SPADE: Inferring Transcriptional Dynamics from Spatial Transcriptomics with Physics-Informed Deep Learning

Anonymous Author(s)
Affiliation
Address
email

Abstract

In situ sequencing–based spatial transcriptomics technologies, such as 10x Genomics Xenium and Vizgen MERSCOPE, have recently emerged as powerful platforms that enable subcellular-resolution mapping of RNA transcripts within intact tissues. While existing computational models developed for pixel-based spatial transcriptomics can be applied to in situ sequencing data, these approaches overlook molecule-level information and thus underutilize the full potential of the high-resolution measurements. Recognizing that post-transcriptional mRNA localization arises from a hybrid process of active transport and diffusion, we hypothesized that the spatial distribution of transcripts relative to the transcription start site encodes information about transcriptional activity within short time windows, offering a new paradigm for inferring transcriptional dynamics. To realize this capability, we present SPADE, a physics- and systems biology–informed deep learning framework that leverages the spatial organization of RNA molecules to infer transcriptional dynamics. SPADE first constructs a trajectory for each cell, ordered along a pseudo-time axis defined by local shifts in molecule distributions, and then employs a recurrent neural network to disentangle RNA synthesis from drift–diffusion processes under a bistate transcriptional regulation model. Extensive evaluations on both simulated and in-house spatial transcriptomics datasets demonstrate that SPADE accurately reveals gene-specific bursting patterns, recovers dynamic transcription rates, and uncovers regulatory delays between genes. As the first framework to estimate temporal variations in transcription rates from static spatial transcript distributions, SPADE establishes a novel paradigm for studying transcriptional dynamics and their underlying biological mechanisms.

1 Introduction

Recent in-situ sequencing-based spatially resolved transcriptomics (SRT) technologies [14, 15, 16], such as 10x Genomics Xenium and Vizgen MERSCOPE, measure the exact locations of mRNA molecules that enable subcellular or even higher resolution quantification of their abundance. By measuring molecule-wise spatial coordinates of transcripts, these platforms provide unprecedented information to study the dynamics of transcription and intracellular transport of mRNA molecules.

Despite this promise, existing computational approaches for SRT data were mainly designed for pixel-based SRT data [3, 11, 12], which quantify gene expression at the cellular level and tend to overlook the fine-grained spatial distributions of individual molecules. The intracellular distribution of mRNA is shaped by three processes: synthesis at transcription start sites, active transport, and diffusion throughout the cytoplasm[6, 1, 2]. These mechanisms jointly determine the number of transcripts and where these transcripts are observed within a cell. Thus, the spatial coordinates of

Submitted to 39th Conference on Neural Information Processing Systems (NeurIPS 2025). Do not distribute.

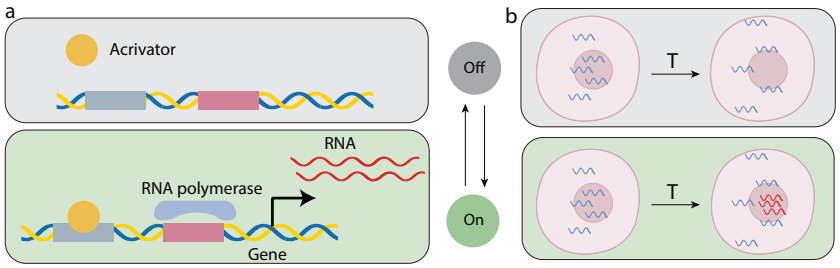

Figure 1: Bi-State transcription model. a) on/off state of transcription. Top: off state, no new transcripts generated; Bottom: on state, RNA molecules generated at the transcription start site. b) RNA dynamics under different states. Top: drift-diffusion of existing molecules from nucleus to cytosol; Bottom: on state, besides the movement of existing molecules, new molecules are generated and will undergo the same diffusion process.

mRNA molecules, ranging from those distributed near the transcriptional start site, reflecting newly synthesized transcripts, to those dispersed in the cytosol, representing transcripts synthesized earlier, provide rich information to study the dynamics of transcription and the subsequent motion of mRNAs.

The spatial distribution of mRNA molecules arises from the combined effects of transcription, active transport, and random diffusion. Consequently, the distance of molecules from the transcriptional start site encodes the integration of these processes, allowing spatial patterns to serve as a short-term temporal proxy of transcriptional dynamics. In this study, building on the dynamic model of mRNA transcription and motion [20], we developed a disentangled deep learning framework that separates the contributions of transcription and active transport, enabling inference of transcriptional rate changes over short time intervals. Unlike existing approaches such as RNA velocity, which estimate transcriptional activity from static omics data at the moment of measurement, our method leverages spatial distributions to predict short-range transcriptional dynamics, thereby capturing regulatory delays and establishing a new paradigm for modeling transcriptional processes.

Based on these ideas, we developed **SPA**tial **D**ynamics inf**E**rence (**SPADE**), a physics- and systems biology-informed deep learning framework to model transcriptional dynamics using in-situ sequencing based SRT data. For each cell, SPADE first infers a pseudo-time trajectory of cells within its local neighborhood by quantifying the shifts of transcript distribution using a Kolmogorov-Smirnov statistic [13]. These time-series sequences are then input to a long-short term memory[8] (LSTM)-based neural network to learn the transition functions between consecutive states. By the systems biology of transcription dynamics, each transition is formed of two components, a transcription *generation module* that models the biosynthesis of mRNA, and a *diffusion module* that models the movements of mRNA molecule. By implementing biological and physical principles into the network architecture, SPADE is able to explicitly learn the underlying factors that drive the transcriptional dynamics.

