PCCP

PAPER

View Article Online
View Journal | View Issue

Cite this: *Phys. Chem. Chem. Phys.*, 2018, **20**, 24099

Estimation of diffusive states from single-particle trajectory in heterogeneous medium using machine-learning methods

Yu Matsuda, *[a] Itsuo Hanasaki, [b] Ryo Iwao, [c] Hiroki Yamaguchi [c] and Tomohide Niimi[c]

Received 23rd April 2018,
Accepted 24th August 2018

DOI: 10.1039/c8cp02566e

rsc.li/pccp

We propose a novel approach to analyze random walks in heterogeneous medium using a hybrid machine-learning method based on a gamma mixture and a hidden Markov model. A gamma mixture and a hidden Markov model respectively provide the number and the most probable sequence of diffusive states from the time series position data of particles/molecules obtained by single-particle/molecule tracking (SPT/SMT) method. We evaluate the performance of our proposed method for numerically generated trajectories. It is shown that our proposed method can correctly extract the number of diffusive states when each trajectory is long enough to be frame averaged. We also indicate that our method can provide an indicator whether the assumption of a medium consisting of discrete diffusive states is appropriate or not based on the available amount of trajectory data. Then, we demonstrate an application of our method to the analysis of experimentally obtained SPT data.

1. Introduction

Single-particle/molecule tracking (SPT/SMT) techniques are widely used to investigate intercellular kinetics and biophysical processes at cell membranes[1–9] and extract non-bulk information on soft-materials.[10–12] As a quantity of interest for SPT/SMT trajectory data, mean-squared displacement (MSD) is widely used. The shape of an MSD indicates the modes of the particle/molecule motion: normal, anomalous, directed, and confined diffusion.[1,7,13] For example, an MSD for a trajectory of a particle/molecule in a homogeneous medium is a linear function of time, and the slope of the MSD indicates the diffusion coefficient. Though the transient behavior of MSD contains further information about particle/molecule motion, the MSD analysis is not convenient for practical use due to scattering of MSD plot induced by the limitation of available trajectory data. Thus, MSD is usually adopted to extract the diffusion coefficient of particles/molecule moving in a homogeneous medium. To analyze SPT/SMT trajectories containing adsorption/desorption motion and diffusion anisotropy, several analysis methods have been developed.[14–20]

In the past few years, new approaches based on Bayesian and machine-learning methods have been developed. Metzner *et al.*[21] developed a Bayesian inference method based on a time-discrete Ornstein–Uhlenbeck process. Their method is specialized for analyzing a Brownian motion with drift and can extract time-dependent statistical parameters. Ott *et al.*[22] applied a hidden Markov model (HMM) to an analysis of two different modes of diffusion: a fast and a slow diffusive state in heterogeneous medium. The drawback of their method is that the number of diffusive states has to be fixed *a priori* in their maximum likelihood approach to an HMM analysis. Persson *et al.*[23] proposed an analysis method, the variational Bayes SPT (vbSPT), based on a variational Bayesian treatment of an HMM.[24–26] Their vbSPT can estimate the number of diffusive states and state transition rates from trajectories in a heterogeneous medium. Their vbSPT is a powerful analysis method when a sufficient amount of trajectory data is available. However, it is often pointed out that a proper choice of a prior distribution is difficult and a large amount of data is required for a reliable estimation of parameters in a variational Bayesian treatment. For example, more than 3000 trajectories (one trajectory consists of steps less than 20) were analyzed for the reliable estimation of the number of diffusive states in the vbSPT.[23] Unfortunately, there remains many situations that limited number of samples is available[27–30] due to photobleaching of the dyes, limited imaging area, and so on. Moreover, since these studies[22,23] provided the discussion in the dimensional form, the users have to adjust parameters to their problem.

