JCTC Journal of Chemical Theory and Computation

pubs.acs.org/JCTC Article

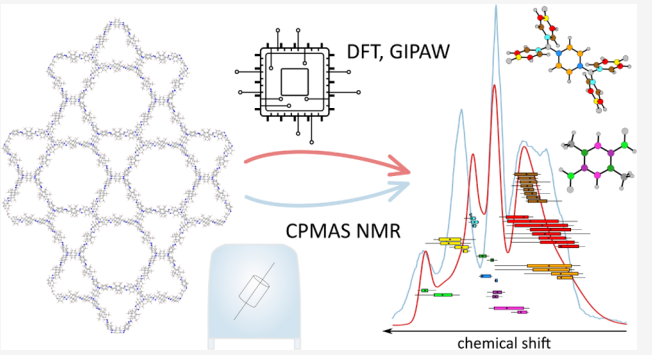
Computational Protocol for the Spectral Assignment of NMR Resonances in Covalent Organic Frameworks

Siebe Vanlommel, Sander Borgmans, C. Vinod Chandran, Sambhu Radhakrishnan, Pascal Van Der Voort, Eric Breynaert, and Veronique Van Speybroeck*

Cite This: https://doi.org/10.1021/acs.jctc.3c01414 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Solid-state nuclear magnetic resonance spectroscopy is routinely used in the field of covalent organic frameworks to elucidate or confirm the structure of the synthesized samples and to understand dynamic phenomena. Typically this involves the interpretation and simulation of the spectra through the assumption of symmetry elements of the building units, hinging on the correct assignment of each line shape. To avoid misinterpretation resulting from library-based assignment without a theoretical basis incorporating the impact of the framework, this work proposes a first-principles computational protocol for the assignment of experimental spectra, which exploits the symmetry of the underlying building blocks for computational feasibility. In this way, this protocol accommodates the validation of previous experimental assignments and can serve to complement new NMR measurements.



1. INTRODUCTION

Covalent organic frameworks (COFs) form a promising type of tunable porous materials which can be tailor-made for applications in heterogeneous catalysis, electrochemistry, sensing, biomedical, and gas storage.[1−9] Characterizing the atomic-level structure of COFs, in particular, their local chemical structure, functionalization and layer stacking, is essential to understand the material's macroscopic properties and to leverage the material toward the applications.[10,11] However, this characterization is not a trivial task in COFs, given the limited long-range order of typical samples. While advanced physicochemical characterization has proved extremely useful,[12−15] experimental methods are sometimes limited to resolving features that are periodic in nature. For example, X-ray diffraction (XRD) may reveal a long-range order and 3D stacking of nodes and linkers and layer stacking in the COF structures, but it cannot provide access to the local chemical environment of the functional groups or assess the extent to which all links in the reticular network have been formed. Therefore, complementary characterization techniques are required, as they can operate at the atomic level and provide atomic level information. In this respect, combining solid-state nuclear magnetic resonance (NMR) spectroscopy with its computational counterpart is an obvious candidate methodology that can provide element-specific information on the local structure of the atomic framework sites and guest species.[16−23] In this work, we take a look at a specific case study of a COF that previously has been well-characterized through a combination of experimental XRD, ^{13}C NMR and computational XRD.[24] Specifically, we calculate the theoretical NMR chemical shifts and propose different methods to compare the NMR spectra to experiment. The incentive for this work was the routine library-based assignment of experimental NMR resonances supported by chemical intuition and (assumed) prior knowledge.[25−30] As we show in this paper, the assumptions on the symmetry of the carbon sites can be wrong, and peaks may be assigned to the wrong resonance in the spectrum. As such, a more systematic and first-principles-based computational method is proposed in order to improve peak assignments and bolster the spectral interpretation. The correct assignment of NMR spectra is a prerequisite to the correct interpretation of spectral changes, for example, during a phase transformation or during catalysis.[31−33] A preliminary version of the methodology described in this paper has previously been applied in relation to the assignment of ^{13}C NMR resonances of two newly synthesized COF systems (TpBpyCOF and TpPpyCOF) with applications in metal-free photocatalysis.[34]

Received: December 25, 2023
Revised: April 2, 2024
Accepted: April 5, 2024

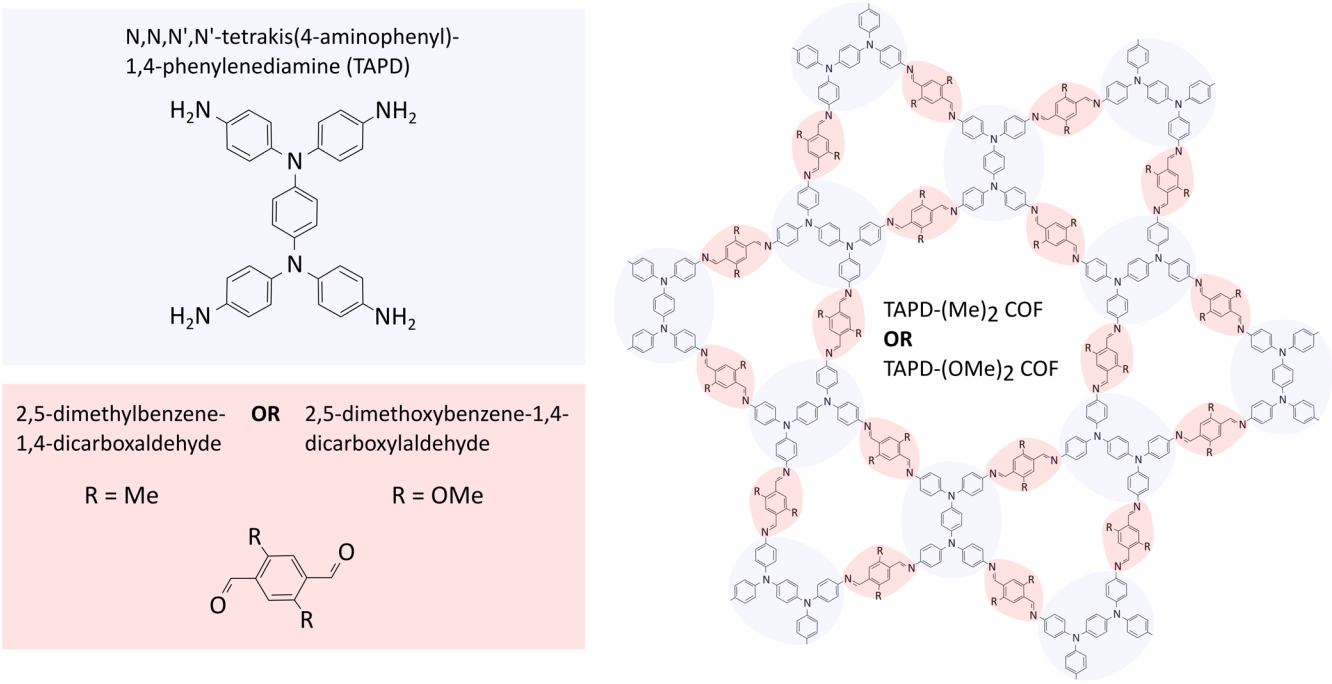


Figure 1. Chemical structure of the building blocks (left) and the final COF lattice (right).

In other work, two newly designed two-dimensional COFs were reported for the production of H_2O_2, whose structure was determined through a combined experimental−

computational approach.[35] These are built up out of *N,N,N′,N′*-tetrakis(4-aminophenyl)-1,4-phenylenediamine nodes (henceforth referred to as TAPD) and either 2,5-dimethylbenzene-1,4-dicarboxaldehyde linkers (referred to as Me) or 2,5-dimethoxybenzene-1,4-dicarboxylaldehyde (referred to as OMe). Figure 1 shows the chemical building blocks and the structure of one layer of the final COF material, exhibiting a Kagome lattice. The full synthesis of the COFs and the characterization of their crystallinity may be found in the original paper.[35] The focus of the current work is the computational modeling of NMR spectra of these COF materials. Here, we present and discuss three different computational methodologies to produce theoretical ^{13}C NMR spectra which can be compared to experimental spectra (Section 2). All methodologies use the same computational chemical shielding data, which are derived from optimized structural models, calculated using density functional theory. The chemical shielding data are derived from a static (0 K) structure, as performing molecular dynamics (MD) for larger sets of structures could become prohibitively expensive. Even though MD plays a huge role in the further development of operando spectroscopic techniques in nanoporous materials,[36−38] we aim to develop a computational protocol that is relatively cheap, hence we prefer static methods over dynamic ones in this case due to the large system size. The larger the system size, the better the methodology works, as increasing the number of atoms in the unit cell will lead to better statistics in the averaging schemes of the methodology and a representation of the true system. The difference between the three proposed methodologies lies in their interpretation of the chemical shifts and the underlying averaging scheme used to emulate the inherent time-averaging of an experimental measurement. Similar approaches have been previously presented for the prediction of powder XRD patterns.[24] The COF systems under investigation, their experimental ^{13}C NMR spectra and the peak assignments are briefly discussed in Section 3.1. Using our three proposed schemes and a dynamic model, we calculate the theoretical ^{13}C NMR spectra of the experimentally reported TAPD-$(Me)_2$ COF and compare their relative efficacy (Section 3.1). Then, we perform a full assignment of the peaks of the experimental TAPD-$(Me)_2$ ^{13}C NMR spectrum and compare this to the previous assignment as was performed in the experimental study (Section 3.2). We show that spectral assignments based on chemical intuition cannot always provide an indisputable assignment of the resonances, therefore, emphasizing the need for computational models to complement experimental data. Furthermore, as NMR is extremely sensitive to the local electronic structure around individual nuclei, computational NMR offers us information on diversity in local structure around carbon atoms and how this influences the chemical shift (Section 3.2). As a second validation, this procedure has also been applied for the spectral assignment of the ^{13}C NMR spectrum of a related material, the TAPD-$(OMe)_2$ COF (Section 3.3). For both COF materials, $^{1}H-^{15}N$ CPMAS NMR spectral measurements were performed with which an assignment of the ^{15}N NMR spectra is provided using the newly developed computational protocol (Section 3.4). ^{15}N NMR typically suffers from the poor natural abundance of ^{15}N (0.368%) requiring very long measurement times to achieve a minimal signal-to-noise ratio, and we discuss the implications of this aspect for the efficacy of the protocol. Finally, as the stacking configuration in 2D COFs is regulated by the local arrangement of neighboring layers, the predictive power toward identifying the layer stacking through ^{13}C NMR spectra is discussed in Section 3.5.

