

Time-Varying Gene Regulatory Networks Inference Using KL Divergence from Single Cell Data

Lingling Zhang

*Department of Mathematics & Statistics
University at Albany, SUNY
Albany, NY, USA
lzhang28@albany.edu*

Yunge Wang

*Department of Mathematics & Statistics
Saint Louis University
St. Louis, MO, USA
yunge.wang@slu.edu*

Tong Si

*Mathematics Department
Culver-Stockton College
Canton, MO, USA
tsi@culver.edu*

Lucas Koch

*Department of Mathematics & Statistics
Saint Louis University
St. Louis, MO, USA
lucas.koch@slu.edu*

Sarah Roberts

*Department of Mathematics & Statistics
Saint Louis University
St. Louis, MO, USA
sarah.roberts@slu.edu*

Haijun Gong*

*Department of Mathematics & Statistics
Saint Louis University
St. Louis, MO, USA
haijun.gong@slu.edu*

*Corresponding author

Abstract—Correct reconstruction of dynamic gene regulatory networks from time-series single-cell RNA sequencing (scRNA-seq) data is essential for understanding biological processes, but remains challenging due to high-dimensionality, sparsity, and temporal heterogeneity. We propose a novel framework that integrates Kullback–Leibler divergence (KL) divergence-based temporal variation measurement with an autoregressive model and different regularization methods to infer time-varying regulatory networks from time-series scRNA-seq data. Partial correlation analysis further refines the sign of interactions (activation or inhibition). Simulation studies on a 10-gene synthetic dataset and a THP-1 monocyte differentiation dataset demonstrate that our approach accurately recovers dynamic network structures and maintains temporal consistency.

Index Terms—Gene Regulatory Networks, KL Divergence, Single-Cell RNA Sequencing, Time-Series Analysis

I. INTRODUCTION

Understanding gene regulatory networks (GRNs) is critical for advancing our knowledge of complex biological systems and the mechanisms underlying various diseases. Single-cell RNA sequencing (scRNA-seq) technologies [1] enable the measurement of gene expression dynamics at single-cell resolution over time. However, inferring dynamic GRNs from time-series scRNA-seq data remains a significant challenge in systems biology, largely due to data sparsity, high dropout rates [2], cellular heterogeneity [3], and the limited number of time points typically available. Various algorithms have been developed to infer GRNs from time-series scRNA-seq data. Approaches based on ordinary or stochastic differential equations aim to model the continuous and dynamic nature of gene regulation over time [4], [5], while probabilistic methods apply machine learning and deep learning techniques to capture the inherent uncertainty in single-cell data [6], [7]. Machine learning-based network inference methods include correlation-based networks [8] and regression-based

approaches. For example, SINCERITIES [9] employs linear regression, while GRNBoost2 [7] uses tree-based regression models. More recently, some deep learning methods, for example DeepMAPS [10] and DeepDRIM [11], have been developed to infer GRNs from scRNA-seq data [10]; however, these approaches require large amounts of data to train the neural network, which poses a significant challenge for time-series scRNA-seq datasets due to their limited size.

Most traditional GRN inference methods assume a stationary network structure, limiting their ability to capture the dynamic regulatory changes that occur during processes such as differentiation, development, or disease progression [12]. A variety of methods have been developed to infer time-varying gene regulatory networks, including dynamic vector autoregressive models [13], [14], heterogeneous and weighted dynamic Bayesian network models [15], [16], dynamic autoregressive Gaussian graphical models [17], and time-varying graphical LASSO approaches [18]. However, these methods were primarily designed for bulk microarray data and are not well-suited to capture the temporal dynamics present in single-cell transcriptomic data.

To overcome these challenges, we propose a novel framework for inferring dynamic gene regulatory networks from time-series scRNA-seq data. Our approach integrates temporal variation quantified by Kullback–Leibler (KL) divergence with an autoregressive modeling framework that employs Lasso and Smoothly Clipped Absolute Deviation (SCAD) penalties for robust network estimation. Following network construction, we apply partial correlation analysis to determine the directionality of regulatory interactions: activation and inhibition. Finally, we validate our method using both a 10-gene synthetic dataset and a real-world time-series scRNA-seq dataset of THP-1 monocyte differentiation. Our results demonstrate that the proposed framework effectively reconstructs dynamic regulatory network structures while preserving temporal consistency across different time points.

