Ab-initio Quality NMR Parameters in Solid-State Materials using a High-Dimensional Neural-Network Representation

Jérôme Cuny,[*,†] Yu Xie,[*,‡] Chris J. Pickard,[¶] and Ali A. Hassanali[*,§]

Laboratoire de Chimie et Physique Quantiques (LCPQ), Université de Toulouse [UPS] and CNRS, 118 Route de Narbonne, F-31062 Toulouse, France, Toulouse, Center for Nanophase Materials Sciences, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, United States, Department of Materials Science & Metallurgy, University of Cambridge, 27 Charles Babbage Road, Cambridge, CB3 0FS, and The Abdus Salaam International Center for Theoretical Physics, Condensed Matter Physics Section, Strada Costiera 11, Trieste, Italy

E-mail: jerome.cuny@irsamc.ups-tlse.fr; yxe@ornl.gov; ahassana@ictp.it

Abstract

Nuclear magnetic resonance (NMR) spectroscopy is one of the most powerful experimental tools to probe the local atomic order of a wide range of solid-state compounds. However, due to the complexity of the related spectra, in particular for amorphous

[*]To whom correspondence should be addressed

[†]Laboratoire de Chimie et Physique Quantiques (LCPQ), Université de Toulouse [UPS] and CNRS, 118 Route de Narbonne, F-31062 Toulouse, France, Toulouse

[‡]Center for Nanophase Materials Sciences, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, United States

[¶]Department of Materials Science & Metallurgy, University of Cambridge, 27 Charles Babbage Road, Cambridge, CB3 0FS

[§]The Abdus Salaam International Center for Theoretical Physics, Condensed Matter Physics Section, Strada Costiera 11, Trieste, Italy

materials, their interpretation in terms of structural information is often challenging. These difficulties can be overcome by combining molecular dynamics simulations to generate realistic structural models with an *ab initio* evaluation of the corresponding nuclear shielding and quadrupolar coupling tensors. However, due to computational constraints, this approach is limited to relatively small system sizes which, for amorphous materials, prevents an adequate statistical sampling of the distribution of the local environments that is required to quantitatively describe the system. In this work, we present an approach to efficiently and accurately predict the NMR parameters of very large systems. This is achieved by using a high-dimensional neural-network representation of NMR parameters that are calculated using an *ab initio* formalism. To illustrate the potential of this approach, we applied this neural-network NMR (NN-NMR) method on the ^{17}O and ^{29}Si quadrupolar coupling and chemical shift parameters of various crystalline silica polymorphs and silica glasses. This approach is, in principal, general and has the potential to be applied to predict the NMR properties of various materials.

1 Introduction

Solid-state nuclear magnetic resonance (SSNMR) spectroscopy is one of the most powerful tools for the structural characterization of amorphous materials.[1–3] In contrast to X-ray or neutron diffraction analysis which require long range atomic order, SSNMR provides information at the local atomic level which allows the probing of the chemical and geometrical disorder that is characteristic of amorphous systems. However, structural disorder is associated with distributions of both the isotropic and anisotropic parts of the various possible NMR interactions that are characteristic of a particular nucleus in a given material. This generally makes it very challenging to unambiguously interpret the experimental data. Despite the development of high-resolution NMR techniques such as two-dimensional multi-quantum magic-angle spinning (MQMAS)[4–7] and dynamic-angle spinning (DAS),[8,9]

the development of very high magnetic field spectrometers and the possibility of isotopic enrichment, a large part of the structural information contained in the spectra can only be partially used. Therefore, in order to extract more information from experimental data so as to better describe the atomic structure of amorphous materials, it is desirable to develop additional tools to complement experimental measurements. This is an urgent need since amorphous materials are present in a number of important devices for industrial applications such as in optical fibres[10] or for the storage of nuclear waste.[11,12] Consequently, significant effort has been put into developing theoretical methods to evaluate NMR tensors from first-principles, in particular from density functional theory (DFT).[13–17] Among them, one can mentioned the widely used *Projector-Augmented Waves* (PAW)[18] and *Gauge-Including Projector-Augmented Waves* (GIPAW)[15] methods for the calculation of quadrupolar coupling and nuclear shielding tensors, respectively. As the minimal requirement for these approaches is a set of atomic coordinates, one can combine them with molecular dynamics (MD) simulations that can provide a large amount of structural data that are representative of a chemical system at a given temperature. By providing both an accurate description of the atomic structure and of the corresponding NMR parameters, this approach allows for a more accurate interpretation of the experimental spectra and a better understanding of the local structure of the probed materials.[19–21] This methodology has been successfully applied to several chemical systems, in particular silicate and aluminosilicate glasses,[22–31] phosphates glasses,[32,33] chalcogenide glasses[34–36] and proton-conducting polymers.[37,38]

The number of atoms that is required to quantitatively represent the statistical distribution of the various chemical environments of an amorphous material is large. For example, Huff *et al.* showed that to obtain a realistic description of the properties of bulk silica, a unit cell consisting of at least 3000 atoms is required.[39] Hence, for this and other chemical systems, MD simulations have generally been performed using several hundreds or thousands of atoms,[40–43] and it is desirable to evaluate the NMR parameters of similar system sizes. However, despite the continuous increase of computational power and the improvement

of software efficiency, DFT based approaches can only handle a limited number of atoms. Therefore, the application of these approaches to amorphous systems is limited to small system sizes of a few hundred atoms.[22,24–27,30,31,36] Although these studies allowed for a better understanding of experimental data, to the best of our knowledge, they were not able to achieve a proper statistical convergence of the studied NMR parameter distributions. These limitations were highlighted by Charpentier *et al.* who proposed an interpolation scheme, the kernel density estimate (KDE), based on an interpolation scheme of a small amount of input data that simulates smooth and converged theoretical spectra.[24] However, this method can have its limitations, in particular for the study of more complex chemical systems and so, it is desirable to develop an approach to handle larger system sizes while maintaining DFT-like accuracy.

For this purpose, we developed a methodology that uses a high-dimensional neural-network architecture to predict the quadrupolar coupling and the chemical shift parameters of large and complex solid-state materials with an *ab initio* quality. This is achieved by training our neural-network using input parameters exclusively calculated from first-principles, in contrast to previously proposed methods based on experimental data and specifically devoted to biological systems.[44,45] In the present case, the structural data describing the atomic environments are provided by MD simulations performed on a finite number of small ($\sim$100 atoms) to middle size ($\sim$300 atoms) unit cells. The corresponding NMR parameters are evaluated for a selected number of frames using an *ab initio* approach. To demonstrate the validity of this method and to test its performance, we have applied it to amorphous silica. Over the last two decades, silica has received an increasing interest from both theoretical and experimental studies,[46–49] which is motivated by both the heterogeneous chemical complexity of the SiO_2 system and by the number of fundamental applications of its derived devices.[46,47,50] For example, silica nano-particles were proved to be an efficient material for various biomedical applications such as carriers of therapeutic agents,[51] and biomedical imaging,[52] or as supporting material for various catalysts.[46,47] The characterization of amor-

phous silica by SSNMR thus provides valuable information about the atomic structure of these complex materials.

The paper is laid out as follows: in section 2, after a brief presentation of the challenges for developing a neural-network for NMR parameters, we describe the approaches we used to produce the input data and how they are further used to generate a NN-NMR. We finish this section by testing its accuracy and transferability on various silica phases. In section 3, the NN-NMR is applied to the characterisation of large amorphous silica slabs of various sizes. Finally, we end with some concluding remarks on our approach and its possible future applications.

2 Computational Methods

2.1 Neural-Network Representation of NMR Parameters

Artificial NN were first developed to investigate the neural signal processing in the brain,[53] and then extended to a wide range of applications such as the fitting of arbitrary functions,[54–59] weather forecasting,[60,61] and processing of chemical data.[62] In the framework of theoretical chemistry, it has been applied to calculate accurate densities of states from limited *ab-initio* data,[63] and more recently to the construction of accurate energy potential surfaces.[64] In our case, the goal of the NN-NMR is to construct a functional relationship between a given atomic configuration and the corresponding NMR parameters, both quadrupolar coupling and chemical shift parameters. A similar procedure has already been integrated in widely used software to accurately predict the ^{1}H, ^{13}C and ^{14}N isotropic chemical shifts properties of large biological systems.[44,45] In these cases, a large amount of experimental structural and NMR data was used to train the neural-network. For amorphous inorganic materials, this kind of approach is not feasible as only distributions of structural and NMR parameters can be obtained experimentally. Moreover, the structure of the macro-molecules were described by a specific set of order parameters such as the sequence of amino-acids

or a set of backbone angles that are not transferable to describe non-biological chemical compounds. Thus, the definition of a set of order parameters that allows for an accurate correlation between a local environment and the corresponding NMR parameters is necessary.