We evaluate SPADE on both simulated datasets and in-house collected in situ sequencing data. Experimental results demonstrate that SPADE accurately recovers dynamic transcription rates, identifies gene-specific bursting patterns, and uncovers regulatory delays between genes. To the best of our knowledge, SPADE is the first framework to infer temporal variations in transcriptional rates from single snapshots of spatial observations. By integrating biophysical principals with disentangled deep learning, SPADE established a new paradigm for studying the dynamical behaviors of transcription. We believe it brings valuable insights for revealing dynamic biological processes from high-resolution spatial data.

2 Background

2.1 RNA dynamics in cells

The observed abundance and spatial localization of RNA molecules are resulted from three processes. *Biosynthesis* generates new mRNA transcripts at transcription start sites, regulated by transcription factors or epigenetic regulators. The newly generated molecules will then be transported from the nucleus to the cytosol, ER, or other subcellular compartments, driven by both *active transport* and

diffusion. The molecules will undergo this drift-diffusion process throughout the cytoplasm until degradation.

A commonly used systems biology model of transcription is the Bi-State model[18, 5, 9] (see Figure 1), where genes switch between "off" and "on" states. In the "off" state, no new molecule will be produced, and the spatial distribution of the transcripts only reflects drift–diffusion of already transcribed molecules. In the "on" state, synthesis and diffusion act jointly. The spatial distributions are shaped by both the synthesis of new mRNAs and the diffusion of newly synthesized and existing mRNAs. Noted, this model captures most patterns of transcriptional dynamics such as cis-transcriptional regulation and transcriptional bursting, which links the spatial distribution of mRNA molecules with their underlying transcription mechanisms.

The biological hypothesis of our model is that the snapshots of transcript spatial distributions encode temporal transcriptional signals: transcripts near the nucleus often correspond to recent generations, while more dispersed molecules are more likely to be related to older events.

2.2 Task definition and formulation

We first introduce the notations and features used in our model. Because transcriptional dynamics are gene-specific, SPADE models each gene independently by using the spatial distribution of its mRNA molecules within each cell. To train a deep disentangling model, we first derive a feature vector from the spatial distribution of transcripts within each cell. For each cell, we discretize the radial distance from the nuclear center to the cell boundary into a series of intervals. Denote the nuclear center as d_0, and the width of each interval as Δd, the boundary of the j-th interval can be defined as:

$$d_j = d_0 + j \cdot \Delta d \tag{1}$$

For each gene, denote the feature vector of a cell as $\mathbf{x}$. Its j-th entry $\mathbf{x}_j$ is the number of mRNA molecules of the gene located between d_{j-1} and d_j. By this definition, the intracellular spatial distribution of the mRNA will be encoded by $\mathbf{x}$.

Based on our biological assumptions, the feature vector $\mathbf{x}_{t_i}$ at any time t_i can be denoted as the summation of two components:

$$\mathbf{x}^{t_i} = \mathbf{m}^{t_i} + \mathbf{g}^{t_i} \tag{2}$$

, where $\mathbf{m}$ denotes the contribution of the drift-diffusion process of the existing molecules and $\mathbf{g}$ represents the contribution of the newly generated transcripts.

By this definition, the dynamics of the system can be represented by how the feature vector evolves between consecutive time points $\mathbf{x}_{t_i}$ to $\mathbf{x}_{t_{i+1}}$:

$$\mathbf{x}_{t_{i+1}} = f(\mathbf{x}_{t_i}) \tag{3}$$

, where f is the transition function that encodes the transcription dynamics following the bistate transcription model and motion of mRNA molecules. The objective of our model is to learn f from the observed spatial patterns of the transcripts.

3 Method

3.1 Overview of SPADE

SPADE is a systems biology-informed deep learning framework designed to infer transcriptional dynamics using in-situ sequencing-based SRT data (see Figure 2). SPADE is based on two computational assumptions: 1) the transcriptional dynamics of one gene in an individual cell can be reconstructed from snapshots of the gene's expression pattern in multiple cells at different states, and 2) the spatial distribution of mRNA molecules can recapitulate the processes of its synthesis and movements. The input of SPADE is the spatial distribution of mRNA molecules, the boundary, and the transcription start site of each cell. The output is the transcription rate in a small time interval for each gene in each cell.

For each cell and each gene, SPADE first infers a pseudo-time trajectory by aligning neighboring cells, achieved by quantifying shifts in the spatial distribution of mRNA molecules across cells (Figure 2a). An LSTM-based recurrent neural network is then trained to model transitions between consecutive states over the trajectory (Figure 2b). Guided by the Bi-State transcription model, each

