TU Delft

Delft University of Technology

Shape-Selectivity Effects of Zeolites on Hydroisomerization of Long-Chain Alkanes

Sharma, S.

DOI
10.4233/uuid:f36da034-5cb3-42ca-a53d-d351f68a9ffa

Publication date
2025

Document Version
Final published version

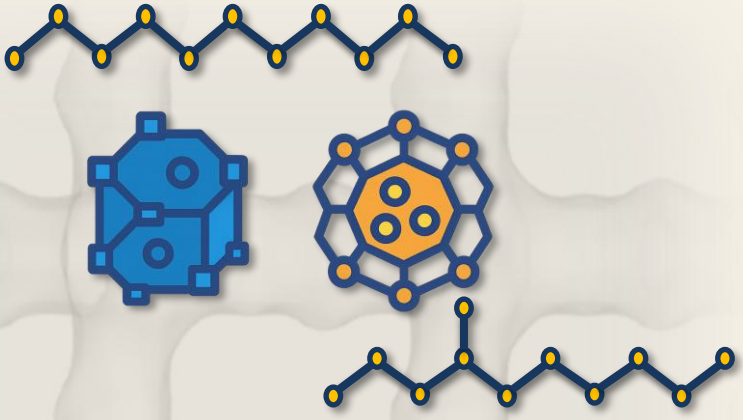
Citation (APA)
Sharma, S. (2025). *Shape-Selectivity Effects of Zeolites on Hydroisomerization of Long-Chain Alkanes*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:f36da034-5cb3-42ca-a53d-d351f68a9ffa

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.
For technical reasons the number of authors shown on this cover page is limited to a maximum of 10.



Shape-Selectivity Effects of Zeolites on Hydroisomerization of Long-Chain Alkanes

Shrinjay Sharma

TU Delft

Shape-Selectivity Effects of Zeolites on Hydroisomerization of Long-Chain Alkanes

Shape-Selectivity Effects of Zeolites on Hydroisomerization of Long-Chain Alkanes

Dissertation

for the purpose of obtaining the degree of doctor
at Delft University of Technology,
by the authority of the Rector Magnificus, prof. dr. ir. T.H.J.J. van der Hagen,
chair of the Board for Doctorates
to be defended publicly on
Monday 12 January 2026 at 12:30 o'clock

by

Shrinjay SHARMA

Master of Science in Mechanical Engineering
Delft University of Technology, the Netherlands
born in Guwahati, India

This dissertation has been approved by the promotors.

Composition of the doctoral committee:

Rector Magnificus,	chairperson
Prof. dr. ir. T.J.H. Vlugt,	Delft University of Technology, *promotor*
Dr. D. Dubbeldam,	University of Amsterdam, *promotor*
Prof. dr. S. Calero,	Eindhoven University of Technology, *promotor*
Independent members:	
Prof. dr. K. Hooman,	Delft University of Technology
Prof. dr. E.A. Pidko,	Delft University of Technology
Prof. dr. S.K. Schnell,	Norwegian University of Science and Technology
Prof. dr. C.M. Simon,	Oregon State University
Dr. ir. M. Ramdin,	Delft University of Technology, reserve member

This work was sponsored by NWO Domain Science for the use of supercomputer facilities. This work is part of the Advanced Research Center for Chemical Building Blocks, ARC-CBBC, which is co-funded and co-financed by the Netherlands Organization for Scientific Research (NWO) and the Netherlands Ministry of Economic Affairs and Climate Policy. The author acknowledges the use of computational resources of the DelftBlue supercomputer, provided by the Delft High Performance Computing Center (https://www.tudelft.nl/dhpc).

UvA

Keywords: Hydroisomerization, Reaction Equilibrium, Breakthrough Curves, and Machine Learning

Copyright © 2025 by S. Sharma

ISBN 000-00-0000-000-0

An electronic copy of this dissertation is available at
https://repository.tudelft.nl/.

This work is dedicated to my family, supervisors, and well wishers.

Contents

1. Introduction **1**
- 1.1. Hydroisomerization . . . 1
- 1.2. Modeling and Simulation Methods for Hydroisomerization . 6
- 1.3. Outline of this thesis . . . 10

2. Segregated Explicit Adsorption Isotherms **13**
- 2.1. Introduction . . . 14
- 2.2. Theory . . . 18
  - 2.2.1. Segregated Ideal Adsorbed Solution Theory (SIAST) . 18
  - 2.2.2. Explicit Isotherms (EI) . . . 21
  - 2.2.3. Segregated Explicit Isotherms (SEI) . . . 24
  - 2.2.4. Breakthrough Curve Model . . . 26
- 2.3. Simulation Details . . . 27
- 2.4. Results and Discussions . . . 30
  - 2.4.1. Adsorption Isotherms . . . 30
  - 2.4.2. Breakthrough Curve Simulations . . . 36
- 2.5. Conclusions . . . 40

3. RUPTURA: Breakthrough Curves Simulation Software **43**
- 3.1. Introduction . . . 44
- 3.2. Pure Component Isotherms . . . 47
  - 3.2.1. Introduction . . . 47
  - 3.2.2. Isotherm models . . . 49
  - 3.2.3. Reduced grand potential . . . 55
- 3.3. Prediction of Mixture Isotherms . . . 57
  - 3.3.1. Introduction . . . 57
  - 3.3.2. Ideal Adsorption Solution Theory (IAST) . . . 59
  - 3.3.3. IAST numerical example . . . 63
  - 3.3.4. Analytic mixture prediction methods . . . 66
  - 3.3.5. Numerical methods to solve the IAST . . . 75
  - 3.3.6. Mixture prediction software . . . 87
  - 3.3.7. IAST implementation in RUPTURA . . . 88
  - 3.3.8. Validation . . . 90
- 3.4. Fixed-Bed Adsorption Simulations . . . 94
  - 3.4.1. Introduction . . . 94
  - 3.4.2. Theory . . . 98
  - 3.4.3. Axial dispersion . . . 101
  - 3.4.4. Mass transfer models . . . 103

3.4.5. Mixture loading prediction . . . 111
3.4.6. Existing breakthrough software . . . 113
3.4.7. Breakthrough model implementation . . . 114
3.4.8. Pulse breakthrough . . . 124
3.4.9. Validation . . . 124
3.4.10. Case Study . . . 128
3.5. Isotherm Fitting . . . 133
3.5.1. Introduction . . . 133
3.5.2. Fitting methods . . . 133
3.5.3. Evaluating isotherm goodness–of–fit . . . 136
3.5.4. Software for isotherm fitting . . . 136
3.5.5. Implementation of the GA-NMA hybrid method . . . 138
3.5.6. Validation . . . 141
3.6. RUPTURA Installation . . . 145
3.6.1. Mac . . . 145
3.6.2. Linux . . . 145
3.6.3. Windows . . . 146
3.7. RUPTURA Input description . . . 146
3.7.1. General . . . 146
3.7.2. Mixture prediction . . . 148
3.7.3. Breakthrough . . . 149
3.7.4. Fitting . . . 151
3.8. RUPTURA Tutorial . . . 151
3.8.1. Introduction . . . 151
3.8.2. Breakthrough of CO_2/N_2 in silicalite . . . 151
3.8.3. Mixture prediction of C_6-isomers in BEA-type zeolite 153
3.8.4. Isotherm fitting . . . 155
3.8.5. Examples . . . 158
3.9. RUPTURA Troubleshooting . . . 159
3.10. Conclusions . . . 160

4. Reaction Equilibrium Distribution in Hydroisomerization 161
4.1. Introduction . . . 162
4.2. Theory . . . 165
4.3. Simulation Details . . . 170
4.4. Results and Discussions . . . 171
4.5. Conclusions . . . 187

5. Thermochemical Properties of Alkanes Using Linear Regression 189
5.1. Introduction . . . 190
5.2. Theory . . . 194
5.3. Results and Discussion . . . 198
5.4. Conclusions . . . 204

6. Computing Entropies for Alkanes Using Linear Regression 207
6.1. Introduction . . . 208

6.2. Theory . . . 210
6.3. Results and Discussion . . . 212
6.4. Conclusions . . . 218

7. Machine Learning for Henry Coefficients of Alkanes in Zeolites 221
7.1. Introduction . . . 222
7.2. Methodology . . . 227
7.3. Results and Discussion . . . 233
7.4. Conclusions . . . 245

8. Conclusions 247

References 251

A. Appendix 291
A.1. Data availability for each chapter . . . 291
A.2. Appendix for Chapter 4 . . . 292
A.2.1. Force field parameters for alkanes . . . 292
A.2.2. Sensitivity analysis for the reaction equilibrium distribution for hydroisomerization of C_7 in MFI-type zeolite 294
A.2.3. Henry coefficients of C_8 isomers in different zeolites . 297
A.2.4. Selectivities of C_8 isomers relative to linear C_8 in the gas phase and different zeolites . . . 299
A.3. Appendix for Chapter 5 . . . 315
A.4. Appendix for Chapter 7 . . . 319
A.4.1. Error analysis of Henry coefficients computed from molecular simulations . . . 319
A.4.2. Comparison between Henry coefficients for alkanes predicted using different machine learning models . . 320
A.4.3. Reaction equilibrium distribution for hydroisomerization of C_{16} isomers . . . 322

Summary 339

Samenvatting 343

Curriculum Vitæ 347

List of Publications 349

Acknowledgements 351

1 Introduction

1.1. Hydroisomerization

Hydroisomerization of alkanes is an essential process for fuel upgrading in the petrochemical and refining industries, where straight-chain alkanes are converted into branched isomers [1]. This transformation significantly increases the research octane number (RON) of gasoline [2, 3], a critical metric for engine performance, and enhances the cold flow properties of diesel and sustainable aviation fuels (SAF) and electro-SAF (eSAF) [4]. Hydroisomerization is often referred to as catalytic dewaxing, particularly in the context of lubricants and diesel refining, because linear alkanes possess high melting points and tend to crystallize as waxes at low temperatures, while the branched isomers exhibit significantly lower melting points due to reduced packing efficiency and steric hindrance [5]. Aromatic compounds, which also improve the RON [6, 7], have major drawbacks such as increased CO emissions [8], soot formation, toxicity [9, 10], and regulatory limitations [4]. Therefore, branched alkanes offer a cleaner and safer pathway to high-performance fuels. Aromatics are especially problematic in aviation fuels, where strict limits are imposed due to adverse effects on combustion efficiency, emissions, and fuel stability [11]. Hydroisomerization is a crucial step in producing high-RON low-aromatic fuels and SAF from renewable feedstocks like hydrotreated vegetable oils and Fischer-Tropsch waxes, where long-chain linear alkanes must be selectively isomerized (Fig 1.1) to meet freezing point and energy density requirements [12]. In parallel, the methanol-to-jet (MTJ) route offers an alternative SAF pathway, where methanol derived from biomass, CO_2, or renewable hydrogen is converted into jet-range hydrocarbons [13].

Zeolite-based catalysts, which are stable at high temperatures, with shape-selective micro/nanopores and tailored acidity, enable efficient isomerization with a certain amount of undesirable cracking [14]. As of 2025, the International Zeolite Association (IZA) has approved approximately 274 zeolite framework types, of which about 55 occur naturally and around 219 are

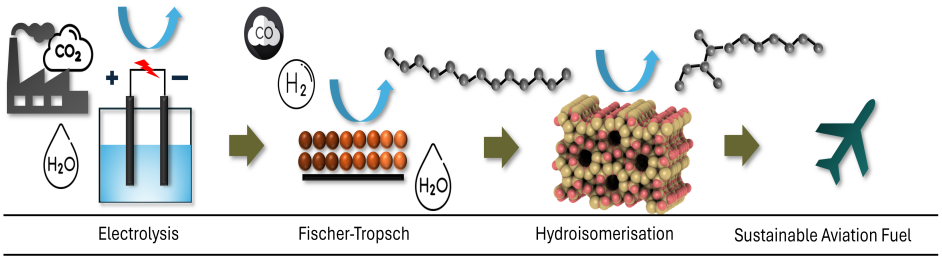


Figure 1.1. Schematic showing production of Sustainable Aviation Fuels (SAFs) from industrial CO_2 emissions. Water electrolysis powered by renewable electricity produces hydrogen, which, together with captured CO_2 undergoes co-electrolysis and produces syngas (CO and H_2O). Further, syngas are converted into long-chain linear alkanes via Fischer–Tropsch (FT) synthesis. These are subsequently isomerized in the shape-selective micro/nanopores of zeolite catalysts to yield high-performance SAFs with improved cold flow properties.

synthetic [15]. Zeolites are crystalline aluminosilicates composed of corner-sharing TO_4 tetrahedra, where T is typically Si or Al [16, 17] (Fig 1.2). These tetrahedra form open three-dimensional frameworks with well-defined rings and channels. Substitution of Si^{4+} by Al^{3+} creates a negative charge on the framework, which is balanced by extra-framework cations such as H^+ or Na^+ located in the pores [17]. The size and geometry of the rings (6-, 8-, 10-, or 12-membered rings) and channels govern the shape-selectivity and catalytic behavior. The zeolite pores act as molecular sieves, selectively adsorbing, diffusing, and transforming molecules based on size, shape, and polarity [17]. The shape-selective behavior of zeolites arises from the geometric constraints imposed by the micropores with diameters typically ranging from 3 to 10 Å [18]. The pore confinements allow zeolites to selectively adsorb alkane isomers and influence adsorption equilibrium, diffusion, and reaction pathways. For instance, medium-pore zeolites such as MFI-type zeolite show pronounced selectivity due to the presence of intersecting channel systems, which restrict the diffusion of bulky isomers. In sharp contrast, large-pore zeolites like FAU-type zeolite offer more open architectures that accommodate a broader range of molecules but may sacrifice some shape-selectivity. Hydroisomerization proceeds via a bifunctional catalytic mechanism involving both metallic and acidic sites in zeolites [14, 19, 20]. Metallic sites, typically containing noble metals like platinum or palladium, facilitate adsorption and dehydrogenation of linear alkanes. This leads to the formation of alkenes. These alkenes are converted into alkyl carbenium ions via protonation in the Brønsted acidic sites [21]. The alkyl carbenium ions are transferred to the metal sites, where alkanes are produced via

hydrogenation [21]. The acid sites inside the zeolites also promote the cracking of alkanes. Cracking reactions involve the cleavage of C–C bonds in hydrocarbons to form smaller molecules. Cracking reactions compete with hydroisomerization and form undesired products with small chain lengths, especially at high acidity or temperature conditions [14].

Studying the reaction kinetics of hydroisomerization provides a conventional approach to assess the influence of shape-selectivity on reaction pathways and product distribution [20, 21, 23–25]. Sazama et al [23] investigated the shape-selectivity effects of MFI-, MOR-, BEA-, and FAU-type zeolites on hydroisomerization of C_6 isomers, revealing that spatial confinements in zeolite channels strongly influence the selectivity. Zeolites featuring 10-member ring channels showed a preference for mono-branched isomers, while the presence of 12-member rings in FAU- and BEA-type zeolites reduced shape-selectivity effects, resulting in both mono- and di-branched isomers [23]. Poursaeidesfahani et al. [21] studied the product shape-selectivity of MEL-, MFI-, and BEA-type zeolites in the hydroconversion of n-heptane (C_7), observing the conversion of linear C_7 to mono-branched isomers, which are subsequently transformed into di-branched isomers with higher free energy barriers for diffusion. Therefore, di-branched isomers are eventually cracked into faster diffusing products [21]. In MFI- and MEL-type zeolites, di-branched isomers with lower diffusion barriers are produced with a higher yield [21]. Agarwal et al. [20] studied shape-selectivity effects of BEA-type zeolite on C_7 hydroconversion using a reactor model. This reactor model can be used to separate the effects of pore structures on adsorption-based selectivities and intra-porous reaction kinetics. Han et al. [24] investigated the influence of zeolites with one-dimensional channel-like pores on the hydroisomerization of C_{12} isomers. These authors observed that the channel dimensions in the range between 4.6×5.7 Å and 5.6×7.0 Å are suitable for the production of di-branched isomers of C_{12}. Tri-branched isomers are favored in channels of dimensions larger than 5.6×7.0 Å [24]. The formation of tri-branched isomers is followed by cracking reactions [24]. Zhang et al. [25] studied shape-selectivity effects of TON-, MTT-, FER-, and MRE-type zeolites in the hydroisomerization of C_{16}. Higher conversion was observed in the FER-type zeolite. TON-, MTT-, and MRE-type zeolites showed higher selectivity towards 2-m-C_{15}. 4-m-C_{15}, and 5-m-C_{15} were formed only in MTT- and MRE-type zeolites. Zhang et al. noted that the absence of multi-branched isomers and high selectivity towards 2-m-C_{15} is due to a pore mouth reaction mechanism involved in the process [25]. According to this mechanism, mono-methyl alkanes can partially enter the pores of one-dimensional zeolites like TON-, MTT-, and AEL-type zeolites [26]. This mechanism also suggests that normal alkanes are predominantly hydroisomerised at the pore mouths. The validity of the pore-mouth mechanism remains contentious, with conflicting interpretations regarding the accessibility of pore entrances to branched alkanes [27]. Maesen et al. showed that both normal and mono-methyl branches can enter the pores of TON-, MTT-,

1

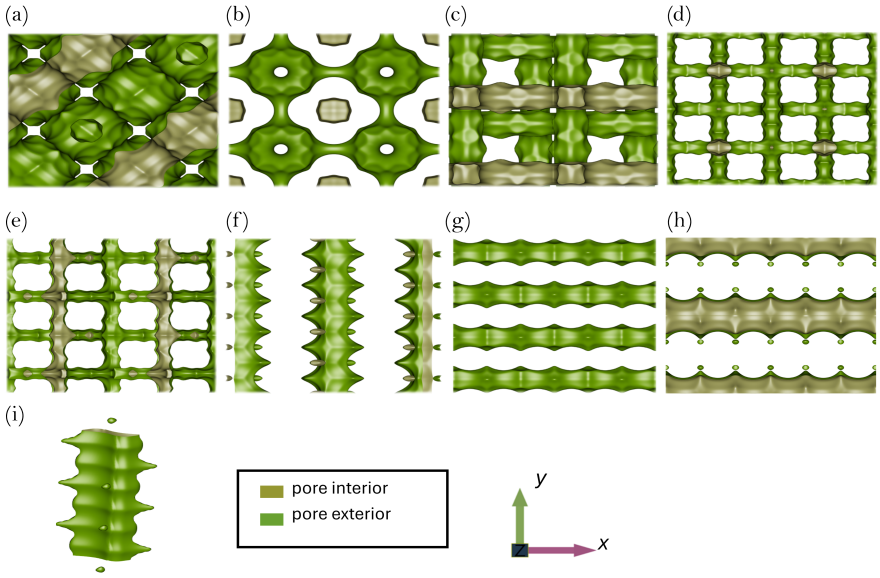


Figure 1.2. Typical representations of pore structures of (a) FAU-, (b) ITQ-29-, (c) BEA-, (d) MEL-, (e) MFI-, (f) MTW-, (g) MRE-, (h) AFI-, and (i) MTT-type zeolites. FAU-type zeolite has cage-like pore structures. ITQ-29-type zeolite also has cage-like pore structures connected by narrow channels. A three-dimensional network of straight channel-like pores in both vertical and horizontal directions is present in BEA- and MEL-type zeolites. MFI-type zeolite has sinusoidal-shaped channels in the horizontal direction and straight channels in the vertical direction. MTW-, MRE, AFI, and MTT-type zeolites contain one-dimensional channel-like pores in vertical and horizontal directions, respectively. In these representations, the horizontal direction is represented by the x-axis and the vertical direction is represented by the y-axis. The iRASPA software [22] was used to generate these images.

and AEL-type zeolites [28]. These authors suggest that the high selectivities for alkanes with terminal methyl groups are due to product shape-selectivity rather than pore mouth catalysis. The higher desorption rate of alkanes with terminal methyl groups leads to shorter residence times and are less prone to react inside the zeolite pores [28]. The other possible mechanism is transition state shape-selectivity [29], which occurs when the geometry of the zeolite pores allows the adsorption and diffusion of reactant and product molecules but restricts the formation of certain transition states due to steric hindrance in the confined space [29]. This leads to the preferential formation of isomers or products whose transition states are geometrically compatible with the zeolite framework. Ernst et al. [29] suggest that the transition states of terminal methyl groups fit better inside TON-, MTT-, and

AEL-type zeolites than alkanes with internal methyl branches. This leads to a higher preference for alkanes with methyl groups present at the end of the hydrocarbon chain. Romero et al. [30] investigated the hydroconversion of n-C_{16} in Pd-FAU, Pd-BEA, Pt-Pd-BEA, and SBA-15 at different conversion levels (4-66%). It was observed that at low conversion, di-branched isomers with methyl groups positioned near the ends of the carbon chain are favored, leading to a cracked product distribution enriched in short-chain alkanes (C_4–C_6) [30]. As conversion increases, di-branched isomers with methyl groups closer to the centre of the chain become more prevalent, thus shifting the cracked product distribution towards intermediate chains (C_8–C_{10}) [30].

Hydroisomerization reactions are reversible, and the reaction product distribution can be close to equilibrium, especially at low conversion [31, 32]. To identify the conditions that can potentially yield optimal reaction products, it is important to comprehend the role of chemical equilibrium in these reactions. Comparing the reaction product distribution at reaction equilibrium with data from experiments or kinetic models helps us to distinguish between the thermodynamic and the kinetic contributions to these reactions. Steijns and Froment [33] compared experimental data on hydroisomerization of C_{10} on Pt/Y-type zeolite to reaction equilibrium data obtained from literature. These researchers have observed that following the hydroconversion process, the fraction of C_{10} in the mixture is larger than the corresponding fraction at equilibrium. The concentrations of the mono- and multi-branched isomers are lower in the actual experiments compared to the thermodynamic equilibrium data [33]. Steijns and Froment [33] have concluded that the conversion of linear alkanes into mono-methyl alkanes is influenced by thermodynamic equilibrium. Estrada-Villagrana et al. [31] modeled the hydroisomerization of a mixture of alkanes (C_6, C_7, C_8, and C_9) in Pt/USY-type zeolite at reaction equilibrium. These authors studied the following cases: (a) there is no restriction on cracking reactions, (b) limiting the cracking of 3,3-m-C_5 and 2,2,3-m-C_5 isomers, and (c) a system without any cracking reaction. The product distribution closest to experimental results was obtained by limiting the cracking of 3,3-m-C_5 and 2,2,3-m-C_5 [31]. The selectivities obtained for the case without cracking reactions differed significantly from the experimental results. This case involved only products due to isomerization. In the experiments, the cracking of longer chain isomers (e.g., C_9) was observed, and the reaction equilibrium was not achieved [31]. However, Estrada-Villagrana et al. emphasized the importance of analysing such systems at reaction equilibrium to enhance selectivity towards isomerization over cracking reactions [31]. The role of the zeolite is shape selective, as the zeolite topology does not influence the acidity of an acid site.

After hydroisomerization, the resulting mixture of linear and branched alkanes must be separated to isolate high-octane isomers used in gasoline blending or feedstocks for further processing. Traditional separation methods, such as distillation, are energy-intensive and often inefficient for

isomers with similar boiling points [34]. In sharp contrast, adsorption-based separation offers a more energy-efficient and selective approach by exploiting subtle differences in molecular shape and size [35]. Narrow pores in zeolites and other microporous materials can preferentially adsorb linear alkanes over branched ones, enabling shape-selective separation [17]. This selectivity is particularly valuable in separating mono-branched from multi-branched or linear isomers, where small differences in molecular geometry play a critical role. The adsorption-based strategies using one-dimensional zeolites such as MTW- and TON-type zeolites can achieve high separation factors for alkane isomers, positioning these zeolites as promising candidates for industrial hydroisomerization [24].

To investigate the pore-confinement effects and reaction pathways outlined in Section 1.1, molecular simulations were used. Modeling of hydroisomerization reactions presents several major challenges due to the vast combinatorial complexity of both the catalyst and reactant spaces. There are millions of possible zeolite frameworks, each with distinct pore topologies and acid site environments, and an even greater number of possible reactant isomers, particularly for long-chain alkanes. For example, C_{16}–C_{25} alkanes relevant to diesel and lubricant applications can have thousands of structural isomers, each with different thermochemical properties. However, reliable thermochemical data, such as enthalpies, entropies, and Gibbs free energies, are typically available only for smaller alkanes (below C_{10}), severely limiting the ability to predict reaction equilibrium distributions for industrially relevant systems. This thesis presents a framework to compute reaction equilibrium distributions for the hydroisomerization of long-chain alkanes.

1.2. Modeling and Simulation Methods for Hydroisomerization

Experimental approaches to hydroisomerization are often time-consuming and resource-intensive, requiring the synthesis and detailed characterization of tailored zeolites, followed by extensive catalytic testing. In contrast, molecular simulations offer a powerful and efficient alternative by enabling rapid screening of zeolite-reactant combinations and providing molecular-level insight into adsorption, diffusion, and reactivity. This approach is especially valuable for long-chain alkanes and hypothetical zeolite structures, for which experimental data are scarce and difficult to obtain. Simulations facilitate property estimation and mechanistic understanding that are often inaccessible through experiments, thus offering a fundamental framework for rational catalyst design and process optimization.

For alkanes, adsorption is primarily governed by weak van der Waals (dispersive) interactions with the pore walls of the zeolite framework [36]. Since alkanes are non-polar and lack functional groups capable of forming hydrogen bonds or electrostatic interactions, the presence of framework

protons (Brønsted acid sites) has a small effect on their adsorption behavior. As a result, alkane uptake is determined largely by the pore geometry and topology of the zeolite, rather than acid site density or placement. All acid sites are equally acidic. Modeling dispersive interactions poses a significant challenge for quantum chemical methods due to their long-range and many-body character. To address this, classical force field–based molecular simulations such as Monte Carlo (MC) and Molecular Dynamics (MD) are widely used [17]. These methods use parameterized force fields to model interactions between the adsorbate molecules and the host framework. The force field parameters are either derived from quantum chemical calculations or tuned to experimental data, such as adsorption isotherms or vapor–liquid equilibria [37]. In these simulations, alkanes are considered as flexible molecules. For the available force fields, united atom models, where non-polar hydrogen atoms are lumped into carbon atoms, are particularly well suited for alkane adsorption. United-atom force fields, specifically optimized for adsorption studies, often yield more accurate and computationally efficient predictions of adsorption properties than all-atom models [38]. These models are suitable for alkane adsorption because the united-atom groups (C, CH, CH_2, CH_3) are charge-neutral, allowing Coulombic interactions to be neglected [39]. The TraPPE united-atom force field [37] is used to describe intramolecular interactions (bond stretching, angle bending, and torsions), while Lennard-Jones potentials model non-bonded interactions [40]. In the TraPPE united-atom force field, bond lengths are constrained to fixed values, as the high frequency modes of bond vibrations or stretching have very small contributions to thermodynamic properties such as densities, vapor-liquid equilibria, and adsorption [37]. Instead, molecular flexibility is introduced only through bond-bending and torsional (dihedral) potentials, which govern conformational changes of alkanes and other chain molecules [37]. This improves computational efficiency (allowing more Monte Carlo cycles and fewer degrees of freedom) while still capturing the essential configurational sampling relevant to fluids and mixtures. Cross-interactions are handled using Lorentz–Berthelot mixing rules [41, 42]. For adsorbent–adsorbate interactions, the TraPPE-zeo force field [43] is used. Further details on the force field parameters can be found in Appendix A.2.1. All-silica zeolites are treated as rigid frameworks, an assumption justified by the negligible influence of framework flexibility on adsorption at infinite dilution [44]. At infinite dilution, small fluctuations in the atomic positions of the framework due to thermal motion (flexibility) do not significantly alter the accessible volume for adsorption. Vlugt and Schenk [44] showed that for both linear and branched alkanes, the Henry coefficients and heats of adsorption differ by less than 10% when comparing rigid and flexible frameworks at low loadings. These minor differences do not significantly affect overall adsorption predictions. Additionally, rigid framework models are attractive due to computational efficiency. When the fit of a molecule inside the zeolite pores is largely constrained, framework flexibility can become important [45].

In this thesis, we do not explicitly include protons inside zeolite pores as we assume identical interactions of the alkane isomers with the protons. The effects of protons become important for alkenes, which react with protons to form alkylcarbenium ions. This is an intermediate step in the hydroisomerization reaction, which is considered to be in equilibrium with the dehydrogenation of alkanes to form alkenes and the formation of corresponding alkane isomers via hydrogenation. At low concentrations, adsorption is well described by Henry's law, with the equilibrium loading proportional to the Henry coefficient, k_H. Henry coefficients for alkanes are typically computed using the Widom test particle insertion method [46], in combination with Configurational-Bias Monte Carlo (CBMC) [47–49]. The Widom test particle insertion method [46] is a simple and efficient technique to compute how easily a molecule can enter a system, especially at low concentrations. It randomly inserts a test molecule into a host material, such as zeolites, and calculates how much energy this insertion would cost. By repeating this several times and averaging the results, this method computes the excess chemical potential or Henry coefficient. CBMC enhances sampling of flexible molecules, such as long-chain alkanes, by growing chains segment-by-segment, applying a bias toward energetically favorable conformations and correcting it with the Rosenbluth factor to maintain detailed balance [50]. To simulate full adsorption isotherms, Grand-Canonical Monte Carlo (GCMC) simulations are used [50]. GCMC operates at constant chemical potential, temperature, and volume, allowing the number of adsorbed molecules to fluctuate, thereby mimicking equilibrium with a gas-phase reservoir. The gas-phase chemical potential is often calculated using equations of state, such as the Peng–Robinson equation [36]. GCMC is useful for multicomponent systems and industrial conditions, though it becomes computationally intensive at high pressures or when electrostatics are significant. Faster alternatives, such as Ideal Adsorbed Solution Theory (IAST) [51, 52] and Real Adsorbed Solution Theory (RAST) [53], predict mixture isotherms using pure component isotherm parameters as input. IAST assumes that the adsorbed phase behaves like an ideal solution, with no interactions between the adsorbate molecules [51, 52, 54]. IAST is based on the principles of homogeneous distribution of adsorbates and the same surface area occupied by a guest molecule in a mixture as in its pure state. It is computationally efficient for screening a large number of adsorbent materials for a specific application [55]. However, due to its limitations in heterogeneous systems [56], this thesis extends the approach using segregated explicit isotherm models that better account for adsorbate-size dependence and heterogeneities of adsorption sites.

Molecular Dynamics (MD) complements Monte Carlo methods by providing time-resolved insights into molecular motion, allowing for the calculation of transport properties such as diffusivity and viscosity. However, MD simulations of long-chain alkanes in zeolites are computationally demanding due to extremely slow diffusion and long correlation times. To overcome this,

rare-event methods such as Transition State Theory (TST) and dynamically corrected TST (dcTST) are employed to estimate diffusivities by computing free energy barriers and rate constants for hopping events across pore constrictions [57–59]. These methods enable the efficient prediction of diffusion behavior even when performing MD simulations is impractical.

Most industrial separation processes are governed by dynamic rather than equilibrium conditions [60]. Evaluating adsorption processes solely based on adsorption isotherms is insufficient. Dynamic properties such as cyclic and long-term stability, competitive adsorption behavior, and mass transfer kinetics must also be characterized to assess adsorbent performance effectively [60]. Transient breakthrough curves, in particular, provide critical insights into the capacity of an adsorbent to selectively separate mixtures or capture specific components under flow conditions [61]. Breakthrough curves describe the dynamic response of an adsorption column, showing how the concentration of a target component in the outlet stream evolves over time as the adsorbent becomes saturated [62, 63]. They are essential for designing and evaluating fixed-bed adsorption processes at realistic flow conditions [64, 65]. To predict such dynamic behavior, fixed-bed adsorption simulations are widely used. These simulations involve solving transport equations that account for advection, axial dispersion, mass transfer between fluid and solid phases, and adsorption kinetics [62]. Numerical methods such as finite difference, finite volume, and finite element approaches are typically used to discretize and solve the resulting system of coupled Partial Differential Equations (PDEs) [66]. Time integration can be performed using either explicit (Runge-Kutta [67]) or implicit schemes (Backward Differencing Formula [67]), depending on the stiffness of the problem. RUPTURA (chapter 3), our open-source software for simulating breakthrough curves, uses the finite difference method with the method of lines framework [68] to discretize the coupled PDEs. A third-order, three-stage explicit method, the Strong Stability Preserving Runge-Kutta (SSP RK(3,3)) [69, 70], is used for time integration.

Accurate thermochemical properties, such as enthalpy, Gibbs free energy, and entropy, are essential for predicting reaction equilibria and evaluating the energy efficiency of chemical processes. However, for long-chain alkane isomers, such data are often unavailable due to the vast number of possible molecular structures. A common strategy to estimate these properties is using group contribution methods [71–73], which relate molecular structure to thermodynamic quantities. These methods are based on an additive approach, in which a molecule is fragmented into predefined functional groups, and the thermochemical properties are predicted by adding the contributions of these groups [72]. Most group contribution methods estimate thermochemical properties using first-order groups or limited second-order corrections, which reduces accuracy for complex hydrocarbons [74]. In recent times, ML models have also gained popularity in predicting thermochemical properties [74–76]. While ML models can

offer improved predictions, they often struggle with extrapolation [75, 77], especially for long-chain alkanes. Linear regression (LR), when based on sufficient and relevant variables, can outperform ML in such cases [77]. A hybrid approach combining LR and ML can also be considered to leverage the advantages of both methods. In this thesis, a user-friendly LR model is developed that uses all possible second-order groups to reliably predict thermochemical properties of alkanes beyond C_{10}.

Predicting Henry coefficients and adsorption isotherms is essential for the design of adsorption-based separation processes, particularly in systems involving complex adsorbates and porous materials. In materials such as metal organic frameworks (MOFs) and zeolites, machine learning (ML) has emerged as a powerful tool for accurately predicting adsorption properties, enabling high-throughput screening without the need for costly molecular simulations or labor-intensive experiments [78–81]. Unlike thermochemical property estimation, linear regression (LR) is generally unsuitable for this task due to the highly non-linear and structure-sensitive nature of adsorption, which cannot be described by simple mathematical models. Consequently, advanced ML methods have become increasingly favored for capturing these complex structure-property relationships. The literature reports various ML approaches, including tree-based models (Random Forest [82], XGBoost [83], CatBoost [84, 85]), kernel methods (Support Vector Regression), deep learning architectures (graph neural networks [86–88]), and transformer based models (TabPFN [89, 90]). These models rely on molecular descriptors such as topological indices, geometric features, chemical properties (e.g., polarizability, molecular volume), and graph-based representations (e.g., atom connectivity, local environments) [91]. For zeolites and other porous materials, structural properties (e.g., pore size, surface area, channel dimensionality) are also included [91]. By learning complex, non-linear interactions between adsorbate and adsorbent features, ML models enable accurate prediction of adsorption properties and support high-throughput screening for shape-selective separations. Having established the simulation framework, the structure of the thesis is outlined below.

1.3. Outline of this thesis

The goal of this thesis is to develop a unified, multiscale modeling workflow to understand and predict product distributions for zeolite-based shape-selective hydroisomerization of long-chain alkanes, and to design efficient strategies for separating these products. This thesis addresses how molecular structures and zeolite pore geometries influence the product distribution of the hydroisomerization reaction at equilibrium, which is important for producing high-octane and sustainable aviation fuels. Fig. 1.3 shows the outline of this thesis. It starts with developing a segregated explicit isotherm model, combining adsorbate-size-dependent explicit isotherm and segregated IAST approaches, to better predict adsorption equilibria in het-

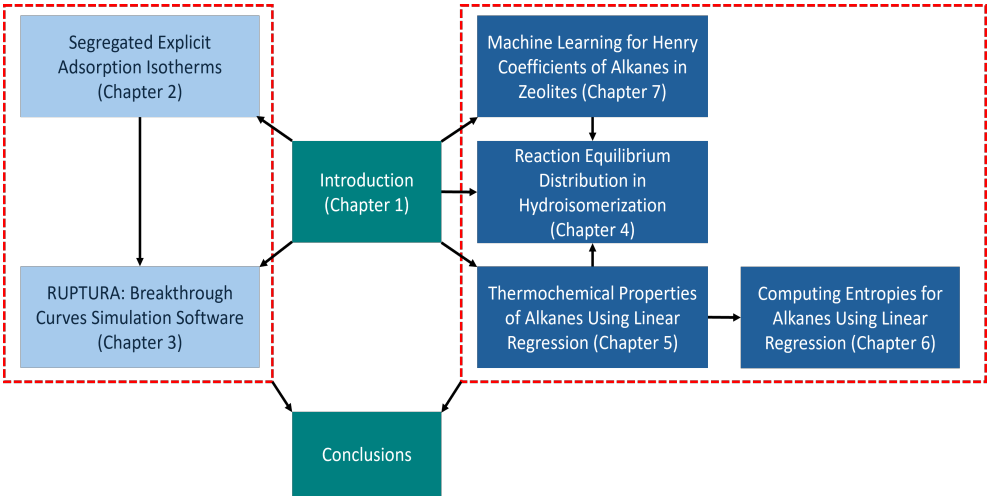


Figure 1.3. Schematic representation of the structure of this thesis. Chapters about the adsorption-based separation (chapters 2 and 3) are shown in light blue on the left side of the figure. Chapters about the zeolite-based catalytic hydroconversion of hydrocarbons (chapters 4, 5, 6, and 7) are shown in white on the right side of the figure.

erogeneous porous materials. Being explicit, this model also enables faster simulations when integrated with breakthrough curve models as discussed in chapter 2. Recognising that real-world separations occur at dynamic conditions, RUPTURA (chapter 3) is developed as open-source software tool to compute breakthrough curves, perform pure component isotherm fitting and predict multicomponent adsorption isotherms using both IAST and explicit isotherms, and the segregated versions of these models. Its open accessibility promotes wider adoption, supports reproducible research, and enables community-driven extensions for adsorption-based separation processes.

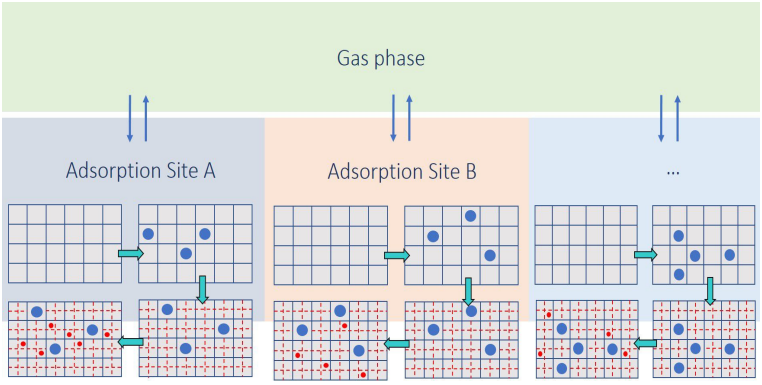
Adsorption insights are further extended to model the reaction equilibrium distribution of alkane isomers during hydroisomerization. The reaction equilibrium distribution plays a key role in determining product selectivities at equilibrium conditions and helps in distinguishing between thermodynamic and kinetic effects on the reaction. This distribution is computed by coupling gas-phase reaction equilibrium with phase equilibrium between the gas and the adsorbed phases. For the gas phase equilibrium, the thermochemical properties of alkanes, Gibbs free energies and enthalpies of formation are required. For alkanes larger than C_{10}, these properties are not readily available in the literature. Also, these group contribution methods do not always predict these properties beyond the chemical accuracy of 1 kcal/mol. Therefore, a second-order group contribution method is

developed to predict the thermochemical properties of long-chain alkanes with chemical accuracy within 1 kcal/mol. The details of this method are discussed in chapter 5. Using the predicted enthalpies and Gibbs free energies, the entropy values are computed for alkanes. Accurate entropy values are essential for designing energy-efficient processes for (iso)alkane handling and separation. Energy efficiency of such processes is best assessed through the second law of thermodynamics, which requires knowledge of exergy destruction, i.e., the product of entropy production and environment temperature. These concepts are discussed in chapter 6. To compute the adsorbed-phase composition for reaction equilibrium at infinite dilution, we need Henry coefficients. Obtaining Henry coefficients from simulations and experiments for all combinations of zeolites and long-chain alkane isomers is computationally expensive and time-consuming due to the large number of potential zeolite structures as well as the millions of structural isomers that exist for long-chain alkanes. Therefore, a Machine Learning (ML) framework is developed as discussed in chapter 7 to predict Henry coefficients for these isomers in one-dimensional zeolites. This model uses molecular descriptors such as chain length, branching, and graph-based representations, enabling accurate and efficient prediction of Henry coefficients for various branching types in alkanes, such as methyl, ethyl, propyl, and isopropyl groups. Different Machine Learning (ML) models, particularly tree-based models (Random Forest [82], Extreme Gradient Boosting [83], Cat Boost [84]), Graph Neural Network (Directed Message Passing Neural Network [87, 92]), and Tabular Prior Fitted Network [89, 90], are evaluated. These models are trained on high-fidelity molecular simulation data and used to uncover variations in Henry coefficients driven by pore confinements and steric effects of guest molecules. Chapter 8 presents the concluding remarks of this thesis, emphasizing that the developed methods enable high-throughput screening and data-driven optimization of zeolite-catalyzed processes. While the primary focus has been on hydroisomerization, the approaches are broadly applicable to a wide range of catalytic and separation systems involving complex hydrocarbon mixtures. Looking ahead, future work can expand the chemical and materials space, adopt advanced sampling and learning strategies, and incorporate kinetics and transport phenomena into integrated reactor models. Overall, this thesis establishes a robust and flexible foundation for scalable, transparent simulation frameworks that bridge molecular-level understanding with process-scale design and optimization.

2

Segregated Explicit Adsorption Isotherms

This chapter is based on the following publication: S. Sharma, M. S. Rigutto, R. Baur, U. Agarwal, E. Zuidema, S. R. G. Balestra, S. Calero, D. Dubbeldam, and T. J. H. Vlugt. "Modelling of adsorbate-size dependent explicit isotherms using a segregated approach to account for surface heterogeneities". *Mol. Phys.* 121 (2023), e2183721.



2.1. Introduction

Adsorption plays an important role in the separation and purification processes of various fluid mixtures [94]. Adsorption-based methods are often used in a wide variety of industrial processes like the separation of hydrocarbon isomers [95], water purification [96], refrigeration [97], CO_2 capture [98], etc. Thermodynamic data like adsorption loadings, heat of adsorption, and heat capacities are crucial in designing adsorption-based processes [99]. Collecting these data for the case of multi-component mixtures using experiments can be very challenging. This is a time-consuming process, as it involves a large number of experiments [100]. Therefore, one has to rely on different modeling techniques to estimate these datasets. Molecular simulations (e.g., Grand-Canonical Monte Carlo [50]) can be used to calculate the adsorption loadings as a function of temperature and pressure. Alternate approaches such as the mixed Langmuir equation [101], Ideal Adsorbed Solution Theory (IAST) [51, 52], and Real Adsorbed Solution theory (RAST) [53] can be used to predict the adsorbed loadings for mixtures. Computing the mixture isotherms for cases consisting of a large number of components using GCMC is often computationally expensive and time-consuming. The same isotherms can be predicted using IAST with much fewer computational resources. The main advantage of IAST is its relative simplicity. The mixture loadings are calculated solely based on the pure component isotherms. The implementation of RAST is also similar to IAST. Unlike IAST, RAST considers a non-ideal adsorbed mixture [102]. The non-ideality of the adsorbed phase is described by the activity coefficients, which depend on the adsorbed loadings [103]. These coefficients are determined using approaches like NRTL [104], Wilson [105], the UNIFAC model [106], etc. The introduction of the activity coefficients leads to a complex system of implicit equations. Often, the activity coefficients are obtained by fitting the respective equations to the experimental loading data [102]. Thus, the predictive nature of the approach is lost in the process.

Due to its simplicity, IAST is often used in high-throughput screening techniques for selecting appropriate adsorbents for a specific purpose [52, 55, 107]. An application where the screening technique is used is the selection of a zeolite or Metal Organic Framework (MOF) for the separation of heptane isomers [55]. The pure component isotherms are obtained either using experiments or molecular simulations. The datasets are then fitted to a suitable adsorption isotherm equation, e.g., the single or multi-site Langmuir isotherm [64], Langmuir-Freundlich isotherm [108], etc. The isotherms of the components in a mixture can be calculated using IAST based on the fitted parameters of the pure component isotherms. For the Langmuir isotherm, the fitted parameters are the saturation loadings and the equilibrium constants.

There are two main assumptions of IAST: (1) It assumes a homogeneous adsorbed phase and disregards the presence of different types of adsorption sites inside the adsorbent. IAST assumes that the adsorbates have equal

access to the available uniform adsorbent surface [56]. This is not always correct as in many cases, there are distinct adsorption sites, and different adsorbates will have their preferred sites for adsorption [56, 109]. (2) IAST considers the adsorbed phase to be an ideal mixture, i.e. the adsorbed molecules interact with each other and the adsorbent using a mean strength of interaction. Deviations from the first assumption can be observed in cases such as the separation of CO_2-CH_4 mixture in DDR-type or ERI-type zeolites [109, 110]. Zeolites like DDR and ERI have cages connected by narrow windows or constrictions. These zeolites are useful for the separation of CO_2 containing mixtures [109]. A significant amount of the adsorbed CO_2 can be found inside the narrow constrictions or the windows. Krishna et al [109] showed that IAST significantly under-predicts the adsorption behavior of the weakly adsorbing CH_4 molecules inside these zeolites. IAST assumes that the adsorption of the CH_4 molecules will be affected by all the adsorbed CO_2 molecules [109, 110]. In reality, the phenomenon of competitive adsorption varies for each type of adsorption site. CH_4 (preferentially adsorbing inside the cages) will be affected by only those CO_2 molecules which are present inside the cages. The second assumption fails in instances where thermodynamic non-ideality occurs in the adsorbed phase. This may occur due to the presence of charged particles in the adsorbent. Krishna et al. [103] showed that for the separation of CO_2 and CH_4 mixtures in NaX zeolite at 300 K, IAST overestimated the CO_2/CH_4 selectivity by ca. 50% [103]. This happens because the CO_2 molecules congregate around the extra-framework cations (Na^+) present inside the zeolites. The CH_4 molecules experience less severe competition from the adsorbed CO_2 molecules at locations which are devoid of these cations [103]. Such non-ideal behavior of the adsorbates cannot be captured by IAST. Since RAST considers non-ideal mixtures, it can predict the selectivity values that are in excellent agreement with the GCMC calculations [103], at the expense of solving a more complicated system of implicit equations compared to IAST.

Several adsorption models account for the surface heterogeneity while calculating the mixture isotherms. Sircar [111] added heterogeneity to the Langmuir isotherm model, for both pure components and mixtures. For mixtures, the model is valid only for components with equal saturation capacities. Valenzuela et al. [112] modified the IAST model by considering an adsorbent with multiple adsorption sites. To account for the surface heterogeneity, a bimodal energy distribution function is defined for each component [112]. This ensures that the preference for adsorption will not be the same at each site. IAST is applied locally to each of these sites, and the overall equilibrium loadings are obtained by integrating over the entire energy distribution for each component. This modification can improve the equilibrium loading values predicted by IAST for a heterogeneous adsorbent. However, care must be taken with the choice of the energy distribution function because it largely affects the results obtained using this approach [113]. This model is called Heterogeneous Ideal Adsorbed Solution Theory

(HIAST) and is valid only when all components follow the same order of the preferred sites for adsorption [112]. Moon and Tien [114] took this approach one step further and developed an empirical procedure to account for the cases where all the components do not follow the same order of preferred sites for adsorption. The disadvantage of these methods is that an additional assumption is required in the form of the energy distribution function for the adsorbent. Ritter et al. [115] also adopted the approach of Moon and Tien [114] to identify the preferred adsorption sites for the components in the mixture, and applied this modification to the dual-site Langmuir isotherm.

Swisher et al. [56] developed a conceptually simpler approach to deal with the issue of segregation. These authors subdivided the adsorbed pore volume into regions where separate competitive adsorption can take place. Each region is considered to be uniform, where there is a separate thermodynamic equilibrium between the gas and the adsorbed phase. IAST is applied to each of these adsorption sites individually. Therefore, an iterative process is required to calculate the equilibrium loadings at each site. The total adsorbed loadings for the individual components are obtained from the sum of the loadings at each site. This approach is known as the Segregated Ideal Adsorbed Solution theory (SIAST) [56]. This method outperforms IAST when the adsorbent is composed of distinct adsorption sites. The adsorbed loadings estimated using SIAST are in excellent agreement with GCMC data [56] for adsorbents with segregated adsorption sites.

The IAST, RAST, and SIAST models involve a set of implicit equations which need to be solved iteratively. No analytic solution is available for IAST, except for binary mixtures with equal saturation loading [116]. Therefore, these methods are relatively slow from a computational point of view. Instead of IAST or SIAST, the use of an analytic method is computationally much more efficient, especially when incorporated into a breakthrough curve or a reactor model. In the numerical breakthrough-curve model, IAST must be solved at every time step and grid point, incurring a high computational cost. Explicit isotherms such as mixed Langmuir, mixed Langmuir-Freundlich, etc., are commonly used to compute the adsorption isotherms of components in a mixture. However, these isotherms may not provide correct results for high loadings (low temperature or high pressure) [20]. One of the key assumptions of the mixed Langmuir isotherm is that the adsorbed molecules are not affected by the presence of the other co-adsorbed molecules [62]. The mixed Langmuir equation is thermodynamically inconsistent if the saturation loadings of all components are not equal [117].

Van Assche et al. [118] developed an explicit multi-component adsorption model which accounts for the size effects of the components. In this chapter, this model will be referred to as Explicit Isotherm (EI). The equations are derived based on the fundamentals of statistical mechanics. The derivation is similar to the Langmuir isotherm using statistical mechanics [62]. The adsorbent is considered as a lattice of identical adsorption sites where the component with the smallest saturation capacity (i.e., the largest component)

is considered to adsorb first. The remaining sites in the lattice are again uniformly subdivided for the component with the next smallest saturation capacity. This process continues until the adsorption of all components is considered. These authors calculate the expressions for the number of possible ways to perform these arrangements and relate these expressions to the chemical potentials of the components, from which a new set of multi-component explicit isotherms is generated. The model can be extended to any number of components and is capable of capturing the Adsorption Preference Reversal (APR) [119, 120]. This means that at low pressures, it favors the adsorption of components with the largest size or the smallest saturation loading and at high pressures, components with smaller size or larger saturation capacity are preferred [119–122]. This phenomenon is known as size entropy [119, 121]. The model reduces to the single-site Langmuir model for the case of pure components. If the saturation capacities of the components are identical, the model transforms into the mixed Langmuir model. Similar to IAST, this model also assumes uniform adsorption sites. Therefore, it cannot be directly used for cases where the adsorbent is composed of distinct adsorption sites and the adsorbent exhibits strong segregation effects.

The goal of this chapter is to extend the EI model to capture the effects of the surface non-uniformity (which is also suggested by Van Assche et al. [118]), validate this for different systems, and check for its influence on the breakthrough curve simulations. We have adopted the approach of Swisher et al. [56] to arrive at an explicit adsorption model for realistic segregated adsorption. The adsorbent is subdivided into distinct adsorption sites. Each site is considered to be uniform. The adsorbed phases in these sites are individually in thermodynamic equilibrium with the gas phase. Instead of applying IAST to these sites, we use the explicit isotherms developed by Van Assche et al. [118]. The sum of the adsorbed loadings at each site yields the overall adsorption of the components in the mixture. We will refer to this model as the Segregated Explicit Isotherm (SEI). A comparative study is performed between the SEI, SIAST, and IAST models and GCMC mixture simulations. It is observed that SEI produces results similar to SIAST. Both SEI and SIAST outperform IAST when strong segregation of adsorption sites is prevalent. The values obtained using SEI and SIAST are consistent with GCMC data.

Most industrial separations take place at dynamic conditions [60]. Fixed-bed adsorption is one of the ways to separate components present in a mixture [62]. Breakthrough curve modeling is used to design and analyze such systems [117]. The calculation of equilibrium loadings is an integral part of the breakthrough curve simulations. Although mixed Langmuir isotherms often provide inaccurate equilibrium loadings, especially at high pressures, these isotherms are commonly used to calculate the equilibrium loadings in breakthrough curve simulations [117, 123]. Van Assche et al. [117] have recently implemented the EI model in a process simula-

tion of a Temperature Swing Adsorption (TSA) system for the separation of a 10-component mixture. These authors are able to generate the breakthrough curves about 2.6 times faster than using IAST. In this study, we incorporate SIAST and SEI into a breakthrough curve model and compare their performance with IAST. This comparison is important because using multi-site adsorption isotherms for pure components leads to much slower IAST calculations. IAST involves two iterative processes: (1) calculation of spreading pressure and (2) calculation of the pure component pressure, which makes the calculations slow for multi-site adsorption. This will be explained in detail in the subsequent sections. As SEI is an explicit model, it will significantly speed up these calculations compared to both SIAST and IAST.

This chapter is organized as follows: the theory and derivations of different models (SIAST, EI and SEI) are provided in the Section 2.2. The simulation details are discussed in the Section 2.3. In the Section 2.4, we compare the adsorption isotherms obtained using IAST, SIAST, and SEI, and validate these isotherms with GCMC mixture simulations. We also analyze the breakthrough curves computed using these three approaches. In Section 2.5, we provide conclusions on the results obtained using IAST, SIAST, and SEI.

2.2. Theory

In this section, we discuss the theory behind the implementation of SIAST and the size-dependent explicit isotherm (EI). We explain the implementation of the segregated approach to the size-dependent explicit isotherm (SEI) and the breakthrough curve modeling approach, which is used to compare the effects of IAST, SIAST, and SEI for dynamic conditions.

2.2.1. Segregated Ideal Adsorbed Solution Theory (SIAST)

Instead of considering the available adsorption volume as a continuous space, the adsorbent material is divided into several distinct adsorption sites. The competitive adsorption at each site takes place separately. These sites can be either uniform or heterogeneous. Swisher et al. [56] have considered uniformity for each adsorption site. Each site is in separate thermodynamic equilibrium with the gas phase, as shown in Fig. 2.1. Since the sites are assumed to be uniform, IAST is applied individually to each site. The implementation of this approach is explained below. For a detailed derivation of the IAST, the reader is referred to Refs.[52, 54]

Consider a mixture with N_c components. A multi-site Langmuir isotherm, which is composed of several single-site Langmuir isotherms, is used to calculate the overall adsorbed loadings for the pure components. At each site, one of these single-site Langmuir isotherms is used to calculate the adsorbed loadings of the pure components. For the i-th component at the

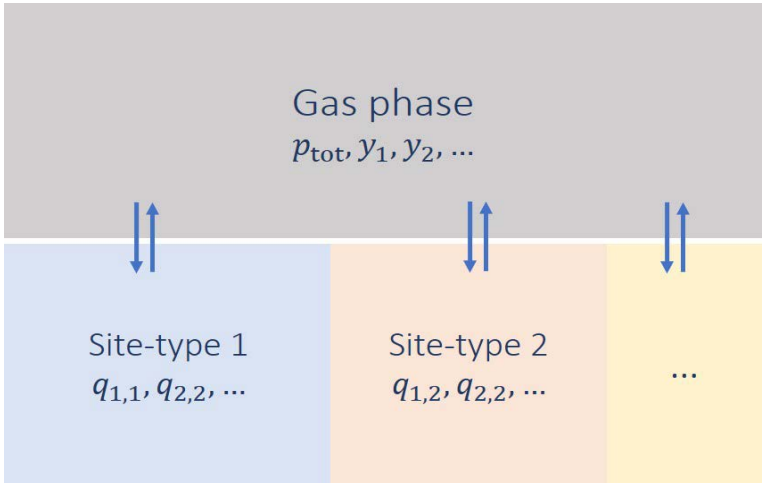


Figure 2.1. Equilibrium of each of the adsorbed phase with the gas phase in the Segregated Ideal Adsorbed Solution Theory (SIAST) model [56]. Each adsorbed phase is in equilibrium with the gas phase. The system is at a constant temperature. The gas phase has a total pressure of p_{tot} and the mole fraction of component i equals y_i. In the adsorbed phase j, the loading of component i is $q_{i,j}$.

j-th site, the Langmuir adsorption isotherm is

$$q_{i,j} = q_{max,i,j} \frac{k_{i,j} p}{1 + k_{i,j} p} \tag{2.1}$$

where $q_{i,j}$ is the adsorbed loading and $q_{max,i,j}$ is the saturation loading of the component i. The units of both $q_{i,j}$ and $q_{max,i,j}$ can be either [mol/(kg framework)] or [molecules/(unit cell)]. $k_{i,j}$ is the equilibrium constant in units of [1/Pa]. The parameters ($q_{max,i,j}$ and $k_{i,j}$) are estimated by fitting the multi-site Langmuir isotherm to the equilibrium loading data obtained using experiments or GCMC simulations. The correspondence between the adsorption sites and the isotherm parameters can be identified using snapshots of molecules present inside the adsorbents, which are obtained using GCMC simulations. In case of only one type of adsorption site, SIAST simply transforms into IAST. In the presence of multiple adsorption sites, adsorption isotherms often show inflection behavior [49, 124, 125]. Inflection behavior can be observed during the adsorption of branched alkanes inside MFI-type zeolite [124]. This zeolite consists of straight and zigzag channels and intersections connecting these channels. Branched alkanes preferentially adsorb at the intersections [124]. It is energetically demanding for these molecules to adsorb inside the channels (possible only at high pressure), which is due to the presence of branches [124]. When a mixture

includes such components with strong affinity for certain adsorption sites, models with a segregated approach, like SIAST, should be implemented to study adsorption.

IAST and SIAST involve the calculation of the spreading pressure (π) for each component of the mixture. It is analogous to pressure but in two dimensions [54]. The spreading pressure (π) of the i-th component at the j-th site is calculated as follows [54]

$$\pi_{i,j} = \frac{RT}{A} \int_0^{p_{\text{pure},i,j}} \frac{q_{\text{pure},i,j}(p)}{p} \, \mathrm{d}p \tag{2.2}$$

where $q_{\text{pure},i,j}(p)$ is the adsorbed loading of the pure component. p is the pressure and the upper limit of the integral, $p_{\text{pure},i,j}$ is the pressure of the pure component in the gas phase required to reach the spreading pressure, $\pi_{i,j}$. It is convenient to work with surface potential (Ψ) which is defined as

$$\Psi_{i,j} = \frac{\pi_{i,j} A}{RT} \tag{2.3}$$

In Eq. 2.2, we substitute $\pi_{i,j}$ with $\Psi_{i,j}$ using Eq. 2.3 and replace $q_{\text{pure},i,j}$ using Eq. 2.1. This leads to

$$\Psi_{i,j} = q_{\text{max},i,j} \ln\left(1 + k_{i,j} p_{\text{pure},i,j}\right) \tag{2.4}$$

According to IAST, the surface potential of the mixture (Ψ_{mix}) is equal to the individual surface potential of the pure components.

$$\Psi_{\text{mix},j} = \Psi_{1,j} = \Psi_{2,j} = \ldots = \Psi_{N_c,j} \tag{2.5}$$

Next, we apply Raoult's law analogy to each adsorption site, which yields

$$p_i = x_{i,j} p_{\text{pure},i,j}\left(\Psi_{i,j}\right) \tag{2.6}$$

In Eq. 2.6, p_i is the partial pressure of component i in the gas phase and $x_{i,j}$ is the mole fraction of component i in the adsorbed phase. At each site j, an independent mass balance is applicable for N_c number of components.

$$\sum_{i=1}^{N_c} x_{i,j} = 1 \tag{2.7}$$

Since the pure component pressures, $p_{\text{pure},i,j}$ cannot be calculated a priori, $\Psi_{i,j}$ in Eq. 2.4 is obtained using iterative schemes such as the Newton-Raphson [126] and the bisection method [126]. The bisection method is used in this work as it can generate values correct up to machine precision, which is not possible in practice with the Newton-Raphson method. Moreover, the Newton-Raphson method is highly sensitive to the initial conditions, and eventually, the algorithm can easily lead to divergence of the solution. The

magnitude of $x_{i,j}$ is calculated using the inverse of Eq. 2.4. In this equation, $p_{\text{pure},i,j}$ is substituted using Eq. 2.6. If the analytic inversion of the surface potential (the inverse of Eq. 2.4) does not provide an explicit expression at each adsorption site, then SIAST will not perform better than IAST.

The total loading on the j-th site equals [54]

$$q_{\text{tot},j} = \left[\sum_{i=1}^{N_c} \frac{x_{i,j}}{q_{\text{pure},i,j}} \right]^{-1} \tag{2.8}$$

The loading of component i is calculated as follows

$$q_{i,j} = x_{i,j} q_{\text{tot},j} \tag{2.9}$$

The total loading for the i-th component (q_i) encompassing all the sites is

$$q_i = \sum_{j=1}^{N_{\text{sites}}} q_{i,j} \tag{2.10}$$

2.2.2. Explicit Isotherms (EI)

The explicit isotherm model developed by Van Assche et al. [118] considers the adsorbent as a uniform lattice subdivided into identical adsorption sites. These adsorption sites can accommodate the largest species present in the mixture. First, we will derive the equations for a binary mixture. Consider that species 1 is the largest species. The lattice is divided into M_1 sites for the adsorption of species 1. The number of ways to adsorb N indistinguishable molecules of species 1 is:

$$\frac{(M_1)!}{(M_1 - N)!N!} \tag{2.11}$$

Figs. 2.2a and 2.2b show the adsorption of N indistinguishable molecules of species 1 on an adsorbent which is subdivided into M_1 uniform lattice sites. Next, species 2 will be adsorbed in the lattice. Species 2 is the smallest of the two components. It is considered to be n times smaller than species 1. After adsorption of species 1, the remaining sites in the lattice $(M_1 - N)$ are further subdivided. Each site is now divided into n parts. Therefore, $(nM_1 - nN)$ sites are available for the adsorption of species 2 (Fig. 2.2c). We rewrite nM_1 as M, which is the total number of available sites for adsorption of species 2 in the absence of species 1. K molecules of species 2 are adsorbed on the remaining sites, which is shown in Fig. 2.2d. The number of ways for this arrangement, considering the indistinguishable nature of the molecules, is

$$\frac{(M - nN)!}{(M - nN - K)!K!} \tag{2.12}$$

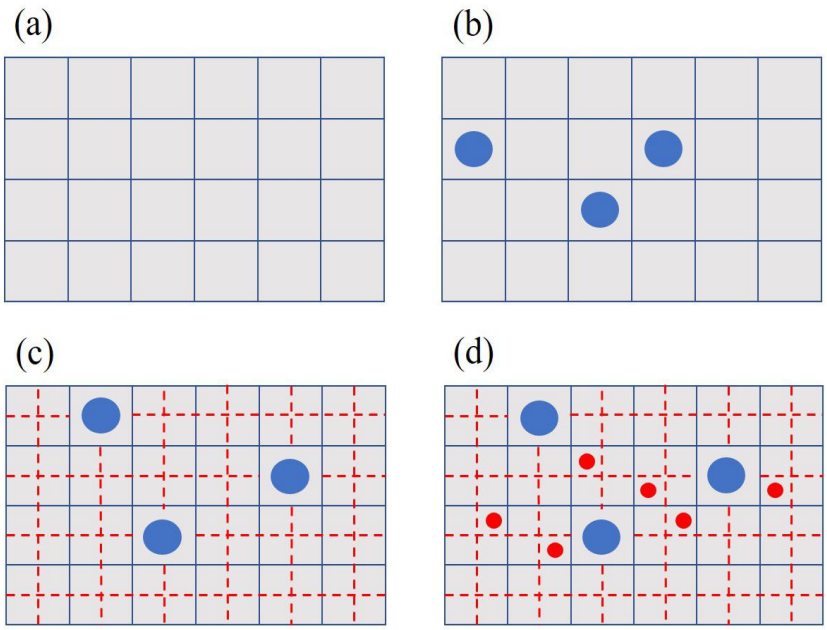


Figure 2.2. Representation of the adsorbent as a lattice where species of different sizes are adsorbed [118]. (a) The adsorbent is divided into M_1 sites. (b) adsorption of N indistinguishable molecules of species 1 is taking place on the lattice, which are represented by the blue circles. (c) After the adsorption of N molecules of species 1, the remaining lattice sites are further divided into $(nM_1 - nN)$ sites. (d) adsorption of K indistinguishable molecules of species 2 is taking place on the lattice. The blue circles represent species 1, and the smaller red circles represent species 2.

The canonical partition function Z for this mixture can be written as [118]

$$Z = \frac{M_1!}{(M_1 - N)!N!} \frac{(M - nN)!}{(M - nN - K)!K!} f_1^N f_2^K \tag{2.13}$$

where f_1 and f_2 are the molecular partition functions for individually adsorbed molecules. The relation between the canonical function and the partial pressures of the components is [62, 118].

$$\frac{\mu_1}{k_B T} = \frac{\mu_1^0}{k_B T} + \ln\left(\frac{p_1}{p^0}\right) = -\left(\frac{\partial \ln(Z)}{\partial N}\right)_{K,M,T} \tag{2.14}$$

$$\frac{\mu_2}{k_B T} = \frac{\mu_2^0}{k_B T} + \ln\left(\frac{p_2}{p^0}\right) = -\left(\frac{\partial \ln(Z)}{\partial K}\right)_{N,M,T} \tag{2.15}$$

where μ_i^0 is the chemical potential of component i in the pure form at pressure (p^0). The reference pressure, p^0 is considered to be 1 bar. p_i is the partial pressure of component i in the mixture. To introduce the equilibrium constants (k_i) for the adsorption of the pure components 1 and 2 in Eqs. 2.14 and 2.15, we use the following equations

$$k_1 = f_1(T) \exp\left(\frac{\mu_1^0}{k_B T}\right) \tag{2.16}$$

$$k_2 = f_2(T) \exp\left(\frac{\mu_2^0}{k_B T}\right) \tag{2.17}$$

We substitute the canonical function Z in Eqs. 2.14 and 2.15 using Eq. 2.13. These equations are further modified by substituting μ_i^0 using Eqs. 2.16 and 2.17 and applying Stirling's approximation to the right-hand side of the equations [118]. The reader is referred to the Supporting Information of Ref. [118] for a detailed derivation. The fractional loadings $\left(q_i/q_{\max,i,j}\right)$ of the components are defined as $\theta_1 = nN/M$ and $\theta_2 = K/M$. Using Eqs. 2.14-2.17, we obtain the following equations for the isotherms:

$$k_1 p_1 = \frac{\theta_1 (1 - \theta_1)^{n-1}}{(1 - \theta_1 - \theta_2)^n} \tag{2.18}$$

$$k_2 p_2 = \frac{\theta_2}{(1 - \theta_1 - \theta_2)} \tag{2.19}$$

Eqs. 2.18 and 2.19 are the implicit forms of the isotherm model. Rearranging these equations leads to the set of explicit binary adsorption isotherms:

$$q_1 = q_{\max,1} \frac{k_1 p_1}{(1 + k_2 p_2)^n + k_1 p_1} \tag{2.20}$$

$$q_2 = q_{\max,2} \frac{k_2 p_2 (1 + k_2 p_2)^{n-1}}{(1 + k_2 p_2)^n + k_1 p_1} \tag{2.21}$$

This model can be extended to an arbitrary number of components (N_c). The following condition is imposed for the saturation capacities:

$$q_{\max,N_c} \geq q_{\max,N_c-1} \geq \cdots \geq q_{\max,2} \geq q_{\max,1} \tag{2.22}$$

Following a similar derivation, for a mixture of N_c components, the isotherm for the i-th component is as follows [118]

$$q_i = (q_{\max,1} k_i p_i) \frac{\left[\prod_{m=1}^{i} \alpha_m\right]}{\beta} \tag{2.23}$$

where

$$\alpha_1 = 1 \tag{2.24}$$

$$\alpha_i = \left[\left(\cdots\left(\left(\left(1 + k_{N_c} p_{N_c}\right)^{n_{N_c}} + k_{N_c-1} p_{N_c-1}\right)^{n_{N_c-1}} + k_{N_c-2} p_{N_c-2}\right)^{n_{N_c-2}} \cdots + k_{i+1} p_{i+1}\right)^{n_{i+1}} + k_i p_i\right]^{n_i - 1} \tag{2.25}$$

$$\beta = \left[\left(\cdots\left(\left(\left(1 + k_N p_{N_c}\right)^{n_{N_c}} + k_{N_c-1} p_{N_c-1}\right)^{n_{N_c-1}} + k_{N_c-2} p_{N_c-2}\right)^{n_{N_c-2}} + \cdots + k_2 p_2\right)^{n_2} + k_1 p_1\right] \tag{2.26}$$

In Eqs. 2.25 and 2.26, the value of n_i for component i is $q_{\max,i}/q_{\max,i-1}$.

2.2.3. Segregated Explicit Isotherms (SEI)

Similar to SIAST, we also consider the equilibrium of each adsorption site separately with the gas phase, as shown in Fig. 2.3. The isotherms derived in section (Explicit Isotherm) are solved separately for adsorption site j.

$$q_{i,j} = \left(q_{\max,i,j} k_{i,j} p_i\right) \frac{\left[\prod_{m=1}^{i} \alpha_{m,j}\right]}{\beta_j} \tag{2.27}$$

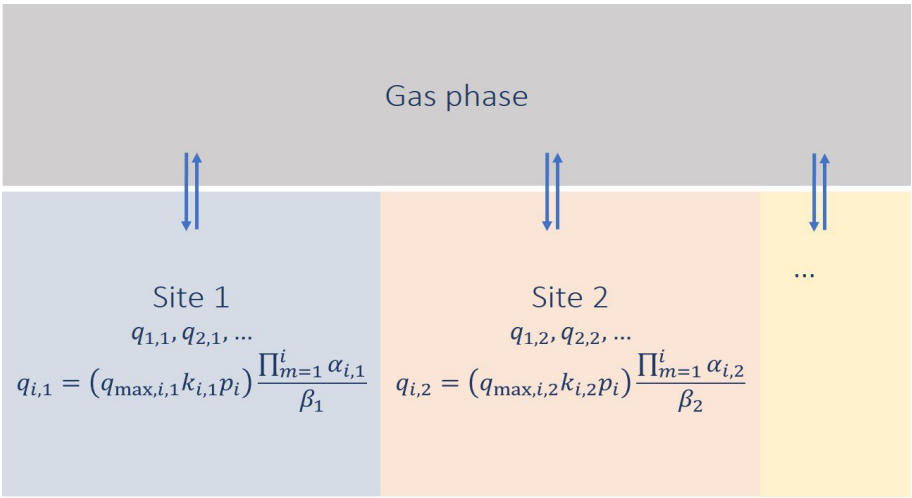


Figure 2.3. Representation of Segregated Explicit Isotherm (SEI) model, which accounts for the separate thermodynamic equilibria of the adsorbed phases with the gas phase at different adsorption sites. Each adsorbed phase is separately in thermodynamic equilibrium with the gas phase, and it is represented by the adsorption isotherm proposed by Van Assche et al. [118] (Eq. 2.23). The gas phase has a total pressure of p_{tot}, and the gas phase mole fractions of the components in the mixture are represented by y_i. In the adsorbed phase j, the loading of the component i is $q_{i,j}$.

The overall loading (q_i) for component i is obtained by summing over the loadings for the component calculated at each adsorption site.

$$q_i = \sum_{j=1}^{N_{\text{sites}}} q_{i,j} \tag{2.28}$$

In Eq. 2.28, N_{sites} is the number of sites available for adsorption, which is the same for all components. The procedure for calculating the equilibrium loadings using SEI is shown in Algorithm 1. In this algorithm, Eqs. 2.23-2.26 are shown recursively.

2.2.4. Breakthrough Curve Model

Breakthrough curves reflect the concentration profiles of the components in a mixture at the outlet of the fixed bed adsorption column [127]. modeling breakthrough curves involves the mass balance for different species present in the mixture. We write the mass transport equation for each component in terms of their partial pressures in the gas phase (p_i). We assume the ideal gas law to be valid. Considering plug flow and isothermal conditions and neglecting axial dispersion, we obtain the following equation for mass transport [55, 128].

$$\frac{\partial p_i(t,z)}{\partial t} = -\frac{\partial\left(v(t,z)\,p_i(t,z)\right)}{\partial z} - RT\left(\frac{1-\varepsilon_b}{\varepsilon_b}\right)\rho_p \frac{\partial q_i(t,z)}{\partial t} \tag{2.29}$$

In Eq. 2.29, p_i is the partial pressure of component i in the gas phase in [Pa], v is the interstitial velocity in [m/s]. The fixed bed void fraction is defined by ε_b, ρ_p is the adsorbent material density in $\left[\text{kg/m}^3\right]$, and z is the position along the column. The interstitial velocity is calculated using the total mass balance. We write the total mass balance in terms of the total pressure (p_{tot}) and impose isobaric conditions in the column [129].

$$\frac{\partial\left(v(t,z)\right)}{\partial z} = -\frac{RT}{p_{\text{tot}}}\left(\frac{1-\varepsilon_b}{\varepsilon_b}\right)\rho_p \sum_{i=1}^{N_c} \frac{\partial q_i(t,z)}{\partial t} \tag{2.30}$$

The term $\frac{\partial q_i(t,z)}{\partial t}$ in Eqs. 2.29 and 2.30 accounts for the mass transfer of the species from the gas phase to the adsorbed phase. It is often modeled using the Linear Driving Force model (LDF) [130, 131]

$$\frac{\partial q_i(t,z)}{\partial t} = k_{M,i}\left(q_{\text{eq},i} - q_i\right) \tag{2.31}$$

The LDF model states that the rate of adsorption is proportional to the amount of adsorbate still required to achieve equilibrium between the gas and the adsorbed phase [132]. In Eq. 2.31, the mass transfer coefficient

Algorithm 1 Procedure for calculating the equilibrium loadings for components in mixtures using Segregated Explicit Isotherm (SEI).

Arrange the components in ascending order based on the value of $q_{\max}$ for the j-th adsorption site.
$q_{\max,1,j} \leq q_{\max,2,j} \cdots \leq q_{\max,N_c-1,j} \leq q_{\max,N_{c,j}}$

Calculate the ratios of $q_{\max,i,j}$ $\left(n_{i,j}\right)$ for each component i
$n_{i,j} = \frac{q_{\max,i,j}}{q_{\max,i-1,j}}$
$n_{1,j} = 1$

$\alpha_{i,j} = \left(\alpha_{i+1,j} + k_{i,j}p_{i,j}\right)^{n_{i,j}}$
$\alpha_{N_c+1,j} = 1$

Equilibrium loading for each component at the j-th site
$q_{i,j} = q_{\max,i,j}k_{i,j}p_{i,j}\frac{1}{\alpha_{1,j}}\prod_{m=1}^{i}\frac{\alpha_{m,j}}{\alpha_{m+1,j}+k_{m,j}p_{m,j}}$

Rearrange the adsorbed loadings in the original order as provided in the input.

Overall adsorbed loading for component i (q_i)
$q_i = \sum_{j=1}^{N_{\text{sites}}} q_{i,j}$

Return q_i

is defined by $k_{M,i}$ in $[1/s]$. $q_{eq,i}$ is the equilibrium loading and q_i is the adsorbed loading of component i. The units of $q_{eq,i}$ and q_i are in [mol/(kg framework)]. To calculate the magnitude of $q_{eq,i}$, an adsorption model must be integrated with the breakthrough curve model. We incorporated both SIAST and SEI into the breakthrough model and made a detailed comparison in terms of accuracy and computational performance.

2.3. Simulation Details

We consider two case studies of adsorption of binary equimolar (50:50) mixtures:

1. carbon dioxide (CO_2) and propane (C_3) on MOR-type zeolite at 300 K [56].
2. butane (n-C_4) and 2-methyl propane (2-m-C_3) on MFI-type zeolite at 400 K.

To calculate the mixture isotherms, we first need to generate the pure component isotherms. Adsorption loadings for the pure components are

calculated using GCMC simulations [50], which are performed using the RASPA2 software [39, 133]. The non-bonded intermolecular interactions (both guest-guest and guest-host) are modeled by Lennard-Jones potentials and electrostatic interactions. The latter are computed using the Ewald summation [134]. The Lennard-Jones parameters and the partial charges are obtained from various sources. For CO_2, these parameters are taken from the work of Garcia-Sanchez et al. [135]. The force field of Dubbeldam et al. [136] is used for C_3, n-C_4, and 2-m-C_3 molecules. To account for the interaction between the adsorbed molecules and the adsorbent, the TraPPE-zeo force field is used [43]. Both zeolites (MOR-type and MFI-type) are pure silicalites and do not contain any extra-framework cations. Short-range interactions were truncated and shifted to zero at 12 Å. CO_2 is considered a rigid molecule, so the bonded interactions are not considered. The bonded interactions for the hydrocarbons are modeled using the united atom TraPPE force fields [37]. In the first study, the simulation box consists of 16 unit cells $(2 \times 2 \times 4)$, and 8 unit cells $(2 \times 2 \times 2)$ in the second case. Ideal gas behavior is assumed for the gas phase. For a pure component in the gas phase, the fugacity is equal to the pressure of the pure component (p). For mixtures, the fugacity of component i is equal to the total pressure multiplied by the gas phase mole-fraction (y_i).

The error bars present in the GCMC simulation data are significantly less than ca. 2%. The datasets are fitted to the multi-site Langmuir isotherms to obtain the isotherm parameters, saturation loadings $\left(q_{\max,i,j}\right)$ and equilibrium constants $\left(k_{i,j}\right)$. To obtain the correct loadings, it is crucial to identify which isotherm parameters correspond to which adsorption site. This is because the preference for the adsorption sites for different components in the mixture may vary. A molecule of an adsorbed species will experience competition only from those molecules of the other species that are adsorbed at the same site. The correspondence between the adsorption sites and the isotherm parameters is identified using simulation snapshots. The data sets for the snapshots are generated during the GCMC simulations performed with the RASPA software [39, 133]. The iRASPA software is used to create the images using these data sets [22].

$q_{\max,i,j}$, $k_{i,j}$, and the partial pressures (p_i) of the components in the gas phase are the main input parameters for the mixture isotherm models. The mixture isotherms are computed using IAST, SIAST and SEI. In IAST and SIAST, the system of implicit equations is solved using the bisection method [126], which enables us to compute the equilibrium loadings correctly up to machine precision. This is essential when IAST or SIAST is incorporated into breakthrough curve models. A small difference in the equilibrium loading can lead to a breakthrough curve quite different from the actual curve. For validation, we have also calculated the adsorbed loadings for the mixtures using GCMC simulations. Comparisons are made between the adsorption isotherms generated using these four methods.

For the breakthrough curve modeling, Eqs. 2.29-2.31 are solved simulta-

neously. Initially, the column is filled with only a non-adsorbing carrier gas (helium). At the start of the breakthrough simulation, the pressure of the carrier gas is equal to the total pressure inside the column $\left(p_{\text{carrier gas}} = p_{\text{tot}}\right)$. The adsorber column is considered to be isothermal and isobaric. The initial conditions are:

$$v(t=0,z) = v_{\text{in}} \tag{2.32}$$
$$p_{\text{tot}}(t=0,z) = p_{\text{tot,in}} \tag{2.33}$$
$$p_{\text{carrier gas}}(t=0,z) = p_{\text{tot}}(t=0,z) \tag{2.34}$$
$$p_i(t=0,z>0) = 0 \tag{2.35}$$
$$q_i(t=0,z) = 0 \tag{2.36}$$

where v is the interstitial velocity and v_{in} is the interstitial velocity at the inlet of the adsorber. p_{tot} is the total pressure of the adsorber column and the total pressure at the inlet is referred to as $p_{\text{tot,in}}$. $p_{\text{carrier gas}}$ is the partial pressure of the carrier gas. p_i and q_i are the partial pressure and adsorbed loadings of the component i, respectively. At the inlet of the column, the partial pressures for each component and the velocity are fixed. At the outlet, the spatial gradients of the partial pressures are considered to be zero. The boundary conditions for the simulations are as follows

$$v(t,z=0) = v_{\text{in}} \tag{2.37}$$
$$p_{\text{tot}}(t,z=0) = p_{\text{tot,in}} \tag{2.38}$$
$$p_i(t,z=0) = y_{i,\text{in}} \times p_{\text{tot,in}} \tag{2.39}$$
$$\frac{\partial p_i(t,z=L)}{\partial z} = 0 \tag{2.40}$$

Eqs. 2.29-2.31 form a system of differential algebraic equations. The method of lines [68] is used to solve these equations numerically. The spatial derivatives present in the partial differential equations (PDEs) are discretized using the Finite Difference Method (FDM) [137]. The advective terms $(\partial(vp_i)/\partial z)$ and $(\partial(v)/\partial z)$ in Eqs. 2.29 and 2.30, respectively, are discretized using a first-order upwind scheme [138]. The Strong Stability Preserving Runge Kutta (SSP-RK(3,3)) method is used for the time integration [69, 139]. This is a third-order scheme and involves three stages in the integration [140].

The input parameters for the simulations are shown in Table 2.1, and the mass transfer coefficients for the components of the mixtures are shown in Table 2.2. The mass transfer coefficients are assumed to be identical since the molecular sizes are comparable. The length of the adsorber column for the first case is considered to be 0.1 m and 0.8 m for the second case. The lengths of the columns are discretized with a uniform grid size, Δz. For the first case, Δz is 0.005 m and 0.05 m for the second case. A time step (Δt = 0.001 s) was chosen for the time integration in both cases.

Table 2.1. Input parameters for the simulation of the breakthrough curve (catalyst particle density, (ρ_p), bed voidage (ε_b), interstitial velocity at the inlet, (v_{in}), and total pressure (p_{tot})) for both cases.

	ρ_p/ [kg/m^3]	ε_b	v_{in}/ [m/s]	p_{tot}/ [bar]
$CO_2 - C_3$	1711.06	0.128	0.1	1.0
n-C_4 − 2-m-C_3	1796.34	0.40	0.1	10.0

Table 2.2. Mass transfer coefficients (k_M) for the components in both the case studies.

	CO_2	C_3	n-C_4	2-m-C_3
k_m [1/s]	0.06	0.06	0.06	0.06

For the simulations, the following High Performance Computing (HPC) facilities were used: (1) Dutch National Supercomputer (Snellius), and (2) a local supercomputer at Delft University of Technology (HAL9000).

2.4. Results and Discussions

In this section, we discuss the adsorption isotherms obtained for pure components and their mixtures in both studies (CO_2-C_3 mixture in MOR-type zeolite at 300 K, and an n-C_4-2-m-C_3 mixture in MFI-type zeolite at 400 K). We compare the results predicted by IAST, SIAST and SEI. For validation, mixture isotherm data using GCMC simulations are generated. The influence of the different models (SIAST and SEI) on the breakthrough curves is also investigated.

2.4.1. Adsorption Isotherms

CO_2-C_3 mixture, MOR-type Zeolite, 300 K

Pure component loadings are calculated using GCMC simulations and fitted to dual-site Langmuir isotherms. This is because there are two possible distinct sites for adsorption inside the MOR-type zeolite. The pure component isotherms are shown in Fig. 2.4 and the corresponding fitted parameters are listed in Table 2.3. In the dual-site Langmuir isotherm, we have two pairs of fitted parameters, which are saturation loading $(q_{max,i,j})$ and equilibrium constant $(k_{i,j})$ at each site. The GCMC data generated in this work are in excellent agreement with the data obtained by Swisher et al. [56]. To identify the correspondence between the pairs of fitted parameters and the adsorption sites, we have generated the snapshots of adsorption of the pure components inside the MOR-type zeolite using the iRASPA software [22].

Figs. 2.5a and 2.5b show typical snapshots of pure CO_2 inside MOR-type at 300 K. The snapshots are taken at 10^3 Pa and 10^9 Pa, respectively.

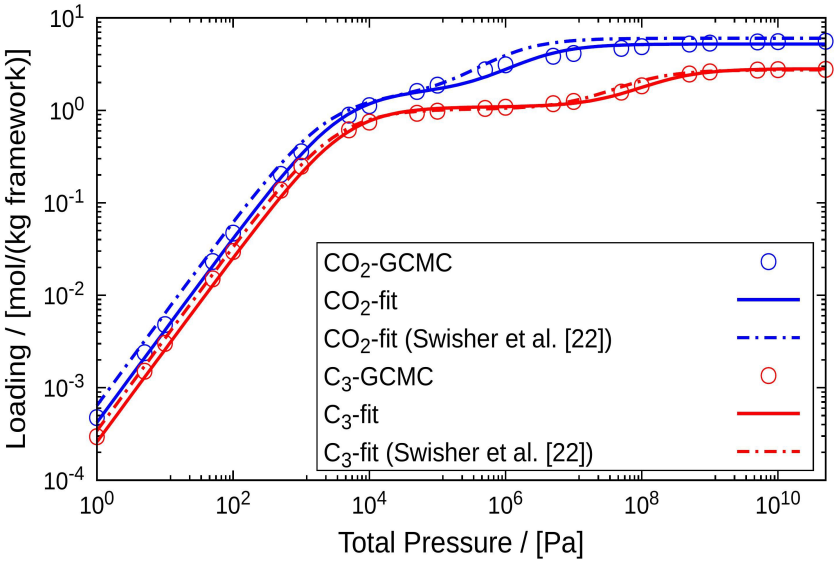


Figure 2.4. Pure component adsorption isotherms of CO_2 and C_3 in MOR-type zeolite at 300 K. Pure component adsorbed loadings calculated using GCMC for the pressure range $\left(10^0 - 5 \times 10^{10}\right)$ Pa are fitted to the dual-site Langmuir isotherms. The empty circles represent GCMC simulation data, and the solid lines are the dual-site Langmuir isotherms fitted to these data sets. The results obtained in this work are in excellent agreement with the data published by Swisher et al. [56].

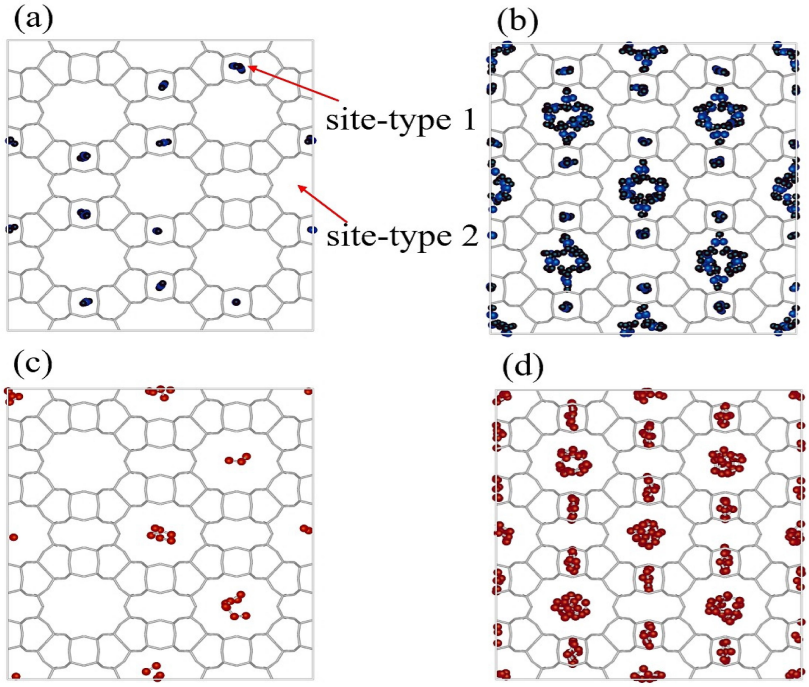


Figure 2.5. Snapshots of adsorption of pure CO_2 and C_3 molecules in MOR-type zeolite at 300 K using GCMC simulations. (a) adsorption of CO_2 at 10^3 Pa, (b) adsorption of CO_2 at 10^9 Pa, (c) adsorption of C_3 at 10^3 Pa and (d) adsorption of C_3 at 10^9 Pa. Site 1 represents the smaller pockets and site 2 represents the larger pockets inside MOR-type zeolite

Table 2.3. Fitted parameters for the adsorption of pure CO_2 and C_3 in MOR-type zeolite at 300 K. Adsorbed loadings obtained from the GCMC simulations for the pressure range $(10^0 - 5 \times 10^{10})$ Pa are fitted using the dual-site Langmuir isotherm. Subscript 1 indicates site 1 (smaller pockets) and subscript 2 indicates site 2 (larger pockets) present inside MOR-type zeolite. (See also Fig. 2.5.)

	k_1 [1/Pa]	$q_{max,1}$ $\left[\frac{\text{mol}}{\text{kg framework}}\right]$	k_2 [1/Pa]	$q_{max,2}$ $\left[\frac{\text{mol}}{\text{kg framework}}\right]$
CO_2	2.617×10^{-4}	1.60	4.859×10^{-7}	3.62
C_3	6.506×10^{-9}	1.73	2.376×10^{-4}	1.09

Table 2.4. Fitted parameters for the adsorption of pure n-C_4 and 2-m-C_3 in MFI-type zeolite at 400 K. Adsorbed loadings obtained from the GCMC simulations for the pressure range of $(10^0 - 5 \times 10^{10})$ Pa are fitted using the dual-site Langmuir isotherm. Subscript 1 indicates site 1 (intersections between the channels), and subscript 2 indicates site 2 (channels) present inside MFI-type zeolite.

	k_1 [1/Pa]	$q_{max,1}$ $\left[\frac{\text{mol}}{\text{kg framework}}\right]$	k_2 [1/Pa]	$q_{max,2}$ $\left[\frac{\text{mol}}{\text{kg framework}}\right]$
n-C_4	2.552×10^{-4}	0.70	2.348×10^{-5}	1.02
2-m-C_3	9.861×10^{-5}	0.70	1.00×10^{-7}	0.90

Similarly, Figs. 2.5c and 2.5d show the snapshots for C_3 molecules at conditions (temperature and pressure) identical to the case of CO_2. At low pressure, CO_2 preferentially adsorbs inside the smaller pockets (site 1) and C_3 adsorbs inside the larger ones (site 2). At higher pressures, once the preferred type of site is filled, the molecules start adsorbing on the remaining types of sites. Therefore, at higher pressures, both CO_2 and C_3 are found in both types of adsorbing pockets. This is clearly shown in Figs. 2.5b and 2.5d. For assigning the pairs of fitted parameters to the adsorption sites, we refer to the smaller pockets as site 1 and the larger ones as site 2. The values of these parameters are shown in Table 2.3.

Next, we compare the mixture isotherms predicted by IAST, SIAST, and SEI and validate the predictions by further comparing with the GCMC simulations for the mixture. Fig. 2.6 shows the isotherms for the mixture of CO_2 and C_3. We can observe the deviations in the adsorption isotherms predicted by the IAST from the GCMC data. The isotherms obtained using SEI and SIAST are in good agreement with the GCMC simulation data. There are very small deviations at higher pressures ($\sim 10^8$ Pa) from the GCMC calculations for both SEI and SIAST, but these deviations are much lower compared to the IAST predictions.

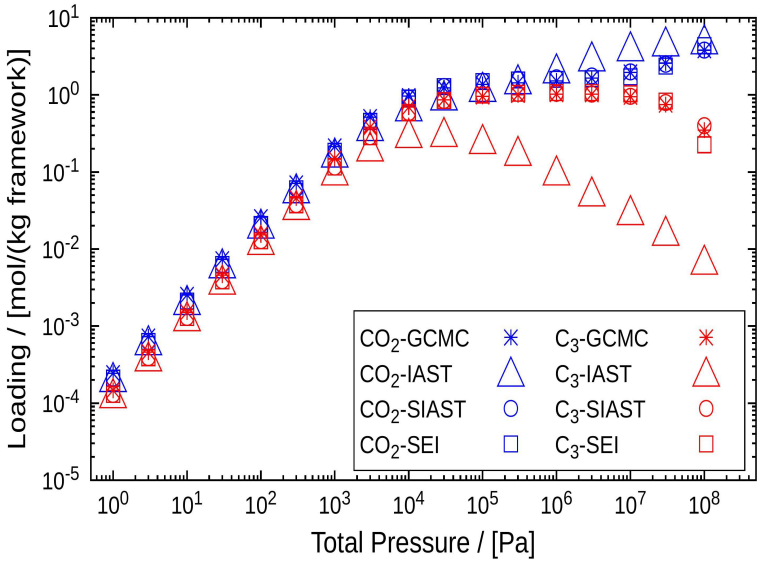


Figure 2.6. Adsorption isotherms of an equimolar mixture CO_2 and C_3 (50:50) in MOR-type zeolite at 300 K. A comparison is made between the adsorbed loadings calculated using GCMC, IAST, SIAST and SEI. The pressure range $(10^2 - 10^8)$ Pa is considered for the calculations performed using IAST, SIAST and SEI. For GCMC simulations, the pressure range is $(10^4 - 10^8)$ Pa. Cross marks represent GCMC calculations, triangles are used for IAST, circles for SIAST and squares for SEI.

The deviations can be explained based on the preference for the adsorption sites. IAST underpredicts the adsorbed loadings for C_3 molecules and overpredicts for CO_2. This is clearly because of the underlying assumption of the presence of uniform adsorbents in the IAST model. This assumption is correct only when the molecules do not have any preference for the adsorption sites. However, in this case, the components favor one site over the other. At higher pressures, the components adsorb inside both pockets, but the magnitudes of the loadings vary depending on which site is preferred more for adsorption. Consequently, the competition between the adsorbates differs at each site. In the larger pockets (site 2), the adsorption of C_3 molecules will be affected by the presence of only those CO_2 molecules that are adsorbed within the same pockets. Assuming that the adsorption of the C_3 molecules will be affected by all the CO_2 molecules, irrespective of the type of site, leads to the underestimation of the values of the adsorbed loading of C_3.

n-C_4-2-m-C_3 mixture, MFI-type Zeolite, 400 K

The same procedure is followed here as in the previous case. We first obtain the pure component isotherms using GCMC simulations and perform curve fitting on these data sets using the dual-site Langmuir isotherms. The pure component isotherms of n-C_4 and 2-m-C_3 are shown in Fig. 2.7. For assigning pairs of fitted parameters to the adsorption sites, we refer to the channels as site 1 and the intersections as site 2. The values of these parameters are shown in Table 2.4.

To identify the preferred sites for adsorption, snapshots of the pure components (n-C_4 and 2-m-C_3) inside the MFI-type zeolites are shown in Fig. 2.8. n-C_4 does not exhibit a strong preference for any of the adsorption sites (Figs. 2.8a and 2.8b). 2-m-C_3 prefers the intersection between the channels in the MFI-type zeolite (Figs. 2.8c and 2.8d). The presence of a branch (methyl group) in the hydrocarbon chain makes it energetically less favorable for 2-m-C_3 to adsorb inside the channels (Fig. 2.8c) [124]. At higher pressures $(\geq 10^6$ Pa), 2-m-C_3 starts adsorbing inside the channels (Fig. 2.8d). At the intersections (site 1) in MFI-type zeolite, the maximum possible loadings are 4 molecules/(unit cell) or ca. 0.70 mol/(kg framework). Therefore, q_{max} is assigned this value at site 1 for both n-C_4 and 2-m-C_3 (Table 2.4).

Fig. 2.9 shows the comparison between the mixture isotherm data, calculated using IAST, SIAST, SEI and GCMC. Even at higher pressures ($\sim 10^8$ Pa), the adsorbed loading values estimated by SIAST and SEI are almost identical to those generated by GCMC simulations. IAST underpredicts the value of the adsorbed loading for 2-m-C_3 at higher pressure $(\geq 10^5$ Pa). The reason for this deviation is again the preference for one adsorption site over the other. For example, at an intersection (site 2), 2-m-C_3 molecules will be affected by the adsorption of only those n-C_4 molecules which are also adsorbed on that site. IAST again fails to capture this effect.

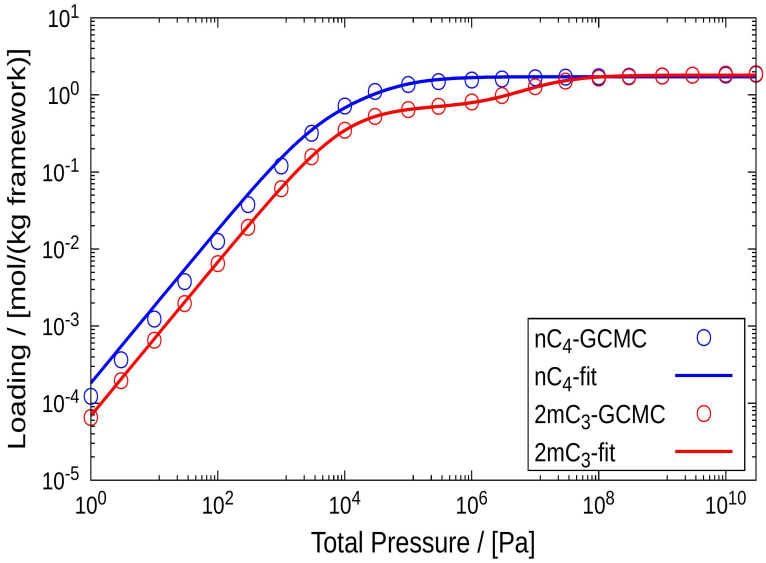


Figure 2.7. Pure component adsorption isotherms of n-C_4 and 2-m-C_3 in MFI-type zeolite at 400K. Pure component adsorbed loadings calculated using GCMC for the pressure range $(10^0 - 5 \times 10^{10})$ Pa are fitted using the dual-site Langmuir isotherms. The empty circles represent GCMC simulation data, and the solid lines are the dual-site Langmuir isotherms fitted to these data sets.

2.4.2. Breakthrough Curve Simulations

The influence of SEI and SIAST on the simulation of breakthrough curves has been studied. Fig. 2.10a shows the breakthrough curves at the exit of the adsorber column for the CO_2-C_3 mixture (10% CO_2, 10% C_3, 80% helium) in MOR-type zeolite at 300 K and 10^5 Pa. The weakly adsorbing component leaves the adsorber first. In the CO_2-C_3 mixture, C_3 is the weakly adsorbing component. It also exhibits a roll-up behavior. This occurs when the molar concentration of the adsorbing species in the gas phase exceeds its inlet concentration ($c_i/c_{\text{in},i} > 1$) [129]. Similarly, Fig. 2.10b shows the breakthrough curves for a mixture of n-C_4 and 2-m-C_3 (10% n-C_4, 10% 2-m-C_3, 80% helium) in MFI-type zeolite at 400K and 10^6 Pa. In this case, 2-m-C_3 is the weakly adsorbing species. Therefore, there is an early onset of the breakthrough curve for 2-m-C_3 compared to n-C_4.

It is clearly observed from Figs. 2.10a and 2.10b that the breakthrough curves calculated using IAST are quite different from those calculated using SIAST and SEI. Stronger differences can be observed in the case of CO_2-C_3. This is due to the highly inaccurate prediction of the adsorbed mixture loadings by IAST. Implementation of SEI and SIAST yields breakthrough curves that coincide with each other for both case studies.

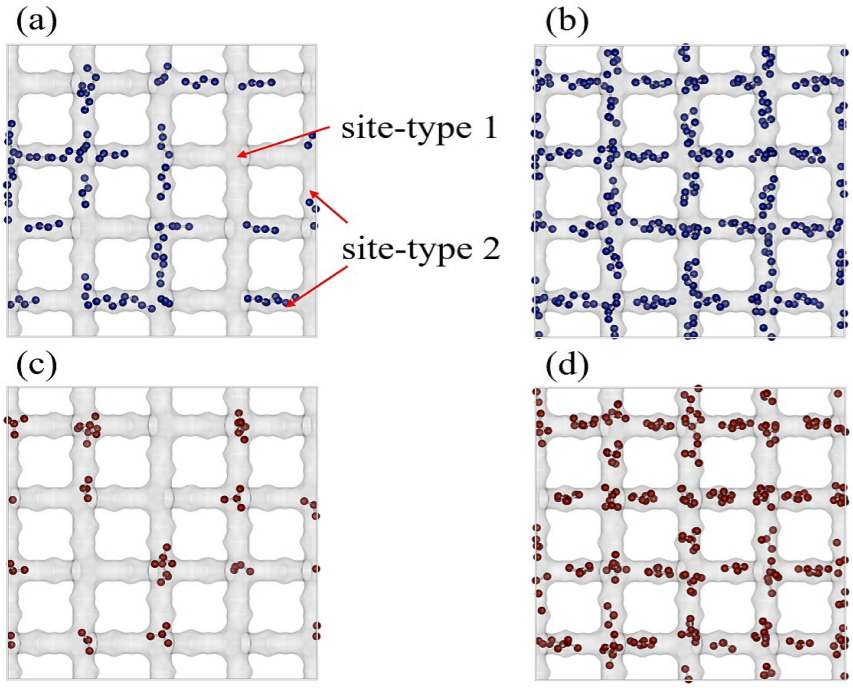


Figure 2.8. Typical snapshots of adsorption of pure C_4 and 2-m-C_3 molecules in MFI-type zeolite at 400 K using GCMC simulations. (a) adsorption of n-C_4 at 10^4 Pa, (b) adsorption of n-C_4 at 10^8 Pa, (c) adsorption of 2-m-C_3 at 10^4 Pa and (d) adsorption of 2-m-C_3 at 10^8 Pa. Site 1 represents the intersections, and site 2 represents the channels inside the MFI-type zeolite.

Table 2.5. Variation in simulation run time (t_{run}) for the breakthrough curve model on implementing SEI, SIAST and IAST. Comparisons are drawn for both case studies: $CO_2 - C_3$ (case 1) and n-C_4 − 2-m-C_3 mixtures (case 2). These simulations were run on a local supercomputer at Delft University of Technology (HAL9000).

	t_{run} [hrs] (case 1)	t_{run} [hrs] (case 2)
IAST	17.4	2.94
SIAST	0.92	0.15
SEI	0.33	0.05

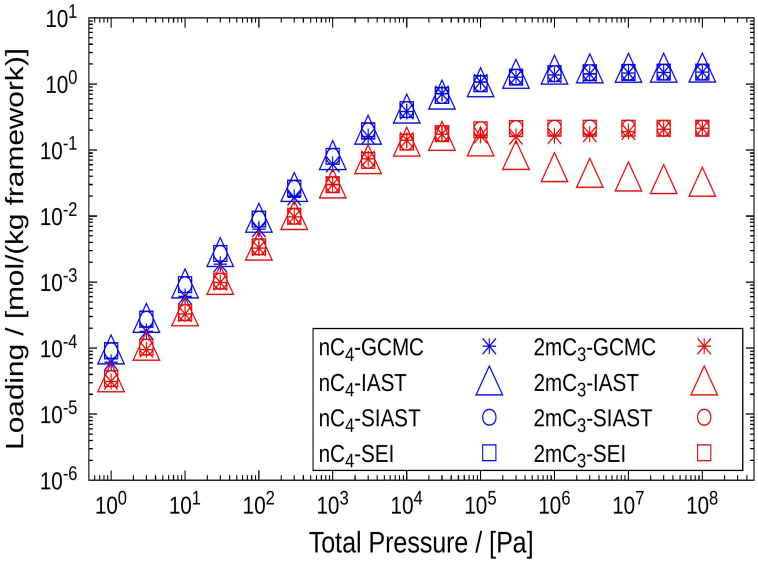


Figure 2.9. Adsorption isotherms of an equimolar mixture n-C_4 and 2-m-C_3 in MFI-type zeolite at 400K. A comparison is drawn between the adsorbed loadings calculated using GCMC, IAST, SIAST and SEI. The pressure range (10^2-10^8) Pa is considered for the calculations performed using IAST, SIAST and SEI. For GCMC simulations, the pressure range is (10^4-10^8) Pa. Cross marks represent GCMC calculations, triangles are used for IAST, circles for SIAST and squares for SEI.

It is important to consider the run time of the simulations. The variation in the run time on implementing these techniques (IAST, SIAST and SEI) to the breakthrough model is shown in Table 2.5. The breakthrough model with SEI implementations leads to the fastest simulations. This is because of the explicit nature of the SEI model. The SIAST model is slower than the SEI model but much faster than the model with IAST implementations. The breakthrough simulations using SEI are about ca. 3 times faster than the simulations using SIAST for both case studies. Again, implementing SIAST leads to calculations ca. 20 times faster than using IAST for both cases. The reason for slower IAST calculations is the use of multi-site adsorption isotherms for the pure components. IAST involves two iterative processes: (1) calculation of the spreading pressure and (2) the pure component pressure (p_{pure}), which is the inverse of the spreading pressure. The use of single-site Langmuir isotherms yields an explicit expression for the pure pressure (p_{pure}) by inverting the spreading pressure, which is not possible for multi-site Langmuir isotherms. In the absence of an explicit expression, one has to adopt an iterative scheme, such as bisection or the Newton-Raphson method, to solve for p_{pure}. This makes the IAST calculations time-consuming. In the

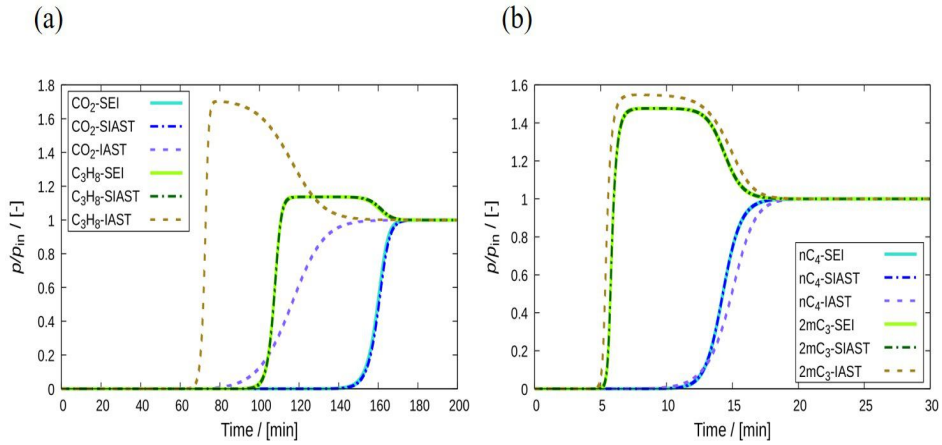


Figure 2.10. Comparison of breakthrough curves obtained on implementing IAST, SIAST and SEI to the breakthrough curve model. Separation of (a) CO_2-C_3 mixture using MOR-type zeolite at 300 K and 10^5 Pa, (b) n-C_4-2-m-C_3 mixture using MFI-type zeolite at 400 K and 10^6 Pa. Each component constitutes 10% of the mixture in the gas phase. The remaining amount is a non-adsorbing carrier gas (helium). Solid lines represent the implementation of SEI. Dashed-dotted lines are used for the SIAST model, and dashed lines are used for the IAST model.

SIAST model, we do not consider the multi-site isotherm at once. Rather, at each site, a part of the multi-site isotherm is used, which is a single-site Langmuir isotherm in itself and for which there is an analytic expression for the inverse of the spreading pressure. As an explicit expression for the pure component pressure is available for each adsorption site, only the spreading pressure (Ψ) is computed iteratively in the SIAST model. This reduces the simulation run-time significantly compared to IAST. That is why the simulation speed for SIAST and SEI differ only by a factor of ca. 3. If the spreading pressure function cannot be inverted to an explicit expression at each site, then SIAST will also involve the above-mentioned two iterative processes. Consequently, SIAST will lose its advantage over IAST in terms of the speed of calculations. The simulation run time shown in Table 2.5 is obtained for IAST and SIAST when the bisection method is used for solving the system of implicit equations.

2.5. Conclusions

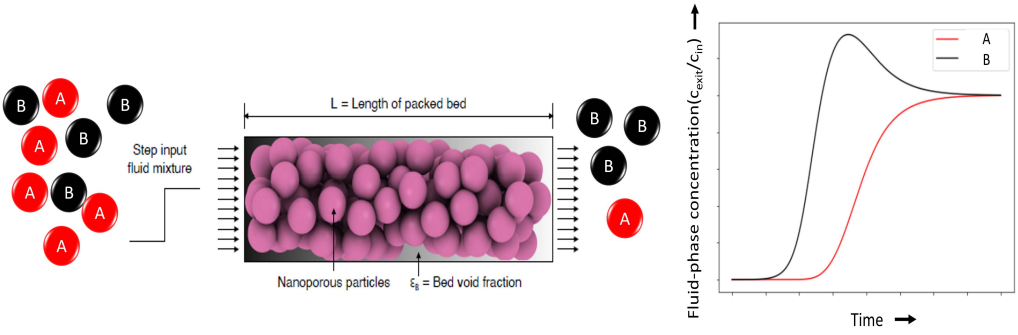
A comparison between Ideal Adsorbed Solution Theory (IAST), Segregated Ideal Adsorbed Solution Theory (SIAST), and Segregated Explicit Langmuir (SEI) has been performed by calculating the adsorption isotherms and breakthrough curves for the case studies: (1) equimolar mixture of CO_2-C_3 in MOR-type zeolite at 300 K, and (2) equimolar mixture of n-C_4-2-m-C_3 mixture in MFI zeolite at 400 K. The mixture isotherms predicted by SEI and SIAST are in excellent agreement with the GCMC simulation data. When the adsorbents have different adsorption sites and the components in the mixture prefer certain sites over the others, SIAST and SEI provide much better predictions of adsorbed loadings than IAST. We have observed this in both case studies. IAST underestimates the adsorption loadings for the less adsorbing components (C_3 in case 1, and 2-m-C_3 in case 2) and over-predicts the counterparts with stronger affinity for adsorption. This is due to the assumption of uniform adsorbents in the IAST model. Due to this assumption, IAST cannot capture the actual competition experienced by different components at each site. In case study 1, CO_2 prefers the smaller pockets inside MOR-type zeolite, whereas C_3 prefers the larger pockets. At higher pressures, when both pockets are filled with CO_2 and C_3, the competitive adsorption in each type of pocket is different due to their preferences. Similar observations are made in the 2^{nd} case study. IAST once more underpredicts the adsorbed loadings of 2-m-C_3 in MFI-type zeolite. 2-m-C_3 preferentially adsorbs at the intersections between the channels inside MFI-type zeolite. At higher total pressure ($p_{tot} > 10^6$ Pa), 2-m-C_3 starts adsorbing inside the channels. In sharp contrast, n-C_4 does not have a strong preference for any of the two sites in MFI-type zeolite (channels and intersections). If the adsorbents have very distinct adsorption sites and the adsorbates prefer a certain site over others, the predicted loadings by IAST will generally not be accurate. SEI and SIAST can provide better estimations

compared to IAST. The major advantage of SEI is that it involves only explicit equations, which makes it computationally much cheaper and improves numerical stability. This is beneficial when equilibrium loadings need to be calculated within a breakthrough curve model, both in terms of speed and accuracy. In this study, we have observed that the breakthrough curve simulations with the SEI model are ca. 3 times faster than the simulations with the SIAST model. However, the major improvement in the run time of the simulations is observed between SIAST and IAST when multi-site isotherms are used for the pure components. The breakthrough model with SIAST implementations is found to be about 20 times faster than implementing IAST. This is because IAST involves two iterative processes: (1) calculation of the spreading pressure (Ψ), and (2) calculation of the pure component pressure (p_{pure}), which is the inverse of the spreading pressure function. In case of SIAST, the inverse of Ψ is generally an explicit function for each type of adsorption site. Hence, no iterations are required to calculate p_{pure}. If the inverse of Ψ at each adsorption site does not lead to an explicit function, SIAST will lose its advantage over IAST in terms of the speed of calculations. The enhancement in the simulation run time on implementing SEI and SIAST will be very beneficial in screening a large number of adsorbents for a certain separation process, which is otherwise very time-consuming on implementing IAST. This comparison is valid only for adsorbents with multiple distinct adsorption sites. For adsorbents with a single type of adsorption site, IAST and SIAST are identical by definition.

3

RUPTURA: Breakthrough Curves Simulation Software

This chapter is based on the following publication: S. Sharma, S. R. Balestra, R. Baur, U. Agarwal, E. Zuidema, M. S. Rigutto, S. Calero, T. J. H. Vlugt, and D. Dubbeldam. "RUPTURA: simulation code for breakthrough, ideal adsorption solution theory computations, and fitting of isotherm models". *Mol. Simul.* 49 (2023), pp. 893–953.



3.1. Introduction

The separation of mixtures is of extreme importance to chemists and chemical engineers [142]. Adsorption-based methods are often implemented for separation of industrial processes such as the separation of hydrocarbons [95], CO_2 capture [98], water purification [96], refrigeration [97], etc. Adsorptive separation processes are divided into two broad categories: (a) continuous flow systems and (b) cyclic batch systems [62]. In a continuous flow system, the fluid phase passes through a fixed bed of solid adsorbents. The capacity of the adsorbents depends on the fluid phase velocity and residence time. To maintain the counter-current contact, the adsorbent must be circulated continuously or the circulation should be mimicked through clever semi-continuous design, e.g. by switching process streams [62]. Such requirements lead to a complex design of the separation system and reduced operational flexibility [62]. An example of such a system is the Simulated Moving Bed (SMB), which is widely used for chromatographic applications [143]. SMB mimics counter-current contact by cyclically shifting feed and product ports across a series of fixed adsorbent columns, avoiding the need for movement of solids while allowing efficient continuous separations [62, 65] Moving bed systems are designed to regenerate adsorbents by another bed by displacement, and this is not a common large-scale industrial adsorption setup. In a moving bed adsorption process, the solid adsorbent physically flows counter-current to the liquid or gas stream, enabling continuous contact and separation. While conceptually effective, true moving beds are rarely implemented due to mechanical challenges such as attrition and solid handling difficulties [62, 65]. In cyclic batch systems, the bed is alternately saturated and regenerated [62]. Based on the method of regenerating the adsorbent, the cyclic systems can be categorized into: (a) Pressure Swing Adsorption (PSA) and (b) Temperature Swing Adsorption (TSA) [144]. In TSA, the adsorber column is regenerated by heating the bed using a hot stream of non-adsorbing gas. This operation is performed at a temperature at which the adsorbed species desorb from the bed and are carried away along with the stream of hot gas. TSA has been widely used for gas drying and volatile organic compound recovery [144]. In PSA, desorption is achieved by lowering the pressure of the column at a constant temperature, which is followed by purging of a non-adsorbing stream of gas to remove the desorbed species [62]. Vacuum Pressure Swing Adsorption (VPSA) is a kind of PSA which is used for hydrogen purification, CO_2 capture, and air separation [144].

PSA is a non-cryogenic gas separation technology that achieves very high purity [145]. Pressure-swing adsorbers are operated in cyclic steady-state consisting of a minimum of two fixed beds with adsorbent, which continuously cycle through four dynamic process steps (Skarstrom cycle [146]): pressurization, adsorption, blow-down, and desorption. The use of more adsorbent columns improves the outlet gas purity and recovery rate because of the possibility of accommodating more pressure equalization

steps in each PSA cycle [147–149]. Pressure equalization is a process where gas leaving the first column being depressurized is used to partially pressurize the second adsorbent column [149]. Apart from improving the recovery and purity, this process also reduces the energy consumption. The frequency of regeneration is high in the case of PSA adsorption systems [62], and it does not have much impact on the adsorbent structures. These systems are designed based on short cycles [62] (ca. seconds to minutes). The frequency of regeneration is a crucial factor in designing TSA-based systems. This is because frequent thermal regeneration processes can adversely affect the adsorbent structure [62]. Therefore, TSA units are typically designed for longer cycles [62] (ca. hours to days) of operation.