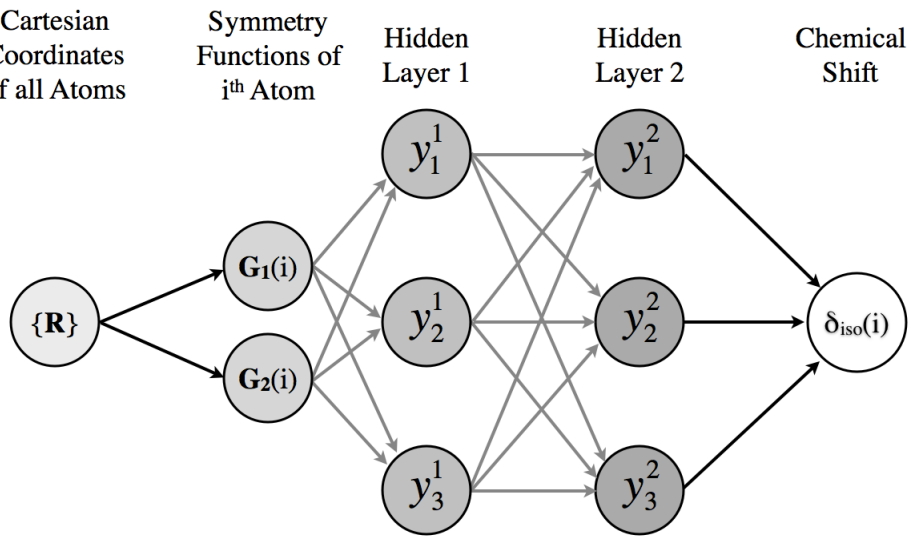


Figure 1: Schematic representation of a small feed-forward NN-NMR composed of two hidden layers of three nodes each. In input, all the Cartesian coordinates of a given frame are transformed into a set of symmetry functions $\{\mathbf{G(i)}\}$ for each atom. In the present scheme, two symmetry functions describe the chemical environment of the i^{th} atom.

To answer these requirements, and to develop an approach that is general and accurate enough to predict the NMR parameters of various solid-state compounds, we propose to generate a feed-forward NN following the same procedure as proposed by Behler and co-workers for the development of accurate potential energy surfaces.[64–67] To do so, we use two sets of calculated input parameters. Firstly, the structural data describing all the possible chemical environments in a given system are provided by MD simulations. These latter are performed on a finite number of small to middle size unit cells at different temperatures. From the resulting snapshots, the chemical environment of each atom is described by a set of generalized symmetry functions as introduced by Behler and Parrinello.[64] The second set of input parameters that is required are the chemical shift and quadrupolar coupling parameters of these atoms that are evaluated for each snapshot using an *ab initio* approach. The neural-network is then trained to associate to a given set of generalized symmetry functions the

corresponding NMR parameters leading to a general functional between the two sets of data as schematically represented in Fig. 1. Once properly trained, this functional, the NN-NMR, can be used to predict the NMR property of any input set of symmetry functions with an *ab initio* quality but at low computational cost. It is worth pointing out that this procedure introduced by Behler and co-workers was successfully applied to the development of NN potentials further allowing the study of various material science problems such as the graphite to diamond transition,[68,69] the high pressure behaviour of sodium,[70] and more recently the phase change material GeTe.[71–76]

2.2 Generation of an NN-NMR for Silica

Input Data. To train our NN-NMR, we considered five crystalline polymorphs of silica : α-quartz,[77] moganite,[78] rutile,[79] coesite[80] and ferrierite.[81] In contrast to the four others, rutile (a high pressure polymorph also called SiO_6 stishovite) is a compound in which the silicon atoms are not in a tetrahedral environment and was thus included to handle over-coordinated silicon and oxygen atoms. Furthermore, as amorphous structures are characterised by disordered local environments that can not be represented by crystalline phases, we also included data from two amorphous structures, noted AM1 and AM2, of 288 atoms each, resulting of two different annealing protocols which are described in the Supporting Information. It is worth pointing out that this number of atoms is of the order of the system sizes generally handle by pure DFT studies on glasses. For each of these structures, MD simulations were performed at three temperatures: 300 K, 700 K and 2000 K for crystalline phases or 1500 K for the amorphous slabs (see Supporting Information for details). A uniform selection of frames along these trajectories leads to a total of 154224 and 77112 independent local environments for oxygen and silicon, respectively.

Thereafter, each of these local environments was characterized by a set of symmetry functions defined using the scheme proposed by Behler and Parrinello.[64] Due to the complexity of the studied amorphous structures, besides the original two body radial and three body

angular symmetry functions, we also introduced some related to four body dihedral angles and atom type, these latter in order to distinguish between different elements. These functions are described in Supporting Information. Each atom is thus characterised by a set of 57 values (a larger set containing 155 symmetry functions was tested and did not lead to any significant improvement of our results), each one associated with one symmetry function, that is a unique fingerprint of the local environment up to 6.0 Å around this atom. To illustrate this assertion, Fig. 2 displays the distribution of two of these values for silicon in the five aforementioned crystalline phases. As can be seen, increasing the temperature expands the range of the values that are visited for each symmetry function. Furthermore, at a given temperature, each compound is characterised by a unique distribution of values showing that these two symmetry functions are already able to achieve a reasonable differentiation between the structures. Thus, the high dimensional formalism we use is expected to lead to an accurate description of the various local environments which could not be achieved using only a small set of order parameters.

In parallel, PAW and GIPAW calculations were conducted to evaluate the ^{17}O and ^{29}Si NMR parameters of the selected frames (see Supporting Information for details on these calculations).[15,18] As ^{29}Si is a spin one half nucleus and most of the experimental studies on silica are performed under magic-angle spinning (MAS) conditions, we chose to concentrate our analysis on ^{17}O δ_{iso}, ^{17}O C_Q and ^{29}Si δ_{iso} that are the most relevant parameters for this particular chemical system. As GIPAW calculations provide nuclear shielding tensors, the isotropic values have to be transformed to be expressed in terms of chemical shift which is the experimentally relevant parameter. In the following discussion, we defined ^{17}O δ_{iso} and ^{29}Si δ_{iso} by using the following formula: $\delta_{iso} = -(\sigma_{iso} - \sigma_{ref})$ with ^{17}O $\sigma_{ref} = 258.1$ ppm and ^{29}Si $\sigma_{ref} = 317.5$ ppm. These values were chosen in order to reproduce the experimental isotropic chemical shift of moganite.

The distributions of ^{29}Si δ_{iso} and ^{17}O δ_{iso} obtained from the PAW and GIPAW calculations for each compound at the three considered temperatures are depicted in Fig. 3 (the ^{17}O C_Q

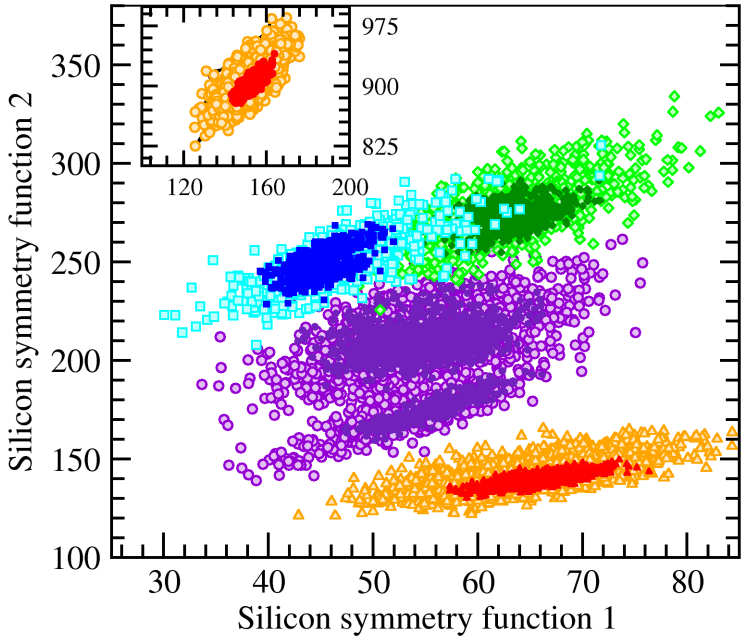


Figure 2: Values associated to two symmetry functions extracted from the frames used for the ^{29}Si δ_{iso} calculations of moganite (green diamonds), α-quartz (red triangles), coesite (blue squares) and ferrierite (violet circles). Values are presented for two temperatures, 300 K (dark colours) and 2000 K (light colours). In the inset, the same data are presented for rutile.

distributions are presented in Supporting Information Fig. S2). Fig. 3 (a) and (b) show that by studying several crystalline polymorphs of silica at different temperatures, we are able to explore a large range of both the ^{29}Si δ_{iso} and ^{17}O δ_{iso} spaces which is mandatory to ensure the accuracy and transferability of the NN-NMR. It is worth pointing out that the ^{29}Si δ_{iso} distributions are divided in two main groups: the high δ_{iso} values between -120 and -60 ppm corresponding to the four coordinated SiO_4 environments and the low δ_{iso} values between -180 and -140 ppm corresponding to the six coordinated SiO_6 ones. Furthermore, we see that the $^{29}Si_{\delta_{iso}}$ distribution of coesite is particularly shifted compared to the other tetrahedral crystalline phases. This comes from a specific local environment around the silicon atoms that is characterized by the four-membered ring depicted in Fig. 4 (a). As expected, the distributions obtained from the amorphous slabs (see Fig. 3 (c) and (d)) share common features with the distribution from the tetrahedral crystalline phases which demonstrates similarity between their local atomic environment. However the spread of the vitreous distributions is much larger than each crystalline structure taken independently which highlights the necessity to include several of them to widen the set of local environments included in the NN-NMR.

