GX-TGVAE: Constraint-Masked Graph Generation for Controllable Molecular Exploration

Anonymous authors
Paper under double-blind review

Abstract

Distributional fidelity in molecular generation alone establishes neither molecule-specific information retention nor controllable out-of-distribution (OOD) exploration. GX-TGVAE connects these capabilities through continuous latent conditioning and constraint-masked atom- and bond-level graph construction. It learns deterministic canonical reconstruction and steers exploration using locally estimated tangent/normal geometry and property gradients. On MOSES, it achieves the highest exact reconstruction among the compared latent baselines and broader structural reach, although distributional fidelity and reconstruction rankings differ across architecture variants. On ZINC, property-guided normal exploration exceeds a random-normal control in Novel Top 5% docking score at the two deeper tested exploration depths. Against the OOD generation baseline MOOD, GX-TGVAE shows high conditional utility within training-similarity bands, but its base sampler lacks sampling mass for fixed-attempt discovery and deep-OOD reach. This distinguishes exploration breadth from reached-region utility. Harvest reallocates sampling budget to selected neighborhoods, improving this deficit and exceeding MOOD's hit yield under a matched evaluation-time budget, while MOOD retains higher upper-tail docking scores. Together, these results identify faithful representation, directional control, and sampling-mass allocation as distinct but complementary design axes for molecular OOD generation.

1 Introduction

Molecular generation requires both faithful representations of individual molecules and ways to manipulate them to explore useful, unfamiliar regions. Distribution-level metrics such as Fréchet ChemNet Distance (FCD) quantify population-level similarity to a reference distribution, but do not reveal latent-to-molecule correspondence, structural breadth, or finite-budget out-of-distribution (OOD) reach. We ask whether **continuous molecular representations can be connected to discrete graph generation with structural constraints and used for faithful reconstruction and controllable OOD exploration**.

Our central methodological contribution is **GX-TGVAE, which connects molecule-specific continuous latent conditioning to constraint-masked primitive graph construction**. GX-TGVAE retains GraphXForm's constraint-aware primitive action formulation (Pirnay et al., 2025) but redesigns the generator around a molecule-level variational autoencoder (VAE), with a graph encoder and a latent-conditioned autoregressive decoder. The latent code is projected to K tokens; the global construction token a_0 cross-attends to them and distributes the conditioned representation through subsequent graph self-attention. Training reconstructs a deterministic canonical construction sequence for each graph; the same latent representation supports geometry-guided exploration. We also show how canonical reconstruction bounds unordered graph likelihood.

We evaluate three questions. First, does the latent retain molecule-specific structure while supporting broad generation? GX-TGVAE leads the compared latent baselines in exact reconstruction on common 10k inputs and all compared baselines on all 7 structural-reach metrics across all 3 generation seeds. Second, can local latent geometry control OOD exploration? Direction controls and anchor-relative gains support this. Third, after useful regions are reached, is discovery limited by utility or sampling mass? MOOD comparison reveals a sampling-mass bottleneck; an evolution strategy

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.

(ES) (Beyer & Schwefel, 2002) and our Harvest extension reallocate sampling, yielding 331.3 versus MOOD's 271.3 hits in the three-seed comparison, while MOOD retains its upper-tail advantage.

2 Related Work

Graph generation and latent interfaces. GraphXForm uses sequential generation with constrained actions and policy optimization; CGVAE combines latent graph generation, valence masking, and property optimization (Pirnay et al., 2025; Liu et al., 2018). The public TGVAE implementation connects a graph encoder to an attention-based simplified molecular-input line-entry system (SMILES) decoder (Nguyen & Karolak, 2025). GX-TGVAE couples molecule-level latent conditioning to constraint-masked primitive graph construction within a redesigned autoregressive generator. Neural selection under graph-state rules admits a neuro-symbolic interpretation, not a separate contribution.

Structural vocabularies and ordering. JT-VAE uses junction trees and graph assembly, MoLeR combines motifs with atom-level generation, and MAGNet assembles and colours structural representations (Jin et al., 2018; Maziarz et al., 2022; Hetzel et al., 2025). GX-TGVAE requires no high-level vocabulary, but MoLeR also has an atom-level fallback, precluding a primitive-generation causal interpretation.

Ordering choices include fixed traces, multiple correct actions, trace marginalization, and learned order. CGVAE uses sampled breadth-first search (BFS) traces in a variational treatment; MoLeR combines canonical ordering with multiple correct next actions (Liu et al., 2018; Maziarz et al., 2022). GraphXForm's single-sequence pretraining does not imply the same canonical rule as GX-TGVAE. Order Matters and Learning-Order ARM address ordering itself (Chen et al., 2021; Wang et al., 2025). GX-TGVAE combines a global continuous latent code with construction from an empty graph, a deterministic canonical representative, and VAE reconstruction; canonical ordering alone is not the distinction.

OOD discovery. MOOD combines score-based graph diffusion with explicit OOD control and property guidance (Lee et al., 2023). These capabilities and its novel-hit benchmark make it a direct comparator for separating finite-budget OOD discovery from reached-region utility.

3 GX-TGVAE

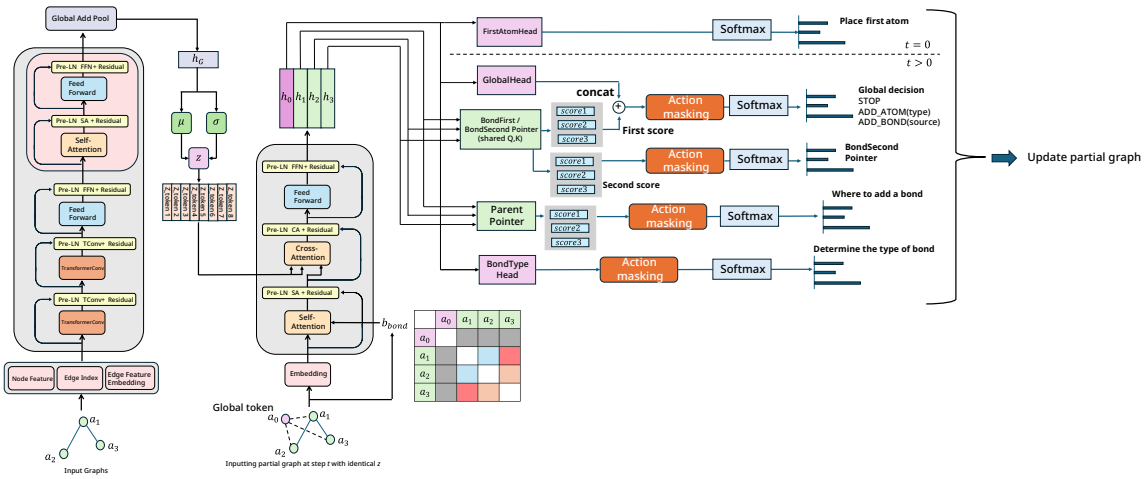


Figure 1: GX-TGVAE, from left to right: encoder, latent interface, decoder, and action selection. TransformerConv (TConv) provides edge-aware local processing. The *persistent global construction token* a_0 directly queries the latent tokens. Block counts are schematic; the final configuration uses E6/S6, an 8-layer decoder, $d_z = 256$, and $K = 8$. Appendix B details the shared BondTypeHead and finite sentinel.

3.1 Latent-conditioned graph construction

Given a molecular graph $G = (V, E)$, edge-aware TransformerConv and global self-attention produce a diagonal Gaussian posterior $q_\phi(z \mid G) = \mathcal{N}(\mu_\phi(G), \operatorname{diag} \sigma^2_\phi(G))$ (Nguyen & Karolak, 2025; Shi et al., 2021; PyTorch Geometric contributors). Training uses reparameterization (Kingma & Welling, 2014), and exploration starts from the posterior mean. The final configuration has an E6/S6 encoder, an 8-layer decoder, $d_z = 256$, and $K = 8$ latent tokens. The linear map $Z(z) = \text{reshape}(W_z z + b_z) \in \mathbb{R}^{K \times d_h}$ is shared across all decoder layers.

Each decoder layer in Figure 1 applies bond-biased self-attention, cross-attention to the latent tokens (Vaswani et al., 2017), and a feed-forward block with pre-norm residual connections. In the final GX-TGVAE configuration (C2_K8), only a_0 directly queries the latent tokens; subsequent self-attention propagates this information to atom states. The *persistent global construction token* a_0 is not a chemical atom. It is present throughout construction, but its hidden state is recomputed from its initial embedding and the current partial graph rather than carried recurrently between steps.

Construction starts by selecting the first atom type for an empty chemical graph. Subsequent actions add an atom (type / parent / bond order), add a bond (source / target / bond order), or terminate. The global decision uses a common index space for termination, atom types, and bond sources. Graph-state rules suppress actions such as self-loops, duplicate bonds, valence violations, and size-bound violations with a finite negative sentinel. This yields a *constraint-masked* distribution before action selection, not a validity guarantee for arbitrary latent codes or sampling conditions. Appendix B details the heads, masks, and AddAtomFirst action.

3.2 Canonical reconstruction

Each graph is mapped to one deterministic canonical sequence $s^* = (a_1^*, \ldots, a_T^*)$ satisfying $F(s^*(G)) = G$. Teacher forcing assigns one target to each prefix and sums cross-entropies over the heads involved:

$$L_{\text{can}}(G, z) = -\log p_\theta(s^*(G) \mid z) = -\sum_{t=1}^{T} \log p_\theta(a_t^* \mid G_{t-1}^*, z). \tag{1}$$

We expand all prefixes into one example per action, sharing the source molecule's latent sample. No prefix subsampling or additional weighting is used. Reconstruction is averaged over actions, while Kullback–Leibler (KL) divergence is averaged over molecules and latent dimensions, giving the implemented objective

$$\widehat{L} = \frac{\sum_{i=1}^{B} L_{\text{can}}(G_i, z_i)}{\sum_{i=1}^{B} T_i} + \frac{\beta(e)}{B d_z} \sum_{i=1}^{B} D_{\text{KL}}(q_\phi(z \mid G_i) \,\|\, \mathcal{N}(0, I)), \tag{2}$$

where $z_i \sim q_\phi$ and $\beta(e)$ is linearly annealed to 1.

Graph likelihood sums probabilities over traces producing the same graph (Liu et al., 2018). With $q_c = p_\theta(s^*(G) \mid z) > 0$, we have $p_\theta(G \mid z) = q_c + r$ and $0 \le r \le 1 - q_c$, hence

$$0 \le L_{\text{can}} + \log p_\theta(G \mid z) \le L_{\text{can}}. \tag{3}$$

Canonical log-probability lower-bounds graph log-likelihood; single-path negative log-likelihood (NLL) upper-bounds graph NLL, with vanishing gap as $q_c \to 1$. Greedy exact reconstruction is not q_c, nor do we equate training and graph-likelihood gradients. Appendix A derives the bounds and distinguishes loss reductions.

3.3 Geometry-guided local latent extrapolation

Starting from an anchor $z_0 = \mu_\phi(G_0)$, we find k-nearest neighbors (kNNs) in a dense bank of training posterior means using the original latent coordinates. Let P_A select coordinates whose mean marginal posterior KL exceeds a threshold. Principal component analysis (PCA) of the centered neighborhood gives $U(z)$, explaining 90% of its variance. With $U_A = \text{orth}(P_A U)$, define

$$P_T = U_A U_A^\top, \qquad P_N = P_A - P_T \tag{4}$$

The initial kNN query excludes the anchor's own bank entry. These are **estimated** spaces, not the true manifold.

For a differentiable target-specific steering surrogate f_ψ, we compute $g = \nabla_z f_\psi(z)$ in the same original coordinates. Writing $\text{unit}(v) = v/\|v\|_2$, the directions are

$$\begin{aligned} v_{\text{ActRaw}} &= \text{unit}(P_A g), \quad v_{\text{ActNorm}} = \text{unit}(P_N g), \\ v_{\text{RandAN}} &= \text{unit}(P_N \epsilon), \qquad v_T = \text{unit}(P_T P_A g), \end{aligned} \quad \epsilon \sim \mathcal{N}(0, I). \tag{5}$$

RandAN draws a fresh direction at each step. A near-zero projected norm raises an error, without fallback.

Conditional geometric justification. Suppose the encoded data concentrate near a C^2 manifold $\mathcal{M}$. At $z \in \mathcal{M}$, let ρ be the length of the **true** normal component of a unit direction v. Under bounded local curvature and sufficiently small $|\alpha|$,

$$\text{dist}(z + \alpha v, \mathcal{M}) \geq \rho|\alpha| - C_{\mathcal{M}}\alpha^2. \tag{6}$$

True normal directions give first-order departure; true tangents give $O(\alpha^2)$. Estimated normals retain a first-order term if their true normal component is sufficiently large (Appendix E, including principal-angle error).

The local-covariance Mahalanobis quadratic form $d^2_{M,\varepsilon}(\delta) = \delta^\top(\Sigma_z + \varepsilon I)^{-1}\delta$ is larger for equal-length steps along lower-variance directions. This does not identify P_N with low density: a rarity interpretation additionally requires distributional assumptions, such as a Gaussian or elliptical approximation. If the expected continuous molecular feature map $F(z) = \mathbb{E}_{G\sim p_\theta(G|z)}[\phi(G)]$ is locally C^2 and the exploration direction satisfies $\|J_F(z)v\| \geq m > 0$, then

$$\|F(z + \alpha v) - F(z)\| \geq m|\alpha| - C_F\alpha^2. \tag{7}$$

This conditional non-collapse guarantees neither uniform sensitivity nor individual molecular OOD. Geometric departure, statistical rarity under additional assumptions, and empirical molecular OOD remain distinct.

Progressive protocol. The fixed reference scale s is calibrated from the local median distance to 50 nearest neighbors using the first 300 anchors of seed 0, excluding self-matches. TanNorm switches to ActNorm after $K_T = 5$ tangent warm-up steps with $\eta/s = 0.25$ ($K_T = 0$ gives ActNorm). The normal phase takes 16 steps of $\delta_N/s = 0.25$, evaluated at $j = 2, 4, 8, 16$, corresponding to cumulative path lengths $\alpha/s = 0.5, 1, 2, 4$. Neighborhoods, PCA, and the required gradients are updated at every step. Evaluation uses a docking oracle distinct from the steering surrogate.

4 Experimental Setup

We evaluate generation, reconstruction, and structural reach on MOSES (Polykovskiy et al., 2020), and OOD exploration on ZINC250k from MOOD's distributed CSV and split indices (Irwin et al., 2012; Lee et al., 2023), training GX-TGVAE separately on each. An internal 10,000-molecule MOSES holdout also supports model selection: architecture comparisons use a fixed 1k subset and baseline comparisons use common canonical 10k inputs. Neither is an independent test. We fixed the architecture before MOOD comparison, emphasizing latent fidelity/use, and apply stochastic weight averaging (SWA) over epochs 90/95/100 (Izmailov et al., 2018).

Generation uses 30,000 attempts per model $\times$ seeds 42/123/2024; the GX-TGVAE architecture comparison uses 1 run with training seed 42 per configuration. Reported variation reflects generation seeds rather than independent training runs. GraphXForm is retrained in a paper-aligned configuration. We distinguish *MAGNet (paper architecture)*, based on Table 3 of the final paper, from *MAGNet (repository-default architecture)*. MAGNet uses only the VAE for reconstruction and includes Flow for prior generation. Appendix H and Appendix D document adjustments, split construction, and baseline provenance. In equations, GX denotes GX-TGVAE.

Distributional evaluation uses MOSES and FCD (Polykovskiy et al., 2020; Preuer et al., 2018). Molecular OOD is measured by maximum Morgan Tanimoto similarity to the ZINC training split used in this study (radius 2, 1,024 bits) (Rogers & Hahn, 2010; RDKit contributors), separately from

identity-based Novelty. Under a common QuickVina2 protocol (Alhossary et al., 2015), the docking score (DS) is $DS = \max(-\text{affinity}, 0)$; higher values indicate better docking scores.

The base fixed-attempt comparison uses 5 targets (parp1, fa7, 5ht1b, braf, jak2) $\times$ seeds 0–4, with 3,000 attempts per configuration and failed outputs retained in the denominator. A candidate passes if it is unique and has similarity < 0.4, quantitative estimate of drug-likeness (QED) > 0.5, and synthetic accessibility (SA) < 5 (Lee et al., 2023; Bickerton et al., 2012; Ertl & Schuffenhauer, 2009). Novel Hit Ratio (NHR) divides passing hits above target-specific DS thresholds by total attempts. Novel Top 5% DS averages the top 150 passing candidates, or all available candidates with their count if fewer exist; empty pools have score and count 0 (Appendix G). Reach/utility comparisons separately use $B = 4$ docking opportunities.

5 Generative and Latent Fidelity

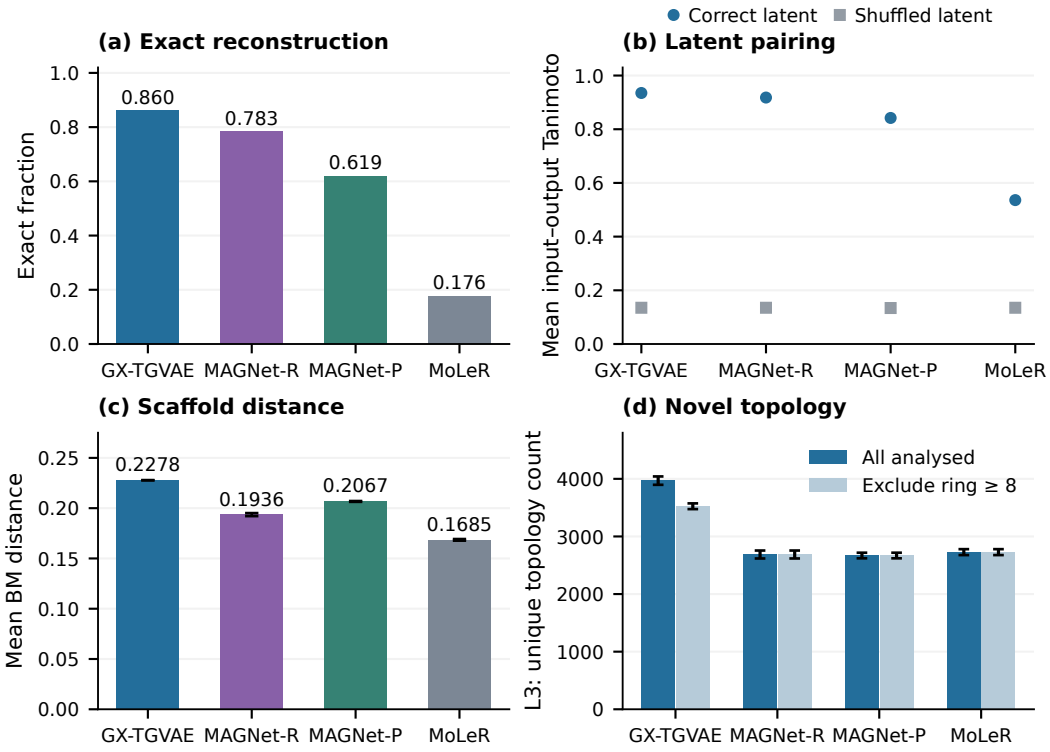


Figure 2: Latent fidelity and structural reach. (a,b) Common 10k reconstruction; MAGNet uses its VAE only. (c,d) 30k attempts $\times$ 3 generation seeds, with sample standard deviations (SDs); MAGNet includes Flow. Distance uses effective-training Bemis–Murcko (BM) scaffolds; L3 counts unique generic topologies absent from training. Large-ring exclusion removes maximum ring sizes ≥ 8. MAGNet-R / -P: repository-default / paper architectures. Full values and denominators: Appendices C–D.

5.1 Distributional generation

In Table 1, GraphXForm has low FCD and low Novelty, whereas TGVAE has high Novelty and lower observed validity. GX-TGVAE occupies a different regime between these, and its FCD ranking relative to MoLeR also differs between Test and TestSF. Checkpoint diagnostics likewise show a model-dependent relationship between improving FCD and changing Novelty (Appendix D). We therefore evaluate distributional fidelity, novelty, molecule-specific latent fidelity, and directed OOD exploration separately: FCD or novelty alone does not establish faithful molecule-specific latents or controllable exploration.