Important factors that determine the economics of PSA units are [150]: (1) high selectivity for the adsorption of one component over the other components present in the gas mixture, (2) high working (adsorption) capacity between the conditions of regeneration and adsorption, (3) mild conditions for regeneration (usually induced by pressure or temperature swings), (4) high stability and resistance against impurities and moisture, and (5) fast adsorption kinetics. These factors most often exclude each other, and chemists and materials scientists attempt to find and rationalize the “sweet spot” for designing adsorbents [150]. In particular, in addition to selectivity, working capacity and recyclability (including the kinetics and energy of regeneration) are also key performance parameters [151]. The working capacity is the difference in loading of the preferentially adsorbed component at the adsorption pressure minus the loading at the purge pressure [152]. The adsorptive delivery should be maximized considering the entire adsorption-desorption cycle [153]. Based on this, an optimal enthalpy of adsorption change can be estimated. Experimental screening of potential adsorbent materials for use in PSA services is time-consuming. Therefore, computational screening of possible adsorbent materials for their performance factors is crucial.

Adsorption processes in fixed-bed columns are influenced by first-order factors (adsorption equilibrium isotherms) and second order factors (kinetics of intra/inter-particle mass/heat transfer, film/heat mass transfer, dispersion, nature of fluid flow, wall heat transfer) [62, 63, 155]. The performance of adsorbents in fixed-bed adsorbers can be evaluated by performing “breakthrough” simulations. A schematic diagram of a fixed-bed adsorber is shown in Fig. 3.1. A fixed bed packed with particles containing porous materials is pressurized and purged with a carrier gas. A fluid is added to the carrier gas, and the change of the component concentrations along the column and at the outlet of the fixed bed are recorded. Depending upon process objectives, a fixed-bed operation may be divided into three types: saturation adsorption, elution, and chromatography [155]. For saturation adsorption, a feed solution is passed through a column packed with adsorbents for the purpose of removal of the preferentially adsorbed component. This process continues until the adsorbents become sufficiently saturated, and the solute removal rate diminishes to the extent that termination of

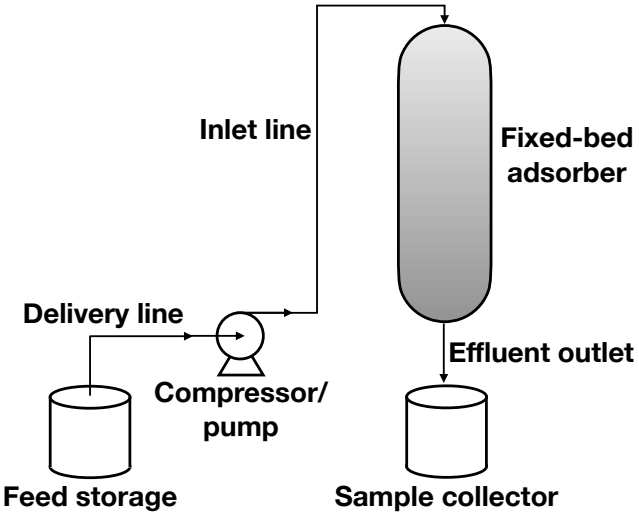


Figure 3.1. Schematic diagram of fixed-bed adsorber. A peristaltic compressor/pump is used to maintain a constant flow rate. In most gas phase adsorbers, the gas enters from the top and flows down through the bed, while in a typical liquid system, the column fills upwards [154]. The bed outlet is connected to a sample collector.

operation becomes necessary. Elution (desorption) is a process in which a solvent is passed through a column of adsorbents saturated with solutes. For chromatographic applications, a pulse of the mixture is introduced into a carrier fluid flowing through a column packed with adsorbents [155]. Pulse breakthrough is affected by both adsorption and desorption.

Despite the importance of breakthrough simulations, there is a lack of open-source software to predict breakthrough curves in fixed-bed adsorbers. In this chapter, we present such a software package named "RUPTURA". The code is available under the MIT license and downloadable from GitHub (https://www.github.com/iraspa/ruptura). Mixture prediction and breakthrough simulations rely on isotherm models that can be obtained by fitting isotherm data. RASPA [39, 156] is a software package for simulating adsorption of molecules in nanoporous materials. The combination with RASPA enables computation of breakthrough curves directly from the pure component adsorption simulations in the grand-canonical ensemble. RUPTURA and RASPA have similar input styles. RUPTURA contains three modules of workflow encountered in this field: (1) the computation of step and pulse breakthrough, (2) the prediction of mixture adsorption (used in the breakthrough equations) based on pure component isotherms, and (3) the fitting of isotherm models on raw (computed or measured) isotherm data. Many isotherm models have been published in the literature [64, 155, 157–160].

We included isotherm models like Langmuir, BET, Henry, Freundlich, Sips, Langmuir-Freundlich, Redlich-Peterson, Toth, Unilan, O'Brien & Myers and Asymptotic Temkin, as well as their multi-site versions or combinations of those. The implemented mixture prediction methods are: Ideal Adsorption Solution Theory (IAST) [51], segregated IAST [56], Explicit Isotherm (EI) model [118], and Segregated Explicit Isotherm (SIAST) model [93]. IAST is computed fast and with near machine precision. The breakthrough simulations include axial dispersion and the Linear Driving Force (LDF) model for mass-transfer [130], and use a numerically stable method called Strong-Stability Preserving Runge-Kutta (SSP-RK) for the numerical integration [69, 70, 139, 140, 161].

We foresee our code being used in (industrial) research for screening adsorbents for separation processes based on PSA and TSA, but also for teaching in chemistry and chemical engineering classes. Hence, in this chapter, we combine the presentation of our code with a review of the underlying theory and methodologies, and a tutorial. The teaching aspect is also the reason why the numerical schemes and fitting procedure are described in detail. The tutorial aspect also implied that we aim to make our code as easy to use as possible, and that the generation of breakthrough pictures and movies is automatic. It is also vital that researchers and students can play around with examples that run in the order of minutes. This is interactive enough to investigate the effect of adsorption, axial dispersion, mass-transfer coefficients, column void fraction, flow velocity, and column length on the breakthrough and separation efficiency. To achieve even higher computational speed, we included isotherm models of explicit nature (EI and SEI) that, although limited to Langmuir behavior, work for any number of components [118].

This chapter is organized as follows. Section 3.2 explains models for pure component isotherms and discusses the methodology to predict mixture results from pure component isotherms in Section 3.3. Such methodologies include IAST, and we detail our implementation and validate it by comparing it to previous work. Section 3.4 contains the theory on breakthrough simulations and our detailed numerical implementation. In Section 3.5, we focus on our genetic-algorithm implementation of isotherm fitting. We close our article with a description of the installation instructions (Section 3.6), the input format and options (Section 3.7), a tutorial (Section 3.8), and a troubleshooting section (Section 3.9). Our main findings are summarized in Section 3.10.

3.2. Pure Component Isotherms

3.2.1. Introduction

Adsorption is a surface process that involves the transfer of a molecule from a bulk fluid to a solid surface. Physical adsorption arises from Van der Waals forces, including dipole–dipole, dipole-induced dipole, London forces, and

possibly hydrogen bonding) [162, 163]. An adsorbate is a molecule adsorbed on the surface of the solid material, and the solid material is referred to as the adsorbent. An adsorption process is the addition of adsorbate to the adsorbent by increasing the adsorptive pressure, while desorption is the removal of adsorbate from the adsorbent by decreasing the adsorptive pressure and/or increasing the temperature [163, 164]. Experimental adsorption data are routinely reported as net or excess amounts adsorbed, while simulations, such as molecular simulations, measure absolute adsorption. The excess adsorbed amount refers to the difference between the actual (absolute) amount adsorbed and the amount that would be present in the same volume at the density of the fluid in the bulk phase [165, 166]. The net adsorbed amount has a reference state that does not require the knowledge of the adsorbent volume, and the solid and adsorbed phases are treated as an entity [167]. Statistical thermodynamic theories and molecular simulations of the adsorption of gases on porous solids are formulated in the language of absolute thermodynamic variables [168].

The fundamental concept in adsorption science is the adsorption isotherm, i.e. the equilibrium relation between the quantity of the adsorbed material and the pressure or concentration in the bulk fluid phase at constant temperature [164]. Mathematically, we can describe absolute adsorption of an adsorbate on an adsorbent with a smooth, continuous function $q(c_i)$ that represents the dependency of the adsorbed phase concentration q_i of a component i on the fluid phase concentrations c_i

$$q_i = q(c_1, c_2, \ldots, c_{N_C}) \qquad \text{at constant } T \tag{3.1}$$

Common units for the loading q include mol/(kg framework) and mg/(mg framework) with $q \geq 0$. For a single-component system, we generally have $\mathrm{d}q/\mathrm{d}p \geq 0$. The adsorption of a component depends not only on the concentration of this component, c_i, but also on the equilibrium concentrations of all other components. Likewise, adsorption from an ideal gas phase can be described by substituting the fluid phase concentrations with partial pressures, p_i, in the gas phase:

$$q_i = q(p_1, p_2, \ldots, p_{N_C}) \qquad \text{at constant } T \tag{3.2}$$

The adsorption equilibrium of a single adsorbate can be described by the adsorption isotherm:

$$q = q(p) \qquad \text{at constant } T \tag{3.3}$$

In general, an increase in the temperature will lead to a decreased amount adsorbed at a given pressure. In case the isotherm model has a well-defined saturation loading q^{sat}, a fractional loading θ can be defined:

$$\theta(p) = \frac{q(p)}{q^{\mathrm{sat}}} \tag{3.4}$$

The Henry coefficient k_H for adsorption is defined as

$$k_H = \lim_{p \to 0} \frac{q(p)}{p} = \lim_{p \to 0} \frac{dq}{dp} \tag{3.5}$$

The Henry coefficient is the slope of the isotherm at very low pressure. In this infinite dilution regime, there are no adsorbate-adsorbate interactions, and adsorption is linearly related to the affinity of the adsorbate. Note that the Henry coefficient depends on temperature.

3.2.2. Isotherm models

Isotherm models are well described in various resources [64, 155, 157–160]. New isotherm models continue to be developed. For example, the new Bingel-Walton isotherm model allows for a continuous, mathematical description of general type V isotherms, which appear in novel flexible MOFs or many water adsorption cases [169]. We will describe some of the isotherm models that are implemented in RUPTURA. The mathematical description for the functional form, the Henry coefficient, the saturation value, and the order of input of the arguments in the code are summarized in Table 3.1. The derived formulas for the inverse of the isotherm are listed in Table 3.2.

Langmuir model The Langmuir equation is the cornerstone of all theories of adsorption. Langmuir (1918) was the first to propose a coherent theory of adsorption onto a flat surface based on a kinetic viewpoint [170]. The assumptions of the Langmuir model are:

1. The surface is homogeneous: all adsorption sites are energetically identical.

2. The adsorption is localised: one molecule per adsorption site (monolayer).

3. There are no lateral interactions between adsorbed molecules.

These assumptions are often true for chemisorption. Using these assumptions, the Langmuir isotherm can be derived as:

$$q(p) = q^{sat} \frac{bp}{1 + bp} \qquad q^{sat} \geq 0, b > 0 \tag{3.6}$$

where q^{sat} is the saturation capacity and b is the coefficient of adsorption representing the affinity of the molecule. In the limit of high pressure, the isotherm will approach q^{sat}. At very low pressures, we obtain Henry's law, with a Henry coefficient of bq^{sat}.

Table 3.1. Isotherm models, Henry and saturation regime, and the order of input of the arguments in RUPTURA.

Model	Equation	Henry	saturation	b_0	b_1	b_2
Langmuir [170]	$q(p) = q^{sat}\frac{bp}{1+bp}$	bq^{sat}	q^{sat}	q^{sat}	b	
n-site Langmuir	$q(p) = \sum_i q_i^{sat}\frac{b_i p}{1+b_i p}$	$\sum_i b_i q^{sat}$	$\sum_i q_i^{sat}$	q^{sat}	b	
BET [171]	$q(p/p^0) = q^{sat}\frac{b(p/p^0)}{(1-c(p/p^0))(1-c+b(p/p^0))}$	$\frac{bq^{sat}}{1-c}$	$\frac{bq^{sat}}{(1-c)(1-c+b)}$	q^{sat}	b	c
Anti-Langmuir [157]	$q(p) = \frac{ap}{1-bp}$	a	✗	a	b	
Henry	$q(p) = ap$	a	✗	a		
Freundlich [172]	$q(p) = ap^{1/\nu}$	✗	✗	a	ν	
Sips [173]	$q(p) = q^{sat}\frac{(bp)^{1/\nu}}{1+(bp)^{1/\nu}}$	✗	q^{sat}	q^{sat}	b	ν
n-site Sips	$q(p) = \sum_i q_i^{sat}\frac{(b_i p)^{1/\nu_i}}{1+(b_i p)^{1/\nu_i}}$	✗	$\sum_i q_i^{sat}$	q^{sat}	b	ν
Langmuir-Freundlich [108]	$q(p) = q^{sat}\frac{bp^{\nu}}{1+bp^{\nu}}$	✗	q^{sat}	q^{sat}	b	ν
n-site Langmuir-Freundlich	$q(p) = \sum_i q_i^{sat}\frac{b_i p_i^{\nu}}{1+b_i p_i^{\nu}}$	✗	$\sum_i q_i^{sat}$	q^{sat}	b	ν
Redlich-Peterson [174]	$q(p) = \frac{ap}{1+bp^{\nu}}$	a	✗	a	b	ν
Toth [175–177]	$q(p) = q^{sat}\frac{bp}{[1+(bp)^{\nu}]^{1/\nu}}$	bq^{sat}	q^{sat}	q^{sat}	b	ν
Unilan	$q(p) = q^{sat}\frac{1}{2\eta}\ln\left[\frac{1+be^{\eta}p}{1+be^{-\eta}p}\right]$	bq^{sat}	q^{sat}	q^{sat}	b	η
O'Brien & Myers [178]	$q(p) = q^{sat}\left[\frac{bp}{1+bp} + \sigma^2\frac{bp(1-bp)}{2(1+bp)^3}\right]$	$q^{sat}b\left(1+\frac{\sigma^2}{2}\right)$	q^{sat}	q^{sat}	b	σ
Quadratic [179, 180]	$q(p) = q^{sat}\frac{bp+2cp^2}{1+bp+cp^2}$	bq^{sat}	$2q^{sat}$	q^{sat}	b	c
Asymptotic Temkin [181, 182]	$q(p) = q^{sat}\frac{bp}{1+bp} + q^{sat}\theta\left(\frac{bp}{1+bp}\right)^2\left(\frac{bp}{1+bp}-1\right)$	bq^{sat}	q^{sat}	q^{sat}	b	θ
Bingel & Walton [169]	$q(p) = q^{sat}\frac{1-\exp[-(a+b)p]}{1+(b/a)\exp[-(a+b)p]}$	aq^{sat}	q^{sat}	q^{sat}	a	b

The bi-Langmuir suggested by Graham [183] is the simplest model for adsorption onto a non-homogeneous surface [157]

$$q(p) = q_1^{sat}\frac{b_1 p}{1+b_1 p} + q_2^{sat}\frac{b_2 p}{1+b_2 p} \tag{3.7}$$

in which the subscripts refer to the two adsorption sites. In general, a heterogeneous surface with several distinct types of homogeneous adsorption sites can be modelled with an n-site model:

$$q(p) = \sum_{i=1}^{n} q_i^{sat}\frac{b_i p}{1+b_i p} \tag{3.8}$$

with a saturation value $\sum_i q_i^{sat}$ and Henry coefficient value $\sum_i b_i q_i^{sat}$. Multi-site models are required to model isotherms with inflections ("kinks"). Molecular simulations were able to explain the molecular origins of inflections in isotherms by examining the locations of molecules as a function of pressure [49, 184].

Anti-Langmuir model In the anti-Langmuir model, with increasing concentration or pressure, the adsorbed loading increases towards infinity [116]. At very low concentration or pressure, this model obeys Henry's law. The anti-Langmuir isotherm expression reads [116, 157]:

$$q(p) = \frac{ap}{1-bp} \qquad 0 \le p \le 1/b,\ b \ge 0 \tag{3.9}$$

Table 3.2. Isotherm models and their inverse. The inverse of the O'Brien & Myers and The Temkin model is too long to write down here, but can easily be computed symbolically using Mathematica. Note that multi-site combinations will always need to be computed numerically.

Model	inverse
Langmuir	$p(q)=\frac{q}{b\left(q^{\text{sat}}-q\right)}$
n-site Langmuir	numerical
BET	$p(q)=\frac{-bq^{\text{sat}}+bq-cq+c^2q\pm\sqrt{4bcq(q-cq)+\left(-bq^{\text{sat}}+bq-cq+c^2q\right)^2}}{2bcq}$
Anti-Langmuir	$p(q)=\frac{q}{(a+bq)}$
Henry	$p(q)=q/a$
Freundlich	$p(q)=\left(\frac{q}{a}\right)^{\nu}$
Sips	$p(q)=\frac{1}{b}\left(\frac{q}{q^{\text{sat}}-q}\right)^{\nu}$
n-site Sips	numerical
Langmuir-Freundlich	$p(q)=\left(\frac{q}{q^{\text{sat}}b-qb}\right)^{1/\nu}$
n-site Langmuir-Freundlich	numerical
Redlich-Peterson	numerical
Toth	$p(q)=\frac{q}{\left(q^{\text{sat}}b-qb\right)^{1/\nu}}$
Unilan	$p(q)=\frac{e^{\eta}\left(\exp\left(\frac{2\eta q}{q^{\text{sat}}}\right)-1\right)}{b\left(e^{2\eta}-\exp\left(\frac{2\eta q}{q^{\text{sat}}}\right)\right)}$
O'Brien & Myers	not shown here
Quadratic	$p(q)=\frac{-bq^{\text{sat}}+bq\pm\sqrt{\left(bq^{\text{sat}}-bq\right)^2+4q\left(2cq^{\text{sat}}-cq\right)}}{2\left(2cq^{\text{sat}}-cq\right)}$
Asymptotic Temkin	not shown here
Bingel & Walton [169]	numerical

In Eq. 3.9, a represents the Henry coefficient k_{H} in mol kg^{-1} Pa^{-1} and b is the equilibrium constant in Pa^{-1}. The anti-Langmuir model does not have a saturation loading, as the maximum loading can increase up to infinity at pressure or concentration equal to $1/b$.

Henry model All isotherms should, in principle, converge to Henry's law at infinite dilution. In Henry regime, the amount adsorbed is proportional to the pressure. The Henry's isotherm model is described by

$$q(p)=ap \tag{3.10}$$

where a is the energetic constant (Henry coefficient), and only depends on temperature.

BET model The Brunauer, Emmett, and Teller (BET) model extends adsorption to multi-layers [171]:

$$q(p) = q^{\text{sat}} \frac{b_s(p)}{(1 - b_l(p))(1 - b_l + b_s(p))} \tag{3.11}$$

where b_s and b_l are the equilibrium constants of adsorption on the bare surface and on a layer of previously adsorbed adsorbates, respectively. Similar to the Langmuir model, it is derived from kinetic adsorption-desorption relations. The assumptions made in this model are:

1. Each molecule in the first adsorbed layer provides an adsorption site for the second layer, and so on.
2. Molecules in the second and subsequent layers are assumed to behave essentially as those in the bulk liquid.

Freundlich model Boedeker proposed the following empirical isotherm equation [172]

$$q(p) = ap^{1/\nu} \tag{3.12}$$

for the adsorption of polar compounds on polar adsorbents, where the exponent $1/\nu$ is smaller than unity. It has been popularized by Freundlich and is therefore known as the Freundlich isotherm. The Freundlich isotherm can be considered a composite of Langmuir isotherms with different b values representing patches of adsorption sites with different adsorption energies [185]. It was shown that summing up a number of Langmuir isotherms leads to Freundlich-type isotherms [186]. The isotherm model can describe adsorption on many heterogeneous surfaces well. However, the model is unable to describe any plateauing trend and also does not have a Henry regime (in fact, the initial slope is infinite). As a result, some authors have mentioned that this isotherm type is unsuitable for the calculation of the reduced grand potential and other thermodynamic properties [187].

Sips model Sips proposed an equation similar in form to the Freundlich equation, but it has a finite limit when the pressure is sufficiently high [173]:

$$q(p) = q^{\text{sat}} \frac{(bp)^{1/\nu}}{1 + (bp)^{1/\nu}} \tag{3.13}$$

In this model, the additional parameter ν is a parameter characterising the heterogeneity of the system. The theoretical basis of the equation is described in the book of Do [64]. There also exists a multi-site form:

$$q(p) = \sum_i q_i^{\text{sat}} \frac{(b_i p)^{1/\nu_i}}{1 + (b_i p)^{1/\nu_i}} \tag{3.14}$$

The Sips isotherm does not have the correct limiting behavior at low pressure, i.e. it does not have a Henry regime (Eq. 3.5 diverges), except when $\nu = 1$.

Langmuir-Freundlich model The Langmuir-Freundlich model is given by [108]

$$q(p) = q^{\text{sat}} \frac{bp^{\nu}}{1 + bp^{\nu}} \tag{3.15}$$

and has the combined form of Langmuir and Freundlich equations. The constant ν is often interpreted as the heterogeneity factor. Values of unity indicate a material with homogeneous binding sites and the isotherm model reduces to the Langmuir model. There also exists a multi-site form

$$q(p) = \sum_i q_i^{\text{sat}} \frac{b_i p^{\nu_i}}{1 + b_i p^{\nu_i}} \tag{3.16}$$

The Langmuir-Freundlich isotherm does not have the correct limiting behavior at low pressure, i.e. it does not have a Henry regime, except when $\nu = 1$.

Redlich-Peterson model The Redlich-Peterson isotherm is an empirical isotherm combining the Langmuir and Freundlich isotherms. The numerator is the same as the Langmuir isotherm and has the advantage of possessing a Henry regime at infinite dilution [174]

$$q(p) = \frac{ap}{1 + bp^{\nu}} \tag{3.17}$$

The parameter a cannot be interpreted as the saturation loading, except when $\nu = 1$ [185].

Toth model The previous isotherm models have their limitations. The Freundlich, Sips, and Langmuir-Freundlich equations are not valid at very low pressures, while the Henry, Freundlich, and Redlich-Peterson models do not have a finite saturation value. One of the empirical equations that is popularly used and satisfies the two end limits is the Toth equation [175–177]:

$$q(p) = q^{\text{sat}} \frac{bp}{[1 + (bp)^{\nu}]^{1/\nu}} \tag{3.18}$$

The Toth equation has been used for fitting data of many adsorbates such as hydrocarbons, carbon oxides, hydrogen sulfide, alcohols on activated carbon, and zeolites [64].

Unilan model The Unilan model derived its name from UNI (Uniform distribution) and LAN (Langmuir local model) [64]. The Unilan model is described by

$$q(p) = q^{\text{sat}} \frac{1}{2\eta} \ln\left[\frac{1 + be^{\eta} p}{1 + be^{-\eta} p}\right] \tag{3.19}$$

η is a measurement of the heterogeneity of the adsorbent. High values indicate a highly heterogeneous system. Being a three-parameter model, the Unilan equation is also very often used to describe many data of hydrocarbons and carbon oxides on activated carbon and zeolites. The Unilan equation has the correct behavior at low and high pressures. In the limit of $\eta = 0$, the Langmuir isotherm is recovered.

O'Brien & Myers The O'Brien and Myers isotherm model is obtained as a truncation to two terms of a series expansion of the adsorption integral equation in terms of the central moments of the adsorption energy distribution [64, 178]

$$q(p) = q^{\text{sat}} \left[\frac{bp}{1+bp} + \sigma^2 \frac{bp\,(1-bp)}{2\,(1+bp)^3} \right] \tag{3.20}$$

where σ is a measure of the width of the adsorption energy distribution.

Quadratic model Statistical thermodynamics suggests that the general form of an isotherm equation should be the ratio of two polynomials of the same degree [180]. The polynomial Langmuir isotherm model, derived from statistical mechanics, reads:

$$q(p) = q^{\text{sat}} \frac{bp + cp^2 + dp^3 + \ldots}{1 + bp + cp^2 + dp^3 + \ldots} \tag{3.21}$$

In practice, the second order isotherm is often used, called the quadratic isotherm model [179, 180]

$$q(p) = q^{\text{sat}} \frac{bp + 2cp^2}{1 + bp + cp^2} \tag{3.22}$$

The loading is convex at low pressures but changes concavity as it saturates, yielding an S-shape, i.e. the isotherm exhibits an inflection point.

Asymptotic approximation to the Temkin model The Temkin isotherm is derived using mean-field arguments and asymptotic approximation [181, 182]

$$q(p) = q^{\text{sat}} \frac{bp}{1+bp} + q^{\text{sat}} \theta \left(\frac{bp}{1+bp} \right)^2 \left(\frac{bp}{1+bp} - 1 \right) \tag{3.23}$$

Here, q^{sat} and b have the same meaning as in the Langmuir isotherm, and θ describes the strength of adsorbate-adsorbate interactions ($\theta < 0$ for attraction).

Bingel and Walton model The Bingel and Walton isotherm model allows for a continuous, mathematical description of general type V isotherms, which appear in novel flexible MOFs or many water adsorption cases [169]. The model is based on the Bass model of innovation diffusion developed in the late 1960s [188]. This model describes the adoption and diffusion of an invention over time and has been widely used in market sales and technology forecasting. Bingel and Walton applied the Bass model to adsorption, where early adopters can be seen as the intrinsic high-affinity adsorption sites of the surface. The word-of-mouth contribution of slower innovations refers to adsorption mechanisms that are mainly driven by molecules of the same species that are already adsorbed and thus present in the adsorbed phase. The Bingel-Walton adsorption equation reads [169]

$$q(p) = q^{\text{sat}} \frac{1 - \exp\left[-(a+b)\,p\right]}{1 + (b/a)\exp\left[-(a+b)\,p\right]} \tag{3.24}$$

where $a > 0$ is the intrinsic adsorption affinity between the adsorbate and the adsorbent, and b is the clustering coefficient describing strong adsorbate–adsorbate interactions. The simple model combines the effects of early adopters and word of mouth, resulting in curves that resemble either the type I adsorption isotherm shape or S-shaped curves. The location of the inflection point is $\log(b/a)/(a+b)$. For small values of a and large values of b, the isotherm shape becomes strongly step-wise. In the limit of $b \to 0$, the isotherm becomes an exponential function:

$$q(p) = q^{\text{sat}}\left(1 - \exp\left[-ap\right]\right) \tag{3.25}$$

while in the limit of $a \to 0$ the adsorption approaches zero. In the limit of $b \to -a$, the isotherm reduces to the Langmuir model

$$q(p) = q^{\text{sat}} \frac{ap}{1 + ap} \tag{3.26}$$

The model has a Henry coefficient of aq^{sat} and a saturation loading of q^{sat}.

3.2.3. Reduced grand potential

Thermodynamics is the study of energy and its transformations, but more generally tries to establish relationships between basic concepts like internal energy U, entropy S, number of particles N, volume V, temperature T, pressure p, and chemical potential μ to describe the behavior of matter and make predictions. For the mixture vapor-liquid equilibrium, the Gibbs-Duhem equation [189]

$$S\mathrm{d}T - V\mathrm{d}p + \sum_i n_i \mathrm{d}\mu_i = 0 \tag{3.27}$$

shows that the independent variables which define the standard states for the components of the mixture are T and p. We can compare the Gibbs-Duhem equation to the differential for a solid material [190]

$$S\mathrm{d}T + \mathrm{d}\Omega + \sum_i n_i \mathrm{d}\mu_i = 0 \tag{3.28}$$

and see that the standard states for mixture adsorption will be determined by T and Ω [190]. The grand potential Ω is the characteristic state function for the grand-canonical ensemble, and the unit of Ω is J/(kg framework). At constant temperature, we have

$$\mathrm{d}\Omega = -S\mathrm{d}T - \sum_i n_i \mathrm{d}\mu_i = -\sum_i n_i \mathrm{d}\mu_i \tag{3.29}$$

which is the "Gibbs adsorption" equation. Replacing chemical potential μ_i by the fugacity f_i

$$\mathrm{d}\Omega = -RT \sum_i n_i \mathrm{d}\ln\left[\frac{f_i}{f_0}\right] \tag{3.30}$$

where f_0 is a reference fugacity to make the argument of the logarithm dimensionless. Replacing fugacity by pressure and integrating for pure-component adsorption from the unadsorbed state at zero pressure, we obtain the grand potential Ω [62, 168, 191, 192]

$$\Omega_i = -RT \int_0^{p_i^*} q_i^*(p)\mathrm{d}\ln(p) \quad \text{or} \quad \Omega_i = -RT \int_0^{p_i^*} \frac{q_i^*(p)}{p}\mathrm{d}p \tag{3.31}$$

where $q_i^*(p)$ is the loading of pure component i given as a function of the pressure. Physically, the grand potential is the free energy change associated with the isothermal immersion of fresh adsorbent in the bulk fluid. The absolute value of the grand potential is the minimum isothermal work necessary to clean the adsorbent [193]. In calculations, it is convenient to introduce a reduced grand potential ψ [190]:

$$\psi\left(p_i^*\right) \equiv -\frac{\Omega_i}{RT} = \int_0^{p^*} \frac{q_i^*(p)}{p}\mathrm{d}p \tag{3.32}$$

The reduced grand potential has units of mol/kg. The integration limit p_i^* is a property of the pure component property, i.e. it is the pressure at a given reduced grand potential. This quantity is known as the sorption pressure (in analogy to the saturation pressure in vapor-liquid equilibrium) and also as the hypothetical pressure. Importantly, the reduced grand potential is defined in terms of absolute adsorption and can be computed from the adsorption isotherm. In Tables 3.3 and 3.4, we list the derived expressions for the reduced grand potential and the sorption pressure (the inverse of the reduced grand potential), respectively, for the various isotherm models.

Table 3.3. Reduced grand potential for isotherm models. The potentials are additive for multi-site models. Note that our expression for the Unilan is the same as Santori et al. [194], but several terms in their expression actually cancel out. Note that the expression for Asymptotic Temkin of Simon et al. [54] contains typos. The dilogarithm function (Li_2) defined by the power series $\mathrm{Li}_2(z) = \sum_{n=1}^{\infty} \frac{z^n}{n^2} \quad |z| < 1$. In RUPTURA, we use the freely available implementation by Alexander Voigt [195]. The hypergeometric function ${}_2F_1$ defined by the power series for $|z| < 1$ [196] ${}_2F_1(a, b; c; z) = \sum_{k=0}^{\infty} \frac{(a)_k (b)_k}{(c)_k} \frac{z^k}{k!}$ where $(q)_k$ is the (rising) Pochhammer symbol. The hypergeometric function can be computed by e.g. Gosper's algorithm [197]. Note that the computation of the reduced grand potential for Redlich-Peterson and Toth isotherms is computationally much more expensive than for the other isotherm models.

Model	reduced grand potential $\psi(p)$	
Langmuir [64]	$\psi(p) = q^{\mathrm{sat}} \ln[1 + bp]$	
n-site Langmuir [64]	$\psi(p) = \sum_i q_i^{\mathrm{sat}} \ln[1 + b_i p]$	
BET [116]	$\psi(p) = q^{\mathrm{sat}} \ln\left[\frac{1+bp-cp}{1-cp}\right]$	
Anti-Langmuir [157]	$\psi(p) = -\frac{a}{b} \ln[1 - bp]$	
Henry	$\psi(p) = ap$	
Freundlich [64]	$\psi(p) = a\nu p^{1/\nu}$	
Sips [64]	$\psi(p) = q^{\mathrm{sat}} \nu \ln\left[1 + (bp)^{1/\nu}\right]$	
n-site Sips	$\psi(p) = q_i^{\mathrm{sat}} \nu \ln\left[1 + (b_i p)^{1/\nu_i}\right]$	
Langmuir-Freundlich [194]	$\psi(p) = \frac{q^{\mathrm{sat}}}{\nu} \ln[1 + bp^{\nu}]$	
n-site Langmuir-Freundlich	$\psi(p) = \sum_i \frac{q_i^{\mathrm{sat}}}{\nu_i} \ln[1 + b_i p^{\nu_i}]$	
Redlich-Peterson [198, 199]	$\psi(p) = ap\,{}_2F_1(1, 1/\nu, 1 + 1/\nu, -bp^{\nu})$	$(bp^{\nu} < 1)$
	$= \frac{a}{\nu b^{1/\nu}} \left[\frac{\pi}{\sin(\pi/\nu)} + \sum_{k=1}^{\infty} (-1)^k \frac{\left(\frac{1}{bp^{\nu}}\right)^{k-1/\nu}}{k - 1/\nu} \right]$	$(bp^{\nu} > 1)$
Toth [194, 200]	$\psi(p) = q^{\mathrm{sat}} bq\ {}_2F_1(1/\nu, 1/\nu, 1 + 1/\nu, -(pb)^{\nu})$	$(bp^{\nu} < 1)$
	$= q^{\mathrm{sat}} \left(\theta - \frac{\theta}{\nu} \ln[1 - \theta^{\nu}] - \sum_{k=1}^{\infty} \frac{\theta^{k\nu+1}}{k\nu[k\nu+1]} \right)$	
Unilan [194]	$\psi(p) = q^{\mathrm{sat}} \frac{1}{2\eta} \left(\mathrm{Li}_2(-be^{-\eta}p) - \mathrm{Li}_2(-be^{\eta}p)\right)$	
O'Brien & Myers [64, 198]	$\psi(p) = q^{\mathrm{sat}} \left[\ln[1 + bp] + \frac{\sigma^2 bp}{2(1+bp)^2} \right]$	
Quadratic [116]	$\psi(p) = q^{\mathrm{sat}} \ln\left[1 + bp + cp^2\right]$	
Asymptotic Temkin [54]	$\psi(p) = q^{\mathrm{sat}} \left(\ln[1 + bp] - \frac{1}{2}\theta\left(\frac{bp}{1+bp}\right)^2 \right)$	
Bingel & Walton [169]	numerical	

3.3. Prediction of Mixture Isotherms

3.3.1. Introduction

A key point in adsorption process development is knowledge on multi-component adsorption equilibria [63]. Experimental methods for (mixture) gas adsorption have recently been reviewed by Shade et al. [201]. Direct

Table 3.4. Sorption pressures for isotherm models. Models with an analytical inverse for the reduced grand potential are an order of magnitude faster than models that have to be numerically inverted.

Model	sorption pressure $p^*(\psi)$
Langmuir	$p^*(\psi) = \frac{1}{b}\left(\exp\left[\psi/q^{\text{sat}}\right]-1\right)$
n-site Langmuir	numerical
BET	numerical
Anti-Langmuir	$p^*(\psi) = \frac{1}{b}\left(1-\exp\left[-\psi b/a\right]\right)$
Henry	$p^*(\psi) = \frac{\psi}{a}$
Freundlich	$p^*(\psi) = \left(\frac{\psi}{a\nu}\right)^{\nu}$
Sips	$p^*(\psi) = \frac{1}{b}\left(\exp\left[\frac{\psi}{\nu q^{\text{sat}}}\right]-1\right)^{\nu}$
n-site Sips	numerical
Langmuir-Freundlich	$p^*(\psi) = \left(\frac{\exp\left[\nu\psi/q^{\text{sat}}\right]-1}{b}\right)^{1/\nu}$
n-site Langmuir-Freundlich	numerical
Redlich-Peterson	numerical
Toth	numerical
Unilan	numerical
O'Brien & Myers	numerical
Quadratic	$p^*(\psi) = \frac{-b\pm\sqrt{b^2-4c+3c\exp\left[\psi/q^{\text{sat}}\right]}}{2c}$
Asymptotic Temkin	numerical
Bingel & Walton	numerical

measurement of mixture adsorption equilibria remains complicated and time-consuming, and mixture adsorption prediction using theoretical models is still the default tool [52]. These theoretical models can be validated by explicit grand-canonical Monte Carlo simulations [50] for mixture adsorption.

One of the commonly used models is the extended Langmuir, which was developed by Butler and Ockrent [202] to describe competitive adsorption. For an N_c component mixture, the adsorption of component i reads [202]

$$q_i = \frac{q_i^{\text{sat}} b_i p_i}{1+\sum_{j=1}^{N_C} b_j p_j} \tag{3.33}$$

The extended Langmuir is only dynamically consistent when all components have the same saturation value [203]. If not, the extended Langmuir is only empirical in nature. Jain and Snoeyink [204] proposed an extension of the Langmuir equation for binary mixtures that is based on the assumption that only a fraction of the adsorption sites that are available for a component

can also be occupied by the other component.

$$q_1 = \frac{\left(q_1^{\text{sat}} - q_2^{\text{sat}}\right) b_1 p_1}{1 + b_1 p_1} + \frac{q_2^{\text{sat}} b_1 p_1}{1 + b_1 p_1 + b_2 p_2} \tag{3.34}$$

$$q_2 = \frac{q_2^{\text{sat}} b_2 p_2}{1 + b_1 p_1 + b_2 p_2} \tag{3.35}$$

Numerous extensions of the Langmuir, Freundlich, Toth, and Redlich-Peterson equations have been developed to model mixture adsorption isotherms [185]. A thermodynamic framework for computing the mixture adsorption is the Ideal Adsorption Solution Theory (IAST) developed by Myers and Prausnitz [51]. For systems following the Langmuir isotherm, IAST yields results identical to the extended Langmuir equation for mixtures, if the saturated amounts are equal [128]. IAST is a predictive model which uses only the pure component data (it does not require any mixture data), is thermodynamically consistent, and is independent of the actual model of physical adsorption [205]. After more than 50 years of its development, IAST remains a benchmark approach for describing mixture adsorption [52]. The applicability of IAST continues to be evaluated on novel materials [103, 206, 207] and for special circumstances, for example, adsorption in the presence of framework deformations [208]. Gharagheizi and Sholl recently evaluated IAST on more than 400 examples in which binary adsorption data and single-component data are available in the same publication [209]. Other implicit multi-component adsorption models are [140]: Vacancy Solution Theory (VST) [210], Real Adsorption Solution Theory (RAST) [211, 212], Spreading Pressure Dependent equation (SPD) [53], Predictive Real Adsorption Solution Theory (PRAST) [213], Multi-component Potential Adsorption Theory (MPAT) [214], Segregated Ideal Adsorbed Solution Theory (SIAST) [56], and Generalized Predictive Adsorbed Solution Theory (GPAST) [215].

3.3.2. Ideal Adsorption Solution Theory (IAST)

Myers and Monson applied solution thermodynamics to adsorption in porous materials, leading to equations similar to those for vapor-liquid equilibria [190]. In IAST calculations, the central quantity is the reduced grand potential ψ_i:

$$\psi_i\left(f_i^*\right) = \int_0^{f^*} \frac{q_i^*(f)}{f} \mathrm{d}f \tag{3.36}$$

The reduced grand potential has units of mol/kg. This quantity is related to the spreading pressure (Π) (or solid-fluid interfacial tension), which is analogous to pressure but in two dimensions [54]. The relation between reduced grand potential and spreading pressure is as follows [54]:

$$\psi\left(f_i^*\right) = \frac{\Pi A}{RT} \tag{3.37}$$

In Eq. 3.37, A is the area of the adsorbent in m^2. The spreading pressure was used in early IAST work based on quasi two-dimensional adsorption at a planar surface using the Gibbs excess formalism [51], but is replaced with the reduced grand potential in the more recent IAST work based on thermodynamics of adsorption in a three-dimensional pore network [190]. Equations written in the language of solution thermodynamics have been derived without any discussion of a dividing surface, Gibbs excess, or spreading pressure and lead to equations similar to those for vapor-liquid equilibria [190].

Following the excellent, detailed descriptions of IAST by Murthi and Snurr [191] and Myers and Monson [190], for a fluid at constant temperature, we have [216]

$$d\mu_i = RT\, d\ln\left[\frac{f_i}{f^0}\right] \tag{3.38}$$

In Eq. 3.38, f^0 is the reference fugacity, which is considered to be equal to 1 bar. Integrating this equation at constant T and Ω from a state of pure i to a state at an arbitrary mole fraction [191]

$$\mu_i(T,\Omega,x) - \mu_i^*(T,\Omega,x) = RT\ln\frac{f_i(T,\Omega,x)}{f_i^*(T,\Omega,x)} \tag{3.39}$$

where the superscript * indicates a property of the pure component evaluated at the mixture T and Ω, and f_i is the fugacity of component i in the adsorbed phase. The proposed definition of an ideal solution for adsorption in porous materials is [190]:

$$\mu_i^{\mathrm{id}}(T,\Omega,x) - \mu_i^{\mathrm{id},*}(T,\Omega,x) \equiv RT\ln x_i \tag{3.40}$$

Eq. 3.40 is equivalent to the equation of the chemical potential of an ideal solution in the bulk. It is the only assumption needed in the Ideal Adsorption Solution Theory (IAST) of Myers and Prausnitz [51]. Using $\mu = \mu^{\mathrm{id}} + \mu^{\mathrm{ex}}$ and subtracting Eq. 3.40 from Eq. 3.39 yields

$$\mu_i^{\mathrm{ex}} = RT\ln\frac{f_i(T,\Omega,x)}{x_i f_i^*(T,\Omega)} \tag{3.41}$$

$$= RT\ln\gamma_i \tag{3.42}$$

Since the argument of the logarithm is defined as the activity coefficient (γ_i) of component i and we have [191]

$$f_i(T,\Omega,x) = \gamma_i x_i f_i^*(T,\Omega) \tag{3.43}$$

The phase equilibrium (iso-fugacity) condition expresses that the fugacity of a component in the gas phase $\left(f_i^{(g)}\right)$ is in equilibrium with the mixture at the specified T and Ω

$$f_i^{(g)}(T,P,y) = \gamma_i x_i f_i^*(T,\Omega) \tag{3.44}$$

This equation may be rewritten in terms of the pressure by replacing the fugacity (f_i) with $y_i\phi_i P$

$$y_i\phi_i P = \gamma_i x_i f_i^* \tag{3.45}$$

where ϕ is the fugacity coefficient in the fluid phase. At equilibrium, the reduced grand potentials of the individual species are the same. The set of equations to be solved in terms of fugacities, assuming an ideal adsorbed solution and γ_i set to unity, is

$$y_i\phi_i p_T = x_i f_i^*(\psi) \tag{3.46}$$

where $f_i^*(\psi)$ is the fugacity at which each pure component is at the same reduced grand potential, ψ, and temperature of the mixture. In Real Adsorbed Solution Theory (RAST), the non-ideal behavior of the adsorbed phase is accounted for by the use of activity coefficients (γ_i) [211].

When gas-phase pressures are sufficiently low, the fugacities in the previous equations may be replaced by pressures. For liquid systems, the same set of equations applies with pressure replaced by concentration [217]. For simplicity, here we will always refer to pressure, with the understanding that all the results will apply to the corresponding liquid system. We note also that using pressure implies the assumption of an ideal gas (for gases) or ideal liquid mixture (for liquids) and that fugacity should replace pressure in a rigorous extension to high-pressure gas systems or non-ideal liquid mixtures.

The $2N_c + 1$ basic equations for IAST are [51, 64]:

$$y_i p_T = p_i = x_i p_i^*(\psi) \qquad i = 1, 2, \ldots, N_c \qquad N_c \text{ equations} \tag{3.47}$$

$$\sum_{i=1}^{N_c} x_i = 1 \qquad 1 \text{ equation} \tag{3.48}$$

$$\psi = \psi_1 = \psi_2 = \cdots = \psi_{N_c} \qquad N_c \text{ equations} \tag{3.49}$$

which can be solved for the $2N_c + 1$ unknowns which includes: (1) N_c values of mole fractions in the adsorbed phase x_i, (2) 1 value of the reduced grand potential ψ, and (3) N_c values of the sorption pressure of the pure component p_i^* that give the same reduced grand potential as that of the mixture.

Eq. 3.47 is analogous to Raoult's law that states that the partial pressure of each component of an ideal mixture of liquids is equal to the vapor pressure of the pure component multiplied by its mole fraction in the mixture. Note that Eqs. 3.47, 3.48, and 3.49 do not contain any information on the vacancy and the amount that has been adsorbed. The specific adsorption area of a given species is inversely proportional to q_i. The total amount adsorbed q_T can be calculated from the Gibbs adsorption isotherm Eq. 3.29 assuming zero mass or volume change upon adsorption, and is given by [190]

$$\frac{1}{q_T} = \sum_{i=1}^{N_c} \frac{x_i}{q_i^*} \tag{3.50}$$

where q_i^* is the adsorbed amount of pure component i at the sorption pressure p_i^*

$$q_i^* = q\left(p_i^*\right) \tag{3.51}$$

Knowing the total amount adsorbed, the amount contributed by component i is given by

$$q_i = x_i q_T \tag{3.52}$$

Eqs. 3.47-3.52 form a set of powerful equations for the IAS theory. If the total pressure and the mole fractions in the gas phase are given, then the unknowns can be computed:

- N_c mole fractions in the adsorbed phase (x_i)
- N_c sorption pressures $\left(p_i^*\right)$
- the total amount adsorbed and the component amount adsorbed

This is the most common use case and occurs, for example, in computing mixture isotherms from pure component isotherms and the use of IAST in fixed-bed adsorbers. However, the inverse problem can be posed as well. In that case, the adsorbed mole fractions x_i and the total adsorbed amount q_T are given, and the following unknowns have to be calculated:

- N_c mole fractions in the gas phase (y_i)
- N_c sorption pressures $\left(p_i^*\right)$
- the total pressure $\left(p_T\right)$

or the adsorbed mole fractions x_i and the total pressure are given, and the following unknowns have to be calculated:

- N_c mole fractions in the gas phase (y_i)
- N_c sorption pressures $\left(p_i^*\right)$
- the total amount adsorbed $\left(q_T\right)$

Note that the correct fundamental thermodynamic variable is the absolute adsorbed amount, and there is only one possible definition of the ideal adsorbed solution [192]. It is also noteworthy that IAST relies on the absolute amount adsorbed, and only one rigorous definition of IAST exists. The form of the adsorption isotherm equation for pure components is arbitrary and can take any form which fits the data best [218]. However, we note that especially the low-pressure data need to be accurately represented, and errors at low pressures lead to large errors in multi-component calculations.

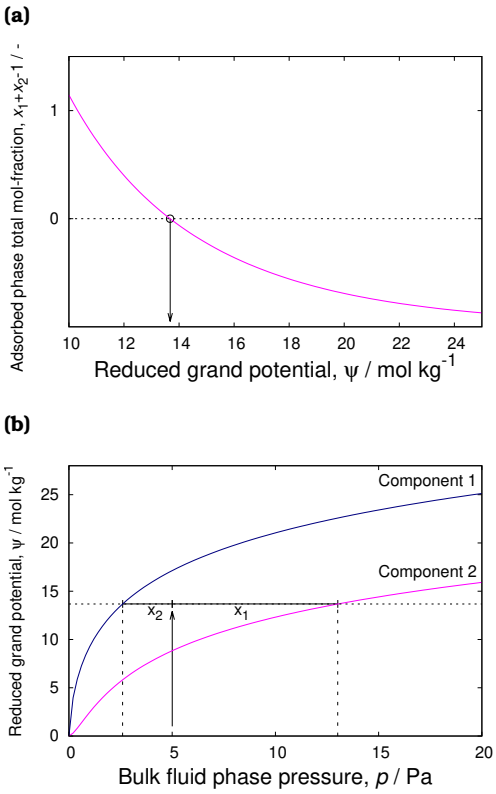


Figure 3.2. IAST algorithm: (a) adsorbed phase mole fraction as a function of the reduced grand potential and (b) graphical representation of the basic IAST relationship. The sum of the adsorbed phase mole fractions is unity at reduced grand potential $\psi = 13.6755$ mol/(kg framework). By using $(\sum_i x_i) - 1$, a root-finding algorithm can be used. Note that $\sum_i x_i$ has a monotonic relation to the reduced grand potential, hence it is also amendable to bi-section methods.

3.3.3. IAST numerical example

To illustrate the IAST algorithm, we consider two pure component Langmuir-Freundlich isotherms:

$$\text{Component 1:} \quad q_{s,1} \frac{b_1 p^{\nu_1}}{1.0 + b_1 p^{\nu_1}} \tag{3.53}$$

$$\text{Component 2:} \quad q_{s,2} \frac{b_2 p^{\nu_2}}{1.0 + b_2 p^{\nu_2}} \tag{3.54}$$

and we list code snippets that can be directly copied and pasted in Mathematica. We first define the parameters of the isotherms:

```
qs1 = 6.3;     qs2 = 5.3;
b1 = 2.0;      b2 = 1.0;
nu1 = 0.75;    nu2 = 1.5;
```

where we assume that $q_{s,i}$ is in units of mol/(kg framework), b_i in units of 1/Pa, pressure in units of Pa, and ν_i dimensionless (Note that $b_i p^{\nu_i}$ is dimensionless). Specifying the fluid-phase mole fractions and total pressure

```
p = 5;
y1 = 0.4;      y2 = 0.6;
```


We first have to find the reduced grand potential ψ that is consistent with the adsorbed phase mole fractions adding up to unity. Filling in (see Table 3.4)

$$p_i^*(\psi) = \frac{1}{b_i^{1/\nu_i}}\left[\exp\left(\frac{\nu_i \psi}{q_i^{\text{sat}}}\right) - 1\right]^{1/\nu_i} \tag{3.55}$$

into Eq. 3.120, and using Mathematica we can numerically solve for `psi`

```
FindRoot[
y1*p/((1.0/(b1^(1.0/nu1)))*
(Exp[nu1 psi/qs1]-1.0)^(1.0/nu1))+
         y2*p/((1.0/(b2^(1.0/nu2)))*
         (Exp[nu2 psi/qs2]-1.0)^(1.0/nu2))-1.0,
         {psi, 10}]
```

The root is psi=13.6755 mol/(kg framework). This process is illustrated in Fig. 3.2(a). Note that the functional relation between the sum of the adsorbed mole fractions and the reduced grand potential is monotonic (which can be exploited by bi-section algorithms). We can obtain the sorption pressures p1 and p2 (that correspond to the reduced grand potential ψ) using

```
p1 = (1/(b1^(1.0/nu1)))*
(Exp[nu1*13.6755/qs1]-1.0)^(1.0/nu1)
p2 = (1/(b2^(1.0/nu2)))*
(Exp[nu2*13.6755/qs2]-1.0)^(1.0/nu2)
```

and we obtain p1=2.59899 and p2=13.0167 Pa. The adsorbed-phase mole fractions x1 and x2 are

```
x1=y1*p/p1
x2=y2*p/p2
```

and we obtain x1=0.769531 and x2=0.230472, graphically depicted in Fig. 3.2(b). With psi and x1 and x2, the total adsorbed amount, qT, can be calculated:

```
qT = 1.0/(0.769531/(qs1*b1*p1^nu1/(1.0+b1*p1^nu1))+
          (0.230472/(qs2*b2*p2^nu2/(1.0+b2*p2^nu2))))
```

leading to qT=5.09176 mol/(kg framework).

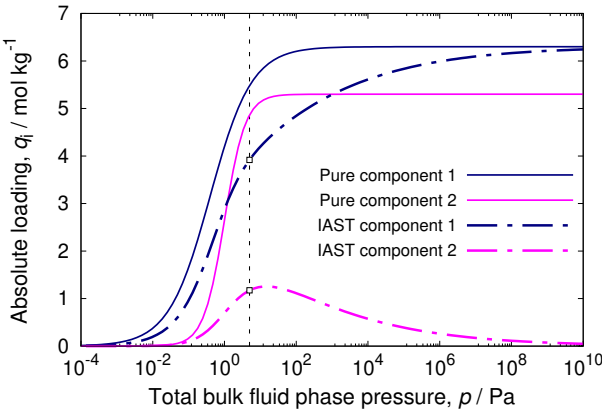


Figure 3.3. Example of IAST prediction for a binary mixture described by Langmuir-Freundlich isotherms. At 5 Pa total pressure and gas-phase mole fractions $y_1 = 0.4$ and $y_2 = 0.6$, we find adsorbed absolute loadings of $q_1 = 3.92$ and $q_2 = 1.17$ mol/kg. Typically, in a mixture, a component with the highest saturation loading will drive the other components out at high pressures.

```
q1 = x1*qT
q2 = x2*qT
```

and we have q1=3.91827 and q2=1.17351 mol/(kg framework). Using the following Mathematica code, we can generate the IAST prediction as a function of pressure

```
pressurebegin = 10^(-4);
pressureend = 10^10;
numberpoints = 100;
For[i = 1, i <= numberpoints, i++,
 p = 10^(((Log10[pressureend] -
           Log10[pressurebegin])*
         (i/(numberpoints - 1.0))) +
          Log10[pressurebegin]);
 root = psi /.
   FindRoot[
    y1*p/((1.0/(b1^(1.0/nu1)))*
    (Exp[nu1 psi/(qs1)] - 1.0)^(1.0/nu1)) +
     y2*p/((1.0/(b2^(1.0/nu2)))*
     (Exp[nu2 psi/(qs2)] - 1.0)^(1.0/
     nu2)) - 1.0, {psi, 10}];
 p1 = (1.0/(b1^(1.0/nu1)))*
      (Exp[nu1*root/qs1] - 1.0)^(1/nu1);
 p2 = (1.0/(b2^(1.0/nu2)))*
      (Exp[nu2*root/qs2] - 1.0)^(1/nu2);
 x1 = y1*P/((1.0/(b1^(1.0/nu1)))*
```

```
      (Exp[nu1*root/qs1] - 1.0)^(1.0/nu1));
x2 = y2*P/((1.0/(b2^(1.0/nu2)))*
      (Exp[nu2*root/qs2] - 1.0)^(1.0/nu2));
qT = 1.0/(x1/(qs1*b1*p1^nu1/(1 + b1*p1^nu1)) +
          x2/(qs2*b2*p2^nu2/(1 + b2*p2^nu2)));
Print[p, " ", x1*qT, " ", x2*qT]]
```

The result is plotted in Fig. 3.3, which shows pure component and mixture isotherms for a binary mixture. Adsorption for the component with the highest saturation loading dominates at high pressures.

3.3.4. Analytic mixture prediction methods

Multi-component Langmuir

In the book of Do [64], the multi-component Langmuir is derived from the IAST equations. Eqs. 3.47 and 3.48 can be combined to yield [219]

$$\sum_{i=1}^{N_C} \frac{p_i}{p_i^*} = \sum_{i=1}^{N_C} x_i = 1 \tag{3.56}$$

This closure equation can be used to reduce the problem to a single non-linear algebraic equation in reduced grand potential ψ:

$$\frac{p_1}{p_1^*(\psi)} + \cdots + \frac{p_{N_C}}{p_{N_C}^*(\psi)} = 1 \tag{3.57}$$

We can then use the expressions for the pure component sorption pressure $\left(p_i^*\right)$ and use these as input in Eq. 3.57. For the single component Langmuir isotherm

$$q_i = q^{\text{sat}} \frac{b_i p_i}{1 + b_i p_i} \tag{3.58}$$

we obtain

$$\psi\left(p^*\right) = \int_0^p \frac{q(p)}{p} \mathrm{d}p = q^{\text{sat}} \int_0^{p^*} \frac{b}{1+bp} \mathrm{d}p = q^{\text{sat}} \ln\left[1 + bp^*\right] \tag{3.59}$$

Inverting this equation, we have

$$p^*(\psi) = \frac{\exp\left[\frac{\psi}{q^{\text{sat}}}\right] - 1}{b} \tag{3.60}$$

Filling this into Eq. 3.57, we obtain

$$\frac{b_1 p_1}{\exp\left(\frac{\psi}{q_1^{\text{sat}}}\right) - 1} + \frac{b_2 p_2}{\exp\left(\frac{\psi}{q_2^{\text{sat}}}\right) - 1} = 1 \tag{3.61}$$

which can be analytically solved for ψ when the saturation capacities q_1^{sat} and q_2^{sat} are equal.

$$\exp\left(\frac{\psi}{q^{sat}}\right) = 1 + b_1 p_1 + b_2 p_2 \tag{3.62}$$

Using this in Eq. 3.60 we find the sorption pressures

$$p_i^*(\psi) = \frac{b_1 p_1 + b_2 p_2}{b_i} \tag{3.63}$$

Substituting the sorption pressures into Raoult's law Eq. 3.47, we obtain the adsorbed phase mole fractions

$$x_i = \frac{b_i p_i}{b_1 p_1 + b_2 p_2} \tag{3.64}$$

The total amount adsorbed is obtained by substituting Eq. 3.58, Eq. 3.63,and Eq. 3.64 into Eq. 3.50

$$q_T = q^{sat} \frac{b_1 p_1 + b_2 p_2}{1 + b_1 p_1 + b_2 p_2} \tag{3.65}$$

and the adsorbed amount contributed by the component i is

$$q_i = q^{sat} \frac{b_i p_i}{1 + b_1 p_1 + b_2 p_2} \quad i = 1, 2 \tag{3.66}$$

which is the extended Langmuir equation (Eq. 3.33) with equal saturation capacities.

Analytic Taylor series expansions of LeVan and Vermeulen [220]

Le Van and Vermeulen [220] derived the explicit isotherms for binary mixtures from the Gibbs adsorption isotherm. These authors considered single-site Langmuir and Freundlich isotherms for the pure components. However, this method can be extended to any kind of pure component isotherms, provided these isotherms have an explicit expression for the reduced grand potential (ψ) in terms of pure component pressure $\left(p_i^*\right)$. This method is applicable for cases where the saturation capacities of the components are close to each other [64]. The derivation is as follows:

According to IAST, the reduced grand potential for pure component i is (ψ_i) and is expressed as

$$\psi_i = \int_0^{p_i^*} \frac{q_i}{p_i} \mathrm{d}p_i \tag{3.67}$$

Also, the reduced grand potential of the mixture $\left(\psi^{\text{mix}}\right)$ is considered to be equal to the pure-component grand potential (ψ_i) at pressure, p_i^* for component i and the temperature of the mixture, i.e.,

$$\psi_i = q_i^{\text{sat}} \ln\left(1 + b_{i,1} p_i^* + b_{i,2}\left(p_i^*\right)^2\right) \tag{3.68}$$

For the ideal adsorbed solution, the partial pressures and the adsorbed phase composition are related by Raoult's law analogy.

$$p_i = x_i p_i^*(\psi) \tag{3.69}$$

In Eq. 3.69, p_i^* is the pure component pressure in equilibrium with an adsorbed phase of the component i at the reduced grand potential (ψ) and temperature of the mixture. p_i is the partial pressure of component i and x_i is the mole fraction in the adsorbed phase. Using the relation $\sum_i^{N_c} x_i = 1$ and substituting x_i using Eq. 3.69 yields

$$\frac{p_1}{p_1^*} + \frac{p_2}{p_2^*} = 1 \tag{3.70}$$

The single component isotherms (Langmuir and Freundlich) are substituted into Eq. 3.67, which on integration yields expressions for pure component pressures $\left(p_1^*,\ p_2^*\right)$. These expressions are used to further substitute p_1^* and p_2^* in Eq. 3.70. The resulting expression, which is a function of p_1, p_2 and ψ is expanded using Taylor series to obtain an explicit expression. The reduced grand potential ψ is differentiated to calculate the binary isotherms.

$$q_i = p_i \frac{\partial \psi}{\partial p_i} \tag{3.71}$$

Le Van and Vermeulen [220] derived the expressions for binary mixtures where both components either obey Langmuir or Freundlich isotherms in their pure form.

1. Langmuir Isotherms

   For pure component i, the reduced grand potential (ψ_i) equals:

$$\psi = q_i^{\text{sat}} \ln\left(1 + b_i p_i^*\right) \tag{3.72}$$

$$p_i^* = \frac{1}{b_i}\left[\exp\left(\frac{\psi}{q_i^{\text{sat}}}\right) - 1\right] \tag{3.73}$$

p_1^* and p_2^* in Eq. 3.70 are substituted using Eq. 3.73, which yields

$$\frac{b_1 p_1}{\exp\left(\frac{\psi}{q_1^{sat}}\right)-1}+\frac{b_2 p_2}{\exp\left(\frac{\psi}{q_2^{sat}}\right)-1}=1 \tag{3.74}$$

Since the reduced grand potentials for the pure components are equal to the mixture (Eq. 3.68), ψ will be used as the grand potential for both the components. To proceed with the derivation, the following parameters are defined:

$$q^{sat}=\frac{q_1^{sat}+q_2^{sat}}{2} \tag{3.75}$$

$$\epsilon=\frac{q_1^{sat}-q_2^{sat}}{2q^{sat}} \tag{3.76}$$

Combining Eqs. 3.75 and 3.76 yields

$$q_1^{sat}=q^{sat}(1+\epsilon) \tag{3.77}$$

$$q_2^{sat}=q^{sat}(1-\epsilon) \tag{3.78}$$

Eq. 3.74 can be expanded using the Taylor series for ψ about $\epsilon=0$.

$$\psi=\psi|_{\epsilon=0}+\frac{\epsilon}{1!}\frac{d\psi}{d\epsilon}\bigg|_{\epsilon=0}+\frac{\epsilon^2}{2!}\frac{d^2\psi}{d\epsilon^2}\bigg|_{\epsilon=0}+\ldots \tag{3.79}$$

Le Van and Vermeulen obtained isotherms for the binary mixture components using the Taylor expansion of Eq. 3.74. The isotherms are calculated using Eq. 3.71. The grand potential (ψ) is as follows

$$\psi=q^{sat}\ln(1+b_1p_1+b_2p_2) \tag{3.80}$$

In Eq. 3.80, q^{sat} is the average saturation loading which depends on the nature of the Taylor expansion (i.e. the number of terms considered in the expansion). These authors have derived explicit isotherms by considering two- and three-term Taylor series expansions of ψ.

(A) Two-term expansion:

$$q^{sat}=\frac{q_1^{sat}b_1p_1+q_2^{sat}b_2p_2}{b_1p_1+b_2p_2} \tag{3.81}$$

$$q_1^{sat}=q^{sat}\frac{b_1p_1}{1+b_1p_1+b_2p_2}+\Delta_{L2} \tag{3.82}$$

where

$$\Delta_{L2} = \left(q_1^{sat} - q_2^{sat}\right) \frac{b_1 b_2 p_1 p_2}{(b_1 p_1 + b_2 p_2)^2} \ln(1 + b_1 p_1 + b_2 p_2) \tag{3.83}$$

(B) Three-term expansion:

$$q^{sat} = \frac{b_1 p_1 q_1^{sat} + b_2 p_2 q_2^{sat}}{b_1 p_1 + b_2 p_2} + 2\frac{\left(q_1^{sat} - q_2^{sat}\right)^2}{q_1^{sat} + q_2^{sat}} \frac{b_1 b_2 p_1 p_2}{(b_1 p_1 + b_2 p_2)^2} \left[\left(\frac{1}{b_1 p_1 + b_2 p_2} + \frac{1}{2}\right) \ln(1 + b_1 p_1 + b_2 p_2) - 1\right] \tag{3.84}$$

$$q_1^{sat} = q^{sat} \frac{b_1 p_1}{1 + b_1 p_1 + b_2 p_2} + \Delta_{L2}(1 + \Delta_{L3}) \tag{3.85}$$

where

$$\Delta_{L3} = \frac{q_1^{sat} - q_2^{sat}}{q_1^{sat} + q_2^{sat}} \frac{1}{b_1 p_1 + b_2 p_2} \left[\frac{(b_2 p_2)^2 + 2(b_2 p_2) - 4(b_1 p_1) - (b_1 p_1)^2}{b_1 p_1 + b_2 p_2} \ln(1 + b_1 p_1 + b_2 p_2) + \frac{3(b_1 p_1)^2 + 4(b_1 p_1) + b_1 b_2 p_1 p_2 - 2(b_2 p_2) - 2(b_2 p_2)^2}{1 + b_1 p_1 + b_2 p_2} \right] \tag{3.86}$$

The multi-component adsorption isotherm for component 2 can be obtained by simply interchanging the subscripts 1 and 2 in the above equations.

2. Freundlich Isotherm

An approach similar to the above case is applied to a binary mixture where the pure components obey the Freundlich isotherm, which is shown below

$$q_i^{sat} = b_i p^{\nu_i} \tag{3.87}$$

The reduced grand potential (ψ) equation in terms of the partial pressure is

$$\left(\frac{b_1}{\nu_1}\right)^{1/\nu_1} p_1 \exp\left(-\frac{1}{\nu_1} \ln \psi\right) + \left(\frac{b_2}{\nu_2}\right)^{1/\nu_2} p_2 \exp\left(-\frac{2}{\nu_2} \ln \psi\right) = 1 \tag{3.88}$$

For unequal Freundlich exponents (ν_i) but very close to each other, the value of ψ can be computed using the Taylor series expansion. The expression for the adsorbed loadings for the components in the mixture is derived using the Taylor series expansion of ψ about ϵ (Eq. 3.79). For Freundlich isotherms, ϵ is

$$\epsilon = \frac{\nu_1 - \nu_2}{\nu_1 + \nu_2} \tag{3.89}$$

In Eq. 3.89, ν_1 and ν_2 are the Freundlich exponents for components 1 and 2, respectively. The expression for the adsorbed loading for component 1 (q_1) is shown below, a detailed derivation for which can be found in Ref. [220].

$$q_1 = \frac{\overline{\nu}\left(\frac{b_1}{\nu_1}\right)^{1/\nu_1} p_1}{\left[\left(\frac{b_1}{\nu_1}\right)^{1/\nu_1} p_1 + \left(\frac{b_2}{\nu_2}\right)^{1/\nu_2} p_2\right]^{1-\overline{\nu}}} + \Delta_{F2} \tag{3.90}$$

where

$$\Delta_{F2} = (\nu_1 - \nu_2) \frac{\left(\frac{b_1}{\nu_1}\right)^{1/\nu_1} p_1 \left(\frac{b_2}{\nu_2}\right)^{1/\nu_2} p_2}{\left[\left(\frac{b_1}{\nu_1}\right)^{1/\nu_1} p_1 + \left(\frac{b_2}{\nu_2}\right)^{1/\nu_2} p_2\right]^{2-\nu}} \ln\left[\left(\frac{b_1}{\nu_1}\right)^{1/\nu_1} p_1 + \left(\frac{b_2}{\nu_2}\right)^{1/\nu_2} p_2\right] \tag{3.91}$$

$$\overline{\nu} = \frac{\nu_1 \left(\frac{b_1}{\nu_1}\right)^{1/\nu_1} p_1 + \nu_2 \left(\frac{b_2}{\nu_2}\right)^{1/\nu_2} p_2}{\left(\frac{b_1}{\nu_1}\right)^{1/\nu_1} p_1 + \left(\frac{b_2}{\nu_2}\right)^{1/\nu_2} p_2} \tag{3.92}$$

Analytic approach of Ilic et al. [221]

Ilic et al. [221] derived an explicit competitive isotherm model for binary mixtures, where the pure components obey quadratic isotherms. For pure components, the reduced grand potential is

$$\psi_i = q_i^{sat} \ln\left(1 + b_{i,1} p_i^* + b_{i,2} \left(p_i^*\right)^2\right) \tag{3.93}$$

The grand potential of the mixture is considered to be equal to the pure-component grand potential (ψ_i) at pressure, p_i^* for component i and the temperature of the mixture, i.e.,

$$\psi = \psi_1 = \psi_2 \tag{3.94}$$

Using Eqs. 3.93, and 3.94 and considering the saturation capacities of the components to be equal $\left(q_1^{\text{sat}} = q_2^{\text{sat}}\right)$, we obtain

$$\ln\left(1 + b_{1,1}p_1^* + b_{1,2}\left(p_1^*\right)^2\right) = \ln\left(1 + b_{2,1}p_1^* + b_{2,2}\left(p_1^*\right)^2\right) \tag{3.95}$$

In Eq. 3.95, $b_{i,j}$ represents the equilibrium constant. This equation can be reformulated using a cubic polynomial, which is defined as a function of the adsorbed mole fraction of component 1 (x_1).

$$F(p_1, p_2, x_1) = A(p_1, p_2)x_1^3 + B(p_1, p_2)x_1^2 + C(p_1, p_2)x_1 + D(p_1, p_2) \tag{3.96}$$

where

$$A(p_1, p_2) = \alpha p_1 + p_1 > 0 \tag{3.97}$$
$$B(p_1, p_2) = p_1(\beta p_1 - 2\alpha) - p_2(1 + \gamma p_2) \tag{3.98}$$
$$C(p_1, p_2) = p_1(\alpha - 2\beta p_1) \tag{3.99}$$
$$D(p_1) = \beta p_1^2 \tag{3.100}$$
$$\alpha = \frac{b_{1,1}}{b_{2,1}} > 0 \tag{3.101}$$
$$\beta = \frac{b_{1,2}}{b_{2,1}} \geq 0 \tag{3.102}$$
$$\gamma = \frac{b_{2,2}}{b_{2,1}} \geq 0 \tag{3.103}$$
$$b_{i,1} > 0 \quad i = 1, 2 \tag{3.104}$$
$$b_{i,2} \geq 0 \quad i = 1, 2 \tag{3.105}$$

The root (x_1) of this cubic polynomial (Eq. 3.96) can be derived analytically by applying the formulae of Cardano [222]. The problem is solved by selecting the physically meaningful root out of the three obtained roots. Once the value of x_1 is computed, the adsorbed loading for component i (q_i) is calculated using the following expression [54]:

$$q_i = \left[\frac{x_1}{q_1^*(p_1^*)} + \frac{1 - x_1}{q_2^*(p_2^*)}\right] x_i \tag{3.106}$$

In Eq. 3.106, x_i is the adsorbed mole fraction of component i, q_i^* is the adsorbed loading for pure component i, and p_i^* is the corresponding gas phase pressure.

Analytic approach of Tarafder and Mazzotti [116]

Tarafder and Mazzotti obtained explicit isotherms for binary mixtures using IAST, where the pure components obey Langmuir-, anti-Langmuir- [157],

Brunauer-Emmett-Teller (BET)-, and quadratic-type adsorption isotherms. These mixture isotherms are valid only when the saturation capacities of the components are identical $\left(q_1^{\text{sat}} = q_2^{\text{sat}} = q^{\text{sat}}\right)$.

Assuming both the fluid and the adsorbed phase as ideal

$$p_i = p_i^* x_i \quad (i = 1, 2) \tag{3.107}$$

p_i is the concentration of component i in the mixture, p_i^* is the pure component concentration, and q_i^{sat} is the saturation loading in Eq. 3.107. Explicit solutions for the adsorbed phase mole fractions (x_1 and x_2) are derived from

$$\psi\left(\frac{p_1}{x_1}\right) = \psi\left(\frac{p_1}{x_2}\right) \tag{3.108}$$

$$x_1 + x_2 = 1 \tag{3.109}$$

The reduced grand potentials (ψ) for component i, computed based on the above-mentioned single-component isotherms, are as follows:

- Langmuir

$$\psi_i = q_i^{\text{sat}} \ln\left(1 + b_i p_i^*\right) \tag{3.110}$$

- Anti-Langmuir

$$\psi_i = -q_i^{\text{sat}} \ln\left(1 - b_i p_i^*\right) \tag{3.111}$$

- BET

$$\psi_i = q_i^{\text{sat}} \ln\left(\frac{1 - \left(b_{i,1} p_i^* + b_{i,2} p_i^*\right)}{1 - b_{i,2} p_i^*}\right) \tag{3.112}$$

- Quadratic

$$\psi_i = q_i^{\text{sat}} \ln\left(1 + b_{i,1} p_i^* + b_{i,2} p_i^{*\,2}\right) \tag{3.113}$$

In Eqs. 3.110 and 3.111, b_i is the equilibrium constant. Similarly, in Eqs. 3.112 and 3.113, $b_{i,j}$ $(j = 1, 2)$ represents the equilibrium constants. The adsorbed mole fraction of component 1 (x_1), and the total adsorbed loading $\left(q_T = q_1 + q_2\right)$ are calculated for different cases depending on the type of single-component isotherms followed by each component in the mixture.

1. Quadratic isotherm (component 1) and BET isotherm (component 2)

$$\frac{x_1}{p_1} = \frac{(b_{1,1}(1-b_{2,2}p_1)-b_{1,2}p_1)}{2(b_{1,1}p_1+b_{2,1}p_1)} + \frac{\sqrt{(b_{1,1}(1-b_{2,2}p_1)-b_{1,2}p_1)^2+4b_{1,2}(b_{1,1}p_1+b_{2,1}p_1)(1-b_{2,2}p_1)}}{2(b_{1,1}p_1+b_{2,1}p_1)} \tag{3.114}$$

$$q_T = q^{\text{sat}}\left[\frac{x_1\left(x_1^2+b_{1,1}p_1x_1+b_{1,2}p_1^2\right)}{\left(b_{1,1}x_1+2\,b_{1,2}p_1\right)p_1} + \frac{\left(x_2-b_{2,2}p_1\right)\left(x_2-b_{2,2}p_1+b_{2,1}p_1\right)}{b_{2,1}p_1}\right]^{-1} \tag{3.115}$$

2. Quadratic isotherm (component 1) and Langmuir isotherm (component 2)

$$\frac{x_1}{p_1} = \frac{(b_{1,1}-b_{1,2}p_1)+\sqrt{(b_{1,1}-b_{1,2}p_1)^2+4b_{1,2}(b_{1,1}p_1+b_2p_1)}}{2(b_{1,1}p_1+b_2p_1)} \tag{3.116}$$

$$q_T = q^{\text{sat}}\left[\frac{x_1\left(x_1^2+b_{1,1}p_1x_1+b_{1,2}p_1^2\right)}{(b_{1,1}x_1+2b_{1,2}p_1)p_1} + \frac{x_2(x_2+b_2p_1)}{b_2p_1}\right]^{-1} \tag{3.117}$$

3. Quadratic isotherm (component 1) and anti-Langmuir isotherm (component 2)

$$\frac{x_1}{p_1} = \frac{(b_{1,1}(1-b_2p_1)-b_{1,2}p_1)}{2(b_{1,1}p_1+b_2p_1)} + \frac{\sqrt{(b_{1,1}(1-b_2p_1)-b_{1,2}p_1)^2+4b_{1,2}(b_{1,1}p_1+b_2p_1)(1-b_2p_1)}}{2(b_{1,1}p_1+b_2p_1)} \tag{3.118}$$

$$q_T = q^{\text{sat}}\left[\frac{x_1\left(x_1^2+b_{1,1}p_1x_1+b_{1,2}p_1^2\right)}{(b_{1,1}x_1+2b_{1,2}p_1)p_1} + \frac{x_2(x_2-b_2p_1)}{b_2p_1}\right]^{-1} \tag{3.119}$$

Analytic approach of Van Assche et al. [118] (Implemented in RUPTURA)

The explicit isotherm (EI) derived by Van Assche et al.[118] to account for adsorbate size effects on multi-component adsorption is explained in chapter 2. The segregated version (SEI), which accounts for adsorbent surface heterogeneity is also explained in chapter 2.

3.3.5. Numerical methods to solve the IAST

The "nested loop algorithm" [219]

Myers and Valenzuela presented a simple method for solving the IAST [219] that is iterative in nature, exploiting the Newton-Raphson method [51, 64, 223]. Tien et al. extended and improved the algorithm in order to reduce the computation time [223, 224]. The method is conceptually elegant, but has slower convergence and is dependent on a good first guess for convergence. This algorithm consists of the calculation of the reduced grand potential (ψ) and then an analytic inversion of the reduced grand potential to obtain expressions for the pure component sorption pressure $\left(p_i^*\right)$ as a function of the reduced grand potential.

Generalising the procedure to solve IAST, named the "nested loop" algorithm by Mangano et al. [225], the total pressure and gas phase composition are specified and a nested problem of a highly non-linear system of equations is solved [64, 219, 223]. Eqs. 3.47 and 3.48 can be combined to yield

$$F(\psi) = \sum_{i=1}^{N_C} x_i - 1 = \sum_{i=1}^{N_C} \frac{y_i p}{p_i^*(\psi)} - 1 = 0 \tag{3.120}$$

which is a function of only the reduced grand potential (ψ). In the nested loop approach, a single implicit equation determines the solution of the IAST. To solve Eq. 3.120 for the reduced grand potential, an iterative method like the Newton-Raphson method can be used [64]. In Newton's method, the derivative of the reduced grand potential can be computed for any isotherm from [225]

$$\frac{\mathrm{d}\psi_i}{\mathrm{d}p_i^*} = \frac{q_i^*\left(p_i^*\right)}{p_i^*} \quad \left(i = 1, 2, \ldots, N_C\right) \tag{3.121}$$

and

$$\frac{\mathrm{d}F(\psi)}{\mathrm{d}\psi} = -\sum_{i=1}^{N_C} \frac{y_i p}{p_i^*(\psi)\, q_i^*\left(p_i^*\right)} \tag{3.122}$$

$p_i^*(\psi)$ can be calculated using Eq. 3.36, either analytically or numerically, depending on the pure component isotherm equation. The resulting proce-

dure is described in the book of Do [64].

$$\psi^{k+1} = \psi^k - \frac{F\left(\psi^{(k)}\right)}{F'\left(\psi^{(k)}\right)} \tag{3.123}$$

$$F\left(\psi^{(k)}\right) = \sum_{i=1}^{N_C} \frac{y_i p}{p_i^*\left(\psi^{(k)}\right)} - 1 \tag{3.124}$$

$$F'\left(\psi^{(k)}\right) = \left[-\sum_{i=1}^{N_C} \frac{y_i p}{\left[p_i^*(\psi)\right] q_i^*} \right]_{\psi=\psi^{(k)}} \tag{3.125}$$

An appropriate selection of the initial condition guarantees convergence [225]:

$$\psi = \min\left[\psi_i(p_i^*)\right] \tag{3.126}$$

The initial guess for ψ from Myers and Valenzuela [219] and Do [64]

$$\psi = \sum_{i=1}^{N_C} y_i \psi(p_i^*) \tag{3.127}$$

does not guarantee convergence strictly [225]. Convergence can be improved by combining Newton with a line search due to the monotonous nature of $F(\psi)$. The nested loop is a very fast numerical method if there exist explicit expressions for the sorption pressures.

FastIAS [205]

The FastIAS approach by O'Brien and Myers expresses the system of $N_C - 1$ equalities as [198, 205]

$$\begin{aligned}
\psi_1\left(p_1^*\right) &= \psi_2\left(p_2^*\right) \\
\psi_2\left(p_2^*\right) &= \psi_3\left(p_3^*\right) \\
&\cdots \\
\psi_{N_C-1}\left(p_{N_C-1}^*\right) &= \psi_{N_C}\left(p_{N_C}^*\right) \\
\sum_{i=1}^{N_C} \frac{p_i}{p_i^*} &= 1
\end{aligned} \tag{3.128}$$

This system constitutes N_C non-linear algebraic equations:

$$G\left(\mathbf{p}^*\right) = 0 \tag{3.129}$$

and equates the reduced grand potentials of successive components. The elements of the vector G are:

$$G_i\left(p^*\right) = \psi_i\left(p_i^*\right) - \psi_{i+1}\left(p_{i+1}^*\right) \qquad i = 1, \ldots, N_C - 1 \tag{3.130}$$

$$G_{N_C}\left(p^*\right) = \sum_{i=1}^{N_C} \frac{p_i}{p_i^*} - 1 \tag{3.131}$$

The N_C-variable Newton-Raphson method may be expressed as for the k-th iteration

$$p^{*(k+1)} \leftarrow p^{*(k)} - \delta^{(k)} \tag{3.132}$$

where δ is determined from the following linear equation

$$\Phi \cdot \delta = G \tag{3.133}$$

O'Brien and Myers added the following update rule to δ [226]

$$p_i^{*(k+1)} = \frac{1}{2} p_i^{*(k)} \qquad \text{if } p_i^{*(k)} + \delta^{(k)} < 0 \tag{3.134}$$

to ensure that the sorption pressures are always positive. The matrix Φ is the Jacobian matrix obtained by differentiating the vector G with respect to the vector p^*.

$$\Phi_{i,j}^{(k)} = \left(\frac{\partial G_i\left(p^{*(k)}\right)}{\partial p_j^{*(k)}} \right)_{p^*{}_{m \neq j}} \tag{3.135}$$

Due to the properties of the reduced grand potential Eq. 3.36, we have

$$\frac{\partial \psi\left(p_i^*\right)}{\partial p_i^*} = \frac{q\left(p_i^*\right)}{p_i^*} \tag{3.136}$$

and the elements of Φ are

$$\Phi_{i,i} = \frac{q\left(p_i^*\right)}{p_i^*} \qquad i = 1, \ldots, N_C - 1 \tag{3.137}$$

$$\Phi_{i,i+1} = -\frac{q\left(p_{i+1}^*\right)}{p_{i+1}^*} \qquad i = 1, \ldots, N_C - 1 \tag{3.138}$$

$$\Phi_{N_C,i} = -\frac{p_i}{\left(p_i^*\right)^2} \qquad i = 1, \ldots, N_C \tag{3.139}$$

This results in a Jacobian that has non-zero elements on the diagonal, the elements just above the diagonal and the last row [64, 205]:

$$\Phi = \begin{bmatrix} \times & \times & & & & \\ & \times & \times & & & \\ & & \times & \times & & \\ & & & \ddots & \ddots & \\ & & & & \times & \times \\ \times & \times & \times & \dots & \times & \times \end{bmatrix} \tag{3.140}$$

Thus, suitable starting values, $p^{*,\text{start}}$, as well as the Jacobian matrix of G with respect to the arguments p^* are needed. The convergence criterion is a relative test of the form

$$\sum_{i=1}^{N_C} \left| \frac{\delta_i^{(k)}}{p_i^{*(k+1)}} \right| < \varepsilon \tag{3.141}$$

The Fast IAS procedure is to solve firstly for the variable p_i^* rather than the reduced spreading pressure in the traditional IAS theory [64]. Solving directly for the sorption pressures of the pure components results in faster computation using the Fast IAS than the IAS theory. Once the spreading pressure is known in the IAS theory, the pure component sorption pressures are computed as the inverse of the integral Eq. 3.36.

Modified Fast IAS [226] (Implemented in RUPTURA)

The modified Fast IAS improves the original version by exploiting the peculiar structure of the equation set [226]:

$$\begin{aligned} \psi_1\left(p_1^*\right) &= \psi_{N_C}\left(p_{N_C}^*\right) \\ \psi_2\left(p_2^*\right) &= \psi_{N_C}\left(p_{N_C}^*\right) \\ &\dots \\ \psi_{N_C-1}\left(p_{N_C-1}^*\right) &= \psi_{N_C}\left(p_{N_C}^*\right) \\ \sum_{i=1}^{N_C} \frac{p_i}{p_i^*} &= 1 \end{aligned} \tag{3.142}$$

The elements of the vector G are:

$$G_i\left(p^*\right) = \psi_i\left(p_i^*\right) - \psi_{N_C}\left(p_{N_C}^*\right) \qquad i = 1, \dots, N_C - 1 \tag{3.143}$$

$$G_{N_C}\left(p^*\right) = \sum_{i=1}^{N_C} \frac{p_i}{p_i^*} - 1 \tag{3.144}$$

The elements of Φ are

$$\Phi_{i,i} = \frac{q\left(p_i^*\right)}{p_i^*} \qquad i = 1, \ldots, N_c - 1 \tag{3.145}$$

$$\Phi_{i,N_c} = -\frac{q\left(p_{N_c}^*\right)}{p_{N_c}^*} \qquad i = 1, \ldots, N_c - 1 \tag{3.146}$$

$$\Phi_{N_c,i} = -\frac{p_i}{\left(p_i^*\right)^2} \qquad i = 1, \ldots, N_c \tag{3.147}$$

In this modified method, the Jacobian matrix Φ is zero everywhere, except for the diagonal terms, the last column, and the last row.

$$\Phi = \begin{bmatrix} \times & & & & & \times \\ & \times & & & & \times \\ & & \times & & & \times \\ & & & \ddots & & \vdots \\ & & & & \times & \times \\ \times & \times & \times & \ldots & \times & \times \end{bmatrix} \tag{3.148}$$

With the peculiar structure of the Jacobian matrix of the above form, it can be reduced to a form such that all the last row of the reduced matrix contains zero elements, except the last element of that row. After the transformation

$$\Phi_{N_c,N_c}^{\text{new}} = \Phi_{N_c,N_c}^{\text{old}} - \sum_{j=1}^{N_c-1} \frac{\Phi_{N_c,j}}{\Phi_{j,j}} \Phi_{j,N_c} \tag{3.149}$$

$$G_{N_c}^{\text{new}} = G_{N_c}^{\text{old}} - \sum_{j=1}^{N_c-1} \frac{\Phi_{N_c,j}}{\Phi_{j,j}} G_j \tag{3.150}$$

and replacing the last term of the last row with $\Phi_{N_c,N_c}^{\text{new}}$ (denoted by a star), we end up with a Jacobian matrix that now has a simple sparse structure

$$\Phi^{\text{new}} = \begin{bmatrix} \times & & & & & \times \\ & \times & & & & \times \\ & & \times & & & \times \\ & & & \ddots & & \vdots \\ & & & & \times & \times \\ & & & & & * \end{bmatrix} \tag{3.151}$$

The result is a coefficient matrix having only a main diagonal and the last column. This system may be solved trivially by back-substitution. The

back-substitution is rendered even simpler since each row has only two entries. The solution for δ is simply [64]

$$\delta_{N_C} = \frac{G^{\text{new}}_{N_C}}{\Phi^{\text{new}}_{N_C,N_C}} \tag{3.152}$$

$$\delta_{N_C-1} = \frac{G_{N_C-1} - \delta_{N_C}\Phi_{N_C-1,N_C}}{\Phi_{N_C-1,N_C-1}} \tag{3.153}$$

$$\delta_{N_C-2} = \frac{G_{N_C-2} - \delta_{N_C}\Phi_{N_C-2,N_C}}{\Phi_{N_C-2,N_C-2}} \tag{3.154}$$

$$\dots \tag{3.155}$$

$$\delta_1 = \frac{G_1 - \delta_{N_C}\Phi_{1,N_C}}{\Phi_{1,1}} \tag{3.156}$$

The FastIAS algorithm is very fast, intrinsically robust, and will converge unless physically unrealistic (negative) values of the sorption pressures p_i^* are obtained [225].

Approach of Choy et al. [227]

Choy et al. [227] developed a numerical method in which, by specifying the adsorption amount (q_i) on the adsorbent in a multi-component system, the corresponding bulk phase concentrations or partial pressures (p_i) can be obtained. This method was developed for binary mixtures and is useful when an analytic expression for the reduced grand potential (ψ) cannot be obtained (e.g. pure components obeying Redlich-Peterson isotherm). However, this method is also applicable to cases where pure components obey other types of isotherms, such as Langmuir, Freundlich, Sips, etc. In addition, the value of the bulk phase concentration of component 2 in pure form $\left(p_2^*\right)$ should be determined using numerical procedures, such as the Newton-Raphson method based on the guessed value for the bulk phase concentration of component 1 in pure form $\left(p_1^*\right)$. The guessing parameter in the numerical solution $\left(p_1^*\right)$ changes in the iterative procedure until the criterion ($\psi_1 = \psi_2$) is met. Therefore, this procedure requires tremendous effort to find the equilibrium concentrations. The numerical integration scheme is shown below. The Redlich-Peterson isotherm is considered for the pure components.

$$\psi = \int_0^p \frac{q}{p}\,\mathrm{d}p = \int_0^p \frac{k_{\text{R}}}{1 + a_{\text{R}}p^{\beta}}\,\mathrm{d}p \tag{3.157}$$

For pure components 1 and 2, the expressions for ψ are:

$$\psi_1 = \int_0^{p_i^*} \frac{k^*_{\text{R},1}}{1 + a^*_{\text{R},1}p^{\beta^*_1}}\,\mathrm{d}p \tag{3.158}$$

$$\psi_2 = \int_0^{p_2^*} \frac{k_{R,2}^*}{1 + a_{R,2}^* p^{\beta_2^*}} \, dp \tag{3.159}$$

$$\frac{1}{q_T} = \frac{x_1}{q_1^*} + \frac{x_2}{q_2^*} \tag{3.160}$$

In Eq. 3.160, q_T is the total loading, and x_1 and x_2 are the mole fractions of the components 1 and 2, respectively. q_1^* and q_2^* are the pure component loadings. Using the values for q_1, and q_2, x_1, x_2, and q_T can be obtained as follows:

$$q_T = q_1 + q_2 \tag{3.161}$$

$$x_1 = \frac{q_1}{q_T} \tag{3.162}$$

$$x_2 = \frac{q_2}{q_T} \tag{3.163}$$

Applying the Redlich-Peterson equation in Eq. 3.160 yields

$$\frac{1}{q_T} = \frac{x_1}{\frac{k_{R,1}^* p_i^*}{1 + a_{R,1}^* {p_i^*}^{\beta_1^*}}} + \frac{x_2}{\frac{k_{R,2}^* p_2^*}{1 + a_{R,2}^* {p_2^*}^{\beta_2^*}}} \tag{3.164}$$

Another numerical method, such as Newton-Raphson, is required to solve Eq. 3.164. In this approach, p_2^* is computed from a given value of p_1^*. Then the values of ψ_1 and ψ_2 are obtained by applying the numerical integration scheme for Eqs. 3.158 and 3.159. The value of p_1^* is optimized till $\psi_1 = \psi_2$. The values of the bulk phase concentration (p_1 and p_2) are obtained using Raoult's law analogy as shown below.

$$p_1 = x_1 p_1^* \tag{3.165}$$

$$p_2 = x_2 p_2^* \tag{3.166}$$

Padé approximants of Frey and Rodrigues [228]

In this study, an explicit solution for IAS theory is developed. It involves fitting the relation between the reduced grand potential (ψ) and the pure component gas phase pressure $\left(p_i^*\right)$ to a three-parameter Padé approximation [228]. This method can be applied to a variety of single-component isotherms [228]. A Padé approximant can be used to represent the pure component pressure $\left(p_i^* = f^{-1}(\psi)\right)$. The accuracy of the expression depends on the kind of approximant employed. An important idea presented

in this publication is to fit the equivalent Padé representation directly to the experimental equilibrium data. One pitfall that can be quickly identified with this method is the inability of the Padé approximation to represent the equilibrium data accurately over sufficiently wide concentration ranges, as required for IAST calculations. The steps involved in this approach are as follows:

1. Determine the functions $\psi\left(p_i^*\right)$ and $p_i^*\left(\psi\right)$ for each components form the single component isotherms. For pure component i obeying Langmuir isotherm, $\psi\left(p_i^*\right)$ equals

$$\psi\left(p_i^*\right) = q_i^{\text{sat}} \ln\left(1 + b_i p_i^*\right) \tag{3.167}$$

   For different values of p_i^*, ψ can be obtained using Eq. 3.167. The expression for the pure component pressure $\left(p_i^*\left(\psi\right)\right)$ is

$$p_i^*\left(\psi\right) = \frac{1}{k_i}\left(\exp\left(\frac{\psi}{q_i^{\text{sat}}}\right) - 1\right) \tag{3.168}$$

2. Fit the function $p_i^*\left(\psi\right)$ to a three-parameter half Padé approximation

$$p_i^* = \frac{\psi}{a_i + f_i\psi + c_i\psi^2} \tag{3.169}$$

   In Eq. 3.169, a_i, f_i and p_i are the coefficients of the three parameter half Padé approximation.

3. Substitute Eq. 3.169 into $\sum_{i=1}^{N_C} x_i = \sum_{i=1}^{N_C} p_i/p_i^* = 1$ which leads to a quadratic formula for ψ:

$$0 = \psi^2 \sum_{i=1}^{N_C} c_i p_i + \psi \sum_{i=1}^{N_C} f_i p_i - \psi + \sum_{i=1}^{N_C} a_i p_i \tag{3.170}$$

4. Determine the explicit multi-component isotherms by combining Eq. 3.170 and $q_i = p_i \frac{\partial \psi}{\partial p_i}$ as shown below.

$$q_i = p_i \frac{\psi^2 c_i + \psi f_i + a_i}{1 - \sum_{i=1}^{N_C} f_i p_i - 2\psi \sum_{i=1}^{N_C} c_i p_i} \tag{3.171}$$

Surface B-splines fitting [229]

In this method, the IAST calculations are performed once to obtain the multi-component adsorption data for each species [229]. These datasets are fitted

with B-splines to determine the set of coefficients involved in the equations for the splines. Once the coefficients are obtained, interpolation methods can be used to obtain the equilibrium loadings at any concentration.

B-splines are also known as basis splines. They are constructed using polynomial functions joined at nodes. For a set of knots or nodes, λ_j (with $j = 0$ to $g + 1$), $g - k$ independent B-splines ($N_{i,k+1}(x)$) of degree k can be constructed [229]. Each spline function (s) is evaluated at $x \in \left[\lambda_0, \lambda_{g+1}\right]$ as follows

$$s(x) = \sum_{i=-k}^{g} a_i N_{i,k+1}(x) \tag{3.172}$$

In Eq. 3.172, a_i are known as B-spline coefficients. $N_{i,k+1}(x)$ is obtained recursively using the following equation.

$$N_{i,k+1}(x) = \frac{x - \lambda_i}{\lambda_{i+k} - \lambda_i} N_{i,k}(x) + \frac{\lambda_{i+k+1} - x}{\lambda_{i+k+1} - \lambda_{i+1}} N_{i+1,k}(x) \tag{3.173}$$

$$N_{\mathrm{i},1} = \begin{cases} 1, & x \in [\lambda_i, \lambda_{i+1}] \\ 0, & x \notin [\lambda_i, \lambda_{i+1}] \end{cases} \tag{3.174}$$

B-splines can be applied to surface approximation using tensor products [229]. Considering a data point (x, y) belonging to $[\lambda_i, \lambda_{i+1}] \times \left[\mu_j, \mu_{j+1}\right]$, the spline ($s$) of degree, k in x, and l in y is as follows

$$s(x, y) = \sum_{i=-k}^{g} \sum_{j=-l}^{h} a_{i,j} N_{i,k+1}(x) M_{j,l+1}(y) \tag{3.175}$$

where $N_{i,k+1}(x)$ and $M_{j,l+1}(y)$ are the normalized B-splines defined on the knots, λ and μ respectively. This surface approximation can be used to calculate the equilibrium loadings in case of multi-component adsorption. For a binary mixture, the adsorption loadings ($q_1(p_1, p_2), q_2(p_1, p_2)$) can be calculated using Eq. 3.175. The B-spline coefficients $\left(a_{i,j}\right)$ are obtained by fitting the surface spline to the data points of the equilibrium loadings as a function of the bulk phase concentrations (p_1 and p_1). This reduces the amount of iterative calculations involved in IAST.

Approach of Landa et al. [198]

The approach of Landa et al. is based on transforming the algebraic system of IAST equations to a system of ODEs with one specified initial value [198]. Starting with

$$\psi\left(p_1^*\right) = \psi_i\left(p_i^*\right) \quad (i = 1, 2, \ldots, N) \tag{3.176}$$

which, after differentiation, reads

$$\frac{\mathrm{d}\psi\left(p_1^*\right)}{\mathrm{d}p_1^*} = \frac{\mathrm{d}\psi_i\left(p_i^*\right)}{\mathrm{d}p_i^*}\frac{\mathrm{d}p_i^*}{\mathrm{d}p_1^*} \quad (i = 2, \ldots, N) \tag{3.177}$$

Eq. 3.177 can be transformed into an integration of a system of $N-1$ decoupled ordinary differential equations

$$\begin{aligned} &\frac{\mathrm{d}p_i^*}{\mathrm{d}p_1^*} = \frac{q_1^*\left(p_1^*\right)}{p_1^*}\frac{p_i^*}{q_i^*\left(p_i^*\right)}, \quad p_i^*(0) = 0 \quad (i = 2, \ldots, N) \\ &\frac{\mathrm{d}p_i^*}{\mathrm{d}p_1^*} = 1, \quad p_i^*(0) = 0 \quad (i = 1) \end{aligned} \tag{3.178}$$

The integration proceeds until the equilibrium condition Eq. 3.120 is met.

The calculation of the reduced grand potential is avoided, as well as the inversion of $\psi(p_i^*)$. The approach does not need a suitable vector of starting values to compute the desired adsorption equilibria.

Segregated IAST (SIAST) - Approach of Swisher et al [56] (Implemented in RUPTURA)

Instead of considering the available adsorption volume as a continuous space, the adsorbent material is divided into several distinct adsorption sites in SIAST. The competitive adsorption at each type of site takes place separately which are separately in thermodynamic equilibrium with the gas phase [56]. These sites are assumed to be uniform, and hence, IAST is applied individually to each type of site. This method is explained in detail in chapter 2.

Direct search minimization methods [194]

The method of Santori et al. [194] is based on the minimization of an objective function representing the iso-reduced grand potential condition. The reduced grand potentials are expressed in terms of molar fractions x_i through a preliminary change of variable from p_i^* to x_i substitution:

$$\mathrm{d}\ln\left[\frac{p_T y_i}{x_i}\right] = -\frac{1}{x_i}\mathrm{d}x_i \tag{3.179}$$

For the dual-site Langmuir, this becomes

$$\psi_i = \int -\left(\frac{q_{s1,i} b_{1,i} \left(\frac{p_T y_i}{x_i}\right)}{x_i + x_i b_{1,i} \left(\frac{p_T y_i}{x_i}\right)} + \frac{q_{s2,i} b_{2,i} \left(\frac{p_T y_i}{x_i}\right)}{x_i + x_i b_{2,i} \left(\frac{p_T y_i}{x_i}\right)} \right) \mathrm{d}x_i \tag{3.180}$$

$$= \int \left(-\frac{q_{s1,i} + q_{s2,i}}{x_i} + \frac{q_{s1,i}}{x_i + p_T y_i b_{1,i}} + \frac{q_{s2,i}}{x_i + p_T y_i b_{2,i}} \right) \mathrm{d}x_i \tag{3.181}$$

$$= -\left(q_{s1,i} + q_{s2,i}\right) \ln\left(x_i\right) + q_{s1,i} \ln\left(b_{1,1} y_i p_T + x_i\right) + q_{s2,i} \ln\left(b_{2,1} y_i p_T + x_i\right) \tag{3.182}$$

Santori et al. [194] derived similar expressions for the Toth, Unilan, and O'Brien–Myers functional forms. Using the derived reduced grand potentials ψ_i, the iso-ψ_i condition can be set as a minimization problem. For the binary case, we have

$$f_{\text{binary}}(x_i) = |\psi(x_1) - \psi(x_2)| \tag{3.183}$$

and the ternary case, we have

$$f_{\text{ternary}}(x_i, x_2) = |\psi(x_1) - \psi(x_2)| + |\psi(x_1) - \psi(x_3)| \tag{3.184}$$

For the more general ternary case, the optimization problem is stated as: Minimize the objective function $f_{\text{ternary}}(x_i, x_2)$ subject to constraints $0 < x_1 < 1 - x_2$ and $0 < x_2 < 1 - x_1$. Santari et al. used Nelder-Mead minimization algorithm [230] operated with a working precision of 10^{-8}, which resulted in final residuals ranging between 10^{-7} and 10^{-9}. The approach avoids pressure inversion and the initial guess, which are the typical requirements of the previous approaches. For binary systems, direct search minimization approach was reported to be faster than the classic IAST equations solution approach up to 19.0 (dual-site Langmuir isotherm) and 22.7 times (Toth isotherm). In ternary systems, this difference decreases to 10.4 (O'Brien and Myers isotherm) times. Compared to the FastIAS approach [205], direct search minimization is up to 4.2 times slower in ternary systems.

Approach of Mirzaei et al. [231]

Mirzaei et al. [231] developed a numerical solution for IAST, especially for cases where an explicit expression for the reduced grand potential (ψ) is not available. Numerical integration is used to calculate ψ. In this approach, the guessing parameter is the adsorbed mole fraction of component 1 (x_1), which is not directly involved in the integration to calculate ψ. The adsorbed mole fraction of component 2 is calculated using

$$\sum_{i=1}^{N_C} x_i = 1 \tag{3.185}$$

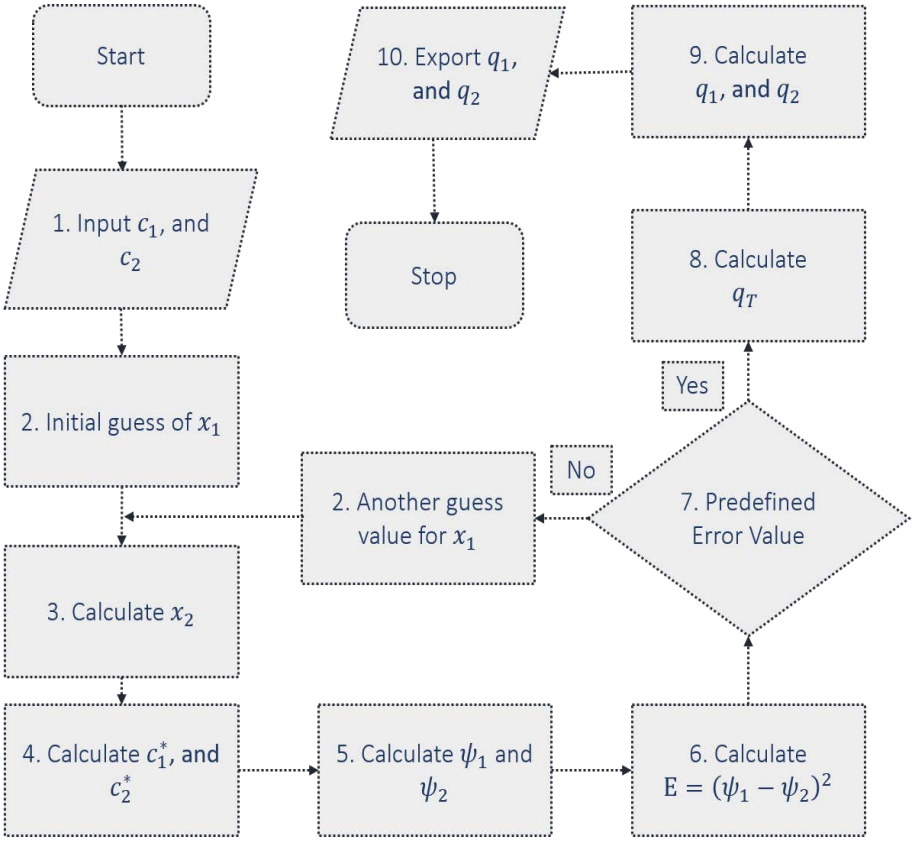


Figure 3.4. Numerical procedure to calculate equilibrium loadings for a binary mixture developed by Mirzaei et al. [231]. The bulk phase concentrations are the input to this approach, and the adsorbed mole fraction for component 1 (x_1) is used as the guessed value. Based on x_1, x_2 is estimated. Further, the grand potentials for components 1 and 2 are calculated separately, and the difference between these two values is computed. If the difference is less than the predefined error value, then the total equilibrium loadings and the loadings for each component are determined. Otherwise, a new guessed value for component 1 is considered, and the steps (3-7) are repeated.

The values of the pure component equilibrium concentrations are estimated as follows, which are also the upper limit in the integration to calculate the grand potential (ψ)

$$p_i = x_i p_i^* \tag{3.186}$$

The grand potentials are calculated separately for both components using

$$\psi_i = \int_0^{p_i^*} \frac{q_i^*(p)}{p}\,\mathrm{d}p \tag{3.187}$$

If the error $(|\psi_1 - \psi_2|)$ is less than the predefined error value, then the total loading is calculated as shown below. Otherwise, the steps mentioned above are repeated with another guessed value for the adsorbed mole fraction of component 1.

$$\frac{1}{q_T = \sum_{i=1}^{N_c} \frac{x_i}{q_i^*}} \tag{3.188}$$

Once, the total equilibrium loading (q_T) is obtained, q_1 and q_2 are calculated as follows

$$q_i = x_i q_T \tag{3.189}$$

The procedure for the numerical method developed by Mirzaei et al. [231] is shown in Fig. 3.4.

3.3.6. Mixture prediction software

Benjamin [232]

M. Benjamin freely provided an Excel spreadsheet and a Java program that carries out the IAST calculations that can be used to calculate the equilibrium distribution of species in a competitive adsorption system, based on the IAST Model. A manuscript containing background information on the IAST and the usefulness of these programs has been published in ES&T [232].

pyIAST [54]

pyIAST [54] is an open-source Python package, pyIAST, to perform IAST calculations for an arbitrary number of components. pyIAST supports several common analytical models to characterize the pure-component isotherms from experimental or simulated data. Alternatively, pyIAST can use numerical quadrature to compute the spreading pressure for IAST calculations by interpolating the pure-component isotherm data. pyIAST can also perform reverse IAST calculations, where one seeks the required gas phase composition to yield a desired adsorbed phase composition.

GAIAST [233]

GAIAST (Genetic Algorithm IAST) [233] is a Fortran 2008+ standard-compliant, free and open source project, aimed at using IAST to predict the loading of the gas mixture on the adsorbed phase, based only on the knowledge of the pure adsorption isotherms of the individual components.

IAST++ [234]

IAST++ [234] is a user-friendly graphic user interface software that can fit the adsorption data to various isotherm models and use the ideal adsorbed solution theory (IAST) to obtain mixture isotherm data. The authors found FastIAS to be quite unstable because it uses the Newton-Raphson method that is highly sensitive to initial conditions and therefore solved the system of equations by minimising the square of each equation and constraints using the downhill simplex (Nelder-Mead) method [230].

pyGAPS [235]

pyGAPS (Python General Adsorption Processing Suite) framework [235] is a versatile and extensible software package which can be used for both routine material characterization as well as complex adsorption isotherm processing. It contains many common characterization methods such as: Brunauer-Emmett-Teller and Langmuir surface area, t and α s plots, pore size distribution calculations (Barrett-Joyner-Halenda, Dollimore-Heal, Horvath-Kawazoe, DFT/NLDFT kernel fitting), isosteric enthalpy calculations, ideal adsorbed solution theory calculations, isotherm modeling and more, as well as the ability to import and store data from Excel, CSV, JSON and SQLite databases. pyGAPS includes the IAST implementation from pyIAST, which has been wrapped for interoperability with pyGAPS isotherm models.

GraphIAST [236]

GraphIAST (Python General Adsorption Processing Suite) framework [236] is a simple, user-friendly program for IAST loading and selectivity predictions for binary gas mixtures based on the Python module pyIAST. The authors developed a graphical user interface resembling commonly known software and made three-dimensional selectivity predictions easily accessible within just a few clicks. The input and output data structure relies on the widely used *.csv format, and isotherm data can be fitted with various established models.

3.3.7. IAST implementation in RUPTURA

FastIAS algorithm

The default IAS algorithm in RUPTURA is FastIAS. The implementation follows the description in Sec. 3.3.5. The original formulation of O'Brien and Myers is defined in terms of a reduced pressure. As a generalization, we use

the sorption pressures and a formulation that is isotherm model agnostic. However, instead of using a convergence test criterion like Eq. 3.141, we use the standard deviation of the obtained reduced grand potentials (using 10^{-13} as default precision).

For efficiency reasons, we cache the computed sorption pressures and reduced grand-potentials (for breakthroughs for each grid point). Since the next computation is very likely close to the current solution, the next solution is usually found within only one or two FastIAS steps. However, the first computation for IAST and breakthrough usually starts at conditions where loadings, sorption pressures, and reduced grand potentials can be very small. We found that at these conditions, using Eq. 3.127 worked better than using Eq. 3.126 for initialization.

Nested-loop algorithm using bisection

Our nested-loop implementation is based on

$$f(\psi) = \sum_{i=1}^{N} \frac{y_i p}{p_i^*(\psi)} = 1 \tag{3.190}$$

Our experience agrees with See et al. [234] that methods like Newton-Raphson do not seem able to numerically reduce the error to machine precision. We note that $\psi_i^*(p)$ is a continuous and, importantly, monotonic function, and therefore also is its inverse $p_i^*(\psi)$. We therefore use bi-section on the reduced grand potential ψ. This automatically guarantees there is a single reduced grand potential ψ that is equal for all components. The decision to reduce or increase ψ is based on whether the sum of the adsorbed mole fractions is larger than unity or not. Additionally, we use the bi-section algorithm to obtain $p_i^*(\psi)$ for the supported isotherm types (see Table 3.1). The bisection algorithm is described in detail in Algorithm 2 for IAST and in Algorithm 3 for inverse computations. The algorithm of Hoseinpoori [237] is also based on bisection, but on $x_1 + x_2$, which is therefore limited to binary components.

For a function f on an interval $[a, b]$, bisection iterates through the following steps:

1. Calculate the midpoint $c = \frac{1}{2}(a + b)$
2. Calculate the function value at the midpoint, $f(c)$
3. If convergence criteria is no met, examine whether $f(c)$ is smaller or larger than unity and replace either $(a, f(a))$ or $(b, f(b))$ with $(c, f(c))$ and continue with the new (smaller) interval.

The method converges linearly with rate $1/2$, i.e. the absolute error is halved at each step. The bisection method is insensitive to the starting values. However, we also cache $p_i^*(\psi)$ for all components, so that subsequent

evaluations use efficient starting values. The bisection method is slow, since (a) bisection itself is slow, (b) it needs the inverse reduced grand potential (numerically), and (c) the nested-loop algorithm requires many iterations. We have implemented this method based on its robustness and accuracy, i.e. it is able to always converge to machine precision. We use a relative tolerance of 10^{-15}. For our use cases in IAST and IAST in fixed-bed adsorption simulations, this was achieved in under 50 steps, even for complex 10-component mixtures.

The FastIAS method is much faster than the nested-loop algorithm using the bisection method. Therefore, it is used as the default method in RUPTURA. We also plan to include the false position method [238] in the next version of RUPTURA.

3.3.8. Validation

[223] and [198] considered a ten-component mixture described by the O'Brien and Myers adsorption isotherm model:

$$q(p) = q^{\text{sat}}\left[\frac{bp}{1+bp} + \sigma^2\frac{bp\,(1-bp)}{2\,(1+bp)^3}\right] \tag{3.191}$$

with reduced grand potential [64, 198]

$$\psi(p) = q^{\text{sat}}\left[\ln\left[1+bp\right] + \frac{\sigma^2 bp}{2\,(1+bp)^2}\right] \tag{3.192}$$

In Table 3.5, we compare our IAST results for a 10-component mixture of [223], also used in Landa et al. [198]. We obtain similar qualitative results, but quantitatively different in the third digit. The tolerance for the results of Moon et al. was 10^{-5}. Our work uses a relative precision of 10^{-15}. Our computed total loading is q_T = 2.26563527 mol/kg, and the sum of the adsorbed phase mole fractions $\sum_i x_i = 1$ is normalized to machine precision, compared to q_T=2.3050 mol/kg and $\sum_i x_i = 0.9912$ for Moon and Tien. The component reduced grand potentials for Moon and Tien are recomputed from the reported partial pressures and adsorbed mole fractions using Eq. 3.192 and hence, the accuracy is limited by the accuracy of the reported adsorbed mole fractions. Nevertheless, the data illustrate a relative precision of 10^{-5}, which is not enough to enforce identical reduced grand potentials ψ_i for all components.

In Table 3.6, we compare our IAST results for a 10-component mixture of [198]. The relative and absolute tolerances for the results of Landau et al. are 10^{-10} and 10^{-12}, respectively. This work uses a relative precision of 10^{-15}. Our computed total loading is q_T = 2.117678 mol/kg, compared to an absolute loading of q_T=2.1142 mol/kg by Landau et al. Note the value of $\psi = 3.76951$ mol/kg of Landa et al., computed from their reported loading data. The error in ψ is here due to the limited precision of the reported loading $q_9 = 0.0010$ mol/kg. Using our value of $q_9 = 0.001042$ in the computation yields $\psi_9 = 3.68052$ mol/kg.

