SUPERFLOW: TWO-STAGE JOINT FLOW MATCHING FOR SUPERCONDUCTOR DOPING

Anonymous authors
Paper under double-blind review

ABSTRACT

Finding effective doping strategies is a difficult but valuable step in superconductor development: both the substitution choice and its doping rate matter, and an unfavorable change can lower the superconducting critical temperature T_c. Locating an optimum experimentally requires synthesizing and characterizing multiple compositions along a substitution trajectory, with substantial time and resource costs. We introduce a benchmark for parent-conditioned optimal-doping prediction. The benchmark combines literature-labeled ICSD structures with ICSD and Materials Project data from 3DSC. We use a large language model to extract experimental T_c values from source papers and correct selected 3DSC labels based on the literature. The benchmark measures doping-rate error and whether a proposed change moves toward an observed optimum. We further propose SuperFlow, a two-stage joint flow-matching framework that predicts doping rate together with crystal structure for a specified substitution axis. A first flow generates a candidate crystal. A goal head reads this result and guides a second joint refinement. SuperFlow achieves the best moving-parent and large-move doping-rate errors and direction accuracies among the evaluated methods on both datasets. Moving-parent doping-rate MAE is 0.11633 on ICSD and 0.08847 on MP. Selected case studies show that predicted doping rate changes are accompanied by local structural changes toward dataset targets. The benchmark and method provide a basis for prioritizing experiments and, with further validation, supporting industrial searches for optimal doping.

1 INTRODUCTION

Chemical substitution tunes the superconducting critical temperature T_c. In cuprates and iron-based superconductors, doping changes electronic states and local structure, producing composition-dependent superconducting phase diagrams (Keimer et al., 2015; Hosono & Kuroki, 2015). Choosing which element to replace, which dopant to introduce, and how much to substitute is therefore consequential: a change can enhance T_c, move a material beyond a favorable regime, or suppress superconductivity. Mapping a substitution series requires repeated synthesis and characterization across doping rates. A model that locates promising doping rates from an existing parent could reduce the number of experiments needed to explore a trajectory (Fig. 1).

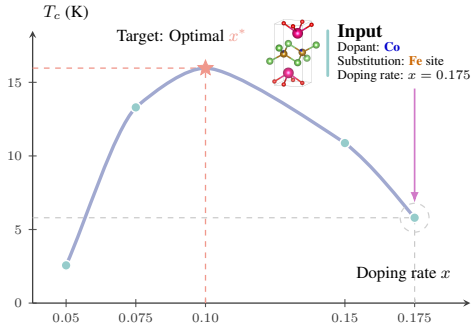

Figure 1: Parent-conditioned optimal doping. Given a parent crystal and a host–dopant substitution axis, the task is to predict a crystal whose doping rate approaches an observed T_c-maximizing doping rate along the same series. The illustrated input uses Co substitution on Fe sites. The target doping rate and temperature curve are unavailable at inference.

Learning this decision requires data that connect composition, structure, and measured superconductivity. SuperCon collects literature-derived superconducting properties (Materials Database Group, 2022), while 3DSC associates such labels with ICSD and Materials Project structures, including approximate structures obtained by artificial doping (Sommer et al., 2023). Composition-based T_c prediction has demonstrated the value of these resources (Stanev et al., 2018; Hamidieh, 2018), but evaluating a proposed doping change requires a different organization: compatible structures must be grouped into

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

substitution trajectories, observed optima must be defined, and predictions must be compared from the same parent states. Moreover, discrepancies between a recorded label and its source composition or crystal form can distort the optimization target.

We address the data problem by revisiting superconductivity evidence in papers associated with ICSD CIF records. A large language model assists article-level extraction and evidence aggregation. We compare accepted literature labels with 3DSC, apply source-grounded corrections to a subset of records, and preserve the distinction between revised labels and additional structural coverage. Combining these records with 3DSC-derived ICSD and MP data, we construct a benchmark for optimal-doping prediction (Fig. 2). Given a parent crystal and a specified host–dopant axis, a model predicts a doping rate and, where supported, a corresponding crystal state. The benchmark reports doping-rate error for all parents, including those already at an observed optimum, and evaluates both doping-rate error and movement direction for parents requiring a change, with additional results for the large-move subset.

To solve this task, we propose SuperFlow, a two-stage joint flow-matching framework built on DMFlow's crystal representation and graph backbone (Wu et al., 2026). Elemental fractions are mapped to the nonnegative sphere, while fractional coordinates retain their periodic geometry. A shared backbone jointly predicts angular, coordinate, and lattice velocities, coupling doping changes to structural motion while preserving the parent's substitution budget. In Stage I, integrating this field from the parent produces an initial crystal proposal. In Stage II, a structure-aware correction (SAC) module pools the proposal's node features and combines them with invariant descriptors of the parent and proposal. Movement, direction, and magnitude readouts define a fixed doping rate goal. The same frozen backbone then starts a second complete integration from the proposal, combining its angular velocity with goal-directed correction while continuing coordinate and lattice updates. A finite-step analysis bounds the refined endpoint's angular distance to the predicted goal in terms of its initial distance and a bound on the base angular speed.

On the moving-parent and large-move subsets of both benchmarks, SuperFlow achieves the lowest doping-rate errors and highest direction accuracies among the evaluated methods. Ablations examine the contributions of structural initialization and the second pass. Selected Cu–O and Fe–As cases show how predicted doping changes accompany local structural changes toward the corresponding dataset targets.

Our contributions are threefold:

1. **A literature-informed optimal-doping benchmark.** We correct selected 3DSC superconductivity labels using original-paper evidence and combine literature-labeled ICSD structures with 3DSC-derived ICSD/MP data. Parent-conditioned substitution trajectories, fixed partitions, and shared metrics support evaluation of doping optimization.
2. **Two-stage SuperFlow with a tracking bound.** We introduce joint crystal proposal and SAC-guided refinement through a shared flow backbone, together with a finite-step bound on angular distance to the predicted goal.
3. **Comparative and structural evaluation.** SuperFlow achieves moving-parent doping-rate MAE of 0.116 (ICSD) and 0.088 (MP), improving over One-round SuperFlow by 10.3% and 19.3%. Case studies show accompanying structural changes.

2 Related work

Superconductivity data and prediction. SuperCon collects experimentally reported superconducting properties and compositions (Materials Database Group, 2022), while ICSD provides experimentally determined inorganic crystal structures (Belsky et al., 2002). Materials Project (MP) complements these resources with structures and properties obtained predominantly through first-principles calculations and relaxation (Jain et al., 2013). By linking SuperCon labels to ICSD and MP structures, including artificially doped structures where exact matches are unavailable, 3DSC enables structure-informed T_c prediction (Sommer et al., 2023). Composition-based models also estimate T_c for supplied materials (Stanev et al., 2018; Hamidieh, 2018; Konno et al., 2021; Roter & Dordevic, 2020). Parent-conditioned doping optimization additionally requires predicting a substitution amount along a specified axis. Our benchmark organizes literature-corrected records around this decision.

Generative models for crystals and superconductors. Generative models address inverse design by producing candidate crystals. Examples include CDVAE (Xie et al., 2022), DiffCSP (Jiao et al., 2023), FlowMM (Miller et al., 2024), and property-conditioned MatterGen (Zeni et al., 2025). Geometric architectures such as SchNet (Schütt et al., 2017) and EGNN (Satorras et al., 2021) learn atomic interactions and symmetry-aware representations. DMFlow extends crystal generation to disordered site compositions. We adopt its representation and backbone (Wu et al., 2026). To steer generation toward desired conditions, models can use classifier-gradient guidance (Dhariwal & Nichol, 2021) or classifier-free guidance (Ho & Salimans, 2022). Guidance has also been developed for general flow matching (Feng et al., 2025). For superconductors, Supercon-Diffusion generates doped formulas under T_c conditioning (Zhong et al., 2024). We evaluate a parent-axis adaptation. SuperFlow performs joint proposal and goal-guided refinement for parent-specific doping optimization.

3 Preliminaries

Crystal representation. Following DMFlow (Wu et al., 2026), we represent a crystal with N sites and D element types as $\mathcal{C} = (\vec{\boldsymbol{L}}, \boldsymbol{S}, \boldsymbol{F})$. Here, $\vec{\boldsymbol{L}} \in \mathbb{R}^{3\times 3}$ is the lattice matrix, $\boldsymbol{S} \in [0,1]^{N\times D}$ contains elemental fractions, and $\boldsymbol{F} \in [0,1)^{N\times 3}$ denotes fractional coordinates. A site may be shared by multiple elements: each row $\boldsymbol{s}_i \in \Delta^{D-1}$ of $\boldsymbol{S}$ records their elemental fractions, with S_{ie} denoting the fraction of element e at site i and $\sum_{e=1}^{D} S_{ie} = 1$. The row $\boldsymbol{f}_i$ of $\boldsymbol{F}$ specifies the position of site i relative to the lattice basis.

Flow matching. We map the elemental fractions $\boldsymbol{S}$ elementwise to $\boldsymbol{U} = \sqrt{\boldsymbol{S}}$, whose rows lie on the nonnegative unit sphere. For host h and dopant d, each eligible site has a site angle $\zeta_i = \arcsin\sqrt{S_{id}/(S_{ih}+S_{id})} \in [0, \pi/2]$, which parameterizes their pair in $\boldsymbol{U}$. Thus $\sin^2\zeta_i$ is the doping rate within that site's host–dopant pair. We integrate $\boldsymbol{z} = (\zeta, \boldsymbol{F}, \boldsymbol{l})$, combining these site angles with periodic fractional coordinates $\boldsymbol{F}$ and lattice parameters $\boldsymbol{l} = \rho(\vec{\boldsymbol{L}})$. Flow matching learns a velocity field $\boldsymbol{v}_\phi(\boldsymbol{z}_t, t)$ by matching the tangent $\boldsymbol{u}_t = \partial_t \boldsymbol{z}_t$ of a training path $\boldsymbol{z}_t$, $t \in [0,1]$ (Lipman et al., 2023; Chen & Lipman, 2024). Integrating $\dot{\boldsymbol{z}}_t = \boldsymbol{v}_\phi(\boldsymbol{z}_t, t)$ then transforms an initial crystal state into a predicted endpoint.

Optimal doping from a parent. We consider substitutional doping, where a dopant element d partially replaces a host element h at eligible crystal sites $\mathcal{I}$. The parent crystal $\mathcal{C}_P$ is the starting structure to be modified and may already contain both elements. The condition $c = (h, d, \mathcal{I})$ specifies the allowed substitution between h and d at eligible sites, while other elemental fractions remain fixed. The parent fixes each eligible site's budget $m_i = S_{P,ih} + S_{P,id} > 0$. The site angles defined above describe redistribution within this budget. Atomic positions and lattice parameters may change along with the composition. The doping rate is $x(\mathcal{C}, h, d) = \mathbf{1}^\top \boldsymbol{S}\boldsymbol{e}_d / \mathbf{1}^\top \boldsymbol{S}(\boldsymbol{e}_h + \boldsymbol{e}_d)$, where $\mathbf{1} \in \mathbb{R}^N$ is the all-ones vector and $\boldsymbol{e}_h, \boldsymbol{e}_d \in \mathbb{R}^D$ are the standard basis vectors for the host and dopant. We write $x_0 = x(\mathcal{C}_P, h, d)$ for the parent doping rate. A substitution series groups compatible crystals with different doping rates along the same substitution axis. Its observed optimum set X^* contains the fractions attaining the highest recorded superconducting critical temperature T_c within that series. Given only $(\mathcal{C}_P, c)$ at inference, the task is to predict a modified crystal $\widehat{\mathcal{C}}$ whose doping rate $\widehat{x} = x(\widehat{\mathcal{C}}, h, d)$ is close to X^*. This may require increasing, decreasing, or preserving x_0. Target doping rates, structures, and T_c values are used only for supervision or evaluation. Accordingly, we condition the flow field on c, writing $\boldsymbol{v}_\phi(\boldsymbol{z}_t, t, c)$, and denote the doping rate decoded from a flow state by $x(\boldsymbol{z})$.

4 Benchmark Construction

Evaluating parent-conditioned doping requires comparable crystal structures with measured critical temperatures across a substitution series. We construct separate ICSD (Belsky et al., 2002) and Materials Project (MP) (Jain et al., 2013) benchmarks from 3DSC (Sommer et al., 2023), supplemented with literature evidence for ICSD.