Table 1: MOSES generation: 30,000 molecules per model and generation seed, averaged over 3 generation seeds. Errors are generation-seed standard deviations (SDs), not training-seed uncertainty. GraphXForm values are from paper-aligned retraining, not reported results from the original paper. MAGNet uses the final generation pipeline including Flow.

Model	Validity	Novelty	FCD/Test ↓	FCD/TestSF ↓
Final GX-TGVAE (C2_K8)	1.000 ± 0.000	0.9790 ± 0.0008	0.9832 ± 0.0029	1.301 ± 0.009
GraphXForm, paper-aligned	1.000 ± 0.000	0.8329 ± 0.0021	0.09340 ± 0.00080	0.5472 ± 0.0004
MoLeR	1.000 ± 0.000	0.9691 ± 0.0009	0.8106 ± 0.0128	1.405 ± 0.008
MAGNet (paper architecture) + Flow	1.000 ± 0.000	0.9784 ± 0.0005	1.565 ± 0.030	2.168 ± 0.035
MAGNet (repository-default architecture) + Flow	1.000 ± 0.000	0.9743 ± 0.0008	1.499 ± 0.013	2.095 ± 0.018
TGVAE, epoch 80, greedy	0.8640 ± 0.0023	0.9911 ± 0.0004	1.409 ± 0.008	1.722 ± 0.014

5.2 Latent fidelity

On the common 1k inputs, GX-TGVAE variants C2_K8 and C4_K8 (all-nodes query) achieve exact reconstruction of 0.8580 and 0.7420, respectively, but FCD/Test of 0.9832 and 0.8717 reverses their ranking. Higher distributional fidelity does not necessarily imply higher latent fidelity.

On common 10k inputs (Figure 2(a,b)), GX-TGVAE leads both MAGNet configurations and MoLeR in exact reconstruction and mean Tanimoto. All use identical canonical inputs, ordering, and a 10,000 denominator. Shuffle, zero, and prior controls yield exact reconstruction of 0 for every model, while shuffled Tanimoto falls near the random-pair floor. The paired mean Tanimoto difference between GX-TGVAE and MAGNet-R is +0.0165 (95% confidence interval (CI) [0.0122, 0.0209]); the CIs against the other 2 models also exclude 0. The common 1k comparison gives the same ranking (Appendix C). This measures cross-model latent-to-structure fidelity, not the isolated causal effect of canonical loss or the latent interface.

5.3 Structural reach

Using 30k samples $\times$ 3 generation seeds per model, we measure nearest-training-neighbor distance on Bemis–Murcko (BM) scaffolds (Bemis & Murcko, 1996) and L3 (Figure 2(c,d)). Distance is 1 minus the maximum scaffold Morgan Tanimoto similarity to effective MOSES training scaffolds (radius 2, 2,048 bits). L3 counts unique topologies absent from training, identified by canonical RDKit generic scaffolds. Excluding molecules without a usable scaffold or fingerprint leaves 29,941–29,986 molecules per set for distance analysis.

GX-TGVAE exceeds all compared baselines on all 7 metrics—mean, p90 / p99, L3, L3 after large-ring exclusion, $d > 0.5$, and $d > 0.7$—across all 3 seeds. Baseline L3 rankings vary across seeds. After excluding maximum ring sizes ≥ 8, GX-TGVAE's L3 remains approximately 1.29 times the largest baseline value; large rings alone do not explain the difference. GX-TGVAE has a low out-of-vocabulary (OOV) molecule fraction of 0.0169$\pm$0.0007 relative to the 381-shape vocabulary, further distinguishing shape OOV from structural reach (Appendix D). This establishes neither a primitive-generation causal effect nor set inclusion between representable structures.

6 Geometry-Guided OOD Exploration

In a comparison of four directions across 3,000 anchors per seed and 5 targets $\times$ 5 seeds, property-guided ActNorm achieves higher Novel Top 5% docking scores than fresh-random RandAN at the two deeper tested exploration depths. In Figure 3, ActNorm minus RandAN is +0.237$\pm$0.098 and

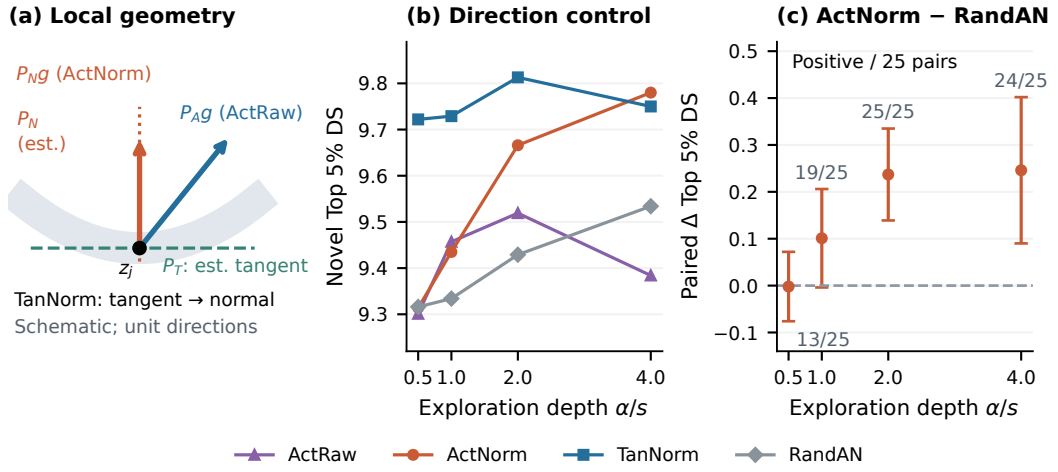


Figure 3: Geometry-guided exploration. (a) Schematic estimated tangent/normal spaces and gradient, not a measured manifold or trajectory. Arrows denote unit directions. (b) Novel Top 5% DS, averaged over seeds within each target and then over 5 targets. (c) Mean ActNorm minus RandAN differences across 25 target–seed pairs, with ±SD. Labels count positive pairs, not CIs or training runs. RandAN draws a fresh random direction at each step. Appendix E and Appendix F give geometric assumptions and the full protocol.

+0.246±0.156 at $\alpha/s = 2, 4$, with positive differences in 25/25 and 24/25 pairs, respectively. The error denotes the standard deviation (SD) of differences across 25 target–seed pairs. This compares configurations with property-gradient direction selection and fresh-random control; it does not identify a pure gradient causal effect.

TanNorm improves docking score over its anchors at $\alpha/s = 2$: the mean $\Delta DS = +1.211$, and the target-averaged fraction of endpoints with $\Delta DS > 0$ is 79.7%. These prespecified endpoint improvements are not trajectory best-of selections and may include reconstruction error or size changes; they do not estimate experimental-activity improvement probabilities.

Without QED/SA gates, increasing TanNorm depth from 0.5 to 4 changes mean maximum similarity from 0.473 to 0.388 and mean DS from 8.784 to 8.913. QED decreases from 0.611 to 0.389, and normalized SA from 0.787 to 0.703. Observed validity is 1.000 across the 400 evaluated GX-TGVAE configurations. These endpoint aggregates imply neither per-molecule monotonic improvement nor preservation of all chemical qualities. Appendix F gives all directions, depths, and anchor-relative analyses.

7 Comparison with MOOD

7.1 Base sampler reveals a sampling-mass bottleneck

Table 2 compares the base samplers over 5 targets × 5 seeds. GX-TGVAE's maxima come from different points in the prespecified 4 arms × 4 depths. **MOOD leads the base sampler on both fixed-attempt discovery metrics.** For total attempts N, unique candidates passing the quality/OOD

Table 2: Fixed-attempt benchmark: 3,000 attempts per configuration across 5 targets and 5 seeds. GX-TGVAE's two maxima come from different settings, not a single operating point.

Metric		GX-TGVAE grid maximum and setting	MOOD, $\lambda = 0.04$
Novel Hit Ratio (%)	3.615	ActNorm, $\alpha/s = 4$	9.120
Novel Top 5% DS	9.813	TanNorm, $\alpha/s = 2$	10.95

gates n_{pass}, and hits n_{hit}, each cell's NHR decomposes, when $n_{\text{pass}} > 0$, as

$$\frac{n_{\text{hit}}}{N} = \underbrace{\frac{n_{\text{pass}}}{N}}_{\text{candidate yield}} \times \underbrace{\frac{n_{\text{hit}}}{n_{\text{pass}}}}_{\text{conditional hit efficiency}}, \quad P(\text{hit}) = P(\text{pass}) \times P(\text{hit} \mid \text{pass}). \tag{8}$$

Passing includes uniqueness, training similarity, QED, and SA, and differs from reach probability below. MOOD at $\lambda = 0.04$ has yield 43.8% and conditional efficiency 21.4%; TanNorm at $\alpha/s = 2$ has 9.3% and 35.0%. Thus GX-TGVAE has higher conditional efficiency, but its main deficit is yield. Products of cell means do not recover mean NHR.

Holding efficiency at 35.0%, matching MOOD's raw hit rate of 9.120% would require yield $y^* = 0.09120/0.350 \simeq 0.261$, or about 26%, in a descriptive counterfactual. This diagnoses the base mass deficit, without forecasting Harvest performance under a different policy and three seeds. Pool-matched Top150 analysis is in Appendix G.4. We next test budget reallocation through Massive Harvest.

7.2 Harvestability-guided Massive Harvest

ES selects among candidate directions; Harvest samples the fixed active-coordinate subspace around selected centers. With GX-TGVAE frozen, a harvestability selector learns local hit rate, rather than center utility, from prior $K = 4$ development data. It selects centers through held-out-seed out-of-fold (OOF) predictions, without docking-score input features. The selector is frozen before fresh Massive Harvest generation and evaluation (Appendix I.1–I.5).

Over 5 targets $\times$ 3 seeds, the learned policy and an oracle-adaptive control allocate 3,000 **evaluation-time docking slots** to selected neighborhoods, matching MOOD's budget. These evaluations reconstruct policies from the saved pool and exclude selector-training costs (Appendix I.6).

Table 3: Massive Harvest at 3,000 evaluation slots, averaged over 5 targets $\times$ 3 seeds (MOOD: $\lambda = 0.04$). Y: candidate yield; H: conditional hit efficiency. This differs from the five-seed base comparison in Table 2.

Strategy	Unique hits	NHR (%)	Y (%)	H (%)	Top 5% DS
TanNorm / learned	256.6	8.553	29.14	28.83	10.308
TanNormES / learned	331.3	11.042	35.80	29.30	10.416
TanNorm / oracle	187.6	6.253	13.03	47.65	10.376
TanNormES / oracle	275.4	9.180	23.06	40.12	10.474
MOOD	271.3	9.044	43.66	21.19	10.939

TanNormES learned yields 331.3 hits / 11.042% NHR, about 22% above MOOD's 271.3 / 9.044%. Differences are positive in 13/15 target–seed cells and 4/5 target means. Non-ES TanNorm learned yields 256.6 hits, below MOOD. The ES configuration has higher candidate yield with similar conditional efficiency (Table 3). Because neighborhood settings also differ, this comparison does not isolate an ES-only causal effect.

Including screening in the same budget, oracle-adaptive TanNormES yields 275.4 hits, slightly above MOOD, but only 8/15 cell differences are positive, with substantial target variation. MOOD retains higher Top 5% DS: 10.939 versus 10.416 for TanNormES learned. Thus improved discovery does not imply upper-tail superiority (Appendix I.5–I.7).

7.3 Reach and utility remain distinct

Across five seeds, base TanNorm ($\alpha/s = 4$) exceeds MOOD ($\lambda = 0.04$) in conditional-median DS summaries in 7 of 8 training-similarity bands; MOOD leads in the deepest band (Figure 4(c,d)). Within-cell medians of unique molecules are averaged over seeds then targets, without QED/SA gates or top-k selection. Residual differences in similarity, size, and target composition limit this to conditional utility of the reached populations.

With $B = 4$, reach C and conditional best-DS expectation Q equally weight 25 target–seed cells. Four depths per GX-TGVAE anchor versus four consecutive MOOD attempts induce different de-

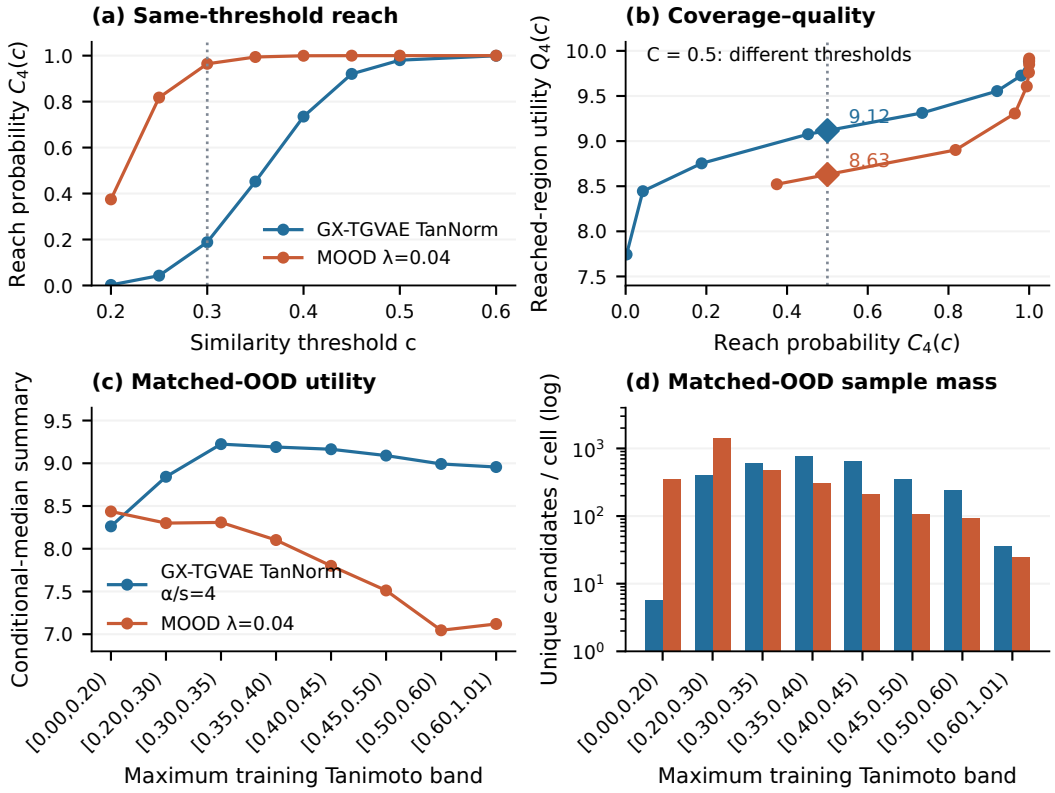


Figure 4: Reach and conditional utility. (a,b) Four depths per GX-TGVAE anchor versus four consecutive MOOD attempts; 25 target–seed cells have equal weight. MOOD leads at $c = 0.3$; diamonds match $C = 0.5$ at different thresholds (0.36 / 0.21). (c,d) TanNorm $\alpha/s = 4$ versus MOOD $\lambda = 0.04$: within-cell DS medians averaged over seeds then targets, and mean unique candidate counts (log scale). All band cells are nonempty; no QED/SA gates or top-k selection. Full definitions and tables: Appendix G.

pendence structures. At the same $c = 0.3$, GX-TGVAE has $(C, Q) = (0.188, 8.75)$ and MOOD $(0.964, 9.31)$: **MOOD is higher on both axes**. Interpolation to $C = 0.5$ gives $c = 0.36, Q = 9.12$ for GX-TGVAE and $c = 0.21, Q = 8.63$ for MOOD, but **the thresholds differ** (Figure 4(a,b)). Reached-trial conditioning also entails selection effects (Appendix G.6). Harvest does not establish superiority on all axes.

8 Discussion, Limitations and Conclusion

Our experiments distinguish three bottlenecks: molecule-specific fidelity, directional control, and sampling mass. GX-TGVAE addresses the first two through latent-conditioned constrained generation and geometry-guided exploration; Harvest allocates mass to promising neighborhoods. The results suggest distinct objectives for point utility and neighborhood harvestability.

Architecture comparisons use one training run; docking scores are proxies for measured activity, and deeper exploration reduces QED/SA. Massive Harvest is exploratory and designed after the pilot, with target-dependent results and a remaining MOOD upper-tail advantage. Future work should address training robustness and chemical quality, prospectively validate harvestability models, and develop geometry-aware local sampling, potentially through information geometry or flow-based latent transport.

GX-TGVAE connects faithful representation, directional control, and sampling-mass allocation in one continuous latent generation framework, providing a foundation for faithful reconstruction and the discovery of useful unfamiliar structures.

AI USE STATEMENT

Generative AI tools were used as research assistants in this work. Their main uses included literature discovery and organization, research discussion and critical feedback, assistance with implementation and debugging, support for examining experimental results and analyses, manuscript review and auditing, writing and language editing, and English translation. For mathematical arguments and proofs, AI tools were used to help organize the presentation and examine the logical consistency of ideas and claims developed by the authors. The central research ideas, methodological decisions, and final scientific interpretations were made by the authors. AI-assisted code, analyses, mathematical arguments, and text were reviewed by the authors and revised or rejected when necessary. The authors take responsibility for the final content of this work.

References

Eddie Aamari and Clément Levrard. Nonasymptotic rates for manifold, tangent space and curvature estimation. *The Annals of Statistics*, 47(1):177–204, 2019. doi: 10.1214/18-AOS1685. URL https://doi.org/10.1214/18-AOS1685.

Amr Alhossary, Stephanus Daniel Handoko, Yuguang Mu, and Chee-Keong Kwoh. Fast, accurate, and reliable molecular docking with QuickVina 2. *Bioinformatics*, 31(13):2214–2216, 2015. doi: 10.1093/bioinformatics/btv082. URL https://academic.oup.com/bioinformatics/article/31/13/2214/195750.

G. W. Bemis and M. A. Murcko. The properties of known drugs. 1. molecular frameworks. *Journal of Medicinal Chemistry*, 39(15):2887–2893, 1996. doi: 10.1021/jm9602928. URL https://pubmed.ncbi.nlm.nih.gov/8709122/.

Hans-Georg Beyer and Hans-Paul Schwefel. Evolution strategies—a comprehensive introduction. *Natural Computing*, 1(1):3–52, 2002. doi: 10.1023/A:1015059928466. URL https://doi.org/10.1023/A:1015059928466.

G. Richard Bickerton, Gaia V. Paolini, Jérémy Besnard, Sorel Muresan, and Andrew L. Hopkins. Quantifying the chemical beauty of drugs. *Nature Chemistry*, 4(2):90–98, 2012. doi: 10.1038/nchem.1243. URL https://doi.org/10.1038/nchem.1243.

Xiaohui Chen, Xu Han, Jiajing Hu, Francisco Ruiz, and Liping Liu. Order matters: Probabilistic modeling of node sequence for graph generation. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pp. 1630–1639, 2021. URL https://proceedings.mlr.press/v139/chen21j.html.

Zhiyuan Chen, Xiaomin Fang, Zixu Hua, Yueyang Huang, Fan Wang, and Hua Wu. HelixMO: Sample-efficient molecular optimization in scene-sensitive latent space. In *2022 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pp. 315–321, 2022. doi: 10.1109/BIBM55620.2022.9995561. URL https://ieeexplore.ieee.org/document/9995561.

B. Efron. Bootstrap methods: Another look at the jackknife. *The Annals of Statistics*, 7(1):1–26, 1979. doi: 10.1214/aos/1176344552. URL https://doi.org/10.1214/aos/1176344552.

B. Efron and R. Tibshirani. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical Science*, 1(1):54–75, 1986. doi: 10.1214/ss/1177013815. URL https://doi.org/10.1214/ss/1177013815.

Peter Ertl and Ansgar Schuffenhauer. Estimation of synthetic accessibility score of drug-like molecules based on molecular complexity and fragment contributions. *Journal of Cheminformatics*, 1:8, 2009. doi: 10.1186/1758-2946-1-8. URL https://link.springer.com/article/10.1186/1758-2946-1-8.

Morten W. Fagerland, Stian Lydersen, and Petter Laake. The McNemar test for binary matched-pairs data: mid-p and asymptotic are better than exact conditional. *BMC Medical Research Methodology*, 13:91, 2013. doi: 10.1186/1471-2288-13-91. URL https://doi.org/10.1186/1471-2288-13-91.

