

# STRUCTEVAL: A BENCHMARK FOR EVALUATING GENERATION, INFERENCE, AND RECONSTRUCTION IN ATOMIC AND CRYSTALLINE STRUCTURES

006 **Anonymous authors**

007 Paper under double-blind review

## ABSTRACT

013 Crystalline materials power technologies from solar conversion to catalysis, yet  
 014 current machine learning evaluations artificially divide infinite lattices and fi-  
 015 nite nanoclusters into separate domains. StructEval unifies these regimes with a  
 016 symmetry-aware, radius-resolved framework that systematically links primitive  
 017 unit cells to nanoparticles across ten industrially critical compounds. Each material  
 018 includes 20+ radii configurations between 0.6–3.0 nm, sampled at 780 quasi-  
 019 uniform orientations, producing nearly 200,000 structures spanning 55 – 11,298  
 020 atoms. StructEval defines three rigorous challenges–unit cell → nanoparticle gener-  
 021 ation, nanoparticle → unit cell resolution, and lattice reconstruction–supported by  
 022 a leakage-free split that isolates orientation interpolation from radius extrapolation.  
 023 This design enables precise measurement of generalization across scale and sym-  
 024 metry. Benchmarking leading generative models reveals severe breakdowns under  
 025 out-of-distribution conditions, exposing a fundamental gap in current architectures.  
 026 By providing a reproducible, geometry-grounded testbed, StructEval establishes  
 027 the foundation for next-generation generative, inference, and reconstruction mod-  
 028 els in crystalline systems. Data and implementations are released at <https://anonymous.4open.science/r/StructEval-ANONYMOUS>.

## 1 INTRODUCTION

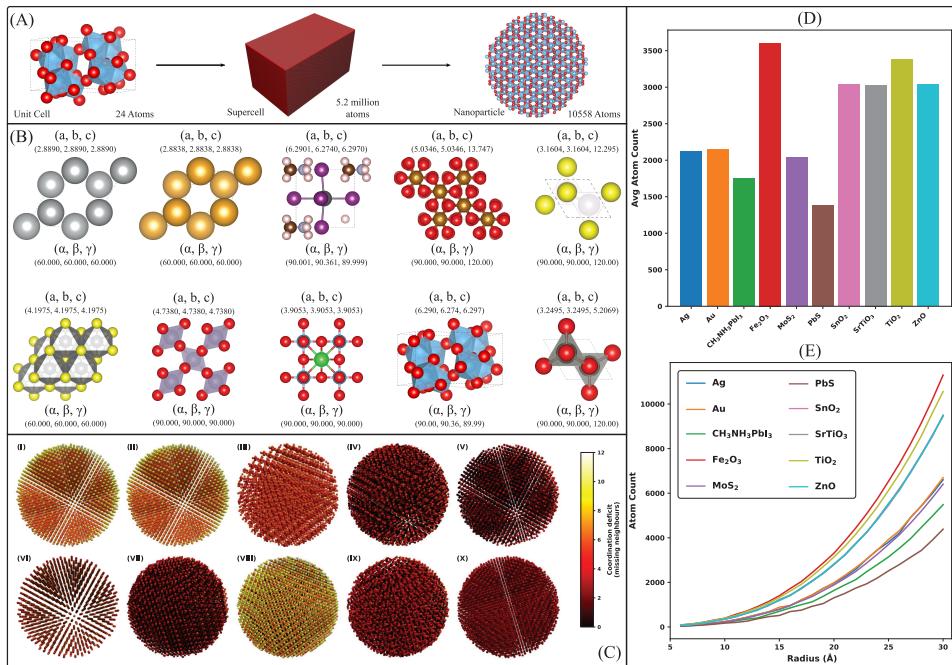
031 Nanostructured materials enable applications from photovoltaics to sensing Simonov & Goodwin  
 032 (2020), with behavior governed both by the periodicity of a primitive unit cell and by the finite  
 033 morphologies that emerge in nanoparticles Cao et al. (2022). These regimes are typically studied  
 034 separately: crystal-growth simulations extend ideal lattices, whereas nanoparticle workflows build  
 035 finite clusters refined by empirical or ab-initio methods Levi & Kotrla (1997). However, such  
 036 simulations are computationally expensive and difficult to scale across compositions, sizes, and  
 037 orientations Surek (2005). First-principles approaches such as DFT Orio et al. (2009) and DFTB  
 038 Elstner & Seifert (2014) provide accurate energetics and reconstructions, but DFT’s cubic cost limits  
 039 large-scale exploration Cohen et al. (2008); DFTB partially alleviates this Liu et al. (2019); Qi et al.  
 040 (2013).

041 Machine learning (ML) offers a scalable alternative Carleo et al. (2019); Karniadakis et al. (2021);  
 042 Alpaydin (2021). GNNs including SchNet Schütt et al. (2018), DimeNet++ Gasteiger et al. (2020),  
 043 TorchMD-Net Thölke & De Fabritiis (2022), and CGNNs Cheng et al. (2021) achieve strong molecular  
 044 and crystalline prediction performance on benchmarks such as QM9 Ruddigkeit et al. (2012);  
 045 Ramakrishnan et al. (2014) and MD17 Chmiela et al. (2017). Equivariant models—E(n)-GNNs  
 046 Satorras et al. (2021), SE(3)-Transformers Fuchs et al. (2020), PaiNN Schütt et al. (2021), SphereNet  
 047 Coors et al. (2018), Equiformer Liao & Smidt (2022), FAENet Duval et al. (2023), and GotenNet  
 048 Aykent & Xia (2025)—improve data efficiency by enforcing geometric symmetries, while multimodal  
 049 approaches Polat et al. (2024); Rollins et al. (2024); Das et al. (2023) integrate diverse inputs.  
 050 Generative models such as CDVAE Xie et al. (2021), GraphDF Luo et al. (2021), UniMat Yang et al.  
 051 (2023), and CrystalFlow Luo et al. (2024), along with recent symmetry-preserving diffusion and flow  
 052 models Jiao et al. (2024); Levy et al. (2025); Kelvinius et al. (2025), advance crystal generation but  
 053 are primarily evaluated on small molecules or bulk periodic crystals. No existing benchmark links

054 primitive-cell inputs to radius- and orientation-controlled nanoparticles or tests generalization under  
 055 realistic shifts in scale and symmetry.

056 StructEval addresses this gap with a symmetry-aware, radius-resolved dataset connecting primitive  
 057 unit cells to systematically generated nanoparticles. For ten technologically relevant materials (Ag,  
 058  $\text{CH}_3\text{NH}_3\text{PbI}_3$ ,  $\text{Fe}_2\text{O}_3$ ,  $\text{MoS}_2$ ,  $\text{PbS}$ ,  $\text{SnO}_2$ ,  $\text{SrTiO}_3$ ,  $\text{TiO}_2$ ,  $\text{ZnO}$ ), the dataset provides each unit  
 059 cell, 25 radii from 0.6–3.0 nm, and 780 quasi-uniform orientations, yielding  $\sim$ 200,000 reproducible  
 060 structures containing 55–11,298 atoms. StructEval defines three tasks: (1) unit-cell  $\rightarrow$  nanoparticle  
 061 generation, (2) nanoparticle  $\rightarrow$  lattice inference, and (3) partial-structure reconstruction. A leakage-  
 062 free split ensures interpolation over orientations (ID) and extrapolation across radii (OOD) are cleanly  
 063 separated. Evaluations of modern generative models—including DiffCSP Jiao et al. (2023), FlowMM  
 064 Miller et al. (2024), FlowLLM Sriram et al. (2024), MatterGen-MP Zeni et al. (2023), and ADiT Joshi  
 065 et al. (2025)—show consistent degradation under OOD shifts, highlighting challenges in symmetry  
 066 preservation, scale generalization, and coherent reconstruction. By unifying bulk crystallography  
 067 with nanoparticle geometry, StructEval provides a rigorous, reproducible platform for benchmarking  
 068 generative and inference models under realistic structural distribution shifts.

069 The remainder of the paper is organized as follows. Section 2 presents a review of prior work on crys-  
 070 talline building blocks and nanostructured morphologies, including first-principles and semi-empirical  
 071 simulation methods, as well as existing benchmarks for crystals and nanomaterials. Section 3 details  
 072 the construction and characteristics of the StructEval dataset. Section 4 discusses the performance of  
 073 the SOTA generative models on StructEval across three distinct tasks. Section 5 outlines the current  
 074 limitations of StructEval dataset. Section 6 summarizes the key findings and proposes directions for  
 075 future research.