Structures and superconductivity labels. For the ICSD branch, we obtain additional labels from papers associated with CIF records. Building on materials-data curation (Himanen et al., 2019)

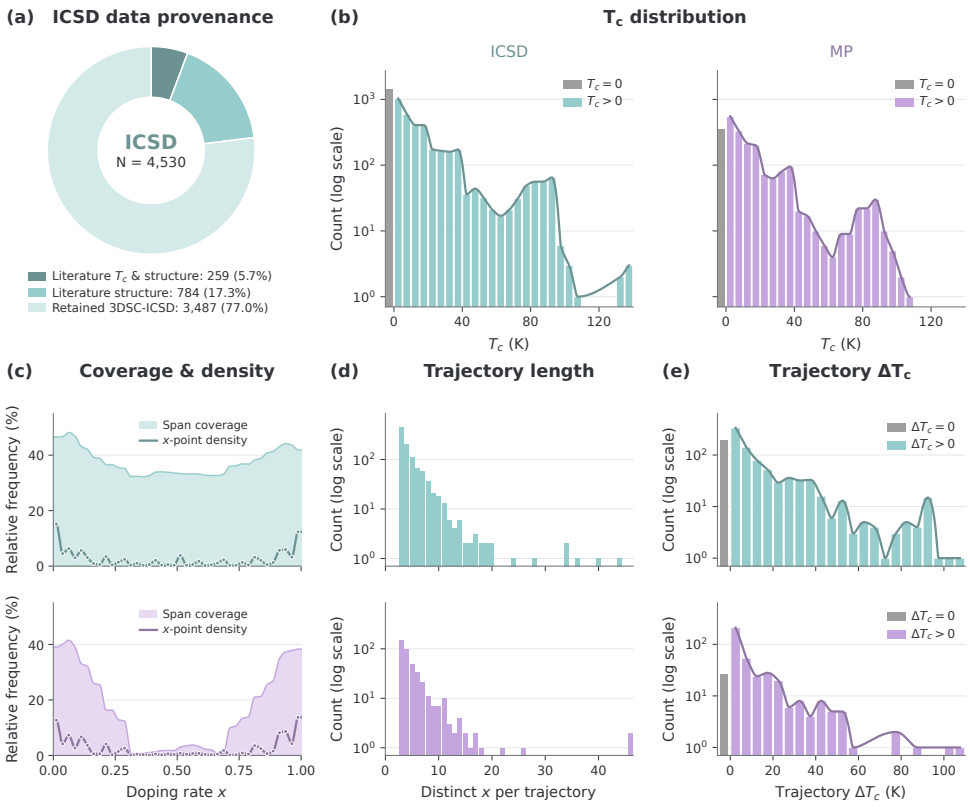

Figure 2: Benchmark overview. (a) Sources of ICSD records. (b) Experimental T_c distributions. (c) Doping-rate coverage and sampling density. (d) Number of distinct doping rates per substitution series. (e) Within-series T_c range, $\Delta T_c = \max T_c - \min T_c$. ICSD and MP appear in the upper and lower rows of (c)–(e), respectively.

and LLM-based information extraction (Dagdelen et al., 2024), we adapt the prompt design of ElementsClaw (Li et al., 2026) to extract and aggregate experimental T_c evidence using GPT-5.4. We retain experimentally supported superconducting and explicit nonsuperconducting reports, excluding unverified cases and negatives inferred from general knowledge. Labels are matched to structures by composition. Literature-supported labels take priority, correcting selected 3DSC entries. Otherwise, we use a unique composition-matched 3DSC label. The resulting ICSD inventory contains 4,530 structural records. The MP branch retains 3DSC's matched or artificially doped structures with composition-associated experimental T_c labels. The two benchmarks are evaluated separately to preserve their distinct structural provenance. The resulting data sources and substitution-series statistics are summarized in Fig. 2. Source counts and representative label corrections are detailed in Appendix A.

Substitution series and parent queries. To turn these labeled structures into doping-optimization queries, we group compatible structures along a common substitution axis. Each resulting series contains at least three distinct doping rates. The doping rates attaining the highest recorded T_c define its observed optimum set X^*. Each structure can serve as a parent $\mathcal{C}_P$, yielding a query $(\mathcal{C}_P, c)$ evaluated against the corresponding X^*. The queries therefore include both parents requiring a doping rate change and those already at an observed optimum. A structure may participate in multiple substitution axes, so structural records and parent queries have different counts.

Linked records, compositions, and source structures are kept in the same partition across substitution series. Split statistics and the evaluation protocol are provided in § 6.1.

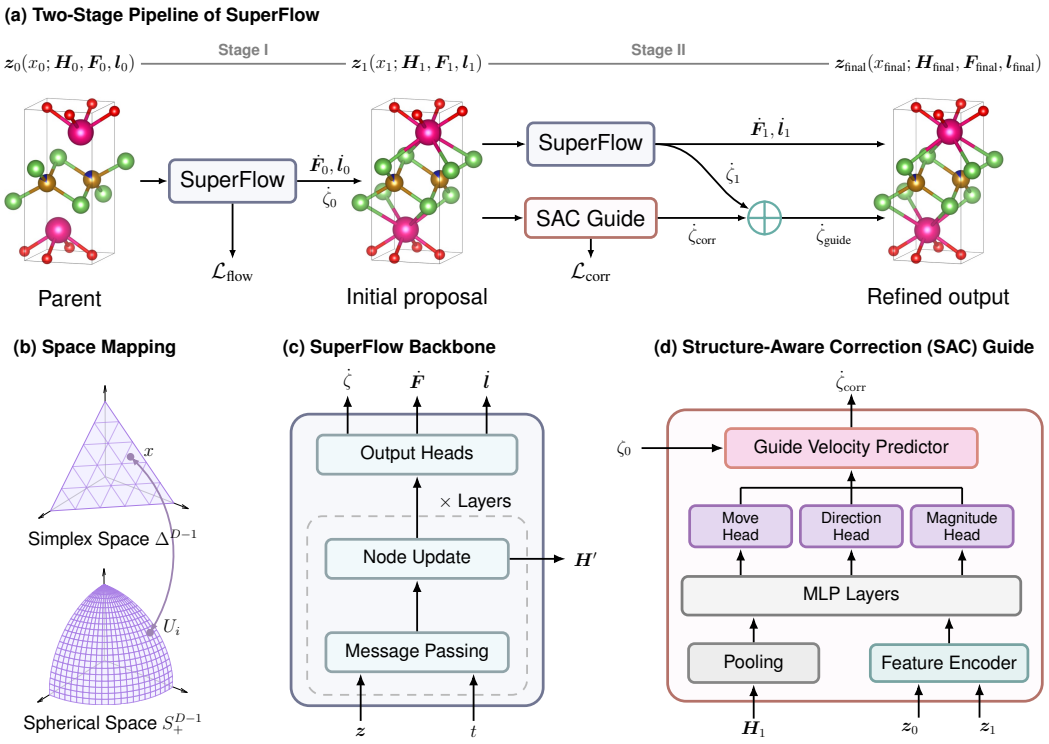

Figure 3: Overview of SuperFlow. (a) $\boldsymbol{z}_0$, $\boldsymbol{z}_1$, and $\boldsymbol{z}_{\text{final}}$ are the parent, proposal, and refined crystal. x, $\boldsymbol{H}$, $\boldsymbol{F}$, and $\boldsymbol{l}$ denote doping rate, node features, fractional coordinates, and lattice parameters. Both passes integrate the displayed velocities. The second starts from $\boldsymbol{z}_1$. The $+$ denotes the weighted angular-velocity combination in Eq. (8). (b) Spherical composition coordinates U_i and their host–dopant angles ζ_i. (c) The backbone predicts joint velocities and computes intermediate features $\boldsymbol{H}'$. (d) $\boldsymbol{H}$ denotes backbone-produced node embeddings, not independently integrated variables. Encoding the first proposal at local time zero gives $\boldsymbol{H}_1$. Pooling yields the graph-level embedding supplied to the second-stage head together with invariant descriptors of $\boldsymbol{z}_0, \boldsymbol{z}_1$. Its readouts set a fixed goal relative to the parent. Correction velocities use the current angles and time. ζ_0 denotes parent angles. $\mathcal{L}_{\text{flow}}$ and $\mathcal{L}_{\text{corr}}$ are the two training objectives.

5 Methodology

Doping optimization must couple changes in doping rate with structural motion. SuperFlow uses a shared joint-flow backbone to generate a proposal and refine it under a goal inferred by SAC, preserving the substitution budgets (Fig. 3).

5.1 Two-stage SuperFlow

SuperFlow backbone. To couple composition and geometry, the backbone operates on the joint state $\boldsymbol{z} = (\zeta, \boldsymbol{F}, \boldsymbol{l})$ of site angles, fractional coordinates, and lattice parameters. On eligible sites, doping-rate evolution preserves the parent's host–dopant budget $m_i = S_{P,ih} + S_{P,id}$ and other elemental fractions remain fixed. The DMFlow-based backbone encodes the evolving crystal and local time to produce intermediate node features $\boldsymbol{H}'$ and joint velocities:

$$(\dot{\zeta}, \dot{\boldsymbol{F}}, \dot{\boldsymbol{l}}) = \text{SuperFlow}_\phi(\boldsymbol{z}, t). \tag{1}$$

The graph is rebuilt at each field evaluation, so changes in doping rate enter the same state used to predict geometric motion. The angle representation preserves the substitution budgets, while coordinates and lattice parameters follow their native constraints. The conversion from raw spherical velocities to angular velocities is given in Appendix B.

Stage I: joint proposal. We train the first pass on paths from parent $\boldsymbol{z}_0$ to an aligned observed-optimum crystal, using

$$\mathcal{L}_{\text{flow}} = \mathcal{L}_\zeta + \lambda_F \mathcal{L}_F + \lambda_l \mathcal{L}_l. \tag{2}$$

Here $\mathcal{L}_\zeta$ is the budget-weighted squared error between predicted and target angular velocities. $\mathcal{L}_F$ and $\mathcal{L}_l$ match coordinate and lattice path velocities. Geometry targets use aligned parent–child sites. Improvement and optimum-to-itself pairs teach movement and preservation. Paths, normalization, and training schedules are detailed in Appendices B and D. At inference, the backbone evaluates $(\dot{\zeta}_0, \dot{\boldsymbol{F}}_0, \dot{\boldsymbol{l}}_0)$ on the evolving first-pass state. Joint integration gives the proposal:

$$\frac{\mathrm{d}\boldsymbol{z}^{\mathrm{I}}(t)}{\mathrm{d}t} = (\dot{\zeta}_0, \dot{\boldsymbol{F}}_0, \dot{\boldsymbol{l}}_0), \qquad \boldsymbol{z}^{\mathrm{I}}(0) = \boldsymbol{z}_0, \quad \boldsymbol{z}_1 = \boldsymbol{z}^{\mathrm{I}}(1). \tag{3}$$

Stage II: structure-aware correction (SAC) training. To predict a refinement goal, we freeze the backbone selected on validation data and encode the complete proposal $\boldsymbol{z}_1$ at local time zero to obtain $\boldsymbol{H}_1$. SAC combines its pooled features with invariant parent–proposal descriptors:

$$(a_m, a_d, a) = \mathrm{Heads}_\psi(\mathrm{Pool}(\boldsymbol{H}_1), \mathrm{FeatureEncoder}(\boldsymbol{z}_0, \boldsymbol{z}_1)). \tag{4}$$

Here Heads_ψ comprises the MLP and three readouts in Fig. 3. a_m, a_d are movement and direction logits, and $a \in [0, \pi/2]$ is an angular displacement magnitude. Feature Encoder constructs auxiliary descriptors, including initial-velocity statistics, as detailed in Appendix B.3.

The readouts are jointly supervised by

$$\mathcal{L}_{\text{corr}} = \lambda_m \mathcal{L}_{\text{move}} + \lambda_d \mathcal{L}_{\text{dir}} + \lambda_a \mathcal{L}_{\text{mag}} + \lambda_e \mathcal{L}_{\text{endpoint}}. \tag{5}$$

Movement and direction use binary classification losses. Magnitude and endpoint use smooth-L_1 supervision in angle space. Direction and magnitude terms apply only to moving parents, whereas movement and endpoint terms use all TRAIN parents. The backbone remains frozen. Appendix B.4 specifies the targets and differentiable training readout.

Stage II: joint inference. The trained readouts define the refinement goal. With $g = \mathbb{I}[\sigma(a_m) \geq \eta_m]$ and parent angle ζ_0, the goal doping rate is

$$x^\dagger = \sin^2\left[\mathrm{clip}_{[0,\pi/2]}\big(\zeta_0 + g\,\mathrm{sgn}_+(a_d)a\big)\right]. \tag{6}$$

The movement logit a_m sets the gate g. a_d selects the sign and a sets the magnitude of the change from the parent doping-rate angle. Here σ is the sigmoid and η_m is selected on VAL. The goal is distributed over eligible sites within their fixed budgets, giving target site values $d_i^\dagger$ and site angles $\zeta_i^\dagger = \arcsin\sqrt{d_i^\dagger/m_i}$, where m_i is the parent's fixed host–dopant site total and $d_i^\dagger/m_i$ is the target doping rate within that total (Appendix B.3). A closed gate retains the parent-doping rate goal and still runs refinement.

The second pass now uses the fixed target angles to refine the first proposal. At each step, SAC computes a correction from the current angles and remaining local time:

$$\dot{\zeta}_{\text{corr}}(\tau) = \frac{\zeta^\dagger - \zeta(\tau)}{\max(1-\tau, \Delta\tau)}, \qquad \Delta\tau = 1/K. \tag{7}$$

The frozen backbone supplies $(\dot{\zeta}_1, \dot{\boldsymbol{F}}_1, \dot{\boldsymbol{l}}_1)$ at the current second-pass state. SAC modifies its angular component:

$$\dot{\zeta}_{\text{guide}} = (1-\beta)\dot{\zeta}_1 + \beta\dot{\zeta}_{\text{corr}}, \qquad 0 < \beta < 1. \tag{8}$$

With local time reset, the second pass integrates from the complete first proposal:

$$\frac{\mathrm{d}\boldsymbol{z}^{\mathrm{II}}(\tau)}{\mathrm{d}\tau} = (\dot{\zeta}_{\text{guide}}, \dot{\boldsymbol{F}}_1, \dot{\boldsymbol{l}}_1), \qquad \boldsymbol{z}^{\mathrm{II}}(0) = \boldsymbol{z}_1, \quad \boldsymbol{z}_{\text{final}} = \boldsymbol{z}^{\mathrm{II}}(1). \tag{9}$$

The target is predicted once, while base and correction velocities are evaluated at each integration step. Coordinate and lattice velocities use the same pre-update state as the angular field. The head supplies no geometric endpoint. The final crystal is decoded from $\boldsymbol{z}_{\text{final}}$, with $\widehat{x} = x(\boldsymbol{z}_{\text{final}})$. Both passes use the projected site-angle integration scheme described in Appendix B.

