Markov Chain Monte Carlo in the Analysis of Single-Molecule Experimental Data

S. C. Kou*, X. Sunney Xie† and Jun S. Liu*

*Department of Statistics, Harvard University, Science Center, Cambridge, MA 02138
†Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA 02138

Abstract. This article provides a Bayesian analysis of the single-molecule fluorescence lifetime experiment designed to probe the conformational dynamics of a single DNA hairpin molecule. The DNA hairpin's conformational change is initially modeled as a two-state Markov chain, which is not observable and has to be indirectly inferred. The Brownian diffusion of the single molecule, in addition to the hidden Markov structure, further complicates the matter. We show that the analytical form of the likelihood function can be obtained in the simplest case and a Metropolis-Hastings algorithm can be designed to sample from the posterior distribution of the parameters of interest and to compute desired estiamtes. To cope with the molecular diffusion process and the potentially oscillating energy barrier between the two states of the DNA hairpin, we introduce a data augmentation technique to handle both the Brownian diffusion and the hidden Ornstein-Uhlenbeck process associated with the fluctuating energy barrier, and design a more sophisticated Metropolis-type algorithm. Our method not only increases the estimating resolution by several folds but also proves to be successful for model discrimination.

INTRODUCTION

Recent technological advances have allowed scientists to make observations on single-molecule dynamics, which was unthinkable just a few decades ago ([13],[25],[23], [19],[12]). Complementary to the traditional experiments done on large ensembles of molecules, single-molecule experiments offer a great potential and many advantages for new scientific discoveries. First, one can directly measure the distributions of molecular properties, rather than relying on the ensemble average. Second, single-molecule experiments allow biochemical processes to be followed in real time and capture transient intermediates, which previously could only be accomplished by synchronizing actions of a large ensemble of molecules. Third, single-molecule trajectories provide detailed dynamic information, which is unavailable from the traditional ensemble experiments. The detailed dynamic information is particularly important for complex biomolecules that have intricate internal structures ([24], [26]). In this article we analyze the single-molecule experimental data on the DNA hairpin kinetics.

A DNA hairpin is a single stranded nucleic acid structure with bases at the two ends complementary to each other so that the intramolecular pairing can form. A DNA hairpin has two states — in the close state, the two ends pair together, while in the open states the pairings are broken [2] (see Figure 1). In a living cell, with the breaking of intermolecular pairing between the two (double helix) DNA strands, the loose strands often form a DNA hairpin structure. DNA hairpin structure participates in many biolog-

CP690, *The Monte Carlo Method in the Physical Sciences*, edited by J. E. Gubernatis

ical functions including, for example, gene regulations [27], DNA recombinations [5], and the facilitation of mutagenic events [22], etc. The hairpin structure can also be a potential antisense drug [20]. Studying the conformation properties of DNA hairpin, such as the conformational fluctuation and energy barrier between the open and close states, hence serves an important model system to understand more complicated biochemical processes.

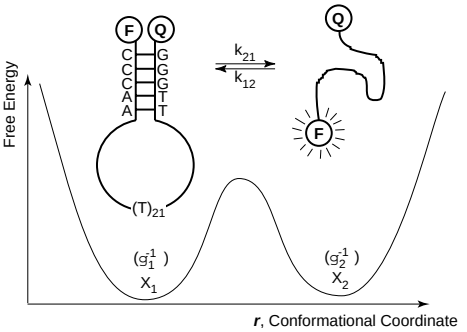

FIGURE 1. The two states of a DNA hairpin. To infer the open and close states, a fluorescence donor and a quencher are attached to the two ends of the DNA hairpin.

In a fluorescence lifetime experiment, the DNA hairpin in a solvent spontaneously switches between the open and close states. A fluorescence donor and a quencher are attached to the two ends of the molecule (see Figure 1). The donor emits photons when it is excited by a laser pulse, and the quencher annihilates the excitation. In the hairpin's close state the quenching is strong, and thus very few photons from the donors are detected; in the open state the quenching is weak, and many photons from the donor are detected. The open/close of the DNA hairpin can hence be inferred indirectly from the detected photon arrivals ([10], [7], [4]).

