Statistical Inference for Markov Jump Processes via Differential Geometric Monte Carlo and the LNA

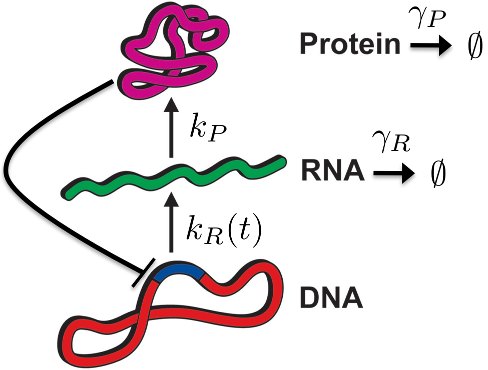
Mark A. Girolami

Department of Statistical Science
University College London

March 2012

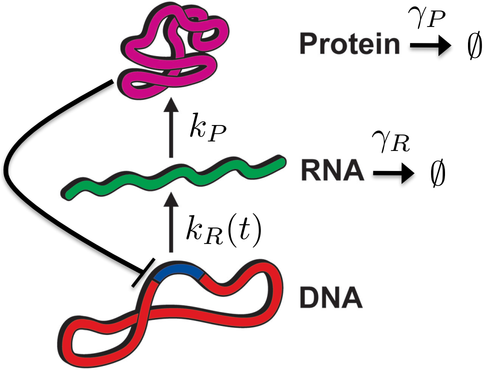
Outline

Single gene expression model



- Gene autoregulation.

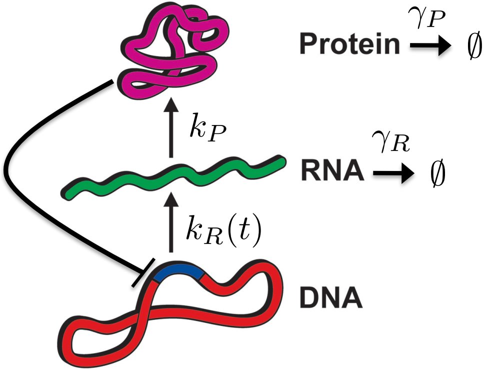
Single gene expression model



- Gene autoregulation.

$$
\begin{aligned}
R1 &: DNA \xrightarrow{k_R(t)} DNA + R \\
R2 &: R \xrightarrow{\gamma_R} \emptyset \\
R3 &: R \xrightarrow{k_P} R + P \\
R4 &: P \xrightarrow{\gamma_P} \emptyset.
\end{aligned}
$$

Single gene expression model

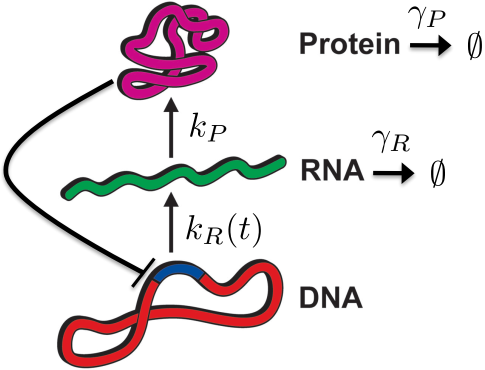


- Gene autoregulation.

$$
\begin{aligned}
R1 &: DNA \xrightarrow{k_R(t)} DNA + R \\
R2 &: R \xrightarrow{\gamma_R} \emptyset \\
R3 &: R \xrightarrow{k_P} R + P \\
R4 &: P \xrightarrow{\gamma_P} \emptyset.
\end{aligned}
$$

- System state $R(t), P(t))$

Single gene expression model



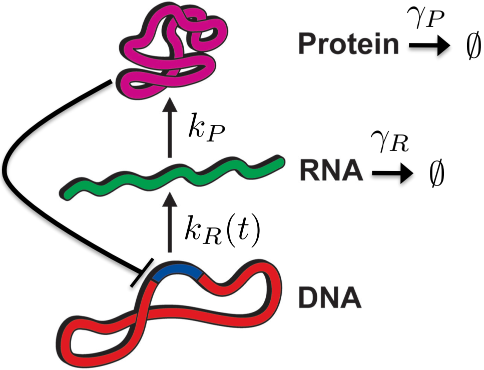
- Gene autoregulation.

$$\begin{aligned} R1 &: DNA \xrightarrow{k_R(t)} DNA + R \\ R2 &: R \xrightarrow{\gamma_R} \emptyset \\ R3 &: R \xrightarrow{k_P} R + P \\ R4 &: P \xrightarrow{\gamma_P} \emptyset. \end{aligned}$$

- System state $R(t), P(t))$

$$\boldsymbol{S} = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix}.$$

Single gene expression model



- Gene autoregulation.

$$\begin{aligned}
R1 : DNA &\xrightarrow{k_R(t)} DNA + R \\
R2 : R &\xrightarrow{\gamma_R} \emptyset \\
R3 : R &\xrightarrow{k_P} R + P \\
R4 : P &\xrightarrow{\gamma_P} \emptyset.
\end{aligned}$$

- System state $R(t), P(t))$

$$\boldsymbol{S} = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix}.$$

- $\boldsymbol{\theta} = (\gamma_R, \gamma_P, k_P, b_0, b_1, b_2, b_3)$

Motivation

- Continuous stochastic process on a discrete state space.

Motivation

- Continuous stochastic process on a discrete state space.
- Transitions happening at random times.

Motivation

- Continuous stochastic process on a discrete state space.
- Transitions happening at random times.
- Transition rates **depend** on current state and **unknown** rate parameters.

