Dynamics on gene networks

Clayton W. Seitz

April 23, 2022

Outline

References

Research Strategy

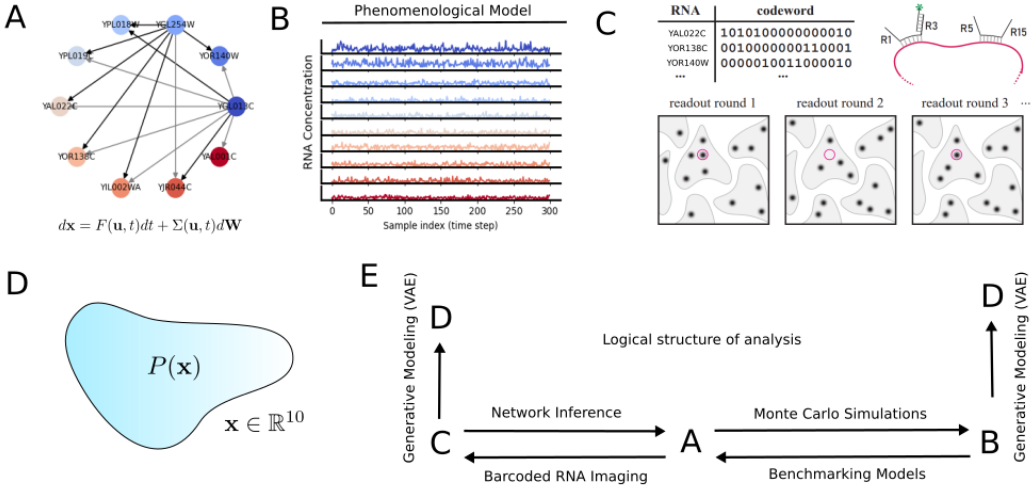


Figure 1: A *10-gene network sampled from Saccharomyces Genome Database (SGD).* **B** *Steady state values from Monte Carlo simulation.* **C** *Barcoding scheme for multiplexed RNA imaging* **D** *Cartoon of a joint distribution in higher dimensions* **E** *Relationships between components*

Simulating gene expression *in-silico*

We can simulate the expression of a gene as a function of the levels of its regulators (TFs), as prescribed by a fixed gene regulatory network (GRN)

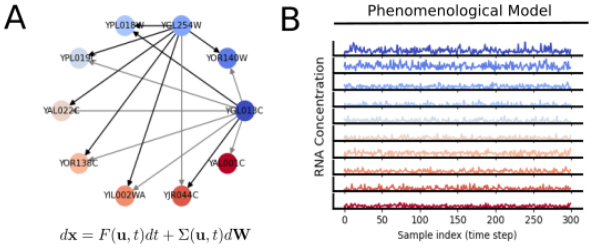
This is done via the chemical Langevin equation (Gillespie, 2000; Dibaeinia 2020):

$$\dot{x}_i = P_i(t) - \lambda_i x_i(t) + q_i \left(\sqrt{P_i(t)}\alpha + \sqrt{\lambda_i x_i}\beta \right)$$

The transcription rate of gene i depends non-linearly on its regulators:

$$P_i = \sum_j p_{ij} + b_i$$

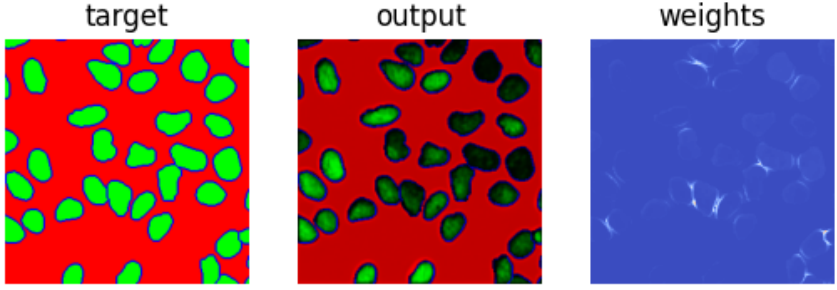
$$p_{ij} = m_{ij} \frac{y_{ij}^{n_{ij}}}{y_{ij}^{n_{ij}} + h_{ij}^{n_{ij}}}$$