Finally, one can highlight that the ^{29}Si δ_{iso} and ^{17}O δ_{iso} distributions are significantly different for both AM1 and AM2. This demonstrates that a single amorphous SiO_2 structure of 288 atoms is not enough to converge such distributions. This drawback is all the more mandatory for more complex chemical compounds, *e.g.* for alkaline doped silica, that are characterised by a larger diversity of local environments. In that case, the simulation of larger unit cells becomes mandatory even to qualitatively describe their chemical structure.

Generation of the NN-NMR. The NN-NMR we employ is formed by two hidden layers of 25 nodes each. Sigmoid activation functions were used in the hidden layers and a linear function was used for the output layers. It was generated using as *training set* half of the aforementioned computed NMR parameters and corresponding sets of symmetry functions. The other half was used as *testing set*, to evaluate, during the fitting procedure,

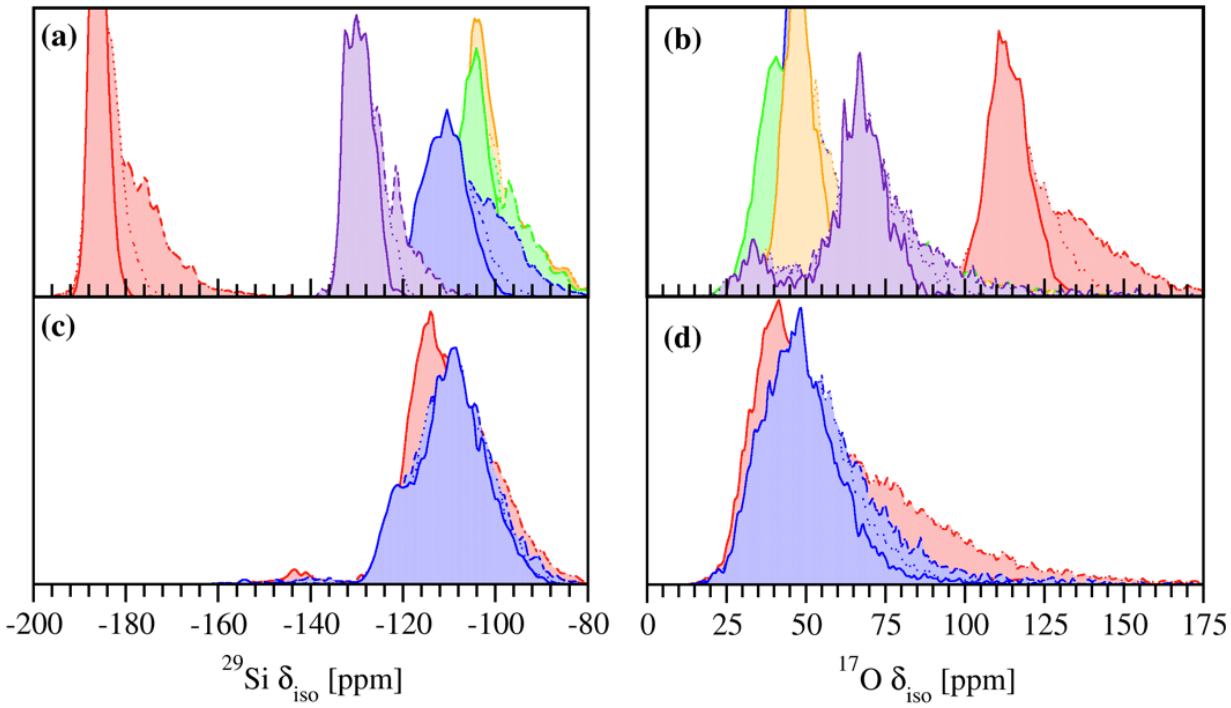


Figure 3: Distributions of ^{29}Si δ_{iso} (a) and ^{17}O δ_{iso} (b) values obtained from simulations at 300 K (plain line), 700 K (dotted line) and 2000 K (dashed line) for rutile (red), moganite (blue), ferrierite (green), α-quartz (orange) and coesite (indigo). Distribution of ^{29}Si δ_{iso} (c) and ^{17}O δ_{iso} (d) values obtained from simulations at 300 K (plain line), 700 K (dotted line) and 1500 K (dashed line) for AM1 (blue) and AM2 (red).

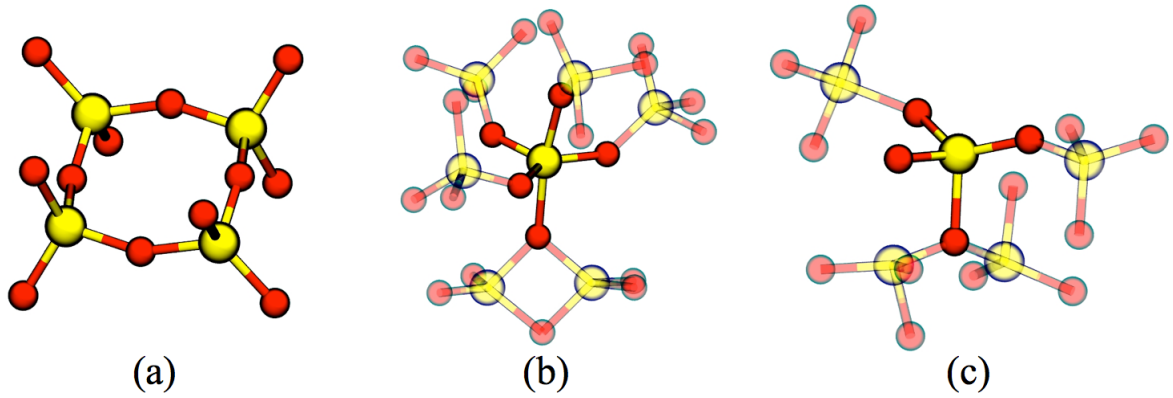


Figure 4: (a) Four-membered ring motif present in coesite. (b) Representative scheme of the triangular bi-pyramid environment of the five-coordinated silicon and the tri-coordinated oxygen atoms in amorphous silica. (c) Representative scheme of a single-coordinated oxygen atom and the corresponding tri-coordinated oxygen atoms found in amorphous silica. Oxygen and silicon atoms are represented by red and yellow spheres, respectively.

the performance of the NN-NMR on data not included in the *training set*. This prevents the fitting procedure from leading to a good fit of the *training set* while having a poor predictive power. The whole of input data at our disposal were thus equally and randomly split into the training and testing sets. Finally, an independent NN-NMR was generated for each NMR parameter for both oxygen and silicon.

To appreciate the capability of the NN-NMR to reproduce the *ab initio* data included in the training and testing sets, we looked at the cumulative histogram of the absolute NN-NMR errors compared to the *ab initio* values. These results are presented in Figure 5 for ^{29}Si and ^{17}O δ_{iso}, and in Fig. S3 of the Supporting Information for ^{17}O C_Q. In addition, Table 1 displays information concerning the mean absolute error obtained between the two sets of data. 80% of the silicon and oxygen atoms in the training and testing sets are characterised by an error in chemical shift that is lower or equal to 1.7 ppm and 2.9 ppm, respectively. The average error between the *ab initio* and NN-NMR values for the ^{29}Si and ^{17}O δ_{iso} parameters in the training set are 1.0 and 1.8 ppm, respectively. When compared to the total visited ^{29}Si and ^{17}O δ_{iso} ranges (which are 124 ppm and 172 ppm for ^{29}Si and ^{17}O δ_{iso}, respectively), this leads to relative errors of $\sim$1.0 % for both parameters. Errors on calculated isotropic chemical shift values due to the DFT approximations are typically 1-2% of the full chemical shift range of a given nucleus which highlights the quality of the the present results. As can be seen in Fig. S3 of the Supporting Information and in Table 1, the same quality is reached for the ^{17}O C_Q parameter. Clearly, these results demonstrate a very good fitting of the input data by the NN-NMR generation procedure which gives strong confidence in the NN-NMR approach. Furthermore, we would like to draw the reader attention on the fact that the fitting procedure is performed on fully converged *ab-initio* calculations, this convergence being easily reached due to the limited size of the systems we used to generate the input data. Thus, the quality of the predicted values does not suffer from an inaccurate convergence of the basis set or *k-point* sampling that can arise when using DFT based approaches on large systems at the limit of computational capabilities.

