

# Spectroscopy-Guided Discovery of Three-Dimensional Structures of Disordered Materials with Diffusion Models

Hyuna Kwon<sup>\*1</sup>, Tim Hsu<sup>†1</sup>, Wenyu Sun<sup>1</sup>, Wonseok Jeong<sup>1</sup>, Fikret Aydin<sup>1</sup>, James Chapman<sup>2</sup>,  
Xiao Chen<sup>1</sup>, Matthew R. Carbone<sup>3</sup>, Deyu Lu<sup>4</sup>, Fei Zhou<sup>1</sup>, and Tuan Anh Pham<sup>‡1,5</sup>

<sup>1</sup>Lawrence Livermore National Laboratory, Livermore, CA 94551, USA

<sup>2</sup>Department of Mechanical Engineering, Boston University, Boston, MA 02215, USA

<sup>3</sup>Computational Science Initiative, Brookhaven National Laboratory, Upton, NY 11973, USA

<sup>4</sup>Center for Functional Nanomaterials, Brookhaven National Laboratory, Upton, NY 11973,  
USA

<sup>5</sup>Laboratory for Energy Applications for the Future, Lawrence Livermore National  
Laboratory, Livermore, CA 94551, USA

## Abstract

The ability to rapidly develop materials with desired properties has a transformative impact on a broad range of emerging technologies. In this work, we introduce a new framework based on the diffusion model, a recent generative machine learning method to predict 3D structures of disordered materials from a target property. For demonstration, we apply the model to identify the atomic structures of amorphous carbons (*a*-C) as a representative material system from the target X-ray absorption near edge structure (XANES) spectra—a common experimental technique to probe atomic structures of materials. We show that conditional generation guided by XANES spectra reproduces key features of the target structures. Furthermore, we show that our model can steer the generative process to tailor atomic arrangements for a specific XANES spectrum. Finally, our generative model exhibits a remarkable scale-agnostic property, thereby enabling generation of realistic, large-scale structures through learning from a small-scale dataset (i.e., with small unit cells). Our work represents a significant stride in bridging the gap between materials characterization and atomic structure determination; in addition, it can be leveraged for materials discovery in exploring various material properties as targeted.

## 1 Introduction

An accelerated materials discovery pipeline with desired properties holds a transformative impact on a broad range of emerging technologies, from energy storage and conversion to ion-selective membranes [1–3]. Traditional approaches for materials discovery involve a time-consuming trial-and-error process that begins with a set of materials from which properties of interest are derived. The understanding gained from this procedure is then used to guide materials development by exploring a new configuration space. Recent advancement in data science approaches provides an alternative and promising means to accelerate materials development [4]. For example, high-throughput materials screening and data mining have led to several breakthroughs in materials development for a broad range of technologies in the past two decades [5, 6]. This approach often relies on simulations techniques, such as first-principles calculations, to rapidly screen a vast space of chemical compositions and structures in order to identify materials with desired properties.

Materials development through inverse design departs from the traditional methods. Instead of deriving properties from structures, target material properties are specified beforehand and materials satisfying these

---

<sup>\*</sup>kwon11@lbl.gov, equal contribution

<sup>†</sup>hsu16@lbl.gov, equal contribution

<sup>‡</sup>pham16@lbl.gov

requirements are inferred [7]. This strategy, however, is generally more challenging. For example, many material systems exhibit multi-to-one mapping in the structure-property space, i.e., multiple distinct structures can result in roughly the same property. It is therefore challenging to learn the inverse mapping even if the forward mapping is fully known. This problem has been the topic of active research in recent years with data-driven approaches, thanks to emerging capabilities of modern machine learning methods [8–11]. A viable direction is to formulate the inverse prediction problem in a probabilistic framework, for instance with a conditional generative model, that generates or samples multiple candidate material structures conditioned on a given property.

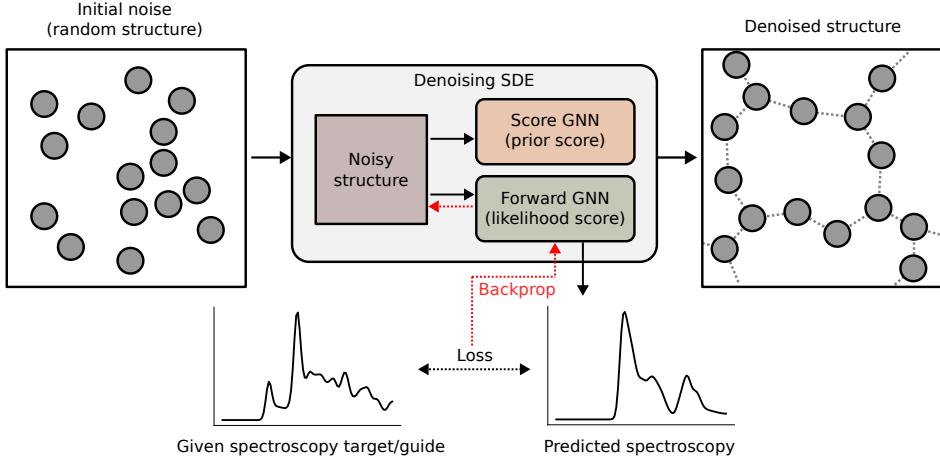
Among the different classes of generative models, the diffusion model [12–14] is particularly well suited for generating structures involving precise three-dimensional (3D) atomic coordinates. The diffusion formulation centers around the idea of denoising perturbed or noisy prior inputs towards realistic distributions based on a learned score function [15, 16] that has the form of a relative distance vector for atomic coordinate data. In addition, when the data space is defined in terms of relative distance vectors between atoms, the score model does not rely on absolute coordinates and is therefore translation-invariant. While there have been numerous works on diffusion model for molecules [17–21], relatively few studies have focused on disordered systems [22, 23]. This family of materials is highly relevant in many emerging applications, such as energy conversion and storage devices, where the materials are often driven away from ideal structures under external driving forces [24, 25]. They can therefore be more complex, often exhibiting richer structural and chemical heterogeneities.

In this study, we demonstrate the application of the diffusion model to identify 3D structures of disordered solid materials from a target property, which in this work is chosen to be spectroscopy to address an outstanding challenge in materials characterization: predicting 3D structures of disordered materials from spectroscopic measurements. We focus on XANES spectroscopy that is widely used to probe local atomic structures, and demonstrate with amorphous carbons (*a*-C) as a representative system. In a similar vein, Comin and Lewis [26] trained a Wasserstein autoencoder [27] to generate amorphous silicon topologies. However, such a model is limited to a fixed number of atoms whereas our diffusion model implemented with graph neural networks (GNN) is scale-agnostic, allowing for generation at any arbitrary scale. We stress that although our present study focuses on single-element *a*-C and XANES spectroscopy, the conditional diffusion model framework can be extended to broader scenarios. For instance, the scale-agnostic property of the diffusion model allows for the generation of large-scale atomic structures based on information learned from a set of diverse small-scale data. The framework can also be naturally extended to multi-element material systems, where the atom type information is also subject to denoising. Finally, multi-target inverse prediction is also straightforward to implement to enable material inverse design for various target properties, as the conditional score function is a simple summation of the prior and likelihood scores from multiple forward models.

Our conditional diffusion model framework shown in Fig. 1 follows a standard structure consisting of a prior score model and a forward model. The former outputs the *prior/unconditional score*, which is used to gradually denoise noisy priors towards realistic samples without bias, while the latter—which predicts per-atom spectral curves in this work—provides the *likelihood/conditioning score* that guides the denoising iterations towards a given target property. It is necessary to emphasize that the target property is not necessarily always achievable (and not all targets are realistic). In this regard, the given spectroscopy target may be more appropriately described as a guide. Importantly, the prior score model and the forward model are decoupled and can be separately trained. We note that the use of a forward model is optional, in which case the generation is simply unconditional. Further details regarding the conditional diffusion model implementation are described in the Methods section.

## 2 Results

We present our results regarding *a*-C 3D structure generation with and without XANES spectroscopy conditioning. In the unconditional case, we focus on the fidelity of the generation, and highlight that the generation is scale-agnostic. In the conditional case, the performance of a structure-to-spectroscopy forward model is briefly discussed, and the spectroscopy-guided generation is validate in two scenarios. The extent of the conditioning effect is also briefly analyzed. Overall, we demonstrate realistic, scale-agnostic generation of *a*-C structures based on spectra inputs/guides.



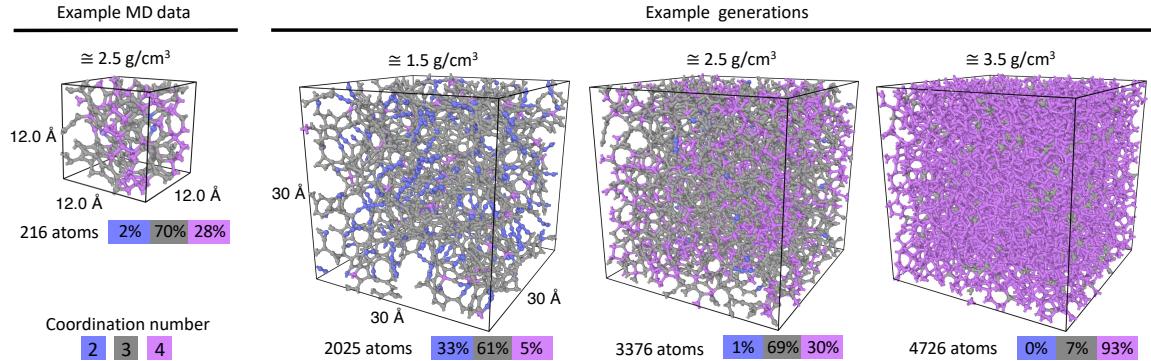
**Figure 1:** The conditional generative/denoising process of amorphous carbon (*a*-C) involves a pair of decoupled (i.e., trained separately) score model and forward model. The score model, based on the diffusion model formalism, provides the prior score, while the forward model, trained to predict spectroscopic curves, is used to provide the likelihood score. The prior score serves to denoise initial random motifs towards realistic topologies without bias, and the likelihood score serves to bias the denoising trajectory towards atomic motifs that satisfy a given spectroscopic target/guide. Notably, the forward model is optional, and the generative process is simply unconditional without a forward model.