5.2 THEORETICAL ANALYSIS: FINITE-STEP BRIDGE TRACKING

For the guided second pass, we bound angular deviation from the fixed SAC target. Let $E_k = \|\zeta_k - \zeta^\dagger\|_m$ be the budget-weighted angular distance to the fixed target after k refinement steps, where $\|u\|_m^2 = \sum_{i\in\mathcal{I}} m_i u_i^2 / \sum_{i\in\mathcal{I}} m_i$.

Proposition 5.1 (Finite-step bridge tracking)**.** *For K projected Euler steps, suppose $0 < \beta < 1$, the target is feasible, the parent budgets and target remain fixed, and the backbone angular field has angular speed at most M in this metric at every step. Then*

$$E_K \le A_K(\beta)(E_0 + M), \qquad A_K(\beta) = \textstyle\prod_{r=1}^{K} \left(1 - \frac{\beta}{r}\right). \tag{10}$$

Here E_0 and E_K measure the proposal's and refined endpoint's angular distances to the SAC target. The bound separates the initial-deviation contribution $A_K E_0$ from the learned-field contribution $A_K M$. Each factor of A_K lies between zero and one, attenuating the initial-deviation term. The backbone can still depend on the evolving coordinates and lattice. M bounds its angular speed along that joint trajectory.

Clipping cannot increase distance to a feasible target. Consequently, each step satisfies $E_{k+1} \le (1 - \beta/(K-k))E_k + (1-\beta)M/K$. Unrolling this recurrence and summing the remaining-step weights gives Eq. (10), an explicit endpoint tracking bound for the implemented sampler. The full derivation is in Appendix C.

6 EXPERIMENTS

We evaluate doping accuracy, structural changes, and the contributions of initialization and refinement.

6.1 EXPERIMENTAL SETUP

Baselines and ablations. **1.** Uniform random samples a doping rate uniformly on $[0, 1]$. We report its exact expected metrics. **2.** Parent-only Direct-x fine-tunes a structurally initialized graph backbone to predict doping rate from the parent and substitution condition. **3.** Supercon-Diffusion-Axis adapts the composition generator of Zhong et al. (2024), fixing non-axis composition and using high-T_c category conditioning. **4.** Tc-gradient-guided SuperFlow jointly fine-tunes the first-stage actor with a scalar T_c regressor, then adds its angular gradient during a single rollout. **5.** One-round SuperFlow removes the second pass. **6.** Train-from-scratch uses random initialization with its documented training schedule. Baseline, ablation, and guidance settings are provided in Appendix E.

Data splits and model selection. We partition parent queries into training, validation, and test sets, keeping linked records, compositions, and source structures in the same partition across substitution series. The test sets contain 1,002 ICSD and 424 MP parents, with full split statistics provided in Appendix A. Supervised doping training uses the benchmark training split. Initialization follows the dataset-specific pretrained lineages in Appendix D. Actor and goal-head checkpoints, together with guidance settings, are selected on validation data, primarily by moving-parent doping-rate MAE after complete sampling. The selected movement threshold is shared across both datasets and all evaluation subsets. Detailed settings are provided in Appendix D.

Evaluation metrics. We evaluate each selected model on three subsets. *All parents* include every test query. *Moving parents* satisfy $\min_{x^*\in X^*} |x_0 - x^*| > 10^{-8}$, while *large moves* require this distance to exceed 0.1. For a predicted doping rate $\widehat{x}$, the error is $e_x = \min_{x^*\in X^*} |\widehat{x} - x^*|$. Its mean gives doping-rate MAE (xMAE), which we report for all three subsets. For moving parents, direction is correct if $(\widehat{x} - x_0)(x^* - x_0) > 0$ for at least one $x^* \in X^*$. We report direction accuracy on the moving-parent and large-move subsets. All methods use identical test queries within each subset.

6.2 RESULTS AND ANALYSIS

Across both datasets, SuperFlow predicts doping rates closer to observed optima and selects the correct movement direction more often than Parent-only Direct-x and Supercon-Diffusion-Axis (Table 1). The second pass further improves One-round SuperFlow on moving parents, including large moves,

indicating that the goal inferred from the first generated crystal provides useful refinement. The pretrained actor also achieves lower moving-parent error than the scratch control under the reported training settings. Tc-gradient-guided SuperFlow gives mixed improvements across datasets and remains less accurate than the two-stage model on moving and large-move cases. These gains come with a preservation tradeoff on ICSD: the complete model has higher all-parent error than One-round SuperFlow and Tc-gradient-guided SuperFlow. Per-parent error and signed-movement distributions are provided in Appendix F.

Table 1: TEST performance on the ICSD and MP benchmarks, including initialization and second-pass ablations. Bold indicates the best result in each metric column for each dataset.

Dataset	Method	All parents xMAE ↓	Moving parents xMAE ↓	Moving parents Dir. (%) ↑	Large moves xMAE ↓	Large moves Dir. (%) ↑
ICSD	Uniform random	0.366	0.427	44.21	0.414	61.13
	Parent-only Direct-x	0.341	0.403	47.08	0.425	59.66
	Supercon-Diffusion-Axis	0.337	0.406	40.72	0.393	58.52
	Train-from-scratch	0.140	0.176	43.64	0.293	57.95
	One-round SuperFlow	0.089	0.130	63.40	0.307	64.20
	Tc-gradient-guided SuperFlow	**0.086**	0.128	62.71	0.306	64.20
	SuperFlow	0.105	**0.116**	**67.53**	**0.253**	**72.73**
MP	Uniform random	0.428	0.432	41.04	0.436	42.07
	Parent-only Direct-x	0.351	0.350	44.82	0.434	37.21
	Supercon-Diffusion-Axis	0.407	0.411	35.06	0.414	31.40
	Train-from-scratch	0.108	0.118	59.15	0.216	60.47
	One-round SuperFlow	0.093	0.110	64.94	0.142	69.77
	Tc-gradient-guided SuperFlow	0.090	0.107	66.16	0.138	70.93
	SuperFlow	**0.082**	**0.088**	**71.95**	**0.120**	**80.23**

6.3 Case studies: doping trajectories and local structure

To examine the structural changes behind the aggregate results, we follow selected ICSD TEST predictions through both passes (Fig. 4) and locate their final doping rates along observed substitution series (Fig. 5). The T_c values belong to dataset records. Generated crystals have no experimentally measured T_c in this evaluation.

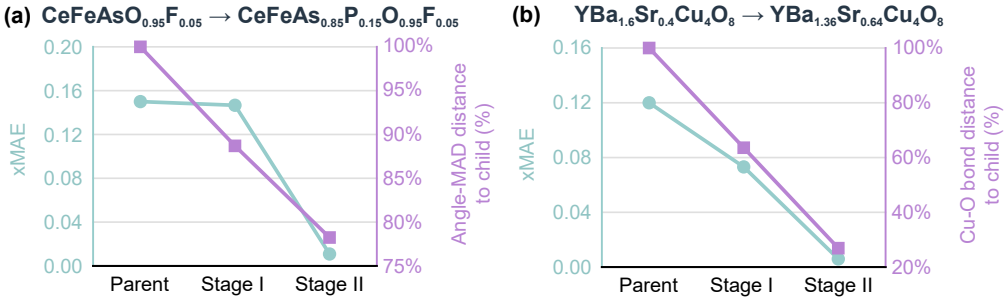

Figure 4: Two-stage errors: Ce–P (left), Y124–Sr (right). Teal: doping-rate error (xMAE). Purple: structural error relative to the target child, normalized by the parent error.

Joint refinement of doping rate and structure. In Fe–As superconductors, higher T_c has been associated with As–Fe–As angles approaching $109.5°$, the regular tetrahedral angle (Lee et al., 2008; Kamihara et al., 2008). In La-based cuprates at fixed doping, the highest T_c is found for flat, square CuO_2 planes (Dabrowski et al., 1996). Motivated by these structural trends, we compare the angle-MAD distance to the optimal child for the iron-based case and the mean planar Cu–O bond-length distance for the cuprate case.

The two cases exhibit different correction patterns. For Ce–P, Stage I makes only a small change in doping-rate error, while Stage II brings the composition much closer to the observed optimum. The

angle-MAD distance decreases across both passes. For Y124–Sr, doping rate and Cu–O bond-length errors already decrease in Stage I and fall further during refinement. The first pass therefore supplies a useful but incomplete proposal in these examples, and the second pass improves both its doping rate and the associated local structural descriptor.

Doping direction along observed trajectories. We next examine doping direction using $\Delta x = x - x_0$, which places each parent at zero (Fig. 5). Two examples start from the same undoped PrFeAsO parent, recorded with the nonsuperconducting label $T_c = 0$. For Fe→Co substitution, SuperFlow predicts a Co doping rate of $\widehat{x} = 0.06198$, close to the observed optimum $x^* = 0.075$, whose recorded T_c is 15.03 K. For Pr→Sr substitution, it predicts $\widehat{x} = 0.24449$ against $x^* = 0.25$, with a recorded optimum T_c of 16.25 K. The doping-rate errors are 0.01302 and 0.00551, respectively.

The As→P example begins with P-free $CeFeAsO_{0.95}F_{0.05}$, whose recorded T_c is 11.8 K. Its prediction $\widehat{x} = 0.16088$ is close to the observed P optimum $x^* = 0.15$ at 21.75 K, a recorded gain of 9.95 K. Conversely, the Mn-containing Nd(O,F) parent has $x_0 = 0.03$ and a recorded T_c of 11.9 K. The model reduces the Mn doping rate to 0.00506, approaching the Mn-free observed optimum at $x^* = 0$ and 49.05 K. Together with the decreasing Rh and Ir examples, this illustrates why both the sign and amount of substitution must be inferred. Circle colors indicate recorded T_c. Diamonds show predicted doping rates, without assigning the observed optimum's temperature to a generated crystal.

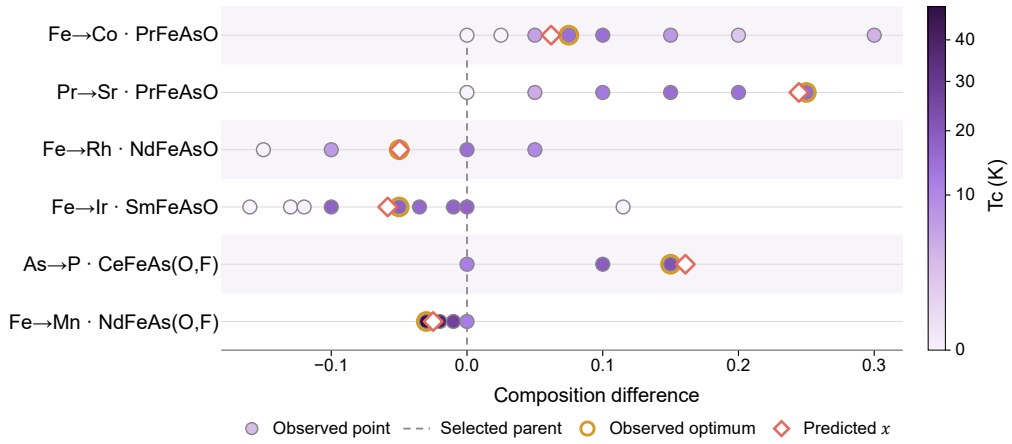

Figure 5: Selected substitution trajectories relative to the input parent, $\Delta x = x - x_0$ (the plotted x_P denotes x_0). Colored circles show observed compositions and their recorded T_c, orange outlines mark observed optima, and diamonds mark predicted doping rates. The vertical dashed line is the selected parent. Each row represents a separate host–dopant axis. Positive and negative displacements increase and decrease the doping rate, respectively.

7 Conclusion

We introduced a literature-informed benchmark for parent-conditioned optimal doping, combining corrections to selected 3DSC labels with ICSD and MP structures organized into substitution series. The benchmark evaluates both doping-rate accuracy and movement toward observed optima from shared parent states. To address this task, SuperFlow jointly evolves doping rate, atomic coordinates, and lattice parameters through an initial proposal and SAC-guided refinement.

SuperFlow achieves the strongest moving-parent and large-move results among the evaluated methods. Ablations assess initialization and refinement, while selected case studies illustrate structural changes accompanying the predicted doping adjustments. A finite-step analysis bounds tracking of the predicted goal. Together, the benchmark and method provide a basis for developing doping-optimization models and prioritizing experimental candidates. Future experimental validation can assess their practical value in locating favorable doping rates for superconductor development.

AI USE STATEMENT

We used Codex to improve the language and readability of the manuscript. We also used GPT-5.4 to extract and synthesize superconductivity evidence from source papers, as described in § 4. The authors take responsibility for the final text, theoretical claims, and reported results, including content developed with AI assistance.

ETHICS STATEMENT

This work studies computational methods for superconductor doping and does not involve human participants or personal data. We do not identify particular ethical concerns arising from this study. Use and disclosure of ICSD-derived data will follow the applicable requirements of the database provider.

REPRODUCIBILITY STATEMENT

We will publicly release the code upon acceptance of the paper. Data construction and evaluation protocols are described in § 4 and 6.1. Training settings and hyperparameters are provided in Appendix D, and the theoretical proof is given in Appendix C. Portions of the data originate from ICSD. Their disclosure and any redistribution will follow ICSD's official requirements and applicable licensing terms. The code release does not imply unrestricted redistribution of these records.

