

# Bayesian parameter estimation for systems biological models of dynamic cellular networks

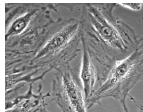
Jarad Niemi<sup>1</sup>, Lingchong You<sup>2</sup> and Mike West<sup>1</sup>

<sup>1</sup> Institute of Statistics & Decision Sciences, <sup>2</sup> Department of Biomedical Engineering, Duke University  
jarad@stat.duke.edu

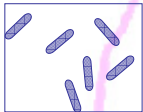
## Overview

Dynamic cellular networks determine production of proteins in the cells in our body. The Rb-E2F network is a key example: it controls cell cycle, proliferation and degradation - a key network in almost all cancers. Gene networks can be modeled using known interactions between activators, repressors and their targets. Emerging technologies reveal single-cell level stochastic variation - complex biological "intrinsic noise" - that plays a major role in network dynamics. We develop statistical models of such single-cell processes - fine-time scale dynamic network models. Model estimation is challenging - requiring informative priors and custom Metropolis proposals. Rb-E2F examples introduce this emerging area of statistical systems biology.

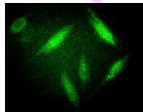
## Data



Identify cells from black and white images

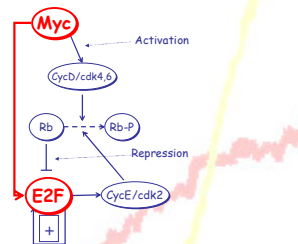


Measure color intensity of phosphorescently marked proteins e.g. E2F and Myc levels



## Models

Biological schematic "model"



Mathematical model

$$\frac{d[E2F^{obs}]}{dt} = k\pi([Myc]) - d[E2F^{obs}]$$

$$\pi(M) = \frac{M^n}{K^n + M^n}$$

- of the red pathway -

Need to account for

- Measurement error
- Sources of stochastic "intrinsic noise"

Novel discrete time statistical models account for both  
Introduce latent noise processes and latent protein levels

$$E2F_t^{obs} = E2F_t + \epsilon_t$$

$$E2F_{t+h} = (1-d)E2F_t + h k \pi(M y \epsilon_t) \exp(\lambda_t) + \delta_t$$

$$\lambda_t \leftarrow AR(1|\phi, \sigma_\lambda^2)$$

$$\epsilon_t \stackrel{iid}{\sim} N(0, \sigma_\epsilon^2)$$

$$\delta_t \stackrel{iid}{\sim} N(0, \sigma_\delta^2)$$

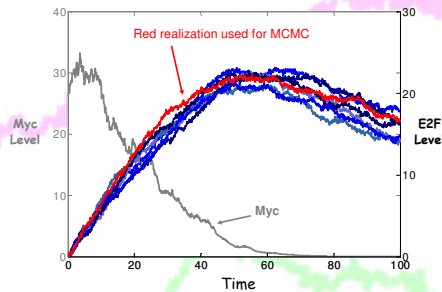
$$h = 1/H$$

Stochastic/statistical models

- of the red pathway -

## Simulations

Stochastic model accurately reflect real data



## MCMC

Priors

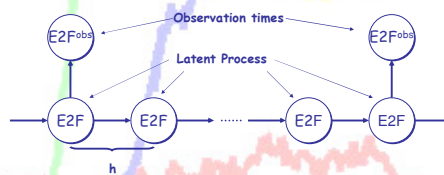
Truncated normals and gammas for key parameters

Gibbs & Metropolis-Hastings

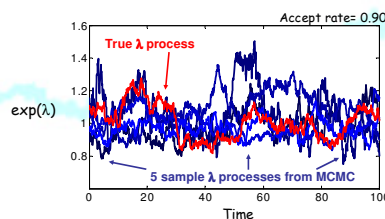
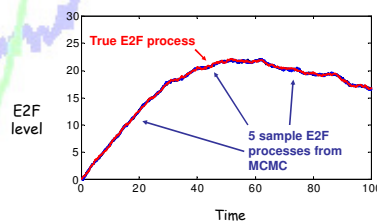
- Full conditionals easy for  $k, d, \phi, \sigma_\epsilon, \sigma_\delta$  and  $\sigma_\lambda$
- Random walk Metropolis-Hastings for  $K$  and  $n$

Forward filtering backward sampling (FFBS)

- FFBS to impute latent E2F protein process and intrinsic noise process  $\lambda$ .
- E2F - FFBS handles missing data
- $\lambda$  process: linearization in  $\lambda$  - Metropolis proposal



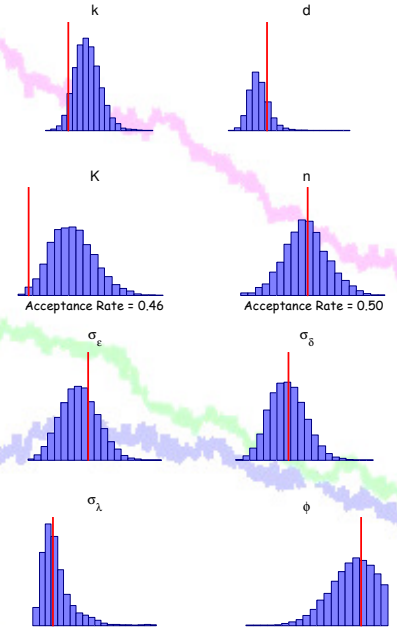
## Latent Process Recovery



## Parameter Estimation

Parameter estimation

Red line indicates true value



## Comments

- Novel stochastic models reflect single cell realities:
  - inherent probabilistic regulatory mechanisms
  - biochemical activation/repression/interaction probabilities
  - systematic intrinsic fluctuations
- Some current foci on:
  - effective MCMC methods including improved likelihood approximations for MH proposals of latent noise processes
  - block updating of parameters
  - expanding the gene network - multiple nodes
  - evaluating the impact of chosen discrete time scale
  - imaging: extraction of real-time data from movies
- Multi-scale models: synthesizing single-cell models with high-throughput genomic data of cell populations

## Some Key References

- Rosenfeld N, Young JW, Alon U, Swain PS and Elowitz MB (2005) Gene regulation at the single cell level, *Science* 307, 1962-1965.
- West and P.J. Harrison (1997) *Bayesian Forecasting and Dynamic Models* (2nd ed.), New York: Springer Verlag
- Wilkinson D (2006 - nearly here) *Stochastic Modelling for Systems Biology*, London: Chapman-Hall/CRC Press
- You L, Hoorlor A and Lin J (2004) Modelling biological systems using Dynetica: A simulator of dynamic networks, *Bioinformatics* 19, 435-436