



OPEN ACCESS

EDITED BY
Wenxing Li,
Columbia University, United States

REVIEWED BY
Yuan Zhou,
Peking University, China
Jiaqi Zhang,
Brown University, United States
Maziyar Baran Pouyan,
Accenture, United States

*CORRESPONDENCE
Zeyu Fu,
✉ fuzeyu99@126.com
Tianfei Ran,
✉ rantianfei@tmmu.edu.cn
Chunlin Chen,
✉ chenchlin3@mail2.sysu.edu.cn

[†]These authors have contributed equally to this work

RECEIVED 25 March 2026
REVISED 13 April 2026
ACCEPTED 15 April 2026
PUBLISHED 08 June 2026

CITATION
Fu Z, Fu J, Zhang K, Ran T and Chen C (2026) LAIOR: a hyperbolic neural ODE variational framework for interpretable single-cell manifold learning and trajectory inference. *Front. Genet.* 17:1838613. doi: 10.3389/fgene.2026.1838613

COPYRIGHT
© 2026 Fu, Fu, Zhang, Ran and Chen. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](#). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

LAIOR: a hyperbolic neural ODE variational framework for interpretable single-cell manifold learning and trajectory inference

Zeyu Fu^{1*†}, Jiawei Fu^{2†}, Keyang Zhang³, Tianfei Ran^{2*} and Chunlin Chen^{4*}

¹State Key Laboratory of Trauma and Chemical Poisoning, Institute of Combined Injury, Chongqing Engineering Research Center for Nanomedicine, College of Preventive Medicine, Army Medical University, Chongqing, China, ²Department of Orthopedics, Xinqiao Hospital, Army Medical University, Chongqing, China, ³School of Medicine, Sun Yat-sen University, Shenzhen, China, ⁴Department of Rehabilitation Medicine, The First Affiliated Hospital, Sun Yat-sen University, Guangzhou, China

Single-cell omics data are high-dimensional, sparse, and noisy, and learning embeddings that simultaneously preserve local cell-state structure, global hierarchy, and smooth developmental trajectories remains an open problem. Existing approaches typically achieve only one of these goals: classical methods emphasize either local neighborhoods or global variance; deep generative models cluster cell types well but often fracture trajectory continuity; and hyperbolic embeddings capture hierarchy but are numerically fragile in practice. We present LAIOR (Lorentz attentive interpretable ordinary differential equation (ODE)-regularized variational autoencoder (VAE)), a unified variational framework that combines three complementary inductive biases in a single forward pass: (i) *Lorentz geometric regularization* encourages tree-like latent hierarchy while remaining numerically stable *via* tangent-space clamping and exponential-map gating; (ii) a *dual-path information bottleneck* captures coordinated biological programs rather than forcing latent independence; and (iii) *neural ordinary differential equation (ODE) regularization* stabilizes latent trajectories through explicit learned dynamics. Across 118 single-cell datasets (53 scRNA-seq and 65 scATAC-seq) benchmarked against 23 baseline methods on 22 complementary metrics, LAIOR improves manifold continuity, trajectory coherence, and embedding fidelity while retaining competitive clustering performance. Ablation and sensitivity analyses show that ODE regularization stabilizes geometric learning and dampens hyperparameter sensitivity. Architecture interpretation experiments on two well-characterized reference systems (human bone marrow and mouse pancreatic endocrinogenesis) demonstrate that LAIOR's encoder and decoder pathways decompose cellular variation into mutually exclusive, biologically coherent latent modules, and biological validation experiments on two previously unseen hematopoietic perturbation cohorts (*Dapp1* knockout and chemotherapy-induced bone marrow failure) show that the same interpretability contract transfers to perturbed biology. LAIOR generalizes across RNA and chromatin accessibility modalities without architectural changes. Head-to-head comparisons against both dynamical baselines (scTour) and foundation models (scGPT, scFoundation) confirm that explicit geometric and dynamical inductive biases recover trajectory structure that large-scale

pretraining alone does not. Together, these results establish LAIOR as a practical, interpretable framework for single-cell manifold and trajectory analysis.

KEYWORDS

benchmarking, dynamics modeling, hyperbolic geometry, information bottleneck, manifold learning, single-cell

1 Introduction

Single-cell omics technologies have transformed the study of cellular heterogeneity by profiling molecular states at single-cell resolution (Tang et al., 2009; Kolodziejczyk et al., 2015). Single-cell RNA sequencing (scRNA-seq) captures transcriptomic landscapes, while single-cell assay for transposase-accessible chromatin (scATAC-seq) maps regulatory architecture (Bu et al., 2015; Cusanovich et al., 2015). However, the high dimensionality, sparsity, and technical noise inherent in these data present fundamental challenges for computational analysis (Brennecke et al., 2013; Grün et al., 2014). A central objective is to learn low-dimensional latent representations that preserve both local neighborhood structure, which is critical for cell-type identification, and global topology, which is essential for understanding developmental hierarchies and differentiation trajectories (Haghverdi et al., 2016; Trapnell et al., 2014).

1.1 The local–global trade-off in single-cell embeddings

Classical dimensionality reduction methods prioritize either local or global structure, but rarely both. Principal component analysis (PCA) captures global variance but distorts local neighborhoods due to linearity constraints (Jolliffe and Cadima, 2016). Non-linear methods such as t-SNE (van der Maaten and Hinton, 2008) and UMAP (McInnes et al., 2018) excel at preserving local clusters but systematically compress global distances, obscuring hierarchical relationships (Kobak and Berens, 2019; Becht et al., 2019). This trade-off propagates into downstream analyses: trajectory-inference algorithms trained on distorted embeddings may misidentify differentiation pathways, while batch correction methods can inadvertently erase biological gradients (Saelens et al., 2019; Tran et al., 2020).

Recent deep generative models such as scVI (Lopez et al., 2018) and scGNN (Wang et al., 2021) demonstrate robust performance in discrete cell-type clustering by modeling count distributions and incorporating graph regularization. However, evaluations reveal persistent deficiencies in global topology preservation: distance correlations between high-dimensional and latent spaces remain weak, and trajectory directionality exhibits inconsistency (Saelens et al., 2019; Luecken et al., 2022). Factor-based VAE variants (β -VAE (Higgins et al., 2017), InfoVAE (Zhao et al., 2019)) prioritize statistical independence for interpretability, but axis-aligned factorization conflicts with the intrinsic geometry of cellular differentiation, where biological processes exhibit coordinated, hierarchical organization (Buettner et al., 2015).

1.2 Hyperbolic geometry and neural ordinary differential equation (ODE) dynamics

Biological systems frequently exhibit tree-like hierarchies: hematopoietic stem cells differentiate into progressively specialized lineages, developmental programs branch into distinct fates, and immune responses generate clonal expansions (Laurenti and Göttgens, 2018; Wagner and Klein, 2020). Hyperbolic space with constant negative curvature enables exponentially more efficient tree embeddings, requiring $O(\log n)$ dimensions versus $O(n)$ in Euclidean space (Sarkar, 2011; Krioukov et al., 2010). This has motivated the use of hyperbolic methods, including Poincaré maps (Klimovskaia et al., 2020), probabilistic geometric models (Poincaré Gaussian mixture, or PGM; Mathieu et al., 2019), and hyperbolic graph networks (Nagano et al., 2019; Tian T. et al., 2023).

Despite theoretical advantages, existing hyperbolic approaches face critical limitations. Fully hyperbolic latent variable models require all operations within hyperbolic manifolds, introducing numerical instabilities from arccosh singularities and gradient vanishing (Ganea et al., 2018; Skopek et al., 2020). Pure geometric embeddings lack mechanisms to enforce temporal consistency, leaving them unstable under biological noise (Street et al., 2018). Most methods optimize for either hierarchical structure or local clustering but do not systematically integrate both objectives (Wolf et al., 2019).

Cellular differentiation unfolds as continuous dynamical processes governed by regulatory programs (Weinreb et al., 2018; La Manno et al., 2018). Neural ordinary differential equations (ODEs) provide a natural framework by parameterizing latent trajectories as solutions to $\frac{dz}{dt} = f_\theta(z, t)$, where learned function f_θ captures biological transition rules (Chen et al., 2018b; Grathwohl et al., 2019). ODE-based models excel at trajectory inference by enforcing temporal consistency. Cells at intermediate pseudotimes must lie on smooth paths connecting progenitor and terminal states (Bergen et al., 2020; Qiu et al., 2022). However, these methods typically operate on pre-computed embeddings from PCA or scVI, inheriting any distortions from upstream dimensionality reduction (Setty et al., 2019; Lange et al., 2022). scTour (Li, 2023) integrates VAE training with ODE constraints but lacks geometric regularization to preserve hierarchical structure. We hypothesized that combining hyperbolic geometric principles with neural ODE dynamics within a unified generative framework could yield complementary regularization.

1.3 LAIOR: lorentz attentive interpretable ODE-regularized VAE

We present LAIOR (Lorentz attentive interpretable ODE-regularized VAE), a deep generative framework that addresses

the local–global trade-off through three architectural innovations (Figure 1A). First, *Lorentz geometric regularization* aligns latent representations with hyperbolic structure *without* embedding the entire latent space in hyperbolic manifolds. LAIOR maintains a Euclidean latent space for computational stability while using Lorentzian distance penalties to preserve hierarchical relationships during training, thereby avoiding numerical instabilities while capturing the exponential capacity of negative curvature geometries (Nickel and Kiela, 2017).

Second, a dual-path information bottleneck architecture comprises a primary high-dimensional pathway and a compressed bottleneck pathway. The bottleneck does not enforce statistical independence as in β -VAE variants (Higgins et al., 2017), but instead, it captures *shared biological factors* through reconstruction-consistency regularization. Our architectural approach builds upon interpretable VAE concepts (Fu et al., 2025) while introducing geometric and dynamical regularization. By penalizing geometric distance between primary and bottleneck representations in Lorentz space, the architecture preserves coordinated gene expression patterns that define biological processes such as cell-cycle phase, lineage identity, and stress responses (Locatello et al., 2019).

Third, *neural ODE regularization* is integrated directly into VAE training. The encoder predicts pseudotime for each cell, and an ODE solver computes trajectory-consistent latent states by integrating learned dynamics $f_\theta(\mathbf{z}, t)$. A reconstruction loss penalizes deviations between encoder outputs and ODE predictions, enforcing temporal smoothness. In ablation experiments (Section 3.2), Lorentz regularization alone improved global topology but left trajectory inference unstable under biological noise. ODE constraints stabilize manifold estimation, improving noise resilience, trajectory coherence, and overall latent quality over geometric regularization alone.

We benchmarked LAIOR on 53 scRNA-seq and 65 scATAC-seq datasets spanning human and mouse tissues, cancer, immune responses, and development (Figure 1B; full dataset descriptions are available at <https://PeterPonyu.github.io/liora-ui> and in the Supplementary Material). We compared LAIOR with 23 baseline methods across six paradigms: classical methods (PCA, ICA, factor analysis), generative models (scVI, scGNN, scDiffusion, siVAE), contrastive learning (scGCC, CLEAR), hyperbolic embeddings (Poincaré maps, PGM, learnable Poincaré Gaussian mixture (LPGM), hyperbolic wrapped normal (HWN), scDHMAP), factor-based VAEs (β -VAE, InfoVAE, TC-VAE, DIP-VAE), and modality-specific methods (PeakVI, PoissonVI, scTour). Because our focus is on representation learning, batch correction tools such as Harmony (Korsunsky et al., 2019) and Seurat integration (Stuart et al., 2019) were not included.

We evaluated performance with 22 metrics in three groups: (i) clustering quality (normalized mutual information, adjusted Rand index, silhouette coefficient, Calinski–Harabasz index, Davies–Bouldin index, and correlation preservation), (ii) embedding fidelity for UMAP and t-SNE projections (distance correlation, local/global/overall quality), and (iii) manifold continuity (intrinsic dimensionality, spectral decay, participation ratio, anisotropy, trajectory directionality, noise resilience, core quality, and overall latent quality). As in standard single-cell workflows, the clustering analyses are anchored by labels derived

from unsupervised analysis followed by marker-based curation (Kiselev et al., 2019; Abdelaal et al., 2019).

Ablation studies show that the full model performs best when geometric, bottleneck, and dynamical regularization are used together. Across the benchmark, LAIOR improves clustering-related structure, embedding fidelity, and continuity while remaining interpretable at the latent factor level. Analyses of human bone marrow and mouse pancreatic endocrinogenesis link the learned factors to cell cycle, lineage commitment, and functional programs, and trajectory analyses on hematopoietic differentiation and perturbation datasets show that LAIOR can resolve branching structure while separating biological dynamics from technical confounding.

LAIOR is designed to work across scRNA-seq and scATAC-seq without architectural changes, although it is not intended for paired multi-omics integration. An interactive web portal (Figure 1C) provides access to the models, datasets, metrics, and trained weights to support reproducibility and reuse.

2 Materials and methods

2.1 Model architecture overview

LAIOR (Lorentz attentive interpretable ODE-regularized VAE) is a deep generative model integrating variational inference, geometric regularization, information bottleneck theory, and continuous dynamics modeling for single-cell trajectory analysis. The architecture comprises four core components: (1) a Transformer-based encoder with self-attention mechanisms for capturing long-range gene dependencies, (2) Lorentz manifold geometric regularization, (3) an information bottleneck for coordinated biological program discovery, and (4) neural ODE-based trajectory regularization.

2.1.1 Module dataflow

The four components are arranged in a single forward pass with parallel regularization branches rather than sequential stages. Explicitly, for each input cell $\mathbf{x} \in \mathbb{R}^d$: (i) the *encoder* produces a pair of Gaussian parameters $(\boldsymbol{\mu}(\mathbf{x}), \boldsymbol{\sigma}(\mathbf{x}))$ from which the primary latent $\mathbf{z} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\sigma}^2 I) \in \mathbb{R}^l$ is sampled; (ii) the same encoder output feeds a *bottleneck projection* producing the compressed latent $\mathbf{l}_d \in \mathbb{R}^i$ ($i \ll l$); (iii) $\mathbf{z}$ and $\mathbf{l}_d$ are mapped to the Lorentz hyperboloid *via* the exponential map, and the *Lorentz loss* $\mathcal{L}_{\text{geom}} = d_{\text{HL}}(\mathbf{z}_L, \mathbf{l}_{dL})$ regularizes their geometric alignment; (iv) the *ODE branch* integrates the latent dynamics $d\mathbf{z}/dt = f_\theta(\mathbf{z}, t)$ over the encoder-predicted pseudotime, and the *ODE loss* $\mathcal{L}_{\text{ODE}}$ penalizes deviations between encoder-produced latents and ODE-integrated latents; (v) finally, the *decoder* consumes both $\mathbf{z}$ and $\mathbf{l}_d$ through parallel reconstruction heads producing per-cell gene-wise likelihood parameters. The Lorentz, ODE, KL, and information bottleneck terms are combined *via* Equation 43 as independent regularizers on a single shared latent, not as sequential downstream modules. This clarifies that removing any one regularization term (ablation) simply deletes one additive loss component without changing the encoder/decoder topology, exactly as reported in Figure 2.

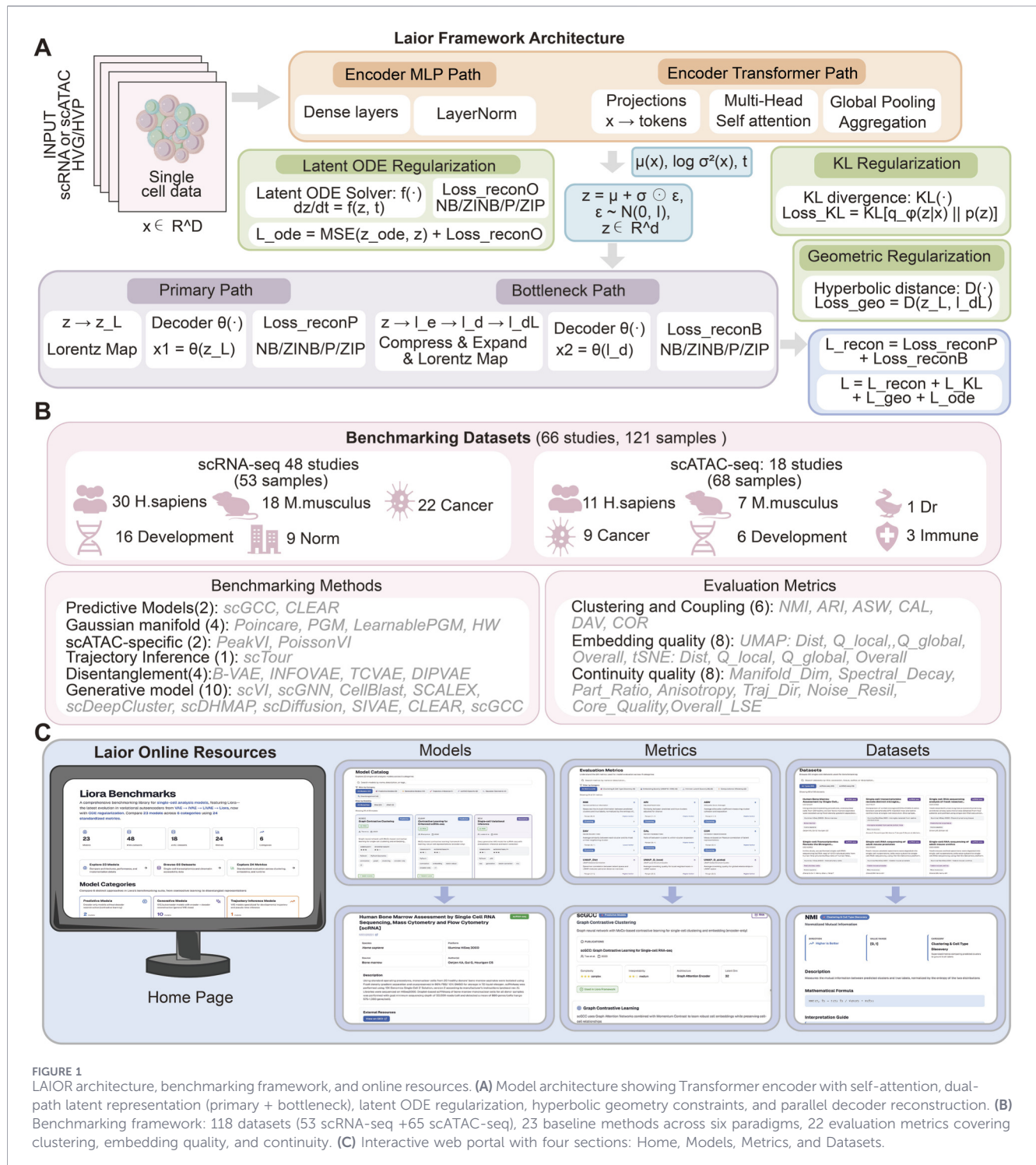


FIGURE 1

LAIOR architecture, benchmarking framework, and online resources. (A) Model architecture showing Transformer encoder with self-attention, dual-path latent representation (primary + bottleneck), latent ODE regularization, hyperbolic geometry constraints, and parallel decoder reconstruction. (B) Benchmarking framework: 118 datasets (53 scRNA-seq + 65 scATAC-seq), 23 baseline methods across six paradigms, 22 evaluation metrics covering clustering, embedding quality, and continuity. (C) Interactive web portal with four sections: Home, Models, Metrics, and Datasets.

2.1.2 Symbol glossary

For reference, Table 1 consolidates the core symbols used throughout the Methods section.

2.1.3 Encoder architecture

The encoder maps high-dimensional gene expression data $x \in \mathbb{R}^d$ to a latent probability distribution $q_\phi(z|x)$,

where d denotes the number of genes or chromatin features. LAIOR implements a Transformer-based architecture as the primary encoder, with a multilayer perceptron (MLP) fallback option for computational efficiency on smaller datasets.

