Hybrid Inference for Stochastic Kinetic Models

Andrew Golightly
School of Mathematics & Statistics
Newcastle University, UK

LICSB, April 2009

Overview

Modelling

Represent a biochemical network with a set of (pseudo-)biochemical reactions:

k species and r reactions with a typical reaction

$$R_i : \quad u_{i1} Y_1 + \ldots + u_{ik} Y_k \quad \xrightarrow{c_i} \quad v_{i1} Y_1 + \ldots + v_{ik} Y_k$$

Stochastic rate constant: c_i

Hazard / instantaneous rate: $h_i(Y, c_i)$ where $Y = (Y_1, \ldots, Y_k)'$ is the current state of the system and

$$h_i(Y, c_i) = c_i \prod_{j=1}^{k} \binom{Y_j}{u_{ij}} = c_i g_i(Y)$$

Modelling cont'd

Some remarks:

- This setup describes a Markov jump process (MJP)
- The effect of reaction R_i is to change the value of each Y_j by $v_{ij} - u_{ij}$
- It can be shown that the time to the next reaction is

$$t \sim \text{Exp}\{h_0(Y, c)\} \qquad \text{where} \qquad h_0(Y, c) = \sum_{i=1}^{k} h_i(Y, c_i)$$

  and the reaction is of type i with probability $h_i(Y, c_i)/h_0(Y, c_i)$
- Hence, the process is easily simulated (and this technique is known as the Gillespie algorithm)

Inference for the Exact Model

Aim: infer the c_i given time-course biochemical data. Following Boys et al. (2008) and Wilkinson (2006):

- Suppose we observe the entire process $\mathbf{Y}$ over $[0, T]$
- The ith unit interval contains n_i reactions with times and types (t_{ij}, k_{ij}), $j = 1, 2, \ldots, n_i$
- Hence, the likelihood for c is

$$\pi(\mathbf{Y}|c) = \left\{ \prod_{i=0}^{T-1} \prod_{j=1}^{n_i} h_{k_{ij}} \left\{ Y(t_{i,j-1}), c_{k_{ij}} \right\} \right\} \exp\left\{ - \int_0^T h_0 \left\{ Y(t), c \right\} \, dt \right\}$$

- So, if $c_i \sim \text{Gamma}(a_i, b_i)$ *a priori* then

$$c_i|\mathbf{Y} \sim \text{Gamma}\left(a_i + r_i \, , \, b_i + \int_0^T g_i \left\{ Y(t) \right\} \, dt \right)$$

where r_i is the no. of type i reactions in $(0, T]$.

Inference for the Exact Model

Aim: infer the c_i given time-course biochemical data. Following Boys et al. (2008) and Wilkinson (2006):

- Suppose we observe the entire process $\mathbf{Y}$ over $[0, T]$
- The ith unit interval contains n_i reactions with times and types (t_{ij}, k_{ij}), $j = 1, 2, \ldots, n_i$
- Hence, the likelihood for c is

$$\pi(\mathbf{Y}|c) = \left\{ \prod_{i=0}^{T-1} \prod_{j=1}^{n_i} h_{k_{ij}} \left\{ Y(t_{i,j-1}), c_{k_{ij}} \right\} \right\} \exp\left\{ - \int_0^T h_0 \left\{ Y(t), c \right\} \, dt \right\}$$

- So, if $c_i \sim \text{Gamma}(a_i, b_i)$ *a priori* then

$$c_i|\mathbf{Y} \sim \text{Gamma}\left(a_i + r_i \, , \, b_i + \int_0^T g_i \left\{ Y(t) \right\} \, dt \right)$$

where r_i is the no. of type i reactions in $(0, T]$.

Inference for the Exact Model

Aim: infer the c_i given time-course biochemical data. Following Boys et al. (2008) and Wilkinson (2006):

- Suppose we observe the entire process $\mathbf{Y}$ over $[0, T]$
- The ith unit interval contains n_i reactions with times and types (t_{ij}, k_{ij}), $j = 1, 2, \ldots, n_i$
- Hence, the likelihood for c is

$$\pi(\mathbf{Y}|c) = \left\{ \prod_{i=0}^{T-1} \prod_{j=1}^{n_i} h_{k_{ij}} \left\{ Y(t_{i,j-1}), c_{k_{ij}} \right\} \right\} \exp\left\{ - \int_0^T h_0 \left\{ Y(t), c \right\} \, dt \right\}$$

- So, if $c_i \sim \text{Gamma}(a_i, b_i)$ *a priori* then

$$c_i|\mathbf{Y} \sim \text{Gamma}\left(a_i + r_i \, , \, b_i + \int_0^T g_i \left\{ Y(t) \right\} \, dt \right)$$

where r_i is the no. of type i reactions in $(0, T]$

Inference for the Exact Model cont'd

Problem: it is not feasible to observe all reaction times and types

- Assume data are observed on a regular grid with

$$Y_{0:T} = \left\{ Y(t) = (Y_1(t), Y_2(t), \ldots, Y_k(t))' : t = 0, 1, 2, \ldots, T \right\}$$

Idea: use a Gibbs sampler to alternate bewteen draws of

1. times and types of reactions in $(0, T]$ conditional on c and the observations,
2. each c_i conditional on the augmented data

Note that step 1 can be performed for each interval $(i, i+1]$ in turn, due to the factorisation of $\pi(\mathbf{Y}|Y_{0:T}, c)$