Algorithm 2 IAST calculation using bi-section. The tolerance value ϵ can be set to 10^{-15}.

function IAST(p, y_i)
  $\psi \leftarrow$ initial value
  $\sum_i x_i \leftarrow \sum_i y_i \times p/p_i^*(\psi)$ ▷ /* Equation 3.47 */
  $\psi_\uparrow \leftarrow \psi$ ▷ /* Initialize the top bracket with the initial ψ value */
  $\psi_\downarrow \leftarrow \psi$ ▷ /* Initialize the bottom bracket with the initial ψ value */
  if $\sum_i x_i > 1$ **then** ▷ /* bottom bracket known, searching for top bracket */
    repeat
      $\psi_\uparrow \leftarrow 2\times \psi_\uparrow$
      $\sum_i x_i \leftarrow \sum_i y_i \times p/p_i^*(\psi_\uparrow)$ ▷ /* Equation 3.47 */
    until $\sum_i x_i < 1$
  else ▷ /* top bracket known, searching for bottom bracket */
    repeat
      $\psi_\downarrow \leftarrow \psi_\downarrow/2$
      $\sum_i x_i \leftarrow \sum_i y_i \times p/p_i^*(\psi_\downarrow)$ ▷ /* Equation 3.47 */
    until $\sum_i x_i > 1$
  end if
  repeat
    $\psi \leftarrow \frac{1}{2}(\psi_\downarrow + \psi_\downarrow)$ ▷ /* Use bisection to find the ψ within the interval */
    $\sum_i x_i \leftarrow \sum_i y_i \times p/p_i^*(\psi)$ ▷ /* Equation 3.47 */
    if $\sum_i x_i > 1$ **then** ▷ /* New range is ψ to $\psi_\uparrow$ */
      $\psi_\downarrow \leftarrow \psi$
    else ▷ /* New range is $\psi_\downarrow$ to ψ */
      $\psi_\uparrow \leftarrow \psi$
    end if
  until $\frac{|\psi_\uparrow - \psi_\downarrow|}{|\psi_\uparrow + \psi_\downarrow|} < \epsilon$ ▷ /* Iterate until $\psi_\uparrow$ and $\psi_\downarrow$ are equal to within precision ϵ */
  $\psi \leftarrow \frac{1}{2}(\psi_\uparrow + \psi_\downarrow)$ ▷ /* Final computed value of the reduced grand potential ψ */
  $\sum_i x_i \leftarrow \sum_i y_i \times p/p_i^*(\psi)$ ▷ /* Mole fractions sum to unity (within precision ϵ), equation 3.48 */
  for $i \in \{1, \ldots, N_c\}$ **do**
    $x_i \leftarrow \frac{y_i p}{p(\psi)}$ ▷ /* Equation 3.47 */
  end for
  $1/q^T \leftarrow \sum_i \frac{x_i}{f(p(\psi))}$ ▷ /* Equation 3.50 */
  for $i \in \{1, \ldots, N_c\}$ **do**
    $q_i \leftarrow x_i q^T$ ▷ /* Equation 3.52 */
  end for
  return ψ, x_i, q_i
end function

Algorithm 3 Inverse computation using bi-section. The tolerance value ϵ can be set to 10^{-15}.

function $p(\psi)$
  $p \leftarrow$ initial value
  $p_\uparrow \leftarrow p$ ▷ /* Initialize the top bracket with the initial p value */
  $p_\downarrow \leftarrow p$ ▷ /* Initialize the bottom bracket with the initial p value */
  if $\psi((p) > 1$ **then** ▷ /* bottom bracket known, searching for top bracket */
    repeat
      $p_\uparrow \leftarrow 2\times\ p_\uparrow$
    until $\psi\left(p_\uparrow\right) < 1$
  else ▷ /* top bracket known, searching for bottom bracket */
    repeat
      $p_\downarrow \leftarrow p_\downarrow/2$
    until $\psi\left(p_\downarrow\right) > 1$
  end if
  repeat
    $p \leftarrow \frac{1}{2}\left(p_\downarrow + p_\downarrow\right)$ ▷ /* Use bisection to find the p within the interval */
    if $\psi\left(p\right) > 1$ **then** ▷ /* New range is p to $p_\uparrow$ */
      $p_\downarrow \leftarrow p$
    else ▷ /* New range is $p_\downarrow$ to p */
      $p_\uparrow \leftarrow p$
    end if
  until $\frac{|p_\uparrow - p_\downarrow|}{|p_\uparrow + p_\downarrow|} < \epsilon$ ▷ /* Iterate until $p_\uparrow$ and $p_\downarrow$ are equal to within precision ϵ */
  $p \leftarrow \frac{1}{2}\left(p_\downarrow + p_\downarrow\right)$
  return p
end function

Table 3.5. Parameters for the O'Brien and Myers adsorption isotherm model considered by Moon and Tien [223] for a 10-component mixture at 300 kPa, q_T=2.3050 mol/kg and $\sum_i x_i = 0.9912$. The tolerance for the results of Moon et al was 10^{-5}. This work uses a relative precision of 10^{-15}. Our computed total loading is $q_T = 2.26563527$ mol/kg, and the sum of the adsorbed phase mole fractions $\sum_i x_i = 1$ is normalized to machine precision.

component	q_i^{sat} [mol/kg]	b_i [Pa^{-1}]	σ_i [-]	p_i [kPa]	y_i [-]	ψ_i [mol/kg] (Ref. [223])	ψ_i [mol/kg] (this work)	x_i [-] (Ref. [223])	x_i [-] (this work)
1	5.0	1.0×10^{-5}	1.2	30.0	0.1	4.261	4.19182774500228	0.3129	0.321867
2	3.0	0.6×10^{-5}	1.1	30.0	0.1	4.191	4.19182774500229	0.0698	0.069755
3	4.0	0.09×10^{-5}	0.8	30.0	0.1	4.194	4.19182774500228	0.0164	0.016412
4	2.0	1.0×10^{-5}	1.2	60.0	0.2	4.191	4.19182774500229	0.0926	0.092559
5	3.5	0.3×10^{-5}	1.0	30.0	0.1	4.191	4.19182774500228	0.0459	0.045873
6	4.0	0.1×10^{-5}	1.1	15.0	0.05	4.201	4.19182774500229	0.0102	0.010241
7	2.0	1.5×10^{-5}	1.2	15.0	0.05	4.192	4.19182774500228	0.0347	0.034710
8	2.5	0.1×10^{-5}	1.15	15.0	0.05	4.170	4.19182774500228	0.0040	0.003953
9	4.0	0.01×10^{-5}	1.0	15.0	0.05	4.145	4.19182774500228	0.0010	0.000980
10	5.5	0.6×10^{-5}	1.0	60.0	0.2	4.192	4.19182774500228	0.4037	0.403652

3.4. Fixed-Bed Adsorption Simulations

3.4.1. Introduction

Most industrial separation processes take place at dynamic conditions [60]. Apart from adsorption isotherms, it is also important to evaluate properties like cyclic stability, long time stability, competitive adsorption, and mass transfer kinetics [60]. Knowledge of these properties is vital for selecting an appropriate adsorbent for a specific separation process. These properties can be obtained using dynamic methods. In most dynamic methods, adsorbents are arranged in fixed beds [239] and the gas mixture is passed through them. Adsorption in fixed beds is carried out in cylindrical columns filled with particles of nanoporous materials. The nanoporous particles are held fixed in place and do not move (Fig. 3.5). Adsorbers are commonly operated in a transient mode. Consequently, the compositions of the gas phase and of the adsorbed molecules inside the adsorbent particles vary with position and time. Transient breakthrough curves are crucial for evaluating the performance of an adsorbent material to separate certain compounds, or to encapsulate a single component [240]. It is also important to acknowledge the fact that the study of breakthrough curves cannot be the sole criterion in designing an adsorption-based system. A full-scale process simulation is required to serve the purpose [241, 242]. Measurements of breakthrough curves through experiments involve several steps ranging from preparing the samples to analysing the results [100]. This procedure can be very time-consuming, especially when breakthrough curves are used to screen a large number of potential adsorbents for a specific separation process [100]. Therefore, breakthrough curve modeling can be very handy in designing adsorption-based separation columns and selecting the best adsorbent material for the process.

The fluid containing the solute flows through the packed bed of adsorbents, usually at a constant mass flow rate. The breakthrough curve has mainly three segments: the unsaturated zone, the mass transfer zone (MTZ) and the saturated zone [60]. These segments are shown in Fig. 3.6. The unsaturated zone is the region where the fluid phase and the adsorbed phase are not in equilibrium. The majority of the adsorption of the components takes place in this zone. The mass transfer zone (MTZ) is the range of the bed where the adsorption process rapidly progresses towards equilibrium. There is a rapid increase in the outlet concentration in this zone, which is influenced by the mass transfer between the gas phase and the adsorbed phase [243], axial dispersion, and heat transfer effects [60]. During operation, this zone moves along the bed from the inlet point to the outlet point. Inside the MTZ, the loading of absolute adsorbates varies from the saturation limit to 0%. When the front edge of the MTZ reaches the end point of the bed, the “breakthrough” occurs (often defined as when the concentration at the outlet reaches a predefined value, e.g., 1%, of the inlet concentration). The saturated zone determines the adsorption capacity of the components [60]. In this zone, the adsorbed phase is in equilibrium with the fluid phase.

Table 3.6. Parameters for the O'Brien and Myers adsorption isotherm model considered by Landau et al. [198] for a 10-component mixture, q_T=2.1142 mol/kg. The relative and absolute tolerances for the results of Landau et al. are 10^{-10} and 10^{-12}, respectively. This work uses a relative precision of 10^{-15}. Our computed total loading is q_T = 2.117678 mol/kg.

component	q_i^{sat} [mol/kg]	b_i [Pa^{-1}]	σ_i [-]	p_i [kPa]	y_i [-]	ψ_i [mol/kg] (Ref. [223])	ψ_i [mol/kg] (this work)	q_i [mol/kg] (Ref. [198])	q_i [mol/kg] (this work)
1	5.0	1.0×10^{-5}	1.2	30.0	0.1	3.68342	3.683573349491790	0.8443	0.845631
2	3.0	0.6×10^{-5}	1.1	3.0	0.01	3.68238	3.683573349491790	0.0192	0.019219
3	4.0	0.09×10^{-5}	0.8	3.0	0.01	3.69288	3.683573349491789	0.0043	0.004325
4	2.0	1.0×10^{-5}	1.2	15.0	0.05	3.68419	3.683573349491789	0.0677	0.067839
5	3.5	0.3×10^{-5}	1.0	60.0	0.2	3.68371	3.683573349491791	0.2467	0.247123
6	4.0	0.1×10^{-5}	1.1	60.0	0.2	3.68317	3.683573349491790	0.1093	0.109459
7	2.0	1.5×10^{-5}	1.2	30.0	0.1	3.68345	3.683573349491791	0.2032	0.203517
8	2.5	0.1×10^{-5}	1.15	60.0	0.2	3.68459	3.683573349491790	0.0446	0.044700
9	4.0	0.01×10^{-5}	1.0	6.0	0.02	3.76951	3.683573349491788	0.0010	0.001042
10	5.5	0.6×10^{-5}	1.0	33.0	0.11	3.68348	3.683573349491788	0.5739	0.574822

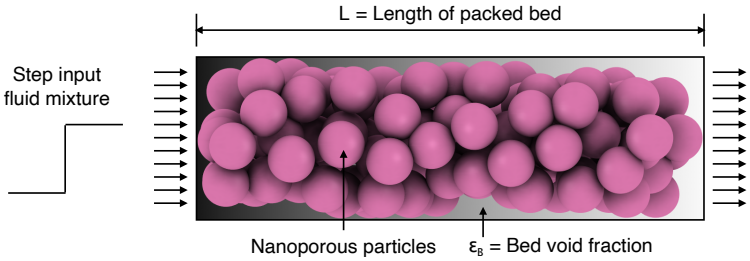


Figure 3.5. Schematic representation of a packed bed adsorber. The packed bed adsorber is a cylindrical column filled with nanoporous adsorbent material (crystalline or amorphous). When a fluid mixture is fed as input to the column, adsorption of various components present in the mixture takes place inside these nanoporous materials.

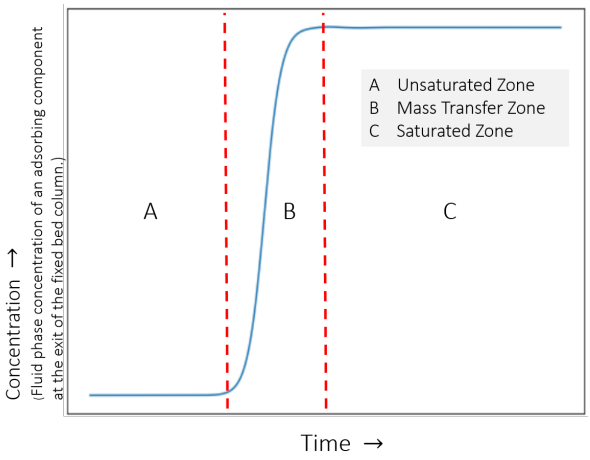


Figure 3.6. Different zones in a breakthrough curve at the exit of the fixed bed column. Fluid phase concentration of an adsorbing component as a function of time at the fixed bed column outlet. The profile is divided into three zones: (a) the unsaturated zone: where the maximum adsorption takes place, (b) mass transfer zone: where adsorption process progresses rapidly towards equilibrium, (c) saturated zone: where no more adsorption takes place as equilibrium between the fluid and the adsorbed phase is achieved.

Therefore, no net mass transfer from the fluid phase to the adsorbed phase takes place.

When the fluid flows across the solid surface, a stagnant film forms along that surface. Molecules from the bulk interstitial phase are transported via axial convection onto the film. Inside the film, there is a velocity gradient in

the direction of the flow but not in the perpendicular direction. Therefore, advection does not occur in the film and molecules diffuse across the stagnant film. The driving force of film diffusion is the concentration gradient located at the interface region between the exterior surface of the adsorbent particles and the bulk solution. The stagnant film thickness is proportional to the Reynolds number of the bulk fluid. The external mass transfer resistance is strongly dependent on the flow conditions, e.g. temperature, pressure, and superficial velocity [244]. The adsorbate molecules transfer between the fluid and the adsorbent particles through several steps [244–246]:

1. transport of adsorbate from bulk fluid to the localized hydrodynamic boundary layer around the adsorbent,
2. interface diffusion between fluid phase and the exterior surface of the adsorbent (often called “external”,“interphase”, or“film" diffusion),
3. intraparticle mass transfer involving pore diffusion and surface diffusion,
4. adsorption of molecules on the inner surface,
5. if the solid is a catalyst: reactions taking place at specific active sites inside the catalyst material,
6. desorption of the molecules from the inner surface,
7. diffusion of the products from the interior of the particle to the pore mouth at the external surface,
8. diffusion of the products from the external particle surface to the bulk fluid (interphase diffusion).

Molecules in the column can move in both axial and radial directions. In case of plug flow behavior, any variation in the radial direction is absent, so the velocity of the fluid, the concentration, and the porosity are assumed to be constant across any cross-section of the bed perpendicular to the axis of the bed. Variation of the velocity in the radial direction can be neglected if the flow is not highly viscous and no considerable boundary layer is formed in that direction. The assumption of neglecting radial gradients is widely accepted [247–249]. The effects of dispersion in all directions are included in the axial dispersion term [62]. To account for the adsorption, we use the Linear Driving Force (LDF) model [130]. In this model, the adsorption of molecules from all directions into adsorbent particles is lumped into a single entity, called the average adsorbed loading (q_i). One can also model the adsorption process by considering variation in the radial direction inside the zeolites. This is possible using the Fick diffusion model. Sircar and Hufton compared the LDF model with the rigorous Fick diffusion model [131]. These authors found that the details regarding the intra-pore diffusion are

lost when breakthrough curves are modelled [131]. Therefore, the authors concluded that in most cases, the LDF model is sufficient to account for the effects of adsorption.

Various mathematical models can be derived from conservation of mass, energy, and momentum and augmented by appropriate transport rate equations and equilibrium isotherms [250]. Mathematical models of fixed-bed columns for adsorption are reviewed by Tien [155, 251], Worch [185, 252], Siahpoosh et al. [253], Xu et al. [246], Shafeeyan et al. [249], Supian et al. [254], Unuabonah et al. [255], and many others. These models differ mainly in the form of the mass transfer rate, the form of the equilibrium isotherm, thermal effects, and the pressure drop along the bed. Fixed-bed column adsorption studies are reviewed by Patel [239] on the removal of various contaminants from wastewater.

Fixed-bed models are either zero-phase, one-phase, two-phase or multiphase resistance models [62, 255]. A zero-phase model ignores mass transfer and assumes instantaneous equilibrium along the column. The one-phase resistance models assume local equilibrium in any cross section along the bed. These models take dispersion and/or diffusion into account but neglect intra-particle diffusion resistance and intra-particle axial dispersion. The one-phase models, known as chemical kinetic models, assume that the mass transfer rate is the difference between two opposing second-order reactions with different rate constants [65], and include [155, 255]: the Thomas model [256], Bohart-Adams model [257], Yoon-Nelson models [258, 259], and Clark model [260]. One-phase models have direct nonlinear mathematical expressions but are not very accurate in describing the breakthrough behavior in a column [261]. Two-phase resistance models take into account both film and intra-particle diffusion. These models are more accurate than one-phase resistance models but have to be solved numerically. Multiphase resistance models take multiple types of diffusion (film diffusion, pore diffusion, and surface diffusion) into account. In the diffusion models, the particle is treated as a homogeneous phase [65]. Note that often surface reaction rates are assumed to be much faster than diffusion. Surface reactions should be considered when these reactions affect the adsorption rate or are the controlling step. Plazinski et al. [262] reviewed sorption kinetic models including surface reaction mechanisms. Medved and Cerny [263] reviewed surface diffusion of particles in porous media.

3.4.2. Theory

To model the fixed bed adsorber column, the following assumptions are considered [253]:

- The fixed bed is tubular, and the adsorbent particles are spherical. The particles are packed uniformly into the fixed bed (no channelling occurs).

- No chemical reaction occurs in the column. This aspect will be considered in our future work.
- The fixed bed is initially filled with carrier gas.
- The gas phase behaves as an ideal gas mixture.
- The process is isothermal adsorption.
- The mass and velocity gradients are negligible in the radial direction of the bed.
- The pressure gradient (if any) is invariant with time and column position and is not affected by the adsorption process.
- The fluid velocity is varying according to the total mass balance along the bed in multi-component systems. The fluid velocity will vary along the length of the column because of the variation in total pressure and/or adsorption of the molecules from the fluid phase.
- The axial dispersion is considered in the bulk phase.
- The carrier gas does not adsorb.
- The physical properties of the gas phase (axial dispersion coefficient and mass transfer coefficient) are those of the feed gas. These properties along with the properties (particle density and bed porosity) of the adsorbent, are constant throughout the column.

For future versions of RUPTURA, we envision including non-ideal gas behavior, liquid phase conditions, non-isothermal breakthroughs, and chemical reactions.

We follow Worch [185] to outline the establishment of the material balance. Considering a differential volume element $dV = A_R dz$, it can be assumed that the amount of adsorbate that is adsorbed onto the adsorbent or accumulated in the void fraction of the volume element must equal the difference between the input and the output of the volume element. Input and output occur by advection and axial dispersion. The general mass balance equation is as follows [185]

$$\dot{N}_{\text{accu.}} + \dot{N}_{\text{ads.}} = \dot{N}_{\text{adv.}} + \dot{N}_{\text{disp.}} \tag{3.193}$$

where $\dot{N}$ represents the change in the amount of adsorbate with time, and the indices indicate the processes: accumulation, adsorption, advection, and dispersion.

- The accumulation term is:

$$\dot{N}_{\text{accu}} = \varepsilon_B dV \frac{\partial c(t,z)}{\partial t} \tag{3.194}$$

where c represents the adsorbate concentration in the fluid phase, z is the distance along the bed length, t denotes time, and ε_B is the bed void fraction.

- The adsorption term is expressed as:

$$\dot{N}_{\text{ads}} = \rho_B \, \mathrm{d}V \frac{\partial \bar{q}(t,z)}{\partial t} \tag{3.195}$$

$$= (1-\varepsilon_B)\rho_P \frac{\partial \bar{q}(t,z)}{\partial t} \tag{3.196}$$

where ρ_B is the bed density, ρ_P is the particle density. The units of both ρ_B and ρ_P are $\left[\mathrm{kg/m^3}\right]$. $\bar{q}$ denotes the average concentration in the adsorbent particle in [mol/(kg framework)], which forms a link between the fluid and solid phase mass balance equations.

- Advection is the transport of substances due to the bulk motion of the fluid in the axial direction and is the difference between the input and output amount per unit time.

$$\dot{N}_{\text{adv}} = u(t,z)A_R c(t,z) - u(t,z+\mathrm{d}z)A_R c(t,z+\mathrm{d}z) \tag{3.197}$$

$$= -A_R \frac{\partial u(t,z)c(t,z)}{\partial z} dz \tag{3.198}$$

$$= -\mathrm{d}V \frac{\partial u(t,z)c(t,z)}{\partial z} \tag{3.199}$$

$$= -\varepsilon_B \mathrm{d}V \frac{\partial v(t,z)c(t,z)}{\partial z} \tag{3.200}$$

where v is the interstitial velocity of the gas phase (related to the superficial gas velocity u by $v = u/\varepsilon_B$).

- Assuming that the axial dispersion can be described by Fick's first law [264], the difference between input and output caused by dispersion is as follows:

$$\dot{N}_{\text{disp}} = \mathcal{D}\varepsilon_B A_R \left[\frac{\partial c(t,z)}{\partial z}\right]_{z+\mathrm{d}z} - \mathcal{D}\varepsilon_B A_R \left[\frac{\partial c(t,z)}{\partial z}\right]_z \tag{3.201}$$

$$= \mathcal{D}\varepsilon_B A_R \frac{\partial c^2(t,z)}{\partial z^2} \mathrm{d}z \tag{3.202}$$

$$= \mathcal{D}\varepsilon_B \mathrm{d}V \frac{\partial c^2(t,z)}{\partial z^2} \tag{3.203}$$

where $\mathcal{D}$ is the axial dispersion coefficient.

Introducing Eqs. 3.194, 3.196, 3.200, and 3.203 into Eq. 3.193, dividing by the control volume $(\mathrm{d}V)$ and the bed void-fraction (ε_B), and applying

the balance equation to each component, we obtain the main governing equation for the fixed-bed model [62, 65].

$$\frac{\partial c_i(t,z)}{\partial t} = -\frac{\partial\left(v(t,z)\,c_i(t,z)\right)}{\partial z} + \mathcal{D}_i\frac{\partial^2 c_i(t,z)}{\partial z^2} - \frac{1-\varepsilon_B}{\varepsilon_B}\rho_P\frac{\partial \overline{q}_i(t,z)}{\partial t} \tag{3.204}$$

For the i-th component, the component mass balance includes an axial dispersion term, an advection flow term, accumulation in the fluid phase, and a source term caused by the adsorption process on the adsorbent particles. If the system is non-isothermal, an additional heat balance is required [62].

3.4.3. Axial dispersion

Models that consider the dispersion are referred to as dispersed-flow models, whereas models that neglect dispersion are termed plug-flow models. In practice, most experiments are set up such that the plug flow assumption is valid. During the fluid flow through the packed bed, axial mixing takes place, which reduces the efficiency of the separation process by broadening the mass transfer zone. Axial mixing is caused by molecular diffusion, turbulent mixing, flow splitting, and rejoining around particles, Taylor dispersion, channelling, and wall effects [62, 265]. These phenomena can be minimized by choosing proper bed and flow conditions. Axial dispersion does depend on the Reynolds number and therefore, changes along the column when appreciable amounts are adsorbed. In particular, the Peclet number describes axial dispersion. The effects of all mechanisms that contribute to axial mixing are lumped into a single effective axial dispersion coefficient $\mathcal{D}_z$. Generally, the contribution from molecular diffusion and Eddy diffusion or turbulence is considered in the formulation of the dispersion coefficient. Coupling theory is used to add the contribution of both terms as shown in the equation stated below [266].

$$\mathcal{D}_z = \gamma D_m + \frac{\lambda d_p v}{1 + C D_m/(d_p v)} \tag{3.205}$$

When intra-particle mass-transfer is neglected (if the product of Reynolds number and Schmidt number, $(Re \cdot Sc) \gg 1$, advection dominates diffusion), $\mathcal{D}_z$ can be estimated using Eq. 3.205.

Table 3.7. Correlations for estimation of axial dispersion coefficient, $\mathcal{D}_z$ in fixed bed adsorbers. D_m is the molecular diffusivity, Sc is the Schmidt number and Re is the Reynolds number. r_p and d_p are the particle radius and particle diameter respectively. v represents velocity and ε_B is the bed porosity.

Dispersion Model	Conditions	Authors
$\mathcal{D}_z = \frac{2r_p v \varepsilon_B}{0.2+0.011Re^{0.48}}$	$10^{-3} < Re < 10^3$	Chung and Wen [267]
$\mathcal{D}_z = 0.7D_m + \frac{5r_p v}{1+4.4D_m/(r_p v)}$	valid for liquids	De Ligny [266]
$\mathcal{D}_z = 0.7D_m + \frac{2r_p v \varepsilon_B}{0.18+0.008Re^{0.59}/(r_p v)}$	-	Rastegar and Gu [268]
$\mathcal{D}_z = 0.73D_m + \frac{0.5d_p v \varepsilon_B}{1+(9.7D_m)/(v\varepsilon_B d_p)}$	$0.008 < Re < 50, 0.0377 < d_p < 0.60$ cm	Edwards and Richardson [269]
$\mathcal{D}_z = \frac{20D_m}{\varepsilon_B} + \frac{0.5D_m ScRe}{\varepsilon_B}$	-	[61, 249]
$\mathcal{D}_z = (0.45 + 0.55\varepsilon_B)D_m + r_p v$	-	Ruthven [62]

Table 3.7 shows various correlations to calculate the axial dispersion coefficient, $\mathcal{D}_z$. $\mathcal{D}_z$ is a function of molecular diffusivity (D_m), particle radius (r_p), fluid phase velocity (v) and bed porosity (ε_B). The molecular diffusivity, D_m used in the expression for axial dispersion coefficient can be calculated either using Molecular Dynamics [50, 270, 271] or the Chapman-Enskog theory [272, 273]:

$$D_m = \frac{AT^{\frac{3}{2}}}{p\sigma_{12}^2\Omega}\left(\frac{1}{M_1} + \frac{1}{M_2}\right)^{0.5} \tag{3.206}$$

In Eq. 3.206, A is an empirical constant of magnitude $1.859{\cdot}10^{-7}$. T is the temperature in [K] and p is the partial pressure of the molecule in [atm]. M is the molar mass in [g/mol]. σ_{12} is the average collision diameter between molecules 1 and 2 in [Å] and Ω is the temperature-dependent collision integral. The expressions for σ_{12} and Ω [253] are as follows.

$$\sigma_{12} = \frac{1}{2}(\sigma_1 + \sigma_2) \tag{3.207}$$

$$\Omega = \left(44.54\left(k_B T/\epsilon_{i,j}\right)^{-4.909} + 1.911\left(k_B T/\epsilon_{i,j}\right)^{-1.575}\right)^{0.1} \tag{3.208}$$

$\sigma_{i,j}$ in Eq. 3.207, and $\epsilon_{i,j}$ in Eq. 3.208 are the effective Lennard-Jones parameters [274, 275] for a mixture of molecules i and j. $\epsilon_{i,j}$ is defined by $\sqrt{\epsilon_i \epsilon_j}$ and k_B is the Boltzmann constant. The expressions for the Lennard-Jones parameters for a molecule i can be estimated as follows.

$$\sigma_i = 0.841 V_c^{1/3} \tag{3.209}$$

$$\epsilon_i = 0.75 k_B T_c \tag{3.210}$$

In Eq. 3.209 and 3.210, V_c and T_c are the critical volume and critical temperature of the molecule i respectively [253]. V_c is in [Å^3] and T_c is in [K].

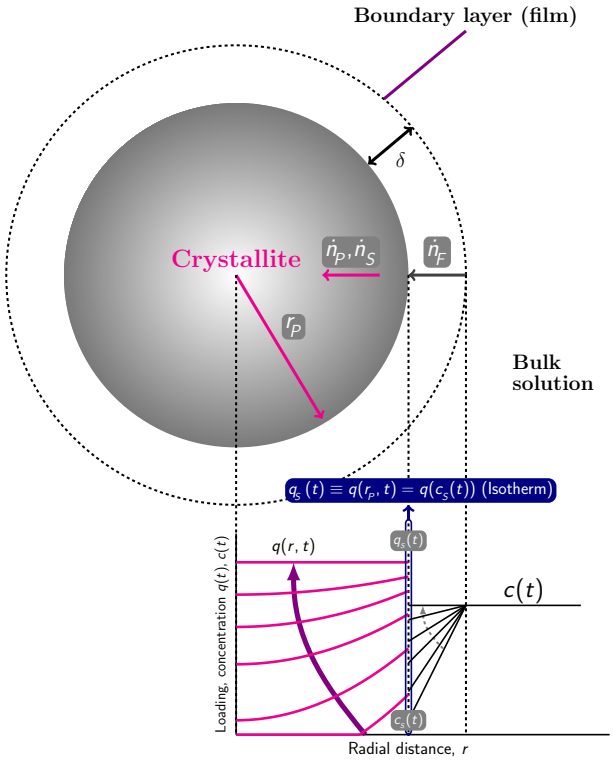


Figure 3.7. Loading and concentration profiles according to the pore diffusion model with external mass transfer resistance. The external mass transfer resistance is modelled using the film flux $\dot{n}_F(t) = K_F\left(c(t) - c_s(t)\right)$, which is the linear difference of the concentration $c(t)$ in the bulk solution, and the concentration $c_s(t)$ at the surface of the crystallite. The flux $\dot{n}_T = \dot{n}_p + \dot{n}_s$ is the sum of the pore diffusion and surface diffusion: $\dot{n}_T = D_p \partial c_p/\partial r + \rho_p D_s \partial q/\partial r$. The continuity of the materials fluxes requires $\dot{n}_F = \dot{n}_T$. The surface is considered to be in equilibrium and the adsorbent loading $q(r_p, t) = q(p(t))$ is calculated from the isotherm with functional form q, where the pressure p is related to the surface concentration c_s via the ideal gas law as $p = RTc_s$.

3.4.4. Mass transfer models

Inside nanoporous materials, molecules can diffuse into the framework via pore diffusion and surface diffusion [157]. Diffusion of the particles in the

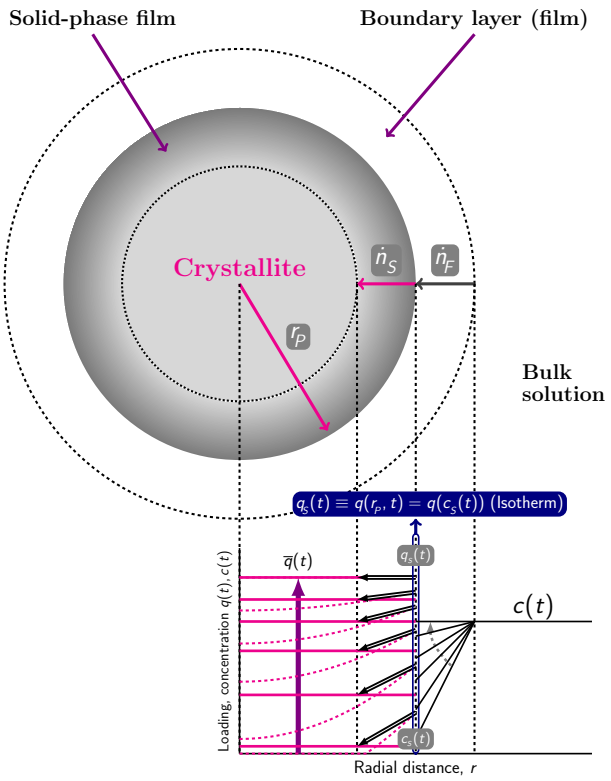


Figure 3.8. Loading and concentration profiles according to the Linear Driving Force (LDF) model with external mass transfer resistance. The LDF model assumes that there is a linear gradient within the solution-side film and within a comparable fictive solid film. The film flux $\dot{n}_F(t) = K_F\left(c(t) - c_s(t)\right)$ is the linear difference of the concentration $c(t)$ in the bulk solution, and the concentration $c_s(t)$ at the surface of the crystallite. The solid flux $\dot{n}_s$ is the linear difference between the equilibrium loading q_s at the surface and the average loading $\bar{q}$: $\dot{n}_s(t) = \rho_P k_s\left(q_s(t) - \bar{q}(t)\right)$, where ρ_P is the particle density, k_s is the intra-particle mass transfer coefficient. The continuity of the materials fluxes requires $\dot{n}_F = \dot{n}_s$. The surface is considered to be in equilibrium and the adsorbent loading $q(r_P, t) = q(p(t))$ is calculated from the isotherm with functional form q, where the pressure p is related to the surface concentration c_s via the ideal gas law as $p = RTc_s$.

pores occurs through two transfer processes depending on the pore size [62]: (1) molecular diffusion, which results from collisions between molecules, dominates in macro pores, (2) Knudsen diffusion occurs for smaller pore sizes due to collisions between molecules and the pore wall. In the surface diffusion mechanism, molecules hop between adsorption sites. It is therefore strongly dependent on concentration. For much more details on diffusion in nanoporous materials, see Ref. [276] and references therein.

The term $\partial\overline{q}(t,z)/\partial t$ in Eq. 3.204 describes adsorption and the overall rate of mass transfer from the fluid phase to the solid phase at time t and column distance z. The adsorbate molecules initially must cross the external film surrounding each adsorbent particle and then diffuse through and along the porous structure of the adsorbent. Many models are developed that differ in the rate-limiting mass transport step [246, 255, 277, 278]: the Linear Driving Force (LDF) model [130], the homogeneous surface diffusion model [279], the film-solid diffusion model with a constant diffusivity [280], the branched pore kinetic-model [281], the pore diffusion model [62, 282], the film-pore diffusion model [283], the diffusion flow-film particle diffusion model [284], the diffusion flow-local equilibrium model [285], the film pore and surface diffusion model [286–289], and the concentration-dependent surface diffusivity model [290] to name but a few. The most frequently applied approximate rate law to simplify the calculation is the LDF approximation [130]. In most cases, the LDF model is a good approximation to the exact but much more complicated surface and pore diffusion models [185].

Self-diffusion is a measure of the mobility of molecules due to their thermal motions, while collective diffusion is a dynamic process, involving many particles, due to their cooperative movements [291]. The collective diffusion is related to the transport diffusion coefficient by the thermodynamic factor [292–294], which can be obtained from the adsorption isotherms. Transport diffusion, taking the thermodynamic effects into account using a Fickian type of approach, can be incorporated in the equipment design equations. The differences between collective and transport diffusion show that there are different viewpoints possible to study diffusion [293, 295, 296]:

- In the Fick approach, the fluxes $\mathbf{N}$ are taken to be linearly dependent on the gradients of the loadings ∇c_i of all species with the "constants" of proportionality being the Fick diffusivity matrix.
- In the Onsager approach, the fluxes are postulated as linear functions of the chemical potential gradients ($\nabla\mu_i$), with the proportionality constants being the Onsager coefficients L_{ij}.
- The Maxwell-Stefan formulation balances diffusive and drag forces, while Fick and Onsager postulate phenomenological flux expressions. The Maxwell-Stefan diffusivity can be related to Fick diffusivity as follows [129, 292]

$$D_F = D_{MS}\Gamma \tag{3.211}$$

In Eq. 3.211, D_{MS} is the Maxwell-Stefan diffusivity and D_F is the Fick diffusivity. Γ is the thermodynamic correction factor, which is derived using chemical potential as the driving force [129]. This factor can be determined from adsorption isotherms after differentiation, as shown below [129].

$$\Gamma \equiv \frac{\partial \ln f}{\partial \ln q} \tag{3.212}$$

In Eq. 3.212, f is the fugacity and q is the adsorbed loading. To understand the underlying principles behind the diffusion process, it is more convenient to describe this phenomenon using Maxwell-Stefan diffusivity [129]. This approach clearly distinguishes between self-diffusivity, diffusivity of a species inside an adsorbent, and diffusivity describing interaction between different species [129].

The viewpoints differ in how to set up the flux-driving force relationship for transport diffusion in nanoporous materials under non-equilibrium conditions [293, 297–299].

In Onsager linear theory, the driving force of diffusion is the gradient of the chemical potential, μ, of the adsorbed particles. The diffusion flux $\mathbf{J}$ is [300]

$$\mathbf{J} = -L\nabla\left(\frac{\mu}{T}\right) = -\frac{L}{T}\frac{\partial \mu}{\partial q}\nabla q \tag{3.213}$$

where L is the Onsager coefficient and q is the adsorption loading. From the practical point of view, the flux $\mathbf{J}$ is usually related to the gradient diffusion coefficient D that depends on adsorption [300]

$$\mathbf{J} = -D\nabla q, \qquad D = D_0 \frac{1}{k_B T}\frac{\partial \mu}{\partial \ln q} \tag{3.214}$$

where $D_0 = kL/q$ does not depend on adsorption and has the meaning of the diffusivity at zero loading. The driving force for diffusion is more correctly expressed in terms of chemical potentials, but Fick's law provides a qualitatively and quantitatively correct representation of adsorption systems as long as the diffusivity is allowed to be a function of the adsorbate concentration [301]. However, for convenience, the diffusivity is often taken as the self-diffusivity and assumed to be a constant. In the limit of zero loading, the self-, collective-, and transport diffusivities become equal.

Film diffusion comprises the transport of the adsorbate from the bulk fluid to the external surface of the adsorbent particle. As long as the state of equilibrium is not reached, the concentration at the external adsorbent surface is always lower than in the bulk fluid due to the continuing adsorption process. As a consequence, a concentration gradient results that extends over a boundary layer of thickness δ (schematically shown in Figure 3.7). The difference between the concentration in the bulk solution, c, and the

concentration at the external surface, c_s, acts as a driving force for the mass transfer through the boundary layer. The mass transfer flux $\dot{n}_F$ for the film diffusion can be derived from Fick's law

$$\dot{n}_F = D_L \frac{dc}{d\delta} \tag{3.215}$$

leading to a film mass transfer equation [185]

$$\frac{d\bar{q}}{dt} = k_F \frac{a_{VR}}{\rho_B} (c - c_s) \tag{3.216}$$

where k_F is the film mass transfer coefficient in $\left[\mathrm{ms}^{-1}\right]$, c_s the concentration at the external surface, and the external surface related to the column (a_{VR}) is

$$a_{VR} = A_S / V_B \tag{3.217}$$

where A_S is the external adsorbent surface area and V_B is the fixed-bed volume. For spherical particles, $a_{VR} = (1 - \varepsilon_B)\, 3/r_p$. In most practical cases, the film diffusion influences only at the beginning of the adsorption process. At later stages of the process, the intra-particle diffusion becomes more important.

The local equilibrium model ignores the mass transfer resistance, and assumes that the mass transfer between solid adsorbent particles and external gas is instantaneous.

$$\frac{\partial q_i}{\partial t} = \frac{\partial q_i^*}{\partial t} \tag{3.218}$$

where q_i is the average amount of component i adsorbed, q_i^* is the amount adsorbed at equilibrium given by the selected adsorption isotherm model. Because of neglecting the mass transfer resistance, the model has limitations in practical applications.

In the Homogeneous Surface Diffusion Model (HSDM) [279, 302], the main diffusion mechanism is surface diffusion. The HSDM assumptions are the following [303]:

- mass balance around a spherical shell element.
- internal mass transfer is only governed by surface diffusion.
- external mass transfer is governed by a linear driving force.
- at the solid/fluid interface, there is continuity between external mass transfer and internal diffusion.
- at the solid/fluid interface, there is equilibrium between adsorbate concentration in the fluid and adsorbate load on the surface.

- the adsorbent particle is assumed to be spherical and homogeneous.

The adsorbent is assumed to be a uniform medium in which the solute adsorbs on the surface and diffuses inwards with an effective diffusivity in accordance with Fick's law. For surface diffusion, the mass transfer rate per unit of surface area, $\dot{n}_s$, is obtained using Fick's law as

$$\dot{n}_s = \rho_p D_s \frac{\partial q}{\partial r} \tag{3.219}$$

where D_s is the surface diffusion coefficient and r is the radial coordinate. Imposing a material balance and assuming D_s to be constant leads to the basic mathematical form of the HSDM for a spherical particle [185, 251, 304]:

$$\frac{\partial q}{\partial t} = \frac{D_s}{r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial q}{\partial r}\right) = D_s\left(\frac{\partial^2 q}{\partial r^2} + \frac{2}{r}\frac{\partial q}{\partial r}\right) \tag{3.220}$$

The homogeneous surface diffusion model is closely related to the homogeneous solid diffusion model and the film-solid diffusion model [305]. In the homogeneous solid diffusion model, all solute within the particle, whether it is in pore fluid or adsorbed by the particle skeleton, is lumped into a single quantity, q [304].

In the Pore Diffusion Model (PDM) [62], the adsorbate transport within the adsorbent particles takes place in the pore fluid, instead of on the surface. For pore diffusion, the mass transfer rate per unit of surface area, $\dot{n}_s$, is again obtained using Fick's law as

$$\dot{n}_p = D_p \frac{\partial c_p}{\partial r} \tag{3.221}$$

where D_p is the pore diffusion coefficient and r is the radial coordinate. In the surface diffusion model, the adsorbent is considered to be homogeneous, and the adsorption equilibrium is assumed to exist only at the outer surface of the adsorbent particle. In the case of pore diffusion, however, the adsorption equilibrium has to be considered at each point of the pore system. In general, it is assumed that there is a local equilibrium between the pore fluid concentration and the solid-phase concentration. Imposing a material balance leads to a more complicated description [185, 304]

$$\rho_p \frac{\partial q}{\partial t} + \varepsilon_p \frac{\partial c_p}{\partial t} = D_p\left(\frac{\partial^2 c_p}{\partial r^2} + \frac{2}{r}\frac{\partial c_p}{\partial r}\right) \tag{3.222}$$

In the Pore and Surface Diffusion Model (PSDM) [286], the adsorbate transport within the adsorbent particles takes place via surface and pore diffusion. The total flux is then given as the sum of the fluxes caused by

surface diffusion and by pore diffusion [185, 304]

$$\dot{n}_T = \dot{n}_S + \dot{n}_P = \rho_P D_S \frac{\partial q}{\partial r} + D_P \frac{\partial c_P}{\partial r} \tag{3.223}$$

$$\frac{\partial q}{\partial t} = D_S \left(\frac{\partial^2 q}{\partial r^2} + \frac{2}{r} \frac{\partial q}{\partial r} \right) + \frac{D_P}{\rho_P} \left(\frac{\partial^2 c_P}{\partial r^2} + \frac{2}{r} \frac{\partial c_P}{\partial r} \right) \tag{3.224}$$

The concentration profiles for film and pore/surface diffusion are schematically shown in Figure 3.7.

The intra-particle diffusion models discussed so far include partial derivatives with respect to time and radial coordinate, which causes an increased effort for the numerical solution. The Linear Driving Force (LDF) model states that the rate of adsorption is proportional to the amount of adsorbate still required to achieve equilibrium [304]. The equation for the flux is approximated by

$$\dot{n}_S = \rho_P k \left(q_S - \bar{q} \right) \tag{3.225}$$

where ρ_P is the particle density, k is the mass transfer coefficient, q_S is the adsorbent loading at the external surface of the adsorbent (which is a function of the bulk composition), and $\bar{q}$ is the mean loading in the particle. The following mass transfer equation can be derived

$$\frac{\mathrm{d}\bar{q}}{\mathrm{d}t} = k \left(q_S - \bar{q} \right) \tag{3.226}$$

The uptake rate of a species into adsorbent particles is proportional to the linear difference between the concentration of that species at the outer surface of the particle q_S (equilibrium adsorption amount) and its average concentration within the particle $\overline{q}$. r_P is the crystallite particle radius.

The LDF model can be derived from a model with a concentration profile within the particle. Assuming a quadratic concentration profile leads directly to the LDF form [306]. Glueckauf [130] derived a general expression to account for the mass transfer of the molecules from the fluid phase into the adsorbent [306]. This expression is valid for arbitrary variation of surface concentration of the adsorbates inside the solid particles [130]. Glueckauf also showed that for larger values of the product of mass transfer coefficient (k) and time (t), the expression is identical to the LDF model [130]. The volume-averaged adsorption amount $\overline{q}$ is given by

$$\overline{q}(t) = \frac{\int_0^{r_P} q(r,t)\, 4\pi r^2\, dr}{\frac{4}{3}\pi r_P^3} = \frac{3}{r_P^3} \int_0^{r_P} q(r,t)\, r^2\, dr \tag{3.227}$$

Taking the derivative, we get

$$\frac{\partial \overline{q}}{\partial t} = \frac{3}{r_P} \int_0^{r_P} \frac{\partial q}{\partial t} r^2 \mathrm{d}r \tag{3.228}$$

Substituting the diffusion equation for spherical geometry [279, 286]

$$\frac{\partial q(r,t)}{\partial t} = D\nabla^2 q(r,t) = \frac{D}{r^2}\frac{\partial}{\partial r}\left[r^2\frac{\partial q}{\partial t}\right] \tag{3.229}$$

into this equation we get

$$\frac{\partial \overline{q}}{\partial t} = \frac{3D}{r_p}\int_0^{r_p}\frac{\partial}{\partial r}\left(r^2\frac{\partial q}{\partial r}\right)\mathrm{d}r = \frac{3D}{r_p}\left[\frac{\partial q}{\partial r}\right]_{r=r_p} \tag{3.230}$$

In Eqs. 3.229 and 3.230, D is the Fickian diffusivity. Next, we consider a generalized concentration profile:

$$q(t,r) = A(t) + B(t)\,r^n \tag{3.231}$$

The linear term has been left out on account of spherical symmetry [306]. $A(t)$ and $B(t)$ can be solved from the two boundary conditions

$$\left(\frac{\partial q}{\partial r}\right)_{r=0} = 0 \tag{3.232}$$

$$q(r=r_p) = q(t,r_p) \tag{3.233}$$

A general solution for the concentration profiles is

$$q(t,r) = \begin{cases} q(t,r_p) - \frac{n+3}{n}\left[q(t,r_p)-\overline{q}\right] + \frac{n+3}{n}\frac{1}{r_p^n}\left[q(t,r_p)-\overline{q}\right]r^n & 0 \le r \le r_p \\ q(t,r_p) & r = r_p \end{cases} \tag{3.234}$$

Taking the derivative of the concentration profiles

$$\left[\frac{\partial q}{\partial r}\right]_{r=r_p} = \frac{n+3}{r_p}\left[q(t,r_p)-\overline{q}\right] \tag{3.235}$$

Finally, substituting the result for $n = 2$ into Eq. 3.230 leads to the LDF equation:

$$\frac{\partial \overline{q}}{\partial t} = \frac{15D}{r_p^2}\left(q_s - \overline{q}\right) \tag{3.236}$$

Therefore, the LDF model can be deduced from assuming a quadratic concentration profile [304, 306]. The precision of the LDF models can be improved by using higher-order models [307].

Although the LDF model was originally developed as a simplified version of the surface diffusion model [185], it can also be related to the pore diffusion model. As shown for the combined surface and pore diffusion mechanism, an effective surface diffusion coefficient can be defined as

$$D_S = D_S + \frac{D_p}{\rho_p}\frac{\partial c_p}{\partial q} \approx D_S + \frac{D_p}{\rho_p}\frac{c_0}{q_0} \tag{3.237}$$

In Eq. 3.237, c_0 and q_0 represent initial ($t = 0$) fluid phase concentration and adsorbed loading of the component. It is also possible to link the mass transfer coefficient (k_s) used in the LDF model with the effective surface diffusion coefficient (D_s) by [306]

$$k_s = \frac{15D^{\text{eff}}}{r_p^2} = \frac{15D_s}{r_p^2} + \frac{15D_p}{r_p^2}\frac{c_0}{\rho_p q_0} \tag{3.238}$$

Surface diffusion as well as pore diffusion can be approximated by the LDF model. In fact, the LDF model can be viewed as a diffusion model where the overall kinetic constant in Eq. 3.226 is a combination of all types of diffusivities, such as the film diffusivity and the intra-particle pore and surface diffusivities. The LDF model can also be extended to include coupled diffusion effects by using the Maxwell-Stefan diffusion formulation [308]. This model reduces the computational time significantly and is the most widely used model to describe the mass transfer kinetics. We have also incorporated this model in our code to account for the mass transfer kinetics. The LDF model includes competitive adsorption processes by combining it with a mixture adsorption prediction model that determines the relation between the concentrations and adsorbed amounts at the external surface.

3.4.5. Mixture loading prediction

In the multi-component breakthrough computation, the loading q_i of species i depends on the concentrations c_i of all N species

$$q_i = q(c_1, c_2, \ldots, c_{N_C}) \tag{3.239}$$

This would require mixture isotherms over all possible ranges of mole fraction. The equilibrium loadings q_i for the components in the mixture are therefore computed from mixture prediction models like IAST using single-component adsorption isotherms. The total average molar loadings of the mixture $(\overline{q}_T)$ within the crystallite is obtained by a summation over all N_C components.

$$\overline{q}_T(t, z) = \sum_{i=1}^{N_C} \overline{q}_i(t, z)\,. \tag{3.240}$$

There are three possibilities to couple IAST equations in adsorptive bed dynamics [194]:

1. The reduced grand potential can be treated as a dependent variable of time and space and added to the system of differential equations, describing the bed dynamics [309].

2. Computation of the adsorption equilibrium separately at each time step.

3. A B-spline approach to pre-compute the equilibrium states [229]. B-splines are also known as basis splines. These are constructed using polynomial functions joined at nodes. For a set of knots or nodes, λ_j, with $j = 0$ to $g + 1$, $g - k$ independent B-splines can be constructed. The summation of such independent splines leads to a spline function. Equilibrium loadings are calculated using IAST for the first time, and then the coefficients involved in the B-spline functions are fitted to the data obtained. Further, interpolations are performed to calculate the equilibrium loadings. This helps in avoiding the iterative calculations involved in IAST when called inside the breakthrough code [229].

This first option results in a strongly non-linear system of differential-algebraic equations which is difficult to solve, computationally expensive and time-consuming [194]. For more than two components, the B-spline approach results in additional multidimensional fitting issues, losing its advantages and limiting the applicability to the binary case [194]. The second option is computationally less expensive than the first one, as the IAST equations are solved separately from the bed dynamics. Unlike the B-spline approach, this method does not have limitations regarding the number of components present in the system. These advantages make the second method a popular approach for calculating equilibrium loading in fixed-bed adsorption problems [194]. Therefore, we have also adopted this approach in our code.

In this work, the IAST equations are solved using the FastIAS and the Nested-Loop bisection method. The latter has the advantage of calculating equilibrium loadings up to machine precision (ca. 10^{-15}). A typical working precision considered by current IAST methods is ca. 10^{-8} [194]. Moreover, we have also incorporated the explicit isotherms mentioned in Section 3.3.4. These isotherms can account for the size effects of the components on the adsorption behavior. Originally, these equations were developed for adsorbents having single sites. We have extended this model to tackle multi-site adsorption [93]. We have used a segregated approach where the adsorbed phase is in equilibrium with the gas phase at each site. This method calculates equilibrium loadings with similar accuracy to the segregated IAST approach by Swisher et al. [56]. The segregated approaches are useful in case of non-uniform adsorption surfaces, where IAST can be inadequate in calculating the correct equilibrium loadings [56].

The explicit isotherms do not require any iterations to calculate the loadings. For integration with the breakthrough model, this method makes the code about 3 times faster than using IAST [93]. Explicit isotherms are similar in speed to using FastIAS with caching, but are not iterative and hence are always converging.

3.4.6. Existing breakthrough software

Different softwares currently available for predicting breakthrough curves:

- Fixed-bed Adsorption Simulation Tool (FAST) [310]: It is a Windows-based software which is free of charge for academic or any other non-commercial purposes. It is licensed under a Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 Unported License. FAST predicts the breakthrough curves for fixed-bed type filters used in water treatment. The solid phase mass transfer can be modelled using a homogeneous surface diffusion model or the LDF model.

- Multi-Flow Inversion of Tracer (MFIT) [311]: It is also a Windows-based open-source software package for calculating breakthrough curves for tracer components present in groundwater. This software is distributed under the Creative Commons Attribution 4.0 License.

- ProSim Dynamic Adsorption Column Simulation (ProSim DAC) [312]: It is a commercial software used for adsorption and regeneration steps in temperature swing adsorption (TSA), pressure swing adsorption (PSA), vacuum and temperature swing adsorption, etc. It finds its application in hydrogen refining, isotopic separation, emission control of volatile organic compounds and solvent recovery.

- 3P Sim [313]: It is a commercial software. 3P Sim is a part of the dynamic sorption analyser instrument, mixSorb L. It can calculate mixture isotherms from pure component isotherms and also evaluate breakthrough curves based on mass and energy balances.

- Aspen Adsorption [314]: Aspen Adsorption is a commercial flowsheet simulator. It is used to design, analyse, simulate, and optimize gas and liquid adsorption systems. It is used for both PSA and TSA processes.

Various commercial numerical platforms have been applied for the modeling of the PSA process [315], such as gPROMS [316], MATLAB [317], and FLUENT [318]. Many research groups have their in-house breakthrough codes, which are not released as open-source codes. These include the work of Krishna and Baur [129], Rodriguez [61, 63], Worch [185], Chung et al. [55], etc. Instead of the LDF model, Krishna and Baur solved a partial differential equation along the crystal diffusion path. The diffusive flux used in the model is provided by Maxwell-Stefan equations [129]. Maxwell-Stefan diffusivity formulation involves a thermodynamic correction factor and binary diffusivities correlating different species. These parameters are evaluated locally along the crystal diffusion path [129]. Chung et al. [55] used breakthrough curve modeling to screen nanoporous materials for hexane and heptane isomer separations. It is an isothermal model with LDF approximation for the mass transfer kinetics of the isomers and Darcy's equation [319] for the pressure drop along the length of the column.

3.4.7. Breakthrough model implementation

Model

The ideal gas law reads

$$c_i = \frac{y_i p_T}{RT} = \frac{p_i}{RT} \tag{3.241}$$

where y_i is the mole fraction of component i in the gas phase, p_T is the total pressure, p_i is the partial pressure, T is the gas temperature, and R the universal gas constant. Assuming ideal gas behavior for the gas phase and isothermal conditions, the partial pressures in the gas phase at position z and time t are obtained by solving the material balance for each component $i = 1, \ldots, N_C$ [62, 65]. Here, the material balance shown in Eq. 3.204 is modified in terms of the partial pressures as shown below.

$$\frac{1}{RT}\frac{\partial p_i(t,z)}{\partial t} = -\frac{1}{RT}\frac{\partial \left(v(t,z)\, p_i(t,z)\right)}{\partial z} + \frac{1}{RT}\mathcal{D}_i \frac{\partial^2 p_i(t,z)}{\partial z^2} - \frac{1-\varepsilon_B}{\varepsilon_B}\rho_P \frac{\partial \overline{q}_i(t,z)}{\partial t} \tag{3.242}$$

Using the LDF approach yields

$$\frac{\partial \overline{q}_i(t,z)}{\partial t} = k_i\left(q_{\mathrm{eq},i} - \overline{q}_i\right) \tag{3.243}$$

Because of adsorption and dispersion, the interstitial velocity v in fixed-bed adsorption is not constant. To calculate the velocity profile, the material balance for the overall mixture is considered by summing Eq. 3.242 over all components:

$$\frac{\partial p_T}{\partial t} = -\frac{\partial \left(v p_T\right)}{\partial z} + \sum_{i=1}^{N_C}\left(\mathcal{D}_i \frac{\partial^2 p_i}{\partial z^2} - RT\frac{1-\varepsilon_B}{\varepsilon_B}\rho_P k_i\left(q_{\mathrm{eq},i} - \overline{q}_i\right)\right) \tag{3.244}$$

The total pressure along the column is invariant with time and has a constant gradient along the length of the column. Therefore, Eq. 3.244 is modified as shown below.

$$p_T \frac{\partial v}{\partial z} = \sum_{i=1}^{N_C}\left(\mathcal{D}_i \frac{\partial^2 p_i}{\partial z^2} - RT\frac{1-\varepsilon_B}{\varepsilon_B}\rho_P k_i\left(q_{\mathrm{eq},i} - \overline{q}_i\right)\right) - v\frac{\partial p_T}{\partial z} \tag{3.245}$$

The pressure gradient $\left(\partial p_T/\partial z = \text{constant}\right)$ can be calculated using the Ergun equation as shown below [320]. The Ergun equation considers both viscous and inertial effects. Apart from the Ergun equation, the Darcy [319] and Carman-Kozeny [321] equations could also be used, but these equations are valid only for viscous flows.

$$\frac{\Delta P}{L} = \frac{150\mu L}{d_p^2}\frac{\left(1-\varepsilon_B\right)^2}{\left(\varepsilon_B\right)^3}v + \frac{1.75 L \rho_f}{d_p}\frac{\left(1-\varepsilon_B\right)}{\left(\varepsilon_B\right)^3}v|v| \tag{3.246}$$

In Eq. 3.246, $\Delta P/L$ is the pressure gradient inside the column of length L. μ is the dynamic viscosity in [Pa s], d_p is the particle diameter in [m], ε_B is the bed porosity. ρ_f is the density of the fluid in [kg/m^3] and v is the superficial velocity in $\left[\mathrm{ms}^{-1}\right]$. Eqs. 3.242, 3.243 and 3.245 are solved together to obtain concentration profiles for all the components inside the column.

- Initial Conditions

  Initially, the column is filled with only carrier gas. The carrier gas does not adsorb and hence, its equilibrium loading is zero. At the start of the breakthrough simulation, the pressure of the carrier gas is equal to the total pressure inside the column $\left(p_{\text{carrier gas}} = p_{\text{t}}\right)$. Also, we have imposed the condition that $\partial p_T/\partial z =$ constant and the constant can be computed from Eq. 3.246. Therefore, at $t = 0$, we have,

$$\sum_i \mathcal{D}_i \frac{\partial^2 p_i(t,z)}{\partial z^2} = \mathcal{D}_{\text{carrier-gas}} \frac{\partial^2 p_{\text{carrier-gas}}(t,z)}{\partial z^2} \tag{3.247}$$

$$= \mathcal{D}_{\text{carrier-gas}} \frac{\partial^2 p_T(t,z)}{\partial z^2} \tag{3.248}$$

$$= 0 \tag{3.249}$$

Using equation 3.245, we obtain

$$\frac{\partial v}{\partial z} = -\frac{1}{p_T(z)}\left[v\frac{\partial p_T}{\partial z}\right] \tag{3.250}$$

$$\frac{\partial v}{v} = -\frac{1}{p_T(z)}\left[\frac{\partial p_T}{\partial z}\right]\partial z \tag{3.251}$$

Integrating both sides yields

$$\int_{v^{\text{in}}}^{v} \frac{\mathrm{d}v}{v} = -\left[\frac{\partial p_T}{\partial z}\right]\int_0^z \frac{1}{p_T(z)}\mathrm{d}z \tag{3.252}$$

$$\ln\left(\frac{v}{v^{\text{in}}}\right) = -\left[\frac{\partial p_T}{\partial z}\right]\int_0^z \frac{1}{\left(p_T + \frac{\partial p_T}{\partial z}z\right)}\mathrm{d}z \tag{3.253}$$

$$\ln\left(\frac{v}{v^{\text{in}}}\right) = \ln\left(\frac{p_T^{\text{in}}}{p_T^{\text{in}} + \frac{\partial p_T}{\partial z}z}\right) \tag{3.254}$$

$$v = v^{\text{in}}\frac{p_T^{\text{in}}}{p_T^{\text{in}} + \frac{\partial p_T}{\partial z}z} \tag{3.255}$$

$$p_T(t,z) = p_T^{\text{in}} + \frac{\partial p_T}{\partial z}z \tag{3.256}$$

Eq. 3.255 shows that at $t = 0$, the velocity is inversely proportional to the position inside the adsorber column. Initially, the pressure of the carrier gas is equal to the total pressure as shown below

$$p_{\text{carrier gas}}(t = 0, z) = p_T(t = 0, z) \tag{3.257}$$

and the partial pressures for all other components are zero inside the column except at the inlet. Also, the adsorption loadings for the components are zero at the beginning.

$$p_i(t = 0, z > 0) = 0 \tag{3.258}$$

$$\overline{q}_i(t = 0, z) = 0 \tag{3.259}$$

- Boundary Conditions

  At the inlet of the column, the partial pressures for each component are fixed. The velocity is also fixed at the inlet. At the outlet, the spatial gradients of the partial pressures are considered zero.

$$v(t, z = 0) = \frac{u^{\text{in}}}{\varepsilon_B} \tag{3.260}$$

Dirichlet boundary condition,

$$p_i(t, z = 0) = y_i^{\text{in}} \cdot p_T^{\text{in}} \tag{3.261}$$

Neumann boundary condition,

$$\frac{\partial p_i(t, z = L)}{\partial z} = 0 \tag{3.262}$$

Instead of using Dirichlet boundary conditions (Eq. 3.261) [322], the Danckwerts boundary condition ($-D_i \frac{\partial p_i(t,z=0)}{\partial z} + v(t, z = 0)p_i(t, z = 0) = v^{\text{in}} p_i^{\text{in}}$) [323] can also be used. If the values of the dispersion coefficients are not very high, then Dirichlet and Danckwerts boundary conditions should produce the same results.

Numerical approach

Fixed bed adsorption involves mass transport equations for the transport of the mixture in the fluid phase. Adsorption of the components from the fluid phase to the adsorbent is modelled using the Linear Driving Force (LDF) model as shown in Eq. 3.243. The LDF model states that the rate of adsorption is proportional to the amount of adsorbate still required to achieve equilibrium [304]. The mass transfer rate is a function of the equilibrium loading of the components, which is obtained either using IAST or explicit

isotherm equations that account for the effects of the size of the molecules [118]. The expression for the velocity is obtained using the equation for the total pressure as shown in Eq. 3.245. These equations form a system of differential algebraic equations involving partial differential equations (PDEs) and non-linear algebraic equations. Different methods can be used to solve this system of equations. Finite difference [137], finite volume [66] or finite element [324] methods are the common numerical approaches used to solve such equations. The method of lines [68] is a popular technique to solve such problems involving PDEs [315]. In this method, the PDEs are converted into ordinary differential equations (ODEs) and eventually, these equations can be solved using already available integration schemes for ODEs [68]. Another approach is to discretize both time and spatial domains simultaneously [315]. Nilchan et al. used the orthogonal collocation method to discretize the time domain and the finite element method for the spatial domain for optimizing pressure swing adsorption systems [325]. Finite Volume Method (FVM) [326] and Finite Element Method (FEM) [327] are also used for discretising the time and spatial domains. In this work, the method of lines will be used to solve the system of differential algebraic equations.

Method of lines

The method of lines is a technique to solve the partial differential equations in which all dimensions except one are discretized [68]. In this way, systems of PDEs can be converted into ODEs and solved using the integration schemes available for ODEs. The method of lines requires the system of PDEs to be a well-posed initial value problem. Elliptical partial differential equations can also be solved using this method, but certain special techniques need to be introduced, like the method of false transients [68, 328]. A common example encountered in many scientific disciplines is a system of partial differential equations with spatial and temporal derivatives. In the method of lines, the spatial derivatives are discretized using numerical discretization schemes like the finite difference method (FDM), the finite volume method (FVM) and the finite element method (FEM). The time derivatives are kept in their continuous form.

Spatial discretization

As mentioned before, spatial derivatives can be discretized using FDM, FVM, and FEM.

- Finite Volume Method (FVM) [66]: The Finite Volume Method is based on the conservation of numerical fluxes at each grid point and involves the integration over the finite volume or the control volume (grid). Many physical laws are conservation laws, especially in the field of fluid mechanics, heat and mass transfer, etc [66]. This makes the FVM a popular method in these disciplines. The values of the dependent variables are stored at the center of the grids or finite volumes.

- Finite Element Method (FEM) [324]: In this method, the region of concern is subdivided into geometrically simple finite-sized elements. Within these elements, the derivatives are approximated as simple functions such as linear or quadratic polynomials. The data for the dependent variables are stored on the nodes of the grids.

- Finite Difference Method (FDM) [137]: It is the most direct method. It is easy to implement and efficient to solve for regular geometries. For complex geometries, the above two methods are preferred. The derivatives are discretized in the form of finite differences, which are derived using a Taylor series expansion. Similar to FEM, FDM also stores the data of the dependent variables at the nodes of the grids.

  For its ease of implementation, FDM is used for the discretization of the spatial domain in this work. Discretization depends on the order of the derivatives. Some examples for first and second order derivatives are shown below in Table 3.8.

Discretization of the governing equations

- Mass Transport Equation: Before diving into the discretization of the entire governing equations, it is important to explore the possibilities for discretising each physical phenomenon individually. The mass transport equation consists of advective, dispersion and adsorption terms. The advective term $\left(\frac{\partial v p_i}{\partial z}\right)$ can be discretized either using the backward difference method or the central difference method [138]:

$$\frac{\partial (v p_i)}{\partial z} = -\frac{v(j)p_i(j) - v(j-1)p_i(j-1)}{\Delta z} \quad \text{Upwind scheme} \tag{3.263}$$

$$\frac{\partial (v p_i)}{\partial z} = -\frac{v(j+1)p_i(j+1) - v(j-1)p_i(j-1)}{2\Delta z} \quad \text{Central-difference scheme} \tag{3.264}$$

  Eqs. 3.263 and 3.264 show the two variations of discretization for the advection term. Index j in these equations represents grid points in the spatial direction. In this case, the backward differencing method is also known as the upwind scheme. Although it is first-order accurate, it is highly suitable for strong advective flows [66, 138]. The central differencing scheme is second-order accurate but leads to an unstable solution or creates non-physical oscillations in the solution [138]. The central differencing scheme does not identify the directionality of the flow [66]. It considers the influence of all the neighboring grid points at the current node. In an advection-dominated flow, the directionality becomes important. For example, in an unidirectional flow, the flux at the current grid point is strongly influenced by the preceding grid

Table 3.8. First and second order derivatives discretized using the Finite Difference Method (FDM) method [137].

Derivatives	Discretization	Scheme
$\left(\frac{\partial u}{\partial z}\right)_j$	$\frac{u(j+1)-u(j)}{\Delta z} + O(\Delta z)$	Forward difference
$\left(\frac{\partial u}{\partial z}\right)_j$	$\frac{u(j)-u(j-1)}{\Delta z} + O(\Delta z)$	Backward difference
$\left(\frac{\partial u}{\partial z}\right)_j$	$\frac{u(j+1)-u(j-1)}{2\Delta z} + O(\Delta z)^2$	Central difference
$\left(\frac{\partial^2 u}{\partial z^2}\right)_j$	$\frac{u(j+1)-2u(j)+u(j-1)}{(\Delta z)^2} + O(\Delta z)^2$	Central difference

point compared to the succeeding one. The upwind scheme takes into account the directionality of the flow [66]. Therefore, the upwind scheme is used for the advective terms in this work. We have used the 1st order upwind scheme. The spatial accuracy can be improved by using a 2nd order upwind scheme. However, the 1st order upwind scheme is easier to converge than the 2nd order scheme. Therefore, we have used the 1st order upwind scheme.

The mass transport equation (Eq. 3.242 in section 3.4.7) is discretized as follows.

$$\frac{\partial p_i}{\partial t} = -\frac{v(j)p_i(j) - v(j-1)p_i(j-1)}{\Delta z} + D_i\frac{p_i(j+1) - 2p_i(j) + p_i(j-1)}{(\Delta z)^2} - RT\left(\frac{1-\varepsilon_B}{\varepsilon_B}\right)\rho_p k_i\left(q_{\mathrm{eq},i}(j) - \overline{q}_i(j)\right) \tag{3.265}$$

- Linear Driving Force (LDF): The LDF method is used to calculate the amount of molecules adsorbed in the adsorbent. At each grid point, the rate of adsorption is a linear function of the equilibrium loading and the actual adsorbed loading at that grid point, which is shown below.

$$\frac{\partial \overline{q}_i}{\partial t} = k_i\left(q_{\mathrm{eq},i}(j) - \overline{q}_i(j)\right) \tag{3.266}$$

In Eq. 3.266, the component mass transfer coefficients (k_i) are considered to be constants.

- Velocity: To calculate the velocity at each grid point the gradient of velocity in Eq. 3.245 is discretized using the 1st order upwind scheme as shown below.

$$\left(\frac{\partial v}{\partial z}\right)_j = \frac{v(j) - v(j-1)}{\Delta z} \tag{3.267}$$

The discretized version of Eq. 3.245 equals.

$$v(j) = v(j-1) - \Delta z \frac{1}{p_T} \sum_{i=1}^{N_C} \left(-D_i \frac{p_i(j+1) - 2p_i(j) + p_i(j-1)}{(\Delta z)^2} \right) \tag{3.268}$$

$$- \Delta z \frac{1}{p_T} \sum_{i=1}^{N_C} \left(RT \left(\frac{1-\varepsilon_B}{\varepsilon_B} \right) \rho_P k_i \left(q_{\mathrm{eq},i}(j) - \overline{q}_i(j) \right) \right)$$

$$- \Delta z \frac{1}{p_T} \left(v(j-1) \frac{\partial p_T}{\partial z} \right) \tag{3.269}$$

Discretization of initial and boundary conditions

- Initial Conditions

Below are the discretized versions of the initial conditions (Eqs. 3.255-3.259) mentioned in section 3.4.7.

$$v(0,j) = v(0,0) \frac{p_T(0,0)}{p_T(0,0) + \frac{\partial p_T}{\partial z} j \cdot \Delta z} \tag{3.270}$$

$$p_T(0,j) = p_T(0,0) + \frac{\partial p_T}{\partial z} j \cdot \Delta z \tag{3.271}$$

$$p_{\text{carrier gas}}(0,j) = p_T(0,j) \tag{3.272}$$

$$p_i(0, j > 0) = 0 \tag{3.273}$$

$$\overline{q}_i(0,j) = 0 \tag{3.274}$$

- Boundary Conditions

Boundary conditions (Eqs. 3.260-3.262) mentioned in section 3.4.7 are discretized as follows.

$$v(t,0) = \frac{u^{\text{in}}}{\varepsilon_B} \tag{3.275}$$

$$p_i(t,0) = y_i^{\text{in}} \cdot p_T^{\text{in}} \tag{3.276}$$

$$p_i(t, N_{\text{grid}}) = p_i(t, N_{\text{grid}} - 1) \tag{3.277}$$

Time integration

Time integration in numerical analysis can be performed in several ways. Explicit and implicit are the two major methods to integrate the temporal derivatives. In explicit methods, the value for the variable at the current step is calculated as a function of the previous time steps [67]. In implicit methods, the present value for the variable depends on the current step. It is

easier to implement an explicit method, but it has constraints due to stability issues [137]. For very stiff differential equations, impractically smaller time steps are required to achieve a stable solution. Implicit methods are always stable and capable of waiving off the smaller time step requirement. However, very large steps are not recommended for the sake of accuracy. Implicit methods can be difficult to implement, especially when equations involve non-linearity. In the breakthrough curve model, the non-linearity occurs due to the calculation of the equilibrium loadings using IAST. IAST itself involves a system of non-linear equations which are solved iteratively (e.g., bisection or the Newton-Raphson method). Further, when IAST is incorporated into the breakthrough curve model, the system of differential equations needs to be solved iteratively, if the implicit method is used for discretization. To explain the integration schemes (implicit and explicit), let us consider the following differential equation

$$\frac{dY}{dt} = F(Y, t) \tag{3.278}$$

This equation can be solved using either implicit or explicit methods as shown below.

$$Y(t + \Delta t) = Y(t) + \Delta t\, F(Y(t)) \quad \text{Explicit method} \tag{3.279}$$

$$Y(t + \Delta t) = Y(t) + \Delta t\, F(Y(t + \Delta t)) \quad \text{Implicit method} \tag{3.280}$$

$$Y(t + \Delta t) = Y(t) + \frac{\Delta t}{2}\,[F(Y(t + \Delta t)) + F(Y(t))] \quad \text{Trapezoidal method [329]} \tag{3.281}$$

Various time integration schemes are available in both implicit and explicit categories. Some common schemes are Forward Euler, Backward Euler, and Crank-Nicolson [330]. Forward Euler is the simplest of all and is an explicit scheme (Eq. 3.279). Backward Euler is a fully implicit scheme (Eq. 3.280). The Crank-Nicolson method is a combination of both the explicit and implicit schemes. It is the finite difference version of the trapezoidal rule (Eq. 3.281). Another possibility is to split $F(Y, t)$ into stiff and non-stiff parts. The stiff part can be solved using implicit methods and the non-stiff part using explicit methods [331, 332] as shown below

$$Y(t + \Delta t) = Y(t) + \Delta t\,[F_1(Y(t + \Delta t)) + F_2(Y(t))] \quad \text{Implicit-Explicit method} \tag{3.282}$$

In Eq. 3.282, for the stiff function $(F_1(Y(t + \Delta t)))$, implicit methods can be used. For the non-stiff function $(F_2(Y(t)))$, explicit methods can be used.

There are several families of time integrators like the Linear Multistep methods [330] and the Runge-Kutta family [67]. In the multistep methods, the derivatives are treated as functions of several timesteps from the past and the future. Examples are Adams Bashfort [67] and Adams Moulton [67]

which are explicit schemes. The Backward Differencing Formula (BDF) is another example of linear multistep methods [67]. The BDF is implicit in nature and a popular method for solving stiff differential equations. Unlike multistep methods, the Runge-Kutta family involves only single time step methods. Rather, the derivatives are solved in several stages or trial steps to increase the order of accuracy. For the intermediate stages, the value of the independent variable at the midpoint of the interval is used. At each intermediate stage, half of the dependent variable calculated at the previous stage is used [330]. Both explicit and implicit Runge-Kutta methods are available in the literature [330, 333]. One of the commonly used Runge-Kutta methods is the 4^{th} order explicit method. Adaptive Runge-Kutta methods also exist where the error is calculated at each time step [334]. Based on this estimation, the time step is varied.

There is a relatively new class of Runge-Kutta Methods which are known as Strong Stability Preserving Runge-Kutta (SSP-RK) [69, 139, 140, 161]. Apart from the stability, these methods also enforce positivity, boundness and preserve monotonicity of the differential equations [140]. These are explicit methods and are stable under Courant–Friedrichs–Lewy (CFL) condition as shown below [140].

$$\frac{u\Delta t}{\Delta z} \leq 1 \tag{3.283}$$

In Eq. 3.283, u is a magnitude of dimension, length/time, Δt is the time step, and Δz is the spatial grid size. $\Delta z/u$ is also known as forward Euler time (Δt_{FE}). According to the CFL criterion, Δt must be less than or equal to Δt_{FE}. There are several variations of SSP-RK (e.g., 3^{rd} order, 5^{th}order) [69]. The SSP-RK(3,3) method is used in RUPTURA, which is shown below. It is a third-order method and involves three stages [140]:

$$Y(t+\Delta t)^{(1)} = Y(t) + \Delta t F(Y(t)) \tag{3.284}$$

$$Y(t+\Delta t)^{(2)} = \frac{3}{4}Y(t) + \frac{1}{4}Y(t+\Delta t)^{(1)} + \frac{1}{4}\Delta t F(Y(t+\Delta t)^{(1)}) \tag{3.285}$$

$$Y(t+\Delta t) = \frac{1}{3}Y(t) + \frac{2}{3}Y(t+\Delta t)^{(2)} + \frac{2}{3}\Delta t F(Y(t+\Delta t)^{(2)}) \tag{3.286}$$

In Eqs. 3.284, 3.285 and 3.286, $Y(t)$ is the dependent variable and $F(Y(t))$ is the derivative of $Y(t)$. In the SSP-RK method, the first step is always a forward Euler step. The derivatives of the dependent variable are solved in several stages, depending on the variation of the SSP-RK method. The 3rd order method is popular because it generates results almost as accurate as its higher order counterparts, and it involves fewer stages [69].

Breakthrough algorithm

Fig. 3.9 shows the schematic of the algorithm to calculate the breakthrough curves. The first step in the algorithm is to specify the initial conditions for

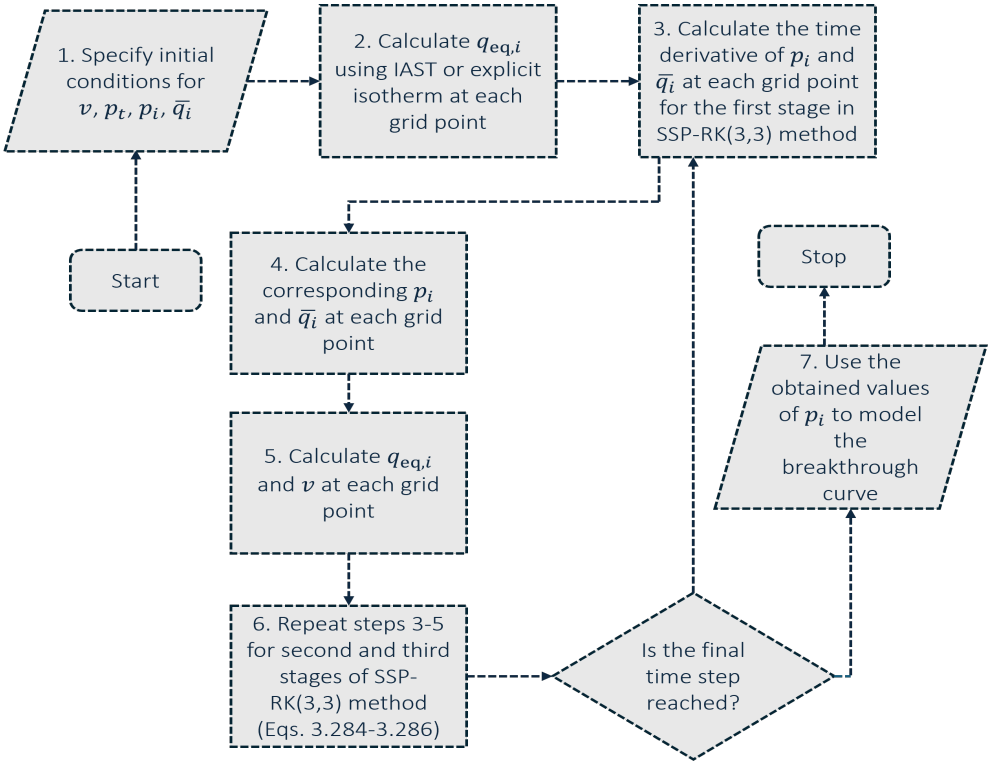


Figure 3.9. Algorithm to calculate the breakthrough curve. It is based on the method of lines (MOL) technique to solve the system of differential algebraic equations. Time integration is done using an explicit method called Strong Stability preserving Runge-Kutta (SSP-RK(3,3)) as shown in Eqs. 3.284-3.286. We first calculate the temporal derivatives of the partial pressures (p_i) and the adsorbed loadings ($\overline{q}_i$). Until the final time step is reached, we update the values of p_i, p_T, $\overline{q}_i$, $q_{eq,i}$ and v at each time step.

partial pressures, total pressure, adsorbed loadings, and velocity. Based on the initial conditions, equilibrium loadings at each grid point are calculated either using IAST or adsorbate size-dependent explicit isotherms [118]. Next, the variables (partial pressure, adsorbed loading, total pressure, velocity) are determined at each time step. The SSP-RK(3,3) method is used for the time integration. This method involves three stages. At each stage, the temporal derivatives of the partial pressures and the adsorption loadings are calculated, which are used to update the new values of partial pressure and adsorbed loadings. Based on the newly estimated partial pressures, the total pressure is calculated at each grid point, after which the equilibrium loading and velocity are calculated. The values obtained at the third stage are considered to be the final values at each time step. The process is

repeated until the final time step is reached.

3.4.8. Pulse breakthrough

RUPTURA also includes pulse-style breakthrough curve modeling. Instead of using a step input for the adsorbing components of the mixture, a pulse input is used. The pulse breakthrough curve allows us to identify whether the adsorbent is able to fractionate different components in a mixture [335]. The breakthrough curves are dome-shaped. The inlet concentrations for the adsorbing components become zero after the specified time period. On reaching the maximum concentration, the breakthrough curve gradually falls back to the zero concentration level. If there is a sufficient time gap between the peaks of the breakthrough curves for the components in a mixture, then it indicates that the adsorbent is suitable for the separation of the components.

All the equations and the conditions remain identical to the step breakthrough model, except for the inlet boundary conditions for the components in the mixture. The modified boundary conditions are shown below.

$$p_i(t, z=0) = \begin{cases} p_i^{\text{in}}, 0 \leq & t < t_{\text{pulse}} \\ 0, & t \geq t_{\text{pulse}} \end{cases} \tag{3.287}$$

$$p_{\text{carrier gas}}(t, z=0) = \begin{cases} p_{\text{carrier gas}}^{\text{in}}, 0 \leq & t < t_{\text{pulse}} \\ p_T^{\text{in}}, & t \geq t_{\text{pulse}} \end{cases} \tag{3.288}$$

In Eqs. 3.287 and 3.288, t_{pulse} is the time period for which the inlet concentrations of the adsorbing components are non-zero. Figs. 3.10a and 3.10b show examples of pulse-style breakthrough curves for the mixture of C_6 isomers: n-hexane (n-C_6), 3-methyl pentane (3-m-C_5), 2,2-dimethyl butane (22-m-C_5), and 2,3-dimethyl butane (2, 3-m-C_4) in BEA-type zeolite at 523K and 12 bar. The dome-shaped breakthrough curves of 3-m-C_5 and 2, 2-m-C_4 are very close to each other, which means the separation of these two isomers is difficult. It is easier to separate n-C_6 and (2, 2-m-C_5) for the operating conditions mentioned in Fig. 3.10. This is because the peaks of their breakthrough curves are separated by a considerable time gap. A longer column (Fig. 3.10b) enhances separation. However, the trade-off is that the retention time inside the bed increases for the adsorbing components. As a result, the amount of a specific component recovered will be lower for the longer column. This can be clearly seen in Fig. 3.10. The peaks of the breakthrough curves in Fig. 3.10b are shorter than those in Fig. 3.10a.

3.4.9. Validation

Jolimaitre et al. [336] considered the separation of mono and di-branched hydrocarbons in MFI-type (silicalite) zeolite. This study includes both experimental analysis and a breakthrough curve model. To validate our model, we

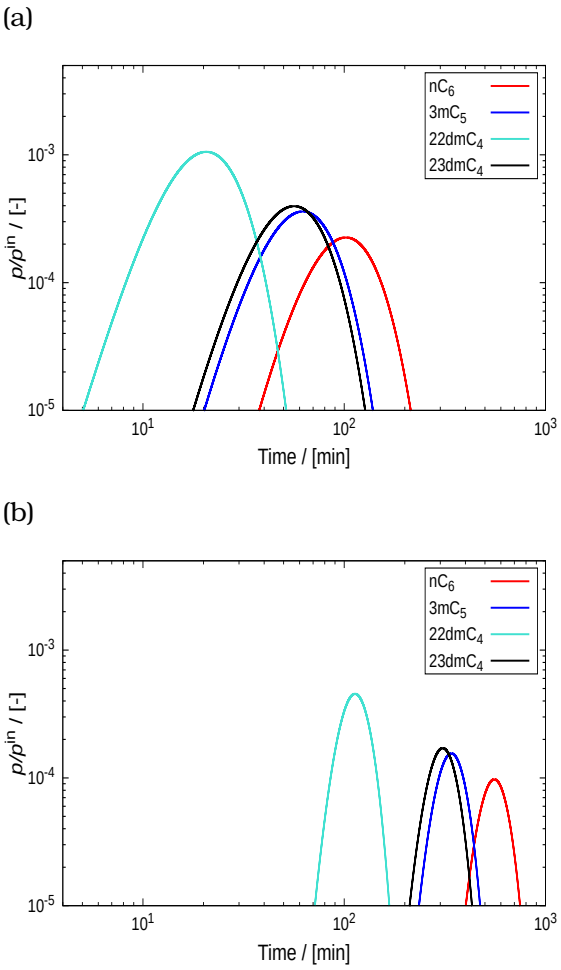


Figure 3.10. Examples of pulse-style breakthrough curve simulation for the mixture of C_6 isomers: n-hexane (n-C_6), 3-methyl pentane (3-m-C_5), 2,2-dimethyl butane (2, 2-m-C_4), and 2,3-dimethyl butane (23-m-C_4) in BEA-type zeolite at 523K and 12 bar. Isothermal and isobaric conditions are imposed on the fixed bed column of length (a) 0.1m and (b) 0.5m. The bed porosity (ε_B) is 0.2. Helium is used as the carrier gas. The inlet mole fractions for the C_6 isomers are considered to be 0.01975. The interstitial velocity (v) at the inlet is 0.019 ms^{-1}. 5s is the pulse time (t_{pulse}).

use the experimental breakthrough curve data generated by Jolimaitre et al. [336]. It involves the separation of 2-methyl butane (2-m-C_4) and 2-methyl pentane (2-m-C_5). The experiments are conducted at 473 K and 5 bar and Nitrogen (N_2) is used as the carrier gas.

Experimental breakthrough conditions for the separation of 2-m-C_4 and 2-m-C_5 in MFI-type zeolite are shown in Table 3.9 [336]. This includes the inlet concentrations of the components in the mixture $\left(c_i^{\text{in}}\right)$, their corresponding partial pressures $\left(p_i^{\text{in}}\right)$ and mole fraction $\left(y_i^{\text{in}}\right)$. The interstitial velocity (v) at the inlet is 0.0197 $\left[\text{ms}^{-1}\right]$ [336] and the bed porosity $\left(\varepsilon_B\right)$ is 0.4.

The axial dispersion coefficients ($\mathcal{D}$) are calculated using the following equation as shown in Table 3.7:

$$\mathcal{D} = \left(0.45 + 0.55\varepsilon_B\right) D_m + 0.5 d_p v \tag{3.289}$$

In Eq. 3.289, D_m is the molecular diffusion coefficient in $\left[\text{m}^2\text{s}^{-1}\right]$, d_p is the particle diameter in [m] and v is the velocity of the fluid phase in $\left[\text{ms}^{-1}\right]$.

The mass transfer coefficient (k) is calculated using Eq. 3.226. This equation shows that k is proportional to the effective diffusion coefficient, $\left(D^{\text{eff}}\right)$ which is obtained as follows [62]:

$$D^{\text{eff}} = \left(\frac{1}{D_m} + \frac{1}{D_K}\right)^{-1} \tag{3.290}$$

In Eq. 3.290, D_m is the molecular diffusion coefficient and D_K is the Knudsen diffusion coefficient. The units of both the diffusivities are in $\left[\text{m}^2\text{s}^{-1}\right]$. D_m is obtained using Eqs. 3.206. The expression for Knudsen diffusivity is shown below.

$$D_K = \frac{d_{\text{pore}}}{3}\sqrt{\frac{8RT}{\pi M}} \tag{3.291}$$

In Eq. 3.291, d_{pore} is the pore diameter, R is the universal gas constant, T is the temperature and M is the molar mass of the component. The values of dispersion coefficients ($\mathcal{D}$) and mass transfer coefficients (k) for 2-m-C_4 and 2-m-C_5 at 473 K and 5 bar total pressure are shown below.

Pure component adsorption isotherms are modelled using single-site Langmuir isotherms. The parameters $\left(q^{\text{sat}}\right)$ and (b) obtained from fitting the Langmuir isotherm to the adsorption loading data are shown below. The adsorption loadings for the pure components are obtained using grand-canonical Monte Carlo (GCMC) simulations [50].

Table 3.9. Experimental conditions for separation of 2-m-C_4, 2-m-C_5. It includes the inlet concentrations $\left(c_i^{\text{in}}\right)$, corresponding partial pressures $\left(p_i^{\text{in}}\right)$ and the mole fractions $\left(y_i^{\text{in}}\right)$.

	c_i^{in}/ [mol m^{-3}]	p_i^{in}/ [bar]	y_i^{in}
2-m-C_4	6.27	0.245	0.049
2-m-C_5	6.60	0.260	0.052

Table 3.10. Values of dispersion coefficient ($\mathcal{D}$) and mass transfer coefficient (k) for 2-m-C_4 and 2-m-C_5. The coefficients are calculated at 473 K and 5 bar for the composition shown in Table 3.9.

	$\mathcal{D}$/ [m^2 s^{-1}]	k/ [s^{-1}]
2-m-C_4	$4.79 \cdot 10^{-5}$	0.056
2-m-C_5	$3.91 \cdot 10^{-5}$	0.051

Table 3.11. Pure component adsorption isotherm parameters. Single-site Langmuir isotherms are used to fit the pure component adsorption loading data obtained using GCMC simulations at 473 K. q^{sat} is the saturation loading and b is the equilibrium constant.

	q^{sat}/ [mol/(kg framework)]	b/ [Pa^{-1}]
2-m-C_4	$0.717 \pm 4 \cdot 10^{-3}$	$2.835 \cdot 10^{-5} \pm 7 \cdot 10^{-7}$
2-m-C_5	$0.699 \pm 2 \cdot 10^{-3}$	$1.271 \cdot 10^{-4} \pm 3 \cdot 10^{-7}$

Fig. 3.11 shows the comparison between the breakthrough curves at the column exit for the mixture of 2-m-C_5 and 2-m-C_4 obtained from our simulation and the experiment performed by Jolimaitre et al. [336]. The breakthrough curves from the simulation and experiment follow similar trends. The discrepancies between the experimental and simulated fronts can occur because of several reasons such as errors in the pure component isotherm data, assumptions involved in IAST/SIAST model for equilibrium calculations, inaccurate estimation of the dispersion or/and diffusion coefficients, use of the LDF model, etc. Further, errors in conducting experiments and the quality of the adsorbent and fluid phase sample can also cause these deviations.

The adsorption of 2-m-C_4 is weaker compared to 2-m-C_5. Therefore, 2-m-C_4 is the first component to exit the column and shows a roll-up behavior. The roll-up refers to the hump that occurs in the breakthrough

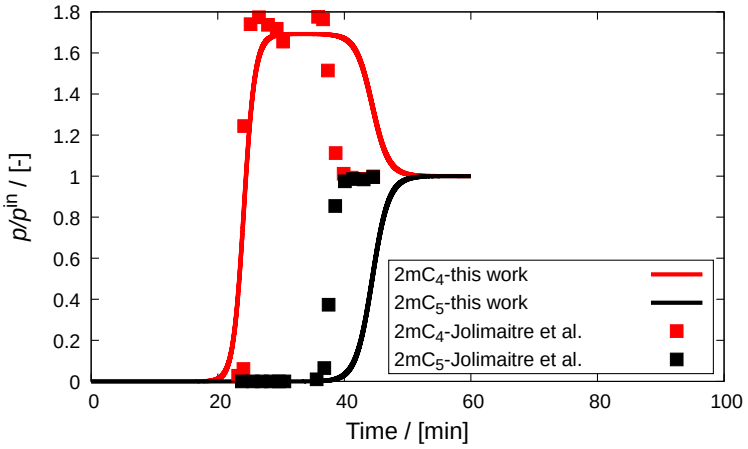


Figure 3.11. Validation of the breakthrough curve model with the experimental results of Jolimaitre et al. [336]. This figure shows the separation of 2-m-C_4 and 2-m-C_5 in MFI-type zeolite at 473 K and 5 bar. The colored squares represent the experimental data and the solid lines represent the results from our simulations.

curve of the weakly adsorbing component before the strongly adsorbing component starts to breakthrough. This behavior occurs because the gas phase concentration of the strongly adsorbing component is continuously decreasing due to its adsorption. As a result, the partial pressure of the weakly adsorbing component increases and at a certain time, it exceeds the inlet partial pressure (i.e., $p_i/p_i^{in} > 1$) [129]. The roll-up falls back to the inlet partial pressure, once the strongly adsorbing component starts to breakthrough, and eventually reaches equilibrium.

3.4.10. Case Study

In this section, we present a case study consisting of fitting of pure component loading data, mixture isotherm prediction, and breakthrough curve simulations. Here, we consider the case of adsorption of CO_2 and C_3H_8 mixture in MOR-type zeolite at 300 K from Ref. [93]. Fitting of pure component loading data and estimation of mixture isotherms, and simulations of breakthrough curves were performed using RUPTURA. The pure component loading data were obtained using the RASPA software [39, 156]. In this study, we also compare the influence of different mixture isotherm prediction models on the adsorption isotherms and the breakthrough curves.

Fig. 3.12a shows pure component isotherms (dual-site Langmuir) fitted to the adsorbed loadings obtained using Grand-Canonical Monte Carlo (GCMC) simulations. The unary isotherms show dual-site behavior because MOR-type zeolite has two types of adsorption sites (pockets) [56, 93]. Dual-

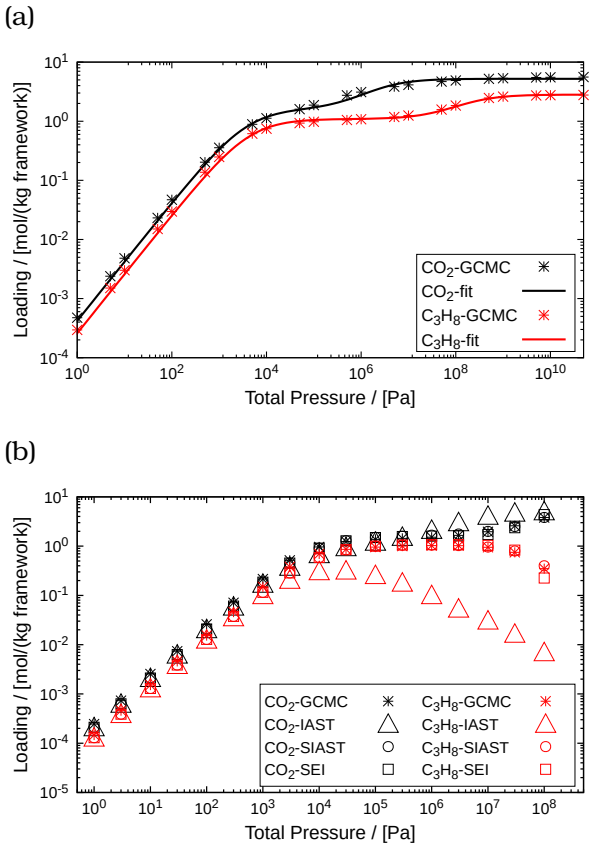


Figure 3.12. Adsorption isotherms for CO_2 and C_3H_8 in MOR-type zeolite at 300 K. (a) Unary isotherms (dual-site Langmuir) which are fitted to the pure component loading data obtained using RASPA, (b) comparison between mixture isotherms for equimolar composition of CO_2 and C_3H_8 computed using IAST, SIAST and SEI.

Table 3.12. Fitted parameters for the adsorption of pure CO_2 and C_3 in MOR-type zeolite at 300K. Adsorbed loadings obtained from the GCMC simulations for the pressure range $(10^0 - 5 \cdot 10^{10})$ Pa are fitted using the dual-site Langmuir isotherm.

	b_1 [1/Pa]	q_1^{sat} [mol/(kg framework)]	b_2 [1/Pa]	q_2^{sat} [mol/(kg framework)]
CO_2	$2.617 \cdot 10^{-4}$ $\pm 3 \cdot 10^{-5}$	1.60 $\pm 5 \cdot 10^{-2}$	$4.859 \cdot 10^{-7}$ $\pm 3 \cdot 10^{-8}$	3.62 $\pm 5 \cdot 10^{-2}$
C_3	$6.506 \cdot 10^{-9}$ $\pm 3 \cdot 10^{-10}$	1.73 $\pm 6 \cdot 10^{-3}$	$2.376 \cdot 10^{-4}$ $\pm 8 \cdot 10^{-6}$	1.09 $\pm 4 \cdot 10^{-3}$

site Langmuir isotherms have two pairs of fitted parameters (equilibrium constant and saturation loading) which are shown in Table 3.12. The MOR-type zeolite has a slightly higher loading of CO_2 compared to C_3H_8. This indicates that the affinity of the structure with CO_2 is higher than that of C_3H_8 and the preference for CO_2 is energetic in nature in this regime. At higher pressure, entropy (packing) effects come into play. Here we see that these effects are also in favor of CO_2 since CO_2 shows a higher adsorption and higher saturation loading in the pure component isotherms.

We can examine what happens when CO_2 and C_3H_8 compete with each other by doing explicit grand-canonical Monte Carlo mixture simulations. Indeed, as shown in Fig. 3.12b, at low loading, CO_2 is adsorbed significantly more than C_3H_8. At high pressures, we note that competition can increase the loading of one species at the expense of another. Thus, in a mixture, the loading of a component can go down with pressure, something that can not happen in single-component isotherms (when the framework is kept rigid). In most scenarios, IAST does an excellent job of predicting the mixture. However, the CO_2 and C_3H_8 in MOR-type zeolite system is an exception caused by segregation effects. In Fig. 3.12b, we compare the mixture isotherms computed using IAST, SIAST, and SEI models for an equimolar mixture of CO_2 and C_3H_8 inside MOR-type zeolite. For this system, IAST shows significant deviations from the GCMC data, whereas SIAST and SEI are in excellent agreement with these data. IAST under-predicts the adsorbed loadings for C_3H_8 and over-predicts for CO_2 adsorption. The reason for the failure in this system is that IAST assumes a uniform adsorbent, but MOR-type zeolite has two distinct types of adsorption sites with CO_2 and C_3H_8 having very different affinities for the two types of sites [56]. Consequently, the competition due to adsorption will not be identical in both of these sites. IAST fails to capture this effect and leads to incorrect predictions of adsorbed loadings.

The segregation effect has a huge influence on the behavior of the breakthrough curves. Fig. 3.13 shows the breakthrough curves for CO_2, and C_3H_8 inside MOR-type zeolite comparing the IAST, SIAST and SEI mixture

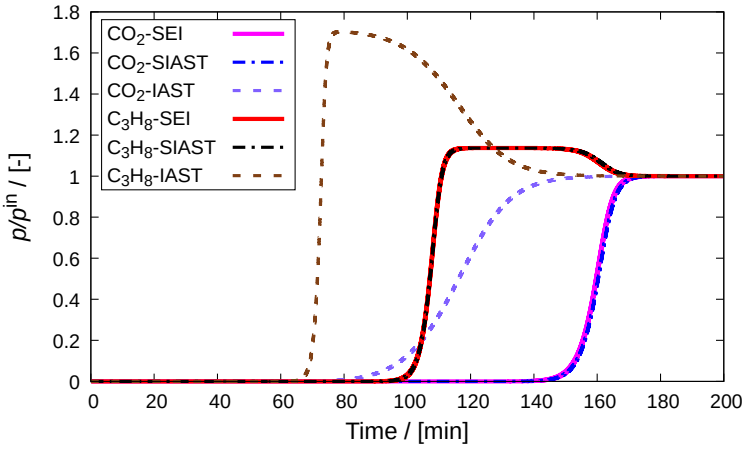


Figure 3.13. Comparison between breakthrough curves obtained on implementing IAST, SIAST, and SEI to the breakthrough curve model. A mixture of CO_2 and C_3H_8 in MOR-type zeolite at 300 K and 10^5Pa is considered. Each of these components constitutes 10% of the gas phase. Helium is used as the carrier gas. The adsorption column is operated in isothermal and isobaric conditions.

predictions. The simulations parameters are: column length $(L) = 0.3$ m, particle density (ρ_p) = 1711.06 kg/m^3, bed void fraction (ε_B) = 0.4, inlet velocity (v^{in})= 0.1 m/s, total pressure (p_T) = 1 bar, and the mass transfer coefficients (k) for CO_2, and C_3H_8 in MOR-type zeolite are taken as 0.06 s^{-1}. The difference in breakthrough behavior between IAST and SIAST is striking. Note that the prediction is based on the same pure component isotherm models, but that the IAST and SIAST differ in the adsorption physics. This example serves as an example that mixture prediction should be validated by explicit grand-canonical Monte Carlo simulations to verify the correct choice of the mixture prediction model. Note that SIAST and SEI lead to almost identical breakthrough curves. SEI is the non-iterative version, and hence has advantages in terms of speed and stability.

RUPTURA provides various options (IAST, SIAST, EI, and SEI) to estimate the mixture equilibrium loadings. One has to know a priori whether the adsorbent is heterogeneous and the adsorbing components prefer certain types of adsorption sites over the others. Knowledge of such information is essential in choosing the right mixture isotherm prediction model. If the adsorbent has distinct adsorption sites, then one should use SIAST [56], or SEI [93]. For uniform adsorbents, one can use IAST or EI. Also, it is important to note that for adsorbents with a single type of adsorption site, SIAST and IAST are identical. Similarly, SEI and EI are identical for uniform adsorbents. One should also be aware that EI and SEI can only be used if

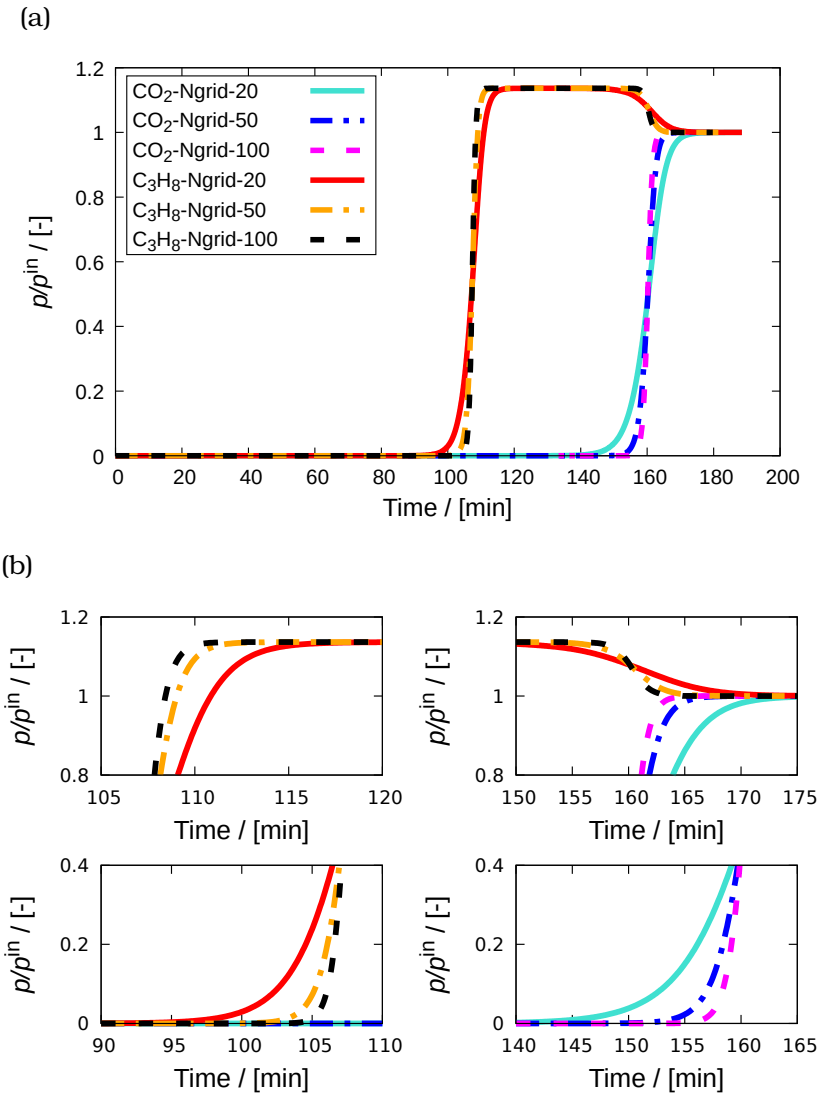


Figure 3.14. Grid independence test on simulation of breakthrough curves for a mixture of CO_2 and C_3H_8 in MOR-type zeolite at 300 K and 10^5 Pa is performed. (a) Breakthrough curves simulated at different spatial grid points ($N_{grid} = 20,\ 50,$ and 100), (b) certain zoomed-in parts of the curves showing discrepancies due to the use of a different number of grid points. Each of these components constitutes 10% of the gas phase. Helium is used as the carrier gas. The adsorption column is operated in isothermal and isobaric conditions. With an increase in the number of grid points, the breakthrough curves slightly change. The variation becomes smaller between $N_{grid} = 50$ and 100.

the pure component isotherms are Langmuir type.

Another important factor to consider is the effect of the number of grid points. A grid independence test is vital to determine the adequate number of grid points necessary for the breakthrough simulation. Here, we present a grid independence test for this case study. Three simulations are performed for this mixture with a different number of spatial grid points ($N_{\text{grid}} = 20,\ 50,$ and 100). In Fig. 3.14, we can observe that the breakthrough curves generated using $N_{\text{grid}} = 20$ are slightly different from those using $N_{\text{grid}} = 50$. The differences in the breakthrough curves become smaller for $N_{\text{grid}} = 50,$ and 100. This indicates that the breakthrough curves undergo changes with increasing N_{grid} to a certain limit. Beyond this point, these curves are invariant to an increase in the number of grid points.

3.5. Isotherm Fitting

3.5.1. Introduction

Curve fitting methods play an important role in many fields that involve system modeling and prediction techniques. Fitting is frequently used when it is necessary to match a complex model with (experimental) data [337, 338]. The determination of a suitable set of parameters for an analytical model (the parametric function) can reveal insights into the underlying physical phenomena. The aim of this work is to accurately fit the pure component isotherm models (as described in Sec. 3.2) to isotherm data obtained from experiments or grand-canonical Monte Carlo simulations. The most common procedure in the literature is the use of analytical functions for model representation. It is also possible to interpolate data in a model-agnostic manner using e.g. cubic splines [339]. Simon, Smit, and Haranczyk provided an interesting discussion on this topic [54]. Both simulation and experimental data have inherent noise, and when using numerical quadrature, there is always the danger of missing some physical information in an analytical model owing to over-fitting or over-interpreting the physical behavior that is actually caused by errors in the data. Recently, [340] designed a method to predict the single-component adsorption of various small adsorbates (Xe, Kr, CH_4, CH_6, and N_2) using machine learning techniques ("multipurpose" multilayer perceptron (MLP)). Using this method, these authors were able to predict thousands of mixing isotherms using a hybrid implementation of MLP-IAST. In RUPTURA, we developed our own implementation of a fitting tool for the isotherm models listed in Table 3.1.

3.5.2. Fitting methods

Linear fitting methods

A widely used technique to fit models to isotherm data is to carry out linear transformations on the isotherm data and then perform linear regression. Some isotherm models, such as Langmuir and Freundlich, can be estimated

from linear regression, since these models can be linearised. Four linear transformations can be applied to the Langmuir isotherm ($q/q^{sat} = (bp/1+bp)$):

1. Double Reciprocal or Lineweaver-Burk transformation [341] ($1/q$ vs. $1/p$ plot, and $1/q = 1/q^{sat} + 1/q^{sat}bp$ transformation): It is strongly biased towards fitting the data in the low concentration range.

2. Reciprocal or Langmuir transformation [342] (p/q vs. p plot, $p/q = 1/bq^{sat} + 1/bp$) tends to amplify the deviations from the fitted equation, highlighting outliers.

3. Eadie-Hofstee transformation [343, 344] (q vs. q/p plot, $q = q^{sat} - q/bp$ transformation). It has some bias toward fitting the data in the low concentration range.

4. Distribution Coefficient or Scatchard transformation [345] (q/p vs. p plot, $q/p = bq^{sat} - bq$). It is biased toward fitting the data in the high concentration range.

Because the error distribution is modified and the weight of each point in the fit is transformation dependent, the result of the fit will depend on the choice of the linearization method [346]. For this reason, the best fit is not the one with the highest correlation coefficient, but the one that predicts an error distribution that corresponds to the original data. Single Langmuir behavior is rare, and in most cases, multiple-site isotherm models are needed. A more general approach is to use non-linear regression [338].

Non-linear Least-Squares (NLLS) fitting methods

In non-linear regression problems, the aim is to find a parametric function

$$y = f\left(x; \{a_k\}_{k=1}^{M}\right) \tag{3.292}$$

which best fits a certain reference data set of points $(x_i, y_i)_{i=1}^{N}$ by minimising an error measure in the fitting. A widely used approach is to use the sum of squared residuals. The problem is thus mathematically defined as follows:

$$\text{WSSE} = \sum_{i=1}^{N} r_i\left(x; \{a_k\}_{k=1}^{M}\right) = \sum_{i=1}^{N} w_i\left(y_i - f(x_i; \{a_k\}_{k=1}^{M})\right)^2 \tag{3.293}$$

$$\phi = \arg_{\{a_k\}_{k=1}^{M}} \min_{x \in \chi} (\text{WSSE}) \tag{3.294}$$

where WSSE stands for Weighted Sum of the Squared Error (or cost function), ϕ is the value in the minimization process, r_i are the residual functions, and w_i are the weights associated with each (x_i, y_i)–point. When the parametric function f is unknown, we have to choose a suitable model with a convenient set of M parameters (a_k), and the error ϕ is minimized in an iterative way.

Gauss-Newton algorithm (GNA) [347]: In the Gauss-Newton method, the sum of squared errors is reduced by assuming that the least squares function is locally quadratic, and the minimum of the quadratic is obtained. The method is based on Newton's method to follow the curvature to the minimum but does not use the Hessian of the residual function (which is computationally expensive), but the Jacobian of the residual functions, $r(\{a_k\}_{k=1}^{M})$.

Steepest descent algorithm (SDA) [347]: In this method, also known as Gradient Descent, the sum of the squared errors is reduced by updating the parameters in the steepest-descent direction. This direction is calculated using the gradient, $-\nabla r(\{a_k\}_{k=1}^{M})$, and is therefore computationally cheap.

Marquardt–Levenberg algorithm (MLA) [348–352]: Developed by Levenberg, Girard, Wynne, Morrison, Marquardt. This method is actually a combination of two other NLLS fitting methods (GNA and SDA): It acts more like an SDA method when the parameters are far from their optimal value, and acts more like the GNA method when the parameters are close to their optimal values. This is the method used by the Gnuplot program, among others. An updated version of this method, which solves the problem of stacked solutions in local minima, is the Trust Region Reflective method [353].

There are other widely used gradient-, Jacobian-, or Hessian-based algorithms to follow a path to find the minimum (whose computational efficiencies depend on the nature of the problem to be optimised): The Conjugate Gradient (CG) method [354], Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm [355–358], and updates (e.g., Limited-memory BFGS [359]), the rational function optimization algorithm (RFO), etc. In general, NLLS fitting is not guaranteed to converge to the global optimum from scratch (the solution with the smallest sum of squared residuals, SSR), and often gets stuck in a local minimum.

Heuristic fitting methods

Heuristic Search methods (Non-Math Optimization algorithms) are not based on finding optimal analytic paths, based on gradients, Jacobians or Hessians, but explore the search space stochastically, keeping track of the 'best' points along the way. Two example approaches are discussed: the Nelder-Mead algorithm and the genetic algorithm.

The Nelder–Mead algorithm (NMA) [230, 360] is also called the downhill simplex method and was first introduced by Spendley, Hext, and Himsworth [360]. It is a heuristic search method based on the construction of a hyper-tetrahedron in M dimensions (a simplex) and how it adapts itself iteratively (by minimising the WSSE value) to the local landscape of the M-dimensional surface for the $\{a_k\}_{k=1}^{M}$ parameters. Each vertex $\mathbf{a}_{j=1}^{M+1}$ of this simplex is a vector with as many dimensions as there are parameters of the function f, i.e., $\mathbf{a}_j = (a_{1,j}, a_{2,j}, \ldots, a_{M,j})$. This is a local method and does not require derivatives.

The Genetic Algorithm (GA) [361, 362] is a type of parallel heuristic search method. This method was described for the first time by De Jong [361] and Mahajan, Kumar, and Porwal [362]. The GA belongs to Evolutionary Algorithms (EA), a family of algorithms for global optimization inspired by biological evolution, and uses biological operators inspired by biological evolution, such as reproduction, mutation, recombination, and selection. It is based on the generation of a population of individuals (each individual is a candidate for the optimization solution) that are ordered according to their fitness. During the process, the population of solutions remains constant, and the worst solutions are replaced (by ranking their goodness–of–fit) by other solutions that are either a mutation or a recombination of better solutions. Some EA are as follows: Genetic Programming [363, 364], Evolution Strategies [365], or Evolutionary Programming [366]. Other nature-inspired algorithms (but not part of the EA family) are: Particle Swarm Optimization (PSO) [367, 368], Ant Colony Optimization (ACO) [369], Cuckoo Search (CS) [370], Grey Wolf Optimization (GWO) and updates [371, 372], or Elephant Herd Optimization (EHO) [373, 374].

GA is the most popular type of EA. In particular, GA represents each individual of the GA population as a string of numbers, usually a binary string called "genotype". Modifications in the genotype imply changes in the fitness. For some problems, it is convenient to use non-binary representations of the genotype.

3.5.3. Evaluating isotherm goodness–of–fit

In RUPTURA, we have used two standard measures for evaluating the goodness-of-fit: the Residual Root Mean Square Error (RMSE), and the correlation coefficient (r). We defined the RMSE as:

$$\text{RMSE} = \left(\frac{\text{WSSE}}{N-M}\right)^{1/2} \tag{3.295}$$

where WSSE is the Weighted Sum of the Squared Error (Eq. 3.293), and N and M are the number of data points and the number of parameters of the model, respectively. The correlation coefficient r (usually r^2) is defined as:

$$r = \frac{\sum_{i=1}^{N}(w_i y_i - \bar{y})(w_i f(x_i) - \bar{f})}{\sqrt{\sum_{i=1}^{N}(w_i y_i - \bar{y})^2 \sum_{i=1}^{N}(w_i f(x_i) - \bar{f})^2}} \tag{3.296}$$

For each optimization iteration, RUPTURA monitors both values.

3.5.4. Software for isotherm fitting

Here we list some software proposed in the literature for isotherm fitting, as well as other general fitting software used for this purpose:

1. ISOTHERM and FIT programs [346, 375, 376]:
   This program is developed by Kinniburgh [375]. To the best of our knowledge, ISOTHERM appears in the literature as one of the first publicly available programs for isotherm fitting. Unfortunately, we have not been able to locate the code to test it.

2. ISOFIT [377]:
   ISOFIT was developed by Matott and Rabideau [377]. It is a program (executables available for Linux and Windows) that utilises a hybrid optimization procedure combining Particle Swarm Optimization with Levenberg–Marquardt non-linear regression. Many tools are available to check the quality of the fit. ISOFIT includes a wide collection of isotherm models (BET, Freundlich, Freundlich with Linear Partitioning, Langmuir-Freundlich, Langmuir, Langmuir with Linear Partitioning, Linear, Polanyi, Polanyi with Linear Partitioning, and Toth isotherms). However, it does not support fitting multisite isotherm models.

3. GAIAST [233]:
   GAIAST is developed by Balestra *et al.* [233]. Written in Fortran code (standard 2003), it has a module for isotherm fitting. The fitting module is coded using a constrained hybrid global optimization procedure (NMA and GA). GA uses a single-precision floating-point format (binary32) to code the genotype. It presents a wide variety of isotherm models, from one to three adsorption sites, as well as some isobar models.

4. IAST++ [234]:
   IAST++ is developed by Lee, Lee, and Kim [234]. It is a Windows user-friendly executable program with a graphical user interface. It contains a module for suggesting isotherm models (among Langmuir, Langmuir-Freundlich, Langmuir Dual Site, Langmuir-Freundlich Dual Site, BET, Quadratic, and Henry) based on the isotherm data entered, as well as visualization of the fitting procedure. If the module does not find an acceptable isotherm model (based on goodness-of-fit), it suggests a quadratic interpolation. In practice, this module requires a fine-tuned selection of ranges in the generation of initial parameter values in order not to get stuck in local minima. This is particularly important for the Langmuir dual-site and Langmuir-Freundlich dual-site models.

5. Non-dedicated software:
   There is a large variety of open-source tools, codes, and libraries for general purposes with which non-linear fits can be performed using different techniques. To mention a few of these: the SciPy Python Library [378] or the GNU Scientific Library (GSL) [379], or widely-used programs like Gnuplot [380], Octave [381]/ MATLAB [317], Origin [382] for NLLS methods, and the PIKAIA software [383] for GA methods.

3.5.5. Implementation of the GA-NMA hybrid method

In RUPTURA, we have implemented a constrained GA-NMA hybrid optimization procedure that combines the GA (for an initial global search for the minimum) with the NMA method (to refine the found best solutions). We employ the hybrid global optimization method that was also used in the GAIAST code, but we have made significant improvements in speed and efficiency. RUPTURA uses hashing tables to map the genes, while GAIAST works with string comparisons.

Genetic Algorithm stage

We have used a concatenation of the 64-bit binary representation (double-precision floating-point format or just binary64) of the M parameters of the Eq. 3.292 to codify the genotype. There is a one-to-one relationship between the genotype and the sequence of parameters $\{a_j\}_{j=1}^{M}$ that is called the "phenotype". The parameter values of the isotherm models are constrained to a range of values to prevent non-physical isotherm parameters (e.g., negative values for the saturation loading are not allowed). These constraints are hard-coded for each isotherm model.

The initial population is based on randomly generated isotherms from a range of initial values for each specified parameter. For example, the saturation values, q^{sat}, are generated between 0.0 to 10.0 [mol/(kg framework)], the b parameter between 1×10^{-10} to $1 \times 10^{+10}$ [Pa^{-1}], and the heterogeneity parameter, ν, anywhere between 0.1 to 2.1 [-]. If the isotherm parameters specified in the input files are non-zero, a refitting process is performed starting from these values.

In each optimization step, k, we sort the population of individuals (called "citizens") according to their fitness and apply five biological operators on their genotype: elitism, crossover, mutation, replacement, and nuclear disaster. These GA operators are shown schematically in Fig. 3.15.

1. Elitism: We allow the best individuals ($\sim 5\%$) of the current generation, k, to pass on unchanged to the next generation. In this way, we ensure that the best individual does not lose fitness in the optimization process.

2. Replacement: The worst ($\sim 5\%$) individuals are replaced by new individuals.

3. Crossover: We mate the population of the current generation (parents) to generate the next generation of individuals (children). We have implemented one-point, two-point, and uniform crossover, but using only the first one is sufficient to converge to a high-quality result. We perform a controlled mating between parents. Two types of crossover were carried out: between elitists (the group that ensures the goodness-of-fit) and between elitists and non-elitists (which provides the global character of the optimization method).