This research was partially supported by the National Institutes of Health under Award Number R15GM148915 (HG) and President's Research Funds.

II. METHODS

Given a time-series single-cell RNA sequencing (scRNA-seq) dataset consisting of m genes and n time points, the number of cells at each time point t_l ($l = 1, 2, \dots, n$) is denoted as s_{t_l} . This time series scRNA-seq dataset can be represented as a collection of gene expression matrices at each time point t : $X \in \mathbb{R}^{s_{t_l} \times m}$, where each matrix X contains the expression levels of m genes across s_{t_l} individual cells at time point t_l .

A. KL-Divergence-Based Temporal Variation Estimation

We assume that at each time point, the expression levels of a specific gene across all single cells follow a probability distribution. Temporal variations in gene expression can be quantified by measuring the dissimilarity between these distributions at consecutive time points. Previous methods like SINCERITIES [9] assume that changes in transcription factor expression directly affect their target genes. While such changes are easily measured in microarray data, scRNA-seq requires more nuanced metrics to capture temporal dynamics at the single-cell level. SINCERITIES uses the Kolmogorov–Smirnov (KS) distance to quantify expression differences over time, but as it only reflects the maximum deviation between cumulative distributions, it may overlook gradual shifts in gene expression.

In this work, we use the Kullback–Leibler (KL) divergence to quantify temporal variation in gene expression. For gene j , the KL divergence between time points t_l and t_{l+1} is defined as:

$$D_{\text{KL}}^j(t_l) = \int_{\Omega} p^j(x, t_{l+1}) \log \left(\frac{p^j(x, t_{l+1})}{p^j(x, t_l)} \right) dx, \quad (1)$$

where $p^j(x, t_l)$ and $p^j(x, t_{l+1})$ are the probability density functions of gene j 's expression at two consecutive time points. While earlier work [9] normalizes this variation by the time interval Δt_l , our results suggest that such normalization is unnecessary, as its influence on the outcome is negligible.

B. Vector Autoregressive Model

Following the strategy used in [9], we formulate gene regulatory network (GRN) inference as a prediction problem, where the temporal shift in a target gene's distribution at time t_{l+1} is predicted based on the shifts of all genes at time t_l . We model this relationship using a first-order vector autoregressive model (VAR(1)), expressed as:

$$D_{\text{KL}}^j(t_{l+1}) = \sum_{p=1}^m \alpha_{p,j} D_{\text{KL}}^p(t_l) + \epsilon, \quad (2)$$

where $\alpha_{p,j}$ quantifies the influence of gene p on gene j .

To investigate how the regulatory network structures are changing at different stages, we adopt a moving window strategy. Within each window, we estimate the regression coefficients α , systematically tracking the evolution of regulatory interactions over time. Given the high dimensionality and inherent sparsity of gene regulatory networks, we next discuss how to apply various regularization methods to infer a sparse network.

C. Time-varying Regulatory Network Inference Algorithm

We use matrix notation to simplify the representation. Let $\mathbf{D}_{\text{KL}}(t) \in \mathbb{R}^{T \times m}$ denote the temporal variation of all m genes across T time points, and let $\alpha_j \in \mathbb{R}^m$ be the vector of regression coefficients associated with gene j . To obtain an optimal and sparse α_j , we solve the following regularized optimization problem:

$$\min_{\alpha_j} \frac{1}{2} \|D_{\text{KL}}^j(t_{l+1}) - \mathbf{D}_{\text{KL}}(t_l)\alpha_j\|^2 + \lambda p(\alpha_j), \quad (3)$$

where $p(\alpha_j)$ is a regularization term, which can be either the Lasso (L_1) penalty or the Smoothly Clipped Absolute Deviation (SCAD) penalty [19], and λ is the regularization parameter that controls the level of sparsity.

Lasso enforces sparsity by shrinking many coefficients to exactly zero, while SCAD reduces estimation bias for large coefficients and encourages smoother transitions in network structure over time. The SCAD penalty function for each coefficient is defined as:

$$p_{\lambda}(\alpha) = \begin{cases} |\alpha|, & \text{if } |\alpha| \leq \lambda \\ \frac{2\gamma\lambda|\alpha| - \alpha^2 - \lambda^2}{2(\gamma-1)}, & \text{if } \lambda < |\alpha| \leq \gamma\lambda \\ \frac{(\gamma+1)\lambda^2}{2}, & \text{if } |\alpha| > \gamma\lambda \end{cases} \quad (4)$$

where, $\lambda (> 0)$ is the regularization parameter controlling sparsity, $\gamma (> 2)$ determines the concavity of the penalty.

