

Electronic Supplementary Information for:

First-principles NMR of oxide glasses boosted by machine learning

Thibault Charpentier^{a,1}

a. Université Paris-Saclay, CEA, CNRS, NIMBE, 91191 Gif-sur-Yvette cedex, France.

S1. Molecular Dynamics Simulations

The classical Molecular Dynamics (MD) simulations of the binary SiO₂-Na₂O glass were performed with the DL_POLY_4 v5.0.0 package.¹ Na-O, Si-O and O-O two-body interactions were parametrized in the Buckingham form for the short range repulsive interactions and partial charges were set to q_O -1.2, q_{Na} +0.6 and q_{Si} +2.4 for the long range electrostatic Coulombic interactions computed with the Ewald summation method with a cutoff radius of 8Å for MD models of 300 atoms (i.e. at half of the simulation unit cell) otherwise 10Å for larger models (600 atoms and more):

$$V_{ij}(r) = -\frac{q_i q_j}{r} + A_{ij} \exp\left\{-\frac{r}{\rho_{ij}}\right\} - \frac{C_{ij}}{r^6} \quad (\text{S1})$$

The parameters were taken from the work of Du et al.² Structural models were prepared using the standard melt-quench procedure starting from a random configuration at experimental density, then melted and equilibrated at 3500K during 100ps in the NVT ensemble. The temperature was decreased to 300K by steps of 100K followed by NVT (30 ps) and NVE (10 ps) simulations at each temperature. A timestep of 0.5fs was used for all simulations. The whole protocol (i.e. starting from a new random configuration) was repeated for each new structural model and for each composition (from 10% to 50% mol. Na₂O)

DFT optimization of the structures and ab-initio MD simulations were performed with the CP2K package.³ using mixed Gaussian and plane-waves basis sets under periodic boundary conditions (PBC). The generalized gradient approximation (GGA) PBE⁴ + DFTD3 functional (where DFTD3 is a correction for accounting Van der Waals dispersion effects)⁵ were used. All element were described with a TZVP Gaussian basis from the built-in MOLOPT basis sets and GTH pseudopotential sets, electronic density and plane-waves were computed with a cutoff energy of 960 Ry with a relative cutoff of 40 Ry (because of the hard description of the sodium atoms including 2s, 2p and 3s orbital in the valence states). aiMD simulations were performed using a timestep of 0.5fs in NVE ensemble, after a thermalization of 4 ps in the NVT ensemble using the CSV thermostat⁶ with a short constant-time (500 fs). Machine Learning Potential (MLP) structures of NS22.5 and NS43.1 glass compositions (22.5 and 43.1 %mol. Na₂O, extracted from Bertani et al. Ref⁷) was used as the initial configuration of aiMD simulations.

1 Corresponding author, thibault.charpentier@cea.fr

S2. Simulation boxes and the NMR database

The classical MD model (300 atoms) compositions and dimensions are given in 1. For each composition, 20 independent simulations were performed and the final configuration extracted from the trajectories at 300K, 1000K, 1500K and 2000K was chosen to form the MD-300K, MD-1000K, MD-1500K and MD-2000K datasets, respectively. Models of 600 atoms were also simulated for the same compositions to form the larger model testing datasets MD2-300K and MD2-1000K. In addition, very large models of 14400 atoms (one for each composition) were generated for assessing the performance of ML-NMR simulation methodology; they are displayed in Figure S1.

Two additional compositions with published NMR data⁸ (^{17}O , ^{29}Si and ^{23}Na), NS22.5 and NS43.1, were also considered. For each composition, three structures were generated with a Machine-Learning Potential (MLP), as described by Bertani et al.⁷ (MLP models, see 2). Structures from an ab-initio Molecular Dynamics (aiMD) trajectory at 300K in the NVE ensemble of 12 ps starting from one MLP model (shown in Figure S2) were extracted every 50fs during the first 2ps (dataset aiMD-300K-a used for training), and every 500fs after 2ps (dataset aiMD-300K-b used for testing).

An overview of the various datasets is given in Table S3. As explained in the main text and in section S3, during the DFT-GIPAW calculations of the NMR interactions with a single point in the reciprocal space, some structural model showed a partial occupation at the Fermi level causing the NMR parameters of some atoms to be erroneous. These models were removed from the database. Note that part of the models were already used in previous works.^{7,9}

Table S1. Classical MD models of the NMR database used in the MD-cellopt, MD-300K, MD-1000K, MD-1500K and MD-2000K datasets.

Composition %mol Na_2O	Unit Cell for 300 atoms	Box size(Å)	Density MD	Density ^(a) cellopt-DFT
10Na	$\text{Si}_{90}\text{Na}_{20}\text{O}_{190}$	16.35	2.289	2.418 (0.020)
20Na	$\text{Si}_{80}\text{Na}_{40}\text{O}_{180}$	16.15	2.383	2.525 (0.020)
30Na	$\text{Si}_{70}\text{Na}_{60}\text{O}_{170}$	15.98	2.466	2.570 (0.019)
40Na	$\text{Si}_{60}\text{Na}_{80}\text{O}_{160}$	15.86	2.532	2.608 (0.011)
50Na	$\text{Si}_{50}\text{Na}_{100}\text{O}_{150}$	15.82	2.560	2.647 (0.013)

(a) Mean values and standard deviations values in parentheses.

Figure S1: large MD structural models of 14400 atoms (with compositions and density values given in Table S1) used for testing the application of the ML-NMR simulation algorithm. Silicon tetrahedra are in blue, sodium atoms are in yellow. The figure was generated with VESTA.¹⁰

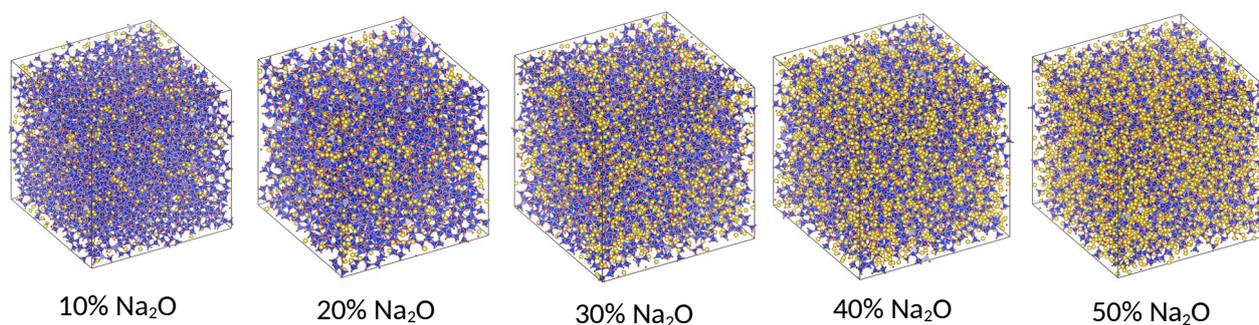


Table S2. MLP-MD models of the NMR database used in the MLP (3 structures per composition) and the aiMD-300K datasets (structures extracted from an aiMD NVE trajectory at 300K).

Composition %mol. Na ₂ O	Unit Cell	Box size (Å) ^(a)	Density ^(a)
22.5	Si ₁₈₆ Na ₁₀₈ O ₄₂₆	22.08 (0.01)	2.240 (0.011)
43.1	Si ₁₃₇ Na ₃₇₈ O ₂₀₈	21.73 (0.14)	2.375 (0.015)

(a) Mean values and standard deviations values in parentheses.

Figure S2 the MLP-MD structural models (with compositions and density values given in Table S2) used for testing the application of the ML-NMR simulation algorithm. Silicon tetrahedra are in blue, sodium atoms are in yellow. The figure was generated with VESTA.³

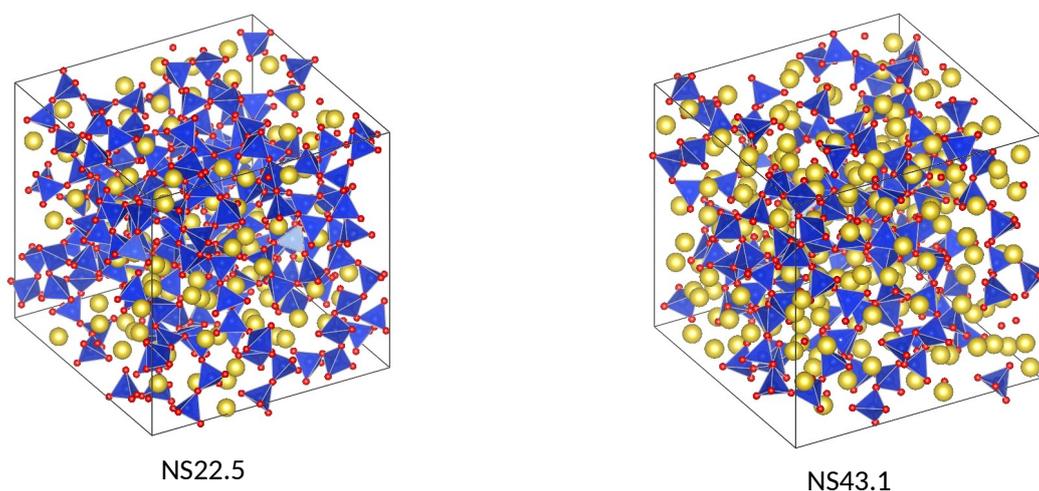


Table S3. Number of structures and atoms in the NMR database.

Learning set	Number of structures	Total number of atoms		
		Si	O	Na
MD-CELLOPT	50	3500	8500	3000
MD-300K	99	6940	16840	5920
MD-1000K	100	7000	17000	6000
MD-1500K	97	5820	16940	5820
MD-2000K	96	6720	16320	5760
MLP	6	969	2412	948
aiMD-300K-a ^(a)	82	13243	32964	12956
aiMD-300K-b ^(b)	18	2907	7236	2844

(a) Structures sampled every $\Delta=50$ fs from 0 ps to 2 ps of the NVE aiMD trajectory at 300K

(b) Structures sampled every $\Delta=500$ fs from 2 ps to 4 ps of the NVE aiMD trajectory at 300K

S.3 DFT-GIPAW calculations with VASP

The NMR properties were computed with the VASP code (version 5.x) ¹¹ using the GIPAW method¹²⁻¹⁵ with the GGA-PBE functional,⁴ a single k-point in the reciprocal space and a kinetic energy cutoff of 550 Ry for the plane-wave expansion. The PAW pseudopotentials of the built-in library (potcar, version 5.4) were chosen for Si, O and Na with the valence state $3s^23p^2$, $2s^22p^4$ and $2s^22p^63s^1$, respectively. For calibrating the computed isotropic magnetic shielding value σ_{iso} into the experimental isotropic chemical shift δ_{iso} , reference samples from Ref.¹⁶ were used: cristobalite SiO_2 , quartz SiO_2 , sodium metasilicate Na_2SiO_3 and sodium disilicate α - and β - $\text{Na}_2\text{Si}_2\text{O}_5$. But for consistency with the glass procedure, the unit cell of each system was duplicated to form a supercell of about 300-400 atoms. Each supercell was subsequently optimized with CP2K using the same parameters as in the DFT optimization of the MD models (optimization of the atomic positions and lattice cell vector lengths but conserving the angles to maintain the space group). The calibration linear regression was chosen to be in the analytical form:

$$\delta_{iso} = -\alpha(\sigma_{iso} - \sigma_{REF}) \quad (\text{S2})$$

with the two parameters α and σ_{REF} to be fitted. The values for ^{29}Si , ^{17}O and ^{23}Na are given in Table S4. Because of the slight dispersion of the NMR parameters in the supercell, the mean value for each crystallographic sites was considered for Eq. S2. Note that systems with a large experimental uncertainty (i.e., with contradictory reported values) were not taken into account (see Ref.¹⁶ for experimental values). Concerning the quadrupolar coupling constant C_Q and biaxiality (or asymmetry parameter) η , they were

computed from the computed from Electric Field Gradient (EFG) tensor \mathbf{V} , after its diagonalization, yielding the principal values (V_{ZZ} , V_{YY} , V_{XX}):

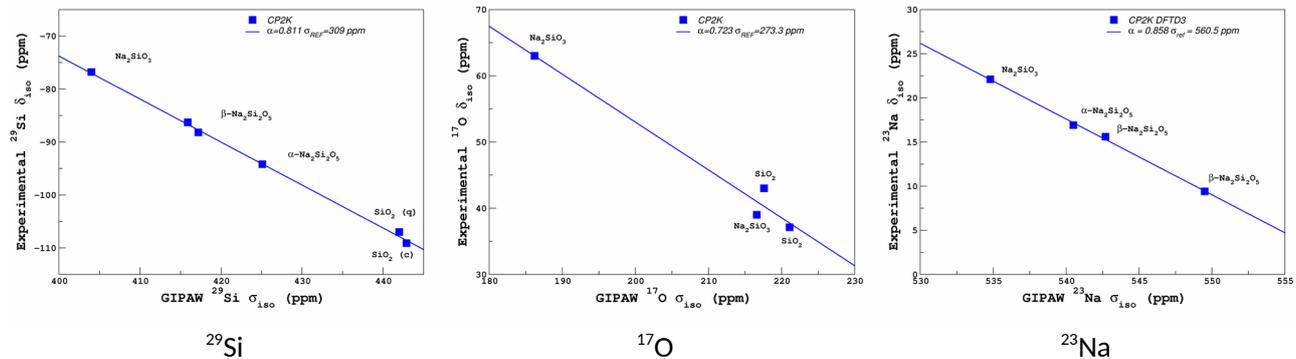
$$C_Q = \frac{eQ}{h} V_{ZZ}, \eta = \frac{V_{YY} - V_{XX}}{V_{ZZ}} \quad \text{with} \quad |V_{XX}| < |V_{YY}| < |V_{ZZ}| \quad (\text{S3})$$

where Q is the nuclear quadrupolar moment (-25.58 mb for ^{17}O and 104 mb for ^{23}Na).

Table S4. Parameters for the linear regression of δ_{iso} against σ_{iso} (Eq. S2)

Atom	Si	O	Na
α	0.811	0.723	0.858
σ_{REF} (ppm)	309	273.3	560.5

Figure S3. Calibration linear regression of ^{29}Si , ^{17}O and ^{23}Na GIPAW-DFT isotropic magnetic shielding values.



