



Single-Cell Environment and Proximal Trajectory Inference using Collaborative Agent Reinforcement Learning

Noah Cohen Kalafut^{1,2} and Daifeng Wang^{1,2,3,*}

¹Department of Computer Sciences, Wisconsin, US. ²Waisman Center, University of Wisconsin-Madison, Wisconsin, US. ³Department of Biostatistics and Medical Informatics, Wisconsin, US

Abstract

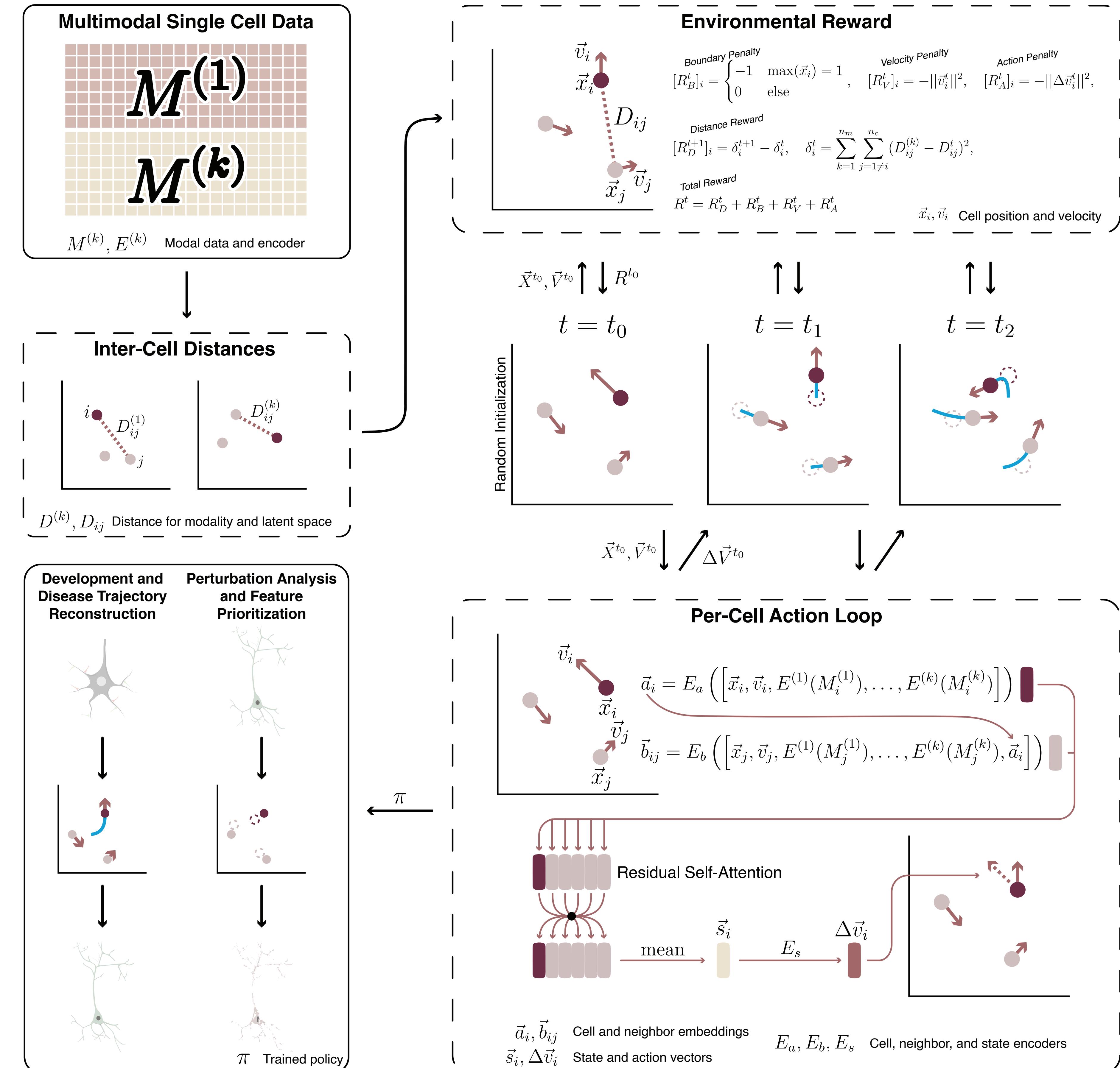
Background: Recent techniques enable functional characterization of single-cells, which allows the study of cellular and molecular mechanisms in complex biological processes including cell development. Many methods have been developed to utilize these single-cell datasets to reveal cell developmental trajectories such as dimensionality reduction and pseudotime. However, these methods generally produce static snapshots of the data, challenging a deeper understanding of the mechanistic dynamics underlying cell development.

