



**HELLENIC REPUBLIC**  
**UNIVERSITY OF IOANNINA**  
**SCHOOL OF ENGINEERING**  
**DEPARTMENT OF MATERIALS SCIENCE AND ENGINEERING**

**MODELING SORPTION AND DIFFUSION OF ALKANES, ALKENES IN POROUS  
MATERIALS USING MOLECULAR SIMULATIONS AND ARTIFICIAL NEURAL  
NETWORKS**

**GKOURRAS ARSENIOS**

**THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY**  
**IOANNINA, 2026**





**ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ  
ΠΑΝΕΠΙΣΤΗΜΙΟ ΙΩΑΝΝΙΝΩΝ  
ΠΟΛΥΤΕΧΝΙΚΗ ΣΧΟΛΗ  
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΕΠΙΣΤΗΜΗΣ ΥΛΙΚΩΝ**

**ΜΟΝΤΕΛΟΠΟΙΗΣΗ ΤΗΣ ΡΟΦΗΣΗΣ ΚΑΙ ΤΗΣ ΔΙΑΧΥΣΗΣ ΑΛΚΑΝΙΩΝ,  
ΑΛΚΕΝΙΩΝ ΣΕ ΠΟΡΩΔΗ ΥΛΙΚΑ ΜΕ ΧΡΗΣΗ ΜΟΡΙΑΚΩΝ ΠΡΟΣΟΜΟΙΩΣΕΩΝ  
ΚΑΙ ΤΕΧΝΗΤΩΝ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ**

**ΓΚΟΥΡΡΑΣ ΑΡΣΕΝΙΟΣ**

**ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ  
ΙΩΑΝΝΙΝΑ, 2026**



*«Η έγκριση της διδακτορικής διατριβής από το Τμήμα Μηχανικών Επιστήμης Υλικών της Πολυτεχνικής Σχολής του Πανεπιστημίου Ιωαννίνων δεν υποδηλώνει αποδοχή των γνώμων του συγγραφέα Ν. 5343/32, άρθρο 202, παράγραφος 2».*

*Date of application of Mr. Arsenios Gourras: 22/10/2020*

*Date of appointment of the Three-Member Advisory Committee: 10/11/2020*

*Supervisor: Leonidas Gergidis, Professor, Department of Materials Science and Engineering, School of Engineering, University of Ioannina*

*Members of the Three-Member Advisory Committee:*

*Alkiviadis Paipetis, Professor, Department of Materials Science and Engineering, School of Engineering, University of Ioannina*

*Antonios Charalampopoulos, Professor, School of Applied Mathematical and Physical Sciences, National Technical University of Athens*

*Date of topic assignment: 10/11/2020*

*MODELING PHYSICAL-CHEMICAL PROCESSES IN POROUS MATERIALS: FROM MOLECULAR SIMULATIONS AND CONTINUUM SIMULATIONS TO ARTIFICIAL INTELLIGENCE*

*Date of topic modification: 25/07/2024*

*MODELING SORPTION AND DIFFUSION OF ALKANES AND ALKENES IN POROUS MATERIALS USING MOLECULAR SIMULATIONS AND ARTIFICIAL NEURAL NETWORKS*

*Appointment of the Seven-Member Examination Committee: 8/10/2025*

*1. Leonidas Gergidis, Professor, Department of Materials Science and Engineering, School of Engineering, University of Ioannina*

*2. Alkiviadis Paipetis, Professor, Department of Materials Science and Engineering, School of Engineering, University of Ioannina*

*3. Antonios Charalampopoulos, Professor, School of Applied Mathematical and Physical Sciences, National Technical University of Athens*

*4. Dimitrios Fotiadis, Professor, Department of Materials Science and Engineering, School of Engineering, University of Ioannina*

*5. Marianthi-Nektaria Barkoula, Professor, Department of Materials Science and Engineering, School of Engineering, University of Ioannina*

*6. Michail Karakasidis, Professor, Department of Materials Science and Engineering, School of Engineering, University of Ioannina*

*7. Nikolaos Kalfagiannis, Assistant Professor, Department of Materials Science and Engineering, School of Engineering, University of Ioannina*

*Approval of the Doctoral Dissertation with grade "Excellent" on 25/02/2026*

*Department Chair*

*Simeon Agathopoulos  
Professor*



*Secretary of the Department*

*Maria Kontou*

Ημερομηνία αίτησης του κ.: Αρσένιου Γκούρρα 22/10/2020

Ημερομηνία ορισμού Τριμελούς Συμβουλευτικής Επιτροπής: 10/11/2020

Επιβλέπων: Λεωνίδα Γεργίδη, Καθηγητής του Τμήματος Μηχανικών Επιστήμης Υλικών της Πολυτεχνικής Σχολής του Πανεπιστημίου Ιωαννίνων

Μέλη Τριμελούς Συμβουλευτικής Επιτροπής:

Αλκιβιάδης Παϊπέτης, Καθηγητής του Τμήματος Μηχανικών Επιστήμης Υλικών της Πολυτεχνικής Σχολής του Πανεπιστημίου Ιωαννίνων

Αντώνιος Χαραλαμπίδης, Καθηγητής της Σχολής Εφαρμοσμένων Μαθηματικών και Φυσικών Επιστημών του Ε.Μ.Π.

Ημερομηνία ορισμού θέματος: 10/11/2020

ΜΟΝΤΕΛΟΠΟΙΗΣΗ ΦΥΣΙΚΩΝ-ΧΗΜΙΚΩΝ ΔΙΕΡΓΑΣΙΩΝ ΣΕ ΠΟΡΩΔΗ ΥΛΙΚΑ: ΑΠΟ ΤΙΣ ΜΟΡΙΑΚΕΣ ΠΡΟΣΟΜΟΙΩΣΕΙΣ ΚΑΙ ΤΙΣ ΠΡΟΣΟΜΟΙΩΣΕΙΣ ΣΥΝΕΧΟΥΣ ΣΤΗΝ ΤΕΧΝΙΚΗ ΝΟΗΜΟΣΥΝΗ

Ημερομηνία τροποποίησης θέματος: 25/07/2024

MODELING SORPTION AND DIFFUSION OF ALKANES, ALKENES IN POROUS MATERIALS USING MOLECULAR SIMULATIONS AND ARTIFICIAL NEURAL NETWORKS

ΔΙΟΡΙΣΜΟΣ ΕΠΤΑΜΕΛΟΥΣ ΕΞΕΤΑΣΤΙΚΗΣ ΕΠΙΤΡΟΠΗΣ : 8/10/2025

1. Λεωνίδα Γεργίδη, Καθηγητής του Τμήματος Μηχανικών Επιστήμης Υλικών της Πολυτεχνικής Σχολής του Πανεπιστημίου Ιωαννίνων,

2. Αλκιβιάδης Παϊπέτης, Καθηγητής του Τμήματος Μηχανικών Επιστήμης Υλικών της Πολυτεχνικής Σχολής του Πανεπιστημίου Ιωαννίνων,

3. Αντώνιος Χαραλαμπίδης, Καθηγητής της Σχολής Εφαρμοσμένων Μαθηματικών και Φυσικών Επιστημών του Ε.Μ.Π.

4. Δημήτριος Φωτιάδης, Καθηγητής του Τμήματος Μηχανικών Επιστήμης Υλικών της Πολυτεχνικής Σχολής του Πανεπιστημίου Ιωαννίνων,

5. Μαριάνθη-Νεκταρία Μάρκουλα, Καθηγήτρια του Τμήματος Μηχανικών Επιστήμης Υλικών της Πολυτεχνικής Σχολής του Πανεπιστημίου Ιωαννίνων

6. Μιχαήλ Καρακασίδης, Καθηγητής του Τμήματος Μηχανικών Επιστήμης Υλικών της Πολυτεχνικής Σχολής του Πανεπιστημίου Ιωαννίνων

7. Νικόλαος Καλφαγιάννης, Επίκουρος Καθηγητής του Τμήματος Μηχανικών Επιστήμης Υλικών της Πολυτεχνικής Σχολής του Πανεπιστημίου Ιωαννίνων.

Έγκριση Διδακτορικής Διατριβής με βαθμό «ΑΡΙΣΤΑ» στις 25/02/2026

Ο Πρόεδρος του Τμήματος



Συμεών Αγαθόπουλος  
Καθηγητής



Η Γραμματέας του Τμήματος



Μαρία Κόντου



*To my  
sister Freda*

*Στην αδερφή  
μου Φρέντα*

## Abstract

Metal organic frameworks (MOFs) constitute a versatile class of porous materials with high structural tunability, offering significant potential across gas storage, separation, catalysis, and pharmaceutical applications. The rapid growth in the number of available MOF structures renders exhaustive experimental evaluation impractical, motivating the increasing adoption of molecular simulations and data-driven methodologies. In this thesis, adsorption and diffusion phenomena of light alkanes, alkenes, and their mixtures in representative MOFs are investigated through Grand Canonical Monte Carlo (GCMC) and Molecular Dynamics (MD) simulations. Sorption isotherms, selectivities, density profiles, and self-diffusion coefficients are computed for different MOFs over wide temperature and pressure ranges, and the resulting predictions are compared with available experimental measurements. Different MOFs have been studied: the Zeolitic Imidazolate Framework (ZIF-8) and the NIIC-20-Bu novel mesoporous material, where NIIC stands for Nikolaev Institute of Inorganic Chemistry, in which the NIIC-20-Bu's synthesis took place. Beyond conventional simulation approaches, artificial neural networks (ANNs) models are developed and trained on combined simulation and experimental datasets to accurately predict adsorption and diffusion behavior with minimal computational cost. This integrated simulation, data approach provides a valuable tool to utilize experimental and simulation data for fast and accurate predictions. Specifically, the sorption and diffusion behavior of small hydrocarbons in ZIF-8 was investigated using GCMC and canonical ensemble (NVT) molecular dynamics simulations. These studies focused on methane, ethane, and ethene, both as pure components and in ternary mixtures across various temperatures. Additional simulations were performed for propane, propene, butane, and butene. To enhance predictive capabilities, ANNs were trained on both experimental and simulation data to estimate sorption isotherms and self-diffusion coefficients. For NIIC-20-Bu, adsorption characteristics of methane, ethane, and their equimolar mixture were explored through GCMC simulations and ANN predictive modeling. Adsorption isotherms were obtained at different temperatures offering insights into the of pure and mixed components within NIIC-20-Bu. Finally, a comparison was conducted to evaluate the performance of NIIC-20-Bu in relation to other popular porous materials, which demonstrated its sorptive superiority.

The outline of this thesis can be summarized below:

**Chapter 1:** Introduction to MOFs (Metal-Organic Frameworks), providing a foundational understanding of MOFs, their structure, properties, and relevance across various applications, particularly in gas storage, separation, and catalysis.

**Chapter 2:** Literature Overview, State of the Art and Thesis Contribution goes deep into existing research, highlighting the progress and challenges in the field of MOFs and offering insights into gaps this thesis aims to address. The chapter is divided into two different approaches, one being the existing framework of experimental work and classical simulations and the other being the incorporation of Machine learning and artificial intelligence into different modelling approaches,

**Chapter 3:** Computational Modeling introduces the computational techniques employed in this study, including Molecular Dynamics (MD), Monte Carlo (MC), and Artificial Neural Networks (ANN), providing the necessary tools to simulate and analyze the behavior of MOFs.

**Chapter 4:** Molecular and Artificial Neural Networks Modeling of Sorption and Diffusion of Small Alkanes, Alkenes and Their Ternary Mixtures in ZIF-8 at Different Temperatures focuses on ZIF-8, a zeolitic imidazolate framework, examining its performance and potential applications through molecular simulations

**Chapter 5:** Adsorption of Methane, Ethane and their Equimolar Mixture in NIIC-20-Bu Metal Organic Framework from Grand Canonical Monte Carlo Simulations and Artificial Neural Networks investigates the properties and performance of the NIIC-20-Bu MOF structure.

**Chapter 6:** Conclusions and Future Work, summarizing key findings and proposing directions for further research to enhance MOF simulations and applications.

## Περίληψη

Τα μέταλλο-οργανικά πλέγματα (MOFs) αποτελούν μια εκτεταμένη κατηγορία πορωδών υλικών με υψηλή δυνατότητα προσαρμογής της δομής τους, προσφέροντας σημαντικές δυνατότητες σε εφαρμογές αποθήκευσης αερίων, διαχωρισμού και κατάλυσης. Η ταχεία αύξηση στον αριθμό των διαθέσιμων δομών MOF καθιστά την πειραματική αξιολόγησή τους μη πρακτική, γεγονός που ενισχύει την αυξανόμενη ανάγκη για υιοθέτηση μοριακών προσομοιώσεων και τεχνικών τεχνητής νοημοσύνης. Στην παρούσα διδακτορική διατριβή διερευνώνται, μέσω μοριακών προσομοιώσεων Monte Carlo (MC) και Μοριακής Δυναμικής (MD), η ρόφηση και η διάχυση μικρών αλκανίων, αλκενίων και των μειγμάτων τους σε αντιπροσωπευτικές MOF δομές. Υπολογίζονται ισόθερμοι ρόφησης, εκλεκτικότητες, πυκνότητες και θέσεις εντοπισμού καθώς και συντελεστές αυτοδιάχυσης των προαναφερθέντων υδρογονανθράκων για διαφορετικά MOFs σε συγκεκριμένο εύρος θερμοκρασιών, πιέσεων και οι προκύπτουσες προβλέψεις συγκρίνονται με διαθέσιμες πειραματικές μετρήσεις. Στην παρούσα διατριβή μελετήθηκαν δύο διαφορετικά MOFs με πρώτο το Zeolitic Imidazolate Framework (ZIF-8), και δεύτερο το NiIC-20-Bu, το οποίο σχεδιάστηκε και συντετέθη στο Nikolaev Institute of Inorganic Chemistry στο Νοβοσιμπίρσκ της Ρωσίας. Πέρα από τις αριθμητικές μελέτες με χρήση τεχνικών μοριακής προσομοίωσης, αναπτύσσονται μοντέλα τεχνητών νευρωνικών δικτύων (ANNs) και εκπαιδεύονται σε διαθέσιμα πειραματικά και σε δεδομένα μοριακών προσομοιώσεων για την ακριβή πρόβλεψη των ισοθέμων ρόφησης και διάχυσης με ελάχιστο υπολογιστικό κόστος. Αυτή η ολοκληρωμένη προσέγγιση μοριακής προσομοίωσης και τεχνητών νευρωνικών δικτύων παρέχει ένα πολύτιμο εργαλείο για γρήγορες και ακριβείς προβλέψεις σχέσεων δομής-ιδιοτήτων. Συγκεκριμένα στην παρούσα διδακτορική διατριβή η συμπεριφορά της ρόφησης και διάχυσης μικρών υδρογονανθράκων (αλκανίων, αλκενίων) στο ZIF-8 διερευνήθηκε χρησιμοποιώντας προσομοιώσεις Monte Carlo στο μεγαλοκανονικό στατιστικό σύνολο (GCMC) και προσομοιώσεις μοριακής δυναμικής στο κανονικό στατιστικό σύνολο (NVT-MD). Για τη βελτίωση και επιτάχυνση των προβλέψεων, τα μοντέλα τεχνητών νευρωνικών δικτύων που αναπτύχθηκαν εκπαιδεύτηκαν με πειραματικά και δεδομένα προσομοιώσεων για την ακριβή πρόβλεψη ισοθέμων ρόφησης και των συντελεστών αυτοδιάχυσης σε πλήθος θερμοκρασιών. Για το NiIC-20-Bu, η ρόφηση του μεθανίου, του αιθανίου και του ισομοριακού μίγματός τους διερευνήθηκε μέσω προσομοιώσεων GCMC και μοντέλων τεχνητών νευρωνικών δικτύων. Οι ισόθερμοι ρόφησης επίσης υπολογίστηκαν σε πολλαπλές θερμοκρασίες, προσφέροντας πληροφορίες για τα καθαρά συστατικά αλλά και το μίγμα του

μεθανίου-αιθανίου εντός του NIIC-20-Bu. Τέλος, πραγματοποιήθηκε μια σύγκριση για την αξιολόγηση της απόδοσης του NIIC-20-Bu σε σχέση με άλλα δημοφιλή πορώδη υλικά που κατέδειξε την ροφητική του υπεροχή.

## Acknowledgments

First and foremost, I would like to express my deepest gratitude to my supervisor, Leonidas N. Gergidis, whose guidance, encouragement, and unwavering support have been instrumental throughout my PhD journey. Professor Gergidis has been more than a mentor, providing insightful feedback, fostering my academic growth, and always pushing me to reach higher standards. Your patience and dedication have been invaluable, and I am immensely thankful for the opportunities you have provided me to expand my knowledge and skills.

I would also like to extend my heartfelt thanks to the members of my thesis committee, Alkiviadis Paipetis, Professor at the Dept. of Materials Science and Engineering and Antonios Charalambopoulos Professor in School of Applied Mathematical and Physical Sciences, for their valuable time, expertise, and constructive feedback, which have greatly improved the quality of my research.

A special thank you to my colleagues and fellow researchers in the Multidisciplinary Multiscale Multiphysics Simulation Team, in particular Andreas Kalogirou and Dimitris Iliopoulos whose collaborations and intellectual exchanges have made this journey more enriching and enjoyable. I am grateful for the collaborative environment and the friendships that have formed along the way.

I would also like to thank Professor Vladimir Fedin (Managing Director at the Russian Academy of Sciences, RAS) and researcher Dr. Denis Samsonenko at the Nikolaev Institute of Inorganic Chemistry of the Siberian Branch of the RAS, Novosibirsk, Russia, for providing the crystal structure of NIIC-20-Bu and for the excellent collaboration.

On a personal note, I wish to thank my family and friends for their unwavering support and encouragement. Especially to my sister Freda I would like to acknowledge her support throughout all my academic journey, your support and sacrifices have given me the strength to continue, without you all of this would never have been possible. To my friends, thank you for your understanding and patience during the demanding times of my research.

Thank you all for being part of this incredible journey.



# Contents

<b>Chapter 1: Introduction to Metal Organic Frameworks .....</b>	<b>1</b>
<b>1.1Metal Organic Frameworks .....</b>	<b>1</b>
<b>1.2Structure of MOFS.....</b>	<b>1</b>
<b>1.3Applications of Metal Organic Framework.....</b>	<b>6</b>
<b>1.3.1 Biological Compounds .....</b>	<b>6</b>
<b>1.3.2 MOFs and Antibiotics .....</b>	<b>7</b>
<b>1.3.3 MOFs and Toxic Pollution .....</b>	<b>8</b>
<b>1.3.4 Gas adsorption in MOFs.....</b>	<b>10</b>
<b>1.3.5 MOFs and Catalysts.....</b>	<b>12</b>
<b>1.4 Types of MOFs .....</b>	<b>15</b>
<b>1.4.1Zeolitic imidazolate frameworks (ZIFs) .....</b>	<b>16</b>
<b>1.4.1.1 Zeolitic imidazolate frameworks-8 (ZIF-8) .....</b>	<b>16</b>
<b>1.4.2 MOFs structures from Nikolaev Institute of Inorganic Chemistry (NIIC) .....</b>	<b>17</b>
<b>1.4.2.1 NIIC 20 Structure.....</b>	<b>19</b>
<b>1.4.2.1 NIIC 20 Bu .....</b>	<b>20</b>
<b>1.5 Molecular Simulations and Machine Learning methods for predicting sorption and diffusion of organic molecules in porous materials.....</b>	<b>21</b>
<b>1.6 Motivation .....</b>	<b>24</b>
<b>References .....</b>	<b>32</b>
<b>Chapter 2: Literature Overview, State of the Art and Thesis Contribution.....</b>	<b>41</b>
<b>2.1 Introduction.....</b>	<b>41</b>
<b>2.2 Classical Simulation techniques .....</b>	<b>42</b>
<b>2.3 Machine Learning approach to Classical Molecular simulations .....</b>	<b>48</b>
<b>2.4 Contribution of this thesis .....</b>	<b>51</b>
<b>References .....</b>	<b>53</b>
<b>Chapter 3: Molecular Modelling .....</b>	<b>57</b>

<b>3.1 Introduction .....</b>	<b>57</b>
<b>3.2 Role of Statistical Mechanics in Materials Design .....</b>	<b>57</b>
<b>3.3 Molecular Dynamics .....</b>	<b>59</b>
<b>3.3.2 Molecular Dynamics Algorithm .....</b>	<b>66</b>
<b>3.3.3 Statistical Ensembles .....</b>	<b>66</b>
<b>3.3.4 Force Field.....</b>	<b>67</b>
<b>3.4 Monte Carlo .....</b>	<b>70</b>
<b>3.4.2 Monte Carlo Algorithm.....</b>	<b>71</b>
<b>3.4.3 Grand Canonical Monte Carlo.....</b>	<b>73</b>
<b>3.5 Artificial Neural Networks (ANNs) .....</b>	<b>74</b>
<b>3.5.1 Artificial Neural Networks architecture.....</b>	<b>76</b>
<b>3.5.2 Activation Functions.....</b>	<b>77</b>
<b>3.5.3 Back propagation Algorithm .....</b>	<b>78</b>
<b>3.5.4 Cross Validation .....</b>	<b>79</b>
<b>References.....</b>	<b>82</b>
<b>Chapter 4: Molecular and Artificial Neural Networks Modeling of Sorption and Diffusion of Small Alkanes, Alkenes and Their Ternary Mixtures in ZIF-8 at Different Temperatures .....</b>	<b>85</b>
<b>4.1 Preface.....</b>	<b>85</b>
<b>4.2 Zeolitic Imidazolate Framework 8.....</b>	<b>86</b>
<b>4.3.1 GCMC and NVT-MD Molecular Simulations .....</b>	<b>88</b>
<b>4.4 Results .....</b>	<b>95</b>
<b>4.4.1 Self-Diffusivity Coefficients for Pure Components and Mixtures Obtained from NVT-MD, ANNs, and MAS PFG NMR Experiments.....</b>	<b>96</b>
<b>4.4.2 Self-Diffusivities for Ternary Mixtures of Methane–Ethane–Ethene.....</b>	<b>100</b>
<b>4.5 Sorption Isotherms of Small Alkanes, Alkenes, and Their Mixtures in ZIF-8.....</b>	<b>103</b>
<b>4.5.1 Sorption Isotherms of Pure Components. ....</b>	<b>104</b>
<b>4.5.2 Sorption Isotherms of Methane-Ethane-Ethene Mixture .....</b>	<b>109</b>

4.6 Conclusions .....	113
References .....	114
<b>Chapter 5 Adsorption of Methane, Ethane and their Equimolar Mixture in NIIC-20-Bu Metal Organic Framework from Grand Canonical Monte Carlo Simulations and Artificial Neural Networks.....</b>	<b>121</b>
5.1 Preface .....	121
5.2 Introduction .....	122
5.3 Molecular Modeling and Simulations .....	124
5.3.1 Molecular Model.....	124
5.3.2 ANNs Methodology.....	127
5.4 Results .....	130
5.4.1 Methane Sorption .....	130
5.4.2 Methane’s sorption in NIIC-20-Bu at different temperatures-Sorption ratio .....	133
5.4.3 Methane’s Sorption in NIIC-20-Bu compared to other porous materials .....	135
5.4.4 Ethane Sorption .....	136
5.4.4 Ethane’s Sorption in NIIC-20-Bu compared to other porous materials .....	138
5.4.5 Methane-Ethane Mixture Sorption.....	139
5.4.6 Methane-Ethane Mixture Sorption selectivity.....	141
5.5 Conclusions .....	143
References .....	145
<b>Chapter 6 Conclusions and Future work.....</b>	<b>151</b>
6.1 Conclusions .....	151
6.1.1 ZIF-8 .....	151
6.1.2 NIIC-20 Bu .....	154
6.2 Future work .....	157
6.2.1 Classical Molecular simulations .....	158
6.2.2 Machine Learning integration.....	159

6.2.3 Continuous mechanics integration.....	161
6.3 Outlook.....	163
References .....	165
Appendix A.....	167
Appendix B.....	171
Appendix C.....	175
Appendix D.....	183
Curriculum vitae (CV) .....	187

## Chapter 1:

Figure 1.1: Typical examples of organic linkers used in the synthesis of MOFs [6].....	2
Figure 1.2: Some common SBUs used in MOF construction (left) [10] / influence of PBUs on the metal–organic framework (right) [7] .....	3
Figure 1.3: Breathing mode of Metal organic Frameworks [8]. .....	3
Figure 1.4: Applications of Metal Organic Frameworks [30].....	5
Figure 1.5: Metal organics Frameworks in Antibiotics [12]. .....	8
Figure 1.6: Metal Organic Frameworks in toxic pollution applications [39]. .....	9
Figure 1.7: CO <sub>2</sub> Capture in Metal Organic Frameworks [41] .....	10
Figure 1.8: Catalysis in Metal Organic Frameworks [51]. .....	13
Figure 1.9: Typical Metal Organics frameworks with their chemical formula and their abbreviations [55]. .....	14
Figure 1.10: Representative types of metal organic networks synthesized by various research groups [57]. .....	15
Figure 1.11: The sodalite topology (left) and narrow six-membered-ring opening through which molecules have to pass (right) [61] .....	17
Figure 1.12: Local coordination and orientation in the case of –tdc <sup>2</sup> linkers, forming the microporous NIIC-10 (left), and in the case of –iph <sup>2</sup> linkers, forming the mesoporous NIIC-20 (right) [63]. .....	18
Figure 1.13: Dodecameric wheel composed of Zn <sup>2+</sup> centers (left); b) the structure of the fundamental nanocell of the NIIC-20-Et framework, with the sphere in the center representing the void space of the cell (middle) [16]; c) the nho-h topology of the overall framework of the metal–organic network NIIC-20-Et (right). Each atom is shown in a different color: zinc (Zn) – green, nitrogen (N) – blue, oxygen (O) – red, carbon (C) – gray; hydrogen (H) is omitted [63]. .....	19
Figure 1.14: Openings of the dodecameric wheels for each material of the NIIC-20-G series. ....	20
Figure 1.15: Number of research papers utilizing classical simulations and machine learning workflows [67]. .....	22
Figure 1.16: Typical Machine Learning workflow incorporated in MOF modelling [68]. .....	23
Figure 1.17: Word cloud representations for the components of a ML model development process. The four main components are database, material representation, target data, and ML algorithm [68]. .....	24

Figure 1.18: Sorption isotherms of short pure alkanes and pure alkenes in silicalite at different temperatures: a) ethane, b) ethene, c) propane, d) propene, e) butane, f) butene. Diagrams show results from the present GCMC simulations, experimental, and simulation results from the literature and the ANN predictions (with continuous line format). .....	25
Figure 1.19: Propane–propene mixture sorption isotherms at 313 and 353 K. ....	26
Figure 1.20: Ethane-Ethene mixture sorption isotherms at 313 and 353 K.....	26
Figure 1.21: Butane–butene, pentane–pentene, hexane–hexene mixture sorption isotherms at 313 K.....	27
Figure 1.22: Self-diffusivity ( $D_s$ ) coefficients for pure alkanes (upper), alkenes (lower) ( $N_C = 2 - 6$ ) at 273, 313, and 373 K obtained from the MD simulations (filled symbols). Experimental MAS PFG NMR results by Dvoyashkina et al. [73] are also shown (open symbols). The experimental sorbate loading range was $0.1 < L < 0.3$ . ANN predictions are shown with x symbol.....	28
Figure 1.23: Density plots for pure ethane, ethene, hexane, hexene at 300 K. Green color is used for the representation of ethane, hexane sorption, while red represents the ethene and hexene sorption. ....	29

## Chapter 2:

Figure 2.1: Effect of time $\delta t$ on thermal conductivity [9]. ....	43
Figure 2.2: Snapshots of the IRMOF-1 and Cu-BTC membrane saturation for $\text{CO}_2/\text{N}_2$ gas mixture at 298 K as a function of time from the constant pressure gradient approach MD simulations. Color code for the gas molecules: C ( $\text{CO}_2$ ), orange; O ( $\text{CO}_2$ ), red; N, blue; $\text{CH}_4$ , green. For the framework atoms: Zn, magenta; Cu, cyan; C, gray; O, ocher [8]. ....	44
Figure 2.3: Adsorption isotherms of pure $\text{CO}_2$ in (a) UTSA-16 UTSA-16(Zn) at 298 K in a pressure range of 0 to 100 KPa and adsorption isotherm of $\text{CO}_2$ from $\text{CO}_2$ - $\text{CH}_4$ mixture ( $\text{CO}_2$ - $\text{CH}_4$ (50:50)) in UTSA-16 and UTSA-16(Zn) [29]. ....	46
Figure 2.4: Isotherm of water adsorption in NU-1000 at 298 K calculated with $5 \times 10^3$ production cycles and variable number $N$ of macrostates used in the calculation of direct transition probabilities. Solid and dashed lines correspond to stable and metastable states, respectively. ....	48
Figure 2.5: Diffusion coefficients (in the log scale) vs temperatures (in the reciprocal scale): (a) longitudinal diffusion; (b) angular diffusion. Results for flexible and rigid MOFs are shown in orange and blue, respectively [30]. ....	49

Figure 2.6 Workflow devised for the derivation and validation of a machine learning potential for H2Al-soc-MOF-1d towards prediction. (a) Construction of the atomistic model for the host/guest system. (b) Generation of the training datasets using ab initio molecular dynamics simulations. (c) Training and validation of the MLP using a continuous neural network. (d) External validation of the trained MLP with its implementation in molecular dynamics simulations and (e) MLP-molecular dynamics and MLP-Monte Carlo predictions [2]......50

Figure 2.7: Schematic view of workflow, material screening using quantum computation and ML. The order and cycle of a pioneering MOF discovery are drawn. Data collection and analysis, atomistic simulation, and data set generation for ML algorithms to discover an advanced MOF structure and its connection with Lab-automation robots to perform experiments are illustrated [1]. .....50

### Chapter 3:

Figure 3.1: Schematic of the synergy between theory, experiment, and simulation.....60

Figure 3.2: Schematic representation of periodic boundary conditions in 2-D .....66

Figure 3.3: MD Algorithm .....68

Figure 3.4: Basic Monte Carlo algorithm. ....73

Figure 3.5: Feedforward Multilayer Perceptron [35].....77

Figure 3.6: Artificial Neural Network architecture.....79

Figure 3.7: Implementation of Back propagation algorithm... .....80

### Chapter 4:

Figure 4.1: Workflow presented in the scope of this research. ....85

Figure 4.2: Number of publications regarding ZIF-8-applications [17]. .....87

Figure 4.3: ZIF-8-unit cell representation. For the representation VESTA software was used. [29]. Different colors are used for representing different atom types: zinc-*grey*, carbon-*brown*, hydrogen-*light pink*, nitrogen- *light blue*. .....89

Figure 4.4: Artificial Neural Network Architecture. ....92

Figure 4.5: Correlation coefficient for the training, testing and for the combined training-testing phases for the sorption ANN predictions. ....94

Figure 4.6: Self-diffusivity coefficients (Ds) for pure alkanes and alkenes at different temperatures. Filled symbols represent results obtained from the present simulation study (sim.) and thick nonfilled symbols represent the MAS PFG NMR experimental measurements (exp.) by Dvoyashkina et al. [43] ANNs predictions are given in thin nonfilled symbols. ....97

Figure 4.7: Self-diffusivity coefficients (Ds) for the methane–ethane– ethene equimolar ternary mixture conducted at different temperatures. Filled symbols represent results obtained from this simulation study (sim.) and nonfilled symbols represent experimental MAS PFG NMR measurements (exp.) by Dvoyashkina et al. [43] The sorbate loading for each component of methane–ethane–ethene mixture was set to 2 molecules/cage following the MAS PFG NMR experimental protocol. .... 100

Figure 4.8: Self-diffusivity (Ds) values for the methane–ethane–ethene sorbed mixture from MD simulations conducted at the 273–373 K range. Individual component sorbate loadings refer to 0.5 (red), 1 (black), 1.5 (blue), and 2 (green with empty symbols shown for reference) molecules/cage. .... 102

Figure 4.9: Three-dimensional contours plot of single particle density distributions for the methane–ethane–ethene sorbed mixture from MD simulations at 293 K. Methane is shown in green, ethane in red, and ethene in blue. The separated contours of density are shown for clarity on the left side, while the actual distributions of the ternary sorbed mixture are given in the right side..... 103

Figure 4.10: Pure Methane GCMC sorption isotherm for the 273-373 K temperature range. ANN predictions are given for intermediate values at 280, 300, 320, 340, and 360 K. .... 104

Figure 4.11: Pure Ethane GCMC sorption isotherm for the 273-373 K temperature range. ANN predictions are given for intermediate values at 280, 300, 320, 340, and 360 K. .... 104

Figure 4.12: Pure Ethene GCMC sorption isotherm for the 273-373 K temperature range. ANN predictions are given for intermediate values at 280, 300, 320, 340, and 360 K. .... 105

Figure 4.13: Relative sorption loadings (rcs) for pure ethane and ethene at T = 273 and 373 K. The ANN rcs predictions are given for T = 280 and 360 K. .... 106

Figure 4.14: GCMC simulation sorption isotherms of short pure alkanes and alkenes at different temperatures: (a) propane, (b) propene, (c) butane, and (d) butene. The ANN predictions are shown in a continuous line format. .... 108

Figure 4.15: GCMC sorption isotherms for ternary methane–ethane–ethene equimolar gas mixture feed at different temperatures ranging from 273 to 373 K (a, b, d–f). ANN predictions for the temperature of 300 K are shown in (c). Pure component GCMC results (open symbols) are shown for reference. .... 110

## Chapter 5:

Figure 5.1: A $2 \times 1 \times 1$ replication of the unit cell of NIIC-20-Bu. Different colors are used for representing different atom types: zinc-green, carbon-gray, oxygen-red, hydrogen-white, nitrogen-blue.....	125
Figure 5.2: Representative architecture of the developed ANN for the sorption predictions. ....	128
Figure 5.3: Correlation coefficient for the training, testing and for the combined training-testing phases. The x-axis represents data used in the sorption training and y-axis the predicted ANN sorption data.....	130
Figure 5.4: GCMC simulation results for sorption isotherms of methane in NIIC-20-Bu at the low-pressure range for five different temperatures (273, 298, 313, 343, 373K) along with experimental measurements from [28]. In continuous line format the prediction from the ANN at four different temperatures are presented (280, 300, 320, 360 K) are shown.....	131
Figure 5.5: GCMC simulation results for methane sorption isotherm at the high-pressure range for five different temperatures (273, 295, 313, 343, 373K) in NIIC-20-Bu along with experimental measurements from [27]. Also, in dashed line format the predictions from the ANN at four different temperatures (280, 300, 320, 360 K) are presented. ....	131
Figure 5.6: Sorption ratio ( $r_S$ ). In the legend box the $T; \backslash\text{vthicksp } T_{\text{ref}}$ is indicated for the GCMC simulation data, experimental measurements and ANN prediction. ....	134
Figure 5.7: GCMC simulation results for sorption isotherms of ethane in NIIC-20-Bu at the low-pressure range for five different temperatures (273, 298, 313, 343, 373 K). In dashed line format the prediction from the ANN at four different temperatures (280, 300, 320, 360 K) are shown. ....	137
Figure 5.8: GCMC simulation results for sorption isotherms of ethane in NIIC-20-Bu at the high-pressure range for five different temperatures (273, 293, 313, 343, 373K). In dashed line format the prediction from the ANN at four different temperatures (280, 300, 320, 360 K) are shown. ....	137
Figure 5.9: Comparison of NIIC-20-Bu, ZIF-8 and Silicalite for methane (a) and (b) for ethane. ....	140
Figure 5.10: SSorption of methane-ethane equimolar mixture in NIIC-20-Bu at temperatures: a) 273, b) 313, and c) 343 K .....	142
Figure 5.11: Selectivity of NIIC-20-Bu for the components of the methane-ethane mixture for $T=273, 313$ and $343$ K. ....	142

## Chapter 6:

Figure 6.1: Molecular Dynamics simulations in NIIC-20 Bu coupled with ANN's prediction at different temperatures and loadings. ....	157
Figure 6.2: Local coordination and orientation of dicarboxylate bridging ligands in NIIC-10 (left) and NIIC-20 (right), leading to the formation of micro- and mesoporous frameworks, respectively [1], [2], [3], [4]. ....	158
Figure 6.3: Comparison of ethane uptake and C <sub>2</sub> H <sub>6</sub> /C <sub>2</sub> H <sub>4</sub> IAST selectivity factor for equimolar C <sub>2</sub> H <sub>6</sub> -C <sub>2</sub> H <sub>4</sub> gas mixture for representative ethane-selective MOFs. The data reported here are marked red [1], [2], [3], [4]. ....	159
Figure 6.4: Performance of the trained SNAP model for ZIF-8. a Learning curves for the RMSE and MAE for energy (left-hand side panel), forces (middle panel) and virial-stress (right-hand side panel). Data are presented for test set A (composed of ~5000 configurations from AIMD simulations) and test set B (composed of ~2000 configurations from classical MD simulations) as a function of the number of configurations in the training set [5]. ....	160
Figure 6.5: A foundation model for materials modelling. Trained only on Materials Project data which consists primarily of inorganic crystals and is skewed heavily towards oxides, MACE-MP-0 is capable of molecular dynamics simulation across a wide variety of chemistries in the solid, liquid and gaseous phases [6]. ....	161
Figure 6.6: The obtained contours of the distribution of the temperature and absolute adsorption in the hydrogen tanks by the changes in the time: (a) The contours of temperature at the time steps of 953s, (b) The contours of temperature at the time steps of 4694s, (c) The contours of absolute adsorption at the time steps of 953s, (d) The contours of absolute adsorption at the time steps of 4694s [7]. ....	162
Figure 6.7: List of publicly available MOF and COF databases [8]. ....	164

# Chapter 1: Introduction to Metal Organic Frameworks

## 1.1 Metal Organic Frameworks

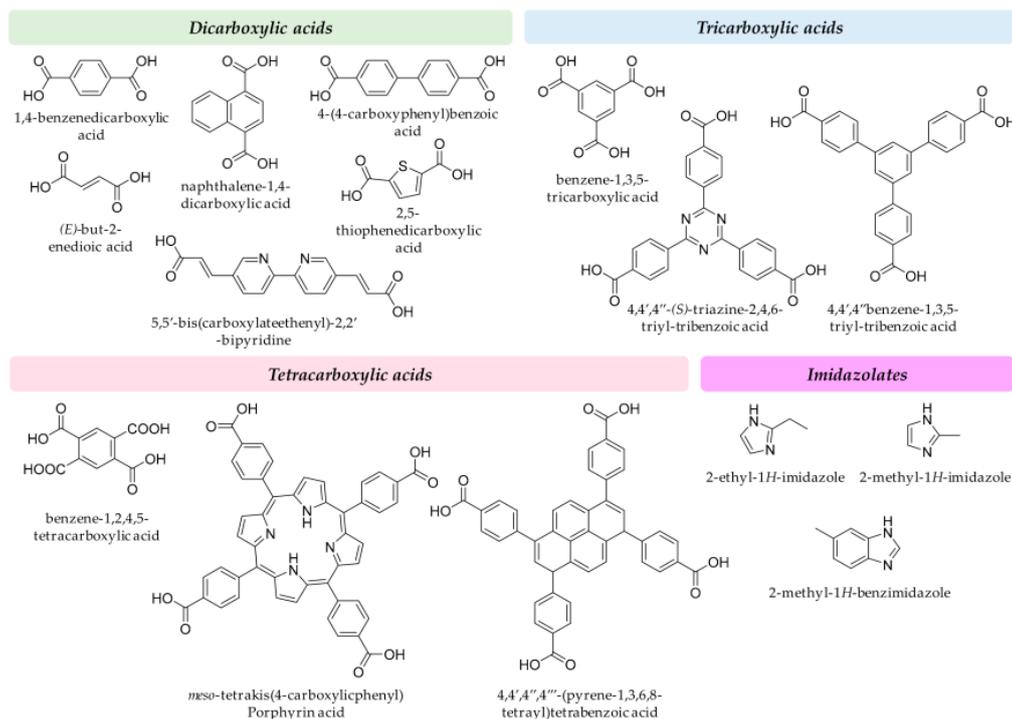
Metal-organic frameworks (commonly known as MOFs or coordination polymers) began gaining high momentum in the 1990s as a promising class of crystalline structures with diverse properties. The birth of metal-organic frameworks dates to the late 1950s [1], but their potential remained largely unnoticed for a long time [2]. In modern materials science, MOFs are a widely studied class of porous materials that combine elements of organic and inorganic chemistry to create one-dimensional, two-dimensional, and three-dimensional topologies. The thousands of available options among inorganic metal ions or metal complexes and organic compounds and the countless combinations between them lead to diverse structures with high design control and unprecedented properties. The result is porous materials with exceptional physical characteristics, such as high pore volumes and very large specific surface areas, a wide range of pore sizes, and good stability. MOFs are utilized in a broad range of chemical applications, including gas storage and separation, sensing technologies, catalysis, and drug storage and delivery [3].

A large part of MOF research, particularly regarding the control of structural stability and the capabilities of each topology, is conducted with the help of computational methods. Through Molecular Dynamics (MD) simulations, information is gathered about the adsorption and diffusion capabilities of gases, as well as their separation, within more favorable timeframes compared to corresponding experiments. When it comes to the adsorption of various gases into MOF pores, Grand Canonical Monte Carlo (GCMC) simulations predict this behavior with remarkable accuracy. These computational procedures are used in combination with experiments, narrowing down the range of candidate materials and guiding experimental efforts toward the most promising MOFs [3].

## 1.2 Structure of MOFs

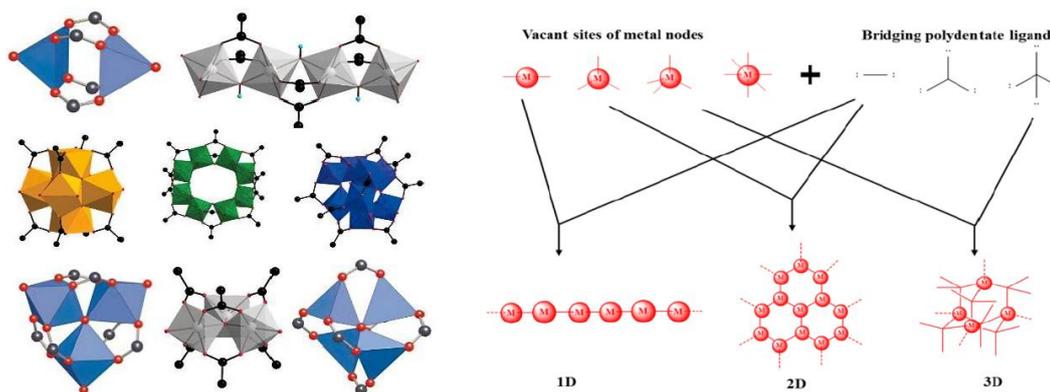
MOFs consist of two types of primary building units (PBUs): metal ions or metal complexes, and organic linkers. The metal ions act as connecting nodes, while the organic chemical compounds bind through them. Typical metal connectors include first-row transition metal ions such as  $\text{Cr}^{3+}$ ,  $\text{Fe}^{3+}$ ,  $\text{Co}^{2+}$ ,  $\text{Zn}^{2+}$ , as well as alkali metals, alkaline earth metals, and rare earth metals [4], [5]. For MOF synthesis, precursor compounds such as nitrates, acetates,

sulfates, chlorides, and metal oxides are commonly used in most synthesis routes except in electrochemical synthesis, where metal rods are employed instead [4]. Regarding the organic nature of MOFs, the organic linkers consist of functional groups such as carboxylic, phosphonic, or sulfonic acids, as well as amine and nitrile groups [4], [5]. These functional groups form coordination bonds with the metal centers or metal complexes at two or more points, depending on whether the bonds are ditopic, tritopic, or polytopic [2]. Some common organic linkers used in MOF design are illustrated in **Figure 1.1**.



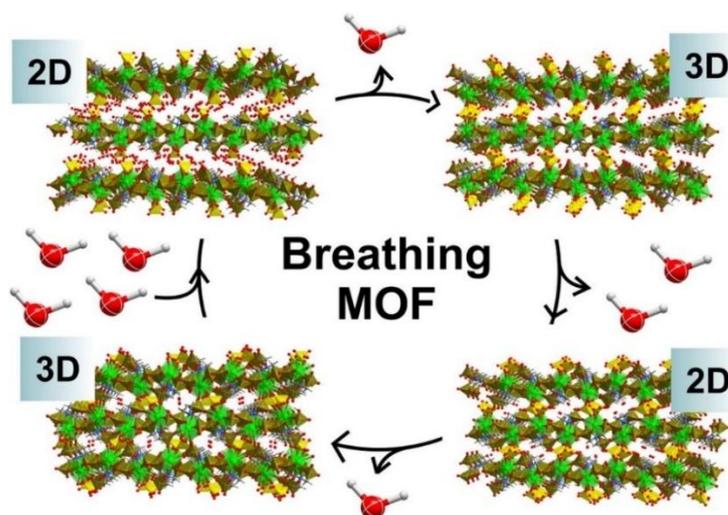
**Figure 1.1:** Typical examples of organic linkers used in the synthesis of MOFs [6].

The organic compounds in MOFs do not bind directly to metal clusters, but rather to metal–oxygen–carbon clusters. These clusters are referred to as secondary building units (SBUs) of a MOF and feature multiple connection points in space, leading to triangular, square, tetrahedral, octahedral, dodecahedral, and many other geometric arrangements. These geometries are a key factor in the vast structural possibilities of MOFs [2], [4]. MOF lattice architecture is impacted by both organic connecting groups and metal ions. An organic linker of the ditopic, tritopic, tetratopic, or polytopic type interacts with the metal ion via one or more labile or unoccupied coordination sites. **Figure 1.2 (left)** depicts typical configurations of metal complexes as SBUs in MOFs. **Figure 1.2 (right)** depicts the degree of influence that the two basic building units have on the metal–organic lattice.



**Figure 1.2:** Some common SBUs used in MOF construction (left) [10] / influence of PBUs on the metal–organic framework (right) [7].

Diving deeper into the geometry and topology of MOFs sheds light to three fundamental components: i) the framework's topology, ii) the inorganic metal centers, and iii) the organic ligands. Each one of these elements contributes to the overall functionality and characteristics of these frameworks, orchestrating a delicate interplay of chemical bonds and spatial arrangement. One of the most intriguing aspects of MOFs lies in their potential for their ability to expand their porosity (breathing mode), surpassing even zeolites in certain design configurations as seen in **Figure 1.3**



**Figure 1.3:** Breathing mode of Metal organic Frameworks [8].

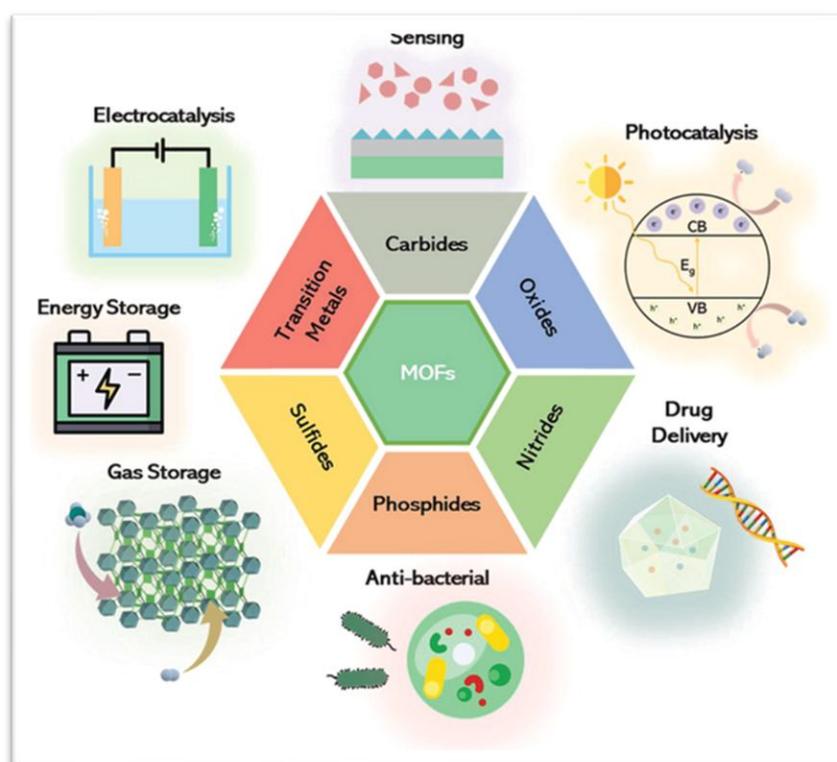
This enhanced porosity is attributed to the enlarged size of the inorganic SBUs, and the dimensional amplification facilitated by sizable metal clusters within the network. These structural nuances not only provide MOFs with larger pores but also enrich them with remarkable versatility and utility across various applications[9], [10], [11], [12]. Thus, the research for MOFs heralds a new era of material science, where the convergence of inorganic

and organic chemistry gives rise to a plethora of possibilities. From catalysis to gas storage, from drug delivery to sensing applications, MOFs are expected to change the landscape of technology and innovation, propelling porous materials into a new realm of industrial applications. Furthermore, the length of the links between inorganic SBUs is intimately related to their high porosity, while longer organic ligands contribute to the enlargement of pore diameter. This interplay between the inorganic and organic constituents underscores the dynamic nature of MOFs, wherein various combinations of metal centers and ligands serve as molecular building blocks, offering a spectrum of possibilities for tailoring their physical and chemical properties [13], [14], [15], [16]. With their extraordinary pore volume and specific surface area, MOFs rise as an unparalleled platform for the immobilization of molecular catalysts on conductive substrates, offering a heightened level of catalytic efficacy. Beyond this, the inherent versatility of MOFs allows for intricate adjustments, facilitating the tailoring of their properties to exhibit enhanced selectivity towards particular chemical reactions.

This inherent adaptability not only broadens the spectrum of potential applications but also underscores the pivotal role MOFs play in advancing catalysis and related domains. Typically, the emergence of MOFs is often perceived as an endeavor to replicate the characteristics of inorganic counterparts, such as zeolites. However, this perception merely scratches the surface of the true essence of MOFs. In reality, it represents a broader conceptualization aimed at integrating (extended) constituents with robust microporosity, thereby ushering in a new era of hybrid material design and creation [17], [18], [19], [20]. Zeolites, while undoubtedly valuable, exhibit limitations in terms of the precise control over functionalization, shape, and pore size, an area where MOFs excel. The quest for ideal constituents is rooted in the desire for materials that can be easily synthesized, maintain consistency, and perform reliably in real world applications. In this regard, MOFs present a promising solution. Their synthetic design principles are inherently straightforward, relying on the meticulous selection of metal centers and ligands to achieve desired properties. Moreover, under specific synthetic conditions, MOFs have the remarkable ability to self-assemble in the solid state, further simplifying their production process. The extensive diversity of elementary components available for MOF synthesis ensures a virtually infinite array of hybrid organic-inorganic combinations, opening doors to a world of possibilities. However, it's worth noting that despite this seemingly boundless potential, a review of literature reveals the existence of recurring structural motifs in MOF preparation. These patterns hint at attempts to predict and

control the architecture of these materials, thereby offering insights into their synthesis and design methodologies [21], [22].

The significance of adsorption cannot be overstated, owing to its straightforward, simplistic, practical, and cost-effective nature. These attributes make it a preferred method, as it minimizes the generation of unwanted byproducts while offering a wide range of recoverable adsorbents. Moreover, adsorption serves as a versatile tool for the removal of various soluble, insoluble, inorganic, and even biological compounds from diverse water systems [23], [24], [25]. In the realm of adsorption, the emergence of MOFs represents a paradigm shift. Considering their remarkable characteristics, MOFs can find application across a great spectrum of fields, showcasing superiority in chemical recognition and separation processes (**Figure 1.4**). Researchers are increasingly drawn to MOFs due to their exceptional adsorption capacity and expansive specific surface area, which offer unparalleled efficiency in capturing target molecules [26], [27], [28], [29].



**Figure 1.4:** Applications of Metal Organic Frameworks [30].

## 1.3 Applications of Metal Organic Frameworks.

### 1.3.1 Biological Compounds

Pan et al. [19] were the first to use a magnetic metal-organic framework composite synthesized by a self-assembly method as an effective adsorbent for the removal of glufosinate, bialaphos, glyphosate, and their primary metabolites, 3-(Methylphosphinico)propionic acid, and aminomethylphosphonic acid. This new material significantly enhances the adsorption capacity of these compounds; furthermore, this innovative approach is based on exploiting the properties of such a composite, with which a powerful analytical method like ultrahigh-performance liquid chromatography-high-resolution mass spectrometry could be combined to achieve separation. The optimization of this technique allowed for a rapid and precise determination of five target compounds that are commonly found in environmental waters. Importantly, the method demonstrated a simultaneous detection of bialaphos, glufosinate, glyphosate and its main metabolites with an impressive recovery rate between 86.2%–104.6% having Relative standard deviation (RSDs) < 10%. Moreover, the linear relationship was also good over a wide concentration range (1.0 mg/L to 100.0 mg/L) of all five target analytes. In particular, the sensitivity was confirmed by the limit of detection ranging from 0.01 mg/L to 0.03 mg/L which revealed that it is highly efficient in measuring trace amounts of impurities in environmental samples hence ensuring its applicability as a tool for monitoring and remediation purposes in cases concerning water or soil contamination. This is an important breakthrough in analytical chemistry that introduces a powerful and versatile means for ecological surveillance and restoration efforts at various sites of pollution.

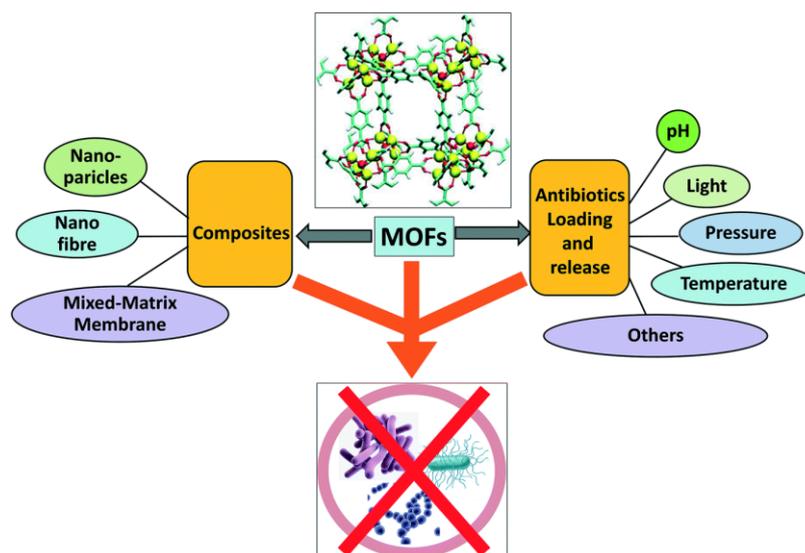
Using the straightforward MLFC method, He et al. [31] created MG@MIL-100-B nanoparticles by introducing functionalized mesopores into MOFs. The researchers used the MG@MIL-100-B composites to extract dopamine, epinephrine, and norepinephrine from rat plasma. The conditions for removal were improved by the evaluated solid phase extraction parameters. While the same variable for norepinephrine was predicted to be 0.02 ng mL<sup>-1</sup>, the detection limits for dopamine and epinephrine (S/N = 3) were reported to be as low as 0.005 ng mL<sup>-1</sup>. The accuracy ranges were 2.84–6.63% (N = 6) for intra-day and 5.70–11.44% (N = 6) for inter-day. Recoveries from confounding tests showed favorable outcomes of 94.40–109.51%. This approach presents a novel sample pre-treatment strategy for clinical treatments and research.

Yamini and Safari et al. [32] created magnetic framework composites with core-shell structure (Fe<sub>3</sub>O<sub>4</sub>@TMU-21), in order to remove trace pyrethroid residues from fruit juice samples using magnetic solid-phase (MSPE). Under ideal circumstances, the range of 0.5–250 µg L<sup>-1</sup> (R<sup>2</sup> = 0.99) yielded optimal linearity, whereas the range of 0.1–0.05 µg L<sup>-1</sup> (based on S/N = 3) was the limit of recognition. The accuracy of the method expressed as a relative standard deviation (RSD) for eliminating and identifying the 100 µg L<sup>-1</sup> pyrethroid residues in the sample solution was reported to be between 3.1 and 4.4%.

This is a big step forward in analytical chemistry, introducing a very effective and versatile instrument for ecological surveillance. This development, which allows for more exact detection and analysis of environmental pollutants, has prompted new research initiatives aimed at comprehending complex ecological relationships, reducing contamination, and providing long-term solutions for ecosystem preservation. Its extensive usefulness includes monitoring air and water quality, measuring biodiversity, and tracking the long-term effects of climate change, making it an important resource for scientists, policymakers, and conservationists around the world [33], [34].

### **1.3.2 MOFs and Antibiotics**

In a significant study, Li et al. [35] evaluated the adsorption properties of ZIF-8 for the extraction of an aqueous combination of two widely used antibiotics (Oxytetracycline hydrochloride and Tetracycline) using ZIF-8. The results showed that ZIF-8 simultaneously absorbed 82.5%-90.7%-90.7%-90.7% of Tetracycline from aqueous solvents. Furthermore, the study reported impressive maximum adsorption capacities of 312.5 mg/g and 303.0 mg/g for oxytetracycline hydrochloride and tetracycline, respectively. The study also reported impressive max. adsorption capacities (312.5-mg/g) and 303.0-mg/g for Oxytetracycline hydrochloride and oxytetracycline, respectively. The adsorption process of these antibiotics follows pseudo-second order kinetics, suggesting a gradual but effective removal mechanism. High correlation coefficient (R<sup>2</sup>): 0.963+0.981+Tetracycline +ZIF-8 adsorption at temperature of 303 K. This suggests a strong synergy between these antibiotics and ZIF-8 adsorbent under specific conditions. This further confirms ZIF-8's potential as an effective antimicrobial removal agent from aqueous systems. These findings offer promising opportunities to address the growing concern of antibiotic pollution and its negative impacts on ecosystems and human life.



**Figure 1.5:** Metal Organic Frameworks in Antibiotics [12].

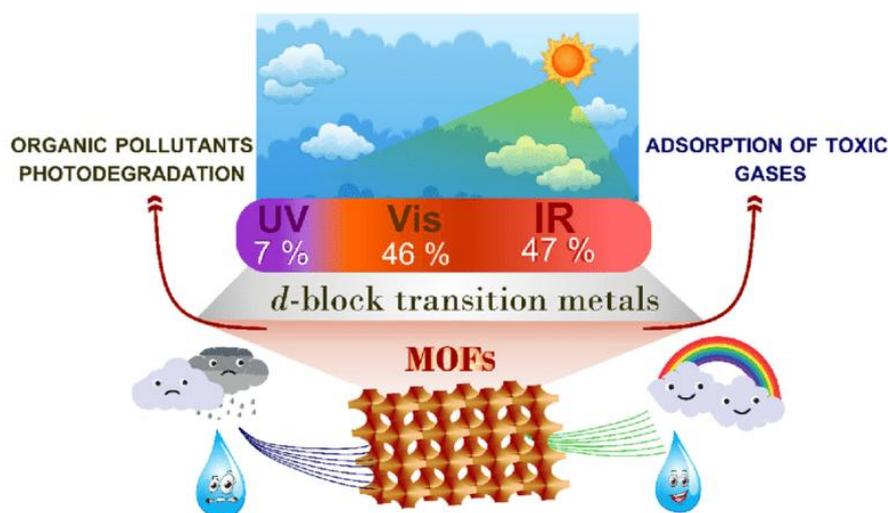
Tran et al. [36], developed a new method for the synthesis of mesoporous (carbon-embedded) zero-valent (Fe) iron sites using a metal-organic framework (MIL-53). This new method was then applied to the remediation of aqueous antibiotics (CIPROFLOXACIN). This research determined that the mesoporous carbon derived from MIL-53 showed great promise as an effective adsorbent for treating aquatic antibiotics, based on thorough evaluations of chemical stability, reusability (up to 5 cycles), and impressive adsorption capacity (90.9 mg/g). These results emphasize the importance of the developed technique in addressing the urgent problem of antibiotic contamination in water systems, showcasing its versatility and efficacy.

### 1.3.3 MOFs and Toxic Pollution

Sun et al. [37] saw the development of a nanostructured Fe-Co based metal-organic framework (MOF-74) adsorbent that was especially designed to remove arsenic from water sources. The ideal molar ratio of Fe/Co in the adsorbent was found by the researchers through rigorous experimentation and analysis to be 2:1, as confirmed by extensive adsorption testing. After analyzing the  $Fe_2Co_1$  MOF-74 using multiple techniques, Sun et al. [37]. revealed some interesting structural characteristics. The diameter of the resultant nanoparticle was determined to be between 60 and 80 nm, indicating a well-tailored nanostructure. The material's large adsorption surface area was highlighted by the impressively high specific surface area of  $147.82 \text{ m}^2/\text{g}$ . Even more, the study reported maximum adsorption capacities for both arsenic species, with values of 292.29 mg/g and 266.52 mg/g observed for As (V) and As (III), respectively. The results signify the exceptional arsenic removal capabilities of the Fe-Co based MOF-74 adsorbent, showcasing its potential as a promising solution for addressing arsenic

contamination in water systems. Such advancements represent crucial steps towards safeguarding human health and environmental integrity in regions affected by arsenic pollution.

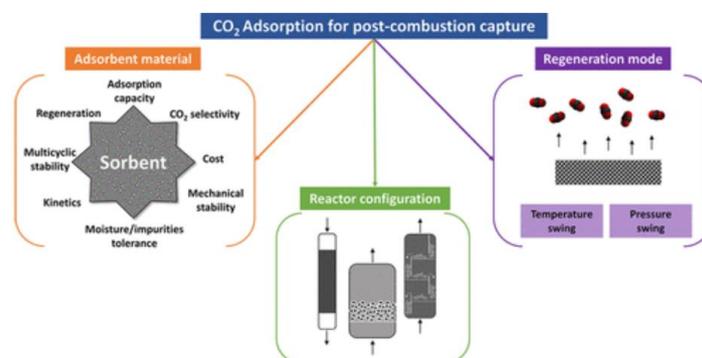
Chowdhury et al. [38], conducted thorough investigation involving the diligent synthesis of MIL-53(Al)-graphene oxide (GO) nanocomposites at a range of mass ratios (from 1% to 25%) of GO to MIL-53(Al). Subsequently, these nanocomposites underwent a thorough assessment to see how well they removed arsenic (As(III)), a common groundwater contaminant. Batch studies were carefully carried out on MIL-53(Al)-GO nanocomposites for As (III) removal from aqueous solutions in order to clarify the adsorption behavior and kinetics under various environmental conditions. Temperature changes (from 298 K to 318 K), starting As (III) concentrations (10 mg/L to 110 mg/L), solution pH levels (2 to 11), and adsorbent dosages (from 0.2 g/L to 3.0 g/L) were all evaluated during the experiment. Surprisingly, the findings revealed that incorporating GO into the MIL-53(Al) framework greatly improved As (III) removal, with the greatest adsorption capacity of 65 mg/g reached at a 3% GO ratio in the nanocomposite. Notably, the adsorption isotherms and kinetics closely followed the Langmuir isotherm and pseudo-second-order models, demonstrating the MIL-53(Al)-GO nanocomposite's resilience and reliability as an arsenic adsorbent. These discoveries have significant implications for the development of sophisticated water filtration technologies, notably in reducing the negative impacts of arsenic poisoning on human health and environmental sustainability.



**Figure 1.6:** Metal Organic Frameworks in toxic pollution applications [39].

### 1.3.4 Gas adsorption in MOFs

In recent years, there has been a concerted effort among researchers to address environmental pollution, driven by the urgent need to tackle pressing environmental issues. Among these concerns, the proliferation of greenhouse gas (GHG) emissions stands out as a significant contributor to global climate change, posing immense challenges to environmental sustainability. Traditional solid adsorbents like zeolites and activated carbons, while widely used, exhibit limitations in their ability to effectively capture and separate CO<sub>2</sub>, a key greenhouse gas. In contrast, Metal-Organic Frameworks (MOFs) have emerged as groundbreaking materials for CO<sub>2</sub> capture and separation due to their increased surface area, tunable pore sizes, and their unique framework structures. This section works to provide an overview and insights into the potential of porous MOFs as solid adsorbents for CO<sub>2</sub> capture and separation across diverse environmental conditions (**Figure 1.7**). However, developing porous materials with a size-selective effect for CO<sub>2</sub> to this day remains a significant challenge, particularly in controlling pore sizes within the narrow range of 3 to 4 Å. Addressing this challenge, Yang et al. [40], introduced a new approach by designing and synthesizing a novel anion-pillared material known as ZU-66 (where ZU stands for Zhejiang University), which incorporates "molecular rotors" to facilitate the separation of CO<sub>2</sub> from other gases such as CH<sub>4</sub> and N<sub>2</sub>. This innovative design conferred ZU-66 with impressive separation selectivity for both CO<sub>2</sub>/N<sub>2</sub> and CO<sub>2</sub>/CH<sub>4</sub>, alongside a high CO<sub>2</sub> capacity of 4.56 mmol g<sup>-1</sup> at 298 K and 1 bar pressure. The remarkable potential of ZU-66 in real-world applications was further validated through actual breakthrough tests, highlighting its efficacy in conducting both CO<sub>2</sub>/N<sub>2</sub> and CO<sub>2</sub>/CH<sub>4</sub> separations. These findings underscore the pivotal role of innovative materials like ZU-66 in advancing the field of CO<sub>2</sub> capture and separation, offering promising solutions to mitigate the adverse impacts of GHG emissions on the environment and human well-being.



