SPACE GROUP CONDITIONAL FLOW MATCHING

Anonymous authors
Paper under double-blind review

ABSTRACT

Inorganic crystals are periodic, highly-symmetric arrangements of atoms in three-dimensional space. Their structures are constrained by the symmetry operations of a crystallographic *space group* and restricted to lie in specific affine subspaces known as *Wyckoff positions*. The frequency an atom appears in the crystal and its rough positioning are determined by its Wyckoff position. Most generative models that predict atomic coordinates overlook these symmetry constraints, leading to unrealistically high populations of proposed crystals exhibiting limited symmetry. We introduce Space Group Conditional Flow Matching, a novel generative framework that samples significantly closer to the target population of highly-symmetric, stable crystals. We achieve this by conditioning the entire generation process on a given space group and set of Wyckoff positions; specifically, we define a conditionally symmetric noise base distribution and a group-conditioned, equivariant, parametric vector field that restricts the motion of atoms to their initial Wyckoff position. Our form of group-conditioned equivariance is achieved using an efficient reformulation of *group averaging* tailored for symmetric crystals. Importantly, it reduces the computational overhead of symmetrization to a negligible level. We achieve state of the art results on crystal structure prediction and de novo generation benchmarks. We also perform relevant ablations.

1 INTRODUCTION

Crystals are solid materials characterized by a periodic arrangement of their constituent atoms. The crystalline structure is fundamentally represented by three components: lattice parameters (defining the geometry of the repeating unit cell), fractional coordinates (specifying the position of each atom within the cell), and the identity of the atom at each location. The discovery of novel crystalline structures is critical for material design and recent progress in generative modeling has demonstrated a promising approach to this problem. However, most existing generative methods overlook key crystallographic properties, the space group and Wyckoff positions, making it challenging for them to generate non-trivial symmetric crystals.

A crystal's space group, a subgroup of the Euclidean group $E(n)$, fully describes the symmetry of the atoms arranged within the unit cell. Beyond its correlation with many optical, electrical, magnetic, and structural properties (Chen et al., 2022; Tang et al., 2019; Malgrange et al., 2014; Yang, 2005), the space group imposes constraints on atomic locations and lattice structure. These manifests in form of Wyckoff positions, which are sets of symmetrically equivalent points within a unit cell. More generally, the Wyckoff positions of a space group partition the unit cell according to the structure of the orbits induced by the group (see fig. 2 for a 2D example).

In this work, we develop a generative model that samples crystals conditioned on a given space group and associated Wyckoff positions. This approach offers two key benefits: (1) it provides greater control over the structure and symmetry of the generated crystals and, and (2) it can leverage the lower-dimensional constraints imposed by Wyckoff positions for improving generation. In contrast to prior methods that incorporate space group information but rely on projection steps to correct atomic placements (Jiao et al., 2024; Levy et al., 2025), our model is designed to inherently preserve the assignment of atoms to their designated Wyckoff positions throughout the generation process.

Our proposed model, *Space Group Conditional Flow Matching* (SGFM), is based on the *Flow Matching* (FM) generative framework (Albergo & Vanden-Eijnden, 2022; Liu et al., 2022; Lipman et al., 2022). We chose the FM framework for two key reasons: it allows us to use customized source

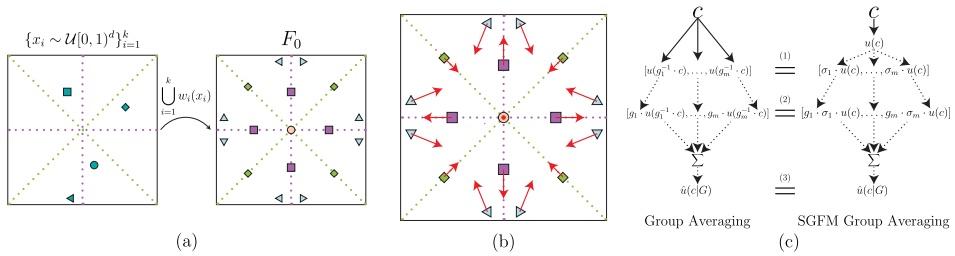

Figure 1: Visualization of the main components of SGFM. (a) Wyckoff position noise prior. General points are sampled randomly and projected according to the conditioned Wyckoff positions. (b) Space group equivariant vector field. The equivariance of the model combined with the G-symmetry of the input crystal ensures that atoms preserve their symmetry structure. (c) Comparison between group averaging and our optimization, heavy arrow implies expensive model forward pass.

distributions and provides known conditions for generating data that respects specified symmetries (Köhler et al., 2020). Based on these advantages, we designed SGFM with two main components: A space group and Wyckoff position conditioned noise prior, which have positive support only for crystal structures that adhere to the symmetry constraints described by the Wyckoff positions; A group conditioned equivariant vector field, which is a single neural network architecture that is able to support arbitrary space group equivariance. Equivariance is achieved through Group Averaging (GA) (Yarotsky, 2022), a symmetrization technique that projects arbitrary functions onto their equivariant versions. Although GA is typically computationally expensive and impractical, we introduce an efficient formulation tailored for symmetric crystals, reducing the computational overhead of the symmetrization operator to a negligible level.

The main contributions of this work are as follows:

- We formalized the problem of symmetric crystal generation in terms of distributional symmetry properties (section 3.1), and extended the conditions introduced by Köhler et al. (2020) to enable flow-based models to sample from such distributions (section 3.2).
- We instantiate this flow model as SGFM (section 3.3), which consists of a noise prior conditional on Wyckoff position along with a space group-equivariant vector field, ensuring that the generated crystals preserve the specified symmetries.
- We propose a novel and efficient implementation of GA for symmetric crystals, equivalent in output to the standard GA but significantly more efficient (see fig. 1 (c)), practically minimizing the computational burden of symmetrization for symmetric crystals.
- SGFM achieves state-of-the-art performance on crystal structure prediction (CSP), a generative task centered on crystal structure generation, and de novo generation (DNG).

2 Preliminaries

Equivariance & Invariance. A function $\varphi : \mathcal{X} \to \mathcal{Y}$ is equivariant with respect to a group G if the action of any group element on the input corresponds to a consistent transformation of the output. Equivariance implies $\varphi(g \cdot x) = g \cdot \varphi(x)$ for all $x \in \mathcal{X}$ and $g \in G$. Invariance is a simplified case of equivariance, with all $g \in G$ mapping to a trivial group action on the output space $\varphi(g \cdot x) = \varphi(x)$. Invariance and equivariance also extend to group products: Let $(g_1, g_2) \in G_1 \times G_2$, φ is $G_1 \times G_2$ equivariant if $\varphi((g_1, g_2) \cdot x) = (g_1, g_2) \cdot \varphi(x)$. Additionally, *$G$-invariant distributions* refer to distributions which have an G-invariant density function. We will use this to construct SGFM.

Crystal Representation. A crystal can be represented using the tuple $c' = (\boldsymbol{L}, \boldsymbol{F}, \boldsymbol{A}) \in \mathcal{C}'$, where $\boldsymbol{L} \in \mathbb{R}^{3\times3}$ is a positive-determinant, invertible matrix defining the lengths, angles, and orientation of the positive-volume unit cell; $\boldsymbol{F} \in [0,1)^{n\times3}$ denotes the fractional coordinates of n atoms within the unit cell; and $\boldsymbol{A} \in \{0,1\}^{n\times h}$ is a one-hot matrix indicating the atom types in $\boldsymbol{F}$ from a set of h atom types. We adopt the space group–conditioned lattice parameterization proposed by Jiao et al. (2024), which replaces $\boldsymbol{L}$ with a rotation-invariant vector $k \in \mathbb{R}^6$ of coefficients of a symmetric matrix basis, and represent a crystal as $c = (k, \boldsymbol{F}, \boldsymbol{A}) \in \mathcal{C}$. Further details provided in appendix C.

Symmetries & Crystals. Our method focuses on how symmetry groups act on crystals. The permutation group S_n acts on c by permuting the rows of $\boldsymbol{F}$ and $\boldsymbol{A}$. Namely, if $\sigma \in S_n$ is represented by a permutation matrix $\boldsymbol{P} \in \{0,1\}^{n\times n}$ then $\sigma \cdot c = (k, \boldsymbol{PF}, \boldsymbol{PA})$. The group of isometries of $\mathbb{R}^3$ known as the Euclidean group $E(3)$, acts on c by applying an orthogonal transformation $\boldsymbol{R} \in O(3)$ and a translation $\tau \in \mathbb{R}^3$ to the fractional coordinates. For a group element $g = (\boldsymbol{R}, \tau) \in E(3)$, the action is defined as $g \cdot c = (k, \boldsymbol{F}\boldsymbol{R}^T + \mathbf{1}_n\tau^T - \lfloor \boldsymbol{F}\boldsymbol{R}^T + \mathbf{1}_n\tau^T \rfloor, \boldsymbol{A})$ where $\mathbf{1}_n \in \{1\}^n$ is a column vector of ones and $\lfloor\cdot\rfloor$ is the element-wise floor function. We further specify the action of the product group on c. Let $(g, \sigma) \in G \times S_n$, we define $(g, \sigma) \cdot c \coloneqq g \cdot (\sigma \cdot c)$. Puny et al. (2021) showed that if $G \leqslant E(3)$ then $g \cdot (\sigma \cdot c) = \sigma \cdot (g \cdot c)$, i.e., the operators commute.

Space Groups. The space group concept formalizes the intrinsic symmetry of a crystal. If $c \in \mathcal{C}$ is a crystal and $G \leqslant E(3)$ is its symmetry space group, then for any $g \in G$ there exist a permutation $\sigma \in S_n$ that satisfies the relation $g \cdot c = \sigma \cdot c$, a property we will denote as *G-symmetry*. In essence, any action of the space group is equivalent to a permutation of the atom positions, fig. 2 visualizes this property using a 2D example. The p4mm symmetry group includes rotations by angles of $\frac{\pi z}{2}$ for $z \in \mathbb{Z}$. The corresponding permutation is invisible, without fabricated labels, because it rearranges the positions of identical shapes. Two Crystals $c_1, c_2 \in \mathcal{C}$ are *Mutually G-Symmetric* if every space group element g corresponds to the same permutation σ on both crystals. Formally, c_1 and c_2 are mutually G-symmetric if $g \cdot c_1 = \sigma \cdot c_1 \iff g \cdot c_2 = \sigma \cdot c_2$. There exist 230 distinct space groups in three-dimensional crystallography. Owing to the intrinsic periodicity of crystal structures, all corresponding subgroups are finite subgroups of the Euclidean group $E(3)$. For non-orthogonal lattice structures, the space group acts on fractional coordinates as elements of the special affine group $SA(3)$, rather than $E(3)$, fig. 5 demonstrates this in 2D. In simpler terms, we apply the group action after mapping every lattice to the cube using $\boldsymbol{L}^{-1}$.