2. METHODOLOGY

The initial atomic geometries were derived from the corresponding (previously performed) force field MD

https://doi.org/10.1021/acs.jctc.3c01414
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

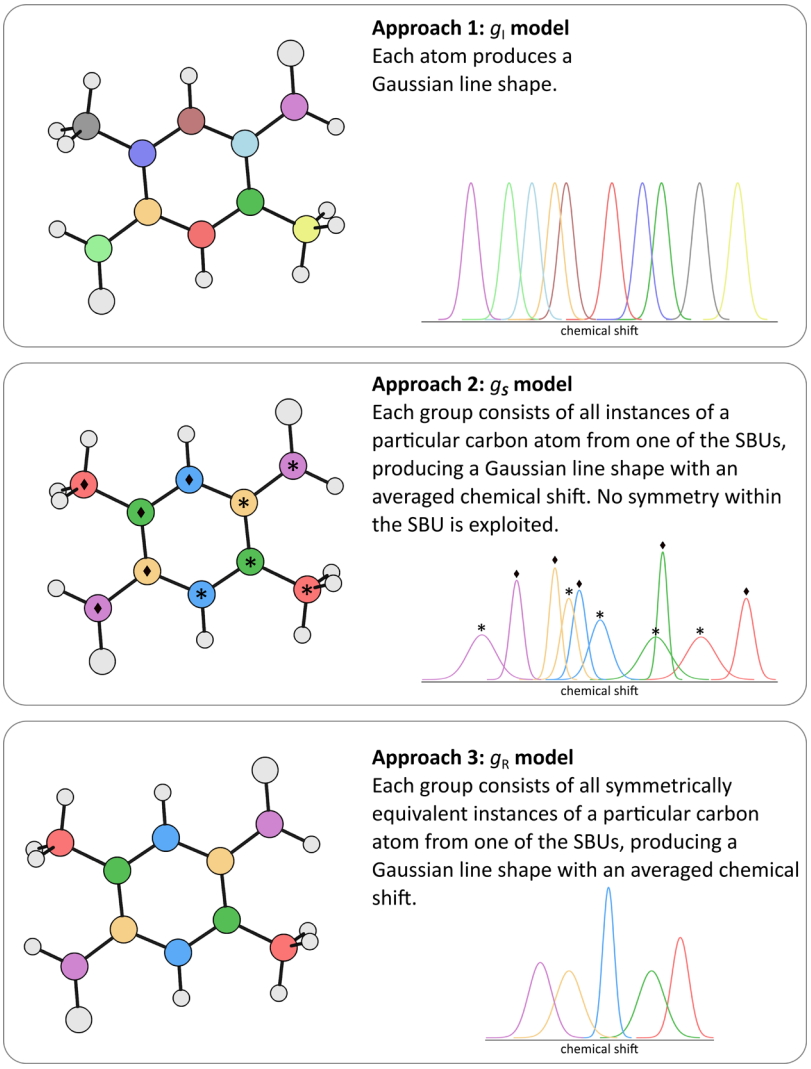


Figure 2. Derivation of Gaussian line shapes in the three approaches summarized using a generic example of a linker SBU. The spectra shown are purely hypothetical for illustration purposes.

simulations by extracting a two layer unit cell from a single snapshot, and fully relaxing the structure using VASP within the PAW formalism.[39,40] The computational details of the FFMD simulations may be found in the Supporting Information of the paper by Krishnaraj et al.[35] (Section S3 therein). NMR chemical shieldings were then obtained through the DFT-GIPAW[41] approach implemented in VASP[42−45] under periodic boundary conditions. Computational settings for geometrical optimization of atomic positions and cell shape and for linear response calculations were as follows: PBE functional[46,47] with DFT-D3(BJ) dispersion correction,[48,49] energy cutoff for plane waves of 500 eV, sampling of the Brillouin zone was limited to the Γ point. All electron density symmetrizations were switched off.

Here, we propose three different ways in which a computational NMR spectrum may be constructed: (i) the uniform model, (ii) the secondary building unit (SBU) average model, and (iii) the symmetrically inequivalent average model. These three approaches represent the extremes of (i) directly using the individual NMR parameters of each carbon atom in the system and (iii) grouping the carbon atoms into chemically equivalent classes and using NMR parameters per carbon class averaged over the entire system, along with the intermediate method of (ii) using NMR parameters for every carbon atom within the SBUs, averaged over all SBUs. In each approach, the spectrum is built up from the summation of Gaussian line shapes (which is a typical choice for spin-$\frac{1}{2}$ systems), determined by two parameters: the position of the peak μ along the chemical shift axis x and the width of the peak σ. In what follows, the three different models are further explained. The spectra that result from these three methods are illustrated for a generic system in Figure 2.

2.1. Approach 1: The Uniform Model. The total spectrum (which is denoted g_I as it constitutes individual atomic sites) is obtained by summing over all atomic sites

$$g_I(x) = A\sum_i \frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{(x-\mu_i)^2}{2\sigma^2}\right) \quad (1)$$

where A denotes an overall amplitude of the spectrum. The index i runs over all of the atoms of the isotope under investigation. In the first approach, the width of the peaks σ is the same for all atoms and can be chosen arbitrarily (or fitted to the experimental spectrum). In a sense, the choice of σ reproduces the intrinsic peak width produced by the NMR

https://doi.org/10.1021/acs.jctc.3c01414
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

nucleus convoluted with broadening effects induced by a variety of factors including field (in)homogeneity, strength of dipolar couplings, magic angle spinning, etc. but excludes chemical shift broadening because every atom in the computational system contributes a single Gaussian resonance to the total spectrum.

2.2. Approach 2: The SBU Average Model. The computational NMR resonance that is calculated for each of the atoms individually in the first approach is not necessarily the resonance that we expect to be present in the experimental spectrum. This is because the system is dynamic, taking each individual atomic site through a plethora of states with varying local geometry. As such, the geometry of a carbon site in a single static structure is not necessarily representative of the average geometry of that site in the true experimental system. This is similar to the previously studied case of ^{27}Al NMR in zeolites,[37] only in this case the nature of the effects that average the resonance is not the presence and mobility of water and extra framework species, but the variations in the local geometry of the structure. Note that it is not even guaranteed that, for example, an MD run takes all of the atomic sites through all the relevant states in a time that is computationally feasible to attain. Therefore, we propose an alternative approach that relies on the symmetry of the system to improve the comparison to the experiment. For this second approach, we group the carbon atoms in the system into symmetrically inequivalent classes and derive one representative average NMR resonance per group. In this way, even in a static model, the Gaussian resonances that make up the total spectrum are composed of multiple resonances on their own in an average manner. As such, through the derivation of an average location and width for each computational resonance, the geometrical diversity within the experimental system can be mimicked. The total spectrum in this second approach, denoted g_S to emphasize the use of symmetry, is composed of a sum over symmetrically inequivalent carbon sites

$$g_S(x) = A\sum_n \frac{1}{\gamma\sigma_n\sqrt{2\pi}}\exp\left(-\frac{(x-\langle\mu\rangle_n)^2}{2(\gamma\sigma_n)^2}\right) \quad (2)$$

where $\langle\mu\rangle_n$ denotes the chemical shift averaged over all atoms belonging to the class n. The classes are chosen as follows: each carbon atom in the nodes and linkers (being the SBUs in the system) that make up the COF represents a single symmetry class. The average NMR resonance for that class is derived by averaging the NMR chemical shielding over all building blocks in the system (effectively, this is an average over all the nodes or all the linkers). The carbon labels are shown in Figure 3, which contains the three building blocks that are used to construct the system under investigation. The TAPD-$(Me)_2$ COF is produced through combining the TAPD node and methylated linkers (denoted as Me), whereas the TAPD-$(OMe)_2$ system is built with methoxy group linkers instead (OMe). The atoms are colored according to the symmetry of the molecules. Note that multiple atoms have the same color, which means that in the building blocks these atoms are symmetrically equivalent. However, in the g_S model, we retain these atoms as being symmetrically inequivalent, because their geometrical environment inside the COF material can be different irrespective of the molecular symmetry of the underlying building blocks. The number of symmetrically inequivalent classes of carbon atoms in the COF material is therefore not the same as in the underlying building blocks. In total, there are 11 symmetrically inequivalent carbon atoms in the nodes and linkers, while we use 40 inequivalent carbon classes in the COF (for the 40 carbon sites that are present in node + linker). This gives us 40 NMR line shapes that are summed in the spectrum g_S. The chemical shift of a carbon class in the node is an average over the six nodes that are present in the unit cell. For a carbon class in the linker, the average comprises 12 linkers. For the widths $\gamma\sigma_n$, we use the natural choice of the standard deviation of the chemical shieldings from their average value $\langle\mu\rangle_n$, scaled by an overall factor of γ. This γ factor tunes the overall width of the spectrum without changing the relative widths of the line shapes, as these are determined by the computational data in this model. We show that this second approach improves the comparison to the experimental spectrum. As an error measure, we take the mean squared error (MSE) of the computational spectrum with respect to the experimental spectrum, normalized by dividing by the integral of the experimental spectrum.

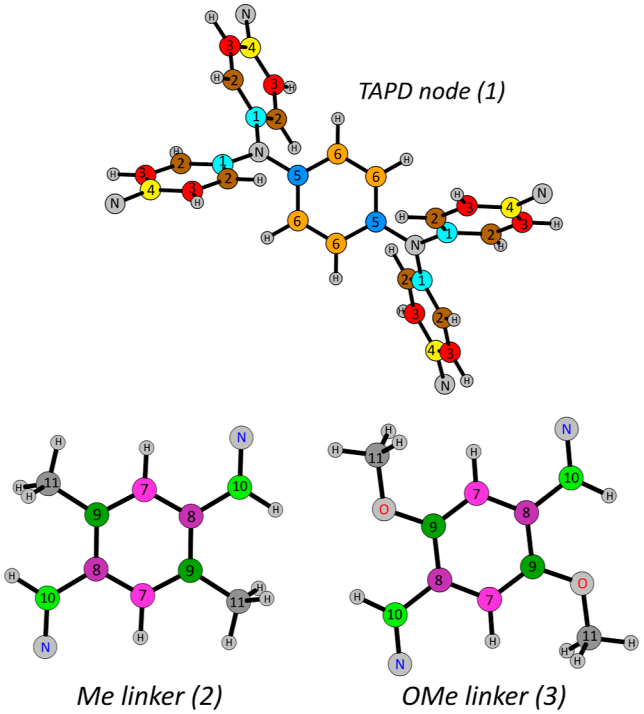


Figure 3. Definitions of the TAPD node, Me linker and OMe linker, where the carbon atoms are colored to show the division into symmetrically inequivalent classes. Nitrogen and oxygen atoms are denoted N and O, respectively.

2.3. Approach 3: The Symmetrically Inequivalent Average Model. As a third option to produce a spectrum, this approach is a variant of the SBU average model, where we reduce the number of symmetry classes by grouping the carbon sites that are chemically equivalent when looking at the symmetry of the nodes and linkers. In that way, the number of carbon classes is drastically reduced to 11. The resonances of the carbon classes are then the average over all atoms with the same color label (same symmetry in the building block), and the line shape width is the standard deviation of all the individual shift values with respect to the average of the symmetry class. We denote the spectrum that is produced in this way g_R, for the reduced number of classes that is used. In this approach, we also have to correct the amplitudes of the

https://doi.org/10.1021/acs.jctc.3c01414
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

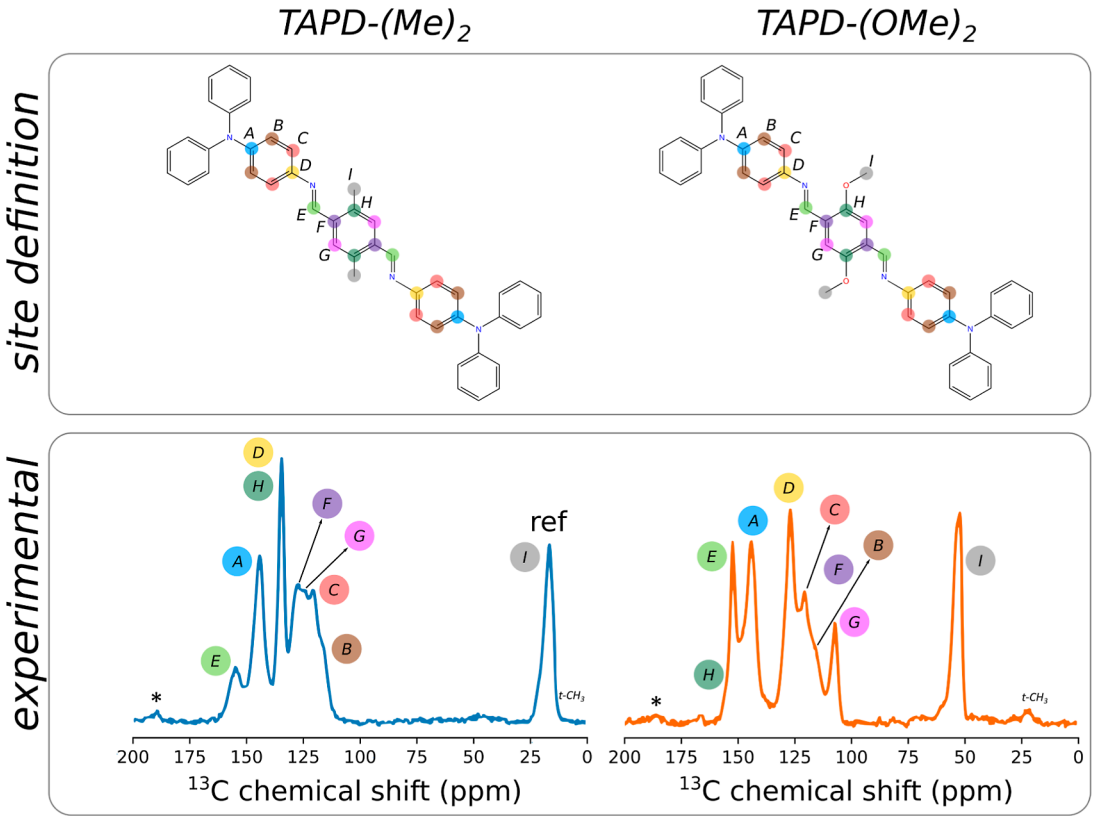


Figure 4. Experimental ^{13}C NMR spectra of the TAPD-$(Me)_2$ and TAPD-$(OMe)_2$ COFs with the assignment of the peaks to specific carbon sites as was performed by the authors of the original paper. Spinning sidebands are indicated with *. Residual terminal Me group carbon contributions are denoted *t-CH*$_3$. The experimental NMR data used for this work is reprinted with permission of ref 35, copyright 2022 ACS.

carbon classes, as not all of the classes comprise the same number of resonances. Therefore, we scale the line shapes with the number of carbon sites belonging to that class (these amplitudes A_n are not parameters that are optimized, because the relative contributions of the line shapes to the spectrum must scale with the number of atoms belonging to that class in the system). The resulting spectrum is therefore

$$g_R(x) = A\sum_n \frac{A_n}{\gamma\sigma_n\sqrt{2\pi}}\exp\left(-\frac{(x-\langle\mu\rangle_n)^2}{2(\gamma\sigma_n)^2}\right) \quad (3)$$

with the sum now running over the 11 carbon classes as indicated by the numbers and colors in Figure 3.