100 Figure 1: Detailed overview of StructEval. (A) Workflow for generating radius-resolved nanoparticles.  
 101 (B) Unit cells of the materials—Ag, Au,  $\text{CH}_3\text{NH}_3\text{PbI}_3$ ,  $\text{Fe}_2\text{O}_3$ ,  $\text{MoS}_2$ ,  $\text{PbS}$ ,  $\text{SnO}_2$ ,  $\text{SrTiO}_3$ ,  
 102  $\text{TiO}_2$ , and  $\text{ZnO}$ —arranged left to right. The lattice constants  $a$ ,  $b$  and  $c$  denote the cell edge lengths, while  
 103 the angles  $\alpha$ ,  $\beta$  and  $\gamma$  specify the inter-edge angles between  $b$ – $c$ ,  $a$ – $c$  and  $a$ – $b$ , respectively. (C) Coordination-deficit heatmaps for the largest nanoparticle of each material (panels I–X follow the  
 104 same material order). (D) Mean atom count per material, illustrating typical nanoparticle sizes. (E)  
 105 Atom count versus nanoparticle radius, revealing how total atom number scales with size.

## 2 RELATED WORK

## 2.1 CRYSTALLINE BUILDING BLOCKS AND NANOSTRUCTURED MORPHOLOGIES

Crystalline unit cells encode the symmetry operations and atomic motifs that ML models must learn to generalize, providing the periodic, coordinative, and orientational patterns from which larger structures derive. When bulk crystals are truncated into finite clusters, their shapes follow orientation-dependent surface energies described by the Gibbs–Wulff theorem Li et al. (2016); Barmparis et al. (2015); Ringe et al. (2013). In StructEval, these effects matter not for their thermodynamic meaning but because changes in radius, orientation, and faceting produce controlled, symmetry-consistent deviations from ideal bulk geometry Kittel & McEuen (2018), forming principled distribution shifts for evaluating geometric generalization. Quantum-confinement-related optical or catalytic phenomena Bera et al. (2010) lie outside the benchmark’s scope; only the crystallographic information needed to produce meaningful radius-resolved nanoclusters is retained. By using deterministic, symmetry-preserving construction rather than simulation-driven morphologies, StructEval offers reproducible and physically grounded structural variation tailored for ML benchmarking Yang et al. (2022).

## 2.2 FIRST-PRINCIPLES AND SEMI-EMPIRICAL SIMULATION METHODS

Physics-based nanoparticle generation methods—including Kohn–Sham DFT Bickelhaupt & Baerends (2000); Yu et al. (2016), linear-scaling approaches such as ONETEP Baer & Head-Gordon (1997); Skylaris et al. (2005), semi-empirical techniques like DFTB Zheng et al. (2005); Spiegelman et al. (2020); Baćić et al. (2020); Kim et al. (2019), and classical or ML-based interatomic potentials Daw & Baskes (1984); Behler & Parrinello (2007); Mahata et al. (2022)—accurately capture energetics and surface reconstructions but face steep size limitations. Standard DFT’s  $\mathcal{O}(N^3)$  scaling restricts simulations to  $\sim 10^3$  atoms, and even reduced-scaling or semi-empirical models cannot reliably generate structures near StructEval’s 11,300-atom upper bound. Because StructEval focuses on geometric and symmetry-based reasoning rather than energetics, fully physics-based generation is unnecessary and computationally infeasible. Instead, deterministic symmetry-preserving construction provides scalable, reproducible structures that maintain crystallographic motifs while enabling systematic variation of radius and morphology. This rationale explains why StructEval employs geometric rather than DFT-, DFTB-, or force-field–driven workflows and highlights that geometric generalization—not simulation fidelity—is the benchmark’s core objective Kurban et al. (2024).

## 2.3 BENCHMARKS FOR CRYSTALS AND NANOMATERIALS

Benchmark datasets have propelled ML advances in quantum chemistry and materials science by supplying standardized structures and labels. Beyond those noted in Section 1, QM7 Blum & Reymond (2009); Rupp et al. (2012) provides atomization energies for small organic molecules, MD22 Chmiela et al. (2023) offers force-labeled molecular dynamics trajectories, and PubChemQC Kim et al. (2025) furnishes millions of ground-state geometries with electronic properties. NablADFT Khrabrov et al. (2022) and QH9 Yu et al. (2024) target Hamiltonian matrix estimation, while Perov-5 Castelli et al. (2012a;b) includes 18,928  $ABX_3$  perovskites and Carbon-24 Pickard (2020) catalogs over 10,000 low-energy carbon frameworks generated by AIRSS and relaxed via DFT. For inorganic solids, the Materials Project underpins MatBench’s fourteen property benchmarks Dunn et al. (2020), and the OC20 Chanussot et al. (2021) and OC22 Tran et al. (2023) datasets provide force-labeled metal and oxide surface structures. LAMBench Peng et al. (2025) aggregates diverse atomic geometries for large-scale pretraining, while CrysMTM Polat et al. (2025) introduces temperature-dependent nanomaterial configurations with multimodal annotations. Despite this breadth, existing datasets focus mainly on bulk crystals or small molecules; radius-graded nanoclusters typically appear only incidentally in catalytic or surface studies, and symmetry-aware or bidirectional tasks remain rare. No public benchmark systematically couples primitive-cell representations to radius- and orientation-resolved nanoclusters across chemistries and size regimes, nor supports tasks such as nanoparticle generation or lattice inference from finite clusters—motivating the development of StructEval.

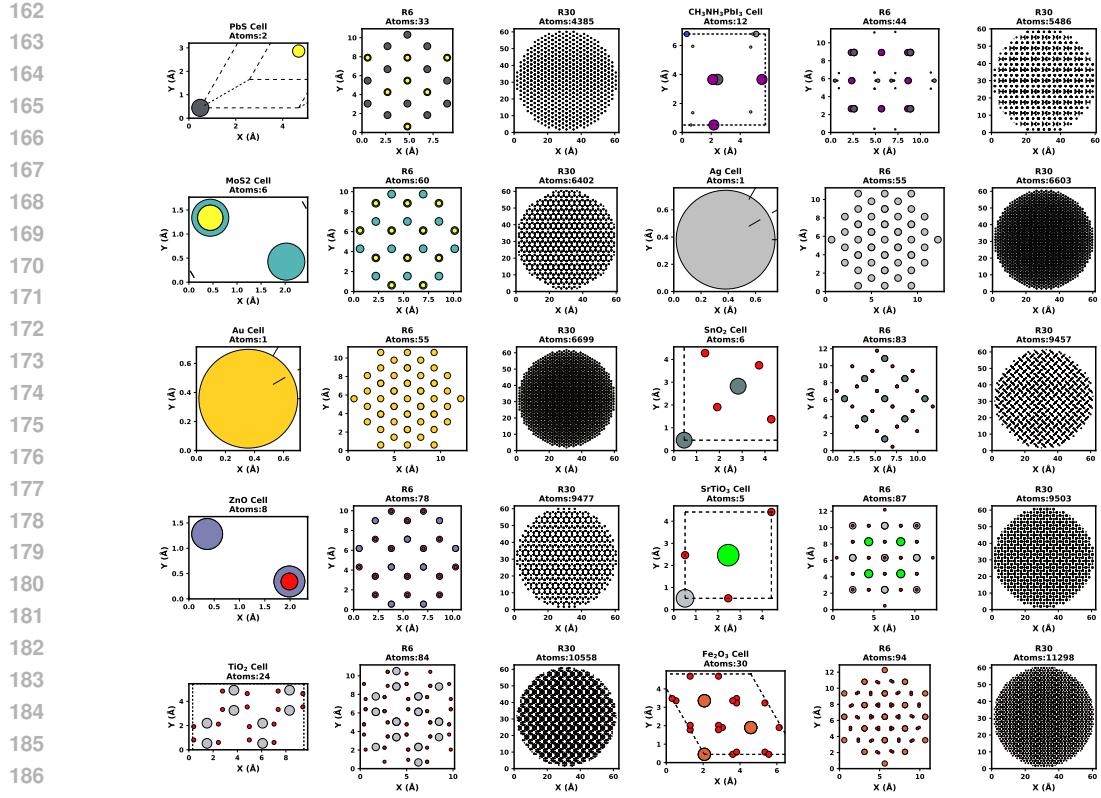


Figure 2: From primitive cell to radius-controlled nanoclusters. For each material in dataset, the panels show—left to right—the primitive unit cell followed by its canonical  $R = 6 \text{ \AA}$  and  $R = 30 \text{ \AA}$  nanoparticles. Materials are arranged from top to bottom in ascending order of the atom count in their  $R30$  cluster, illustrating how coordination environments and bulk-like cores emerge with increasing radius. All views share a common Ångström scale. Atom colours follow the conventional CPK palette: Ag (light grey), Au (gold), C (dark grey), H (white), N (blue), O (red), S (yellow), Fe (orange-brown), Mo (teal), Sn (silver), Sr (green), Ti (slate grey), Zn (forest green), Pb (dark grey-blue), and I (purple).

