npj | computational materials

Article

Published in partnership with the Shanghai Institute of Ceramics of the Chinese Academy of Sciences

Article in Press

https://doi.org/10.1038/s41524-026-02384-4

Generative inverse estimation of 3D atomic coordination from near-edge spectra via equivariant diffusion models

Check for updates

Ren Okubo[1], Yu Fujikata[1,2], Izumi Takahara[1] & Teruyasu Mizoguchi[1]

Extracting 3D atomic coordinates from spectroscopic data is a longstanding inverse problem. We present an equivariant diffusion model that generates the distribution of site-specific 3D structures directly from near-edge spectra (ELNES/XANES). Trained on Si–O crystals, the model achieves radial accuracy comparable to Extended X-ray Absorption Fine Structure (EXAFS) (bond-distance MAE ~ 0.04 Å) but with superior coordination number precision (errors < 2.0% vs. EXAFS ~ 20%). Crucially, it reconstructs full 3D geometries including bond angles, overcoming the limitations of 1D radial distribution analysis. The model shows encouraging, composition-dependent transferability to amorphous systems despite being trained exclusively on crystalline lattices. Application to experimental O *K*-edge spectra from *α*-quartz, cristobalite and stishovite demonstrates feasibility for experimental spectra. This generative approach goes beyond conventional template matching, which retrieves structures from predefined references, by enabling the exploration of candidate 3D structures within a learned structural space and providing a step toward automated, quantitative, 3D local-structure inference from spectroscopic data.

The creation of novel materials underpins innovation across electronics, aerospace, automotive engineering, energy technologies, and healthcare[1–3]. At the heart of this progress lies precise knowledge of local three-dimensional (3D) atomic coordination: bond lengths, coordination numbers, and bond angles govern macroscopic properties in decisive ways. Achieving atomic-level control over these structural features has therefore become a central goal in materials development, driving demand for analytical methodologies capable of resolving 3D coordination environments with atomic-scale precision.

Representative diffraction-based methods for structure determination include X-ray diffraction (XRD) and neutron diffraction (ND). These techniques measure periodicity based on the average atomic positions according to the Bragg diffraction condition. While such measurements reflect long-range order, they are less sensitive to medium-range and short-range order. In addition to the Bragg reflections that constitute XRD and ND, total scattering, which also includes diffuse scattering arising from atomic thermal vibrations, defects, and local structure, and its Fourier transform, the pair distribution function (PDF), contain information on short-range and medium-range order compared with the diffraction patterns described above. In particular, the PDF directly provides information on interatomic pair distances. Neutron and synchrotron total-scattering experiments combined with Reverse Monte Carlo (RMC) modeling are established and powerful approaches for constructing experimentally constrained three-dimensional atomistic models of disordered materials, providing access not only to interatomic distances but also to angular information. However, RMC can suffer from problems such as the lack of sufficient scientific or physical constraints and the non-uniqueness of the resulting structures arising from the measured data used. In addition, total scattering and PDF represent ensemble averages over sites within a structure, and site specificity is therefore lost.

In contrast to diffraction-based approaches, in the imaging domain, scanning transmission electron microscopy (STEM) represents one of the most effective approaches for atomic-resolution structural analysis. In STEM, a focused electron probe is scanned across a specimen to form images by detecting transmitted or scattered electrons. While atomic-resolution images can be obtained by converging the probe to dimensions smaller than atomic column spacing, the resulting data are limited to two-dimensional projections. Although TEM/STEM tomography can reconstruct three-dimensional information by acquiring images at multiple tilt angles, beam-induced damage, limited accessible tilt ranges, and reconstruction artifacts often compromise the completeness and accuracy of structural analysis, particularly for complex materials[4,5].

Complementary to imaging, spectroscopic techniques provide access to local coordination information. Extended energy-loss fine structure

[1]Institute of Industrial Science, The University of Tokyo, Komaba, Meguro, Japan. [2]Science & Innovation Center, Mitsubishi Chemical Corporation, Kanagawa, Japan. e-mail: teru@iis.u-tokyo.ac.jp

(EXELFS) in electron microscopy and its X-ray counterpart, extended X-ray absorption fine structure (EXAFS), have long served as cornerstone methods for local structure determination[6–10]. These techniques analyze oscillations in spectra arising from scattering of excited electrons or X-rays by the surrounding atomic environment. Fourier transformation yields radial distribution functions (RDFs) from which interatomic distances and coordination numbers can be extracted. In conventional least-squares fitting, EXAFS achieves impressive bond distance "fitting" precision (typically 0.01–0.02 Å for known structural models)[11–13]. Furthermore, by using reverse Monte Carlo (RMC) or molecular dynamics (MD)-based modeling, it is possible to obtain structural models that are consistent with EXAFS data and thereby infer three-dimensional atomic arrangements and angular information. However, the accuracy and uniqueness of such models strongly depend on the quality of the experimental data, the available constraints, the treatment of noise, and the underlying interatomic interaction model; therefore, sufficiently high accuracy is not always guaranteed[14–16]. For coordination-number estimation by EXAFS, an uncertainty of approximately 20% is often cited as a typical guideline[17,18], mainly because of the strong correlation between coordination number and Debye–Waller factors. The reliability of this estimation can be improved by using an ab initio Debye–Waller factor approach, in which Debye–Waller factors are independently evaluated based on DFT and are not treated as free fitting parameters. However, the actual accuracy depends on the system, structural model, and data quality, and the extent of the improvement has not been systematically established as a general quantitative value. Moreover, EXELFS and EXAFS require data acquisition over an extended high-energy range beyond the absorption edge, which is experimentally challenging and time-consuming, particularly in electron microscopy where beam damage and specimen drift become critical concerns.

In contrast to extended fine structure, the near-edge region of core-level excitation spectra—electron energy-loss near-edge structure (ELNES) in electron microscopy and X-ray absorption near-edge structure (XANES) in X-ray spectroscopy—can be acquired over a narrower energy range, making measurements more experimentally feasible and efficient. These near-edge spectra originate from excitation of core electrons at specific atomic sites and are highly sensitive to local coordination geometry, bond angles, and elemental composition[7,19]. The spectral features encode rich information about the three-dimensional arrangement of atoms surrounding the excited site. However, extracting explicit three-dimensional atomic coordinates from ELNES/XANES remains a fundamental challenge: interpretation has traditionally relied on qualitative comparison with reference spectra or first-principles calculations, limiting throughput and requiring substantial expertise.

Recognizing these challenges, researchers have increasingly turned to data-driven approaches to extract structural information from ELNES and XANES[20–26]. Machine learning methods can learn complex patterns from large datasets to predict properties efficiently. In 2017, Timoshenko et al.[14] used supervised machine learning to predict structural descriptors of metal nanoparticles, such as size, morphology, and coordination number, from XANES spectra. Following this work, several studies[27–29] reported the prediction of structural descriptors from XANES. For example, Kiyohara et al[30,31]. developed neural network models to predict coordination numbers and radial distribution functions from ELNES/XANES for silicone oxide crystal system. More recently, in addition to the reports predicting the oxidation state of copper from EELS and XAS spectra[32,33], Takahara et al.[34] demonstrated prediction of ground-state electronic structures from core-loss spectra. While these advances have been significant, existing methods have largely been limited to predicting averaged local-structure descriptors and have not yet achieved generation of full three-dimensional atomic coordinates at specific excitation sites.

Recent advances in generative modeling offer new possibilities for structure prediction. Diffusion models, in particular, have achieved groundbreaking results across diverse domains including image generation, computer vision, and natural language processing, and their application to atomic structure generation is advancing rapidly[35–38]. Kwon et al.[39] recently demonstrated conditional generation of amorphous carbon structures using a diffusion model with XANES spectra as conditioning inputs. Their approach successfully generated structures whose statistical descriptors, such as radial distribution functions and bond angle distributions, were consistent with those of target amorphous structures, thereby demonstrating the potential of spectroscopy-guided generative modeling for atomic structure generation. However, their study focused on bulk amorphous carbon in large supercells, and the generated structures were evaluated using global or ensemble-level descriptors, including Steinhardt order parameters, bond angle distributions, radial distribution functions, Voronoi cell volume distributions, and per-atom potential energies. Such metrics assess whether the generated structural ensemble is statistically plausible, but they do not directly evaluate whether a generated structure reproduces a specific target structure at the atomic-coordinate level. Therefore, the site-specific prediction of three-dimensional local atomic coordinates around individual excitation sites from ELNES/XANES spectra remains an open challenge.

To address this gap, we have developed a conditional diffusion model that generates the distribution of three-dimensional atomic coordinates in the local coordination environment around individual excited atomic directly from ELNES/XANES spectra (Fig.1). Our approach utilizes a diffusion model with an equivariant graph neural network architecture, explicitly preserving the rotational and translational symmetries inherent to atomic structures. Unlike previous work that generated global structural descriptors, our method generates explicitly candidates of three-dimensional positions of atoms surrounding the excitation site, enabling quantitative evaluation at the atomic level.

We focus on binary Si–O compounds as a model system for several compelling reasons. First, Si–O materials are of significant industrial and technological importance: silica and silicon oxides are fundamental in semiconductor manufacturing, lithium-ion battery anodes, catalysis, and optical devices[40–42]. Second, the O *K*-edge in Si–O compounds exhibit rich spectral features that are highly sensitive to local coordination environments. Third, the system benefits from numerous well-characterized crystal structures with diverse coordination motifs, providing an ideal dataset for training and validation. These characteristics make the Si–O system an optimal platform for demonstrating and rigorously evaluating site-specific three-dimensional coordination generation from ELNES/XANES spectra. Systematic differences can arise between experimental and calculated spectra owing to energy-axis shifts, differences in spectral shape and intensity, many-body and self-energy effects, limitations of the core-hole approximation, and instrumental noise and broadening. Therefore, in principle, it would be desirable to train the model using experimental spectra. However, for Si–O systems, there is no experimental spectral dataset with sufficient size and diversity for model training. In this study, we therefore constructed a dataset of calculated spectra and used it to train the model as an approximate data source for learning the correlation between local structures and spectral features.

Results

We trained a diffusion model on 2055 computed O K-edge spectra from 273 Si–O crystal structures, enabling direct generation of three-dimensional atomic coordinates conditioned on spectroscopic input. The crystal structures were randomly split into training, validation, and test sets in an 8:1:1 ratio with a random seed of 42, ensuring that all local coordination environments from the same crystal structure were assigned to the same split. The training dataset encompassed a compositional range of SiO_x ($x = 0.5–4.0$), though dominated by SiO_2 structures (1836 of 2055 spectra, Supplementary Fig. 1), reflecting the natural prevalence of silica polytypes in crystalline databases. Computational details for the O K-edge spectra are provided in the Methods.

To rigorously evaluate our method's inferential power and transferability, we designed a three-tier validation strategy. First, we assessed quantitative accuracy on held-out crystalline test structures, comparing performance against conventional template-matching approaches. Second,

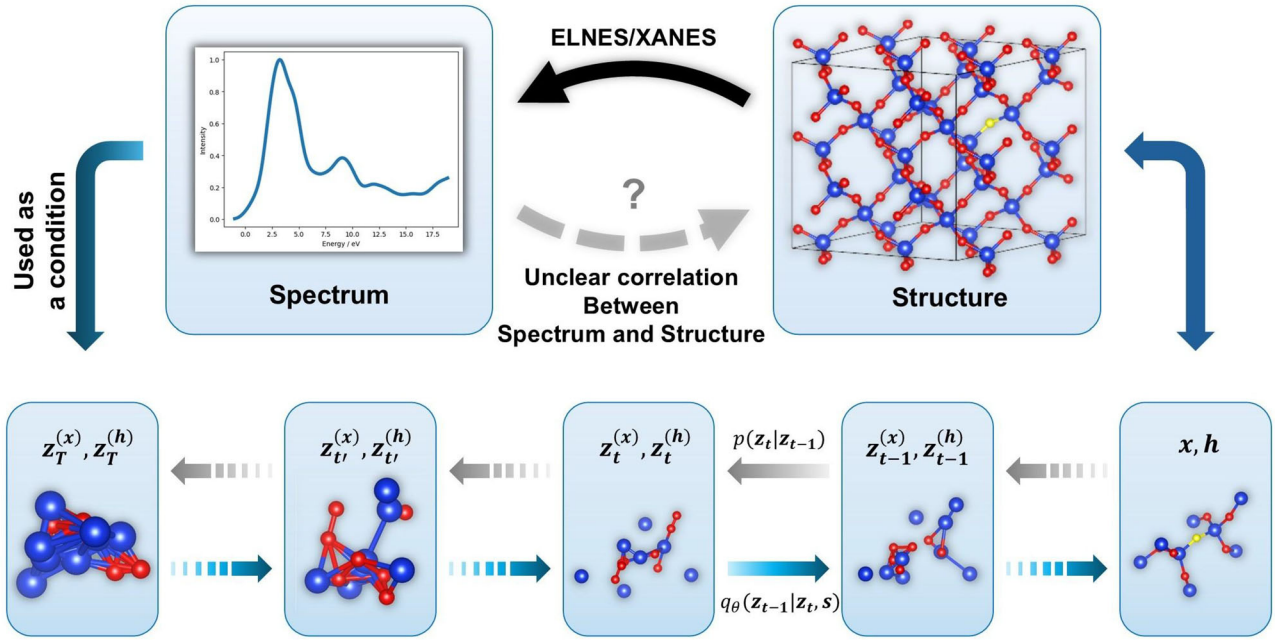


Fig. 1 | Overview of spectral-to-structure prediction approach. Our generative diffusion model, conditioned on experimental or calculated ELNES/XANES spectra, directly generates complete three-dimensional atomic coordinates around the absorbing atom. The denoising process begins with Gaussian noise z_T and progressively refines atomic coordinates, x, and node features h—comprising atomic species a—, and through timesteps ($z_{t'}$(x), $z_{t'}$(h), z_t(x), z_t(h), z_{t-1}(x), z_{t-1}(h)…) using an equivariant graph neural network, ultimately producing physically realistic 3D atom coordinates consistent with the input spectral condition.

we challenged the model with amorphous SiO_x systems ($x = 0.25-2.0$, prepared separately as described in "Methods") and further examined generation from an ensemble-averaged amorphous spectrum to assess whether statistical local-structure information can be extracted from averaged spectral inputs. Finally, we applied the crystalline-trained model to experimental O *K*-edge spectra from *α*-quartz, cristobalite and stishovite to demonstrate real-world applicability. Together, these benchmarks demonstrate that site-specific three-dimensional atomic coordination can be inferred directly from near-edge spectroscopic data, highlighting the potential of this approach to provide complementary local structural information alongside established structure-determination methods.