Motivation

- Continuous stochastic process on a discrete state space.
- Transitions happening at random times.
- Transition rates **depend** on current state and **unknown** rate parameters.
- Markov property, $p(\boldsymbol{x}_t|\boldsymbol{x}_{t-1},\ldots,\boldsymbol{x}_0) = p(\boldsymbol{x}_t|\boldsymbol{x}_{t-1})$

Motivation

- Continuous stochastic process on a discrete state space.
- Transitions happening at random times.
- Transition rates **depend** on current state and **unknown** rate parameters.
- Markov property, $p(\boldsymbol{x}_t|\boldsymbol{x}_{t-1},\ldots,\boldsymbol{x}_0) = p(\boldsymbol{x}_t|\boldsymbol{x}_{t-1})$
- $\boldsymbol{x}_t = (x_{1,t},\ldots,x_{N,t})^T$ system state at time t for N random variables.

Motivation

- Continuous stochastic process on a discrete state space.
- Transitions happening at random times.
- Transition rates **depend** on current state and **unknown** rate parameters.
- Markov property, $p(\boldsymbol{x}_t|\boldsymbol{x}_{t-1}, \ldots, \boldsymbol{x}_0) = p(\boldsymbol{x}_t|\boldsymbol{x}_{t-1})$
- $\boldsymbol{x}_t = (x_{1,t}, \ldots, x_{N,t})^T$ system state at time t for N random variables.
- State change vectors $\boldsymbol{s}_j = (s_{1,j}, \ldots, s_{N,j})^T, \; j \in \{1, \ldots, M\}$

Motivation

- Continuous stochastic process on a discrete state space.
- Transitions happening at random times.
- Transition rates **depend** on current state and **unknown** rate parameters.
- Markov property, $p(\boldsymbol{x}_t | \boldsymbol{x}_{t-1}, \ldots, \boldsymbol{x}_0) = p(\boldsymbol{x}_t | \boldsymbol{x}_{t-1})$
- $\boldsymbol{x}_t = (x_{1,t}, \ldots, x_{N,t})^T$ system state at time t for N random variables.
- State change vectors $\boldsymbol{s}_j = (s_{1,j}, \ldots, s_{N,j})^T, \ j \in \{1, \ldots, M\}$
- Transition rates, $f_j(\boldsymbol{x}, \boldsymbol{\theta})dt$ for state change j in the interval $[t, t + dt)$.

Motivation

- Continuous stochastic process on a discrete state space.
- Transitions happening at random times.
- Transition rates **depend** on current state and **unknown** rate parameters.
- Markov property, $p(\mathbf{x}_t|\mathbf{x}_{t-1},\ldots,\mathbf{x}_0) = p(\mathbf{x}_t|\mathbf{x}_{t-1})$
- $\mathbf{x}_t = (x_{1,t},\ldots,x_{N,t})^T$ system state at time t for N random variables.
- State change vectors $\mathbf{s}_j = (s_{1,j},\ldots,s_{N,j})^T,\ j \in \{1,\ldots,M\}$
- Transition rates, $f_j(\mathbf{x},\boldsymbol{\theta})dt$ for state change j in the interval $[t, t+dt)$.
- Master Equation

$$\frac{dp(\mathbf{x},t|\mathbf{x}_0,t_0)}{dt} = \sum_{j=1}^{M}\left[f_j(\mathbf{x}-\mathbf{s}_j,\boldsymbol{\theta},t)p(\mathbf{x}-\mathbf{s}_j,t|\mathbf{x}_0,t_0) - f_j(\mathbf{x},\boldsymbol{\theta},t)p(\mathbf{x},t|\mathbf{x}_0,t_0)\right].$$

Motivation

- Continuous stochastic process on a discrete state space.
- Transitions happening at random times.
- Transition rates **depend** on current state and **unknown** rate parameters.
- Markov property, $p(\boldsymbol{x}_t|\boldsymbol{x}_{t-1}, \dots, \boldsymbol{x}_0) = p(\boldsymbol{x}_t|\boldsymbol{x}_{t-1})$
- $\boldsymbol{x}_t = (x_{1,t}, \dots, x_{N,t})^T$ system state at time t for N random variables.
- State change vectors $\boldsymbol{s}_j = (s_{1,j}, \dots, s_{N,j})^T, \ j \in \{1, \dots, M\}$
- Transition rates, $f_j(\boldsymbol{x}, \boldsymbol{\theta})dt$ for state change j in the interval $[t, t + dt)$.
- Master Equation

$$\frac{dp(\boldsymbol{x}, t|\boldsymbol{x}_0, t_0)}{dt} = \sum_{j=1}^{M} \left[f_j(\boldsymbol{x} - \boldsymbol{s}_j, \boldsymbol{\theta}, t) p(\boldsymbol{x} - \boldsymbol{s}_j, t|\boldsymbol{x}_0, t_0) - f_j(\boldsymbol{x}, \boldsymbol{\theta}, t) p(\boldsymbol{x}, t|\boldsymbol{x}_0, t_0) \right].$$

- Intractible, simulate trajectories by a Stochastic Simulation Algorithm.

Diffusion Approximation

- Informal derivation, τ-leaping:

Diffusion Approximation

- Informal derivation, τ-leaping:
- Choose $\tau > 0$ such that:

$$
\begin{aligned}
f_j(\boldsymbol{x}_{t'}, \boldsymbol{\theta}) \approx f_j(\boldsymbol{x}_t, \boldsymbol{\theta}), &\quad \forall t' \in [t, t+\tau],\ \forall j \in [1, M] \qquad (1)\\
f_j(\boldsymbol{x}_t, \boldsymbol{\theta})\tau \gg 1, &\quad \forall j \in [1, M] \qquad (2)
\end{aligned}
$$

Diffusion Approximation

- Informal derivation, τ-leaping:
- Choose $\tau > 0$ such that:

$$
\begin{aligned}
f_j(\boldsymbol{x}_{t'}, \boldsymbol{\theta}) \approx f_j(\boldsymbol{x}_t, \boldsymbol{\theta}), &\quad \forall t' \in [t, t+\tau],\ \forall j \in [1, M] \qquad (1)\\
f_j(\boldsymbol{x}_t, \boldsymbol{\theta})\tau \gg 1, &\quad \forall j \in [1, M] \qquad (2)
\end{aligned}
$$

- Conditions (1) and (2) can be satisfied if $x_i \gg 1$.