REFERENCES

A. Belsky, M. Hellenbrandt, V. L. Karen, and P. Luksch. New developments in the inorganic crystal structure database (ICSD): accessibility in support of materials research and design. *Acta Crystallographica Section B: Structural Science*, 58(3):364–369, 2002. doi: 10.1107/S0108768102006948. URL https://doi.org/10.1107/S0108768102006948.

Ricky T. Q. Chen and Yaron Lipman. Flow matching on general geometries. In *The Twelfth International Conference on Learning Representations*. OpenReview.net, 2024. URL https://openreview.net/forum?id=g7ohDlTITL.

B. Dabrowski, Z. Wang, K. Rogacki, J. D. Jorgensen, R. L. Hitterman, J. L. Wagner, B. A. Hunter, P. G. Radaelli, and D. G. Hinks. Dependence of superconducting transition temperature on doping and structural distortion of the CuO_2 planes in $La_{2-x}m_xCuO_4$ (M = Nd, Ca, Sr). *Physical Review Letters*, 76(8):1348–1351, 1996. doi: 10.1103/PhysRevLett.76.1348. URL https://doi.org/10.1103/PhysRevLett.76.1348.

John Dagdelen, Alexander Dunn, Sanghoon Lee, Nicholas Walker, Andrew S Rosen, Gerbrand Ceder, Kristin A Persson, and Anubhav Jain. Structured information extraction from scientific text with large language models. *Nature communications*, 15(1):1418, 2024.

Prafulla Dhariwal and Alexander Nichol. Diffusion models beat GANs on image synthesis. In *Advances in Neural Information Processing Systems*, volume 34, 2021. URL https://proceedings.neurips.cc/paper/2021/hash/49ad23d1ec9fa4bd8d77d02681df5cfa-Abstract.html.

Ruiqi Feng, Chenglei Yu, Wenhao Deng, Peiyan Hu, and Tailin Wu. On the guidance of flow matching. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267, pp. 16993–17029. PMLR, 2025. URL https://proceedings.mlr.press/v267/feng25s.html.

Justin Gilmer, Samuel S. Schoenholz, Patrick F. Riley, Oriol Vinyals, and George E. Dahl. Neural message passing for quantum chemistry. In *Proceedings of the 34th International Conference on Machine Learning*, volume 70, pp. 1263–1272. PMLR, 2017. URL https://proceedings.mlr.press/v70/gilmer17a.html.

Kam Hamidieh. A data-driven statistical model for predicting the critical temperature of a superconductor. *Computational Materials Science*, 154:346–354, 2018. doi: 10.1016/j.commatsci.2018.07.052. URL https://doi.org/10.1016/j.commatsci.2018.07.052.

Lauri Himanen, Amber Geurts, Adam Stuart Foster, and Patrick Rinke. Data-driven materials science: status, challenges, and perspectives. *Advanced Science*, 6(21):1900808, 2019.

Jonathan Ho and Tim Salimans. Classifier-free diffusion guidance. *arXiv preprint arXiv:2207.12598*, 2022. URL https://arxiv.org/abs/2207.12598.

Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. In *Advances in Neural Information Processing Systems*, volume 33, 2020. URL https://proceedings.neurips.cc/paper/2020/hash/4c5bcfec8584af0d967f1ab10179ca4b-Abstract.html.

Hideo Hosono and Kazuhiko Kuroki. Iron-based superconductors: current status of materials and pairing mechanism. *Physica C: Superconductivity and its Applications*, 514:399–422, 2015. doi: 10.1016/j.physc.2015.02.020. URL https://doi.org/10.1016/j.physc.2015.02.020.

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

Rui Jiao, Wenbing Huang, Peijia Lin, Jiaqi Han, Pin Chen, Yutong Lu, and Yang Liu. Crystal structure prediction by joint equivariant diffusion. In *Advances in Neural Information Processing Systems*, volume 36, pp. 17464–17497. Curran Associates, Inc., 2023. doi: 10.52202/075280-0767. URL https://proceedings.neurips.cc/paper_files/paper/2023/hash/38b787fc530d0b31825827e2cc306656-Abstract.html.

Yoichi Kamihara, Takumi Watanabe, Masahiro Hirano, and Hideo Hosono. Iron-based layered superconductor la [o1-x f x] feas (x= 0.05- 0.12) with t c= 26 k. *Journal of the American Chemical Society*, 130(11):3296–3297, 2008.

B. Keimer, S. A. Kivelson, M. R. Norman, S. Uchida, and J. Zaanen. From quantum matter to high-temperature superconductivity in copper oxides. *Nature*, 518:179–186, 2015. doi: 10.1038/nature14165. URL https://doi.org/10.1038/nature14165.

Tomohiko Konno, Hodaka Kurokawa, Fuyuki Nabeshima, Yuki Sakishita, Ryo Ogawa, Iwao Hosako, and Atsutaka Maeda. Deep learning model for finding new superconductors. *Physical Review B*, 103(1):014509, 2021.

Chul-Ho Lee, Akira Iyo, Hiroshi Eisaki, Hijiri Kito, Maria Teresa Fernandez-Diaz, Toshimitsu Ito, Kunihiro Kihou, Hirofumi Matsuhata, Markus Braden, and Kazuyoshi Yamada. Effect of structural parameters on superconductivity in fluorine-free $LnFeAsO_{1-y}$ (Ln = La, Nd). *Journal of the Physical Society of Japan*, 77(8):083704, 2008. doi: 10.1143/JPSJ.77.083704. URL https://doi.org/10.1143/JPSJ.77.083704.

Mingze Li, Yu Rong, Songyou Li, Lihong Wang, Jiacheng Cen, Liming Wu, Anyi Li, Zongzhao Li, Qiuliang Liu, Rui Jiao, Tian Bian, Pengju Wang, Hao Sun, Jianfeng Zhang, Ji-Rong Wen, Deli Zhao, Shifeng Jin, Tingyang Xu, and Wenbing Huang. Agentic fusion of large atomic and language models to accelerate superconductor discovery. *arXiv preprint arXiv:2604.23758*, 2026. doi: 10.48550/arXiv.2604.23758. URL https://arxiv.org/abs/2604.23758.

Yaron Lipman, Ricky T. Q. Chen, Heli Ben-Hamu, Maximilian Nickel, and Matt Le. Flow matching for generative modeling. In *The Eleventh International Conference on Learning Representations*. OpenReview.net, 2023. URL https://openreview.net/forum?id=PqvMRDCJT9t.

Materials Database Group. MDR SuperCon Datasheet Ver.220808, 2022. URL https://doi.org/10.48505/nims.3837.

Benjamin Kurt Miller, Ricky T. Q. Chen, Anuroop Sriram, and Brandon M. Wood. FlowMM: Generating materials with Riemannian flow matching. In Ruslan Salakhutdinov, Zico Kolter, Katherine Heller, Adrian Weller, Nuria Oliver, Jonathan Scarlett, and Felix Berkenkamp (eds.),

Proceedings of the 41st International Conference on Machine Learning, volume 235 of *Proceedings of Machine Learning Research*, pp. 35664–35686. PMLR, 2024. URL https://proceedings.mlr.press/v235/miller24a.html.

B Roter and SV Dordevic. Predicting new superconductors and their critical temperatures using machine learning. *Physica C: Superconductivity and its applications*, 575:1353689, 2020.

Víctor Garcia Satorras, Emiel Hoogeboom, and Max Welling. E(n) equivariant graph neural networks. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139, pp. 9323–9332. PMLR, 2021. URL https://proceedings.mlr.press/v139/satorras21a.html.

Kristof T. Schütt, Pieter-Jan Kindermans, Huziel E. Sauceda, Stefan Chmiela, Alexandre Tkatchenko, and Klaus-Robert Müller. SchNet: A continuous-filter convolutional neural network for modeling quantum interactions. In *Advances in Neural Information Processing Systems*, volume 30, 2017. URL https://proceedings.neurips.cc/paper/2017/hash/303ed4c69846ab36c2904d3ba8573050-Abstract.html.

Timo Sommer, Roland Willa, Jörg Schmalian, and Pascal Friederich. 3DSC: A dataset of superconductors including crystal structures. *Scientific Data*, 10:816, 2023. doi: 10.1038/s41597-023-02721-y. URL https://doi.org/10.1038/s41597-023-02721-y.

Valentin Stanev, Corey Oses, A. Gilad Kusne, Efrain Rodriguez, Johnpierre Paglione, Stefano Curtarolo, and Ichiro Takeuchi. Machine learning modeling of superconducting critical temperature. *npj Computational Materials*, 4:29, 2018. doi: 10.1038/s41524-018-0085-8. URL https://doi.org/10.1038/s41524-018-0085-8.

Liming Wu, Rui Jiao, Qi Li, Mingze Li, Songyou Li, Shifeng Jin, and Wenbing Huang. DMFlow: Disordered materials generation by flow matching. In *32nd SIGKDD Conference on Knowledge Discovery and Data Mining – AI for Sciences Track*, 2026.

Tian Xie, Xiang Fu, Octavian-Eugen Ganea, Regina Barzilay, and Tommi Jaakkola. Crystal diffusion variational autoencoder for periodic material generation. In *The Tenth International Conference on Learning Representations*. OpenReview.net, 2022.

Claudio Zeni, Robert Pinsler, Daniel Zügner, Andrew Fowler, Matthew Horton, Xiang Fu, Zilong Wang, Aliaksandra Shysheya, Jonathan Crabbé, Shoko Ueda, Roberto Sordillo, Lixin Sun, Jake Smith, Bichlien Nguyen, Hannes Schulz, Sarah Lewis, Chin-Wei Huang, Ziheng Lu, Yichi Zhou, Han Yang, Hongxia Hao, Jielan Li, Chunlei Yang, Wenjie Li, Ryota Tomioka, and Tian Xie. A generative model for inorganic materials design. *Nature*, 639:624–632, 2025. doi: 10.1038/s41586-025-08628-5. URL https://www.nature.com/articles/s41586-025-08628-5.

Chengquan Zhong, Jingzi Zhang, Yuelin Wang, Yanwu Long, Pengzhou Zhu, Jiakai Liu, Kailong Hu, Junjie Chen, and Xi Lin. High-performance diffusion model for inverse design of high T_c superconductors with effective doping and accurate stoichiometry. *InfoMat*, 6(5):e12519, 2024. doi: 10.1002/inf2.12519. URL https://doi.org/10.1002/inf2.12519.

CONTENTS OF APPENDIX

A Dataset construction and label evidence **14**

A.1 Structural corpus and literature labels 14

A.2 Benchmark partitions and case definitions 14

A.3 Source-paper checks of label discrepancies 15

B Model architecture and implementation **16**

B.1 State representation and doping rate readout 16

B.2 SuperFlow Backbone 17

B.3 Structure-Aware Correction (SAC) 17

B.4 Training objectives 19

B.5 Two-stage sampling 20

C Theoretical analysis and proof **21**

C.1 One-step recurrence and finite-step bound 21

D Training settings and hyperparameters **22**

D.1 Structural pretraining 22

D.2 First-stage joint flow 22

D.3 Goal-head training and second-pass settings 23

E Baseline and ablation details **24**

E.1 Uniform random 24

E.2 Parent-only Direct-x 24

E.3 Supercon-Diffusion-Axis 24

E.4 Tc-gradient-guided SuperFlow 25

E.5 One-round SuperFlow 25

E.6 Train-from-scratch control 25

F Supplementary results **26**

F.1 Doping-rate error 26

F.2 Signed movement toward an observed optimum 27

A Dataset construction and label evidence

This appendix distinguishes the structural corpus used for pretraining from the ICSD and MP benchmarks used for doping prediction. We describe literature-label collection, benchmark partitions, and representative source-paper checks in turn.

A.1 Structural corpus and literature labels

The structural pretraining corpus contains 91,976 structures from the ICSD 2017 CIF collection after filtering, each with at most 100 atomic sites. The corpus is split into 82,779 training and 9,197 validation structures at a 9:1 ratio. Structural pretraining learns a flow from random initial states to crystal structures, with settings given in Appendix D. Literature-label collection supplies the downstream benchmark described below.

GPT-5.4 assists article-level evidence extraction and aggregation using a prompt adapted from ElementsClaw (Li et al., 2026). Labels are retained when supported by experimental superconductivity or explicit nonsuperconductivity reports. Unverified claims and negatives inferred from general knowledge are excluded from the literature-label pool.

A.2 Benchmark partitions and case definitions

A parent query consists of a starting crystal and a specified substitution axis. The same structural record can therefore give rise to multiple queries. Table 2 summarizes the fixed query partitions and evaluation subsets for both benchmarks.

Table 2: Fixed benchmark partitions. Each parent query has one assigned target association, including identity cases. Moving-parent and large-move subsets follow the doping-rate distance criteria in § 6.1.

Dataset	Split	Trajectories	Parent queries	Moving parents	Large moves
ICSD	TRAIN	603	2,933	2,069	830
	VAL	210	1,051	742	175
	TEST	197	1,002	582	176
MP	TRAIN	243	1,273	964	196
	VAL	81	424	316	53
	TEST	81	424	328	86

The 4,530-record ICSD benchmark combines three sources of structures and labels: 1,140 literature-labeled local structures, 214 local structures using formula-associated 3DSC labels, and 3,176 structures constructed or reused through the 3DSC pipeline. Within the first group, 259 records have a different label from the corresponding archived 3DSC composition, 97 agree, and 784 lack a same-composition entry. These counts describe where structures and their labels originate. Fig. 2 instead groups records by their contribution relative to 3DSC: 259 have revised labels, 784 add structural coverage for compositions absent from 3DSC, and the remaining $97 + 214 + 3{,}176 = 3{,}487$ retain 3DSC-consistent labels.

