Branched Schrödinger Bridge Matching

Sophia Tang[1], Yinuo Zhang[2], Alexander Tong[3,4], Pranam Chatterjee[1,5 †]

[1]Department of Computer and Information Science, University of Pennsylvania
[2]Center of Computational Biology, Duke-NUS Medical School
[3]Mila, Quebec AI Institute, [4]Université de Montréal
[5]Department of Bioengineering, University of Pennsylvania
[†]Correspondence to: pranam@seas.upenn.edu

Motivation

1. **Schrödinger Bridge Matching Frameworks are Limited to Single Continuous Trajectories.** Predicting how a population evolves between an initial and final state is central to many problems in generative modeling, from simulating perturbation responses to modeling cell fate decisions. Existing approaches, such as Flow Matching and Schrödinger Bridge Matching, effectively learn mappings between two distributions by modeling a single stochastic path but are restricted to modeling evolution along a single path.
2. **Branching Dynamics is Governed by Energy-Minimizing Trajectories.** To define more complex systems where the optimal dynamics cannot be accurately captured by minimizing the standard squared Euclidean cost in entropic OT, the Generalized Schrödinger Bridge (GSB) problem introduces an additional nonlinear state-cost.
3. **Branching is Central to Modeling Cellular Perturbation Responses.** When a homogeneous cell population undergoes a perturbation, such as gene knockouts or drug treatments, it frequently induces fate bifurcation as the cell population splits into multiple phenotypically distinct outcomes or commits to divergent cell fates

Solving the Branched Generalized SB Problem with Conditional Stochastic Optimal Control (CondSOC)

$$\min_{\{u_{t,k},g_{t,k}\}_{k=0}^K} \underbrace{\mathbb{E}_{(\mathbf{x}_0,\mathbf{x}_{1,0})\sim\pi_{0,1,0}}\int_0^1\left\{\mathbb{E}_{p_{t|0,1,0}}\left[\frac{1}{2}\|u_{t,0}(X_{t,0})\|^2+V_t(X_{t,0})\right]w_{t,0}\right.}_{\text{primary path}}+\underbrace{\sum_{k=1}^K\mathbb{E}_{(\mathbf{x}_0,\mathbf{x}_{1,k})\sim\pi_{0,1,k}}\int_0^1\mathbb{E}_{p_{t|0,1,k}}\left[\frac{1}{2}\|u_{t,k}(X_{t,k})\|^2+V_t(X_{t,k})\right]w_{t,k}\Big\}dt}_{\text{secondary branches}} \quad \text{s.t.} \begin{cases} dX_{t,k}=u_{t,k}(X_{t,k})dt+\sigma dB_t \\ X_0=\mathbf{x}_0,\ X_{1,k}=\mathbf{x}_{1,k} \\ w_{0,k}=\delta_{k=0},\ w_{1,k}=w_{1,k} \end{cases}$$

Branched CondSOC as the Sum of Unbalanced CondSOC Objectives

Branched Energy Loss Function to Optimize Branched Drift and Growth Networks

$$\mathcal{L}_{\text{energy}}(\theta,\phi)=\int_0^1\mathbb{E}_{\{p_{t,k}\}_{k=0}^K}\Big\{\underbrace{\left[\frac{1}{2}\|u_{t,0}^\theta(\mathbf{x}_{t,0})\|^2+V_t(\mathbf{x}_{t,0})\right]w_{t,0}^\phi}_{\text{primary path}}+\underbrace{\sum_{k=1}^K\left[\frac{1}{2}\|u_{t,k}^\theta(\mathbf{x}_{t,k})\|^2+V_t(\mathbf{x}_{t,k})\right]w_{t,k}^\phi}_{K\text{ secondary branches}}\Big\}dt \quad \text{s.t.}\ w_{t,0}^\phi=1+\int_0^t g_{s,1}^\phi(\mathbf{x}_{s,1})ds,\quad w_{t,k}^\phi=\int_0^t g_{s,k}^\phi(\mathbf{x}_{s,k})ds$$