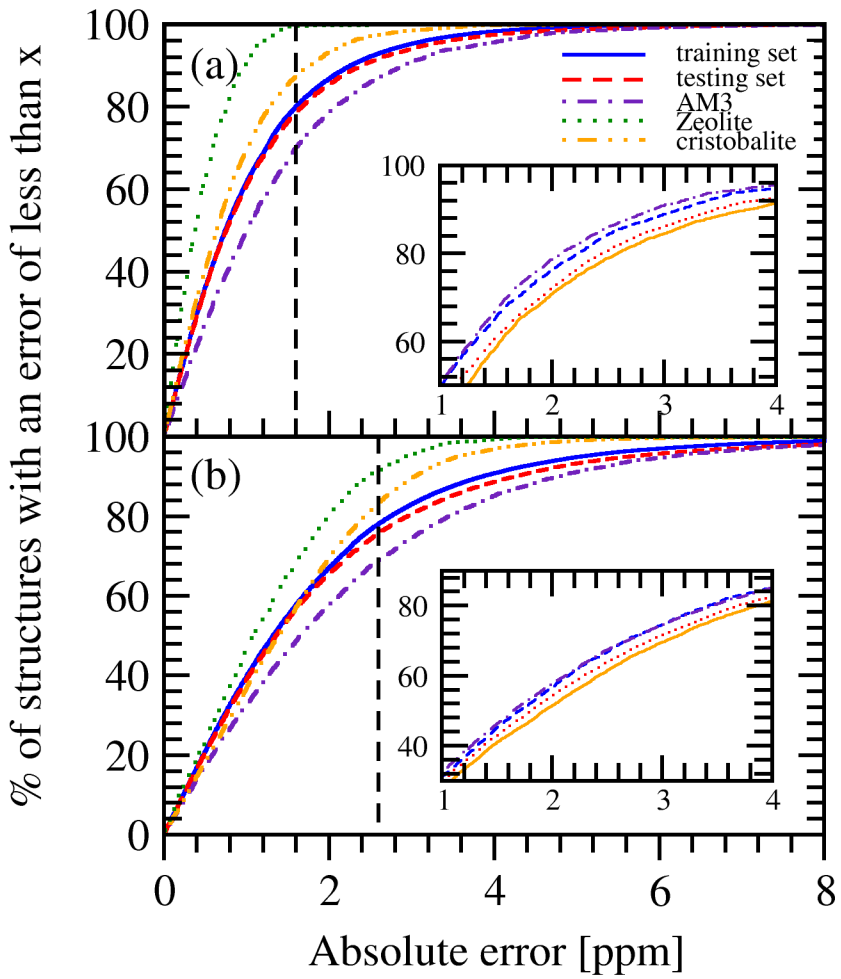


Figure 5: Cumulative histograms of the absolute NN-NMR errors compared to the *ab initio* values in both the training (blue plain) and testing (red dashed) sets for ^{29}Si δ_{iso} (a) and ^{17}O δ_{iso} (b). The dotted dashed indigo, dotted green and double dotted dashed orange lines present the predictive results for AM3, zeolite ZSM-11 and α-cristobalite, respectively. The vertical dashed black lines indicate the absolute error for which 80 % of the structures are below. The insets present the accuracy of the prediction for AM3 when the NN-NMR is produced from 60% (dashed blue), 20% (dotted red) and 5% (plain orange) of the total set of data (dotted dashed violet).

Table 1: Mean absolute error (in ppm for δ_{iso} and in MHz for C_Q) for the training (Roman) and testing (*italic*) sets for the generation of the NN-NMR and for the predicted δ_{iso} and C_Q values for α-cristoballite, zeolite ZSM-11 and AM3. These compounds were not included in the generation of the NN-NMR and thus represent purely predictive results.

	NN-NMR	α-cristoballite	Zeolite ZSM-11	AM3
^{29}Si δ_{iso}	1.0 / *1.1*	0.80	0.50	1.3
^{17}O δ_{iso}	1.8 / *1.9*	1.6	1.2	2.2
^{17}O C_Q	0.14 / *0.14*	0.07	0.06	0.15

NN-NMR Test. To demonstrate the predictive ability of the presently generated NN-NMR, we applied it to three compounds that were not included in the training and testing sets: the α-cristoballite,[82] the zeolite ZSM-11,[83] and a third amorphous silica structure of 432 atoms, called AM3 (details of its generation are provided in Supporting Information). For each of these three structures, MD simulations were performed at 300 K and 700 K. On a randomly selected number of frames, both *ab initio* and NN-NMR evaluation of the NMR parameters were performed and compared. The results are presented in Figure 5 for ^{29}Si and ^{17}O δ_{iso}, and in Fig. S3 fo the Supporting Information for ^{17}O C_Q. Table 1 displays information concerning the average error obtained between the two sets of data. For the two crystalline phases, the quality of the prediction is very good as the profile of the cumulative histograms of the errors is better than the ones of the testing and training sets. It is worth pointing out that the ability of the present NN-NMR to accurately predict the NMR parameters of zeolite ZSM-11 is extremely encouraging as only one zeolite-type structure was used to train the NN-NMR. This is all the more interesting that zeolites represent a class of silica materials with very specific structural properties, *e.g.* large cavities that allow to accommodate various molecular aggregates used for catalytic, oil refining and decontamination applications,[84–88] it is thus rather encouraging to see that our approach is able to describe them correctly.

The quality of the prediction for AM3 compared to the *ab initio* results is also good, with average errors of 1.3 and 2.2 ppm for ^{29}Si and ^{17}O δ_{iso} despite the fact that the AM3 phase is totally independent from AM1 and AM2. This accuracy is achieved because the NN-NMR was trained by including a diversity of crystalline and amorphous structures that allows for an efficient exploration of the local environments, *i.e.* of the set of symmetry functions describing the SiO_2 chemical system. Figure 6 displays the absolute NN-NMR prediction error on the ^{29}Si δ_{iso} and ^{17}O δ_{iso} parameters as a function of the calculated GIPAW values. These two pictures show that there is no asymmetries in the NN-NMR prediction error over the visited GIPAW ^{29}Si and ^{17}O chemical shift ranges. Indeed, although some

regions are more visited than others, the maximum errors hardly vary along them. This demonstrates that the various chemical environments are described with similar accuracy. The only exception for the ^{29}Si δ_{iso} distribution is around -140/150 ppm that displays a slightly higher error due to the fewer number of data in this particular region. For the ^{17}O δ_{iso} distribution, the region above 80 ppm also displays a slightly higher error for the same reason as it corresponds to the tail of the distribution (see below). Overall, one has to keep in mind that when comparing *ab initio* calculated NMR parameters with experimental measurements, a difference lower than 10% is generally considered as accurate. Consequently, our NN-NMR results clearly do not decrease the quality of the *ab initio* data and can thus be considered as highly accurate.

We finally checked how the quality of the NN-NMR varies as a function of the number of input data in the training set. Thus, to give a guide line of how many structures are required for training a high quality NN-NMR, we generated a set of NN-NMR by progressively decreasing the number of input local environments used for the generation. We generated three new NN-NMR from 60%, 20% and 5% of the initial set of local environments and applied them on the AM3 phase. The comparison with the *ab initio* results are presented in the two insets of Fig. 5. Clearly, the accuracy of the NN-NMR only starts to decrease at 20% of the amount of input data. This suggests that the amount of input data we included in the NN-NMR generation was large enough to capture the different chemical environments visited in the various phases. However, as this amount is not known *a priori*, it is necessary to check it along the course of the generation as it likely varies with the complexity of the studied chemical system. In our particular case, as the number of input data is not a source of inaccuracy, improving the NN-NMR would require to include more independent structures in the generation procedure, in particular, more amorphous slabs. It worth pointing out that the actual exploration of the chemical environments of this system is largely satisfactory as among the twenty-two AM3 frames we used, only 31 silicon and 3 oxygen atoms display symmetry functions values that are out of the range of the generated NN-NMR. Thus, the