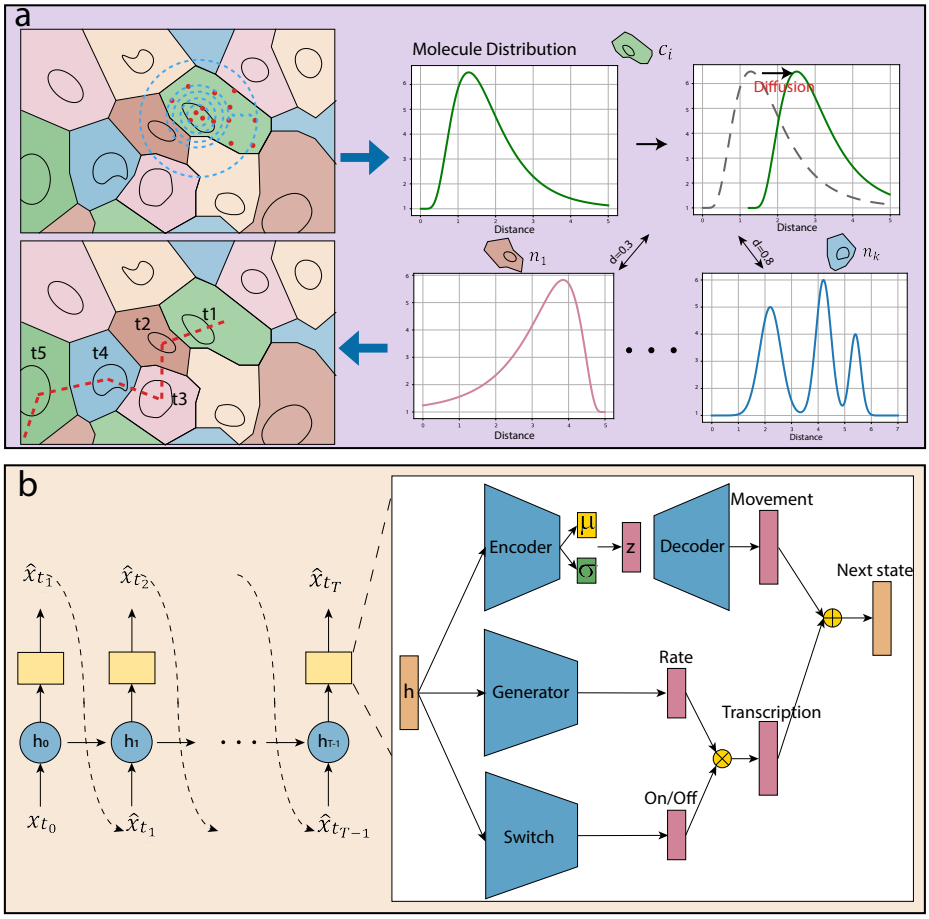

Figure 2: SPADE framework. a) Construction of pseudo-time trajectories. top left: molecules in the starting cell; right: starting from cell c_i, estimate possible next state after diffusion, and compare with neighbors to derive transition probabilities; bottom left: random walk from the starting cell. b) Fitting the trajectories with a sequencing model, disentangling each state transition into molecule movement and generation components using systems biology constraints.

transition is decomposed into a *generation module* that captures mRNA synthesis and a *movement module* that models drift–diffusion of existing molecules. By implementing systems biology model into its neural network architecture, SPADE explicitly disentangles the observed spatial distribution of mRNA molecules by their synthesis and movement.

3.2 Inference of pseudo-time trajectory in the neighborhood of each cell

A central challenge in learning transcriptional dynamics from omics data is the absence of true time-series measurements. Prior studies have shown that snapshots of cells at different stages can be aligned to approximate temporal progression [4, 10]. Building on this concept, we hypothesize that cells with similar spatial distributions of an mRNA molecule can be aligned along a pseudo-time trajectory. SPADE implements this idea by constructing a pseudo-trajectory for each gene and each cell, by aligning the spatial distribution of the mRNA molecule in the cells within a spatial neighborhood, and fitting them with a deep disentangling model to reconstruct transcription rate, transport, and diffusion dynamics from the nucleus to the cytosol of the mRNA in each cell.

To construct the pseudo-time trajectory, for each cell c_i and a given gene, we first identify its k-nearest neighbors $n_1, n_2, \cdots, n_k \in N(c_i)$ by the spatial coordinates of the cell centers. We then estimate the empirical cumulative distribution function (CDF) $\hat{F}(c_i)$ of the transcripts of the gene and simulate

possible next states $\{\hat{F}_s(c_i)\}_{s=1}^{3}$ by diffusing the molecules for 1–3 steps. For each neighbor n_j, we also compute the CDF and then derive a dissimilarity score by employing the Kolmogorov–Smirnov (KS) distance:

$$d_{ij} = \min_{s=1,2,3} KS\big(\hat{F}(n_j), \hat{F}_s(c_i)\big). \tag{4}$$

A smaller d_{ij} indicates a higher likelihood that c_i transitions to n_j. Transition probabilities are then computed via a softmax transformation:

$$p_{ij} = \frac{\exp(-\alpha \cdot d_{ij})}{\sum_{l=1}^{k} \exp(-\alpha \cdot d_{il})}. \tag{5}$$

, where α is a hyperparameter that controls the sharpness of the transition probabilities. Using these probabilities, we perform random walks of length T from N randomly chosen starting cells, producing an $N \times T$ collection of pseudo-trajectories. Each row corresponds to the approximate time-series of a single cell and serves as training input for SPADE. As illustrated in Figure 2a, local transcript distributions are compared via KS distance, and the resulting transition matrix is used to sample candidate trajectories.

3.3 Modeling Transitions

SPADE models the transition between consecutive states $\mathbf{x}^{t_i} \mapsto \mathbf{x}^{t_{i+1}}$ by decomposing it into two components: (1) the drift–diffusion of existing molecules and (2) the generation of new molecules. An LSTM-based recurrent neural network parameterizes both modules, embedding historical information in the hidden state h_{t_i} (see Figure 2b).