S.4 Computation of SOAP descriptors

The SOAP descriptors were computed using an in-house code (written in C++11). First, the structure is analysed to determine the list of neighbours of each of atom within the cutoff radius r_{cut} so that a list of all pairs of atoms can be established. For each pair of neighbouring atoms ij linked by the vector \mathbf{r}_{ij} , the partial SOAP descriptors $c_{nlm}^{ij} = c_{nlm}(\mathbf{r}_{ij})$ are computed:

$$c_{nlm}(\mathbf{r}_{ij}) = h_{nl}(r_{ij}) \times Y_{lm}(\theta_{ij}, \phi_{ij}) \quad (\text{S4})$$

where r_{ij} is the inter-atomic distance and (θ_{ij}, ϕ_{ij}) are the polar and azimuthal angles (also denoted as \hat{r}_{ij} in literature). Note that throughout this work we are working with real spherical harmonics (RSH), so that no complex conjugate quantities appears. The functions $h_{nl}(r)$ are resulting from the Gaussian smoothing in space:

$$h_{nl}(r) = \frac{1}{(2\pi\sigma^2)^{3/2}} \times f_c(r) \times \int_0^{r_{\text{cut}}} 4\pi u^2 du R_{nl}(u) i_l\left(\frac{ur}{\sigma}\right) \exp\left\{-\frac{u^2+r^2}{2\sigma^2}\right\} \quad (\text{S5})$$

where $f_c(r) = 1/2 \{ \cos(\pi r/r_{cut}) + 1 \}$ is the smoothing function at the cutoff radius, σ is the Gaussian smoothing parameter (also denoted as σ_{SOAP} in the main text for sake of clarity), $R_{nl}(u)$ are the radial basis functions and $i_l(x)$ denotes the modified spherical Bessel function of the first kind. For sake of efficiency, the $h_{nl}(r)$ are pre-computed on a r -grid on $[0, r_{cut}]$ and interpolated with cubic-spline for any requested r value. Alternatively, taking benefit of the choice of the spherical Bessel function as the radial functions:

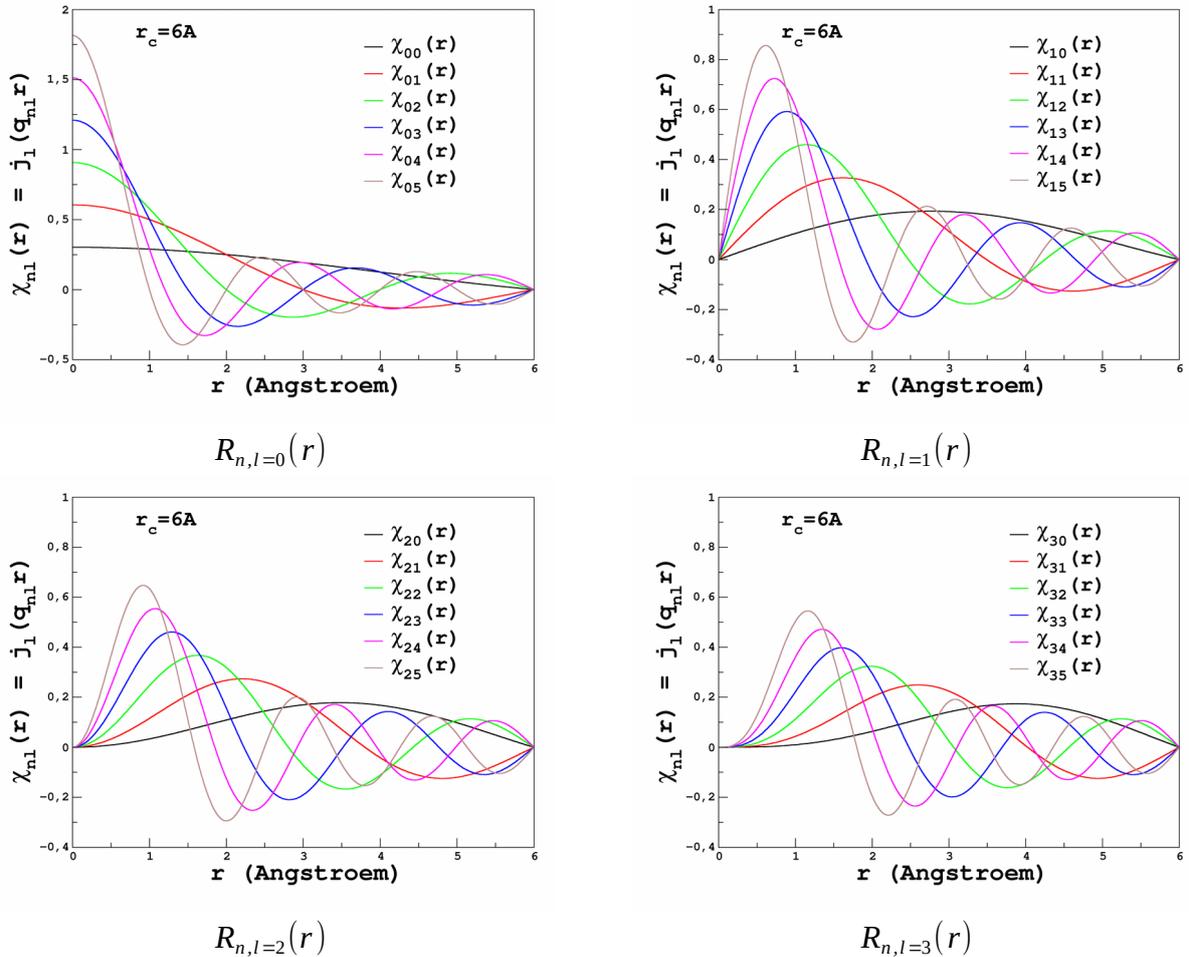
$$R_{nl}(r) = j_l(\alpha_{nl} r/r_{cut}) \quad (S6)$$

where α_{nl} is the n^{th} root of the Bessel function so that $R_{nl}(r_{cut}) = 0$ for any n and l values ($n \leq l$) as shown in Figure S4. Another advantage of using spherical Bessel functions (Eq. S6) is the usage of the fast Discrete Hankel Transform (DHT) for computing the integral Eq. S5. Finally, the SOAP descriptors of an atom i are calculated as

$$c_{nlm}^{\mu} = \sum_{j \in N_i^{\mu}} c_{nlm}(r_{ij}) \quad (S7)$$

where N_i^{μ} is the set of neighbouring atoms of kind μ of the central atom i .

Figure S4. The spherical Bessel function used as radial functions (Eq. S6) for computing the SOAP descriptors (Eq. S4 and Eq. S5).



Typical CPU times for computing the SOAP descriptors on a single Intel Core i7 processor are shown in Figure S5 for various system sizes. We observe an excellent linearity with the number of atoms but with an increase in the slope when augmenting the cutoff-radius value. For the largest system studied, all descriptors could be computed in less than 2s (without any parallelization). The efficiency of the procedure implemented in our code allows calculating the descriptors on a MD trajectory (typically 100 000 structures on a 1 ns trajectory), as discussed and illustrated in the main text. Parallelization is planned for future work.

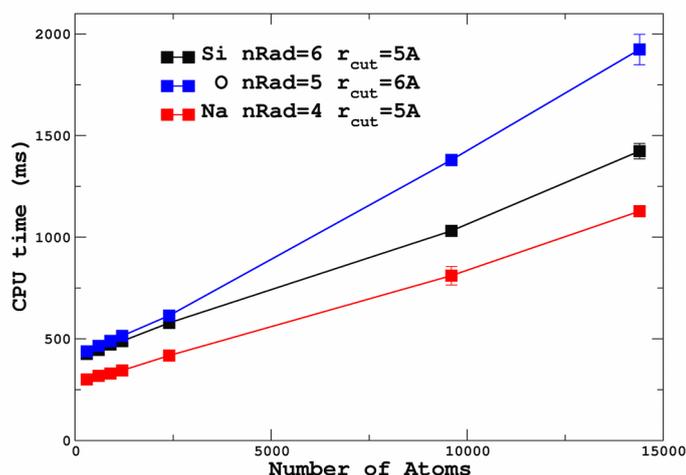


Figure S5. CPU time for the calculations of the SOAP descriptors (single Intel CORE i7) with respect to the number of atoms using $L_{\max}=4$, and $\sigma_{\text{SOAP}}=0.4 \text{ \AA}$ in Eq. S5.

S.5 Optimization of SOAP descriptors by LSSVR

For Si, O and Na atom, the descriptors are computed and the prediction of the isotropic magnetic shielding σ_{iso} with the least-square support vector regression (LSSVR, see text for description) was used to determine the optimal values of the following (hyper)parameters (maximum angular momentum was fixed to $L_{\max}=4$): the maximum number of radial function N_{MAX} (also denoted nRad), the Gaussian smoothing σ_{SOAP} and cutoff radius r_{cut} . We show here as an illustration the results obtained with the MD-300K and MD-1000K dataset using the optimized value of $\sigma_{\text{SOAP}} = 0.4 \text{ \AA}$. For each set of values of the hyperparameters, the convergence of σ_{iso} mean absolute error (MAE) with respect to the Nyström size (N_{ξ} , see text) was calculated. Results are shown below for each atom for cutoff radius values ranging from 3 Å to 6 Å.

The convergence curves for the MD-300K dataset are shown in Figure S6 (^{29}Si), Figure S7 (^{17}O) and Figure S8 (^{23}Na); for the MD-1000K dataset in Figure S9 (^{29}Si), Figure S10 (^{17}O) and Figure S11 (^{23}Na).

The agreement between the LSSVR predictions with the DFT-GIPAW values are shown for the MD-300K, MD-1000K, MD-1500K and MD-2000K datasets in Figure S12, Figure S13 and Figure S14.

Transferability tests between datasets are given in Tables S5, S6 and S7.

Figure S6. Convergence of ^{29}Si σ_{iso} MAE (in ppm) with the Nyström size of the LSSVR algorithm with the MD-300K dataset.

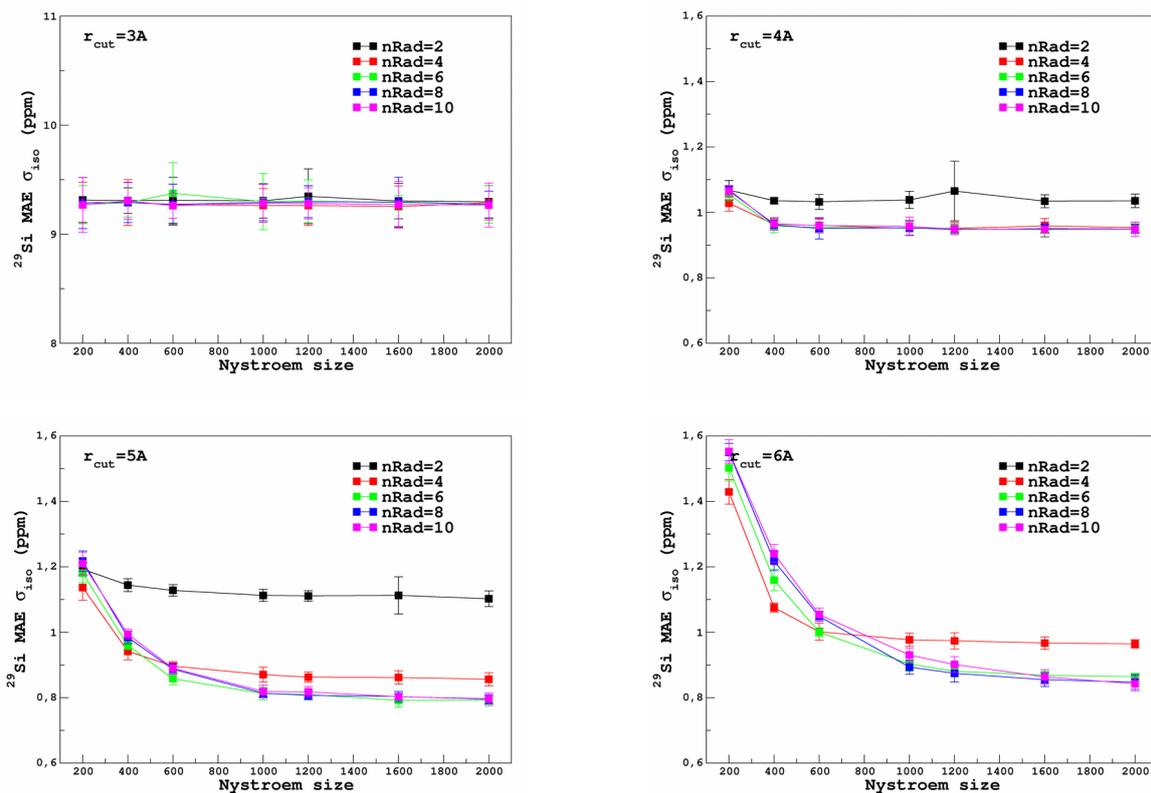


Figure S7. Convergence of ^{17}O σ_{iso} MAE (in ppm) with to the Nyström size of the LSSVR algorithm with the MD-300K dataset.

