

# Parameter inference for small biochemical systems

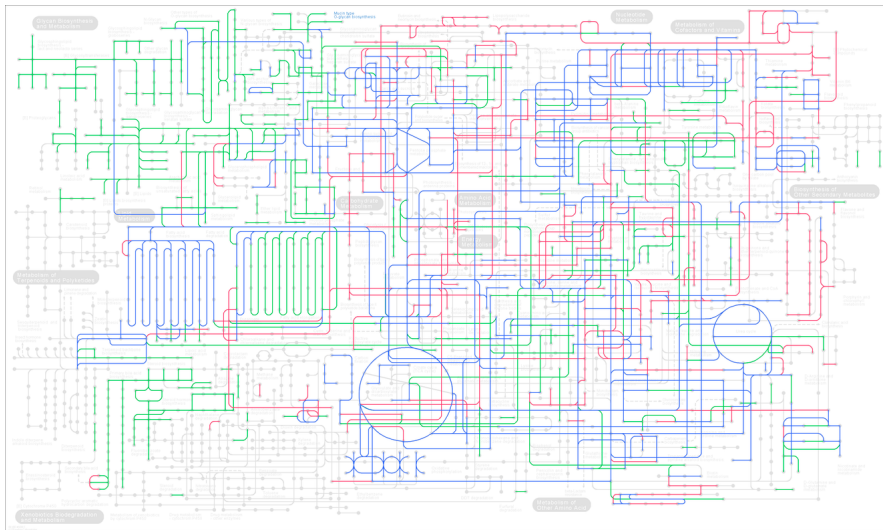
Eric Kernfeld<sup>1</sup>

<sup>1</sup>University of Washington Department of Statistics

# Paper

Darren Wilkinson's "Parameter inference for stochastic kinetic models of bacterial gene regulation," a book chapter in [1].

# Problem



# Problem

## Cast of Characters

- $r_j$  is a customer behavior pattern in a restaurant.
- $R_j(t)$  counts occurrences of  $r_j$  up to time  $t$ .
- $c_j$  is the rate at which  $r_j$  happens.
- $X(t)$  is the restaurant state: customers, orders, cash register.
- $\mathcal{D}_t$  is an incomplete observation of  $X(t)$  with error.
- $\tau$  governs measurement error.
- $\theta$  is  $\tau$  and  $c$  together.

# Problem

$r_1$ : 2 customers, hotcakes  $\xrightarrow{c_1}$  2 customers, 2 hotcakes, +\$5.65

# Problem

$r_1$ : 2 customers, hotcakes  $\xrightarrow{c_1}$  2 customers, 2 hotcakes, +\$5.65

$$R_1(t) \sim PP(c_1 \binom{X_1(t)}{2} \binom{X_2(t)}{1})$$

# Problem

$r_1$ : 2 customers, hotcakes  $\xrightarrow{c_1}$  2 customers, 2 hotcakes, +\$5.65

$$R_1(t) \sim PP(c_1 \binom{X_1(t)}{2} \binom{X_2(t)}{1})$$

$r_2$ : no prereqs  $\xrightarrow{c_2}$  1 customer

# Problem

$r_1$ : 2 customers, hotcakes  $\xrightarrow{c_1}$  2 customers, 2 hotcakes, +\$5.65

$$R_1(t) \sim PP(c_1 \binom{X_1(t)}{2} \binom{X_2(t)}{1})$$

$r_2$ : no prereqs  $\xrightarrow{c_2}$  1 customer

$$R_1(t) \sim PP(c_2)$$



# Problem

$r_1$ : 2 customers, hotcakes  $\xrightarrow{c_1}$  2 customers, 2 hotcakes, +\$5.65

$$R_1(t) \sim PP(c_1 \binom{X_1(t)}{2} \binom{X_2(t)}{1})$$

$r_2$ : no prereqs  $\xrightarrow{c_2}$  1 customer

$$R_1(t) \sim PP(c_2)$$

$r_2$ : sick customer, healthy customer, hotcakes  $\xrightarrow{c_3}$  2 sick customers

