Constrained Diffusion for Accelerated Structure Relaxation of Inorganic Solids with Point Defects

Anonymous Author(s)
Affiliation
Address
`email`

Abstract

Point defects affect material properties by altering electronic states and modifying local bonding environments. However, high-throughput first-principles simulations of point defects are costly due to large simulation cells and complex energy landscapes. To this end, we propose a generative framework for simulating point defects, overcoming the limits of costly first-principles simulators. By leveraging a primal-dual algorithm, we introduce a constraint-aware diffusion model which outperforms existing constrained diffusion approaches in this domain. Across six defect configuration settings for Bi_2Te_3, the proposed approach provides state-of-the-art performance generating physically grounded structures.

1 Introduction

Point defects play an important role in determining the properties of crystalline materials [1–4]. Recent work on additive manufacturing (AM) of thermoelectric materials reveals a more complex relationship between processing conditions and point defects than in traditional methods [5–7]. They are exploited to tailor the structural, electronic, and thermal properties of advanced materials for technological applications.

Layered chalcogenides, a promising material class, are highly susceptible to point defects due to their mixed bonding with weak interlayer van der Waals interactions and strong intralayer covalent bonds, which critically influence electronic and phononic transport properties [8, 9]. A well-known example is bismuth telluride (Bi_2Te_3), which has been widely studied for its high thermoelectric efficiency in the low-temperature regime [10–13]. The performance of Bi_2Te_3 is strongly governed by defect concentration and configuration, and preliminary AM work has shown that laser processing parameters can change thermoelectric properties to optimize efficiency and carrier transport. [6, 7].

Point defects are hard to probe directly and are typically identified indirectly via combined techniques. Traditionally, density functional theory (DFT) have been used to complement experimental methods for defect identification [14] which is computationally expensive, making high-throughput studies computationally prohibitive. This situation highlights the need for alternative approaches, such as surrogate models, to supplement the DFT calculations. In related domains, deep learning methods have successfully overcome these barriers [15]. Generative models enable digital twins for defect prediction, with diffusion models achieving state-of-the-art material generation and inverse design [16–18]. Yet, these models excel at generating realistic data, struggle with strict physical constraints. Without reliability, designs risk impracticality and may hinder transition to production.

For other scientific applications, these challenges have been addressed through physics-aware generative processes. Simple constraints can be injected into sampling [19–21], but complex ones often require integrating costly simulators with high overhead [22, 23]. In this case, runtime demands may limit utility, making rejection sampling more efficient. For physics-informed generation of

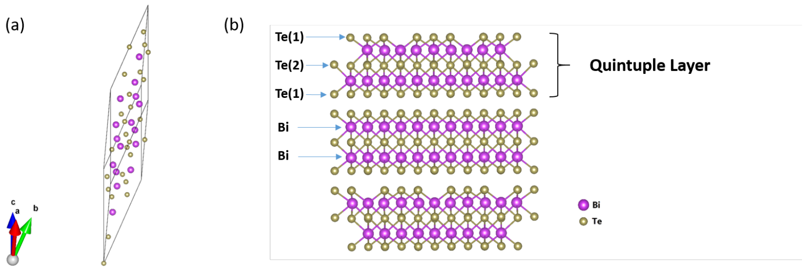

Figure 1: **(a)** A 2×2×2 supercell of Bi_2Te_3 for DFT simulations. **(b)** Another supercell representation, highlighting quintuple layer, structure Te(1)–Bi–Te(2)–Bi–Te(1) with vacancy defect.

materials like Bi_2Te_3, first-principle simulators [24] are challenging, as constraints lack closed forms and evaluations scale exponentially with system size.

Contributions. To address these existing challenges, this paper makes the following contributions: **(1)** It extends previous constrained generation approaches to handle complex constraints modeled by tractable functions and neural surrogates, leveraging a primal-dual algorithm. **(2)** It provides the first study which explores constrained diffusion to supplement the DFT-based point defect simulation of a thermoelectric material (Bi_2Te_3). **(3)** Through rigorous evaluation across six defect configurations of interest, we provide state-of-the-art performance for generative modeling of physically realistic predictions of layered chalcogenides.