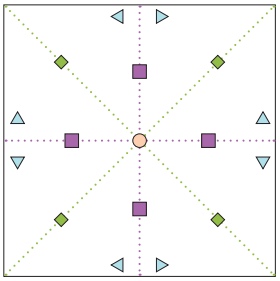

Figure 2: A 2D example of a unit cell with p4mm symmetry. Applying a group element to this set permutes "atoms" of the same type (shape and color). The symmetry divides the unit cell (black box) into four Wyckoff position: the center, horizontal and vertical coordinate axes, diagonal axes, and general position (denoted by white space).

Wyckoff Positions. Intuitively, Wyckoff positions of a space group G indicate regions with specific symmetry properties. Atoms in general position occupy the least symmetric position in the crystal, appear most frequently in the unit cell, and enjoy the fewest restrictions on their coordinates. Meanwhile, atoms in one of the several special positions occupy a region of higher symmetry, appear less frequently in the unit cell, and are restricted to lie in low-dimensional affine subspaces. More formally, a Wyckoff position w, is defined by a set of $|w| = m$ affine projections $\{(\boldsymbol{V}_i, \tau_i)\}_{i=1}^m$ onto the corresponding affine subspace, with $\boldsymbol{V}_i \in \mathbb{R}^{3\times3}$ and $\tau_i \in \mathbb{R}^3$. We denote projection of $f \in [0,1)^3$ onto each of these m affine subspaces by,

$$w(f) \coloneqq \{\boldsymbol{V}_i f + \tau_i - \lfloor \boldsymbol{V}_i f + \tau_i \rfloor\}_{i=1}^m . \tag{1}$$

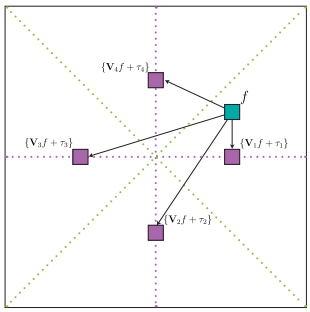

Figure 3: Projection of a coordinate f using Wyckoff position w.

If w is a Wyckoff postion of a space group G, for every $y \in w(f)$ we have the property $w(f) = G \cdot y$ where $G \cdot y \coloneqq \{g \cdot y \mid g \in G\}$ denotes the *orbit* of y under G. This is visualized in fig. 3; it illustrates how f is mapped to an orbit induced by G through w. Each affine transformation $(\boldsymbol{V}_i, \tau_i)$ identifies with a subgroup $G' \leqslant G$ that fixes points on its corresponding image. Namely, $y_i = \boldsymbol{V}_i f + \tau_i - \lfloor \boldsymbol{V}_i f + \tau_i \rfloor$ for some $f \in [0,1)^3$ if and only if $G' = \{g \in G \mid g \cdot y_i = y_i\}$. That means that $G' = G_{y_i}$, the *site-symmetry* (stabilizer) group of y_i. We define the crystal c (with fractional coordinates $\boldsymbol{F}$) to be $\mathcal{W}$*-constructable* with respect to $\mathcal{W} = (w_1, \ldots, w_k)$ if there exist k points $\left\{x_i \in [0,1)^3\right\}_{i=1}^k$ in the unit cell such that $\boldsymbol{F} = \bigcup_{i=1}^k w_i(x_i)$ up to a permutation.

Flow Matching (FM) is a generative modeling framework that transform samples from a simple base distribution p_0 into a complex target distribution p_1 using a time-dependent diffeomorphic

map, called a flow, $\psi : [0,1] \times \mathcal{X} \to \mathcal{X}$. This flow is defined through the differential equation:

$$\frac{\mathrm{d}}{\mathrm{d}t}\psi_t(x) = v_t(\psi_t(x)), \quad \psi_0(x) = x \tag{2}$$

where $v_t : [0,1] \times \mathcal{X} \to \mathcal{X}$ is a vector field governs the evolution of the ψ_t. The flow induces a time-dependent probability density path $p_t : [0,1] \times \mathcal{X} \to \mathbb{R}$ starting at p_0 and ending with p_1. FM trains a parametric approximation u_t of the true vector field v_t (Lipman et al., 2024) by solving a regression objective. A conditional flow $\psi(\cdot \,|\, y) : [0,1] \times \mathcal{X} \to \mathcal{X}$ transports the entire base distribution to a single target point $y \in \mathcal{X}$ and is governed by vector field $v_t(\cdot \,|\, y)$, which is easy to compute, unlike $v_t(\cdot)$. Lipman et al. (2022) demonstrated that optimizing u_t to match $v_t(\cdot \,|\, y)$ with regression leads to the same optimum as matching the marginal vector field v_t.

3 METHOD

This section presents our proposed model, SGFM, a flow matching-based generative approach designed to sample crystal structures conditioned on specified space groups and Wyckoff positions. We start by formalizing the problem and defining the target distribution we aim to sample from in section 3.1. Then, section 3.2 outlines sufficient conditions for a flow model to sample from this distribution. Section 3.3 provides a detailed overview of the model, including its key components, the noise prior and vector field. Finally, section 3.4 presents the training details of SGFM.

3.1 PROBLEM DEFINITION

Given a finite set of crystals $\{c_1, \ldots, c_m\}$, each associated with a space group and Wyckoff positions $(G_i, \mathcal{W}_i)$ and drawn from an unknown target distribution q, our goal is to design a generative model that samples crystals $c \sim p_1$ such that $p_1 \approx q$. To incorporate the structural information encoded by G and $\mathcal{W}$, we factorize the distribution as $p_1(c) = p_1(c \mid G, \mathcal{W})q(G, \mathcal{W})$, enabling us to model the crystal distribution conditionally on G and $\mathcal{W}$. While G and $\mathcal{W}$ are jointly sampled from empirical distribution, the proposed generative model focuses on sampling crystals from the conditional distribution $c \sim p_1(\cdot \mid G, \mathcal{W})$, which satisfies the following properties:

- $p(\cdot \,|\, G, \mathcal{W})$ is a *G-symmetric* distribution, meaning that $p(c \,|\, G, \mathcal{W}) > 0$ if c is G-symmetric. The distribution assigns positive probability exclusively to crystals for which G is their corresponding space group.
- $p(\cdot \,|\, G, \mathcal{W})$ is a *$\mathcal{W}$-constructable* distribution, meaning that $p(c \,|\, G, \mathcal{W}) > 0$ if c is $\mathcal{W}$-constructable. p only supports crystals with fractional coordinates constructible by $\mathcal{W}$.

Our analysis is based upon the general conditions for invariant sampling in flow models (Köhler et al., 2020; Rezende et al., 2019). Next, we will briefly revisit and extend those results to our setting.

3.2 THEORETICAL ANALYSIS

We now present the theoretical concepts underpinning the development of our generative model. The following theorem establishes the conditions under which a flow-based model can sample from a G-invariant distribution. This result has been proven in prior work by Köhler et al. (2020) and Song et al. (2023). For completeness, we state the theorem here and provide a concise version of its proof in appendix A.1, as it serves as a foundational component for our theoretical analysis.

Theorem 3.1. *The probability path $p_t(x)$ defined by a flow generated by a G-equivariant vector field u_t from a G-invariant prior p is G-invariant for all $t \in [0,1]$.*

The proof of this theorem consists of two parts. First, we demonstrate that the flow $\psi_t(x)$, is G-equivariant. Second, we show $p_t(x)$ is G-invariant. Building on this proof, we derive the conditions under which a flow-based model can sample from a distribution that is both G-symmetric and $\mathcal{W}$-constructable. Achieving this requires extending the standard framework with two modifications:

1. Introducing a noise prior that is itself G-symmetric and $\mathcal{W}$-constructable.
2. Ensuring that the vector field model u_t is equivariant with respect to the space group G and the permutation group S_n. To do so we extend the vector field (and corresponding flow) equivariance to the group product $G \times S_n$.

Algorithm 1 Sample $\boldsymbol{F}_0 \sim p_0(\cdot|G, \mathcal{W})$

Input: $\mathcal{W} = \{w_1, \ldots, w_k\}$ s.t. w_i is a Wyckoff position of the space group G.
Output: $\boldsymbol{F}_0 \in [0,1)^{n\times 3}$ s.t. $n = \sum_{i=1}^{k} |w_i|$.
set $\boldsymbol{F}_0 = [\,]$
for $i = 1$ to k **do**
  Sample $x \sim \mathcal{U}\,[0,1]^3$
  $\boldsymbol{F}_0 = \text{Concatenate}([\boldsymbol{F}_0, w_i(x)])$
end for
return $\boldsymbol{F}_0$

Theorem 3.2. *The probability path $p_t(x)$ defined by a flow generated by a $G\times S_n$ equivariant vector field u_t from a G-symmetric and $\mathcal{W}$-constructable prior p is G-symmetric and $\mathcal{W}$-constructable for all $t \in [0,1]$.*

This theorem follows directly from the lemma below (proof in appendix A.2) , which establishes that if the initial point c_0 of a $G \times S_n$ equivariant flow is a G-symmetric crystal structure, any point along the flow will be mutually G-symmetric with c_0. Furthermore, we demonstrate that the flow preserves the site-symmetry structure of c_0. This implies that if c_0 is $\mathcal{W}$-constructable, then $\psi_t(c_0)$ is also $\mathcal{W}$-constructable. Figure 1 (b) visualizes the core idea behind the theorem, illustrating how the equivariant vector field constrains atoms to move solely within the image of their Wyckoff position.

Lemma 3.3. *Let ψ_t be a $G \times S_n$ equivariant flow and $c \in \mathcal{C}$ be G-symmetric and $\mathcal{W}$-constructable, then $\psi_t(c)$ is G-symmetric and $\mathcal{W}$-constructable.*

3.3 SGFM

In this section, we introduce the key components of SGFM, with emphasis on the prior model and the learned vector field architecture. We explain how the previously outlined conditions are concretely implemented. The main focus lies in the interaction between the method and the crystal's fractional coordinates, due to their strong dependence on the space group and Wyckoff positions.