Multi-Stage Training of BranchSBM

Stage 1: Learn Energy-Minimizing Neural Interpolant

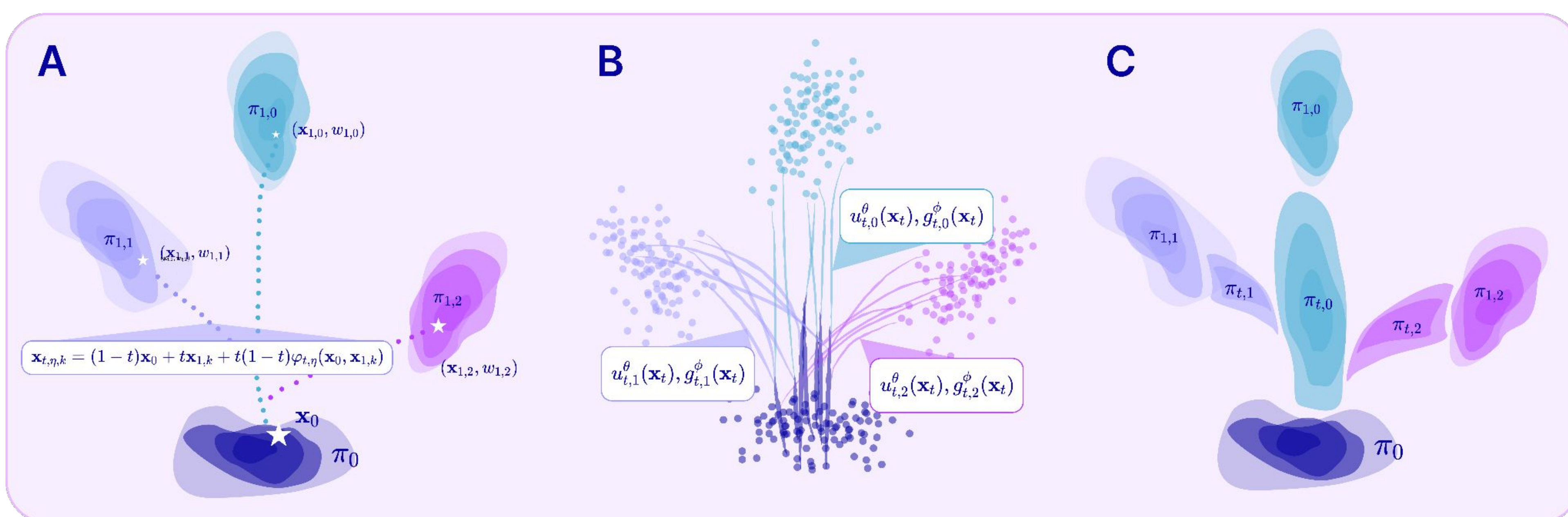
$$\mathbf{x}_{t,\eta,k}=(1-t)\mathbf{x}_0+t\mathbf{x}_{1,k}+t(1-t)\varphi_{t,\eta}(\mathbf{x}_0,\mathbf{x}_{1,k})$$
$$\dot{\mathbf{x}}_{t,\eta,k}=\mathbf{x}_1-\mathbf{x}_0+t(1-t)\dot{\varphi}_{t,\eta}(\mathbf{x}_0,\mathbf{x}_{1,k})+(1-2t)\varphi_{t,\eta}(\mathbf{x}_0,\mathbf{x}_{1,k})$$

Trajectory Loss

$$\mathcal{L}_{\text{traj}}(\eta)=\sum_{k=0}^K\int_0^1\mathbb{E}_{(\mathbf{x}_0,\mathbf{x}_{1,k})\sim\pi_{0,1,k}}\left[\frac{1}{2}\|\dot{\mathbf{x}}_{t,\eta,k}\|_2^2+V_t(\mathbf{x}_{t,\eta,k})\right]dt$$

Stage 2: Match Parameterized Velocity Field

$$\mathcal{L}_{\text{flow}}(\theta)=\sum_{k=0}^K\int_0^1\mathbb{E}_{(\mathbf{x}_0,\mathbf{x}_{1,k})\sim\pi_{0,1,k}}||\dot{\mathbf{x}}_{t,\eta,k}^\star-u_{t,k}^\theta(\mathbf{x}_{t,k})||_2^2dt$$



Stage 3: Learn Optimal Growth Networks

$$\mathcal{L}_{\text{growth}}(\phi)=\mathcal{L}_{\text{energy}}(\theta,\phi)+\mathcal{L}_{\text{match}}(\phi)+\mathcal{L}_{\text{mass}}(\phi)+\sum_{k=0}^K\|g_{t,k}^\phi\|_2^2$$