A flat trajectory has the same recorded T_c at every doping rate, so every parent in it is already at an observed optimum. The ICSD splits contain 82, 34 and 77 such trajectories in TRAIN, VAL and TEST, while MP has 16, 6 and 5. The corresponding nonflat counts are 521/176/120 and 227/75/76. All-parent evaluation includes flat trajectories. The training subsets used by the joint flow and SAC goal head are distinguished below.

Trajectory construction and training pairs. ICSD and MP are evaluated separately. Trajectories are constructed from compatible crystal frameworks and a fixed substitution axis, with at least three distinct doping rates. Linked records, compositions, and source structures are assigned together to TRAIN, VAL, or TEST.

Each query has one assigned parent–target association, with the target drawn from the observed T_c maxima of its series. Joint-flow training uses the assigned target crystal and retains the parent itself

when it is already optimal. SAC supervises the doping rate change toward the nearest member of X^* to x_0, as specified in Appendix B.4. Evaluation retains the full set X^* and uses the nearest-optimum error and any-optimum direction rule in § 6.1.

The joint-flow training stream samples lower-to-optimum and optimum-to-itself pairs from nonflat trajectories. Here, lower parents have a recorded T_c below the series maximum, while peak parents already attain that maximum. The SAC goal head uses the complete TRAIN query set, including flat trajectories, for its endpoint and movement supervision. Table 2 reports the complete query counts alongside the moving-parent subsets.

A.3 Source-paper checks of label discrepancies

Table 3 compares the archived 3DSC labels with the adopted literature labels and identifies the supporting source-paper evidence.

Table 3: Ten source-paper label comparisons, with temperatures in kelvin. "Literature label" is the adopted value in our frozen inventory. Zero is a negative-label code for the stated conditions, not a measurement of a transition at zero kelvin. References and evidence locations follow the table.

Composition	3DSC	Literature label	Source-paper evidence	Ref.
$CuCrO_2$	132	0	Antiferromagnetic semiconductor. The 24 K anomaly is T_N.	[S1]
$HgBa_2Ca_4Cu_5O_{12.32}$	0	101	As-prepared Hg-1245, $\delta = 0.32(3)$: $T_c = 101$ K.	[S2]
$Sr_2CuO_2F_2$	44.33	0	Stoichiometric tetragonal T' phase is nonsuperconducting.	[S3]
K_3C_{60}	0	19	Single-phase fcc material: measured $T_c = 19.3$ K. Stored value is rounded.	[S4]
Rb_4C_{60}	29	0	Insulating tetragonal A_4C_{60} phase, distinct from superconducting A_3C_{60}.	[S5]
$YBa_{1.78}Sr_{0.22}Cu_4O_8$	0	76	Refined $x = 0.11$ crystal. $T_{c,\mathrm{eff}}$ falls from 77.7 K by about 1.7 K.	[S6]
$YBa_{1.38}Sr_{0.62}Cu_4O_8$	0	80.5	Refined $x = 0.31$ crystal: $T_{c,\mathrm{eff}} = 80.5$ K.	[S6]
$NbSe_3$	3.74	0	Pure material shows no superconductivity above 0.38 K at ambient pressure.	[S7]
Li_2CsC_{60}	12.0	0	The cubic fulleride is explicitly reported as nonsuperconducting.	[S8]
LaFeAsO	10.2	0	The undoped parent is nonsuperconducting. F-substituted samples are distinct.	[S9]

Composition and condition checks. The Hg-1245 source assigns 101 K to the $\delta = 0.32$ main phase and approximately 114 K to a minority phase. The latter is not the label of the listed structure. The two Sr-substituted Y124 entries are distinct refined compositions from the same study: the low-substitution single crystal has an effective transition near 76 K, while the higher-substitution crystal has 80.5 K. The paper separately reports polycrystalline values, which are not substituted for these single-crystal values. For K_3C_{60}, the source value is 19.3 K and the stored label is 19 K. Negative examples distinguish stoichiometry, bulk phase and measurement conditions: Rb_4C_{60} is an insulating fulleride, stoichiometric $Sr_2CuO_2F_2$ differs from its excess-fluorine superconducting derivative, Li_2CsC_{60} is identified as a nonsuperconducting cubic fulleride, and the LaFeAsO comparison concerns the undoped parent rather than the F-substituted superconducting samples. These label assignments apply to the reported compositions, phases, and measurement conditions.

Original references and evidence locations.

S1. A. Maignan, C. Martin, R. Frésard, V. Eyert, E. Guilmeau, S. Hébert, M. Poienar and D. Pelloquin. *On the strong impact of doping in the triangular antiferromagnet $CuCrO_2$*. Solid State Communications 149, 962–967 (2009). See the abstract and transport/magnetic measurements. doi:10.1016/i.ssc.2009.02.026, open manuscript.

S2. Q. Huang, O. Chmaissem, J. J. Capponi, C. Chaillout, M. Marezio, J. L. Tholence and A. Santoro. *Neutron powder diffraction study of the crystal structure of $HgBa_2Ca_4Cu_5O_{12+\delta}$ at room temperature and at 10 K*. Physica C 227, 1–9 (1994). The abstract gives composition, main-phase and minority-phase transitions. doi:10.1016/0921-4534(94)90349-2.

S3. J. L. Kissick, C. Greaves, P. P. Edwards, V. M. Cherkashenko, E. Z. Kurmaev, S. Bartkowski and M. Neumann. *Synthesis, structure, and XPS characterization of the stoichiometric*

phase $Sr_2CuO_2F_2$. Physical Review B 56, 2831–2835 (1997). See Section IV, p. 2835. doi:10.1103/PhysRevB.56.2831.

S4. P. W. Stephens, L. Mihaly, P. L. Lee, R. L. Whetten, S.-M. Huang, R. Kaner, F. Diederich and K. Holczer. *Structure of single-phase superconducting K_3C_{60}*. Nature 351, 632–634 (1991). The abstract identifies the fcc superconducting phase and its 19.3 K transition. The local CIF cites this structural study. doi:10.1038/351632a0.

S5. G. Klupp, K. Kamarás, N. M. Nemes, C. M. Brown and J. Leão. *Static and dynamic Jahn–Teller effect in the alkali metal fulleride salts A_4C_{60} (A=K,Rb,Cs)*. Physical Review B 73, 085415 (2006). See the introduction and infrared/neutron results. doi:10.1103/PhysRevB.73.085415, author-hosted full text.

S6. J. Karpinski, S. M. Kazakov, M. Angst, A. Mironov, M. Mali and J. Roos. *Influence of Sr substitution on the structure, charge distribution, and critical temperature of $Y(Ba_{1-x}Sr_x)_2Cu_4O_8$ single crystals*. Physical Review B 64, 094518 (2001). Tables I and II identify the refined compositions and structures. Section III.A gives the single-crystal transition temperatures. doi:10.1103/PhysRevB.64.094518.

S7. K. Kawabata and M. Ido. *Evidence for extrinsic superconductivity in $NbSe_3$*. Solid State Communications 44, 1539–1542 (1982). The abstract specifies the 0.38 K ambient-pressure measurement floor and identifies the weak superconducting response as a boundary effect. doi:10.1016/0038-1098(82)90673-1.

S8. I. Hirosawa, K. Prassides, J. Mizuki, K. Tanigaki, M. Gevaert, A. Lappas and J. K. Cockcroft. *Orientational disorder of C_{60} in Li_2CsC_{60}*. Science 264, 1294–1297 (1994). The abstract explicitly identifies Li_2CsC_{60} as nonsuperconducting and describes its face-centered-cubic structure. The local CIF cites this same structural study. doi:10.1126/science.264.5163.1294, indexed original abstract.

S9. C. de la Cruz et al. *Magnetic order close to superconductivity in the iron-based layered $LaO_{1-x}F_xFeAs$ systems*. Nature 453, 899–902 (2008). The abstract and neutron-diffraction study explicitly distinguish nonsuperconducting, undoped LaOFeAs from superconducting F-substituted material. Table 1 gives the tetragonal parent structure. This supports the undoped composition/phase assignment, rather than identifying the local CIF as the same measured specimen. doi:10.1038/nature07057, open manuscript.

B Model Architecture and Implementation

This appendix follows the modules in Fig. 3: state representation, SuperFlow Backbone, and SAC, followed by their training objectives and two-stage sampling. Each module is described by its inputs, computation, and outputs. Training schedules and numerical settings are in Appendix D.

B.1 State Representation and Doping Rate Readout

The physical crystal is $\mathcal{C} = (\vec{\boldsymbol{L}}, \boldsymbol{S}, \boldsymbol{F})$, with site-composition vectors $\boldsymbol{s}_i$. The spherical representation is $\boldsymbol{U} = \sqrt{\boldsymbol{S}}$, with inverse $\boldsymbol{S} = \boldsymbol{U} \odot \boldsymbol{U}$ elementwise. The native lattice vector $\boldsymbol{l} = \rho(\vec{\boldsymbol{L}})$ contains three lengths and three transformed angles in the DMFlow parameterization. The flow state is $\boldsymbol{z} = (\zeta, \boldsymbol{F}, \boldsymbol{l})$. Decoding applies ρ^{-1} to the lattice parameters. Additional attributes describing positional disorder are inherited from the parent and held fixed during both passes. Thus the final decoded primary state is $\widehat{\mathcal{C}} = (\rho^{-1}(\boldsymbol{l}_{\text{final}}), \boldsymbol{U}_{\text{final}} \odot \boldsymbol{U}_{\text{final}}, \boldsymbol{F}_{\text{final}})$.

For positive-budget eligible sites, the site-angle parameterization is

$$(U_{ih}, U_{id}) = \sqrt{m_i}(\cos\zeta_i, \sin\zeta_i), \qquad \zeta_i \in [0, \pi/2]. \tag{11}$$

The row U_i is a spherical vector, while ζ_i is its host–dopant angle. Fixed parent budgets and fixed elemental fractions complete the reconstruction of $\boldsymbol{U}$ from ζ.

Inactive elements and noneligible sites retain their parent site compositions, so $S_{ih} + S_{id} = m_i$ and each row of $\boldsymbol{S}$ remains normalized. Define $Q = \sum_{i=1}^{N}(S_{P,ih} + S_{P,id})$, $Q_{\mathcal{I}} = \sum_{i\in\mathcal{I}} m_i$, and $D_{\text{fix}} = \sum_{i\notin\mathcal{I}} S_{P,id}$. The doping rate readout is

$$x(\boldsymbol{z}) = \frac{D_{\text{fix}} + \sum_{i\in\mathcal{I}} S_{id}}{Q} = \frac{D_{\text{fix}} + \sum_{i\in\mathcal{I}} U_{id}^2}{Q} = \frac{D_{\text{fix}} + \sum_{i\in\mathcal{I}} m_i \sin^2\zeta_i}{Q}. \tag{12}$$

When all host–dopant sites are eligible, $D_{\text{fix}} = 0$ and $Q = \sum_{i \in \mathcal{I}} m_i$. The goal head uses the scalar parent doping-rate angle $\zeta_0 = \arcsin \sqrt{x_0}$. The site-specific parent angles are denoted by $\zeta_{0,i}$.

B.2 SuperFlow Backbone

The module in Fig. 3(c) maps the current state and local time to joint velocities:

$$(\dot{\zeta}, \dot{\boldsymbol{F}}, \dot{\boldsymbol{l}}) = \mathrm{SuperFlow}_{\phi}(\boldsymbol{z}, t). \tag{13}$$

The fixed substitution condition is suppressed in this module notation.

Message Passing and Node Update. The graph is built from the crystal reconstructed from $\boldsymbol{z}$. Periodic coordinate differences, site-composition information, lattice features, time, and substitution metadata enter the DMFlow-based encoder. Message passing aggregates graph-neighborhood information (Gilmer et al., 2017). The DMFlow-specific updates (Wu et al., 2026) produce intermediate node features $\boldsymbol{H}'$ and final site embeddings $\boldsymbol{h}_i = \mathrm{Enc}_{\phi}(\boldsymbol{z}, t, c)_i$. For SAC, the frozen encoder is evaluated on $\boldsymbol{z}_1$ at local time zero, giving $\boldsymbol{H}_1 = \mathrm{Enc}_{\phi}(\boldsymbol{z}_1, 0, c)$. Table 4 summarizes the shared encoder configuration.

Table 4: Shared graph encoder architecture.

