Charles universitry

Faculty of Science

Study programme: Chemistry

Branch of study: Chemistry

Daniel Willimetz

Theoretical investigation of ^{27}Al chemical shifts dependence on water amount and temperature in zeolite MFI

Teoretické studium závislosti ^{27}Al chemických posunů na množství vody v kanálech a na teplotě v zeolitu MFI

Bachelor's thesis

Supervisor: **RNDr. Lukáš Grajciar, Ph. D.**
Consultant: **Dr. rer. nat. Andreas Erlebach**

Prague, 2023

Prohlašuji, že jsem závěrečnou práci zpracoval samostatně a že jsem uvedl všechny použité informační zdroje a literaturu. Tato práce ani její podstatná část nebyla předložena k získání jiného nebo stejného akademického titulu.

Praha, 17. 05. 2023   Daniel Willimetz

Abstract:

This bachelor thesis explores ^{27}Al NMR spectra in zeolite MFI and investigates the impact of several factors on the chemical shielding values, including water loading, temperature, and the relative positions of aluminum pairs. Various machine learning-based methods for calculating chemical shift are evaluated. Molecular dynamics simulations with neural network potentials are used to simulate experimental conditions. Neural network potentials offer a highly efficient means of calculating energy with a significant speed-up of approximately 1000 times faster than density functional theory, while maintaining a high level of accuracy. This study is the first to examine ^{27}Al NMR under *operando* conditions, with a focus on the experimentally relevant amount of water.

Keywords:

^{27}Al NMR, machine learning, zeolites, MFI, *operando* conditions, neural network potentials

Abstrakt:

Tato bakalářská práce zkoumá ^{27}Al NMR spektra v zeolitu MFI a zkoumá vliv několika faktorů na chemické stínění, včetně obsahu vody, teploty a relativních poloh hliníkových párů. Hodnoceny jsou různé metody výpočtu chemického posunu založené na strojovém učení. Molekulové dynamické simulace s využitím neuronových sítí jsou použity k simulaci experimentálních podmínek. Neuronové sítě nabízejí vysokou účinnost výpočtu energie s významným zrychlením, asi 1000krát rychlejší než teorie funkcionálu hustoty, při zachování vysoké úrovně přesnosti. Tato studie je první, která zkoumá ^{27}Al NMR za *operando* podmínek, s důrazem na experimentálně relevantní množství vody.

Klíčová slova:

^{27}Al NMR, strojové učení, zeolity, MFI, *operando* podmínky, potenciály založen=neuronových sítí

Abbreviations

2(5)p – Two(five)-parameter equation for calculating shielding

ASE – Atomic simulation environment

BAS – Brønsted acid site

DFT – Density functional theory

HF – Harthree–Fock

IZA – International zeolite database

KRR – Kernel ridge regression

MD – Molecular dynamics

MFI – Mordenite framework inverted

ML - Machine learning

NMR – Nuclear magnetic resonance

(N)NNN – (Next-)next-nearest neighbor

NNP – Neural network potentials

PES – Potential energy surface

UC – Unit cell

ZSM-5 – Zeolite Socony Mobil–5

Contents

1 Introduction and Motivation

Zeolites are highly useful microporous aluminosilicate structures employed in diverse industrial applications, such as catalysis, adsorption, or separation. ZSM-5 is one of the most widely utilized zeolites in the petrochemical industry, serving as a catalyst for hydrocarbon transformations. The catalytic activity of zeolites is primarily determined by their chemical composition, with structural characteristics such as the shape of the channels, the accessible pore volume, or the position of aluminum atoms within the framework also contributing to their catalytic properties. Specifically, the chemical composition of a zeolite plays a vital role in determining its suitability for different types of catalysis. The substitution of silicon atoms with aluminum atoms in the framework generates a negative charge that must be compensated by cations, including H_3O^+, Co^{2+}, or Na^+. The type of compensating ion can also influence the properties of the zeolite.[1]

The position of aluminum atoms within the zeolite framework plays a crucial role in determining the catalytic properties, therefore it is essential to identify a reliable method for assigning aluminum atoms to specific sites within the structure. X-ray crystallography is unable to effectively distinguish between aluminum and silicon atoms, which limits its usefulness for this purpose. Several studies have proposed the use of nuclear magnetic resonance (NMR) as a viable approach for aluminum siting and structure determination.[2–5]

Solid-state NMR spectra are notoriously difficult to interpret due to broadened signals and significant signal overlap. To gain a better understanding of these spectra, theoretical calculations of NMR parameters are necessary. Theoretical studies using first-principles calculations and molecular dynamics (MD) simulations can provide valuable insights into the behavior of zeolites. Due to the large size of unit cells in zeolites, theoretical electronic structure calculations are computationally demanding. The widely used density functional theory (DFT) method is routinely used for single structures or for very short dynamical simulations but may not be adequate for studying dynamic effects properly. A cost-efficient alternative to DFT are machine learning (ML) techniques such as neural network potentials (NNP), which can significantly accelerate not only dynamical calculations but also property predictions, such as NMR shielding values.[6,7]

The study of ^{27}Al NMR spectroscopy can provide detailed information about the local

environment of aluminum atoms in the zeolite framework, which is useful for understanding zeolite properties such as adsorption or catalysis. Multiple studies have attempted to assign experimental spectra to specific positions in the framework.[4,5,8] Figure 1 indicates that there is inconsistency in the calculated chemical shifts among the three studies, which aimed to assign the experimental ^{27}Al NMR to specific T-sites in ZSM-5.
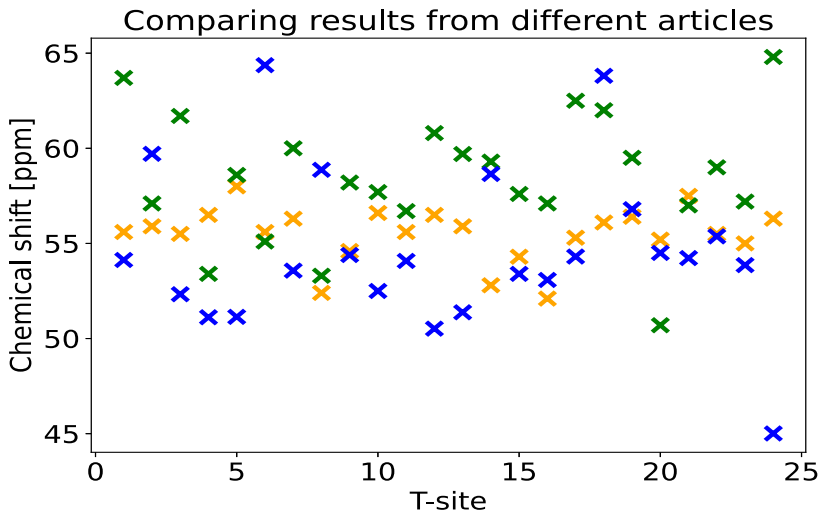

Figure 1: Calculated chemical shifts of 24 T-sites in ZSM-5 reported by three studies with different calculation methods. Blue points are data reported by Dib et al.,[8] orange from Holzinger et al.[5] and green from Sklenak et al.[4]

Standard DFT computations offer valuable information about the interpretation of spectra. However, they rely on approximate models, which are constructed based on numerous assumptions, such as neglecting the presence of water molecules or charge compensating cation. Therefore the experimental conditions are not fully replicated. The present thesis aims to evaluate the feasibility of assigning T-sites to their corresponding peaks in experimental ^{27}Al NMR spectra and to also address the mentioned limitations by testing several parameters that are commonly overlooked in the theoretical calculations of ^{27}Al chemical shifts. Specifically, the effects of water loading, temperature, and the presence of Al pairs are examined to determine their impact. This is achieved by utilizing NNP and other ML derivated techniques allowing highly efficient MD simulations at *operando* conditions.

2 Theoretical Introduction

2.1 Zeolites

Zeolites are crystalline structures that consist of tetrahedral units TO_4, where the T-atom can be either silicon or another heteroatom such as gallium, titanium, or more commonly, aluminum. This thesis considers only the aluminum atoms as the heteroatoms in the lattice. Positions, where either silicon or aluminum atoms are located, are called T-sites and are crystallographically distinguishable. The number of T-sites varies depending on the type of zeolite framework. The position of aluminum atoms in a zeolite structure can have an impact on its properties, which has been demonstrated experimentally, such as in the case of the methanol-to-olefins reaction in ZSM-5.[9] Zeolites are composed of TO_4 tetrahedra that form cavities and channels with diameters up to 20 Å. These channels can have different shapes and sizes, which make zeolites excellent for shape-selective catalysis.[1] The International Zeolite Association (IZA) database lists a total of 248 distinct zeolite structure types that have been identified and characterized so far.[10]

In zeolites, the introduction of aluminum atoms creates a tetrahedral AlO_4^- unit, which possesses a negative charge that must be compensated by a cation. When this cation is a proton (H^+), it can form an acidic group Si-(OH)-Al known as a Brønsted acid site (BAS). Such a BAS is depicted in Figure 2 and such zeolites are termed acid(ic) zeolites. When water molecules are present within the zeolite structure, the proton that is bound to oxygen may dissociate from the framework and generate an H_3O^+ cation. This acidic property plays a significant role in various industrial applications of zeolites.[1]

The number of aluminum atoms in the framework is limited by the Löwenstein rule,[11] which states that the two aluminum atoms must be separated at least by one silicon atom. As a result, the minimum Si/Al ratio in zeolites is equal to 1. However, there is no defined upper limit for the Si/Al ratio and it can theoretically be infinitely high.

2.1.1 MFI

The IZA database provides a three-letter abbreviation for each zeolite framework.[10] The MFI type framework is widely used in various industries, such as petrochemistry, gas
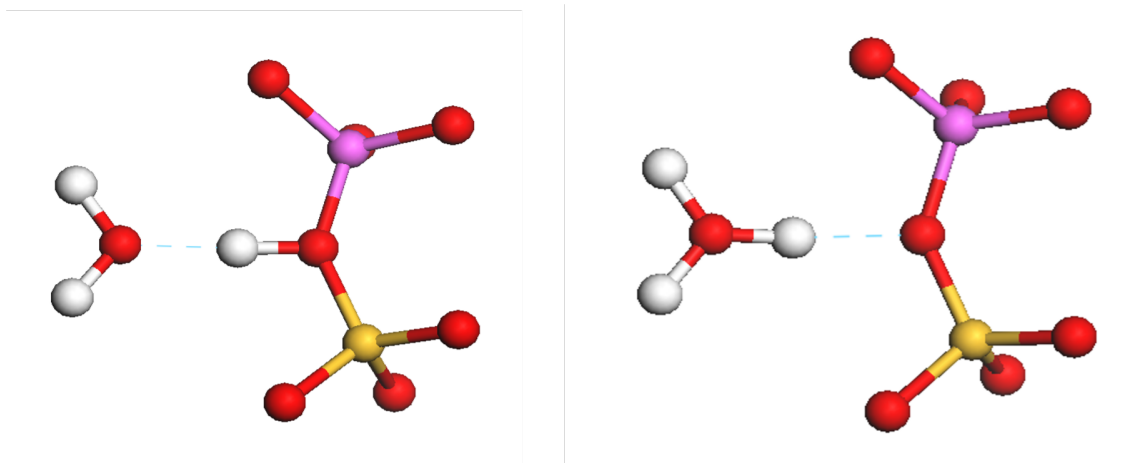
Figure 2: BAS in two states – not solvated proton (left) and solvated proton creating H_3O^+ (right). Oxygen is labeled with red color, aluminum with purple, silicon with yellow and hydrogen with white.

purification, and the reduction of NO_x. ZSM-5, an MFI type zeolite, is widely regarded as one of the most useful zeolites due to its capability to isomerize hydrocarbons in oil processing.[1] The MFI framework is considered to be one of the most intricate types of zeolite structures made up of pentasil units and containing at least 12 distinct T-sites. The framework is formed through the arrangement of these units into a pentasil chain, as illustrated in Figure 3.[1] MFI can adopt a structure with either an orthorhombic or monoclinic symmetry, which determines the number of distinct T-sites. All T-sites depending on cell symmetry are displayed in Figure 4.