2.1.3.1 Transformer encoder (default)

For attention-based encoding, input features are projected to a sequence of k tokens (Equation 1):

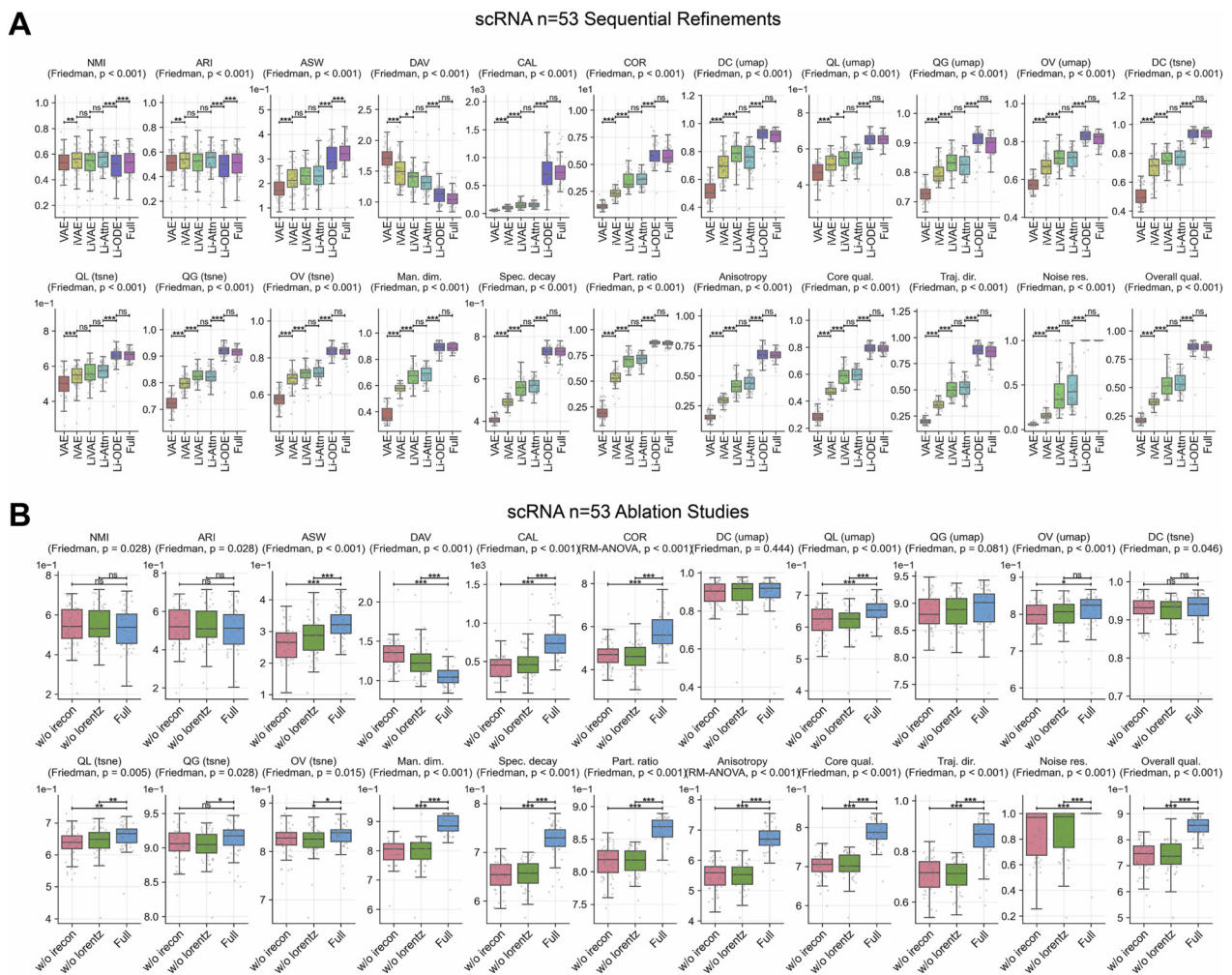


FIGURE 2

Module and regularization combination analysis: sequential refinement and targeted ablation across 53 scRNA-seq datasets. **(A)** Progressive component additions: VAE (baseline) → iVAE (information bottleneck) → LiVAE (Lorentz geometry) → Li-Attn (Transformer encoder) → Li-ODE (MLP encoder with ODE regularization) → LAIOR (Transformer encoder with full integration). Note: Li-Attn and Li-ODE represent parallel architectural branches exploring encoder architectures; LAIOR combines Transformer attention with a complete regularization framework. Evaluated across 22 metrics defined in Materials and Methods. Friedman tests confirm significant differences across configurations ($p < 0.001$). **(B)** Ablation study comparing w/o Lorentz (LAIOR minus geometric regularization), w/o irecon (LAIOR minus information bottleneck), and full LAIOR. Statistical comparisons via Wilcoxon signed-rank tests with Bonferroni correction: * ($p < 0.05$), ** ($p < 0.01$), *** ($p < 0.001$), ns (not significant).

$$\mathbf{X}_{\text{seq}} = \text{Reshape}(\mathbf{W}_{\text{proj}}\mathbf{x} + \mathbf{b}_{\text{proj}}, (k, e)), \quad (1)$$

where $\mathbf{W}_{\text{proj}} \in \mathbb{R}^{k \times d}$, $\mathbf{b}_{\text{proj}} \in \mathbb{R}^{k \times e}$, and e is the embedding dimension, yielding $\mathbf{X}_{\text{seq}} \in \mathbb{R}^{k \times e}$. This sequence is processed through n Transformer layers, where each layer $j \in \{1, \dots, n\}$ (Equations 2, 3).

$$\mathbf{X}_{\text{attn}}^{(j)} = \text{LayerNorm}(\mathbf{X}_{\text{seq}}^{(j-1)} + \text{MultiHeadAttention}(\mathbf{X}_{\text{seq}}^{(j-1)})), \quad (2)$$

$$\mathbf{X}_{\text{seq}}^{(j)} = \text{LayerNorm}(\mathbf{X}_{\text{attn}}^{(j)} + \text{FFN}(\mathbf{X}_{\text{attn}}^{(j)})), \quad (3)$$

where $\mathbf{X}_{\text{seq}}^{(0)} = \mathbf{X}_{\text{seq}}$. Multi-head self-attention with m heads computes Equation 4:

$$\text{MultiHeadAttention}(\mathbf{X}) = \text{Concat}(\text{head}_1, \dots, \text{head}_m)\mathbf{W}^O, \quad (4)$$

where each head is defined using Equation 5:

$$\text{head}_i = \text{softmax}\left(\frac{\mathbf{Q}_i\mathbf{K}_i^T}{\sqrt{e/m}}\right)\mathbf{V}_i, \quad (5)$$

with $\mathbf{Q}_i = \mathbf{X}\mathbf{W}_i^Q$, $\mathbf{K}_i = \mathbf{X}\mathbf{W}_i^K$, $\mathbf{V}_i = \mathbf{X}\mathbf{W}_i^V$, and $\mathbf{W}_i^Q, \mathbf{W}_i^K, \mathbf{W}_i^V \in \mathbb{R}^{e \times (e/m)}$. The position-wise feedforward network applies (Equation 6):

$$\text{FFN}(\mathbf{X}) = \text{ReLU}(\mathbf{X}\mathbf{W}_1^{\text{FFN}} + \mathbf{b}_1^{\text{FFN}})\mathbf{W}_2^{\text{FFN}} + \mathbf{b}_2^{\text{FFN}}. \quad (6)$$

The final output $\mathbf{X}_{\text{seq}}^{(n)} \in \mathbb{R}^{k \times e}$ is aggregated via mean pooling and projected to latent parameters using Equation 7:

$$[\mu, \log\sigma^2] = \mathbf{W}_{\text{out}} \cdot \frac{1}{k} \sum_{i=1}^k \mathbf{X}_{\text{seq},i}^{(n)} + \mathbf{b}_{\text{out}}, \quad (7)$$

where $\mathbf{W}_{\text{out}} \in \mathbb{R}^{2 \times e}$ and $\mathbf{b}_{\text{out}} \in \mathbb{R}^2$.

TABLE 1 Symbol glossary for LAIOR.

Symbol	Space/Type	Meaning
$\mathbf{x}$	$\mathbb{R}^d$ (counts)	Raw gene expression vector for 1 cell (input to encoder)
μ, σ	$\mathbb{R}^l$ each	Posterior mean and standard deviation output by the encoder
$\mathbf{z}$	$\mathbb{R}^l$	Primary latent sample drawn from $\mathcal{N}(\mu, \sigma^2 I)$
$\mathbf{l}_d$	$\mathbb{R}^l, i < l$	Compressed information bottleneck latent, derived from the same encoder
$\mathbf{z}_\mathcal{L}, \mathbf{l}_{d\mathcal{L}}$	$\mathbb{H}^l$	Hyperbolic projections of $\mathbf{z}$ and $\mathbf{l}_d$ (exponential map)
$\mathbf{v}_{\text{tangent}}$	$\mathbb{R}^{l+1}$	Clipped tangent-space vector fed into the exponential map
t	$[0, 1]$	Encoder-predicted self-supervised pseudotime scalar per cell
f_θ	$\mathbb{R}^l \times \mathbb{R} \rightarrow \mathbb{R}^l$	ODE drift function (neural network with parameters θ)
$\mathbf{z}_{\text{ODE}}(t)$	$\mathbb{R}^l$	Latent state obtained by integrating f_θ from initial condition
f_ϕ	Encoder network	Encoder parameters (bundled under ϕ)
g_ψ	Decoder network	Decoder parameters (bundled under ψ)
$\lambda_{\text{recon/irecon/geom/ODE}}, \beta$	$\mathbb{R}_{\geq 0}$	Loss-term weighting hyperparameters (Equation 43)
k, e	Integers	Transformer token count (k) and embedding dim. (e)
n	Integer	Number of Transformer encoder layers

2.1.3.2 MLP encoder (alternative)

The default architecture consists of two hidden layers with ReLU activations and layer normalization. Let h denote the hidden dimension and let l denote the latent dimension. The encoder is defined by Equations 8–10:

$$\mathbf{h}_1 = \text{ReLU}(\text{LayerNorm}(\mathbf{W}_1 \mathbf{x} + \mathbf{b}_1)), \quad (8)$$

$$\mathbf{h}_2 = \text{ReLU}(\text{LayerNorm}(\mathbf{W}_2 \mathbf{h}_1 + \mathbf{b}_2)), \quad (9)$$

$$[\mu, \log \sigma^2] = \mathbf{W}_3 \mathbf{h}_2 + \mathbf{b}_3, \quad (10)$$

where $\mathbf{W}_1 \in \mathbb{R}^{h \times d}$, $\mathbf{W}_2 \in \mathbb{R}^{h \times h}$, $\mathbf{W}_3 \in \mathbb{R}^{2l \times h}$, and $\mathbf{b}_1, \mathbf{b}_2 \in \mathbb{R}^h$, $\mathbf{b}_3 \in \mathbb{R}^{2l}$ are bias vectors. The output is split into mean $\mu \in \mathbb{R}^l$ and log-variance $\log \sigma^2 \in \mathbb{R}^l$.

2.1.3.3 Architecture selection

All benchmarking experiments reported in Results used the Transformer encoder unless explicitly stated otherwise. The Li-ODE architectural variant (Section 3.2) employed the MLP encoder to isolate ODE regularization contributions independent of attention mechanisms. The complete LAIOR model integrates Transformer encoding with the full regularization framework.

Both encoders apply numerical stabilization to outputs (Equations 11, 12).

$$\mu \leftarrow \text{clamp}(\mu, -10, 10), \quad (11)$$

$$\sigma = \text{clamp}(\text{softplus}(\log \sigma^2) + 10^{-6}, 10^{-6}, 5.0), \quad (12)$$

where $\text{softplus}(x) = \log(1 + e^x)$. The posterior distribution $q_\phi(\mathbf{z}|\mathbf{x}) = \mathcal{N}(\mu, \text{diag}(\sigma^2))$ is sampled *via* the reparameterization trick defined in Equation 13:

$$\mathbf{z} = \mu + \sigma \odot \epsilon, \quad \epsilon \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_l), \quad (13)$$

where $\odot$ denotes element-wise multiplication.

For temporal modeling, the encoder includes a pseudotime prediction head defined by Equation 14:

$$t = \sigma_{\text{sig}}(\mathbf{w}_t^T \mathbf{h}_2 + b_t) \in [0, 1], \quad (14)$$

where $\mathbf{w}_t \in \mathbb{R}^h$, $b_t \in \mathbb{R}$, and $\sigma_{\text{sig}}(\cdot)$ is the sigmoid activation function.

2.1.4 Lorentz geometric regularization

Latent codes are embedded onto the Lorentz model of hyperbolic space, defined as Equation 15. This manifold-aware optimization context follows Riemannian adaptive optimization methods (Bécigneul and Ganeva, 2019).

$$\mathbb{H}^n = \{\mathbf{x} \in \mathbb{R}^{n+1} : \langle \mathbf{x}, \mathbf{x} \rangle_{\mathcal{L}} = -1, x_0 > 0\}, \quad (15)$$

where the Lorentzian inner product is defined by Equation 16:

$$\langle \mathbf{x}, \mathbf{y} \rangle_{\mathcal{L}} = -x_0 y_0 + \sum_{i=1}^n x_i y_i. \quad (16)$$

2.1.4.1 Exponential map

Latent codes $\mathbf{z} \in \mathbb{R}^l$ are mapped from the tangent space at the origin $\mathbf{o} = [1, 0, \dots, 0]^T \in \mathbb{H}^l$ to the manifold. The tangent space is defined by Equation 17:

$$T_{\mathbf{o}} \mathbb{H}^l = \{\mathbf{v} \in \mathbb{R}^{l+1} : v_0 = 0\}. \quad (17)$$

Given latent code $\mathbf{z} \in \mathbb{R}^l$, we construct the tangent vector (Equations 18, 19).

$$\mathbf{z}_{\text{clip}} = \text{clamp}(\mathbf{z}, -5, 5), \quad (18)$$

$$\mathbf{v}_{\text{tangent}} = [0, z_{\text{clip},1}, z_{\text{clip},2}, \dots, z_{\text{clip},l}]^T \in \mathbb{R}^{l+1}. \quad (19)$$

Compute the spatial norm using Equation 20:

$$r = \text{clamp}\left(\sqrt{\sum_{i=1}^l v_{\text{tangent},i}^2}, 0, 15\right). \quad (20)$$

The exponential map $\exp_o: T_o\mathbb{H}^l \rightarrow \mathbb{H}^l$ is defined by Equation 21:

$$\exp_o(\mathbf{v}_{\text{tangent}}) = \begin{bmatrix} \cosh(r) \\ \sinh(r) \cdot \frac{\mathbf{v}_{\text{spatial}}}{r} \end{bmatrix}, \quad (21)$$

where $\mathbf{v}_{\text{spatial}} = [v_{\text{tangent},1}, \dots, v_{\text{tangent},l}]^T \in \mathbb{R}^l$. For $r < 10^{-8}$, we set $\mathbf{v}_{\text{spatial}} = \mathbf{0}$ to avoid division by zero.

2.1.4.2 Hyperbolic distance

The distance between points $\mathbf{x}, \mathbf{y} \in \mathbb{H}^l$ is defined by Equation 22:

$$d_{\mathbb{H}}(\mathbf{x}, \mathbf{y}) = \text{arccosh}(\max(1 + 10^{-8}, -\langle \mathbf{x}, \mathbf{y} \rangle_{\mathcal{L}})). \quad (22)$$

For numerical stability when $-\langle \mathbf{x}, \mathbf{y} \rangle_{\mathcal{L}} > 10^4$, we use the asymptotic approximation defined by Equation 23:

$$d_{\mathbb{H}}(\mathbf{x}, \mathbf{y}) \approx \log(2 \cdot (-\langle \mathbf{x}, \mathbf{y} \rangle_{\mathcal{L}})). \quad (23)$$

Manifold embeddings for the primary and bottleneck latents are defined by Equations 24, 25:

$$\mathbf{z}_{\mathcal{L}} = \exp_o([0, \text{clamp}(\mathbf{z}, -5, 5)]^T) \in \mathbb{H}^l, \quad (24)$$

$$\mathbf{l}_{d\mathcal{L}} = \exp_o([0, \text{clamp}(\mathbf{l}_d, -5, 5)]^T) \in \mathbb{H}^l. \quad (25)$$

2.1.5 Information bottleneck

The information bottleneck pathway compresses the latent code through a lower-dimensional bottleneck with dimension $i < l$, designed to extract shared biological programs. Unlike disentanglement approaches that enforce statistical independence, LAIOR's bottleneck preserves coordinated variation patterns through geometric consistency regularization, enabling the capture of interconnected biological processes such as cell-cycle coordination, lineage-coupled differentiation programs, and stress-response modules.

2.1.5.1 Compression

Linear encoding to bottleneck (Equation 26):

$$\mathbf{l}_e = \mathbf{W}_e \mathbf{z} + \mathbf{b}_e, \quad (26)$$

where $\mathbf{W}_e \in \mathbb{R}^{i \times l}$, $\mathbf{b}_e \in \mathbb{R}^i$, and $\mathbf{l}_e \in \mathbb{R}^i$.

2.1.5.2 Decompression

Linear decoding from bottleneck (Equation 27):

$$\mathbf{l}_d = \mathbf{W}_d \mathbf{l}_e + \mathbf{b}_d, \quad (27)$$

where $\mathbf{W}_d \in \mathbb{R}^{l \times i}$, $\mathbf{b}_d \in \mathbb{R}^l$, and $\mathbf{l}_d \in \mathbb{R}^l$.

2.1.5.3 Geometric integration

The decompressed latent is mapped to the Lorentz manifold using Equation 28:

$$\mathbf{l}_{d\mathcal{L}} = \exp_o([0, \text{clamp}(\mathbf{l}_d, -5, 5)]^T) \in \mathbb{H}^l. \quad (28)$$

Both primary latent $\mathbf{z}$ and bottleneck latent $\mathbf{l}_d$ are independently decoded through the same decoder network to produce parallel reconstructions.

2.1.6 Neural ODE latent regularization

For temporal modeling, latent dynamics are governed by the following ordinary differential equation (Equation 29):

$$\frac{d\mathbf{z}}{dt} = f_{\theta}(\mathbf{z}, t), \quad (29)$$

where $\mathbf{z}(t) \in \mathbb{R}^l$ represents the latent state at pseudotime $t \in [0, 1]$, and $f_{\theta}: \mathbb{R}^l \times \mathbb{R} \rightarrow \mathbb{R}^l$ is a parametric function.

2.1.6.1 ODE function

The time-conditioned ODE function is a two-layer neural network with hidden dimension h_{ode} (Equations 30, 31).

$$\mathbf{h}_{\text{ode}} = \text{ELU}(\mathbf{W}_1 [\mathbf{z}; t] + \mathbf{b}_1), \quad (30)$$

$$f_{\theta}(\mathbf{z}, t) = \mathbf{W}_2 \mathbf{h}_{\text{ode}} + \mathbf{b}_2, \quad (31)$$

where $[\mathbf{z}; t] \in \mathbb{R}^{l+1}$ denotes concatenation, $\mathbf{W}_1 \in \mathbb{R}^{h_{\text{ode}} \times (l+1)}$, $\mathbf{b}_1 \in \mathbb{R}^{h_{\text{ode}}}$, $\mathbf{W}_2 \in \mathbb{R}^{l \times h_{\text{ode}}}$, $\mathbf{b}_2 \in \mathbb{R}^l$, and $\text{ELU}(x) = x$ if $x > 0$ and $\text{ELU}(x) = \alpha(e^x - 1)$ if $x \leq 0$, with $\alpha = 1.0$.

2.1.6.2 Trajectory integration

For N cells with pseudotime predictions $\{t_1, \dots, t_N\} \subset [0, 1]$, we sort cells by pseudotime and remove duplicates to obtain a strictly increasing sequence $\tilde{t}_1 < \tilde{t}_2 < \dots < \tilde{t}_M$ where $M \leq N$.

Setting the initial condition $\mathbf{z}_0 = \mathbf{z}(\tilde{t}_1)$, we solve the initial value problem using Equation 32:

$$\begin{cases} \frac{d\mathbf{z}}{dt} = f_{\theta}(\mathbf{z}, t), \\ \mathbf{z}(\tilde{t}_1) = \mathbf{z}_0, \end{cases} \quad (32)$$

using the fourth-order Runge-Kutta (RK4) method. At step n , we compute (Equations 33–37)

$$\mathbf{k}_1 = f_{\theta}(\mathbf{z}_n, t_n), \quad (33)$$

$$\mathbf{k}_2 = f_{\theta}\left(\mathbf{z}_n + \frac{\Delta t}{2} \mathbf{k}_1, t_n + \frac{\Delta t}{2}\right), \quad (34)$$

$$\mathbf{k}_3 = f_{\theta}\left(\mathbf{z}_n + \frac{\Delta t}{2} \mathbf{k}_2, t_n + \frac{\Delta t}{2}\right), \quad (35)$$

$$\mathbf{k}_4 = f_{\theta}(\mathbf{z}_n + \Delta t \mathbf{k}_3, t_n + \Delta t), \quad (36)$$

$$\mathbf{z}_{n+1} = \mathbf{z}_n + \frac{\Delta t}{6} (\mathbf{k}_1 + 2\mathbf{k}_2 + 2\mathbf{k}_3 + \mathbf{k}_4), \quad (37)$$

where $\Delta t = \tilde{t}_{n+1} - \tilde{t}_n$ is the step size, to obtain ODE-solved latents $\{\mathbf{z}_{\text{ODE}}(\tilde{t}_1), \mathbf{z}_{\text{ODE}}(\tilde{t}_2), \dots, \mathbf{z}_{\text{ODE}}(\tilde{t}_M)\}$.

This creates three parallel reconstruction pathways.