2 Preliminaries

Score-based Diffusion Models [25, 26], and equivalently denoising diffusion probabilistic models [27, 28], model data distributions by introducing a forward noising process and training a neural network to approximate its reverse. In the *forward process*, a clean sample $\boldsymbol{x}_0 \sim p_{\text{data}}(\boldsymbol{x}_0)$ is gradually perturbed through a sequence of intermediate distributions $\{p_t(\boldsymbol{x}_t)\}_{t=0}^{T}$. This is achieved using a Gaussian noise kernel with a predefined variance schedule $\bar{\alpha}_t$, which increases with t. As $t \to T$, the distribution of $\boldsymbol{x}_t$ converges to a standard Gaussian, such that $p_T(\boldsymbol{x}_T) \approx \mathcal{N}(0, I)$. This process facilitates the training of a score network $\boldsymbol{s}_\theta(\boldsymbol{x}_t, t)$, which learns the score function $\boldsymbol{s}_\theta(\boldsymbol{x}_t, t) \approx \nabla_{\boldsymbol{x}_t} \log p_t(\boldsymbol{x}_t|\boldsymbol{x}_0)$ by error in the predicted score estimate:

$$\min_{\theta} \mathop{\mathbb{E}}_{t\sim[T,1],\ \boldsymbol{x}_0\sim p_{\text{data}}} \left[\|\boldsymbol{s}_\theta(\boldsymbol{x}_t, t) - \nabla_{\boldsymbol{x}_t} \log p_t(\boldsymbol{x}_t|\boldsymbol{x}_0)\|_2^2\right] \tag{1}$$

The trained score network $\boldsymbol{s}_\theta(\boldsymbol{x}_t, t)$ is then used in the *reverse process* to iteratively reconstruct data samples from the training distribution p_{data}. At each timestep t, the score function is applied to the reverse process, transitioning Gaussian noise $\boldsymbol{x}_T \sim \mathcal{N}(0, I)$ to high fidelity samples.

Projected Diffusion Models [19] build from Langevin Dynamics employed by score-based diffusion models, interpreting the sampling process as an optimization with respect to a series of intermediate probability density functions (Equation (2a)). In doing so, the formulation enabled a natural extension to constrained sampling problems by enforcing constraint adherence on the sample:

$$\underset{\boldsymbol{x}_T,\ldots,\boldsymbol{x}_1}{\text{minimize}} \sum_{t=T,\ldots,0} -\log p_t(\boldsymbol{x}_t|\boldsymbol{x}_0) \tag{2a}$$

$$\text{s.t.:} \quad \boldsymbol{x}_T, \ldots, \boldsymbol{x}_0 \in \mathbf{C}. \tag{2b}$$

Constraints can then be enforced by extending the reverse process with Projected Langevin Dynamics, such each update step is projected to the nearest feasible point by $\mathcal{P}_{\mathbf{C}}(\boldsymbol{x}) = \arg\min_{\boldsymbol{y}\in\mathbf{C}} \|\boldsymbol{y} - \boldsymbol{x}\|_2^2$:

$$\boldsymbol{x}_t^{(i+1)} = \mathcal{P}_{\mathbf{C}}\big(\boldsymbol{x}_t^{(i)} + \gamma_t \boldsymbol{s}_\theta(\boldsymbol{x}_t^{(i)}, t) + \sqrt{2\gamma_t}\epsilon\big) \tag{3}$$

where γ_t is the step size decreasing with t, and ϵ is Gaussian noise.

3 Constrained Diffusion for Point Defect Generation

Generating relaxed atomic structures with targeted defects in Bi_2Te_3 is a challenging problem, as the combinatorial diversity of defect placements creates a vast configuration space where DFT becomes computationally prohibitive. To address this, we propose a *generative diffusion framework* that samples Bi_2Te_3 structures under prescribed stoichiometries. Given target counts of Bi and Te atoms in a supercell, as shown in Figure 1, the model generates diverse arrangements that explore physically meaningful defects while retaining diffusion sampling's inherent randomness. However, diffusion alone does not ensure physical plausibility, motivating the adoption of three constraint types.

1. *Geometric Constraints.* Ensure physically-meaningful interatomic distances and periodic boundary conditions. Physical validity requires a lower bound on interatomic distances to prevent atomic collapse. For consistency with DFT simulations, generated structures must also satisfy periodic boundary constraints to align with reference distributions. Thus,
$$\|r_i - r_j\| \;\geq\; d_{\min}, \;\; \forall i \neq j; \qquad f_i \in [0,1]^3, \tag{4}$$
where $d_{\min}$ is the minimum allowable interatomic distance between arbitrary atoms r_i and r_j, and f_i is the fractional coordinate constrained to $[0,1]^3$ to enforce periodic boundary conditions.
2. *Distributional Constraints.* Align radial distributions (within 5Å) with reference data; This reflects structural patterns, such as relative intensities of coordination shells. This yeilds,
$$D_{[0,R_c]}(p_{\text{gen}},\, p_{\text{ref}}) \leq \varepsilon, \tag{5}$$
where $D_{[0,R_c]}$ measures discrepancy between distributions over distances in $[0, R_c]$, p denotes the empirical distribution of pairwise distances in generated and reference structures.
3. *Force Minimization.* Stable structures are defined by low total forces, rendering force minimization critical for physical plausibility. As shown in Section 4, stochastic generations often yield high forces. We therefore introduce a *soft* force-minimization constraint, adding a $\|\text{force}(r)\|^2$ prediction term from a surrogate [29] to the objective (2a), optimized at sampling.