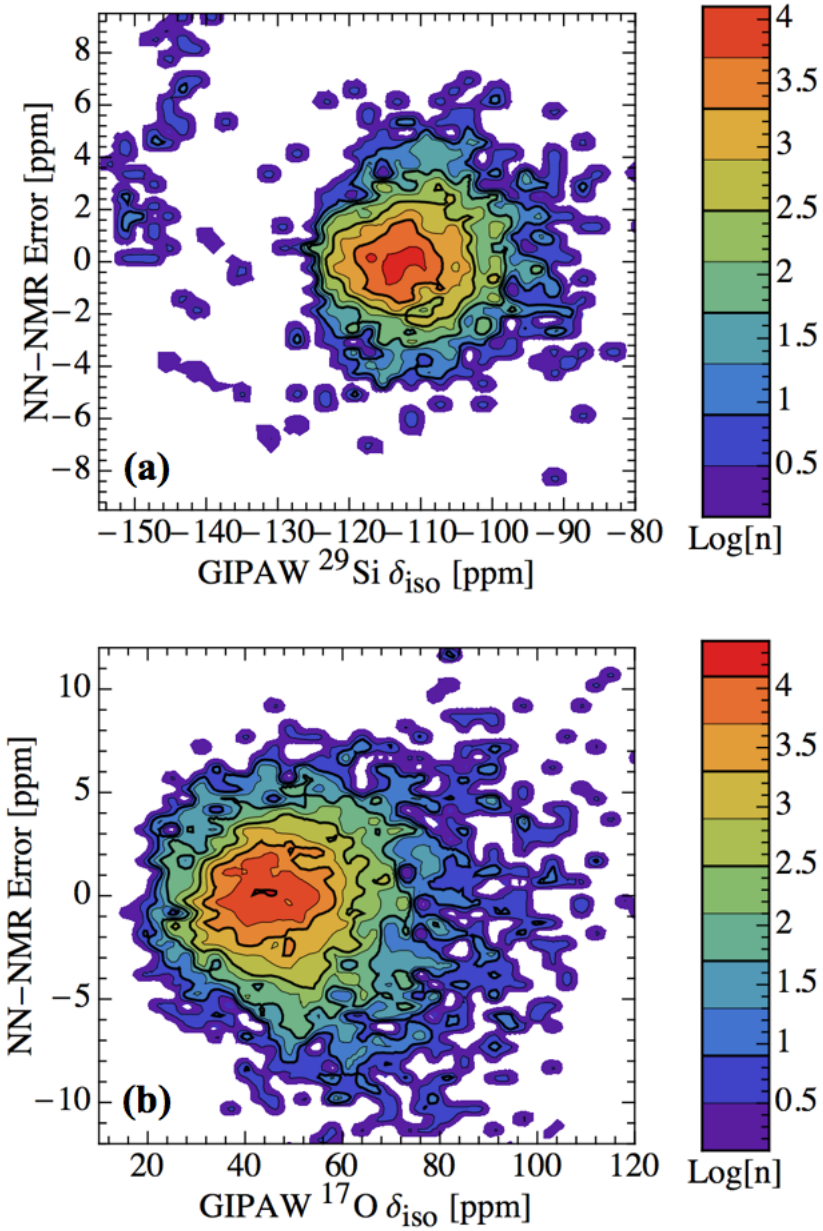


Figure 6: Distribution of the absolute NN-NMR prediction error on the ^{29}Si δ_{iso} (a) and ^{17}O δ_{iso} (b) parameters as a function of the GIPAW values. Data are presented for the AM3 amorphous phase at 300 and 700 K.

local environments of the large amorphous structures are well described by the data set constructed from many small systems. This demonstrates the validity of our approach, as it shows that an NN-NMR can be produced quickly and accurately using a relatively small number of input local environments. Once generated, the application of the NN-NMR requires a few seconds and around thirty seconds per frame for a 2592 and a 10368 atoms unit cell, respectively. More importantly, it behaves almost linearly with the number of atoms within the unit cell. Hence, the NN-NMR approach overcomes the computational bottleneck involving the *ab initio* evaluation of NMR parameter in large systems.

3 Application and Discussion

Vitreous silica materials are generally characterised by cavities that can span several tens of angstroms in length and that can tune their physico-chemical properties.[89–91] Thus, an accurate modelling of such materials requires large system sizes and as a first application of the NN-NMR method, we tried to evaluate the convergence of the ^{29}Si δ_{iso}, ^{17}O δ_{iso} and ^{17}O C_Q distributions in amorphous silica slabs as function of the size of the simulated system and of the annealing protocol used for its generation. Pure DFT based approaches are limited to answer such questions as this requires the use of both large system sizes and application to a large number of slabs which would lead to a tremendous computational cost.

To proceed, we used three new slabs of 2592 atoms each generated with different annealing protocols (see Supporting Information for details) and referred to as AM4, AM5 and AM6. A last structure, named AM7, containing 10368 atoms was also produced. To the best of our knowledge, this is the largest silica structures ever studied using an *ab initio* quality procedure for the calculation of the NMR parameters. For each structure, molecular dynamics simulations were conducted at 300 K and 500 K leading to trajectories on which the NN-NMR was applied to a number of snapshots. The predicted distributions at 300 K are plotted in Fig. 7. For the three different parameters, the curves are very similar. This

demonstrates that the distribution we obtain are independent of the annealing protocol used to generate the amorphous structure, at least when amorphous slabs of several thousand of atoms are used. Furthermore, the curves of Fig. 7 demonstrate that an almost perfect convergence of the distributions is obtained for systems containing ∼2000 atoms. This good convergence of the NMR parameters has to be compared to the convergence of the structural parameters characteristic of the different slabs. One quantity that is generally looked at to evaluate such a convergence is the radial distribution functions (*g(r)*). The O-O, Si-O and Si-Si *g(r)* of AM4, AM5, AM6 and AM7 are presented in Fig. 8. The four slabs clearly display similar radial distribution functions which support the assertion that slabs containing ∼2000 atoms are structurally well converged and thus lead to similar distribution of NMR parameters.

To test that small unit cells are not able to converge the ^{29}Si δ_{iso}, ^{17}O δ_{iso} and ^{17}O C_Q silica distributions, we applied the NN-NMR on the AM1 and AM2 phases at 300 K. The corresponding curves are provided in Fig. S4 of the Supporting Information. It is worth pointing out that the NN-NMR approach allows to properly converge these curves which is hardly feasible using an *ab initio* approach due to computational limitations. The AM1 and AM2 distributions display clear differences with the ones of AM7 which demonstrates that small unit cells are not sufficient to statistically converge the proportion of local environments characteristic of the SiO_2 chemical system. This is further highlighted in their radial distribution functions presented in Fig. S5 in the Supporting Information, which also show differences with the curves of Fig. 8, in particular the Si-Si *g(r)*.

In Fig. 7 (a), the ^{29}Si δ_{iso} distributions are presented along with the experimental spectrum obtained by Clark *et al.*[92] This latter is very close to the simulated spectra, the only difference being a small broadening of the distribution in all the simulated spectra compared to the experimental one. This demonstrates that the NN-NMR approach retains the *ab initio* quality of the calculated NMR parameters and thus the NN-NMR results can be confidently compared to experimental data.

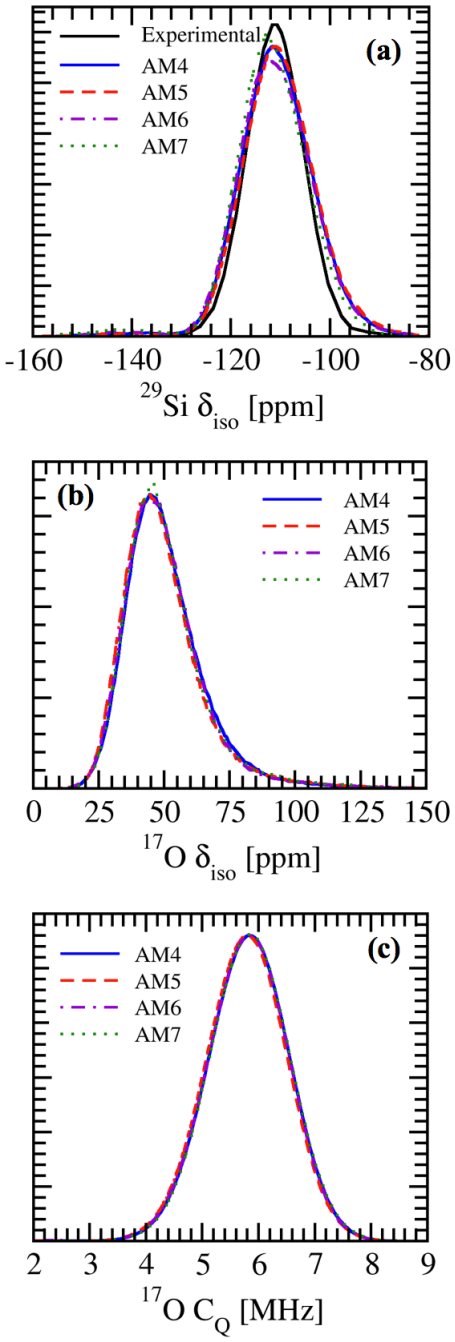


Figure 7: Distribution of ^{29}Si δ_{iso} (a), ^{17}O δ_{iso} (b) and ^{17}O C_Q (c) obtained from the AM4, AM5, AM6 and AM7 structures at 300 K. For ^{29}Si δ_{iso}, the distribution can be directly compared to experimental data. Picture (a) also contains the experimental spectrum obtained by Clark *et al.*[92]

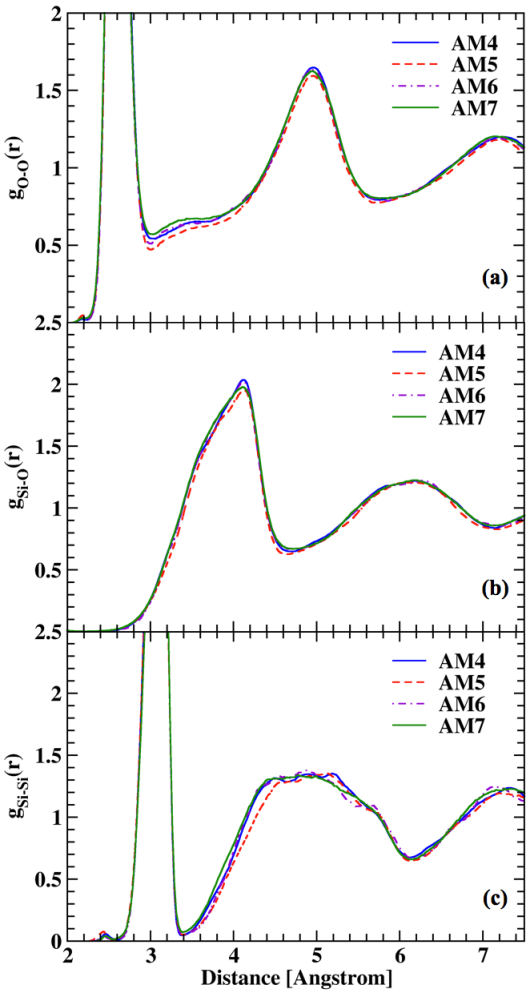


Figure 8: Radial distribution functions for O-O (a), Si-O (b) and Si-Si (c) of the AM4, AM5, AM6 and AM7 phases obtained at 300 K.