[a] *Department of Modern Mechanical Engineering, Waseda University, 3-4-1 Ookubo, Shinjuku-ku, Tokyo 169-8555, Japan. E-mail: y.matsuda@aoni.waseda.jp*
[b] *Department of Mechanical Systems Engineering, Tokyo University of Agriculture and Technology, Naka-cho 2-24-16, Koganei, Tokyo 184-8588, Japan*
[c] *Department of Micro-Nano Mechanical Science and Engineering, Nagoya University, Furo-cho, Chikusa, Nagoya, 464-8603, Japan*

In this study, we extend these HMM approaches by introducing a gamma mixture model (GMM).[31] In other word, we propose a hybrid machine learning method based on a GMM and an HMM. Our method is simple and can be easily implemented in a postprocessing program for SPT/SMT. The number of diffusive states is estimated by the GMM with the expectation–maximization (EM) algorithm[32] because the GMM is a more suitable classification algorithm than an HMM. An HMM originally estimate the most probable path of hidden states, where the number of the states are given as a prior knowledge.[32] By combining GMM and HMM, we can estimate most probable sequence of diffusive states from the trajectory data by HMM based on the number of the states estimated by the GMM in our method. We apply our proposed method to a numerically generated Brownian motion to investigate the performance. Since the previous studies such as ref. 20, 22 and 23 and our present method assumed that a medium consists of discrete diffusive states, we also discuss the statistical validity of the assumption from the given number of data points of trajectories. Then, the experimentally obtained SPT data is analyzed as a demonstration.

2. Numerical model of Brownian motion

We assume that a particle is a single point; that is, the position of a maker and the center of diffusion are indistinguishable. Since the positions of the maker and the diffusion center cannot be distinguished in an ordinary optical microscope due to the diffraction limit,[1,3,4] this assumption is reasonable in most SPT/SMT experiments. We consider the trajectory of a particle in a medium consisting of K media of the diffusion coefficients $\hat{D}_m$ ($m = 1, 2, 3, \ldots, K$), where $\hat{D}_m$ may be expressed as a function of position or time. The diffusion coefficients as functions of position and time correspond to a spatially heterogeneous medium such as a surface having adsorption sites[22,33–35] and to a temporally heterogeneous medium whose property is temporally controlled,[36,37] respectively. We also assume that the particle isotropically diffuses in each medium and the effect of interfaces between the media on the particle motion is negligible. The trajectory is generated by the Wiener process described by the overdamped Langevin equation as follows

$$\frac{\mathrm{d}\hat{\boldsymbol{r}}(\hat{t})}{\mathrm{d}\hat{t}} = \sqrt{2\hat{D}_m}\boldsymbol{\xi}(\hat{t}), \tag{1}$$

where $\hat{\boldsymbol{r}}(\hat{t})$ is the position vector of the particle at time $\hat{t}$. The variables $\boldsymbol{\xi}(\hat{t})$ is Gaussian random with $\mathbb{E}[\xi_i(\hat{t})] = 0$, $\mathrm{var}[\xi_i(\hat{t})] = 1$, and $\mathbb{E}[\xi_i(\hat{t})\xi_i(\hat{t}')] = \delta(\hat{t} - \hat{t}')$, where $\mathbb{E}[\cdots]$ and $\mathrm{var}[\cdots]$ indicate expectation and variance, respectively. The subscripts, $i = 1, 2, 3$, corresponding to components of axes x, y, z, respectively. We introduce the characteristic time τ and length $\sqrt{D_1\tau}$. Then, eqn (1) is nondimensionalized as

$$\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t} = \sqrt{2D_{\mathrm{R},m}}\boldsymbol{\xi}, \tag{2}$$

where the $\boldsymbol{r}$ and t are nondimensionalized quantities, and $D_{\mathrm{R},m}$ is the ratio of D_m to D_1 (the subscript R indicates ratio). Eqn (2) is numerically solved by the finite difference method expressed by

$$\boldsymbol{r}(t_{j+1}) = \boldsymbol{r}(t_j) + \sqrt{2\Delta t D_{\mathrm{R},m}}\boldsymbol{\xi}(t_j), \tag{3}$$

where t_j is the time of j-th time step, and Δt is the discretized time (*i.e.*, $t_j = j\Delta t$). The dimensional time step $\tau\Delta t$ corresponds to the time interval between each frame of SPT/SMT experiment. In this study, $\tau = 1$ to simplify the discussion. To obtain particle trajectories having a few diffusive states, $D_{\mathrm{R},m}$ is varied in the calculation of eqn (3). We treat $D_{\mathrm{R},m}$ as a function of time varying with given state transition probability, because it is difficult to obtain the trajectories under the controlled condition with treating $D_{\mathrm{R},m}$ as a function of position. Since the squared displacement is the only feature and does not contain the position information of the particle, this treatment of $D_{\mathrm{R},m}$ is reasonable in our method.