Generation of 3D coordination candidates from test spectra

We first evaluated the model's fundamental capability: can it accurately generate candidates of three-dimensional atomic coordinates around excitation sites when conditioned on ELNES/XANES spectra? For each test spectrum, we generated structures by sampling initial coordinates $\boldsymbol{X}_{initial}$, and atomic species $\boldsymbol{A}_{initial}$ from a standard normal distribution, conditioning on the spectrum $\boldsymbol{S}_{condition}$, running the reverse diffusion process and sampling candidate structures from spectrum-conditioned distribution $\boldsymbol{p}(\boldsymbol{x}, \boldsymbol{h}|\boldsymbol{s})$. As generation targets, we considered coordination environments extending to third-nearest neighbors from the excited oxygen site.

Figure 2 presents quantitative validation of the model's structural inference accuracy across all test spectra. In general, conditional diffusion models provide a distribution of multiple structural candidates that are consistent with the input condition, and structural inference from ELNES/XANES spectra is inherently a non-unique inverse problem. On the other hand, the site-specific O K-edge ELNES spectra considered in this study impose relatively strong structural constraints on the local coordination environment around the excited oxygen atom, thereby substantially restricting the allowable local structural space. Therefore, in this study, we used the RMSD relative to the target structure as a practical metric for quantifying the reproducibility of local geometries consistent with the conditioning spectrum. Following alignment of the generated and target structures using the procedure described in the Methods section, RMSD was calculated separately for atoms within the first, second, and third coordination shells; the results are shown in Fig. 2a. The overall mean RMSD across all atoms within the first three coordination shells is 0.678 Å (68 pm). More revealing, however, is the analysis focusing on the vicinity of the excited oxygen atom: the MAE of the average Si–O bond length around the excited oxygen atom was 0.041 Å (4.1 pm), which is on a length scale comparable to the fitting precision for average bond lengths often discussed in EXAFS analyses (0.01–0.02 Å). This is remarkable considering that EXAFS achieves such precision through a range of fitting approaches including conventional least-squares fitting, RMC, and MD-based methods, namely, iterative refinement against pre-defined structural models, whereas our approach generates atomic coordinates de novo from spectral data alone.

Beyond radial accuracy, the model demonstrates exceptional performance in coordination number prediction, a well-known limitation of conventional EXAFS analysis, which typically suffers from ~20% uncertainties due to correlation effects between coordination number and Debye-Waller factors in conventional least-fitting, although ab initio Debye–Waller factor approaches can partially mitigate this limitation[43,44]. Our model achieves coordination number prediction errors of only 1.94% and 0.82% for the first and second shell, representing more than one order of magnitude improvement over EXAFS capabilities.

Most critically, unlike any radial distribution technique, our diffusion model simultaneously generates full three-dimensional atomic arrangements, including angle—information accessible by EXAFS/EXELFS only under limited and favorable conditions. We note that the alignment procedure (Methods) does not guarantee the globally optimal rotation matrix, and slightly smaller RMSDs may be achievable with alternative alignments. Nonetheless, the achieved accuracy across multiple metrics—comparable radial precision to EXAFS (0.041 Å vs. 0.01–0.02 Å), order-of-magnitude superior coordination number accuracy (1.94% vs. ~20%), and unique 3D coordinate inference capability—positions spectral-to-structure generation as a powerful complement to, and in some aspects an improvement over, traditional local structure determination methods.

Figure 2b shows coordination environments corresponding to three representative cases marked by red circles in Fig. 2a. The case A is the highest accuracy in test data from SiO_2 crystal structure. The RMSD of the generated

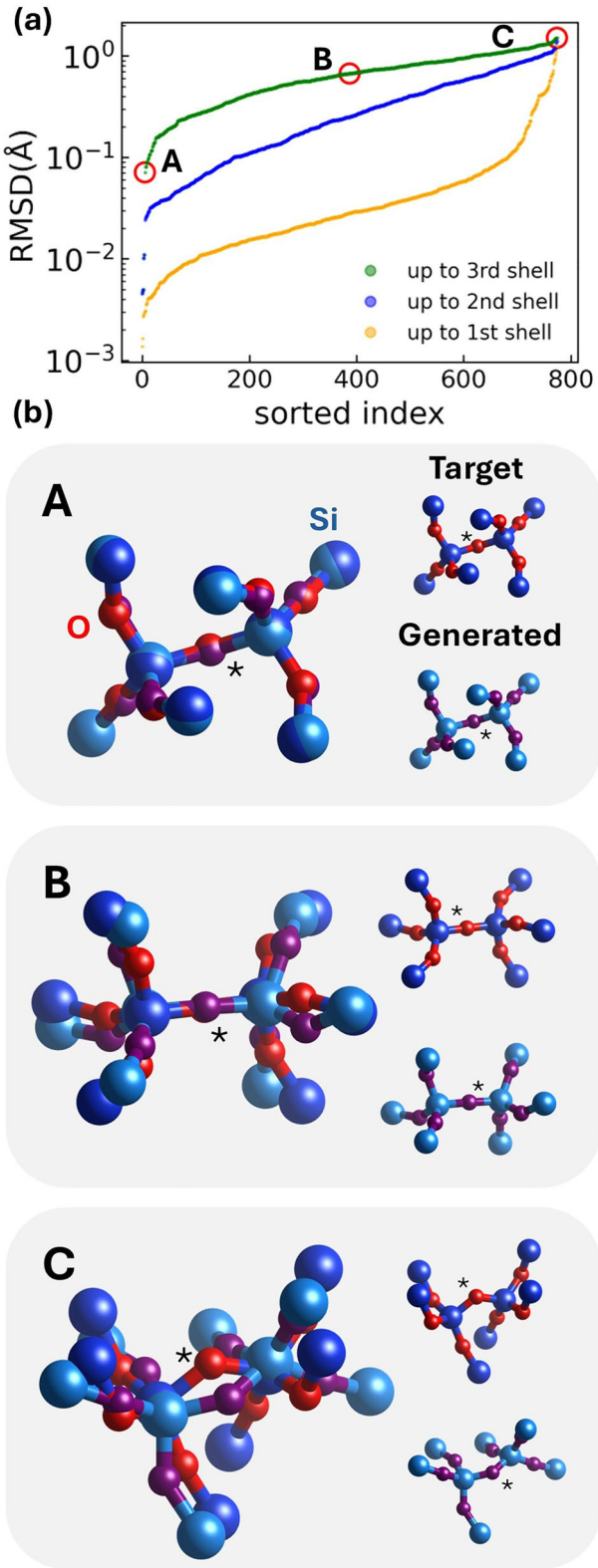


Fig. 2 | Root-mean-square distance (RMSD) for Results of Conditional Generation. a RMSD between target and generated structure of all the test data for 1st, 2nd, and 3rd shells. RMSD distribution across test local coordination environments. The vertical axis shows RMSD on a logarithmic scale, and the horizontal axis shows the index obtained by sorting environments in ascending order of RMSD. **b** Structures corresponding to the plot A, B, C in (**a**). The target structure is shown in the upper right of each frame and the generated structure in the lower right; an overlay of these two structures is shown to the left. The asterisk-marked site corresponds to the excited atomic site. Blue and red spheres correspond to Si and oxygen, respectively.

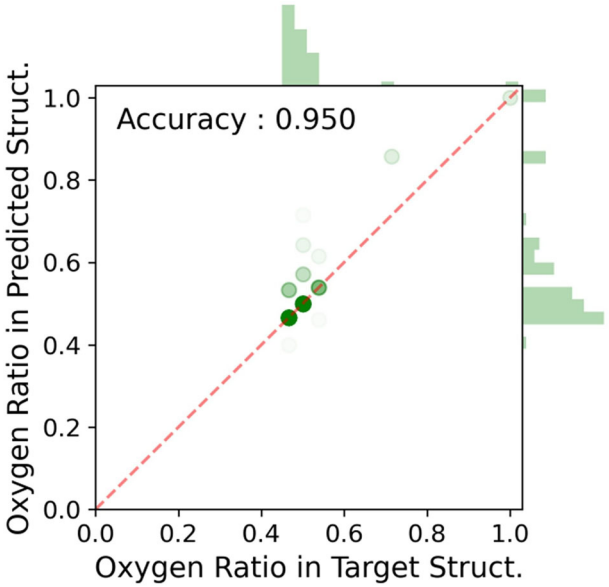


Fig. 3 | Prediction accuracy of elemental species. The accuracy is defined as the proportion of test cases where the oxygen fraction matches exactly.

structure relative to the target was 0.071 Å, indicating that not only distances but also all structural parameters including bond angles agree with the target with high accuracy. Absolute error of average Si–$O_{excited}$ bond distance is 0.015 Å, and absolute error of average Si–$O_{excited}$-Si bond angle is 0.859°. Visual inspection of the aligned structures confirms that the central excited O site, the first-shell Si, the second-shell O, and even the third-shell Si almost completely overlap between the target and generated structures.

For the median-accuracy case B (RMSD 0.671 Å), predictions are highly accurate up to second-nearest neighbors, with Si–O bond lengths well reproduced. The absolute error of average Si–$O_{excited}$ bond distance is 0.007 Å, and absolute error of average Si–$O_{excited}$-Si bond angle is 7.360°. The bond angles from the second- to third-nearest neighbors deviate from those of the target; nonetheless, even the third-shell oxygens are inferred at relatively close coordinates. In the lowest-accuracy case C (RMSD 1.509 Å), the generated structure exhibits the largest overall deviation from the target structure. Notably, the averaged local structural descriptors around the excited oxygen site are still reproduced reasonably well, with an absolute error of 0.008 Å for the average Si–$O_{excited}$ bond distance and 9.265° for the average Si–$O_{excited}$-Si bond angle. Therefore, the increased RMSD in this case cannot be explained simply by inaccuracies in these averaged local geometric quantities. Instead, it likely reflects positional deviations of atoms over a broader spatial range, including regions farther from the excited oxygen site. This interpretation is consistent with the distance-dependent analysis, which showed that atoms closer to the excited oxygen site tend to be generated more accurately than those farther away.

Next, we quantitatively evaluated the inference accuracy of elemental species. Figure 3 shows the oxygen fraction in generated versus target structures. We defined accuracy as the proportion of test cases where the oxygen fraction matches exactly. The proposed model achieved an accuracy of 0.950. To contextualize this value, we compared the model performance with two dummy baselines that do not extract any information from the input ELNES/XANES spectra. The first dummy baseline assumes equal probabilities for Si and O, corresponding to an uninformed prediction of an oxygen fraction of 0.5 for all test cases. This baseline achieved an accuracy of 0.168. The second dummy baseline accounts for the compositional bias of

the present training dataset. Because the dataset is biased toward SiO_2-based local structures, many local environments consist of seven O atoms and eight Si atoms, making 7/15 the modal oxygen fraction in the training dataset. A dummy classifier that always predicts this modal value achieved an accuracy of 0.781. The accuracy of the proposed model is higher than both baseline values, suggesting that the high accuracy cannot be explained solely by the compositional bias of the dataset. Rather, the model appears to extract element-specific information from the input ELNES/XANES spectra that contributes to the inference of the oxygen fraction. We also examined relationships among inference accuracy, compound composition, and training sample frequency. When inferred elemental species did not match targets, coordinate RMSDs also tended to be large. Regarding compositional dependence, coordination environments from SiO_2—the most abundant composition in the dataset—showed high success rates for species generation, whereas less frequent compositions such as Si_2O_5 and Si_3O_8 exhibited lower accuracy.