## 2.1 Unconditional generation

The fidelity of the structure generation is first examined without considering the conditioning effect. We emphasize that our generation model is scale-agnostic, i.e., the model can learn from small-scale molecular dynamics (MD) trajectories and consequently generate similar configurations at any arbitrary scale. This is enabled by the node-centric (local) predictions of the underlying graph neural network architecture. The denoising process therefore amounts to shifting each atom’s coordinate towards realistic topologies based on its local environment without the need to observe global variables. This scale-agnostic feature is particularly useful for generating large domains encompassing diverse topologies, which can be costly to sample via MD simulations. In this regard, the large-scale generated structures can be interpreted as having “stitched” or combined the small but diverse MD-simulated motifs. Such a generation is important for examining material properties where a large supercell with diverse local structures is required for yielding sufficient statistics. Examples include predicting hydrogen diffusion [28], gas selectivity [29], interfacial interactions [30], and graphitization [31] in amorphous systems and disordered interfaces. We intend to investigate scaled generations as part of future effort.

The scale-agnostic feature is demonstrated in Fig. 2, where the generated *a*-C configurations have a supercell 2.5 times larger than that of the original training data. As expected, the generated *a*-C topology highly depends on the density (or the number of atoms per unit volume) at which the structure is generated. Specifically, *a*-C motif tends to be *sp*- or *sp*<sup>2</sup>-hybridized at low densities, and *sp*<sup>2</sup>- or *sp*<sup>3</sup>-hybridized at high densities. In fact, the generation is always (implicitly) conditioned on the density (input cell size and number of particles within) even without a spectral target. For simplicity, the term *unconditional* used throughout this paper refers to generation without a spectral target. We note that the coordination number (CN) is used to estimate the hybridization. Specifically, we define two atoms as bonded if their distance is less than a cutoff of 1.84 Å, which is determined based on the C-C radial distribution function (RDF) of the structures in the training set.

The unconditionally generated structures are realistic (Fig. 3), adhering to the training data distributions in terms of the various structural and thermodynamic properties considered in this work: Steinhardt order parameters [32], bond angles, RDFs, Voronoi cell volumes, and per-atom potential energies [33, 34]. The Steinhardt parameter in particular stands out for its detailed insights into the local symmetry of atomic arrangements, complementing other descriptors like RDF and bond angle distribution. First, as shown in Fig. 3a, the scatter plot of the Steinhardt features ( $\bar{q}_4$ ,  $\bar{q}_6$ ) indicates that the scatter points from the generated *a*-C structures roughly match the distribution of the training data, but importantly they do not coincide. This indicates that



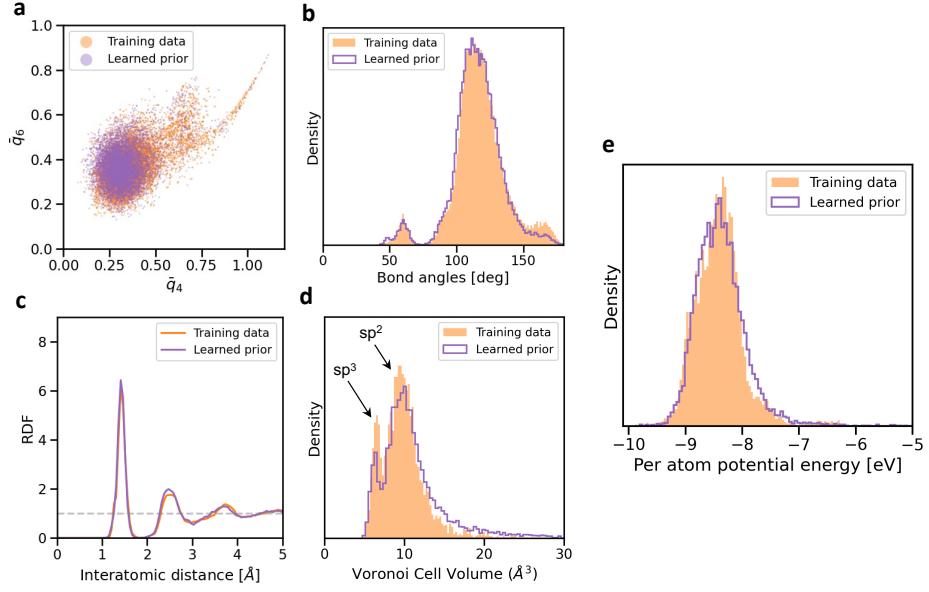
**Figure 2:** The generation of the  $a$ -C structures is scale-agnostic, and the generated topologies highly depend on density (as expected). While the leftmost panel displays a representative training data point with a density of  $\approx 2.5 \text{ g/cm}^3$ , it is essential to note that the subsequent larger cells on the right are not exclusive products of this single  $2.5 \text{ g/cm}^3$  cell. The generation can be at any arbitrary scale because the denoising process amounts to shifting each atom's location towards realistic topology based on its local environment. Consistent with the MD training data, the  $a$ -C topologies generated at low density typically have a small coordination number (CN) of 2 and 3, and at high density the CN is typically 3 and 4.

the generated topologies are similar to, but not exact copies of, the training atom motifs. Second, the bond angle, RDF, and Voronoi cell volume distributions further validate that the generated configurations are realistic, resembling the configurations of the training set (Fig. 3b-d). Finally, Fig. 3e shows that the per-atom potential energy distribution obtained from the carbon Gaussian Approximation Potential (GAP) [34] of the generation largely matches that of the training data. This observation suggests that the generated structures are not only geometrically realistic, but also energetically realistic. Importantly, the training dataset consists of 30 snapshots (of various cell sizes) of MD-simulated  $a$ -C, totaling 7,463 carbons. Overall, the  $a$ -C generation is realistic, exhibiting quantities highly similar, but not identical, to that of the MD data. It is viable to subject the generation to downstream atomistic simulations thanks to the generated atoms having realistic energies.

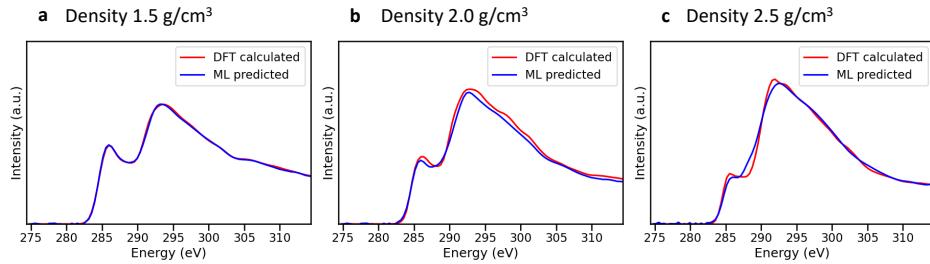
## 2.2 Conditional generation

Since the conditional generation formulation requires a forward prediction model, we trained a forward model to predict XANES spectra of  $a$ -C configurations from their atomic structures (Supplemental Fig. 1). Illustrated in Fig. 4, this model is sufficiently accurate on three validation  $a$ -C structures that yield distinctively different XANES spectra (at a density of 1.5, 2.5, and  $3.5 \text{ g/cm}^3$ , respectively). Both the forward model and the unconditional score model are based on the same GNN architecture, and further details regarding the forward model are described in the Methods section. We point out that in our prior work [35] we trained a forward model based on a multi-layer perceptron (MLP) architecture with Many Body Tensor Representation (MBTR) descriptors as the input, and the current GNN exhibits a marginal yet discernible performance improvement over the MLP-based model.

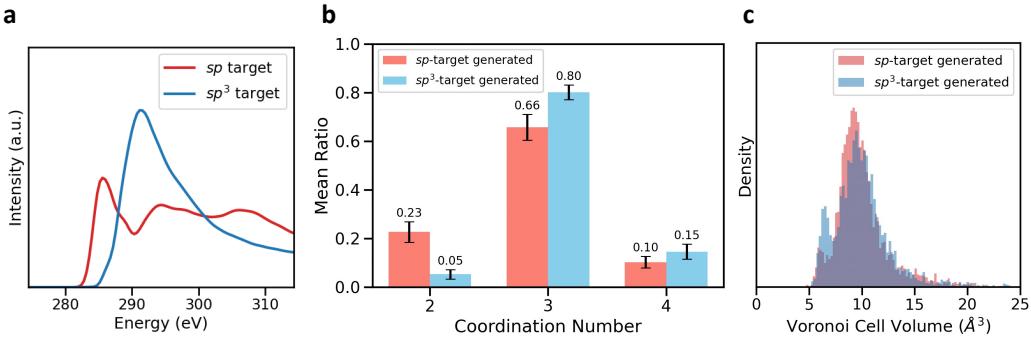
We now validate the guided generation result of our conditional diffusion model. As previously stated, even without a given spectroscopy target, the  $a$ -C generation outcome depends on the user-specified density, to the extent that  $sp/sp^2$  carbons dominate at low densities, and  $sp^2/sp^3$  carbons dominate at high densities. In other words, density implicitly imposes a strong conditioning effect on the structural generation. Therefore, to investigate the conditioning effect of the spectroscopy guide alone, we provided two distinctively different spectral inputs while fixing the density ( $2.0 \text{ g/cm}^3$ ). One of the spectral guides is an averaged spectral curve from all  $sp$  carbons, and the other from all  $sp^3$  carbons in the training set. Effectively, these spectral inputs serve to guide the generation towards  $sp$  or  $sp^3$  topologies in structures that would otherwise be dominated by  $sp^2$  carbons at  $2.0 \text{ g/cm}^3$ . As shown in Fig. 5, the difference in the two spectral inputs/targets certainly impacts the generation outcomes. Namely, generation guided by the  $sp$  target expectedly results in a higher portion of two-fold coordination carbons, while the  $sp^3$  target raises the portion of those with a four-fold coordination.