Inference for the Exact Model cont'd

Problem: it is not feasible to observe all reaction times and types

- Assume data are observed on a regular grid with

$$Y_{0:T} = \left\{ Y(t) = (Y_1(t), Y_2(t), \ldots, Y_k(t))' : t = 0, 1, 2, \ldots, T \right\}$$

Idea: use a Gibbs sampler to alternate bewteen draws of

1. times and types of reactions in $(0, T]$ conditional on c and the observations,
2. each c_i conditional on the augmented data

Note that step 1 can be performed for each interval $(i, i+1]$ in turn, due to the factorisation of $\pi(\mathbf{Y}|Y_{0:T}, c)$

Inference for the Exact Model cont'd

Problem: it is not feasible to observe all reaction times and types

- Assume data are observed on a regular grid with

$$Y_{0:T} = \left\{ Y(t) = (Y_1(t), Y_2(t), \ldots, Y_k(t))' : t = 0, 1, 2, \ldots, T \right\}$$

Idea: use a Gibbs sampler to alternate bewteen draws of

1. times and types of reactions in $(0, T]$ conditional on c and the observations,
2. each c_i conditional on the augmented data

Note that step 1 can be performed for each interval $(i, i+1]$ in turn, due to the factorisation of $\pi(\mathbf{Y}|\, Y_{0:T}, c)$

Difficulties

- These techniques do not scale well to problems of realistic size and complexity...
- True process is discrete and stochastic — stochasticity is vital — what about discreteness?
- Treating molecule numbers as continuous and performing exact inference for the resulting approximate model appears to be promising..
- From the literature:
  - Approximations via moment closure (Gillespie & Golightly (2009), Pete Milner's talk etc)
  - A diffusion approximation (Golightly & Wilkinson (2009), Ruttor et al. (2009), Heron et al. (2007), etc)

Difficulties

- These techniques do not scale well to problems of realistic size and complexity...
- True process is discrete and stochastic — stochasticity is vital — what about discreteness?
- Treating molecule numbers as continuous and performing exact inference for the resulting approximate model appears to be promising..
- From the literature:
  - Approximations via moment closure (Gillespie & Golightly (2009), Pete Milner's talk etc)
  - A diffusion approximation (Golightly & Wilkinson (2009), Ruttor et al. (2009), Heron et al. (2007), etc)

Diffusion Approximation

- Consider an infinitesimal time interval $(t, t + dt]$. Let

  $dR(t) =$ the r-vector of the number of reaction events in $(t, t + dt]$

  $S =$ the $k \times r$ *net effect* matrix so that

  $dY(t) = S\,dR(t)$, the amount the system state should be updated by

- The ith element of $dR(t)$ is Poisson$(h_i(Y(t), c_i)dt)$ and so

  $$\mathrm{E}\{dR(t)\} = h(Y(t), c)\,dt, \qquad \mathrm{Var}\{dR(t)\} = \mathrm{diag}\{h(Y(t), c)\}\,dt$$

  where $h(Y(t), c) = (h_1(Y(t), c_1), \ldots, h_r(Y(t), c_r))'$

- Plainly,

  $$dR(t) = h(Y(t), c)\,dt + \mathrm{diag}\left\{\sqrt{h(Y(t), c)}\right\} dW(t)$$

  $$\Rightarrow dY(t) = S\,h(Y(t), c)\,dt + \sqrt{S\,\mathrm{diag}\{h(Y(t), c)\}\,S'}\,dW(t)$$

  since $dY(t) = S\,dR(t)$

Diffusion Approximation

- Consider an infinitesimal time interval $(t, t + dt]$. Let

  $dR(t) =$ the r-vector of the number of reaction events in $(t, t + dt]$

  $S =$ the $k \times r$ *net effect* matrix so that

  $dY(t) = S\, dR(t)$, the amount the system state should be updated by

- The ith element of $dR(t)$ is Poisson$(h_i(Y(t), c_i)dt)$ and so

  $$\mathrm{E}\{dR(t)\} = h(Y(t), c)\, dt, \qquad \mathrm{Var}\{dR(t)\} = \mathrm{diag}\{h(Y(t), c)\}\, dt$$

  where $h(Y(t), c) = (h_1(Y(t), c_1), \ldots, h_r(Y(t), c_r))'$

- Plainly,

$$dR(t) = h(Y(t), c)\, dt + \mathrm{diag}\left\{\sqrt{h(Y(t), c)}\right\} dW(t)$$

$$\Rightarrow dY(t) = S\, h(Y(t), c)\, dt + \sqrt{S\, \mathrm{diag}\{h(Y(t), c)\}\, S'}\, dW(t)$$

since $dY(t) = S\, dR(t)$

Diffusion Approximation

- Consider an infinitesimal time interval $(t, t + dt]$. Let

  $dR(t) =$ the r-vector of the number of reaction events in $(t, t + dt]$

  $S =$ the $k \times r$ *net effect* matrix so that

  $dY(t) = S\, dR(t)$, the amount the system state should be updated by

- The ith element of $dR(t)$ is Poisson$(h_i(Y(t), c_i)dt)$ and so

  $$\mathrm{E}\{dR(t)\} = h(Y(t), c)\, dt, \qquad \mathrm{Var}\{dR(t)\} = \mathrm{diag}\{h(Y(t), c)\}\, dt$$

  where $h(Y(t), c) = (h_1(Y(t), c_1), \ldots, h_r(Y(t), c_r))'$

- Plainly,

  $$dR(t) = h(Y(t), c)\, dt + \mathrm{diag}\left\{\sqrt{h(Y(t), c)}\right\} dW(t)$$

  $$\Rightarrow dY(t) = S\, h(Y(t), c)\, dt + \sqrt{S\, \mathrm{diag}\{h(Y(t), c)\}\, S'}\, dW(t)$$

  since $dY(t) = S\, dR(t)$

Diffusion Approximation – Inference

Work with the Euler discretisation

$$\Delta Y(t) = S\,h(Y(t), c)\,\Delta t + \sqrt{S\,\mathrm{diag}\{h(Y(t), c)\}\,S'}\,\Delta W(t)$$
$$\Delta W(t) \sim N(0, I\Delta t)$$