Noise Prior. According to theorem 3.2, for the generated distribution to satisfy the conditions outlined in section 3.1, the noise prior must also satisfy the same constraints. Algorithm 1 presents a noise prior sampling pseudocode (with 2D visualizations in fig. 1 (a)) that generates initial fractional coordinates compliant with these requirements. The sampling procedure for the lattice parameters and atom types is described in section 3.4 for the reader's convenience. The algorithm iterates over the set of Wyckoff positions and samples orbits induced by G by projecting random points from the unit cell. It follows directly from the algorithm's construction that $\boldsymbol{F}_0$ is $\mathcal{W}$-constructable and the following lemma further establishes that $\boldsymbol{F}_0$ is also G-symmetric. The proof (appendix A.3) relies on the fact that the action of a group element on an orbit defines a bijection.

Lemma 3.4. *Let $G \leqslant E(3)$ be a space group, $\mathcal{W}$ a corresponding set of Wyckoff positions and $\boldsymbol{F}_0 \in [0,1)^{n\times 3}$ a $\mathcal{W}$-constructable set of points in the unit cell, then $\boldsymbol{F}_0$ is G-symmetric.*

Space Group Conditional Vector Field. As noted previously, u_t must be $G \times S_n$ equivariant in order for the flow to be G-symmetric and $\mathcal{W}$-constructable. The challenge lies in using a single u_t model across crystals with varying space groups. Since some space groups (with non-orthogonal lattice structure) act on the fractional coordinates with special affine structured transformations, (see fig. 5), using an $E(3) \times S_n$-equivariant model is inadequate. To address these limitations, we adopt Group Averaging (GA) (Yarotsky, 2022), a symmetrization operator that projects a backbone model onto the space of G-equivariant functions. It is defined as:

$$\hat{u}(c \,|\, G) = \sum_{g \in G} g \cdot u(g^{-1} \cdot c). \tag{3}$$

Applying GA to enforce space group symmetry addresses the previously mentioned limitations. Specifically, if the backbone u is S_n equivariant , then $\hat{u}(\cdot \,|\, G)$ is $G \times S_n$ equivariant (Puny et al., 2021). Additionally, GA is not limited to subgroups of $E(3)$ and support all finite groups. Importantly, Puny et al. (2021) showed that symmetrization preserves the expressive power of the original model. However, a drawback of GA is its computational burden: directly applying eq. (3) increases

the number of evaluations of u by a factor of $|G|$, which can be as large as 192 for 3D space groups (the average space group size in the MP-20 dataset is ~ 45). To mitigate this computationally intensive formulation, we leverage the fact that the inputs to $\hat{u}(\cdot \mid G)$ are G-symmetric crystals, allowing us to derive an efficient and equivalent formulation of GA specific to this case.

Lemma 3.5. *Let $c \in C$ be a crystal, G its space group, and u an S_n equivariant vector field. Then, eq.* (3) *can be equivalently rewritten as follows:*

$$\hat{u}(c \mid G) = \sum_{g \in G} g \cdot \sigma_{g^{-1}|c} \cdot u(c) \tag{4}$$

Where $\sigma_{g^{-1}|c} \in S_n$ satisfy the equation $\sigma_{g^{-1}|c} \cdot c = g^{-1} \cdot c$.

The formulation presented in eq. (4) requires only a single evaluation of u, which dramatically improves the model efficiency. Figure 1 (c) compares between eq. (3) and eq. (4) and visualize the efficiency gain. Furthermore, computing $\sigma_{g^{-1}|c}$ is computationally efficient, since we can decompose the problem according to the orbits of c, determined by $\mathcal{W}$ (appendix A.3). At inference time, these permutations only need to be computed once for c_0, since theorem 3.2 guarantees that the flow preserves G-symmetry structure. During training, permutations are computed only once during preprocessing for every data point. Table 1 compares the training and generation runtimes between SGFM, a non-equivariant variant (no symmetrization), and standard GA highlighting the efficiency gains of our GA implementation compared to the standard GA, and demonstrating that its computational cost is comparable to using a backbone without symmetrization. Further details of this comparison are in appendix H.

Table 1: Training and generation time comparison of different vector field models.

Model	Training		Generation
	Batch size	Time (s)	Time (s)
SGFM	64	28.2	17.81
Non-Equivariant	64	26.3	16.39
GA	1	600	-

3.4 Training SGFM

This section provides an overview of the SGFM training process. Let $c_1 \in \mathcal{C}$ be a crystal from the training set with a corresponding space group G and Wyckoff positions $\mathcal{W}$. We will denote $c_0 \sim p_0(\cdot|G, \mathcal{W}, c_1)$ a sample from the conditional noise prior, $c_t = \psi_t(c_0 \mid c_1)$ the conditional flow where $c_t = (k_t, \boldsymbol{F}_t, \boldsymbol{A}_t)$, $v_t(c_t|c_1) = (v_t^k(c_t|c_1), v_t^{\boldsymbol{F}}(c_t|c_1), v_t^{\boldsymbol{A}}(c_t|c_1))$ is the conditional vector field and $\hat{u}_t(c_t|G) = (\hat{u}_t^k(c_t|G), \hat{u}_t^{\boldsymbol{F}}(c_t|G), \hat{u}_t^{\boldsymbol{A}}(c_t|G))$ is the prediction of the $G \times S_n$ equivariant vector field parametric model.

Lattice Parameters. As noted in section 2, we represent lattice parameters using the group-conditioned form from (Jiao et al., 2024), where $k \in \mathbb{R}^6$ encodes the basis coefficients of a 3D symmetric matrix constrained to G-specific subspaces. To sample $k_0 \in \mathbb{R}^6$, we first draw coefficients $k' \sim \mathcal{N}(0, \boldsymbol{I})$ and apply a group condition mask: $k_0 = k' \odot m(G)$, where $m(G) \in \{0, 1\}^6$ is a group-dependent binary mask that zeros out the irrelevant coefficients. k_t is computed as a linear interpolation of k_0 and k_1, $k_t = (1 - t)k_0 + tk_1$ and the corresponding component of the conditional vector field is $v_t^k(c_t|c_1) = k_1 - k_0$. Since the group action does not directly act on the lattice parameterization we need to apply the $m(G)$ on $\hat{u}_t^k(c_t|G)$, both in training and after each generation step. The lattice optimization objective is:

$$\mathcal{L}^k(\theta) = \mathbb{E}_{t, q(c_1), p_0(c_0|G)} \left\| \hat{u}_t^k(c_t|G) \odot m(G) - (k_1 - k_0) \right\|_2^2 \tag{5}$$

Atom Types. For the DNG task, which involves predicting atom types, we follow the modeling approach introduced in Miller et al. (2024), where atom types are represented using a $\{-1, 1\}$ binary format instead of standard one-hot encoding. Specifically, $A_1 \in \{0, 1\}^{n \times h}$ is converted into its binary representation $\tilde{A}_1 \in \{-1, 1\}^{n \times \lceil \log_2 h \rceil}$. To ensure G-symmetry in the initial sample c_0, atom types must be consistent within each orbit. Accordingly, we sample initial Gaussian noise $\mathcal{N}(0, 1)^{\lceil \log_2 h \rceil}$ per orbit and broadcast it to all atoms within that orbit to sample $\tilde{A}_0 \in \{-1, 1\}^{n \times \lceil \log_2 h \rceil}$. We define $\tilde{A}_t = (1 - t)\tilde{A}_0 + t\tilde{A}_0$ and $v_t^A(c_t|c_1) = \tilde{A}_1 - \tilde{A}_0$. The atom types optimization objective is:

$$\mathcal{L}^A(\theta) = \mathbb{E}_{t, q(c_1), p_0(c_0|G)} \left\| \hat{u}_t^A(c_t|G) - (\tilde{A}_1 - \tilde{A}_0) \right\|_2^2 \tag{6}$$

Our GA formulation (eq. (4)) ensures that $\hat{u}_t^A(c_t \mid G)$ is G-invariant, meaning the atom type vector field is consistent across orbits, as required. During inference, we apply the sign function to convert the continuous atom type predictions into their binary representation.

Fractional Coordinates. Algorithm 1 describes a general procedure for sampling fractional coordinates that are both G-symmetric and $\mathcal{W}$-constructable. To ensure G-symmetry of the conditional flow, the initial coordinates $\boldsymbol{F}_0 \sim p_0(\cdot, |, G, \mathcal{W})$ must be G-symmetric with $\boldsymbol{F}_1$. This requires that the order of elements and operators in $\mathcal{W}$ match that used to generate $\boldsymbol{F}_1$, which we precompute during preprocessing using the PyXtal library (Fredericks et al., 2021). We adopt the flat torus geometry of the unit cell, following the approach proposed by Miller et al. (2024), and define the conditional flow over the fractional coordinates -

$$\psi_t(\boldsymbol{F}_0|\boldsymbol{F}_1) = \boldsymbol{F}_0 + t \cdot \log_{\boldsymbol{F}_0}(\boldsymbol{F}_1) \tag{7}$$

Where $\log_{(\cdot)}(\cdot)$ (eq. (12)) is the element-wise logarithmic map over the flat tori. In appendix B we demonstrate: (1) the conditional vector field $\log_{\boldsymbol{F}_0}(\boldsymbol{F}_1)$ is G-equivariant but with respect to a different representation of G. Let $g \in G$ then $\log_{g\cdot\boldsymbol{F}_0}(g \cdot \boldsymbol{F}_1) = g^\star \log_{\boldsymbol{F}_0}(\boldsymbol{F}_1)$ where $g^\star$ is defined by the homomorphism $(\boldsymbol{R}, \tau) \mapsto \boldsymbol{R}$; (2) $\psi_t(\boldsymbol{F}_0|\boldsymbol{F}_1)$ is mutually G-symmetric with $\boldsymbol{F}_0$ and $\boldsymbol{F}_1$, hence G-symmetric and $\mathcal{W}$-constructable. The fractional coordinates optimization objective is:

$$\mathcal{L}^{F}(\theta) = \mathbb{E}_{t,q(c_1),p_0(c_0|G,\mathcal{W})} \left\| \hat{u}_t^{\boldsymbol{F}}(c_t|G) - \log_{\boldsymbol{F}_0}(\boldsymbol{F}_1) \right\|_2^2 \tag{8}$$

Combining all the components we obtain SGFM training objective:

$$\mathcal{L}^{\text{SGFM}}(\theta) = \lambda_k \mathcal{L}^{k}(\theta) + \lambda_{\boldsymbol{F}} \mathcal{L}^{\boldsymbol{F}}(\theta) + \lambda_{\boldsymbol{A}} \mathcal{L}^{\boldsymbol{A}}(\theta) \tag{9}$$

where $\lambda_k, \lambda_{\boldsymbol{F}}, \lambda_{\boldsymbol{A}} \in \mathbb{R}^+$ are hyperparameters.