**Figure 3:** The generated prior amorphous topologies are realistic based on comparisons of (a) Steinhardt parameters [32], (b) bond angle distribution, (c) radial distribution, (d) Voronoi cell volume distribution, and (e) per-atom GAP potential energy distribution [33, 34] against the training data. The generated structures consist of 29 cubic cells ( $14^3 \text{ \AA}$ ) of  $a\text{-C}$ , totaling 7,256 atoms.



**Figure 4:** The forward XANES spectroscopy regression model is sufficiently accurate on unseen validation  $a\text{-C}$  configurations. Note that the forward model predicts per-atom spectral curves, which are then averaged to give rise to global spectra that are compared to DFT calculations. The comparison is shown for a validation set of three structures at densities (a) 1.5, (b) 2.0, and (c) 2.5  $\text{g/cm}^3$ .

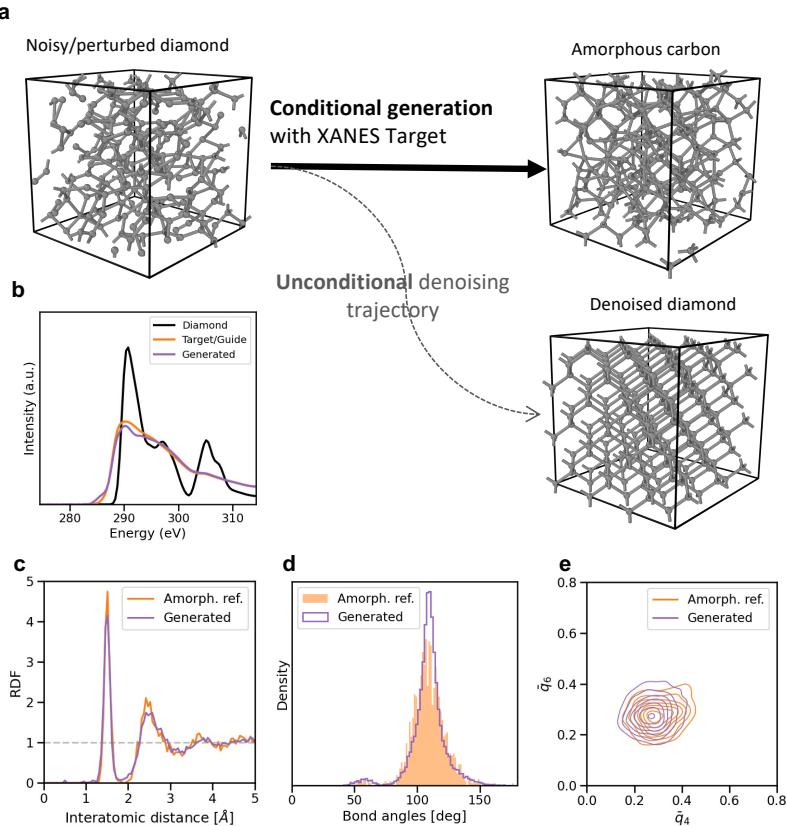


**Figure 5:** Demonstration of guided generation at the density of  $2.0 \text{ g/cm}^3$  based on two distinctively different spectral guides/targets. As shown in the figure, a difference in (a) target XANES spectra certainly impacts the (b) coordination numbers and (c) Voronoi cell volume distribution of the generated *a*-C structures. Note that for each spectral input, 16 structure generations were performed, totaling 3,456 atoms. The standard deviation (across different generations) in the coordination number ratio is denoted by the error bar in (b). We also emphasize that *sp*<sup>2</sup> carbons (which have a coordination number of three) are still prevalent in both scenarios due to the strong conditioning effect of the density ( $2.0 \text{ g/cm}^3$ ), at which *sp*<sup>2</sup> hybridization is the typical topology.

Further, the *sp*-guided generation resulted in a stronger signature in the Voronoi cell volume distribution for the *sp* topology [36], while the *sp*<sup>3</sup>-guided generation resulted in a stronger signature for the *sp*<sup>3</sup> topology. Overall, this demonstration suggests that we can effectively steer the generative process towards structures that are better aligned with the provided spectral input. For those interested in the precise matching of coordination number ratios, we repeated this procedure using density  $1.5 \text{ g/cm}^3$ , as illustrated in Supplemental Fig. 2.

It is crucial to point out that by design a XANES target might not be always fully achievable in the context of our conditional generative model. In this regard, it is more appropriate to consider XANES spectra as a guide for our diffusion model, where the generative process is biased towards certain topologies based on a given spectral target. As evident in Fig. 5b, the majority of the generated carbons remain *sp*<sup>2</sup> despite the “target” for *sp* and *sp*<sup>3</sup> motifs. This is not at all an indication of failure for the generative model. In fact, at  $2.0 \text{ g/cm}^3$ , the typical topology is indeed *sp*<sup>2</sup>, which the model observed during training. Therefore the model also imposes the prior knowledge about what should be realistically expected at  $2.0 \text{ g/cm}^3$ . Mathematically the generative process involves a balance between two terms: the unconditional term for achieving realistic samples and the guide term for biasing towards a target (further details in Methods). If all the generated carbons have *sp* or *sp*<sup>3</sup> topology at  $2.0 \text{ g/cm}^3$ , then the resulting structures are very likely unphysical (such structures are also never observed in the MD training dataset).

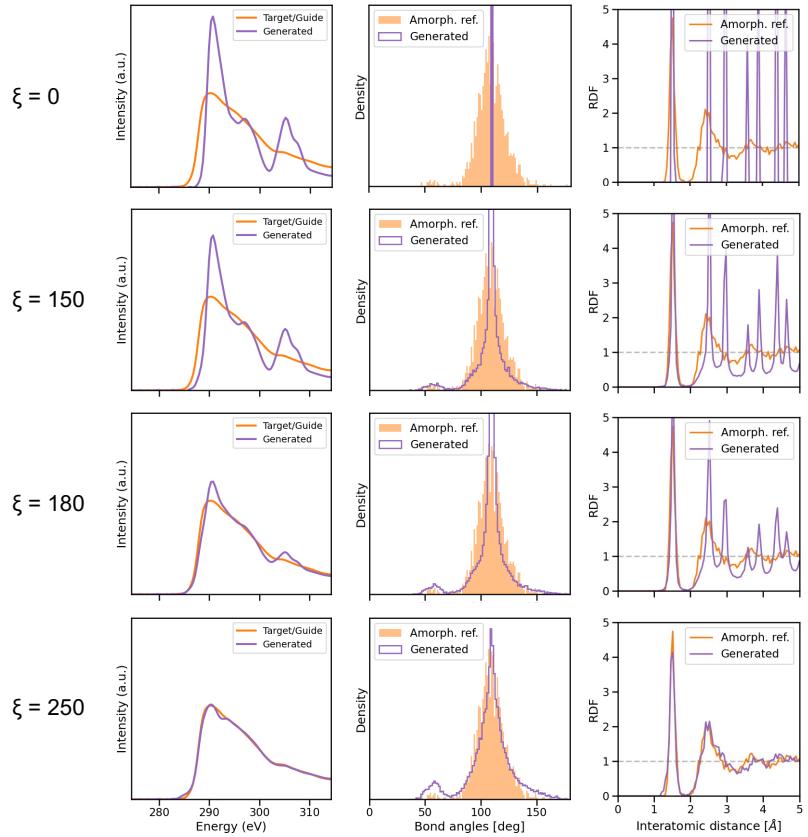
As another demonstration of guided generation, we initialized a heavily perturbed diamond structure that would by default be denoised into perfect diamond via the unconditional denoising trajectory. We note that a separate generative model was trained in this case, where the training dataset includes only the perfect diamond structure and an amorphous structure of the same density ( $3.5 \text{ g/cm}^3$ ). As shown in Fig. 6a-b, by introducing a spectral target for the amorphous structure, the guided generation resulted in amorphous topology instead of the default diamond. Further, the generated amorphous structures are realistic in terms of RDF, bond angle distributions, and Steinhardt parameters (Fig. 6c-d). The choice of demonstrating the contrast between *a*-C and diamond is to further validate the conditioning effect of spectroscopy-guided generation, particularly for generating materials that may be considered completely different classes, e.g., amorphous vs. crystalline. However, generating crystalline motifs (at arbitrary scales) is likely an entirely different application scope. While our diffusion model can denoise heavily perturbed diamond structure into a virtually noiseless state, we did not thoroughly investigate the feasibility of denoising fully random structures into crystalline materials. We leave such a task for future work. The perturbed diamond structures were initialized by applying Gaussian noise  $\epsilon \sim \mathcal{N}(0, 0.4^2)$  to the atomic coordinates (unit in Å) of diamond structures. Lastly, as a bonus example, we also demonstrate the guided generation for *a*-C vs. buckyballs (Supplemental Fig. 3). Moreover, through conditional generation, we observed the generation of structures with a diamond-like appearance from heavily noisy diamond input. The guide provided by conditional generation allows for further refinement of this structure,



**Figure 6:** Demonstration of spectroscopy-guided generation where a heavily perturbed diamond structure can result in either perfect denoised diamond topology by default (unconditionally), or amorphous topology with the use of a spectral guide. (a) Visualization of the difference in the generation outcome with and without the spectral guide. (b) Comparison between the given target spectrum and the spectrum predicted from the generated amorphous structure, along with a diamond spectrum for reference. The generated amorphous structures are geometrically realistic compared to an MD-simulated reference structure (“amorph. ref.”) in terms of (c) RDF, (d) bond angle distribution, and (e) Steinhardt parameters.

closely aligning it with the diamond structure (Supplemental Fig. 4).