Diffusion Approximation

- Informal derivation, τ-leaping:
- Choose $\tau > 0$ such that:

$$\begin{aligned} f_j(\boldsymbol{x}_{t'}, \boldsymbol{\theta}) &\approx f_j(\boldsymbol{x}_t, \boldsymbol{\theta}), \quad \forall t' \in [t, t+\tau],\ \forall j \in [1, M] \quad (1) \\ f_j(\boldsymbol{x}_t, \boldsymbol{\theta})\tau &\gg 1, \qquad\qquad\qquad \forall j \in [1, M] \quad (2) \end{aligned}$$

- Conditions (1) and (2) can be satisfied if $x_i \gg 1$.
- (1) implies that the number of transitions to states j are independently Poisson distributed with mean $f_j(\boldsymbol{x}_t, \boldsymbol{\theta})\tau$.

Diffusion Approximation

- Informal derivation, τ-leaping:
- Choose $\tau > 0$ such that:

$$f_j(\boldsymbol{x}_{t'}, \boldsymbol{\theta}) \approx f_j(\boldsymbol{x}_t, \boldsymbol{\theta}), \quad \forall t' \in [t, t+\tau], \; \forall j \in [1, M] \tag{1}$$

$$f_j(\boldsymbol{x}_t, \boldsymbol{\theta})\tau \gg 1, \qquad \forall j \in [1, M] \tag{2}$$

- Conditions (1) and (2) can be satisfied if $x_i \gg 1$.
- (1) implies that the number of transitions to states j are independently Poisson distributed with mean $f_j(\boldsymbol{x}_t, \boldsymbol{\theta})\tau$.
- (2) implies that the number of transitions can be reasonably approximated by a Normal distribution.

Diffusion Approximation

- Informal derivation, τ-leaping:
- Choose $\tau > 0$ such that:

$$f_j(\boldsymbol{x}_{t'}, \boldsymbol{\theta}) \approx f_j(\boldsymbol{x}_t, \boldsymbol{\theta}), \quad \forall t' \in [t, t+\tau], \; \forall j \in [1, M] \tag{1}$$

$$f_j(\boldsymbol{x}_t, \boldsymbol{\theta})\tau \gg 1, \qquad \forall j \in [1, M] \tag{2}$$

- Conditions (1) and (2) can be satisfied if $x_i \gg 1$.
- (1) implies that the number of transitions to states j are independently Poisson distributed with mean $f_j(\boldsymbol{x}_t, \boldsymbol{\theta})\tau$.
- (2) implies that the number of transitions can be reasonably approximated by a Normal distribution.
- Langevin Equation

$$d\boldsymbol{x}_t = \boldsymbol{S}\boldsymbol{f}(\boldsymbol{x}_t, \boldsymbol{\theta})dt + \frac{1}{\sqrt{\Omega}}\boldsymbol{S}\sqrt{\mathrm{diag}(\boldsymbol{f}(\boldsymbol{x}_t, \boldsymbol{\theta}))}dB_t \tag{3}$$

Linear Noise Approximation

- **Assume** that $\boldsymbol{x} = \phi + \frac{1}{\sqrt{\Omega}}\boldsymbol{\xi}$
- ϕ are deterministic and $\boldsymbol{\xi}$ stochastic variables.

Linear Noise Approximation

- **Assume** that $\boldsymbol{x} = \boldsymbol{\phi} + \frac{1}{\sqrt{\Omega}}\boldsymbol{\xi}$
- $\boldsymbol{\phi}$ are deterministic and $\boldsymbol{\xi}$ stochastic variables.
- Rewrite the transition functions and Taylor expand around $\boldsymbol{\phi}$

$$f_j(\boldsymbol{x}, \boldsymbol{\theta}) = f_j\left(\boldsymbol{\phi} + \frac{1}{\sqrt{\Omega}}\boldsymbol{\xi}, \boldsymbol{\theta}\right) = f_j(\boldsymbol{\phi}, \boldsymbol{\theta}) + \frac{1}{\sqrt{\Omega}}\sum_{i=1}^{N}\frac{\partial f_j(\boldsymbol{\phi}, \boldsymbol{\theta})}{\partial \phi_i}\xi_i + O(\Omega^{-1})$$

Linear Noise Approximation

- **Assume** that $\boldsymbol{x} = \boldsymbol{\phi} + \frac{1}{\sqrt{\Omega}}\boldsymbol{\xi}$
- ϕ are deterministic and $\boldsymbol{\xi}$ stochastic variables.
- Rewrite the transition functions and Taylor expand around ϕ

$$f_j(\boldsymbol{x}, \boldsymbol{\theta}) = f_j\left(\boldsymbol{\phi} + \frac{1}{\sqrt{\Omega}}\boldsymbol{\xi}, \boldsymbol{\theta}\right) = f_j(\boldsymbol{\phi}, \boldsymbol{\theta}) + \frac{1}{\sqrt{\Omega}} \sum_{i=1}^{N} \frac{\partial f_j(\boldsymbol{\phi}, \boldsymbol{\theta})}{\partial \phi_i} \xi_i + O(\Omega^{-1})$$

- Replace in the diffusion retain $O(1)$ terms for $d\phi$.