Drift–diffusion of existing molecules. We describe the dynamics of $\mathbf{m}$ using a drift–diffusion process:

$$d\mathbf{m}(t) = \boldsymbol{\mu}(\mathbf{m}(t))\, dt + \boldsymbol{\sigma}(\mathbf{m}(t))\, dB_t, \tag{6}$$

, where $\boldsymbol{\mu}$ corresponds to *active transport* (constant velocity) and $\boldsymbol{\sigma}$ accounts for *diffusion* (stochastic Brownian motion). Using Euler's approximation, the update after a time interval Δt would be:

$$\mathbf{m}(t + \Delta t) = \mathbf{m}(t) + \boldsymbol{\mu}(\mathbf{m}(t))\Delta t + \boldsymbol{\sigma}(\mathbf{m}(t))B_t. \tag{7}$$

In our framework, this process will be embedded into the LSTM latent space. Given a feature sequence $\{\mathbf{x}^{t_0}, \ldots, \mathbf{x}^{t_T}\}$ derived from the trajectories in last section, the hidden state is updated as:

$$h^{t_i} = LSTMCell(h^{t_{i-1}}, \mathbf{x}^{t_i}). \tag{8}$$

From the hidden state, we employ a variational auto-encoder (VAE) inspired encoder-decoder module to learn the molecule movement process during the time interval between t_i and t_{i+1}. Specifically, the latent representation of the drift and diffusion components are embedded via encoder networks μ_θ and σ_θ:

$$z^{t_i} \sim \mathcal{N}(\mu_\theta(h^{t_i}), \sigma_\theta(h^{t_i})), \tag{9}$$

$$\hat{\mathbf{m}}^{t_{i+1}} = \mathbf{x}^{t_i} + f_M(z^{t_i}), \tag{10}$$

, where f_M is a decoder mapping the latent variable to the original feature space. This allows the output state at the current time step to approximate the molecule movement component of the next time step.

Generation of new molecules. New transcripts are assumed to appear near the nuclear center within a certain time, making $\mathbf{g}^{t_{i+1}}$ a vector with a single non-zero entry at the first dimension. Following our bistate transcription model, two variables will be involved in this process: (1) the on/off state of the transcription, denoted by a Bernoulli variable α^{t_i}, and (2) the transcription rate at the on state, denoted as v^{t_i}. Based on the hidden state, a *switch* module is employed to learn the on/off state, while a *generation* module will be used to infer the transcription rate:

$$\alpha^{t_i} \sim \text{Bernoulli}(f_{G_1}(h^{t_i})), \tag{11}$$

$$v^{t_i} = \text{ReLU}(f_{G_2}(h^{t_i})), \tag{12}$$

$$\hat{\mathbf{g}}^{t_{i+1}} = \alpha^{t_i} \cdot v^{t_i} \cdot \mathbf{e}_1, \tag{13}$$

, where f_{G_1} and f_{G_2} are fully connected networks denoting the *switch* module and the *generation* module, respectively, $\mathbf{e}_1$ is a one-hot vector indicating the nuclear compartment, and ReLU ensures positivity of the transcription rate.

Transition. The overall state update is then given by the summation of these two components:

$$\hat{\mathbf{x}}^{t_{i+1}} = \hat{\mathbf{m}}^{t_{i+1}} + \hat{\mathbf{g}}^{t_{i+1}}. \tag{14}$$

Location-aware learning. Transcriptional activity of the same gene can vary across tissue regions due to spatial heterogeneity, arising from differences in cell type composition, differentiation states, and microenvironmental factors[19]. To account for this, SPADE incorporates a spatial embedding module that conditions the initial hidden state on the cell's spatial coordinates. Since the LSTM propagates information recursively through hidden states, the initial state h_{t_0} will influence subsequent dynamics. Rather than a random initialization, we map the spatial coordinates (C_x, C_y, C_z) into the hidden space via a fully connected embedding layer:

$$h^{t_0} = f_s(C_x, C_y, C_z), \tag{15}$$

, where f_s denotes the embedding function. By conditioning h^{t_0} on spatial locations, SPADE learns location-specific patterns of the transcriptional dynamics, enabling it to better capture spatial heterogeneity within tissues.

Training objective. The overall loss function combines a reconstruction term with regularization informed by the bistate transcription model. Given a sequence with T time steps, the reconstruction loss penalizes the difference between the predicted and observed feature vectors:

$$L_{\text{recon}} = \frac{1}{T} \sum_{i=1}^{T} \|\hat{\mathbf{x}}^{t_i} - \mathbf{x}^{t_i}\|_2^2, \tag{16}$$

, where each prediction is decomposed as $\hat{\mathbf{x}}^{t_i} = \hat{\mathbf{m}}^{t_i} + \hat{\mathbf{g}}^{t_i}$.

Since $\hat{\mathbf{x}}^{t_i}$ is the sum of two components, training only with L_{recon} may lead to trivial solutions. To address this, we introduce a regularization loss on the Bernoulli transcription state α^{t_i}:

$$L_{\text{reg}} = \frac{1}{T} \sum_{i=1}^{T} \left(\alpha^{t_i} + \beta \left|\alpha^{t_i} - \alpha^{t_{i-1}}\right|\right), \tag{17}$$

, where the first term controls the total fraction of time transcription is "on", and the second penalizes excessive on/off switching to enforce temporal smoothness.

The final objective is a weighted combination of the two:

$$L = L_{\text{recon}} + \lambda L_{\text{reg}}, \tag{18}$$

, where λ balances reconstruction accuracy with biological plausibility. After training, the inferred transcription rates v^{t_i} provide dynamic estimates of gene activity in each local neighborhood.