Setting	Value
Graph layers / hidden width	6 / 512
Time embedding / Fourier frequencies	256 / 128
Graph connectivity	Fully connected
Atom-count embedding capacity	129

Output Heads. Site embeddings supply the composition and coordinate heads. The pooled graph feature $\boldsymbol{h}_{\text{graph}} = \mathrm{Pool}(\{\boldsymbol{h}_i\})$ supplies the lattice head. Pooling is defined below. The raw outputs are

$$\bar{\boldsymbol{v}}_i^{\boldsymbol{U}} = f_{\phi,\boldsymbol{U}}(\boldsymbol{h}_i), \qquad \bar{\boldsymbol{v}}_i^{\boldsymbol{F}} = f_{\phi,\boldsymbol{F}}(\boldsymbol{h}_i), \qquad \bar{\boldsymbol{v}}^{\boldsymbol{l}} = f_{\phi,\boldsymbol{l}}(\boldsymbol{h}_{\text{graph}}). \tag{14}$$

Each $f_{\phi,\cdot}$ is a learned output map. The composition and coordinate rows form the matrices $\bar{\boldsymbol{V}}^{\boldsymbol{U}}$ and $\bar{\boldsymbol{V}}^{\boldsymbol{F}}$. Conversion to native velocities, tangent projection, and angular readout produce $\boldsymbol{v}_{\phi} = (\dot{\zeta}_{\phi}, \boldsymbol{V}_{\phi}^{\boldsymbol{F}}, \boldsymbol{v}_{\phi}^{\boldsymbol{l}})$. We express the resulting field in the native coordinates of the flow state:

$$\boldsymbol{v}_{\phi}(\boldsymbol{z}, t, c) = \mathcal{T}_{\boldsymbol{z}}\left(\bar{\boldsymbol{V}}^{\boldsymbol{U}}, \bar{\boldsymbol{V}}^{\boldsymbol{F}}, \bar{\boldsymbol{v}}^{\boldsymbol{l}}\right), \tag{15}$$

where $\mathcal{T}_{\boldsymbol{z}}$ denotes these fixed conversion, projection, and angular readout operations. Writing $V_{\phi,ie}^{\boldsymbol{U}}$ for entry (i, e) of the converted spherical composition field $\boldsymbol{V}_{\phi}^{\boldsymbol{U}}$, the internal angular velocity is

$$\dot{\zeta}_{\phi,i} = \frac{U_{ih} V_{\phi,id}^{\boldsymbol{U}} - U_{id} V_{\phi,ih}^{\boldsymbol{U}}}{m_i}, \qquad i \in \mathcal{I}. \tag{16}$$

Only positive-budget sites enter this expression. The corresponding constrained spherical velocity is $\dot{U}_{ih} = -\sqrt{m_i} \sin \zeta_i \, \dot{\zeta}_{\phi,i}$ and $\dot{U}_{id} = \sqrt{m_i} \cos \zeta_i \, \dot{\zeta}_{\phi,i}$. Thus the composition head supplies angular motion while the coordinate and lattice heads remain active. We suppress the fixed condition c in subsequent field evaluations.

B.3 Structure-Aware Correction (SAC)

The module in Fig. 3(d) first combines its two feature branches:

$$(a_m, a_d, a) = \mathrm{Heads}_{\psi}(\mathrm{Pool}(\boldsymbol{H}_1), \mathrm{FeatureEncoder}(\boldsymbol{z}_0, \boldsymbol{z}_1)). \tag{17}$$

Here Heads_{ψ} comprises the MLP Layers and the Move, Direction, and Magnitude Heads shown in the figure. Their outputs feed the Guide Velocity Predictor. The following paragraphs follow this computation order.

Pooling. The input is the proposal node-feature matrix $\boldsymbol{H}_1$, computed by the frozen backbone at local time zero. Pooling concatenates the mean and sum of its site features:

$$\mathrm{Pool}(\boldsymbol{H}_1) = \left[\frac{1}{N} \sum_{i=1}^{N} \boldsymbol{h}_{1,i} \,\middle\|\, \sum_{i=1}^{N} \boldsymbol{h}_{1,i} \right]. \tag{18}$$

Here $\|$ denotes feature concatenation. Mean pooling summarizes average site features, while sum pooling retains their accumulated contribution. The output is a 1024-dimensional graph feature. The same pooling operation produces $\boldsymbol{h}_{\mathrm{graph}}(\boldsymbol{z}, t, c)$ inside the backbone.

Feature Encoder. The inputs are $\boldsymbol{z}_0$ and $\boldsymbol{z}_1$, together with the fixed substitution metadata and cached initial angular velocities of the frozen backbone. The module computes the auxiliary vector $\boldsymbol{s}_\omega$ listed in Table 5 and returns its standardized values:

$$\mathrm{FeatureEncoder}(\boldsymbol{z}_0, \boldsymbol{z}_1) = \mathrm{Std}_{\boldsymbol{s}_\omega}(\boldsymbol{s}_\omega). \tag{19}$$

Feature extraction and standardization are deterministic. For either feature block, $\mathrm{Std}(q) = (q - \mu_{\mathrm{TRAIN}})/\max(\sigma_{\mathrm{TRAIN}}, 0.01)$ is componentwise, with fixed TRAIN statistics reused at inference. Initial angular-velocity statistics are computed directly from the frozen field at the parent.

Let $x_1 = x(\boldsymbol{z}_1)$, $\dot{\zeta}_{0,i} = \dot{\zeta}_{\phi,i}(\boldsymbol{z}_0, 0)$, and $\langle f \rangle_m = \sum_{i \in \mathcal{I}} m_i f_i / \sum_{i \in \mathcal{I}} m_i$. The velocity statistics use the original parent at initial time, whereas the graph embedding uses the first generated state. Both input blocks are computed once between the two passes.

Table 5: The 13 scalar inputs to the second-stage goal head. N is the retained site count and $\epsilon_b = 10^{-8}$.

Index	Scalar	Meaning
1	x_0	Parent doping rate
2	x_1	First-round generated doping rate
3	$x_1 - x_0$	Signed first-round doping rate change
4	$\lvert x_1 - x_0 \rvert$	Absolute first-round doping rate change
5	ζ_0	Parent doping-rate angle
6	$\arcsin \sqrt{x_1}$	First-round doping-rate angle
7	$\mu_{\dot{\zeta}} = \langle \zeta_0 \rangle_m$	Weighted mean initial angular velocity
8	$\langle \lvert \dot{\zeta}_0 \rvert \rangle_m$	Weighted mean absolute initial angular velocity
9	$\sqrt{\langle (\dot{\zeta}_0 - \mu_{\dot{\zeta}})^2 \rangle_m}$	Weighted initial angular-velocity standard deviation
10	$\sum_{i \in \mathcal{I}} m_i / \lvert \mathcal{I} \rvert$	Mean eligible-site host–dopant site budget
11	$N/64$	Normalized site count
12	$\mathbb{I}[x_0 < \epsilon_b]$	Lower doping-rate boundary flag
13	$\mathbb{I}[x_0 > 1 - \epsilon_b]$	Upper doping-rate boundary flag

MLP Layers. The inputs are the pooled graph feature and the Feature Encoder output. The graph block is standardized, concatenated with the already-standardized auxiliary block, and passed to the learned MLP:

$$\begin{aligned} \boldsymbol{y} &= \left[\mathrm{Std}_{\mathrm{graph}}(\mathrm{Pool}(\boldsymbol{H}_1)) \,\|\, \mathrm{FeatureEncoder}(\boldsymbol{z}_0, \boldsymbol{z}_1) \right], \\ \boldsymbol{h}_{\mathrm{goal}} &= \mathrm{MLP}_\psi(\boldsymbol{y}). \end{aligned} \tag{20}$$

Each input block is standardized once. The graph and auxiliary blocks have dimensions 1024 and 13, giving a 1037-dimensional input. The shared output $\boldsymbol{h}_{\mathrm{goal}} \in \mathbb{R}^{128}$ feeds all three readouts. Table 6 specifies the MLP and output architecture.

Table 6: Goal-head architecture.

Setting	Value
Head widths	$1037 \to 256 \to 128 \to 3$ outputs
Hidden operations	LayerNorm, SiLU, dropout 0.1, then SiLU
Magnitude transform	$(\pi/2)$ sigmoid

Move Head, Direction Head, and Magnitude Head. The three readouts are

$$a_m = \boldsymbol{w}_m^\top \boldsymbol{h}_{\text{goal}} + b_m, \qquad a_d = \boldsymbol{w}_d^\top \boldsymbol{h}_{\text{goal}} + b_d, \qquad a = \frac{\pi}{2}\sigma(\boldsymbol{w}_a^\top \boldsymbol{h}_{\text{goal}} + b_a). \tag{21}$$

They give movement and direction logits and an angular magnitude in $[0, \pi/2]$. The trainable parameters ψ comprise the head trunk and readouts. The encoder parameters ϕ remain fixed.

Guide Velocity Predictor. The module takes the readouts (a_m, a_d, a), the parent angle and budgets, and the first proposal's site compositions. Its fixed-goal calculation is

$$g = \mathbb{I}[\sigma(a_m) \geq \eta_m], \qquad x^\dagger = \sin^2\Big[\text{clip}_{[0,\pi/2]}\big(\zeta_0 + g\,\text{sgn}_+(a_d)a\big)\Big]. \tag{22}$$

Here $\text{sgn}_+(u)$ is $+1$ for $u \geq 0$ and -1 otherwise.

The bridge targets are constructed from the first generated site compositions, using the total budget Q, eligible budget $Q_{\mathcal{I}}$, and fixed dopant amount D_{fix} defined above. For $i \in \mathcal{I}$, let $d_{1,i} = (\boldsymbol{S}_1)_{id}$, $D_1 = \sum_{i\in\mathcal{I}} d_{1,i}$, and

$$D_{\text{goal}} = \text{clip}_{[0,Q_{\mathcal{I}}]}(Qx^\dagger - D_{\text{fix}}), \qquad \delta_D = D_{\text{goal}} - D_1. \tag{23}$$

The implementation interpolates toward the pure-dopant or pure-host endpoint:

$$d_i^\dagger = \begin{cases} d_{1,i} + \dfrac{\delta_D}{Q_{\mathcal{I}} - D_1}(m_i - d_{1,i}), & \delta_D > 0, \\ d_{1,i} + \dfrac{\delta_D}{D_1} d_{1,i}, & \delta_D < 0, \\ d_{1,i}, & \delta_D = 0, \end{cases} \qquad \zeta_i^\dagger = \arcsin\sqrt{d_i^\dagger/m_i}. \tag{24}$$

Here $m_i = S_{P,ih} + S_{P,id}$ is the fixed host–dopant site total at site i, so $d_i^\dagger/m_i$ is the normalized target doping rate. Each denominator is positive in its applicable branch. This preserves $0 \leq d_i^\dagger \leq m_i$ and $\sum_{i\in\mathcal{I}} d_i^\dagger = D_{\text{goal}}$ without forcing equal site ratios. Elemental fractions at noneligible sites remain fixed. The realized goal doping rate is $(D_{\text{fix}} + D_{\text{goal}})/Q$. It equals $x^\dagger$ whenever the head goal is feasible. The implementation uses small denominator and square-root floors.

The target angles $\zeta^\dagger$ are stored once. Given the current second-pass angles and time, the module outputs the correction velocity

$$\dot{\zeta}_{\text{corr}}(\tau) = \frac{\zeta^\dagger - \zeta(\tau)}{\max(1-\tau, \Delta\tau)}, \qquad \Delta\tau = 1/K. \tag{25}$$

This velocity is recomputed at each integration step. No coordinate or lattice endpoint is predicted by SAC.

B.4 Training objectives

The two training objectives supervise the joint-flow backbone and the SAC MLP/readouts, respectively. Optimization schedules and hyperparameters remain in Appendix D.

Joint-flow objective $\mathcal{L}_{\text{flow}}$. The parent-to-target angle path is

$$\zeta_i(t) = \zeta_{0,i} + q(t, x_0)(\zeta_i^* - \zeta_{0,i}), \qquad u_i^\zeta(t) = \partial_t \zeta_i(t). \tag{26}$$

To adjust path progress near the doping rate boundaries while preserving both endpoints, we use the angular progress function

$$q(t, x_0) = t + b(x_0)t(1-t), \qquad b(x_0) = \mathrm{clip}_{[0,1]} \frac{\delta - \min(x_0, 1-x_0)}{\delta}. \tag{27}$$

For parents away from the doping rate boundaries, $b(x_0) = 0$ and $q(t, x_0) = t$. Near a boundary, the added term advances more of the displacement early in the path while retaining $q(0, x_0) = 0$ and $q(1, x_0) = 1$. The boundary width δ is specified in Appendix D.

The angular target is the derivative of Eq. (26), $u_i^\zeta(t) = \partial_t q(t, x_0)(\zeta_i^* - \zeta_{0,i})$. For the aligned coordinate and lattice endpoints, the DMFlow product-manifold path gives

$$\begin{aligned} \boldsymbol{F}_t &= \mathrm{Exp}_{\boldsymbol{F}_P}(t\xi_{\boldsymbol{F}}), & \xi_{\boldsymbol{F}} &= \mathrm{Log}_{\boldsymbol{F}_P}(\boldsymbol{F}^*), & \boldsymbol{U}_t^{\boldsymbol{F}} &= \partial_t \boldsymbol{F}_t, \\ \boldsymbol{l}_t &= (1-t)\boldsymbol{l}_P + t\boldsymbol{l}^*, & & & \boldsymbol{u}_t^{\boldsymbol{l}} &= \boldsymbol{l}^* - \boldsymbol{l}_P. \end{aligned} \tag{28}$$

The coordinate maps interpolate periodic displacements on the fractional-coordinate torus. Lattice interpolation is in native parameter space. The angles in Eq. (26) reconstruct $\boldsymbol{U}_t$ for graph encoding, while the integrated state is $\boldsymbol{z}_t = (\zeta_t, \boldsymbol{F}_t, \boldsymbol{l}_t)$. Unlike the random-source structural pretraining path, this fine-tuning path starts from the parent.