Table 2: CSP Results over a collection of datasets. MR denotes match rate. Best results are bolded

Model	MP-20		MPTS-52		Perov-5		Carbon-24		Alex-MP-20	
	MR (%) ↑	RMSE ↓	MR (%) ↑	RMSE ↓	MR (%) ↑	RMSE ↓	MR (%) ↑	RMSE ↓	MR (%) ↑	RMSE ↓
CDVAE	33.90	.1045	5.34	.2106	45.31	.1138	17.09	.2969	-	-
FlowMM	61.39	.0566	17.54	.1726	53.15	.0992	23.47	.4122	-	-
OMatG	69.83	.0741	27.38	.1970	83.06	.3753	-	-	72.50	.1260
DiffCSP++	80.27	.0295	46.29	.0896	98.44	.0430	-	-	-	-
GCFM (ours)	**82.74**	**.0288**	**51.79**	**.0827**	**98.57**	**.0188**	**55.02**	**.0952**	**81.98**	**.0243**

4 EXPERIMENTS

The experiments can be divided into two sections: *Crystal Structure Prediction* (CSP) implies predicting the fractional coordinates and lattice parameters given atom types. We will show that conditioning on the correct space groups and Wyckoff positions, enabled by SGFM, has a significantly positive effect on the quality of predicted fractional coordinates and lattice parameters. Wyckoff positions are not known prima facia, so we use a method to predict them. We test on five datasets and perform extensive ablation studies to assess our method. In the second task *De Novo Generation* (DNG), we generate the atom types along with the fractional coordinates and lattice parameters.

Datasets. We evaluate our method on five datasets: *MP*-20 (Jain et al., 2013), with 45,231 diverse crystals from the Materials Project; *MPTS*-52, a time-ordered variant with $40,476$ crystals featuring larger unit cells; and *Alex-MP*-20, a large-scale set of $607,684$ crystals combining MP-20 and Alexandria data (Schmidt et al., 2022a;b). We also assess CSP on two unit-test style datasets: *Perov*-5 (Castelli et al., 2012), with $18,928$ perovskites sharing a common structure but varying atom types, and *Carbon*-24 (Pickard, 2020), containing $10,153$ carbon crystals with diverse structures.

Baselines. We compared SGFM to several state-of-the-art baselines. Methods that do not incorporate space group information in their generation process include *CDVAE* (Xie et al., 2021), *ADiT* (Joshi et al., 2025), *FlowMM* (Miller et al., 2024), *FlowLLM* (Sriram et al., 2024), and *OMatG* (Hoellmer et al., 2025). In contrast, *SymmCD* (Levy et al., 2025), *DiffCSP++* (Jiao et al., 2024), *WyFormer* (Kazeev et al., 2025), and *SGEquiDiff* (Chang et al., 2025) explicitly incorporate space group information. Additional details on each baseline are provided in appendix C.

Model Details. To model $\hat{u}_t$, we adopt the architecture used in Miller et al. (2024), which utilizes *EGNN* (Satorras et al., 2022) to handle fractional coordinates. The model applies sinusoidal embeddings to the fractional coordinates, ensuring invariance to lattice translations in addition to the space group equivariance. A description of the architecture and the hyperparameters used in each experiment are provided in appendix G. For improved sampling quality, we apply inference anti-annealing (Yim et al., 2023; Bose et al., 2023) that adjusts the prediction velocity during generation.

CRYSTAL STRUCTURE PREDICTION

The generative task in CSP requires sampling from the conditional target distribution $c \sim q(\cdot|\boldsymbol{A})$, where $\boldsymbol{A}$ denotes a predefined atom type composition. This conditioning implies that during both training and generation $\boldsymbol{A}_t = \boldsymbol{A}$ for all $t \in [0, 1]$, effectively ignoring the loss term $\mathcal{L}^{\boldsymbol{A}}$ and the atom type component $\hat{u}_t^A(c_t|G)$. For evaluation, a crystal structure is generated for each entry in the test set—which provides empirical samples of a space group, Wyckoff positions, and atom types—and then compared against the corresponding ground truth structure using pymatgen `StructureMatcher` (Ong et al., 2013) with same threshold values as in Jiao et al. (2024). We report two metrics: the match rate (MR), defined as the fraction of generated structures that successfully match their ground truth counterparts, and the RMSE, averaged over all matched pairs. We conduct CSP across all datasets, with results summarized in table 2. Two key conclusions emerge: first, incorporating space group and Wyckoff position information significantly enhances structure generation performance; second, among the methods that leverage this additional information, SGFM consistently outperforms others across all shared datasets, achieving state-of-the-art results. Comparing CSP accuracy as a function of generation steps (fig. 4), we observe that SGFM reaches near-optimal accuracy within just 50–100 steps, whereas DiffCSP++ converges more slowly, requiring up to 1000 steps to approach its best performance—while still showing a notable gap in match rate compared to SGFM.

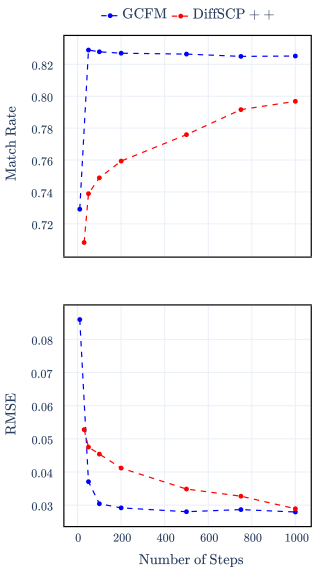

Figure 4: MR (up) and RMSE (down) as a function of generation steps on MP-20.

We ran an experiment to evaluate the impact of SGFM's components. Table 3 compares the CSP evaluations of several models: SGFM, another using a non-equivariant vector field with a uniform noise prior, and a third using a non-equivariant vector field with a conditional noise prior. The results show that starting with a conditional noise prior offers some advantage on its own. Its combination with the equivariant model, however, yields a significant improvement. The missing combination was not evaluated because the uniform distribution does not provide the initial conditions useful for an equivariant vector field.

Table 3: Noise prior vs. model type comparison. The reported result is MR on the MP-20 dataset.

Noise Prior \Model	Equivariant	Non-Equivariant
Wyckoff Conditional	**82.74**	68.16
Uniform	-	64.49

Obtaining the ground truth space group and Wyckoff positions corresponding to a given atom type composition can be challenging. To address this, we leverage *CSPML* (Kusaba et al., 2022), a metric learning-based model that, given an atom type composition, retrieves a similar composition from a template set—along with its associated space group and Wyckoff positions. Specifically, for each crystal in the test set, we match its atom type composition $\boldsymbol{A}$ to a template composition $\boldsymbol{A}'$, from the training set. We then use the space group and Wyckoff positions of $\boldsymbol{A}'$ as conditioning inputs for SGFM. The generated structure is subsequently compared to the ground truth structure from the test set. Table 4 presents a summary of our evaluation using CSPML and a comparison to baselines.

Table 4: CSPML evaluation. The reported metric is MR.

Model	MP-20	MPTS-52	Perov-5
CSPML	70.51	36.98	51.84
DiffCSP++	**70.58**	**37.17**	52.17
SGFM	70.13	35.09	**54.10**

DE NOVO GENERATION

We evaluate the ability of our generative model to discover thermodynamically stable and novel crystals, identify the validity of the generated samples, and investigate divergences in distributions of properties. The results and baselines are shown in table 5. Next, we will describe the various evaluation metrics. *Validity %* defines two different heuristics that realsitic crystals should satsify. *Structural validity* implies that the pairwise atomic distances of a crystal's atoms are all greater than 0.5Å. *Compositional validity* implies that a crystal has a neutral charge according to so-called SMACT (Davies et al., 2019) rules. The *properties* that we consider for computing divergences include ρ the atomic density defined by number of atoms divided by unit cell volume, elem (airity) the number of unique elements in a crystal, and cn (coordination number) or the the number of bonds

Table 5: DNG evaluation. Models were trained on the MP-20 dataset. NFE refers to the number of generation steps per sample. We evaluated all the S.U.N. metrics, except those marked with *.

Model	NFE	Validity (%) ↑		Property ↓			Stable (%) ↑ S.U.N (%) ↑		Stable (%) ↑ S.U.N (%) ↑	
		Structural	Composition	d_ρ	d_{elem}	d_{cn}	$E_{\text{hull}} < 100$ meV/Atom		$E_{\text{hull}} < 0$ meV/Atom	
CDVAE	5000	**100.00**	86.70	0.688	0.278	-	-	-	-	-
ADiT	500	99.74	**92.14**	-	-	-	**72.0**	27.4	13.0	4.6
FlowMM	500	96.86	83.24	0.075	0.079	0.443	31.2	19.7	4.6	2.3
FlowLLM	250	99.81	89.05	0.660	0.090	-	67.9	21.9	14.2	3.6
OMatG	680	95.05	82.84	**0.060**	**0.017**	0.165	44.4	23.7	6.6	2.2
SymmCD	1000	90.34	85.81	0.230	0.400	-	-	-	-	-
DiffCSP++ (empirical)	1000	99.94	85.12	0.235	0.374	-	31.4	21.1	7.2	4.0
DiffCSP++ (Wyformer)	1000	99.66	80.34	0.670	0.098	-	-	-	-	3.8*
SGEquiDiff	1000	99.25	86.16	0.193	0.209	-	-	25.8*	-	-
SGFM (empirical)	500	99.87	86.81	0.075	0.181	**0.076**	64.1	**30.3**	**14.6**	**6.9**
SGFM (Wyformer)	500	99.87	84.76	0.237	0.233	-	48.4	22.6	10.6	4.7

per atom on average. We report the Wasserstein divergence between the test set and a structurally and compositionally valid subset of 1000 generated samples. Finally, we the *thermodynamic stability*, *novelty*, and *uniqueness* of generated crystals. Thermodynamic stability implies a structure is at or near a local minima in composition space. This requires a short explanation which can be read in appendix D. We then compute the uniqueness and novelty of each stable crystal (*S.U.N.*) against other generations and the train and validation set, respectively using `StructureMatcher` (Ong et al., 2013) with default settings. We trained SGFM on the MP-20 dataset, including an atom type prediction module, and generated structures from each of our configurations for evaluation. The configurations include using Wyckoff positions taken from the train set, denoted *empirical*, and from the output of *Wyformer* (Kazeev et al., 2025), with an eponymous denotation. All systems were evaluated with $10,000$ samples, except DiffCSP++ (empirical) that uses only $1,000$ samples. DiffCSP++ (Wyformer) (Kazeev et al., 2025) and SGEquiDiff (Chang et al., 2025) are reported results with slightly different density functional theory settings and only 100 relaxations, respectively.