$$d\boldsymbol{\phi}_t = \boldsymbol{Sf}(\boldsymbol{\phi}_t, \boldsymbol{\theta})dt$$

Linear Noise Approximation

- **Assume** that $\boldsymbol{x} = \phi + \frac{1}{\sqrt{\Omega}}\boldsymbol{\xi}$
- ϕ are deterministic and ξ stochastic variables.
- Rewrite the transition functions and Taylor expand around ϕ

$$f_j(\boldsymbol{x}, \boldsymbol{\theta}) = f_j\left(\phi + \frac{1}{\sqrt{\Omega}}\boldsymbol{\xi}, \boldsymbol{\theta}\right) = f_j(\phi, \boldsymbol{\theta}) + \frac{1}{\sqrt{\Omega}}\sum_{i=1}^{N}\frac{\partial f_j(\phi, \boldsymbol{\theta})}{\partial \phi_i}\xi_i + O(\Omega^{-1})$$

- Replace in the diffusion retain $O(1)$ terms for $d\phi$.

$$d\phi_t = \boldsymbol{Sf}(\phi_t, \boldsymbol{\theta})dt$$

- Neglect any terms higher than $O(\frac{1}{\sqrt{\Omega}})$ for $d\xi$

$$d\xi_t = \boldsymbol{SJ_f}(\phi, \boldsymbol{\theta})\xi dt + \boldsymbol{S}\sqrt{\text{diag}(\boldsymbol{f}(\phi_t, \boldsymbol{\theta}))}dB_t$$

Linear Noise Approximation

- **Assume** that $\boldsymbol{x} = \phi + \frac{1}{\sqrt{\Omega}}\boldsymbol{\xi}$
- ϕ are deterministic and $\boldsymbol{\xi}$ stochastic variables.
- Rewrite the transition functions and Taylor expand around ϕ

$$f_j(\boldsymbol{x}, \boldsymbol{\theta}) = f_j\left(\phi + \frac{1}{\sqrt{\Omega}}\boldsymbol{\xi}, \boldsymbol{\theta}\right) = f_j(\phi, \boldsymbol{\theta}) + \frac{1}{\sqrt{\Omega}} \sum_{i=1}^{N} \frac{\partial f_j(\phi, \boldsymbol{\theta})}{\partial \phi_i} \xi_i + O(\Omega^{-1})$$

- Replace in the diffusion retain $O(1)$ terms for $d\phi$.

$$d\phi_t = \boldsymbol{Sf}(\phi_t, \boldsymbol{\theta})dt$$

- Neglect any terms higher than $O(\frac{1}{\sqrt{\Omega}})$ for $d\boldsymbol{\xi}$

$$d\boldsymbol{\xi}_t = \boldsymbol{SJ_f}(\phi, \boldsymbol{\theta})\boldsymbol{\xi} dt + \boldsymbol{S}\sqrt{\mathrm{diag}(\boldsymbol{f}(\phi_t, \boldsymbol{\theta}))} dB_t$$

- which is a linear SDE with analytic solution

Likelihood for the Linear Noise Approximation

$$p(\boldsymbol{x}^{(TS)}|\boldsymbol{\theta}) \propto \mathcal{N}(\boldsymbol{\mu}(\boldsymbol{\theta}), \boldsymbol{\Sigma}(\boldsymbol{\theta}))$$

Likelihood for the Linear Noise Approximation

$$p(\boldsymbol{x}^{(TS)}|\boldsymbol{\theta}) \propto \mathcal{N}(\boldsymbol{\mu}(\boldsymbol{\theta}), \boldsymbol{\Sigma}(\boldsymbol{\theta}))$$

- $\boldsymbol{\mu}(\boldsymbol{\theta}) = (\phi_{t1}, \ldots, \phi_{tn})^T$ a nN vector with solutions of the MRE.

Likelihood for the Linear Noise Approximation

$$p(\boldsymbol{x}^{(TS)}|\boldsymbol{\theta}) \propto \mathcal{N}(\boldsymbol{\mu}(\boldsymbol{\theta}), \boldsymbol{\Sigma}(\boldsymbol{\theta}))$$

- $\boldsymbol{\mu}(\boldsymbol{\theta}) = (\phi_{t1}, \ldots, \phi_{tn})^T$ a nN vector with solutions of the MRE.
- $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ a $nN \times nN$ block matrix with blocks $\boldsymbol{\Sigma}(\boldsymbol{\theta})^{i,j}$ $N \times N$

$$\boldsymbol{\Sigma}(\boldsymbol{\theta})^{i,j} = \begin{cases} \boldsymbol{V}_{t_i} & \text{if } i = j \\ cov(\boldsymbol{x}_{t_i}, \boldsymbol{x}_{t_j}) = \boldsymbol{\Sigma}(\boldsymbol{\theta})^{i,j-1}\boldsymbol{\Phi}(t_{j-1}, t_j)^T & \end{cases}$$

Likelihood for the Linear Noise Approximation

$$p(\boldsymbol{x}^{(TS)}|\boldsymbol{\theta}) \propto \mathcal{N}(\boldsymbol{\mu}(\boldsymbol{\theta}), \boldsymbol{\Sigma}(\boldsymbol{\theta}))$$

- $\boldsymbol{\mu}(\boldsymbol{\theta}) = (\phi_{t1}, \ldots, \phi_{tn})^T$ a nN vector with solutions of the MRE.
- $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ a $nN \times nN$ block matrix with blocks $\boldsymbol{\Sigma}(\boldsymbol{\theta})^{i,j}$ $N \times N$

$$\boldsymbol{\Sigma}(\boldsymbol{\theta})^{i,j} = \begin{cases} \boldsymbol{V}_{t_i} & \text{if } i = j \\ cov(\boldsymbol{x}_{t_i}, \boldsymbol{x}_{t_j}) = \boldsymbol{\Sigma}(\boldsymbol{\theta})^{i,j-1} \boldsymbol{\Phi}(t_{j-1}, t_j)^T & \end{cases}$$