Fig. 1 shows the sample trajectory under the conditions of $t = 1.0 \times 10^{-2}$, $D_{\mathrm{R},1} = 1.0$, and $D_{\mathrm{R},2} = 5.0$. The total number of time steps is 50. The locations of a particle are classified by color according to the diffusive states.

3. Machine-learning methods of SMT/SPT data

3.1 Gamma mixture model

After experimentally/numerically obtaining trajectory data, a squared displacement $d(t_j)$ at the time of j-th time step is calculated as

$$d(t_j) = |\boldsymbol{r}(t_j) - \boldsymbol{r}(t_{j-1})|^2. \tag{4}$$

Considering ensemble average of the squared displacement, one can obtain mean squared displacement (MSD). We use only the squared displacement $d(t_j)$ as the feature in our proposed method; thus, our method can be applied to multi-dimensional trajectories without any modification. Since the probability distribution of $\boldsymbol{r}(t_j) - \boldsymbol{r}(t_{j-1})$ follows a Gaussian distribution (see eqn (3)), the probability distribution of $d(t_j)$ follows a gamma distribution. Then, the number of diffusive states and

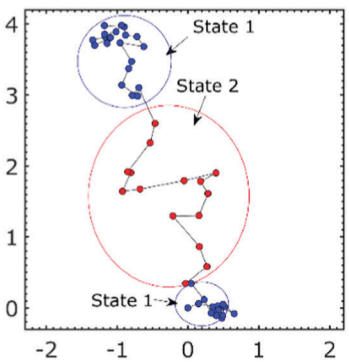

Fig. 1 Sample trajectory of Brownian motion with two diffusive states (state 1 and 2) under the conditions of $\Delta t = 1.0 \times 10^{-2}$, $D_{\mathrm{R},1} = 1.0$, and $D_{\mathrm{R},2} = 5.0$. The total number of time steps is 50.

its diffusion coefficients are estimated from the squared displacement $d(t_j)$ using a univariate GMM.[31]

We apply the EM algorithm to find maximum likelihood solutions for models having unobserved latent variables. Here, we introduce a GMM following the introduction procedure of a well-known mixture distribution, a Gaussian mixture model.[32] The univariate gamma distribution with a shape parameter α and an inverse scale parameter β is written as

$$G(x|\alpha,\beta)=\frac{x^{\alpha-1}}{\Gamma(\alpha)\beta^{\alpha}}\exp\left(-\frac{x}{\beta}\right), \tag{5}$$

where $\Gamma(x)$ is the gamma function. In the gamma mixture distribution, the probability distribution $p(x)$ is expressed as a linear superposition of gamma distributions as follows:

$$p(x)=\sum_{k=1}^{K}\pi_k G(x|\alpha_k,\beta_k), \tag{6}$$

where π_k are the mixing coefficients with $0 \le \pi_k \le 1$ and $\sum_{k=1}^{K}\pi_k = 1$. The number of gamma distributions having different parameters is represented as K, which also corresponds to the number of diffusive states in our method. Here, we introduce latent variables of a K-dimensional binary (0 or 1) random variable $\boldsymbol{z} = \{z_1, z_2,\ldots,z_K\}$ with $\sum_{k=1}^{K} z_k = 1$. This means that the probability of an observation x comes from z_k is represented by $p(z_k = 1) = \pi_k$. Since z is a binary random variable, the distribution $p(z_k = 1)$ can be written in the form