The three approaches, along with how they distinguish groups of atoms and their use of symmetry, are summarized in Figure 2. In the first approach, there are four parameters that can be changed independently to minimize the MSE between the computational g_I spectrum and the experimental spectrum: the overall amplitude A, the width of the Gaussians σ and the two parameters a and b in the linear shift-shielding relation $\delta_{iso} = a\sigma_{iso} + b$ that is used to convert the calculated isotropic shielding values σ_{iso} to isotropic chemical shifts δ_{iso}. In the second approach (g_S), these parameters are the same, except that each symmetry class now has its own characteristic width that is calculated as the standard deviation on the data in each class, so that the parameter we tune is an overall scaling factor γ that scales all of the widths equally in order to retain their relative magnitudes. Similar to the fitting of the σ parameter in the uniform model, it makes sense to fit γ to the experimental spectrum, as in the experimental case, there are broadening effects other than chemical shift broadening. All computational NMR spectra that we produce are quantitative, in the sense that the integral of the line shapes is proportional to the amount of carbon atoms in the system, which have that particular NMR resonance. In the experimental CPMAS spectrum, this is not necessarily so. We refer the reader to Section S1 of the Supporting Information for a brief discussion on this property. Finally, the third approach has the same parameters to be fit as those of the second approach.

As stated before, the static model takes into account only a single geometrically optimized configuration of the system. We can also take into account motion of the atoms by taking a set of snapshots of the system from a (force field or first principles) MD run, calculating the NMR chemical shifts of the atoms in the system through DFT-GIPAW, and averaging the NMR spectrum over time by time-averaging the chemical shifts for each individual site. We can then use the time-averaged NMR chemical shifts to produce a spectrum similar to the static NMR chemical shift values. To assess whether the static g_S approach succeeds in producing chemical shifts that mimic the experimental geometrical diversity, we will also derive a spectrum using the g_S approach but with chemical shifts originating from a dynamic model. This spectrum is denoted $\overline{g}_S$, with the $\overline{g}$ emphasizing that the chemical shieldings used in the model are time-averages derived from a force field MD run. Generally this dynamic approach takes a lot of computational resources (either for developing a suitable force field to perform MD or through first-principles MD). We refer the reader to Section S2 of the Supporting Information for a comparison between a spectrum produced with time-averaged chemical shifts versus chemical shifts derived from a static

https://doi.org/10.1021/acs.jctc.3c01414
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

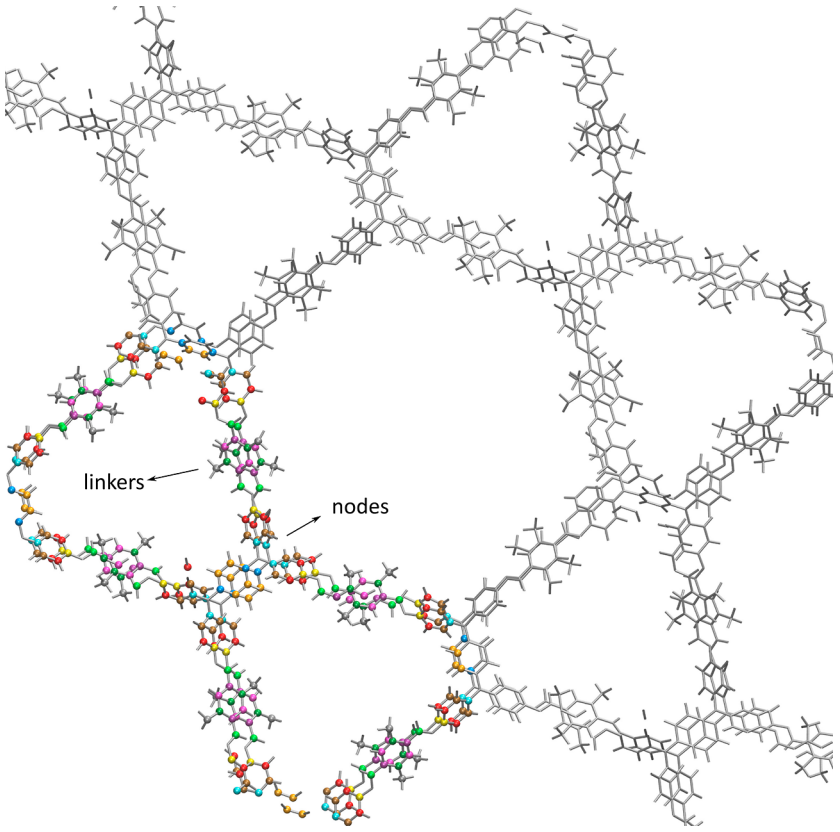


Figure 5. Two layers of TAPD-$(Me)_2$ COF material in the inverted stacking scheme which was previously found to be exhibited experimentally. The carbon atoms contained within the unit cell used in the calculations are colored according to their corrected symmetry class defined in this work.

model. The results validate the approach of averaging over carbon atoms as a way to mimic averaging chemical shifts over time.

3. RESULTS

3.1. Overview of the Three Models for the TAPD-$(Me)_2$ COF. As illustrated in the original experimental peak assignment of Figure 4, several atoms are lumped together in the labeling. This is a simplification that does not reflect that the three carbon atoms that are bound to the same nitrogen atom (experimental carbon site *A* in Figure 4) are not symmetrically equivalent. Two additional symmetrically inequivalent classes must be introduced to include all carbon atoms, as introduced in Figure 3. To illustrate the connectivity in the TAPD-$(Me)_2$ framework, Figure 5 shows part of the extended framework (all atoms of which are colored gray) as well as the unit cell used in the calculations, for which the carbon atoms are colored according to their correct symmetry class as defined in Figure 3. The framework of the TAPD-$(OMe)_2$ COF, which will be discussed later, is not shown as the connectivity is essentially the same, except for the methyl groups (CH_3) being replaced with methoxy groups (OCH_3). In what follows, carbon labels that were originally introduced in the experimental assignment of the spectrum will be denoted with letters *A* through *I*. The extended (corrected) carbon labels introduced in this work are denoted with numbers 1 through 11 (Figure 3), in order to clearly differentiate them from the experimental labeling.

With these corrected symmetry labels, the performance of our three approaches, defined in Section 2, can be assessed. Figure 6 shows the computational spectra obtained using the three approaches (g_I, g_S, and g_R) plotted against the experimental spectrum of the TAPD-$(Me)_2$ COF.

Before we look at the assignments of the resonances, we first consider how the three models compare. Visually, all three approaches produce a spectrum that is close to the experimental TAPD-$(Me)_2$ spectrum. Whether we sum Gaussian line shapes at locations based on individual atomic chemical shifts (g_I) or whether we do this using average chemical shifts (g_S or g_R): the overall location of the peaks is similar to experiment. There are, however, some subtle yet important differences. The relative error which is reported in the bottom right table of Figure 6 shows that the g_S model spectrum is the closest (smallest MSE error) to the experimental spectrum, followed closely by the g_R model, which has only a +3% higher MSE error. The g_I approach however gives a significantly higher MSE error with respect to experiment, with a value being +27% higher than the g_S model. This shows the merit of averaging chemical shifts based on the symmetry of the underlying building blocks: both models succeed partly in implicitly taking into account the mobility of the atoms. Looking at the γ parameter then, we see that the overall width of the Gaussians in the g_I model is 1.84 ppm (recall that the initial line width in the spectrum is arbitrarily chosen to be 1.0 ppm, therefore the γ parameter being 1.84 means the overall width is 1.84 ppm for all line shapes in this approach). For the g_S and the g_R models, the initial line widths of the spectrum (before fitting γ) are not chosen arbitrarily, rather they are the standard deviation of the chemical shifts in each symmetry class with respect to the average value. This

https://doi.org/10.1021/acs.jctc.3c01414
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

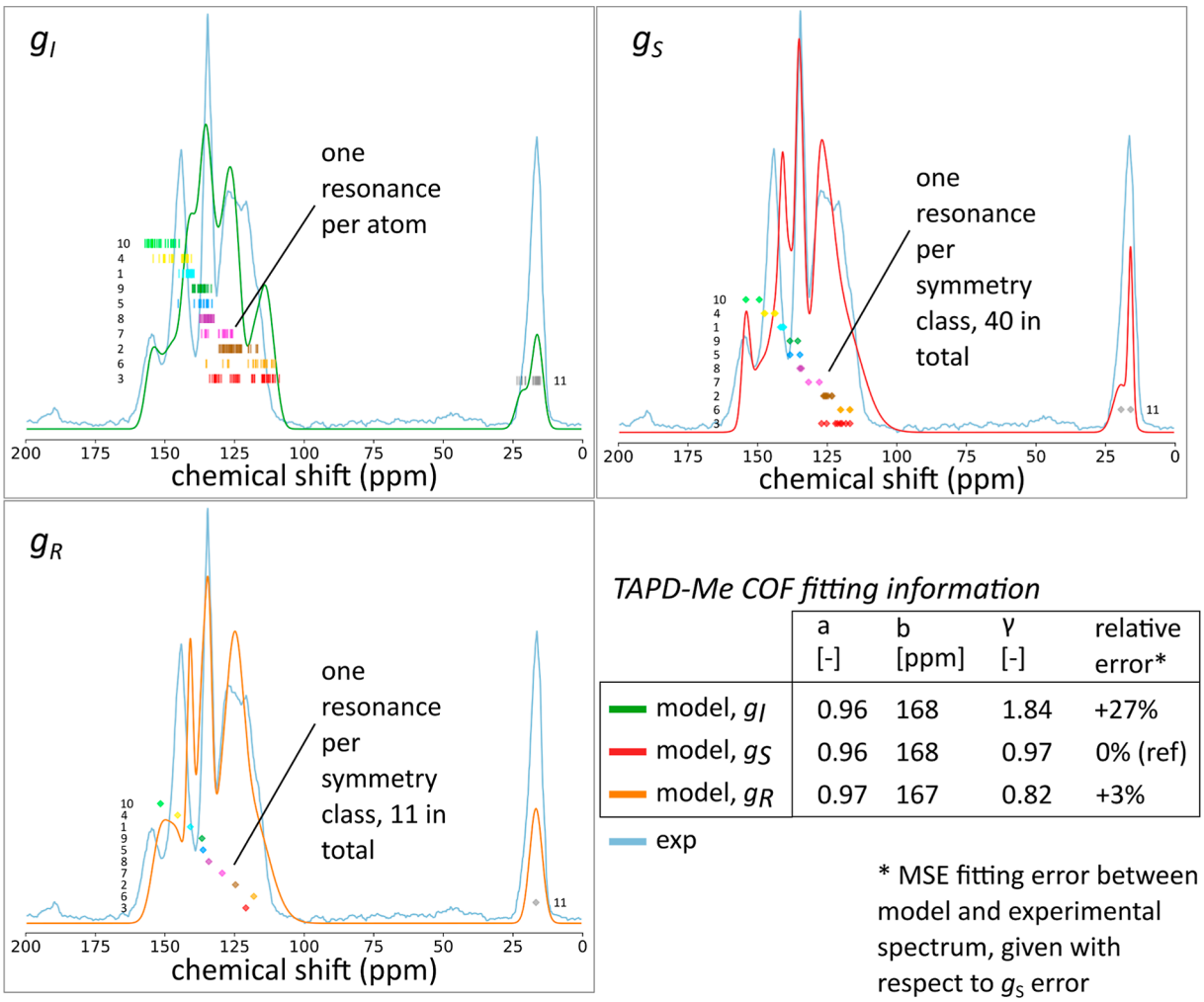


Figure 6. Experimental and computational g_I, g_S, g_R^1 ^{1}H-decoupled ^{13}C CPMAS NMR spectra of the TAPD-$(Me)_2$ COF. The bottom right panel shows the parameters fitted to minimize the MSE between the experimental and the model spectrum. The parameters a and b denote the fitting parameters of the linear shift-shielding relation ($\delta_{iso} = a\sigma_{iso} + b$) and γ is the overall multiplier of the widths used in the spectra. The overall amplitude A is not shown, as it carries no relevant information.

