Periodic Materials Generation using Text-Guided Joint Diffusion Model

Kishalay Das[1], Subhojyoti Khastagir[1], Pawan Goyal[1], Seung-Cheol Lee[2], Satadeep Bhattacharjee[2], Niloy Ganguly[1]

[1]Indian Institute of Technology Kharagpur, India, [2] Indo Korea Science and Technology Center, Bangalore, India

Periodic Crystal Materials

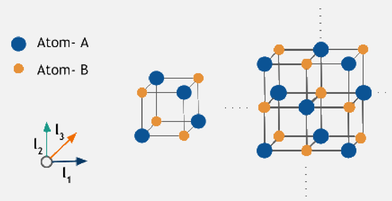


Figure: Structure of Crystalline Materials

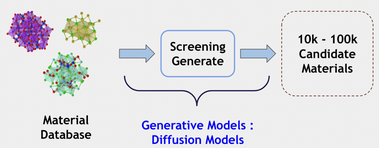
- **Crystal Material M = (A, X, L)**
- **Feature Matrix (A)**: Atomic Feature Set of the Material.
- **Coordinate Matrix (X)** : Atomic Fractional Coordinate Positions.
- **Lattice Matrix (L)** : How a unit cell repeats itself in the 3D space.

- we can represent its infinite periodic structure as

$$\hat{C} = \{\hat{c}_i | \hat{c}_i = c_i + \sum_{j=1}^{3} k_j l_j\}; \ \hat{X} = \{\hat{x}_i | \hat{x}_i = x_i\} \tag{1}$$

where $k_1, k_2, k_3, i \in Z, 1 \leq i \leq n$.

Problem Formulation



New Material Generation:
Given a set of stable and valid 3D crystal structures, generate new crystal materials which are valid and stable in nature and pose desired chemical properties.

Challenges

- None of these existing SOTA models learns the **joint distribution** of atom coordinates, types, and lattice structure of the material.
- SE(3)-equivariant GNNs as backbone denoising network → largely relies on messages passing around the local neighborhood of the atoms → fail to incorporate global structural knowledge into the diffusion process.
- SOTA models are unconditional by design → limited utility in real-world scenarios.
- In a realistic setup, users would input key material details, such as chemical formula, space group, symmetry, bond lengths, and properties, ensuring the generated structure aligns with these specifications.

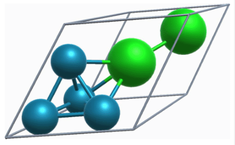
Text Guided Material Generation

BaPd2 is Cubic structured and crystallizes in the cubic Fd-3m space group. Ba(1) is bonded in a 12-coordinate geometry to twelve equivalent Pd(1) atoms. All Ba(1)-Pd(1) bond lengths are 3.37 Å. Pd(1) is bonded to six equivalent Ba(1) and six equivalent Pd(1) atoms to form a mixture of face, edge, and corner-sharing cuboctahedra. All Pd(1)-Pd(1) bond lengths are 2.88 Å. Its formation energy per atom is -0.578, band gap is 0.0, energy above hull is 0.0. Generate the material

Detailed Description by Robocrystallographer

Below is a description of a bulk material. The chemical formula is BaPd2. The elements are Ba, Pd. The formation energy per atom is negative. The band gap is zero. The energy above the convex hull is zero. The space group number is 227. The crystal system is cubic.Generate the material

Short Prompt by Users



Unit Cell Structure

Figure: Detailed textual description generated by Robocrystallographer, less-detailed prompts by domain experts, and crystal unit cell structure of **$BaPd_2$**.

Text-Guided Diffusion Model for Material Generation

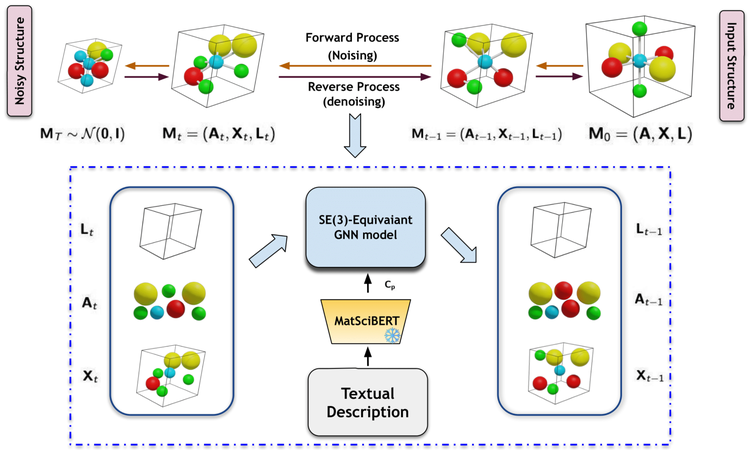


Figure: Model Architecture of our proposed text guided diffusion model TGDMat.

Diffusion on *L* Leverage Denoising Diffusion Probabilistic Model (DDPM)

$$q(L_t|L_0) = \mathcal{N}(L_t|\sqrt{\bar{\alpha}_t}L_0,\ (1-\bar{\alpha}_t)\mathbf{I}) \tag{2}$$

$$p(L_{t-1}|M_t, C_p) = \mathcal{N}\{L_{t-1}|\mu^L(M_t, C_p), \beta_t \frac{(1-\bar{\alpha}_{t-1})}{(1-\bar{\alpha}_t)}\mathbf{I}\}$$

$$\mu^L(M_t, C_p) = \frac{1}{\sqrt{\alpha_t}}\left(L_t - \frac{1-\alpha_t}{\sqrt{1-\bar{\alpha}_t}}\hat{\epsilon}^L(M_t, C_p, t)\right) \tag{3}$$

$$\mathcal{L}_{lattice} = \mathbb{E}_{\epsilon^L, t\sim\mathcal{U}(1,T)}\|\epsilon^L - \hat{\epsilon}^L\|_2^2 \tag{4}$$