Figure 3: Building units of MFI framework type – pentasil unit (left) and pentasil chain (right).[1]

Figure 5 illustrates the complete MFI framework which consists of multiple channels, including the largest one with a circumference of 10 T-atoms. The unit cell of MFI has
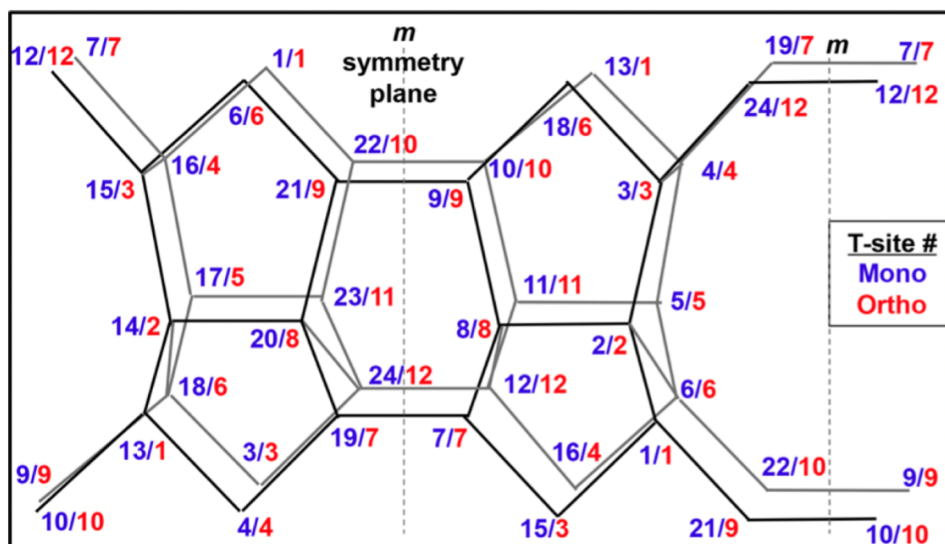
Figure 4: T-site labels located in MFI at orthorhombic (red) or monoclinic (blue) cell.[12]

an accessible volume of almost 10 %,[10] which can be utilized to calculate the optimal number of water molecules that can fit in its channels. Based on the assumption that the density of water is $1 \text{ g} \cdot \text{cm}^{-3}$, the optimal water loading for MFI zeolite would be 17 water molecules per unit cell, given that the total volume of the MFI zeolite cell is approximately 5300 Å^3.

Figure 5: MFI framework displayed from two different angles.[1,10]

2.2 NMR

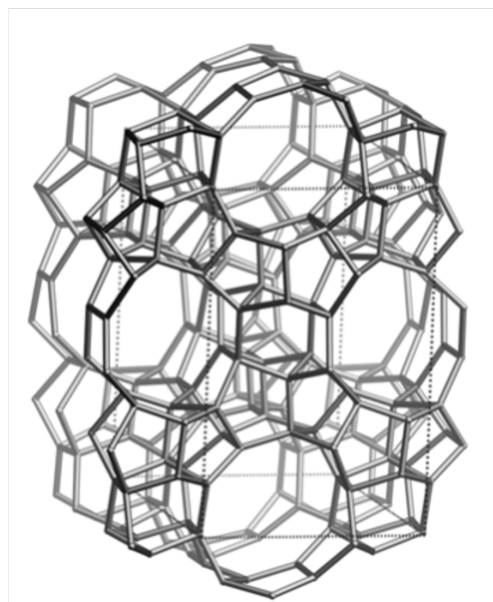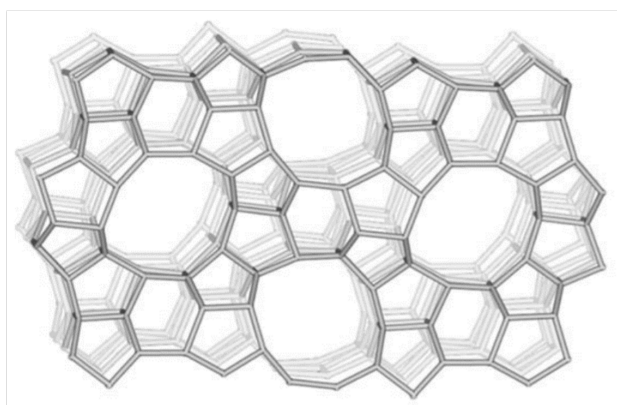Nuclear magnetic resonance is based on the principle that certain nuclei possess spin I. The spin gives rise to $2I + 1$ degenerate spin states. The application of a constant magnetic field causes the degeneracy to be split into multiple energy states. The nuclei experience an effective magnetic field B_{eff} that depends on their chemical surroundings and is described by the following Equation (1):

$$B_{\text{eff}} = B_0(1 - \sigma), \tag{1}$$

where σ is the shielding constant and B_0 is the applied magnetic field. Information about the local environment can be extracted from the chemical shielding through the manipulation of the nuclear spin. The absolute measurement of chemical shielding is challenging, and thus, chemical shift is used instead as a relative measure in parts per million (ppm) with respect to a known compound as a reference. Different chemical environments result in different chemical shifts. For instance, the chemical shift of aluminum in zeolites depends on the coordination number. For aluminum in a tetrahedral configuration, the chemical shift is around 60 ppm, whereas for octahedral coordination, it is around 0 ppm, and for an aluminum with the coordination number of five, it is around 20 ppm.[1]

^{27}Al nucleus has spin $\frac{5}{2}$, which means that it is a quadrupolar nucleus. Therefore the spectral peaks are broadened due to the quadrupolar interaction, which can be described by two parameters. The magnitude of this interaction is denoted by the coupling constant C_{Q}, while the asymmetry of the interaction is characterized by the asymmetry parameter η.

Symmetric environments usually have lower values of C_{Q}, which leads to a smaller broadening of NMR signals.[1] The electron distribution around NMR-active nuclei is typically non-spherical, causing the shielding to depend on the orientation of the molecule. In solid-state NMR, shielding is represented by a 3×3 matrix-shaped tensor σ displayed in Equation (2):[13]

$$\sigma = \begin{pmatrix} \sigma_{xx} & \sigma_{xy} & \sigma_{xz} \\ \sigma_{yx} & \sigma_{yy} & \sigma_{yz} \\ \sigma_{zx} & \sigma_{zy} & \sigma_{zz} \end{pmatrix}, \tag{2}$$

where x, y, z are the spatial axes with z being parallel to magnetic field B_0. The shielding tensor is commonly described by two values: the isotropic chemical shift and the asymmetry.[13] The isotropic shielding σ_{iso}, which determines the position of NMR peaks, is calculated by Equation (3):

$$\sigma_{\text{iso}} = \frac{1}{3}(\sigma_{xx} + \sigma_{yy} + \sigma_{zz}). \tag{3}$$

Asymmetry η, which dictates the shape of peaks in NMR spectra, is defined by Equation (4):

$$\eta = \frac{\sigma_{xx} - \sigma_{yy}}{\sigma_{zz} + \sigma_{\text{iso}}}. \tag{4}$$

2.2.1 NMR in zeolites

Solid-state NMR is a valuable technique to gain insight into the local structure of elements in solids, such as zeolites. Specifically, ^{29}Si NMR is a useful tool to investigate silicate materials, as it can provide information on the Si-O-Si angles, which are known to directly correlate with chemical shifts. ^{17}O NMR is mainly used to study the catalytically active sites in zeolites, as the oxygen atoms in the framework act as the acidic centers.[14]

Several studies have proposed that ^{27}Al NMR spectroscopy can be utilized to identify and assign aluminum atoms to distinct T-sites in zeolite structures.[2–5] Numerous research papers have reported the successful assignment of specific T-sites to different peaks in ^{27}Al NMR spectra in zeolites such as ZSM-5[4] or FER.[3] Most studies that have calculated the ^{27}Al chemical shift in zeolites have done so without considering the presence of a counter ion that would compensate for the negative charge of the AlO_4^- unit. Instead, a uniform positive background charge is assumed. It has been suggested that the solvated countercations do not interact with AlO_4^- tetrahedra and that neither the solvent nor the countercations affect the shielding.[4]

Experimental NMR is typically not performed on dehydrated zeolites because they

produce broad spectral peaks. Without water, the acidic proton cannot solvate, which causes the Al-O bond to stretch and produces a strong quadrupolar interaction.[15] This can make NMR features invisible or difficult to discern. Figure 6 displays the experimental ^{27}Al NMR spectra of zeolite H-ZSM-5 at different hydration degrees.[16] In addition to neglecting the presence of water and countercations, various assumptions are commonly made in the evaluation of chemical shielding in computational studies, such as relying on a single minimal energy structure that is expected to represent the system of interest fully. However, the validity of these assumptions is not fully established, and some have already been challenged.[7,17]
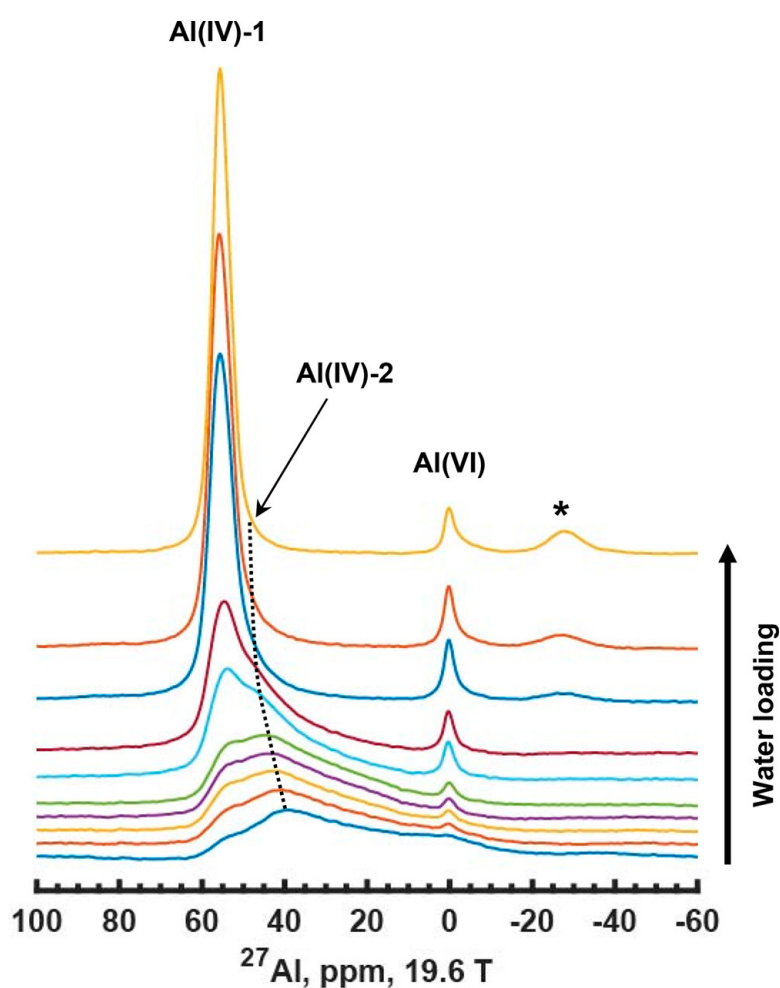

Figure 6: Experimental ^{27}Al NMR spectra of zeolite H-ZSM-5 at different water loadings. Two major peaks are observed corresponding to Al atoms at different coordination environment.[16]

One of the most influential studies on ^{27}Al NMR was conducted by Lippmaa in 1986. This study proposed a simple linear correlation between aluminum chemical shifts and

average Al-O-Si angle displayed in Equation (5):[18]

$$\delta_{CS}(\text{Al}) = -0.50 \cdot \alpha(\text{Al-O-Si}) + 132 \text{ ppm} \tag{5}$$

where $\delta_{CS}(\text{Al})$ is the chemical shift of aluminum and $\alpha(\text{Al-O-Si})$ is the angle between atoms Al-O-Si (i.e., TOT angle), which is typically between 130 ° to 150 °. This method of calculating chemical shifts has been used in numerous articles.[5,19] The validity of a simple linear correlation between TOT angles and the [27]Al chemical shift has been debated, particularly for highly siliceous zeolites, suggesting Si/Al ratio also plays a role in chemical shifts.[4,17] Also, the research has shown that another aluminum atom can cause a change in the chemical shift of up to 4 ppm.[20]