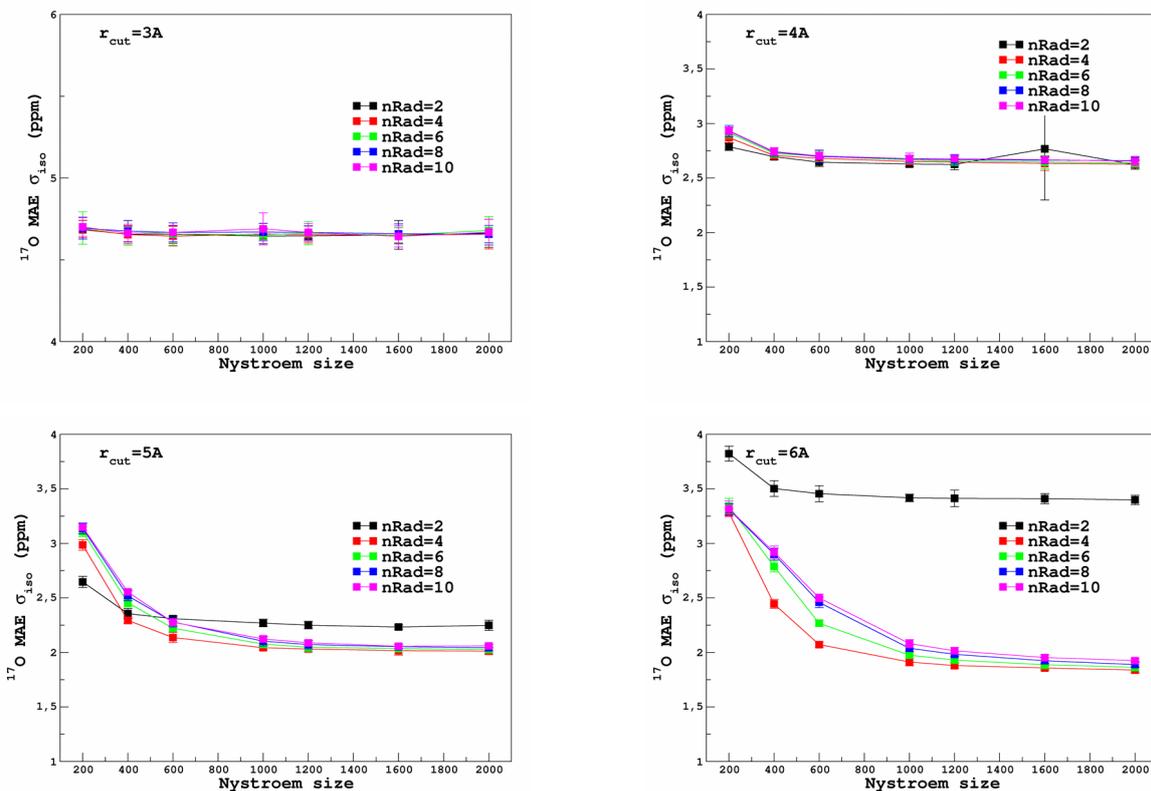


Figure S8. Convergence of ^{29}Si σ_{iso} MAE (in ppm) with the Nyström size of the LSSVR algorithm with the MD-300K dataset.

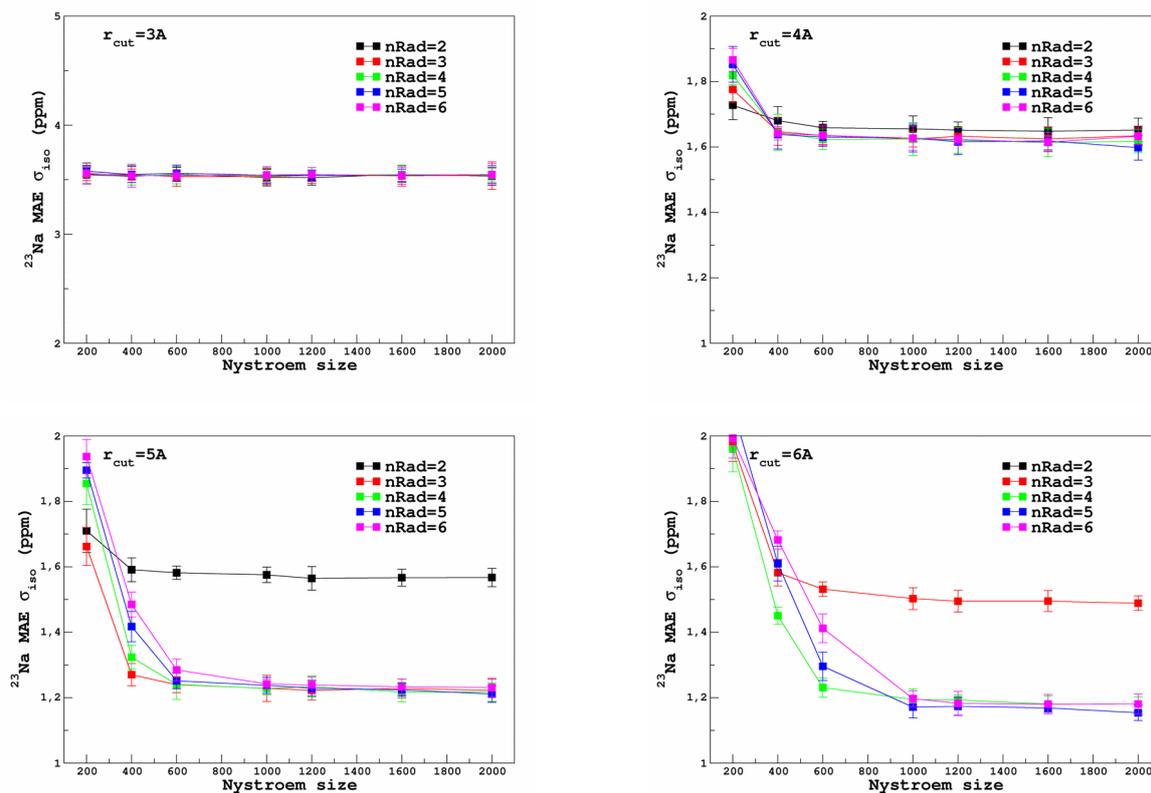


Figure S9. Convergence of the ^{29}Si σ_{iso} MAE (in ppm) with the Nyström size of the LSSVR algorithm with the MD-1000K dataset.

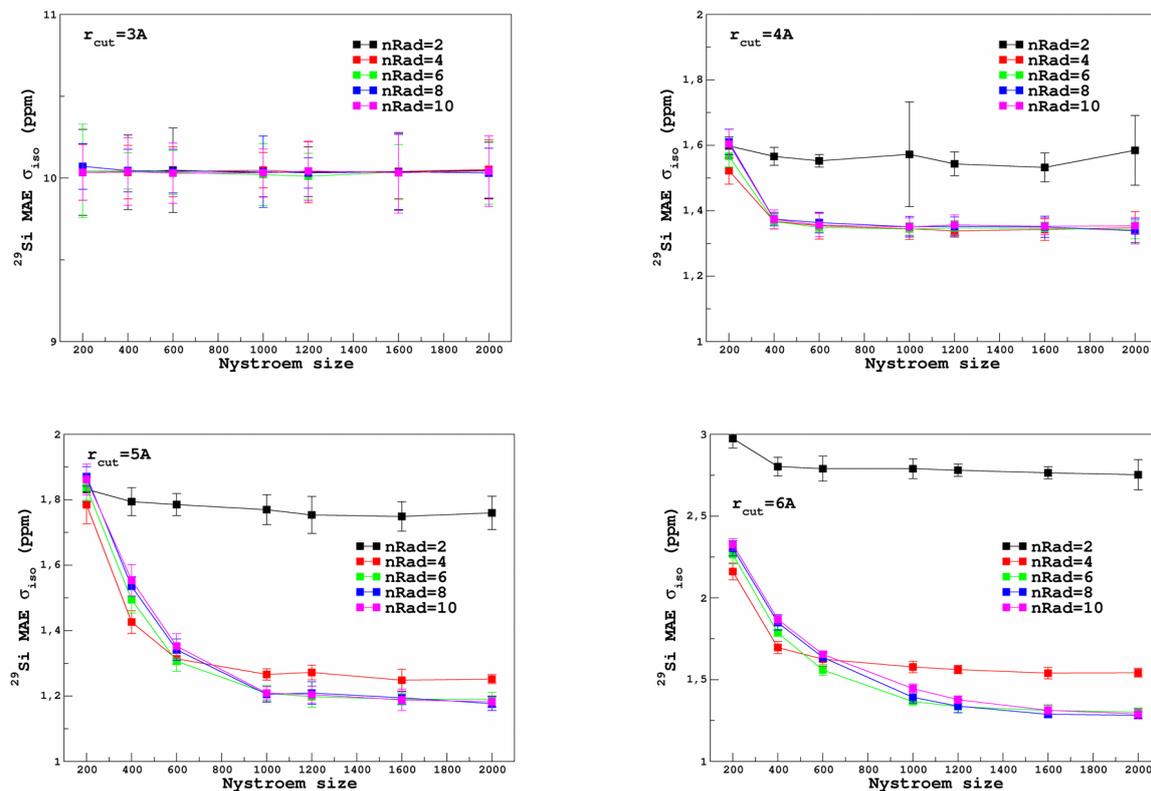


Figure S10. Convergence of the ^{17}O σ_{iso} MAE (in ppm) with the Nyström size of the LSSVR algorithm with the MD-1000K dataset.

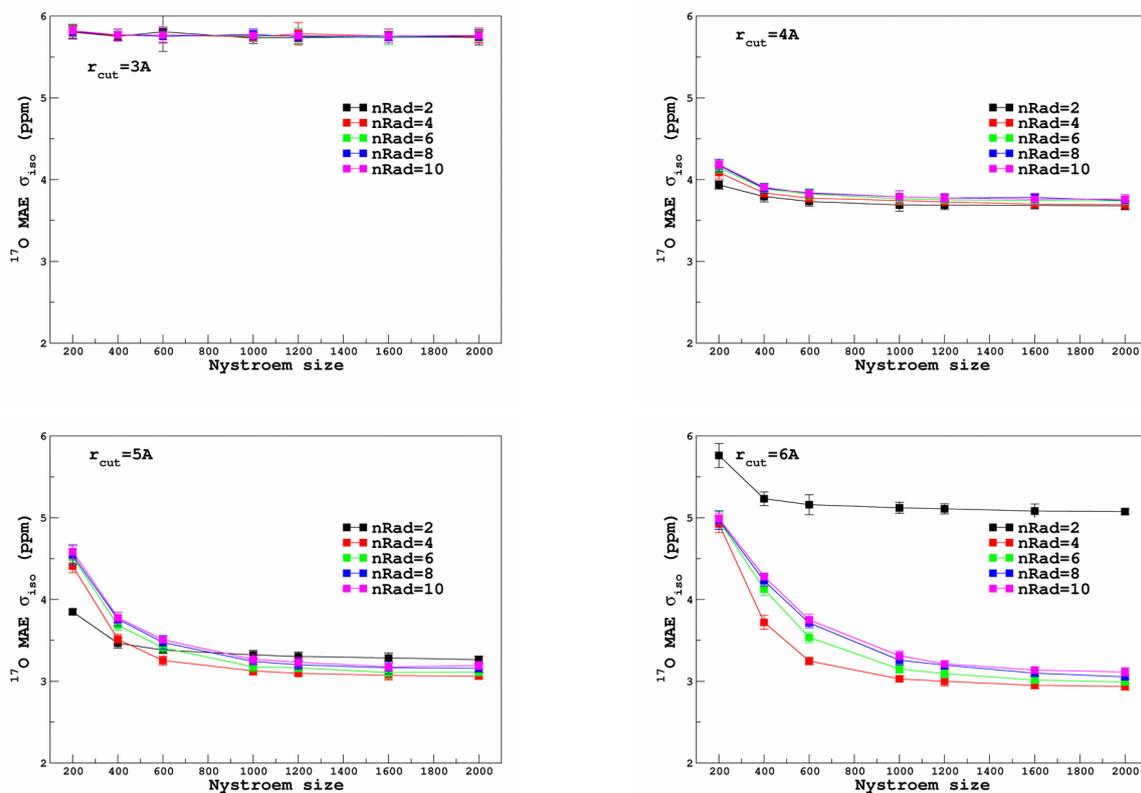


Figure S11. Convergence of the ^{23}Na σ_{iso} MAE (in ppm) with the Nyström size of the LSSVR algorithm with the MD-1000K dataset.

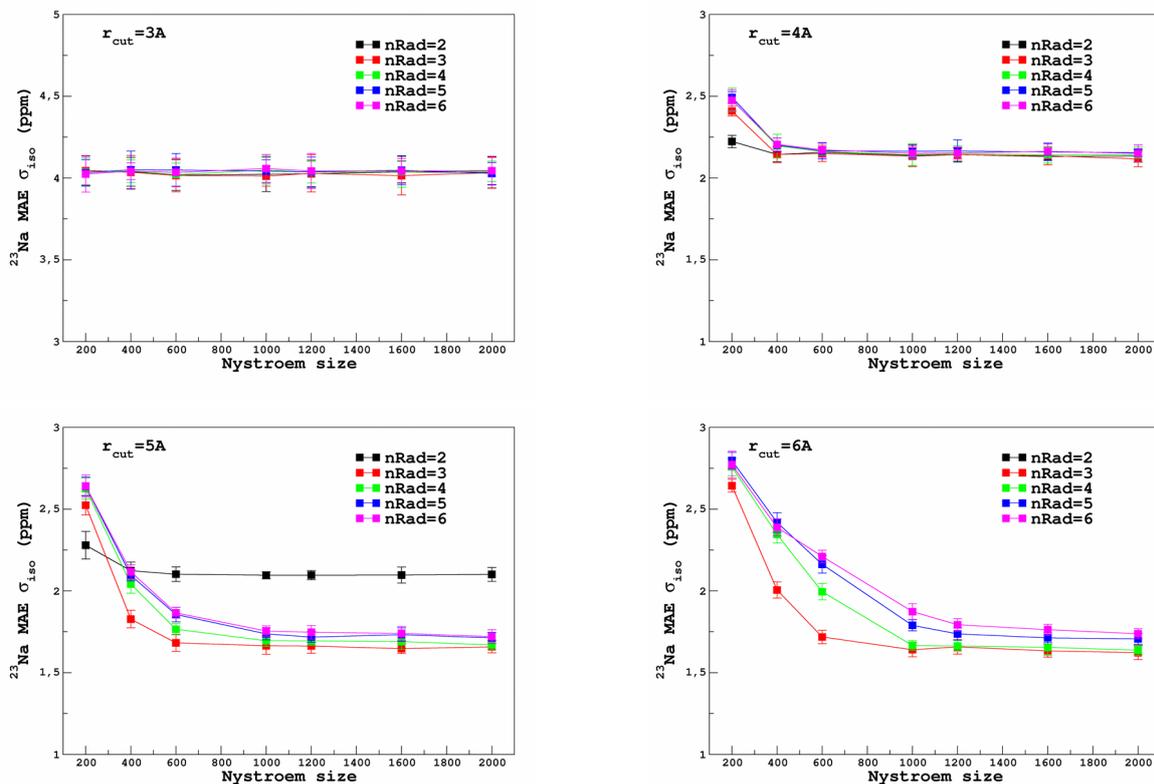
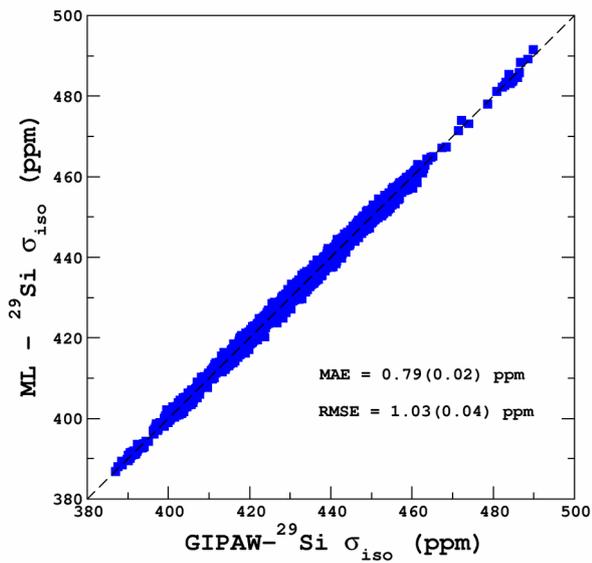
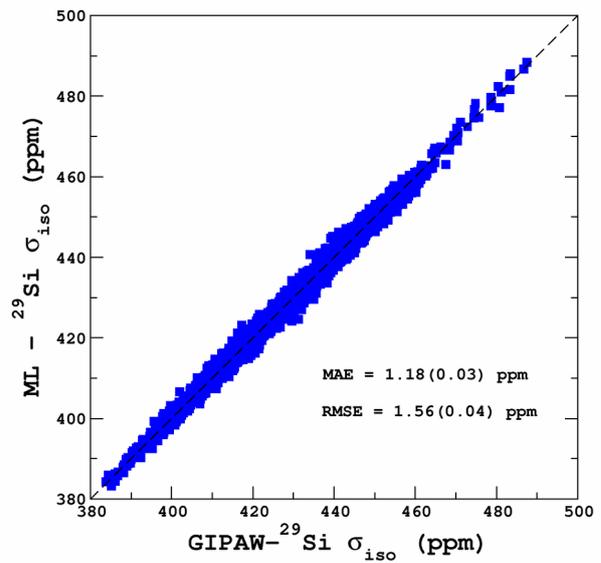