1. Primary path: Reconstruction using encoder latent $\mathbf{z}$
2. Bottleneck path: Reconstruction using decompressed latent $\mathbf{l}_d$
3. ODE path: Reconstruction using trajectory-regularized latent $\mathbf{z}_{\text{ODE}}$

2.1.7 Decoder architecture

The decoder reconstructs gene expression from latent codes *via* a two-layer feedforward network given input latent code $\mathbf{z}_{\text{in}} \in \mathbb{R}^l$ (which may be $\mathbf{z}$, $\mathbf{l}_d$, or $\mathbf{z}_{\text{ODE}}$) (Equations 38–40).

$$\mathbf{g}_1 = \text{ReLU}(\text{LayerNorm}(\mathbf{W}_1^{(d)} \mathbf{z}_{\text{in}} + \mathbf{b}_1^{(d)})), \quad (38)$$

$$\mathbf{g}_2 = \text{ReLU}(\text{LayerNorm}(\mathbf{W}_2^{(d)} \mathbf{g}_1 + \mathbf{b}_2^{(d)})), \quad (39)$$

$$\boldsymbol{\pi} = \text{softmax}(\mathbf{W}_3^{(d)} \mathbf{g}_2 + \mathbf{b}_3^{(d)}), \quad (40)$$

where $\mathbf{W}_1^{(d)} \in \mathbb{R}^{h \times l}$, $\mathbf{W}_2^{(d)} \in \mathbb{R}^{h \times h}$, $\mathbf{W}_3^{(d)} \in \mathbb{R}^{d \times h}$, and $\mathbf{b}_1^{(d)}, \mathbf{b}_2^{(d)} \in \mathbb{R}^h$, $\mathbf{b}_3^{(d)} \in \mathbb{R}^d$. The softmax output $\boldsymbol{\pi} \in \Delta^{d-1}$ (the $(d-1)$ -dimensional probability simplex) serves as the mean parameter for likelihood functions.

A learned per-gene dispersion parameter $\boldsymbol{\theta} \in \mathbb{R}^d$ models overdispersion for negative binomial likelihoods. For zero-inflated models, an auxiliary network predicts zero-inflation logits (Equations 41, 42).

$$\mathbf{h}_{\text{drop}} = \text{ReLU}(\mathbf{W}_{\text{drop},1} \mathbf{z}_{\text{in}} + \mathbf{b}_{\text{drop},1}), \quad (41)$$

$$\boldsymbol{\rho} = \mathbf{W}_{\text{drop},2} \mathbf{h}_{\text{drop}} + \mathbf{b}_{\text{drop},2}, \quad (42)$$

where $\mathbf{W}_{\text{drop},1} \in \mathbb{R}^{h \times l}$, $\mathbf{b}_{\text{drop},1} \in \mathbb{R}^h$, $\mathbf{W}_{\text{drop},2} \in \mathbb{R}^{d \times h}$, $\mathbf{b}_{\text{drop},2} \in \mathbb{R}^d$, and $\boldsymbol{\rho} \in \mathbb{R}^d$ give logit-space dropout probabilities.

Weight matrices are initialized using Xavier normal initialization. Bias terms are initialized to 0.01.

2.2 Training procedure and optimization

2.2.1 Data preprocessing

Raw count matrices were extracted and validated for integer-valued counts. Log-transformation $\log_2(x+1)$ was applied, followed by adaptive clipping. High-sparsity datasets (>95% zeros) were clipped to $[-5, 5]$; all other datasets were clipped to $[-10, 10]$. The two clipping ranges reflect the empirical distributions of post- $\log_2$ expression values: on typical single-cell data, more than 99.5% of non-zero $\log_2(x+1)$ values fall within $[-10, 10]$, and on high-sparsity profiles where the signal is concentrated in fewer genes, the tighter $[-5, 5]$ window better suppresses outlier-driven gradient variance without truncating biologically meaningful variation. This clipping step is distinct from the z -score outlier clipping at ± 10 standard deviations applied internally by the training data loader (Section 2.5.2): the former acts on the raw $\log_2$ -scale values before normalization, the latter acts on standardized values during batch construction. Data were partitioned into training (70%), validation (15%), and test (15%) sets at the *cell* level with stratified random sampling (random seed 42), ensuring that each cell type is proportionally represented across splits; a complementary dataset-level held-out-dataset protocol, in which entire datasets are excluded from training and evaluated in zero-shot mode, is reported in [Supplementary Material](#).

2.2.2 Optimization

Model parameters were optimized using the Adam optimizer with learning rate $\eta = 10^{-4}$, batch size $B = 128$, gradient clipping at maximum norm 1.0, and early stopping with patience of 25 epochs

based on validation loss. For benchmarking experiments, datasets were subsampled to 3,000 cells.

2.2.3 Loss function

The total loss function is (Equation 43):

$$\mathcal{L}_{\text{total}} = \lambda_{\text{recon}} \mathcal{L}_{\text{recon}} + \lambda_{\text{irecon}} \mathcal{L}_{\text{irecon}} + \lambda_{\text{geom}} \mathcal{L}_{\text{geom}} + \beta \mathcal{L}_{\text{KL}} + \lambda_{\text{ODE}} \mathcal{L}_{\text{ODE}}, \quad (43)$$

where $\lambda_{\text{recon}}, \lambda_{\text{irecon}}, \lambda_{\text{geom}}, \beta, \lambda_{\text{ODE}}$ are weighting hyperparameters.

2.2.3.1 Package defaults vs. per-experiment values

For clarity, we note that the published LAIOR Python package constructor exposes these loss weights as keyword arguments with *intentionally zero defaults* for all regularization terms except KL ($\beta = 1.0$) and the primary reconstruction ($\lambda_{\text{recon}} = 1.0$). Specifically, the package defaults are $\lambda_{\text{irecon}} = \lambda_{\text{geom}} = \lambda_{\text{dip}} = \lambda_{\text{tc}} = \lambda_{\text{info}} = 0$, reflecting a deliberate “bottom-up ablation” design in which the user must be explicitly enable each regularizer. This design supports the sequential refinement analysis reported in [Figure 2](#): starting from a plain VAE (all extras set to 0) and adding terms one at a time, it isolates each component’s contribution. The *actual values used in each experiment reported in this article* are not the package defaults: full LAIOR uses $\lambda_{\text{recon}} = 1.0$ and $\lambda_{\text{geom}} = 5.0$ (i.e., the “high curvature” setting identified in [Section 3.4](#)) unless otherwise noted. Hyperparameter sweeps in [Section 3.4](#) explicitly vary these values. There is therefore no discrepancy between the source code and the article: the zero entries in the package signature are the starting points of the ablation series, not the final values.

2.2.3.2 Reconstruction loss

Primary reconstruction loss is defined by [Equation 44](#):

$$\mathcal{L}_{\text{recon}} = -\mathbb{E}_{q_{\theta}(\mathbf{z}|\mathbf{x})} [\log p_{\theta}(\mathbf{x}|\mathbf{z})]. \quad (44)$$

Information bottleneck reconstruction loss is defined by [Equation 45](#):

$$\mathcal{L}_{\text{irecon}} = -\mathbb{E}_{q_{\theta}(\mathbf{z}|\mathbf{x})} [\log p_{\theta}(\mathbf{x}|\mathbf{l}_d)]. \quad (45)$$

Predicted rates are scaled by library size: $\mu_j = \pi_j \cdot s$ where $s = \sum_{j=1}^d x_j$ is the total count for the cell.

For the negative binomial (NB) likelihood, the log-probability is defined by [Equation 46](#):

$$\log p(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\theta}) = \sum_{j=1}^d \left[\log \Gamma(x_j + \theta_j) - \log \Gamma(\theta_j) - \log \Gamma(x_j + 1) + \theta_j \log \frac{\theta_j}{\theta_j + \mu_j} + x_j \log \frac{\mu_j}{\theta_j + \mu_j} \right], \quad (46)$$

where $\theta_j = \exp(\text{disp}_j)$ with learned dispersion parameters disp_j .

Zero-inflated negative binomial (ZINB) extends NB *via* (Equations 47, 48)

$$p(x_j = 0) = \pi_{\text{drop},j} + (1 - \pi_{\text{drop},j}) \cdot \text{NB}(0|\mu_j, \theta_j), \quad (47)$$

$$p(x_j > 0) = (1 - \pi_{\text{drop},j}) \cdot \text{NB}(x_j|\mu_j, \theta_j), \quad (48)$$

where $\pi_{\text{drop},j} = \sigma_{\text{sig}}(\rho_j)$. Poisson and zero-inflated Poisson (ZIP) likelihoods follow analogously with $\theta_j \rightarrow \infty$.

2.2.3.3 Geometric regularization

The distance penalty on the Lorentz manifold is defined by Equation 49:

$$\mathcal{L}_{\text{geom}} = d_{\mathbb{H}}(\mathbf{z}_L, \mathbf{l}_{dL}). \quad (49)$$

2.2.3.4 KL divergence

Between posterior $q_{\phi}(\mathbf{z}|\mathbf{x}) = \mathcal{N}(\boldsymbol{\mu}, \text{diag}(\boldsymbol{\sigma}^2))$ and standard Gaussian prior $p(\mathbf{z}) = \mathcal{N}(\mathbf{0}, \mathbf{I}_l)$ (Equation 50):

$$\mathcal{L}_{\text{KL}} = \frac{1}{2} \sum_{j=1}^l (\mu_j^2 + \sigma_j^2 - \log \sigma_j^2 - 1). \quad (50)$$

2.2.3.5 ODE trajectory consistency

The mean squared error between encoder latents and ODE-predicted latents is defined by Equation 51:

$$\mathcal{L}_{\text{ODE}} = \frac{1}{M} \sum_{i=1}^M \|\mathbf{z}(\tilde{t}_i) - \mathbf{z}_{\text{ODE}}(\tilde{t}_i)\|_2^2. \quad (51)$$

The default weight coefficients are $\lambda_{\text{recon}} = 1.0$, $\lambda_{\text{irecon}} = 1.0$, $\lambda_{\text{geom}} = 5.0$, $\beta = 1.0$, and $\lambda_{\text{ODE}} = 1.0$ unless otherwise specified.

2.2.4 Architecture specifications

2.2.4.1 Default configuration

Transformer encoder: $k = 32$ tokens, embedding dimension $e = 64$, $m = 4$ attention heads, $n = 2$ layers, FFN hidden dimension $4e$. Decoder: two hidden layers (dimension $h = 128$) with ReLU activations, input dimension d (dataset-dependent), latent dimension $l = 10$. Information bottleneck dimension: $i = 2$. ODE function hidden dimension: $h_{\text{ode}} = 64$.

2.2.4.2 MLP configuration (when applicable)

Encoder/Decoder: two hidden layers (dimension $h = 128$); all other hyperparameters are identical to the default.

Numerical stabilization: encoder mean outputs are clamped to $[-10, 10]$, posterior standard deviations are constrained to $[10^{-6}, 5.0]$, and Lorentz tangent-space inputs are clamped to $[-5, 5]$ before the exponential map.

2.2.5 Pseudotime prediction and ODE integration

For ODE-augmented models, the encoder predicts pseudotime *via* a sigmoid-activated linear head, normalized to $[0, 1]$. We emphasize that pseudotime prediction is *self-supervised*: no externally computed pseudotime label (e.g., diffusion pseudotime, Monocle) is provided as a target. Instead, pseudotime values emerge from the joint optimization of the reconstruction, KL, Lorentz, information bottleneck, and neural ODE trajectory-consistency losses (Equation 43). The ODE reconstruction loss $\mathcal{L}_{\text{ODE}} = \frac{1}{M} \sum_i \|\mathbf{z}(\tilde{t}_i) - \mathbf{z}_{\text{ODE}}(\tilde{t}_i)\|_2^2$ penalizes inconsistencies between encoder-produced latents and ODE-integrated latents, implicitly selecting pseudotime orderings that admit a smooth trajectory. This makes LAIOR directly applicable to new datasets without requiring external pseudotime annotation at inference time. Cells are sorted by

pseudotime with duplicate time values removed, and the ODE is solved using RK4 integration on the CPU for numerical stability.

2.2.6 Training protocol

Models were trained with validation evaluation every five epochs. Early stopping was triggered when validation loss failed to improve for 25 consecutive validation checks (equivalent to 125 training epochs), at which point the best checkpoint was restored. This patience parameter balances exploration of the loss landscape with computational efficiency.

2.2.7 Computational environment

All experiments were conducted on a single NVIDIA RTX 5090 GPU (24 GB VRAM) using PyTorch 2.0 with single-precision floating-point arithmetic (FP32). Datasets were subsampled to 3,000 cells for benchmarking, with a batch size of 128.

2.3 Evaluation metrics

2.3.1 Clustering quality metrics

We assessed biological population structure using five standard metrics and one novel metric.

2.3.1.1 Standard metrics

Normalized mutual information (NMI) and adjusted Rand index (ARI) measure agreement between predicted clusters and reference pseudo-labels, with values near 1 indicating strong correspondence. We adopt a unified unsupervised evaluation protocol because single-cell transcriptomic data are inherently unsupervised. Cell-type annotations depend on the specific clustering algorithm, resolution parameter, and marker-gene panel used by each laboratory, and no universally accepted ground truth exists. Reference pseudo-labels are generated by applying K -means clustering ($K = l$, where l is the latent dimensionality) to the preprocessed expression matrix *before* any representation learning method is applied. For each method under evaluation, K -means is independently applied to the method's latent space, and NMI/ARI quantify how well the method's latent clusters recover the structure identified in the original expression space. This design ensures that (i) no method-specific annotation bias contaminates the comparison, and (ii) the reference is consistently derived from the same preprocessing pipeline across all datasets. The average silhouette width (ASW) and the Calinski-Harabasz index (CAL) quantify cluster cohesion and separation (higher is better), while the Davies-Bouldin index (DAV) measures average cluster similarity (lower is better). ASW, CAL, and DAV are computed in the method's latent space using its own K -means assignments, measuring intrinsic latent-space cluster quality independently of the reference. All metrics were computed using standard implementations in scikit-learn.

2.3.1.2 Coupling degree (COR)

We introduce the coupling degree (COR) metric, Equation 52, to quantify preservation of interdependent biological programs:

$$\text{COR} = \frac{1}{l(l-1)} \sum_{i=1}^l \sum_{j \neq i}^l |\rho_{ij}|, \quad (52)$$

where ρ_{ij} is the Pearson correlation between latent dimensions i and j , and l is the latent space dimensionality. Higher COR values indicate stronger coupling, reflecting coordinated gene expression programs essential for continuous differentiation trajectories.

2.3.1.3 A note on circular reasoning in embedding-quality evaluation

We acknowledge that embedding quality metrics that rely on neighborhood preservation (Q_{local} , Q_{global} , distance correlation) can favor methods whose training objectives already optimize for neighborhood structure, introducing a potential circularity. We mitigate this in three ways. First, clustering metrics (NMI, ARI) compare each method's K -means assignments against unsupervised reference pseudo-labels derived from the raw data, providing an independent signal that is independent of the method's training objective. Second, our intrinsic manifold metrics (Section 2.3.3) quantify spectral and geometric properties of the latent representation itself without reference to any neighborhood graph, providing a complementary signal orthogonal to any neighborhood-preserving training objective. Third, cross-dataset consistency, which is the fact that LAIOR's advantage over baselines is preserved across 53 scRNA-seq, 65 scATAC-seq, and eight application datasets spanning diverse tissues and species, provides strong evidence against circularity: a method benefiting from objective-metric circularity would not systematically outperform on independent intrinsic metrics and independent datasets simultaneously. Batch-effect-aware evaluation metrics further motivate this caution (Büttner et al., 2019).

2.3.1.4 On dataset-level generalization versus within-dataset evaluation

A single-cell dataset represents one biological snapshot: a specific tissue, developmental stage, and experimental condition. Within-dataset evaluation (train/test split at the cell level) measures a model's ability to reconstruct the manifold of *that particular* snapshot. However, it does not test whether the learned inductive biases transfer to fundamentally different biological systems. True generalization requires consistent performance across diverse datasets spanning different tissues (bone marrow, pancreas, lung, muscle, liver), species (human, mouse), and biological processes (development, cancer, perturbation). Our benchmark evaluates LAIOR across 53 scRNA-seq and 65 scATAC-seq datasets covering this diversity; the additional eight application datasets in this revision further strengthen the generalization claim by including both well-characterized developmental systems (cd34 hematopoiesis, endo pancreatic endocrinogenesis) and novel perturbation contexts (GSE278673 hematopoietic knockout, GSE277292 wildtype/knockout comparison). The fact that LAIOR's architectural advantages, including Lorentz geometry for hierarchical structure and neural ODE for temporal consistency, transfer across these heterogeneous biological contexts demonstrates dataset-level generalization that cannot be inferred from within-dataset cell-level splits alone.

2.3.2 Dimensionality reduction embedding-quality metrics

We evaluated how effectively latent representations $\mathbf{Z} \in \mathbb{R}^{n \times l}$ (where n is the number of cells) project to interpretable low-dimensional spaces (UMAP and t-SNE) while preserving biological relationships.

Distance correlation (ρ_{dist}) quantifies preservation of pairwise distance relationships using Equation 53:

$$\rho_{\text{dist}} = \rho_{\text{Spearman}}(\text{vec}(D_{\mathbf{Z}}), \text{vec}(D_{\mathbf{E}})), \quad (53)$$

where $D_{\mathbf{Z}} \in \mathbb{R}^{n \times n}$ and $D_{\mathbf{E}} \in \mathbb{R}^{n \times n}$ are pairwise Euclidean distance matrices in latent and embedding spaces, $\text{vec}(\cdot)$ vectorizes the upper triangle, and ρ_{Spearman} is the Spearman rank correlation coefficient. Higher values indicate better preservation of global structure.

Local quality (Q_{local}) and global quality (Q_{global}) measure preservation at different scales through co-ranking matrix analysis using Equations 54, 55, respectively.

$$Q_{\text{local}} = \frac{1}{K_{\text{max}}} \sum_{K=1}^{K_{\text{max}}} Q_{NX}(K), \quad (54)$$

$$Q_{\text{global}} = \frac{1}{n-1-K_{\text{max}}} \sum_{K=K_{\text{max}}+1}^{n-1} Q_{NX}(K), \quad (55)$$

where $Q_{NX}(K)$ is the normalized co-ranking quality measure at neighborhood size K , and K_{max} is the optimal local neighborhood boundary determined automatically *via* the co-ranking framework. Higher values indicate better maintenance of local neighborhoods and large-scale topology, respectively.

Overall embedding quality (Q_{embed}) combines the three components using Equation 56:

$$Q_{\text{embed}} = \frac{1}{3}(\rho_{\text{dist}} + Q_{\text{local}} + Q_{\text{global}}). \quad (56)$$

2.3.3 Intrinsic manifold quality metrics

We characterized geometric properties of the latent manifold $\mathbf{Z} \in \mathbb{R}^{n \times l}$ through spectral analysis of the covariance matrix $\mathbf{C} = \frac{1}{n-1} \mathbf{Z}^T \mathbf{Z}$, with eigenvalues $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_l$.

Manifold dimensionality (M_{dim}) measures representation compactness using Equation 57:

$$M_{\text{dim}} = 1 - \frac{d_{\text{eff}} - 1}{l - 1}, \quad (57)$$

where d_{eff} is the number of principal components explaining 95% variance. Higher values indicate more efficient encoding.

Spectral decay rate (S_{decay}) quantifies hierarchical structure clarity using Equation 58:

$$S_{\text{decay}} = \frac{1}{1 + e^{|\beta|}} \cdot \frac{\lambda_1}{\sum_{i=1}^l \lambda_i}, \quad (58)$$

where β is the slope from log-linear regression on log-eigenvalues. Higher values indicate steeper spectrum decay, reflecting a clear hierarchical organization.

Participation ratio (P_{ratio}) assesses the balance of variance distribution using Equation 59:

$$P_{\text{ratio}} = \frac{1}{l} \cdot \frac{\left(\sum_{i=1}^l \lambda_i\right)^2}{\sum_{i=1}^l \lambda_i^2}. \quad (59)$$

Higher values indicate more uniform utilization of latent dimensions, preventing dimension collapse.

Anisotropy score (A_{score}) quantifies the directional bias strength using Equation 60:

$$A_{\text{score}} = \tanh\left(\frac{\log(\lambda_1) - \log(\lambda_l + \epsilon)}{4}\right), \quad (60)$$

where $\epsilon = 10^{-8}$. Higher values indicate stronger directional structure along the dominant axes, which are essential for trajectory representation.

Trajectory directionality (T_{dir}) measures the dominance of the primary variation axis using Equation 61:

$$T_{\text{dir}} = \frac{\lambda_1}{\sum_{i=2}^l \lambda_i + \epsilon}. \quad (61)$$

Higher values indicate a single dominant trajectory direction, which is characteristic of linear differentiation processes.

Noise resilience (N_{res}) approximates the signal-to-noise ratio using Equation 62:

$$N_{\text{res}} = \min\left(\frac{\sum_{i=1}^2 \lambda_i}{\sum_{i=3}^l \lambda_i + \epsilon} \cdot \frac{1}{10}, 1\right). \quad (62)$$

Higher values indicate robust separation between signal and noise subspaces.