The correlation between the TOT angle and aluminum chemical shift has been theoretically explained. As the TOT angle increases, the hybridization of the oxygen atom changes, resulting in an increase in the s orbital component and a contraction of bonding orbitals. This causes an increase in electron density on aluminum and oxygen atoms, leading to an increase in chemical shielding and a decrease in chemical shift.[21] It has been suggested that the bond distance between aluminum and oxygen does not have a significant effect compared to the TOT angle.[21]

2.3 Machine learning and potential energy surface sampling

Machine learning (ML) is a subfield of artificial intelligence that involves the use of algorithms to analyze and learn from data, in order to make predictions based on patterns in the data. There are two main types of machine learning: supervised learning and unsupervised learning. In supervised learning, a model is trained using data that contains both input and output values (labels), while unsupervised learning is trained using only input values.[22] Among the popular machine learning algorithms are neural networks and kernel ridge regression (KRR). ML techniques have been increasingly utilized in various fields, including data processing,[23] healthcare[24] or in scientific research including sampling of the potential energy surface (PES).[25]

The total potential energy of an atomic system can provide information about many of its properties. Local minima in the PES correspond to (meta)stable atomic struc-

tures, such as pristine or defective zeolites and surfaces. The curvature of the PES correlates with the force constants and vibrational and phonon frequencies, while saddle points on the surface correspond to the activation energies of chemical reactions and phase transitions.[22] Due to the high computational cost of calculating the PES with *ab initio* methods, particularly for large systems or molecular dynamics simulations, neural network-based methods have been developed to model the PES with similar accuracy.[26]

Neural networks are a type of machine learning algorithms that are inspired by biological neurons. They consist of multiple layers, including an input layer (e.g., molecular structure) and an output layer (e.g., potential energy), with several hidden layers in between. These layers consist of interconnected modules, known as neurons, which process inputs from the previous layer. The inputs are processed using parameters, weights, and activation functions.[25] Neural network potentials (NNP) are machine learning models that are trained to model the PES of atomic systems with *ab initio* accuracy. They are trained using data from density functional theory (DFT) calculations, and can only be as accurate as the reference DFT method used to train them. NNPs have found use in simulations of large systems, such as zeolites, with energy errors as low as a few meV per atom.[6]

Kernel ridge regression (KRR) is a supervised machine learning method that combines the ridge regression algorithm with the kernel trick. Ridge regression is a method used to estimate model coefficients that fit the training data best. These coefficients are estimated by minimizing the cost function, which consists of the sum of squared residuals and a regularization term with parameter λ and the square of model coefficients. The cost function is displayed by the following Equation (6):

$$\sum_{i=1}^{k}(y_i - \hat{y}_i)^2 = \sum_{i=1}^{k}(y_i - w \cdot x_i)^2 + \lambda \cdot w^2,\qquad(6)$$

where y_i is reference value of training data, $\hat{y}_i$ is the predicted value, w represents model coefficients, x_i is the input value of training data (features), k is the total number of samples in the training set and λ is a parameter determining the bias. An optimal λ is determined by grid search to achieve the best possible trade-off between model variance

and bias. The kernel trick is a method used to perform linear regression by mapping the original features to a higher dimensional space. This method allows for more complex and flexible models to be constructed. The transforming of the data is illustrated in Figure 7. KRR has gained popularity in the scientific community due to its simplicity and flexibility. It has been widely used for predicting material properties from descriptors, as well as generating density functionals.[22] KRR has also been shown to perform competently in calculation of solid state NMR parameters in aluminosilicate glasses.[27]
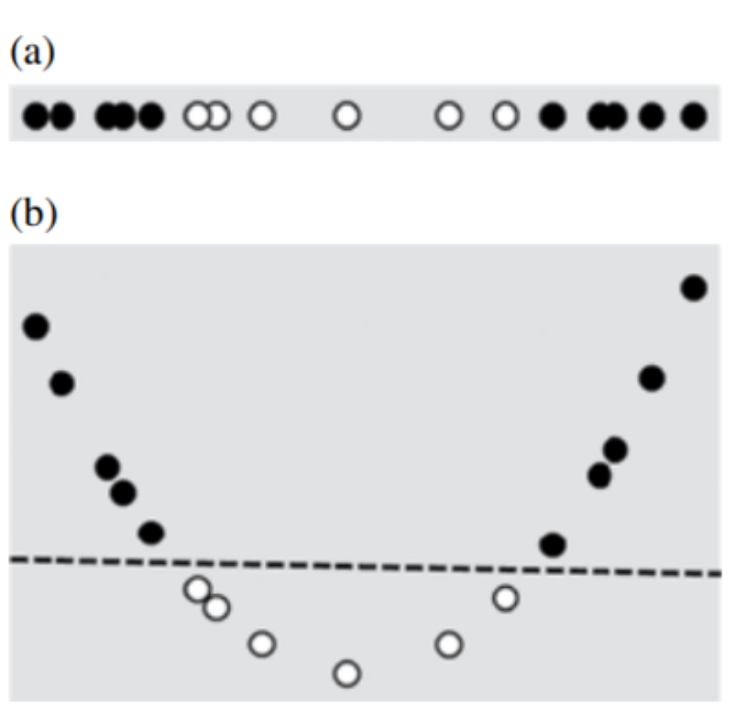

Figure 7: Visualization of the kernel trick. (a) original set of data and (b) transformed data into higher dimension separable by a hyperplane.[22]

2.4 Density functional theory

The stationary Schrödinger equation is a fundamental tool for determining the properties of an atomic system. There are three popular approximations to solving this equation: Hartree-Fock (HF) method, post-HF methods, and Density Functional Theory (DFT). The (post)-HF methods rely on $3N$ variables, where N represents the number of electrons, while DFT depends on only three spatial variables. Due to their high computational cost scaling at least as $\mathcal{O}(N^3)$, HF and post-HF methods are not always practical for large systems, such as zeolites. Therefore, DFT has become a popular method for calculating electronic structure and atomic properties. DFT employs the electron density as a

substitute for many-electron wave functions. The electron density ρ is defined by Equation (7):[28]

$$\rho\left(\mathbf{r}\right) = N \int d\mathbf{r}_2 \int d\mathbf{r}_3 \ldots \int d\mathbf{r}_N \left|\Psi\left(\mathbf{r}, \mathbf{r}_2, \ldots, \mathbf{r}_N\right)\right|^2 \qquad (7)$$

where Ψ is the wavefunction, N is the number of electrons and $\mathbf{r}$ is a spatial variable. The validity of using electron density instead of many-electron wave functions in calculating atomic properties is justified by the Hohenberg-Kohn theorems.[29]

The energy of a system, E_{DFT}, can be calculated by expressing it as a function of electron density, which can be mathematically represented by Equation (8):[28]

$$E_{\mathrm{DFT}} = T_{\mathrm{S}}\left[\rho\right] + E_{\mathrm{ne}}\left[\rho\right] + J\left[\rho\right] + E_{\mathrm{xc}}[\rho], \qquad (8)$$

where T_{S} is the kinetic term, E_{ne} is the potential energy term between nucleus and electron, J is the Coulomb term and E_{xc} is the exchange-correlation functional. The terms T_{S}, E_{ne}, and J in the Equation (8) can be calculated exactly, while the exchange-correlation functional term E_{xc} cannot be precisely determined. Various DFT methods differ in their approaches to calculating this E_{xc} functional. It is worth noting that we tacitly assumed the validity of the Born-Oppenheimer approximation, which assumes that the nuclei are stationary due to their much larger mass compared to electrons.

2.5 Molecular dynamics

Molecular dynamics is a simulation method used to study the behavior of many-body systems, implementing the evolution of the atomic system according to laws of motion described by Newton. While this method can be applied to most systems, it is important to consider quantum effects when dealing with light atoms or high-frequency vibrational motion.[30]

The molecular dynamics algorithm can be simplified into several steps. Firstly, the forces acting on each atom are computed by evaluating the interaction potentials. To improve the efficiency of the simulation, a cut-off radius is introduced, which limits the number of interacting atoms. Then the forces are numerically integrated using Newton's second law of motion and the Verlet algorithm to propagate positions and velocities of atoms in time. The total kinetic energy of the system fluctuates which changes the

instantaneous temperature. The temperature is a function of the velocity v_i and mass m_i of each atom and can be defined using Equation (9):[30]

$$T = \sum_{i=1}^{N} \frac{m_i v_i^2}{k_\text{B} N_\text{f}} \,,$$ (9)

where k_B is the Boltzmann constant and N_f are the degrees of freedom which can be calculated as $3N - 3$ with N being the number of atoms. To replicate experimental conditions, it is essential to maintain a constant temperature during molecular dynamics simulations. Several thermostats have been developed for this purpose, including the Berendsen, Andersen, and Nosé-Hoover thermostats.[30]

3 Methods

3.1 Creating the models

The MFI structure was obtained from the IZA database,[10] and its unit cell parameters were adjusted to the values that have been demonstrated to be optimal for the framework containing aluminum.[31] Optimal cell volume was determined by the lowest energy model. Modified cell parameters are:

a = 20.20 Å

b = 19.85 Å

c = 13.22 Å.

These parameters result in a total volume of 5300 Å^3. Using Materials studio,[32] one silicon atom was replaced with aluminum and a hydrogen atom was manually added to the oxygen atom bound to the aluminum, creating a Si-O(H)-Al group that represents a Brønsted acid site (BAS). The MFI unit cell contains 12 distinct T-sites, each connected to four oxygen atoms, resulting in a total of 48 possible BAS. For models with a lower Si/Al ratio, the same modeling procedure was followed as for the unit cell containing one aluminum atom.

Water molecules were manually placed to provide immediate access to the BAS for low water loading (i.e., 1-3 water molecules). For higher water content, the Sorption module in Materials Studio[32] was used with the COMPASS27 force field to place the water molecules.

3.2 Calculation details and NNP

Neural network potentials (NNP) were employed to perform molecular dynamics simulations and geometry optimizations. These potentials were developed specifically for zeolite systems containing water molecules, and were constructed using the SchNetPack library[33] written in the Python programming language.[34] The training and testing data for the neural networks were obtained using DFT employing the exchange-correlation functional SCAN+D3.[35] The Python library Atomic Simulation Environment (ASE)[36] was utilized to analyze structural features, such as the positions of framework atoms and water molecules.

3.3 Optimization of geometry

Three optimization algorithms (BFGS, FIRE, and MDMIN) were evaluated to determine the optimal method for geometry optimization. BFGS was found to offer the best performance and was therefore used for all subsequent geometry optimizations. To achieve convergence, a maximum force threshold of 1 meV·Å^{-1} per atom was applied. A radius cutoff of 6 Å was used to limit the interaction range included in the calculations. The optimizations were conducted under constant volume conditions.

3.4 MD simulation

The MD simulation was performed for a time period of 1 ns with a time step of 0.5 fs. To facilitate the analysis of various structural characteristics, a snapshot of the system was saved every 100 steps, leading to a total of 20 000 structures per one 1ns simulation. The Nosé-Hoover thermostat was employed to maintain a temperature of 350 K throughout the simulation.[37] The radius cutoff was set at 6 Å. The simulation was conducted in the NVT ensemble, where the number of atoms, temperature, and volume were held constant.