# Problem

$r_1$ : 2 customers, hotcakes  $\xrightarrow{c_1}$  2 customers, 2 hotcakes, +\$5.65

$$R_1(t) \sim PP(c_1 \binom{X_1(t)}{2} \binom{X_2(t)}{1})$$

$r_2$ : no prereqs  $\xrightarrow{c_2}$  1 customer

$$R_1(t) \sim PP(c_2)$$

$r_2$ : sick customer, healthy customer, hotcakes  $\xrightarrow{c_3}$  2 sick customers

$$R_3(t) \sim PP(c_3 \binom{X_1(t)}{1} \binom{X_2(t)}{1} \binom{X_3(t)}{1})$$

Assume pattern  $r_j$  occurs independently of other patterns except when the contents of the restaurant changes.

# Problem

Given  $\mathcal{D}_t$ , find  $c$ .

# Problem

Given  $\mathcal{D}_t$ , find  $c$ . Also, switching to ODE's is not allowed: stochastic models are biologically necessary ([2] and references therein).

# Problem

Likelihood:

$$P(X(\vec{t}), \vec{t} | c) = \prod_{i=1}^n c_{\nu_i} \prod_{j=1}^u \binom{X_j(t_{i-1})}{p_{\nu_i j}} \exp \left( -c_{\nu_i} \int_0^T \binom{X_j(t)}{p_{\nu_i j}} dt \right)$$

# Other work

- Approximate MJP with SDE [3, 4, 5]
- Method of moments [6, 7]
- Variational inference with mean-field approximation [8]
- Approximate Bayesian Computation (ABC) and MCMC [9, 10, 11, 12]
- EM with a sample average replacing the E-step expectation or similar [13, 14, 15, 16, 17]

# Likelihood Free MCMC

To produce a chain of samples from  $P(\theta|D)$ , using a proposal  $q(\theta^*|\theta)$ , accept with probability  $p_{rej}(\theta^*|\theta) \equiv \min\{1, A\}$  if

$$A = \frac{q(\theta, x|\theta^*, x^*)}{q(\theta^*, x^*|\theta, x)} \times \frac{P(\theta^*, x^*|\mathcal{D})}{P(\theta, x|\mathcal{D})}$$

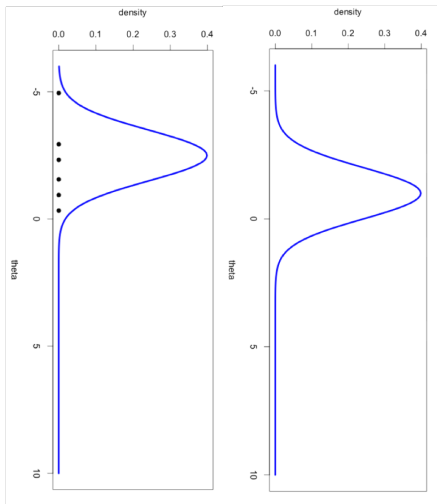
# Likelihood Free MCMC

$$\begin{aligned} & \frac{q(\theta^*, x^* | \theta, x)}{q(\theta, x | \theta^*, x^*)} \times \frac{P(x|\theta)}{P(x^*|\theta^*)} \frac{P(\theta)}{P(\theta^*)} \frac{P(\mathcal{D}|x, \theta)}{P(\mathcal{D}|x^*, \theta^*)} \\ &= \frac{f(\theta^*|\theta)}{f(\theta|\theta^*)} \frac{P(x^*|\theta^*)}{P(x|\theta)} \times \frac{P(x|\theta)}{P(x^*|\theta^*)} \frac{P(\theta)}{P(\theta^*)} \frac{P(\mathcal{D}|x, \theta)}{P(\mathcal{D}|x^*, \theta^*)} \\ &= \frac{f(\theta^*|\theta)}{f(\theta|\theta^*)} \times \frac{P(\theta)}{P(\theta^*)} \frac{P(\mathcal{D}|x, \theta)}{P(\mathcal{D}|x^*, \theta^*)}. \end{aligned}$$