Hence, transition densities are approximated as Gaussian with

$$Y(t+\Delta t)|Y(t), c \sim \mathrm{N}\left(Y(t) + S\,h(Y(t), c)\,\Delta t\,,\; S\,\mathrm{diag}\{h(Y(t), c)\}\,S'\Delta t\right)$$

So, if data are observed on a fine grid, $t_0 < t_1 < \ldots < t_n$,

$$\pi(c|\cdot) \propto \pi(c) \times \prod_{i=1}^{n} \pi\left(Y(t_i)|Y(t_{i-1}), c\right)$$
$$= \text{prior} \times \text{likelihood under the Euler scheme}$$

Diffusion Approximation – Inference

Work with the Euler discretisation

$$\Delta Y(t) = S\,h(Y(t), c)\,\Delta t + \sqrt{S\,\mathrm{diag}\{h(Y(t), c)\}\,S'}\,\Delta W(t)$$
$$\Delta W(t) \sim N(0, I\Delta t)$$

Hence, transition densities are approximated as Gaussian with

$$Y(t+\Delta t)|\,Y(t), c \sim \mathrm{N}\left(Y(t) + S\,h(Y(t), c)\,\Delta t\,,\; S\,\mathrm{diag}\{h(Y(t), c)\}\,S'\Delta t\right)$$

So, if data are observed on a fine grid, $t_0 < t_1 < \ldots < t_n$,

$$\pi(c|\,\cdot) \propto \pi(c) \times \prod_{i=1}^{n} \pi\left(Y(t_i)|\,Y(t_{i-1}), c\right)$$
$$= \text{prior} \times \text{likelihood under the Euler scheme}$$

Difficulties

- Typically inter-observation times are too large to be used as a time-step in the Euler approximation
- One solution is to augment low frequency data with latent observations to allow the Euler approximation to become accurate
- MCMC can then be used to sample the joint posterior of latent observations and parameters (see Golightly & Wilkinson (2005,2008))
- For low copy number scenarios, ignoring inherent discreteness seems unacceptable...

Difficulties

- Typically inter-observation times are too large to be used as a time-step in the Euler approximation
- One solution is to augment low frequency data with latent observations to allow the Euler approximation to become accurate
- MCMC can then be used to sample the joint posterior of latent observations and parameters (see Golightly & Wilkinson (2005,2008))
- For low copy number scenarios, ignoring inherent discreteness seems unacceptable...

Difficulties

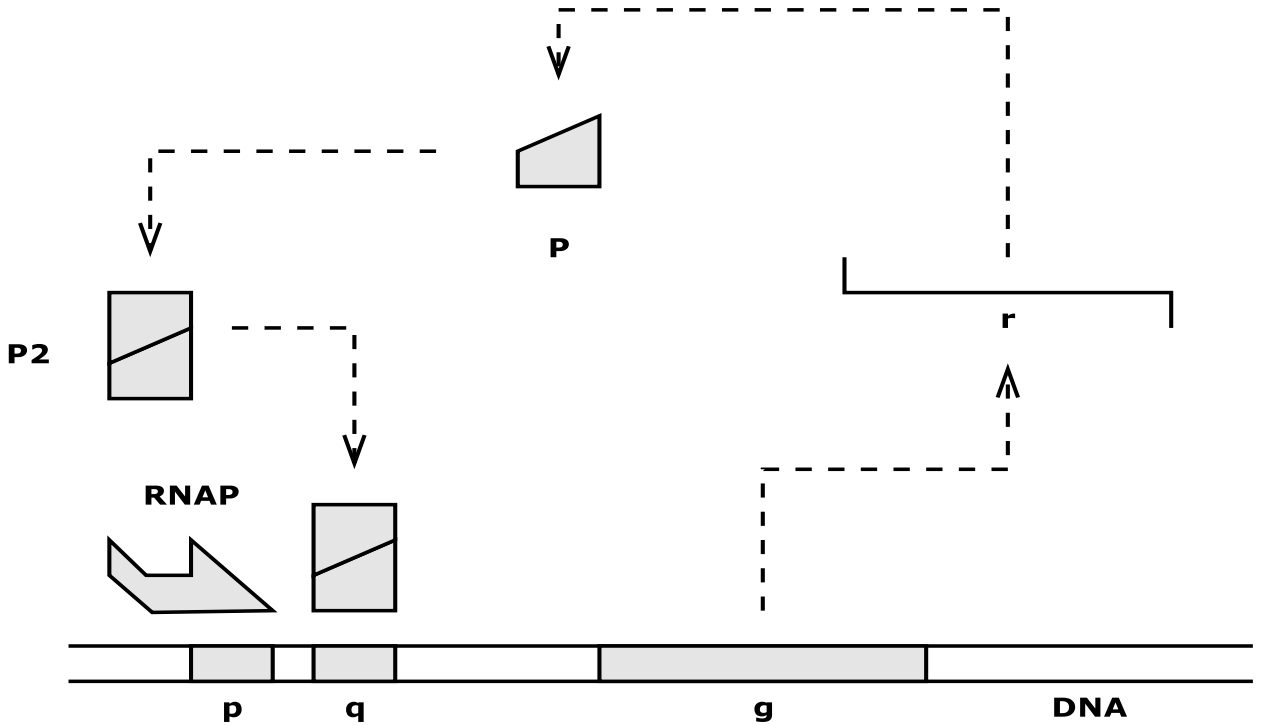
- Typically inter-observation times are too large to be used as a time-step in the Euler approximation
- One solution is to augment low frequency data with latent observations to allow the Euler approximation to become accurate
- MCMC can then be used to sample the joint posterior of latent observations and parameters (see Golightly & Wilkinson (2005,2008))
- For low copy number scenarios, ignoring inherent discreteness seems unacceptable...