Algorithm 1 Dynamic regulatory network inference

Input: Time-stamped single-cell data matrix X

Output: Time-varying gene regulatory networks.

Step 1: Temporal Variation Calculation

- Randomly sample single cells at each time point.
- Compute temporal variations $D_{\text{KL}}^j(t_l)$ using KL-divergence.
- Repeat n times.

Step 2: Network Structure Learning

- Construct sliding windows between consecutive time points and build VAR(1) model for each gene.
- Solve the optimization problem to learn optimal α :

$$\min_{\alpha_j} \frac{1}{2} \|D_{\text{KL}}^j(t_{l+1}) - \mathbf{D}_{\text{KL}}(t_l)\alpha_j\|^2 + \lambda p(\alpha_j)$$

Step 3: Regulatory relationship inference

- Compute Spearman partial correlations and get the sign representing activation/inhibition.
-

Algorithm 1 outlines the procedure for inferring a time-varying gene regulatory network from time series scRNA-seq data. We first randomly sample single cells and compute temporal variations across time points for all genes using KL divergence, repeating this process n times. Next, we apply a VAR(1) model with different regularization techniques, including LASSO and SCAD, to infer a sparse, directed gene regulatory network. Finally, we compute partial correlations to determine whether each regulatory interaction represents activation (positive sign) or inhibition (negative sign).

III. RESULTS

In this section, we apply Algorithm 1 to reconstruct time-varying gene regulatory networks from time-series scRNA-seq data using KL divergence, VAR(1) model and different regularization methods (LASSO and SCAD).

A. Datasets and Parameters

Two datasets are used to evaluate the performance of our method:

- **10-Gene Dataset:** An *in silico* time-series dataset simulated from *Escherichia coli* GRNs, consisting of 100 single cells per time point across 8 time points.
- **THP-1 Dataset:** A real-world single-cell dataset capturing the differentiation of THP-1 human myeloid leukemia cells into macrophages, spanning 8 time points (0, 1, 6, 12, 24, 48, 72, 96 hours) with 120 cells per time point.

The LASSO and SCAD implementations are carried out using the `ncvreg` package in R. The optimal regularization parameter λ is selected via 10-fold cross-validation.

B. Evaluation Metrics

To evaluate network inference performance, we adopt the following metrics:

- **AUROC** (Area Under the Receiver Operating Characteristic Curve): Measures the ability to correctly predict regulatory edges.
- **Similarity Score:** Measures temporal stability by comparing edge overlaps between networks at adjacent time points.

C. Applications

We randomly sample 80% of single cells 100 times and compute the KL divergence-based temporal variation of all genes across time points. Following Algorithm 1, we reconstruct the time-varying regulatory network using various regularization methods and evaluate inference performance using AUROC and similarity scores.

To visually illustrate the inferred dynamic gene regulatory networks, we display representative dynamic regulatory network structures derived using SCAD and LASSO regularization under the forward KL divergence. Fig. 1 and Fig. 2 present six time-varying networks generated with LASSO and SCAD penalties (with $\gamma = 3$). In each subfigure, directed edges indicate regulatory relationships: solid black arrows represent activation, while red dashed arrows denote inhibition. These visualizations highlight how network topology evolves over time and differ across regularization strategies.

The dynamic networks inferred using LASSO (Fig. 1) tend to be more fragmented and exhibit greater variability across time points. Although LASSO occasionally identifies more edges at certain time points, these connections often fluctuate abruptly, indicating lower temporal consistency. In contrast, the SCAD-based networks (Fig. 2) are more structured, capturing key regulatory interactions while maintaining temporal coherence. Moreover, SCAD recovers stable hub genes whose regulatory roles persist over consecutive time points, reflecting

its ability to promote smooth network transitions. So, SCAD regularization under KL divergence strikes an effective balance between network sparsity and temporal smoothness, yielding biologically interpretable GRNs with stable and consistent dynamic patterns.