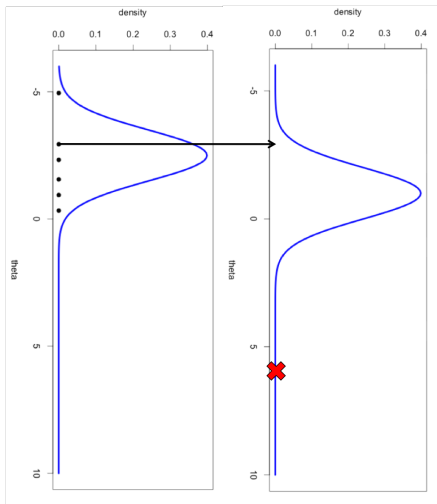
This approach due to Marjoram et al. (paper title: “MCMC Without Likelihoods”) [18].



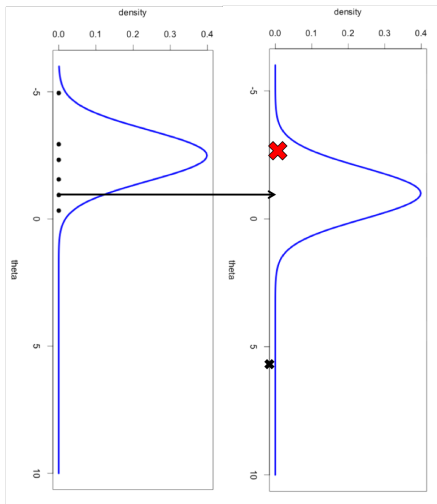
# Likelihood Free Particle MCMC



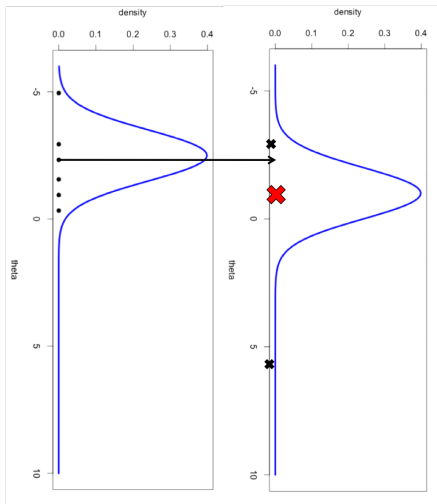
# Likelihood Free Particle MCMC



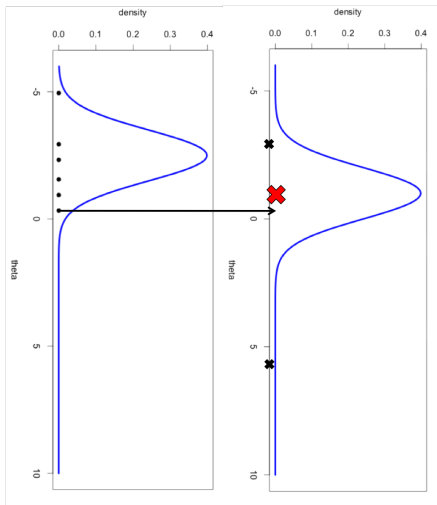
# Likelihood Free Particle MCMC



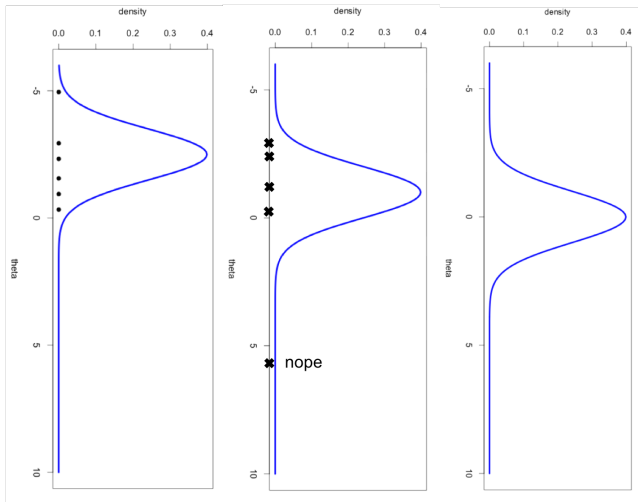
# Likelihood Free Particle MCMC



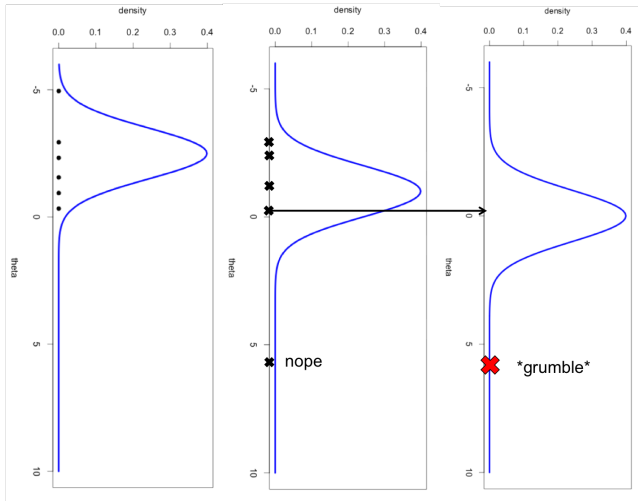
# Likelihood Free Particle MCMC



# Likelihood Free Particle MCMC



# Likelihood Free Particle MCMC



# Likelihood Free Particle MCMC





# Likelihood Free Particle MCMC

Given a hidden continuous-time Markov process  $\{x_t\}_{t=0}^T$  with:

- Unknown parameters  $\theta$  and known initial state  $x_0$

- Data points  $\mathcal{D}_{t_i}$  at times  $t_i$ ,  $i \in \{1, \dots, I\}$

- A simple, tractable error model  $P(\mathcal{D}_{t_i} | x_{t_i}, \theta)$

- A simulator for paths of  $x$  given  $\theta$

- An array  $B_0$  of 1,000,000 samples from a prior on  $\theta, x_0$

- Empty arrays  $B_i$  of the same length