Comparison with conventional template matching

Having demonstrated accurate coordinate generation, we next asked whether this approach offers advantages over established analysis methods. Various methods have been used to extract structural information from ELNES/XANES spectra, including template matching, linear combination fitting (LCF), non-negative matrix factorization (NMF), multivariate curve resolution–alternating least squares (MCR-ALS), forward modeling, and neural-network-based approaches. These methods are useful for separating mixed components, extracting latent spectral components, correlating spectra with structural descriptors, or verifying spectral consistency with assumed structural models. However, the objective of the present study is to infer the three-dimensional local coordination environment itself around the excited site. From this perspective, template matching, which provides concrete structural candidates from a library of known structures, is the conventional method most directly comparable to the output of the present model. Therefore, we adopted template matching as the primary baseline in this study. We directly compared our generative model against template matching to quantify performance differences.

In template matching, we computed the cosine similarity between each test spectrum and all reference spectra in the database (the training split and the validation split of our Si–O crystal dataset), then assigned the structure corresponding to the minimum cosine similarity. To enable fair comparison, we quantified structural similarity using the Smooth Overlap of Atomic Positions (SOAP) descriptor[45]. We embedded both template-matched and model-generated structures into SOAP vectors and computed their cosine similarity to the target structure. Because the dataset was split at the crystal-structure level, spectra derived from the same crystal structures as the test spectra were not included in the reference database.

The result of the comparison is shown in Fig. 4. The horizontal axis represents the similarity between the matched spectrum and the target spectrum, whereas the vertical axis represents the difference score in structural similarity. Detailed definitions of these axes are provided in the Methods section. Points in green indicate cases where template matching achieves higher accuracy, whereas points in purple indicate superior performance by our generative model. Because the dataset contains a large fraction of SiO_2 structures, test spectra derived from SiO_2 crystal structures are evaluated under conditions in which template matching can relatively easily identify structurally similar candidates from the reference library. Indeed, template matching often provides accurate predictions under these conditions. Nevertheless, our model achieves a level of structural accuracy comparable to that of template matching, as reflected by the distribution of the structural-similarity difference score being centered around zero in Fig. 4. In addition, our model generated candidate structures with higher accuracy than template matching in 38.84% of the test cases. These results indicate that the proposed generative approach can predict structures with accuracy comparable to template matching, while also providing improved predictions in some cases. Importantly, unlike template matching, which is restricted to selecting pre-existing structures from the reference library, the proposed method explores plausible structural space conditioned on the input spectrum and directly generates candidate structures that are not limited to known database entries.

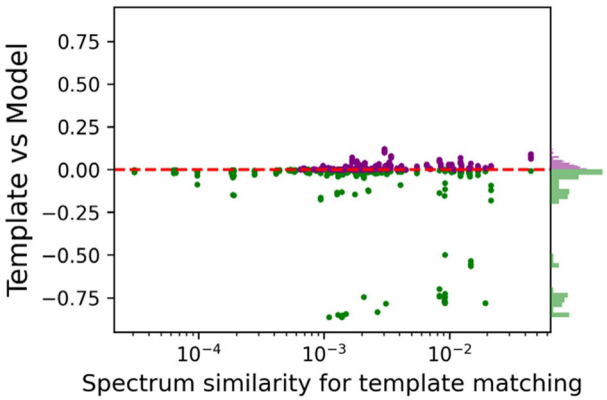


Fig. 4 | Comparison with conventional template matching. The horizontal axis shows the mean 1 - cosine similarity between the best-matching spectrum from the template database and the target spectrum, where larger values indicate more difficult cases for template matching. The vertical axis shows the structural similarity gap score, defined as the difference in SOAP-based cosine similarity to the target structure. Purple points indicate cases where our generative model achieves higher structural similarity to the target; green points indicate cases where template matching performs better. The right panel shows the distribution of similarity gaps. Our model outperforms template matching in 38.84% of test cases.

Robustness evaluation for the model

Experimental ELNES/XANES spectra inevitably contain measurement noise, energy-axis shifts, and broadening that reflects the instrumental resolution and sample conditions. Such perturbations can be one of the major sources of distributional differences when a model trained on calculated spectra is applied to experimental data. Therefore, in this study, we evaluated the robustness of the model by artificially introducing realistic perturbations expected in experimental spectra into the calculated spectra. The details of the procedure used to introduce each perturbation are described in the "Methods" section. After applying the perturbations to the test data used in Fig. 2, structure generation was performed using the model trained on unperturbed data. The changes in geometric accuracy evaluation as a function of the magnitude of each perturbation are summarized in Fig. 5 and Table 1.

Figure 5a shows the change in RMSD upon the addition of noise to the conditioning spectra. The RMSD values obtained using noisy spectra did not exhibit noticeable degradation compared with those obtained using the noise-free theoretical spectra. This result suggests that the proposed model has a certain degree of robustness against random noise in the input spectra. In particular, the model does not appear to rely excessively on small pointwise fluctuations or local high-frequency components in the conditioning spectra, but instead may extract structural information from more robust spectral features distributed over a broader energy range. In some cases, the RMSD values slightly decreased after noise addition; however, this should be regarded as a small variation associated with the stochastic nature of the generative inference rather than as evidence that noise systematically improves the prediction accuracy.

Figure 5b shows the generation results obtained using spectra broadened under conditions different from those used for the training dataset. While the Gaussian broadening applied to the training data was $\sigma = 0.50$ eV, the generation accuracy decreased when test data with larger σ values were used. In particular, for $\sigma = 0.90$ eV, the mean RMSD increased by approximately 0.31 Å compared with the case of $\sigma = 0.50$ eV. This degradation can be attributed to the systematic difference in spectral shape between the training and test data caused by the change in broadening, as well as to the decrease in the spectral information content due to increased

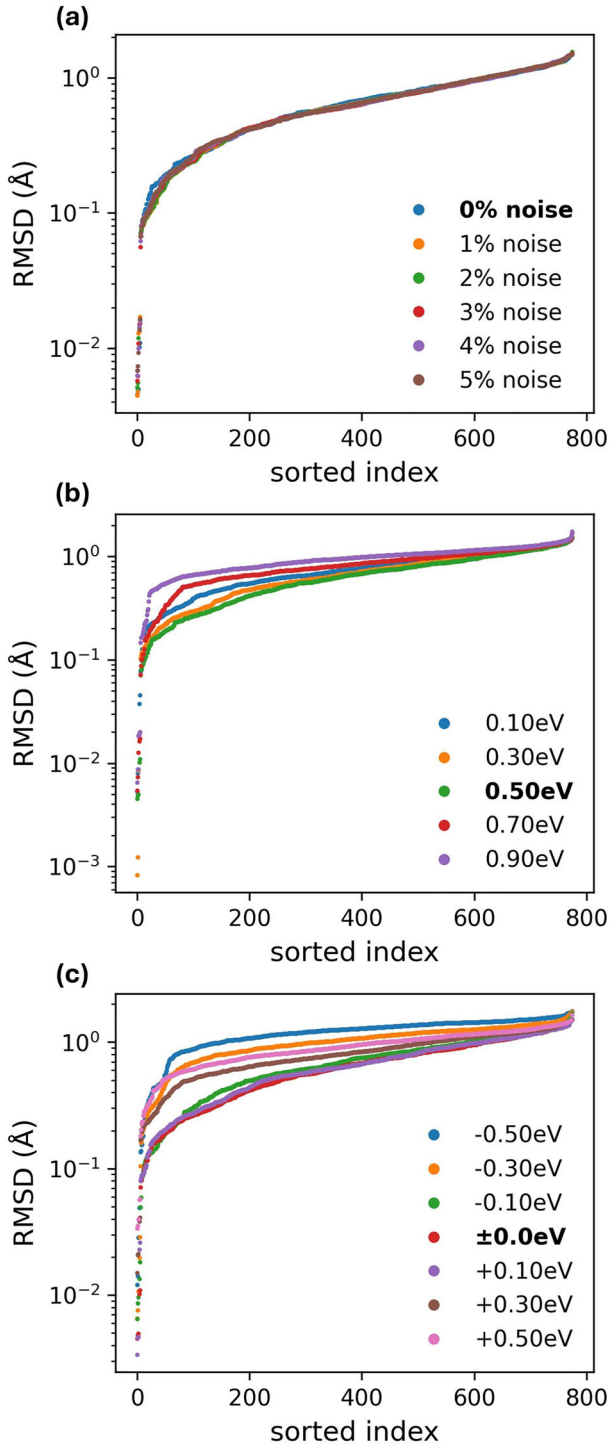


Fig. 5 | RMSD for the generation under perturbed spectrum condition. In all panels, values in bold indicate evaluation results obtained under unperturbed spectral conditions using the same test dataset as that used to generate the structures shown in Fig. 2. **a** RMSD for the generation under the noise added spectrum condition. Each label means the noise strength added to spectrum. **b** RMSD for the generation under the condition of spectrum with different broadening. Each label means the sigma value for gaussian broadening. **c** RMSD for the generation with energy-shifted spectrum. Each label indicates the shift magnitude.

broadening. On the other hand, when a smaller σ value was used, the generation accuracy was also lower than that for $\sigma = 0.50$ eV, but the magnitude of the degradation was smaller than that observed for larger σ values. This is likely because, although a systematic difference from the training data was introduced, the decrease in broadening made the peak structures more distinct, allowing spectral features related to the local structure to be relatively preserved.

Table 1 | Geometric accuracy evaluation of generated structures under perturbed spectrum conditions

Perturbation Type		Median – RMSD, ×10^{-2} Å			Mean Absolute Error	
		Up to 1st shell	Up to 2nd shell	Up to 3rd shell	Average Si–O* Distance, ×10^{-2} Å	Average Si–O*–Si Angle, deg
Noise	0%	2.82	25.09	67.09	4.13	6.79
	1%	2.66	25.90	64.33	4.19	7.19
	2%	2.65	25.91	63.68	4.25	7.18
	3%	2.68	24.86	62.36	4.13	6.75
	4%	2.70	24.84	63.25	3.92	6.79
	5%	2.79	25.32	63.76	4.67	7.13
Smearing	**0.1 eV**	5.23	31.84	75.42	6.38	11.05
	0.3 eV	4.32	28.31	69.16	8.09	10.21
	0.5 eV	2.82	25.09	67.09	4.13	6.79
	0.7 eV	4.84	38.72	84.20	5.14	10.04
	0.9 eV	7.09	52.38	97.45	6.58	13.78
Energy Shift	**-0.5 eV**	24.14	80.12	127.03	13.52	32.97
	-0.3 eV	16.78	61.52	106.71	9.33	24.32
	-0.1 eV	5.08	31.73	74.26	6.37	11.13
	±0.0 eV	2.82	25.09	67.09	4.13	6.79
	+0.1 eV	4.07	27.77	65.86	4.96	9.01
	+0.3 eV	7.76	37.98	83.17	6.87	13.39
	+0.5 eV	10.71	47.54	94.33	6.98	19.21

Underlined values indicate the evaluation results obtained under unperturbed spectral conditions using the same test dataset as that used to generate the structures shown in Fig. 2. Values in the "Perturbation Type" column denote, respectively, the noise amplitude relative to the maximum intensity of each spectrum, the sigma value used for Gaussian smearing, and the magnitude of the energy shift.

Figure 5c shows the change in RMSD as a function of the energy shift. Compared with the other two types of perturbations, the decrease in generation accuracy was more pronounced for the energy shift. When the spectrum was shifted by −0.5 eV, the mean RMSD increased by approximately 0.60 Å compared with the case without an energy shift. This result suggests that the model strongly utilizes information related to peak positions in generating local coordination environments and that the consistency of the energy axis has a significant impact on the accuracy of conditional generation.

In addition, the MAE of the average bond length between the excited oxygen atom and coordinating Si atoms changed by at most approximately 0.094 Å depending on the presence or absence of the energy shift. In contrast, the MAE of the average bond angle involving the excited oxygen atom and coordinating Si atoms reached 32.97° for an energy shift of −0.5 eV, which is 26.18° larger than that obtained without an energy shift. These results indicate that interatomic distance information can be estimated relatively stably, whereas the reproducibility of more subtle geometrical features, such as angles involving three-body correlations, is more susceptible to degradation.

Overall, the proposed model exhibited a certain degree of robustness against perturbations such as noise, changes in broadening, and energy shifts, although the generation accuracy decreased compared with the unperturbed condition. This result also supports the interpretation that the model effectively utilizes spectral features such as peak positions, relative intensities, and spectral shapes. At the same time, it indicates that the domain gap between calculated and experimental spectra remains an

important factor limiting the predictive performance. In particular, energy-axis calibration is suggested to be an important preprocessing step for generating three-dimensional local coordination environments, including angular information, from ELNES/XANES spectra. Because this result indicates that consistency in the energy alignment of the input spectra is important for conditional structure generation, we further examined whether training with energy-shift-augmented spectra could improve robustness against energy-axis misalignment; the results are provided in Supplementary Text 3 and Supplementary Fig. 4.