Method: To address this, we have developed scEPTIC-RL (single-cell Environment and Proximal Trajectory Inference using Collaborative Reinforcement Learning), a multi-agent reinforcement learning model to recapitulate the dynamic progression of cells during development. scEPTIC-RL takes single-cell data, either single or multimodality, and trains a collaborative reinforcement learning model that governs cell-cell dynamic interactions driving development. Particularly, it models single cells as individual agents which coordinate progression on a latent space through interacting with neighboring cells. The trained model can further prioritize cellular features and in-silico predict the dependencies of cell development from feature perturbations (e.g., gene knockdown). We apply scEPTIC-RL to both simulation and real-word single-cell multiomics datasets including brain development and cancers, revealing potential novel mechanistic insights on gene expression and regulation in those complex developmental processes.

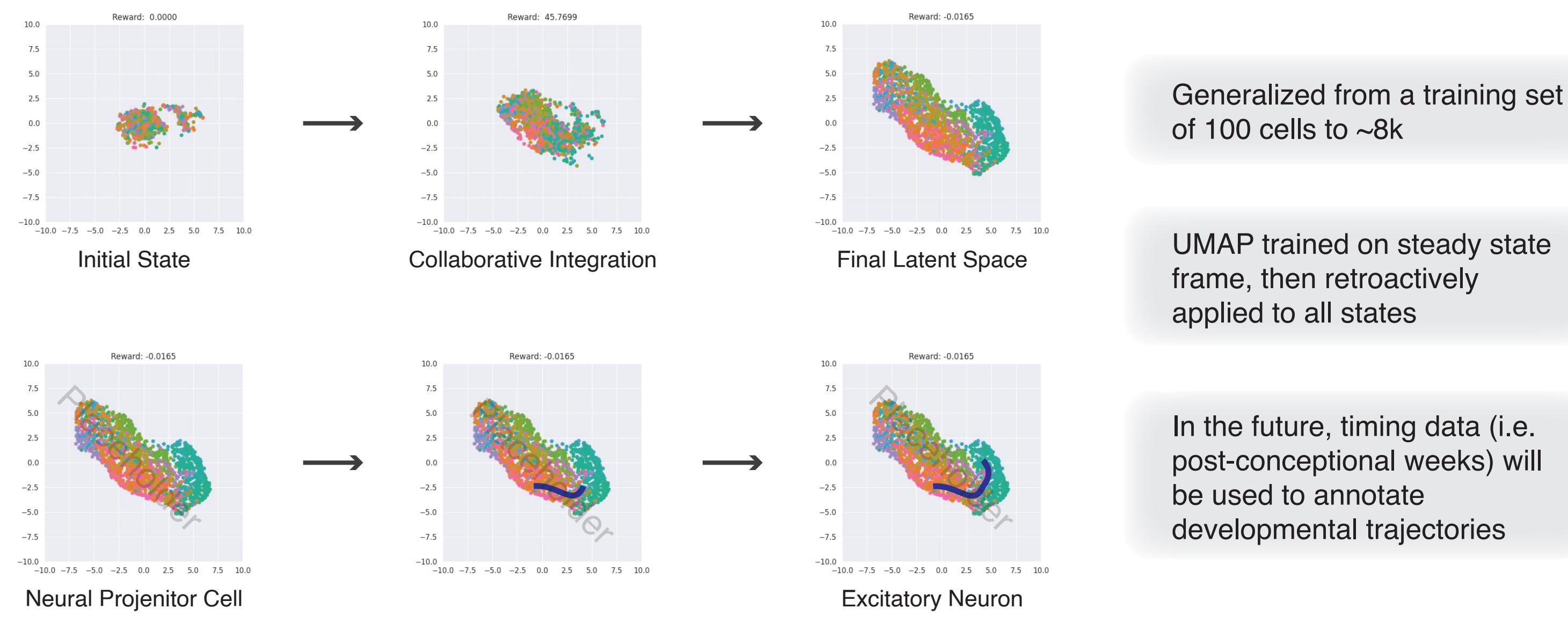
Datasets

Citation	Dataset Description	Cells	Features RNA-seq	Features ATAC-seq
Liu et al. [1]	Simulation data consisting of 3 distinct cell types	300	1,000	2,000
Trevino et al. [2]	Developing human brain data over 21 post-conceptual weeks	8,981	34,104	19,836
Cao and Gao [3]	Single-nucleus profiling of the adult mouse cortex	9,190	28,930	241,757