results in the γ parameter having a value closer to 1.0, showing that this standard deviation is a natural and logical choice for the line widths. Choosing a single line width for all carbon resonances as in g_I is therefore not very realistic, and the models g_S and g_R, with variable line shapes per carbon class, more closely reflect the real situation, in line with their variable local environment. The value for γ of 0.97 for the g_S model means that barely any correction to the width is necessary to minimize the MSE error between the g_S model and experiment. For the g_R model, however, the line width is corrected downward, with the γ value being 0.82. This means that we are grouping the carbon atoms together too much, in turn leading to a larger spread on the chemical shifts with respect to the average value, which increases the line width of the line shape. It is also worth looking at the methyl resonance (carbon class 11, gray). For the g_I and g_S models, this resonance is asymmetrical, which is also the case in the experimental spectrum. For the g_R model, however, there is only a single line shape to produce this resonance, and therefore by definition it is symmetrical, in disagreement with experiment. It seems that the g_S model, which treats every atom in the underlying building blocks as representing a separate symmetry class, performs best in reproducing the spectrum by averaging some information away while retaining enough diversity. The g_R model averages away too much information, resulting in line shapes that are too wide and too symmetrical. The g_I model on the other hand does not average any information and produces line shapes that are not broad enough compared to the experimental data. Additionally, the large spread on the chemical shifts clouds the assignment of line shapes, as will be discussed later.

3.2. Corrected Assignment of the TAPD-$(Me)_2$ ^{13}C NMR Spectrum Using the g_S Approach. Focusing now on the ^{13}C NMR spectrum, using the SBU average approach (g_S) due to its optimal reproduction of the experimental spectrum, Figure 7 compares the previously shown experimental assignment (top) to the computational g_S spectrum (bottom). All carbon classes are present in the experimental assignment, and their assigned location can be compared to our calculations, except for carbon class A, which must be split into three classes, namely 1, 5, and 6.

The resonance with the highest chemical shift near 154 ppm is the one with carbon label 10, which is assigned correctly (class E). The next peak at 141 ppm is attributed to carbon sites of label 1, which is correct except that there are two other types of carbon atoms, which are incorrectly given the same

https://doi.org/10.1021/acs.jctc.3c01414
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

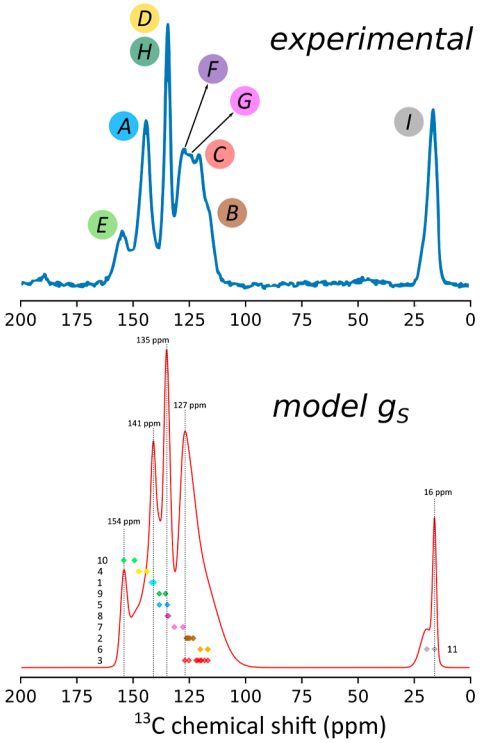


Figure 7. Comparison between previous experimental assignment of the TAPD-$(Me)_2$ COF versus g_S model assignment of this COF. Chemical shift values indicated on the bottom plot are not the only line shapes of the spectrum, but are merely present as a guide to the reader's eye.

label *A*. Additionally, the peak at 141 ppm is also in part due to atoms with carbon label 4 (yellow), which experimentally was incorrectly attributed to the peak near 135 ppm in the spectrum. The central peak near 135 ppm is not due to carbon label 4, rather it is due to label 9 (dark green, assigned correctly), and due to label 5, which is not present in the experimental classification. Next, there are resonances that largely overlap and contribute to the peak with a maximum near 127 ppm and showing a broad shoulder until 100 ppm. The higher ppm values of this peak are due to carbon labels 8 and 7 (highest shift values in the peak, correctly assigned experimentally). Then, there is the contribution of carbon label 2, and last label 6 and label 3. Label 6 is missing in the experimental assignment. Label 2 was assumed to contribute to the rightmost part of the peak with lowest ppm values, which is incorrect. Rather, label 6 contributes the most near 125 ppm, with some contribution also coming from label 3. The rightmost part of the peak is due to label 6 and label 3. Lastly, the methyl resonance is of course the only contributor for the peaks near 16 ppm, where we note an asymmetrical shoulder toward 25 ppm, which is also present in the g_S model.

Furthermore, the data show that assigning the features of the experimental spectrum to specific carbon classes can be ambiguous. Even though carbon sites may be chemically equivalent in the underlying building block, they can still show a large difference in the chemical shift. Even if the chemical shift is averaged out over multiple building blocks in the system, the resulting chemical shift can strongly differ, depending on the exact location of the carbon site in the building block. The striking example of this is the assignment of carbon label 3. The resonances originating from these sites may be found in a much broader part of the spectrum, which comprises the entire range of chemical shifts, where labels 2 and 6 are also found, meaning that they overlap. Clearly, the chemical shifts of carbon sites that, based on the symmetry of the underlying building blocks, are equivalent can be found in a larger range than was expected experimentally because experimentally these were assigned to one specific location in the spectrum. The origin of this large range in which the chemical shifts may be found is geometrical in nature, following the flexibility of the imine linkages, and we can show this by taking a look at the local geometry around the

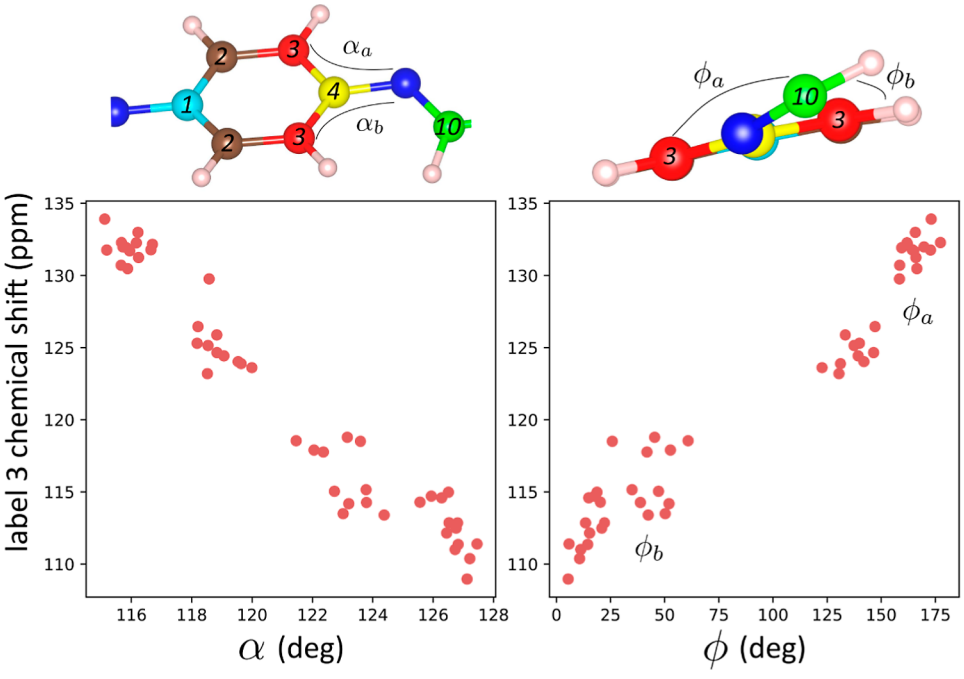


Figure 8. Correlation between two geometrical properties defined for the carbon sites 3 with their theoretical chemical shift.

https://doi.org/10.1021/acs.jctc.3c01414
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

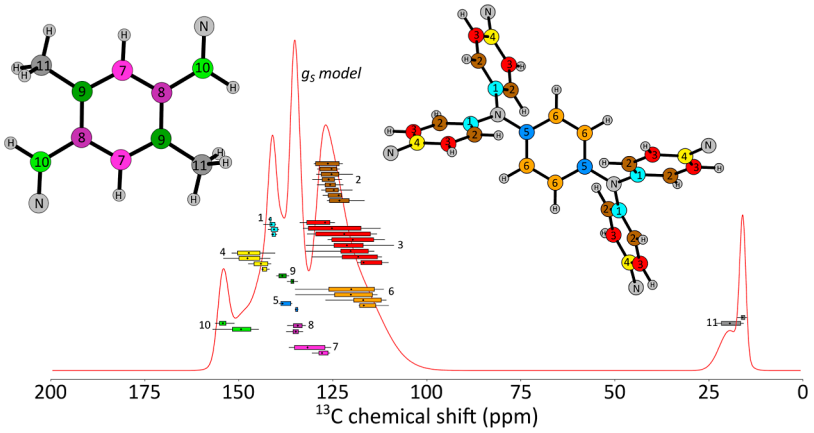


Figure 9. Model g_S spectrum of the TAPD-$(Me)_2$ COF with boxplots to show the underlying distribution of the chemical shifts. The average chemical shift is plotted with a diamond symbol (♦).

carbon sites with label 3 (red). Figure 8 shows the chemical shifts of the carbon atoms within carbon label 3, plotted against two geometrical parameters: the C3−C4−N angle (α) and the C3−C4−N−C10 dihedral angle (ϕ). Both of these may be defined for any carbon atom with label 3, and they both have a strong correlation with the chemical shift of the carbon site. This also translates to the averages: the carbon classes within label 3 can have, on average, a larger or smaller angle α (or, conversely, a smaller or larger dihedral ϕ), which will affect the average chemical shift of the class. The geometrical diversity in the system leads to a large spread on the chemical shifts, and some carbon classes, such as the class with label 3, may contribute to different parts of the spectrum. This illustrates the fact that while all C3 sites are chemically equivalent in the underlying building block, they have their own unique geometrical environment within the COF framework and can show a broad range of chemical shifts. This results in the fact that spectral features cannot necessarily be assigned to single carbon classes; rather, one must consider the possibility of significant overlap between carbon classes. This possibility is generally not considered in the assignment of carbon spectra in experimental studies, which again emphasizes the need to complement spectral interpretation with theoretical calculations. The same effect could be described for label 2 and the angle C2−C1−N, for example.

The previous discussion leads us to the fact that it is interesting to show the diversity in the g_S model of the underlying chemical shifts of the line shapes. For this reason, one could also consider the distribution of chemical shifts per line shape through a boxplot, as illustrated for the TAPD-$(Me)_2$ COF in Figure 9. Such plots are more informative than just plotting the average chemical shifts as they give an idea of the underlying distribution of chemical shifts that produce the average. Additionally, it is immediately clear which carbon sites reside in more geometrically (and therefore potentially chemically) diverse framework sites.