$$d\mathbf{x} = F(\mathbf{u}, t)dt + \Sigma(\mathbf{u}, t)d\mathbf{W}$$

Training on BBBC039 U2OS Cells

BBBC039: 200 images, 160 train + 40 validation, 256 x 256 random crop

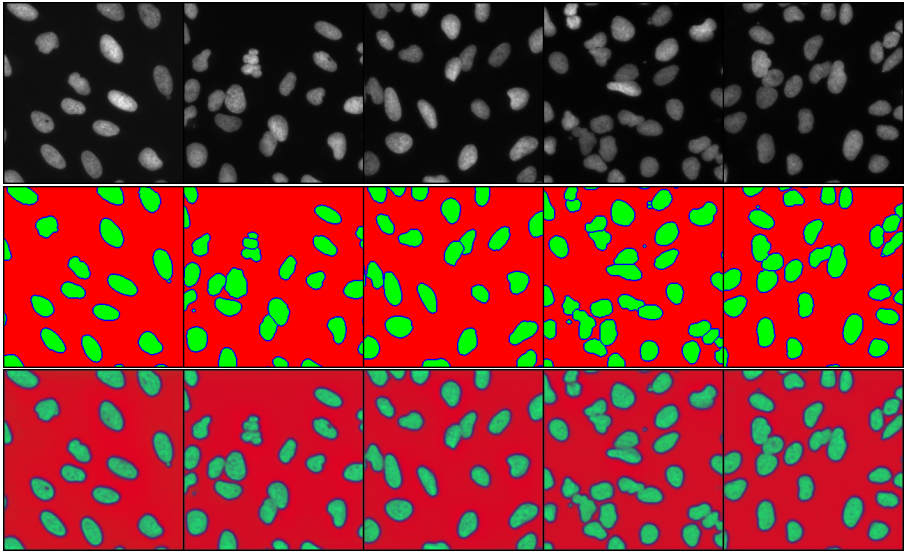


We train a 3-channel semantic segmentation model with **weighted** cross-entropy loss:

$$\mathcal{L} = \sum_{i,j} w_{ij} \log p_{ij}(\tilde{x}) = \sum_{i,j} w_{ij} \log \frac{\exp(-s_{ij}(\tilde{x}))}{\sum_{x \in \chi} \exp(-s_{ij}(\tilde{x}))}$$

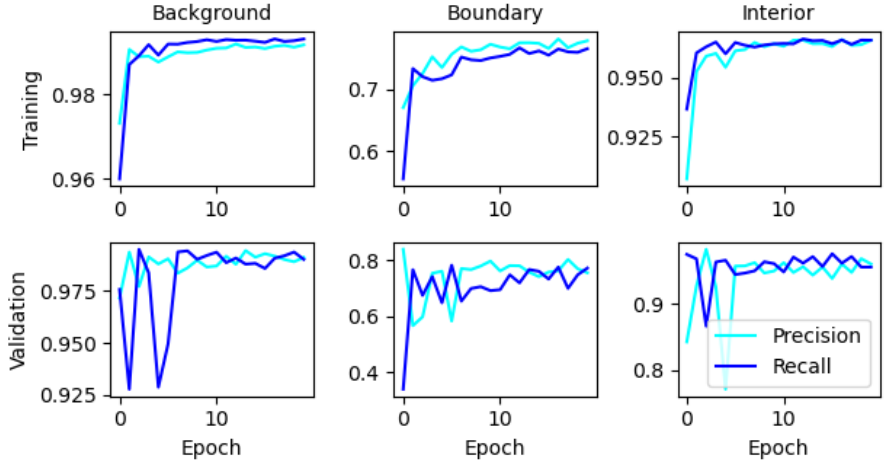
p_{ij} is the probability the model assigns a pixel to the true class $\tilde{x} \in \{\mathrm{a}, \mathrm{b}, \mathrm{c}\}$

Training on BBBC039 U2OS Cells



Training on BBBC039 U2OS Cells

Learning rate $\eta = 0.01$, Batch-size $B = 5$ (32 train iterations, 8 validation)



A scalable algorithm for inferring gene regulatory networks

Interest in reverse-engineering whole-genome interaction networks from simultaneous measurements of the expression levels of all (or at least most) genes in many samples, under a common set of experimental conditions

This algorithm applies when it is valid to assume that the system is in a steady state. For systems out of equilibrium, we need inference algorithms designed to operate on time-series data

Typically gene expression data have low sampling rates and relatively small amount of data. Moreover, GRNs have a high number of genes with complex, nonlinear regulatory mechanisms

Experimental considerations

Gene interactions are inferred from gene expression data. RNA-seq has single cell-specificity and time resolution but lacks spatial resolution and data is noisy

FISH techniques have single-cell specificity, spatial resolution, less noisy, but multiplexing is difficult and cells are fixed

High cost of multiplexing precludes acquisition of time-resolved data is single cell studies, which is important when statistics of the genes of interest are not stationary (circadian rhythms, cell-cycle, drug-treatment.

Transcription is not necessarily Poisson-like and has been shown to have switching behavior (transcriptional bursts). This has important implications for our models

Even when we can collect single-cell time-series data, data collected at a time point will contain extra variability due to asynchrony of cells within a population (in terms of progression through a biological process)

The Phi-Mixing Coefficient

Interest in reverse-engineering whole-genome interaction networks from simultaneous measurements of the expression levels of all (or at least most) genes in many samples, under a common set of experimental conditions

This algorithm applies when it is valid to assume that the system is in a steady state. For systems out of equilibrium, we need inference algorithms designed to operate on time-series data

Typically gene expression data have low sampling rates and relatively small amount of data. Moreover, GRNs have a high number of genes with complex, nonlinear regulatory mechanisms

Important topics

Models range from networks of a few genes with detailed dynamical models to very large networks with coarse statistical description.

- Linear dynamics of small networks (deterministic, stochastic)
- Nonlinear dynamics of small networks (deterministic, stochastic) - Bintu model
- Inferring network structure - Phi-Mixing Coefficient
- Inferring network structure from linear dynamics - Hidden Markov Models
- Inferring nonlinear network structure from empirical data - ?
- Simulating stochastic dynamics (Gillespie algorithm)
- Simulating stochastic nonlinear dynamics (Michaelis-Menten kinetics, SERGIO)
- Transcriptional bursting - switching behavior of gene promoter

References I