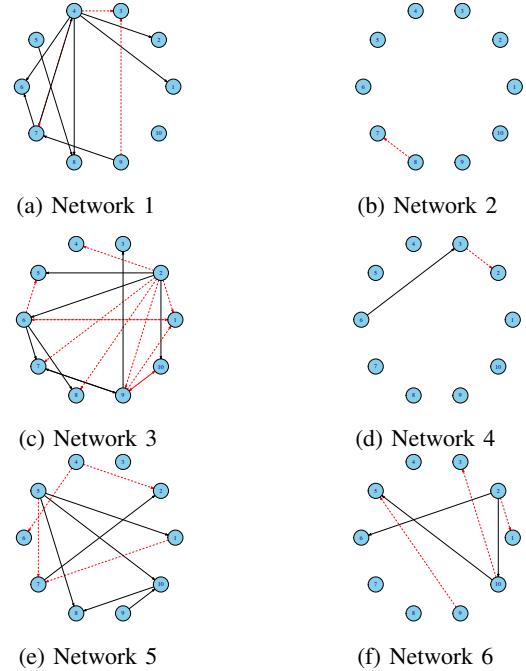


Fig. 1: Time-varying regulatory networks using KL divergence and LASSO penalty from the 10-Genes Dataset.

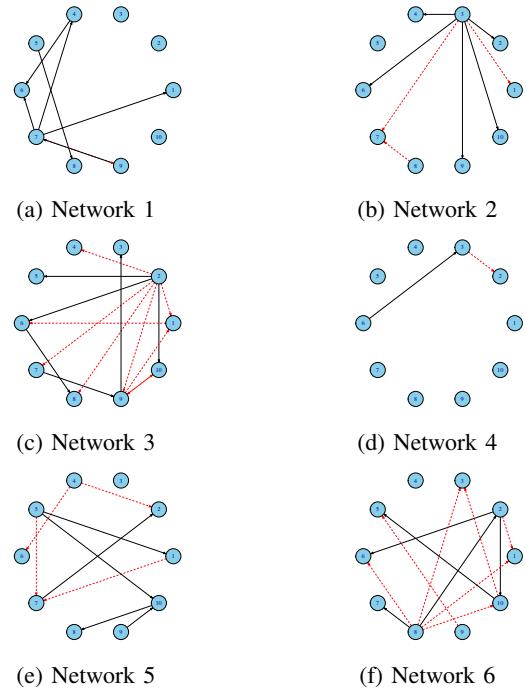


Fig. 2: Time-varying regulatory networks using KL divergence and SCAD ($\gamma = 3$) penalty from the 10-Genes Dataset.

To evaluate the reliability and stability of the inferred networks, we repeat the entire pipeline five times for each regularization method. For each run, we compute the AUROC and network similarity scores across time points. The results are averaged to obtain mean performance curves, with standard deviations indicating variability. Fig. 3 and Fig. 4 illustrate the AUROC and similarity scores for networks inferred using LASSO and SCAD under KL divergence. Additionally, to investigate the impact of regularization strength in SCAD, we plot performance curves across different values of the tuning parameter γ , highlighting its influence on both accuracy and temporal smoothness.

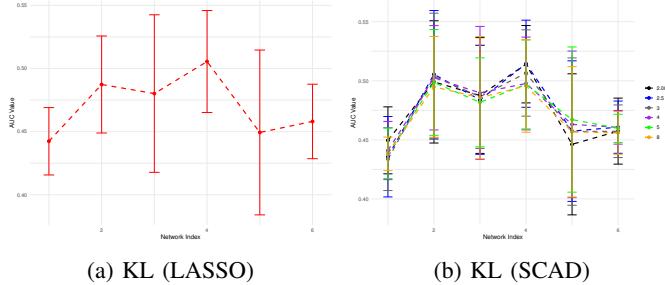


Fig. 3: AUROC values for LASSO and SCAD-based network inference on the 10-gene dataset under KL divergence.