- Fisher Information

$$FI(\boldsymbol{\theta})_{m,n} = \frac{\partial \boldsymbol{\mu}(\boldsymbol{\theta})^T}{\partial \theta_m} \boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}) \frac{\partial \boldsymbol{\mu}(\boldsymbol{\theta})}{\partial \theta_n} + \frac{1}{2} \operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}) \frac{\partial \boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial \theta_m} \boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}) \frac{\partial \boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial \theta_n} \right)$$

Likelihood for the Linear Noise Approximation

$$p(\boldsymbol{x}^{(TS)}|\boldsymbol{\theta}) \propto \mathcal{N}(\boldsymbol{\mu}(\boldsymbol{\theta}), \boldsymbol{\Sigma}(\boldsymbol{\theta}))$$

- $\boldsymbol{\mu}(\boldsymbol{\theta}) = (\phi_{t1}, \ldots, \phi_{tn})^T$ a nN vector with solutions of the MRE.
- $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ a $nN \times nN$ block matrix with blocks $\boldsymbol{\Sigma}(\boldsymbol{\theta})^{i,j}$ $N \times N$

$$\boldsymbol{\Sigma}(\boldsymbol{\theta})^{i,j} = \begin{cases} \boldsymbol{V}_{t_i} & \text{if } i = j \\ cov(\boldsymbol{x}_{t_i}, \boldsymbol{x}_{t_j}) = \boldsymbol{\Sigma}(\boldsymbol{\theta})^{i,j-1}\boldsymbol{\Phi}(t_{j-1}, t_j)^T \end{cases}$$

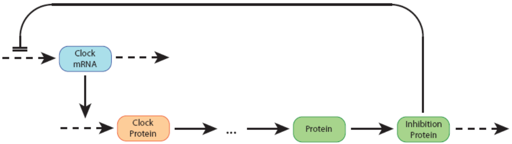
- Fisher Information

$$FI(\boldsymbol{\theta})_{m,n} = \frac{\partial \boldsymbol{\mu}(\boldsymbol{\theta})^T}{\partial \theta_m} \boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}) \frac{\partial \boldsymbol{\mu}(\boldsymbol{\theta})}{\partial \theta_n} + \frac{1}{2} \operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}) \frac{\partial \boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial \theta_m} \boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}) \frac{\partial \boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial \theta_n} \right)$$

- Augment the MRE for ϕ with the lower triangular elements of $\boldsymbol{V}$ and solve the augmented system with forward sensitivity analysis.

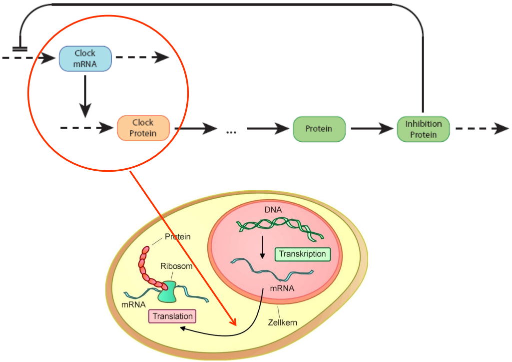
Mechanistic Models

Mechanistic modelling

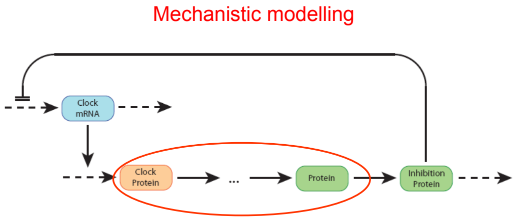


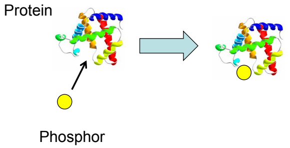
Mechanistic Models

Mechanistic modelling



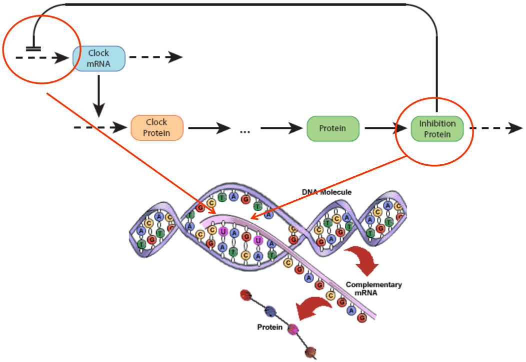
Mechanistic Models





Mechanistic Models

Mechanistic modelling



Mechanistic Models

Mechanistic modelling

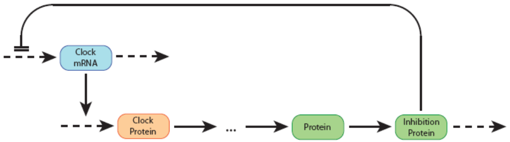


Figure: A diagram of the Goodwin Circadian Oscillator model network.

$$
\begin{aligned}
\frac{dx_1}{dt} &= \frac{k_1}{1+x_n^{\rho}} - m_1 x_1 \\
\frac{dx_2}{dt} &= k_2 x_1 - m_2 x_2 \\
&\vdots \\
\frac{dx_n}{dt} &= k_n x_{n-1} - m_n x_n
\end{aligned}
$$

Mechanistic Models

Mechanistic modelling

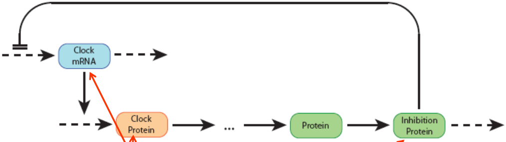


Figure: A diagram of the Goodwin Circadian Oscillator model network.