Leon Hetzel, Johanna Sommer, Bastian Rieck, Fabian Theis, and Stephan Günnemann. MAGNet: Motif-agnostic generation of molecules from scaffolds. In *International Conference on Learning Representations*, 2025. URL https://proceedings.iclr.cc/paper_files/paper/2025/hash/43d1d3bdd92204c96fa4ac3c578f6a33-Abstract-Conference.html.

John J. Irwin, Teague Sterling, Michael M. Mysinger, Erin S. Bolstad, and Ryan G. Coleman. ZINC: A free tool to discover chemistry for biology. *Journal of Chemical Information and Modeling*, 52(7):1757–1768, 2012. doi: 10.1021/ci3001277. URL https://doi.org/10.1021/ci3001277.

Pavel Izmailov, Dmitrii Podoprikhin, Timur Garipov, Dmitry Vetrov, and Andrew Gordon Wilson. Averaging weights leads to wider optima and better generalization. In *Proceedings of the Thirty-Fourth Conference on Uncertainty in Artificial Intelligence*, 2018. URL https://auai.org/uai2018/proceedings/papers/313.pdf.

Wengong Jin, Regina Barzilay, and Tommi Jaakkola. Junction tree variational autoencoder for molecular graph generation. In *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pp. 2323–2332, 2018. URL https://proceedings.mlr.press/v80/jin18a.html.

Daniel N. Kaslovsky and François G. Meyer. Non-asymptotic analysis of tangent space perturbation. *Information and Inference: A Journal of the IMA*, 3(2):134–187, 2014. doi: 10.1093/imaiai/iau004. URL https://doi.org/10.1093/imaiai/iau004.

Diederik P. Kingma and Max Welling. Auto-encoding variational bayes. In *International Conference on Learning Representations*, 2014. URL https://arxiv.org/abs/1312.6114.

Can Koban, Gökçe Uludoğan, Elif Ozkirimli, and Arzucan Özgür. ConGen: Targeted molecule generation through contrastive learning and latent optimization. *Molecular Informatics*, 45(5): e70038, 2026. doi: 10.1002/minf.70038. URL https://onlinelibrary.wiley.com/doi/10.1002/minf.70038.

Seul Lee, Jaehyeong Jo, and Sung Ju Hwang. Exploring chemical space with score-based out-of-distribution generation. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pp. 18872–18892, 2023. URL https://proceedings.mlr.press/v202/lee23f.html.

Qi Liu, Miltiadis Allamanis, Marc Brockschmidt, and Alexander L. Gaunt. Constrained graph variational autoencoders for molecule design. In *Advances in Neural Information Processing Systems*, volume 31, 2018. URL https://proceedings.neurips.cc/paper/2018/hash/b8a03c5c15fcfa8dae0b03351eb1742f-Abstract.html.

Youzhi Luo, Keqiang Yan, and Shuiwang Ji. GraphDF: A discrete flow model for molecular graph generation. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pp. 7192–7203, 2021. URL https://proceedings.mlr.press/v139/luo21a.html.

Krzysztof Maziarz, Henry Jackson-Flux, Pashmina Cameron, Finton Sirockin, Nadine Schneider, Nikolaus Stiefl, Marwin Segler, and Marc Brockschmidt. Learning to extend molecular scaffolds with structural motifs. In *International Conference on Learning Representations*, 2022. URL https://openreview.net/forum?id=ZTsoE8G3GG.

Quinn McNemar. Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika*, 12:153–157, 1947. doi: 10.1007/BF02295996. URL https://doi.org/10.1007/BF02295996.

Rocío Mercado, Tobias Rastemo, Edvard Lindelöf, Günter Klambauer, Ola Engkvist, Hongming Chen, and Esben Jannik Bjerrum. Graph networks for molecular design. *Machine Learning: Science and Technology*, 2(2):025023, 2021. doi: 10.1088/2632-2153/abcf91. URL https://doi.org/10.1088/2632-2153/abcf91.

Trieu Nguyen and Aleksandra Karolak. Transformer graph variational autoencoder for generative molecular design. *Biophysical Journal*, 124(22):3867–3875, 2025. doi: 10.1016/j.bpj.2025.01.022. URL https://doi.org/10.1016/j.bpj.2025.01.022.

Jonathan Pirnay, Jan G. Rittig, Alexander B. Wolf, Martin Grohe, Jakob Burger, Alexander Mitsos, and Dominik G. Grimm. GraphXForm: Graph transformer for computer-aided molecular design. *Digital Discovery*, 4(4):1052–1065, 2025. doi: 10.1039/D4DD00339J. URL https://doi.org/10.1039/D4DD00339J.

Daniil Polykovskiy, Alexander Zhebrak, Benjamin Sanchez-Lengeling, Sergey Golovanov, Oktai Tatanov, Stanislav Belyaev, Rauf Kurbanov, Aleksey Artamonov, Vladimir Aladinskiy, Mark

Veselov, Artur Kadurin, Simon Johansson, Hongming Chen, Sergey Nikolenko, Alán Aspuru-Guzik, and Alex Zhavoronkov. Molecular sets (MOSES): A benchmarking platform for molecular generation models. *Frontiers in Pharmacology*, 11:565644, 2020. doi: 10.3389/fphar.2020.565644. URL https://www.frontiersin.org/journals/pharmacology/articles/10.3389/fphar.2020.565644/full.

Kristina Preuer, Philipp Renz, Thomas Unterthiner, Sepp Hochreiter, and Günter Klambauer. Fréchet ChemNet Distance: A metric for generative models for molecules in drug discovery. *Journal of Chemical Information and Modeling*, 58(9):1736–1741, 2018. doi: 10.1021/acs.jcim.8b00234. URL https://pubmed.ncbi.nlm.nih.gov/30118593/.

PyTorch Geometric contributors. PyTorch Geometric: TransformerConv documentation. URL https://pytorch-geometric.readthedocs.io/en/latest/generated/torch_geometric.nn.conv.TransformerConv.html.

RDKit contributors. RDKit: Open-source cheminformatics. URL https://www.rdkit.org.

David Rogers and Mathew Hahn. Extended-connectivity fingerprints. *Journal of Chemical Information and Modeling*, 50(5):742–754, 2010. doi: 10.1021/ci100050t. URL https://pubmed.ncbi.nlm.nih.gov/20426451/.

Chence Shi, Minkai Xu, Zhaocheng Zhu, Weinan Zhang, Ming Zhang, and Jian Tang. GraphAF: a flow-based autoregressive model for molecular graph generation. In *International Conference on Learning Representations*, 2020. URL https://openreview.net/forum?id=S1esMkHYPr.

Yunsheng Shi, Zhengjie Huang, Shikun Feng, Hui Zhong, Wenjing Wang, and Yu Sun. Masked label prediction: Unified message passing model for semi-supervised classification. In *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence*, pp. 1548–1554, 2021. doi: 10.24963/ijcai.2021/214. URL https://www.ijcai.org/proceedings/2021/214.

C. Spearman. The proof and measurement of association between two things. *The American Journal of Psychology*, 15(1):72–101, 1904. doi: 10.2307/1412159. URL https://doi.org/10.2307/1412159.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. In *Advances in Neural Information Processing Systems*, volume 30, 2017. URL https://proceedings.neurips.cc/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html.

Zhe Wang, Jiaxin Shi, Nicolas Heess, Arthur Gretton, and Michalis Titsias. Learning-order autoregressive models with application to molecular graph generation. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 64282–64301, 2025. URL https://proceedings.mlr.press/v267/wang25cn.html.

A Canonical Reconstruction and Graph Likelihood

A.1 Ordering and trace ambiguity in sequential graph generation

An unordered graph has no unique construction order. Multiple complete action sequences can yield the same graph G, differing in atom insertion order, the timing of ring-closing bonds, or the order of symmetric atoms. Write the set of complete sequences as

$$\mathcal{S}(G) = \{s : F(s) = G\}$$

where F deterministically maps a complete action sequence to its final graph. For a latent-conditioned autoregressive generator, unordered graph likelihood is

$$p_\theta(G \mid z) = \sum_{s \in \mathcal{S}(G)} p_\theta(s \mid z) \tag{A1}$$

Methods differ in their treatment of trace marginalization and ordering. CGVAE uses a variational lower bound through BFS traces and sampling-based approximations; this does not imply that every training configuration always uses multiple traces (Liu et al., 2018). MoLeR can supervise multiple correct next actions for the current partial molecule, which is compatible with canonical ordering (Maziarz et al., 2022). GraphAF and GraphDF use fixed ordered representations, but this alone does not define a unique canonical representative for each graph (Shi et al., 2020; Luo et al., 2021). We position GraphINVENT as a sequential graph generator (Mercado et al., 2021). Order Matters infers order as a latent variable, while Learning-Order ARM learns an ordering policy (Chen et al., 2021; Wang et al., 2025).

GX-TGVAE does not propose a general solution to the ordering problem. For each graph, it fixes a deterministic canonical construction map

$$C : G \mapsto s^*(G) \in \mathcal{S}(G) \tag{A2}$$

and reconstructs that representative trace from a molecular latent code. This reduces VAE training to standard autoregressive teacher forcing without enumerating, sampling, or marginalizing $\mathcal{S}(G)$.

A.2 Assumptions on canonical targets

We assume that the canonicalizer satisfies the following conditions.

1. **Determinism:** The same molecular graph yields the same canonical action sequence.
2. **Completeness:** The sequence is complete, from the first-atom action to termination.
3. **Round-trip:** $F(s^*(G)) = G$.
4. **Feasibility consistency:** At each teacher-forced prefix, the target action is consistent with the implemented structural mask.

The implementation constructs targets using RDKit canonical atom ranking and BFS. It starts from an empty chemical graph, adds the first atom, and constructs the graph sequentially. Appendix B details the action encoding.

After Kekulization, the implementation computes RDKit canonical atom ranks, starts from the lowest-ranked atom, and processes each atom's neighbors in rank order. Edges to unvisited atoms yield ADD_ATOM; unprocessed edges between visited atoms yield ADD_BOND; STOP ends the sequence. The graph contains atom types and bond orders represented by the action vocabulary, without independent actions for charge or stereochemistry. Checking molecular identity after preprocessing differs from validating the canonicalizer under arbitrary atom permutations.

A.3 Canonical log-probability lower-bounds graph log-likelihood

Define the canonical path probability as

$$q_c := p_\theta(s^*(G) \mid z) > 0$$

and the total probability of other sequences reaching the same graph as

$$r := \sum_{s \in \mathcal{S}(G) \setminus \{s^*(G)\}} p_\theta(s \mid z) \geq 0$$

Then

$$p_\theta(G \mid z) = q_c + r. \tag{A3}$$

Since the total probability over all sequences is at most 1,

$$0 \leq r \leq 1 - q_c. \tag{A4}$$

It follows that

$$p_\theta(G \mid z) \geq q_c$$

and taking logarithms gives

$$\boxed{\log p_\theta(G \mid z) \geq \log p_\theta(s^*(G) \mid z)} \tag{A5}$$

Thus, canonical-path log-probability is a computable lower bound on unordered graph log-likelihood; reversing the sign makes single-path NLL an upper bound on graph NLL. The following is our derivation for this fixed canonical representative.

Writing canonical path NLL and graph NLL as

$$L_{\text{can}} = -\log q_c, \qquad L_G = -\log(q_c + r)$$

the gap is

$$\begin{aligned} L_{\text{can}} - L_G &= \log \frac{q_c + r}{q_c} \\ &= \log \left(1 + \frac{r}{q_c}\right) \geq 0. \end{aligned} \tag{A6}$$

Furthermore, $L_G \geq 0$ gives

$$\boxed{0 \leq L_{\text{can}} - L_G \leq L_{\text{can}}} \tag{A7}$$

which is equivalent to Equation (3) in the main text.

A.3.1 A gap bound from probability concentration

If the probability assigned to the canonical path satisfies

$$q_c \geq 1 - \varepsilon, \qquad 0 < \varepsilon < 1$$

then Equation (A4) implies

$$0 \leq p_\theta(G \mid z) - q_c = r \leq \varepsilon. \tag{A8}$$

Moreover,

$$0 \leq L_{\text{can}} - L_G \leq L_{\text{can}} \leq -\log(1-\varepsilon). \tag{A9}$$

Probability concentration on the canonical path is therefore sufficient to bring the single-path objective close to graph likelihood in both probability and NLL space.

This does **not** state that maximizing canonical path probability is generally equivalent to maximizing graph likelihood: the probability r of other valid traces also contributes to graph likelihood.

A.4 Teacher forcing and faithful reconstruction

For the canonical sequence

$$s^*(G) = (a_1^*, \ldots, a_T^*)$$

the autoregressive factorization is

$$p_\theta(s^*(G) \mid z) = \prod_{t=1}^{T} p_\theta(a_t^* \mid G_{t-1}^*, z). \tag{A10}$$

Therefore,

$$L_{\text{can}}(G, z) = -\sum_{t=1}^{T} \log p_\theta(a_t^* \mid G_{t-1}^*, z). \tag{A11}$$

Single-target cross-entropy increases the target action's conditional probability at the canonical teacher prefix. Although it does not directly optimize the probabilities of all traces for the same graph, it directly serves the VAE reconstruction objective of **recovering a unique teacher trajectory with high probability from latent** z.

Greedy exact reconstruction is not the path probability q_c in Equation (A10). The argmax can match the target at every step and yield exact reconstruction even when the product of stepwise probabilities is small. We maintain this distinction throughout the main text and appendices.

A.5 Prefix expansion and training loss reduction

Each molecule's canonical trajectory is expanded into one prefix example per action, from the first action through STOP. All prefixes of each selected molecule are used, without prefix subsampling or additional prefix weighting. Prefixes from the same molecule share the sampled latent z_i. Prefix batching passes multiple teacher-forced partial graphs to the decoder together; it does not change the probabilistic definition of the sequence.

For ADD_ATOM, cross-entropies for Level 0, parent, and bond type are summed; ADD_BOND sums those for Level 0, target, and bond type. The first atom uses a dedicated head, and STOP uses only Level 0. The sum over these conditional heads is the NLL of one action; it is not averaged again over the number of heads.

Let $L_{\text{can}}(G_i, z_i)$ be molecule i's canonical path NLL and T_i its action count. The batch reconstruction reduction is

$$L_{\text{recon}} = \frac{\sum_i L_{\text{can}}(G_i, z_i)}{\sum_i T_i}. \tag{A12}$$

KL is averaged over molecules and latent dimensions, giving the training objective

$$\widehat{L} = L_{\text{recon}} + \frac{\beta(e)}{Bd_z} \sum_{i=1}^{B} D_{\text{KL}}(q_\phi(z \mid G_i) \| \mathcal{N}(0, I)). \tag{A13}$$

where $\beta(e)$ is linearly annealed. Equation (A11) is therefore sequence-level NLL, Equation (A12) is the implemented action-average reduction, and Equation (A13) is the full VAE training loss. The latter two should not be interpreted as having the same normalization as molecule-level graph NLL or an ELBO.

A.6 Implementation checks for canonical targets

Preprocessing includes a round-trip check that reconstructs the graph from its action sequence and compares canonical SMILES, and a check for consistency between target actions and masks at teacher prefixes. Both are enabled in the final launch configurations. The current mask-consistency check covers Level 0, ADD_ATOM parents, and ADD_BOND targets. Bond-type masks are applied separately in the decoder, so the existence of this check is not an exhaustive validation of all heads.

These code-level checks differ from an empirical demonstration of canonical invariance under arbitrary atom-index permutations. We do not report unverified check success rates or molecule counts here.

B Architecture and Implementation Details

B.1 Encoder and latent representation

Figure B1 illustrates the repeated encoder structure.

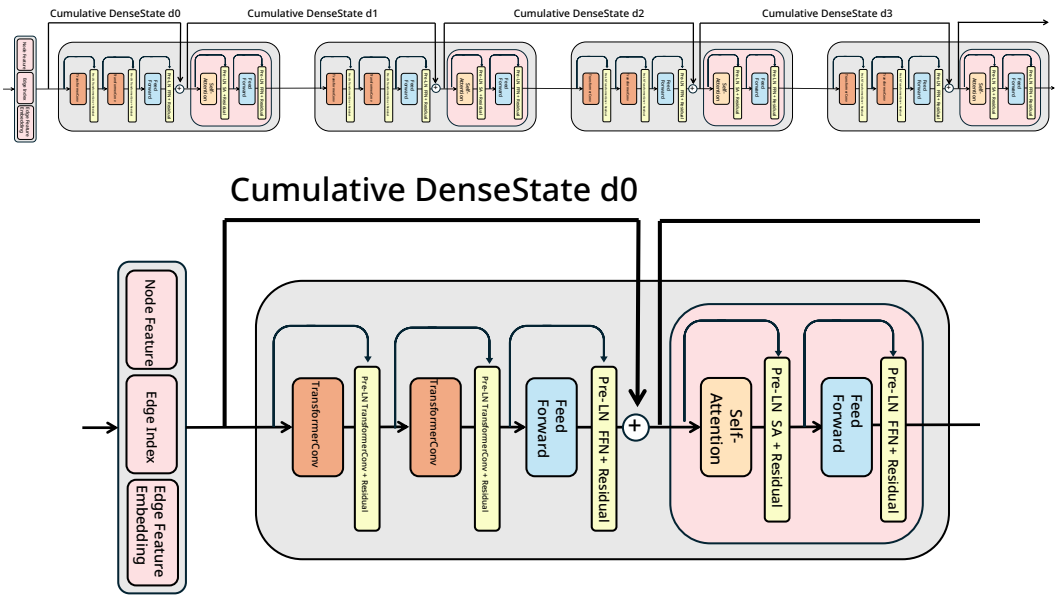


Figure B1: Encoder local / global processing and cumulative dense states. Top: the repeated structure. Bottom: an enlarged view of the input and first repeated unit. Each local block applies TransformerConv twice. The displayed block count is schematic; the final configuration is E6/S6. Dense updates operate on local block outputs, followed by global self-attention. Real-atom representations are sum-pooled after the final LayerNorm.

GX-TGVAE encodes molecular graphs using edge-aware local message passing and global self-attention. The final E6/S6 configuration has 6 local encoder blocks and 6 global self-attention blocks. Each local block contains two TransformerConv operations and a feed-forward transformation; E6

therefore does not mean six TransformerConv calls. UniMP and the PyTorch Geometric documentation provide the mechanism and implementation references, respectively (Shi et al., 2021; PyTorch Geometric contributors).

The final encoder uses cumulative dense connections over local-block outputs. Each block applies global self-attention after its local processing and dense update. After the final LayerNorm, we sum-pool real-atom representations. From this graph-level representation we obtain

$$\mu_\phi(G), \quad \log \sigma^2_\phi(G)$$

and define

$$q_\phi(z \mid G) = \mathcal{N}(\mu_\phi(G), \operatorname{diag} \sigma^2_\phi(G))$$

The latent dimension is $d_z = 256$. Training uses Gaussian reparameterization samples (Kingma & Welling, 2014), whereas latent diagnostics and exploration anchors use posterior means.

The use of a graph encoder to obtain a latent representation is inspired by TGVAE (Nguyen & Karolak, 2025). Our encoder configuration, depth, pooling, and dense connections are not claimed to match the original TGVAE settings. Appendix H discusses the public TGVAE implementation and the wrapper used in our baseline experiments.

Molecular graph features. The encoder receives atom-type indices, mapped to learned 256-dimensional embeddings, and a bidirectional sparse edge list. Each edge carries a bond-order index (single / double / triple: 0 / 1 / 2), mapped to a learned 64-dimensional embedding supplied to TransformerConv. Aromatic bonds are Kekulized before graph conversion. No separate charge, chirality, aromaticity, or degree feature is concatenated to the encoder input; decoder degree embeddings are a separate mechanism.

B.2 Latent-conditioning tokens and persistent global token a_0

A linear projection maps the latent code to

$$Z(z) = \text{reshape}(W_z z + b_z) \in \mathbb{R}^{K \times d_h}$$

The final configuration uses $K = 8$ and $d_h = 256$. Decoder states include chemical atom states and a_0, which is not a chemical atom. The *persistent global construction token* a_0 serves as:

1. The sole construction token in the initial state, before any chemical atom exists.
2. A global query that directly reads the latent-conditioning tokens.
3. A global hub / readout that participates in self-attention with the entire partial graph.
4. A construction token present at every decoding step.