Terminal Set Constraints. Projection-based constraint enforcement can be very effective in many applications. Unlike post-processing, these techniques yield higher-fidelity outputs to the learned distribution, as the score function refines samples displaced from high-density regions. Yet, projected diffusion models rely on a stringent assumption: *Constraints can be meaningfully imposed on any* $\boldsymbol{x}_t \sim p_t(\boldsymbol{x}_t)$ *for an arbitrary* t [19].

Often, it may be reasonable to assume that constraints can be modeled at intermediate states (e.g., path planning, where even noisy samples form valid trajectories) [30]. However, this assumption does not hold in many scientific applications such as ours, where complex dynamics require surrogates that are unreliable on noisy samples. This challenge, studied in training-free guidance [31–33], can lead to "misaligned gradients," only matching the true value as $t \to 0$, leading to compounding errors [34, 35]. Alternatively, finetuning on noisy data is often ineffective: **(1)** Labels are not available for our constraints due to simulators requirements [24], which are not only prohibitive due to runtime but also as the DFT process fails on increasingly noisy samples. **(2)** Even when labels are available, gradient inaccuracies remain a concern [36].

Hence, designing a constraint evaluation surrogate robust to high noise levels is prohibitive for complicated constraint sets. A more promising strategy is to relax the intermediate feasibility assumption. To facilitate this, our framework treats the objective in Equation (2) as a *multi-stage optimization problem.* While Equation (2b) enforces $\boldsymbol{x}_t \in \mathbf{C}$, from a multi-stage perspective only the terminal state $\boldsymbol{x}_0$ must be feasible. Thus, intermediate constraints serve as a conservative proxy for final feasibility, while earlier states may remain unconstrained if the process yields $\boldsymbol{x}_0 \in \mathbf{C}$.

Hence, we impose the constraints only on the final minimization of $-\log p_{\text{data}}(\boldsymbol{x}_0)$. This can then be interpreted as terminal set constraints drawn from Model Predictive Control theory [37], where feasibility is often only enforced on final decision variables. This removes dependence on intermediate feasibility assumptions as constraints are only imposed on clean samples within p_{data}.

Primal-Dual Projection Algorithm. As neural surrogates provide differentiable loss functions, it is most reasonable to solve the projection mapping using a gradient-based approach. To facilitate this, we adopt the Augmented Lagrangian representation of the constraint problem following Liang et al., incorporating constraints into the objective through Lagrange multipliers [38], and yielding:
$$\mathcal{L}(\boldsymbol{y}, \boldsymbol{x}_0^{(i)}; \lambda, \mu) := \|\boldsymbol{y} - \boldsymbol{x}_0^{(i)}\|_2^2 + \sum_j^N \lambda_j c_j(\boldsymbol{y}) + \sum_k^N \frac{\mu_k}{2} c_k(\boldsymbol{y})^2 \tag{6}$$
where $\lambda = \{\lambda_1, \ldots, \lambda_N\}$ and $\mu = \{\mu_1, \ldots, \mu_N\}$ are dual variables associated with N constraint violation functions $c = \{c_1, \ldots, c_N\}$ and $\boldsymbol{y}$ is the solution optimized. Ensuring the strongest

Method	Config	RMSD ↓	RDF ↓	Force (eV/Å) ↓
Conditional DM	16Bi + 21Te	1.86 ± 0.33	65.85 ± 19.75	1.85×10^9
	16Bi + 22Te	1.93 ± 0.36	82.11 ± 29.54	3.49×10^9
Projected DM	16Bi + 21Te	2.54 ± 0.55	1.11 ± 0.77	8.35×10^2
	16Bi + 22Te	1.41 ± 0.83	0.78 ± 0.55	7.20×10^1
Post-proc DM	16Bi + 21Te	2.30 ± 0.40	27.25 ± 28.30	8.45×10^6
	16Bi + 22Te	1.14 ± 0.65	21.82 ± 28.09	9.29×10^6
Ours	16Bi + 21Te	$\mathbf{0.90} \pm 0.38$	$\mathbf{0.30} \pm 0.11$	$\mathbf{7.88 \times 10^{-2}}$
	16Bi + 22Te	$\mathbf{1.02} \pm 0.70$	$\mathbf{0.35} \pm 0.18$	$\mathbf{1.44 \times 10^{-1}}$

Table 1: Empirical comparison across stoichiometric configurations. Extended table in Appendix C.

relaxation of this projection, the Lagrangian dual is used to optimize the dual variables [39]:

$$\arg\max_{\lambda,\,\mu} \left(\arg\min_{\boldsymbol{y}} \big(\mathcal{L}(\boldsymbol{y}, \boldsymbol{x}_0^{(i)}; \lambda, \mu) \big) \right) \tag{7}$$

This converges to satisfy the distance and distributional constraints while simultaneously minimizing the force and remaining close to the original input. Following Equation (3), we begin applying this as our projection operator at $t = 0$. More details are provided in Appendix B.