Diffusion on *A* Leverage Discrete Denoising Diffusion Probabilistic Model (D3PM)

$$q(a_t|a_{t-1}) = Cat(a_t; p = a_{t-1}Q_t) \tag{5}$$

$$[Q_t]_{i,j} := q(a_t = i \mid a_{t-1} = j) := \begin{cases} 1, & \text{if } i = j = [MASK]. \\ 1-\beta_t, & \text{if } i = j \neq [MASK] \\ \beta_t, & \text{if } i \neq j = [MASK]. \end{cases} \tag{6}$$

$$\mathcal{L}_{type} = \mathcal{L}_{VB} + \lambda\mathcal{L}_{CE} \tag{7}$$

Diffusion on *X* Atomic fractional coordinates in crystal material lives in quotient space $\mathbb{R}^{N\times3}/\mathbb{Z}^{N\times3}$ induced by the crystal periodicity → noise sample from Wrapped Normal (WN) distribution to ***X*** and during backward diffusion leverage Score Matching Networks.

$$q(X_t \mid X_0) = \mathcal{N}_W(X_t \mid X_0, \sigma_t^2\mathbf{I});\ X_t = f_w(X_0 + \sigma_t\epsilon^X) \tag{8}$$

$$\mathcal{L}_{coord} = \mathbb{E}_{\substack{X_t\sim q(X_t|X_0) \\ t\sim\mathcal{U}(1,T)}}\|\nabla_{X_t}\log q(X_t|X_0) - \hat{\epsilon}^X(M_t, C_{p,t})\|_2^2 \tag{9}$$

Text Guided Denoising Network We extract the [CLS] token embedding from the material's textual description and project it into a contextual embedding C_p, which is then used to guide the denoising process in the equivariant GNN model. We fuse textual representation C_p into input atom feature h_i^0 as

$$h_i^0 = \rho\ \{\ f_{atom}(a_i)\ ||\ f_{pos}(t)\ ||\ C_p \tag{10}$$

Crystal Structure Prediction (CSP)

Method	# Samples	Perov-5		Carbon-24		MP-20	
		Match Rate ↑	RMSE ↓	Match Rate ↑	RMSE ↓	Match Rate ↑	RMSE ↓
CDVAE	1	45.31	0.1138	17.09	0.2969	33.90	0.1045
	20	88.51	0.0464	88.37	0.2286	66.95	0.1026
DiffCSP	1	52.02	0.0760	17.54	0.2759	51.49	0.0631
	20	**98.60**	0.0128	88.47	0.2192	77.93	0.0492
TGDMat (Short)	1	56.54	0.0583	24.13	0.2424	52.22	0.0597
	20	98.25	0.0137	88.28	0.2252	80.97	0.0443
TGDMat (Long)	1	**90.46**	**0.0203**	**44.63**	**0.2266**	**55.15**	**0.0572**
	20	98.59	**0.0072**	**95.27**	**0.1534**	**82.02**	**0.0483**

Table: Summary of results on *CSP* task.

Random Material Generation (Gen)

Dataset	Method	Validity ↑		Coverage ↑		Property Statistics (EMD) ↓		
		Compositional	Structural	COV-R	COV-P	# Element	ρ	$\mathscr{E}$
Perov-5	CDVAE	98.59	**100**	99.45	98.46	0.0628	0.1258	0.0264
	CDVAE+	98.45	99.8	99.53	99.09	0.0609	0.1276	0.0223
	DiffCSP	**98.85**	**100**	99.74	98.27	0.0128	0.1110	0.0263
	DiffCSP+	98.44	**100**	99.85	98.53	0.0119	0.1070	0.0241
	TGDMat(Short)	98.28	**100**	99.7	99.24	0.0108	0.0947	0.0257
	TGDMat(Long)	98.63	**100**	**99.83**	**99.52**	**0.0090**	**0.0497**	**0.0187**
Carbon-24	CDVAE	-	**100**	99.8	83.08	-	0.1407	0.285
	CDVAE+	-	**100**	99.8	84.76	-	0.1377	0.266
	DiffCSP	-	**100**	**99.9**	97.27	-	0.0805	0.082
	DiffCSP+	-	**100**	**99.9**	97.33	-	0.0763	0.085
	TGDMat(Short)	-	**100**	99.8	91.77	-	0.0681	0.087
	TGDMat(Long)	-	**100**	**99.9**	92.43	-	**0.043**	**0.063**
MP-20	CDVAE	86.70	**100**	99.15	99.49	1.432	0.6875	0.2778
	CDVAE+	87.42	**100**	99.57	99.81	0.972	0.6388	0.2977
	DiffCSP	83.25	**100**	99.71	99.76	0.3398	0.3502	0.1247
	DiffCSP+	85.07	**100**	99.8	99.89	0.3122	0.3799	0.1355
	TGDMat(Short)	86.60	**100**	99.79	99.88	0.3337	**0.3296**	**0.1154**
	TGDMat(Long)	**92.97**	**100**	**99.89**	99.95	**0.2890**	0.3382	0.1189

Table: Summary of results on *Gen* task.

Takeaways

- First to explore text-guided diffusion for material generation, bridging language understanding and material structure generation.
- Unlike prior models, TGDMat conducts joint diffusion on lattices, atom types, and coordinates, enhancing its ability to accurately capture the crystal geometry.
- Additionally, incorporating global structural knowledge through textual descriptions at each denoising step improves TGDMat's ability to generate plausible materials with valid and stable structures.

Contact and Paper Information

Paper

Github