### 3 DATASET CONSTRUCTION AND TASK DEFINITION

#### 3.1 CRYSTAL STRUCTURES AND NANOPARTICLE GENERATION

The STRUCTEVAL suite spans elemental solids, perovskites, transition-metal dichalcogenides, and binary oxides. The selected compounds include Silver (Ag) King (2002b), Gold (Au) King (2002a), methylammonium lead iodide ( $\text{CH}_3\text{NH}_3\text{PbI}_3$ ) Walsh et al. (2019), hematite ( $\text{Fe}_2\text{O}_3$ ) Finger & Hazen (1980), molybdenum disulfide ( $\text{MoS}_2$ ) Wyckoff (1963a); Grau-Crespo & Lopez-Cordero (2002), galena ( $\text{PbS}$ ) Wyckoff (1963b), cassiterite ( $\text{SnO}_2$ ) Baur et al. (1971), strontium titanate ( $\text{SrTiO}_3$ ) Mitchell & Carpenter (2000), anatase titanium dioxide ( $\text{TiO}_2$ ) Horn et al. (1972), and zincite zinc oxide ( $\text{ZnO}$ ) Wyckoff (1963c). Primitive unit cells were extracted from CIFs and serve as the basis for nanoparticle generation and symmetry-aware modeling tasks. Lattice parameters, space groups, and atomic positions are provided in Appendix A, while Figure 2 summarizes dataset structure and key statistics.

**Supercell construction.** Let  $N = 60$ . We define the periodic lattice of atomic sites as

$$\Lambda = \left\{ n_1 \mathbf{a}_1 + n_2 \mathbf{a}_2 + n_3 \mathbf{a}_3 + \mathbf{r}_j \mid n_1, n_2, n_3 \in \{0, \dots, N-1\}, j \in \{1, \dots, M\} \right\}, \quad (1)$$

where  $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3 \in \mathbb{R}^3$  are the primitive lattice vectors of the unit cell, and  $\{\mathbf{r}_j\}_{j=1}^M \subset \mathbb{R}^3$  are the basis positions of the  $M$  atoms inside each unit cell. The integers  $n_1, n_2, n_3 \in \{0, 1, \dots, N-1\}$

216 serve as lattice indices, specifying the unit cell position along the three lattice directions, and may  
 217 equivalently be viewed as elements of  $\mathbb{Z}_N$ . Replicating the unit cell  $N$  times along each direction  
 218 yields an  $N \times N \times N$  supercell, and the set  $\Lambda$  in equation 1 collects all atomic sites contained within  
 219 it.

220

221 **Reference atom.** Choose a central reference position  $\mathbf{r}_0 \in \Lambda$  (e.g., the image of  $\mathbf{r}_1$  in the central  
 222 unit cell defined by equation 1).

223

224 **Nanoparticle definition.** For a given cutoff radius  $R$ , define the nanoparticle as

225 
$$P(R) = \left\{ \mathbf{x} \in \Lambda \mid \|\mathbf{x} - \mathbf{r}_0\|_2 \leq R \right\}, \quad (2)$$

227 i.e., all lattice sites within a sphere of radius  $R$  around  $\mathbf{r}_0$ .

228

229 **Radius sampling.** Select  $K = 25$  radii on a uniform grid:

230 
$$R_k = R_{\min} + (k - 1) \Delta R, \quad k = 1, \dots, K, \quad (3)$$

231 with end-points and step size

232 
$$R_{\min} = 0.6 \text{ nm}, \quad R_{\max} = 3.0 \text{ nm}, \quad \Delta R = \frac{R_{\max} - R_{\min}}{K - 1} = 0.1 \text{ nm}. \quad (4)$$

233 Relationships captured by the equations 2–4 define the complete set  $\{P(R_k)\}_{k=1}^{25}$ , spanning quantum-  
 234 confined clusters ( $R \leq 1.0$  nm) to near-bulk fragments ( $R \geq 2.0$  nm).

235

### 236 3.2 QUASI-UNIFORM ROTATIONS

237

238 Deterministic, quasi-uniform rotations ensure that each nanoparticle orientation is represented without  
 239 randomness, enabling reproducible data augmentation and unbiased evaluation of rotation-equivariant  
 240 models. By constructing a finite, low-discrepancy *set* on  $\text{SO}(3)$  that includes the identity and  
 241 approximates uniform coverage, one can systematically generate rotated copies of any atomic  
 242 configuration Yershova et al. (2010). Let  $\mathcal{P} = \{\mathbf{r}_1, \dots, \mathbf{r}_n\} \subset \mathbb{R}^3$  denote a finite point cloud (with  $n$   
 243 points). Our goal is to build a finite set  $\mathcal{G} \subset \text{SO}(3)$  containing the identity that provides near-uniform,  
 244 reproducible sampling of rotation space. The construction proceeds in three steps:

245

246 **Axis–angle parametrization.** Every rotation  $R \in \text{SO}(3)$  is represented as

247 
$$R = R(\hat{\mathbf{u}}, \phi), \quad (5)$$

248 where  $\hat{\mathbf{u}} \in \mathbb{S}^2$  is a unit axis and  $\phi \in [0, \pi]$  is the rotation angle, with the equivalence  $(\hat{\mathbf{u}}, \phi) \sim$   
 249  $(-\hat{\mathbf{u}}, 2\pi - \phi)$  ensuring a unique representation. Using Rodrigues’ formula Dai (2015), this rotation  
 250 is expressed as

251 
$$R(\hat{\mathbf{u}}, \phi) = I_3 \cos \phi + (1 - \cos \phi) \hat{\mathbf{u}} \hat{\mathbf{u}}^\top + \sin \phi [\hat{\mathbf{u}}]_\times, \quad (6)$$

252 where  $I_3$  is the  $3 \times 3$  identity matrix,  $\hat{\mathbf{u}} = (u_1, u_2, u_3)^\top$ , and

253 
$$[\hat{\mathbf{u}}]_\times = \begin{pmatrix} 0 & -u_3 & u_2 \\ u_3 & 0 & -u_1 \\ -u_2 & u_1 & 0 \end{pmatrix} \quad (7)$$

254 is the skew-symmetric matrix implementing the linear map  $\mathbf{v} \mapsto \hat{\mathbf{u}} \times \mathbf{v}$ , which guarantees that  
 255  $R(\hat{\mathbf{u}}, \phi)$  in equation 6 is orthogonal and has determinant 1.

256

257 **Quasi-Uniform Grid on  $\text{SO}(3)$ .** We define a deterministic sampling scheme via the Hopf fibration  
 258 Lyons (2003):

259 
$$(\hat{\mathbf{u}}, \psi) \longmapsto R(\hat{\mathbf{u}}, 2\psi), \quad \hat{\mathbf{u}} \in \mathbb{S}^2, \psi \in [0, \pi).$$

260 To discretize this space, we proceed as follows.

261

262 1. **Fibonacci lattice on  $\mathbb{S}^2$ .** For  $k = 0, \dots, N_{\text{axis}} - 1$  set

263 
$$z_k = 1 - \frac{2k + 1}{N_{\text{axis}}}, \quad r_k = \sqrt{1 - z_k^2}, \quad \varphi_g = \frac{1 + \sqrt{5}}{2}, \quad \varphi_k = 2\pi k \varphi_g^{-1},$$

264

265 and define the axis  $\hat{\mathbf{u}}_k = (r_k \cos \varphi_k, r_k \sin \varphi_k, z_k)$ .