Match Loss

$$\mathcal{L}_{\text{match}}(\phi)=\sum_{k=0}^K\mathbb{E}_{p_{1,k}}\left(w_{1,k}^\phi(\mathbf{x}_{1,k})-w_{1,k}^\star\right)^2$$

Mass Loss

$$\mathcal{L}_{\text{mass}}(\phi)=\int_0^1\mathbb{E}_{\{p_{t,k}\}_{k=0}^K}\left(\sum_{k=0}^K w_{t,k}^\phi(\mathbf{x}_{t,k})-w_t^{\text{total}}\right)^2dt$$

Stage 4: Joint Training of Velocity and Growth Networks

$$\mathcal{L}_{\text{joint}}(\theta,\phi)=\mathcal{L}_{\text{energy}}(\theta,\phi)+\mathcal{L}_{\text{match}}(\phi)+\mathcal{L}_{\text{mass}}(\phi)+\mathcal{L}_{\text{recons}}(\theta)$$

Multi-Path Navigation of 3D LiDAR Manifolds

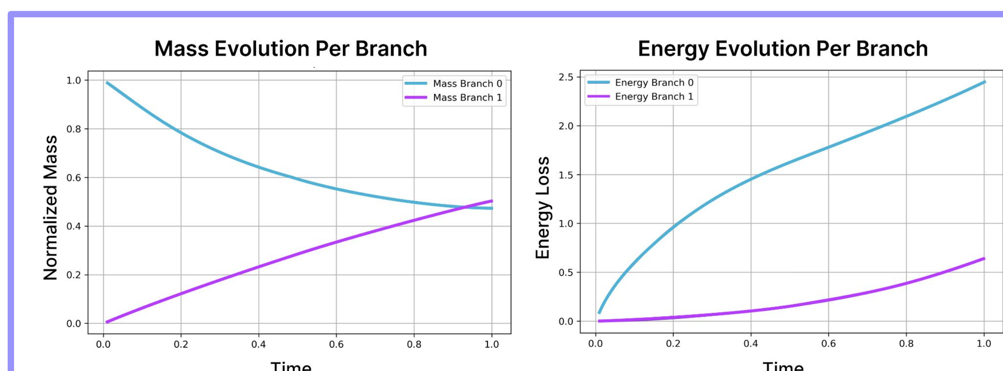


Table 1: Benchmark of BranchSBM against single-branch SBM on multi-path surface navigation. Wasserstein distances (W1 and W2) between the reconstructed and ground-truth distributions with 100 Euler steps at time t = 1 from validation samples in the initial distribution. Results are averaged over 5 independent runs.

Model	$\mathcal{W}_1$ (↓)	$\mathcal{W}_2$ (↓)
Single Branch SBM	$0.975_{\pm 0.009}$	$1.285_{\pm 0.007}$
BranchSBM	$\mathbf{0.239}_{\pm 0.001}$	$\mathbf{0.309}_{\pm 0.003}$