4 Results and discussion

4.1 Solvation and adsorption energy

The role of proton solvation is significant in ^{27}Al NMR as keeping the proton bound to the framework leads to the broadening of spectral peaks, thereby making it difficult to interpret the experimental spectra. In this thesis, solvation is defined as follows: the proton is considered solvated when the distance of oxygen from the water molecules to the proton R_{w} is smaller than the distance of oxygen from the framework to the proton R_{f}. The distances mentioned in the solvation definition can be visualized in Figure 8.
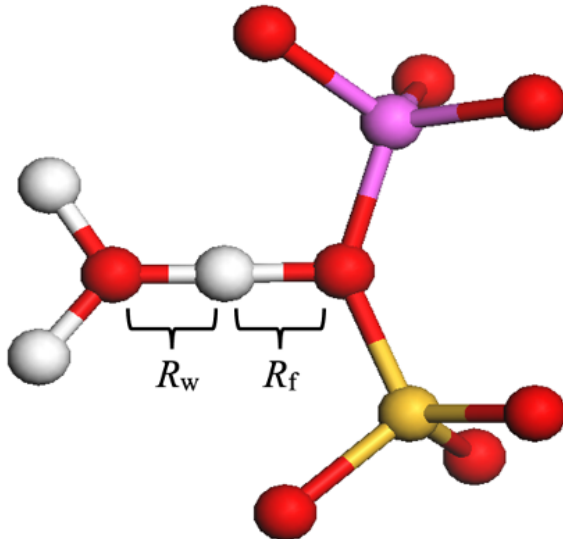

Figure 8: Visualization of the distances used in solvation analysis, where R_{w} represents the distance between the oxygen of water molecule to the proton, and R_{f} represents the distance of the framework oxygen to the proton.

The solvation percentage was determined using Equation (10):

$$\text{Solvation}[\%] = \frac{N_{\text{solv}}}{N_{\text{total}}} \cdot 100 \text{ \%}, \tag{10}$$

where N_{solv} is the number of structures in which the proton is solvated and N_{total} is the total number of structures generated from the MD simulation. This value can be used to estimate the acidity of specific Brønsted acid sites (BAS), as the acidic BAS have higher solvation percentages. The solvation can be a valuable parameter for investigating the energy of water adsorption. Water adsorption energy E_{ads} is the amount of energy that

is released when a water molecule is adsorbed onto the zeolite framework. This energy can be calculated using Equation (11):

$$E_{\text{ads}} = E_{\text{fr,w}} - E_{\text{fr}} - E_{\text{w}}, \tag{11}$$

where $E_{\text{fr,w}}$ is the potential energy of a zeolite system with a water molecule, E_{fr} is the energy of the framework without water molecule and E_{w} is the energy of the water molecule in a gas phase. Figure 9 shows the relationship between the adsorption energy and the solvation percentage of all possible BAS with 1 Al atom per unit cell. This correlation is consistent with previous studies, which state that BAS with higher acidity have more exothermic adsorption energy,[38] justifying the use of the solvation as an important parameter of the model.
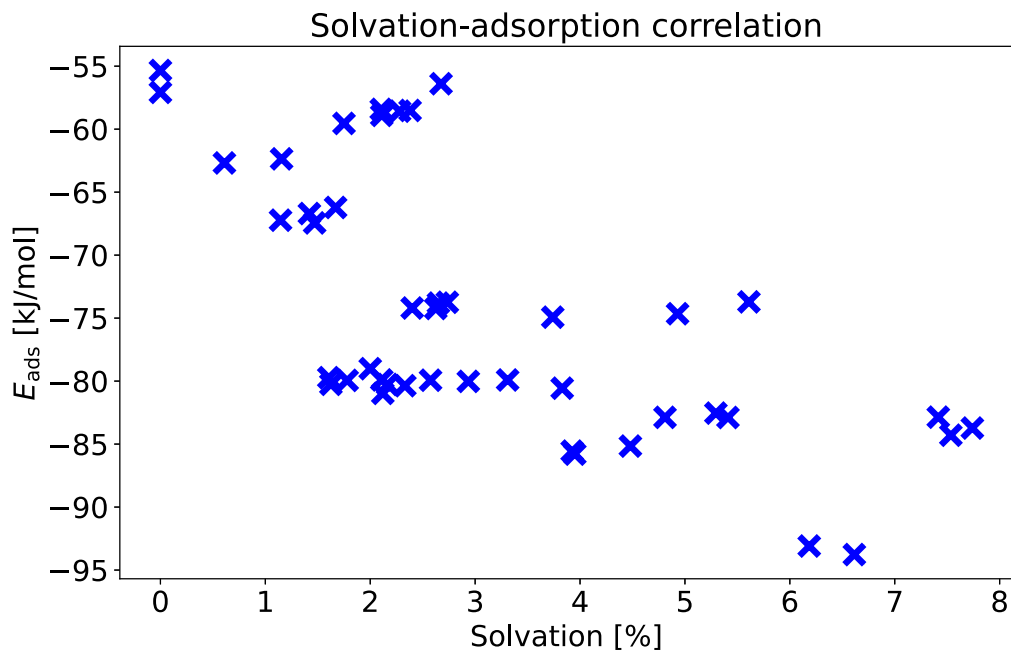

Figure 9: Adsorption energy dependence on solvation percentage for the model with 1 water molecule and 1 Al per unit cell.

4.2 Chemical shift models and their performance

There are multiple methods to calculate chemical shift, each with varying levels of accuracy and computational cost. The most computationally expensive approach is to use *ab initio* methods, such as DFT, to calculate chemical shift. Although this is a very accurate method, the high computational cost makes it unsuitable for employing it for large number of structures. In contrast, a simpler method was proposed by Lippmaa,[18] which relates the TOT angle with the chemical shift, as described in the Equation (5).

Several alternative methods for calculating chemical shift have been developed beyond the Lippmaa method. These include the 2-parameter equation (2p), as denoted in Equation (12), the 5-parameter equation (5p), as denoted in Equation (13), based on bond lengths d and bond angles α:

$$\sigma = 180.60 \cdot d(\text{Al-O}) + 0.80 \cdot \alpha(\text{Al-O-Si}) + 64.15 \text{ ppm} \tag{12}$$

$$\begin{aligned}\sigma = \ &185.34 \cdot d(\text{Al-O}) - 15.6 \cdot (\max[d(\text{Al-O})] - \min[d(\text{Al-O})]) + 2.94 \ \cdot d(\text{Al-Si}) \\ &+ 0.73 \cdot \alpha(\text{Al-O-Si}) - 1.40 \cdot \alpha(\text{O-Al-O}) + 209.33 \text{ ppm,}\end{aligned} \tag{13}$$

and the Kernel Ridge Regression (KRR) approach.[7] To transform chemical shielding σ into chemical shift δ, a calibration Equation (14) was used:[7]

$$\delta = -1.110 \cdot \sigma + 609.7 \text{ ppm} \tag{14}$$

To assess the reliability of various methods for calculating shielding, they were compared with results obtained from DFT using PBE functional. Due to the high computational cost of DFT calculations, only a limited number of structures were selected for a testing set. Specifically, the testing set comprised structures with a single aluminum atom per unit cell, and for each water loading (0, 1, 2, 3, and 17 water molecules per unit cell), two Brønsted acid sites (BAS) were chosen to cover the entire range of chemical shifts obtained from the KRR calculations. The chosen BAS are listed in the Table 1. For each selected BAS, 10 structures were selected from the MD trajectory with equal time intervals between them, resulting in total of 100 structures in the testing set. The testing set includes structures that contain O4 oxygen, which exhibit high instability when a

water molecule is introduced to the framework, as it cannot be solvated due to its intraframework hydrogen bond with the framework oxygen and its position which prohibits the access of a water molecule. This property of O4 oxygen causes the KRR method to significantly overestimate chemical shift values compared to DFT calculations, with an offset of approximately 8 ppm. However, the trend in chemical shift values obtained from the KRR method is relatively consistent with DFT calculations, as shown in Figure 10. The overestimation could be attributed to the absence of similar configurations to O4 in the KRR training set.

Table 1: Selected BAS for each water loading to test the accuracy of the different methods

Water loading	1st BAS	2nd BAS
0	T4O4	T6O10
1	T9O21	T7O18
2	T9O15	T7O18
3	T9O15	T7O17
17	T9O15	T11O16

To evaluate the performance of each model, the error was calculated as the absolute difference between the chemical shift obtained from DFT and that predicted by the model. The mean absolute error (MAE) for each water loading is summarized in the Table 2.
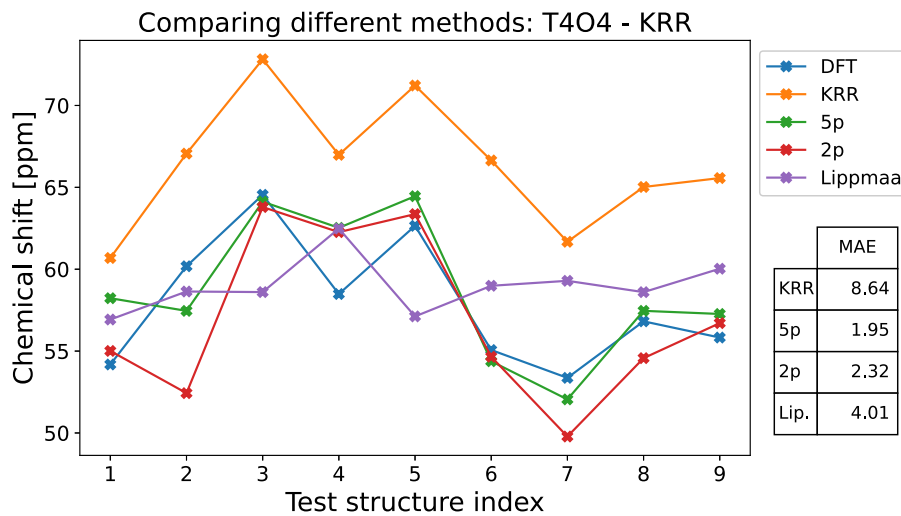

Figure 10: Comparing different chemical shifts calculation method with DFT for T4O4 BAS without water with MAE values in ppm.

Table 2: Mean absolute error in ppm of methods compared to DFT at different water loadings

Water loading	Lippmaa	2p	5p	KRR
0	4.01	2.34	2.21	5.61
1	3.19	1.60	2.06	1.84
2	3.63	1.11	1.74	1.18
3	3.96	1.85	2.04	1.32
17	3.03	1.90	2.13	1.27

The high error of KRR for lower water loading can be explained by reviewing the testing structures for each chosen BAS individually. Selected testing structures are shown in Figure 11. Despite the high error associated with the chemical shift calculated from KRR, it exhibits similar trends to the DFT-calculated chemical shifts with an average offset of approximately 2-3 ppm. However, this offset diminishes at higher water loading. At the maximum water loading of 17 water molecules per unit cell, KRR surpasses all other methods by more than 0.5 ppm.