For each time point (for  $i \in \{1, \dots, I\}$ ):

- Until  $B_i$  is full:

  - Draw  $(\theta^*, x_{t_{i-1}}^*)$  from  $B_{i-1}$  or a KDE of its contents

  - Using  $(\theta^*, x_{t_{i-1}}^*)$ , simulate up to  $x_{t_i}^*$ , the state at time  $t_i$

  - Set  $A = \min(1, \frac{P(\mathcal{D}_{t_i} | x_{t_i}^*, \theta^*)}{P(\mathcal{D}_{t_i} | x_{t_i}, \theta)})$

  - With probability  $A$ , overwrite  $(\theta, x_{t_i})$  with  $(\theta^*, x_{t_i}^*)$

  - After burn-in and thinning, add  $(\theta, x_{t_i})$  to  $B_i$

# Likelihood Free Particle MCMC

$$\frac{P(x_{t_i}^* | x_{t_{i-1}}^*, \theta^*, \mathcal{D}_{t_{i-1}})}{P(x_{t_i} | x_{t_{i-1}}, \theta, \mathcal{D}_{t_{i-1}})} \frac{P(x_{t_{i-1}}^*, \theta^* | \mathcal{D}_{t_{i-1}})}{P(x_{t_{i-1}}, \theta | \mathcal{D}_{t_{i-1}})} \times$$
$$\frac{P(x_{t_i} | x_{t_{i-1}}, \theta, \mathcal{D}_{t_{i-1}})}{P(x_{t_i}^* | x_{t_{i-1}}^*, \theta^*, \mathcal{D}_{t_{i-1}})} \frac{P(x_{t_{i-1}}, \theta | \mathcal{D}_{t_{i-1}})}{P(x_{t_{i-1}}^*, \theta^* | \mathcal{D}_{t_{i-1}})} \frac{P(\mathcal{D}_{t_i} | x_{t_i}, \theta)}{P(\mathcal{D}_{t_i} | x_{t_i}^*, \theta^*)}$$