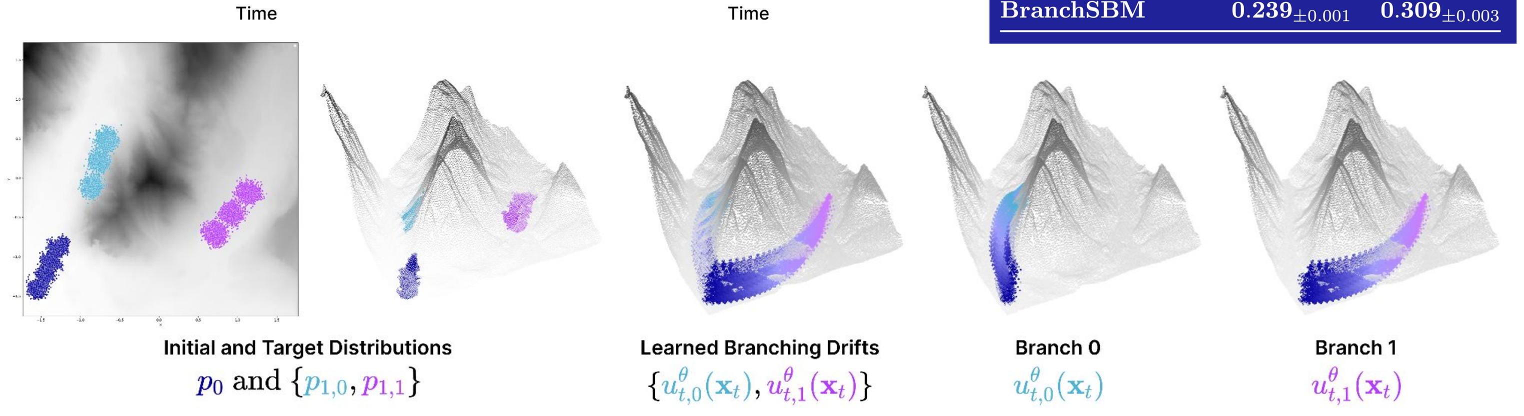


Fig 2. Application of BranchSBM on Learning Branched Paths on a LiDAR Manifold. (Top) Plot of weight (left) and energy (right) evolution of each branch over time. **(Bottom)** Plots of the initial and target distributions, learned interpolants, and learned branched trajectories on the LiDAR manifold.

Modeling Cell Fate Differentiation

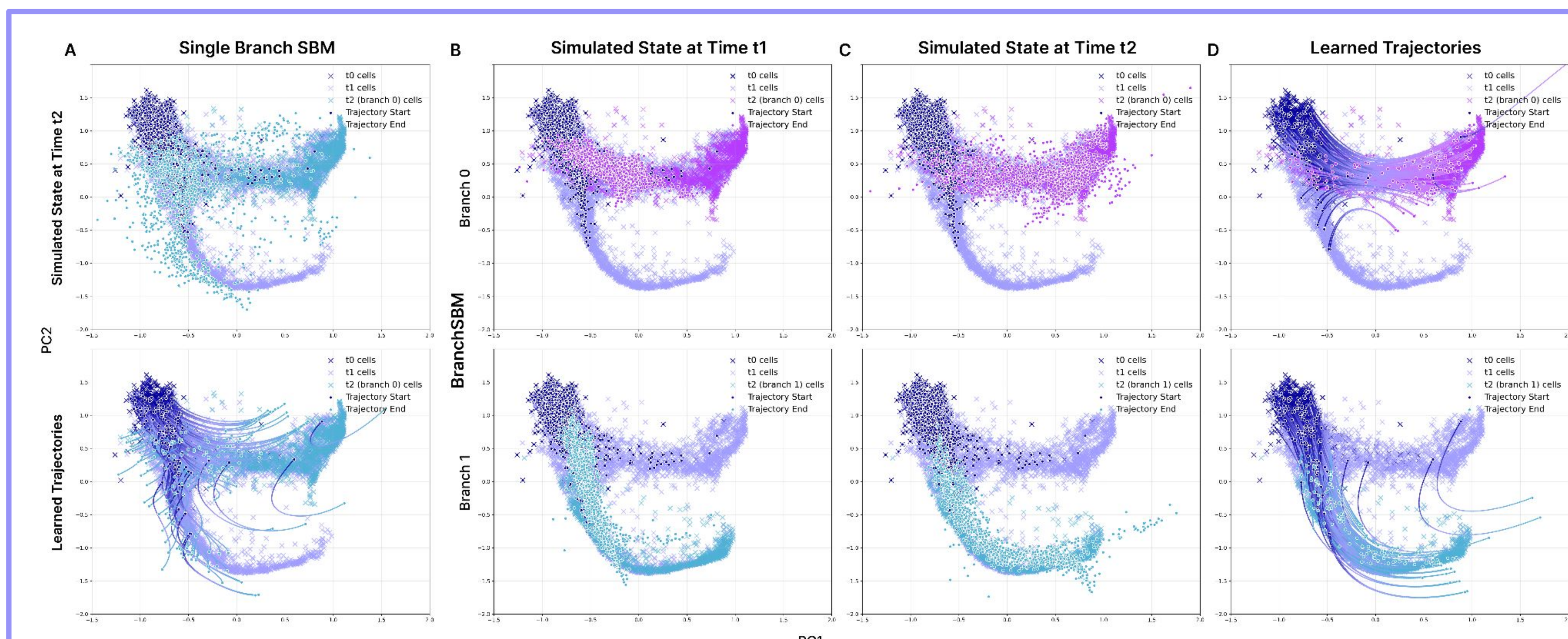


Fig 3. Application of BranchSBM on Modeling Differentiating Single-Cell Population Dynamics. Mouse hematopoiesis scRNA-seq data is provided for three time points. Simulated states (top) and trajectories (bottom) at time using single-branch SBM. Simulated states with BranchSBM at **(B)** t1 and **(C)** t2. **(D)** Learned trajectories on validation samples

