HORIZONTAL FLOW MATCHING FOR SINGLE-CELL SPATIOTEMPORAL DYNAMICS

Anonymous authors
Paper under double-blind review

ABSTRACT

Learning the spatiotemporal dynamics of cells from spatial transcriptomics requires connecting snapshots of different cells acquired in different coordinate frames. A rotation or translation of a section can appear as cell movement and distort the inferred correspondence. We introduce horizontal flow matching, which learns changes in tissue organization, molecular state, and cell abundance up to shared rigid motion. The dynamics follow a continuity–reaction equation whose action assigns no cost to shared rigid motion, and the resulting horizontal projector removes this component from spatial prediction errors during learning and from predicted velocities during generation. An equivalent endpoint problem jointly estimates correspondence and relative pose in Wasserstein–Fisher–Rao geometry and yields analytic paths whose velocity and growth serve as regression targets. Regressing these targets recovers the horizontal dynamics under ideal population conditions. Our model, HORIZON, uses an equivariant transformer to predict velocity and growth from the current tissue, so that a tissue can be generated continuously from a single initial snapshot. On a simulated tissue observed in perturbed coordinate frames, alignment improves reconstruction and projection reduces rigid drift. On four biological time courses, HORIZON attains lower reconstruction error than state-of-the-art methods across spatial, molecular, and abundance metrics, with consistent gains when interpolating held-out snapshots.

1 INTRODUCTION

Development and regeneration change the composition of a tissue, the molecular state of its cells, and their spatial arrangement at the same time. A generative model of these coupled changes would describe how a tissue changes continuously from one observed stage to the next, a step toward virtual tissues (Bunne et al., 2024; Cao et al., 2026). Spatial transcriptomics records cell positions together with gene expression, but the measurement destroys the sample (Ståhl et al., 2016; Chen et al., 2022; Moses & Pachter, 2022). A time course therefore consists of snapshots of different cells, and the dynamics connecting them are not observed (Weinreb et al., 2018; Schiebinger et al., 2019). Specifically, spatial time courses introduce an additional ambiguity: tissue sections collected at different times are generally observed in independent coordinate frames. A generative model must therefore distinguish changes in the tissue itself from changes in the frame.

Dynamic optimal transport (OT) connects distributions by the path of least kinetic action (Benamou & Brenier, 2000; Villani, 2009), and flow matching learns such paths by regressing velocity targets along conditional interpolants (Lipman et al., 2023; Liu et al., 2023a). Trajectory inference applies these ideas to single-cell time courses (Tong et al., 2020; Huguet et al., 2022). When abundance changes, unbalanced OT also allows mass to be created and destroyed (Chizat et al., 2018c; Liero et al., 2018), and WFR-FM regresses velocity and growth jointly along closed-form Wasserstein–Fisher–Rao (WFR) paths (Peng et al., 2026a). However, these formulations describe dynamics in the recorded coordinates and do not separate tissue deformation from a shared change of frame.

This ambiguity is specific to spatial measurements, since sections collected at different times are treated separately and can differ by rotations and translations unrelated to biological change (Zeira et al., 2022; Clifton et al., 2023). Such frame differences affect both learning and generation: they can distort cell correspondence, make rigid frame motion appear as cellular motion, and induce rigid drift when the learned velocity is integrated. Pre-aligning the sections addresses the first two issues but

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

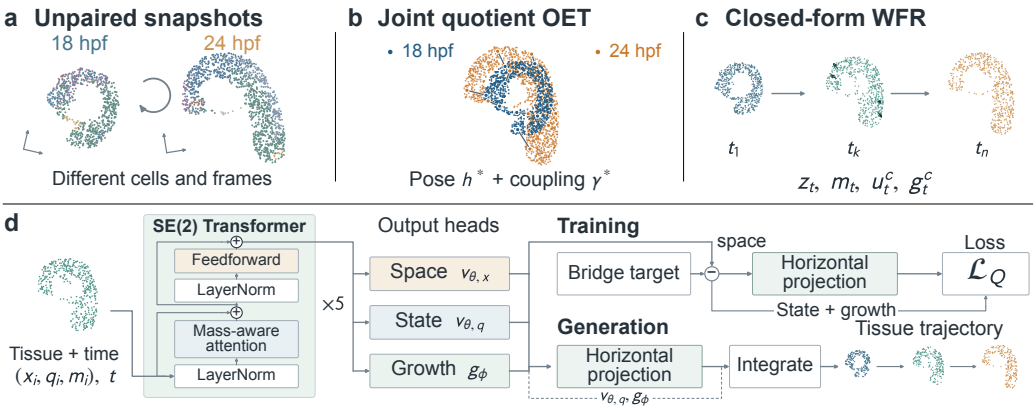

Figure 1: **Horizontal flow matching with HORIZON. (a)** Snapshots contain different cells in unrelated frames. **(b)** Quotient optimal entropy transport (qOET) jointly estimates a rigid pose and an unbalanced correspondence between cells. **(c)** Closed-form WFR bridges give intermediate states, masses, velocities, and growth rates. **(d)** An equivariant transformer learns velocity and growth from these states, and the horizontal projector is applied both in the spatial loss and to the velocity used for integration. Panels (a–c) use the 18–24 hpf zebrafish interval.

does not constrain rigid components that may reappear during generation. stVCR jointly learns rigid alignment and neural ODE dynamics by simulation (Peng et al., 2026b; Chen et al., 2018), while STORIES avoids a shared frame by comparing within-section distances but does not learn spatial velocity (Huizing et al., 2026). Quotient-space diffusion removes rigid motion from updates of fixed atomic structures (Xu et al., 2026), and rotation-invariant dynamic OT removes it from transport between fixed-mass measures (Meugang Toukam et al., 2026). Tissue time courses additionally involve different cells, changing mass, and coupled spatial and molecular dynamics. This setting therefore requires a notion of motion relative to the tissue itself that is compatible with unbalanced transport and used consistently in the training paths, regression loss and integrated velocity.

We introduce *horizontal flow matching* to learn tissue dynamics up to shared rigid motion. Starting from WFR continuity–reaction dynamics, we assign no cost to a moving rigid frame, which yields a horizontal projector that removes the best-fitting translation and rotation from the spatial velocity. The same projector is used in both training and generation, so the model captures changes in tissue organization while factoring out unidentifiable frame motion. We show that the resulting dynamic problem is equivalent to jointly estimating endpoint correspondence and relative pose, yielding closed-form WFR paths whose velocity and growth provide regression targets without ODE solves; projected regression recovers the horizontal dynamics at the population optimum. Our model, HORIZON, combines these targets with an equivariant transformer to generate continuous tissue trajectories from the initial snapshot alone (Figure 1). In our experiments, endpoint alignment improves reconstructed tissue shape, while horizontal projection suppresses rigid drift during generation. Our contributions are threefold:

- **A learning principle.** Horizontal flow matching defines tissue dynamics up to shared rigid motion, using a single projector consistently in spatial regression and generation while learning molecular change and cellular populatoin growth.
- **A geometric foundation.** The quotient WFR dynamic problem reduces to endpoint alignment, with qOET jointly estimating relative pose and correspondence; projected regression recovers the resulting horizontal dynamics at the population optimum.
- **A model and evaluation.** HORIZON combines joint alignment, closed-form WFR bridge targets, and equivariant prediction. We evaluate it on a benchmark with hidden frame perturbations and on four biological time courses, separating the effects of alignment and projection.

2 RELATED WORK

Transport and flow matching for cell dynamics. Dynamic optimal transport describes the path between two distributions through a continuity equation and a kinetic action (Benamou & Brenier, 2000; Villani, 2009), and unbalanced transport extends the description to populations whose mass changes (Chizat et al., 2018a;c; Liero et al., 2016; Kondratyev et al., 2016). Flow matching learns a velocity field by regressing targets along conditional paths (Lipman et al., 2023; Liu et al., 2023a; Albergo & Vanden-Eijnden, 2023; Tong et al., 2024a). In single-cell biology, transport connects snapshots of different cells (Schiebinger et al., 2019; Lavenant et al., 2024; Klein et al., 2025), dynamical models learn the connecting fields (Tong et al., 2020; Huguet et al., 2022; Neklyudov et al., 2023; 2024), and growth enters through unbalanced or branching formulations (Fischer et al., 2019; Sha et al., 2024; Zhang et al., 2025; Pariset et al., 2023; Tong et al., 2024b). WFR-FM learns movement and growth from closed-form WFR geodesics (Peng et al., 2026a; Liero et al., 2018), and related methods fit velocity and growth jointly or adapt paths to the data geometry (Wang et al., 2025; Kapuśniak et al., 2024). Our formulation penalizes only spatial motion relative to the tissue.

Dynamics and registration in spatial transcriptomics. stVCR pairs an initial OT-based alignment with rigid transformations that are optimized jointly with neural ODE dynamics (Peng et al., 2026b). STORIES learns a cell fate potential with fused Gromov–Wasserstein (FGW) transport (Mémoli, 2011; Vayer et al., 2020), in which molecular discrepancies form the linear cost and a quadratic term compares within-section spatial distances (Huizing et al., 2026). These distances do not change under rigid motion of a section, so STORIES needs no common frame, but it also learns no spatial velocity. NicheFlow models transitions between spatial microenvironments (Sakalyan et al., 2025). FlowCloud decodes point clouds from a neural ODE latent space trained with Chamfer and Sinkhorn losses and needs rigid pre-alignment when sections are rotated by more than 30° (Liu et al., 2026). Registration methods such as PASTE, PASTE2, GPSA, and STalign align observed sections (Zeira et al., 2022; Liu et al., 2023b; Jones et al., 2023; Clifton et al., 2023), building on Procrustes analysis (Gower, 1975; Umeyama, 1991). In each case the frame is fixed outside the dynamics.

Equivariance and quotient geometry. Equivariant generative models constrain how outputs transform when the input frame changes (Köhler et al., 2020; Satorras et al., 2021; Hoogeboom et al., 2022; Klein et al., 2023). Quotient-space methods identify states related by a symmetry, as in shape spaces, quotient diffusion, flow matching on permutation orbits, and rotation-invariant transport (Kendall, 1984; Xu et al., 2026; Wang et al., 2026; Meugang Toukam et al., 2026). Meugang Toukam et al. (2026) derive a Procrustes–Wasserstein distance from a rotation-invariant Benamou–Brenier action between probability measures. Theorem 2 establishes the equivalence between the dynamic and static formulations of unbalanced WFR transport with molecular state; the solution supplies the training targets of Section 4. Equivariance makes the model consistent across frames, but an equivariant model can still predict a shared rigid motion, which the projector removes.

3 TISSUE DYNAMICS IN A MOVING COORDINATE FRAME

As summarized in Figure 1, the method aligns consecutive snapshots and matches their cells (panel b), follows closed-form WFR bridges between matched cells to obtain targets (panel c), and regresses these targets while ignoring shared rigid motion, which is also removed during generation (panel d). This section derives the projector from the WFR action and shows that the resulting dynamic problem is equivalent to aligning the endpoints.

3.1 EVOLVING TISSUE MEASURES

A cell state $z = (x, q)$ pairs a position $x \in \mathbb{R}^{d_x}$ with a molecular representation $q \in \mathbb{R}^{d_q}$. A snapshot is the weighted point cloud $\mu = \sum_i a_i \delta_{(x_i, q_i)}$, where $a_i > 0$ is the abundance represented by state i. A tissue is thus a finite nonnegative measure on $\mathcal{Z} = \mathbb{R}^{d_x} \times \mathbb{R}^{d_q}$, and its total mass is its total abundance. We compare states through $d^2_\Lambda((x, q), (x', q')) = \lambda_x \|x - x'\|^2 + \lambda_q \|q - q'\|^2$, where $\lambda_x, \lambda_q > 0$ weight spatial and molecular displacement.

Let ρ_t denote the tissue at time t, $u = (u_x, u_q)$ its velocity in position and molecular state, and g its relative growth rate. The continuity–reaction equation balances transport against changes in

abundance. On an interval rescaled to $[0, 1]$, the WFR distance selects the fields joining $\rho_0 = \mu_0$ to $\rho_1 = \mu_1$ with least action:

$$\mathrm{WFR}^2_\delta(\mu_0, \mu_1) = \inf_{\rho,u,g} \frac{1}{2} \int_0^1 \int_{\mathcal{Z}} \left(\lambda_x \|u_x\|^2 + \lambda_q \|u_q\|^2 + \delta^2 g^2\right) \mathrm{d}\rho_t \mathrm{d}t, \quad \partial_t \rho + \nabla_x \cdot (\rho u_x) + \nabla_q \cdot (\rho u_q) = g\rho, \tag{1}$$

where $\delta > 0$ controls the cost of mass change. This action penalizes every spatial displacement, including a translation or rotation of the whole tissue, which is exactly what a change of coordinate frame between two sections produces.

3.2 Removing frame motion from the action

The group $G = \mathrm{SE}(d_x)$ of proper rotations and translations acts on position only, $h \cdot (x, q) = (Rx + b, q)$ for $h = (R, b)$. The pushforward $h_\# \mu$ moves every cell of μ by the same rigid motion. We write $[\mu] := \{h_\# \mu : h \in G\}$ for the orbit of μ, i.e., the equivalence class of tissue measures that differ only by a shared rigid motion. Measures in the same orbit differ only by a shared rigid transformation of their spatial coordinates, and therefore preserve relative spatial geometry, scale, chirality, molecular states, and weights. A moving rigid frame has velocity $a_t + \Omega_t x$ with $\Omega_t^\top = -\Omega_t$. We add this field to the spatial transport in Equation (1) without charging for it:

$$\overline{\mathcal{A}}_\delta(\mu_0, \mu_1) = \inf_{\rho,u,g,a,\Omega} \frac{1}{2} \int_0^1 \int \left(\lambda_x \|u_x\|^2 + \lambda_q \|u_q\|^2 + \delta^2 g^2\right) \mathrm{d}\rho_t \mathrm{d}t, \tag{2}$$
$$\text{s.t.} \quad \partial_t \rho + \nabla_x \cdot \{\rho(u_x + a + \Omega x)\} + \nabla_q \cdot (\rho u_q) = g\rho, \quad \rho_0 = \mu_0,\ \rho_1 = \mu_1.$$

Only motion relative to the frame now contributes to the spatial cost, while molecular velocity and growth are penalized as before.

3.3 Horizontal velocity and its projector

Fix the total spatial velocity $v = u_x + a + \Omega x$ at one time. The action penalizes only $u_x = v - a - \Omega x$, so the optimal frame minimizes $\int \|v - a - \Omega x\|^2 \mathrm{d}\rho$ over translations a and skew matrices Ω. This is a weighted least-squares fit of a rigid field to v. The residual of the fit is orthogonal to every rigid field in $L^2(\rho)$, and we call it the horizontal part of v. The fit depends on the mass $M = |\rho|$, the centroid $\bar{x} = M^{-1} \int x \mathrm{d}\rho$, the mean velocity $\bar{v} = M^{-1} \int v \mathrm{d}\rho$, the inertia $S = \int (x - \bar{x})(x - \bar{x})^\top \mathrm{d}\rho$, and the cross-moment $C = \int (v - \bar{v})(x - \bar{x})^\top \mathrm{d}\rho$.

Proposition 1 (Closed-form horizontal projection)**.** *The least-squares rigid component of v is $a^\star = \bar{v} - \Omega^\star \bar{x}$, where the minimum-norm skew solution satisfies $\Omega^\star S + S\Omega^\star = C - C^\top$. Consequently*

$$P^H_\rho v(x, q) = v(x, q) - \bar{v} - \Omega^\star (x - \bar{x}) \tag{3}$$

is orthogonal in $L^2(\rho)$ to every rigid velocity field, and molecular velocity and growth are unaffected.

In the planar case of our experiments, $\Omega^\star = \omega^\star J$, where J is the rotation by a right angle and $\omega^\star = \int (x - \bar{x})^\top J^\top (v - \bar{v}) \mathrm{d}\rho \,/ \int \|x - \bar{x}\|^2 \mathrm{d}\rho$ is the angular momentum about the centroid divided by the moment of inertia (Appendix A.1). Substituting the optimal frame into Equation (2) leaves the ordinary continuity equation in v and replaces the spatial kinetic term by $\lambda_x \left\| P^H_{\rho_t} v \right\|^2$. A path whose spatial velocity is horizontal at every time is called a *horizontal lift* of the corresponding path of tissue shapes.

3.4 An equivalent endpoint problem

Solving Equation (2) directly requires optimizing over paths, but the action also has a static form. Define the distance between endpoint orbits by

$$\overline{\mathrm{WFR}}_\delta([\mu], [\nu]) := \inf_{h \in G} \mathrm{WFR}_\delta(\mu, h_\# \nu). \tag{4}$$

A single pose acts on the whole target measure, so optimizing h cannot move cells individually. Because G acts by isometries, the value does not depend on which representatives are chosen.

Theorem 2 (The equivalence between the dynamic and static formulations)**.** *For nonzero finite atomic measures on $\mathcal{Z}$ and $G = \mathrm{SE}(d_x)$, the infimum in Equation* (4) *is attained, $\overline{\mathrm{WFR}}_\delta$ is a metric on the orbit space, and $\overline{\mathcal{A}}_\delta(\mu_0, \mu_1) = \min_{h\in G} \mathrm{WFR}^2_\delta(\mu_0, h_\#\mu_1) = \overline{\mathrm{WFR}}^2_\delta([\mu_0], [\mu_1])$.*

The action identity follows from two bounds (Appendix A.2). In its co-moving rigid frame, any admissible path becomes an ordinary WFR path of the same cost ending at a representative of $[\mu_1]$, which gives the lower bound. Conversely, a rigid frame at zero cost carries the WFR geodesic to the optimally posed target $h^\star{}_\#\mu_1$ back to μ_1, which gives the upper bound.