5 Related Work

There is a growing body of literature about generative models for inorganic crystals. We focus here on works with similar inductive biases, namely explicit utilization of Wyckoff positions. We first consider works that generate atomic coordinates. Cao et al. (2024) created an autoregressive model that generates crystals sequentially in Wyckoff position's lexicographic order. Jiao et al. (2024); Levy et al. (2025) produced diffusion models that both represent crystals within the asymmetric unit, a memory-efficient formulation that contains just one representative per orbit. Neither of these methods utilize space group equivariance and both require projection steps to keep atomic coordinates within the target Wyckoff positions. A concurrently developed diffusion model by Chang et al. (2025) also utilizes the asymmetric unit; however, it does utilize space group equivariance via group averaging. Working in the asymmetric unit does not allow for our efficient reformulation in eq. (4). As written, Levy et al. (2025) do not address the crystal structure prediction problem. There are also a class of models that generate coarse-grained Wyckoff positions alone, ignoring explicit atomic coordinates. (Zhu et al., 2024; Kazeev et al., 2025) both take this approach, inspired by regression methods (Goodall & Lee, 2020; Goodall et al., 2022). These models synergize with ours and generate Wyckoff positions for SGFM to use during de novo generation in section 4. Further discussion of other relevant work is left for appendix E.

6 Conclusions

In this work, we introduced SGFM, a FM based generative model for crystal structures, conditioned on space group and Wyckoff positions. By design, SGFM produces crystals that satisfy symmetry constraints, relying on sufficient conditions we formulated over the noise prior and vector field. We also implemented an efficient group averaging method, enabling the incorporation of space group equivariance into the vector field model with minimal overhead. Evaluated on both CSP and DNG tasks, SGFM achieved state-of-the-art performance. Future directions include extending the model to an unconditional generation setting, where space group and Wyckoff positions are also generated rather than specified.

REFERENCES

Michael S Albergo and Eric Vanden-Eijnden. Building normalizing flows with stochastic interpolants. *arXiv preprint arXiv:2209.15571*, 2022.

Michael S. Albergo, Nicholas M. Boffi, and Eric Vanden-Eijnden. Stochastic interpolants: A unifying framework for flows and diffusions, 2023. URL https://arxiv.org/abs/2303.08797.

Luis Barroso-Luque, Muhammed Shuaibi, Xiang Fu, Brandon M Wood, Misko Dzamba, Meng Gao, Ammar Rizvi, C Lawrence Zitnick, and Zachary W Ulissi. Open materials 2024 (omat24) inorganic materials dataset and models. *arXiv preprint arXiv:2410.12771*, 2024.

Avishek Joey Bose, Tara Akhound-Sadegh, Kilian Fatras, Guillaume Huguet, Jarrid Rector-Brooks, Cheng-Hao Liu, Andrei Cristian Nica, Maksym Korablyov, Michael Bronstein, and Alexander Tong. Se (3)-stochastic flow matching for protein backbone generation. *arXiv preprint arXiv:2310.02391*, 2023.

Zhendong Cao, Xiaoshan Luo, Jian Lv, and Lei Wang. Space group informed transformer for crystalline materials generation, 2024. URL https://arxiv.org/abs/2403.15734.

Ivano E Castelli, David D Landis, Kristian S Thygesen, Søren Dahl, Ib Chorkendorff, Thomas F Jaramillo, and Karsten W Jacobsen. New cubic perovskites for one-and two-photon water splitting using the computational materials repository. *Energy & Environmental Science*, 5(10):9034–9043, 2012.

Rees Chang, Angela Pak, Alex Guerra, Ni Zhan, Nick Richardson, Elif Ertekin, and Ryan P Adams. Space group equivariant crystal diffusion. *arXiv preprint arXiv:2505.10994*, 2025.

Lei Chen, Chandan Setty, Haoyu Hu, Maia G. Vergniory, Sarah E. Grefe, Lukas Fischer, Xinlin Yan, Gaku Eguchi, Andrey Prokofiev, Silke Paschen, Jennifer Cano, and Qimiao Si. Topological semimetal driven by strong correlations and crystalline symmetry. *Nature Physics*, 18 (11):1341–1346, September 2022. ISSN 1745-2481. doi: 10.1038/s41567-022-01743-4. URL http://dx.doi.org/10.1038/s41567-022-01743-4.

Ricky T. Q. Chen and Yaron Lipman. Flow matching on general geometries, 2024. URL https://arxiv.org/abs/2302.03660.

Daniel W Davies, Keith T Butler, Adam J Jackson, Jonathan M Skelton, Kazuki Morita, and Aron Walsh. Smact: Semiconducting materials by analogy and chemical theory. *Journal of Open Source Software*, 4(38):1361, 2019.

Scott Fredericks, Kevin Parrish, Dean Sayre, and Qiang Zhu. Pyxtal: A python library for crystal structure generation and symmetry analysis. *Computer Physics Communications*, 261:107810, April 2021. ISSN 0010-4655. doi: https://doi.org/10.1016/j.cpc.2020.107810. URL http://www.sciencedirect.com/science/article/pii/S0010465520304057.

Colin W Glass, Artem R Oganov, and Nikolaus Hansen. Uspex—evolutionary crystal structure prediction. *Computer physics communications*, 175(11-12):713–720, 2006.

Rhys EA Goodall and Alpha A Lee. Predicting materials properties without crystal structure: Deep representation learning from stoichiometry. *Nature Communications*, 11(1):1–9, 2020.

Rhys EA Goodall, Abhijith S Parackal, Felix A Faber, Rickard Armiento, and Alpha A Lee. Rapid discovery of stable materials by coordinate-free coarse graining. *Science Advances*, 8 (30):eabn4117, 2022.

Philipp Hoellmer, Thomas Egg, Maya M. Martirossyan, Eric Fuemmeler, Zeren Shui, Amit Gupta, Pawan Prakash, Adrian Roitberg, Mingjie Liu, George Karypis, Mark Transtrum, Richard G. Hennig, Ellad B. Tadmor, and Stefano Martiniani. Open materials generation with stochastic interpolants, 2025. URL https://arxiv.org/abs/2502.02582.

Anubhav Jain, Shyue Ping Ong, Geoffroy Hautier, Wei Chen, William Davidson Richards, Stephen Dacek, Shreyas Cholia, Dan Gunter, David Skinner, Gerbrand Ceder, and Kristin A. Persson. The Materials Project: A materials genome approach to accelerating materials innovation. *APL Materials*, 1(1):011002, 07 2013. ISSN 2166-532X. doi: 10.1063/1.4812323. URL https://doi.org/10.1063/1.4812323.

Rui Jiao, Wenbing Huang, Yu Liu, Deli Zhao, and Yang Liu. Space group constrained crystal generation, 2024. URL https://arxiv.org/abs/2402.03992.

Chaitanya K. Joshi, Xiang Fu, Yi-Lun Liao, Vahe Gharakhanyan, Benjamin Kurt Miller, Anuroop Sriram, and Zachary W. Ulissi. All-atom diffusion transformers: Unified generative modelling of molecules and materials, 2025. URL https://arxiv.org/abs/2503.03965.

Nikita Kazeev, Wei Nong, Ignat Romanov, Ruiming Zhu, Andrey Ustyuzhanin, Shuya Yamazaki, and Kedar Hippalgaonkar. Wyckoff transformer: Generation of symmetric crystals, 2025. URL https://arxiv.org/abs/2503.02407.

Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

Scott Kirklin, James E Saal, Bryce Meredig, Alex Thompson, Jeff W Doak, Muratahan Aykol, Stephan Rühl, and Chris Wolverton. The open quantum materials database (oqmd): assessing the accuracy of dft formation energies. *npj Computational Materials*, 1(1):1–15, 2015.

Jonas Köhler, Leon Klein, and Frank Noé. Equivariant flows: Exact likelihood generative learning for symmetric densities. In *Proceedings of the 37th International Conference on Machine Learning*, ICML'20. JMLR.org, 2020.

Jonas Köhler, Michele Invernizzi, Pim De Haan, and Frank Noé. Rigid body flows for sampling molecular crystal structures. In *International Conference on Machine Learning*, pp. 17301–17326. PMLR, 2023.

Georg Kresse and Jürgen Furthmüller. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Physical review B*, 54(16):11169, 1996.

Minoru Kusaba, Chang Liu, and Ryo Yoshida. Crystal structure prediction with machine learning-based element substitution. *Computational Materials Science*, 211:111496, 2022.

Daniel Levy, Siba Smarak Panigrahi, Sékou-Oumar Kaba, Qiang Zhu, Kin Long Kelvin Lee, Mikhail Galkin, Santiago Miret, and Siamak Ravanbakhsh. Symmcd: Symmetry-preserving crystal generation with diffusion models, 2025. URL https://arxiv.org/abs/2502.03638.

Peijia Lin, Pin Chen, Rui Jiao, Qing Mo, Cen Jianhuan, Wenbing Huang, Yang Liu, Dan Huang, and Yutong Lu. Equivariant diffusion for crystal structure prediction. In *Forty-first International Conference on Machine Learning*, 2024. URL https://openreview.net/forum?id=VRv8KjJNuj.

Yaron Lipman, Ricky TQ Chen, Heli Ben-Hamu, Maximilian Nickel, and Matthew Le. Flow matching for generative modeling. In *The Eleventh International Conference on Learning Representations*, 2022.