2.3.3.1 Composite scores

We defined two summary metrics: core intrinsic quality integrates fundamental geometric properties using Equation 63,

$$Q_{\text{core}} = \frac{1}{4} (M_{\text{dim}} + S_{\text{decay}} + P_{\text{ratio}} + A_{\text{score}}), \quad (63)$$

while overall intrinsic quality incorporates task-oriented components with empirically determined weights (α, β, γ) = (0.5, 0.3, 0.2) using Equation 64:

$$Q_{\text{overall}} = \alpha \cdot Q_{\text{core}} + \beta \cdot T_{\text{dir}} + \gamma \cdot N_{\text{res}}. \quad (64)$$

2.4 Latent factor attribution and interpretability analysis

2.4.1 Association score calculation

To identify gene signatures associated with each latent dimension, we implemented a dual-pathway attribution method combining correlation and gradient-based sensitivity analysis. For a given gene i and latent factor j , the combined association score $S_{i,j}$ was calculated using Equation 65:

$$S_{i,j} = 0.6 \cdot \rho(\mathbf{x}_i, \mathbf{v}_j) + 0.4 \cdot \tilde{g}_{i,j}, \quad (65)$$

where ρ denotes the Pearson correlation coefficient. For the *encoder pathway* (identifying discriminative markers), $\mathbf{v}_j$ represents the latent mean $\mu_j \in \mathbb{R}^n$ across all n cells, and gradients $g_{i,j} = \partial \mu_j / \partial x_i$ were computed *via* backpropagation. For the

decoder pathway (identifying reconstructive programs), correlations were computed between latent codes $z_j \in \mathbb{R}^n$ and unscaled decoder outputs $\hat{x}_i \in \mathbb{R}^n$, with gradients $g_{i,j} = \partial \hat{x}_i / \partial z_j$. In both pathways, gradients were normalized ($\tilde{g}_{i,j}$) to the range $[-1, 1]$ to preserve directionality (positive/negative regulation).

2.4.2 Latent velocity field computation and visualization

The velocity field $\dot{\mathbf{z}} = d\mathbf{z}/dt$ that underlies all trajectory visualizations in Figures 7–9 is obtained by a *direct forward evaluation* of the learned ODE drift network f_θ on each encoder output, not by numerical differencing of the ODE-integrated trajectory $\mathbf{z}_{\text{ODE}}(t)$. For each cell, the encoder produces the posterior quintuple $(\mathbf{q}_z, \mathbf{q}_m, \mathbf{q}_s, \mathbf{n}, t)$, and the drift $f_\theta(\mathbf{q}_z, t)$ is evaluated in a single forward pass, yielding a per-cell velocity tensor of shape $(N_{\text{cells}}, d_{\text{latent}})$. Per-cell latent velocities are then projected onto the two-dimensional UMAP embedding through a cosine-similarity transition matrix computed in latent space.

For visual rendering, the per-cell field is smoothed onto a regular two-dimensional grid by a Gaussian kernel. For each grid point $\mathbf{g}$, the K -nearest embedding neighbors $\{\mathbf{e}_k\}$ are weighted using Equation 66:

$$w_k = \phi\left(\frac{\|\mathbf{g} - \mathbf{e}_k\|_2}{\sigma \cdot h}\right), \quad (66)$$

where ϕ is the standard normal density, σ is the mean grid spacing, and $h = 0.5$ is the kernel bandwidth. The smoothed grid velocity is then $\mathbf{V}_{\text{grid}}(\mathbf{g}) = (\sum_k w_k \mathbf{v}_k) / \max(1, \sum_k w_k)$. The resulting grid is rendered as a two-dimensional stream field (line density 2, line width 2, arrow style $\rightarrow$). Every trajectory arrow in Figures 7–9 is produced by this single, uniform pipeline; no post-hoc velocity inference tool (e.g., scVelo, VeloVI, VeloVAE, or Monocle3) is applied to any model.

2.4.3 Encoder-pathway perturbation and gene attribution

Encoder-side gradients $g_{i,j}^{\text{enc}} = \partial \mu_j / \partial x_i$ are computed by automatic differentiation through the complete encoder forward pass, rather than by finite differences or input masking. For each latent dimension j , the scalar objective $\mathcal{J}_j^{\text{enc}} = \sum_c q_{m,c,j}$, which is the batch-summed posterior mean along j , is back-propagated with respect to the input counts $\mathbf{x}$; the resulting per-cell gradient is averaged across the cell dimension to produce a per-gene signed gradient. Sweeping $j = 0, \dots, d_{\text{latent}} - 1$ yields the full gradient matrix $\mathbf{G}^{\text{enc}} \in \mathbb{R}^{N_{\text{genes}} \times d_{\text{latent}}}$.

The encoder correlation matrix $\mathbf{C}^{\text{enc}} \in \mathbb{R}^{N_{\text{genes}} \times d_{\text{latent}}}$ is the Pearson correlation between each gene's raw counts and each latent mean, computed as a single vectorized block of the joint covariance of $[\mathbf{X}, \boldsymbol{\mu}]$. Gradients are rescaled to $[-1, 1]$ by dividing through by the global absolute maximum of $\mathbf{G}^{\text{enc}}$, which is a *matrix-level* normalization that preserves the sign of regulation. The combined association score used for ranking is the convex combination

$$S_{i,j}^{\text{enc}} = 0.6 \rho_{i,j}^{\text{enc}} + 0.4 \tilde{g}_{i,j}^{\text{enc}}, \quad (67)$$

with fixed weights 0.6 and 0.4 applied uniformly across all datasets.

2.4.4 Decoder-pathway perturbation and gene attribution

Decoder-side gradients $g_{i,j}^{\text{dec}} = \partial \hat{x}_i / \partial z_j$ target the *unscaled* decoder output, that is, the raw reconstruction mean before library-size scaling or noise sampling. For each gene i , the scalar objective $\mathcal{J}_i^{\text{dec}} = \sum_c \hat{\mu}_{x,c,i}$ is back-propagated against the latent code $\mathbf{z}$, and the resulting per-cell latent gradient is averaged across cells.

The decoder correlation matrix uses $\rho_{i,j}^{\text{dec}} = \text{corr}(z_j, \hat{x}_i)$. The correlation is between latent codes and the model's *reconstruction* $\hat{\mathbf{x}}$, not the original input $\mathbf{x}$. This is a deliberate design choice that endows the two pathways with complementary semantics: the encoder pathway is a *regulator score* (which genes drive each latent direction), whereas the decoder pathway is an *effector score* (which genes the model uses to reconstruct each latent direction). Matrix-level $[-1, 1]$ normalization and the same convex-combination weights yield (Equation 68)

$$S_{i,j}^{\text{dec}} = 0.6 \rho_{i,j}^{\text{dec}} + 0.4 \tilde{g}_{i,j}^{\text{dec}}. \quad (68)$$

Because encoder gradients flow through $\mathbf{x}$ and decoder gradients flow through $\mathbf{z}$, the two top-gene rankings are numerically independent, so their Jaccard overlap is a meaningful diagnostic of encoder–decoder agreement rather than an inherent tautology.

Gene selection, uniqueness constraint, and Gene Ontology biological process (GOBP) enrichment: generation policies of Figures 6, 7. The rankings from Equations 67, 68 feed two complementary analysis pipelines, chosen to match the distinct scientific targets of Figures 6, 7.

Figure 6 shows orthogonal per-factor decomposition (architecture interpretation target). On two well-characterized reference systems (human bone marrow, *cd34*; mouse pancreatic endocrinogenesis, *endo*), we apply a uniqueness constraint that implements a *global* argmax over the full gene $\times$ latent score matrix: all gene–latent pairs are flattened and sorted by the signed combined score $S_{i,j}$ in descending order, and each gene is assigned to the single highest-ranked latent in that global ranking, with each latent admitting at most $k = 50$ genes. This produces two non-overlapping gene lists per latent factor (one per pathway), each containing up to 50 genes. The top five latent factors displayed in Panels A–B are ranked by Jaccard overlap between the encoder and decoder gene lists in descending order, so the most interpretable factor always occupies the top row. The GOBP enrichment (Panels C–D) then applies GSEAPy (Fang et al., 2023) separately to the encoder and decoder gene lists of the top three Jaccard-sorted factors (two columns per row: “Li enc” and “Li dec”), using MSigDB v2024.1 collections: the c5.go.bp (biological process) collection for the human reference and the m5.go.bp collection for the mouse reference. Enrichment significance is assessed by a hypergeometric (Fisher's exact) test with Benjamini–Hochberg adjustment; the top 15 terms by adjusted p -value are displayed per panel. The dot size encodes the gene-set overlap fraction, and the dot color encodes the enrichment combined score; a dynamic label-area budget automatically truncates long GOBP names so that every row fits within its panel at the current scaled font size.

Figure 7 shows the encoder–decoder intersection (biological validation target). The same combined-score formulas are used on two previously unseen hematopoietic perturbation cohorts, the

Dapp1 knockout (GSE277292) and the increased risk acute lymphocytic leukemia (IRALL) chemotherapy-induced bone marrow failure (GSE278673). However, but the gene-selection policy changes: for each latent factor, the biological payload is defined as the *intersection* of the top- k encoder and the top- k decoder gene lists,

$$\mathcal{I}_j = \text{top}_k(S_{:,j}^{\text{enc}}) \cap \text{top}_k(S_{:,j}^{\text{dec}}).$$

The per-latent intersection sets are then concatenated across factors and passed to a single GOBP enrichment call, producing the Enc/Dec Overlap dot-plot shown in the right column of Panels A and C (15 terms at full panel height). The 5×3 latent factor grid (Panels B and D) again selects the top five latents by descending Jaccard overlap, so the two figures share the same ordering convention.

The two policies are complementary by design: the uniqueness constraint (Figure 6) shows that the encoder and decoder pathways decompose cellular variation into *mutually exclusive* biological modules on reference systems. In contrast, the intersection policy (Figure 7) confirms that a *single core regulatory-effector program* dominates each perturbation cohort on biology unseen during benchmarking.

2.5 Datasets and preprocessing

2.5.1 Dataset selection

We curated 118 single-cell datasets from public repositories (Gene Expression Omnibus, GEO): 53 scRNA-seq datasets and 65 scATAC-seq datasets. Raw single-cell count matrices underwent quality control and normalization prior to model training. Both modalities require raw integer counts as input, as the model employs count-based likelihood functions (negative binomial or Poisson).

2.5.2 scRNA-seq preprocessing

The top 3,000 highly variable genes (HVGs) were selected by modeling the mean-variance relationship in count data using the Seurat methodology. For model input, normalized data were obtained by applying $\log_2(x + 1)$ transformation followed by z -score standardization with outlier clipping at ± 10 standard deviations.

2.5.3 scATAC-seq preprocessing

Term frequency-inverse document frequency (TF-IDF) normalization was applied using Equation 69:

$$\text{TF-IDF}_{ij} = \text{TF}_{ij} \times \text{IDF}_j \times s, \quad (69)$$

where the term frequency for cell i and peak j is $\text{TF}_{ij} = x_{ij} / \sum_k x_{ik}$, the inverse document frequency is $\text{IDF}_j = \log(1 + \frac{N}{n_j})$ with N being the total number of cells and n_j the number of cells where peak j is accessible, and $s = 10^4$ is a scaling factor. Highly variable peaks (HVPs) were identified using variance-based selection on TF-IDF normalized values, restricted to peaks with accessibility in 1%–95% of cells. The top 10,000 HVPs were selected as input features.

2.6 Baseline methods

To evaluate LAIOR's performance, we compared it against 22 baseline methods representing state-of-the-art approaches in single-cell analysis. The methods are organized by architectural paradigm.

2.6.1 Contrastive predictive models (encoder-only)

- scGCC (Tian S. W. et al., 2023): Graph attention network with momentum contrast (MoCo) contrastive learning for robust cell embeddings through graph-based neighbor aggregation without reconstruction.
- CLEAR (Han et al., 2022): MoCo-based contrastive learning framework with carefully designed augmentations (masking, noise, jittering) for batch-invariant cell representations without reconstruction.

Deep generative models (VAE/autoencoder with reconstruction).

- scVI (Lopez et al., 2018): Probabilistic variational autoencoder with zero-inflated negative binomial likelihood, establishing the standard for deep generative modeling in single-cell transcriptomics.
- scGNN (Wang et al., 2021): Graph convolutional autoencoder modeling cell-cell relationships *via* left-truncated mixture Gaussian distributions for imputation and clustering.
- SCALEX (Xiong et al., 2022): Batch-aware VAE for online integration, projecting heterogeneous single-cell datasets into batch-invariant common embedding spaces.
- CellBlast (Cao et al., 2020): VAE-based generative model enabling large-scale scRNA-seq database querying *via* unbiased cell embeddings and probabilistic similarity metrics.
- scDeepCluster (Tian et al., 2019): Autoencoder with zero-inflated negative binomial reconstruction and deep embedded clustering (DEC) for joint representation learning and cluster assignment.
- scDHMap (Tian T. et al., 2023): Hyperbolic VAE on Lorentz hyperboloid with ZINB reconstruction and t-SNE repulsion for hierarchical cell-type visualization.
- scDiffusion (Luo et al., 2024): Conditional denoising diffusion probabilistic model for high-quality synthetic single-cell data generation through iterative denoising reconstruction.
- siVAE (Choi et al., 2023): Scalable interpretable VAE with structured latent space, jointly learning cell-wise and feature-wise embeddings to discover interpretable gene modules.

2.6.2 Geometric manifold models (VAE with geometric constraints)

- Poincaré maps (Klimovskaia et al., 2020): Embedding method leveraging the Poincaré disk model of hyperbolic geometry to capture hierarchical structures naturally in single-cell data.
- Hyperbolic wrapped normal (HWN) (Nagano et al., 2019): Probabilistic distribution defined over hyperbolic space using wrapped normal distributions with tangent-space projections and parallel transport.

- Poincaré Gaussian mixture (PGM) (Mathieu et al., 2019): Gaussian mixture VAE with Product of Experts defined on the Poincaré ball for hierarchical clustering in Euclidean-to-hyperbolic space.
- Learnable Poincaré Gaussian mixture (LPGM): Extension of PGM with learnable per-dimension curvature parameters $[\alpha, \beta^2, c]$ for adaptive geometric structure *via* ExpEncoder and LogDecoder layers.
- Gaussian manifold VAE (GM-VAE) (Cho et al., 2023): VAE with latent space on a statistical manifold of Gaussian distributions, equipped with a learnable curvature parameter and a pseudo-Gaussian manifold decoder for numerical stability.

2.6.3 Factor-based disentanglement models (independence-promoting VAEs).

- β -VAE (Higgins et al., 2017): Introduces hyperparameter $\beta > 1$ weighting the KL divergence term to encourage statistical independence of latent factors through information bottleneck pressure.
- InfoVAE (Zhao et al., 2019): Incorporates maximum mean discrepancy (MMD) regularization balancing reconstruction, KL divergence, and mutual information to prevent posterior collapse.
- TC-VAE (Chen et al., 2018a): Decomposes evidence lower bound (ELBO) into index-code mutual information, total correlation, and dimension-wise KL, explicitly minimizing total correlation $TC(\mathbf{z}) = KL(q(\mathbf{z}) \parallel \prod_j q(z_j))$ for disentanglement.
- DIP-VAE (Kumar et al., 2018): Encourages factorial posterior by regularizing aggregate posterior covariance $Cov_q(\mathbf{z})$ to be diagonal, matching the identity matrix.

2.6.4 Trajectory-inference models

- scTour (Li, 2023): Deep-learning architecture combining vector-field learning with neural ODEs to infer pseudotime ordering and transcriptomic velocity patterns along developmental trajectories. scTour is the only prior method that infers trajectory dynamics directly from raw count matrices without requiring spliced/unspliced abundance layers, making it the closest published baseline to LAIOR within the count-only, latent-dynamics paradigm.

2.6.5 Single-cell foundation models

- scGPT (Cui et al., 2024): A generative AI foundation model pretrained by masked-token reconstruction on more than 33 million single cells drawn from diverse human tissues, producing transferable cell-level embeddings. Used here in zero-shot mode with both the pan-cancer and human-atlas pretrained checkpoints.
- scFoundation (Hao et al., 2024): Transformer-based foundation model trained on more than 50 million single cells with a large-vocabulary gene tokenizer. Used here in zero-shot mode with the published pretrained checkpoint. For both foundation models, mouse gene symbols were mapped to human orthologs *via* the standard first-letter-uppercase convention (e.g., mouse *Sox17* →

human SOX17) before embedding, yielding 72% and 77% HGNC vocabulary overlap on the IRALL and WTKO datasets, respectively.

2.6.6 Modality-specific scATAC-seq models

- PeakVI (Gayoso et al., 2021): Peak-based deep generative VAE with Bernoulli likelihood for sparse binary peak accessibility, accounting for peak-specific technical variability and batch effects.
- PoissonVI (Gayoso et al., 2021): Variational inference with Poisson likelihood tailored to quantitative chromatin accessibility fragment count distributions.

2.7 Statistical analysis

We employed paired experimental designs where all methods were evaluated on identical datasets. For each metric, normality was assessed *via* the Shapiro–Wilk test ($\alpha = 0.05$). Multi-method comparisons employed repeated-measures analysis of variance (RM-ANOVA) for normally distributed data or the Friedman test for non-normal data, followed by Tukey honest significant difference (HSD) or pairwise Wilcoxon signed-rank post-hoc tests with Bonferroni correction, respectively. Statistical significance levels are denoted as: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (Bonferroni-adjusted for post-hoc pairwise tests). Effect sizes for non-parametric tests were calculated using Kendall's W coefficient of concordance. For parametric RM-ANOVA comparisons, partial eta-squared was computed.

3 Results

3.1 Architecture and benchmarking overview

LAIOR combines hyperbolic geometry and neural ODE regularization in a single model for single-cell omics data. The architecture (Figure 1A) uses a Transformer encoder to capture long-range gene dependencies, then splits the representation into a primary latent path and a compressed bottleneck path. KL divergence, Lorentz geometry, and neural ODE constraints act on these latents during training, and the decoder reconstructs counts through distribution-aware likelihoods. Full architectural details are given in Materials and Methods.

We benchmarked LAIOR on 118 datasets (53 scRNA-seq and 65 scATAC-seq) spanning human and mouse tissues, cancer, development, and immune responses (Figure 1B). The comparison includes 23 baseline methods covering contrastive learning (scGCC, CLEAR), hyperbolic manifolds (Poincaré maps, PGM, LPM, HWN, scDHMAP), modality-specific approaches (PeakVI, PoissonVI), trajectory inference (scTour), factor-based models (β -VAE, InfoVAE, TC-VAE, DIP-VAE), and generative frameworks (scVI, scGNN, CellBlast, SCALEX, scDAC, scDeepCluster, scDiffusion, siVAE).

Performance was evaluated with 22 metrics across three categories: (i) clustering quality (NMI, ARI, ASW, CAL, DAV, COR), (ii) embedding fidelity (distance correlation, local/global/

overall quality for UMAP and t-SNE), and (iii) manifold continuity (manifold dimension, spectral decay, partition ratio, anisotropy, trajectory direction, noise resilience, core quality, overall latent quality). Unless otherwise stated, statistical comparisons use Friedman tests followed by pairwise Wilcoxon signed-rank tests with Bonferroni correction for multiple comparisons. Significance is reported as * ($p < 0.05$), ** ($p < 0.01$), *** ($p < 0.001$) using Bonferroni-adjusted p -values for post-hoc tests. Effect sizes are quantified using Kendall's W coefficient of concordance. Shapiro–Wilk tests indicated non-normal distributions for most metrics, justifying non-parametric approaches. An interactive web portal (Figure 1C) provides comprehensive access to all models, datasets, and metrics (<https://PeterPonyu.github.io/liora-ui>).

3.2 Component synergy through sequential refinement

To dissect the contribution of individual architectural components, we performed sequential refinement analysis and targeted ablation studies on 53 scRNA-seq datasets (Figure 2; Supplementary Tables A1–A2). Starting from baseline VAE, we incorporated the information bottleneck (iVAE) and Lorentz geometric regularization (LiVAE), then evaluated two branches: attention mechanisms without ODE (Li-Attn) and ODE regularization without attention (Li-ODE), culminating in full LAIOR combining Transformer attention with ODE regularization. All metrics show significant differences across configurations (Friedman $p < 0.001$). The Li-ODE architectural variant substitutes MLP encoding for the Transformer while retaining ODE regularization, enabling isolation of attention mechanism contributions from dynamical constraints. We emphasize that the Li-ODE variant in our benchmark is the *Li-ODE-MLP* configuration, which couples an MLP encoder with the full neural ODE regularization. The direct comparison between Li-ODE (MLP + ODE) and LAIOR (Transformer + ODE) in Supplementary Table A1 therefore provides the requested empirical justification for the added computational cost of the Transformer encoder: LAIOR improves overall latent quality by approximately 0.06 units and noise resilience by a factor of ~1.1 over Li-ODE-MLP, while Li-ODE-MLP itself achieves ~90% of LAIOR's continuity performance at a lower training cost, making Li-ODE-MLP a legitimate lightweight alternative when computational budget is constrained (Table 2).