270 2. **Irrational angle increments.** For  $m = 0, \dots, M_{\text{ang}} - 1$ , let  
 271

$$272 \quad \psi_m = 2\pi m \varphi_g^{-1} \bmod 2\pi.$$

273 3. **Rotation-set construction.** Define  $R_{k,m} = R(\hat{\mathbf{u}}_k, 2\psi_m)$  and let  $\mathcal{G}$  be the set consisting of  
 274 the identity and all  $R_{k,m}$ . Because  $\psi_0 = 0$  yields  $R = I$ , we count identity once, hence  
 275

$$276 \quad |\mathcal{G}| = 1 + N_{\text{axis}}(M_{\text{ang}} - 1).$$

277 **Quasi-Uniform Coverage (sketch).** As  $N_{\text{axis}}, M_{\text{ang}} \rightarrow \infty$ , the induced empirical measure of the  
 278 above low-discrepancy set provides increasingly uniform coverage of  $\text{SO}(3)$  with respect to the Haar  
 279 measure (in the sense of vanishing discrepancy), leveraging the equidistribution of the Fibonacci  
 280 lattice on  $\mathbb{S}^2$  Stanley (1975) and the incommensurate twist increments John (1998).  
 281

282 **Rigid-body action on the point cloud.** Let  $\mathcal{P} = \{\mathbf{r}_1, \dots, \mathbf{r}_n\} \subset \mathbb{R}^3$  be a finite point cloud and  
 283  $\mathbf{c} = \frac{1}{n} \sum_{i=1}^n \mathbf{r}_i$  be its center of mass. For each  $R \in \mathcal{G}$ , define the rigidly rotated configuration  
 284

$$285 \quad \mathcal{P}_R = \{R(\mathbf{r}_i - \mathbf{c}) + \mathbf{c} \mid i = 1, \dots, n\}, \quad (8)$$

287 which preserves all inter-point distances while re-centering about  $\mathbf{c}$ . To ensure numerical robustness,  
 288 a tolerance  $\varepsilon > 0$  is used when applying equation 8: near-zero angles are snapped, degenerate axes  
 289 are regularized, and infinitesimal perturbations break floating-point degeneracies. This deterministic  
 290 augmentation yields a near-uniform sampling of orientations from  $\mathcal{G}$  without introducing stochastic  
 291 noise.

292 **Configuration diversity.** Although individual nanoparticles may exhibit high symmetry, the Carte-  
 293 sian product of (i) ten distinct chemistries/crystal families, (ii) 25 radii spanning quantum-confined to  
 294 near-bulk regimes, and (iii)  $|\mathcal{G}| = 780$  deterministic orientations produces  $\approx 200,000$  unique atomic  
 295 configurations, capturing rich variation in surface terminations, coordination statistics, facet exposure,  
 296 and global morphology—especially for non-cubic/anisotropic lattices.  
 297

### 298 3.3 DATA SPLITS

300 For each material  $m$ , we enumerate all nanoparticle instances as

$$301 \quad \mathcal{G}_m = \{(R, b) \mid R \in \{R6, \dots, R30\}, b \in \{0, \dots, 780\}\},$$

303 covering every radius–orientation pair. We treat radii R6–R24 as in-distribution and R25–R30 as  
 304 out-of-distribution, then apply a random split (70%/10%/20%) to the in-distribution subset to obtain  
 305  $\text{Train}_m$ ,  $\text{Val}_m$ , and  $\text{IDTest}_m$ , with  $\text{OODTest}_m$  defined by the held-out radii. Grouping all orientations  
 306  $(R, b)$  together ensures rotation-invariant partitions and prevents orientation leakage.

307 **Physical grounding and scope.** Nanoparticles are generated by deterministic spherical truncation of  
 308 experimentally characterized bulk crystals, preserving local crystallography while breaking global  
 309 periodicity. We omit thermodynamic or electronic labels and avoid large-scale ab-initio relaxations to  
 310 keep the benchmark geometry- and symmetry-focused, fully reproducible at  $\approx 200,000$  structures.  
 311 Because DFT scales cubically and our clusters reach  $\approx 10^3$ – $10^4$  atoms, exhaustive first-principles  
 312 labeling is infeasible; instead, we rely on a symmetry-aware geometric pipeline that cleanly supports  
 313 ID/OOD splits. **On rotations.** Rotations are central to defining orientation generalization: for each  
 314 radius  $R_k$ , we construct non-overlapping train/ID/OOD  $\text{SO}(3)$  grids with decreasing angular spacing  
 315 ( $9^\circ$ ,  $6^\circ$ ,  $3^\circ$ ) using a greedy quaternion scheme and store rotated copies indexed by their quaternion.  
 316 We do not apply Kabsch or ICP alignment, since alignment removes the requirement to predict  
 317 absolute orientation and would inflate scores for models that produce consistently rotated—but not  
 318 correctly oriented—structures.

### 319 3.4 TASKS

321 The dataset spans three tasks evaluated with metrics that capture both geometric and crystallographic  
 322 fidelity. Core geometric measures include RMSD, Hausdorff distance Rucklidge (1997), convex-hull  
 323 volume ratio error, radial-distribution-function divergence, and validity rate. Task 2 additionally  
 evaluates lattice-parameter error, space-group accuracy, and their joint accuracy. Task 3 further

324 includes surface-atom recall and coordination-histogram divergence. All tasks rely on the material-  
 325 specific splits  $\text{Train}_m$ ,  $\text{Val}_m$ ,  $\text{IDTest}_m$ , and  $\text{OODTest}_m$  to separately assess orientation interpolation  
 326 and radius extrapolation. Figure 3 summarizes the three tasks, with detailed metric definitions  
 327 provided in Appendix B.  
 328

329 **Task 1: Unit Cell → Nanoparticle Generation.** From a primitive unit cell  $u_m$  and a target radius  
 330  $R$ , the model must generate realistic nanoparticles that faithfully reproduce both the infinite lattice  
 331 periodicity and the finite-size surface morphology. Mastering this capability is essential for producing  
 332 computationally efficient nanostructure models that accurately capture surface-driven phenomena,  
 333 thereby enhancing predictive accuracy in material design:

$$f_1 : (u_m, R) \mapsto P \subset \mathbb{R}^3. \quad (9)$$

334 The mapping in equation 9 is assessed with the metrics  $\text{RMSD}(P, P^*)$ ,  $d_{\text{Haus}}(P, P^*)$ ,  $\Delta V_{\text{hull}}(P, P^*)$ ,  
 335  $E_{\text{RDF}}(P, P^*)$ , and  $V_R(P)$ .  
 336

337 **Task 2: Nanoparticle → Lattice Inference.** Local geometric information extracted from finite  
 338 nanoparticles is used to infer their lattice constants and symmetry group, exploring whether such  
 339 short-range cues suffice to reconstruct extended periodicity and crystallographic invariants. Achieving  
 340 this capability is vital for bridging localized observations with global structural order, thereby enabling  
 341 accurate property predictions and reliable materials-design workflows:  
 342

$$f_2 : P \mapsto (\ell, g), \quad \ell = (a, b, c, \alpha, \beta, \gamma), \quad g \in \mathcal{S}, \quad (10)$$

343 so that equation 10 maps each nanoparticle  $P$  to its lattice parameters  $\ell$  and a space-group label  
 344  $g$  drawn from the set  $\mathcal{S}$  of crystallographic groups. Evaluation metrics are then  $\text{RMSE}(\ell, \ell^*)$ ,  
 345  $\mathbb{W}[g = g^*]$ , and the joint accuracy  $\mathbb{W}[(\ell, g) = (\ell^*, g^*)]$ , all assessed with respect to the ground-truth  
 346 pair  $(\ell^*, g^*)$ .  
 347

348 **Task 3: Lattice Reconstruction.** Partial occlusion of nanoparticles—whether applied at random  
 349 or biased toward surface atoms—tests the model’s capability to reconstruct the complete atomic  
 350 arrangement from incomplete inputs. Such evaluation reveals whether the model can infer missing  
 351 structural information and preserve chemical fidelity, a critical requirement for reliable nanoparticle  
 352 modeling and downstream materials-design applications:  
 353

$$f_3 : P_{\text{mask}} \mapsto P, \quad P_{\text{mask}} = M \cdot P, \quad (11)$$

354 where  $M$  in Eq. (11) is a masking operator that removes a fraction  $p$  of atoms. Evaluation metrics  
 355 include  $\text{RMSD}(P, P^*)$ ,  $R_{\text{surf}}(P, P^*)$ ,  $D_{\text{KL}}(h(P) \parallel h(P^*))$ , and  $V_R(P, P^*)$ .  
 356

## 357 4 EXPERIMENTS

358 Multiple generative models—such as DiffCSP, FlowMM, FlowLLM, MatterGen-MP, and ADiT—are  
 359 evaluated on all three tasks using the splits defined in Section 3.4. Implementation details and  
 360 hyperparameters are provided in Appendix C, with comprehensive task-specific results in Appendix D.  
 361 To clarify task solvability, we note that Task 1 (unit cell → nanoparticle) is deterministic under  
 362 spherical truncation of the periodic lattice; Task 2 (nanoparticle → lattice inference) remains solvable  
 363 because local neighborhoods preserve crystallographic invariants such as coordination shells, bond-  
 364 angle distributions, and stoichiometry; and Task 3 (partial reconstruction) has a clear lower bound,  
 365 as oracle-style baselines that fill masked regions via nearest-neighbor lattice completion achieve  
 366 substantially lower errors than learned models, indicating significant remaining headroom.  
 367