The input states at each step are constructed from the learned initial embedding for a_0 and the current partial graph. Real atoms use atom-type and degree embeddings. "Persistent" means that a token with this role is present at every step; its hidden embedding is not carried over from the previous step as in an RNN.

The decoder has 8 layers, each applying bond-biased self-attention, latent cross-attention, and a feed-forward transformation in that order. Every sublayer uses pre-norm and a residual connection, and the stack ends with LayerNorm. The general attention mechanism follows the Transformer (Vaswani et al., 2017). The conditioner containing the latent projection and cross-attention is shared across layers.

In the final `a0_only` configuration, only a_0 queries the latent tokens. Information acquired through cross-attention in that layer propagates to chemical atom states through graph self-attention in subsequent layers. Under `all_nodes`, every node queries the latent tokens. Changing the query scope itself adds no parameters.

B.3 Autoregressive action space

Generation starts from an empty chemical graph. The initial `ADD_ATOM_FIRST` (AddAtomFirst) action uses `FirstAtomHead` to select an atom type from a_0 and create the first real atom. Later

actions add a new atom, add a bond between existing atoms, or STOP. Each action updates the partial graph, and node states are recomputed at the next step.

Let N denote the number of real atoms and N_{atom} the number of atom types. Indices below are zero-based; real-atom indices exclude a_0.

Level 0 concatenates the STOP and atom-type logits from the global head with bond-source scores computed from a_0 and real-atom representations. ADD_ATOM conditionally selects a parent and bond type; ADD_BOND selects a target and bond type. BondSecond selects the target, not the bond type. The shared bond-type head computes logits from a_0 for both branches. BondSecond marks the source using a learned source-marking embedding and uses the same query/key projections as bond-source selection.

Thus, the training losses for ADD_ATOM and ADD_BOND each sum cross-entropies for three conditional decisions. STOP uses only Level 0, and the initial atom uses only FirstAtomHead. Prefix batching and reduction follow Appendix A.5.

B.4 Structural action masking

Immediately before action selection, we apply masks that depend on the current partial graph.

- Level 0 suppresses atom addition when the configured maximum atom count is reached or all atoms have saturated valence. It also suppresses bond sources with no remaining valence or no eligible target.
- Parent selection suppresses existing atoms with saturated valence.
- Bond-target selection suppresses self-loops, duplicate bonds, and targets with saturated valence.
- Bond-type selection suppresses orders exceeding the smaller remaining valence of the two endpoints. ADD_ATOM also uses the allowed valence of the new atom type.

Masked logits are replaced by the finite sentinel -10^4. This is not a mathematical $-\infty$ mask and does not guarantee exactly zero probability for invalid actions under arbitrary latent inputs or sampling conditions. Generation applies softmax / sampling to masked logits, while reconstruction evaluation uses greedy selection for the corresponding decisions. We therefore describe the mechanism as *constraint-masked* generation and distinguish it from observed validity.

This design admits a neuro-symbolic interpretation: neural latent representations guide structural choices, while graph-state rules suppress inappropriate candidates before action selection. This is an architectural interpretation of GX-TGVAE, not a separate novel contribution.

B.5 Final configuration and checkpoint provenance

The final architecture is C2_K8. Table 5 lists the main settings shared between MOSES and ZINC and the differences arising from their data vocabularies.

The MOSES atom vocabulary is C, N, O, F, S, Cl, Br; the ZINC vocabulary is C, N, O, F, P, S, Cl, Br, I. Both use single, double, and triple bonds, without independent ionization or chirality actions.

The final encoder uses 6 local blocks and 6 global self-attention blocks, and the decoder uses 8 layers. C4_K8 changes the latent-query scope from the global token to all nodes while keeping $K = 8$ and the parameter count of MOSES C2_K8 unchanged (19,300,434). Appendix C describes architecture selection.

The final checkpoint uses SWA over epochs 90 / 95 / 100 (Izmailov et al., 2018). Floating-point tensors are arithmetically averaged; non-floating tensors use the last checkpoint's values.

Training settings. Both final models use AdamW with constant learning rate 1.41×10^{-4}, betas $(0.9, 0.999)$, epsilon 10^{-8}, weight decay 0, and gradient-norm clipping at 1.0. Batch size is 128 molecules, with 3,000 updates per epoch for MOSES and 388 for ZINC. These are configured training epochs rather than full passes over all molecules. The KL coefficient at training epoch e, starting at 1, is $\min(1, e/50)$; it remains 1 after the 50-epoch warmup. The objective uses canonical teacher

Table 4: GX-TGVAE action decisions and their implementation. Level 0 concatenates Global outputs and bond-source scores into a common index space; it does not select every argument of every action jointly through a single head.

Role	Head / pointer	Index / condition
First atom	`FirstAtomHead`	Atom type $0, \ldots, N_{\text{atom}} - 1$; initial step only
Global	Global head	STOP: 0; addition of atom type k: $1 + k$
Bond source	Bond-source pointer	Level 0 index of source i: $1 + N_{\text{atom}} + i$
Parent	Parent pointer	ADD_ATOM parent $0, \ldots, N - 1$
Bond target	Bond-target pointer	ADD_BOND target $0, \ldots, N - 1$; uses source marking
Bond type	Bond-type head	Shared by ADD_ATOM / ADD_BOND; 0 / 1 / 2 denote single / double / triple

Table 5: Final GX-TGVAE configuration. Encoder E6/S6 and decoder D8 refer to the effective module settings.

Setting	MOSES	ZINC250k
Encoder / decoder	E6/S6 / D8	E6/S6 / D8
Latent dimension / tokens	256 / 8	256 / 8
Hidden width / attention heads	256 / 8	256 / 8
Feed-forward width / dropout	1024 / 0.1	1024 / 0.1
Cross-attention scope	`a0_only`	`a0_only`
Encoder readout	Sum over real atoms	Sum over real atoms
Atom types / maximum atoms	7 / 27	9 / 38
Parameters	19,300,434	19,302,486
Training seed / epochs	42 / 100	42 / 100
SWA source epochs	90, 95, 100	90, 95, 100

forcing with no label smoothing; prefix batching and loss reduction are specified in Appendix A.5. Dropout is 0.1, training lasts 100 epochs with seed 42, and EMA is disabled. Periodic checkpoints are saved every 5 epochs; the final evaluation model is the fixed SWA of epochs 90, 95, and 100.

B.6 Components inherited from GraphXForm and GX-TGVAE modifications

We build on GraphXForm's graph-state autoregressive construction, chemical graph action representation, and structural-constraint approach (Pirnay et al., 2025). We do not claim these as GX-TGVAE-specific novelties.

GX-TGVAE connects a molecular VAE latent representation to canonical reconstruction from an empty graph, combining AddAtomFirst, the persistent global construction token a_0, latent cross-attention at every decoder layer, a multi-token latent interface, and latent geometry operators.

C Architecture Selection and Latent Fidelity

C.1 Selection principle

The final architecture was selected on MOSES before examining OOD / MOOD results. Selection emphasized molecule-specific latent fidelity as well as FCD, and fixed the SWA window to epochs 90 / 95 / 100. The general SWA method is due to Izmailov et al. (2018); the window and architecture choices here are conditions of our experiments.

We held out 10,000 molecules from the MOSES training pool as internal validation. A fixed subset of 1,000 inputs was used for paired architecture evaluation, and canonical 10,000 inputs for baseline

parity. These are not independent test sets: the 1k set is a subset of the 10k set. Appendix H.1 gives split counts and identities.

In this section's tables, GX-TGVAE denotes final C2_K8, MAGNet-P denotes MAGNet (paper architecture), and MAGNet-R denotes MAGNet (repository-default architecture). MAGNet latent reconstruction uses the VAE encoder–decoder without Flow.

C.2 Scope $\times K$ comparison

Table 6 reports greedy decoding from posterior means for the same 1,000 inputs. K is the latent-token count; scope specifies whether only a_0 or all nodes query the latent representation through cross-attention.

Table 6: Exact reconstruction on the same 1,000 inputs. a0_only and all_nodes denote cross-attention scope.

Scope	$K = 4$	$K = 8$	Δ_K
a0_only	0.6510	0.8580	+0.2070
all_nodes	0.6190	0.7420	+0.1230

Increasing $K = 4$ to 8 improves exact reconstruction under both scopes. The difference between the improvements is

$$I = (0.742 - 0.619) - (0.858 - 0.651) = -0.0840 \pm 0.0173$$

Here, $\pm$ denotes the standard error from 10,000 bootstrap resamples of paired inputs (Efron, 1979), not the SD across generation or training seeds. This interaction is a descriptive comparison, not a preregistered formal test.

Table 7 compares C2_K8 (a0_only) and C4_K8 (all_nodes) in the final window. The prespecified criteria for adopting C4 were a positive paired exact-reconstruction difference with $z \geq 3$ and lower internal-validation FCD. The exact difference was -0.1160, with $z = -8.94$, failing the first criterion; better FCD alone did not justify adoption. Test FCD was evaluated after selection. The reversed rankings by exact reconstruction and FCD motivate separate evaluation of distributional and latent fidelity.

Table 7: Final-window architecture comparison. Differences are C4 minus C2.

Metric	C2 K8	C4 K8	Δ
exact reconstruction	0.8580	0.7420	-0.1160
posterior Tanimoto	0.9334	0.8812	-0.0522
FCD/InternalVal	1.1897	1.0548	C4 better
FCD/Test (30k×3)	0.9832	0.8717	C4 better

C.3 Latent-use controls

The correct condition supplies each input's posterior mean; shuffle permutes latent assignments across inputs; zero supplies a zero vector; and prior supplies a prior sample. Every condition uses the same 1,000 inputs. T is Morgan Tanimoto similarity between input and output (radius 2, 2,048 bits), and ΔT_{latent} is the difference in mean Tanimoto between correct and shuffle. We use saved aggregate values, not differences between rounded displays.

The random-pair Tanimoto floor was approximately 0.1355. Shuffle and prior were near this floor, zero was also far below correct, and exact reconstruction was 0 for every control. Thus, reconstruction under the correct condition depends on molecule–latent pairing.

Table 8: 1k latent-use controls. Validity is 1.000 for every model and condition. Exact reconstruction is 0 for all shuffle / zero / prior controls.

Model	Exact	Correct T	Shuffle T	Zero T	Prior T	ΔT
GX-TGVAE	0.858	0.933	0.136	0.169	0.135	0.7977
MAGNet-R	0.774	0.914	0.134	0.174	0.139	0.7798
MAGNet-P	0.630	0.848	0.134	0.153	0.140	0.7142
MoLeR	0.180	0.542	0.135	0.153	0.133	0.4062

C.4 10k parity and paired comparisons

Table 9 reports results on the same 10,000 canonical inputs. All 4 models are evaluated on 10,000 inputs, and each of the 3 GX-TGVAE-versus-baseline comparisons uses 10,000 pairs.

Table 9: Latent fidelity on 10,000 canonical inputs. MAGNet uses the VAE alone.

Model	Exact	Correct T	Shuffle T	ΔT
GX-TGVAE	0.860	0.935	0.135	0.7998
MAGNet-R	0.783	0.918	0.135	0.7838
MAGNet-P	0.619	0.842	0.134	0.7085
MoLeR	0.176	0.536	0.135	0.4010

For McNemar's test (McNemar, 1947), let b count inputs reconstructed exactly only by GX-TGVAE and c those reconstructed exactly only by the baseline. We use the two-sided exact test based on the binomial distribution of discordant pairs (Fagerland et al., 2013). The 95% CI for the Tanimoto difference uses the 2.5 / 97.5 percentiles of 10,000 paired bootstrap resamples at the input-pair level (Efron & Tibshirani, 1986) (resampling seed 42). These CIs describe input-set uncertainty for fixed trained models, not training-seed uncertainty.

Table 10: 10k paired comparison: McNemar discordant counts and paired-bootstrap mean Tanimoto differences.

Comparison	b/c	p	ΔT [95% CI]
GX-TGVAE − MAGNet-R	1683 / 912	2.25×10^{-52}	+0.0165 [0.0122, 0.0209]
GX-TGVAE − MAGNet-P	3066 / 659	$<10^{-300}$	+0.0923 [0.0873, 0.0974]
GX-TGVAE − MoLeR	7024 / 181	$<10^{-300}$	+0.3986 [0.3925, 0.4048]

The corresponding 1k comparison gives GX-TGVAE minus MAGNet-R: $b/c = 176/92$, $p = 3.21 \times 10^{-7}$, mean Tanimoto difference $+0.0192$ (95% CI [0.0053, 0.0335]); GX minus MAGNet-P: 297/69, $p = 8.11 \times 10^{-35}$, $+0.0851$ ([0.0690, 0.1014]); and GX minus MoLeR: 693/15, $p = 5.63 \times 10^{-183}$, $+0.3918$ ([0.3729, 0.4107]). The 1k and 10k comparisons are not independent replications.

C.5 Scaffold-frequency stratification

We stratified the 10k inputs by Bemis–Murcko scaffold (Bemis & Murcko, 1996) frequency in the effective MOSES training split. Acyclic molecules share a common label in the analysis implementation. Counts across the 6 strata in Table 11 sum to 10,000. The 1,569 inputs with frequency 0 have scaffolds absent from the training split. Table SEs are $\sqrt{p(1-p)/n}$, not variation across seeds.

The overall GX-TGVAE minus MAGNet-R exact-reconstruction difference is approximately 0.077, and approximately 0.096 for held-out scaffolds, indicating that GX-TGVAE retains latent-to-structure fidelity in the held-out condition. Spearman correlations (Spearman, 1904) between frequency and exact reconstruction are small for all models ($|\rho|$ approximately 0.04 or less), although GX-TGVAE

Table 11: All scaffold-frequency strata in the 10k set. Errors for exact reconstruction are binomial SEs, not seed SDs.

Train freq.	n	GX-TGVAE	MAGNet-R	MAGNet-P	MoLeR
0 (held-out)	1569	0.820 ± 0.010	0.724 ± 0.011	0.547 ± 0.013	0.161 ± 0.009
1-2	1070	0.856 ± 0.011	0.782 ± 0.013	0.601 ± 0.015	0.192 ± 0.012
3-10	1425	0.853 ± 0.009	0.779 ± 0.011	0.612 ± 0.013	0.178 ± 0.010
11-100	2290	0.881 ± 0.007	0.803 ± 0.008	0.670 ± 0.010	0.183 ± 0.008
101-1000	1963	0.877 ± 0.007	0.824 ± 0.009	0.664 ± 0.011	0.178 ± 0.009
1001+	1683	0.856 ± 0.009	0.765 ± 0.010	0.582 ± 0.012	0.164 ± 0.009

also has a statistically significant correlation at 10k. We therefore do not claim independence from scaffold frequency. Table 12 displays values computed from unrounded results.

Table 12: Spearman correlation between $\log(1+\text{scaffold frequency})$ and exact reconstruction. Small correlations do not establish independence.

Model	ρ	p
GX-TGVAE	0.03214	0.001307
MAGNet-R	0.04006	0.000062
MAGNet-P	0.03711	0.000205
MoLeR	-0.00340	0.734178

C.6 Molecular-size dependence

Table 13 reports correlations with heavy-atom count. For GX-TGVAE, the correlation remains near 0 at 10k, while the other models have weak negative correlations. This is a descriptive analysis, not a preregistered primary test.

Table 13: Spearman correlation between heavy-atom count and exact reconstruction (10k).

Model	ρ	p
GX-TGVAE	-0.009	0.359
MAGNet-R	-0.079	3.4×10^{-15}
MAGNet-P	-0.068	1.0×10^{-11}
MoLeR	-0.064	2.1×10^{-10}

C.7 Latent activity diagnostics

Based on marginal KL in the final MOSES window (threshold 0.01), the active-dimension counts are 48/256 for C2_K8 and 49/256 for C4_K8. Active-dimension count alone does not explain the latent-fidelity difference between scopes. This MOSES diagnostic and the 41 active coordinates fixed for ZINC exploration in Appendix E.1 concern different models and datasets. Active count is not a headline performance metric.

C.8 Independent-test latent-fidelity robustness

We repeated the reconstruction and latent-use evaluation of Appendix C.4 on 10,000 fixed molecules from the MOSES official test set, unused for training or model selection and with no canonical graph overlap with training. All 4 models received the same inputs, with fixed checkpoints, no reselection, and unchanged canonicalization, greedy decoding, and metric definitions.

The 3 out-of-vocabulary inputs for MAGNet-R / MAGNet-P remain in the Exact denominator of 10,000 as reconstruction failures. Mean Tanimoto is evaluated on 9,997 inputs with valid outputs for

Table 14: Latent fidelity on 10,000 independent test molecules. ΔT is Correct T minus Shuffle T before rounding. The Exact denominator is 10,000 for every model.

Model	Exact	Correct T	Shuffle T	ΔT
GX-TGVAE	0.8605	0.9336	0.1345	0.7991
MAGNet-R	0.7987	0.9225	0.1341	0.7883
MAGNet-P	0.6284	0.8464	0.1336	0.7128
MoLeR	0.1725	0.5347	0.1343	0.4004

these models and on 10,000 for GX-TGVAE / MoLeR. These denominators apply to correct / shuffle / zero / prior within each model.

Both Exact and Correct T preserve the internal-10k ranking: GX-TGVAE, MAGNet-R, MAGNet-P, then MoLeR. All models reproduce a large Tanimoto drop from correct to shuffled latents, with zero exact reconstruction under shuffle / zero / prior. This confirms dependence on correct molecule–latent pairing for the fixed models on independent inputs; it does not evaluate training-seed variability.

D Distributional and Structural Generation

D.1 MOSES generation protocol

We evaluate 30,000 generated molecules from each model's fixed checkpoint with generation seeds 42 / 123 / 2024. Tables 15 and 16 report 14 selected standard MOSES metrics (Polykovskiy et al., 2020). We separate distributional metrics, including FCD (Preuer et al., 2018), from our auxiliary scaffold diagnostics defined below. The $\pm$ values are sample SDs over 3 generation seeds, not training-seed uncertainty.

GraphXForm-P denotes the paper-aligned configuration (8 heads, dropout 0.1, epoch 500, 1.5M batches, sampling temperature 1.0). GX-TGVAE uses greedy decoding from C2_K8 SWA, and TGVAE uses greedy decoding at epoch 80. MAGNet-P and MAGNet-R denote MAGNet (paper architecture) and MAGNet (repository-default architecture), respectively; their prior generation includes post-hoc Flow.

Table 15: Standard MOSES metrics (a). The 14 reported metrics use means $\pm$ sample SDs over 30k samples across 3 generation seeds. Custom scaffold metrics are separated in Table 17.

Metric	GX-TGVAE	GraphXForm-P	MoLeR
valid	1.0000 ± 0.0000	1.0000 ± 0.0000	1.0000 ± 0.0000
unique@10k	1.0000 ± 0.0001	0.9991 ± 0.0002	0.9997 ± 0.0001
FCD/Test	0.9832 ± 0.0029	0.0934 ± 0.0008	0.8106 ± 0.0128
FCD/TestSF	1.3009 ± 0.0088	0.5472 ± 0.0004	1.4052 ± 0.0075
SNN/Test	0.4965 ± 0.0001	0.6066 ± 0.0008	0.5176 ± 0.0000
SNN/TestSF	0.4806 ± 0.0003	0.5698 ± 0.0004	0.4967 ± 0.0005
Frag/Test	0.9978 ± 0.0001	0.9995 ± 0.0000	0.9870 ± 0.0004
Frag/TestSF	0.9968 ± 0.0000	0.9981 ± 0.0000	0.9863 ± 0.0005
Scaf/Test	0.8049 ± 0.0062	0.9253 ± 0.0010	0.7973 ± 0.0112
Scaf/TestSF	0.1666 ± 0.0063	0.1161 ± 0.0028	0.1568 ± 0.0104
IntDiv	0.8562 ± 0.0001	0.8570 ± 0.0003	0.8584 ± 0.0002
IntDiv2	0.8511 ± 0.0001	0.8509 ± 0.0002	0.8524 ± 0.0002
Filters	0.9595 ± 0.0023	0.9916 ± 0.0004	0.9674 ± 0.0015
Novelty	0.9790 ± 0.0008	0.8329 ± 0.0021	0.9691 ± 0.0009

Auxiliary scaffold diagnostics. The following 3 metrics are auxiliary diagnostics introduced in this study, not official MOSES metrics. They are computed directly from generated SMILES and training scaffold sets, rather than by transforming aggregate MOSES values. RDKit extracts each molecule's Bemis–Murcko scaffold (Bemis & Murcko, 1996), whose canonical

Table 16: Standard MOSES metrics (b). The 14 reported metrics use means ± sample SDs over 30k samples across 3 generation seeds. Custom scaffold metrics are separated in Table 17.