Finally, we briefly analyze the extent to which guided generation adheres to a given spectral target. Mathematically the conditional generative process is based on the conditional score function, which is the summation of two terms: the unconditional score points towards realistic samples, and the likelihood score points towards samples that satisfy the given target. The relative weights of the two terms can be adjusted with a tuning parameter  $\xi$ , as shown in Equation 13 and detailed in Section 4.2. A larger  $\xi$  translates to more weight for the likelihood term, or a more strict emphasis for the target to be satisfied, while  $\xi = 0$  signifies unconditional generation. One practical application of tuning the value of  $\xi$  is to relax the target constraint such that more diverse samples can be generated. Based on the same experiment for generating *a*-C vs. diamond, the value of  $\xi$  is shown to clearly impact the extent to which the generated structure is amorphous, as illustrated in detailed in Fig. 7. As  $\xi$  increases from zero, the generation transitions from the unconditional regime, where diamond is generated by default, to the conditional regime, where *a*-Cs are generated. Notably, the predicted spectrum from the generated carbons at  $\xi = 180\text{--}250$  closely matches the given target spectrum. This analysis is another validating evidence of conditional generation, particularly as a tunable knob for flexible inverse design of materials.



**Figure 7:** Based on the experiment for generating  $a\text{-C}$  vs. diamond, the tuning parameter  $\xi$  can significantly impact the extent to which the generated structure is amorphous. Each row corresponds to a different value of  $\xi$ . The XANES spectrum (which is used as the target), the bond angle distribution, and the RDF from a reference amorphous structure are fixed in all rows. As  $\xi$  increases, the generation transitions from the unconditional regime, where diamond is generated by default, to the conditional regime, where  $a\text{-Cs}$  are generated.

### 3 Conclusions

We propose a conditional generative framework based on a diffusion model to predict 3D atomic structures from spectroscopic data. This framework can be generalized to different targets that fit to a broader field of materials discovery by inverse design. First, we show that the generative model exhibits a remarkable scale-agnostic property, enabling it to learn from small-scale MD trajectories and generate large-scale structures. This versatility is useful for scenarios where accurate prediction of material properties is contingent on the generation of sufficiently large domains with diverse topologies. Second, our conditional generation process, influenced by spectroscopy input, offers control over the composition and structure of generated molecules. Therefore, one can guide the generative process to produce specific atomic arrangements, thereby tailoring materials for desired properties and applications. Lastly, through extensive fidelity assessments, we demonstrated the fidelity of our generative model. Specifically, the prior model accurately replicates key structural features observed in training data, while our posterior model yields structures that closely align with target data, including spectral predictions and structural characteristics.

Our work represents a significant stride in bridging the gap between materials characterization and atomic structure determination. In particular, by leveraging advanced generative models, our approach offer a promising solution to address to longstanding challenges in reverse engineering atomic structures from complex spectroscopic data. Our methodology holds promise for a multitude of future endeavors. For example, beyond the current scope, it can be applied to multiple targets, including X-ray Photoelectron Spectroscopy, pair distribution function, and others, opening new avenues for predicting atomic structures from multi-modal materials characterizations. Finally, our framework can be leveraged for materials discovery, not only in the realm of spectroscopy but also in exploring various material properties as targeted. Examples include mechanical strengths of materials under extreme conditions, electrocatalytic activity of catalysts for hydrogen production, and permeability of ion-selective membranes, and so on.

## 4 Methods

The methodologies employed in this work are described below, including disordered/amorphous structure generation via unconditional diffusion model, its extension to conditional/guided diffusion model formulation with the use of a forward model, how the models are trained, their GNN architectures, and the spectroscopy simulations performed for gathering data.

### 4.1 Unconditional diffusion model

We approach the task of amorphous/disordered structure generation with a machine learned generative model, namely the diffusion model [12–14]. In what may be considered the standard formulation, the diffusion model establishes a mapping between “clean”, original data distribution  $p_0(\mathbf{x}_0)$  and a latent distribution  $p_t(\mathbf{x}_t)$  by iteratively applying (Gaussian) noises to  $\mathbf{x}_0$ , over a fictitious time quantity  $t$ , towards what is typically a unit Gaussian distribution  $\mathcal{N}(0, \mathbf{I})$ . This *forward noising* process can be written as  $p_t(\mathbf{x}_t|\mathbf{x}_0) = \mathcal{N}(\alpha_t \mathbf{x}_0, \sigma_t^2 \mathbf{I})$  and implemented as  $\mathbf{x}_t = \alpha_t \mathbf{x}_0 + \sigma_t \epsilon$ ,  $\epsilon \sim \mathcal{N}(0, \mathbf{I})$ , where  $(\alpha_t, \sigma_t)$  are noise schedule functions that dictate how noise is applied with respect to  $t$ . Typically, the span of  $t$  is set to  $(0, 1)$ , over which  $\alpha_t$  monotonically decreases from 1 to 0, and  $\sigma_t$  monotonically increases from 0 to 1. The latent distribution  $p_t(\mathbf{x}_t)$  therefore becomes a unit isotropic Gaussian at  $t = 1$ . The same forward noising process can also be described as Langevin dynamics with the following stochastic differential equation (SDE):

$$d\mathbf{x}_t = f(t)\mathbf{x}_t dt + g(t)d\mathbf{w}_t, \quad (1)$$

$$f(t) = \frac{d \log \alpha_t}{dt}, g^2(t) = \frac{d \sigma_t^2}{dt} - 2 \frac{d \log \alpha_t}{dt} \sigma_t^2, \quad (2)$$

where  $f(t)$  and  $g(t)$  are drift and diffusion coefficients related to  $(\alpha_t, \sigma_t)$ , and  $\mathbf{w}_t$  is the standard Wiener process.

To generate a realistic data point, one first samples from the unit Gaussian latent distribution  $\mathbf{x}_t \sim \mathcal{N}(0, \mathbf{I})$  at  $t = 1$  and then maps it back to the original data space at  $t = 0$  via the *reverse denoising* process. While the forward noising operation is trivial, the reverse denoising operation is not and requires a learned model. It is

described as the following SDE:

$$d\mathbf{x}_t = [f(t)\mathbf{x}_t - g^2(t)\nabla_{\mathbf{x}_t} \log p_t(\mathbf{x}_t)] dt + g(t)d\mathbf{w}_t, \quad (3)$$

$$\approx [f(t)\mathbf{x}_t - g^2(t)s_\theta(\mathbf{x}_t, t)] dt + g(t)d\mathbf{w}_t, \quad (4)$$

$$= \left[ f(t)\mathbf{x}_t + \frac{g^2(t)}{\sigma_t} \epsilon_\theta(\mathbf{x}_t, t) \right] dt + g(t)d\mathbf{w}_t, \quad (5)$$

where  $\nabla_{\mathbf{x}_t} \log p_t(\mathbf{x}_t)$  is the so-called score function and in this work is approximated by the score model  $s_\theta(\mathbf{x}_t, t)$  with parameters  $\theta$ . The score function can also be approximated using  $\epsilon_\theta$ , which learns to predict the noise  $\epsilon$  in noisy  $\mathbf{x}_t$ , and is related to  $s_\theta$  by  $\epsilon_\theta = -\sigma_t s_\theta$ . The reverse denoising SDE can then be solved using well known techniques such as the Euler-Maruyama method [37].

While the above formulation of the diffusion model is generally applicable to many data types, it is modified in this work to account for atomic structures in periodic cell boundaries. Note that an atomic structure may be generally represented as a tuple of the atomic coordinates and atom types  $(\mathbf{r}, \mathbf{z})$ , along with unit cell dimensions  $\mathbf{A} = (\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3) \in (\mathbb{R}^3, \mathbb{R}^3, \mathbb{R}^3)$  if the structure is periodic. However, the amorphous carbon systems studied in this work are single-element only (carbon) and therefore  $\mathbf{z}$  is omitted, leaving  $\mathbf{x} = \mathbf{r}$ . Since  $\mathbf{r}$  lies in a unit cell with periodic boundaries, it is sensible to apply the “variance-exploding” scheme for the noise schedule, where  $\alpha_t$  is fixed to 1 and  $\sigma_t$  monotonically increases from 0 to some maximum value  $k$ . This means that the drift coefficient becomes zero  $f(t) = 0$ . The updated forward noising SDE is then

$$d\mathbf{x}_t = g(t)d\mathbf{w}_t, \quad g^2(t) = \frac{d\sigma_t^2}{dt}, \quad (6)$$

and its corresponding reverse denoising SDE is

$$d\mathbf{x}_t = -g^2(t)\nabla_{\mathbf{x}_t} \log p_t(\mathbf{x}_t)dt + g(t)d\mathbf{w}_t. \quad (7)$$

Importantly, due to the periodic boundaries, the noisy, perturbed structure becomes asymptotically a uniformly random structure. Therefore, to sample a realistic structure, one instead initializes a set of uniformly random coordinates  $\mathbf{r}$  inside a periodic unit cell—rather than Gaussian-distributed  $\mathbf{r}$ —and then applies the reverse denoising SDE. The structure generation algorithm is listed in Algorithm 1. In this work we define the noise schedule as  $(\alpha_t, \sigma_t) = (1, kt)$ , where the span of  $t$  is still set to  $(0, 1)$ . Notably, the structure generation algorithm is *scale-agnostic*, allowing for users to specify the unit cell dimensions  $\mathbf{A}$  and structure density  $D$ , which influence the number of atoms  $N_a$  for the generation. In practice,  $D$  should have a realistic value, similar to the densities of the structures in the training dataset.