4 Experiments

Baselines. We compare our approach against three representative baselines: **(1)** conditional diffusion models [40], where constraints are incorporated during training and sampling to bias the denoising process toward feasible regions. **(2)** post-processing optimization [41], where the final output of an unconstrained diffusion model is projected onto the constraint set in a single step. **(3)** projected diffusion models [19], which enforce constraints at every denoising step. This represents the current state-of-the-art but relies on meaningful constraint evaluation under noisy configurations.

Evaluation. We target four categories of defects that frequently occur in Bi_2Te_3 structure systems: Te vacancies, Bi vacancies, Bi on Te anti-site, and Te on Bi anti-site. Additional details on the problem setup and results are provided in Appendix C. Our evaluation consists of the following metrics:

- **RMSD (Root Mean Square Deviation)**: Computes positional deviation between generations and closest reference structures by matching atoms via a Hungarian matching algorithm.
- **RDF (Radial Distribution Function)**: Compares radial distributions for all generations within 5Å to assess global geometric consistency. This indicates the accuracy of positions and heights of nearest neighbor shell peaks under a common defect.
- **Force (eV/Å)**: Reports the total force as computed by a neural surrogate [29]. Lower force indicates generations closer to stable states and more suitable for downstream simulations.

Interpretation. As shown in Table 1, our method achieves substantially lower deviation from reference structures and reduced total forces. Existing approaches observe key limitations: **(1)** conditional diffusion preserves structural characteristics but allows overlapping atoms, undermining physical realism and leading to extremely high total force; **(2)** post-processing, while improving over the conditional model, struggles to impose the highly nonconvex constraints in a single step, as the final output often deviates from the constraint set significantly; **(3)** projected diffusion provides the strongest baseline, but misaligned gradients at higher noise results in higher RDF deviation. In contrast, our method preserves diffusion dynamics while ensuring feasibility. By applying constraints on the final Langevin optimization, gradient inaccuracies are avoided (unlike projected diffusion models), while providing a sufficient number of constrained diffusion steps to ensure feasibility (unlike post-processing schemes). We achieve the lowest RMSD in nearly all settings, the highest RDF similarity (by more than 2x), and lower total forces by several orders of magnitude.

5 Conclusion

Motivated by the real-world importance of thermoelectric defect structure modeling, this paper addresses existing barriers by integrating constraints into the sampling process through tractable functions and neural surrogate models. Building upon a primal–dual projection algorithm, this work provides effectively enforces feasibility while addressing challenges with misaligned gradients. Across six representative defect configurations in Bi_2Te_3, our approach provides state-of-the-art performance in producing physically realistic structures, highlighting the significance of our framework for complex material systems.

References

[1] Jing-Feng Li, Wei-Shu Liu, Li-Dong Zhao, and Min Zhou. High-performance nanostructured thermoelectric materials. *NPG Asia Materials*, 2(4):152–158, 2010.

[2] Xianli Su, Ping Wei, Han Li, Wei Liu, Yonggao Yan, Peng Li, Chuqi Su, Changjun Xie, Wenyu Zhao, Pengcheng Zhai, et al. Multi-scale microstructural thermoelectric materials: transport behavior, non-equilibrium preparation, and applications. *Advanced Materials*, 29(20):1602013, 2017.

[3] Anuj Goyal, Prashun Gorai, Haowei Peng, Stephan Lany, and Vladan Stevanović. A computational framework for automation of point defect calculations. *Computational Materials Science*, 130:1–9, 2017.

[4] Irea Mosquera-Lois, Seán R Kavanagh, Aron Walsh, and David O Scanlon. Identifying the ground state structures of point defects in solids. *npj Computational Materials*, 9(1):25, 2023.

[5] Jiang Bi, Zeqi Liu, Bo Li, Shide Li, Zhuoyun Yang, Mikhail Dmitrievich Starostenkov, and Guojiang Dong. Additive manufacturing of thermoelectric materials: materials, synthesis and manufacturing: a review. *Journal of Materials Science*, 59(2):359–381, 2024.

[6] Ryan Welch, Dean Hobbis, Andrew J Birnbaum, George Nolas, and Saniya LeBlanc. Nano-and micro-structures formed during laser processing of selenium doped bismuth telluride. *Advanced Materials Interfaces*, 8(15):2100185, 2021.