# Likelihood Free Particle MCMC

$$\frac{P(x_{t_i}^* | x_{t_{i-1}}^*, \theta^*, \mathcal{D}_{t_{i-1}})}{P(x_{t_i} | x_{t_{i-1}}, \theta, \mathcal{D}_{t_{i-1}})} \frac{P(x_{t_{i-1}}^*, \theta^* | \mathcal{D}_{t_{i-1}})}{P(x_{t_{i-1}}, \theta | \mathcal{D}_{t_{i-1}})} \times$$
$$\frac{P(x_{t_i} | x_{t_{i-1}}, \theta, \mathcal{D}_{t_{i-1}})}{P(x_{t_i}^* | x_{t_{i-1}}^*, \theta^*, \mathcal{D}_{t_{i-1}})} \frac{P(x_{t_{i-1}}, \theta | \mathcal{D}_{t_{i-1}})}{P(x_{t_{i-1}}^*, \theta^* | \mathcal{D}_{t_{i-1}})} \frac{P(\mathcal{D}_{t_i} | x_{t_i}, \theta)}{P(\mathcal{D}_{t_i} | x_{t_i}^*, \theta^*)}$$

$$P(x_{t_i} | x_{t_{i-1}}, \theta, \mathcal{D}_{t_{i-1}}) P(x_{t_{i-1}}, \theta | \mathcal{D}_{t_{i-1}}) P(\mathcal{D}_{t_i} | x_{t_i}, \theta)$$

# Likelihood Free Particle MCMC

$$\frac{P(x_{t_i}^* | x_{t_{i-1}}^*, \theta^*, \mathcal{D}_{t_{i-1}})}{P(x_{t_i} | x_{t_{i-1}}, \theta, \mathcal{D}_{t_{i-1}})} \frac{P(x_{t_{i-1}}^*, \theta^* | \mathcal{D}_{t_{i-1}})}{P(x_{t_{i-1}}, \theta | \mathcal{D}_{t_{i-1}})} \times$$
$$\frac{P(x_{t_i} | x_{t_{i-1}}, \theta, \mathcal{D}_{t_{i-1}})}{P(x_{t_i}^* | x_{t_{i-1}}^*, \theta^*, \mathcal{D}_{t_{i-1}})} \frac{P(x_{t_{i-1}}, \theta | \mathcal{D}_{t_{i-1}})}{P(x_{t_{i-1}}^*, \theta^* | \mathcal{D}_{t_{i-1}})} \frac{P(\mathcal{D}_{t_i} | x_{t_i}, \theta)}{P(\mathcal{D}_{t_i} | x_{t_i}^*, \theta^*)}$$

$$P(x_{t_i} | x_{t_{i-1}}, \theta, \mathcal{D}_{t_{i-1}}) P(x_{t_{i-1}}, \theta | \mathcal{D}_{t_{i-1}}) P(\mathcal{D}_{t_i} | x_{t_i}, \theta)$$

$$= P(x_{t_i}, x_{t_{i-1}}, \theta | \mathcal{D}_{t_{i-1}}) P(\mathcal{D}_{t_i} | x_{t_i}, \theta, x_{t_{i-1}}, \mathcal{D}_{t_{i-1}})$$

$$= P(x_{t_i}, x_{t_{i-1}}, \theta | \mathcal{D}_{t_i}) P(\mathcal{D}_{t_i} | \mathcal{D}_{t_{i-1}})$$