Metric	MAGNet-P+Flow	MAGNet-R+Flow	TGVAE
valid	1.0000 ± 0.0000	1.0000 ± 0.0000	0.8640 ± 0.0023
unique@10k	0.9997 ± 0.0002	0.9997 ± 0.0002	1.0000 ± 0.0001
FCD/Test	1.5653 ± 0.0297	1.4990 ± 0.0133	1.4092 ± 0.0080
FCD/TestSF	2.1682 ± 0.0354	2.0946 ± 0.0182	1.7218 ± 0.0135
SNN/Test	0.4900 ± 0.0006	0.4992 ± 0.0007	0.4804 ± 0.0001
SNN/TestSF	0.4728 ± 0.0003	0.4816 ± 0.0002	0.4669 ± 0.0003
Frag/Test	0.9966 ± 0.0002	0.9961 ± 0.0001	0.9949 ± 0.0001
Frag/TestSF	0.9937 ± 0.0003	0.9934 ± 0.0002	0.9934 ± 0.0001
Scaf/Test	0.8209 ± 0.0042	0.8400 ± 0.0071	0.6795 ± 0.0124
Scaf/TestSF	0.1113 ± 0.0049	0.1049 ± 0.0158	0.2261 ± 0.0047
IntDiv	0.8629 ± 0.0004	0.8605 ± 0.0003	0.8504 ± 0.0001
IntDiv2	0.8569 ± 0.0004	0.8542 ± 0.0003	0.8455 ± 0.0001
Filters	0.9772 ± 0.0004	0.9814 ± 0.0004	0.9760 ± 0.0005
Novelty	0.9784 ± 0.0005	0.9743 ± 0.0008	0.9911 ± 0.0004

SMILES defines its identity. Specifically, the output of `GetScaffoldForMol` is represented by `MolToSmiles(canonical=True)`. These diagnostics add no genericization, custom stereochemistry removal, or largest-fragment selection.

Index every molecule occurrence in the generated file. Let I contain occurrences that can be parsed and have a scaffold extracted, and let $N_{\text{scaf}} = |I|$. **Duplicate molecules remain separate occurrences in the denominator.** Empty scaffolds of acyclic molecules are retained as a common `[ACYCLIC]` category. Only parsing or extraction failures are excluded from I. With scaffold identity s_i, define the set of unique generated scaffolds as

$$\mathcal{S}_{\text{gen}} = \{s_i : i \in I\}$$

The training set $\mathcal{S}_{\text{train}}$ uses the same extraction and acyclic conventions. The reference for this table is **all 1,584,663 molecules in the official MOSES training set**, unlike the effective training split used for scaffold-frequency analysis in C and structural reach in D.4.

For $N_{\text{scaf}} > 0$, define

$$\text{ScafNovMol} = \frac{1}{N_{\text{scaf}}} \sum_{i \in I} \mathbf{1}[s_i \notin \mathcal{S}_{\text{train}}],$$

$$\text{ScafNovUniq} = \frac{|\mathcal{S}_{\text{gen}} \setminus \mathcal{S}_{\text{train}}|}{|\mathcal{S}_{\text{gen}}|}, \qquad \text{ScafUniq} = \frac{|\mathcal{S}_{\text{gen}}|}{N_{\text{scaf}}}$$

These are molecule-weighted scaffold novelty, unique-scaffold novelty, and scaffold uniqueness, respectively; the implemented `scaffold_uniqueness` corresponds to ScafUniq. When $N_{\text{scaf}} = 0$, all 3 metrics are undefined and are not imputed as 0.

Metrics are computed per seed, followed by the mean and sample SD over 3 seeds. In Table 17, N_{scaf} is 30,000 for each seed for GX-TGVAE, GraphXForm-P, MoLeR, and both MAGNet configurations. For TGVAE, it is 25,841 / 25,969 / 25,947 for seeds 42 / 123 / 2024, respectively. The denominator is thus not necessarily the full 30k attempts.

These training-scaffold diagnostics are auxiliary and are not headline results in the main text. Scaffold-novelty fractions differ from nearest-neighbor distance, distant-tail counts, and L3 unique generic topology counts in D.4. Retaining acyclic molecules as one category here also differs from excluding outputs with undefined scaffold distance in D.4.

Table 17: Our auxiliary scaffold diagnostics, separate from standard MOSES metrics and headline structural reach. Mean ± sample SD over 30k samples across 3 seeds.

Model	ScafNovMol	ScafNovUniq	ScafUniq
GX-TGVAE	0.5942 ± 0.0010	0.7419 ± 0.0024	0.7729 ± 0.0015
GraphXForm-P	0.3091 ± 0.0015	0.4588 ± 0.0013	0.6335 ± 0.0028
MoLeR	0.4803 ± 0.0017	0.6956 ± 0.0024	0.6524 ± 0.0020
MAGNet-P+Flow	0.5636 ± 0.0026	0.7524 ± 0.0028	0.6861 ± 0.0028
MAGNet-R+Flow	0.5386 ± 0.0041	0.7476 ± 0.0018	0.6552 ± 0.0039
TGVAE	0.5928 ± 0.0012	0.7491 ± 0.0014	0.7598 ± 0.0018

D.2 Checkpoint trajectories

In the checkpoint diagnostic, GraphXForm's FCD/Test changes from 0.2898 at epoch 200 to 0.1851 at the best checkpoint, while Novelty changes from 0.8796 to 0.7840. This series uses 10,000 samples at temperature 1.0 from the official-native configuration, distinct from the paper-aligned 30k by 3-seed setting in Table 15.

For TGVAE with 12,000 samples and greedy decoding, FCD/Test improves from 2.4530 at epoch 20 to 1.4905 at epoch 80, while Novelty remains nearly unchanged at 0.9922 versus 0.9912. These within-model trajectories are auxiliary observations of how distribution fitting affects novelty in a model-dependent manner, not a universal FCD–Novelty trade-off law. Series with different sample counts are not used as matched inter-model performance comparisons.

D.3 MAGNet raw and Flow priors

The MAGNet (Hetzel et al., 2025) prior-generation baseline uses the final pipeline including trained post-hoc Flow. Flow aligns the prior sampling distribution and is not inserted into the reconstruction path for an input molecule. Table 18 illustrates why this distinction matters.

Table 18: MAGNet raw prior versus post-hoc Flow prior.

Variant	FCD raw	FCD +Flow	Scaf raw	Scaf +Flow
MAGNet-P	5.980	1.565	0.469	0.821
MAGNet-R	4.288	1.499	0.656	0.840

MOSES validity can accept disconnected SMILES. MAGNet's MOSES validity is 1.000, whereas the connected-validity sidecar is approximately 0.873–0.878 for the paper architecture and 0.882–0.884 for the repository-default architecture. Appendix H.5 discusses validity semantics and implementation.

D.4 Structural reach over three generation seeds

We use results from a common scaffold pipeline applied to the existing 30k generations across 3 seeds for final GX-TGVAE, MoLeR, and both MAGNet configurations. Distance is 1 minus maximum Morgan Tanimoto similarity between a generated molecule's Bemis–Murcko scaffold and effective-training-split scaffolds (radius 2, 2,048 bits). The training cache contains 375,026 BM scaffold identities; fingerprintable scaffolds serve as the distance reference. This is scaffold distance, not whole-molecule fingerprint distance. Excluding acyclic and other outputs with undefined scaffold distance gives denominators of 29,941–29,986.

L3 counts *unique topologies* absent from the 64,404 training topologies after atom types and bond orders are abstracted with RDKit's `MakeScaffoldGeneric`. The distant tails $d > 0.5$ and $d > 0.7$ count *molecules*. We preserve these different units and do not divide L3 by the distance-analysis denominator to form a fraction. Table 19 gives aggregates, and Tables 20 and 21 give seed-specific values for all 12 sets.

Table 19: Structural reach: mean ± sample SD over 3 seeds. Both MAGNet configurations include Flow.

Metric	GX-TGVAE	MoLeR	MAGNet-P	MAGNet-R
Mean distance	0.2278 ± 0.0001	0.1685 ± 0.0008	0.2067 ± 0.0004	0.1936 ± 0.0015
p90	0.5142 ± 0.0012	0.4544 ± 0.0019	0.5000 ± 0.0000	0.4897 ± 0.0003
p99	0.6925 ± 0.0030	0.6064 ± 0.0056	0.6837 ± 0.0017	0.6621 ± 0.0040
L3	3,969.0 ± 72.7	2,728.0 ± 51.4	2,669.0 ± 48.7	2,686.7 ± 69.0
L3, ring-excluded	3,524.3 ± 49.7	2,728.0 ± 51.4	2,669.0 ± 48.7	2,686.7 ± 69.0
d>0.5	3,296.3 ± 14.0	1,621.0 ± 51.1	2,848.7 ± 24.2	2,508.7 ± 36.2
d>0.7	266.0 ± 13.5	49.3 ± 8.1	229.3 ± 1.5	161.3 ± 12.1

Table 20: Seed-specific distance statistics for 12 generation sets. Distance n counts molecules with defined scaffold distance.

Model	Seed	Distance n	Mean	p90	p99
GX-TGVAE	42	29984	0.2279	0.5146	0.6957
GX-TGVAE	123	29986	0.2278	0.5128	0.6897
GX-TGVAE	2024	29984	0.2277	0.5152	0.6923
MoLeR	42	29946	0.1690	0.4545	0.6093
MoLeR	123	29946	0.1675	0.4524	0.6000
MoLeR	2024	29941	0.1689	0.4561	0.6100
MAGNet-P	42	29977	0.2066	0.5000	0.6857
MAGNet-P	123	29972	0.2065	0.5000	0.6825
MAGNet-P	2024	29974	0.2071	0.5000	0.6829
MAGNet-R	42	29977	0.1922	0.4894	0.6667
MAGNet-R	123	29965	0.1932	0.4898	0.6596
MAGNet-R	2024	29971	0.1953	0.4898	0.6600

Table 21: Seed-specific topology and distant-molecule counts for 12 generation sets. Both MAGNet configurations include Flow.

Model	Seed	L3	L3 ring-excluded	d>0.5	d>0.7
GX-TGVAE	42	4035	3571	3310	281
GX-TGVAE	123	3891	3472	3282	255
GX-TGVAE	2024	3981	3530	3297	262
MoLeR	42	2746	2746	1617	53
MoLeR	123	2670	2670	1572	40
MoLeR	2024	2768	2768	1674	55
MAGNet-P	42	2664	2664	2830	229
MAGNet-P	123	2720	2720	2840	231
MAGNet-P	2024	2623	2623	2876	228
MAGNet-R	42	2608	2608	2527	175
MAGNet-R	123	2737	2737	2532	157
MAGNet-R	2024	2715	2715	2467	152

GX-TGVAE exceeds the compared baselines in all 21 observations across 7 metrics and 3 seeds. Baseline L3 rankings vary by seed, so we do not rank MoLeR and the two MAGNet configurations against one another. The 3 generation seeds do not establish general statistical superiority that accounts for training-seed variation.

D.5 Large-ring exclusion

After excluding molecules with maximum ring size at least 8, GX-TGVAE's seed-averaged count of novel generic topologies decreases as

$$3969.0 \rightarrow 3524.3$$

but remains approximately 1.29 times the largest baseline value, 2728.0. All 3 baseline models have 0 large-ring molecules across all 3 seeds, so their L3 counts are unchanged by exclusion. This control does not support attributing GX-TGVAE's structural-reach difference solely to large-ring generation; it does not exclude all other artifacts.

D.6 Shape-vocabulary OOV

We decomposed generated molecules using MAGNet's training-derived 381-shape vocabulary. OOV in Table 22 is the fraction of molecules containing at least one shape outside that vocabulary. We follow the implementation's term "shape"; Appendix H.2 relates it to the scaffold representation in the final MAGNet paper.

Table 22: OOV relative to the 381-shape vocabulary. Errors are sample SDs over 3 generation seeds. Older GX-TGVAE and W-D are historical controls.

Set	Analyzed	OOV fraction
Final GX-TGVAE, 3 seeds	30,000 / seed	0.0169 ± 0.0007
older GX-TGVAE champion	30,000 / seed	0.0237 ± 0.0007
MoLeR	30,000 / seed	0.0014 ± 0.0001
W-D λ=2 + Flow	30,000	0.0652
MOSES train control	30,000	0.0000
MOSES test control	10,000	0.0004

Final GX-TGVAE has an OOV fraction of 1.69%. OOV molecule fraction and L3 topology count have different units, so we do not infer a causal contribution of OOV to structural reach. The older GX champion and W-D $\lambda = 2$+Flow are historical controls, not new contributions in the main text.

For MAGNet's own seed 42 outputs, the paper architecture has 3,573 decomposition failures out of 30,000 molecules, leaving a denominator of 26,427 and 0 OOV molecules. The repository-default architecture has 3,351 failures, a denominator of 26,649, and 2 OOV molecules. Because these denominators differ from those of other models, these OOV rates serve as vocabulary controls rather than a model ranking.

E Geometry and Local Extrapolation

E.1 Active latent space and local PCA

The exploration anchor is

$$z_0 = \mu_\phi(G_0)$$

A dense latent bank is constructed from training-molecule posterior means. Coordinates whose mean marginal posterior KL exceeds a threshold are selected as active dimensions, with orthogonal projection P_A. The final ZINC protocol fixes the mask of 41 active coordinates obtained at threshold 0.01 in an existing precheck of 10,000 molecules. The dense bank for local-neighborhood search differs from the set used to estimate this active mask.

We obtain the kNNs of the current position z using Euclidean distance in the original latent coordinates, center the neighbors, and apply PCA. Let $U(z)$ span the principal subspace explaining 90% of local variance. Projection into the active space and reorthogonalization yield

$$U_A(z) = \text{orth}(P_A U(z))$$

The tangent and normal projectors are

$$P_T = U_A U_A^\top, \qquad P_N = P_A - P_T. \tag{E1}$$

The SVD of the centered neighborhood matrix and reorthogonalization after active projection use float64; the latter retains components with singular values above 10^{-8}. This is a local-PCA estimator, not a direct observation of the true data-manifold tangent or normal space (Kaslovsky & Meyer, 2014). Let P_A be an orthogonal coordinate projector in the ambient latent space, and let the nonzero columns of U_A form an orthonormal basis within its range. Then P_T and P_N orthogonally decompose the active space. Zero-padding columns in the implementation do not contribute to either projector.

We refer to Aamari and Levrard for background on local tangent estimation and local manifold regularity (Aamari & Levrard, 2019). Their estimator, assumptions, and guarantees do not transfer directly to our bank-based PCA. We derive the conditional geometric relationships below.

E.2 First-order manifold departure

Proposition E.1 (Local normal displacement). Let $\mathcal{M} \subset \mathbb{R}^{d_z}$ be a C^2 embedded manifold and $z \in \mathcal{M}$. Decompose a unit vector v as

$$v = v_T + v_N, \qquad v_T \in T_z\mathcal{M}, \quad v_N \in N_z\mathcal{M}, \quad \rho = \|v_N\|$$

If the second fundamental form / curvature is bounded near z, there exists a constant $C_\mathcal{M} > 0$ such that, for sufficiently small $|\alpha|$,

$$\boxed{\text{dist}(z + \alpha v, \mathcal{M}) \geq \rho|\alpha| - C_\mathcal{M}\alpha^2} \tag{E2}$$

In particular, for a pure normal direction, $\rho = 1$,

$$\text{dist}(z + \alpha v_N, \mathcal{M}) = |\alpha| + O(\alpha^2), \tag{E3}$$

whereas for a pure tangent direction,

$$\text{dist}(z + \alpha v_T, \mathcal{M}) = O(\alpha^2). \tag{E4}$$

Proof. Choose local orthonormal coordinates with $z = 0$ and tangent space $\mathbb{R}^q \times \{0\}$. Locally, the C^2 manifold is the graph

$$\mathcal{M} = \{(u, h(u)) : u \in \mathbb{R}^q\}$$

with $h(0) = 0$ and $Dh(0) = 0$, hence

$$\|h(u)\| \leq C\|u\|^2$$

Comparing the normal components of $\alpha v = (\alpha v_T, \alpha v_N)$ and a manifold point $(u, h(u))$, the manifold's normal displacement is second order, while the query's normal component is first order with length $\rho|\alpha|$. To turn this into a distance lower bound, let $a = |\alpha|$ and consider the local chart containing z. Since $z \in \mathcal{M}$, distance is at most a, and points farther than $2a$ from the query can be excluded when proving the lower bound. For all remaining points $(u, h(u))$, $\|u\| \leq 3a$, so

$$\|(\alpha v_T, \alpha v_N) - (u, h(u))\| \geq \|\alpha v_N - h(u)\| \geq \rho a - 9Ca^2.$$

For sufficiently small a, all such nearby points lie in this chart. Taking the infimum gives Equation (E2). For a pure normal direction, the upper bound a and the lower bound imply Equation (E3). For a pure tangent direction, comparison with $(\alpha v_T, h(\alpha v_T)) \in \mathcal{M}$ bounds distance by $O(\alpha^2)$; together with nonnegativity, this gives Equation (E4).

This proposition states the geometric fact that a normal component produces first-order departure from the local manifold. It does not directly guarantee anything about global density or molecular novelty.

E.3 Estimated normal and principal-angle error

The implemented $\hat{v}_N$ is a unit estimated-normal direction obtained from local PCA. Define its true-normal projection ratio as

$$\rho = \|P_{N_z\mathcal{M}} \hat{v}_N\|$$

Then Proposition E.1 directly gives

$$\text{dist}(z + \alpha \hat{v}_N, \mathcal{M}) \geq \rho|\alpha| - C\alpha^2 \tag{E5}$$

We state the conditions explicitly. Let $\widehat{N} = \text{range}(P_N)$ be the estimated normal subspace and $N = N_z\mathcal{M}$ the true normal subspace. Assume they lie in the same ambient Euclidean space and $1 \leq r = \dim \widehat{N} \leq \dim N$. Let $Q_{\widehat{N}}$ and Q_N be orthonormal bases, and define the largest principal angle $\theta \in [0, \pi/2]$ by

$$\cos\theta = \sigma_{\min}(Q_N^\top Q_{\widehat{N}}).$$

For any unit vector $\hat{v}_N = Q_{\widehat{N}} b$ with $\|b\| = 1$,

$$\rho = \|P_N^{\text{true}} \hat{v}_N\| = \|Q_N^\top Q_{\widehat{N}} b\| \geq \sigma_{\min}(Q_N^\top Q_{\widehat{N}}) = \cos\theta, \qquad P_N^{\text{true}} = Q_N Q_N^\top.$$

Equivalently, one can use the one-sided projector-error condition $\|(I - P_N^{\text{true}})P_N\|_2 \leq \sin\theta$ from the estimated-normal space. Substituting into Equation (E5) retains a positive first-order lower bound when $\cos\theta > 0$. If $r > \dim N$, an estimated direction may vanish under true-normal projection, so we do not claim a positive lower bound for all directions in that case.

We cannot observe the true manifold and therefore do not estimate θ. Equation (E5) serves as a **conditional theoretical link for interpreting the implementation**.

E.4 Local Mahalanobis interpretation

Write the local covariance as

$$\Sigma_z = U\Lambda U^\top, \qquad \Lambda = \text{diag}(\lambda_1, \ldots, \lambda_d)$$

where U is a complete orthonormal eigenbasis of the ambient space and $\lambda_i \geq 0$. To address rank deficiency and numerical instability, take $\varepsilon > 0$ and define the regularized local Mahalanobis distance by

$$d^2_{M,\varepsilon}(\delta) = \delta^\top (\Sigma_z + \varepsilon I)^{-1} \delta \tag{E6}$$

Writing $c_i = u_i^\top \delta$ in the PCA basis gives

$$\boxed{d^2_{M,\varepsilon}(\delta) = \sum_i \frac{c_i^2}{\lambda_i + \varepsilon}} \tag{E7}$$

Thus, a step of Euclidean length $|\alpha|$ along PCA eigendirection u_i has

$$d^2_{M,\varepsilon}(\alpha u_i) = \frac{\alpha^2}{\lambda_i + \varepsilon}. \tag{E8}$$

For equal Euclidean displacement, lower-variance directions have greater standardized local distance. Interpreting this as statistical rarity, or low local density, requires additional distributional assumptions. Examples are a Gaussian approximation with δ measured from its center, or an elliptical approximation whose density decreases with Mahalanobis radius. A large quadratic form for a step from an arbitrary starting point does not alone imply lower density at its endpoint. Because GX-TGVAE's P_N is the orthogonal complement of the retained high-variance tangent PCs within the active space, normal-like movement can be interpreted as emphasizing locally low-variance components.