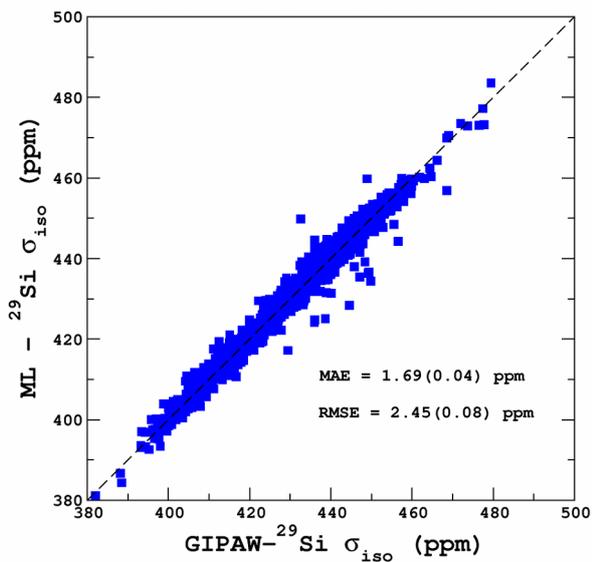
Figure S12. LSSVR versus DFT-GIPAW ^{29}Si σ_{iso} values (Nyström size 2000).



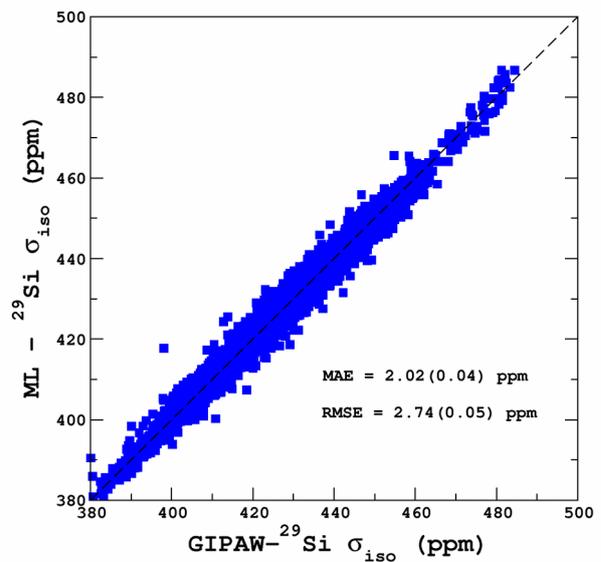
MD-300K



MD-1000K

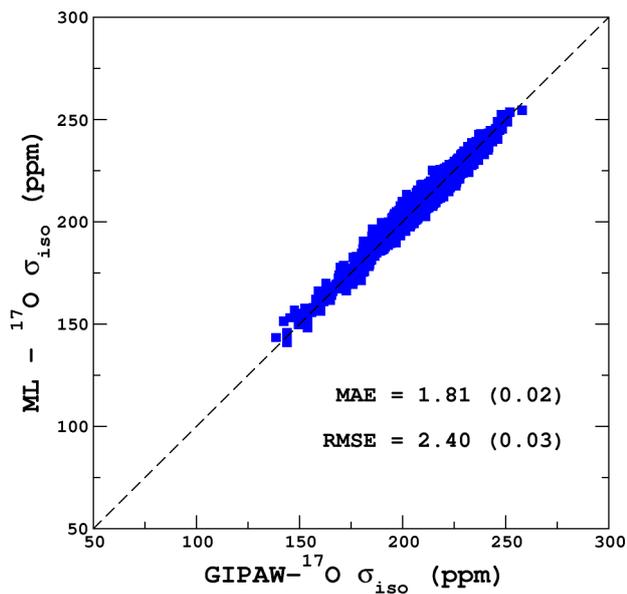


MD-1500K

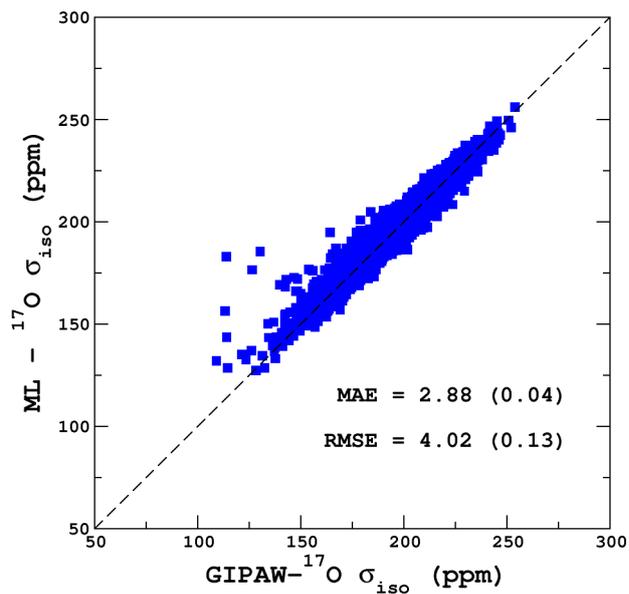


MD-2000K

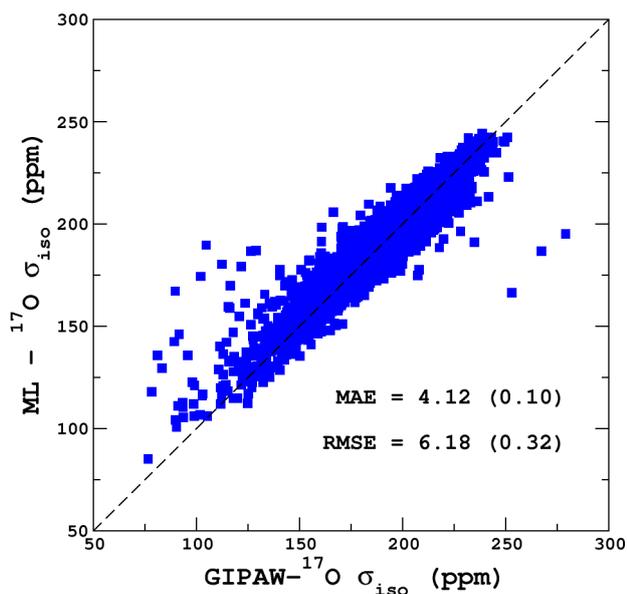
Figure S13. LSSVR versus DFT-GIPAW ^{17}O σ_{iso} (Nyström size 4000).



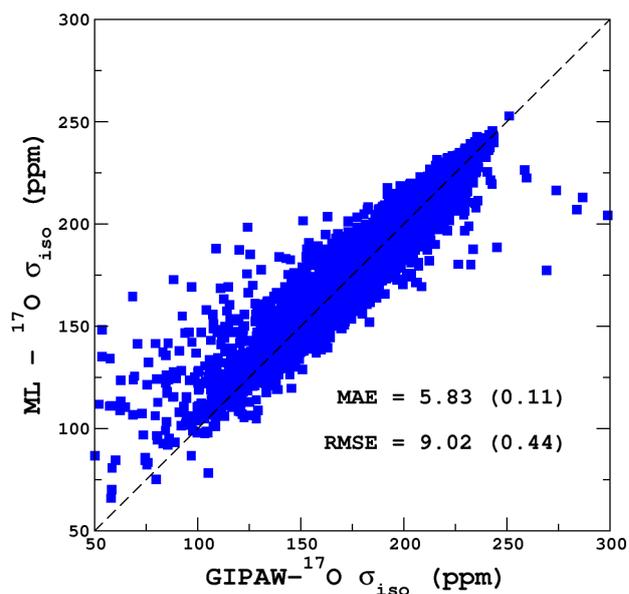
MD-300K



MD-1000K

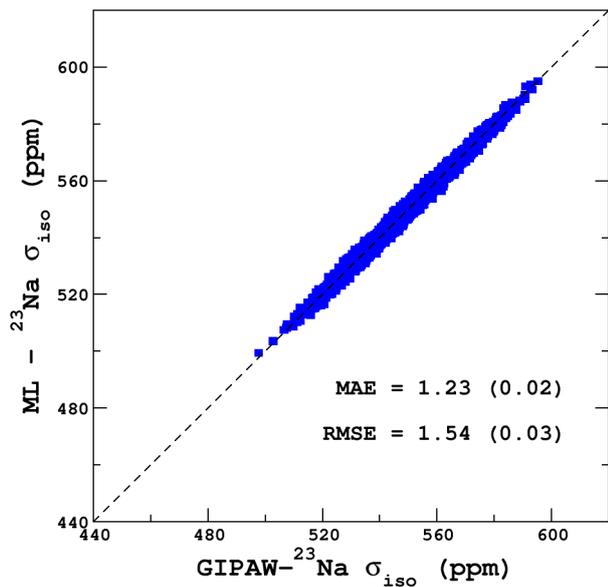


MD-1500K

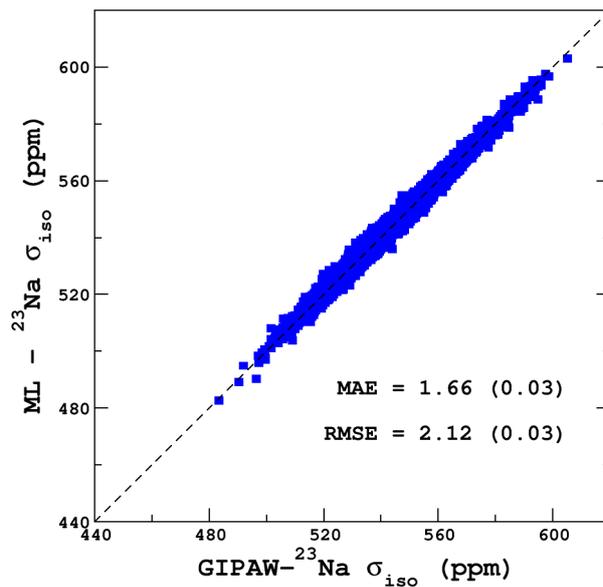


MD-2000K

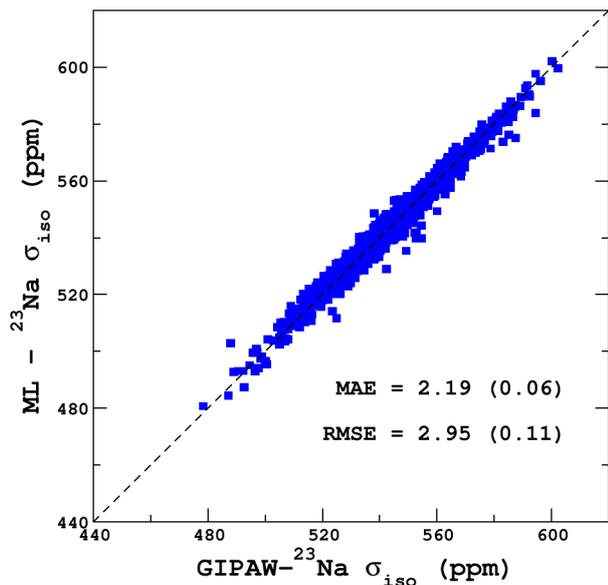
Figure S14. LSSVR versus DFT-GIPAW ^{23}Na σ_{iso} (Nyström size 1000).



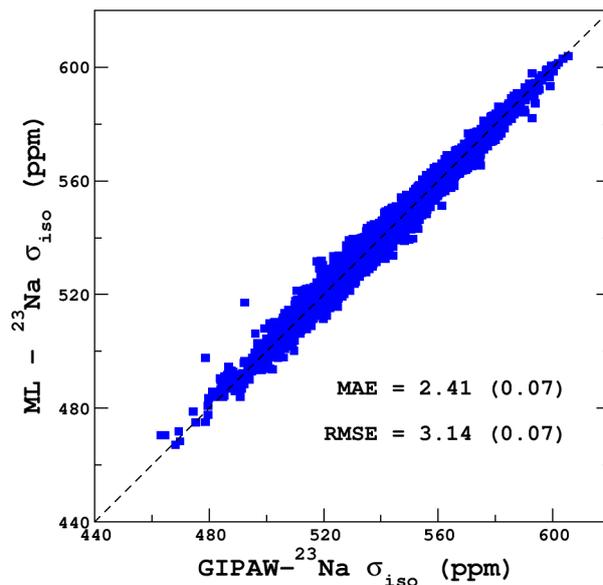
MD-300K



MD-1000K



MD-1500K



MD-2000K

Table S5. Prediction errors of the ^{29}Si isotropic magnetic shielding between different datasets.