4 Horizontal flow matching

By Theorem 2, aligning the endpoints solves the dynamic problem. This section makes the solution explicit and turns it into a regression loss.

4.1 Joint alignment

Alignment and correspondence depend on each other. The relative pose determines which cells are close, and the correspondence determines how the sections should be aligned, so we estimate the two together. For snapshots $\mu = \sum_i a_i\delta_{z_i}$ and $\nu = \sum_j b_j\delta_{z'_j}$, let $\gamma_{ij} \geq 0$ be the mass transported from state i to state j. Quotient optimal entropy transport (qOET) solves

$$(h^\star, \gamma^\star) \in \operatorname*{arg\,min}_{h\in G,\, \gamma\geq 0} 2\delta^2 \left[\left\langle c_\delta(z_i, hz'_j), \gamma_{ij} \right\rangle + \mathrm{KL}(\gamma\mathbf{1} \mid a) + \mathrm{KL}(\gamma^\top\mathbf{1} \mid b) \right], \tag{5}$$

with $c_\delta(z, z') = -2\log\cos(d_\Lambda(z, z')/(2\delta))$ when $d_\Lambda(z, z') < \pi\delta$ and $c_\delta = \infty$ otherwise. The first term is the transport cost after alignment, and the two Kullback–Leibler terms let the marginals of the coupling deviate from the observed abundances, which accounts for the creation and destruction of mass (Liero et al., 2018; Chizat et al., 2018b). At a fixed pose, the unregularized minimum equals $\mathrm{WFR}^2_\delta(\mu, h_\#\nu)$ (Liero et al., 2018). Minimizing over the pose as well gives $\overline{\mathrm{WFR}}^2_\delta$ and hence, by Theorem 2, the dynamic action. A single pose is shared by all pairs, whereas the coupling may split the mass of one cell among several destinations.

4.2 Bridges and training targets

The coupling pairs the endpoints but does not specify the intermediate states. Along the coupling, the WFR geodesic between two measures splits into independent paths of single weighted points, which we call bridges, so it suffices to describe one pair. Write $\gamma = \gamma^\star$. A pair (i, j) with $\gamma_{ij} > 0$ receives a source mass and a target mass,

$$m_0 = a_i \frac{\gamma_{ij}}{(\gamma\mathbf{1})_i}, \qquad m_1 = b_j \frac{\gamma_{ij}}{(\gamma^\top\mathbf{1})_j}, \tag{6}$$

which differ when abundance changes along the pair. Set $z_0 = z_i$ and $\bar{z}_1 = h^\star z'_j$. The bridge moves one weighted point, called a carrier, from z_0 to $\bar{z}_1$ while its mass changes from m_0 to m_1. With $\vartheta = d_\Lambda(z_0, \bar{z}_1)/(2\delta) < \pi/2$ the half-angle between the two endpoints, the WFR geodesic and its derivatives are

$$m_t = (1-t)^2 m_0 + t^2 m_1 + 2t(1-t)\sqrt{m_0 m_1}\cos\vartheta, \qquad g^c_t = \dot{m}_t/m_t, \tag{7}$$

$$\alpha_t = \vartheta^{-1}\operatorname{atan2}(t\sqrt{m_1}\sin\vartheta,\ (1-t)\sqrt{m_0} + t\sqrt{m_1}\cos\vartheta),$$

$$z_t = (1-\alpha_t)z_0 + \alpha_t\bar{z}_1, \qquad u^c_t = \frac{\sqrt{m_0 m_1}\sin\vartheta}{\vartheta m_t}(\bar{z}_1 - z_0). \tag{8}$$

The carrier moves along the segment from z_0 to $\bar{z}_1$ at a speed that depends on its mass change, and α_t is the fraction of the segment covered at time t. The velocity $u^c_t = \dot{z}_t$ and the growth rate g^c_t are available in closed form, so no finite differences are needed (Appendix A.4 treats $\vartheta = 0$). A source state that is not transported, $(\gamma\mathbf{1})_i = 0$, decays in place with mass $(1-t)^2 a_i$, and an untransported target state, $(\gamma^\top\mathbf{1})_j = 0$, appears at its aligned position with mass $t^2 b_j$. Summing all components gives the intermediate tissue at time t, which interpolates between μ and $h^\star{}_\#\nu$.

4.3 Regressing velocity and growth

The network predicts $(v_{\theta,x}, v_{\theta,q}, g_\phi)[\rho](z,t)$ from the current tissue and time. A training step draws one interval and one local time t, then samples bridge states $z_i = (x_i, q_i)$ with masses m_i, velocities $u_i = (u_i^x, u_i^q)$, and growth rates g_i from the intermediate tissue at that time. Because all sampled states belong to the same tissue, we fit one rigid field to all of their spatial errors $e_i = v_{\theta,x}(z_i,t) - u_i^x$, $(a_B, \Omega_B) = \arg\min_{a,\Omega^\top=-\Omega} \sum_i m_i \|e_i - a - \Omega x_i\|^2$, and define the loss as

$$\mathcal{L}_Q = \mathbb{E} \sum_i m_i \big[\lambda_x \|e_i - a_B - \Omega_B x_i\|^2 + \lambda_q \|v_{\theta,q}(z_i,t) - u_i^q\|^2 + \delta^2 |g_\phi(z_i,t) - g_i|^2 \big]. \quad (9)$$

The spatial term applies the horizontal norm of Section 3.3 to the regression residual, while molecular velocity and growth are fitted directly. The rigid field must be shared across cells, because a separate rigid field per cell could absorb any residual and would remove the spatial supervision entirely. The next results explain why regressing sampled bridges recovers the intended dynamics even when several bridges pass through the same state.

4.4 Population consistency

Optimal bridges have a horizontal aggregate velocity (Lemma 3), and the projected regression loss recovers this aggregate (Theorem 4). Proofs are in Appendices A.6 and A.7, respectively.

Optimal alignment gives horizontal target dynamics. Several bridges may pass through one state at time t, and the aggregate field averages their velocities weighted by current mass. For an exact qOET solution, this field contains no shared rigid motion.

Lemma 3 (Optimally aligned geodesics are horizontal)**.** *Let $h^\star$ attain Equation* (4) *and let $\widetilde{\rho}_t$ be the WFR geodesic from μ_0 to $h^\star{}_\#\mu_1$ obtained by marginalizing the bridges of an exact qOET solution, with spatial aggregate velocity $\widetilde{u}_x$. Then $\widetilde{u}_x = P^H_{\widetilde{\rho}_t}\widetilde{u}_x$ in $L^2(\widetilde{\rho}_t)$ for almost every t; that is, $\widetilde{\rho}_t$ is a horizontal lift of a quotient geodesic.*

Intuitively, removing a nonzero rigid component would lower the action without changing the endpoint shapes, which would contradict optimality. The projector therefore leaves the ideal target unchanged. A field learned from finite data may still contain a rigid component, and since the loss in Equation (9) does not penalize it, we apply the projector during generation to remove it.

Bridge regression recovers the aggregate field. In ordinary conditional flow matching, regression averages the conflicting bridge velocities that pass through a state (Lipman et al., 2023; Tong et al., 2024a). The following identity shows that this remains true when spatial errors are measured up to a rigid field. Here Z_t is a bridge state sampled according to mass, $\mathfrak{se}(d_x)$ denotes the rigid fields $a + \Omega x$, and a tissue is regular when it has positive mass and no nonzero rigid field vanishes on its support.

Theorem 4 (Horizontal conditional flow matching identity)**.** *Fix t, let U^c be the conditional bridge velocity with aggregate $\bar{U}(z) = \mathbb{E}[U^c \mid Z_t = z]$, and assume ρ_t is regular. For every candidate spatial field v,*

$$\inf_{\xi \in \mathfrak{se}(d_x)} \mathbb{E} \|v(Z_t) - U^c - \xi(Z_t)\|^2 = \inf_{\xi \in \mathfrak{se}(d_x)} \mathbb{E} \big\|v(Z_t) - \bar{U}(Z_t) - \xi(Z_t)\big\|^2 + C_t, \quad (10)$$

where C_t does not depend on v or ξ. Hence every population minimizer satisfies $P^H_{\rho_t} v = P^H_{\rho_t} \bar{U}$.

Together, these results identify the qOET geodesic's horizontal and molecular velocities and growth under exact population regression. Integration recovers this path when the continuity equation is well posed (Corollary A.4), and adjacent optimal paths concatenate (Proposition A.5). The guarantees require a global qOET optimum, regular tissues, and exact regression over all fields. The implementation uses nonconvex pose search, sampled batches, and a fixed number of carriers (Appendix B.1).

5 The HORIZON model

HORIZON solves qOET once per interval, trains a transformer on the resulting bridges, and integrates the learned fields from the first snapshot (Algorithm 1).

Computing supervision. We solve Equation (5) by alternating between the pose and the coupling. For a fixed pose, the coupling solver works on a sparse set of candidate pairs and adds improving pairs from the rest until the relative primal–dual gap over the full support falls below 10^{-4}, which certifies the coupling without storing the full cost matrix. Because the joint problem is nonconvex, we start from several initial poses and accept a pose update only if it lowers the objective after the coupling is re-solved. The certificate thus applies to the coupling at the chosen pose but does not guarantee that the pose is globally optimal (Appendix B).

Predicting from the current tissue. The network is a transformer over the unordered set of cells in the current tissue (Vaswani et al., 2017; Lee et al., 2019). Each cell is embedded from its molecular state, time, and position relative to the tissue moments, not its own weight. Each pair of cells adds two invariant features to the attention logits: the dot product of their centered positions and the signed area these positions span. Residual attention blocks update the embeddings, so every prediction depends on the surrounding population. Adding $\log m_j$ to the attention logit of cell j weights its contribution by the abundance it represents. Splitting a weight among coincident copies therefore leaves the output unchanged, so predictions depend only on the tissue measure. Since distances, dot products, and signed areas are unchanged by translation and proper rotation, all hidden features are invariant. The spatial head outputs invariant coefficients that multiply a basis of directions attached to the tissue: the centered position $x - \bar{x}$, its rotation $J(x - \bar{x})$, the covariance-transformed direction $\Sigma(x - \bar{x})$, and pairwise differences aggregated over the population. Every basis vector rotates with the tissue, so the predicted spatial velocity rotates as well. Separate heads output molecular velocity and growth, which are invariant (Proposition A.6). The construction follows equivariant point-cloud networks (Satorras et al., 2021), and equivariance lets bridges from different source frames train a single model.

Training. Each update passes a batch of weighted bridge states from one interval and one time through the transformer and evaluates Equation (9). Attention and the batch projector both operate on this sampled population. On the real datasets we train for 9,000 Adam (Kingma & Ba, 2015) updates (Table 3).

Generating a trajectory. Generation starts from all cells of the initial snapshot. At every RK4 stage, the network evaluates the current tissue, and the projector is recomputed from its mass, centroid, and inertia. Positions then follow $P_\rho^H v_{\theta,x}$, molecular states follow $v_{\theta,q}$, and log weights follow g_ϕ, with a maximum step of 0.25 in normalized time. The number of carriers stays fixed, and generation uses neither endpoint correspondences nor restarts from later observations.

6 Experiments

Data. We use axolotl brain regeneration, chicken heart development, and Drosophila and zebrafish embryogenesis (Wei et al., 2022; Mantri et al., 2021; Qiu et al., 2024; Liu et al., 2022), each with four or five snapshots and 3,550–46,209 cells in total. Expression is represented by 32 principal components fitted on each series. Coordinates keep their recorded orientation and origin, and mass is the cell count relative to the initial snapshot. All retained cells are used in qOET and in evaluation, and the architecture and optimizer settings are the same for all datasets (Table 3).

Baselines and protocols. We compare with WFR-FM, stVCR, STORIES, and NicheFlow on the same processed inputs. HORIZON and the first three baselines generate the whole trajectory from the initial snapshot, whereas NicheFlow restarts each interval from its observed source (adjacent teacher forcing). We therefore also evaluate HORIZON with these restarts, using the same trained model. Because restarts provide observed intermediate states, Table 2 reports the two protocols separately (Appendix D). STORIES predicts only molecular states, and NicheFlow does not predict mass and is run on both raw and qOET-aligned coordinates.

Metrics. Sliced Wasserstein distance (SW) (Bonneel et al., 2015) and maximum mean discrepancy (MMD) (Gretton et al., 2012) compare molecular and joint spatial–molecular distributions, spatial SW and symmetric Chamfer distance compare geometry, and relative error (RE) measures total abundance. We use two kinds of spatial scores to separate the effects of alignment and projection.

Table 1: **Simulated tissue under hidden changes of coordinate frame.** Mean ± standard error over completed fits (n = completed/attempted, severe regime). *Direct* spatial SW is scored against intrinsic ground truth without alignment, measuring sensitivity to the frames used for training; *quotient* metrics are computed after $\mathrm{SE}(2)$ alignment. The indented rows use the same network and training settings as HORIZON. *Registered, no projector* aligns the snapshots in advance with the qOET poses and then trains and integrates without the projector, *no quotient* does the same on the raw snapshots, and *no growth* sets the reaction field to zero.

		Direct spatial SW		Severe regime			
Method	n	Canonical	Severe	Quotient sp. SW	Quotient joint SW	Gene SW	Mass RE (%)
HORIZON (ours)	9/9	0.111±0.004	0.115±0.003	0.042±0.002	0.037±0.001	0.070±0.005	0.02±0.01
registered, no projector	9/9	–	0.127±0.004	0.039±0.002	0.034±0.001	0.073±0.007	0.02±0.01
no quotient (raw, no projector)	9/9	–	1.246±0.093	0.113±0.017	0.121±0.039	0.195±0.065	0.23±0.04
no growth (reaction set to zero)	9/9	–	0.117±0.003	0.045±0.003	0.039±0.002	0.085±0.004	9.32±0.00
WFR-FM	9/9	0.223±0.165	1.213±0.027	0.110±0.008	0.142±0.016	0.298±0.028	0.14±0.03
stVCR	9/9	0.115±0.015	0.249±0.021	0.119±0.006	0.100±0.005	0.160±0.012	2.18±0.20

Table 2: **Reconstruction of real spatial time courses.** Lower is better. Scores average targets after the initial snapshot, with a common evaluator and proper $\mathrm{SE}(2)$ alignment. *first*: generation from the initial snapshot. *adj. TF*: each interval starts from its observed source. Rules separate protocols; bold marks the best within each. HORIZON (first) and WFR-FM average three seeds, other rows use one seed, and HORIZON (adj. TF) reuses its seed-17 model. Dashes mark outputs not predicted or scored (Appendix D).

Dataset	Method	Protocol	Joint SW	Joint MMD	Spatial SW	Chamfer	Gene SW	Gene MMD	Mass RE (%)
Axolotl	HORIZON	first	**0.063**	**0.003**	**0.240**	**0.367**	**0.122**	**0.018**	**0.09**
	WFR-FM	first	1.441	0.167	5.907	2.860	1.171	0.214	146.3
	stVCR	first	1.040	0.218	4.294	1.008	0.248	0.044	1.30
	STORIES	first	–	–	–	–	1.087	0.207	–
	HORIZON	adj. TF	**0.054**	**0.001**	**0.224**	**0.323**	**0.046**	**0.006**	–
	NicheFlow raw	adj. TF	0.201	0.025	0.843	0.360	0.279	0.060	–
	NicheFlow aligned	adj. TF	0.270	0.038	1.055	0.383	0.345	0.076	–
Chicken	HORIZON	first	**0.134**	**0.037**	**0.199**	**0.392**	**0.597**	**0.098**	**0.45**
	WFR-FM	first	0.187	0.064	0.540	0.576	0.751	0.135	1.56
	stVCR	first	0.576	0.205	1.806	2.079	1.518	0.262	99.2
	STORIES	first	–	–	–	–	1.476	0.271	–
	HORIZON	adj. TF	**0.061**	**0.014**	**0.081**	0.224	0.252	**0.034**	–
	NicheFlow raw	adj. TF	0.064	0.025	0.198	0.150	**0.217**	0.042	–
	NicheFlow aligned	adj. TF	0.079	0.035	0.268	**0.149**	0.231	0.047	–
Drosophila	HORIZON	first	**0.035**	**0.013**	**0.043**	**0.125**	**0.107**	**0.023**	**0.10**
	WFR-FM	first	0.074	0.031	0.064	0.152	0.157	0.032	0.23
	stVCR	first	0.228	0.112	0.836	0.669	0.427	0.110	11.5
	STORIES	first	–	–	–	–	0.797	0.241	–
	HORIZON	adj. TF	**0.018**	**0.005**	**0.029**	**0.103**	**0.081**	**0.017**	–
	NicheFlow raw	adj. TF	0.059	0.030	0.158	0.140	0.160	0.045	–
	NicheFlow aligned	adj. TF	0.088	0.046	0.185	0.142	0.175	0.051	–
Zebrafish	HORIZON	first	**0.071**	**0.009**	**0.206**	**0.358**	**0.289**	**0.058**	**0.51**
	WFR-FM	first	1.035	0.160	4.049	3.142	1.441	0.374	45.0
	stVCR	first	1.862	0.331	7.394	7.855	1.862	0.377	99.5
	STORIES	first	–	–	–	–	0.831	0.215	–
	HORIZON	adj. TF	**0.028**	**0.003**	**0.060**	**0.184**	**0.108**	**0.023**	–
	NicheFlow raw	adj. TF	0.135	0.029	0.537	0.264	0.243	0.058	–
	NicheFlow aligned	adj. TF	0.141	0.029	0.561	0.222	0.251	0.052	–

Quotient metrics first align each prediction to its target by a proper $\mathrm{SE}(2)$ pose (Umeyama, 1991), using the same evaluator for every method, and thus measure error in tissue shape. *Direct* metrics skip this alignment and so also penalize rigid drift. Because they require the intrinsic frame, they are reported only in simulation.