The two and five parameter equations have been found to be more accurate in calculating chemical shift compared to the simple correlation between chemical shift and TOT angle proposed by Lippmaa.[18] This observation indicates that the chemical shift of tetrahedral Al atoms in zeolite framework is influenced not only by the TOT angle but also by the Al-O bond length. These findings contradict a study done by Liu et al.[21] that suggests that bond distance is not as significant as TOT angle in determining chemical shift. However, the KRR method has been demonstrated to be the most reliable method for calculating chemical shift for hydrated zeolites. Hence, the KRR method is employed for computation unless stated otherwise.
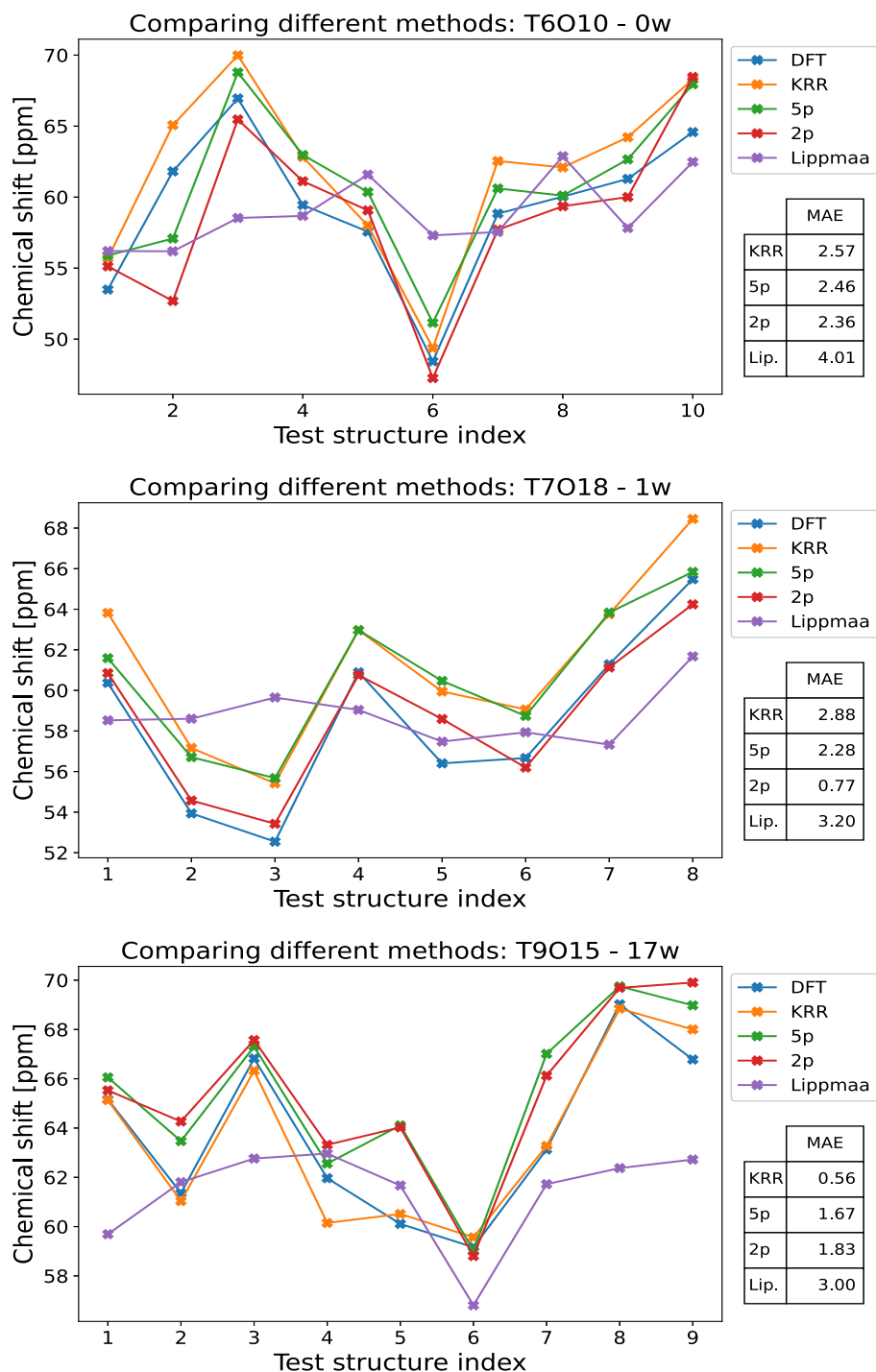
Figure 11: Chemical shifts for individual testing structures for a specific BAS at different water loadings with MAE values in ppm. BAS and water loading (from top): 0 water molecules/UC, BAS T6O10
1 water molecule/UC, BAS T7O18
17 water molecules/UC, BAS T9O15

4.3 Time-averaging chemical shift values

Calculating chemical shift typically involves obtaining a single structure that is expected to correspond to the ground state of the structure, which is the arrangement of atoms that corresponds to the global minimum energy state on the potential energy surface. This is expected to be achieved by optimizing the geometry of the reasonable initial structure chosen by the expert. To verify the validity of this approach, all structures in the molecular dynamics (MD) simulation of a model without water and with 1 Al atom per unit cell at 350 K were locally optimized using NNP, and the chemical shift was calculated using the 2-parameter equation. Figure 12 shows the chemical shift of the optimized structures as a function of time during the MD simulation.
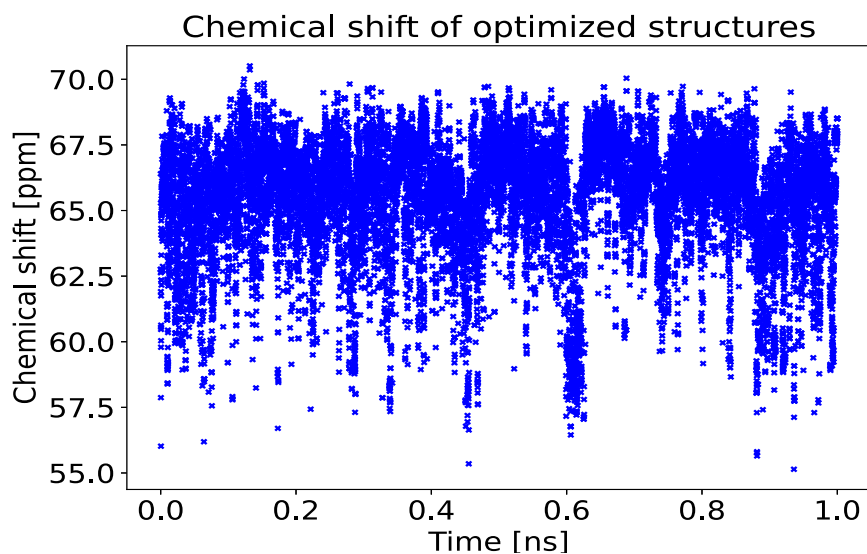

Figure 12: Chemical shift of the locally optimized structures of the MD trajectory of a model without water and with 1 Al atom per unit cell at 350 K.

The calculated chemical shift of locally optimized structures is highly dependent on the starting structure, and thus it cannot be reliably reproduced. These results are consistent with what has been proposed by Hoffman et al.[39] for MFI zeolite.

In contrast to using local optimization to obtain structures and calculate their chemical shifts, a plot of the computed chemical shift over time was generated using the entire MD simulation of the same system. This plot is shown in Figure 13. Although the chemical shift values vary over a wide range, the moving average value remains relatively constant.
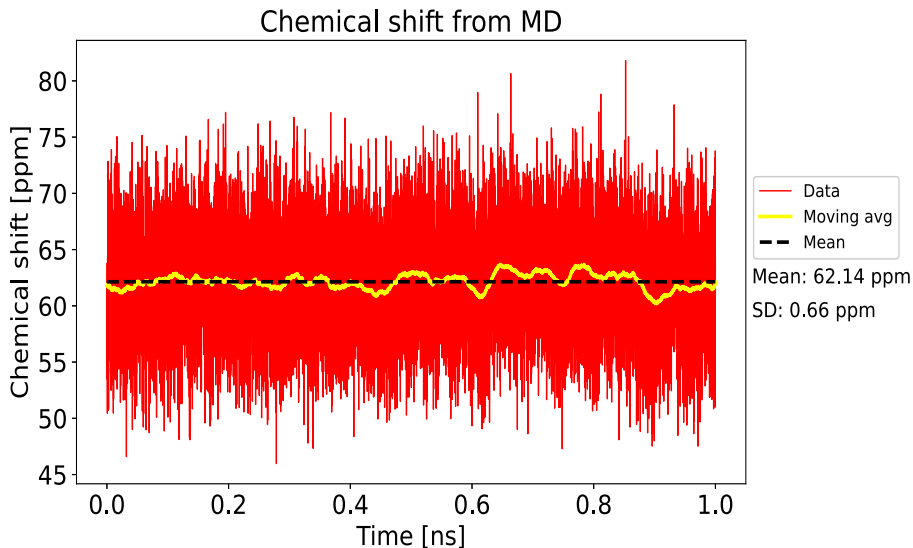
Figure 13: Chemical shift over time of the MD trajectory with moving average with a window size of 25 ps. The average value of chemical shift is 62.14 ppm and the standard deviation of the moving average values is 0.66 ppm.

4.4 Effects of different factors on chemical shift

4.4.1 Water loading

Chemical shift was investigated at various water loadings, including 0, 1, 2, 3, or 17 water molecules per unit cell. The 17 water molecules represent hydrated zeolites, which were determined based on the accessible volume of MFI zeolite from the IZA database[10] and the required amount of water molecules to achieve a water density of $1 \text{ g} \cdot \text{cm}^{-3}$. This amount of water molecules is consistent with the findings of Holzinger et al.,[5] who reported a water loading of 15 ± 1 water molecules per unit cell in their experimental investigation of H-ZSM-5 sample with a Si/Al ratio of 50.

There are four oxygen atoms connected to every T-atom in the zeolite framework, resulting in four possible Brønsted acid site (BAS) configurations for every aluminum atom. The distribution of BAS populations is governed by the Boltzmann probability distribution, which states that the most stable configurations are the most likely to be populated. The probability of each configuration can be calculated using Equation (15):

$$\frac{p_i}{p_j} = e^{\frac{(E_j - E_i)}{k_B T}},\tag{15}$$

where p_i is the probability of a state i with energy E_i, k_B is the Boltzmann constant and

29

T is the temperature.

By calculating the Boltzmann probability factor for all four BAS configurations and normalizing them so the sum of Boltzmann probability factors is equal to one, a normalized Boltzmann probability factor is obtained that was used to average the chemical shift values of the different BAS configurations. This Boltzmann averaging of the four BAS configurations connected to each T-atom yields a single value for the chemical shift, which is expected to be a valid approximation of the experimentally measured value for a specific T-site.

The plot presented in Figure 14 shows the variation in chemical shift values as a function of water loading for all twelve T-sites. The high chemical shift values observed in the dehydrated state (i.e., at 0 water molecules per unit cell) can be partially explained by the offset of the KRR method at low water loading, as demonstrated in section 4.2. This is made evident by examining the results from the other methods, as shown in the Figure 15, where it can be observed that there is no general decrease in chemical shift from 0 to 1 water molecule. In Table 3 are values of chemical shift for each water loading from different methods of calculating chemical shift.
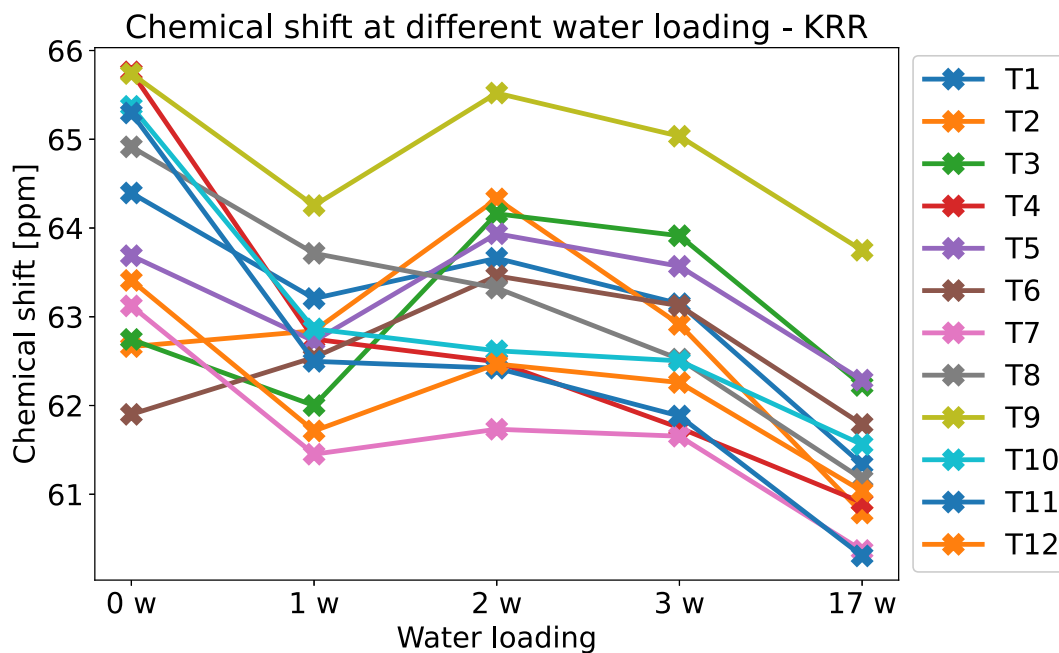