Testing set	MD-300K		MD-1000K		MD-1500K		MD-2000K	
Training set	MAE (ppm)	RMSE (ppm)	MAE (ppm)	RMSE (ppm)	MAE (ppm)	RMSE (ppm)	MAE (ppm)	RMSE (ppm)
MD-300K	0.7	0.7	1.2	2.6	1.8	14.9	2.3	9.7
MD-1000K	0.8	1.1	1.0	1.6	1.6	13.6	1.9	7.2
MD-1500K	1.0	1.5	1.2	2.5	1.5	12.3	2.0	7.1
MD-2000K	1.0	1.8	1.3	2.8	1.7	14.0	1.7	5.5

Table S6. Prediction errors of the ^{23}Na isotropic magnetic shielding values between different datasets.

Testing set	MD-300K		MD-1000K		MD-1500K		MD-2000K	
Training set	MAE (ppm)	RMSE (ppm)	MAE (ppm)	RMSE (ppm)	MAE (ppm)	RMSE (ppm)	MAE (ppm)	RMSE (ppm)
MD-300K	1.0	1.7	1.7	4.7	2.3	10.5	2.9	17.6
MD-1000K	1.3	2.6	1.4	3.2	2.1	8.1	2.4	10.3
MD-1500K	1.4	3.0	1.7	4.5	1.8	6.0	2.4	9.4
MD-2000K	1.5	3.6	1.7	4.8	2.1	8.2	2.1	7.3

Table S7. Prediction errors of the ^{17}O isotropic magnetic shielding values between the different MD training sets

Testing set	MD-300K		MD-1000K		MD-1500K		MD-2000K	
Training set	MAE (ppm)	RMSE (ppm)	MAE (ppm)	RMSE (ppm)	MAE (ppm)	RMSE (ppm)	MAE (ppm)	RMSE (ppm)
MD-300K	1.6	4.5	3.1	22.0	6.3	9583	9.4	1476
MD-1000K	2.2	8.9	2.5	12.1	5.2	9524	7.6	1382
MD-1500K	3.1	17.1	3.4	21.4	4.9	9500	7.1	1330
MD-2000K	4.3	36.8	4.5	38.5	6.0	9485	6.7	1270

S.6 Learning the EFG tensor

Results of application of the LSSVR algorithm to the prediction the quadrupolar parameters C_Q , η and P_Q are shown in Figure S15 for ^{23}Na , and Figure S16 for ^{17}O .

The optimization of the SOAP descriptors (cutoff radius r_{cut} and number of radial basis function $n_{\text{Rad}}=N_{\text{MAX}}$ with $\sigma_{\text{SOAP}}=0.4 \text{ \AA}$) for the prediction of the EFG tensor components with the LRR λ -SOAP algorithm are shown in Figure S17. The agreement between the LSSVR predictions with the DFT-GIPAW values are shown for the MD-300K, MD-1000K datasets in Figure S18.

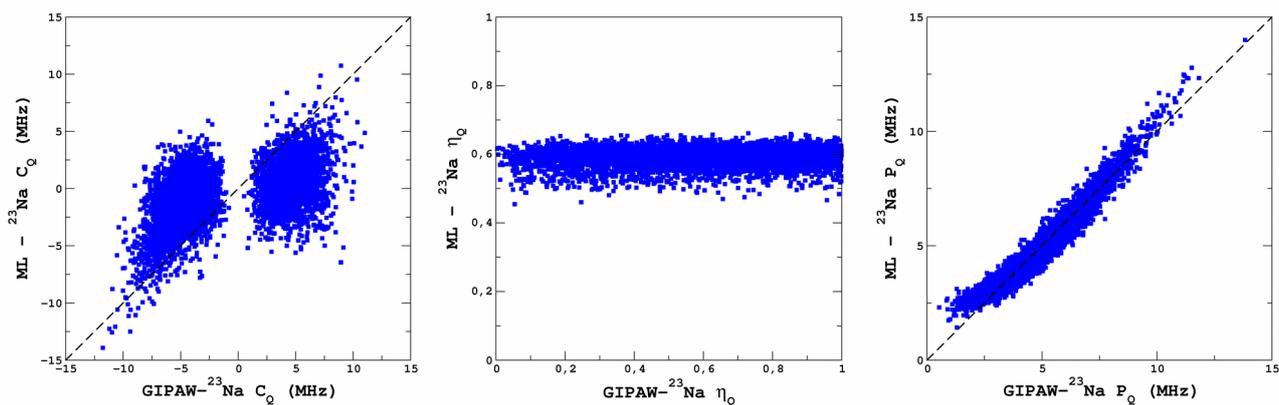


Figure S15. LSSVR versus DFT-GIPAW ^{23}Na quadrupolar parameters for the MD-300K dataset.

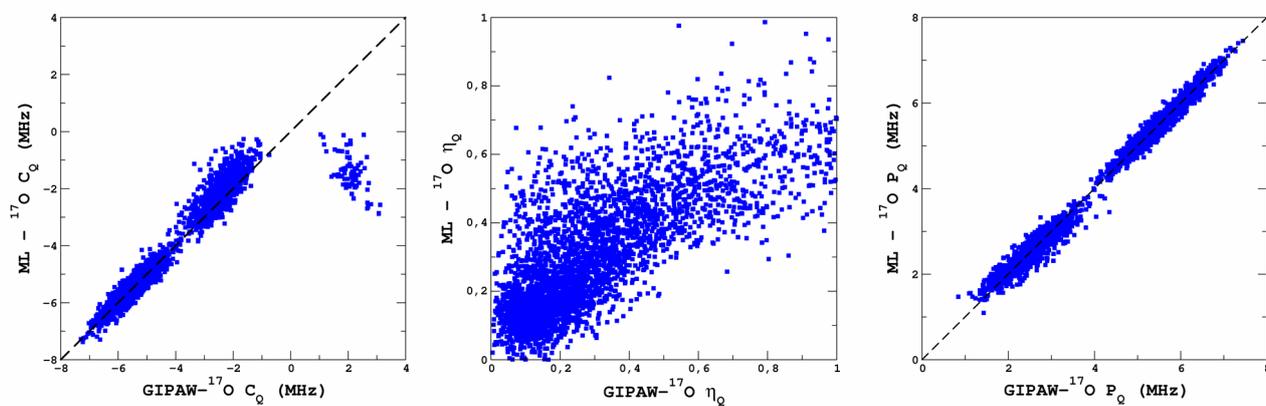
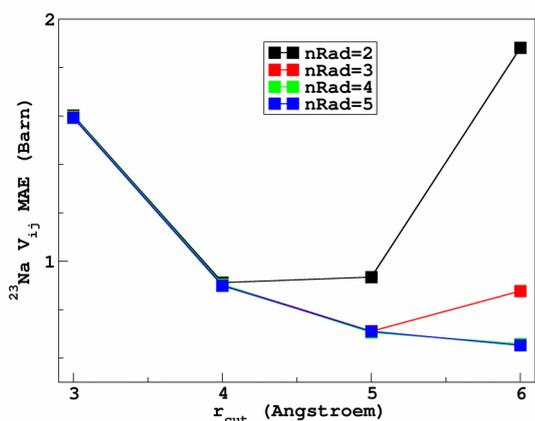
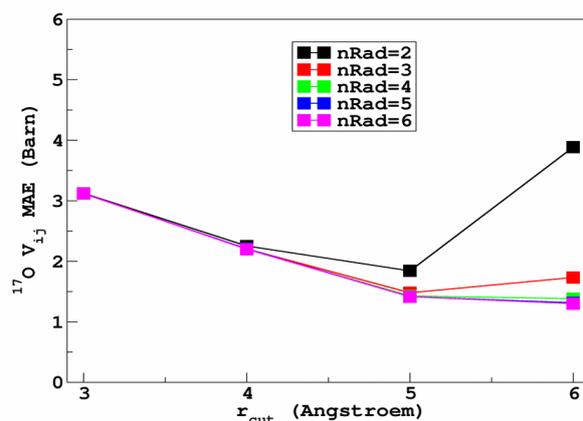


Figure S16. LSSVR versus DFT-GIPAW ^{17}O quadrupolar parameters for the MD-300K dataset.

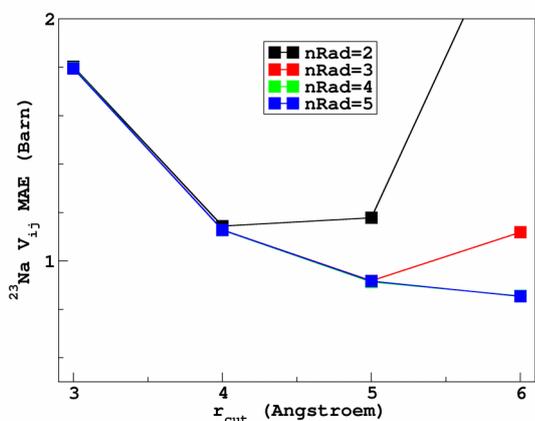
Figure S17. Variation of the mean absolute error (MAE) of the EFG components predicted with LRR λ -SOAP with respect to the cutoff radius of the SOAP descriptor, for different values of the number of radial functions (n_{Rad}) using the MD-300K and MD-1000K datasets.



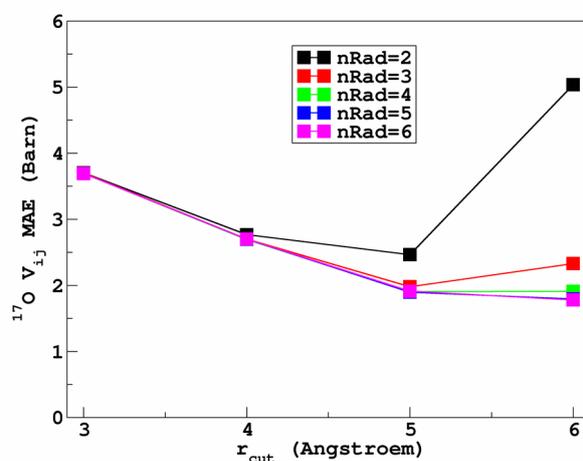
^{23}Na EFG, MD-300K



^{17}O EFG, MD-300K

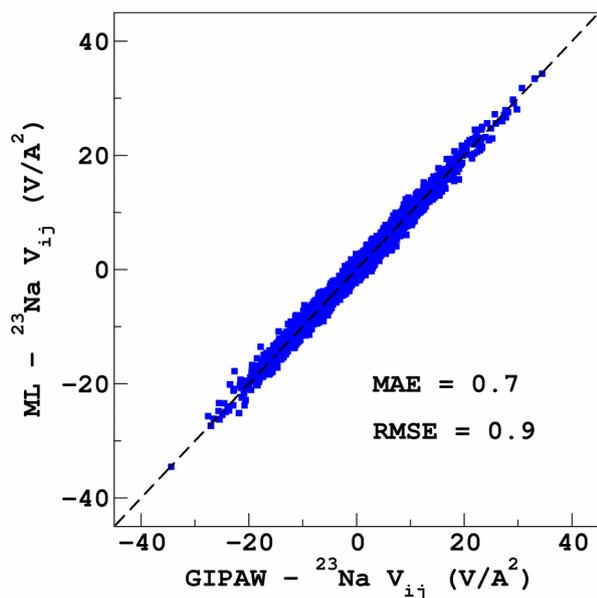


^{23}Na EFG, MD-1000K

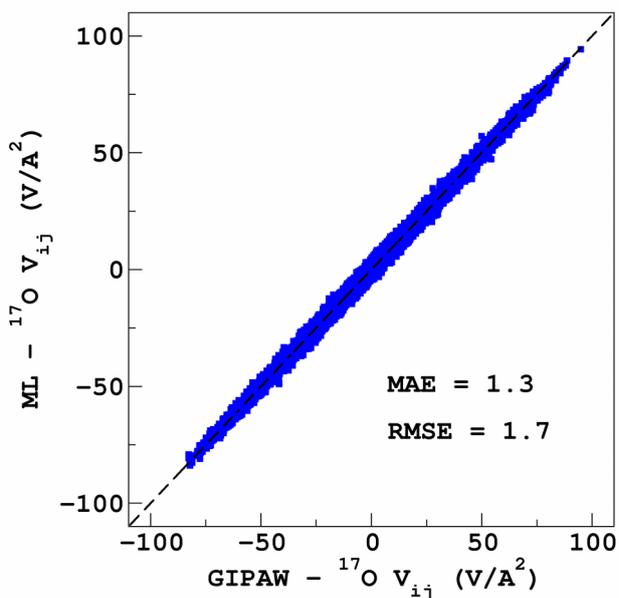


^{17}O EFG, MD-1000K

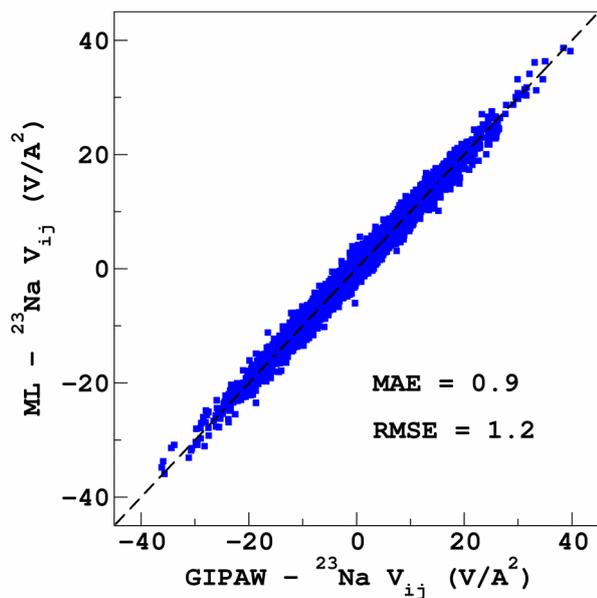
Figure S18. LRR λ -SOAP versus DFT-GIPAW ^{23}Na and ^{17}O EFG components (in spherical form) for MD-300K and MD-1000K datasets.