6.1 ROBUSTNESS TO CHANGES OF COORDINATE FRAME

A simulated tissue bends, contracts, and changes mass and molecular state. Each of three intrinsic trajectories has six snapshots and is observed in one canonical, three mildly perturbed, and three severely perturbed frames. The applied poses are hidden from training, and Table 1 scores the five snapshots after the first.

Sensitivity to the frame. As the perturbation grows from canonical to severe, the direct spatial SW of HORIZON rises only from 0.111 to 0.115, while that of WFR-FM grows more than fivefold and

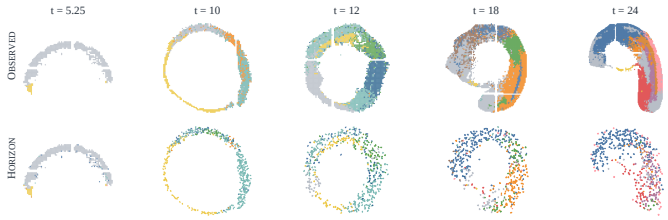

Figure 2: **Reconstructing zebrafish development from the initial snapshot.** Observations are above and the seed 17 generation below; time is in hours post fertilization. Predictions use no restarts and are aligned for evaluation. Points carry equal mass, so point count shows abundance; colors show observed annotations above and annotations decoded from predicted expression below (Appendix J).

that of stVCR more than doubles. HORIZON also has lower quotient spatial and joint errors than both baselines, so its advantage extends to tissue shape.

Contributions of alignment and projection. Aligning the snapshots in advance with the qOET poses but omitting the projector gives a quotient spatial SW of 0.039. This is close to the 0.042 of the full model and far below the 0.113 obtained without either ingredient, so alignment accounts for the gain in reconstructed shape. The projector targets rigid drift, which evaluation alignment removes, so its effect appears in the direct metrics. In the paired control it reduces direct spatial SW from 0.127 to 0.116 and lowers direct spatial SW in 7 of 9 paired variants and direct Chamfer in all 9, while the paired difference in quotient spatial SW stays within one standard error of zero. Setting growth to zero raises mass RE from 0.02% to 9.3% with little change in spatial error (Appendices C and F give the full benchmark and sensitivity).

6.2 RECONSTRUCTION AND INTERPOLATION OF REAL TISSUES

Reconstruction from the initial snapshot. In Table 2, HORIZON has lower mean error than WFR-FM and stVCR on all seven metrics and all four datasets, and lower molecular error than STORIES. Under adjacent teacher forcing, it is more accurate in 44 of 48 comparisons with the two NicheFlow variants, and the only exceptions are chicken's Chamfer and gene SW. Figure 2 shows the generated changes in organization and abundance, and mean mass RE stays below 0.51% in every series. Spatial gains are largest on axolotl and zebrafish and smallest on Drosophila, whose frames are nearly aligned. An ablation without alignment and projection, which converges on two series, raises spatial SW from 0.199 to 0.448 on chicken but only from 0.043 to 0.049 on Drosophila (Table 9).

Interpolation at unobserved times. We hold out each interior snapshot before coupling and training and score the omitted target. HORIZON completes all 11 folds, whereas WFR-FM and stVCR each fail one zebrafish fold due to unconstrained growth. On the folds each baseline completes, HORIZON has lower error in 48 of 60 comparisons with WFR-FM and 59 of 60 with stVCR. Its spatial SW and Chamfer are lower on every fold, and all twelve losses to WFR-FM involve molecular or joint metrics (Appendix G). A HORIZON fit, including qOET, takes 12–20 minutes on one GPU, 2.6–11.4 times faster than stVCR and 20–30 times faster than NicheFlow (Appendix D).

7 DISCUSSION AND CONCLUSION

Horizontal flow matching learns tissue dynamics consistently across coordinate frames by combining joint alignment with horizontal projection. Alignment provides frame-aware training paths, while the projector removes rigid motion from both regression and generation. The resulting fields jointly describe tissue organization, molecular state, and abundance, and provide readouts of biological programs and coupling affinities (Appendix I).

Limitations. The quotient formulation removes any collective rigid motion that is indistinguishable from a change of coordinate frame, so genuinely rigid tissue motion cannot be recovered from these observations alone. In addition, the current model is fitted separately to each time course, leaving generalization across specimens to future work. Finally, scalability remains a practical limitation, since both coupling construction and attention have quadratic complexity in the number of cells.

AI USE STATEMENT

The authors defined the research objective and core idea. We used generative AI tools to assist with the mathematical derivations, implementation refinement, numerical experiments, figure preparation, and manuscript editing. We reviewed all AI-assisted work. The authors checked the derivations and proofs step by step, verified AI-assisted code against independent numerical checks, checked every figure against the underlying results, and reviewed all edited text for correctness. The authors interpreted the results, verified the final claims, and take responsibility for the final content of this work, including text, claims, and artifacts produced with the aid of generative AI.

REPRODUCIBILITY STATEMENT

The anonymous repository at `https://anonymous.4open.science/r/horizon-47F6` contains the code, processed inputs, qOET teachers, and reference checkpoints. Complete proofs of all stated results, together with the assumptions under which they hold, are given in Appendix A. Appendix B relates these population results to the finite implementation and specifies the qOET solver, the network, training and generation (Algorithm 1), preprocessing, and the evaluation metrics, and Table 3 lists the settings for each dataset. Appendix C describes the frame-perturbation benchmark, including the seeds that generate every view, and Appendix D gives the baseline configurations, evaluation protocols, and runtimes. Appendices F and G report the sensitivity and timepoint holdout studies.

REFERENCES

Michael S. Albergo and Eric Vanden-Eijnden. Building normalizing flows with stochastic interpolants. In *International Conference on Learning Representations*, 2023.

Jean-David Benamou and Yann Brenier. A computational fluid mechanics solution to the Monge–Kantorovich mass transfer problem. *Numerische Mathematik*, 84:375–393, 2000.

Nicolas Bonneel, Julien Rabin, Gabriel Peyré, and Hanspeter Pfister. Sliced and Radon Wasserstein barycenters of measures. *Journal of Mathematical Imaging and Vision*, 51:22–45, 2015.

Charlotte Bunne, Yusuf Roohani, Yanay Rosen, et al. How to build the virtual cell with artificial intelligence: Priorities and opportunities. *Cell*, 187(25):7045–7063, 2024. doi: 10.1016/j.cell.2024.11.015.

Natalie Cao, Yifan Lu, and Xiaojie Qiu. Towards predictive virtual embryos with genomics and AI. *Nature Methods*, 23(9):1666–1670, 2026. doi: 10.1038/s41592-026-03055-4.

Ao Chen, Sha Liao, Mengnan Cheng, et al. Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball-patterned arrays. *Cell*, 185(10):1777–1792.e21, 2022.

Ricky T. Q. Chen, Yulia Rubanova, Jesse Bettencourt, and David Duvenaud. Neural ordinary differential equations. In *Advances in Neural Information Processing Systems*, volume 31, 2018.

Lénaïc Chizat, Gabriel Peyré, Bernhard Schmitzer, and François-Xavier Vialard. An interpolating distance between optimal transport and Fisher–Rao metrics. *Foundations of Computational Mathematics*, 18:1–44, 2018a.

Lénaïc Chizat, Gabriel Peyré, Bernhard Schmitzer, and François-Xavier Vialard. Scaling algorithms for unbalanced optimal transport problems. *Mathematics of Computation*, 87(314):2563–2609, 2018b. doi: 10.1090/mcom/3303.

Lénaïc Chizat, Gabriel Peyré, Bernhard Schmitzer, and François-Xavier Vialard. Unbalanced optimal transport: Dynamic and Kantorovich formulations. *Journal of Functional Analysis*, 274(11):3090–3123, 2018c.

Kalen Clifton, Manjari Anant, Gohta Aihara, Lyla Atta, Osagie K. Aimiuwu, Justus M. Kebschull, Michael I. Miller, Daniel Tward, and Jean Fan. STalign: Alignment of spatial transcriptomics data using diffeomorphic metric mapping. *Nature Communications*, 14(1):8123, 2023. doi: 10.1038/s41467-023-43915-7.

Marco Cuturi. Sinkhorn distances: Lightspeed computation of optimal transport. In *Advances in Neural Information Processing Systems*, volume 26, 2013.

David S. Fischer, Anna K. Fiedler, Eric M. Kernfeld, Ryan M. J. Genga, Aimée Bastidas-Ponce, Mostafa Bakhti, Heiko Lickert, Jan Hasenauer, Rene Maehr, and Fabian J. Theis. Inferring population dynamics from single-cell RNA-sequencing time series data. *Nature Biotechnology*, 37: 461–468, 2019.

John C. Gower. Generalized Procrustes analysis. *Psychometrika*, 40(1):33–51, 1975.

Arthur Gretton, Karsten M. Borgwardt, Malte J. Rasch, Bernhard Schölkopf, and Alexander Smola. A kernel two-sample test. *Journal of Machine Learning Research*, 13:723–773, 2012.

Emiel Hoogeboom, Víctor Garcia Satorras, Clément Vignac, and Max Welling. Equivariant diffusion for molecule generation in 3D. In *International Conference on Machine Learning*, pp. 8867–8887, 2022.

Guillaume Huguet, Daniel Sumner Magruder, Alexander Tong, Oluwadamilola Fasina, Manik Kuchroo, Guy Wolf, and Smita Krishnaswamy. Manifold interpolating optimal-transport flows for trajectory inference. In *Advances in Neural Information Processing Systems*, volume 35, pp. 29705–29718, 2022.

Geert-Jan Huizing, Jules Samaran, Daniele Capocefalo, Anna Audit, Gabriel Peyré, and Laura Cantini. STORIES: Learning cell fate landscapes from spatial transcriptomics using optimal transport. *Nature Methods*, 23:522–531, 2026. doi: 10.1038/s41592-025-02855-4.

Andrew Jones, F. William Townes, Didong Li, and Barbara E. Engelhardt. Alignment of spatial genomics data using deep Gaussian processes. *Nature Methods*, 20:1379–1387, 2023.

Kacper Kapuśniak, Peter Potaptchik, Teodora Reu, Leo Zhang, Alexander Tong, Michael Bronstein, Avishek Joey Bose, and Francesco Di Giovanni. Metric flow matching for smooth interpolations on the data manifold. In *Advances in Neural Information Processing Systems*, volume 37, pp. 135011–135042. Curran Associates, Inc., 2024. doi: 10.52202/079017-4291.

David G. Kendall. Shape manifolds, Procrustean metrics, and complex projective spaces. *Bulletin of the London Mathematical Society*, 16(2):81–121, 1984.

Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In *International Conference on Learning Representations*, 2015.

Dominik Klein, Giovanni Palla, Marius Lange, Michal Klein, Zoe Piran, Manuel Gander, Laetitia Meng-Papaxanthos, Michael Sterr, Lama Saber, Changying Jing, Aimée Bastidas-Ponce, Perla Cota, Marta Tarquis-Medina, Shrey Parikh, Ilan Gold, Heiko Lickert, Mostafa Bakhti, Mor Nitzan, Marco Cuturi, and Fabian J. Theis. Mapping cells through time and space with moscot. *Nature*, 638:1065–1075, 2025.

Leon Klein, Andreas Krämer, and Frank Noé. Equivariant flow matching. In *Advances in Neural Information Processing Systems*, volume 36, pp. 59886–59910, 2023.

Jonas Köhler, Leon Klein, and Frank Noé. Equivariant flows: Exact likelihood generative learning for symmetric densities. In *International Conference on Machine Learning*, pp. 5361–5370, 2020.

Stanislav Kondratyev, Léonard Monsaingeon, and Dmitry Vorotnikov. A new optimal transport distance on the space of finite Radon measures. *Advances in Differential Equations*, 21(11/12): 1117–1164, 2016.

Hugo Lavenant, Stephen Zhang, Young-Heon Kim, and Geoffrey Schiebinger. Toward a mathematical theory of trajectory inference. *Annals of Applied Probability*, 34(1A):428–500, 2024.

Juho Lee, Yoonho Lee, Jungtaek Kim, Adam R. Kosiorek, Seungjin Choi, and Yee Whye Teh. Set transformer: A framework for attention-based permutation-invariant neural networks. In *International Conference on Machine Learning*, pp. 3744–3753, 2019.

Matthias Liero, Alexander Mielke, and Giuseppe Savaré. Optimal transport in competition with reaction: The Hellinger–Kantorovich distance and geodesic curves. *SIAM Journal on Mathematical Analysis*, 48(4):2869–2911, 2016.

Matthias Liero, Alexander Mielke, and Giuseppe Savaré. Optimal entropy-transport problems and a new Hellinger–Kantorovich distance between positive measures. *Inventiones Mathematicae*, 211: 969–1117, 2018.

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

Chang Liu et al. Spatiotemporal mapping of gene expression landscapes and developmental trajectories during zebrafish embryogenesis. *Developmental Cell*, 57(10):1284–1298.e5, 2022. doi: 10.1016/j.devcel.2022.04.009.

Xingchao Liu, Chengyue Gong, and Qiang Liu. Flow straight and fast: Learning to generate and transfer data with rectified flow. In *International Conference on Learning Representations*, 2023a.

Xinhao Liu, Ron Zeira, and Benjamin J. Raphael. Partial alignment of multislice spatially resolved transcriptomics data. *Genome Research*, 33(7):1124–1132, 2023b.

Yinbo Liu, Keyang Ye, Wenshan Sun, Handi Gao, and Tian Tian. FlowCloud: Learning continuous spatiotemporal dynamics from unpaired sparse point cloud snapshots. In *International Conference on Machine Learning*, 2026.

Madhav Mantri, Gaetano J. Scuderi, Roozbeh Abedini-Nassab, Michael F. Z. Wang, David McKellar, Hao Shi, Benjamin Grodner, Jonathan T. Butcher, and Iwijn De Vlaminck. Spatiotemporal single-cell RNA sequencing of developing chicken hearts identifies interplay between cellular differentiation and morphogenesis. *Nature Communications*, 12:1771, 2021. doi: 10.1038/s41467-021-21892-z.

Facundo Mémoli. Gromov–Wasserstein distances and the metric approach to object matching. *Foundations of Computational Mathematics*, 11:417–487, 2011.

Kevine Meugang Toukam, Max von Renesse, and Johannes Storn. Dynamical optimal transport with $\mathfrak{so}(d)$-invariance: From theory to computation. *arXiv preprint arXiv:2607.16782*, 2026. URL https://arxiv.org/abs/2607.16782.

Lambda Moses and Lior Pachter. Museum of spatial transcriptomics. *Nature Methods*, 19:534–546, 2022.

Kirill Neklyudov, Rob Brekelmans, Daniel Severo, and Alireza Makhzani. Action matching: Learning stochastic dynamics from samples. In *International Conference on Machine Learning*, pp. 25858–25889, 2023.

Kirill Neklyudov, Rob Brekelmans, Alexander Tong, Lazar Atanackovic, Qiang Liu, and Alireza Makhzani. A computational framework for solving Wasserstein Lagrangian flows. In *International Conference on Machine Learning*, pp. 37461–37485, 2024.

Matteo Pariset, Ya-Ping Hsieh, Charlotte Bunne, Andreas Krause, and Valentin De Bortoli. Unbalanced diffusion Schrödinger bridge. *arXiv preprint arXiv:2306.09099*, 2023. doi: 10.48550/arXiv.2306.09099.

Qiangwei Peng, Zihan Wang, Junda Ying, Yuhao Sun, Qing Nie, Lei Zhang, Tiejun Li, and Peijie Zhou. WFR-FM: Simulation-free dynamic unbalanced optimal transport. In *International Conference on Learning Representations*, 2026a. URL https://arxiv.org/abs/2601.06810.

Qiangwei Peng, Peijie Zhou, and Tiejun Li. stVCR: Spatiotemporal dynamics of single cells. *Nature Methods*, 23:542–553, 2026b. doi: 10.1038/s41592-026-03010-3.

Gabriel Peyré and Marco Cuturi. Computational optimal transport: With applications to data science. *Foundations and Trends in Machine Learning*, 11(5–6):355–607, 2019. doi: 10.1561/2200000073.

Xiaojie Qiu et al. Spatiotemporal modeling of molecular holograms. *Cell*, 187(26):7351–7373.e61, 2024. doi: 10.1016/j.cell.2024.10.011.