Figure 14: Boltzmann averaged values of chemical shift for each T-site in the MFI framework with 1 Al atom per unit cell and different water loadings. The chemical shift was calculated by the KRR method.
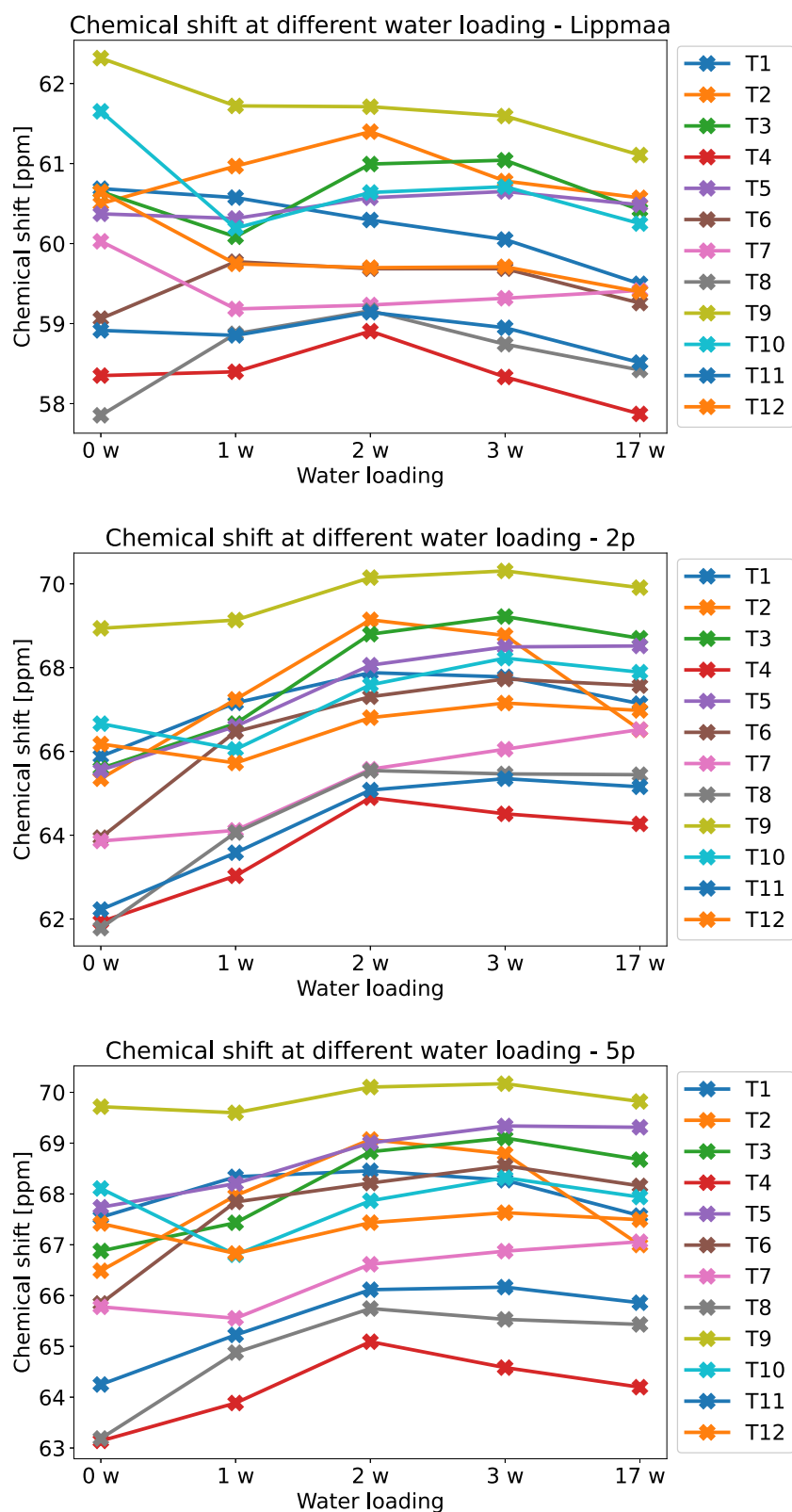
Figure 15: Chemical shift values for each T-site in the MFI framework with 1 Al atom per unit cell and different water loadings. The chemical shift was calculated by different methods: Lipppmaa, 2p and 5p.

Table 3: Average values of chemical shift with ranges in ppm at different water loadings from different methods

Water loading	Lippmaa	2p	5p	KRR
0	60.1 ± 2.2	65.4 ± 3.6	66.4 ± 3.3	63.8 ± 1.9
1	60.1 ± 1.7	66.1 ± 3.1	66.7 ± 2.9	62.9 ± 1.4
2	60.3 ± 1.4	67.5 ± 2.6	67.6 ± 2.5	63.6 ± 1.9
3	60.0 ± 1.6	67.4 ± 2.9	67.4 ± 2.8	63.3 ± 1.7
17	59.5 ± 1.6	67.1 ± 2.8	67.0 ± 2.8	62.0 ± 1.7

The most significant change in chemical shift calculated by KRR is observed between models with 3 and 17 water molecules per unit cell. This change can be attributed to the increasing number of water molecules, as solvation percentages of 3 and 17 water molecules are practically the same, both with over 99 %. This observation suggests a connection between the number of water molecules and chemical shift for all T-sites. However, the magnitude of this connection varies for each specific T-site, indicating that the effect of the number of water molecules on chemical shift is site-dependent. Other methods for calculating chemical shifts do not exhibit this effect, as can be seen from Figures 15 and Table 3. However, they are less robust than KRR in this range of water loading, as shown in section 4.2.

4.4.2 Al pairs

The introduction of an additional aluminum atom within the zeolite framework may result in the formation of an aluminum pair. The aluminum pair is characterized by two aluminum atoms in close proximity to one another, to the extent that they can interact and influence each other. However, if these two atoms are positioned at a sufficient distance, their behavior will resemble that of isolated aluminum atoms.[40]

The T-atom closest to the aluminum atom is referred to as the nearest neighbor, followed by the next-nearest neighbor and so on. However, the closest two aluminum atoms can only be in a next-nearest neighbor (NNN) position due to the Löwenstein rule,[11] which prohibits the nearest neighbor position. The relative positions of aluminum pairs with NNN and next-next-nearest neighbor (NNNN) positions are illustrated in Figure 16.
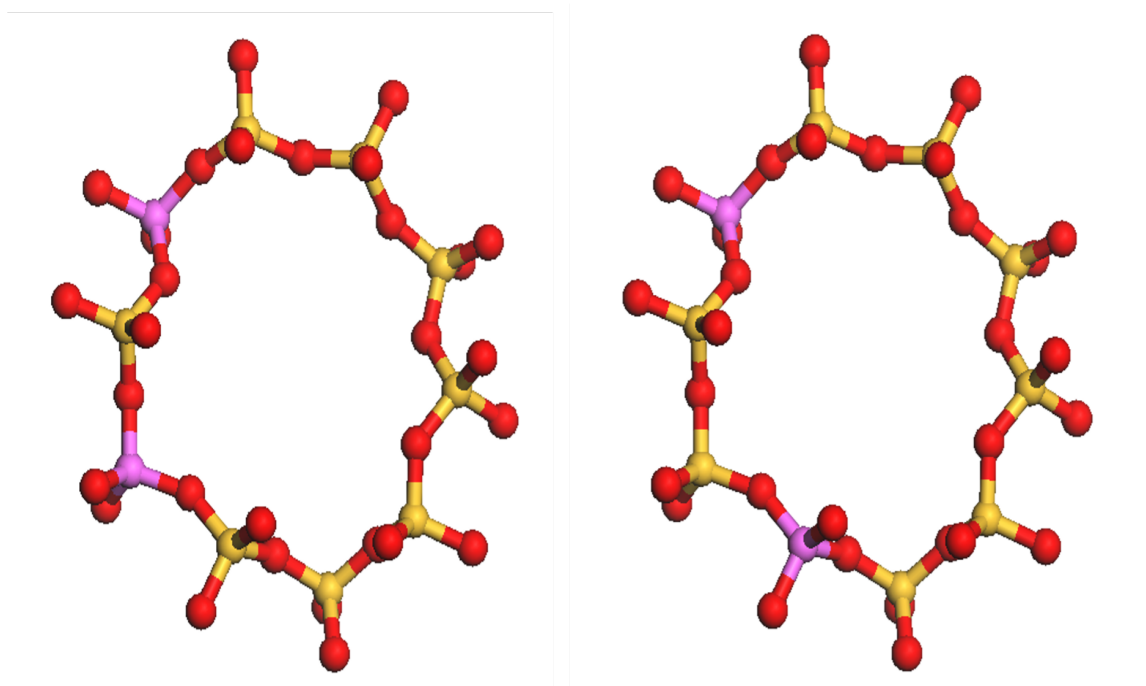

Figure 16: Al pair in the NNN positions (left) and NNNN position (right) in one ring.

To investigate the effect of introducing an additional aluminum atom to the framework, a total of 20 models containing aluminum pairs were created. At least one of the T-site in each model was T12, which was selected due to its position at the intersections of channels and its reputation as the most stable position for aluminum atoms in the MFI framework.[12] The T12O8 BAS was selected as the initial BAS due to its stability across

various water loadings. Subsequently, an additional aluminum atom was inserted into the framework at either the next-nearest-neighbor (NNN) or the next-next-nearest-neighbor (NNNN) position relative to the initial T12O8 BAS. This resulted in the creation of 12 NNN and 8 NNNN aluminum pairs. Each pair was then loaded with varying numbers of water molecules per unit cell, specifically 0, 1, 2, 3, or 17 water molecules.

The KRR method was employed to calculate the chemical shift of each aluminum atom. The resulting T12 aluminum chemical shift was plotted as a function of the distance between aluminum atoms, as shown in Figure 17. In the dehydrated model, the chemical shift change can be as large as 3 ppm, and it is more prominent in NNN aluminum pairs. A similar trend is observed for all water loadings. Děděcek et al.[20] also reached a similar conclusion, reporting that the change in the chemical shift in the Al-Si-Al pair can be as high as 4 ppm.
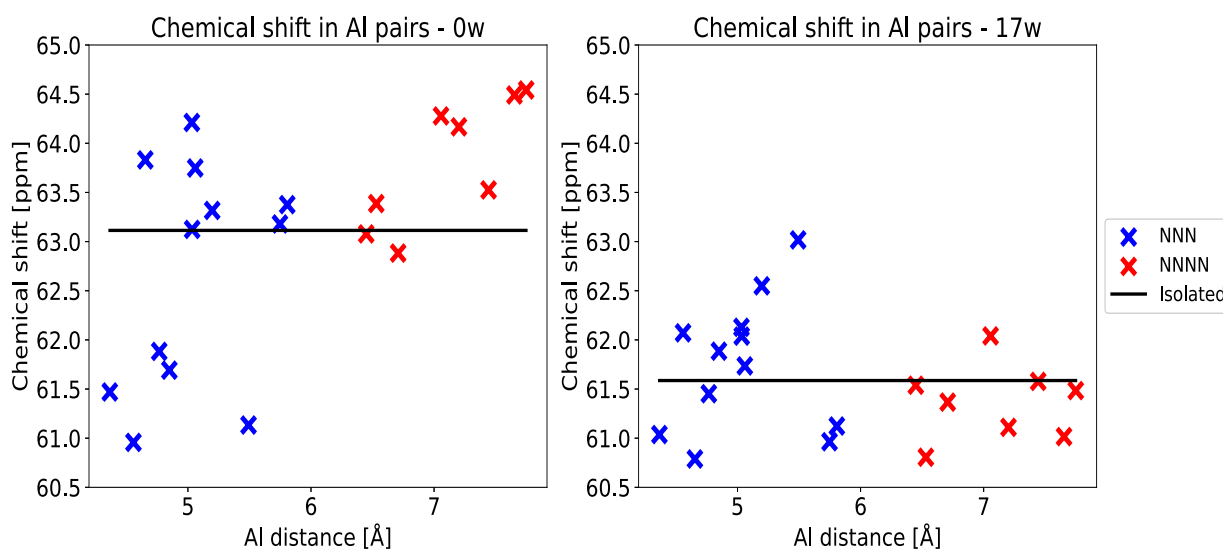

Figure 17: Calculated chemical shift of the aluminum at T12 position being part of the aluminum pair for 0 (left) and 17 (right) water molecules per unit cell. Black line represents chemical shift value for isolated T12O8. Blue data points are from NNN pairs and red are from NNNN pairs.

There appears to be no direct correlation between the interatomic distance separating aluminum atoms and their chemical shifts, as the chemical shift can vary in either a positive or negative direction. Furthermore, the magnitude of the chemical shift alteration appears to be distinct for each individual aluminum pair.