4 Experiments

We evaluate SPADE on both synthetic datasets and in-house generated Xenium SRT data. The synthetic data-based experiments allow quantitative assessment of the framework's capability and accuracy in recovering known transcription rate and mRNA movement, while the real data-based experiments demonstrate the biological interpretability and discoveries enabled by SPADE. Our evaluation focuses on three aspects: (1) accuracy in reconstructing transcriptional rates and changes, such as transcription bursting, (2) ability to disentangle the synthesis and movement of mRNA in explaining the spatial distribution of the molecules, and (3) biological insights such as delayed transcriptional regulatory relations.

4.1 Synthetic data and evaluation metrics

To systematically evaluate SPADE, we simulated the transcription using the Bi-State model and diffusion of mRNA molecules using Fick's law of diffusion within individual cells under different kinetic parameters. Each scenario varied in transcriptional on/off states, transcription rates, and diffusion and transport velocities, thereby creating diverse dynamic patterns. The simulated data provides ground-truth dynamic parameters, enabling quantitative evaluation of SPADE's ability to recover transcriptional dynamics. Detailed simulation design and parameter settings are available in Appendix A.

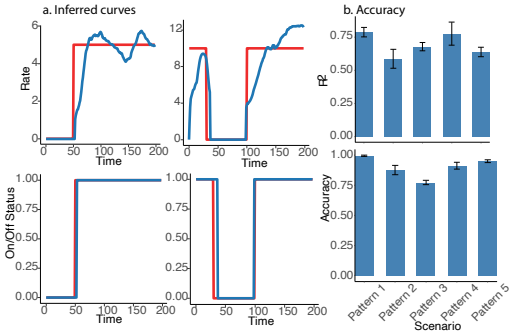

Figure 3: Simulated experiments. Top: transcription rate (red = ground truth, blue = SPADE prediction). Bottom: on/off state.

We assess the performance of SPADE using two metrics: (1) the accuracy of on/off state inferred at each time step, and (2) the coefficient of determination (R^2) between predicted and ground-truth transcription rates. As shown in Figure 3, SPADE accurately captures both the shifts between on and off states and the continuous transcription rates across all the scenarios. The predicted on/off states closely matched ground truth, and transcription rate estimates achieved R^2 scores above 0.6 across all settings, with higher performance in simpler kinetic setups. These results demonstrate that SPADE can faithfully disentangle the generation and movement processes from static spatial distributions of the observed mRNA molecules. Results on all other scenarios are available in Appendix B Figure 7

4.2 Trajectory

Constructing pseudo-trajectories for each cell–gene pair is a key step in SPADE, as the sampled sequences provide the training data for dynamics inference. To validate that these pseudo-trajectories capture meaningful sequential patterns, we visualized sampled trajectories from real spatial transcriptomics data. Figure 4 shows an example trajectory at a local spatial region (illustrated by red arrows), and Figure 5 visualizes both the spatial distributions of the mRNA molecules in each cell and the density of mRNA molecules with respect to the distance to the transcription start site.

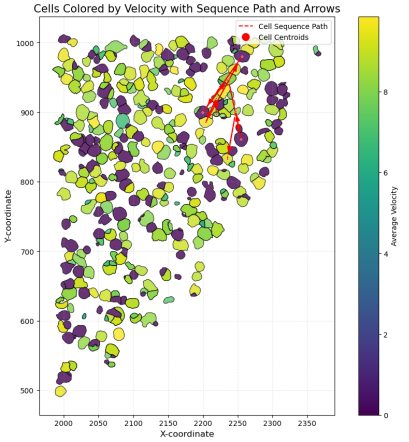

Figure 4: Example pseudo-trajectory at a local region.

Examining the molecule-level patterns along the trajectory (Figure 5 top) shows that successive time steps recapitulate the expected outward diffusion of transcripts from the nucleus. Consistently, the radial density plots (Figure 5 bottom) revealed smooth shifts in transcript distributions, confirming that the sampled trajectories capture biologically interpretable temporal progressions. More visualizations are available in Figure 9.

These results demonstrated that our diffusion-based random-walk method effectively leverages local neighborhood information to construct pseudo-time series while preserving coherent spatiotemporal patterns, thus providing reliable input sequences for model training.

4.3 SPADE reveals delay effect patterns

In cis-transcriptional regulation, transcription factors (TFs) activate or suppress downstream targets with an inherent time delay because of the time required for their own transcription and translation [7]. The transcription rate in a small time interval inferred by SPADE first time enables the direct detection of delay effects in cis-regulation using static data.

We evaluated known TF–target pairs by testing the lagged dependencies between their inferred transcription rate with Granger causality test[17]. As shown in Figure 6b, SPADE captures clear delay patterns in which TF activation curves lead to the corresponding target gene responses (see