[7] Cagri Oztan, Ryan Welch, and Saniya LeBlanc. Additive manufacturing of bulk thermoelectric architectures: a review. *Energies*, 15(9):3121, 2022.

[8] Jeffrey D Cain, Eve D Hanson, Fengyuan Shi, and Vinayak P Dravid. Emerging opportunities in the two-dimensional chalcogenide systems and architecture. *Current Opinion in Solid State and Materials Science*, 20(6):374–387, 2016.

[9] AI Romanenko, GE Chebanova, Tingting Chen, Wenbin Su, and Hongchao Wang. Review of the thermoelectric properties of layered oxides and chalcogenides. *Journal of Physics D: Applied Physics*, 55(14):143001, 2021.

[10] Yu Pan, Tian-Ran Wei, Chao-Feng Wu, and Jing-Feng Li. Electrical and thermal transport properties of spark plasma sintered n-type bi 2 te 3- x se x alloys: the combined effect of point defect and se content. *Journal of Materials Chemistry C*, 3(40):10583–10589, 2015.

[11] JJ Shen, LP Hu, TJ Zhu, and XB Zhao. The texture related anisotropy of thermoelectric properties in bismuth telluride based polycrystalline alloys. *Applied Physics Letters*, 99(12), 2011.

[12] Yu Pan and Jing-Feng Li. Thermoelectric performance enhancement in n-type bi2 (tese) 3 alloys owing to nanoscale inhomogeneity combined with a spark plasma-textured microstructure. *NPG Asia Materials*, 8(6):e275–e275, 2016.

[13] Chunmei Tang, Zhicheng Huang, Jun Pei, Bo-Ping Zhang, Peng-Peng Shang, Zhihang Shan, Zhiyue Zhang, Haiyun Gu, and Kaibin Wen. Bi 2 te 3 single crystals with high room-temperature thermoelectric performance enhanced by manipulating point defects based on first-principles calculation. *RSC advances*, 9(25):14422–14431, 2019.

[14] Christoph Freysoldt, Blazej Grabowski, Tilmann Hickel, Jörg Neugebauer, et al. First-principles calculations for point defects in solids. *Rev. Mod. Phys.*, 86:253–305, Mar 2014.

[15] Sehyun Chun, Sidhartha Roy, Yen Thi Nguyen, Joseph B Choi, Holavanahalli S Udaykumar, and Stephen S Baek. Deep learning for synthetic microstructure generation in a materials-by-design framework for heterogeneous energetic materials. *Scientific reports*, 10(1):13307, 2020.

[16] Rongzhi Dong, Nihang Fu, Edirisuriya MD Siriwardane, and Jianjun Hu. Generative design of inorganic compounds using deep diffusion language models. *The Journal of Physical Chemistry A*, 128(29):5980–5989, 2024.

[17] Junkil Park, Aseem Partap Singh Gill, Seyed Mohamad Moosavi, and Jihan Kim. Inverse design of porous materials: a diffusion model approach. *Journal of Materials Chemistry A*, 12(11): 6507–6514, 2024.

[18] Izumi Takahara, Kiyou Shibata, and Teruyasu Mizoguchi. Generative inverse design of crystal structures via diffusion models with transformers. *arXiv preprint arXiv:2406.09263*, 2024.

[19] Jacob K Christopher, Stephen Baek, and Nando Fioretto. Constrained synthesis with projected diffusion models. *Advances in Neural Information Processing Systems*, 37:89307–89333, 2024.

[20] Chaoran Cheng, Boran Han, Danielle C Maddix, Abdul Fatir Ansari, Andrew Stuart, Michael W Mahoney, and Yuyang Wang. Gradient-free generation for hard-constrained systems. *arXiv preprint arXiv:2412.01786*, 2024.

[21] Utkarsh Utkarsh, Pengfei Cai, Alan Edelman, Rafael Gomez-Bombarelli, and Christopher Vincent Rackauckas. Physics-constrained flow matching: Sampling generative models with hard constraints. *arXiv preprint arXiv:2506.04171*, 2025.

[22] Ye Yuan, Jiaming Song, Umar Iqbal, Arash Vahdat, and Jan Kautz. Physdiff: Physics-guided human motion diffusion model. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 16010–16021, 2023.

[23] Stefano Zampini, Jacob K Christopher, Luca Oneto, Davide Anguita, and Ferdinando Fioretto. Training-free constrained generation with stable diffusion models. *arXiv preprint arXiv:2502.05625*, 2025.

[24] Paolo Giannozzi, Stefano Baroni, Nicola Bonini, Matteo Calandra, Roberto Car, Carlo Cavazzoni, Davide Ceresoli, Guido L Chiarotti, Matteo Cococcioni, Ismaila Dabo, et al. Quantum espresso: a modular and open-source software project for quantumsimulations of materials. *Journal of physics: Condensed matter*, 21(39):395502, 2009.