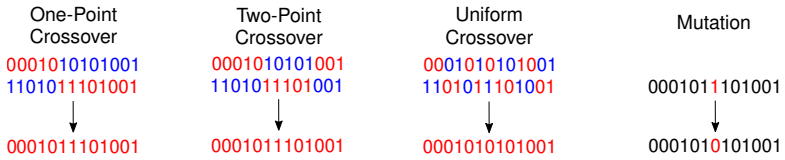


Figure 3.15. Schematic representation of GA operators: crossover and mutation. The crossover operator is analogous to the crossover that happens during sexual reproduction in biology, and it combines the genotype of two parents to generate new offspring. We implemented three types of crossover operators: with one and two points, and the uniform crossover. The mutation operator is activated before the children are added to the next generation.

4. Mutation: Some of the children, depending on the mutation rate value, do not pass unaltered to the next generation: We apply small changes in their genotype (see Fig. 3.15).

5. Nuclear disaster: Only elitists are passed on to the next generation. The rest of the individuals are replaced by new individuals. This operator operates at a very low rate and is designed to escape from local minima. When this operator is activated, we break the cycle of operators and jump directly to the next generation, $j+1$.

At the end of the operator cycle, we update the children to become parents and move on to the next generation, $j+1$. During the optimization process, RUPTURA prints the values of the genotype, phenotype (parameters), as well as the RMSE and the squared correlation coefficient, r^2, on the screen. As an example, for a step $j = 114$ of a Langmuir-Freundlich isotherm model fitting, it would look like:

```
mol: 0  step: 114  Fitness: 0.061057  R^2: 0.990488  Similarity: 2191/4096   Finishing: 80/100
number of parameters: 3
genotype: 0011111111110011000000000000000000000000000000000000000000000101100 parameter: 1.1875
genotype: 0011111011111111111111111111111111111111111111111111111111111111111 parameter: 3.0518e-05
genotype: 0011111111111011000011011001000011001101100111011100111000101000 parameter: 1.38162
```

When the GA reached an optimal genotype ($r^2 \to 1$ and RMSE $\to 0$ or RUPTURA has reached the maximum number of steps allowed), the candidate for the global minimum is refined during the NMA stage.

NMA (simplex) stage

We start with the parameter values corresponding to the phenotype of the best individual in the GA population. At each iteration k, the $M+1$ vertices of the simplex, Δ_k, are converging to the minimum, and we sort and label the vertices according to Eq. 3.294, such that:

$$\phi(\mathbf{a}_1) \leq \phi(\mathbf{a}_2) \leq \ldots \leq \phi(\mathbf{a}_{M+1}), \tag{3.297}$$

being $\mathbf{a}_1$ the best, and $\mathbf{a}_{M+1}$ the worst. In addition, we also calculate the centroid of the best M vertices, i.e., avoiding the worst:

$$\overline{\mathbf{a}} = \sum_{i=1}^{M} \frac{\mathbf{a}_i}{M} \tag{3.298}$$

Following the version of [384] of the Nelder–Mead algorithm, we updated the vertices by performing four possible operations: reflection, expansion, contraction, and shrinkage. Each of these operations is determined by a scalar: $\alpha = 1$, $\beta = 2$, $\gamma = 1/2$, and $\delta = 1/2$, respectively. We have used a version in which these scalars are constant, but there are modern implementations in which these parameters are adaptable [385]. In Fig. 3.16 we show the algorithm of a single iteration step. In more detail, the four operations are

1. Reflection: Once we have ordered the vertices according to Eq. 3.297, and we have calculated the centroid of Eq. 3.298, we proceed to calculate the reflected point, $\mathbf{a}_r$:

$$\mathbf{a}_r = \overline{\mathbf{a}} + \alpha(\overline{\mathbf{a}} + \mathbf{a}_{M+1}) \tag{3.299}$$

   We evaluate this point, $\phi(\mathbf{a}_r)$. If $\phi(\mathbf{a}_r) \leq \phi(\mathbf{a}_r) < \phi(\mathbf{a}_M)$, then we accept the reflection point: $\mathbf{a}_{M+1} = \mathbf{a}_r$.

2. Expansion: If $\phi(\mathbf{a}_r) \leq \phi(\mathbf{a}_1)$ then we calculate the expansion point, $\mathbf{a}_e$:

$$\mathbf{a}_e = \overline{\mathbf{a}} + \beta(\mathbf{a}_r - \overline{\mathbf{a}}) \tag{3.300}$$

   We evaluate this expansion point, $\phi(\mathbf{a}_e)$. If $\phi(\mathbf{a}_e) < \phi(\mathbf{a}_r)$ then we accept $\mathbf{a}_{M+1} = \mathbf{a}_e$, otherwise we accept the reflection point: $\mathbf{a}_{M+1} = \mathbf{a}_r$.

3. Contraction.
   a) Outside: If $\phi(\mathbf{a}_M) \leq \phi(\mathbf{a}_r) < \phi(\mathbf{a}_{M+1})$ we perform an outside contraction:

$$\mathbf{a}_{oc} = \overline{\mathbf{a}} + \gamma(\mathbf{a}_r - \overline{\mathbf{a}}) \tag{3.301}$$

      and we evaluate the new point, $\phi(\mathbf{a}_{oc})$. If $\phi(\mathbf{a}_{oc}) \leq \phi(\mathbf{a}_r)$ then we accept the outside contraction point: $\mathbf{a}_{M+1} = \mathbf{a}_{oc}$, else we perform a shrinkage operation.

   b) Inside: If $\phi(\mathbf{a}_r) \geq \phi(\mathbf{a}_{M+1})$ we perform an inside contraction:

$$\mathbf{a}_{ic} = \overline{\mathbf{a}} - \gamma(\mathbf{a}_r - \overline{\mathbf{a}}) \tag{3.302}$$

      and we evaluate the new point, $\phi(\mathbf{a}_{ic})$. If $\phi(\mathbf{a}_{ic}) \leq \phi(\mathbf{a}_{M+1})$ then we accept the inside contraction point: $\mathbf{a}_{M+1} = \mathbf{a}_{ic}$, else we perform a shrinkage operation.

4. Shrinkage: We generate the new vertices:

$$\mathbf{v}_i = \mathbf{a}_1 + \delta(\mathbf{a}_i - \mathbf{a}_1), \qquad \forall i = 2, \ldots, M+1. \tag{3.303}$$

   These will be the new vertices, $\mathbf{a}_{i=1}^{M+1}$, in the next iteration.

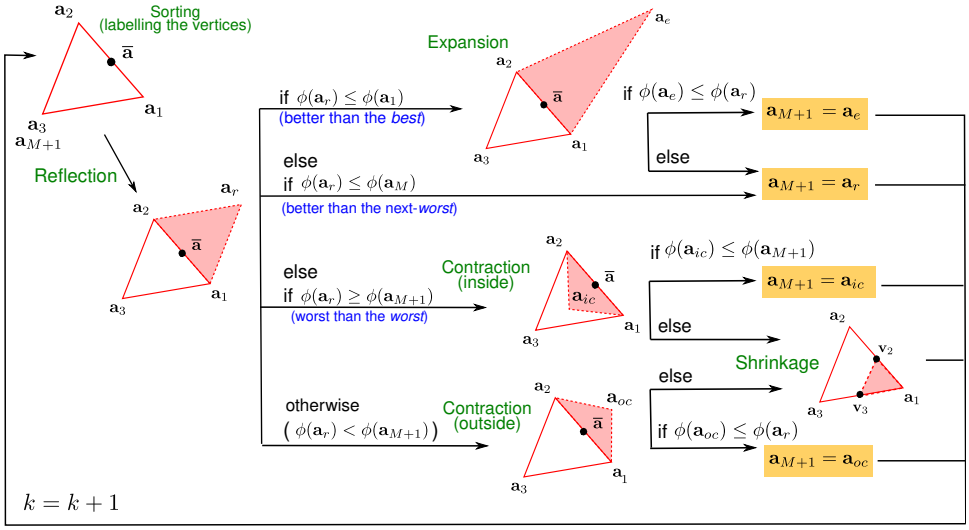


Figure 3.16. Flowchart of one step, k, in the Nelder-Mead algorithm [230] for a parametric function, f, with two parameters. In this way, the simplex is a triangle, and the visualization is easier. The four operations and the sorting stage are in green labels.

3.5.6. Validation

To validate the RUPTURA fitting function and compare it with other widely used codes (ISOFIT, IAST++, pyIAST, GAIAST, and Gnuplot; see subsection 3.5.4) we fit the adsorption isotherms of n-C_7 in BEA zeolite at 552 K. The isotherm exhibits a behavior that can be modelled by many kinds of models: Langmuir, Langmuir-Freundlich, Sips, Toth, etc. We have chosen the Toth model of a single adsorption site (three parameters: q^{sat}, b, and ν). For this relatively simple isotherm shape, all codes are able to accurately fit the isotherm model to the data, as shown in Fig. 3.17. The obtained goodness-of-fits were $r^2 > 0.998$, and RMSE < 0.025 [mol/(kg framework)], even from scratch. ISOFIT++ does not incorporate the Toth model. The fitting of the isotherm using the Langmuir model using the IAST++ code gave reasonable results (RMSE = 0.025 [mol/(kg framework)]). RUPTURA takes 1.1 s to do this fit using a single-core Intel Core i7-10700K CPU at 3.80GHz.

A much more challenging case is the adsorption isotherm of o-xylene in MAF-X8 at 300 K. The isotherm shows, see Fig. 3.18, two steep slopes at 1×10^{-6} and 10 Pa around an inflection ('kink'). By adding a third site to the model, it fits well to a third bend of the isotherm. Adding a fourth site does not qualitatively improve the model. We have modelled the isotherm using the Langmuir-Freundlich model with two adsorption sites. The dual-site Langmuir-Freundlich model has six parameters (q_1^{sat}, b_1, ν_1, q_2^{sat}, b_2, and ν_2). In general, for an increased number of model

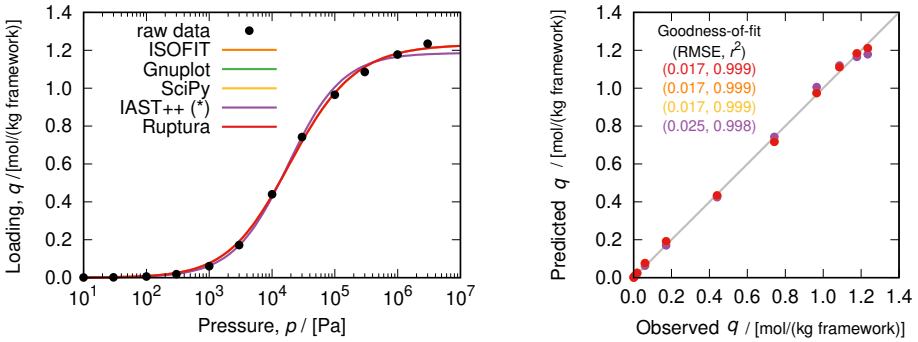


Figure 3.17. (Left) Adsorption of n-C_7 in BEA zeolite at 552 K and some fitted models using ISOFIT, Gnuplot, IAST++, and RUPTURA for a Toth model of a single adsorption site. (Right) Predicted loading vale vs. observed loading values. The goodness-of-fit in colors (same code color of the Left panel) for each code. The RMSE values are in [mol/(kg framework)] units.

parameters, it is more likely that local minima will appear, and global optimization algorithms become necessary.

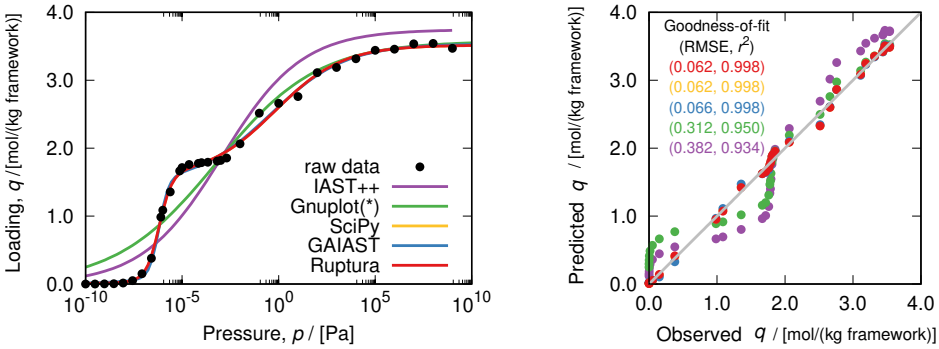


Figure 3.18. (Left) Adsorption of o-xylene in MAF-X8 zeolite at 300 K and some fitted models using IAST++, Gnuplot (from scratch), GAIAST, and RUPTURA for a dual-site Langmuir-Freundlich model. (Right) Predicted loading values vs. observed loading values. The goodness-of-fit for each code in colors defined in the left panel. The RMSE values are in [mol/(kg framework)] units.

GAIAST and RUPTURA predict equivalent values (the fitting algorithms are very similar), but RUPTURA is much more optimized and performs the same fitting in ten times less time, with goodness-of-fit values of RMSE ≤ 0.066 [mol/(kg framework)] and $r^2 = 0.998$. However, the optimizations from scratch using pyIAST (not shown in the Figure), Gnuplot (from scratch), or

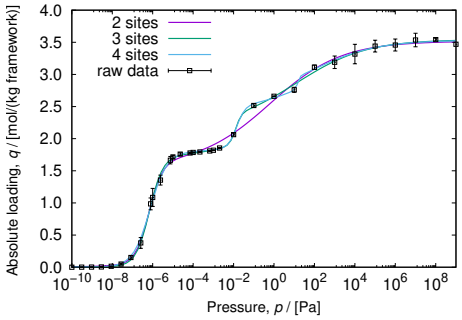


Figure 3.19. The adsorption isotherm prediction for two, three, and four sites Langmuir-Freundlich isotherm models for *o*-xylene in MAF-X8. Without the experimental error information for the loading data is difficult to establish what kind of model is more realistic.

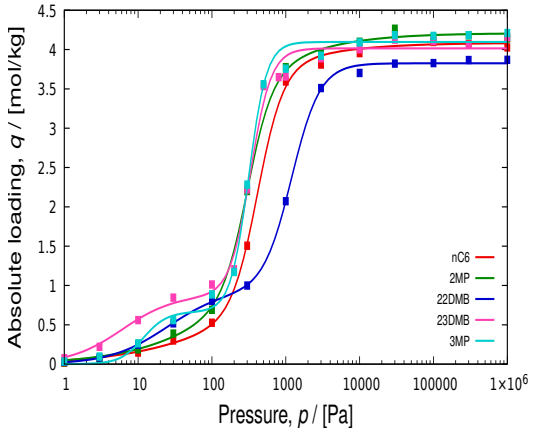


Figure 3.20. The C_6-alkane-isomers adsorption isotherms in MAF-6 computed from grand-canonical Monte Carlo simulations fitted with dual- and triple-site Langmuir-Freundlich isotherm models. The peculiar shape of these isotherms in the low-pressure region makes this system challenging to perform a reliable fitting.

IAST++, with RMSE ≥ 0.312 [mol/(kg framework)] were unable to accurately fit the model. That is, the error is too large for the model to represent the isotherm data well and to be used in IAST or breakthrough calculations. Only if the starting values are chosen close to the solution is Gnuplot able to generate models with high goodness of fit (see Table 3.13). RUPTURA takes 2.6 s to do this fit, and GAIAST, 26 s. If we consider a Langmuir-Freundlich with three sites (nine parameters) for the model in the fitting procedure, RUPTURA takes 6.279 s, obtaining RMSE = 0.0385 [mol/(kg framework)], and $r^2 = 0.9993$. However, erroneous results could be obtained by over-fitting a simple isotherm with an unnecessarily complex model. Regardless of this issue, RUPTURA can quickly optimize difficult isotherms (with multiple minima) as shown in Fig. 3.19.

A particularly difficult test case is alkane C_6 isomers in MAF-6 at 298 K with a relatively small amount of points, but with small error bars. Fig. 3.20 shows that RUPTURA produces excellent fits using dual- and triple-site Langmuir-Freundlich models. The shape of the isotherms, combined with smooth and sharply rising slopes over many orders of magnitude of pressure, poses a challenge for the fit. Note that using a small amount of points over a large pressure range (equally distributed in log-scale) is a very common scenario to explore the overall shape of the isotherms.

Lastly, we used Python SciPy Library to test the RUPTURA code (see Listing 3.1). Scipy (`scipy.optimize.curve_fit`) uses the Trust Region Reflective (TRR) method for constrained problems [386, 387]. The TRR method is an evolution of the Levenberg-Marquardt algorithm, but with global convergence. In Figs. 3.17 and 3.18, as well as Table 3.13, you can see that RUPTURA gives goodness-of-fit comparable to TRR.

Listing 3.1: Python script used to fit isotherms with Toth and dual-site Langmuir-Freundlich models using the SciPy Library.

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from scipy.optimize import curve_fit

def Toth(P, q, b, n):
    return q * b * P / pow(1 + pow(b*P,n),(1.0/n))

def DSLangmuirFreundlich(P, q1, b1, n1, q2, b2, n2):
    return q1 * b1 * pow(P,n1) / (1 + b1 * pow(P,n1)) + q2 * b2 * pow(P,n2)
    / (1 + b2 * pow(P,n2))

df = pd.read_csv('Results.dat-MAF-x8-P1-Repeat-433K-o-xylene', sep=' ')
popt, pcov = curve_fit(DSLangmuirFreundlich,
                       df['P'].values, df['L'].values,
                       p0=(1.0,1.0,1.0,1.0,1.0,1.0),
                       bounds=([0,1.0E-20,0.1,0,1.0E-20,0.1],
                       [100,1.0E+20,10.0,100,1.0E+20,10.0]))
print(*popt)
```

Table 3.13. Fitted parameters for the adsorption isotherm (dual-site Langmuir-Freundlich) of o-xylene in MAF-x8-P1 zeolite at 433K using different methods. In this table, unlike in Fig. 3.18, the Gnuplot fit has been performed using initial values close to the final solution, to avoid being stuck in a local minimum.

	q_1^{sat}/ [mol/(kg framework)]	b_1/ [Pa^{-1}]	ν_1/ [-]	q_2^{sat}/ [mol/(kg framework)]	b_2/ [Pa^{-1}]	ν_2/ [-]	RMSE/ [mol/(kg framework)]
RUPTURA	1.9377	1.1132	0.2476	1.5793	1.0485×10^8	1.2963	0.062
SciPy	1.9583	1.1334	0.2439	1.5617	2.8218×10^8	1.3654	0.062
Gnuplot	1.9554	1.1305	0.2438	1.5642	2.4394×10^8	1.3553	0.064
GAIAST	2.0059	1.1785	0.2337	1.5215	3.2212×10^9	1.5360	0.066

3.6. RUPTURA Installation

For the basic functionality, only a C++11 compiler is needed. However, automatic picture generation is based on `gnuplot` [380], and automatic movie generation is based on `gnuplot` and `ffmpeg` [388]. FFmpeg needs to have support for `HEVC/H.265` and/or `H.264` video encoding.

The default color scheme of Gnuplot has several issues: the sequence of colors is hard to see for color-blind people, and yellow is hardly visible against a white background. We therefore use the color scheme published in the book 'Gnuplot in action', stylesheet listing 12.7 [389]. All the required software tools are available as open-source software.

3.6.1. Mac

Some additional software tools need to be installed.

First, the `Xcode Command Line Tools` need to be installed. The easiest way to install Xcode Command Line Tools is by installing Homebrew, the popular package manager for macOS. When you install Homebrew, you will be offered the option of installing Xcode Command Line Tools. Go to the website https://brew.sh and copy and paste the installation command into a terminal. The command line tools include `git`, a version control system. Using Homebrew, we can also install `gnuplot` and `ffmpeg`.

```
brew install gnuplot
brew install ffmpeg
```

Next, download RUPTURA using `git` and compile the code

```
git clone https://github.com/iraspa/ruptura
cd ruptura
cd src
make
```

3.6.2. Linux

Some additional software tools need to be installed. For Ubuntu-based systems

```
sudo apt install git build-essential gnuplot ffmpeg
```

Next, download RUPTURA using `git` and compile the code:

```
git clone https://github.com/iraspa/ruptura
cd ruptura
cd src
make
```


3.6.3. Windows

Some additional software tools need to be installed.

- git
  The official build is available for download on the Git website. Just go to https://git-scm.com/download/win and the download will start automatically.

- gnuplot
  Go to https://sourceforge.net/projects/gnuplot/files/gnuplot/ to download gnuplot, and install it via the binary installer. Preferably install gnuplot in `C:\Program Files\gnuplot`, or alternatively add the directory `C:\Program Files\gnuplot\bin` to the `PATH` variable.

- ffmpeg
  Go to https://github.com/BtbN/FFmpeg-Builds/releases, and download `ffmpeg-master-latest-win64-gpl.zip`. Unzip, and preferably install ffmpeg in `C:\Program Files\ffmpeg-master-latest-win64-gpl`, or alternatively add the directory `C:\Program Files\ffmpeg-master-latest-win64-gpl\bin` to the `PATH` environment variable.

Next, download RUPTURA using `git`

```
git clone https://github.com/iraspa/ruptura
```

A binary executable `src/ruptura.exe` for Windows is already provided.

3.7. RUPTURA Input description

3.7.1. General

All input is case-insensitive and white space does not matter. Font `keyword` means the literal text, [a|b] means the input of either `a` or `b`, [real] means the expected number is a floating point value, [int] means an integer is expected, and [string] means a string is expected. Strings are not allowed to contain spaces (replace these by dashes or underscores). Numbering of indices starts at zero.

Simulation types

- `SimulationType` [`MixturePrediction` | `Breakthrough` | `Fitting`]
  Selects the type of the simulation. [`MixturePrediction` selects the computation of a mixture based on pure component isotherms, the pressure-range, and the fluid-phase mole fractions. `Breakthrough` selects the computation of a mixture adsorbing in a fixed-bed adsorber. `Fitting` selects the computation of the best fit of an analytic isotherm model to raw adsorption data.

Component information

- `Component` [int] MoleculeName [string]
  The index of the component and the name of the component. The component name is used in the automatically generated plot-files and movies.

- `Filename` [string]
  The name of a file associated with this component. Used in fitting to specify the raw data of the isotherm.

- `CarrierGas` [yes|no]
  Specifies whether this component is the carrier gas in a breakthrough simulation. There can be only one carrier-gas and a breakthrough simulation *must* contain a carrier-gas. For the carrier gas there is no need to specify an isotherm model (the carrier gas is assumed not to adsorb).

- `GasPhaseMolFraction` [real]
  The dimensionless gas-phase mole fraction of the component. The values of all the components should sum up to unity, but if not, the code gives a warning and continues with the explicitly normalized mole fractions.

- `MassTransferCoefficient` [real]
  The mass-transfer coefficient of this component used in the Linear Driving Force (LDF) model. Units: 1/s.

- `AxialDispersionCoefficient` [real]
  The coefficient of axial mixing attributed to a diffusion-like process. Units: m^2/s.

- `NumberOfIsothermSites` [int]
  The number of isotherm sites associated with this component. This keyword is followed by a list of isotherm models and parameters chosen from (See Table 3.1 for definition of the models):

  - `Langmuir` [real] [real]
    Parameter b_0 in mol kg^{-1}, parameter b_1 in Pa^{-1}.

- `Anti-Langmuir` [real] [real]
  Parameter b_0 in mol kg^{-1} Pa^{-1}, parameter b_1 in Pa^{-1}.
- `BET` [real] [real] [real]
  Parameter b_0 in mol kg^{-1}, b_1 dimensionless, parameter b_2 dimensionless.
- `Henry` [real]
  Parameter b_0 in mol kg^{-1} Pa^{-1}.
- `Freundlich` [real] [real]
  Parameter b_0 in mol kg^{-1} Pa^{-1}, parameter b_1 dimensionless.
- `Sips` [real] [real] [real]
  Parameter b_0 in mol kg^{-1}, parameter b_1 in Pa^{-1}.
- `Langmuir-Freundlich` [real] [real] [real]
  Parameter b_0 in mol kg^{-1}, parameter b_1 in Pa^{-1}, parameter b_2 dimensionless.
- `Redlich-Peterson` [real] [real] [real]
  Parameter b_0 in mol kg^{-1} Pa^{-1}, parameter b_1 in Pa^{-1}, parameter b_2 dimensionless.
- `Toth` [real] [real] [real]
  Parameter b_0 in mol kg^{-1}, parameter b_1 in Pa^{-1}, parameter b_2 dimensionless.
- `Unilan` [real] [real] [real]
  Parameter b_0 in mol kg^{-1}, parameter b_1 in Pa^{-1}, parameter b_2 dimensionless.
- `O'Brien & Myers` [real] [real] [real]
  Parameter b_0 in mol kg^{-1}, parameter b_1 in Pa^{-1}, parameter b_2 dimensionless.
- `Quadratic` [real] [real] [real]
  Parameter b_0 in mol kg^{-1}, parameter b_1 in Pa^{-1}, parameter b_2 in Pa^{-2}.
- `Temkin` [real] [real] [real]
  Parameter b_0 in mol kg^{-1}, parameter b_1 in Pa^{-1}, parameter b_2 dimensionless.
- `Bingel & Walton` [real] [real] [real]
  Parameter b_0 in mol kg^{-1}, parameter b_1 in Pa^{-1}, parameter b_2 in Pa^{-1}.

3.7.2. Mixture prediction

The mixture prediction module predicts, using various possible theoretical models, the adsorbed phase mole fractions based on pure component isotherms only.

- `MixturePredictionMethod` [`IAST` | `SIAST` | `EI` | `SEI`]
  The method used to predict the mixture adsorption based on pure component isotherm information. `IAST` denotes the Ideal Adsorption Solution Theory (IAST) [51], and `SIAST` is segregated IAST where IAST is applied to sites individually [56]. `EI` is the explicit isotherm model (not iterative and hence very fast) that is applicable to Langmuir models only [93, 118]. `SEI` is the segregated explicit isotherm model. These models are explained in Secs. 3.3.4 and 3.3.5. Default: IAST.

- `IASTMethod` [`FastIAS` | `Bisection`]
  The method used to compute the IAST mixture prediction. Default: FastIAS.

- `PressureStart` [real]
  The lowest pressure of the range of pressures to be evaluated. Must be smaller than `PressureEnd`. Units: Pa.

- `PressureEnd` [real]
  The highest pressure of the range of pressures to be evaluated. Must be larger than `PressureStart`. Units: Pa.

- `NumberOfPressurePoints` [int]
  The number of points equally spaced in log-scale or linear scale.

- `PressureScale` [`log` | `linear`]
  The scale of the range of pressures, either logarithmic scale (`log`) or linear (`linear`).

3.7.3. Breakthrough

The breakthrough module allows the computation and evaluation of the performance of a fixed bed adsorber. A fixed bed packed with particles containing a porous material is pressurized and purged with a carrier gas. A (mixture-) fluid is added to the carrier gas, resulting in a step-wise change of the inlet concentration. In chromatographic separation processes, a pulse-wise change of the inlet concentrations is used. The change of the component concentrations along the column and at the outlet of the fixed bed are recorded.

Simulation duration

- `BreakthroughType` [`step` | `pulse`]
  Selects a step breakthrough initial condition at the inlet, or a pulse condition as explained in Secs. 3.4.7 and 3.4.8.

- `Pulselength` [real]
  The length in time of the pulse condition at the inlet. Units: s, default: 10.0.

- `NumberOfTimeSteps` [int| `auto`]
  Explicitly set the number of time steps of the breakthrough computation or (only for step-breakthrough) let the program determine when the breakthrough computation is converged (`auto`).

- `TimeStep` [real]
  The time step used in the numerical integration scheme. Units: s, default: 0.0005.

- `PrintEvery` [int]
  How frequently to write status information to the screen. Default: 10000.

- `WriteEvery` [int]
  How frequently to write data to the output files. Default: 10000.

Column properties

- `DisplayName` [string]
  The name of the column, or material, that will be displayed in the output files (plots and movies).

- `ColumnVoidFraction` [real]
  The typical value of the fixed-bed porosity or bulk void fraction is ε_B=0.38-0.40. Units: -, default: 0.4.

- `ParticleDensity` [real]
  The particle density ρ_P (density of the adsorbent grain) includes the inner porosity but excludes the void fraction of the bed. Units: kg/m^3, default 1000.0.

- `TotalPressure` [real]
  The total pressure at which to compute the breakthrough. Units: Pa, default: `1e6`.

- `PressureGradient` [real]
  The gradient of the pressure along the column. Units: Pa/m, default: 0.

- `ColumnEntranceVelocity` [real]
  The interstitial velocity of the fluid at the entrance of the column. Units: m/s, default: 0.1.

- `ColumnLength` [real]
  The length of the column. Units: m, default: 0.3.

Integration settings

- `NumberOfGridPoints` [int]
  The number of grid points used in the spatial discretization of the column. Default: 100.

3.7.4. Fitting

The fitting module allows the fitting of isotherm models on raw, computed, or measured adsorption isotherms.

- `ColumnPressure` [int]
  The index of the column in the raw data file containing the information on the pressure.

- `ColumnLoading` [int]
  The index of the column in the raw data file containing the information on the absolute adsorption amount.

3.8. RUPTURA Tutorial

3.8.1. Introduction

The tutorial examples are designed to be fast. The binary mixture breakthrough example of CO_2/N_2 in silicalite runs in a matter of seconds. This allows interactive computations for teaching purposes, where students can easily play around with the column properties and investigate how the separation would be influenced by the gas-phase mole fractions, the bed-void fraction, the length of the column, axial dispersion and mass-transfer coefficients, etc. The second example of C_6-isomers in BEA-type zeolite is more advanced but still runs in one or two minutes. Both examples make use of the explicit isotherm model developed by Van Assche et al. [118] (Sec. 2.2.2) for the prediction of the mixture isotherms.

3.8.2. Breakthrough of CO_2/N_2 in silicalite

The directory `tutorial/Silicalite-CO2-N2/breakthrough` contains the input file `simulation.input`:

```
SimulationType           Breakthrough

// Column settings
DisplayName              Silicalite
Temperature              313.0           // [K]
ColumnVoidFraction       0.4             // [-]
ParticleDensity          1144.03         // [kg/m^3]
TotalPressure            2.5e6           // [Pa]
PressureGradient         0.0             // [Pa/m]
ColumnEntranceVelocity   0.1             // [m/s]
ColumnLength             0.3             // [m]

// Run settings
NumberOfTimeSteps        auto
PrintEvery               1000
WriteEvery               200
TimeStep                 0.01            // [s]
NumberOfGridPoints       30
```

```
MixturePredictionMethod ExplicitIsotherm

Component 0 MoleculeName               Helium
            GasPhaseMolFraction        0.9            // [-]
            CarrierGas                 yes

Component 1 MoleculeName               CO2
            GasPhaseMolFraction        0.05           // [-]
            MassTransferCoefficient    0.06           // [1/s]
            AxialDispersionCoefficient 0.0            // [m^2/s]
            NumberOfIsothermSites      1
            Langmuir                   2.858 1.089e-5 // [mol/(kg framework)] [1/Pa]

Component 2 MoleculeName               N2
            GasPhaseMolFraction        0.05           // [-]
            MassTransferCoefficient    0.06           // [1/s]
            AxialDispersionCoefficient 0.0            // [m^2/s]
            NumberOfIsothermSites      1
            Langmuir                   2.094 0.111e-5 // [mol/(kg framework)] [1/Pa]
```


The simulation type is set to `breakthrough`, the column properties and run settings are specified, and lastly, the information on all the components. The component information starts with the keyword `Component` followed by the component index (starting from zero) and the name of the component. The properties set after this line refer to the component denoted by the `Component` keyword. A component requires isotherm information, except for the carrier gas. There must be a component labelled as carrier-gas, and there can be only one. The isotherm information for the carrier gas is implicitly a Langmuir isotherm with the affinity parameter set to zero. For a component that is not the carrier gas, the number of isotherm sites needs to be specified, followed by a list of isotherm models along with the corresponding parameters. A component also needs a gas-phase mole fraction and mass-transfer coefficient, while the axial dispersion coefficient is optional.

The code is executed on Mac/Linux by typing in the terminal

```
./run
```

On Windows, you can run it from the Windows console using the `run.bat` batch script, or by double-clicking on it. After a few seconds, the program is finished, and many new files have been created

```
column.data                make_movie_Pnorm          plot_column_P
component_0_Helium.data    make_movie_Pt             plot_column_Pnorm
component_1_CO2.data       make_movie_Q              plot_column_Pt
component_2_N2.data        make_movie_Qeq            plot_column_Q
make_graphs                make_movie_V              plot_column_Qeq
make_movie_Dpdt            make_movies               plot_column_V
make_movie_Dqdt            plot_column_Dpdt          plot_breakthrough
make_movie_P               plot_column_Dqdt
```

The `column.data` and `component` files are the files with the actual simulation data. `make_movies` creates movies to visualize the variation of different quantities along the length of the column. `Dpdt` and `Dqdt` represent time derivatives of the partial pressures and the adsorbed loadings of each component, respectively. Movies are also created for the variation of the partial pressures (`P`), normalized pressure (ratio of partial pressure at the column exit to that at the inlet: `Pnorm`), total pressure (`Pt`), adsorbed loadings (`Q`), equilibrium loading (`Qeq`), and interstitial velocity (`V`) along the

length of the fixed bed column. The `plot_column` files are used by the movie scripts to create figures for these measured quantities (using gnuplot). On Windows, the `make_movies` and `make_graphs` scripts have the `bat` extension. By running these two scripts, the figures (`breakthrough.pdf` and `breakthrough_dimensionless.pdf`) and movies (files with `.mp4` extension and file names starting with `column_movie`) will be created:

```
breakthrough.pdf                  column_movie_Pnorm.mp4
breakthrough_dimensionless.pdf    column_movie_Pt.mp4
column_movie_Dpdt.mp4             column_movie_Q.mp4
column_movie_Dqdt.mp4             column_movie_Qeq.mp4
column_movie_P.mp4                column_movie_V.mp4
```


You could also run the individual make-movie scripts if you are interested only in some of them. The figures are in `pdf` files, the movies are in `h265` format for Mac/Linux and `h264` on windows, with resolution `1200x800` by default. On Mac/Linux, command line options can be specified to modify the defaults

```
./make_movies -e 5 -w 1600 -h 1200 -q 18 -l
```

The `-e` option stands for 'every' and allows a periodic sampling only using every 5 data points in the example case. The `-w` and `-h` change the width and height of the movie, while `-q` determines the quality of the movies. The quality range is 0–51, where 0 is lossless and 51 is the worst quality possible. The default in RUPTURA is 18, which is visually nearly lossless. The `-l` option stands for legacy and changes the movie format `h265` to the older `h264`. On Windows, these options can be specified in the Windows console as

```
make_movies 5 1600 1200 18
```

Note the fixed order of the options, and the format on Windows is `h264` as `h265` is not supported out-of-the-box under Windows.

The output figures are shown in Fig. 3.21a. The breakthrough graph plots the normalized concentration at the exit of the column as a function of time. Two versions of the plots are available: (1) as a function of time in seconds, (2) as a function of dimensionless time. Fig. 3.21b shows a single frame of the movie, here the loading of the components adsorbed in the crystallites along the column.

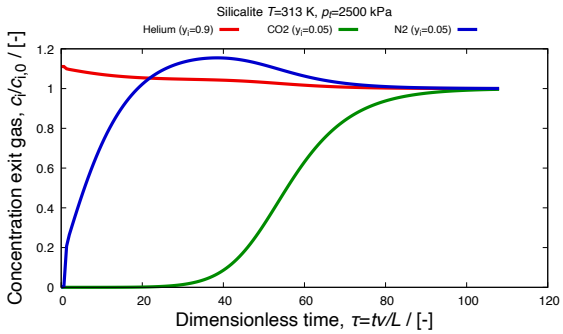
3.8.3. Mixture prediction of C_6-isomers in BEA-type zeolite

For the C_6-isomers in BEA-type zeolite example, we can test how well the explicit Langmuir models mixture adsorption performs compared to IAST. The directory `tutorial/BEA-C6/mixture_prediction` contains the input file `simulation.input`:

```
SimulationType          MixturePrediction

ColumnName              BEA
Temperature             552.0           // [K]
PressureStart           1e3             // lowest pressure
PressureEnd             1e7             // highest pressure
NumberOfPressurePoints  100             // number of points equally spaced
PressureScale           log             // [log, or linear]
```

(a)



(b)

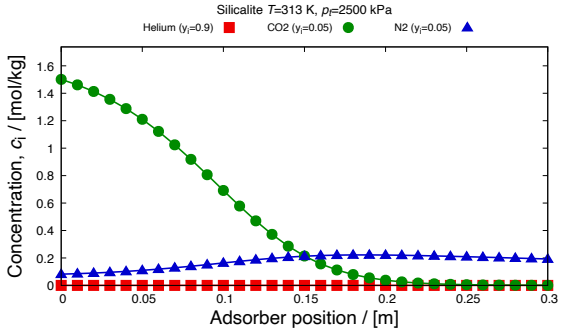


Figure 3.21. Breakthrough predictions of an equimolar mixture of CO_2-N_2 in silicalite at 313K and $2.5 \cdot 10^6$Pa: (a) breakthrough curve plot (file `breakthrough_dimensionless.pdf`) (b) movie frame of the loading inside the column (file `column_movie_Q.mp4`).

```
MixturePredictionMethod  ExplicitLangmuir // [IAST, SIAST, or ExplicitLangmuir]

Component 0 MoleculeName              Helium
            GasPhaseMolFraction       0.96            // [-]
            CarrierGas                yes

Component 1 MoleculeName              nC6
            GasPhaseMolFraction       0.01            // [-]
            NumberOfIsothermSites     1
            Langmuir                  1.393 3.307e-05 // [mol/(kg framework)] [1/Pa]

Component 2 MoleculeName              C5m3
            GasPhaseMolFraction       0.01            // [-]
            NumberOfIsothermSites     1
            Langmuir                  1.559 1.809e-05 // [mol/(kg framework)] [1/Pa]

Component 3 MoleculeName              C4m2m2
            GasPhaseMolFraction       0.01            // [-]
            NumberOfIsothermSites     1
            Langmuir                  1.714 5.451e-06 // [mol/(kg framework)] [1/Pa]

Component 4 MoleculeName              C4m2m3
            GasPhaseMolFraction       0.01            // [-]
            NumberOfIsothermSites     1
            Langmuir                  1.678 1.522e-05 // [mol/(kg framework)] [1/Pa]
```

Note that the input is very similar in structure. The major difference is that the simulation type is set to `MixturePrediction`. Using the option `MixturePredictionMethod` we can control the specific method used in the mixture prediction. Minor differences are that the mass-transfer and axial dispersion coefficients are not needed. To compute the mixture isotherm, some info on the pressure range, i.e. begin and end-pressure, and the number of pressure points is required. These points can be spread equidistant in normal or in log-scale.

After running, the output directory contains the following plot files.

```
component_0_Helium.data     make_graphs
component_1_nC6.data        plot_mixture
component_2_C5m3.data       plot_mixture_mol_fractions
component_3_C4m2m2.data     plot_pure_components
component_4_C4m2m3.data
```

Files with `.data` extension and file names starting with component consist of equilibrium loading data for the specified temperature and the range of pressure. In Windows, the script `make_graphs` has the `bat` extension. After running that script, we obtain the plots in PDF format

```
mixture_prediction.pdf
mixture_prediction_mol_fractions.pdf
pure_component_isotherms.pdf
```

`mixture_prediction.pdf` file consists of the plots of mixture adsorption isotherms for all the components present in the mixture in the same unit specified for saturation loadings (`q_sat`) in the input file. `mixture_prediction_mol_fractions.pdf` file consists of the same plots in the form of mole fractions. `pure_component_isotherms.pdf` shows the plots for the isotherms for the pure components. In Fig. 3.22, we can compare the pure component isotherms with the mixture isotherm predictions. In the mixture, we can see the effect of competitive adsorption.

3.8.4. Isotherm fitting

To obtain the isotherm model parameters, we can use the fitting module. The directory `tutorial/MAF-X8-xylenes` contains the input file

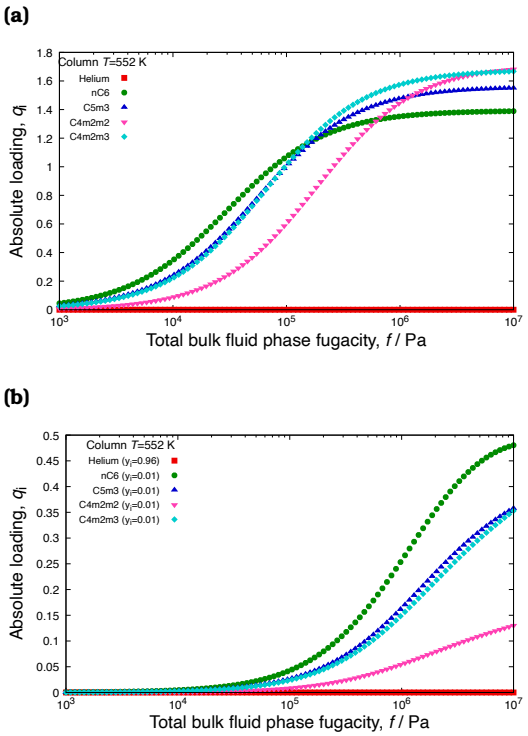


Figure 3.22. Mixture predictions of a C_6 isomer mixture in BEA- type zeolite at 552 K using explicit isotherm model developed by Van Assche et al. [118]: (a) pure components (file `pure_component_isotherms.pdf`), (b) mixture prediction (file `mixture_prediction.pdf`).

simulation.input:

```
SimulationType              Fitting

DisplayName                 MAF-X8
ColumnPressure              3
ColumnLoading               8
PressureScale               log

Component 0 MoleculeName                p-xylene
            FileName                    Results.dat-MAF-x8-P1-Repeat-433K-p-xylene
            NumberOfIsothermSites       2
            Langmuir-Freundlich         0.0 0.0 0.0
            Langmuir-Freundlich         0.0 0.0 0.0
```


The result of the fit is printed at the end of the output.

```
    number of isotherm sites:  2
    Langmuir-Freundlich isotherm
        q_sat: 1.71616                  // [mol/(kg framework)]
        b:     3.54859e+09              // [1/Pa]
        nu:    1.47094                  // [-]
    Langmuir-Freundlich isotherm
        q_sat: 1.63797                  // [mol/(kg framework)]
        b:     12.3201                  // [1/Pa]
        nu:    0.515807                 // [-]
```

File names starting with Results.dat consist of the equilibrium loading data computed using grand-canonical Monte Carlo simulations.

```
Results.dat-MAF-x8-P1-Repeat-433K-benzene
Results.dat-MAF-x8-P1-Repeat-433K-ethylbenzene
Results.dat-MAF-x8-P1-Repeat-433K-m-xylene
Results.dat-MAF-x8-P1-Repeat-433K-o-xylene
Results.dat-MAF-x8-P1-Repeat-433K-p-xylene
Results.dat-MAF-x8-P1-Repeat-433K-toluene
make_graphs
plot_fit_component_0_p-xylene
```

Again, using the make_graphs script, we can create the PDFs of the figures, e.g.,