$$
\begin{aligned}
\frac{dx_1}{dt} &= \frac{k_1}{1 + x_n^{\rho}} - m_1 x_1 \\
\frac{dx_2}{dt} &= k_2 x_1 - m_2 x_2 \\
&\vdots \\
\frac{dx_n}{dt} &= k_n x_{n-1} - m_n x_n
\end{aligned}
$$

Mechanistic Models

Mechanistic modelling

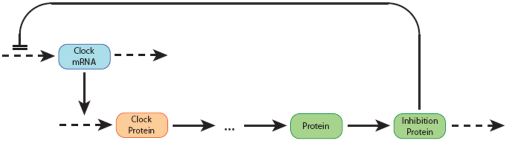


Figure: A diagram of the Goodwin Circadian Oscillator model network.

$$
\begin{aligned}
\frac{dx_1}{dt} &= \frac{k_1}{1+x_n^{\rho}} - m_1 x_1 \\
\frac{dx_2}{dt} &= k_2 x_1 - m_2 x_2 \\
&\vdots \\
\frac{dx_n}{dt} &= k_n x_{n-1} - m_n x_n
\end{aligned}
$$

Mechanistic Models

Mechanistic modelling

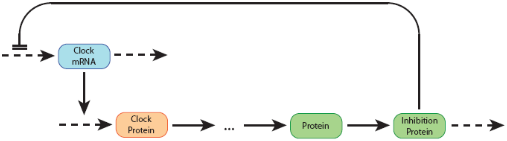


Figure: A diagram of the Goodwin Circadian Oscillator model network.

Dependence on kinetic parameters → θ

$$
\begin{aligned}
\frac{dx_1}{dt} &= \frac{k_1}{1 + x_n^{\rho}} - m_1 x_1 \\
\frac{dx_2}{dt} &= k_2 x_1 - m_2 x_2 \\
&\vdots \\
\frac{dx_n}{dt} &= k_n x_{n-1} - m_n x_n
\end{aligned}
$$

Mechanistic Models

Necessary condition for stable oscillations

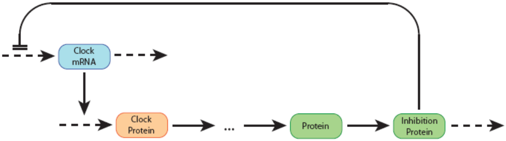


Figure: A diagram of the Goodwin Circadian Oscillator model network.

$$
\begin{aligned}
\frac{dx_1}{dt} &= \frac{k_1}{1 + x_n^{\rho}} - m_1 x_1 \\
\frac{dx_2}{dt} &= k_2 x_1 - m_2 x_2 \\
&\vdots \\
\frac{dx_n}{dt} &= k_n x_{n-1} - m_n x_n
\end{aligned}
$$

> 2

> 8

Mechanistic Models

Example: log likelihood landscape

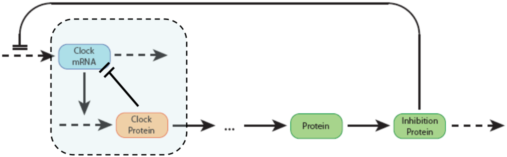


Figure: A diagram of the Goodwin Circadian Oscillator model network.

$$
\begin{aligned}
\frac{dx_1}{dt} &= \frac{k_1}{1 + x_n^{\rho}} - m_1 x_1 \\
\frac{dx_2}{dt} &= k_2 x_1 - m_2 x_2 \\
&\vdots \\
\frac{dx_n}{dt} &= k_n x_{n-1} - m_n x_n
\end{aligned}
$$

Mechanistic Models

Example: log likelihood landscape

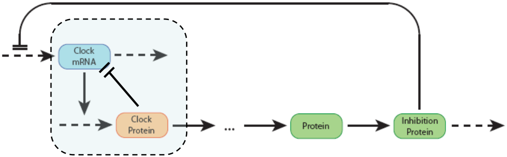


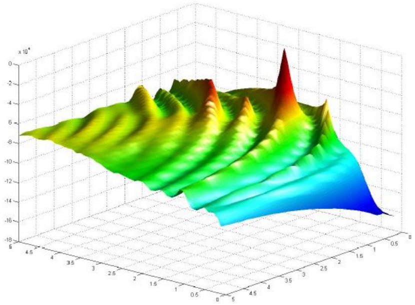
Figure: A diagram of the Goodwin Circadian Oscillator model network.

$$
\begin{aligned}
\frac{dx_1}{dt} &= \frac{k_1}{1 + x_n^{\rho}} - m_1 x_1 \\
\frac{dx_2}{dt} &= k_2 x_1 - m_2 x_2 \\
&\vdots \\
\frac{dx_n}{dt} &= k_n x_{n-1} - m_n x_n
\end{aligned}
$$

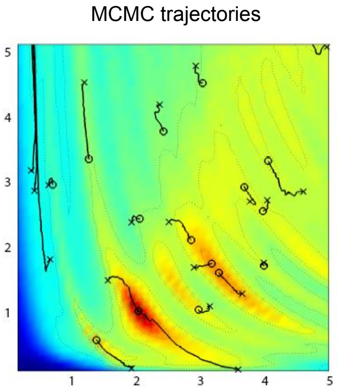
2 parameters free,
others fixed

Mechanistic Models

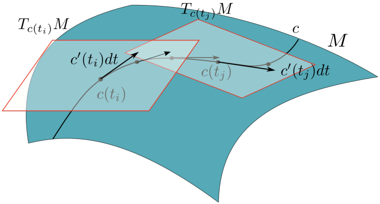
Posterior landscape $P(\boldsymbol{\theta}|\mathcal{M},\mathcal{D})$