Kristiyan Sakalyan, Alessandro Palma, Filippo Guerranti, Fabian J. Theis, and Stephan Günnemann. Modeling microenvironment trajectories on spatial transcriptomics with NicheFlow. In *Advances in Neural Information Processing Systems*, volume 38, pp. 114430–114473, 2025. doi: 10.52202/085713-3453.

Víctor Garcia Satorras, Emiel Hoogeboom, and Max Welling. E(n) equivariant graph neural networks. In *International Conference on Machine Learning*, pp. 9323–9332, 2021.

Geoffrey Schiebinger, Jian Shu, Marcin Tabaka, Brian Cleary, Vidya Subramanian, Aryeh Solomon, Joshua Gould, Siyan Liu, Stacie Lin, Peter Berube, et al. Optimal-transport analysis of single-cell gene expression identifies developmental trajectories in reprogramming. *Cell*, 176(4):928–943.e22, 2019. doi: 10.1016/j.cell.2019.01.006.

Yutong Sha, Yuchi Qiu, Peijie Zhou, and Qing Nie. Reconstructing growth and dynamic trajectories from single-cell transcriptomics data. *Nature Machine Intelligence*, 6:25–39, 2024. doi: 10.1038/s42256-023-00763-w.

Patrik L. Ståhl, Fredrik Salmén, Sanja Vickovic, et al. Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science*, 353(6294):78–82, 2016.

Alexander Tong, Jessie Huang, Guy Wolf, David Van Dijk, and Smita Krishnaswamy. TrajectoryNet: A dynamic optimal transport network for modeling cellular dynamics. In *International Conference on Machine Learning*, pp. 9526–9536, 2020.

Alexander Tong, Kilian Fatras, Nikolay Malkin, Guillaume Huguet, Yanlei Zhang, Jarrid Rector-Brooks, Guy Wolf, and Yoshua Bengio. Improving and generalizing flow-based generative models with minibatch optimal transport. *Transactions on Machine Learning Research*, 2024a.

Alexander Tong, Nikolay Malkin, Kilian Fatras, Lazar Atanackovic, Yanlei Zhang, Guillaume Huguet, Guy Wolf, and Yoshua Bengio. Simulation-free Schrödinger bridges via score and flow matching. In *International Conference on Artificial Intelligence and Statistics*, volume 238 of *Proceedings of Machine Learning Research*, pp. 1279–1287, 2024b.

Shinji Umeyama. Least-squares estimation of transformation parameters between two point patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(4):376–380, 1991.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. In *Advances in Neural Information Processing Systems*, volume 30, 2017.

Titouan Vayer, Laetitia Chapel, Rémi Flamary, Romain Tavenard, and Nicolas Courty. Fused Gromov–Wasserstein distance for structured objects. *Algorithms*, 13(9):212, 2020.

Cédric Villani. *Optimal Transport: Old and New*. Springer, 2009.

Dongyi Wang, Yuanwei Jiang, Zhenyi Zhang, Xiang Gu, Peijie Zhou, and Jian Sun. Joint velocity-growth flow matching for single-cell dynamics modeling. In *Advances in Neural Information Processing Systems*, volume 38, pp. 178180–178216, 2025. doi: 10.52202/085713-5363.

Sinan Wang, Jinjin He, Shenyifan Lu, Ruicheng Wang, Greg Turk, and Bo Zhu. Generative modeling with orbit-space particle flow matching. *ACM Transactions on Graphics*, 45(4):1–27, 2026. doi: 10.1145/3811342.

Xiaoyu Wei et al. Single-cell Stereo-seq reveals induced progenitor cells involved in axolotl brain regeneration. *Science*, 377(6610):eabp9444, 2022. doi: 10.1126/science.abp9444.

Caleb Weinreb, Samuel Wolock, Betsabeh K. Tusi, Merav Socolovsky, and Allon M. Klein. Fundamental limits on dynamic inference from single-cell snapshots. *Proceedings of the National Academy of Sciences*, 115(10):E2467–E2476, 2018.

Yixian Xu, Yusong Wang, Shengjie Luo, Kaiyuan Gao, Tianyu He, Di He, and Chang Liu. Quotient-space diffusion models. In *International Conference on Learning Representations*, 2026. URL https://arxiv.org/abs/2604.21809.

Ron Zeira, Max Land, Alexander Strzalkowski, and Benjamin J. Raphael. Alignment and integration of spatial transcriptomics data. *Nature Methods*, 19(5):567–575, 2022. doi: 10.1038/s41592-022-01459-6.

Zhenyi Zhang, Tiejun Li, and Peijie Zhou. Learning stochastic dynamics from snapshots through regularized unbalanced optimal transport. In *International Conference on Learning Representations*, 2025.

A PROOFS AND SUPPLEMENTARY RESULTS

We use the notation of Sections 3–4. Continuity equations are understood in the weak sense, and admissible frame coefficients are integrable in time. Whenever a horizontal projection is used, the measure has positive mass and finite spatial second moment, and the projected field belongs to $L^2(\rho)$. These conditions make the weighted moments and the rigid-frame flow well defined. The distinction between the population results and their numerical implementation is given in Appendix B.1.

A.1 PROOF OF PROPOSITION 1

Let ρ have mass $M > 0$, and let $\bar{x}, \bar{v}, S, C$ be the weighted moments defined in Section 3.3. The least-squares rigid component of v is $a^\star + \Omega^\star x$, where

$$a^\star = \bar{v} - \Omega^\star \bar{x}, \qquad \Omega^\star S + S\Omega^\star = C - C^\top.$$

The minimum-norm skew solution defines $P_\rho^H v = v - \bar{v} - \Omega^\star(x - \bar{x})$, the orthogonal projection onto the complement of the rigid fields in $L^2(\rho)$.

Proof. Consider the convex quadratic objective

$$F(a, \Omega) = \int \|v - a - \Omega x\|^2 \, \mathrm{d}\rho, \qquad \Omega^\top = -\Omega.$$

For any translation perturbation b, stationarity gives

$$0 = \frac{\mathrm{d}}{\mathrm{d}\epsilon} F(a + \epsilon b, \Omega)\Big|_{\epsilon=0} = -2b^\top \int (v - a - \Omega x)\mathrm{d}\rho.$$

Hence $a = \bar{v} - \Omega\bar{x}$. Writing $y = x - \bar{x}$ and $w = v - \bar{v}$ reduces the objective to $F(\Omega) = \int \|w - \Omega y\|^2 \, \mathrm{d}\rho$. For any skew matrix K,

$$\begin{aligned} 0 &= -\frac{1}{2} \frac{\mathrm{d}}{\mathrm{d}\epsilon} F(\Omega + \epsilon K)\Big|_{\epsilon=0} \\ &= \int (w - \Omega y)^\top K y \mathrm{d}\rho = \mathrm{tr}\left((C^\top + S\Omega)K\right). \end{aligned}$$

Orthogonality to all skew matrices means that $C^\top + S\Omega$ is symmetric. Using $S^\top = S$ and $\Omega^\top = -\Omega$ gives

$$C^\top + S\Omega = C - \Omega S, \qquad \Omega S + S\Omega = C - C^\top.$$

These normal equations are sufficient because F is convex. Their minimum-norm solution is given by the pseudoinverse of $\mathcal{L}_S : \Omega \mapsto \Omega S + S\Omega$ on $\mathfrak{so}(d_x)$. If S is positive definite, this operator is invertible: in an eigenbasis of S, its action on the (i, j) entry is multiplication by $s_i + s_j > 0$. In the singular case, any two minimizing rigid fields agree ρ-almost everywhere, so the projected field remains unique.

For the residual $r = v - a^\star - \Omega^\star x$, the two stationarity conditions imply

$$\int r \mathrm{d}\rho = 0, \qquad \int r^\top K x \mathrm{d}\rho = 0 \quad (K^\top = -K).$$

It follows that $\langle r, b + Kx \rangle_{L^2(\rho)} = 0$ for every rigid field, which proves the claimed orthogonal projection.

In two dimensions, write $\Omega^\star = \omega^\star J$ with J the quarter turn. Differentiating $F(\omega J)$ gives

$$\omega^\star = \frac{\int (Jy)^\top w \mathrm{d}\rho}{\int \|Jy\|^2 \, \mathrm{d}\rho} = \frac{\int (x - \bar{x})^\top J^\top (v - \bar{v}) \mathrm{d}\rho}{\int \|x - \bar{x}\|^2 \, \mathrm{d}\rho},$$

when the denominator is positive; the minimum-norm choice is zero when it vanishes. A measure is regular when no nonzero rigid field vanishes ρ-almost everywhere. In two dimensions, this is equivalent to positive mass and a positive denominator in the preceding expression. □

A.2 Proof of Theorem 2

For nonzero finite atomic measures on $\mathcal{Z}$ and $G = \mathrm{SE}(d_x)$, the infimum in Equation (4) is attained, $\overline{\mathrm{WFR}}_\delta$ is a metric on the orbit space, and

$$\overline{\mathcal{A}}_\delta(\mu_0, \mu_1) = \min_{h \in G} \mathrm{WFR}_\delta^2(\mu_0, h_\# \mu_1) = \overline{\mathrm{WFR}}_\delta^2([\mu_0], [\mu_1]).$$

Proof. Write $d = \mathrm{WFR}_\delta$. Since G acts only on the spatial coordinates and preserves d_Λ, it acts by isometries of d.

Existence of an optimal pose. Let $h_n = (R_n, b_n)$ be a minimizing sequence. The rotations lie in the compact group $SO(d_x)$. If the translations are bounded, a subsequence converges to $h^\star$, and continuity of WFR on finite atomic measures gives

$$d^2(\mu, {h^\star}_\# \nu) = \inf_{h \in G} d^2(\mu, h_\# \nu).$$

If the translations are unbounded, pass to a subsequence with $\|b_n\| \to \infty$. Both supports are finite, so all cross-distances eventually exceed the transport cut locus $\pi\delta$. The optimal coupling then vanishes and

$$d^2(\mu, h_{n\#}\nu) = 2\delta^2(|\mu| + |\nu|).$$

If the infimum is smaller than this value, such a subsequence cannot be minimizing. If the infimum equals this value, any fixed translation that separates the supports beyond the cut locus attains it. Thus an optimizer exists in both cases.

Metricity. For representatives changed by $g_0, g_1 \in G$, isometry gives

$$d(g_{0\#}\mu, h_\# g_{1\#}\nu) = d(\mu, (g_0^{-1} h g_1)_\# \nu).$$

The map $h \mapsto g_0^{-1} h g_1$ is a bijection of G, so the infimum depends only on the orbits. Nonnegativity follows from that of d, and symmetry follows by replacing h with h^{-1}. For ϵ-optimal poses h_1, h_2 for (μ, ν) and (ν, η),

$$\begin{aligned}
\overline{\mathrm{WFR}}_\delta([\mu], [\eta]) &\le d(\mu, (h_1 h_2)_\# \eta) \\
&\le d(\mu, h_{1\#}\nu) + d(h_{1\#}\nu, (h_1 h_2)_\# \eta) \\
&= d(\mu, h_{1\#}\nu) + d(\nu, h_{2\#}\eta) \\
&\le \overline{\mathrm{WFR}}_\delta([\mu], [\nu]) + \overline{\mathrm{WFR}}_\delta([\nu], [\eta]) + 2\epsilon.
\end{aligned}$$

Letting $\epsilon \downarrow 0$ proves the triangle inequality. Finally, if the quotient distance is zero, attainment and metricity of d imply $\mu = {h^\star}_\# \nu$, hence $[\mu] = [\nu]$.

Lower bound for the dynamic action. Take an admissible tuple (ρ, u, g, a, Ω) in Equation (2). Let $G_t(x) = R_t x + b_t$ solve

$$\partial_t G_t(x) = a_t + \Omega_t G_t(x), \qquad G_0 = I,$$

and put $\widetilde{\rho}_t = (G_t^{-1})_\# \rho_t$. For $f \in C_c^1(\mathcal{Z})$, define $\psi_t(x, q) = f(G_t^{-1} x, q)$. The chain rule gives

$$\partial_t \psi_t + \nabla_x \psi_t \cdot (a_t + \Omega_t x) = 0.$$

Testing the weak continuity equation against ψ_t therefore yields

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} \int f \mathrm{d}\widetilde{\rho}_t &= \int \big[\partial_t \psi_t + \nabla_x \psi_t \cdot (u_x + a_t + \Omega_t x) + \nabla_q \psi_t \cdot u_q + g\psi_t\big] \mathrm{d}\rho_t \\
&= \int \big[\nabla_x f \cdot \widetilde{u}_x + \nabla_q f \cdot \widetilde{u}_q + \widetilde{g} f\big] \mathrm{d}\widetilde{\rho}_t,
\end{aligned}$$

where

$$\widetilde{u}_x(x, q, t) = R_t^\top u_x(G_t x, q, t), \quad \widetilde{u}_q(x, q, t) = u_q(G_t x, q, t), \quad \widetilde{g}(x, q, t) = g(G_t x, q, t).$$

Thus the pulled-back fields satisfy the ordinary continuity–reaction equation. Since R_t is orthogonal,

$$\begin{aligned}
&\int \big(\lambda_x \|\widetilde{u}_x\|^2 + \lambda_q \|\widetilde{u}_q\|^2 + \delta^2 \widetilde{g}^2\big) \mathrm{d}\widetilde{\rho}_t \\
&\qquad = \int \big(\lambda_x \|u_x\|^2 + \lambda_q \|u_q\|^2 + \delta^2 g^2\big) \mathrm{d}\rho_t.
\end{aligned}$$

The pulled-back path has endpoints μ_0 and $(G_1^{-1})_{\#}\mu_1$. Its action is consequently at least $d^2(\mu_0, (G_1^{-1})_{\#}\mu_1)$, and taking the infimum over admissible tuples gives

$$\overline{\mathcal{A}}_\delta(\mu_0, \mu_1) \geq \min_{h \in G} d^2(\mu_0, h_{\#}\mu_1).$$

Upper bound. Fix $h \in G$ and an ambient WFR geodesic $\widetilde{\rho}_t$ joining μ_0 to $h_{\#}\mu_1$. Since $\mathrm{SE}(d_x)$ is connected, choose a smooth rigid path G_t from I to h^{-1} and set $\rho_t = G_{t\#}\widetilde{\rho}_t$. Define the relative fields by

$$u_x(G_t x, q, t) = R_t \widetilde{u}_x(x, q, t), \quad u_q(G_t x, q, t) = \widetilde{u}_q(x, q, t), \quad g(G_t x, q, t) = \widetilde{g}(x, q, t),$$

and place $\dot{G}_t \circ G_t^{-1} = a_t + \Omega_t x$ in the frame term. The preceding change of variables, read in reverse, proves feasibility. Moreover, $\rho_0 = \mu_0$, $\rho_1 = \mu_1$, and the action equals $d^2(\mu_0, h_{\#}\mu_1)$. Minimizing over h gives the reverse inequality. □

A.3 Frame covariance

Proposition A.1 (Covariance under changes of representatives)**.** *For $g_0, g_1 \in G$, changing the endpoint representatives to $g_{0\#}\mu_0, g_{1\#}\mu_1$ sends every optimal relative pose h to $g_0 h g_1^{-1}$ and preserves its OET coupling. In particular,*

$$\operatorname*{Argmin}_h \mathrm{WFR}_\delta(g_{0\#}\mu_0, h_{\#}g_{1\#}\mu_1) = g_0 \operatorname*{Argmin}_h \mathrm{WFR}_\delta(\mu_0, h_{\#}\mu_1) g_1^{-1}. \tag{11}$$

For a spatial field, with $(g_* v)(gx, q) = Rv(x, q)$ when $g = (R, b)$,

$$P^H_{g_{\#}\rho}(g_* v) = g_*(P^H_\rho v).$$

Proof. Ground-metric invariance gives, for every pair of endpoint states,

$$d_\Lambda(g_0 z_i, (g_0 h g_1^{-1}) g_1 z'_j) = d_\Lambda(z_i, h z'_j).$$

All transport costs are unchanged, as are the weights and KL penalties. The map $h \mapsto g_0 h g_1^{-1}$ is bijective, which proves the assertion about poses and couplings in both directions.

For the projector, the transformed moments are

$$\bar{x}' = R\bar{x} + b, \qquad \bar{v}' = R\bar{v}, \qquad S' = RSR^\top, \qquad C' = RCR^\top.$$

The minimum-norm skew solution consequently transforms as $\Omega' = R\Omega R^\top$: it solves the transformed Sylvester equation, and conjugation preserves the Frobenius norm. Substitution gives

$$\begin{aligned}(P^H_{g_{\#}\rho} g_* v)(gx, q) &= Rv(x, q) - R\bar{v} - (R\Omega R^\top)\big(Rx + b - (R\bar{x} + b)\big) \\ &= R\big(v(x, q) - \bar{v} - \Omega(x - \bar{x})\big).\end{aligned}$$

This is the claimed covariance. The pose identity is set-valued and does not require a unique optimizer. For symmetric tissues, different minimizers may give different representative-level paths, each optimal in the quotient. Selecting a unique equivariant lift would additionally require an equivariant choice among those minimizers. □

A.4 Analytic WFR bridges

Proposition A.2 (Bridge construction and marginal dynamics)**.** *Fix a pose h and an optimal unregularized OET coupling γ between $\mu = \sum_i a_i \delta_{z_i}$ and $h_{\#}\nu$. For every positive edge, assign the two masses in Equation* (6) *and use the bridge in Equations* (7)–(8)*. Add the pure decay and creation components for zero rows and columns. Their sum has endpoints $\mu, h_{\#}\nu$, and the mass-weighted aggregate velocity and growth satisfy the continuity–reaction equation. The resulting path is an ambient WFR geodesic.*

Proof. Endpoint masses. Write $r_i = \sum_j \gamma_{ij}$ and $s_j = \sum_i \gamma_{ij}$. The OET coupling has soft marginal constraints, so r and s need not equal a and b. Instead, the two masses assigned to each positive edge satisfy

$$\sum_j m_{0,ij} = \sum_j a_i \frac{\gamma_{ij}}{r_i} = a_i \quad (r_i > 0), \qquad \sum_i m_{1,ij} = \sum_i b_j \frac{\gamma_{ij}}{s_j} = b_j \quad (s_j > 0).$$

A zero row contributes a decay component of initial mass a_i, and a zero column contributes a creation component of terminal mass b_j. Consequently, the full collection of components has the prescribed endpoint measures. It is this pair of mass allocations, or semi-coupling, that restores the endpoints (Liero et al., 2018; Peng et al., 2026a).