### 368 4.1 TASK 1: UNIT CELL TO NANOPARTICLE GENERATION

369 Task 1 evaluates nanoparticle generation from a primitive unit cell and target radius equation 9,  
 370 requiring models to capture both lattice periodicity and finite-size morphology. As shown in Table 1,  
 371 no method excels across normalized RMSD, RDF KL divergence, and volume ratio in either ID  
 372 or OOD settings: MatterGen-MP attains the lowest RMSD but suffers from large volume ratios;  
 373 FlowLLM achieves the best  $\text{RDF}_{\text{KL}}$  but with high RMSD; and DiffCSP and FlowMM provide mid-  
 374 range RMSD while lagging on other metrics. OOD performance is evaluated using leakage-free splits  
 375

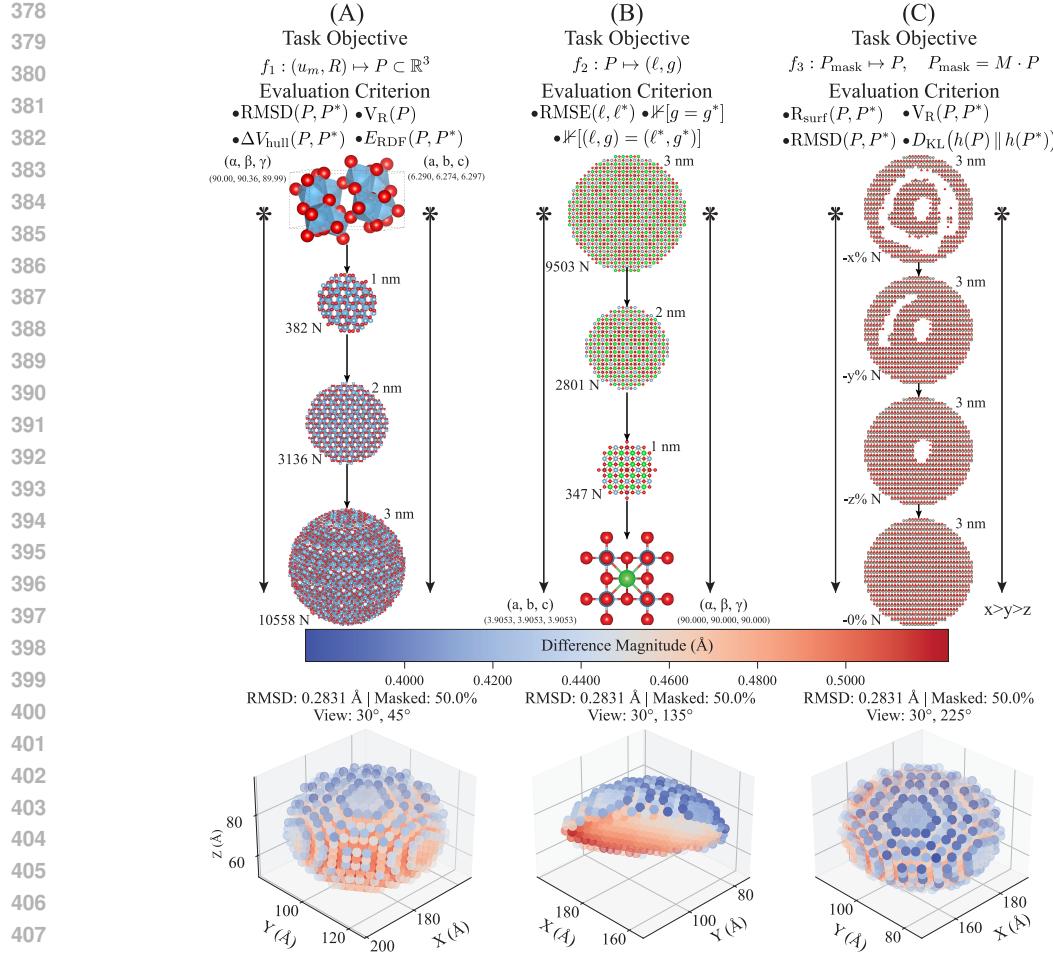


Figure 3: Overview of benchmark tasks and lattice reconstruction experiment result. (A) shows generation of nanoparticles starting from the unit cell. (B) includes the generating unit cell configurations from the nanoparticles. Lastly, (C) illustrates lattice reconstruction from partially masked configurations. N is the atom count in the system. First material is  $\text{TiO}_2$ , second is  $\text{SrTiO}_3$  while the last one is  $\text{ZnO}$ . Scatter plots in the bottom are different angle reconstruction results from CDVAE model for a single material on OOD set with masked 50% of the original atoms.

that exclude entire radius ranges and deploy disjoint orientation grids, ensuring models are tested on true scale and orientation extrapolation. The consistent OOD degradation across architectures indicates limitations in scale awareness, multi-scale feature modeling, and long-range symmetry handling. Training times further reveal large efficiency gaps—FlowLLM is orders of magnitude slower per epoch, while faster models such as MatterGen-MP and DiffCSP do not achieve the best overall accuracy. These results underscore fundamental trade-offs between geometric fidelity, physical realism, and computational efficiency, and point to the need for explicit radius conditioning, hierarchical representations, and stronger geometric inductive biases.

## 4.2 TASK 2: NANOPARTICLE TO LATTICE INFERENCE

Task 2 evaluates lattice-parameter and space-group inference from finite nanoparticles equation 10, testing whether local geometry is sufficient to recover global crystallographic invariants. As shown in Table 2, models such as DiffCSP and ADiT achieve relatively low lattice-parameter RMSEs but perform poorly on symmetry classification, with all space-group accuracies below 40% and joint accuracy effectively 0%. For example, ADiT attains an ID RMSE of 55.240 Å, 29.6% space-group accuracy, and 0.0% joint accuracy; OOD results remain similar at 55.246 Å, 29.1%, and 0.0%.

432 Because each material’s space group is constant across radii, these accuracies reflect per-material  
 433 label learning, while the stable RMSE shows that lattice-parameter prediction generalizes to larger  
 434 clusters. The persistent 0% joint accuracy highlights the core difficulty: nanoparticles preserve  
 435 only local symmetry, losing global periodicity under spherical truncation, making simultaneous  
 436 recovery of lattice parameters and exact space group inherently ambiguous. OOD generalization  
 437 remains limited, and training-time disparities further emphasize model constraints—FlowLLM is  
 438 computationally prohibitive without accuracy gains, whereas faster models like DiffCSP and ADiT  
 439 still fail to infer symmetry reliably. Overall, current architectures partly recover unit-cell geometry  
 440 but cannot translate local structure into correct global symmetry, which is precisely the challenge  
 441 Task 2 is designed to expose.

### 442 443 4.3 TASK 3: LATTICE RECONSTRUCTION

444 Task 3 evaluates a model’s ability to reconstruct full nanoparticle structures from partially observed  
 445 inputs equation 11, using random and surface-biased masking to test whether missing atoms can be  
 446 recovered while maintaining structural and chemical fidelity. As shown in Table 3, performance is  
 447 uniformly poor: RMSD remains high, surface-atom recall is near zero, and validity rates rarely exceed  
 448 0%. Even the strongest models—FlowMM and FlowLLM—offer only marginal improvements,  
 449 failing to restore missing surface atoms or preserve local chemistry; RDF KL divergence also  
 450 stays large across all methods, indicating that none recover the underlying crystalline order from  
 451 incomplete data. MatterGen-MP and ADiT show slightly lower divergence in isolated cases, but no  
 452 model demonstrates consistent or robust reconstruction. Training-time comparisons further reveal  
 453 substantial inefficiencies: ADiT is faster but not more accurate, while FlowLLM is both slow and  
 454 ineffective. Overall, greater model capacity or compute does not overcome the inherent difficulty  
 455 of nanoparticle completion, and the persistent failures highlight why Task 3 serves as a stringent  
 456 benchmark for partial-structure recovery; an illustrative reconstruction error map is shown in Figure 3.

## 457 458 5 LIMITATIONS

460 While StructEval bridges bulk symmetry and finite nanoclusters, it has limitations: it covers only  
 461 ten materials currently; our clusters are crystallographically grounded fragments—deterministic  
 462 truncations of experimentally validated bulk structures—but they are not thermodynamically validated  
 463 or passivated, and they omit temperature, disorder, and environmental effects. This is intentional:  
 464 by decoupling geometric/symmetry reasoning from energetic confounders, the benchmark provides  
 465 a clean, deterministic testbed for generation, inference, and reconstruction under strict ID/OOD  
 466 controls. We acknowledge that experimental nanoparticles can undergo surface reconstructions and  
 467 environment-dependent stabilization.