Let $A \triangleq$ close, and $B \triangleq$ open. The simplest model is a continuous-time two-state Markov chain ([15], [16], [17]), which can be depicted as

$$A \underset{k_{21}}{\overset{k_{12}}{\rightleftarrows}} B, \tag{1}$$

where k_{12} and k_{21} represent the transition rates between the two states. Let $\gamma(t)$ denote the decay rate of the hidden fluorescence state, which takes values a and b, respectively, at states A and B. The Fokker-Planck equation gives the transition matrix of this two-state model:

$$P(t) = e^{Qt} = \begin{pmatrix} \pi_1 + \pi_2 e^{-kt} & \pi_2(1-e^{-kt}) \\ \pi_1(1-e^{-kt}) & \pi_2 + \pi_1 e^{-kt} \end{pmatrix},$$

where $k = k_{12} + k_{21}$, $(\pi_1, \pi_2) = (\frac{k_{21}}{k_{12}+k_{21}}, \frac{k_{12}}{k_{12}+k_{21}})$, and $Q = \begin{pmatrix} -k_{12} & k_{12} \\ k_{21} & -k_{21} \end{pmatrix}$.

However, the γ process cannot be observed directly; instead, one can observe the photon arrival time t and a corresponding fluorescence decay time τ (with respect to its excitation pulse in the pulse train). The photon arrival time t follows a doubly stochastic

Poisson process with the arrival rate inversely proportional to $\gamma(t)$. If a photon arrives at time t, its fluorescence decay time τ has an exponential distribution with rate $\gamma(t)$. The thick line in Figure 2 depicts the unobservable two-state Markov chain corresponding to the open-close of the DNA hairpin. Each vertical bar represents the arrival of a photon. The height of each bar represents the corresponding decay time τ of that recorded photon.

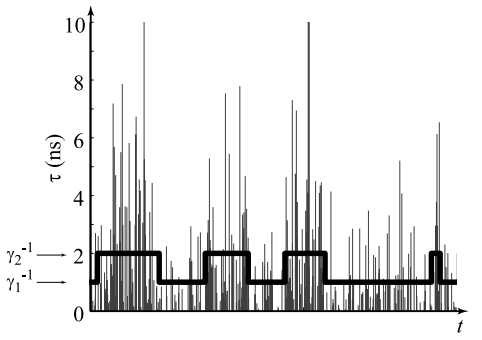

FIGURE 2. The data structure in the single-molecule lifetime experiments. The photon arrival times t are represented by the vertical bars, whose heights represent the decay time τ.

In addition to the hidden Markov structure of $\gamma(t)$, the *unobservable* trajectories of Brownian diffusion of the hairpin molecule further complicates the inference (see Section 3). Furthermore, it has also been argued that the two-state model is too simplistic to reflect the nature because the energy barrier between the two states may fluctuate dynamically or there may be sub-states within each of the two states and these substates may communicate at different rates. With the current data resolution and existing inference methods, discerning different models and assessing their fit to the experimental data have remained difficult [18].

In order to successfully cope with both the experimental and the modeling complexity, we use a Bayesian data augmentation approach (Tanner and Wong, 1987, [21]), which has advantages over the conventional method-of-moment type approaches widely used in the field in many aspects including: a) a better time resolution; b) a broader range of accessible time scales; c) a much better accuracy in extracting model dependent parameters. We expect that the general strategies developed here can be widely applied to other single-molecule experiments. Our analysis here shows a significant improvement in estimation accuracy for several physical parameters of interest and provides a strong statistical evidence to favor the more complex model that allows for the fluctuation of the energy barrier between the two states over the simple two-state model. Section 2 details the two-state statistical model and the Bayesian analysis via a Metropolis-type algorithm. Section 3 introduces the data augmentation approach to handle the experimental complications. Section 4 considers models beyond the two-state case and discusses model assessment. Section 5 analyzes experimental data. Section 6 concludes the paper and provides some further discussion.