The two loss components, averaged over training pairs and $t \sim \mathcal{U}(0,1)$, are

$$\mathcal{L}_\zeta = \mathbb{E}\left[\frac{\sum_{i\in\mathcal{I}} m_i(\dot{\zeta}_{\phi,i}(\boldsymbol{z}_t, t) - u_i^\zeta(t))^2}{\sum_{i\in\mathcal{I}} m_i}\right], \tag{29}$$

$$\mathcal{L}_G = \mathbb{E}_{\text{aligned}}\left[\lambda_F \|\boldsymbol{V}_\phi^{\boldsymbol{F}}(\boldsymbol{z}_t, t) - \boldsymbol{U}_t^{\boldsymbol{F}}\|_{\text{site}}^2 + \lambda_l \|\boldsymbol{v}_\phi^{\boldsymbol{l}}(\boldsymbol{z}_t, t) - \boldsymbol{u}_t^{\boldsymbol{l}}\|_{\text{dim}}^2\right]. \tag{30}$$

The two unweighted geometry terms define $\mathcal{L}_F$ and $\mathcal{L}_l$ in Eq. (2), so $\mathcal{L}_{\text{flow}} = \mathcal{L}_\zeta + \mathcal{L}_G$. Here $\|\boldsymbol{B}\|_{\text{site}}^2 = \|\boldsymbol{B}\|_{\text{F}}^2/(3N)$ and $\|\boldsymbol{b}\|_{\text{dim}}^2 = \|\boldsymbol{b}\|_2^2/6$. Geometry supervision is applied to pairs with a valid composition-preserving site alignment. The coordinate norm averages over sites and the three fractional-coordinate components. The lattice norm averages over its parameter dimensions. Geometry output heads receive their direct gradients. On shared parameters, we project away a geometry-gradient component opposing the doping-rate gradient and cap its norm relative to that gradient. This prioritizes the doping rate objective while learning joint structural motion.

Nonflat trajectories provide both improvement pairs and optimum-to-itself pairs. The latter have zero endpoint displacement and train the field at observed optima. A late training phase also uses cached generated TRAIN states as recovery starts toward the same targets.

SAC objective $\mathcal{L}_{\text{corr}}$. The labels refer to the parent and its nearest observed optimum. Relative to the parent angle, let

$$\Delta^* = \arcsin\sqrt{x^*} - \zeta_0, \quad m^* = \mathbb{I}[|\Delta^*| > \epsilon_{\text{move}}], \quad d^* = \mathbb{I}[\Delta^* > 0], \quad a^* = |\Delta^*|. \tag{31}$$

Here ϵ_{move} is the numerical movement tolerance specified in Appendix D, and x^* is the nearest member of X^* to x_0. The first generated state is an input to this decision. The supervised displacement remains relative to the known parent. For head training, define the differentiable goal-angle prediction

$$\zeta_{\text{soft}} = \mathrm{clip}_{[0,\pi/2]}\big(\zeta_0 + \sigma(a_m)\tanh(a_d)a\big). \tag{32}$$

The stage-II head objective is

$$\begin{aligned} \mathcal{L}_{\text{corr}} = \lambda_m \,\mathrm{BCE}_{\kappa_m}(a_m, m^*) + \lambda_d \mathbb{E}_{m^*=1}\,\mathrm{BCE}_{\kappa_d}(a_d, d^*) \\ + \lambda_a \mathbb{E}_{m^*=1}\ell_{\text{H}}(a, a^*) + \lambda_e \ell_{\text{H}}(\zeta_{\text{soft}}, \arcsin\sqrt{x^*}). \end{aligned} \tag{33}$$

Class weights κ_m, κ_d are computed on TRAIN, and ℓ_{H} is smooth-L_1. Optimization updates the head, with no backpropagation through either full rollout.

B.5 Two-stage sampling

Both passes use the same selected backbone and the same parent-defined substitution constraints. The first pass integrates the base field from the parent:

$$\frac{\mathrm{d}\boldsymbol{z}_t}{\mathrm{d}t} = \boldsymbol{v}_\phi(\boldsymbol{z}_t, t, c), \qquad \boldsymbol{z}_{t=0} = \boldsymbol{z}_0. \tag{34}$$

First pass and SAC evaluation. At inference, the first pass integrates the joint field over a complete local time interval:

$$\boldsymbol{z}_1 = \mathcal{R}_\phi(\boldsymbol{z}_0, c, 0 \to 1), \qquad \dot{\zeta}_i = \dot{\zeta}_{\phi,i}, \quad \dot{\boldsymbol{F}} = \boldsymbol{V}_\phi^{\boldsymbol{F}}, \quad \dot{\boldsymbol{l}} = \boldsymbol{v}_\phi^{\boldsymbol{l}}. \tag{35}$$

Site angles are clipped to $[0, \pi/2]$, and the structural state is projected onto its product manifold after each Euler update. After obtaining the complete proposal $\boldsymbol{z}_1$, the frozen backbone computes $\boldsymbol{H}_1$ at local time zero. Pooling, Feature Encoder, MLP Layers, and the three heads are evaluated once. Guide Velocity Predictor then sets the fixed target $\zeta^\dagger$ using the goal and lift above.

Second pass. Reset local time to zero and start from the full $\boldsymbol{z}_1$. At each step the frozen backbone supplies $(\dot{\zeta}_1, \dot{\boldsymbol{F}}_1, \dot{\boldsymbol{l}}_1)$ and Guide Velocity Predictor updates $\dot{\zeta}_{\mathrm{corr}}$. Combine them as

$$\dot{\zeta}_{\mathrm{guide}} = (1-\beta)\dot{\zeta}_1 + \beta\dot{\zeta}_{\mathrm{corr}}. \tag{36}$$

At numerical step k, with $\tau_k = k/K$, the update is

$$\zeta_{k+1} = \mathrm{clip}_{[0,\pi/2]}\Big[\zeta_k + \Delta\tau\big((1-\beta)\dot{\zeta}_1(\boldsymbol{z}_k, \tau_k) + \beta\dot{\zeta}_{\mathrm{corr}}(\tau_k)\big)\Big]. \tag{37}$$

Here the subscript in $\dot{\zeta}_1$ identifies the second pass, while k indexes its numerical steps. The sampler reconstructs $U_{ih} = \sqrt{m_i}\cos\zeta_i$ and $U_{id} = \sqrt{m_i}\sin\zeta_i$ after updating the angles. Elemental fractions outside the substitution axis remain fixed. Coordinates and lattice parameters are updated jointly from the same pre-update state. The second pass starts at the complete $\boldsymbol{z}_1$ with local time zero and ends at $\boldsymbol{z}_{\mathrm{final}}$. The backbone uses the same angular field $\dot{\zeta}_\phi$ in both passes. It is denoted $\dot{\zeta}_1$ within refinement.

C Theoretical analysis and proof

This section analyzes the angular part of the stated sampler in exact arithmetic. The graph field may depend on all components $\boldsymbol{z} = (\zeta, \boldsymbol{F}, \boldsymbol{l})$. No correctness assumption is made about the direction of its base angular velocity. The original parent, editable-site set and positive budgets m_i remain fixed throughout the pass. The goal head is evaluated once between the two passes. Its lifted target vector remains fixed throughout refinement, including when the movement gate selects the parent doping rate.

Here k indexes integration steps within the second pass. Its initial complete state is the first proposal $\boldsymbol{z}_1$. During the second pass, the target angle vector $\zeta^\dagger$ and parent budgets are fixed. On positive-budget eligible sites define

$$\|u\|_m^2 = \frac{\sum_{i\in\mathcal{I}} m_i u_i^2}{Q_{\mathcal{I}}}, \qquad Q_{\mathcal{I}} = \sum_{i\in\mathcal{I}} m_i, \qquad E_k = \left\|\zeta_k - \zeta^\dagger\right\|_m. \tag{38}$$

For $\Delta\tau = 1/K$ and $\tau_k = k/K$, the actual projected Euler form is

$$\zeta_{k+1} = \Pi\left[\zeta_k + \frac{\beta}{K-k}(\zeta^\dagger - \zeta_k) + \Delta\tau(1-\beta)\dot{\zeta}_\phi(\boldsymbol{z}_k, \tau_k)\right], \tag{39}$$

where Π clips each active angle to $[0, \pi/2]$.

C.1 One-step recurrence and finite-step bound

At the Euler evaluation times $\tau_k = k/K$, $k = 0, \ldots, K-1$, the remaining time satisfies $1 - \tau_k = (K-k)/K \geq \Delta\tau$. Consequently,

$$\frac{\Delta\tau\beta}{\max(1-\tau_k, \Delta\tau)} = \frac{\beta}{K-k}, \tag{40}$$

which yields Eq. (39) exactly. The target lies in the clipping box, and scalar clipping satisfies $|\Pi_i(y_i) - \zeta_i^\dagger| \leq |y_i - \zeta_i^\dagger|$. Multiplication by positive weights, summation and a square root establish the same inequality in $\|\cdot\|_m$.

Proof of Prop. 5.1. Write $a_k = 1 - \beta/(K-k)$ and $v_k = \left\| \dot{\zeta}_\phi(\boldsymbol{z}_k, \tau_k) \right\|_m$. Since $0 < \beta < 1$, every a_k is nonnegative. Subtracting the target before projection, then applying nonexpansiveness and the triangle inequality, gives

$$E_{k+1} \le a_k E_k + \Delta\tau(1-\beta)v_k. \tag{41}$$

Define $A_0 = 1$ and $A_m = \prod_{r=1}^{m}(1-\beta/r)$, consistent with Eq. (10). A perturbation inserted at step k is multiplied by A_{K-1-k} through the remaining updates. Unrolling Eq. (41) therefore gives the trajectory-specific bound

$$E_K \le A_K E_0 + \frac{1-\beta}{K} \sum_{k=0}^{K-1} A_{K-1-k} v_k. \tag{42}$$

To replace step velocities by a common bound, observe

$$(m+1)A_{m+1} - mA_m = (1-\beta)A_m. \tag{43}$$

Summing over $m = 0, \ldots, K-1$ yields $KA_K = (1-\beta)\sum_{m=0}^{K-1} A_m$. With $v_k \le M$, the second term of Eq. (42) is at most $A_K M$. Hence

$$E_K \le A_K(\beta)E_0 + A_K(\beta)M = A_K(\beta)(E_0 + M), \qquad A_K(\beta) = \prod_{r=1}^{K}\left(1 - \frac{\beta}{r}\right), \tag{44}$$

which is Eq. (10). □

For fixed $0 < \beta < 1$, $\log A_K(\beta) = \sum_{r=1}^{K} \log(1-\beta/r) \le -\beta \sum_{r=1}^{K} 1/r$, so $A_K(\beta) \to 0$. Thus, for a fixed initial state and goal and a base-velocity bound M uniform in K, the endpoint goal-distance bound tends to zero as the number of integration steps increases.

D Training settings and hyperparameters

D.1 Structural pretraining

We follow DMFlow (Wu et al., 2026) for de novo generation (DNG) pretraining of the joint composition, coordinate, and lattice flow. The structural corpus is described in Appendix A. Encoder dimensions are given in Table 4, and the pretraining configuration is listed in Table 7.

Table 7: Structural pretraining configuration. Batch size is per process and is additionally limited by graph-size budgets.

Setting	Value
Epochs	2,000
Batch size / batch site budget / batch edge budget	8 / 384 / 12,000
Optimizer / weight decay	AdamW / 0
Initial / minimum learning rate	3×10^{-4} / 3×10^{-6}
Schedule / gradient clipping	Cosine / 0.5
$\boldsymbol{U}$ / $\boldsymbol{F}$ / $\boldsymbol{l}$ loss coefficients	2,000 / 400 / 1
Seed / validation interval	42 / 5 epochs

D.2 First-stage joint flow

The first-stage joint-flow model is initialized from the dataset-specific doping-rate flow model and then trained to update composition and geometry together. During warmup, only the coordinate and lattice output heads are trained. Subsequent epochs also update the final graph block, composition head, and axis-conditioning parameters. Earlier graph blocks remain frozen. Table 8 specifies the schedule and coefficients for the paths and losses in Appendix B. Numerically identical geometry targets are treated as identity pairs.

The nonflat TRAIN pools contain 2,069 lower and 554 peak cases on ICSD, and 964 lower and 241 peak cases on MP. Flat TRAIN trajectories contribute 310 and 68 additional cases to the full inventories, respectively. They are excluded from this joint-flow sampling stream. MP's selected epoch precedes recovery-state training. Sampling is balanced across trajectories within each stream.

Table 8: First-stage joint fine-tuning settings. Selected epochs are determined by complete VAL rollout.

Setting	Value
Epoch budget / geometry-output warmup	20 / 5
Batch size	64
Geometry-output / shared learning rate	10^{-4} / 10^{-6}
Optimizer / weight decay	AdamW / 0
Gradient clipping	1.0 per parameter group
Nominal sampled examples per epoch, ICSD / MP	2,933 / 1,273
Ordinary pair sampling	Nonflat lower / peak, 50:50
Recovery-state collection	After epoch 15, original TRAIN only
Final-five-epoch sampling	25% recovery, remaining lower / peak, 50:50
Selected epoch, ICSD / MP	20 / 5
Evaluation and inference	200 projected Euler steps
Boundary width δ	0.1
Coordinate / lattice loss weights	400 / 1
Shared geometry-gradient norm cap	10% of doping-rate-gradient norm
Seed	42

D.3 Goal-head training and second-pass settings

Table 9 summarizes goal-head optimization and second-pass sampling. The head uses the full TRAIN inventories of 2,933 ICSD and 1,273 MP cases. In Eq. (33), $(\lambda_m, \lambda_d, \lambda_a, \lambda_e) = (1, 1, 2, 1)$. Movement and direction BCE positive weights are the TRAIN negative/positive count ratios. Direction and magnitude terms are evaluated only on moving cases. Smooth-L_1 uses transition parameter 0.05. The training movement label uses the angular tolerance $\epsilon_{\text{move}} = 10^{-8}$ radians in Eq. (33). Evaluation subsets use the doping-rate distance thresholds in § 6.1.

Table 9: Goal head and second-round settings.