3.3. Corrected Assignment of the TAPD-$(OMe)_2$ ^{13}C NMR Spectrum Using the g_S Approach. With the previous discussions in mind, we turn our attention to the TAPD-$(OMe)_2$ COF and immediately consider the g_S model together with the boxplot representation to get the correct assignment of the carbon spectrum and subsequently compare to experiment (the g_I and g_R model spectra may be found in Section S3). Figure 10 shows the experimental spectrum and its assignment (top), together with the g_S model spectrum and

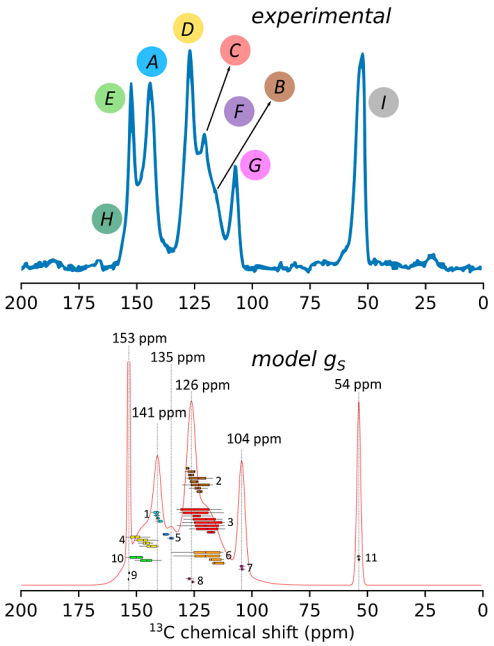


Figure 10. Comparison between previous experimental assignment of the TAPD-$(OMe)_2$ COF versus g_S model assignment of this COF. In this figure the boxplots describing the underlying chemical shift distribution is also shown.

corrected assignment with boxplots to show the diversity in underlying chemical shifts (bottom). Again, aside from the missing carbon classes, there is a clear discrepancy between several carbon assignments. The leftmost resonance near 153 ppm is indeed caused by carbon label 9 (correctly assigned experimentally). Immediately next to that, in between the 153 and 141 ppm peaks, carbon label 10 produces the spectrum, together with the misassigned carbon label 4 (similar to the TAPD-$(Me)_2$ COF). Next, carbon label 1 is correctly assigned to the peak near 141 ppm, complemented by contributions from the missing carbon class 5 to the same peak, as the neighboring peak at lower ppm values. The central peak at 126 ppm and its broad shoulder toward 100 ppm is similar to the TAPD-$(Me)_2$ COF spectrum: carbon label 2 mostly produces the central peak at 126 ppm, with contributions from label 3 and label 8. As such, label 8 is misassigned as well in this case (dark purple dot in the experimental spectrum). Again the

https://doi.org/10.1021/acs.jctc.3c01414
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

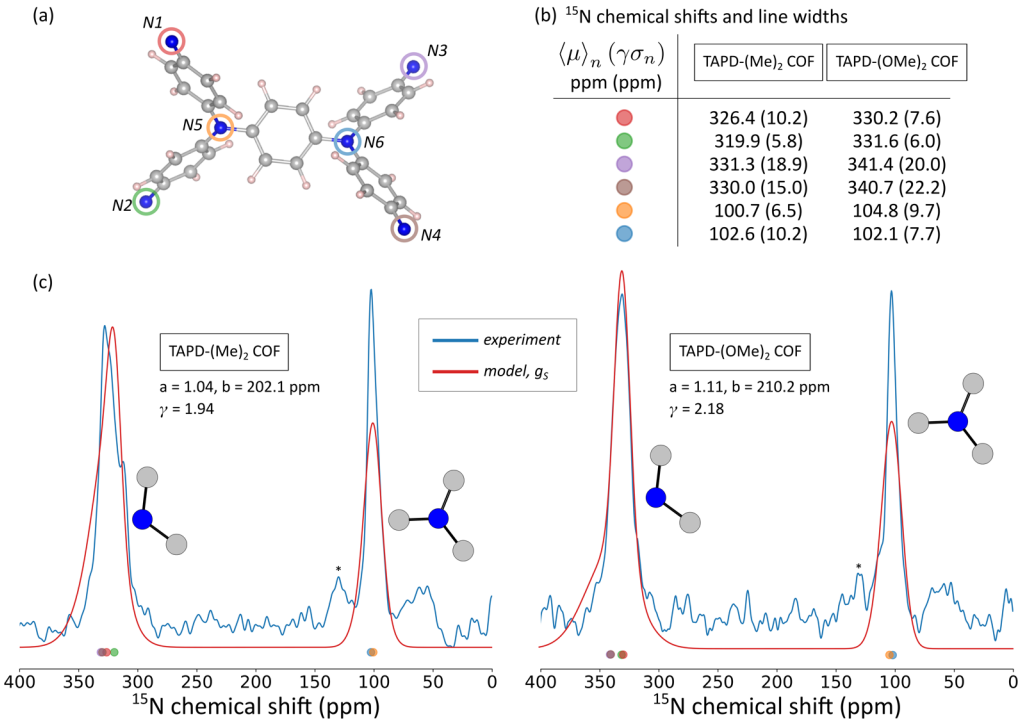


Figure 11. Application of the g_S protocol for the high signal-to-noise ratio ^{15}N NMR spectra of the TAPD-$(Me)_2$ and TAPD-$(OMe)_2$ COFs. (a) Definition of nitrogen classes based on the TAPD node. (b) Nitrogen classes chemical shifts and corresponding line widths in the g_S spectra. (c) ^{15}N experimental and model spectra (g_S). Spinning side bands are indicated with an asterisk (*).

label 2 and 3 contributions are misplaced; however, they do not produce distinct resonances. Similar to TAPD-$(Me)_2$ COF, TAPD-$(OMe)_2$ COF exhibits a lot of geometrical diversity around those sites, leading to large distributions in chemical shifts, even for the average values used in the g_S model. The carbon sites at label 7 are correctly assigned to the peak at 104 ppm, which is nicely separated from the rest of the spectrum. Lastly, the methoxy peak is located near 54 ppm. Clearly, the observations from the TAPD-$(OMe)_2$ and TAPD-$(Me)_2$ COF are general: experimental assignments obscure the underlying overlap of NMR contributions, following an underestimation of the diversity in chemical shifts that originates in geometric diversity. The misassignments that follow from this are exacerbated by potentially omitted (but relevant) carbon classes. The observations made here underline the main message that a unique and correct interpretation of NMR spectra of complex materials cannot be done based solely on experimental data. In such systems, complementing experimental data with a theoretical analysis that correctly accounts for structural diversity and possible dynamical effects, allows for a unique assignment and can also assess the broadness of some peaks.

3.4. ^{15}N NMR and Comment on Noisy Spectra. As the NMR calculations also provide the nitrogen shieldings, the same procedures can be repeated to produce a computational ^{15}N NMR spectrum for the TAPD-$(Me)_2$ and TAPD-$(OMe)_2$ COFs, for which new experimental NMR measurements were performed. The methodology for the collection of the ^{15}N NMR data is detailed in Section S4 of the Supporting Information. The experiments were repeated two times for each of the two COFs: once with 92,160 transients for TAPD-$(Me)_2$ COF and 82,944 transients for TAPD-$(OMe)_2$ COF, resulting in a lower signal-to-noise ratio and a second time with 261,120 and 467,520 transients for respectively TAPD-$(Me)_2$ and TAPD-$(OMe)_2$, resulting in a higher signal-to-noise ratio. After collection, a broadening filter of 100 Hz was applied to the spectra. All spectra at different measurement times and both with and without artificial broadening are shown in Section S4 of the Supporting Information. Here, the 100 Hz broadened spectra with the longest measurement times were used for both COFs.

The nitrogen classes are defined through the same procedure as the carbon classes, except that all nitrogen sites are included in the node, and no nitrogen sites are retained in the linker. This gives us six distinct nitrogen classes, as shown in Figure 11a. Note that the very low natural abundance of ^{15}N (0.368%) often results in the necessity for extremely long experiment durations. As such, it is often necessary to aim for reasonable, rather than optimal, signal-to-noise ratios. We do note that there is still a lot of noise present in the ^{15}N NMR experimental spectrum. The significant amount of noise leads to strong local variations in the amplitude in the spectra, which can be problematic for the parameter fitting procedure in our methodology. We have seen that in cases with a higher amount of noise, the fitting will often lead to broader line shapes (γ considerably larger than 1) in an attempt to replicate the noisy amplitude variations in between physical resonances which leads to a high MSE error. This error is minimized by the model through a broader line shape that on average has a lower MSE error than a strongly varying computational spectrum. This is definitely something to keep in mind in the case of noisy spectra, as it might get rid of spectral details that are otherwise relevant to the system. Additionally, if artificial broadening is introduced to reduce these strong noisy oscillations (as is the case here), the protocol will consequently significantly broaden the computational spectrum to match this.

https://doi.org/10.1021/acs.jctc.3c01414
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

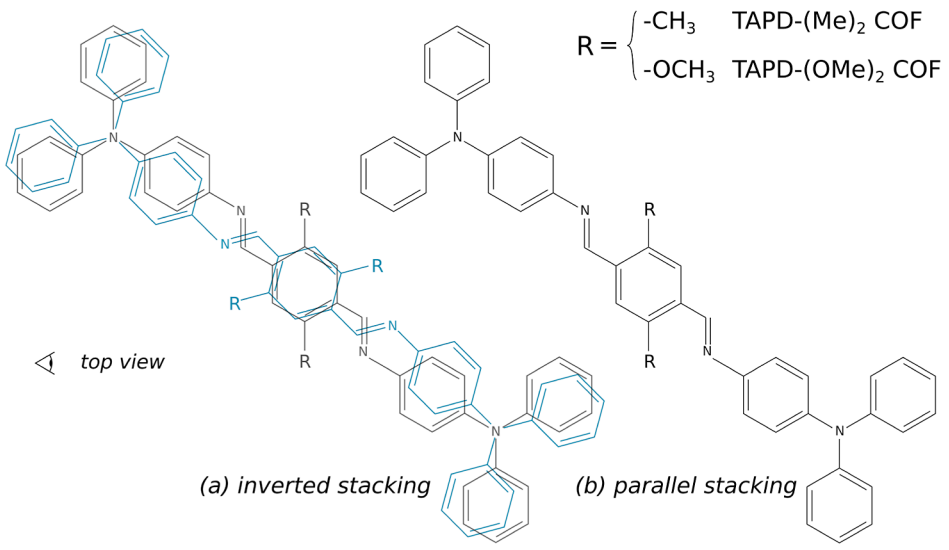


Figure 12. Definition of inverted and parallel stacking. These perfectly aligned structures are the ones that are used as starting structural models for the geometrical optimization.

The class chemical shifts $\langle\mu\rangle_n$ and line widths $\gamma\sigma_n$ for the nitrogen line shapes are shown in Figure 11b and the resulting spectra are shown in Figure 11c. The line widths of conjugated nitrogen atoms can also be high, up to 18.9 ppm in the TAPD-$(Me)_2$ case and 22.2 ppm in the TAPD-$(OMe)_2$ case. This is consistent with the large spread and line widths that are found for the carbon label 3 class. The geometric diversity affects not only the carbon but also the nitrogen chemical shifts. Overall, the experimental ^{15}N NMR spectrum is reproduced nicely with the g_S method for both the TAPD-$(Me)_2$ as well as the TAPD-$(OMe)_2$ COF. By our g_S model, it is predicted that some variation in the experimental line shape between 300 and 350 ppm (two-coordinate nitrogens) is not due to noise, rather there are physical differences in the nitrogen sites contributing to that line shape. This is the case because the average chemical shifts (plotted as round colored dots at the bottom of the spectra) are not all in the exact same location. Additionally, the width of those line shapes can be significantly larger than for the nitrogens bonded to three carbon sites. Not all variations in the experimental spectrum necessarily originated in noise. As is the case for the ^{13}C NMR spectrum, atoms that we may expect to be chemically equivalent may reside in geometrically distinct sites due to the mobility and geometric diversity in the system. A line shape in the NMR spectrum may consist of multiple line shapes at slightly different chemical shifts. As such, a feature in a line shape that may look like noise (especially in noisy spectra) may be of physical origin. This gives us information about the local structure around those nitrogen sites bonded to two carbon sites, which might be more structurally diverse than that seen in the experimental spectrum alone. This is additional information that the experimental nitrogen spectrum alone does not provide.