A single transported component. Let z_0, z_1 be its endpoints and let $\vartheta = d_\Lambda(z_0, z_1)/(2\delta) < \pi/2$. In the WFR cone, a weighted state has radial coordinate $\sqrt{m}$, and the squared cone distance between these two states is

$$d^2_{\mathrm{cone}} = m_0 + m_1 - 2\sqrt{m_0 m_1}\cos\vartheta.$$

The WFR action of the corresponding single-component geodesic is $2\delta^2 d^2_{\mathrm{cone}}$ under the convention of Equation (1). The straight segment in the cone plane is represented by

$$A_t = (1-t)\sqrt{m_0} + t\sqrt{m_1}e^{i\vartheta}, \qquad m_t = |A_t|^2, \qquad \phi_t = \arg A_t.$$

Expanding $|A_t|^2$ gives Equation (7). Moreover,

$$\dot\phi_t = \frac{\mathrm{Im}(\overline{A_t}\dot A_t)}{|A_t|^2} = \frac{\sqrt{m_0 m_1}\sin\vartheta}{m_t},$$
$$\dot z_t = \frac{\dot\phi_t}{\vartheta}(z_1 - z_0) = \frac{\sqrt{m_0 m_1}\sin\vartheta}{\vartheta m_t}(z_1 - z_0).$$

Here $z_t = (1 - \phi_t/\vartheta)z_0 + (\phi_t/\vartheta)z_1$. Differentiating the mass gives

$$\dot m_t = -2(1-t)m_0 + 2tm_1 + 2(1-2t)\sqrt{m_0 m_1}\cos\vartheta, \qquad g^c_t = \dot m_t/m_t. \tag{12}$$

At $\vartheta = 0$, the continuous limit is

$$\alpha_t = \frac{t\sqrt{m_1}}{(1-t)\sqrt{m_0} + t\sqrt{m_1}}, \qquad \frac{\sin\vartheta}{\vartheta} = 1.$$

Pure decay and creation have masses $(1-t)^2 m_0$ and $t^2 m_1$ and growth rates $-2/(1-t)$ and $2/t$, respectively. The rates are used only at positive-mass times. Their weighted squared growth is $4m_0$ and $4m_1$, so the singular zero-mass endpoint contributes no divergence to the action.

The marginal continuity equation. Index all components by ℓ and put $\rho_t = \sum_\ell m_\ell(t)\delta_{z_\ell(t)}$. Define $\bar u_t(z)$ and $\bar g_t(z)$ by averaging u^c_ℓ and g^c_ℓ over components at z, with weights proportional to $m_\ell(t)$. For any $f \in C^1_c(\mathcal{Z})$,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\int f\,\mathrm{d}\rho_t &= \sum_\ell m_\ell(t)\big[\nabla f(z_\ell(t))\cdot u^c_\ell(t) + g^c_\ell(t) f(z_\ell(t))\big] \\ &= \int \big[\nabla f\cdot\bar u_t + \bar g_t f\big]\mathrm{d}\rho_t.\end{aligned}$$

This is the weak continuity–reaction equation, including when several components occupy the same state.

Optimality. For a positive edge, stationarity of the unregularized OET objective gives

$$c_{ij} + \log(r_i/a_i) + \log(s_j/b_j) = 0, \qquad \sqrt{m_{0,ij}m_{1,ij}}\cos\vartheta_{ij} = \gamma_{ij}.$$

Writing $J(\gamma)$ for the bracketed objective in Equation (5), it follows that

$$\begin{aligned}J(\gamma) &= \sum_{ij}\gamma_{ij}\big[c_{ij} + \log(r_i/a_i) + \log(s_j/b_j)\big] + |\mu| + |\nu| - 2\sum_{ij}\gamma_{ij} \\ &= |\mu| + |\nu| - 2\sum_{ij}\gamma_{ij}.\end{aligned}$$

The sums involving logarithms run over positive edges. Summing the cone actions of all transported, decay, and creation components gives exactly $2\delta^2 J(\gamma) = \mathrm{WFR}^2_\delta(\mu, h_\#\nu)$ by the OET representation (Liero et al., 2018). Conditional Jensen inequalities bound the action of $(\rho, \bar u, \bar g)$ above by this sum. Its endpoint constraints bound the same action below by $\mathrm{WFR}^2_\delta(\mu, h_\#\nu)$. Both bounds are therefore equal, and the marginal path minimizes the WFR action on $[0, 1]$. □

A.5 EMPIRICAL HORIZONTAL PROJECTION

Proposition A.3 (Consistency of empirical horizontal projection)**.** *Let (ρ_B, e_B) be weighted empirical populations whose mass, spatial and residual first moments, matrices S_B, C_B, and residual second moment converge to those of (ρ, e). Suppose $\mathcal{L}_S$ is invertible on $\mathfrak{so}(d_x)$. Then the fitted coefficients (a_B, Ω_B) and the minimum weighted losses converge to their population counterparts. In particular, this holds almost surely for i.i.d. mass-weighted samples of a fixed residual field with finite spatial and residual second moments.*

Proof. The empirical normal equations are

$$a_B = \bar{e}_B - \Omega_B \bar{x}_B, \qquad \mathcal{L}_{S_B}(\Omega_B) = C_B - C_B^\top.$$

Moment convergence implies convergence of $\mathcal{L}_{S_B}$ in operator norm. Since $\mathcal{L}_S$ is invertible on the finite-dimensional skew subspace, $\mathcal{L}_{S_B}$ is invertible for all sufficiently large B, and continuity of inversion yields

$$\Omega_B = \mathcal{L}_{S_B}^{-1}(C_B - C_B^\top) \longrightarrow \mathcal{L}_S^{-1}(C - C^\top) = \Omega^\star, \qquad a_B \longrightarrow a^\star.$$

If $Q_B = \int \|e_B\|^2 \, \mathrm{d}\rho_B$, expanding the quadratic objective at its minimizer gives

$$\min F_B = Q_B - M_B \|\bar{e}_B\|^2 - \mathrm{tr}(\Omega_B^\top C_B).$$

Every term converges by the assumed moment convergence and the coefficient convergence just proved. For i.i.d. samples from ρ/M with weight M/B, the strong law supplies these moment limits; integrability of the cross-moments follows from Cauchy–Schwarz. The same conclusion applies to stratified or importance-weighted sampling whenever its weighted moments satisfy the stated convergence assumptions. □

A.6 PROOF OF LEMMA 3

Let $h^\star$ attain Equation (4)*, and let $\widetilde{\rho}_t$ be the bridge marginal of an exact qOET solution from μ_0 to $h^\star{}_\#\mu_1$, with spatial aggregate velocity $\widetilde{u}_x$. Then*

$$\widetilde{u}_x = P^H_{\widetilde{\rho}_t} \widetilde{u}_x \quad \text{in } L^2(\widetilde{\rho}_t) \text{ for almost every } t.$$

Proof. By Proposition A.2, the bridge marginal is an ambient WFR geodesic. Its action is $D^2 = \overline{\mathrm{WFR}}_\delta^2([\mu_0], [\mu_1])$, since the endpoint pose is globally optimal. Decompose

$$\widetilde{u}_x = P^H_{\widetilde{\rho}_t} \widetilde{u}_x + \xi_t, \qquad \xi_t(x) = a_t + \Omega_t x.$$

On any measurable set of times where the fitted coefficients are bounded, place ξ_t in the free frame term and subtract it from the relative spatial velocity. The total velocity and hence the path are unchanged. Orthogonality of the projection reduces the action by precisely the squared norm of this rigid component on that set.

More explicitly, choose the minimum-norm coefficients and let $E_n = \{t \in [1/n, 1 - 1/n] : \|a_t\| + \|\Omega_t\|_{\mathrm{F}} \le n\}$. The frame $\mathbf{1}_{E_n}\xi_t$ is integrable, and the tuple with relative velocity $\widetilde{u}_x - \mathbf{1}_{E_n}\xi_t$ is feasible between μ_0 and $h^\star{}_\#\mu_1$. Theorem 2 and the Pythagorean identity give

$$D^2 \le D^2 - \frac{\lambda_x}{2} \int_{E_n} \int \|\xi_t\|^2 \, \mathrm{d}\widetilde{\rho}_t \mathrm{d}t \le D^2.$$

The integral is zero for every n. The sets E_n increase to $(0, 1)$ up to a null set, so $\xi_t = 0$ in $L^2(\widetilde{\rho}_t)$ for almost every t. Thus the original bridge marginal is already horizontal. □

A.7 PROOF OF THEOREM 4

Fix t and normalize the mass-weighted bridge measure to a probability law. Let $U^c \in L^2$ be the spatial bridge velocity, $\bar{U}(z) = \mathbb{E}[U^c \mid Z_t = z]$, and assume ρ_t is regular. For every $v \in L^2(\rho_t)$,

$$\inf_{\xi \in \mathfrak{se}(d_x)} \mathbb{E} \|v(Z_t) - U^c - \xi(Z_t)\|^2 = \inf_{\xi \in \mathfrak{se}(d_x)} \mathbb{E} \left\|v(Z_t) - \bar{U}(Z_t) - \xi(Z_t)\right\|^2 + C_t,$$

$$C_t = \mathbb{E} \left\|U^c - \bar{U}(Z_t)\right\|^2.$$

Consequently, every population minimizer satisfies $P^H_{\rho_t} v = P^H_{\rho_t} \bar{U}$.

Proof. Set $\eta = U^c - \bar{U}(Z_t)$. By conditional expectation, $\mathbb{E}[\eta \mid Z_t] = 0$. For a fixed rigid field ξ, let $r(Z_t) = v(Z_t) - \bar{U}(Z_t) - \xi(Z_t)$. Since r is a function of Z_t and all fields have finite second moments,

$$\mathbb{E}\langle r(Z_t), \eta\rangle = \mathbb{E}\big\langle r(Z_t), \mathbb{E}[\eta \mid Z_t]\big\rangle = 0.$$

Expanding the square now gives

$$\begin{aligned}\mathbb{E}\left\|v(Z_t) - U^c - \xi(Z_t)\right\|^2 &= \mathbb{E}\left\|r(Z_t) - \eta\right\|^2 \\ &= \mathbb{E}\left\|r(Z_t)\right\|^2 + \mathbb{E}\left\|\eta\right\|^2.\end{aligned}$$

The second term is independent of v and ξ, so taking the infimum over the rigid subspace proves the identity. With $M_t = |\rho_t|$, Proposition 1 identifies the remaining term as

$$\inf_{\xi} \mathbb{E}\left\|v(Z_t) - \bar{U}(Z_t) - \xi(Z_t)\right\|^2 = M_t^{-1}\left\|P^H_{\rho_t}(v - \bar{U})\right\|^2_{L^2(\rho_t)}.$$

It is nonnegative and vanishes for $v = \bar{U}$. Every minimizer must therefore satisfy the claimed projected equality. Applying the same conditional-expectation calculation to the molecular and growth targets, without the infimum over rigid fields, identifies their minimizers with their respective conditional means. □

A.8 Population recovery

Corollary A.4 (Population recovery of quotient geodesics)**.** *Assume a globally optimal unregularized qOET solution and regular tissues along its bridge marginal $\rho_t^\star$. Let $\mathbb{E}_t$ denote expectation under the normalized mass-weighted bridge law and let $M_t = |\rho_t^\star|$. Suppose the population version of Equation* (9)*,*

$$\begin{aligned}\mathcal{L}_{\mathrm{pop}} = \int_0^1 M_t\Big[\lambda_x \inf_{\xi \in \mathfrak{se}(d_x)} \mathbb{E}_t\left\|v_x(Z_t,t) - U_x^c - \xi(Z_t)\right\|^2 \\ + \lambda_q \mathbb{E}_t\left\|v_q(Z_t,t) - U_q^c\right\|^2 + \delta^2 \mathbb{E}_t|g(Z_t,t) - G^c|^2\Big]\mathrm{d}t,\end{aligned} \tag{13}$$

is minimized exactly over all square-integrable fields. For fields that depend on the population, this loss evaluates them at $\rho_t^\star$. Every minimizer induces the same projected spatial velocity, molecular velocity, and growth $\rho_t^\star$-almost everywhere for almost every t, and $\rho_t^\star$ is a weak solution of

$$\partial_t \rho_t + \nabla_x \cdot \big(\rho_t P^H_{\rho_t} v_{\theta,x}\big) + \nabla_q \cdot \left(\rho_t v_{\theta,q}\right) = g_\phi \rho_t. \tag{14}$$

Writing $D = \overline{\mathrm{WFR}}_\delta([\mu_0],[\mu_1])$, this path has action D^2 and

$$\overline{\mathrm{WFR}}_\delta([\rho_s^\star],[\rho_t^\star]) = (t-s)D \qquad (0 \le s < t \le 1),$$

with $\rho_1^\star = h^\star{}_\#\mu_1 \in [\mu_1]$. If Equation (14) *has a unique solution with initial measure μ_0 in the class under consideration, exact population integration recovers $\rho_t^\star$.*

Proof. Let $(\bar{u}_x, \bar{u}_q, \bar{g})$ be the mass-weighted aggregate bridge fields. Theorem 4 and its molecular and growth counterparts write the population loss as a constant plus

$$\int_0^1 \int \left[\lambda_x \left\|P^H_{\rho_t^\star}(v_x - \bar{u}_x)\right\|^2 + \lambda_q \left\|v_q - \bar{u}_q\right\|^2 + \delta^2 |g - \bar{g}|^2\right]\mathrm{d}\rho_t^\star \mathrm{d}t.$$

The aggregate fields attain zero excess loss. Since the weights $\lambda_x, \lambda_q, \delta^2$ are positive, every exact minimizer must make each squared residual zero almost everywhere. Lemma 3 then gives

$$P^H_{\rho_t^\star} v_x = \bar{u}_x, \qquad v_q = \bar{u}_q, \qquad g = \bar{g} \quad \rho_t^\star\text{-almost everywhere, for almost every } t.$$

Proposition A.2 supplies the weak continuity equation for the aggregate fields. Substituting these equalities proves Equation (14) on the same path.

The ambient bridge path is a constant-speed WFR geodesic of length D. For $0 \le s < t \le 1$, the quotient metric is bounded by the ambient one, so

$$\begin{aligned}D &\le \overline{\mathrm{WFR}}_\delta([\rho_0^\star],[\rho_s^\star]) + \overline{\mathrm{WFR}}_\delta([\rho_s^\star],[\rho_t^\star]) + \overline{\mathrm{WFR}}_\delta([\rho_t^\star],[\rho_1^\star]) \\ &\le sD + (t-s)D + (1-t)D = D.\end{aligned}$$

Equality of the sums forces equality in each term. Its action is D^2 by the bridge construction and horizontality. To obtain the prescribed terminal representative μ_1, choose a smooth rigid path H_t from I to $(h^\star)^{-1}$ and set $\widehat{\rho}_t = H_{t\#}\rho_t^\star$. The construction in Theorem 2 puts the frame velocity into the free term, giving a feasible tuple for Equation (2) with the same action. Finally, $\rho_t^\star$ is a solution with initial measure μ_0; uniqueness, when assumed, identifies it with the exactly integrated population solution. □

A.9 Multiple snapshots

Proposition A.5 (Multi-snapshot quotient action)**.** *For $t_0 < \cdots < t_K$, let $\overline{\mathcal{E}}$ be the action of Equation* (2) *integrated over physical time. For nonzero finite atomic orbit constraints $\rho_{t_k} \in [\mu_k]$,*

$$\inf_{\rho_{t_k} \in [\mu_k]} \overline{\mathcal{E}}(\rho) = \sum_{k=0}^{K-1} \frac{\overline{\mathrm{WFR}}_\delta^2([\mu_k], [\mu_{k+1}])}{t_{k+1} - t_k}. \tag{15}$$

Each summand is attained by a time-rescaled quotient geodesic. Consecutive segments admit representatives meeting at their shared snapshot, and, when the tissues along the segments are regular, their horizontal lifts concatenate.