In the ^{29}Si δ_{iso} distributions of the amorphous phases (see Fig. 7 (a) and also Fig. 3 (c)) there appears a small signal around -140 ppm that is absent in the experimental spectrum. To elucidate the origin of this signal, we examined the correlation between the isotropic chemical shift of the silicon atoms and their coordination number with the surrounding oxygen atoms. This result is shown in Fig. 9 (a) for the AM7 amorphous phase at 300 K (the same plots obtained for the AM4 phase at 300 and 500 K are provided in Supporting Information). First of all, this picture reveals the absence of six-coordinated silicon environments in this vitreous structure. This is consistent with the absence of chemical shift values lower than -150 ppm as observed in stishovite. As expected, the great majority of the local environments corresponds to four-coordinated silicon and two-coordinated oxygen atoms. However, the ^{29}Si δ_{iso} peak at $\sim$-140 ppm in the silicon distribution comes from five-coordinated silicon atoms as displays in Fig. 9 (a). In Fig. 4 (b), a representative scheme of such a chemical feature is depicted and can be described as a distorted triangular bi-pyramidal environment. Furthermore, as also revealed by Fig. 4 (b), this over-coordinated silicon is associated with a tri-coordinated oxygen atom. The NMR signature of these latter is provided by Fig. 9 (b) that shows a non-negligible number of such chemical environments. However, their ^{17}O δ_{iso} is not distinct from the two-coordinated oxygen atoms as their signature is encompassed in the high chemical shift tail of the ^{17}O δ_{iso} distribution. Furthermore, their ^{17}O quadrupolar coupling constant is also not distinguishable from the main distribution as it falls in the same range as the two-coordinated oxygen atoms. It is worth pointing out that experimentally, five-coordinated silicon atoms have never been detected under ambient pressure conditions. Pressures higher than 8-10 GPa are generally required to stabilised this chemical feature. Therefore, their existence in our simulations, although in minority, is attributed to the BKS potential we used to generate the amorphous slabs. Previous studies performed using this BKS force-field showed that it over estimates the amount of five-coordinated silicon compared to *ab initio* MD.[40,43,93] It is however interesting to note that Stebbins evidenced such local environment under ambient pressure in $K_2Si_4O_9$ using ^{29}Si SSNMR and that the spectrum he

obtained is similar to the distributions presented in Fig. 7 (a).[94] This clearly demonstrates the ability of the NN-NMR approach to accurately describe different coordination-states of silicon. Lastly, Fig. 9 (a), (b) and (c) show a tail in the low-coordination region that correspond to transient states. When looking specifically for these low coordinated atoms, we found only one dangling oxygen atom the structure of which is depicted in Fig. 4 (c).

To go further into the ability of the NN-NMR to accurately handle different coordination states of silicon we evaluated the ^{29}Si δ_{iso} distribution of the stishovite from a MD simulation at 300 K. The signal appears in the low chemical shift region and is centered around -170 ppm. Of course, due the specific chemical environment of silicon in stishovite, this separation with the four coordinated phases is expected and has already been observed in previous ^{29}Si SSNMR experiments of stishovite.[95] In this work, Xue *et al.* performed measurements under ambient pressure and temperature conditions of a stishovite sample prepared at 9.5 GPa. They showed that the ^{29}Si NMR signal of their sample was 83.0 ppm lower than the one obtained for an amorphous four coordinated silica impurity. Our calculations confirm these observations (see above for the results on amorphous structures) as we calculate a difference of about 70-80 ppm. This demonstrates the ability of the NN-NMR to accurately handle different coordination states of the silicon atoms.

4 Conclusions

In conclusion, we present in this work an approach to predict *ab initio* NMR parameters using a high dimensional neural-network formalism. Once properly trained on a set of representative structures, it allows for an accurate and efficient prediction of the NMR parameters of large system sizes with an accuracy that is very close to *ab-initio* quality. Furthermore, it can also be applied to a very large number of small to middle size unit cells which can not be achieved using DFT based approaches due to the tremendous cost in computational power that requires. It thus answers the requirement recently expressed in a review by

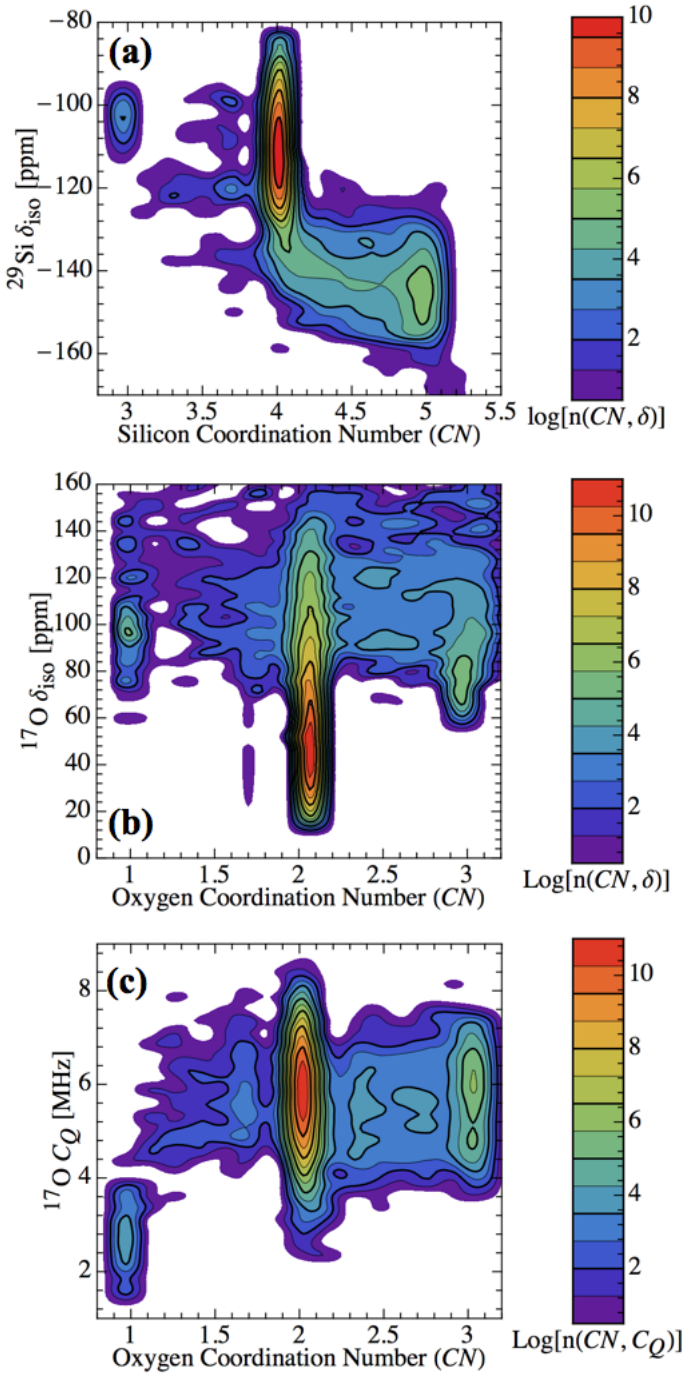


Figure 9: Correlation between ^{29}Si δ_{iso} (a), ^{17}O δ_{iso} (b) ^{17}O C_Q (c) and the coordination number of the corresponding silicon and oxygen atoms obtained from the AM7 structure at 300 K. The coordination number of an atom was defined using the following switching function: $CN(r) = (1-(\frac{r-1.6}{0.25})^{12})/(1-(\frac{r-1.6}{0.25})^{16})$. Similar plots obtained for the AM4 amorphous structure at 300 and 500 K are provided in Supporting Information for comparison.

Charpentier and co-worker to increase the size of the studied systems to be able to further extend the range of potential applications of the MD-GIPAW methodology.[21] The NN-NMR approach was tested and validated on the ^{17}O and ^{29}Si quadrupolar coupling and chemical shift parameters of various crystalline and amorphous silica phases. In this latter case, several studies have shown that very good correlations can be found between NMR parameters and local geometric parameters which allows for an efficient mapping between SSNMR spectra and structural properties.[23,24,28] However, this is not true for multicomponent glasses for which such correlations are more difficult to define.[96,97] The NN-NMR approach will thus be extremely valuable to investigate the chemistry of such complex systems. Indeed, it is, in principle, general and can be applied to any chemical system and any NMR parameter such as the *J-coupling* or the anisotropic part of the chemical shift tensor. In combination with *ab initio* molecular dynamics approaches or accurate force-fields simulations, it has the potential to study finite temperature properties of complex materials such as amorphous phases or surfaces. It can also be coupled to isothermal-isobaric ensemble simulations to performed pressure dependent studies which require large system sizes. In the spirit of the NMR-crystallography approach,[98] it could be also used to refine large an complex structures or, in combination with efficient minimisation algorithms, to generate a range of possible structures corresponding to a given distribution of NMR parameters. In addition, improvement in the NN-NMR generation could be done by exploring the local environments of a given chemical system not by using molecular dynamics simulations only but also by using efficient structure prediction algorithms such as the Ab-Initio Random Structure Searching (AIRSS).[99]

Supporting Information Available

Description of four body dihedral angles and atomic types based symmetry functions. Details on the molecular dynamics simulations performed on all the crystalline and amorphous structures studied in this work as well as details on the PAW and GIPAW calculations of

the nuclear shieldings and quadrupolar coupling constants. ^{17}O C_Q distributions obtained from the PAW calculations and cumulative histograms of the absolute NN-NMR error in the prediction of ^{17}O C_Q. Radial distribution functions and converged ^{29}Si δ_{iso}, ^{17}O δ_{iso} and ^{17}O C_Q NN-NMR distributions for AM1 and AM2 and the coordination number analysis of the AM4 phase at 300 K and 500 K are also provided. This material is available free of charge via the Internet at `http://pubs.acs.org/`.