The basis of P_N need not coincide exactly with the covariance's bottom eigenvectors. Equation (E7) specifies the weighting of a general displacement expanded in the PCA basis. This Mahalanobis quadratic form is interpretive; it is not used to set final exploration step sizes or determine molecular OOD. The main text therefore describes **the relationship between subspace decomposition by retained PCA components and inverse-variance weighting in Mahalanobis distance**, rather than equating normal directions with Mahalanobis distance.

E.5 Decoder non-collapse as a conditional statement

Geometric departure in latent space does not constitute molecular exploration unless it changes the decoded chemistry. Since discrete graphs are not readily differentiable, consider the expectation of continuous molecular features $\phi(G) \in \mathbb{R}^k$:

$$F(z) = \mathbb{E}_{G \sim p_\theta(G|z)}[\phi(G)] \tag{E9}$$

Proposition E.2 (Directional decoder response). Suppose F is C^2 near z, a unit direction v satisfies

$$\|J_F(z)v\| \geq m > 0$$

and the Jacobian is locally Lipschitz:

$$\|J_F(z+u) - J_F(z)\| \leq H\|u\|$$

Then, for sufficiently small $|\alpha|$,

$$\boxed{\|F(z+\alpha v) - F(z)\| \geq m|\alpha| - \frac{H}{2}\alpha^2} \tag{E10}$$

Proof. The fundamental theorem of calculus gives

$$F(z+\alpha v) - F(z) = \int_0^\alpha J_F(z+tv)v\, dt.$$

Adding and subtracting $J_F(z)v$, then applying the triangle inequality and the Jacobian Lipschitz bound, yields

$$\|F(z+\alpha v) - F(z) - \alpha J_F(z)v\| \le \int_0^{|\alpha|} Ht\,dt = \frac{H}{2}\alpha^2.$$

The reverse triangle inequality then gives Equation (E10).

This proposition is a **lower bound conditional on a non-collapse assumption**. We have not established a uniform $m > 0$ across all active-normal directions. A Lipschitz upper bound alone also permits a constant decoder, so the lower-sensitivity condition matters when explaining extrapolation.

We measure final molecular OOD empirically using the decoded molecule's maximum Morgan Tanimoto similarity to the ZINC training split used in this study, not Equation (E10). A lower bound on changes in expected features does not guarantee changes in individual decoded molecules or OOD reach. Equation (7) in the main text corresponds to $C_F = H/2$.

E.6 Geometry → statistical rarity → molecular OOD

We thus distinguish three stages:

first-order local manifold departure → local statistical rarity → empirical molecular OOD

- First: a conditional geometric explanation based on a manifold hypothesis and a normal component.
- Second: standardized local distance from local covariance / Mahalanobis geometry, interpreted as statistical rarity under additional distributional assumptions.
- Third: empirical observations from Tanimoto / scaffold metrics after decoding.

This diagram is not a chain of unconditional implications. The first stage requires manifold and direction assumptions, the second requires distributional assumptions, and the third is checked independently on generated molecules. Conditional decoder response does not identify these three concepts with one another.

E.7 Dense bank, local scale, and update rule

The dense bank comprises posterior means from the internal training pool. The final ZINC lock uses 203,653 encodable rows out of 203,654 inputs, yielding 202,397 unique bank entries indexed by cached molecule keys. The main protocol uses $k = 50$ neighbors and PCA retained variance 0.90. Initial localization excludes the anchor's own bank entry by row index, not by a small-distance criterion. The same index is no longer excluded after the first step.

The reference scale s is fixed by calibration on the first 300 anchors of seed 0, excluding self-entries and using the local median distance to 50 nearest neighbors in the dense bank. The final protocol uses $s = s_{\text{ref}} = 3.828$; step sizes are not renormalized by the local scale measured at each step. The normal step is

$$\delta_N/s = 0.25$$

We evaluate $j \in \{2, 4, 8, 16\}$ out of 16 steps. Cumulative normal-phase path length is $\alpha_j = j\delta_N$, giving 4 checkpoints at $\alpha_j/s \in \{0.5, 1, 2, 4\}$. Because direction changes at each step, α_j is not the straight-line distance from the anchor. Neighbors, local PCA, and required gradients are re-estimated at each step. Differentiation includes input standardization inside the predictor; gradients and geometry use the same original latent coordinates.

TanNorm first takes $K_T = 5$ tangent steps with $\eta/s = 0.25$, then switches to the active-normal property direction.

E.8 Role of historical geometry diagnostics

Active-normal projection is motivated by avoiding movement along inactive coordinates and exploring within an active subspace that the decoder may use. Active-coordinate selection itself does not guarantee the directional lower sensitivity in Proposition E.2. Geometry from earlier pilots or different latent objectives is not mixed with results or method-selection evidence for the final protocol. Comparisons of alternative objectives such as WAE are not new contributions of this paper.

F Geometry-Guided OOD Exploration

F.1 Experimental unit

Using final GX-TGVAE trained on ZINC, we evaluated 5 protein targets (parp1, fa7, 5ht1b, braf, jak2), 5 generation seeds (0–4), and 3,000 anchors per target–seed cell. Anchors are sampled uniformly from the ZINC training split used in this study, with initial position $z_0 = \mu_\phi(G_0)$. Appendix H.1 describes split provenance and preprocessing.

The steering surrogate is a differentiable model that predicts a property score from a latent code. We use

$$g(z) = \nabla_z f_\psi(z)$$

Final evaluation uses the common QuickVina docking protocol (Alhossary et al., 2015), not surrogate predictions. The reference scale, local bank, step count, and progressive protocol follow Appendix E.

Steering surrogate and training. For each target we train a separate MLP with widths $256 \to 512 \to 512 \to 1$, GELU and dropout 0.1 after each hidden layer, and a linear scalar output. The input is the frozen ZINC encoder's posterior mean. The regression target is $y = \text{clip}(DS, 0, 20)\, QED\, SA_n/20$, using the positive docking score, QED, and already normalized SA stored in the distributed ZINC label table. Thus steering differentiates a composite property prediction, not a prediction of DS alone. The 5 target-specific predictors are shared across the GX-TGVAE direction policies.

We use a 95/5 training/internal-validation group split by cleaned canonical molecular identity, with seed 42. This predictor split is separate from the generator's internal holdout. Each latent coordinate is standardized inside the network using the mean and population standard deviation over distinct training latent vectors, with 10^{-6} added to the standard deviation. The target is not standardized. We minimize mean squared error with Adam, learning rate 10^{-3}, weight decay 10^{-4}, and batch size 1024, for at most 200 epochs. Internal-validation RMSE selects the checkpoint; early stopping uses patience 20 and a minimum RMSE improvement of 10^{-6}. The training seed is 42. The official test split is used only for reporting predictor metrics, not fitting or epoch selection.

During steering, the predictor is in evaluation mode (dropout disabled); differentiation of the scalar network output with respect to the raw latent input includes the internal standardization. Summing the independent batch outputs before differentiation gives each sample's own gradient. Final molecule evaluation uses the docking oracle, not this composite surrogate score.

Held-out predictive evaluation. Table 23 reports predictive accuracy on 23,398 held-out examples with no canonical molecular identity shared with predictor training or validation. Predictive accuracy on these in-distribution held-out molecules is not an OOD calibration guarantee.

Table 23: Target-specific surrogate accuracy on 23,398 overlap-free held-out examples. Epoch is selected on internal validation. Scores evaluate the composite regression target.

Target	Epoch	RMSE	MAE	Pearson	Spearman
parp1	63	0.03936	0.02812	0.7753	0.7953
fa7	64	0.03703	0.02606	0.7063	0.7419
5ht1b	41	0.04646	0.03568	0.6925	0.6938
braf	38	0.03982	0.02850	0.7704	0.7917
jak2	45	0.03808	0.02722	0.7647	0.7881

Neighborhood computation. At each progressive step, we compute exact Euclidean distances to the latent bank in batches of 256 queries and retrieve the 50 nearest neighbors. An available anchor-bank index is excluded. Centered-neighborhood SVD and reorthogonalization after active projection use float64; tangent dimension is selected by the 90% variance rule. This is an exact bank scan, not an approximate-nearest-neighbor index. The 16-step normal trajectory is run once and decoded at the 4 prescribed depths; TanNorm additionally has 5 tangent steps. Reported fixed-attempt budgets therefore do not imply equal wall-clock costs across methods.

F.2 Direction policies

Using the projectors P_A, P_T, P_N from Appendix E.1, we compare the directions

$$\begin{aligned} v_{ActRaw} &= \text{unit}(P_A g), \\ v_{ActNorm} &= \text{unit}(P_N g), \\ v_{RandAN} &= \text{unit}(P_N \epsilon), \quad \epsilon \sim \mathcal{N}(0, I), \\ v_T &= \text{unit}(P_T P_A g). \end{aligned}$$

ActRaw uses the active-space property gradient; ActNorm uses the active-normal property gradient. RandAN draws a new random active-normal direction at each step. TanNorm switches from a tangent phase to ActNorm.

We also used a persistent random active-normal policy, RandANP, in displacement diagnostics, but it is not a primary arm of the final 5 by 5 full-docking benchmark. The ActNorm–RandAN comparison supports direction selection relative to the tested random control, not a pure causal effect of the gradient alone.

F.3 Fixed-depth primary results

Tables 24 and 25 report the prespecified depths $\alpha/s \in \{0.5, 1, 2, 4\}$. The hit-ratio denominator is 3,000 attempts per target–seed cell. Pool gates and Top5 follow Appendix G.1. We average over 5 seeds within each target, then equally over 5 targets.

Table 24: Novel Hit Ratio (%). The denominator is 3,000 attempts per target–seed cell.

Arm	$\alpha/s = 0.5$	1	2	4
ActRaw	2.264	2.700	2.621	2.341
ActNorm	2.283	2.673	3.512	3.615
TanNorm	3.308	3.199	3.239	3.052
RandAN	2.203	2.324	2.751	3.205

Table 25: Novel Top 5% DS, averaged over seeds within targets, then over 5 targets.

Arm	$\alpha/s = 0.5$	1	2	4
ActRaw	9.301	9.457	9.519	9.384
ActNorm	9.314	9.435	9.666	9.780
TanNorm	9.722	9.729	9.813	9.750
RandAN	9.316	9.334	9.429	9.534

F.4 Direction control

Table 26 gives paired ActNorm minus RandAN differences in Novel Top 5% DS for the same target and seed. Differences are positive in 25/25 pairs at $\alpha/s = 2$ and 24/25 at $\alpha/s = 4$. We interpret this as direction-selection evidence on an upper-tail benchmark at deeper settings.

Tables 27 and 28 report existing results without gates or top-k selection: molecule-level median DS is computed within each target–seed similarity band before taking ActNorm minus RandAN.

Table 26: Target–seed paired differences, ActNorm minus RandAN. Positive counts positive differences among 25 pairs.

α/s	Δ Top5 DS	Pair SD	Positive
0.5	−0.002	0.074	13/25
1.0	+0.101	0.105	19/25
2.0	+0.237	0.098	25/25
4.0	+0.246	0.156	24/25

At $\alpha/s = 2, 4$, the mean paired difference is positive in all 7 reported bands. This does not imply improvement for every pair or molecule. Pair SD is the sample SD of the 25 differences.

Table 27: Within-band median DS differences, ActNorm minus RandAN (a). Each band has 25 target–seed pairs.

α/s	Similarity band	Δ median DS	Pair SD	Positive
0.5	[0.20,0.30)	+0.160	0.166	19/25
0.5	[0.30,0.35)	+0.070	0.131	15/25
0.5	[0.35,0.40)	+0.090	0.096	17/25
0.5	[0.40,0.45)	+0.026	0.096	9/25
0.5	[0.45,0.50)	+0.052	0.081	15/25
0.5	[0.50,0.60)	+0.076	0.083	18/25
0.5	[0.60,1.01)	+0.010	0.050	5/25
1.0	[0.20,0.30)	+0.302	0.223	20/25
1.0	[0.30,0.35)	+0.258	0.195	23/25
1.0	[0.35,0.40)	+0.226	0.095	25/25
1.0	[0.40,0.45)	+0.114	0.104	19/25
1.0	[0.45,0.50)	+0.150	0.091	24/25
1.0	[0.50,0.60)	+0.142	0.081	23/25
1.0	[0.60,1.01)	+0.092	0.069	22/25

Table 28: Within-band median DS differences, ActNorm minus RandAN (b). Each band has 25 target–seed pairs.

α/s	Similarity band	Δ median DS	Pair SD	Positive
2.0	[0.20,0.30)	+0.470	0.266	24/25
2.0	[0.30,0.35)	+0.400	0.149	25/25
2.0	[0.35,0.40)	+0.396	0.118	25/25
2.0	[0.40,0.45)	+0.374	0.122	25/25
2.0	[0.45,0.50)	+0.330	0.138	25/25
2.0	[0.50,0.60)	+0.302	0.105	25/25
2.0	[0.60,1.01)	+0.180	0.088	24/25
4.0	[0.20,0.30)	+0.552	0.410	21/25
4.0	[0.30,0.35)	+0.610	0.219	25/25
4.0	[0.35,0.40)	+0.560	0.163	25/25
4.0	[0.40,0.45)	+0.564	0.149	25/25
4.0	[0.45,0.50)	+0.532	0.178	25/25
4.0	[0.50,0.60)	+0.470	0.161	25/25
4.0	[0.60,1.01)	+0.338	0.335	19/25

F.5 Anchor-relative improvement

We docked the same anchors under the common protocol and computed endpoint differences

$$\Delta DS_i = DS(G_i^{explore}) - DS(G_i^{anchor})$$

using attempt pairs for which both anchor and endpoint have valid docking scores. Unlike the primary fixed-attempt benchmark, this conditional comparison excludes missing pairs from its denominator.

Within each target, we pooled successful pairs across 5 seeds, computed the mean, median, and improved fraction, then averaged each statistic equally across 5 targets. The Median summary in Table 29 is neither the median of a pool across targets nor a seed average of target–seed medians. Valid pair counts for the reported conditions are 74,999 for TanNorm 2, 74,993 for TanNorm 4, 74,998 for ActRaw 4, 74,998 for ActNorm 4, and 75,000 for RandAN 4.

Table 29: Anchor-relative comparison. Mean, median, and improved fraction are computed within each target and then averaged equally over 5 targets.

Arm	α/s	Mean ΔDS	Median summary	Improved
TanNorm	2	+1.211	+1.28	79.7%
TanNorm	4	+1.189	+1.32	78.3%
ActRaw	4	+1.312	+1.42	81.7%
ActNorm	4	+0.852	+0.98	76.7%
RandAN	4	+0.375	+0.36	62.3%

At TanNorm $\alpha/s = 2$, mean improvement is positive for all 5 targets; the target average is $+1.211$, and the improved fraction is 79.7%. This paired evidence indicates that high-DS endpoints do more than simply inherit good anchor scores.

F.6 Depth and gate-free properties

Aggregate changes from prespecified TanNorm depth 0.5 to 4 are

$$\text{mean max train similarity} : 0.473 \to 0.388,$$

$$\text{mean DS} : 8.784 \to 8.913,$$

$$\text{QED} : 0.611 \to 0.389, \qquad SA_n : 0.787 \to 0.703$$

QED and SA follow established molecular property metrics (Bickerton et al., 2012; Ertl & Schuffenhauer, 2009). Here, $SA_n = (10 - SA)/9$ transforms SA so that larger values are better.

This aggregation groups trajectories by arm, target, seed, and canonical anchor key. Later values overwrite duplicate records for the same key and depth, and endpoints are pooled across all targets and seeds. DS averages valid scores; each property averages available values. TanNorm has 74,935 DS observations at depth 0.5 and 74,773 at depth 4, distinct from the attempt-pair denominators in Appendix F.5. Within the evaluated range, aggregate mean DS at the deeper endpoint does not decline relative to the shallower endpoint, while QED and the synthetic-accessibility proxy decrease. We do not claim monotonic improvement for individual molecules across intermediate depths.

Observed molecular validity for the GX-TGVAE exploration configurations is 1.000. This is an empirical observation for the evaluated settings, not a validity guarantee across the entire latent space from the constraint-masked decoder.

G MOOD Comparison and Reach–Utility Decomposition

G.1 Fixed-attempt benchmark

The MOOD-style benchmark (Lee et al., 2023) fixes $N = 3,000$ attempts for each configuration, target, and seed. The first occurrence of each canonical molecule key is treated as the unique candidate. The eligible pool contains those candidates with valid docking scores satisfying

$$\text{max Tanimoto}(G, D_{train}) < 0.4, \quad QED > 0.5, \quad SA < 5$$

where D_{train} is the ZINC training split used in this study. In the implemented normalized SA, this corresponds to $SA_n > 5/9$; all inequalities are strict.

For docking affinity a, let $DS = \max(-a, 0)$. A Novel Hit is an eligible candidate exceeding the target threshold: 10.0 for parp1, 8.5 for fa7, 8.7845 for 5ht1b, 10.3 for braf, and 9.1 for jak2.[1] Novel Hit Ratio divides the hit count by 3,000. Failed generation or docking, duplicates, and gate failures remain in the attempt denominator.

Novel Top 5% DS averages the top 150 eligible candidates (5% of the attempt budget). If the pool contains fewer than 150 candidates, all available candidates are used; if it is empty, the score is 0. We record the candidate count when the pool is insufficient. This specifies edge-case handling in our evaluation implementation, rather than claiming that the original MOOD paper specifies every branch.

G.2 Primary fixed-budget results

Table 30 reports MOOD results. GX-TGVAE's maxima within the comparison grid are 3.615% hit ratio for ActNorm at $\alpha/s = 4$ and Top5 DS of 9.813 for TanNorm at $\alpha/s = 2$; these are different operating points. MOOD at $\lambda = 0.04$ exceeds both maxima on their respective metrics.

Table 30: MOOD fixed-attempt benchmark with 3,000 attempts per target–seed cell.

λ	Hit ratio (%)	Top5 DS
0	5.333	10.194
0.01	7.784	10.721
0.04	9.120	10.950
0.09	4.915	9.638
0.16	0.043	7.114

G.3 Candidate yield and conditional hit efficiency

For each evaluation unit with $n_{pass} > 0$,

$$\frac{n_{hit}}{N} = \frac{n_{pass}}{N} \times \frac{n_{hit}}{n_{pass}} \tag{G1}$$

We call n_{pass}/N candidate yield and n_{hit}/n_{pass} conditional hit efficiency. If $n_{pass} = 0$, raw hit is 0 and conditional efficiency is undefined. An empty pool is not treated as an observed efficiency of 0; conditional aggregation uses nonempty pools.

Table 31 aggregates representative points over 5 targets. Since an average of products generally differs from a product of averages, the displayed means must not be multiplied to recompute raw hit rates. This is a descriptive decomposition locating most of the MOOD–TanNorm gap on the yield side.

Table 31: Candidate yield and conditional hit efficiency. Each column aggregates over targets and seeds. Multiplying the displayed means does not recompute raw hit rates.

Setting	Yield	Conditional	Product	Raw hit
MOOD 0.04	43.8%	21.4%	9.4%	9.120%
MOOD 0.01	45.5%	17.4%	7.9%	7.784%
ActNorm 4	18.2%	20.0%	3.6%	3.615%
TanNorm 2	9.3%	35.0%	3.3%	3.239%
ActRaw 4	5.8%	39.8%	2.3%	2.341%
RandAN 4	25.0%	12.8%	3.2%	3.205%

[1] These target-specific thresholds are also listed in the official MOOD implementation's evaluate.py.

Yield counterfactual. Under the counterfactual assumption that TanNorm's conditional hit efficiency at $\alpha/s = 2$ remains 35.0%, the yield needed to match MOOD's raw hit rate of 9.120% at $\lambda = 0.04$ is

$$y^* = \frac{0.09120}{0.350} \approx 0.261$$

This is approximately 2.8 times the current 9.3%. It is a descriptive counterfactual quantifying the yield increase required to close the current gap through yield alone, not a prediction of future performance. Conditional efficiency may also change when yield changes.