Modeling Drug-Induced Perturbation Responses

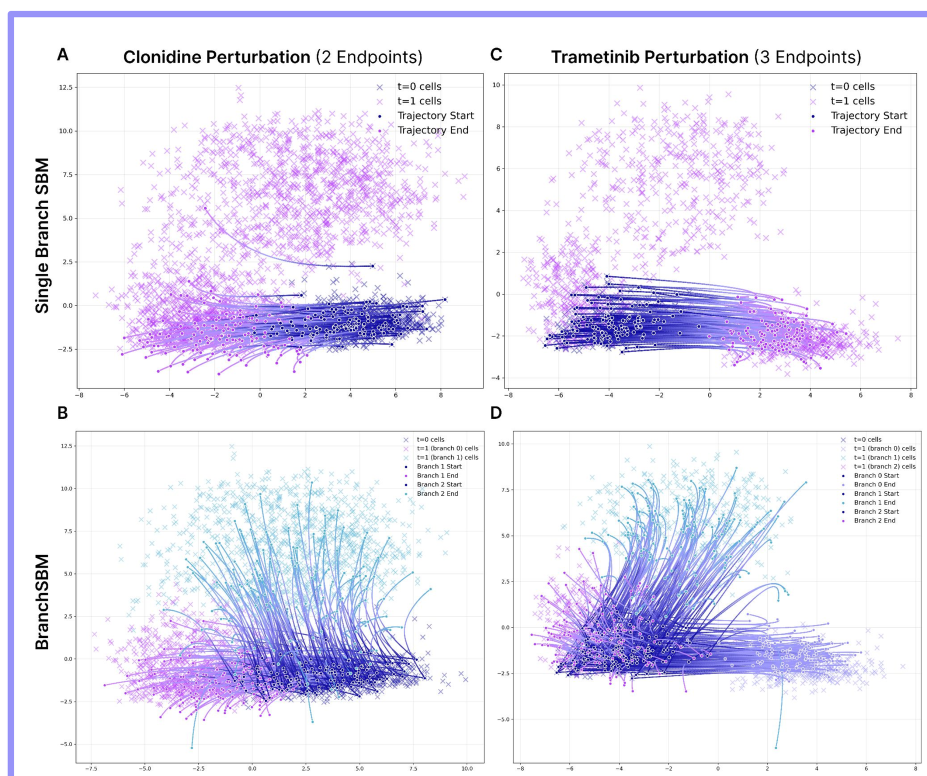


Table 2: Results for Clonidine Perturbation Modeling for Increasing Principal Component Dimensions. Maximum-mean discrepancy (MMD) across all PCs and Wasserstein distances (W1 and W2) of top 2 PCs between ground truth and reconstructed distributions at t = 1 simulated from the validation data at t = 0. Results for single-branch SBM (50 PCs) and BranchSBM (2 branches) were averaged over 5 independent runs.

Model	RBF-MMD (↓)	$\mathcal{W}_1$ (↓)	$\mathcal{W}_2$ (↓)
Single Branch SBM (50 PCs)	$0.279_{\pm 0.024}$	$5.124_{\pm 0.509}$	$6.149_{\pm 0.463}$
BranchSBM			
50 PCs	$0.065_{\pm 0.001}$	$\mathbf{1.076}_{\pm 0.085}$	$\mathbf{1.224}_{\pm 0.097}$
100 PCs	$\mathbf{0.053}_{\pm 0.002}$	$1.832_{\pm 0.174}$	$2.037_{\pm 0.174}$
150 PCs	$0.083_{\pm 0.001}$	$1.722_{\pm 0.064}$	$1.931_{\pm 0.035}$

Table 3: Results for Trametinib Perturbation Modeling

Model	RBF-MMD (↓)	$\mathcal{W}_1$ (↓)	$\mathcal{W}_2$ (↓)
Single Branch SBM	$0.246_{\pm 0.013}$	$5.428_{\pm 0.234}$	$6.426_{\pm 0.186}$
BranchSBM	$\mathbf{0.053}_{\pm 0.001}$	$\mathbf{0.838}_{\pm 0.061}$	$\mathbf{0.973}_{\pm 0.050}$

Fig 4. Comparison Between BranchSBM and Single-Branch SBM. Scaling BranchSBM across 50, 100, and 150 principal components and three branches outperforms single-branch SBM.

