

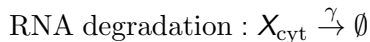
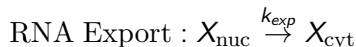
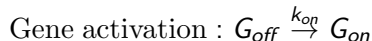
Bayesian inference of transcription dynamics from population snapshots of single-molecule RNA FISH in single cells

Clayton W. Seitz

August 1, 2022

A spatial model for induced gene expression

Let X represent an arbitrary RNA transcript of an induced gene G . Assume two promoter states (on and off)



Raw data collected post induction can be used to infer parameters

$$\theta = (k_{\text{on}}, k_{\text{off}}, k_t, k_{\text{exp}}, \gamma)$$

Bayesian inference of model parameters

It is well-known that using just means and variances gives poor estimates of the model parameters (Munsky et al. PNAS 2018)

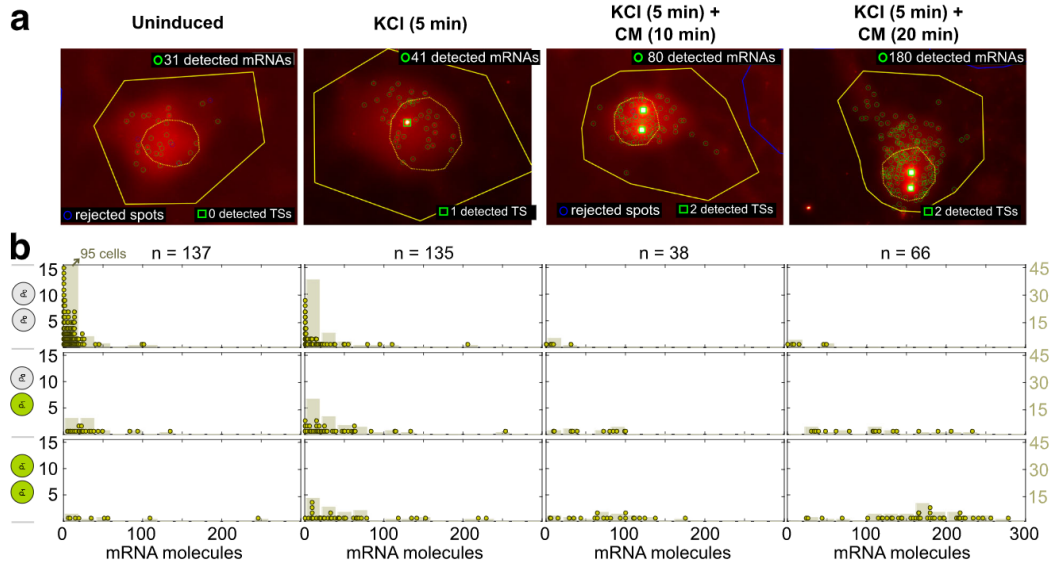
Let $\theta = (k_{on}, k_{off}, k_t, k_{exp}, \gamma)$. Using Bayes Rule:

$$P(\theta|X) = \frac{P(X|\theta)P(\theta)}{\int P(X|\theta)P(\theta)} \propto P(X|\theta)P(\theta)$$

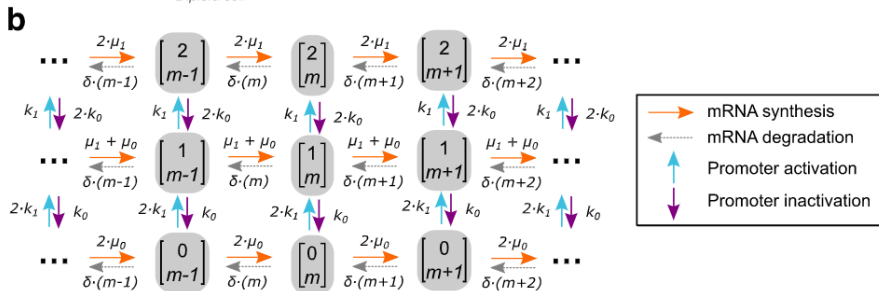
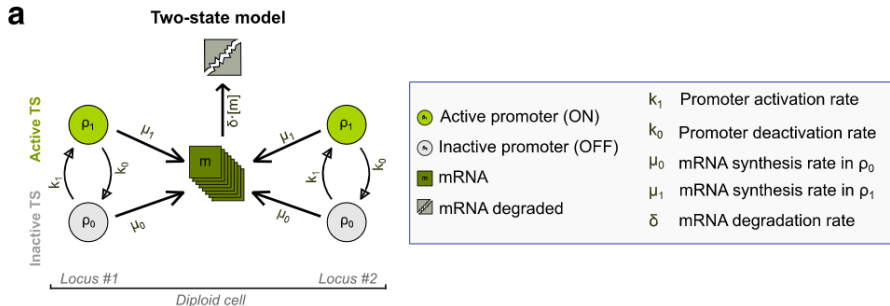
Can infer θ if we know the likelihood $P(X|\theta)$ and specify a prior $P(\theta)$