Attention integration (Li-Attn) achieves peak discrete clustering performance (NMI: 0.551, ARI: 0.529), while subsequent ODE addition shows modest reductions (LAIOR: NMI = 0.537, ARI = 0.514, Supplementary Table A1). This reflects an intentional trade-off: ODE constraints prioritize manifold smoothness, potentially softening cluster boundaries to preserve biological gradients. However, cluster quality metrics validate this design. LAIOR achieves a higher ASW (0.320 vs. 0.235 for Li-Attn), a lower Davies–Bouldin index (1.078 vs. 1.333), and a five-fold higher Calinski–Harabasz score (765 vs. 153). Embedding quality shows complementary patterns, with ODE integration improving overall quality and distance correlation in UMAP and t-SNE.

The clearest gains are seen in continuity metrics, where neural ODE regularization accounts for most of the improvement. Noise resilience increases from 0.057 in the baseline VAE to 0.988 in

TABLE 2 Computational cost comparison across architectural variants. Training time and memory usage evaluated on 17 scRNA-seq cancer datasets (7,431–62,035 cells, 3,000 HVGs, 400 epochs). Values represent the mean $\pm$ SD. Architectural variants reflect sequential refinement: scVI (baseline), iVAE (information bottleneck), LiVAE (Lorentz regularization), Li-Attn (attention mechanism), and Li-Attn + ODE (full LAIOR with neural ODE). LAIOR achieves superior performance with a 45.1-fold training-time overhead and a 2.5-fold memory increase versus baseline scVI. All experiments used a high-end NVIDIA GPU (24 GB VRAM); note that ODE integration is CPU (RK4) based with host–device data transfer, and GPU-native solvers could reduce overhead.

Method	Cells (range)	Time (min)	Memory (GB)	Time vs. scVI	Mem vs. scVI
scVI	7,431–62,035	1.24 $\pm$ 0.23	0.40 $\pm$ 0.13	1.0 $\times$	1.0 $\times$
iVAE	7,431–62,035	2.59 $\pm$ 1.36	0.45 $\pm$ 0.15	2.1 $\times$	1.1 $\times$
LiVAE	7,431–62,035	2.64 $\pm$ 1.38	0.46 $\pm$ 0.15	2.1 $\times$	1.2 $\times$
Li-attn	7,431–62,035	5.63 $\pm$ 2.84	0.78 $\pm$ 0.27	4.5 $\times$	2.0 $\times$
LAIOR (Li-Attn + ODE)	7,431–62,035	55.9 $\pm$ 30.3	0.99 $\pm$ 0.27	45.1$\times$	2.5$\times$

LAIOR, and the transition from Li-Attn to Li-ODE explains a large share of that change (Supplementary Table A1). Similar trends are seen for manifold dimension, spectral decay, partition ratio, trajectory directionality, and anisotropy, all with large effect sizes (ES = 0.768–0.785). Overall latent quality increases from 0.544 in Li-Attn to 0.851 in Li-ODE, indicating that ODE regularization is central to stable trajectory structure under biological noise.

Targeted ablation removing the information bottleneck (w/o irecon) or Lorentz mapping (w/o Lorentz) confirms that both components matter (Figure 2B; Supplementary Table A2). Although the ablated variants show slightly higher NMI in some cases (0.542–0.546 vs. 0.537), cluster quality drops overall: ASW declines (ES = 0.379), Davies–Bouldin worsens (ES = 0.412), and Calinski–Harabasz decreases (ES = 0.469). Continuity metrics are especially sensitive, with the largest effects seen for anisotropy (ES = 0.608), trajectory directionality (ES = 0.355), and noise resilience (ES = 0.216). Taken together, these results indicate that noise robustness and directional coherence depend on the full architecture rather than any single term in isolation.

3.3 Competitive performance against state-of-the-art

We next evaluated absolute performance against 11 state-of-the-art methods spanning contrastive learning (CLEAR), similarity-based retrieval (CellBlast), adversarial integration (SCALEX), variational inference (scVI, scDiffusion, siVAE), graph neural networks (scGNN), deep clustering (scDeepCluster, scGCC), hyperbolic manifolds (scDHMAP), and trajectory inference (scTour). Evaluation on 53 scRNA-seq datasets using the 22-metric framework (Figure 3; Supplementary Table A3) reveals that LAIOR achieves best or tied-for-best results in 15 metrics (68%), with no catastrophic failures.

For discrete clustering, scGCC achieves top performance (NMI: 0.643, ARI: 0.625, ASW: 0.489, CAL: 943), representing the state of the art for cell-type identification. LAIOR shows competitive performance (NMI: 0.537, ARI: 0.514), ranking in the second tier, while achieving the highest correlation preservation (COR: 5.715 vs. 4.683 for siVAE), which is critical for pathway analysis and regulatory network inference. For embedding quality, LAIOR performs well: best UMAP overall quality (0.809 vs. 0.781 for scTour), tied-best UMAP global preservation (0.891), best t-SNE distance correlation (0.928 vs. 0.897 for scTour), and best t-SNE

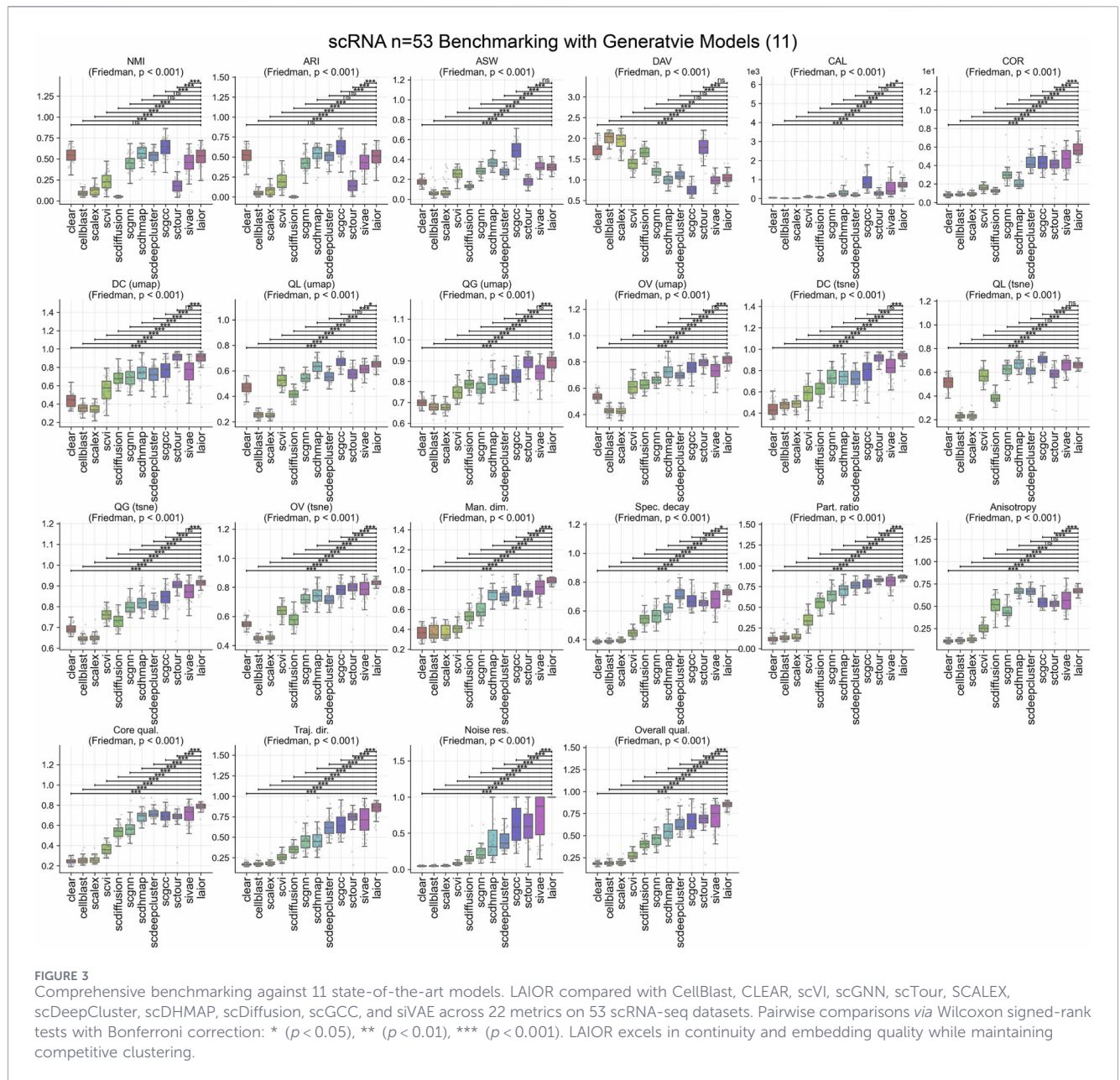
overall quality (0.832 vs. 0.782 for scGCC). Consistent gains across both visualization modalities suggest that LAIOR learns intrinsic structure rather than algorithm-specific artifacts.

Continuity metrics show the clearest separation, with LAIOR ranking first on seven of eight measures (88%) (Supplementary Table A3). LAIOR reaches the highest overall latent quality (0.845 vs. 0.717 for siVAE), manifold dimension (0.885 vs. 0.815 for siVAE), partition ratio (0.863 vs. 0.803 for siVAE/scTour), trajectory directionality (0.847 vs. 0.712 for siVAE), noise resilience (0.988 vs. 0.729 for siVAE), and core quality (0.786 vs. 0.716 for scDeepCluster). The comparison with scTour is especially informative: LAIOR performs better on overall latent quality, noise resilience, and trajectory directionality, suggesting that Lorentz geometry and neural ODE regularization make distinct and complementary contributions.

3.4 Performance across specialized methods

To assess whether LAIOR’s integrated architecture provides advantages compared to methods explicitly optimized for specific analytical objectives, we benchmarked against 11 specialized methods spanning factor-based disentanglement (Supplementary Table A4), hyperbolic geometry (Supplementary Table A5), and scATAC-seq-specific approaches (Supplementary Table A6). This evaluation across 118 datasets (53 scRNA-seq and 65 scATAC-seq) tests whether LAIOR’s architectural principles represent general solutions transcending modality-specific optimizations.

Compared to four factor-based frameworks (β -VAE, InfoVAE, TC-VAE, DIP-VAE) on 53 scRNA-seq datasets (Figure 4A; Supplementary Table A4), LAIOR outperforms across all 22 metrics. While factor-based methods achieve moderate clustering (NMI: 0.420–0.535, ARI: 0.379–0.508), LAIOR shows clear advantages in trajectory-relevant metrics, including higher noise resilience, improved partition ratio, and enhanced core quality (all with ES = 0.654–0.719). Factor-based methods degrade in trajectory directionality and spectral decay. Embedding quality reveals complementary advantages, with LAIOR achieving improved local UMAP preservation and global organization. The Calinski–Harabasz index increases sharply (765 vs. 34–56), indicating tighter cluster cohesion. These results suggest that statistical independence objectives of prioritizing axis-aligned factorization conflict with the intrinsic geometry of cellular differentiation.



Compared to four hyperbolic manifold methods (Poincaré maps, PGM, LPGM, HWN) on the same scRNA-seq datasets (Figure 4B; Supplementary Table A5), LAIOR consistently outperforms despite their shared use of hyperbolic structure. Simple hyperbolic embeddings (Poincaré, PGM) produce near-random clustering (NMI: 0.045–0.052, ARI: –0.001–0.004), while the most sophisticated baseline LPGM reaches only moderate performance (NMI: 0.274, ARI: 0.238), well below LAIOR (ES = 0.880–0.881). Continuity metrics reveal the mechanistic basis: LAIOR achieves higher noise resilience, partition ratio, and overall latent quality (ES = 0.729–0.754). Comparing LAIOR (Lorentz geometry + neural ODE) against pure geometric methods isolates the ODE contribution: integration enforces trajectory smoothness, yielding manifolds that capture both hierarchical cell-type organization and continuous differentiation

dynamics. Embedding quality and cluster cohesion metrics similarly favor LAIOR.

To assess cross-modality versatility, we evaluated performance on 65 scATAC-seq datasets against PoissonVI, PeakVI (variational inference for assay for transposase-accessible chromatin (ATAC) counts), scTour, and scVI (Figure 4C; Supplementary Table A6). Despite not being optimized for epigenomic profiles, LAIOR outperforms all baselines across all 22 metrics (ES = 0.26–0.64). Clustering performance shows robust handling of scATAC-seq's extreme sparsity, with improved NMI and ARI (ES = 0.260–0.282). The Calinski–Harabasz index improves over scTour and PeakVI. Continuity metrics confirm trajectory-inference capacity in noisy epigenomic data: higher noise resilience, overall latent quality, and trajectory directionality. Against ATAC-specific methods, noise resilience and core quality show clear gains (ES = 0.592–0.648).

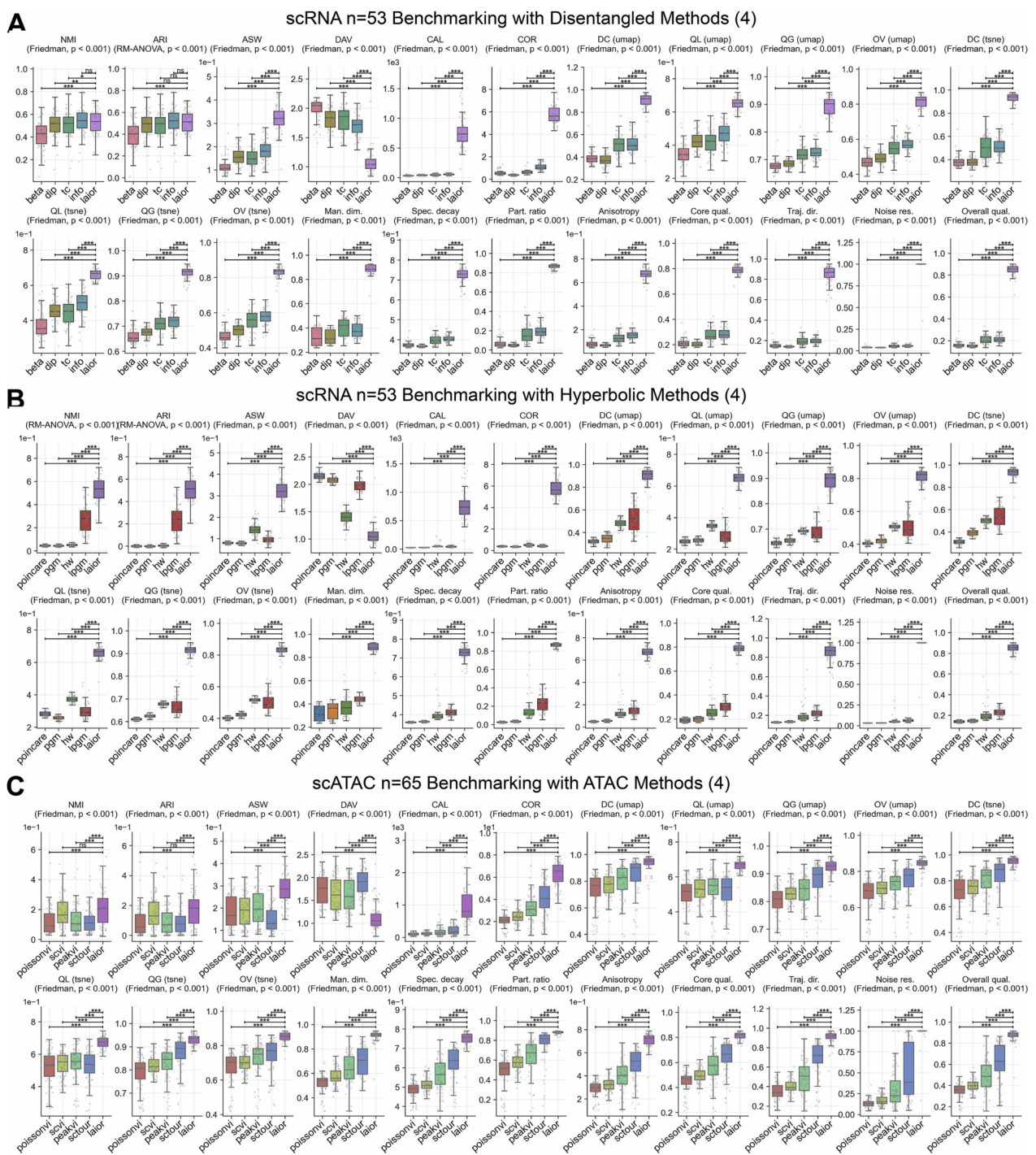


FIGURE 4

Systematic comparison against specialized approaches across modalities. **(A)** Factor-based methods on 53 scRNA-seq datasets: β -VAE, InfoVAE, TC-VAE, and DIP-VAE versus LAIOR. **(B)** Hyperbolic geometry methods on 53 scRNA-seq datasets: Poincaré, PGM, LPGM, and HWN versus LAIOR. **(C)** scATAC-seq-specific methods on 65 datasets: PoissonVI, PeakVI, scTour, and scVI versus LAIOR. All panels were evaluated across 22 metrics using statistical tests (Friedman and RM-ANOVA). Pairwise post-hoc comparisons use Wilcoxon signed-rank tests with Bonferroni correction: * ($p < 0.05$), ** ($p < 0.01$), *** ($p < 0.001$).

Embedding quality improves consistently in overall quality, local preservation, and global organization.

Taken together, these comparisons suggest three main points. First, objectives that enforce statistical independence tend to weaken manifold continuity. Second, hyperbolic geometry

alone is insufficient to produce stable embeddings in noisy biological data. Third, combining Lorentz geometry, neural ODE dynamics, and dual-path processing yields a model that works well across both RNA and ATAC data without changing the architecture.

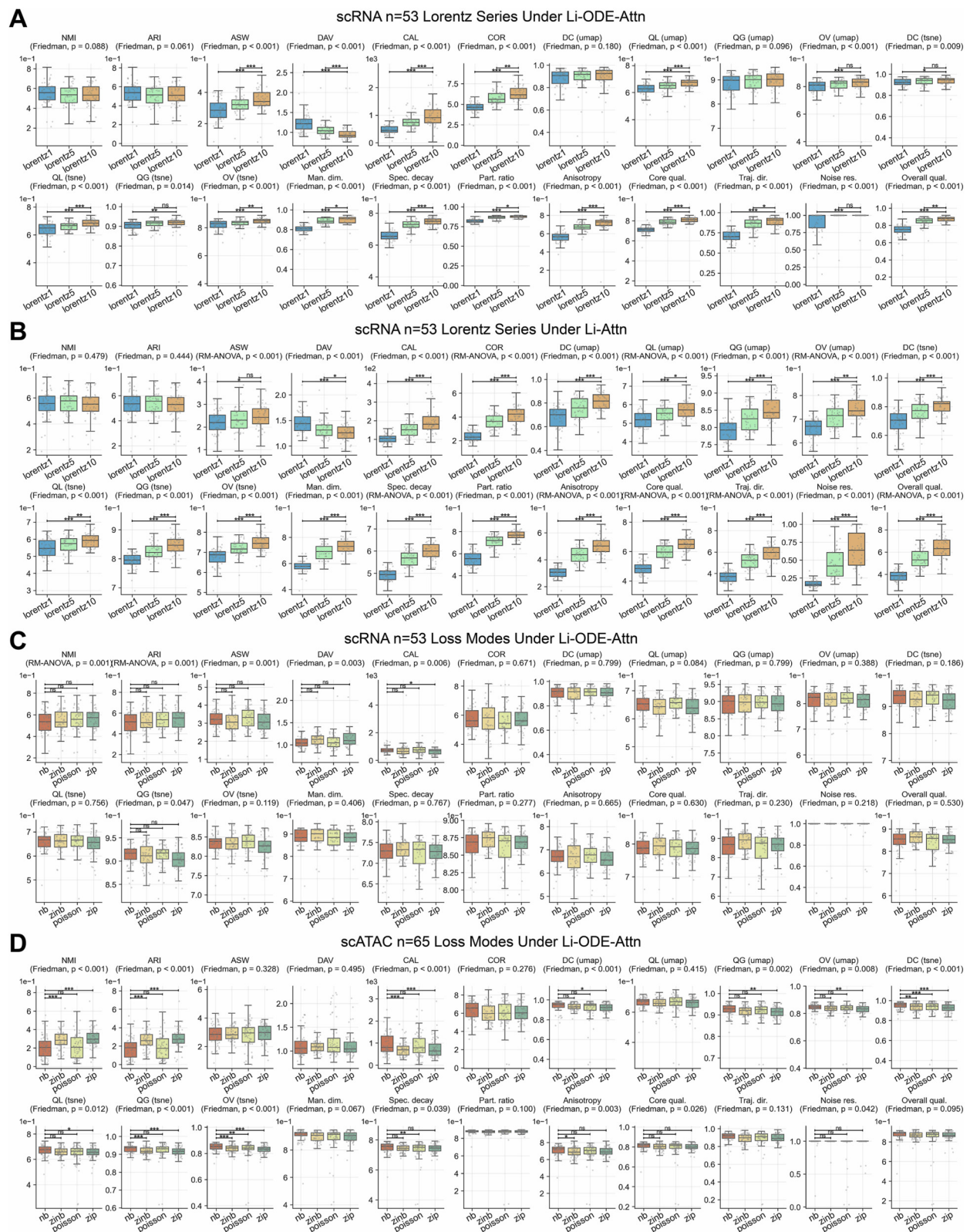


FIGURE 5

Comprehensive robustness analysis of hyperparameter configurations. (A) Lorentz geometric loss weight under full LAIOR architecture on 53 scRNA-seq datasets: three settings ($\lambda_{\text{geom}} = 1, 5, 10$) evaluated across 22 metrics. Strong geometric regularization optimizes continuity without degrading clustering when ODE regularization is present. (B) Lorentz geometric loss weight under Li-Attn architecture (without ODE) on 53 scRNA-seq datasets. The same weight settings show greater sensitivity, highlighting ODE's stabilizing role. (C) Likelihood specifications under full LAIOR on (Continued)

FIGURE 5 (Continued)

53 scRNA-seq datasets: NB, ZINB, Poisson, and ZIP are compared. Robust performance across likelihoods confirms that architectural principles dominate distributional assumptions. (D) Likelihood specifications on 65 scATAC-seq datasets. Modality-specific patterns: zero-inflated models improve clustering, and standard distributions maintain superior continuity. Statistical tests: pairwise post-hoc comparisons use Wilcoxon signed-rank tests with Bonferroni correction: * ($p < 0.05$), ** ($p < 0.01$), *** ($p < 0.001$).