Difficulties

- Typically inter-observation times are too large to be used as a time-step in the Euler approximation
- One solution is to augment low frequency data with latent observations to allow the Euler approximation to become accurate
- MCMC can then be used to sample the joint posterior of latent observations and parameters (see Golightly & Wilkinson (2005,2008))
- For low copy number scenarios, ignoring inherent discreteness seems unacceptable...

Hybrid Inference – Motivation

Toy Prokaryotic Auto-Regulation:



Hybrid Inference – Motivation

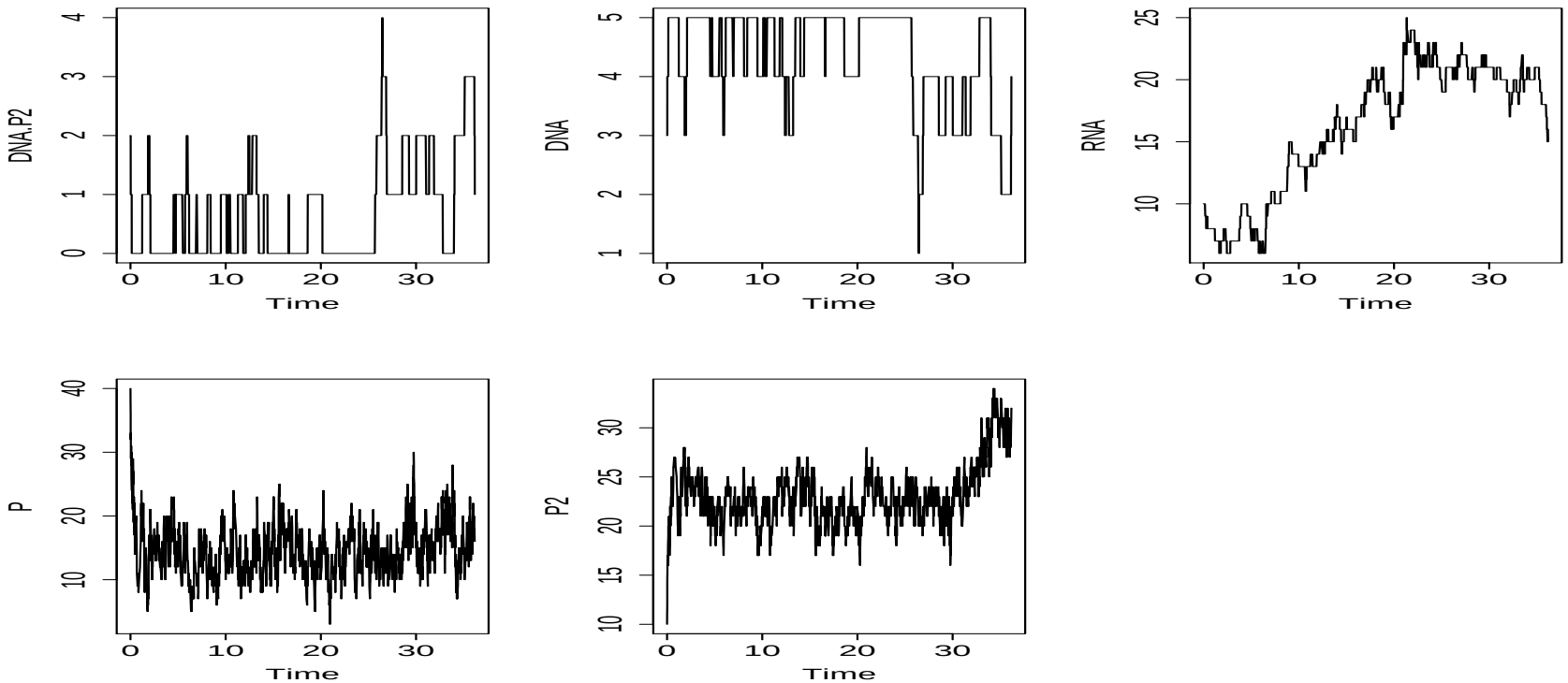
Toy Prokaryotic Auto-Regulation:

R_1 :	$\mathrm{DNA} + \mathrm{P}_2$	$\longrightarrow$	$\mathrm{DNA} \cdot \mathrm{P}_2$	Repression
R_2 :	$\mathrm{DNA} \cdot \mathrm{P}_2$	$\longrightarrow$	$\mathrm{DNA} + \mathrm{P}_2$	
R_3 :	DNA	$\longrightarrow$	$\mathrm{DNA} + \mathrm{RNA}$	Transcription
R_4 :	RNA	$\longrightarrow$	$\mathrm{RNA} + \mathrm{P}$	Translation
R_5 :	2P	$\longrightarrow$	P_2	Dimerisation
R_6 :	P_2	$\longrightarrow$	2P	
R_7 :	RNA	$\longrightarrow$	$\emptyset$	Degradation
R_8 :	P	$\longrightarrow$	$\emptyset$	