[25] Yang Song and Stefano Ermon. Generative modeling by estimating gradients of the data distribution. *Advances in neural information processing systems*, 32, 2019.

[26] Yang Song, Jascha Sohl-Dickstein, Diederik P Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. *arXiv preprint arXiv:2011.13456*, 2020.

[27] Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Advances in neural information processing systems*, 33:6840–6851, 2020.

[28] Chicago Y Park, Michael T McCann, Cristina Garcia-Cardona, Brendt Wohlberg, and Ulugbek S Kamilov. Random walks with tweedie: A unified framework for diffusion models. *arXiv preprint arXiv:2411.18702*, 2024.

[29] Ilyes Batatia, David P Kovacs, Gregor Simm, Christoph Ortner, and Gábor Csányi. Mace: Higher order equivariant message passing neural networks for fast and accurate force fields. *Advances in neural information processing systems*, 35:11423–11436, 2022.

[30] Jinhao Liang, Jacob K Christopher, Sven Koenig, and Ferdinando Fioretto. Simultaneous multi-robot motion planning with projected diffusion models. *arXiv preprint arXiv:2502.03607*, 2025.

[31] Haotian Ye, Haowei Lin, Jiaqi Han, Minkai Xu, Sheng Liu, Yitao Liang, Jianzhu Ma, James Y Zou, and Stefano Ermon. Tfg: Unified training-free guidance for diffusion models. *Advances in Neural Information Processing Systems*, 37:22370–22417, 2024.

[32] Jiwen Yu, Yinhuai Wang, Chen Zhao, Bernard Ghanem, and Jian Zhang. Freedom: Training-free energy-guided conditional diffusion model. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 23174–23184, 2023.

[33] Lewis Cornwall, Joshua Meyers, James Day, Lilly S Wollman, Neil Dalchau, and Aaron Sim. Training-free guidance of diffusion models for generalised inpainting, 2024.

[34] Yifei Shen, Xinyang Jiang, Yifan Yang, Yezhen Wang, Dongqi Han, and Dongsheng Li. Understanding and improving training-free loss-based diffusion guidance. *Advances in Neural Information Processing Systems*, 37:108974–109002, 2024.

[35] Philipp Vaeth, Alexander M Fruehwald, Benjamin Paassen, and Magda Gregorova. Gradcheck: Analyzing classifier guidance gradients for conditional diffusion sampling. *arXiv preprint arXiv:2406.17399*, 2024.

[36] Bahjat Kawar, Roy Ganz, and Michael Elad. Enhancing diffusion-based image synthesis with robust classifier guidance. *arXiv preprint arXiv:2208.08664*, 2022.

[37] James Blake Rawlings, David Q Mayne, Moritz Diehl, et al. *Model predictive control: theory, computation, and design*, volume 2. Nob Hill Publishing Madison, WI, 2020.

[38] Stephen P Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

[39] Ferdinando Fioretto, Pascal Van Hentenryck, Terrence WK Mak, Cuong Tran, Federico Baldo, and Michele Lombardi. Lagrangian duality for constrained deep learning. In *Joint European conference on machine learning and knowledge discovery in databases*, pages 118–135. Springer, 2020.

[40] Jonathan Ho and Tim Salimans. Classifier-free diffusion guidance. *arXiv preprint arXiv:2207.12598*, 2022.

[41] Giorgio Giannone, Akash Srivastava, Ole Winther, and Faez Ahmed. Aligning optimization trajectories with diffusion models for constrained design generation. *Advances in neural information processing systems*, 36:51830–51861, 2023.

[42] Prafulla Dhariwal and Alexander Nichol. Diffusion models beat gans on image synthesis. *Advances in neural information processing systems*, 34:8780–8794, 2021.

[43] Adham Hashibon and Christian Elsässer. First-principles density functional theory study of native point defects in bi$_2$te$_3$. *Phys. Rev. B*, 84:144117, Oct 2011. doi: 10.1103/PhysRevB.84.144117.

[44] Pauli Virtanen and others. SciPy 1.0: Fundamental algorithms for scientific computing in python. *Nature Methods*, 17:261–272, 2020.

[45] Batuhan Yildirim and Hamish Galloway Brown. by256/rdfpy: rdfpy-v1.0.0, March 2021. URL `https://doi.org/10.5281/zenodo.4625675`.