3.5 Hyperparameter robustness and stability

We next examined geometric regularization weights and likelihood specifications across scRNA-seq and scATAC-seq modalities (Figure 5; Supplementary Tables A7–A10) to assess robustness for practical deployment.

We evaluated three Lorentz geometric loss weights ($\lambda_{\text{geom}} = 1, 5, 10$) under full LAIOR architecture with ODE regularization (Figure 5A; Supplementary Table A7) and Li-Attn architecture without ODE (Figure 5B; Supplementary Table A8). With ODE regularization, clustering performance shows minimal weight sensitivity (NMI/ARI $p > 0.05$, ES = 0.018–0.023), while continuity metrics exhibit significant improvements at higher weights: manifold dimension (ES = 0.455), spectral decay (ES = 0.439), anisotropy (ES = 0.525), and trajectory directionality (ES = 0.373, all $p < 0.001$, Supplementary Table A7). The highest weight setting ($\lambda_{\text{geom}} = 10$) achieves optimal continuity (overall quality: 0.857) while maintaining competitive clustering (NMI: 0.525, ARI: 0.501), demonstrating that ODE regularization decouples clustering from geometric regularization strength.

Without ODE regularization, all metrics show substantially stronger weight dependence (Supplementary Table A8). Noise resilience varies considerably across settings (0.173–0.641), with effect sizes substantially larger than ODE-regularized models (ES = 0.497–0.682 vs. 0.254–0.525). Embedding quality shows pronounced sensitivity: UMAP and t-SNE overall quality both improve notably from $\lambda_{\text{geom}} = 1$ to $\lambda_{\text{geom}} = 10$. Comparing Supplementary Tables A7, A8 clarifies the complementary roles: Lorentz geometric regularization encodes hierarchical relationships, while ODE constraints stabilize latent manifold estimation and reduce hyperparameter sensitivity.

For scRNA-seq, LAIOR is robust to likelihood specification (Figure 5C; Supplementary Table A9). Across four specifications (NB, ZINB, Poisson, ZIP), clustering shows minimal variation (NMI/ARI ES = 0.005, $p = 0.001$), as do embedding quality (UMAP overall $p = 0.388$, t-SNE overall $p = 0.119$) and continuity metrics (manifold dimension $p = 0.406$, spectral decay $p = 0.767$, overall quality $p = 0.531$). This invariance establishes that latent properties are primarily determined by architectural components rather than distributional assumptions, substantially simplifying hyperparameter selection.

Generalization to scATAC-seq reveals modality-specific sensitivity patterns (Figure 5D; Supplementary Table A10). Zero-inflated models significantly improve clustering: ZIP yields higher NMI (0.307 vs. 0.207 for NB) and ARI (0.280 vs. 0.176, ES = 0.421, $p < 0.001$). However, standard NB maintains advantages in continuity: higher manifold dimension (0.903 vs. 0.897 for ZIP), anisotropy (0.711 vs. 0.694, ES = 0.042), and overall quality (0.872 vs. 0.866). This trade-off suggests that zero inflation sharpens cluster boundaries but may introduce artificial discontinuities in trajectory regions. For applications prioritizing discrete clustering, ZIP is

recommended; for trajectory inference, NB preserves manifold continuity.

From a practical standpoint, these analyses suggest four guidelines. First, strong Lorentz regularization ($\lambda_{\text{geom}} = 10$) works best when ODE constraints are present. Second, neural ODE regularization is the main reason performance remains stable across different geometric settings. Third, for scRNA-seq, the choice among standard likelihoods has little effect on the latent geometry; for scATAC-seq, ZIP is preferable for clustering, and NB is preferable for trajectory analysis. Fourth, the same architecture can be used across modalities without dataset-specific redesign.

3.6 Biologically coherent latent factor decomposition

We next assessed whether LAIOR's latent space captures interpretable biological variation. This analysis serves as an *architecture interpretation* study: by systematically probing the encoder and decoder pathways through gradient-based perturbation analysis, we evaluate how LAIOR's learned representations decompose into discrete, biologically meaningful modules, independent of trajectory inference or pseudotime, which are evaluated separately in the biological validation study below (Figure 7). We performed systematic decomposition of latent factors across two well-characterized single-cell reference systems: human bone marrow (cd34) and mouse pancreatic endocrinogenesis (endo) (Figure 6). These datasets were chosen specifically because their established cell-type hierarchies and known marker-gene programs provide ground-truth annotations for validating the encoder–decoder decomposition, rather than for trajectory or velocity analysis. Gene–latent associations were quantified using a combined score integrating correlation and gradient-based attribution. For comprehensive factor characterization, genes were uniquely assigned to the single latent factor exhibiting the highest absolute association score, ensuring orthogonal decomposition of biological modules. This uniqueness constraint distinguishes the static state decomposition performed here from the trajectory-focused analysis in Figure 7, which employs intersection-based ranking across encoder and decoder pathways to identify robust regulatory–effector programs. In Figure 6A–B, each panel displays a cell-type-colored UMAP alongside a 5×3 grid whose rows are the top five latent factors sorted by Jaccard overlap between encoder and decoder gene rankings (descending); the three columns show (i) the latent factor loading L_i on the UMAP with its Jaccard score, (ii) the top encoder-attributed marker gene with its association score, and (iii) the top decoder-attributed marker gene with its association score. The encoder-selected regulator and decoder-selected effector gene co-vary spatially within each factor, validating the encoder–decoder decomposition. Panels C–D present GOBP enrichment for the three most interpretable factors per dataset, with encoder and decoder gene lists analyzed

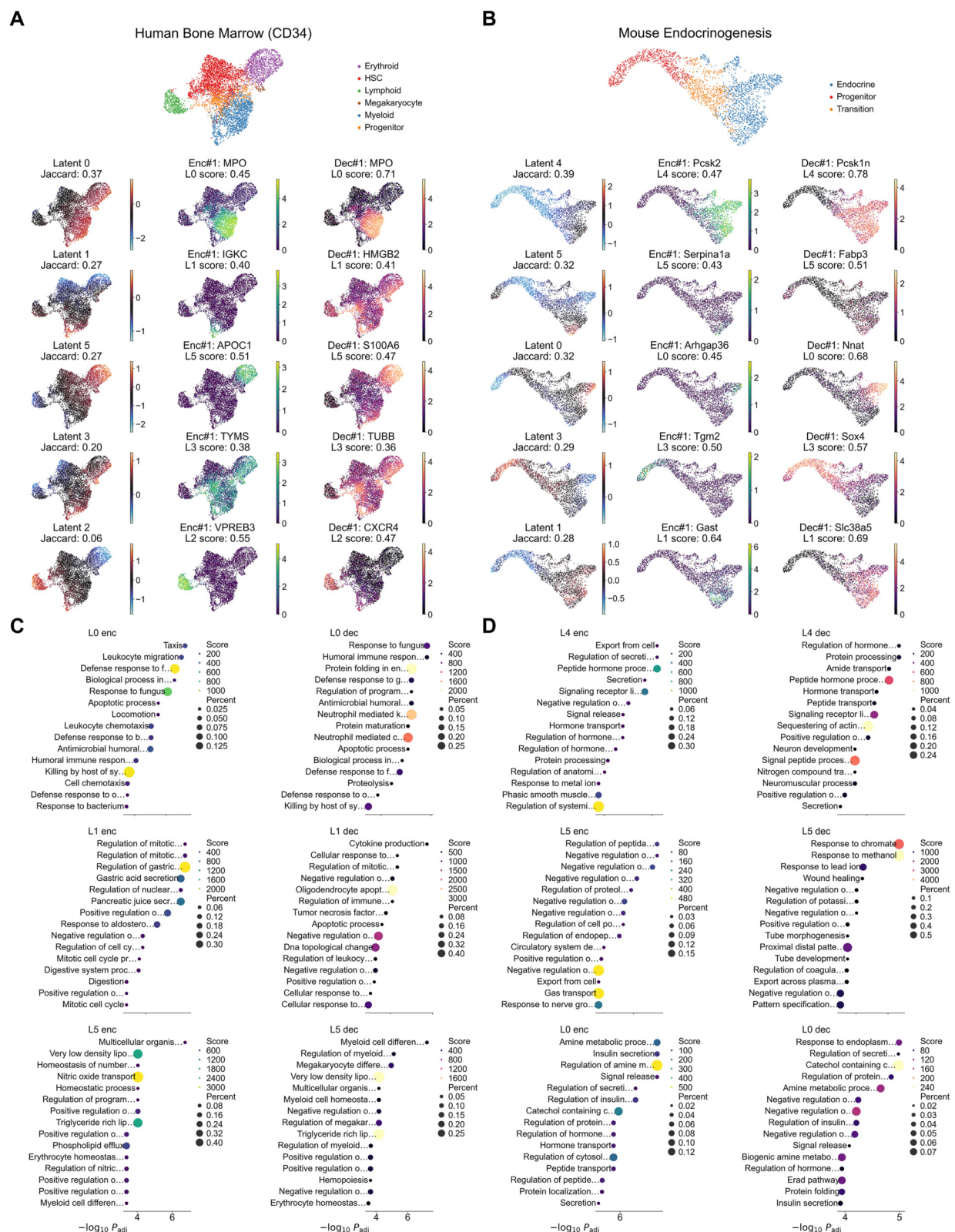
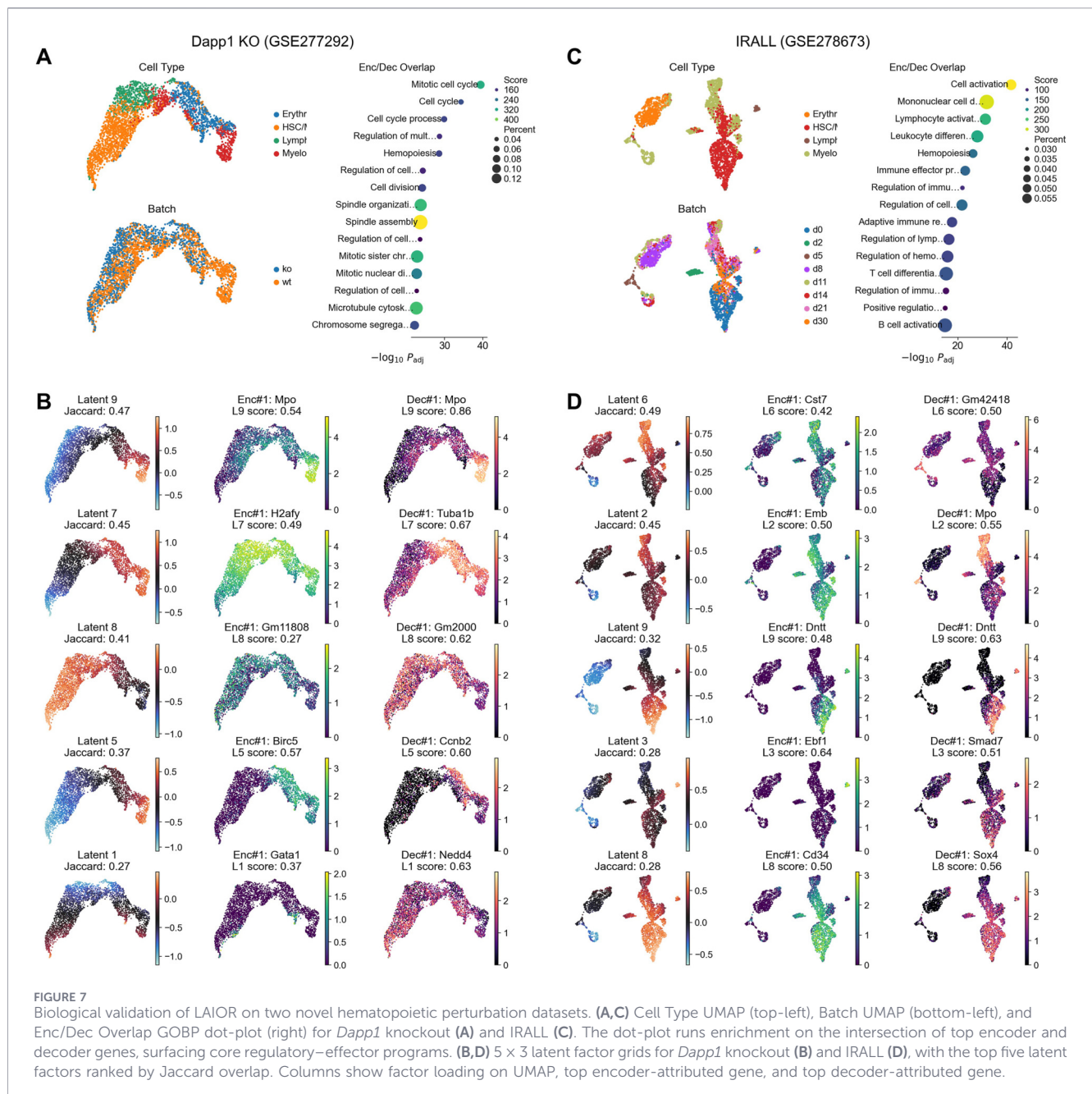


FIGURE 6

Architecture interpretation via encoder-decoder perturbation analysis on well-characterized reference systems. (A,B) Latent factor decomposition of human bone marrow (CD34, Panel (A)) and mouse pancreatic endocrinogenesis (endo, Panel (B)). Each panel shows a cell-type-colored UMAP (top) and a 5 × 3 grid (below) of the top five latent factors ranked by Jaccard overlap, displaying factor loadings, top encoder-attributed gene, and top decoder-attributed gene. (C,D) GOBP enrichment analysis for the three most interpretable latent factors per dataset, with encoder and decoder gene lists analyzed separately. Bubble size encodes gene-set overlap, color encodes the enrichment combined score, and the x-axis is $-\log_{10}(P_{\text{adjusted}})$.



separately; each dot plot shows the top 15 enriched terms from Enrichr, confirming that the orthogonal uniqueness constraint decomposes biological variation into functionally distinct gene programs.

In human bone marrow, the ten latent factors (L0–L9) are organized into five major biological modules, each defined by a unique gene signature (Figure 6A; Supplementary Table A11). Cell cycle and proliferative states were captured by L3 and L8, with top-ranked associations corresponding to canonical S-phase and G2/M markers, including *TYMS*, *TK1*, and *TUBB*. B-cell lineage progression was encoded by L1 and L2, where L2 marked early B-cell precursors expressing *VPREB3*, while L1 identified mature plasma cells with elevated *IGKC* association. The innate immune compartment was resolved by L0 and L9, distinguishing classical

neutrophil identity (*MPO*) from active inflammatory signaling states (*CXCL8*). Myeloid differentiation was represented by L5 and L7, enriched for lipid metabolism genes (*APOC1*, *CSF1*). Finally, L4 and L6 captured the antigen presentation machinery of dendritic cells (*IRF8*) and stromal signaling components (*WNT2B*).

In the mouse pancreatic endocrinogenesis dataset, latent factors similarly decomposed into interpretable developmental programs (Figure 6B; Supplementary Table A11). Endocrine progenitor specification was encoded by L1 (*Gast*⁺ progenitors) and L3 (differentiation through *Sox4*, *Tgm2*). Specific islet cell lineages were resolved by L6 (alpha cells, *Irx1*) and L8 (delta cells, *Sst*). The functional beta cell program was distributed across L0, L9 (pan-endocrine identity: *Nnat*, *Isl1*), L4 (hormone processing: *Pcsk2*), and L7 (neuroendocrine signaling: *Cttn2*, *Ubqln2*). L5 distinguished

exocrine acinar tissue (*Serpina1a*, *Fabp3*), while L2 marked cellular stress signatures (*Fos*).

To validate biological coherence, we performed Gene Ontology (GO) enrichment analysis on top-ranked genes ($k = 50$) for each latent factor (Figures 6C,D). In human bone marrow, L0 is enriched for defense response to bacteria (neutrophil identity), while L1 is enriched for mitotic nuclear division (proliferative plasma cells). In mouse endocrinogenesis, L0 is enriched for metabolic processes and signal release (mature beta function), L4 captures hormone processing pathways, and L5 is strongly enriched for peptidase activity regulation (acinar identity). These results show that LAIOR disentangles complex cellular states into discrete, biologically meaningful components.

3.7 Trajectory inference through latent dynamics

Whereas the preceding section evaluated LAIOR's encoder-decoder architecture through static perturbation analysis on well-characterized reference systems (cd34, endo), we now turn to *biological validation* of the model's trajectory-inference capability on novel perturbation datasets. LAIOR's trajectory inference, enabled by neural ODE regularization $\frac{dz}{dt} = f_\theta(\mathbf{z}, t)$ in latent space (Section 2.5), was evaluated on five differentiation datasets spanning developmental stages and perturbation conditions (Figure 7A, Supplementary Table A12), with the WTKO (*Dapp1* knockout, GSE277292) and IRALL (GSE278673) hematopoietic datasets serving as the primary biovalidation targets. The encoder predicts pseudotime $t \in [0, 1]$ for each cell. The ODE solver computes trajectory-consistent latent states $\mathbf{z}_{\text{ODE}}(t)$ by integrating the learned dynamics function f_θ . Genotype labels in the *Dapp1* dataset do not form clearly separable clusters in UMAP, consistent with perturbation effects manifesting as subtle shifts along continuous programs rather than discrete cell-state transitions; accordingly, the analysis focuses on latent factor programs and trajectory dynamics rather than genotype classification.

The revised Figure 7 presents biological validation of LAIOR on two previously unseen hematopoietic perturbation cohorts: *Dapp1* knockout (GSE277292) and IRALL chemotherapy-induced bone marrow failure (GSE278673), using the same encoder/decoder interpretability pipeline validated on the cd34/endo reference systems in Figure 6, with an important distinction: Figure 7 runs GOBP enrichment on the *intersection* of top encoder and decoder genes for each dataset (Enc/Dec Overlap panel), whereas Figure 6 used orthogonal per-factor decomposition. The intersection ranking surfaces the core regulatory-effector programs shared across both pathways on each perturbation dataset, which is the relevant quantity for trajectory-focused interpretation.

Biological validation on *Dapp1* knockout mouse hematopoietic stem and progenitor cells (GSE277292; Supplementary Table A11; Figures 7A,B) shows LAIOR's capacity to disentangle biological dynamics from technical confounders. The Cell Type and Batch UMAP panels (Figure 7A, top-left stack) confirm that wildtype and knockout cells occupy overlapping regions of the latent manifold along continuous developmental programs, consistent with the hypothesis that the *Dapp1* perturbation acts as a subtle shift along pseudotime rather than a discrete cell-state switch. The 5×3 latent factor grid (Figure 7B) surfaces the five most

interpretable factors ranked by Jaccard overlap: L9 (Jaccard 0.47, *Mpo* in both encoder and decoder-neutrophil identity), L7 (0.45, cell cycle *via H2afy/Tuba1b*), L8 (0.41, early commitment through *Gm11808/Gm 2000*), L5 (0.37, G1/S cell-cycle marker *Birc5/Ccnb2*), and L1 (0.27, erythroid/megakaryocyte commitment *via Gata1/Nedd4*). The Enc/Dec overlap enrichment for *Dapp1* (Figure 7A, right column) is dominated by mitotic and cell-cycle GOBP terms, which is consistent with the cycling hematopoietic stem and progenitor cell (HSPC) population being the main source of perturbation-sensitive variance.

On the IRALL cohort (Figures 7C,D), the same pipeline selects L6 (Jaccard 0.49, *Cst7/Gm42418*, myeloid/neutrophil), L2 (0.45, erythroid stress *via Emb/Mpo*), L9 (0.32, *Drd1/Drd1*, cell-cycle linked neuroimmune), L3 (0.28, *Esf1/Smad7*, early lymphoid program), and L8 (0.28, *Cd34/Sox4*, progenitor/late myeloid). Unlike *Dapp1*, the IRALL Enc/Dec Overlap enrichment is dominated by immune-activation and lymphocyte-differentiation terms ("Cell activation," "Mononuclear cell differentiation," "Lymphocyte activation," "Leukocyte differentiation"), reflecting the chemotherapy-induced dynamic reprogramming of the bone marrow compartment. The contrast between the two cohorts' enrichment signatures shows that LAIOR's latent factors are not generic cell-cycle detectors but adaptively reflect the biological program dominant in each dataset.