Generally we have to resort to Monte Carlo methods to find $P(X|\theta)$

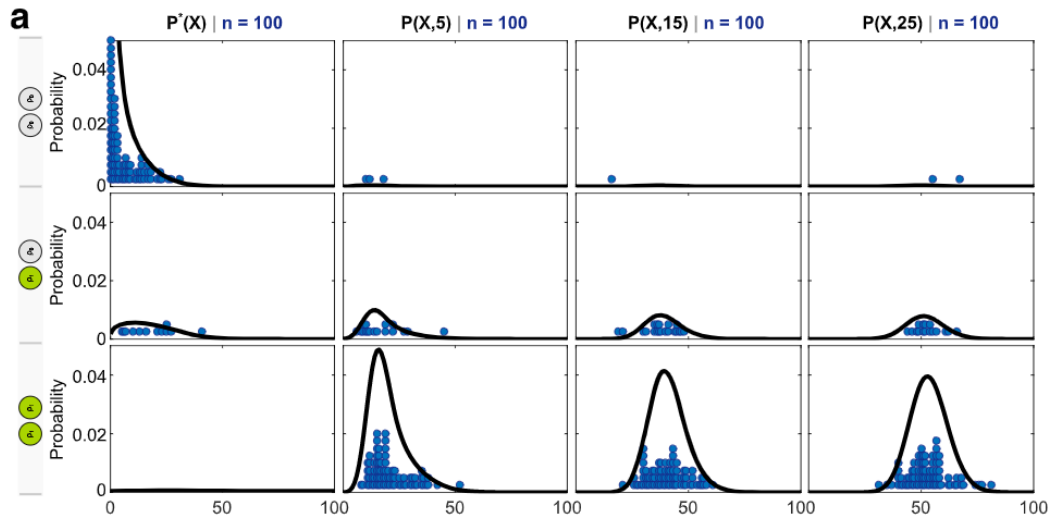
smFISH data of Npas4 mRNA in primary neurons after membrane depol.



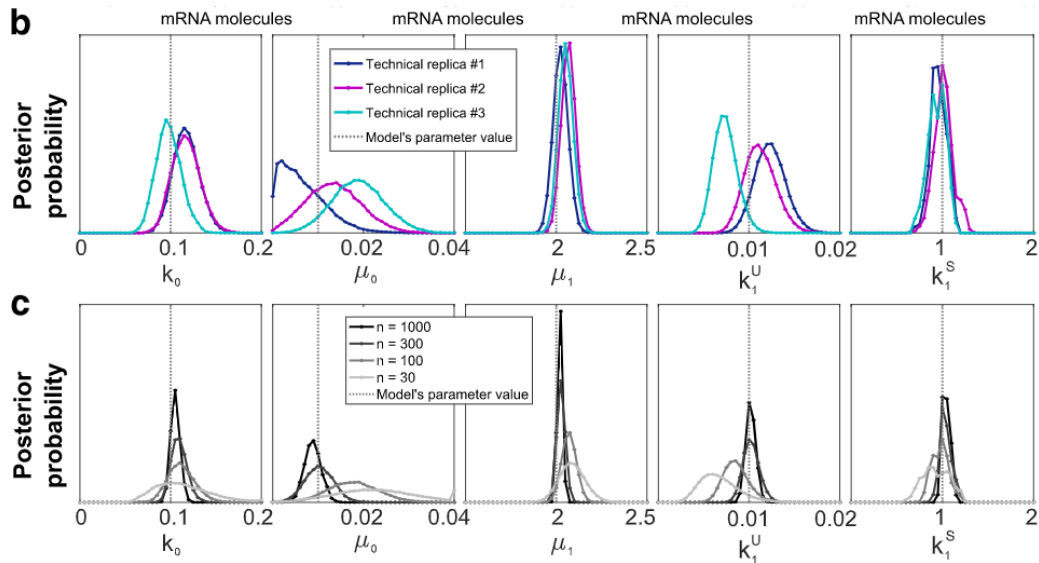
Two-state model of gene expression



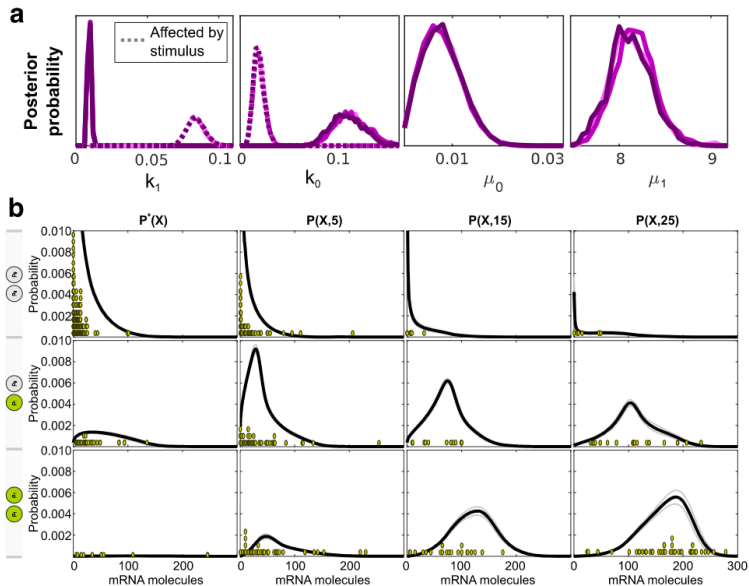
Validating BayFish on synthetic smFISH data



Validating BayFish on synthetic smFISH data



Model selection using BayFish and information criteria



Bayesian posterior distribution for (k_1, k_0) -stimulus model run on Npas4 data