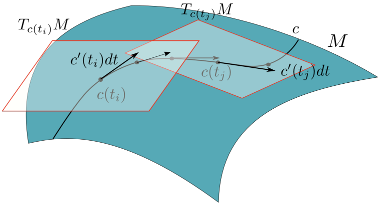
Mechanistic Models



Geometric Concepts in MCMC

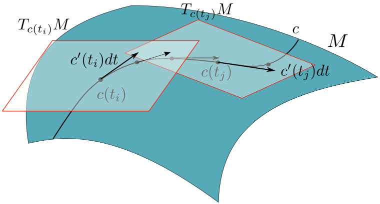


Geometric Concepts in MCMC



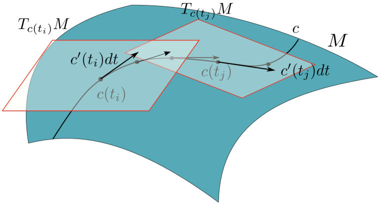
- Tangent space - local metric defined by $\delta\boldsymbol{\theta}^{\mathsf{T}}\mathbf{G}(\boldsymbol{\theta})\delta\boldsymbol{\theta} = g_{kl}\delta\theta_k\delta\theta_l$

Geometric Concepts in MCMC



- Tangent space - local metric defined by $\delta\boldsymbol{\theta}^{\mathsf{T}}\mathbf{G}(\boldsymbol{\theta})\delta\boldsymbol{\theta} = g_{kl}\delta\theta_k\delta\theta_l$
- Christoffel symbols - characterise Levi-Civita connection on manifold

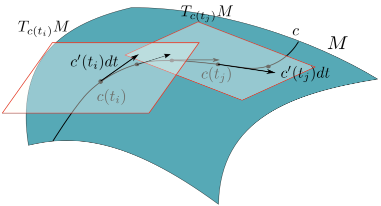
Geometric Concepts in MCMC



- Tangent space - local metric defined by $\delta\boldsymbol{\theta}^{\mathsf{T}}\mathbf{G}(\boldsymbol{\theta})\delta\boldsymbol{\theta} = g_{kl}\delta\theta_k\delta\theta_l$
- Christoffel symbols - characterise Levi-Civita connection on manifold

$$\Gamma^i_{kl} = \frac{1}{2}\sum_m g^{im}\left(\frac{\partial g_{mk}}{\partial\theta^l} + \frac{\partial g_{ml}}{\partial\theta^k} - \frac{\partial g_{kl}}{\partial\theta^m}\right)$$

Geometric Concepts in MCMC

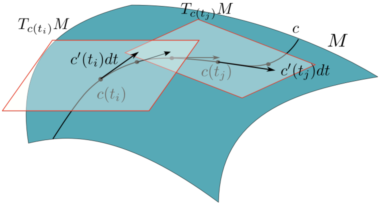


- Tangent space - local metric defined by $\delta\boldsymbol{\theta}^{\mathsf{T}}\mathbf{G}(\boldsymbol{\theta})\delta\boldsymbol{\theta} = g_{kl}\delta\theta_k\delta\theta_l$
- Christoffel symbols - characterise Levi-Civita connection on manifold

$$\Gamma^i_{kl} = \frac{1}{2}\sum_m g^{im}\left(\frac{\partial g_{mk}}{\partial\theta^l} + \frac{\partial g_{ml}}{\partial\theta^k} - \frac{\partial g_{kl}}{\partial\theta^m}\right)$$

- Geodesics - shortest path between two points on manifold

Geometric Concepts in MCMC



- Tangent space - local metric defined by $\delta\boldsymbol{\theta}^{\mathsf{T}}\mathbf{G}(\boldsymbol{\theta})\delta\boldsymbol{\theta} = g_{kl}\delta\theta_k\delta\theta_l$
- Christoffel symbols - characterise Levi-Civita connection on manifold

$$\Gamma^i_{kl} = \frac{1}{2}\sum_m g^{im}\left(\frac{\partial g_{mk}}{\partial\theta^l} + \frac{\partial g_{ml}}{\partial\theta^k} - \frac{\partial g_{kl}}{\partial\theta^m}\right)$$

- Geodesics - shortest path between two points on manifold

$$\frac{d^2\theta^i}{dt^2} + \sum_{k,l}\Gamma^i_{kl}\frac{d\theta^k}{dt}\frac{d\theta^l}{dt} = 0$$

Journal of the
Royal Statistical Society
SERIES B
Statistical
Methodology

J. R. Statist. Soc. B (2011)
73, *Part* 2, *pp.* 123–214

Riemann manifold Langevin and Hamiltonian Monte Carlo methods

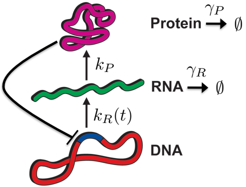
Mark Girolami and Ben Calderhead

University College London, UK

[*Read before* The Royal Statistical Society *at a meeting organized by the* Research Section *on Wednesday, October 13th, 2010*, Professor D. M. Titterington *in the Chair*]

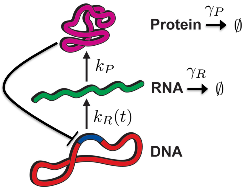
Summary. The paper proposes Metropolis adjusted Langevin and Hamiltonian Monte Carlo sampling methods defined on the Riemann manifold to resolve the shortcomings of existing Monte Carlo algorithms when sampling from target densities that may be high dimensional and exhibit strong correlations. The methods provide fully automated adaptation mechanisms that circumvent the costly pilot runs that are required to tune proposal densities for Metropolis–Hastings or indeed Hamiltonian Monte Carlo and Metropolis adjusted Langevin algorithms. This allows for highly efficient sampling even in very high dimensions where different scalings may be required for the transient and stationary phases of the Markov chain. The methodology proposed exploits the Riemann geometry of the parameter space of statistical models and thus automat-

Single gene expression model



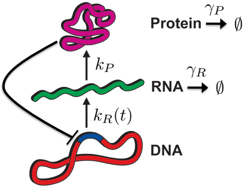
- Single gene expression model with negative feedback.