BAYESIAN ANALYSIS OF THE TWO-STATE MODEL

Let $Y(t)$ be the total number of photon arrivals up to time t. In an infinitesimal time interval $(t, t+dt)$, the probability of observing a photon is proportional to $\gamma^{-1}(t)\,dt$. Denoting $\triangle Y_t = Y(t+dt) - Y(t)$, we have

$$P\{\triangle Y_t = 1 | \gamma_t\} = A_0(t)\gamma^{-1}(t)\,dt \quad (2)$$

$$[\tau | \triangle Y_t = 1, \gamma_t] \sim \gamma(t)\exp(-\gamma(t)\tau) \quad (3)$$

where $A_0(t)$ is the photon arrival intensity at time t. We first treat $A_0(t)$ as a constant over time: $A_0(t) \equiv A_0$. In real experiments, the photon intensity $A_0(t)$ may also be stochastic, and this additional complexity will be addressed in Section 3.

Let $0 = t_0 < t_1 < \cdots < t_n$ be the observed photon arrival times, and let τ_i be the corresponding fluorescence decay time. The pairs $\{(t_i, \tau_i)\}_{i=0}^n$ are collected through the fluorescence lifetime experiments. We note that the likelihood consists of two parts: (i) the contribution from the observed photons at time t_i and the corresponding τ_i; and (ii) the contribution from the time interval (t_i, t_{i+1}), in which no photon arrives. By employing an infinitesimal discretization technique and matrix computation, Kou, Xie and Liu (2003, [8]) showed that the likelihood of observing $\{(t_i, \tau_i)\}_{i=0}^n$ is

$$L(\mathbf{t}, \tau | \theta) = (\pi_1, \pi_2) D_0 E \left(\prod_{i=0}^{n-1} e^{(Q-E)\triangle t_i} D_{i+1} E \right) \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad (4)$$

where $\triangle t_i = t_{i+1} - t_i$, and the matrices $E = \mathrm{diag}(A_0/a, A_0/b)$, $D_i = \mathrm{diag}(ae^{-a\tau_i}, be^{-b\tau_i})$. We note that formula (4) is applicable to any finite-state hidden Markov process model, such as the two-by-two model of [18].

The likelihood function (4) has five free parameters $\theta = (a, b, \pi_1, k, A_0)$ with the constraints that (i) $a > b > 0$, (ii) $0 \le \pi_1 \le 1$,, (iii) $k > 0$, and (iv) $A_0 > 0$. Let $\eta(\theta)$ denote the prior distribution on the parameters. The posterior distribution $P(\theta | \mathbf{t}, \tau) \propto \eta(\theta) L(\mathbf{t}, \tau | \theta)$. The inference on the parameters (e.g., k) can be represented by summarizing statistics from this distribution. For example, the posterior mean $\hat{k}$ of k can be used as an estimate of the true k:

$$\hat{k} = \int k P(\theta | \mathbf{t}, \tau) d\theta.$$

Since analytical computations of this type are infeasible, we design a Metropolis-type algorithm (Metropolis *et al.* 1953 [11], Hastings 1970 [6]) to simulate from $P(\theta | \mathbf{t}, \tau)$:

- Given a and b, (i) draw x from $\Gamma(\frac{1}{c_1}, ac_1)$ and y from $\Gamma(\frac{1}{c_2}, bc_2)$, where c_1 and c_2 are two tuning parameters; and (ii) let $a' = \max(x, y)$, $b' = \min(x, y)$.
- Given π_1, draw π_1' from the beta distribution $B(c_3\pi_1, c_3(1-\pi_1))$.
- Given k, draw k' from $\Gamma(\frac{1}{c_4}, kc_4)$. The mean and variance of $\frac{k'}{k}$ are 1 and c_4 respectively, thus letting c_4 finely tune the perturbation.
- Given A_0, draw A_0' from $\Gamma(\frac{1}{c_5}, A_0 c_5)$, whose mean is A_0, and variance is controlled by c_5

To test the efficacy of the sampling scheme, we generated the γ process from the two-state model, then $\{(t_i, \tau_i)\}_{i=0}^{n}$ according to (2) and (3). With a flat prior on the parameters, we apply the MCMC algorithm on the simulated data to draw samples from the posterior distribution. Figure 3 summarizes the posterior distributions for the parameters of interest, where the vertical bars represent the true values. The algorithm runs quite fast: A total number of 10,000 samples took less than two minutes to draw on a Pentium 4 PC.