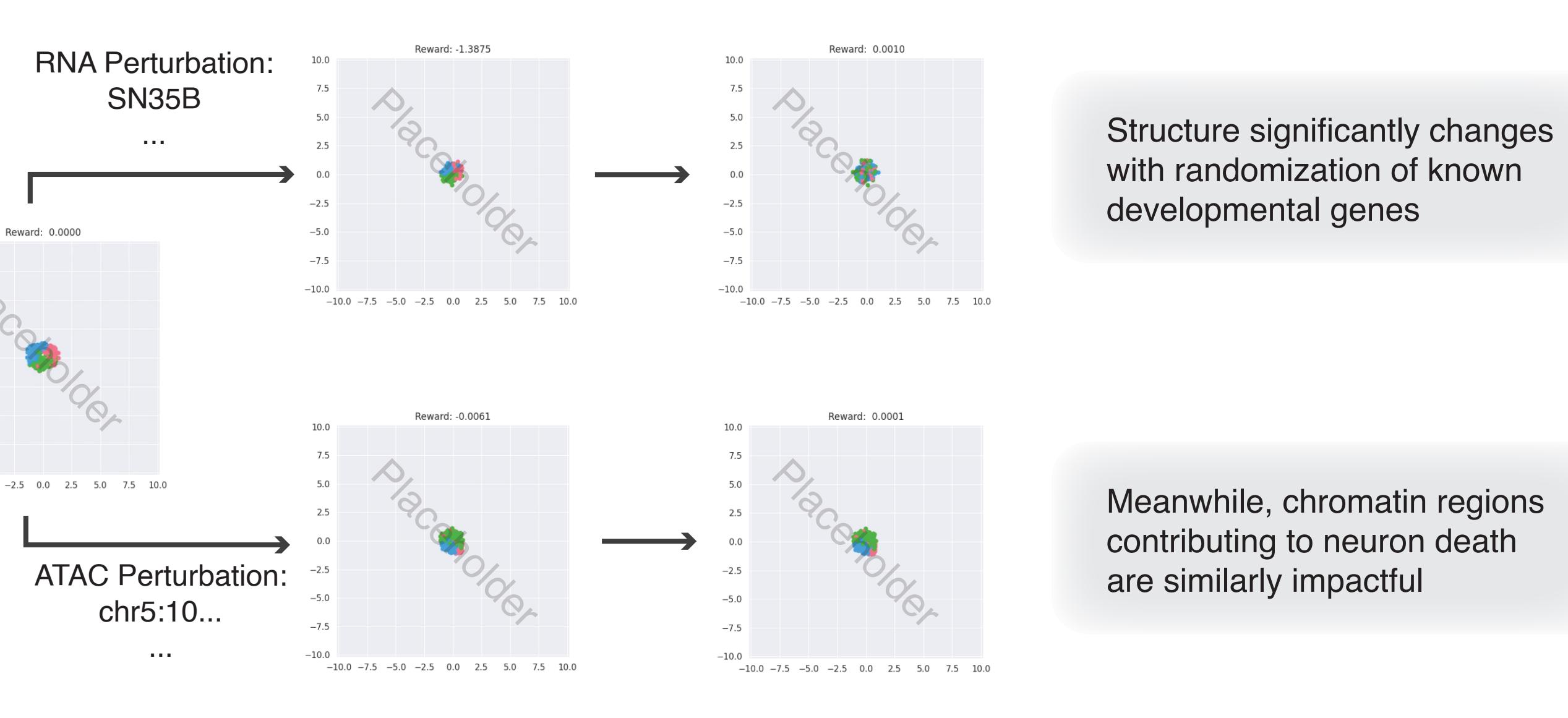
scEPTIC-RL Model Usage and Architecture



Recovery of Brain Developmental Trajectories



Feature Prioritization in the Mouse Cortex



Algorithm

Algorithm 1 scEPTIC-RL

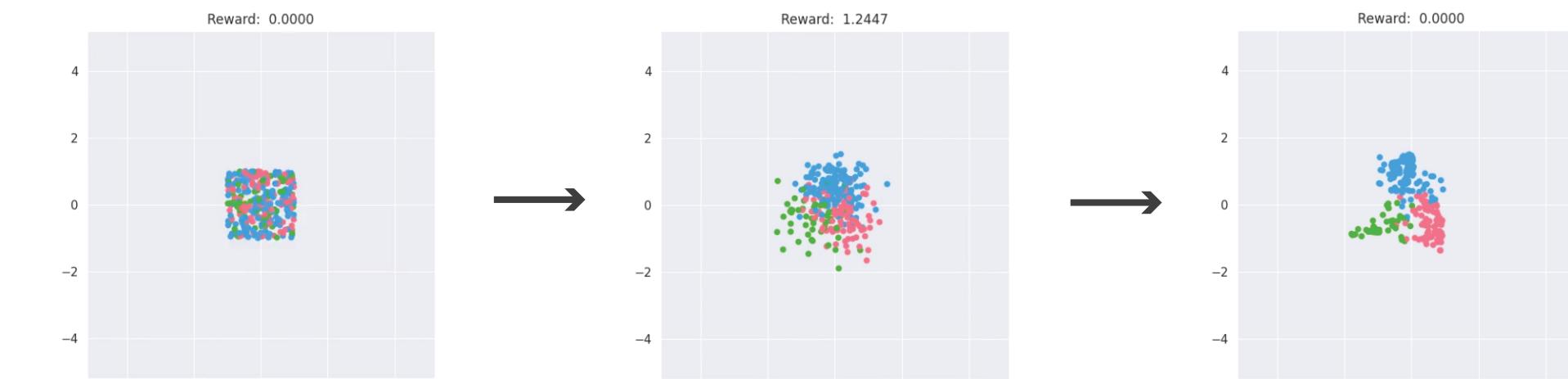
```

Input:  $M^{(1)}, M^{(2)}, \dots, M^{(n_m)}$ 
Parameters: #max.timesteps, #update.timesteps, #epochs,  $\gamma$ 
Output:  $\pi_\theta$ 
1: Randomly initialize environment  $e$  and model  $\pi_\theta$ 
2:  $t \leftarrow 1$ 
3: for #max.timesteps do
4:    $s^t \leftarrow e$ 's state
5:    $a^t \sim \pi_\theta(s^t)$ 
6:    $R^t, \text{terminal} \leftarrow$  perform actions  $a^t$  on environment  $e$ 
7:   Record  $(s^t, a^t, R^t)$  into memory
8:   if #update.timesteps|t then
9:     Compute episode rewards using  $\gamma$  decay
10:    for #epochs do
11:      Compute  $L_{\text{CLIP}}(\theta)$  using PPO
12:      Backpropagate on  $\pi_\theta$  from memory
13:    end for
14:   end if
15:   if terminal then
16:     Reset environment
17:   end if
18:    $t \leftarrow t + 1$ 
19: end for

```

$$L_{\text{CLIP}}(\theta) = \text{MIN} \left(\frac{\pi_\theta(a^t | s^t)}{\pi_{\theta_{\text{old}}}(a^t | s^t)} \hat{A}^t, \text{CLIP} \left(\frac{\pi_\theta(a^t | s^t)}{\pi_{\theta_{\text{old}}}(a^t | s^t)}, 1 - \epsilon, 1 + \epsilon \right) \hat{A}^t \right)$$

Simulation Data



scEPTIC-RL is able to effectively generalize from a 50 cell training set to 300 cells and reconstruct a cell type-separated latent space with no annotation data provided.

Conclusions

- ★ Effectively separates cells by phenotype, including cell type
- ★ Reconstructs cell developmental and disease progression trajectories through reliable generalization
- ★ Prioritizes genes and cell features important to development and phenotype

Overall, scEPTIC-RL provides reliable trajectory reconstructions for cell progression and disease development in an easily distributable and highly generalizable manner. Future applications include extension to further datasets as well as data imputation.