Yaron Lipman, Marton Havasi, Peter Holderrieth, Neta Shaul, Matt Le, Brian Karrer, Ricky T. Q. Chen, David Lopez-Paz, Heli Ben-Hamu, and Itai Gat. Flow matching guide and code, 2024. URL https://arxiv.org/abs/2412.06264.

Xingchao Liu, Chengyue Gong, and Qiang Liu. Flow straight and fast: Learning to generate and transfer data with rectified flow. *arXiv preprint arXiv:2209.03003*, 2022.

Cécile Malgrange, Christian Ricolleau, and Michel Schlenker. *Symmetry and physical properties of crystals*. Springer, 2014.

Amil Merchant, Simon Batzner, Samuel S Schoenholz, Muratahan Aykol, Gowoon Cheon, and Ekin Dogus Cubuk. Scaling deep learning for materials discovery. *Nature*, pp. 1–6, 2023.

Mila AI4Science, Alex Hernandez-Garcia, Alexandre Duval, Alexandra Volokhova, Yoshua Bengio, Divya Sharma, Pierre Luc Carrier, Michał Koziarski, and Victor Schmidt. Crystal-GFN: sampling crystals with desirable properties and constraints. *AI for Accelerated Materials Design Workshop (NeurIPS)*, 2023.

Benjamin Kurt Miller, Ricky T. Q. Chen, Anuroop Sriram, and Brandon M Wood. Flowmm: Generating materials with riemannian flow matching, 2024. URL https://arxiv.org/abs/2406.04713.

Shyue Ping Ong, William Davidson Richards, Anubhav Jain, Geoffroy Hautier, Michael Kocher, Shreyas Cholia, Dan Gunter, Vincent L Chevrier, Kristin A Persson, and Gerbrand Ceder. Python materials genomics (pymatgen): A robust, open-source python library for materials analysis. *Computational Materials Science*, 68:314–319, 2013.

John P Perdew, Kieron Burke, and Matthias Ernzerhof. Generalized gradient approximation made simple. *Physical review letters*, 77(18):3865, 1996.

Chris J. Pickard. Airss data for carbon at 10gpa and the c+n+h+o system at 1gpa. *Materials Cloud Archive*, 2020.0026/v1, 2020. doi: 10.24435/materialscloud:2020.0026/v1.

Chris J Pickard and RJ Needs. Ab initio random structure searching. *Journal of Physics: Condensed Matter*, 23(5):053201, 2011.

Omri Puny, Matan Atzmon, Heli Ben-Hamu, Edward J Smith, Ishan Misra, Aditya Grover, and Yaron Lipman. Frame averaging for invariant and equivariant network design. *Tenth International Conference on Learning Representations (ICLR)*, 2021.

Danilo Jimenez Rezende, Sébastien Racanière, Irina Higgins, and Peter Toth. Equivariant hamiltonian flows, 2019. URL https://arxiv.org/abs/1909.13739.

James E Saal, Scott Kirklin, Muratahan Aykol, Bryce Meredig, and Christopher Wolverton. Materials design and discovery with high-throughput density functional theory: the open quantum materials database (oqmd). *Jom*, 65(11):1501–1509, 2013.

Victor Garcia Satorras, Emiel Hoogeboom, and Max Welling. E(n) equivariant graph neural networks, 2022. URL https://arxiv.org/abs/2102.09844.

Jonathan Schmidt, Noah Hoffmann, Hai-Chen Wang, Pedro Borlido, Pedro JMA Carriço, Tiago FT Cerqueira, Silvana Botti, and Miguel AL Marques. Large-scale machine-learning-assisted exploration of the whole materials space. *arXiv preprint arXiv:2210.00579*, 2022a.

Jonathan Schmidt, Hai-Chen Wang, Tiago FT Cerqueira, Silvana Botti, and Miguel AL Marques. A dataset of 175k stable and metastable materials calculated with the pbesol and scan functionals. *Scientific Data*, 9(1):64, 2022b.

Yuxuan Song, Jingjing Gong, Minkai Xu, Ziyao Cao, Yanyan Lan, Stefano Ermon, Hao Zhou, and Wei-Ying Ma. Equivariant flow matching with hybrid probability transport, 2023. URL https://arxiv.org/abs/2312.07168.

Anuroop Sriram, Benjamin Kurt Miller, Ricky T. Q. Chen, and Brandon M. Wood. Flowllm: Flow matching for material generation with large language models as base distributions, 2024. URL https://arxiv.org/abs/2410.23405.

Feng Tang, Hoi Chun Po, Ashvin Vishwanath, and Xiangang Wan. Comprehensive search for topological materials using symmetry indicators. *Nature*, 566(7745):486–489, February 2019. ISSN 1476-4687. doi: 10.1038/s41586-019-0937-5. URL http://dx.doi.org/10.1038/s41586-019-0937-5.

Hai-Chen Wang, Silvana Botti, and Miguel AL Marques. Predicting stable crystalline compounds using chemical similarity. *npj Computational Materials*, 7(1):12, 2021.

Peter Wirnsberger, George Papamakarios, Borja Ibarz, Sébastien Racanière, Andrew J Ballard, Alexander Pritzel, and Charles Blundell. Normalizing flows for atomic solids. *Machine Learning: Science and Technology*, 3(2):025009, 2022.

Zhenqin Wu, Bharath Ramsundar, Evan N Feinberg, Joseph Gomes, Caleb Geniesse, Aneesh S Pappu, Karl Leswing, and Vijay Pande. Moleculenet: a benchmark for molecular machine learning. *Chemical science*, 9(2):513–530, 2018.

Tian Xie, Xiang Fu, Octavian-Eugen Ganea, Regina Barzilay, and Tommi S Jaakkola. Crystal diffusion variational autoencoder for periodic material generation. In *International Conference on Learning Representations*, 2021. URL https://arxiv.org/abs/2110.06197.

Jiashi Yang. *An introduction to the theory of piezoelectricity*. Springer, 2005.

Mengjiao Yang, KwangHwan Cho, Amil Merchant, Pieter Abbeel, Dale Schuurmans, Igor Mordatch, and Ekin Dogus Cubuk. Scalable diffusion for materials generation. *arXiv preprint arXiv:2311.09235*, 2023.

Dmitry Yarotsky. Universal approximations of invariant maps by neural networks. *Constructive Approximation*, 55(1):407–474, 2022.

Jason Yim, Andrew Campbell, Andrew YK Foong, Michael Gastegger, José Jiménez-Luna, Sarah Lewis, Victor Garcia Satorras, Bastiaan S Veeling, Regina Barzilay, Tommi Jaakkola, et al. Fast protein backbone generation with se (3) flow matching. *arXiv preprint arXiv:2310.05297*, 2023.

Claudio Zeni, Robert Pinsler, Daniel Zügner, Andrew Fowler, Matthew Horton, Xiang Fu, Sasha Shysheya, Jonathan Crabbé, Lixin Sun, Jake Smith, et al. Mattergen: a generative model for inorganic materials design. *arXiv preprint arXiv:2312.03687*, 2023. URL https://arxiv.org/abs/2312.03687.

Ruiming Zhu, Wei Nong, Shuya Yamazaki, and Kedar Hippalgaonkar. Wycryst: Wyckoff inorganic crystal generator framework. *Matter*, 7(10):3469–3488, 2024.

A Proofs

A.1 Proof of theorem 3.1

Proof. The proof has two main parts. First, we will show that the flow ψ_t defined by the G-equivariant vector field u_t is G-equivariant. Then, we will use this property to demonstrate that the resulting probability path p_t is G-invariant. As a reminder, the flow $\psi : [0,1] \times \mathcal{X} \to \mathcal{X}$ is defined by the following ODE:

$$\frac{d}{dt}\psi_t(x) = u_t(\psi_t(x)) \tag{10}$$

$$\psi_0(x) = x \tag{11}$$

To demonstrate that ψ_t is equivariant, we will show that two functions, $\varphi_t(x) \coloneqq \psi_t(g \cdot x)$ and $\phi_t(x) = g \cdot \psi_t(x)$ (for arbitrary $g \in G$) satisfy the same ODE with identical initial conditions.

$$\frac{d}{dt}\varphi_t(x) = \frac{d}{dt}\psi_t(g \cdot x) = u_t(\psi_t(g \cdot x)) = u_t(\varphi_t(x))$$
$$\varphi_0(x) = \psi_0(g \cdot x) = g \cdot x$$

$$\frac{d}{dt}\phi_t(x) = \frac{d}{dt}g \cdot \psi_t(x) = g \cdot \frac{d}{dt}\psi_t(x) = g \cdot u_t(\psi_t(x)) = u_t(g \cdot \psi_t(x)) = u_t(\phi_t(x))$$
$$\phi_0(x) = g \cdot \psi_0(x) = g \cdot x$$

Where the forth equality uses the G-equivariance of u_t. We can therefore conclude $\psi_t(g \cdot x) = g \cdot \psi_t(x)$ for every $x \in \mathcal{X}, g \in G$ and $t \in [0,1]$, which prove that ψ_t is G-equivariant.

It remains to show that p_t defines an invariant probability path.

$$\begin{aligned} p_t(g \cdot x) &= p_0(\psi_t^{-1}(g \cdot x)) \det \left[\frac{\partial \psi_t^{-1}}{\partial x}(g \cdot x) \right] \\ &= p_0(\psi_t^{-1}(x)) \det \left[\frac{\partial \psi_t^{-1}}{\partial x}(g \cdot x) \right] \\ &= p_0(\psi_t^{-1}(x)) \det \left[g \cdot \frac{\partial \psi_t^{-1}}{\partial x}(x) \cdot g^{-1} \right] \\ &= p_0(\psi_t^{-1}(x)) \det \left[\frac{\partial \psi_t^{-1}}{\partial x}(x) \right] = p_t(x) \end{aligned}$$

The second equality follows from the G-equivariance of ψ_t^{-1} and the G-invariance of p_0. The third equality is a consequence of the definition of the Jacobian matrix for equivariant functions, and the final equality relies on standard properties of the determinant.