## 468 469 6 CONCLUSION AND FUTURE WORK

471 Existing benchmarks emphasize either periodic crystals or small molecules, leaving the transitional  
 472 bulk-to-nanoscale regime largely unaddressed. StructEval fills this gap with a radius-resolved,  
 473 symmetry-aware dataset linking primitive unit cells to nanoparticles sampled across 25 radii and 780  
 474 quasi-uniform orientations, supporting three bidirectional tasks—generation, inference, and recon-  
 475 struction—under strict leakage-free splits that probe both interpolation and extrapolation. Although  
 476 the benchmark focuses on real-space geometric and symmetry reasoning, it remains fully compatible  
 477 with XRD-style structure-solution tasks Guo et al. (2025), as each nanoparticle configuration can be  
 478 used to simulate nanocrystalline diffraction patterns. Across tasks, state-of-the-art models perform  
 479 reasonably in-distribution but degrade sharply out-of-distribution, exhibiting increased recon-  
 480 struction errors, symmetry misclassifications, and completion failures. Since DFT scales cubically with  
 481 system size, direct simulation of large nanoparticles is impractical, and surrogate models that fail  
 482 to generalize cannot reliably replace it for predicting lattice parameters or morphologies. These  
 483 findings highlight the limitations of relying solely on local geometric cues and motivate approaches  
 484 with explicit symmetry constraints, multi-scale hierarchies, and targeted augmentation. Designed  
 485 as a modular platform with planned extensions to molecular crystals, perovskites, polymers, and  
 multi-phase systems, StructEval unifies bulk and nanoscale regimes and establishes a foundation for  
 developing models capable of capturing the full structural complexity of crystalline nanomaterials.

486 REFERENCES  
487

- 488 Ethem Alpaydin. *Machine Learning*. MIT press, 2021.
- 489 Sarp Aykent and Tian Xia. GotenNet: Rethinking Efficient 3D Equivariant Graph Neural Networks.  
490 In *The Thirteenth International Conference on Learning Representations*, 2025. URL <https://openreview.net/forum?id=5wxCQDtbMo>.
- 491
- 492 Vladimir Bačić, Thomas Heine, and Agnieszka Kuc. Analytical approach to phonon calculations in  
493 the scc-dftb framework. *The Journal of Chemical Physics*, 153(14), 2020.
- 494
- 495 Roi Baer and Martin Head-Gordon. Sparsity of the density matrix in kohn-sham density functional  
496 theory and an assessment of linear system-size scaling methods. *Physical Review Letters*, 79(20):  
497 3962, 1997.
- 498
- 499 Georgios D Barmparis, Zbigniew Lodzianna, Nuria Lopez, and Ioannis N Remediakis. Nanopar-  
500 ticle shapes by using wulff constructions and first-principles calculations. *Beilstein Journal of*  
501 *Nanotechnology*, 6(1):361–368, 2015.
- 502
- 503 W. H. Baur, R. A. Sass, et al. The rutile structure of sno<sub>2</sub>. *Acta Crystallographica Section B*, 27:2133,  
504 1971.
- 505
- 506 Jörg Behler and Michele Parrinello. Generalized neural-network representation of high-dimensional  
507 potential-energy surfaces. *Physical Review Letters*, 98(14):146401, 2007.
- 508
- 509 Debasis Bera, Lei Qian, Teng-Kuan Tseng, and Paul H Holloway. Quantum dots and their multimodal  
510 applications: a review. *Materials*, 3(4):2260–2345, 2010.
- 511
- 512 F Matthias Bickelhaupt and Evert Jan Baerends. Kohn-sham density functional theory: predicting  
513 and understanding chemistry. *Reviews in Computational Chemistry*, pp. 1–86, 2000.
- 514
- 515 L. C. Blum and J.-L. Reymond. 970 million druglike small molecules for virtual screening in the  
516 chemical universe database GDB-13. *J. Am. Chem. Soc.*, 131:8732, 2009.
- 517
- 518 Shuang Cao, Ning Sui, Peng Zhang, Tingting Zhou, Jinchun Tu, and Tong Zhang. TiO<sub>2</sub> nanostructures  
519 with different crystal phases for sensitive acetone gas sensors. *Journal of Colloid and Interface  
520 Science*, 607:357–366, 2022.
- 521
- 522 Giuseppe Carleo, Ignacio Cirac, Kyle Cranmer, Laurent Daudet, Maria Schuld, Naftali Tishby, Leslie  
523 Vogt-Maranto, and Lenka Zdeborová. Machine learning and the physical sciences. *Reviews of  
524 Modern Physics*, 91(4):045002, 2019.
- 525
- 526 Ivano E Castelli, David D Landis, Kristian S Thygesen, Søren Dahl, Ib Chorkendorff, Thomas F  
527 Jaramillo, and Karsten W Jacobsen. New cubic perovskites for one-and two-photon water splitting  
528 using the computational materials repository. *Energy & Environmental Science*, 5(10):9034–9043,  
529 2012a.
- 530
- 531 Ivano E Castelli, Thomas Olsen, Soumendu Datta, David D Landis, Søren Dahl, Kristian S Thygesen,  
532 and Karsten W Jacobsen. Computational screening of perovskite metal oxides for optimal solar  
533 light capture. *Energy & Environmental Science*, 5(2):5814–5819, 2012b.
- 534
- 535 Lowik Chanussot, Abhishek Das, Siddharth Goyal, Thibaut Lavril, Muhammed Shuaibi, Morgane  
536 Riviere, Kevin Tran, Javier Heras-Domingo, Caleb Ho, Weihua Hu, et al. Open catalyst 2020  
537 (oc20) dataset and community challenges. *Acs Catalysis*, 11(10):6059–6072, 2021.
- 538
- 539 Jiucheng Cheng, Chunkai Zhang, and Lifeng Dong. A geometric-information-enhanced crystal graph  
540 network for predicting properties of materials. *Communications Materials*, 2(1):92, 2021.
- 541
- 542 Stefan Chmiela, Alexandre Tkatchenko, Huziel E Saucedo, Igor Poltavsky, Kristof T Schütt, and  
543 Klaus-Robert Müller. Machine learning of accurate energy-conserving molecular force fields.  
544 *Science Advances*, 3(5):e1603015, 2017.
- 545
- 546 Stefan Chmiela, Valentin Vassilev-Galindo, Oliver T Unke, Adil Kabylda, Huziel E Saucedo, Alexan-  
547 dre Tkatchenko, and Klaus-Robert Müller. Accurate global machine learning force fields for  
548 molecules with hundreds of atoms. *Science Advances*, 9(2):eadf0873, 2023.