```
isotherms_fit_p-xylene.pdf
```

The result is plotted in Fig. 3.23. It shows the initial random starting isotherm and the final fit result after GA optimization for adsorption of p-xylene in MAF-X8- type zeolite at 433K.

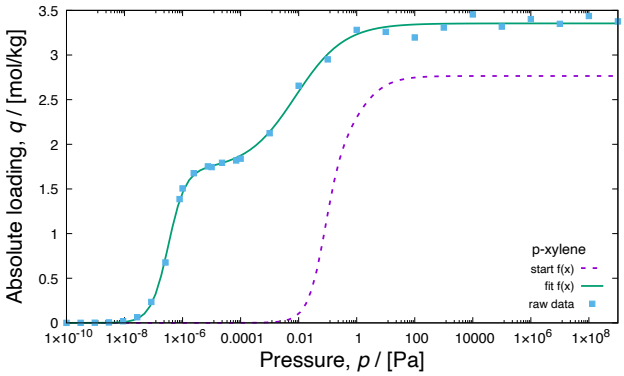


Figure 3.23. Fitting a dual-site Langmuir isotherm model to the pure component isotherm data for the adsorption of p-xylene in MAF-X8-type zeolite at 433K. The pure component adsorbed loadings are computed using grand-canonical Monte Carlo simulations. The fitting is performed using genetic algorithm optimization, which is shown in the file `isotherms_fit_p-xylene.pdf`.

Table 3.14. List of mixture prediction, breakthrough, and fitting examples provided with RUPTURA. For breakthrough simulations, in addition to the N_C-components, there is a carrier-gas component present. The timings are done on a Mac Studio using an M1-chip.

structure	mixture	N_c	N_{grid}	Δt [s]	p_i [kPa]	$\tau = tv/L$ [-]	time	p_i [kPa]	$\tau = tv/L$ [-]	time
MOR	CO_2-C_3H_8	2	100	0.0005	125.0	270.8	1m36	1250.0	53.2	0m25
JUC-77	xylenes	4	100	0.0005	25.0	201.8	4m02	250.0	43.5	0m58
MIL-125-NH2	xylenes	4	100	0.0005	25.0	355.5	7m34	250.0	77.2	2m07
CoBDP	xylenes	4	100	0.0005	25.0	395.3	8m08	250.0	68.5	1m50
MFI	xylenes	4	100	0.0005	25.0	1025.2	10m13	250.0	162.5	2m08
MIL-47	xylenes	4	100	0.0005	25.0	504.5	11m18	250.0	95.0	2m37
MAF-X8	xylenes	4	100	0.0005	25.0	643.2	13m28	250.0	99.0	2m37
CoBDP	alkanes-C_6	5	100	0.0005	20.0	489.3	13m33	200.0	86.2	2m51
BEA	alkanes-C_7	5	100	0.0005	1.2	873.24	17m10	12.0	331.7	6m57
fe2bdp3	alkanes-C_6	5	100	0.0005	20.0	3520.0	105m58	200.0	376.7	12m41
ZIF-77	alkanes-C_5-C_6-C_7	15	100	0.0005	20.0	269.3	28m03	200.0	52.8	5m44

3.8.5. Examples

RUPTURA comes with many examples included. These examples have three parts: (1) fitting the pure component isotherm data with different isotherm models to obtain isotherm parameters (e.g., equilibrium constant (b), saturation loading (q^{sat}), etc), (2) prediction of mixture isotherms with the fitted parameters as input, (3) breakthrough curve simulations with the fitted parameters as input. The examples range from a simple binary mixture

up to a 15-component C_5-C_6-C_7 alkane mixture. The xylene examples have been selected for displaying qualitatively different adsorption behavior.

The examples are listed in Table 3.14 along with typical run times for breakthrough simulations, here on a Mac Studio using an M1-chip. Most examples run in a few minutes. However, keep in mind that run-times depend on the total pressure. Increasing the pressure shifts the breakthrough to shorter dimensionless times, and decreasing the pressure shifts the breakthrough to longer dimensionless times. Also, for lower pressure or stiff systems, a smaller time-step might be in order. This could significantly increase the computational time needed for the breakthrough computation. Other factors that influence the run times are the details of the isotherm models and isotherm shapes.

3.9. RUPTURA Troubleshooting

The curve fitting does not visually match well to the isotherm data. Run the fitting several times to make sure the solution is stable. If the problem remains, then it is the isotherm model and/or the number of isotherm sites. Increase the number of sites if inflections are not represented well.

Breakthrough error: pressure gradient is too large (negative outlet pressure). The given pressure and pressure gradient result in negative pressure further down the column. This is nonphysical and indicates that the pressure gradient is chosen too high, and/or the pressure is too low.

Instability or erratic breakthrough curves. The time step might be too high, reduce the time step and check for improvement. Check the values of mass transfer and axial dispersion coefficients for physical validity. Check whether using the reference IAS method using bisection (`IASTMethod Bisection`) makes a difference. Try this first in a mixture prediction calculation, since this is fast. If you find a difference, then there is a numerical instability in the implemented FastIAS method.

Breakthrough curves with sharp transitions or spikes. Use more grid points in the spatial discretization of the column.

The breakthrough curve depends on the time step and number of grid points. The breakthrough curves are correct in the limit of a very small time step and a large number of grid points. Generally, you will have to choose the time-step and number of grid points such that a further decrease in time-step or increase in the number of grid points will not (visually) change the breakthrough curves. Note that always using a very small time step and a large number of grid points is computationally expensive.

No convergence or hangs Use recommended units, like mol/kg for loading, and Pascal for pressure. Other units could lead to very small or very large numbers.

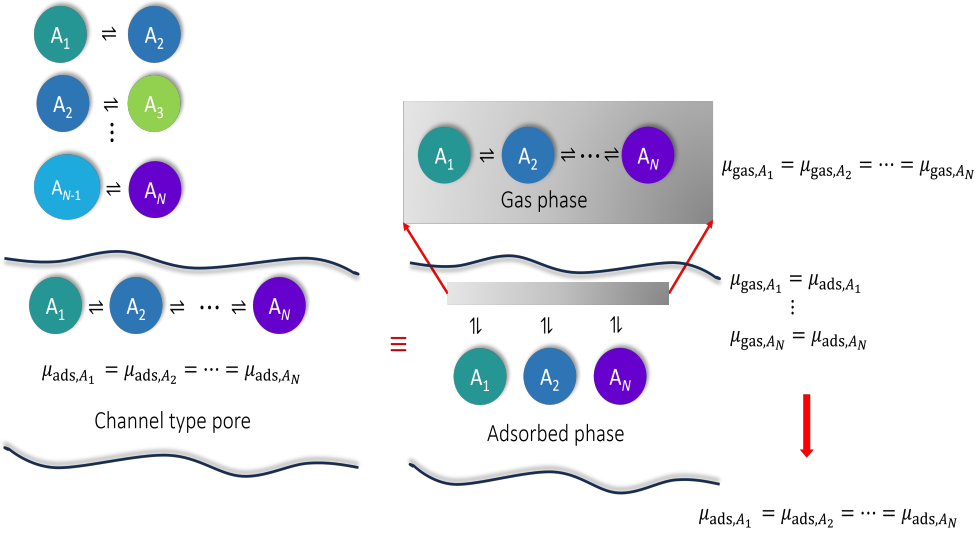
3.10. Conclusions

We presented the RUPTURA code (https://github.com/iraspa/ruptura), which is a freely available, open-source package for the computation of breakthrough curves, mixture adsorption, and fitting of isotherm models to raw adsorption data. RUPTURA contains three modules of workflow encountered in this field: (1) the computation of step/pulse breakthrough, (2) the prediction of mixture adsorption (used in the breakthrough equations) based on pure component isotherms, and (3) the fitting of isotherm models on raw (computed or measured) isotherm data. We included isotherm models like Langmuir, BET, Henry, Freundlich, Sips, Langmuir-Freundlich, Redlich-Peterson, Toth, Unilan, O'Brian & Myers, Asymptotic Temkin, and Bingel & Walton, including their multi-site versions or combinations of these isotherms. The mixture prediction methods implemented include Ideal Adsorption Solution Theory (IAST), segregated IAST, and explicit Langmuir methods. IAST is computed fast and at machine precision. The breakthrough simulations include axial dispersion and the Linear Driving Force (LDF) model for mass transfer, and have excellent numerical stability through the use of Strong-Stability Preserving Runge-Kutta (SSP-RK) integrators. RUPTURA is freely available (MIT license) and should be viewed as a demonstration code that may be useful for researchers working in the field and for teaching in chemistry and chemical engineering classes. For future versions of RUPTURA, we envision including support for non-ideal gas behavior, liquid phase conditions, and chemical reactions. Non-isothermal operation is an important aspect that needs to be considered. Variation in temperature along the column can be significant for certain cases, such as adsorption of CO_2. Therefore, in the next version of RUPTURA, we plan to include non-isothermal operations. Also, the code will be modified to simulate adsorption columns with more than one type of adsorbent. This will be useful in separating multiple components inside the column.

4

Reaction Equilibrium Distribution in Hydroisomerization

This chapter is based on the following publication: S. Sharma, M. S. Rigutto, E. Zuidema, U. Agarwal, R. Baur, D. Dubbeldam, and T. J. H. Vlugt. "Understanding shape selectivity effects of hydroisomerization using a reaction equilibrium model". *J. Chem. Phys.* 160 (2024), p. 214708.



4.1. Introduction

In transitioning towards fuels and chemicals from renewable sources, platforms that provide clean hydrocarbon liquid energy carriers derived from carbon dioxide directly or via bio-components can play an important role [391–393]. For sustainable aviation fuel and low-carbon gas-oil or lubricants, isoalkanes with a high degree of branching are the preferred constituents [394]. Hence, zeolite-catalysed shape-selective hydroisomerization [395], often called catalytic dewaxing, will be a key step in the production of such sustainable fuels and lubricants, similar to its current application in refineries and Gas-To-Liquid technology [396, 397].

Hydroisomerization is the process of converting normal alkanes into branched isomers in the presence of H_2 [1]. In hydroisomerization reactions, linear alkanes are adsorbed and dehydrogenated at the metal sites of the zeolites [21]. This leads to the formation of alkenes, which migrate to the acid sites of the zeolites. These alkenes form alkylcarbenium ions via protonation, which are transferred to the metal sites where alkanes are produced through hydrogenation reactions [21]. Hydroisomerization reactions are generally performed in a special category of nanoporous materials called zeolites [398]. Zeolites are important for the selective production and separation of branched hydrocarbons [399] due to their ability to sort molecules primarily based on size exclusion [17]. This is due to a regular pore structure of molecular dimensions. The maximum size of the molecular species that can enter the zeolite pores is controlled by the channel dimensions [17]. It is very important to understand the role of the shapes and sizes of channels present in different zeolites for determining the selectivity of the hydroisomerization process [400]. This is because hydroisomerization reactions produce a large number of isomers and cracked products. The number of isomers formed increases rapidly with an increase in the carbon chain length [401]. Therefore, a detailed analysis of the reaction product distribution in experiments is not always possible.

Hydroisomerization reactions are reversible in nature. To identify conditions that optimize reaction product yields, it is important to comprehend the role of chemical equilibrium in these reactions. Estrada-Villagrana et al [31] modeled the hydroisomerization of a mixture of alkanes (C_6, C_7, C_8, and C_9) in Pt/USY-type zeolite at reaction equilibrium. These authors studied the following cases: (1) hydroisomerization with no restriction to cracking reactions; (2) limiting the cracking of 3,3-m-C_5 and 2,2,3-m-C_5 isomers, and (3) a system without any cracking reaction. The reaction product distribution closest to experimental results [402] was obtained by limiting the cracking of 3,3-m-C_5 and 2,2,3-m-C_5 isomers[31]. The selectivities obtained for the case without any cracking reaction (Case 3) differed significantly from the experimental results, which involved only reaction products due to isomerization. In the experiments [402], the cracking of longer chain isomers (e.g., C_9) was observed [31]. However, Estrada-Villagrana et al. [31] emphasized the importance of analysing such systems (Case 3) at reaction equilibrium

to enhance selectivity towards isomerization over cracking reactions. Also, comparing reaction equilibrium distribution with data from experiments or kinetic models helps us to distinguish between thermodynamic and kinetic contributions to the reaction product distribution. Steijns and Froment [33] performed experiments of hydroisomerization of C_{10} molecules on Pt/Y-type zeolite and compared the results to reaction equilibrium data [403] obtained from literature. These authors have observed that following the hydroconversion process, the fraction of n-C_{10} in the mixture is larger than the corresponding fraction at equilibrium. The concentrations of the di-branched and multi-branched isomers are always lower than the thermodynamic equilibrium data [33]. Steijns and Froment [33] have concluded that for a mixture of isomers with the same degree of branching, thermodynamic equilibrium is reached at medium conversion (ca. 40%-70%). For a mixture of isomers with varying degrees of branching, equilibrium is never reached. This aligns with the widely accepted carbenium ion mechanism [33]. In this work, reaction equilibrium distributions of hydroisomerization of C_7 and C_8 isomers are studied which will serve as the first step towards understanding the reaction product distribution of hydroisomerization which is influenced by adsorption of reactants, reaction kinetics, diffusion of molecules inside zeolites and desorption of reaction products [17, 395]. The aim of this study is to understand the shape-selectivity effects of FAU-, ITQ-29-, BEA-, MEL-, MFI-, MTW-, and MRE-type zeolites (Fig. 1.2 in chapter 1.2) on hydroisomerization of C_7 and C_8 isomers at reaction equilibrium. This study does not account for reaction kinetics and diffusion of isomers inside zeolites. Cracking reactions, being irreversible [20] in nature, are also excluded from this analysis. For modeling of cracking reactions, the reader is referred to Ref. [404]. Future investigations will aim at understanding the reaction product distributions for the isomerization of long-chain alkanes and to which extent chemical equilibrium is applicable.

Achieving chemical reaction equilibrium in experiments is a challenging task, as hydroisomerization reactions are always accompanied by irreversible cracking reactions. An alternative pathway to study reaction equilibria is by performing Reaction Ensemble Monte Carlo (RxMC) simulations [405–407] using a classical force field-based description of hydrocarbon-zeolite and hydrocarbon-hydrocarbon interactions. Grand-Canonical Monte Carlo (GCMC) simulations coupled with the reaction ensemble can be used to compute the chemical equilibrium distribution [408, 409]. Another way to compute the adsorbed phase reaction equilibrium distribution is by using the RxMC method in the constant pressure Gibbs ensemble (GE-NPT) [408, 410]. The reaction ensemble trial moves involve the removal of randomly selected reactant molecules and the insertion of reaction product molecules simultaneously [407, 411]. The efficiency of this algorithm is low for systems with high molecule densities and a large number of independent reactions [407, 411]. Therefore, these methods can become computationally very time-consuming for studying the hydroisomerization of alkanes, which

consists of a large number of possible reactions, especially for long-chain hydrocarbons. There is a faster alternative to compute reaction equilibrium distributions when chemical reactions and phase equilibrium occur simultaneously [31, 412]. This method has several requirements: (1) Each species should achieve phase equilibrium between all phases; (2) The reactions must be in chemical equilibrium in each phase. For a system at constant temperature and pressure, the Gibbs free energy of the system is minimum [412] at equilibrium. The chemical reaction equilibrium distribution can be obtained by minimizing the molar Gibbs free energy of the system with respect to variations that are consistent with the stoichiometry of the chemical reactions [412]. For mono-molecular reactions, this can be achieved by simply equating the Gibbs free energies or the chemical potentials of the components present in the reaction equilibrium mixture. A detailed explanation of this method is provided in Section 2 of this study. Here, we impose chemical equilibrium in the gas phase for the hydroisomerization of alkanes and phase equilibrium of the isomers between the gas and the adsorbed phase. This automatically mimics the reaction equilibrium conditions in the adsorbed phase. Therefore, the adsorbed phase distribution is the reaction equilibrium distribution in this case. At reaction equilibrium, the chemical equilibrium of alkenes as intermediates may play an important role. In this study, we consider only the reaction equilibrium of the alkane isomers involved in the reactions. This is due to the lack of availability of thermodynamic data for the corresponding alkene intermediates. We consider equilibrium between dehydrogenation of alkanes into alkenes, protonation of alkenes to form alkylcarbenium ions and formation of corresponding alkane isomers via hydrogenation. The reaction equilibrium distributions, both at infinite dilution and high pressures, are investigated. The gas phase equilibrium distribution at infinite dilution is computed using the ideal gas chemical potential [413–415], and the corresponding distribution in the adsorbed phase is calculated using Henry's law.

This chapter is organized as follows: The important concepts and the theory behind the method to compute reaction equilibrium distributions are discussed in Section 4.2. The simulation techniques used in this study are explained in detail in Section 4.3. In Section 4.4, the shape-selectivity effects of different zeolites on the Henry coefficients and the reaction equilibrium distribution of C_7 isomers at both infinite dilution and finite loading conditions are analysed. It is observed that multi-branched isomers are favored in cage-like structures (ITQ-29-type zeolite) and mono-branched isomers are favored in channel-like structures. At high pressures, the gas phase composition is computed by considering deviations from ideal gas behavior. The adsorbed phase distribution at finite loadings is obtained using Ideal Adsorbed Solution Theory (IAST) [51, 54]. It is observed that the gas phase equilibrium distribution is not significantly influenced by the gas phase pressure, especially at temperatures equal to or above 500 K. Therefore, variations in the selectivity in the adsorbed phase are influ-

enced by the pure component adsorption isotherm parameters which are used as input for mixture isotherm calculations. In Section 4.5, we discuss conclusions on how the shape-selectivity of various zeolites influences the hydroisomerization reaction. The force field parameters required for the calculations of the Henry coefficients are listed in Appendix A.2. Results and raw data (Henry coefficients and adsorbed phase selectivities) pertaining to C_8 isomers are also included in Appendix A.2. The chemical thermodynamic data for C_7 and C_8 isomers obtained from the thermodynamic tables by Scott [416] are listed in the Supporting Information (SI2.xlsx) of Ref. [390]. The file SI2.xlsx also contains data for the gas phase distribution of C_7 and C_8 isomers at reaction equilibrium. Additionally, the Henry coefficients, gas phase selectivity, adsorbed phase selectivity data at infinite dilution, the pure component fugacity coefficients, and the gas phase distribution at finite loadings for these isomers are also included in Supporting Information (SI2.xlsx) of Ref. [390]. The nomenclature for all of these isomers is listed in the worksheet Intro of SI2.xlsx [390].

4.2. Theory

In the context of chemical reaction equilibria, chemical equilibrium is defined by the summation of the product of stoichiometric coefficients and the chemical potentials of each component equating to zero [417].

$$\sum_{r=1}^{N_r} \nu_r \mu_r + \sum_{p=1}^{N_p} \nu_p \mu_p = 0 \tag{4.1}$$

Here, ν is the stoichiometric coefficient of each component participating in the reaction and μ is the chemical potential for each of these components. The subscript r indicates reactants and p denotes reaction products. For mono-molecular reactions, this equation can be simplified as the chemical potentials of the reactant and reaction product become equal:

$$\mu_r = \mu_p \tag{4.2}$$

This leads to a computationally faster method in calculating reaction equilibrium distribution compared to RxMC simulations. This study primarily concentrates on hydroisomerization reactions, which are mono-molecular in nature. The modeling framework explained here remains valid for other heterogeneous mono-molecular reactions. Consider a system of such reactions involving hydrocarbon chains (A_1, A_2, $\cdots$, A_N).

$$A_1 \rightleftharpoons A_2;\ A_2 \rightleftharpoons A_3;\ \cdots;\ A_{N-1} \rightleftharpoons A_N$$

According to Eq. 4.2, we have

$$\mu_{ads,A_1} = \mu_{ads,A_2} = \mu_{ads,A_3} = \cdots = \mu_{ads,A_N} \tag{4.3}$$

This scenario is analogous to a fluid mixture undergoing adsorption, where the gas phase composition reaches chemical equilibrium. By enforcing gas phase equilibrium for a mixture comprising of components (A_1, A_2, $\cdots$, A_N), we obtain

$$\mu_{gas,A_1} = \mu_{gas,A_2} = \mu_{gas,A_3} = \cdots = \mu_{gas,A_N} \tag{4.4}$$

At adsorption equilibrium, the chemical potential of the components in the gas phase and the adsorbed phase are equal, so we have,

$$\begin{aligned} &\mu_{ads,A_1} = \mu_{gas,A_1}; \quad \mu_{ads,A_2} = \mu_{gas,A_2}; \quad \mu_{ads,A_3} = \mu_{gas,A_3}; \quad \cdots; \\ &\mu_{ads,A_N} = \mu_{gas,A_N} \end{aligned} \tag{4.5}$$

Combining Eq. 4.4 with Eq. 4.5 yields Eq. 4.3. This is shown schematically in Fig. 4.1, which shows that the reaction equilibrium distribution in the adsorbed phase can be ascertained by considering a gas-phase composition at chemical equilibrium at a specific temperature and pressure, coupled to the corresponding mixture adsorption isotherms of the individual components involved in the reactions [408, 409]. In the scenario of infinite dilution, the mixture adsorption isotherms can be replaced with Henry's law.

To determine the gas mixture at chemical equilibrium, the chemical potentials for each of these components present in the mixture are required. At low gas phase pressure, the chemical potentials of the components in the gas phase mixture can be approximated as the ideal chemical potential ($\mu_{id,i}$) [413, 418].

$$\mu_{id,i} = RT\ln\left(\frac{\rho_0}{q_{0,i}}\right) + RT\ln\left(\frac{\rho_i}{\rho_0}\right) \tag{4.6}$$

In Eq. 5.3, ρ_0 denotes the reference density which is chosen to be 1 molecule/Å. $q_{0,i}$ is the isolated molecule partition function [413–415] and ρ_i is the number density of the component i. R is the universal gas constant and T is the absolute temperature. The quantity $q_{0,i}$ can be computed using quantum chemical software such as Gaussian [419]. The term $RT\ln(\rho_0/q_{0,i})$ in Eq. 5.3 is the reference chemical potential, $\mu_{ref,i}$. This quantity can be calculated using thermodynamic tables such as Scott's tables [416] and the JANAF tables [420]. $\mu_{ref,i}$, calculated using Scott's tables, equals

$$\mu_{ref,i} = [G_{0,i}(T) - H_{0,i}(T_{ref})] - D_{0,i} \tag{4.7}$$

In Scott's tables, the reference temperature, T_{ref}, is 0 K. In Eq. 4.7, $G_{0,i}(T)$ denotes the standard Gibbs free energy of component i at a specified temperature T. $H_{0,i}(0\ \text{K})$ represents the standard enthalpy of component (i) at 0 K. $D_{0,i}$ is the atomization energy computed using Scott's table as follows.

$$D_{0,i} = \sum_{j=1}^{N_{atoms,i}} a_j \Delta_f H_{0,j}(0\ \text{K}) - \Delta_f H_{0,i}(0\ \text{K}) \tag{4.8}$$

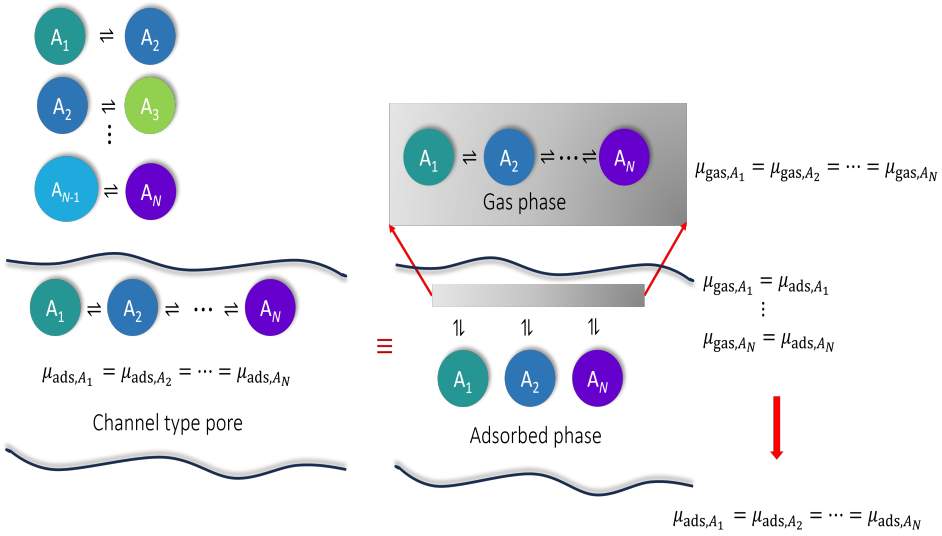


Figure 4.1. Schematic representation of the modeling framework to compute reaction equilibrium loading in the adsorbed phase. Chemical reaction equilibrium is imposed for the system consisting of A_1, A_2,$\cdots$, and A_N-type molecules in the gas phase and phase equilibrium between the gas and the adsorbed phase counter-parts. This establishes chemical equilibrium between the molecules in the adsorbed phase.

In Eq. 4.8, a_j is the number of atoms of type j present in molecule type i. $\Delta_f H_{0,i}$ is the enthalpy of formation of i. $\Delta_f H_{0,i}$ can be obtained from experiemtal data [421, 422], group contribution methods [71, 423], and quantum chemistry calculations [424–427]. For mono-molecular reactions, equating the ideal gas chemical potentials (Eq. 5.3) of the components leads to the ratios of the densities of the reactant and the reaction product in each reaction as shown below.

$$\frac{\rho_1}{\rho_2} = \frac{q_{0,1}}{q_{0,2}} \tag{4.9}$$

The subscripts 1 and 2 in Eq. 4.9 represent the reactant and reaction product in the reaction. For a system of N reacting components, Eq. 4.9 can be extended as follows.

$$\frac{\rho_2}{\rho_1} = \frac{q_{0,2}}{q_{0,1}}; \; \frac{\rho_3}{\rho_1} = \frac{q_{0,3}}{q_{0,1}}; \; \cdots; \; \frac{\rho_N}{\rho_1} = \frac{q_{0,N}}{q_{0,1}} \tag{4.10}$$

The mole fractions in the gas phase, y_i, are calculated using the ratios of the gas phase densities.

$$y_i = \frac{\rho_i}{\sum_{n=1}^{N} \rho_n} \tag{4.11}$$

It is important to note that the error propagation of the gas phase distribution due to uncertainties in the ideal gas chemical potentials of C_7 and C_8 isomers is usually negligible. At infinite dilution, the quantity of molecules at reaction equilibrium in the adsorbed phase is determined using Henry's law,

$$q_{\mathrm{load},i} = k_{\mathrm{H},i} P_{\mathrm{tot}} y_i \tag{4.12}$$

Here, $q_{\mathrm{load},i}$ is the amount of adsorbed loading of component i, $k_{\mathrm{H},i}$ is the Henry coefficient of component i, and P_{tot} is the total gas phase pressure of the reacting mixture. The mole fractions in the adsorbed phase x_i are computed using

$$x_i = \frac{q_{\mathrm{load},i}}{\sum_{n=1}^{N} q_{\mathrm{load},n}} = \frac{k_{\mathrm{H},i} y_i}{\sum_{n=1}^{N} k_{\mathrm{H},n} y_n} = \frac{k_{\mathrm{H},i} \rho_i}{\sum_{n=1}^{N} k_{\mathrm{H},n} \rho_n} \tag{4.13}$$

At higher pressures, the chemical potential of a real gas is considered, which is the sum of the ideal gas chemical potential and an excess chemical potential μ_{ex}:

$$\begin{aligned} \mu_i &= RT \ln\left(\frac{\rho_0}{q_{0,i}}\right) + RT \ln\left(\frac{\rho_i}{\rho_0}\right) + \mu_{\mathrm{ex}} \\ &= RT \ln\left(\frac{\rho_0}{q_{0,i}}\right) + RT \ln\left(\frac{\beta y_i P_{\mathrm{tot}}}{\rho_0}\right) + RT \ln(\phi_i) \end{aligned} \tag{4.14}$$

In Eq. 4.14, P_{tot} is the total pressure of the mixture. ϕ_i is the fugacity coefficient of component i in the gas phase. If the total pressure is not too high (typically < 10 bar), the fugacity coefficients of the components in a mixture can be estimated using the Lewis and Randall rule [428].

$$f_i = \phi_{\mathrm{pure},i} y_i P_{\mathrm{tot}} \tag{4.15}$$

According to the Lewis and Randall rule, the fugacity f_i of a component i in a mixture is equal to the product of the total pressure P_{tot}, mole fraction y_i in the gas phase, and the corresponding pure component fugacity coefficient $\phi_{\mathrm{pure},i}$ at the same temperature and total pressure. Replacing ϕ_i in Eq. 4.14 using Eq. 4.15 leads to

$$\mu_i = RT \ln\left(\frac{\rho_0}{q_{0,i}}\right) + RT \ln\left(\frac{\beta y_i P_{\mathrm{tot}} \phi_{\mathrm{pure},i}}{\rho_0}\right) \tag{4.16}$$

The ratio of the mole fractions of the reactant and the reaction product in a mono-molecular reaction is computed by combining Eqs. 4.2 and 4.16:

$$\frac{y_2}{y_1} = \frac{q_{0,2}}{q_{0,1}} \frac{\phi_{\mathrm{pure},1}}{\phi_{\mathrm{pure},2}} \tag{4.17}$$

The ratio of the densities of the reaction product and the reactant is equal to the ratio of the corresponding mole fractions:

$$\frac{\rho_2}{\rho_1} = \frac{q_{0,2}}{q_{0,1}} \frac{\phi_{\text{pure},1}}{\phi_{\text{pure},2}} \tag{4.18}$$

For a system of N reacting components, Eq. 4.18 can be extended as follows.

$$\frac{\rho_2}{\rho_1} = \frac{q_{0,2}}{q_{0,1}} \frac{\phi_{\text{pure},1}}{\phi_{\text{pure},2}}; \quad \frac{\rho_3}{\rho_1} = \frac{q_{0,3}}{q_{0,1}} \frac{\phi_{\text{pure},1}}{\phi_{\text{pure},3}}; \quad \cdots; \quad \frac{\rho_N}{\rho_1} = \frac{q_{0,N}}{q_{0,1}} \frac{\phi_{\text{pure},1}}{\phi_{\text{pure},N}} \tag{4.19}$$

The ratio of the densities (Eq. 4.18) at higher pressure is expressed as the product of the ratio of the densities at infinite dilution (Eq. 4.9) and the inverse of the ratio of the fugacity coefficients of the pure components. At higher pressures, the equilibrium loadings are calculated using the Ideal Adsorbed Solution Theory (IAST) [51, 54]. The mole fractions in the adsorbed phase, denoted by x_i are calculated using

$$x_i = \frac{q_{\text{load},i}}{\sum_{n=1}^{N} q_{\text{load},i}} \tag{4.20}$$

The selectivity of a component in the gas phase is quantified as [429]:

$$s_i^{\text{gas}} = \frac{y_i}{\left(\sum_{n=1}^{N} y_n\right) - y_i} \tag{4.21}$$

To provide insight into the reaction equilibrium distribution in the adsorbed phase, the reaction selectivity is defined as [429]:

$$s_i^{\text{ads}} = \frac{x_i}{\left(\sum_{n=1}^{N} x_n\right) - x_i} \tag{4.22}$$

The selectivity of a component in both gas phase (Eq. 4.21) and adsorbed phase (Eq. 4.22) is defined as the ratio of the mole fraction of the component to the sum of the mole fractions of all other components present in the same phase.

$$s_{\text{rel},i}^{\text{gas}} = \frac{s_i^{\text{gas}}}{s_{\text{ref}}^{\text{gas}}} \tag{4.23}$$

$$s_{\text{rel},i}^{\text{ads}} = \frac{s_i^{\text{ads}}}{s_{\text{ref}}^{\text{ads}}} \tag{4.24}$$

To compare the selectivity of a component relative to another, the term relative selectivity is defined as the ratio of the absolute selectivity of that component to a reference component as shown in Eqs. 4.23 and 4.24 for the gas $\left(s^{\text{gas}}_{\text{rel},i}\right)$ and the adsorbed phase $\left(s^{\text{ads}}_{\text{rel},i}\right)$ respectively. In this study, n-C_7 and n-C_8 molecules are chosen as reference components for computing relative selectivities.

4.3. Simulation Details

To determine the ratios of the densities of the isomers at chemical equilibrium, the reference chemical potential $\mu_{\text{ref},i}$ (Eq. 4.7) is a crucial parameter. To compute this, thermodynamic tables by Scott [416] are used. The detailed raw data sets essential for these calculations, which are obtained from these tables, are available in the Supporting Information (SI2.xlsx) of Ref. [390] for both C_7 and C_8 isomers.

At infinite dilution, the adsorbed loadings are calculated using Henry's law (Eq. 4.12). The equilibrium loading of each isomer is proportional to the corresponding Henry coefficient k_{H}. Henry coefficients are computed using the Widom test particle insertion method [46] combined with the Configurational-Bias Monte Carlo (CBMC) method [47–49]. Alkanes are modelled using the united-atom model [430]. This model is found to be accurate while being computationally cheap [37] because the united-atoms (C, CH, CH_2, and CH_3) are charge-neutral and the Coulomb interactions can be omitted [39]. TraPPE united-atom force field [37] is used for the intramolecular bonded interactions, which include bond-stretching, bond-bending, and torsion interactions. All parameters necessary for intramolecular bonded potentials are listed in Appendix A.2 (tables A.3 and A.4). The non-bonded interactions between the adsorbent and the adsorbate, as well as the intramolecular non-bonded interactions, are modelled using Lennard-Jones interactions [40]. For non-bonded inter-adsorbate interactions, the Lennard-Jones parameters for each component are obtained from Dubbeldam *et al.* [136], which are listed in Appendix A.2 (tables A.1). The TraPPE-zeo force field [43] is used for the non-bonded interactions between the adsorbate and the adsorbent. All silica zeolites are considered as rigid structures [44]. The influence of zeolite flexibility is negligible on adsorption processes, especially at infinite dilution [44]. The Lennard-Jones parameters for zeolite atoms are listed in Appendix A.2 (table A.2). Lennard-Jones interactions are truncated and shifted at 12 Å without applying tail corrections. To account for these interactions between different types of atoms, the Lorentz-Berthelot [41, 42] mixing rules are used. The force field parameters used in this study are extensively validated and can perfectly reproduce experimental adsorption results [43, 136]. Table 4.1 lists the number of unit cells in the simulation box for each zeolite, along with the respective dimensions and the void fractions of these zeolites. The results for the hydroisomerization of C_7 isomers in these zeolites are included in Section 4.4, while the results pertaining to

C_8 isomers are presented in Appendix A.2.

To calculate the ratios of the gas phase densities of the isomers at high pressures, the computations require pure component fugacity coefficients corresponding to the total pressure of the mixture (Eq. 4.19). These values are determined using the Peng-Robinson equation of state [431] which requires the magnitudes of critical temperature, critical pressure, and acentric factors as input. These critical parameters are listed in the worksheet Critical_consts of the Supporting Information (SI2.xlsx) of Ref. [390]. The adsorbed phase equilibrium loadings at high pressures are calculated using Ideal Adsorbed Solution Theory (IAST). Pure component adsorbed loadings at various pressures (10^1 Pa - 10^8 Pa) are computed using the Grand-Canonical Monte Carlo (GCMC) simulations coupled with the CBMC-algorithm using the RASPA software [39, 133]. For details on computing adsorption isotherms using GCMC simulations, the reader is referred to Refs. [432, 433]. The pure component adsorbed loadings are fitted to the single-site Langmuir-type isotherm equation. The fitted parameters (Table 4.2) are used as input in calculating adsorbed phase loadings of mixtures. The effect of higher pressure on the adsorption of C_7 isomers in MTW-type zeolite is studied using the same number of unit cells for the simulation box as specified in Table 4.1. The pure component isotherm parameters for different isomers in MTW-type zeolite at 500 K are shown in Table 4.2.

4.4. Results and Discussions

The considered zeolites exhibit diverse pore shapes and sizes, such as cage-like structures (FAU- and ITQ-type zeolites), channels with intersections (MEL-, MFI-, and BEA-type zeolites), and one-dimensional channel-type zeolites (MRE- and MTW-type zeolites). We extensively analyse shape-selectivity effects of the zeolites at infinite dilution for C_7 isomers. Additionally, the effect of pressure on the equilibrium distribution of C_7 isomers inside MTW-type zeolite has been studied. The temperature was varied from 400 K to 700 K. The effects of FAU-, ITQ-29-, BEA-, MEL-, MFI-, MTW-, and MRE-type zeolites on the hydroisomerization of C_7 isomers are discussed. The shape-selectivity effects on hydroisomerization of C_8 isomers are also studied. The results pertaining to C_8 isomers are shown in Figs. A.5-A.19 in Appendix A.2. Additionally, the numerical values of the thermochemical properties of all isomers ranging from C_7 to C_8 isomers listed by Scott [416] are included in the Supporting Information (SI2.xlsx) of Ref. [390].

At infinite dilution, the equilibrium distribution of reactions in the adsorbed phase is governed by the interplay between the gas phase distribution and the Henry coefficients. In Fig. 4.2, the selectivities of C_7 isomers relative to n-C_7 at gas phase equilibrium are shown for temperatures ranging from 400 K to 700 K. 2-m-C_6 and 3-m-C_6 exhibit the highest selectivity for all isomers. There is a temperature-dependent shift favoring the selectivity of 3-m-C_6 over 2-m-C_6 at higher temperatures. 3-e-C_5 displays the smallest

Table 4.1. Total number of unit cells for different zeolites used in the simulations to compute the Henry coefficients of C_7 and C_8 isomers, along with the dimensions and void fractions of these zeolites [18].

zeolite	number of unit cells	unit cell dimension/ [Å]			void fraction/ [-]
		a	*b*	*c*	
FAU-type	8 (2×2×2)	24.35	24.35	24.35	0.49
ITQ-29-type	27 (3×3×3)	11.92	11.92	11.92	0.403
BEA-type	4 (2×2×1)	12.66	12.66	26.41	0.41
MEL-type	8 (2×2×2)	20.27	20.27	13.46	0.28
MFI-type	8 (2×2×2)	20.09	19.74	13.14	0.29
MTW-type	36 (2×9×2)	25.55	5.26	12.12	0.21
MRE-type	20 (5×2×2)	8.26	14.56	20.31	0.16

Table 4.2. Pure component Langmuir-type isotherm parameters for adsorption of C_7 isomers in MTW-type zeolite at 500 K obtained by fitting to the results from CBMC simulations in the Grand-Canonical ensemble.

isomer	saturation loading/ [mol/(kg zeolite)]	Langmuir constant/ [1/Pa]
n-C_7	0.471	1.12×10^{-3}
2-m-C_6	0.476	4.97×10^{-4}
3-m-C_6	0.479	1.99×10^{-4}
3-e-C_5	0.620	9.16×10^{-6}
2,2-m-C_5	0.496	4.96×10^{-6}
2,3-m-C_5	0.647	3.09×10^{-5}
2,4-m-C_5	0.644	4.74×10^{-5}
3,3-m-C_5	0.504	2.80×10^{-6}
2,2,3-m-C_4	0.486	1.20×10^{-6}

selectivity for the mono-branched isomers. At 400 K, 2,2-m-C_5 has the highest selectivity for di-branched molecules. At elevated temperatures (600 K and 700 K), 2,3-m-C_5 has the largest mole fraction for all di-branched isomers. The selectivity of 2,4-m-C_5 decreases relative to 2,3-m-C_5 with increasing temperatures. For the di-branched isomers, 3,3-m-C_5 is the least favored molecule. 2,2,3-m-C_4 has the smallest selectivity compared to the other isomers.

Fig. 4.3 shows the Henry coefficients for C_7 isomers in FAU- and ITQ-29-type zeolites for a temperature range of 400 K to 700 K. FAU-type zeolite shows small variations in adsorption selectivity of C_7 isomers. This can be attributed to its large pore diameter $(7.4\ \text{Å})$ [18]. The geminal alkanes (2,2-m-C_5 and 3,3-m-C_5) have the lowest Henry coefficients for di-branched isomers due to the steric hindrance posed by the close proximity of two

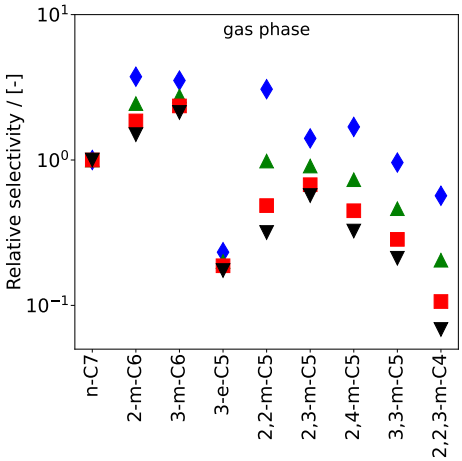


Figure 4.2. Relative selectivities of different C_7 isomers at reaction equilibrium in the gas phase at infinite dilution. The reaction equilibrium distribution is computed at ♦ 400 K, ▲ 500 K, ■ 600 K, and ▼ 700 K. The raw data is listed in the worksheet lnq_iC7 of the Supporting Information (SI2.xlsx file) of Ref. [390].

methyl groups during adsorption. ITQ-29- and FAU-type zeolites show similar affinity for adsorption because of the cage-like pore structures present in these zeolites. Fig. 4.3b shows an increase in the Henry coefficients with an increasing degree of branching in ITQ-29-type zeolite. Such variations are also present in FAU-type zeolite, but the larger pore size of this zeolite significantly mitigates this effect. 2,2-m-C_5 and 3,3-m-C_5 are the least favored di-branched isomers in both ITQ-29 and FAU-type zeolite environments.

BEA-type (Fig. 4.4a) and MEL-type (Fig. 4.4b) zeolites exhibit strikingly similar trends in the variation of Henry coefficients for C_7 isomers. These three-dimensional zeolite structures feature straight channel-type pores with intersections. The channels present in BEA-type zeolite (7.6 × 6.4 and 5.5 × 5.5 Å) [434] have larger diameters than those in MEL-type zeolite (5.3 × 5.4 Å) [18]. An increase in the degree of branching leads to a decrease in Henry coefficients in both zeolites. n-C_7 is the most favored isomer in the channel-type pores of BEA- and MEL-type zeolites. 3-e-C_5, characterized by the presence of an ethyl group, has the lowest Henry coefficient compared to the other mono-branched isomers. 2,2-m-C_5, 3,3-m-C_5, and 2,2,3-m-C_4 are the least adsorbed isomers in both zeolites. The proximity of multiple methyl branches in these isomers reduces the preference for adsorption inside the channel-type pores. 2,4-m-C_5 is the di-branched isomer with the highest affinity for adsorption in both BEA- and MEL-type zeolites. This is attributed to the considerable separation between the two methyl

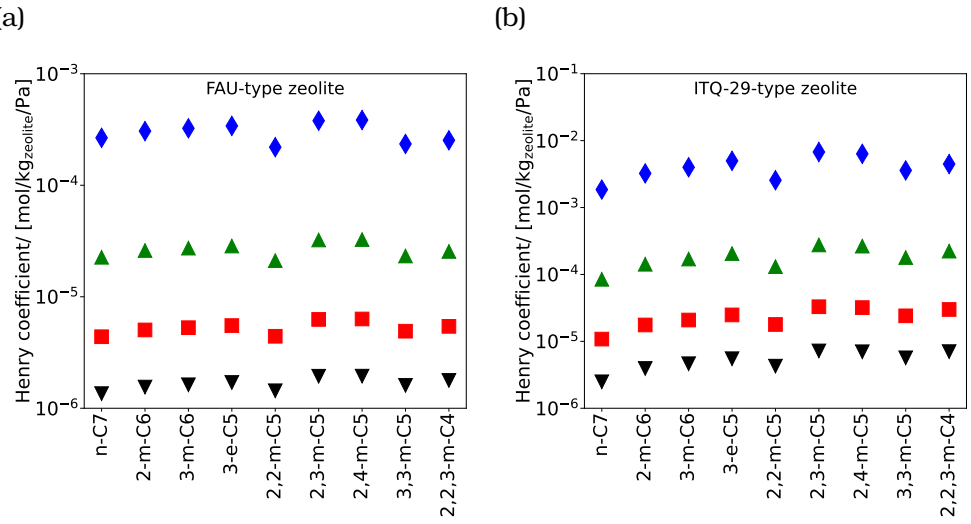


Figure 4.3. Henry coefficients of C_7 isomers in (a) FAU- and (b) ITQ-29-type zeolites at ♦ 400 K, ▲ 500 K, ■ 600 K, and ▼ 700 K. The raw data is listed in the worksheets xi_iC7_400K, xi_iC7_500K, xi_iC7_600K, and xi_iC7_700K of the Supporting Information (SI2.xlsx file) of Ref. [390].

groups in this particular isomer. In Fig. 4.4c, the Henry coefficients for C_7 isomers in MFI-type zeolites are listed. MFI-type zeolite is also a three-dimensional zeolite structure consisting of channel-like pores connected by intersections [18]. Unlike BEA- and MEL-type zeolites, MFI-type zeolite is characterized by both sinusoidal and straight channels [18]. Branched isomers preferentially adsorb at channel intersections of MFI-type zeolite [49, 124], while linear alkanes (n-C_7) lack such preferences [49, 124]. Similar to BEA- and MEL-type zeolites, MFI-type zeolite also displays decreasing adsorption affinity with an increasing degree of branching. Nevertheless, the preferences for di-branched isomers differ, favoring 2,2-m-C_5 due to the optimal fit of its two methyl groups at channel intersections (Fig. 4.5a). As the separation between the methyl groups expands (2,3-m-C_5 and 2,4-m-C_5), the probability of either branch attempting to fit into the channel-type pores increases (Figs. 4.5b and 4.5c), resulting in an unfavorable scenario. This is because of the smaller diameter of the channels compared to the intersections. 3,3-m-C_5 also exhibits preferential adsorption in the MFI-type zeolite, but with a lower Henry coefficient compared to 2,2-m-C_5, which can be attributed to the higher steric hindrance experienced by 3,3-m-C_5.

Figs. 4.6a and 4.6b show the Henry coefficients for various C_7 isomers in MTW- and MRE-type zeolites, which feature one-dimensional channel-like pore structures. An overall reduction in the magnitude of Henry coefficients

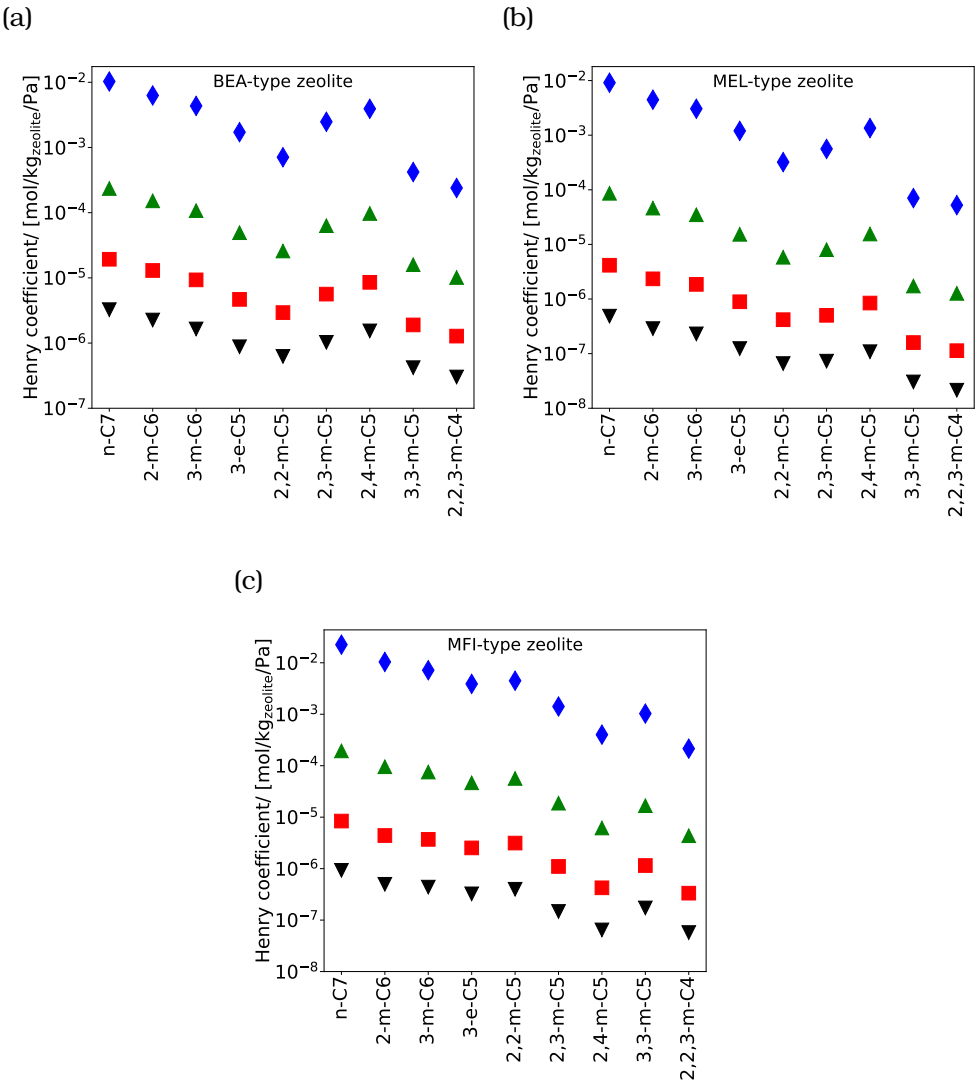


Figure 4.4. Henry coefficients of C_7 isomers in (a) BEA-, (b) MEL-, (c) MFI-type zeolites at ♦ 400 K, ▲ 500 K, ■ 600 K, and ▼ 700 K. The raw data is listed in the worksheets xi_iC7_400K, xi_iC7_500K, xi_iC7_600K, and xi_iC7_700K of the Supporting Information (SI2.xlsx file) of Ref. [390].

is observed with an increasing degree of branching in both MTW- and MRE-type zeolites. This is consistent with other zeolites with channel-like pore structures (BEA-, MEL-, and MFI-type zeolites). The decrease is sharper

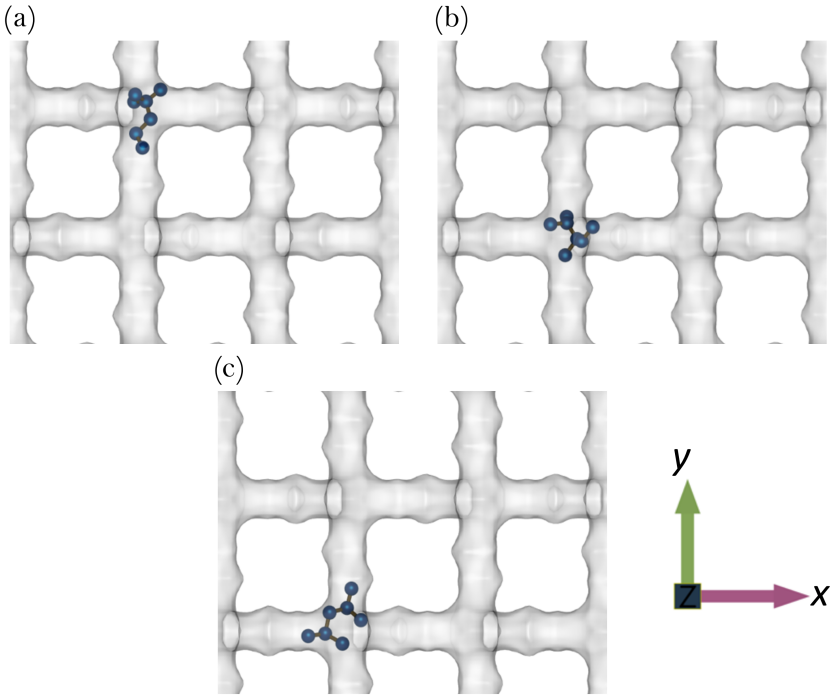


Figure 4.5. Typical snapshots of adsorption of (a) 2,2-m-C_5, (b) 2,3-m-C_5, and (c) 2,4-m-C_5 isomers in MFI-type zeolite at infinite dilution and 500 K. These snapshots are obtained from Monte Carlo simulations of a single molecule inside MFI-type zeolite. In these snapshots, the sinusoidal-shaped channels of MFI-type zeolite are located in the horizontal direction, and the straight channels are located in the vertical direction. The iRASPA software [22] was used to generate these images. For di-branched alkanes, MFI favors 2,2-m-C_5 compared to 2,3-m-C_5 and 2,4-m-C_5 because the two methyl groups in 2,2-m-C_5 optimally fit at the channel intersections. As the methyl groups are spaced further apart for 2,3-m-C_5 and 2,4-m-C_5, one methyl group increasingly enters the narrow channels, leading to less favorable configurations.

than that observed in BEA-, MEL-, and MFI-type zeolites, due to smaller pore diameters and the absence of channel intersections. Inside both MTW- and MRE-type zeolites, 3-e-C_5 is the least favored mono-branched isomer. The variations in Henry coefficients in MTW-type zeolite closely resemble those observed in BEA- and MEL-type zeolites. 2,2-m-C_5 and 3,3-m-C_5 are the least preferred di-branched isomers due to steric hindrance posed by the close proximity of the two methyl branches. Similar to MTW-type zeolite, MRE-type zeolite also exhibits decreasing Henry coefficients with

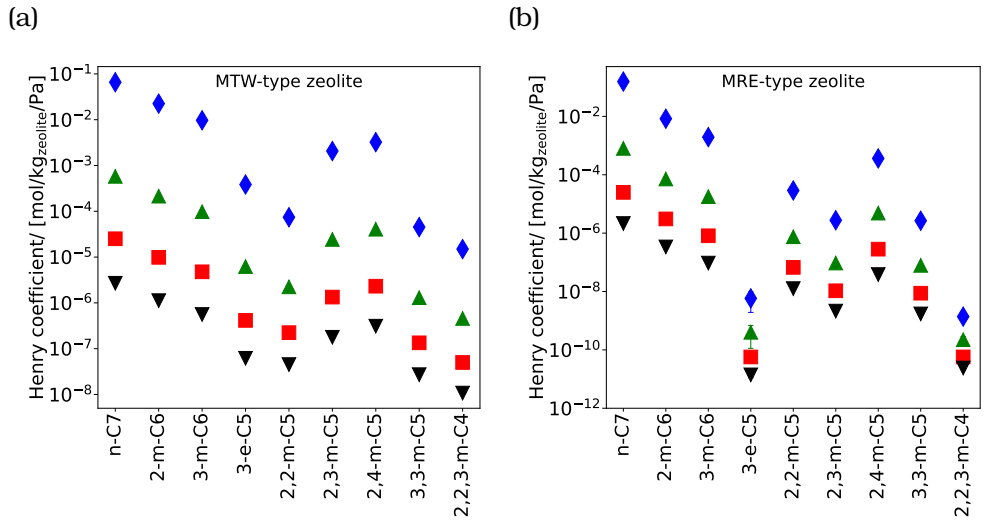


Figure 4.6. Henry coefficients of C_7 isomers in (a) MTW- and (b) MRE-type zeolites at ♦ 400 K, ▲ 500 K, ■ 600 K, and ▼ 700 K. The raw data is listed in the worksheets xi_iC7_400K, xi_iC7_500K, xi_iC7_600K, and xi_iC7_700K of the Supporting Information (SI2.xlsx file) of Ref. [390].

increasing branching. Larger variations in magnitudes of Henry coefficients can be observed in MRE-type zeolite compared to MTW-type zeolite. Inside MRE-type zeolite, 2,3-m-C_5 is less favored than the 2,2-m-C_5 isomer, which is attributed to the corrugations present in the channel-like pores. These corrugations are formed by a combination of peaks and crests, which create regions with varying channel diameters. The higher probability of accommodating both methyl groups in a region with a peak favors the preferential adsorption of 2,2-m-C_5 (Fig. 4.7a) compared to 2,3-m-C_5 (Fig. 4.7b) in MRE-type zeolite. 2,4-m-C_5 (Fig. 4.7c) is the most favored di-branched isomer because of the possibility to fit the two methyl branches in the regions where the peaks of the channel corrugations are present.

Adsorption of multi-branched isomers is favored in zeolites with cage-like structures (ITQ-29-type zeolite). Channel-like pore structures, particularly one-dimensional ones without intersections, such as in MTW- and MRE-type zeolites, favor the adsorption of linear and mono-branched isomers. The temperature-dependent variations in the Henry coefficients are similar in all zeolites considered in this study. Each isomer shows a reduced affinity for adsorption at elevated temperatures. Apart from the Henry coefficients, the equilibrium distribution of hydroisomerization reactions in the adsorbed phase is also influenced by the mole fractions of the isomers in the gas phase (Eq. 4.13). At infinite dilution, the gas phase distribution relies solely on

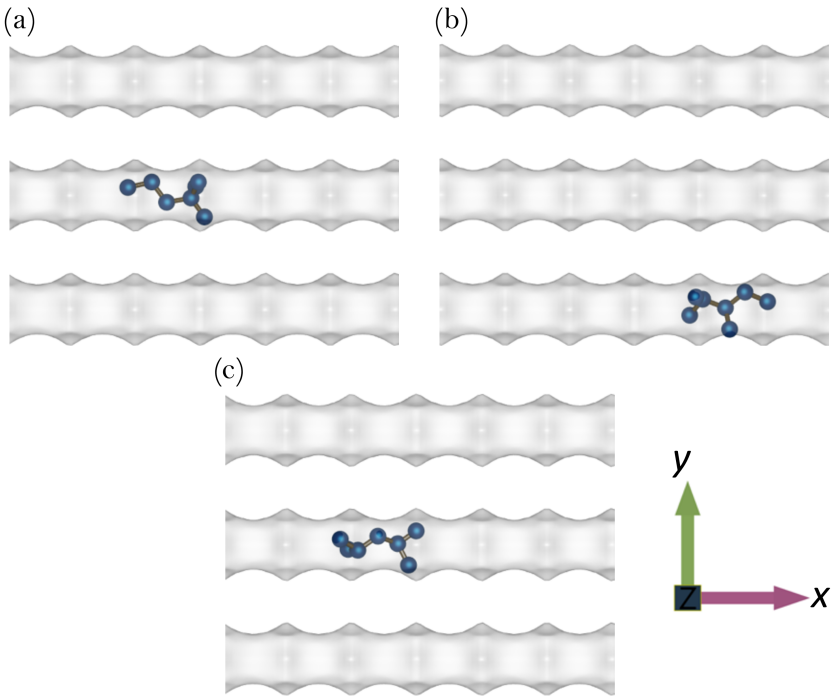


Figure 4.7. Typical snapshots of adsorption of (a) 2,2-m-C_5, (b) 2,3-m-C_5, and (c) 2,4-m-C_5 isomers in MRE-type zeolite at infinite dilution condition and 500 K. These snapshots are obtained from Monte Carlo simulations of a single molecule inside MRE-type zeolite. In these snapshots, one dimensional channels of MRE-type zeolites are shown in the horizontal direction. The iRASPA software [22] was used to generate these images. In MRE-type zeolites, 2,2-m-C_5 is preferentially adsorbed compared to 2,3-m-C_5 because the methyl groups of 2,2-m-C_5 are more likely to fit near a corrugation peak. For 2,3-m-C_5, there is a high probability for one of the methyl groups to be present at one of the crests of the channel corrugations. 2,4-m-C_5 is most favored, as its widely spaced branches offer less steric hindrance and these branches align best with the corrugation peaks.

the ideal gas reference chemical potential or the isolated molecule partition function (Eq. 5.3). The equilibrium distribution in the adsorbed phase for a system of mono-molecular reactions is determined by the combined influences of the gas phase selectivity and the Henry coefficients. For FAU-type zeolite, the variations in the adsorbed phase selectivity are similar to those observed in the gas phase selectivity, as shown in Fig. 4.8a. This is due to a lack of variation in the magnitudes of the Henry coefficients of C_7 isomers in FAU-type zeolite, irrespective of the degree of branching of

these isomers. This is because of the larger pore size in FAU-type zeolite. Similar variations in selectivities can be observed in ITQ-29-type zeolite to a certain extent (Fig. 4.8b. The preference for 2,2-m-C_5 is consistently lower than that for 2,3-m-C_5 and 2,4-m-C_5 in this zeolite, despite 2,2-m-C_5 being favored over the other two isomers in the gas phase at 400 K. This discrepancy is attributed to the larger influence of the small magnitudes of Henry coefficient compared to the corresponding gas phase selectivity. This is caused by the smaller channel diameters which connect the cage-type pores in ITQ-29-type zeolite (4.1 Å) [18] compared to FAU-type zeolite (7.4 Å) [18]. At temperatures of 500 K and above, 2,3-m-C_5 exhibits a higher selectivity than 2,4-m-C_5 in both FAU and ITQ-29-type zeolites. This is favored by both gas phase selectivity and Henry coefficients. ITQ-29- and FAU-type zeolites show lower selectivity for 2,2,3-m-C_4 compared to the other isomers. Although 2,2,3-m-C_4 preferentially adsorbs in these zeolites, the lower gas phase selectivity plays a dominant role in determining its fraction in the adsorbed phase at reaction equilibrium.

We examine reaction equilibrium distributions of C_7 isomers in zeolites featuring three-dimensional channel-like pore structures. The selectivities of C_7 isomers relative to n-C_7 in BEA-type zeolite show an overall decrease with increasing degree of branching (Fig. 4.9a). This is influenced by the variations in Henry coefficients of the C_7 isomers in BEA-type zeolite. At 400K, 2-m-C_6 is favored over n-C_7 and 3mC_6. As temperature increases, the selectivities of n-C_7, 2-m-C_6, and 3-m-C_6 become nearly equal. 3-e-C_5 is the least favored mono-branched isomer due to low gas phase selectivity and a lower Henry coefficient. At low temperatures (500 K and lower), 2,4-m-C_5 is the most preferred di-branched isomer primarily due to a higher Henry coefficient. At elevated temperatures (600 K and 700 K), 2,4-m-C_5 and 2,3-m-C_5 exhibit almost identical selectivities. 3,3-m-C_5 has the lowest selectivity compared to the other di-branched isomers, attributed to the steric hindrances leading to a reduced adsorption affinity. The variations in the selectivities in MEL-type zeolite (Fig. 4.9b) are similar to those in BEA-type zeolite. 22mC_5 shows a higher selectivity than 2,3-m-C_5 at 400 K. At higher temperatures, the difference between the Henry coefficients of these two isomers becomes less significant. The higher selectivity of 2,3-m-C_5 compared to 2,2-m-C_5 is mainly caused by the gas phase selectivity. Similar to BEA-type zeolite, MEL-type zeolite also favors 2,4-m-C_5 as the most preferred di-branched isomer at lower temperatures (400 K and 500 K). At elevated temperatures (700 K), 2,4-m-C_5 and 2,3-m-C_5 show nearly identical selectivities because these isomers have Henry coefficients of comparable magnitudes. This is favored by the slightly lower gas phase selectivity of 2,4-m-C_5 compared to 2,3-m-C_5 at temperatures higher than or equal to 500 K. Selectivities of C_7 isomers relative to n-C_7 in MFI-type zeolite also show a decreasing trend in the reaction equilibrium distribution with an increasing degree of branching (Fig. 4.9c). The selectivity of di-branched isomers is highly influenced by Henry coefficients in MFI-type zeolite. 2,2-

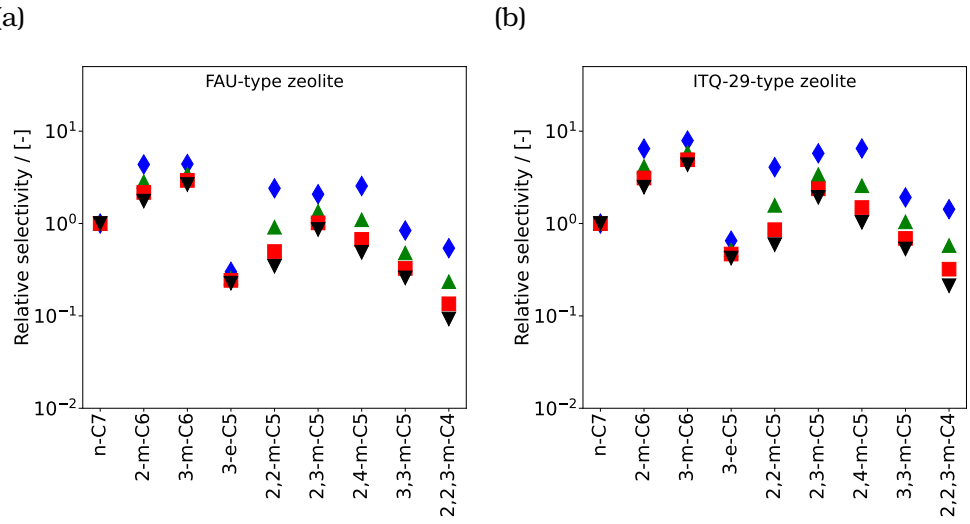


Figure 4.8. Relative selectivities of different C_7 isomers at reaction equilibrium in (a) FAU-, (b) ITQ-29-type zeolites at infinite dilution. The reaction equilibrium distribution is computed at ♦ 400 K, ▲ 500 K, ■ 600 K, and ▼ 700 K. The raw data is listed in the worksheets xi_iC7_400K, xi_iC7_500K, xi_iC7_600K, and xi_iC7_700K of the Supporting Information (SI2.xlsx file) of Ref. [390].

m-C_5 is the most favored and 2,4-m-C_5 the least favored isomers. This aligns closely with the Henry coefficients of these two isomers in MFI-type zeolite. The two methyl groups in 2,2-m-C_5 can be optimally accommodated at the intersections of the channel-type pores in MFI-type zeolite. The increasing probability of accommodating one of the methyl groups of 2,4-m-C_5 inside the channels makes the adsorption of this isomer less favorable than 2,2-m-C_5.

Figs. 4.10a and 4.10b show the relative selectivities of C_7 isomers in MTW and MRE-type zeolites. In MTW-type zeolite, the selectivity decreases from n-C_7 to 3mC_6 at temperatures larger than or equal to 500 K. The decreasing trend in the Henry coefficients of these isomers plays a dominant role in deciding the corresponding selectivities. 2,2-m-C_5 has a lower Henry coefficient compared to 3-e-C_5 at all temperatures. The low gas phase selectivity for 3-e-C_5 influences its lower selectivity in the adsorbed phase compared to 2,2-m-C_5. Both the Henry coefficient and gas phase selectivity favor 2,4-m-C_5 compared to other di-branched isomers at 400 K in MTW-type zeolite. The selectivity drops at higher temperatures because of the decrease in the gas phase selectivity compared to 2,3-m-C_5. The relative selectivities in MRE-type zeolite are similar to the corresponding variations in Henry

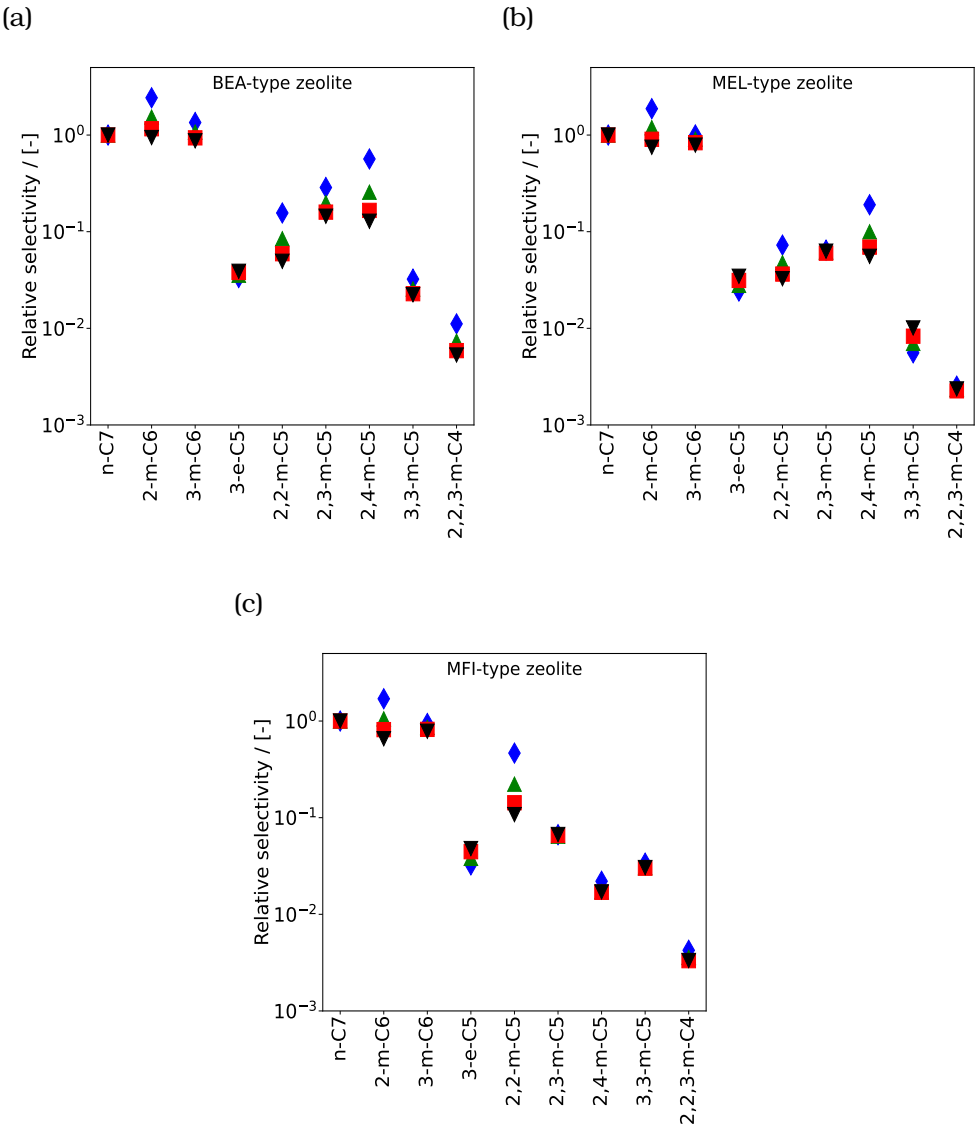


Figure 4.9. Relative selectivities of different C_7 isomers at reaction equilibrium in (a) BEA-, (b) MEL-, (c) MFI-type zeolites at infinite dilution. The reaction equilibrium distribution is computed at ♦ 400 K, ▲ 500 K, ■ 600 K, and ▼ 700 K. The raw data is listed in the worksheets xi_iC7_400K, xi_iC7_500K, xi_iC7_600K, and xi_iC7_700K of the Supporting Information (SI2.xlsx file) of Ref. [390].

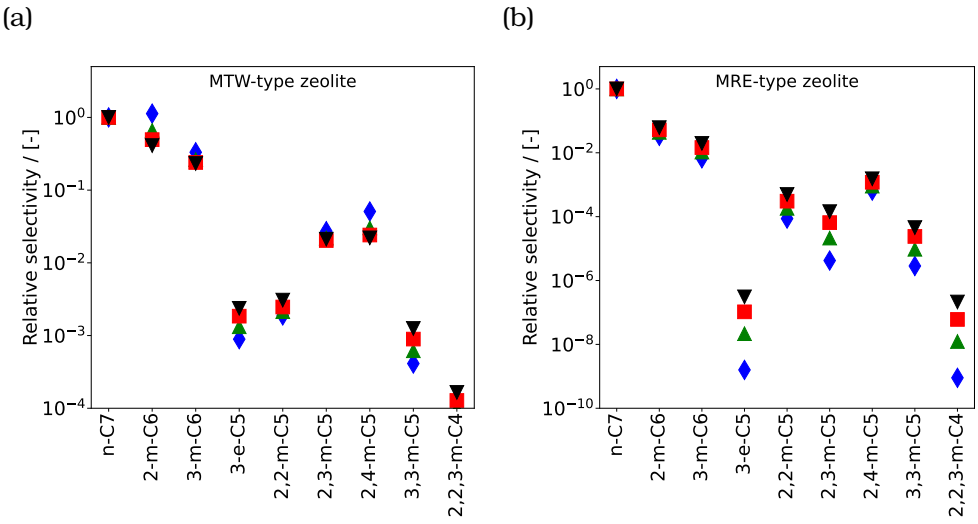


Figure 4.10. Relative selectivities of different C_7 isomers at reaction equilibrium in (a) MTW- and (b) MRE-type zeolites at infinite dilution. The reaction equilibrium distribution is computed at ♦ 400 K, ▲ 500 K, ■ 600 K, and ▼ 700 K. The raw data is listed in the worksheets xi_iC7_400K, xi_iC7_500K, xi_iC7_600K, and xi_iC7_700K of the Supporting Information (SI2.xlsx file) of Ref. [390].

coefficients. The smaller pore diameters in MRE-type zeolite lead to a larger influence of Henry coefficients compared to the gas phase distribution at chemical equilibrium.

The effects of Henry coefficients on adsorbed phase equilibrium distribution are not prominent in zeolites with larger pore sizes such as FAU-type zeolite. While multi-branched isomers are preferably adsorbed in ITQ-29-type zeolite, such influence on the adsorbed phase reaction equilibrium distribution decreases due to reduced gas phase selectivity. All di-branched isomers exhibit higher selectivities compared to n-C_7 in ITQ-29-type zeolite. In BEA-, MEL-, and MFI-type zeolites, n-C_7, 2-m-C_6, and 3-m-C_6 have very similar selectivities. Additionally, the mono-branched isomers except for 3-e-C_5 have higher selectivity compared to the di-branched isomers in these zeolites. In zeolites with narrow channel-like pores such as MRE- and MTW-type zeolites, the affinity of C_7 isomers for adsorption plays an important role in determining adsorbed phase selectivity at reaction equilibrium. MTW- and MRE-type zeolites favor linear isomers (n-C_7) the most. 2,2,3-m-C_4 is the least favored isomer in all zeolites considered in this study. In this study, it is observed that the variations in selectivities between the linear and the branched isomers are significantly influenced by the pore diameter of the

zeolite. The channel distributions and the shape of the channels play an important role in determining the differences in the selectivities of isomers with similar degree of branching. To quantify the impact of uncertainties in gas phase thermochemical properties and Henry coefficients on reaction equilibrium distributions, a Monte Carlo-based uncertainty propagation was performed (Appendix A.2.2). The results show that the predicted mole fractions and selectivities are robust, with Gaussian-like distributions centered around the actual values. A detailed description of this procedure is provided in Appendix A.2.2.

The equilibrium distribution of C_7 isomers in BEA-type zeolite is compared to the results from experiments conducted by Agarwal et al [20] (Fig. 4.11). These authors have shown a complete reaction network of hydroisomerization of C_7 isomers and possible hydrocracking reactions. It is important to note that the computation of reaction product distribution at chemical equilibrium does not include any cracked products. This is because of the irreversible nature of the cracking reactions. At temperatures smaller than 500 K, negligible conversion of n-C_7 takes place (ca. 3%) [20]. There is very high conversion at temperatures above 581.55 K and the reaction product distribution is mainly composed of cracked products (2-m-C_3 and C_3 molecules) [20]. At intermediate temperatures in the range 543.95 K to 569.02 K, the reaction product distribution is comparable to the molar composition at reaction equilibrium computed in this study (Fig. 4.11). The distribution obtained from experiments at 556.48 K is closest to the equilibrium distribution. This indicates that at medium conversion (543.95 K to 569.02 K), thermodynamic equilibrium is approached for isomers with a similar degree of branching [33].

Fig. 4.12 shows the distribution of gas phase reaction equilibrium for the hydroisomerization of C_7 isomers at finite loadings. The gas phase distribution remains relatively constant with increasing pressure. Specifically, at 400 K, deviations from infinite dilution conditions are observed at ca. 3 bar, which is due to phase transition. As temperature increases, these deviations become less pronounced. At 500 K, the gas phase distribution deviates from ideal gas behavior at ca. 10 bar, but at higher temperatures (600 K and 700 K), these discrepancies become nearly negligible (Figs. 4.12c and 4.12d). Apart from the reference chemical potential (Eq. 5.3), the gas phase distribution also depends on the ratios of the corresponding fugacity coefficients. The pure component fugacity coefficients of C_7 isomers at temperatures ranging from 400 K, to 500 K and pressures ranging from 1×10^1 Pa to 3×10^6 Pa are listed in the worksheet Fugacity_iC7 of the Supporting Information SI2.xlsx [390]. There are negligible variations in fugacity coefficients of C_7 isomers at temperatures (500 K or above). Therefore, the fugacity ratios remain close to unity. This leads to negligible deviations in the gas phase distribution with variations in pressure.

The adsorbed phase equilibrium distribution is influenced by the adsorption affinity of the isomers towards the adsorbate. Quantitatively, this can

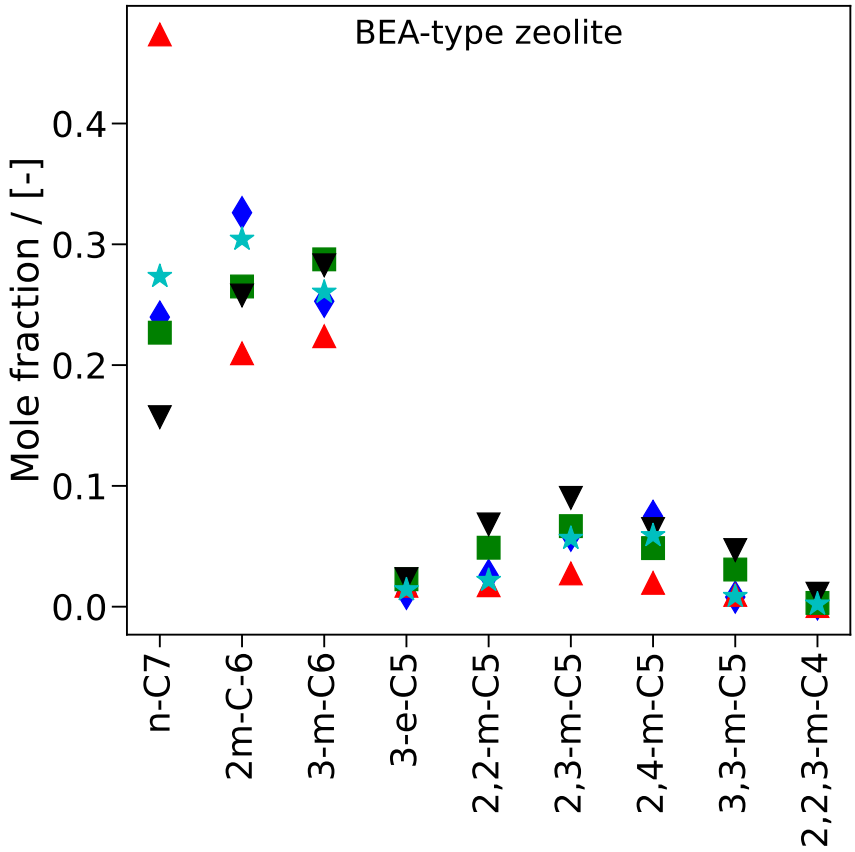


Figure 4.11. Analysis of reaction product distribution of C_7 isomers in BEA-type zeolite obtained from experiments by Agarwal et al [20] (▲543.95 K, ■556.48 K, and ▼569.02 K) and computed at chemical equilibrium in this work (♦500 K and ★600 K). Cracking reactions are not included in computing reaction equilibrium distribution because of the irreversible nature of these reactions. The mole fractions of the reaction products at temperatures (▲543.95 K, ■556.48 K, and ▼569.02 K) are also calculated without including the cracked products for comparing with the reaction equilibrium distributions (♦500 K and ★600 K).

be understood from the pure component adsorption isotherm parameters, which are used as input for computing mixture adsorption isotherms. Fig. 4.13a shows the pure component adsorption isotherms of C_7 isomers in MTW-type zeolite at 500 K. The isomers with a longer main chain (e.g., n-C_7) are preferred at low pressures. At high pressures, the shorter chains are pre-

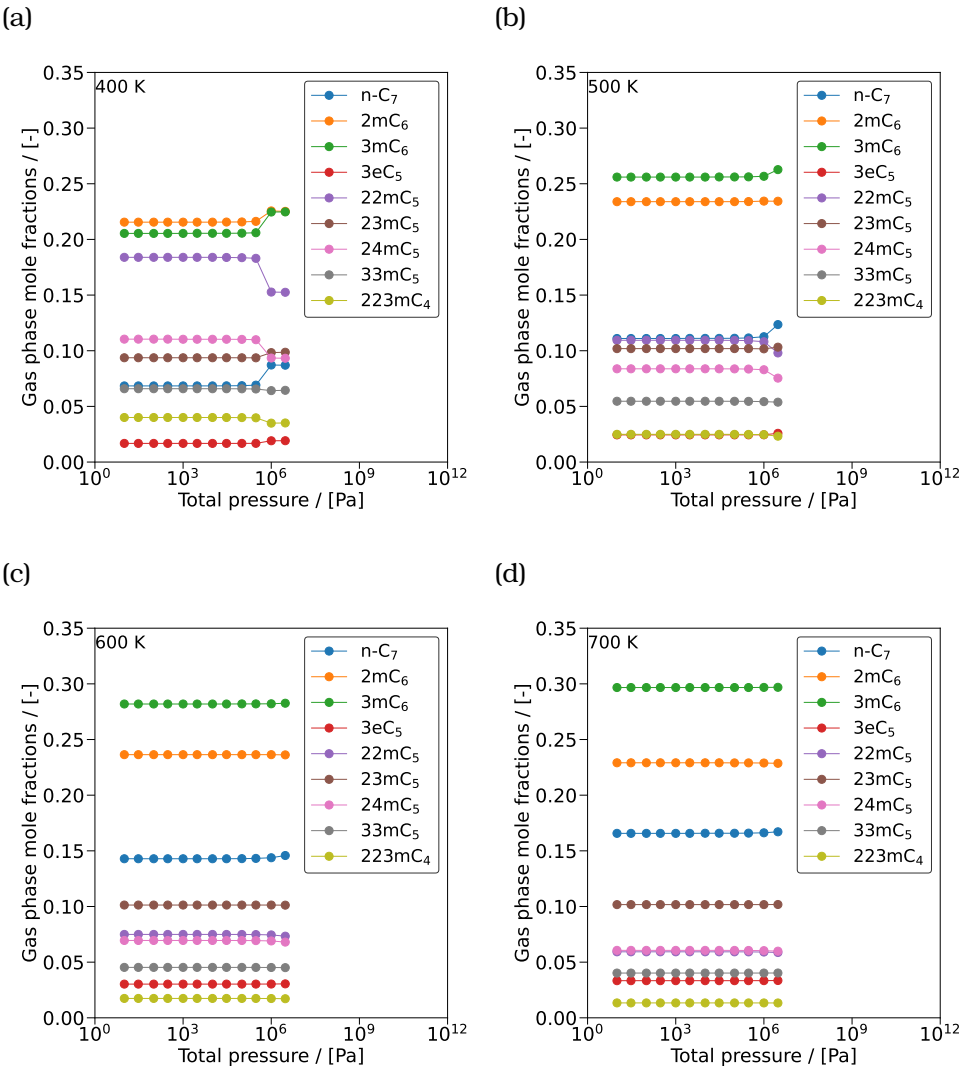


Figure 4.12. Effect of gas phase pressure on the equilibrium distribution of C_7 isomers at temperatures (a) 400 K, (b) 500 K, (c) 600 K, and (d) 700 K. The raw data is listed in the worksheet lnQ_iC7_high_press of the Supporting Information (SI2.xlsx file) of Ref. [390].

ferred because of a lack of available space for adsorption. This phenomenon is known as size entropy [119–121]. Mixture adsorption isotherms for C_7 isomers in MTW-type zeolite at 500 K are shown in Fig. 4.13b. The corresponding gas phase composition is considered to be in chemical equilibrium

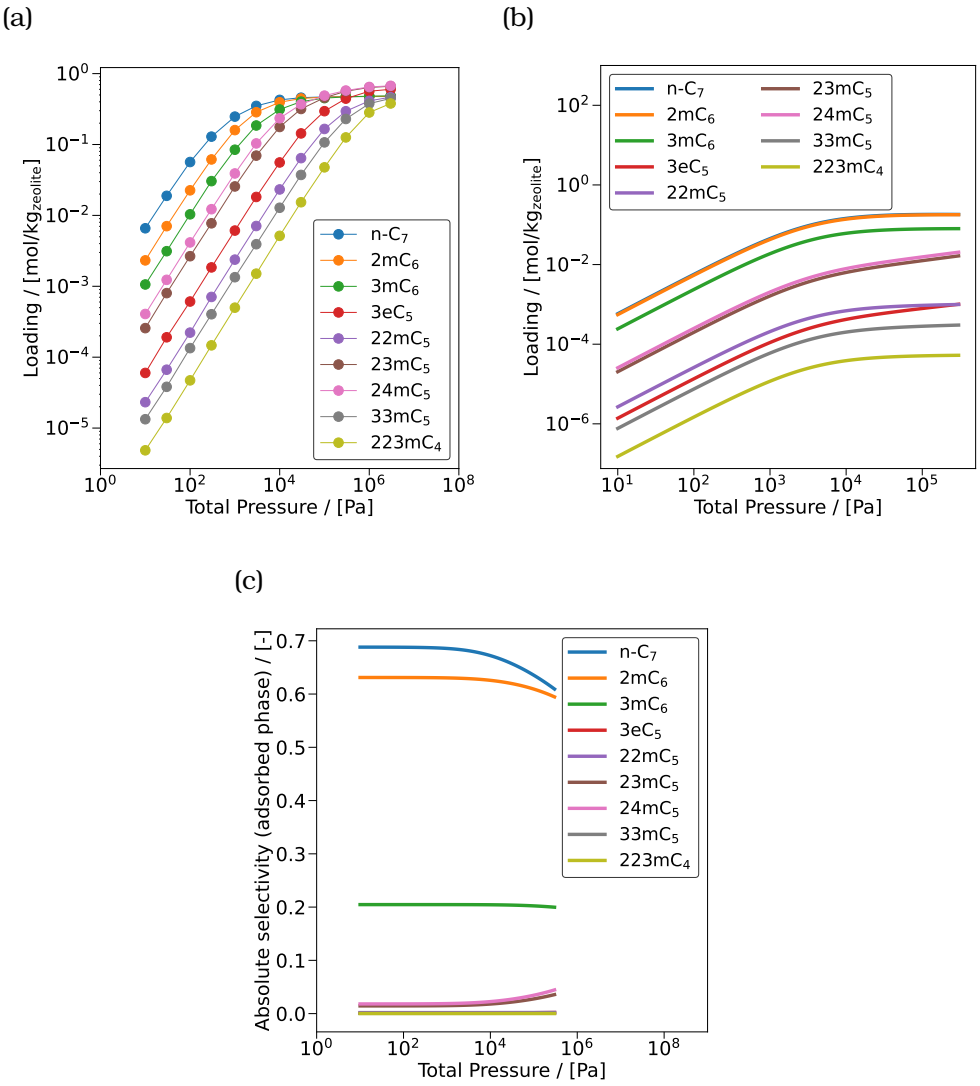


Figure 4.13. (a) Pure component adsorption isotherms, (b) mixture adsorption isotherms computed using IAST, and (c) reaction equilibrium distribution (absolute selectivities) in the adsorbed phase of C_7 isomers in MTW-type zeolite at 500 K. The pure component isotherms are fitted to single-site Langmuir-type isotherms.

at 500 K. These mixture isotherms are equivalent to the adsorbed phase distribution. The corresponding selectivities of the C_7 isomers relative are

illustrated in Fig. 4.13c. The absolute selectivities of n-C_7, 2-m-C_6, and 3-m-C_6 decrease with increasing pressure, whereas the selectivities of the multi-branched isomers increase with an increase in pressure. This is again due to size entropy effects [119–121].

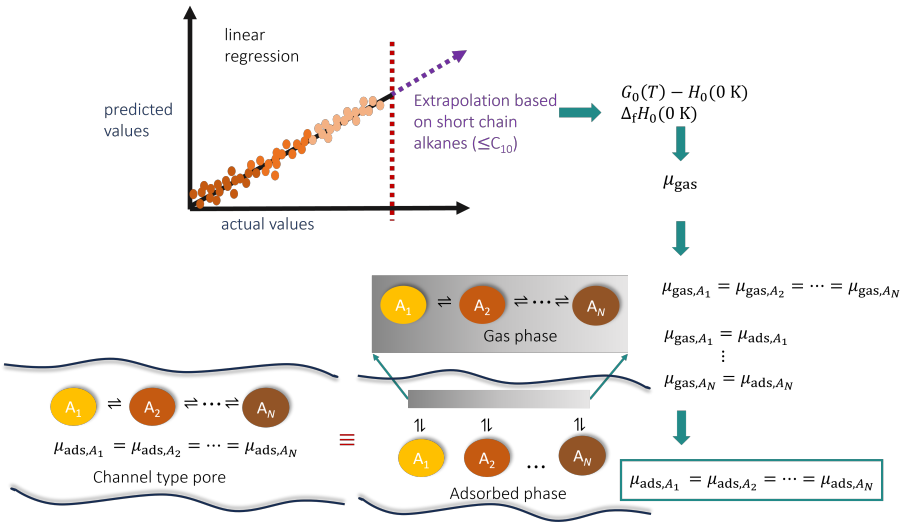
4.5. Conclusions

Shape-selectivity effects of zeolites on C_7 and C_8 isomerization are addressed using a reaction equilibrium model. In this work, we neglect framework flexibility, reaction kinetics, and cracking reactions, which are important when developing a detailed kinetic model. The equilibrium distribution in the gas phase relies on the isolated molecule partition functions. These are computed using the thermodynamic tables by Scott [416]. The selectivity in the adsorbed phase is determined by the combined effects of the gas-phase distribution and the Henry coefficients. In zeolites with larger pore sizes (FAU-type zeolite), Henry coefficients have a small impact on the adsorbed phase distribution at reaction equilibrium. Multi-branched isomers are favored in zeolites with cage-like structures (ITQ-29-type zeolite). Di-branched isomers consistently exhibit higher selectivity than n-C_7 in ITQ-29-type zeolite. BEA-, MEL-, and MFI-type zeolites exhibit similar selectivities for n-C_7, 2-m-C_6, and 3-m-C_6, with a general preference for mono-branched isomers over di-branched isomers. In zeolites with narrow one-dimensional channel-like pores (MRE- and MTW-type zeolites), variations in Henry coefficients play a crucial role in deciding the adsorbed phase reaction equilibrium distribution. MTW- and MRE-type zeolites favor n-C_7 the most. 2,2,3-m-C_4 is the least favored isomer in all zeolites considered in this study. This is due to the presence of multiple methyl branches in close proximity in this isomer, causing steric hindrance. To account for the influence of pressure, non-ideal gas behavior must be considered. The gas-phase distribution is influenced by the ratios of the isolated molecule partition function and the inverse of the ratios of the fugacity coefficients of the components. Increasing pressure hardly affects the gas phase distribution of C_7 isomers at higher temperatures (500 K - 700 K). This is because the values of the fugacity coefficients of the isomers are relatively similar. The selectivities at high pressures are primarily governed by the parameters of the pure component isotherms, which serve as inputs for the mixture isotherm model (e.g., IAST). Future work will focus on understanding the reaction product distribution of long-chain alkanes and assessing the extent to which chemical equilibrium is achieved within zeolite pores.

5

Thermochemical Properties of Alkanes Using Linear Regression: Application to Hydroisomerization

This chapter is based on the following publication: S. Sharma, J. J. Sleijfer, J. op de Beek, S. van der Zeeuw, D. Zorzos, S. Lasala, M. S. Rigutto, E. Zuidema, U. Agarwal, R. Baur, S. Calero, D. Dubbeldam, and T. J. H. Vlugt. "Prediction of Thermochemical Properties of Long-Chain Alkanes Using Linear Regression: Application to Hydroisomerization". *J. Phys. Chem. B* 128 (2024), pp. 9619–9629.



5.1. Introduction

In transitioning towards fuels and chemicals from renewable sources, platforms that provide clean hydrocarbon liquid energy carriers from carbon dioxide either directly or via bio-components can play an important role [391]. For sustainable aviation fuel and low-carbon gas-oil or lubricants, iso-alkanes with a high degree of branching are the preferred constituents [394]. Hence, shape-selective zeolite-catalyzed hydroisomerization [395], often called catalytic dewaxing, is likely to become a key step in the production of iso-alkanes, as it currently is for the classical analogue products [396, 397]. To design processes, catalysts and equipment for handling, reacting and separating (iso)alkanes and their mixtures, a detailed understanding of the thermochemical properties of these compounds, such as enthalpy, Gibbs free energy, entropy, heat capacity, and fugacity, is necessary [436]. To compute a reliable product distribution from hydroisomerization reactions, accurate prediction of thermochemical properties of alkanes is of utmost importance [390]. Another interesting application is alkane-based phase change materials for thermal energy storage systems [437].

Thermochemical properties of all isomers up to C_{10} reported by Scott [416]. The work of Scott lists Gibbs free energies ($G_0 - H_0(0\ \mathrm{K})$), enthalpies ($H_0 - H_0(0\ \mathrm{K})$), absolute entropies S_0, Gibbs free energies of formation $\Delta_f G_0$, enthalpies of formation $\Delta_f H_0$, and constant pressure heat capacities c_{p_0} of these isomers at temperatures (0-1500) K. These properties are obtained from a correlation developed using statistical mechanics, which has been trained on experimental data [422]. The experimental data for thermochemical properties of different alkanes available in literature are reported in the NIST chemistry webbook [421]. A limited amount of experimental data is available for alkanes longer than C_{10}. [438, 439]. Thermochemical properties of linear alkanes till C_{20} are listed in Refs. [403, 440]. Group contribution methods are commonly used to predict thermochemical properties of long-chain alkanes [72]. These methods are additive in nature, where the structure of a molecule is fragmented into functional groups and the thermochemical properties are estimated by summing up the contributions of each functional group present in the molecule [72]. The contribution of an individual group remains the same in every molecule in which it appears [72]. A wide range of group contribution methods for thermochemical properties of organic molecules exists in literature [441] which includes methods developed by Benson et al [71], Joback and Reid [73], Constantinou and Gani [423], Marrero and Gani [442], Hukkerikar et al [443], Albahri and Aljasmi [444], and Domalski and Hearing [445]. Benson's group additivity method [71] is commonly used to predict $\Delta_f H_0$ and S_0 [74]. Yaw's hand book [446] lists the entropies of formation $\Delta_f S_0$, $\Delta_f H_0$, $\Delta_f G_0$, and c_{p_0} in gas, liquid, and solid phases for a wide range of alkanes till C_{100}. The properties in this handbook are either collected from experimental data available in literature or predicted using Joback and Reid's method [73], especially for long-chain alkanes. The Design Institute for Physical Properties (DIPPR)

database [447] also lists the thermochemical properties of a large number of alkanes obtained from quantum chemical calculations, experiments, and group contributions like Benson's method [71]. An alternative to predict thermochemical properties is using Machine Learning (ML) models. Yalamanchi et al predicted $\Delta_f H_0$ of alkanes, alkenes, and alkynes at 298.15 K using a Support Vector Regression (SVR) model, which provided better prediction than Benson's group additivity when compared with experimental data [74]. Trinh et al also used an SVR model to predict $\Delta_f H_0$ of a wide variety of organic compounds. The training dataset was obtained from the DIPPR database [76]. Aldosari et al. predicted $\Delta_f S_0$ $\Delta_f H_0$ and c_{p_0} of hydrocarbons using SVR, v-SVR, and Random Forest Regression (RFR) algorithms [75]. Alternatively, one could also consider a High-Dimensional Model Representation (HDMR) for the longer hydrocarbons [448]. As group contribution methods are more common, we opted for this approach.

Most of the group contribution methods available in literature predict thermochemical properties for a wide range of hydrocarbons. These methods are not always accurate due to either considering only first-order groups (-CH_3, -CH_2, -CH, -C) or combining first-order group contributions with very few second-order groups [74]. Fig. 5.1 shows typical examples of second-order groups (a) CH_2(CH3)(CH) and (b) CH_2(CH)(CH), which consider the interaction between the central atom (here CH_2) and the neighboring united atoms. The united atoms inside the brackets are the first-order neighbouring groups. Considering only the first-order groups leads to less reliable predictions for highly branched isomers. Increasing the number of second-order groups in a group contribution method provides a better prediction of the thermochemical properties. Thermochemical property predictions using Scott's correlation [422] based on statistical mechanics are in very good agreement with experimental data for chains up to C_{10} isomers. This correlation is very complex and involves many different types of functions and fitting parameters, which makes it hard to apply for long-chains ($> C_{10}$) that were not in the training set. ML-based predictions have the potential to provide better predictions compared to group contribution methods. However, ML could perform poorly in case of extrapolation [75, 77], which is required for long-chain alkanes due to the absence of experimental data. If a sufficient number of independent variables is considered and the output is linearly dependent on these variables, Linear Regression (LR) will perform better than the ML models in the extrapolated region [77]. In this study, a user-friendly LR model is developed, where the occurrences of all possible second-order groups are considered as independent variables to accurately predict the thermochemical properties of alkanes longer than C_{10}.

Here, we aim to predict the thermochemical properties $(G_0 - H_0(0\ \mathrm{K}))$, $(H_0 - H_0(0\ \mathrm{K}))$, $\Delta_f G_0$, and $\Delta_f H_0$ of alkanes longer than C_{10} isomers using LR with second-order groups as molecule descriptors. The training dataset includes the ideal gas thermochemical properties of C_1 to C_{10} isomers at

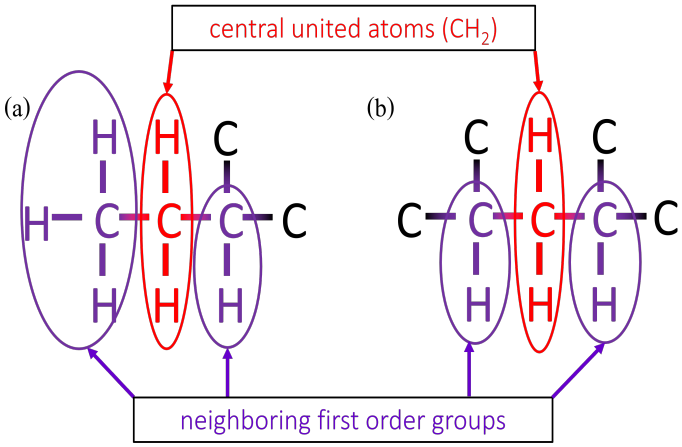


Figure 5.1. Typical examples of second-order groups (a) $CH_2(CH_3)(CH)$ and (b) $CH_2(CH)(CH)$ with CH_2 as the central united atom. The united atoms inside the brackets are the neighbouring first-order groups. In first-order group contribution methods, only the central united atom is considered, which is CH_2 in cases (a) and (b). In both cases, first-order group contributions will be identical and lead to inaccurate prediction of thermochemical properties. Unlike first-order group contribution methods, second-order group contribution methods consider the interactions between the central atom (here CH_2) and the neighbouring groups, which leads to more accurate prediction of thermochemical properties.

5

temperatures ranging from (0-1000) K, listed by Scott [416]. LR is performed at each temperature separately. To account for the effect of temperature, a quadratic polynomial as a function of temperature is fitted to the coefficients of the second-order groups for each thermochemical property. The predictions using the second-order group contribution method outperform the first-order group contribution methods and exceed the chemical accuracy of 1 kcal/mol. This is because the second-order groups include the interactions between the neighbouring groups of atoms. This study also aims towards using the predicted thermochemical properties $\Delta_f H_0$ at 0 K and $(G_0(T) - H_0\,(0\text{ K}))$ to compute the reaction equilibrium distribution of hydroisomerization of alkanes longer than C_{10} isomers at that temperature.

To optimize the yield of branched isomers in hydroisomerization, it is important to understand the reaction product distribution at chemical equilibrium [31, 390]. Hydroisomerization reactions involve adsorption of linear alkanes and dehydrogenation of these alkanes in the metal sites of the zeolites forming alkenes [1]. Protonation of alkenes takes place at the acid sites of the zeolites to form alkylcarbenium ions [21]. These ions are transferred to the metal sites where alkanes are produced via hydrogenation. This indi-

cates that alkenes as intermediates can play an important role at reaction equilibrium. However, due to the absence of alkenes in the final product distribution and lack of experimental data of thermochemical properties of alkenes, the reaction equilibrium of only alkanes is considered in this study. In our previous study, Sharma et al [390] studied the shape selectivity effects of zeolites on the reaction equilibrium distribution of hydroisomerization of C_7 and C_8 isomers. It was shown that the reaction equilibrium distribution of alkanes is useful for understanding the shape selectivity effects of zeolites on hydroisomerization of alkanes. The reaction equilibrium distribution of this reaction is determined by establishing reaction equilibrium in the gas phase and phase equilibrium between the gas and the adsorbed phase [390]. For applications such as the production of sustainable aviation fuels, long-chain alkanes (e.g., C_{16}) with a high degree of branching [394] are desirable because of high energy density, low freezing point, and good thermal stability [11]. The LR model is used to predict $\Delta_f H_0$ at 0 K and $(G_0 - H_0\,(0\text{ K}))$ at a specified temperature which are used to compute the gas phase distribution, and classical force field-based simulations are used to quantify the interactions between isomers and the zeolite. To automate the workflow and handling of a large number of isomers, alkanes are represented using SMILES strings [449, 450]. A Python function is developed to generate SMILES strings from the IUPAC names of alkanes which is required as an input in the source code for the LR model (Supporting Information SI2.py of Ref. [435]). The reaction equilibrium distribution inside constraining pore zeolites differs significantly from the gas phase distribution [390]. At infinite dilution, the reaction equilibrium distribution is strongly influenced by Henry coefficients of alkanes in constraining pore zeolites such as MTW-type zeolite [390] with pore diameters 5.6×6.0 Å [18]. The computation of Henry coefficients using classical force field-based Monte Carlo simulations requires interaction terms as input to account for both bonded and non-bonded interactions of alkanes and also non-bonded interactions between the zeolite atoms and the alkanes. A source code (Supporting Information SI3.py of Ref. [435]) for automated force field file generation for use with the RASPA2 software [39, 133] is developed to avoid the manual entry of a large number of interaction terms, which is especially useful for long-chain alkanes. This code also uses the Python function to generate SMILES strings for alkanes. Using the values for $\Delta_f H_0$ at 0 K and $(G_0 - H_0\,(0\text{ K}))$ at 500 K obtained from the LR model and Henry coefficients of alkanes calculated using classical Monte Carlo simulations in the RASPA2 software [39, 133], the reaction equilibrium distributions of hydroisomerization of C_{10} and C_{14} isomers in MTW-type zeolite at infinite dilution are computed. The reaction equilibrium distribution of C_{10} isomers in MTW-type zeolite computed using the thermochemical properties obtained from our LR model and the training dataset are in very good agreement. This suggests that the thermochemical properties predicted using the LR model can be reliably used to compute the reaction equilibrium distribution of hydroisomerization of long-chain

alkanes. In this study, hydroisomerization of C_{14} is shown as an example for long-chain alkanes. In future studies, the reaction equilibrium distribution of hydroisomerization of alkanes longer than C_{14} will be analysed in constraining pore zeolites such as MTW- and MRE-type zeolites.

This chapter is organized as follows: The important concepts and the theory behind linear regression, and simulation details are provided in Section 5.2. Our main results are discussed in Section 5.3. It is observed that LR with second-order group contributions outperforms methods based on first-order group contributions. The variations in the thermochemical properties due to the differences in branching patterns of isomers are well captured by this method, in sharp contrast to the predictions using group contribution methods available in the literature. The coefficients of the occurrences of the second-order groups of each thermochemical property are fitted using temperature-dependent quadratic polynomials. In Section 5.4, conclusions on the performance of the LR model in predicting different thermochemical properties of long-chain alkanes and the use of these properties for computing reaction equilibrium distribution of hydroisomerization reactions are discussed. The predicted thermochemical properties using this LR model can be reliably used to compute the reaction equilibrium distribution of hydroisomerization of alkanes longer than C_{10}. The training dataset for the thermochemical properties of all isomers ranging from C_1 to C_{10} is listed in the Supporting Information SI1.xlsx of Ref. [435]. In the Supporting Information SI2.py [435], the source code for the automatic generation of force field files to compute Henry coefficients in the RASPA2 software [39, 133] is included. The source code for the LR model is provided in SI3.py (Ref. [435]). SI4.xlsx [435] contains the predicted thermochemical properties of all isomers till C_{14} molecules with temperature ranging from (0-1000) K. The coefficients of the occurrences of the second-order groups obtained using LR and the corresponding temperature-dependent quadratic polynomial fits are also listed in SI4.xlsx. The reaction equilibrium distribution data of hydroisomerization of C_{10} and C_{14} molecules in the gas phase and MTW-type zeolite at 500 K are tabulated in SI4.xlsx [435]. This includes the ideal gas chemical potentials, Henry coefficients, and selectivities of these isomers in both the gas and the adsorbed phase. Appendix A.3 contains figures showing $\Delta_f H_0$ values from Scott's tables, our LR model, and the DIPPR database. Appendix A.3 also shows the variations of the coefficients of second-order groups with C as the central atom for $(G_0 - H_0\,(0\text{ K}))$ at different temperatures and the reaction equilibrium distribution imposed on the gas phase for C_{10} isomers.

5.2. Theory

A linear regression model is used to predict thermochemical properties of alkanes longer than C_{10} with the occurrences of the first or the second-order groups as independent variables, as shown below. This linear regression is

performed using the SciPy library in Python [451]:

$$y = \alpha_0 + \sum_{k=1}^{N_{\text{descriptor}}} \alpha_k x_k \tag{5.1}$$

In Eq. 5.1, y is the thermochemical property predicted using LR, x_k is the independent variable which refers to the occurrence of a first or a second-order group k in a molecule, and $N_{\text{descriptor}}$ is the total number of such independent variables or descriptors. α_0 is the intercept of the equation for LR. α_k refers to the coefficients of the independent variables. In first-order group contribution methods, the occurrences of the united atoms CH_4, CH_3, CH_2, CH, and C are considered as independent variables. In second-order group contribution methods, the combination of the central united atom and its nearest neighbors is considered as an independent variable. There are 69 second-order groups for alkanes, which are listed in the Excel worksheet Second_order_grps of the Supporting Information (SI1.xlsx) of Ref. [435]. In addition to these groups, CH_4 is also included in the second-order group contribution to predict the thermochemical properties of methane.

The training dataset contains isomers in the range C_1-C_{10}. This dataset includes the occurrences of only 46 second-order groups. The remaining second-order groups are present in alkanes of chain length longer than C_{10}. The list of the second-order groups that are not present in the training dataset, along with the smallest isomers which contain these groups, are shown in the worksheet Second_order_grps of the Supporting Information (SI1.xlsx) of Ref. [435]. To account for the contribution of these groups in the longer chains, these groups are approximated by a similar second-order group present in the training dataset. For example, the second-order group $C(C)(CH_2)(CH_2)(CH_2)$ is approximated by $C(C)(CH_2)(CH_2)(CH_3)$. The list of all such approximations is also included in the worksheet Second_order_grps of the Supporting Information (SI1.xlsx) of Ref. [435]. These second-order groups correspond to highly branched isomers which are unlikely to form inside constraining pore zeolites like MTW-type zeolite during hydroisomerization of alkanes. This is due to constraining pores present in MTW-type zeolite and steric hindrance caused by the proximity of the branches present in these isomers [390]. Therefore, approximate predictions of the thermochemical properties of such alkanes are sufficient to compute the reaction equilibrium distribution of all hydroisomerization reactions inside zeolites.

Linear regressions for the thermochemical properties are performed at a specific temperature, and the coefficients of the second-order groups (α_k) obtained from this model are specific to that temperature. To include the effect of temperature, these coefficients are refitted to a temperature-dependent quadratic polynomial.

$$\alpha = A \times T^2 + B \times T + C \tag{5.2}$$

In Eq. 5.2, A, B, and C are constants and T is the temperature in K. The values of the coefficients for the thermochemical properties $\Delta_f H_0$, $\Delta_f G_0$,

$(G_0 - H_0(0\ \text{K}))$, and $(H_0 - H_0(0\ \text{K}))$ are listed in the worksheets DHf0_coeff, DGf0_coeff, G0-H0(0K)_coeff, and H0-H0(0K)_coeff of the Supporting Information (SI2.xlsx) of Ref. [435] respectively. The corresponding temperature-dependent polynomials are listed in the worksheets DHf0_coeff_poly, DGf0_coeff_poly, G0-H0(0K)_coeff_poly, and H0-H0(0K)_coeff_poly of SI4.xlsx of the Supporting Information of Ref. [435]. The source code to predict thermochemical properties of alkanes using LR is provided in the Supporting Information (SI2.py) of Ref.[435]. To handle a large number of isomers during linear regression, alkanes are represented as SMILES strings [449, 450]. This code includes a function to generate SMILES strings for alkanes with a maximum of 5-carbon alkyl branches. The thermochemical properties, $\Delta_f H_0$ at 0 K and $(G_0 - H_0(T_{\text{ref}}))$ at a specified temperature are used to compute ideal gas chemical potentials [413, 414, 418] which are further used to compute the reaction equilibrium distribution of hydroisomerization of alkanes. The ideal gas chemical potential of component i equals [413, 414]

$$\begin{aligned}\mu_{\text{id},i} &= \mu_{\text{ref},i} + RT\ln\left(\frac{\rho_i}{\rho_0}\right)\\ &= (G_{0,i}(T) - H_{0,i}(T_{\text{ref}})) - D_{0,i} + RT\ln\left(\frac{\rho_i}{\rho_0}\right)\end{aligned} \tag{5.3}$$

In Eq. 5.3, $\mu_{\text{ref},i}$ is the reference chemical potential of component i, ρ_i is the number density of component i, and ρ_0 is the reference density which is chosen to be 1 molecule/ Å^3. T_{ref} is the reference temperature, which is 0 K in this study. $D_{0,i}$ is the atomization energy [418] which is

$$D_{0,i} = a_{\text{C}}\Delta_f H_{0,\text{C}}(0\ \text{K}) + a_{\text{H}}\Delta_f H_{0,\text{H}}(0\ \text{K}) - \Delta_f H_{0,i}(0\ \text{K}) \tag{5.4}$$

In Eq. 5.4, a_{C} is the number of C atoms and a_{H} is the number of H atoms present in the alkane isomer i. $\Delta_f H_{0,\text{C}}$, $\Delta_f H_{0,\text{H}}$, and $\Delta_f H_{0,i}$ are the enthalpies of formation of the C atom, the H atom, and alkane isomer i, respectively, at 0 K. The value of $\Delta_f H_{0,\text{C}}$ is 711.185 kJ/mol and $\Delta_f H_{0,\text{H}}$ is 216.035 kJ/mol at 0 K which are obtained from the JANAF tables [420]. $\Delta_f H_{0,i}$ values of alkanes are provided in the Supporting Information (SI1.xlsx) of Ref. [435].

The reaction equilibrium distribution of hydroisomerization reactions in zeolites can be computed by imposing reaction equilibrium in the gas phase and phase equilibrium between the gas and the adsorbed phase [31, 390, 408]. At reaction equilibrium, the chemical potentials of the reactants and the reaction products are equal in the gas phase [390]. Equating the ideal gas chemical potentials of the reactants and the reaction products leads to the gas phase reaction equilibrium distribution at infinite dilution. The equilibrium loadings in the adsorbed phase are computed using Henry's law at infinite dilution and mixture adsorption isotherm models such as Ideal Adsorbed Solution Theory (IAST) [51, 54] at finite loadings. At high temperatures (>= 500 K), the effect of pressure on the reaction equilibrium distribution of hydroisomerization of alkanes is negligible [390]. High

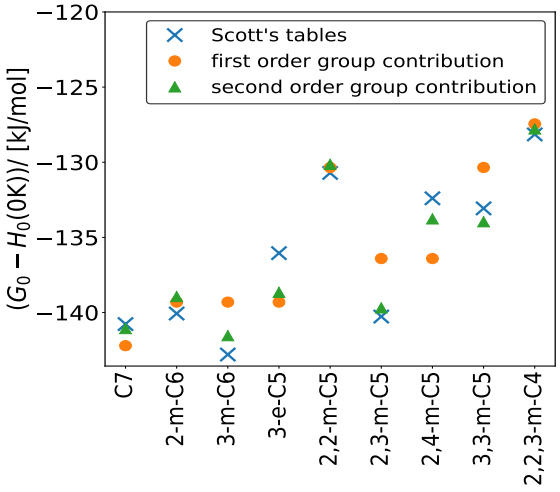


Figure 5.2. Comparison between the thermochemical property (G_0−H_0(0 K)) predicted using the LR model with first and second-order group contribution methods for C_7 isomers at 298.15 K. The predictions using the second-order group contributions are in very good agreement with the data from the Scott tables [416].

temperature leads to negligible variations in the gas phase distribution of alkane isomers with pressure and a decrease in the amount of molecules adsorbing in the zeolites [390]. Therefore, only infinite dilution is considered in this study. For details on computing the reaction equilibrium distribution of hydroisomerization, the reader is referred to chapter 4.

The required Henry coefficients are computed using the Widom test particle insertion method [46] combined with the Configurational-Bias Monte Carlo (CBMC) method [48, 49, 452, 453] in the RASPA2 software [39, 133]. Alkanes are modelled using a united-atom model [430] and the Coulomb interactions are neglected because alkanes are non-polar [39]. The intramolecular non-bonded interactions of alkanes and the intermolecular non-bonded interactions between the alkanes and the zeolite atoms are modelled using the Lennard-Jones interactions [40]. The Lennard-Jones parameters for alkanes are obtained from Dubbeldam et al [136]. The Lennard-Jones parameters for the zeolite atoms are taken from the TraPPE-zeo force field [43]. The intramolecular bonded interactions, which include bond-stretching, bond-bending, and torsion interactions, are obtained from Refs. [37, 454] Both bonded and non-bonded interaction parameters are listed in the Excel worksheet forcefield_param of the Supporting Information SI1.xlsx of Ref. [435]. Files containing force field parameters and the list of bonded and non-bonded interactions are required as input in the RASPA2 software [39, 133].

The number of intramolecular interactions increases tremendously with increasing chain lengths of alkanes (e.g., n-C_{14} contains 91 intramolecular interactions). Therefore, a Python code for automatic generation of force field files for alkanes is provided in the Supporting Information (SI3.py) of Ref. [435]. This code requires alkanes to be represented as SMILES strings [449, 450]. Therefore, this code also includes the Python function to generate SMILES strings for alkanes. All Lennard-Jones interactions are truncated and shifted at 12 Å without applying tail corrections. The number of unit cells in the simulation box of MTW-type zeolite is $2 \times 15 \times 2$ for C_{10} isomers and $2 \times 25 \times 2$ for C_{14} isomers. The zeolite is considered to be a rigid structure as the effect of zeolite flexibility is negligible on adsorption processes, especially at infinite dilution [44]. The selectivity of a component (Eqs. 4.21 and 4.22 in chapter 4) in both gas phase and adsorbed phase is defined as the ratio of the mole fraction y_i of the component to the sum of the mole fractions of all other components present in the same phase. To compare the selectivity of a component relative to another, the term relative selectivity ($s_{\text{rel},i}$) is defined as the ratio of the absolute selectivity of that component to a reference component as shown in Eq. 4.23 and 4.24 in chapter 4 for both gas and adsorbed phase. In this study, n-C_{10} and n-C_{14} molecules are chosen as reference components for computing relative selectivities.

5.3. Results and Discussion

Fig. 5.2 shows the comparison between $(G_0 - H_0(0\ \text{K}))$ of C_7 isomers at 400 K predicted by the LR model using first and second-order group contribution methods. The predicted values are also compared with the data obtained from the tables by Scott [416]. The thermochemical properties obtained using the second-order groups as descriptors are in very good agreement with the training dataset (Fig. 5.2) These predictions are much better than those obtained using the first-order group contributions. This is because the influence of the neighbouring groups of atoms is neglected in the first-order group contribution, which is clearly shown in Fig. 5.2.

The Mean Absolute Errors (MAEs) for the prediction of the thermochemical properties of alkanes using the first and the second-order group contributions are listed in Table 5.1. The use of the second-order groups as descriptors leads to smaller MAEs compared to the first-order group contributions. Therefore, only second-order group contributions are considered in this study. For comparison, the values of $\Delta_f H_0$ at 400 K for a few selected C_{10} isomers obtained from the training dataset [416] and computed using the LR model are listed in table 5.2. The values computed using the LR model and those obtained from the training dataset are in excellent agreement.