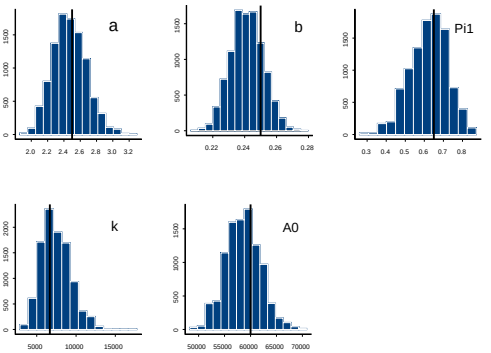

FIGURE 3. Histograms of the posterior samples. Vertical bar in each panel is the true value.

DATA AUGMENTATION FOR BROWNIAN DIFFUSIONS

The DNA hairpin in the experiment is placed in a focal volume illuminated by a laser beam. The laser excites the donor dye on the DNA hairpin molecule so that the dye releases photons from time to time. At the same time, the DNA hairpin molecule also diffuses in the focal volume, which results in a time-varying nonconstant laser illuminating intensity $A_0(t)$. We can write $A_0(t) = A_0\alpha(t)$ with $\alpha(t) = \exp\left(-\frac{B_x^2(t)+B_y^2(t)}{2w_{xy}^2} - \frac{B_z^2(t)}{2w_z^2}\right)$, where the known constants w_{xy} and w_z specify the x-y and z axes of the ellipsoidal focal volume, and $(B_x(t), B_y(t), B_z(t))$ is the physical location of the molecule described by a standard three-dimensional Brownian motion with known diffusion constant. In the presence of diffusion, (2) and (3) are changed to

$$P\{\triangle Y_t = 1 \mid \gamma_t, \alpha_t\} = A_0\alpha(t)\gamma^{-1}(t)\,dt \tag{5}$$

$$[\tau \mid \triangle Y_t = 1, \gamma_t, \alpha_t] \sim \gamma(t)\exp(-\gamma(t)\tau) \tag{6}$$

The conditioning on $\alpha(t)$ changes the likelihood to a conditional likelihood

$$L(\mathbf{t}, \tau|\theta, \alpha_t) = (\pi_1, \pi_2)D_0E_0\left(\prod_{i=0}^{n-1} e^{(Q-E_i)\triangle t_i}D_{i+1}E_{i+1}\right)\begin{pmatrix}1\\1\end{pmatrix}. \tag{7}$$

where $E_i = \mathrm{diag}(\alpha(t_i)A_0/a,\, \alpha(t_i)A_0/b)$. With a prior distribution $\eta(\theta)$ on the parameters, the posterior distribution of θ given the observations $\{(t_i,\tau_i)\}_{i=0}^n$ is

$$P(\theta|\mathbf{t},\tau) \propto \int \eta(\theta)L(\mathbf{t},\tau|\theta,\alpha_t)P(\alpha_t)\,d(\alpha_t), \tag{8}$$

where $P(a_t)$ denotes the transition density of the illuminating process $\alpha(t)$.

Since it is infeasible to integrate out the Brownian diffusion (B_x,B_y,B_z) analytically, we need to sample from the joint posterior distribution of θ and (B_x,B_y,B_z),

$$P(\theta,B_x,B_y,B_z|\mathbf{t},\tau) \propto \eta(\theta)L(\mathbf{t},\tau|\theta,\alpha_t)P(B_x)P(B_y)P(B_z), \tag{9}$$

where $P(B_x)$ denotes the transition density of Brownian motion. This way the hidden Brownian diffusion is effectively marginalized out. Starting from an initial θ and (B_x,B_y,B_z), we iteratively draw θ conditioning on the diffusion (B_x,B_y,B_z) and draw (B_x,B_y,B_z) conditioning on θ. The sampling of θ conditioning on (B_x,B_y,B_z) is achieved by the algorithm outlined in Section 2. The sampling of (B_x,B_y,B_z) conditioning on θ can be achieved by updating the diffusion chain component by component. To efficiently compute the likelihood in the component-wise updating, we used a *forward-backward* recursion: backward compute the partial sums in the likelihood, forward sample the diffusion chain one component at a time. The computational cost of this forward-backward recursion is only twice the number of the observed photon arrivals.

BEYOND TWO-STATE: THE CONTINUOUS DIFFUSIVE MODEL

It has been observed that for certain molecules (other than the DNA hairpin) the two-state model (1) is not accurate enough to describe the conformational details [18]. This phenomenon of "dynamic disorder" motivates models beyond the two-state. The 2×2 model in [18] is such an attempt and can be analyzed by the same Bayesian data augmentation approach outlined in the previous two sections. The 2×2 model can be further generalized to a continuous diffusive model, which needs additional effort. In this model a continuous stochastic control process is introduced, which "controls" the transition rates as follows:

$$A \underset{k_{21}e^{-x(t)}}{\overset{k_{12}e^{-x(t)}}{\rightleftarrows}} B, \tag{10}$$

where $x(t)$ satisfies the Ornstein-Uhlenbeck equation $dx_t = -\lambda x_t dt + \sqrt{2\xi\lambda}dW_t$. Intuitively, this can be seen as a result of a stochastically fluctuating energy barrier between the two states (see Figure 4), where the Ornstein-Uhlenbeck process captures the fluctuation [1]. Although some debates have been set forth, there is no clear evidence as to whether the continuous diffusive model is definitively more appropriate than the simple two-state model for the DNA hairpin.

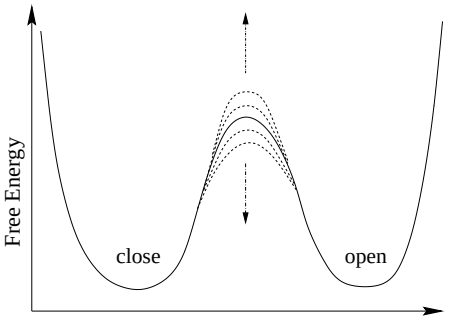

FIGURE 4. The fluctuating energy barrier between the two states of the DNA hairpin. The Ornstein-Uhlenbeck process models this dynamic oscillation.

By employing the discretization and matrix techniques, we obtain the closed-form conditional likelihood giving both $\alpha(t)$ and $x(t)$:

$$L(\mathbf{t},\tau|\theta,\alpha_t,x_t) = (\pi_1,\pi_2)D_0E_0\left(\prod_{i=0}^{n-1} e^{(Q_i-E_i)\triangle t_i}D_{i+1}E_{i+1}\right)\begin{pmatrix}1\\1\end{pmatrix}, \qquad (11)$$

where $Q_i \triangleq \begin{pmatrix} -k_{12}\exp(-x(t_i)) & k_{12}\exp(-x(t_i)) \\ k_{21}\exp(-x(t_i)) & -k_{21}\exp(-x(t_i)) \end{pmatrix}$. The posterior distribution of the parameters (θ,λ,ξ) given the observations $\{(t_i,\tau_i)\}$ is hence

$$P(\theta,\lambda,\xi|\mathbf{t},\tau) \propto \int\int \eta(\theta,\lambda,\xi)L(\mathbf{t},\tau|\theta,\alpha_t,x_t)P(\alpha_t)P(x_t|\lambda,\xi)\,d(\alpha_t)\,d(x_t), \qquad (12)$$

where η denotes the prior distribution on the parameters, and $P(x_t|\lambda,\xi)$ is the transition density of the Ornstein-Uhlenbeckprocess $x(t)$.

We again use the data augmentation approach to impute $x(t)$ and $\mathbf{B}(t) = (B_x(t),B_y(t),B_z(t))$ and use the Metropolis-type algorithm to accomplish the path integral in (12). To improve the Monte Carlo efficiency, we let $\phi = \sqrt{\xi\lambda}$ and work on the transformed parameters (λ,ϕ), which are less correlated than the original (λ,ξ). The joint distribution of (θ,λ,ϕ) and $(\mathbf{B}(t),x(t))$ is

$$\begin{aligned} & P(\theta,\lambda,\phi,B_x,B_y,B_z,x_t \mid \mathbf{t},\tau) \\ \propto\ & \eta'(\theta,\lambda,\phi)L(\mathbf{t},\tau \mid \theta,\alpha_t,x_t)P(B_x)P(B_y)P(B_z)P(x_t \mid \lambda,\frac{\phi^2}{\lambda}), \end{aligned} \qquad (13)$$

where η' is the prior distribution on (θ,λ,ϕ).

To further improve the computation efficiency, we introduce a scale transformation to update ϕ and x_t jointly, i.e.,

$$(\lambda,\phi,\theta,B_x,B_y,B_z,x_t) \to (\lambda,s\phi,\theta,B_x,B_y,B_z,sx_t),$$

where s is a scalar. We first propose s from the gamma distribution $\Gamma(s;1/c,c)$, and then accept the proposed s with probability: $r = \min\left\{1,\frac{\Gamma(s^{-1};1/c,c)p(s)s}{\Gamma(s;1/c,c)p(1)}\right\}$. This move is

important because of the high correlation between (λ, ϕ) and x_t — once the process x_t is given, the distribution of (λ, ϕ) is very tightly concentrated on its mode and vice versa for x_t due to the huge chain length. Figure 5 compares the autocorrelation of the samples with and without the scale move.

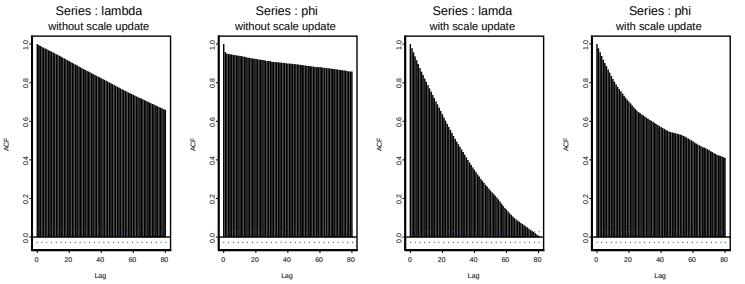

FIGURE 5. Autocorrelations of the posterior samples with and without the scale-transformation update. The left two panels do not have the scale update, while the right two have.

Since the stationary distribution of the control process x_t is $N(0, \xi)$, we note that the two-state model is actually a degenerate case of the diffusive model: it corresponds to ξ being 0, which suggests that after applying the algorithm to a given data set, we can look at the estimated value of $\sqrt{\xi}$. If the value is very close to 0, it then provides a strong indication that the two-state model is perhaps sufficient to explain the data.

EXAMPLES

Fitting a single trajectory with the two-state model

We analyzed a data set with 784 (t_i, τ_i) pairs obtained by the Xie lab. Due to technological limitations, their experiments have some additional complications such as the arrival of *background* photons, and the time-wrapping and negative reading of the machine-recorded τ, which can be easily accommodated by modifying the E_i and D_i matrices in (7) respectively. Applying the data augmentation method with the backward-forward updating on the modified likelihood, we obtained 5000 posterior samples, which can be used to derive the posterior distribution of any parameter of interest. Figure 6 shows the posterior distribution of $1/k$, which is termed the *decay-time constant* and indicates the energy barrier between the open and close states. Thus, the point estimate of $1/k$ is $109\mu s$ and its 90% probability interval is $[58, 220]\mu s$.

In some of the previous approaches, arrival times were first "binned" together and then used to fit certain moment equations for estimating parameters of interest ([14], [3]). Because of the binning, these approaches suffer a significant loss of time resolution. For the same data set that we analyzed, the methods based on "binning" have a maximum time resolution of 280 microseconds (μs). Furthermore, it is extremely difficult, if not impossible, for these methods to provide a measure of uncertainty of the estimates.

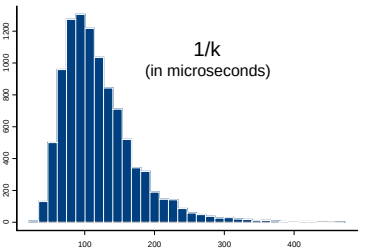

FIGURE 6. Distribution of decay time constant.

The diffusive model

We first applied the diffusive model to the data set (with 784 observation pairs) analyzed in the previous subsection. From Figure 7, we observe that the estimated values of $\sqrt{\xi}$ is very close to 0, which implies that for this data set the two-state model is quite a reasonable approximation.

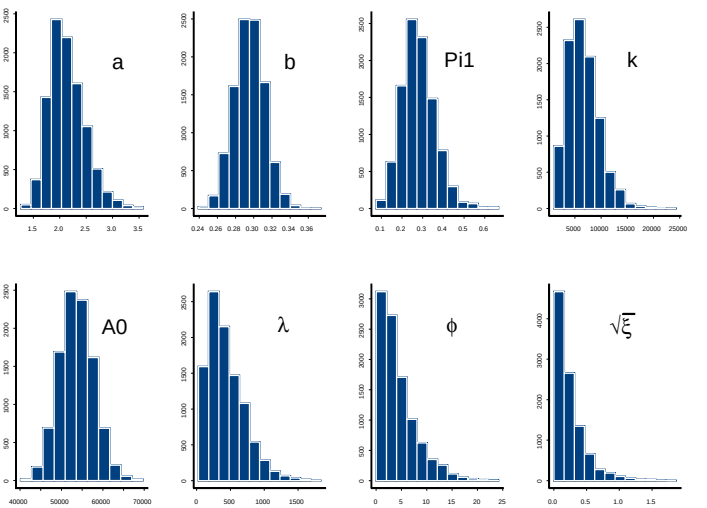

FIGURE 7. Histograms of posterior samples of the parameters of interest based on one data set.

Next we applied the algorithm to analyze the 50 DNA hairpin data sets obtained by the Xie lab. Comparing the estimates shown in Figure 8 with those obtained from a single data set, one clearly sees that with more information available the estimates become much sharper. Furthermore, the posterior samples of $\sqrt{\xi}$ is significantly different from

0, indicating that the two-state model does not fit the data.

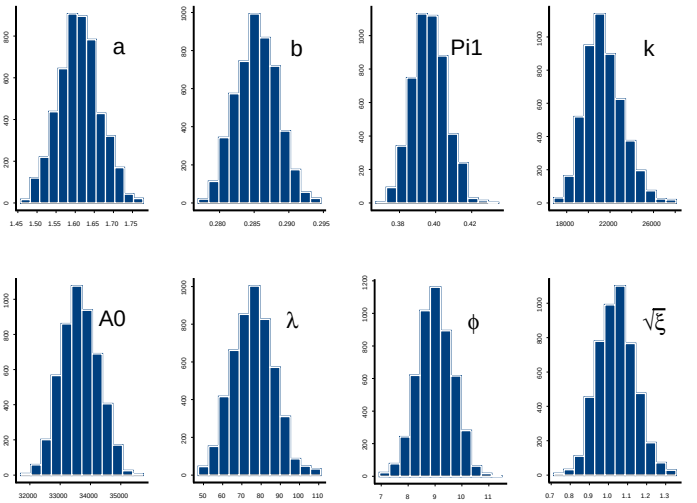

FIGURE 8. Histogram of 5000 posterior samples for parameters of interest based on 50 data sets.

The analysis here shows that for a short trajectory (such as the one with 784 observations) the two-state model is approximately fine; however, to describe the long run behavior of the DNA hairpin, the two-state model is insufficient, which indicates that the energy barrier between the close and open of the DNA hairpin has more complex behavior than the simple static picture depicted in the two-state model. The fluctuation of the energy barrier in this case may be due to conformational flexibility in other parts of the DNA molecule.

DISCUSSION

Although MCMC approaches illustrated in this article have found wide acceptance in the statistics community [9], their use for statistical estimation problems in other scientific disciplines is relatively uncommon. In the past, many researchers in physical sciences do not feel the necessity of delicate and efficient statistical inference methods in that the size of their data on the ensemble is often overwhelmingly large and ad-hoc methods such as moment-matching would be more than sufficient to provide needed information. The single-molecule experiments enabled by the advance of modern technology, as well as many large-scale genomics experiments, seem to have altered the landscape. As shown in this article, the Bayesian analysis provides much sharper estimates of the parameters associated with DNA hairpin dynamics compared with the existing moment-matching and binning methods.

We have discussed three important issues for efficiently analyzing single-molecule data: (a) synthesizing various stochastic models for conformational dynamics, (b) deriving likelihoods associated with these stochastic processes; (c) solving the experimental complications such as molecular diffusion and time-wrapping. Data augmentation techniques, aided with Markov chain Monte Carlo methods prove to be a very powerful tool. It not only is conceptually simple — the idea of augmenting the hidden processes is very intuitive — but also provides a viable means to circumvent the analytical intractability.

ACKNOWLEDGMENTS

The authors thank Haw Yang and Long Cai for helpful discussions and providing the single-molecule data. This work is supported in part by NSF grant DMS-0204674, Harvard University Clark-Cooke Fund and an NIH grant.

REFERENCES

1. Agmon, N. and Hopfi eld, J. J. (1983). *J. Chem. Phys.*, 78, 6947-6959.
2. Bonnet, G., Krichevsky, O. and Libchaber, A. (1998). *Proc. Natl. Acad. Sci.*, 95, 8602-8606.
3. Brown, F. L. and Silbey, R. J. (1998). *J. Chem. Phys.*, 108, 7434-7450.
4. Eggeling, C., Fries, J., Brand, L., Günther, R., and Seidel C. (1998). *Proc. Natl. Acad. Sci.*, 95, 1556-1561.
5. Froelich-Ammon, S., Gale, K. and Osheroff, N. (1994). *J. Biol. Chem.*, 269, 7719-7725.
6. Hastings, W. K. (1970). *Biometrika*, 57, 97-109.
7. Jia, Y., Sytnik, A., Li, L., Vladimirov, S., Cooperman, B. and Hochstrasser, R. (1997). *Proc. Natl. Acad. Sci.*, 94, 7932-7936.
8. Kou, S. C., Xie, X. S. and Liu, J. S. (2003). Bayesian analysis of single molecule experimental data. Preprint. Harvard University.
9. Liu, J. S. (2001). *Monte Carlo Strategies in Scientific Computing*. Springer, New York.
10. Lu, H. P., Xun, L. and Xie, X. S. (1998). *Science*, 282, 1877-1882.
11. Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. and Teller, E. (1953). *J. Chem. Phys.*, 21, 1087-1091.
12. Moerner, W. (2002). *J. Phys. Chem. B*, 106 (5), 910-927.
13. Nie, S. and Zare, R. (1997). *Ann. Rev. Biophys. Biomol. Struct.*, 26, 567-596.
14. Pfluegl, W., Brown, F. L. and Silbey, R. J. (1998). *J. Chem. Phys.*, 108, 6876-6883.
15. Reilly, P. D and Skinner, J. L. (1993). *Phys. Rev. Lett.*, 71, 4257-4260.
16. Reilly, P. D and Skinner, J. L. (1994a). *J. Chem. Phys.*, 101, 959-964.
17. Reilly, P. D and Skinner, J. L. (1994b). *J. Chem. Phys.*, 101, 965-973.
18. Schenter, G. K., Lu, H. P. and Xie, X. S. (1999). *J. Phys. Chem. A*, 103, 10477-10488.
19. Tamarat, P., Maali, A., Lounis B. and Orrit, M. (2000). *J. Phys. Chem. A*, 104 (1), 1-16.
20. Tang, J., Temsamani, J. and Agrawal, S. (1993). *Nucleic Acids Res.*, 21, 2729-2735.
21. Tanner, M. A. and Wong, W. H. (1987). *J. Amer. Statist. Assoc.*, 82, 528-540.
22. Trinh, T. and Sinden, R. (1993). *Genetics*, 134, 409-422.
23. Weiss, S. (2000). *Nature Struct. Biol.*, 7 (9), 724 - 729
24. Xie, X. S. and Lu, H. P. (1999). *J. Bio. Chem.*, 274, 15967-15970.
25. Xie, X. S. and Trautman, J. K. (1998). *Ann. Rev. Phys. Chem.*, 49, 441-480
26. Yang, H., Luo, G., Karnchanaphanurach, P., Louise, T.-M., Xun, L. and Xie, X. S. (2002). Single-molecule protein dynamics on multiple time scales probed by electron transfer. Department of Chemistry and Chemical Biology, Harvard University. Preprint.
27. Zazopoulos, E., Lalli, E., Stocco, D., and Sassone-Corsi, P. (1997). *Nature*, 390, 311-315.