# Questions? I



Bernardo, J.M., Bayarri, M.J., Berger, J.O., Dawid, A.P., Heckerman, D., Smith, A.F.M., West, M.:  
Ninth Valencia international meeting on Bayesian statistics,  
Benidorm, Spain, 03-08.06.2010.  
Oxford U.P., Oxford (2012)



Reinker, S., Altman, R., Timmer, J.:  
Parameter estimation in stochastic biochemical reactions.  
IEE Proceedings-Systems Biology **153**(4) (2006) 168–178



Golightly, A., Wilkinson, D.J.:  
Bayesian inference for stochastic kinetic models using a diffusion  
approximation.  
Biometrics **61**(3) (2005) 781–788

# Questions? II



Golightly, A., Wilkinson, D.J.a.:

Bayesian parameter inference for stochastic biochemical network models using particle markov chain monte carlo.

Interface Focus **1**(6) (10 2011) 807–820



Fearnhead, P., Giagos, V., Sherlock, C.:

Inference for reaction networks using the linear noise approximation.

Biometrics **70**(2) (2014) 457–466



Milner, P., Gillespie, C.S., Wilkinson, D.J.:

Moment closure based parameter inference of stochastic kinetic models.

Statistics and Computing **23**(2) (2013) 287–295

# Questions? III



Zechner, C., Ruess, J., Krenn, P., Pelet, S., Peter, M., Lygeros, J., Koeppl, H.:

Moment-based inference predicts bimodality in transient gene expression.

Proceedings of the National Academy of Sciences **109**(21)  
(2012) 8340–8345



Opper, M., Sanguinetti, G.:

Variational inference for markov jump processes.

In: Advances in Neural Information Processing Systems. (2008)  
1105–1112



Owen, J., Wilkinson, D.J., Gillespie, C.S.:

Likelihood free inference for markov processes: a comparison.  
arXiv preprint arXiv:1410.0524 (2014)

# Questions? IV

 Owen, J., Wilkinson, D.J., Gillespie, C.S.:

Scalable inference for markov processes with intractable likelihoods.

Statistics and Computing (2014) 1–12

 Hobolth, A., Stone, E.A.:

Simulation from endpoint-conditioned, continuous-time markov chains on a finite state space, with applications to molecular evolution.

The annals of applied statistics **3**(3) (2009) 1204

 Zechner, C., Unger, M., Pelet, S., Peter, M., Koepl, H.:

Scalable inference of heterogeneous reaction kinetics from pooled single-cell recordings.

Nature methods **11**(2) (2014) 197–202



# Questions? V



Gupta, A., Rawlings, J.B.:

Comparison of parameter estimation methods in stochastic chemical kinetic models: Examples in systems biology.

*AIChE Journal* **60**(4) (2014) 1253–1268



Srivastava, R., Rawlings, J.B.:

Parameter estimation in stochastic chemical kinetic models using derivative free optimization and bootstrapping.

*Computers & chemical engineering* **63** (2014) 152–158



Bayer, C., Moraes, A., Tempone, R., Vilanova, P.:

An efficient forward-reverse expectation-maximization algorithm for statistical inference in stochastic reaction networks.

*arXiv preprint arXiv:1504.04155* (2015)

# Questions? VI



Horváth, A., Manini, D.:

Parameter estimation of kinetic rates in stochastic reaction networks by the em method.

In: BioMedical Engineering and Informatics, 2008. BMEI 2008. International Conference on. Volume 1., IEEE (2008) 713–717



Daigle, B.J., Roh, M.K., Petzold, L.R., Niemi, J.:

Accelerated maximum likelihood parameter estimation for stochastic biochemical systems.

BMC bioinformatics **13**(1) (2012) 68



Marjoram, P., Molitor, J., Plagnol, V., Tavaré, S.:

Markov chain monte carlo without likelihoods.

Proceedings of the National Academy of Sciences **100**(26) (2003) 15324–15328

# Questions? VII