Proof. Let $\Delta_k = t_{k+1} - t_k$ and $\tau = (t - t_k)/\Delta_k$. For a trajectory restricted to this interval, the fields in local time are $u^\tau = \Delta_k u$ and $g^\tau = \Delta_k g$, with the same scaling for the frame field. Changing variables in the action gives

$$\mathcal{E}_{[t_k, t_{k+1}]} = \frac{1}{\Delta_k} \mathcal{E}_{[0,1]}.$$

Theorem 2 bounds the right-hand side below by $\overline{\mathrm{WFR}}_\delta^2([\mu_k], [\mu_{k+1}])/\Delta_k$. Summing proves the lower bound in Equation (15).

For the upper bound, choose a minimizing quotient geodesic on each unit interval and rescale it to physical time. At a shared snapshot, the incoming endpoint and outgoing initial measure belong to the same orbit. Push the entire outgoing segment forward by a fixed group element that makes these two measures coincide. This leaves its action and terminal orbit unchanged. Repeating over the intervals gives a continuous measure path attaining the sum. The weak continuity equation holds on the concatenation because the boundary terms at each shared snapshot cancel. For regular horizontal lifts, Proposition A.1 preserves horizontality under these fixed changes of representative, proving the last assertion. □

A.10 Architecture covariance

Proposition A.6 (Architecture covariance)**.** *For $h = (R, b) \in \mathrm{SE}(2)$ and $hz = (Rx + b, q)$, the transformer satisfies*

$$\begin{aligned} v_{\theta,x}[h_\#\rho](hz, t) &= R v_{\theta,x}[\rho](z, t), \\ v_{\theta,q}[h_\#\rho](hz, t) &= v_{\theta,q}[\rho](z, t), \qquad g_\phi[h_\#\rho](hz, t) = g_\phi[\rho](z, t). \end{aligned}$$

Its empirical horizontal field is covariant:

$$\big(P^H_{h_\#\rho} v_{\theta,x}[h_\#\rho]\big)(hz) = R\big(P^H_\rho v_{\theta,x}[\rho]\big)(z).$$

Splitting a positive-weight carrier into coincident copies with positive weights summing to the original weight leaves the attention operator and the predicted fields unchanged, up to floating-point summation.

Proof. Rigid transformations. Let $y_i = x_i - \bar{x}$. Under h, centered positions and pairwise differences transform by R, while the molecular states, time, and total mass are unchanged. Since $R \in SO(2)$,

$$\|Ry_i\| = \|y_i\|, \qquad (Ry_i)^\top (Ry_j) = y_i^\top y_j, \qquad (Ry_i)^\top J (Ry_j) = y_i^\top J y_j.$$

The mass-normalized covariance transforms as $\Sigma' = R\Sigma R^\top$, so its trace, determinant, and the scalar $y_i^\top \Sigma y_i$ are also invariant. Hence the initial embeddings and pairwise scalar features are invariant.

For one attention head, write the logit before its mass bias as ℓ_{ij} and the scalar-channel value as V_j. Its output and normalizer are

$$Z_i = \sum_j m_j e^{\ell_{ij}}, \qquad A_i = \frac{\sum_j m_j e^{\ell_{ij}} V_j}{Z_i}, \qquad \log Z_i.$$

Invariant input features produce invariant queries, keys, values, and logits, so both A_i and $\log Z_i$ are invariant. The residual and pointwise feedforward operations preserve this property. Induction over the blocks therefore proves invariance of all hidden scalar features and of the molecular and growth heads.

Every basis vector of the spatial head transforms by R. For the centroid-relative and covariance bases this follows from

$$y_i' = Ry_i, \qquad Jy_i' = RJy_i, \qquad \Sigma' y_i' = R\Sigma y_i,$$

where $RJ = JR$ for proper planar rotations. Pairwise differences and their quarter turns obey the same rule, and the coefficients used to aggregate them are invariant. The spatial output is thus covariant. Proposition A.1 then proves covariance of its horizontal projection.

Splitting a carrier. Replace (z_j, m_j) by copies $(z_j, m_{j,r})$ with $m_{j,r} > 0$ and $\sum_r m_{j,r} = m_j$. This preserves the tissue measure, its total mass, centroid, and covariance. No node input depends on the individual weight, so all copies have the original node embedding and identical queries, keys, and values. For every query i,

$$\sum_r m_{j,r} e^{\ell_{ij}} = m_j e^{\ell_{ij}}, \qquad \sum_r m_{j,r} e^{\ell_{ij}} V_j = m_j e^{\ell_{ij}} V_j.$$

The split leaves both the numerator and denominator of attention, and therefore its log-normalizer, unchanged. Copies used as queries likewise receive identical updates. Induction over blocks preserves the equality of their embeddings. The output heads depend on these embeddings, moment-based directions, and attention-weighted pairwise directions; each such weighted sum is unchanged by the same identities. This proves the splitting assertion. □

B Implementation and reproducibility details

B.1 From population guarantees to the implementation

The implementation differs from the setting of the population results in Appendix A in three ways. First, the pose search in Equation (5) is nonconvex, so the solver certifies the coupling at the selected pose but not the pose itself. Second, the loss in Equation (9) fits the rigid field on each sampled batch, whereas Equation (13) fits it to the population residual. Proposition A.3 gives convergence for a fixed residual field as batch moments converge, but at finite batch size the fit can be biased, and because attention is also computed over the sampled population, the batch size affects the evaluated field as well as the gradient variance. Third, generation uses a fixed set of carriers. When a characteristic representation of the population equation exists, its positive-weight carriers satisfy

$$\dot{X}_t = P^H_{\rho_t} v_{\theta,x}, \qquad \dot{Q}_t = v_{\theta,q}, \qquad \frac{\mathrm{d}}{\mathrm{d}t} \log w_t = g_\phi. \tag{16}$$

Splitting mass can require several carriers at the same initial state with weights summing to the original mass, and pure creation can require components with zero initial mass. A fixed set of carriers therefore reproduces the population marginal of Corollary A.4 only when it can represent these characteristics, and the implemented RK4 scheme is a finite-carrier discretization of Equation (16). The analytic bridge derivatives use local time, and training divides velocity and growth by the physical interval length.

Algorithm 1 Horizontal flow matching with HORIZON

Require: measures $\{\widehat{\mu}_k\}_{k=0}^{K}$, WFR scale δ, ground weights λ_x, λ_q
1: **for** $k = 0, \dots, K-1$ **do**
2: solve joint $\mathrm{SE}(2)$ qOET (Eq. 5) for (h_k, γ_k) and form the source and target masses of Eq. 6
3: construct transported, decay, and creation bridge components
4: **for** optimization step $s = 1, \dots, S$ **do**
5: sample interval k, one shared local time t, and B cell states from the bridges
6: evaluate analytic (z_t, m_t, u_t^c, g_t^c) (Eqs. 7–8)
7: evaluate equivariant v_θ and invariant g_ϕ
8: fit the shared rigid field (a_B, Ω_B) by weighted least squares
9: update (θ, ϕ) with Equation (9)
10: **infer:** integrate $\dot{X} = P_\rho^H v_{\theta,x}$, $\dot{Q} = v_{\theta,q}$, $\dot{w} = g_\phi w$; recompute P_ρ^H at every stage

B.2 NUMERICAL AND EXPERIMENTAL SETTINGS

Solving qOET. For each interval, a deterministic weighted coreset of at most 1,024 cells provides initial poses for several optimization starts; the coupling solve then uses all cells. At a fixed pose, neighbours under the ground metric initialize a sparse edge set, accelerated projected gradients solve the restricted OET problem, and blockwise scans add every edge with a reduced-cost violation. Column generation stops when the relative primal–dual gap over the full support falls below 10^{-4}, and cut-locus masking excludes forbidden transport. At a fixed coupling, iteratively reweighted least squares Procrustes (Gower, 1975; Umeyama, 1991) proposes a proper rotation and translation; two such updates alternate with coupling solves, and each is accepted only if the joint objective decreases. The lowest objective over initial poses is retained, and all 15 intervals in the real datasets meet the certificate. Couplings remain fixed during neural training.

Network. Node inputs are the molecular PCs, the time encoding, the centred radius and covariance quadratic form of the cell, the log total mass, and the trace and determinant of the mass-weighted covariance; no individual weight is a node input. Pairwise features are radial distances, centred dot products, and signed centred areas. The transformer has five residual blocks with 384 channels, eight attention heads, QK normalization, SwiGLU, and LayerScale. As a numerical check of Proposition A.6, we evaluate each released checkpoint in double precision on up to 256 observed cells with unequal weights and split 64 of them into coincident copies with weight fractions 0.3 and 0.7. Across the four datasets, the largest change in spatial velocity, molecular velocity, or growth at any original or copied cell is below 1.5×10^{-13}, whereas the largest spatial velocities are 7.6–42. Duplicating the same cells without splitting their weight, which changes the measure, changes the spatial velocity by 0.65–6.2.

Training. The optimizer is Adam with cosine decay of the learning rate to 2×10^{-5}. Training cycles through shuffled intervals, draws local time by scrambled stratified sampling, and evaluates an exponential moving average of the weights. Fitting times for every method are measured on one NVIDIA GeForce RTX 4090 D (24 GB). Baseline source revisions are `11c7ae997` for WFR-FM, `26aa79a63` for stVCR, and `61c79f445` for NicheFlow.

Preprocessing. Raw counts are library-normalized to the global median, transformed with $\log(1 + x)$, standardized over at most 2,000 prevalence-filtered high-variance genes, and mapped to a joint 32-dimensional PCA for each dataset. Spatial coordinates keep their recorded orientation and origin and are divided by the median, across snapshots, of the within-snapshot median 30-neighbour distance. The zebrafish series starts at 5.25 hpf because the retained sections then share a common embryonic anatomy, whereas the 3.3 hpf sample contains a distinct extra-embryonic structure. stVCR uses its published sampling schedule, and STORIES 0.1.1 uses its published normalization and hyperparameters on all four series.

Metric definitions. SW is weighted sliced W_2 with 128 fixed random projections. MMD is the square root of biased weighted RBF MMD^2 with a pooled median-distance bandwidth. Chamfer is the square root of the average of the two probability-weighted nearest-neighbour squared distances. Joint metrics use the same scaled (x, q) ground blocks as the method. Before every quotient metric, a

Table 3: **Shared HORIZON configuration.** One architecture and optimizer are used on every real dataset: 9,000 updates, width 384, five blocks, eight heads, feed-forward width 768, initial learning rate 1.5×10^{-3} with cosine decay to 2×10^{-5}, EMA 0.995, gradient clip 100, $\delta = 10$, $\lambda_x = 1$, and projected RK4 maximum normalized step 0.25. B is the number of bridge states per update. Figures use seed 17, and Table 7 also includes seeds 29 and 43.

Dataset	Times	Full cells per snapshot	B	λ_q	Reference seed	Time input
Axolotl	2, 5, 10, 15, 20	7,668, 8,106, 9,440, 9,676, 11,319	1,024	0.15	17	piecewise Fourier
Chicken	4, 7, 10, 14	147, 528, 908, 1,967	768	0.05	17	piecewise Fourier
Drosophila	5.6, 6.3, 7, 9.1, 10.5	1,263, 1,248, 1,039, 1,078, 1,110	768	0.05	17	piecewise Fourier
Zebrafish	5.25, 10, 12, 18, 24	563, 1,036, 2,081, 3,048, 5,271	1,024	0.05	17	piecewise Fourier

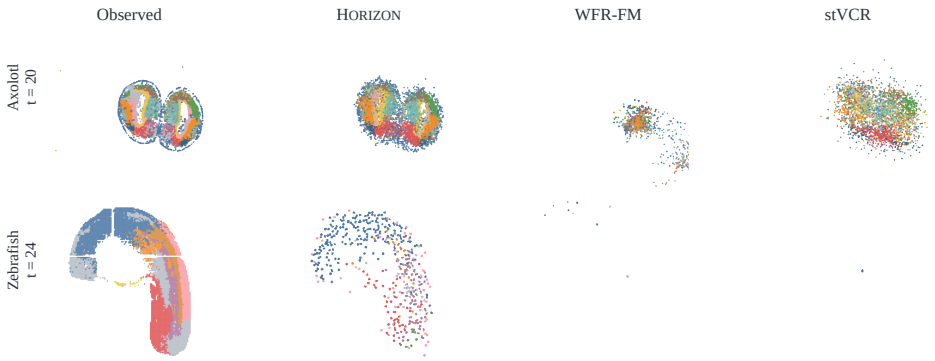

Figure 3: **Spatial comparison at the final snapshot.** All methods use the same inputs and display canvas, and predictions are integrated from the first snapshot and aligned by the evaluation pose. Each displayed point carries equal mass, so point count shows total mass; colour shows the shared molecular PCs.

scorer independent of qOET searches 72 proper rotation angles, refines eight candidates, and profiles translation through weighted centroids using spatial Chamfer. The same alignment is applied to every method.

C DETAILS OF THE FRAME-PERTURBATION BENCHMARK

Setup. The benchmark starts from the Rectangle trajectory distributed with stVCR. Three independently sampled intrinsic trajectories use seeds 1103, 2207, and 3301, and frame seeds 401, 503, and 607 generate the mild and severe proper-$\mathrm{SE}(2)$ perturbations. The 21 training datasets comprise three canonical, nine mild, and nine severe views, and each frame seed is paired with the canonical view of its intrinsic trajectory. Each dataset contains six snapshots and retains half of the available cell states, with a minimum of 128. A separate 51-frame intrinsic trajectory provides the ground truth for visualization, and the reported metrics average the five snapshots after the first.

The qOET solves use an entropic warm start with regularization $(0.5, 0.2, 0.08, 0.03)$ and 500 Sinkhorn iterations per stage (Cuturi, 2013; Peyré & Cuturi, 2019) at tolerance 10^{-8}, followed by an unregularized APG solve of up to 50,000 iterations with relative gap tolerance 10^{-4}. WFR-FM uses 3,000 updates and stVCR 500 epochs, their published defaults for this trajectory. The *no quotient* ablation fixes the pose to the identity, keeps the coupling solver and transformer unchanged, and trains and integrates without the projector. The *no growth* ablation evaluates the trained HORIZON model with its reaction field set to zero.

Results. Across the three regimes and 11 metrics, HORIZON has lower mean error than WFR-FM in 32 of 33 comparisons and than stVCR in 29 of 33. HORIZON and stVCR complete all 21 fits and WFR-FM 20. Tables 4–6 report all three regimes, and Appendix F reports the sensitivity to δ and λ_q.

Registration without projection. For the *registered, no projector* ablation in Table 1, the interval poses h_k from qOET are composed to align all snapshots with the first frame, as in the aligned NicheFlow comparison. WFR bridges are then built at the identity pose, and the HORIZON architecture is trained with the same optimizer and integrated with the same step size, both without the

Table 4: Frame-perturbation benchmark: direct and quotient sliced-Wasserstein metrics (mean±standard error over successful runs; n reports successes/attempts; lower is better).

Method	Regime	n	Direct W_x	Direct $W_{x,g}$	$W_x^{\mathcal{Q}}$	$W_{x,g}^{\mathcal{Q}}$
HORIZON (ours)	Canonical	3/3	0.111 ± 0.004	0.079 ± 0.005	$0.040 \pm 9.84 \times 10^{-4}$	$0.035 \pm 9.90 \times 10^{-4}$
	Mild	9/9	0.114 ± 0.004	0.082 ± 0.003	0.039 ± 0.002	0.037 ± 0.002
	Severe	9/9	0.115 ± 0.003	0.081 ± 0.002	0.042 ± 0.002	0.037 ± 0.001
WFR-FM	Canonical	2/3	0.223 ± 0.165	0.295 ± 0.222	0.219 ± 0.162	0.293 ± 0.221
	Mild	9/9	0.337 ± 0.007	0.227 ± 0.004	0.075 ± 0.003	0.087 ± 0.005
	Severe	9/9	1.213 ± 0.027	0.786 ± 0.016	0.110 ± 0.008	0.142 ± 0.016
stVCR	Canonical	3/3	0.115 ± 0.015	0.084 ± 0.009	0.066 ± 0.001	0.062 ± 0.002
	Mild	9/9	0.119 ± 0.005	0.087 ± 0.003	0.067 ± 0.001	0.063 ± 0.001
	Severe	9/9	0.249 ± 0.021	0.176 ± 0.012	0.119 ± 0.006	0.100 ± 0.005

Table 5: Frame-perturbation benchmark: direct and quotient Chamfer and joint-MMD metrics (mean±standard error over successful runs; n reports successes/attempts; lower is better).