Application to amorphous materials

A critical question for practical applicability is whether a model trained exclusively on crystalline structures can be generalized to amorphous materials. This question has particular technological relevance: amorphous Si–O materials are used across diverse fields including glass, catalysis, electronics, biomedicine, and surface coating[40,42,46]. Moreover, amorphous SiO_x is regarded as a promising anode material for lithium-ion batteries due to high capacity and silicon resource abundance, with oxygen concentration strongly correlated to battery performance[41]. Fundamental understanding of local coordination in amorphous SiO_x could therefore significantly impact materials design across these applications. To investigate this, we constructed amorphous SiO_x structures ($x = 0.25, 0.5, 0.75, 1.0, 1.25, 1.5, 1.75, 2.0$) via classical molecular dynamics simulations using the Tersoff potential[47,48]. Preparing the cell consisting of around 200 atoms, we performed cooling from molten states and subsequently applied structural relaxation to obtain amorphous configurations[49]. For each oxygen site in the resulting structures, we computed the core-level excitation spectra under the same conditions described below and used these spectra as conditioning inputs during generation as described in Methods. We did not include any amorphous spectra in the training dataset and ensured that the model was trained solely on crystalline Si–O compounds.

Figure 6a, b shows, for amorphous $SiO_{2.0}$ and $SiO_{0.25}$, the RMSD of the three-dimensional coordinates in conditional generation as well as the target and generated structures. The results for all other compositions $x = 0.5, 0.75, 1.0, 1.25, 1.5$, and 1.75 are summarized in Supplementary Fig. 6 and Supplementary Table 1. It is important to note that the model was trained exclusively on crystalline structures; the amorphous configurations provide a stringent test of transferability beyond the crystalline training domain. For $SiO_{2.0}$ and $SiO_{0.25}$, the mean first-shell RMSDs were 0.173 and 0.474 Å, respectively. However, the results also reveal a clear composition dependence in generation accuracy. The mean first-shell RMSD increases from 0.173 Å for amorphous $SiO_{2.0}$ to 0.343 Å for amorphous $SiO_{1.5}$, indicating that the prediction accuracy already deteriorates when the composition moderately deviates from the SiO_2-rich domain. The degradation becomes more pronounced for oxygen-deficient compositions, as illustrated by the mean first-shell RMSD of 0.474 Å for amorphous $SiO_{0.25}$. These results show that the present model does not exhibit uniform generalizability across the entire Si–O compositional space and indicate the need to expand the dataset for underrepresented non-SiO_2 compositions.

For amorphous $SiO_{2.0}$, inferences reproduced atoms near the excited oxygen site with relatively high accuracy, and in many cases successfully captured local motifs such as oxygen coordinated by two Si atoms. The elemental species inference accuracy for amorphous $SiO_{2.0}$ was 84.4% (Supplementary Fig. 5). For amorphous $SiO_{0.25}$, the RMSD distribution exhibited bimodality: one group below 0.3 Å and another around 1 Å. The low-RMSD group corresponds to coordination environments with few atoms near the excited oxygen, a consequence of using a 2.0 Å cutoff radius and including up to third-nearest neighbors. In $SiO_{0.25}$, Si–Si bonds are sometimes not counted as neighboring atoms, leading to smaller coordination environments. While RMSD values were similar across compositions, visual inspection revealed that inferences for $SiO_{2.0}$ typically captured overall structural shapes, whereas those for $SiO_{0.25}$ often failed to reproduce characteristic local motifs such as Si–Si bonds and oxygen coordinated by three Si atoms. The elemental species inference accuracy for $SiO_{0.25}$ was notably lower at 16.4%, with a tendency to overestimate oxygen content.

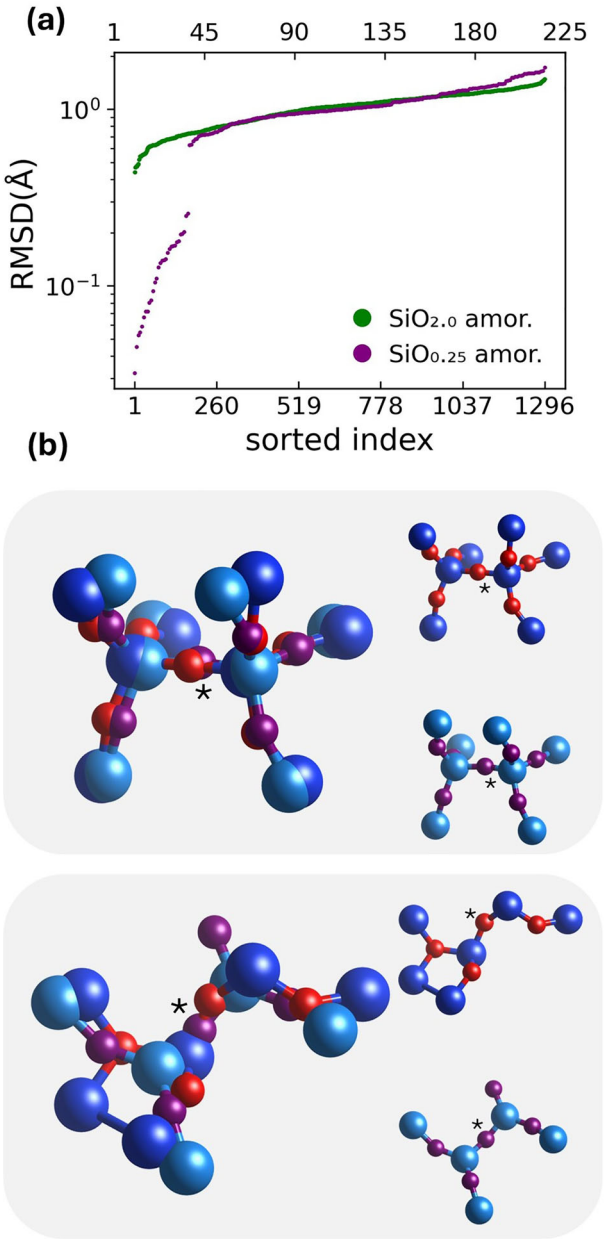


Fig. 6 | Conditional generation for amorphous (amor). a Coordination evaluation for $SiO_{2.0}$ and $SiO_{0.25}$ amor. Upper x-axis: $SiO_{0.25}$ amor. Lower x-axis: $SiO_{2.0}$ amor. **b** Target structures and predicted structures from the spectrum of $SiO_{2.0}$ and $SiO_{0.25}$ amor. The target structure is shown in the upper right of each frame and the generated structure in the lower right; an overlay of these two structures is shown to the left. The asterisk-marked site corresponds to the excited atomic site. Blue and red spheres correspond to Si and oxygen, respectively.

Despite having no exposure to disordered structures during training, the model successfully generated reasonable three-dimensional atomic coordinates for amorphous $SiO_{2.0}$ as shown in Fig. 6b, demonstrating a promising degree of transferability. The larger RMSD observed for $SiO_{0.25}$ reflects the more heterogeneous local coordination environments at this composition, which deviate more substantially from the crystalline training distribution. Overall, these results indicate that the model can generate reasonable local structures for amorphous $SiO_{2.0}$, despite being trained only on crystalline structures. At the same time, the reduced accuracy for $SiO_{1.5}$ and more oxygen-deficient compositions shows that this transferability is composition-dependent rather than uniform across the full Si–O compositional space.

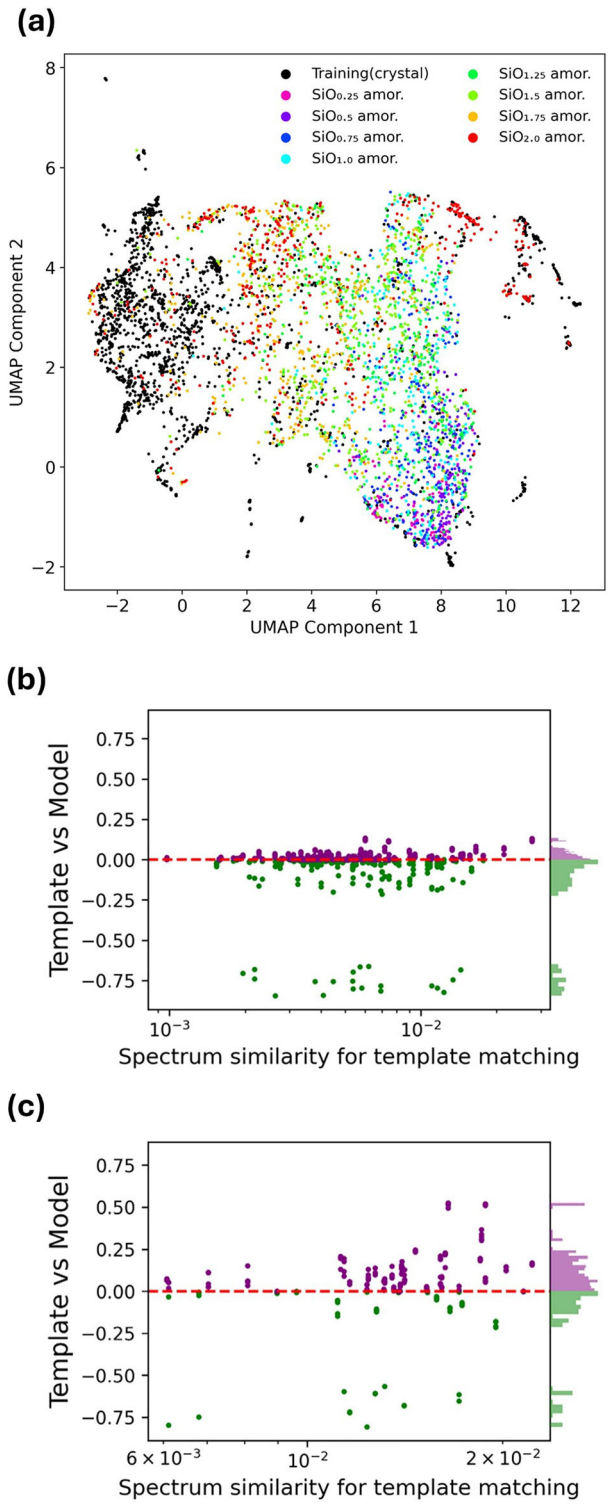


Fig. 7 | Relationship between model prediction and data distribution. a Spectrum data distribution created by UMAP. **b** Comparison between template matching and model prediction for $SiO_{2.0}$ amorphous. **c** Comparison between template matching and model prediction for $SiO_{0.25}$ amorphous.

Understanding performance differences across compositions

To understand these compositional differences, we visualized training and amorphous spectra via UMAP[50] (Fig. 7(a)). Amorphous spectra show composition-dependent distributions, shifting continuously as oxygen content varies. Training spectra cluster on the left, often near $SiO_{2.0}$ and $SiO_{1.75}$ amorphous spectra. Indeed, approximately 90% of training data came from SiO_2 materials, with the remaining ~10% from compositions such as SiO, Si_2O_3, and SiO_4 (Supplementary Fig. 21). This trend is consistent with the higher accuracy observed for amorphous $SiO_{2.0}$, whose spectra lie close to the training distribution, and with the reduced accuracy for oxygen-deficient compositions such as $SiO_{0.25}$, whose spectra are far from the available training examples.

Figure 7b, c compares our model against template matching for amorphous $SiO_{2.0}$ and $SiO_{0.25}$ amorphous spectra, respectively. For $SiO_{2.0}$, both methods perform well due to the similarity of those spectra to the training data, with our model achieving higher similarity in 53.6% of test cases.

For $SiO_{0.25}$, where spectra lie far from the training distribution, our model nevertheless achieved higher similarity in 70.2% of test cases. This suggests that, despite the limited representation of non-SiO_2 compositions in the training dataset, the model may still capture part of the spectrum–structure correlations associated with minority non-SiO_2 local motifs and can outperform template matching even in challenging cases where similar spectra are absent from the reference database.

Overall, in the prediction of amorphous structures using a model trained on crystal-structure-derived data, the proposed method achieved higher structural similarity than template matching in a certain fraction of cases, including over half of the test cases for amorphous $SiO_{2.0}$ and over 70% for amorphous $SiO_{0.25}$. Although template matching also provides a meaningful retrieval-based baseline, these results indicate that directly generating candidate atomic structures can be effective even for disordered systems whose structures are not explicitly included in the crystalline reference data. However, the shell-wise RMSD analysis shows that the generation accuracy degrades as the composition deviates from the SiO_2-rich region, reflecting the strong bias of the present training dataset toward SiO_2-derived local structures. Thus, the present model demonstrates encouraging transferability mainly for amorphous structures whose local environments are close to the training distribution, but it does not yet establish broad generalizability across the entire Si–O compositional space. Importantly, this limitation also indicates a clear direction for improvement: expanding and rebalancing the training dataset, together with targeted data augmentation of underrepresented oxygen-deficient and Si-rich local motifs, is expected to improve generation quality for minority compositions and to enhance the generality of the proposed spectrum-to-structure generation framework.

Generation from an ensemble-averaged amorphous spectrum

In the amorphous-structure analysis described above, site-specific O K-edge spectra calculated for individual oxygen sites were used as conditioning inputs. However, in practical ELNES/XANES measurements, the measured spectrum often contains contributions from multiple nonequivalent local environments and is therefore closer to an ensemble-averaged spectrum than to a strictly site-specific spectrum. We therefore examined how the present model responds to an ensemble-averaged spectral input by using amorphous SiO_2 as a representative disordered system.

For this analysis, the site-specific O K-edge spectra calculated for all oxygen sites in the amorphous SiO_2 structure were averaged to construct a single ensemble-averaged spectrum, which was then used as the conditioning input for repeated structure generation. The purpose of this analysis was not to reconstruct a unique local environment corresponding to a particular oxygen site, but to evaluate whether the distribution of generated local structures reflects statistical information contained in the original amorphous structure. Details of the generation procedure and statistical analysis are provided in Supplementary Text 7.