---

**Algorithm 1** Unconditional disordered structure generation via Euler-Maruyama method

---

**Require:** Periodic unit cell  $\mathbf{A} = (\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3)$ , density  $D$ , learned score model  $s_\theta$ , noise schedule  $g(t)$ , and timestep size  $\Delta t$  (which has a negative value).  
Determine number of atoms  $N_a$  from  $\mathbf{A}$  and  $\rho$   
 $\mathbf{x} \leftarrow \mathcal{U}(\Omega)$  for  $N_a$  atoms, where  $\Omega$  is the domain inside  $\mathbf{A}$   
 $t \leftarrow 0.999$   
**repeat**  
     $\Delta\mathbf{x} \leftarrow -g^2(t)s_\theta(\mathbf{x}, t)\Delta t + g(t)\epsilon, \quad \epsilon \sim \mathcal{N}(0, |\Delta t|\mathbf{I})$   
     $\mathbf{x} \leftarrow \mathbf{x} + \Delta\mathbf{x}$   
     $t \leftarrow t + \Delta t$   
**until**  $t \approx 0$

---

## 4.2 Conditional diffusion model

The theory of the diffusion model, based on an underlying score function, can be quite naturally extended for conditional generation. First, the score function introduced in the previous section is written as  $\nabla_{\mathbf{x}_t} \log p_t(\mathbf{x}_t)$  and is regarded as the unconditional, unguided, or prior score function, i.e., without conditioning bias or guide during the generation/denoising process. Then, consider a conditioning or posterior score function

$\nabla_{\mathbf{x}_t} \log p_t(\mathbf{x}_t|y)$  that is conditioned on some target or guide property  $y$ . Importantly, this posterior score can be decomposed via Bayes' rule into

$$\nabla_{\mathbf{x}_t} \log p_t(\mathbf{x}_t|y) = \nabla_{\mathbf{x}_t} \log p_t(\mathbf{x}_t) + \nabla_{\mathbf{x}_t} \log p_t(y|\mathbf{x}_t) \quad (8)$$

$$\approx s_\theta(\mathbf{x}, t) + \nabla_{\mathbf{x}_t} \log p_t(y|\mathbf{x}_t). \quad (9)$$

It is now apparent that conditional generation enabled by the diffusion model can be very flexible, since computing the posterior score amounts to simple addition of the decoupled prior score  $\nabla_{\mathbf{x}_t} \log p_t(\mathbf{x}_t)$ —approximated by  $s_\theta$ —and the likelihood score  $\nabla_{\mathbf{x}_t} \log p_t(y|\mathbf{x}_t)$ . Further, computing the likelihood score entails taking the gradient of (the log of) a probabilistic forward prediction model  $p_t(y|\mathbf{x}_t)$ . For different conditional quantities  $y$  such as spectral signatures, Steinhardt parameters, band gap, and so on, different forward models can be combined with the prior model in a plug-and-play fashion. In the case of a deterministic forward model  $y = F(\mathbf{x})$ , converting it into the probabilistic counterpart can be straightforward by adding Gaussian noises of variance  $\hat{\sigma}^2$  to the forward prediction:  $y = F(\mathbf{x}) + \epsilon$ ,  $\epsilon \sim \mathcal{N}(0, \hat{\sigma}^2 \mathbf{I})$ . The resulting likelihood score is then

$$\nabla_{\mathbf{x}_t} \log p_t(y|\mathbf{x}_t) = -\frac{1}{2\hat{\sigma}^2} \nabla_{\mathbf{x}_t} \|y - F(\mathbf{x}_t)\|^2, \quad (10)$$

which can be interpreted as pointing to lower values of the L2 loss between the forward prediction  $F(\mathbf{x}_t)$  and a given target  $y$ . In terms of implementation, the conditional generation process is virtually identical to Algorithm 1 except for the input score model, which is replaced by  $\nabla_{\mathbf{x}_t} \log p_t(\mathbf{x}_t|y)$ .

However, there is still a caveat with the conditional generation formulation: evaluating the forward prediction  $F(\mathbf{x}_t)$  involves noisy input  $\mathbf{x}_t$  but the forward model may be ill-defined with respect to noisy, highly perturbed  $\mathbf{x}_t$ , which at  $t = 1$  is complete noise without any meaningful signal. For some forward models, this does not pose as an issue, such as the case of atomic descriptors/fingerprints, e.g., Steinhardt[38], ACSF[39], SOAP[40], MBTR[41] order parameters, where the structure-to-descriptor forward mapping still has a valid output when  $\mathbf{x}_t$  is highly noisy or even completely random. The forward model developed in this work, however, is a graph network model trained to predict spectroscopic curves off of atomic motifs in disordered or amorphous carbons, which still exhibit certain topological order and are far from completely random. Our forward model output is therefore ill-defined and likely holds no meaningful gradient information with noisy input. To address this, Chung et al.[42] proposed to inverse the relationship  $\mathbf{x}_t = \alpha_t \mathbf{x}_0 + \sigma_t \epsilon$  and estimate the “clean” version of data  $\hat{\mathbf{x}}_0$  from noisy  $\mathbf{x}_t$  as

$$\hat{\mathbf{x}}_0 = \frac{1}{\alpha_t} (\mathbf{x}_t - \sigma_t \epsilon_\theta(\mathbf{x}_t, t)) \quad (11)$$

$$= \frac{1}{\alpha_t} (\mathbf{x}_t + \sigma_t^2 s_\theta(\mathbf{x}_t, t)), \quad (12)$$

followed by forward model prediction, resulting in the following likelihood score:

$$\nabla_{\mathbf{x}_t} \log p_t(y|\mathbf{x}_t) = -\rho \nabla_{\mathbf{x}_t} \|y - F(\hat{\mathbf{x}}_0(\mathbf{x}_t))\|^2 = -\frac{\xi}{\|y - F(\cdot)\|} \nabla_{\mathbf{x}_t} \|y - F(\hat{\mathbf{x}}_0(\mathbf{x}_t))\|^2, \quad (13)$$

where the stepsize  $\rho$  is another implementation practice by Chung et al.[42], and is defined as a normalized quantity  $\rho = \xi / \|y - F(\hat{\mathbf{x}}_0(\mathbf{x}_t))\|$ , with  $\xi$  being a tuning parameter. In this work, we have found the conditional generation with the spectroscopy forward model to be viable only with this “estimation trick”.

### 4.3 Forward model

As previously stated, the forward model used in this work for the purpose of conditional or guided generation of disordered carbons is a spectroscopy prediction model. Specifically, it is a GNN that outputs per-atom spectroscopic curves based on local atomic motifs. In this representation, the atomic structure is represented as a graph, wherein nodes encapsulate information pertaining to atomic species and atomic positional coordinates, while edges encode atomic bonding information in the form of vectorized interatomic distances. The GNN is subjected to training on a dataset containing atomistic structures along with their corresponding XANES spectra.

## 4.4 Model training

The score model  $s_\theta(\mathbf{x}_t, t)$  is a GNN trained with dataset  $\mathcal{D}_s = \{(\mathbf{x}_0, \mathbf{A})^{(i)}\}$  that consists of MD snapshots containing atomic coordinates  $\mathbf{x}_0$  and unit cell dimensions  $\mathbf{A}$ , with atom type information  $\mathbf{z}$  omitted due to the dataset being single-element. The score is learned via the denoising score matching loss

$$\begin{aligned} \mathbf{x}_t &= \alpha_t \mathbf{x}_0 + \sigma_t \epsilon, \quad \epsilon \sim \mathcal{N}(0, \mathbf{I}), \\ \mathcal{L}_{\text{DSM}} &= \mathbb{E}_{\mathbf{x}_0, \epsilon, t} [\|s_\theta(\mathbf{x}_t, t) \sigma_t + \epsilon\|^2], \end{aligned} \quad (14)$$

where  $\mathbf{A}$  is implicitly accounted for by the atomic graph representation. Algorithm 2 lists the pseudocode for training the score model.

---

**Algorithm 2** Score model training

---

**Require:** Training dataset  $\mathcal{D}_s$ , noise schedule  $(\alpha_t, \sigma_t)$ , initial model parameters  $\theta$ , another set of model parameters  $\theta'$  for EMA, EMA decay rate  $\beta$ , gradient descent optimizer `Optim`, and learning rate  $\eta$ .