4.4.3 Temperature

The effect of temperature is often neglected in studies that calculate ^{27}Al chemical shift because the calculations are typically based on structures at the minimum of the potential energy surface (PES). Temperature effect was investigated on the ^{27}Al chemical shift in the T5 position of the framework. The temperature range of 250 K to 500 K was considered with a step of 50 K. The T5 position was selected due to its high stability and location at the intersection of the channels. All possible BAS were modeled and subjected to MD simulations at different temperatures with varying water loading. The chemical shift of different BAS was averaged using the Boltzmann probability distribution at various temperatures. The results are presented in Figure 18 that illustrates the chemical shift as a function of temperature.
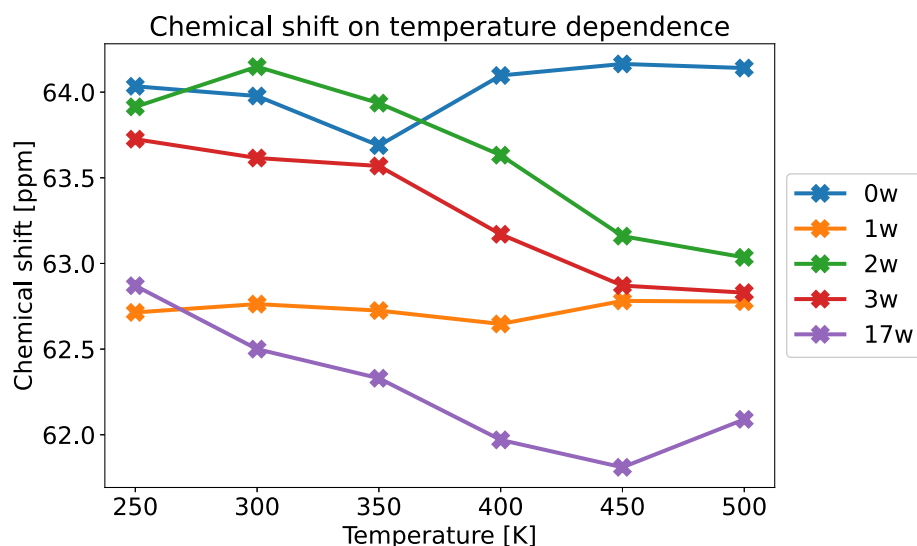

Figure 18: Chemical shift of aluminum at T5 position at different temperatures for different water loadings.

In the absence of water molecules, the zeolite model exhibits negligible changes in chemical shift with temperature. The largest change is observed between 300 K and 350 K, where the chemical shift increases by negligible 0.2 ppm. Similarly, for the model with 1 water molecule per unit cell, the chemical shift remains almost constant across the temperature range. However, for the models with higher water loadings of 2 and 3 water molecules, a significant and steady decrease in chemical shift is observed with increasing temperature. The model with 17 water molecules per unit cell exhibits a general decrease in chemical shift with a slight increase from 450 K to 500 K.

The observed phenomenon may be partially explained by variations in the solvation percentage with temperature. Figure 19 depicts the relationship between temperature and solvation percentage for various water loadings. For the model with 1 water molecule, the solvation percentage exhibits a slight increase of less than 1 %, which is a negligible change and could be due to small simulation errors. In contrast, for the model with 2 water molecules, the change in solvation percentage is significant, decreasing by more than 30 % from 250K to 500K. A similar trend is observed for the model with 3 water molecules, although the change is not as pronounced. The proton in the 17 water molecules per unit cell model remains fully solvated throughout the entire temperature range.

Through the comparison of the solvation and chemical shift dependencies on temperature shown in Figure 18 and Figure 19, it is observed that there are similar trends, indicating a direct correlation between the solvation percentage and chemical shift. It can be concluded that the temperature effect can be disregarded for dehydrated zeolites and models with only one water molecule per unit cell, as any change in chemical shift is solely due to the change in solvation. However, the observed decrease in chemical shift with increasing temperature in the models with 17 water molecules cannot be explained by solvation effects. This is because at such high water loadings, the proton is essentially fully solvated at all temperatures, rendering the applied definition of solvation less informative in explaining this observed trend.

The relationship between the solvation effect and chemical shift is complex, as the former is primarily determined by the number of water molecules present in the unit cell. As previously noted in section 4.4.1, the chemical shift decreases with an increasing number of water molecules, which is contrary to the presented effect of solvation percentage on the chemical shift.

The observed solvation effects on the chemical shift are specific to a given BAS and cannot be generalized to all BAS with 1 Al atom per unit cell or even to BAS connected to one T-site. The lack of correlation between solvation percentage and chemical shift for all BAS supports this conclusion. A plot of chemical shift as a function of solvation percentage for 1 water molecule per unit cell systems is presented in Figure 20. This suggests that other factors, such as the local environment and position of the Al atom, play a stronger role in determining the chemical shift than the solvation effect alone.
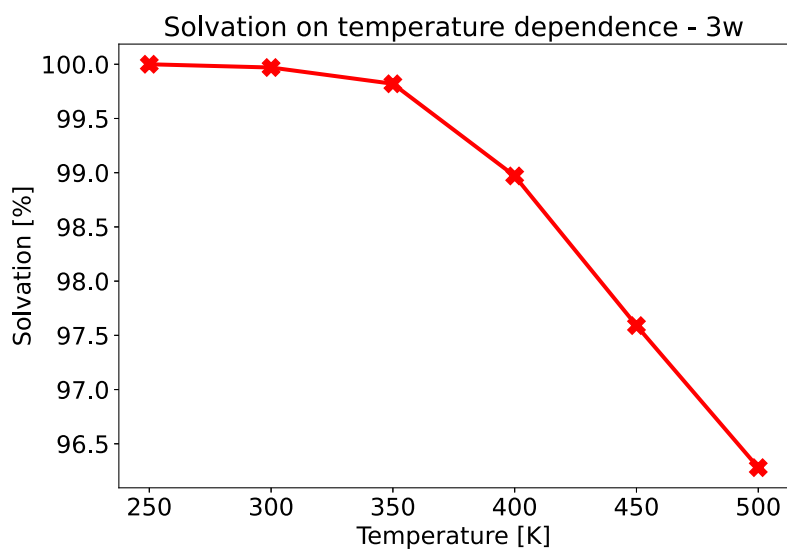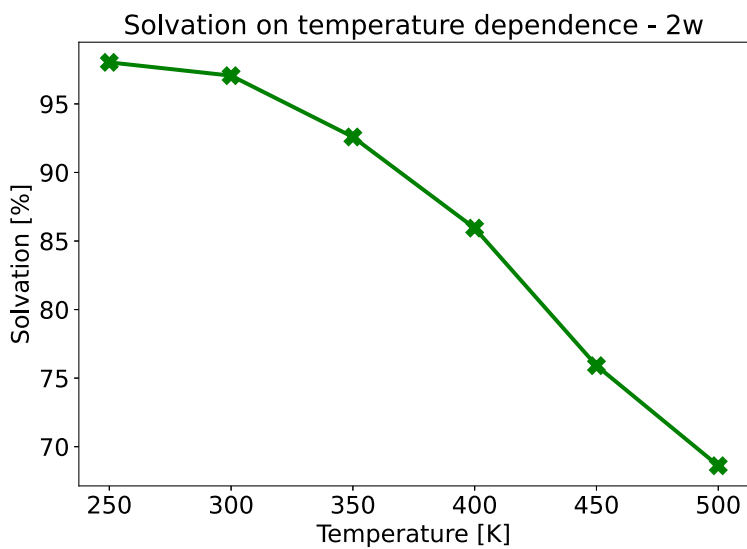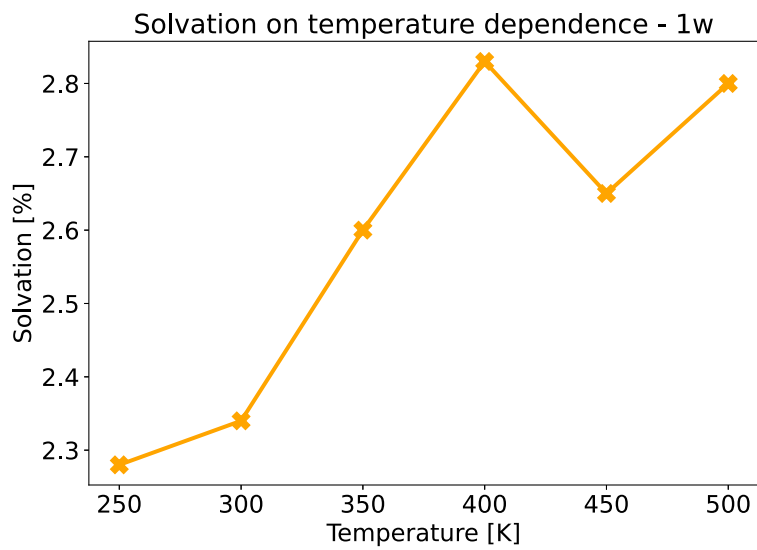
Figure 19: Solvation dependence on temperature for 1, 2 and 3 water molecules per unit cell.
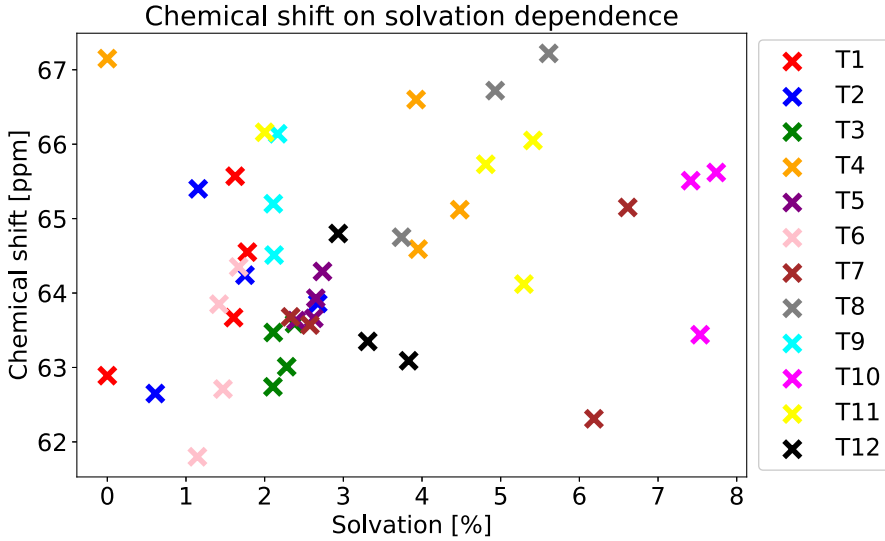
Figure 20: Chemical shift dependence on solvation percentage for all possible BAS with 1 Al atom per unit cell in MFI. BAS are colored based on what T-site they are connected to.

4.5 Comparing to experiment

As described in section 2.2, the experimental NMR measurements can only be obtained from fully hydrated zeolites. Therefore, to attribute the experimental NMR peaks to specific T-sites, the calculated chemical shifts at hydrated zeolites are considered. The chemical shift values for each T-site at water loading of 17 water molecules per unit cell calculated by the KRR method are provided in Table 4.

Table 4: Computed chemical shift for each T-site in hydrated zeolites using the KRR method.

T-site	σ(17w) [ppm]	T-site	σ(17w) [ppm]
T1	61.33	T7	60.37
T2	60.79	T8	61.17
T3	62.23	T9	63.75
T4	60.91	T10	61.56
T5	62.29	T11	60.31
T6	61.79	T12	61.04

[27]Al is a quadrupolar nucleus, resulting in a significant broadening of spectral peaks in ZSM-5. Full line widths are ranging from 0.9 to 2.3 ppm at 14.1 T.[5] The broadening

and overlapping of signals make it difficult to assign each experimental signal to a specific calculated chemical shift value obtained by the KRR method. This is due to the fact that most T-sites would exhibit overlapping signals, with the exception of T9, which has the highest chemical shift value.