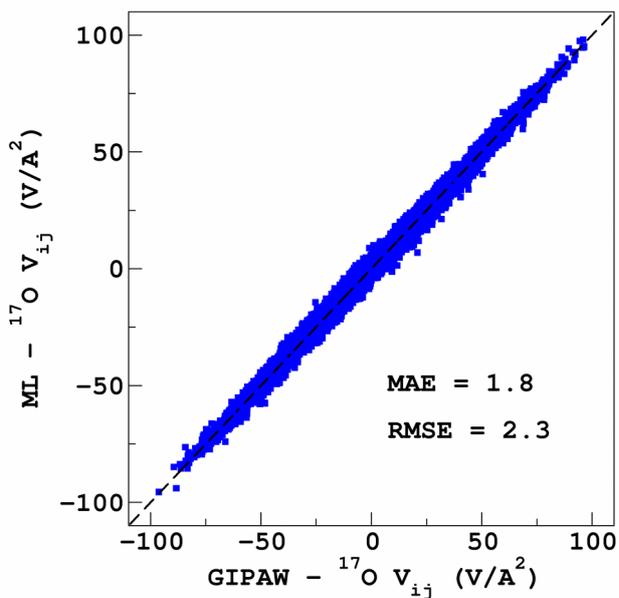
^{23}Na MD-300K



^{17}O MD-300K



^{23}Na MD-1000K



^{17}O MD-1000K

S.7 Transferability of ML predictors

The accuracy of the ML algorithms trained on 300 atoms datasets (with merged MD-300K and MD-1000K datasets) for the prediction of the isotropic magnetic shielding and EFG tensor components are shown in Figure S19 and Figure S20, respectively.

Similarly, accuracy of learning from the initial part of an aiMD trajectory (2 ps) to extrapolate to full MD trajectory (12 ps) are illustrated in Figure S21 (NS22.5) and Figure S22 (NS43.1) for the isotropic magnetic shielding and in and for the EFG tensor.

Figure S19. LSSVR versus DFT-GIPAW ^{29}Si , ^{17}O and ^{23}Na σ_{iso} of the MD2-300K+1000K (600 atoms) datasets. The LSSVR was trained with the MD-300K+1000K datasets (300 atoms).

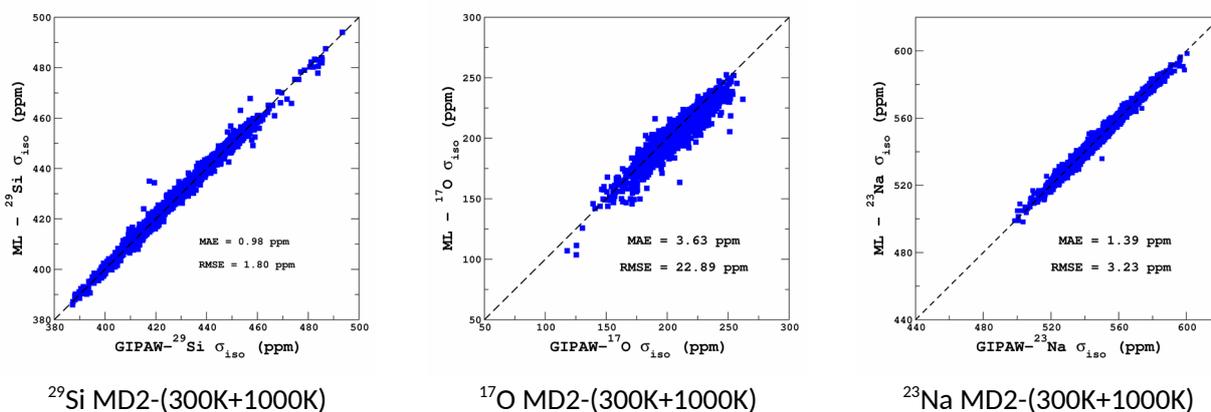


Figure S20. LRR λ -SOAP versus DFT-GIPAW ^{23}Na and ^{17}O EFG components for the MD2-300K+1000K (600 atoms) datasets. The LRR was trained with the MD-300K+1000K datasets (300 atoms).

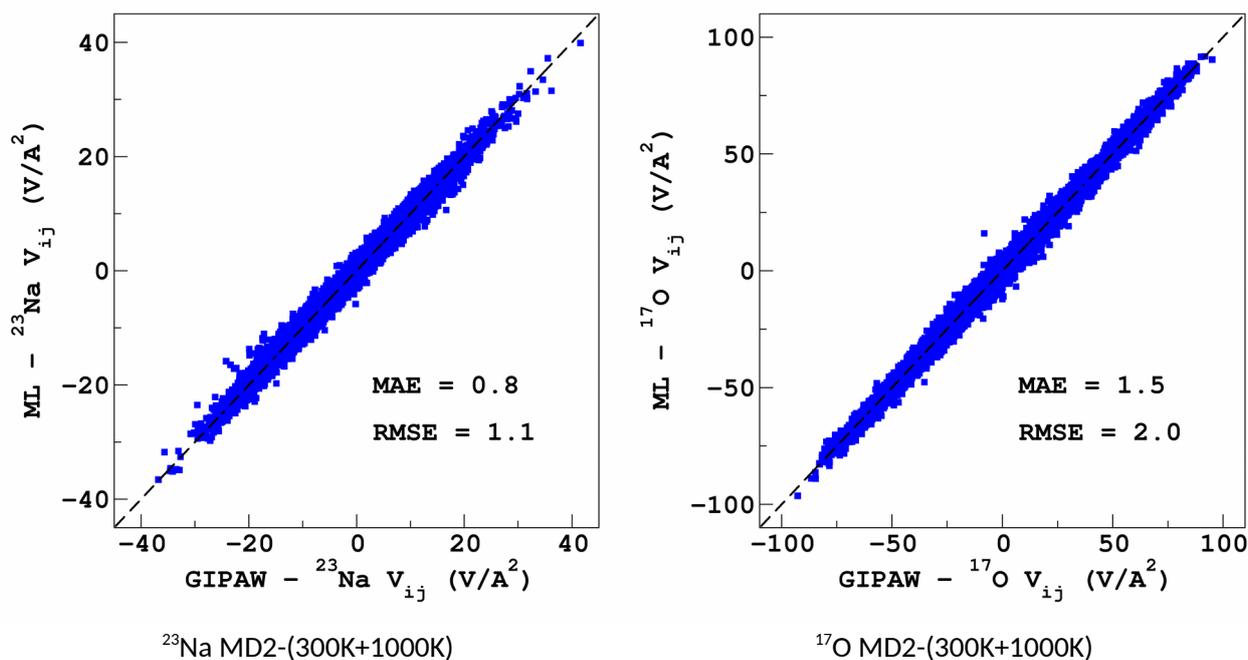


Figure S21. LSSVR versus DFT-GIPAW ^{29}Si , ^{17}O and ^{23}Na σ_{iso} (NS22.5). LSSVR was trained with the aiMD-300K-a dataset and tested on the aiMD-300K-a (Top) and aiMD-300K-b (Bottom) datasets. .

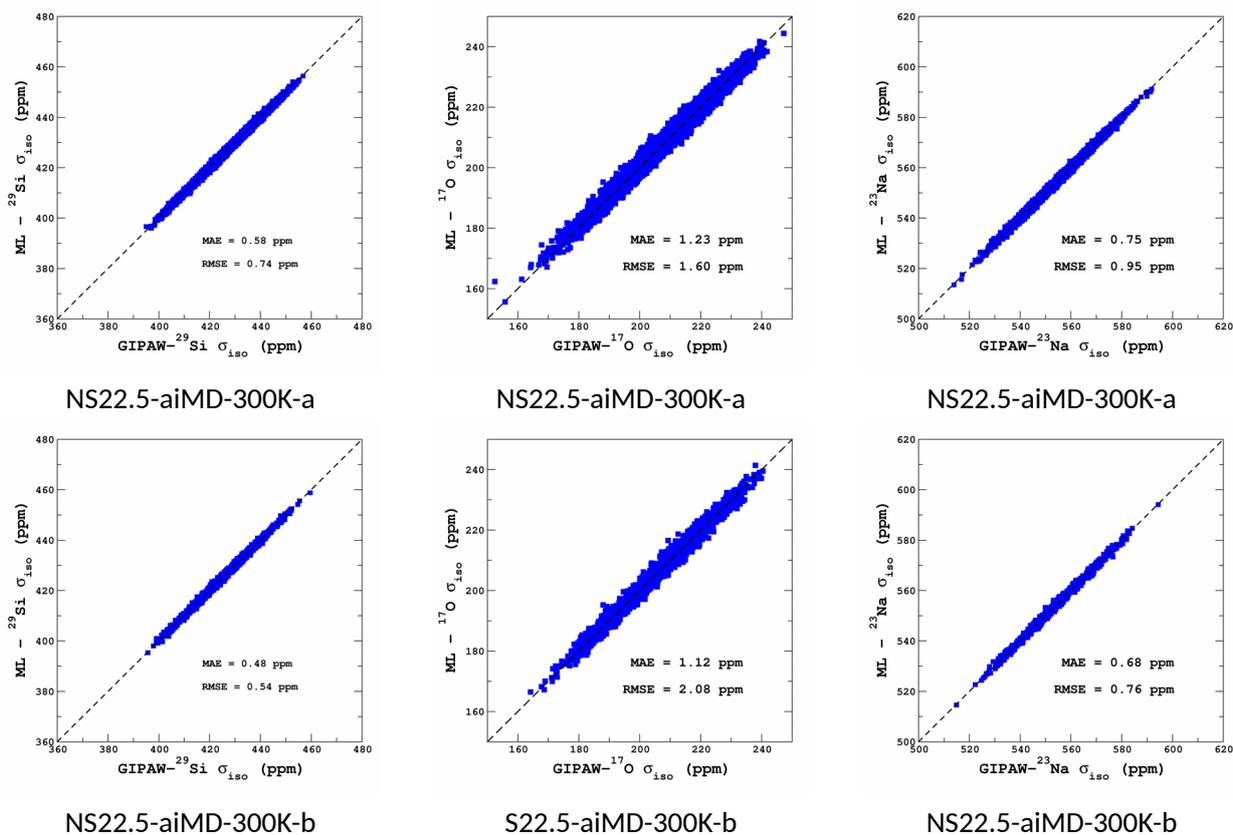


Figure S22. LSSVR versus DFT-GIPAW ^{29}Si , ^{17}O and ^{23}Na σ_{iso} (NS43.1). LSSVR was trained with the aiMD-300K-a dataset and tested on aiMD-300K-a (Top) and aiMD-300K-b (Bottom) datasets.

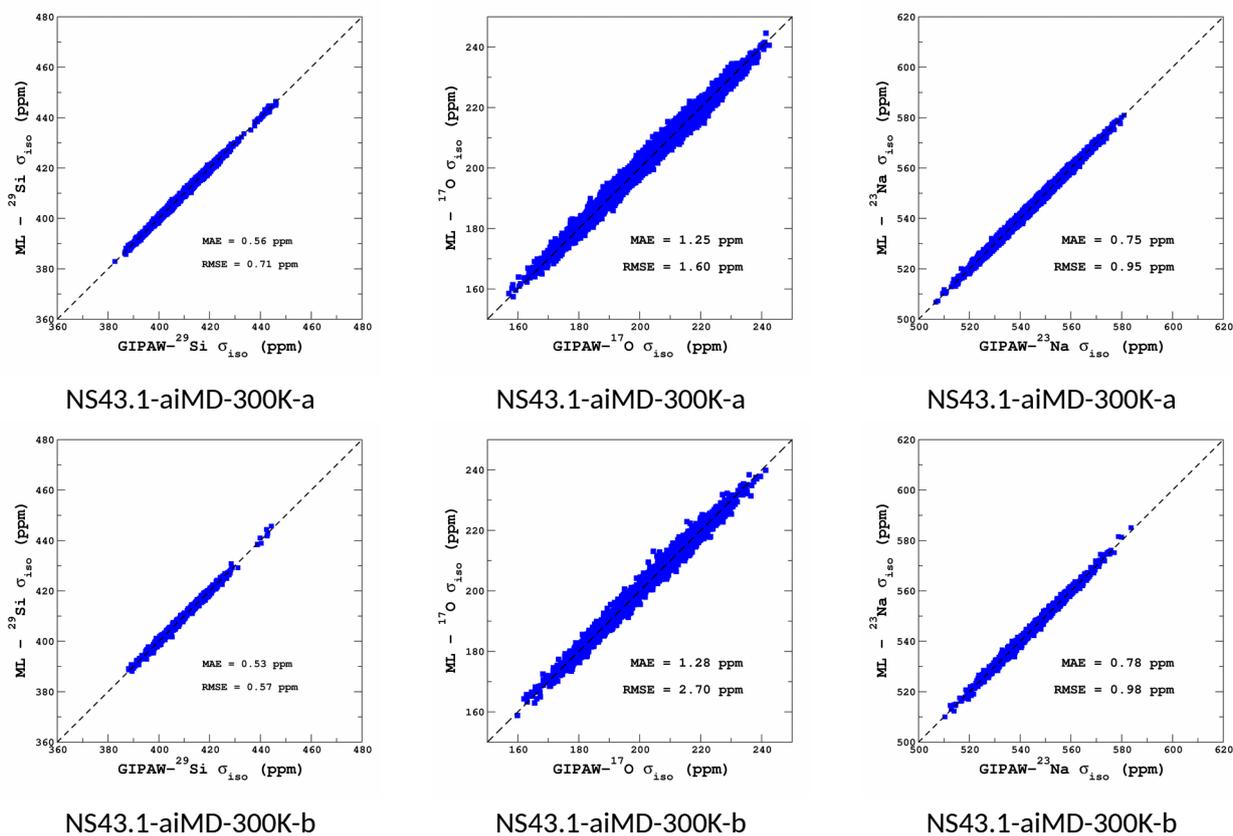
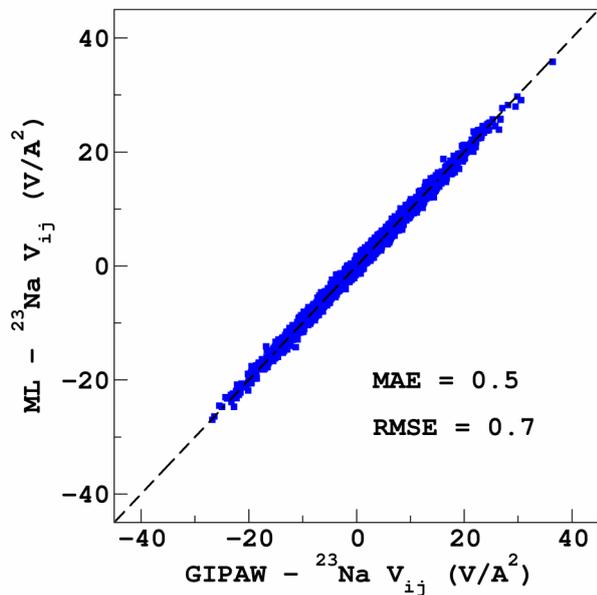
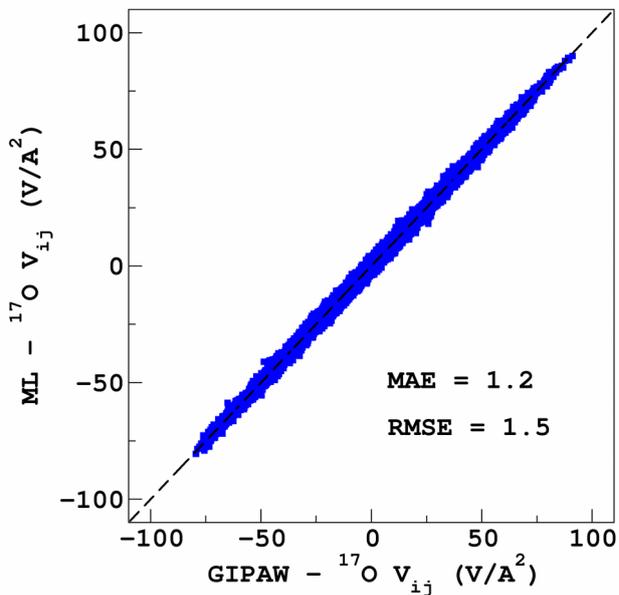