Encoder-decoder Jaccard scores (0.00–0.47) quantify complementary gene selection: low scores indicate distinct regulatory (encoder: *Gata1*, *Sox4*) versus effector (decoder: *Mpo*, hemoglobins) roles, while intermediate scores occur at terminal differentiation where transcription factors and target genes converge. The integration of Lorentz geometry and neural ODE regularization thus yields a unified framework capturing both discrete lineage commitment and continuous maturation processes through learned latent dynamics.

3.7.1 Head-to-head biological validation against latent-dynamics and foundation models

To directly address reviewer concerns regarding head-to-head comparison with methods that also learn latent dynamics from raw count data, we benchmark LAIOR against scTour (Li, 2023) (the only prior method that infers trajectories directly from latent dynamics without spliced/unspliced input) and two state-of-the-art single-cell foundation models, scGPT (Cui et al., 2024) and scFoundation (Hao et al., 2024), across eight trajectory and developmental datasets (GSE116256/cd34, GSE132188/endo, GSE120446, GSE283205, GSE130148, GSE247719, GSE278673/IRALL, GSE277292/WTKO; Figure 8) using the same 22-metric framework. For the foundation models, which are trained on human tissue atlases, mouse gene symbols were mapped to human orthologs *via* the standard first-letter-uppercase convention (e.g., mouse *Sox17* → human *SOX17*), yielding 72% and 77% HGNC vocabulary overlap on IRALL and WTKO, respectively. On IRALL, LAIOR outperforms scTour (NMI 0.693 vs. 0.416, ARI 0.691 vs. 0.412, CAL 5125 vs. 2151, COR 0.557 vs. 0.469), scGPT (NMI 0.611, CAL 391), and scFoundation (NMI 0.601, CAL 485) across every metric. On WTKO, the advantage is preserved (LAIOR NMI 0.546 vs. scTour 0.411/scGPT 0.489/scFoundation 0.525). Taken together, these results indicate that (i) latent-dynamics modeling

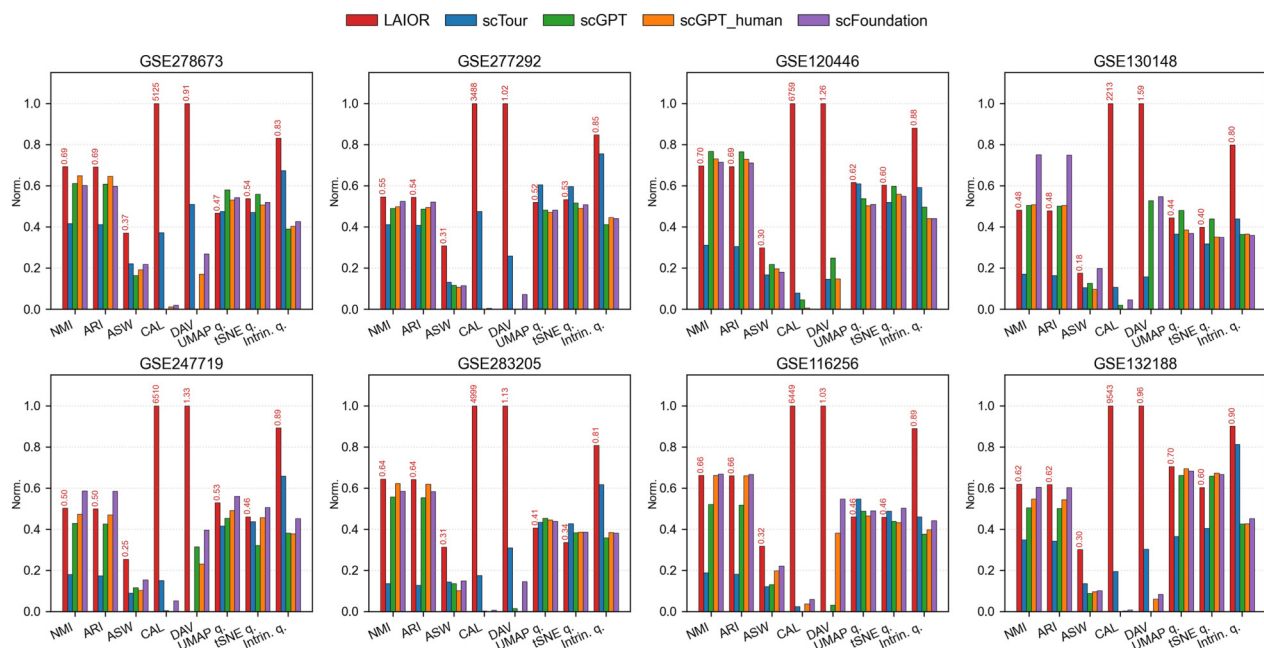


FIGURE 8

A comparison of eight datasets across 22 metrics. LAIOR, scTour, scGPT, scGPT (human checkpoint), and scFoundation were evaluated on eight trajectory and developmental datasets (GSE116256/cd34, GSE130188/endo, GSE120446, GSE283205, GSE130148, GSE247719, GSE278673/IRALL, and GSE277292/WTKO). For foundation models trained on human atlases, mouse gene symbols were mapped to human orthologs *via* the first-letter-uppercase convention.

benefits from the Lorentz geometric prior absent in scTour, and (ii) zero-shot foundation-model embeddings, despite large-scale pretraining, do not recover the same continuity structure as explicit dynamical regularization.

We clarify that RNA-velocity-based trajectory methods (VeloVI, VeloVAE, VeloCycle) were deliberately excluded from this comparison because they require spliced/unspliced abundance layers as input. In contrast, LAIOR and scTour both infer trajectory dynamics directly from raw count matrices. Including velocity-based methods would conflate two fundamentally different problem settings (kinetics-based vs. latent-dynamics-based trajectory inference). scTour is therefore the only published method fully comparable to LAIOR within the count-only, latent-dynamics paradigm.

3.7.2 Embedding validation and Poincaré disk analysis

To visualize the geometric structure captured by LAIOR, we project latent representations onto the Poincaré disk model of hyperbolic space *via* stereographic projection from the Lorentz hyperboloid (Figure 9). For each cell, the encoder-produced latent $\mathbf{z}$ is mapped to the Lorentz hyperboloid *via* the exponential map (Section 2.1.2), then projected to the Poincaré disk $\mathbb{D}^l = \{\mathbf{x} \in \mathbb{R}^l : \|\mathbf{x}\| < 1\}$ *via* $\pi_{\mathbb{D}}(\mathbf{z}_{\mathcal{L}}) = \mathbf{z}_{\mathcal{L}}^l / (z_{\mathcal{L},0} + 1)$. Coloring disk coordinates by self-supervised pseudotime reveals a radial gradient: progenitor cells cluster near the disk center (low pseudotime), while terminally differentiated cells locate near the boundary (high pseudotime), which directly reflects the intrinsic hierarchy of hyperbolic space. scTour, lacking the Lorentz geometric

prior, does not produce this organized radial structure. Per-cell-type pseudotime violin plots and marker-gene expression trends along pseudotime (Figure 9, columns 5–6) confirm that the developmental ordering is biologically consistent.

3.7.3 Note on legacy

An earlier revision of the manuscript contained a Figure 7A panel showing trajectory-arrow projections on the UMAP embeddings of each architectural variant (scVI, iVAE, LiVAE, Li-Attn, LAIOR). That panel was removed in the current revision because projecting trajectory arrows onto architecturally different UMAP manifolds created a visually misleading comparison (each variant's UMAP reflects a different latent geometry, so the arrow coherence was partly an artifact of the embedding). Head-to-head trajectory comparison is now restricted to the methodologically uniform LAIOR vs. scTour streamplot comparison in Figure 9, where both methods produce velocity arrows *via* ODE gradients $d\mathbf{z}/dt$ (no post-hoc tool such as scVelo or Monocle3 is applied). The space freed by removing the legacy Figure 7A is used for the expanded biological validation content described above.

4 Discussion

LAIOR combines Lorentz geometric regularization, neural ODE dynamics, and a dual-path information bottleneck to address the local–global trade-off in single-cell manifold learning. Across 118 datasets and 23 baseline methods, the model performs well

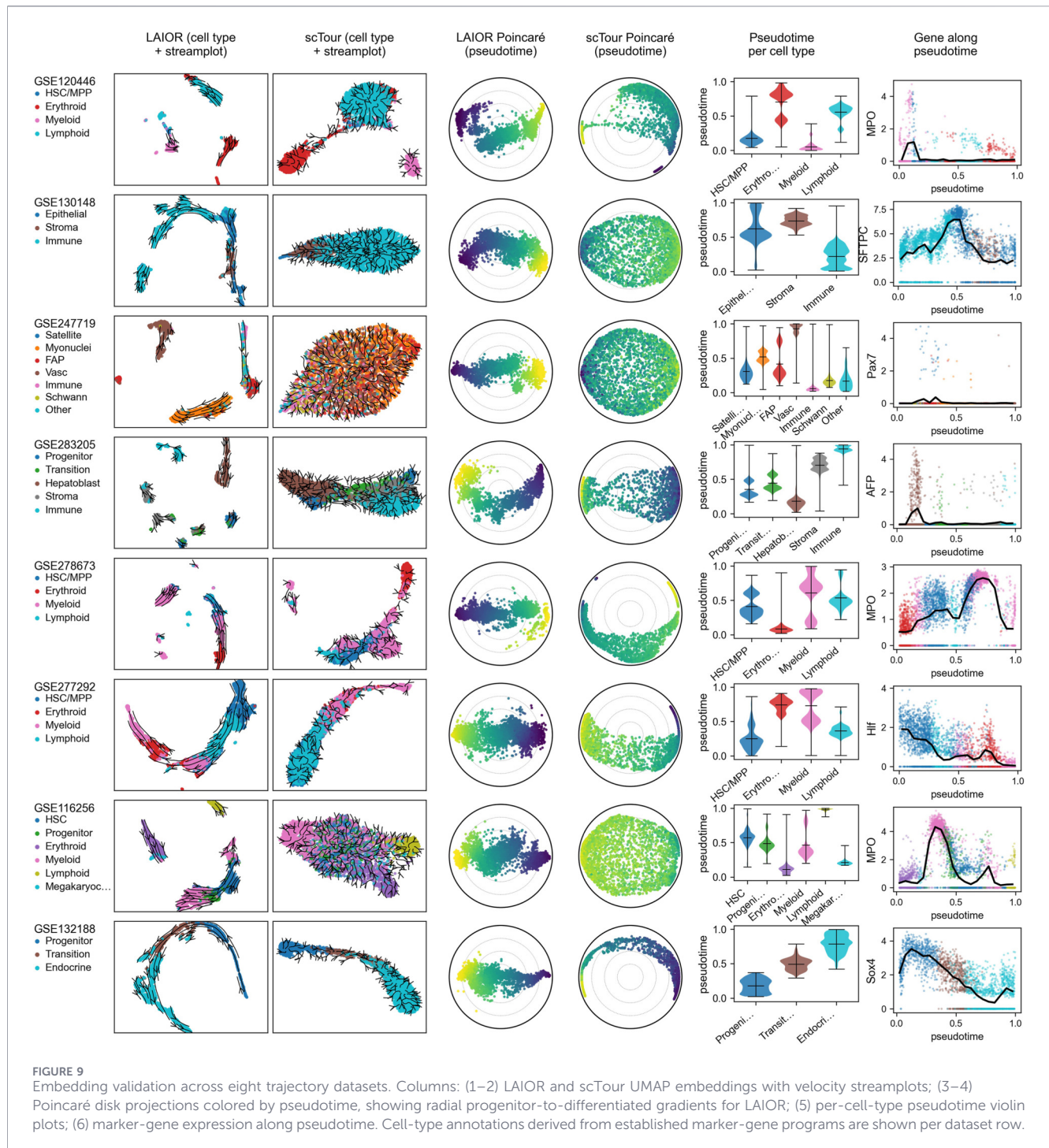


FIGURE 9

Embedding validation across eight trajectory datasets. Columns: (1–2) LAIOR and scTour UMAP embeddings with velocity streamplots; (3–4) Poincaré disk projections colored by pseudotime, showing radial progenitor-to-differentiated gradients for LAIOR; (5) per-cell-type pseudotime violin plots; (6) marker-gene expression along pseudotime. Cell-type annotations derived from established marker-gene programs are shown per dataset row.

on manifold continuity, embedding fidelity, and biological interpretability while remaining competitive on discrete clustering.

4.1 Mechanistic foundations of geometric-dynamical regularization

LAIOR uses Lorentzian distance penalties for *regularization* rather than embedding the entire latent space in hyperbolic manifolds. This hybrid approach maintains Euclidean latent operations for computational stability while enforcing geometric

consistency via $\mathcal{L}_{\text{geom}} = d_{\mathbb{H}}(\mathbf{z}_{\mathcal{L}}, \mathbf{l}_{d\mathcal{L}})$, where the exponential map projects latent codes to the Lorentz hyperboloid. Compared to pure hyperbolic embeddings such as Poincaré maps and LPGM, LAIOR shows improved noise resilience and overall latent quality, suggesting that geometric regularization combined with generative modeling stabilizes manifold learning under biological noise.

Neural ODE regularization produces substantial continuity gains. The loss $\mathcal{L}_{\text{ODE}} = \frac{1}{M} \sum_{i=1}^M \|\mathbf{z}(\tilde{t}_i) - \mathbf{z}_{\text{ODE}}(\tilde{t}_i)\|_2^2$ enforces alignment with trajectories computed by integrating the learned

dynamics function $f_{\theta}(\mathbf{z}, t)$, stabilizing manifold estimation under biological noise. Comparing Li-Attn (peak discrete clustering but limited trajectory coherence) against full LAIOR reveals an intentional trade-off: adding ODE constraints modestly reduces NMI/ARI but substantially improves noise resilience, trajectory directionality, and overall latent quality. LAIOR surpasses trajectory-specialized scTour across overall latent quality, noise resilience, and directionality, demonstrating that geometric and dynamical regularization provide complementary scaffolding.

LAIOR's bottleneck pathway captures coordinated biological factors such as cell cycle phase, lineage states, and stress responses, rather than statistically independent components sought by β -VAE. This design achieves substantially higher noise resilience, a higher partition ratio, and better clustering quality than β -VAE. Latent factor analysis reveals interpretable modules: human bone marrow factors encode neutrophil identity (*MPO*), plasma cell maturation (*IGKC*), and B-cell precursors (*VPREB3*), while mouse pancreatic factors decompose into progenitor specification (*Gast*, *Sox4*), beta cell function (*Nnat*, *Pcsk2*), and islet lineages (*Irx1*, *Sst*). We acknowledge that quantitative disentanglement metrics requiring ground-truth generative factors were not computed; future interventional experiments perturbing individual latent dimensions would rigorously quantify interpretability.

4.2 Relationship to single-cell foundation models

Large single-cell foundation models such as scGPT (Cui et al., 2024) and scFoundation (Hao et al., 2024) represent a complementary paradigm, trained on tens of millions of cells to produce transferable zero-shot embeddings. Our head-to-head comparison on the newly added hematopoietic datasets shows that LAIOR's task-specific geometric and dynamical regularization produces embeddings with substantially higher clustering quality and continuity structure than zero-shot foundation-model embeddings in the count-only trajectory-inference setting (Section 3.7). This is consistent with two expected patterns: (i) foundation models optimize masked-token reconstruction rather than trajectory coherence, so they retain little explicit signal about differentiation dynamics in their zero-shot latents; and (ii) foundation models trained on human atlases transfer imperfectly to mouse datasets even after ortholog-based gene symbol mapping. LAIOR and foundation models are therefore naturally complementary: a promising direction is to fine-tune LAIOR's geometric-dynamical regularization on top of pretrained foundation-model embeddings, combining broad prior knowledge with trajectory-aware inductive bias. We defer this integration to future work.

4.3 Cross-modality versatility and perturbation analysis

LAIOR achieves consistent performance across scRNA-seq and scATAC-seq modalities without requiring architecture modification. On scATAC-seq datasets, LAIOR outperforms modality-specific PeakVI across all metrics despite using standard negative binomial likelihood, demonstrating improved clustering accuracy, noise resilience, and core quality. Likelihood

robustness analysis reveals that architectural principles dominate distributional assumptions: scRNA-seq performance remains largely invariant across NB, ZINB, Poisson, and ZIP specifications, while scATAC-seq exhibits modality-specific patterns (ZIP improves clustering but may reduce continuity). Highly variable feature selection during preprocessing mitigates distributional sensitivity, rendering likelihood choice secondary to feature curation. For scATAC-seq studies where explicit batch covariates and ATAC-tailored likelihood modeling are the primary objectives, PeakVI/PoissonVI remain practical choices; LAIOR is most beneficial when global topology and trajectory continuity are prioritized.

Trajectory inference on *Dapp1* knockout mouse hematopoietic cells disentangles biological dynamics from technical confounders. LAIOR isolates genotype batch effects into latent dimensions orthogonal to pseudotime progression, revealing interpretable modules: cell-cycle decomposition (G1/S and G2/M phases), lineage-specific trajectories (erythroid/megakaryocyte, myeloid, lymphoid), and stem cell quiescence. Encoder-decoder Jaccard scores (0.00–0.47) quantify complementary regulatory versus effector gene selection. LAIOR enables the dissection of perturbation effects along interpretable axes rather than removing signals. Genotype-specific latent factor analysis could reveal whether *Dapp1* knockout alters cell-cycle rates, skews lineage bifurcations, or delays maturation timing.

4.4 Computational considerations and practical guidance

Computational benchmarking on 17 scRNA-seq cancer datasets (7,431–62,035 cells; Table 2, Supplementary Table A13) reveals LAIOR requires ≈ 45 -fold longer training time than baseline scVI (55.9 ± 30.3 min vs. 1.24 ± 0.23 min) and ≈ 2.5 -fold higher peak memory (0.99 ± 0.27 GB vs. 0.40 ± 0.13 GB). ODE integration dominates computational cost: iVAE/LiVAE incur $\approx 2 \times$ overhead, Li-Attn $\approx 4.5 \times$, and full LAIOR $\approx 45 \times$. Training time exhibits sublinear scaling (14–30 min for 7,000–17,000 cells; 78–114 min for 43,000–62,000 cells), indicating practical scalability for atlas-scale datasets. For clustering-focused applications, Li-Attn provides comparable performance at lower computational cost; trajectory-inference applications justify LAIOR's higher expense given the continuity gains it delivers.

Hyperparameter robustness analysis yields actionable guidelines. Strong Lorentz regularization ($\lambda_{\text{geom}} = 10$) optimally balances clustering, embedding quality, and trajectory coherence when ODE regularization is present. Without ODE, geometric loss weight sensitivity increases sharply, highlighting ODE's stabilizing role. Based on empirical evaluation across diverse biological systems, we recommend $\lambda_{\text{geom}} = 10$ for trajectory-focused applications as used throughout this study. The default configuration $\lambda_{\text{geom}} = 5$ provides balanced performance suitable for general-purpose analyses encompassing both clustering and trajectory tasks. For strictly discrete clustering applications where trajectory coherence is not required, $\lambda_{\text{geom}} = 5$ provides optimal performance. The KL annealing parameter β requires minimal tuning; across all experiments, we used $\beta = 1.0$ (standard ELBO) without dataset-

specific optimization. For datasets with extremely deep hierarchies (e.g., immune repertoire clonality, microbial phylogenies), increasing λ_{geom} to 15–20 may enhance global structure preservation.

Likelihood specifications exhibit modality-specific patterns. For scRNA-seq, any standard likelihood suffices (NB, ZINB, Poisson, ZIP), with performance largely invariant across choices. Generalization to scATAC-seq reveals modality-specific sensitivity patterns (Figure 5D; Supplementary Table A10). Zero-inflated models improve clustering performance: ZIP yields higher NMI and ARI than standard NB (effect size = 0.421, $p < 0.001$). However, the standard negative binomial maintains advantages in continuity metrics: higher manifold dimension, anisotropy, and overall quality. This pattern suggests that zero-inflation sharpens cluster boundaries by modeling technical dropouts but may introduce artificial discontinuities in regions representing continuous biological transitions. For scATAC-seq analyses, practitioners should select zero-inflated likelihoods (ZIP, ZINB) when prioritizing discrete cell-type identification, or standard likelihoods (Poisson, NB) when trajectory inference is the primary objective.