Numerous research studies have attempted to assign experimental NMR peaks to specific T-sites by utilizing Lippmaa's correlation between TOT angle and chemical shift in Equation (5).[2,5,19] The chemical shifts at 17 water molecules per unit cell calculated by the Lippmaa's approximation are in Table 5. Despite the existence of some signal overlap, experimental NMR peaks could be assigned to a small group of T-sites, according to these results. Holzinger et al.[5] also reached a similar conclusion by employing Lippmaa's correlation and angles obtained from X-ray diffraction to calculate the [27]Al chemical shift for each T-site. They assigned experimental ^{27}Al NMR peaks to a groups of T-sites through this method. However, as demonstrated in section 4.2, chemical shift calculated solely from the TOT angle is not appropriate for the determination of [27]Al chemical shift.

Table 5: Computed chemical shift for each T-site in hydrated zeolites using the Lippmaa method.

T-site	$\sigma(17w)$ [ppm]	T-site	$\sigma(17w)$ [ppm]
T1	59.50	T7	59.41
T2	60.57	T8	58.42
T3	60.42	T9	61.11
T4	57.87	T10	60.25
T5	60.48	T11	58.51
T6	59.26	T12	59.40

Holzinger et al.[5] and Sklenak et al.[4] are the only studies that have computed chemical shift values for every T site and compared them to experimental ^{27}Al NMR in H-MFI type zeolites. Both studies performed their calculations and measurements on H-ZSM-5. Therefore a direct comparison of the calculated chemical shift values by KRR method is not feasible due to the fact that ZSM-5 has a monoclinic cell structure. This is in contrast to the orthorhombic cell structure of MFI, which contains only 12 T-sites as opposed to ZSM-5's 24 T-sites. To compare the results from their experimental data to

the results obtained by KRR, one can examine the range of chemical shift values. The range of experimental values provided by Sklenak is about 10 ppm in H-ZSM-5, while Holzinger's work provides a range spanning approximately 6 ppm. The acquired range from combining MD simulation with the KRR method of calculating chemical shift is approximately 3.5 ppm, as displayed in section 4.4.1. One can relate the high range of experimental values from Sklenak's study to the low Si/Al ratio in their samples. Although the authors stated that their samples did not contain aluminum pairs based on the use of Co^{2+} exchange experiments, it is highly unlikely that this is the case. The Co^{2+} exchange method is effective only when the aluminum atoms are located within the same ring, thereby allowing for the possibility of NNNN pairs being present.[41]

To test the hypothesis that the high range of experimental values provided by Sklenak may be attributed to the low Si/Al ratio in their samples, a series of 45 models were constructed. These 45 models contained randomly placed aluminum atoms with Si/Al ratios ranging from 15 to 23 and were subsequently loaded with 17 water molecules per unit cell, to accurately mimic the conditions of hydrated experimental samples. The resulting range of calculated chemical shift values was found to be 11 ppm, which closely approximates the experimental range reported in Sklenak's study. Based on these findings it is suggested that high range of experimental ^{27}Al chemical shift can be attributed to presence of close Al atoms in the framework, especially for low Si/Al models.

Holzinger et al.[5] have conducted a thorough study using ZSM-5 samples with a Si/Al ratio of 140, which has a high probability of containing only isolated Al atoms. The range of experimental values observed was slightly over 6 ppm, which is closer to the 3.5 ppm reported in this thesis. However, the higher range reported in Holzinger's work may still be attributed to a small presence of aluminum pairs.

5 Conclusion

The comparative analysis of the performance of various methods of calculating chemical shift with respect to the reference DFT values demonstrated the superior performance of the 2-parameter equation over the simplistic correlation between the TOT angle and chemical shift originally proposed by Lippmaa.[18] These results suggest that the bond length of aluminum and oxygen is a crucial factor in determining the chemical shift of ^{27}Al NMR in zeolites, contrary to findings from Liu et. al.[21] However, the most advanced machine learning approach, the KRR method, was found to outperform all other tested methods for hydrated zeolites by more than 0.5 ppm. For lower water loadings, the experimentally less important scenario, the KRR performance deteriorates, however, it does exhibit similar trends to DFT with some offset that is dependent on both the BAS and water loading.

The data revealed that the quantity of water present in the zeolite system has an impact on the observed chemical shift, with the magnitude of the effect varying depending on the specific T-site. This observation strongly suggests that the presence of water molecules has a significant impact on the local aluminum environment, contradicting one of the rationales for using a background charged and dehydrated models.[4]

It was proposed that for lower water loadings (i.e., 0-3 water molecules) ^{27}Al chemical shift is not directly affected by temperature, but rather through solvation percentage, which is observed to decrease with increasing temperature. In addition, it has been shown that in the fully hydrated zeolite framework the ^{27}Al chemical shift generally decreases with increasing temperature. This behavior cannot be attributed to solvation change with temperature, as at high water loadings (i.e., 17 water molecules per unit cell), the proton is fully solvated at all temperatures. Other factors, such as water dynamics or proton mobility, may contribute to this observed phenomenon and require further investigation for a complete understanding.

The Si/Al ratio has been demonstrated to play a crucial role in the determination of chemical shift values, as the presence of a pair of aluminum in the unit cell can cause a change in the chemical shift value by more than 3 ppm from the value for a single isolated aluminum.

The comparison to the previous experimental and computational studies is not straightforward as H-ZSM-5 zeolite is commonly used in these studies, which has a different symmetry than MFI. Further, comparing to experimental chemical shift values can be challenging due to the possibility of the presence of aluminum pairs in experimental ZSM-5 samples. Hence, it is concluded that the full assignment of calculated chemical shift values to experimental measurements is expected to be difficult due to the narrow separation between the calculated chemical shift shifts, which would lead to spectral overlap in the experimental data.

References

(1) Čejka, J.; Russel, M. E.; Nachtigall, P., *Zeolites in Catalysis: Properties and Applications*; Royal Society of Chemistry, 2017.

(2) Holzinger, J.; Nielsen, M.; Beato, P.; Brogaard, R. Y.; Buono, C.; Dyballa, M.; Falsig, H.; Skibsted, J.; Svelle, S. *J. Phys. Chem. C* **2019**, *123*, 7831–7844.

(3) Dědeček, J.; Lucero, M. J.; Li, C.; Gao, F.; Klein, P.; Urbanova, M.; Tvaruzkova, Z.; Sazama, P.; Sklenak, S. *J. Phys. Chem. C* **2011**, *115*, 11056–11064.

(4) Sklenak, S.; Dědeček, J.; Li, C.; Wichterlova, B.; Gabova, V.; Sierka, M.; Sauer, J. *Phys. Chem. Chem. Phys.* **2009**, *11*, 1237–1247.

(5) Holzinger, J.; Beato, P.; Lundegaard, L. F.; Skibsted, J. *J. Phys. Chem. C* **2018**, *122*, 15595–15613.

(6) Erlebach, A.; Nachtigall, P.; Grajciar, L. *NPJ Comput. Mater.* **2022**, *8*, 174.

(7) Lei, C.; Erlebach, A.; Brivio, F.; Grajciar, L.; Tošner, Z.; Heard, C.; Nachtigall, P. **2023**, preprint available at https://doi.org/10.26434/chemrxiv-2023–ggxfl.

(8) Dib, E.; Mineva, T.; Veron, E.; Sarou-Kanian, V.; Fayon, F.; Alonso, B. *J. Phys. Chem. Lett.* **2018**, *9*, 19–24.

(9) Liang, T.; Chen, J.; Qin, Z.; Li, J.; Wang, P.; Wang, S.; Wang, G.; Dong, M.; Fan, W.; Wang, J. *ACS Catal.* **2016**, *6*, 7311–7325.

(10) Baerlocher, C.; McCusker, L. B. Database of zeolite structures. `http://www.iza-structure.org/databases/`.

(11) Loewenstein, W. *Am. Mineral.* **1954**, *39*, 92–96.

(12) Knott, B. C.; Nimlos, C. T.; Robichaud, D. J.; Nimlos, M. R.; Kim, S.; Gounder, R. *ACS Catal.* **2018**, *8*, 770–784.

(13) Duer, M. J., *The Basics of Solid-State NMR*; John Wiley I & Sons, Ltd, 2001.

(14) Profeta, M.; Mauri, F.; Pickard, C. J. *J. Am. Chem. Soc.* **2003**, *125*, 541–548.

(15) Zhao, Z.; Xu, S.; Hu, M. Y.; Bao, X.; Peden, C. H. F.; Hu, J. *J. Phys. Chem. C* **2015**, *119*, 1410–1417.

(16) Chen, K.; Gan, Z.; Horstmeier, S.; White, J. L. *J. Am. Chem. Soc.* **2021**, *143*, 6669–6680.

(17) Kučera, J.; Nachtigall, P. *Microporous Mesoporous Mater.* **2005**, *85*, 279–283.

(18) Lippmaa, E.; Samoson, A.; Magi, M. *J. Am. Chem. Soc.* **1986**, *108*, 1730–1735.

(19) Van Bokhoven, J. A.; Koningsberger, D. C.; Kunkeler, P.; van Bekkum, H.; Kentgens, A. P. M. *J. Am. Chem. Soc.* **2000**, *122*, 12842–12847.

(20) Dědeček, J.; Sklenak, S.; Li, C.; Wichterlová, B.; Gábová, V.; Brus, J.; Sierka, M.; Sauer, J. *J. Phys. Chem. C* **2009**, *113*, 1447–1458.

(21) Liu, Y.; Nekvasil, H.; Tossell, J. *J. Phys. Chem. A* **2005**, *109*, 3060–6.

(22) Mueller, T.; Kusne, A. G.; Ramprasad, R., *Machine Learning in Materials Science*; John Wiley & Sons, Ltd, 2016.

(23) Qiu, J.; Wu, Q.; Ding, G.; Xu, Y.; Feng, S. *J. Adv. Signal Process.* **2016**.

(24) Sahiner, B.; Pezeshk, A.; Hadjiiski, L. M.; Wang, X.; Drukker, K.; Cha, K. H.; Summers, R. M.; Giger, M. L. *Med. Phys.* **2019**, *46*, DOI: `https://doi.org/10.1002/mp.13264`.

(25) Strieth-Kalthoff, F.; Sandfort, F.; Segler, M. H. S.; Glorius, F. *Chem. Soc. Rev.* **2020**, *49*, 6154–6168.

(26) Behler, J.; Parrinello, M. *Phys. Rev. Lett.* **2007**, *98*, 146401.

(27) Chaker, Z.; Salanne, M.; Delaye, J.-M.; Charpentier, T. *Phys. Chem. Chem. Phys.* **2019**, *21*, 21709–21725.

(28) Jensen, F., *Introduction to Computational Chemistry*; John Wiley and Sons, Inc., 2006.

(29) Hohenberg, P.; Kohn, W. *Phys. Rev.* **1964**, *136*, B864–B871.

(30) Frenkel, D.; Smit, B., *Understanding Molecular Simulation*; Academic Press, Inc.,2001.

(31) Frk, L.; Erlebach, A. Unpublished results, 2022.

(32) BIOVIA, Dassault Systèmes, BIOVIA Materials Studio, 2021.

(33) Schütt, K. T.; Kessel, P.; Gastegger, M.; Nicoli, K. A.; Tkatchenko, A.; Müller, K. R. *J. Chem. Theory Comput.* **2019**, *15*, 448–455.

(34) Van Rossum, G.; Drake Jr, F. L., *Python reference manual*; Centrum voor Wiskunde en Informatica Amsterdam, 1995.

(35) Saha, I.; Erlebach, A.; Nachtigall, P.; Heard, C. J.; Grajciar, L. **2023**, preprint available at https://doi.org/10.26434/chemrxiv-2022–d1sj9.

(36) Hjorth Larsen, A. et al. *J. Phys. Condens. Matter.* **2017**, *29*, 273002.

(37) Nosé, S. *J. Chem. Phys.* **1984**, *81*, 511–519.

(38) Trachta, M.; Bulánek, R.; Bludský, O.; Rubeš, M. *Sci. Rep.* **2022**, *12*, 7301.

(39) Hoffman, A.; DeLuca, M.; Hibbitts, D. *J. Phys. Chem. C* **2019**, *123*, 6572–6585.

(40) Dědeček, J.; Tabor, E.; Sklenak, S. *ChemSusChem* **2019**, *12*, 556–576.

(41) Dědeček, J.; Kaucký, D.; Wichterlová, B.; Gonsiorová, O. *Phys. Chem. Chem. Phys.* **2002**, *4*, 5406–5413.