Several group contribution methods are available in literature such as Bensons's [71], Constantinou and Gani's [423], Joback's [73], and Domalski and Hearing's [445] methods which either consider first-order group contributions or a combination of first-order groups and a few second-order

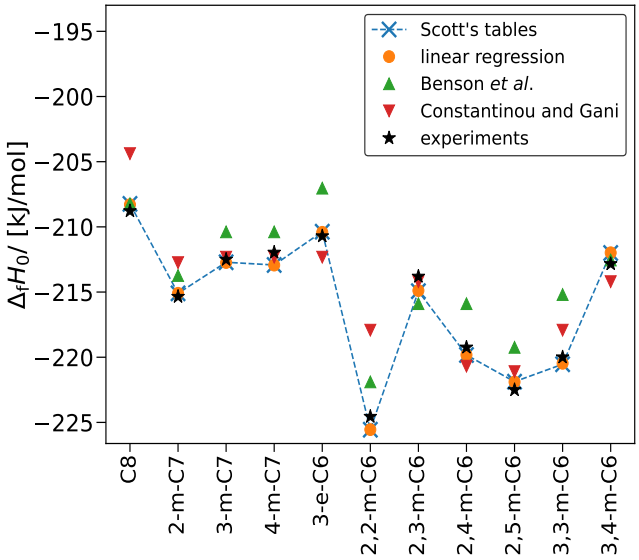


Figure 5.3. Prediction of $\Delta_f H_0$ for C_8 isomers at 298.15 K using the LR model, Benson et al's group additivity method [71], Constantinou and Gani's group contribution [423], the Scott tables [416], and the experimental data listed by Scott [422]. The predictions using the LR model are in excellent agreement with the Scott tables and the data from the experiments [422] For the $\Delta_f H_0$ values of C_8 isomers at 298.15 K, using the experimental data as the reference, the Mean Absolute Errors (MAEs) are the smallest for our LR model (0.62 kJ/mol) and the Scott tables (0.64 kJ/mol). This is followed by Constantinou and Gani's group contribution method (2.04 kJ/mol), and Benson et al's group additivity method (2.37 kJ/mol) respectively. The dashed blue line through the data points obtained from the Scott tables is a guide to the eye.

Table 5.1. Mean Absolute Errors (MAEs) of the thermochemical properties $\Delta_f G_0$, $\Delta_f H_0$, $(G_0 - H_0(0\ \text{K}))$, $(H_0 - H_0(0\ \text{K}))$ predicted using linear regression (LR) with the first and the second-order group contributions.

Thermochemical Property	MAE (1st order)/ [kJ/mol]	MAE (2nd order)/ [kJ/mol]
$\Delta_f G_0$	7.529	1.029
$\Delta_f H_0$	5.834	0.152
$(G_0 - H_0(0\ \text{K}))$	3.087	1.012
$(H_0 - H_0(0\ \text{K}))$	1.041	0.181

Table 5.2. Comparison between the values of enthalpy of formation $\Delta_f H_0$ for C_{10} isomers at 400 K obtained from the training dataset [416] and the linear regression model.

Isomer	$\Delta_f H_0$/ [kJ/mol]	
	Training data	Linear regression data
n-C_{10}	-265.47	-265.47
4-m-C_9	-270.04	-270.07
4-e-C_8	-267.48	-267.49
2,2-m-C_8	-282.55	-282.56
3-e-4-m-C_7	-266.60	-266.53
2,2,5-m-C_7	-287.19	-287.09
3,3-e-C_6	-266.10	-266.08
4-e-3,3-m-C_6	-261.75	-261.76
2,2,3,5-m-C_6	-284.68	-284.57
3-e-2,2,4-m-C_5	-261.37	-261.37
2,2,3,4,4-m-C_5	-261.54	-261.54

groups. Fig. 5.3 shows the variations in $\Delta_f H_0$ of a few C_9 isomers at 298.15 K predicted using the LR model, Benson's [71], and Constantinou and Gani's methods [423]. The properties predicted using Benson's and Constantinou and Gani's methods are computed using the SPLIT software by AmsterChem [455]. Fig. 5.3 also includes data obtained from the Scott tables [416], and Yaws' handbook [446]. Yaws' handbook [446] uses Joback's group contribution method [73] for long-chain isomers. Benson's [71] and Constantinou and Gani's [423] group contribution methods are not always able to distinguish between isomers based on the positions and the types of branches these isomers possess. For example, the Constantinou and Gani's method [423] provided nearly identical values of $\Delta_f H_0$ for 3-m-C_8, 4-m-C_8, 3-e-C_7, and 4-e-C_7 whereas experimental data from Scott tables clearly shows variations due to the presence of the methyl and the ethyl groups as branches in these isomers. Similarly, Benson's method [71] does not distinguish between

2,3-m-C_7, 2,4-m-C_7, and 2,5-m-C_7. Such variations are well captured by our LR model and are necessary to determine the selectivities of reaction products in hydroisomerization of alkanes [390]. $\Delta_f H_0$ values of a few C_9 and C_{10} isomers at 298.15 K obtained from the Scott tables and predicted by LR are also compared with those listed in the DIPPR database [447] as shown in Fig. A.20 in Appendix A.3. $\Delta_f H_0$ of C_{10} isomers listed in the DIPPR database [447] are computed using the Domalski and Hearing's method [445]. The values predicted using our LR model are in excellent agreement with the experimental data from the Scott tables. The values obtained from the DIPPR database [447] are also in good agreement with the Scott tables [416] with small deviations for 4-m-C_8, 2-m-C_9, and 3-m-C_9.

LR is performed at a specific temperature for each thermochemical property. To account for the effect of temperature on the thermochemical properties and compute these properties at any temperature, the coefficients which correspond to the occurrence of the second-order groups are fitted to a temperature-dependent quadratic polynomial (Eq. 5.2). Fig. A.21 in Appendix A.3 shows the variations in the magnitudes of the coefficients of the occurrences of the second-order groups for $(G_0 - H_0(0\ \mathrm{K}))$ with C as the centre atom present in the training dataset. The fitted coefficients using the quadratic polynomial are in excellent agreement with those predicted using the LR model. The values of the coefficients are different at a specific temperature (Fig. A.21 in Appendix A.3). In case of first-order group contributions, the variations will be identical as each group has the same central united atom C. Therefore, combining these coefficients into a single coefficient to reduce the number of independent variables or simply using first-order group contributions will lead to erroneous predictions of the thermochemical properties. This clearly indicates the need for a second-order group contribution method.

The predicted thermochemical properties $(G_0 - H_0(0\ \mathrm{K}))$ at a specified temperature and $\Delta_f H_0$ at 0 K for alkanes longer than C_{10} are used to compute the ideal gas chemical potentials which are further used in calculating the reaction equilibrium distribution of hydroisomerization of long-chain alkanes. Fig. A.22 in Appendix A.3 shows the reaction product distribution of C_{10} isomers in the gas phase at infinite dilution and 500 K. The reaction product distribution obtained using LR is in very good agreement with the training dataset (Fig. A.22 in Appendix A.3). In both cases, the variations in the selectivities of C_{10} isomers relative to n-C_{10} are similar. The gas phase distribution and Henry coefficients are used to compute the reaction equilibrium distribution in MTW-type zeolite at infinite dilution and 500 K (Fig. 5.4). For mono-branched isomers, 4-p-C_7 and 4-ip-C_7 have the smallest preferences in MTW-type zeolite as shown in Fig. 5.4b. This indicates that isomers with branches longer than propyl or isopropyl groups, such as butyl and pentyl groups, will have very low selectivities in MTW-type zeolite and can be excluded from the reaction product distribution. Such isomers can be eliminated from the calculation of the reaction equilibrium distribution of

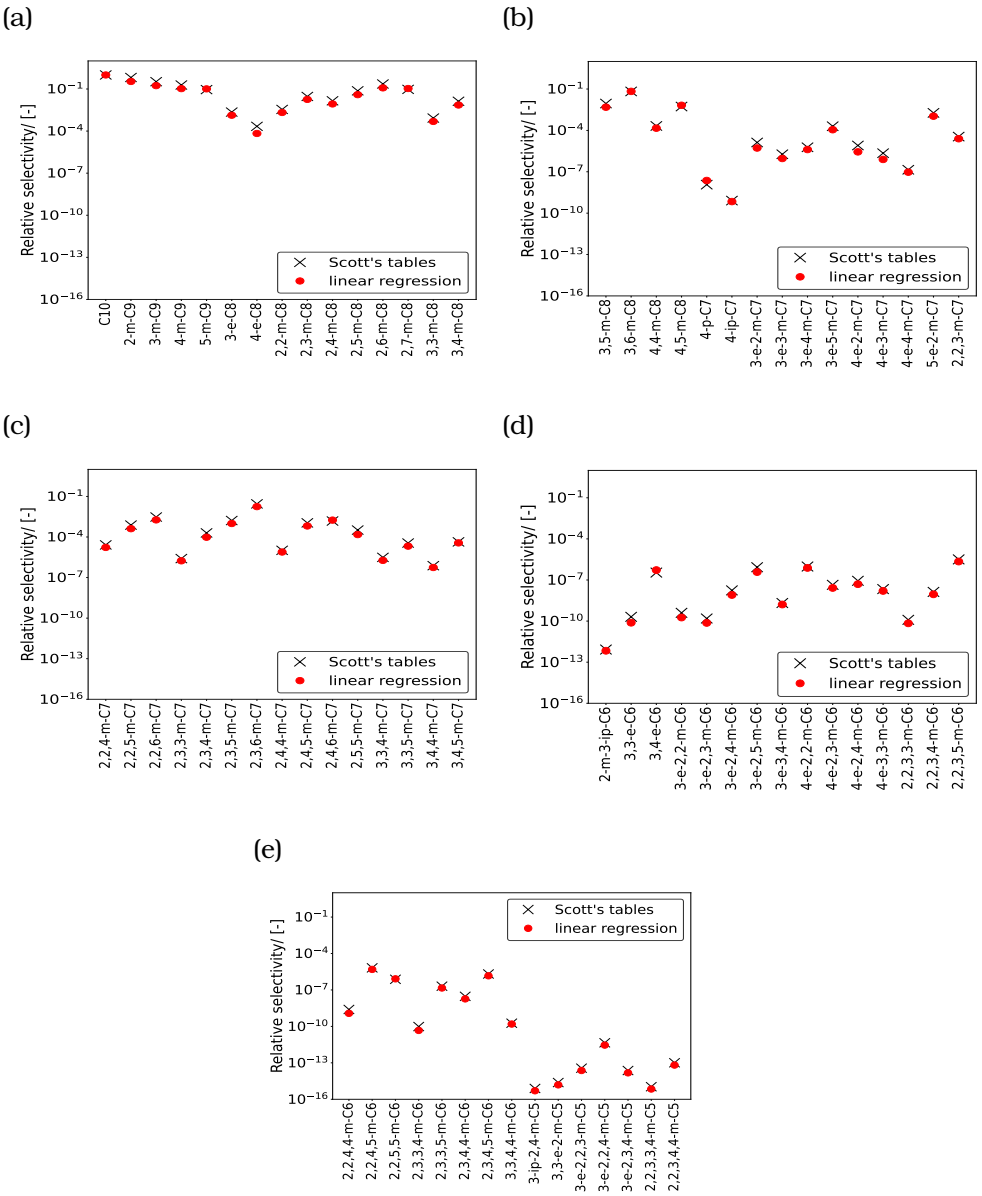


Figure 5.4. Selectivities of C_{10} isomers relative to n-C_{10} at reaction equilibrium in MTW-type zeolite at infinite dilution and 500 K. The gas phase reaction equilibrium distribution required in computing the adsorbed phase reaction equilibrium distribution are obtained using the Scott tables (black crosses) and the LR model (red filled circles). The raw data is listed in the Excel worksheet xi_iC10_500K of the Supporting Information (SI4.xlsx file) of Ref. [435].

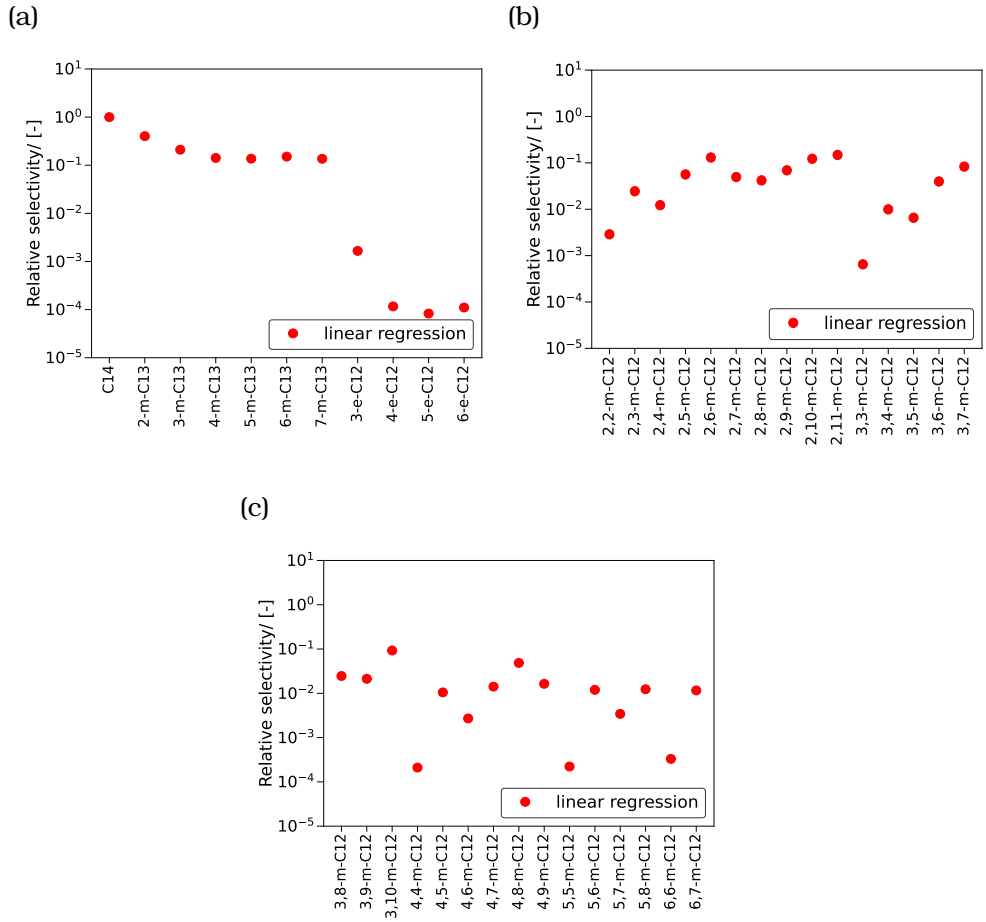


Figure 5.5. Selectivities of different C_{14} isomers relative to n-C_{14} at reaction equilibrium in MTW-type zeolite at infinite dilution and 500 K. The LR model is used to predict $(G_0 - H_0\,(0\text{ K}))$ at 500 K and $\Delta_f H_0$ at 0 K which are used to compute the ideal gas chemical potentials of C_{14} isomers. These chemical potentials are used to compute the gas phase reaction equilibrium distribution of hydroisomerization of C_{14}. The raw data is listed in the Excel worksheet xi_iC14_500K of the Supporting Information (SI4.xlsx file) of Ref. [435].

hydroisomerization of long-chain alkanes. The ethyl-trimethyl isomers (3-e-2,2,3-m-C_5, 3-e-2,2,4-m-C_5, and 3-e-2,3,4-m-C_5) and pentamethyl isomers (2,2,3,3,4-m-C_5 and 2,2,3,4,4-m-C_5) have the least preference compared to all other isomers. This indicates that isomers with multiple branches present in the vicinity of each other are not preferably formed inside constraining pore zeolites such as MTW-type zeolite. Reaction product distributions obtained using both LR and Scott tables provide very similar variations in selectivities. This suggests that the thermochemical properties obtained using LR are reliable and can be used in computing the reaction equilibrium distribution for hydroisomerization of long-chain alkanes.

Fig. 5.5 shows the reaction product distribution of C_{14} isomers in MTW-type zeolite at infinite dilution and 500 K. Mono-methyl branched isomers have a higher preference compared to the mono-ethyl and di-methyl isomers. This is due to the smaller pore diameters present in MTW-type zeolites (5.6 × 6.0 Å) [18]. The order of magnitude of relative selectivities varies for di-branched isomers, which depend on the proximity of the methyl branches. The geminal alkanes (2,2-m-C_{12}, 3,3-m-C_{12}, 4,4-m-C_{12}, 5,5-m-C_{12}, and 6,6-m-C_{12}) have the least selectivity due to the steric hindrance posed by the presence of two methyl groups at the same position in the alkane chain. Isomers with methyl groups far apart (2,10-m-C_{12} and 2,11-m-C_{12}) have higher selectivities compared to other di-methyl isomers. In future studies, such criteria will be used to identify relevant isomers for computing the reaction equilibrium distribution of alkanes longer than C_{14}. Such filtering of isomers is necessary because the number of isomers increases enormously for alkanes longer than C_{14}, and it is difficult to consider each isomer experimentally.

5.4. Conclusions

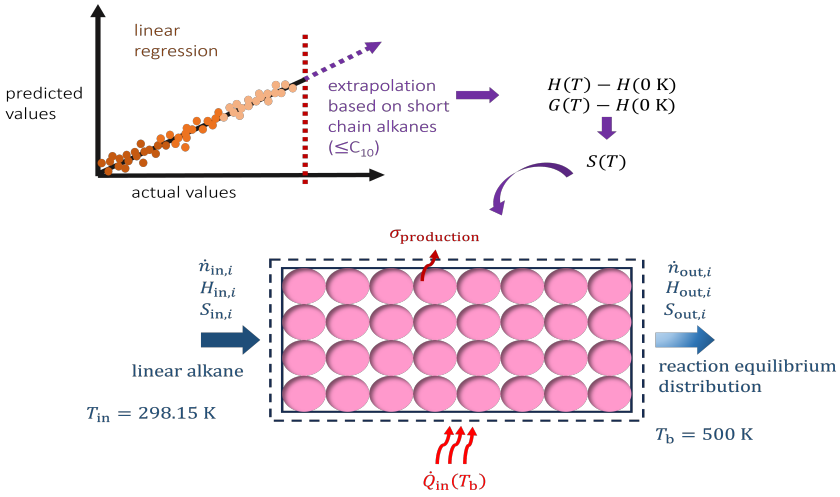
In this work, we developed a linear regression model based on 69 systematically enumerated second-order group contributions to predict the thermochemical properties $\Delta_f H_0$, $\Delta_f G_0$, $(G_0 - H_0(0\ \mathrm{K}))$, and $(H_0 - H_0(0\ \mathrm{K}))$ of alkanes. The model achieves excellent agreement with the Scott tables [416], with errors well below the threshold of chemical accuracy and a maximum mean absolute error of only 1.03 kJ/mol for $\Delta_f G_0$. Although the use of 69 second-order groups may seem extensive, this is feasible due to the comprehensive availability of high-quality reference data in the Scott tables, and it enables a much more accurate description of structural effects in branched alkanes compared to first-order group methods and classical approaches such as those of Benson [71], Joback [73], and Constantinou and Gani [423]. The model successfully reproduces equilibrium distributions for hydroisomerization of C_{10} and C_{14} isomers in MTW-type zeolites, in combination with automatically computed Henry coefficients from RASPA2 [39, 133], and reveals important selectivity rules such as the suppression of isomers with long branches and the higher selectivity of isomers with widely

separated branches. These results confirm that the LR model is a powerful tool for reaction equilibrium analysis of long-chain alkanes in zeolites and can be extended to heavier alkanes, where automated generation of force field and input files will be essential. For other thermophysical properties such as critical temperature, critical pressure, acentric factor, and liquid density, the limited availability of experimental data makes a fully exhaustive second-order scheme less practical. In such cases, the Constantinou-Gani framework [423], possibly expanded with selected second-order groups, provides an attractive compromise. Recent work has demonstrated that combining the second-order approximation strategy of the Constantinou-Gani method with a sensitivity-guided selection of groups, inspired by the approach of Sharma et al. [435], yields compact yet chemically meaningful group sets that preserve predictive accuracy for both thermochemical and critical properties. This hybrid strategy offers a balanced bridge between accuracy and simplicity [456].

6

Computing Entropies for Alkanes Using Linear Regression: Application to Hydroisomerization

This chapter is based on the following publication: S. Sharma, R. Baur, M. S. Rigutto, E. Zuidema, U. Agarwal, S. Calero, D. Dubbeldam, and T. J. H. Vlugt. "Computing Entropy for Long-Chain Alkanes Using Linear Regression: Application to Hydroisomerization". *Entropy* 26 (2024), p. 1120.



6.1. Introduction

Entropy data are important for a wide range of hydrocarbon applications, from industrial processes [75] to environmental assessments [458]. Entropy changes during chemical reactions like hydroisomerization, combustion, and cracking have a large influence on the feasibility of such reactions [459]. Iso-alkanes with high branching are preferred over linear ones in sustainable aviation fuel and lubricants [394], making shape-selective zeolite-catalyzed hydroisomerization [17, 395], also known as catalytic dewaxing, important for the production of these alkanes [396, 397]. Understanding thermochemical properties, such as entropies and enthalpies, is essential for designing efficient processes and equipment for (iso)alkane handling and separation [436]. Process energy efficiency is best evaluated through the second law of thermodynamics [459, 460]. The second law efficiency requires calculating the exergy destruction, which equals the entropy production multiplied by the environmental temperature [459, 460]. The exergy destruction is the useful work that is destroyed due to irreversibilities in a process [461].

Thermochemical properties, including absolute entropies (S_0) for isomers up to C_{10} from 0 to 1500 K, have been reported by Scott [416]. Experimental data for the thermochemical properties of various alkanes are available in the literature and can be accessed through the NIST Chemistry Webbook [421]. However, there is limited experimental data for alkanes longer than C_{10}. Several group contribution methods exist in the literature to compute thermochemical properties, which include those by Benson et al. [71], Constantinou and Gani [423], Joback and Reid [73], Marrero and Gani [442], Hukkerikar et al. [443], Albahri and Aljasmi [444], and Domalski and Hearing [445]. Yaw's Handbook [446] and the Design Institute for Physical Properties (DIPPR) database [447] also list thermochemical properties for many long-chain alkanes, which are either obtained from experiments or group contribution methods. Machine Learning (ML) models have emerged as an alternative for predicting thermochemical properties. Aldosari et al. have successfully used Support Vector Regression (SVR) [462], ν-SVR [463], and Random Forest Regression (RFR) [82] algorithms to predict $\Delta_f S_0$, $\Delta_f H_0$, and $c_{p,0}$ for hydrocarbons [75]. For longer hydrocarbons, High-Dimensional Model Representation (HDMR) could be considered as another approach to predict these thermochemical properties [448].

Most of the group contribution methods in literature are often inaccurate because these methods only consider first-order groups (CH_3, CH_2, CH, and C) or combine with a few second-order groups [74, 423]. Second-order group contributions account for interactions between neighboring groups, improving thermochemical property predictions [435]. Increasing the number of second-order groups improves the predictions of thermochemical properties because these groups also account for interactions between the nearest neighboring groups of atoms [435]. Scott's correlation [422], based on statistical mechanics, provides accurate predictions up to C_{10} isomers when validated with experiments. However, this correlation is complex and

difficult to apply to long-chain alkanes (> C_{10}), as it requires numerous functions and fitting parameters. While Machine Learning models offer promising predictive power, these models may struggle with extrapolation for long-chain alkanes [75, 77]. In such cases, Linear Regression (LR) can outperform ML models when the independent variables are accurate and the output has a linear relationship with these variables [77].

In chapter 5 of this thesis, we developed a user-friendly Linear Regression (LR) model based on second-order group contributions to predict thermochemical properties like Gibbs free energies ($G_0 - H_0(0\ \text{K})$), enthalpies ($H_0 - H_0(0\ \text{K})$), Gibbs free energies of formation $\Delta_f G_0$, and enthalpies of formation $\Delta_f H_0$ for alkanes longer than C_{10} for temperatures in the range (0-1000) K [435]. The second-order groups consider the interactions between the central atom and the neighboring groups. These groups are solely determined based on the topology of the alkanes with SMILES strings as input. Therefore, no 3D structures of alkanes were used in our LR model. Here, we calculated absolute entropies S_0 from ($G_0 - H_0(0\ \text{K})$) and ($H_0 - H_0(0\ \text{K})$), predicted using the LR model. The entropy values are in excellent agreement with both Scott's data [412] and experimental values [422]. Our method effectively captures the variations in entropy in isomers due to differences in branching. The S_0 values are used to compute the entropy production during hydroisomerization of alkanes in different zeolites at reaction equilibrium. Understanding the reaction product distribution at equilibrium is key to optimizing branched isomer yield in hydroisomerization [31]. Hydroisomerization involves adsorption and subsequent dehydrogenation of linear alkanes at the metal sites of zeolites, forming alkenes [1]. These alkenes are protonated at the acid sites, resulting in the formation of alkyl carbenium ions [21]. Alkanes are regenerated from these alkylcarbenium ions via hydrogenation [21]. Hydroisomerization is accompanied by cracking reactions depending on the operating temperature of the reactor. This study is relevant for conditions where the hydroisomerization reaction approaches chemical equilibrium with a negligible amount of cracked products. Cracking reactions, being irreversible [20], must be excluded from the reaction equilibrium distribution study. Calculating reaction equilibrium distributions for hydroisomerization reactions also provides insights into the shape selectivity effects of zeolites during these processes [390]. The calculation of entropy production in an equilibrium reactor involves steady-state mass and energy balances. The reaction equilibrium distribution provides the amount of isomers at the reactor outlet, while the energy balance determines the heat input required for the reactions. We analyzed the heat input and the entropy production for C_7 isomers in FAU-, ITQ-29-, BEA-, MEL-, MFI-, MTW-, and MRE-type zeolites at 500 K, observing small variations in these zeolites due to differences in reaction equilibrium distributions. Both heat input and entropy production increase with alkane chain length and higher temperatures due to rising enthalpies and absolute entropies.

This chapter is organized as follows: Section 6.2 presents the key equations for computing the absolute entropies of alkanes, the mass balance in the equilibrium reactor, the energy balance for calculating the heat input, and the entropy balance for determining the entropy production. Section 6.3 discusses the results, showing that the computed entropies are highly accurate. A minimal variation in entropy production and heat input is observed for different zeolites during the hydroisomerization of linear alkanes. Both heat input and entropy production increase with longer alkane chains and higher temperatures. Section 6.4 provides conclusions on the accuracy and usefulness of the entropy data for alkanes and the effects of zeolites, chain length and temperature on the entropy production for hydroisomerization at reaction equilibrium. The Supporting Information SI.xlsx of Ref. [457] contains the computed entropy values derived from the enthalpies and Gibbs free energies predicted by our LR model for isomers ranging from C_1 to C_{14}, as well as the comparison between the absolute entropies computed using our model [435] and those predicted using group contribution methods by Benson et al. [71] and Constantinou and Gani [423].

6.2. Theory

The absolute entropy S_0 is computed as follows [459]

$$S_0 = \left(\frac{1}{T}\right)[(H_0 - H_0\,(0\text{ K})) - (G_0 - H_0\,(0\text{ K}))] \tag{6.1}$$

where $(H_0 - H_0\,(0\text{ K}))$ is the enthalpy at temperature T relative to 0 K. 0 K is considered as the reference temperature in this study. $(G_0 - H_0\,(0\text{ K}))$ is the Gibbs free energy at temperature T relative to 0 K. These thermochemical properties are predicted using our LR model based on second-order group contributions for hydrocarbons [435]. The computed absolute entropies S_0 for alkane isomers $(C_1 - C_{14})$ at temperatures (0-1000) K are provided in the Supporting Information SI.xlsx. S_0 values are used to calculate the entropy production in a hydroisomerization reactor at reaction equilibrium (Fig. 6.1), where ideal gas behavior is assumed. Calculating the entropy production requires molar flow rates at the inlet and the outlet of the reactor and the rate of heat input. Both heat input and entropy production are computed per unit mole of fluid mixture at the inlet. A steady state molar balance is considered in the reactor as we assume that all isomerization reactions are at equilibrium:

$$\sum_{i=1}^{N_{\text{comp}}} \dot{n}_{\text{in},i} = \sum_{i=1}^{N_{\text{comp}}} \dot{n}_{\text{out},i} \tag{6.2}$$

where $\dot{n}_{\text{in},i}$ and $\dot{n}_{\text{out},i}$ are the molar flow rates of component i at the reactor column inlet and outlet respectively. At the inlet, only the linear alkane is

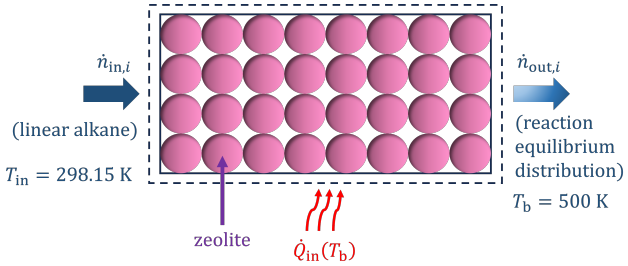


Figure 6.1. Schematic representation of a reactor for hydroisomerization of linear alkanes at chemical equilibrium. The feed at the column inlet ($\dot{n}_{\text{in},i}$) is at 298.15 K. The heat input $\left(\dot{Q}_{\text{in},i}\right)$ and the reactions take place at 500 K.

considered at 298.15 K. The composition at the outlet is obtained from the reaction equilibrium distribution for hydroisomerization of linear alkanes at infinite dilution in the zeolite. The reaction equilibrium distribution is obtained by imposing a gas phase reaction equilibrium for the alkane isomers and a simultaneous phase equilibrium between the gas and the adsorbed phase for each component [390]. This satisfies the reaction equilibrium distribution in the zeolites [408]. A similar approach was also adopted by Hansen et al. [410] to study the influence of silicalite-1 pores on the reaction equilibrium distribution of the propene metathesis reaction. The adsorbed phase loadings for these isomers are obtained using Henry's law. The Henry coefficients are computed using Widom's test particle insertion method [46, 50] in the RASPA software [39, 133, 464]. The mole fractions in the adsorbed phase depend on the Henry coefficients and the gas phase mole fractions. The total pressure cancels out from the equation [390]. For further details on this method, the reader is referred to Refs. [390, 435]. The gas phase reaction equilibrium distribution is computed using $(G_0 - H_0\,(0\text{ K}))$ at the operating temperature T and $\Delta_{\text{f}}H_{0,i}\,(0\text{ K})$ at 0 K [390]. For alkanes longer than C_{10} isomers, these properties are obtained using our LR model [435]. The rate of heat input $\dot{Q}_{\text{in}}$ to the reactor is computed from the steady state energy balance for the reactor [459].

$$\dot{Q}_{\text{in}} = \sum_{i=1}^{N_{\text{comp}}} \dot{n}_{\text{out},i} H_i - \sum_{i=1}^{N_{\text{comp}}} \dot{n}_{\text{in},i} H_i \tag{6.3}$$

where H_i is the enthalpy of component i which is computed as follows [459].

$$H_i\,(T) = \Delta_{\text{f}}H_{0,i}\,(0\text{ K}) + (H_i\,(T) - H_i\,(0\text{ K})) \tag{6.4}$$

In Eq. 6.4, $\Delta_{\text{f}}H_{0,i}\,(0\text{ K})$ is the enthalpy of formation of component i at 0 K. The entropy production $\sigma_{\text{production}}$ is computed using the steady-state entropy balance for the reactor [459].

$$\sigma_{\text{production}} = \sum_{i=1}^{N_{\text{comp}}} \dot{n}_{\text{out},i} S_{\text{out},i} - \sum_{i=1}^{N_{\text{comp}}} \dot{n}_{\text{in},i} S_{\text{in},i} - \frac{\dot{Q}_{\text{in}}}{T_{\text{b}}} \tag{6.5}$$

$S_{\text{in},i}$ and $S_{\text{out},i}$ in Eq. 6.5 are the absolute entropies of component i at the inlet and the outlet of the reactor. $S_{\text{in/out},i}$ depends on the fluid composition at the inlet and outlet of the reactor [459]. The heat input and the reactions take place at temperature T_{b}.

$$S_{\text{in/out},i} = S_{0,i} - R \ln\left(\frac{x_{\text{in/out},i} P}{P_{\text{ref}}}\right) \tag{6.6}$$

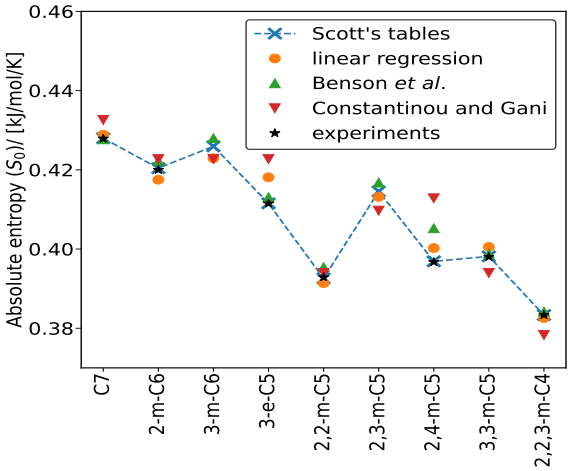
where $S_{0,i}$ is the absolute entropy for an isolated molecule of component i, R is the universal gas constant, $x_{\text{in/out},i}$ are the mole fractions of component i at the inlet or outlet of the reactor. P is the operating pressure of the reactor, and P_{ref} is the reference pressure, which is considered as 1 bar in this study. The term $R \ln(P/P_{\text{ref}})$ drops out in the entropy balance (Eq. 6.5) because the total molar flow rate at the inlet and outlet is equal (Eq. 6.2). For all cases, the operating temperature T_{b} is 500 K.

6.3. Results and Discussion

Fig. 6.2 compares the absolute entropies of C_7 (Fig. 6.2a) and C_8 (Fig. 6.2b) isomers at 298.15 K, calculated using our LR model [435], the group contribution methods of Benson et al. [71] and Constantinou and Gani [423], as well as data obtained from Scott's tables [416] and experimental results [422]. The data in Scott's tables were obtained by fitting statistical mechanics-based correlations to experimental measurements [422]. The entropies predicted by our LR model [435] show excellent agreement with both Scott's tables and experimental data. Unlike the group additivity method by Benson et al. [71] and the group contribution approach by Constantinou and Gani [423], our LR model successfully distinguishes between different isomers based on the number, types, and positions of the branches. The Mean Absolute Errors (MAEs) of $(H_0 - H_0\,(0\text{ K}))$ and $(G_0 - H_0\,(0\text{ K}))$, predicted by our linear regression model, are 1.012 kJ/mol and 0.181 kJ/mol, respectively [435]. The MAE for the product of temperature and absolute entropy S_0 is 1.03 kJ/mol, which exceeds the chemical accuracy of 4.184 kJ/mol. The dataset comparing the absolute entropies computed by our method and those predicted by group contribution methods of Benson et al. [71] and Constantinou and Gani [423] for C_{14} isomers at 298.15 K is included in the Excel worksheet S0_comparison of the Supporting Information SI.xlsx.

The composition at the reactor column outlet represents the reaction equilibrium distribution for hydroisomerization of alkanes in different zeolites at infinite dilution. Fig. 6.3 shows the reaction equilibrium distri-

(a)



(b)

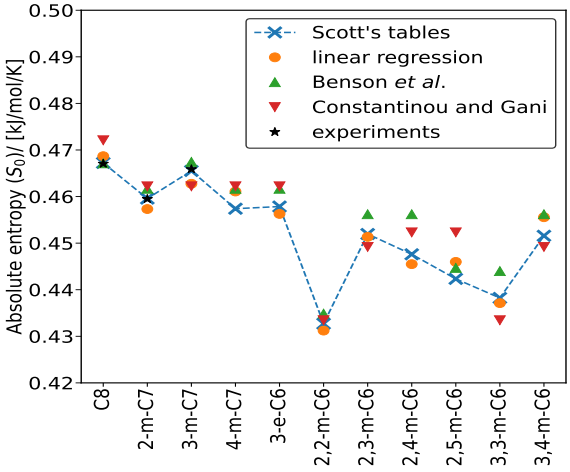


Figure 6.2. Computation of absolute entropies S_0 for (a) C_7 and (b) C_8 isomers at 298.15 K using our LR model, group contribution methods by Benson et al. [71], and Constantinou and Gani [423], the Scott's tables [416], and the experimental data listed by Scott [422]. The predictions using the LR model are in excellent agreement with Scott's tables and the data from the experiments [422]. The dashed blue line through the data points obtained from Scott's tables is a guide to the eye.

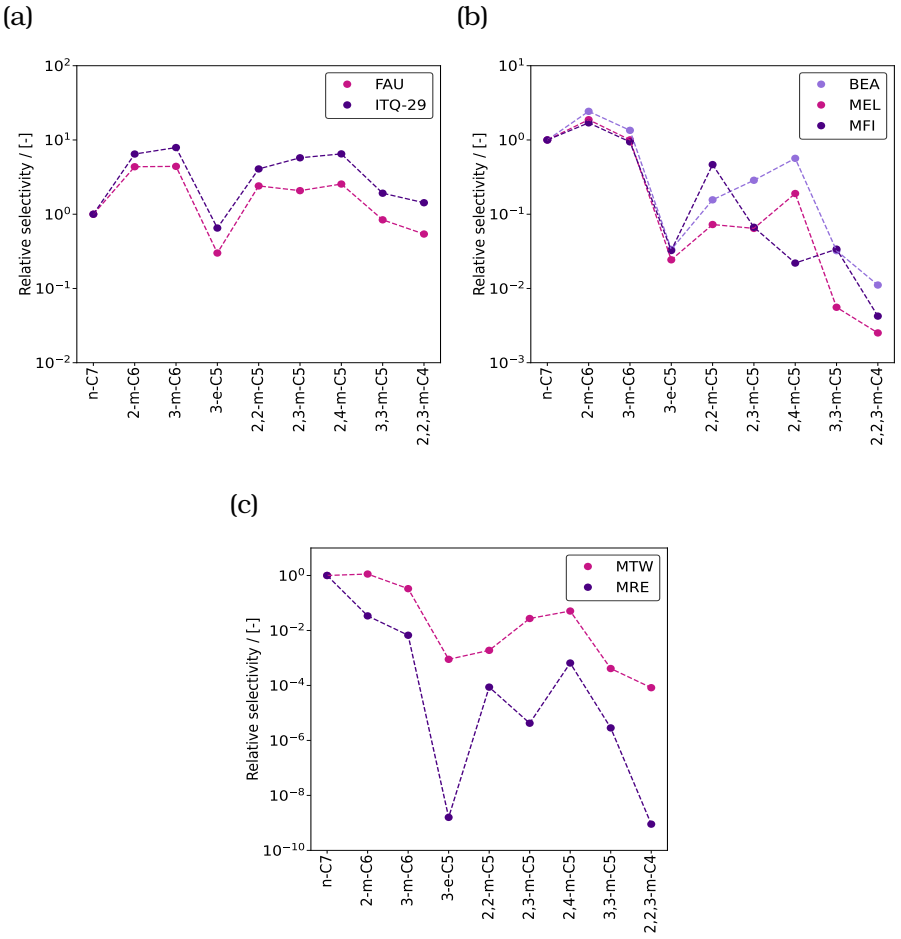


Figure 6.3. Selectivities of C_7 isomers relative to n-C_7 at reaction equilibrium in (a) FAU- and ITQ-29-type zeolites, (b) BEA-, MEL-, and MFI-type zeolites, and (c) MTW- and MRE-type zeolites at infinite dilution and 500 K. The absolute selectivity is defined as the mole fraction of a certain component divided by the sum of the mole fractions of all other components [429]. The relative selectivity refers to the ratio of the absolute selectivity of a specific isomer to that of the reference isomer, which is C_7 in this case. The dashed lines through the data points are a guide to the eye.

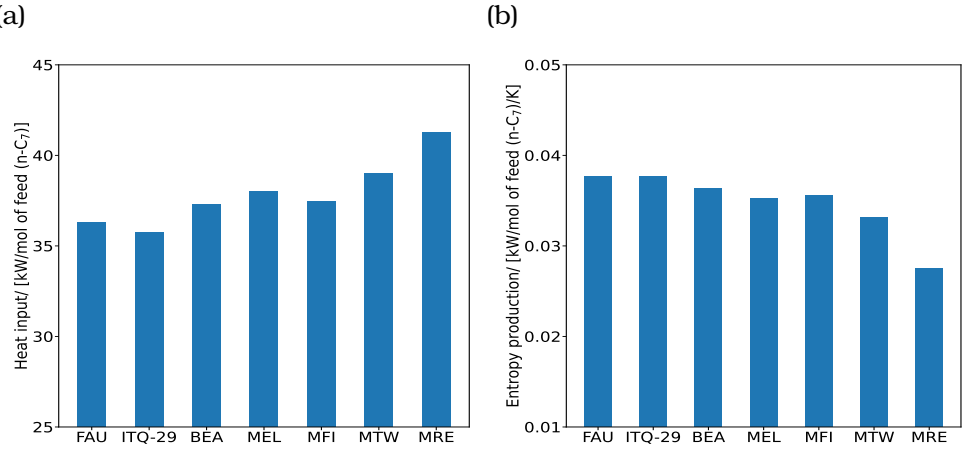


Figure 6.4. Variations of (a) heat input $\dot{Q}_{\text{in}}$ and (b) entropy production $\sigma_{\text{production}}$ for hydroisomerization of C_7 isomers at reaction equilibrium in FAU-, ITQ-29-, BEA-, MEL-, MFI-, MTW-, and MRE-type zeolites. The reactions take place at 500 K and infinite dilution in the zeolite.

butions for the hydroisomerization of C_7 isomers in FAU-, ITQ-29-, BEA-, MEL-, MFI-, MTW-, and MRE-type zeolites at 500 K. These zeolites are classified into three categories based on the shape and size of the pores, which show distinct selectivity for alkanes. FAU- and ITQ-29-type zeolites have cage-like pore structures [18]. BEA-, MEL-, and MFI-type zeolites feature three-dimensional channels with intersections [18]. MTW- and MRE-type zeolites contain one-dimensional pore structures [18]. The data sets for the reaction equilibrium distributions are included in the Excel sheet xi_C7_500K of the Supporting Information SI.xlsx. In FAU- and ITQ-29-type zeolites (Fig. 6.3a), the adsorbed phase selectivity is primarily influenced by the gas phase thermochemical properties $\Delta_f H_0\,(0\text{ K})$ and $(G_0 - H_0\,(0\text{ K}))$. Dimethyl isomers (2,2-m-C_5, 2,3-m-C_5, and 2,4-m-C_5) are favored over n-C_7 inside these zeolites. Due to larger pore diameters, the influence of pore structure on selectivities is small. FAU- and ITQ-29-type zeolites can accommodate molecules as large as 11.24 Å and 11.05 Å respectively [18]. BEA-, MEL-, and MFI-type zeolites with 3D channel-like pores connected via intersections show similar selectivities for n-C_7 and the monomethyl isomers (2-m-C_6 and 3-m-C_6), which are preferentially formed compared to dimethyl-, trimethyl-, and ethyl-branched isomers (Fig. 6.3b). In BEA-, MEL-, and MFI-type zeolites, variations in selectivities are influenced by both gas phase thermochemical properties and Henry coefficients. In MTW- and MRE-type zeolites, n-C_7 has the highest selectivity, followed by 2-m-C_6 and 3-m-C_6, with variations influenced largely by the Henry coefficients (Fig.

6.3c) because of the smaller diameters and one-dimensional nature of the pores [390]. The overall differences in selectivities between the linear and the branched isomers are mainly determined by the pore diameters of the zeolites. For isomers with identical degrees of branching, the selectivities are determined by the distributions and the shapes of the pore structures.

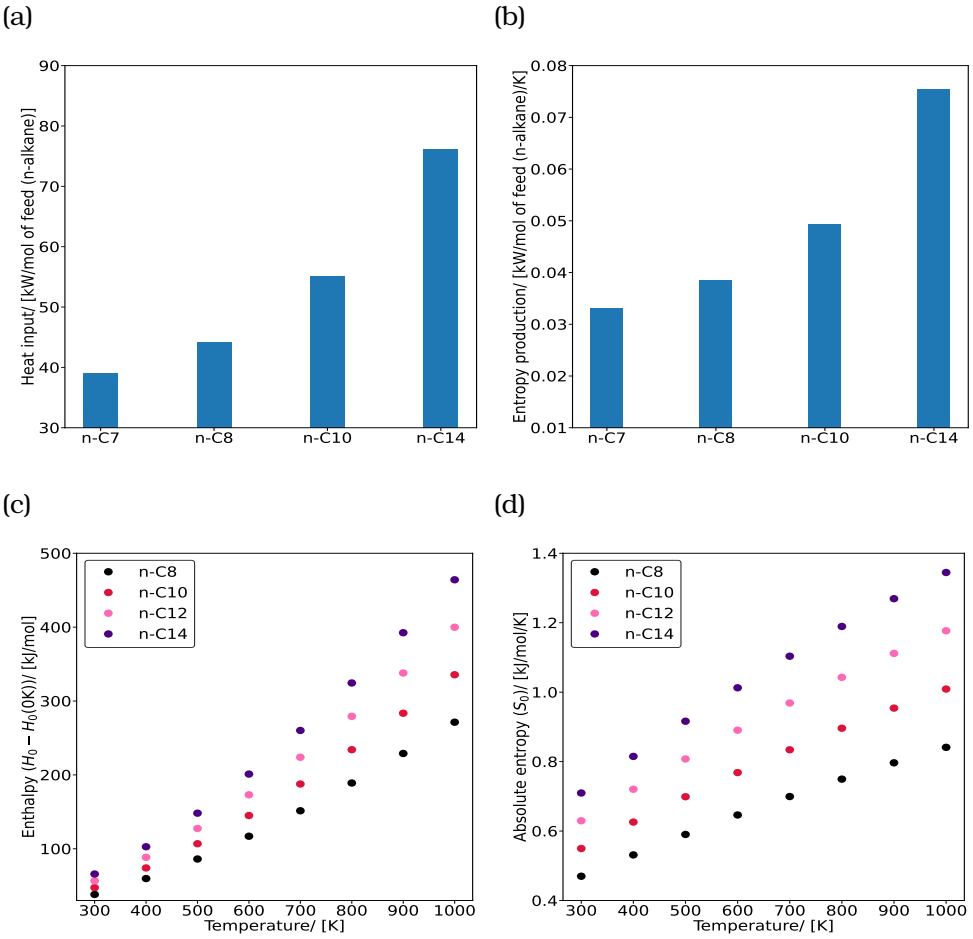


Figure 6.5. Effect of chain length on (a) heat input $\dot{Q}_{\text{in}}$ and (b) entropy production $\sigma_{\text{production}}$ for hydroisomerization of C_7, C_8, C_{10}, and C_{14} isomers at reaction equilibrium in MTW-type zeolite. The reactions take place at 500 K and infinite dilution. Variations of (c) enthalpies $H_0 - H_0\,(0\text{ K})$ and (d) absolute entropies S_0 with temperatures in the range (300-1000) K for n-C_8, n-C_{10}, n-C_{12}, and n-C_{14} isomers.

Fig. 6.4 shows the variations in the rates of heat input (Fig. 6.4a) and

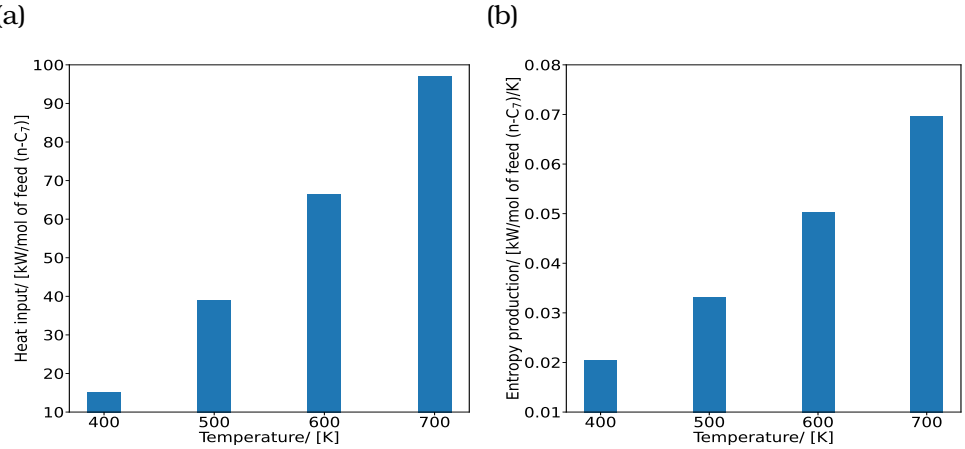


Figure 6.6. Effect of temperature on (a) heat input $\dot{Q}_{\text{in}}$ and (b) entropy production $\sigma_{\text{production}}$ for hydroisomerization of C_7 isomers at reaction equilibrium in MTW-type zeolite at 400, 500, 600, and 700 K. The reactions take place in the zeolite at infinite dilution.

entropy production (Fig. 6.4b) for hydroisomerization of C_7 isomers in an equilibrium reactor for different zeolites. Both heat input and entropy production show small variations with the zeolites, influenced by the unique reaction equilibrium distributions in each type of these zeolites. The rate of heat input slightly increases from large pore zeolites (FAU- and ITQ-29-types) to narrower 1D zeolites (MTW- and MRE-types) (Fig. 6.4a). This is because n-C_7, 2-m-C_6, and 3-m-C_6 isomers, which have higher enthalpies compared to the multi-branched isomers, are formed in large proportions in these 1D zeolites at reaction equilibrium. The entropy production decreases with decreasing pore size of the zeolites due to smaller values of entropies for these linear and mono-branched alkanes at reaction equilibrium in these zeolites. This may vary when cracking reactions are involved.

Hydroisomerization of linear alkanes, n-C_7, n-C_8, n-C_{10}, and n-C_{14} are studied in MTW-type zeolite at 500 K and infinite dilution. With increasing chain length, both heat input to the reactor (Fig. 6.5a) and entropy production (Fig. 6.5b) increase because of the rapid increase in enthalpies and entropies with temperature for long-chain alkanes as shown in Figs. 6.5c and 6.5d respectively. The enthalpies increase with chain length due to the larger number of carbon-hydrogen (C-H) and carbon-carbon (C-C) bonds, each adding to the total enthalpy of the molecule. Longer alkanes also have higher absolute entropies compared to the shorter ones because the larger number of atoms allows for more molecular arrangements [465]. This leads to an increase in entropy production during the hydroisomer-

ization of longer alkanes. Fig. 6.6a shows an increase in the rate of heat

input with increasing temperature. The internal energy increases at higher temperatures, which leads to higher enthalpy values. This increases the rate of heat requirement in the reactor column. The absolute entropies of alkanes increase with temperature because higher temperatures provide more energy to the molecules to occupy a larger number of configurations. This leads to an increase in the rate of entropy production with temperature (Fig. 6.6b).

6.4. Conclusions

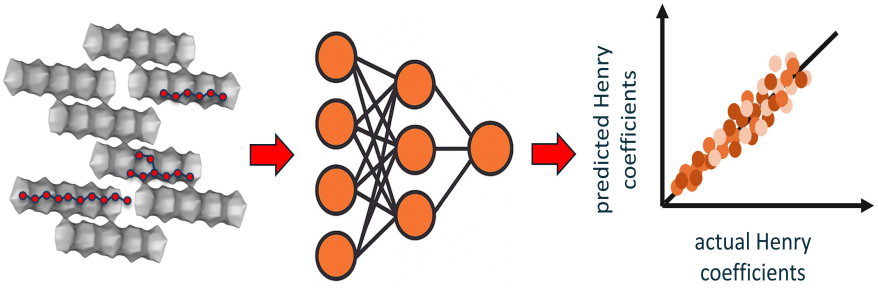
The absolute entropies (S_0) of alkanes computed from $(H_0 - H_0(0\ \text{K}))$ and $(G_0 - H_0(0\ \text{K}))$ as predicted by our LR model show excellent agreement with Scott's tables [416] and experimental data [422]. Our model effectively captures entropy variations based on the type, number, and positions of branches in isomers, outperforming group contribution methods like those by Benson et al. [71] and Constantinou and Gani [423]. The computed absolute entropies will be valuable for designing and optimizing alkane-based processes with a focus on second law efficiency. Unique reaction product distributions are observed in the hydroisomerization of C_7 isomers in different zeolites. Large-pore zeolites, such as FAU- and ITQ-29-type zeolites favor dimethyl isomers compared to n-C_7, which is mainly due to $(G_0 - H_0(0\ \text{K}))$ and $\Delta_f H_0\,(0K)$. In BEA-, MEL-, and MFI-type zeolites, normal alkanes and monomethyl isomers show similar selectivities and are preferred over dimethyl and trimethyl isomers for C_7 and C_8 alkanes. In narrow-pore zeolites (MRE- and MTW-types), n-C_7 is preferred, followed by mono-branched isomers. Small variations in heat input and entropy production are observed for different zeolites due to distinct reaction product distributions, influenced by zeolite pore shape and size. These trends may shift when cracking reactions are considered. As the alkane chain length increases, both heat input and entropy production increase. This is due to the increase in enthalpies and entropies with temperature for longer chains. A larger number of C-H and C-C bonds in long-chain alkanes increases the internal energy, leading to a larger enthalpy and, consequently, higher heat input. The increased number of atoms in long-chain alkanes allows for more molecular conformations, leading to higher absolute entropies. At higher temperatures, increased molecular energy leads to more accessible conformations, raising both heat input and entropy production. At higher temperatures, molecules gain more energy to occupy a larger number of configurations, causing an increase in entropy and, therefore, an increase in entropy production. This work demonstrates the accuracy of our LR model in predicting absolute entropies of alkanes and offers insights into heat requirements and entropy production during the formation of branched

isomers in zeolite-catalyzed hydroisomerization, which supports improved design and optimization of this process. The computed entropies will also be relevant for improving energy efficiencies in other processes involving hydrocarbons.

7

Machine Learning for Henry Coefficients of Alkanes in Zeolites

This chapter is based on the following publication: S. Sharma[1], P. Yang[1], Y. Liu, K. Rossi, P. Bai, M. S. Rigutto, E. Zuidema, U. Agarwal, R. Baur, S. Calero, D. Dubbeldam, and T. J. H. Vlugt. "Machine Learning-Based Predictions of Henry Coefficients for Long Chain Alkanes in One Dimensional Zeolites: Application to Hydroisomerization". *J. Phys. Chem. C* (2025). Manuscript accepted.



[1]These authors contributed equally to this work.

7.1. Introduction

In transitioning towards producing fuels and chemicals from renewable sources, platforms that deliver clean hydrocarbon liquid energy carriers, either directly from carbon dioxide or via bio-based intermediates, are expected to play a significant role [391]. For applications such as sustainable aviation fuel, low-carbon gas oils and lubricants, iso-alkanes with a high degree of branching are the preferred components due to favorable combustion and flow properties [394]. Consequently, shape-selective zeolite-catalyzed hydroisomerization, commonly referred to as catalytic dewaxing, is going to be a crucial step in the production of branched alkanes, just as it is in the manufacturing of conventional petroleum-derived analogues [395].

Predicting selectivities at reaction equilibrium for hydroisomerization, particularly at low conversion, requires accurate knowledge of both gas-phase thermochemical properties and adsorption behavior inside zeolite pores. In our recent study on shape selectivity of zeolites on hydroisomerization [390] (chapter 4), we demonstrated that combining the gas-phase Gibbs free energies and the enthalpies of formation and Henry coefficients inside zeolite pores enables reliable prediction of isomer selectivities at reaction equilibrium. This can be very useful for determining the product distribution for long-chain alkanes in zeolites, which is difficult to obtain from experiments. While thermochemical properties of long-chain alkanes in the gas phase can be reliably predicted using our previously developed linear regression model based on second-order group contributions [435], obtaining accurate Henry coefficients still remains a major challenge. Henry coefficients are essential for understanding adsorption-based shape selectivities of branched and linear alkanes in zeolite frameworks, providing insight into how molecular structures influence zeolite adsorption. Henry coefficients are defined as the slope of the adsorption isotherm at low loading in units of moles/kg framework/Pa. Henry coefficients are usually obtained from force field-based molecular simulations [50], which are computationally demanding, particularly for a large number of isomers for long-chain alkanes. This makes large-scale screening impractical, especially when enumerating hundreds of thousands of isomers [466]. This calls for the development of robust and accurate Machine Learning (ML) models for quick and reliable predictions of Henry coefficients.

As the alkane chain length increases, the number of possible structural isomers grows exponentially [466] due to the large number of ways carbon atoms can be arranged. For instance, C_4 has only two isomers, C_{10} has 75, and C_{20} has over 366,000 isomers [466]. This rapid combinatorial explosion presents a significant challenge for systematic studies involving long-chain alkanes. To efficiently explore this large chemical space, it is essential to implement automated enumeration of isomers and generation of the corresponding SMILES (Simplified Molecular Input Line Entry System) strings [449]. SMILES representations are essential for simulations, property predictions, and machine learning workflows. Several tools are

available for isomer enumeration, each with distinct strengths and limitations. PubChemPy [467, 468] supports chemical searches but not full isomer enumeration. MOLGEN (closed-source) and MAYGEN (open-source) [469] reliably generate comprehensive sets of structural isomers. Surge is another open-source isomer generator which uses a canonical path method for enumeration [470]. ENU [466], a graph theory-based tool optimized for acyclic alkane isomers, is particularly efficient for long-chain hydrocarbons. As shown in Table 7.1, the SMILES strings generated by ENU [466] do not always follow IUPAC naming conventions. This complicates the systematic classification of isomers, an important step in understanding zeolite shape selectivity for hydroisomerization. To address this issue, we have developed an isomer enumeration code in C++ and Python to generate exhaustive lists of alkane isomers containing methyl, ethyl, propyl, and isopropyl branches. Isomers with branches larger than propyl or isopropyl groups are excluded, as such bulky substituents are unlikely to adsorb in the narrow pores (ca. 5-6 Å) [18] of one-dimensional zeolites.

Table 7.1. Comparison of SMILES strings obtained from ENU software [466] and the SMILES strings [449] corresponding to IUPAC nomenclature [471] for different isomers.

IUPAC abbreviation	**SMILES (ENU software)**	**SMILES (IUPAC)**
2,3,9-m-7-e-C_{10}	CCC(CCCC(C)C(C)C)CC(C)C	CC(C)C(C)CCCC(CC)CC(C)C
2,3,8-m-6-p-C_{10}	CCCC(CCC(C)C(C)C)CC(C)C	CC(C)C(C)CCC(CCC)CC(C)C

Henry coefficients for alkanes are typically computed using the Widom particle insertion method [46] combined with the Configurational Bias Monte Carlo (CBMC) algorithm [452, 453]. This method becomes computationally expensive for large numbers of structural isomers of long-chain alkanes, for a large number of zeolite frameworks. ML models have recently emerged as a promising alternative for predicting Henry coefficients. These models have gained significant importance for accelerating materials discovery and optimization for a range of applications, such as gas separation, adsorption-based storage, and catalyst design [91]. Recent reviews have highlighted the transformative potential of AI in zeolite discovery and adsorption modeling, including generative design, property prediction, and simulation acceleration [472, 473]. In porous materials such as Metal Organic Frameworks (MOFs) and zeolites, ML has proven to be highly effective in predicting adsorption properties, thereby enabling high-throughput screening without resorting to computationally expensive molecular simulations or time-consuming experiments [78–81]. Efficient ML models have been developed to predict temperature-dependent Henry coefficients and adsorption selectivities for various adsorbates in MOFs, facilitating rapid evaluation for gas separation applications with small molecules such as CO_2, CH_4, and H_2 capture [474]. A comprehensive study by Gharagheizi and Sholl [475] systematically evaluates the accuracy of IAST in predicting binary gas adsorption for a

wide range of porous materials and conditions. Daou et al [476] combined ML with Ideal Adsorbed Solution Theory (IAST) to efficiently screen silica and cationic zeolites for short-chain alkane separation. These authors used ML-predicted isotherms to identify trends in capture storage and selectivity trends. Beyond adsorption selectivities, other studies have examined the underlying structure-property relationships. In particular, Rzepa et al [91] revealed that adsorption enthalpy and entropy correlate linearly in various zeolite–adsorbate systems, governed largely by the degree of molecular confinement. Liu et al [477] have developed a three-dimensional Convolutional Neural Network (CNN) [478] framework ZeoNet designed to predict Henry coefficients for n-octadecane adsorption in over 330,000 zeolite structures using efficient grid-based volumetric representations. This method significantly outperforms traditional geometric-descriptor models and achieves near-simulation accuracy while being orders of magnitude faster [477]. Despite these advances, systematic studies of the entire long-chain alkane space, especially in the context of zeolite-based adsorption, remain insufficiently investigated. The main challenge lies in the large number of zeolites (ca. 10^6 hypothetical structures [479]) and the millions of possible isomers for long-chain alkanes [466]. Computing Henry coefficients for such a large number of alkane–zeolite combinations using molecular simulations is practically not feasible. Our work addresses this gap by exploring ML-based prediction of Henry coefficients for long-chain alkane isomers in one-dimensional zeolites, specifically MTT-, MTW-, MRE-, and AFI-type zeolites (Fig. 1.2 in chapter 1) because of the simplicity of the structures, their relevance for hydroisomerization reactions, and the absence of complicated window effects [57, 480]. While this study focuses on training and evaluating models within individual zeolite frameworks, future work will explore cross-framework generalization to assess model transferability for different pore geometries and confinement environments. Prediction of Henry coefficients of alkanes is particularly important for hydroisomerization applications, where adsorption-based shape selectivity governs catalytic performance.

A major challenge here arises from activity cliffs [481, 482], which are sharp discontinuities in structure-property relationships, where minor structural changes yield disproportionately large differences in molecular properties. Activity cliffs occur because adsorption in narrow-pore zeolites is highly sensitive to alkyl branching. These cliffs are particularly problematic for ML models, which often assume a smooth structure–property landscape [482]. Though extensively studied in medicinal chemistry, activity cliffs have not been addressed in the context of hydrocarbon adsorption to the best of our knowledge. In the case of alkane adsorption in zeolites, activity cliffs are observed in Henry coefficients due to effects such as branching position, symmetry, and confinement-induced steric interactions. A small structural change like a methyl group shift from one position to the next in branched alkanes can significantly alter its fit in one-dimensional zeolite pores, leading to sharp variations in Henry coefficients. For example, 2,3-dimethylpentane

and 2,4-dimethylpentane (2,3-m-C_5 and 2,4-m-C_5) isomers in MRE-type zeolite at 500 K differ by orders of magnitude in Henry coefficients (9.44×10^{-8} and 4.85×10^{-6} mol/kg framework/Pa, respectively) [390]. This is due to the shape of the zeolite pores. The corrugations in the zeolite pores arise from alternating peaks and crests caused by the arrangement of the zeolite atoms [390]. This results in variations in channel diameter over the length of the pore (Fig. 7.1). The likelihood of the methyl branches in 2,4-m-C_5 fitting into two separate peaks is higher than for 2,3-m-C_5 (Fig. 7.1). This is due to the larger separation between the methyl groups in 2,4-m-C_5, which allows for more favorable adsorption in the corrugated pore structure. The most widely used metric to identify such cliffs is the Structure-Activity Landscape Index (*SALI*) [483], defined as the ratio of the absolute property difference and a structural distance metric. Typical examples of structural distance metrics are the Tanimoto coefficient [484, 485] and the Levenshtein distance [486] between SMILES strings. Alternative approaches are two- and three-dimensional activity cliff maps [483, 487], Matched Molecular Pair (MMP) analysis [488], and gradient-based landscape profiling [489]. Several recent studies have emphasized the importance of explicitly identifying and incorporating activity cliff data into training pipelines through structural diversity sampling, active learning, and contrastive learning strategies [490–492]. We also quantify activity cliffs for long-chain alkanes using Levenshtein distance-based SALI indices and systematically analyze the impact on ML model accuracy. This provides not only a performance benchmark for conventional and graph-based models but also a strategy for improving predictive power in chemically diverse and cliff-prone datasets.

In this chapter, the performance of several ML models, including Random Forest (RF) [82], Extreme Gradient Boosting (XGB) [83], Cat Boost (CB) [84, 85], Tabular Prior Data Fitted Network (TabPFN) [89, 90], and Directed Message Passing Neural Network (D-MPNN) [92] are compared to predict Henry coefficients for alkane isomers in MTT-, MTW-, MRE-, and AFI-type zeolites at 523 K. RF [82] is an ensemble learning method that constructs multiple decision trees using bootstrapped subsets of the training data and averages the predictions to improve accuracy and mitigate over-fitting. It introduces additional randomness by selecting a subset of features at each split, making it highly robust to noise and capable of capturing complex feature interactions. XGB [83] is a gradient boosting algorithm that builds trees sequentially, with each new tree learning to correct the errors of its predecessors. It incorporates advanced regularization techniques, efficient handling of missing data, and optimized tree-pruning strategies, making it exceptionally fast, scalable, and highly accurate for structured datasets. CB [84, 85] is a gradient boosting framework, specifically designed to handle categorical features natively, eliminating the need for extensive preprocessing such as one-hot encoding. It uses ordered boosting to reduce over-fitting and prevent target leakage, and it constructs symmetric trees, which both accelerate training and enhance model generalization. TabPFN [89, 90] is

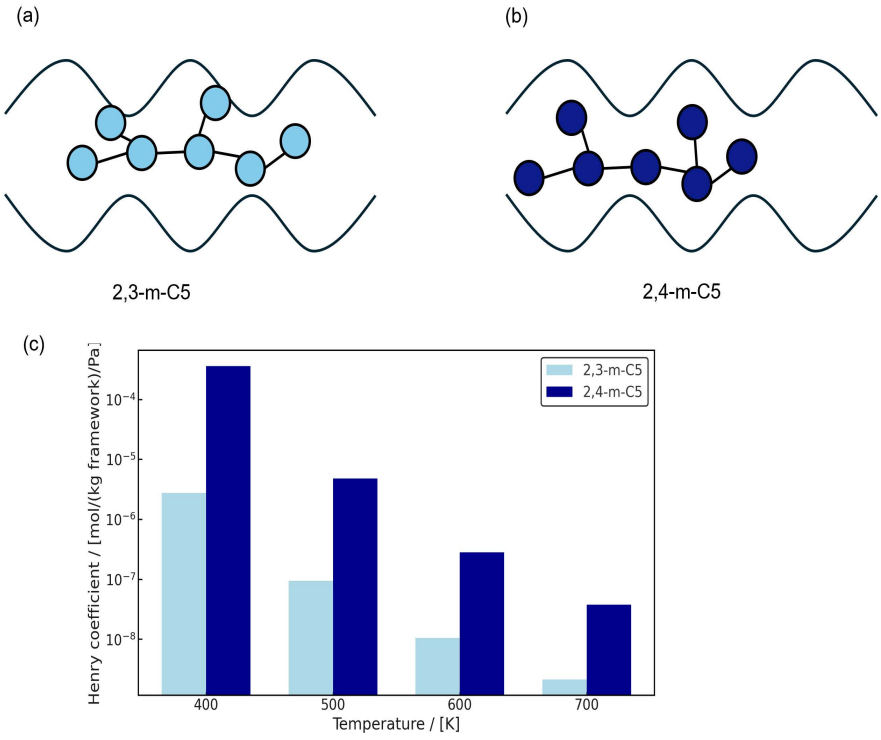


Figure 7.1. Schematic representation of the adsorption of (a) 2,3-m-C_5 and (b) 2,4-m-C_5 isomers in MRE-type zeolite at infinite dilution conditions. The zeolite pore corrugations arising from alternating peaks and crests [390] induce variations in channel diameter. The larger separation between the methyl groups in 2,4-m-C_5 enables more favourable adsorption compared to 2,3-m-C_5. (c) Henry coefficients of 2,3-m-C_5 and 2,4-m-C_5 in MRE-type zeolite for the temperature range 400–700 K. Despite having structural similarities, the Henry coefficients of these isomers differ by orders of magnitude, a typical example of activity cliff.

a transformer-based model, pre-trained on millions of synthetic data and fine-tuned to perform Bayesian inference on small, structured datasets. It enables near-instant, probabilistic predictions with no gradient-based training, making it especially powerful for tabular data where interpretability and uncertainty quantification are key. D-MPNN [92] is a graph neural network architecture that represents molecules as graphs derived from the SMILES strings, where atoms are nodes and bonds are directed edges. By directed message passing, bond-level interactions are captured and encoded

into molecular representations, which are subsequently processed by a Feedforward Neural Network (FNN) to predict molecular properties, offering a powerful approach for capturing intricate chemical information. For D-MPNN, the Chemprop package is used [86–88]. For training the ML models, Henry coefficients are computed for 1110 isomers in each zeolite. The dataset consists of linear (C_1 to C_{30}) and branched (C_4 to C_{20}) alkanes. There are 30 linear, 70 mono-methyl, 435 di-methyl, and 29 tri-methyl isomers, with the remainder consisting of multi-branched isomers containing ethyl, methyl, propyl, and isopropyl groups. The corresponding datasets can be found in the Supporting Information SI2 folder of Ref. [493]. Both D-MPNN and TabPFN provide better predictions (i.e., exhibiting a larger correlation coefficient, R^2) than the other models. In comparison to TabPFN, D-MPNN provides more accurate predictions for isomers with small Henry coefficients. The effect of active learning is analyzed to select structurally diverse isomers from the training set. This is an efficient way of exploring chemical space to identify diverse molecules for the training dataset and helps in achieving better predictions with fewer training data points. Activity cliff analysis was conducted on the Henry coefficients of alkanes in MTT-type zeolites. Oversampling isomers associated with high activity cliffs in the training set led to a modest improvement in predictive performance, increasing the R^2 value by approximately 4% from 0.76 to 0.79. A more in-depth analysis is needed to achieve substantial enhancements in model accuracy. The predicted Henry coefficients, together with thermochemical properties of alkanes obtained from our previously developed linear regression model [435], are used to calculate the reaction equilibrium distribution of C_{16} isomers in MTW-type zeolite. The resulting distributions indicate that linear, mono-methyl, di-methyl, tri-methyl, tetra-methyl, and mono-ethyl substituted isomers are the most favoured groups of isomers in this zeolite.

This chapter is organized as follows: Section 2 contains simulation details for Henry coefficients of alkanes in zeolites, important concepts and algorithms behind isomer enumeration, ML models, active learning, activity cliffs, and reaction equilibrium distribution. Our main results are discussed in Section 3, which includes a comparison between different ML models for predicting Henry coefficients, the use of active learning to efficiently design the training dataset, the activity cliff analysis for alkanes in MTT-type zeolite, and the reaction equilibrium distribution of C_{16} isomers in one-dimensional zeolites. Section 4 provides concluding remarks on the performance of these ML models, room for future improvement for better predictive power, and their application to zeolite shape selectivity for hydroisomerization.

7.2. Methodology

Alkane isomers are generated for structures containing up to (iso)propyl groups. We did not exclude isomers based on their potential reaction pathways. For instance, some isomers may crack rapidly and are unlikely to

appear in the product distribution. Isomers with larger branches are excluded due to the very unfavorable formation free energy of such alkanes in narrow-pore zeolites [435]. The scripts to generate these isomers are provided in the Supporting Information SI1_py.py and SI1_cpp.cpp, and the lists of generated isomers are available in the Supporting Information SI1_isolist.xlsx. The procedure for isomer generation is shown in Algorithm 4.

Henry coefficients for the training and testing data sets are calculated using the Widom test particle insertion method [46], in combination with the Configurational-Bias Monte Carlo (CBMC) technique [47, 48] using the RASPA2 software [39, 133, 464]. In the Widom insertion test particle method, a virtual alkane molecule is randomly inserted into the zeolite simulation box at infinite dilution, and the excess chemical potential is estimated from the ensemble average of the Boltzmann factor of the Rosenbluth weight. The Henry coefficient (k_H) is computed using:

$$k_H = \lim_{P \to 0} \frac{q}{P} = \frac{1}{\rho_{framework} RT} \frac{\langle W_{zero} \rangle}{\langle W_{IG} \rangle} \tag{7.1}$$

In Eq. 7.1, q is the amount adsorbed per unit mass of the adsorbent and P is the pressure of the adsorbate in the fluid phase. W_{zeo} is the Rosenbluth weight of the alkane inside the zeolite pores and W_{IG} is the Rosenbluth weight of the isolated alkane molecule in the ideal gas phase. $\langle\ \rangle$ denotes ensemble averages. $\rho_{framework}$ is the zeolite framework density, R is the universal gas constant, and T is the temperature. Alkanes are modeled as united-atoms [430], which offer a favourable balance between accuracy and computational efficiency [37]. In this model, united atoms C, CH, CH_2, and CH_3 are treated as charge-neutral, allowing the omission of Coulombic interactions [39]. The intramolecular bonded interactions are described using the TraPPE united-atom force field [37]. Non-bonded interactions, both between adsorbent and adsorbate as well as inside adsorbate molecules, are modelled using Lennard-Jones potentials [40]. Lennard-Jones parameters for alkanes are obtained from Dubbeldam et al [136], while adsorbent–adsorbate interactions are described using the TraPPE-zeo force field [43]. In this work, all silica zeolites are treated as rigid frameworks [44], since zeolite flexibility has a very small effect on adsorption behaviour, particularly at infinite dilution. Lennard-Jones interactions are truncated and shifted at 12 Å without tail corrections. Cross-interactions between different atom types are handled using the Lorentz–Berthelot mixing rules [41, 42]. For each case, we performed 5,000,000 Monte Carlo cycles. In these simulations, a cycle consists of N trial moves, where N is the number of molecules, with a minimum of 20 steps. This ensures that, on average, one Monte Carlo trial move (successful or unsuccessful) is attempted for each molecule per cycle. The number of trial positions during the growth of the alkane chain is considered to be 10. Further simulation details are provided in the Appendix A.2 of this thesis. Table 7.2 summarizes details on the simulation box for each zeolite,

Algorithm 4 Procedure to generate structural isomers of alkanes with methyl, ethyl, propyl, and isopropyl branches. The isomer generation scripts, SI1_python.cpp and SI1_cpp.cpp can be found in the folder SI1 of the Supplementary Information.

1: Define a set of atomic units representing alkyl groups.
   - C – carbon atom in the main chain
   - (C) – methyl group
   - (CC) – ethyl group
   - (CCC) – propyl group
   - (C(C)C) – isopropyl group
2: Generate a list of these atomic units based on the specified input for total chain length and number of different types of branches.
3: Generate all possible permutations of the atomic unit list, ensuring the first two and the last positions are fixed with C atoms, as branching is not allowed at the terminal or near-terminal carbons.
4: Discard any permutation with more than two consecutive branches which are impossible to form as two out of four bonds of C atoms are connected to the neighbouring C atoms in the main chain.
5: For each remaining permutation, create two lists: one indicating the positions of branches in the main chain, and the other specifying the corresponding types of branches (methyl, ethyl, propyl, and isopropyl) at those positions. These lists are generated in both forward and backward directions. The position list with the smallest number at the first point of difference is chosen. If two or more side branches are in equivalent positions, the lowest number is assigned to the one which will come first in the name.
6: Discard permutations with branches at chemically invalid positions. For example, methyl groups cannot be attached to the terminal C atoms, and ethyl groups cannot be placed at the first, second, and last atoms in the main chain.
7: Generate IUPAC names [471] and corresponding SMILES strings [449] for the valid structures.
8: Identify and eliminate any duplicate entries.

such as the number of unit cells, box dimensions, and void fractions. A sufficiently large simulation box is used to eliminate finite-size effects, which is particularly important for long-chain alkanes confined in narrow zeolite pores. The Python scripts provided in Ref. [435] are used to automatically convert SMILES strings to input force field files for RASPA2 software.

To predict Henry coefficients using ML, both descriptor- and graph neural network-based models are used. For descriptor-based ML models, three main types of descriptors are used: (1) total chain length, which is the total

Table 7.2. Total number of unit cells for MTT-, MTW-, MRE-, and AFI-type zeolites used in the simulations to compute the Henry coefficients for alkanes, along with the dimensions and void fractions of these zeolites. The pore dimensions are obtained from the International Zeolite Association (IZA) [18] and the void fractions are obtained from the iRASPA software [22].

zeolite	number of unit cells	unit cell dimension/ [Å]			void fraction/ [-]
		a	*b*	*c*	
MTW-type	162 (2 × 27 × 3)	25.55	5.26	12.12	0.24
MTT-type	162 (27 × 2 × 3)	5.26	22.03	11.38	0.095
MRE-type	108 (18 × 3 × 2)	8.26	14.56	20.31	0.17
AFI-type	153 (3 × 3 × 17)	13.83	13.83	8.58	0.29

number of carbon atoms in the isomer, (2) main chain length, which is the total number of carbon atoms in the isomer excluding the branches, and (3) the number of methyl, ethyl, propyl, or isopropyl groups at each position in the main chain. These descriptors are input features for the RF, XGB, CB, and TabPFN models. The RF model is implemented using the Scikit-learn library [494], while the CB and XGB models are implemented using the CatBoost [85] and XGBoost [83] libraries, respectively. The D-MPNN model requires only SMILES strings as input. In this study, we used Chemprop [86–88], a software package for message passing neural networks, to implement the D-MPNN. Molecular graphs are constructed using atom and bond features generated by RDKit [495]. Atom-level features include atomic number, degree, formal charge, chiral tag, number of hydrogens, hybridization, aromaticity, and atomic mass. Bond-level features include indicators such as bond existence, bond type, conjugation, indicator of ring, and stereochemistry. As an example, the SMILES string and the descriptors for three example isomers are shown in Table 7.3. To ensure consistency and reproducibility, the same random splits of the datasets are used in all models. The dataset was divided into training, validation, and test sets (0.72:0.08:0.2). The data in the validation set was interchanged with data from the training set to identify the best split. The test set was never part of the training.

To evaluate the size of the training set needed to achieve satisfactory accuracy in predicting Henry coefficients for a diverse range of alkane isomers in zeolites, we compared a random selection strategy and an active learning strategy [496] for building successively larger training datasets. The active learning algorithm uses Gaussian process regression with a marginalized graph kernel (GPR-MGK) [496] as the surrogate model. At each iteration, a GPR-MGK model is constructed from the current training set to estimate the predictive uncertainties of the remaining molecules, and the molecule with the largest uncertainty is added to the training set until 50 new isomers are identified. To provide some intuition about the selections by the active

Table 7.3. Typical descriptors for a few alkane isomers (3-e-2-m-C_6, 4-p-C_7, and 2-m-4-p-C_7) used in training ML models to predict Henry coefficients for alkanes in zeolites. e, m, p, and ip stand for ethyl, methyl, propyl, and isopropyl groups. Total chain (total number of C atoms in the isomer), main chain (total number of C atoms in the main chain of the isomer), and number of different types of branches at each C atom in the main chain are used for descriptor-based ML models (Random Forest, Extreme Gradient Boosting, Cat Boost, and Tabular Prior Fitted Network). SMILES strings are used as input features in the Directed Message Passing Neural Network.

Isomer	SMILES	Total chain	Main chain	Methyl positions					Ethyl positions					Propyl positions				
				2	3	4	...	19	2	3	4	...	19	2	3	4	...	19
3-e-2-m-C_6	CC(C)C(CC)CCC	9	6	1	0	0	...	0	0	1	0	...	0	0	0	0	...	0
4-p-C_7	CCCC(CCC)CCC	10	7	0	0	0	...	0	0	0	0	...	0	0	0	1	...	0
2-m-4-p-C_7	CC(C)CC(CCC)CCC	11	7	1	0	0	...	0	0	0	0	...	0	0	0	1	...	0

learning algorithm, the initial 50 molecular structures selected by active learning from the training sets, comprising linear alkanes (C_1–C_{30}) and methyl-branched alkanes (C_4–C_{20}), as well as linear alkanes (C_1–C_{30}) and methyl-, ethyl-, propyl-, and isopropyl-branched alkanes (C_4–C_{20}) are shown in Figs. S10 a and b in the Supporting Information SI3.pdf. Following each active learning iteration, the D-MPNN and the TabPFN models were retrained on the expanded training set. To ensure robust training and prevent overfitting across the different dataset sizes, the D-MPNN model was trained with a consistent set of hyperparameters. We employed a dropout rate of 0.2 in the 3-layer feed-forward network and an early stopping protocol with a patience of 25 epochs based on the validation loss. Furthermore, a batch size of 50 was used to ensure stable gradient updates and prevent instabilities from uneven final batches in the dataloader. The model performance at each iteration was evaluated on the same full validation and test sets as described above.

The Structure Activity Landscape Index (*SALI*) is used to quantify activity cliffs in isomers for Henry coefficients. This index is a pairwise score that captures the magnitude of the property change with respect to the distance of two compounds in the chemical space [481].

$$SALI = \frac{|-\ln(k_{H,i}) - (-\ln(k_{H,j}))|}{d_{i,j}} \tag{7.2}$$

In Eq. 7.2, $d_{i,j}$ is the structural distance between a pair of isomers i and j. In this work, Levenshtein distance [486] between SMILES strings is used as $d_{i,j}$. This distance quantifies dissimilarities between two strings by counting the minimum number of single-character edits, which include insertions, deletions, or substitutions, required to transform one string into the other. Pairwise activity cliffs measured using *SALI* are computed for the Henry coefficients of alkane isomers adsorbed in MTT-type zeolites at 523 K. Based on the *SALI* values, the dataset is divided into low and high activity cliff subsets. 70% of the low cliff data and 20% of the high cliff data are added to the training set randomly. Oversampling is performed by duplicating the high-cliff data. A comparison is made between the predictions of TabPFN models with and without oversampling of high-activity cliff data. The detailed procedure is presented in Algorithm 5.

The predicted Henry coefficients for C_{16} isomers at 523 K are used to compute the reaction equilibrium distribution for hydroisomerization of linear C_{16} into branched isomers at infinite dilution in MTW-, MTT-, MRE-, and AFI-type zeolites. The chemical reaction equilibrium distribution is obtained by imposing a gas-phase reaction equilibrium for the alkane isomers and a simultaneous phase equilibrium between the gas and the adsorbed phase for each component [390, 435]. This satisfies the reaction equilibrium distribution in the zeolites [408], which is typically valid at low conversions of hydrocarbons. The adsorbed phase loadings for these isomers are modelled using Henry's law. The Henry coefficients are predicted using

Algorithm 5 Procedure to analyze the effect of oversampling of high activity cliff data on the performance of the TabPFN model. As a case study, isomers containing only methyl groups in MTT-type zeolite are considered.

1: Compute the pairwise structural distance $d_{i,j}$ between isomers present in the original training set using Levenshtein distance.
2: Select isomer pairs with $d_{i,j}$ below a predefined threshold. For isomers containing only methyl groups in MTT-type zeolite, a Levenshtein distance threshold of three was applied to define structural similarity.
3: Compute the Structure Activity Landscape Index ($SALI$) for the selected isomer pairs.
4: Identify the maximum value of SALI computed for these pairs.
5: Select isomer pairs with SALI larger than 30% of the maximum value of SALI. Label these pairs as high activity cliff pairs. Divide the original dataset into high and low cliff data.
6: Add 20% of the high activity cliff data to the training set.
7: Oversample this subset by duplicating rows.
8: Compare the performance of TabPFN model with and without oversampling of the high cliff data in the training dataset.

the ML models. Henry coefficients are predicted using the D-MPNN model for MTT- and MRE-type zeolites, while the TabPFN model is used for the MTW- and AFI-type zeolites. For further details on this method, the reader is referred to Refs [390, 435]. The gas-phase reaction equilibrium distribution is computed using the Gibbs free energy ($G_0(T) - H_0(0\ \mathrm{K})$) at the operating temperature T (523 K) and enthalpy of formation ($\Delta_f H_0(0\ \mathrm{K})$) at 0 K [390]. These properties are predicted using our linear regression model [435] based on a second-order group contribution method. The reaction equilibrium distribution is characterized by the selectivities of the isomers relative to the linear alkane in the adsorbed phase. Absolute selectivities are defined as the ratio of the mole fraction of a component to the sum of the mole fractions of all other components in the adsorbed phase [429]. Relative selectivities compare the absolute selectivity of a specific isomer to that of the corresponding linear alkane, providing a measure of preferential adsorption. For further details, the reader is referred to Eqs. 4.21-4.24 in chapter 4.

7.3. Results and Discussion

Fig. 7.2a shows coefficients of determination, R^2 of D-MPNN, RF, XGB, CB, and TabPFN for predicting the negative logarithm of Henry coefficients, $-\ln(k_H)$, for linear (C_1–C_{30}) and methyl-branched (C_4–C_{20}) alkanes in MTT-, MTW-, MRE-, and AFI-type zeolites at 523 K. Error bars are small for most alkane isomers, indicating reliable estimates of the Henry coefficients. Small error bars are essential because such precision ensures that the training

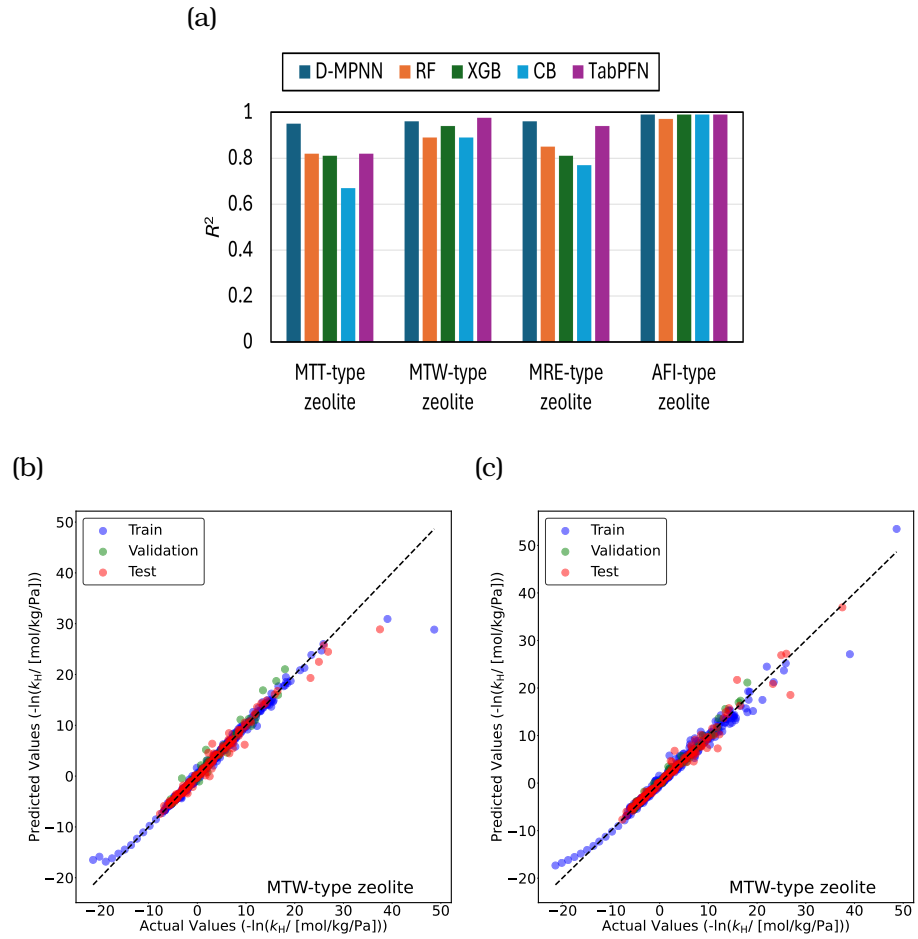


Figure 7.2. (a) Coefficients of determination (R^2) for Directed Message Passing Neural Network (D-MPNN), Random Forest (RF), Extreme Gradient Boosting (XGB), Cat Boost (CB), and Tabular Prior Fitted Network (TabPFN) models predicting the negative logarithm of Henry coefficients, $-\ln(k_H)$, for linear alkanes (C_1–C_{30}) and methyl-branched alkanes (C_4–C_{20}) in MTT-, MTW-, MRE-, and AFI-type zeolites at 523 K. The unit of k_H is moles/kg framework/Pa. Models are trained separately for each zeolite. Parity plots for predictions of $-\ln(k_H)$ in MTW-type zeolite at 523 K using (b) TabPFN and (c) D-MPNN models. Blue circles indicate training isomers, and green circles represent validation isomers. The random seed is varied to identify the most suitable split between the training and the validation datasets. Red circles represent the test isomers, which are never part of the training set. The standard deviations for the actual values of $-\ln(k_H)$ are too small to plot.

data supplied is reliable and accurately represents the underlying trends, thereby improving model generalization. For highly branched isomers, the error bars can be relatively large compared to the Henry coefficients, due to poor fit inside the zeolite pores and challenges in achieving convergence in Monte Carlo simulations. These highly branched isomers are of limited practical relevance, as they occur only in negligible amounts in the product distribution of hydroisomerization reactions. Since operating pressures are at most 20 bar, Henry coefficients on the order of 10^{-40} or 10^{-80} will both lead to virtually zero adsorption, making the exact value irrelevant. Further details on the error bars are provided in Appendix A.4. $-\ln(k_H)$ is chosen instead of k_H because Henry coefficients of alkanes vary orders of magnitude. ML models perform better on an evenly distributed dataset where the target variable has a Gaussian-like distribution [497]. This can be achieved using $-\ln(k_H)$ instead of k_H. D-MPNN and TabPFN perform better than the RF, XGB, and CB models. In AFI-type zeolite, all ML models achieve high accuracy with R^2 values exceeding ca. 0.98. TabPFN outperforms other models for MTW-type zeolites with 0.98 R^2 and D-MPNN delivers superior performance in MTT- and MRE-type zeolites with R^2 values of 0.95 and 0.96 respectively. Figs. 7.2b and 7.2c show the parity plots predicted by TabPFN and D-MPNN for linear (C_1–C_{30}) and methyl-branched alkanes (C_4–C_{20}) in MTW-type zeolite at 523 K. These plots provide predictions for isomers present in the training (blue circles), validation (green circles), and test (red circles) sets. For alkanes larger than C_{20}, predictions by both TabPFN and D-MPNN deviate from the actual values due to lack of training data in this range. D-MPNN performs better than TabPFN for low Henry coefficients or high values for $-\ln(k_H)$. These are usually highly branched isomers. For applications such as hydroisomerization, highly branched isomers will not form inside the narrow pores of one-dimensional zeolites such as MTW-, MTT-, and MRE-type zeolites [390, 435]. The parity plots for MTT-, MRE- and AFI-type zeolites are included in Appendix A.4.

Fig. 7.3a shows the coefficient of determination, R^2 for D-MPNN, RF, XGB, CB, and TabPFN for predicting $-\ln(k_H)$ for linear (C_1–C_{30}) and methyl-, ethyl-, propyl-, and isopropyl-branched (C_4–C_{20}) alkanes in MTT-, MTW-, MRE-, and AFI-type zeolites at 523 K. For MTT-type zeolite, XGB performs slightly better (larger R^2) than the other ML models. For MTW- and AFI-type zeolites, TabPFN provides better predictions with R^2 values of ca. 0.96 and 0.94, respectively. D-MPNN performs better for MRE-type zeolite with R^2 ca. 0.95. Figs. 7.3b and 7.3c show the parity plots for $-\ln(k_H)$ predicted by TabPFN and D-MPNN for linear (C_1–C_{30}) and methyl-, ethyl-, propyl-, and isopropyl-branched alkanes (C_4–C_{20}) in MTW-type zeolite. Similar to the dataset with all methyl groups, D-MPNN performs better at low Henry coefficients compared to TabPFN when alkanes with ethyl, propyl, and isopropyl branches are introduced to the dataset. Each model provides reasonable predictions. It is difficult to claim a single best model for predicting Henry coefficients. TabPFN and D-MPNN have shown more consistency in predictions

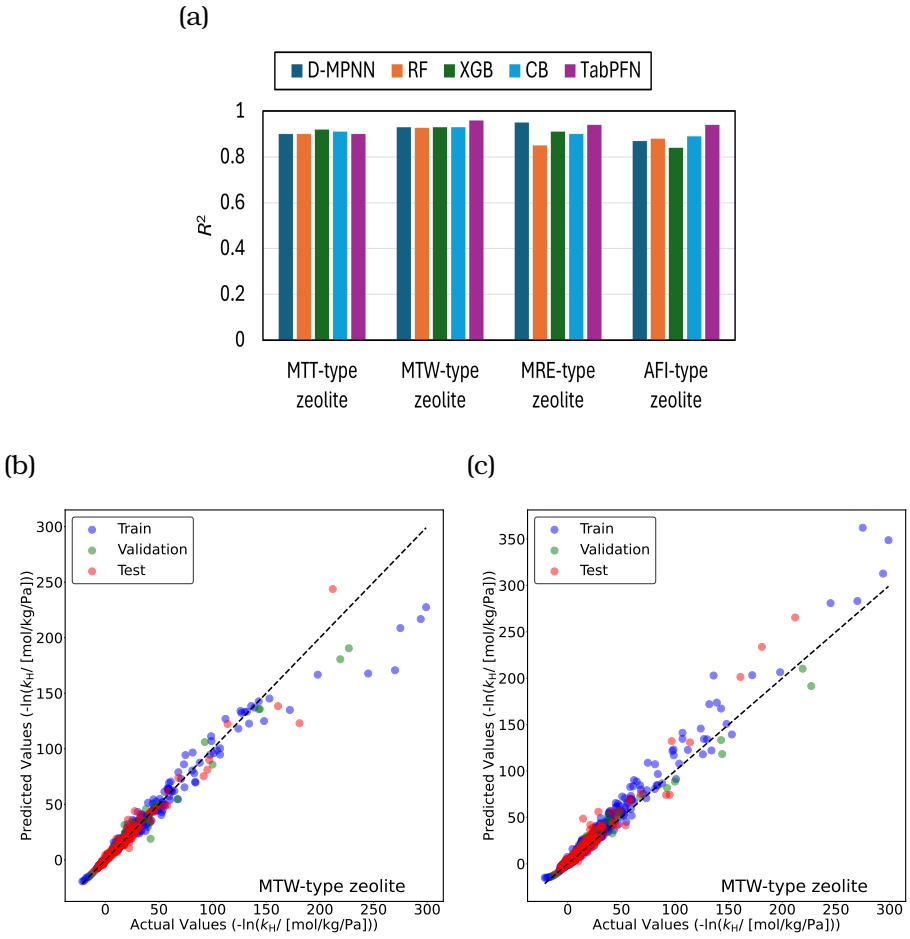


Figure 7.3. (a) Coefficients of determination, R^2 for Directed Message Passing Neural Network (D-MPNN), Random Forest (RF), Extreme Gradient Boosting (XGB), Cat Boost (CB), and Tabular Prior Fitted Network (TabPFN) models predicting the negative logarithm of Henry coefficients, $-\ln(k_H)$, for linear alkanes (C_1–C_{30}) and methyl-, ethyl-, propyl-, and isopropyl-branched alkanes (C_4–C_{20}) in MTT-, MTW-, MRE-, and AFI-type zeolites at 523 K. ML models are trained separately for each zeolite. Parity plots for $-\ln(k_H)$ predictions in MTW-type zeolite at 523 K using (b) TabPFN and (c) D-MPNN models. Blue circles indicate training isomers, and green circles represent validation isomers. The random seed is varied to identify the most suitable split between the training and the validation datasets. Red circles represent the test isomers, which are never part of the training set. The standard deviations for the actual values of $-\ln(k_H)$ are too small to plot.

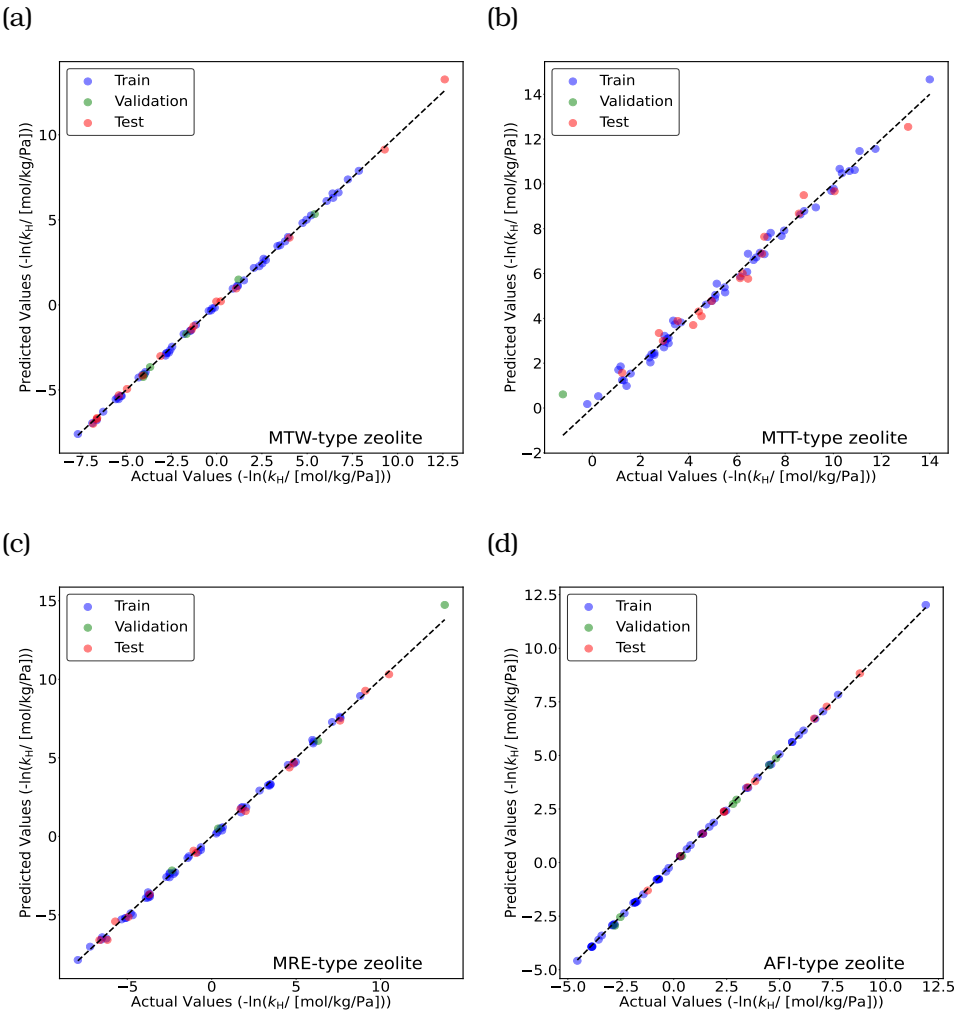


Figure 7.4. Parity plots for predicting -ln(k_H) for mono-methyl alkanes (C_4–C_{20}) in (a) MTW-, (b) MTT-, (c) MRE-, and (d) AFI-type zeolites at 523 K using the TabPFN model. The training datasets contain linear (C_1-C_{30}) and methyl-, ethyl-, propyl-, and isopropyl-branched (C_4-C_{20}) alkanes. Blue circles indicate training isomers, and green circles represent validation isomers. The random seed is varied to identify the most suitable split between the training and the validation datasets. Red circles represent the test isomers, which are never part of the training set. The standard deviations for the actual values of -ln(k_H) are too small to plot.

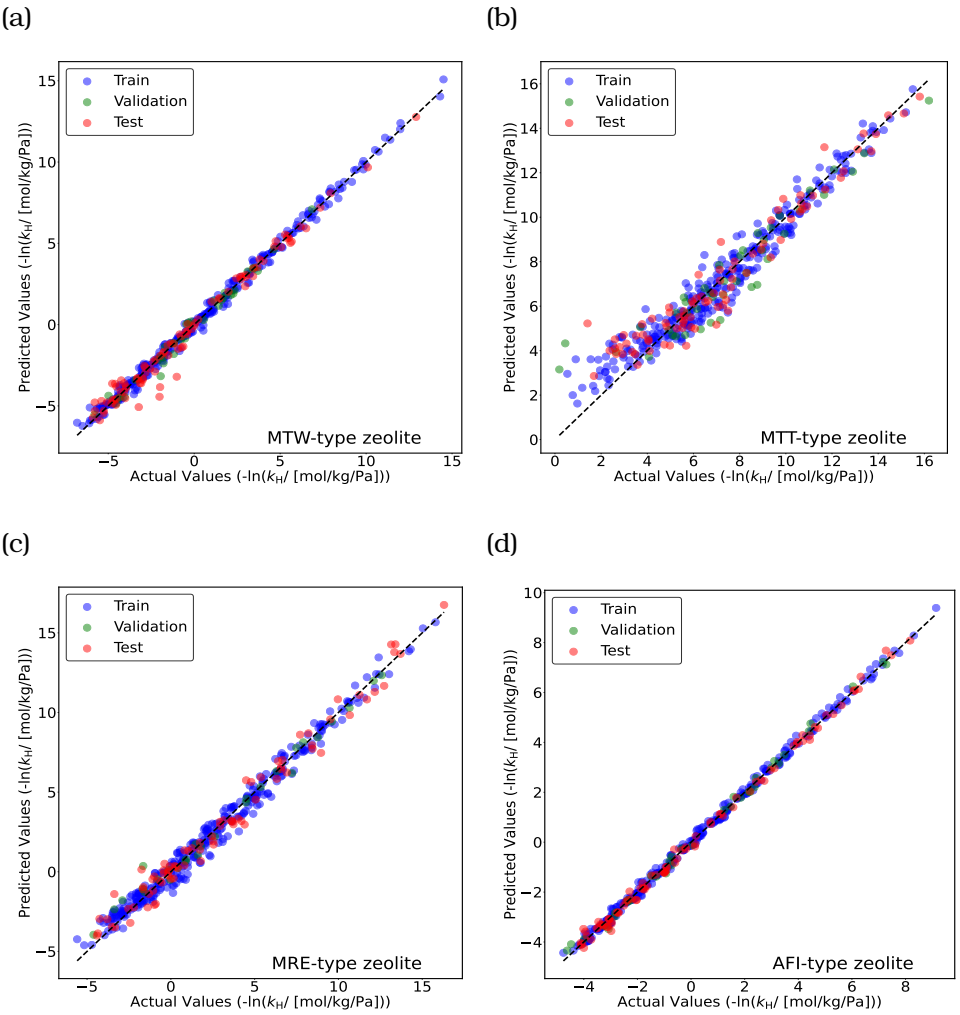


Figure 7.5. Parity plots for predicting -ln(k_H), for di-methyl alkanes (C_5–C_{20}) in (a) MTW-, (b) MTT-, (c) MRE-, and (d) AFI-type zeolites at 523 K using the TabPFN model. he training datasets contain linear (C_1-C_{30}) and methyl-, ethyl-, propyl-, and isopropyl-branched (C_4-C_{20}) alkanes. Blue circles indicate training isomers, and green circles represent validation isomers. The random seed is varied to identify the most suitable split between the training and the validation datasets. Red circles represent the test isomers, which are never part of the training set. The standard deviations for the actual values of -ln(k_H) are too small to plot.

compared to the other tree-based models (Figs. 7.2a and 7.3a). The parity plots for MTT-, MRE- and AFI-type zeolites are included in the Supporting Information SI3.pdf (Appendix A.4). The parity plots for mono-methyl and di-methyl alkanes obtained using the TabPFN model are shown in Figs. 7.4 and 7.5. TabPFN provides excellent predictions for mono- and di-methyl alkanes in MTW-, MTT-, MRE-, and AFI-type zeolites ($R^2 > 0.92$). To further improve the predictability of the models, a larger dataset is required, and structurally diverse isomers need to be included in the training set using advanced methods such as active learning.

In Figs. 7.6 and 7.7, the influence of training set size on the predictive accuracy of the TabPFN and D-MPNN models for $-\ln(k_H)$ is shown for linear (C_1–C_{30}) and branched (methyl-, ethyl-, propyl-, and isopropyl-, C_4–C_{20}) alkanes in MTW- and MTT-type zeolites at 523 K, comparing active learning and random selection strategies. For the TabPFN model (Fig. 7.6), active learning consistently resulted in higher predictive accuracy for both zeolite types. In case of MTW, model performance stabilized beyond 250 training points, with R^2 reaching ca. 0.975, whereas random selection exhibited substantial fluctuations, yielding R^2 values of only ca. 0.88 even at 500 training points (Figs. 7.6a and 7.6b). A similar trend was observed for MTT, where active learning achieved stable predictions of $-\ln(k_H)$ beyond 250 training points, with R^2 of ca. 0.92 and RMSE of ca. 10, clearly outperforming random selection (Figs. 7.6c and 7.6d). In contrast, a random selection of 600 training points in MTT yielded an R^2 of only 0.67. For the D-MPNN model (Fig. 7.7), the primary advantage of active learning was the increased stability and consistency of the training process. While the mean performance of random selection occasionally caught or exceeded that of active learning at certain data sizes due to statistical chance, its performance was unreliable, characterized by large fluctuations and high variance (indicated by the large error bars in Fig. 7.7 especially 650 data set size for MTT-type zeolite). In contrast, the active learning provided a more monotonic and reliable improvement, having either comparable performance in the larger dataset size or better performance in the smaller data set size. For example, in the MTW-type zeolite (Figs. 7.7a and 7.7b), active learning steadily converged to a stable R^2 of 0.95, whereas the random selection strategy remained erratic throughout the process. A direct comparison between the two models reveals a key difference in their data efficiency. The TabPFN model demonstrated remarkable performance even with very small training sets. For instance, in the MTW zeolite, TabPFN achieved an R^2 of approximately 0.95 with only 150 training points (Fig. 7.6a). In contrast, the D-MPNN model required around 350 data points to stably reach a similar level of accuracy (Fig. 7.7a). This trend suggests that while both models can ultimately achieve high predictive accuracy, TabPFN is significantly more data-efficient, making it a strong candidate for scenarios where labeled data is particularly scarce. This higher data efficiency can be partly attributed to the strong prior knowledge encoded in TabPFN's

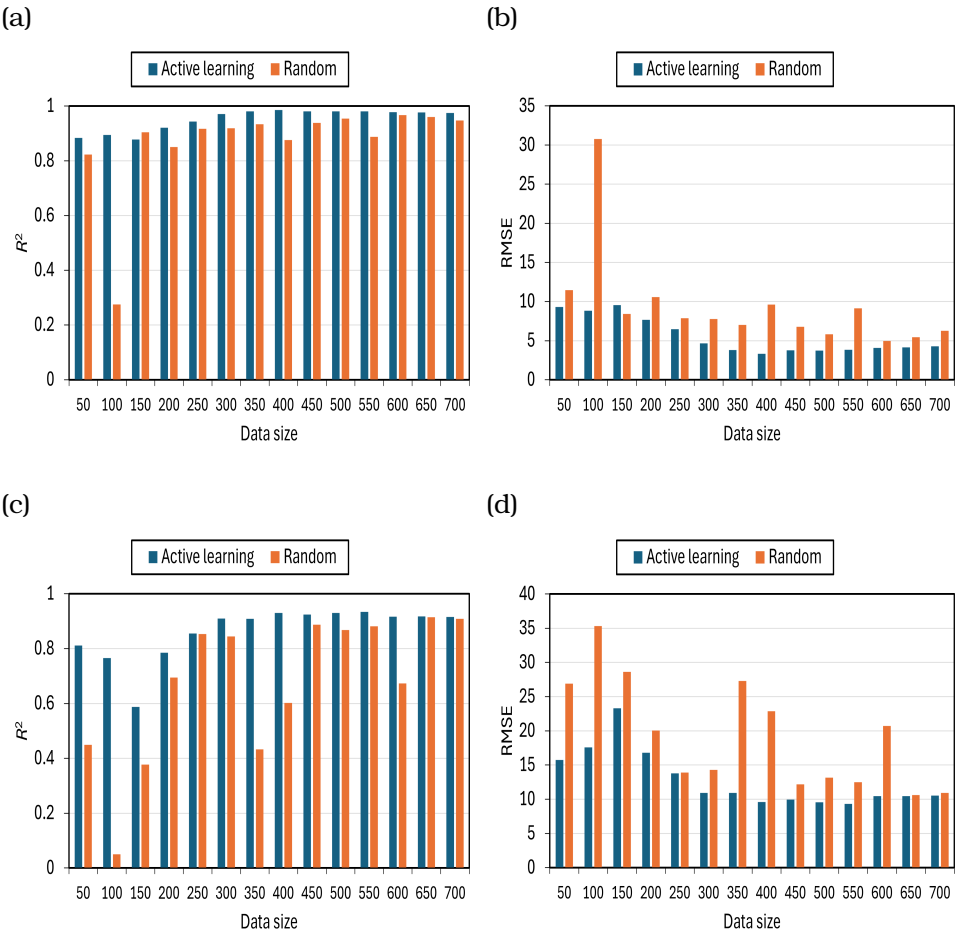


Figure 7.6. Test accuracies of the TabPFN model as a function of training set size, comparing an active learning strategy (blue bars) and a random selection strategy (orange bars). R^2 and RMSE values are shown for models trained on the negative logarithm of Henry coefficients for linear (C_1–C_{30}) and methyl-, ethyl-, propyl-, and isopropyl-branched (C_4–C_{20}) alkanes in MTW-type (a, b) and MTT-type (c, d) zeolites at 523 K.

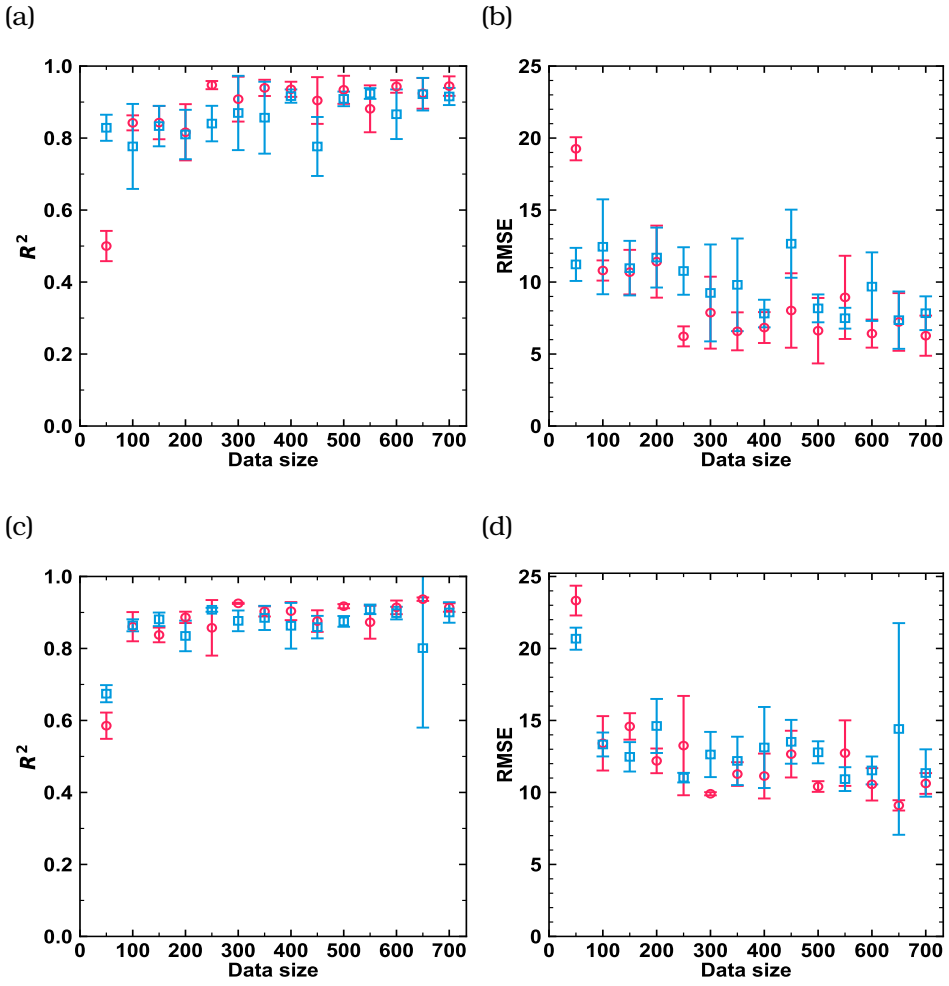


Figure 7.7. Test accuracies of the D-MPNN model as a function of training set size, comparing an active learning strategy (red circles) and a random selection strategy (blue squares). R^2 and RMSE values are shown for models trained on the negative logarithm of Henry coefficients for linear (C_1–C_{30}) and methyl-, ethyl-, propyl-, and isopropyl-branched (C_4–C_{20}) alkanes in MTW-type (a, b) and MTT-type (c, d) zeolites at 523 K.

Table 7.4. Pronounced activity cliffs in Henry coefficients for alkane isomer pairs with small variations in branching positions in MTT-type zeolite at 523 K.

Isomer	Henry coefficient/ [mol/kg/Pa]	Isomer	Henry coefficient/ [mol/kg/Pa]
5,5-m-C_{10}	1.54×10^{-4}	5,6-m-C_{10}	3.42×10^{-6}
4,4-m-C_{14}	1.38×10^{-3}	4,5-m-C_{14}	8.66×10^{-5}
2,3,3-m-C_6	7.73×10^{-10}	2,3,5-m-C_6	1.51×10^{-7}

Transformer architecture [90]. Overall, for both the TabPFN and the D-MPNN models, the active learning strategy delivered a more systematic and robust path to high performance. It was characterized by lower variance and a more reliable improvement in R^2 and RMSE values, making it a superior strategy for efficiently training models with limited labeled data compared to the less reliable, high-variance outcomes of random selection. A notable exception to this trend occurs at the smallest training set size of 50 for the D-MPNN model, where random selection exhibits a higher initial accuracy. This phenomenon can also be found for models trained on linear (C_1–C_{30}) and methyl-branched (C_4–C_{20}) alkanes, as shown in Figs. S9 in the Supporting Information SI3.pdf. This is a known phenomenon in active learning referred to as the cold start problem [498]. At this early stage, the active learning strategy prioritizes exploration by selecting diverse, high-uncertainty samples to efficiently map the feature space. While crucial for long-term performance, these samples can be initially challenging for the neural networks to learn. Conversely, random selection can, by chance, sample a more homogeneous cluster of simpler alkanes, leading to a deceptively strong initial performance.

Figs. 7.8a and 7.8b show the parity plots for $-\ln(k_H)$ predicted by TabPFN and D-MPNN for linear (C_1 – C_{30}) and methyl-branched alkanes (C_4 – C_{20}) in MTT-type zeolites. Large deviations in low predicted Henry coefficients are observed with TabPFN compared to D-MPNN. This is more pronounced for highly branched isomers with large activity cliffs. These discrepancies arise due to the narrow, highly elliptical pores of the zeolite, causing significant differences in Henry coefficients even for minor structural variations in isomers. These cliffs violate the foundational assumption of many ML models, i.e., similar structures should yield similar properties. This assumption serves as the foundation for most quantitative structure–activity relationship (QSAR) [482] and ML methods in chemistry.

Table 7.5 shows the effect of oversampling of high activity cliff [481] data on TabPFN predictions in MTT-type zeolites. Three train/test splits were used by varying the random state. For each train-test split, 20% of the high cliff isomers were retained in the training set. Oversampling this subset resulted in a modest improvement in predictive accuracy, with an average increase in R^2 values by approximately 4% as shown in Table 7.5. Further

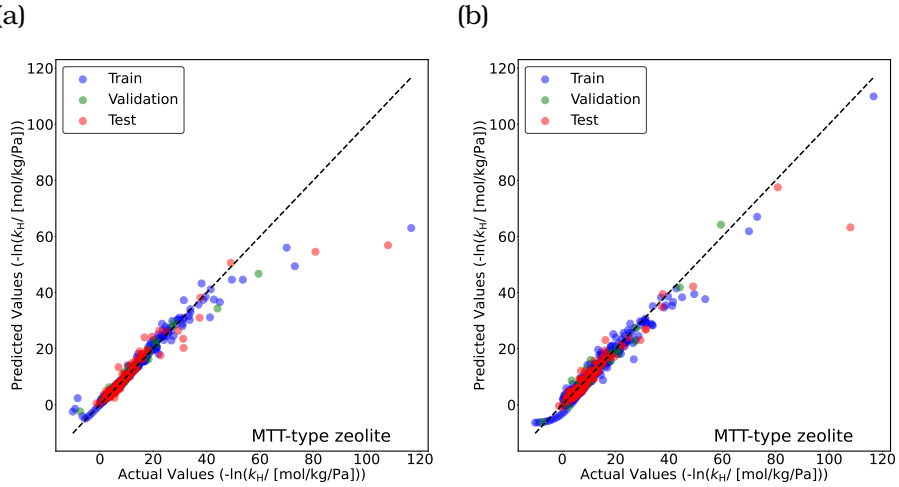


Figure 7.8. Parity plots for the negative logarithm of Henry coefficients, $-\ln(k_H)$ for linear alkanes (C_1–C_{30}) and methyl-branched alkanes (C_4–C_{20}) in MTT-type zeolite at 523 K predicted by (a) Tabular Prior Fitted Network (TabPFN) and (b) Directed Message Passing Neural Network (D-MPNN) models. Blue circles indicate training isomers, and green circles represent validation isomers. The random seed is varied to identify the most suitable split between the training and the validation datasets. Red circles represent the test isomers, which are never part of the training set. The standard deviations for the actual values of $-\ln(k_H)$ are too small to plot.

7

Table 7.5. Effect of oversampling 20% high activity cliff data on the prediction accuracy (R^2) of TabPFN for methyl-branched and linear alkanes in MTT-type zeolite at 523 K at different random states, representing different train-test splits and the corresponding average R^2. There is an average 3.74% increase in R^2 by oversampling 20% high activity cliff data.

dataset	R^2 at different random states				% increase in R^2
	0	42	90	average	
without oversampling	0.73	0.77	0.78	0.76	3.74
with oversampling	0.75	0.82	0.80	0.79	

investigation, including the application of more advanced techniques such as contrastive learning [499, 500], may be necessary to better address the challenges posed by activity cliffs. On the other hand, Henry coefficients for isomers exhibiting high activity cliffs may also be sensitive to the treatment of the rigidity of the zeolite framework.

Fig. 7.9a shows the variations in average Henry coefficients for different

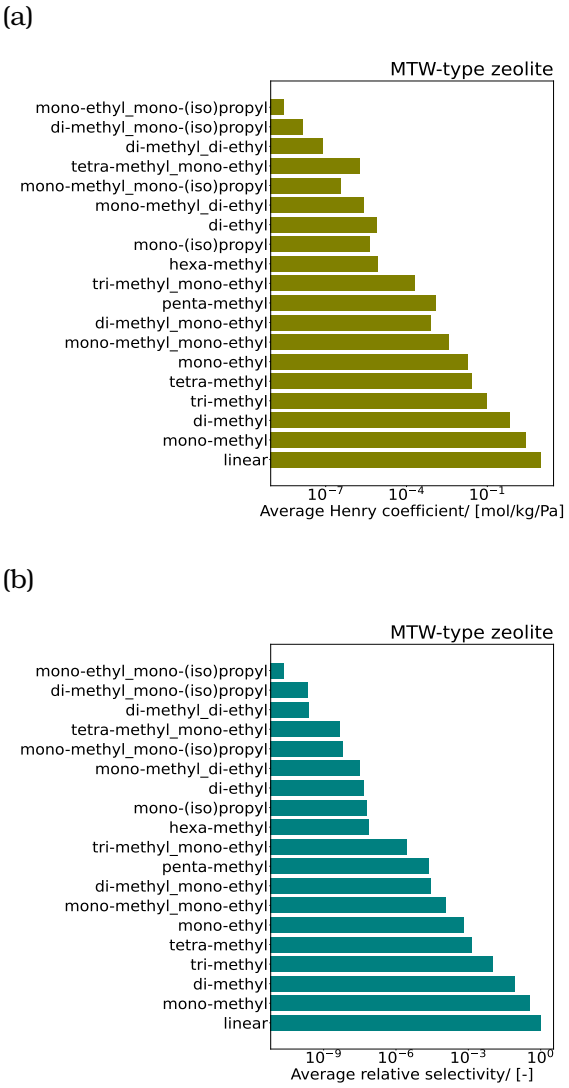


Figure 7.9. (a) Average Henry coefficients for different categories of C_{16} isomers in MTW-type zeolite at 523 K predicted using the TabPFN model. For each category, the Henry coefficients are averaged over all the isomers belonging to that category. (b) Average selectivities of different categories of C_{16} isomers relative to linear C_{16} at reaction equilibrium in MTW-type zeolite at 523 K. For mono-methyl isomers, 2-m-C_{15} has the largest Henry coefficient and the relative selectivity. Similarly, 2,13-m-C_{14} and 2,6,10-m-C_{13} have the largest selectivities for di- and tri-methyl isomers, respectively.

groups of C_{16} isomers in MTW-type zeolite at 523 K where the coefficients are averaged in each category. As a narrow-pore zeolite, MTW-type zeolite [18] preferentially adsorbs linear C_{16}, followed by isomers with mono-, di-, tri-, and tetra-methyl groups, as well as mono-ethyl substitutions. In sharp contrast, highly branched isomers, particularly those containing both mono-ethyl and mono-(iso)propyl groups, exhibit significantly smaller adsorption affinity. A similar trend is observed in the relative selectivities of C_{16} isomers for hydroisomerization at reaction equilibrium (Fig. 7.9b). In MTW, the adsorption strength, as reflected by the Henry coefficients, plays a more dominant role than the gas-phase thermochemical properties in determining the reaction equilibrium distribution. In MTT-, MRE-, and AFI-type zeolites, isomers with mono-, di-, tri-, and tetra-methyl and mono-ethyl branching are also the most favored groups in terms of relative selectivity for hydroisomerization of linear C_{16} at reaction equilibrium. Bar plots summarizing the average Henry coefficients and relative selectivities for C_{16} isomer groups in these zeolites are provided in SI3.pdf (Appendix A.4). The reaction equilibrium distribution of linear, mono-methyl (2-m-C_{15} to 8-m-C_{15}) and di-methyl (2,2-m-C_{14} to 2,13-m-C_{14}) isomers of C_{16} are shown in Fig. S11 in SI3.pdf (Fig. A.33 in Appendix A.4). Accurate prediction of Henry coefficients is essential for reliable computation of reaction equilibrium distributions. Therefore, a high coefficient of determination, R^2 serves as a critical performance metric for model evaluation. This work will be further extended to enhance the predictive performance of the ML models by incorporating larger datasets and using advanced techniques such as active learning that include a more diverse set of alkane isomers in the training set. Additionally, a generalized ML framework will be developed, which will be capable of predicting Henry coefficients for various alkanes in different types of zeolite structures.

7.4. Conclusions