$$p(\boldsymbol{z})=\prod_{k=1}^{K}\pi_k^{z_k}. \tag{7}$$

The conditional distribution of x at given $\boldsymbol{z}$ is also expressed as

$$p(x|\boldsymbol{z})=\prod_{k=1}^{K}G(x|\alpha_k,\beta_k)^{z_k}. \tag{8}$$

When we consider a data set $\boldsymbol{X} = (x_1, x_2,\ldots,x_N)^{\mathrm{T}}$ consisting of N observations, the corresponding latent variables can be written by an $N \times K$ matrix of $\boldsymbol{Z}$ in which with n-th row is given by z_n^{T}, where the superscript T means transpose. Using eqn (7) and (8), the likelihood for the data set $\{\boldsymbol{X},\boldsymbol{Z}\}$ is expressed as

$$p(\boldsymbol{X},\boldsymbol{Z}|\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\pi})=\prod_{n=1}^{N}\prod_{k=1}^{K}\pi_k^{z_{nk}}G(x_n|\alpha_k,\beta_k)^{z_{nk}}, \tag{9}$$

where $\boldsymbol{\alpha} = (\alpha_1, \alpha_2,\ldots,\alpha_K)^{\mathrm{T}}$, $\boldsymbol{\beta} = (\beta_1, \beta_2,\ldots,\beta_K)^{\mathrm{T}}$, $\boldsymbol{\pi} = (\pi_1, \pi_2,\ldots,\pi_K)^{\mathrm{T}}$, and z_{nk} is n, k element of $\boldsymbol{Z}$. By maximizing the likelihood, we can find the best model to fit the data set. The log likelihood is written as

$$\ln p(\boldsymbol{X},\boldsymbol{Z}|\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\pi})=\sum_{n=1}^{N}\sum_{k=1}^{K}z_{nk}\{\ln\pi_k+\ln G(x_n|\alpha_k,\beta_k)\}, \tag{10}$$

The direct maximization of eqn (10) is difficult, because the $\boldsymbol{Z}$ is unobserved variables. Therefore, we consider the maximization problem of the expectation of eqn (10) with respect to the posterior distribution of the latent variables $\boldsymbol{Z}$. The posterior distribution $p(\boldsymbol{Z}|\boldsymbol{X},\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\pi})$ is obtained from eqn (7) and (8) with Bayes' theorem as

$$p(\boldsymbol{Z}|\boldsymbol{X},\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\pi})\propto\prod_{n=1}^{N}\prod_{k=1}^{K}\pi_k^{z_{nk}}G(x_n|\alpha_k,\beta_k)^{z_{nk}}. \tag{11}$$

Then the expectation of z_{nk} with respect to the posterior distribution is calculated as

$$\begin{aligned}\mathbb{E}[z_{nk}] &= \frac{\sum_{z_n} z_{nk}\prod_{l}[\pi_l G(x_n|\alpha_l,\beta_l)]^{z_{nl}}}{\sum_{z_n}\prod_{m}[\pi_m G(x_n|\alpha_m,\beta_m)]^{z_{nm}}}\\ &= \frac{\pi_k G(x_n|\alpha_k,\beta_k)}{\sum_m \pi_m G(x_n|\alpha_m,\beta_m)} =: \gamma(z_{nk}).\end{aligned} \tag{12}$$

Using eqn (12), the expectation of the log likelihood is expressed as

$$\begin{aligned}\mathbb{E}_{\boldsymbol{Z}}[\ln p(\boldsymbol{X},\boldsymbol{Z}|\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\pi})] &= \sum_{\boldsymbol{Z}} p(\boldsymbol{Z}|\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\pi})\ln p(\boldsymbol{X},\boldsymbol{Z}|\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\pi})\\ &= \sum_{n=1}^{N}\sum_{z_n} p(\boldsymbol{z}_n|\boldsymbol{x}_n,\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\pi})\ln p(\boldsymbol{x}_n,\boldsymbol{z}_n|\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\pi})\\ &= \sum_{n=1}^{N}\sum_{k=1}^{K}\gamma(z_{nk})\{\ln\pi_k+\ln G(x_n|\alpha_k,\beta_k)\}.\end{aligned} \tag{13}$$