Moreover, the γ parameter is much higher than that in the models for the carbon spectra. This is normal; artificial broadening was added to the ^{15}N NMR spectra. While the noise is mitigated in the broadened spectrum, the actual peaks in the signal are broadened as well. This leads to a larger discrepancy between the width of the line shape caused by chemical shift broadening as well as artificial broadening, compared to the model values for the widths, which are purely due to chemical shift broadening.

3.5. Comparing Stacking Schemes and Mobility. As weak noncovalent interactions govern the stacking configurations, the relative layer position is inherently prone to disorder.[10,11,24,50−52] As such, COF layers can be stacked in many different ways that can vary in interlayer distance, layer offset, functional group location, functional group orientation (in-plane or out-of-plane), and so on. Previously, structural models of distinct stacking schemes were proposed and ordered as a function of their likelihood to reproduce the experimental PXRD pattern.[24] By comparing the experimental spectra to the computational XRD counterpart, it was revealed that the TAPD-$(Me)_2$ and TAPD-$(OMe)_2$ COF layers are primarily stacked in the inverted configuration as shown in Figure 12. In this inverted stacking scheme, the functional groups (denoted **R** in the figure) are placed on alternating carbon atoms of the central benzene ring of the linkers when going from one layer to the next. In a parallel stacking scheme, these functional groups are placed on the same carbon atom.

However, the XRD method that was previously used to identify the stacking scheme[35] requires medium- to long-range order to obtain clear signals. As the method is not sensitive on the atomic length scale, there is little local information that may be gathered through this technique.[50] Moreover, COF samples typically lack long-range order, making structure determination through XRD more difficult and ambiguous.[53,54] Additionally, if different regions of the sample have different stacking schemes, such a structural determination from XRD becomes virtually impossible. Theoretically, in such a case, it would still be possible to use NMR to attempt to assign the stacking schemes that are present. While solid-state NMR spectra do reflect long-range order, they do not intrinsically rely on the presence of this order to yield well-resolved spectra. Thus, in those (more complicated) cases where multiple stacking schemes are present in the sample, the NMR spectrum would be the sum of the spectra of different stacking schemes and one would be able to sum theoretical spectra and fit the results to the experimental spectrum. In this way, it might be possible to still resolve the stacking scheme in a regime in which the XRD method fails. In this section, we show that it could also be possible to use NMR to identify the

https://doi.org/10.1021/acs.jctc.3c01414
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

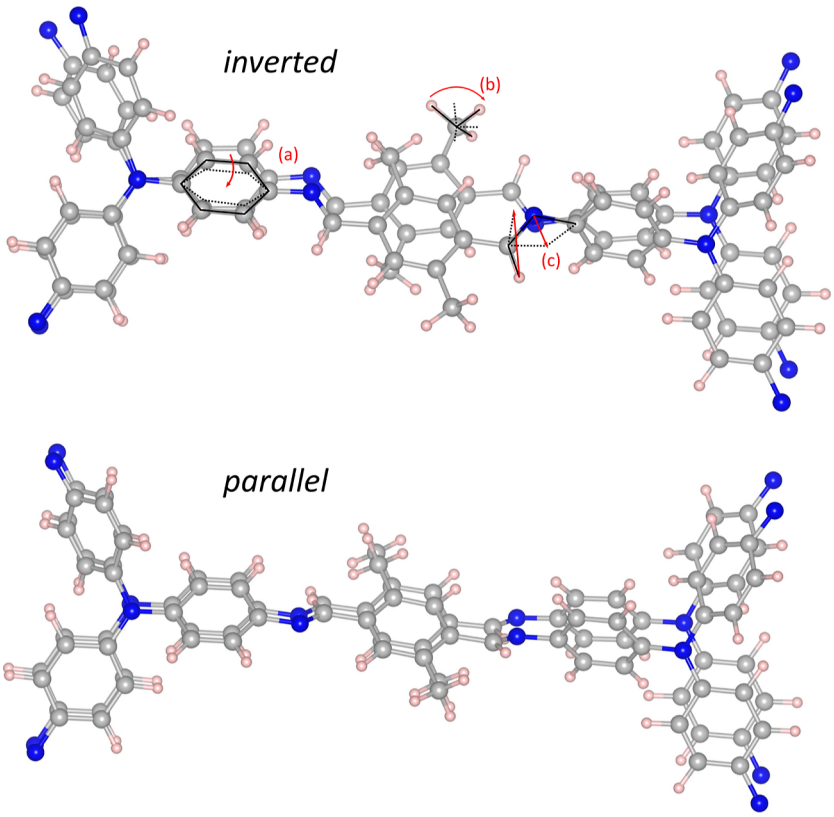


Figure 13. Cutout of the optimized inverted (top) and parallel (bottom) TAPD-$(Me)_2$ models. Letters (a−c) indicate displacements/rotations that may occur in these structures.

stacking scheme, at least in the simple case, where a single-stacking scheme dominates.

Considering again the TAPD-$(Me)_2$ system, Figure 13 illustrates the atomic positions in the statically optimized parallel structure, which are largely similar to those of the perfectly stacked starting structure. The 0 K energetic minimum for the parallel stacking scheme is found close to a situation with perfectly aligned layers, as schematically shown in Figure 12. In contrast, the atoms are displaced far more during static optimization of the inverted structure and the resultant structure lies further from being perfectly aligned. This is related to mobility of the structure. The more motional freedom the linkers and parts of the linkers have, the more diverse the local geometry around carbon sites is. Figure 13 shows some types of mobility that can be present that have a large effect on the local structure other than displacement of one entire COF layer with respect to the other. The fact that in the inverted stacking scheme there is more mobility could be explained by the larger interlayer distance. The distance between the linkers in adjacent layers of the COF is 3.42 Å in the parallel case where it is 3.87 Å in the inverted case in the optimized structures.

With this larger mobility in the inverted case in mind, consider the data in Figure 14, which shows the g_S line widths of the carbon line shapes for the parallel stacking model plotted against the line widths in the inverted stacking model for the TAPD-$(Me)_2$ COF. In this case, we have fixed the γ parameter to be exactly 1.0, so that the widths of the line shapes exactly reflect the structural diversity without additionally fitting these to the experimental spectrum, which could lead to unwanted and unphysical conclusions. The line widths of the carbon class line shapes are consistently higher in the

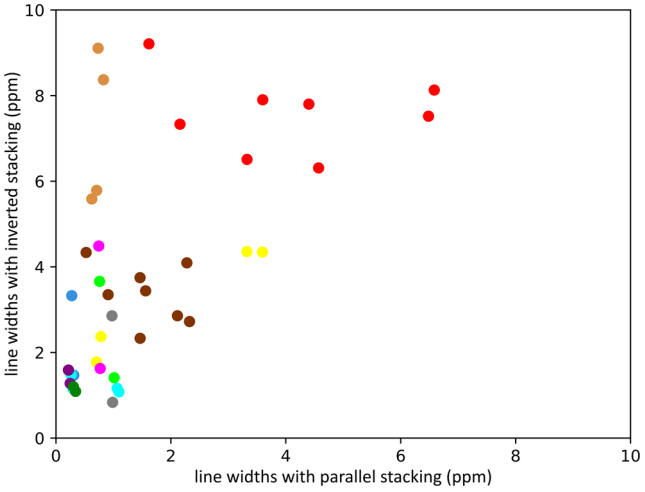


Figure 14. Comparison of the line widths of the carbon class line shapes between the inverted and parallel stacking schemes for the TAPD-$(Me)_2$ COF. The line widths are not scaled (γ = 1).

inverted stacking scheme compared to the parallel stacking scheme. Within the same carbon class, the carbon sites show more similar chemical shifts if the COF is stacked parallelly. The local geometry around the carbon sites in the parallel model is much more homogeneous than for the inverted stacking model, as can be seen in Figure 13 where part of the structurally optimized inverted and parallel structures are shown. From the line widths it is already clear that the parallel model will probably not be able to explain the experimental NMR spectrum. The homogeneity of the structure leads to spectral line widths that are implausibly low when they are

https://doi.org/10.1021/acs.jctc.3c01414
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

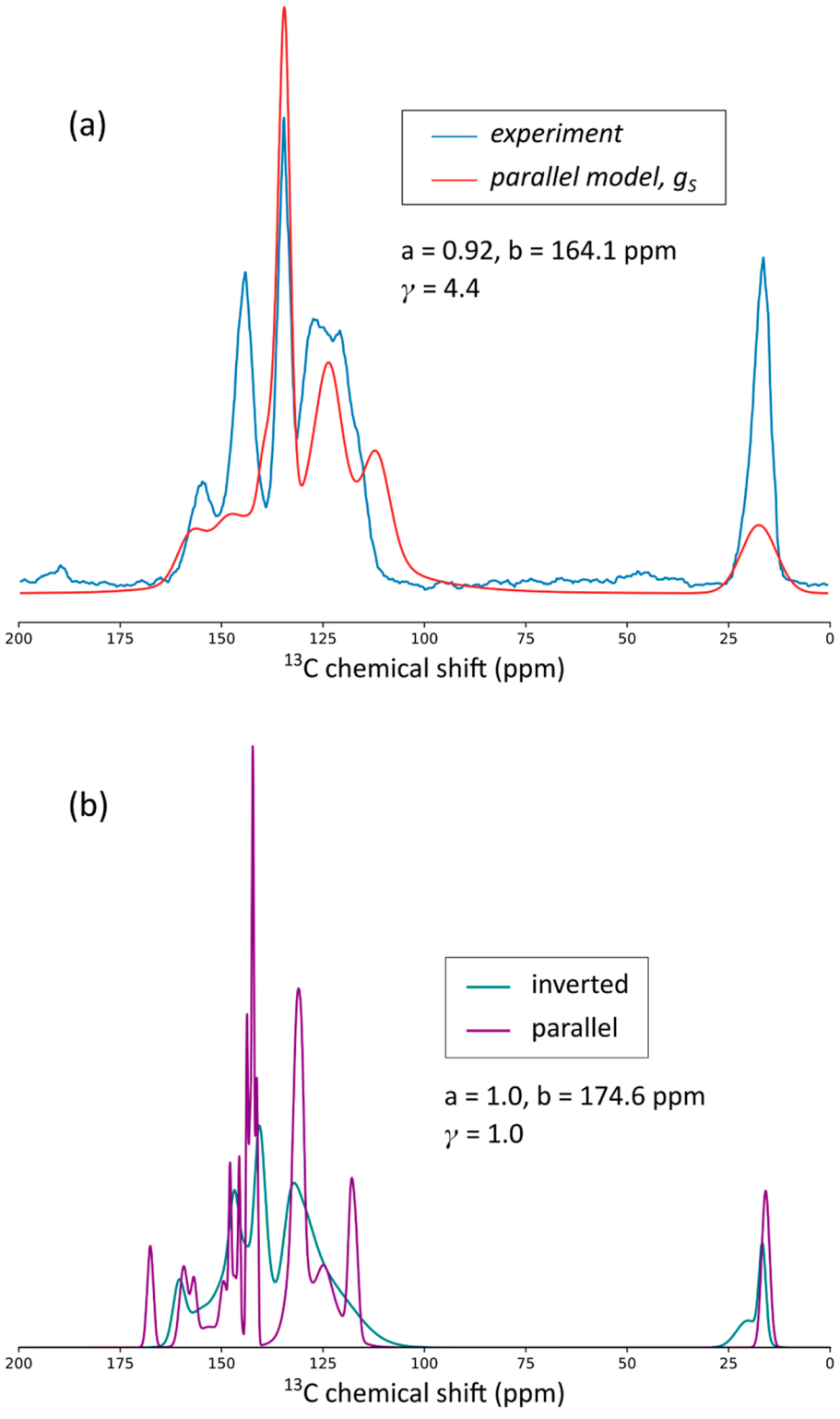


Figure 15. Fitting of the computational parallel stacking NMR data to the ^{13}C CPMAS NMR spectrum of the parallel model shows that the parallel model is unlikely to be present experimentally. (a) Attempting to fit the parallel g_S TAPD-$(Me)_2$ model spectrum to the experimental ^{13}C CPMAS NMR spectrum shows that a parallelly stacked model cannot explain the data. (b) A comparison between the inverted and parallel stacking schemes of the TAPD-$(Me)_2$ COF for the ^{13}C CPMAS NMR spectrum using the g_S approach with generic parameters indicated on the plot. The line widths as they emerge from the parallel model are far lower than for the inverted stacking model, reflecting the symmetry that is present in the parallel model.

compared to the experimental situation. Nevertheless, we tried to fit the ^{13}C NMR spectrum of the parallel model to the experimental TAPD-$(Me)_2$ NMR spectrum. The results are shown in Figure 15a. Other than the missing peaks, it is clear that the parameter that controls the overall width of the spectrum γ = 4.4 is unnaturally high, given the fact that in the inverted model γ = 0.97 remained close to 1 (see the table in

https://doi.org/10.1021/acs.jctc.3c01414
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

Figure 6). The model with the parallel stacking scheme, therefore, cannot explain the ^{13}C NMR data.