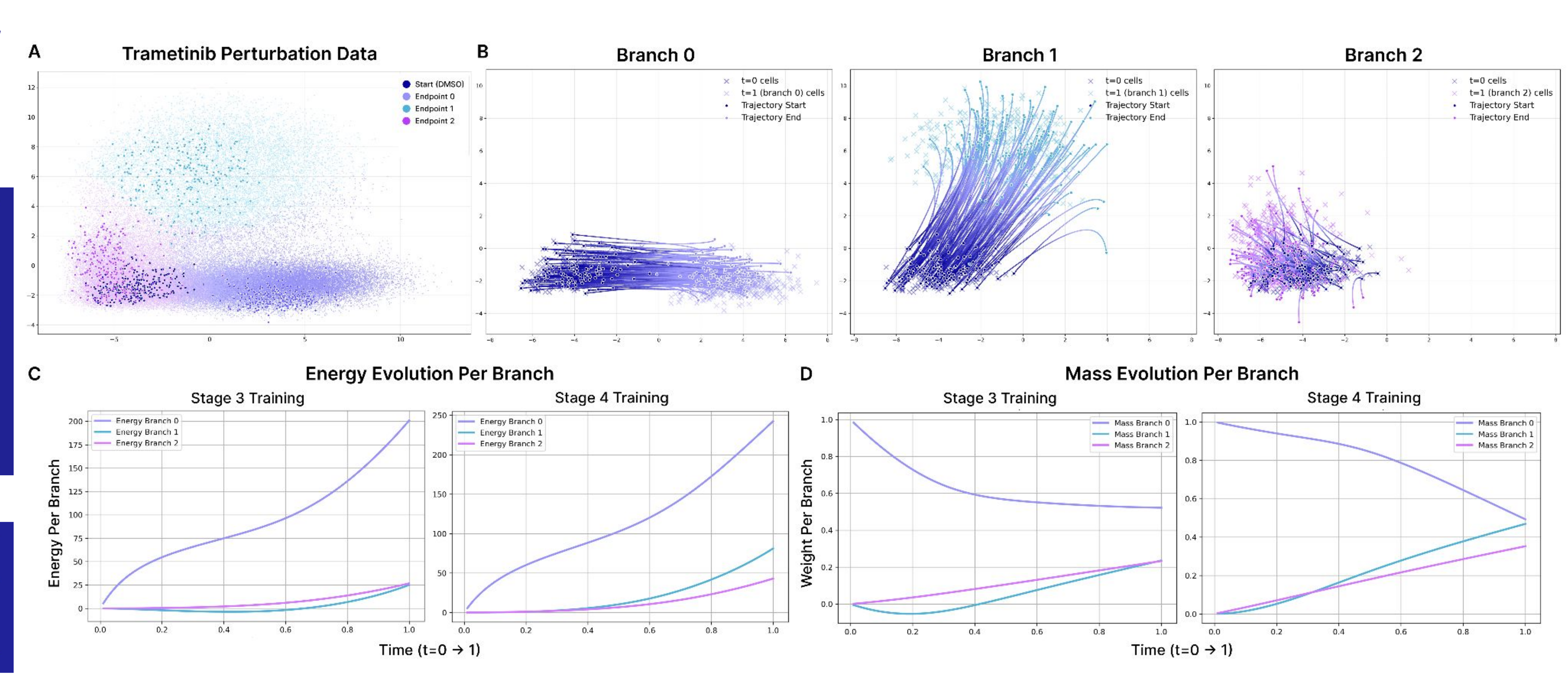


Fig 5. Results for Trametinib Perturbation Modeling with BranchSBM. (A) Clustered Trametinib perturbation data. **(B)** Three branches learned by BranchSBM. **(C)** Evolution of energy and mass.

Conclusions

1. We define the Branched Generalized Schrödinger Bridge problem and introduce **BranchSBM**, a novel matching framework that learns optimal branched trajectories from an initial distribution to multiple target distributions.
2. We derive the **Branched Conditional Stochastic Optimal Control (CondSOC)** problem as the **sum of Unbalanced CondSOC objectives** and leverage a multi-stage training algorithm to learn the optimal branching drift and growth fields that transport mass along a branched trajectory
3. We demonstrate the unique capability of BranchSBM to model dynamic branching trajectories while matching multiple target distributions across various problems, including **3D navigation over LiDAR manifolds**, **modeling differentiating single-cell population dynamics**, and **predicting heterogeneous cell states after perturbation**.

Preprint