**repeat**

```
    Sample  $(\mathbf{x}_0, \mathbf{A}) \sim \mathcal{D}$ 
    Sample  $t \sim \mathcal{U}(0.001, 0.999)$ 
    Sample  $\epsilon \sim \mathcal{N}(0, \mathbf{I})$ 
     $\mathbf{x}_t \leftarrow \alpha_t \mathbf{x}_0 + \sigma_t \epsilon, \quad \epsilon \sim \mathcal{N}(0, \mathbf{I})$ 
     $\mathcal{L}_{\text{DSM}} \leftarrow \mathbb{E}_{\mathbf{x}_0, \epsilon, t} [\|s_\theta(\mathbf{x}_t, t) \sigma_t + \epsilon\|^2]$ 
     $\theta \leftarrow \text{Optim}(\mathcal{L}_{\text{DSM}}, \theta, \eta)$ 
     $\theta' \leftarrow \beta \theta' + (1 - \beta) \theta$ 
```

**until** convergence

---

The forward model is also a GNN, trained with dataset  $\mathcal{D}_F = \{(\mathbf{x}_0, \mathbf{A}, y)^{(i)}\}$  that additionally has per-atom spectroscopic curves  $y$ . The spectroscopic prediction is learned via a simple mean-squared-error (MSE) loss between the predicted output  $\hat{y}$  and  $y$ :  $\mathcal{L}_{\text{MSE}} = \text{MSE}(\hat{y}, y)$ .

For both the score model and the forward model, exponential moving average (EMA) was used to facilitate training. Rectified Adam (RAdam) was used as the gradient descent optimizer without learning rate or weight decay. For every sampled minibatch of the structure snapshots, the atomic coordinates are randomly rotated in order for the score, (forward) model to learn rotational equivariance (invariance). The training parameters are detailed in Table 1. All other parameters, if unspecified, default to implementations in PyTorch 1.11.0 [43] and PyTorch-Geometric 2.0.4 [44], which were used for the GNN implementations and training.

## 4.5 GNN architecture

Both the score model and the forward model share the same GNN architecture design, which consists of three components: an Encoder, a Processor (for graph message-passing), and a Decoder. The input atomic graph to these operations is denoted as  $(V, E)$ , where  $V$  is a set node features  $\{\mathbf{h}_i\}$  for node atom  $i$ , and  $E$  is a set of edge/bond features  $\{\mathbf{e}_{ij}\}$  from node  $i$  to node  $j$ . The set of edges  $E$  is determined by a cutoff distance  $r_c$ , within which a pair of nodes form an edge connection.

In the Encoder stage, the node features are computed as a combined embedding of the atom type  $\mathbf{z}$  and the fictitious time  $t$ :

$$\mathbf{h}_i \leftarrow M_H(\text{onehot}(\mathbf{z}_i)) + M_T(\text{GaussianRFF}(t)), \quad (15)$$

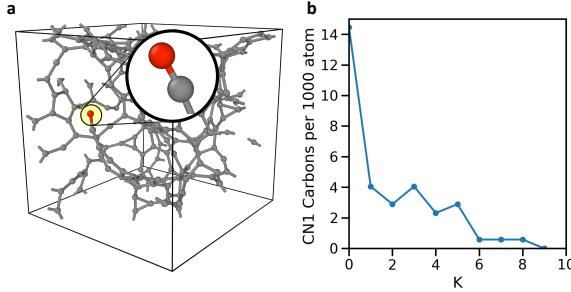
where  $M_H$  and  $M_T$  are multilayer perceptrons (MLPs) that embed the one-hot representation of  $\mathbf{z}$  and the Gaussian Fourier features [45] (GaussianRFF) of  $t$ , respectively. Note that in this work,  $\mathbf{z}$  is simply a dummy variable because the atom type information in the amorphous carbon dataset is omitted, but the Encoder is defined as above so that it can be adjusted to account for different atom types. The edge feature  $\mathbf{e}_{ij}$  is the concatenation of the edge vectors  $\mathbf{r}_{ij} = \mathbf{r}_j - \mathbf{r}_i$  and the corresponding edge lengths  $\bar{r}_{ij}$ , followed by an MLP  $M_E$ :

$$\mathbf{e}_{ij} \leftarrow M_E(\mathbf{r}_{ij} \oplus \bar{r}_{ij}), \quad (16)$$

where  $\oplus$  designates concatenation of features.  $M_H$ ,  $M_E$ , and  $M_T$  consist of two dense linear layers, with SiLU activation function after the first layer and layer normalization after the second.

**Table 1:** Training and model parameters.

| Name                                | Value       |               |
|-------------------------------------|-------------|---------------|
|                                     | Score model | Forward model |
| Number of model updates             | 1,000,000   | 400,000       |
| Learning rate $\eta$                | 0.001       | 0.0001        |
| EMA decay rate $\beta$              | 0.999       | 0.999         |
| Minibatch size                      | 64          | 8             |
| Node, edge, time feature dimensions | 256         | 256           |
| Number of graph convolutions        | 5           | 3             |



**Figure 8:** Unphysical CN1 (coordination number of 1) carbons can be effectively reduced by  $K$  cycles of restart sampling. An example of CN1 carbon is shown in (a). Roughly 6–10 cycles of restart sampling can effectively minimize occurrences of CN1 carbons (b).

We used the Processor from MeshGraphNets [46] for implementing graph convolutions. We refer to the original paper for further details of MeshGraphNets. There is a main difference between our implementation and the original work: the original work considers two sets of edges (mesh-space edges and world-space edges), whereas there is only one in our work.

The Decoder is a simple MLP  $M_O$  that maps the node features into prediction output:

$$\hat{y}_i \leftarrow M_O(\mathbf{h}_i). \quad (17)$$

For the score model, the output is the approximated score function, and for the forward model, the output is the spectroscopic signature.  $M_O$  has the same architecture as that of  $M_H$ ,  $M_E$ , and  $M_T$ , except that there is no layer normalization at the end.

Model parameters are listed in Table 1.

## 4.6 Restart sampling

Since the generative model in this work is not constrained by physics rule, there is always a chance for the model to generate unphysical topologies that are extremely unlikely or virtually impossible. Such “bad” samples occur since there are always errors associated with any fitted score model. In this work one main instance of unphysical  $a$ -C topology is carbons with a coordination number of only one (CN1), which can exist in our generation results using the usual SDE sampler. To address this, we have found the *restart sampling* method [47] to be effective in avoiding CN1 carbon generation (Fig. 8). In brief, the restart sampling method entails multiple cycles of reapplying forward noising and reverse denoising operations between two points in the denoising time schedule. In this work, the two points are  $t = 0, 0.3$ . Note that we apply SDE denoising, but the original restart sampling technique is based on ODE (ordinary differential equation) denoising.

## 4.7 X-ray absorption spectral simulations

We extracted a dataset consisting of 58 atomic configurations of *a*-C at different densities from molecular dynamics (MD) simulations with the Gaussian approximation potential (GAP), which has been shown to provide a good description of structural properties of *a*-C at the room temperature [34]. Each configuration consists of 216 atoms in a simple-cubic lattice with varying densities of 1.5, 2.0, 2.5, 3.0, and 3.5 g/cm<sup>3</sup> to sample a wide range of local carbon structures. Collectively, this dataset encompassed a total of 12,528 local carbon sites. We then calculated K-edge spectra of these atomic sites using Density Functional Theory (DFT) with the core-hole approximation (XCH) method [48]. All computational procedures were carried out utilizing the QuantumESPRESSO package [49]. To efficiently sample the Brillouin zone, we employed the Shirley reduced basis method [50]. The Perdew–Burke–Ernzerhof (PBE) formulation of the generalized-gradient approximation (GGA) was adopted [51]. We represented electronic wavefunctions and charge density using a plane-wave basis set with kinetic energy cutoffs of 30 Ry and 240 Ry, respectively. All our calculations utilized ultrasoft pseudopotentials [52]. For the excited atom, we employed a modified pseudopotential with one electron removed from the 1s core orbital.

Within the XCH approximation, we approximated the first core-excited state through constrained-occupancy DFT, which involved incorporating both the core-hole pseudopotential and the excited electron positioned at the conduction band minimum. Transition probabilities at each energy level were computed using Fermi's Golden Rule and then convoluted with a uniform Gaussian broadening of 0.2 eV. To align the calculated spectrum with experimental data [53], we matched the first major peak of the graphite spectrum and applied the same constant shift to align the spectra of all carbon sites within *a*-C systems. It is necessary to note that other methods can also be used to compute XANES spectra, such as those based on time-dependent DFT or more computationally intensive Bethe–Salpeter calculations, which explicitly consider the interaction between the electron and core-hole [54–56].

## Acknowledgements

This work was carried out under the auspices of the US Department of Energy by Lawrence Livermore National Laboratory (LLNL) under contract No. DE-AC52-07NA27344. We acknowledge support from LLNL Laboratory Directed Research and Development (LDRD) Program Grant No. 22-ERD-014. This research used the Theory and Computation facility of the Center for Functional Nanomaterials, which is a US Department of Energy Office of Science User Facility at Brookhaven National Laboratory under Contract DE-SC0012704. M.R.C., and D.L. acknowledge the support by the U.S. Department of Energy, Office of Science, Office Basic Energy Sciences, under Award FWPPS-030. F.Z. was supported by LLNL LDRD Grant 23-SI-006. J.C acknowledges support from the College of Engineering and the Department of Mechanical Engineering at Boston University.

## Author Contributions

H.K., T.H., and T.A.P. supervised the research. F.A. carried out MD simulations of amorphous carbons. W.S. and W. J. computed the XANES spectra of amorphous carbons. F.Z. and T.H. developed the concept of conditional diffusion models. H.K., T.H., and T.A.P. wrote the manuscript with inputs from all authors.

## Data Availability

All data required to reproduce this work can be requested by contacting the corresponding authors.