Figure 8a shows the site-specific spectra calculated for individual oxygen sites and the corresponding ensemble-averaged spectrum. Representative structures generated from this averaged spectrum are shown in Fig. 8b. These structures should be interpreted as samples from the distribution conditioned on the same ensemble-averaged spectrum, rather than as reconstructions of specific oxygen sites in the amorphous structure. To evaluate the generated structural distribution, we calculated the average

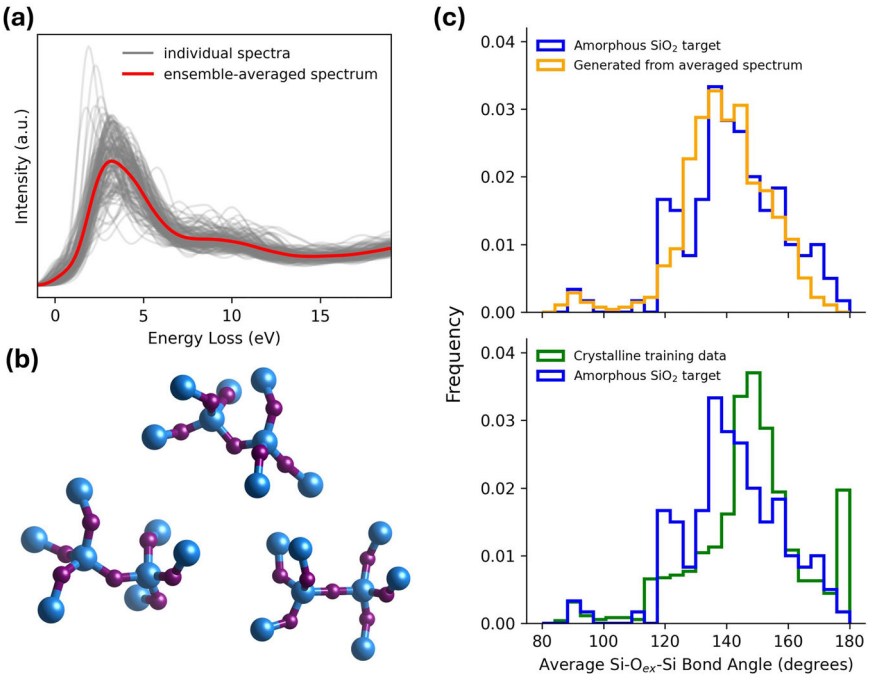


Fig. 8 | Generation from an ensemble-averaged amorphous SiO_2 spectrum. a Gray lines indicate the individual O *K*-edge spectra calculated for individual oxygen sites in the amorphous SiO_2 structure, and the red line indicates the ensemble-averaged spectrum. **b** Representative examples of local structures generated using the ensemble-averaged spectrum as the conditioning input. **c** Comparison of Si–$O_{excited}$-Si bond angle distributions. Upper panel compares the amorphous SiO_2 target distribution with the distribution obtained from structures generated using the ensemble-averaged spectrum. The lower panel compares the amorphous SiO_2 target distribution with the distribution obtained from the crystalline training dataset.

Si–$O_{excited}$–Si bond angle for each generated structure and compared the resulting distribution with that of the original amorphous SiO_2 structure and that of the crystalline training dataset.

As shown in Fig. 8c, the bond-angle distribution obtained from the generated structures was closer to the amorphous SiO_2 target distribution than to the crystalline training-set distribution. Quantitatively, the one-dimensional Wasserstein distance between the amorphous target and generated distributions was 3.39°, whereas that between the amorphous target and crystalline training-set distributions was 5.33°. A bootstrap analysis provided the same trend, as described in Supplementary Text 7. These results suggest that the ensemble-averaged spectrum retains statistical information on local coordination geometries and that the present generative model can use this information to sample local structures compatible with the averaged spectral input. At the same time, this result does not imply that an ensemble-averaged spectrum uniquely determines individual local environments; reconstruction of specific oxygen-site structures from such averaged spectra remains intrinsically non-unique.

Applicability to experimental spectrum

Finally, we examined the applicability of the proposed model to experimental spectra. Here, we consider conditional generation using experimental O *K*-edge ELNES/XANES spectra of three SiO_2 polymorphs: *α*-quartz, cristobalite and stishovite. *α*-quartz was selected in this study as it provides a well-characterized reference system[30,51,52]. Cristobalite was additionally examined as another SiO_2 polymorph with a tetrahedral Si–O network, providing an experimental test beyond α-quartz within tetrahedrally coordinated silica polymorphs. Stishovite, a high-pressure polymorph of SiO_2 with a markedly different local Si coordination environment, was also examined because its very high density and hardness make it relevant to fundamental studies of advanced ceramics and deep-Earth science[53–55]. Importantly, in each of these crystalline SiO_2 polymorphs, the oxygen sites are crystallographically equivalent. Therefore, although the measured experimental O *K*-edge spectra are ensemble-averaged over equivalent oxygen sites, they can be effectively regarded as representative site-specific spectra for the oxygen site in each structure and can be used as conditioning inputs for site-specific coordination-environment generation. To avoid including spectra corresponding to the target structures in the training data, the calculated spectra of α-quartz, cristobalite and stishovite were excluded from the full dataset. The remaining 2051 spectra were then split into training and validation datasets at a ratio of 9:1, and the trained model was used to predict local coordination environments conditioned on the experimental spectra.

Figure 9a–c shows the experimental and calculated O *K*-edge spectrum of *α*-quartz, the radial distribution function centered on the excited oxygen, and the target versus generated structures. The RMSD between the generated and target coordinates was 0.614 Å. In the RDF, the peak positions within approximately 3 Å of the excited oxygen largely coincided, whereas the peak around 4 Å, corresponding to the third-nearest neighbors, was not reproduced with full fidelity. This is consistent with previous work using neural networks to predict RDFs from XANES/ELNES[52], which also reported deviations in the 3–5 Å range. That study identified the near-edge rise region as the primary contributor to predicting average bond lengths and angles. In our case, although pronounced noise was observed in the experimental spectrum and several peaks in the post-edge region of the calculated spectrum were difficult to identify in the experimental spectrum, likely because of noise and/or broadening, the near-edge rise regions aligned well (Fig. 9a). This suggests that, as in previous studies, the near-edge rise played a major role in our model's distance and angle predictions. We anticipate improved accuracy through finer energy resolution in calculations and smoothing of experimental spectra.

Next, the generation results for cristobalite and stishovite are shown in Fig. 10. For cristobalite, the experimental O *K*-edge spectrum exhibited substantially smoother spectral features than the calculated spectrum, indicating a non-negligible difference in spectral broadening and/or spectral resolution between the experimental and calculated spectra. Under this condition, the generated structure reproduced the overall tetrahedral local coordination motif around the excited oxygen atom, which is characteristic of tetrahedrally coordinated silica polymorphs. The RMSD between the target and generated structures was 0.730 Å. The absolute error of the average Si–O bond length around the excited oxygen atom was 0.0231 Å, whereas that of the average Si–O–Si bond angle was 28.843°. These results indicate that the local bond-distance information was reproduced reasonably well, while the detailed angular information was less accurately recovered. This behavior is consistent with the robustness analysis, which showed that changes in spectral broadening and energy-axis alignment can degrade the reproducibility of subtle geometrical features such as bond angles.

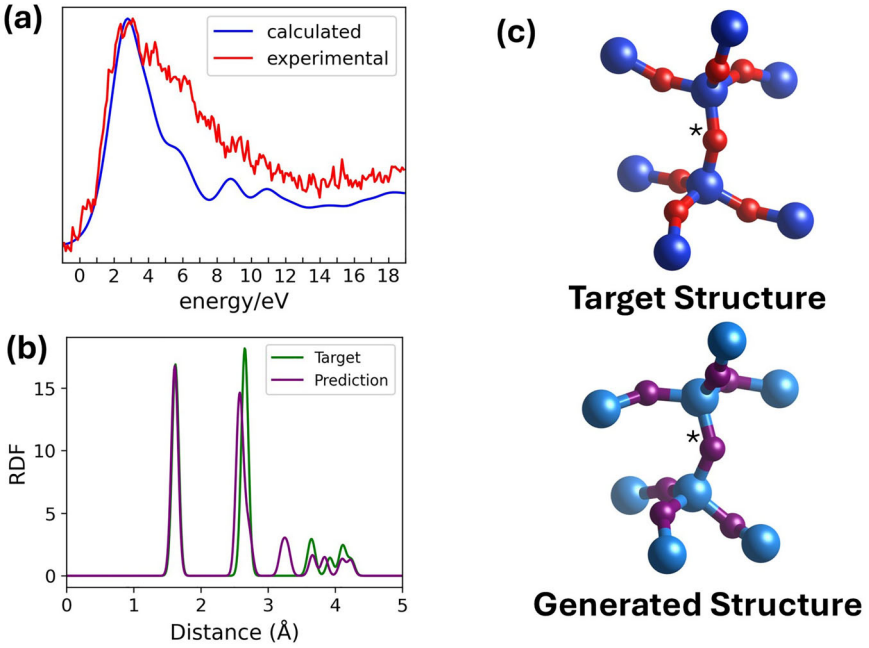


Fig. 9 | Conditional generation with experimental spectrum from α-quartz. a Experimental Spectrum of alpha quartz. **b** Radial distribution function of target structure and predicted structure. **c** Target Structure and generated Structure. The asterisk-marked site corresponds to the excited atomic site. Blue and red spheres correspond to Si and oxygen, respectively.

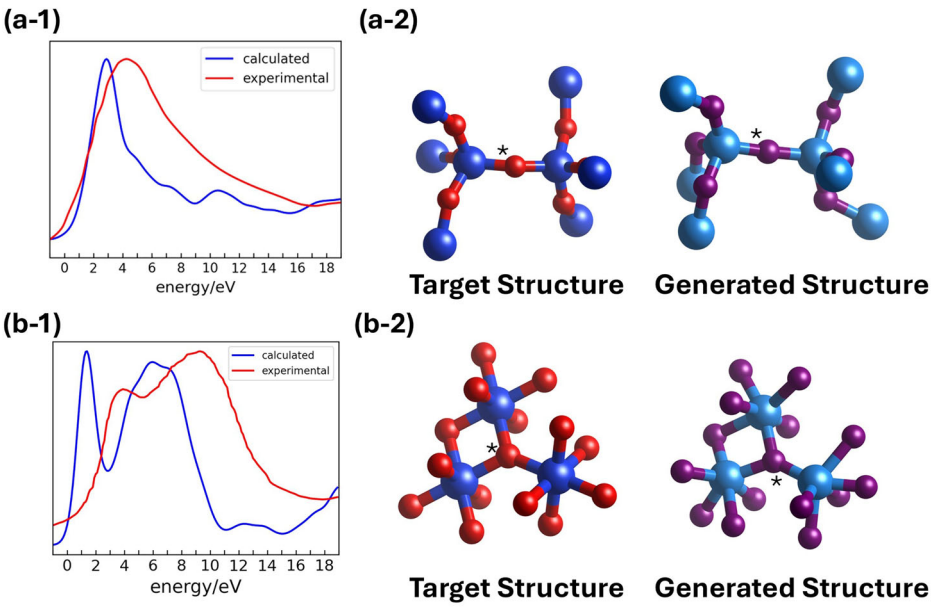


Fig. 10 | Conditional generation with experimental spectrum from cristobalite and stishovite. a-1 Experimental spectrum of cristobalite. **a-2** The local substructures around the excited oxygen atom in the target structure and generated structure. The asterisk-marked site corresponds to the excited atomic site. Blue and red spheres correspond to Si and oxygen, respectively. **b-1** Experimental spectrum of stishovite. **b-2** The local substructures around the excited oxygen atom in the target structure and generated structure, showing atoms up to the 2nd shell. The asterisk-marked site corresponds to the excited atomic site. Blue and red spheres correspond to Si and oxygen, respectively.

We next examined stishovite, which provides a more challenging test because its local coordination environment is distinct from the tetrahedral networks of α-quartz and cristobalite. In the comparison between the experimental and calculated spectra shown in Fig. 10(b-1), two major peaks were observed in both spectra, whereas a certain shift in the peak positions was also observed. In addition, a difference in the broadening width of the first peak was identified. The RMSD between the target and generated structures was 1.078 Å, which was larger than the mean RMSD shown in Fig. 2. One possible reason for this is that the local coordination environment of stishovite contains a relatively large number of atoms to be generated, resulting in a larger number of structural degrees of freedom to be inferred.

On the other hand, as shown in Figs. 10(b-2), when focusing on the structure near the excited oxygen atom in the generated structure, an octahedral structure in which six oxygen atoms coordinate to Si, which is characteristic of stishovite, was generated. The MAE of the average bond length between the excited oxygen atom and its coordinating Si atoms was 0.0498 Å, and the MAE of the average bond angle was 0.0285°, indicating that the local coordination motif around the excited oxygen atom was generated with high accuracy. Therefore, these results suggest that, even for stishovite, which has a different local coordination motif, the corresponding local coordination environment can be reproduced to some extent from the experimental spectrum.