To get an impression of how different the inverted and parallel spectra look, we show both the parallel and inverted g_S spectra in Figure 15b. The line shape of each carbon or nitrogen class has its "natural" line width (the standard deviation of the chemical shifts belonging to the respective class with respect to their average, so $\gamma = 1$). We choose a generic shift-shielding relation with $a = 1$, and $b = 174.6$ ppm together with a generic amplitude of $A = 1$ to plot the spectra. We also want to draw attention to the asymmetry of the isolated methyl resonance: while this resonance manifests as asymmetrical in the experimental spectrum, the parallel model produces a symmetric resonance. From all these considerations, and the fact that the inverted spectrum is in near-perfect agreement with experiment, it is clear that the parallel stacking configuration, if present, is negligible with respect to the inverted stacking configuration. This case shows that NMR could also be used to determine stacking configurations in COFs.

4. CONCLUSIONS

In this paper, we propose three approaches to construct computational NMR spectra for COFs through the usage of quantum mechanical static structural models. All models start from the same structural model and the chemical shielding data but approach it in a different way, ranging from no averaging over shieldings to maximal averaging. It is clear that all three approaches could be deemed viable alternatives for the production of a model carbon spectrum, as each approach produces a spectrum that is in decent agreement with the experimental spectra. In reality, there is mobility of the framework, and each carbon site is dynamic. What we should aim to mimic is an ensemble average over configurations that is representative of the dynamic COF structure. During the collection of the NMR signal, carbon sites propagate through multiple possible geometries, and therefore, the actual NMR resonance that a single carbon site produces is an average resonance. One could resort to MD for the propagation of the system either at the DFT level of theory or using a force field. The first, however, is expensive for systems of this size. The computational cost could still be a problem in the dynamic approach with force fields because the force field still has to be derived from more expensive calculations, which is in itself not always easy in complex materials. In any case, given the fact that COFs can be placed in different stacking configurations and with different orientations for the linkers, it can become impractical to perform MD runs for all possible configurations, whereas it might still be feasible to use static calculations for all candidate structures. Within this work, we have shown that a static approach can yield a great deal of information without the need to perform explicit MD simulations. Instead, the inherent spatiotemporal averaging of an experimental measurement can be mimicked by an appropriate configurational averaging technique for the chemical shifts, grounded in a particular grouping strategy of atomic sites. The three approaches g_I, g_S, g_R each have their own specific merit and downsides:

- The uniform model (g_I) uses a single line shape per atomic site. Since we are using a static model, it is therefore not very physical that each carbon site produces a single line shape with one specific chemical shift. In this sense, the model is not a good approximation for the ensemble average that we aim to represent. However, the spectrum that it produces can visually still be similar to the experimental spectrum, therefore it could serve as a quick first check that requires little postprocessing.
- The symmetrically inequivalent average model (g_R) is, in a sense, the opposite of the g_I model, because it averages out a lot of information where the g_I model does not average anything. The atomic sites are grouped based on the symmetry of the nodes and linkers. However, the symmetry of the building block inside the COF material can be lower than the isolated molecule. When the chemical shift is overly averaged, information can be lost, as carbon sites that are not equivalent inside the framework are treated as being equivalent.
- The SBU average model (g_S) is an intermediate between the g_I and the g_R model, as it still uses symmetry to group the carbon sites into inequivalent classes. However, in this model there is more freedom, as the classes of the g_R model are partitioned. Each atom in the node or linker represents an inequivalent class, and the average is taken over the nodes/linkers present in the system. With this model, we approximate the ensemble average the closest, without the need to perform expensive MD simulations at the DFT level of theory.

Furthermore, we hypothesize that, for some systems, there exist better choices for the g_S model classes than just basing the classes on the symmetry of the underlying building blocks. For example in the case of COFs, if there is reason to believe that carbon atoms across different layers will produce a separate resonance in the NMR spectrum, it will be better not to average the chemical shifts over these different building blocks across layers. Rather, in that case, it would be best to simply increase the number of classes (and, if necessary to obtain good statistics, increase the unit cell size) and only average the chemical shifts over classes within the same layer of COF material. However, increasing the number of classes further in this system is not beneficial in this case, as the number of carbon atoms per class in such a modified g_S model becomes too small to retain decent statistics in this case. The principle of increasing the number of carbon classes to limit the averaging of chemical shifts to carbon atoms residing within the same layer can also be applied to the g_R model, the resulting spectrum of which may be found in Section S5. Looking beyond the materials discussed in this work, an optimal choice for the different classes in the g_S model hinges on the (in)equivalence of the atomic sites within the material, originating from the invariant structural features of the material.

Similar to systems where significant geometrical differences are expected to be present across layers, in cases in which some form of turbostratic disorder is present, multiple regions of the COF can have different space groups. In that case, some carbon classes that by default in the g_S model are treated to be equivalent are better kept as separate classes. In a case such as COF-5,[55] it is possible that considering the different existing symmetries of the (different regions of the) framework is preferable over defining the carbon classes based on the symmetry of the building block.

We also showed that the assignments of carbon spectra that are routinely performed by experimentalists are not always fully

https://doi.org/10.1021/acs.jctc.3c01414
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

correct. Whether this results from missing carbon classes stemming from an oversimplification of the structural models involved, or assumptions about the spectra concerning dynamics of the species studied, it is clear that performing a DFT calculation provides relatively cheap and necessary complementary information for unique peak assignments. These methods can be readily applied to check existing assignments in the literature as well as for future studies, where uncertainty in the spectral interpretation may arise due to missing information, misinterpretation of mobility, or noisy spectra.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jctc.3c01414.

Comment on quantitative vs. nonquantitative experimental spectra, internal validation: dynamic vs. static model within g_S approach, g_I and g_R models for the TAPD-$(OMe)_2$ COF, experimental NMR details and comparison between low/high signal-to-noise ratio of ^{15}N NMR spectra, g_R model with no averaging across layers, structural models, input files and Python scripts are provided through the following Github repository: https://github.com/molmod/copacof (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Veronique Van Speybroeck − *Center for Molecular Modeling, Ghent University, 9052 Ghent, Belgium;* orcid.org/0000-0003-2206-178X; Email: Veronique.VanSpeybroeck@UGent.be

Authors

Siebe Vanlommel − *Center for Molecular Modeling, Ghent University, 9052 Ghent, Belgium*

Sander Borgmans − *Center for Molecular Modeling, Ghent University, 9052 Ghent, Belgium;* orcid.org/0000-0002-9187-4293

C. Vinod Chandran − *NMRCoRe, NMR/X-Ray Platform for Convergence Research, B-3001 Leuven, Belgium;* orcid.org/0000-0003-4074-8833

Sambhu Radhakrishnan − *NMRCoRe, NMR/X-Ray Platform for Convergence Research, B-3001 Leuven, Belgium;* orcid.org/0000-0002-0274-2759

Pascal Van Der Voort − *Department of Chemistry, Ghent University, 9000 Ghent, Belgium;* orcid.org/0000-0002-1248-479X

Eric Breynaert − *NMRCoRe, NMR/X-Ray Platform for Convergence Research, B-3001 Leuven, Belgium;* orcid.org/0000-0003-3499-0455

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jctc.3c01414

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The computational resources and services used in this work were provided by the Flemish Supercomputer Center, funded by Ghent University, Fund for Scientific Research-Flanders (FWO), and the Flemish Government department EWI. NMRCoRe acknowledges the Flemish government, department EWI for infrastructure investment via the Hermes Fund (AH.2016.134) and for financial support as International Research Infrastructure (I001321N: Nuclear Magnetic Resonance Spectroscopy Platform for Molecular Water Research) via FWO Vlaanderen.

■ REFERENCES

(1) Côté, A. P.; Benin, A. I.; Ockwig, N. W.; O'Keeffe, M.; Matzger, A. J.; Yaghi, O. M. Porous, Crystalline, Covalent Organic Frameworks. *Science* **2005**, *310*, 1166−1170.

(2) Geng, K.; He, T.; Liu, R.; Dalapati, S.; Tan, K. T.; Li, Z.; Tao, S.; Gong, Y.; Jiang, Q.; Jiang, D. Covalent Organic Frameworks: Design, Synthesis, and Functions. *Chem. Rev.* **2020**, *120*, 8814−8933.

(3) Esrafili, A.; Wagner, A.; Inamdar, S.; Acharya, A. P. Covalent Organic Frameworks for Biomedical Applications. *Adv. Healthcare Mater.* **2021**, *10*, 2002090.

(4) Sharma, R. K.; Yadav, P.; Yadav, M.; Gupta, R.; Rana, P.; Srivastava, A.; Zbořil, R.; Varma, R. S.; Antonietti, M.; Gawande, M. B. Recent Development of Covalent Organic Frameworks (COFs): Synthesis and Catalytic (Organic-Electro-Photo) Applications. *Mater. Horiz.* **2020**, *7*, 411−454.

(5) Zhao, X.; Pachfule, P.; Thomas, A. Covalent Organic Frameworks (COFs) for Electrochemical Applications. *Chem. Soc. Rev.* **2021**, *50*, 6871−6913.

(6) Wu, M.-X.; Yang, Y.-W. Applications of Covalent Organic Frameworks (COFs): From Gas Storage and Separation to Drug Delivery. *Chin. Chem. Lett.* **2017**, *28*, 1135−1143.

(7) Fan, H.; Mundstock, A.; Feldhoff, A.; Knebel, A.; Gu, J.; Meng, H.; Caro, J. Covalent Organic Framework−Covalent Organic Framework Bilayer Membranes for Highly Selective Gas Separation. *J. Am. Chem. Soc.* **2018**, *140*, 10094−10098.

(8) Tran, Q. N.; Lee, H. J.; Tran, N. Covalent Organic Frameworks: From Structures to Applications. *Polymers* **2023**, *15*, 1279.

(9) Wang, H.; Wang, H.; Wang, Z.; Tang, L.; Zeng, G.; Xu, P.; Chen, M.; Xiong, T.; Zhou, C.; Li, X.; Huang, D.; Zhu, Y.; Wang, Z.; Tang, J. Covalent Organic Framework Photocatalysts: Structures and Applications. *Chem. Soc. Rev.* **2020**, *49*, 4135−4165.

(10) Rawat, K. S.; Borgmans, S.; Braeckevelt, T.; Stevens, C. V.; Van Der Voort, P.; Van Speybroeck, V. How the Layer Alignment in Two-Dimensional Nanoporous Covalent Organic Frameworks Impacts Its Electronic Properties. *ACS Appl. Nano Mater.* **2022**, *5*, 14377−14387.

(11) Mähringer, A.; Medina, D. D. Taking Stock of Stacking. *Nat. Chem.* **2020**, *12*, 985−987.

(12) Guo, H.; Zhang, L.; Xue, R.; Ma, B.; Yang, W. Eyes of Covalent Organic Frameworks: Cooperation between Analytical Chemistry and COFs. *Rev. Anal. Chem.* **2019**, *38*, 20170023.

(13) Nguyen, H. L. Reticular Design and Crystal Structure Determination of Covalent Organic Frameworks. *Chem. Sci.* **2021**, *12*, 8632−8647.