- 5 species $\mathrm{DNA}, \mathrm{DNA} \cdot \mathrm{P}_2, \mathrm{RNA}, \mathrm{P}, \mathrm{P}_2$ and 8 reactions with rate constants $c = (c_1, \ldots, c_8)'$
- Note that DNA and $\mathrm{DNA} \cdot \mathrm{P}_2$ are deterministically related

Hybrid Inference – Motivation

Synthetic data simulated via the Gillespie algorithm:



Hybrid Inference – Motivation

Comments:

- Numbers of DNA and DNA $\cdot P_2$ are in $\{0, 1, 2, 3, 4, 5\}$
- Reactions that change numbers of DNA and DNA $\cdot P_2$ must occur fairly infrequently

Therefore:

- Treat numbers of DNA and DNA $\cdot P_2$ as discrete – label these as slow
- Treat numbers of RNA, P and P_2 as continuous – label these as fast
- Label any reaction that changes the state of the slow species as slow and the remaining ones as fast

How can we perform inference within this framework?

Hybrid Inference – Motivation

Comments:

- Numbers of DNA and DNA $\cdot P_2$ are in $\{0, 1, 2, 3, 4, 5\}$
- Reactions that change numbers of DNA and DNA $\cdot P_2$ must occur fairly infrequently

Therefore:

- Treat numbers of DNA and DNA $\cdot P_2$ as discrete – label these as slow
- Treat numbers of RNA, P and P_2 as continuous – label these as fast
- Label any reaction that changes the state of the slow species as slow and the remaining ones as fast

How can we perform inference within this framework?

Hybrid Simulation

See for example, Salis & Kaznessis (2005):

1. Initialise the system, set $t := 0$
2. Calculate the fast reaction hazards, numerically integrate the SDE for the fast reactions over $(t, t + \Delta t]$, giving a sample path for the fast species over $(t, t + \Delta t]$
3. Using the slow reaction hazards, decide whether or not a slow reaction has happened in $(t, t + \Delta t]$
4. If no slow reaction has occurred, set $t := t + \Delta t$ and update the fast species to their proposed values at t
5. If one slow reaction has occurred, identify the time t_1 and type, set $t = t_1$ and update the system to t_1
6. If more than one slow reaction has occurred, reduce Δt and goto step 2
7. If $t < T_{max}$, return to step 2

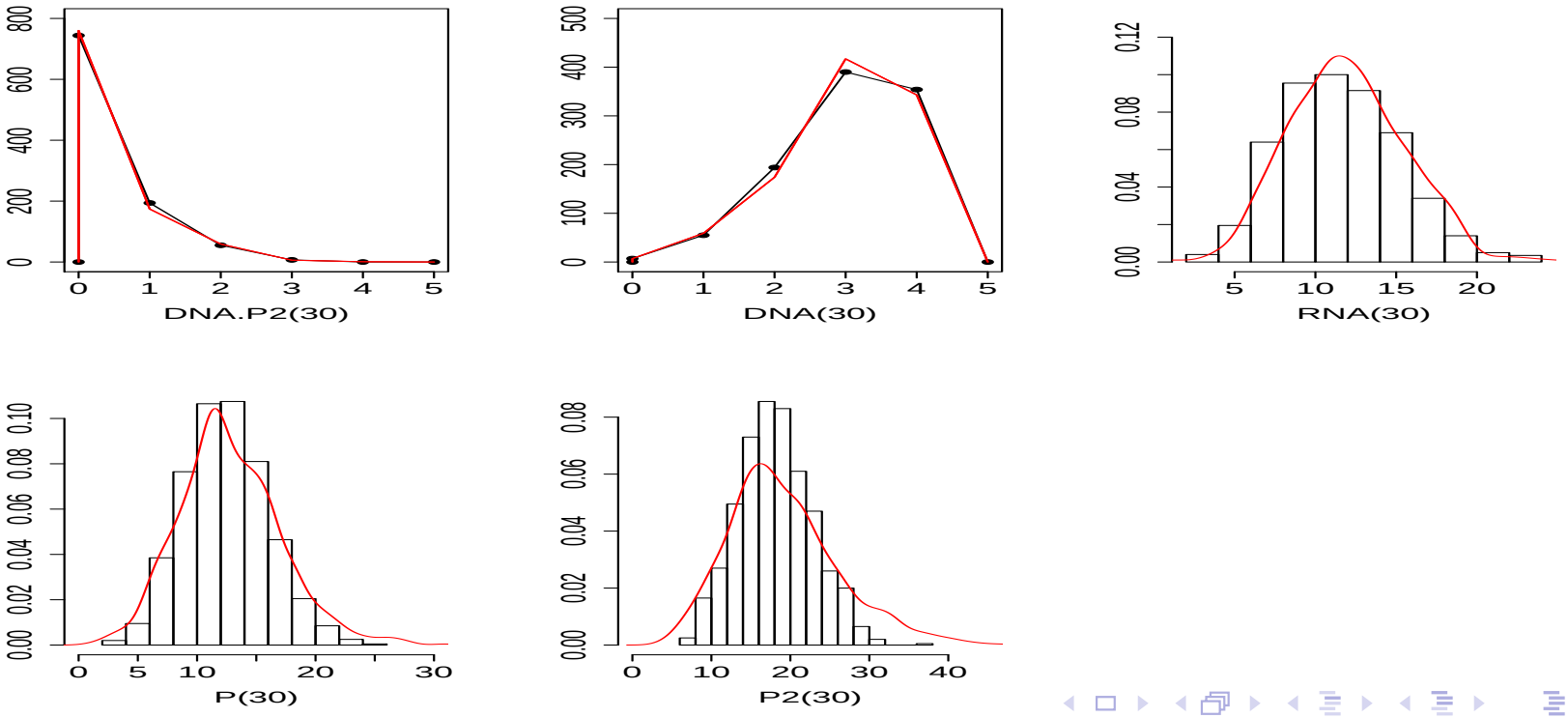
Hybrid Simulation cont'd

Remarks:

- The method is faster than Gillespie's exact method, since we use a time-discretisation for the fast species and we control the size of the time-step!
- Other hybrid simulation techniques are possible:
  - Discrete/ODE methods (see Kiehl, Mattheyses & Simmons (2004))
  - The maximal timestep method (see Puchalka & Kierzek (2004)) combines exact updating procedures for slow species with τ-leaping for the rest
- We can use the simulator inside an MCMC algorithm to make inference for c...

Performance of the Hybrid Simulator

Toy autoreg system – distributions of DNA, $\text{DNA} \cdot \text{P}_2$, RNA, P, P_2 at time 30 using 1000 simulations, red = hybrid with $\Delta t = 0.5$ and an Euler time step of 0.1, black = Gillespie. Computational cost scales as 1.2 : 1 in favour of the hybrid scheme



Bayesian Filtering

Suppose we have noisy observations $X_{0:(i-1)} = \{X(t) : t = 0, 1, \ldots, i-1\}$ where

$$X(t) = Y(t) + \epsilon, \qquad \epsilon \sim \mathrm{N}(0, \Sigma)$$

Goal: generate a sample from $\pi\left[c, Y(i)|X_{0:i}\right]$ given a new datum $X(i)$

$$\pi\left[c, Y(i)|X_{0:i}\right] \propto \int \pi\left[c, Y(i-1)|X_{0:i-1}\right] \pi\left[Y_{(i-1,i]}|c\right] \pi\left[Y(i)|X(i)\right] dY_{[i-1,i)}$$

where $Y_{(i-1,i]} = \{Y(t) : t \in (i-1, i]\}$ is the latent path in $[i-1, i]$

Idea: if we can sample $\pi\left[c, Y(i-1)|X_{0:i-1}\right]$ then we can use MCMC to sample the target $\pi\left[c, Y(i)|X_{0:i}\right]$

Bayesian Filtering

Suppose we have noisy observations $X_{0:(i-1)} = \{X(t) : t = 0, 1, \ldots, i-1\}$ where

$$X(t) = Y(t) + \epsilon, \qquad \epsilon \sim \mathrm{N}(0, \Sigma)$$

Goal: generate a sample from $\pi\left[c, Y(i)|X_{0:i}\right]$ given a new datum $X(i)$

$$\pi\left[c, Y(i)|X_{0:i}\right] \propto \int \pi\left[c, Y(i-1)|X_{0:i-1}\right] \pi\left[Y_{(i-1,i]}|c\right] \pi\left[Y(i)|X(i)\right] dY_{[i-1,i)}$$

where $Y_{(i-1,i]} = \{Y(t) : t \in (i-1, i]\}$ is the latent path in $[i-1, i]$

Idea: if we can sample $\pi\left[c, Y(i-1)|X_{0:i-1}\right]$ then we can use MCMC to sample the target $\pi\left[c, Y(i)|X_{0:i}\right]$

Bayesian Filtering

Suppose we have noisy observations $X_{0:(i-1)} = \{X(t) : t = 0, 1, \ldots, i-1\}$ where

$$X(t) = Y(t) + \epsilon, \qquad \epsilon \sim \mathrm{N}(0, \Sigma)$$

Goal: generate a sample from $\pi\left[c, Y(i)|X_{0:i}\right]$ given a new datum $X(i)$

$$\pi\left[c, Y(i)|X_{0:i}\right] \propto \int \pi\left[c, Y(i-1)|X_{0:i-1}\right] \pi\left[Y_{(i-1,i]}|c\right] \pi\left[Y(i)|X(i)\right] dY_{[i-1,i)}$$

where $Y_{(i-1,i]} = \{Y(t) : t \in (i-1, i]\}$ is the latent path in $[i-1, i]$

Idea: if we can sample $\pi\left[c, Y(i-1)|X_{0:i-1}\right]$ then we can use MCMC to sample the target $\pi\left[c, Y(i)|X_{0:i}\right]$

A Particle Approach

MCMC scheme:

- Propose $(c^*, Y(i-1)^*)' \sim \pi[\cdot|X_{0:i-1}]$
- Draw $Y^*_{(i-1,i]} \sim \pi[\cdot|c]$ using the hybrid simulator
- Accept/reject with probability

$$\min\left\{1, \frac{\pi[Y(i)^*|X(i)]}{\pi[Y(i)|X(i)]}\right\}$$

Comments:

- Since the hybrid simulator is used as a proposal process, we don't need to evaluate its associated likelihood
- Since $\pi[c, Y(i-1)|X_{0:i-1}]$ does not have analytic form (typically) we approximate this density with a cloud of points or particles, hence the term particle filter

A Particle Approach

MCMC scheme:

- Propose $(c^*, Y(i-1)^*)' \sim \pi[\cdot|X_{0:i-1}]$
- Draw $Y^*_{(i-1,i]} \sim \pi[\cdot|c]$ using the hybrid simulator
- Accept/reject with probability

$$\min\left\{1, \frac{\pi[Y(i)^*|X(i)]}{\pi[Y(i)|X(i)]}\right\}$$

Comments:

- Since the hybrid simulator is used as a proposal process, we don't need to evaluate its associated likelihood
- Since $\pi[c, Y(i-1)|X_{0:i-1}]$ does not have analytic form (typically) we approximate this density with a cloud of points or particles, hence the term particle filter

Toy Application Revisited

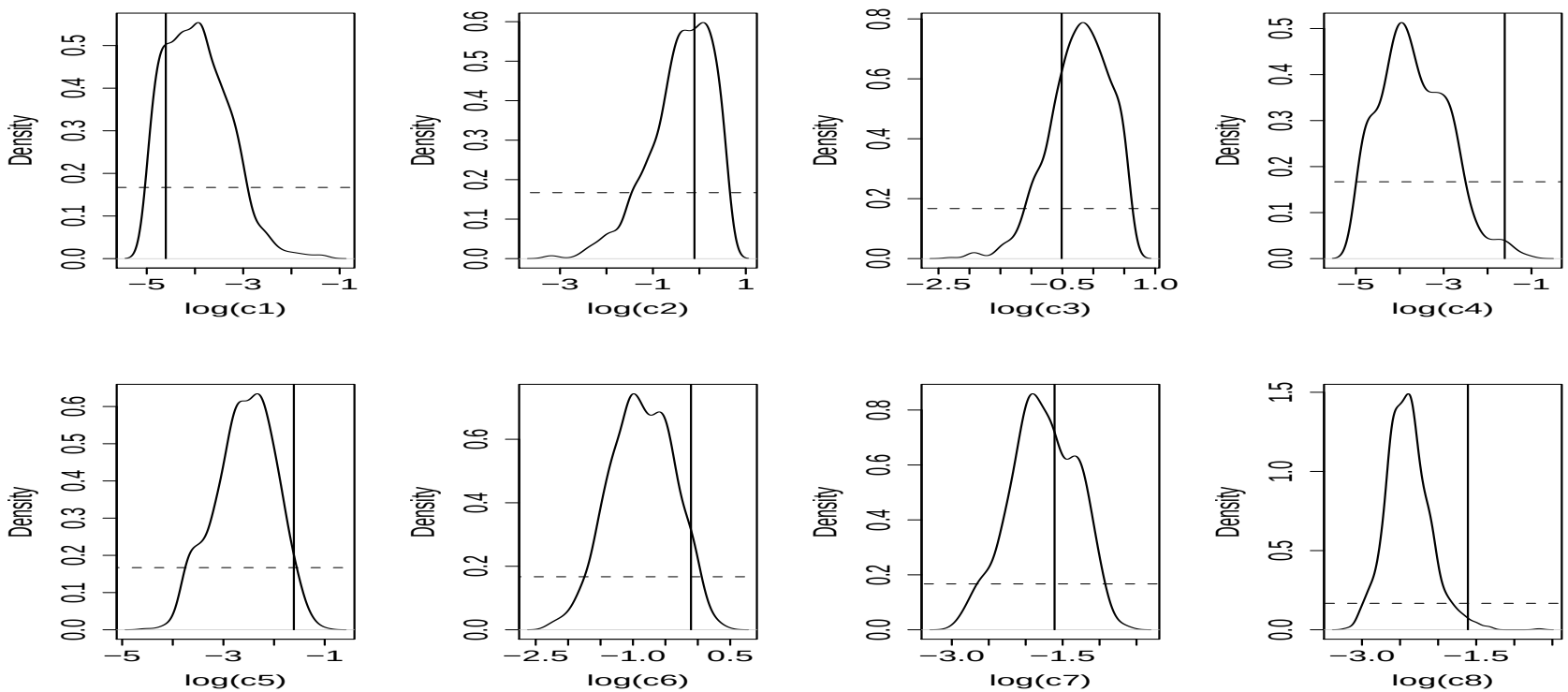
Prokaryotic Auto-Regulation:

R_1 :	$\mathrm{DNA} + \mathrm{P}_2$	$\longrightarrow$	$\mathrm{DNA} \cdot \mathrm{P}_2$	Repression
R_2 :	$\mathrm{DNA} \cdot \mathrm{P}_2$	$\longrightarrow$	$\mathrm{DNA} + \mathrm{P}_2$	
R_3 :	DNA	$\longrightarrow$	$\mathrm{DNA} + \mathrm{RNA}$	Transcription
R_4 :	RNA	$\longrightarrow$	$\mathrm{RNA} + \mathrm{P}$	Translation
R_5 :	2P	$\longrightarrow$	P_2	Dimerisation
R_6 :	P_2	$\longrightarrow$	2P	
R_7 :	RNA	$\longrightarrow$	$\emptyset$	Degradation
R_8 :	P	$\longrightarrow$	$\emptyset$	

- 50 observations simulated on $[0, 49]$ via the Gillespie algorithm
- Add a realisation of a standard Gaussian random variable to each observation
- Rate constants are $c = (0.01, 0.8, 0.6, 0.2, 0.2, 0.9, 0.2, 0.2)'$, take Uniform $\mathrm{U}(-5, 1)$ priors for $\log(c_i)$
- Run the particle filter to recover these