Template matching was also performed for these experimental spectra. The similarities of the matched structures to the target structures were 0.964, 0.903, and 0.799 for α-quartz, cristobalite and stishovite, respectively. In contrast, the similarities of the structures generated by the present model were 0.985, 0.969, and 0.801 for α-quartz, cristobalite and stishovite,

respectively. Thus, the proposed model achieved a higher similarity than template matching for α-quartz and cristobalite, and a similarity comparable to template matching for stishovite.

Although the numerical similarity scores for stishovite were almost identical, the corresponding atomic structures showed a clear qualitative difference. As shown in Supplementary Fig. 8, the structure selected by template matching did not reproduce characteristic local motifs of stishovite, such as Si-centered octahedral coordination and oxygen atoms coordinated by three Si atoms. This result indicates a limitation of conventional template matching, which is intrinsically a retrieval-based approach and therefore selects the closest structure from a predefined reference database, even when the selected structure does not contain the essential local motifs of the target material. In contrast, the present model directly generates atomic structures conditioned on the input spectrum and can explore structural configurations beyond a discrete set of database entries. Therefore, even when the global similarity score is comparable to that of template matching, the proposed method can provide more physically meaningful local structural predictions by recovering characteristic coordination environments and local motifs.

In the experimental cases examined above, systematic differences, such as broadening, noise, and energy shifts, were observed between the experimental and calculated spectra. This reflects the general difficulty of achieving complete quantitative agreement between calculated and experimental spectra. Because differences in energy alignment and spectral broadening were particularly evident for the experimental stishovite spectrum, we also performed an additional generation test using alternative spectral preprocessing and a model trained on a broadening-matched training dataset; the results are summarized in Supplementary Text 5 and Supplementary Fig. 9. This additional test further confirmed that the local coordination motif of stishovite near the excited oxygen atom can be recovered, although accurate reconstruction of more distant coordination shells remains challenging. Nevertheless, the present model successfully reproduced the overall shape of the three-dimensional local coordination environment in the case of α-quartz, the tetrahedral local coordination motif around the excited oxygen atom in the case of cristobalite, and the local structural motif around the excited oxygen atom in the case of stishovite. Therefore, the correlation between spectral features and structure is considered to be shared to some extent between experimental and calculated spectra, suggesting the applicability of the present model to experimental spectra within this range.

Discussion

This work establishes that diffusion models can generate the distribution of site-specific three-dimensional atomic coordinates directly from ELNES/XANES spectra, a capability that opens new avenues for materials characterization.

The quantitative accuracy achieved by our model merits careful evaluation in the context of established structure determination techniques. For first-shell bond distances, our mean absolute error of the average Si–O bond distance of 0.041 Å is only 2–4 times larger than the typical "fitting" precision of EXAFS analysis (0.01–0.02 Å), despite addressing a fundamentally different—and arguably more challenging—inverse problem. EXAFS achieves its impressive precision through conventional least-squares fitting, namely, iterative refinement against pre-defined structural models, whereas our approach generates candidate atomic coordinates de novo from spectral data alone, without any structural priors or candidate models. More significantly, our model demonstrates exceptional performance in coordination number prediction, achieving errors of only 1.94% (first shell) and 0.82% (second shell). This represents more than one order of magnitude improvement over EXAFS analysis, where coordination number uncertainties of ~20% are widely recognized as a fundamental limitation arising from strong correlations between coordination number and Debye–Waller factors. This dramatic improvement transforms coordination number determination from a well-known "pitfall" of EXAFS analysis into a reliable quantitative measurement[43].

Most critically, our method uniquely provides candidates of complete three-dimensional atomic positions including bond angles and dihedral angles—structural information that EXAFS cannot directly access and EXELFS obtains only through complex multiple- scattering analysis requiring significant structural assumptions. This combination of capabilities—comparable radial accuracy, superior coordination number precision, and unique 3D coordinate generation—positions our approach not merely as an alternative to EXAFS/EXELFS, but as a complementary tool that addresses key limitations of traditional methods while maintaining competitive performance in their areas of strength.

The method is immediately applicable to several important classes of materials: (1) systems with crystallographically equivalent sites, exemplified here by α-quartz, cristobalite and stishovite, where all oxygen atoms share identical coordination environments; (2) two-dimensional materials and layered compounds where equivalent sites are arrayed along specific crystallographic directions[56]; (3) grain boundaries and interfaces exhibiting local symmetry; and (4) materials amenable to emerging depth-resolved spectroscopy techniques. For these systems, spectrum-to-structure prediction provides a powerful analytical tool that complements conventional characterization methods. In particular, when used in conjunction with local total-scattering/PDF-RMC analyses, the present ELNES/XANES-based generative method may provide complementary local and potentially site- and chemistry-sensitive structural information that can help constrain or interpret local coordination environments.

While acquiring truly site-resolved spectra from individual atoms distributed in three dimensions remains experimentally challenging, the measurement landscape is rapidly evolving. Notable progress has been made using state-of-the-art STEM-EELS, and emerging techniques such as depth-resolved ELNES using high-convergence electron beams are beginning to provide quasi-site-specific information along the depth direction[57]. Beyond electron microscopy, synchrotron-based XANES measurements offer complementary capabilities: while typically providing ensemble-averaged spectra, recent advances in focused X-ray beams and scanning techniques enable spatial resolution approaching tens of nanometers. In practical measurements, ELNES/XANES spectra often contain contributions from multiple non-equivalent local sites rather than representing a single site-specific coordination environment. To examine the model response to such averaged spectral inputs, we tested mixed spectra constructed as weighted averages of two ELNES spectra associated with different average Si–O–Si bond angles around the excited oxygen atom (Supplementary Fig. 10). The average bond angle in the generated structures varied continuously between approximately 139° and 180° according to the spectral mixing ratio, indicating that interpolation in the ELNES spectral space can be reflected by the model as interpolation in the structural-descriptor space. Although further validation is required for other descriptors, this result suggests that the model can generate plausible local-structure candidates even for conditioning spectra not explicitly included in the training data. In addition, the amorphous-structure results in the present study indicate that, even in materials with substantial local disorder, site-specific ELNES spectra can still contain useful information on the local coordination environment for generative structure inference. In addition, the ensemble-averaged amorphous-spectrum analysis presented in the Results section suggests that the proposed model can extract statistical local-structure information from averaged spectral inputs, although individual local environments cannot be uniquely reconstructed from such spectra. The additional tests using energy-shift augmentation and a broadening-matched training dataset further highlight the importance of spectral preprocessing and training-data preparation when applying the model to experimental spectra.

Beyond immediate applications, several clear pathways exist for enhancing model performance and broadening scope. Expanding training datasets to encompass diverse compositions, particularly off-stoichiometric and minority phases, would improve inferences across the full compositional range. Incorporating multi-edge information—for example, combining O *K*-edge with Si *L*-edge and/or *K*-edge, could provide

complementary constraints to enhance both coordinate accuracy and elemental species identification. Extension to ternary and higher-order systems would enable analysis of the broader range of technologically relevant materials. Integration of uncertainty quantification would allow the model to indicate inference confidence, particularly valuable when extrapolating beyond training distributions.

The observation that the model achieved higher SOAP-based structural similarity than conventional template matching in 38.84% of crystalline test cases and over 70% of amorphous cases, combined with quantitative accuracy approaching EXAFS precision for bond distances and exceeding EXAFS performance for coordination numbers, highlights the strength of the present approach for spectrum-to-structure inference. In contrast to template matching, which retrieves the most similar pre-existing structure from a database, our generative model directly produces explicit three-dimensional coordinate candidates from the input spectrum. This capability allows the model to propose comparable coordination environments that are not restricted to structures already present in the database, and may therefore help identify unexpected local structures, defects, or intermediate phases. The amorphous tests suggest that the model partially captures spectrum–structure correspondences associated with local structural motifs in the crystalline training data and that some of these correspondences can transfer to disordered local environments. However, this transferability is strongly composition-dependent. The improved performance for amorphous SiO_2-rich structures is consistent with their proximity to the SiO_2-dominated training distribution, whereas the degraded first-shell accuracy for amorphous $SiO_{1.5}$ and more oxygen-deficient compositions indicates that the current model does not yet provide broad generalizability across the full Si–O compositional space. This limitation identifies a clear route for improvement: expanding and rebalancing the training dataset, particularly by increasing oxygen-deficient and Si-rich local environments and by applying data augmentation to minority structural motifs, is expected to improve prediction accuracy for underrepresented compositions and enhance the generality of the spectrum-to-structure generation framework.

Looking forward, integration with high-throughput experimental workflows could enable rapid screening of local coordination in combinatorial materials libraries. Combined with active learning strategies, where the model identifies high-uncertainty inferences to guide targeted measurements, this could create closed-loop characterization systems that accelerate materials discovery. The approach is also directly applicable to operando studies, where tracking local coordination changes during battery cycling, catalytic reactions, or phase transformations could provide unprecedented insight into dynamic processes at the atomic scale.

Methods

Dataset preparation, spectrum calculations, and graph representation

This study targets 273 Si–O compounds registered in the Materials Project[58,59]. Si–O compounds are widely used in glass, batteries, and the semiconductor industry, and their three-dimensional (3D) local coordination environments are closely related to optical properties, ion conductivity, and dielectric characteristics; thus, structural analysis is of industrial importance[40–42,46,60]. Despite the industrial importance of Si–O compounds, the number of reported experimental spectra for these compounds is still insufficient. Therefore, in this study, we used theoretical spectra obtained from first-principles calculations as an approximate data source for learning the relationship between local structures and spectral features. For 2055 symmetrically unique oxygen sites, we constructed a dataset that pairs the ELNES/XANES spectrum at each site with the local coordination structure around that site, and we computed the O *K*-edge. The O *K*-edge lies near 530 eV and can be measured using electron microscopy or soft X-rays. Because the O *K*-edge mainly reflects information on O-*p* orbitals in the conduction band and on O-Si hybridization in addition to oxygen itself, it is well suited as the model spectrum used in this work.

Electronic-structure calculations for the O *K*-edge were performed using CASTEP[61] based on the plane-wave pseudopotential method, and the oxygen *K*-edge spectra were computed with OptaDOS[62]. For structural relaxation, we employed the GGA–PBE exchange–correlation functional. The plane-wave cutoff energy was uniformly set to 500 eV, and k-points were chosen using a Monkhorst–Pack grid[63] such that the k-point density was 0.03 $Å^{-1}$. Spin polarization was not considered. By generating the pseudopotential for the excited oxygen "on the fly," we calculated the excited-state electronic structure while accounting for the core-hole effect. The calculation was performed using an excited-electron and core-hole (XCH) setup, in which the excited O atom was described by a core-hole pseudopotential ({1s1}), while the total cell remained charge-neutral (charge = 0). To avoid interactions between core holes, we used supercells. In spectral calculations, with the internally calculated Fermi energy set to 0, the number of unoccupied bands was adjusted so that states were included up to 25 eV above the Fermi level in addition to the occupied states. Broadening was applied using a fixed-width Gaussian function with $\sigma = 0.5$ eV. Each spectrum was spline-interpolated to 200 points (0.1 eV steps, $-1 \sim 19$ eV) and normalized to max intensity value.

We treated, as the coordination environment corresponding to each spectrum, the atomic species and coordinates of atoms up to the third-nearest neighbors from the excited oxygen site, and set these as the targets for generation by the model. Neighboring atoms were defined as sites within a cutoff radius (=2.0 Å) from each atomic site. Data in which no atoms existed in the first-nearest shell—i.e., environments defined by only a single atom—were excluded from the dataset.

The crystal structures were randomly split into training, validation, and test sets in an 8:1:1 ratio with a random seed of 42, ensuring that all local coordination environments from the same crystal structure were assigned to the same split.

We also generated amorphous structural data for SiO_x ($x = 0.25, 0.5, 0.75, \ldots, 2.0$) via molecular dynamics (MD) using the Tersoff potential[47–49]. After structural relaxation, ELNES/XANES spectra were computed for these structures under the same conditions as above, and the resulting data were used as test cases to assess the model's applicability to amorphous systems. In the amorphous simulation, totally 45, 75, 90, 100, 125, 144, 140, and 144 O *K*-edges from oxygen sites were calculated for $x = 0.25, 0.5, 0.75, 1.0, 1.25, 1.5, 1.75, 2.0$ SiO_x compounds, respectively.

To evaluate the applicability of the model to experimental spectra, three-dimensional local coordination environments were generated from the experimental O K-edge spectra of α-quartz, cristobalite and stishovite. The experimental O K-edge ELNES spectrum of α-quartz was measured using an aberration-corrected STEM equipped with a monochromator (JEM-ARM200F, JEOL Ltd.) with an energy resolution of 30 meV. Because the O K-edge is broadened by approximately 0.1 eV due to the core-hole lifetime, the instrumental energy resolution of 30 meV is considered sufficiently small.

Cristobalite was additionally examined as another SiO_2 polymorph with a tetrahedral Si–O network, providing an experimental XANES-based test beyond α-quartz within tetrahedrally coordinated silica polymorphs. The experimental O K-edge XANES spectrum of cristobalite was digitized from the spectrum reported by Brown Jr. et al.[64].