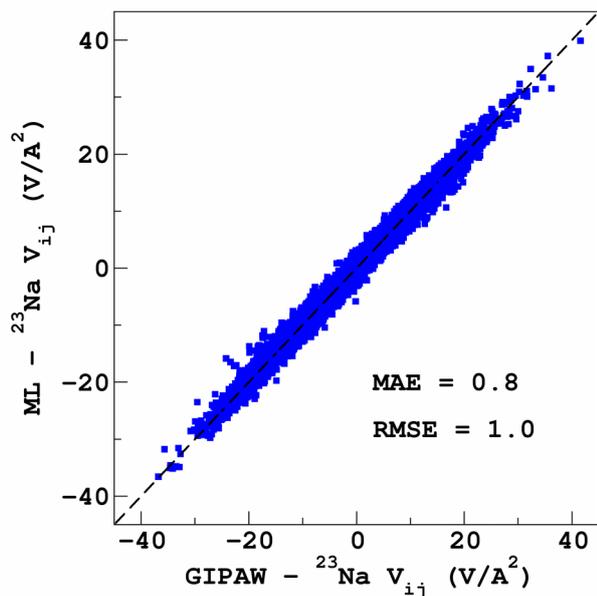
Figure S23. LRR λ -SOAP versus DFT-GIPAW ^{23}Na and ^{17}O EFG components (NS22.5). LRR was trained using the aiMD-300K-a dataset and tested on aiMD-300K-a (Top) and aiMD-300K-b (Bottom) datasets.



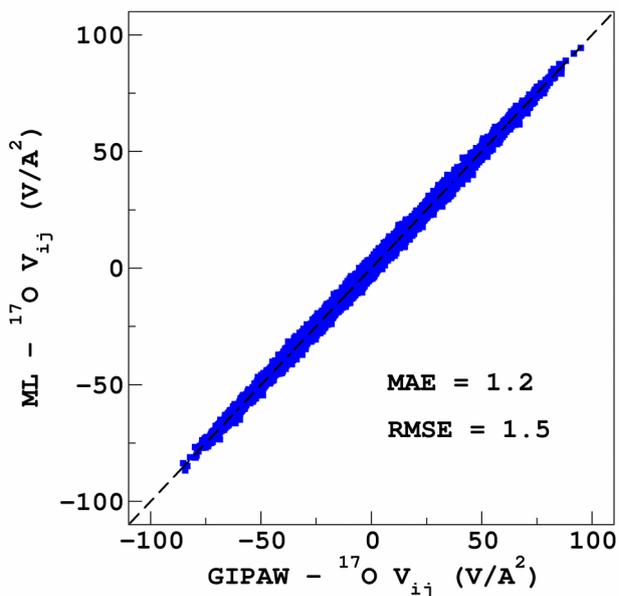
NS22.5-aiMD-300K-a



NS22.5-aiMD-300K-a

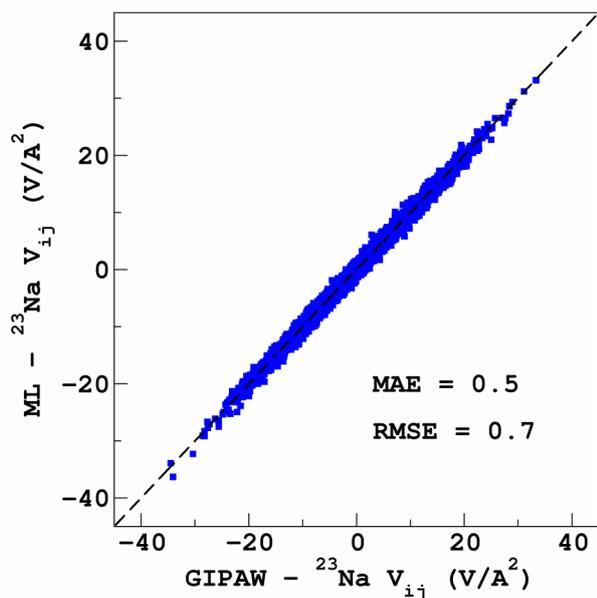


NS22.5-aiMD-300K-b

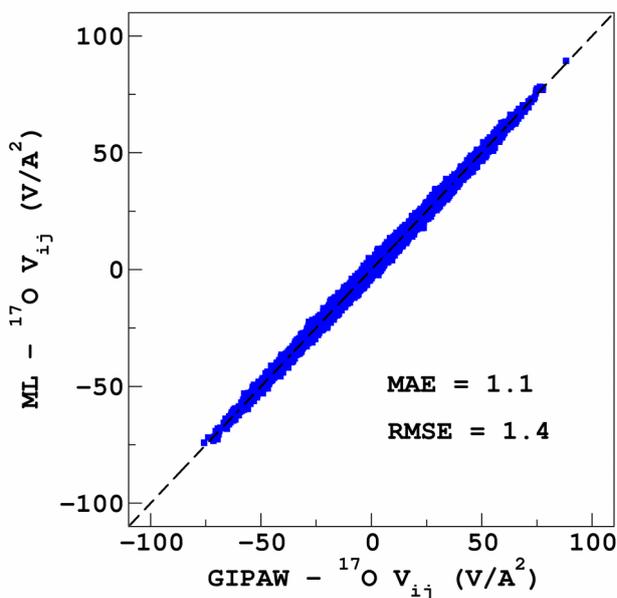


NS22.5-aiMD-300K-b

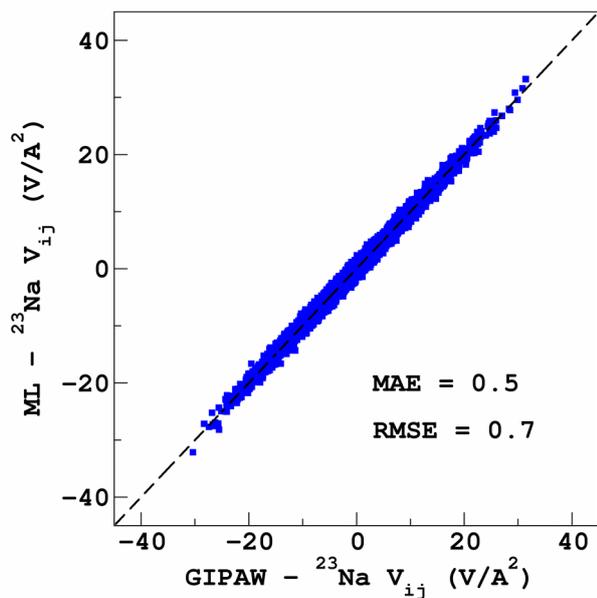
Figure S24. λ -SOAP LRR versus DFT-GIPAW ^{23}Na and ^{17}O EFG components (NS43.1) . LRR was trained with the aiMD-300K-a dataset and tested on aiMD-300K-a (Top) and aiMD-300K-b (Bottom) datasets.



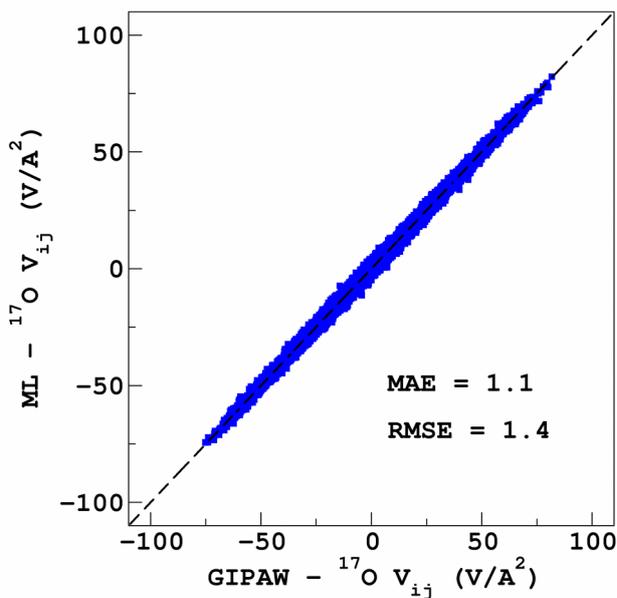
NS43.1-aiMD-300K-a



S43.1-aiMD-300K-a



NS43.1-aiMD-300K-b



NS43.1-aiMD-300K-b

S.8 Simulation of ^{23}Na MQMAS NMR spectra

Figure S25. ML-NMR simulation of ^{23}Na MQMAS NMR spectra with MD-300K and large models (using EFG scaling factor, see text).

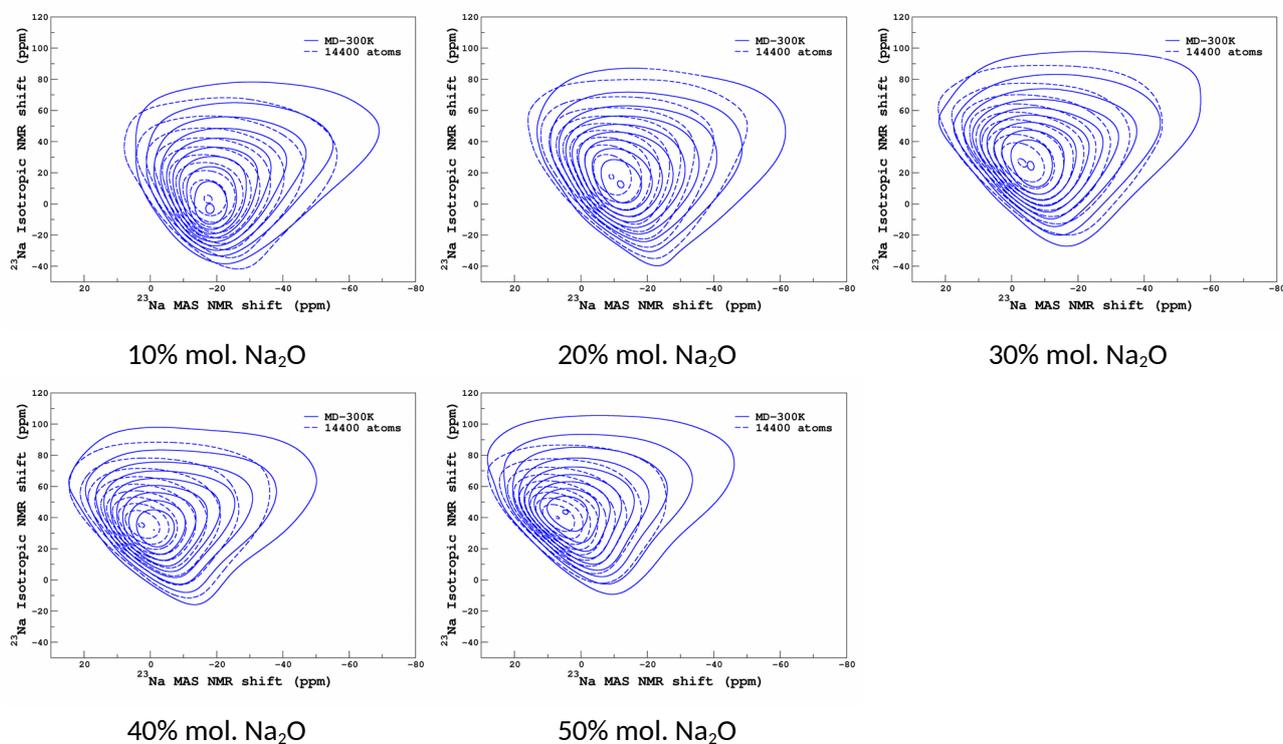
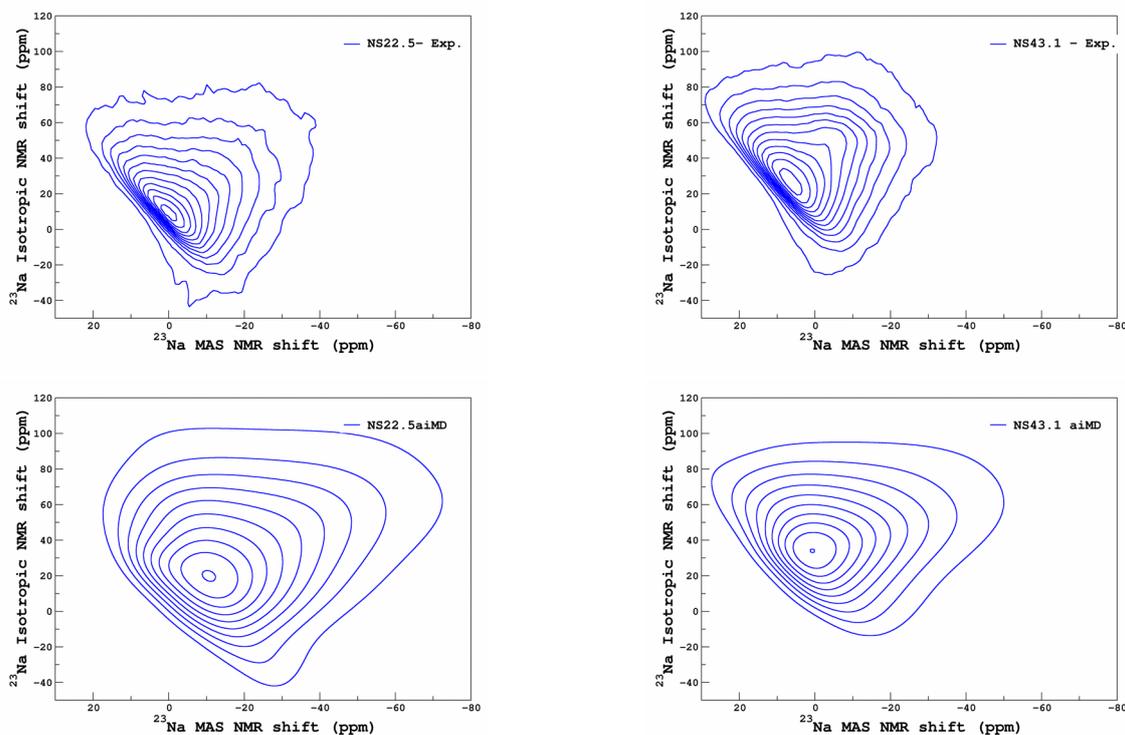


Figure S26. Comparison between ML-NMR simulation of ^{23}Na MQMAS NMR spectra with aiMD-300K (using EFG scaling factor, see text) and experimental data.