A.2 Proof of theorem 3.3

Proof. The first part of the proof, which involves showing that ψ_t is G-symmetric, is straightforward and follows directly from the equivariance properties of ψ_t. Since ψ_t is $G \times S_n$ (and the group actions commute) it trivial to see that it is equivariant to each of the groups separately. Let the $g \in G$ then:

$$g \cdot \psi_t(c) = \psi_t(g \cdot c) = \psi_t(\sigma \cdot c) = \sigma \cdot \psi_t(c)$$

Where the first equality follows from the G-equivariance of ψ_t, the second holds because c is G-symmetric, and the final equality follows from the S_n-equivariance of ψ_t. From the above equation, we also conclude that c and $\psi_t(c)$ are mutually G-symmetric. Now, let $g' \in G_{f_i}$, meaning the g' belongs to the site-symmetry group of f_i, the i^{th} fractional coordinate of c. Since $g' \in G$ there exist a permutation $\sigma' \in S_n$ s.t $g' \cdot c = \sigma' \cdot c$. Moreover, because $g' \in G_{f_i}$, the permutation must fix the index i, $\sigma'(i) = i$. From the previous part of the proof, we know that $g' \cdot \psi_t(c) = \sigma' \cdot \psi_t(c)$ and since $\sigma'(i) = i$, it follows that $g' \in G_{f_i'}$, where f_i' is the i^{th} fractional coordinate of $\psi_t(c)$. Therefore, $\psi_t(c)$ retains the same site-symmetry structure as c, and is thus also $\mathcal{W}$-constructable.

□

A.3 Proof of theorem 3.4

let $g \in G$, our goal is to show that there exists a permutation $\sigma \in S_n$ such that $g \cdot \boldsymbol{F}_0 = \sigma \cdot \boldsymbol{F}_0$. Since F_0 is $\mathcal{W}$-constructable it can be written as a union of orbits under the action of G. Focusing on a single orbit generated by w_i, and denote it as $\boldsymbol{F}_0^{w_i}$ we can observe that $g \cdot \boldsymbol{F}_0^{w_i} = \sigma' \cdot \boldsymbol{F}_0^{w_i}$ for some $\sigma' \in S_{|w_i|}$. This holds because the action of a group element on an orbit is a bijection. Repeating this process for each orbit contained in $\boldsymbol{F}_0$ yields a construction for the permutation σ. □

A.4 Proof of theorem 3.5

Proof.

$$\hat{u}_t(c|G) = \sum_{g \in G} g \cdot u_t(g^{-1} \cdot c) = \sum_{g \in G} g \cdot u_t(\sigma_{g^{-1}|c} \cdot c) = \sum_{g \in G} g \cdot \sigma_{g^{-1}|c} \cdot u_t(c)$$

the second equation comes from the G-symmetry of c and the last comes from the S_n equivariance of u_t. □

B Conditional Flow on Fractional Coordinates

Lemma B.1. *Let G be a space group, the flat tori logarithmic map $\log_{\boldsymbol{F}_0}(\boldsymbol{F}_1)$ is $G \times S_n$ equivariant.*

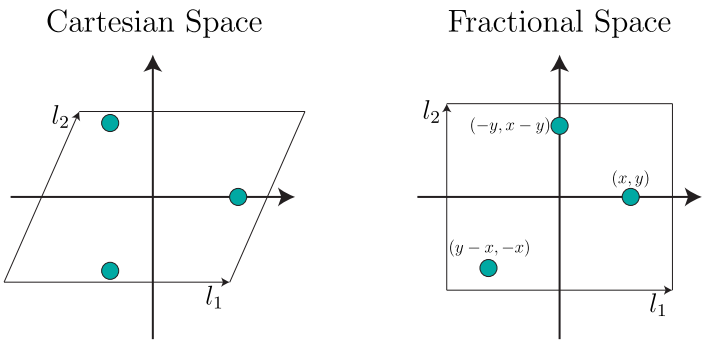

Figure 5: 2D visualization of the deformed geometry induced by moving to fractional coordinates in a non-orthogonal lattices basis. This example demonstrates how a 3-fold rotational space group becomes a set of special affine transformations when acting on fractional coordinates.

Proof. let $g \in G$ such that $g = (\boldsymbol{R}, \tau)$. Since the orthogonal components of G maps the crystal to itself, it preserve the lattice structure. combining with the following lemma (which is expressed with respect to a single point) we get that logarithmic maps is equivariant with respect to the space group and that the representation the acts on the output domain includes only the orthogonal part without the translation. The S_n equivariance is trivial for an element wise function. □

Lemma B.2. *Let $g \in G$ such that $g = (\boldsymbol{R}, \tau)$. if $\boldsymbol{R}$ maps $\mathbb{Z}^3$ to itself $\boldsymbol{R} \log_x(y) = \log_{g \cdot x}(g \cdot y)$.*

Proof. Let $\log_x(y) = v$ that means that exist $z \in \mathbb{Z}^3$ s.t $v = y - x + z$ where $v \in [-\frac{1}{2}, \frac{1}{2})^3$. Now lets assume $\log_{g \cdot x}(g \cdot y) = v'$, that means that there exist $z' \in \mathbb{Z}^3$ s.t $v' = g \cdot y - g \cdot x + z'$. plugging in $g = (\boldsymbol{R}, \tau)$ results in $v' = \boldsymbol{R}(y - x) + z'$. combining both equations we get that $v' = \boldsymbol{R}v + z''$ (because $\boldsymbol{R}z \in \mathbb{Z}^3$). Since $v' \in [-\frac{1}{2}, \frac{1}{2})^3$ and $\|v\| = \|\boldsymbol{R}v\|$ we conclude that $v' = \boldsymbol{R}v$.

□

Lemma B.3. *The conditional flow $\psi_t(\boldsymbol{F}_0|\boldsymbol{F}_1)$ is G-symmetric and $\mathcal{W}$-constructable.*

Proof.

$$\begin{aligned} g \cdot \psi_t(\boldsymbol{F}_0|\boldsymbol{F}_1) &= (\boldsymbol{F}_0 + t \log_{\boldsymbol{F}_0}(\boldsymbol{F}_1))R^T + \mathbf{1}_n \tau^T \\ &= \boldsymbol{F}_0 \boldsymbol{R}^T + \mathbf{1}_n \tau^T + t \log_{\boldsymbol{F}_0}(\boldsymbol{F}_1) R^T \\ &= g \cdot \boldsymbol{F}_0 + t \log_{g \cdot \boldsymbol{F}_0}(g \cdot \boldsymbol{F}_1) \\ &= \sigma \cdot \boldsymbol{F}_0 + t \log_{\sigma \cdot \boldsymbol{F}_0}(\sigma \cdot \boldsymbol{F}_1) \\ &= \sigma \cdot \psi_t(\boldsymbol{F}_0|\boldsymbol{F}_1) \end{aligned}$$

The fact that $\psi_t(\boldsymbol{F}_0|\boldsymbol{F}_1)$ is $\mathcal{W}$-constructable follows directly from the proof in appendix A.2 and the fact the $\psi_t(\boldsymbol{F}_0|\boldsymbol{F}_1)$ is mutually G-symmetric with $\boldsymbol{F}_0$ and $\boldsymbol{F}_1$. □

C Lattice Representation

The lattice matrix $\boldsymbol{L} \in \mathbb{R}^{3\times3}$ characterizes the geometry of the unit cell. When $\boldsymbol{L}$ corresponds to a physically valid lattice, i.e., it has positive volume, it is invertible and can be decomposed to the product $\boldsymbol{L} = \boldsymbol{Q} \exp(\boldsymbol{S})$ where $\boldsymbol{Q} \in \mathbb{R}^{3\times3}$ is an orthogonal matrix and $\boldsymbol{S} \in \mathbb{R}^{3\times3}$ is a symmetric

matrix. Representing the lattice parameters via $\boldsymbol{S}$ enjoys the benefits of orthogonal invariance (any orthogonal transformation is added to $\boldsymbol{Q}$), which makes this representation invariant to any space group operations. Jiao et al. (2024) suggested representing $\boldsymbol{S}$ using the coefficients of the following basis -

$$\boldsymbol{B}_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \boldsymbol{B}_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \boldsymbol{B}_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix},$$
$$\boldsymbol{B}_4 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \boldsymbol{B}_5 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}, \boldsymbol{B}_6 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

This basis enables clustering of the crystallographic space groups based on the basis coefficients used to represent $\boldsymbol{S}$. Table 6 summarizes the lattice and coefficient constraints for each crystal family type.

Table 6: Relationship between the lattice shape and the constraint of the symmetric bases.

Crystal Family	Space Group No.	Lattice Shape	Constraint of Symmetric Bases
Triclinic	$1 \sim 2$	No Constraint	No Constraint
Monoclinic	$3 \sim 15$	$\alpha = \gamma = 90°$	$k_1 = k_3 = 0$
Orthorhombic	$16 \sim 74$	$\alpha = \beta = \gamma = 90°$	$k_1 = k_2 = k_3 = 0$
Tetragonal	$75 \sim 142$	$\alpha = \beta = \gamma = 90°$ $a = b$	$k_1 = k_2 = k_3 = 0$ $k_4 = 0$
Hexagonal	$143 \sim 194$	$\alpha = \beta = 90°, \gamma = 120°$ $a = b$	$k_2 = k_3 = 0, k_1 = -log(3)/4$ $k_4 = 0$
Cubic	$195 \sim 230$	$\alpha = \beta = \gamma = 90°$ $a = b = c$	$k_1 = k_2 = k_3 = 0$ $k_4 = k_5 = 0$

D Density Functional Theory

Crystals exist in competition for stability between alternatives with the same composition. If one plots energy against composition, the lowest energy structures form a *convex hull*. We say a crystal is thermodynamically stable if it is near or below this convex hull. Since we do not know all structures, there is epistemic uncertainty in this characterization. The difference between the energy of a crystal and this convex hull is denoted E_{hull}. We report $E_{\text{hull}} < 100$ meV/atom and $E_{\text{hull}} < 0$ meV/atom rates for stability metrics. These values are computed by prerelaxation with a machine learning interatomic potential (Barroso-Luque et al., 2024) followed by relaxation and energy evaluation using density functional theory.

For the stability metrics, we applied the Vienna ab initio simulation package (VASP) (Kresse & Furthmüller, 1996) to compute relaxed geometries and ground state energies at a temperature of 0 K and pressure of 0 atm. We used the default settings from the Materials Project (Jain et al., 2013) known as the `MPRelaxSet` with the PBE functional (Perdew et al., 1996) and Hubbard U corrections. These correspond with the settings that our prerelaxation network OMat24 (Barroso-Luque et al., 2024) was trained on, so prerelaxation should reduce DFT energy, up to fitting error.