- 540 Aron J Cohen, Paula Mori-Sánchez, and Weitao Yang. Insights into current limitations of density  
 541 functional theory. *Science*, 321(5890):792–794, 2008.
- 542
- 543 Benjamin Coors, Alexandru Paul Condurache, and Andreas Geiger. Spherenet: Learning spherical  
 544 representations for detection and classification in omnidirectional images. In *Proceedings of the*  
 545 *European Conference on Computer Vision (ECCV)*, pp. 518–533, 2018.
- 546
- 547 Jian S Dai. Euler–rodrigues formula variations, quaternion conjugation and intrinsic connections.  
 548 *Mechanism and Machine Theory*, 92:144–152, 2015.
- 549
- 550 Kishalay Das, Pawan Goyal, Seung-Cheol Lee, Satadeep Bhattacharjee, and Niloy Ganguly. Crys-  
 551 mmnet: multimodal representation for crystal property prediction. In *Uncertainty in Artificial*  
 552 *Intelligence*, pp. 507–517. PMLR, 2023.
- 553
- 554 Murray S Daw and Michael I Baskes. Embedded-atom method: Derivation and application to  
 555 impurities, surfaces, and other defects in metals. *Physical Review B*, 29(12):6443, 1984.
- 556
- 557 Alexander Dunn, Qi Wang, Alex Ganose, Daniel Dopp, and Anubhav Jain. Benchmarking materials  
 558 property prediction methods: the matbench test set and automatminer reference algorithm. *npj*  
 559 *Computational Materials*, 6(1):138, 2020.
- 560
- 561 Alexandre Agm Duval, Victor Schmidt, Alex Hernández-García, Santiago Miret, Fragkiskos D  
 562 Malliaros, Yoshua Bengio, and David Rolnick. Faenet: Frame averaging equivariant gnn for  
 563 materials modeling. In *International Conference on Machine Learning*, pp. 9013–9033. PMLR,  
 564 2023.
- 565
- 566 Marcus Elstner and Gotthard Seifert. Density functional tight binding. *Philosophical Transactions*  
 567 *of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 372(2011):20120483,  
 568 2014.
- 569
- 570 Matthias Fey and Jan Eric Lenssen. Fast graph representation learning with pytorch geometric. *arXiv*  
 571 preprint *arXiv:1903.02428*, 2019.
- 572
- 573 L. W. Finger and R. M. Hazen. Crystal structure and isothermal compression of  $\text{fe}_2\text{o}_3$ ,  $\text{cr}_2\text{o}_3$ , and  
 574  $\text{v}_2\text{o}_3$  to 50 kbars. *Journal of Applied Physics*, 51:5362–5367, 1980. doi: 10.1063/1.327451.
- 575
- 576 Fabian Fuchs, Daniel Worrall, Volker Fischer, and Max Welling. Se (3)-transformers: 3d roto-  
 577 translation equivariant attention networks. *Advances in Neural Information Processing Systems*,  
 578 33:1970–1981, 2020.
- 579
- 580 Johannes Gasteiger, Shankari Giri, Johannes T Margraf, and Stephan Günemann. Fast and  
 581 uncertainty-aware directional message passing for non-equilibrium molecules. *arXiv preprint*  
 582 *arXiv:2011.14115*, 2020.
- 583
- 584 R. Grau-Crespo and R. Lopez-Cordero.  $\text{Mos}_2$  structural properties. *Phys. Chem. Chem. Phys.*, 4:  
 585 4078, 2002.
- 586
- 587 Gabe Guo, Tristan Luca Saidi, Maxwell W Terban, Michele Valsecchi, Simon JL Billinge, and Hod  
 588 Lipson. Ab initio structure solutions from nanocrystalline powder diffraction data via diffusion  
 589 models. *Nature Materials*, pp. 1–9, 2025.
- 590
- 591 M. Horn, C. R. Meagher, et al. Structure of anatase  $\text{tio}_2$ . *Zeitschrift für Kristallographie*, 136:273,  
 592 1972.
- 593
- 594 Rui Jiao, Wenbing Huang, Peijia Lin, Jiaqi Han, Pin Chen, Yutong Lu, and Yang Liu. Crystal  
 595 structure prediction by joint equivariant diffusion. *Advances in Neural Information Processing*  
 596 *Systems*, 36:17464–17497, 2023.
- 597
- 598 Rui Jiao, Wenbing Huang, Yu Liu, Deli Zhao, and Yang Liu. Space group constrained crystal  
 599 generation. *arXiv preprint arXiv:2402.03992*, 2024.
- 600
- 601 Sarah John. Lattice functions, wavelet aliasing, and so (3) mappings of orthonormal filters. *Journal*  
 602 *of Mathematical Physics*, 39(1):569–593, 1998.

- 594 Chaitanya K Joshi, Xiang Fu, Yi-Lun Liao, Vahe Gharakhanyan, Benjamin Kurt Miller, Anuroop  
 595 Sriram, and Zachary W Ulissi. All-atom diffusion transformers: Unified generative modelling of  
 596 molecules and materials. *arXiv preprint arXiv:2503.03965*, 2025.
- 597
- 598 George Em Karniadakis, Ioannis G Kevrekidis, Lu Lu, Paris Perdikaris, Sifan Wang, and Liu Yang.  
 599 Physics-informed machine learning. *Nature Reviews Physics*, 3(6):422–440, 2021.
- 600 Filip Ekström Kelvinius, Oskar B Andersson, Abhijith S Parackal, Dong Qian, Rickard Armiento,  
 601 and Fredrik Lindsten. Wyckoffdiff—a generative diffusion model for crystal symmetry. *arXiv  
 602 preprint arXiv:2502.06485*, 2025.
- 603
- 604 Kuzma Khrabrov, Ilya Shenbin, Alexander Ryabov, Artem Tsypin, Alexander Telepov, Anton  
 605 Alekseev, Alexander Grishin, Pavel Strashnov, Petr Zhilyaev, Sergey Nikolenko, et al. nabladft:  
 606 Large-scale conformational energy and hamiltonian prediction benchmark and dataset. *Physical  
 607 Chemistry Chemical Physics*, 24(42):25853–25863, 2022.
- 608
- 609 Byung-Hyun Kim, Jolla Kullgren, Matthew J Wolf, Kersti Hermansson, and Peter Broqvist. Mul-  
 610 tiscale modeling of agglomerated ceria nanoparticles: interface stability and oxygen vacancy  
 611 formation. *Frontiers in Chemistry*, 7:203, 2019.
- 612
- 613 Sungewan Kim, Jie Chen, Tiejun Cheng, Asta Gindulyte, Jia He, Siqian He, Qingliang Li, Benjamin A  
 614 Shoemaker, Paul A Thiessen, Bo Yu, et al. Pubchem 2025 update. *Nucleic Acids Research*, 53  
 615 (D1):D1516–D1525, 2025.
- 616
- 617 H. W. King. *CRC Handbook of Chemistry and Physics*. CRC Press, 83 edition, 2002a. Standard  
 618 phase data for gold (Au).
- 619
- 620 H. W. King. *CRC Handbook of Chemistry and Physics*. CRC Press, 83 edition, 2002b. Standard  
 621 phase data for silver (Ag).
- 622
- 623 Charles Kittel and Paul McEuen. *Introduction to solid state physics*. John Wiley & Sons, 2018.
- 624
- 625 Mustafa Kurban, Can Polat, Erchin Serpedin, and Hasan Kurban. Enhancing the electronic properties  
 626 of tio2 nanoparticles through carbon doping: An integrated dftb and computer vision approach.  
 627 *Computational Materials Science*, 244:113248, 2024.
- 628
- 629 Andrea C Levi and Miroslav Kotrla. Theory and simulation of crystal growth. *Journal of Physics:  
 630 Condensed Matter*, 9(2):299, 1997.
- 631
- 632 Daniel Levy, Siba Smarak Panigrahi, Sékou-Oumar Kaba, Qiang Zhu, Kin Long Kelvin Lee, Mikhail  
 633 Galkin, Santiago Miret, and Siamak Ravankhsh. Symmcid: Symmetry-preserving crystal  
 634 generation with diffusion models. *arXiv preprint arXiv:2502.03638*, 2025.
- 635
- 636 Rongjin Li, Xiaotao Zhang, Huanli Dong, Qikai Li, Zhigang Shuai, and Wenping Hu. Gibbs–curie–  
 637 wulff theorem in organic materials: a case study on the relationship between surface energy and  
 638 crystal growth. *Advanced Materials*, 28(8):1697–1702, 2016.
- 639
- 640 Yi-Lun Liao and Tess Smidt. Equiformer: Equivariant graph attention transformer for 3d atomistic  
 641 graphs. *arXiv preprint arXiv:2206.11990*, 2022.
- 642
- 643 Hongsheng Liu, Gotthard Seifert, and Cristiana Di Valentin. An efficient way to model complex  
 644 magnetite: Assessment of scc-dftb against dft. *The Journal of chemical physics*, 150(9), 2019.
- 645
- 646 Xiaoshan Luo, Zhenyu Wang, Qingchang Wang, Jian Lv, Lei Wang, Yanchao Wang, and Yan-  
 647 ming Ma. Crystalfow: A flow-based generative model for crystalline materials. *arXiv preprint  
 648 arXiv:2412.11693*, 2024.
- 649
- 650 Youzhi Luo, Keqiang Yan, and Shuiwang Ji. Graphdf: A discrete flow model for molecular graph  
 651 generation. In *International Conference on Machine Learning*, pp. 7192–7203. PMLR, 2021.
- 652
- 653 David W Lyons. An elementary introduction to the hopf fibration. *Mathematics Magazine*, 76(2):  
 654 87–98, 2003.