A Related Work

Constraint Conditioning [40, 42] enables controllable generation by incorporating a conditioning variable to bias the posterior which is sampled. Accomplished either explicitly through the addition of classifier-derived gradient signals or implicitly by training the denoiser with specific conditioning labels, these techniques provide soft guidance towards particular subdistributions. Notably, these methods have been applied to improve constraint adherence in a variety of applications including [fill in references]. Yet, while constraint conditioning can improve feasibility rates in specific cases, it is unreliable when exact satisfaction is required. Particularly when constraint sets are complex, conditioning methods have been shown ineffective in providing viable outputs, a challenge that is demonstrated in our empirical analysis.

Post-Processing Optimization [41] provides an alternative approach which injects constraints following the denoising process. These approaches leverage a generative model to produce a starting structure $\boldsymbol{x}_0$, after which domain-specific corrections are applied. For material structure generations, this often consists of running DFT simulators [24] to relax the atomic configurations, ensuring stability as the structure is refined to reach equilibrium. Yet, these approaches are inherently void of distributional information, and the optimization procedure may drive the samples away from the learned distribution. While the original candidate structure will fall within p_{data}, without access to the learned score function or likelihood estimates, post-processing can degrade sample quality, resulting in an output distribution that is constrained but no longer resembles the training data.

B Augmented Lagrangian Method

In our settings, projection is required to ensure a collection of physical constraints. Unlike standard convex projection operators, these constraints are often highly non-convex and some are evaluated through surrogate models, such as pretrained MACE model. This makes exact projection onto feasible set intractable. To overcome this, we rely on the Augmented Lagrangian Method (ALM) [38]. Instead of attempting to solve the constrained problem in a closed form, Lagrangian relaxation converts these into differentiable residuals. By combining linear multipliers λ and quadratic penalties μ, ALM provides a mechanism to gradually enforce feasibility.

Let $\boldsymbol{x}_0$ denote the current sample produced by the diffusion process, and let $\boldsymbol{y}$ be the projected candidate. We aim to find $\boldsymbol{y}$ that remains close to $\boldsymbol{x}_t$ while reducing constraint violations. The approach introduces a relaxed objective by embedding constraint residuals into the optimization problem:

$$\mathcal{L}(\boldsymbol{y}, \boldsymbol{x}_0^{(i)}; \lambda, \mu) := \|\boldsymbol{y} - \boldsymbol{x}_0^{(i)}\|_2^2 + w \cdot \|\mathrm{force}(\boldsymbol{y})\|_2^2 + \sum_j^N \lambda_j c_j(\boldsymbol{y}) + \sum_k^N \frac{\mu_k}{2} c_k(\boldsymbol{y})^2 \tag{8}$$

where $\{c_1, \ldots, c_N\}$ denotes a series of differentiable violation residual of constraints. Additionally, we treat the force prediction provided by MACE [29] as a minimization term (scaled by w) rather than as a hard constraint, due to runtime considerations and surrogate predictive accuracy. This formulation transforms the constrained projection into a differentiable optimization problem that can be solved iteratively alongside diffusion sampling. Algorithm 1 summarizes the procedure: starting from the diffusion output $\boldsymbol{x}_t$, we iteratively compute constraint residuals, update the augmented objective, take a gradient step, and adjust multipliers and penalties.

In our implementation, the residual vector $\tilde{\phi}(y)$ collects violations from three major classes of constraints: (i) geometric constraints, including minimum interatomic distances and periodic boundaries; (ii) distributional constraints, capturing radial distribution alignment; and (iii) force minimization, based on surrogate predicted forces.

Algorithm 1: Augmented Lagrangian Projection

```
Input: $\boldsymbol{x}_t$, Lagrange multiplier: $\lambda$, quadratic penalty: $\mu$,
       scaling constant: $\alpha$, step size: $\gamma$, tolerance: $\delta$
$\boldsymbol{y} \leftarrow \boldsymbol{x}_t$
while $\sum_i c_i(\boldsymbol{y}) < \delta$ do
    for $j \leftarrow 1$ to max_inner_iter do
        $\mathcal{L}_{\mathrm{ALM}} \leftarrow$
            $\|\boldsymbol{x}_t - \boldsymbol{y}\|_2^2 + \sum_j^N \lambda_j c_j(\boldsymbol{y}) + \sum_k^N \frac{\mu_k}{2} c_k(\boldsymbol{y})^2$
        $\boldsymbol{y} \leftarrow \boldsymbol{y} - \gamma \nabla_{\boldsymbol{y}} \mathcal{L}_{\mathrm{ALM}}$
    $\lambda \leftarrow \lambda + \mu \sum_j^N \lambda_j c_j(\boldsymbol{y});\quad \mu \leftarrow \min(\alpha\mu,\ \mu_{\max})$
$\boldsymbol{x}_{t-\Delta} \leftarrow \boldsymbol{y}$
return $\boldsymbol{x}_{t-\Delta}$
```

C Evaluation Setup

Our model is trained on full DFT relaxation trajectories. The training dataset consists of approximately 13,000 structures from DFT trajectories. Each structure is represented by atom types and corresponding 3D coordinates in Cartesian space. Our generative model starts from Gaussian noise and progressively denoises toward physically relaxed structures under the targeted defect configurations. Each of the six stoichiometric configurations was evaluated with 100 generated samples, resulting in 600 structures per method. These six configurations were chosen because they are the most prevalent in our training trajectories and serve as representative cases of the four major defect categories considered [43]. All methods are evaluated under a shared random seeds to ensure comparability.