Setting	Value
Epoch budget / batch size	20 / 256
Optimizer / learning rate / weight decay	AdamW / 10^{-3} / 10^{-4}
Gradient clipping	5.0
Checkpoint evaluations	Epochs 5, 10, 20, complete 200+200 VAL rollout
Selected epoch, ICSD / MP	10 / 20
Selection priority	Moving parents xMAE, all-parent xMAE, earlier epoch
Direction / movement rule	Hard sign / threshold 0.15
Bridge weight β, ICSD / MP	0.375 / 0.300
Second-round steps / time interval	200 / $[0, 1]$
Seed	42

After freezing the selected first-stage model and SAC goal head, we select one shared positive movement threshold on VAL from $\{0.001, 0.005, 0.01, 0.02, 0.05, 0.10, 0.15, 0.20, 0.25\}$. Relative to the zero-threshold reference, whose movement gate is always open, an eligible candidate allows at most 0.5% higher all-parent xMAE, 1% higher moving or large-move xMAE, and one percentage point lower moving or large-move direction accuracy on each dataset. Among eligible candidates, selection minimizes the mean relative all-parent xMAE across datasets, then mean relative moving xMAE, then threshold. This selects 0.15, which is used for all reported TEST metrics. Low-probability cases retain a goal at the parent doping rate but still undergo the full second-round joint integration.

E BASELINE AND ABLATION DETAILS

E.1 UNIFORM RANDOM

Doping rates are uniformly distributed on $[0, 1]$. The reported metrics are exact case-weighted expectations. Tied optima use the same nearest-optimum error and any-optimum direction definitions as the learned methods.

E.2 PARENT-ONLY DIRECT-x

This baseline uses the DMFlow-based crystal graph encoder with a 64-bin doping-rate head. The head accepts two graph-feature inputs, both supplied with the same parent embedding, together with the host–dopant condition. Thus doping rate is predicted directly from the parent without a generated proposal. The graph model, head and conditioning parameters are optimized with AdamW at 3×10^{-4}, default weight decay 0.01, cosine decay to 3×10^{-6} and gradient clipping 0.5. The budget is ten epochs and 1,600 sampled examples per epoch, with batch size one on each of two processes. Selection uses moving VAL doping-rate error. ICSD and MP select epochs six and three. The objective is equally weighted target-bin cross-entropy and L_1 error of the bin expectation.

The head predicts a categorical distribution over $B = 64$ equally spaced doping rate support points $c_k = k/(B-1)$, $k = 0, \ldots, B-1$. Given logits b_k, the continuous prediction and nearest-support training label are

$$p_k = \mathrm{softmax}(\boldsymbol{b})_k, \qquad \widehat{x} = \sum_{k=0}^{B-1} p_k c_k, \qquad k^* = \mathrm{round}\big((B-1)x^*\big). \tag{45}$$

For a target doping rate x^*, the per-example objective is

$$\mathcal{L}_{\mathrm{direct}} = -\log p_{k^*} + |\widehat{x} - x^*|. \tag{46}$$

E.3 SUPERCON-DIFFUSION-AXIS

We adapt the formula generator to a fixed parent and substitution axis by conditioning on a high-T_c category and holding non-axis composition to the parent during sampling. The generated host and dopant coefficients determine the predicted doping rate.

The original three-channel representation describes the integer, first decimal and second decimal digits of each stoichiometric coefficient, with ten one-hot positions per digit. We use an explicit vocabulary of 86 elements. When a formula coefficient exceeds the encoding range, all coefficients are divided by a common factor, preserving their ratios. Coefficients are multiplied by 100 and rounded to integers before digit extraction, giving a consistent two-decimal representation. The resulting tensor has three channels, 86 element rows and ten digit columns.

The original context-conditioned U-Net is retained with feature width 256. We train its standard DDPM noise-prediction objective (Ho et al., 2020) using Adam, batch size 128 and initial learning rate 10^{-3}. A training-noise-loss plateau scheduler reduces this rate by a factor of 0.1 after ten stagnant epochs. The diffusion schedule has 400 steps with linearly spaced noise variances from 10^{-4} to 0.02, and context dropout is 0.1. Temperature conditioning uses three classes covering the full benchmark: $T_c \leq 40$ K, $40 < T_c < 77$ K, and $T_c \geq 77$ K. Zero labels belong to the first class.

During reverse diffusion, the nonaxis entries are replaced at each step by the parent matrix at the matching noise level. At the final step, hard digit decoding produces host/dopant coefficients. If both decode to zero, a soft digit-expectation readout supplies their ratio. Reported predictions use the highest class, $T_c \geq 77$ K, with no additional classifier-free amplification or property-based reranking. For the reported ICSD comparison, full VAL generated doping-rate error selects between epochs 100 and 300. MP selects among epochs 100, 300 and 1,000. The selected epochs are 300 for ICSD and 100 for MP. All training compositions come from the same benchmark TRAIN split.

E.4 Tc-gradient-guided SuperFlow

Motivated by gradient-based guidance (Dhariwal & Nichol, 2021; Feng et al., 2025), this baseline uses the following task-specific angular update:

$$\dot{\zeta}_{\mathrm{guided},k} = \dot{\zeta}_\phi(\boldsymbol{z}_k, t_k) + \alpha\ \mathrm{Norm}\Big(\nabla_\zeta \widehat{T}_c(\boldsymbol{z}_k)\Big)\,. \tag{47}$$

Here $\widehat{T}_c$ denotes the standardized temperature prediction, α controls guidance strength, and Norm denotes the scaling and clipping calibrated on TRAIN examples below.

The baseline starts from the selected first-stage joint SuperFlow and adds a freshly initialized $1024 \to 256 \to 1$ SiLU Tc head. The participating backbone, axis-conditioning parameters, and composition, coordinate, and lattice output heads are fine-tuned jointly with this temperature head for ten epochs using AdamW, learning rate 10^{-5}, zero weight decay, batch size 64 and seed 42. The flow-matching stream retains the lower/optimum-parent balance and geometric-gradient protection described in Appendix D. Tc supervision uses each observed TRAIN structure's own label, including zero-coded labels. Records with the same preprocessed structure-file path and temperature label are deduplicated. This gives 2,708 ICSD and 1,179 MP regression records. No endpoint Tc label is assigned to an interpolated or generated state.

Tc labels are standardized by the TRAIN mean and standard deviation, with the latter floored at 1 K. The regression loss is smooth-L_1 with unit transition width. A fixed set of 32 training calibration examples sets the regression-loss weight so that the initial shared-backbone Tc-gradient norm is one quarter of the doping-rate-loss gradient norm, excluding the output heads. This weight is then fixed. The resulting coefficients are 0.870433 for ICSD and 0.0346583 for MP. Shared/Tc-head and geometry-output gradients are clipped separately to norm one. The property encoder uses local time zero during both Tc training and guidance, and reads the current generated $\boldsymbol{S}/\boldsymbol{F}/\boldsymbol{l}$ state at every sampling step.

For each saved checkpoint, the same calibration examples set s_{grad} to the 90th percentile of the Euclidean norm of the active-site angular gradient of standardized Tc. With $\boldsymbol{g}_{T_c} = \nabla_\zeta \widehat{T}_c$, guidance uses

$$\mathrm{Norm}(\boldsymbol{g}_{T_c}) = \frac{\boldsymbol{g}_{T_c}/s_{\mathrm{grad}}}{\max(1, \|\boldsymbol{g}_{T_c}/s_{\mathrm{grad}}\|_2)}. \tag{48}$$

Only eligible angles receive this gradient. The ordinary budget-preserving projection and $\boldsymbol{S}/\boldsymbol{F}/\boldsymbol{l}$ integration are retained. The single-pass sampler has 200 steps. VAL selects among trained epochs $\{2, 5, 10\}$ and guidance strengths $\{0.03, 0.1, 0.3\}$ by moving xMAE, then all-parent xMAE, smaller strength and earlier epoch. This selects epoch 10 on ICSD and epoch 2 on MP, both with strength 0.03. All reported results for this baseline use the full-label training stream.

E.5 One-round SuperFlow

One-round SuperFlow uses the selected first-stage flow model and a single 200-step rollout. It removes the generated-state goal head and second pass, while retaining the same first-stage checkpoint and evaluation cases.

E.6 Train-from-scratch control

This ablation trains the first-stage joint-flow model from random initialization for 60 epochs with batch size 16. It uses a 20-epoch output/condition warmup at learning rate 3×10^{-4}, followed by unfreezing at learning rate 10^{-5}. The selected epochs are 40 for ICSD and 30 for MP. It is evaluated by a single 200-step rollout and does not attach the second-round goal head. This configuration is reported separately from One-round SuperFlow.

F SUPPLEMENTARY RESULTS

F.1 DOPING-RATE ERROR

Fig. 6 complements the averages in Table 1 with the corresponding TEST error distributions for all seven methods. For each parent, the plotted quantity is

$$e_x = \min_{x^* \in X^*} |\widehat{x} - x^*|. \tag{49}$$

It is an absolute difference in the doping rate, rather than a relative percentage error. The evaluation subsets are defined in § 6.1, with query counts in Table 2. All methods use the same queries within a panel.

The white circle marks the median, the thick black segment spans the 25th–75th percentiles, and the black diamond marks the mean. For learned methods, violin shapes use Gaussian kernel density estimates restricted to the observed value range. Each violin is independently normalized to the same maximum width. Width therefore shows relative density within a distribution, not cohort size or a directly comparable probability across methods. Uniform random uses the exact case-weighted distribution induced by $\widehat{x} \sim \mathrm{Uniform}[0, 1]$, rather than a density of per-case expected errors. Summary markers use the underlying distributions, not the smoothed contours.

Many predictions have low error, but the distributions retain high-error tails. Consistent with Table 1, SuperFlow lowers moving- and large-move xMAE relative to One-round SuperFlow on both datasets. On ICSD, its all-parent xMAE is higher than those of One-round SuperFlow and Tc-gradient-guided SuperFlow, whereas on MP it is lower.

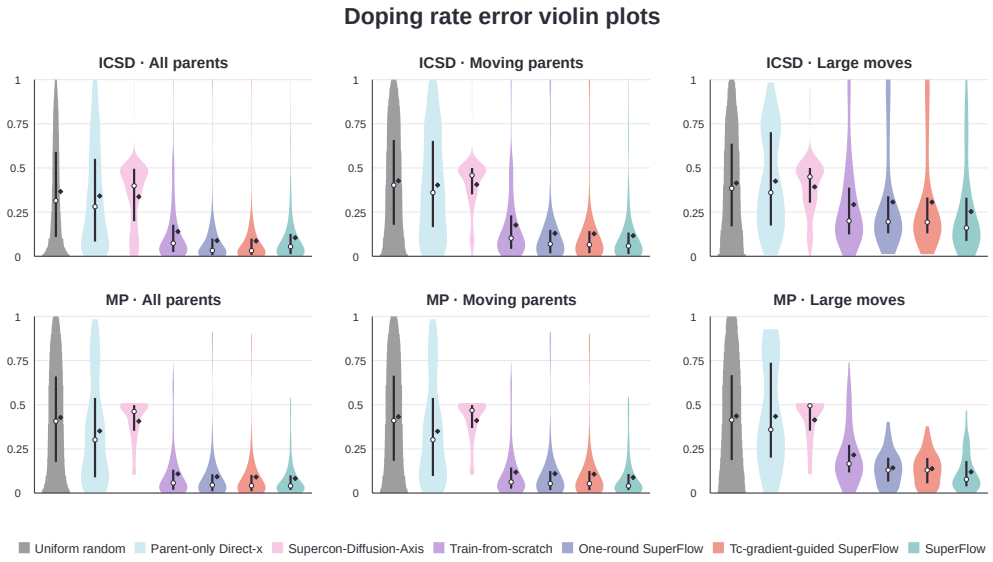

Figure 6: TEST doping-rate-error distributions. Rows show ICSD and MP. Columns show all parents, moving parents and large moves. Lower errors are better. Colors identify the seven methods in Table 1. Marker and density conventions are defined in Appendix F.

F.2 SIGNED MOVEMENT TOWARD AN OBSERVED OPTIMUM

Fig. 7 displays a continuous diagnostic of the direction metric. For each moving parent, we plot

$$s_x = \max_{x^* \in X^*} \left[\operatorname{sgn}(x^* - x_0)(\widehat{x} - x_0) \right]. \tag{50}$$

Positive values denote a movement in the direction of at least one observed optimum, negative values denote a movement away from every optimum, and zero denotes no doping rate change. When tied optima lie on both sides of the parent, this definition gives $s_x = |\widehat{x} - x_0|$, matching the "any optimum" direction rule in § 6. Already-optimal parents are excluded because their desired movement direction is undefined.

The fraction of cases with $s_x > 0$ gives the reported direction accuracy for each learned method. Uniform random uses the corresponding exact probability. The diamond is the mean signed movement, not direction accuracy. A positive movement can overshoot an optimum, so a larger positive value does not necessarily indicate a smaller doping-rate error. The two figures should therefore be read together. Kernel smoothing near zero is only a display device. Accuracy is computed from the underlying predictions. SuperFlow has the highest direction accuracy in the moving and large-move cohorts on both datasets, but negative movements remain.



Figure 7: TEST signed-movement distributions. Rows show ICSD and MP. Columns show moving parents and large moves. The horizontal zero line separates movements toward an observed optimum from movements away from all optima. Colors, summary markers and density normalization follow Fig. 6.

Scope of the results. The benchmark targets the highest-recorded-T_c compositions within observed substitution series. The tracking bound characterizes refinement toward the predicted goal, while doping accuracy and local structural changes are evaluated against the dataset targets. The second pass improves moving-parent and large-move accuracy on both datasets (Table 1). The generated crystals provide candidates for further experimental validation.