Fig. 3 shows that networks inferred using LASSO under KL divergence yield moderate AUROC scores across different stages, exhibiting consistent patterns. In contrast, SCAD consistently achieves relatively higher AUROC values, especially at lower values of the tuning parameter γ . Fig. 4 presents the similarity curves across time points, highlighting the temporal consistency of the inferred networks. Both LASSO and SCAD produce comparable similarity trends, indicating a degree of structural continuity over time. However, the performance of SCAD is more sensitive to the tuning parameter γ , with larger values generally promoting smoother transitions between time points, while smaller values introduce greater variability and dynamic flexibility. This tunable behavior allows SCAD to better capture transient or context-specific regulatory interactions when needed. These results are consistent with the visualizations shown in Fig. 1 and Fig. 2, where LASSO-based networks exhibit more fragmented and less coherent temporal patterns, while SCAD-based networks maintain smoother and more stable regulatory structures over time. The similarity curves further validate the ability of SCAD to preserve meaningful temporal continuity, especially under appropriate tuning of the parameter γ . In summary, KL divergence proves effective in capturing dynamic regulatory patterns in the 10-gene dataset. Incorporating SCAD regularization further improves inference performance, particularly when the tuning parameter γ is appropriately chosen to balance network sparsity and temporal smoothness.

Next, we apply the proposed framework to analyze the timestamped single-cell qRT-PCR data and reconstruct the gene regulatory network (GRN) underlying THP-1 monocytic cell

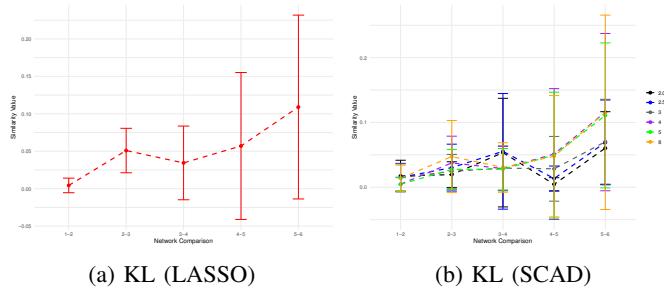


Fig. 4: Similarity curves for LASSO and SCAD-based network inference on the 10-gene dataset under KL divergence.

differentiation into macrophages. The dataset comprises 960 THP-1 cells collected at eight time points, with expression measured for 45 transcription factors (TFs) [20]. However, only 20 TFs overlap with a known gene regulatory network [21], allowing AUROC evaluation to be performed on this subset, while network inference is conducted on all 45 TFs.

Following the same protocol as with the 10-gene dataset, we repeat the full analysis five times for each KL divergence type and regularization method. In each run, we compute the AUROC and network similarity scores across time points to evaluate both inference accuracy and temporal stability of the reconstructed networks. The AUROC results obtained using LASSO and SCAD regularization under KL divergence are presented in Fig. 5, while the corresponding network similarity scores are shown in Fig. 6.

Fig. 5 (a) shows the AUROC scores using LASSO vary across time points without a clear upward or downward trend, indicating fluctuating accuracy in capturing regulatory interactions. The similarity curves (Fig. 6 (a)) also show considerable variation, suggesting that LASSO tends to generate less stable dynamic networks over time. In contrast, SCAD yields patterns (Fig. 5 (b) and Fig. 6 (b)) consistent with those observed in the 10-gene dataset. Lower values of the tuning parameter (e.g., $\gamma = 2.001, \gamma = 2.5$) lead to higher AUROC scores, reflecting improved accuracy in network reconstruction. Higher γ values (e.g., $\gamma = 5, \gamma = 8$), while yielding slightly lower AUROC, result in greater temporal coherence, as indicated by smoother transitions in the similarity curves.

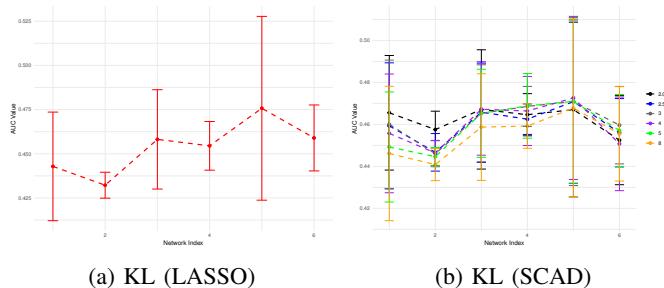


Fig. 5: AUROC values for LASSO and SCAD-based network inference on the THP-1 dataset.