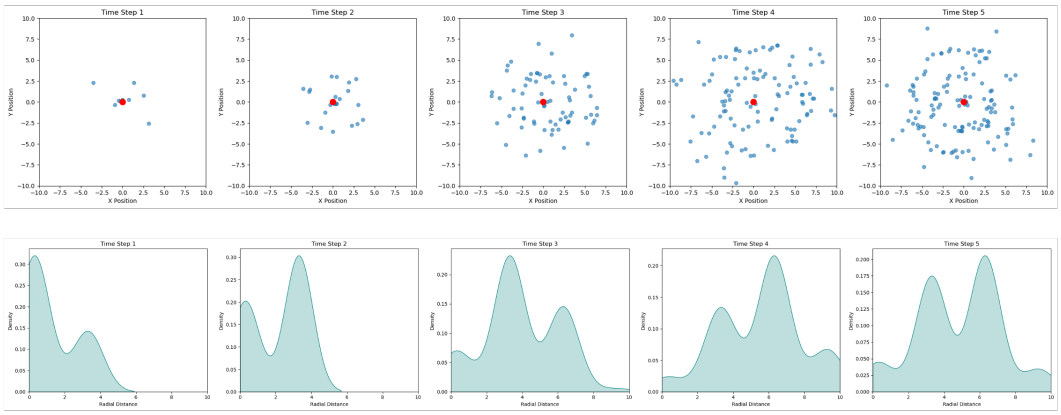

Figure 5: Intracellular RNA molecule distributions (top) and corresponding density plots (bottom) of the selected trajectory.

more examples in Figure 8). In addition, we aggregated the results of each TF-target pair across all cells. The boxplots of p-values of the Granger causality (Figure 6a) suggested that the known TF–target pairs consistently exhibit stronger causal signals than randomly sampled TF-target pairs.

These findings indicate that transcription rates inferred by SPADE over short time intervals can capture regulatory delays, thereby providing evidence of causal relationships between TFs and their targets.

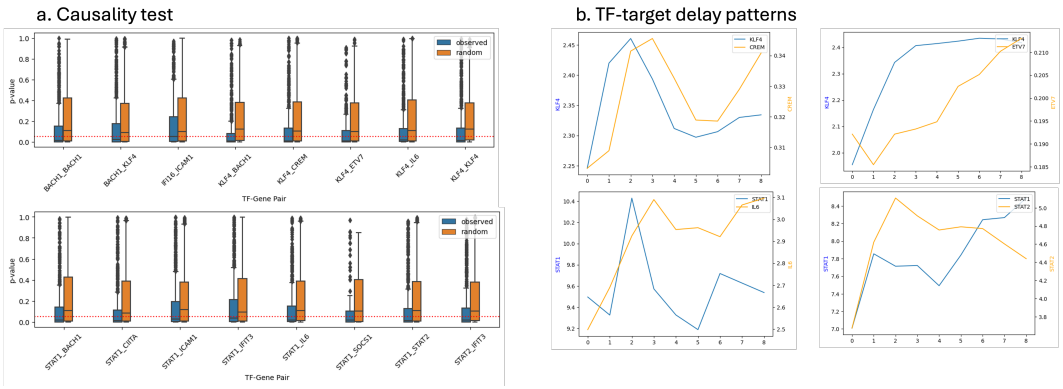

Figure 6: Transcriptional delay effects detected by SPADE. a) Causality test results on transcription rates of TF-target pairs. b) Examples of transcription rate curves of TF-target pairs inferred by SPADE.

5 Conclusion

In this work, we introduced SPADE, a systems biology–informed deep learning framework for inferring transcriptional rates and mRNA dynamics from static spatial transcriptomics data. By leveraging molecule-level spatial distributions, SPADE reconstructs pseudo-trajectories, disentangles transcriptional activity into generation and drift–diffusion processes, and embeds spatial heterogeneity through location-aware modeling. Experiments on both synthetic and real-world datasets demonstrate that SPADE accurately recovers transcriptional rates, reliably infers transcriptional on/off states, and uncovers regulatory delays between transcription factors and their targets.

Our analysis also underscores the broader potential of coupling biophysical constraints with deep learning to learn the dynamics of biological processes using static snapshots of omics data. Future extensions will incorporate additional modalities, such as epigenomic and proteomic data. Overall, SPADE establishes a foundation for AI-driven discovery of causal and dynamic mechanisms in complex biological systems.

References

[1] Reem Abouward and Giampietro Schiavo. Walking the line: mechanisms underlying directional mrna transport and localisation in neurons and beyond. *Cellular and Molecular Life Sciences*, 78(6):2665–2681, 2021.

[2] Josep Biayna and Gabrijela Dumbović. Decoding subcellular rna localization one molecule at a time. *Genome Biology*, 26(1):45, 2025.

[3] Hao Chen, Dongshunyi Li, and Ziv Bar-Joseph. Scs: cell segmentation for high-resolution spatial transcriptomics. *Nature methods*, 20(8):1237–1243, 2023.

[4] Zhanlin Chen, William C King, Aheyon Hwang, Mark Gerstein, and Jing Zhang. Deepvelo: Single-cell transcriptomic deep velocity field learning with neural ordinary differential equations. *Science advances*, 8(48):eabq3745, 2022.

[5] Shasha Chong, Chongyi Chen, Hao Ge, and X Sunney Xie. Mechanism of transcriptional bursting in bacteria. *Cell*, 158(2):314–326, 2014.

[6] Sulagna Das, Maria Vera, Valentina Gandin, Robert H Singer, and Evelina Tutucci. Intracellular mrna transport and localized translation. *Nature reviews Molecular cell biology*, 22(7):483–504, 2021.

[7] Feng He, Jan Buer, An-Ping Zeng, and Rudi Balling. Dynamic cumulative activity of transcription factors as a mechanism of quantitative gene regulation. *Genome Biology*, 8(9):R181, 2007.

[8] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997.