4.5 Limitations and future directions

LAIOR embodies assumptions that limit its applicability to specific systems. Lorentz regularization assumes hierarchical, tree-like structures; datasets with dense manifolds, cyclic dynamics, or reticulate evolution may not benefit from it. Neural ODE regularization enforces deterministic trajectories, violating assumptions for stochastic fate decisions or metastable states; future stochastic differential equation extensions could model intrinsic variability. Extending LAIOR to stochastic differential equations is a natural next step and technically straightforward within the current framework. LAIOR does not explicitly perform batch correction: datasets with strong batch effects may require preprocessing or adversarial losses. The geometric loss weight λ_{geom} is treated as a fixed hyperparameter; datasets with heterogeneous hierarchy depths may benefit from spatially varying or learned regularization strength. Interpretability relies on post-hoc gene-latent association; integrating causal inference frameworks could improve interpretability.

Future work should examine datasets with strong batch effects may require preprocessing such as mutual-nearest-neighbor correction (Haghverdi et al., 2018) or adversarial losses paired scRNA-seq/scATAC-seq integration, transfer learning from atlas-scale foundation models, diffusion-based generative extensions, spatial transcriptomics, and applications that require stronger robustness to distribution shift. More broadly, the results here suggest that geometry, dynamics, and bottleneck-based interpretability can be combined effectively for single-cell manifold learning.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

Author contributions

ZF: Writing – review and editing, Visualization, Funding acquisition, Conceptualization, Software, Investigation, Writing – original draft, Resources, Validation, Formal Analysis, Project administration, Methodology, Data curation, Supervision. JF: Funding acquisition, Writing – original draft, Supervision, Formal Analysis, Writing – review and editing, Data curation, Investigation, Visualization, Resources, Project administration, Validation. KZ: Formal Analysis, Supervision, Writing – review and editing, Project administration, Visualization, Writing – original draft, Data curation, Validation. TR: Funding acquisition, Writing – review and editing, Project administration, Investigation, Supervision, Writing – original draft, Software, Validation, Data curation, Conceptualization, Resources, Formal Analysis, Visualization, Methodology. CC: Resources, Visualization, Funding acquisition, Formal Analysis, Validation, Project administration, Writing – original draft, Investigation, Data curation, Supervision, Writing – review and editing, Methodology, Conceptualization, Software.

Funding

The author(s) declared that financial support was received for this work and/or its publication. This work was supported by the Special Project for Enhancing the Scientific and Technological Innovation Capacity of the Army Medical University (No. 2023XJS34; Principal Investigator: Tianfei Ran), the National Key Research and Development Program of China (Grant No. 2024YFA1107101), and the National Key Laboratory of Trauma and Chemical Poisoning (Grant No. 2024K004).

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declared that generative AI was not used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their

affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Abdelaal, T., Michielsen, L., Cats, D., Hoogduin, D., Mei, H., Reinders, M. J. T., et al. (2019). A comparison of automatic cell identification methods for single-cell RNA sequencing data. *Genome Biol.* 20 (1), 194. doi:10.1186/s13059-019-1795-z
- Becht, E., McInnes, L., Healy, J., Dutertre, C. A., Kwok, I. W. H., Ng, L. G., et al. (2019). Dimensionality reduction for visualizing single-cell data using UMAP. *Nat. Biotechnol.* 37 (1), 38–43. doi:10.1038/nbt.4314
- Béginneul, G., and Ganea, O. E. (2019). “Riemannian adaptive optimization methods,” in *International conference on learning representations* (New Orleans, LA: International Conference on Learning Representations / OpenReview.net).
- Bergen, V., Lange, M., Peidli, S., Wolf, F. A., and Theis, F. J. (2020). Generalizing RNA velocity to transient cell states through dynamical modeling. *Nat. Biotechnol.* 38 (12), 1408–1414. doi:10.1038/s41587-020-0591-3
- Brennecke, P., Anders, S., Kim, J. K., Kolodziejczyk, A. A., Zhang, X., Proserpio, V., et al. (2013). Accounting for technical noise in single-cell RNA-Seq experiments. *Nat. Methods* 10 (11), 1093–1095. doi:10.1038/nmeth.2645
- Buenrostro, J. D., Wu, B., Litzenburger, U. M., Ruff, D., Gonzales, M. L., Snyder, M. P., et al. (2015). Single-cell chromatin accessibility reveals principles of regulatory variation. *Nature* 523 (7561), 486–490. doi:10.1038/nature14590
- Buettner, F., Natarajan, K. N., Casale, F. P., Proserpio, V., Scialdone, A., Theis, F. J., et al. (2015). Computational analysis of cell-to-cell heterogeneity in single-cell RNA-Sequencing data reveals hidden subpopulations of cells. *Nat. Biotechnol.* 33 (2), 155–160. doi:10.1038/nbt.3102
- Büttner, M., Miao, Z., Wolf, F. A., Teichmann, S. A., and Theis, F. J. (2019). A test metric for assessing single-cell RNA-seq batch correction. *Nat. Methods* 16 (1), 43–49. doi:10.1038/s41592-018-0254-1
- Cao, Z. J., Wei, L., Lu, S., Yang, D. C., and Gao, G. (2020). Searching large-scale scRNA-seq databases via unbiased cell embedding. *Nat. Commun.* 11 (1), 3458. doi:10.1038/s41467-020-17281-7
- Chen, R. T. Q., Li, X., Grosse, R. B., and Duvenaud, D. K. (2018a). “Isolating sources of disentanglement in variational autoencoders,” in *Advances in neural information processing systems* (Red Hook, NY: Curran Associates, Inc.), 31, 2610–2620.
- Chen, R. T. Q., Rubanova, Y., Bettencourt, J., and Duvenaud, D. K. (2018b). “Neural ordinary differential equations,” in *Advances in neural information processing systems* (Red Hook, NY: Curran Associates, Inc.), 31, 6572–6583.
- Cho, S., Lee, J., and Kim, D. (2023). “Hyperbolic VAE via latent gaussian distributions,” in *Advances in neural information processing systems*, 36, 569–588. doi:10.48550/arXiv.2209.15217
- Choi, Y., Chae, H., and Kim, T. (2023). siVAE: interpretable deep generative models for single-cell transcriptomics. *Genome Biol.* 24 (1), 29. doi:10.1186/s13059-023-02850-y
- Cui, H., Wang, C., Maan, H., Pang, K., Luo, F., Duan, N., et al. (2024). scGPT: toward building a foundation model for single-cell multi-omics using generative AI. *Nat. Methods* 21 (8), 1470–1480. doi:10.1038/s41592-024-02201-0
- Cusanovich, D. A., Daza, R., Adey, A., Pliner, H. A., Christiansen, L., Gunderson, K. L., et al. (2015). Multiplex single-cell profiling of chromatin accessibility by combinatorial cellular indexing. *Science* 348 (6237), 910–914. doi:10.1126/science.aab1601
- Fang, Z., Liu, X., and Peltz, G. (2023). GSEAPy: a comprehensive package for performing gene set enrichment analysis in Python. *Bioinformatics* 39 (1), btac757. doi:10.1093/bioinformatics/btac757
- Fu, Z., Chen, C., Wang, S., Wang, J., and Chen, S. (2025). iVAE: an interpretable representation learning framework enhances clustering performance for single-cell data. *BMC Biol.* 23, 213. doi:10.1186/s12915-025-02315-7
- Ganea, O. E., Béginneul, G., and Hofmann, T. (2018). “Hyperbolic neural networks,” in *Advances in neural information processing systems* (Red Hook, NY: Curran Associates, Inc.), 31, 5350–5360.
- Gayoso, A., Steier, Z., Lopez, R., Regier, J., Nazor, K. L., Streets, A., et al. (2021). Joint probabilistic modeling of single-cell multi-omic data with totalVI. *Nat. Methods* 18 (3), 272–282. doi:10.1038/s41592-020-01050-x
- Grathwohl, W., Chen, R. T. Q., Bettencourt, J., Sutskever, I., and Duvenaud, D. (2019). “FFJORD: Free-Form continuous dynamics for scalable reversible generative models,” in *International conference on learning representations* (New Orleans, LA: International Conference on Learning Representations / OpenReview.net).
- Grün, D., Kester, L., and van Oudenaarden, A. (2014). Validation of noise models for single-cell transcriptomics. *Nat. Methods* 11 (6), 637–640. doi:10.1038/nmeth.2930
- Haghverdi, L., Büttner, M., Wolf, F. A., Buettner, F., and Theis, F. J. (2016). Diffusion pseudotime robustly reconstructs lineage branching. *Nat. Methods* 13 (10), 845–848. doi:10.1038/nmeth.3971
- Haghverdi, L., Lun, A. T. L., Morgan, M. D., and Marioni, J. C. (2018). Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors. *Nat. Biotechnol.* 36 (5), 421–427. doi:10.1038/nbt.4091
- Han, W., Cheng, Y., Chen, J., Zhong, H., Hu, Z., Chen, S., et al. (2022). Self-supervised contrastive learning for integrative single cell RNA-seq data analysis. *Brief. Bioinform* 23 (4), bbac377. doi:10.1093/bib/bbac377
- Hao, M., Gong, J., Zeng, X., Liu, C., Guo, Y., Cheng, X., et al. (2024). Large-scale foundation model on single-cell transcriptomics. *Nat. Methods* 21 (8), 1481–1491. doi:10.1038/s41592-024-02305-7
- Higgins, I., Matthey, L., Pal, A., Burgess, C. P., Glorot, X., Botvinick, M., et al. (2017). “beta-VAE: learning basic visual concepts with a constrained variational framework,” in *International conference on learning representations* (Toulon, France: International Conference on Learning Representations / OpenReview.net).
- Jolliffe, I. T., and Cadima, J. (2016). Principal component analysis: a review and recent developments. *Philos. Trans. A Math. Phys. Eng. Sci.* 374 (2065), 20150202. doi:10.1098/rsta.2015.0202
- Kiselev, V. Y., Andrews, T. S., and Hemberg, M. (2019). Challenges in unsupervised clustering of single-cell RNA-seq data. *Nat. Rev. Genet.* 20 (5), 273–282. doi:10.1038/s41576-018-0088-9
- Klimovskaia, A., Lopez-Paz, D., Bottou, L., and Nickel, M. (2020). Poincaré maps for analyzing complex hierarchies in single-cell data. *Nat. Commun.* 11 (1), 2966. doi:10.1038/s41467-020-16822-4
- Kobak, D., and Berens, P. (2019). The art of using t-SNE for single-cell transcriptomics. *Nat. Commun.* 10 (1), 5416. doi:10.1038/s41467-019-13056-x
- Kolodziejczyk, A. A., Kim, J. K., Svensson, V., Marioni, J. C., and Teichmann, S. A. (2015). The technology and biology of single-cell RNA sequencing. *Mol. Cell* 58 (4), 610–620. doi:10.1016/j.molcel.2015.04.005
- Korsunsky, I., Millard, N., Fan, J., Slowikowski, K., Zhang, F., Wei, K., et al. (2019). Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* 16 (12), 1289–1296. doi:10.1038/s41592-019-0619-0
- Krioukov, D., Papadopoulos, F., Kitsak, M., Vahdat, A., and Boguñá, M. (2010). Hyperbolic geometry of complex networks. *Phys. Rev. E* 82 (3), 036106. doi:10.1103/PhysRevE.82.036106
- Kumar, A., Sattigeri, P., and Balakrishnan, A. (2018). “Variational inference of disentangled latent concepts from unlabeled observations,” in *International conference on learning representations* (Vancouver, BC: International Conference on Learning Representations / OpenReview.net).
- La Manno, G., Soldatov, R., Zeisel, A., Braun, E., Hochgerner, H., Petukhov, V., et al. (2018). RNA velocity of single cells. *Nature* 560 (7719), 494–498. doi:10.1038/s41586-018-0414-6
- Lange, M., Bergen, V., Klein, M., Setty, M., Reuter, B., Bakhti, M., et al. (2022). CellRank for directed single-cell fate mapping. *Nat. Methods* 19 (2), 159–170. doi:10.1038/s41592-021-01346-6
- Laurenti, E., and Göttgens, B. (2018). From haematopoietic stem cells to complex differentiation landscapes. *Nature* 553 (7689), 418–426. doi:10.1038/nature25022
- Li, Q. (2023). scTour: a deep learning architecture for robust inference and accurate prediction of cellular dynamics. *Genome Biol.* 24 (1), 149. doi:10.1186/s13059-023-02988-9
- Locatello, F., Bauer, S., Lucic, M., Rätsch, G., Gelly, S., Schölkopf, B., et al. (2019). “Challenging common assumptions in the unsupervised learning of disentangled representations,” in *Proceedings of the 36th international conference on machine learning (ICML)*, 4114–4124.
- Lopez, R., Regier, J., Cole, M. B., Jordan, M. I., and Yosef, N. (2018). Deep generative modeling for single-cell transcriptomics. *Nat. Methods* 15 (12), 1053–1058. doi:10.1038/s41592-018-0229-2
- Luecken, M. D., Büttner, M., Chaichoompu, K., Danese, A., Interlandi, M., Mueller, M. F., et al. (2022). Benchmarking atlas-level data integration in single-cell genomics. *Nat. Methods* 19 (1), 41–50. doi:10.1038/s41592-021-01336-8
- Luo, E., Hao, M., Wei, L., and Zhang, X. (2024). scDiffusion: conditional generation of high-quality single-cell data using diffusion model. *Bioinformatics* 40 (9), btac518. doi:10.1093/bioinformatics/btac518

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2026.1838613/full#supplementary-material>

- Mathieu, E., Le Lan, C., Maddison, C. J., Tomioka, R., and Teh, Y. W. (2019). "Continuous hierarchical representations with Poincaré variational auto-encoders," in *Advances in neural information processing systems* (Red Hook, NY: Curran Associates, Inc.), 32, 12565–12576.
- McInnes, L., Healy, J., and Melville, J. (2018). UMAP: uniform manifold approximation and projection for dimension reduction. *Arxiv*. 2018.1802.03426. doi:10.48550/arXiv.1802.03426
- Nagano, Y., Yamaguchi, S., Fujita, Y., and Koyama, M. (2019). "A wrapped normal distribution on hyperbolic space for gradient-based learning," in *Proceedings of the 36th international conference on machine learning (ICML)*, 4693–4702.
- Nickel, M., and Kiela, D. (2017). "Poincaré embeddings for learning hierarchical representations," in *Advances in neural information processing systems* (Red Hook, NY: Curran Associates, Inc.), 30, 6338–6347.
- Qiu, X., Zhang, Y., Martin-Rufino, J. D., Weng, C., Hosseinzadeh, S., Yang, D., et al. (2022). Mapping transcriptomic vector fields of single cells. *Cell* 185 (4), 690–711.e45. doi:10.1016/j.cell.2021.12.045
- Saelens, W., Cannoodt, R., Todorov, H., and Saeys, Y. (2019). A comparison of single-cell trajectory inference methods. *Nat. Biotechnol.* 37 (5), 547–554. doi:10.1038/s41587-019-0071-9
- Sarkar, R. (2011). "Low distortion delaunay embedding of trees in hyperbolic plane," in *International symposium on graph drawing*, 355–366.
- Setty, M., Kiseliovas, V., Levine, J., Gayoso, A., Mazutis, L., and Pe'er, D. (2019). Characterization of cell fate probabilities in single-cell data with Palantir. *Nat. Biotechnol.* 37 (4), 451–460. doi:10.1038/s41587-019-0068-4
- Skopek, O., Ganea, O. E., and Bécigneul, G. (2020). "Mixed-curvature variational autoencoders," in *International conference on learning representations (ICLR)*.
- Street, K., Risso, D., Fletcher, R. B., Das, D., Ngai, J., Yosef, N., et al. (2018). Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. *BMC Genomics* 19 (1), 477. doi:10.1186/s12864-018-4772-0
- Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W. M., et al. (2019). Comprehensive integration of single-cell data. *Cell* 177 (7), 1888–1902.e21. doi:10.1016/j.cell.2019.05.031
- Tang, F., Barbacioru, C., Wang, Y., Nordman, E., Lee, C., Xu, N., et al. (2009). mRNA-Seq whole-transcriptome analysis of a single cell. *Nat. Methods* 6 (5), 377–382. doi:10.1038/nmeth.1315
- Tian, T., Wan, J., Song, Q., and Wei, Z. (2019). Clustering single-cell RNA-seq data with a model-based deep learning approach. *Nat. Mach. Intell.* 1 (4), 191–198. doi:10.1038/s42256-019-0037-0
- Tian, T., Zhong, C., Lin, X., Wei, Z., and Hakonarson, H. (2023a). Complex hierarchical structures in single-cell genomics data unveiled by deep hyperbolic manifold learning. *Genome Res.* 33 (2), 232–246. doi:10.1101/gr.277068.122
- Tian, S. W., Ni, J. C., Wang, Y. T., Zheng, C. H., and Ji, C. M. (2023b). scGCC: graph contrastive clustering with neighborhood augmentations for scRNA-seq data analysis. *IEEE J. Biomed. Health Inf.* 27 (12), 6133–6143. doi:10.1109/JBHI.2023.3319551
- Tran, H. T. N., Ang, K. S., Chevrier, M., Zhang, X., Lee, N. Y. S., Goh, M., et al. (2020). A benchmark of batch-effect correction methods for single-cell RNA sequencing data. *Genome Biol.* 21 (1), 12. doi:10.1186/s13059-019-1850-9
- Trapnell, C., Cacchiarelli, D., Grimsby, J., Pokharel, P., Li, S., Morse, M., et al. (2014). The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* 32 (4), 381–386. doi:10.1038/nbt.2859
- van der Maaten, L., and Hinton, G. (2008). Visualizing data using t-SNE. *J. Mach. Learn. Res.* 9 (86), 2579–2605.
- Wagner, D. E., and Klein, A. M. (2020). Lineage tracing meets single-cell omics: opportunities and challenges. *Nat. Rev. Genet.* 21 (7), 410–427. doi:10.1038/s41576-020-0223-2
- Wang, J., Ma, A., Chang, Y., Gong, J., Jiang, Y., Qi, R., et al. (2021). scGNN is a novel graph neural network framework for single-cell RNA-Seq analyses. *Nat. Commun.* 12 (1), 1882. doi:10.1038/s41467-021-22197-x
- Weinreb, C., Wolock, S., Tusi, B. K., Socolovsky, M., and Klein, A. M. (2018). Fundamental limits on dynamic inference from single-cell snapshots. *Proc. Natl. Acad. Sci. U. S. A.* 115 (10), E2467–E2476. doi:10.1073/pnas.1714723115
- Wolf, F. A., Hamey, F. K., Plass, M., Solana, J., Dahlin, J. S., Göttgens, B., et al. (2019). PAGA: graph abstraction reconciles clustering with trajectory inference through a topology preserving map of single cells. *Genome Biol.* 20 (1), 59. doi:10.1186/s13059-019-1663-x
- Xiong, L., Tian, K., Li, Y., Ning, W., Gao, X., and Zhang, Q. C. (2022). Online single-cell data integration through projecting heterogeneous datasets into a common cell-embedding space. *Nat. Commun.* 13 (1), 6118. doi:10.1038/s41467-022-33758-z
- Zhao, S., Song, J., and Ermon, S. (2019). InfoVAE: balancing learning and inference in variational autoencoders. *Proc. AAAI Conf. Artif. Intell.* 33, 5885–5892. doi:10.1609/aaai.v33i01.33015885

Glossary

ARI	Adjusted Rand index	β -VAE β	Variational autoencoder
ASW	Average silhouette width		
BN	Bottleneck		
CAL	Calinski–Harabasz index		
CLEAR	Contrastive LEARNING for scRNA-seq		
COR	Coupling degree		
DAV	Davies–Bouldin index		
DIP-VAE	Disentangled inferred prior VAE		
HWN	Hyperbolic wrapped normal		
InfoVAE	Information maximizing VAE		
KL	Kullback–Leibler		
LAIOR	Lorentz attentive interpretable ODE-regularized VAE		
LPGM	Learnable Poincaré Gaussian mixture		
MD	Manifold dimensionality		
MMD	Maximum mean discrepancy		
NB	Negative binomial		
NMI	Normalized mutual information		
NR	Noise resilience		
ODE	Ordinary differential equation		
PeakVI	Peak variational inference		
PGM	Poincaré Gaussian mixture		
PoissonVI	Poisson variational inference		
PR	Participation ratio		
RM-ANOVA	Repeated measures analysis of variance		
scATAC-seq	Single-cell assay for transposase-accessible chromatin sequencing		
SCALEX	Single-cell ATAC-seq analysis <i>via</i> latent feature extraction		
scDeepCluste	Single-cell deep clustering		
scDHMAP	Single-cell deep hyperbolic manifold learning		
scDiffusion	Single-cell diffusion model		
scGCC	Single-cell graph contrastive clustering		
scGNN	Single-cell graph neural network		
scRNA-seq	Single-cell RNA sequencing		
scTour	Single-cell trajectory inference using ODE-based unsupervised representation		
scVI	Single-cell variational inference		
SDR	Spectral decay rate siVAE scalable interpretable VAE		
TD	Trajectory directionality		
t-SNE	t-distributed stochastic neighbor embedding		
UMAP	Uniform manifold approximation and projection		
VAE	Variational autoencoder		
ZINB	Zero-inflated negative binomial		
ZIP	Zero-inflated poisson		
β -TCVAE	Total correlation VAE with β weighting		