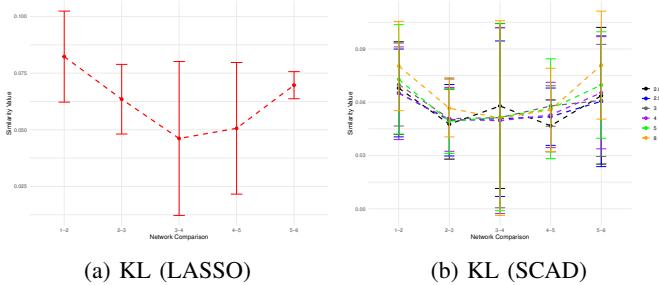


Fig. 6: Similarity curves for LASSO and SCAD-based network inference on the THP-1 dataset.

Overall, these results highlight SCAD's ability to balance GRN accuracy with dynamic smoothness more effectively than LASSO, which often produces sparser but more temporally variable network structures.

IV. CONCLUSIONS AND DISCUSSIONS

In this study, we proposed a novel and flexible framework for inferring time-varying gene regulatory networks from time-series single-cell RNA sequencing data. Our method integrates Kullback–Leibler (KL) divergence to quantify temporal gene expression changes, a first-order vector autoregressive model, VAR(1), to capture regulatory dependencies across time, and both LASSO and Smoothly Clipped Absolute Deviation (SCAD) regularization techniques to enforce the sparsity of network structure. To further identify the directionality and sign of regulatory interactions, we employed Spearman partial correlation analysis to classify edges as activation or inhibition.

We evaluated the performance of our framework on two datasets. Extensive experiments demonstrated that our approach effectively captures biologically meaningful dynamics in gene regulation. In particular, SCAD regularization, when combined with carefully selected tuning parameters, consistently outperformed LASSO in terms of AUROC scores and temporal smoothness of inferred networks.

Our future work will extend this framework to support additional divergence measures (e.g., f-divergence [22]), higher-order VAR processes, nonlinear gene interactions, and the integration of multi-omics data. Moreover, the current framework does not account for dropout events or change-points; future work will incorporate some missing value imputation [23] and change-point detection [24] methods to enhance model robustness and accuracy.