[9] Anton JM Larsson, Per Johnsson, Michael Hagemann-Jensen, Leonard Hartmanis, Omid R Faridani, Björn Reinius, Åsa Segerstolpe, Chloe M Rivera, Bing Ren, and Rickard Sandberg. Genomic encoding of transcriptional burst kinetics. *Nature*, 565(7738):251–254, 2019.

[10] Shengyu Li, Pengzhi Zhang, Weiqing Chen, Lingqun Ye, Kristopher W Brannan, Nhat-Tu Le, Jun-ichi Abe, John P Cooke, and Guangyu Wang. A relay velocity model infers cell-dependent rna velocity. *Nature biotechnology*, 42(1):99–108, 2024.

[11] Y Long, KS Ang, M Li, KLK Chong, R Sethi, C Zhong, et al. Spatially informed clustering, integration, and deconvolution of spatial transcriptomics with graphst. nat commun 2023; 14 (1): 1155.

[12] Yahui Long, Kok Siong Ang, Raman Sethi, Sha Liao, Yang Heng, Lynn van Olst, Shuchen Ye, Chengwei Zhong, Hang Xu, Di Zhang, et al. Deciphering spatial domains from spatial multi-omics with spatialglue. *Nature Methods*, 21(9):1658–1667, 2024.

[13] Raul HC Lopes. Kolmogorov-smirnov test. In *International encyclopedia of statistical science*, pages 718–720. Springer, 2011.

[14] Sergio Marco Salas, Louis B Kuemmerle, Christoffer Mattsson-Langseth, Sebastian Tismeyer, Christophe Avenel, Taobo Hu, Habib Rehman, Marco Grillo, Paulo Czarnewski, Saga Helgadottir, et al. Optimizing xenium in situ data utility by quality assessment and best-practice analysis workflows. *Nature Methods*, pages 1–11, 2025.

[15] Lambda Moses and Lior Pachter. Museum of spatial transcriptomics. *Nature methods*, 19(5):534–546, 2022.

[16] Xuyu Qian, Kyle Coleman, Shunzhou Jiang, Andrea J Kriz, Jack H Marciano, Chunyu Luo, Chunhui Cai, Monica Devi Manam, Emre Caglayan, Abbe Lai, et al. Spatial transcriptomics reveals human cortical layer and area specification. *Nature*, pages 1–11, 2025.

[17] Ali Shojaie and Emily B Fox. Granger causality: A review and recent advances. *Annual Review of Statistics and Its Application*, 9(1):289–319, 2022.

[18] Yaolai Wang, Tengfei Ni, Wei Wang, and Feng Liu. Gene transcription in bursting: a unified mode for realizing accuracy and stochasticity. *Biological Reviews*, 94(1):248–258, 2019.

[19] Jing Xiao, Xinyang Yu, Fanlin Meng, Yuncong Zhang, Wenbin Zhou, Yonghong Ren, Jingxia Li, Yimin Sun, Hongwei Sun, Guokai Chen, et al. Integrating spatial and single-cell transcriptomics reveals tumor heterogeneity and intercellular networks in colorectal cancer. *Cell Death & Disease*, 15(5):326, 2024.

[20] Alireza Yazdani, Lu Lu, Maziar Raissi, and George Em Karniadakis. Systems biology informed deep learning for inferring parameters and hidden dynamics. *PLoS computational biology*, 16(11):e1007575, 2020.

A Simulation Design

To quantitatively evaluate SPADE with ground-truth parameters, we simulated the transcriptional dynamics of RNA molecules within single cells based on a bistate (on/off) model combined with drift–diffusion processes.

Cellular environment. Each cell was modeled as a circular domain of fixed radius, within which molecules undergo directed drift and Brownian diffusion until degradation. Simulations were performed at a time resolution of 0.1 for a total of $T = 20$ time units.

Transcriptional dynamics. RNA synthesis followed a bistate process: during the *on* state, transcripts are generated at a constant rate and initiated near the nuclear center; during the *off* state, no new molecules are produced. In all cases, synthesized molecules will undergo the drift–diffusion dynamics.

Simulation scenarios. To capture a range of transcriptional behaviors, we designed five scenarios varying in the sequence of on/off states, their durations, and transcription rates (Table 1).

Table 1: Simulation scenarios used to benchmark SPADE.

Scenario	State sequence	Rate	Durations
1	Off $\rightarrow$ On	5	50, 150
2	Off $\rightarrow$ On $\rightarrow$ Off	5	50, 100, 50
3	On $\rightarrow$ Off $\rightarrow$ On	5	50, 100, 50
4	Off $\rightarrow$ On $\rightarrow$ Off	10	80, 80, 40
5	On $\rightarrow$ Off $\rightarrow$ On	10	30, 70, 100

These scenarios were chosen to reflect diverse transcriptional patterns, including different on/off switch patterns, and variable transcriptional intensities. The resulting data provides ground-truth dynamic parameters for benchmarking SPADE's inference of transcriptional states and rates.