## Code Availability

The source code for this work is available at <https://github.com/LLNL/graphite>.

## Competing interests

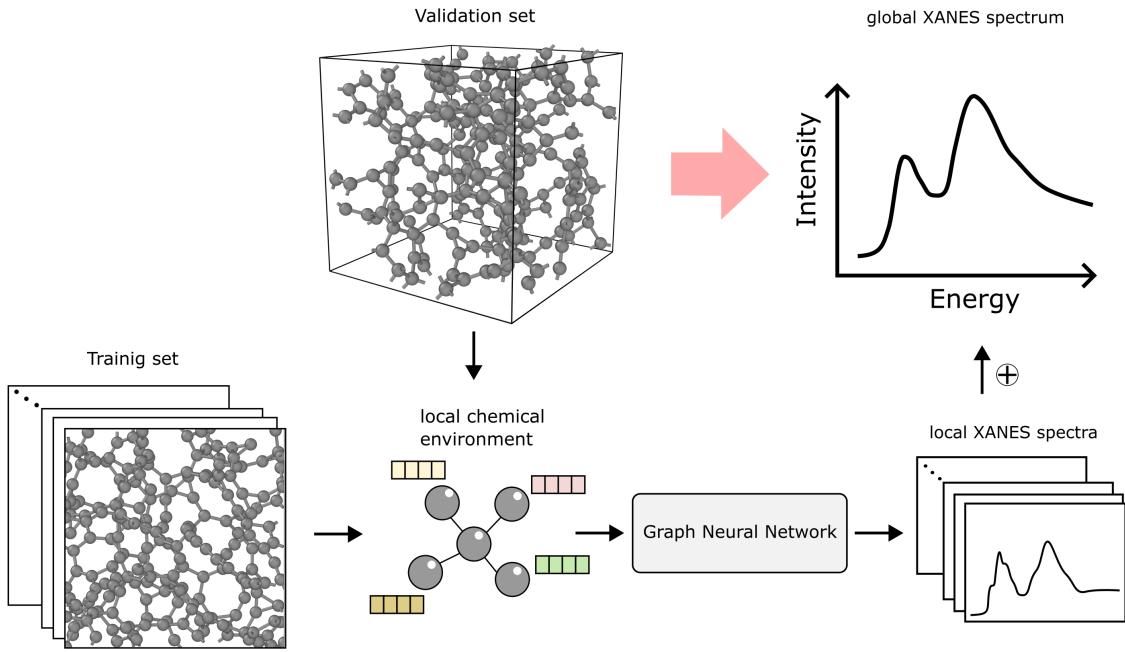
The authors state that there is no conflict of interest.

## References

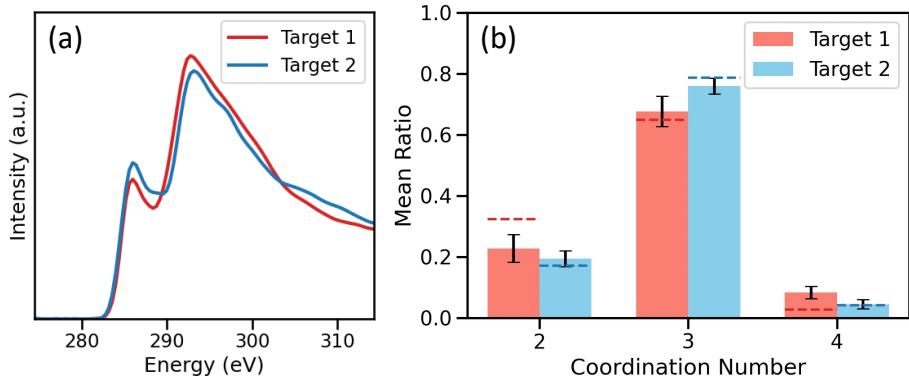
1. Simon, P. & Gogotsi, Y. Materials for electrochemical capacitors. *Nature materials* **7**, 845–854 (2008).
2. Nitta, N., Wu, F., Lee, J. T. & Yushin, G. Li-ion battery materials: present and future. *Materials today* **18**, 252–264 (2015).
3. Aluru, N. R. *et al.* Fluids and Electrolytes under Confinement in Single-Digit Nanopores. *Chemical reviews* **123**, 2737–2831 (2023).
4. Correa-Baena, J.-P. *et al.* Accelerating materials development via automation, machine learning, and high-performance computing. *Joule* **2**, 1410–1420 (2018).
5. Jain, A. *et al.* A high-throughput infrastructure for density functional theory calculations. *Computational Materials Science* **50**, 2295–2310 (2011).
6. Jain, A. *et al.* Commentary: The Materials Project: A materials genome approach to accelerating materials innovation. *APL materials* **1** (2013).
7. Sanchez-Lengeling, B. & Aspuru-Guzik, A. Inverse molecular design using machine learning: Generative models for matter engineering. *Science* **361**, 360–365 (2018).
8. Noh, J. *et al.* Inverse design of solid-state materials via a continuous representation. *Matter* **1**, 1370–1384 (2019).
9. Yao, Z. *et al.* Inverse design of nanoporous crystalline reticular materials with deep generative models. *Nature Machine Intelligence* **3**, 76–86 (2021).
10. Kim, B., Lee, S. & Kim, J. Inverse design of porous materials using artificial neural networks. *Science advances* **6**, eaax9324 (2020).
11. Noh, J., Gu, G. H., Kim, S. & Jung, Y. Machine-enabled inverse design of inorganic solid materials: promises and challenges. *Chemical Science* **11**, 4871–4881 (2020).
12. Sohl-Dickstein, J., Weiss, E., Maheswaranathan, N. & Ganguli, S. *Deep unsupervised learning using nonequilibrium thermodynamics* in *International Conference on Machine Learning* (2015), 2256–2265.
13. Ho, J., Jain, A. & Abbeel, P. Denoising diffusion probabilistic models. *Advances in Neural Information Processing Systems* **33**, 6840–6851 (2020).
14. Song, Y. *et al.* Score-based generative modeling through stochastic differential equations. *arXiv preprint arXiv:2011.13456* (2020).
15. Hyvärinen, A. & Dayan, P. Estimation of non-normalized statistical models by score matching. *Journal of Machine Learning Research* **6** (2005).
16. Vincent, P. A connection between score matching and denoising autoencoders. *Neural computation* **23**, 1661–1674 (2011).
17. Hoogeboom, E., Satorras, V. G., Vignac, C. & Welling, M. *Equivariant diffusion for molecule generation in 3d* in *International conference on machine learning* (2022), 8867–8887.
18. Jing, B., Corso, G., Chang, J., Barzilay, R. & Jaakkola, T. Torsional diffusion for molecular conformer generation. *Advances in Neural Information Processing Systems* **35**, 24240–24253 (2022).
19. Vignac, C. *et al.* Digress: Discrete denoising diffusion for graph generation. *arXiv preprint arXiv:2209.14734* (2022).
20. Xu, M. *et al.* Geodiff: A geometric diffusion model for molecular conformation generation. *arXiv preprint arXiv:2203.02923* (2022).
21. Weiss, T. *et al.* Guided Diffusion for Inverse Molecular Design (2023).

22. Xie, T., Fu, X., Ganea, O.-E., Barzilay, R. & Jaakkola, T. Crystal diffusion variational autoencoder for periodic material generation. *arXiv preprint arXiv:2110.06197* (2021).
23. Zheng, S. *et al.* Towards Predicting Equilibrium Distributions for Molecular Systems with Deep Learning. *arXiv preprint arXiv:2306.05445* (2023).
24. Pham, T. A., Ping, Y. & Galli, G. Modelling heterogeneous interfaces for solar water splitting. *Nature materials* **16**, 401–408 (2017).
25. Wood, B. C., Heo, T. W., Kang, S., Wan, L. F. & Li, S. Beyond idealized models of nanoscale metal hydrides for hydrogen storage. *Industrial & Engineering Chemistry Research* **59**, 5786–5796 (2020).
26. Comin, M. & Lewis, L. J. Deep-learning approach to the structure of amorphous silicon. *Physical Review B* **100**, 094107 (2019).
27. Tolstikhin, I., Bousquet, O., Gelly, S. & Schoelkopf, B. Wasserstein auto-encoders. *arXiv preprint arXiv:1711.01558* (2017).
28. Chapman, J. *et al.* Hydrogen in disordered titania: connecting local chemistry, structure, and stoichiometry through accelerated exploration. *Journal of Materials Chemistry A* **11**, 8670–8683 (2023).
29. Jiang, S. *et al.* Molecular dynamics simulations of gas selectivity in amorphous porous molecular solids. *Journal of the American Chemical Society* **135**, 17818–17830 (2013).
30. Gaikwad, P. S., Kowalik, M., Jensen, B. D., Van Duin, A. & Odegard, G. M. Molecular dynamics modeling of interfacial interactions between flattened carbon nanotubes and amorphous carbon: Implications for ultra-lightweight composites. *ACS applied nano materials* **5**, 5915–5924 (2022).
31. Li, K. *et al.* ReaxFF molecular dynamics simulation for the graphitization of amorphous carbon: a parametric study. *Journal of chemical theory and computation* **14**, 2322–2331 (2018).
32. Steinhardt, P. J., Nelson, D. R. & Ronchetti, M. Bond-orientational order in liquids and glasses. *Physical Review B* **28**, 784 (1983).
33. Caro, M. A. *GAP interatomic potential for amorphous carbon (2.0) [Data set]*
34. Deringer, V. L. & Csányi, G. Machine learning based interatomic potential for amorphous carbon. *Physical Review B* **95**, 094203 (2017).
35. Kwon, H. *et al.* Harnessing Neural Networks for Elucidating X-ray Absorption Structure–Spectrum Relationships in Amorphous Carbon. *The Journal of Physical Chemistry C* **127**, 16473–16484 (2023).
36. Jana, R., Savio, D., Deringer, V. L. & Pastewka, L. Structural and elastic properties of amorphous carbon from simulated quenching at low rates. *Modelling and Simulation in Materials Science and Engineering* **27**, 085009 (2019).
37. Mao, X. The truncated Euler–Maruyama method for stochastic differential equations. *Journal of Computational and Applied Mathematics* **290**, 370–384 (2015).
38. Steinhardt, P. J., Nelson, D. R. & Ronchetti, M. Bond-orientational order in liquids and glasses. *Physical Review B* **28**, 784 (1983).
39. Behler, J. Atom-centered symmetry functions for constructing high-dimensional neural network potentials. *The Journal of chemical physics* **134** (2011).
40. Bartók, A. P., Kondor, R. & Csányi, G. On representing chemical environments. *Physical Review B* **87**, 184115 (2013).
41. Huo, H. & Rupp, M. Unified representation of molecules and crystals for machine learning. *arXiv preprint arXiv:1704.06439* (2017).
42. Chung, H., Kim, J., Mccann, M. T., Klasky, M. L. & Ye, J. C. *Diffusion Posterior Sampling for General Noisy Inverse Problems in The Eleventh International Conference on Learning Representations* (2022).
43. Paszke, A. *et al.* Pytorch: An imperative style, high-performance deep learning library. *Advances in neural information processing systems* **32**, 8026–8037 (2019).
44. Fey, M. & Lenssen, J. E. Fast graph representation learning with PyTorch Geometric. Preprint at <https://arxiv.org/abs/1903.02428> (2019).