REFERENCES

- [1] S. S. Potter, “Single-cell rna sequencing for the study of development, physiology and disease,” *Nature Reviews Nephrology*, vol. 14, no. 8, pp. 479–492, 2018.
- [2] P. V. Kharchenko, L. Silberstein, and D. T. Scadden, “Bayesian approach to single-cell differential expression analysis,” *Nature methods*, vol. 11, no. 7, pp. 740–742, 2014.
- [3] A. Wagner, A. Regev, and N. Yosef, “Revealing the vectors of cellular identity with single-cell genomics,” *Nature biotechnology*, vol. 34, no. 11, pp. 1145–1160, 2016.
- [4] H. Matsumoto, H. Kiryu, C. Furusawa, M. S. Ko, S. B. Ko, N. Gouda, T. Hayashi, and I. Nikaido, “Scode: an efficient regulatory network inference algorithm from single-cell rna-seq during differentiation,” *Bioinformatics*, vol. 33, no. 15, pp. 2314–2321, 2017.
- [5] H. Matsumoto and H. Kiryu, “Scoup: a probabilistic model based on the ornstein–uhlenbeck process to analyze single-cell expression data during differentiation,” *BMC bioinformatics*, vol. 17, pp. 1–16, 2016.
- [6] V. A. Huynh-Thu, A. Irrthum, L. Wehenkel, and P. Geurts, “Inferring regulatory networks from expression data using tree-based methods,” *PloS one*, vol. 5, no. 9, p. e12776, 2010.
- [7] T. Moerman, S. Aibar Santos, C. Bravo González-Blas, J. Simm, Y. Moreau, J. Aerts, and S. Aerts, “Grnboost2 and arboreto: efficient and scalable inference of gene regulatory networks,” *Bioinformatics*, vol. 35, no. 12, pp. 2159–2161, 2019.
- [8] T. E. Chan, M. P. Stumpf, and A. C. Babtie, “Gene regulatory network inference from single-cell data using multivariate information measures,” *Cell systems*, vol. 5, no. 3, pp. 251–267, 2017.
- [9] N. Papili Gao, S. M. Ud-Dean, O. Gandrillon, and R. Gunawan, “Sin- certies: inferring gene regulatory networks from time-stamped single cell transcriptional expression profiles,” *Bioinformatics*, vol. 34, no. 2, pp. 258–266, 2018.
- [10] Q. Song, M. Ruffalo, and Z. Bar-Joseph, “Using single cell atlas data to reconstruct regulatory networks,” *Nucleic Acids Research*, vol. 51, no. 7, pp. e38–e38, 2023.
- [11] J. Chen, C. Cheong, L. Lan, X. Zhou, J. Liu, A. Lyu, W. K. Cheung, and L. Zhang, “Deepdrim: a deep neural network to reconstruct cell-type-specific gene regulatory network using single-cell rna-seq data,” *Briefings in bioinformatics*, vol. 22, no. 6, p. bbab325, 2021.
- [12] Z. Bar-Joseph, A. Gitter, and I. Simon, “Studying and modelling dynamic biological processes using time-series gene expression data,” *Nat. Rev. Genet.*, vol. 13, no. 8, pp. 552–564, Jul. 2012.
- [13] A. Fujita, J. R. Sato, H. M. Garay-Malpartida, P. A. Morettin, M. C. Sogayar, and C. E. Ferreira, “Time-varying modeling of gene expression regulatory networks using the wavelet dynamic vector autoregressive method,” *Bioinformatics*, vol. 23, no. 13, pp. 1623–1630, 2007.
- [14] A. Fujita, J. R. Sato, H. M. Garay-Malpartida, R. Yamaguchi, S. Miyano, M. C. Sogayar, and C. E. Ferreira, “Modeling gene expression regulatory networks with the sparse vector autoregressive model,” *BMC systems biology*, vol. 1, pp. 1–11, 2007.
- [15] H. Richards, Y. Wang, T. Si, H. Zhang, and H. Gong, “Intelligent learning and verification of biological networks,” *Advances in Artificial Intelligence, Computation, and Data Science: For Medicine and Life Science*, pp. 3–28, 2021.
- [16] F. Dondelinger, S. Lèbre, and D. Husmeier, “Non-homogeneous dynamic Bayesian networks with Bayesian regularization for inferring gene regulatory networks with gradually time-varying structure,” *Machine Learning*, vol. 90, no. 2, pp. 191–230, Feb. 2013.
- [17] E. C. Wit and A. Abbruzzo, “Inferring slowly-changing dynamic gene-regulatory networks,” *BMC bioinformatics*, vol. 16, pp. 1–11, 2015.
- [18] D. Hallac, Y. Park, S. Boyd, and J. Leskovec, “Network inference via the time-varying graphical lasso,” *KDD*, vol. 2017, pp. 205–213, Aug. 2017.
- [19] J. Fan and R. Li, “Variable selection via nonconcave penalized likelihood and its oracle properties,” *Journal of the American statistical Association*, vol. 96, no. 456, pp. 1348–1360, 2001.
- [20] T. Kouno, M. de Hoon, J. C. Mar, Y. Tomaru, M. Kawano, P. Carninci, H. Suzuki, Y. Hayashizaki, and J. W. Shin, “Temporal dynamics and transcriptional control using single-cell gene expression analysis,” *Genome biology*, vol. 14, pp. 1–12, 2013.
- [21] Y. Tomaru, C. Simon, A. R. Forrest, H. Miura, A. Kubosaki, Y. Hayashizaki, and M. Suzuki, “Regulatory interdependence of myeloid transcription factors revealed by matrix rna-seq analysis,” *Genome biology*, vol. 10, pp. 1–13, 2009.
- [22] S. M. Ali and S. D. Silvey, “A general class of coefficients of divergence of one distribution from another,” *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 28, no. 1, pp. 131–142, 1966.
- [23] W.-S. Liu, T. Si, A. Kriauciunas, M. Snell, and H. Gong, “Bidirectional f-divergence-based deep generative method for imputing missing values in time-series data,” *Stats*, vol. 8, no. 1, p. 7, 2025.
- [24] T. Si, Y. Wang, L. Zhang, E. Richmond, T.-H. Ahn, and H. Gong, “Multivariate time series change-point detection with a novel pearson-like scaled bregman divergence,” *Stats*, vol. 7, no. 2, pp. 462–480, 2024.