Stishovite is a high-pressure polymorph of SiO_2 with a rutile-type structure, and it has a distinctive coordination environment in which oxygen is threefold coordinated within the SiO_2 composition. Therefore, it is considered suitable for evaluating the applicability of the model to experimental spectra. The experimental O K-edge ELNES spectrum of stishovite was obtained by degitizing the O K-edge spectrum of stishovite reported by Sharp et al. from the published figure[65]. These experimental spectra were aligned by setting the energy corresponding to the threshold intensity $I_{\mathrm{threshold}}$, calculated using the following equation, to 0:

$$I_{\mathrm{threshold}} = I_{\mathrm{base}} + 0.1 \times (I_{\mathrm{peak}} - I_{\mathrm{base}})$$

The search window for I_{peak} was set to 520–540 eV. I_{base} was defined as the maximum intensity within the search window, and I_{base} was defined as the median intensity in the lower 15% energy range of the search window.

No additional broadening was applied, and the intensities were normalized by the maximum intensity in the same manner as for the training dataset. The experimental spectra were also converted into 200-dimensional vectors over the range from −1 to 19 eV using the same interpolation procedure as that used for the training spectra.

A coordination environment containing N atoms was represented as a fully connected graph $\boldsymbol{G} = (\boldsymbol{X}, \boldsymbol{A}, \boldsymbol{S})$ in tuple form. Here, $\boldsymbol{X} = (\boldsymbol{x_1}, \boldsymbol{x_2}, , \boldsymbol{x_N}) \in \mathbb{R}^{N\times 3}$ denotes the cartesian coordinates of each atom with the excited oxygen site taken as the origin; $\boldsymbol{A} \in \{0,1\}^{N\times 2}$ encodes the elemental species of each atom using a one-hot representation, under the assumption that the elemental species are limited to Si and O; and $\boldsymbol{S} = (\boldsymbol{s}, \boldsymbol{0}, , \boldsymbol{0}) \in \mathbb{R}^{N\times 200}$ holds the spectral information. In this notation, $\boldsymbol{s}$ is the spectral vector at the excited oxygen site obtained from the calculations described above, and $\boldsymbol{0}$ denotes a zero vector of the same size as $\boldsymbol{s}$. The first indices of $\boldsymbol{X}$, $\boldsymbol{A}$, and $\boldsymbol{S}$ are assigned to excited oxygen; there are no constraints on how the remaining indices are assigned. In structure generation from spectra, the number of atoms N constituting the local coordination environment was assumed to be known.

Perturbation of calculated spectra with realistic noise, energy shifts and broadening

To evaluate the robustness of the model against differences between experimental and calculated spectra, we prepared test datasets in which noise, energy shifts, and changes in broadening were separately applied to the calculated spectra in the test data. For noise addition, Gaussian noise with a standard deviation corresponding to 1, 2, 3, 4, or 5% of the maximum intensity of each test spectrum was added. For the energy-shifted datasets, each spectrum was shifted along the energy axis by ±0.1, ±0.3, or ±0.5 eV. After the shift, each spectrum was sampled over the range from −1 to 19 eV at 0.1 eV intervals to obtain a 200-dimensional spectral vector. For the datasets with modified broadening, spectra were generated by applying Gaussian smoothing with $\sigma = 0.10$, 0.30, 0.70, or 0.90 eV, using the Gaussian smoothing condition of the training dataset, $\sigma = 0.50$ eV, as the reference.

Model architecture

An equivariant graph neural network (GNN) that incorporates the inductive bias of equivariance was used because it is suitable for molecular and crystalline structure generation[36,37,66,67]. In EGNN[66], the message-passing mechanism learns message vectors that encode interactions between neighboring nodes and aggregates them to update each node's coordinates $\boldsymbol{x}$ and features $\boldsymbol{h}$. In our implementation, we concatenate the spectral vector $\boldsymbol{s}$ to the node features $\boldsymbol{h}$, so that the learned messages—and consequently the updates of $\boldsymbol{x}$ and $\boldsymbol{h}$—are conditioned on the spectral information. The EGNN consisted of five layers. In each layer, the neural networks used to update the coordinate vector $\boldsymbol{x}$, update the node features $\boldsymbol{h}$, and output the message vector $\boldsymbol{m}$ were each implemented as a three-layer MLP, with a hidden-layer dimension of 1024 in all cases. The message vector was set to 256 dimensions.

Training and generation processes

For the forward diffusion process, following Hoogeboom et al. (2022), we defined the diffusion according to:

$$x_t = \alpha_t x_0 + \sigma_t \epsilon,$$

$$\alpha_t = (1-2s)\left(1-\left(\frac{t}{T}\right)^2\right)^p + s,$$

$$\sigma_t = \sqrt{1-\alpha_t^2}.$$

Here, we set $s = 1.0\times10^{-5}$ and $p = 2$. T denotes the total number of timesteps in the diffusion process, and we used $T = 1000$. The above diffusion process was applied to the one-hot-encoded element-type vectors and the Cartesian coordinates. For both the element-type vectors and coordinate vectors, the prediction target of the GNN was ϵ, and the mean squared error (MSE) was used as the loss function. Training was performed for 3000 epochs with batch size 1. We used AdamW as the optimizer with a learning rate of 1.0×10^{-4} and a weight decay of 1.0×10^{-12}. During training, we updated the model parameters θ_t, and at inference time we used the exponential moving average (EMA) parameters θ_t^{EMA}. The EMA parameters were computed as:

$$\theta_t^{\mathrm{EMA}} = \beta\theta_{t-1}^{\mathrm{EMA}} + (1-\beta)\theta_t,$$

with $\beta = 0.999$. For classifier-free guidance, during training, the spectral data were replaced with zero vectors of the same size with a probability of 10%. During inference, ϵ_{cfg} was predicted according to the following equation:

$$\epsilon_{\mathrm{cfg}} = (\gamma+1)\times\epsilon_{\mathrm{cond}} - \gamma\times\epsilon_{\mathrm{uncond}}$$

Here, ϵ_{cond} and $\epsilon_{\mathrm{uncond}}$ denote the noise predictions conditioned on the spectral vector and on the zero vector, respectively. γ is the classifier-free guidance factor, and we set $\gamma = 1$. The element-type vectors reverse-diffused to timestep 0 were decoded into one-hot vectors by setting the dimension with the maximum value to 1.

Evaluation of coordination environments

Next, we describe the evaluation methodology for the generated coordination environments. The generated coordinates were assessed using Root Mean Square Distance. The RMSD was calculated according to the following equation:

$$RMSD = \sqrt{\frac{1}{N}\sum_i^N |\boldsymbol{r}_i^{target} - \boldsymbol{r}_i^{generated}|^2}$$

where N is the number of atoms in the structure, $\mathrm{RMSD} = \sqrt{\frac{1}{\mathrm{N}}\sum_{\mathrm{i}}^{\mathrm{N}} |\boldsymbol{r}_{\mathrm{i}}^{\mathrm{target}} - \boldsymbol{r}_{\mathrm{i}}^{\mathrm{generated}}|^2}$ and $\boldsymbol{r}_i^{\mathrm{generated}}$ are the Cartesian coordinates of the atomic site with index i in the target and generated structures, respectively. Specifically, the first shell consists of atoms directly bonded to the excited oxygen atom, the second shell consists of atoms directly bonded to the atoms in the first shell, and the third shell consists of atoms directly bonded to the atoms in the second shell. Here, two atoms were considered to be bonded when their interatomic distance was within the cutoff radius of 2.0 Å. Because the model architecture is an equivariant graph neural network, the orientation of the generated graph and the ordering of node indices are not uniquely determined. Therefore, prior to coordinate evaluation, we perform alignment between the target structure and the generated structure. For alignment, we use the Kabsch algorithm[68,69], which computes the rotation matrix that minimizes the RMSD between two point sets when the corresponding point pairs are known. However, because the index order in the generated graph is not fixed, the correspondence between points is unknown. Therefore, we performed alignment using an iterative closest point-like procedure based on the Hungarian and Kabsch algorithms.

The detailed procedure is as follows. First, the centroids of both structures were translated to the origin. A rotation matrix generated randomly, with the identity matrix used in the first iteration, was then applied to the generated structure to rotate it. Subsequently, the correspondence between atomic pairs in the target and generated structures was determined using the Hungarian algorithm, and the Kabsch algorithm was applied according to this correspondence. This procedure yields the rotation matrix R and translation vector t that minimize the RMSD. The generated structure was then aligned using the obtained R and t, after which the assignment of

atomic pairs by the Hungarian algorithm and the alignment by the Kabsch algorithm were repeated. After the procedure was repeated for the predefined number of min_iter iterations, convergence was determined when the relative improvement in RMSD between successive iterations fell below the tolerance threshold. This procedure was repeated for the specified number of restarts, and the alignment with the smallest final RMSD among all restarts was adopted as the optimal alignment. The specific parameter values were set to min_iter = 10 and restarts = 32.

The average bond length between the excited oxygen atom and Si atoms around it, as well as the average bond angle involving the excited oxygen atom and two coordinating Si atoms, were also used for evaluation. For the target structure, the average Si–O bond length and Si–O–Si bond angle around the excited oxygen atom were calculated using each Si atom located within the cutoff distance of 2.0 Å from the excited oxygen atom. For the generated structure, the correspondence of site indices between the target and generated structures obtained from the alignment procedure described above was used. Specifically, the atoms in the generated structure corresponding to those used for calculating the averages in the target structure were extracted, and the average bond length and average bond angle were calculated in the same manner. The differences in the average bond length and average bond angle between the target and generated structures were evaluated using the mean absolute error.

The coordination number (CN) of the first shell was defined as the number of nearest-neighbor atoms around the excited oxygen atom, whereas the CN of the second shell was defined as the total number of atoms in the first and second nearest-neighbor shells. The coordination number error was calculated according to the following equation, based on the definition of relative uncertainty commonly used to evaluate the uncertainty in coordination-number estimation in EXAFS analysis:

$$\text{CN error} = \frac{|\text{CN}_{\text{target}} - \text{CN}_{\text{generated}}|}{\text{CN}_{\text{target}}}$$

For structure inference from experimental spectra, the atomic radial distribution g(R) centered on the excited oxygen atom was calculated according to the following equation:

$$g(R) = \frac{dn(R)}{4\pi R^2 \rho dR}$$

Here, R denotes the distance from the excited oxygen atom. In addition, $dn(R)$ represents the number of atoms located within the range from R to $R + dR$, where $dR = 0.01$ Å. Considering that ELNES reflects the local environment around the excited atom, ρ was defined as the atomic number density within a radius of 5 Å centered on the excited oxygen atom. The obtained atomic radial distribution was smoothed using one-dimensional Gaussian smoothing with $\sigma = 0.05$ Å and used as the radial distribution function (RDF).

Comparison with template matching

For template matching, the training and validation datasets were used as the reference database. The similarity between ELNES/XANES spectra was evaluated using cosine similarity. Cases in which spectra derived from the same crystal structure were matched were excluded.

The target structure, the structure generated by the model, and the structure predicted by template matching were each represented as vectors using SOAP descriptors. The structural similarity between the target and model-generated structures, $S_{target-model}$, and that between the target and template-matching-predicted structures, $S_{\text{target}-\text{template}}$, were quantified using cosine similarity and compared to evaluate the accuracy of the present method relative to template matching. In the comparison plots with template matching shown in Figs. 4 and 7, the vertical axis represents the difference score obtained by subtracting $S_{target-template}$ from $S_{\text{target}-\text{model}}$.

The horizontal axis represents one minus the similarity between the matched spectrum and the target spectrum in template matching.

For the SOAP descriptor parameters, we used species = [Si, O], r_cut = 7, n_max = 15, l_max = 10, and sigma = 0.5.

Data availability

The structural data used for constructing the dataset and for spectrum calculations were obtained from Materials Project (http://www.materialsproject.org)[58,59]. The dataset of calculated O-*K* spectra is available upon request to corresponding authors.

Code availability

The code used in this study is publicly available in the authors' GitHub repository (https://github.com/Ren-Okubo/ELNESDiff).