Acknowledgement

The authors gratefully acknowledge Michele Parrinello for providing the original neural-network program and for useful discussions and reading of the manuscript. The authors would also like to thank Michele Ceriotti for his help with the creation of the pictures. Computational time for this work was provided by the Swiss National Supercomputing Center (CSCS). The authors also thank the supercomputing facility of Toulouse III University, CALMIP, for the allocation of computer resources (P1320).

References

(1) MacKenzie, K. J. D.; Smith, M. E. *Multinuclear Solid-State NMR of Inorganic Materials*; Pergamon: Oxford, 2002; pp 201–616.

(2) Bakhmutov, V. I. *Solid-State NMR in Materials Science: Principles and Applications*; CRC Press, 2011; pp 231–257.

(3) Wasylishen, R. E., Ashbrook, S. E., Wimperis, S., Eds. *NMR of Quadrupolar Nuclei in Solid Materials*; Wiley-Blackwell, 2012.

(4) Frydman, L.; Harwood, J. S. *J. Am. Chem. Soc.* **1995**, *117*, 5367–5368.

(5) Medek, A.; Harwood, J. S.; Frydman, L. *J. Am. Chem. Soc.* **1995**, *117*, 12779–12787.

(6) Fernandez, C.; Amoureux, J. *Solid State Nucl. Magn. Reson.* **1996**, *5*, 315–321.

(7) Amoureux, J.-P.; Fernandez, C.; Steuernagel, S. *J. Magn. Reson., Ser A* **1996**, *123*, 116–118.

(8) Chmelka, B. F.; Mueller, K. T.; Pines, A.; Stebbins, J.; Wu, Y.; Zwanziger, J. W. *Nature* **1989**, *339*, 42–43.

(9) Farnan, I.; Grandinetti, P. J.; Baltisberger, J. H.; Stebbins, J. F.; Werner, U.; Eastman, M. A.; Pines, A. *Nature* **1992**, *358*, 31–35.

(10) Le Bourhis, E. *Glass; Mechanics and Technology*; Wiley-VCH Verlag GmbH & Co. KGaA, 2007; Chapter 4, pp 39–51.

(11) Sales, B. C.; Boatner, L. A. *Science* **1984**, *226*, 45–48.

(12) Lutze, W.; Malow, G.; Ewing, R. C.; Jercinovic, M. J.; Keil, K. *Nature* **1985**, *314*, 252–255.

(13) Mauri, F.; Pfrommer, B. G.; Louie, S. G. *Phys. Rev. Lett.* **1996**, *77*, 5300–5303.

(14) Sebastiani, D.; Parrinello, M. *J. Phys. Chem. A* **2001**, *105*, 1951–1958.

(15) Pickard, C. J.; Mauri, F. *Phys. Rev. B: Condens. Matter* **2001**, *63*, 245101.

(16) Thonhauser, T.; Ceresoli, D.; Mostofi, A. A.; Marzari, N.; Resta, R.; Vanderbilt, D. *J. Chem. Phys.* **2009**, *131*, 101101–4.

(17) Laskowski, R.; Blaha, P. *Phys. Rev. B: Condens. Matter* **2012**, *85*, 035132.

(18) Blöchl, P. E. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1994**, *50*, 17953–17979.

(19) Thibault, C. *Solid State Nucl. Magn. Reson.* **2011**, *40*, 1–20.

(20) Bonhomme, C.; Gervais, C.; Babonneau, F.; Coelho, C.; Pourpoint, F.; Azaï, T.; Ashbrook, S. E.; Griffin, J. M.; Yates, J. R.; Mauri, F.; Pickard, C. J. *Chem. Rev.* **2012**, *112*, 5733–5779.

(21) Charpentier, T.; Menziani, M. C.; Pedone, A. *RSC Adv.* **2013**, *3*, 10550–10578.

(22) Charpentier, T.; Ispas, S.; Profeta, M.; Mauri, F.; Pickard, C. J. *J. Phys. Chem. B* **2004**, *108*, 4147–4161.

(23) Benoit, M.; Profeta, M.; Mauri, F.; Pickard, C. J.; Tuckerman, M. E. *J. Phys. Chem. B* **2005**, *109*, 6052–6060.

(24) Charpentier, T.; Kroll, P.; Mauri, F. *J. Phys. Chem. C* **2009**, *113*, 7917–7929.

(25) Ispas, S.; Charpentier, T.; Mauri, F.; Neuville, D. *Solid State Sci.* **2010**, *12*, 183–192.

(26) Pedone, A.; Charpentier, T.; Menziani, M. C. *Phys. Chem. Chem. Phys.* **2010**, *12*, 6054–6066.

(27) Soleilhavoup, A.; Delaye, J.-M.; Angeli, F.; Caurant, D.; Charpentier, T. *Magn. Reson. Chem.* **2010**, *48*, S159–S170.

(28) Angeli, F.; Villain, O.; Schuller, S.; Ispas, S.; Charpentier, T. *Geochim. Cosmochim. Acta* **2011**, *75*, 2453–2469.

(29) Angeli, F.; Villain, O.; Schuller, S.; Charpentier, T.; de Ligny, D.; Bressel, L.; Wondraczek, L. *Phys. Rev. B: Condens. Matter* **2012**, *85*, 054110.

(30) Gambuzzi, E.; Pedone, A.; Menziani, M. C.; Angeli, F.; Caurant, D.; Charpentier, T. *Geochim. Cosmochim. Acta* **2014**, *125*, 170–185.

(31) Gambuzzi, E.; Pedone, A.; Menziani, M. C.; Angeli, F.; Florian, P.; Charpentier, T. *Solid State Nucl. Magn. Reson.* **2015**, *68–69*, 31–36.

(32) Forler, N.; Vasconcelos, F.; Cristol, S.; Paul, J.-F.; Montagne, L.; Charpentier, T.; Mauri, F.; Delevoye, L. *Phys. Chem. Chem. Phys.* **2010**, *12*, 9053–9062.

(33) Vasconcelos, F.; Cristol, S.; Paul, J.-F.; Delevoye, L.; Mauri, F.; Charpentier, T.; Caër, G. L. *J. Phys.: Condens. Matter* **2013**, *25*, 255402.

(34) Kibalchenko, M.; Yates, J. R.; Pasquarello, A. *J. Phys.: Condens. Matter* **2010**, *22*, 145501.

(35) Kibalchenko, M.; Yates, J. R.; Massobrio, C.; Pasquarello, A. *J. Phys. Chem. C* **2011**, *115*, 7755–7759.

(36) Sykina, K.; Bureau, B.; Le Polles, L.; Roiland, C.; Deschamps, M.; Pickard, C. J.; Furet, E. *Phys. Chem. Chem. Phys.* **2014**, *16*, 17975–17982.

(37) Lee, Y. J.; Bingöl, B.; Murakhtina, T.; Sebastiani, D.; Meyer, W. H.; Wegner, G.; Spiess, H. W. *J. Phys. Chem. B* **2007**, *111*, 9711–9721.

(38) Kins, C. F.; Dudenko, D.; Sebastiani, D.; Brunklaus, G. *Macromolecules* **2010**, *43*, 7200–7211.

(39) Huff, N. T.; Demiralp, E.; Çagin, T.; Goddard III, W. A. *J. Non-Cryst. Solids* **1999**, *253*, 133–142.

(40) Tangney, P.; Scandolo, S. *J. Chem. Phys.* **2002**, *117*, 8898–8904.

(41) Dávila, L. P.; Caturla, M.-J.; Kubota, A.; Sadigh, B.; Diáz de la Rubia, T.; Shackelford, J. F.; Risbud, S. H.; Garofalini, S. H. *Phys. Rev. Lett.* **2003**, *91*, 205501.

(42) Trachenko, K.; Dove, M. T. *Phys. Rev. B: Condens. Matter* **2003**, *67*, 064107.

(43) Liang, Y.; Miranda, C. R.; Scandolo, S. *Phys. Rev. B: Condens. Matter* **2007**, *75*, 024205.

(44) Meiler, J. *J. Biomol. NMR* **2003**, *26*, 25–37.

(45) Shen, Y.; Bax, A. *J. Biomol. NMR* **2007**, *38*, 289–302.