(14) Ma, T.; Kapustin, E. A.; Yin, S. X.; Liang, L.; Zhou, Z.; Niu, J.; Li, L.-H.; Wang, Y.; Su, J.; Li, J.; Wang, X.; Wang, W. D.; Wang, W.; Sun, J.; Yaghi, O. M. Single-Crystal X-ray Diffraction Structures of Covalent Organic Frameworks. *Science* **2018**, *361*, 48−52.

(15) Giovine, R.; Volkringer, C.; Trébosc, J.; Amoureux, J.-P.; Loiseau, T.; Lafon, O.; Pourpoint, F. NMR Crystallography to Probe the Breathing Effect of the MIL-53(Al) Metal−Organic Framework using Solid-State NMR Measurements of ^{13}C−^{27}Al Distances. *Acta Crystallogr., Sect. C: Struct. Chem.* **2017**, *73*, 176−183.

(16) Reif, B.; Ashbrook, S. E.; Emsley, L.; Hong, M. Solid-State NMR Spectroscopy. *Nat. Rev. Methods Primers* **2021**, *1*, 2.

(17) Bryce, D. L. NMR Crystallography: Structure and Properties of Materials from Solid-State Nuclear Magnetic Resonance Observables. *IUCrJ* **2017**, *4*, 350−359.

(18) Rehman, S. U.; Xu, S.; Xu, H.; Tao, T.; Li, Y.; Yu, Z.; Ma, K.; Xu, W.; Wang, J. The Role of NMR in Metal Organic Frameworks: Deep Insights into Dynamics, Structure and Mapping of Functional Groups. *Mater. Today Adv.* **2022**, *16*, 100287.

(19) Southern, S. A.; Bryce, D. L. Recent Advances in NMR Crystallography and Polymorphism. *Annu. Rep. NMR Spectrosc.* **2021**, *102*, 1−80.
(20) Xu, Y.; Southern, S. A.; Szell, P. M. J.; Bryce, D. L. The Role of Solid-State Nuclear Magnetic Resonance in Crystal Engineering. *CrystEngComm* **2016**, *18*, 5236−5252.
(21) Klein, P.; Pashkova, V.; Thomas, H. M.; Whittleton, S. R.; Brus, J.; Kobera, L.; Dedecek, J.; Sklenak, S. Local Structure of Cationic Sites in Dehydrated Zeolites Inferred from ^{27}Al Magic-Angle Spinning NMR and Density Functional Theory Calculations. A Study on Li-Na-and K-Chabazite. *J. Phys. Chem. C* **2016**, *120*, 14216−14225.
(22) Mlekodaj, K.; Dedecek, J.; Pashkova, V.; Tabor, E.; Klein, P.; Urbanova, M.; Karcz, R.; Sazama, P.; Whittleton, S. R.; Thomas, H. M.; Fishchuk, A. V.; Sklenak, S. Al Organization in the SSZ-13 Zeolite. Al Distribution and Extraframework Sites of Divalent Cations. *J. Phys. Chem. C* **2019**, *123*, 7968−7987.
(23) Sardo, M.; Santos, S. M.; Babaryk, A. A.; López, C.; Alkorta, I.; Elguero, J.; Claramunt, R. M.; Mafra, L. Diazole-Based Powdered Cocrystal Featuring a Helical Hydrogen-Bonded Network: Structure Determination from PXRD, Solid-State NMR and Computer Modeling. *Solid State Nucl. Magn. Reson.* **2015**, *65*, 49−63.
(24) Borgmans, S.; Rogge, S. M. J.; De Vos, J. S.; Stevens, C. V.; Van Der Voort, P.; Van Speybroeck, V. Quantifying the Likelihood of Structural Models through a Dynamically Enhanced Powder X-Ray Diffraction Protocol. *Angew. Chem., Int. Ed.* **2021**, *60*, 8913−8922.
(25) Wang, Y.; Liu, H.; Pan, Q.; Wu, C.; Hao, W.; Xu, J.; Chen, R.; Liu, J.; Li, Z.; Zhao, Y. Construction of Fully Conjugated Covalent Organic Frameworks via Facile Linkage Conversion for Efficient Photoenzymatic Catalysis. *J. Am. Chem. Soc.* **2020**, *142*, 5958−5963.
(26) Zhu, Y.; Wan, S.; Jin, Y.; Zhang, W. Desymmetrized Vertex Design for the Synthesis of Covalent Organic Frameworks with Periodically Heterogeneous Pore Structures. *J. Am. Chem. Soc.* **2015**, *137*, 13772−13775.
(27) Sun, B.; Wang, D.; Wan, L. Template Synthesis of Imine-Based Covalent Organic Framework Core-Shell Structure and Hollow Sphere: A Case of COFTTA-DHTA. *Sci. China Chem.* **2017**, *60*, 1098−1102.
(28) Lyu, H.; Diercks, C. S.; Zhu, C.; Yaghi, O. M. Porous Crystalline Olefin-Linked Covalent Organic Frameworks. *J. Am. Chem. Soc.* **2019**, *141*, 6848−6852.
(29) Zhang, G.; Tsujimoto, M.; Packwood, D.; Duong, N. T.; Nishiyama, Y.; Kadota, K.; Kitagawa, S.; Horike, S. Construction of a Hierarchical Architecture of Covalent Organic Frameworks via a Postsynthetic Approach. *J. Am. Chem. Soc.* **2018**, *140*, 2602−2609.
(30) Kumar, S.; Abdulhamid, M. A.; Dinga Wonanke, A. D.; Addicoat, M. A.; Szekely, G. Norbornane-Based Covalent Organic Frameworks for Gas Separation. *Nanoscale* **2022**, *14*, 2475−2481.
(31) Li, S.; Lafon, O.; Wang, W.; Wang, Q.; Wang, X.; Li, Y.; Xu, J.; Deng, F. Recent Advances of Solid-State NMR Spectroscopy for Microporous Materials. *Adv. Mater.* **2020**, *32*, 2002879.
(32) Hoffmann, H.; Debowski, M.; Müller, P.; Paasch, S.; Senkovska, I.; Kaskel, S.; Brunner, E. Solid-State NMR Spectroscopy of Metal−Organic Framework Compounds (MOFs). *Materials* **2012**, *5*, 2537−2572.
(33) Wang, W.; Xu, J.; Deng, F. Recent Advances in Solid-State NMR of Zeolite Catalysts. *Natl. Sci. Rev.* **2022**, *9*, nwac155.
(34) Debruyne, M.; Borgmans, S.; Radhakrishnan, S.; Breynaert, E.; Vrielinck, H.; Leus, K.; Laemont, A.; De Vos, J.; Rawat, K. S.; Vanlommel, S.; Rijckaert, H.; Salemi, H.; Everaert, J.; Vanden Bussche, F.; Poelman, D.; Morent, R.; De Geyter, N.; Van Der Voort, P.; Van Speybroeck, V.; Stevens, C. V. Engineering of Phenylpyridine- and Bipyridine-Based Covalent Organic Frameworks for Photocatalytic Tandem Aerobic Oxidation/Povarov Cyclization. *ACS Appl. Mater. Interfaces* **2023**, *15*, 35092−35106.
(35) Krishnaraj, C.; Sekhar Jena, H.; Bourda, L.; Laemont, A.; Pachfule, P.; Roeser, J.; Chandran, C. V.; Borgmans, S.; Rogge, S. M. J.; Leus, K.; Stevens, C. V.; Martens, J. A.; Van Speybroeck, V.; Breynaert, E.; Thomas, A.; Van Der Voort, P. Strongly Reducing (Diarylamino)benzene-Based Covalent Organic Framework for Metal-Free Visible Light Photocatalytic H_2O_2 Generation. *J. Am. Chem. Soc.* **2020**, *142*, 20107−20116.
(36) Grajciar, L.; Heard, C. J.; Bondarenko, A. A.; Polynski, M. V.; Meeprasert, J.; Pidko, E. A.; Nachtigall, P. Towards *Operando* Computational Modeling in Heterogeneous Catalysis. *Chem. Soc. Rev.* **2018**, *47*, 8307−8348.
(37) Vanlommel, S.; Hoffman, A. E. J.; Smet, S.; Radhakrishnan, S.; Asselman, K.; Chandran, C. V.; Breynaert, E.; Kirschhock, C. E. A.; Martens, J. A.; Van Speybroeck, V. How Water and Ion Mobility Affect the NMR Fingerprints of the Hydrated JBW Zeolite: A Combined Computational-Experimental Investigation. *Chem.—Eur. J.* **2022**, *28*, No. e202202621.
(38) Lei, C.; Erlebach, A.; Brivio, F.; Grajciar, L.; Tošner, Z.; Heard, C.; Nachtigall, P. The need for operando modelling of ^{27}Al NMR in zeolites. 2023, *chemrxiv.org ePrint archive*, https://chemrxiv.org/engage/chemrxiv/article-details/6407582d6642bf8c8f1fb191 (accessed June 6, 2023).
(39) Blöchl, P. E. Projector Augmented-Wave Method. *Phys. Rev. B* **1994**, *50*, 17953−17979.
(40) Kresse, G.; Joubert, D. From Ultrasoft Pseudopotentials to the Projector Augmented-Wave Method. *Phys. Rev. B* **1999**, *59*, 1758−1775.
(41) Pickard, C. J.; Mauri, F. All-Electron Magnetic Response with Pseudopotentials: NMR Chemical Shifts. *Phys. Rev. B* **2001**, *63*, 245101.
(42) Kresse, G.; Furthmüller, J. Efficient Iterative Schemes for Ab Initio Total-Energy Calculations Using a Plane-Wave Basis Set. *Phys. Rev. B* **1996**, *54*, 11169−11186.
(43) Kresse, G.; Furthmüller, J. Efficiency of Ab-Initio Total Energy Calculations for Metals and Semiconductors Using a Plane-Wave Basis Set. *Comput. Mater. Sci.* **1996**, *6*, 15−50.
(44) Kresse, G.; Hafner, J. Ab Initio Molecular-Dynamics Simulation of the Liquid-Metal-Amorphous-Semiconductor Transition in Germanium. *Phys. Rev. B* **1994**, *49*, 14251−14269.
(45) Kresse, G.; Hafner, J. Ab Initio Molecular Dynamics for Liquid Metals. *Phys. Rev. B* **1993**, *47*, 558−561.
(46) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865−3868.
(47) Zhang, Y.; Yang, W. Comment on "Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1998**, *80*, 890.
(48) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the Damping Function in Dispersion Corrected Density Functional Theory. *J. Comput. Chem.* **2011**, *32*, 1456−1465.
(49) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate Ab Initio Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104.
(50) Pütz, A. M.; Terban, M. W.; Bette, S.; Haase, F.; Dinnebier, R. E.; Lotsch, B. V. Total Scattering Reveals the Hidden Stacking Disorder in a 2D Covalent Organic Framework. *Chem. Sci.* **2020**, *11*, 12647−12654.
(51) Zhang, Y.; Položij, M.; Heine, T. Statistical Representation of Stacking Disorder in Layered Covalent Organic Frameworks. *Chem. Mater.* **2022**, *34*, 2376−2381.
(52) Huang, J.; Shin, S.-J.; Tolborg, K.; Ganose, A. M.; Krenzer, G.; Walsh, A. Room-Temperature Stacking Disorder in Layered Covalent-Organic Frameworks from Machine-Learning Force Fields. *Mater. Horiz.* **2023**, *10*, 2883−2891.
(53) Waller, P. J.; Gándara, F.; Yaghi, O. M. Chemistry of Covalent Organic Frameworks. *Acc. Chem. Res.* **2015**, *48*, 3053−3063.
(54) Zhou, G.; Yang, T.; Huang, Z. Structure Determination of a Low-Crystallinity Covalent Organic Framework by Three-Dimensional Electron Diffraction. *Commun. Chem.* **2023**, *6*, 116.
(55) Pope, G.; Vazquez, D.; Uribe-Romo, F.; Harper, J. K. Characterizing the Turbostratic Disorder in COF-5 with NMR Crystallography. 2020, *chemrxiv.org ePrint archive, chemrxiv.org/engage/chemrxiv/article-details/60c7508b842e655b46db3a3a* (accessed June 6, 2023).

https://doi.org/10.1021/acs.jctc.3c01414
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX