Bayesian analysis of single-molecule experimental data

S.C. Kou, S. Xie, J. Liu

Mingwei Tang

June 15, 2015

Overview

Molecule: DNA hairpin

- Single-stranded nucleic acid with two ends
- Two states: closed and open
- Transitions between two state

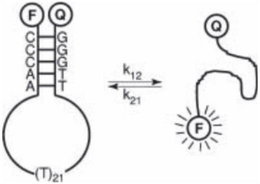


Figure: closed(left),open(right)

Molecule: DNA hairpin

- Single-stranded nucleic acid with two ends
- Two states: closed and open
- Transitions between two state

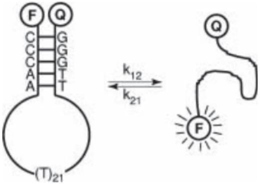
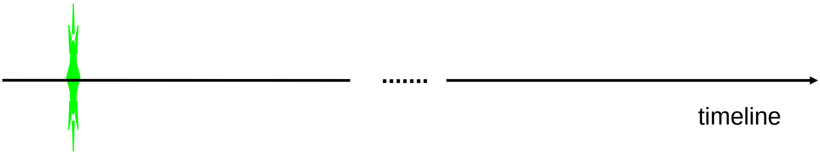
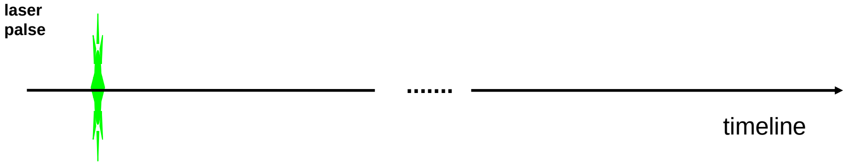


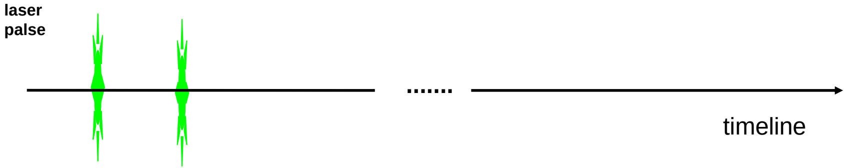
Figure: closed(left),open(right)

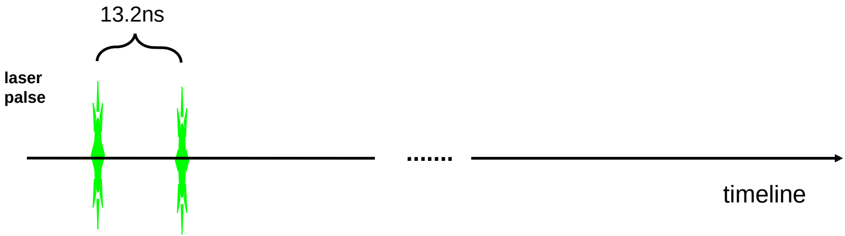
- Question: How often does the transition happen?
- The state can not be observed

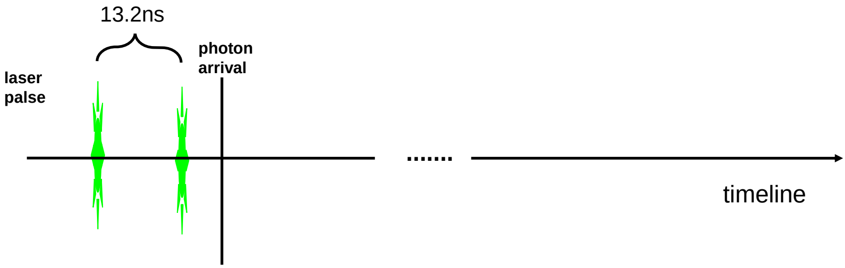

timeline

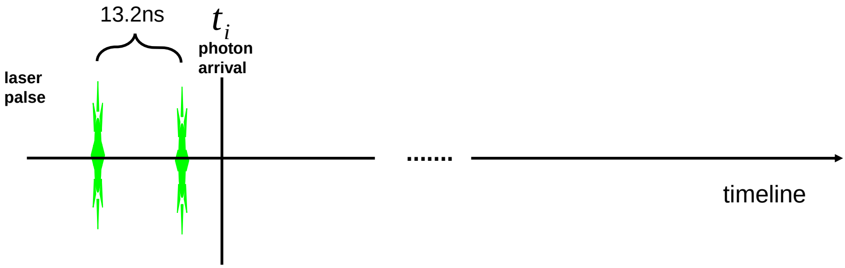

timeline

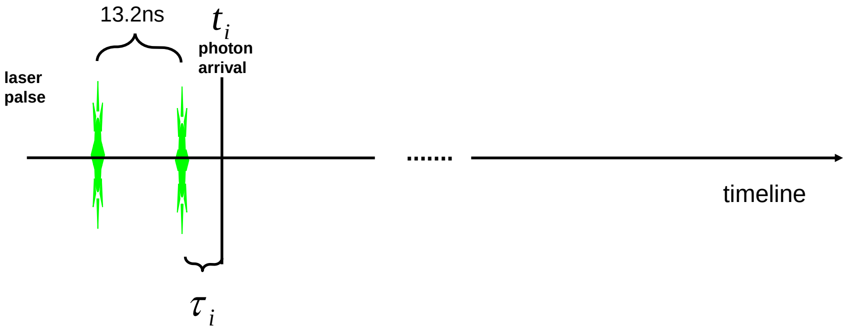

laser
palse
......
timeline

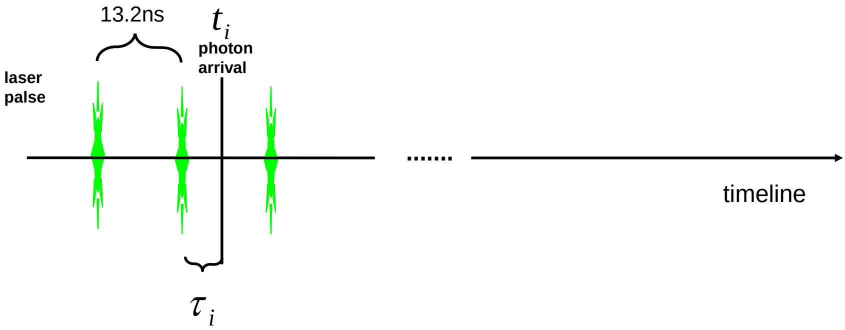

laser
palse
......
timeline

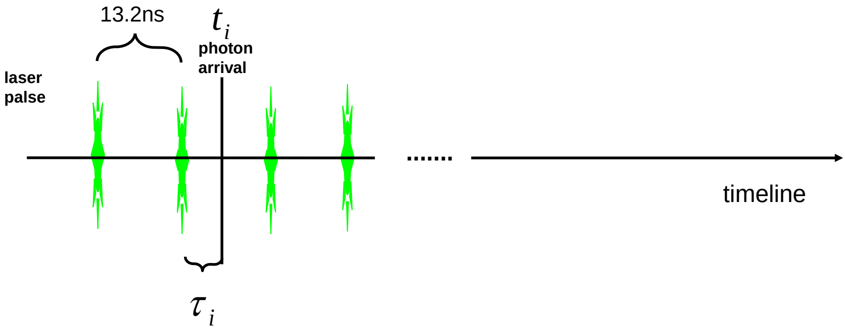

13.2ns
laser
palse
....
timeline

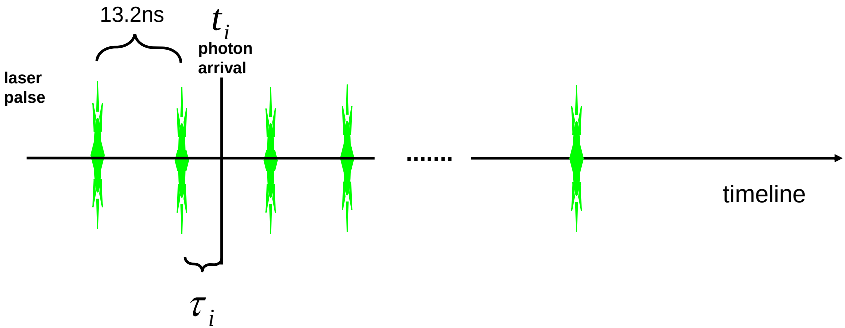

13.2ns
photon
arrival
laser
palse
....
timeline


13.2ns
t_i
photon
arrival
laser
palse
......
timeline

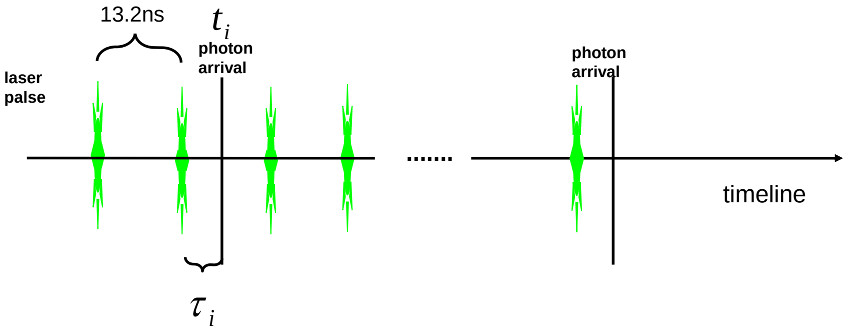

13.2ns
t_i
photon
arrival
laser
palse
......
timeline
τ_i


13.2ns
t_i
photon
arrival
laser
palse
......
timeline
τ_i

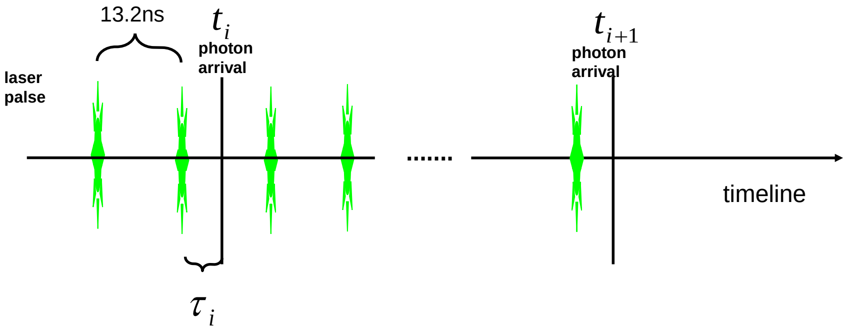

13.2ns
t_i
photon
arrival
laser
palse
......
timeline
τ_i



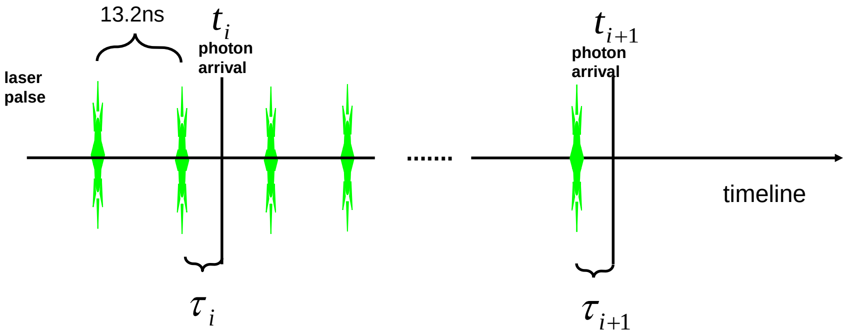
13.2ns
t_i
photon
arrival
laser
palse
timeline
τ_i

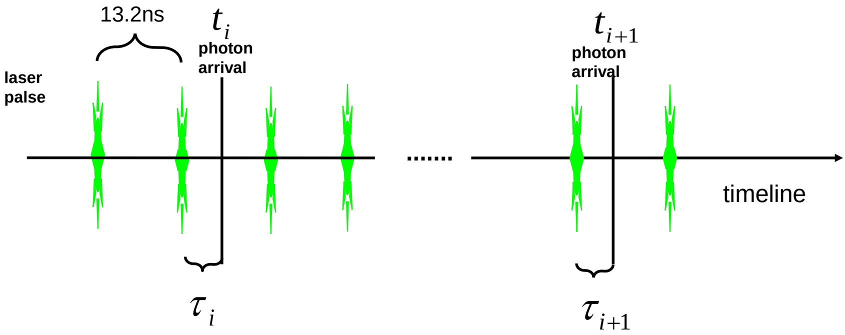


13.2ns
t_i
photon arrival
laser palse
photon arrival
......
timeline
τ_i

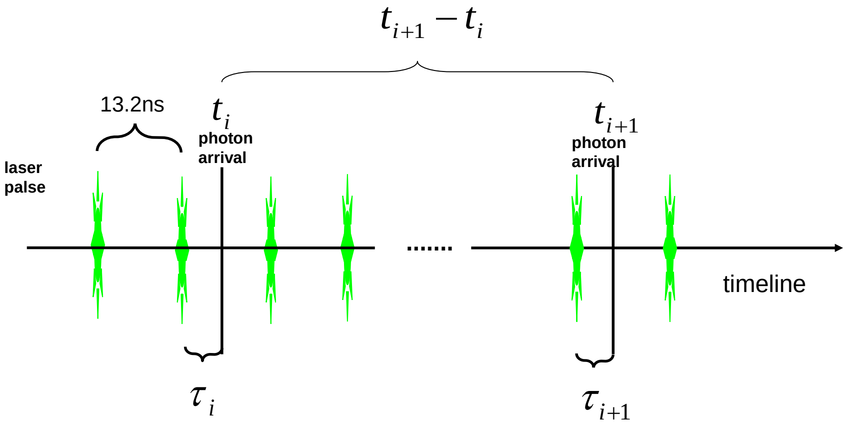


13.2ns
t_i
photon arrival
t_{i+1}
photon arrival
laser pulse
....
timeline
τ_i

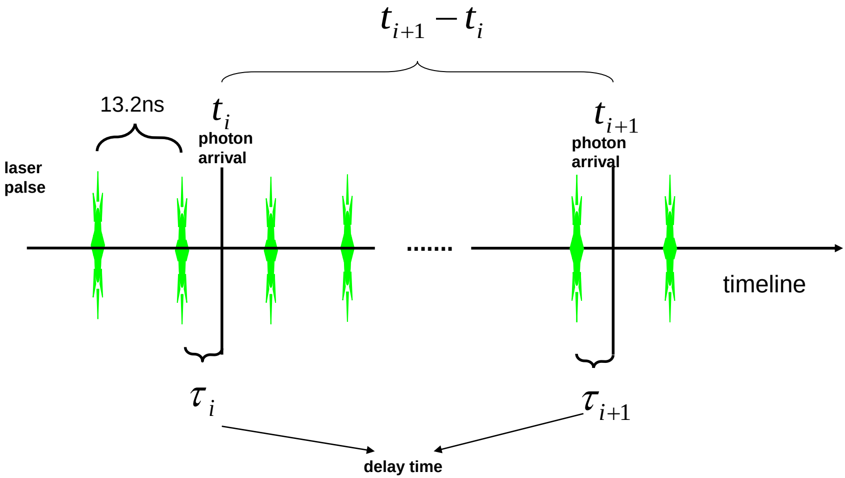

13.2ns
t_i
photon
arrival
laser
palse
t_{i+1}
photon
arrival
....
timeline
τ_i
τ_{i+1}



13.2ns
t_i
photon arrival
t_{i+1}
photon arrival
laser palse
......
timeline
τ_i
τ_{i+1}



$t_{i+1} - t_i$
13.2ns
t_i
photon
arrival
t_{i+1}
photon
arrival
laser
palse
....
timeline
τ_i
τ_{i+1}



$t_{i+1} - t_i$
13.2ns
t_i
photon arrival
t_{i+1}
photon arrival
laser palse
....
timeline
τ_i
τ_{i+1}
delay time

Flurescence lifetime experiments

- The arrival time and delay time depends on the DNA state
  - Closed state: less arrivals and shorter delay time
- Goal:
  1. Model the state transition
  2. Make inference on the parameters related to photon arrival rate and state transition rate

Model 1: Two-state model

Model 1: Two-state model

- The transition : continuous-time Markov chain
  Infinitesimal generator

$$\mathbf{Q} = \begin{pmatrix} -k_{12} & k_{12} \\ k_{21} & -k_{21} \end{pmatrix}$$

Model 1: Two-state model

- The transition : continuous-time Markov chain
  Infinitesimal generator

$$\mathbf{Q} = \begin{pmatrix} -k_{12} & k_{12} \\ k_{21} & -k_{21} \end{pmatrix}$$

- At $t = 0$, start from stationary distribution
  $\pi = (\pi_1, \pi_2) = (\frac{k_{21}}{k_{12} + k_{12}}, \frac{k_{12}}{k_{12} + k_{12}})$
- Use $k = k_{12} + k_{21}$ and π_1 for the transition parameter

Model 1: Two-state model

- The transition : continuous-time Markov chain
  Infinitesimal generator

$$\mathbf{Q} = \begin{pmatrix} -k_{12} & k_{12} \\ k_{21} & -k_{21} \end{pmatrix}$$

- At $t = 0$, start from stationary distribution
  $\pi = (\pi_1, \pi_2) = (\dfrac{k_{21}}{k_{12} + k_{12}}, \dfrac{k_{12}}{k_{12} + k_{12}})$
- Use $k = k_{12} + k_{21}$ and π_1 for the transition parameter
- State variable $\gamma(t)$: $\gamma(t) = \begin{cases} a & \text{Open state at time } t \\ b & \text{Closed state at time } t \end{cases}$
  where $a > b > 0$

Data Oberseved $(\mathbf{t}, \boldsymbol{\tau})$

- Photon arrival time t_i, $t_0 < t_1 < \cdots < t_n$
  - Counting process from non-homogeneous Poisson Process
  - Rate $\lambda(t) = A_0/\gamma(t)$
  - $A_0 > 0$: Photon arrival intensity
- Delay time τ_i associated with, t_i $\tau_0, \ldots, \tau_n$
  - $[\tau_i|\gamma(t_i)] \sim \text{Exp}(\gamma(t_i))$

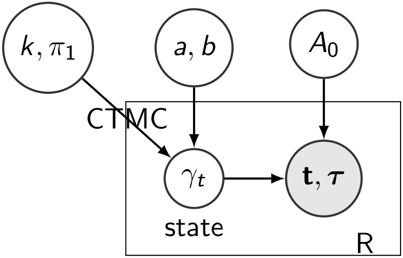
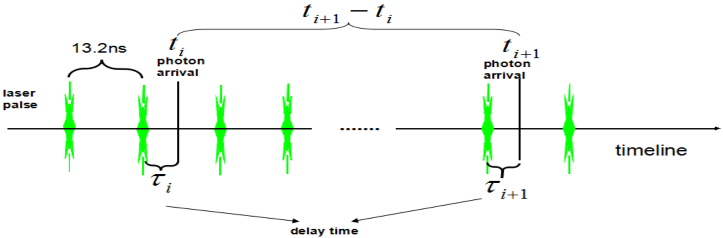


Figure: Generative View of the model

Likehood calculation

- Y denote the number of arrivals at time t.
  $\Delta Y_t = Y(t+dt) - Y(t)$
- Likelihood construction $L(\mathbf{t}, \boldsymbol{\tau}, \gamma|\theta)$
  - Assumption: $t_{i+1} - t_i \perp\!\!\!\perp \tau_i | \gamma(t_i)$
  - arrival time t_i $\quad P(\Delta Y_{t_i} = 1 | \gamma_{t_i}) = \frac{A_0}{\gamma(t_i)} dt$
  - delay time τ_i $\quad P(\tau_i | \Delta Y_{t_i} = 1, \gamma_{t_i}) = \gamma(t_i) \exp(-\gamma(t_i)\tau_i)$
  - no photon arrives in (t_i, t_{i+1}):

$$P\left(Y_{t_{i+1}}^{-} - Y_{t_i} = 0, \gamma(t_{i+1}) | \gamma(t_i)\right)$$



No arrival probability

Theorem

Let Y_t denotes the total number of arrivals at interval $[0, t)$. Then

$$P\left(Y^{-}_{t_{i+1}} - Y_{t_i} = 0, \gamma(t_{i+1}) | \gamma(t_i)\right)$$
$$= [\exp(\mathbf{Q} - \mathbf{H})(t_{i+1} - t_i)]_{(\gamma(t_i), \gamma(t_{i+1}))}$$

where $\mathbf{H} = diag(A_0/a, A_0/b)$ *rate for the arrival time*

- Intuition: Kolmogorov forward equation and ODE

Goal: Inference on parameters

- Parameters $\boldsymbol{\theta} = (a, b, \pi_1, k, A_0)$
- Likelihood function

$$
\begin{aligned}
L(\mathbf{t}, \boldsymbol{\tau}|\theta) &= \sum_{\gamma} L(\mathbf{t}, \boldsymbol{\tau}, \gamma|\theta) \\
&= (\pi_1, \pi_2)\mathbf{D}_0\mathbf{H}\left[\prod_{i=0}^{n-1} \exp\{(\mathbf{Q}-\mathbf{H})(t_{i+1}-t_i)\}\mathbf{D}_{i+1}\mathbf{H}\right]\begin{pmatrix} 1 \\ 1 \end{pmatrix}
\end{aligned}
$$

where $\mathbf{D}_i = \mathrm{diag}(a\exp(-a\tau_i), b\exp(-b\tau_i))$ density for the delay time

Posterior sampling by MCMC

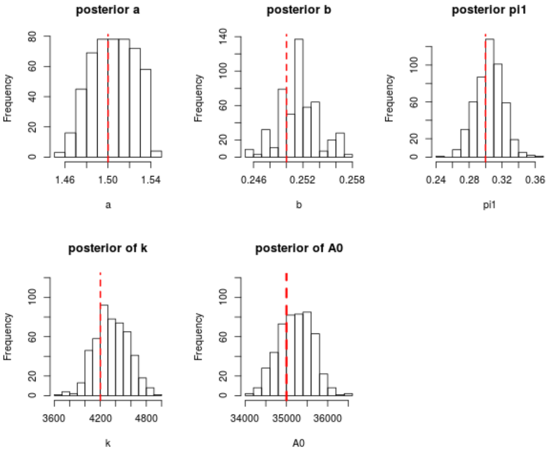
- $\eta(\boldsymbol{\theta})$ be the prior distribution
- Posterior distribution

$$P(\boldsymbol{\theta}|\mathbf{t},\boldsymbol{\tau}) \propto \eta(\boldsymbol{\theta})L(\mathbf{t},\boldsymbol{\tau}|\boldsymbol{\theta})$$

- Direct sampling is impossible
- The posterior can be sampled by Metropolis-Hasting algorithm

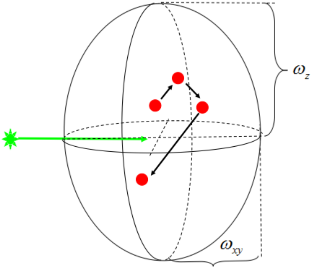
Simulations

- 5000 iterations, throw first 2500, draw a sample every 5 iterations
- the posterior sample covers the true parameter



A question with A_0

- Constant photon arrival intensity?
- The DNA molecule will move in the focal volume
- The arrival intensity varies with molecule location



- Use $A(t) = A_0\alpha(t) \quad \alpha(t) \in (0, 1]$

- $(B_x(t), B_y(t), B_z(t))$ position at time t.

$$\alpha(t) = \exp\left\{-\frac{B_x^2(t) + B_y^2(t)}{2w_{xy}^2} - \frac{B_z^2(t)}{2w_z^2}\right\}$$

- Motion of the Molecule: Brownian motion
  - Use a three independent Brownian motion $(B_x(t), B_y(t), B_z(t))$ to model the location
  - $dB_x(t) = \sigma dW_t$
- w_{xy}, w_z are known

- Arrival time t_i $\qquad P(\Delta Y_{t_i} = 1 | \gamma_{t_i}, \alpha_{t_i}) = A(t_i)/\gamma(t_i)$

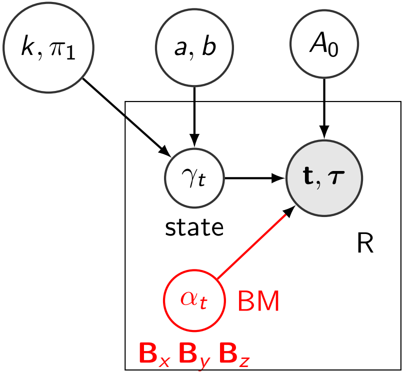


Figure: Generative view of the two-state Model with Brownian motion

Likelihood contruction

- Approximation: $\alpha(t) = \alpha(t_i)$ for $t \in (t_i, t_{i+1})$
- Conditioning on $\alpha(t)$: substitute A_0 with $A(t_i) = A_0\alpha(t_i)$

Likelihood contruction

- Approximation: $\alpha(t) = \alpha(t_i)$ for $t \in (t_i, t_{i+1})$
- Conditioning on $\alpha(t)$: substitute A_0 with $A(t_i) = A_0\alpha(t_i)$

$$L(\mathbf{t}, \boldsymbol{\tau}|\theta, \alpha(t))$$

$$= (\pi_1, \pi_2)\mathbf{D}_0\mathbf{H}_0 \left[\prod_{i=0}^{n-1} \exp\{(\mathbf{Q} - \mathbf{H}_i)(t_{i+1} - t_i)\}\mathbf{D}_{i+1}\mathbf{H}_{i+1}\right] \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

where $H_i = \begin{pmatrix} A(t_i)/a & 0 \\ 0 & A(t_i)/b \end{pmatrix}$

- Posterior distribution has the form

$$P(\theta|\mathbf{t},\boldsymbol{\tau}) \propto \int \eta(\theta)L(\mathbf{t},\tau|\theta,\alpha(t))P(\alpha(t))d(\alpha(t))$$
$$= \int \eta(\theta)L(\mathbf{t},\tau|\theta,\alpha(t))P(\mathbf{B}(t))d(\mathbf{B}(t))$$

- Posterior distribution has the form

$$P(\theta|\mathbf{t},\boldsymbol{\tau}) \propto \int \eta(\theta)L(\mathbf{t},\tau|\theta,\alpha(t))P(\alpha(t))d(\alpha(t))$$
$$= \int \eta(\theta)L(\mathbf{t},\tau|\theta,\alpha(t))P(\mathbf{B}(t))d(\mathbf{B}(t))$$

- Method: Data augmentation
  - Draw θ conditioning on current diffusion (B_x, B_y, B_z)

$$\boldsymbol{\theta} \sim [\boldsymbol{\theta}|\mathbf{B},\mathbf{t},\boldsymbol{\tau}] \propto \eta(\boldsymbol{\theta})L(\mathbf{t},\boldsymbol{\tau}|\boldsymbol{\theta},\alpha_t)$$

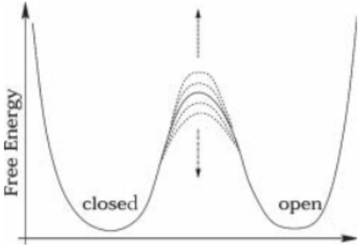
  - Draw the diffusion (B_x, B_y, B_z) conditioning on the current value of $\boldsymbol{\theta}$,

$$[B_x, B_y, B_z|\boldsymbol{\theta},\mathbf{t},\boldsymbol{\tau}] \sim L(\mathbf{t},\boldsymbol{\tau}|\boldsymbol{\theta},\alpha(t))P(B_x)P(B_y)P(B_z)$$

Model2: Continuous diffusive model

State transition: non-homogeneous CTMC

- Intuition: Transition depends on energy barrier x_t



- The energy barrier changes with time

Model the change of Energy barrier

- x_t modeled by Ornstein-Uhlenbeck process $\lambda > 0, \xi > 0$

$$\mathrm{d}x_t = -\lambda x_t \mathrm{d}t + \sqrt{2\lambda\xi}\,\mathrm{d}W_t$$

Model the change of Energy barrier

- x_t modeled by Ornstein-Uhlenbeck process $\lambda > 0, \xi > 0$

$$dx_t = -\lambda x_t dt + \sqrt{2\lambda\xi} dW_t$$

- Continuous diffusive model:
  The transition rate is no longer constant

$$\mathbf{Q}(t) = \begin{pmatrix} -k_{12}\exp(-x(t)) & k_{12}\exp(-x(t)) \\ k_{21}\exp(-x(t)) & -k_{21}\exp(-x(t)) \end{pmatrix}$$

- At $t = 0$, the OU-process starts at stationary distribution

$$x_0 \sim N(0, \xi)$$

A generative view of the model

Continuous diffusion model

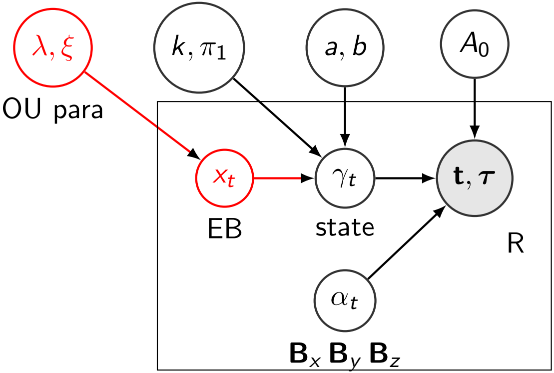


Figure: Generative View of the model

Posterior for continuous diffusive model

- Likelihood construction: Approximation
  $\mathbf{Q}(t) = \mathbf{Q}(t_i), t \in (t_i, t_{i+1})$

  $L(\mathbf{t}, \boldsymbol{\tau} | \theta, \alpha(t), \mathbf{x_t})$

  $$= (\pi_1, \pi_2)\mathbf{D}_0\mathbf{H}_0 \left[\prod_{i=0}^{n-1} \exp\{(\mathbf{Q}(t_i) - \mathbf{H}_i)(t_{i+1} - t_i)\}\mathbf{D}_{i+1}\mathbf{H}_{i+1} \right] \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Posterior for continuous diffusive model

- Likelihood construction: Approximation
  $\mathbf{Q}(t) = \mathbf{Q}(t_i), t \in (t_i, t_{i+1})$

$$L(\mathbf{t}, \boldsymbol{\tau} | \theta, \alpha(t), \mathbf{x_t})$$

$$= (\pi_1, \pi_2)\mathbf{D}_0\mathbf{H}_0 \left[\prod_{i=0}^{n-1} \exp\{(\mathbf{Q}(t_i) - \mathbf{H}_i)(t_{i+1} - t_i)\}\mathbf{D}_{i+1}\mathbf{H}_{i+1} \right] \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

- Posterior distribution

$$P(\boldsymbol{\theta}, \lambda, \xi | \mathbf{t}, \boldsymbol{\tau}) \propto$$
$$\int\int \eta'(\boldsymbol{\theta}, \lambda, \xi) L(\mathbf{t}, \boldsymbol{\tau} | \theta, \alpha_t, \mathbf{x_t}) \mathbf{P}(\alpha_\mathbf{t}) \mathbf{P}(\mathbf{x_t} | \lambda, \xi) \mathbf{d}(\alpha_\mathbf{t}) \mathbf{d}(\mathbf{x_t})$$

- Method: Data augmentation

Sampling Steps

1. Sample parameter $\boldsymbol{\theta}$

$$\boldsymbol{\theta} \sim [\boldsymbol{\theta}|\lambda, \xi, \mathbf{B}, x_t, \mathbf{t}, \boldsymbol{\tau}] \propto \eta'(\boldsymbol{\theta}, \lambda, \xi) L(\mathbf{t}, \boldsymbol{\tau}|\theta, \alpha_t, x_t)$$

2. Sample diffusion parameter λ, ξ

$$(\lambda, \xi) \sim [\lambda, \xi|\boldsymbol{\theta}, \mathbf{B}, x_t, \mathbf{t}, \boldsymbol{\tau}] \propto \eta'(\boldsymbol{\theta}, \lambda, \xi) P(x_t|\lambda, \xi)$$

3. Sample the the Brownian motion path

$$\mathbf{B} \sim [\mathbf{B}|\boldsymbol{\theta}, \lambda, \xi, x_t, \mathbf{t}, \boldsymbol{\tau}] \propto L(\mathbf{t}, \boldsymbol{\tau}|\theta, \alpha_t, x_t) P(\mathbf{B})$$

4. Sample the energy barrier path

$$x(t) \sim [x_t|\boldsymbol{\theta}, \lambda, \xi, \mathbf{B}, \mathbf{t}, \boldsymbol{\tau}] \propto L(\mathbf{t}, \boldsymbol{\tau}|\theta, \alpha_t, x_t) P(x_t|\lambda, \xi)$$

Association with two states model

$$dx_t = -\lambda x_t dt + \sqrt{2\lambda\xi} dW_t$$

If $\xi \simeq 0$

- The stationary distribution $N(0, \xi)$ will degenerate to 0
- The SDE has solution $x_t = 0$

Association with two states model

$$dx_t = -\lambda x_t dt + \sqrt{2\lambda\xi} dW_t$$

If $\xi \simeq 0$

- The stationary distribution $N(0, \xi)$ will degenerate to 0
- The SDE has solution $x_t = 0$
- The infinitesimal $Q(t)$

$$\mathbf{Q}(t) = \begin{pmatrix} -k_{12} & k_{12} \\ k_{21} & -k_{21} \end{pmatrix}$$

Association with two states model

$$dx_t = -\lambda x_t dt + \sqrt{2\lambda\xi} dW_t$$

If $\xi \simeq 0$

- The stationary distribution $N(0, \xi)$ will degenerate to 0
- The SDE has solution $x_t = 0$
- The infinitesimal $Q(t)$

$$\mathbf{Q}(t) = \begin{pmatrix} -k_{12} & k_{12} \\ k_{21} & -k_{21} \end{pmatrix}$$

- Exactly the two-state model!

Model Comparision

Model Comparision

1. By checking the value of ξ
   $\mathbf{H}_0 : \xi = 0$ two-state model
   $\mathbf{H}_1 : \xi > 0$ continuous diffusion model

Model Comparision

1. By checking the value of ξ
   $\mathbf{H}_0 : \xi = 0$ two-state model
   $\mathbf{H}_1 : \xi > 0$ continuous diffusion model
2. By comparing Bayes factor

$$\mathrm{BF} = \frac{P(\mathbf{t}, \boldsymbol{\tau} | M_1)}{P(\mathbf{t}, \boldsymbol{\tau} | M_2)}$$

   where M_1 is the two state model, M_2 is the continuous diffusive model

Details on priors and other parameters

1. Prior issues
   1.1 Informative prior for $\boldsymbol{\theta} = (a, b, \pi_1, k, A_0)$
   - $a \sim \Gamma(2, 1)$
   - $b \sim \Gamma(1.5625, 1.5625)$
   - $\pi_1 \sim \text{beta}(0.89, 0.89)$
   - $\pi_1 \sim \text{Exp}(1/40000)$
   - $A_0 \sim \Gamma(1.96, 5.6 \times 10^{-5})$

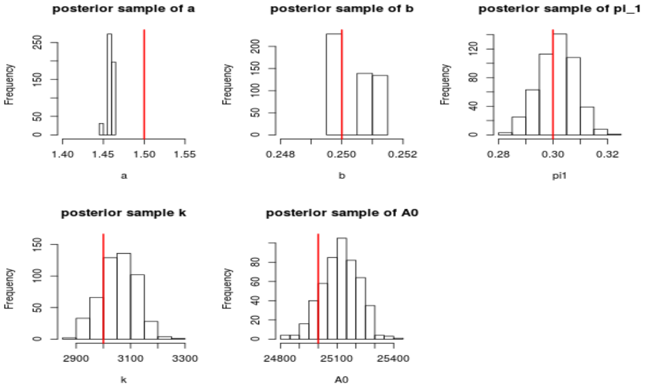
   1.2 Less information for λ,ξ.
   - $\lambda \sim \Gamma(40, 0.5)$
   - $\xi \sim \Gamma(2, 1)$
2. Other parameters
   - Brownian motion parameters: $w_{xy} = 310, w_z = 1760$
   - BM constant σ^2 is not given, set as 1000

Experiment 1: Simulated datasets

- Simulate 50 sequences of $(\mathbf{t}, \boldsymbol{\tau})$s
- Each sequence is simulated from two-state BM model with $t_{max} < 0.05$
- The number of observations in each datasets varies from $1000 \sim 5000$
- Run both two-state model and continuous diffusion model for 5000 iterations

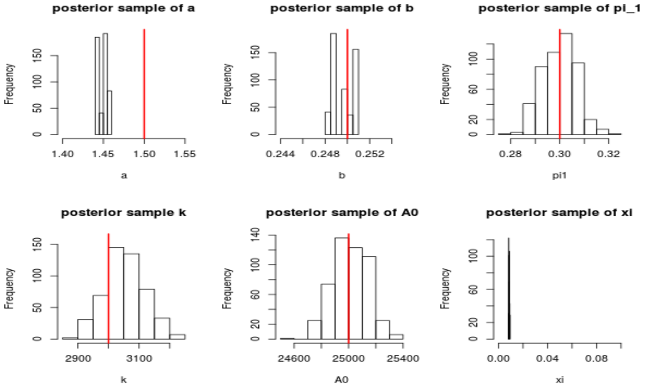
two-state BM model

- Posterior samples for (a, b, π_1, k, A_0)



Continuous diffusive model

- Posterior samples for $(a, b, \pi_1, k, A_0, \xi)$



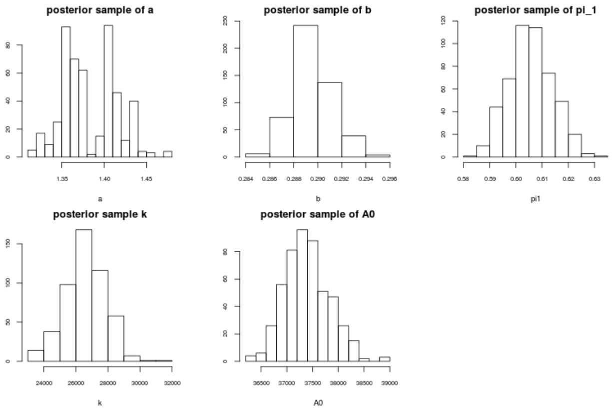
BF $= 0.99$, no significant difference between the two model

Experiment 2: Real data

- 50 real datasets from Xie's lab at Harvard University
- Each contains a sequence of 1815 pairs of (t_i, τ_i)
- Run both two-state model and continuous diffusion model for 5000 iterations

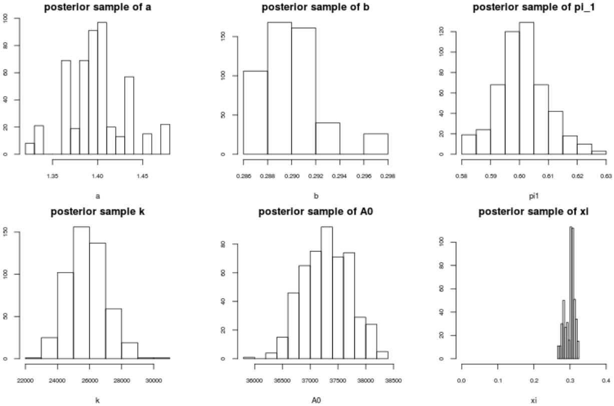
two-state BM model

- posterior samples (a, b, π_1, k, A_0)



Continuous diffusive model

- Posterior samples for $(a, b, \pi_1, k, A_0, \xi)$



- Comparing posterior mean

para	prior	twostateBM	Con-diff
a	$\Gamma(1, 0.5)$	1.367	1.405
b	$\Gamma(1.56, 1.56)$	0.289	0.289
π_1	$\text{Beta}(0.89, 0.89)$	0.604	0.605
k	$\text{Exp}(1/4000)$	26744	25830
A_0	$\Gamma(1.96, \frac{7}{12500})$	37421	37235

- BF $= 0.023$, evidence for continuous diffusive model

Summary

- Fluorescence experiment
- Two models: (Likelihood function)
  - two-state model: CTMC transition
    Two state model with BM
  - Continuous diffusion model: OU-process for energy barrier
- Sampling from posterior distribution
  - Metropolis-hasting algorithm
  - Data Augmentation algorithm
- Model selection:
  - By ξ
  - Bayes factor
- Experiment : continuous diffusive model fits better on the real data

Discussion

1. Pros
   - First Bayesian model to study s single-molecule experiment
   - Can incorporate many conditions in the experiment (BM, OU for EB)
   - Can be extended to model other counting process with latent structure
   - The computation cost for each iteration is $\mathcal{O}(n)$
2. Cons
   - Low efficiency in component-wise update in the Brownian motion path
   - Sensitive to prior (λ, ξ)
   - Some other models between two-states model and continuous diffusive model

Thank you!

Componentwise update

For $i = 0, 1, \ldots, n$

Componentwise update

For $i = 0, 1, \ldots, n$

1. Propose a new location $B_i' = (B_x', B_y', B_z')$ for the ith time point $B(t_i) = (B_x(t_i), B_y(t_i), B_z(t_i))$
   Calculate α' at time t_i

Componentwise update

For $i = 0, 1, \ldots, n$

1. Propose a new location $B'_i = (B'_x, B'_y, B'_z)$ for the ith time point $B(t_i) = (B_x(t_i), B_y(t_i), B_z(t_i))$
   Calculate α' at time t_i
2. Calculate M-H ratio

$$r = \frac{L(\mathbf{t}, \boldsymbol{\tau} | \theta, \alpha'_t) P(B'_x) P(B'_y) P(B'_z) T(B' \to B(t_i))}{L(\mathbf{t}, \boldsymbol{\tau} | \theta, \alpha_t) P(B_x) P(B_y) P(B_z) T(B \to B'(t_i))}$$

3. Sample $U \sim U(0, 1)$.
   Update $B(t_i)$ to B' when $U < r$

The diffusion Path

- Identifiability issues
  - The path of $\alpha(t)$ can be is related conditional likelihood. We can find a path will high posterior probability, if $(B_x(t_0), B_y(t_0), B_z(t_0))$ is fixed.
  - Notice $\alpha(t) = \exp\left\{-\frac{B_x^2(t)+B_y^2(t)}{2w_{xy}^2} - \frac{B_z^2(t)}{2w_z^2}\right\}$.
  - The underlying Brownian path is not identifiable. Multiple paths for the sample $\boldsymbol{\alpha}_t$

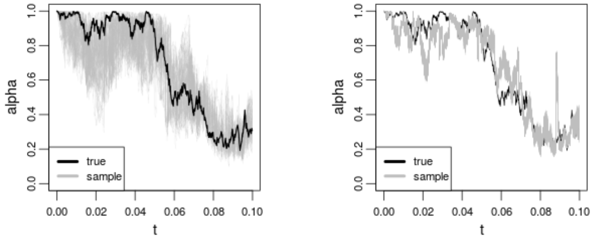
The diffusion path

Problems related to component-wise update

- Low acceptance rate
  - $r = \frac{L(\mathbf{t}, \boldsymbol{\tau}|\theta, \alpha_t')}{L(\mathbf{t}, \boldsymbol{\tau}|\theta, \alpha_t)} \cdot \frac{P(B_x')P(B_y')P(B_z')}{P(B_x)P(B_y)P(B_z)}$ if using symmetric proposed density function
  - $P(\mathbf{B}_x) = (2\pi)^{n/2} \exp\left[-\frac{1}{2\sigma^2} \sum_{i=0}^{n-1} \frac{[B_x(t_{i+1}) - B_x(t_i)]^2}{\Delta t_i}\right]$
  - $L(\mathbf{t}, \boldsymbol{\tau}|\theta, \alpha_t)$not sensitive to α, $P(B_x), P(B_y), P(B_z)$ sensitive to the change of B_x, B_y, B_z
  - Easily stuck in a "smooth" Brownian motion path

two-stage update

1. stage one:
   Propose a change $B'_x \sim N\left(B_x(t_{i-1}), \sigma^2(t_i - t_{i-1})\right)$
   $u \sim U(0,1)$. Accept the change is $u \leq \dfrac{L(\mathbf{t}, \boldsymbol{\tau}|\theta, \alpha'_t)}{L(\mathbf{t}, \boldsymbol{\tau}|\theta, \alpha_t)}$
2. stage two: M-H method with component-wise update



Why bayesian method?

Why bayesian method?

- Tradition methods:
  - Method of Moments
  - Maximum likelihood estimation directly
  - EM

Why bayesian method?

- Tradition methods:
  - Method of Moments
  - Maximum likelihood estimation directly
  - EM
- Why Bayesian?
  - Closed from likelihood function
  - The model can be written as a generative model
  - Informative prior

Computation cost

	simple twostate	twostateBM	cont-diff
#ofparas	5	5	7
#ofupdates/iter	5	3n+6	4n+11
cpu-cost(naive)	$\mathcal{O}(n)$	$\mathcal{O}(n^2)$	$\mathcal{O}(n^2)$
cpu-cost(opt)	$\mathcal{O}(n)$	$\mathcal{O}(n)$	$\mathcal{O}(n)$

Backward-forward algorithm

- Likelihood function
$$(\pi_1, \pi_2)\left[\prod_{i=0}^{n-1} \mathbf{D}_i \mathbf{H}_i \exp\{(\mathbf{Q} - \mathbf{H}_i)(t_{i+1} - t_i)\}\right] \mathbf{D}_n \mathbf{H}_n \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$
- Compute *backwards* a sequence of matrices K_i by recursion

$$\begin{cases} \mathbf{K}_{n+1} = \mathbf{I}, \\ \mathbf{K}_n = \mathbf{D}_n \mathbf{H}_n, \\ \mathbf{K}_i = \mathbf{D}_i \mathbf{H}_i \exp\{(\mathbf{Q} - \mathbf{H}_i)\Delta t_i\} \mathbf{K}_{i+1}, \quad i = n-1, \ldots, 1, 0 \end{cases}$$

- Forward calculation
  1. Propose a change $\mathbf{B}' = (B'_x, B'_y, B'_z)$ for the ith time point $(B_x(t_i), B_y(t_i), B_z(t_i))$, calculate α'_{t_i},$\mathbf{H}'$ based on (B'_x, B'_y, B'_z).
  2. Compute

$$\mathbf{R} = \begin{cases} \mathbf{D}_i\mathbf{H}_i \exp\{(\mathbf{Q} - \mathbf{H}_i)\Delta t_i\} & \text{if } i < n, \\ \mathbf{D}_n\mathbf{H}_n & \text{if } i = n, \end{cases}$$

$$\mathbf{S} = \begin{cases} \mathbf{D}_i\mathbf{H}'_i \exp\{(\mathbf{Q} - \mathbf{H}'_i)\Delta t_i\} & \text{if } i < n, \\ \mathbf{D}_n\mathbf{H}'_n & \text{if } i = n, \end{cases}$$

and

$$L(\mathbf{t}, \boldsymbol{\tau} | \boldsymbol{\theta}, \boldsymbol{\alpha}'_t) = \mathbf{v}_i\mathbf{S}\mathbf{K}_{i+1} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \text{ and}$$

$$L(\mathbf{t}, \boldsymbol{\tau} | \boldsymbol{\theta}, \boldsymbol{\alpha}_t) = \mathbf{v}_i\mathbf{R}\mathbf{K}_{i+1} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

- Compute MH ratio

$$r = \frac{L(\mathbf{t}, \boldsymbol{\tau}|\boldsymbol{\theta}, \boldsymbol{\alpha}'_t)P(\mathbf{B}'_x)P(\mathbf{B}'_y)P(\mathbf{B}'_z)T(\mathbf{B}'_i \rightarrow \mathbf{B}(t_i))}{L(\mathbf{t}, \boldsymbol{\tau}|\boldsymbol{\theta}, \boldsymbol{\alpha}_t)P(\mathbf{B}_x)P(\mathbf{B}_y)P(\mathbf{B}_z)T(\mathbf{B}_i \rightarrow \mathbf{B}'(t_i))},$$

  where $T(\cdot \rightarrow \cdot)$ is the transition density of the proposal distribution.
- Generate $u \sim \text{Uniform}(0, 1)$.
  If $u < \min(1, r)$, then update $\mathbf{B}(t_i)$ to $\mathbf{B}'$ and $\mathbf{v}_{i+1} = \mathbf{v}_i\mathbf{S}$.
  Otherwise, keep $\mathbf{B}(t_i)$ unchanged and $\mathbf{v}_{i+1} = \mathbf{v}_i\mathbf{R}$

Sensitivity issues

Likelihood function:

- Sensitive to a, b since $\boldsymbol{\tau}$ mainly contains information for a, b
- Not sensitive to π, k, A_0
- Not sensitive to the $\alpha(t)$ path and OU path x_t

Why combining multiple datasets

1. Observed sequence (t_i, τ_i) not i.i.d
2. Brownian motion model, as $t \to \infty$, $\alpha(t) \to 0$ with high probability
3. Identifiability issues for the energy barrier path