This chapter presents an ML framework for predicting Henry coefficients of long-chain alkanes in one-dimensional zeolites using both descriptor- and graph-based models, as well as software for the fast enumeration of hydrocarbon isomers. For the evaluated models, TabPFN and D-MPNN demonstrate consistently high predictive performance, particularly for linear and moderately branched isomers. TabPFN provides excellent predictions for mono- and di-methyl alkanes in MTW-, MTT-, MRE-, and AFI-type zeolites. D-MPNN exhibits better accuracy for isomers with low Henry coefficients, which are typically highly branched and difficult to adsorb in narrow-pore zeolites. Challenges remain in accurately predicting Henry coefficients for highly branched isomers, primarily due to the presence of activity cliffs and limited training data diversity. Active learning strategies and oversampling of high-cliff data led to modest improvements in model accuracy, suggesting that these techniques have the potential to improve

model performance. Under-prediction of adsorption for bulky isomers may not significantly impact practical applications such as hydroisomerization, since such isomers are sterically excluded from narrow zeolite pores. The predicted Henry coefficients, combined with gas-phase thermodynamic properties derived from a second-order group contribution model, enable the computation of reaction equilibrium distributions for hydroisomerization processes. This integrated approach provides valuable insights into the thermodynamic versus kinetic contributions to isomer selectivity, aiding the development of lumped kinetic models and catalyst design. Future work will focus on expanding the dataset by incorporating more chemically diverse isomers, including those with ethyl, propyl, and isopropyl branches, using active learning and other advanced sampling techniques. The development of a general ML framework for predicting Henry coefficients of a broad range of alkanes in different zeolite topologies is envisioned. This will facilitate high-throughput screening and rational catalyst design for hydrocarbon upgrading applications. Another promising direction is the development of a two-step ML model: in the first step, an ML classifier determines whether an isomer can fit in a zeolite (finite Henry coefficient) or is effectively excluded (Henry coefficient approaching zero), while in the second step, a regression model for $-\ln(k_H)$ is trained only on the isomers that fit.

8

Conclusions

This thesis presents a comprehensive, multiscale investigation into the shape-selective hydroisomerization of long-chain alkanes in zeolites, focusing on how molecular structure, pore topology, and thermodynamics jointly influence adsorption and chemical reaction equilibria. By integrating molecular simulation, thermodynamic modeling, linear regression, and machine learning, this work offers both a mechanistic understanding and practical tools for catalyst design and process optimization. Molecular simulations are essential for studying adsorption in zeolites because the large number of alkane isomers and zeolite structures renders experimental characterization impractical, time-consuming, and costly. Force field-based simulations, such as classical Monte Carlo methods, offer an efficient and scalable alternative to quantum mechanical calculations, which are typically too computationally intensive for large systems. In addition, these simulations provide a fast and systematic approach to generating high-quality training data for machine learning models used in screening adsorption properties across a wide chemical space. Pure component adsorption isotherms for alkanes in zeolites can be efficiently computed using Grand-Canonical Monte Carlo (GCMC) simulations combined with Configurational-Bias Monte Carlo (CBMC), which effectively samples flexible alkane conformations in confined pores. However, simulating multicomponent adsorption remains computationally demanding and is not feasible for on-the-fly calculations in dynamic simulations such as breakthrough curve modeling. As a result, surrogate models or thermodynamic approximations are often needed to enable fast and accurate prediction of mixture adsorption during process simulations. A key contribution is the development of a segregated explicit isotherm (SEI) model, which captures adsorbate-size-dependent behavior in materials with heterogeneous adsorption sites. Merging the strengths of segregated Ideal Adsorbed Solution Theory (SIAST) and explicit mixture isotherms, SEI improves the accuracy and speed of adsorption predictions, particularly in dynamic simulations like breakthrough curves. To make these models accessible, the open-source software RUPTURA was developed, offering a

flexible platform for simulating fixed-bed adsorption and separation processes involving complex hydrocarbon mixtures. To assess how the zeolite pore structure influences hydroisomerisation equilibrium, a reaction equilibrium model was introduced that combines gas-phase thermodynamics with adsorption behavior. This model reveals how pore topology affects the thermodynamic stabilization of isomers, shifting equilibrium product distributions and enabling shape-selective catalysis. Capturing these equilibrium limits is important for differentiating the influence of thermodynamics from the kinetics of hydroisomerization. Gas-phase thermochemical properties of alkanes, enthalpies of formation and Gibbs energies are needed to compute the gas-phase equilibrium distribution. Due to the unavailability of these thermochemical properties for long-chain alkanes, a second-order group contribution model was developed to predict enthalpies and Gibbs free energies using linear regression. The absolute entropy was then obtained from these values using standard thermodynamic relations. These thermochemical properties enabled the computation of chemical reaction equilibrium distributions and also supported the evaluation of second-law efficiency via exergy destruction, a key metric for assessing the energy performance of hydrocarbon upgrading processes. In addition to the gas-phase thermochemical properties, Henry coefficients are essential for computing the reaction equilibrium distribution in hydroisomerization. Computing Henry coefficients for long-chain alkanes is challenging due to the large space of alkane isomers and zeolite structures. To overcome this, machine learning models were developed to predict Henry coefficients. Descriptor-based (Random Forest, Extreme Gradient Boost, Cat Boost, and TabPFN) and graph-based (D-MPNN) approaches were shown to complement each other, with TabPFN excelling for simpler isomers and D-MPNN for highly branched, low-adsorbing structures. While activity cliffs and data sparsity remain challenges, the use of active learning and targeted oversampling demonstrated potential for improving the predictive performance of the ML models. Importantly, the under-prediction of adsorption for bulky isomers is less relevant for hydroisomerisation, since these isomers are sterically hindered from entering the narrow zeolite pores. When combined with gas-phase thermodynamics, the predicted Henry coefficients enable rapid computation of equilibrium product distributions, supporting lumped kinetic modeling and rational catalyst design for hydroisomerization. This thesis introduces an integrated framework ranging from molecular-scale thermodynamic and adsorption modeling to data-driven predictions and process-level insights. By addressing key challenges, including surface heterogeneity, missing thermochemical data, and isomeric complexity, it provides a foundation for scalable, accurate modeling of hydroisomerisation at equilibrium conditions. The combination of physically interpretable models with machine learning opens pathways for high-throughput screening and data-guided process optimization. Beyond hydroisomerization, the approaches developed here are applicable to other catalytic and separation processes involving complex

hydrocarbon mixtures.

Overall, this thesis bridges the gap between molecular-level insights and process-scale applications, advancing both the fundamental understanding and practical engineering of zeolite-catalyzed hydroisomerisation. The models and tools developed here lay a strong foundation for data-driven design of selective, energy-efficient catalytic processes. Looking ahead, several avenues for future research can be built on the foundation laid in this thesis. Expanding the dataset to include a broader range of alkane isomers, particularly those with ethyl, propyl, and isopropyl branches, will improve the robustness of the models and make them more general. To overcome challenges related to data sparsity and activity cliffs, advanced sampling techniques such as uncertainty-aware active learning and contrastive learning may be used to selectively explore under-represented and high-error regions of the chemical space. Extending the machine learning framework to cover more zeolite topologies and conditions will enable high-throughput screening for hydroisomerisation and related hydrocarbon upgrading processes. Machine learning models can also be developed to directly predict complete pure component isotherms, or a hybrid approach can be pursued where maximum adsorption capacities are predicted via ML and combined with Henry coefficients to reconstruct full isotherms in a physically meaningful way. ML models for mixture adsorption isotherms can be trained to capture the competitive behaviour of hydrocarbons in complex zeolite structures, particularly important for simulating real-world mixtures where multicomponent effects dominate. Beyond equilibrium modeling, future efforts can focus on developing comprehensive reactor models that integrate adsorption, diffusion, and reaction kinetics. This can involve extending RUPTURA to support reactor simulations, incorporating reaction mechanisms such as hydrocracking as quasi-equilibrium or kinetically-controlled steps. Such an extension would allow for the simulation of fixed-bed or catalytic reactors at realistic operating conditions and provide deeper insights into catalyst performance beyond thermodynamic limits. Tuning zeolite frameworks for desired product distributions remains a central goal in catalyst design. Future work can focus on systematically linking structural features such as pore diameter, connectivity, and tortuosity to selectivity outcomes, supported by simulations and ML-based surrogate models. Moving beyond rigid framework approximations toward flexible framework modeling will provide more accurate predictions for adsorption- and diffusion-based properties, especially for branched or bulky isomers where pore breathing and structural relaxation can influence transport and selectivity. Finally, continued development and sharing of open-source software tools are essential for fostering transparency, reproducibility, and community-driven innovation in adsorption and catalytic process modeling. Tools like RUPTURA provide a flexible foundation that can be adapted to new chemistries, materials, and process models. As the field moves increasingly toward integrated, data-informed design strategies, such platforms will be critical for accelerat-

ing discovery and bridging the gap between molecular understanding and scalable process development.

Bibliography

[1] M. L. Gunawan, T. H. Novita, F. Aprialdi, D. Aulia, A. S. F. Nanda, C. B. Rasrendra, Z. Addarojah, D. Mujahidin, and G. T. M. Kadja. "Palm-oil transformation into green and clean biofuels: Recent advances in the zeolite-based catalytic technologies". *Bioresour. Technol. Rep.* (2023), p. 101546.

[2] C. Morley. "A fundamentally based correlation between alkane structure and octane number". *Combust. Sci. Technol.* 55 (1987), pp. 115–123.

[3] N. Pasadakis, V. Gaganis, and C. Foteinopoulos. "Octane number prediction for gasoline blends". *Fuel Process. Technol.* 87 (2006), pp. 505–509.

[4] A. Verkehrswende. *ANALYSIS-Defossilising aviation with e-SAF-An introduction to technologies, policies, and markets for*. 2024. URL: https://www.agora-verkehrswende.org/publications/scaling-power-to-x-fuels-in-transport.

[5] G. C. A. S. Bruce C. Gates James R. Katzer. *Chemistry of Catalytic Processes*. 1st ed. Chicago: McGraw-Hill, 1979.

[6] S. Baek, R. Field, M. Goldstone, P. Kirk, J. Lester, and R. Perry. "A review of atmospheric polycyclic aromatic hydrocarbons: sources, fate and behavior". *Water, Air, Soil Pollut.* 60 (1991), pp. 279–300.

[7] J. Biswas and I. Maxwell. "Octane enhancement in fluid catalytic cracking: I. Role of ZSM-5 addition and reactor temperature". *Appl. Catal.* 58 (1990), pp. 1–18.

[8] G. Karavalakis, D. Short, D. Vu, R. Russell, M. Hajbabaei, A. Asa-Awuku, and T. D. Durbin. "Evaluating the effects of aromatics content in gasoline on gaseous and particulate matter emissions from SI-PFI and SIDI vehicles". *Environ. Sci. Technol.* 49 (2015), pp. 7021–7031.

[9] J. B. Heywood. *Internal combustion engine fundamentals*. 2nd ed. New York: McGraw-Hill Education, 2018.

[10] M. Frenklach. "Reaction mechanism of soot formation in flames". *Phys. Chem. Chem. Phys.* 4 (2002), pp. 2028–2037.

[11] J. Holladay, Z. Abdullah, and J. Heyne. *Sustainable aviation fuel: Review of technical pathways*. Tech. rep. DOE EERE; Pacific Northwest National Lab.(PNNL), Richland, WA, United States. DOE EERE; Pacific Northwest National Lab.(PNNL), Richland, WA (United States), 2020.

[12] J. A. Okolie, D. Awotoye, M. E. Tabat, P. U. Okoye, E. I. Epelle, C. C. Ogbaga, F. Güleç, and B. Oboirien. "Multi-criteria decision analysis for the evaluation and screening of sustainable aviation fuel production pathways". *Iscience* 26 (2023), p. 106944.

[13] A. Elwalily, E. Verkama, F. Mantei, A. Kaliyeva, A. Pounder, J. Sauer, and F. Nestler. "Sustainable aviation fuel production via the methanol pathway: a technical review". *Sustainable Energy Fuels* (2025).

[14] J. Weitkamp. "Catalytic hydrocracking—mechanisms and versatility of the process". *ChemCatChem* 4 (2012), pp. 292–306.

[15] International Zeolite Association – Structure Commission. *Database of Zeolite Structures*. Accessed on 23 September 2025. URL: https://europe.iza-structure.org/IZA-SC/ftc_table.php.

[16] D. Breck. *Zeolite Molecular Sieves: Structure, Chemistry, and Use*. 1st ed. NewYork: John Wiley & Sons, 1973.

[17] B. Smit and T. L. Maesen. "Molecular simulations of zeolites: adsorption, diffusion, and shape selectivity". *Chem. Rev.* 108 (2008), pp. 4125–4184.

[18] C. Baerlocher, L. B. McCusker, and D. H. Olson. *Atlas of zeolite framework types*. 6th ed. Amsterdam: Elsevier, 2007.

[19] K. Cheng, L. I. van der Wal, H. Yoshida, J. Oenema, J. Harmel, Z. Zhang, G. Sunley, J. Zečević, and K. P. de Jong. "Impact of the spatial organization of bifunctional metal–zeolite catalysts on the hydroisomerization of light alkanes". *Angew. Chem.* 132 (2020), pp. 3620–3628.

[20] U. Agarwal, M. S. Rigutto, E. Zuidema, A. P. J. Jansen, A. Poursaeidesfahani, S. Sharma, D. Dubbeldam, and T. J. H. Vlugt. "Kinetics of zeolite-catalyzed heptane hydroisomerization and hydrocracking with CBMC-modeled adsorption terms: Zeolite Beta as a large pore base case". *J. Catal.* 415 (2022), pp. 37–50.

[21] A. Poursaeidesfahani, M. F. de Lange, F. Khodadadian, D. Dubbeldam, M. Rigutto, N. Nair, and T. J. H. Vlugt. "Product shape selectivity of MFI-type, MEL-type, and BEA-type zeolites in the catalytic hydroconversion of heptane". *J. Catal.* 353 (2017), pp. 54–62.

[22] D. Dubbeldam, S. Calero, and T. J. H. Vlugt. "iRASPA: GPU-accelerated visualization software for materials scientists". *Mol. Simul.* 44 (2018), pp. 653–676.

[23] P. Sazama, J. Pastvova, D. Kaucky, J. Moravkova, J. Rathousky, I. Jakubec, and G. Sadovska. “Does hierarchical structure affect the shape selectivity of zeolites? Example of transformation of n-hexane in hydroisomerization”. *J. Catal.* 364 (2018), pp. 262–270.

[24] Y. Han, J. Yuan, M. Xing, J. Cao, Z. Chen, L. Zhang, Z. Tao, Z. Liu, A. Zheng, X. Wen, *et al.* “Shape selectivity of zeolite for hydroisomerization of long-chain alkanes”. *New J. Chem.* 47 (2023), pp. 1401–1412.

[25] M. Zhang, Y. Chen, L. Wang, Q. Zhang, C.-W. Tsang, and C. Liang. “Shape selectivity in hydroisomerization of hexadecane over Pt supported on 10-ring zeolites: ZSM-22, ZSM-23, ZSM-35, and ZSM-48”. *Ind. Eng. Chem. Res.* 55 (2016), pp. 6069–6078.

[26] W. Souverijns, J. A. Martens, G. F. Froment, and P. A. Jacobs. “Hydrocracking of isoheptadecanes on Pt/H–ZSM-22: an example of pore mouth catalysis”. *J. Catal.* 174 (1998), pp. 177–184.

[27] V. M. Akhmedov and S. H. Al-Khowaiter. “Recent advances and future aspects in the selective isomerization of high n-Alkanes”. *Catal. Rev.* 49 (2007), pp. 33–139.

[28] T. L. M. Maesen, M. Schenk, T. J. H. Vlugt, J. P. De Jonge, and B. Smit. “The shape selectivity of paraffin hydroconversion on TON-, MTT-, and AEL-type sieves”. *J. Catal.* 188 (1999), pp. 403–412.

[29] S. Ernst, J. Weitkamp, J. A. Martens, and P. A. Jacobs. “Synthesis and shape-selective properties of ZSM-22”. *Appl. Catal.* 48 (1989), pp. 137–148.

[30] D. Romero, M. Rigutto, and E. J. Hensen. “Mechanistic aspects of n-hexadecane hydroconversion: Impact of di-branched isomers on the cracked products distribution”. *Fuel* 358 (2024), p. 130264.

[31] A. D. Estrada-Villagrana and C. De La Paz-Zavala. “Application of chemical equilibrium for hydrocarbon isomerization analysis”. *Fuel* 86 (2007), pp. 1325–1330.

[32] J. H. Gary, G. E. Handwerk, and M. J. Kaiser. *Petroleum refining: technology and economics. New York: M.* 5th ed. Boca Raton: CRC Press, 2007. DOI: https://doi.org/10.4324/9780203907924.

[33] M. Steijns, G. Froment, P. Jacobs, J. Uytterhaeven, and J. Weitkamp. “Hydroisomerization and hydrocracking. 2. Product distributions from n-decane and n-dodecane”. *Ind. Eng. Chem. Prod. Res. Dev.* 20 (1981), pp. 654–660.

[34] W. Yang, X. Yang, Y. Wang, R. Hou, Q. Gong, and Y. Pan. “Pervaporation separation of C6 alkane isomers by Al-bttotb membrane”. *J. Membr. Sci.* 661 (2022), p. 120916.

[35] C. Y. Chuah, H. Lee, and T.-H. Bae. “Recent advances of nanoporous adsorbents for light hydrocarbon (C_1–C_3) separation”. *Chem. Eng. J.* 430 (2022), p. 132654.

[36] D. Dubbeldam. “Computer-simulation of adsorption and diffusion of hydrocarbons in zeolites”. PhD Thesis. Universiteit van Amsterdam, 2005. URL: https://pid.uba.uva.nl/ark:/88238/b19900200386102051 31.

[37] M. G. Martin and J. I. Siepmann. “Transferable potentials for phase equilibria. 1. United-atom description of n-alkanes”. *J. Phys. Chem. B* 102 (1998), pp. 2569–2577.

[38] M. D. Macedonia and E. J. Maginn. “A biased grand canonical Monte Carlo method for simulating adsorption using all-atom and branched united atom models”. *Mol. Phys.* 96 (1999), pp. 1375–1390.

[39] D. Dubbeldam, A. Torres-Knoop, and K. S. Walton. “On the inner workings of Monte Carlo codes”. *Mol. Simul.* 39 (2013), pp. 1253–1292.

[40] J. Lennard and I. Jones. “On the determination of molecular fields.—I. From the variation of the viscosity of a gas with temperature”. *Proc. R. Soc. London, Ser. A* 106 (1924), pp. 441–462.

[41] H. A. Lorentz. “Ueber die anwendung des satzes vom virial in der kinetischen theorie der gase”. *Ann. Phys.* 248 (1881), pp. 127–136.

[42] D. Berthelot. “Sur le mélange des gaz”. *C. R.* 126 (1898), p. 15.

[43] P. Bai, M. Tsapatsis, and J. I. Siepmann. “TraPPE-zeo: transferable potentials for phase equilibria force field for all-silica zeolites”. *J. Phys. Chem. C* 117 (2013), pp. 24375–24387.

[44] T. J. H. Vlugt and M. Schenk. “Influence of framework flexibility on the adsorption properties of hydrocarbons in the zeolite silicalite”. *J. Phys. Chem. B* 106 (2002), pp. 12757–12763.

[45] S. Caro-Ortiz, E. Zuidema, M. Rigutto, D. Dubbeldam, and T. J. Vlugt. “Effects of framework flexibility on the adsorption and diffusion of aromatics in MFI-type zeolites”. *The Journal of Physical Chemistry C* 124 (2020), pp. 24488–24499.

[46] B. Widom. “Some topics in the theory of fluids”. *J. Chem. Phys.* 39 (1963), pp. 2808–2812.

[47] T. J. H. Vlugt, M. Martin, B. Smit, J. Siepmann, and R. Krishna. “Improving the efficiency of the configurational-bias Monte Carlo algorithm”. *Mol. Phys.* 94 (1998), pp. 727–733.

[48] T. J. H. Vlugt. “Efficiency of parallel CBMC simulations”. *Mol. Simul.* 23 (1999), pp. 63–78.

[49] T. J. H. Vlugt, R. Krishna, and B. Smit. “Molecular simulations of adsorption isotherms for linear and branched alkanes and their mixtures in silicalite”. *J. Phys. Chem. B* 103 (1999), pp. 1102–1118.

[50] D. Frenkel and B. Smit. *Understanding molecular simulation: from algorithms to applications*. 3rd ed. London: Academic Press, 2023.

[51] A. L. Myers and J. M. Prausnitz. "Thermodynamics of mixed-gas adsorption". *AIChE J.* 11 (1965), pp. 121–127.

[52] K. Walton and D. Sholl. "Predicting multicomponent adsorption: 50 years of the Ideal Adsorbed Solution Theory". *AIChE J.* 61 (2015), pp. 2757–2762.

[53] O. Talu and I. Zwiebel. "Multicomponent adsorption equilibria of nonideal mixtures". *AIChE J.* 32 (1986), pp. 1263–1276.

[54] C. Simon, B. Smit, and M. Haranczyk. "pyIAST: Ideal Adsorbed Solution Theory (IAST) Python Package". *Comput. Phys. Commun.* 200 (2016), pp. 364–380.

[55] Y. G. Chung, P. Bai, M. Haranczyk, K. T. Leperi, P. Li, H. Zhang, T. C. Wang, T. Duerinck, F. You, and J. T. Hupp. "Computational screening of nanoporous materials for hexane and heptane isomer separation". *Chem. Mater.* 29 (2017), pp. 6315–6328.

[56] J. A. Swisher, L.-C. Lin, J. Kim, and B. Smit. "Evaluating mixture adsorption models using molecular simulation". *AIChE J.* 59 (2013), pp. 3054–3064.

[57] D. Dubbeldam and B. Smit. "Computer simulation of incommensurate diffusion in zeolites: Understanding window effects". *J. Phys. Chem. B* 107 (2003), pp. 12138–12152.

[58] D. Dubbeldam, E. Beerdsen, T. J. H. Vlugt, and B. Smit. "Molecular simulation of loading-dependent diffusion in nanoporous materials using extended dynamically corrected transition state theory". *J. Chem. Phys.* 122 (2005).

[59] E. Beerdsen, B. Smit, and D. Dubbeldam. "Molecular simulation of loading dependent slow diffusion in confined systems". *Phys. Rev. Lett.* 93 (2004), p. 248301.

[60] *Dynamic Sorption Methods*. https://www.dynamicsorption.com/dynamic-sorption-method/. accessed on 03-11-2022.

[61] F. A. da Silva, J. A. Silva, and A. E. Rodrigues. "A general package for the simulation of cyclic adsorption processes". *Adsorption* 5 (1999), pp. 229–244.

[62] D. Ruthven. *Principles of Adsorption & Adsorption Processes*. 1st ed. New York: John Wiley & Sons, Inc., 1984.

[63] A. E. Rodrigues. "Perspectives on Adsorption. What else? A personal view". *Fluid Phase Equilib.* 564 (2023), p. 113614.

[64] D. D. Do. *Adsorption Analysis: Equilibria And Kinetics (With* CD Containing Computer MATLAB Programs). 1st ed. Ser. Chem. Eng. London: World Scientific Publishing, 1998.

[65] R. Yang. *Gas Separation by Adsorption Processes*. 1st ed. London: Imperial College Press, 1987.

[66] H. Versteeg and W. Malalasekera. *An introduction to computational fluid dynamics: the finite volume method*. 2nd ed. Glasgow: Pearson education, 2007.

[67] J. Butcher. *Numerical methods for ordinary differential equations*. 3rd ed. West Sussex: John Wiley & Sons, 2016.

[68] W. E. Schiesser. *The numerical method of lines: integration of partial differential equations*. 2nd ed. San Diego: Academic Press, 2012.

[69] D. I. Ketcheson. "Highly efficient strong stability-preserving Runge–Kutta methods with low-storage implementations". *SIAM J. Sci. Comp.* 30 (2008), pp. 2113–2136.

[70] S. Gottlieb and L. Gottlieb. "Strong stability preserving properties of Runge-Kutta time discretization methods for linear constant coefficient operators". *J. Sci. Comput.* 18 (2003), pp. 83–109.

[71] S. W. Benson, F. Cruickshank, D. Golden, G. R. Haugen, H. O'neal, A. Rodgers, R. Shaw, and R. Walsh. "Additivity rules for the estimation of thermochemical properties". *Chem. Rev.* 69 (1969), pp. 279–324.

[72] R. Gani. "Group contribution-based property estimation methods: advances and perspectives". *Curr. Opin. Chem. Eng.* 23 (2019), pp. 184–196.

[73] K. G. Joback and R. C. Reid. "Estimation of pure-component properties from group-contributions". *Chem. Eng. Commun.* 57 (1987), pp. 233–243.

[74] K. K. Yalamanchi, V. C. Van Oudenhoven, F. Tutino, M. Monge-Palacios, A. Alshehri, X. Gao, and S. M. Sarathy. "Machine learning to predict standard enthalpy of formation of hydrocarbons". *J. Phys. Chem. A* 123 (2019), pp. 8305–8313.

[75] M. N. Aldosari, K. K. Yalamanchi, X. Gao, and S. M. Sarathy. "Predicting entropy and heat capacity of hydrocarbons using machine learning". *Energy and AI* 4 (2021), p. 100054.

[76] C. Trinh, D. Meimaroglou, S. Lasala, and O. Herbinet. "Machine Learning for the prediction of the thermochemical properties (enthalpy and entropy of formation) of a molecule from its molecular descriptors". *Comput.-Aided Chem. Eng.* 51 (2022), pp. 1471–1476.

[77] E. S. Muckley, J. E. Saal, B. Meredig, C. S. Roper, and J. H. Martin. "Interpretable models for extrapolation in scientific machine learning". *Digital Discovery* 2 (2023), pp. 1425–1435.

[78] D. Hewitt, T. Pope, M. Sarwar, A. Turrina, and B. Slater. "Machine learning accelerated high-throughput screening of zeolites for the selective adsorption of xylene isomers". *Chem. Sci.* 13 (2022), pp. 13178–13186.

[79] I.-T. Sung, Y.-H. Cheng, C.-M. Hsieh, and L.-C. Lin. “Machine Learning for Gas Adsorption in Metal–Organic Frameworks: A Review on Predictive Descriptors”. *Ind. Eng. Chem. Res.* 64 (2025), pp. 1859–1875.

[80] K. M. Jablonka, D. Ongari, S. M. Moosavi, and B. Smit. “Big-data science in porous materials: materials genomics and machine learning”. *Chem. Rev.* 120 (2020), pp. 8066–8129.

[81] M. Petković, J. M. Vicent-Luna, V. Menkovski, and S. Calero. *Zeolite adsorption property prediction using deep learning*. arXiv e-prints arXiv–2403. 2024. URL: https://arxiv.org/html/2403.12659v1.

[82] L. Breiman. “Random Forests”. *Mach. Learn.* 45 (2001), pp. 5–32.

[83] T. Chen and C. Guestrin. “XGBoost: A Scalable Tree Boosting System”. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM. 2016, pp. 785–794. URL: https://doi.org/10.1145/2939672.2939785.

[84] A. V. Dorogush, V. Ershov, and A. Gulin. *CatBoost: gradient boosting with categorical features support*. arXiv preprint. 2018. URL: https://arxiv.org/abs/1810.11363.

[85] L. Prokhorenkova, G. Gusev, A. Vorobev, A. V. Dorogush, and A. Gulin. “CatBoost: Unbiased Boosting with Categorical Features”. *Advances in Neural Information Processing Systems 31 (NeurIPS 2018)*. NeurIPS. 2018. URL: https://arxiv.org/abs/1706.09516.

[86] E. Heid and W. H. Green. “Machine learning of reaction properties via learned representations of the condensed graph of reaction”. *J. Chem. Inf. Model.* 62 (2021), pp. 2101–2110.

[87] E. Heid, K. P. Greenman, Y. Chung, S.-C. Li, D. E. Graff, F. H. Vermeire, H. Wu, W. H. Green, and C. J. McGill. “Chemprop: a machine learning package for chemical property prediction”. *J. Chem. Inf. Model.* 64 (2023), pp. 9–17.

[88] W. Jin, R. Barzilay, and T. Jaakkola. “Multi-objective molecule generation using interpretable substructures”. *37th International Conference on Machine Learning*. PMLR. 2020. URL: https://arxiv.org/abs/2002.03244.

[89] N. Hollmann, S. Müller, K. Eggensperger, and F. Hutter. “TabPFN: A transformer that solves small tabular classification problems in a second”. *The Eleventh International Conference on Learning Representations*. ICLR. 2023. DOI: https://doi.org/10.48550/arXiv.2207.01848.

[90] N. Hollmann, S. Müller, L. Purucker, A. Krishnakumar, M. Körfer, S. B. Hoo, R. T. Schirrmeister, and F. Hutter. "Accurate predictions on small data with a tabular foundation model". *Nature* 637 (2025), pp. 319–326.

[91] C. Rzepa, D. Dabagian, D. W. Siderius, H. W. Hatch, V. K. Shen, J. Mittal, and S. Rangarajan. "Elucidating Thermodynamically Driven Structure–Property Relations for Zeolite Adsorption Using Neural Networks". *JACS Au* 4 (2024), pp. 4673–4690.

[92] K. Yang, K. Swanson, W. Jin, C. Coley, P. Eiden, H. Gao, A. Guzman-Perez, T. Hopper, B. Kelley, M. Mathea, A. Palmer, V. Settels, T. Jaakkola, K. Jensen, and R. Barzilay. "Analyzing learned molecular representations for property prediction". *J. Chem. Inf. Model.* 59 (2019), pp. 3370–3388.

[93] S. Sharma, M. S. Rigutto, R. Baur, U. Agarwal, E. Zuidema, S. R. G. Balestra, S. Calero, D. Dubbeldam, and T. J. H. Vlugt. "Modelling of adsorbate-size dependent explicit isotherms using a segregated approach to account for surface heterogeneities". *Mol. Phys.* 121 (2023), e2183721.

[94] C. L. Cavalcante. "Industrial adsorption separation processes: Fundamentals, modeling and applications". *Lat. Am. Appl. Res.* 30 (2000), pp. 357–364.

[95] P. Pullumbi, F. Brandani, and S. Brandani. "Gas separation by adsorption: technological drivers and opportunities for improvement". *Curr. Opin. Chem. Eng.* 24 (2019), pp. 131–142.

[96] Q. U. Jiuhui. "Research progress of novel adsorption processes in water purification: a review". *J. Environ. Sci.* 20 (2008), pp. 1–13.

[97] M. A. Alghoul, M. Y. Sulaiman, B. Z. Azmi, K. Sopian, and M. A. Wahab. "Review of materials for adsorption refrigeration technology". *Anti-Corros. Methods Mater.* 54 (2007), pp. 225–229.

[98] C. H. Yu, C. H. Huang, and C. S. Tan. "A review of CO_2 capture by absorption and adsorption". *Aerosol Air Qual. Res.* 12 (2012), pp. 745–769.

[99] B. Crittenden and W. J. Thomas. *Adsorption technology and design*. 1st ed. Woburn: Butterworth-Heinemann, 1998.

[100] A. Poursaeidesfahani, E. Andres-Garcia, M. de Lange, A. Torres-Knoop, M. Rigutto, N. Nair, F. Kapteijn, J. Gascon, D. Dubbeldam, and T. J. H. Vlugt. "Prediction of adsorption isotherms from breakthrough curves". *Microporous Mesoporous Mater.* 277 (2019), pp. 237–244.

[101] E. C. Markham and A. F. Benton. "The adsorption of gas mixtures by silica". *J. Am. Chem. Soc.* 53 (1931), pp. 497–507.

[102] A. Erto, A. Lancia, and D. Musmarra. "A Real Adsorbed Solution Theory model for competitive multicomponent liquid adsorption onto granular activated carbon". *Microporous and Mesoporous Materials* 154 (2012), pp. 45–50.

[103] R. Krishna and J. M. van Baten. "How Reliable Is the Ideal Adsorbed Solution Theory for the Estimation of Mixture Separation Selectivities in Microporous Crystalline Adsorbents?" *ACS Omega* 6 (2021), pp. 15499–15513.

[104] H. Renon and J. M. Prausnitz. "Local compositions in thermodynamic excess functions for liquid mixtures". *AIChE J.* 14 (1968), pp. 135–144.

[105] G. M. Wilson. "Vapor-liquid equilibrium. XI. A new expression for the excess free energy of mixing". *J. Am. Chem. Soc.* 86 (1964), pp. 127–130.

[106] A. Fredenslund, R. L. Jones, and J. M. Prausnitz. "Group-contribution estimation of activity coefficients in nonideal liquid mixtures". *AIChE J.* 21 (1975), pp. 1086–1099.

[107] R. Krishna. "Methodologies for screening and selection of crystalline microporous materials in mixture separations". *Sep. Purif. Technol.* 194 (2018), pp. 281–300.

[108] R. A. Koble and T. E. Corrigan. "Adsorption isotherms for pure hydrocarbons". *Ind. Eng. Chem.* 44 (1952), pp. 383–387.

[109] R. Krishna and J. M. van Baten. "Segregation effects in adsorption of CO_2-containing mixtures and their consequences for separation selectivities in cage-type zeolites". *Sep. Purif. Technol.* 61 (2008), pp. 414–423.

[110] R. Krishna and J. M. van Baten. "Influence of segregated adsorption on mixture diffusion in DDR zeolite". *Chem. Phys. Lett.* 446 (2007), pp. 344–349.

[111] S. Sircar. "Role of adsorbent heterogeneity on mixed gas adsorption". *Ind. Eng. Chem. Res.* 30 (1991), pp. 1032–1039.

[112] D. P. Valenzuela, A. L. Myers, O. Talu, and I. Zwiebel. "Adsorption of gas mixtures: effect of energetic heterogeneity". *AIChE J.* 34 (1988), pp. 397–402.

[113] K. Wang, S. Qiao, and X. Hu. "On the performance of HIAST and IAST in the prediction of multicomponent adsorption equilibria". *Sep. Purif. Technol.* 20 (2000), pp. 243–249.

[114] H. Moon and C. Tien. "Adsorption of gas mixtures on adsorbents with heterogeneous surfaces". *Chem. Eng. Sci.* 43 (1988), pp. 2967–2980.

[115] J. A. Ritter, S. J. Bhadra, and A. D. Ebner. “On the use of the dual-process Langmuir model for correlating unary equilibria and predicting mixed-gas adsorption equilibria”. *Langmuir* 27 (2011), pp. 4700–4712.

[116] A. Tarafder and M. Mazzotti. “A Method for Deriving Explicit Binary Isotherms Obeying the Ideal Adsorbed Solution Theory”. *Chem. Eng. Technol.* 35 (2012), pp. 102–108.

[117] T. R. C. Van Assche, G. V. Baron, and J. F. M. Denayer. “Adsorption size Effects for Langmuir systems in process simulators: case study comparing explicit Langmuir-based models and FASTIAS”. *Ind. Eng. Chem. Res.* 60 (2021), pp. 12092–12099.

[118] T. R. C. Van Assche, G. V. Baron, and J. F. M. Denayer. “An explicit multicomponent adsorption isotherm model: accounting for the size-Effect for components with Langmuir adsorption behavior”. *Adsorption* 24 (2018), pp. 517–530.

[119] R. Krishna and J. M. van Baten. “Entropy-based separation of linear chain molecules by exploiting differences in the saturation capacities in cage-type zeolites”. *Sep. Purif. Technol.* 76 (2011), pp. 325–330.

[120] Z. Du, G. Manos, T. J. H. Vlugt, and B. Smit. “Molecular simulation of adsorption of short linear alkanes and their mixtures in silicalite”. *AIChE J.* 44 (1998), pp. 1756–1764.

[121] R. Krishna, B. Smit, and S. Calero. “Entropy effects during sorption of alkanes in zeolites”. *Chem. Soc. Rev.* 31 (2002), pp. 185–194.

[122] E. García-Pérez, I. Torrens, S. Lago, D. Dubbeldam, T. J. H. Vlugt, T. L. Maesen, B. Smit, R. Krishna, and S. Calero. “Elucidating alkane adsorption in sodium-exchanged zeolites from molecular simulations to empirical equations”. *Appl. Surf. Sci.* 252 (2005), pp. 716–722.

[123] S. Sircar. “Comments on practical use of Langmuir gas adsorption isotherm model”. *Adsorption* 23 (2017), pp. 121–130.

[124] T. J. H. Vlugt, W. Zhu, F. Kapteijn, J. Moulijn, B. Smit, and R. Krishna. “Adsorption of linear and branched alkanes in the zeolite silicalite-1”. *J. Am. Chem. Soc.* 120 (1998), pp. 5599–5600.

[125] R. Krishna, B. Smit, and T. J. H. Vlugt. “Sorption-induced diffusion-selective separation of hydrocarbon isomers using silicalite”. *J. Phys. Chem. A* 102 (1998), pp. 7727–7730.

[126] R. L. Burden and J. D. Faires. *Numerical Analysis*. 9th ed. Mathematics Series. Boston: PWS-Kent Publishing Company, 2010.

[127] W. Kast. *Adsorption aus der Gasphase. Ingenieurwissenschaftliche Grundlagen und technische Verfahren, Verlag Chemie*. 1988.

[128] R. Yang. *Adsorbents: Fundamentals and Applications*. 1st ed. Hoboken: John Wiley & Sons, 2003.

[129] R. Krishna and R. Baur. “Modelling issues in zeolite based separation processes”. *Sep. Purif. Technol.* 33 (2003), pp. 213–254.

[130] E. Glueckauf. “Theory of chromatography. Part 10.—Formulæ for diffusion into spheres and their application to chromatography”. *Trans. Faraday Soc.* 51 (1955), pp. 1540–1551.

[131] S. Sircar and J. Hufton. “Why does the Linear Driving Force model for adsorption kinetics work?” *Adsorption* 6 (2000), pp. 137–147.

[132] A. Özdural. “2.48 - Modeling Chromatographic Separation”. *Comprehensive Biotechnology*. Ed. by M. Moo-Young. 2nd ed. Burlington: Academic Press, 2011, pp. 681–695. ISBN: 978-0-08-088504-9. DOI: https://doi.org/10.1016/B978-0-08-088504-9.00407-4.

[133] D. Dubbeldam, S. Calero, D. E. Ellis, and R. Q. Snurr. “RASPA: molecular simulation software for adsorption and diffusion in flexible nanoporous materials”. *Mol. Simul.* 42 (2016), pp. 81–101.

[134] P. P. Ewald. “Die Berechnung optischer und elektrostatischer Gitterpotentiale”. *Ann. Phys.* 369 (1921), pp. 253–287.

[135] A. Garcia-Sanchez, C. O. Ania, J. B. Parra, D. Dubbeldam, T. J. H. Vlugt, R. Krishna, and S. Calero. “Transferable force field for carbon dioxide adsorption in zeolites”. *J. Phys. Chem. C* 113 (2009), pp. 8814–8820.

[136] D. Dubbeldam, S. Calero, T. J. H. Vlugt, R. Krishna, T. L. M. Maesen, and B. Smit. “United atom force field for alkanes in nanoporous materials”. *J. Phys. Chem. B* 108 (2004), pp. 12301–12313.

[137] J. D. Anderson and J. Wendt. *Computational fluid dynamics*. 1st ed. Vol. 206. New York: McGraw-Hill, 1995.

[138] H. P. Langtangen and S. Linge. *Finite difference computing with PDEs: a modern software approach*. 1st ed. Cham: Springer Nature, 2017.

[139] C. W. Shu and S. Osher. “Efficient implementation of essentially non-oscillatory shock-capturing schemes”. *J. Comp. Phys.* 77 (1988), pp. 439–471.

[140] H. Landa. “Development of an efficient method for simulating fixed-bed adsorption dynamics using Ideal Adsorbed Solution Theory”. http://d-nb.info/1128726513/34. PhD thesis. Otto-von-Guericke-Universität Magdeburg, 2016.

[141] S. Sharma, S. R. Balestra, R. Baur, U. Agarwal, E. Zuidema, M. S. Rigutto, S. Calero, T. J. H. Vlugt, and D. Dubbeldam. “RUPTURA: simulation code for breakthrough, ideal adsorption solution theory computations, and fitting of isotherm models”. *Mol. Simul.* 49 (2023), pp. 893–953.

[142] A. de Haan and H. Bosch. *Industrial Separation Processes; Fundamentals*. 1st ed. Berlin/Boston: De Gruyter, 2013.

[143] A. E. Rodrigues. *Simulated Moving Bed Technology: Principles, Design and Process Applications*. 1st ed. Kidlington: Butterworth-Heinemann, 2015.

[144] N. Jiang, Y. Shen, B. Liu, D. Zhang, Z. Tang, G. Li, and B. Fu. "CO_2 capture from dry flue gas by means of VPSA, TSA and TVSA". *J. CO_2 Util.* 35 (2020), pp. 153–168.

[145] P. Luis. "Chapter 8 - Hybrid processes based on membrane technology". *Fundamental Modelling of Membrane Systems*. Ed. by P. Luis. Amsterdam: Elsevier, 2018, pp. 301–343. ISBN: 978-0-12-813483-2.

[146] D. Ruthven, S. Farooq, and K. Knaebel. *Pressure Swing Adsorption*. 1st ed. New York: VCH Publishers Inc., 1994.

[147] R. Canevesi, C. Borba, A. da Silva, and C. Grande. "Towards a design of a pressure swing adsorption unit for small scale biogas upgrading at". *Energy Procedia* 158 (2019), pp. 848–853.

[148] M. Yavary, H. Ebrahim, and C. Falamaki. "The effect of number of pressure equalization steps on the performance of pressure swing adsorption process". *Chem. Eng. Process.* 87 (2015), pp. 35–44.

[149] C. A. Grande. "Advances in pressure swing adsorption for gas separation". *Int. Schol. Res. Not.* 2012 (2012), p. 982934.

[150] M. Oschatz and M. Antonietti. "A search for selectivity to enable CO_2 capture with porous adsorbents". *Energy Environ. Sci.* 11 (2018), pp. 57–70.

[151] R. Kumar. "Pressure Swing Adsorption Process: Performance Optimum and Adsorbent Selection". *Ind. Eng. Chem. Res.* 33 (1994), pp. 1600–1605.

[152] M. Ho, G. Allinson, and D. Wiley. "Reducing the Cost of CO_2 Capture from Flue Gases Using Pressure Swing Adsorption". *Ind. Eng. Chem. Res.* 47 (2008), pp. 4883–4890.

[153] S. Bhatia and A. Myers. "Optimum Conditions for Adsorptive Storage". *Langmuir* 22 (2006), pp. 1688–1700.

[154] M. K. Al Mohammed, M. Danish, I. M. Khan, I. H. Ali, M. Hasan, and A. El Jery. "Continuous fixed bed CO_2 adsorption: Breakthrough, column efficiency, mass transfer zone". *Processes* 8 (2020), p. 1233.

[155] C. Tien. *Introduction to Adsorption: Basics, Analysis, and Applications*. 1st ed. Amsterdam: Elsevier, 2019.

[156] D. Dubbeldam, S. Calero, D. E. Ellis, and R. Q. Snurr. "RASPA: Molecular Simulation Software for Adsorption and Diffusion in Flexible Nanoporous Materials". *Mol. Simul.* 42 (2016), pp. 81–101.

[157] G. Guiochon, A. Felinger, and D. Shirazi. *Fundamentals of Preparative and Nonlinear Chromatography*. 2nd ed. Amsterdam: Academic Press, 2006.

[158] O. Moradi. “Thermodynamics of Interfaces”. *Thermodynamics - Interaction Studies - Solids, Liquids and Gases*. Rijeka: InTechOpen, 2011. Chap. 8.

[159] N. Ayawei, A. Ebelegi, and D. Wankasi. “Modelling and Interpretation of Adsorption Isotherms”. *J. Chem.* 2017 (2017), p. 3039817.

[160] C. Girish. “Various isotherm models for multicomponent adsorption: A review”. *Int. J. Res. Civ. Eng. Technol.* 8 (2017), pp. 80–86.

[161] C.-W. Shu. “Total-variation diminishing time discretizations”. *IAM J. Sci. Statist. Comput.* 9 (1988), pp. 1073–1084.

[162] R. Winterton. “Van der Waals forces”. *Contemp. Phys.* 11 (1970), pp. 559–574.

[163] J. Condon. *Surface area and porosity determinations by physisorption: measurements and theory*. 1st ed. Amsterdam: Elsevier Science, 2006.

[164] A. Dabrowski. “Adsorption - from theory to practice”. *Adv. Colloid Interface Sci.* 93 (2001), pp. 135–224.

[165] S. Sircar. “Excess properties and thermodynamics of multicomponent gas adsorption”. *J. Chem. Soc., Faraday Trans.* 81 (1985), pp. 1527–1540.

[166] R. Pini. “Interpretation of net and excess adsorption isotherms in microporous adsorbents”. *Microporous Mesoporous Mater.* 187 (2014), pp. 40–52.

[167] S. Gumma and O. Talu. “Net Adsorption: A Thermodynamic Framework for Supercritical Gas Adsorption and Storage in Porous Solids”. *Langmuir* 26 (2010), pp. 17013–17023.

[168] A. Myers. “Thermodynamics of adsorption in porous materials”. *AIChE J.* 48 (2002), pp. 145–160.

[169] L. Bingel and K. Walton. “Surprising Use of the Business Innovation Bass Diffusion Model To Accurately Describe Adsorption Isotherm Types I, III, and V”. *Langmuir* 39 (2023), pp. 4475–4482.

[170] I. Langmuir. “The adsorption of gases on plane surfaces of glass, mica and platinum”. *J. Am. Chem. Soc.* 40 (1918), pp. 1361–1403.

[171] F. Gritti and G. Guiochon. “New thermodynamically consistent competitive adsorption isotherm in RPLC”. *J. Colloid Interface Sci.* 1 (2003), pp. 43–59.

[172] C. C. Boedeker. “Über das Verhältnis zwischen Masse und Wirkung beim Kontakt ammoniakalscher Flüssigkeiten mit Ackererde und mit kohlensaurem Kalk”. *Z.A. Pflanzenbau* 7 (1859), pp. 48–58.

[173] R. Sips. “On the Structure of a Catalyst Surface”. *J. Chem. Phys.* 16 (1948), pp. 490–495.

[174] O. Redlich and D. Peterson. “A useful adsorption isotherm”. *J. Phys. Chem.* 63 (1959), p. 1024.

[175] J. Toth. “State equations of the solid-gas interface layers”. *Acta Chim. Acad. Sci. Hung.* 69 (1971), pp. 311–328.

[176] J. Toth. “Uniform Interpretation of Gas/Solid Adsorption”. *Adv. Colloid Interface Sci.* 55 (1995), pp. 1–239.

[177] J. Toth. *Adsorption: Theory, Modelling, and Analysis*. 1st ed. New York: Marcel Dekker Inc, 2002.

[178] J. A. O’Brien and A. L. Myers. “Physical adsorption of gases on heterogeneous surfaces: Series expansion of isotherms using central moments of the adsorption energy distribution”. *J. Chem. Soc. Faraday Trans.* 80 (1984), pp. 1467–1477.

[179] T. Hill. *An Introduction to Statistical Thermodynamics*. 3rd ed. New York: Dover Publications, 1986.

[180] B. Lin, Z. Ma, S. Golshan-Shirazi, and G. Guiochon. “Study of the Representation of Competitive Isotherms and of the Intersection between Adsorption Isotherms”. *J. Chromatogr.* 475 (1989), pp. 1–11.

[181] V. Tempkin. “Kinetics of ammonia synthesis on promoted iron catalyst”. *Acta Phys. Chim. USSR* 12 (1940), pp. 327–356.

[182] C. Simon, J. Kim, L.-C. Lin, R. Martin, M. Haranczyk, and B. Smit. “Optimizing Nanoporous Materials for Gas Storage”. *Phys. Chem. Chem. Phys.* 16 (2014), pp. 5499–5513.

[183] D. Graham. “The characterization of physical adsorption systems”. *J. Phys. Chem.* 57 (1953), pp. 665–669.

[184] B. Smit and T. Maesen. “Commensurate freezing of alkanes in the channels of a zeolite”. *Nature* 374 (1995), pp. 42–44.

[185] E. Worch. *Adsorption Technology in Water Treatment*. 1st ed. Berlin/-Boston: Walter de Gruyter GmbH & Co, 2012.

[186] W. Weber Jr and F. DiGiano. *Process dynamics in environmental systems*. 1st ed. New York: John Wiley & Sons, 1996.

[187] O. Talu and A. L. Myers. “Rigorous thermodynamic treatment of gas adsorption”. *AIChE J.* 34 (1988), pp. 1887–1893.

[188] F. Bass. “A New Product Growth for Model Consumer Durables”. *Manag. Sci.* 15 (1969), pp. 215–227.

[189] H. DeVoe. *Thermodynamics and Chemistry*. 1st ed. Hoboken: Prentice-Hall, Inc., 2015.

[190] A. L. Myers and P. A. Monson. “Physical adsorption of gases: the case for absolute adsorption as the basis for thermodynamic analysis”. *Adsorption* 20 (2014), pp. 591–622.

[191] M. Murthi and R. Snurr. “Effects of Molecular Siting and Adsorbent Heterogeneity on the Ideality of Adsorption Equilibria”. *Langmuir* 20 (2004), pp. 2489–2497.

[192] S. Brandani, E. Mangano, and M. Luberti. “Net, excess, and absolute adsorption in mixed gas adsorption”. *Adsorption* 23 (2017), pp. 569–576.

[193] A. Myers. “Thermodynamics of adsorption”. *Chemical Thermodynamics for Industry*. Ed. by T. Letcher. 1st ed. Cambridge: Royal Society of Chemistry, 2004. Chap. 21, pp. 243–253.

[194] G. Santori, M. Luberti, and H. Ahn. “Ideal Adsorbed Solution Theory solved with direct search minimisation”. *Comput. Chem. Eng.* 71 (2014), pp. 235–240.

[195] A. Voigt. *Comparison of methods for the calculation of the real dilogarithm regarding instruction-level parallelism*. (arXiv:2201.01678). 2022. DOI: 10.48550/ARXIV.2201.01678.

[196] R. Vidunas. “Algebraic transformations of Gauss hypergeometric functions”. *Funkcial. Ekvac.* 52 (2009), pp. 139–180.

[197] W. Koepf. “Gosper’s Algorithm”. *Hypergeometric Summation: An Algorithmic Approach to Summation and Special Function Identities*. London: Springer London, 2014, pp. 79–101.

[198] H. Landa, D. Flockerzi, and A. Seidel-Morgenstern. “A Method for Efficiently Solving the IAST Equations with an Application to Adsorber Dynamics”. *AIChE J.* 59 (2012), pp. 1263–1277.

[199] A. Seidel, G. Reschke, S. Friedrich, and D. Gelbin. “Equilibrium adsorption of two-component organic solutes from aqueous solutions on activated carbon”. *Adsorption Sci. Technol.* 3 (1986), pp. 189–199.

[200] A. Myers and D. Valenzuela. “Adsorption from bisolute systems on active carbon”. *J. Water Pollut. Control. Fed.* 45 (1973), pp. 2463–2479.

[201] D. Shade, B. Bout, D. Sholl, and K. Walton. “Opening the Toolbox: 18 Experimental Techniques for Measurement of Mixed Gas Adsorption”. *Ind. Eng. Chem. Res.* 61 (2022), pp. 2367–2391.

[202] J. Butler and C. Ockrent. “Studies in electrocapillarity. Part III. The surface tensions of solutions containing two surface-active solutes”. *J. Phys. Chem.* 34 (1930), pp. 2841–2859.

[203] D. Broughton. “Adsorption isotherms for binary gas mixtures”. *Ind. Eng. Chem.* 40 (1948), pp. 1506–1508.

[204] J. Jain and V. Snoeyink. “Adsorption from bisolute systems on active carbon”. *J. Water Pollut. Control. Fed.* 45 (1973), pp. 2463–2479.

[205] J. O’Brien and A. Myers. “Rapid calculations of multicomponent adsorption equilibria from pure isotherm data”. *Ind. Eng. Chem. Process. Des. Dev.* 24 (1985), pp. 1188–1191.

[206] H. Chen and D. Sholl. "Examining the Accuracy of Ideal Adsorbed Solution Theory without Curve-Fitting Using Transition Matrix Monte Carlo Simulations". *Langmuir* 23 (2007), pp. 6431–6437.

[207] N. Cessford, N. Seaton, and T. Düren. "Evaluation of Ideal Adsorbed Solution Theory as a Tool for the Design of Metal–Organic Framework Materials". *Ind. Eng. Chem. Res.* 51 (2012), pp. 4911–4921.

[208] G. Fraux, A. Boutin, and F.-X. Coudert. "On the use of the IAST method for gas separation studies in porous materials with gate-opening behavior". *Adsorption* 24 (2018), pp. 233–241.

[209] F. Gharagheizi and D. Sholl. "Comprehensive Assessment of the Accuracy of the Ideal Adsorbed Solution Theory for Predicting Binary Adsorption of Gas Mixtures in Porous Materials". *Ind. Eng. Chem. Res.* 61 (2022), pp. 727–739.

[210] S. Suwanayuen and R. P. Danner. "Vacancy Solution Theory of Adsorption From Gas Mixtures". *AIChE J.* 26 (1980), pp. 76–83.

[211] E. Costa, J. Sotelo, G. Calleja, and C. Marron. "Adsorption of binary and ternary hydrocarbon-gas mixtures on activated carbon - experimental-determination and theoretical prediction of the ternary equilibrium data". *AIChE J.* 27 (1981), pp. 5–12.

[212] G. Gamba, R. Rota, G. Storti, S. Carrà, and M. Morbidelli. "Vacancy Solution Theory of Adsorption From Gas Mixtures". *AIChE J.* 35 (1989), pp. 959–966.

[213] M. Sakuth, J. Meyer, and J. Gmehling. "Measurement and prediction of binary adsorption equilibria of vapors on dealuminated Y-zeolites (DAY)". *Chem. Eng. Process.* 37 (1998), pp. 267–277.

[214] S. Bartholdy, M. G. Bjorner, E. Solbraa, A. Shapiro, and G. M. Kontogeorgis. "Measurement and prediction of binary adsorption equilibria of vapors on dealuminated Y-zeolites (DAY)". *Ind. Eng. Chem. Res.* 52 (2013), pp. 11552–11563.

[215] A. Ladshaw, S. Yiacoumi, and C. Tsouris. "A Generalized Procedure for the Prediction of Multicomponent Adsorption". *AIChE J.* 61 (2015), pp. 2600–2610.

[216] J. Smith and H. van Ness. *Introduction to Chemical Engineering Thermodynamics*. 4th ed. New York: McGraw-Hill, 1987.

[217] C. J. Radke and J. M. Prausnitz. "Thermodynamics of multi-solute adsorption from dilute liquid solutions". *AIChE J.* 18 (1972), pp. 761–768.

[218] E. Richter, W. Schutz, and A. L. Myers. "Effect of adsorption equation on prediction of multicomponent adsorption equilibria by the Ideal Adsorbed Solution Theory". *Chem. Eng. Sci.* 22 (1989), pp. 1609–1616.

[219] A. L. Myers and D. Valenzuela. “Computer algorithm and graphical method for calculating adsorption equilibria of gas mixtures”. *J. Chem. Eng. Jpn.* 19 (1986), pp. 392–396.

[220] M. LeVan and T. Vermeulen. “Binary Langmuir and Freundlich isotherms for Ideal Adsorbed Solutions”. *J. Phys. Chem.* 85 (1981), pp. 3247–3250.

[221] M. Ilic, D. Flockerzi, and A. Seidel-Morgenstern. “A thermodynamically consistent explicit competitive adsorption isotherm model based on second-order single component behaviour”. *J. Chromatogr. A* 1217 (2010), pp. 2132–2137.

[222] G. Cardano and T. Witmer. *The Great Art; Or, The Rules of Algebra. Translated and Edited by T. Richard Witmer*. 1st ed. Cambridge: MIT Press, 1968.

[223] H. Moon and C. Tien. “Further work on multicomponent adsorption equilibria calculations based on the Ideal Adsorbed Solution Theory”. *Ind. Eng. Chem. Res.* 26 (1987), pp. 2042–2047.

[224] J. H. Yun, H. C. Park, and H. Moon. “Multicomponent adsorption calculations based on adsorbed solution theory”. *Korean J. Chem. Eng.* 13 (1996), pp. 246–254.

[225] E. Mangano, D. Friedrich, and S. Brandani. “Robust algorithms for the solution of the Ideal Adsorbed Solution Theory equations”. *AIChE J.* 61 (2014), pp. 981–991.

[226] J. O’Brien and A. Myers. “A comprehensive technique for equilibrium calculations in adsorbed mixtures: the generalized FastIAS Method”. *Ind. Eng. Chem. Res.* 27 (1988), pp. 2085–2092.

[227] K. Choy, J. Porter, and G. McKay. “Single and Multicomponent Equilibrium Studies for the Adsorption of Acidic Dyes on Carbon from Effluent”. *Langmuir* 20 (2004), pp. 9646–9656.

[228] D. D. Frey and A. E. Rodrigues. “Explicit Calculation of Multicomponent Equilibria for Ideal Adsorbed Solutions”. *AIChE J.* 40 (1994), pp. 182–186.

[229] J. C. Santos, Y. Cheng, M. M. Dias, and A. E. Rodrigues. “Surface B-splines fitting for speeding up the simulation of adsorption processes with IAS model”. *Comput. Chem. Eng.* 35 (2011), pp. 1186–1191.

[230] J. Nelder and R. Mead. “A Simplex Method for Function Minimization”. *Comput. J.* 7 (1965), pp. 308–313.

[231] A. Mirzaei, Z. Chen, F. Haghighat, and L. Yerushalmi. “Enhanced adsorption of anionic dyes by surface fluorination of zinc oxide: A straightforward method for numerical solving of the Ideal Adsorbed Solution Theory (IAST)”. *Chem. Eng. J.* 30 (2017), pp. 407–418.

[232] M. Benjamin. “New Conceptualization and Solution Approach for the Ideal Adsorbed Solution Theory (IAST)”. *Environ. Sci. Technol.* 43 (2009), pp. 2530–2536.

[233] S. R. G. Balestra, R. Bueno-Perez, R. Aguilar, and S. Calero. *GAIAST: A Fortran code with Genetic Algorithms (GA) for an Adsorbed Solution Theory (IAST)*. http://dx.doi.org/10.5281/zenodo.596674. 2016.

[234] S. Lee, J. H. Lee, and J. Kim. “User-friendly graphical user interface software for Ideal Adsorbed Solution Theory calculations”. *Korean J. Chem. Eng.* 35 (2018), pp. 214–221.

[235] P. Iacomi and P. Llewellyn. “pyGAPS: a Python-based framework for adsorption isotherm processing and material characterisation”. *Adsorption* 25 (2019), pp. 1533–1542.

[236] E. Dautzenberg, S. S. van Hurne, M. Smulders, and L. C. P. M. de Smet. “GraphIAST: A graphical user interface software for Ideal Adsorption Solution Theory (IAST) calculations”. *Comput. Phys. Commun.* 280 (2022), p. 108494.

[237] S. Hoseinpoori. “Steady-state modeling of co-adsorption phenomena in direct air carbon-capture system using the IAST”. http://hdl.handle.net/10589/169486. PhD Thesis. Politecnico di Milano and RWTH Aachen University, 2020.

[238] J. L. Chabert, E. Barbin, J. Borowczyk, M. Guillemot, and A. Michel-Pajus. *A history of algorithms: from the pebble to the microchip*. 1st ed. Vol. 23. Springer, 1999.

[239] H. Patel. “Fixed-bed column adsorption study: a comprehensive review”. *Appl. Water Sci.* 9 (2019), pp. 1–17.

[240] J. Silva, A. Ferreira, P. Mendes, A. Cunha, K. Gleichmann, and A. E. Rodrigues. “Adsorption Equilibrium and Dynamics of Fixed Bed Adsorption of CH_4/N_2 in Binderless Beads of 5A Zeolite”. *Ind. Eng. Chem. Res.* 54 (2015), pp. 6390–6399.

[241] N. Wilkins, A. Rajendran, and S. Farooq. “Dynamic column breakthrough experiments for measurement of adsorption equilibrium and kinetics”. *Adsorption* 27 (2021), pp. 397–422.

[242] A. Farmahini, S. Krishnamurthy, D. Friedrich, S. Brandani, and L. Sarkisov. “Performance-based screening of porous materials for carbon capture”. *Chem. Rev.* 121 (2021), pp. 10666–10741.

[243] A. Gabelman. “Adsorption basics: part 1”. *Chem. Eng. Prog.* 113 (2017), pp. 48–53.

[244] R. Klaewkla, M. Arend, and W. Hoelderich. “A Review of Mass Transfer Controlling the Reaction Rate in Heterogeneous Catalytic Systems”. *Mass Transfer-Advanced Aspects*. Ed. by H. Nakajima. London: INTECH Open Access Publisher, 2011. Chap. 29.

[245] H. Fogler. *Elements of Chemical Reaction in Engineering*. 3rd ed. Naucalpan de Juárez: Prentice-Hall of India, 1999.

[246] Z. Xu, J.-G. Cai, and B.-C. Pan. "Mathematically modeling fixed-bed adsorption in aqueous systems". *J. Zhejiang Univ. Sci. A* 14 (2013), pp. 155–176.

[247] J. G. Jee, H. J. Park, S. J. Haam, and C. H. Lee. "Effects of nonisobaric and isobaric steps on O2 pressure swing adsorption for an aerator". *Ind. Eng. Chem. Res.* 41 (2002), pp. 4383–4392.

[248] M. Kim, J. Moon, C. Lee, H. Ahn, and W. Cho. "Effect of heat transfer on the transient dynamics of temperature swing adsorption process". *Korean J. Chem. Eng.* 21 (2004), pp. 703–711.

[249] M. S. Shafeeyan, W. M. A. W. Daud, and A. Shamiri. "A review of mathematical modeling of fixed-bed columns for carbon dioxide adsorption". *Chem. Eng. Res. Des.* 92 (2014), pp. 961–988.

[250] K. Hwang, J. Jun, and W. Lee. "Fixed-bed adsorption for bulk component system non-equilibrium, non-isothermal and non-adiabatic model". *Chem. Eng. Sci.* 50 (1995), pp. 813–825.

[251] C. Tien. *Adsorption Calculations and Modeling*. 1st ed. Boston: Butterworth-Heinemann, 1994.

[252] E. Worch. "Fixed-bed adsorption in drinking water treatment: a critical review on models and parameter estimation". *J. Water Supply Res. Technol. AQUA* 57 (2008), pp. 171–183.

[253] M. Siahpoosh, S. Fatemi, and A. Vatani. "Mathematical Modeling of Single and Multi-Component Adsorption Fixed Beds to Rigorously Predict the Mass Transfer Zone and Breakthrough Curves". *Iran. J. Chem. Chem. Eng.* 28 (2009), pp. 25–44.

[254] N. M. Supian, N. A. Halif, and N. Rusli. "A Mini-Review of Coupled Convection-Diffusion Equations in a Fixed-Bed Adsorption". *IOP Conf. Ser.: Mater. Sci. Eng.* 767 (2019), p. 012031.

[255] E. Unuabonah, M. Omorogie, and N. Oladoja. "5 - Modeling in Adsorption: Fundamentals and Applications". *Composite Nanoadsorbents*. Amsterdam: Elsevier, 2019, pp. 85–118.

[256] H. Thomas. "Heterogeneous ion exchange in a flowing system". *J. Am. Chem. Soc.* 66 (1944), pp. 1664–1666.

[257] G. Bohart and E. Adams. "Some aspects of the behavior of charcoal with respect to chlorine". *J. Am. Chem. Soc.* 42 (1920), pp. 523–544.

[258] Y. Yoon and J. Nelson. "Application of gas adsorption kinetics I. A theoretical model for respirator cartridge service life". *Am. Ind. Hyg. Assoc. J.* 45 (1984), pp. 509–516.

[259] E. Unuabonah, M. El-Khaiary, B. Olu-Owolabi, and K. Adebowale. "Predicting the dynamics and performance of a polymer–clay based composite in a fixed bed system for the removal of lead (II) ion". *Chem. Eng. Res. Des.* 90 (2012), pp. 1105–1115.

[260] R. Clark. "Evaluating the cost and performance of field-scale granular activated carbon systems". *Environ. Sci. Technol.* 21 (1987), pp. 573–580.

[261] L. Lapidus and N. Amundson. "The effect of longitudinal diffusion in ion exchange and chromatographic columns". *J. Phys. Chem.* 56 (1952), pp. 984–988.

[262] W. Plazinski, W. Rudzinskia, and A. Plazinska. "Theoretical models of sorption kinetics including a surface reaction mechanism: A review". *Adv. Colloid Interface Sci.* 152 (2009), pp. 2–13.

[263] I. Medved and R. Cerny. "Surface diffusion in porous media: A critical review". *Microporous Mesoporous Mater.* 142 (2011), pp. 405–422.

[264] A. Fick. "Üeber diffusion". *Ann. Phys.* 170 (1855), pp. 59–86.

[265] J. C. Knox. "Predictive Simulation of Gas Adsorption in Fixed-Beds and Limitations due to the Ill-Posed Danckwerts Boundary Condition". https://louis.uah.edu/uah-dissertations/108/. PhD Thesis. Swiss School of Business Research, 2016.

[266] C. de Ligny. "Coupling between diffusion and convection in radial dispersion of matter by fluid flow through packed beds". *Chem. Eng. Sci.* 25 (1970), pp. 1177–1181.

[267] S. Chung and C. Wen. "Longitudinal dispersion of liquid flowing through fixed and fluidized beds". *AIChE J.* 14 (1968), pp. 857–866.

[268] S. Rastegar and T. Gu. "Empirical correlations for axial dispersion coefficient and Peclet number in fixed-bed columns". *J. Chromatogr. A* 1490 (2017), pp. 133–137.

[269] M. Edward and J. Richardson. "Gas dispersion in packed beds". *Chem. Eng. Sci.* 23 (1968), pp. 109–123.

[270] I. N. Tsimpanogiannis, O. A. Moultos, L. F. M. Franco, M. B. de M. Spera, M. Erdős, and I. Economou. "Self-diffusion coefficient of bulk and confined water: a critical review of classical molecular simulation studies". *Mol. Simul.* 45 (2019), pp. 425–453.

[271] A. T. Celebi, S. H. Jamali, A. Bardow, T. J. H. Vlugt, and O. A. Moultos. "Finite-size effects of diffusion coefficients computed from molecular dynamics: a review of what we have learned so far". *Mol. Simul.* 47 (2021), pp. 831–845.

[272] E. L. Cussler. *Diffusion: Mass transfer in fluid systems*. 3rd ed. New York: Cambridge University Press, 2009.

[273] R. Bird. "Transport phenomena". *Appl. Mech. Rev.* 55 (2002), R1–R4.

[274] J. E. Jones. “On the determination of molecular fields.—I. From the variation of the viscosity of a gas with temperature”. *Proc. R. soc. Lond. Ser. A-Contain. Pap. Math. Phys. Character.* 106 (1924), pp. 441–462.

[275] J. E. Jones. “On the determination of molecular fields.—II. From the equation of state of a gas”. *Proc. R. soc. Lond. Ser. A-Contain. Pap. Math. Phys. Character.* 106 (1924), pp. 463–477.

[276] J. Kärger, D. Ruthven, and D. Theodorou. *Diffusion in nanoporous materials*. 1st ed. Weinheim: John Wiley & Sons, 2012.

[277] J. Crittenden, N. Hutzler, D. Geyer, J. Oravitz, and G. Friedman. “Transport of organic compounds with saturated groundwater flow: model development and parameter sensitivity”. *Water Resour. Res.* 22 (1986), pp. 271–284.

[278] A. Kapoor, R. Yang, and C. Wong. “Surface-Diffusion”. *Catal. Rev. - Sci. Eng.* 31 (1989), pp. 129–214.

[279] J. Rosen. “Kinetics of a Fixed Bed System for Solid Diffusion into Spherical Particles”. *J. Chem. Phys.* 20 (1952), pp. 387–394.

[280] A. Mathews and W. Weber Jr. “Effects of external mass transfer and intraparticle diffusion on adsorption in slurry reactors”. *AIChE Symp. Ser.* 73 (1976), pp. 91–98.

[281] R. Peel, A. Benedek, and C. Crowe. “A Branched Pore Kinetic-Model for Activated Carbon Adsorption”. *AIChE J.* 27 (1981), pp. 26–32.

[282] Y. S. S. Zhang. “Study on protein adsorption kinetics to a dye–ligand adsorbent by the pore diffusion model”. *J. Chromatogr. A* 964 (2002), pp. 35–46.

[283] D. Ko, J. Porter, and G. McKay. “Film-pore diffusion model for the fixed-bed sorption of copper and cadmium ions onto bone char”. *Water Res.* 35 (2001), pp. 3876–3886.

[284] E. Worch. “Modelling the solute transport under nonequilibrium conditions on the basis of mass transfer equations”. *J. Contam. Hydrol.* 68 (2004), pp. 97–120.

[285] M. Rahman, F. Amiri, and E. Worch. “Application of the mass transfer model for describing non-equilibrium transport of HOCs through natural geosorbents”. *Water Res.* 37 (2003), pp. 4673–4684.

[286] T. Weber and R. Chakravorti. “Pore and solid diffusion models for fixed-bed adsorbers”. *AIChE J.* 20 (1974), pp. 228–238.

[287] H. Spahn and E. Schlunder. “The scale-up of activated carbon columns for water purification, based on results from batch tests—I: theoretical and experimental determination of adsorption rates of single organic solutes in batch tests”. *Chem. Eng. Sci.* 30 (1975), pp. 529–537.

[288] V. Brauch and E. Schlunder. "The scale-up of activated carbon columns for water purification, based on results from batch tests—II: theoretical and experimental determination of breakthrough curves in activated carbon columns". *Chem. Eng. Sci.* 30 (1975), pp. 539–548.

[289] O. Levenspiel. "Chemical reaction engineering". *Ind. Eng. Chem. Res.* 38 (1999), pp. 4140–4143.

[290] X. Yang, S. Otto, and B. Al-Duri. "Concentration-dependent surface diffusivity model (CDSDM): numerical development and application". *Chem. Eng. J.* 94 (2003), pp. 199–209.

[291] A. Rudin and P. Choi. *The Elements of Polymer Science and Engineering: An Introductory Text for Engineers and Chemists*. 3rd ed. Oxford, UK: Academic Press, 2012.

[292] H. A. Kooijman and R. Taylor. "Estimation of Diffusion-Coefficients in Multicomponent Liquid-Systems". *Ind. Eng. Chem. Res.* 30 (1991), pp. 1217–1222.

[293] D. Theodorou, R. Snurr, and A. Bell. "Molecular Dynamics and Diffusion in Microporous Materials". *Comprehensive Supramolecular Chemistry*. Ed. by G. Alberti and T. Bein. 1st ed. Vol. 7. Oxford: Pergamon Oxford, 1996. Chap. 18, pp. 507–548.

[294] R. Krishna. "Diffusion in porous crystalline materials". *Chem. Soc. Rev.* 41 (2012), pp. 3099–3118.

[295] R. Krishna and R. Taylor. "Multicomponent Mass Transfer-Theory and Applications". *Handbook of Heat and Mass Transfer*. 3rd ed. Houston: Gulf Publishing Co., 1986. Chap. 7.

[296] R. Krishna and J. M. van Baten. "Investigating the validity of the Bosanquet formula for estimation of diffusivities in mesopores". *Chem. Eng. Sci.* 69 (2012), pp. 684–688.

[297] H. Ramanan and S. Auerbach. "Molecular Dynamics and Diffusion in Microporous Materials". *Fluid Transport in Nanoporous Materials*. Ed. by W. C. Conner and J. Fraissard. Proceedings of the NATO Advanced Study Institute. Dordrecht: Springer, 2006, pp. 93–124.

[298] R. Krishna and J. van Baten. "Describing binary mixture diffusion in carbon nanotubes with the Maxwell-Stefan equations. An investigation using molecular dynamics simulations." *Ind. Eng. Chem. Res.* 45 (2006), pp. 2084–2093.

[299] D. Dubbeldam and R. Snurr. "Recent developments in the molecular modeling of diffusion in nanoporous materials". *Mol. Simul.* 33 (2007), pp. 305–325.

[300] D. Valkovska and K. Danov. "Determination of bulk and surface diffusion coefficients from experimental data for thin liquid film drainage". *J. Colloid Interface Sci.* 223 (2000), pp. 314–316.

[301] M. LeVan, G. Carta, and M. Carmen. “Adsorption and Ion Exchange”. *Perry’s Chemical Engineers’ Handbook*. Ed. by R. Perry and D. Green. New York: McGraw-Hill, 2008. Chap. 16.

[302] D. Hand, J. Crittenden, and W. Thacker. “Simplified models for design of fixed-bed adsorption systems”. *J. Environ. Eng.* 110 (1984), pp. 440–456.

[303] S. Baup, C. Jaffre, D. Wolbert, and A. Laplanche. “Adsorption of pesticides onto granular activated carbon: determination of surface diffusivities using simple batch experiments”. *Adsorption* 6 (2000), pp. 219–228.

[304] A. Ozdural. “2.51 - Modeling Chromatographic Separation”. *Comprehensive Biotechnology*. 3rd ed. Oxford: Pergamon, 2017, pp. 754–772.

[305] B. Al-Duri. “Adsorption Modelling and Mass Transfer”. *Use of Adsorbents for the Removal of Pollutants from Wastewaters*. Ed. by G. McKay. Boca Raton: CRC Press, 1996. Chap. 7, pp. 133–173.

[306] C. H. Liaw, J. S. P. Wang, R. A. Greenkorn, and K. C. Chao. “Kinetics of Fixed-Bed Adsorption - New Solution”. *AIChE J.* 25 (1979), pp. 376–381.

[307] J. Alvarez-Ramirez, G. Fernandez-Anaya, F. Valdes-Parada, and J. Ochoa-Tapia. “Physical consistency of generalized Linear Driving Force models for adsorption in a particle”. *Ind. Eng. Chem. Res.* 44 (2005), pp. 6776–6783.

[308] R. Krishna. “A Maxwell-Stefan-Glueckauf description of transient mixture uptake in microporous adsorbents”. *Sep. Purif. Technol.* 191 (2018), pp. 392–399.

[309] J. Mota and A. Rodrigo. “Calculations of multicomponent adsorption-column dynamics combining the potential and Ideal Adsorbed Solution Theories”. *Ind. Eng. Chem. Res.* 39 (2000), pp. 2459–2467.

[310] A. Sperlich, S. Schimmelpfennig, B. Baumgarten, A. Genz, G. Amy, E. Worch, and M. Jekel. “Predicting anion breakthrough in Granular Ferric Hydroxide (GFH) adsorption filters”. *Water Res.* 42 (2008), pp. 2073–2082.

[311] J. Bodin. “MFIT 1.0. 0: Multi-Flow Inversion of Tracer breakthrough curves in fractured and karst aquifers”. *Geosci. Model Dev.* 13 (2020), pp. 2905–2924.

[312] A. Kane, S. Giraudet, J.-B. Vilmain, and P. Le Cloirec. “Intensification of the temperature-swing adsorption process with a heat pump for the recovery of dichloromethane”. *J. Environ. Chem. Eng.* 3 (2015), pp. 734–743.

[313] A. Möller, R. Eschrich, C. Reichenbach, J. Guderian, M. Lange, and J. Möllmer. "Dynamic and equilibrium-based investigations of CO_2-removal from CH_4-rich gas mixtures on microporous adsorbents". *Adsorption* 23 (2017), pp. 197–209.

[314] *Aspen Adsorption*. accessed on 14-12-2022. URL: https://www.aspentech.com/en/products/pages/aspen-adsorption.

[315] R. Zhang, Y. Shen, Z. Tang, W. Li, and D. Zhang. "A Review of Numerical Research on the Pressure Swing Adsorption Process". *Processes* 10 (2022), p. 812.

[316] *gPROMS*. https://www.psenterprise.com/products/gproms. accessed on 14-12-2022.

[317] *MATLAB*. https://nl.mathworks.com/products/matlab.html. accessed on 14-12-2022.

[318] *Ansys Fluent: Fluid simulation software*. https://www.ansys.com/products/fluids/ansys-fluent. accessed on 14-12-2022.

[319] H. Darcy. *Les fontaines publiques de la ville de Dijon: Exposition et application des principes à suivre et des formules à employer dans les questions de distribution d'eau: Ouvrage terminé par un appendice relatif aux fournitures d'eau de plusieurs villes, au filtrage des eaux et à la fabrication des tuyaux de fonte, de plomb, de tôle et de bitume*. 1st ed. Vol. 2. Paris: V. Dalmont, 1856.

[320] S. Ergun. "Fluid flow through packed columns". *Chem. Eng. Prog.* 48 (1952), pp. 89–94.

[321] P. C. Carman. "Fluid flow through a granular bed". *Trans. Inst. Chem. Eng. London* 15 (1937), pp. 150–156.

[322] A. H. D. Cheng and D. T. Cheng. "Heritage and early history of the boundary element method". *Eng. Ana. Bound. Elmnts.* 29 (2005), pp. 268–302.

[323] H. V. Mott and Z. A. Green. "On Danckwerts' boundary conditions for the plug-flow with dispersion/reaction model". *Chem. Eng. Comm.* 202 (2015), pp. 739–745.

[324] N. Ottosen and H. Petersson. *Introduction to Finite Element Method*. 1st ed. New York: Prentice Hall, 1992.

[325] S. Nilchan and C. C. Pantelides. "On the optimisation of periodic adsorption processes". *Adsorption* 4 (1998), pp. 113–147.

[326] P. Cruz, J. C. Santos, F. D. Magalhães, and A. Mendes. "Simulation of separation processes using finite volume method". *Comput. Chem. Eng.* 30 (2005), pp. 83–98.

[327] E. S. Kikkinides and R. T. Yang. "Effects of bed pressure drop on isothermal and adiabatic adsorber dynamics". *Chem. Eng. Sci.* 48 (1993), pp. 1545–1555.

[328] W. Schiesser. *Computational mathematics in engineering and applied science: ODEs, DAEs, and PDEs*. 1st ed. Boca Raton: CRC press, 1994.

[329] A. Iserles. *A first course in the numerical analysis of differential equations*. 2nd ed. Cambridge: Cambridge University Press, 2008.

[330] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes in Fortran 90: The Art of Scientific Computing*. 3rd ed. New York: Cambridge University Press, 1996.

[331] J. Verwer and B. Sommeijer. "An Implicit - Explicit Runge - Kutta - Chebyshev Scheme for Diffusion - Reaction Equations". *SIAM J. Sci. Comput.* 25 (2004), pp. 1824–1835.

[332] P. Deuflhard, E. Hairer, and J. Zugck. "One-step and extrapolation methods for differential-algebraic systems". *Numer. Math.* 51 (1987), pp. 501–516.

[333] J. Butcher. "Implicit Runge-Kutta processes". *Math. Comput.* 18 (1964), pp. 50–64.

[334] W. Press and S. Teukolsky. "Adaptive stepsize Runge-Kutta integration". *Comput. Phys.* 6 (1992), pp. 188–191.

[335] A. Torres-Knoop and D. Dubbeldam. "Exploiting Large-Pore Metal-Organic Frameworks for Separations through Entropic Molecular Mechanisms". *Chem. Phys. Chem.* 16 (2015), pp. 2046–2067.

[336] E. Jolimaitre, K. Ragil, M. Tayakout-Fayolle, and C. Jallut. "Separation of mono and dibranched hydrocarbons on silicalite". *AIChE J.* 48 (2002), pp. 1927–1937.

[337] M. Caceci and W. Cacheris. "Fitting curves to data". *Byte* 9 (1984), pp. 340–362.

[338] H. Motulsky and L. Ransnas. "Fitting curves to data using nonlinear regression: a practical and non-mathematical review". *The FASEB Journal* 1 (1987), pp. 365–374.

[339] I. Jaraíz-Arroyo, A. Martin-Calvo, J. Gutiérrez-Sevillano, C. Barranco, N. Diaz-Diaz, and S. Calero. "OCEAN: An Algorithm to Predict the Separation of Biogas Using Zeolites". *Ind. Eng. Chem. Res.* 59 (2020), pp. 7212–7223.

[340] R. Anderson and D. Gómez-Gualdróna. "Deep learning combined with IAST to screen thermodynamically feasible MOFs for adsorption-based separation of multiple binary mixtures". *J. Chem. Phys.* 154 (2021), p. 234102.

[341] H. Lineweaver and D. Burk. "The Determination of Enzyme Dissociation Constants". *J. Am. Chem. Soc.* 56 (1934), pp. 658–666.

[342] J. Dowd and D. Riggs. "A comparison of estimates of Michaelis-Menten kinetic constants from various linear transformations". *J. Biol. Chem.* 240 (1965), pp. 863–869.

[343] G. Eadie. “The inhibition of cholinesterase by physostigmine and prostigmine”. *J. Bio. Chem.* 146 (1942), pp. 85–93.

[344] B. Hofstee. “Non-Inverted Versus Inverted Plots in Enzyme Kinetics”. *Nature* 184 (1959), pp. 1296–1298.

[345] G. Scatchard. “The attractions of proteins for small molecules and ions”. *Ann. N. Y. Acad. Sci.* 51 (1949), pp. 660–672.

[346] D. G. Kinniburgh. “General purpose adsorption isotherms”. *Environ. Sci. Technol.* 20 (1986), pp. 895–904.

[347] Á. Magreñán and I. Argyros. *A Contemporary Study of Iterative Methods*. 1st ed. London: Elsevier, 2018.

[348] K. Levenberg. “A method for the solution of certain non-linear problems in least squares”. *Quart. Appl. Math.* 2 (1944), pp. 164–168.

[349] A. Girard. “Excerpt from Revue d’optique théorique et instrumentale”. *Rev. Opt* 37 (1958), pp. 225–241.

[350] C. Wynne. “Lens Designing by Electronic Digital Computer: I”. *Proc. Phys. Soc.* 73 (1959), pp. 777–787.

[351] D. Morrison. “Methods for nonlinear least squares problems and convergence proofs”. *JPL Seminar Proc. on tracking programs and orbit determination*. 1960.

[352] D. Marquardt. “An Algorithm for Least-Squares Estimation of Nonlinear Parameters”. *J. Soc. Ind. Appl. Math.* 11 (1963), pp. 431–441.

[353] D. Sorensen. “Newton’s Method with a Model Trust Region Modification”. *SIAM J. Numer. Anal.* 19 (1982), pp. 409–426.

[354] M. Hestenes and E. Stiefel. “Methods of conjugate gradients for solving”. *J. Res. Nat. Bur. Stand.* 49 (1952), p. 409.

[355] C. Broyden. “The Convergence of a Class of Double-rank Minimization Algorithms 1. General Considerations”. *IMA J. Appl. Math.* 6 (1970), pp. 76–90.

[356] R. Fletcher. “A new approach to variable metric algorithms”. *Comput. J.* 13 (1970), pp. 317–322.

[357] D. Goldfarb. “A family of variable-metric methods derived by variational means”. *Math. Comput.* 24 (1970), pp. 23–26.

[358] D. Shanno. “Conditioning of quasi-Newton methods for function minimization”. *Math. Comput.* 24 (1970), pp. 647–656.

[359] D. Liu and J. Nocedal. “On the limited memory BFGS method for large scale optimization”. *Math. Program.* 45 (1989), pp. 503–528.

[360] W. Spendley, G. R. Hext, and F. R. Himsworth. “Sequential Application of Simplex Designs in Optimisation and Evolutionary Operation”. *Technometrics* 4 (1962), pp. 441–461.

[361] K. De Jong. “An Analysis of the Behavior of a Class of Genetic Adaptive Systems”. https://hdl.handle.net/2027.42/4507. PhD thesis. University of Michigan, 1975.

[362] M. Mahajan, S. Kumar, and R. Porwal. “Applying genetic algorithm to increase the efficiency of a data flow-based test data generation approach”. *ACM SIGSOFT Software Engineering Notes* 37 (2012), pp. 1–5.

[363] J. Koza and R. Poli. “A Genetic Programming Tutorial”. *Introductory Tutorials in Optimization, Search and Decision Support*. 1st ed. New York: Springer, 2003. Chap. 8, pp. 143–185.

[364] H. Hirsh, W. Banzhaf, J. Koza, C. Ryan, L. Spector, and C. Jacob. “Genetic Programming”. *IEEE Intell. Syst.* 15 (2000), pp. 74–84.

[365] M. Emmerich, O. Shir, and H. Wang. “Evolution Strategies”. *Handbook of Heuristics*. 1st ed. Cham: Springer International Publishing, 2018, pp. 1–31.

[366] A. Eiben and J. Smith. *Introduction to Evolutionary Computing*. 1st ed. Natural Computing Series. Verlag/Berlin/Heidelberg: Springer, 2003, pp. 1–300.

[367] J. Kennedy and R. Eberhart. “Particle Swarm Optimization”. *Proceedings of ICNN’ 95 - International Conference on Neural Networks*. Honolulu: IEEE, 1995.

[368] Y. Shi and R. Eberhart. “A modified particle swarm optimizer”. *1998 IEEE International Conference on Evolutionary Computation Proceedings. IEEE World Congress on Computational Intelligence*. Anchorage: IEEE, 1998.

[369] M. Dorigo, M. Birattari, and T. Stutzle. “Ant colony optimization”. *IEEE Comput. Intell. Mag.* 1 (2006), pp. 28–39.

[370] X. Yang and S. Deb. “Cuckoo search via Lévy flights”. *2009 World congress on nature & biologically inspired computing (NaBIC)*. IEEE. Coimbatore, 2009, pp. 210–214.

[371] S. Mirjalili, S. Mirjalili, and A. Lewis. “Grey wolf optimizer”. *Adv. Eng. Software* 69 (2014), pp. 46–61.

[372] K. Li, S. Li, Z. Huang, M. Zhang, and Z. Xu. “Grey Wolf Optimization algorithm based on Cauchy-Gaussian mutation and improved search strategy”. *Sci. Rep.* 12 (2022).

[373] G. Schütz and S. Trimper. “Elephants can always remember: Exact long-range memory effects in a non-Markovian random walk”. *Phys. Rev. E* 70 (2004).

[374] A. Gopinath and K. Aravamudan. “A novel, initial guess free optimization algorithm for estimating parameters of batch kinetics model used to simulate adsorption of pollutant molecules in aqueous streams”. *J. Mol. Liq.* 275 (2019), pp. 510–522.

[375] D. G. Kinniburgh. “ISOTHERM. A computer program for analyzing adsorption data”. *British Geological Survey, Wallingford* (1985). Report WD/ST/85/02.

[376] D. Kinniburgh. “FIT: User Guide”. *British Geological Survey, Wallingford* (1985). Report WD/93/23.

[377] L. Matott and A. Rabideau. “ISOFIT – A program for fitting sorption isotherms to experimental data”. *Environ. Model. Softw.* 23 (2008), pp. 670–676.

[378] P. Virtanen, R. Gommers, T. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. van der Walt, M. Brett, J. Wilson, K. Millman, N. Mayorov, A. Nelson, E. Jones, R. Kern, E. Larson, C. Carey, I. Polat, Y. Feng, E. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. Quintero, C. Harris, A. Archibald, A. Ribeiro, F. Pedregosa, P. van Mulbregt, and SciPy 1.0 Contributors. “SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python”. *Nat. Methods* 17 (2020), pp. 261–272.

[379] B. Gough. *GNU scientific library reference manual*. 3rd ed. Devon: Network Theory Ltd., 2009.

[380] W. Thomas, K. Colin, and many others. *GNUPLOT 4.6: an interactive plotting program*. http://gnuplot.sourceforge.net/. 2013.

[381] J. W. Eaton, D. Bateman, S. Hauberg, and R. Wehbring. *GNU Octave version 8.1.0 manual: a high-level interactive language for numerical computations*. 2023. URL: https://octave.org/doc/v8.1.0/.

[382] O. Corporation. *Origin(Pro), 2020b*. OriginLab Corporation, Northampton, MA, USA. 2020.

[383] P. Charbonneau. “Genetic algorithms in astronomy and astrophysics”. *Astrophys. J., Suppl. Ser.* 101 (1995), p. 309.

[384] J. C. Lagarias, J. A. Reeds, M. H. Wright, and P. E. Wright. “Convergence Properties of the Nelder-Mead Simplex Method in Low Dimensions”. *SIAM J. Optim.* 9 (1998), pp. 112–147.

[385] F. Gao and L. Han. “Implementing the Nelder-Mead simplex algorithm with adaptive parameters”. *Comput. Optim. Appl.* 51 (2010), pp. 259–277.

[386] R. Byrd, R. Schnabel, and G. Shultz. “A Trust Region Algorithm for Nonlinearly Constrained Optimization”. *SIAM J. Numer. Anal.* 24 (1987), pp. 1152–1170.

[387] M. Branch, T. Coleman, and Y. Li. “A Subspace, Interior, and Conjugate Gradient Method for Large-Scale Bound-Constrained Minimization Problems”. *SIAM J. Sci. Comput.* 21 (1999), pp. 1–23.

[388] S. Tomar. “Converting video formats with FFmpeg”. *Linux Journal* 2006 (2006), p. 10.

[389] P. Janert. *Gnuplot in action: understanding data with graphs*. 2nd ed. Greenwich: Manning Publications Co., 2015.

[390] S. Sharma, M. S. Rigutto, E. Zuidema, U. Agarwal, R. Baur, D. Dubbeldam, and T. J. H. Vlugt. “Understanding shape selectivity effects of hydroisomerization using a reaction equilibrium model”. *J. Chem. Phys.* 160 (2024), p. 214708.

[391] S. van Bavel, S. Verma, E. Negro, and M. Bracht. “Integrating CO_2 electrolysis into the gas-to-liquids–power-to-liquids process”. *ACS Energy Lett.* 5 (2020), pp. 2597–2601.

[392] V. Kumaravel, J. Bartlett, and S. C. Pillai. “Photoelectrochemical conversion of carbon dioxide (CO_2) into fuels and value-added products”. *ACS Energy Lett.* 5 (2020), pp. 486–519.

[393] H. W. Ryu, D. H. Kim, J. Jae, S. S. Lam, E. D. Park, and Y.-K. Park. “Recent advances in catalytic co-pyrolysis of biomass and plastic waste for the production of petroleum-like hydrocarbons”. *Bioresour. Technol.* 310 (2020), p. 123473.

[394] H. Calis, W. Lüke, I. Drescher, and A. Schütze. “Synthetic Diesel Fuels”. *Handbook of Fuels: Energy Sources for Transportation* (2021), pp. 161–200.

[395] B. Smit and T. L. Maesen. “Towards a molecular understanding of shape selectivity”. *Nature* 451 (2008), pp. 671–678.

[396] M. S. Rigutto, R. van Veen, and L. Huve. “Zeolites in hydrocarbon processing”. *Stud. Surf. Sci. Catal.* (2007), p. 855.

[397] P. Mäki-Arvela, T. A. Kaka khel, M. Azkaar, S. Engblom, and D. Y. Murzin. “Catalytic hydroisomerization of long-chain hydrocarbons for the production of fuels”. *Catalysts* 8 (2018), p. 534.

[398] N. Viswanadham, R. Kamble, S. K. Saxena, and M. Garg. “Studies on octane boosting of industrial feedstocks on Pt/H-BEA zeolite”. *Fuel* 87 (2008), pp. 2394–2400.

[399] A. Soualah, J. L. Lemberton, L. Pinard, M. Chater, P. Magnoux, and K. Moljord. “Hydroisomerization of long-chain n-alkanes on bifunctional Pt/zeolite catalysts: Effect of the zeolite structure on the product selectivity and on the reaction mechanism”. *Appl. Catal., A* 336 (2008), pp. 23–28.

[400] A. Poursaeidesfahani. “Zeolite-based separation and production of branched hydrocarbons”. PhD thesis. Delft University of Technology, 2019. URL: https://doi.org/10.4233/uuid:770acc49-69e0-448d-8869-9dd01aca7e19.

[401] M. Schenk. “Shape selectivity in zeolites”. PhD thesis. University of Amsterdam, 2003. URL: https://hdl.handle.net/11245/1.212114.

[402] J. F. Denayer, G. V. Baron, G. Vanbutsele, P. A. Jacobs, and J. A. Martens. “Evidence for alkylcarbenium ion reaction intermediates from intrinsic reaction kinetics of C_6–C_9 n-alkane hydroisomerization and hydrocracking on Pt/H–Y and Pt/USY Zeolites”. *J. Catal.* 190 (2000), pp. 469–473.

[403] D. R. Stull, E. F. Westrum, and G. C. Sinke. *The Chemical Thermodynamics of Organic Compounds*. 1st ed. New York: John Wiley & Sons Inc, 1969.

[404] M. D. Macedonia and E. J. Maginn. “Impact of confinement on zeolite cracking selectivity via Monte Carlo integration”. *AIChE J.* 46 (2000), pp. 2504–2517.

[405] W. R. Smith and B. Triska. “The reaction ensemble method for the computer simulation of chemical and phase equilibria. I. Theory and basic examples”. *J. Chem. Phys.* 100 (1994), pp. 3019–3027.

[406] J. K. Johnson, A. Z. Panagiotopoulos, and K. E. Gubbins. “Reactive Canonical Monte Carlo: a new simulation technique for reacting or associating fluids”. *Mol. Phys.* 81 (1994), pp. 717–733.

[407] A. Poursaeidesfahani, R. Hens, A. Rahbari, M. Ramdin, D. Dubbeldam, and T. J. H. Vlugt. “Efficient application of Continuous Fractional Component Monte Carlo in the reaction ensemble”. *J. Chem. Theory Comput.* 13 (2017), pp. 4452–4466.

[408] I. Matito-Martos, A. Rahbari, A. Martin-Calvo, D. Dubbeldam, T. J. H. Vlugt, and S. Calero. “Adsorption equilibrium of nitrogen dioxide in porous materials”. *Phys. Chem. Chem. Phys.* 20 (2018), pp. 4189–4199.

[409] D. O. Wasik, A. Martín-Calvo, J. J. Gutiérrez-Sevillano, D. Dubbeldam, T. J. H. Vlugt, and S. Calero. “Enhancement of formic acid production from carbon dioxide hydrogenation using metal-organic frameworks: Monte Carlo simulation study”. *Chem. Eng. J.* 467 (2023), p. 143432.

[410] N. Hansen, S. Jakobtorweihen, and F. J. Keil. “Reactive Monte Carlo and Grand-Canonical Monte Carlo simulations of the propene metathesis reaction system”. *J. Chem. Phys.* 122 (2005), p. 164705.

[411] R. Hens. “Molecular simulation of phase and reaction equilibria: Software and algorithm development”. PhD thesis. Delft University of Technology, 2020. URL: https://doi.org/10.4233/uuid:41c32a8f-2db7-4091-abb5-4d6a5e596345.

[412] S. I. Sandler. *Chemical, biochemical, and engineering thermodynamics*. 4th ed. New Jersey: John Wiley & Sons, 2017.

[413] R. Hens, A. Rahbari, S. Caro-Ortiz, N. Dawass, M. Erdös, A. Poursaeidesfahani, H. S. Salehi, A. T. Celebi, M. Ramdin, O. A. Moultos, D. Dubbeldam, and T. J. H. Vlugt. "Brick-CFCMC: Open source software for Monte Carlo simulations of phase and reaction equilibria using the Continuous Fractional Component method". *J. Chem. Inf. Model.* 60 (2020), pp. 2678–2682.

[414] H. M. Polat, H. S. Salehi, R. Hens, D. O. Wasik, A. Rahbari, F. De Meyer, C. Houriez, C. Coquelet, S. Calero, D. Dubbeldam, O. A. Moultos, and T. J. H. Vlugt. "New features of the open source Monte Carlo software Brick-CFCMC: thermodynamic integration and hybrid trial moves". *J. Chem. Inf. Model.* 61 (2021), pp. 3752–3757.

[415] A. Rahbari, R. Hens, M. Ramdin, O. A. Moultos, D. Dubbeldam, and T. J. H. Vlugt. "Recent advances in the Continuous Fractional Component Monte Carlo methodology". *Mol. Simul.* 47 (2021), pp. 804–823.

[416] D. W. Scott. *Chemical thermodynamic properties of hydrocarbons and related substances: properties of the alkane hydrocarbons, C_1 through C_{10}, in the ideal gas state from 0 to 1500 K.* 1st ed. Washington DC: US Department of the Interior, Bureau of Mines, 1974.

[417] W. R. Smith and R. W. Missen. *Chemical reaction equilibrium analysis: theory and algorithms.* 1st ed. New York: Wiley, 1982.

[418] H. M. Polat, F. de Meyer, C. Houriez, O. A. Moultos, and T. J. H. Vlugt. "Solving chemical absorption equilibria using free energy and quantum chemistry calculations: methodology, limitations, and new open-source software". *J. Chem. Theory Comput.* 19 (2023), pp. 2616–2629.

[419] 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. J. Montgomery, J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, 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, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox. "Gaussian 09, Revision D. 01, Gaussian, Inc., Wallingford CT". *Gaussian 09*. 9th ed. Gaussian Inc Wallingford, 2009. URL: http://www.%20gaussian.%20com.

[420] M. W. Chase. *NIST-JANAF thermochemical tables*. 4th ed. Vol. 9. https://janaf.nist.gov/. Washington DC: American Chemical Society, 1998.

[421] P. J. Linstrom and W. G. Mallard. "The NIST Chemistry WebBook: A chemical data resource on the internet". *J. Chem. Eng. Data* 46 (2001), pp. 1059–1063.

[422] D. W. Scott. "Correlation of the chemical thermodynamic properties of alkane hydrocarbons". *J. Chem. Phys.* 60 (1974), pp. 3144–3165.

[423] L. Constantinou and R. Gani. "New group contribution method for estimating properties of pure compounds". *AIChE J.* 40 (1994), pp. 1697–1710.

[424] L. A. Curtiss, K. Raghavachari, P. C. Redfern, and J. A. Pople. "Assessment of Gaussian-2 and density functional theories for the computation of enthalpies of formation". *J. Chem. Phys.* 106 (1997), pp. 1063–1079.

[425] A. Nicolaides, A. Rauk, M. N. Glukhovtsev, and L. Radom. "Heats of formation from G2, G2 (MP2), and G2 (MP2, SVP) total energies". *J. Phys. Chem.* 100 (1996), pp. 17460–17464.

[426] C. J. Cramer. *Essentials of computational chemistry: theories and models*. 2nd ed. Chichester: John Wiley & Sons, 2013.

[427] V. Van Speybroeck, R. Gani, and R. J. Meier. "The calculation of thermodynamic properties of molecules". *Chem. Soc. Rev.* 39 (2010), pp. 1764–1779.

[428] J. B. Ott and J. Boerio-Goates. *Chemical thermodynamics: advanced applications*. 1st ed. Vol. 2. Cornwall: Academic Press, 2000.

[429] O. Levenspiel. *Chemical reaction engineering*. 3rd ed. New York: John Wiley & Sons, 1998.

[430] J.-P. Ryckaert and A. Bellemans. "Molecular dynamics of liquid alkanes". *Faraday Discuss. Chem. Soc.* 66 (1978), pp. 95–106.

[431] D.-Y. Peng and D. B. Robinson. "A new two-constant equation of state". *Ind. Eng. Chem. Fundam.* 15 (1976), pp. 59–64.

[432] D. Nicholson and N. G. Parsonage. *Computer simulation and the statistical mechanics of adsorption*. 2nd ed. New York: Academic Press, 1988.

[433] B. Smit. "Grand-Canonical Monte Carlo simulations of chain molecules: adsorption isotherms of alkanes in zeolites". *Mol. Phys.* 85 (1995), pp. 153–172.

[434] T. O. Bok, E. P. Andriako, E. E. Knyazeva, and I. I. Ivanova. "Engineering of zeolite BEA crystal size and morphology via seed-directed steam assisted conversion". *RSC Adv.* 10 (2020), pp. 38505–38514.

[435] S. Sharma, J. J. Sleijfer, J. op de Beek, S. van der Zeeuw, D. Zorzos, S. Lasala, M. S. Rigutto, E. Zuidema, U. Agarwal, R. Baur, S. Calero, D. Dubbeldam, and T. J. H. Vlugt. "Prediction of Thermochemical Properties of Long-Chain Alkanes Using Linear Regression: Application to Hydroisomerization". *J. Phys. Chem. B* 128 (2024), pp. 9619–9629.

[436] T. M. Letcher. *Chemical thermodynamics for industry*. 1st ed. Cambridge: Royal Society of Chemistry, 2004.

[437] H. Peng, D. Zhang, X. Ling, Y. Li, Y. Wang, Q. Yu, X. She, Y. Li, and Y. Ding. "n-Alkanes phase change materials and their microencapsulation for thermal energy storage: a critical review". *Energy Fuels* 32 (2018), pp. 7262–7293.

[438] F. M. Fraser and E. J. Prosen. "Heats of combustion of liquid n-hexadecane, 1-hexadecene, n-decylbenzene, n-decylcyclohexane, n-decylcyclopentane, and the variation of heat of combustion with chain length". *J. Res. Natl. Bur. Stand.* 55 (1955), pp. 329–333.

[439] E. Prosen, K. Pitzer, and F. Rossini. "Heats and free energies of formation of the paraffin hydrocarbons, in the gaseous state, to 1500 degree". *Natl. Bur. Stand* 34 (1945), p. 403.

[440] F. D. Rossini. *Selected Values of Properties of Hydrocarbons: Prepared as Part of the Work of the American Petroleum Institute Research Project 44*. 1st ed. Washington, D.C.: US Department of Commerce, National Bureau of Standards, 1947, pp. 175–356.

[441] R. Morales-Rodriguez. *Thermodynamics: Fundamentals and Its Application in Science*. 1st ed. Rijeka: InTech, 2012.

[442] J. Marrero and R. Gani. "Group-contribution based estimation of pure component properties". *Fluid Phase Equilibria* 183 (2001), pp. 183–208.

[443] A. S. Hukkerikar, R. J. Meier, G. Sin, and R. Gani. "A method to estimate the enthalpy of formation of organic compounds with chemical accuracy". *Fluid Phase Equilib.* 348 (2013), pp. 23–32.

[444] T. A. Albahri and A. F. Aljasmi. "SGC method for predicting the standard enthalpy of formation of pure compounds from their molecular structures". *Thermochim. Acta* 568 (2013), pp. 46–60.

[445] E. S. Domalski and E. D. Hearing. "Estimation of the Thermodynamic Properties of Hydrocarbons at 298.15 K". *J. Phys. Chem. Ref. Data* 17 (1988), pp. 1637–1678.

[446] C. L. Yaws. *Yaws' Handbook of Thermodynamic Properties for Hydrocarbons and Chemicals*. 1st ed. New York: Knovel, 2009. URL: https://app.knovel.com/hotlink/toc/id:kpYHTPHC09/yaws-handbook-thermodynamic.

[447] J. C. Bloxham, M. E. Redd, N. F. Giles, T. A. Knotts IV, and W. V. Wilding. “Proper use of the DIPPR 801 database for creation of models, methods, and processes”. *J. Chem. Eng. Data* 66 (2020), pp. 3–10.

[448] M. Y. Hayes, B. Li, and H. Rabitz. “Estimation of molecular properties by high-dimensional model representation”. *J. Phys. Chem. A* 110 (2006), pp. 264–272.

[449] D. Weininger. “SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules”. *J. Chem. Inf. Comput. Sci.* 28 (1988), pp. 31–36.

[450] D. Weininger, A. Weininger, and J. L. Weininger. “SMILES. 2. Algorithm for generation of unique SMILES notation”. *J. Chem. Inf. Comput. Sci.* 29 (1989), pp. 97–101.

[451] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, İ. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, and SciPy 1.0 Contributors. “SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python”. *Nat. Methods* 17 (2020), pp. 261–272. DOI: 10.1038/s41592-019-0686-2.

[452] J. I. Siepmann and D. Frenkel. “Configurational Bias Monte Carlo: a new sampling scheme for flexible chains”. *Mol. Phys.* 75 (1992), pp. 59–70.

[453] J. J. De Pablo, M. Laso, J. I. Siepmann, and U. W. Suter. “Continuum-configurational-bias Monte Carlo simulations of long-chain alkanes”. *Mol. Phys.* 80 (1993), pp. 55–63.

[454] S. K. Nath and R. Khare. “New forcefield parameters for branched hydrocarbons”. *J. Chem. Phys.* 115 (2001), pp. 10837–10844.

[455] J. van Baten. *AmsterChem: tailor made engineering software solutions*. https://www.amsterchem.com/. Accessed on 24.07.2024.

[456] Z. Li, L. Constantinou, R. Baur, D. Dubbeldam, S. Calero, S. Sharma, M. Rigutto, P. Dey, and T. J. H. Vlugt. “Second-order Group Contribution Method for T_c, P_c, ω, ΔG_f^0, ΔH_f^0 and Liquid Densities of Linear and Branched Alkanes”. *Mol. Phys.* (2025). Manuscript accepted.

[457] S. Sharma, R. Baur, M. S. Rigutto, E. Zuidema, U. Agarwal, S. Calero, D. Dubbeldam, and T. J. H. Vlugt. “Computing Entropy for Long-Chain Alkanes Using Linear Regression: Application to Hydroisomerization”. *Entropy* 26 (2024), p. 1120.

[458] P. Ruelle and U. W. Kesselring. “Aqueous solubility prediction of environmentally important chemicals from the mobile order thermodynamics”. *Chemosph.* 34 (1997), pp. 275–298.

[459] M. J. Moran, H. N. Shapiro, D. D. Boettner, and M. B. Bailey. *Fundamentals of engineering thermodynamics*. 7th ed. Hoboken, USA: John Wiley & Sons, 2010.

[460] M. Kanoğlu, Y. A. Çengel, and İ. Dinçer. *Efficiency evaluation of energy systems*. 1st ed. New York, USA: Springer Science & Business Media, 2012.

[461] S. Kjelstrup, D. Bedeaux, E. Johannessen, and J. Gross. *Non equilibrium thermodynamics for engineers*. 2nd ed. Singapore: World Scientific Publishing Co. Pte. Ltd, 2010.

[462] H. Drucker, C. J. Burges, L. Kaufman, A. Smola, and V. Vapnik. “Support vector regression machines”. *Advances in Neural Information Processing Systems* 9 (1996), pp. 155–161.

[463] B. Schölkopf. “New support vector algorithms”. *Neural Computation* 12 (2000), pp. 1083–1121.

[464] Y. A. Ran, S. Sharma, S. R. G. Balestra, Z. Li, S. Calero, T. J. H. Vlugt, R. Q. Snurr, and D. Dubbeldam. “RASPA3: A Monte Carlo code for computing adsorption and diffusion in nanoporous materials and thermodynamics properties of fluids”. *J. Chem. Phys.* 161 (2024), p. 114106.

[465] P. Perrot. *A to Z of Thermodynamics*. 1st ed. Guildford, UK: Oxford University Press, USA, 1998.

[466] S. R. Rieder, M. P. Oliveira, S. Riniker, and P. H. Hünenberger. “Development of an open-source software for isomer enumeration”. *J. Cheminf.* 15 (2023), p. 10.

[467] M. Swain. *PubChemPy: A Python wrapper for the PubChem PUG REST API*. https://github.com/mcs07/PubChemPy. 2023.

[468] S. Kim, P. A. Thiessen, E. E. Bolton, J. Chen, G. Fu, A. Gindulyte, L. Han, J. He, S. He, B. A. Shoemaker, J. Wang, B. Yu, and S. H. Bryant. “PubChem substance and compound databases”. *Nucleic Acids Res.* 44 (2016), pp. D1202–D1213.

[469] M. A. Yirik, M. Sorokina, and C. Steinbeck. “MAYGEN: an open-source chemical structure generator for constitutional isomers based on the orderly generation principle”. *J. Cheminf.* 13 (2021), pp. 1–14.

[470] B. D. McKay, M. A. Yirik, and C. Steinbeck. “Surge: a fast open-source chemical graph generator”. *J. Cheminf.* 14 (2022), p. 24.

[471] H. A. Favre and W. H. Powell. *Nomenclature of organic chemistry: IUPAC recommendations and preferred names 2013*. Cambridge: Royal Society of Chemistry, 2013. DOI: 10.1039/9781849733069.

[472] M. Wu, S. Zhang, and J. Ren. “AI-empowered digital design of zeolites: Progress, challenges, and perspectives”. *APL Materials* 13 (2025).

[473] E. Xie, X. Wang, J. I. Siepmann, H. Chen, and R. Q. Snurr. “Generative AI for Design of Nanoporous Materials: Review and Future Prospects”. *Digital Discovery* 4 (2025).

[474] X. Yu, S. Choi, D. Tang, A. J. Medford, and D. S. Sholl. “Efficient models for predicting temperature-dependent Henry’s constants and adsorption selectivities for diverse collections of molecules in metal–organic frameworks”. *J. Phys. Chem. C* 125 (2021), pp. 18046–18057.

[475] F. Gharagheizi and D. S. Sholl. “Comprehensive assessment of the accuracy of the ideal adsorbed solution theory for predicting binary adsorption of gas mixtures in porous materials”. *Ind. Eng. Chem. Res.* 61 (2021), pp. 727–739.

[476] A. S. Daou, H. Fang, S. E. Boulfelfel, P. I. Ravikovitch, and D. S. Sholl. “Machine Learning and IAST-Aided High-Throughput Screening of Cationic and Silica Zeolites for Alkane Capture, Storage, and Separations”. *J. Phys. Chem. C* 128 (2024), pp. 6089–6105.

[477] Y. Liu, G. Perez, Z. Cheng, A. Sun, S. C. Hoover, W. Fan, S. Maji, and P. Bai. “ZeoNet: 3D convolutional neural networks for predicting adsorption in nanoporous zeolites”. *J. Mater. Chem. A* 11 (2023), pp. 17570–17580.

[478] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. “Gradient-Based Learning Applied to Document Recognition”. *Proc. IEEE* 86 (1998), pp. 2278–2324.

[479] M. W. Deem, R. Pophale, P. A. Cheeseman, and D. J. Earl. “Computational discovery of new zeolite-like materials”. *J. Phys. Chem. C* 113 (2009), pp. 21353–21360.

[480] D. Dubbeldam, S. Calero, T. Maesen, and B. Smit. “Understanding the Window Effect in Zeolite Catalysis”. *Angew. Chem., Int. Ed.* 42 (2003), pp. 3624–3626.

[481] M. Aldeghi, D. E. Graff, N. Frey, J. A. Morrone, E. O. Pyzer-Knapp, K. E. Jordan, and C. W. Coley. “Roughness of molecular property landscapes and its impact on modellability”. *J. Chem. Inf. Model.* 62 (2022), pp. 4660–4671.

[482] A. Cherkasov, E. N. Muratov, D. Fourches, A. Varnek, I. I. Baskin, M. Cronin, J. Dearden, P. Gramatica, Y. C. Martin, R. Todeschini, V. Consonni, V. E. Kuz’min, R. Cramer, R. Benigni, C. T. Yang, J. Rathman, L. Terfloth, J. Gasteiger, A. Richard, and A. Tropsha. “QSAR modeling: where have you been? Where are you going to?” *J. Med. Chem.* 57 (2014), pp. 4977–5010.

[483] R. Guha and J. H. Van Drie. “Structure- activity landscape index: identifying and quantifying activity cliffs”. *J. Chem. Inf. Model.* 48 (2008), pp. 646–658.

[484] T. T. Tanimoto. *An elementary mathematical theory of classification and prediction*. 1st ed. New York: International Business Machines Corporation, 1958.

[485] J. W. Godden, L. Xue, and J. Bajorath. “Combinatorial preferences affect molecular similarity/diversity calculations using binary fingerprints and Tanimoto coefficients”. *J. Chem. Inf. Comp. Sci.* 40 (2000), pp. 163–166.

[486] V. I. Levenshtein. “Binary codes capable of correcting deletions, insertions, and reversals”. *Soviet Physics Doklady*. Vol. 10. Soviet Union. 1966, pp. 707–710.

[487] A. M. Wassermann, M. Wawer, and J. Bajorath. “Activity landscape representations for structure- activity relationship analysis”. *J. Med. Chem.* 53 (2010), pp. 8209–8223.

[488] E. Griffen, A. G. Leach, G. R. Robb, and D. J. Warner. “Matched molecular pairs as a medicinal chemistry tool: miniperspective”. *J. Med. Chem.* 54 (2011), pp. 7739–7750.

[489] R. Gómez-Bombarelli, J. N. Wei, D. Duvenaud, J. M. Hernández-Lobato, B. Sánchez-Lengeling, D. Sheberla, J. Aguilera-Iparraguirre, T. D. Hirzel, R. P. Adams, and A. Aspuru-Guzik. “Automatic chemical design using a data-driven continuous representation of molecules”. *ACS Cent. Sci.* 4 (2018), pp. 268–276.

[490] D. Stumpfe and J. Bajorath. “Exploring activity cliffs in medicinal chemistry: miniperspective”. *J. Med. Chem.* 55 (2012), pp. 2932–2942.

[491] Y. Cao, T. Jiang, and T. Girke. “A maximum common substructure-based algorithm for searching and predicting drug-like compounds”. *Bioinformatics* 24 (2008), pp. i366–i374.

[492] K. Huang, T. Fu, W. Gao, Y. Zhao, Y. Roohani, J. Leskovec, C. W. Coley, C. Xiao, J. Sun, and M. Zitnik. “Artificial intelligence foundation for therapeutic science”. *Nat. Chem. Biol.* 18 (2022), pp. 1033–1036.

[493] S. Sharma, P. Yang, Y. Liu, K. Rossi, P. Bai, M. S. Rigutto, E. Zuidema, U. Agarwal, R. Baur, S. Calero, D. Dubbeldam, and T. J. H. Vlugt. “Machine Learning-Based Predictions of Henry Coefficients for Long Chain Alkanes in One Dimensional Zeolites: Application to Hydroisomerization”. *J. Phys. Chem. C* (2025). Manuscript accepted.

[494] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and É. Duchesnay. "Scikit-learn: Machine Learning in Python". *J. Mach. Learn. Res.* 12 (2011), pp. 2825–2830.

[495] A. P. Bento, A. Hersey, E. Félix, G. Landrum, A. Gaulton, F. Atkinson, L. J. Bellis, M. De Veij, and A. R. Leach. "An open source chemical structure curation pipeline using RDKit". *J. Cheminf.* 12 (2020), pp. 1–16.

[496] Y. Xiang, Y.-H. Tang, Z. Gong, H. Liu, L. Wu, G. Lin, and H. Sun. "Efficient Exploration of Chemical Compound Space Using Active Learning for Prediction of Thermodynamic Properties of Alkane Molecules". *J. Chem. Inf. Model.* 63 (2023), pp. 6515–6524.

[497] G. James, D. Witten, T. Hastie, R. Tibshirani, *et al. An introduction to statistical learning*. 2nd ed. Cham: Springer, 2013.

[498] L. Chen, Y. Bai, S. Huang, Y. Lu, B. Wen, A. Yuille, and Z. Zhou. "Making Your First Choice: To Address Cold Start Problem in Medical Active Learning". *Medical Imaging with Deep Learning*. Vol. 227. Proceedings of Machine Learning Research. PMLR, 2024, pp. 496–525. URL: https://proceedings.mlr.press/v227/chen24a.html.

[499] W. X. Shen, C. Cui, X. Su, Z. Zhang, A. Velez-Arce, J. Wang, X. Shi, Y. Zhang, J. Wu, Y. Z. Chen, and M. Zitnik. *Activity Cliff-Informed Contrastive Learning for Molecular Property Prediction*. ChemRxiv preprint. 2024. URL: https://chemrxiv.org/engage/chemrxiv/article-details/6703d9c351558a15ef5b9e06.

[500] O. K. Shirekar, A. Singh, and H. Jamali-Rad. *Self-attention message passing for contrastive few-shot learning*. arXiv preprint. 2023. URL: https://arxiv.org/abs/2210.06339.

[501] M. P. Allen and D. J. Tildesley. *Computer simulation of liquids*. 2nd ed. Oxford: Oxford university press, 2017.

[502] E. G. Lewars. *Computational Chemistry. Introduction to the Theory and Applications of Molecular and Quantum Mechanics*. 2nd ed. New York: Springer, 2011.

[503] J. Caro. "Zeolites and mesoporous materials as advanced functional material". *Studies in Surface Science and Catalysis*. Vol. 154. Elsevier, 2004, pp. 80–93.

[504] P. Bai, M. Y. Jeon, L. Ren, C. Knight, M. W. Deem, M. Tsapatsis, and J. I. Siepmann. "Discovery of optimal zeolites for challenging separations and chemical transformations using predictive materials modeling". *Nat. Commun.* 6 (2015), p. 5912.

[505] S. Gooijer, S. Capelo, S. Sharma, S. Giancola, J. R. Galan-Mascaros, T. J. H. Vlugt, D. Dubbeldam, J. M. Vicent-Luna, and S. Calero. "TAMOF-1 for Capture and Separation of the main Flue Gas Components". *J. Mater. Chem. A* (2025), pp. 16879–16892.

[506] Y. A. Ran, J. Tapia, S. Sharma, P. Bai, S. Calero, T. J. H. Vlugt, and D. Dubbeldam. "Implementation of energy and force grids in molecular simulations of porous materials". *Mol. Phys.* (2025), e2542486.

[507] Z. Li, L. Constantinou, R. Baur, D. Dubbeldam, S. Calero, S. Sharma, M. Rigutto, P. Dey, and T. J. H. Vlugt. "Review of Group Contribution Methods for Prediction of Thermodynamic Properties of Long-chain Hydrocarbons". *Mol. Phys.* (2025). Manuscript accepted.

A

Appendix

A.1. Data availability for each chapter

Sample simulation input and output files for each chapter in this thesis (except for chapter 1 and chapter 8) are available in the 4TU data repository to ensure the clarity and reproducibility of the data. The corresponding DOIs for each chapter are provided below:

Chapter 2:
https://doi.org/10.4121/9f0aec2f-8905-42ae-945d-32fbf4262969.v1
Chapter 3:
https://doi.org/10.4121/4943dc0f-d908-4215-89ce-b9a6163930d2.v1
Chapter 4:
https://doi.org/10.4121/33ad5d25-14b5-4490-b308-fc09097e89f7.v1
Chapter 5:
https://doi.org/10.4121/f905ae7e-2950-4a9b-91b4-b592e4c788e9.v1
Chapter 6:
https://doi.org/10.4121/e855547e-384c-4802-9faa-00981a99419b.v1
Chapter 7:
https://doi.org/10.4121/7fdc96f0-69ff-4c1a-be2b-8aa6a0a812cb.v1

A

A.2. Appendix for Chapter 4

The following items are presented in this Appendix:

1. Force field parameters for alkanes (section: A.2.1)
2. Sensitivity analysis for the reaction equilibrium distribution for hydroisomerization of C_7 in MFI-type zeolite (section: A.2.2)
3. Henry coefficients of C_8 isomers in different zeolites (section: A.2.3)
4. Selectivities of C_8 isomers relative to n-C_8 in the gas phase and different zeolites (section:A.2.4)

A.2.1. Force field parameters for alkanes

Table A.1 lists the Lennard-Jones parameters for the united atoms present in alkanes. These parameters are used to compute intramolecular and intermolecular interactions of alkane chains and are taken from Dubbeldam et al [136]. To account for the interactions between the zeolite atoms and alkane molecules, the TraPPE-zeo force field [43] is used. The Lennard-Jones parameters for O and Si are listed in Table A.2. For interactions between different types of atoms, the Lennard-Jones parameters are calculated using the Lorentz-Berthelot mixing rules [501]. The TraPPE united atom force field is used for the intramolecular bonded interactions present in alkanes. These interactions are divided into bond-stretching, bond-bending, and torsion potentials. TraPPE force field uses a fixed C-C bond length which is equal to 1.54 Å [43]. The bond-bending potential is described by a harmonic potential [502]:

$$U_{\text{bend}} = \frac{1}{2} k_{\text{bend}} (\theta - \theta_0)^2 \tag{A.1}$$

In Eq. A.1, k_{bend} is the force constant for the bending potential and θ is the bending angle. The values for the reference angles θ_0 are listed in Table A.3 [37]. These values depend on the type of united atom present at the center of the bend. The torsion potential is described by the three-cosine dihedral equation [502].

$$U_{\text{torsion}} = c_0 + c_1 [1 + \cos(\phi)] + c_2 [1 - \cos(2\phi)] + c_3 [1 + \cos(3\phi)] \tag{A.2}$$

The coefficients (c_0, c_1, c_2, c_3) in Eq. A.2 for different types of torsions are listed in Table A.4 [37]. Table A.4 lists the values of the coefficients for different torsions described by the three-cosine dihedral equation. The TraPPE force field [37] parameters for the torsion type CH_x-C-C-CH_y are not available. In this study, this torsion type is obtained from Ref. [454].

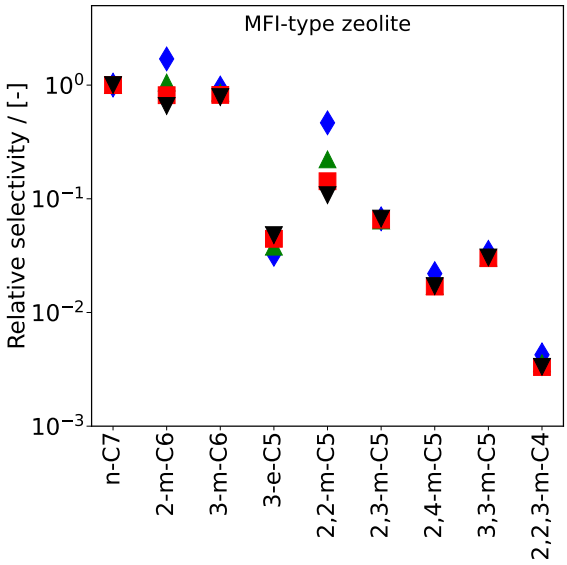


Figure A.1. Selectivities of C_7 isomers relative to n-C_7 at reaction equilibrium in the gas phase at infinite dilution. The reaction equilibrium distribution is computed at ♦ 400 K, ▲ 500 K, ■ 600 K, and ▼ 700 K. The raw data is listed in the worksheet (lnq_iC7) of the Supporting Information (SI2.xlsx file) of Ref. [390].

A

Table A.1. Lennard-Jones force field parameters for united atoms to account for intramolecular and intermolecular non-bonded interactions in alkanes. These parameters are taken from Dubbeldam et al [136].

united atom	ϵ/k_B/ [K]	σ/ [Å]
CH_3	108.0	3.76
CH_2	56.0	3.96
CH	17.0	4.67
C	0.8	6.38

Table A.2. Lennard-Jones force field parameters for zeolite atoms to account for intermolecular non-bonded interactions between zeolite atoms and alkanes. These parameters are taken from Bai et al [43]. For interactions between different types of atoms, the Lennard-Jones parameters are determined using the Lorentz–Berthelot mixing rules [41, 42].

atom	ϵ/k_B/ [K]	σ/ [Å]
O	53.0	3.30
Si	22.0	2.30

Table A.3. Values of the reference angles θ_0 and the force constants k_{bend} for different bend types.

bend type	θ_0/ [deg]	k_{bend}/k_B/ [K/rad^2]
CH_x—CH_2—CH_y	114.0	62500
CH_x—CH—CH_y	112.0	62500
CH_x—C—CH_y	109.47	62500

Table A.4. Values of the coefficients in the TraPPE united atom torsion potential [37, 454] for different torsion types.

torsion type	c_0/k_B/ [K]	c_1/k_B/ [K]	c_2/k_B/ [K]	c_3/k_B/ [K]
CH_x-CH_2-CH_2-CH_y	0	355.03	-68.19	791.32
CH_x-CH_2-CH-CH_y	-251.06	428.73	-111.85	441.27
CH_x-CH_2-C-CH_y	0	0	0	461.29
CH_x-CH-CH-CH_y	-251.06	428.73	-111.85	441.27
CH_x-CH-C-CH_y	0	0	0	1635.7
CH_x-C-C-CH_y	0	0	0	1635.7

A.2.2. Sensitivity analysis for the reaction equilibrium distribution for hydroisomerization of C_7 in MFI-type zeolite

Fig. A.1 shows the relative selectivities of C_7 isomers in MFI-type zeolite at chemical reaction equilibrium. As discussed in Chapter 4, selectivities

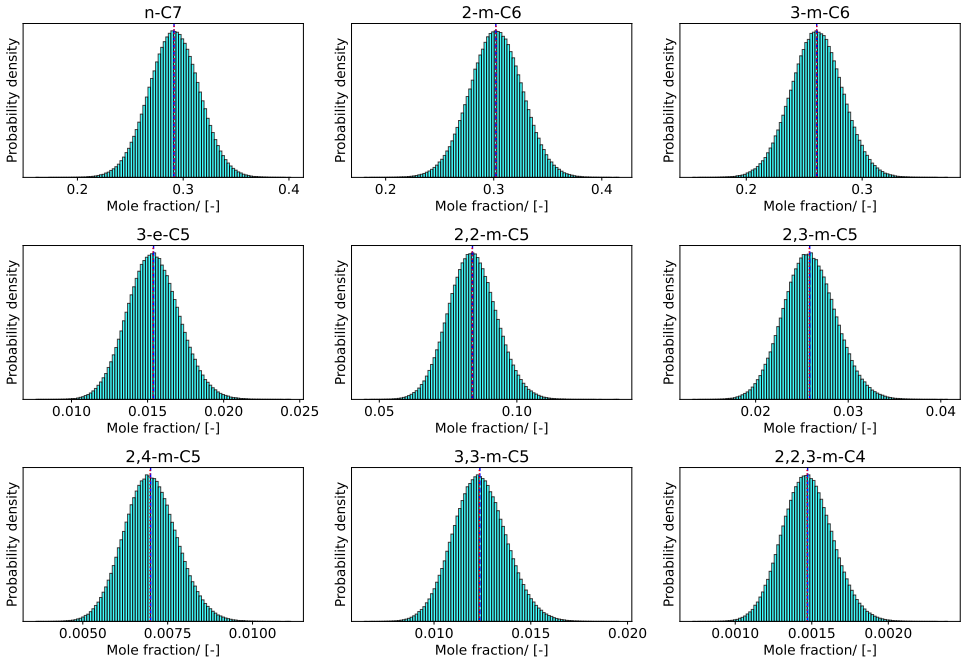


Figure A.2. Probability distributions of variations in adsorbed-phase mole fractions for C_7 isomers at chemical reaction equilibrium in MFI-type zeolite at 500 K, accounting for uncertainties in the standard enthalpy of formation ($\Delta_f H_0$), Gibbs free energy difference ($G_0(T) - H_0(0\ \mathrm{K})$), and Henry coefficient (k_H).

for Hydroisomerization of n-alkanes at reaction equilibrium are derived from Henry coefficients, Gibbs free energies, and standard enthalpies of formation (at 0 K) in the gas phase. It is important to recognize that these thermodynamic and adsorption properties inherently carry uncertainties, regardless of whether they are obtained from experiments, quantum mechanical calculations, or molecular simulations. For example, the Gibbs free energy is generally difficult to determine with an accuracy better than the chemical accuracy of 1 kcal/mol [424]. Similarly, Henry coefficients exhibit uncertainty both in experimental measurements [503] and in molecular simulations [17]. In the latter, two primary sources of uncertainty are present: (1) statistical noise due to the finite length of the simulation, and (2) model limitations, such as approximations in the force field, assumptions of a rigid zeolite framework [44], and uncertainties in force field parameters. To evaluate the impact of these uncertainties on the predicted equilibrium behavior, we performed a sensitivity analysis. Specifically, we examined how uncertainties in the Gibbs free energy, enthalpy of formation, and Henry

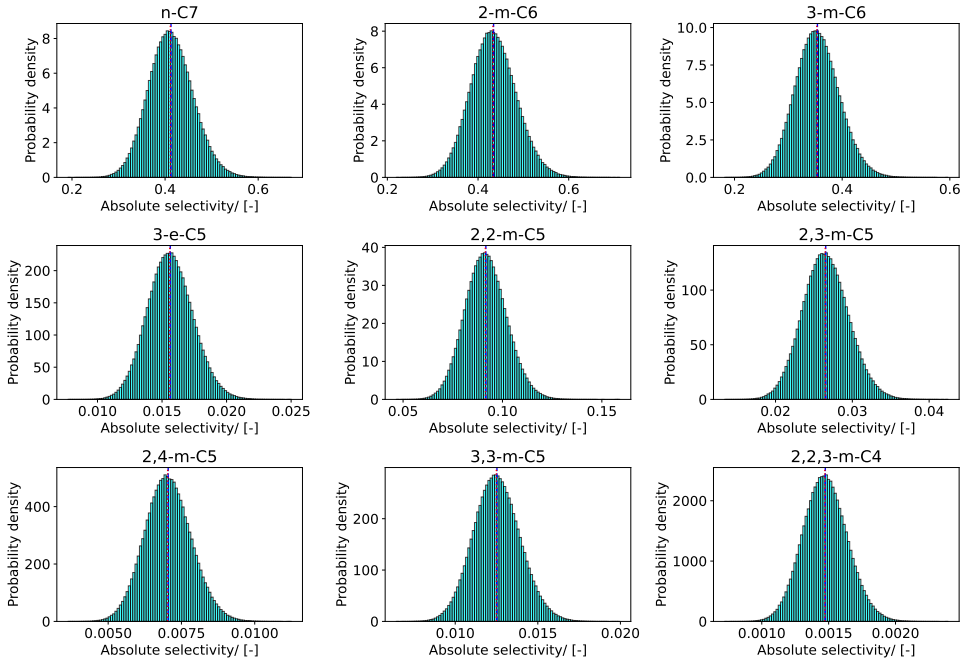


Figure A.3. Probability distributions of variations in absolute selectivities for C_7 isomers at chemical reaction equilibrium in MFI-type zeolite at 500 K, accounting for uncertainties in the standard enthalpy of formation ($\Delta_f H_0$), Gibbs free energy difference ($G_0(T) - H_0$ (0 K)), and Henry coefficient (k_H).

coefficients affect the computed adsorbed-phase loadings and absolute selectivities. This analysis was conducted using Monte Carlo simulations that incorporate probabilistic variations in the input parameters. In this approach, uncertainties in $\Delta_f H_0$ and $G_0(T) - H_0(0\ \mathrm{K})$ are modeled as Gaussian distributions with standard deviations of 1 kcal/mol, consistent with chemical accuracy. Uncertainties in Henry coefficients, k_H, are modeled with standard deviations equal to 10% of the predicted values, an assumption justified by the demonstrated accuracy of the force fields used for simulating zeolite–hydrocarbon interactions [43, 136]. For each of the 10^6 Monte Carlo steps, a new set of perturbed input values is sampled, and the corresponding equilibrium composition is computed using Eqs. 4.10–4.24. The resulting probability distributions of mole fractions (Fig. A.2) and absolute selectivities (Fig. A.3) exhibit approximately Gaussian shapes centered around the mean values. These mean values (blue dashed lines) are closely aligned with the unperturbed predictions (actual values) (red dotted lines), indicating that the response of the system to input uncertainty is limited. Fig. A.4 shows the absolute selectivities of C_7 isomers at chemical reaction equilibrium

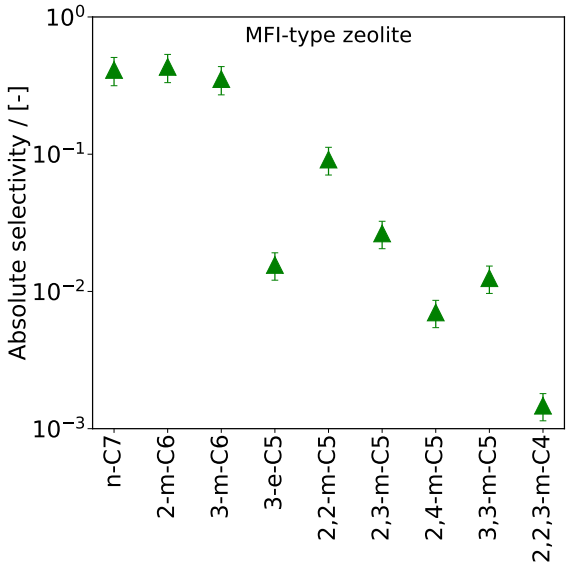


Figure A.4. Absolute selectivities of C_7 isomers at reaction equilibrium in the gas phase at infinite dilution. The reaction equilibrium distribution is computed at ▲ 500 K. The raw data is listed in the worksheet (lnq_iC7) of the Supporting Information (SI2.xlsx file) of Ref. [390]. The error bars represent two times the standard deviations obtained from the sensitivity analysis.

in MFI-type zeolite at 500 K. The error bars in the figure represent twice the standard deviations for each isomer, as determined from the sensitivity analysis. The small error bars indicate that the computed selectivities and loadings are robust with respect to the expected uncertainties in gas phase thermochemistry and Henry coefficients, reinforcing the reliability of the model predictions at realistic conditions.

A.2.3. Henry coefficients of C_8 isomers in different zeolites

Figs. A.5-A.6 show the variations in Henry coefficients for C_8 isomers in FAU- and ITQ-29-type zeolites. FAU- and ITQ-29-type zeolites favor multi-branched isomers over mono-branched isomers. These variations are less significant in FAU-type zeolite because of the larger channel diameters (7.4 Å) [18] which connects the cage-type pores. Figs. A.7-A.9 show the variations in the Henry coefficients for different C_8 isomers in BEA-, MEL-, and MFI-type zeolites. BEA- and MEL-type zeolites show similar trends in Henry coefficients for mono- and di-branched isomers, which decrease with increasing degree of branching in both zeolites. For the mono-branched isomers, 3-e-C_6 has the least affinity for adsorption in both BEA- and

A

MEL-type zeolites. 2,2-m-C_6, 3,3-m-C_6, and 3-e-3-m-C_5 have the smallest preference for adsorption compared to the other di-branched isomers. The tri-branched isomers have a lower affinity for adsorption in MEL-type zeolites compared to the di-branched isomers. This is because of the smaller pore diameters in MEL-type zeolite (5.3 × 5.4 Å) [18]. In MFI-type zeolite, 2,2-m-C_6 and 3,3-m-C_6 have larger Henry coefficients compared to other di-branched isomers. This is because the methyl branches can optimally fit into the intersections of MFI-type zeolite. As the separation distance between the methyl groups increases (2,4-m-C_6 versus 2,5-m-C_6), the likelihood for either branch attempting to fit into the channel-like pores increases, which is not favored in MFI-type zeolite. The tri-branched and the tetra-branched isomers are the least favored isomers in MFI-type zeolite due to the smaller pore diameters (sinusoidal channels: 5.1 × 5.5 Å and straight channels: 5.3 × 5.6 Å) [18] which leads to steric hindrance in presence of multiple branches in close proximity. Fig. A.10 shows the Henry coefficients for C_8 isomers in MTW-type zeolite. Similar to C_7 isomers, an overall decrease in the Henry coefficients with increasing degree of branching is observed for C_8 isomers. The presence of two methyl groups or a combination of a methyl and an ethyl group does not favor the adsorption of 2,2-m-C_6, 3,3-m-C_6, 3-e-2-m-C_5, and 3-e-3-m-C_5 in the one-dimensional pores of MTW-type zeolite. MRE-type zeolite also shows a decrease in Henry coefficients as the degree of branching increases (Fig. A.11). There is a sharp decrease in the Henry coefficients compared to those in MTW-type zeolite. This is because MRE-type zeolite has comparatively smaller channel diameters (5.6 × 5.6 Å) than MTW-type zeolite (6.0 × 5.6 Å) [18]. The presence of an ethyl group does not favor the adsorption of 3-e-2-m-C_5 and 3-e-3-m-C_5 isomers in the MRE-type zeolite. Two branches attached to the adjacent carbon atoms in the main chain are not preferred in MRE-type zeolites. This can be attributed to the corrugations present inside the channels. Typical snapshots of MRE-type zeolite are shown in Fig. 4.7 of section 4.4. These corrugations are periodic combinations of peaks and crests which lead to variations in diameter inside the channels. The effects of corrugations can be quantified using the pore bumpiness factor [504]. This factor is defined as the difference between the maximum diameter of a sphere that can be included inside the channel and the maximum diameter of a sphere that can freely diffuse inside the channel [504]. The bumpiness factor for MRE-type zeolite is 0.77 Å. When two branches are adjacent to each other in the isomers, one of these has a higher probability of being fitted in the region of a smaller channel diameter. This explains why, in addition to 3-e-2-m-C_5, the molecules 2,3-m-C_6 and 3,4-m-C_6 also have comparatively lower Henry coefficients.

A.2.4. Selectivities of C_8 isomers relative to linear C_8 in the gas phase and different zeolites

Fig. A.12 shows the selectivities of C_8 isomers relative to n-C_8 at reaction equilibrium in the gas phase. n-C_8 and the mono-branched isomers exhibit little variations in the magnitudes of selectivities. 3-e-2-m-C_5 and 3-e-3-m-C_5 isomers have the smallest selectivities compared to other di-branched isomers. 2,2,3,3-m-C_4 is the least preferred isomer compared to all C_8 isomers. These observed trends in the selectivities at reaction equilibrium in the gas phase solely rely on the ratios of the isolated molecule partition functions of these isomers in the ideal gas phase.

Figs. A.13-A.19 show the selectivities of C_8 isomers relative to n-C_8 at reaction equilibrium in FAU-, ITQ-29-, BEA-, MEL-, MFI-, MTW-, and MRE-type zeolites. The variations in the selectivities are influenced by the combined effects of the isolated molecule partition functions in the gas phase and the Henry coefficients of these isomers in the zeolites. The influence of Henry coefficients on the reaction equilibrium distribution in the FAU-type zeolite is not very noticeable (Fig. A.13). This is due to the presence of larger pore diameters (7.4 Å) [18] in FAU-type zeolite. Multi-branched isomers preferentially adsorb in ITQ-29-type zeolite (Fig. A.6). However, the variations in Henry coefficients are not very large in this zeolite. Therefore, the adsorbed phase distribution (Fig. A.14) at reaction equilibrium closely resembles the gas phase distribution (Fig. A.12). In BEA-, MEL-, and MFI-type zeolites, the mono-branched isomers are preferred at reaction equilibrium distribution. For zeolites with narrow channel-like pores such as MRE- and MTW-type zeolites, the variations in Henry coefficients are very important in deciding which isomers are preferentially formed at reaction equilibrium. MTW- and MRE-type zeolites prefer n-C_8 the most. In all zeolites considered in this study, 2,2,3,3-m-C_4 is the least favored isomer at reaction equilibrium.

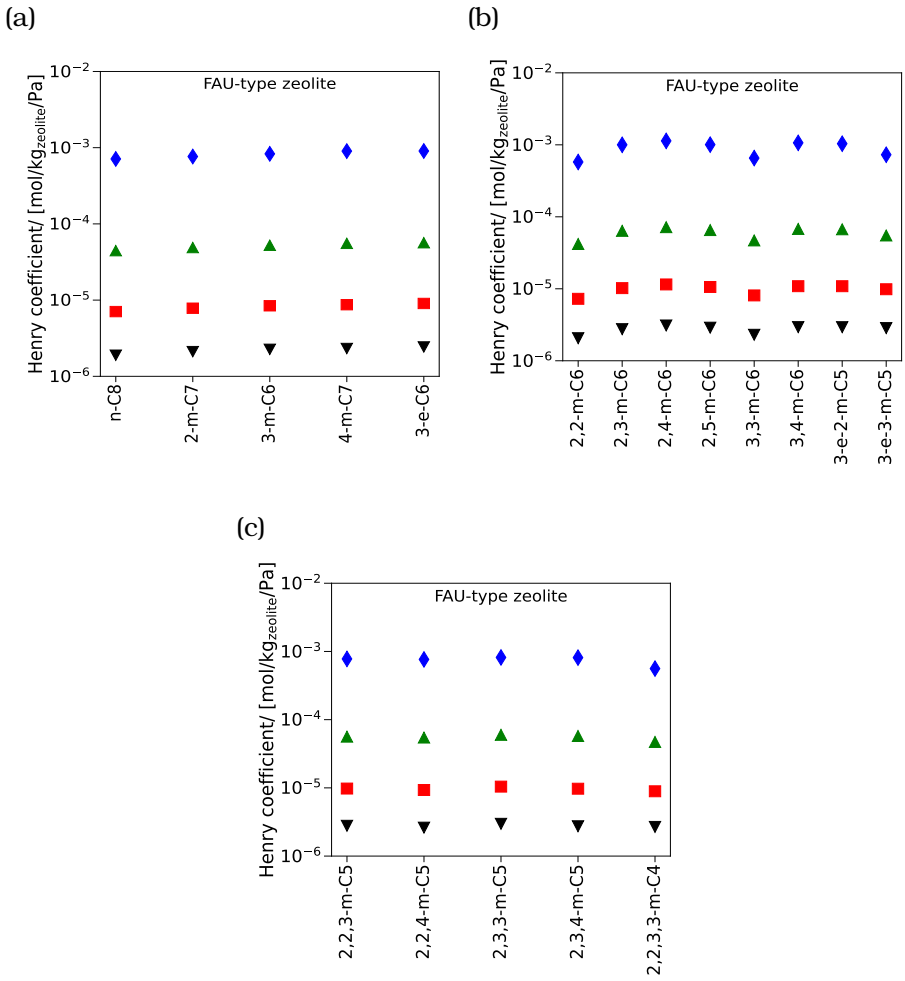


Figure A.5. Henry coefficients of C_8 isomers in FAU-type zeolite at ♦ 400 K, ▲ 500 K, ■ 600 K, and ▼ 700 K. The raw data is listed in the worksheets (xi_iC8_400K, xi_iC8_500K, xi_iC8_600K, and xi_iC8_700K) of the Supporting Information (SI2.xlsx file) of Ref. [390].

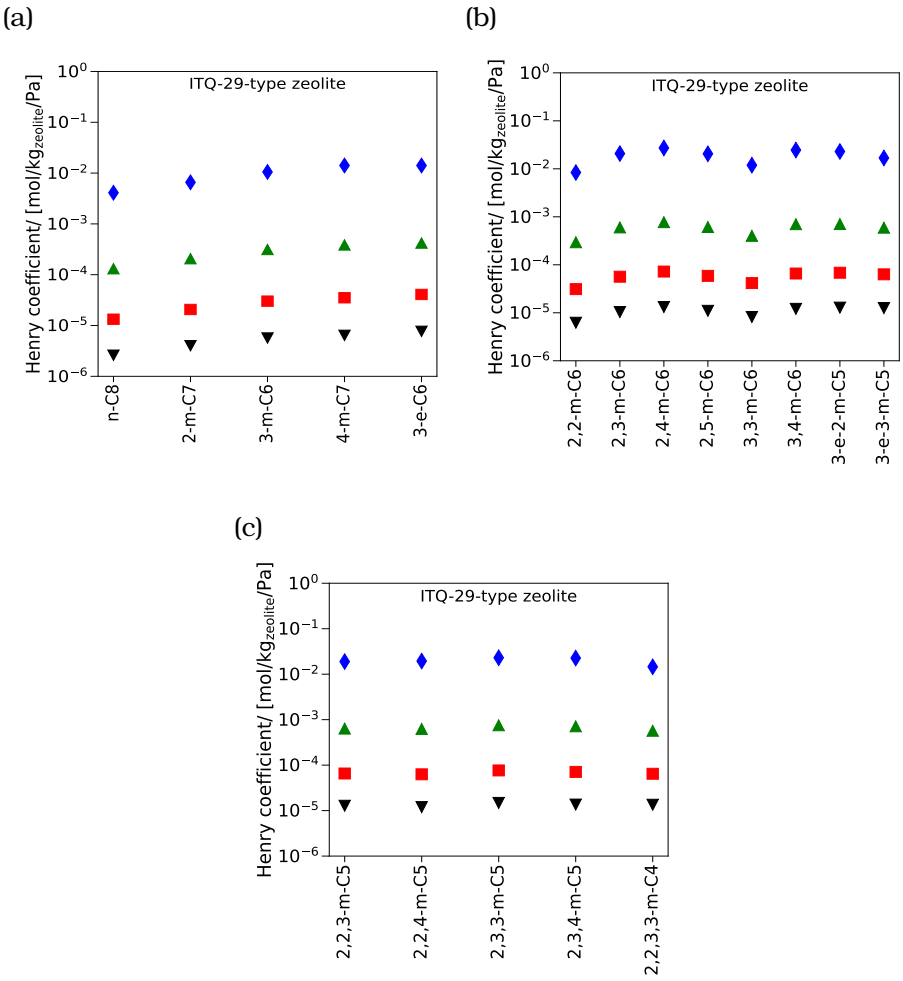


Figure A.6. Henry coefficients of C_8 isomers in ITQ-29-type zeolite at ♦ 400 K, ▲ 500 K, ■ 600 K, and ▼ 700 K. The raw data is listed in the worksheets (xi_iC8_400K, xi_iC8_500K, xi_iC8_600K, and xi_iC8_700K) of the Supporting Information (SI2.xlsx file) of Ref. [390].

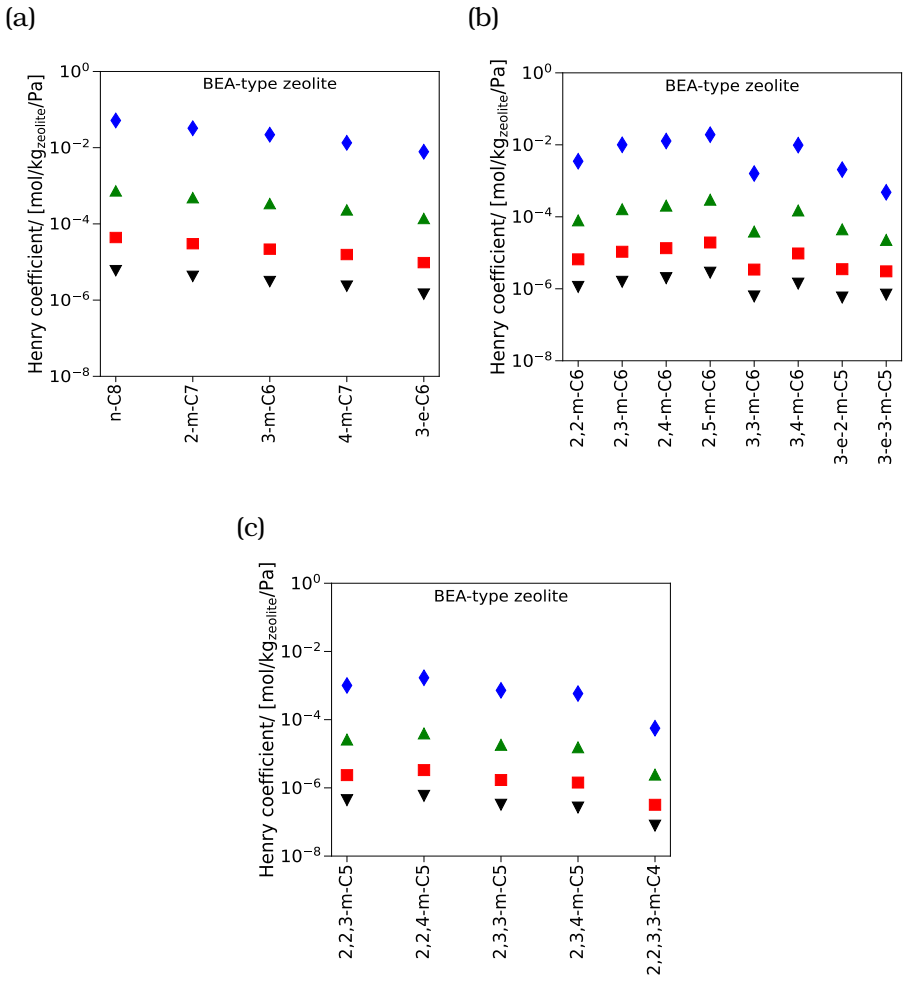


Figure A.7. Henry coefficients of C_8 isomers in BEA-type zeolite at ♦ 400 K, ▲ 500 K, ■ 600 K, and ▼ 700 K. The raw data is listed in the worksheets (xi_iC8_400K, xi_iC8_500K, xi_iC8_600K, and xi_iC8_700K) of the Supporting Information (SI2.xlsx file) of Ref. [390].

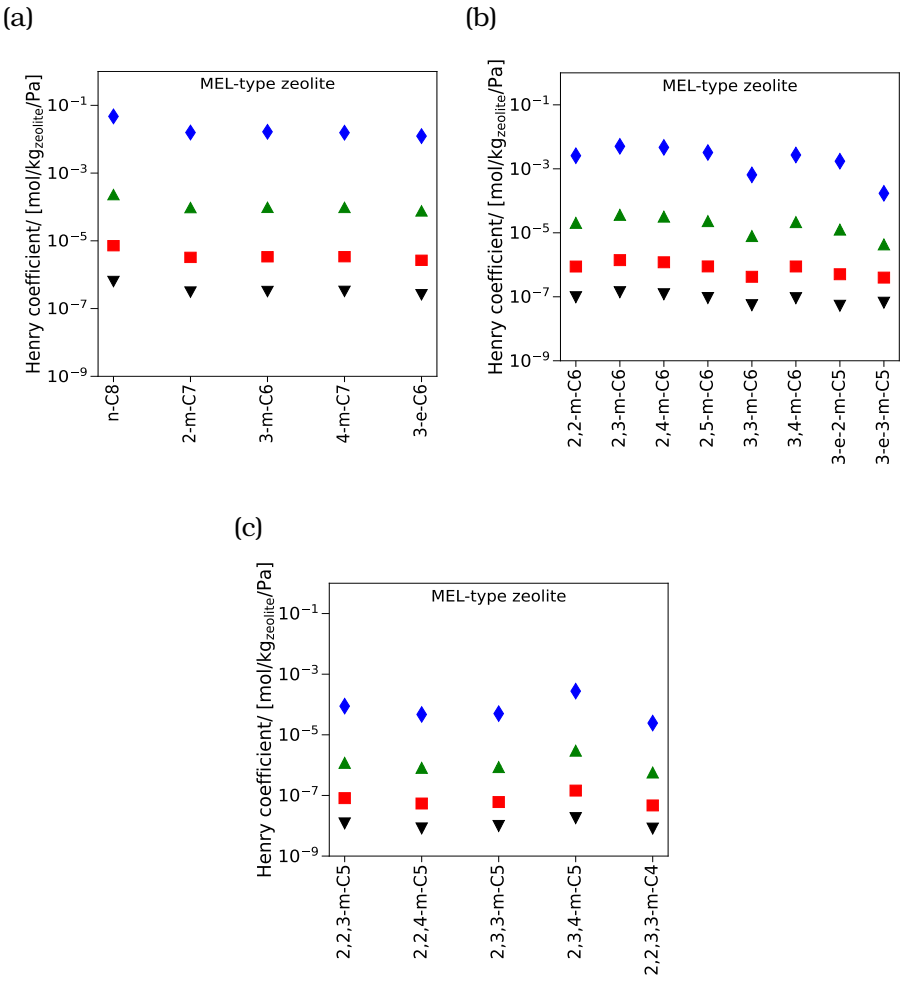


Figure A.8. Henry coefficients of C_8 isomers in MEL-type zeolite at ♦ 400 K, ▲ 500 K, ■ 600 K, and ▼ 700 K. The raw data is listed in the worksheets (xi_iC8_400K, xi_iC8_500K, xi_iC8_600K, and xi_iC8_700K) of the Supporting Information (SI2.xlsx file) of Ref. [390].

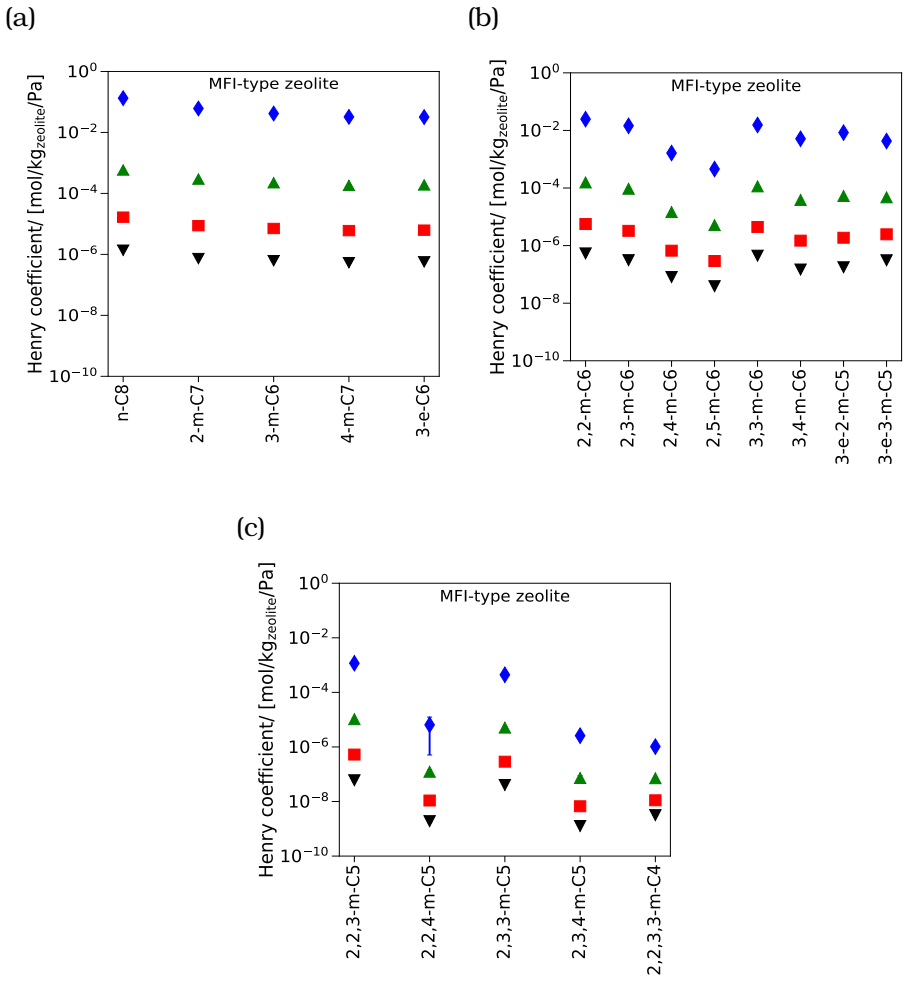


Figure A.9. Henry coefficients of C_8 isomers in MFI-type zeolite at ♦ 400 K, ▲ 500 K, ■ 600 K, and ▼ 700 K. The raw data is listed in the worksheets (xi_iC8_400K, xi_iC8_500K, xi_iC8_600K, and xi_iC8_700K) of the Supporting Information (SI2.xlsx file) of Ref. [390].

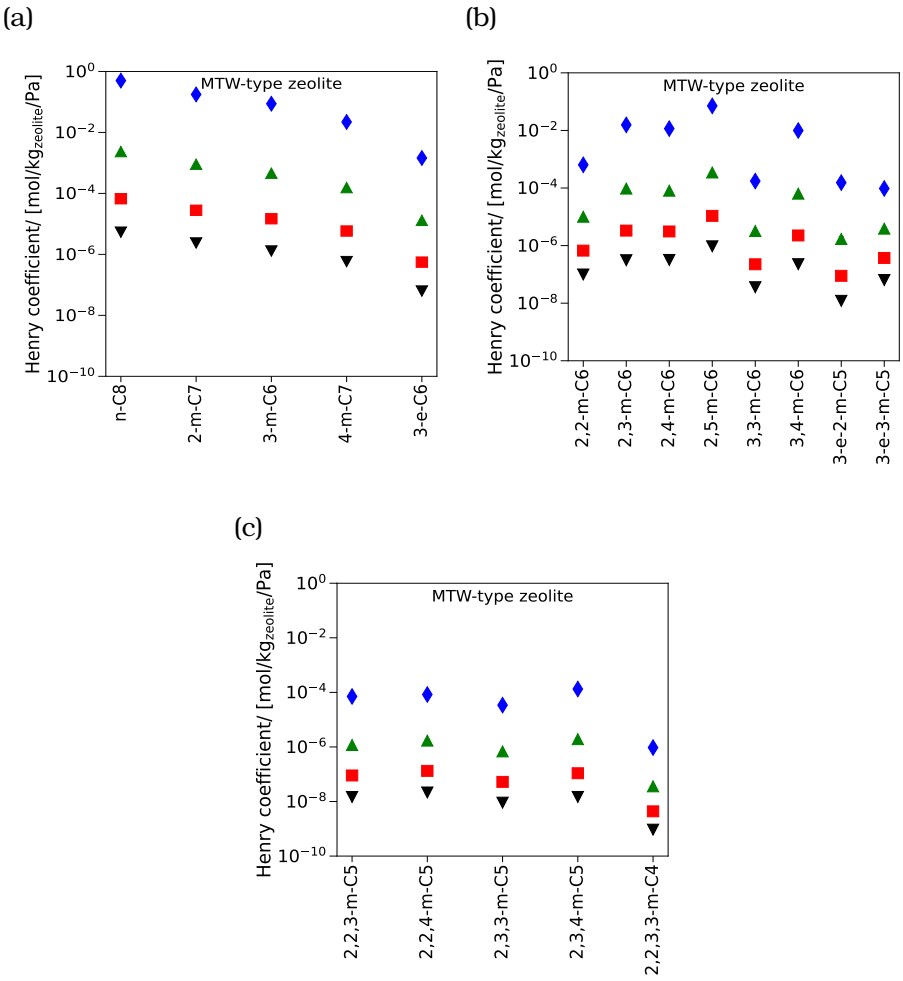


Figure A.10. Henry coefficients of C_8 isomers in MTW-type zeolite at ♦ 400 K, ▲ 500 K, ■ 600 K, and ▼ 700 K. The raw data is listed in the worksheets (xi_iC8_400K, xi_iC8_500K, xi_iC8_600K, and xi_iC8_700K) of the Supporting Information (SI2.xlsx file) of Ref. [390].

A

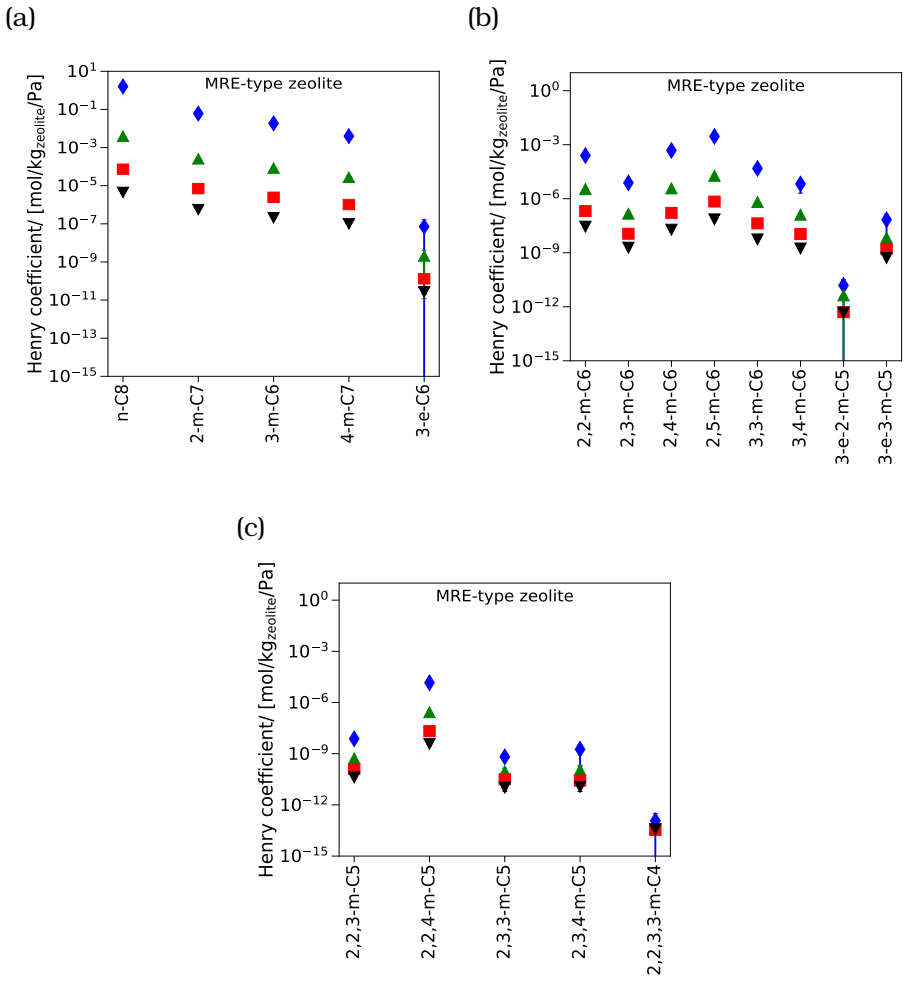


Figure A.11. Henry coefficients of C_8 isomers in MRE-type zeolite at ♦ 400 K, ▲ 500 K, ■ 600 K, and ▼ 700 K. The raw data is listed in the worksheets (xi_iC8_400K, xi_iC8_500K, xi_iC8_600K, and xi_iC8_700K) of the Supporting Information (SI2.xlsx file) of Ref. [390].

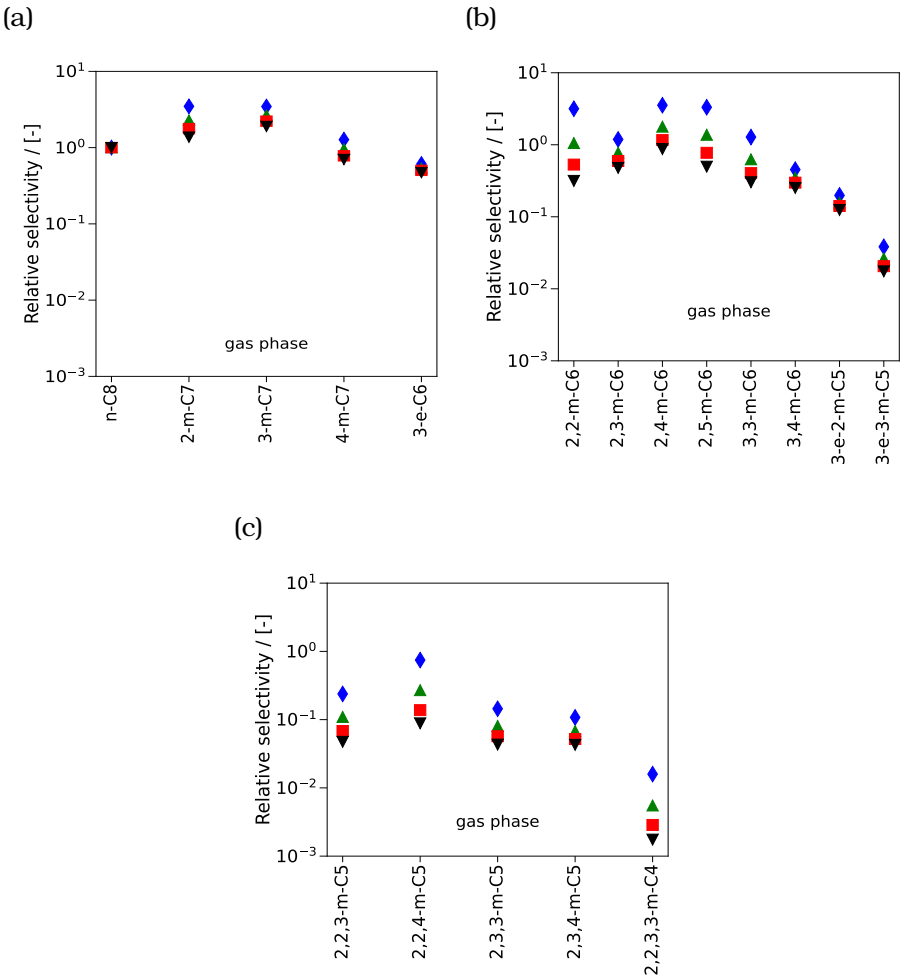


Figure A.12. Selectivities of C_8 isomers relative to n-C_8 at reaction equilibrium in the gas phase at infinite dilution. The reaction equilibrium distribution is computed at ♦ 400 K, ▲ 500 K, ■ 600 K, and ▼ 700 K. The raw data is listed in the worksheet (lnq_iC8) of the Supporting Information (SI2.xlsx file) of Ref. [390].

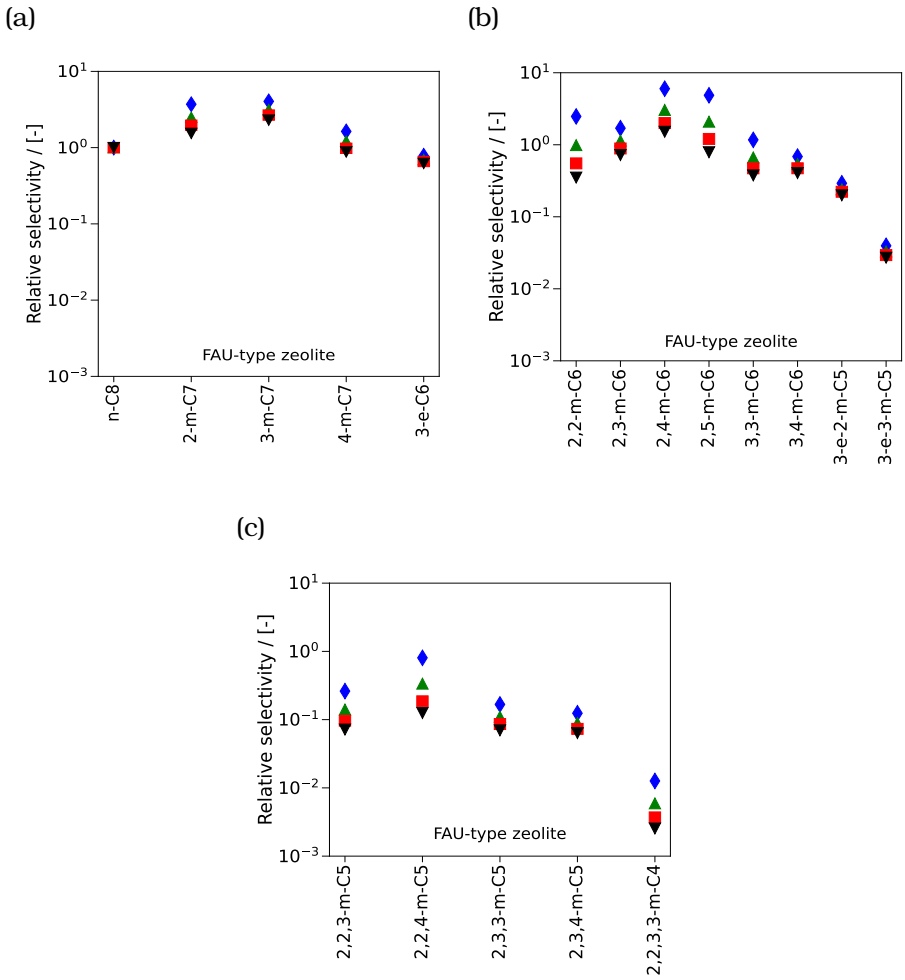


Figure A.13. Selectivities of C_8 isomers relative to n-C_8 at reaction equilibrium in FAU-type zeolite at infinite dilution. The reaction equilibrium distribution is computed at ♦ 400 K, ▲ 500 K, ■ 600 K, and ▼ 700 K. The raw data is listed in the worksheets (xi_iC8_400K, xi_iC8_500K, xi_iC8_600K, and xi_iC8_700K) of the Supporting Information (SI2.xlsx file) of Ref. [390].

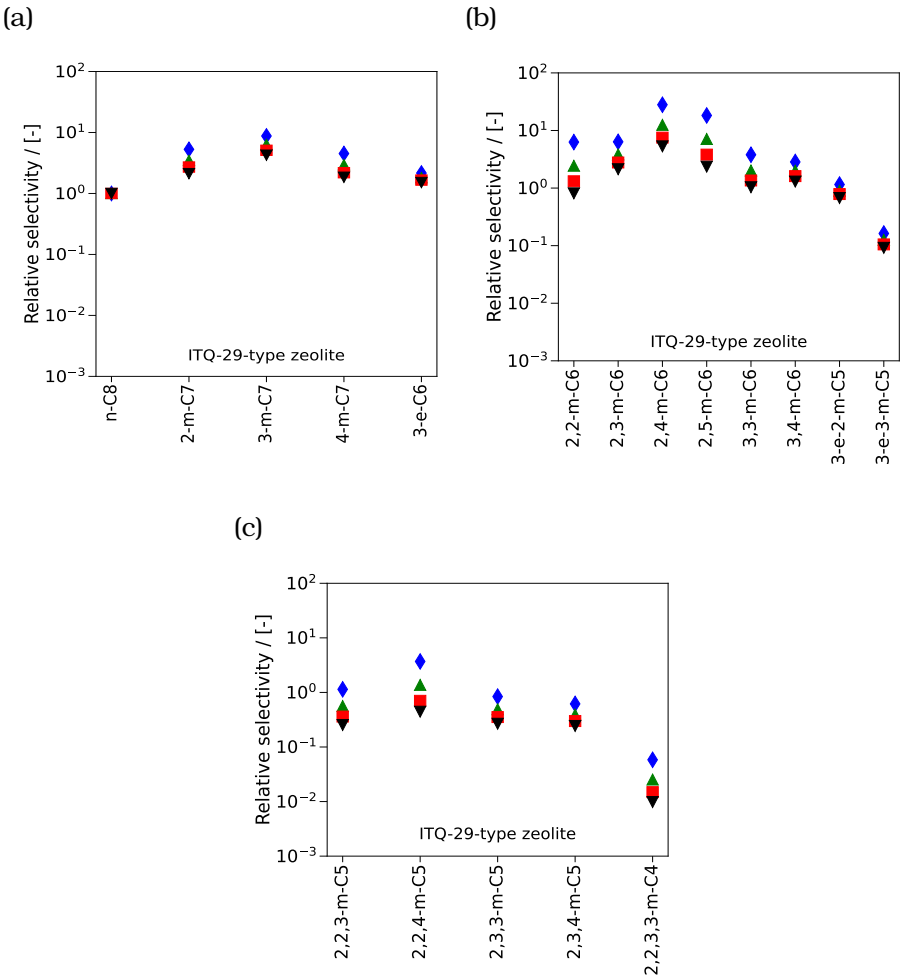


Figure A.14. Selectivities of C_8 isomers relative to n-C_8 at reaction equilibrium in ITQ-29-type zeolite at infinite dilution. The reaction equilibrium distribution is computed at ♦ 400 K, ▲ 500 K, ■ 600 K, and ▼ 700 K. The raw data is listed in the worksheets (xi_iC8_400K, xi_iC8_500K, xi_iC8_600K, and xi_iC8_700K) of the Supporting Information (SI2.xlsx file) of Ref. [390].

A

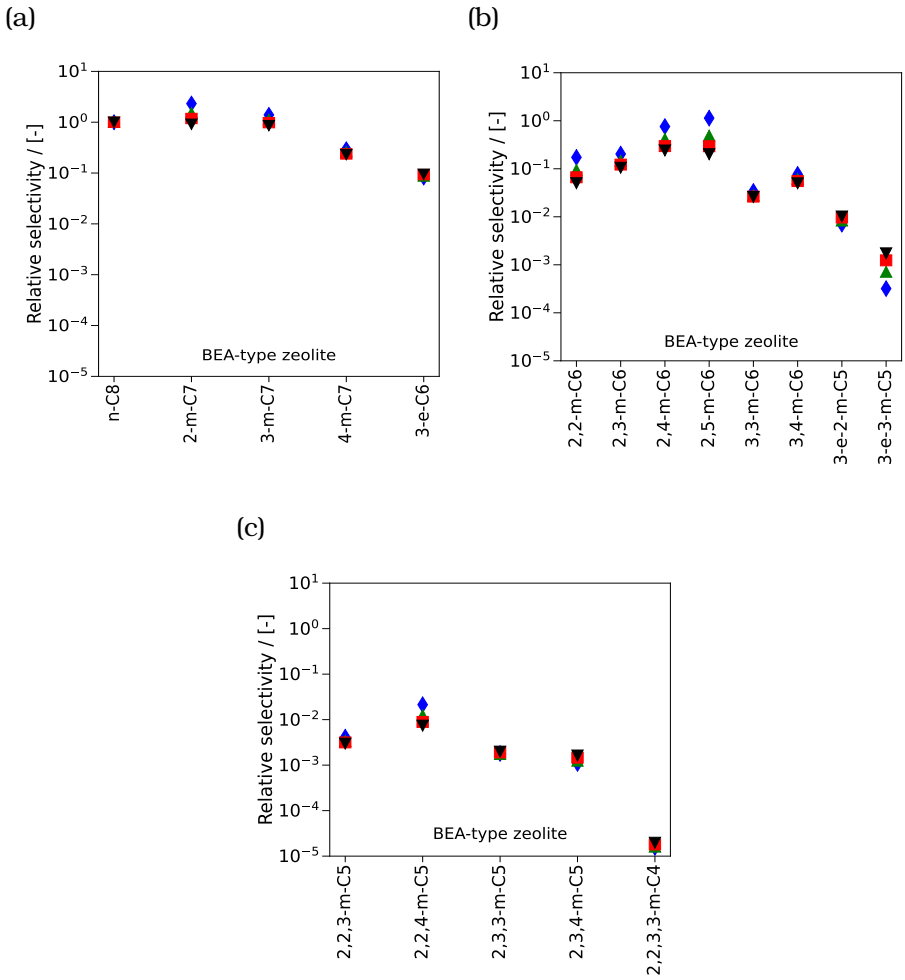


Figure A.15. Selectivities of C_8 isomers relative to n-C_8 at reaction equilibrium in BEA-type zeolite at infinite dilution. The reaction equilibrium distribution is computed at ♦ 400 K, ▲ 500 K, ■ 600 K, and ▼ 700 K. The raw data is listed in the worksheets (xi_iC8_400K, xi_iC8_500K, xi_iC8_600K, and xi_iC8_700K) of the Supporting Information (SI2.xlsx file) of Ref. [390].

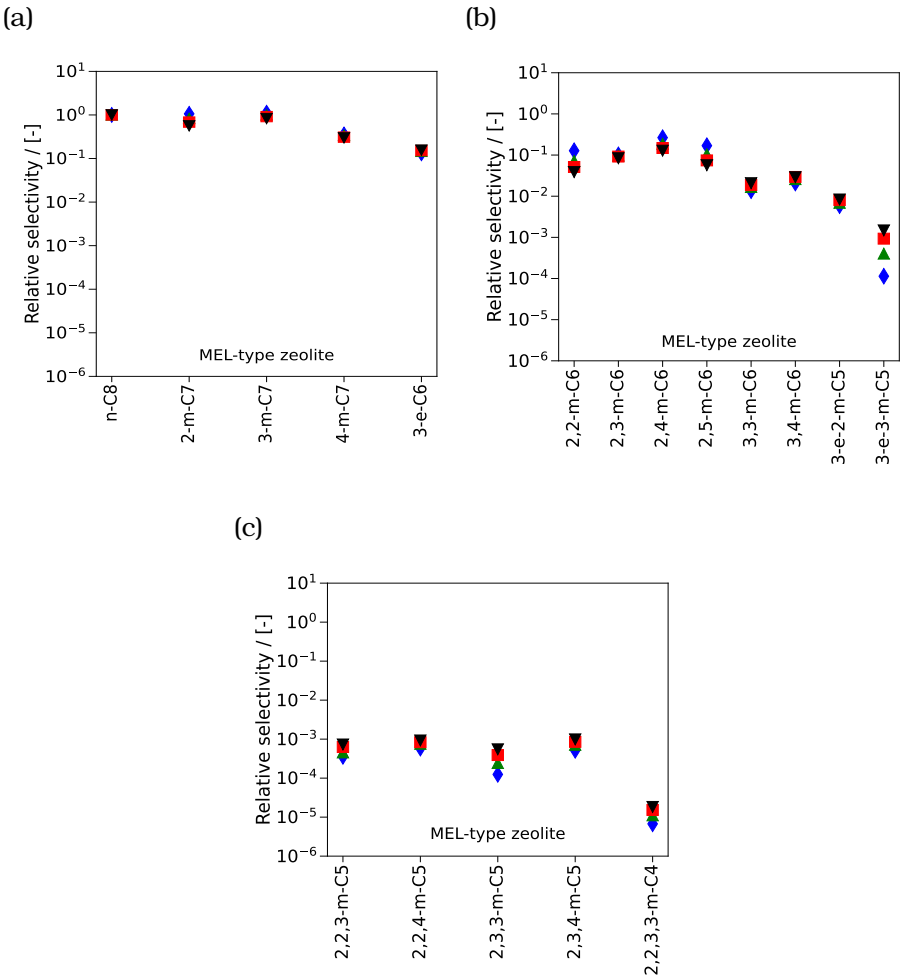


Figure A.16. Selectivities of C_8 isomers relative to n-C_8 at reaction equilibrium in MEL-type zeolite at infinite dilution. The reaction equilibrium distribution is computed at ♦ 400 K, ▲ 500 K, ■ 600 K, and ▼ 700 K. The raw data is listed in the worksheets (xi_iC8_400K, xi_iC8_500K, xi_iC8_600K, and xi_iC8_700K) of the Supporting Information (SI2.xlsx file) of Ref. [390].

A

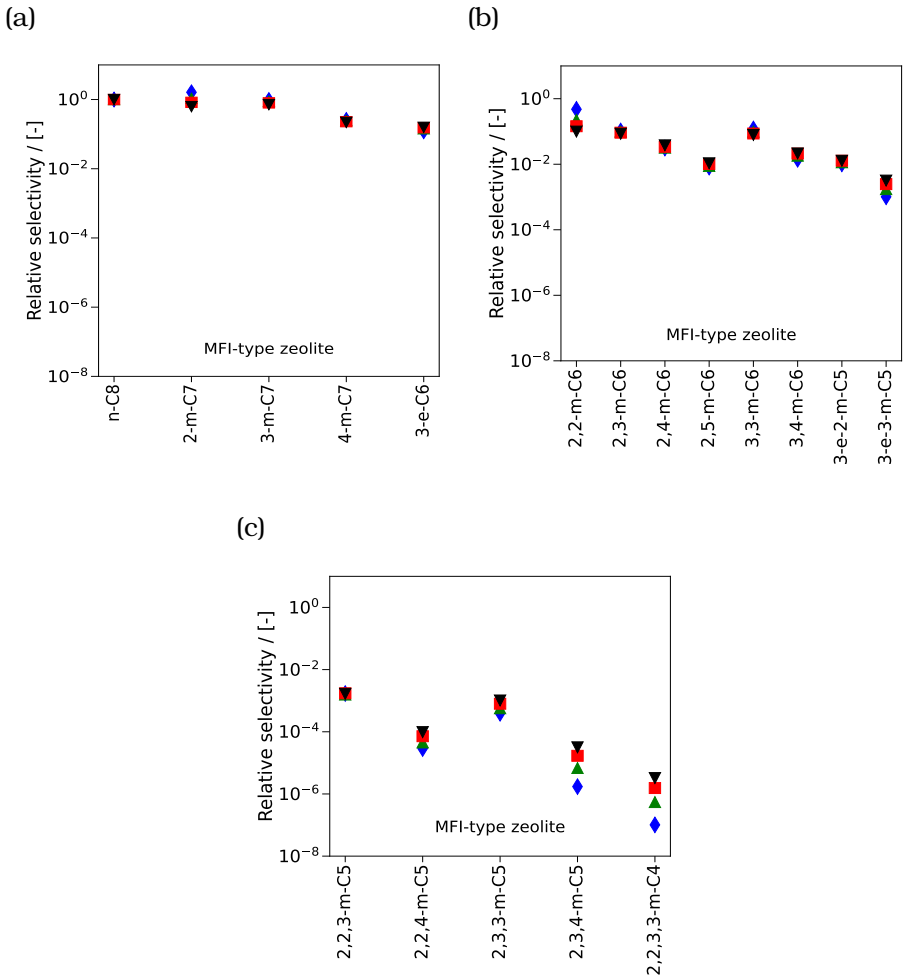


Figure A.17. Selectivities of C_8 isomers relative to n-C_8 at reaction equilibrium in MFI-type zeolite at infinite dilution. The reaction equilibrium distribution is computed at ♦ 400 K, ▲ 500 K, ■ 600 K, and ▼ 700 K. The raw data is listed in the worksheets (xi_iC8_400K, xi_iC8_500K, xi_iC8_600K, and xi_iC8_700K) of the Supporting Information (SI2.xlsx file) of Ref. [390].

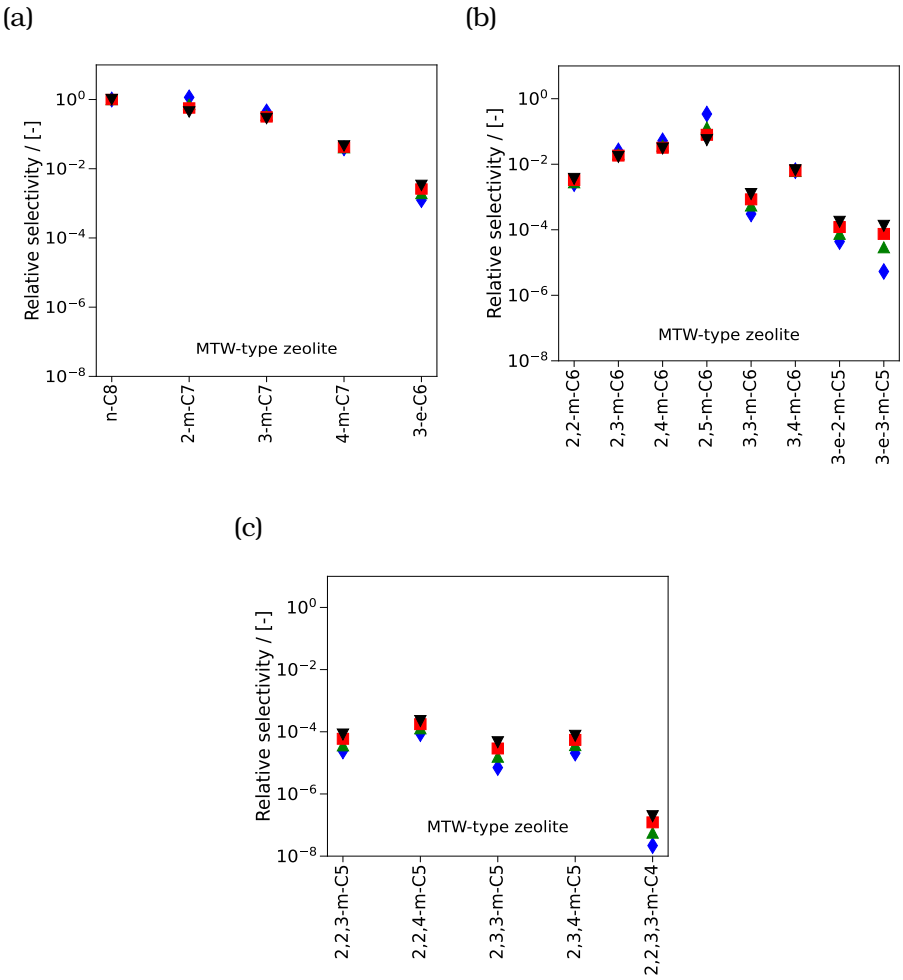


Figure A.18. Selectivities of C_8 isomers relative to n-C_8 at reaction equilibrium in MTW-type zeolite at infinite dilution. The reaction equilibrium distribution is computed at ♦ 400 K, ▲ 500 K, ■ 600 K, and ▼ 700 K. The raw data is listed in the worksheets (xi_iC8_400K, xi_iC8_500K, xi_iC8_600K, and xi_iC8_700K) of the Supporting Information (SI2.xlsx file) of Ref. [390].

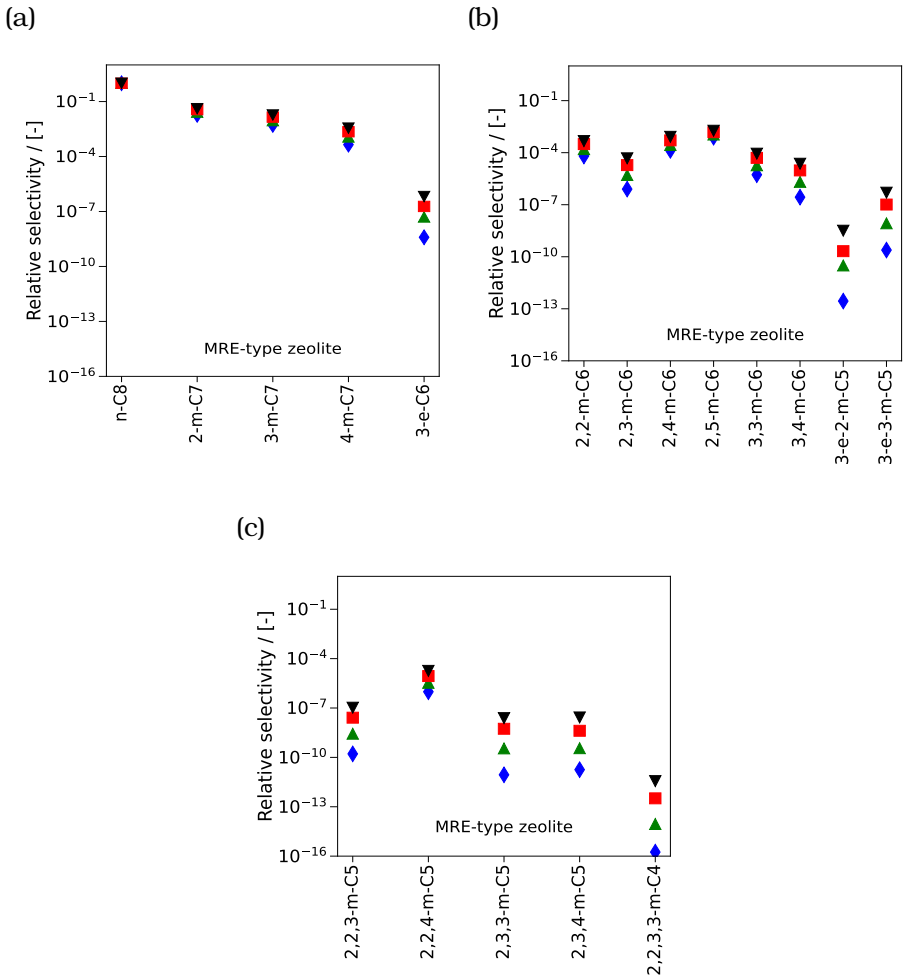


Figure A.19. Selectivities of C_8 isomers relative to n-C_8 at reaction equilibrium in MRE-type zeolite at infinite dilution. The reaction equilibrium distribution is computed at ♦ 400 K, ▲ 500 K, ■ 600 K, and ▼ 700 K. The raw data is listed in the worksheets (xi_iC8_400K, xi_iC8_500K, xi_iC8_600K, and xi_iC8_700K) of the Supporting Information (SI2.xlsx file) of Ref. [390].

A.3. Appendix for Chapter 5

The following items are presented in this Supporting Information:

- Comparison between enthalpies of formation ($\Delta_f H_0$) at 298.15 K obtained from Linear Regression (LR), the Scott tables [416], and the DIPPR database [447].
- Variations of the coefficients of second-order groups with C as the central atom for Gibbs free energy ($G_0 - H_0$(0 K)) at different temperatures.
- Reaction equilibrium distribution of C_{10} isomers relative to n-C_{10} in the gas phase at infinite dilution and 500 K.

Fig. A.20 shows the comparison between $\Delta_f H_0$ of C_9 and C_{10} isomers at 298.15 K obtained from the Scott tables [416], predictions by our Linear Regression (LR) model, and those listed in the DIPPR database [447]. $\Delta_f H_0$ of C_{10} isomers (Fig. A.20) listed in the DIPPR database [447] are computed using the method by Domalski and Hearing [445]. The values predicted using our LR model are in excellent agreement with the experimental data from the Scott tables. The values obtained from the DIPPR database [447] are also in good agreement with the Scott tables [416] with small deviations for 4-m-C_8, 2-m-C_9, and 3-m-C_9.

Fig. A.21 shows the variations in the coefficients of the second-order groups for $(G_0 - H_0(0\text{ K}))$ with C as the center atom present in the training dataset at different temperatures. The fitted coefficients using the quadratic polynomial are in excellent agreement with those predicted using the LR model. In case of first-order group contributions, all entries will be identical as each group has the same central united atom C. Therefore, combining these coefficients into a single coefficient to reduce the number of independent variables or simply using first-order group contributions will lead to erroneous predictions of the thermochemical properties. This clearly indicates the need for a second-order group contribution method.

Fig. A.22 shows the reaction equilibrium distribution of C_{10} isomers relative to n-C_{10} imposed in the gas phase at infinite dilution and 500 K. The reaction equilibrium distribution obtained using the Scott tables [416] and our LR model is in very good agreement. The gas phase distribution combined with phase equilibrium between the gas and the adsorbed phases provide the reaction equilibrium distribution inside the pores of MTW-type zeolite for the hydroisomerization reaction (Fig. 5.4 of section 5.3).

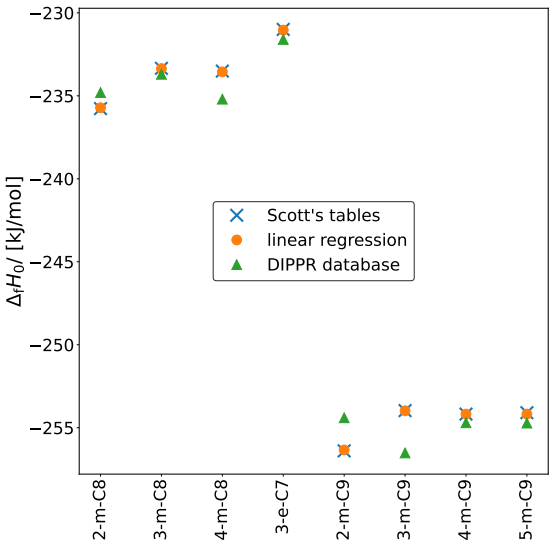


Figure A.20. Comparison between the predictions of $\Delta_f H_0$ at 298.15 K for C_9 and C_{10} isomers using our Linear Regression (LR) model and the data obtained from the DIPPR database [447]. The predictions using our Linear Regression model are in excellent agreement with the data obtained from the Scott tables [416].

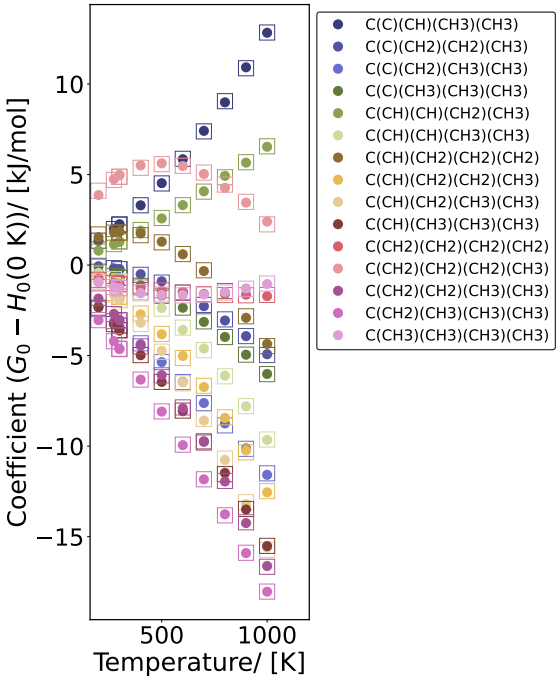


Figure A.21. Variations of the coefficients of second-order groups with C as the central atom at different temperatures. This plot includes all second-order groups with C as the central atom present in the training dataset [416] for Gibbs free energy ($G_0 - H_0$(0 K)). The colored circles represent the coefficients of the second-order groups obtained using linear regression and the square symbols are for those obtained using the temperature-dependent quadratic polynomial fit. The quadratic polynomial is fitted to the coefficients obtained using the LR model.

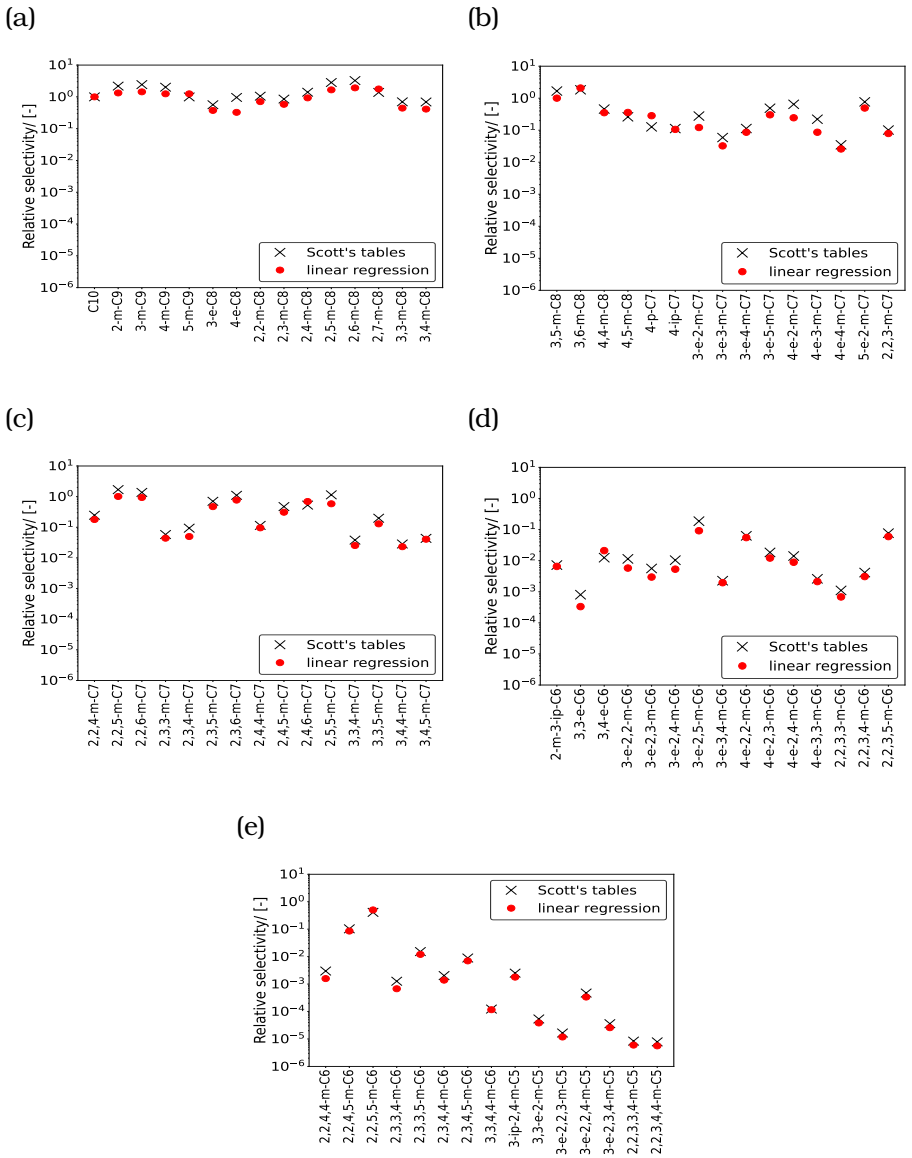


Figure A.22. Gas phase distribution of C_{10} isomers relative to n-C_{10} at reaction equilibrium in the gas phase at infinite dilution and 500 K. Both Scott tables (black crosses) and our Linear Regression (LR) model (red filled circles) are used to predict Gibbs free energies ($G_0 - H_0(0\ \mathrm{K})$) at 500 K and and enthalpies of formation ($\Delta_f H_0$) at 0 K which are used to compute the ideal gas chemical potentials of C_{10} isomers. These chemical potentials are used to compute the gas phase reaction equilibrium distribution of hydroisomerization of C_{10} [390]. The chemical reaction equilibrium distribution obtained using the Scott tables and the LR model is in very good agreement. The raw data is listed in the Excel worksheet xi_iC10_500K of the Supporting Information SI4.xlsx of Ref. [435].

A.4. Appendix for Chapter 7

The following items are presented in this Appendix:

1. Error analysis of Henry coefficients computed from molecular simulations
2. Parity plots for predictions of the negative logarithm of Henry coefficients for linear (C_1-C_{30}) and methyl-branched (C_4-C_{20}) alkanes in MTW-, MTT-, MRE-, and AFI-type zeolites at 523 K using Random Forest (RF), Extreme Gradient Boosting (XGB), Cat Boost (CB), Tabular Prior Fitted Network (TabPFN), and Directed Message Passing Neural Network (D-MPNN).
3. Parity plots for predictions of the negative logarithm of Henry coefficients for linear (C_1-C_{30}) and methyl-, ethyl-, propyl-, and isopropyl-branched (C_4-C_{20}) alkanes in MTW-, MTT-, MRE-, and AFI-type zeolites at 523 K using Random Forest (RF), Extreme Gradient Boosting (XGB), Cat Boost (CB), Tabular Prior Fitted Network (TabPFN), and Directed Message Passing Neural Network (D-MPNN).
4. Reaction equilibrium distribution for hydroisomerization of C_{16} isomers in MTW-type zeolite at 523 K.
5. Bar plots showing average Henry coefficients and average selectivities for different categories of C_{16} isomers relative to linear C_{16} at reaction equilibrium for hydroisomerization in MTW-, MTT-, MRE-, and AFI-type zeolites at 523 K.

A.4.1. Error analysis of Henry coefficients computed from molecular simulations

While error bars are generally useful, including these on the parity plots for CBMC-derived Henry coefficients is not meaningful in this context. For linear and moderately branched alkanes, the CBMC simulations yield small error bars (Table A.5) due to efficient sampling and high insertion probabilities. The errors in RASPA2 are computed as the 95% confidence interval. For long-chain and highly branched alkanes in narrow-pore zeolites, the error bars may become disproportionately large (Table A.6). This is a consequence of the poor fit of these molecules inside the zeolite channels. The low values in Table A.6 indicate that, at the experimental pressures (ca. 20 bar), these molecules will not adsorb appreciably. Therefore, whether the Henry coefficient is on the order of 10^{-40} or 10^{-80} is irrelevant at 20 bar, they will not fit in the channels regardless. These large uncertainties do not reflect inherent physical variability, but rather the limitations of the Configurational-Bias Monte Carlo method in these extreme cases. Moreover, such bulky isomers are typically of limited relevance for adsorption in narrow-pore zeolites because these isomers simply do not fit well in the zeolite pores,

A

Table A.5. Henry coefficients and corresponding errors of some linear and mono methyl branched alkanes in MTT-type zeolite at 523 K.

Isomer	Henry coefficient/ [mol/kg/framework]	Error/ [mol/kg/framework]
C_{14}	1.26×10^{-2}	4.78×10^{-4}
3-m-C_8	1.53×10^{-5}	1.11×10^{-7}
6-m-C_{11}	6.97×10^{-4}	3.88×10^{-5}

Table A.6. Henry coefficients and corresponding errors of some highly branched alkanes in MTT-type zeolite at 523 K.

Isomer	Henry coefficient/ [mol/kg/framework]	Error/ [mol/kg/framework]
4,4-e-2,2,3,3-m-6-ip-C_9	1.64×10^{-86}	5.04×10^{-55}
4,6-e-2-m-4,5-ip-C_9	2.68×10^{-128}	4.56×10^{-86}
3,3-e-2,4-m-5,5-p-C_8	1.76×10^{-130}	2.78×10^{-62}

and including the respective error bars would distract from the overall trends in the data.

A.4.2. Comparison between Henry coefficients for alkanes predicted using different machine learning models

Parity plots for the negative logarithm of Henry coefficients -ln(k_H) for linear (C_1-C_{30}) and methyl-branched (C_4-C_{20}) alkanes predicted by RF [82], XGB [83], CB [84, 85], TabPFN [89, 90], and D-MPNN [86–88, 92] are shown in Figs. A.23-A.26 for MTW-, MTT-, MRE-, and AFI-type zeolites at 523 K. Similarly, parity plots for -ln(k_H) for linear (C_1-C_{30}) and methyl-, ethyl-, propyl-, and isopropyl-branched (C_4-C_{20}) alkanes are shown in Figs. A.27-A.30 for MTW-, MTT-, MRE-, and AFI-type zeolites at 523 K. In most cases, TabPFN and D-MPNN provide better predictions (larger R^2) compared to the other ML models.

In Fig. A.31, the effect of training set size on the accuracy of the D-MPNN model in predicting -ln(k_H) for linear (C_1–C_{30}) and methyl-branched alkanes (C_4–C_{20}) alkanes in MTW- and MTT-type zeolites at 523 K was tested for the active learning and random selection strategies. The initial 50 molecular structures selected by active learning from the training sets, comprising linear alkanes (C_1–C_{30}) and methyl-branched alkanes (C_4–C_{20}), as well as linear alkanes (C_1–C_{30}) and methyl-, ethyl-, propyl-, and isopropyl-branched alkanes (C_4–C_{20}) are shown in Figs. A.32.

Activity cliffs present a significant challenge in predicting Henry coefficients of alkanes in zeolites, as minor changes in molecular structure can cause large, order of magnitude variations in adsorption behavior. We ex-

Table A.7. Henry coefficients (k_H) of alkanes in zeolites: actual values and predictions from TabPFN model without and with oversampling of high activity cliff isomers.

Isomer	k_H simulation/ [mol/kg/Pa]	k_H predicted/ [mol/kg/Pa]	
		without oversampling	with oversampling
2,7-m-C_{16}	6.185×10^{-3}	3.314×10^{-3}	3.496×10^{-3}
2-m-C_8	2.327×10^{-5}	2.759×10^{-5}	2.621×10^{-5}
3,11-m-C_{18}	5.404×10^{-2}	4.003×10^{-2}	5.308×10^{-2}
4,4-m-C_{16}	5.265×10^{-3}	4.669×10^{-3}	5.133×10^{-3}
2,4-m-C_{18}	3.072×10^{-2}	1.782×10^{-2}	1.793×10^{-2}
3,9-m-C_{11}	8.246×10^{-6}	1.579×10^{-5}	1.389×10^{-5}
3-m-C_{14}	1.629×10^{-3}	2.828×10^{-3}	2.751×10^{-3}
7,8-m-C_{16}	3.357×10^{-4}	1.201×10^{-3}	8.177×10^{-4}
2,2-m-C_9	1.877×10^{-5}	2.153×10^{-5}	2.006×10^{-5}
3,7-m-C_{16}	6.064×10^{-2}	1.379×10^{-2}	1.496×10^{-2}
3,6-m-C_{10}	3.744×10^{-6}	6.667×10^{-6}	6.645×10^{-6}
2,3,5,11,13-m-C_{15}	1.015×10^{-7}	1.944×10^{-8}	2.477×10^{-8}
4,4-m-C_{17}	2.475×10^{-2}	1.147×10^{-2}	1.182×10^{-2}
2,6-m-C_{17}	1.754×10^{-1}	8.092×10^{-2}	9.009×10^{-2}
3,14-m-C_{16}	1.507×10^{-3}	2.321×10^{-3}	2.014×10^{-3}
4,7-m-C_{10}	1.120×10^{-6}	2.227×10^{-6}	2.000×10^{-6}
2-m-C_{16}	4.953×10^{-2}	5.447×10^{-2}	5.275×10^{-2}
8,10-m-C_{18}	8.732×10^{-3}	1.748×10^{-2}	1.680×10^{-2}
2,12-m-C_{13}	2.906×10^{-4}	1.803×10^{-4}	2.004×10^{-4}
C_{11}	8.278×10^{-4}	9.070×10^{-4}	8.803×10^{-4}
2,8-m-C_{16}	1.784×10^{-3}	2.297×10^{-3}	2.109×10^{-3}
2,4-m-C_9	5.612×10^{-6}	5.057×10^{-6}	5.258×10^{-6}
2,2,4,5-m-C_6	1.343×10^{-8}	1.268×10^{-10}	1.498×10^{-10}
2,10-m-C_{18}	4.625×10^{-1}	1.349×10^{-1}	1.404×10^{-1}
3,4-m-C_9	6.550×10^{-7}	5.632×10^{-7}	5.800×10^{-7}
4-m-C_9	3.492×10^{-5}	1.558×10^{-5}	1.607×10^{-5}
5,8-m-C_{15}	1.081×10^{-4}	2.512×10^{-4}	1.984×10^{-4}
4,4,6,6-m-C_9	2.399×10^{-15}	1.586×10^{-9}	3.597×10^{-10}
6-m-C_{18}	2.032×10^{-1}	2.740×10^{-1}	2.384×10^{-1}
3,9-m-C_{12}	1.814×10^{-5}	3.101×10^{-5}	2.698×10^{-5}
8,9-m-C_{18}	1.703×10^{-3}	6.603×10^{-3}	5.272×10^{-3}
2,2,5,5,6,9-m-C_{14}	3.365×10^{-10}	9.575×10^{-14}	1.480×10^{-13}
2,2,4,4,7-m-C_8	8.419×10^{-17}	2.512×10^{-14}	5.188×10^{-15}
3,5-m-C_8	1.564×10^{-6}	1.294×10^{-6}	1.365×10^{-6}
2,8-m-C_{14}	2.159×10^{-4}	3.627×10^{-4}	3.166×10^{-4}
4-m-C_{13}	1.563×10^{-3}	7.999×10^{-4}	8.473×10^{-4}
2,13-m-C_{17}	5.371×10^{-3}	2.026×10^{-2}	1.792×10^{-2}
3,5,6,6,10-m-C_{12}	1.174×10^{-8}	4.924×10^{-10}	7.424×10^{-10}
2,2,4,11,11-m-C_{12}	1.618×10^{-8}	4.459×10^{-9}	7.135×10^{-9}
2,9-m-C_{11}	4.610×10^{-5}	6.231×10^{-5}	5.211×10^{-5}
2,8-m-C_{17}	4.534×10^{-3}	5.363×10^{-3}	4.898×10^{-3}
3,13-m-C_{17}	2.695×10^{-3}	4.725×10^{-3}	4.276×10^{-3}
3,8-m-C_{15}	2.505×10^{-4}	4.814×10^{-4}	4.309×10^{-4}
4,5-m-C_9	8.971×10^{-7}	8.748×10^{-7}	9.059×10^{-7}
3,10-m-C_{14}	1.267×10^{-4}	2.531×10^{-4}	2.105×10^{-4}
3,7-m-C_{15}	1.116×10^{-2}	5.460×10^{-3}	5.858×10^{-3}
4,5-m-C_{12}	6.873×10^{-6}	1.265×10^{-5}	1.235×10^{-5}
3,5-m-C_{10}	1.213×10^{-5}	8.660×10^{-6}	9.023×10^{-6}
4,14-m-C_{17}	1.726×10^{-3}	4.744×10^{-3}	3.597×10^{-3}
5,10-m-C_{15}	1.542×10^{-3}	9.916×10^{-4}	1.010×10^{-3}
3,4-m-C_{11}	2.496×10^{-6}	4.008×10^{-6}	3.792×10^{-6}
6-m-C_{17}	7.619×10^{-2}	1.061×10^{-1}	1.008×10^{-1}
7,8-m-C_{15}	8.732×10^{-5}	3.336×10^{-4}	2.531×10^{-4}
2,2,4,7-m-C_9	7.476×10^{-9}	2.158×10^{-9}	2.529×10^{-9}
3,3-m-C_{12}	3.800×10^{-4}	1.100×10^{-4}	1.139×10^{-4}
2,14-m-C_{18}	3.061×10^{-1}	8.094×10^{-2}	8.388×10^{-2}

amined the effect of oversampling high activity cliff isomers in the training set, as described in Section 2 of the main text. Oversampling 20% of these isomers yielded a modest 4% increase in R^2. Table A.7 summarizes the improvements in Henry coefficient predictions for these isomers obtained using the TabPFN model. Expanding the dataset with additional isomers will enhance predictive performance, and complementary approaches such as contrastive learning should be explored.

A.4.3. Reaction equilibrium distribution for hydroisomerization of C_{16} isomers

Selectivities of C_{16} isomers relative to n-C16 at reaction equilibrium in MTW-type zeolite are shown in Fig. A.33. Mono-methyl isomers show larger selectivities than di-methyl isomers. The high selectivity of 2-m-C_{15} relative to n-C_{16} is attributed to its larger gas-phase mole fraction. For di-methyl alkanes, isomers with methyl groups positioned farther apart (e.g., 2,6-m-C_{14} to 2,13-m-C_{14}) are favored compared to those with closely spaced methyl groups (e.g., 2,2-m-C_{14} to 2,5-m-C_{14}), likely due to reduced steric hindrance in the zeolite pores. The raw data for the reaction equilibrium distribution of C_{16} isomers are provided in SI2_HC.xlsx. Average Henry coefficients for different groups of C_{16} isomers and average selectivities of these groups relative to linear C_{16} in MTW-, MTT-, MRE-, and AFI-type zeolites are shown in Figs. A.34-A.37. Linear, mono-, di-, tri-, and tetra-methyl, and mono-ethyl isomers are the most favorable groups in terms of relative selectivities in all zeolites considered in this study. Highly branched isomers such as those with both ethyl and propyl groups adsorb poorly in these zeolites (Figs. A.34-A.37), reflected by low Henry coefficients.

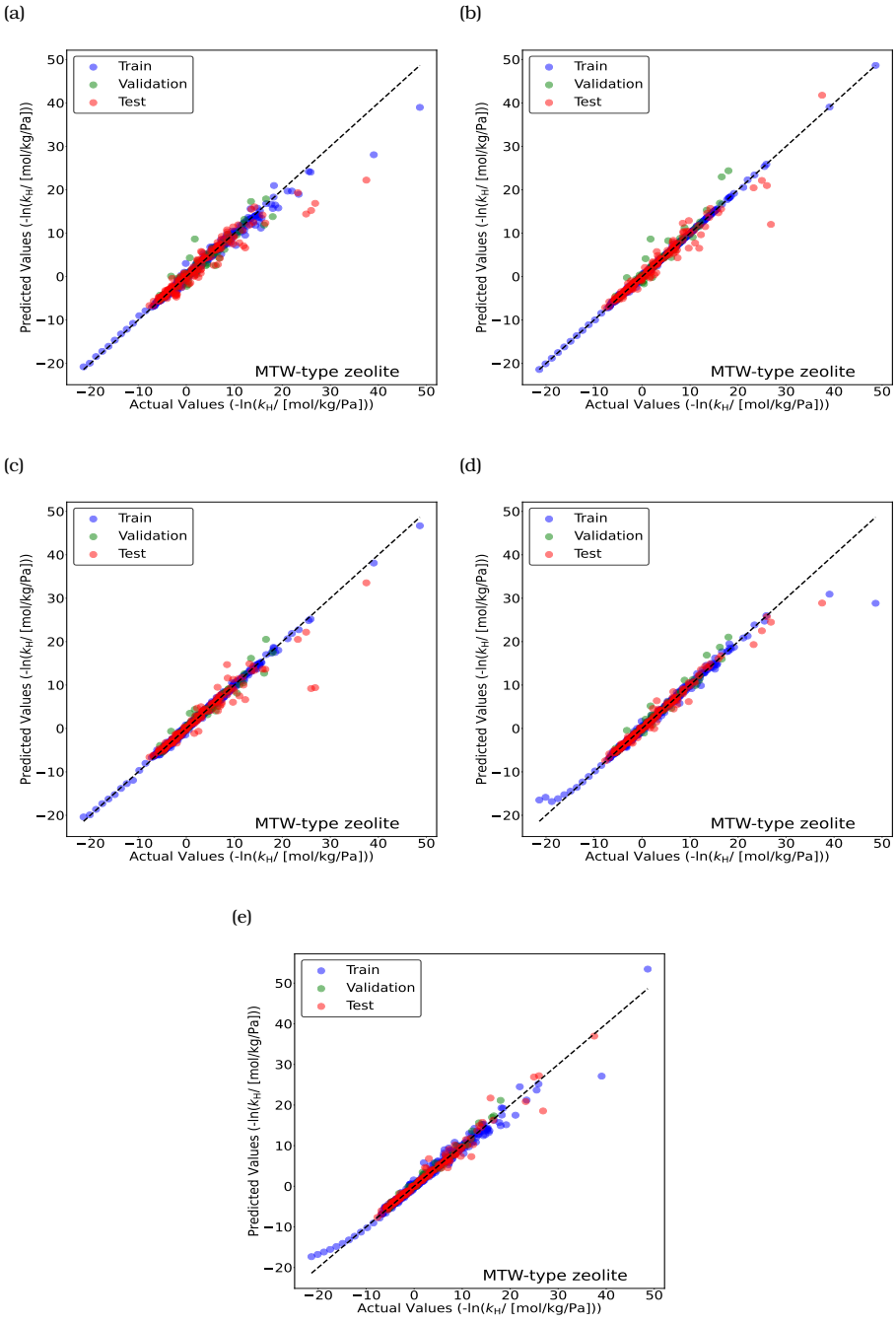


Figure A.23. Parity plots for predicting the negative logarithm of Henry coefficients -ln(k_H) for linear alkanes (C_1–C_{30}) and methyl-branched alkanes (C_4–C_{20}) in MTW-type zeolite at 523 K using (a) RF, (b) XGB, (c) CB, (d) TabPFN, and (e) D-MPNN models. Blue circles indicate training isomers and green circles represent validation isomers. The random seed is varied to identify the most suitable split between the training and the validation datasets. Red circles represent the test isomers which are never part of the training set. The standard deviations for the actual values of -ln(k_H) are too small to plot.

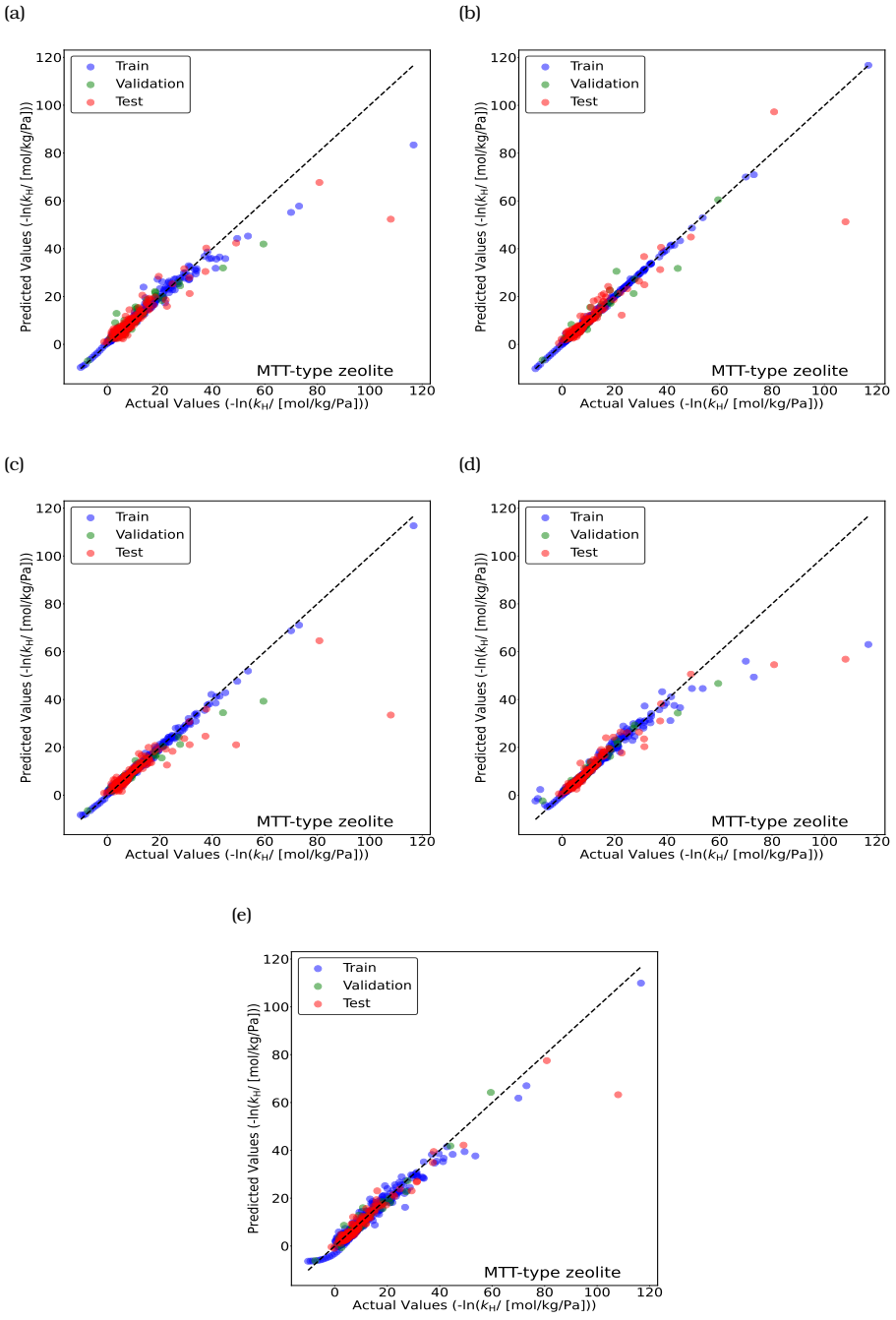


Figure A.24. Parity plots for predicting the negative logarithm of Henry coefficients -ln(k_H) for linear alkanes (C_1–C_{30}) and methyl-branched alkanes (C_4–C_{20}) in MTT-type zeolite at 523 K using (a) RF, (b) XGB, (c) CB, (d) TabPFN, and (e) D-MPNN models. Blue circles indicate training isomers and green circles represent validation isomers. The random seed is varied to identify the most suitable split between the training and the validation datasets. Red circles represent the test isomers which are never part of the training set. The standard deviations for the actual values of -ln(k_H) are too small to plot.

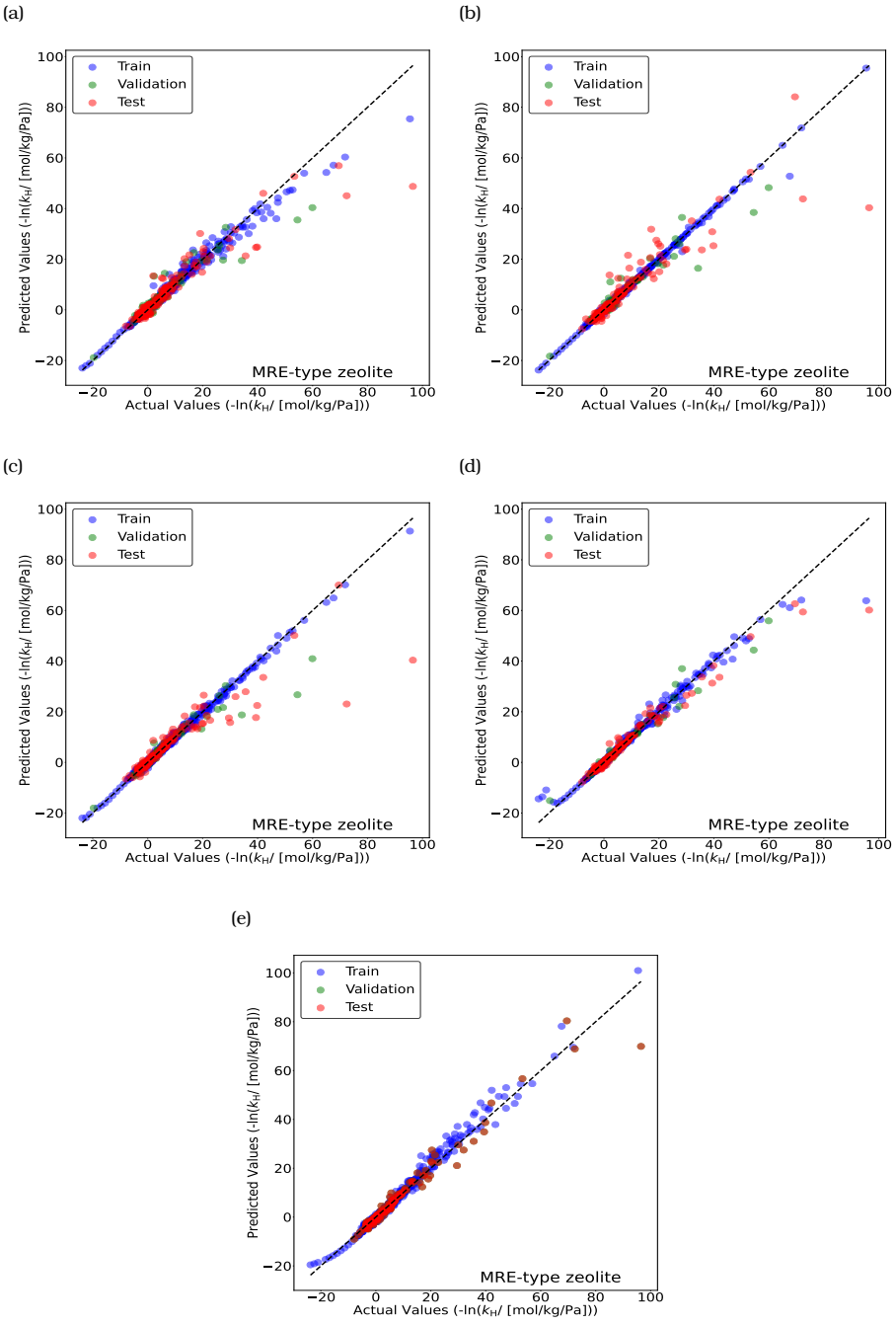


Figure A.25. Parity plots for predicting the negative logarithm of Henry coefficients -ln(k_H) for linear alkanes (C_1–C_{30}) and methyl-branched alkanes (C_4–C_{20}) in MRE-type zeolite at 523 K using (a) RF, (b) XGB, (c) CB, (d) TabPFN, and (e) D-MPNN models. Blue circles indicate training, and green circles represent validation isomers. The random seed is varied to identify the most suitable split between the training and the validation datasets. Red circles represent the test isomers, which are never part of the training set. The standard deviations for the actual values of -ln(k_H) are too small to plot.

A

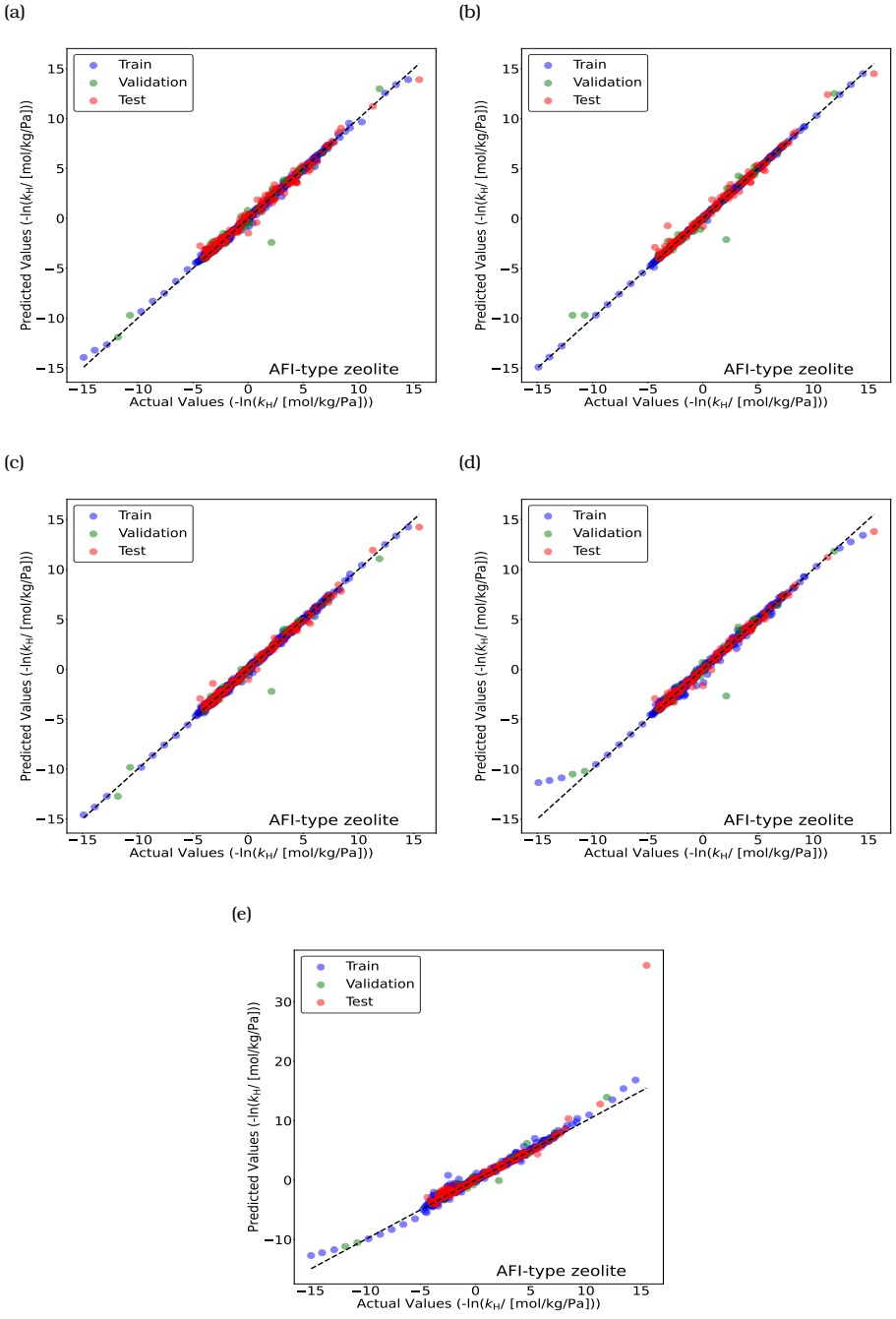


Figure A.26. Parity plots for predicting the negative logarithm of Henry coefficients -ln(k_H) for linear alkanes (C_1–C_{30}) and methyl-branched alkanes (C_4–C_{20}) in AFI-type zeolite at 523 K using (a) RF, (b) XGB, (c) CB, (d) TabPFN, and (e) D-MPNN models. Blue circles indicate training isomers and green circles represent validation isomers. The random seed is varied to identify the most suitable split between the training and the validation datasets. Red circles represent the test isomers which are never part of the training set. The standard deviations for the actual values of -ln(k_H) are too small to plot.

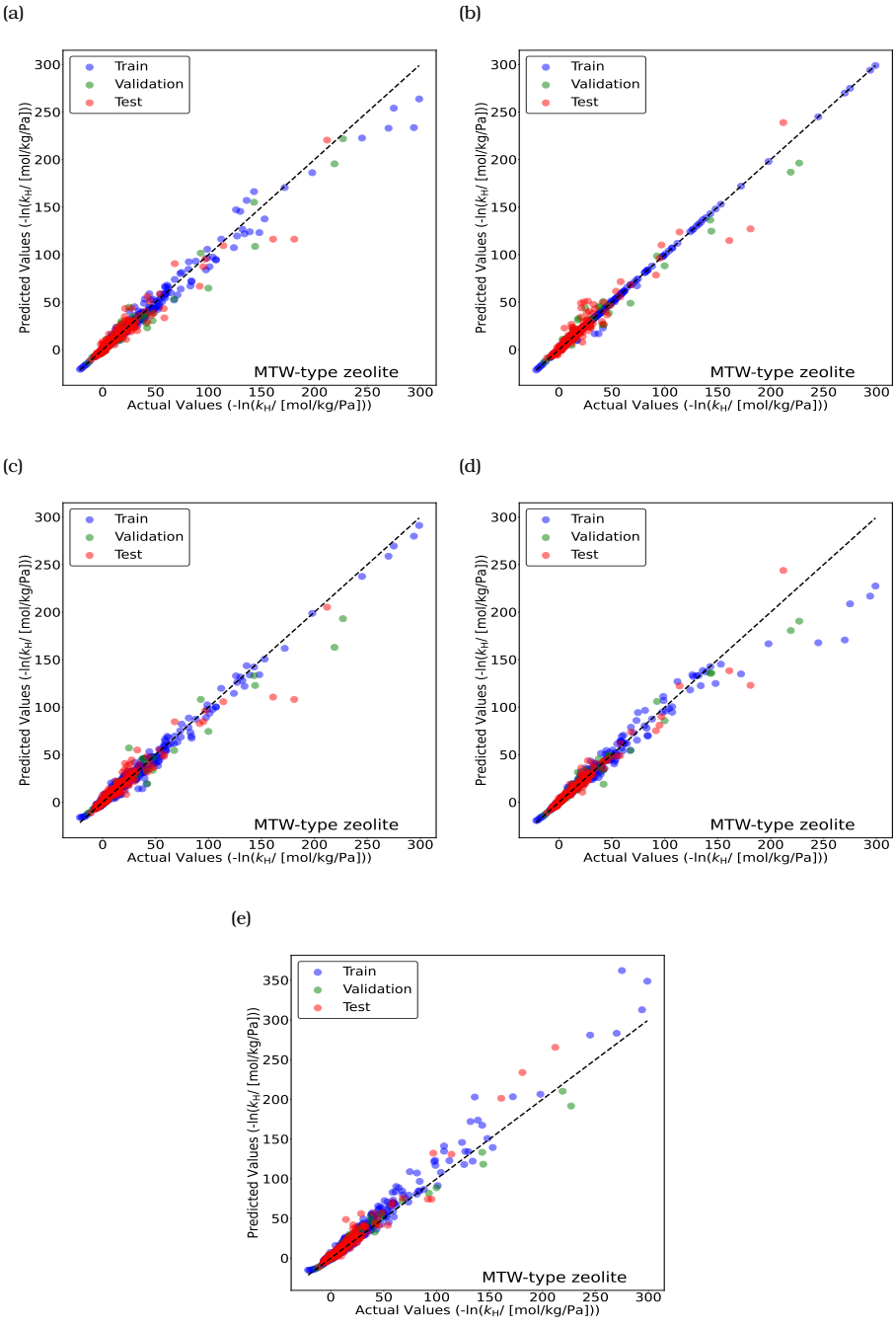


Figure A.27. Parity plots for predicting -ln(k_H), for linear (C_1–C_{30}) and methyl-, ethyl-, propyl-, and isopropyl-branched alkanes (C_4–C_{20}) in MTW-type zeolite at 523 K using (a) RF, (b) XGB, (c) CB, (d) TabPFN, and (e) D-MPNN models. Blue circles indicate training isomers and green circles represent validation isomers. The random seed is varied to identify the most suitable split between the training and the validation datasets. Red circles represent the test isomers which are never part of the training set. The standard deviations for the actual values of -ln(k_H) are too small to plot.

A

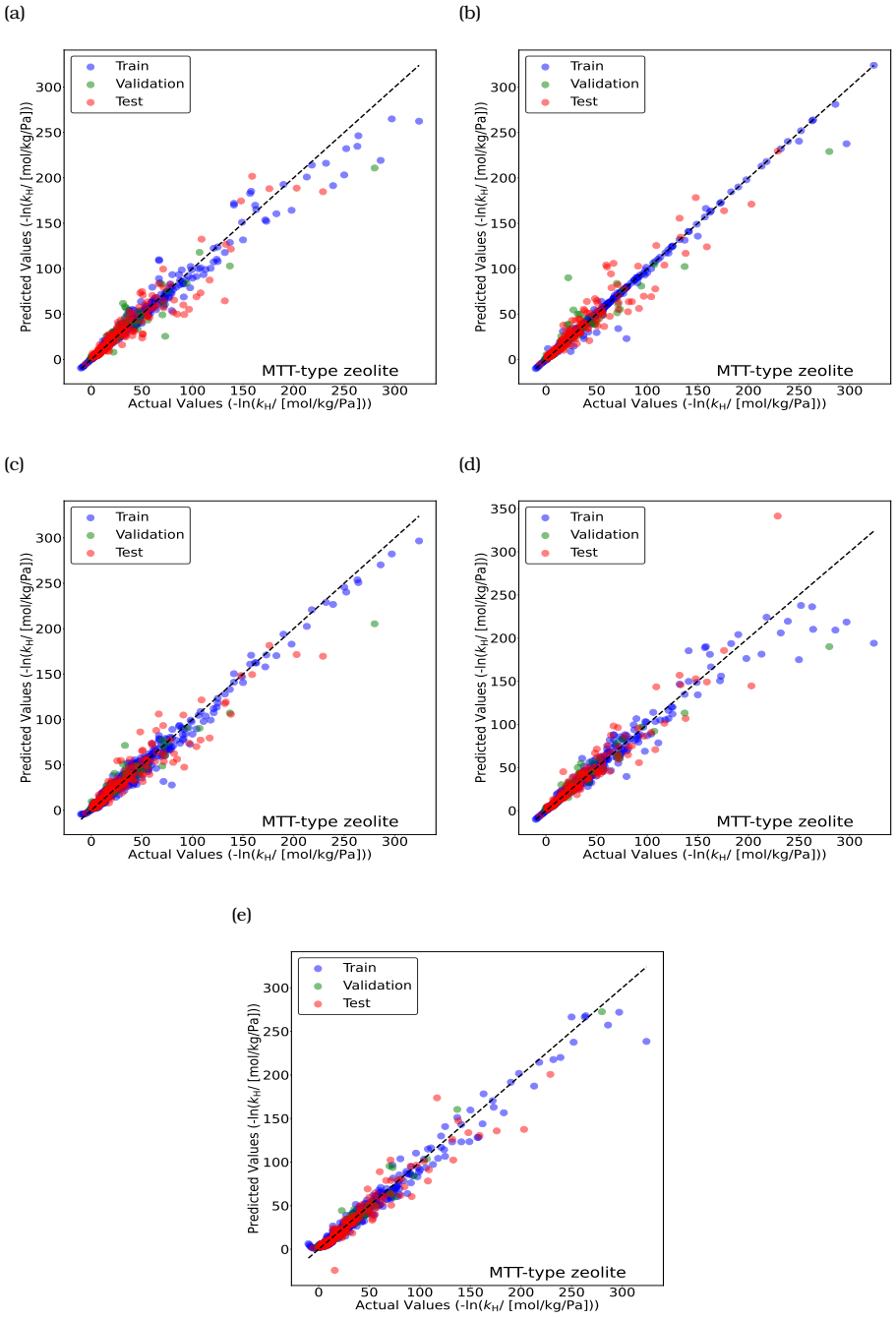


Figure A.28. Parity plots for predicting -ln(k_H), for linear (C_1–C_{30}) and methyl-, ethyl-, propyl-, and isopropyl-branched alkanes (C_4–C_{20}) in MTT-type zeolite at 523 K using (a) RF, (b) XGB, (c) CB, (d) TabPFN, and (e) D-MPNN models. Blue circles indicate training isomers and green circles represent validation isomers. The random seed is varied to identify the most suitable split between the training and the validation datasets. Red circles represent the test isomers which are never part of the training set. The standard deviations for the actual values of -ln(k_H) are too small to plot.

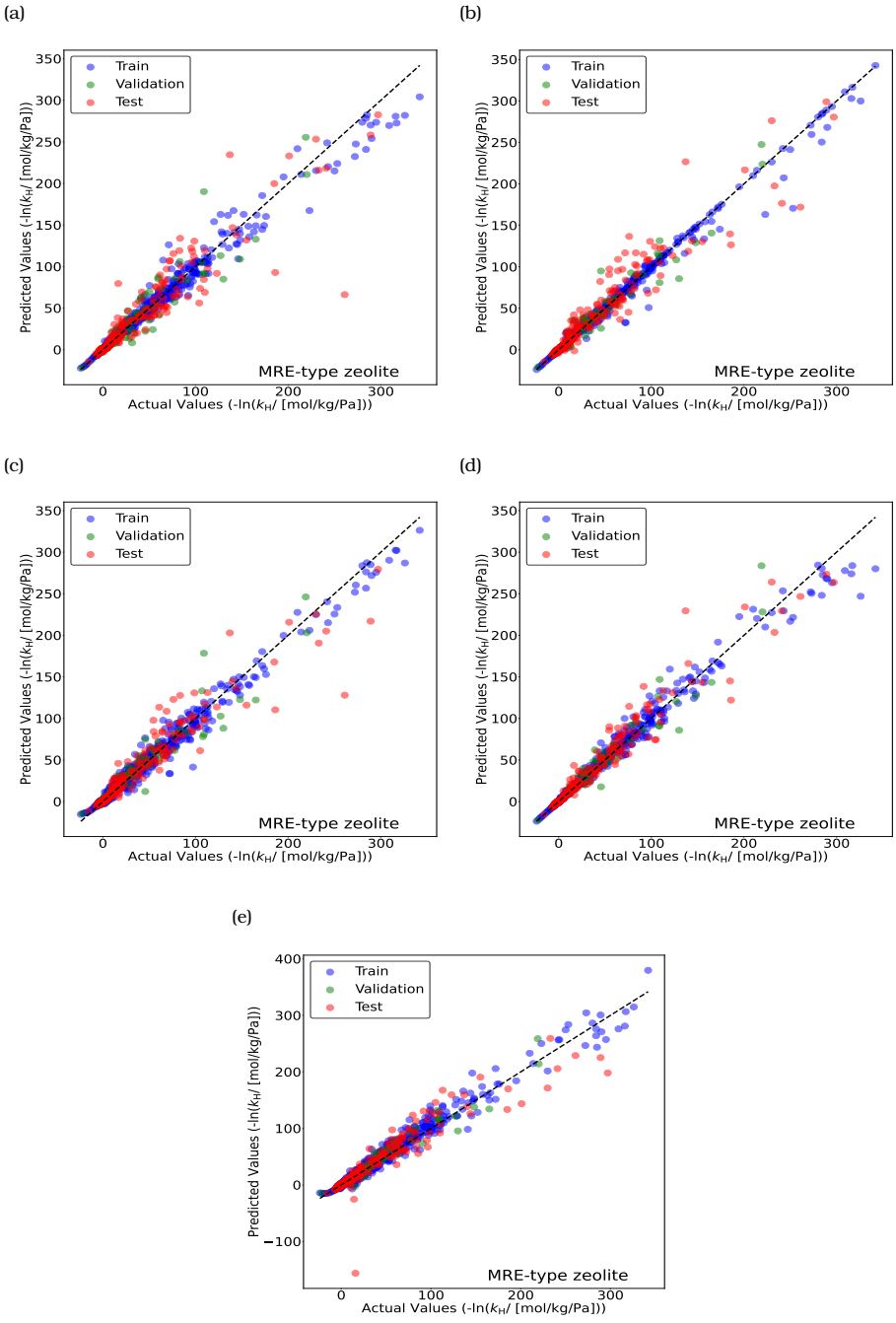


Figure A.29. Parity plots for predicting -ln(k_H), for linear (C_1–C_{30}) and methyl-, ethyl-, propyl-, and isopropyl-branched alkanes (C_4–C_{20}) in MRE-type zeolite at 523 K using (a) RF, (b) XGB, (c) CB, (d) TabPFN, and (e) D-MPNN models. Blue circles indicate training isomers and green circles represent validation isomers. The random seed is varied to identify the most suitable split between the training and the validation datasets. Red circles represent the test isomers, which are never part of the training set. The standard deviations for the actual values of -ln(k_H) are too small to plot.

A

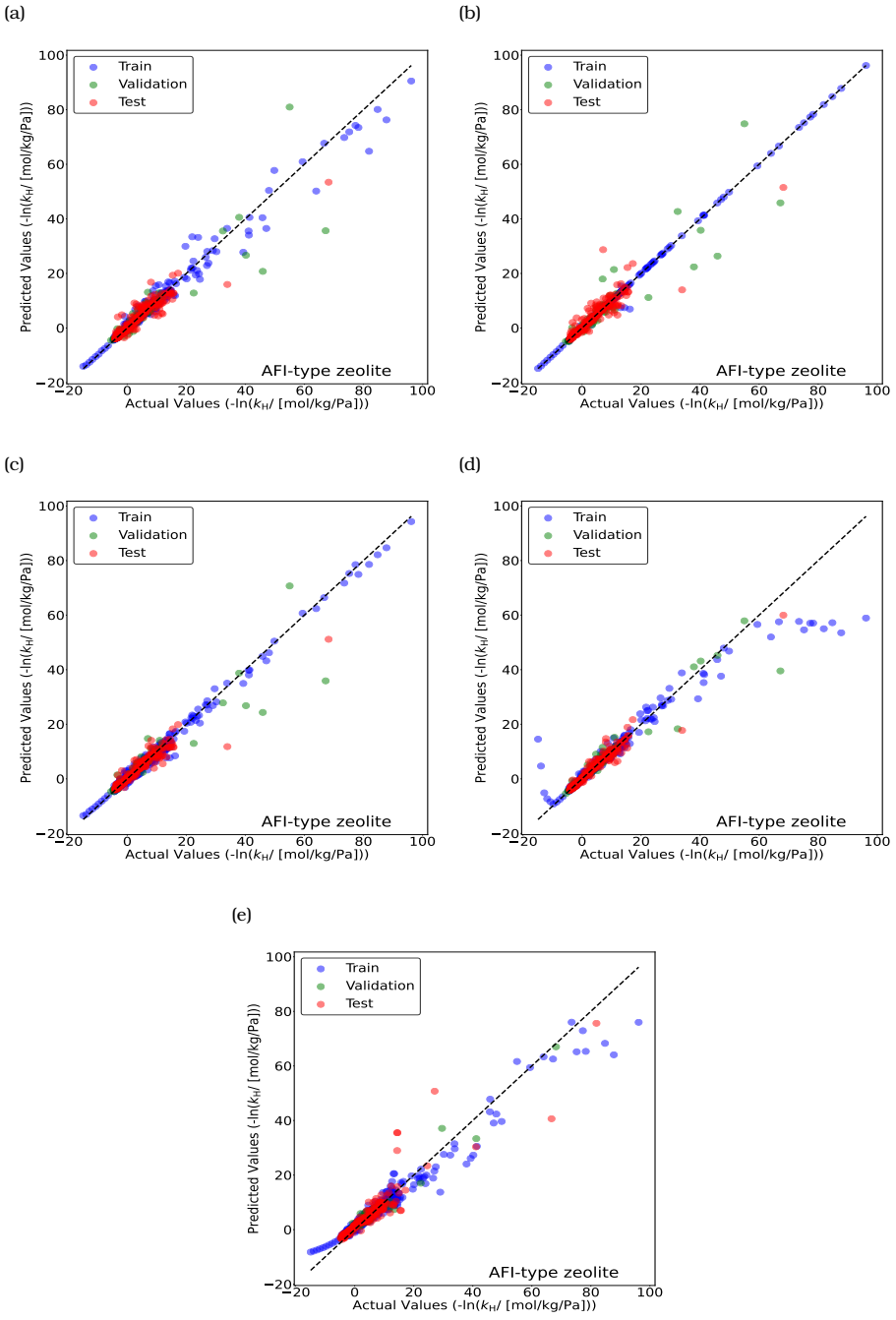


Figure A.30. Parity plots for predicting -ln(k_H), for linear (C_1–C_{30}) and methyl-, ethyl-, propyl-, and isopropyl-branched alkanes (C_4–C_{20}) in AFI-type zeolite at 523 K using (a) RF, (b) XGB, (c) CB, (d) TabPFN, and (e) D-MPNN models. Blue circles indicate training isomers and green circles represent validation isomers. The random seed is varied to identify the most suitable split between the training and the validation datasets. Red circles represent the test isomers which are never part of the training set. The standard deviations for the actual values of -ln(k_H) are too small to plot.

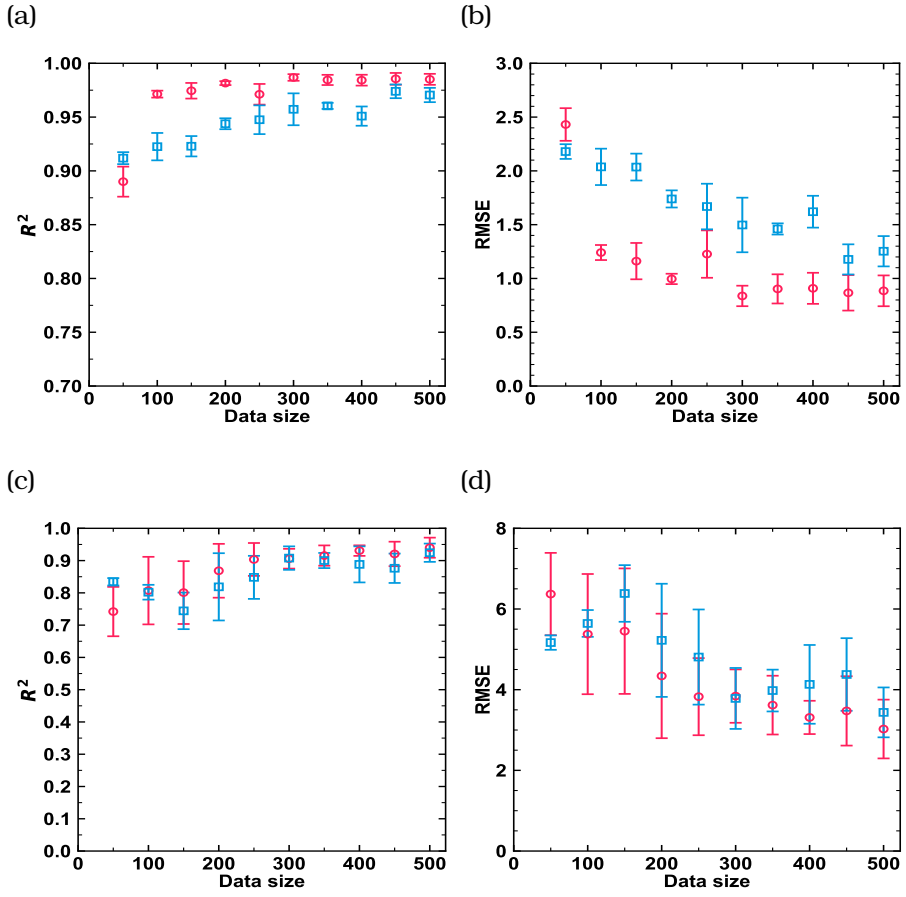


Figure A.31. Test accuracies of the D-MPNN model as a function of training set size, comparing an active learning strategy (red circles) and a random selection strategy (blue squares). R^2 and RMSE values are shown for models trained on the negative logarithm of Henry coefficients for linear (C_1–C_{30}) and methyl-branched (C_4–C_{20}) alkanes in MTW-type (a, b) and MTT-type (c, d) zeolites at 523 K.

A

(a) (b)

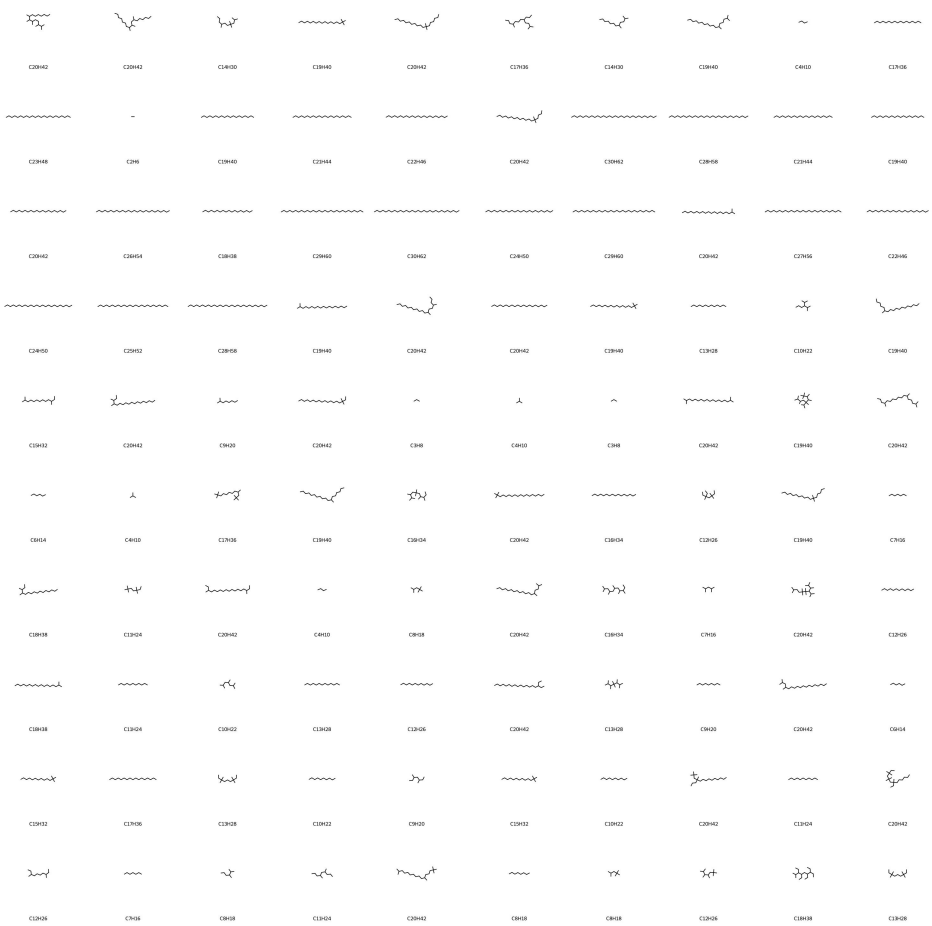


Figure A.32. Initial 50 molecular structures selected by active learning from a training set of (a) linear (C_1–C_{30}) and methyl-branched alkanes (C_4–C_{20}) and (b) linear (C_1–C_{30}) and methyl-, ethyl-, propyl-, and isopropyl-branched alkanes (C_4–C_{20}).

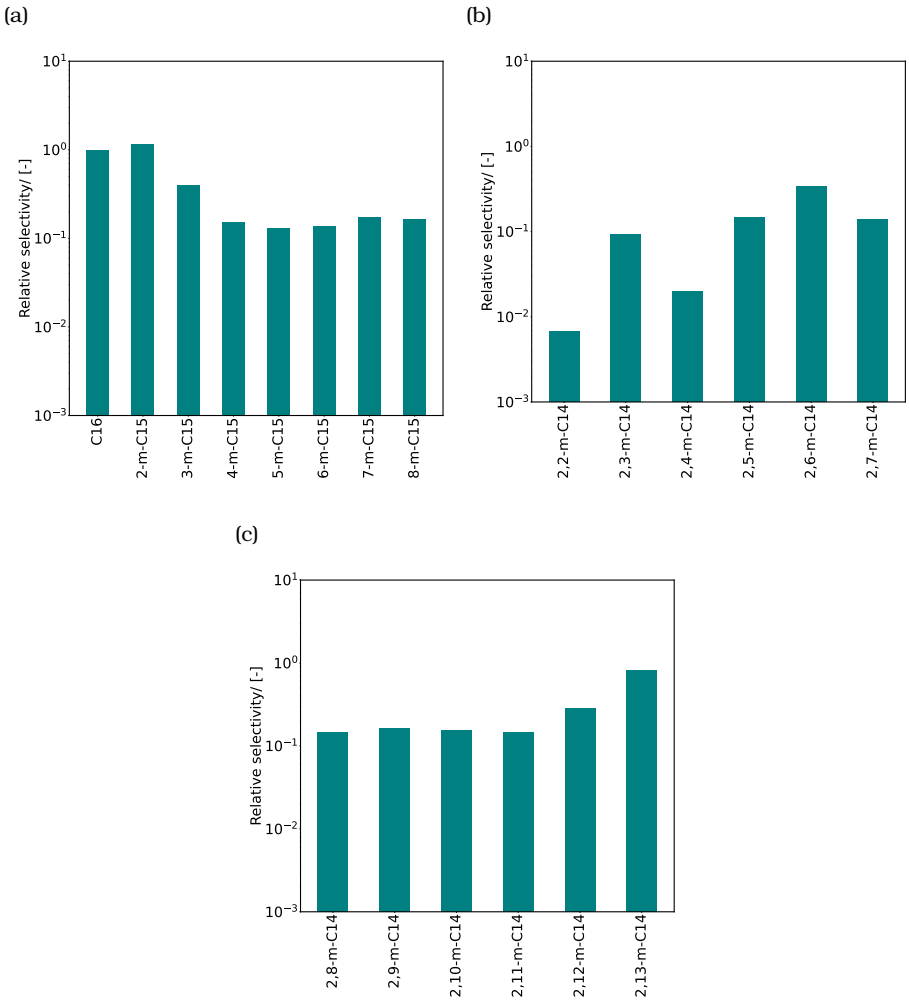


Figure A.33. Selectivities of (a) n-C_{16} and mono-methyl (2-m-C_{15} - 8-m-C_{15}), (b) di-methyl (2,2-m-C_{14} - 2,7-m-C_{14}), and (c) di-methyl (2,8-m-C_{14} - 2,13-m-C_{14}) isomers relative to n-C_{16} at reaction equilibrium in MTW-type zeolite at infinite dilution and 523 K. The absolute selectivities (Eq. 4.22) are defined as the mole fraction of a certain component divided by the sum of the mole fractions of all other components [429]. The relative selectivities (Eq. 4.24) refer to the ratio of the absolute selectivity of a specific isomer to that of the reference isomer, which is C_{16} in this case.

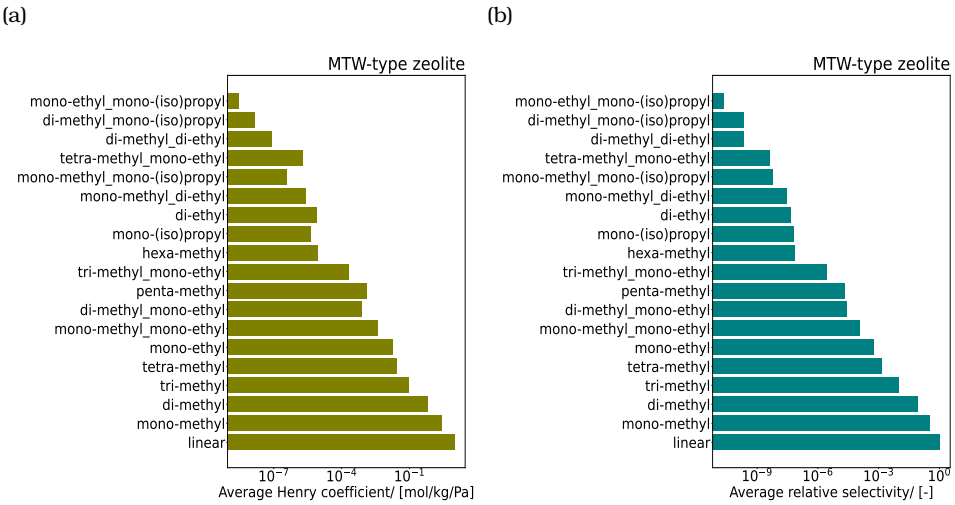


Figure A.34. (a) Average Henry coefficients for different categories of C_{16} isomers in MTW-type zeolite at 523 K predicted using the TabPFN model. (b) Average selectivities of different categories of C_{16} isomers relative to linear C_{16} at reaction equilibrium in MTW-type zeolite at 523 K. For each category, the Henry coefficients and the selectivities are averaged over all the isomers belonging to that category.

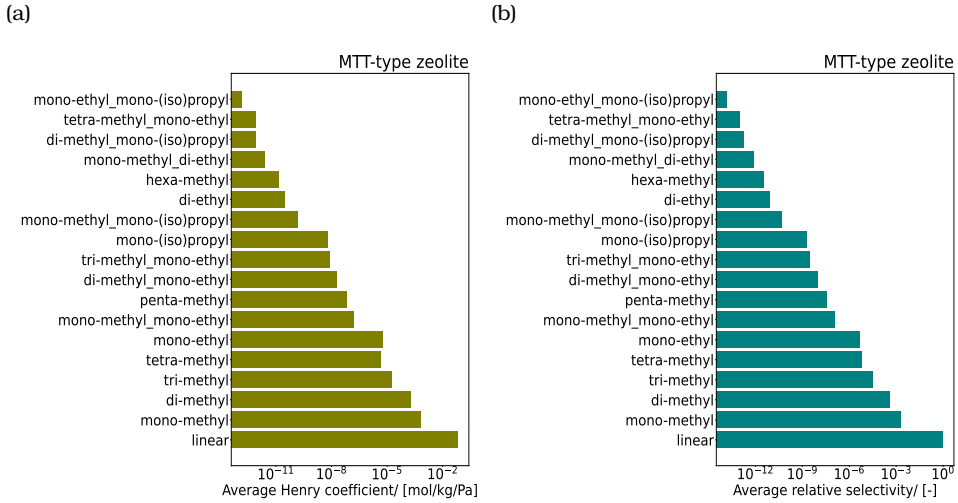


Figure A.35. (a) Average Henry coefficients for different categories of C_{16} isomers in MTT-type zeolite at 523 K predicted using the TabPFN model. (b) Average selectivities of different categories of C_{16} isomers relative to linear C_{16} at reaction equilibrium in MTT-type zeolite at 523 K. For each category, the Henry coefficients and the selectivities are averaged over all the isomers belonging to that category.

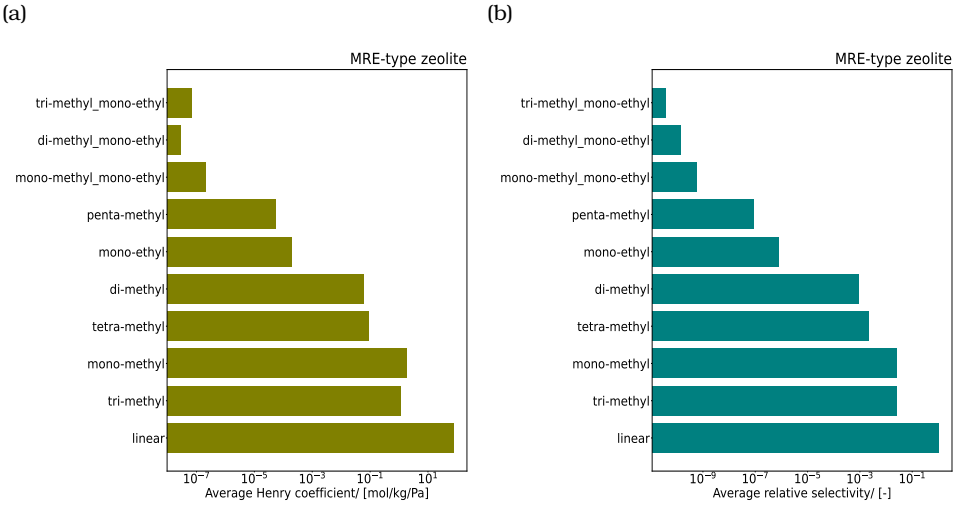


Figure A.36. (a) Average Henry coefficients for different categories of C_{16} isomers in MRE-type zeolite at 523 K predicted using the TabPFN model. (b) Average selectivities of different categories of C_{16} isomers relative to linear C_{16} at reaction equilibrium in MRE-type zeolite at 523 K. For each category, the Henry coefficients and the selectivities are averaged over all the isomers belonging to that category.

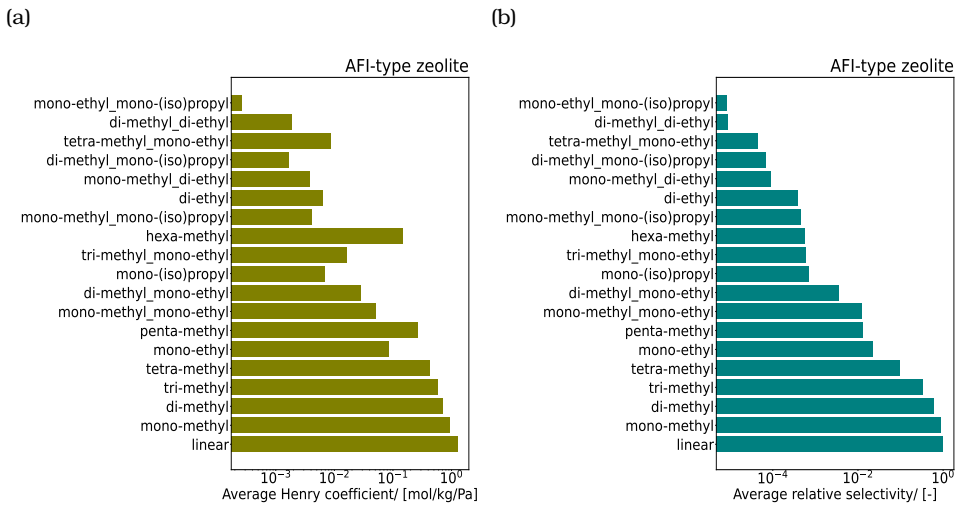


Figure A.37. (a) Average Henry coefficients for different categories of C_{16} isomers in AFI-type zeolite at 523 K predicted using the TabPFN model. (b) Average selectivities of different categories of C_{16} isomers relative to linear C_{16} at reaction equilibrium in AFI-type zeolite at 523 K. For each category, the Henry coefficients and the selectivities are averaged over all the isomers belonging to that category.

Summary

This thesis presents a multiscale modeling framework to investigate the shape-selectivity effects of zeolites on hydroisomerization of long-chain alkanes, combining adsorption thermodynamics, reaction equilibrium modeling, linear regression, and machine learning. The ultimate goal is to understand how molecular structure and pore topology determine the performance of zeolite-based catalysts for selective hydrocarbon upgrading. Accounting for surface heterogeneities in adsorption isotherms is essential for accurately modeling separation processes, where the competitive adsorption of alkane isomers at different adsorption sites strongly influences reaction equilibrium and selectivity. Variations in adsorption-site affinity for different isomers play an important role in determining product distributions of hydroisomerization. In chapter 2, a Segregated Explicit Isotherm (SEI) equation model is developed that captures adsorbate size-dependent behavior in materials with heterogeneous surface sites. The main advantage of SEI is that it uses only explicit equations, which enhances numerical stability and significantly accelerates computing speed, particularly when equilibrium loadings must be calculated in breakthrough curve simulations. In our study, SEI-based breakthrough simulations were about three times faster than those using Segregated Ideal Adsorbed Solution Theory (SIAST). The most significant speedup, however, was observed when comparing SIAST to IAST for multi-site isotherms: SIAST was approximately 20 times faster than IAST. This is because IAST requires two iterative steps: (1) computing the spreading pressure, and (2) inverting the spreading pressure function to obtain the pure component pressure. In sharp contrast, SIAST typically uses explicit expressions for the inverse of spreading pressure at each adsorption site, eliminating the need for iterations. To ensure these advanced isotherm and breakthrough models are accessible and easily usable for practical simulations, RUPTURA, an open-source software, is introduced in chapter 3. RUPTURA contains three core modules commonly used in adsorption process modeling: (1) simulation of step and pulse breakthrough experiments, (2) prediction of mixture adsorption based on pure component isotherms, and (3) fitting of isotherm models to raw adsorption data, either computed from GCMC simulations or experimentally measured. This software includes a wide range of isotherm models—such as Langmuir, BET, Henry, Freundlich, Sips, Langmuir-Freundlich, Redlich-Peterson, Toth, Unilan, O'Brien & Myers, Asymptotic Temkin, and Bingel & Walton—along with the multi-site extensions of these isotherms and hybrid formulations. For mixture adsorption prediction, IAST, segregated IAST, and explicit Langmuir-

based approaches are implemented in RUPTURA. IAST calculations are performed rapidly and with machine precision. Breakthrough curve simulations incorporate axial dispersion and use the Linear Driving Force (LDF) model for mass transfer, ensuring high numerical stability through Strong-Stability Preserving Runge-Kutta (SSP-RK) integrators. RUPTURA is designed for fast and flexible evaluation of fixed-bed separation processes involving mixtures of hydrocarbons and other molecules. While accurate adsorption isotherms are an important foundation, but cannot solely capture the full thermodynamics of hydroisomerization; incorporating the reaction equilibrium distribution is essential for a complete understanding of product selectivity. The reaction equilibrium distribution determines the thermodynamic limits of isomer yields at imposed operating conditions. Accurate prediction of this distribution is essential for optimizing catalyst selectivity and process efficiency. Chapter 4 focuses on calculating the reaction equilibrium distribution of alkane isomers in zeolites, combining gas-phase chemical equilibrium with phase equilibrium between the gas and the adsorbed phases. This framework demonstrates how shape selectivity alters the equilibrium product distributions in zeolite pores. The shape selectivity of zeolites for C_7 and C_8 hydroisomerization is evaluated. Gas-phase reaction equilibrium distributions are derived from the Gibbs energy of formation, i.e., isolated-molecule partition functions, while adsorption mole fractions are computed using Henry coefficients. In large-pore zeolites (e.g., FAU-type zeolite), adsorption has a small impact on equilibrium, whereas in narrow-pore zeolites (e.g., MTW- and MRE-type zeolites), Henry coefficients strongly affect the product distribution. Mono-branched isomers are generally favored compared to di-branched ones in medium-pore structures (e.g., MFI- and BEA-type zeolites), while multi-branched isomers are stabilized in cage-like frameworks (e.g., ITQ-29-type zeolite). Steric hindrance disfavors highly branched isomers like 2,2,3-m-C_4 in all zeolites. At elevated pressures and temperatures (500–700 K), deviations from the ideal gas phase hardly affect the reaction equilibrium distribution. Adsorption contributions are extremely important in determining the reaction product distribution at high pressures. Understanding these trends at equilibrium is crucial for optimizing hydroisomerization selectivity and guiding catalyst design. The required thermochemical data for long-chain alkanes are limited in the literature. Yet, such data are needed for accurately determining reaction product distributions at chemical equilibrium, especially in processes like hydroisomerization. Chapter 5 addresses this gap by developing a linear regression model based on second-order group contributions to predict key thermochemical properties: enthalpy of formation ($\Delta_f H_0$), Gibbs energy of formation ($\Delta_f G_0$), enthalpy ($H_0 - H_0(0\text{ K})$), and Gibbs energy ($G_0 - H_0(0\text{ K})$). The model exceeds chemical accuracy, with a maximum mean absolute error of just 1.03 kJ/mol for $\Delta_f G_0$, and outperforms established methods such as group contribution methods by Benson et al, Joback et al, and Constantinou–Gani. Temperature dependence is captured using quadratic

fits for the group contributions. The predicted gas-phase thermochemical properties, when coupled with Henry coefficients computed from the RASPA2 software, yield reaction equilibrium distributions that closely match those based on the Scott tables for $\Delta_f H_0$ and $(G_0 - H_0(0\ \mathrm{K}))$. For C_{10} and C_{14} isomers in MTW-type zeolites, the model indicates that isomers with long or closely spaced branches show low selectivity and have negligible contribution to the reaction product distribution. This framework provides a robust, scalable approach to evaluating equilibrium distributions of long-chain alkanes in zeolites. Accurate knowledge of thermochemical properties, $H_0 - H_0(0\ \mathrm{K})$ and $G_0 - H_0(0\ \mathrm{K})$, is also important in computing entropies, which are needed in calculating exergy destruction. Evaluating energy efficiency through the second law of thermodynamics requires quantifying exergy destruction. Exergy destruction represents the useful work lost due to irreversibilities and is a key metric for identifying inefficiencies in chemical processes. Chapter 6 extends the thermodynamic analysis of hydroisomerization by computing absolute entropies and second-law efficiencies using a Linear Regression (LR) model based on second-order group contributions. The computed entropies closely agree with the entropy values obtained from the Scott tables and experiments, capturing variations due to the number, type, and position of branches in alkane isomers. Using this model, the entropy production and heat input are analyzed for the hydroisomerization of C_7 and C_8 alkanes in various zeolites. Reaction product distributions vary with zeolite pore structure. Large pores in FAU- and ITQ-29-type zeolites favor multi-branched isomers, while narrow pores in MTW- and MRE-type zeolites prefer linear and mono-branched isomers. These differences lead to variations in entropy production and exergy destruction across zeolite types. For larger alkane chain lengths, both enthalpy and entropy increase due to greater molecular complexity, leading to higher heat input and entropy production. At elevated temperatures, increased molecular motion enhances configurational entropy, thereby increasing entropy production and exergy loss. This work demonstrates the ability of our LR model to accurately predict absolute entropies and provides valuable insight into the thermodynamic efficiency of hydroisomerization. These results are not only relevant for catalyst and process optimization, but can also be useful in energy analysis for a broader range of hydrocarbon-based processes. In addition to gas-phase thermochemical properties, Henry coefficients are essential for determining the reaction equilibrium distribution in hydroisomerization. Due to the large number of isomers for long-chain alkanes and the diversity of zeolite structures, computing Henry coefficients solely through molecular simulations becomes impractical. Due to this complexity and computational cost, an efficient predictive approach is needed to complement traditional molecular simulations. Chapter 7 presents a machine learning framework for predicting Henry coefficients in one-dimensional zeolites, combining descriptor-based (TabPFN) and graph-based (D-MPNN) models with automated isomer enumeration tools. TabPFN performs well

for linear and moderately-branched isomers, while D-MPNN excels for low-adsorbing, highly-branched structures. Despite challenges posed by activity cliffs and limited data for bulky isomers, active learning and oversampling improve performance. In hydroisomerization, under-predicting the Henry coefficients of sterically excluded isomers is not critical because the very low Henry coefficients indicate negligible adsorption in the narrow pores of the zeolite and, therefore, little influence on the reaction equilibrium distribution. By integrating ML-predicted Henry coefficients with thermodynamic data from a second-order group contribution model, this framework enables efficient computation of reaction equilibrium distributions, offering insights into thermodynamic versus kinetic control and supporting catalyst and process design. Overall, this thesis presents a unified approach that connects adsorption, reaction thermodynamics, and molecular design to support the rational development of shape-selective catalysts for the hydroisomerization reaction.

Samenvatting

Dit proefschrift presenteert een multiscale methode om de shape selectivity van zeolieten bij hydroisomerisatie van alkanen met lange ketens te onderzoeken, waarbij adsorptiethermodynamica, reactie-evenwichtsmodellering, lineaire regressie en machine learning worden gecombineerd. Het uiteindelijke doel is om te begrijpen hoe de moleculaire structuur en de topologie van zeoliet poriën de prestaties van zeoliet gebaseerde katalysatoren voor selectieve koolwaterstofupgrading bepalen. Rekening houden met oppervlakte heterogeniteiten van adsorptie isothermen is essentieel voor het nauwkeurig modelleren van katalytische en scheidingsprocessen, omdat competitieve adsorptie van alkaanisomeren op verschillende adsorptieplaatsen het reactie-evenwicht en de selectiviteit sterk beïnvloedt. Variaties in de affiniteit van de adsorptieplaatsen voor verschillende isomeren spelen een belangrijke rol bij het bepalen van de productverdelingen van hydroisomerisatie. In hoofdstuk 2 wordt een Segregated Explicit Isotherm (SEI) model ontwikkeld dat adsorbaatgrootte-afhankelijk gedrag in materialen met heterogene oppervlakteplaatsen modelleert. Het belangrijkste voordeel van SEI is dat uitsluitend expliciete vergelijkingen nodig zijn. Dit verbetert de numerieke stabiliteit en versnelt de berekeningen, vooral wanneer adsorptie evenwichten moeten worden berekend in zogenaamde breakthrough simulaties. In onze studie zijn SEI-gebaseerde breakthrough simulaties ongeveer drie keer sneller dan simulaties die gebruikmaakten van de Segregated Ideal Adsorbed Solution Theory (SIAST). De meest significante versnelling werd echter waargenomen bij een vergelijking van SIAST met IAST voor multisite isothermen: SIAST was ongeveer 20 keer sneller. Dit komt doordat IAST twee iteratieve stappen vereist: (1) het berekenen van de spreidingsdruk, en (2) het inverteren van de spreidingsdruk functie. SIAST gebruikt echter doorgaans expliciete uitdrukkingen voor de inverse van de spreidingsdruk op elke adsorptielocatie, waardoor iteraties niet nodig zijn. Om ervoor te zorgen dat deze geavanceerde isotherm- en breakthrough modellen toegankelijk en eenvoudig te gebruiken zijn in praktische proces simulaties, wordt in hoofdstuk 3 de open-source software RUPTURA geïntroduceerd. RUPTURA bevat drie kernmodules die veel worden gebruikt bij het modelleren van adsorptieprocessen: (1) simulaties van zogenaamde “step” en “pulse” breakthrough experimenten, (2) voorspelling van mengseladsorptie op basis van isothermen van zuivere componenten, en (3) het fitten van isotherm modellen aan ruwe adsorptiedata, berekend op basis van grand-canonical Monte Carlo simulaties of experimenteel gemeten. De RUPTURA software omvat een breed scala aan isothermmodellen, zoals Langmuir, BET, Henry, Freundlich, Sips,

Langmuir-Freundlich, Redlich-Peterson, Toth, Unilan, O'Brien & Myers, Asymptotic Temkin en Bingel & Walton en eveneens multisite uitbreidingen van deze isothermen en hun hybride formuleringen. Voor de voorspelling van mengseladsorptie zijn IAST, SIAST en expliciete Langmuir-gebaseerde benaderingen geïmplementeerd in RUPTURA. IAST-berekeningen worden snel en met machineprecisie uitgevoerd. Simulaties van breakthrough curves omvatten axiale dispersie en gebruiken het Linear Driving Force (LDF) model voor massaoverdracht, hetgeen een hoge numerieke stabiliteit garandeert via Strong-Stability Preserving Runge-Kutta (SSP-RK)-integratoren. RUPTURA is ontworpen voor snelle en flexibele evaluatie van scheidingsprocessen met een vast absorptie bed waarbij mengsels van koolwaterstoffen en andere moleculen betrokken zijn. Hoewel nauwkeurige adsorptie isothermen essentiële input zijn, is er meer nodig om de volledige thermodynamica van hydroisomerisatie te beschrijven. Het correct meenemen van de evenwichtsverdeling van chemische reacties is essentieel voor een volledig begrip van de productselectiviteit. De evenwichtsverdeling van de reactie bepaalt de thermodynamische grenzen van isomeeropbrengsten onder opgelegde bedrijfsomstandigheden. Nauwkeurige voorspelling van deze verdeling is essentieel voor het optimaliseren van de katalysatorselectiviteit en proceseffici$\ddot{e}$ntie. Hoofdstuk 4 richt zich op de berekening van de evenwichtsverdeling van de reactie van alkaanisomeren in zeolieten, waarbij het chemische evenwicht in de gasfase wordt gecombineerd met het fase-evenwicht tussen de gasfase en de geadsorbeerde fase. Dit laat zien hoe vormselectiviteit de evenwichtsverdeling van de producten in zeolietporiën beïnvloedt. De vormselectiviteit van zeolieten voor C_7 en C_8 hydroisomerisatie werd hiermee geëvalueerd. De evenwichtsverdelingen van de chemische reactie in de gasfase worden afgeleid van de Gibbs-vormingsenergie (bepaalde door de partitiefuncties van geïsoleerde moleculen), terwijl molfracties in de geadsorbeerde fase worden berekend met behulp van Henry coëfficiënten. In zeolieten met grote poriën (bijv. zeoliet van het FAU-type) heeft adsorptie een geringe invloed op het reactie evenwicht, terwijl in zeolieten met smalle poriën (bijv. zeolieten van het MTW- en MRE-type) de Henry coëfficiënten de productverdeling sterk beïnvloeden. In structuren met middelgrote poriën (bijvoorbeeld zeolieten van het type MFI en BEA) genieten monovertakte isomeren over het algemeen de voorkeur boven bi-vertakte isomeren. Meervoudig vertakte isomeren worden gestabiliseerd in kooiachtige structuren (bijvoorbeeld zeoliet van het type ITQ-29). Sterische hindering is in alle zeolieten nadelig voor sterk vertakte isomeren zoals 2,2,3-m-C_4. Bij hoge drukken en temperaturen (500-700 K) blijft het gasfase-evenwicht vrijwel ideaal, waardoor de relatieve bijdrage van adsorptie aan de vorming van de reactieproductverdeling wordt versterkt. Inzicht in deze trends bij chemisch evenwicht is cruciaal voor het optimaliseren van de selectiviteit van hydroisomerisatie en het sturen van het katalysatorontwerp. De vereiste thermochemische gegevens voor lang alkanen zijn helaas beperkt aanwezig in de literatuur. Toch zijn dergelijke gegevens nodig om de verdeling van reactieproducten

bij chemisch evenwicht nauwkeurig te bepalen, met name voor processen zoals hydroisomerisatie. Hoofdstuk 5 pakt deze lacune aan door een lineair regressiemodel te ontwikkelen op basis van bijdragen van tweede-orde groepen om belangrijke thermochemische eigenschappen te voorspellen: de vormingsenthalpie ($\Delta_f H_0$), de Gibbs vormings energie ($\Delta_f G_0$), de enthalpie ($H_0 - H_0(0\ \mathrm{K})$) en de Gibbs energie ($G_0 - H_0(0\ \mathrm{K})$). Het model overtreft de chemische nauwkeurigheid, met een maximale gemiddelde absolute fout van ca. 1 kJ/mol voor $\Delta_f G_0$, en presteert beter dan bestaande methoden zoals de groepsbijdragemethoden van Benson et al., Joback et al. en Constantinou-Gani. De temperatuurafhankelijkheid wordt beschreven met behulp van kwadratische functie voor de groepsbijdragen. De voorspelde thermochemische eigenschappenin de gasfase, in combinatie met Henry coëfficiënten berekend met de RASPA2 software, leveren reactie-evenwichtsverdelingen op die goed overeenkomen met de thermodynamische tabellen van Scott voor $H_0 - H_0(0\ \mathrm{K})$ en $G_0 - H_0(0\ \mathrm{K})$. Voor C_{10}- en C_{14}-isomeren in MTW-type zeoliet voorspelt het model dat isomeren met grote of dicht bij elkaar gelegen vertakkingen een lage selectiviteit hebben en een verwaarloosbare bijdrage leveren aan de verdeling van de reactieproducten bij chemisch evenwicht. Dit model biedt een robuuste en schaalbare benadering voor het evalueren van evenwichtsverdelingen van lange alkanen in zeolieten. Nauwkeurige kennis van thermochemische eigenschappen ($H_0 - H_0(0\ \mathrm{K})$ en $G_0 - H_0(0\ \mathrm{K})$) zijn ook nodig voor het berekenen van entropieën, die weer nodig zijn bij het berekenen van de zogenaamde exergievernietiging. Het evalueren van energie-efficiëntie via de tweede hoofdwet van de thermodynamica vereist het kwantificeren van exergievernietiging. Exergievernietiging is de nuttige arbeid die verloren gaat door onomkeerbare processen en is een belangrijke maatstaf voor het identificeren van inefficiënties in (chemische) processen. Hoofdstuk 6 breidt de thermodynamische analyse van hydroisomerisatie uit door absolute entropieën en de tweede hoofdwet efficiëntie te berekenen met behulp van een lineair regressiemodel gebaseerd op bijdragen van tweede-orde groepen. De berekende entropieën komen nauwkeurig overeen met de entropiewaarden verkregen uit de tabellen en experimenten van Scott, waarbij veranderingen worden waargenomen als gevolg van het aantal, type en de positie van vertakkingen in alkaanisomeren. Met behulp van dit model worden de entropieproductie en warmte-input geanalyseerd voor de hydroisomerisatie van C_7- en C_8 alkanen in verschillende zeolieten. De verdeling van reactieproducten varieert met de poriënstructuur van de zeoliet. Grote poriën in FAU- en ITQ-29-type zeolieten geven de voorkeur aan meervoudig vertakte isomeren, terwijl smalle poriën in MTW- en MRE-type zeolieten lineaire en mono vertakte isomeren prefereren. Deze verschillen leiden tot variaties in entropieproductie en exergievernietiging tussen verschillende zeolieten. Bij langere alkaanketens nemen zowel de enthalpie als de entropie toe door de grotere moleculaire complexiteit, hetgeen leidt tot een hogere warmte-input en entropieproductie. Bij hogere temperaturen nemen de moleculaire beweging de configurationele entropie

toe, waardoor de entropieproductie en het exergieverlies toenemen. Dit werk toont aan dat het lineaire regressie model absolute entropieën nauwkeurig kan voorspellen en het biedt waardevolle inzichten in de thermodynamische efficiëntie van hydroisomerisatie. Deze resultaten zijn niet alleen relevant voor katalysator- en procesoptimalisatie, maar kunnen ook nuttig zijn bij energieanalyse voor een breder scala aan koolwaterstof gebaseerde processen. Naast thermochemische eigenschappen in de gasfase zijn Henry coëfficiënten essentieel voor het bepalen van de evenwichtsverdeling van de chemische reactie bij hydroisomerisatie. Vanwege het grote aantal isomeren voor alkanen met lange ketens en de diversiteit aan zeolietstructuren is het berekenen van Henry coëfficiënten uitsluitend via moleculaire simulaties niet haalbaar. Vanwege deze complexiteit en benodigde rekentijd is een efficiënte voorspellende aanpak nodig als aanvulling op deze moleculaire simulaties. Hoofdstuk 7 presenteert een machine learning model voor het voorspellen van Henry coëfficiënten in eendimensionale zeolieten, waarbij descriptor-gebaseerde (TabPFN) en graph-gebaseerde (D-MPNN) modellen worden gecombineerd met een tool voor de geautomatiseerde enumeratie van alkaan isomeren. TabPFN presteert goed voor lineaire en matig vertakte isomeren, terwijl D-MPNN excelleert voor laag-adsorberende en sterk vertakte alkanen. Ondanks uitdagingen door zogenaamde activity cliffs en beperkte gegevens voor complexe isomeren, worden de prestaties verbeterd door active learning en oversampling. Bij hydroisomerisatie is het onderschatten van de Henry coëfficiënten van sterisch uitgesloten isomeren niet cruciaal, omdat de zeer lage Henry coëfficiënten wijzen op een verwaarloosbare adsorptie in de nauwe poriën van het zeoliet en daarom weinig invloed hebben op de evenwichtsverdeling van de chemische reactie. Door machine learning voorspelde Henry coëfficiënten te combineren met thermodynamische gegevens van een tweede-orde groepsbijdrage model kunnen we het chemisch evenwicht van hydroisomerisatie nauwkeurig voorspellen. Dit kan verder inzicht bieden in thermodynamische versus kinetische controle en het detailleerde ontwerp van zowel katalysator als proces. Hierdoor presenteert dit proefschrift een uniforme aanpak die adsorptie, reactiethermodynamica en moleculair ontwerp verbindt om de ontwikkeling van vorm selectieve katalysatoren voor hydroisomerisatie te ondersteunen.

Curriculum Vitæ

Shrinjay Sharma was born on 5 January 1996 in Guwahati, India. In 2014, he began his B-Tech in Mechanical Engineering at the National Institute of Technology Silchar, India. After graduating in 2018, he commenced his MSc studies in Mechanical Engineering (specialization in Energy and Process Technology) at Delft University of Technology, the Netherlands. For his MSc thesis, supervised by Prof.dr. Earl Goetheer (TNO and TU Delft), he investigated the Dynamic Electrochemical Promotion of Catalysis in the context of Fischer–Tropsch synthesis. His research focused on modeling the effect of electrochemically controlled transport of promoter ions on product selectivity, offering mechanistic insights into how external fields influence catalytic performance. In 2021, he continued his academic journey as a PhD candidate at the Process and Energy Department of TU Delft, under the supervision of Prof.dr.ir. Thijs J.H. Vlugt (Delft University of Technology), Dr. David Dubbeldam (University of Amsterdam), and Prof.dr. Sofia Calero (Eindhoven University of Technology), in collaboration with Shell's Energy Transition Campus Amsterdam. His PhD research centres on understanding the shape-selectivity effects of zeolites on the hydroisomerization of long-chain alkanes. To study this phenomenon, he uses a multiscale modeling approach combining molecular simulations, reaction equilibrium modeling, breakthrough curve analysis, and machine learning. His work uses Monte Carlo simulations to compute Henry coefficients and adsorption isotherms of alkane isomers in zeolites. These data are fed into thermodynamic models to predict equilibrium compositions and selectivity trends. He uses linear regression and advanced machine learning methods, including graph neural networks, to predict thermochemical properties of long-chain alkanes, which is otherwise a computationally expensive task for millions of isomers in a large number of zeolites. This dissertation presents the findings of his research.

List of Publications

1. S. Sharma, M. S. Rigutto, R. Baur, U. Agarwal, E. Zuidema, S. R. G. Balestra, S. Calero, D. Dubbeldam, and T. J. H. Vlugt. "Modelling of adsorbate-size dependent explicit isotherms using a segregated approach to account for surface heterogeneities". *Mol. Phys.* 121 (2023), e2183721

2. S. Sharma, S. R. Balestra, R. Baur, U. Agarwal, E. Zuidema, M. S. Rigutto, S. Calero, T. J. H. Vlugt, and D. Dubbeldam. "RUPTURA: simulation code for breakthrough, ideal adsorption solution theory computations, and fitting of isotherm models". *Mol. Simul.* 49 (2023), pp. 893–953

3. S. Sharma, M. S. Rigutto, E. Zuidema, U. Agarwal, R. Baur, D. Dubbeldam, and T. J. H. Vlugt. "Understanding shape selectivity effects of hydroisomerization using a reaction equilibrium model". *J. Chem. Phys.* 160 (2024), p. 214708

4. S. Sharma, J. J. Sleijfer, J. op de Beek, S. van der Zeeuw, D. Zorzos, S. Lasala, M. S. Rigutto, E. Zuidema, U. Agarwal, R. Baur, S. Calero, D. Dubbeldam, and T. J. H. Vlugt. "Prediction of Thermochemical Properties of Long-Chain Alkanes Using Linear Regression: Application to Hydroisomerization". *J. Phys. Chem. B* 128 (2024), pp. 9619–9629

5. S. Sharma, R. Baur, M. S. Rigutto, E. Zuidema, U. Agarwal, S. Calero, D. Dubbeldam, and T. J. H. Vlugt. "Computing Entropy for Long-Chain Alkanes Using Linear Regression: Application to Hydroisomerization". *Entropy* 26 (2024), p. 1120

6. S. Sharma[1], P. Yang [1], Y. Liu, K. Rossi, P. Bai, M. S. Rigutto, E. Zuidema, U. Agarwal, R. Baur, S. Calero, D. Dubbeldam, and T. J. H. Vlugt. "Machine Learning-Based Predictions of Henry Coefficients for Long Chain Alkanes in One Dimensional Zeolites: Application to Hydroisomerization". *J. Phys. Chem. C* (2025). Manuscript accepted.

[1]These authors contributed equally to this work.

Publications not included in this thesis:

1. U. Agarwal, M. S. Rigutto, E. Zuidema, A. P. J. Jansen, A. Poursaeidesfahani, S. Sharma, D. Dubbeldam, and T. J. H. Vlugt. “Kinetics of zeolite-catalyzed heptane hydroisomerization and hydrocracking with CBMC-modeled adsorption terms: Zeolite Beta as a large pore base case”. *J. Catal.* 415 (2022), pp. 37–50

2. Y. A. Ran, S. Sharma, S. R. G. Balestra, Z. Li, S. Calero, T. J. H. Vlugt, R. Q. Snurr, and D. Dubbeldam. “RASPA3: A Monte Carlo code for computing adsorption and diffusion in nanoporous materials and thermodynamics properties of fluids”. *J. Chem. Phys.* 161 (2024), p. 114106

3. S. Gooijer, S. Capelo, S. Sharma, S. Giancola, J. R. Galan-Mascaros, T. J. H. Vlugt, D. Dubbeldam, J. M. Vicent-Luna, and S. Calero. “TAMOF-1 for Capture and Separation of the main Flue Gas Components”. *J. Mater. Chem. A* (2025), pp. 16879–16892

4. Y. A. Ran, J. Tapia, S. Sharma, P. Bai, S. Calero, T. J. H. Vlugt, and D. Dubbeldam. “Implementation of energy and force grids in molecular simulations of porous materials”. *Mol. Phys.* (2025), e2542486

5. Z. Li, L. Constantinou, R. Baur, D. Dubbeldam, S. Calero, S. Sharma, M. Rigutto, P. Dey, and T. J. H. Vlugt. “Review of Group Contribution Methods for Prediction of Thermodynamic Properties of Long-chain Hydrocarbons”. *Mol. Phys.* (2025). Manuscript accepted.

6. Z. Li, L. Constantinou, R. Baur, D. Dubbeldam, S. Calero, S. Sharma, M. Rigutto, P. Dey, and T. J. H. Vlugt. “Second-order Group Contribution Method for T_c, P_c, ω, ΔG_f^0, ΔH_f^0 and Liquid Densities of Linear and Branched Alkanes”. *Mol. Phys.* (2025). Manuscript accepted.

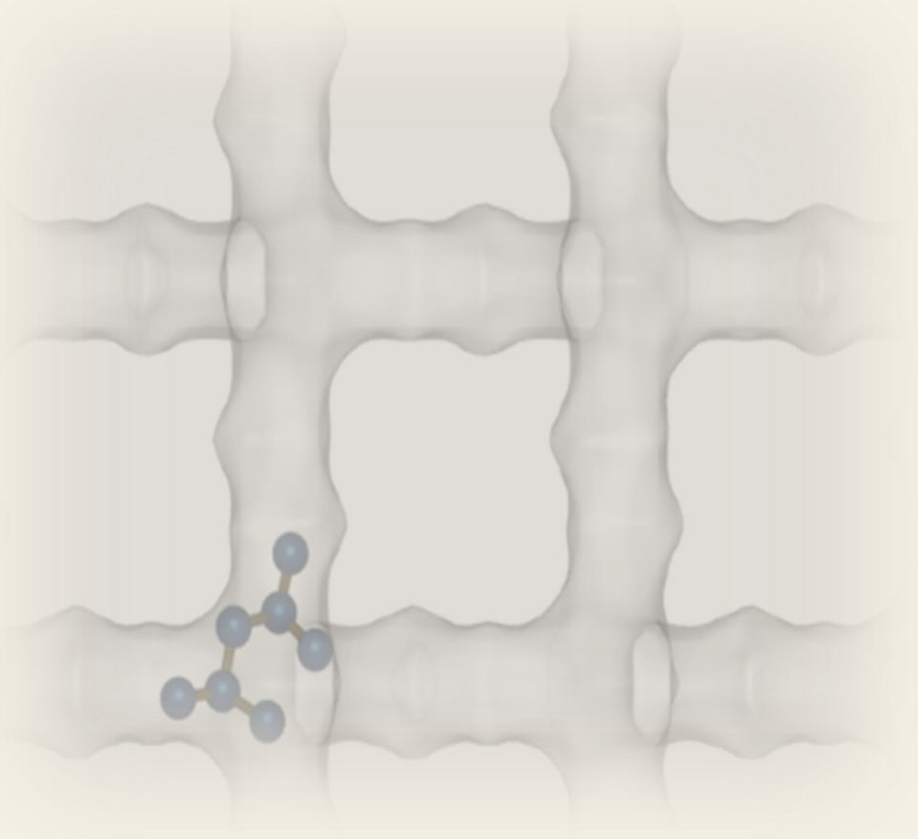
Acknowledgements

I am truly grateful to my promoters Prof.dr.ir. Thijs J.H. Vlugt, Dr. David Dubbeldam, and Prof.dr. Sofia Calero for their guidance, encouragement, and steady support throughout this journey. Their mentorship has shaped both my scientific thinking and personal development in profound ways. Whether through brainstorming sessions, thoughtful feedback, or simply creating space for my ideas to grow, they have consistently provided opportunities for collaboration and recognition. I deeply appreciate the trust they placed in me and the freedom they gave me to explore and contribute. I have also been fortunate to collaborate with our industry partners from Shell, Dr. Marcello Rigutto, Dr. Richard Baur, Dr. Erik Zuidema, and Dr. Umag Agarwal. Our interactions offered valuable perspective on the relevance and application of this research. I also extend my thanks to Dr. Salvador Balestra, Ping Yang, Dr. Kevin Rossi, Dr. Peng Bai, Yachan Liu, Youri Ran, Sussanne Gooijer, Joseph Tapia, Dr. Jose M. Vicent-Luna, and Dr. Silvia Lasala for their scientific input and collaboration during this project. Throughout my PhD, I also had the opportunity to work with a group of bright bachelor and master students, Josh J. Sleijfer, Jeroen Op de Beek, Stach van der Zeeuw, Daniil Zorzos, Ziyan Li, and Erik Johnsson. Each project was unique, and I learned something new from every collaboration. Their contributions added real depth to this work, and I am thankful for their enthusiasm and dedication. This project was made possible through the support of the ARC-CBBC consortium and NWO. I am grateful to the ARC-CBBC support office, especially Hannah Thuijs, Anita den Heijer, and former members Masja Spijkstra, Jeroen Meijer, and Marijke Badings, for their invaluable assistance. The computational work relied on the generous resources provided by the Snellius and DelftBlue supercomputers, and I appreciate the access facilitated by the Delft High Performance Computing Centre.

A special thanks to my office mates and close colleagues, Georgia, Mert, Darshan, Jelle, Parsa, Panji, and Sebastiaan, whose companionship transformed daily routines, coffee breaks, walks, and travels into lasting memories. My friendly darts and pool rivalry with Georgia brought much-needed fun to everyday life. I thoroughly enjoyed making liquid nitrogen ice cream with Mert and Mengmeng, traveling with Darshan to New York City, going on a gator safari with Jelle in Orlando, and hiking the Grand Canyon with Georgia and Panji. Thank you, each one of you, for always keeping the mood light and the banter sharp. I am deeply grateful to all my colleagues from the Process & Energy department who made this journey more engaging,

collaborative, and enjoyable. I want to acknowledge Dr.ir. Mahinder Ramdin, Dr.ir. Tim Nissjen, and Dr. Othonas Moultos for the engaging discussions that shaped both my thinking and my work. Our discussions were always thought-provoking and often exactly what I needed.

I have been incredibly fortunate to have friends who made this country feel like a home away from home. I am deeply grateful to Jayant, Deep, Shruti, Shaheen, Maitry, Moorchana, Iftikhar, Shubhashis, Dwijashish, Barsha, Prantik, Mridul, Nitish, Ajay, Keerthi, Sakina, Anand, Avishek, Nishant, Varun, Nivya, Saloni, and many more whose names would make the list wonderfully long. Their warmth, kindness, and shared moments transformed the unfamiliar into something comforting and familiar. From celebrations to quiet dinners, from long conversations to simple companionship, they have been a steady anchor through the highs and lows of this journey. These memories are ones I will carry with me long after this chapter comes to a close. Lastly, I want to express my heartfelt thanks to my parents, grandparents, uncles, aunts, and cousins. Their unwavering support and encouragement have been the constant behind every step I took. This journey would not have been possible without the love and belief they have always shown me.



TU Delft