G.4 Pool-matched Top5

With fixed $k = 150$, a larger pass pool allows selection from a more extreme upper quantile. For each target–seed cell, we set

$$m = \min(n_{pass}^{GX}, n_{pass}^{MOOD})$$

sample m candidates without replacement from each pool, and apply the same top-150 rule. The existing implementation excludes pairs with $m < 150$ from this auxiliary analysis. All 4 comparisons in Table 32 retain 25/25 pairs, so this rule causes no missing pairs here.

For each pair, we average scores over repeated subsampling, then average over seeds and targets in that order. The implementation defaults are 400 repetitions and random generator seed 20260911. Saved results do not include a complete snapshot of runtime arguments, so these defaults are distinguished from run provenance. The table's m is the rounded mean of pair-specific values, not one fixed value shared by all pairs.

Table 32: Pool matching against MOOD λ=0.04. m is the seed-then-target average of pairwise minimum pool sizes, rounded to an integer.

GX-TGVAE setting	Mean m	GX-TGVAE	MOOD	Δ
TanNorm 2	278	9.813	9.258	+0.554
TanNorm 4	220	9.750	8.904	+0.846
ActNorm 4	546	9.780	10.053	−0.272
RandAN 4	750	9.534	10.383	−0.849

Pool matching does not always favor GX-TGVAE: MOOD remains stronger against ActNorm and RandAN. This auxiliary analysis examines differences in selection opportunity and does not replace the primary fixed-attempt benchmark.

G.5 Matched-OOD bands

We compare unique candidates by maximum-training-Tanimoto band after removing the hard similarity gate, QED / SA gates, and top-k selection. The *bandwise conditional-median summary* takes molecule-level median DS within each target–seed cell, averages over nonempty seeds, and then averages over targets. It is not the median of all pooled molecules. Candidate counts are averaged over seeds within each target, then over targets.

All 25 cells in all 8 bands in Table 33 are nonempty for both methods. However, GX-TGVAE has only 5.68 candidates on average in the deepest band. At sim < 0.3, sample mass is approximately 1,751 for MOOD and 399 for GX-TGVAE; at sim < 0.2, it is approximately 350 versus 6. GX-TGVAE has a higher conditional-median summary under some conditions in comparable OOD bands that it reaches, whereas MOOD allocates more sample mass to deep OOD. This difference in mass allocation does not imply absent GX-TGVAE decoder support.

G.6 Coverage and reached-region utility

We match docking opportunities at $B = 4$. For GX-TGVAE TanNorm, one trajectory comprises the same anchor's prespecified depths $\alpha/s \in \{0.5, 1, 2, 4\}$; for MOOD at $\lambda = 0.04$, one block comprises 4 consecutive attempts in generation order within each target–seed cell. MOOD's 3,000

Table 33: All 8 bands for TanNorm α/s=4 and MOOD λ=0.04. n is the seed-then-target mean unique-candidate count per 3,000 attempts. DS is the bandwise conditional-median summary. All 25/25 target–seed cells are nonempty in every band for both methods.

Band	GX-TGVAE n	MOOD n	GX-TGVAE DS	MOOD DS
[0.00,0.20)	5.68	349.72	8.262	8.436
[0.20,0.30)	393.04	1401.04	8.842	8.300
[0.30,0.35)	592.48	473.92	9.224	8.308
[0.35,0.40)	748.00	307.12	9.190	8.102
[0.40,0.45)	644.52	211.56	9.164	7.800
[0.45,0.50)	341.88	105.16	9.090	7.512
[0.50,0.60)	238.52	93.24	8.992	7.046
[0.60,1.01)	35.48	24.52	8.956	7.120

attempts yield 750 blocks. GX-TGVAE trajectories are grouped by target, seed, and canonical anchor key, retaining the last value for duplicate records at the same key and depth. GX-TGVAE has 2,999 trajectories for seeds 2 / 3 of each target and 3,000 otherwise, totaling 74,990. MOOD has 18,750 blocks across 25 cells. Matching B aligns opportunities per unit, not the total number of units or their within-unit dependence structure.

Common within-cell metrics and aggregation. Let $n_{t,s}$ be the number of units in cell (t, s), and let $E_{i,t,s}(c)$ contain unit i's candidates with valid DS and similarity below c.

$$R_{i,t,s}(c) = \mathbf{1}[|E_{i,t,s}(c)| > 0], \qquad n^{\text{reach}}_{t,s}(c) = \sum_i R_{i,t,s}(c).$$

For reached units, compute $M_{i,t,s}(c) = \max_{b \in E_{i,t,s}(c)} DS_b$, and define

$$C_{B,t,s}(c) = \frac{n^{\text{reach}}_{t,s}(c)}{n_{t,s}}, \qquad Q_{B,t,s}(c) = \frac{1}{n^{\text{reach}}_{t,s}(c)} \sum_{i:R_{i,t,s}(c)=1} M_{i,t,s}(c)$$

The conditional-utility expression is defined only for cells with $n^{\text{reach}}_{t,s}(c) > 0$. Both methods use the same equal weighting of 25 cells:

$$C_B(c) = \frac{1}{25} \sum_{t,s} C_{B,t,s}(c), \qquad Q_B(c) = \frac{1}{25} \sum_{t,s} Q_{B,t,s}(c).$$

Coverage includes cells with zero reach. We report Q_B only when $Q_{B,t,s}$ is defined in all 25 cells; we neither impute empty cells as 0 nor change the denominator to include only nonempty cells. At all 8 thresholds in Table 36, 25/25 cells are nonempty for both methods. This is not a pooled average over all reached units.

Table 34: Same threshold c=0.3. Both methods average equally over 25 target–seed cells.

Method	C	Q
GX-TGVAE TanNorm	0.188	8.75
MOOD λ=0.04	0.964	9.31

Table 35: Matched coverage C=0.5, linearly interpolated on the common 25-cell mean curves. Thresholds differ between methods.

Method	Threshold c	Q
GX-TGVAE TanNorm	0.36	9.12
MOOD λ=0.04	0.21	8.63

Table 36: All 8 observed coverage–quality points. Both methods have 25/25 nonempty cells at every threshold. B=4 with equal weighting of 25 cells.

c	GX-TGVAE C	GX-TGVAE Q	MOOD C	MOOD Q
0.2	0.0023	7.7437	0.3746	8.5220
0.25	0.0427	8.4448	0.8176	8.9016
0.3	0.1883	8.7535	0.9645	9.3061
0.35	0.4522	9.0765	0.9941	9.6063
0.4	0.7349	9.3124	0.9995	9.7614
0.45	0.9205	9.5542	0.9999	9.8491
0.5	0.9803	9.7263	1.0000	9.8865
0.6	0.9994	9.8851	1.0000	9.9149

At the same threshold $c = 0.3$, MOOD has both higher coverage and higher conditional best-of utility. At coverage 0.5, GX-TGVAE has Q of 9.12 versus MOOD's 8.63, but their thresholds differ at approximately 0.36 and 0.21, respectively.

For matched coverage, we linearly interpolate Q and the threshold using coverage as the horizontal coordinate between adjacent observed grid points. Both endpoints must have defined Q; we do not interpolate across undefined points or extrapolate. The threshold intervals bracketing $C = 0.5$ are [0.35, 0.40] for GX-TGVAE and [0.20, 0.25] for MOOD. Figure 6 shows the coverage–quality curves. Equal coverage does not imply equal OOD depth, and conditional selection and molecular-composition differences remain.

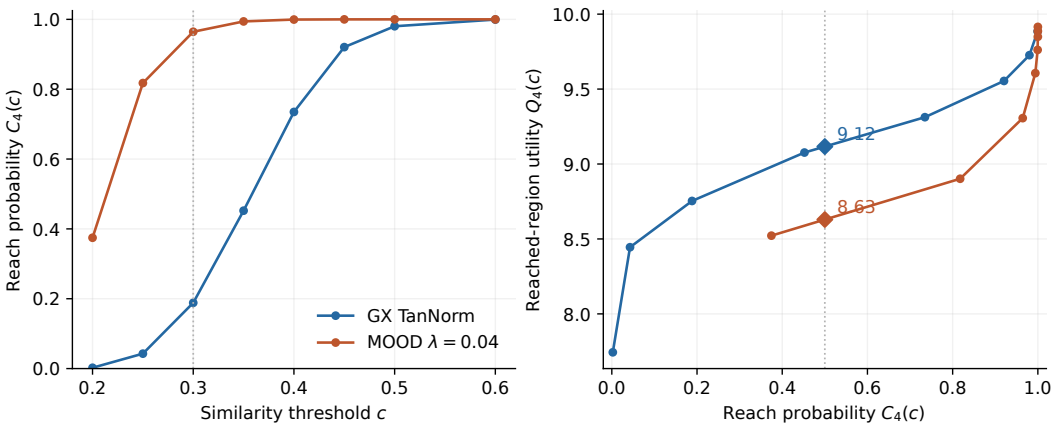


Figure 6: Coverage–quality curves with the common 25-cell aggregation. Left: reach probability versus threshold. Right: reached-region utility versus reach probability. Points are the 8 observed thresholds; lines connect adjacent points. Diamonds on the right are linear interpolants at $C = 0.5$, with different thresholds for the two methods. GX-TGVAE uses TanNorm and MOOD uses $\lambda = 0.04$, both with $B = 4$.

G.7 Docking stochasticity

Under QuickVina's `--cpu 5` setting, scores vary even for identical inputs and seeds. The existing duplicate-pair audit estimates single-run noise SD at approximately 0.231 kcal/mol under an independent-error approximation (Appendix H.3). Best-of-4 selection can introduce upward noise bias. Matching B controls the number of observations but does not ensure identical noise distributions, molecule dependence, or correlations across points; symmetry or cancellation of bias does not follow.

Observed differences, including matched-OOD summaries, should be read alongside this noise scale. Large sample size alone does not exclude noise effects. We apply neither noise-model corrections nor new uncertainty analyses in this comparison.

H Reproducibility and Evaluation Audits

H.1 Dataset splits

From the official MOSES Full training pool of 1,584,663 molecules (Polykovskiy et al., 2020), we fixed 10,000 molecules as internal validation using seed 42, leaving an effective training set of 1,574,663. The sets of saved SMILES rows have train / validation intersection 0. The 1k architecture comparison uses the first 1,000 entries in the same split's index sequence; baseline parity uses canonical 10k inputs. Official test / test_scaffolds serve as generation-metric references and are not used for architecture selection.

ZINC inputs originate from MOOD's distributed `zinc250k.csv` and `valid_idx_zinc250k.json` (Lee et al., 2023). We do not describe our split as an independent "ZINC official train set." MOOD indices partition the CSV's 249,455 rows into 224,568 training and 24,887 test rows. Neutralization, resanitization, stereochemistry removal, and canonical SMILES conversion remove 10,859 training and 1,194 test rows, leaving 213,709 and 23,693 saved rows.

The training side contains 212,398 unique graph keys and 1,311 duplicate rows. We select 10,000 graph keys for validation and remove all rows carrying those keys from training, forming a group split. Internal training has 203,654 rows; validation has 10,000 rows, one per key. These counts do not sum to 213,709 because 55 additional duplicate rows assigned to validation are collapsed to one row per key. Internal train / validation graph-key intersection is 0.

The preprocessed full train and test sets overlap on 295 unique graph keys, reflecting the non-injective preprocessing map. Keeping only unique test keys absent from training yields a clean test set of 23,378 with intersection 0 against full training. The ordinary and clean test sets are not interchangeable.

Split identifiers, file checksums, and manifests are maintained in separate reproducibility materials.

H.2 Checkpoint and baseline provenance

Final GX-TGVAE checkpoints for both MOSES and ZINC use SWA over epochs 90 / 95 / 100; Appendix B gives their configurations. Architecture comparisons generally use 1 run with training seed 42. We do not infer training-seed variability from the 3 seeds for 30k generation or the 5 seeds for the OOD benchmark. Correspondence between runtime configurations and the public release is maintained in a reproducibility manifest outside the paper.

MAGNet. MAGNet (paper architecture) follows Table 3 of the final ICLR 2025 paper (Hetzel et al., 2025). We distinguish "shape" in earlier versions and the public implementation from "scaffold" in the final paper. In our baseline configuration, `shape_gnn_dim` is adjusted from 128 to 129 to satisfy dimension-construction requirements. MAGNet (repository-default architecture) uses 189.

Paper / repository-default layer counts are 2 / 4 for the encoder, 3 / 2 for the hierarchical graph, 2 / 1 for the latent module, and 4 / 2 for the shape encoder. Both have latent dimension 100, leaf and shape RNN hidden dimensions of 256 each, and sum node aggregation. "Repository-default" denotes the architecture choice, not a reset of all training hyperparameters to repository defaults.

Both configurations share VAE training settings: batch size 64, learning rate 0.0003071057367777374, decay 0.9801, and gradient clipping 3. KL weight starts at 0 and, after step 2000, increases by 0.0005 every 2500 steps up to 0.01. Flow uses batch size 1024 and hidden dimension 512. Trained Flow is included for prior generation and excluded for reconstruction. These are the settings actually used in our baseline experiments, not a claim of matching every condition in the original paper.

TGVAE. The TGVAE (Nguyen & Karolak, 2025) baseline corresponds to a saved configuration with 8 encoder and 8 decoder layers, dropout 0.0 / 0.1, 43,431,451 parameters, and greedy decoding at epoch 80. The inspected vendored latent sampler uses `torch.rand_like`, while the GX-TGVAE training wrapper with `--reparam normal` replaces its forward method with one using `torch.randn_like`. The saved training log for the target baseline records application of the

normal sampler; the final evaluation log with the same experiment identifier corresponds to greedy decoding from the stated checkpoint. Complete binding between the checkpoint and the full runtime source snapshot remains unverified. We distinguish the original model identity, public implementation, and experimental baseline settings.

For checkpoint selection, we generated 12,000 candidates from each of 20 saved snapshots at epochs 5–100, securing 10,000 valid molecules even under validity near 86%. Candidates were compared using FCD @10,000 against internal validation held out from MOSES training; epoch 80 ranked first among checkpoints meeting the uniqueness eligibility condition. Decoding conditions at that checkpoint were then compared with the same internal-validation FCD, selecting greedy decoding. MOSES test / test_scaffolds are reference sets for final reporting, not selection metrics for checkpoints or decoding conditions.

GraphXForm and MoLeR. The main GraphXForm (Pirnay et al., 2025) condition is paper-aligned: 8 heads, dropout 0.1, and epoch 500 (1.5M batches). The official-native configuration with 16 heads, dropout 0, and best epoch 986 (3.0M batches) is a separate variant; its checkpoint diagnostic in Appendix D.2 is not mixed with the main condition. MoLeR (Maziarz et al., 2022) uses greedy generation, and its latent-use evaluation is a paired comparison on common canonical inputs.

H.3 Docking protocol and stochasticity

Final docking fixes QuickVina2 (Alhossary et al., 2015) to `--cpu 5`, exhaustiveness 1, num_modes 10, and seed 42. Ligand conformers use a shared cache indexed by canonical SMILES keys.

Across 13,539 duplicate pairs in which the same input molecule appeared in existing runs, a fixed seed did not ensure identical scores. Table 37 reports the existing aggregate. We do not assume that these duplicates are independent samples from one noise distribution shared uniformly across all molecules and targets.

Table 37: Existing duplicate-pair audit (13,539 pairs). Docking-score differences are in kcal/mol.

Statistic	Value
exact score match	0.6512
$\lvert\Delta\rvert > 0.1$	0.2515
$\lvert\Delta\rvert > 0.5$	0.0557
$\lvert\Delta\rvert > 1.0$	0.0186
max $\lvert\Delta\rvert$	3.90
signed mean difference	−0.0036
pair-difference SD	0.3273

Approximating the two errors as independent and identically distributed, the pair-difference SD gives a single-run scale of

$$\sigma_{dock} = 0.3273/\sqrt{2} \approx 0.231 \text{ kcal/mol}$$

This is a conditional scale estimate and does not establish the absence of systematic bias or heteroscedasticity. We do not treat single-molecule differences around 0.1 kcal/mol as strong evidence, nor conclude that noise cannot change an aggregate conclusion merely because it is aggregated.

Within a plan sharing the same score lookup, identical target–molecule keys refer to the same memoized realization. Score differences for an identical molecule cancel within that scope, but not for the same molecule docked independently in another run or for differences between distinct molecules. Sharing a conformer cache differs from sharing a docking realization.

H.4 Hardware and random-number reproducibility

The experimental A100 MIG environment contained both 3g.40gb and 4g.40gb profiles. The existing audit found different CUDA `torch.randn` outputs across profiles under the same seed, while outputs matched within a profile under the checked conditions. Specifying only the GPU type does not fix the profile, so subsequent runs fixed 3g.40gb where required.

RNG kernel / grid mapping dependent on GPU multiprocessor count is a candidate explanation. However, binding the exact experimental PyTorch / CUDA source versions to kernel paths remains open; we do not draw a general causal conclusion excluding all other nondeterminism. Treating the same seed as the same random numbers also requires recording the hardware execution environment's identity.

H.5 Validity and software semantics

MOSES validity (Polykovskiy et al., 2020) primarily checks RDKit parsing and sanitization, so disconnected molecules containing . can be valid. MAGNet +Flow has MOSES validity 1.000 but connected-validity sidecars around 0.87–0.88, with approximately 3.5k–3.8k disconnected outputs per 30k. GX-TGVAE adds atoms and bonds to a single connected partial graph, and no disconnected generation was observed in this evaluation. Numerically equal validity need not have identical semantics; we do not use validity as the central evidence of superiority over baselines.

RDKit and PyG TransformerConv software identities follow existing references (RDKit contributors; PyTorch Geometric contributors). Saved metadata for the 10k paired analysis record RDKit 2026.03.3. We do not assign this version to all training, generation, or docking runs. Complete run-specific bindings for RDKit, PyTorch, PyG, CUDA, MOSES revisions, and QuickVina binary hashes remain open. Inspecting the current environment alone cannot establish the historical environment.

H.6 Structural-reach reproducibility

Structural reach in Appendix D.4 evaluates the same generated sets from 4 models and 3 generation seeds with a common pipeline. Distance uses the effective training split, unlike the full official training set used by the auxiliary scaffold diagnostics in Appendix D.1. Metric definitions, seed-specific denominators, cache identities, and execution environments are linked in the reproducibility materials.

H.7 Denominators and transcription

The units of analysis differ as follows.

- Latent fidelity: common 1k or 10k inputs, retaining the corresponding pair counts.
- MOSES generation: 30k attempts across 3 generation seeds, with each metric's valid / unique or other denominator.
- Structural distance: molecules with defined scaffold distance. L3 counts unique topologies; shape OOV uses decomposable molecules.
- Fixed-budget docking: 3,000 attempts. Failures, duplicates, and gate failures remain in the hit-ratio denominator.
- Anchor-relative comparison: pairs with valid anchor and endpoint scores.
- Matched-OOD: medians of unique molecules within each band, averaged over seeds and then targets.
- Coverage: GX-TGVAE trajectories with 4 depths and MOOD blocks with 4 attempts. Coverage and conditional utility are both computed within common target–seed cells and then averaged equally across 25 cells (Appendix G.6).

Standard MOSES metrics and our auxiliary scaffold diagnostics appear in separate tables; both report means and sample SDs for 30k generations across 3 generation seeds. Inter-model comparisons use aligned reference sets and denominators for each metric.

H.8 Remaining reproducibility limits

This manuscript links saved numerical results to their denominators and model variants and separates the scopes of canonical likelihood, geometry, and empirical OOD claims. Open items remain in the correspondence between complete runtime source snapshots and the public release, full binding of TGVAE checkpoints to runtime source snapshots, software / binary identities, and scaffold-cache

construction provenance. The coverage–quality comparison uses a common 25-cell aggregation for both methods.

These checks do not imply retraining or regeneration of results. They delimit what can be established without guessing unverified implementation conditions. Unresolved citations and individual verification records remain in audit ledgers outside the manuscript.