As shown in eqn (13), the expectation of the log likelihood consists of the two terms, $\gamma(z_{nk})\ln\pi_k$ and $\gamma(z_{nk})\{\ln G(x|\alpha_k,\beta_k)\}$. First, we maximize the expectation of the log likelihood with respect to the mixing coefficients π_k. Since there is the constraint condition $\sum_{k=1}^{K}\pi_k = 1$, we introduce a Lagrange multiplier λ and consider the following Lagrange function $\Lambda(\boldsymbol{\pi},\lambda)$:

$$\Lambda(\boldsymbol{\pi},\lambda)=\sum_{n=1}^{N}\sum_{k=1}^{K}\gamma(z_{nk})\ln\pi_k+\lambda\left(\sum_{k=1}^{K}\pi_k-1\right) \tag{14}$$

Then, taking the derivative of $\Lambda(\boldsymbol{\pi},\lambda)$ with respect to π_k equal to 0, we can obtain

$$\sum_{n=1}^{N}\frac{\gamma(z_{nk})}{\pi_k}=-\lambda. \tag{15}$$

Multiplying eqn (15) by π_k and summing over k with the condition of $\sum_{k=1}^{K}\pi_k = 1$, the equation, $N = -\lambda$, is obtained. Then, eqn (15) can be written as

$$\pi_k=\frac{1}{N}\sum_{n=1}^{N}\gamma(z_{nk}). \tag{16}$$

To maximize the term $\gamma(z_{nk})\{\ln G(x|\alpha_k,\beta_k)\}$ in eqn (13), the partial derivatives of it with respect to α_k and β_k are set to be 0 as follows,

$$\frac{\partial}{\partial\alpha_k}\sum_{n=1}^{N}\sum_{k=1}^{K}\gamma(z_{nk})\ln G(x_n|\alpha_k,\beta_k)=0, \tag{17}$$

$$\frac{\partial}{\partial \beta_k}\sum_{n=1}^{N}\sum_{k=1}^{K}\gamma(z_{nk})\ln G(x_n|\alpha_k,\beta_k)=0. \tag{18}$$

Eqn (18) can be easily solved as

$$\beta_k=\frac{1}{\alpha_k}\frac{\sum_{n=1}^{N}\gamma(z_{nk})x_n}{\sum_{n=1}^{N}\gamma(z_{nk})}. \tag{19}$$

Using eqn (19), α_k is expressed as

$$\ln\alpha_k-\psi(\alpha_k)=\ln\frac{\sum_{n=1}^{N}\gamma(z_{nk})x_n}{\sum_{n=1}^{N}\gamma(z_{nk})}-\frac{\sum_{n=1}^{N}\gamma(z_{nk})\ln x_n}{\sum_{n=1}^{N}\gamma(z_{nk})}, \tag{20}$$

where ψ is the digamma function. Since eqn (20) cannot be analytically solved, α_k is numerically calculated. Now we can calculate the parameters for GMM using the EM algorithm: first, the initial values for parameters $\boldsymbol{\alpha}^{\mathrm{old}}$, $\boldsymbol{\beta}^{\mathrm{old}}$, $\boldsymbol{\pi}^{\mathrm{old}}$ are chosen. Second, as the expectation (E) step, $\gamma(z_{nk})$ is calculated from eqn (12) based on $\boldsymbol{\alpha}^{\mathrm{old}}$, $\boldsymbol{\beta}^{\mathrm{old}}$, $\boldsymbol{\pi}^{\mathrm{old}}$. Third, as the maximization (M) step, $\boldsymbol{\alpha}^{\mathrm{new}}$, $\boldsymbol{\beta}^{\mathrm{new}}$, $\boldsymbol{\pi}^{\mathrm{new}}$ are calculated from eqn (20), (19) and (16), respectively. The maximum likelihood parameters can be obtained by the iterative calculation of the E and M steps.