S.9 Simulation of ^{17}O MAS and MQMAS NMR spectra

Figure S27. ML-NMR simulation of ^{17}O MQMAS NMR spectra with MD-300K (using EFG scaling factor, see text).

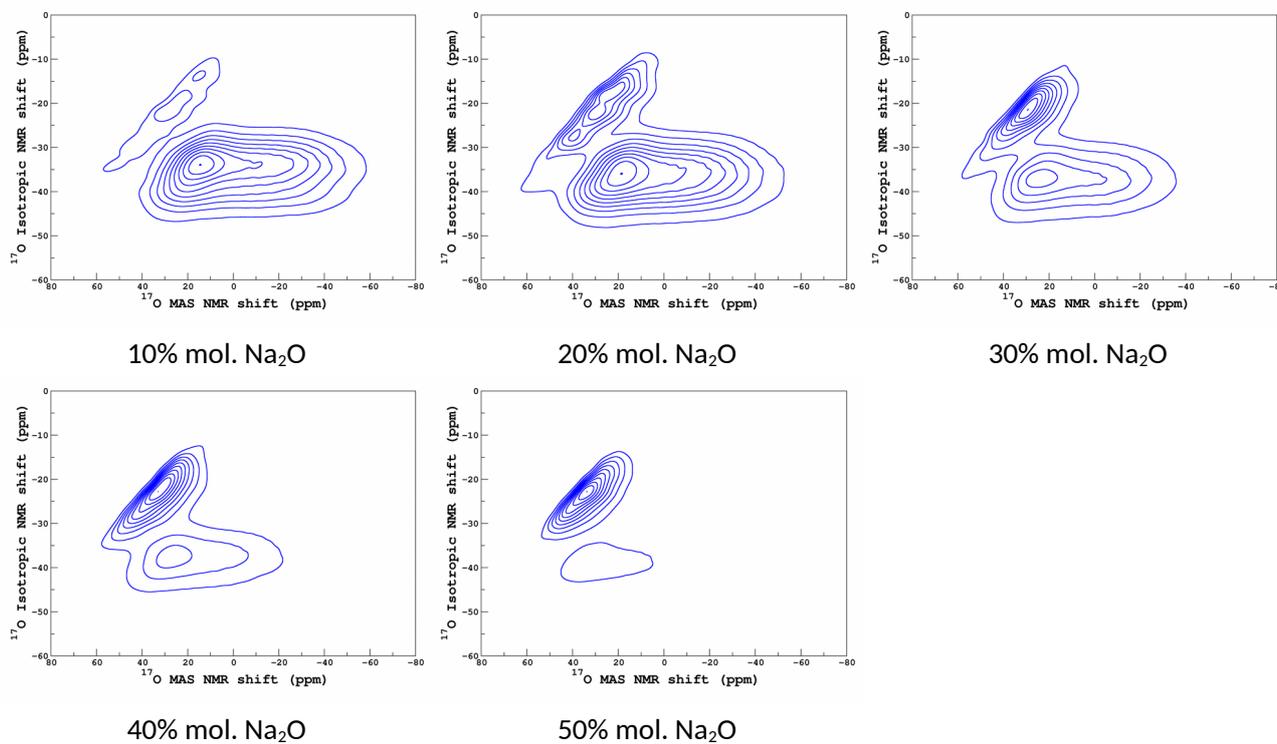
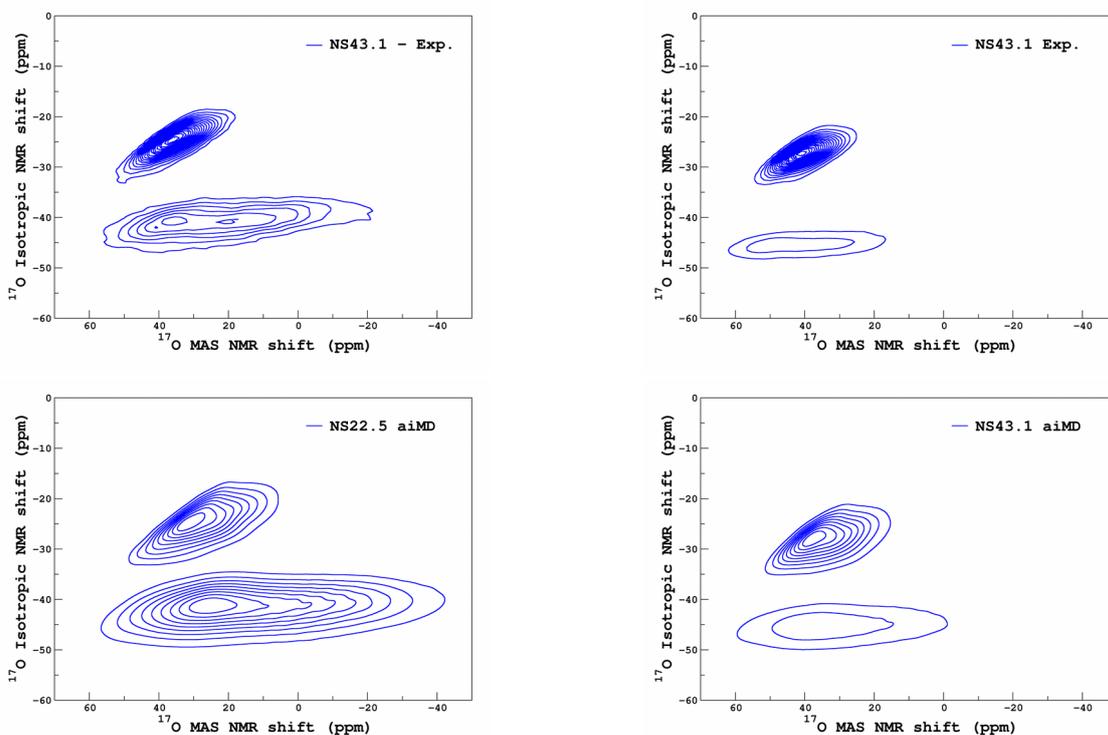
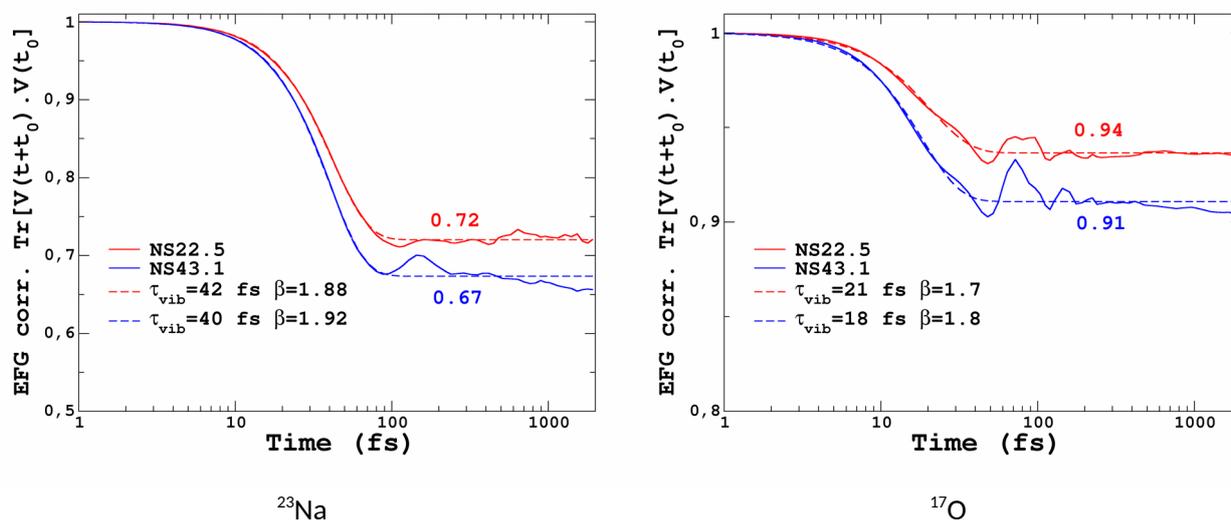


Figure S28. Comparison between ML-NMR simulation of ^{17}O MQMAS NMR spectra with aiMD-300K (using EFG scaling factor, see text) and experimental data.



S10. EFG tensor correlation functions

Figure S29. ^{23}Na and ^{17}O EFG correlation functions $G_{\text{EFG}}(\tau) = \langle \mathbf{V}(\tau+t_0) \mathbf{V}(t_0) \rangle_{t_0}$ of the NS22.5 and NS43.1 glasses (aiMD trajectory at 300K). Dashed lines represents fits with a single stretched exponential decay. Values of the plateau are indicated.



Notes and References

1. Guest MF, Elena AM, Chalk ABG. DL_POLY - A performance overview analysing, understanding and exploiting available HPC technology. *Mol Simul.* 2021;47(2-3):194-227. <https://doi.org/10.1080/08927022.2019.1603380>
2. Du J, Cormack AN. The medium range structure of sodium silicate glasses: a molecular dynamics simulation. *J Non-Cryst Solids.* 2004;349:66-79. <https://doi.org/10.1016/j.jnoncrysol.2004.08.264>
3. Kühne TD, Iannuzzi M, Del Ben M, et al. CP2K: An electronic structure and molecular dynamics software package - Quickstep: Efficient and accurate electronic structure calculations. *J Chem Phys.* 2020;152(19):194103. <https://doi.org/10.1063/5.0007045>
4. Perdew JP, Burke K, Wang Y. Generalized gradient approximation for the exchange-correlation hole of a many-electron system. *Phys Rev B.* 1996;54(23):16533. <https://doi.org/10.1103/PhysRevB.54.16533>
5. 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(15):154104. <https://doi.org/10.1063/1.3382344>
6. Bussi G, Donadio D, Parrinello M. Canonical sampling through velocity rescaling. *J Chem Phys.* 2007;126(1):014101. <https://doi.org/10.1063/1.2408420>
7. Bertani M, Charpentier T, Faglioni F, Pedone A. Accurate and Transferable Machine Learning Potential for Molecular Dynamics Simulation of Sodium Silicate Glasses. *J Chem Theory Comput.* 2024;20(3):1358-1370. <https://doi.org/10.1021/acs.jctc.3c01115>
8. Angeli F, Villain O, Schuller S, Ispas S, Charpentier T. Insight into sodium silicate glass structural organization by multinuclear NMR combined with first-principles calculations. *Geochim Cosmochim Acta.* 2011;75(9):2453-2469. <https://doi.org/10.1016/j.gca.2011.02.003>
9. Chaker Z, Salanne M, Delaye J-M, Charpentier T. NMR shifts in aluminosilicate glasses via machine learning. *Phys Chem Chem Phys.* 2019. <https://doi.org/10.1039/C9CP02803J>
10. Momma K, Izumi F. VESTA 3 for three-dimensional visualization of crystal, volumetric and morphology data. *J Appl Crystallogr.* 2011;44(6):1272-1276. <https://doi.org/10.1107/S0021889811038970>

11. Hafner J, Kresse G. The Vienna AB-Initio Simulation Program VASP: An Efficient and Versatile Tool for Studying the Structural, Dynamic, and Electronic Properties of Materials. In: Gonis A, Meike A, Turchi PEA, eds. *Prop. Complex Inorg. Solids*. Boston, MA: Springer US; 1997:69–82. https://doi.org/10.1007/978-1-4615-5943-6_10
12. Vasconcelos F, Wijs GA de, Havenith RWA, Marsman M, Kresse G. Finite-field implementation of NMR chemical shieldings for molecules: Direct and converse gauge-including projector-augmented-wave methods. *J Chem Phys*. 2013;139(1):014109. <https://doi.org/10.1063/1.4810799>
13. Pickard CJ, Mauri F. All-electron magnetic response with pseudopotentials: NMR chemical shifts. *Phys Rev B*. 2001;63(24):245101. <https://doi.org/10.1103/PhysRevB.63.245101>
14. Charpentier T. The PAW/GIPAW approach for computing NMR parameters: A new dimension added to NMR study of solids. *Solid State Nucl Magn Reson*. 2011;40(1):1–20. <https://doi.org/10.1016/j.ssnmr.2011.04.006>
15. Bonhomme C, Gervais C, Babonneau F, et al. First-Principles Calculation of NMR Parameters Using the Gauge Including Projector Augmented Wave Method: A Chemist's Point of View. *Chem Rev*. 2012;112(11):5733–5779. <https://doi.org/10.1021/cr300108a>
16. Charpentier T, Ispas S, Profeta M, Mauri F, Pickard CJ. First-Principles Calculation of ^{17}O , ^{29}Si , and ^{23}Na NMR Spectra of Sodium Silicate Crystals and Glasses. *J Phys Chem B*. 2004;108(13):4147–4161. <https://doi.org/10.1021/jp0367225>