We did *not* make any guesses about oxidation states! This deviates from the Materials Project which does make those guesses. For this reason, our energy above hull calculations for structures that need to consider oxidation state are slightly high, implying that we might be under-predicting stability. This applies to any stability result we calculated. We expect it to also be a negligible effect.

The results from DiffCSP++ WyFormer in table 5 were computed by Kazeev et al. (2025) and differ slightly from ours. Specifically, they run a multiple relaxations to avoid errors that come from using a poor initial guess before relaxation. Since we prerelax with OMat24 we expect that double relaxation is unnecessary. Consult their work for further details, but we believe the differences are negligible for this purpose.

E Related Work

As a continuation from section 5, we discuss other related work. We still limit the focus to the most-relevant parts of this large body of literature.

Our method resembles non-deep learning based methods that propose structures using Wyckoff positions as inductive bias (Glass et al., 2006; Pickard & Needs, 2011) and refine the atomic positions using density functional theory. This field is known as high-throughput screening of inorganic crystals and it is responsible for generating several important datasets of stable materials (Saal et al., 2013; Kirklin et al., 2015; Wang et al., 2021; Schmidt et al., 2022a;b). Recently, those searches have been sped up by machine learning interatomic potentials that closely approximate density functional theory (Merchant et al., 2023).

Now we take a step further away conceptually to discuss methods that are tangentially related to ours. *Crystal-GFN* (Mila AI4Science et al., 2023) is a G-flow network that uses the space group, but does not consider Wyckoff positions. Several other works generate crystals without considering multiple types of atom (Wirnsberger et al., 2022), or molecule (Köhler et al., 2023). Additionally, there is a large and growing cannon of generative models for materials that do not have general space group equivariance (Xie et al., 2021; Yang et al., 2023; Zeni et al., 2023; Miller et al., 2024; Sriram et al., 2024; Lin et al., 2024; Joshi et al., 2025; Hoellmer et al., 2025).

F Baselines

We provide additional context on the core approach behind each baseline we compared against: *CDVAE*, integrates a diffusion model with a variational autoencoder for crystal structure generation; *ADiT*, which use latent-based diffusion model and train on additional information from the QM9 (Wu et al., 2018) dataset; *FlowMM*, an application of Riemannian Flow Matching (Chen & Lipman, 2024) that incorporates non-trivial geometries in the crystal representation space; *FlowLLM*, combines FlowMM with a Large Language model that uses as base distribution generator. *OMatG*, leverages Stochastic Interpolants (Albergo et al., 2023) for material generation; *SymmCD*, operates on the asymmetric unit and incorporates Wyckoff positions as part of the generative process; *DiffCSP++*, a diffusion-based model that conditions on space groups and projects each denoising step through Wyckoff position transformations; *WyFormer*, which employs an autoregressive model to generate Wyckoff positions (conditioned on space group) and subsequently uses DiffCSP++ model for full structure generation to predict the structure; and finally, *SGEquiDiff* a diffusion based model that enforce space group equivariance while working on the asymmetric unit.

G Model Details

G.1 Architecture

In this section, we present a comprehensive overview of our vector field model $\hat{u}_t(\cdot \mid G)$, along with the hyperparameters employed during training and generation across all experiments. The model takes as input a crystal $c = (k, \boldsymbol{F}, \boldsymbol{A})$, where $f^i \in \mathbb{R}^3$ denotes the i^{th} fractional coordinate in $\boldsymbol{F}$, and $a^i \in \{0, 1\}^h$ represents the i^{th} atom type indicator vector in $\boldsymbol{A}$. The forward computation of s

layers model $\hat{u}(c, |, G)$ is defined by the following set of equations:

$$
\begin{aligned}
a^i_{embed} &= \phi^a(a^i) \\
t_{embed} &= \text{SinusoidalTimeEmbedding}(t) \\
h^i_{(0)} &= \phi^{\text{embed}}(\left[a^i_{embed}, t_{embed}\right]) \\
m^{ij}_{(l)} &= \phi^{\text{edge}}_{(l)}(\left[h^i_{(l-1)}, h^j_{(l-1)}, k, \text{SinusoidalEmbedding}(\log_{f^i}(f^j)), \frac{L^T L \log_{f^i}(f^j)}{\left\|L^T L \log_{f^i}(f^j)\right\|}\right]) \\
m^i_{(l)} &= \frac{1}{n}\sum_{j=1}^{n} m^{ij}_{(l)} \\
h^i_{(l)} &= \phi^{\text{node}}_{(l)}(\left[h^i_{(l-1)}, m^i_{(l)}\right]) \\
u^k_t(c_t) &= \phi^k(\frac{1}{n}\sum_{j=1}^{n} h^i_{(s)}) \\
(u^{\boldsymbol{F}}_t(c_t))^i &= \phi^{\boldsymbol{F}}(h^i_{(s)}) \\
(u^{\boldsymbol{A}}_t(c_t))^i &= \phi^{\boldsymbol{A}}(h^i_{(s)}) \\
u_t(c_t) &= (u^k_t(c_t), u^{\boldsymbol{F}}_t(c_t), u^{\boldsymbol{A}}_t(c_t)) \\
\hat{u}_t(c_t \,|\, G) &= \sum_{g \in G} g \cdot \sigma_{g^{-1}|c} \cdot u_t(c_t)
\end{aligned}
$$

We denote d as the hidden dimension of the model, d_t as the Sinusoidal Time Embedding dimension, and d_s as the Sinusoidal Embedding dimension. Next, we list the learnable modules constructing the model and denote their input and output dimension as $x \to y$. ϕ^a is a linear layer $h \to d$, ϕ^{embed} is a linear layer $d + d_t \to d$ dimension d, $\phi^{\text{edge}}_{(l)}$ is 2-layer Multi-layer Perceptron (MLP) $2d + d_t + 9 \to d$, $\phi^{\text{node}}_{(l)}$ is a 2-layer Multi-layer Perceptron (MLP) $2d \to d$, ϕ^k is a linear layer $d \to 6$, $\phi^{\boldsymbol{F}}$ is a linear layer $d \to 3$ and $\phi^{\boldsymbol{A}}$ is a linear layer $d \to h$. The last equation represents the group averaging presented in eq. (4). The flat tori logarithmic map is defined by the equation:

$$
\log_{f^i}(f^j) = \frac{1}{2\pi}\,\text{atan2}(\left[\sin(f^j - f^i), \cos(f^j - f^i)\right]) \tag{12}
$$

Table 7 summarize the hyperparameters used to train our SGFM models. Note that the same configuration was applied uniformly across all datasets and tasks. The hyperparameter search was performed on MP-20 (CSP) and the resulting settings were adopted for all other experiments.

Table 7: Hyperparameter details for all the models reported in the paper. Hyperparameter (bottom row) search was conducted on the MP-20 dataset.

Dataset	Number of Layers	d	d_t	d_s	Activation	Layer Norm
CSP						
MP-20	8	512	256	128	SiLU	✓
MPTS-52	8	512	256	128	SiLU	✓
Carbon-24	8	512	256	128	SiLU	✓
Perov-5	8	512	256	128	SiLU	✓
Alex-MP-20	8	512	256	128	SiLU	✓
DNG						
MP-20	8	512	256	128	SiLU	✓
Hyperparameter Range						
MP-20	$\{6, 7, 8, 9\}$	$\{128, 256, 512\}$	$\{128, 256\}$	$\{128, 256\}$	-	-

Table 8: Training hyperparameter details for all the models reported in the paper. Hyperparameter (bottom row) search was conducted per experiment.

Dataset	Batch Size/GPU	Learning Rate	Epochs	$\lambda_{\boldsymbol{F}}$	$\lambda_{\boldsymbol{A}}$	λ_k
			CSP			
MP-20	64	0.0005	5000	100	-	1
MPTS-52	32	0.0005	5000	100	-	1
Carbon-24	64	0.0005	8000	100	-	1
Perov-5	256	0.0005	1000	100	-	1
Alex-MP-20	32	0.0005	1250	100	-	1
			DNG			
MP-20	64	0.0007	5000	100	1	1
			Hyperparameter Range			
-	-	$\{0.0002, 0.0005, 0.0007\}$	$\{1000, 1250, 2000, 5000, 8000\}$	$\{1, 10, 50, 100\}$	$\{1, 10, 50, 100\}$	$\{1, 10, 50, 100\}$

G.2 Training & Generation

All of our models were trained using the `ADAM` optimizer (Kingma & Ba, 2014) on 8 NVIDIA A100 GPUs. Table 8 outlines the training configuration for each model, including the ranges explored during hyperparameter search. We employed a cosine annealing learning rate schedule with a minimum learning rate of 0.00001. As described in section 4, we applied inference anti-annealing to enhance generation quality. This technique modifies the vector field by scaling it with a time-dependent function $s(t) = 1 + s't$, where $s' \in \mathbb{R}^+$ is a hyperparameter. We defined separate annealing parameters for each crystal component: $s'_{\boldsymbol{F}}$ and s'_k (no annealing was applied to atom type prediction). For the CSP experiments, we set $s'_{\boldsymbol{F}} = 3$, $s'_k = 3$, and for DNG, we used $s'_{\boldsymbol{F}} = 5$, $s'_k = 3$. All datasets have $60/20/20$ train/validation/test split except Alex-MP-20 that has $80/10/10$ split.

H Running Time Ablations

This experimental ablation study aims to evaluate the efficiency of SGFM by comparing its performance during both training and generation against two baseline models: (1) a non-equivariant variant where the vector field does not incorporate group symmetry, and (2) a standard GA implementation as defined in eq. (3). For each model, we measured the time required to train a single epoch on MP-20, as well as the time needed to generate a batch of 64 (with 500 generation steps). The training time was averaged over 10 epochs, while the generation time was averaged over 100 batches. Training was conducted on an NVIDIA RTX8000 using 8 GPUs, while generation was performed on a single NVIDIA A10 GPU. The results are summarized in table 1. Due to memory constraints, the standard GA model was limited to a maximal batch size of 1 per GPU. As training under these conditions was not feasible, we do not report generation timings for this model. The results highlight a significant efficiency gap between the SGFM implementation of GA and the standard version, while showing minimal difference compared to the non-equivariant model.

I Large Language Models

We utilized of large language models (LLMs) to assist with language refinement and proofreading. No content, ideas, or analyses were produced by these tools. The usage was quite limited.