B Additional example results of the experiments

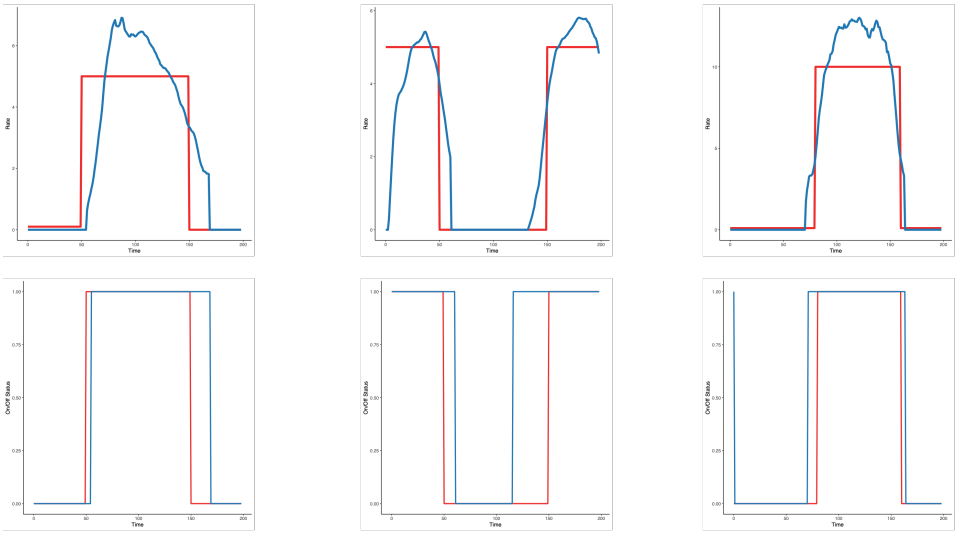

Figure 7: Inferred transcription rates and states of simulated scenarios which are not shown in main text.

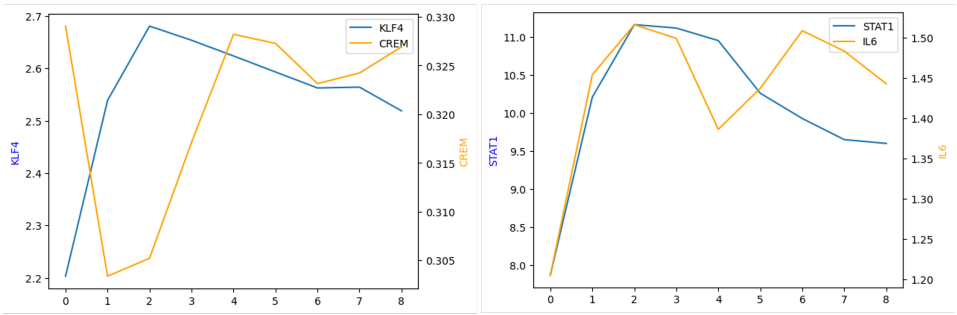

Figure 8: Additional examples of detected delay effects.

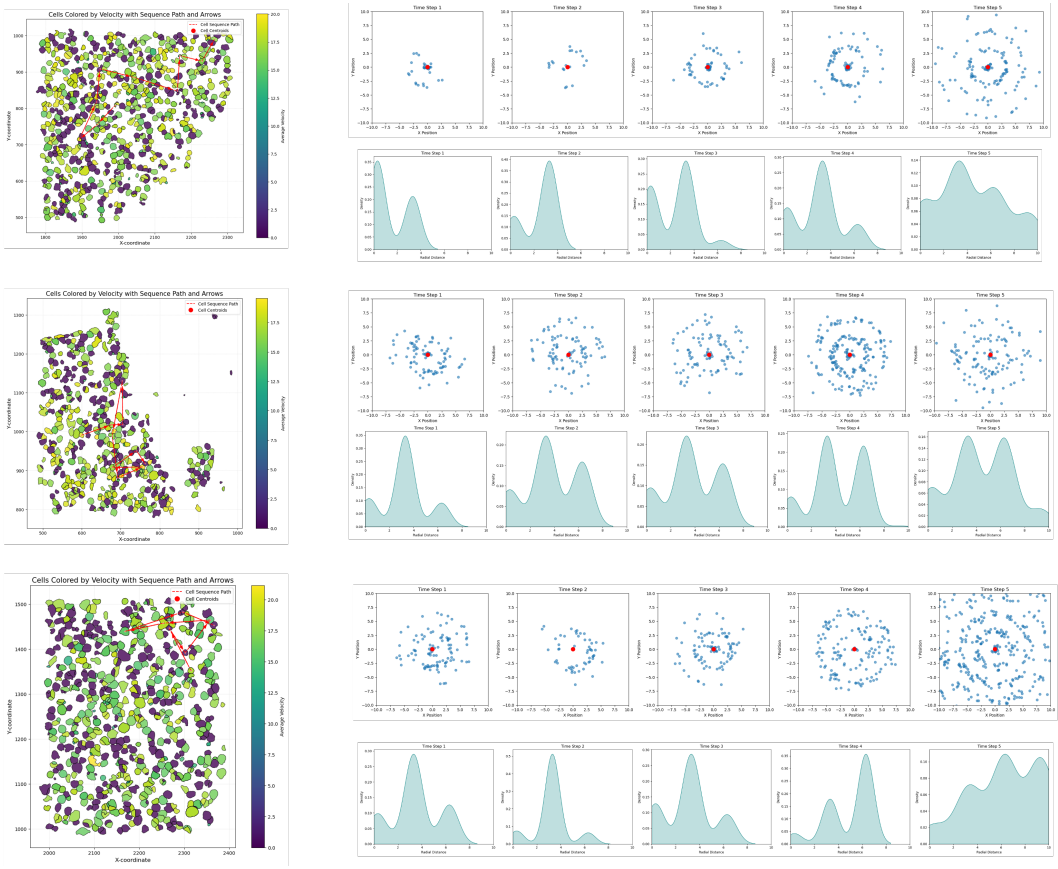

Figure 9: More examples showing the molecule distributions of sampled pseudo-trajectories.