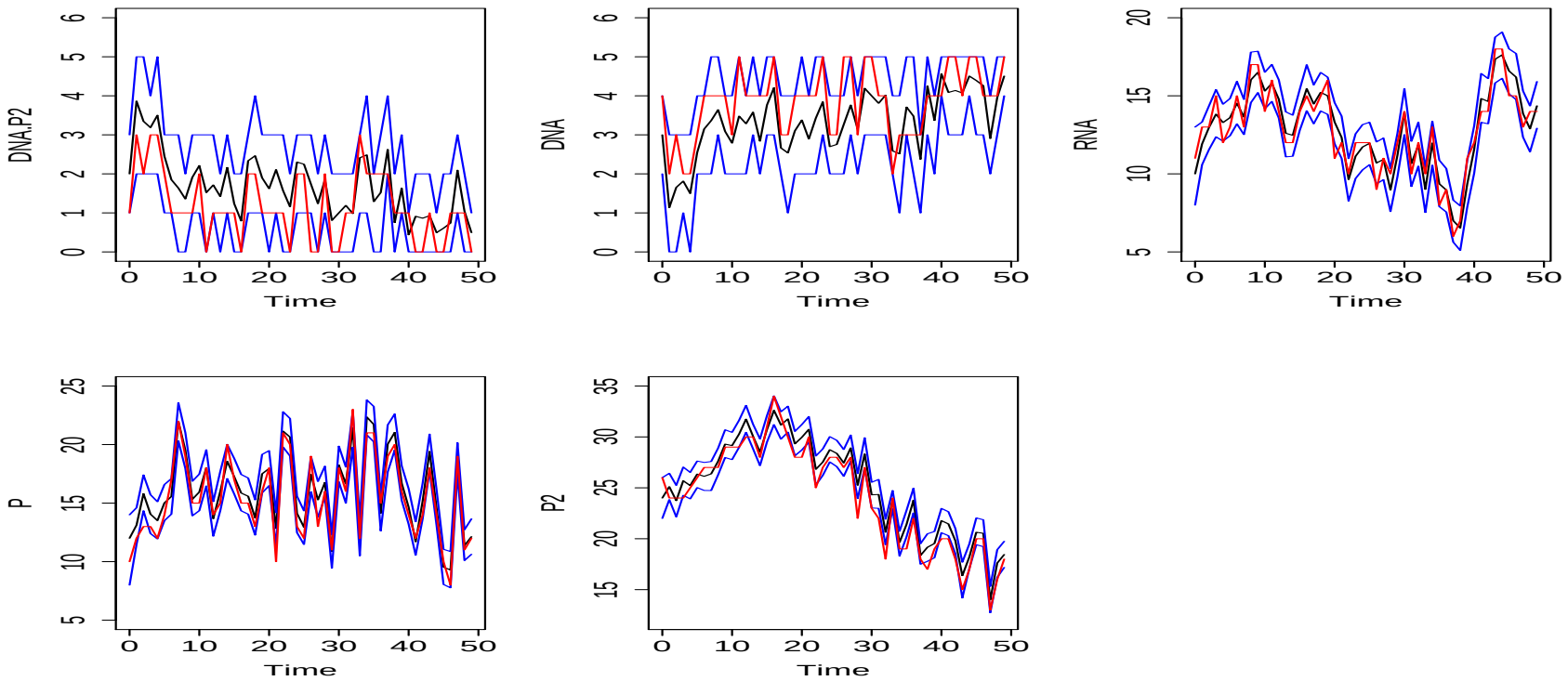
Results – 20,000 particles

Marginal posterior densities for each $\log(c_i)$, priors are indicated by the dotted line



Results – 20,000 particles

Flitered means (black), upper and lower 2.5% quantiles (blue) for $Y(t)$. True values are indicated by the red line



Summary

- Inferring rate constants that govern discrete stochastic kinetic models is computationally challenging
- It appears promising to consider an approximation of the model and perform exact inference using the approximate model
- A hybrid forwards simulator (or indeed any forwards simulator) can be used as a proposal process inside a particle filter
- Assessing the performance of the inference scheme, making comparisons with existing methods and extensions to partial observation remains of interest

Boys, R. J., Wilkinson, D.J. and T.B.L. Kirkwood (2008). Bayesian inference for a discretely observed stochastic kinetic model. *Statistics and Computing 18*, 125–135

Golightly, A. and D. J. Wilkinson (2009). Markov chain Monte Carlo algorithms for SDE parameter estimation. *To appear in Lawrence et al, Learning in Computational Systems Biology*. MIT press.

Golightly, A. and C. S. Gillespie (2009). Bayesian inference for generalized stochastic population growth models with application to aphids. *In submission*.

Heron, E. A., Finkenstadt, B. and D. A. Hand (2007). Bayesian inference for dynamic transcriptional regulation; the Hes1 system as a case study. *Bioinformatics 23*, 2596–2603.

Ruttor, A., Sanguinetti, G. and M. Opper (2009). Approximate inference for stochastic reaction systems. *To appear in Lawrence et al, Learning in Computational Systems Biology*. MIT press.

Salis, H. and Y. Kaznessis (2005). Accurate hybrid simulation of a system of coupled biochemical reactions. *Journal of Chemical Physics 122*, 054103

Wilkinson, D. J. (2006). *Stochastic Modelling for Systems Biology*. Chapman & Hall/CRC Press.

Contact details...

email: a.golightly@ncl.ac.uk
www: http://www.mas.ncl.ac.uk/~nag48/