- 648 Avik Mahata, Tanmoy Mukhopadhyay, and Mohsen Asle Zaeem. Modified embedded-atom method  
 649 interatomic potentials for al-cu, al-fe and al-ni binary alloys: From room temperature to melting  
 650 point. *Computational Materials Science*, 201:110902, 2022.
- 651
- 652 Benjamin Kurt Miller, Ricky TQ Chen, Anuroop Sriram, and Brandon M Wood. Flowmm: Generating  
 653 materials with riemannian flow matching. In *Forty-first International Conference on Machine  
 654 Learning*, 2024.
- 655 R. H. Mitchell and M. A. Carpenter. Physics and chemistry of  $\text{SrTiO}_3$  perovskites. *Physics and  
 656 Chemistry of Minerals*, 27:583, 2000.
- 657
- 658 Maylis Orio, Dimitrios A Pantazis, and Frank Neese. Density functional theory. *Photosynthesis  
 659 Research*, 102:443–453, 2009.
- 660 Anyang Peng, Chun Cai, Mingyu Guo, Duo Zhang, Chengqian Zhang, Antoine Loew, Linfeng  
 661 Zhang, and Han Wang. Lambench: A benchmark for large atomic models. *arXiv preprint  
 662 arXiv:2504.19578*, 2025.
- 663
- 664 Chris J Pickard. Airss data for carbon at 10gpa and the c+ n+ h+ o system at 1gpa. *(No Title)*, 2020.
- 665 Can Polat, Mustafa Kurban, and Hasan Kurban. Multimodal neural network-based predictive  
 666 modeling of nanoparticle properties from pure compounds. *Machine Learning: Science and  
 667 Technology*, 5(4):045062, 2024.
- 668
- 669 Can Polat, Erchin Serpedin, Mustafa Kurban, and Hasan Kurban. Crysmtm: a multiphase,  
 670 temperature-resolved, multimodal dataset for crystalline materials. *Machine Learning: Science  
 671 and Technology*, 6(3):030603, 2025.
- 672 Tingting Qi, Charles W Bauschlicher Jr, John W Lawson, Tapan G Desai, and Evan J Reed. Compari-  
 673 son of reaxff, dftb, and dft for phenolic pyrolysis. 1. molecular dynamics simulations. *The Journal  
 674 of Physical Chemistry A*, 117(44):11115–11125, 2013.
- 675
- 676 Raghunathan Ramakrishnan, Pavlo O Dral, Matthias Rupp, and O Anatole Von Lilienfeld. Quantum  
 677 chemistry structures and properties of 134 kilo molecules. *Scientific Data*, 1(1):1–7, 2014.
- 678 Emilie Ringe, Richard P Van Duyne, and Laurence D Marks. Kinetic and thermodynamic modified  
 679 wulff constructions for twinned nanoparticles. *The Journal of Physical Chemistry C*, 117(31):  
 680 15859–15870, 2013.
- 681
- 682 Zachary A Rollins, Alan C Cheng, and Essam Metwally. Molprop: Molecular property prediction  
 683 with multimodal language and graph fusion. *Journal of Cheminformatics*, 16(1):56, 2024.
- 684 William J Rucklidge. Efficiently locating objects using the hausdorff distance. *International Journal  
 685 of computer vision*, 24:251–270, 1997.
- 686
- 687 Lars Ruddigkeit, Ruud Van Deursen, Lorenz C Blum, and Jean-Louis Reymond. Enumeration of 166  
 688 billion organic small molecules in the chemical universe database gdb-17. *Journal of Chemical  
 689 Information and Modeling*, 52(11):2864–2875, 2012.
- 690
- 691 M. Rupp, A. Tkatchenko, K.-R. Müller, and O. A. von Lilienfeld. Fast and accurate modeling of  
 692 molecular atomization energies with machine learning. *Physical Review Letters*, 108:058301,  
 693 2012.
- 694
- 695 Victor Garcia Satorras, Emiel Hoogeboom, and Max Welling. E (n) equivariant graph neural networks.  
 In *International Conference on Machine Learning*, pp. 9323–9332. PMLR, 2021.
- 696
- 697 Kristof Schütt, Oliver Unke, and Michael Gastegger. Equivariant message passing for the prediction  
 698 of tensorial properties and molecular spectra. In *International Conference on Machine Learning*,  
 699 pp. 9377–9388. PMLR, 2021.
- 700
- 701 Kristof T Schütt, Huziel E Sauceda, P-J Kindermans, Alexandre Tkatchenko, and K-R Müller.  
 Schnet—a deep learning architecture for molecules and materials. *The Journal of Chemical Physics*,  
 148(24), 2018.

- 702 Arkadiy Simonov and Andrew L Goodwin. Designing disorder into crystalline materials. *Nature*  
 703 *Reviews Chemistry*, 4(12):657–673, 2020.
- 704
- 705 Chris-Kriton Skylaris, Peter D Haynes, Arash A Mostofi, and Mike C Payne. Introducing onetep:  
 706 Linear-scaling density functional simulations on parallel computers. *The Journal of Chemical*  
 707 *Physics*, 122(8), 2005.
- 708 Fernand Spiegelman, Nathalie Tarrat, Jérôme Cuny, Leo Dontot, Evgeny Posenitskiy, Carles Martí,  
 709 Aude Simon, and Mathias Rapacioli. Density-functional tight-binding: basic concepts and applica-  
 710 tions to molecules and clusters. *Advances in Physics: X*, 5(1):1710252, 2020.
- 711 Anuroop Sriram, Benjamin Miller, Ricky TQ Chen, and Brandon Wood. Flowllm: Flow matching  
 712 for material generation with large language models as base distributions. *Advances in Neural*  
 713 *Information Processing Systems*, 37:46025–46046, 2024.
- 714
- 715 Richard P Stanley. The fibonacci lattice. *The Fibonacci Quarterly*, 13(3):215–232, 1975.
- 716
- 717 Thomas Surek. Crystal growth and materials research in photovoltaics: progress and challenges.  
*Journal of Crystal growth*, 275(1-2):292–304, 2005.
- 718
- 719 Philipp Thölke and Gianni De Fabritiis. Torchmd-net: equivariant transformers for neural network  
 720 based molecular potentials. *arXiv preprint arXiv:2202.02541*, 2022.
- 721
- 722 Richard Tran, Janice Lan, Muhammed Shuaibi, Brandon M Wood, Siddharth Goyal, Abhishek Das,  
 723 Javier Heras-Domingo, Adeesh Kolluru, Ammar Rizvi, Nima Shoghi, et al. The open catalyst 2022  
 724 (oc22) dataset and challenges for oxide electrocatalysts. *ACS Catalysis*, 13(5):3066–3084, 2023.
- 725
- 726 Aron Walsh, elds22, Federico Brivio, and Jarvist Moore Frost. Wmd-group/hybrid-perovskites:  
 727 Collection 1 (v1.0). <https://doi.org/10.5281/zenodo.2641358>, 2019. Hybrid per-  
 728 ovskite  $\text{CH}_3\text{NH}_3\text{PbI}_3$  structural data.
- 729
- 730 R. W. G. Wyckoff. *Crystal Structures Volume 1*. Interscience Publishers, 1963a.
- 731
- 732 R. W. G. Wyckoff. *Crystal Structures Volume 1*. Interscience Publishers, 1963b.
- 733
- 734 R. W. G. Wyckoff. *Crystal Structures Volume 1*. Interscience Publishers, 1963c.
- 735
- 736 Tian Xie, Xiang Fu, Octavian-Eugen Ganea, Regina Barzilay, and Tommi Jaakkola. Crystal diffusion  
 737 variational autoencoder for periodic material generation. *arXiv preprint arXiv:2110.06197*, 2021.
- 738
- 739 Ruo Xi Yang, Caitlin A McCandler, Oxana Andriuc, Martin Siron, Rachel Woods-Robinson,  
 740 Matthew K Horton, and Kristin A Persson. Big data in a nano world: a review on computa-  
 741 tional, data-driven design of nanomaterials structures, properties, and synthesis. *ACS Nano*, 16  
 742 (12):19873–19891, 2022.
- 743
- 744 Sherry Yang, KwangHwan Cho, Amil Merchant, Pieter Abbeel, Dale Schuurmans, Igor Mordatch, and  
 745 Ekin Dogus Cubuk. Scalable diffusion for materials generation. *arXiv preprint arXiv:2311.09235*,  
 746 2023.
- 747
- 748 Anna Yershova, Swati Jain, Steven M Lavalle, and Julie C Mitchell. Generating uniform incremental  
 749 grids on so (3) using the hopf fibration. *The International Journal of Robotics Research*, 29(7):  
 750 801–812, 2010.
- 751
- 752 Haiyang Yu, Meng Liu, Youzhi Luo, Alex Strasser, Xiaofeng Qian, Xiaoning Qian, and Shuiwang  
 753 Ji. Qh9: A quantum hamiltonian prediction benchmark for qm9 molecules. *Advances in Neural*  
 754 *Information Processing Systems*, 36, 2024.
- 755
- 756 Haoyu S Yu, Shaohong L Li, and Donald G Truhlar. Perspective: Kohn-sham density functional  
 757 theory descending a staircase. *The Journal of Chemical Physics*, 145(13), 2016.
- 758
- 759 Claudio Zeni, Robert Pinsler, Daniel Züigner, Andrew Fowler, Matthew Horton, Xiang Fu, Sasha  
 760 Shysheya, Jonathan Crabbé, Lixin Sun, Jake Smith, et al. Mattergen: a generative model for  
 761 inorganic materials design. *arXiv preprint arXiv:2312.03687*, 2023.
- 762
- 763 Guishan Zheng, Stephan Irle, and Keiji Morokuma. Performance of the dftb method in comparison to  
 764 dft and semiempirical methods for geometries and energies of c20–c86 fullerene isomers. *Chemical*  
 765 *Physics Letters*, 412(1-3):210–216, 2005.