I Evolution-Strategy Search and Massive Harvest

This inference-time extension addresses the candidate-mass deficit of the base trajectory sampler while keeping the GX-TGVAE encoder, decoder, and property surrogate fixed. ES selects one search direction from several candidates, a $4 \to 1$ operator; Harvest produces multiple outputs around one center, a $1 \to K$ operator. After the K=4 pilot, we fixed the selector and Massive Harvest evaluation rules, then generated a K=1024 pool with new perturbations. The extension is exploratory and was designed after the base benchmark.

I.1 Geometry-aware direction selection by ES

Candidates and updates. Let P_A be the active-coordinate projector and $U_T(z)$ an orthonormal tangent basis estimated from the current neighborhood. We use $P_T = U_T U_T^\top$ and $P_N = P_A - P_T$. ActES proposes directions in the active space; NormES uses the active-normal space. TanNormES takes $K_T = 5$ tangent warm-up steps followed by a normal phase of $T = 16$ steps. Neighborhoods and property gradients are updated at each step.

We use a simple $(1, 4)$-style evolution strategy (ES) (Beyer & Schwefel, 2002) with population $M = 4$ and selection $\mu = 1$. Candidate 0 is the corresponding base arm's analytic/property-gradient direction; the other three are normalized Gaussian vectors projected into the relevant subspace. The probe radius is $\rho/s = 0.5$, the update step is $\delta/s = 0.25$, and $s = 3.828$ is fixed. For each direction v_i, we greedily decode $z + \rho v_i$ and select a direction using

$$F_i = R(\widehat{f}_i) + R(1 - \text{sim}_i) - 5[0.5 - \text{QED}_i]_+ - 5[5/9 - SA_{n,i}]_+.$$

Here R maps the ranks of valid candidates within the population to $[-0.5, 0.5]$, assigning average ranks to ties. If only one candidate is valid, its rank terms are 0; invalid candidates have fitness $-\infty$. The composite property surrogate $\widehat{f}$ is defined in Appendix F and is distinct from the docking oracle. Similarity is maximum Tanimoto to the official training molecules. We update $z \leftarrow z + \delta v_{i^*}$, or remain at the current point if all candidates are invalid. We do not learn a covariance matrix or impose a hard OOD threshold in the fitness.

Neighborhood estimation and comparison scope. Exact nearest-neighbor search and neighborhood SVD determine a tangent basis explaining 90% of the variance; it is reorthogonalized after projection into the active space. To reproduce the existing trajectories, TanNorm uses 50 neighbors and TanNormES uses 256. Each retains its original direction RNG settings. These comparisons therefore evaluate complete search configurations rather than isolate ES with neighborhood settings held fixed. Probe decoding and surrogate evaluation incur additional computation, but direction selection does not use docking feedback.

I.2 ES and non-ES trajectory results

We use five targets (5ht1b, braf, fa7, jak2, parp1), three generation seeds, and 3,000 anchors per cell. Table 38 distinguishes all trajectory checkpoints, including the initial point, from the four checkpoints at $\alpha/s = 0.5, 1, 2, 4$. Reach is the fraction of anchors with at least one qualifying molecule; conditional utility averages the best DS within reached anchors. Full trajectories contain 22 points for TanNorm / TanNormES, including the five-step warm-up, and 17 for the other arms. Full trajectories and the four-point budget do not have equal output budgets.

TanNormES has a higher full-trajectory any-hit rate than TanNorm, but conditional best DS does not improve uniformly. At the four matched checkpoints and $c = 0.4$, TanNormES does not exceed TanNorm in reach or utility. We therefore evaluate ES together with subsequent Harvest allocation

Table 38: ES and base-arm trajectory evaluation (means over 15 cells). Full includes all checkpoints, the initial point, and warm-up; $B = 4$ uses four fixed depths. C, Q denote reach and conditional best DS at similarity threshold 0.4; C_u also requires QED/SA gates. p_{hit} is the fraction of anchors whose trajectory contains at least one hit.

Arm	p_{hit}	C_{full}	Q_{full}	C_u	C_4	Q_4
ActRaw	0.228	0.816	9.464	0.604	0.563	8.918
ActES	0.330	0.881	9.502	0.777	0.674	8.912
ActNorm	0.251	0.897	9.447	0.714	0.747	8.828
NormES	0.272	0.881	9.200	0.793	0.768	8.687
TanNorm	0.329	0.938	10.024	0.709	0.740	9.324
TanNormES	0.436	0.949	9.939	0.836	0.730	9.075

rather than interpret it as an improvement at every depth or on every utility measure. ES selection is also not evidence that estimated normal directions depart from the true molecular manifold.

I.3 K=4 Harvest pilot

Local latent generation has precedents: HelixMO generates multiple Gaussian perturbations of selected candidates (Chen et al., 2022), while ConGen decodes neighborhoods around optimized latent points (Koban et al., 2026). Here, Harvest samples at a fixed radius in a fixed active-coordinate subspace around geometry-guided search centers of the frozen GX-TGVAE, retaining every output without best-of selection; neighborhood hit rates subsequently inform center selection and budget allocation.

From TanNorm / TanNormES centers z_j at normal-phase steps $j = 8, 12, 16$, we generate $K = 4$ neighbors using

$$z^H_{j,k} = z_j + \rho_H v_k, \qquad v_k = \frac{P_A r_k}{\|P_A r_k\|}, \quad r_k \sim \mathcal{N}(0, I), \quad \rho_H / s = 0.5$$

where P_A projects onto the fixed 41 active coordinates identified from the ZINC latent bank. Sampling takes place in this active-coordinate subspace around each center, rather than along a center-specific tangent basis or with an adaptively optimized radius. Each sample is greedily decoded and retained, without fitness-based best-of selection. This gives a 2×2 design: search with or without ES, each with or without Harvest.

The 3,000 anchors in each cell give 9,000 centers. Center-only evaluation has 9,000 outputs; adding all four neighbors gives 45,000. Table 39 reports scaling at this natural output budget. Intermediate $K = 1, 2, 3$ values average over all combinations of the four neighbors.

Table 39: Natural-budget scaling within the K=4 pilot. For each center, K added neighbors are averaged over all $\binom{4}{K}$ combinations, then equally over 15 cells. Centers are also outputs, giving a budget of $9,000(1 + K)$.

Search	K	Outputs	Unique hits	NHR (%)	Top 5% DS
TanNorm	0	9000	282.3	3.137	10.619
TanNorm	1	18000	563.5	3.131	10.971
TanNorm	2	27000	844.2	3.127	11.155
TanNorm	3	36000	1124.5	3.124	11.277
TanNorm	4	45000	1404.4	3.121	11.366
TanNormES	0	9000	620.6	6.896	10.815
TanNormES	1	18000	1041.2	5.784	11.039
TanNormES	2	27000	1461.2	5.412	11.176
TanNormES	3	36000	1880.7	5.224	11.274
TanNormES	4	45000	2299.7	5.110	11.350

Unique hits increase at the natural budget, but so does the number of outputs. At a fixed budget of 3,000, we sample 200 anchor blocks without replacement, each containing three centers and 12

neighbors. Averaging 400 subsamples, neither search configuration reaches MOOD's performance (Table 40). Thus, the K=4 pilot alone does not close the fixed-budget deficit.

Table 40: K=4 pilot comparison at 3,000 slots. GX-TGVAE results average 400 subsamples of anchor blocks. The Massive Harvest selector is not used.

Search	Unique hits	NHR (%)	Top 5% DS
TanNorm	93.6	3.120	9.778
TanNormES	152.9	5.097	10.084
MOOD	271.3	9.044	10.939

Table 41 reports each neighbor's DS change relative to its center. Many neighbors have similar scores, but both increases and decreases occur; we do not assume locally flat utility.

Table 41: Neighbor-minus-center ΔDS in the K=4 pilot, using valid paired docking evaluations. Not every neighbor improves on its center.

Search	j	Mean	Median	10th pct.	90th pct.	$\Delta DS \geq 0$ (%)
TanNorm	8	-0.018	0.000	-1.700	1.700	51.9
TanNorm	12	-0.020	0.000	-1.900	1.900	52.2
TanNorm	16	-0.032	0.000	-2.100	2.100	52.3
TanNormES	8	-0.004	0.000	-1.600	1.600	52.3
TanNormES	12	-0.007	0.000	-1.700	1.700	52.1
TanNormES	16	-0.010	0.000	-1.800	1.800	51.8

I.4 Pilot-trained harvestability selector

For large-scale local sampling, the frequency of useful neighboring outputs matters alongside the center's own score. We define harvestability conceptually as

$$h(z_c) = \mathbb{E}_{\delta z \sim q_H(\cdot|z_c)}[\mathbf{1}\{\text{decode}(z_c + \delta z) \text{ is a hit}\}]$$

and train on the K=4 pilot's local hit rate. This target is the fraction of hits among four outputs, not the deduplicated unique-hit count at $K = 1024$.

Features are the center's surrogate score, training similarity, QED, normalized SA, and one-hot depth $j = 8, 12, 16$. The center's docking score is not an input feature. For each search $\times$ target, we fit linear ridge regression

$$\min_{w,b} \|Xw + b\mathbf{1} - y\|_2^2 + \|w\|_2^2$$

with $\lambda = 1$ and an unregularized intercept. The four continuous features are standardized using the fit subset's means and population standard deviations; a standard deviation below 10^{-12} is replaced by 1. Predictions are clipped to $[0, 1]$.

Three-fold out-of-fold (OOF) prediction fits on 18,000 centers from two generation seeds and predicts 9,000 centers from the remaining seed. For each target $\times$ seed $\times$ search, we select the top 32 predicted centers, breaking ties by ascending anchor index and then j. Following the within-pilot OOF results in Table 42, we fixed these features, ridge coefficient, folds, top32 selection, and the evaluation rules below.

Table 42: OOF prediction and top32 local hit rates on K=4 development data (means over 15 cells). Random32 is a random-selection control in the same pilot.

Search	Spearman	Top32 hit (%)	Random32 hit (%)
TanNorm	0.160	8.44	3.28
TanNormES	0.152	10.47	4.79

K=4 is development data: the selector itself was designed after inspecting the pilot results. It was not retuned using the new K=1024 perturbations or outcomes. OOF avoids fitting on the evaluated

seed's own pilot outcomes; it does not establish generalization to unseen targets or complete independence of molecular identities across seeds. Docking-derived labels used to fit the selector must be distinguished from center docking scores used as selection features.

I.5 K=1024 POOL AND LOCAL CAPACITY

Three center-selection rules. Alongside the learned selector, we fixed two oracle controls. (i) The hit-stratified control randomly selects 32 centers that are themselves hits, with quotas of 11/11/10 for $j = 8, 12, 16$; remaining hit centers fill any shortfall. (ii) Screen150 uses the same 150 anchor indices per seed across targets and searches, evaluates their 450 centers at three depths, and retains all hit centers. The first diagnoses local capacity around oracle-selected hit centers; the second supports budget comparison including screening costs.

The union across selection rules contains 2,550 unique centers, each producing 1,024 outputs for a total of 2,611,200 slots. There are 2,602 selection memberships because a center may belong to multiple rules. Perturbation RNGs are fixed by center identity and use a different sequence from the K=4 pilot. All 2,550 regenerated centers match their original canonical molecular identities. New perturbations can still decode previously observed molecules; the new pool is not a test set of entirely unseen molecular identities.

Measuring capacity. Using each center's fixed generation order, we count unique molecules, unique hits, and unique hit BM scaffolds for prefixes $K = 4, 16, 32, 64, 128, 256, 512, 1024$. Centers receive equal weight within each selection rule; centers belonging to multiple rules enter each corresponding summary. Table 43 includes the hit-stratified and Screen150 controls.

Table 43: Local capacity per center at $K = 1024$, weighting selected centers equally within each selection rule. Hit BM counts unique nonempty Bemis–Murcko scaffolds among hits. The aggregation unit differs from cell means at the 3,000-slot budget.

Selection	Search	Unique mol.	Unique hits	Hit BM	NHR (%)
Learned	TanNorm	848.7	75.19	71.97	7.34
Learned	TanNormES	874.1	102.03	96.89	9.96
Screen150	TanNorm	802.8	68.17	66.15	6.66
Screen150	TanNormES	861.1	101.19	95.99	9.88
Hit-stratified	TanNorm	814.2	65.09	62.55	6.36
Hit-stratified	TanNormES	849.1	78.58	74.50	7.67

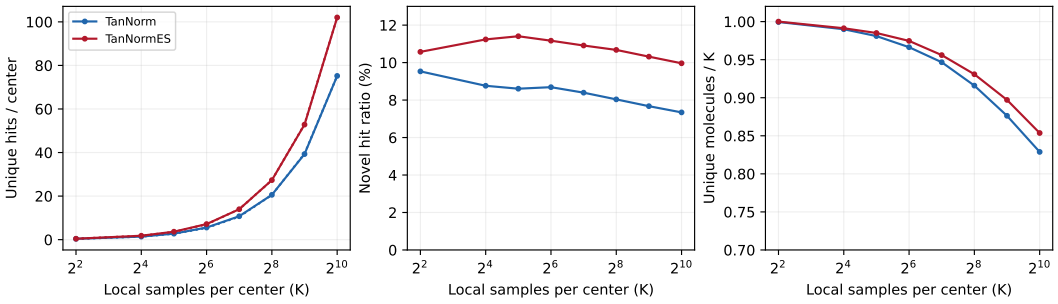


Figure 7: Capacity curves for centers chosen by the learned selector. Left: unique hits per center. Middle: NHR with all sampling slots in the denominator. Right: unique molecules / K. Dashed lines on the left show random-subset means from the same pool.

At $K = 1024$, learned TanNormES centers yield means of 874.1 unique molecules, 102.03 unique hits, and 96.89 unique hit BM scaffolds. The hit count exceeds TanNorm's 75.19, and the results are not merely repetitions of a few scaffolds. TanNormES prefix NHR declines gradually from 10.57% at $K = 4$ to 9.96% at $K = 1024$. Selected neighborhoods thus contain substantial useful output mass, but duplicates and changing hit density do not support an infinite, flat, or spherical productive region.

To assess order dependence, the saved analysis also draws K outputs without replacement from the same pool over 32 repetitions (Table 44). Prefix and random-subset means are close for learned TanNormES at $K = 64, 256, 512$, suggesting that high capacity is not explained solely by the initial member order. These repetitions measure within-pool sampling variation, not training variability or uncertainty across independent pools.

Table 44: Mean unique hits per center: prefix / random-subset. Random uses 32 repetitions of sampling K outputs without replacement from the same 1,024-output pool, not confidence intervals across independent experiments.

Selection	Search	$K = 64$	$K = 256$	$K = 512$
Learned	TanNorm	5.56 / 5.37	20.57 / 20.51	39.29 / 39.49
Learned	TanNormES	7.15 / 7.04	27.33 / 27.31	52.83 / 52.97
Screen150	TanNorm	5.53 / 5.35	19.57 / 19.43	37.04 / 36.71
Screen150	TanNormES	7.39 / 7.31	27.75 / 27.73	53.52 / 53.22
Hit-stratified	TanNorm	4.83 / 5.05	18.32 / 18.57	35.03 / 34.93
Hit-stratified	TanNormES	5.86 / 5.84	21.95 / 21.84	41.73 / 41.68

I.6 Fixed evaluation budget and target-level results

Budget definition. We reconstruct policies with 3,000 evaluation slots from the completed pool using the order and allocation fixed beforehand. The learned policy spends 32 slots on the selected centers and 2,968 on neighbors. Rank-ordered round-robin allocation assigns 93 neighbors to 24 centers and 92 to eight centers. The oracle-adaptive policy spends 450 slots on screening, including non-hit centers, and 2,550 on the neighborhoods of all selected hit centers. Screen150 yields 4–62 hit centers, requiring at most 638 neighbors per center, within the pool capacity.

The 3,000 slots constitute an **evaluation-time docking budget**. Duplicate outputs and failures consume slots, and score-cache reuse does not reduce the denominator. This is neither lifetime oracle cost nor a wall-clock-matched comparison: it excludes pilot selector training, probe decoding during search, and generation of the full pool. MOOD uses 3,000 attempts at $\lambda = 0.04$ for the same five targets $\times$ three seeds.

Hits and duplicates. Eligible candidates are unique by canonical identity and satisfy training similarity < 0.4, QED > 0.5, and normalized SA $> 5/9$. DS negates docking affinity and clips it below at 0. Hits are eligible candidates exceeding the existing target-specific benchmark thresholds: 5ht1b, 8.7845; braf, 10.3; fa7, 8.5; jak2, 9.1; parp1, 10.0. Unique hits are deduplicated across all 3,000 outputs of each strategy $\times$ target $\times$ seed, including center–neighbor and between-center duplicates. We retain the first occurrence of each identity. Capacity tables deduplicate within each center instead, so summing per-center unique hits does not reproduce the 3,000-slot count.

Top 5% DS averages the top 150 molecules in the eligible unique pool, or all available candidates if fewer than 150 exist. Invalid or missing docking scores are non-hits and contribute 0 to DS aggregation. NHR, candidate yield, and conditional efficiency are computed per cell and averaged with equal weights over 15 cells. The product of mean yield and mean conditional efficiency does not recover mean NHR.

As shown in Main Table 3 and Tables 45–46, TanNormES learned yields 331.3 hits versus MOOD's 271.3. The paired difference averages 59.9 hits with a between-cell SD of 77.5; it is positive in 13/15 cells and 4/5 target means. MOOD is higher on braf. Oracle-adaptive TanNormES yields 275.4 hits, meeting the primary exploratory criterion fixed before pool generation (higher mean hits and NHR than MOOD), but the difference is 4.1 hits with SD 108.3 and is positive in only 8/15 cells. These SDs describe target–seed variation, not significance tests.

Non-ES TanNorm yields 256.6 hits with learned selection and 187.6 with oracle selection, both below MOOD's mean. Among learned policies, the ES configuration raises candidate yield from 29.14% to 35.80%, with similar conditional efficiencies of 28.83% and 29.30%. Among oracle policies, yield rises from 13.03% to 23.06% while conditional efficiency falls from 47.65% to 40.12%. This

Table 45: Unique hits at 3,000 slots in all 15 target–seed cells. L: learned selector; O: oracle-adaptive; TN: TanNorm; ES: TanNormES. MOOD uses the same seed subset.

Target	Seed	TN-L	ES-L	TN-O	ES-O	MOOD
5ht1b	0	500	752	273	468	581
5ht1b	1	444	723	203	394	564
5ht1b	2	666	759	255	388	588
braf	0	124	120	114	153	217
braf	1	78	207	68	145	178
braf	2	181	109	158	134	194
fa7	0	53	88	84	141	51
fa7	1	88	108	109	128	40
fa7	2	124	128	165	189	52
jak2	0	251	373	214	348	290
jak2	1	220	377	188	279	305
jak2	2	311	368	263	365	296
parp1	0	303	272	232	304	247
parp1	1	193	298	248	370	229
parp1	2	313	287	240	325	238

Table 46: Unique hits by target (means over three seeds). Abbreviations follow Table 45.

Target	TN-L	ES-L	TN-O	ES-O	MOOD
5ht1b	536.7	744.7	243.7	416.7	577.7
braf	127.7	145.3	113.3	144.0	196.3
fa7	88.3	108.0	119.3	152.7	47.7
jak2	260.7	372.7	221.7	330.7	297.0
parp1	269.7	285.7	240.0	333.0	238.0

shows the importance of allocating sufficient candidate mass alongside high hit density. Differences between search configurations prevent attributing the mass increase solely to ES.

I.7 Completeness and interpretation

All 2,611,200 slots have an evaluation status. The 98 docking-failure slots remain in the denominator as non-hits. Scores are reused for the same target and canonical molecule, so slot counts are not counts of distinct molecules or newly executed docking calls. Properties use similarity to the official ZINC training set and the common QED/SA definitions.

These results update the fixed-budget hit-yield deficit observed for the base trajectory sampler. TanNormES with the learned policy exceeds MOOD in unique hits under the same evaluation budget. MOOD's Top 5% DS of 10.939 remains higher than learned TanNormES's 10.416 and oracle TanNormES's 10.474, so the extension does not establish superiority on every utility metric. Learned and oracle policies differ in both center selection and screening cost; their difference cannot be decomposed into a selector-only effect.

Treating center score and neighborhood productivity as distinct objectives helped improve budget allocation. The next steps are to broaden the scope of a fixed policy and design sampling around nonuniform neighborhood structure. Information geometry and flow-based transport are candidates for future work, not methods implemented or validated here.