Metric computation For pairwise RMSD similarity, given a generated structure and the set of relaxed references under the same stoichiometric composition. Before measuring, both structures are centered, and atoms are matched by species. Within each species, we form the pairwise Eucliudean distance matrix between generated and reference coordinates and solve an optimal one to one assignment using the Hungarian algorithm[44]. The final RMSD is taken under this assignment and reported at the structure level of the sampling set.

The RDF characterizes the radial number density of neighbors, i.e., the probability density of finding another atom at distance r[45]. We compute $g(r)$ for generated structures under a 5Å local cutoff. Among relaxed references of the same composition, we then identify the nearest reference in RDF space by minimizing the RMSD between RDF profiles. The result is recorded as the sample's distributional deviation. This procedure provides a global distribution check that is independent of the previous RMSD assignment.

To assess proximity to locally stable states, we estimate total forces for all generated samples with a pretrained surrogate and summarize each structure by the magnitude of its total forces. Lower values indicate that the generated configuration lies closer to an equilibrium basin and is more suitable as an input to downstream relaxation. We report mean total forces over the sampled set to enable comparison across methods.

Method	Config	RMSD ↓	RDF ↓	Force (eV/Å) ↓
Conditional DM	16Bi + 21Te	1.86 ± 0.33	65.85 ± 19.75	1.85×10^{9}
	16Bi + 22Te	1.93 ± 0.36	82.11 ± 29.54	3.49×10^{9}
	18Bi + 22Te	2.10 ± 0.38	87.90 ± 39.29	1.64×10^{9}
	14Bi + 26Te	1.22 ± 0.60	40.63 ± 41.77	1.28×10^{9}
	14Bi + 24Te	0.75 ± 0.81	27.41 ± 44.46	5.52×10^{8}
	13Bi + 24Te	0.95 ± 0.51	22.17 ± 33.53	3.90×10^{8}
Projected DM	16Bi + 21Te	2.54 ± 0.55	1.11 ± 0.77	8.35×10^{2}
	16Bi + 22Te	1.41 ± 0.83	0.78 ± 0.55	7.20×10^{1}
	18Bi + 22Te	1.85 ± 1.17	1.45 ± 1.62	5.85×10^{3}
	14Bi + 26Te	1.43 ± 0.78	1.00 ± 0.67	5.04×10^{1}
	14Bi + 24Te	2.35 ± 0.45	0.94 ± 0.49	5.50×10^{1}
	13Bi + 24Te	2.93 ± 0.39	1.44 ± 0.38	1.17×10^{2}
Post-proc DM	16Bi + 21Te	2.30 ± 0.40	27.25 ± 28.30	8.45×10^{6}
	16Bi + 22Te	1.14 ± 0.65	21.82 ± 28.09	9.29×10^{6}
	18Bi + 22Te	1.41 ± 0.45	63.27 ± 35.58	1.69×10^{8}
	14Bi + 26Te	1.16 ± 0.62	50.00 ± 35.32	6.54×10^{7}
	14Bi + 24Te	2.17 ± 0.38	37.37 ± 28.29	1.61×10^{7}
	13Bi + 24Te	2.75 ± 0.34	59.24 ± 28.05	1.76×10^{7}
Ours	16Bi + 21Te	0.90 ± 0.38	0.30 ± 0.11	7.88×10^{-2}
	16Bi + 22Te	1.02 ± 0.70	0.35 ± 0.18	1.44×10^{-1}
	18Bi + 22Te	1.22 ± 0.42	0.47 ± 0.22	3.00×10^{-1}
	14Bi + 26Te	1.00 ± 0.39	0.28 ± 0.19	1.67×10^{-1}
	14Bi + 24Te	0.68 ± 0.56	0.30 ± 0.11	1.80×10^{-1}
	13Bi + 24Te	0.99 ± 0.49	0.36 ± 0.06	1.93×10^{-1}

Table 2: Complete experimental results across all six stoichiometric configurations in four defect categories: Te vacancies (16Bi + 22Te, 16Bi + 21Te), Bi vacancies (14Bi + 24Te, 13Bi + 24Te), Bi in Te antisites (18Bi + 22Te), and Te in Bi antisites (14Bi + 26Te).