(46) Bergna, H. E., Roberts, W. O., Eds. *Colloidal Silica : Fundamentals and Applications*; CRC Press, 2006; pp 711–852.

(47) Papirer, E., Ed. *Adsorption on Silica Surfaces*; Marcel Dekker, 2000; pp 665–746.

(48) Saika-Voivod, I.; Sciortino, F.; Grande, T.; Poole, P. H. *Philos. Trans. R. Soc. A-Math. Phys. Eng. Sci.* **2005**, *363*, 525–535.

(49) Takada, A. *Advances in Glass and Optical Materials II*; John Wiley & Sons, Inc., 2006; pp 179–187.

(50) Pagliaro, M. *Silica-Based Materials for Advanced Chemical Applications*; RSC Publishing, 2009.

(51) Li, Z.; Barnes, J. C.; Bosoy, A.; Stoddart, J. F.; Zink, J. I. *Chem. Soc. Rev.* **2012**, *41*, 2590–2605.

(52) Vivero-Escoto, J. L.; Huxford-Phillips, R. C.; Lin, W. *Chem. Soc. Rev.* **2012**, *41*, 2673–2685.

(53) McCulloch, W. S.; Pitts, W. *Bull. Math. Biophys.* **1943**, *5*, 115–133.

(54) Bishop, C. M. *Neural Networks for Pattern Recognition*; Oxford University Press: Oxford, 1995.

(55) Haykin, S. *Neural Networks and Learning Machines*, 3rd ed.; Prentice Hall: New York, 2009.

(56) Hornik, K.; Stinchcombe, M.; White, H. *Neural Netw.* **1989**, *2*, 359–366.

(57) Hornik, K. *Neural Netw.* **1991**, *4*, 251–257.

(58) Attali, J.-G.; Pages, G. *Neural Netw.* **1997**, *10*, 1069–1081.

(59) Scarselli, F.; Chung Tsoi, A. *Neural Netw.* **1998**, *11*, 15–37.

(60) Abhishek, K.; Singh, M.; Ghosh, S.; Anand, A. *Procedia Technology* **2012**, *4*, 311–318.

(61) Litta, A. J.; Mary Idicula, S.; Mohanty, U. C. *International Journal of Atmospheric Sciences* **2013**, *2013*, 14.

(62) Gasteiger, J.; Zupan, J. *Angew. Chem. Int. Ed. Engl.* **1993**, *32*, 503–527.

(63) Pickard, C. J.; Payne, M. C.; Brown, L. M.; Gibbs, M. N. In *ELECTRON MICROSCOPY AND ANALYSIS 1995*; Cherns, D., Ed.; IOP PUBLISHING LTD, 1995; Vol. 147; pp 211–214.

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

(65) Behler, J. *Phys. Chem. Chem. Phys.* **2011**, *13*, 17930–17955.

(66) Handley, C. M.; Behler, J. *Eur. Phys. J. B* **2014**, *87*, 1–16.

(67) Behler, J. *Int. J. Quantum Chem.* **2015**, *115*, 1032–1050.

(68) Khaliullin, R. Z.; Eshet, H.; Kühne, T. D.; Behler, J.; Parrinello, M. *Phys. Rev. B: Condens. Matter* **2010**, *81*, 100103.

(69) Khaliullin, R. Z.; Eshet, H.; Kühne, T. D.; Behler, J.; Parrinello, M. *Nature Mater.* **2011**, *10*, 693–697.

(70) Eshet, H.; Khaliullin, R. Z.; Kühne, T. D.; Behler, J.; Parrinello, M. *Phys. Rev. Lett.* **2012**, *108*, 115701.

(71) Sosso, G. C.; Behler, J.; Bernasconi, M. *Phys. Status Solidi B* **2012**, *249*, 1880–1885.

(72) Sosso, G. C.; Miceli, G.; Caravati, S.; Behler, J.; Bernasconi, M. *Phys. Rev. B: Condens. Matter* **2012**, *85*, 174103.

(73) Sosso, G. C.; Donadio, D.; Caravati, S.; Behler, J.; Bernasconi, M. *Phys. Rev. B: Condens. Matter* **2012**, *86*, 104301.

(74) Sosso, G. C.; Colombo, J.; Behler, J.; Del Gado, E.; Bernasconi, M. *J. Phys. Chem. B* **2014**, *118*, 13621–13628.

(75) Campi, D.; Donadio, D.; Sosso, G. C.; Behler, J.; Bernasconi, M. *J. Appl. Phys.* **2015**, *117*, 015304.

(76) Sosso, G. C.; Salvalaglio, M.; Behler, J.; Bernasconi, M.; Parrinello, M. *J. Phys. Chem. C* **2015**, *119*, 6428–6434.

(77) D'Amour, H.; Denner, W.; Schulz, H. *Acta Cryst. B* **1979**, *35*, 550–555.

(78) Miehe, G.; Graetsch, H. *Eur. J. Mineral.* **1992**, *4*, 693–706.

(79) Baur, W. H.; Khan, A. A. *Acta Cryst. B* **1971**, *27*, 2133–2139.

(80) Araki, T.; Zoltai, T. *Z. Kristall.* **1969**, *129*, 381–387.

(81) Morris, R. E.; Weigel, S. J.; Henson, N. J.; Bull, L. M.; Janicke, M. T.; Chmelka, B. F.; Cheetham, A. K. *J. Am. Chem. Soc.* **1994**, *116*, 11849–11855.

(82) Downs, R. T.; Palmer, D. C. *Am. Mineral.* **1994**, *79*, 9–14.

(83) Toby, B. H.; Eddy, M. M.; Fyfe, C. A.; Kokotailo, G. T.; Strobl, H.; Cox, D. E. *J. Mater. Res.* **1988**, *3*, 563–569.

(84) Corma, A.; Diaz-Cabanas, M. J.; Martinez-Triguero, J.; Rey, F.; Rius, J. *Nature* **2002**, *418*, 514–517.

(85) Ray, S.; Vasudevan, S. *Inorg. Chem.* **2003**, *42*, 1711–1719.

(86) Choi, M.; Na, K.; Kim, J.; Sakamoto, Y.; Terasaki, O.; Ryoo, R. *Nature* **2009**, *461*, 246–249.

(87) Shang, J.; Li, G.; Singh, R.; Xiao, P.; Liu, J. Z.; Webley, P. A. *J. Phys. Chem. C* **2013**, *117*, 12841–12847.

(88) Primo, A.; Garcia, H. *Chem. Soc. Rev.* **2014**, *43*, 7548–7561.

(89) Armatas, G. S.; Kanatzidis, M. G. *Nature* **2006**, *441*, 1122–1125.

(90) Lin, Y.-S.; Haynes, C. L. *J. Am. Chem. Soc.* **2010**, *132*, 4834–4842.

(91) Kim, M.-H.; Na, H.-K.; Kim, Y.-K.; Ryoo, S.-R.; Cho, H. S.; Lee, K. E.; Jeon, H.; Ryoo, R.; Min, D.-H. *ACS Nano* **2011**, *5*, 3568–3576.

(92) Clark, T. M.; Grandinetti, P. J.; Florian, P.; Stebbins, J. F. *Phys. Rev. B: Condens. Matter* **2004**, *70*, 064202.

(93) Trave, A.; Tangney, P.; Scandolo, S.; Pasquarello, A.; Car, R. *Phys. Rev. Lett.* **2002**, *89*, 245504.

(94) Stebbins, J. F. *Nature* **1991**, *351*, 638–639.

(95) Xue, X.; Stebbins, J. F.; Kanzaki, M. *Phys. Chem. Minerals* **1993**, *19*, 480–485.

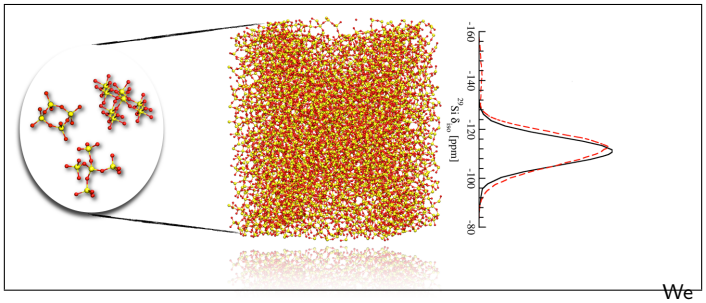
(96) Pedone, A.; Gambuzzi, E.; Menziani, M. C. *J. Phys. Chem. C* **2012**, *116*, 14599–14609.

(97) Gambuzzi, E.; Charpentier, T.; Menziani, M. C.; Pedone, A. *Chem. Phys. Lett.* **2014**, *612*, 56–61.

(98) Harris, R. K., Wasylishen, R. E., Duer, M. J., Eds. *NMR Crystallography*; John Wiley & Sons, 2009; pp 1–50.

(99) Pickard, C. J.; Needs, R. J. *Phys. Rev. Lett.* **2006**, *97*, 1–4.

Graphical TOC Entry



We present an approach that allows to theoretically evaluate the NMR parameters of very large systems while keeping an *ab initio* quality. This is achieved by using a high-dimensional neural-network representation of NMR parameters calculated using an *ab initio* formalism on a limited number of small representative systems. We demonstrate the efficiency and accuracy of our approach on silica.