**Figure 1.7:** CO<sub>2</sub> Capture in Metal Organic Frameworks [41].

Huelsenbeck et al. [42], conducted a comprehensive evaluation of the morphological control of the anisotropic  $[\text{Zn}_2(\text{NDC})_2(\text{DABCO})]_n$  Metal-Organic Framework (MOF) and its subsequent adsorption characteristics for the separation of  $\text{CO}_2/\text{CH}_4$  gas mixture. Employing modulators to influence the morphology of MOF crystals, the researchers observed an increase in film orientation for the rod-shaped morphology, indicating the favorable nature of MOF morphology for applications requiring thin-film formation and large surface areas. Furthermore, Alivand et al. [43]. investigated the impact of different ratios of cluster/modulator (Y) and cluster/ligand (X) on the gas uptake of  $\text{MIL-101@M-X-Y}$ , along with its textural features and reaction yield. Their research demonstrated a substantial improvement in the surface area and pore volume of the  $\text{MIL-101@M-0.5-Y}$  series, synthesized with a Cr: H2BDC ratio of 1:2, compared to the conventional 1:1 ratio. For instance, in the  $\text{MIL-101@M-0.5-0.5}$  case, the surface area increased to  $3596 \text{ m}^2/\text{g}$  and the pore volume increased to  $1.65 \text{ cm}^3/\text{g}$ , representing increments of 23.8% and 27.9%, respectively, over the conventional  $\text{MIL-101}(\text{Cr})$  sample. Using an in-house setup at 1 to 35 bar and 298 K, researchers estimated the  $\text{H}_2\text{S}$  and  $\text{CO}_2$  adsorption isotherms of all synthesized constituents. Notably, the  $\text{MIL-101@M-0.5-0.5}$  exhibited an impressive adsorption capacity for  $\text{H}_2\text{S}$  ( $7.63 \text{ mmol/g}$ ) and  $\text{CO}_2$  ( $3.16 \text{ mmol/g}$ ) at 1 bar, marking a substantial increase of 59.3% and 44.9% compared to the conventional  $\text{MIL-101}(\text{Cr})$ . This enhancement in gas adsorption capacity, particularly for polar  $\text{H}_2\text{S}$  molecules, was attributed to both the increased formation of unsaturated  $\text{Cr}^{3+}$  sites and the enhanced textural features of  $\text{MIL-101@M-0.5-0.5}$ , facilitating more effective communication at low pressures. In addition to its exceptional gas uptake capacity and textural features, the modified  $\text{MIL-101@M-0.5-0.5}$  also exhibited a remarkable reaction yield of 71.9%, significantly higher than the conventional  $\text{MIL-101}(\text{Cr})$  at 59.7%. These findings underscore the potential of tailored MOFs in achieving superior gas separation performance and highlight their relevance in addressing environmental challenges related to gas emissions and pollution.

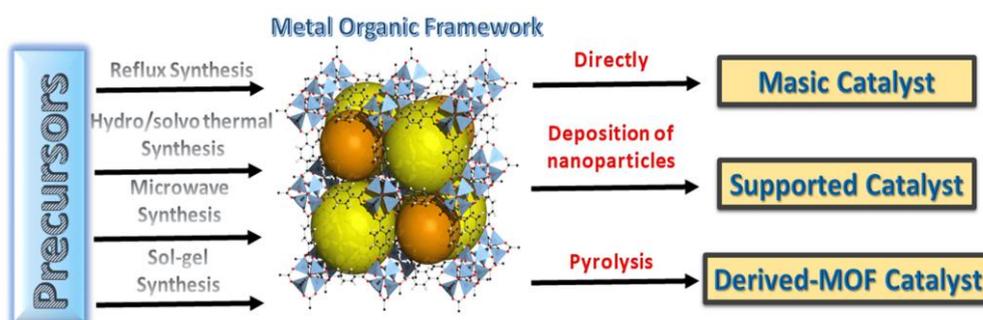
In a research study [44], a porous, three-dimensional Co(II)-metal organic framework,  $[\{\text{Co}(\text{TCPB})0.5(\text{H}_2\text{O})\}.\text{DMF}]_n$  (abbreviated as MOF1) was created by combining a hard, tetratopic ligand, H4TCPB (1,2,4,5-Tetrakis(4-carboxyphenyl)benzene). Gas adsorption research showed that MOF1 has a remarkable adsorption property for  $\text{CO}_2$  compared to other gases such as  $\text{N}_2$ , Ar, and  $\text{H}_2$ , with an impressive isosteric heat of adsorption ( $Q_{st}$ ) value of  $35.4 \text{ kJ/mol}$ . The significantly elevated  $Q_{st}$  value was attributed to improved interaction between  $\text{CO}_2$  molecules and Co(II) metal sites located within the one-dimensional channels of MOF1. Zeng et al. [45] investigated the separation of  $\text{Rn}/\text{N}_2$  and  $\text{Rn}/\text{O}_2$  mixtures at 298 K using a

variety of typical MOFs. Among them, HKUST-1, ZIF-11, ZIF-12, and IRMOF-62 emerged as the most promising candidates for Rn separation, exhibiting selectivities of approximately 2500 and 1200 for Rn/N<sub>2</sub> and Rn/O<sub>2</sub> mixtures, respectively, at 1 atm and 298 K. Moreover, Cheng et al. [46], reported the production of MOF-1, which showed significant potential for selective gas adsorption owing to its delaminated nanosheets with a porous structure. Yoon et al. [47] introduced a novel method involving the reduction of mesoporous MOF incorporated with Cu(I) ions using Na<sub>2</sub>SO<sub>3</sub> as a reducing agent. The incorporation of Cu(I) resulted in a notable increase in the working capacity for C<sub>3</sub>H<sub>6</sub> (1.3 mmol/g) at 10-100 kPa and a fourfold increase in C<sub>3</sub>H<sub>6</sub>/C<sub>3</sub>H<sub>8</sub> selectivity compared to the parent MOF. Under dynamic mixture flow conditions, the Cu(I)-loaded MOF effectively separated C<sub>3</sub>H<sub>6</sub> from C<sub>3</sub>H<sub>8</sub>. Furthermore, Wang et al. [48] developed a unique heterometallic MOF, [In<sub>6</sub>O<sub>3</sub>Tb<sub>3</sub>O(CBDA)<sub>3</sub>].18DMF.3H<sub>2</sub>O (referred to as In/Tb-CBDA), using the heterometallic cooperative crystallization (HCC) approach. In/Tb-CBDA exhibited high selective adsorption for C<sub>2</sub>H<sub>6</sub>/CH<sub>4</sub>, C<sub>3</sub>H<sub>8</sub>/CH<sub>4</sub>, and CO<sub>2</sub>/CH<sub>4</sub> at ambient temperature, as confirmed by theoretical ideal adsorption solution theory calculations. These pioneering studies underscore the potential of MOFs in diverse applications, ranging from gas separation to selective adsorption, offering promising avenues for addressing environmental and industrial challenges.

### 1.3.5 MOFs and Catalysts

A common strategy employed to enhance the stability and catalytic efficiency of MOFs involves integrating catalytically active yet unstable nanoparticles within their porous structures. One critical application area for such enhanced MOFs is in proton exchange membrane fuel cells, where the oxygen reduction reaction (ORR) holds paramount importance. Due to the sluggish kinetics of ORR, necessitating high loading of Pt catalysts to expedite the reaction, the resultant increase in cost poses a significant hurdle to the widespread commercial adoption of fuel cells. In response, a novel approach has emerged, leveraging MOFs as templates and precursors to fabricate hierarchical porous structured metal nanoparticles/carbon composites through pyrolysis in an inert atmosphere [49], [50]. This study presents a comprehensive overview of the advantages and limitations associated with employing MOFs as catalysts. The compositional makeup of MOF frameworks and their pore structures exert significant influence over catalytic performance, enabling functionalities such as shape-selective and bifunctional catalysis. Nonetheless, further research endeavors are imperative to elucidate the precise catalytic sites within MOFs, quantify reaction/transport processes, and ascertain intrinsic catalytic reaction rates. The inherent heterogeneity of MOFs, coupled with

the presence of defective structures not readily discernible through conventional X-ray diffraction crystallography, poses challenges that impede progress in this domain. Therefore, comprehensive insights into the dynamics of MOFs, identification of catalytic sites, and elucidation of intrinsic kinetics of catalytic reactions are essential to advance optimal catalyst synthesis protocols. Addressing challenges pertaining to MOF stability and regeneration, particularly concerning node-linker bonds, requires a deeper understanding of MOF behavior under catalytic conditions. It is imperative to probe these issues comprehensively and devise strategies to mitigate processes leading to deactivation, thus paving the way for enhanced stability and sustained catalytic performance. Efforts in this direction will not only bolster our understanding of MOF catalysis but also facilitate the development of robust and efficient catalysts for diverse industrial applications (**Figure 1.8**).



**Figure 1.8:** Catalysis in Metal Organic Frameworks [51].

Yang et al. [52], devised a new method for synthesizing a catalyst comprising  $\gamma$ -Fe nanoparticles encapsulated within a porous N-doped carbon matrix, employing a NaOH-assisted and metal-organic framework (MOF)-derived approach. The presence of Fe atoms doped within the MOF structure played a pivotal role in stabilizing the resulting  $\gamma$ -Fe nanoparticles. Remarkably, the catalysts exhibited exceptional catalytic activity, particularly in facilitating the four-electron transfer pathway for oxygen reduction reaction (ORR) in alkaline solutions, showcasing promising potential for various applications. In a separate study, Gao et al. [53], developed a wet-chemical approach to fabricate well-dispersed and ultrafine CrPd nanoparticles (NPs) supported on a MIL101-NH<sub>2</sub> substrate. The resulting Cr<sub>0.4</sub>Pd<sub>0.6</sub>/MIL-101-NH<sub>2</sub> catalyst demonstrated remarkable catalytic activity, achieving a turnover frequency of approximately 2009 mol H<sub>2</sub> mol Pd<sup>-1</sup> h<sup>-1</sup> at 323 K for hydrogen production from aqueous formic acid solutions. This innovative synthetic technique holds promise for immobilizing

ultrafine metal NPs on functionalized MOFs and expanding its applicability to other metallic systems, especially in aqueous solution-based reactions.

Wu et al. [54], have presented a unique method for immobilizing various ruthenium complexes, such as  $\text{RuCl}_3$ ,  $[\text{RuCpCl}_2]_2$  (where Cp represents pentamethylcyclopentadienyl), and  $[\text{Ru}(\text{C}_6\text{Me}_6)\text{Cl}_2]_2$  (where  $\text{C}_6\text{Me}_6$  designates hexamethylbenzene), onto an azolium-based MOF by post-synthetic metalation. The resulting series of effective heterogenized ruthenium catalysts, known as Rux-NHC-MOF (with x ranging from 1 to 3 and NHC denoting N-heterocyclic carbene), demonstrated outstanding catalytic performance in the hydrogenation of  $\text{CO}_2$  to formic acid. Notably, the  $\text{Ru}_3$ -NHC-MOF catalyst with the electron-donating capability of the  $\text{C}_6\text{Me}_6$  ligand from the  $[\text{Ru}(\text{C}_6\text{Me}_6)\text{Cl}_2]_2$  complex had the highest activity level. Under specified reaction conditions, including a total pressure of 8 MPa ( $\text{H}_2/\text{CO}_2 = 1$ ) and the addition of  $\text{K}_2\text{CO}_3$  additive in N,N-Dimethylformamide solvent, the catalyst attained a high turnover number (TON) value of up to 3803 at  $120^\circ\text{C}$ . Moreover, Under specific reaction conditions, including a total pressure of 8 MPa ( $\text{H}_2/\text{CO}_2 = 1$ ) and the presence of  $\text{K}_2\text{CO}_3$  additive in N,N-Dimethylformamide solvent, the catalyst achieved a high turnover number (TON) value up to 3803 at  $120^\circ\text{C}$ . Moreover, the heterogenized  $\text{Ru}_3$ -NHC-MOF catalyst exhibited efficient recovery by filtration without significant loss of catalytic activity, underscoring its potential for practical applications.

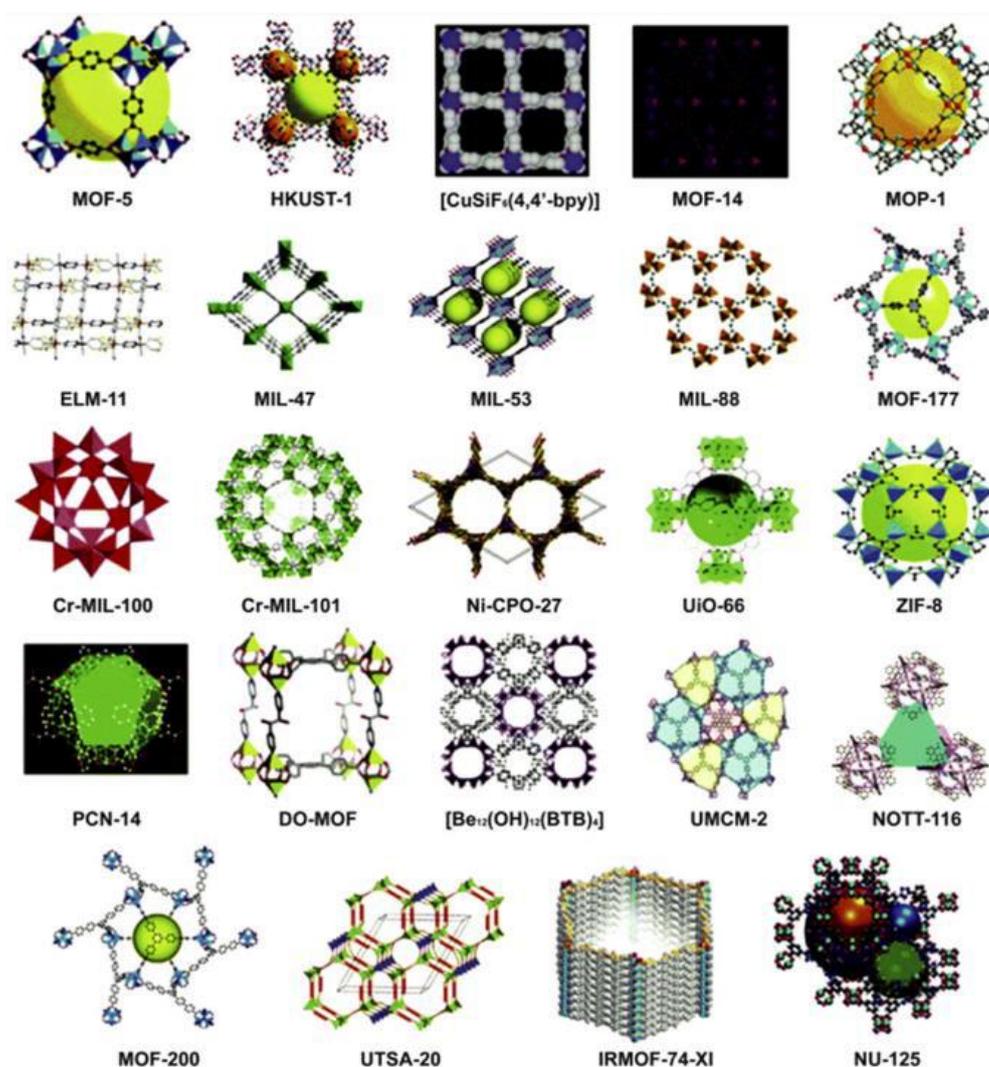
As research on MOFs has greatly expanded, numerous frameworks have been extensively studied, with the most recognisable examples shown in **Figure 1.9**.

No.	Designation	Formula	Abbreviation interpretation	
1	MOF-74	$\text{Zn}_2\text{DOT}$	Metal-Organic Frameworks	
2	MOF-101	$\text{Cu}_2(\text{BDC-Br})_2(\text{H}_2\text{O})_2$		
3	MOF-177	$\text{Zn}_4\text{O}(\text{BTB})_2$		
4	MOF-235	$[\text{Fe}_3\text{O}(\text{BDC})_3(\text{DMF})_3][\text{FeCl}_4](\text{DMF})_3$	IsoReticular Metal-Organic Frameworks	
5	MOF-253	$\text{Al}(\text{OH})(\text{BPYDC})$		
6	IRMOF-1 (MOF-5)	$\text{Zn}_4\text{O}(\text{BDC})_3 \cdot 7\text{DEF} \cdot 3\text{H}_2\text{O}$		
7	IRMOF-16	$\text{Zn}_4\text{O}(\text{TPDC})_3 \cdot 17\text{DEF} \cdot 2\text{H}_2\text{O}$		
8	UiO-66	$\text{Zr}_6\text{O}_4(\text{BDC})_6$	Universitetet i Oslo	
9	UiO-67	$\text{Zr}_6\text{O}_4(\text{BPDC})_6$		
10	UiO-68	$\text{Zr}_6\text{O}_4(\text{TPDC})_6$	Materials of Institut Lavoisier	
11	MIL-53	$\text{Al}(\text{OH})(\text{BDC})$		
12	MIL-53(Al)-NH <sub>2</sub>	$\text{Al}(\text{OH})(\text{BDC-NH}_2)$		
13	MIL-88A	$\text{Fe}_3\text{O}(\text{MeOH})_3(\text{O}_2\text{CCH}=\text{CHCO}_2)_3 \cdot \text{MeCO}_2 \cdot n\text{H}_2\text{O}$		
14	MIL-88-Fe	$\text{Fe}_3\text{O}(\text{MeOH})_3(\text{O}_2\text{C}(\text{CH}_2)_2\text{CO}_2)_3 \cdot \text{AcO}(\text{MeOH})_{4.5}$		
15	MIL-88B-4CH <sub>3</sub>	$2\text{Fe}_3\text{O}(\text{OH})(\text{H}_2\text{O})_2(\text{BDC-Me}_2)_3$		
16	MIL-100-Fe	$\text{Fe}^{\text{III}}_3\text{O}(\text{H}_2\text{O})_2\text{F}(\text{BTC})_2 \cdot n\text{H}_2\text{O}$		
17	MIL-101	$\text{Cr}_3\text{O}(\text{H}_2\text{O})_2\text{F}(\text{BDC})_2 \cdot n\text{H}_2\text{O}$		
18	HKUST-1 (MOF-199)	$\text{Cu}_3(\text{BTC})_2$		Hong Kong University of Science and Technology
19	LIC-1	$\text{Cd}_2(\text{BDC-NH}_2)_3(\text{DMF})_4$		
20	ZIF-8	$\text{Zn}(\text{MIM})_2$	Zeolite Imidazolate Framework	
21	ZIF-90	$\text{Zn}(\text{FIM})_2$		
22	CPL-2	$\text{Cu}_2(\text{PZDC})_2(4,4'\text{-BPY})$	Coordination Polymer with pillared Layer structure	
23	F-MOF-1	$[\text{Cu}(\text{HFBBAA})(\text{phen})_2](\text{H}_2\text{HFBBAA})_2(\text{H}_2\text{O})(\text{HCO}_2)$		Fluorinated Metal-Organic Framework
24	MOP-1	$\text{Cu}_{24}(\text{m-BDC})_{24}(\text{DMF})_{14}(\text{H}_2\text{O})_{10}$	Metal-Organic Polyhedra	

**Figure 1.9:** Typical Metal Organic frameworks with their chemical formula and their abbreviations [55].

## 1.4 Types of MOFs

The continuously expanding field of MOFs has attracted significant attention from the scientific community in the fields of chemistry and materials science, with over 90,000 MOF structures identified so far, according to the Cambridge Structural Database. Some characteristic types of MOFs are illustrated in **Figure 1.10**. Alongside the vast number of different MOFs, there are numerous methods for classifying them. One widely adopted classification approach, which will be briefly discussed below based on their constituent building units [56]. While there are several families of MOFs in this thesis two types will be covered in the next subsections. Specifically, zeolitic imidazolate frameworks (ZIF) and Nikolaev Institute of Inorganic Chemistry frameworks (NIIC).



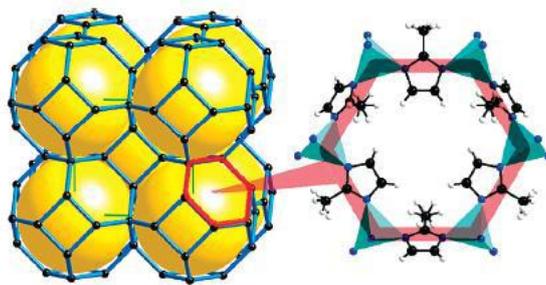
**Figure 1.10:** Representative types of metal organic networks synthesized by various research groups [57].

## 1.4.1 Zeolitic imidazolate frameworks (ZIFs)

Zeolitic imidazolate frameworks (ZIFs) are a subcategory of metal-organic structures that adopt the topology of conventional zeolites, while retaining the hybrid organic–inorganic nature characteristic of MOFs. They consist of tetrahedrally coordinated transition metal complexes such as Fe, Cu, Co, and Zn, which are connected through imidazolate linkages of the form M–Im–M with an angle of approximately  $145^\circ$  (where M is the tetrahedrally coordinated metal ion and Im the imidazolate linker), via self-assembly mechanisms [58], [59]. Their topology mirrors that of aluminosilicate zeolites, as the metal ions mimic the role of silicon and the imidazolate anions act as bridging agents, resembling the role of oxygen in a zeolitic structure [59], [60]. However, the coordination ability between metal centers and organic linkers is the key feature that distinguishes ZIFs from conventional zeolites, granting them superior lattice flexibility. Equally notable is their high chemical and thermal stability. This class of MOFs is quite popular for both gas, small alcohol storage and separation applications, with a strong emphasis on molecular simulations. Thanks to their biocompatibility, ZIFs are also gaining popularity in the medical field, particularly for cancer therapy and bone treatment applications [56].

### 1.4.1.1 Zeolitic imidazolate frameworks-8 (ZIF-8)

The framework of ZIF-8 adopts a sodalite (SOD) topology, as illustrated in **Figure 1.11**. ZIF-8 features micropores with diameters of approximately  $3.4 \text{ \AA}$ , allowing for the adsorption of small gas molecules, while the pore cavities have a diameter of  $11.6 \text{ \AA}$ . The material exhibits thermal stability up to  $400^\circ\text{C}$  and possesses a high surface area, typically ranging from  $900$  to  $1600 \text{ m}^2/\text{g}$ . Additionally, ZIF-8 is hydrophobic, which further enhances its performance in selective gas adsorption applications [61], [62]. ZIF-8 is an esteemed part of the MOF family, with excellent features in gas storage, catalysis, drug delivery, and sensing applications. Metal ions or clusters coupled with organic ligands form MOFs, which have very porous structures and large surface areas. ZIF-8 consists of zinc ions ( $\text{Zn}^{2+}$ ) coupled with imidazolate ligands, forming a three-dimensional network of interconnecting nanopores.



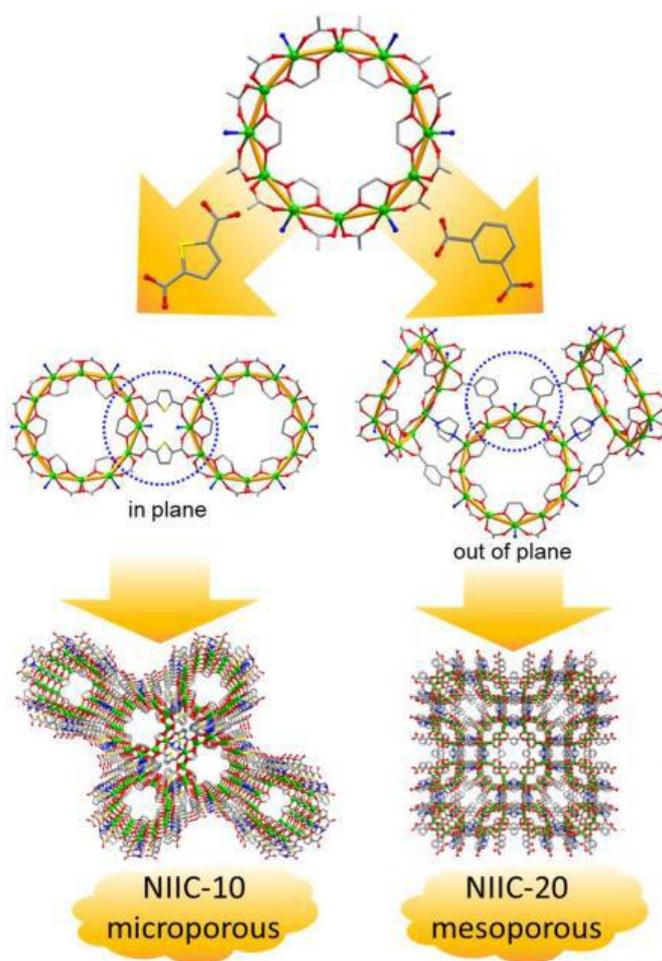
**Figure 1.11:** The sodalite topology (left) and narrow six-membered-ring opening through which molecules have to pass (right) [61].

### 1.4.2 MOFs structures from Nikolaev Institute of Inorganic Chemistry (NIIC)

Lysova et al. [63], [64], [74] created two families of 3D isorecticular MOF structures, named NIIC-10 and NIIC-20. The acronym “NIIC” originates from the Nikolaev Institute of Inorganic Chemistry, where these two types of materials were synthesized. NIIC-10 features a microporous structure with channel sizes ranging from 1.8 to 4.9 Å, while NIIC-20 has a mesoporous structure with larger pores of 2.5 nm in diameter. These two new MOF series are based on carboxylate-based building units of the form  $Zn_{12}(RCOO)_{12}(G)_6$ , arranged in a dodecameric wheel-like structure, where  $(RCOO)_{12}$  represents either 2,5-thiophenedicarboxylate (tdc) in the case of NIIC-10, or isophthalate (iph) in the case of NIIC-20. The  $(G)_6$  component refers to the glycol fragment, which decorates the inner gate of the wheel-shaped architecture [64].

In their research, Lysova et al. [63] initially discovered noteworthy properties in the NIIC-10 material series, particularly regarding their adsorption selectivity for hydrocarbons. Specifically, NIIC-10 exhibited a preference for saturated over unsaturated hydrocarbons, such as ethane over ethylene and cyclohexane over benzene, a behavior that is quite the opposite of the typical trend, where unsaturated hydrocarbons are more readily adsorbed due to interactions with  $\pi$ -orbitals. Lysova et al. [63] concluded that the glycol fragments inside the channels play a key role in this phenomenon by facilitating numerous Van der Waals interactions between the saturated substrates and the material’s surface. This unique property could be exploited in ethylene production, using NIIC-10 MOFs to separate ethane impurities from ethylene, as an alternative to the traditional distillation technique. However, such an application may face practical limitations, since porous materials typically trap unsaturated hydrocarbons [63].

To improve the practical applicability of the NIIC-10 MOFs, the research team of Lysova et al. [63] successfully modified the previous series by changing the dicarboxylic linker group (RCOOH)<sub>12</sub> from  $-tdc^{2-}$  to  $-iph^{2-}$ , creating the NIIC-20 material family of materials. Since the angle of the isophthalic carboxylate group ( $\varphi = 119.1^\circ$ ) is larger than that of the 2,5-thiophenedicarboxylate ( $\varphi = 114.7^\circ$ ), the self-assembly of the crystals differs dramatically. As illustrated in **Figure 1.12** of Lysova et al. [63], the  $\{Zn\}_{12}$  wheels align in parallel when connected with  $-tdc^{2-}$  linkers, whereas with  $-iph^{2-}$  linkers, arched structures are formed and the orientation is out of plane. This linker modification leads to a mesoporous structure. In this way, NIIC-20 retains the fundamental dodecameric wheel structure and the attractive ethane-over-ethylene adsorption capability, while also offering greater adsorption capacity for small gases due to its mesoporous nature [63], [64].



**Figure 1.12:** Local coordination and orientation in the case of  $-tdc^{2-}$  linkers, forming the microporous NIIC-10 (left), and in the case of  $-iph^{2-}$  linkers, forming the mesoporous NIIC-20 (right) [63].

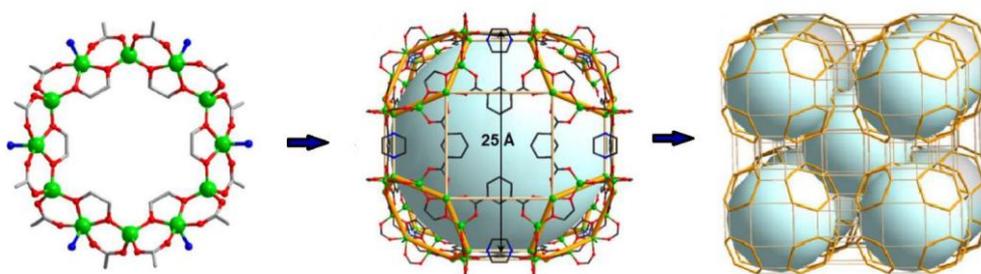
### 1.4.2.1 NIIC 20 Structure

The general chemical formula of the NIIC-20-G MOF family is represented as:  $Zn_{12}(iph)_6(dabco)_3(G)_6$  (where dabco = 1,4-diazabicyclo [2.2.2]octane). Each of the five different porous materials in the NIIC-20 family series is distinguished by the type of deprotonated diol group (G) it contains:

- NIIC-20-Et contains ethylene glycol (EtO<sub>2</sub>)
- NIIC-20-Pr contains 1,2-propanediol (PrO<sub>2</sub>)
- NIIC-20-Bu contains 1,2-butanediol (BuO<sub>2</sub>)
- NIIC-20-Pe contains 1,2-pentanediol (PeO<sub>2</sub>)
- NIIC-20-Gl contains glycerol (GlO<sub>2</sub>)

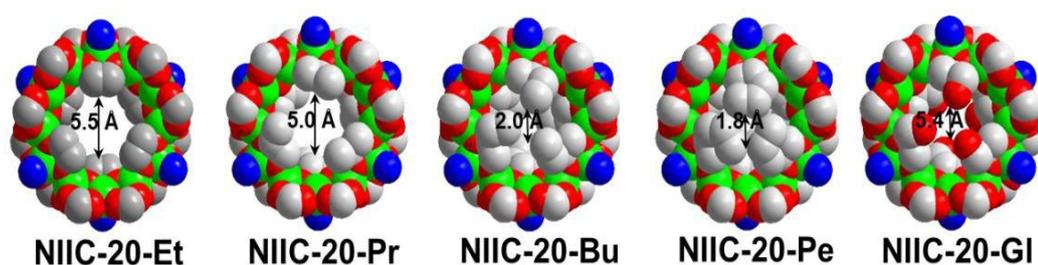
In general, the five different materials have similar structure and composition, so it is not necessary to analyze them individually. NIIC-20-Et will be used as a representative example [63].

The structure of the dodecanuclear wheel consists of two types of crystallographically independent Zn(II) centers. One tetrahedral Zn(II) center coordinates with two oxygen atoms from two iph<sup>2-</sup> anions and two oxygen atoms from two deprotonated glycol molecules. The second type of Zn(II) center has a square-pyramidal geometry, coordinating with two oxygen atoms from two iph<sup>2-</sup> anions, two oxygen atoms from one deprotonated glycol molecule, and one nitrogen atom from a dabco molecule [64]. Twelve Zn(II) centers, alternating between these two types, are connected by twelve carboxylate groups on the outer side and six glycol dianions on the inner side, forming the dodecanuclear wheel with an internal diameter of 5.5 Å (in the case of NIIC-20-Et) (**Figure 1.13**) [63], [64].



**Figure 1.13:** Dodecameric wheel composed of Zn<sup>2+</sup> centers (left); b) the structure of the fundamental nanocell of the NIIC-20-Et framework, with the sphere in the center representing the void space of the cell (middle) [16]; c) the nho-h topology of the overall framework of the metal–organic network NIIC-20-Et (right). Each atom is shown in a different color: zinc (Zn) – green, nitrogen (N) – blue, oxygen (O) – red, carbon (C) – gray; hydrogen (H) is omitted [63].

The organic glycol substituents influence the pores connecting to the cages, affecting at the same time their aperture size and chemical functionality. As the carbon chain length increases, the pore opening decreases, while the hydrophobic character of the framework increases. When glycerol is used as the organic substituent, hydroxyl group interactions and hydrogen bond formation are introduced, enhancing the hydrophilicity of the mesoporous structure. Based on this logic, the material with the largest pore opening among the five is NIIC-20-Et, which contains the smallest ethylene glycol molecules. Conversely, the addition of alkyl-substituted glycols, as in NIIC-20-Pr, NIIC-20-Bu, and NIIC-20-Pe, significantly reduces the  $Zn_{12}$  wheel opening to 1.8 Å, increasing hydrophobicity [63], [64]. The pore openings of the dodecanuclear wheels for the five different NIIC-20-G networks are shown in **Figure 1.14**.



**Figure 1.14:** Openings of the dodecameric wheels for each material of the NIIC-20-G series.

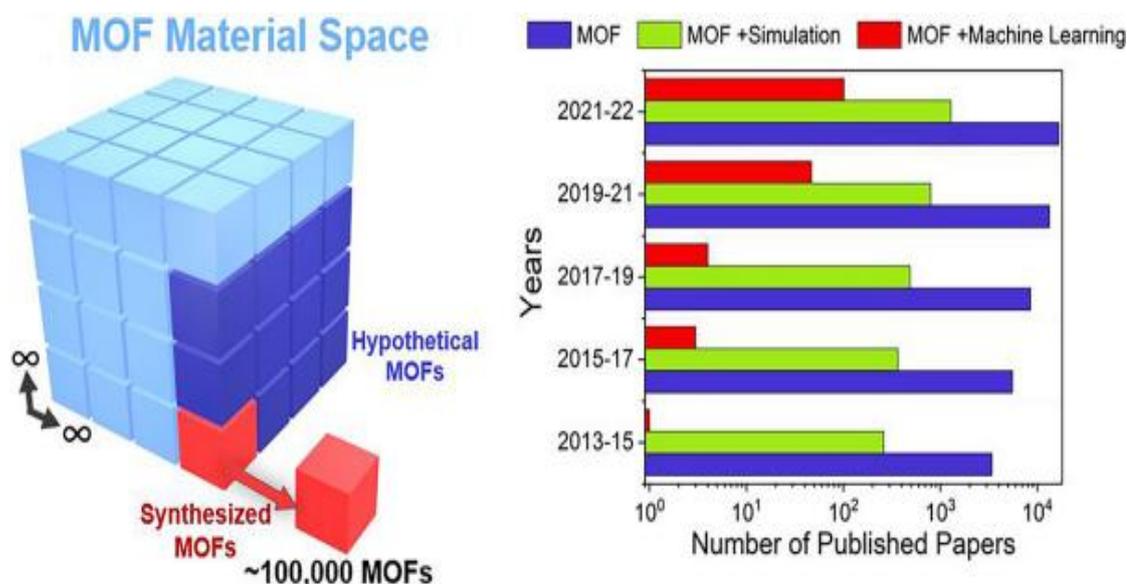
#### 1.4.2.1 NIIC 20 Bu

In this PhD thesis, the NIIC-20-Bu structure was studied. Among the five different networks, NIIC-20-Bu has the second smallest dodecanuclear window opening, due to the folding of the 1,2-butanediol groups. Nevertheless, Lysova et al. [63] observed that NIIC-20-Bu shows very promising performance in the separation of alkane/alkene gases. It demonstrated the highest ethane/ethylene adsorption selectivity among the five mesoporous structures. Lysova et al. [63] concluded that the ideal balance between the geometry of its windows and the satisfactory hydrophobicity of the NIIC-20-Bu structure makes it ideal for intramolecular alkane–alkene interactions, and more specifically, for ethylene exclusion. This characteristic, together with the enhanced adsorption capacity of this MOF, is the subject of the current thesis. It was investigated using various molecular modeling and ANNs, the basic principles of which are discussed in the following chapters.

## 1.5 Molecular Simulations and Machine Learning methods for predicting sorption and diffusion of organic molecules in porous materials

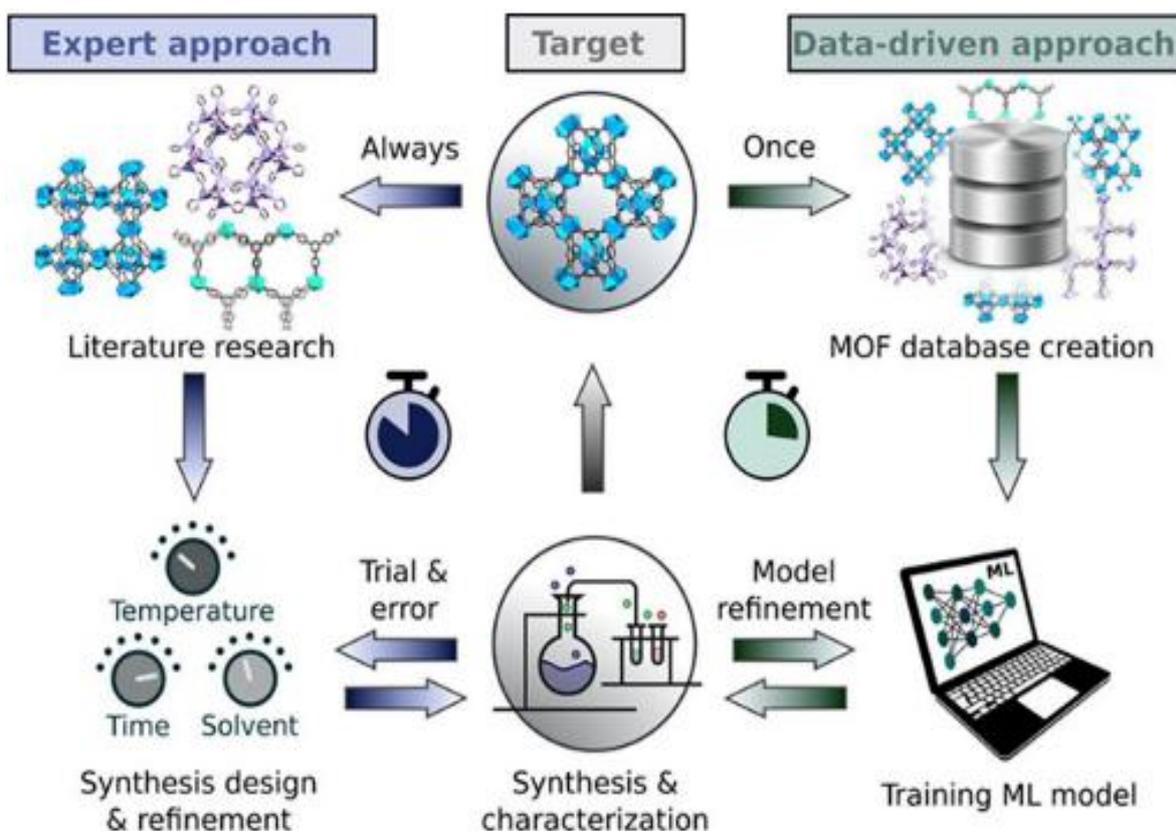
Molecular simulations methods, such as Monte Carlo (MC), equilibrium molecular dynamics (EMD), and nonequilibrium molecular dynamics (NEMD) simulations, have become indispensable tools for calculating gas adsorption and separation within MOFs. Initially, these simulations were limited to a few MOFs, but recent advancements, particularly in high-throughput computational screening (HTCS) techniques, have enabled the exploration of a much larger pool of MOFs [65], [66]. The primary objective behind employing HTCS methods is to efficiently assess the suitability of numerous structures for specific applications and to pinpoint the most promising candidates for further experimental investigation. Consequently, the advancement of computational resources is expected to play a crucial role in expanding the scope of HTCS studies, generating even more extensive datasets.

However, the exponential growth in the number of MOFs poses a significant challenge. Conducting molecular simulations for every conceivable material would be impractical. Therefore, there is a pressing need for data-driven approaches that can streamline the process of material discovery, minimizing the reliance on exhaustive simulations (particularly atomistic, molecular) and experiments. Even if it were feasible to simulate every material and application, the sheer volume and complexity of the resulting data would present a formidable obstacle without the aid of data science techniques. This challenge is often likened to finding a needle in a haystack, as traditional methods rely heavily on empirical, theoretical, or simulation data, often employing brute-force approaches or expert intuition, which are inherently limited. In contrast, the integration of data science into material research represents a paradigm shift, enabling a systematic approach to extracting and processing data from various sources, including experimental, theoretical, and simulation data, to establish machine learning (ML)-based workflows. This holistic approach holds the promise of revolutionizing material discovery by leveraging the power of data analytics to uncover hidden insights and accelerate the identification of novel materials with tailored properties (**Figure 1.15**).



**Figure 1.15:** Number of research papers utilizing classical simulations and machine learning workflows [67].

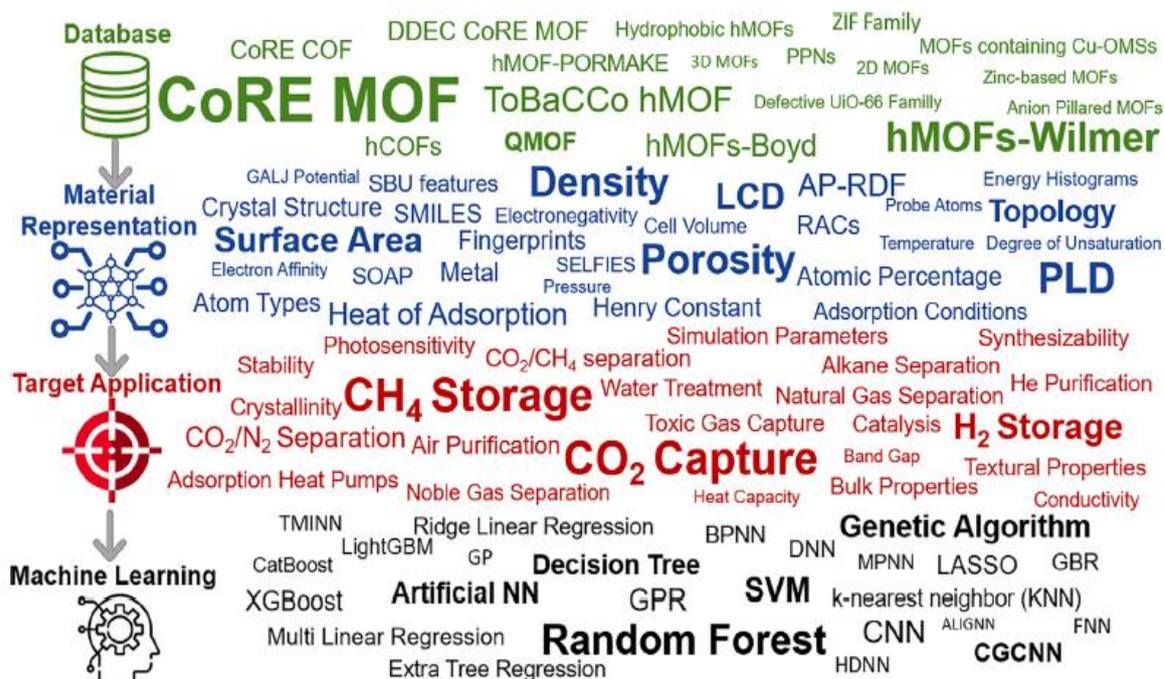
Driven by the immense potential of integrating ML techniques into the computational modeling of MOFs, this PhD thesis aims to underscore various approaches that amalgamate molecular simulations with ML methods. Our focus is on accurately using both the properties and potentials of MOFs across diverse processes, including gas storage, separation, catalysis, as well as predicting synthesizability, guest accessibility, and stability. We commence by outlining the evolutionary trajectory of computational modeling of MOFs, tracing its journey from molecular simulations of a select few materials to the adoption of high-throughput computational screening (HTCS) approaches and, ultimately, the incorporation of ML-guided modeling techniques. Furthermore, we delve into recent pioneering studies that have contributed to material design by introducing novel material representations, curating expansive material databases, employing innovative learning methodologies, and crafting ML workflows.



**Figure 1.16:** Typical Machine Learning workflow incorporated in MOF modelling [68].

In recent years, there has been an increase in interest in using machine learning and artificial intelligence techniques to model the sorption and diffusion of hydrocarbons, such as alkanes, alkenes, and their combinations, within MOFs [69]. These computational methods provide a useful tool for unraveling the complicated interactions between hydrocarbons and MOF structures, revealing adsorption mechanisms, guest-host interactions, and diffusion paths. Machine learning algorithms can detect detailed patterns and correlations in the massive volumes of experimental and computational data available on MOFs and hydrocarbon adsorption, resulting in accurate predictions of sorption isotherms, diffusion coefficients, and breakthrough behaviors. Furthermore, these models can capture the effects of pore size, surface chemistry, and framework flexibility on hydrocarbon absorption and transport in MOFs, allowing for the design of MOF materials customized to specific hydrocarbon storage and separation applications. Researchers are expanding our understanding of hydrocarbon-MOF interactions by combining machine learning with MOF chemistry, paving the door for the development of next-generation materials with unrivaled performance in energy, environmental, and industrial applications. **Figure 1.17** illustrates the four key components of

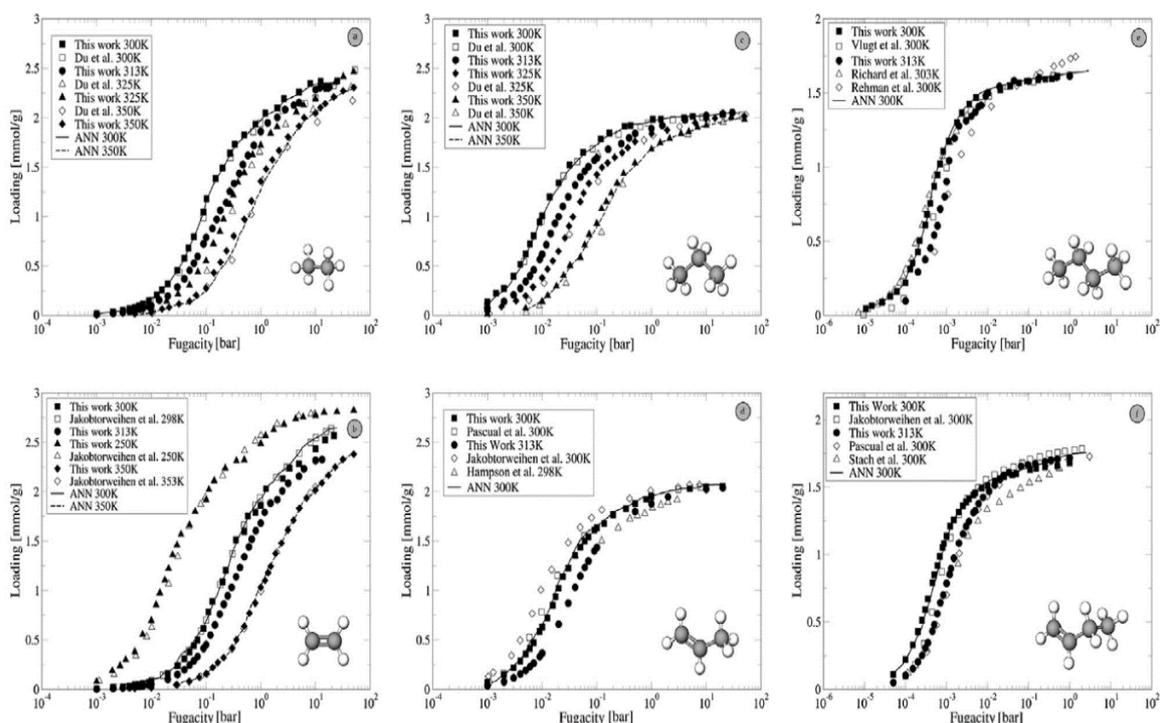
ML model development: databases, material features, target data, and ML algorithms. All these components combined create the basis for any ML/AI or neural network modelling workflow.



**Figure 1.17:** Word cloud representations for the components of a ML model development process. The four main components are database, material representation, target data, and ML algorithm [68].

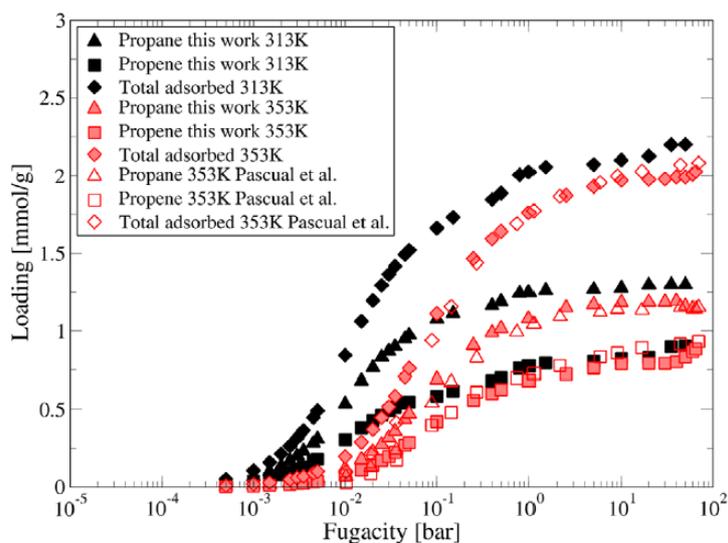
## 1.6 Motivation

This PhD thesis was ignited by our previous work in where modeling [70], [71] and ANNs models were used for predicting sorption and diffusion of organic molecules in porous materials. In particular, in this research endeavor [72] Molecular dynamics (MD) in canonical (NVT) statistical ensemble and grand canonical-Monte Carlo (GCMC) simulations along with artificial neural networks (ANNs) techniques were used for the study of diffusion and sorption characteristics of small alkanes, alkenes, and their mixtures in the zeolite-silicalite. Sorption isotherms along with their respective predictions can be shown in **Figures 1.18, 1.19, 1.20, 1.21** and self-diffusion coefficients **Figure 1.22** of alkanes, alkenes and the respective alkane–alkene mixtures (consisting of the same number of carbon atoms) in silicalite. The findings were directly compared with recent magic-angle spinning pulsed field-gradient nuclear magnetic resonance experimental diffusivity measurements and were found in good agreement. Furthermore, new results during this research were provided for the alkane–alkene systems.



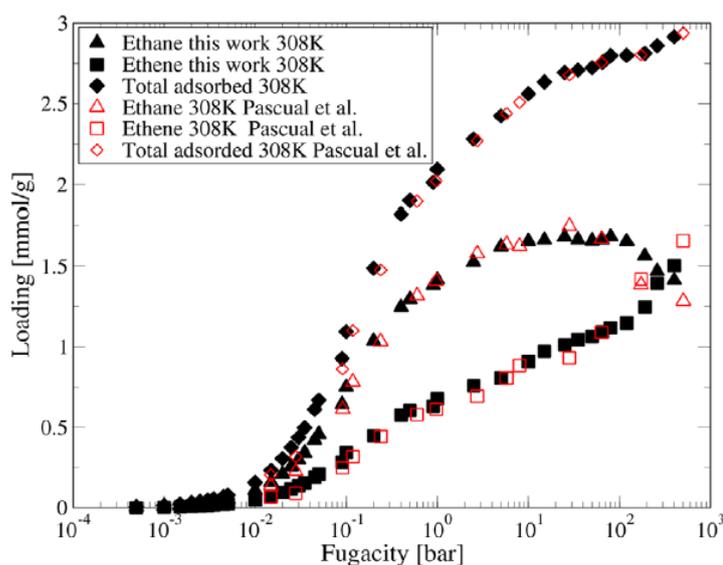
**Figure 1.18:** Sorption isotherms of short pure alkanes and pure alkenes in silicalite at different temperatures: a) ethane, b) ethene, c) propane, d) propene, e) butane, f) butene. Diagrams show results from the present GCMC simulations, experimental, and simulation results from the literature and the ANN predictions (with continuous line format).

In **Figure 1.18** the results from the pure components a) ethane, b) ethene, c) propane, d) propene, e) butane, f) butene) are shown alongside with available experimental data as well as the predictions from the ANNs. As it is evident from the sorption isotherms our model captured the trend and were in good agreement with the experimental data from literature. Also, a significant achievement was the utilization of ANNs for the accurate prediction of the sorption of pure components which demonstrated the capabilities of the AI methodology which at that time was one of the first applications of ANN in the field of porous materials.



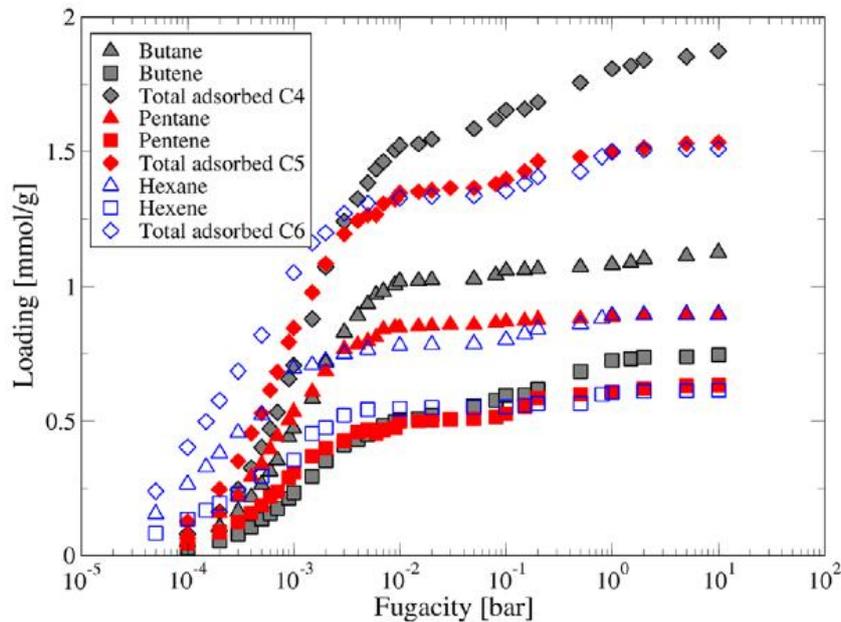
**Figure 1.19:** Propane–propene mixture sorption isotherms at 313 and 353 K.

In **Figure 1.19** the sorption isotherm of the binary mixture of propane-propene was calculated. The propane–propene mixture isotherms were computed at 313 K and additionally at 353 K for comparison with other simulation data. The results from the two simulations at 353 K are in good agreement. It was observed that throughout the fugacity range propane molecules are sorbed in higher amounts than the propene molecules. Around the fugacity value of 100 bar the total adsorbed quantity was around 2 mmol/g at 353 K and 2.25 mmol/g at 313 K. It is also clear that the lower the temperature (313 K) the higher the hydrocarbon amount sorbed.



**Figure 1.20:** Ethane-Ethene mixture sorption isotherms at 313 and 353 K.

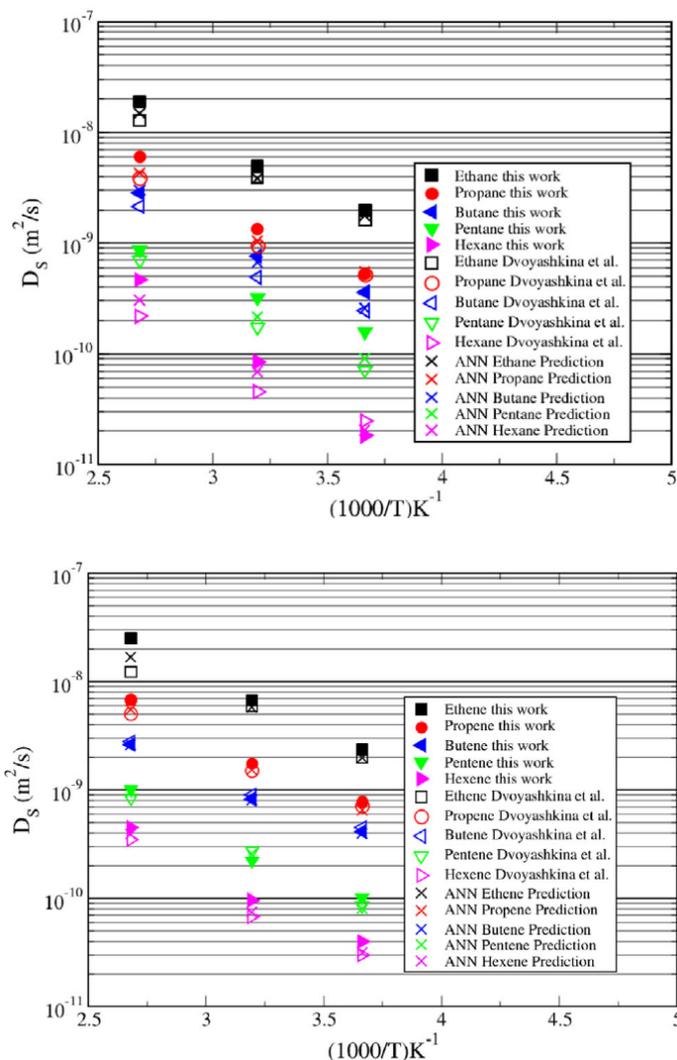
The sorption isotherm of the binary mixture of ethane-ethene was calculated and is shown in **Figure 1.20**. For the ethane-ethene mixture it was observed that for low, intermediate, and for higher fugacity values (up to 200 bar) ethane molecules are adsorbed in greater amounts compared to their ethene counterpart. A crossover to higher ethene sorbed quantities compared to ethane is evident for fugacity values beyond 200 bar as can be seen in **Figure 1.20** and captured well by the two simulations. The higher sorbed amounts are located between 2.75 and 3 mmol/g for fugacity values just below the extreme fictitious pressure of 500 bar.



**Figure 1.21:** Butane–butene, pentane–pentene, hexane–hexene mixture sorption isotherms at 313 K.

In **Figure 1.21** the sorption isotherms for three different systems at 313K are shown (butane-butene, pentane-pentene, hexane-hexene). The results showed that both butane and butene in lower fugacity values ( $< 5 \times 10^{-4}$  bar) are sorbed in similar amounts. For intermediate values of fugacity and especially in the range of ( $5 \times 10^{-3} - 4 \times 10^{-2}$  bar) the ratio of the butane over butene sorbed amounts is 1.78. Similar behavior is exhibited by the pentane–pentene mixture isotherms at 313 K. It was observed in **Figure 1.21** that pentane molecules are adsorbed in higher amounts compared to the pentene counterpart reaching a total adsorbed value of 1.53 mmol/g at 10 bar. Finally, the hexane–hexene sorption isotherms at 313 K is presented in **Figure 1.21** presenting a similar behavior with the pentane–pentene mixture isotherms. The hexane molecules are sorbed in higher amounts than the hexene molecules

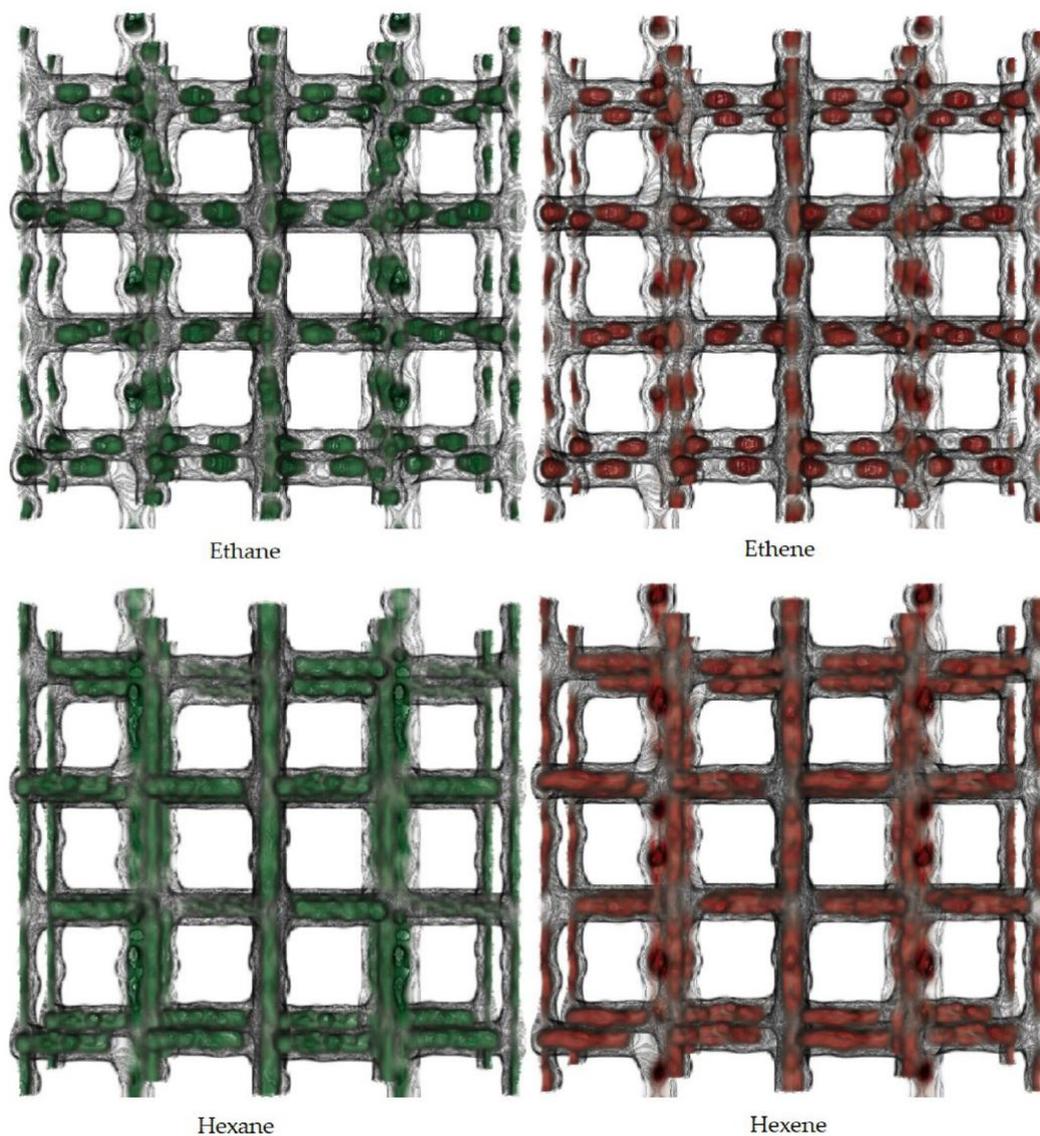
(hexane amount sorbed over hexene amount sorbed is 1.46) reaching a total saturation value of 1.51 mmol/g at 10 bar with the ratio.



**Figure 1.22:** Self-diffusivity ( $D_s$ ) coefficients for pure alkanes (upper), alkenes (lower) ( $N_C = 2 - 6$ ) at 273, 313, and 373 K obtained from the MD simulations (filled symbols). Experimental MAS PFG NMR results by Dvoyashkina et al. [73] are also shown (open symbols). The experimental sorbate loading range was  $0.1 < L < 0.3$ . ANN predictions are shown with x symbol.

In **Figure 1.22** the self-diffusion coefficients for pure alkanes and alkenes are shown at different temperatures, along with experimental data from Dvoyashkina et al. [73]. The MD simulations captured the trend and showed qualitative agreement with the experimental measurements. One of the most important types of information shared in this graph is the predictions made by the ANN, where it's quite clear that the predictions are influenced by both the simulation data but also the experimental ones.

Finally, the structural characteristics of pure alkane, alkene, and alkane-alkene mixtures within the silicalite framework were examined. Through computational analyses, distinct sorption domains and binding site preferences within these confined systems were revealed. This study highlighted the complex interactions between molecular species and the framework's pores, offering valuable insights into confined molecular behavior with relevance to applications such as gas separation and storage in porous materials



**Figure 1.23:** Density plots for pure ethane, ethene, hexane, hexene at 300 K. Green color is used for the representation of ethane, hexane sorption, while red represents the ethene and hexene sorption.

The results presented in this chapter are included in the article that resulted from the research titled “*Modeling Sorption and Diffusion of Alkanes, Alkenes, and their Mixtures in Silicalite: From MD and GCMC Molecular Simulations to Artificial Neural Networks*” by Gurras and Gergidis (Adv. Theory Simul., 4, 2000210, 202) doi: 10.1002/adts.202000210

In light of the findings discussed above, the research direction is now being extended to include the study of different materials, with a particular focus on Metal-Organic Frameworks. Specifically, ZIF-8 and the recently introduced NIIC-20-Bu have been identified as promising candidates for further investigation due to their potential in advanced sorption and separation processes

Metal-Organic Frameworks (MOFs) attracted significant attention in recent years due to their remarkable structural versatility and potential applications in various fields including gas storage, separation, and catalysis. Particularly, the investigation of alkane, alkene sorption and self-diffusion within MOF frameworks stands as a pivotal area of research, crucial for advancing our understanding of host-guest interactions at the molecular level. The sorption and diffusion behaviors of small alkanes, such as methane, ethane, ethene, and propane, and their mixtures within MOFs are of great importance owing to their relevance in numerous industrial processes.

Understanding the fundamental mechanisms governing the sorption and diffusion of these molecules in MOFs is essential for the design and optimization of MOF-based materials tailored for specific applications. This doctoral thesis endeavors to investigate the detailed coupling mechanisms between MOF structure and alkane/alkene sorption/self-diffusion kinetics through a hybrid approach utilizing molecular simulations and ANNs. MD simulations provide a detailed atomistic understanding of the dynamic behavior of hydrocarbons within MOFs, identifying and analyzing the key factors governing diffusion behavior. Monte Carlo simulations complement MD by offering insights into the thermodynamic aspects of alkane sorption, aiding in the prediction of equilibrium uptake capacities. In conjunction with molecular modelling, ANNs serve as powerful predictive tools capable of capturing complex nonlinear relationships between MOF structure and alkane/alkene sorption/diffusion. By training ANNs on extensive datasets generated from molecular simulations and experimental data found in literature, accurate predictions of sorption isotherms and self-diffusion coefficients can be attained, boosting materials screening and design processes. Drawing upon

the findings and discussions presented thus far, the primary objectives of this research are outlined as follows:

1. Investigate the influence of MOF topology, pore size, and functionalization on alkane sorption and diffusion kinetics.
2. Develop robust ANN models for predicting sorption and self-diffusion coefficients in diverse MOF systems.
3. Validate the predictive capabilities of the developed ANN models against experimental and simulation data, where available.
4. Gain insights into the underlying molecular mechanisms governing hydrocarbon sorption and diffusion in MOFs through comprehensive analysis of simulation results.

## References

- [1] E. A. Tomic, "Thermal stability of coordination polymers," *J Appl Polym Sci*, vol. 9, no. 11, pp. 3745–3752, 1965. doi: 10.1002/app.1965.070091121
- [2] A. Schoedel, "Chapter 2 - Secondary building units of MOFs," in *Metal-Organic Frameworks for Biomedical Applications*, M. Mozafari, Ed., Woodhead Publishing, 2020, pp. 11–44. doi: <https://doi.org/10.1016/B978-0-12-816984-1.00003-2>.
- [3] I. Erucar and S. Keskin, "High-Throughput Molecular Simulations of Metal Organic Frameworks for CO<sub>2</sub> Separation: Opportunities and Challenges," *Front Mater*, vol. Volume 5-2018, 2018, doi: 10.3389/fmats.2018.00004.
- [4] S. Soni, P. Bajpai, and C. Arora, "A review on metal-organic framework: synthesis, properties and application," *Characterization and Application of Nanomaterials*, vol. 2, Nov. 2018, doi: 10.24294/can.v2i2.551.
- [5] L. Jiao, J. Y. R. Seow, W. S. Skinner, Z. U. Wang, and H.-L. Jiang, "Metal–organic frameworks: Structures and functional applications," *Materials Today*, vol. 27, pp. 43–68, 2019, doi: <https://doi.org/10.1016/j.mattod.2018.10.038>.
- [6] P. Rocío-Bautista, I. Taima-Mancera, J. Pasán, and V. Pino, "Metal-Organic Frameworks in Green Analytical Chemistry," *Separations*, vol. 6, no. 3, 2019, doi: 10.3390/separations6030033.
- [7] V. F. Yusuf, N. I. Malek, and S. K. Kailasa, "Review on Metal–Organic Framework Classification, Synthetic Approaches, and Influencing Factors: Applications in Energy, Drug Delivery, and Wastewater Treatment," *ACS Omega*, vol. 7, no. 49, pp. 44507–44531, Dec. 2022, doi: 10.1021/acsomega.2c05310.
- [8] R. F. Mendes and F. A. Almeida Paz, "Dynamic breathing effect in metal-organic frameworks: Reversible 2D-3D-2D-3D single-crystal to single-crystal transformation," *Inorganica Chim Acta*, vol. 460, pp. 99–107, 2017, doi: <https://doi.org/10.1016/j.ica.2016.09.037>.
- [9] J.-R. Li, R. J. Kuppler, and H.-C. Zhou, "Selective gas adsorption and separation in metal–organic frameworks," *Chem. Soc. Rev.*, vol. 38, no. 5, pp. 1477–1504, 2009, doi: 10.1039/B802426J.

- [10] H.-L. Jiang, B. Liu, T. Akita, M. Haruta, H. Sakurai, and Q. Xu, "Au@ZIF-8: CO Oxidation over Gold Nanoparticles Deposited to Metal–Organic Framework," *J Am Chem Soc*, vol. 131, no. 32, pp. 11302–11303, 2009, doi: 10.1021/ja9047653.
- [11] D. Danaci, R. Singh, P. Xiao, and P. A. Webley, "Assessment of ZIF materials for CO<sub>2</sub> capture from high pressure natural gas streams," *Chemical Engineering Journal*, vol. 280, pp. 486–493, 2015, doi: <https://doi.org/10.1016/j.cej.2015.04.090>.
- [12] C. Byrne, A. Ristić, S. Mal, M. Opresnik, and N. Zabukovec Logar, "Evaluation of ZIF-8 and ZIF-90 as Heat Storage Materials by Using Water, Methanol and Ethanol as Working Fluids," *Crystals (Basel)*, vol. 11, no. 11, 2021, doi: 10.3390/cryst11111422.
- [13] C. Hu, J.-D. Xiao, X.-D. Mao, L.-L. Song, X.-Y. Yang, and S.-J. Liu, "Toughening mechanisms of epoxy resin using aminated metal-organic framework as additive," *Mater Lett*, vol. 240, pp. 113–116, 2019, doi: <https://doi.org/10.1016/j.matlet.2018.12.123>.
- [14] X.-S. Xing, Z.-H. Fu, N.-N. Zhang, X.-Q. Yu, M.-S. Wang, and G.-C. Guo, "High proton conduction in an excellent water-stable gadolinium metal–organic framework," *Chem. Commun.*, vol. 55, no. 9, pp. 1241–1244, 2019, doi: 10.1039/C8CC08700H.
- [15] I. Mirkovic, L. Lei, D. Ljubic, and S. Zhu, "Crystal Growth of Metal–Organic Framework-5 around Cellulose-Based Fibers Having a Necklace Morphology," *ACS Omega*, vol. 4, no. 1, pp. 169–175, 2019, doi: 10.1021/acsomega.8b02332.
- [16] H. Yang *et al.*, "Metal–organic framework coated titanium dioxide nanorod array p–n heterojunction photoanode for solar water-splitting," *Nano Res*, vol. 12, no. 3, pp. 643–650, 2019, doi: 10.1007/s12274-019-2272-4.
- [17] F. Ke *et al.*, "Facile fabrication of magnetic metal–organic framework nanocomposites for potential targeted drug delivery," *J. Mater. Chem.*, vol. 21, no. 11, pp. 3843–3848, 2011, doi: 10.1039/C0JM01770A.
- [18] Q. Yang, C.-C. Yang, C.-H. Lin, and H.-L. Jiang, "Metal–Organic-Framework-Derived Hollow N-Doped Porous Carbon with Ultrahigh Concentrations of Single Zn Atoms for Efficient Carbon Dioxide Conversion," *Angewandte Chemie International Edition*, vol. 58, no. 11, pp. 3511–3515, 2019, doi: <https://doi.org/10.1002/anie.201813494>.

- [19] F. Yang, M. Wu, Y. Wang, S. Ashtiani, and H. Jiang, "A GO-Induced Assembly Strategy To Repair MOF Nanosheet-Based Membrane for Efficient H<sub>2</sub>/CO<sub>2</sub> Separation," *ACS Appl Mater Interfaces*, vol. 11, no. 1, pp. 990–997, 2019, doi: 10.1021/acsami.8b19480.
- [20] N. Marshall *et al.*, "Polythiophene Doping of the Cu-Based Metal–Organic Framework (MOF) HKUST-1 Using Innate MOF-Initiated Oxidative Polymerization," *Inorg Chem*, vol. 58, no. 9, pp. 5561–5575, 2019, doi: 10.1021/acs.inorgchem.8b03465.
- [21] J. Pang *et al.*, "Exploring the sandwich antibacterial membranes based on UiO-66/graphene oxide for forward osmosis performance," *Carbon N Y*, vol. 144, pp. 321–332, 2019, doi: <https://doi.org/10.1016/j.carbon.2018.12.050>.
- [22] L. Li, S. Zhang, J. P. Ruffley, and J. K. Johnson, "Energy Efficient Formaldehyde Synthesis by Direct Hydrogenation of Carbon Monoxide in Functionalized Metal–Organic Frameworks," *ACS Sustain Chem Eng*, vol. 7, no. 2, pp. 2508–2515, 2019, doi: 10.1021/acssuschemeng.8b05413.
- [23] S. Wang *et al.*, "Framework disorder and its effect on selective hysteretic sorption of a T-shaped azole-based metal–organic framework," *IUCrJ*, vol. 6, no. 1, pp. 85–95, Jan. 2019, doi: 10.1107/S2052252518015749.
- [24] L. Zhang *et al.*, "Low-Cost and High-Performance Microporous Metal–Organic Framework for Separation of Acetylene from Carbon Dioxide," *ACS Sustain Chem Eng*, vol. 7, no. 1, pp. 1667–1672, 2019, doi: 10.1021/acssuschemeng.8b05431.
- [25] S. Nandi, E. Sharma, V. Trivedi, and S. Biswas, "Metal–Organic Framework Showing Selective and Sensitive Detection of Exogenous and Endogenous Formaldehyde," *Inorg Chem*, vol. 57, no. 24, pp. 15149–15157, 2018, doi: 10.1021/acs.inorgchem.8b02411.
- [26] J. Cheng *et al.*, "Self-assembly of 2D-metal–organic framework/graphene oxide membranes as highly efficient adsorbents for the removal of Cs<sup>+</sup> from aqueous solutions," *RSC Adv.*, vol. 8, no. 71, pp. 40813–40822, 2018, doi: 10.1039/C8RA08410F.
- [27] M. Lu, L. Li, S. Shen, D. Chen, and W. Han, "Highly efficient removal of Pb<sup>2+</sup> by a sandwich structure of metal–organic framework/GO composite with enhanced stability," *New J. Chem.*, vol. 43, no. 2, pp. 1032–1037, 2019, doi: 10.1039/C8NJ05091K.

- [28] H. Qiu *et al.*, “Fabrication of agricultural waste supported UiO-66 nanoparticles with high utilization in phosphate removal from water,” *Chemical Engineering Journal*, vol. 360, pp. 621–630, 2019, doi: <https://doi.org/10.1016/j.cej.2018.12.017>.
- [29] A. A. Pribylov, K. O. Murdmaa, O. V Solovtsova, and M. K. Knyazeva, “Methane adsorption on various metal-organic frameworks and determination of the average adsorption heats at supercritical temperatures and pressures,” *Russian Chemical Bulletin*, vol. 67, no. 10, pp. 1807–1813, 2018, doi: [10.1007/s11172-018-2293-2](https://doi.org/10.1007/s11172-018-2293-2).
- [30] T. Villenoisy *et al.*, “Principles of Design and Synthesis of Metal Derivatives from MOFs,” *Adv Mater*, vol. 35, p. e2210166, Apr. 2023, doi: [10.1002/adma.202210166](https://doi.org/10.1002/adma.202210166).
- [31] X. He, Y. Yu, and Y. Li, “Facile synthesis of boronic acid-functionalized magnetic metal–organic frameworks for selective extraction and quantification of catecholamines in rat plasma,” *RSC Adv*, vol. 8, pp. 41976–41985, Jan. 2018, doi: [10.1039/C8RA07356B](https://doi.org/10.1039/C8RA07356B).
- [32] Y. Yamini and M. Safari, “Magnetic Zink-based metal organic framework as advance and recyclable adsorbent for the extraction of trace pyrethroids,” *Microchemical Journal*, vol. 146, pp. 134–141, 2019, doi: <https://doi.org/10.1016/j.microc.2018.12.059>.
- [33] S. Wang *et al.*, “Two-Dimensional Charge-Separated Metal-Organic Framework for Hysteretic and Modulated Sorption,” *Inorg Chem*, vol. 52, Jan. 2013, doi: [10.1021/ic301781n](https://doi.org/10.1021/ic301781n).
- [34] Y. Wang *et al.*, “Effect of Functional Groups on the Adsorption of Light Hydrocarbons in fmj-type Metal–Organic Frameworks,” *Cryst Growth Des*, vol. 19, no. 2, pp. 832–838, Feb. 2019, doi: [10.1021/acs.cgd.8b01403](https://doi.org/10.1021/acs.cgd.8b01403).
- [35] N. Li, L. Zhou, X. Jin, G. Owens, and Z. Chen, “Simultaneous removal of tetracycline and oxytetracycline antibiotics from wastewater using a ZIF-8 metal organic-framework,” *J Hazard Mater*, vol. 366, pp. 563–572, 2019, doi: <https://doi.org/10.1016/j.jhazmat.2018.12.047>.
- [36] T. Van Tran *et al.*, “MIL-53 (Fe)-directed synthesis of hierarchically mesoporous carbon and its utilization for ciprofloxacin antibiotic remediation,” *J Environ Chem Eng*, vol. 7, no. 1, p. 102881, 2019, doi: <https://doi.org/10.1016/j.jece.2019.102881>.
- [37] J. Sun, X. Zhang, A. Zhang, and C. Liao, “Preparation of Fe–Co based MOF-74 and its effective adsorption of arsenic from aqueous solution,” *Journal of Environmental Sciences*, vol. 80, pp. 197–207, 2019, doi: <https://doi.org/10.1016/j.jes.2018.12.013>.

- [38] T. Chowdhury, L. Zhang, J. Zhang, and S. Aggarwal, "Removal of Arsenic(III) from Aqueous Solution Using Metal Organic Framework-Graphene Oxide Nanocomposite," *Nanomaterials*, vol. 8, no. 12, 2018, doi: 10.3390/nano8121062.
- [39] Y. C. López *et al.*, "Transition metal-based metal-organic frameworks for environmental applications: a review," *Environ Chem Lett*, vol. 19, no. 2, pp. 1295–1334, 2021, doi: 10.1007/s10311-020-01119-1.
- [40] L. Yang *et al.*, "Anion Pillared Metal-Organic Framework Embedded with Molecular Rotors for Size-Selective Capture of CO<sub>2</sub> from CH<sub>4</sub> and N<sub>2</sub>," *ACS Sustain Chem Eng*, vol. 7, no. 3, pp. 3138–3144, 2019, doi: 10.1021/acssuschemeng.8b04916.
- [41] T. Ghanbari, F. Abnisa, and W. M. A. Wan Daud, "A review on production of metal organic frameworks (MOF) for CO<sub>2</sub> adsorption," *Science of The Total Environment*, vol. 707, p. 135090, 2020, doi: <https://doi.org/10.1016/j.scitotenv.2019.135090>.
- [42] L. Huelsenbeck *et al.*, "Modulating and Orienting an Anisotropic Zn-Based Metal Organic Framework for Selective CH<sub>4</sub>/CO<sub>2</sub> Gas Separation," *Crystals (Basel)*, vol. 9, no. 1, 2019, doi: 10.3390/cryst9010020.
- [43] M. S. Alivand, M. Shafiei-Alavijeh, N. H. M. H. Tehrani, E. Ghasemy, A. Rashidi, and S. Fakhraie, "Facile and high-yield synthesis of improved MIL-101(Cr) metal-organic framework with exceptional CO<sub>2</sub> and H<sub>2</sub>S uptake; the impact of excess ligand-cluster," *Microporous and Mesoporous Materials*, vol. 279, pp. 153–164, 2019, doi: <https://doi.org/10.1016/j.micromeso.2018.12.033>.
- [44] S. S. Dhankhar and C. M. Nagaraja, "Construction of a 3D porous Co(ii) metal-organic framework (MOF) with Lewis acidic metal sites exhibiting selective CO<sub>2</sub> capture and conversion under mild conditions," *New J. Chem.*, vol. 43, no. 5, pp. 2163–2170, 2019, doi: 10.1039/C8NJ04947E.
- [45] X. Wei *et al.*, "Poly(deep eutectic solvent)-functionalized magnetic metal-organic framework composites coupled with solid-phase extraction for the selective separation of cationic dyes," *Anal Chim Acta*, vol. 1056, pp. 47–61, 2019, doi: <https://doi.org/10.1016/j.aca.2018.12.049>.
- [46] K. Cheng, Y. Li, Z. Gao, F. Chen, C. You, and B. Sun, "Two-dimensional metal organic framework for effective gas absorption," *Inorg Chem Commun*, vol. 101, pp. 27–31, 2019, doi: <https://doi.org/10.1016/j.inoche.2018.12.003>.

- [47] J. W. Yoon, A.-R. Kim, M. J. Kim, T.-U. Yoon, J.-H. Kim, and Y.-S. Bae, “Low-temperature Cu(I) loading on a mesoporous Metal–Organic framework for adsorptive separation of C<sub>3</sub>H<sub>6</sub>/C<sub>3</sub>H<sub>8</sub> mixtures,” *Microporous and Mesoporous Materials*, vol. 279, pp. 271–277, 2019, doi: <https://doi.org/10.1016/j.micromeso.2018.12.041>.
- [48] X. Zeng, F. Chen, and D. Cao, “Screening metal-organic frameworks for capturing radioactive gas Rn in indoor air,” *J Hazard Mater*, vol. 366, pp. 624–629, 2019, doi: <https://doi.org/10.1016/j.jhazmat.2018.12.042>.
- [49] S. Wang, B. Ye, C. An, J. Wang, and Q. Li, “Synergistic effects between Cu metal–organic framework (Cu-MOF) and carbon nanomaterials for the catalyzation of the thermal decomposition of ammonium perchlorate (AP),” *J Mater Sci*, vol. 54, no. 6, pp. 4928–4941, 2019, doi: [10.1007/s10853-018-03219-4](https://doi.org/10.1007/s10853-018-03219-4).
- [50] S. Choi and M. Oh, “Well-Arranged and Confined Incorporation of PdCo Nanoparticles within a Hollow and Porous Metal–Organic Framework for Superior Catalytic Activity,” *Angewandte Chemie International Edition*, vol. 58, no. 3, pp. 866–871, 2019, doi: <https://doi.org/10.1002/anie.201812827>.
- [51] G. Valdebenito, M. González-Carvajal, L. Santibañez, and P. Cancino, “Metal–Organic Frameworks (MOFs) and Materials Derived from MOFs as Catalysts for the Development of Green Processes,” *Catalysts*, vol. 12, no. 2, 2022, doi: [10.3390/catal12020136](https://doi.org/10.3390/catal12020136).
- [52] J.-J. Yang *et al.*, “Metal-Organic Framework Derived N/C Supported Austenite Nanoparticles as Efficient Oxygen Reduction Catalysts,” *ChemNanoMat*, vol. 5, no. 4, pp. 525–530, 2019, doi: <https://doi.org/10.1002/cnma.201800629>.
- [53] D. Gao *et al.*, “CrPd nanoparticles on NH<sub>2</sub>-functionalized metal-organic framework as a synergistic catalyst for efficient hydrogen evolution from formic acid,” *Chemical Engineering Journal*, vol. 361, pp. 953–959, 2019, doi: <https://doi.org/10.1016/j.cej.2018.12.158>.
- [54] C. Wu, F. Irshad, M. Luo, Y. Zhao, X. Ma, and S. Wang, “Ruthenium Complexes Immobilized on an Azolium Based Metal Organic Framework for Highly Efficient Conversion of CO<sub>2</sub> into Formic Acid,” *ChemCatChem*, vol. 11, no. 4, pp. 1256–1263, 2019, doi: <https://doi.org/10.1002/cctc.201801701>.

- [55] M. Safaei, M. M. Foroughi, N. Ebrahimpour, S. Jahani, A. Omid, and M. Khatami, "A review on metal-organic frameworks: Synthesis and applications," *TrAC Trends in Analytical Chemistry*, vol. 118, pp. 401–425, 2019, doi: <https://doi.org/10.1016/j.trac.2019.06.007>.
- [56] D. Wang, H. Yao, J. Ye, Y. Gao, H. Cong, and B. Yu, "Metal-Organic Frameworks (MOFs): Classification, Synthesis, Modification, and Biomedical Applications," *Small*, vol. 20, no. 47, p. 2404350, 2024, doi: <https://doi.org/10.1002/smll.202404350>.
- [57] S. Dutt, A. Kumar, and S. Singh, "Synthesis of Metal Organic Frameworks (MOFs) and Their Derived Materials for Energy Storage Applications," *Clean Technologies*, vol. 5, pp. 140–166, Jan. 2023, doi: [10.3390/cleantechnol5010009](https://doi.org/10.3390/cleantechnol5010009).
- [58] A. Gkourras and L. N. Gergidis, "Molecular and Artificial Neural Networks Modeling of Sorption and Diffusion of Small Alkanes, Alkenes and Their Ternary Mixtures in Zif-8 at Different Temperatures," *J Phys Chem B*, vol. 126, no. 29, pp. 5582–5594, 2022.
- [59] B. Chen, Z. Yang, Y. Zhu, and Y. Xia, "Zeolitic imidazolate framework materials: Recent progress in synthesis and applications," *J. Mater. Chem. A*, vol. 2, Jul. 2014, doi: [10.1039/C4TA02984D](https://doi.org/10.1039/C4TA02984D).
- [60] Y. V. Kaneti *et al.*, "Strategies for Improving the Functionality of Zeolitic Imidazolate Frameworks: Tailoring Nanoarchitectures for Functional Applications," *Advanced Materials*, vol. 29, no. 38, p. 1700213, 2017, doi: <https://doi.org/10.1002/adma.201700213>.
- [61] H. Bux, F. Liang, Y. Li, J. Cravillon, M. Wiebcke, and J. Caro, "Zeolitic Imidazolate Framework Membrane with Molecular Sieving Properties by Microwave-Assisted Solvothermal Synthesis," *J Am Chem Soc*, vol. 131, no. 44, pp. 16000–16001, Nov. 2009, doi: [10.1021/ja907359t](https://doi.org/10.1021/ja907359t).
- [62] M. C. McCarthy, V. Varela-Guerrero, G. V Barnett, and H.-K. Jeong, "Synthesis of Zeolitic Imidazolate Framework Films and Membranes with Controlled Microstructures," *Langmuir*, vol. 26, no. 18, pp. 14636–14641, Sep. 2010, doi: [10.1021/la102409e](https://doi.org/10.1021/la102409e).
- [63] A. A. Lysova, D. G. Samsonenko, K. A. Kovalenko, A. S. Nizovtsev, D. N. Dybtsev, and V. P. Fedin, "A Series of Mesoporous Metal-Organic Frameworks with Tunable Windows Sizes and Exceptionally High Ethane over Ethylene Adsorption Selectivity," *Angewandte Chemie - International Edition*, vol. 59, no. 46, pp. 20561–20567, Nov. 2020, doi: [10.1002/anie.202008132](https://doi.org/10.1002/anie.202008132).

- [64] A. A. Lysova *et al.*, “Tuning the Molecular and Cationic Affinity in a Series of Multifunctional Metal-Organic Frameworks Based on Dodecanuclear Zn(II) Carboxylate Wheels,” *J Am Chem Soc*, vol. 141, no. 43, pp. 17260–17269, Oct. 2019, doi: 10.1021/jacs.9b08322.
- [65] Q. Shuai, W.-L. Li, C. Zhao, and J. Yu, “Recent advances of computational simulations on carbon capture in MOFs,” *Mater Today Commun*, p. 110050, 2024. doi.org/10.1016/j.mtcomm.2024.110050
- [66] M. Shi, G. Jiang, X. Cheng, X. Zhou, and J. Du, “The capture performance of An-MOF for fission gases: Insight from DFT and AIMD calculations,” *Microporous and Mesoporous Materials*, vol. 363, p. 112838, 2024. doi.org/10.1016/j.micromeso.2023.112838
- [67] H. Demir, H. Daglar, H. C. Gulbalkan, G. O. Aksu, and S. Keskin, “Recent advances in computational modeling of MOFs: From molecular simulations to machine learning,” *Coord Chem Rev*, vol. 484, p. 215112, 2023, doi: <https://doi.org/10.1016/j.ccr.2023.215112>.
- [68] Y. Luo *et al.*, “MOF Synthesis Prediction Enabled by Automatic Data Mining and Machine Learning,” *Angewandte Chemie International Edition*, vol. 61, no. 19, p. e202200242, 2022, doi: <https://doi.org/10.1002/anie.202200242>.
- [69] V. Dao and J. K. Singh, “Accelerating In Silico Discovery of Metal–Organic Frameworks for Ethane/Ethylene and Propane/Propylene Separation: A Synergistic Approach Integrating Molecular Simulation, Machine Learning, and Active Learning,” *ACS Appl Mater Interfaces*, vol. 16, no. 6, pp. 6971–6987, 2024, doi: 10.1021/acsami.3c14505.
- [70] L. N. Gergidis, D. N. Theodorou, and H. Jobic, “Dynamics of n-Butane–Methane Mixtures in Silicalite, Using Quasielastic Neutron Scattering and Molecular Dynamics Simulations,” *J Phys Chem B*, vol. 104, no. 23, pp. 5541–5552, Jun. 2000, doi: 10.1021/jp0000073.
- [71] L. N. Gergidis and D. N. Theodorou, “Molecular Dynamics Simulation of n-Butane–Methane Mixtures in Silicalite,” *J Phys Chem B*, vol. 103, no. 17, pp. 3380–3390, Apr. 1999, doi: 10.1021/jp983680p.
- [72] A. Gurras and L. N. Gergidis, “Modeling Sorption and Diffusion of Alkanes, Alkenes, and their Mixtures in Silicalite: From MD and GCMC Molecular Simulations to Artificial Neural Networks,” *Adv Theory Simul*, vol. 4, no. 3, p. 2000210, 2021, doi: <https://doi.org/10.1002/adts.202000210>.

- [73] N. Dvoyashkina *et al.*, “Alkane/alkene mixture diffusion in silicalite-1 studied by MAS PFG NMR,” *Microporous and Mesoporous Materials*, vol. 257, pp. 128–134, 2018, doi: <https://doi.org/10.1016/j.micromeso.2017.08.015>.
- [74] D. Iliopoulos, Modeling of adsorption and diffusion of low-molecular-weight hydrocarbons and their mixtures in metal–organic porous materials using molecular simulations, M.S. thesis, Dept. of Materials Science and Engineering, University of Ioannina, Ioannina, Greece, 2024

# Chapter 2: Literature Overview, State of the Art and Thesis Contribution

## 2.1 Introduction

In this chapter, the most recent advances in numerical simulation and modeling strategies as they apply to Metal-Organic Frameworks (MOFs) will be presented. This chapter covers a wide range of simulation strategies, starting with traditional molecular simulations and progressing to the rapidly growing fields of machine learning (ML) and artificial intelligence (AI)-driven techniques [1], [2], [3], [4]. The overriding purpose of this literature review chapter is to demonstrate how these computational techniques both classical and data-driven have increased our understanding of MOFs, ranging from their structural features and behavior under varied situations to their prospective applications in sectors as diverse as gas storage, catalysis, and separation technologies [5], [6], [7].

Molecular simulations have long been crucial to the research of MOFs, providing a way to predict, analyze, and lead experimental results. Classical simulation approaches, including molecular dynamics (MD) and Monte Carlo (MC) simulations, have proven useful in predicting structural-properties relations at the molecular level. These approaches allow for a thorough analysis of adsorption processes, guest-host interactions, and gas and liquid transport inside the unit cells of MOF structures. Classical MD and MC techniques [8], [9] have offered significant insights into MOF thermodynamic characteristics, enabling the prediction of adsorption isotherms, selectivity, and diffusion coefficients which are crucial for process engineers to get a hold on. The capacity to replicate these characteristics under varied pressure and temperature, has led to substantial breakthroughs in the design and optimization of MOFs for gas separation and catalytic processes [10], [11], [12].

Despite the benefits they offer, traditional simulation approaches have inherent limits, especially when working with complicated or large-scale systems due to the increasing number of system parameters. The computing expense of accurately simulating massive MOF structures or highly dynamic systems can be prohibitively expensive, and the reliance on predetermined force fields in classical simulations might restrict the accuracy of predictions for new materials. To address these issues, researchers have increasingly resorted to quantum mechanical (QM) approaches like as density functional theory (DFT) [13], [14], [15], [16], which provide a more thorough and basic explanation of MOFs' electronic structure. DFT

simulations have proved critical for understanding MOFs' electrical, optical, and catalytic capabilities, particularly when identifying active catalytic sites or investigating the electronic interactions between metal nodes and organic linkers [17], [18], [19], [20], [21]. It should be noted that DFT simulations can be computationally inefficient or even prohibitive, especially for MOF frameworks consisting of more than 200 atoms.

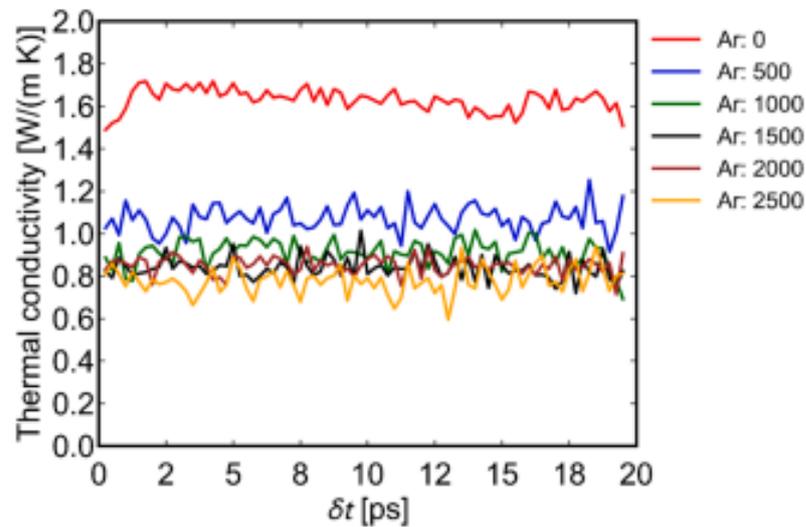
More recently, the incorporation of machine learning (ML) and artificial intelligence (AI) into molecular simulations has created new opportunities for expediting the identification and optimization of MOFs. These methodologies, which rely on big data sets and complex algorithms, have been used to predict MOF characteristics with exceptional speed and precision. ML models may be trained on current experimental or simulation data to predict MOF adsorption behavior, stability, or catalytic activity, avoiding time-consuming classical simulations. In addition to property prediction, ML and AI are increasingly being utilized for inverse design, which allows computers to recommend new MOF structures customized to specific applications based on desired performance parameters [4], [22], [23].

The combination of classical molecular simulations with ML and AI approaches is a significant new approach for the study of small hydrocarbons sorbed in ZIF-8 as demonstrated from our previous work [24]. Researchers may now investigate a considerably broader spectrum of MOF architectures and behaviors by combining the deep atomistic insights afforded by classical and quantum simulations with machine learning's predictive capabilities. This hybrid method has already shown promising results, notably in the fields of gas separation and storage, where ML-driven predictions were utilized to quickly screen hundreds of MOF candidates for high-performance materials [25], [26], [27].

## 2.2 Classical Simulation techniques

In this research endeavor Ito et al. [9] calculated the thermal conductivity of HKUST-1 loaded with water, methanol and ethanol this methodology is of great significance because the calculation of thermal conductivity requires the use of Non-Equilibrium Molecular Dynamics (NEMD). In order to calculate the thermal conductivity (**Figure 2.1**) of MOFs using the Green-Kubo formula and equilibrium molecular dynamics (EMD), the heat autocorrelation function (HCACF) must be truncated at a certain period. Few investigations have been undertaken to provide a technique for truncating HCACFs for MOFs, making it difficult to judge the trustworthiness of MOF thermal conductivities computed using EMD. Previous research focused on total thermal conductivity and did not identify the role of individual components. The authors developed a technique to determine the truncation time of an HCACF

and use it to compute the thermal conductivity of MOFs. After establishing the method's validity, they divided the instantaneous heat flux into distinct components to calculate thermal conductivities. The suggested approach is effective for truncating HCACFs and estimating thermal conductivity of complicated materials like MOFs.

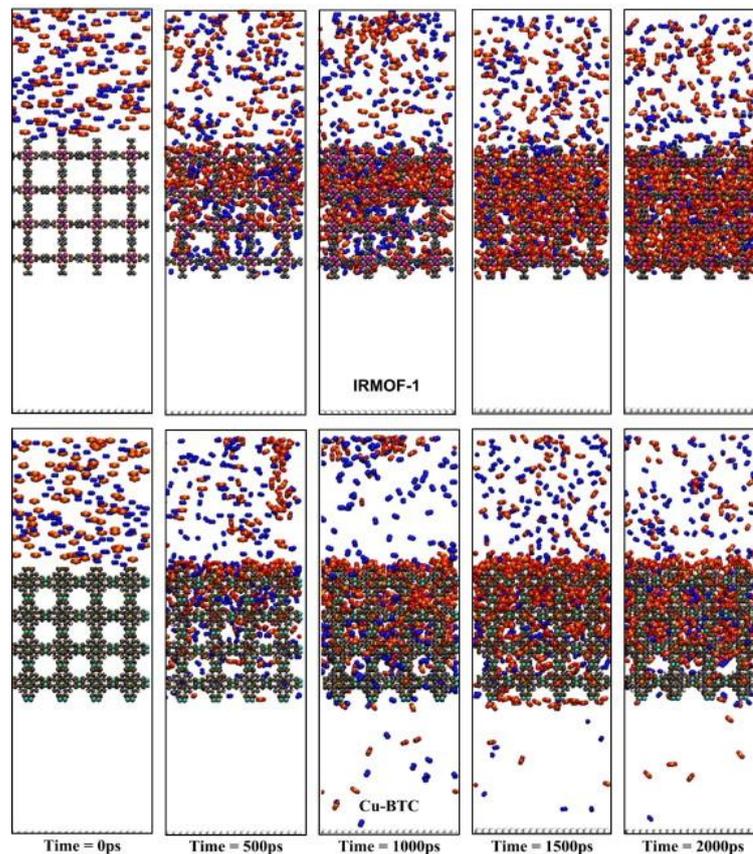


**Figure 2.1:** Effect of timestep  $\delta t$  on thermal conductivity [9].

This work used EMD simulations to analyze the thermal transport characteristics of HKUST-1 loaded with Ar atoms. The authors also explored ways for truncating HCACF for more accurately estimating the MOF's thermal conductivity using the Green-Kubo formula. The study found that existing approaches for determining the cutoff time for integrating HCACFs were not adequate for the MOF HCACF profile [28]. They also developed a novel approach for truncating MOF HCACFs and found that the parameters utilized had no major effect on the outcomes. The thermal conductivity was split into its constituents, and their contributions to the instantaneous heat flux were analyzed. Loading Ar atoms led to decreased thermal conductivity. Despite increasing the amount of loaded Ar atoms, the virial term in heat flux remained dominating over thermal conductivity. Atomic interactions have a greater impact on thermal conductivity than kinetic contributions. However, the decomposed thermal conductivities created uncertainty when determining MOF thermal conductivities using EMD simulations. The suggested approach can address HCACF truncation and determine thermal conductivity of complicated materials like MOFs.

In a different paper Liu et al. [8] investigated the membrane separation of  $\text{CO}_2/\text{N}_2$  and  $\text{CO}_2/\text{CH}_4$  mixtures in five Zr-MOFs (Zr-fum, UiO-66, DUT-52, Zr-cca, and UiO-67) with comparable ligand structures and varying pore-limiting diameters (PLD) ranging from 0.36 nm to 0.59 nm. The gas permeation and separation capabilities are assessed using the CGD-MD

technique. The study found that CO<sub>2</sub> preferentially sorbs over N<sub>2</sub> and CH<sub>4</sub>, leading to the separation of the gas mixtures. Meanwhile, the PLD of MOFs is an important element in penetration. MOFs with bigger PLD exhibit better permeability but worse selectivity for CO<sub>2</sub>/N<sub>2</sub> and CO<sub>2</sub>/CH<sub>4</sub> separation. Calculating free-energy profiles revealed that bigger PLDs result in a reduced energy barrier for gas transport via MOF pores, leading to improved CO<sub>2</sub> separation. The results were compared to the GCMC + EMD methodology, which is a standard method for predicting gas adsorption and diffusion in MOFs. The CGD-MD technique is more reliable than the GCMC+EMD approach because it takes into account component interactions and probable mass transfer resistance at membrane surfaces. This work reveals microscopic insights into CO<sub>2</sub> separation in Zr-MOF membranes, indicating their potential for gas separation. These observations are illustrated in **Figure 2.2**, which shows snapshots of the gas mixture entering the MOF structure and how it evolves over time.

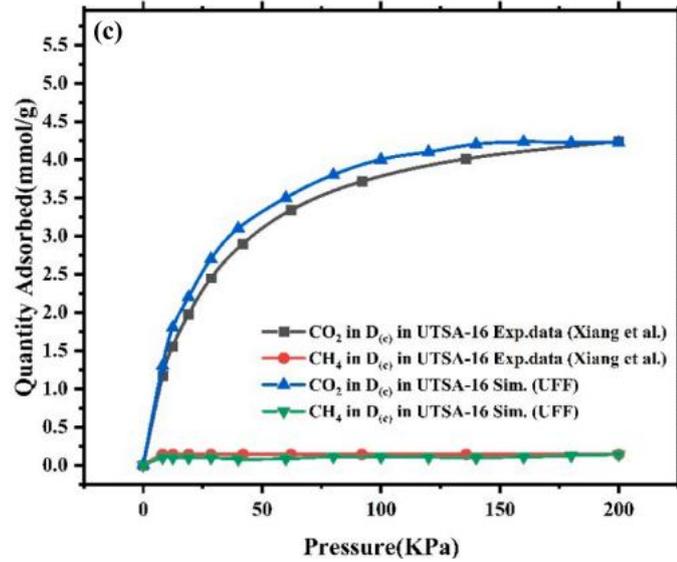


**Figure 2.2:** Snapshots of the IRMOF-1 and Cu-BTC membrane saturation for CO<sub>2</sub>/N<sub>2</sub> gas mixture at 298 K as a function of time from the constant pressure gradient approach MD simulations. Color code for the gas molecules: C (CO<sub>2</sub>), orange; O (CO<sub>2</sub>), red; N, blue; CH<sub>4</sub>, green. For the framework atoms: Zn, magenta; Cu, cyan; C, gray; O, ocher [8].

In addition, this work used molecular modeling to evaluate the gas separation performance of five Zr-MOFs: Zr-fum, UiO-66, DUT-52, Zrcca, and UiO-67. Zr-MOFs' permeation selectivity for CO<sub>2</sub>/N<sub>2</sub> and CO<sub>2</sub>/CH<sub>4</sub> diminishes as their PLD increases, according

to CGD-MD calculations. This is because when steric hindrance decreases, the N<sub>2</sub> and CH<sub>4</sub> gases diffuse faster. Zr-fum has reduced permeation selectivity for CO<sub>2</sub>/N<sub>2</sub> compared to UiO-66, as most CO<sub>2</sub> molecules are adsorbed at the adsorption site. Zr-MOFs exhibit greater permeation selectivity for CO<sub>2</sub>/N<sub>2</sub> compared to CO<sub>2</sub>/CH<sub>4</sub>. Gas diffusion properties in Zr-MOFs were examined using the umbrella sample approach. The energy barrier of Zr-MOF pores decreases with increasing pore size. The energy barrier for three gases is provided as CO<sub>2</sub>>N<sub>2</sub>>CH<sub>4</sub>, which is determined by the interaction between gas molecules and Zr-MOFs. This results in a greater diffusion coefficient for CH<sub>4</sub> and N<sub>2</sub> than for CO<sub>2</sub>. The energy barrier for gases in Zr-fum has the opposite effect due to steric hindrance from the small pore and stiff structure of MOF. The CGD-MD approach yields similar findings to the GCMC+EMD method, except for diffusion selectivity in Zr-fum. The CGD-MD approach uses concentration gradients to overcome steric hindrance in gas molecules with comparable diameters to MOFs. The GCMC+EMD approach cannot precisely assess the gas separation performance of MOF membranes under heavy gas loading.

Sadeghi et al. [29], aimed to validate earlier experimental results and investigate the CO<sub>2</sub> adsorption performance of UTSA-16(Zn) by modifying its primary adsorption sites with alkali and alkaline earth metal cations. The simulation campaign required the use of GCMC and MD simulations to investigate the CO<sub>2</sub> adsorption capacity of freshly developed MOFs. They calculated the Radial Distribution Function (RDF) to determine the principal CO<sub>2</sub> adsorption sites in UTSA-16(Zn). Their data show that the potassium cation (K<sup>+</sup>) is the major adsorption site. They explore how substituting K<sup>+</sup> with different metal cations, including Li<sup>+</sup>, Na<sup>+</sup>, Rb<sup>+</sup>, Be<sup>2+</sup>, Mg<sup>2+</sup>, and Ca<sup>2+</sup>, affects the MOF's CO<sub>2</sub> adsorption capability. Their findings show that Be<sup>2+</sup> and Mg<sup>2+</sup> have an advantageous effect on CO<sub>2</sub> uptake in MOF due to their increased electronegativity compared to other metal cations. Be-UTSA-16(Zn) and Mg-UTSA-16(Zn) outperform UTSA-16(Zn) in CO<sub>2</sub> adsorption across diverse intake feed conditions. When pure CO<sub>2</sub> is supplied, the novel MOFs enhance CO<sub>2</sub> adsorption capacity by 8.81% and 5.37%, respectively, compared to UTSA-16(Zn). These findings offer useful insights into the CO<sub>2</sub> adsorption behavior of recently synthesized MOFs and suggest techniques for creating effective CO<sub>2</sub> capture materials. The sorption isotherm referring to the above observations can be seen in **Figure 2.3**.



**Figure 2.3:** Adsorption isotherms of pure CO<sub>2</sub> in (a) UTSA-16 UTSA-16(Zn) at 298 K in a pressure range of 0 to 100 KPa and adsorption isotherm of CO<sub>2</sub> from CO<sub>2</sub>- CH<sub>4</sub> mixture (CO<sub>2</sub>-CH<sub>4</sub> (50:50)) in UTSA-16 and UTSA-16(Zn) [29].

To summarize, experimental data is utilized to confirm the simulated force field's accuracy and applicability. The CO<sub>2</sub> adsorption behavior of UTSA-16 and UTSA-16(Zn) from pure CO<sub>2</sub> and a combination of CO<sub>2</sub> and CH<sub>4</sub> was found to be more consistent with UFF results. The study uses GCMC and MD simulations to investigate the CO<sub>2</sub> adsorption characteristics of UTSA-16(Zn) when its K adsorption sites are replaced with alkali or alkaline earth metals. The important findings of this study may be summarized as follows:

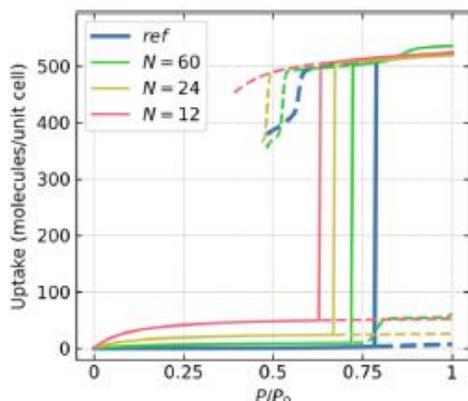
- Zn in UTSA-16(Zn) enhances MOF-adsorbate interactions and increases CO<sub>2</sub> adsorption performance over UTSA-16 alone.
- The investigation reveals that UTSA-16(Zn) is a promising adsorbent for CO<sub>2</sub> capture, but not for SO<sub>2</sub>, NO<sub>2</sub>, or H<sub>2</sub>S. III. K<sup>+</sup> sites in UTSA-16(Zn) exhibit higher CO<sub>2</sub> adsorption compared to Zn.
- The replacement of K<sup>+</sup> cations with Li<sup>+</sup>, Na<sup>+</sup>, Rb<sup>+</sup>, Be<sup>2+</sup>, Mg<sup>2+</sup>, or Ca<sup>2+</sup> in UTSA-16(Zn) reveals that Li<sup>+</sup>, Rb<sup>+</sup>, and Ca<sup>2+</sup> are not efficient in improvement of CO<sub>2</sub> adsorption capacity. Additionally, Na<sup>+</sup> shows a comparable behavior to K<sup>+</sup> in CO<sub>2</sub> adsorption.
- Furthermore, Be<sup>2+</sup> and Mg<sup>2+</sup> exhibit higher CO<sub>2</sub> adsorption capacity compared to UTSA-16(Zn).
- The electronegativity of the metal atoms in MOFs affects their interaction with adsorbate molecules. Alkali metal cations (Li<sup>+</sup>, Na<sup>+</sup>, and K<sup>+</sup>) have lower electronegativities than alkaline earth metal cations (Be<sup>2+</sup>, Mg<sup>2+</sup>, and Ca<sup>2+</sup>).

Generally, this leads to weaker interactions and lower adsorption strengths for alkali metal cations.

Cai et al. [6] used molecular dynamics simulation and density functional theory calculations to investigate CO<sub>2</sub> adsorption and diffusion processes in five distinct systems. Functionalized isorecticular MOFs (IRMOF-1–5), considering the impact of functionalized linkers affect the adsorption capacity of functionalized MOFs. The results demonstrated that CO<sub>2</sub> adsorption is governed by two elements: binding energy and porosity of MOFs. The porosity of the MOFs plays an important role in IRMOF-5, resulting in the lowest level of CO<sub>2</sub> uptake. CO<sub>2</sub> has the highest potential mean force (PMF) at the interface with functionalized MOFs, corresponding to the highest CO<sub>2</sub> density distribution. IRMOF-3 with functionalized linker –NH<sub>2</sub> exhibits the maximum CO<sub>2</sub> adsorption because of its greater porosity and binding energy. IRMOF-5 with the functionalized linker OC<sub>5</sub>H<sub>11</sub> had the lowest CO<sub>2</sub> diffusivity and maximum binding energy, but the lowest CO<sub>2</sub> uptake rate. Among the five simulated functionalized MOFs, IRMOF-3 is an effective CO<sub>2</sub> adsorbent, while IRMOF-5 can segregate CO<sub>2</sub> from other gases, making it useful for constructing CO<sub>2</sub> capture devices. This work demonstrated the use of MD modeling and DFT to evaluate the impact of functionalized linkers on CO<sub>2</sub> adsorption and diffusion in MOFs. CO<sub>2</sub> diffusivities decline in this order: IRMOF-1, IRMOF-2, IRMOF-3, IRMOF-4, and IRMOF-5. The functionalized MOFs showed different binding energies for CO<sub>2</sub>, supporting this conclusion. CO<sub>2</sub> uptake decreases in the following order: IRMOF-3, IRMOF-4, IRMOF-2, IRMOF-1, and IRMOF-5. The binding energy and porosity of MOFs determine their ability to adsorb CO<sub>2</sub>. IRMOF-5 has the lowest CO<sub>2</sub> uptake due to its high adsorbent porosity. CO<sub>2</sub>'s PMF is strongest near the interface, corresponding to the highest CO<sub>2</sub> density distribution. CO<sub>2</sub> molecules travel away from functionalized MOFs, resulting in the greatest PMF in the CO<sub>2</sub>/IRMOF-5 system due to restricted room for movement.

Mazur et al. [5] employed a grand canonical transition matrix Monte Carlo (GC-TMMC) methodology and proposed an efficient interpolation scheme that significantly reduces the number of required simulations while maintaining accuracy of the results. Through the example of water adsorption in three MOFs: MOF-303, MOF-LA2–1, and NU-1000, they show that the extrapolation of the free energy landscape allows for prediction of the adsorption properties over a continuous range of pressure and temperature. This innovative and versatile method provides rich thermodynamic information, enabling rapid, large-scale computational screening of sorbents for adsorption, applicable for a variety of sorbents and gases. This work includes a modified version of the RASPA2 code with a ghost swap move implementation, as

well as a Python library aimed at reducing user input for evaluating data obtained from TMMC simulations, given the robust applicative potential of the proposed technique.



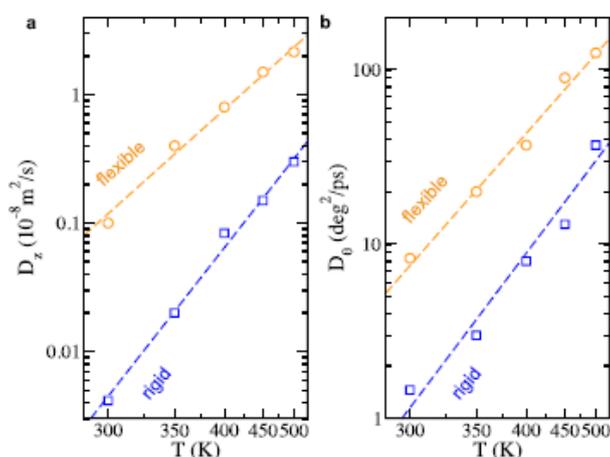
**Figure 2.4:** Isotherm of water adsorption in NU-1000 at 298 K calculated with  $5 \times 10^3$  production cycles and variable number  $N$  of macrostates used in the calculation of direct transition probabilities. Solid and dashed lines correspond to stable and metastable states, respectively.

The paper presents and investigates the practical use of the TMMC technique in the NVT + ghost swap strategy to compute water adsorption isotherms in MOFs. This approach interpolates transition probabilities to decrease processing costs while preserving accurate findings. The TMMC approach offers more thermodynamic insight into adsorbing systems compared to GCMC simulations. This method increases precision and decreases computation time for water adsorption by avoiding simulation concerns with local energy minima. For the first time, probability interpolation with temperature extrapolation was utilized to provide high-quality findings at temperatures higher than those directly simulated. This work includes a modified RASPA2 code with a ghost swap move implementation to make the approach more accessible and applicable. The researchers also offer a Python utility to simplify data analysis from TMMC simulations, reducing user input. This package facilitates probability interpolation and MPD extrapolation, as well as automated computation of isotherms (including data standardization via AIF file).

## 2.3 Machine Learning approach to Classical Molecular simulations

Moving on the next subsection which highlights the usage of Machine learning and Artificial intelligence, Zheng et al. [30] demonstrated how atomistic simulations can benefit from a machine-learning model trained on quantum chemistry data. Simulating  $\text{CO}_2$  chemisorption in MOFs with coordinatively unsaturated metal sites requires adding flexibility. Their study shows that  $\text{CO}_2$  diffusion in a flexible Mg-MOF-74 structure is significantly faster than in a rigid one, refuting earlier calculations that assumed rigid MOFs. This study utilized the Deep

Potential-Smooth Edition (DeepPot-SE) model [31] for developing ML potential. The model has the following features: (1) chemical species and atom coordinates are the only inputs, (2) it can be used for systems of different sizes, (3) the designed descriptor preserves translation, rotation, and permutation symmetries, (4) it is continuously differentiable at the cutoff radius (taken as  $\sim 6$  Å in this work), and (5) it is interfaced with TensorFlow and Largescale Atomic/Molecular Massively Parallel Simulator (LAMMPS). Due to the unavailability of credible force fields for flexible MOFs, most GCMC and MD simulations have focused on rigid MOFs. The flexibility of MOFs can impact gas loading and diffusivity due to corrections for zero-point and thermal energies caused by thermal vibrations.

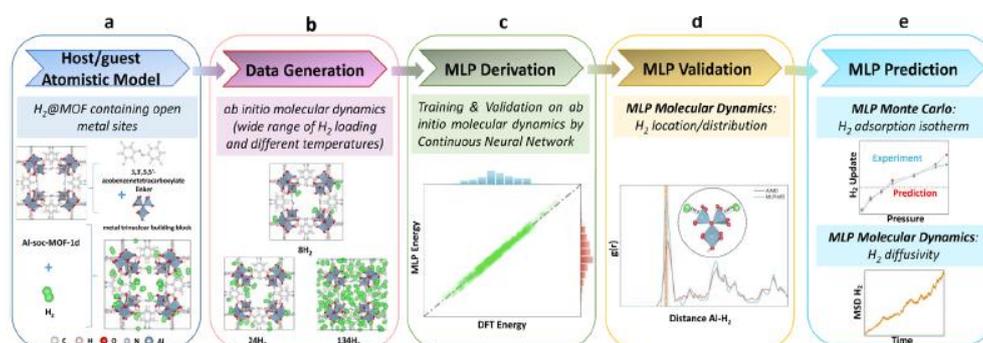


**Figure 2.5:** Diffusion coefficients (in the log scale) vs temperatures (in the reciprocal scale): (a) longitudinal diffusion; (b) angular diffusion. Results for flexible and rigid MOFs are shown in orange and blue, respectively [30].

The theoretical understanding of  $\text{CO}_2$  diffusivity in flexible MOFs is limited. We used an MLP trained on DFT-based MD trajectories to simulate  $\text{CO}_2$  diffusion in Mg-MOF-74. Including the flexibility of Mg-MOF-74 can significantly accelerate  $\text{CO}_2$  diffusion compared to assuming it is rigid. This study highlights that the calculated  $\text{CO}_2$  diffusion in rigid MOFs may differ dramatically from experimental data, particularly when chemisorption occurs.

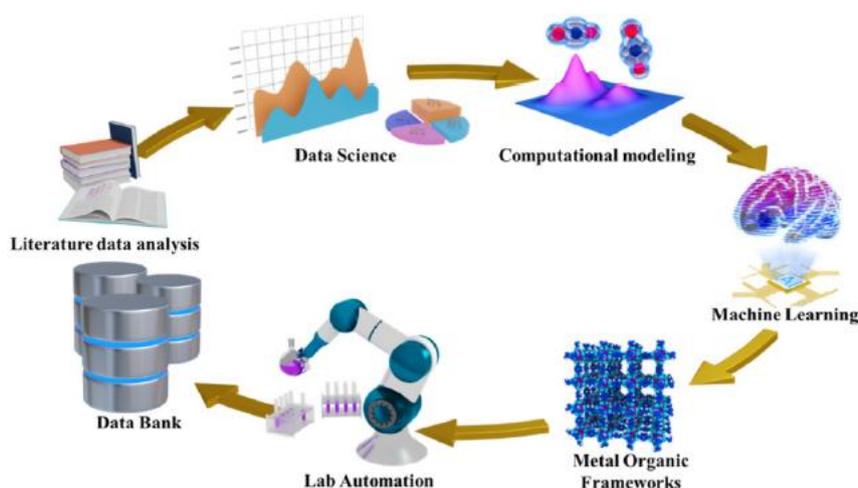
Another recent study is by Liu et al. [2] where their work is the first to precisely characterize the interactions between the difficult  $\text{H}_2$  guest molecule and MOFs with OMS. This unique technique for assessing MOFs containing OMS for  $\text{H}_2$  adsorption and low-pressure capture of different compounds, despite their highly polarizable nature, paves the way for more accurate and efficient *in silico* assessments. The goal of this study is to create a deep neural network-based MLP to study  $\text{H}_2$  adsorption and diffusion in a MOF with OMS. The Al-soc-MOF-1d platform was chosen as a proof-of-concept. It consists of a 6-connected metal trinuclear building block and a 4-connected rectangular-planar organic ligand, 3,30,5,50-

azobenzenetetracarboxylate linker. This MOF is a benchmark material with open metal sites in the Al oxo-trimer nodes and a moderate unit cell size (432 atoms and 10 178 Å<sup>3</sup> cell volume). Its In- and Fe-versions have previously been shown to be effective for H<sub>2</sub> adsorption.



**Figure 2.6** Workflow devised for the derivation and validation of a machine learning potential for H<sub>2</sub>@Al-soc-MOF-1d towards prediction. (a) Construction of the atomistic model for the host/guest system. (b) Generation of the training datasets using ab initio molecular dynamics simulations. (c) Training and validation of the MLP using a continuous neural network. (d) External validation of the trained MLP with its implementation in molecular dynamics simulations and (e) MLP-molecular dynamics and MLP-Monte Carlo predictions [2].

Another study that incorporates ML and AI for CO<sub>2</sub> capture is presented by Mashhadimoslem et al. [1]. They explored methods for training ML algorithms, such as text mining from scientific articles and formula processing for MOFs based on their chemical characteristics. Digitizing scientific documents can improve the efficiency of synthesizing improved MOFs for future ML research. This review essay concludes with a vision for pioneering MOF synthesis pathways for CO<sub>2</sub> collection.



**Figure 2.7:** Schematic view of workflow, material screening using quantum computation and ML. The order and cycle of a pioneering MOF discovery are drawn. Data collection and analysis, atomistic simulation, and data set generation for ML algorithms to discover an advanced MOF structure and its connection with Lab-automation robots to perform experiments are illustrated [1].

## 2.4 Contribution of this thesis

This PhD thesis advances the understanding and application of MD, MC simulations, and ANNs to complex material design and performance prediction challenges. The major contributions derived from this research work are summarized as follows:

1. **An Integrated Approach of MD, MC, and ANNs for MOFs**  
MD and MC simulations have traditionally been used alone or in combination to investigate the thermodynamic and kinetic properties of MOFs. This study blends MD and MC approaches with ANNs models, yielding a hybrid computational methodology that improves the prediction power and efficiency. The incorporation of ANNs into the MD-MC framework is a novel approach aiming to overcome the computational limitations (CPU time and memory consumption) imposed by the conventional simulations of large-scale physical systems modeled in long-time scales. This thesis introduces ANNs models and suggests novel approaches that significantly expedite the prediction of sorption and diffusion characteristics of light hydrocarbons in MOFs under varied situations, particularly in gas adsorption.
2. **Enhanced Predictive Models for Gas Adsorption and Separation.**  
One of the most difficult aspects in MOF research is accurately predicting the adsorption and separation performance of gases, especially in energy-critical applications like carbon capture, hydrogen storage, and methane purification. This thesis improves forecasting capacities in these domains by using ANNs based on MD, MC simulations and available experimental measurements. These models outperform standard simulation approaches in terms of prediction speed. The ability to swiftly screen and forecast MOF properties for various gases and situations speeds up the material discovery process, making it an effective tool for creating MOFs with tailored functionality.
3. **New insights about hydrocarbon adsorption and hydrocarbon dynamics**  
This study provides new insights into the link between hydrocarbon adsorption behavior by simulating small hydrocarbon dynamics in detail. The use of ANNs enables the examination of datasets obtained from MD and MC simulations, revealing previously unknown trends and patterns in MOF behavior. These findings not only improve our fundamental understanding of MOFs, but they also give essential information for enhancing material performance in practical applications.

4. A framework for high-throughput screening of MOFs  
The methodology presented in this thesis adds to the larger field of materials science by offering a solid framework for high-throughput MOF screening. The use of MD, MC, and neural networks allows for rapid evaluation of large MOF datasets, discovering ideal candidates for specific applications without the need for time-consuming and computationally expensive classical simulations. This methodology considerably improves the productivity of the MOF discovery pipeline, especially for applications that demand simultaneous optimization of various features, such as selectivity, capacity, and stability under working circumstances.
5. Generalization of Neural Network Models for various MOF applications.  
While numerous prior investigations have focused on applying neural networks to specific MOFs or gas systems, this thesis broadens the approach by creating transferable models that may be used with a variety of MOFs and gas combinations. Advanced machine learning approaches, such as ANNs, allow models to adapt to various structures and different systems. This versatility broadens the scope of the suggested methodology, making it a significant tool for researchers working in a variety of MOF research fields, such as catalysis.
6. Contributions to the Development of Computational Tools for MOF Research  
This thesis not only introduces a novel technique to study sorption and diffusion of hydrocarbons in MOFs, but it also contributes to the overall development of computational tools in the field. The inclusion of the integrated MD-MC-neural network framework on open-source platforms makes these tools available to a broader research community. This study sets the way for future improvements in MOF modeling by simplifying the use of machine learning approaches in conjunction with traditional simulation methods. It also accelerates the pace of discovery in this quickly changing field.

In closing remarks, this PhD thesis contributes to both the theoretical and practical aspects of MOF research in terms of computational advances. The advancement of computational techniques for analyzing MOF characteristics and behaviors opens up and sheds light to new avenues for the design and optimization of MOFs and porous materials in general for a variety of commercial applications but also for the advancement of the field in terms of research applications

## References

- [1] H. Mashhadimoslem *et al.*, “Computational and Machine Learning Methods for CO<sub>2</sub> Capture Using Metal–Organic Frameworks,” *ACS Nano*, vol. 18, no. 35, pp. 23842–23875, 2024, doi: 10.1021/acsnano.3c13001.
- [2] S. Liu *et al.*, “Machine learning potential for modelling H<sub>2</sub> adsorption/diffusion in MOFs with open metal sites,” *Chem. Sci.*, vol. 15, no. 14, pp. 5294–5302, 2024, doi: 10.1039/D3SC05612K.
- [3] H. Demir and S. Keskin, “A New Era of Modeling MOF-Based Membranes: Cooperation of Theory and Data Science,” *Macromol Mater Eng*, vol. 309, no. 1, p. 2300225, 2024. doi.org/10.1002/mame.202300225
- [4] S. Zhang, Y. He, Z. Zhang, and C. Zhong, “Machine learning aided investigation on the structure-performance correlation of MOF for membrane-based He/H<sub>2</sub> separation,” *Green Chemical Engineering*, 2024. doi.org/10.1016/j.gce.2024.01.005
- [5] B. Mazur, L. Firlej, and B. Kuchta, “Efficient Modeling of Water Adsorption in MOFs Using Interpolated Transition Matrix Monte Carlo,” *ACS Appl Mater Interfaces*, vol. 16, no. 19, pp. 25559–25567, 2024, doi: 10.1021/acsam.4c02616.
- [6] S. Cai, L. Yu, E. Huo, Y. Ren, X. Liu, and Y. Chen, “Adsorption and Diffusion Properties of Functionalized MOFs for CO<sub>2</sub> Capture: A Combination of Molecular Dynamics Simulation and Density Functional Theory Calculation,” *Langmuir*, vol. 40, no. 13, pp. 6869–6877, 2024, doi: 10.1021/acs.langmuir.3c03782.
- [7] A. Granja-DelRio and I. Cabria, “Grand Canonical Monte Carlo simulations of hydrogen and methane storage capacities of two novel Al-nia MOFs at room temperature,” *Int J Hydrogen Energy*, vol. 50, pp. 685–696, 2024. doi.org/10.1016/j.ijhydene.2023.08.023
- [8] F. Cabrales-Navarro, J. Gómez-Ballesteros, and P. Balbuena, “Molecular dynamics simulations of metal-organic frameworks as membranes for gas mixtures separation,” *J Memb Sci*, vol. 428, pp. 241–250, Sep. 2013, doi: 10.1016/j.memsci.2012.10.058.
- [9] H. Ito, K. Fujiwara, and M. Shibahara, “Decomposition and determination of thermal conductivity of MOFs with fluid molecules via equilibrium molecular dynamics,” *Int J Heat Mass Transf*, vol. 228, p. 125676, 2024, doi: https://doi.org/10.1016/j.ijheatmasstransfer.2024.125676.
- [10] L. zhang, Q. Li, Z. Lu, and X. Wang, “Understanding Adsorption and Separation behavior of shorter chain alkane mixtures in Zeolitic Imidazolate Frameworks by molecular simulations,”

- Procedia Comput Sci*, vol. 4, pp. 1193–1202, 2011, doi: <https://doi.org/10.1016/j.procs.2011.04.128>.
- [11] X. Gong, Y. Wang, and T. Kuang, “ZIF-8-Based Membranes for Carbon Dioxide Capture and Separation,” *ACS Sustain Chem Eng*, vol. 5, no. 12, pp. 11204–11214, 2017, doi: [10.1021/acssuschemeng.7b03613](https://doi.org/10.1021/acssuschemeng.7b03613).
- [12] K. Li *et al.*, “Zeolitic Imidazolate Frameworks for Kinetic Separation of Propane and Propene,” *J Am Chem Soc*, vol. 131, no. 30, pp. 10368–10369, 2009, doi: [10.1021/ja9039983](https://doi.org/10.1021/ja9039983).
- [13] V. C. Anadebe, V. I. Chukwuike, K. C. Nayak, E. E. Ebenso, and R. C. Barik, “Combined electrochemical, atomic scale-DFT and MD simulation of Nickel based metal organic framework (Ni-MOF) as corrosion inhibitor for X65 pipeline steel in CO<sub>2</sub>-saturated brine,” *Mater Chem Phys*, vol. 312, p. 128606, 2024. doi.org/10.1016/j.matchemphys.2023.128606
- [14] M. U. C. Braga, G. H. Perin, L. H. de Oliveira, and P. A. Arroyo, “DFT calculations for adsorption of H<sub>2</sub>S and other natural gas compounds on M-BTC MOF clusters,” *Adsorption*, pp. 1–17, 2024. doi: [10.21203/rs.3.rs-3098135/v1](https://doi.org/10.21203/rs.3.rs-3098135/v1)
- [15] Q. Shuai, W.-L. Li, C. Zhao, and J. Yu, “Recent advances of computational simulations on carbon capture in MOFs,” *Mater Today Commun*, p. 110050, 2024. doi.org/10.1016/j.mtcomm.2024.110050
- [16] M. Shi, G. Jiang, X. Cheng, X. Zhou, and J. Du, “The capture performance of An-MOF for fission gases: Insight from DFT and AIMD calculations,” *Microporous and Mesoporous Materials*, vol. 363, p. 112838, 2024. doi.org/10.1016/j.micromeso.2023.112838
- [17] H.-L. Jiang, B. Liu, T. Akita, M. Haruta, H. Sakurai, and Q. Xu, “Au@ZIF-8: CO Oxidation over Gold Nanoparticles Deposited to Metal–Organic Framework,” *J Am Chem Soc*, vol. 131, no. 32, pp. 11302–11303, 2009, doi: [10.1021/ja9047653](https://doi.org/10.1021/ja9047653).
- [18] J.-R. Li, R. J. Kuppler, and H.-C. Zhou, “Selective gas adsorption and separation in metal–organic frameworks,” *Chem. Soc. Rev.*, vol. 38, no. 5, pp. 1477–1504, 2009, doi: [10.1039/B802426J](https://doi.org/10.1039/B802426J).
- [19] A. Phan, C. J. Doonan, F. J. Uribe-Romo, C. B. Knobler, M. O’Keeffe, and O. M. Yaghi, “Synthesis, Structure, and Carbon Dioxide Capture Properties of Zeolitic Imidazolate Frameworks,” *Acc Chem Res*, vol. 43, no. 1, pp. 58–67, 2010, doi: [10.1021/ar900116g](https://doi.org/10.1021/ar900116g).
- [20] D. Fairen-Jimenez *et al.*, “Flexibility and swing effect on the adsorption of energy-related gases on ZIF-8: combined experimental and simulation study,” *Dalton Trans.*, vol. 41, no. 35, pp. 10752–10762, 2012, doi: [10.1039/C2DT30774J](https://doi.org/10.1039/C2DT30774J).

- [21] M. Haeri-Nejad and M. Eic, "Use of molecular simulation for evaluating adsorption equilibrium of inhalation anaesthetic agents on metal–organic frameworks," *Can J Chem Eng*, vol. 102, no. 4, pp. 1646–1657, 2024. doi.org/10.1002/cjce.25134
- [22] V. Daoo and J. K. Singh, "Accelerating In Silico Discovery of Metal–Organic Frameworks for Ethane/Ethylene and Propane/Propylene Separation: A Synergistic Approach Integrating Molecular Simulation, Machine Learning, and Active Learning," *ACS Appl Mater Interfaces*, vol. 16, no. 6, pp. 6971–6987, 2024. doi: 10.1021/acsami.3c14505
- [23] N. A. Chikaodili *et al.*, "Physisorption or chemisorption: Insight from AI computing model based on DFT, MC/MD-simulation for prediction of MOF-based inhibitor adsorption on Cu in brine solution," *Comput Theor Chem*, p. 114730, 2024. doi.org/10.1016/j.comptc.2024.114730
- [24] A. Gurras and L. N. Gergidis, "Modeling Sorption and Diffusion of Alkanes, Alkenes, and their Mixtures in Silicalite: From MD and GCMC Molecular Simulations to Artificial Neural Networks," *Adv Theory Simul*, vol. 4, no. 3, p. 2000210, 2021, doi: <https://doi.org/10.1002/adts.202000210>.
- [25] H. C. Gulbalkan, A. Uzun, and S. Keskin, "Combining computational screening and machine learning to explore MOFs and COFs for methane purification," *Appl Phys Lett*, vol. 124, no. 20, 2024. doi.org/10.1021/acsami.2c08977
- [26] S. Achour and Z. Hosni, "ML-driven models for predicting CO<sub>2</sub> uptake in metal–organic frameworks (MOFs)," *Can J Chem Eng*, 2024. doi.org/10.1002/cjce.25509
- [27] Z. Deng and L. Sarkisov, "Engineering Machine Learning Features to Predict Adsorption of Carbon Dioxide and Nitrogen in Metal–Organic Frameworks," *The Journal of Physical Chemistry C*, 2024. doi.org/10.1021/acs.jpcc.4c01692
- [28] H. Matsubara, G. Kikugawa, M. Ishikiriyama, S. Yamashita, and T. Ohara, "Equivalence of the EMD- and NEMD-based decomposition of thermal conductivity into microscopic building blocks," *J Chem Phys*, vol. 147, no. 11, p. 114104, Sep. 2017, doi: 10.1063/1.4990593.
- [29] M. Sadeghi, F. Esmailzadeh, D. Mowla, and A. Zandifar, "Improving CO<sub>2</sub> capture in UTSA-16(Zn) via alkali and alkaline earth metal Introduction: GCMC and MD simulations study," *Sep Purif Technol*, vol. 338, p. 126534, 2024, doi: <https://doi.org/10.1016/j.seppur.2024.126534>.
- [30] B. Zheng, G. X. Gu, C. dos Santos, R. Neumann Barros Ferreira, M. Steiner, and B. Luan, "Simulating CO<sub>2</sub> diffusivity in rigid and flexible Mg-MOF-74 with machine-learning force fields," *APL Machine Learning*, vol. 2, no. 2, p. 026115, May 2024, doi: 10.1063/5.0190372.

- [31] B. Zheng *et al.*, “Quantum Informed Machine-Learning Potentials for Molecular Dynamics Simulations of CO<sub>2</sub>’s Chemisorption and Diffusion in Mg-MOF-74,” *ACS Nano*, vol. 17, no. 6, pp. 5579–5587, 2023, doi: 10.1021/acsnano.2c11102.

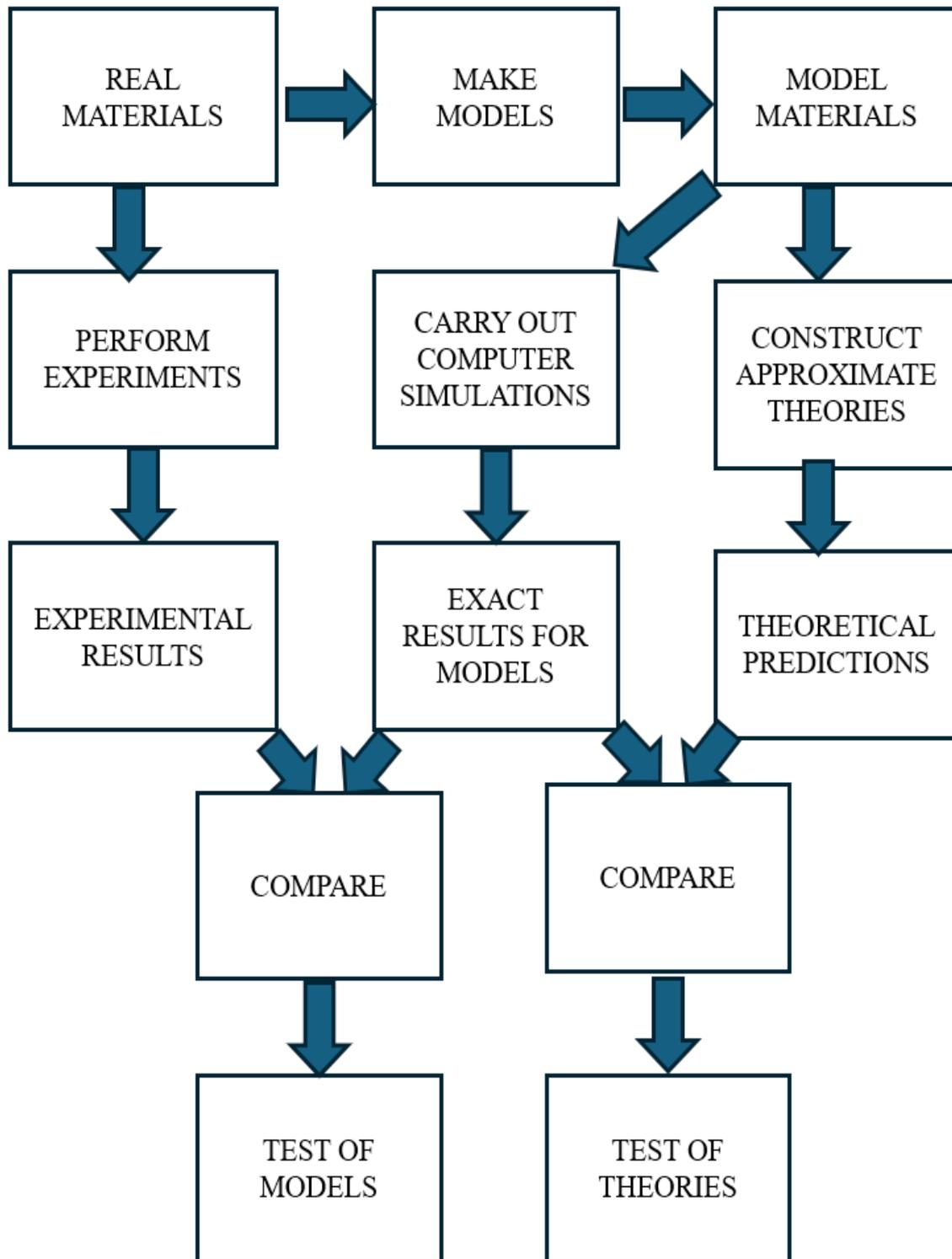
# Chapter 3: Molecular Modelling

## 3.1 Introduction

This chapter provides a background on the theoretical foundations and key computational approaches utilized in this thesis. The primary objective is to introduce and explain the computational techniques employed in the study, with a particular emphasis on Monte Carlo (MC) simulations, specifically the grand canonical Monte Carlo (GCMC) method, and Molecular Dynamics (MD) simulations. These methods are widely used for studying the thermodynamic and transport properties of complex systems at the atomic and molecular levels. Additionally, the chapter will present the techniques and methodologies used for Artificial Neural Network (ANN) modeling, which plays a crucial role in analyzing and predicting system behavior based on computational data. A detailed discussion of the fundamental principles underlying each of these approaches will be provided, along with an explanation of their implementation and relevance to the research objectives. By outlining these methodologies, this chapter aims to establish a solid theoretical framework that will support the subsequent analysis and findings presented in this thesis.

## 3.2 Role of Statistical Mechanics in Materials Design

As shown in **Figure 3.1**, statistical mechanics constitutes the connecting link between information at the molecular level and the macroscopic behavior of a material system. Knowing the chemical composition of a molecule or a group of molecules in a material system, quantum mechanics can determine the potential energy as a function of the molecular configuration (i.e., the set of positions of all nuclei that make up the system). Quantum mechanics can also identify low-energy molecular configurations, thus providing information about molecular geometry (bond lengths, bond angles, torsional angles). Taking this fundamental information on molecular geometry and energies as input, statistical mechanics can predict properties as functions of the macroscopic state of the material system, which consists of the molecules under study. The macroscopic state is defined by variables such as temperature, pressure, composition, mechanical stresses, electric field, etc.



**Figure 3.1:** Schematic of the synergy between theory, experiment, and simulation.

The methods of statistical mechanics can generally be divided into two categories: analytical theories and computational simulations. Although statistical mechanics provides a rigorous formalism for predicting macroscopic properties from molecular geometry and interactions, the exact analytical solution of this formalism is impossible for all but the simplest molecular

models (e.g., a monatomic ideal gas). Statistical mechanical theories introduce approximations either in the molecular model used (representation of molecular shape, spatial arrangement of molecules, and their mutual interactions) or in the mathematical formulation of the problem (e.g., neglecting correlations between the positions of pairs of interacting molecules at large separations), or both. In this way, the theories lead to closed-form systems of algebraic, differential, or integral equations that can be solved analytically or numerically. When the approximations are made with care and physical intuition, they can provide satisfactory solutions at low computational cost. Molecular simulations, on the other hand, can be regarded as techniques for the numerical solution of the full statistical-mechanical formulation, given a model for molecular geometry and interactions. Simulations usually proceed by generating (sampling) a large number of microscopic configurations of the system under study, whose characteristics are analyzed to extract information about the mechanisms responsible for macroscopic behavior and to estimate properties in a manner analogous to that followed in experiments. For this reason, simulations are often referred to as “computational experiments.” When performed correctly, they can yield accurate solutions within the limits of controllable numerical error. Simulations are specialized to specific systems, in contrast to theories, which are usually developed for broad classes of systems. The most important limitations of simulations are mainly associated with their high computational requirements, which hinder studies over large time and length scales. In many cases, however, significant improvements can be achieved through the development of new, more efficient algorithms.

When simulations are applied to the same molecular model as theories, they can reveal shortcomings and deficiencies in the approximations involved in the theories. For this reason, simulations constitute excellent tools for testing theories. Comparisons between simulation results and experimental data under identical macroscopic conditions can reveal deficiencies in the model used in simulations to represent a material system. This synergy among theory, simulation, and experiment (**Figure 3.1**) is evident in modern research based on statistical mechanics

### **3.3 Molecular Dynamics**

Molecular Dynamics (MD) is a molecular modeling method of simulating and investigating a range of time-evolving properties belonging to a system of particles[1], [2], [3], [4], [5]. To determine the microscopic temporal variations of a particle ensemble, the system is described by Newton’s classical equations of motion, which are solved numerically under the influence of boundary conditions appropriate for the symmetry or geometry of the system. Based on

classical mechanics, MD provides insight into the microscopic dynamic behavior of the individual atoms that make up a given system by computing the positions and velocities (momenta) of the constituent particles [1], [2], [6]. We could compare the MD technique to a "computational microscope," whose findings, with the help of statistical mechanics, can be utilized as a sampling method from a statistical mechanical ensemble and determine macroscopic equilibrium concepts. Such concepts include average thermodynamic quantities (pressure, temperature, volume, etc.), structure, and free energies [1], [2], [6]. Of course, to achieve a realistic prediction of the microscopic behavior of a system, it is essential not only to model the dynamic behavior of the molecules but also to simulate their interactions. Theoretically, molecular interactions and structural characteristics can be predicted using quantum mechanics. However, in practice, many of the problems addressed by computational simulations cannot be solved using this principle, primarily due to its extremely high computational cost [7], [8]. In MD, intermolecular interactions within a molecular system are described in a more simplified manner through the introduction of specific input parameters and an interaction potential, commonly known as a force field. Unlike the quantum mechanical approach, MD force fields disregard the electronic degrees of freedom of molecules, considering only the motion of their nuclei. This simplification follows the Born-Oppenheimer approximation, which allows for the separation of nuclear and electronic motions, thereby enabling the expression of a system's energy as a function of nuclear coordinates alone.

Force fields are based on two fundamental assumptions: additivity and transferability. The additivity assumption allows the total energy of any system to be treated as the sum of different potential energy contributions with distinct physical meanings, such as electrostatic forces, dispersion forces, and bond deformations. The transferability assumption, on the other hand, suggests that potential energy functions developed for a small molecular system can be applied to a much broader range of molecules, provided they share similar chemical groups. The accuracy of a Molecular Dynamics simulation strongly depends on the quality of the force field used to describe particle interactions. A well-parameterized force field ensures reliable predictions of molecular behavior, whereas inaccuracies in the interaction potential can lead to significant deviations between simulated and real-world results [2], [6]. MD is a well-established field with a long history of development, evolving into an indispensable computational tool applicable across numerous scientific disciplines, including chemistry, biology, and physics. It has the capability to predict, with great detail, the microscopic dynamic behavior of various types of systems, including gases, liquids, solids, surfaces, and molecular complexes [1], [6], [9], [10].

MD simulations are in many respects similar to real experiments. For example, in an experiment we prepare the material we intend to study. We place the sample in an apparatus connected to a measuring instrument, e.g. a thermometer, a manometer, or a viscometer, and we measure the property of interest over a certain period of time. If the measurements are subject to statistical noise, as is most often the case, we increase the number of measurements and take a time average, thereby obtaining more accurate estimates. In an MD simulation we follow the same procedure. First, we prepare the sample; that is, we choose the physical system and the model that describes it as reliably as possible. Thus, for a system consisting of  $N_{particles}$ , we solve Newton's equations until the properties of the system no longer change with time. Once the system has reached equilibrium, we perform the measurement. In reality, of course, many of the common errors that occur in a computational "experiment" are similar to the errors that can occur in a real experiment. For example, the sample may not have been prepared properly, the measurement may be too short, during its duration the system may undergo some irreversible change, or finally we may not be measuring what we think we are measuring. It should be emphasized at this point that the study of molecular dynamics focuses on classical systems consisting of a large number of particles. By the term classical we mean that the motion of the nuclei of the particles obeys the laws of classical mechanics. This is an exceptionally good approximation for a wide range of materials. Only in the case of translational or rotational motion of light atoms or molecules (He, H<sub>2</sub>, D<sub>2</sub>), or of vibrational motion with frequency  $\nu$  such that  $h\nu \gg k_B T$ , would quantum effects become important.

In the following sections, the fundamental approach of this computational simulation method will be analyzed in detail. This will include an overview of the MD algorithm, the different types of statistical ensembles, and the core principles of interaction potentials.

### 3.3.1 Equations of Motion

In this chapter we will deal with the mathematical description of the dynamics of a classical system. Our goal is the formulation of the equations of motion for a system of  $N$  interacting particles characterized by a potential energy function  $V$ . These equations can be written in various ways. However, the most fundamental formulation is provided by the Lagrangian equations of motion.

$$\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_k} \right) - \frac{\partial L}{\partial q_k} = 0 \quad (3.1)$$

where  $q_k$ ,  $k = 1, \dots, N$ , are the generalized coordinates that describe the system. Grouped together, they are referred to as a vector  $\mathbf{q}$ . The Lagrangian function  $L(\mathbf{q}, \dot{\mathbf{q}}, t)$  is defined as:

$$L = K - V \quad (3.2)$$

where  $K(\mathbf{q}, \dot{\mathbf{q}}, t)$  and  $V(\mathbf{q}, \dot{\mathbf{q}}, t)$  are the kinetic and potential energy, respectively. In the general case,  $K$  and  $V$  are functions of the generalized coordinates, their time derivatives, and time. Equation 3.1 is based on Hamilton's principle of least action (Hamilton's variational principle). According to this principle, the motion of a system from time  $t_1$  to time  $t_2$ , in which all forces except possibly constraint forces are derived from a generalized scalar potential that may be a function of the coordinates, the velocities, and time, follows the trajectory for which the line integral, or action integral, is minimized and has a stationary point.

$$S_{cl} = \int_{t_1}^{t_2} L(\mathbf{q}, \dot{\mathbf{q}}, t) dt \quad (3.3)$$

By the term "stationary point" for a line integral we mean that the integral has the same value along any other neighboring trajectory that differs from the first by an infinitesimal displacement. The conjugate momentum  $p_k$  of the generalized coordinate  $q_k$  is defined as:

$$p_k = \frac{\partial L(\mathbf{q}, \dot{\mathbf{q}}, t)}{\partial \dot{q}_k} \quad (3.4)$$

The Hamiltonian function  $H(\mathbf{q}, \mathbf{p}, t)$  is defined, in the Lagrangian formulation, as:

$$H(\mathbf{q}, \mathbf{p}, t) = \sum_k \dot{q}_k p_k - L(\mathbf{q}, \dot{\mathbf{q}}, t) \quad (3.5)$$

If the Lagrangian  $L$  does not depend explicitly on time, but only through  $\mathbf{q}$  and  $\dot{\mathbf{q}}$ , then  $H$  is a constant of motion and corresponds to the total energy of the system.

A second formulation of the equations of motion is that of Hamilton. Starting from the Hamiltonian function, a differential description of the trajectory is given by the relations:

$$\dot{q}_k = \frac{\partial H}{\partial p_k}, \dot{p}_k = -\frac{\partial H}{\partial q_k} \quad (3.6)$$

If the Hamiltonian function  $H$  does not depend explicitly on time but only through  $\mathbf{p}$  and  $\mathbf{q}$ , it can easily be shown that  $H$  is a constant of motion and equal to the total energy. The system is then characterized as conservative.

The third and final formulation that we will refer to is that of Newton. For a system described by the Cartesian coordinates  $\mathbf{r}_i$  of its individual particles, a differential description of the motion is given by the equations:

$$m_i \ddot{\mathbf{r}}_i = \mathbf{F}_i \quad (3.7)$$

$$\mathbf{F}_i = -\nabla_{\mathbf{r}_i} V \quad (3.8)$$

where  $m_i$ ,  $\mathbf{r}_i$ , and  $\mathbf{F}_i$  are the mass, the position vector, and the total force acting on particle  $i$ , respectively, and  $V(\mathbf{r}_1, \dots, \mathbf{r}_N)$  is the potential energy function of the system. This formulation follows very easily from Lagrange's equation (3.2) if we use the usual definition of  $K$

$$K = \sum_{i=1}^N \frac{\mathbf{p}_i^2}{2m_i} \quad (3.9)$$

The potential energy function  $V(\mathbf{r}_1, \dots, \mathbf{r}_N)$  is often expressed as a sum of contributions from one-body, two-body, three-body, etc., interactions:

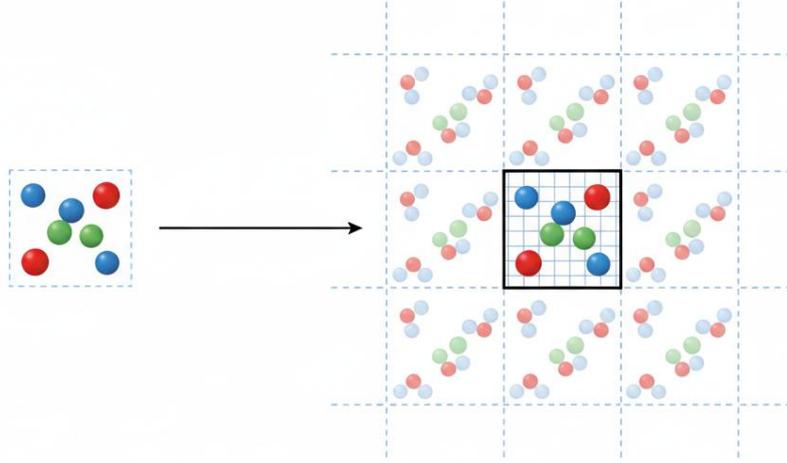
$$V = \sum_i v_1(\mathbf{r}_i) + \sum_i \sum_{j>i} v_2(\mathbf{r}_i, \mathbf{r}_j) + \sum_i \sum_{j>i} \sum_{k>j>i} v_3(\mathbf{r}_i, \mathbf{r}_j, \mathbf{r}_k) + \dots \quad (3.10)$$

In all the above formulations, the coordinates  $\mathbf{q}$  or  $\mathbf{r}$  are assumed to be independent variables. The equations must be modified if there are constraints between the degrees of freedom of the system. In the Newtonian formulation, the presence of constraints between the coordinates of the system's centers leads to an equation for the force  $\mathbf{F}_i$  of the form:

$$\mathbf{F}_i = -\nabla_{\mathbf{r}_i} V + \mathbf{g}_i \quad (3.11)$$

where  $\mathbf{g}_i$  is the force due to the constraints acting on particle  $i$ . The calculation of  $\mathbf{g}_i$  given the constraint equations is described below, where the algorithm of constrained dynamics is developed.

Regardless of how they are formulated, the equations of motion are time-reversible. In an isolated system of interacting particles, the quantities  $H$  (total energy),  $\mathbf{P} = \sum_i \mathbf{p}_i$  (total linear momentum), and  $\mathbf{L} = \sum_i \mathbf{r}_i \times \mathbf{p}_i$  (total angular momentum) are conserved. On the other hand, in a system of particles contained in a box with periodic boundary conditions (**Figure 3.2**), which evolves in time subject to interactions between the particles and their images,  $H$  and  $\mathbf{P}$  are conserved, but not angular momentum  $\mathbf{L}$ . Each time a particle crosses one face of the box and enters through the opposite face, a discontinuous change is produced in the total angular momentum of the box. MD simulations are usually carried out in a box with periodic boundary conditions, under the macroscopic constraints of constant  $N$ ,  $V$ ,  $E$ , or  $P$ ; that is, by sampling the probability density of the phase space of the microcanonical statistical ensemble, with the additional constraint that the total momentum remains constant.



**Figure 3.2:** Schematic representation of periodic boundary conditions in 2-D.

Let us consider an MD simulation of a system in which all interactions between the various centers are included in the potential energy function  $V(\mathbf{r}_1, \dots, \mathbf{r}_N)$ . This problem is solved by the numerical integration of differential equations of the form:

$$\ddot{\mathbf{r}}_i = \frac{1}{m_i} \mathbf{F}_i(\mathbf{r}_1, \dots, \mathbf{r}_N), \quad i = 1, 2, \dots, N \quad (3.12)$$

with given initial conditions  $\mathbf{r}_i(\mathbf{0})$  and  $\dot{\mathbf{r}}_i(\mathbf{0})$ . This is an initial-value problem, which is transformed into a system of coupled first-order ordinary differential equations by the substitution:

$$\dot{\mathbf{r}}_i = \mathbf{v}_i, \quad i = 1, 2, \dots, N \quad (3.13)$$

$$\dot{\mathbf{v}}_i = \frac{1}{m_i} \mathbf{F}_i(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N), \quad i = 1, 2, \dots, N \quad (3.14)$$

with initial conditions  $\mathbf{r}_i(\mathbf{0})$ ,  $\mathbf{v}_i(\mathbf{0})$ . With  $\mathbf{y} \equiv [\mathbf{r} \ \mathbf{v}]^T$ , the above system has the general form

$$\dot{\mathbf{y}} = \mathbf{f}(t, \mathbf{y}) \quad (3.15)$$

with given  $\mathbf{y}(\mathbf{0})$ . As we know from initial-value problems in chemical engineering, there exists a large number of finite-difference methods for solving the above equation. In general, one proceeds along the independent variable (time) with finite steps  $\delta t$ . The step size may be constant or may be adapted during the numerical solution. The numerical expression for determining  $\mathbf{y}(t)$  is called an  $n$ -th order method when the error is

$$\epsilon = |\mathbf{y}^{numerical} - \mathbf{y}^{exact}| = O[(\delta t)^{n+1}] \quad (3.16)$$

When choosing an algorithm for integrating the above equations, we must pay attention to certain characteristics that distinguish MD from other initial-value problems. Some of these are listed below.

- The part of MD that requires the most computational time is the evaluation of the forces acting on the particles. A good criterion for the efficiency of an MD algorithm is the ratio

$$\frac{\text{Number of Forces Evaluated}}{\text{Simulation Time}}$$

to be as small as possible. The speed of all other calculations performed in the algorithm is not important, since the entire computation is dominated by the calculation of the forces  $\mathbf{F}_i$ . From Equation (3.16) it is clear that the algorithm should not require many force evaluations per integration time step. Thus, popular methods such as the fourth-order Runge–Kutta–Gill 4<sup>th</sup> order method are rarely used in MD because they do not satisfy this criterion. Also, from (3.16) it follows that the algorithm should allow the largest possible integration time step  $\delta t$ . There is, of course, a trade-off between the number of equations that are evaluated and the size of the step that can be used.

- The algorithm must be stable. This means that the error  $\epsilon$  should not grow rapidly with increasing  $\delta t$ . The problem of stability is important in “stiff” initial-value problems that involve two or more characteristic times. MD simulations of molecular systems exhibit stiffness with respect to the integration step due to the wide spectrum of frequencies or characteristic times that appear. The short characteristic times of bond-stretching and bond-angle vibrations coexist with the long times of translational motion and all kinds of long-range collective motions. In general, stiffness in initial-value problems can be eliminated by using implicit (non-explicit) algorithms; however, such algorithms are not preferred in MD because they require many force evaluations.

- The integration algorithm should require little memory.

- The algorithm must be accurate. This means that the error  $\epsilon$  should remain small for sufficiently large  $\delta t$ . Considering the need for an algorithm to describe the dynamical trajectory as accurately as possible, one can see that the problems solved in MD, being highly nonlinear, are characterized by great sensitivity to initial conditions. Thus, two trajectories that are initially very close diverge rapidly.

### 3.3.2 Molecular Dynamics Algorithm

The simplest version of a Molecular Dynamics algorithm follows these steps:

- I. The various parameters serve as the input to the program (e.g., initial temperature, number of molecules, density, time step, etc.).
- II. The system is initialized: selection of the initial positions and velocities of the molecules in the system.
- III. In this phase, the classical rules of Newtonian motion are applied. This step, in combination with the previous one, forms the "core" of the simulation technique and is repeated until the predetermined duration of the simulation is reached, thereby tracking the temporal evolution of the system.
- IV. After completing the central simulation loop, the average values of the desired quantities are computed and printed, and the simulation ends [10]

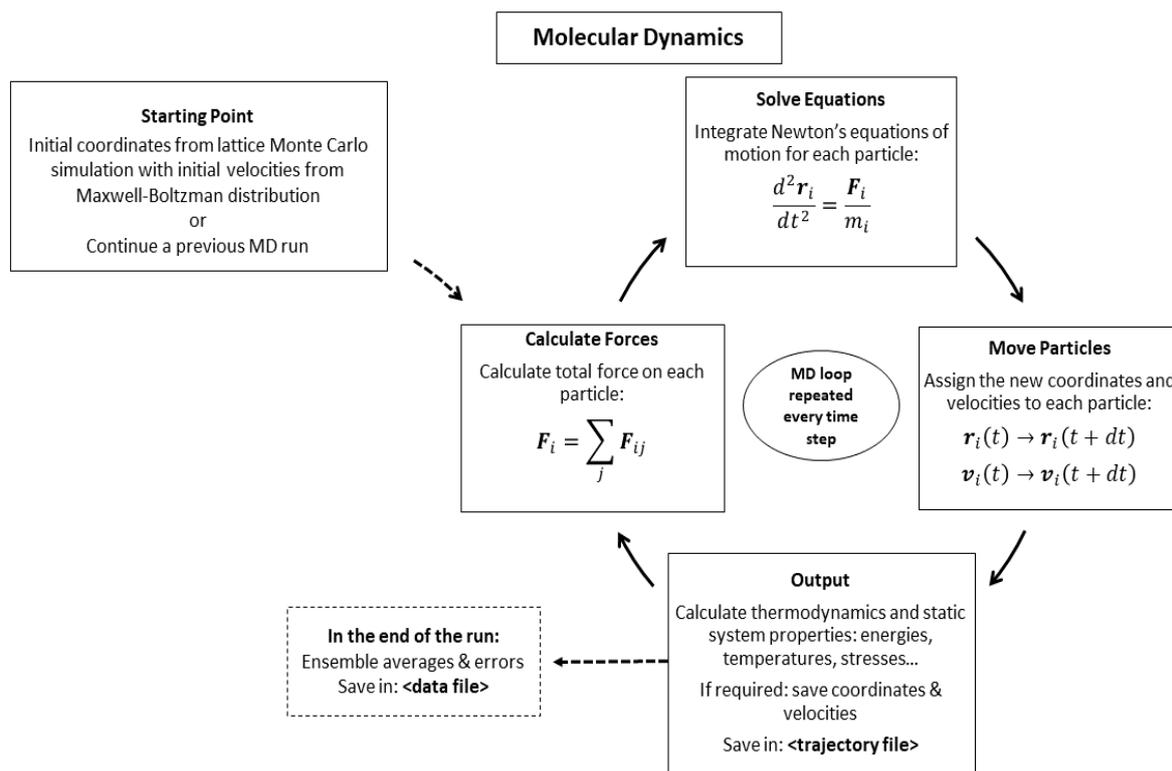


Figure 3.3: MD algorithm.

### 3.3.3 Statistical Ensembles

As established in the previous chapters, during a Molecular Dynamics simulation, the microscopic details of the system are linked to physical properties, such as thermodynamic

equilibrium properties, through statistical mechanics. Specifically, according to the concept of the Gibbs ensemble, many microscopic configurations of a very large system lead to the same macroscopic properties, implying that knowledge of the exact motion of each molecule in a system is not required to predict its properties. Thus, among a large number of identical systems, macroscopic properties are formed as ensemble averages [1], [11].

Some characteristic examples of thermodynamic variables that define statistical ensembles at constant values include volume ( $V$ ), pressure ( $P$ ), temperature ( $T$ ), chemical potential ( $\mu$ ), and the number of particles, as well as the microscopic velocities ( $v_i$ ) and positions ( $r_i$ ). In conclusion, the idea of a statistical ensemble is based on an idealization in which multiple replicas of a system are created, which are analyzed simultaneously. Each of these replicas represents a possible state in which the actual system could exist. These microstates of the system are determined by the boundary conditions of the system [10], [11].

The main statistical ensembles are as follows [12], [13], [14], [15], [16]:

- Microcanonical ensemble ( $NVE$ ): The total number of particles ( $N$ ), the total volume ( $V$ ), and the total energy of the system ( $E$ ) are constant.
- Canonical ensemble ( $NVT$ ): In this ensemble, the total number of particles ( $N$ ), the total volume ( $V$ ), and the total temperature of the system ( $T$ ) are constant.
- Isothermal-isobaric ensemble ( $NPT$ ): In this type of statistical ensemble, the total temperature ( $T$ ), the number of particles ( $N$ ), and the pressure ( $P$ ) of the system are constant.
- Isoenthalpic-isobaric ensemble ( $NPH$ ): The constant macroscopic properties of this ensemble are the total number of particles ( $N$ ), the total pressure ( $P$ ), and the total enthalpy ( $H$ ).
- Grand canonical ensemble ( $\mu VT$ ): The constant macroscopic properties of this ensemble are the chemical potential ( $\mu$ ), the volume ( $V$ ), and the temperature ( $T$ )

### 3.3.4 Force Field

The interaction potentials or force fields (FFs) are the cornerstone and the most important variable of molecular mechanics, providing information about the interactions within a system of molecules, whether through bonding and angle forces between atoms or non-contact forces [17], [18]. As a mathematical expression, the force field consists of an analytical form of the intermolecular potential energy  $U(r_1, r_2, \dots, r_N)$ , which includes a set of parameters derived from experimental data or mathematical calculations. In the first case, experimental data are

obtained using techniques such as neutron, X-ray, and electron diffraction, NMR, Raman spectroscopy, infrared and neutron spectroscopy, etc. Alternatively, semi-empirical quantum mechanical calculations or ab-initio calculations are employed to compute the parameters of a force field [18]. It is worth noting that FFs model the real potential primarily through the use of simple or more complex interaction potentials, where molecules are ensembles of atoms bound by forces (such as harmonic potentials). The modeling of the real potential using FFs is designed in such a way that it allows the study of physical systems within the required time and spatial scales on modern computational systems, without losing critical details about the structure-property relationships of materials, as well as the phenomena under investigation [18]. Among the numerous interaction potentials of varying complexity, a typical form still applies, which is expressed by Equation 3.17 [19], [20]:

$$U = \sum_{bonds} \frac{1}{2} k_b (r - r_0)^2 + \sum_{angles} \frac{1}{2} k_a (\theta - \theta_0)^2 + \sum_{torsions} \frac{V_n}{2} [1 + \cos(n\phi - \delta)] + \sum_{\substack{improper \\ dihedrals}} k_\xi (\xi - \xi_0)^2 + \sum_{LJ} 4\epsilon_{ij} \left( \frac{\sigma_{ij}^{12}}{r_{ij}^{12}} - \frac{\sigma_{ij}^6}{r_{ij}^6} \right) + \sum_{charges} \frac{q_i q_j}{r_{ij}}. \quad (3.17)$$

Bond interactions are classified into four main types: bond length, bond angle, bond torsion, and sometimes generalized or improper dihedral angles. These four categories of interactions are represented by the first four terms in the equation in the corresponding order in Equation 3.17 [18], [19]. The first term is a harmonic function that governs the bond length and is used to calculate the energy of bond stretching. However, this term does not account for the possibility of bond breaking, making it unsuitable for calculations when a chemical reaction is taking place. The second term in Equation 3.17 represents the energy of bond angle distortion between successive atoms. In both of these terms,  $k_b$  and  $k_a$  are constants depending on the type of atoms involved in the bond. The third parameter in the same equation is responsible for calculating the rotational energy of bonds in two planes around a central bond, essentially calculating the torsional dihedral energy [19]. This is a sinusoidal relationship where  $\phi$  the torsion angle,  $\delta$  is the minimum value of the potential  $n$  determines the number of minima or maxima between 0 and  $2\pi$  and  $V_n$  determines the height of the potential barrier. The last type of intramolecular interaction is given by the fourth term in the equation, which provides the energy for improper dihedral angles, meant for stabilizing groups, such as in the case of aromatic rings being constrained to a planar geometry [17].

The non-bonded interactions in a typical force field must also include Van der Waals (VDW) and electrostatic forces, which are represented by the fifth and sixth terms in Equation 3.17, respectively. In the case of intermolecular Van der Waals interactions, the fifth term employs the Lennard-Jones potential, which accurately predicted weak VDW bonds between noble gases and has since been used for a vast number of elements and interactions. The Lennard-Jones potential has the following form:

$$V_{LJ}(r) = 4\epsilon \left[ \left( \frac{\sigma}{r} \right)^{12} - \left( \frac{\sigma}{r} \right)^6 \right], \quad (3.18)$$

where  $r$  is the distance between the atoms, and  $\epsilon$  and  $\sigma$  are parameters dependent on the known properties of the material [21]. The  $\epsilon$  or  $\epsilon_{ij}$  represents the atomic attractive strength, and  $\sigma_{ij}$  the distance at which the potential becomes zero. As we can see in the diagram of **Figure 3.3**, the  $r^{-12}$  term represents the repulsion of atoms at very short distances, and the  $r^{-6}$  term the attractive force between two atoms that are far apart. The coefficients  $\sigma_{ij}$  and  $\epsilon_{ij}$  can be calculated using the Lorentz-Berthelot relation [18], [19]

$$\sigma_{ij} = \frac{1}{2} (\sigma_{ii} + \sigma_{jj}) \quad \text{and} \quad \epsilon_{ij} = \sqrt{(\epsilon_{ii} \cdot \epsilon_{jj})}. \quad (3.19)$$

As for the sixth and final term of Equation 3.17, for electrostatic interactions, Coulomb's law is followed for calculating the electrical contribution to the total energy [19]. Based on this, the total electrostatic interaction energy between point charges  $q_i, q_j$  at positions  $r_i, r_j$  is given by:

$$U = \frac{1}{4\pi\epsilon_0} \sum_{i<j} \frac{q_i q_j}{|r_i - r_j|} = \frac{1}{8\pi\epsilon_0} \sum_{i \neq j} \frac{q_i q_j}{|r_i - r_j|}, \quad (3.20)$$

Where  $\epsilon_0$  is the electric permittivity in free space. In the first form of the expression, all pairs are counted once (e.g., the interaction between 1 and 2 is counted, not between 2 and 1), while in the second form, all interactions are counted (the interaction between 1 and 2 is counted, as well as the interaction between 2 and 1), and the total is divided by two to compensate for double counting. Of course, for a very large or infinite system, and given that electrostatic interactions are long-range, the calculation of the electrostatic energy becomes, if not incredibly laborious, practically impossible, and therefore additional numerical calculations are required. In the case of porous adsorption systems, the Ewald method [22] is used.

## 3.4 Monte Carlo

### 3.4.1 Introduction

The use of randomness to obtain definitive answers is a revolutionary idea that can be traced back to the early 18th century with the famous 'Buffon's Needle.' This refers to the study by the French scientist Georges-Louis Leclerc de Buffon, who used random, repeated tosses of a needle on a lined surface to estimate the value of  $\pi$ . However, such experiments address previously understood deterministic problems, which is not in line with modern simulation, which solves deterministic problems by finding a probabilistic equivalent [23].

The Monte Carlo method (MC) was conceptualized after the completion of the “Manhattan Project” in the late stages of World War II by Neumann, Ulam, and Metropolis[24]. It emerged within the context of neutron diffusion research in fissile materials. The method relies on various computational algorithms that yield results through extensive random sampling. It serves as a valuable tool for solving physical and mathematical problems based on percolation theory, essentially embodying any approach that utilizes random numbers to address complex probability problems or even deterministic problems [25].

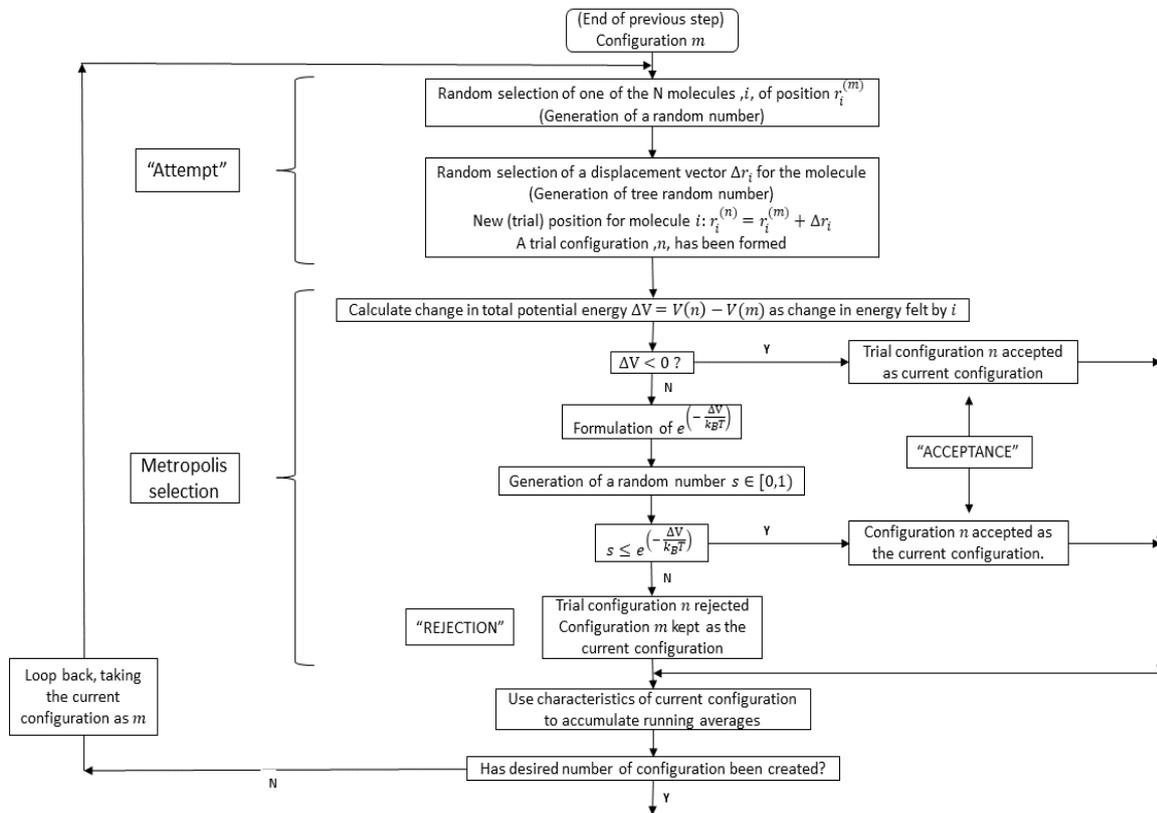
While the basic principle of the Monte Carlo method is that it relies on random number sampling experiments, which are then applied to a more limited number of numerical operations, in practice, things are somewhat different. For the convenience of testing and re-execution, the numbers introduced into a Monte Carlo simulation are not solely the result of uncertainty, but also 'pseudorandom' sequences, constructed in such a way as to present a high degree of randomness [26]. All these different approaches lead to a much shorter simulation, ideal for large systems[26], which produces results with statistically acceptable error. By increasing the number of repetitions within realistic time frames, this error can actually be reduced, ensuring more accurate calculations [27], [28].

In practice, the Monte Carlo method:

- Finds extensive use in problems where the prediction of future results depends on input variables and random 'noise', each having its own distinct distribution.
- It is a valuable tool when the physical problem is mathematically complex, and it is difficult to derive an analytical solution from it.
- Provides a means to conduct predictions and the underlying probability of observing those results [27].

### 3.4.2 Monte Carlo Algorithm

In the context of this thesis, the Metropolis Monte Carlo method is used. MC simulations of this form can be performed in various ensembles such as NVT, NPT, and  $\mu$ VT. In the case of adsorption simulations, they can be performed in the Gibbs ensemble, where there is an equilibrium of adsorption between the total liquid phase and the internal adsorbent. Two other statistical ensembles that are used are the  $\mu$ 1N2PT and the grand-canonical ensemble [22]. The selected statistical ensemble will be discussed below. Based on different thermodynamic conditions, characteristic configurations, or microstates of the system are taken, using the pseudorandom number sampling technique mentioned above. The steps of a typical Monte Carlo algorithm are depicted in **Figure 3.4**.



**Figure 3.4:** Basic Monte Carlo algorithm.

For the execution of a trial move, the MC algorithm first selects a random molecule  $i$  at position  $\mathbf{r}_i^{(m)}$  in state  $m$ . To this molecule, it assigns a random displacement vector  $\Delta \mathbf{r}_i$ . Thus, the position of the molecule in the new state  $n$  is:

$$\mathbf{r}_i^{(n)} = \mathbf{r}_i^{(m)} + \Delta \mathbf{r}_i \quad (3.21)$$

$$\Delta V = V(n) - V(m) \quad (3.22)$$

In the case where:

- $\Delta V < 0$ : The move is accepted, and state  $n$  becomes the current state, taking it into account when calculating the system's average state.
- $\Delta V \geq 0$ : A new inequality of the form must be satisfied:

$$s \leq e^{\left(\frac{-\Delta V}{k_B T}\right)}, \quad (3.23)$$

Where  $s$  is a random number in the range  $[0,1)$ . When this relation holds, the move is accepted, and state  $n$  becomes the current state. Otherwise, the move is rejected, and state  $m$  remains unchanged.

After accepting an MC move, the steps outlined are repeated until the required number of microstates for calculating the average state has been fulfilled.

In this way, a new trial state  $n$  is created. The following is the computational acceptance/rejection method of Metropolis, which is based on calculating the potential energy exerted on the molecule in the form of a charge. It is carried out as follows:

We observe that the MC algorithm initially attempts to take a step by selecting a random molecule  $i$  at position  $r_i(m)$  in state  $m$  and assigning it a random displacement vector  $\Delta r_i$ . Thus, the position of the molecule in the new state  $n$  is given by  $r_i(n) = r_i(m) + \Delta r_i$ . This is followed by the Metropolis acceptance/rejection computational method, which relies on computing the potential energy experienced by the molecule in the form of charge, as given by the equation:

$$\Delta V = V(n) - V(m) \quad (3.24)$$

The term " $\Delta V < 0$ " indicates that the change in potential energy is negative, and therefore the movement is accepted. In the case where the change in potential energy is non-positive ( $\Delta V \geq 0$ ), a new inequality must be satisfied in the form of:

$$s \leq e^{\left(\frac{-\Delta V}{k_B T}\right)}, \quad (3.25)$$

where  $s$  is a random number belonging to the interval  $[0,1)$ . When inequality (3.25) holds, the movement is accepted, and state  $n$  becomes the current state. Otherwise, the movement is rejected, and state  $m$  remains unchanged. In the case of accepting a Monte Carlo move, the algorithm's steps are repeated until the required number of states for computing the average state has been collected.

### 3.4.3 Grand Canonical Monte Carlo

In systems like the one in this thesis, the goal is to calculate the average number of adsorbed molecules as a function of the external conditions of temperature and pressure of the adsorbent or reservoir with which they are in contact. Approaching the problem theoretically, we could apply the Molecular Dynamics method with the appropriate statistical ensemble (e.g., grand-canonical ensemble) and simulate the contact of the adsorbent with a gas.

However, the MD method is not ideal for simulating such a complex system, primarily due to the required system equilibration time. While in a real experiment, the equilibration of the adsorbent/adsorbed gas system takes minutes to hours, depending on the type of gas under study, in an MD simulation, one minute of experimental time corresponds to  $10^9$  seconds for a computer [13]. Another significant problem is the focus on a large part of the MD simulation being on the gas phase of the system, which is not directly relevant to the calculation and consumes significant CPU time. In contrast, the focus area is the bulk properties of the system, which differ from the properties of the interface between the gas phase and the adsorbent. The only way to reduce the influence of these interfacial properties with this method is to increase the system size, which is not ideal, including the duration problem [15].

In the grand-canonical Monte Carlo, the temperature, volume, and chemical potential ( $\mu, V, T$ ) are kept constant, and the adsorbed gas is in equilibrium with the gas in the reservoir, as the reservoir and the adsorbent have the same temperature and chemical potential. Thus, the interfacial property problem is eliminated, and the gas in contact with the adsorbent is considered part of the reservoir, applying  $\mu$  and  $T$ . Knowing these static values, the equilibrium concentration inside the adsorbent can be determined. For these reasons, GCMC is the ideal statistical ensemble for adsorption studies.

Monte Carlo simulations in the grand-canonical statistical ensemble ( $\mu VT$ ) use a direct generalization of the canonical ensemble ( $NVT$ ). The simulations are carried out under constant chemical potential ( $\mu$ ), volume ( $V$ ), and temperature ( $T$ ), while the number of molecules ( $N$ ) fluctuates. To calculate a property  $A$  (such as energy, pressure, etc.), the following relation is used:

$$\langle \mathcal{A} \rangle_{\mu VT} = \frac{\sum_{N=0}^{\infty} \frac{1}{N!} (Vz)^N \int ds \mathcal{A}(s) \exp(-\beta \mathcal{V}(s))}{Q_{\mu VT}} \quad (3.26)$$

Where  $N$  is the total number of particles,  $z = e^{\beta\mu/\Lambda^3}$  ( $\Lambda$  is the de Broglie wavelength and the inverse temperature  $\beta = 1/k_B T$ ),  $\mathcal{V}(s)$  is the potential energy, and  $\mathcal{A}(s)$  is the value of the property under study in the current microstate of the system  $s = (s_1, s_2, \dots, s_N)$ . The exponential  $\exp(-\beta\mathcal{V}(s))$  is called the Boltzmann factor and is used to ensure the preference of lower-energy configurations and to sample higher-energy states with reduced probability. The integral is introduced for calculating the average state of the system for all possible arrangements of the  $N$  particles. In the denominator of Equation (3.26), the grand canonical partition function is introduced, which is defined as follows:

$$Q_{\mu VT} = \sum_N \frac{1}{N! h^{3N}} \left( \frac{\mu N}{k_B} \right) \int r dr dp \exp\left(-\frac{\mathcal{H}}{k_B T}\right) \quad (3.27)$$

Where  $H$  is the Hamiltonian of the system, which represents the total energy,  $h$  is the Planck constant,  $k_B$  is the Boltzmann constant, and  $r, p$  are the positions and momenta of the particles, respectively.

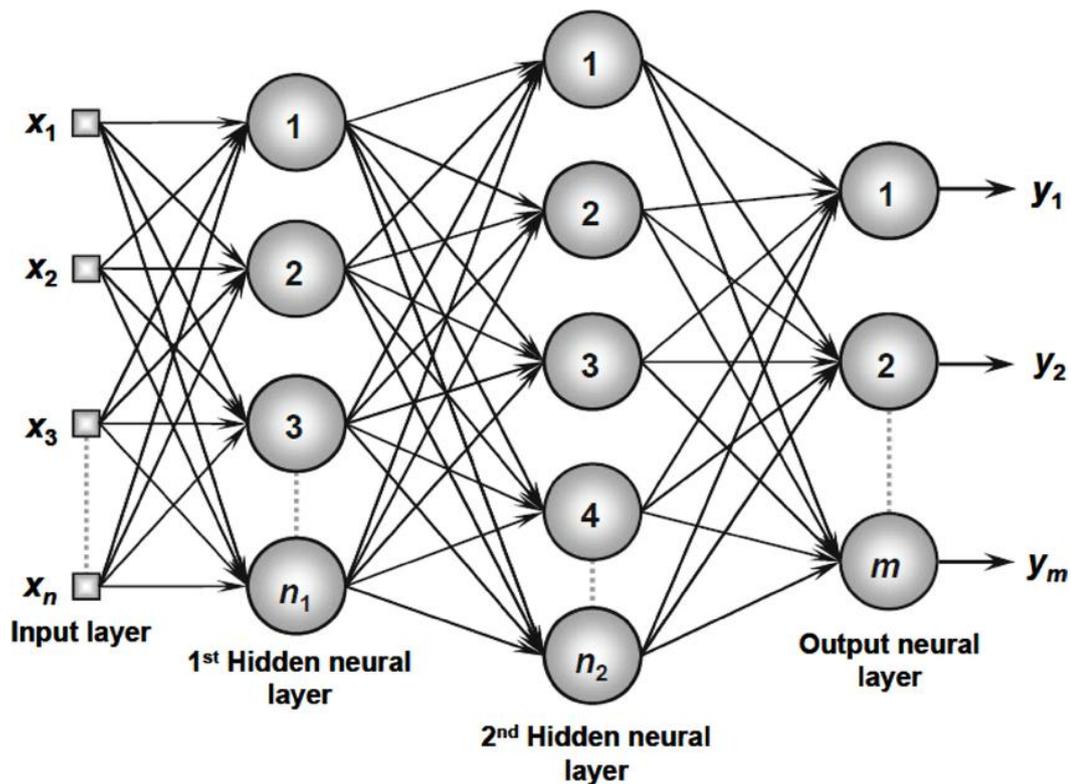
### 3.5 Artificial Neural Networks (ANNs)

Inspired by the biological neural networks inherent in the brain, Artificial Neural Networks (ANNs) are a novel class of AI techniques. ANNs, which were first envisioned as one of the first AI systems influenced by biology, imitate the structure and functional features of biological neural networks to carry out intricate computational tasks. The fundamental ANN, which is made up of linked nodes called neurons arranged into layers, functions as a crude model of the neural architecture of the brain. Because it is flexible and adaptable to different issue domains, this strong yet basic paradigm is widely used in a wide range of AI software solutions. The framework shown is a feedforward multilayer perceptron (FFMLP), a basic kind of artificial neural network (ANN) architecture that has been the basis for many other architectural extensions and variants. The FFMLP is made up of an input layer, an output layer, and one or more hidden layers. Interconnected neurons make up each layer. The FFMLP has been greatly improved and refined since its creation, including cutting-edge methods like gradient descent and backpropagation to maximize learning and performance.

Artificial neurons, also known as nodes, are the building blocks of all neural networks. These artificial neurons are made up of components that resemble ordinary human nerve cells. Parts of the node are named after their biological counterparts, and their purposes are very similar. Each of these components of ANN is responsible for certain jobs and mathematical functions. Dendrites are utilized to accept inputs. A cell body, also known as SOMA, processes the

information via activation functions. Axons translate the received input into outputs and convey them to synapses, which are the connections between neurons [29].

An artificial neural network (ANN) is made up of neurons, which are the basic building blocks of the network's architecture and the basic parts that process information. These neurons are connected across layers in order to build an Artificial Neural Network. A multilayer perceptron has at least one hidden layer, adding complexity and the ability to do complicated computations and pattern recognition. In contrast, a single-layer perceptron connects the input layer directly to the output layer. The capacity of a multilayer perceptron to learn and generalize from input data is improved by the hidden layers, which make it easier to extract higher-level characteristics and abstract complicated relationships within the data. The addition of hidden layers allows the ANN to represent nonlinear interactions and capture subtle patterns, increasing its usability and applicability across a wide range of tasks and domains. These hidden layers provide the so-called depth in deep learning and are depicted in **Figure 3.5**



**Figure 3.5:** Feedforward Multilayer Perceptron [35].

Each of these layers can be modified in accordance with the objective's size. The number of inputs in the input layer can be changed based on the data. For example, every pixel in a picture can be utilized as an input for image recognition. The quantity of outputs is determined by the desired outcome.

The number of hidden layers and how many nodes are contained within each hidden layer is more challenging to estimate. A general rule of thumb is to have equal or fewer nodes than in the input layer but more than in the output layer. Usually, testing a few various parameters and selecting the most effective result is the ideal way to go. However, it should be noted that the more hidden layers and nodes there are, the more difficult tasks the ANN can do, but this also means that more computer power is required to complete the jobs[30]. **Figure 3.5** shows a multilayer perceptron feedforward neural network. Various neural network topologies can be formed by combining multiple layers and back feeding the output data. In a feedforward neural network, information only flows in one direction, however in a backpropagation network, information moves both ways, forming a feedback circuit that allows the network to learn on its own.

### **3.5.1 Artificial Neural Networks architecture**

According to Nelson et al.[31], there are infinitely many ways to organize a neural network, but only a few models are commonly used. A neural network's organization can be characterized by their neurodynamics and architecture. Neurodynamics are the qualities of an artificial neuron, which include the following [32]:

- combination of inputs
- production of outputs
- a transfer (activation) function
- weight schemes, i.e. weight initialization and weight learning algorithms.

These attributes can be applied to the entire network on a system level. Network architecture, also known as network topology, defines the organization of a network and its key properties, which are:

- types of interconnections among artificial neurons (henceforth referred to as just neurons)
- number of neurons and
- number of layers

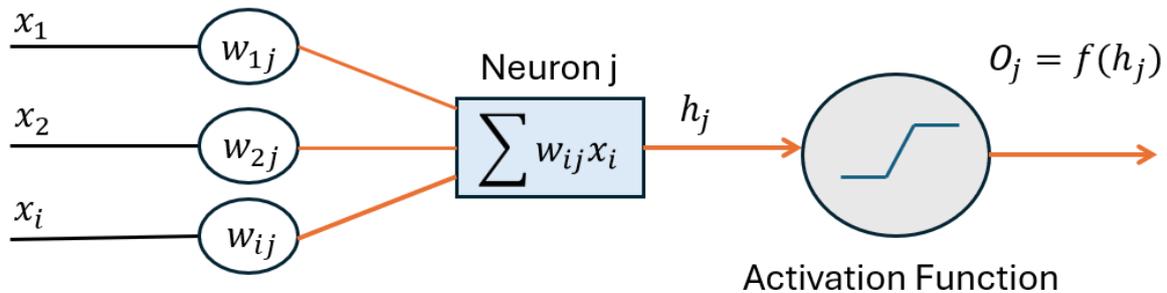
The input layer of an ANN serves as a buffer, transmitting data to the subsequent layers. Preprocessing inputs may be necessary for ANNs, which solely work with numerical data. To use the ANN, the input data may need to be scaled and converted to numerical form. Haynes et al. [33] study describes an ANN real estate price simulator program that used a binary encoding for qualitative data on the availability of residential property attributes. The output layer of an ANN is similar to the input layer but carries information from the network to the

outside world. Output data often requires post-processing to make it understandable and useable outside of the network. Post-processing can vary from basic output scaling to complex hybrid systems.

### 3.5.2 Activation Functions

The activation function determines a neuron's output based on a sum of its weighted inputs. Neurons in the hidden layer have nonlinear transfer functions, which contribute to the neural network ability to adapt to nonlinearities. In **Figure 3.6**, neuron  $j$ 's output is presented after the transfer function  $f$  maps its weighted inputs from neurons 1 to  $i$  as follows:

$$Q_j = f_j(\sum_i w_{ij} x_i) \quad (3.28)$$



**Figure 3.6:** Artificial Neural Network architecture.

A transfer function converts real values to a domain typically restricted by 0 to 1 or -1 to 1. Bounded activation functions are often referred to as squashing functions. Early ANN models, such as the perceptron, utilized a basic threshold function (also called a step-function, hard-limiting activation, or Heaviside function):

- threshold:  $f(x) = 0$  if  $x < 0$ , 1 otherwise.

Current ANN models mostly utilize sigmoid transfer functions. Masters [34] defined a sigmoid function as a continuous, real-valued function with a positive derivative and a restricted range.

Example of a sigmoid function are:

- logistic  $f(x) = \frac{1}{1+e^{-x}}$
- hyperbolic tangent:  $f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$

The logistic function remains the most commonly applied in ANN models due to the ease of computing its derivative: Kalman et al.[35] propose that the hyperbolic tangent function is the

optimal transfer function. Master's et al. [34] found that the design of a function has little influence on a network but can significantly impact training speed. Other typical transfer functions are:

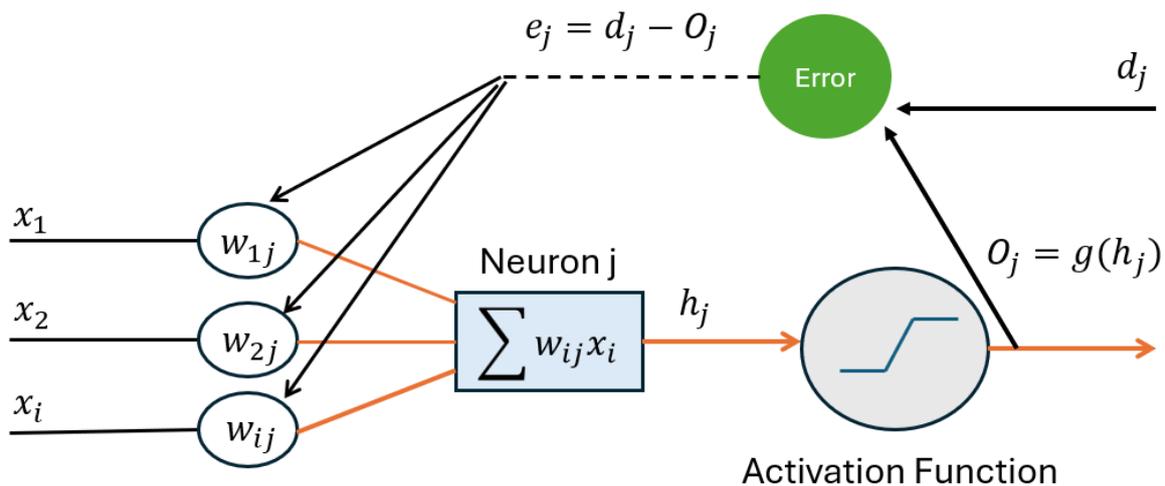
- linear:  $f(x) = x$  Which is normally used in a simple input/output model
- Gaussian:  $f(x) = e^{-x/2}$

It is impossible to train an artificial neural network to achieve the theoretical limit values of sigmoid functions. Values near the boundaries should be considered as having reached them. In a logistic function with limits of 0 to 1, neurons are completely engaged at 0.9 and switched off around 0.1.

### 3.5.3 Back propagation Algorithm

An ANN's starting weights are often chosen at random or via an algorithm. The learning method adjusts the weights based on the error between the network output and the desired output. The learning algorithm aims to minimize the error to an acceptable level. Back-propagation is the most widely used learning algorithm for multilayer neural networks.

The backpropagation (BP) method is a variation of the delta rule designed for networks with hidden layers. It is the most extensively used learning algorithm among artificial neural network researchers. Its popularity is attributed to its straightforward design and implementation.



**Figure 3.7:** Implementation of Back propagation algorithm.

### 3.5.4 Cross Validation

Cross-validation is a resampling procedure used to evaluate machine learning models on a limited data sample and increase the efficiency of the model in terms of training and testing. The method involves a single parameter, denoted as “k,” which determines the number of groups into which a data sample is divided. Consequently, this technique is commonly known as k-fold cross-validation. When a specific value for k is selected, it can replace the variable “k” in the model reference—for instance, using k=10 results in 10-fold cross-validation. Cross-validation is commonly used in applied machine learning to assess how well a model performs on unseen data. It involves using a limited sample to estimate the model’s expected performance in general, beyond the training data. This method is popular because it provides a less biased estimate of model skill compared to other techniques, such as a simple train/test split. Additionally, k-fold cross-validation evaluates the model design rather than specific training instances, as it re-trains the same model design with different training sets.

In a prediction problem, a model is often provided with a dataset of known data on which to train (the training dataset), as well as a collection of unknown data (or first seen data) against which to test the model [36]. The goal of cross-validation is to test the model's ability to predict new data that was not used in its estimation, in order to identify problems such as overfitting or selection bias[37] and to provide insight into how the model will generalize to an independent dataset. In this research k-fold cross validation was utilized on the models to get the findings. This approach resamples machine learning models for analysis purposes. This method minimizes bias caused by random data allocation to models.

The mechanics of cross-validation are simple, but the arguments for using it with neural networks are more complicated. The goal of the classification issue is to identify a set of neural network weights and bias values such that the input data produces output values that closely match the target values. A simple way would be to train the neural network with all 30 available data elements. However, this strategy is likely to identify weights and bias values that match the data exceptionally, possibly with 100 percent accuracy but when presented with a new, previously unseen piece of input data, the neural network will most likely predict very poorly. This problem is known as overfitting. To avoid overfitting, the available data is divided into two sets: a training set (typically 80 percent to 90 percent of the data) for finding good weights and bias values, and a test set (10 percent to 20 percent of the data) for evaluating the quality of the resulting neural network. The simplest version of cross-validation divides the available data at random into a single training set and a single test set. This is commonly referred to as

hold-out validation. However, the hold-out strategy is relatively uncertain because an unlucky split of the available data may result in an unsuccessful neural network. One possibility is to do hold-out validation numerous times. This is known as repeated subsampling validation. Nonetheless this strategy carries some risk since, while improbable, some data items may be used exclusively for training and never for testing, or vice versa. The aim behind k-fold cross-validation is to divide all available data points into roughly equal-sized groups. Each set is utilized exactly once as a test set, with the remaining data serving as a training set. A basic implementation of the k-fold cross validation can be found in **Appendix B**

## References

- [1] A. Chronos, M. J. D. Rushton, C. Jiang, and L. H. Tsoukalas, “Nuclear wasteform materials: Atomistic simulation case studies,” *Journal of nuclear materials*, vol. 441, no. 1–3, pp. 29–39, 2013.
- [2] H. Fehske, R. Schneider, and A. Weisse, *Computational many-particle physics*, vol. 739. Springer, 2007.
- [3] D. C. Rapaport, *The art of molecular dynamics simulation*. Cambridge university press, 2004.
- [4] S. Nosé, “A unified formulation of the constant temperature molecular dynamics methods,” vol. 81, no. 1, pp. 511–519, Jul. 1984, doi: 10.1063/1.447334.
- [5] S. Nosé, “A molecular dynamics method for simulations in the canonical ensemble,” *Mol Phys*, vol. 52, no. 2, pp. 255–268, 1984, doi: 10.1080/00268978400101201.
- [6] M. E. Tuckerman and G. J. Martyna, “Understanding modern molecular dynamics: Techniques and applications,” 2000, *ACS Publications*.
- [7] L. Monticelli and D. P. Tieleman, “Force fields for classical molecular dynamics,” *Biomolecular simulations: Methods and protocols*, pp. 197–213, 2013.
- [8] K. Lin and Z. Wang, “Multiscale mechanics and molecular dynamics simulations of the durability of fiber-reinforced polymer composites,” *Commun Mater*, vol. 4, no. 1, p. 66, 2023, doi: 10.1038/s43246-023-00391-2.
- [9] C. A. Krishna, S. Yashonath, A. Schüring, S. Fritzsche, and J. Kärger, “Influence of the Methane–Zeolite A Interaction Potential on the Concentration Dependence of Self-diffusivity,” *Adsorption Science & Technology*, vol. 29, no. 6, pp. 553–567, 2011.
- [10] B. Smit, *Understanding molecular simulation: from algorithms to applications*. Academic Press, 1996.
- [11] V. V. Kulagina, S. V. Eremeev, and A. I. Potekaev, “The Molecular-Dynamics Method for Different Statistical Ensembles,” *Russian Physics Journal*, vol. 48, pp. 122–130, 2005.
- [12] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, “Equation of state calculations by fast computing machines,” *J Chem Phys*, vol. 21, no. 6, pp. 1087–1092, 1953.
- [13] D. Frenkel, “Introduction to monte carlo methods,” *NIC series*, vol. 23, pp. 29–60, 2004.

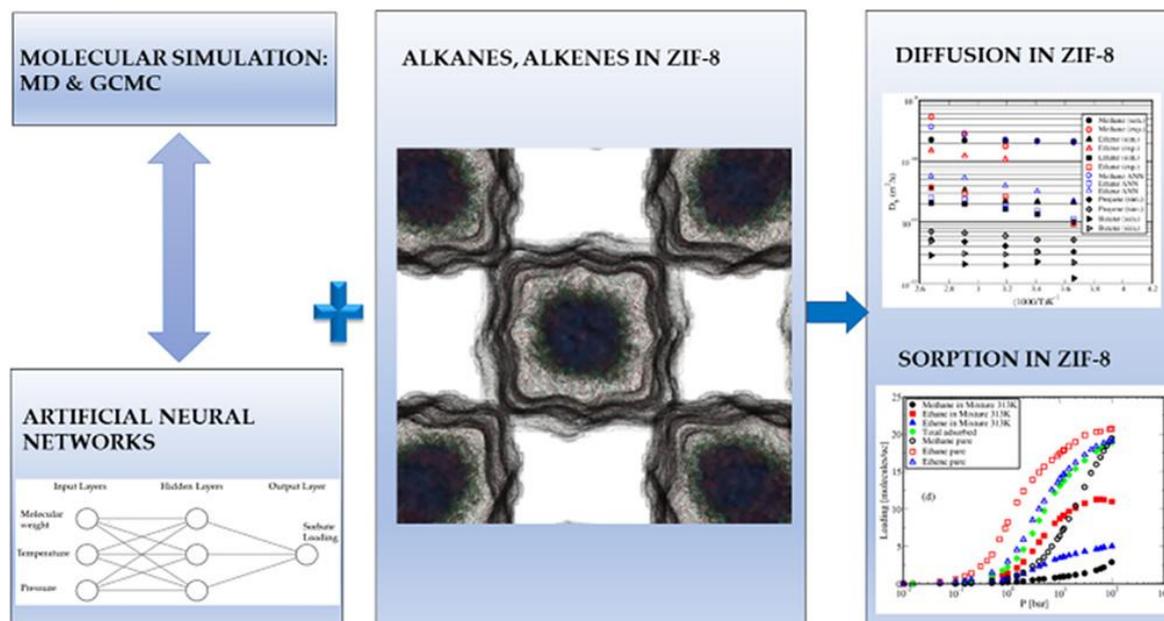
- [14] D. C. G. Jones, *Atomic physics*. Routledge, 2018.
- [15] D. A. McQuarrie, “Statistical Mechanics Harper Collins Publishers,” *New York, NY*, 1976.
- [16] R. Swendsen, *An introduction to statistical mechanics and thermodynamics*. Oxford University Press, USA, 2020.
- [17] S. Patodia, A. Bagaria, and D. Chopra, “Molecular dynamics simulation of proteins: a brief overview,” *J Phys Chem Biophys*, vol. 4, no. 6, p. 1, 2014.
- [18] L. Zheng, A. A. Alhossary, C.-K. Kwok, and Y. Mu, “Molecular dynamics and simulation,” 2019.
- [19] G. Kanagalingam, S. Schmitt, F. Fleckenstein, and S. Stephan, “Data scheme and data format for transferable force fields for molecular simulation,” *Sci Data*, vol. 10, no. 1, p. 495, 2023.
- [20] J. Zhang, “Canonical dual theory applied to a Lennard-Jones potential minimization problem,” *arXiv: Optimization and Control*, 2011, [Online]. Available: <https://api.semanticscholar.org/CorpusID:15112546>
- [21] J. B. Adams, “Bonding energy models,” *Encyclopedia of materials: science and technology*, pp. 763–767, 2001.
- [22] D. Dubbeldam, A. Torres-Knoop, and K. S. Walton, “On the inner workings of Monte Carlo codes,” *Mol Simul*, vol. 39, no. 14–15, pp. 1253–1292, 2013.
- [23] R. L. Harrison, “Introduction to monte carlo simulation,” in *AIP conference proceedings*, 2010, p. 17.
- [24] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, “Equation of state calculations by fast computing machines,” *J Chem Phys*, vol. 21, no. 6, pp. 1087–1092, 1953.
- [25] C. L. Brooks, “Computer simulation of liquids,” *J Solution Chem*, vol. 18, no. 1, p. 99, 1989, doi: 10.1007/BF00646086.
- [26] D. Earl and M. Deem, “Monte Carlo Simulations,” *Methods Mol Biol*, vol. 443, pp. 25–36, Jul. 2008, doi: 10.1007/978-1-59745-177-2\_2.
- [27] M. Gervasio and K. Lu, “Monte Carlo simulation modeling of nanoparticle–polymer cosuspensions,” *Langmuir*, vol. 35, no. 1, pp. 161–170, 2018.

- [28] J. Gils *et al.*, “Monte Carlo for Protein Structures,” Jul. 2023.
- [29] B. G. Buchanan, “A (very) brief history of artificial intelligence,” *AI Mag*, vol. 26, no. 4, p. 53, 2005.
- [30] B. M. Wilamowski, “Neural network architectures and learning algorithms,” *IEEE Industrial Electronics Magazine*, vol. 3, no. 4, pp. 56–63, 2009, doi: 10.1109/MIE.2009.934790.
- [31] M. McCord-Nelson and W. T. Illingworth, *A practical guide to neural nets*. Addison-Wesley Longman Publishing Co., Inc., 1991.
- [32] C. N. W. Tan, “Artificial neural networks: A financial tool as applied in the Australian market,” Bond University, 1997.
- [33] J. Ahire, *Artificial Neural Networks: The brain behind AI*. Lulu Press, Incorporated, 2018. [Online]. Available: <https://books.google.be/books?id=RHBTDwAAQBAJ>
- [34] T. Masters, *Practical neural network recipes in C++*. USA: Academic Press Professional, Inc., 1993.
- [35] B. L. Kalman and S. C. Kwasny, “Why tanh: choosing a sigmoidal function,” in *[Proceedings 1992] IJCNN International Joint Conference on Neural Networks*, 1992, pp. 578–581.
- [36] S. M. Piryonesi and T. E. El-Diraby, “Data Analytics in Asset Management: Cost-Effective Prediction of the Pavement Condition Index,” *Journal of Infrastructure Systems*, vol. 26, no. 1, p. 4019036, 2020, doi: 10.1061/(ASCE)IS.1943-555X.0000512.
- [37] D. M. Allen, “The Relationship Between Variable Selection and Data Augmentation and a Method for Prediction,” *Technometrics*, vol. 16, no. 1, pp. 125–127, 1974, doi: 10.1080/00401706.1974.10489157.



# Chapter 4: Molecular and Artificial Neural Networks Modeling of Sorption and Diffusion of Small Alkanes, Alkenes and Their Ternary Mixtures in ZIF-8 at Different Temperatures

## 4.1 Preface



**Figure 4.1:** Workflow presented in the scope of this research.

In this work Molecular Dynamics (MD) simulations in the NVT statistical ensemble were conducted to study the self-diffusion of normal alkanes, alkenes, and their mixtures at different temperatures and sorbate loadings in ZIF-8 as it is depicted in **Figure 4.1**. Molecular modeling and computer simulation techniques have been used as important tools for the investigation of diffusion and sorption phenomena in microporous materials. This starts with detailed quantitative models of atomic-level structure and interactions and proceeds to derive thermodynamic and transport properties based on the principles of dynamics and statistical mechanics. Molecular simulations (Molecular Dynamics and Monte Carlo) of materials in general [1], [2] and particularly of microporous material–sorbate systems have witnessed unprecedented growth in the last 40 years [3]. Additionally, GCMC simulations were utilized to describe sorption characteristics of the aforementioned hydrocarbons. Complementary simulations relative to pure alkanes and alkenes and for alkane–alkene mixtures sorbed in the ZIF-8 framework also took place, expanding the knowledge of diffusion and sorption phenomena in the microporous material. It should be mentioned that additional comparisons

take place with results obtained from microimaging methods [4], [5] providing additional justifications and insights regarding diffusion and sorption phenomena in ZIF-8.

In conjunction with the molecular-based simulation results and experimental measurements, Machine Learning (ML) techniques allow the inferring of mathematical models to describe qualitatively and quantitatively structure–property relations from available data. Potentially this can provide reliable estimations/predictions of adsorption and diffusion properties of sorbates in porous structures such as zeolites-MOFs-ZIFs with reduced computational cost by orders of magnitude when compared to molecular simulation approaches [6], [7] In particular, Artificial Neural Networks (ANNs) constitute a subset of ML and are at the heart of Deep Learning (DL) algorithms. They can be classified into different types such as Perceptron, Feedforward, or Multi- Layer Perceptron (MLPs), Convolutional (CNNs), and Recurrent (RNNs). ANNs are information processing models inspired by the nervous system of the human brain, mimicking the way that biological neurons communicate and interact through signals. They can be considered as a collection of simple computational nodes that are interconnected. Neurons are the building blocks of the ANNs architecture, and each node receives a set of inputs from various sources, performs calculations based on these inputs, and generates an output. The predictive capabilities of ANNs have been proven very powerful when sufficient data sets are available, providing quick predictions with minimum computational cost. They find numerous scientific and industrial applications and are used in a wide range of fields such as computer vision, image classification, pattern recognition, natural language processing, noise reduction, prediction, and forecasting [8], [9], [10].

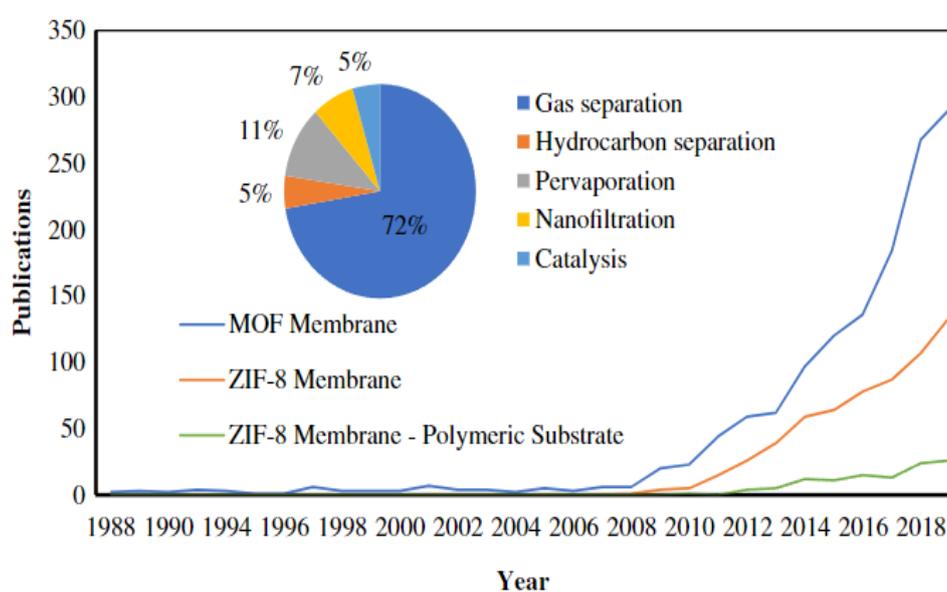
The results presented in this chapter are included in the article that resulted from the research titled “*Molecular and Artificial Neural Networks Modeling of Sorption and Diffusion of Small Alkanes, Alkenes and Their Ternary Mixtures in ZIF-8 at Different Temperatures,*” published in The Journal of Physical Chemistry B, 126(29), 5582–5594 (2022), doi: 10.1021/acs.jpccb.2c03478

## 4.2 Zeolitic Imidazolate Framework 8

Zeolitic Imidazolate Framework-8 (ZIF-8) is an esteemed part of the metal-organic frameworks (MOF) family, with excellent features in gas storage, catalysis, drug delivery, and sensing applications. Metal ions or clusters coupled with organic ligands form MOFs, which have very porous structures and large surface areas. ZIF-8 consists of zinc ions ( $\text{Zn}^{2+}$ ) coupled with imidazolate ligands, forming a three-dimensional network of interconnecting nanopores [11],

[12]. The unique design of ZIF-8 confers extraordinary porosity, as seen by a huge surface area and well-defined pore structure. These holes may selectively adsorb gas molecules based on their size, shape, and chemical properties, making ZIF-8 an excellent choice for gas storage, separation, and purification applications. ZIF-8 has significant functions in gas storage and capture. Its porous structure allows for the adsorption of gases including hydrogen, methane, and carbon dioxide, making it a suitable candidate for fuel storage and carbon capture technologies. Notably, in the field of hydrogen storage for fuel cell vehicles [13], ZIF-8 has the potential to increase storage capacity and efficiency.

Furthermore, ZIF-8 has shown potential in catalytic applications [14]. Its large surface area and tunable pore structure allow for efficient substrate binding and reaction amplification. ZIF-8-based catalysts can be tuned for a variety of chemical transformations by inserting catalytically active sites into the framework, enabling applications in organic synthesis, pollutant degradation and energy conversion, as well as gas storage and catalysis [15], [16]. Its adjustable pore size and surface chemistry allow for the encapsulation and regulated release of medicinal chemicals, making it a viable platform for targeted drug delivery systems. Furthermore, its selective adsorption qualities make it suitable for high-sensitivity and high-selectivity sensing applications that detect specific molecules or analytes. In broad terms, ZIF-8 is a great breakthrough in the field of porous materials, providing a versatile platform with a wide range of applications. Its unique combination of porosity, tunability, and usefulness makes it a great asset in tackling numerous technological difficulties and developing sustainable solutions for the future.



**Figure 4.2:** Number of publications regarding ZIF-8-applications [17].

**Figure 4.2** depicts a substantial rise in interest in using ZIF-8 for a variety of applications, indicating a significant increase in research and development resources devoted to exploring its possibilities. ZIF-8's adaptability and effectiveness have been increasingly apparent in recent years, prompting further research in a variety of sectors [18], [19]. Among the core applications, gas separation and pervaporation stand out as key areas of focus, owing to ZIF-8's exceptional characteristics and performance in these fields. Gas separation [20], which accounts for a significant number of ZIF-8 applications, demonstrates the material's capacity to selectively adsorb and separate gas molecules based on size, shape, and affinity. It is quite clear that the interest for using ZIF-8 for many applications has grown exponentially in the last years. The main applications for ZIF8 would be gas separation and pervaporation [21]. Gas separation has great potential for a variety of applications, including greenhouse gas capture and industrial gas purification [22], [23], [24]. Similarly, pervaporation, another notable use, takes advantage of ZIF-8's unique properties to enable the selective transport of molecules over membranes, providing a promising option for liquid separation processes. The growing interest and investment in these applications underscores ZIF-8's potential to transform gas and liquid separation technologies, accelerating progress in sustainability, energy efficiency, and industrial processes [17].

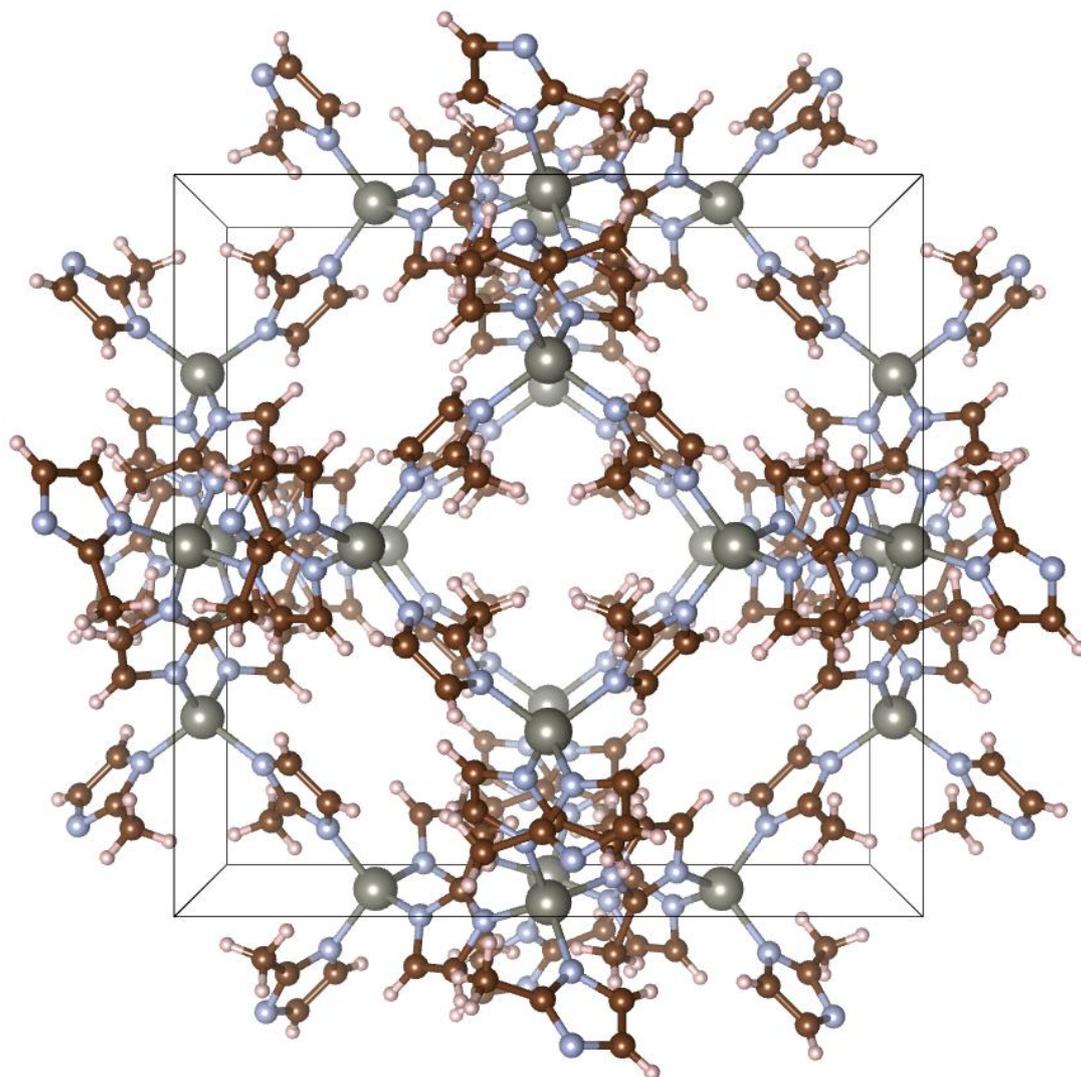
### **4.3 Molecular and ANN Modelling**

This subchapter outlines the adopted modelling approach within this thesis's scope for simulating ZIF-8. Computational modelling is crucial for understanding and predicting MOF behavior under various conditions, such as various temperatures, pressure and different sorbed loading. Moreover, the details about both Molecular simulations and Artificial Neural Networks will be presented in detail. Additionally, this study will present and compare the results obtained from simulation data and neural network predictions with experimental data to validate the efficiency of the simulation techniques employed. Through this comparative analysis, the reliability and accuracy of the simulations will be assessed, greatly enhancing our understanding of the behavior of ZIF-8 and its applicability in practical scenarios.

#### **4.3.1 GCMC and NVT-MD Molecular Simulations**

The unit cell of the ZIF-8 framework was constructed from powder X-ray diffraction data (PXRD) [25] as it is shown in **Figure 4.3**. It contains six four membered rings and eight six-membered rings that correspond to SOD zeolite-type topology [26]. The cubic framework of ZIF-8 is composed of zinc and 2-methylimidazolate linkers with a lattice constant of 16.991

Å. The simulation box used in the present study consists of eight-unit cells created by a  $2 \times 2 \times 2$  repetition in the  $x$ ,  $y$ , and  $z$  directions, respectively. The normal alkane and alkene molecules are flexible molecules, especially when the chain length becomes large, and are modeled by the TraPPE-United Atom force field [27], [28]. The  $\text{CH}_i$  ( $i = 1, 2$ , and  $3$ ) entities of the hydrocarbon molecules are considered as individual united atom dispersive interaction centers.



**Figure 4.3:** ZIF-8-unit cell representation. For the representation VESTA software was used. [29]. Different colors are used for representing different atom types: zinc-grey, carbon-brown, hydrogen-light pink, nitrogen- light blue.

The alkane and alkene bond lengths were assumed rigid and constrained to 1.54 and 1.33 Å, respectively. Interactions with the framework atoms of ZIF-8 were modeled using GenericMOFs force field (as implemented in RASPA2 [30]. GenericMOFs is a mixed force

field based on the Universal Force Field (UFF) [31] and Dreiding force-field [32]. The cross-interaction potential parameters were obtained through Lorentz–Berthelot relations and the cutoff distance for the dispersive interactions were set to  $r_c = 12 \text{ \AA}$ . Under the molecular simulation framework, GCMC and NVT-MD numerical computations were conducted for the qualitative and quantitative description of sorption and diffusion of small alkanes, alkenes such as methane, ethane, ethene and of their ternary mixtures in ZIF-8. For a number of important reasons that came out during the present investigation relative to (i) the coherence of the study, (ii) the extraction of significant missing physical information, (iii) the implementation of ANNs toward more generic predictions based on enriched data sets, especially in the case where experimental measurements are heterogeneous, limited, scarce or even unavailable, and (iv) consistency and validity checking of the adopted computational approaches, additional NVTMD and GCMC molecular simulations were conducted for other small alkane and alkene molecules, such as propane, propene, butane, and butene. In the course of GCMC simulations, three types of MC moves were applied, which allowed the insertion, deletion, and the translation of hydrocarbon molecular entities in ZIF-8. Each trial move is selected and attempted with equal probability. The total MC steps, necessary for the adequate sampling of the relative phase space, during the production stage of the simulations were ranging from  $10^6$  to  $20 \times 10^6$  strongly depending on the external pressure applied and hence with the actual hydrocarbon sorbate loading in the ZIF-8 sorbent material. Regarding GCMC calculations and toward simulating sorption isotherms the adoption of a rigid framework is a common and reasonable approximation, especially for molecules that do not fit tightly in cage openings or cause substantial framework distortion. Thus, the GCMC simulation production runs were carried out with rigid ZIF-8. The MD simulations were conducted in the canonical statistical ensemble mimicking this way physicochemical systems with constant number of molecules, system volume and temperature (NVT). The Verlet algorithm is adopted for the integration of Newton’s equations of motion. The solution of the aforementioned initial value problem is necessary to be connected with the statistical mechanics principles and transport phenomena at molecular level for the description of the actual hydrocarbon sorbate mobility in confinement. The integration step  $dt$  was set equal to 1 fs. The Nosé–Hoover chain thermostat was used in the MD simulations. The total duration of the MD production runs was between 30 and 80 ns depending on the loading for each simulation and proved to be adequate for the system equilibration and for the estimation of the self-diffusivity coefficients [33]. In contrast to GCMC calculations, in the MD approach a flexible ZIF-8 sorbent was considered in order to mimic in a realistic way the breathing mode of the aperture which connects the sorption cages.

The rigidity of ZIF-8 in conjunction with the 3.4 Å small size openings have been proven unrealistic to simulate the actual diffusion process, which necessitates and invokes a jump mechanism along the adjacent cages. The flexible framework was modeled using the proper bond stretching and angle bending parameters related to framework atoms following the work by Krokidas et al. [34] The molecular simulations have been conducted using RASPA2 molecular simulation software[30]. The computational time for the individual production runs was ranging from 2 to 20 days depending on the sorbate loading. All computations have been carried out with a personal computer having an AMD Ryzen 2700X processor equipped with 16 GB of memory running under a Ubuntu 16.04 LTS (Xenial Xerus) operating system.

### 4.3.2 Predictive Modeling Using Artificial Neural Networks

The very basic ANNs architecture encapsulates (i) the input layer, (ii) a number of hidden layers, (iii) the output layer, (iv) an activation function, and (v) a set of weights. In a more formal and mathematical description, a neural network can be expressed as:

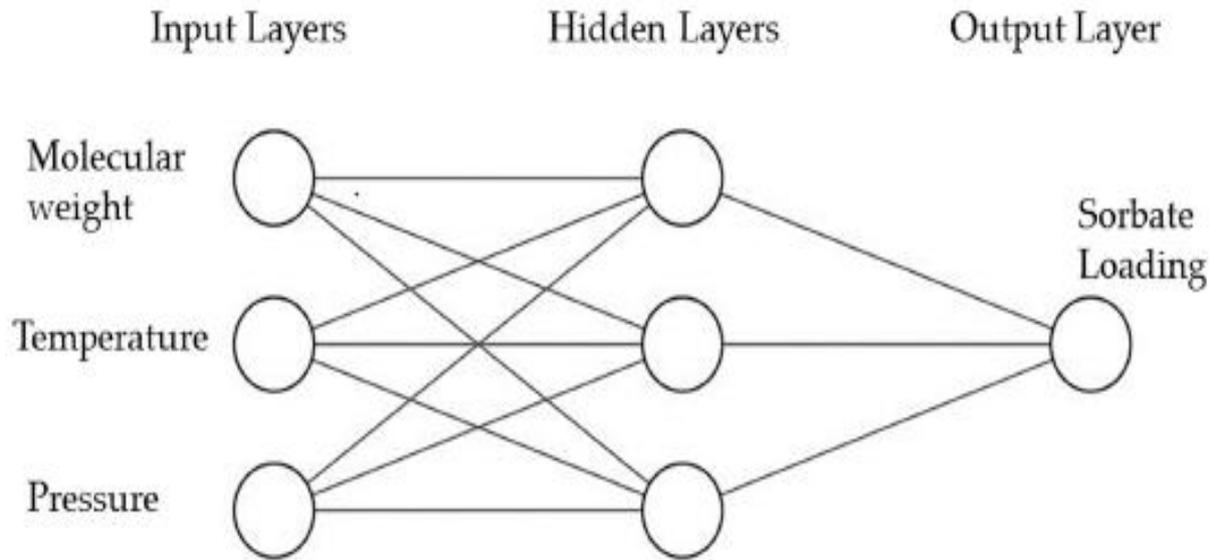
$$u_k = \sum_j w_{kj} \chi_j \quad (4.1)$$

$$y_k = \phi(u_k - \theta_k) \quad (4.2)$$

The quantities involved represent the inputs  $\chi_j$ , the set of weights  $w_{kj}$ , the bias  $\theta_k$  of the system, and the output  $y_k$  where  $\phi$  is the activation function. In the present work, different ANNs were constructed, including one for the sorption and one for the diffusion of pure components following the spirit of our previous predictive modeling work [35] For the case of ternary mixtures, two additional ANNs, which take into account the different types of the sorbent molecules in the mixture, have been developed in order to describe sorption and diffusion in ZIF-8. For the case of mixture sorption characteristics, the predictions are based only to the GCMC data of the present work due to the lack of experimental measurements or other simulation results reported in the relevant literature.

Regarding the sorption predictive modeling the computational approach uses as an input layer the molecular weight ( $M_w$ ), the temperature ( $T$ ), and the pressure ( $P$ ). Consequently, the input vector  $\mathbf{x}_j^{sorption} = (x_1 = M_w, x_2 = T, x_3 = P)_j$  is created. Index  $j$  represents a particular sample from the total number  $N$  of samples that make up the input state space  $X$  that was created from data sets originating from the present GCMC computations. In a

representative ANN architecture **Figure 4.4** for the pure component's sorption predictions developed in this study is given.



**Figure 4.4:** Artificial Neural Network Architecture.

It is worth mentioning that the number of input variables of the actual problem can potentially affect the difficulty by increasing the complexity of the adopted ANNs modeling procedure. The connection in a mathematical quantitative relation of the number of variables with the amount of input data sets is not an easy task and is more or less system specific. Trial and error procedures have been used in the present investigation in order to overcome such difficulties following our previous work [35]. The number of input variables is not the most important factor affecting the accuracy of predictions made by the ANNs. From a theoretical perspective, of significant importance is the amount of data sets (quantity of data used in the training state) along with the number of hidden layers and this should be prioritized in order to have reliable predictions. The hidden layers used in this study were three and were chosen via a trial-and-error procedure and found to be optimal for the cases studied. The number of hidden layers was chosen following the procedure presented by Mas et al. [36] The output layer of the network consists of a 1-D output vector containing the sorbed amounts  $c_{\text{sorption}}$  in units of molecules per unit cell. The activation function adopted for the sorption prediction is the widely used logistic function. This sigmoid-type function has the ability to describe problems with nonlinear behavior.

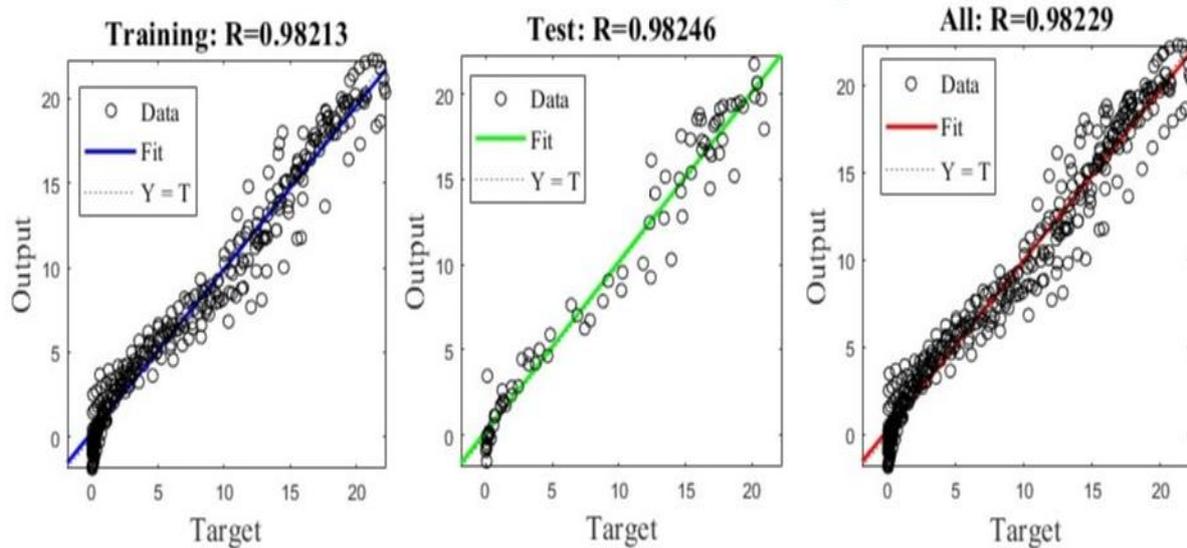
For the ANNs diffusion predictions the input vector invokes the sorbate loading  $c_{\text{sorbate}}$  and has the form  $\mathbf{x}_j^{\text{diffusion}} = (x_1 = M_w, x_2 = T, x_3 = C_{\text{sorbate}})_j$ . The output layer is 1-D output vector containing the self-diffusivities  $D_s$ . In this case, the activation function was the Rectified Linear Unit (ReLU) [37] and is analogous to half-wave rectification used in electrical engineering. ReLU is a commonly used activation function in ANNs defined as the positive part of its argument.

Following our previous work [35], the learning algorithm adopted for the developed ANNs during the training phase was the Back Propagation (BackProp), which potentially allows the reduction of the amount of noise captured by the ANN during the modeling procedure [38], [39] Bayesian regularization is applied in order to avoid overfitting of the number of data used during the training and predictive phases [40]. The set of weights ( $w$ ) follows a random initialization process and is corrected after each epoch of the training state. The epochs were in the range of 1000–1200 which proved to be sufficient in this study. It should be noted that the increase of the epochs is not advised under all circumstances due to the high risk of overfitting the predicted data as shown by Sihna et al. [41] The error of the performed calculations was measured via the Mean Squared Error loss function (MSE) which is a widely used error function for ANNs. MSE has the same units of measurement as the square of the quantity being estimated (units of diffusivity squared or sorbate loading squared).

In the present predictive modeling procedure, an in-house code was developed based on MATLAB's Neural Net Fit library [42]. The ANNs models are provided in the **Appendix C** of the work. It is necessary to bear in mind that molecular simulation and experimental measurements with the modern computer resources can potentially provide useful data sets for ANNs significantly upgrading their reliability, accuracy, and predictive performance with important mutual benefits toward the solution of difficult technological problems. It should be also stressed out that despite the fact that ANNs can provide accurate results it is wise to keep in mind that is a “data driven” methodology, which does not take into account the actual physics and the underlying physicochemical phenomena. Consequently, no physical reasoning would be given or attempted to be given based solely on the ANN computational approach.

For the sorption prediction (pure components and mixtures cases) only data sets from the current GCMC simulations were used. For the diffusion study data sets were enriched by including the MAS PFG NMR measurements of Dvoyashkina et al. [43] Regarding the computer time requirements the training phase of the ANNs for both cases (diffusion and

sorption) required 8 to 10 s to execute, while the prediction phase needed very short execution times ( $t < 1$  s) in order to complete. A representative correlation coefficient for the training, testing phases, and for the combined training-testing is presented in **Figure 4.5**.



**Figure 4.5:** Correlation coefficient for the training, testing and for the combined training-testing phases for the sorption ANN predictions.

The final implementation of the ANNs of the present work (diffusion ANNs and sorption ANNs) were based on simulation data at 273, 293, 313, 343, and 373 K following the MAS PFG NMR [43] experimental protocol. For all the developed ANNs consistency-reliability tests have been performed. Nevertheless, in a proof-of-concept framework it was possible to directly compare ANNs predictions with additional simulation results at intermediate temperatures, sorbate loadings or pressures which showed reasonable agreement for the sorbate/ZIF-8 systems and proved the robustness of the adopted ANNs approach.

It should be mentioned that the training phase of the ANNs for alkanes or alkenes sorption and diffusion required 8 to 10 s to execute on average. This procedure for the pure component sorption prediction was completed with an MSE value of  $2.24325 \times 10^{-3}$  (molecules/uc)<sup>2</sup> and correlation coefficients (i) for the training phase  $R = 0.98213$ , (ii) for the testing phase for  $R = 0.98246$ , and (iii) for the combined trained-testing  $R = 0.98229$ . For the case of pure component diffusion, the procedure completed with an MSE value of  $2.28914 \times 10^{-3}$  (m<sup>2</sup>/s)<sup>2</sup> and correlation coefficients (i) for the training phase  $R = 0.98226$ , (ii) for the testing phase  $R = 0.98352$ , and (iii) for the combined trained-testing  $R = 0.98289$ . In Table 4.1 the relevant information is summarized for all ANNs developed for pure components and mixtures.

**Table 4.1.** Representative Correlation Coefficients (R) for Training (Train), Testing (Test) and Combined (Comb) Phases along with MSE Values for the ANNs Developed for Sorption and Diffusion of Pure Components and Ternary Mixtures Sorbed in ZIF-8. MSE units are (molecules/uc)<sup>2</sup> for sorption and (m<sup>2</sup>/s)<sup>2</sup> for diffusion predictions.

ANNs	R <sub>train</sub>	R <sub>test</sub>	R <sub>comb</sub>	MSE
Sorption (Pure components)	0.98213	0.98246	0.98229	$2.24325 \times 10^{-3}$
Sorption (Mixture)	0.96624	0.96912	0.96723	$2.35487 \times 10^{-3}$
Diffusion (Pure Component)	0.98226	0.98352	0.98289	$2.28914 \times 10^{-3}$
Diffusion (Mixture)	0.97256	0.97951	0.97896	$2.29495 \times 10^{-3}$

## 4.4 Results

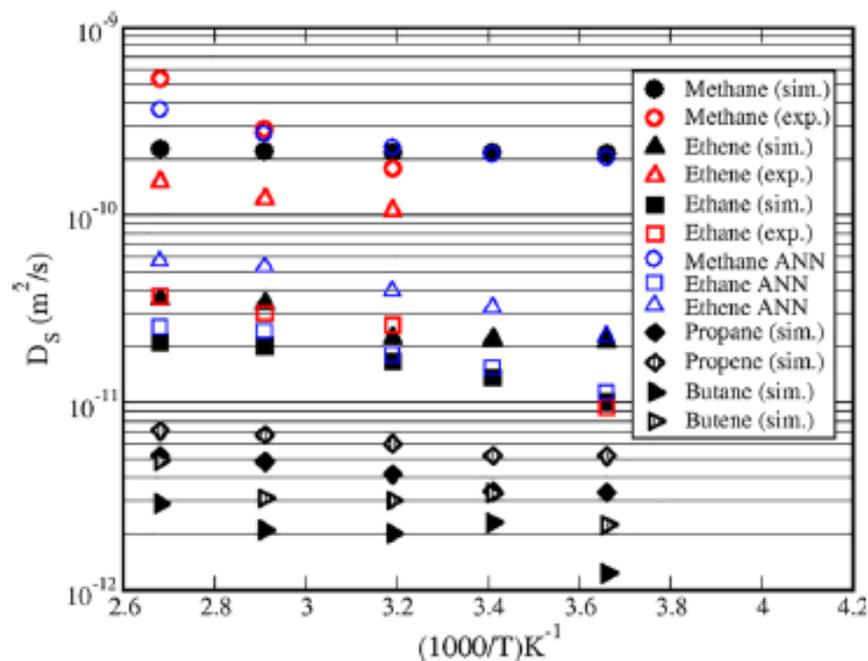
In this section, we present the simulation results obtained from Grand Canonical Monte Carlo (GCMC) and NVT-MD simulations, supplemented by predictions generated by Artificial Neural Networks (ANNs). These findings provide both quantitative and qualitative insights into the sorption and transport properties of small alkanes, alkenes, and ternary mixtures within the ZIF-8 framework. By integrating the results of rigorous computational simulations with the prediction capabilities of ANNs, we hope to provide a thorough knowledge of the sorption and transport behavior of these compounds within the porous structure of ZIF-8. Through multidimensional analysis, we hope to shed light on key factors influencing sorption selectivity, diffusion kinetics, and molecular interactions, thereby contributing to the advancement of knowledge in this field and informing potential ZIF-8 applications in gas separation, purification, and storage technologies.

#### 4.4.1 Self-Diffusivity Coefficients for Pure Components and Mixtures Obtained from NVT-MD, ANNs, and MAS PFG NMR Experiments

The self-diffusivity coefficients can be calculated [44], [45], [46] from the longtime linear part (Fickian) of the mean-squared displacement with respect to time, as can be extracted from the actual small alkanes and alkenes molecular trajectories in ZIF-8.

The MD simulations findings for pure alkanes (from methane to butane) and alkenes (from ethene to butene) are shown in **Figure 4.6** as a function of the upscaled inverse temperature ( $1000/T$ ) along with the ANN predictions and the MAS PFG NMR measurements of Dvoyashkina et al. [43] for direct comparisons. The sorbate loading for all pure alkanes and alkenes was set to six molecules/cage following the MAS PFG NMR experimental protocol. It should be noted that in the experimental work of Dvoyashkina et al. [43] methane, ethane, and ethene, along with their equimolar ternary mixture, have been studied. In the present simulation work, pure propane, propene, butane, and butene molecules have been additionally studied in order to provide a more generic information regarding the diffusion and sorption of hydrocarbons in ZIF-8.

In general, the MD simulations results are in agreement with the experimental measurements following more or less their qualitative or quantitative trends. The decrease of the self-diffusivity coefficients with the decrease of the temperature is the general behavior for the alkane and alkene hydrocarbon sorbates in ZIF-8 as it is apparent in **Figure 4.6**. The ANNs results based on the simulation and experimental data of all alkanes-alkenes give predictions bounded by the MAS PFG NMR measured values and the computed values from the NVT-MD simulations.



**Figure 4.6:** Self-diffusivity coefficients ( $D_s$ ) for pure alkanes and alkenes at different temperatures. Filled symbols represent results obtained from the present simulation study (sim.) and thick nonfilled symbols represent the MAS PFG NMR experimental measurements (exp.) by Dvoyashkina et al. [43] ANNs predictions are given in thin nonfilled symbols.

The methane molecules with the lowest molecular weight among the hydrocarbons used in this work have the highest self-diffusivity values as it is evident from simulation and experimental findings. The simulation results are located in the  $[2 \times 10^{-10}, 3 \times 10^{-10}]$  m<sup>2</sup>/s interval, being practically constant (a very weak almost infinitesimal declining tendency can be detected for the simulation results) as temperature decreases. Methane's  $D_{s, \text{exp}}$  linear inverse temperature dependence depicted in the semilogarithmic scale of the diagram, implies Arrhenius exponential decrease for the self-diffusivity with the temperature's decrease and is being captured by the ANN predictions. The MAS PFG NMR data are limited to three temperatures (373, 343, and 313 K), and despite the fact that the  $D_s^{\text{sim}}, D_s^{\text{exp}}$  values for 343 and 313 K are in close agreement, the exhibited experimental steep reduction for the self-diffusivity of methane is mainly affected by the  $D_s^{\text{exp}}$  (373 K) value. The experimentally measured self-diffusivity value decreases by a factor of 3 from  $D_s^{\text{exp}}$  (373 K) =  $5.36 \times 10^{-10}$  m<sup>2</sup>/s to  $D_s^{\text{exp}}$  (313 K) =  $1.76 \times 10^{-10}$  m<sup>2</sup>/s. The calculated values from the present simulation work are lying in the self-diffusivity window reported by Hertag et al. [47] ( $D_s \in [10^{-10}, 2 \times 10^{-9}]$  m<sup>2</sup>/s), which has been obtained for the diffusion of methane sorbed at various loadings in ZIF-8 at 293 K and for different force fields.

Ethane's and ethene's diffusion is slower compared to methane molecules in all studied temperatures, as is evident in **Figure 4.6**. Interestingly, ethene exhibits higher self-diffusivities compared to ethane, which is in accordance with the independent MAS PFG NMR measurements reported in the research [43], [48] Concerning the pure ethene component based on the Dvoyashkina et al[43] measurements, the relative self-diffusivity between 373 and 313 K attains values  $D_s^{exp}(373\text{ K})/D_s^{exp}(313\text{ K}) = (1.51 \times 10^{-10}\text{ m}^2/\text{s})/(1.06 \times 10^{-10}\text{ m}^2/\text{s}) = 1.43$  for the experimental measurements and  $D_s^{sim}(373\text{ K})/D_s^{sim}(313\text{ K}) = (3.56 \times 10^{-11}\text{ m}^2/\text{s})/(2.21 \times 10^{-11}\text{ m}^2/\text{s}) = 1.61$  for the simulation data, which are in good agreement despite the fact that the simulation underestimates the individual self-diffusivity values. The ANN predictions based on the simulation and experimental data are in good qualitative–quantitative agreement with experimental measurements and NVT-MD computations. A calculation for the self-diffusivity ratio based on the ANN predictions gives  $D_s^{ANN}(373\text{ K})/D_s^{ANN}(313\text{ K}) = (5.66 \times 10^{-11}\text{ m}^2/\text{s})/(3.92 \times 10^{-11}\text{ m}^2/\text{s}) = 1.44$ . It should be noted that the representative diffusivity ratio calculations and direct comparisons between NVT-MD, MAS PFG NMR, and ANN results are limited to the temperatures of 373 and 313 K since the experimental measurements are lacking the diffusivity data for 273 and 293 K temperatures.

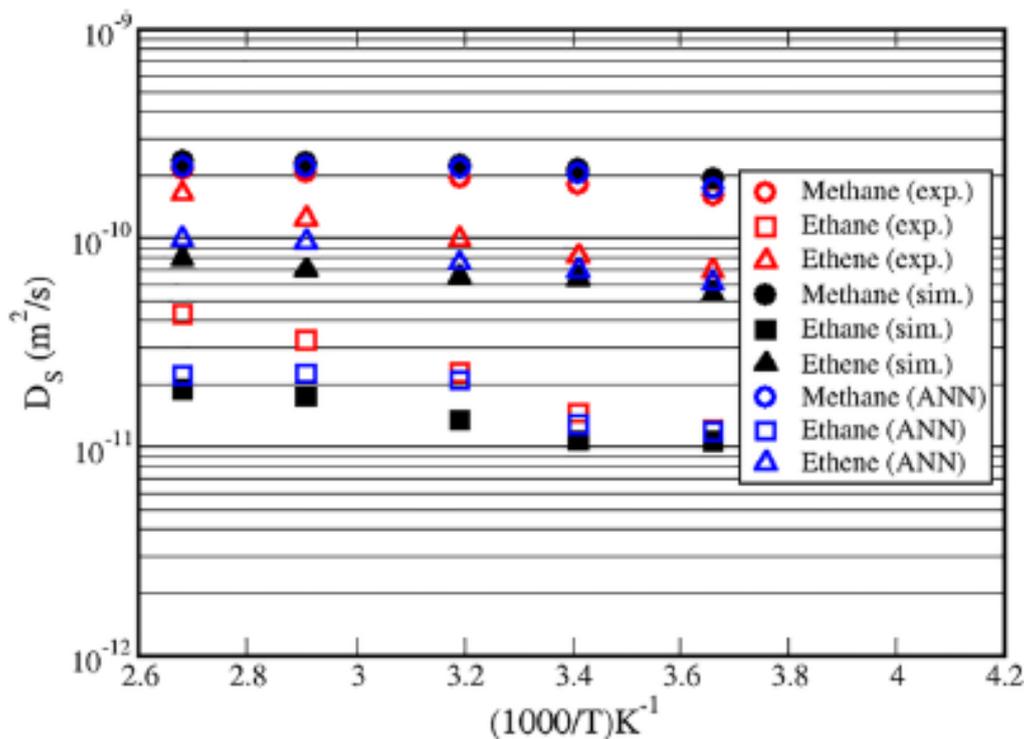
Simulation results for the sluggish ethane molecule are in qualitative and quantitative agreement when compared to NMR measurements. An explicit calculation of the diffusivity ratio between 373 K and for the available experimental diffusivity value at 273 K gives  $D_s^{exp}(373\text{ K})/D_s^{exp}(273\text{ K}) = (3.7 \times 10^{-11}\text{ m}^2/\text{s})/(9.5 \times 10^{-12}\text{ m}^2/\text{s}) = 3.9$ , for the simulation data we have  $D_s^{sim}(373\text{ K})/D_s^{sim}(273\text{ K}) = (2.10 \times 10^{-11}\text{ m}^2/\text{s})/(1.01 \times 10^{-11}\text{ m}^2/\text{s}) = 2.1$  and for the ANN predictions we have  $D_s^{ANN}(373\text{ K})/D_s^{ANN}(273\text{ K}) = (2.51 \times 10^{-11}\text{ m}^2/\text{s})/(1.14 \times 10^{-11}\text{ m}^2/\text{s}) = 2.2$ .

These calculations reveal a significant decrease of self-diffusivity with the decrease of temperature but the NVT-MD results and ANN predictions (based on data from all alkanes and alkenes over all temperatures studied) do not support the diffusivity reduction by a 3.9 factor found in the MAS PFG NMR experiments. It should be noted that experimental IR microimaging self-diffusivity measurements reported for pure ethane at 298 K and different sorbate loadings (0.2, 2, and 3.6 molecules/cage) [49] were close to  $10^{-11}\text{ m}^2/\text{s}$  and for 4 molecules/cage was around  $1.5 \times 10^{-11}\text{ m}^2/\text{s}$  [4]. These values are in agreement with the MD simulation value of  $1.37 \times 10^{-11}\text{ m}^2/\text{s}$  obtained at 293 K and with an individual ANN prediction  $1.47 \times 10^{-11}\text{ m}^2/\text{s}$  at 298 K both at sorbate loading of 6 molecules/cage.

Complementary NVT-MD computations have been performed for pure linear hydrocarbons including propane, propene, butane and butene at the same 6 molecules/cage sorbate loading also presented in **Figure 4.6**. As the temperature decreases, the decrease of self-diffusivity is the general trend for the linear chains with three and four carbon atoms. It is clear that as the molecular weight of alkane increases the self-diffusivity decreases, resulting in  $D_s^{butane} < D_s^{propane} < D_s^{ethane} < D_s^{methane}$ . Butane is the hydrocarbon molecule with the lowest  $D_s$  values in the  $[10^{-12}, 3 \times 10^{-12}]$  m<sup>2</sup>/s range for the temperatures studied. A similar behavior is exhibited by the alkene molecules with  $D_s^{butene} < D_s^{propene} < D_s^{ethene}$ . In addition, as shown for ethene and confirmed for propene and butene the alkenes have higher self-diffusion coefficients than their alkane counterparts. The ANNs predictions for the cases of propane, propene, butane, and butene are not presented (but are provided in the **Appendix D**), since the respective data points used in the training state solely originate from the present MD simulations due to the lack of experimental measurements for these molecules. At this point it is necessary to note that the reduction of self-diffusivity as the molecular weight of the hydrocarbon chain sorbed in ZIF-8 increases has been supported by IR microimaging (IRM) experimental measurements which were further analyzed with Transition State Theory (TST) calculations at 298 K [5], [49]. From our MD simulations and for 6 molecules/cage sorbate loading at 293 K, the self-diffusivity of propane is 1 order of magnitude lower than ethane's, with  $D_s^{sim,ethane}(293\text{ K})/D_s^{sim,propane}(293\text{ K}) \approx 4$  and is similar to the behavior for ethene and propene, where  $D_s^{sim,ethene}(293\text{ K})/D_s^{sim,propene}(293\text{ K}) \approx 4$ . These reductions are a consequence of the larger size of the C3 hydrocarbon. It should be also noted that for propane and propene in ZIF-8 the microimaging results at 298 K and for sorbate loadings lower than 6 molecules/cage (the loading of present work) revealed a dramatic decrease of the transport and corrected diffusivity values (even by 4 order of magnitude) attaining values close to  $10^{-15}$  m<sup>2</sup>/s. The significant drop is not captured by the MD simulation results. This can be attributed to the idealized ZIF-8 framework, which is modeled without defects, cracks or degradation effects that can be present in a real sample. The force field interaction parameters along with the modeling characteristics of the window's breathing mode can also play a significant role especially when the hydrocarbons fall in the range where the effective ZIF-8 aperture size becomes more selective being commensurate to critical guest sizes of 4–5 Å.

#### 4.4.2 Self-Diffusivities for Ternary Mixtures of Methane–Ethane–Ethene.

Additional NVT-MD simulations have been performed to study the diffusion of methane, ethane and ethene in their equimolar mixture at the temperature range of 273–373 K in ZIF-8. The sorbate loading for each component of methane–ethane–ethene mixture was set to 2 molecules/ cage following the MAS PFG NMR experimental protocol. It is clear from the self-diffusivities reported in **Figure 4.7** that MD simulations are in reasonable agreement with MAS PFG NMR measurements[43]



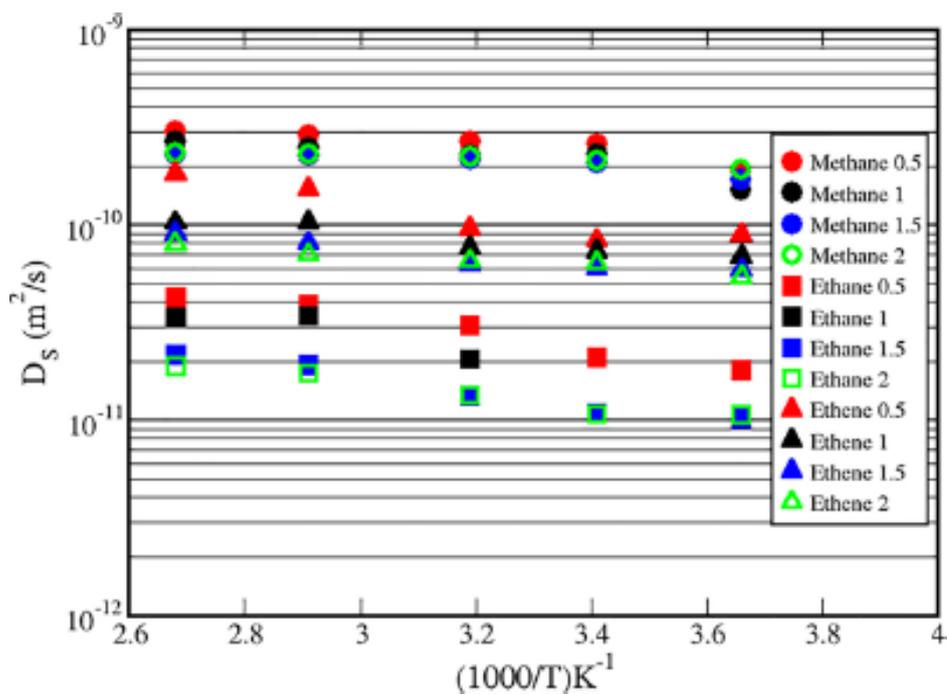
**Figure 4.7:** Self-diffusivity coefficients ( $D_s$ ) for the methane–ethane– ethene equimolar ternary mixture conducted at different temperatures. Filled symbols represent results obtained from this simulation study (sim.) and nonfilled symbols represent experimental MAS PFG NMR measurements (exp.) by Dvoyashkina et al. [43] The sorbate loading for each component of methane–ethane–ethene mixture was set to 2 molecules/cage following the MAS PFG NMR experimental protocol.

It is also evident that among the three light hydrocarbons in the mixture, methane exhibits the largest  $D_s^{methane,sim}(373\text{ K}) = 2.32 \times 10^{-10}\text{ m}^2/\text{s}$  value, which compares well with  $D_s^{methane,exp}(373\text{ K}) = 2.13 \times 10^{-10}\text{ m}^2/\text{s}$ . Ethane is the molecule with the lowest long distance mobility with  $D_s^{ethane,sim}(T)$  values bounded in the  $[1.06 \times 10^{-11}, 1.88 \times 10^{-11}]\text{ m}^2/\text{s}$  range defined by the lower and upper temperatures of 273 and 373 K, respectively. It is an interesting fact that the hydrocarbon mixture components exhibit diffusive characteristics inherited from their pure sorbate behavior. At the lowest studied temperature of 273 K simulation self-

diffusivity values are showing the following quantitative relations:  $D_s^{methane,sim} \approx 18 D_s^{ethane,sim}$ ,  $D_s^{methane,sim} \approx 3.5 D_s^{ethene,sim}$ , and  $D_s^{ethene,sim} \approx 5.1 D_s^{ethane,sim}$  and are directly comparable and in reasonable agreement with the MAS PFG NMR reported values that give  $D_s^{methane,exp} \approx 13.3 D_s^{ethane,exp}$ ,  $D_s^{methane,exp} \approx 2.3 D_s^{ethene,exp}$  and  $D_s^{ethene,exp} \approx 5.8 D_s^{ethane,exp}$  respectively. At the highest studied temperature of 373 K, self-diffusivity value simulation predictions give  $D_s^{methane,sim} \approx 12.3 D_s^{ethane,sim}$ ,  $D_s^{methane,sim} \approx 2.9 D_s^{ethene,sim}$ , and  $D_s^{ethene,sim} \approx 4.2 D_s^{ethane,sim}$ , with the respective values from the experiments being approximately 5.0, 1.3, and 3.8. These findings support a potential diffusivity driven separation ability of the three studied hydrocarbons in ZIF-8. The simulation results overestimate the diffusivity separation factor compared to MAS PFG NMR measurements.

In order to investigate the self-diffusivity dependence on the loading additional NVT-MD simulations have been carried out for equimolar methane–ethane–ethene sorbed mixture with the individual component sorbate loading set to 1, 2, and 3 molecules/uc (equivalent to 0.5, 1, and 1.5 molecules/cage), which represent systems having a total loading of 3, 6, and 9 molecules/uc.

As can be seen from the MD results presented in **Figure 4.8** the increase of the individual and inevitably of the total loading in the sorbed mixture does not significantly affect the self-diffusivity of methane. Ethene and ethane self-diffusivity values show a clear sorbate loading dependence. For ethane as the individual methane loading increases from 0.5 to 2 molecules per cage the diffusivity reduces by a factor of 1.7 at 273 K and by 2.3 at 373 K, while the average decrease over the different temperatures studied is close to 2.1. Ethene in all loadings and temperatures exhibits higher molecular mobility than ethane. The increase of the sorbed loading affects the diffusivity, and a drop by a factor of 1.64 at 273 K and by 2.27 at 373 K was calculated as ethane-sorbed loading increases from 0.5 to 2 molecules per cage (equivalent to 1.5 and 6 molecules per cage total loading of the mixture, respectively). For ethene, an average of the diffusivity loading driven reduction factor over the different temperatures is calculated close to 1.7.

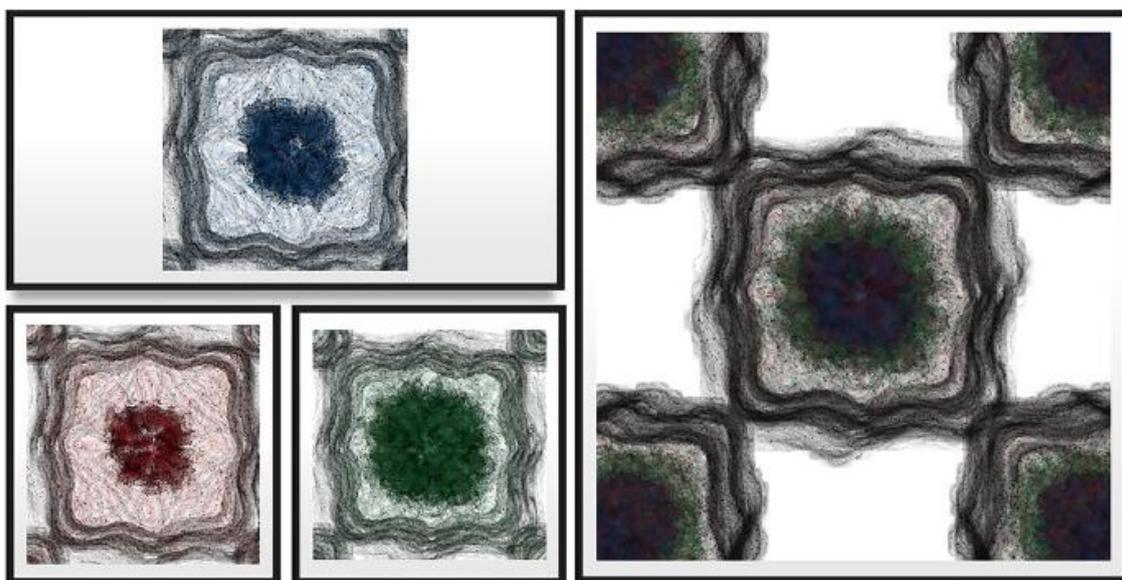


**Figure 4.8:** Self-diffusivity ( $D_s$ ) values for the methane–ethane–ethene sorbed mixture from MD simulations conducted at the 273–373 K range. Individual component sorbate loadings refer to 0.5 (red), 1 (black), 1.5 (blue), and 2 (green with empty symbols shown for reference) molecules/cage.

As mentioned, intracrystalline mobility depends on the actual molecular hydrocarbon molecular weight and is higher for methane, with ethane being the sluggish molecule. It is clear that the actual molecular size of the molecules plays a crucial role on the self-diffusivity value resulting in the inequality  $D_s^{methane} > D_s^{ethene} > D_s^{ethane}$ . Beyond that, the siting preferences and the actual proximity of the molecules in motion with the ZIF-8 apertures can affect the intracrystalline motion and the associated hopping mechanisms and jump rates.

The molecular long-distance traversal necessitates motion along different cages. This intercage motion can be significantly affected and facilitated by the preferential localization of sorbate molecules close to the narrow apertures. Calculations from the actual molecule positions ( $r$ ) of the sorbed single particle density  $\rho_k^{(1)}(r)$  distributions [30], [33], [35], [50] ( $k$  stands for the particular alkane or alkene) along with three dimensional (3D) contouring depicted for a representative ternary mixture at 293 K in **Figure 4.9** reveal that ethene and particular ethane molecules are strongly localized at the cages' central regions. On the other hand, methane density contours are clearly more extended and closer to the openings connecting adjacent cages. Frequent jump events occurring between different cages can be significantly promoted by methane's residence in the vicinity of apertures and can be reflected to the higher self-diffusivity values. In addition, the more extended density clouds of ethane

compared to ethene at the central regions can potentially lead to significant reduction of molecular mobility and consequently of self-diffusivity.



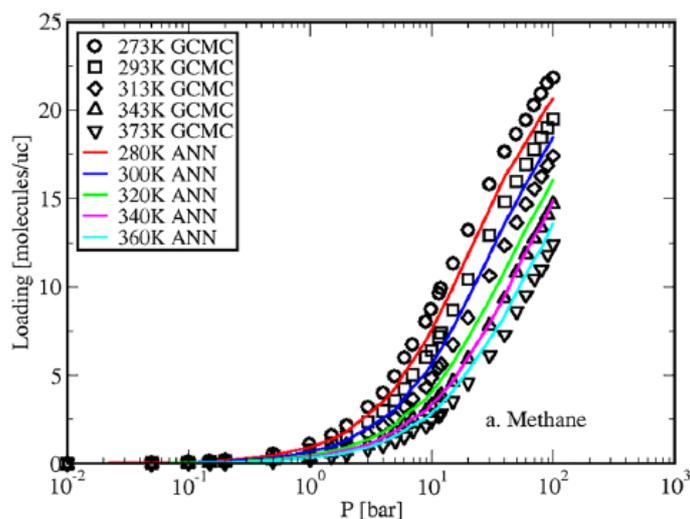
**Figure 4.9:** Three-dimensional contours plot of single particle density distributions for the methane–ethane–ethene sorbed mixture from MD simulations at 293 K. Methane is shown in green, ethane in red, and ethene in blue. The separated contours of density are shown for clarity on the left side, while the actual distributions of the ternary sorbed mixture are given in the right side.

#### **4.5 Sorption Isotherms of Small Alkanes, Alkenes, and Their Mixtures in ZIF-8.**

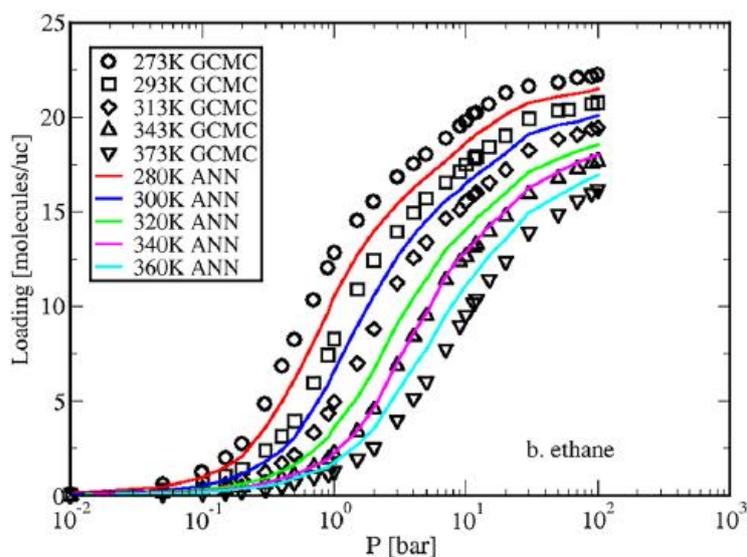
The sorption isotherms from the GCMC simulations along with the ANNs predictions are presented for pure alkanes and alkenes including methane, ethane, ethene, propane, propene, butane, and butene along with a methane–ethane–ethene equimolar gas mixture feed for the 273–373 K range, in accordance with the systems presented in previous sections and for a pressure range between  $10^{-2}$  and  $10^2$  bar. The ANNs results are shown for 280, 300, 320, 340, and 360 K, which are intermediate temperature values with respect to those calculated from the GCMC simulation temperature protocol. These intermediate temperature ANN predictions will be proven to be consistent following the qualitative and quantitative trends of the GCMC sorption isotherms.

### 4.5.1 Sorption Isotherms of Pure Components.

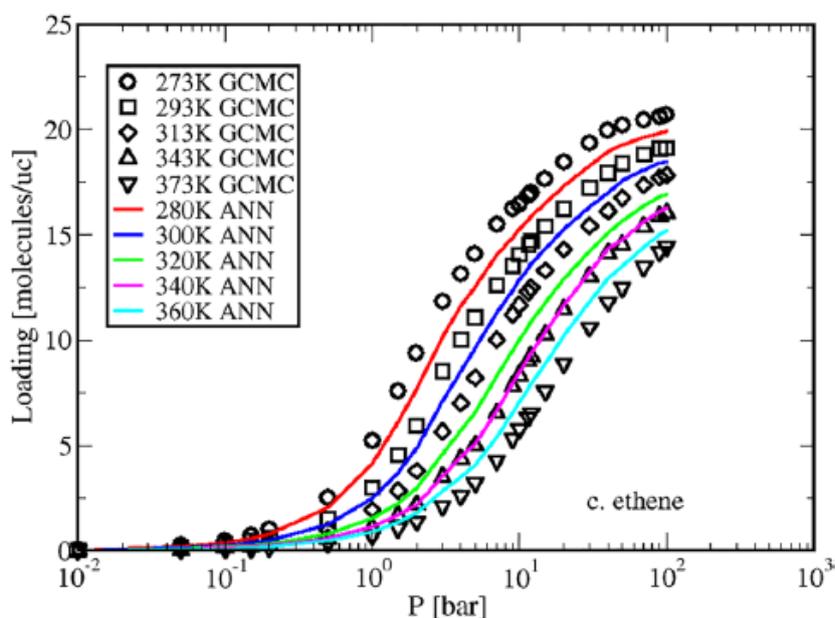
The sorption isotherms of pure methane, ethane, and ethene are depicted in **Figure 4.10**, **Figure 4.11**, **Figure 4.12**. For the  $[10^{-2}, 10^2]$  bar window methane, ethane, and ethene sorbed molecules follow a rather monotonic increase. It is also clear that the higher the temperature, the lower the number of methane molecules sorbed in ZIF-8 for a given pressure value. The ANN predictions follow the qualitative and quantitative trends of the GCMC isotherms at different temperatures.



**Figure 4.10:** Pure Methane GCMC sorption isotherm for the 273-373 K temperature range. ANN predictions are given for intermediate values at 280, 300, 320, 340, and 360 K.



**Figure 4.11:** Pure Ethane GCMC sorption isotherm for the 273-373 K temperature range. ANN predictions are given for intermediate values at 280, 300, 320, 340, and 360 K.



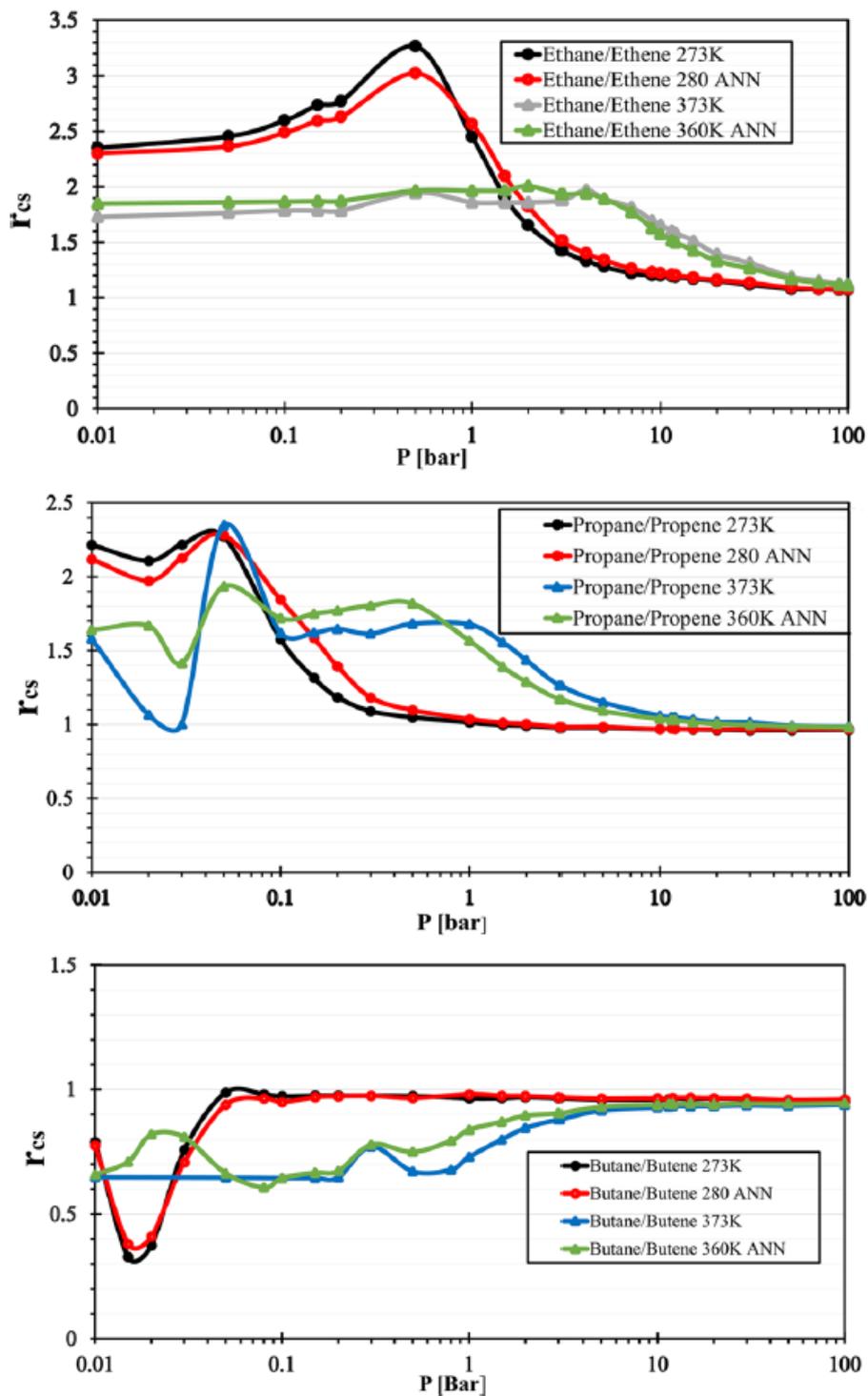
**Figure 4.12:** Pure Ethene GCMC sorption isotherm for the 273-373 K temperature range. ANN predictions are given for intermediate values at 280, 300, 320, 340, and 360 K.

For the sorption of methane in ZIF-8, depicted in **Figure 4.10**, it is clear that the saturation of the loading necessitates elevated pressures beyond the aforementioned pressure window. Simulations above the  $10^2$  bar, although feasible for in silico “experiments”, might be of minor importance or unrealistic for the actual experimental conditions. The sorbate loading at 273 K and 100 bar for methane molecules sorbed in ZIF-8 was calculated close to 22 molecules/uc. The sorbed amount ratio between the highest (373 K) and the lowest temperature (273 K) for 100 bar is approximately equal to 1.8.

The first molecules with carbon–carbon bonds, which include ethane and ethene are also shown in **Figure 4.11**, **Figure 4.12**. At first glance it can be observed that ethane and ethene molecules can be sorbed in ZIF-8 at comparable loadings for the higher-pressure window of [10, 100] bar. The loadings at the maximum pressure studied were  $c_s^{ethane}(273\text{ K}, 100\text{ bar}) = 22.2$   $c_s^{ethene}(273\text{ K}, 100\text{ bar}) = 20.7$  molecules/uc, as can be seen in **Figure 4.11**, **Figure 4.12**.

In order to have more detailed information regarding the sorption in all pressure values ( $P$ ) and temperatures, the ratio of the sorbed number of alkane molecules over alkene molecules, defined as  $r_{cs}(T, P) = c_s^{alkane,pure}(T, P) / c_s^{alkene,pure}(T, P)$ , was computed. GCMC and ANN computations of  $r_{cs}$  for pure ethane and ethene are shown in **Figure 4.13** for

representative temperatures, revealing some interesting qualitative and quantitative characteristics.



**Figure 4.13:** Relative sorption loadings ( $r_{cs}$ ) for pure ethane and ethene at  $T = 273$  and  $373$  K. The ANN  $r_{cs}$  predictions are given for  $T = 280$  and  $360$  K.

It is possible for  $r_{cs}$  to attain values ranging from 2.4 to 3.4 at 273 K and 1.7 to 2 at 373 K. At the temperature of 273 K and for low pressure values in the [0.01, 1] bar range,

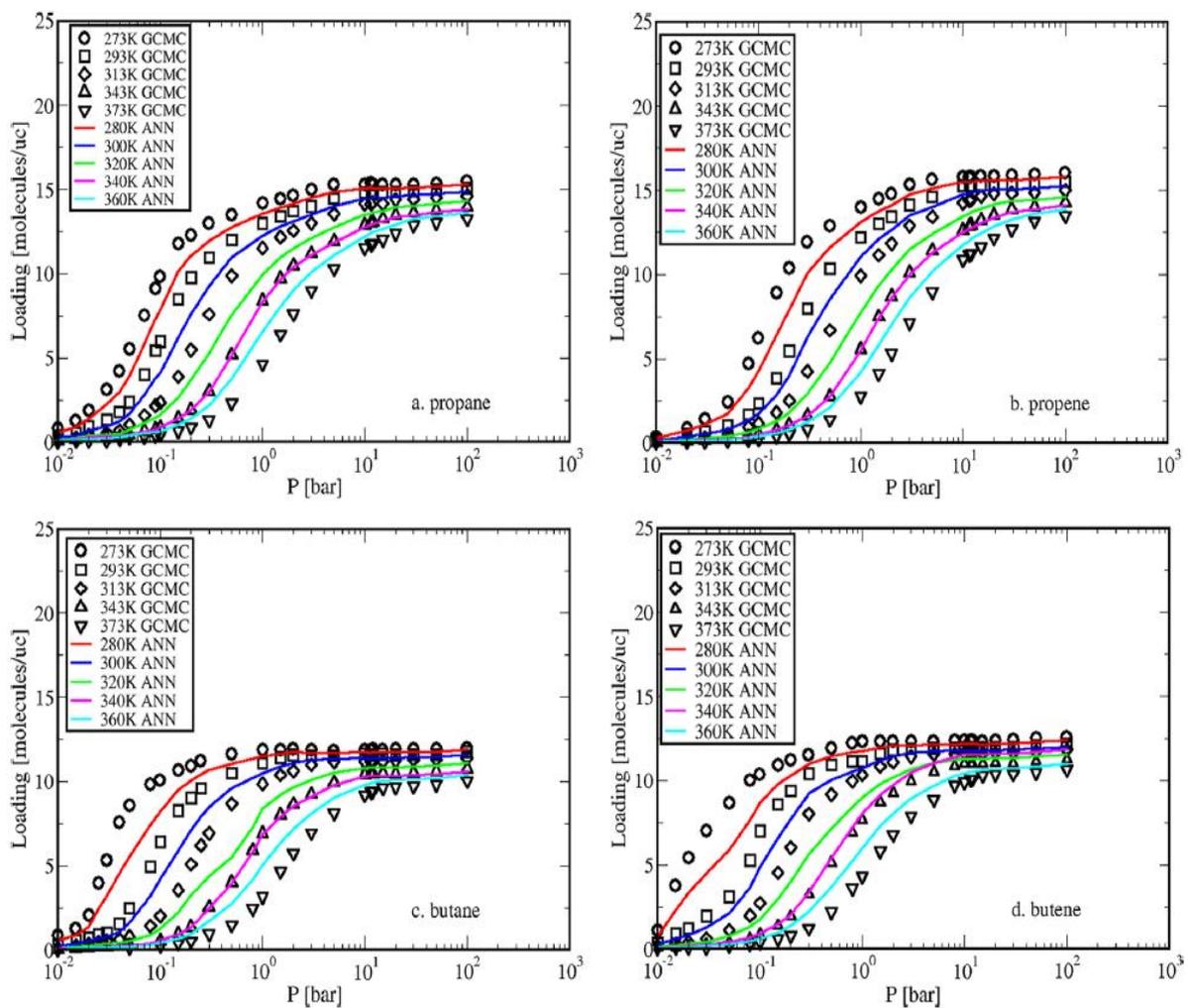
which includes the Henry's law region, ethane can be sorbed in higher amounts compared to ethene in ZIF-8, and consequently, the  $r_{cs}$  (273 K, P) curve shows an increasing tendency for pressures up to 0.5 bar. This increasing tendency is not followed by the  $r_{cs}$  (373 K, P) curve, which resembles a rather constant function as we move from the low pressures to the pressure of 4 bar, despite the fact that ethane is adsorbed in higher amounts compared to ethene ( $r_{cs} \approx 1.8$ ). It is worth mentioning that there is a crossover pressure around 1.7 bar. Beyond the aforementioned pressure, the  $r_{cs}$  (373 K, P) curve values exceed the values of  $r_{cs}$  (273 K, P). The increase of pressure beyond 0.5 bar at 273 K and 4 bar at 373 K triggers the decrease of  $r_{cs}$  values, which results in the final equalization of molecules sorbed for pressure values beyond 70 bar at the aforementioned temperatures.

The ANN predictions, which are based on data over all temperatures and molecular species, closely follow the ethane and ethene sorption preference trends and values not only regarding the sorption isotherms, but in addition the qualitative and quantitative characteristics of  $r_{cs}$ , as can be seen in **Figures 4.10, 4.11, 4.12**.

The isotherms for pure propane and pure propene sorption in ZIF-8 are shown in **Figure 4.14**, and for pressure values higher than 10 bar, they exhibit a plateau region indicating sorbate loading saturation. The highest saturation value can be obtained for the lower temperature of 273 K, with the propane saturation loading value being 15.5 molecules/uc, while for propene it is close to 16 molecules/uc. In the [0.01, 1] bar range, propane can be sorbed in higher amounts compared to propene, as indicated by the GCMC simulation data (273 K) and ANN predictions (280 K) with the  $r_{cs}$  maximum value located around 2.2, as can be seen in **Figure 4.13**. For the pressure range of [0.1, 1] bar,  $r_{cs}$  drastically decreases to one, which indicates that the ZIF-8 material does not show any particular sorption preference of pure propane over pure propene at the temperature of 273 K. This behavior remains practically unchanged, with pressure values further increasing from 10 to 100 bar. Interestingly the sorption performance of ZIF-8 differentiates at elevated temperatures and in particular for 373 K. For the pressure range of [0.1, 1] bar, ZIF-8 exposes a steady preference behavior for the pure ethane with  $r_{cs}$  bound on the [1.6, 1.8] interval. Higher pressures beyond 10 bar are necessary in order to have  $r_{cs} = 1$ , which is indicative for pure components and high pressures of ZIF-8 that does not have any particular sorption preference for propane over propene.

For butane and butene, the longest hydrocarbon molecules studied, the sorbate saturation can be achieved for even lower pressure values compared to propane and propene,

as is evident from **Figure 4.14c, d**. In particular, the saturation loading value is around 12 molecules/uc for butane and 12.5 molecules/uc for butene at  $T = 273$  K, with the plateau value starting above 0.4 bar. The saturation loading is lower for elevated temperatures, and the saturation pressure threshold is being shifted to higher pressure values above 10 bar for  $T = 373$  K. The saturation loading for butane is close to 10 molecules/uc, while for butene it is slightly higher and close to 10.7 molecules/uc. The ANNs predictions for the intermediate temperatures capture the tendencies and sorbed quantities reasonably well and are located within the GCMC results for neighboring temperatures.



**Figure 4.14:** GCMC simulation sorption isotherms of short pure alkanes and alkenes at different temperatures: (a) propane, (b) propene, (c) butane, and (d) butene. The ANN predictions are shown in a continuous line format.

The calculation of  $r_{CS}$  in the case of pure butane and pure butene, which is depicted in **Figure 4.13**, reveals some interesting differences with respect to the cases studied so far. The

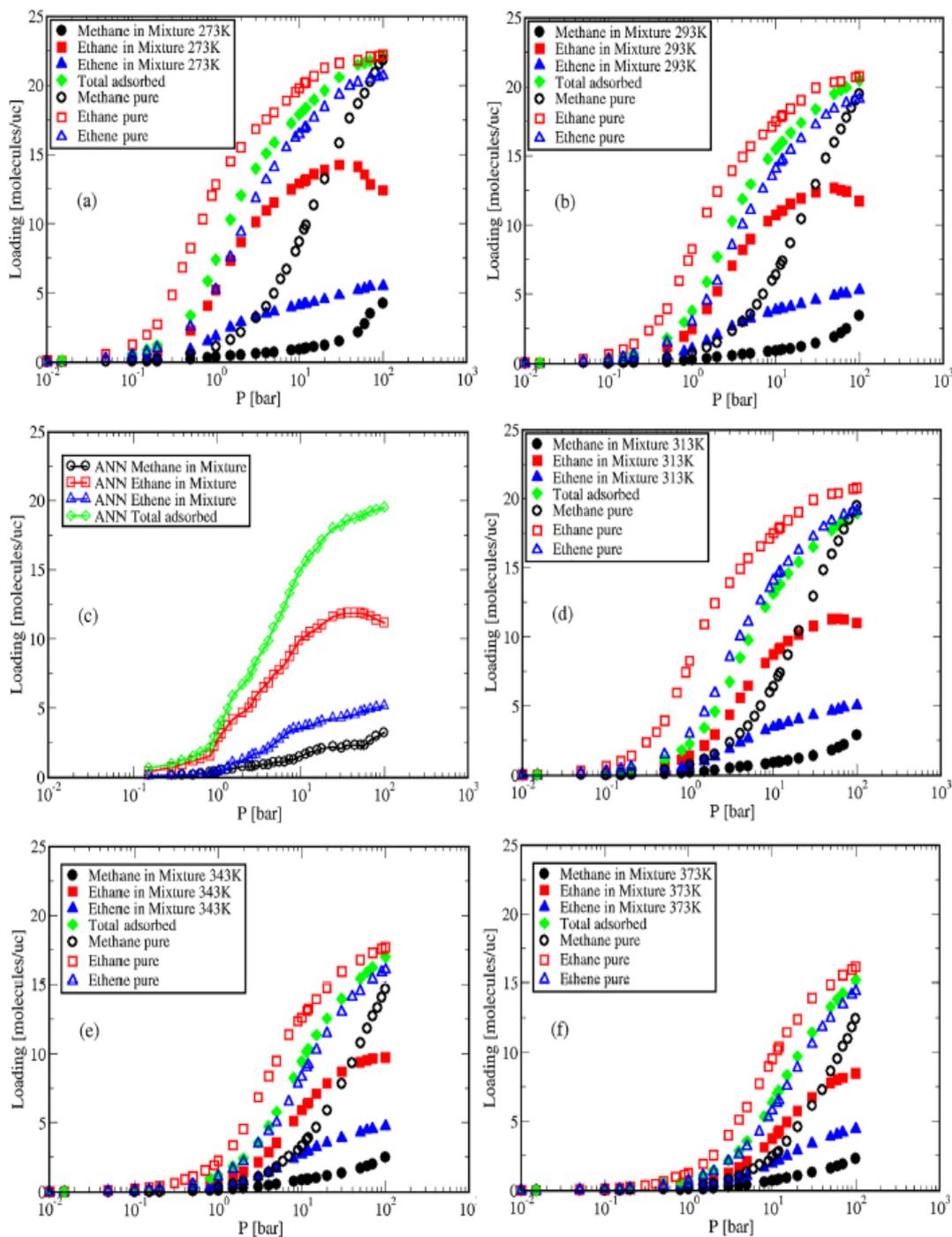
first striking difference is that  $r_{CS}$  remains constant, attaining values equal to 1 at the lower temperatures, 273 K for the GCMC computations and 280 K for the ANN predictions, throughout the extended pressure window of [0.05, 100] bar. Under these physical conditions, ZIF-8 shows no particular adsorption preference for butane over butene in contrast to other alkane–alkene pairs investigated in this work. The second significant difference is that for elevated temperatures at 360 and 373 K the ZIF-8 sorbent shows a clear butene preference over butane from the very low pressures up to the 1 bar pressure. This can be attributed to the more favorable attractive dispersive interactions with the ZIF-8 framework that butene molecules can experience compared to their butane counterparts. The aforementioned preference gradually vanishes with the increase of the pressure from 1 to 10 bar. For high pressures, no sorption preference of butane over butene can be detected. It should be pointed out that the ANN predictions closely follow the GCMC findings quite well, despite the  $r_{CS}$  overshoot at 360 K and pressure around 0.3 bar.

Taking into account the aforementioned GCMC computations and ANNs predictions it can be concluded that in the majority of cases with pure components the number of alkane molecules sorbed in the unit cell of ZIF-8 is higher compared to the respective alkene for the [1, 100] bar pressure window except butane and butene molecules which show the opposite behavior mainly at elevated temperatures.

#### **4.5.2 Sorption Isotherms of Methane-Ethane-Ethene Mixture**

In **Figure 4.15 a, b, d, f**, the sorption isotherms of methane, ethane, and ethene in their equimolar gas mixture feed at different temperatures (a representative ANN sorption isotherm prediction at 300 K is also included in **Figure 4.15, c**) are presented along with the pure components sorption data which serve as reference curves for direct comparisons.

At first glance it can be observed that the sorption characteristics of the mixture significantly differentiate from the pure hydrocarbon's behavior. A competition between the mixture constituents for sorption is evident. In particular is reflected to the lower number of hydrocarbon molecules from the gas mixture which are actually sorbed compared to their pure component behavior.



**Figure 4.15:** GCMC sorption isotherms for ternary methane–ethane–ethene equimolar gas mixture feed at different temperatures ranging from 273 to 373 K (a, b, d–f). ANN predictions for the temperature of 300 K are shown in (c). Pure component GCMC results (open symbols) are shown for reference.

The available free space for occupation by the hydrocarbon mixture components hand in hand with the increase in pressure facilitates the monotonic increase of their sorbed amounts in ZIF-8. The number of ethane and ethene molecules sorbed in the unit cell shows a constant increase in low, intermediate and high values of pressure while the methane molecules are sorbed in considerable amounts ( $\geq 2$  methane molecules/uc  $\equiv$  1 methane molecule/cage) only for pressures beyond 50 bar at 273 K which is an entropy driven behavior. This characteristic pressure for methane's sorbate loading activation depends on temperature and increases to 60, 65, 70, and 85 bar for the temperatures of 293, 313, 343, and 373 K, respectively. ANN prediction at 300 K gives a pressure value around 60.5 bar.

It is interesting the fact that as pressure increases the ethane sorbed molecules start to fall off as can be seen for ethane's sorption isotherms at 273, 293, and 313 K. This sudden loading decrease potentially provides essential space to the other mixture constituents for sorption. A closer inspection of the aforementioned isotherms not only for ethane but also for ethene and methane can lead to another interesting sorption characteristic which is associated with the pressure threshold signaling the reduction of ethane molecules sorbed. In particular, at 273 K and for ethane's threshold pressure which is around 40 bar the methane's isotherms in the semilogarithmic scale representation of **Figure 4.15** exhibits a clear change of curve's slope defining two distinct linear regions. A linear with low slope value in the [1, 40] bar pressure window and a second linear region with increased slope in the [40, 100] bar interval. This abrupt change on methane's isotherm, which is also evident for 293, 300 (ANN also captured the change), and 313 K, is not followed by ethene's isotherms. Therefore, it can be concluded that primarily the methane molecules at the higher pressures seek residence in the limited space created by the decrease of the ethane molecules sorbed in ZIF-8. The aforementioned pressure threshold corresponds to methane loading of 1.5 molecules/uc and is mainly unaffected by the increase of the temperature from 273 to 313 K, something that can be related to the available space for sorption in the two distinct cages of the unit cell. For the elevated temperatures of 343 and 373 K, the maximum applied 100 bar is not adequate to trigger the aforementioned reduction of the sorbed ethane molecules.

The ZIF-8 molecular sorbent shows its strong preference for ethane molecules. It is clear that ethane molecules for intermediate and higher-pressure values are sorbed in significantly higher number of molecular entities compared to ethene and methane molecules in their ternary mixture establishing the loading inequality relation  $c_s^{ethane} > c_s^{ethene} > c_s^{methane}$ . The loading  $c_s$  refers to the number of molecules sorbed in the unit cell of

ZIF-8. For the hydrocarbon mixture the sorption selectivity is defined as  $S_{i,j} = \frac{x_i/x_j}{y_i/y_j}$  ( $i \neq j$ , with  $i, j$  referring to methane, ethane, ethene) and which in the case of the present equimolar gas mixture feed can be reduced to  $S_{i,j} = \frac{x_i}{x_j} = \frac{c_s^i}{c_s^j}$  and can be readily calculated from the actual number of molecules sorbed  $c_s^i, c_s^j$  of hydrocarbons  $i$  and  $j$  in ZIF-8, respectively.

Representative calculations from the GCMC simulations at the temperature of 273 K give for the pressure of 10 bar  $S_{ethane,ethene}(10\text{ bar}) = 3.2$  and  $S_{ethane,methane}(10\text{ bar}) = 14.1$ . At the higher temperature of 373 K, the respective calculations give  $S_{ethane,ethene}(10\text{ bar}) = 1.9$  and  $S_{ethane,methane}(10\text{ bar}) = 5.4$ . In addition, ANN predictions at 300 K and the same pressure give  $S_{ethane,ethene}^{ANN}(10\text{ bar}) = 2.7$  and  $S_{ethane,methane}^{ANN}(10\text{ bar}) = 10.9$ . The highest selectivities monitored for ethane over methane and for ethane over ethene were close to 17 and 3.2, respectively, for pressures in the [2, 10] bar window and at 273 K. The GCMC and ANN results reveal a significant sorption selectivity of ethane over ethene and methane in ZIF-8 which can be achieved in a range of pressures and for different temperatures above and below the room temperature. The selective behavior of ethane over methane in ZIF-8 at the temperature of 298 K in equimolar methane–ethane–propane ternary mixture has been manifested in the simulation findings of Zhang et al. [51]. In their study the longer hydrocarbon was propane, and it was sorbed in significant higher amounts compared to other mixture components. Ethane was sorbed in considerable amounts higher than a methane molecule. In addition, a similar reduction of the longer chain molecules (propane) sorbed was found beyond a threshold pressure promoting additional sorption of ethane and methane molecules. It can be concluded that in ZIF-8 the molecular weight and consequently the length of the hydrocarbon molecule plays a significant role in promoting the selective adsorption of higher molecular weight molecules over those with lower and this behavior is enthalpy driven.

The total adsorbed number of molecules in ZIF-8 with the increase of the temperature reduces monotonically in the pressure window of [1, 100] bar, as can be seen in **Figure 4.15**. For the representative 10 bar pressure, it reduces gradually attaining values 17.9, 15.5, 14.66 (ANN prediction), 13.1, 9.4, and 6.4 for 273, 293, 300, 313, 343, and 373 K, respectively. It should be noted that the ANN predicted isotherm (**Figure 4.15c**) at 300 K follows the qualitative and quantitative trends revealed from the GCMC simulations at 273–373 K.

## 4.6 Conclusions

In the present simulation work, the sorption isotherms and the self-diffusion coefficients for pure small alkanes and alkenes along with representative ternary mixtures sorbed in ZIF-8 have been studied using molecular simulations and ANNs computational methodologies. In particular, the NVT-MD simulations for pure methane, ethane, ethene, propane, propene, butane, butene, and methane–ethane–ethene equimolar gas mixture feed are in agreement with what is available in the literature MAS PFG NMR measurements following the experimental qualitative and quantitative trends at the temperature range of 273–373 K. The decrease of the self-diffusivity coefficients with the decrease of the temperature is the general behavior for the pure alkane and alkene hydrocarbon sorbates, followed by the vast majority of the methane–ethane–ethene mixture components. Experimental, simulation, and ANNs approaches proved that the sorbed pure hydrocarbons with lower molecular weights have higher diffusivities, which translates to  $D_s^{methane} > D_s^{ethene} > D_s^{ethane} > D_s^{propene} > D_s^{propane} > D_s^{butene} > D_s^{butane}$ . Interestingly in the studied ternary mixtures methane proved to be the fastest moving molecule in the cage-like ZIF-8 environment compared to ethene and to the sluggish ethane molecule.

The sorption isotherms from the GCMC simulations along with the ANNs predictions have been computed and presented for pure alkane and alkenes and for representative mixtures also at the 273 – 373 K temperature range and for pressures in the  $[10^{-2}, 10^2]$  bar window. For the ternary mixture ethane is preferentially sorbed over ethene in the cages of ZIF-8. Pressures approaching 100 bar are necessary for triggering methane’s sorption but still a very limited number (<3) of methane molecules are sorbed in the unit cell of ZIF-8.

Based on the present molecular simulation data and experimental measurements the adopted ANNs modeling procedure in particular aspects of the current simulation work showed its strong capability in predicting the self-diffusion coefficients and the sorption characteristics of pure hydrocarbon and hydrocarbon mixtures in ZIF-8. ANNs modeling predictive capabilities applied to molecular systems can serve as a versatile and reliable tool, in conjunction with molecular simulation methods and experimental measurements for fast computations using moderate computer resources.

## References

- [1] D. Frenkel and B. Smit, *Understanding molecular simulation : from algorithms to applications. 2nd ed.*, 1996. doi: 10.1063/1.881812.
- [2] M. Kotelyanskii and D. Theodorou, *Simulation Methods for Polymers*. 2004. doi: 10.1201/9780203021255.fmatt.
- [3] J. Kärger and D. M. Ruthven, “Diffusion in nanoporous materials: fundamental principles, insights and challenges,” *New J. Chem.*, vol. 40, no. 5, pp. 4027–4048, 2016, doi: 10.1039/C5NJ02836A.
- [4] C. Chmelik, “Characteristic features of molecular transport in MOF ZIF-8 as revealed by IR microimaging,” *Microporous and Mesoporous Materials*, vol. 216, pp. 138–145, 2015, doi: <https://doi.org/10.1016/j.micromeso.2015.05.008>.
- [5] C. Chmelik, R. Gläser, J. Haase, S. Hwang, and J. Kärger, “Application of microimaging to diffusion studies in nanoporous materials,” *Adsorption*, vol. 27, no. 5, pp. 819–840, 2021, doi: 10.1007/s10450-020-00279-4.
- [6] T. Zhou, Z. Song, and K. Sundmacher, “Big Data Creates New Opportunities for Materials Research: A Review on Methods and Applications of Machine Learning for Materials Design,” *Engineering*, vol. 5, no. 6, pp. 1017–1026, 2019, doi: <https://doi.org/10.1016/j.eng.2019.02.011>.
- [7] Z. Wang, Y. Zhou, T. Zhou, and K. Sundmacher, “Identification of optimal metal-organic frameworks by machine learning: Structure decomposition, feature integration, and predictive modeling,” *Comput Chem Eng*, vol. 160, p. 107739, 2022, doi: <https://doi.org/10.1016/j.compchemeng.2022.107739>.
- [8] D. J. Amit and D. J. Amit, *Modeling brain function: The world of attractor neural networks*. Cambridge university press, 1989. doi: 10.1017/CBO9780511623257
- [9] J. Vijayamohanan, A. Gupta, O. Noakoasteen, S. K. Goudos, and C. G. Christodoulou, “Source detection with multi-label classification,” *IEEE Open Journal of Signal Processing*, 2023. doi: 10.1109/OJSP.2023.3280854.
- [10] K. Suzuki, *Artificial neural networks: methodological advances and biomedical applications*. BoD–Books on Demand, 2011.

- [11] A. Phan, C. J. Doonan, F. J. Uribe-Romo, C. B. Knobler, M. O’Keeffe, and O. M. Yaghi, “Synthesis, Structure, and Carbon Dioxide Capture Properties of Zeolitic Imidazolate Frameworks,” *Acc Chem Res*, vol. 43, no. 1, pp. 58–67, 2010, doi: 10.1021/ar900116g.
- [12] D. Fairen-Jimenez *et al.*, “Flexibility and swing effect on the adsorption of energy-related gases on ZIF-8: combined experimental and simulation study,” *Dalton Trans.*, vol. 41, no. 35, pp. 10752–10762, 2012, doi: 10.1039/C2DT30774J.
- [13] L. Li, S. Zhang, J. P. Ruffley, and J. K. Johnson, “Energy Efficient Formaldehyde Synthesis by Direct Hydrogenation of Carbon Monoxide in Functionalized Metal–Organic Frameworks,” *ACS Sustain Chem Eng*, vol. 7, no. 2, pp. 2508–2515, 2019, doi: 10.1021/acssuschemeng.8b05413.
- [14] J.-R. Li, R. J. Kuppler, and H.-C. Zhou, “Selective gas adsorption and separation in metal–organic frameworks,” *Chem. Soc. Rev.*, vol. 38, no. 5, pp. 1477–1504, 2009, doi: 10.1039/B802426J.
- [15] H.-L. Jiang, B. Liu, T. Akita, M. Haruta, H. Sakurai, and Q. Xu, “Au@ZIF-8: CO Oxidation over Gold Nanoparticles Deposited to Metal–Organic Framework,” *J Am Chem Soc*, vol. 131, no. 32, pp. 11302–11303, 2009, doi: 10.1021/ja9047653.
- [16] K. Li *et al.*, “Zeolitic Imidazolate Frameworks for Kinetic Separation of Propane and Propene,” *J Am Chem Soc*, vol. 131, no. 30, pp. 10368–10369, 2009, doi: 10.1021/ja9039983.
- [17] P. D. Sutrisna, N. Prasetya, N. F. Himma, and I. G. Wenten, “A mini-review and recent outlooks on the synthesis and applications of zeolite imidazolate framework-8 (ZIF-8) membranes on polymeric substrate,” *Journal of Chemical Technology & Biotechnology*, vol. 95, no. 11, pp. 2767–2774, 2020, doi: <https://doi.org/10.1002/jctb.6433>.
- [18] Q. Yang, C.-C. Yang, C.-H. Lin, and H.-L. Jiang, “Metal–Organic-Framework-Derived Hollow N-Doped Porous Carbon with Ultrahigh Concentrations of Single Zn Atoms for Efficient Carbon Dioxide Conversion,” *Angewandte Chemie International Edition*, vol. 58, no. 11, pp. 3511–3515, 2019, doi: <https://doi.org/10.1002/anie.201813494>.
- [19] F. Yang, M. Wu, Y. Wang, S. Ashtiani, and H. Jiang, “A GO-Induced Assembly Strategy To Repair MOF Nanosheet-Based Membrane for Efficient H<sub>2</sub>/CO<sub>2</sub> Separation,” *ACS Appl Mater Interfaces*, vol. 11, no. 1, pp. 990–997, 2019, doi: 10.1021/acsami.8b19480.

- [20] A. A. Pribylov, K. O. Murdmaa, O. V Solovtsova, and M. K. Knyazeva, "Methane adsorption on various metal-organic frameworks and determination of the average adsorption heats at supercritical temperatures and pressures," *Russian Chemical Bulletin*, vol. 67, no. 10, pp. 1807–1813, 2018, doi: 10.1007/s11172-018-2293-2.
- [21] X. Gong, Y. Wang, and T. Kuang, "ZIF-8-Based Membranes for Carbon Dioxide Capture and Separation," *ACS Sustain Chem Eng*, vol. 5, no. 12, pp. 11204–11214, 2017, doi: 10.1021/acssuschemeng.7b03613.
- [22] J. Cheng *et al.*, "Self-assembly of 2D-metal-organic framework/graphene oxide membranes as highly efficient adsorbents for the removal of Cs<sup>+</sup> from aqueous solutions," *RSC Adv.*, vol. 8, no. 71, pp. 40813–40822, 2018, doi: 10.1039/C8RA08410F.
- [23] S. Nandi, E. Sharma, V. Trivedi, and S. Biswas, "Metal-Organic Framework Showing Selective and Sensitive Detection of Exogenous and Endogenous Formaldehyde," *Inorg Chem*, vol. 57, no. 24, pp. 15149–15157, 2018, doi: 10.1021/acs.inorgchem.8b02411.
- [24] S. Wang *et al.*, "Framework disorder and its effect on selective hysteretic sorption of a T-shapedazole-based metal-organic framework," *IUCrJ*, vol. 6, no. 1, pp. 85–95, Jan. 2019, doi: 10.1107/S2052252518015749.
- [25] K. S. Park *et al.*, "Exceptional chemical and thermal stability of zeolitic imidazolate frameworks," *Proceedings of the National Academy of Sciences*, vol. 103, no. 27, pp. 10186–10191, 2006, doi: 10.1073/pnas.0602439103.
- [26] Ch. Baerlocher and L. B. McCusker, "Database of Zeolite Structures" <https://www.iza-structure.org/>
- [27] M. G. Martin and J. I. Siepmann, "Transferable Potentials for Phase Equilibria. 1. United-Atom Description of n-Alkanes," *J Phys Chem B*, vol. 102, no. 14, pp. 2569–2577, 1998, doi: 10.1021/jp972543+.
- [28] C. D. Wick, M. G. Martin, and J. I. Siepmann, "Transferable Potentials for Phase Equilibria. 4. United-Atom Description of Linear and Branched Alkenes and Alkylbenzenes," *J Phys Chem B*, vol. 104, no. 33, pp. 8008–8016, 2000, doi: 10.1021/jp001044x.
- [29] K. Momma and F. Izumi, "VESTA: a three-dimensional visualization system for electronic and structural analysis," *J Appl Crystallogr*, vol. 41, no. 3, pp. 653–658, 2008, doi: <https://doi.org/10.1107/S0021889808012016>.

- [30] D. E. E. David Dubbeldam Sofia Calero and R. Q. Snurr, “RASPA: molecular simulation software for adsorption and diffusion in flexible nanoporous materials,” *Mol Simul*, vol. 42, no. 2, pp. 81–101, 2016, doi: 10.1080/08927022.2015.1010082.
- [31] A. K. Rappe, C. J. Casewit, K. S. Colwell, W. A. I. I. I. Goddard, and W. M. Skiff, “UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations,” *J Am Chem Soc*, vol. 114, no. 25, pp. 10024–10035, 1992, doi: 10.1021/ja00051a040.
- [32] S. L. Mayo, B. D. Olafson, and W. A. Goddard, “DREIDING: a generic force field for molecular simulations,” *J Phys Chem*, vol. 94, no. 26, pp. 8897–8909, 1990, doi: 10.1021/j100389a010.
- [33] L. N. Gergidis and D. N. Theodorou, “Molecular Dynamics Simulation of n-Butane–Methane Mixtures in Silicalite,” *J Phys Chem B*, vol. 103, no. 17, pp. 3380–3390, 1999, doi: 10.1021/jp983680p.
- [34] P. Krokidas, M. Castier, and I. G. Economou, “Computational Study of ZIF-8 and ZIF-67 Performance for Separation of Gas Mixtures,” *The Journal of Physical Chemistry C*, vol. 121, no. 33, pp. 17999–18011, 2017, doi: 10.1021/acs.jpcc.7b05700.
- [35] A. Gurras and L. N. Gergidis, “Modeling Sorption and Diffusion of Alkanes, Alkenes, and their Mixtures in Silicalite: From MD and GCMC Molecular Simulations to Artificial Neural Networks,” *Adv Theory Simul*, vol. 4, no. 3, p. 2000210, 2021, doi: <https://doi.org/10.1002/adts.202000210>.
- [36] J. F. Mas and J. J. Flores, “The application of artificial neural networks to the analysis of remotely sensed data,” *Int J Remote Sens*, vol. 29, no. 3, pp. 617–663, 2008, doi: 10.1080/01431160701352154.
- [37] L. L. Y. Yanhui and E. K. George, “Dying ReLU and Initialization: Theory and Numerical Examples,” *Commun Comput Phys*, vol. 28, no. 5, pp. 1671–1706, 2020, doi: <https://doi.org/10.4208/cicp.OA-2020-0165>.
- [38] J.-R. Zhang, J. Zhang, T.-M. Lok, and M. R. Lyu, “A hybrid particle swarm optimization–back-propagation algorithm for feedforward neural network training,” *Appl Math Comput*, vol. 185, no. 2, pp. 1026–1037, 2007, doi: <https://doi.org/10.1016/j.amc.2006.07.025>.

- [39] A. Van Ooyen and B. Nienhuis, "Improving the convergence of the back-propagation algorithm," *Neural Networks*, vol. 5, no. 3, pp. 465–471, 1992, doi: [https://doi.org/10.1016/0893-6080\(92\)90008-7](https://doi.org/10.1016/0893-6080(92)90008-7).
- [40] D. Burden Frank and Winkler, "Bayesian Regularization of Neural Networks," in *Artificial Neural Networks: Methods and Applications*, D. J. Livingstone, Ed., Totowa, NJ: Humana Press, 2009, pp. 23–42. doi: 10.1007/978-1-60327-101-1\_3.
- [41] S. Sinha, T. N. Singh, V. K. Singh, and A. K. Verma, "Epoch determination for neural network by self-organized map (SOM)," *Comput Geosci*, vol. 14, no. 1, pp. 199–206, 2010, doi: 10.1007/s10596-009-9143-0.
- [42] MATLAB, 9.7.0.1190202 (R2017b). Natick, Massachusetts: The MathWorks Inc., 2017.
- [43] N. Dvoyashkina, D. Freude, S. S. Arzumanov, and A. G. Stepanov, "Monitoring the Diffusivity of Light Hydrocarbons in a Mixture by Magic Angle Spinning Pulsed Field Gradient NMR: Methane/Ethane/Ethene in ZIF-8," *The Journal of Physical Chemistry C*, vol. 121, no. 45, pp. 25372–25376, 2017, doi: 10.1021/acs.jpcc.7b09335.
- [44] L. N. Gergidis, D. N. Theodorou, and H. Jobic, "Dynamics of n-Butane–Methane Mixtures in Silicalite, Using Quasielastic Neutron Scattering and Molecular Dynamics Simulations," *J Phys Chem B*, vol. 104, no. 23, pp. 5541–5552, Jun. 2000, doi: 10.1021/jp0000073.
- [45] D. Dubbeldam and R. Q. Snurr, "Recent developments in the molecular modeling of diffusion in nanoporous materials," *Mol Simul*, vol. 33, no. 4–5, pp. 305–325, 2007, doi: 10.1080/08927020601156418.
- [46] H. Jobic and D. N. Theodorou, "Quasi-elastic neutron scattering and molecular dynamics simulation as complementary techniques for studying diffusion in zeolites," *Microporous and Mesoporous Materials*, vol. 102, no. 1, pp. 21–50, 2007, doi: <https://doi.org/10.1016/j.micromeso.2006.12.034>.
- [47] L. Hertäg *et al.*, "Diffusion of CH<sub>4</sub> and H<sub>2</sub> in ZIF-8," *J Memb Sci*, vol. 377, no. 1, pp. 36–41, 2011, doi: <https://doi.org/10.1016/j.memsci.2011.01.019>.
- [48] C. Chmelik, D. Freude, H. Bux, and J. Haase, "Ethene/ethane mixture diffusion in the MOF sieve ZIF-8 studied by MAS PFG NMR diffusometry," *Microporous and Mesoporous Materials*, vol. 147, no. 1, pp. 135–141, 2012, doi: <https://doi.org/10.1016/j.micromeso.2011.06.009>.

- [49] C. Chmelik and J. Kärger, “The predictive power of classical transition state theory revealed in diffusion studies with MOF ZIF-8,” *Microporous and Mesoporous Materials*, vol. 225, pp. 128–132, 2016, doi: <https://doi.org/10.1016/j.micromeso.2015.11.051>.
- [50] J.-P. Hansen and I. R. McDonald, *Theory of simple liquids: with applications to soft matter*. Academic press, 2013.
- [51] L. Zhang, Q. Li, Z. Lu, and X. Wang, “Understanding Adsorption and Separation behavior of shorter chain alkane mixtures in Zeolitic Imidazolate Frameworks by molecular simulations,” *Procedia Comput Sci*, vol. 4, pp. 1193–1202, 2011, doi: <https://doi.org/10.1016/j.procs.2011.04.128>.



# **Chapter 5: Adsorption of Methane, Ethane and their Equimolar Mixture in NIIC-20-Bu Metal Organic Framework from Grand Canonical Monte Carlo Simulations and Artificial Neural Networks**

## **5.1 Preface**

In this chapter the adsorption of methane, ethane and their equimolar mixture in the newly synthesized mesoporous NIIC-20-Bu Metal-Organic Framework (MOF) were investigated utilizing molecular simulations and Artificial Neural Networks. To the best of our knowledge, this is the first computational study of small alkanes sorbed in this particular novel MOF. Grand Canonical Monte Carlo simulations provided the adsorption isotherms of the alkanes in NIIC-20-Bu at different temperatures. The simulation findings were compared with existing experimental sorption measurements showing reasonable quantitative and qualitative agreement. Predictive models based on Artificial Neural Networks are developed incorporating simulation data and available experimental measurements in the training phase to predict the sorption isotherms of methane, ethane and their equimolar mixture in NIIC-20-Bu mesoporous material with the minimum computational cost. The present findings highlight the potential applications of NIIC-20-Bu as efficient adsorbent material in many applications.

As part of this research, the NIIC-20-Bu structure was thoroughly examined due to its unique properties and potential applications in gas separation. Among the five different networks studied, NIIC-20-Bu stands out because it possesses the second smallest twelve-nuclear window opening. This structural characteristic arises from the folding of the 1,2-butanediol groups, which partially restricts the pore openings. Despite this limitation, previous studies have highlighted the promising performance of NIIC-20-Bu in the selective separation of alkane/alkene gas mixtures, making it an attractive candidate for industrial applications.

Notably, NIIC-20-Bu exhibited the highest ethane/ethylene adsorption selectivity among the five mesoporous structures analyzed. This suggests that the material has a strong preference for ethane molecules over ethylene, which is a crucial factor in refining and petrochemical processes where the efficient separation of these gases is required. The superior selectivity observed in NIIC-20-Bu is attributed to a combination of its well-balanced window geometry and its moderate hydrophobicity. These structural features contribute to the selective interactions between the framework and the adsorbed gas molecules, allowing for enhanced

separation between ethane and ethylene. The ideal balance of the pore architecture, along with the hydrophobic nature of NIIC-20-Bu, plays a key role in facilitating intermolecular interactions between alkanes and alkenes. Specifically, the material exhibits a tendency to exclude ethylene while preferentially adsorbing ethane, making it a promising candidate for applications in gas purification and separation technologies. This property is particularly valuable in industries where high-purity ethylene is required for polymer production and other chemical processes.

Overall, the unique combination of pore geometry, hydrophobicity, and selective adsorption properties makes NIIC-20-Bu a highly suitable material for advanced gas separation applications. Future studies could further explore its adsorption behavior under varying conditions, such as different temperatures and pressures, to optimize its performance and broaden its potential applications in the field of gas separation and purification.

The results presented in this chapter are included in the article that resulted from the research titled “Adsorption of Methane, Ethane and Their Equimolar Mixture in NIIC-20-Bu Metal-Organic Framework from Grand Canonical Monte Carlo Simulations and Artificial Neural Networks,” published in *Advanced Theory and Simulations* (2025), e00695, doi: 10.1002/adts.202500695

## 5.2 Introduction

Metal-organic frameworks (MOFs) have gained significant attention and have emerged as an important class of microporous solids over the past two decades. They are extensively investigated due to their remarkable structural and functional diversity and physicochemical properties [1], [2], [3]. The nearly infinite combinations of metal nodes and organic linkers have led to the synthesis of over 100,000 experimental MOFs. As a result of their structural and functional diversity, metal-organic frameworks hold great potential for various applications [4]. They have shown significant potential as gas or general adsorbents [5], [6], [7], [8], luminescent sensors [9], and materials with catalytic, optical, and magnetic properties [10], [11], [12]. This versatility has presented opportunities for their use across a broad spectrum of applications.

In addition, the demand for efficient and selective gas storage and separation technologies has escalated due to increasing global energy consumption and environmental concerns. In this context, the separation and storage of light hydrocarbons, particularly methane and ethane, are of paramount importance. Methane is the primary component of natural gas, a

crucial fuel source, and also a potent greenhouse gas, making its purification and storage a key challenge for clean energy and climate change mitigation [13], [14]. The separation of ethane from methane and other hydrocarbons is essential for the petrochemical industry, as ethane serves as a feedstock for producing plastics and other chemicals. Traditional separation methods like cryogenic distillation are energy-intensive and costly, creating a strong motivation for developing alternative materials for hydrocarbon gas separation. MOFs have emerged as highly promising adsorbents for these applications [15]. Their unique properties, including high surface areas, tunable pore sizes, and modifiable surface chemistry, allow for precise control over host-guest interactions [16]. This has led to a surge in research the last decade, with numerous papers exploring the selective sorption behavior of methane, ethane, and their mixtures in various MOFs. Both experimental and computational approaches, such as Grand Canonical Monte Carlo (GCMC), Molecular Dynamics (MD), Density Functional Theory (DFT) simulations and Machine Learning (ML) techniques, have been instrumental in unraveling the underlying mechanisms of hydrocarbons storage and selective sorption [17], [18]. High-throughput computational campaigns (including MOFs screening studies) have been particularly successful in identifying promising MOFs for hydrocarbons storage and separation from vast databases of hypothetical and experimental structures, guiding the rational design of next-generation materials [19], [20], [21], [22], [23]. These simulation studies often focus on the subtle differences in interactions between the MOF pore walls and the hydrocarbon molecules, revealing how factors like pore size, shape, and functional groups can be fine-tuned to achieve high selectivity, even in complex mixture feeds. In particular, over the past several years, ML has substantially transformed the way of MOF discovery, design and synthesis, improving their physicochemical performance on any given application. Driven by the abundant data from experiments or brute-force simulations, ML techniques not only efficiently and accurately predict MOF properties but also quantitatively derive structure–property relationships [24]. In recent studies [25], [26] the synthesis of two isostructural series of modular metal-organic frameworks (MOFs) named NIIC-10 (microporous) and NIIC-20 (mesoporous) has been reported. NIIC stands for Nikolaev Institute of Inorganic Chemistry, where the aforementioned synthesis took place. These newly synthesized MOFs are based on dodecanuclear wheel-shaped carboxylate building blocks in  $Zn_{12}(RCOO)_{12}(G)_6$ , where RCOO represents either 2,5-thiophenedicarboxylate (tdc) in NIIC-10 or isophthalate (iph) in NIIC-20. The glycolate moiety (G) decorates the inner portal of the wheel-shaped structure. In particular, the NIIC-20 series exhibit an open mesoporous framework and demonstrate high adsorption capacity for small alkanes [27], [28]. Additionally, they exhibit exceptional

selectivity in adsorbing ethane over ethylene [26]. This unique preference for the adsorption of saturated hydrocarbons over unsaturated ones is quite unusual and can be attributed to the specific arrangement of the glycol-decorated portals. These portals facilitate numerous van der Waals interactions between saturated substrates and the surface of the NIIC-20-G material.

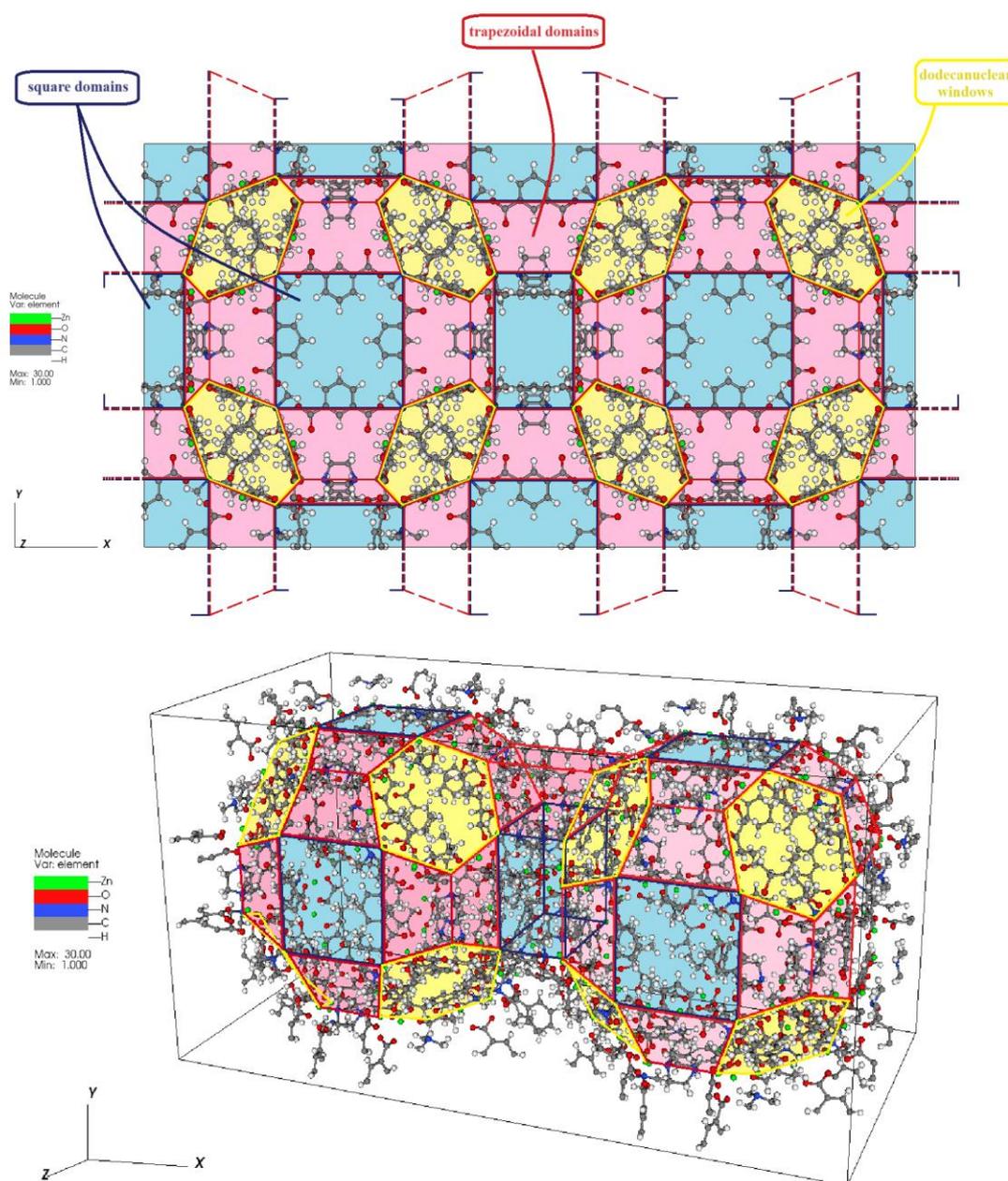
The objective of the present work is the simulation study of adsorption of methane, ethane and their equimolar mixtures in NIIC-20-Bu at various temperature and pressure conditions aiming to the potential applications of the NIIC-20 sorbent structures for efficient and enhanced small molecular weight hydrocarbon storage. Grand Canonical Monte Carlo (GCMC) simulations and Artificial Neural Networks (ANNs) are adopted for the description and calculation of qualitative and quantitative sorption characteristics of light hydrocarbons in NIIC-20-Bu. By combining molecular-based simulation results and experimental measurements, ANNs techniques can be employed to develop predictive models that accurately describe structure-property relationships. This approach has the potential to provide reliable estimations and predictions of adsorption and diffusion properties of hydrocarbon sorbates in porous structures, such as MOFs and Zeolites [29], [30]. Importantly, ANNs techniques offer the advantage of significantly reducing computational costs compared to traditional molecular simulation approaches and bridge the gap between simulation and experiments, making them more efficient for large-scale predictions [31], [32].

## 5.3 Molecular Modeling and Simulations

### 5.3.1 Molecular Model

The NIIC-20 MOFs family consists of five mesoporous frameworks denoted as NIIC-20-G with the general chemical formula  $Zn_{12}(i\text{-bdc})_6(G)_6(\text{dabco})_3$ , where *i*-bdc stands for isophthalate, G stands for one of the following: ethylenediolate (Et), 1,2-propylenediolate (Pr), 1,2-buthylenediolate (Bu), 1,2-pentylenediolate (Pe), glycerolate (Gl) and dabco is the 1,4-diazabicyclo[2.2.2]octane which is also part of the framework structure [25], [26]. The NIIC-20 MOFs feature a unique and complex nbo-h topology, consisting of dodecanuclear windows  $Zn_{12}(\text{RCOO})_{12}(G)_6$  building units with apertures of 2 Å in the case of NIIC-20-Bu), large nanopores-nanocages (25 Å) and intersecting channels with tetragonal (3.5 Å × 3.5 Å) and trapezoidal openings (6 Å × 3.5 Å). The primary components of all these domains are dodecanuclear building units connected to each other through isophthalate ligands (iph<sup>2-</sup>anions) and dabco molecules.

The unit cell of the NIIC-20-Bu framework was constructed from single crystal X-ray diffraction (SCXRD) study [25], provided by the research group at Nikolaev Institute of Inorganic Chemistry. In the present study, a simulation box comprising of eight unit cells was utilized, achieved through a  $2 \times 2 \times 2$  repetition in  $x$ ,  $y$ , and  $z$ -directions of the unit cell, respectively. A  $2 \times 1 \times 1$  replication of the unit cell is depicted in **Figure 5.1** in two and three dimensions highlighting the different regions, characteristic domains and chemical entities.



**Figure 5.1:** A  $2 \times 1 \times 1$  replication of the unit cell of NIIC-20-Bu. Different colors are used for representing different atom types: zinc-green, carbon-gray, oxygen-red, hydrogen-white, nitrogen-blue.

The methane molecule along with the methyl groups (-CH<sub>3</sub>) of ethane molecule were modeled as individual united atom dispersive interaction centers. The interactions with the framework atoms of NIIC-20 were simulated using the GenericMOFs force field (a hybrid force field based on the Universal Force Field (UFF)[33] and the Dreiding force field [34]), which was implemented in RASPA simulation software [35]. The cutoff distance for the dispersive interactions was set at  $r_c = 12 \text{ \AA}$ , which means that interactions beyond this distance are not taken into account in the course of the simulations. Methane, ethane and their equimolar gas mixture sorption in NIIC-20-Bu was qualitatively and quantitatively studied using the molecular simulation framework, specifically through Grand Canonical Monte Carlo (GCMC) numerical computations.

In the course of the GCMC simulations three types of Monte Carlo (MC) moves were used to explore the phase space of the sorbent-sorbate physical system. These moves include the insertion, deletion, and translation of sorbate molecules, and they are typically applied to deal with the positions and identities of molecules in the mesoporous NIIC-20-Bu framework. In this study each trial move is selected and attempted with equal probability. In the context of GCMC calculations and the simulation of sorption isotherms, it is common and reasonable to employ a rigid framework approximation, particularly for molecules that do not fit tightly within or significantly distort the sorbent's framework. Consequently, the production runs of GCMC simulations were conducted using a rigid NIIC-20-Bu framework.

This GCMC simulation procedure involved a series of distinct stages: minimization, relaxation, equilibration, and production. Each stage serves a specific purpose in preparing the system and then collecting statistically robust data. Energy minimization of the initial structure took place, ensuring that the simulation starts from a stable and realistic configuration, followed by a short relaxation period (50,000 GCMC steps) as a pre-processing/initialization step. After initialization, the equilibration simulation stage took place for thousands of MC steps, ranging from  $10^5$  to  $5 \times 10^5$  depending on the pressure/sorbate loading value, allowing the system to reach equilibrium. For the actual production phase, the number of GCMC steps was chosen to ensure comprehensive sampling of the system's configurational space, particularly for complex mesoporous MOF structures with diverse pore environments. A preliminary analysis was conducted to determine the optimal number of steps. For NIIC-20-Bu and its hydrocarbon systems (methane, ethane, and their mixture) at low to intermediate sorbate loadings, a smaller number of steps ( $5 \times 10^5$  to  $10^6$ ) was sufficient. However, for higher pressures and sorbed amounts, a larger number of steps ( $5 \times 10^6$ ) was necessary to ensure adequate sampling and

thermodynamic equilibrium (for a typical well equilibrated run of  $5 \times 10^6$  steps the relative difference for the methane sorbate loading value (at 1 atm, 273 K) with respect to the sorbate loading value of a long run with  $10^7$  steps was below 0.007%). This approach balances computational efficiency with the need for reliable data.

The pressure and temperature values were selected to cover a wide range of conditions relevant to typical gas storage and potential separation applications following previous experimental work [27], [28]. The specific pressures were chosen to accurately map out the full sorption isotherm, from the initial Henry's law region to the saturation plateau. This allows for a thorough understanding of the sorption mechanism and the accurate determination of key parameters like saturation capacity. The temperatures were selected to investigate the effect of thermal energy on sorption, which is crucial for evaluating the performance of the NIIC-20-Bu mesoporous MOF under different operational conditions.

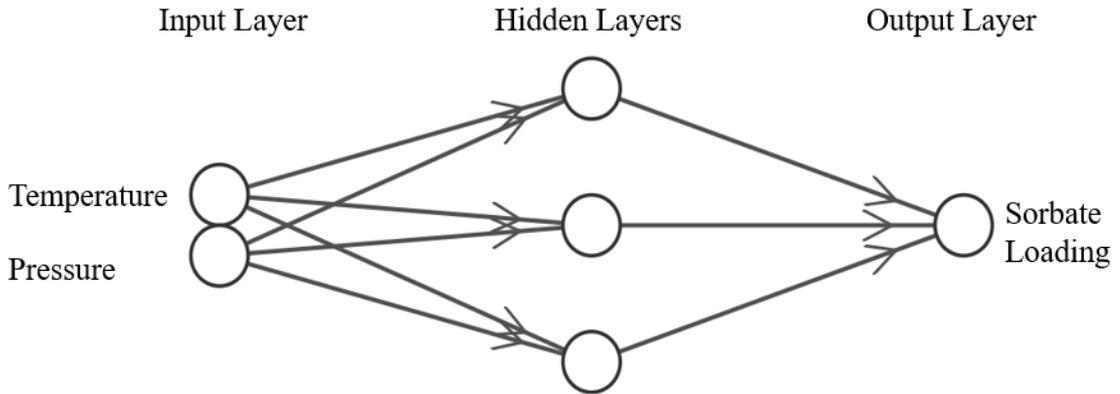
The individual production runs required computational times ranging from 5 to 40 days, depending on the applied pressure and consequently the sorbate loading. These computations were performed on a modest computer system equipped with an AMD Ryzen 5950X processor and 32 GB of memory. The computer operated under the Ubuntu 16.04 LTS (Xenial Xerus) operating system. The molecular simulations in this study were conducted using the RASPA2 molecular simulation software [35].

### **5.3.2 ANNs Methodology**

ANNs are a subset of ML techniques and form the foundation of Deep Learning (DL) algorithms. ANNs consist of interconnected nodes or "neurons" organized in layers, with each neuron performing a specific computation. Through the training process, ANNs learn to recognize patterns, make predictions, and solve complex problems by adjusting the weights and biases of the connections between neurons. Deep Learning, powered by ANNs, has been successful in various fields such as computer vision, natural language processing, and speech recognition [36], [37], [38], owing to its ability to automatically extract and learn hierarchical representations from large data sets. While ANNs have taken the world by storm, their usage in combination with molecular simulations to provide reliable predictions and insights regarding the respective processes have started to flourish only recently [29], [39], [40].

A similar approach to this ANNs methodology framework was used in our previous endeavors [29], [30], where the sorption and diffusion of small alkanes in Zeolite Silicalite and MOF ZIF-8 were studied. The predictive modeling computational approach for sorption uses

the temperature ( $T$ ) and pressure ( $p$ ) in the input layer. Consequently, the input vector  $\chi_j^{Sorption} = (x_1 = T, x_2 = p)_j$  is created. Index  $j$  represents a particular sample from the total number  $N$  of samples which make up the input state space  $X$  that was created from data sets originating from the present GCMC computations. To build the dataset for the sorption ANN all available simulation outputs and experimental measurements were concatenated into one pooled dataset. No additional weighting or scaling factors were applied to favor experimental over simulated points (or vice versa) and each sample, regardless of origin, contributes equally to the learning procedure of the ANN. The fully normalized dataset ( $N$  total samples) was randomly permuted with a fixed seed to guarantee reproducibility. Afterwards, the allocation of data samples took place with: 70% of samples used for training, 15% for validation and 15 % held out for the final test set. In **Figure 5.2** a representative ANNs architecture specifically for the sorption predictions developed in this study is shown.

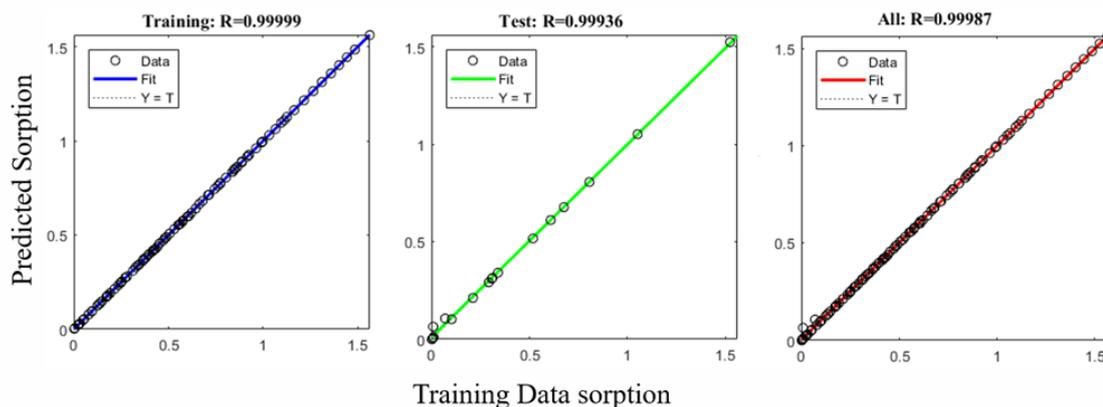


**Figure 5.2:** Representative architecture of the developed ANN for the sorption predictions.

The neural network employed in this study is a feed-forward architecture consisting of three hidden layers, each comprising three neurons per layer. The selection process for the ANN architecture adopted a trial-and-error approach, which was found to be optimal for the sorption predictions. This approach aligns with procedures demonstrated by separate studies [41], and it is worth noting that a similar network architecture has been employed in previous research efforts. The hidden layers utilize the logistic (sigmoid-type) activation function, which is extensively used in predictive modeling with ANNs due to its capacity to capture complex, non-linear relationships observed in sorption behavior. The output layer consists of a one-dimensional output vector corresponding to the sorbate loading, denoted as  $c_{sorbate}$ . During training, the back-propagation algorithm was used as the learning method, enhanced with a Bayesian regularization scheme. This probabilistic regularization mechanism balances error minimization and weight magnitude, thereby reducing overfitting without requiring manual

tuning of regularization strength. It also contributes to improved robustness and predictive accuracy by mitigating the impact of noise captured during the modeling procedure [42], [43]. The training process began with random initialization of the weight set ( $\mathbf{w}$ ), which was iteratively adjusted at the end of each epoch. The number of epochs used in this study was 1000, determined through iterative experimentation using values from 1000-1200 for optimal performance. Although this range was found to yield satisfactory performance, it is important to recognize that increasing the number of epochs does not universally guarantee better results and may increase the risk of overfitting, as shown in previous studies [44], [45]. To further manage this risk, no explicit early stopping criteria were enforced; however, training was automatically halted if the validation error failed to improve over six consecutive training cycles. Model performance was evaluated using the Mean Squared Error (MSE) loss function, which quantifies the average squared difference between predicted and actual values. In **Appendix D** file a typical MSE vs. epochs diagram is given. The MSE metric provides a robust measure of predictive accuracy and is widely utilized in ANNs-based modeling frameworks.

For the predictive modeling procedure, an in-house code was developed utilizing the Neural Net Fit library in MATLAB [46] and the model (with the associated model parameters) is given in **Appendix D**. This code implementation allowed for the customization and adaptation of the neural network model to suit the specific requirements and objectives of the study. In terms of computer time requirements, the training phase of the ANN took approximately 8 to 10 seconds to execute. On the other hand, the prediction phase involved significantly shorter execution times, typically less than 1 second. This highlights the efficiency of the developed models in generating predictions once the training phase is completed. The correlation coefficient for the training, testing and for the combined training-testing is shown in **Figure 5.3**. It is crucial to note that the developed ANNs utilize distinct training data sets for sorption. These data sets consist of information obtained from the present molecular simulation work augmented by the experimental measurements acquired from [27], [28]. This combination gives an excellent route for a more robust and accurate predictive modeling approach.



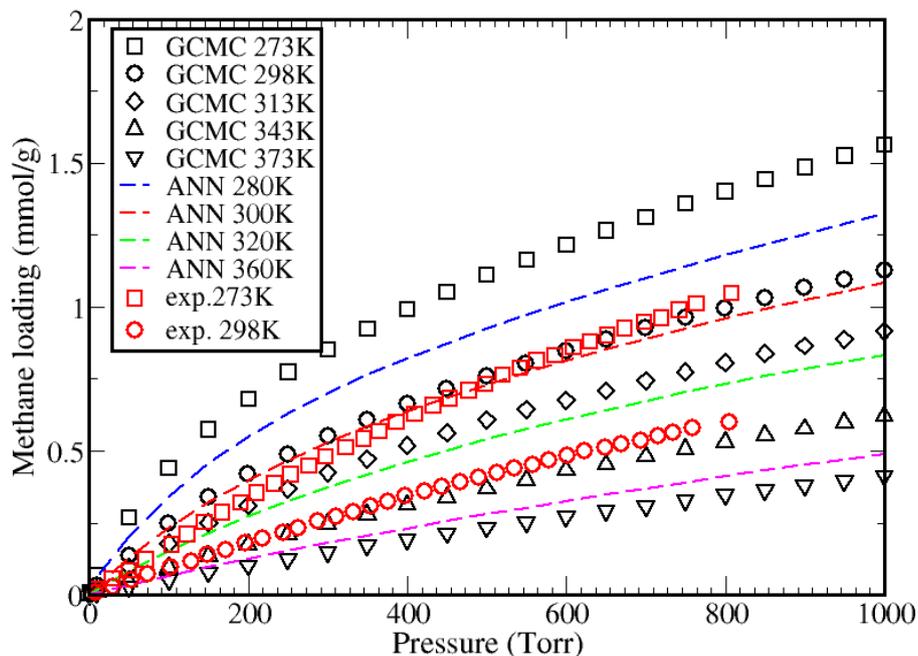
**Figure 5.3:** Correlation coefficient for the training, testing and for the combined training-testing phases. The x-axis represents data used in the sorption training and y-axis the predicted ANN sorption data.

## 5.4 Results

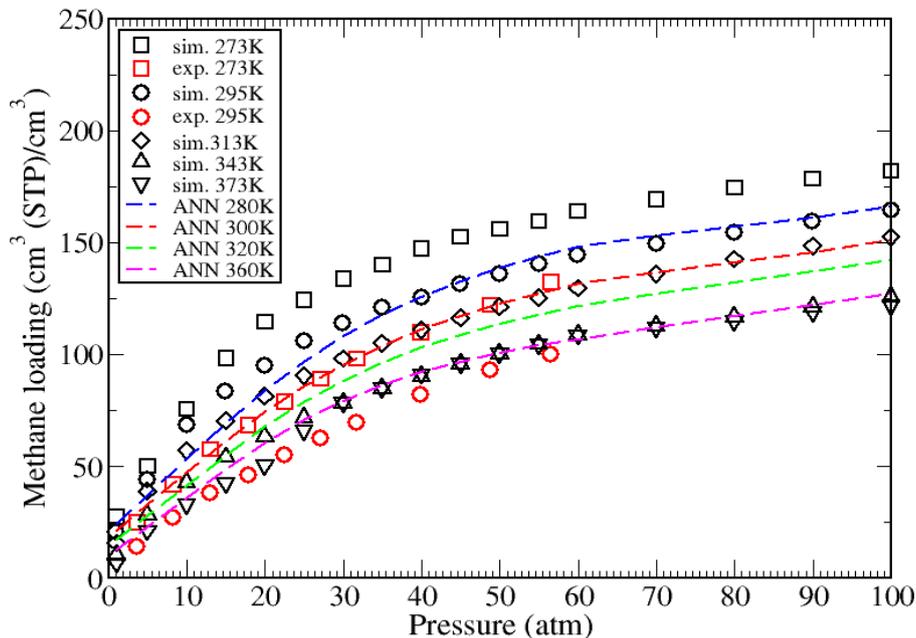
### 5.4.1 Methane Sorption

The sorption isotherms have been computed using GCMC simulations for different temperatures varied from  $T=273$  to 373 K and are depicted in **Figure 5.4** for pressures up to 1000 Torr and in **Figure 5.5** for pressure values up to 100 atm. The experimental sorption measurements at 273 and 298 K for pressures up to 800 Torr acquired from [28] are shown in Figure 5.4. The experimental measurements at 273 K and 295 K for pressures beyond 800 Torr and up to 100 atm acquired from [27] are presented in **Figure 5.5**. It is important to emphasize that the presentation of the results in two different pressure ranges follows the data presentation figure format adopted in the aforementioned experimental work.

The applied pressure and temperature play a crucial role in the actual sorption process. The pressure increase leads to an increase in methane's sorbate loading. The sorbed amount can be also controlled with the proper adjustment of the system's temperature. The experimental measurements acquired from [27], [28], as well as the GCMC simulation data and ANNs predictions share the aforementioned qualitative characteristics as can be seen in **Figures. 5.4, 5.5**. The higher the temperature the lower the amount sorbed. Adsorption of gas increases order and therefore is unfavorable. At higher temperatures, the entropic effect predominates and adsorption of gas decreases. At lower temperatures, the reverse is true. Upon initial review, the GCMC simulation results demonstrate an overestimation of methane sorption relative to the experimental measurements conducted at 273 and 295 K. This can be attributed to the force-field adopted and the rigid framework used in the present study.



**Figure 5.4:** GCMC simulation results for sorption isotherms of methane in NIIC-20-Bu at the low-pressure range for five different temperatures (273, 298, 313, 343, 373K) along with experimental measurements from [28]. In continuous line format the prediction from the ANN at four different temperatures are presented (280, 300, 320, 360 K) are shown.



**Figure 5.5:** GCMC simulation results for methane sorption isotherm at the high-pressure range for five different temperatures (273, 295, 313, 343, 373K) in NIIC-20-Bu along with experimental measurements from [27]. Also, in dashed line format the predictions from the ANN at four different temperatures (280, 300, 320, 360 K) are presented.

It should be noted that generic force fields, while broadly applicable, often fail to capture the nuanced electronic environments prevalent within MOFs. This includes the specific hybridization states of carbon atoms and the significant ion polarization effects that occur around hard-Lewis acid metal centers. Such inaccuracies lead to an artificial inflation of interaction strengths between adsorbate molecules (methane and ethane) and the MOF/porous framework, resulting in higher predicted adsorption capacities than experimentally observed [47]. Furthermore, the assumption of a rigid framework neglects the dynamic and responsive nature of many MOFs. These materials can undergo significant structural changes, such as pore contraction, hindered access due to mobile pendant groups, or non-monotonic deformation, in response to gas adsorption. This omission in simulations leads to an overprediction of accessible pore volume and available adsorption sites [48].

In particular, for the low pressure range (pressures up to 1000 Torr) the GCMC simulation results and ANN predictions (based on a training set that included data from the GCMC simulations and experimental measurements from the 3.7 to 56.6 atm at all temperatures studied) share the same qualitative behaviour and exhibit reasonable quantitative agreement. The increase of temperature, results in a non-linear gradual increase of the methane amount sorbed in the NIIC-20-Bu framework. The ANN predictions at the intermediate temperatures ( $T=280, 300, 320, 360$  K) are located between the GCMC isotherms at the respective upper and lower bounding temperatures, as anticipated. In particular, the  $T=300$  K ANN prediction curve lies between the 298 and 313 K GCMC simulation points in close proximity to the simulation points at  $T = 298$  K as can be observed in **Figure 5.4**. The ANN prediction at 360 K closely follows the more linear shaped isotherm data at 343 and 373 K, lying in between.

The isotherms for pressures greater than 1000 Torr=1.3158 atm are depicted in **Figure 5.5**. The isotherms display similar qualitative and quantitative characteristics with the behavior at the low pressure regime. The gradual increase of the sorbed amount with pressure increase is evident while the temperature decrease guarantees higher sorbate loadings. For high values of pressure (beyond 30 atm) the amounts of methane sorbed equalize for the two distinct and significantly different temperatures of 343 and 373 K. While feasible for our *in silico* 'experiments,' simulations exceeding 60 atm might be of minor importance under actual experimental conditions. Indicatively, for methane at 56.6 atm the experimentally measured uptake is  $132.2 \text{ cm}^3(\text{STP})/\text{cm}^3$  while the GCMC value at 55 atm is  $159.5 \text{ cm}^3(\text{STP})/\text{cm}^3$  both at 273 K. Similarly, at 295 K the amounts measured experimentally and acquired from the

molecular simulation are  $100.0 \text{ cm}^3(\text{STP})/\text{cm}^3$  (at 56.6 atm) and  $140.1 \text{ cm}^3(\text{STP})/\text{cm}^3$  (at 55 atm), respectively. For the higher pressure value of 100 atm used in our simulations the methane amounts are 181.8 (273 K), 164.3 (295 K), 152.2 (313 K), 125.5 (343), 121.6 (373 K)  $\text{cm}^3(\text{STP})/\text{cm}^3$ .

The sorption ANN model takes into account the experimental measurements [27], [28] during the ANN's phases (training-testing) and the actual predictions are depicted in **Figure 5.5** and for indicative values in **Table 5.1**. The ANN predictions are in qualitative and quantitative agreement at the whole range of pressures and temperatures studied.

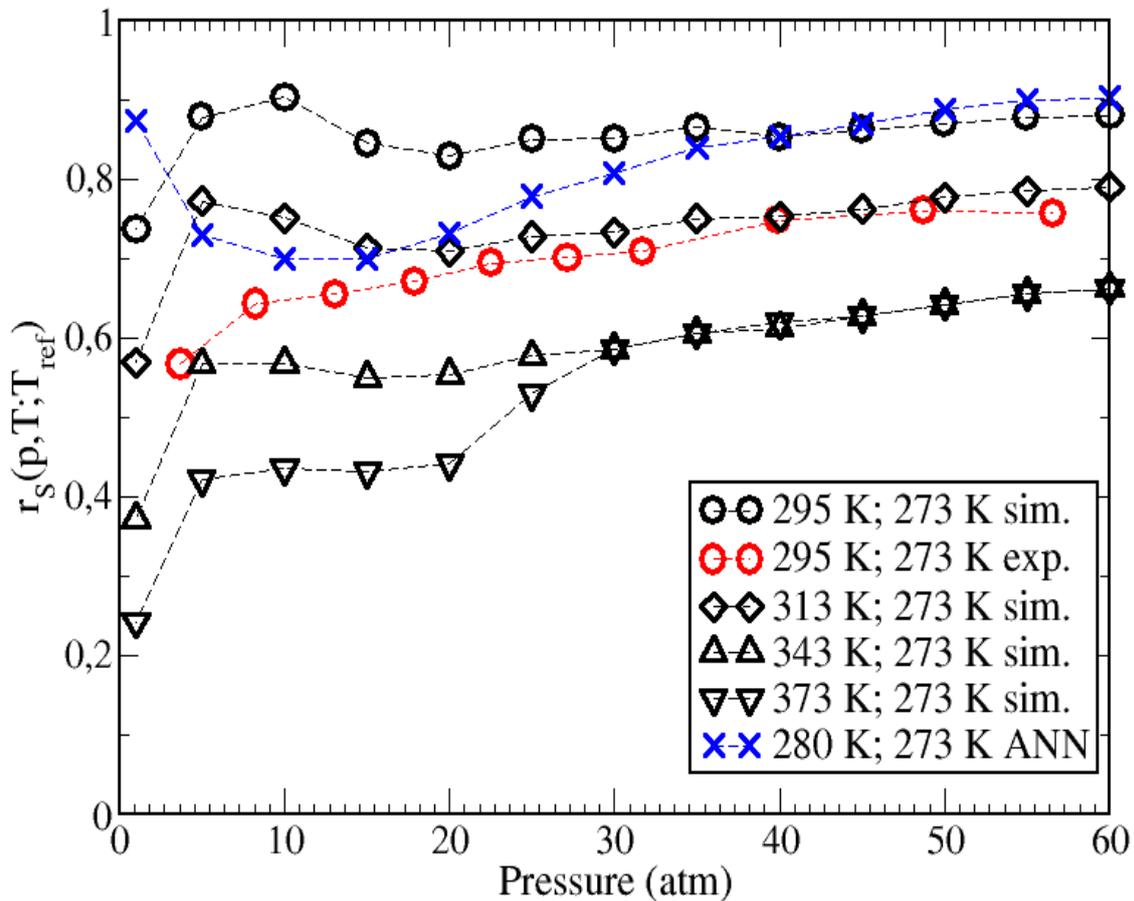
**Table 5.1** Methane loading from GCMC simulations, ANN predictions and Experimental measurements. For the depicted ANN loading prediction, the mean value (55.8 atm) from GCMC (55 atm) and experimental pressures (56.6 atm) was used. In addition, ANN sorbate loading predicted values were monitored at 280 and 300 K which are the closest to 273 and 295 K temperatures used in GCMC simulations and Experiments.

Methodology	Temperature [K]	Pressure [atm]	Methane loading [ $\text{cm}^3$ (STP)/ $\text{cm}^3$ ]
GCMC	273	55	159.5
ANN	280	55.8	143.2
Experiment	273	56.6	132.2
GCMC	295	55	140.1
ANN	300	55.8	127.5
Experiment	295	56.6	100.0

#### 5.4.2 Methane's sorption in NIIC-20-Bu at different temperatures-Sorption ratio

Having in our disposal the dependence of the sorbed loading of methane with the pressure as it is recorded in the sorption isotherm and for different temperatures, additional quantitative comparisons can be made through the calculation of sorbate ratio  $r_S(p, T; T_{ref}) = r_S(p, T) = \frac{c_{sorbate}(p, T)}{c_{sorbate}(p, T_{ref})}$ , where  $c_{sorbate}(p, T)$ ,  $c_{sorbate}(p, T_{ref})$  are the sorbate concentrations at pressure  $p$  and temperatures  $T$ ,  $T_{ref} = 273$  K, respectively. The temperature of 273 K was chosen as reference, because at this temperature the most intense adsorption takes place. It should be noted that the above ratio is calculated for the GCMC simulation data, the

ANN predictions and for the experimental measurements from [27]. In particular, for the experimental data the measurements at 273 K were used as the reference while for the simulation data at 295, 313, 343 K and ANN prediction at 280 K the reference was the GCMC simulation data at 273 K.



**Figure 5.6:** Sorption ratio ( $r_s$ ). In the legend box the  $T; T_{ref}$  is indicated for the GCMC simulation data, experimental measurements and ANN prediction.

Regarding the ratios of sorbed methane concentrations for the lower pressure values in the vicinity of 1 atm, the temperature increase (above the reference temperature of 273 K) results in a reduction of the  $r_s(p, T)$  ratio. In particular, the  $r_s$  attains the following values 0.72, 0.58, 0.38, 0.24 at the respective temperatures of 295, 313, 343, 373 K. The increase of pressure beyond 1 atm leads to an increase in the  $r_s$  ratio for the GCMC simulation data. More specifically, this applies to pressure value up to 5 atm. Further pressure increase leads to plateau values for  $r_s(60 \text{ atm}, 295 \text{ K}) = 0.84$  and  $r_s(60 \text{ atm}, 313 \text{ K}) = 0.8$ . Slight differentiation from the aforementioned behavior is evident for the  $r_s$ -curves at 343 and 373 K, which practically remain constant beyond 5 atm and almost up to the pressure value of 20 atm. The increase of

pressure values beyond 20 atm triggers the increase of  $r_s$  (the increase is higher for the  $r_s$  curve at  $T = 373$  K compared to  $T = 343$  K curve). Additionally, for  $p > 30$  atm an equalization of the slightly increasing concentration ratio is observed for the  $r_s$  curve of 343 and 373 K ( $r_s(p > 30$  atm, 373 K)  $\approx r_s(p > 30$  atm, 343 K)). The slight upward trend of the aforementioned  $r_s$ -curves leads to a ratio value of 0.67 for the higher studied pressure of 60 atm.

Relative differentiation is observed for the sole experimental data available ( $r_s(p, 295\text{K})$ ) depicted in **Figure. 5.6** of the functional dependency of concentration ratio to pressure. A slight increasing tendency of  $r_s$  in the pressure range [3,40] atm, can be observed. The 0.73 plateau value for the concentration ratio is obtained for pressures above 40 atm. Regarding the ANN prediction of  $r_s$  slight deviations are evident when compared to simulation data and experimental measurements. In particular, for pressures up to 10 atm, a downward trend starting from  $r_s(1\text{atm}, 280\text{K}) = 0.88$  ending at  $r_s(10\text{atm}, 280\text{K}) = 0.7$  is clear. For pressure values greater than 10 atm the value of the ratio increases up to  $r_s(50\text{ atm}, 280\text{K}) = 0.88$  and for further increase of pressure in the [50,60] atm range the plateau value  $r_s(60\text{ atm}, 280\text{K}) = 0.91$  is attained. The two aforementioned  $r_s$  values are higher when compared to the simulation data for  $r_s(p > 50\text{ atm}, 295\text{K})$  as expected given that the increase of temperature from 280K to 295 K reduces the sorption capacity.

### 5.4.3 Methane's Sorption in NIIC-20-Bu compared to other porous materials

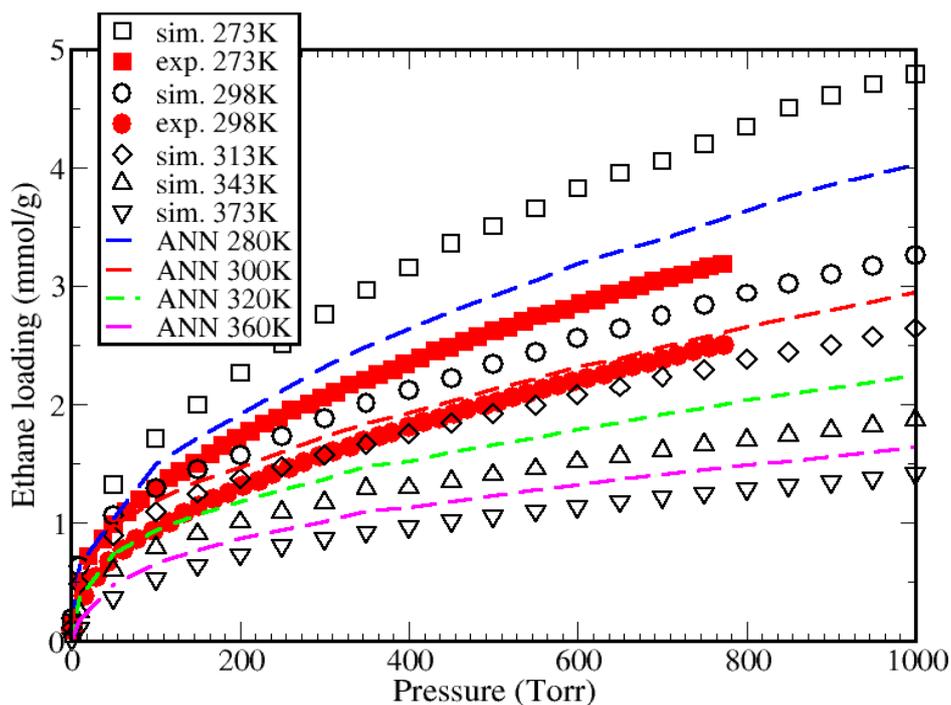
It is noteworthy that the voluminous NIIC-20-Bu aperture (with 25 Å approximate cage diameter) can host a larger number of methane molecules compared to other MOF or Zeolite type materials. For the representative physical system under  $p = 10$  atm and  $T = 273$  K the methane's sorbed amount in NIIC-20-Bu is around  $75.6\text{ cm}^3(\text{STP})/\text{cm}^3$  which is equivalent to 3.8 mmol/g. These values are significantly higher when compared to 3.2 mmol/g for methane sorbed in ZIF-8 (contains cages of diameter 11.6 Å which are connected through 6-ring windows of 3.4 Å) [30]. The improvement of methane's mmol/g sorption amount for NIIC-20-Bu over ZIF-8 is valued at 18.8%.

Similar comparison can be made for methane sorption in the narrow pores (straight and sinusoidal channels with elliptical cross sections with principal axes roughly  $5.4\text{ Å} \times 5.6\text{ Å}$  and  $5.1\text{ Å} \times 5.5\text{ Å}$ ) of zeolite silicalite which is the aluminum free form of Zeolite Socony Mobil-5 (ZSM-5) for the same pressure (10 atm) but for the slight different temperature of 300 K. The  $T = 300\text{ K}$  and  $p = 10\text{ atm}$  ANN prediction depicted with red dashed line in **Figure. 5.5** (a more detailed discussion regarding the ANN results will be given in the sequel) is close to 47.1

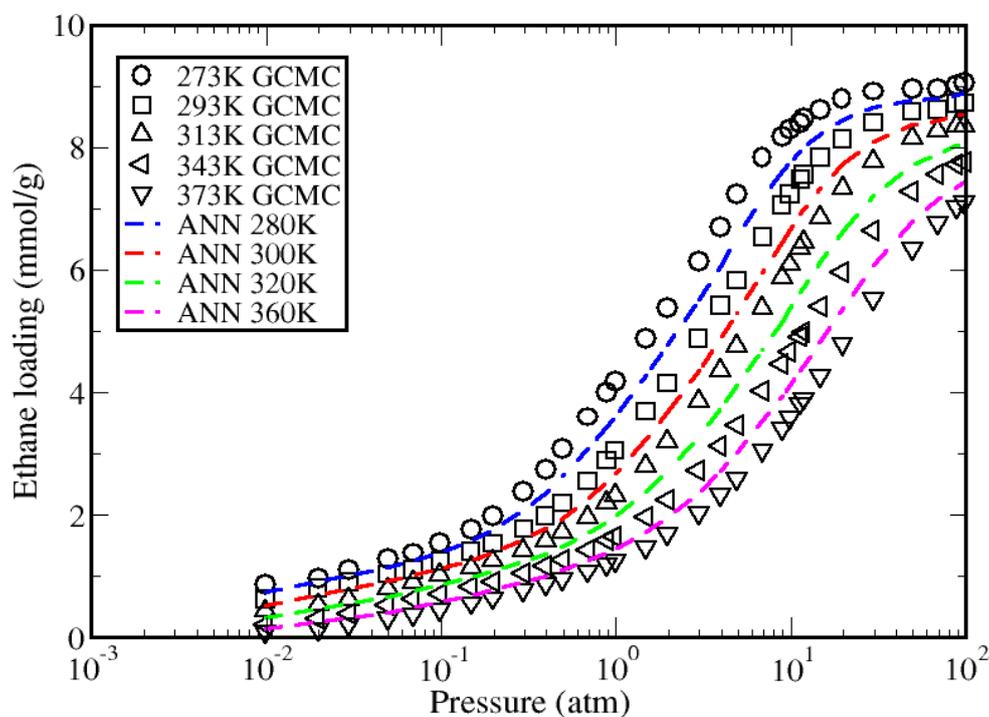
$\text{cm}^3(\text{STP})/\text{cm}^3$  or 2.4 mmol/g which are significantly higher at the same temperature and pressure conditions when compared to methane's sorption amount of 1.8 mmol/g in silicalite [49]. The improvement of methane's mmol/g sorption amount for NIIC-20-Bu over silicalite is valued at 33.3%. The aforementioned calculations demonstrate the high methane adsorption capacity of NIIC-20-Bu sorbent. It is clear that the saturation of the methane loading necessitates elevated pressures beyond 55 atm which was the maximum applied pressure in the experimental work reported in [27].

#### 5.4.4 Ethane Sorption

The sorption isotherms of ethane are depicted in **Figure 5.7** for pressures up to 1000 Torr and in **Figure 5.8** for intermediate to higher (1 to 100 atm) pressure values and have been computed using GCMC simulations for different temperatures varied from  $T=273$  to 373 K. The experimental sorption measurements at 273 and 293 K from [28] are also presented for **Figure 5.7** to have a direct comparison between simulation and experimental results. In addition, ANN predictions are also included. The interplay of pressure and temperature is vital in determining the sorption process's efficacy. In parallel with the findings regarding methane sorption in the previous subsection, analogous trends are observed in the sorption behavior of ethane. Elevated pressure enhances ethane sorption, directly influencing the amount of sorbate loading. Meanwhile, adjusting the system's temperature to different values allows for precise control over the amount of ethane sorbed. The higher the temperature the lower the ethane sorbed amount in NIIC-20-Bu. The simulation results of ethane sorption (**Figure 5.7**) tend to overestimate the sorbed amount, similar the GCMC sorption isotherms at low (**Figure 5.4**) and high (**Figure 5.5**) pressures of methane in NIIC-20-Bu. The sorbed amount for the higher studied pressure of 100 atm is close to 9 (as calculated from GCMC simulation), 8.7 (GCMC), 8.5 (ANN) and 8.3 (GCMC) mmol (sorbate)/(g of sorbent) at 273, 293, 300 and 313 K, respectively, which is significantly higher, in fact four times higher, when compared to the ethane's sorbed amount in zeolite silicalite at similar pressure and for 300 K temperature [29] which demonstrates superior adsorption capacity of NIIC-20-Bu framework.



**Figure 5.7:** GCMC simulation results for sorption isotherms of ethane in NIIC-20-Bu at the low-pressure range for five different temperatures (273, 298, 313, 343, 373 K). In dashed line format the prediction from the ANN at four different temperatures (280, 300, 320, 360 K) are shown.

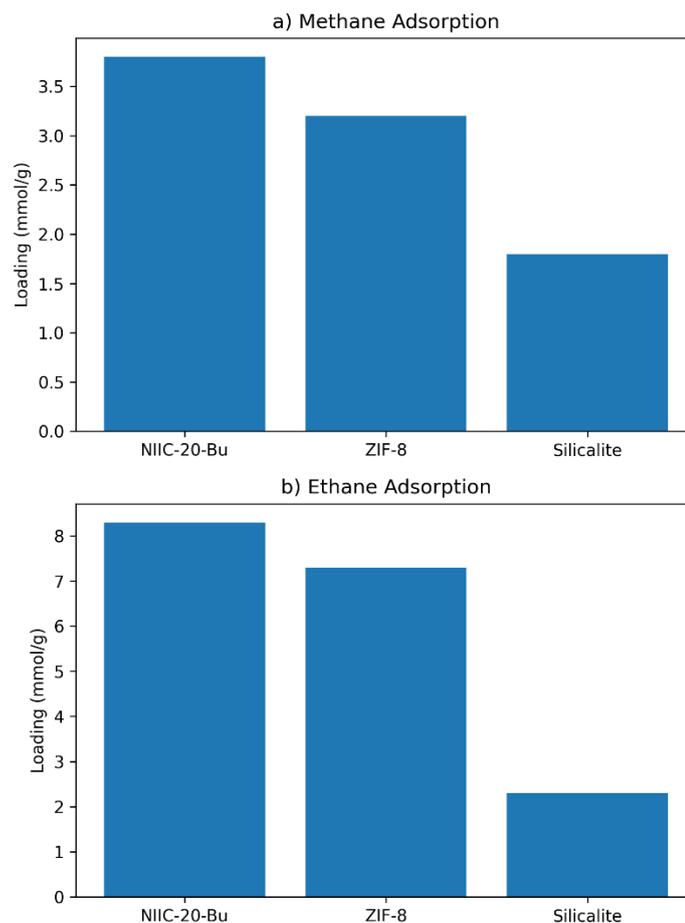


**Figure 5.8:** GCMC simulation results for sorption isotherms of ethane in NIIC-20-Bu at the high-pressure range for five different temperatures (273, 293, 313, 343, 373K). In dashed line format the prediction from the ANN at four different temperatures (280, 300, 320, 360 K) are shown.

#### 5.4.4 Ethane's Sorption in NIIC-20-Bu compared to other porous materials

For the representative physical system under  $p = 10$  atm and  $T=273$  K the ethane's sorbed amount in NIIC-20-Bu is around 8.3 mmol/g. These values are significantly higher when compared to 7.3 mmol/g for ethane sorbed in ZIF-8 [30]. The improvement of ethane's sorption capacity in mmol/g of the NIIC-20-Bu over ZIF-8 is valued at 13.7% at  $T=273$  K.

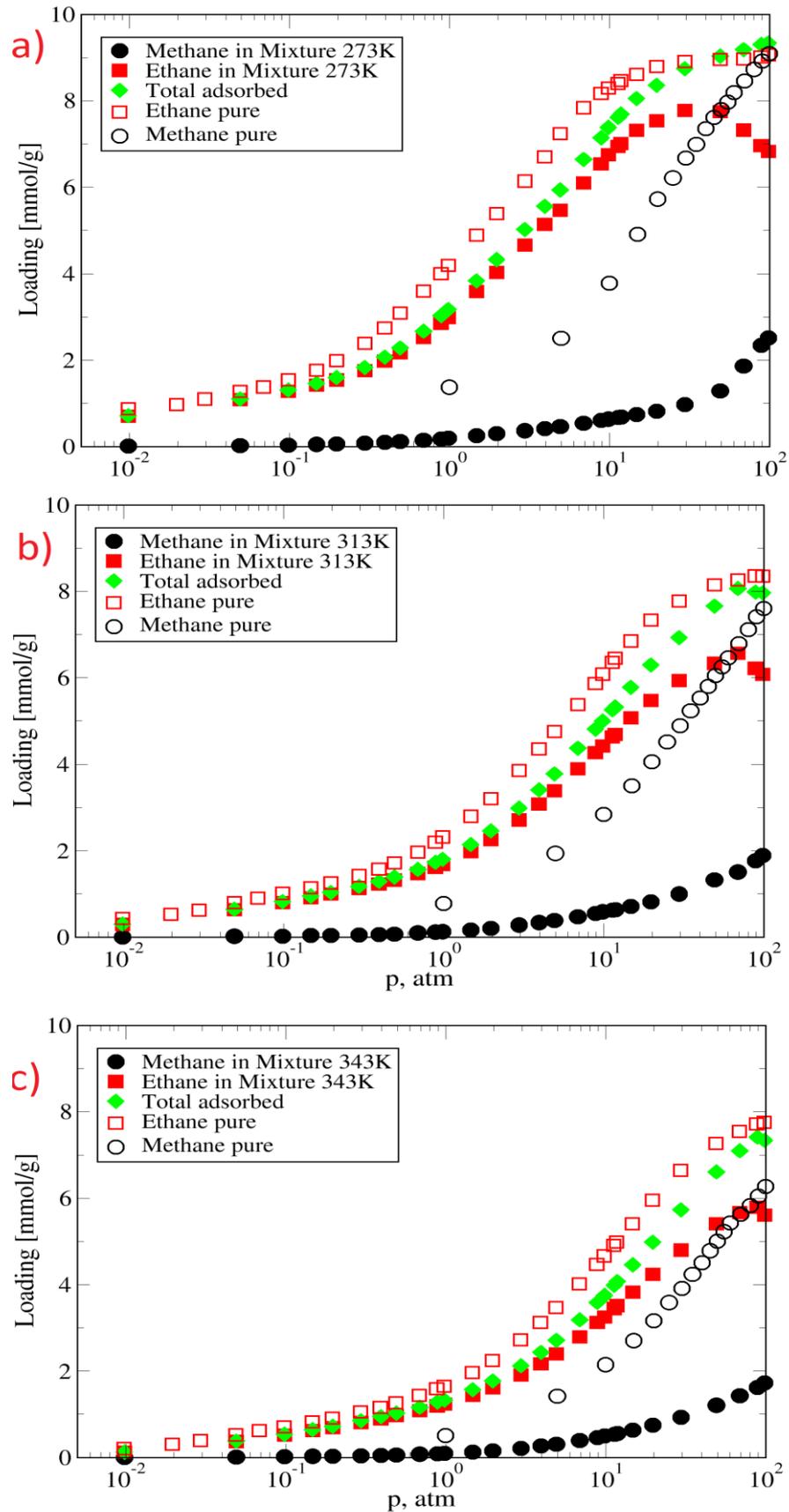
Similar comparisons can be made for ethane sorption in the narrow pores of zeolite silicalite for the same pressure (10 atm) but for the slightly different temperature of 300 K. The ANN NIIC-20-Bu prediction depicted with the red dashed line in **Figure 5.8** at  $T = 300$  K and  $p=10$  atm, is close to 6.7 mmol/g which is significantly higher when compared to ethanes's sorption amount of 2.3 mmol/g in silicalite [29] at the same temperature and pressure conditions. The improvement of ethane's sorption capacity in mmol/g of the NIIC-20-Bu over silicalite is valued at 191.3% at  $p=10$  atm,  $T=300$  K. The aforementioned calculations demonstrate the high adsorption capacity of NIIC-20-Bu sorbent with respect to ethane sorbate. In **Figure 5.9** the comparison of the different frameworks is presented.



**Figure 5.9:** Comparison of NIIC-20-Bu, ZIF-8 and Silicalite for methane (a) and (b) for ethane.

### 5.4.5 Methane-Ethane Mixture Sorption

Additional GCMC simulations have been conducted for the computations of equimolar methane-ethane gas mixture isotherms at three different temperatures  $T=273, 313, 343$  K and are shown in **Figure 5.10**. In addition to the mixture isotherms, the dataset also includes the total adsorption amount and the individual sorption amount values for both pure methane and ethane. These additional data points serve as references, providing further insight into the overall adsorption behavior of the system and allowing for a comparative analysis of the sorption tendencies of methane and ethane in the equimolar mixture. Initially, it can be observed that the sorption characteristics of the mixture significantly deviate from the behavior of the pure hydrocarbons. A competition between the mixture constituents for sorption is evident. In particular, that is reflected in the lower number of hydrocarbon molecules from the gas mixture which are actually sorbed compared to their pure component behavior. The NIIC-20-Bu mesoporous sorbent shows its strong preference for ethane over methane molecules. Ethane molecules from low to high pressure values are sorbed in significantly higher amounts compared to methane in their binary mixture. The available free space for occupation by the hydrocarbon mixture components along with the increase in pressure facilitates the monotonic increase of their sorbed amounts in NIIC-20-Bu. The ethane and methane amounts sorbed per sorbent gram show a constant increase with pressure increase while the methane molecules are sorbed in considerable amounts beyond 30 atm (e.g. 273 K) which is an entropy driven behavior (**Figure 5.10**).

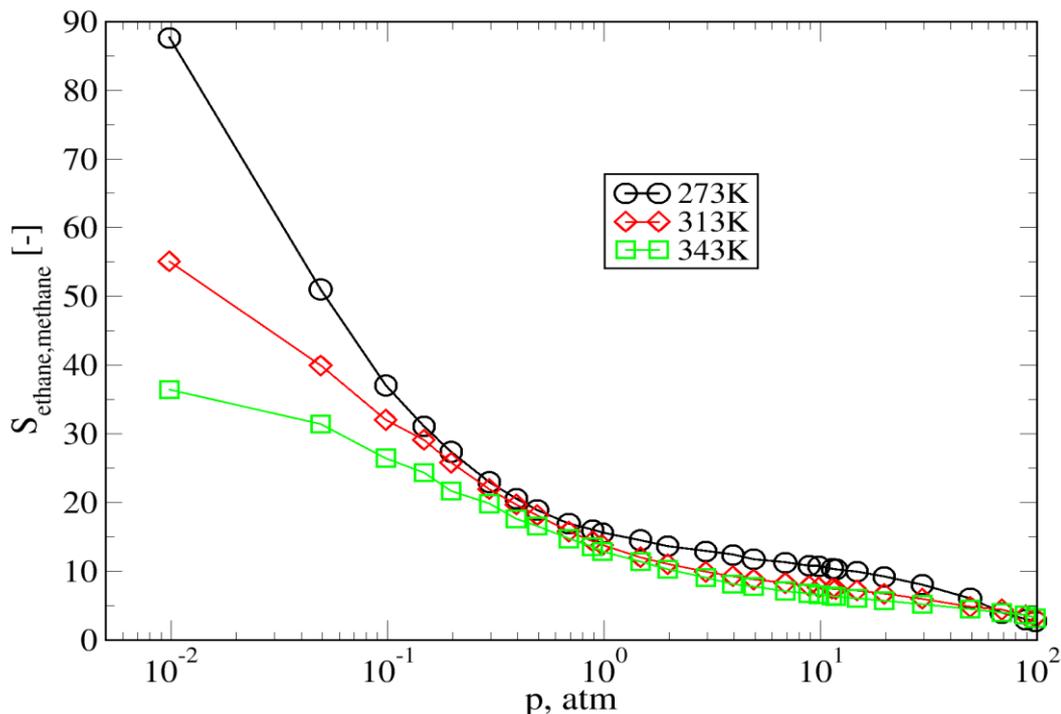


**Figure 5.10:** Sorption of methane-ethane equimolar mixture in NIIC-20-Bu at temperatures: a) 273, b) 313, and c) 343 K.

It is important to highlight that as pressure increases the ethane sorbed amount starts to decrease as can be seen for ethane's sorption isotherms at 273, 313, and 343 K. This sudden loading decrease potentially provides essential space to methane for sorption. A closer inspection of the aforementioned isotherms not only for ethane but also for methane can lead to another interesting sorption characteristic which is associated with the pressure threshold signaling the reduction of ethane amount sorbed. In particular, at 273 K and for ethane's threshold pressure which is around 30 atm the methane's isotherm in the semilogarithmic scale representation of **Figure 5.10a** exhibits significantly augmented sorption loadings as the pressure value increases beyond the threshold value. Therefore, it can be concluded that primarily the methane molecules at the higher pressures seek residence in the limited space created by the decrease of the ethane molecules sorbed in NIIC-20-Bu. The aforementioned pressure threshold corresponds to methane loading approximately equal to 1.0 mmol/g and is very slightly affected by the increase of the temperature to 313 K ( $\approx 1.5$  mmol/g) and 343 K ( $\approx 1.6$  mmol/g), which can be related to the available space for sorption. At the higher studied pressure close to 100 atm the sorbed amount of ethane attains the values of 6.8, 6.2, 5.6 mmol/g while for methane are 2.5, 1.9, 1.7 mmol/g at the aforementioned temperatures, respectively.

#### 5.4.6 Methane-Ethane Mixture Sorption selectivity

For the hydrocarbon mixture the sorption selectivity is defined as  $S_{i,j} = \frac{x_i/x_j}{y_i/y_j}$ , where  $i$  refers to ethane and  $j$  to methane,  $x$  represents the sorbed molar fraction and  $y$  the gas molar fraction. This definition is used in scientific literature [50] and is a direct measure of the separation efficiency of a material. In GCMC simulations, the number of particles of each component in the simulation box directly gives the molar fractions in the adsorbed phase, while the bulk gas phase composition is set by the input chemical potentials. In the case of the present equimolar gas mixture feed, selectivity can be reduced to  $S_{i,j} = \frac{x_i}{x_j} = \frac{c_s^i}{c_s^j}$  and can be readily calculated from the actual amount sorbed  $c_s^i, c_s^j$  of hydrocarbons  $i$  and  $j$  in NIIC-20-Bu, respectively. The selectivity ( $S_{ethane,methane}$ ) calculation is depicted in **Figure 5.11**.



**Figure 5.11:** Selectivity of NIIC-20-Bu for the components of the methane-ethane mixture for  $T=273, 313$  and  $343$  K.

The selectivity calculation offers valuable insights into the intricate interplay between temperature and pressure. Selectivity is promoted with the temperature increase especially in the pressure range  $[0.01$  to  $0.1]$  atm, while the increase of pressure causes the systematic lowering of the selectivity of methane-ethane mixture in NIIC-20-Bu as can be seen in **Figure 5.11**. In particular, lower temperatures, exemplified by  $T = 273$  K, exhibit heightened selectivity across the pressure spectrum, implying a more pronounced propensity for methane and ethane separation. Specifically, by increasing the pressure from  $0.01$  to  $1$  atm selectivity values drop from  $87$  to  $15.5$ . It should be noted that even for higher pressures ( $p \in [1, 10]$  atm) selectivity although declining reaches values above  $10$ . This heightened selectivity suggests a more favorable thermodynamic landscape for differentiation between the two molecules, potentially attributed to reduced kinetic energy and increased intermolecular forces at lower temperatures. Conversely, as temperature escalates to  $313$  and  $343$  K, selectivity gradually wanes, indicative of diminished discriminative capabilities in separating methane and ethane. This decline aligns with established thermodynamic principles wherein elevated temperatures foster greater molecular mobility and intermolecular collisions, mitigating the discerning power of the separation process. Such nuanced temperature-pressure-selectivity relationships underscore the imperative of tailored process design and optimization strategies in diverse

industrial contexts, where precise control over separation parameters is paramount for maximizing efficiency and product purity.

## 5.5 Conclusions

The sorption of methane, ethane and their equimolar gas mixture feed sorbed in the mesoporous NIIC-20-Bu framework has been studied using GCMC simulations and ANNs computational methodologies. The sorption isotherms of methane and ethane for low and high pressures from GCMC simulations have been computed at different temperatures ranging from 273 to 373 K and compared with existing pure component experimental measurements showing reasonable qualitative agreement. The GCMC simulation results in an overestimate of the amount of methane or ethane sorbed in NIIC-20-Bu framework. This can be attributed to the force-field adopted and the rigid framework used. Additionally, hybrid data sets created from the aforementioned experimental measurements and the present GCMC computations have been used to develop ANNs predictive modeling approaches for pure methane and ethane sorption in NIIC-20-Bu providing fast and reliable estimations in wide range of temperatures and pressures. This hybrid approach allows the ANN to make predictions that are both physically plausible and experimentally relevant, effectively serving as a generalized regression tool to fill in the gaps where either experimental data is scarce or simulations are too computationally expensive. The simulation results confirmed the strong adsorption capacity of NIIC-20-Bu for methane and ethane. When compared with other porous structures (silicalite, ZIF-8), the findings highlighted the promising prospects this mesoporous material offers for applications in the storage of light hydrocarbons.

GCMC simulations conducted for an equimolar methane-ethane gas mixture feed showed that the sorption characteristics of the mixture in NIIC-20-Bu significantly differentiate from the pure hydrocarbons' behavior. A competition between the mixture constituents for sorption was evident. The NIIC-20-Bu mesoporous sorbent showed its strong preference for ethane over methane molecules. Ethane molecules from low to high pressure values are sorbed in significantly higher amounts compared to methane in their binary mixture. The sorption selectivity calculation revealed that the pressure increase from 0.01 to 1 atm drives the ethane over methane selectivity values to drop from the remarkable value of 87 to 15.5. It is very interesting that even for pressures between 1 and 10 atm the NIIC-20-Bu framework continues to be a preferential adsorbent of ethane over methane with selectivity values above 10.

A key consideration in employing Artificial Neural Networks (ANNs) is that while their modeling procedure can yield accurate outcomes, the methodology operates as a "black box," lacking insight into the fundamental physics and physicochemical phenomena. To realize the full predictive potential of ANNs, it is imperative to integrate them with supplementary molecular simulations and experiments that explore a wider range of physical conditions/parameters and more complex systems. The computational power of ANNs can thus serve to enhance and validate the invaluable knowledge gained from explicit simulation and experimental methodologies.

## References

- [1] J.-R. Li, R. J. Kuppler, and H.-C. Zhou, "Selective gas adsorption and separation in metal-organic frameworks," *Chem. Soc. Rev.*, vol. 38, no. 5, pp. 1477–1504, 2009, doi: 10.1039/B802426J.
- [2] N. Stock and S. Biswas, "Synthesis of Metal-Organic Frameworks (MOFs): Routes to Various MOF Topologies, Morphologies, and Composites," *Chem Rev*, vol. 112, no. 2, pp. 933–969, Feb. 2012, doi: 10.1021/cr200304e.
- [3] H.-L. Jiang, B. Liu, T. Akita, M. Haruta, H. Sakurai, and Q. Xu, "Au@ZIF-8: CO Oxidation over Gold Nanoparticles Deposited to Metal-Organic Framework," *J Am Chem Soc*, vol. 131, no. 32, pp. 11302–11303, Aug. 2009, doi: 10.1021/ja9047653.
- [4] R. Freund et al., "The current status of MOF and COF applications," *Angewandte Chemie International Edition*, vol. 60, no. 45, pp. 23975–24001, 2021. doi: 10.1002/anie.202106259
- [5] K. Sumida et al., "Carbon Dioxide Capture in Metal-Organic Frameworks," *Chem Rev*, vol. 112, no. 2, pp. 724–781, Feb. 2012, doi: 10.1021/cr2003272.
- [6] M. P. Suh, H. J. Park, T. K. Prasad, and D.-W. Lim, "Hydrogen Storage in Metal-Organic Frameworks," *Chem Rev*, vol. 112, no. 2, pp. 782–835, Feb. 2012, doi: 10.1021/cr200274s.
- [7] J.-B. Lin et al., "A scalable metal-organic framework as a durable physisorbent for carbon dioxide capture," *Science (1979)*, vol. 374, no. 6574, pp. 1464–1469, 2021. doi: 10.1126/science.abi7281
- [8] P. A. Demakov, S. S. Volynkin, D. G. Samsonenko, V. P. Fedin, and D. N. Dybtsev, "A selenophene-incorporated metal-organic framework for enhanced CO<sub>2</sub> uptake and adsorption selectivity," *Molecules*, vol. 25, no. 19, p. 4396, 2020. doi: 10.3390/molecules25194396
- [9] Y. Zhang, S. Yuan, G. Day, X. Wang, X. Yang, and H.-C. Zhou, "Luminescent sensors based on metal-organic frameworks," *Coord Chem Rev*, vol. 354, pp. 28–45, 2018. doi: 10.1016/j.ccr.2017.06.007
- [10] C. Wang, T. Zhang, and W. Lin, "Rational synthesis of noncentrosymmetric metal-organic frameworks for second-order nonlinear optics," *Chem Rev*, vol. 112, no. 2, pp. 1084–1104, 2012. doi: 10.1021/cr200252n

- [11] S. M. J. Rogge et al., “Metal–organic and covalent organic frameworks as single-site catalysts,” *Chem. Soc. Rev.*, vol. 46, no. 11, pp. 3134–3184, 2017, doi: 10.1039/C7CS00033B.
- [12] G. Mínguez Espallargas and E. Coronado, “Magnetic functionalities in MOFs: from the framework to the pore,” *Chem. Soc. Rev.*, vol. 47, no. 2, pp. 533–557, 2018, doi: 10.1039/C7CS00653E.
- [13] C. Duan, S. Gao, and X.-M. Chen, “Special issue on the 8th Chinese Coordination Chemistry Conference,” 2019, Elsevier.
- [14] X.-J. Xie, M.-Y. Zhou, H. Zeng, W. Lu, and D. Li, “Pore Engineering in Metal–Organic Frameworks for Enhanced Hydrocarbon Adsorption and Separation,” *Acc Mater Res*, vol. 6, no. 2, pp. 195–209, Feb. 2025, doi: 10.1021/accountsmr.4c00336.
- [15] C. Jiang et al., “Recent advances in metal–organic frameworks for gas adsorption/separation,” *Nanoscale Adv.*, vol. 4, no. 9, pp. 2077–2089, 2022, doi: 10.1039/D2NA00061J.
- [16] S. Vandenbrande, T. Verstraelen, J. J. Gutiérrez-Sevillano, M. Waroquier, and V. Van Speybroeck, “Methane Adsorption in Zr-Based MOFs: Comparison and Critical Evaluation of Force Fields,” *The Journal of Physical Chemistry C*, vol. 121, no. 45, pp. 25309–25322, Nov. 2017, doi: 10.1021/acs.jpcc.7b08971.
- [17] Y. Zhou, X. Zhang, T. Zhou, and K. Sundmacher, “Computational Screening of Metal–Organic Frameworks for Ethylene Purification from Ethane/Ethylene/Acetylene Mixture,” *Nanomaterials*, vol. 12, no. 5, 2022, doi: 10.3390/nano12050869.
- [18] P. Z. Moghadam, Y. G. Chung, and R. Q. Snurr, “Progress toward the computational discovery of new metal–organic framework adsorbents for energy applications,” *Nat Energy*, vol. 9, no. 2, pp. 121–133, 2024. doi: 10.1038/s41560-023-01417-2
- [19] C. Min et al., “Large-scale computational screening of metal-organic framework membranes for ethane/ethylene separation,” *Acta Chimi Sin*, vol. 80, no. 9, p. 1277, 2022. doi: 10.6023/A22040186
- [20] S. Han and J. Kim, “Design and Screening of Metal–Organic Frameworks for Ethane/Ethylene Separation,” *ACS Omega*, vol. 8, no. 4, pp. 4278–4284, Jan. 2023, doi: 10.1021/acsomega.2c07517.

- [21] S. Zhang et al., “Molecular simulation of CH<sub>4</sub> and CO<sub>2</sub> adsorption behavior in coal physicochemical structure model and its control mechanism,” *Energy*, vol. 285, p. 129474, Nov. 2023, doi: 10.1016/j.energy.2023.129474.
- [22] C. Baykasoglu, H. Mert, and C. U. Deniz, “Grand canonical Monte Carlo simulations of methane adsorption in fullerene pillared graphene nanocomposites,” *J Mol Graph Model*, vol. 106, p. 107909, 2021.
- [23] X. Wu, Z. Cao, X. Lu, and W. Cai, “Prediction of methane adsorption isotherms in metal–organic frameworks by neural network synergistic with classical density functional theory,” *Chemical Engineering Journal*, vol. 459, p. 141612, 2023. doi: 10.1016/j.cej.2023.141612
- [24] H. Tang, L. Duan, and J. Jiang, “Leveraging machine learning for metal–organic frameworks: a perspective,” *Langmuir*, vol. 39, no. 45, pp. 15849–15863, 2023. doi: 10.6023/A22040186
- [25] A. A. Lysova et al., “Tuning the molecular and cationic affinity in a series of multifunctional metal–organic frameworks based on dodecanuclear Zn (II) carboxylate wheels,” *J Am Chem Soc*, vol. 141, no. 43, pp. 17260–17269, 2019. doi: 10.1021/jacs.9b08322.
- [26] A. A. Lysova, D. G. Samsonenko, K. A. Kovalenko, A. S. Nizovtsev, D. N. Dybtsev, and V. P. Fedin, “A series of mesoporous metal-organic frameworks with tunable window sizes and exceptionally high ethane over ethylene adsorption selectivity,” *Angewandte Chemie International Edition*, vol. 59, no. 46, pp. 20561–20567, 2020. doi: 10.1002/anie.202008132
- [27] A. A. Lysova, K. A. Kovalenko, D. N. Dybtsev, S. N. Klyamkin, E. A. Berdonosova, and V. P. Fedin, “Hydrocarbon adsorption in a series of mesoporous metal-organic frameworks,” *Microporous and Mesoporous Materials*, vol. 328, p. 111477, 2021. doi: 10.1016/j.micromeso.2021.111477
- [28] A. A. Lysova, K. A. Kovalenko, A. S. Nizovtsev, D. N. Dybtsev, and V. P. Fedin, “Efficient separation of methane, ethane and propane on mesoporous metal-organic frameworks,” *Chemical Engineering Journal*, vol. 453, p. 139642, 2023. doi: 10.1016/j.cej.2022.139642

- [29] A. Gurras and L. N. Gergidis, “Modeling sorption and diffusion of alkanes, alkenes, and their mixtures in silicalite: from MD and GCMC molecular simulations to artificial neural networks,” *Adv Theory Simul*, vol. 4, no. 3, p. 2000210, 2021. doi: 10.1002/adts.202000210
- [30] A. Gkourras and L. N. Gergidis, “Molecular and Artificial Neural Networks Modeling of Sorption and Diffusion of Small Alkanes, Alkenes and Their Ternary Mixtures in Zif-8 at Different Temperatures,” *J Phys Chem B*, vol. 126, no. 29, pp. 5582–5594, 2022. doi: 10.1021/acs.jpcc.2c03478
- [31] T. Zhou, Z. Song, and K. Sundmacher, “Big Data Creates New Opportunities for Materials Research: A Review on Methods and Applications of Machine Learning for Materials Design,” *Engineering*, vol. 5, no. 6, pp. 1017–1026, 2019, doi: <https://doi.org/10.1016/j.eng.2019.02.011>.
- [32] Z. Wang, Y. Zhou, T. Zhou, and K. Sundmacher, “Identification of optimal metal-organic frameworks by machine learning: Structure decomposition, feature integration, and predictive modeling,” *Comput Chem Eng*, vol. 160, p. 107739, 2022. doi: 10.1016/j.compchemeng.2022.107739
- [33] A. K. Rappé, C. J. Casewit, K. S. Colwell, W. A. Goddard III, and W. M. Skiff, “UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations,” *J Am Chem Soc*, vol. 114, no. 25, pp. 10024–10035, 1992. doi: 10.1021/ja00051a040
- [34] S. L. Mayo, B. D. Olafson, and W. A. Goddard, “DREIDING: a generic force field for molecular simulations,” *Journal of Physical chemistry*, vol. 94, no. 26, pp. 8897–8909, 1990. doi: 10.1021/j100389a010.
- [35] 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*, vol. 42, no. 2, pp. 81–101, 2016. doi: 10.1080/08927022.2015.1010082
- [36] C. Christodoulou and M. Georgiopoulos, “Applications of neural networks in electromagnetics. Artech House,” Inc.: Norwood, MA, USA, 2000.
- [37] K. Suzuki, *Artificial neural networks: methodological advances and biomedical applications*. BoD–Books on Demand, 2011.
- [38] C. M. Bishop and N. M. Nasrabadi, *Pattern recognition and machine learning*, vol. 4, no. 4. Springer, 2006. doi: 10.1117/1.2819119

- [39] X. Cheng, Y. Liao, Z. Lei, J. Li, X. Fan, and X. Xiao, "Multi-scale design of MOF-based membrane separation for CO<sub>2</sub>/CH<sub>4</sub> mixture via integration of molecular simulation, machine learning and process modeling and simulation," *J Memb Sci*, vol. 672, p. 121430, 2023, doi: <https://doi.org/10.1016/j.memsci.2023.121430>.
- [40] H. R. S. Anna, A. G. Barreto Jr, F. W. Tavares, and M. B. de Souza Jr, "Machine learning model and optimization of a PSA unit for methane-nitrogen separation," *Comput Chem Eng*, vol. 104, pp. 377–391, 2017. doi: [doi.org/10.1016/j.compchemeng.2017.05.006](https://doi.org/10.1016/j.compchemeng.2017.05.006)
- [41] J. F. Mas and J. J. Flores, "The application of artificial neural networks to the analysis of remotely sensed data," *Int J Remote Sens*, vol. 29, no. 3, pp. 617–663, 2008. doi: [7.062798741e-15](https://doi.org/10.1080/01447270802798741)
- [42] J.-R. Zhang, J. Zhang, T.-M. Lok, and M. R. Lyu, "A hybrid particle swarm optimization–back-propagation algorithm for feedforward neural network training," *Appl Math Comput*, vol. 185, no. 2, pp. 1026–1037, 2007. doi: [10.1016/j.amc.2006.07.025](https://doi.org/10.1016/j.amc.2006.07.025).
- [43] F. Burden and D. Winkler, "Bayesian Regularization of Neural Networks," *Methods Mol Biol*, vol. 458, pp. 23–42, Nov. 2009, doi: [10.1007/978-1-60327-101-1\\_3](https://doi.org/10.1007/978-1-60327-101-1_3).
- [44] S. V. Alavi Nezhad Khalil Abad, M. Yilmaz, D. Jahed Armaghani, and A. Tugrul, "Prediction of the durability of limestone aggregates using computational techniques," *Neural Comput. Appl.*, vol. 29, no. 2, pp. 423–433, Jan. 2018, doi: [10.1007/s00521-016-2456-8](https://doi.org/10.1007/s00521-016-2456-8).
- [45] S. Wang, Z. Xu, T. Stratford, B. Li, Q. Zeng, and J. Su, "Machine learning approach for analysing and predicting the modulus response of the structural epoxy adhesive at elevated temperatures," *J Adhes*, vol. 99, pp. 1–19, Nov. 2023, doi: [10.1080/00218464.2023.2183851](https://doi.org/10.1080/00218464.2023.2183851).
- [46] V. MATLAB, "9.3. 0.713579 (R2017b); The Mathworks," Inc.: Natick, MA, USA, 2017.
- [47] H. W. B. Teo, A. Chakraborty, and S. Kayal, "Evaluation of CH<sub>4</sub> and CO<sub>2</sub> adsorption on HKUST-1 and MIL-101 (Cr) MOFs employing Monte Carlo simulation and comparison with experimental data," *Appl Therm Eng*, vol. 110, pp. 891–900, 2017. doi: [10.1016/j.applthermaleng.2016.08.126](https://doi.org/10.1016/j.applthermaleng.2016.08.126)
- [48] R. Babarao, S. Dai, and D. Jiang, "Effect of pore topology and accessibility on gas adsorption capacity in zeolitic- imidazolate frameworks: bringing molecular simulation close

to experiment,” *The Journal of Physical Chemistry C*, vol. 115, no. 16, pp. 8126–8135, 2011. doi: 10.1021/jp1117294.

[49] R. Q. Snurr, R. L. June, A. T. Bell, and D. N. Theodorou, “Molecular simulations of methane adsorption in silicalite,” *Mol Simul*, vol. 8, no. 1–2, pp. 73–92, 1991. doi: 10.1080/08927029108022468

[50] H. Aljamaan, M. Al Ismail, and A. R. Kavscek, “Experimental investigation and Grand Canonical Monte Carlo simulation of gas shale adsorption from the macro to the nano scale,” *J Nat Gas Sci Eng*, vol. 48, pp. 119–137, 2017. doi: 10.1016/j.jngse.2016.12.024.

[51] L. N. Gergidis and D. N. Theodorou, “Molecular Dynamics Simulation of n-Butane-Methane Mixtures in Silicalite,” *J Phys Chem B*, vol. 103, no. 17, pp. 3380–3390, 1999. doi: 10.1021/jp983680p

[52] L. N. Gergidis, D. N. Theodorou, and H. Jovic, “Dynamics of n-Butane- Methane Mixtures in Silicalite, Using Quasielastic Neutron Scattering and Molecular Dynamics Simulations,” *J Phys Chem B*, vol. 104, no. 23, pp. 5541–5552, 2000. doi: 10.1021/jp0000073

# Chapter 6 Conclusions and Future work

## 6.1 Conclusions

At this stage of the PhD thesis, a summary section is provided that highlights the critical aspects of all research items investigated and also integrates them into a brief narrative that reflects the overall progression of the work. This section expands on the key findings from each study. It clarifies how the individual components interrelate, showing how each line of investigation builds upon the previous ones. In doing so, it demonstrates how the collected results converge toward the central aims of the thesis, articulating the broader impact of the research, its contribution to the field, and its implications for future studies.

### 6.1.1 ZIF-8

Based on all previous simulation activity presented in the earlier sections, a summary of the key findings regarding the ZIF-8 MOF is provided in the following paragraphs. This summary brings together the main observations made throughout the study and offers a clearer picture of how the material behaved across the different simulations. It highlights the recurring trends that appeared, the general responses of ZIF-8 under the various conditions examined, and the broader insights that can be drawn when all results are viewed collectively. By connecting these points, the summary helps outline the overall understanding gained about ZIF-8 and clarifies its relevance within the wider goals of the thesis.

The sorption isotherms and self-diffusion coefficients of small alkanes, alkenes, and selected ternary mixtures in ZIF-8 were examined using molecular simulations and ANN models. NVT-MD results for methane, ethane, ethene, propane, propene, butane, butene, and an equimolar methane–ethane–ethene mixture agree with MAS PFG NMR data across 273–373 K, capturing the expected decrease in diffusivity with lower temperature. Experimental, simulation, and ANN results consistently show higher diffusivities for lighter hydrocarbons, following,  $D_s(\text{methane}) > D_s(\text{ethene}) > D_s(\text{ethane}) > D_s(\text{propene}) > D_s(\text{propane}) > D_s(\text{butene}) > D_s(\text{butane})$ . In the ternary mixture, methane remains the fastest species in ZIF-8, while ethane is the slowest.

GCMC sorption isotherms and ANN predictions were produced for the same species and mixtures over 273–373 K and  $10^{-2}$ – $10^2$  bar. In the ternary mixture, ethane is favored over ethene, and methane uptake remains minimal even near 100 bar.

Combining simulations, experiments, and ANN modeling demonstrated that ANNs accurately predict both self-diffusion and sorption behavior in ZIF-8. Their integration with molecular simulations and experiments offers an efficient and reliable framework for studying hydrocarbon sorption and transport with modest computational cost.

Below a set of important highlights are presented to highlight the simulations performed on the framework of ZIF-8.

- Pure-component diffusion: Alkenes typically outperform same-carbon alkanes (methane > ethene > propene > propane > butene > butane); diffusivity decreases with size.
- Diffusivity magnitudes: C<sub>4</sub> species drop to  $\sim 10^{-12}$  m<sup>2</sup>/s, which creates gaps that allow for kinetic separations, while methane is  $\sim 10^{-10}$  m<sup>2</sup>/s.
- Temperature dependence:  $D_s$  increases Arrhenius-like with  $T$  (273–373 K); ANNs pick up on and replicate the trend within the pressure and temperature domain.
- Ternary mixture mobility: NMR and simulations are in good agreement with only minor absolute differences, with methane remaining the most mobile component in equimolar methane–ethane–ethene.
- Kinetic selectivities:  $D_s$  ratios compress with loading and  $T$  and are highest at low  $T$ . Ethane and ethene slow down more strongly than methane, which is least sensitive to loading.
- Why methane moves more quickly: Larger species stay in cage interiors for longer; residence probability maps indicate methane near apertures/windows which means more inter-cage hops.
- Sorption of pure components: methane Low affinity at low  $T$ , affects mixture behavior at moderate pressures; requires high  $P$  to approach saturation.
- Ethane vs. ethene uptake:  $P$  and  $T$  decrease preference, with ethane being preferred at low-to-mid  $P$  and low  $T$ .
- Features of C<sub>3</sub>/C<sub>4</sub> saturation: Butane/butene saturate at much lower  $P$ ; unit-cell capacities are approximately 12–16 molecules/uc; propane/propene saturate above  $\sim 10$  bar (273 K).
- Butane vs. butene crossover: They are comparable at room temperature; however, at higher  $T$ , butene may be preferred at low  $P$  due to the subtle interplay of size, shape, and framework dynamics.

- Selectivity for mixture sorption: ZIF-8 favors ethane over ethene and methane in equimolar feeds, particularly at lower  $T$ , selectivity decreases as sites become saturated.

Below a table is presented with comparisons about the Molecular Dynamics simulations presented in this research.

**Table 6.1:** Comparison of Molecular Dynamics simulations of pure components and ternary mixture.

Aspect	Pure components	Ternary mixture (CH <sub>4</sub> -C <sub>2</sub> H <sub>6</sub> -C <sub>2</sub> H <sub>4</sub> , equimolar)
Ordering of self-diffusivity ( $D_s$ )	CH <sub>4</sub> > C <sub>2</sub> H <sub>4</sub> > C <sub>2</sub> H <sub>6</sub> > C <sub>3</sub> H <sub>6</sub> > C <sub>3</sub> H <sub>8</sub> > C <sub>4</sub> H <sub>8</sub> > C <sub>4</sub> H <sub>10</sub>	Ordering preserved for CH <sub>4</sub> -C <sub>2</sub> H <sub>6</sub> -C <sub>2</sub> H <sub>4</sub> : CH <sub>4</sub> fastest; C <sub>2</sub> H <sub>4</sub> > C <sub>2</sub> H <sub>6</sub>
Typical magnitudes	CH <sub>4</sub> ~ (2–3) × 10 <sup>-10</sup> m <sup>2</sup> /s (293–373 K); C <sub>4</sub> ~ (1–3) × 10 <sup>-12</sup> m <sup>2</sup> /s	CH <sub>4</sub> remains ~2 × 10 <sup>-10</sup> m <sup>2</sup> /s at high $T$ ; C <sub>2</sub> species reduced by ~1.7–2.3 × vs pure (loading dependent)
Loading sensitivity	CH <sub>4</sub> weak dependence; C <sub>2</sub> -C <sub>4</sub> slow down significantly with loading	Same trend; competition further slows C <sub>2</sub> species vs CH <sub>4</sub>
Temperature dependence	Arrhenius-like ( $D_s \uparrow$ with $T$ )	Arrhenius-like behavior preserved; ratios compress as $T \uparrow$
Siting Preference	CH <sub>4</sub> near apertures/windows; larger species in cage interior	Co-adsorbates modestly enhance window residence for CH <sub>4</sub>

Another table is presented with the comparison of the GCMC simulations presented in this research.

**Table 2:** Comparison of GCMC simulations of pure components and ternary mixture.

Aspect	Pure components	Ternary mixture (CH <sub>4</sub> –C <sub>2</sub> H <sub>6</sub> –C <sub>2</sub> H <sub>4</sub> , equimolar)
Adsorption preference	Ethane > Ethene > Methane (at low–mid $P$ , low $T$ )	Ethane strongly preferred; selectivities $S_{\text{ethane/ethene}} \approx 2.5\text{--}3.5$ (273–300 K); $S_{\text{ethane/methane}} \approx 10\text{--}14$ (273 K)
Methane threshold	High $P$ required to approach saturation at low $T$	Mixture methane loading remains low until a pressure threshold ( $\sim 40\text{--}85$ bar, increasing with $T$ ), then rises more steeply
Isotherm features	Ethane rises rapidly; methane relatively flat then increases; capacities $\sim 12\text{--}16$ mol/uc (C <sub>3</sub> /C <sub>4</sub> reach saturation earlier)	At low $T$ , ethane can slightly decline as methane rises; methane isotherm shows a slope ‘kink’ above $\sim 40$ bar
Temperature sensitivity	Higher $T$ reduces loadings and preferences	Mixture selectivities decrease with $T$ ; methane threshold shifts to higher $P$
Saturation	Heavier species saturate earlier	Total capacity shared across species

### 6.1.2 NIIC-20 Bu

Based on all previous simulation activity presented in the earlier sections, a more complete summary of the main findings regarding the NIIC-20 Bu MOF is provided in the following paragraphs. This summary brings together the most important observations made throughout the work and highlights the points that appeared consistently across the different stages of the study. It explains the general trends that were seen, how the material behaved under various

conditions, and what overall picture emerges when all results are viewed together. In this way, the summary offers a clear and unified understanding of the NIIC-20 Bu MOF and its role within the broader aims of the thesis.

The sorption of methane, ethane, and their equimolar mixture in the mesoporous NIIC-20-Bu framework was examined using GCMC simulations and ANN models. Sorption isotherms for both gases were computed from 273 to 373 K at low and high pressures and compared with available experimental data, showing reasonable qualitative agreement but with GCMC generally overpredicting uptake because of the chosen force field and rigid framework. Hybrid datasets combining experimental and GCMC results were then used to train ANNs for fast, reliable sorption predictions across broad temperature and pressure ranges. This approach yields physically consistent and experimentally relevant estimates while reducing the need for additional experiments or costly simulations. NIIC-20-Bu exhibited strong methane and ethane adsorption and compared favorably with other porous materials such as silicalite and ZIF-8.

For the equimolar methane–ethane mixture, GCMC simulations revealed competitive adsorption with a strong ethane preference. Ethane uptake remained significantly higher than methane across the entire pressure range. Selectivity decreased from 87 at 0.01 atm to 15.5 at 1 atm, yet remained above 10 even up to 10 atm, indicating sustained ethane selectivity.

While ANNs provide accurate predictions, their black-box nature limits mechanistic insight. Their full value is realized when combined with molecular simulations and experiments that broaden the explored physical conditions, enabling ANNs to complement and reinforce the understanding gained from explicit computational and experimental studies.

Below, a set of key highlights is presented to summarize the simulations performed on the NIIC-20 Bu framework. These points capture the most relevant observations and provide a clear overview of the main outcomes that emerged throughout the study

- Pure-component sorption (using experiments and ANN): For both CH<sub>4</sub> and C<sub>2</sub>H<sub>6</sub>, uptake rises with pressure and falls with temperature. GCMC tends to overestimate absolute capacities vs experiments, but trends match; ANN fits reproduce isotherm families quickly.
- High-pressure ethane: C<sub>2</sub>H<sub>6</sub> sorption is higher than silicalite under comparable conditions, reaching ~9/8.7/8.5/8.3 mmol g<sup>-1</sup> at 273/293/300/313 K at 100 atm.

Benchmarking of materials: At 300 K and 10 atm: NIIC-20-C<sub>2</sub>H<sub>6</sub> ≈ 6.7 mmol g<sup>-1</sup> vs silicalite 2.3 mmol g<sup>-1</sup> Bu CH<sub>4</sub> ≈ 2.4 mmol g<sup>-1</sup> vs silicalite 1.8 mmol g<sup>-1</sup> (+~33%).

- Strong ethane preference in mixture sorption (equimolar CH<sub>4</sub>–C<sub>2</sub>H<sub>6</sub>): Due to site competition, mixture loadings are non-additive (lower than the sum of pure-component uptakes); ethane predominates at all pressures and temperatures, although selectivity decreases as *P* and *T* increase.
- Pressure threshold: Ethane loading can drop at about 30 atm at 273 K; above this point, the uptake slope of methane rises. The loadings at ~100 atm are approximately C<sub>2</sub>H<sub>6</sub> ≈ 6.8/6.2/5.6 and CH<sub>4</sub> ≈ 2.5/1.9/1.7 mmol g<sup>-1</sup> at 273/313/343 K, respectively.
- Selectivity vs. temperature and pressure: Equimolar selectivity S(C<sub>2</sub>H<sub>6</sub>/CH<sub>4</sub>) is very high at low *P* (e.g., ~87 at 0.01 atm, 273 K) and decreases with pressure (e.g., ~15.5 at 1 atm, 273 K), staying >10 up to ~10 atm; selectivity decreases as temperature rises.

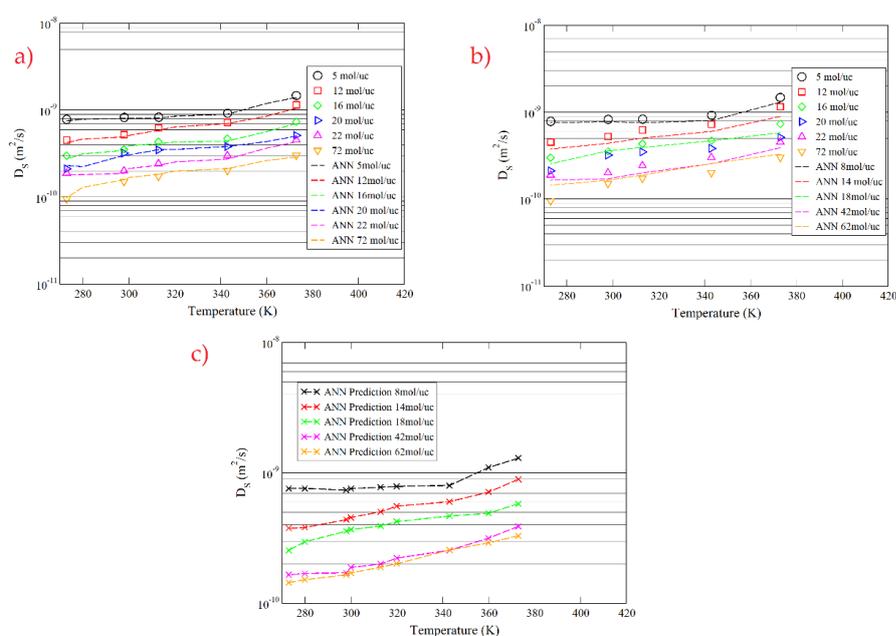
Below a table is presented with a comparison between the pure components and their mixture

**Table 3:** Comparison of GCMC simulations of pure components and their binary mixtures

Aspect	Pure components (CH <sub>4</sub> , C <sub>2</sub> H <sub>6</sub> )	Equimolar mixture (CH <sub>4</sub> –C <sub>2</sub> H <sub>6</sub> )
Adsorption preference	Ethane > Methane across 273–373 K and broad <i>p</i> ; stronger framework affinity.	Strong ethane preference persists; selectivity S(C <sub>2</sub> H <sub>6</sub> /CH <sub>4</sub> ) high at low <i>P/T</i> and decreases with ↑ <i>P</i> and ↑ <i>T</i> (e.g., >10 up to ~10 atm at 273 K).
Uptake magnitudes (representative)	At 100 atm: C <sub>2</sub> H <sub>6</sub> ≈ 9/8.7/8.5/8.3 mmol g <sup>-1</sup> at 273/293/300/313 K; CH <sub>4</sub> lower (e.g., ~3–4 mmol g <sup>-1</sup> at 10 atm, 273–300 K).	At 100 atm: C <sub>2</sub> H <sub>6</sub> ≈ 6.8/6.2/5.6 and CH <sub>4</sub> ≈ 2.5/1.9/1.7 mmol g <sup>-1</sup> at 273/313/343 K.
Pressure dependence	Monotonic increase of uptake with Pressure for both gases.	Around ~30 atm at 273 K, C <sub>2</sub> H <sub>6</sub> loading can decline; CH <sub>4</sub> uptake slope increases beyond this ‘space-making’ threshold.
Temperature dependence	Uptake decreases with <i>T</i> for both components.	Selectivity and ethane loading weaken with <i>T</i> ; CH <sub>4</sub> ‘activation’ threshold shifts to higher <i>P</i> at higher <i>T</i> .
Total capacity	Sum of individual pure uptakes reflects independent occupancy.	Total mixture uptake is lower than the sum of pure uptakes (non-additive) due to competitive site occupation.

## 6.2 Future work

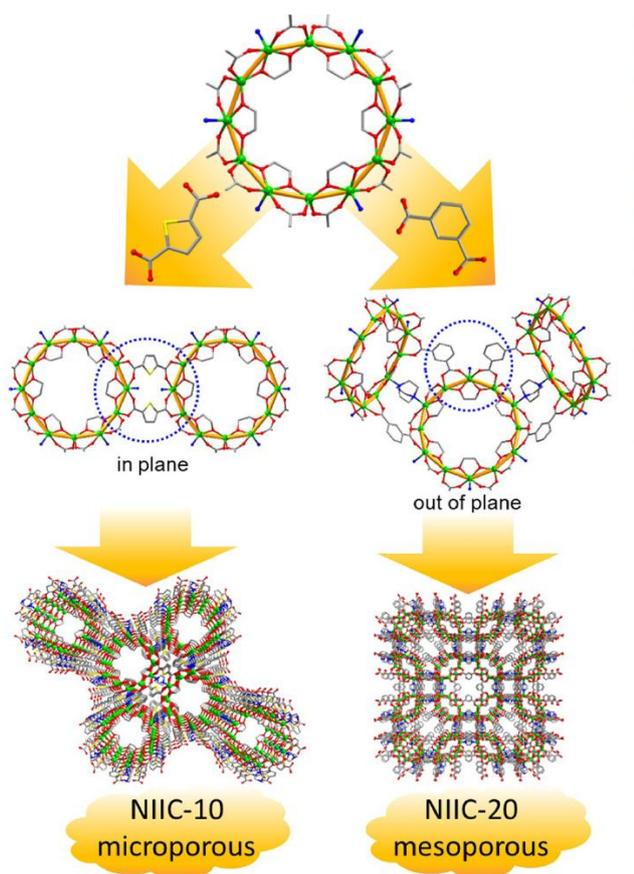
The study of porous materials, and MOF's in particular, presents an exceptionally wide range of possible research directions, from basic structural characterization to sophisticated functional applications. The study of molecular diffusion processes is one of the most important of these since transport phenomena in porous channels have a significant impact on how well they work in applications like gas separation and catalysis. Because of this material's distinct pore topology and adjustable chemical environment, the diffusion behavior of alkanes within the NIIC-20Bu framework has been the focus of much research. In the future, there are a number of ways to further develop and broaden this research using both advanced machine learning approaches and improved classical simulation techniques. Utilizing these methodologies can give great insights into the intricate interactions between the forces that control adsorption and diffusion by continuously improving computational models to represent ever-more-realistic features of framework behavior and molecule interactions. In order to capture the mechanistic complexities and quantitative trends in diffusivity, these studies have combined traditional molecular simulations with contemporary machine-learning-based techniques, such as ANN's. Another aspect of this approach would be to simulate the diffusion of small alkanes and alkenes in the NIIC-20 Bu combined with ANN's. Such a dual-method approach has already been initialized and allows for the validation of computational predictions as shown in **Figure 6.1**.



**Figure 6.1:** Molecular Dynamics simulations in NIIC-20 Bu coupled with ANN's prediction at different temperatures and loadings.

### 6.2.1 Classical Molecular simulations

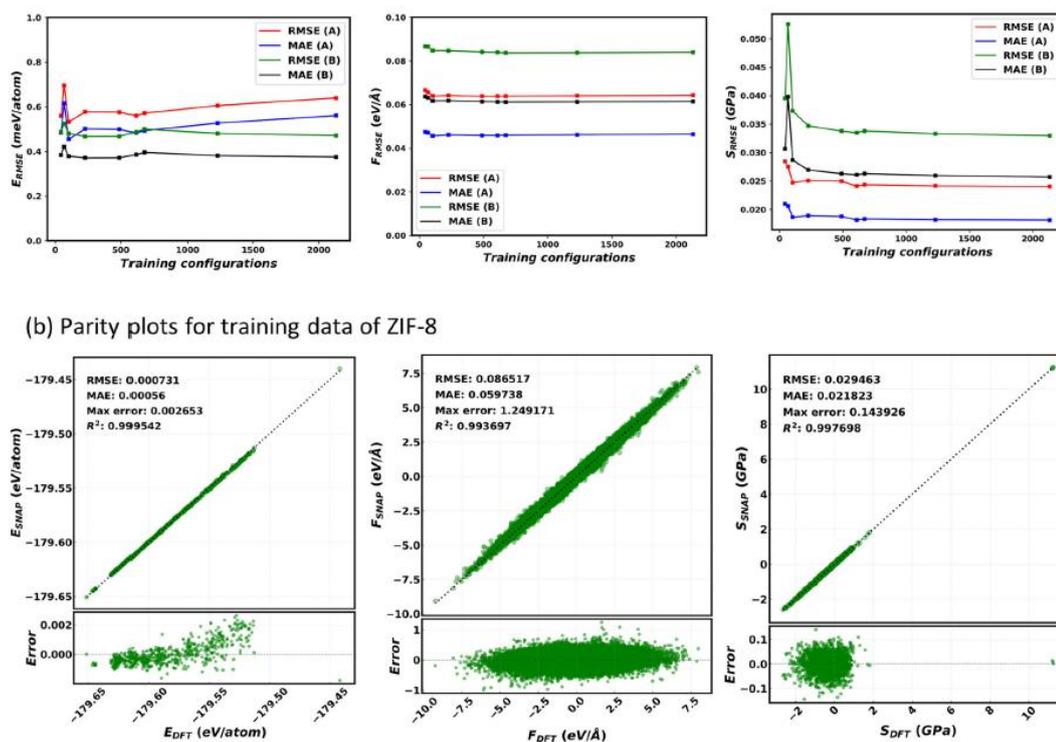
Future developments of the present work can focus on a wider variety of materials in the NIIC family [1], [2], [3], [4] as well as other classes of metal-organic frameworks that show promising structural and transport properties with regard to classical Molecular Dynamics and Monte Carlo simulations. The simulations' scope can be expanded to include longer-chain hydrocarbon molecules like pentane and hexane, which add more conformational degrees of freedom and interaction dynamics, in addition to investigating different framework chemistries. Additionally, research on mixtures of long-chain and short-chain hydrocarbons may provide information on competitive or synergistic transport processes inside the porous framework. These simulations would enable a systematic understanding of how molecular structure and composition influence diffusion and adsorption behavior in porous media. Within this framework, key transport properties such as self-diffusivity and transport diffusivity can be rigorously computed for both pure components and complex multicomponent systems, including binary and ternary mixtures.



**Figure 6.2:** Local coordination and orientation of dicarboxylate bridging ligands in NIIC-10 (left) and NIIC-20 (right), leading to the formation of micro- and mesoporous frameworks, respectively [1], [2], [3], [4].



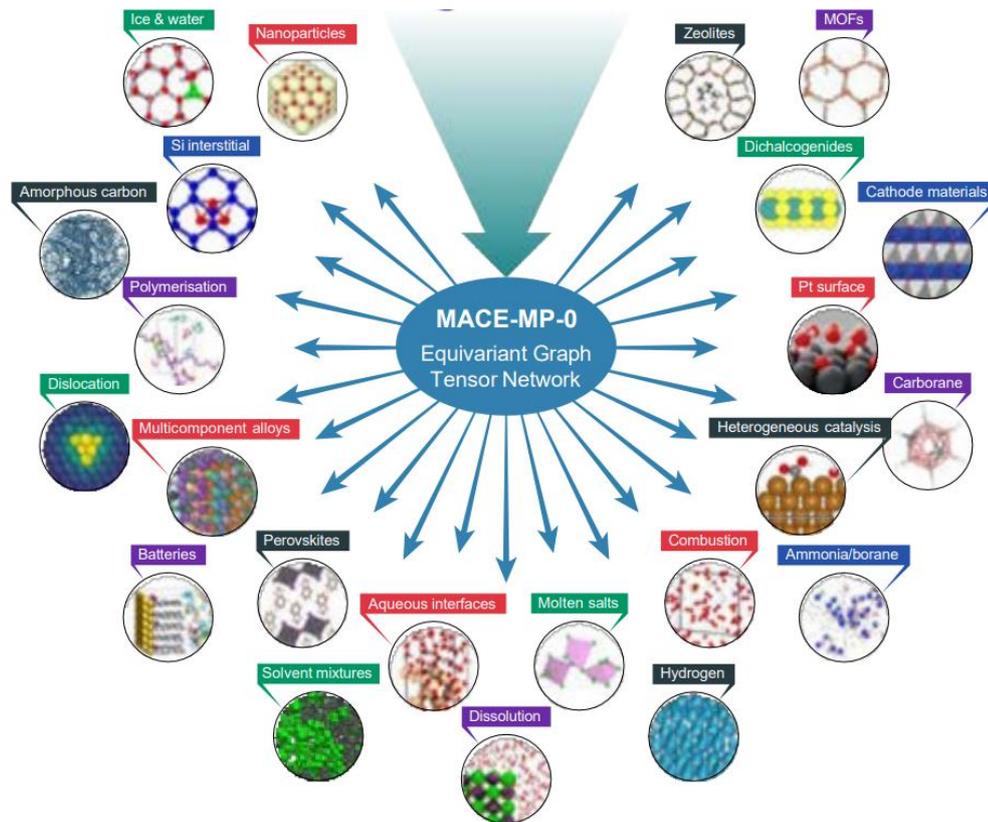
are as accurate as DFT but computationally cheaper than a DFT simulation. This research Demonstrates a general, scalable strategy for building accurate MLPs for flexible MOFs. It provides a practical alternative to neural-network potentials and opens the way to study MOF flexibility, diffusion, adsorption, and reactions at realistic time and length scales, requires fewer training data for the initial dataset and can be trained quickly on a laptop. Performance of this research can be seen in **Figure 6.4**.



**Figure 6.4:** Performance of the trained SNAP model for ZIF-8. A) Learning curves for the RMSE and MAE for energy (left-hand side panel), forces (middle panel) and virial-stress (right-hand side panel). Data are presented for test set A (composed of  $\sim 5000$  configurations from AIMD simulations) and test set B (composed of  $\sim 2000$  configurations from classical MD simulations) as a function of the number of configurations in the training set [5].

In addition to supervised learning frameworks for property prediction, another compelling avenue involves the development and training of machine learning interatomic potentials. In particular, models such as the MACE [6] architecture allow for the construction of highly accurate force fields by training on ab initio datasets, such as those obtained from density functional theory (DFT) calculations. These ML potentials can then be fine-tuned and validated through their application in classical molecular dynamics simulations, enabling a significant reduction in computational cost while retaining quantum-level accuracy. Such approaches hold great potential for accelerating the exploration of complex chemical and material spaces and for bridging the gap between electronic structure methods and large-scale

atomistic simulations. The mace foundation model can be used on a variety of applications as can be seen in **Figure 6.5**.



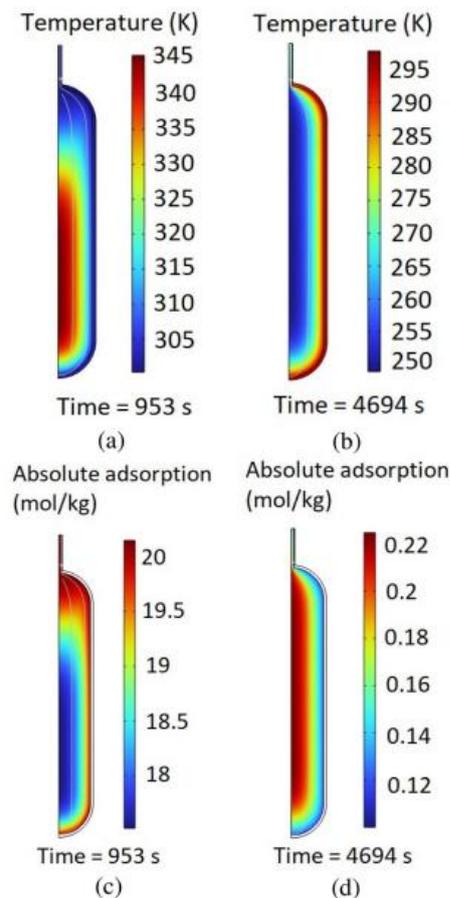
**Figure 6.5:** A foundation model for materials modelling. Trained only on Materials Project data which consists primarily of inorganic crystals and is skewed heavily towards oxides, MACE-MP-0 is capable of molecular dynamics simulation across a wide variety of chemistries in the solid, liquid and gaseous phases [6].

Usually, the MACE architecture is applied via the following workflow: first define the system (materials, composition, and thermodynamic conditions such as temperature and pressure), determine the phase through an appropriate density specification (gas or liquid), generate training and test datasets via DFT calculations, fine-tune the foundation model to optimize performance, and finally run benchmark simulations to assess the accuracy of the fine-tuned model.

### 6.2.3 Continuous mechanics integration

The use of finite element method (FEM) and finite volume method (FVM) techniques to model the macroscale behavior of porous materials presents another, and intriguing, direction for future research. This method provides a potent way to convert complicated atomistic-level data into predictions at the continuum scale that are directly applicable to engineering applications, including local diffusivity, adsorption energetics, molecular

orientation, and mechanical properties. FEM can be used to simulate complex macroscopic scenarios such as mass transfer, deformation under pressure, or coupled thermal and fluid transport in heterogeneous media by integrating outputs from Monte Carlo simulations or molecular dynamics, such as spatially-resolved transport coefficients or stress-strain responses. The investigation of how molecular-level processes collectively affect bulk properties like effective diffusivity, permeability, selectivity, and mechanical robustness is made possible by this multiscale modeling framework, which bridges the gap between length and time scales. When designing and optimizing materials for use in separation technologies, catalytic reactors, or membrane systems—where performance is determined by both microstructural features and operating conditions—such predictive capabilities are especially helpful. Finally, combining FEM/FVM [7] (**Figure 6.6**) and atomistic simulations improves our basic comprehension of material behavior at all scales and offers a strong computational approach to direct experimental work and speed up the development of new materials.



**Figure 6.6:** The obtained contours of the distribution of the temperature and absolute adsorption in the hydrogen tanks by the changes in the time: (a) The contours of temperature at the time steps of 953s, (b) The contours of temperature at the time steps of 4694s, (c) The

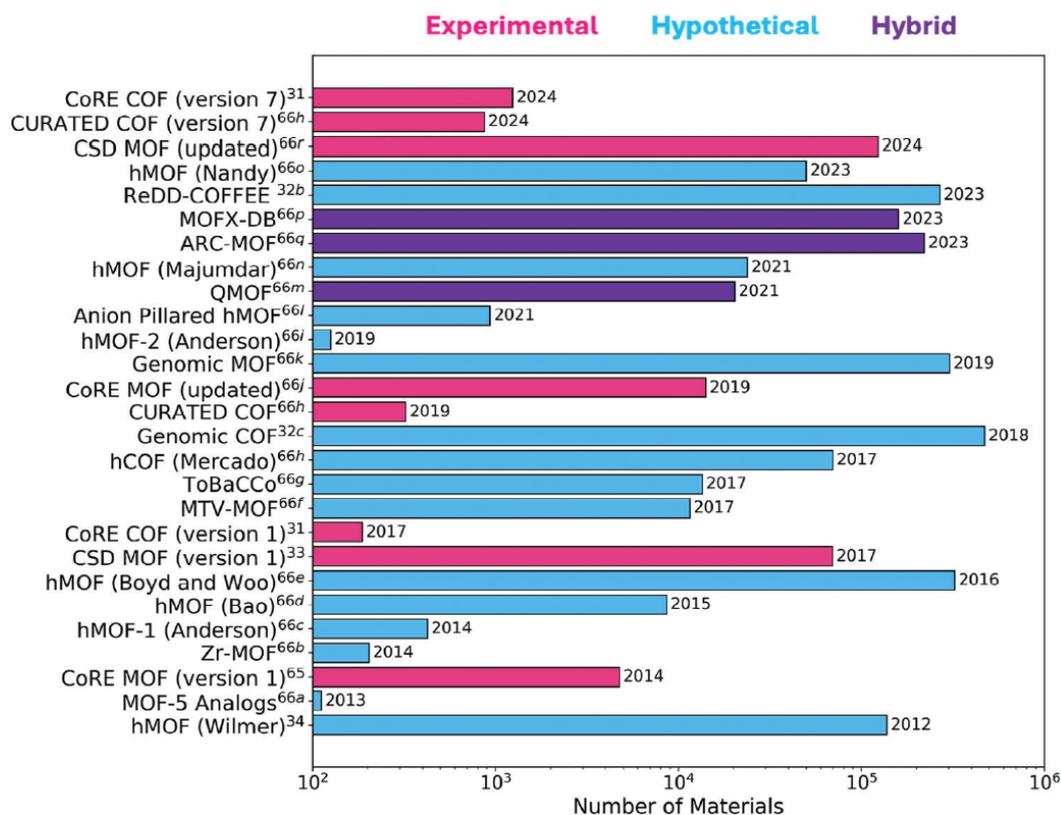
contours of absolute adsorption at the time steps of 953s, (d) The contours of absolute adsorption at the time steps of 4694s [7].

## 6.3 Outlook

This study demonstrated how artificial intelligence, and classical molecular simulations can work together to calculate transport properties in specific metal-organic frameworks and convert those predictions into useful, process-level metrics. By combining data-driven approaches with physics-based models, we were able to obtain accurate estimates of quantities like sorption isotherms/characteristics and self-diffusion coefficients. In addition, it was possible to efficiently map these estimates onto important performance metrics. A thorough yet computationally effective method for screening materials, ranking candidates, and directing further experimental validation is offered by this integrated approach.

MOFs, in particular, and the overall field of porous materials are at a particularly dynamic period especially if we consider that the Nobel prize in Chemistry 2025 was awarded to Susumu Kitagawa, Richard Robson and Omar Yaghi. They have developed a new form of molecular architecture. In their constructions, metal ions function as cornerstones that are linked by long organic (carbon-based) molecules. Together, the metal ions and molecules are organized to form crystals that contain large cavities

Evidently, the rapid synthesis of new frameworks (**Figure 6.7**) continues to accelerate, encompassing not only experimentally realized structures but also a growing library of in silico-generated hypothetical topologies and hybrid architectures that strategically combine multiple chemistries to achieve tailored functionalities [8]. This expanding design space is enabling researchers to explore structure–property relationships with greater precision and to target frameworks optimized for specific applications, ranging from gas storage and separations to catalysis and advanced electronic or optical materials. All new computational workflows can also be applied to enhance the speed in which all possible candidates can be screened.



**Figure 6.7:** List of publicly available MOF and COF databases [8].

Longer timescales are now possible due to a change of the scale and fidelity of simulations of these materials brought about by the expansion of access to high-performance computing. Rapid developments in AI, including surrogate modeling, transfer learning, machine-learned interatomic potentials, and active-learning workflows, are also speeding up discovery by lowering computational costs without sacrificing accuracy. When combined, these advancements close the long-standing gap between experiments and simulations. More reliable predictions are now supported by standardized datasets, principled model validation, and explicit error estimates, while automated benchmarking and AI-assisted inverse design enable focused synthesis and testing. Thus, a strong, scalable, and increasingly essential framework for assessing MOF performance can be utilized by the scientific community.

## References

- [1] A. A. Lysova *et al.*, “Tuning the Molecular and Cationic Affinity in a Series of Multifunctional Metal-Organic Frameworks Based on Dodecanuclear Zn(II) Carboxylate Wheels,” *J Am Chem Soc*, vol. 141, no. 43, pp. 17260–17269, Oct. 2019, doi: 10.1021/jacs.9b08322.
- [2] A. A. Lysova, K. A. Kovalenko, D. N. Dybtsev, S. N. Klyamkin, E. A. Berdonosova, and V. P. Fedin, “Hydrocarbon adsorption in a series of mesoporous metal-organic frameworks,” *Microporous and Mesoporous Materials*, vol. 328, Dec. 2021, doi: 10.1016/j.micromeso.2021.111477.
- [3] A. A. Lysova, K. A. Kovalenko, A. S. Nizovtsev, D. N. Dybtsev, and V. P. Fedin, “Efficient separation of methane, ethane and propane on mesoporous metal-organic frameworks,” *Chemical Engineering Journal*, vol. 453, Feb. 2023, doi: 10.1016/j.cej.2022.139642.
- [4] A. A. Lysova, D. G. Samsonenko, K. A. Kovalenko, A. S. Nizovtsev, D. N. Dybtsev, and V. P. Fedin, “A Series of Mesoporous Metal-Organic Frameworks with Tunable Windows Sizes and Exceptionally High Ethane over Ethylene Adsorption Selectivity,” *Angewandte Chemie - International Edition*, vol. 59, no. 46, pp. 20561–20567, Nov. 2020, doi: 10.1002/anie.202008132.
- [5] A. Sharma and S. Sanvito, “Quantum-accurate machine learning potentials for metal-organic frameworks using temperature driven active learning,” *NPJ Comput Mater*, vol. 10, no. 1, Dec. 2024, doi: 10.1038/s41524-024-01427-y.
- [6] I. Batatia *et al.*, “A foundation model for atomistic materials chemistry,” Mar. 2024, [Online]. Available: <http://arxiv.org/abs/2401.00096>. doi: 10.48550/arXiv.2401.00096
- [7] H. Pourrahmani, M. H. Mohammadi, B. Pourhasani, A. Gharehghani, M. Moghimi, and J. Van herle, “Simulation and optimization of the impacts of metal-organic frameworks on the hydrogen adsorption using computational fluid dynamics and artificial neural networks,” *Sci Rep*, vol. 13, no. 1, Dec. 2023, doi: 10.1038/s41598-023-45391-x.
- [8] H. Daglar, H. C. Gulbalkan, G. O. Aksu, and S. Keskin, “Computational Simulations of Metal–Organic Frameworks to Enhance Adsorption Applications,” 2024, *John Wiley and Sons Inc.* doi: 10.1002/adma.202405532.



# Appendix A

Example of force field in RASPA

# general rule for shifted vs truncated

shifted

# general rule tailcorrections

no

# number of defined interactions

20

# type interaction

Zn1   lennard-jones 0.42 2.7   // D. Dubbeldam, K.S. Walton, D.E. Ellis, R.Q.  
Snurr, Angew. Chem. Int. Ed. 2007, 46, 4496-4499.

O1    lennard-jones 700.0 2.98   // idem

O2    lennard-jones 70.5 3.11   // idem

C1    lennard-jones 47.0 3.74   // idem

C2    lennard-jones 47.86 3.47   // idem

C3    lennard-jones 47.86 3.47   // idem

H1    lennard-jones 7.65 2.85   // idem

He    lennard-jones 10.9 2.64   // J.O. Hirschfelder et al., Molecular Theory of  
Gases and Liquids, Wiley, New York, 1954, p. 1114.

O\_co2   lennard-jones 85.671 3.017   // A. Garcia-Sanchez et al., J. Phys. Chem.  
C 2009, 113, 8814-8820.

C\_co2   lennard-jones 29.933 2.745   // idem

N\_n2    lennard-jones 38.298 3.306   // A. Martin-Calvo et al. , Phys. Chem.  
Chem. Phys. 2011, 13, 11165-11174.

N\_com   none                    // idem

O\_o2    lennard-jones 53.023 3.045   // A. Martin-Calvo et al. , Phys. Chem.  
Chem. Phys. 2011, 13, 11165-11174.

O\_com   none                    //

```

Ar      lennard-jones 124.070 3.38 // A. Martin-Calvo et al. , Phys. Chem. Chem.
Phys. 2011, 13, 11165-11174.

CH4     lennard-jones 158.5 3.72 // M. G. Martin et al., J. Chem. Phys. 2001,
114, 7174-7181.

CH3     lennard-jones 108.0 3.76 // D. Dubbeldam et al., J. Phys. Chem. B,
108(33), 12301-12313

CH2     lennard-jones 56.0 3.96 // idem

CH      lennard-jones 17.0 4.67 // idem

C       lennard-jones 0.8 6.38 // idem

# general mixing rule for Lennard-Jones
Lorentz-Berthelot

```

Example of input script in RASAP2

```

SimulationType      MonteCarlo
NumberOfCycles      25000
NumberOfInitializationCycles 5000
PrintEvery          1000

Forcefield          ExampleZeolitesForceField
RemoveAtomNumberCodeFromLabel yes

Framework 0
FrameworkName MFI_SI
UnitCells 2 2 2
HeliumVoidFraction 0.29
ExternalTemperature 300.0
ExternalPressure 1e4 1e5

```

ComputeNumberOfMoleculesHistogram yes  
WriteNumberOfMoleculesHistogramEvery 5000  
NumberOfMoleculesHistogramSize 1100  
NumberOfMoleculesRange 80

ComputeEnergyHistogram yes  
WriteEnergyHistogramEvery 5000  
EnergyHistogramSize 400  
EnergyHistogramLowerLimit -110000  
EnergyHistogramUpperLimit -20000

Component 0 MoleculeName methane  
MoleculeDefinition ExampleDefinitions  
TranslationProbability 0.5  
ReinsertionProbability 0.5  
SwapProbability 1.0  
CreateNumberOfMolecules 0



## Appendix B

%% K-fold validation: implementation

load fisheriris

X1 = meas;

y1 = categorical(species);

k = 5;

fitFcn1 = @(Xtr,ytr) fitcsvm(Xtr, ytr, 'KernelFunction','rbf','Standardize',true);

predFcn1 = @(mdl,Xte) predict(mdl, Xte);

out1 = kfold\_validate(X1, y1, k, fitFcn1, predFcn1);

fprintf('[Classification] Mean accuracy: %.3f ± %.3f\n', out1.meanScore, out1.stdScore);

% ----- regression-----%

n = 300; D = 5;

X2 = randn(n,D);

true\_w = (1:D)';

y2 = X2\*true\_w + 0.5\*randn(n,1);

k = 10;

fitFcn2 = @(Xtr,ytr) fitrlinear(Xtr, ytr, 'Learner','leastquares','Regularization','ridge');

predFcn2 = @(mdl,Xte) predict(mdl, Xte);

rmse = @(yt,yh) sqrt(mean((yt - yh).^2));

out2 = kfold\_validate(X2, y2, k, fitFcn2, predFcn2, @(yt,yh) -rmse(yt,yh));

fprintf('[Regression] Mean RMSE: %.3f\n', -out2.meanScore);

```

% ===== Implementation =====

function out = kfold_validate(X, y, k, fitFcn, predFcn, metricFcn)
% K-fold cross-validation.
% X: NxD, y: N×1, k: folds
% fitFcn: @(Xtr,ytr)->model
% predFcn: @(model,Xte)->yhat
% metricFcn: @(ytrue,yhat)->scalar (higher is better)

if nargin < 6
    isClass = iscategorical(y) || islogical(y) || numel(unique(y)) < 10;
    if isClass
        metricFcn = @(yt, yh) mean(yt(:) == cast(yh(:), class(yt)));
    else
        metricFcn = @(yt, yh) -sqrt(mean((yt(:) - yh(:)).^2));
    end
else
    isClass = iscategorical(y) || islogical(y) || numel(unique(y)) < 10;
end

if isClass
    cv = cvpartition(y, 'KFold', k);
else
    cv = cvpartition(numel(y), 'KFold', k);
end

scores = zeros(k,1);
indices = struct('train', cell(k,1), 'test', cell(k,1));
preds = cell(k,1);

```

```
for i = 1:k
    tr = training(cv, i);
    te = test(cv, i);

    mdl = fitFcn(X(tr,:), y(tr));
    yhat = predFcn(mdl, X(te,:));
    preds{i} = yhat;

    scores(i) = metricFcn(y(te), yhat);
    indices(i).train = find(tr);
    indices(i).test = find(te);
end

out.scores = scores;
out.meanScore = mean(scores);
out.stdScore = std(scores);
out.preds = preds;
out.indices = indices;
out.cv = cv;
end
```



# Appendix C

Sorption, diffusion ANNs models

The sorption and diffusion models (necessitate MATLAB 2017b or later) along with their parameters can be found in the following github repository:  
[https://github.com/lgergidi/ANNsZIF8\\_22.git](https://github.com/lgergidi/ANNsZIF8_22.git)

Sorption ANN model parameters

```
function [Y,Xf,Af] = myNeuralNetworkFunction(X,~,~)
%MYNEURALNETWORKFUNCTION neural network simulation function.
%
% Generated by Neural Network Toolbox function genFunction,
%
% [Y] = myNeuralNetworkFunction(X,~,~) takes these arguments:
%
% X = 1xTS cell, 1 inputs over TS timesteps
% Each X{1,ts} = 3xQ matrix, input #1 at timestep ts.
%
% and returns:
% Y = 1xTS cell of 1 outputs over TS timesteps.
% Each Y{1,ts} = 1xQ matrix, output #1 at timestep ts.
%
% where Q is number of samples (or series) and TS is the number of timesteps.
%#ok<*RPMT0>
% ===== NEURAL NETWORK CONSTANTS =====
% Input 1
x1_step1.xoffset = [16.04;273;0.01];
x1_step1.gain = [0.0475285171102662;0.02;0.02000200020002];
x1_step1.ymin = -1;
```

```

% Layer 1

b1 = [-
0.26190952842227410713;0.07445866754430043033;6.3299435088749946843];

IW1_1 = [0.8152138028487788679 -0.085230331892812108419
0.18919735772240178773;0.30306734048520966862 -0.087069964422032103402
0.10253287737659551315;0.83225911003602748295 -0.1088156432252801592
6.7495764548917644987];

% Layer 2

b2 = [-2.0773690968695488479;-1.1471915027775787532;1.9245717467179490612];

LW2_1 = [-13.530160716856393321 6.0541831440064459713
11.485931414020351227;-2.3638843386926793322 6.0301419927547454947
0.47123351474552244156;2.5165421990286915843 -7.1261269543425198592 -
0.68508417302842083974];

% Layer 3

b3 = [1.904208205899081463;-0.81123955649487977659;2.3023356897495648177];

LW3_2 = [2.2023000238229406555 -6.8257774127769001993 -
2.8269153810853797637;0.21903031917332205225 1.8044397205630564596
1.2938236832811376154;2.0143462529436679809 -5.7268651237070704596 -
2.4882317499548292083];

% Layer 4

b4 = -5.0293971223908835455;

LW4_3 = [-3.5518494601463177496 1.2688942682178039334
8.7334881300699827023];

% Output 1

y1_step1.ymin = -1;

y1_step1.gain = 0.0900737337955242;

y1_step1.xoffset = 0.0027375;

% ===== SIMULATION =====

% Format Input Arguments

isCellX = iscell(X);

```

```

if ~isCellX
X = {X};
end
% Dimensions
TS = size(X,2); % timesteps
if ~isempty(X)
Q = size(X{1},2); % samples/series
else
Q = 0;
end
% Allocate Outputs
Y = cell(1,TS);
% Time loop
for ts=1:TS
% Input 1
Xp1 = mapminmax_apply(X{1,ts},x1_step1);

% Layer 1
a1 = tansig_apply(repmat(b1,1,Q) + IW1_1*Xp1);

% Layer 2
a2 = tansig_apply(repmat(b2,1,Q) + LW2_1*a1);

% Layer 3
a3 = tansig_apply(repmat(b3,1,Q) + LW3_2*a2);

% Layer 4
a4 = repmat(b4,1,Q) + LW4_3*a3;

```

```

% Output 1
Y{1,ts} = mapminmax_reverse(a4,y1_step1);
end
% Final Delay States
Xf = cell(1,0);
Af = cell(4,0);
% Format Output Arguments
if ~isCellX
    Y = cell2mat(Y);
end
end
% ===== MODULE FUNCTIONS =====
% Map Minimum and Maximum Input Processing Function
function y = mapminmax_apply(x,settings)
    y = bsxfun(@minus,x,settings.xoffset);
    y = bsxfun(@times,y,settings.gain);
    y = bsxfun(@plus,y,settings.ymin);
end
% Sigmoid Symmetric Transfer Function
function a = tansig_apply(n,~)
    a = 2 ./ (1 + exp(-2*n)) - 1;
end
% Map Minimum and Maximum Output Reverse-Processing Function
function x = mapminmax_reverse(y,settings)
    x = bsxfun(@minus,y,settings.ymin);
    x = bsxfun(@rdivide,x,settings.gain);
    x = bsxfun(@plus,x,settings.xoffset);

```

```

end

Diffusion ANN model parameters

function [y1] = myNeuralNetworkFunction(x1)

%MYNEURALNETWORKFUNCTION neural network simulation function.

%

% Generated by Neural Network Toolbox function genFunction,

%

% [y1] = myNeuralNetworkFunction(x1) takes these arguments:

% x = 3xQ matrix, input #1

% and returns:

% y = 1xQ matrix, output #1

% where Q is the number of samples.

%#ok<*RPMT0>

% ===== NEURAL NETWORK CONSTANTS =====

% Input 1

x1_step1.keep = [1 2];

x1_step2.xoffset = [16.04;273];

x1_step2.gain = [0.0475285171102662;0.02];

x1_step2.ymin = -1;

% Layer 1

b1 = [-2.4248711305964278928;0;-2.4248711305964278928];

IW1_1 = [1.8562135078827242118 -1.560279274089645396;1.5286930660659518999
1.8823117461679610418;-1.413995089584613174 1.9699283963206886128];

% Layer 2

b2 = [2.0191493984303714448;0;-2.0191493984303714448];

LW2_1 = [-1.1347286731301977625 1.6318265055225207316
0.35566471209844552348;-1.1620618903342725403 1.375497334813064354 -
0.91355543791092652661;-
1.3050159087562120952 0.51538664878635154221 1.4519897979420182121];

```

```

% Layer 3
b3 = [-2.0191493984303714448;0;-2.0191493984303714448];
LW3_2 = [1.1918744327641783709 -0.13814346072545208299
1.6239815312877159403;-0.25194910484058130606 -0.95308987156819502395
1.7621309935603204355;-
0.35279487144938437293 1.1503619332987566892 1.6214707811977033813];

% Layer 4
b4 = -0.75299833536039884052;
LW4_3 = [-0.47897698583290782004 0.37927568014821799736 -
0.7363386690489752695];

% Output 1
y1_step1.ymin = -1;
y1_step1.gain = 3739925575.48105;
y1_step1.xoffset = 1.23e-12;

% ===== SIMULATION =====

% Dimensions
Q = size(x1,2); % samples

% Input 1
xp1 = removeconstantrows_apply(x1,x1_step1);
xp1 = mapminmax_apply(xp1,x1_step2);

% Layer 1
a1 = relu1_apply(repmat(b1,1,Q) + IW1_1*xp1);

% Layer 2
a2 = relu1_apply(repmat(b2,1,Q) + LW2_1*a1);

% Layer 3
a3 = relu1_apply(repmat(b3,1,Q) + LW3_2*a2);

% Layer 4
a4 = repmat(b4,1,Q) + LW4_3*a3;

% Output 1

```

```

y1 = mapminmax_reverse(a4,y1_step1);
end
% ===== MODULE FUNCTIONS =====
% Map Minimum and Maximum Input Processing Function
function y = mapminmax_apply(x,settings)
y = bsxfun(@minus,x,settings.xoffset);
y = bsxfun(@times,y,settings.gain);
y = bsxfun(@plus,y,settings.ymin);
end
% Remove Constants Input Processing Function
function y = removeconstantrows_apply(x,settings)
y = x(settings.keep,:);
end
% Sigmoid Symmetric Transfer Function
function a = relu1_apply(n,~)

end
% Map Minimum and Maximum Output Reverse-Processing Function
function x = mapminmax_reverse(y,settings)
x = bsxfun(@minus,y,settings.ymin);
x = bsxfun(@rdivide,x,settings.gain);
x = bsxfun(@plus,x,settings.xoffset);
end

```



## Appendix D

```
function [Y,Xf,Af] = neural_function(X,~,~)
%NEURAL_FUNCTION neural network simulation function.
%
% Generated by Neural Network Toolbox function genFunction, 11-Feb-2024 18:18:56.
%
% [Y] = neural_function(X,~,~) takes these arguments:
%
% X = 1xTS cell, 1 inputs over TS timesteps
% Each X{1,ts} = 2xQ matrix, input #1 at timestep ts.
%
% and returns:
% Y = 1xTS cell of 1 outputs over TS timesteps.
% Each Y{1,ts} = 1xQ matrix, output #1 at timestep ts.
%
% where Q is number of samples (or series) and TS is the number of timesteps.
%#ok<*RPMT0>
% ===== NEURAL NETWORK CONSTANTS =====
% Input 1
x1_step1.xoffset = [273;1];
x1_step1.gain = [0.02;0.002002002002002];
x1_step1.ymin = -1;
% Layer 1
b1 = [0.29372403915256645179;1.168963935211287497;-0.63043411593191966968];
IW1_1 = [0.64595314279793702994 -
0.13577401717141457382;0.51370859289760351096
0.54567330850816675447;0.67468342698607708385 -0.76707892294666546462]
% Layer 2
b2 = [0.26686979744083066945;-
0.076899921495714868014;0.93984531382767866337];
LW2_1 = [-0.046069688994832144746 -1.5913991916304348617
0.68249805678155350819;1.657016980344600432 -1.9686046168582362892 -
```

```
1.0094375606056764649;0.92524242214932683748 -0.45915679357184491227  
0.59654493424610388619];
```

```
% Layer 3
```

```
b3 = [0.23715647737689560248;-  
0.77833200084014342313;0.2464710255355543056];
```

```
LW3_2 = [0.78941508173748964339 0.011484096762404357589  
0.70543594961297606538;1.7731106774893432565 2.5131542685953247585  
1.525152788596395359;-0.49137938974157158167 -0.93095437171452932734  
2.331145683121673251]
```

```
% Layer 4
```

```
b4 = -0.16392364421578986544;
```

```
LW4_3 = [-1.0303964359961819142 -1.4049772492216947573 -  
1.7715889268308566695];
```

```
% Output 1
```

```
y1_step1.ymin = -1;
```

```
y1_step1.gain = 1.27958659157346;
```

```
y1_step1.xoffset = 0.000524244;
```

```
% ===== SIMULATION =====
```

```
% Format Input Arguments
```

```
isCellX = iscell(X);
```

```
if ~isCellX
```

```
    X = {X};
```

```
end
```

```
% Dimensions
```

```
TS = size(X,2); % timesteps
```

```
if ~isempty(X)
```

```
    Q = size(X{1},2); % samples/series
```

```
else
```

```
    Q = 0;
```

```
end
```

```
% Allocate Outputs
```

```
Y = cell(1,TS);
```

```

% Time loop
for ts=1:TS
    % Input 1
    Xp1 = mapminmax_apply(X{1,ts},x1_step1);
    % Layer 1
    a1 = tansig_apply(repmat(b1,1,Q) + IW1_1*Xp1);
    % Layer 2
    a2 = tansig_apply(repmat(b2,1,Q) + LW2_1*a1);
    % Layer 3
    a3 = tansig_apply(repmat(b3,1,Q) + LW3_2*a2);
    % Layer 4
    a4 = repmat(b4,1,Q) + LW4_3*a3;
    % Output 1
    Y{1,ts} = mapminmax_reverse(a4,y1_step1);
end
% Final Delay States
Xf = cell(1,0);
Af = cell(4,0);
% Format Output Arguments
if ~isCellX
    Y = cell2mat(Y);
end
end
% ===== MODULE FUNCTIONS =====
% Map Minimum and Maximum Input Processing Function
function y = mapminmax_apply(x,settings)
    y = bsxfun(@minus,x,settings.xoffset);
    y = bsxfun(@times,y,settings.gain);
    y = bsxfun(@plus,y,settings.ymin);
end
% Sigmoid Symmetric Transfer Function

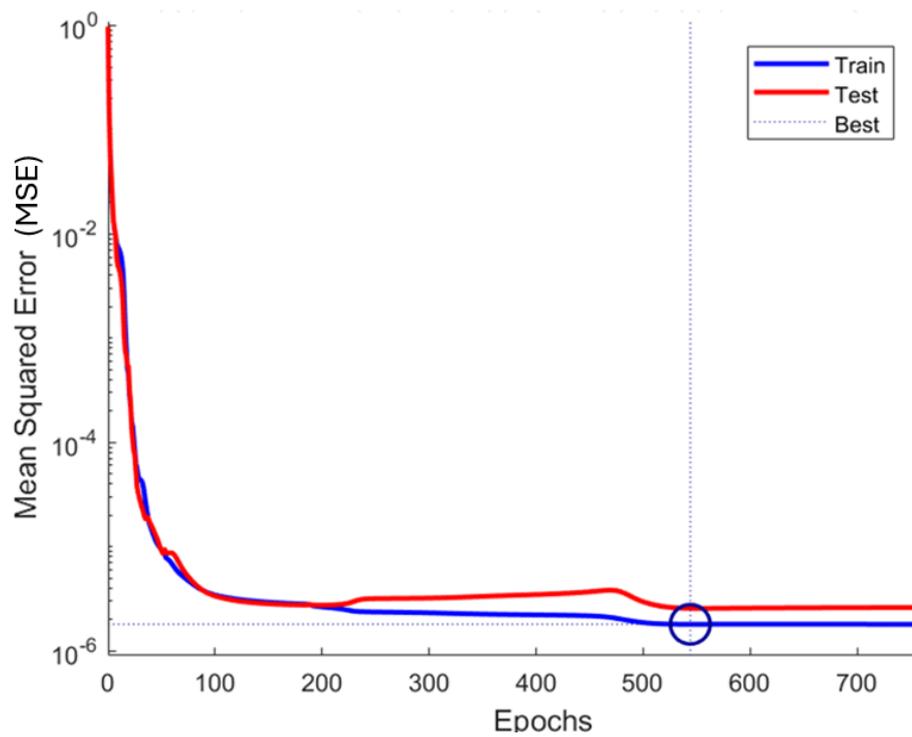
```

```

function a = sigmoid_apply(n,~)
    a = 2 ./ (1 + exp(-2*n)) - 1;
end
% Map Minimum and Maximum Output Reverse-Processing Function
function x = mapminmax_reverse(y,settings)
    x = bsxfun(@minus,y,settings.ymin);
    x = bsxfun(@rdivide,x,settings.gain);
    x = bsxfun(@plus,x,settings.xoffset);
end

```

**Best Train Performance at Epoch 544, MSE = 1,7992 10<sup>-6</sup>**



**Figure:** Training and Testing performance measured by MSE

# Curriculum vitae (CV)

Gkourras Arsenios

[LinkedIn Profile](#)  
[Github Repository](#)

Email: [arseniosgkk@gmail.com](mailto:arseniosgkk@gmail.com)  
Mobile: +306984739125  
Address: Duivelsers 7, Halle Belgium 1500

## SKILLS SUMMARY

---

- **Technical Skills:** FEA/FVM/CFD Simulations, Molecular Dynamics, Monte Carlo Simulations, Simulation Tools, Gas Separation, Porous Materials, Thermodynamics, Thermal Simulation, CAD Tools, Mechanical Design, Machine Learning (AI/ML), Digital Image Processing, Data Science, R&D
- **Programming Languages:** Python, PHP, C/C++, C#, JAVA, MATLAB, R
- **Programming Frameworks:** Scikit, OPENCV, TensorFlow, Keras, OpenCV, SciPy, Pandas, Tkinter
- **Software Packages:** AutoCAD, SolidWorks, Inventor, Catia ANSYS, STAR CCM+, RASPA2, Avogadro, Comsol, LAAMPS, OPENFoam
- **Platforms:** Linux, Windows, Arduino, Raspberry
- **Soft Skills:** Leadership, Problem Solving, Public Speaking, Multi-disciplinary, Adaptable, Self-motivated, Communicative
- **Languages:** English(Fluent), Greek (Native) Dutch (Intermediate)

## EDUCATION

---

- **University of Ioannina Department of Materials Science and Engineering** Ioannina, Greece  
*PhD* 11/2019 -  
*Doctoral Thesis Title: Modelling of sorption and diffusion of alkanes, alkenes in porous materials using molecular simulations and artificial neural networks*
- **University of Ioannina Department of Materials Science and Engineering** Ioannina, Greece  
*Master of Science - Advanced Materials* 11/2017 - 11/2019  
*Master Thesis Title: Sorption and diffusion study of small alkanes-alkenes and their mixtures in zeolites: From molecular simulations to neural networks.*
- **Technological Institute of Thessaly Department of Mechanical Engineering** Larissa, Greece  
*Bachelor Degree - Energy Department* 09/2013 - 09/ 2017  
*Bachelor Thesis Title: Investigation of the capabilities of OpenFOAM software in internal flow simulations*

## EXPERIENCE

---

- **SCREEN Semiconductor solutions** Belgium  
*Scientific reseacher-Simulations Engineer (Full-time)* 03/2024 -
  - **Molecular Modelling:** Employed state-of-the-art computational methodologies such as Machine learning interatomic potentials on complicated narrow-space semiconductor structures.
  - **Kinetic Monte Carlo:** Simulated etching rates with atomic scale accuracy in realistic process time scales.
  - **FEM Modelling:** Applied Finite Element modelling for post CMP process
  - **Experimental validation:** Conducted experiments on various process(wafer cleaning, post CMP, Defect control) to verify and validate simulation results
- **PV-Auto-Scout** Ioannina  
*Mechanical Engineer- Lead Software Developer (Full-time)* 09/2019 - 06/2022
  - **Characterization of photovoltaic modules using image processing and infrared thermography:** Established a robust methodology using image processing and infrared thermography for the comprehensive assessment of photovoltaic modules.
  - **Software development for the identification of defective PV modules:** Developed Python and OpenCV-based software for efficient identification of defective PV modules, integrating seamlessly into the software-hardware environment.
  - **Neural Network Modelling:** Successfully applied Neural Network techniques for training and testing the image processing methodology.
  - **Algorithm evaluation report:** Led on-field tests with cross-functional teams to evaluate image processing results, achieving a correlation success rate exceeding 90%. Followed a rigorous design, implementation, and verification procedure.
- **Materials I Laboratory** Ioannina  
*Teacher Assistant (Part-time)* 06/2019 - 06/2023
  - **IR/UVSpectroscopy:** Guided pre-graduate students through a comprehensive IR and UV Spectroscopy exercise, fostering a deep understanding of molecular absorption principles and techniques, empowering them to confidently apply their knowledge in real-world scenarios.

**Internship at the Laboratory of Heat Transfer**  
*Laboratory Personnel (Full Time)*

Larissa  
04/2017 - 09/2017

- **Research:** Engaged in lab projects with a strong RD team and Application Scientists with a computational simulation emphasis, particularly in the domain of fluid mechanics simulations for various case studies. Supported students in the ongoing experiments, facilitated data uploads from different experimental groups, and developed midterm mini-tests for the class.
- **Managed Academic Laboratory Equipment:** Oversaw the setup and preparation of experimental equipment for each academic exercise conducted during the semester.

**Internship at the Laboratory of Renewable Energy Sources**  
*Laboratory Personnel (Part-time)*

Larissa  
06/2016 - 02/2017

- **Research with MSc and PhD students:** I collaborated closely with fellow MSc and PhD students-researchers, offering valuable assistance to their respective projects
- **Organize Laboratory Equipment:** I meticulously prepared the experimental equipment utilized for each experiment conducted throughout the entire semester

**PROJECTS**

---

- **Thermansys - Thermal Imaging Software :** (Work in progress) Led design and development of an advanced thermal imaging software for comprehensive inspection of composite materials, utilizing state-of-the-art algorithms with scientific computing, including lock-in pulsed phased thermography. Enabled real-time analysis, accurate defect detection, and high-resolution visualization. Utilized Python and non-destructive testing tools.
- **PV-AUTO SCOUT Computer Vision:** AI-driven software using thermal imaging for precise fault detection in PV modules. Offers real-time analysis, remote accessibility, and comprehensive reporting for optimized maintenance strategies and system performance. Utilizes Python, Image processing, Keras, Tensorflow, Non-Destructive Testing, and OpenCV.
- **Multidisciplinary Multiscale Multiphysics Simulation Team Website (Web Development):** Created and still maintain the following website (<http://www.materials.uoi.gr/simlab/>) Tools: PHP, HTML, CSS
- **Calculus Calculator:** Developed code during my Ph.D. Code can be used to solve derivatives, limits, integrals, ODEs and PDEs

**PUBLICATIONS**

---

- **Article:** A. Gkourras, D. Iliopoulos, L. N. Gergidis, D. G. Samsonenko, and V. P. Fedin, " Adsorption of Methane, Ethane and Their Equimolar Mixture in NiIC-20-Bu Metal-Organic Framework from Grand Canonical Monte Carlo Simulations and Artificial Neural Networks." Adv. Theory Simul. (2025).
- **Article:** A. Gkourras, L. N. Gergidis, Molecular and artificial neural networks modeling of sorption and diffusion of small alkanes, alkenes and their ternary mixtures in ZIF-8 at different temperatures, The Journal of Physical Chemistry B 126 (29) (2022) 5582-5594.
- **Article:** Angelos, Ntaflou, Arsenios Gkourras, L. N. Gergidis, Alkiviadis S. Paipetis, Comparison of pulsed, pulse-phase, lock in and standard deviation thermography techniques for nondestructive evaluation of composite materials., Composites Meet Sustainability – Proceedings of the 20th European Conference on Composite Materials, ECCM20. 26-30 June, 2022, Lausanne, Switzerland
- **Article:** Gurras, A. and Gergidis, L.N. (2021), Modeling Sorption and Diffusion of Alkanes, Alkenes, and their Mixtures in Silicalite: From MD and GCMC Molecular Simulations to Artificial Neural Networks. Adv. Theory Simul. 2000210
- **Article:** Gurras, A., Gergidis, L., Mytafides, C., Tzounis, L., & Paipetis, A. S. (2021). Automated detection-classification of defects on photo-voltaic modules assisted by thermal drone inspection. In MATEC Web of Conferences (Vol. 349, p. 03015). EDP Sciences.

**CONFERENCES**

---

- Oral Presentation on the EMRS Fall Meeting 0 September, 2025
- Attended The IRASPA /RASPA workshop - 03 February, 2023
- Oral and Paper presentation European Conference on Composite Materials (ECCM20). June 26-30, 2022
- Poster Presentation for the 13th Panhellenic Scientific Chemical Engineering Conference. 2-4 June, 2022
- Attended The IRASPA /RASPA workshop - 20 January, 2022
- Attended the Online workshop on Computational Materials Science - 18 December, 2021
- Oral and Paper presentation International Conference of Engineering Against Failure (ICEAF VI). June 23-25, 2021

**CERTIFICATIONS**

---

- The Complete Python Developer Course (Udemy)
- The Complete Android Oreo Developer Course (Udemy)

**VOLUNTEER WORK**

---

- **Trainer's assistant: Elite Fight Club Ioannina** Ioannina, Greece  
Conducted Muay Thai and Kick-Boxing classes for children ages (8-14). 01/2018 - 09/2023