Single gene expression model



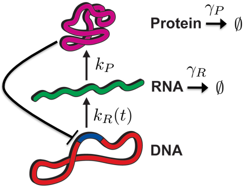
- Single gene expression model with negative feedback.
- System state $(R(t), P(t))^T$ models the population of RNA and protein.

Single gene expression model



- Single gene expression model with negative feedback.
- System state $(R(t), P(t))^T$ models the population of RNA and protein.
- $k_R(P, t) = (b_0 \exp(-b_1(t - b_2)^2) + b_3)/(1 + (P/H)^{n_H})$
- $H = b_3 k_P/(2\gamma_R\gamma_P), \quad n_H = 1/2$

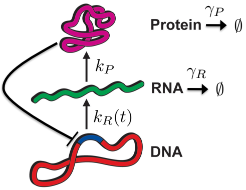
Single gene expression model



- Single gene expression model with negative feedback.
- System state $(R(t), P(t))^T$ models the population of RNA and protein.
- $k_R(P, t) =$ $(b_0 \exp(-b_1(t - b_2)^2) + b_3)/(1 + (P/H)^{n_H})$
- $H = b_3 k_P/(2\gamma_R\gamma_P), \quad n_H = 1/2$

$$\boldsymbol{S} = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix}$$

Single gene expression model



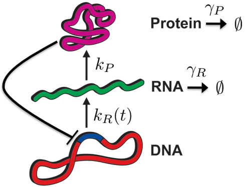
- Single gene expression model with negative feedback.
- System state $(R(t), P(t))^T$ models the population of RNA and protein.
- $k_R(P,t) = (b_0 \exp(-b_1(t-b_2)^2) + b_3)/(1+(P/H)^{n_H})$
- $H = b_3 k_P/(2\gamma_R\gamma_P), \quad n_H = 1/2$

$$\boldsymbol{S} = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix}$$

- Deterministic MRE $\phi = (\phi_R, \phi_P)^T$

$$\begin{aligned} d\phi_R/dt &= k_R(\phi_P, t) - \gamma_R\phi_R \\ d\phi_P/dt &= k_P\phi_R - \gamma_P\phi_P \end{aligned}$$

Single gene expression model



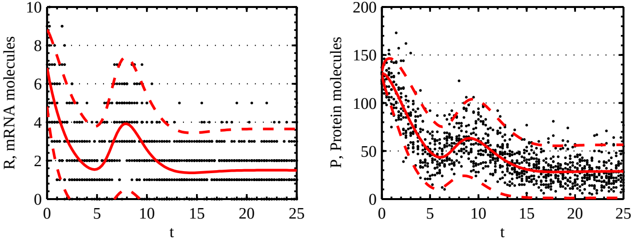
- Single gene expression model with negative feedback.
- System state $(R(t), P(t))^T$ models the population of RNA and protein.
- $k_R(P, t) = (b_0 \exp(-b_1(t - b_2)^2) + b_3)/(1 + (P/H)^{n_H})$
- $H = b_3 k_P/(2\gamma_R\gamma_P), \quad n_H = 1/2$

$$\boldsymbol{S} = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix}$$

- Deterministic MRE $\phi = (\phi_R, \phi_P)^T$

$$\begin{aligned} d\phi_R/dt &= k_R(\phi_P, t) - \gamma_R\phi_R \\ d\phi_P/dt &= k_P\phi_R - \gamma_P\phi_P \end{aligned}$$

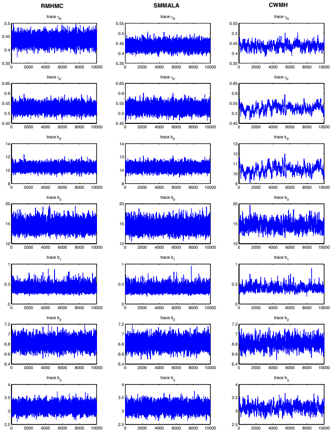
Simulated Data



- Simulated data generated with SSA.
- 10 independent sample paths for each time point.
- Parameters set to

γ_R	γ_P	k_P	b_0	b_1	b_2	b_3
0.44	0.52	10.0	15.0	0.40	7.0	3.0

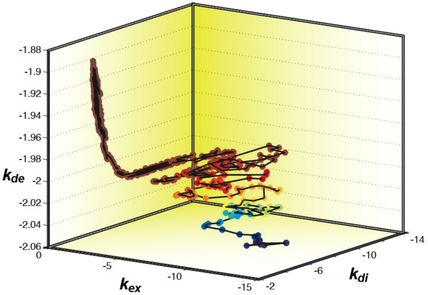
Trace Plots



Effective Sample Size

10,000 posterior samples	γ_R	γ_P	k_P	b_0	b_1	b_2	b_3
RMHMC	6532	6593	6614	5112	5384	6595	6642
SMMALA	2990	3270	3454	3124	3164	3316	3195
CWMH	201	71	73	465	339	420	239

Nonlinear Dynamic System - Circadian Clock Gene Control



Conclusions

- MJP common tool to describe many phenomena in physical and life sciences.
- LNA provides a useful approximation in appropriate operational regimes.
- Decouples deterministic and stochastic characteristics of model.
- Statistical inference remains a formidable challenge over such models.
- Exploitation of schoolboy differential geometry in MCMC provides effective inference tool.
- Phil.Trans paper describes a number of larger scale scenarios.