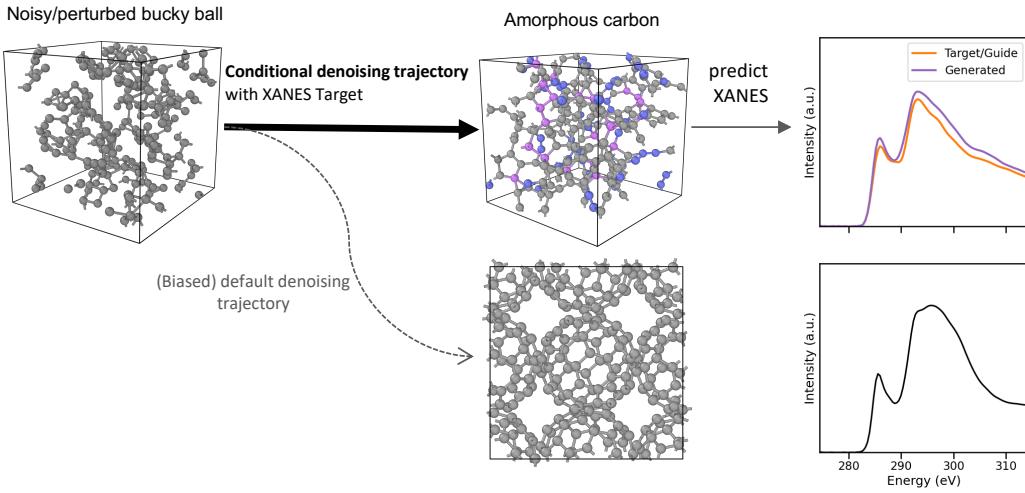
45. Tancik, M. *et al.* Fourier features let networks learn high frequency functions in low dimensional domains. *Advances in Neural Information Processing Systems* **33**, 7537–7547 (2020).
46. Pfaff, T., Fortunato, M., Sanchez-Gonzalez, A. & Battaglia, P. W. Learning mesh-based simulation with graph networks. *arXiv preprint arXiv:2010.03409* (2020).
47. Xu, Y. *et al.* Restart Sampling for Improving Generative Processes. *arXiv preprint arXiv:2306.14878* (2023).
48. Prendergast, D. & Galli, G. X-ray absorption spectra of water from first principles calculations. *Phys. Rev. Lett.* **96**, 215502 (2006).
49. Giannozzi, P. *et al.* QUANTUM ESPRESSO: A modular and open-source software project for quantum simulations of materials. *J. Phys. Condens. Matter* **21**, 395502 (2009).
50. Prendergast, D. & Louie, S. G. Bloch-state-based interpolation: An efficient generalization of the Shirley approach to interpolating electronic structure. *Phys. Rev. B* **80**, 235126 (2009).
51. Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865 (1996).
52. Vanderbilt, D. Soft self-consistent pseudopotentials in a generalized eigenvalue formalism. *Phys. Rev. B* **41**, 7892 (1990).
53. Sainio, S. *et al.* Trends in carbon, oxygen, and nitrogen core in the X-ray absorption spectroscopy of carbon nanomaterials: a guide for the perplexed. *J. Phys. Chem. C* **125**, 973–988 (2020).
54. Shirley, E. L. Ab initio inclusion of electron-hole attraction: Application to X-ray absorption and resonant inelastic X-ray scattering. *Phys. Rev. Lett.* **80**, 794 (1998).
55. Vinson, J., Kas, J. J., Vila, F., Rehr, J. & Shirley, E. Theoretical optical and X-ray spectra of liquid and solid H<sub>2</sub>O. *Phys. Rev. B* **85**, 045101 (2012).
56. Lopata, K., Van Kuiken, B. E., Khalil, M. & Govind, N. Linear-response and real-time time-dependent density functional theory studies of core-level near-edge X-ray absorption. *J. Chem. Theory Comput.* **8**, 3284–3292 (2012).



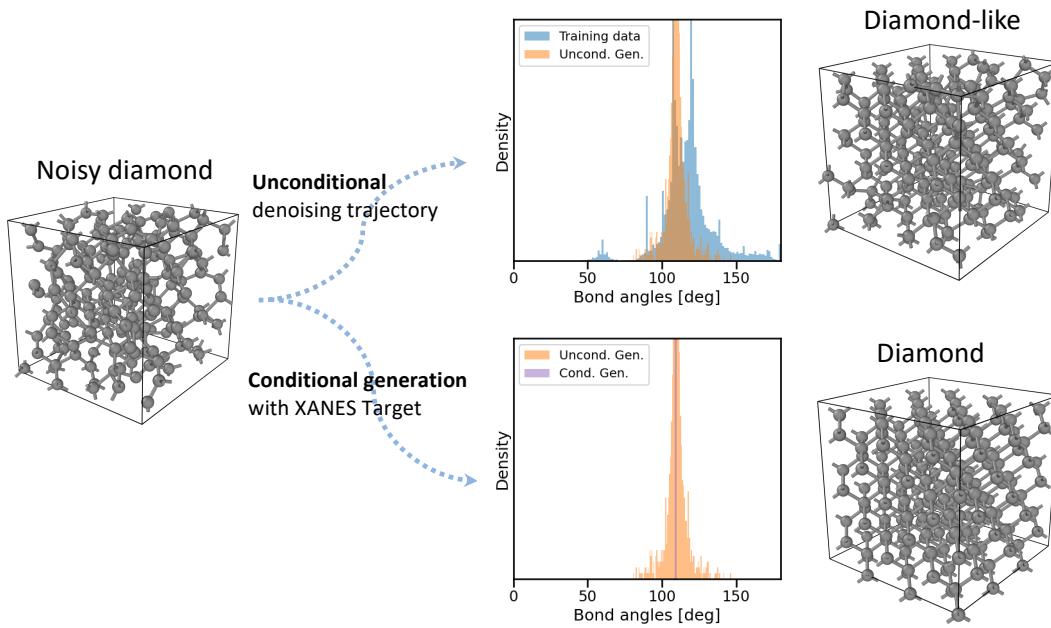
**Supplementary Figure 1:** Schematic of forward model. This schematic elucidates the operation of our forward model, which receives local chemical environment data as input and employs an equivariant graph neural network for the production of local XANES spectra of individual atom. Subsequently, for validation purposes, we predict the local XANES spectra of all atoms in structures that were omitted from the training dataset. These locally predicted spectra are cumulatively summed to generate the global XANES spectra, allowing for a quantitative assessment of their concordance with the DFT computed global XANES spectrum.



**Supplementary Figure 2:** (a) Spectroscopy of two distinct structures with a density of  $1.5 \text{ g/cm}^3$ . Despite having the same density, their differing topologies result in distinct spectroscopic profiles. (b) Coordination number ratios of amorphous carbon structures generated through conditional generation based on the spectroscopies in panel (a). Dashed lines represent target coordination numbers, derived from real structures corresponding to the spectroscopies in panel (a). The coordination number ratios of the generated structures closely align with these target values, validating the fidelity of the conditional generation process.



**Supplementary Figure 3:** The generation process, starting from an initial noisy structure, unfolds in two distinct transformations: (1) conditional denoising trajectory, influenced by target XANES spectroscopy, and (2) default denoising trajectory, driven by an unconditional diffusion model. Comparisons of the generated amorphous structures against target data, including spectroscopy predictions is shown on the right panel.



**Supplementary Figure 4:** The top row shows the generation of diamond-like structure from initial heavily noisy diamond structure via unconditional generation, and the bottom row illustrates the refined structure achieved through conditional generation with the guide based on the diamond's spectrum. The conditional generation process significantly enhances the resemblance of the resulting structure to the ideal diamond structure. This showcases the utility of conditional generation in refining generated structures with specific spectroscopic targets.