Received: 16 February 2026; Accepted: 15 October 2026;
Published online: 12 November 2026

References

1. Lu, Z. Computational discovery of energy materials in the era of big data and machine learning: a critical review. *Mater. Rep. Energy* **1**, 100047 (2021).
2. Tehrani, M. Advanced electrical conductors: an overview and prospects of metal nanocomposite and nanocarbon based conductors. *Phys. Status Solidi* **218**, 2000704 (2021).
3. Francis, A. Biological evaluation of preceramic organosilicon polymers for various healthcare and biomedical engineering applications: a review. *J. Biomed. Mater. Res. B Appl. Biomater.* **109**, 744–764 (2021).
4. Wang, C. et al. Three-dimensional atomic structure of grain boundaries resolved by atomic-resolution electron tomography. *Matter* **3**, 1999–2011 (2020).
5. Midgley, P. A. & Dunin-Borkowski, R. E. Electron tomography and holography in materials science. *Nat. Mater.* **8**, 271–280 (2009).
6. Newville, M. Fundamentals of XAFS. *Rev. Miner. Geochem.* **78**, 33–74 (2014).
7. Rehr, J. J. & Albers, R. C. Theoretical approaches to X-ray absorption fine structure. *Rev. Mod. Phys.* **72**, 621–654 (2000).
8. de Groot, F. High-resolution X-ray emission and X-ray absorption spectroscopy. *Chem. Rev.* **101**, 1779–1808 (2001).
9. Mizoguchi, T. A review of simulation, characterization, and informatics in inorganic materials. 133, 713–734. https://doi.org/10.2109/jcersj2.25121 (2025).
10. Tatsumi, K. et al. Distribution of solute atoms inβ- and spinelSi6−zAlzOzN8−zby AlK-edge x-ray absorption near-edge structure. *Phys. Rev. B Condens. Matter Mater. Phys.* **71**, 033202 (2005).
11. Lawniczak-Jablonska, K., Libera, J. & Iwanowski, R. J. EXAFS determination of local atomic structure of selected transition metals in CdSe matrix. *J. Alloy. Compd.* **286**, 89–92 (1999).
12. Pettifer, R. F., Mathon, O., Pascarelli, S., Cooke, M. D. & Gibbs, M. R. J. Measurement of femtometre-scale atomic displacements by X-ray absorption spectroscopy. *Nature* **435**, 78–81 (2005).
13. Levina, A., Armstrong, R. S. & Lay, P. A. Three-dimensional structure determination using multiple-scattering analysis of XAFS: applications to metalloproteins and coordination chemistry. *Coord. Chem. Rev.* **249**, 141–160 (2005).
14. Timoshenko, J. & Frenkel, A. I. Probing structural relaxation in nanosized catalysts by combining EXAFS and reverse Monte Carlo methods. *Catal. Today* **280**, 274–282 (2017).
15. Iesari, F., Di Cicco, A. & Okajima, T. Effect of data quality on results of Reverse Monte Carlo analysis of EXAFS data. *Radiat. Phys. Chem.* **215**, 111352 (2024).
16. Wei, X. et al. Enhancing the reliability of Reverse Monte Carlo simulations of metallic glass structure by imposing strict constraints from partial pair correlation functions. *Comput. Mater. Sci.* **244**, 113169 (2024).

17. Mierzwa, B. EXAFS as a tool for studies of bimetallic PdCo nanocluster structure. *J. Alloy. Compd.* **362**, 178–188 (2004).
18. Penner-Hahn, J. E. Characterization of "spectroscopically quiet" metals in biology. *Coord. Chem. Rev.* **249**, 161–177 (2005).
19. Ikeno, H. & Mizoguchi, T. Basics and applications of ELNES calculations. *Microscopy* **66**, 305–327 (2017).
20. Chen, P.-Y., Shibata, K., Hagita, K., Miyata, T. & Mizoguchi, T. Prediction of the ground-state electronic structure from core-loss spectra of organic molecules by machine learning. *J. Phys. Chem. Lett.* **14**, 4858–4865 (2023).
21. Guda, A. A. et al. Understanding X-ray absorption spectra by means of descriptors and machine learning algorithms. *Npj Comput. Mater.* **7**, 203 (2021).
22. Li, Z. et al. Highly sensitive 2D X-ray absorption spectroscopy via physics informed machine learning. *Npj Comput. Mater.* **10**, 128 (2024).
23. Ni, H. et al. Quantifying atomically dispersed catalysts using deep learning assisted microscopy. *Nano Lett.* **23**, 7442–7448 (2023).
24. Mousavi, M. S. S., Pofelski, A., Teimoori, H. & Botton, G. A. Alignment-invariant signal reality reconstruction in hyperspectral imaging using a deep convolutional neural network architecture. *Sci. Rep.* **12**, 17462 (2022).
25. Mizoguchi, T. Data-driven ELNES/XANES analysis: predicting spectra, unveiling structures and quantifying properties. *Microscopy* dfaf038 (2025).
26. Na Narong, T., Zachko, Z. N., Torrisi, S. B. & Billinge, S. J. L. Interpretable multimodal machine learning analysis of X-ray absorption near-edge spectra and pair distribution functions. *Npj Comput. Mater.* **11**, 98 (2025).
27. Carbone, M. R., Yoo, S., Topsakal, M. & Lu, D. Classification of local chemical environments from X-ray absorption spectra using supervised machine learning. *Phys. Rev. Mater*. **3**, 033604 (2019).
28. Torrisi, S. B. et al. Random forest machine learning models for interpretable X-ray absorption near-edge structure spectrum-property relationships. *npj Comput. Mater.* **6**, 109 (2020).
29. Zheng, C., Chen, C., Chen, Y. & Ong, S. P. Random forest models for accurate identification of coordination environments from X-ray absorption near-edge structure. *Patterns* **1**, 100013 (2020).
30. Kiyohara, S., Tsubaki, M. & Mizoguchi, T. Learning excited states from ground states by using an artificial neural network. *Npj Comput. Mater.* **6**, 68 (2020).
31. Mizoguchi, T. & Kiyohara, S. Machine learning approaches for ELNES/XANES. *Microscopy* **69**, 92–109 (2020).
32. Gleason, S. P., Lu, D. & Ciston, J. Prediction of the Cu oxidation state from EELS and XAS spectra using supervised machine learning. *Npj Comput. Mater.* **10**, 221 (2024).
33. Lee, B. et al. Predictive autoencoder-transformer model of Cu oxidation state from EELS and XAS spectra. Preprint at *arXiv* https://doi.org/10.48550/arXiv.2601.11509 (2026).
34. Takahara, I., Uesugi, F., Kimoto, K., Shibata, K. & Mizoguchi, T. Toward the atomic-level analysis of ground-state electronic structures of solid materials via prediction from core-loss spectra: a computational study in Si. *J. Phys. Chem. C Nanomater. Interfaces* **128**, 13500–13507 (2024).
35. Ho, J., Jain, A. & Abbeel, P. Denoising diffusion probabilistic models. Preprint at *arXiv* https://doi.org/10.48550/arXiv.2006.11239 (2020).
36. Xu, M. et al. GeoDiff: a geometric diffusion model for molecular conformation generation. Preprint at *arXiv* https://doi.org/10.48550/arXiv.2203.02923 (2022).
37. Hoogeboom, E., Satorras, V. G., Vignac, C. & Welling, M. Equivariant diffusion for molecule generation in 3D. Preprint at *arXiv* https://doi.org/10.48550/arXiv.2203.17003 (2022).
38. Alverson, M. et al. Generative adversarial networks and diffusion models in material discovery. *Digit. Discov.* **3**, 62–80 (2024).
39. Kwon, H. et al. Spectroscopy-guided discovery of three-dimensional structures of disordered materials with diffusion models. *Mach. Learn. Sci. Technol.* **5**, 045037 (2024).
40. McLean, B. & Yarovsky, I. Structure, properties, and applications of silica nanoparticles: Recent theoretical modeling advances, challenges, and future directions. *Small* **20**, e2405299 (2024).
41. Deng, W. et al. Recent advances of high performance SiO (0<x<2) anodes. *Mater. Rep. Energy* **4**, 100270 (2024).
42. Ferry, D. K. & Rode, D. L. Physical and electrical properties of silica. *Appl. Phys. Rev*. **12**, 011304 (2025).
43. Ravel, B. & Kelly, S. D. The difficult chore of measuring coordination by EXAFS. In *Proc. AIP Conference*. https://doi.org/10.1063/1.2644458 (AIP, 2007).
44. Kathyola, T. A., Willneff, E. A., Willis, C. J., Dowding, P. J. & Schroeder, S. L. M. Reactive CaCO3 formation from CO2 and methanolic Ca(OH)2 dispersions: transient methoxide salts, carbonate esters and sol-gels. *ACS Phys. Chem. Au* **4**, 555–567 (2024).
45. Bartók, A. P., Kondor, R. & Csányi, G. On representing chemical environments. Preprint at *arXiv* https://doi.org/10.48550/arXiv.1209.3140 (2012).
46. Zhu, W., Zheng, G., Cao, S. & He, H. Thermal conductivity of amorphous SiO2 thin film: A molecular dynamics study. *Sci. Rep.* **8**, 10537 (2018).
47. Tersoff, J. Modeling solid-state chemistry: interatomic potentials for multicomponent systems. *Phys. Rev. B Condens. Matter* **39**, 5566–5568 (1989).
48. Tersoff, J. Empirical interatomic potential for silicon with improved elastic properties. *Phys. Rev. B Condens. Matter* **38**, 9902–9905 (1988).
49. Fujikata, Y., Sugisawa, H. & Mizoguchi, T. Machine learning-based XANES analysis for predicting the local structure and valence in amorphous silicon suboxides. *Phys. Chem. Chem. Phys.* https://doi.org/10.1039/d5cp03408f (2026).
50. McInnes, L., Healy, J. & Melville, J. UMAP: uniform manifold approximation and projection for dimension reduction. Preprint at *arXiv* https://doi.org/10.48550/arXiv.1802.03426 (2018).
51. Liao, K. et al. Revealing spatial distribution of Al-coordinated species in a phase-separated aluminosilicate glass by STEM-EELS. *J. Phys. Chem. Lett.* **11**, 9637–9642 (2020).
52. Kiyohara, S. & Mizoguchi, T. Radial distribution function from X-ray absorption near edge structure with an artificial neural network. *J. Phys. Soc. Jpn.* **89**, 103001 (2020).
53. Nishiyama, N. et al. Synthesis of nanocrystalline bulk SiO2 stishovite with very high toughness. *Scr. Mater.* **67**, 955–958 (2012).
54. Takaichi, G. et al. Limited stability of hydrous SiO2 stishovite in the deep mantle. *Earth Planet. Sci. Lett.* **640**, 118790 (2024).
55. Nishiyama, N. et al. Fracture-induced amorphization of polycrystalline SiO2 stishovite: a potential platform for toughening in ceramics. *Sci. Rep.* **4**, 6558 (2014).
56. Suenaga, K. & Koshino, M. Atom-by-atom spectroscopy at graphene edge. *Nature* **468**, 1088–1090 (2010).
57. Kim, M.-S. et al. Charge disproportionation at twisted SrTiO3 bilayer interface driven by local atomic registry. *ACS Nano* **19**, 39714–39724 (2025).
58. Jain, A. et al. Commentary: The Materials Project: a materials genome approach to accelerating materials innovation. *APL Mater.* **1**, 011002 (2013).
59. Horton, M. K. et al. Accelerated data-driven materials science with the Materials Project. *Nat. Mater.* **24**, 1522–1532 (2025).
60. Griscom, D. L. Optical properties and structure of defects in silica glass. *J. Ceram. Soc. Jpn.* **99**, 923–942 (1991).
61. Clark, S. J. et al. First principles methods using CASTEP. *Z. Kristallogr. Cryst. Mater.* **220**, 567–570 (2005).
62. Morris, A. J., Nicholls, R. J., Pickard, C. J. & Yates, J. R. OptaDOS: a tool for obtaining density of states, core-level and optical spectra from

electronic structure codes. *Comput. Phys. Commun.* **185**, 1477–1485 (2014).
63. Monkhorst, H. J. & Pack, J. D. Special points for Brillouin-zone integrations. *Phys. Rev. B* **13**, 5188–5192 (1976).
64. Brown, G. E. Jr, Waychunas, G. A., Stohr, J. & Sette, F. Near-edge structure of oxygen in inorganic oxides : effect of local geometry and cation type. *J. Phys., Colloq.* **47**, C8–685 (1986).
65. Sharp, T. et al. Distinction between six- and fourfold coordinated silicon in SiO2 polymorphs via electron loss near edge structure (ELNES) spectroscopy. *Phys. Chem. Miner.* **23**, 17–24 (1996).
66. Satorras, V. G., Hoogeboom, E. & Welling, M. E(n) Equivariant graph neural networks. Preprint at *arXiv* https://doi.org/10.48550/arXiv.2102.09844, 9323–9332 (2021).
67. Wang, J. et al. E(n)-Equivariant Cartesian tensor message passing interatomic potential. *Nat. Commun.* **15**, 7607 (2024).
68. Kabsch, W. A solution for the best rotation to relate two sets of vectors. *Acta Crystallogr. A* **32**, 922–923 (1976).
69. Kabsch, W. A discussion of the solution for the best rotation to relate two sets of vectors. *Acta Crystallogr. A* **34**, 827–828 (1978).

Acknowledgements

We would like to acknowledge Prof. Kiyou Shibata, Nagoya University for his helpful discussion. This study was supported by Japan Science and Technology Agency (JST) (Nos. JPMJAX24DB and JPMJBS2418), the Ministry of Education, Culture, Sports, Science and Technology (MEXT) (Nos. 25K24645, 26K22608 and 26K01205), and New Energy and Industrial Technology Development Organization (NEDO), Japan. Some computations were carried out using the computer resource offered by Research Institute for Information Technology, Kyushu University.

Author contributions

R.O., I.T., and T.M. designed the research. R.O. performed the calculations and analyzed the data. Y.F. calculated the amorphous Si–O structures, and R.O., I.T., Y.F., and T.M. wrote the manuscript.

Funding

The authors received no specific funding for this work.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41524-026-02384-4.

Correspondence and requests for materials should be addressed to Teruyasu Mizoguchi.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2026

This Article in Press is shared early to give you faster access to new research. It is citable and carries a permanent DOI. The final edited version will replace it automatically.