Method	Regime	n	Direct Chamfer	Direct joint MMD	Quotient Chamfer	Quotient joint MMD
HORIZON (ours)	Canonical	3/3	0.116 ± 0.005	0.023 ± 0.002	$0.096 \pm 1.91 \times 10^{-4}$	$0.006 \pm 2.58 \times 10^{-4}$
	Mild	9/9	0.117 ± 0.004	0.024 ± 0.001	$0.095 \pm 6.91 \times 10^{-4}$	$0.006 \pm 5.75 \times 10^{-4}$
	Severe	9/9	0.119 ± 0.004	$0.023 \pm 9.84 \times 10^{-4}$	$0.096 \pm 5.36 \times 10^{-4}$	$0.006 \pm 3.90 \times 10^{-4}$
WFR-FM	Canonical	2/3	0.273 ± 0.164	$0.020 \pm 1.37 \times 10^{-5}$	0.272 ± 0.163	$0.019 \pm 3.48 \times 10^{-4}$
	Mild	9/9	0.227 ± 0.009	0.085 ± 0.002	0.127 ± 0.003	0.021 ± 0.001
	Severe	9/9	0.721 ± 0.026	0.311 ± 0.007	0.148 ± 0.009	0.033 ± 0.006
stVCR	Canonical	3/3	0.100 ± 0.002	0.030 ± 0.005	$0.094 \pm 9.63 \times 10^{-4}$	$0.017 \pm 4.39 \times 10^{-4}$
	Mild	9/9	$0.101 \pm 9.16 \times 10^{-4}$	0.031 ± 0.002	0.094 ± 0.001	$0.018 \pm 6.46 \times 10^{-4}$
	Severe	9/9	0.266 ± 0.024	0.055 ± 0.004	0.137 ± 0.004	0.028 ± 0.002

projector. The ablation is compared with a HORIZON control retrained alongside it on the nine severe variants with the same architecture and hyperparameters, and the control matches the main benchmark to within 0.004 on every spatial metric. Paired over variants, the full model lowers direct spatial SW by 0.011 ± 0.003, direct joint SW by 0.006 ± 0.002, and direct Chamfer by 0.013 ± 0.003, while the difference in quotient spatial SW is 0.003 ± 0.003. Evaluation alignment removes the accumulated rigid offset, so the effect of the projector appears in the direct metrics.

D RUNTIME AND BASELINE COMPARISONS

All comparisons use the same processed inputs and evaluator. Table 7 adds the seed standard deviations to the free-generation block of Table 2, and Table 8 gives the fitting times.

Comparison with NicheFlow. When HORIZON generates from the first snapshot while NicheFlow restarts every interval, HORIZON still has lower error than the better NicheFlow variant in 15 of 24 comparisons at seed 17 (14 of 24 for three-seed means, the difference being chicken's spatial SW). Aligning NicheFlow's inputs with the qOET poses gives worse results than the raw inputs on three of the four series.

Removing the quotient on real data. Table 9 repeats the *no quotient* ablation of the frame-perturbation benchmark on the real time courses. The WFR coupling is solved with the pose fixed to the identity by dense APG with the same continuation schedule, a budget of $30{,}000$ iterations, and a target relative gap of 10^{-4}. The network of Table 3 is trained on these bridges and integrated at the same step size, both without the projector. On chicken, removing the quotient more than doubles spatial SW and gene SW and quadruples the mass error, whereas on Drosophila, whose frames are nearly aligned, spatial SW changes only from 0.043 to 0.049. On zebrafish (last interval, 18 to 24 hpf, $3{,}048 \times 5{,}271$ cells) and axolotl (first interval), the identity-pose coupling reached neither the 10^{-4} certificate nor a relaxed 10^{-3} certificate within $150{,}000$ iterations, whereas joint optimization of pose and coupling certifies all fifteen intervals at 10^{-4}. With the pose fixed, misaligned pairs require long-range transport, and the coupling problem becomes ill-conditioned.

NicheFlow configuration on chicken. NicheFlow fixes one microenvironment budget from the largest snapshot and samples every snapshot to that budget without replacement, which the 147-cell chicken D4 slice cannot supply at the published grid. A finer grid does not help, because the smallest

Table 6: Frame-perturbation benchmark: gene-space and finite-mass metrics (mean±standard error over successful runs; n reports successes/attempts; lower is better).

Method	Regime	n	W_g	Gene MMD	Mass RE
HORIZON (ours)	Canonical	3/3	0.076 ± 0.008	0.011 ± 0.002	$2.68 \times 10^{-4} \pm 1.06 \times 10^{-4}$
	Mild	9/9	0.084 ± 0.008	0.012 ± 0.002	$2.03 \times 10^{-4} \pm 4.93 \times 10^{-5}$
	Severe	9/9	0.070 ± 0.005	$0.010 \pm 9.51 \times 10^{-4}$	$1.92 \times 10^{-4} \pm 5.57 \times 10^{-5}$
WFR-FM	Canonical	2/3	0.484 ± 0.298	0.078 ± 0.007	$7.10 \times 10^{-4} \pm 9.89 \times 10^{-5}$
	Mild	9/9	0.204 ± 0.011	0.068 ± 0.004	$6.23 \times 10^{-4} \pm 4.72 \times 10^{-5}$
	Severe	9/9	0.298 ± 0.028	0.115 ± 0.009	$0.001 \pm 3.00 \times 10^{-4}$
stVCR	Canonical	3/3	0.127 ± 0.008	0.077 ± 0.002	0.035 ± 0.005
	Mild	9/9	0.136 ± 0.004	0.081 ± 0.002	0.035 ± 0.003
	Severe	9/9	0.160 ± 0.012	0.132 ± 0.015	0.022 ± 0.002

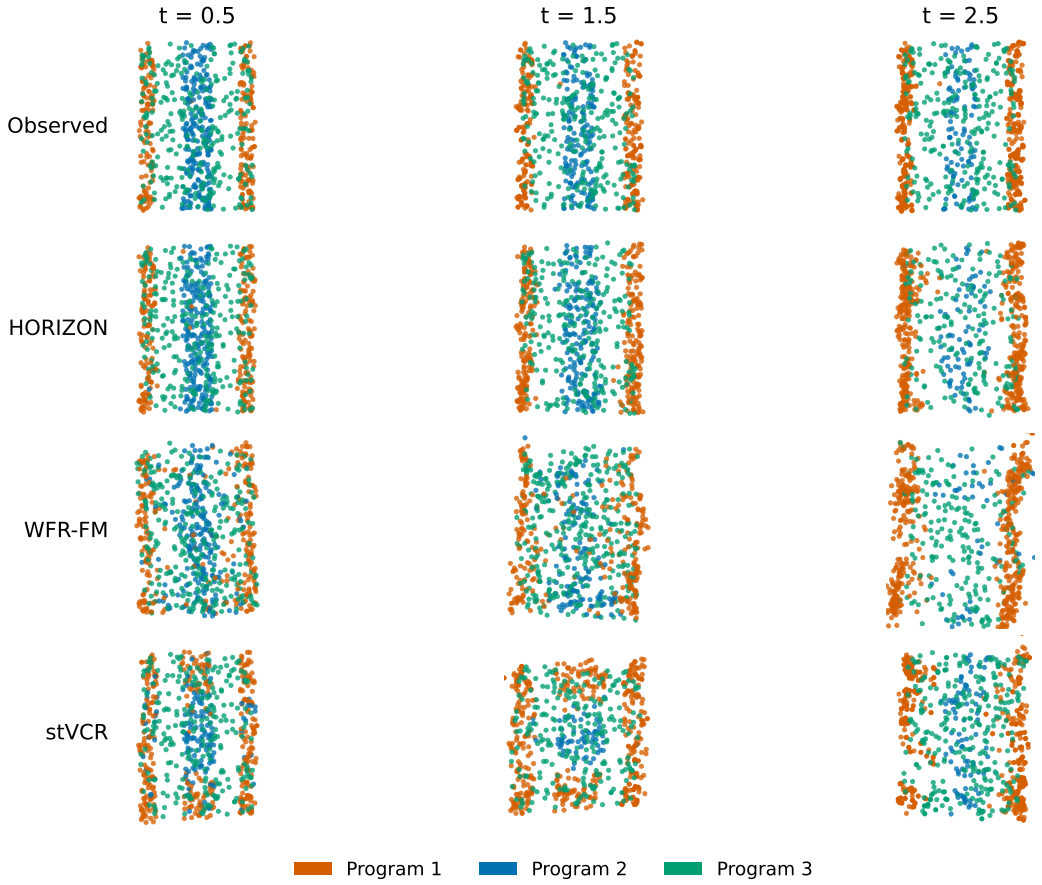

Figure 4: **Trajectories under severe frame perturbations.** Observed snapshots and predictions of HORIZON, WFR-FM, and stVCR at three times for the first severe variant. Colors show the three molecular programs, and all panels share the canvas of the observations after evaluation alignment.

snapshot saturates at one microenvironment per cell while the budget set by the largest keeps growing. We therefore coarsen the grid to the finest feasible spacing, $dx/dy = 0.27/0.36$ against the published $0.15/0.20$ at the same aspect ratio, which lowers the budget from 432 to 138 microenvironments. In addition, at the published radius the two earliest chicken snapshots yield a modal microenvironment of one cell, so we raise the radius to 0.30, the smallest value at which every chicken snapshot reaches the neighbourhood size that the published radius already yields on Drosophila. Both changes are confined to chicken. Table 11 reports both NicheFlow input configurations with their generation-seed variation and fitting times.

STORIES under both evaluators and both protocols. STORIES reports molecular metrics through its own evaluator, whose normalization differs from the shared evaluator, and its standard protocol is adjacent teacher forcing. Table 2 evaluates free generation from the first snapshot; Table 10

Table 7: **Free-generation results of Table 2 with seed variation.** Lower is better. HORIZON and WFR-FM report mean ± sample standard deviation over three seeds, and stVCR uses one seed.

Dataset	Method	Joint SW	Joint MMD	Spatial SW	Chamfer	Gene SW	Gene MMD	Mass RE (%)
Axolotl	HORIZON	**0.0629 ± 0.0049**	**0.0034 ± 0.0008**	**0.2400 ± 0.0223**	**0.3667 ± 0.0103**	**0.1218 ± 0.0037**	**0.0178 ± 0.0005**	**0.0912 ± 0.0173**
	WFR-FM	1.4411 ± 0.1103	0.1669 ± 0.0294	5.9074 ± 0.4223	2.8603 ± 0.4860	1.1715 ± 0.2544	0.2137 ± 0.0466	146.2522 ± 67.1410
	stVCR	1.0405	0.2180	4.2944	1.0078	0.2482	0.0436	1.2956
Chicken	HORIZON	**0.1341 ± 0.0071**	**0.0367 ± 0.0011**	**0.1994 ± 0.0269**	**0.3916 ± 0.0099**	**0.5971 ± 0.0067**	**0.0982 ± 0.0079**	**0.4523 ± 0.1870**
	WFR-FM	0.1873 ± 0.0343	0.0637 ± 0.0178	0.5397 ± 0.1989	0.5761 ± 0.1281	0.7513 ± 0.0141	0.1352 ± 0.0049	1.5590 ± 0.6506
	stVCR	0.5765	0.2048	1.8059	2.0790	1.5178	0.2617	99.1742
Drosophila	HORIZON	**0.0352 ± 0.0119**	**0.0130 ± 0.0066**	**0.0426 ± 0.0023**	**0.1248 ± 0.0017**	**0.1065 ± 0.0025**	**0.0226 ± 0.0011**	**0.0980 ± 0.0281**
	WFR-FM	0.0741 ± 0.0140	0.0314 ± 0.0082	0.0643 ± 0.0031	0.1519 ± 0.0010	0.1570 ± 0.0039	0.0319 ± 0.0008	0.2326 ± 0.0172
	stVCR	0.2280	0.1119	0.8357	0.6686	0.4273	0.1096	11.4701
Zebrafish	HORIZON	**0.0707 ± 0.0039**	**0.0093 ± 0.0011**	**0.2058 ± 0.0142**	**0.3585 ± 0.0146**	**0.2886 ± 0.0140**	**0.0582 ± 0.0014**	**0.5060 ± 0.0452**
	WFR-FM	1.0349 ± 0.2718	0.1603 ± 0.0285	4.0486 ± 1.0621	3.1423 ± 1.2365	1.4411 ± 0.2543	0.3743 ± 0.1079	44.9785 ± 8.1872
	stVCR	1.8620	0.3312	7.3944	7.8549	1.8623	0.3766	99.4662

Table 8: **Fitting time on one GPU, in seconds.** HORIZON fit includes the one-time sparse qOET solve and simulation-free neural regression. stVCR uses its published training procedure. Speedup compares complete HORIZON fitting with stVCR.

	HORIZON			stVCR	stVCR/ours
Dataset	qOET	NN	fit	NN	fit speedup
Axolotl	794.1	298.3	1,092.4	12,424.4	11.4×
Chicken	465.5	281.2	746.7	1,950.3	2.6×
Drosophila	361.4	440.3	801.7	5,704.7	7.1×
Zebrafish	929.3	288.8	1,218.2	4,752.6	3.9×

reports both protocols under both evaluators, all with seed 17 and the configuration of Appendix B. Under every combination HORIZON has lower gene SW and MMD than STORIES. On its own Sinkhorn objective, STORIES is lower on axolotl under teacher forcing and on chicken and zebrafish under both protocols.

E SPARSE QOET SCALING

Table 12 compares the blockwise column-generation solver with dense APG on identical problems at a fixed pose. Agreement is measured by objective and endpoint marginals because an optimal coupling need not be unique. At $8{,}192^2$ support the sparse solver preserves the objective to 7.7×10^{-6} relative error and both marginals to 1.5×10^{-4} relative ℓ_1 error, while using 11.3-fold less peak memory and 10.5-fold less wall clock. The largest full zebrafish interval, $3{,}048 \times 5{,}271$, is 1.68-fold faster and uses 38% less peak memory. All sparse rows meet the same certificate over full support, with tolerance 10^{-4}, as the qOET solutions on real data.

F SENSITIVITY TO THE WFR SCALE AND THE MOLECULAR WEIGHT

Table 13 varies the WFR scale and the molecular weight on three severe variants, one per intrinsic seed, with all other settings held fixed. We test $\delta \in \{5, 10, 20\}$ and $\lambda_q \in \{0.1, 0.25, 0.5\}$, varying one parameter at a time. Quotient spatial SW stays within 0.038–0.052 and joint SW within 0.034–0.043, compared with 0.044 and 0.038 at the reference setting. Mass RE is 0.06% at $\delta = 5$ and rounds to 0.00% at $\delta = 20$. Gene SW varies more across variants than across settings.

G TIMEPOINT HOLDOUT STRESS TEST

Each HORIZON fold uses the full-time-course settings with 9,000 training steps and the same preprocessing, and recomputes qOET across the gap left by the removed snapshot. WFR-FM uses its published 3,000 updates with seed 11, the same budget as in the main comparison. Both methods exclude the held-out snapshot from coupling and training, score only that target, and stop generation at its time.

Table 9: **Removing the quotient on real data.** Lower is better. *no quotient* fixes the pose to the identity and removes the projector; the architecture, optimizer, step size, and evaluator match HORIZON, and both rows generate from the first snapshot. The HORIZON rows repeat Table 2.

Dataset	Method	Joint SW	Joint MMD	Spatial SW	Chamfer	Gene SW	Gene MMD	Mass RE (%)
Chicken	HORIZON	**0.134**	**0.037**	**0.199**	**0.392**	**0.597**	**0.098**	**0.45**
	no quotient	0.242	0.070	0.448	0.527	1.234	0.205	2.06
Drosophila	HORIZON	**0.035**	**0.013**	**0.043**	**0.125**	**0.107**	0.023	**0.10**
	no quotient	0.050	0.021	0.049	0.127	0.111	0.023	0.14

Table 10: **STORIES under its own evaluator and under the shared evaluator.** Lower is better; seed 17 throughout. STORIES' evaluator uses its normalization and Sinkhorn objective; the shared evaluator is the one used in Table 2, where the free-generation STORIES row appears. Bold marks the best entry per column and dataset.

		STORIES evaluator			Shared evaluator	
Dataset	Method	$\text{Sinkhorn}_{\epsilon=.1}$	Gene SW	Gene MMD	Gene SW	Gene MMD
Axolotl	HORIZON (first)	0.2132	**0.1248**	**0.0184**	**0.1178**	**0.0173**
	STORIES (first)	0.2709	1.1093	0.2101	1.0867	0.2071
	STORIES (adj. TF)	**0.1847**	0.6373	0.1332	0.6525	0.1376
Chicken	HORIZON (first)	0.4407	**0.6750**	**0.1089**	**0.6019**	**0.0969**
	STORIES (first)	**0.2654**	1.4526	0.2705	1.4761	0.2706
	STORIES (adj. TF)	0.2748	1.0530	0.2257	1.0900	0.2289
Drosophila	HORIZON (first)	**0.4213**	**0.1150**	**0.0247**	**0.1037**	**0.0214**
	STORIES (first)	0.5939	0.8224	0.2535	0.7969	0.2407
	STORIES (adj. TF)	0.5696	0.5413	0.1579	0.5609	0.1604
Zebrafish	HORIZON (first)	0.1100	**0.3251**	**0.0669**	**0.2861**	**0.0578**
	STORIES (first)	0.0907	0.8332	0.2158	0.8309	0.2155
	STORIES (adj. TF)	**0.0813**	0.3833	0.1069	0.4342	0.1219

Table 14 lists every fold. The 48 of 60 comparisons in which HORIZON has lower error than WFR-FM split into 15 of 18 on axolotl, 12 of 12 on chicken, 9 of 18 on Drosophila, and 12 of 12 on zebrafish, so most of the losses occur on Drosophila, whose frames are nearly aligned. WFR-FM fails the zebrafish 18-hpf fold when learned growth drives log weights beyond the finite range, and stVCR fails the 10-hpf fold when learned growth eliminates the population.

H POPULATION MASS OVER TIME

Figure 5 shows how the learned reaction field tracks total mass. Chicken and zebrafish grow by factors of 13.38 and 9.36, axolotl reaches 1.48, and Drosophila contracts before recovering.

I ADDITIONAL BIOLOGICAL ANALYSES

To assess whether the couplings preserve biological annotations, we compute label affinity weighted by transported mass, $\sum_{ij} \gamma_{ij} \mathbf{1}[y_i = y_j] / \sum_{ij} \gamma_{ij}$, and divide it by the probability of matching labels under independent transported marginals. Mean lift is 5.85 for fine labels in axolotl, 2.24 for fine labels in zebrafish, 2.82 for tissue labels in Drosophila, and 1.17 for region labels in chicken. Lift for axolotl injury side ranges from 1.92 to 1.97.

The program analyses complement these coupling summaries. Chicken, Drosophila, and zebrafish panels use global weighted means of matched high-variance genes along the generated trajectory. Chicken programs use HAND1/2, MYH1F/MYH1B, collagen/PCOLCE, and HBZ. Drosophila programs use Mhc/actins and Hsp83. Zebrafish programs use FOXA, TBX/mesp, myosin light chain, NANOG, keratin, and HOX sets. For axolotl, identities of injured and uninjured cohorts at 2 DPI are propagated by the Lagrangian flow.

Table 11: **NicheFlow with generation-seed variation and fitting time.** Lower is better; all rows use the shared evaluator. HORIZON generates from the first snapshot with the seed-17 model, and NicheFlow restarts every interval from its observed source (Table 2 compares both methods under this protocol). NicheFlow is trained for 100,000 steps, and its values are mean ± sample standard deviation over five generation seeds. Its aligned input uses the proper $\mathrm{SE}(2)$ qOET poses.

Dataset	Method	Protocol	Joint SW	Joint MMD	Spatial SW	Chamfer	Gene SW	Gene MMD	Fit (h)
Axolotl	HORIZON	first	0.0605	0.0029	0.2292	0.3616	0.1178	0.0173	0.30
	NicheFlow raw	adj. TF	0.201±0.005	0.025±0.001	0.843±0.023	0.360±0.003	0.279±0.008	0.060±0.001	9.36
	NicheFlow aligned	adj. TF	0.270±0.001	0.038±0.000	1.055±0.007	0.383±0.004	0.345±0.004	0.076±0.001	9.82
Chicken	HORIZON	first	0.1328	0.0356	0.1855	0.3898	0.6019	0.0969	0.21
	NicheFlow raw	adj. TF	0.064±0.002	0.025±0.001	0.198±0.005	0.150±0.002	0.217±0.007	0.042±0.001	6.91
	NicheFlow aligned	adj. TF	0.079±0.001	0.035±0.001	0.268±0.004	0.149±0.001	0.231±0.007	0.047±0.001	6.80
Drosophila	HORIZON	first	0.0489	0.0206	0.0447	0.1229	0.1037	0.0214	0.22
	NicheFlow raw	adj. TF	0.059±0.011	0.030±0.006	0.158±0.004	0.140±0.001	0.160±0.002	0.045±0.000	6.57
	NicheFlow aligned	adj. TF	0.088±0.012	0.046±0.006	0.185±0.003	0.142±0.001	0.175±0.002	0.051±0.001	6.75
Zebrafish	HORIZON	first	0.0700	0.0090	0.2072	0.3682	0.2861	0.0578	0.34
	NicheFlow raw	adj. TF	0.135±0.002	0.029±0.001	0.537±0.009	0.264±0.004	0.243±0.005	0.058±0.001	7.04
	NicheFlow aligned	adj. TF	0.141±0.002	0.029±0.001	0.561±0.008	0.222±0.003	0.251±0.009	0.052±0.001	6.94

Table 12: **Certified sparse qOET scaling at a fixed pose.** Dense and column-generated solves use the same 10^{-4} full-gap tolerance. Memory is peak allocated GPU GiB; $\Delta J/J$ and marginal error compare the two converged solutions.

Support	Dense		Sparse		Speedup	Agreement
	sec	GiB	sec	GiB		$\Delta J/J$ / marg.
$1{,}024^2$	10.3	0.14	11.2	0.08	0.92×	1.1e–8 / 2.8e–6
$2{,}048^2$	25.7	0.55	17.8	0.16	1.44×	9.7e–7 / 1.0e–4
$4{,}096^2$	110.2	2.18	30.1	0.33	3.66×	1.6e–5 / 2.0e–4
$8{,}192^2$	649.1	8.69	62.0	0.77	10.48×	7.7e–6 / 1.5e–4
$3{,}048 \times 5{,}271$	290.8	2.18	173.0	1.35	1.68×	3.7e–5 / 1.2e–4

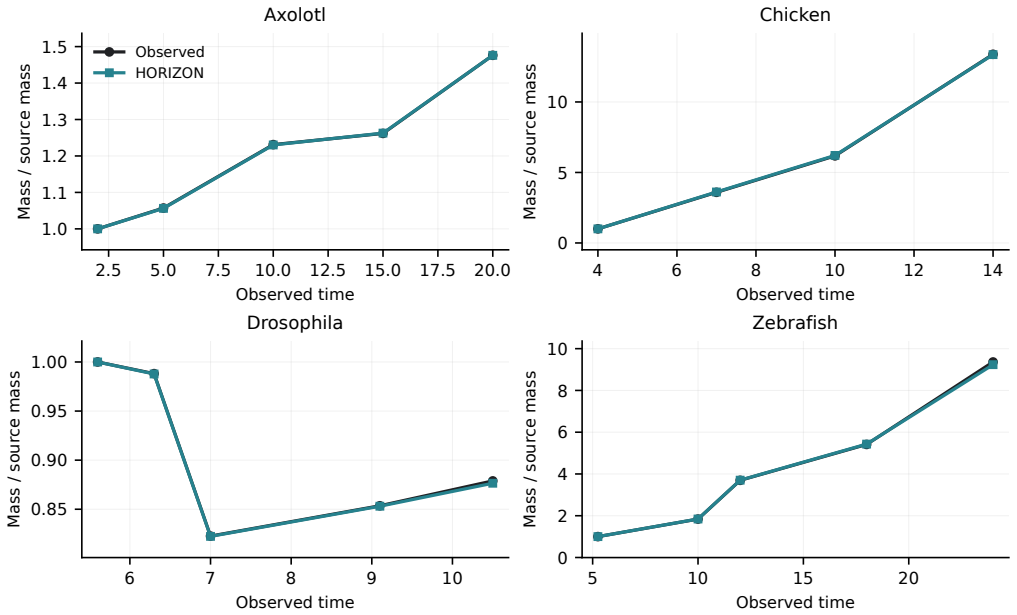

Figure 5: **Population growth across all real time courses.** Observed cell count ratios and HORIZON predictions are normalized to the first snapshot. Curves show the predictive mean over three seeds with sample standard deviation. The model follows both expansion and contraction; mean relative error over target snapshots ranges from 0.091% to 0.506%.

Table 13: **Sensitivity to the WFR scale δ and the molecular weight λ_q** on the severe regime of the frame-perturbation benchmark (mean $\pm$ standard error over three variants, one per intrinsic seed; n = completed runs).

Configuration	n	Direct sp. SW	Quotient sp. SW	Quotient joint SW	Gene SW	Mass RE (%)
$\delta = 10$, $\lambda_q = 0.25$ (reference)	3	0.120 ± 0.003	0.044 ± 0.004	0.038 ± 0.004	0.086 ± 0.019	0.03 ± 0.01
$\delta = 5$	3	0.081 ± 0.010	0.038 ± 0.001	0.034 ± 0.002	0.069 ± 0.005	0.06 ± 0.01
$\delta = 20$	3	0.126 ± 0.009	0.052 ± 0.004	0.043 ± 0.003	0.066 ± 0.006	0.00 ± 0.00
$\lambda_q = 0.1$	3	0.107 ± 0.011	0.038 ± 0.001	0.041 ± 0.005	0.107 ± 0.021	0.02 ± 0.01
$\lambda_q = 0.5$	3	0.123 ± 0.002	0.043 ± 0.002	0.036 ± 0.001	0.072 ± 0.004	0.01 ± 0.00

Table 14: **Results for individual holdout folds.** Lower is better. Each row is an independently trained fold.

Dataset	Held out	Method	Spatial SW	Chamfer	Gene SW	Gene MMD	Joint SW	Joint MMD
Axolotl	5 DPI	HORIZON	**1.1144**	**0.4706**	1.1985	0.2801	**0.4030**	**0.07918**
		WFR-FM	2.0499	0.8889	**1.1741**	0.2824	0.5615	0.09006
		stVCR	5.2877	0.7662	1.2062	**0.2793**	1.2498	0.2958
	10 DPI	HORIZON	**1.1916**	**0.7883**	2.2785	0.4573	**0.6121**	**0.1125**
		WFR-FM	3.2498	2.6579	**1.8388**	**0.3835**	0.8631	0.1261
		stVCR	5.0112	1.3247	2.3118	0.4996	1.3296	0.2670
	15 DPI	HORIZON	**1.6507**	**0.5153**	**2.8597**	**0.6156**	**0.7702**	**0.1659**
		WFR-FM	5.4997	1.4054	2.9679	0.6321	1.5301	0.2545
		stVCR	3.6716	0.6001	3.1105	0.6536	1.1758	0.2660
Chicken	D7	HORIZON	**0.3626**	**0.3540**	**0.8776**	**0.1885**	**0.1726**	**0.09744**
		WFR-FM	0.3676	0.3695	1.0686	0.2188	0.1986	0.1011
		stVCR	1.7355	1.8389	1.6706	0.2659	0.5704	0.2562
	D10	HORIZON	**0.5036**	**0.4958**	**0.6841**	**0.1625**	**0.1894**	**0.09546**
		WFR-FM	0.9197	0.8415	1.0346	0.2262	0.2362	0.1155
		stVCR	1.8973	1.8462	1.6085	0.2929	0.5879	0.2338
Drosophila	6.3	HORIZON	**0.1929**	**0.1624**	0.2587	0.07411	**0.1746**	0.08301
		WFR-FM	0.2057	0.1627	**0.2446**	**0.06971**	0.1807	**0.08290**
		stVCR	1.2277	1.0658	0.3127	0.08753	0.2863	0.1022
	7	HORIZON	**0.2458**	**0.1668**	**0.4997**	**0.1240**	0.1597	0.08124
		WFR-FM	0.2506	0.1679	0.5339	0.1365	**0.1032**	**0.05360**
		stVCR	0.8154	0.4151	0.6412	0.1611	0.2739	0.1488
	9.1	HORIZON	**0.06340**	**0.1485**	0.6278	0.2091	0.2019	0.1106
		WFR-FM	0.06917	0.1551	**0.5175**	**0.1882**	**0.1796**	**0.1074**
		stVCR	1.2712	0.5366	0.6520	0.2247	0.3649	0.2103
Zebrafish	10 hpf	HORIZON	**0.7822**	**0.7900**	**0.7003**	**0.2378**	**0.2528**	**0.07266**
		WFR-FM	5.9320	5.4335	1.5381	0.3999	1.5076	0.4154
		stVCR	failed: learned growth eliminated the population					
	12 hpf	HORIZON	**0.6474**	**0.5200**	**0.7653**	**0.2055**	**0.2367**	**0.05822**
		WFR-FM	1.7982	1.6627	1.1956	0.2820	0.5129	0.1069
		stVCR	6.2865	6.7300	1.7878	0.3491	1.5713	0.3018
	18 hpf	HORIZON	**1.9884**	**1.3160**	**0.9462**	**0.2680**	**0.4896**	**0.1417**
		WFR-FM	failed: non-finite growth during generation					
		stVCR	5.1215	5.3531	1.8569	0.5967	1.2778	0.2929

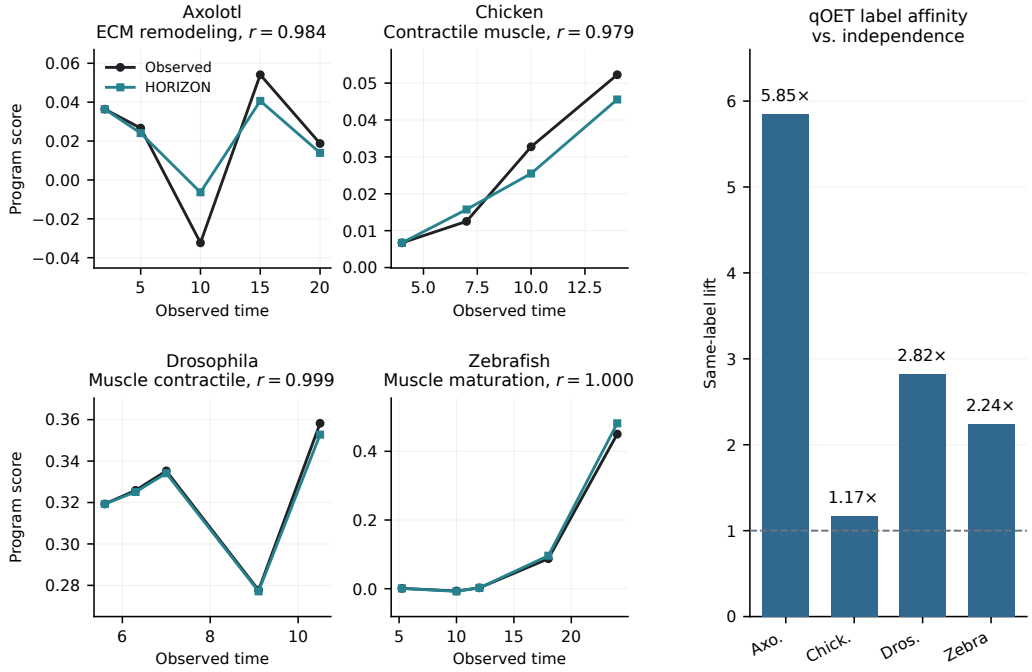

Figure 6: **Biological program dynamics and coupling structure.** The four panels on the left compare observed and predicted program trajectories. The panel on the right summarizes label affinity weighted by transported mass, averaged over intervals and relative to independent marginals.

Contrasts in reactive glia (VIM/GFAP/SLC1A3/SOX2/JUN), neuronal maturation (DCX/MAP2/GAD1/GAD2/SYP), and collagen remodeling are compared with observed contrasts between injury sides. The qOET lift for injury labels supports this cohort interpretation. Correlations use interval summaries after the initial snapshot, with $n = 4, 3, 4, 4$ for axolotl, chicken, Drosophila, and zebrafish. Across these program sets, correlations span 0.948–0.989, 0.979–1.000, 0.989–0.999, and 0.993–1.000, respectively.

J ADDITIONAL SPATIAL TRAJECTORIES

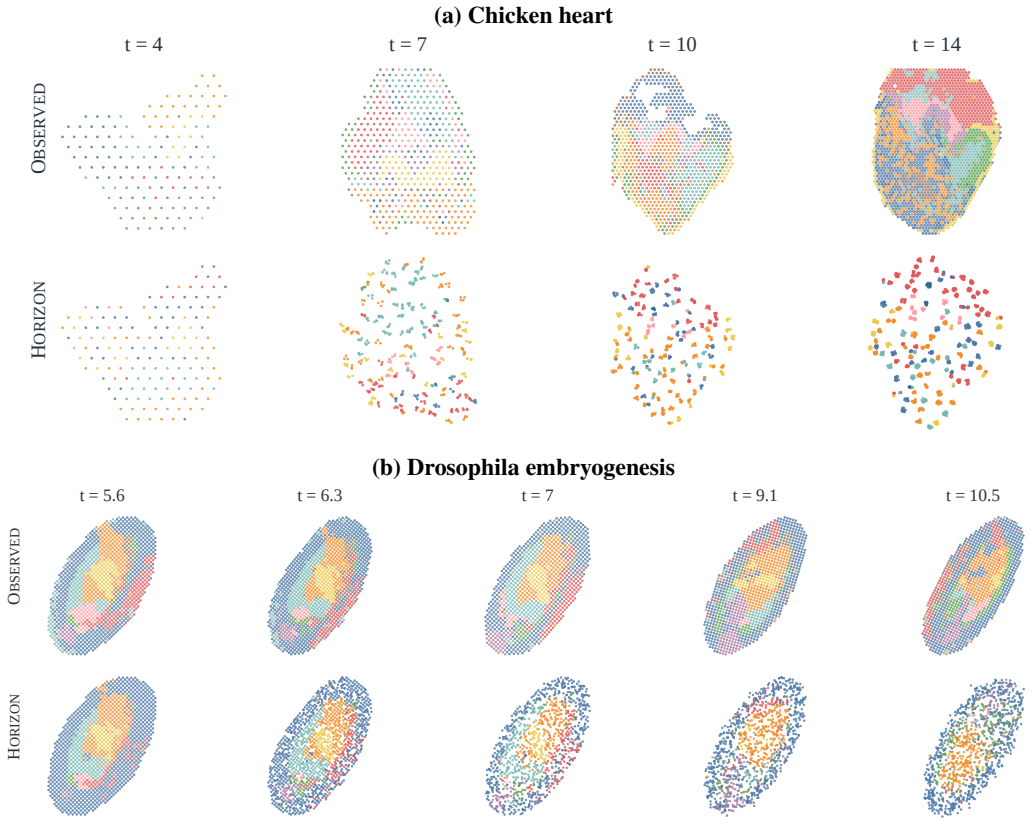

Figure 7: **Additional generated time courses.** Zebrafish is shown in Figure 2. Observations are on top and HORIZON predictions on the bottom. Each displayed point carries equal mass, so point count shows abundance. Colors show observed annotations on top and annotations decoded from the predicted molecular PCs on the bottom.

Continuous trajectory animations and comparisons at observed times for all four datasets are included in the supplementary material.