The number of gamma distributions having different parameters or the number of diffusive states K is extracted based on the Bayesian information criterion (BIC) represented as

$$\mathrm{BIC}(K)=-2\ln p(\boldsymbol{X}|\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\pi},K)+m\ln N, \tag{21}$$

where m is the number of estimated parameters. We adopt K to minimize BIC in eqn (21) as the number of diffusive states to avoid the over-fitting of GMM to the data. The ratios of diffusion coefficients, $D_{\mathrm{R},m}$, can be calculated as the mean values of each gamma distribution. When the EM algorithm converge to a local solution or some of the mixing coefficients are degenerate ($\pi_k \ll 1$), we seek the maximum likelihood solution by randomly resetting initial conditions of the parameters.

3.2 Hidden Markov model

After extracting the number of diffusion states, an HMM is applied to the sequential data of the squared displacements $d(t)$ to extract the most probable path of the diffusive sates. In this study, we use an HMM based on the maximum likelihood approach. Latent variables are also considered in the HMM. An observation $d(t_j)$ is generated through an emission probability (conditional distribution of the observation from a specific state)[32] from the corresponding latent variable as in the case of the GMM. The latent variable transfers from one state to another following the transition probability. The transition and emission probabilities are estimated by the Baum–Welch algorithm.[32] Then, using the estimated transition and emission probabilities, the most likely path is calculated by the Viterbi algorithm.[32] The detail of these algorithm is described in text books such as ref. 32. Since libraries of HMMs are widely distributed (ex. Statistics and Machine Learning Toolbox of Matlab[38] and hmmlearn for Python[39]), one can easily implement an HMM in their code. In the following calculation, we use the Statistics and Machine Learning Toolbox Ver. 11.2 of Matlab R2017a. We compare the results by Matlab with those by Python 3.5.2, showing good agreement between both results.

4. Results and discussions

We validate our proposed method using numerically generated trajectories (see eqn (3)) in which the number and sequence of diffusive states are given and known in advance. Then, we apply our method to experimentally obtained SPT data as a demonstration.

4.1 Classification performance of GMM for trajectories with localization error

We first compare our method with the existing methods.[20,23] In ref. 20 they compare their method with vbSPT[23] under several conditions. They numerically generated trajectories of 1000 tracks of 1000 frames. The conditions are given in dimensional form as $\hat{D}_1 = 0.015\ \mu\mathrm{m}^2\ \mathrm{s}^{-1}$, $\hat{D}_2 = 0.06\ \mu\mathrm{m}^2\ \mathrm{s}^{-1}$, and $\Delta\hat{t} = 4.0\times10^{-2}$ s in ref. 20. The lifetimes of each state are equal, and hence the fractions of each state are equal. The localization error of the particle is set at 20 nm. On this condition, the correctness of the estimation for the number of diffusive states are estimated and reported as 0.78 in ref. 20 and 0.8 by vbSPT.[23]

Following the numerical condition used in ref. 20, we numerically generate the trajectories of 1000 tracks of 1000 frames under the conditions of $D_{\mathrm{R},1} = 1.0$, $D_{\mathrm{R},2} = 4.0$, $\Delta t = 4.0\times10^{-2}$, and the localization error of 0.16, where the characteristic time and length are $\tau = 1$ s and 0.122 μm. A frame average is used to reduce the scattering of data in ref. 20. We also consider a frame average for the squared displacements at the time of j-th time step as follows

$$\bar{d}(t_j)=\frac{1}{2L+1}\sum_{l=-L}^{L}d(t_{j+l})=\frac{1}{2L+1}\sum_{l=-L}^{L}\left|\boldsymbol{r}(t_{j+l})-\boldsymbol{r}(t_{j+l-1})\right|^2, \tag{22}$$

where $2L+1$ is the number of averaged frames and $2L+1$ of the squared displacements are averaged. After calculating and averaging the squared displacement for each trajectory, we estimate the number of diffusive states using all trajectories. It is assumed that the trajectories are independent and identically distributed. In this study, we calculate the correctness of estimation for the number of diffusive states calculated from 50 independent trials. The correctness is 1.0 for $L = 1$ and $L = 2$. These L are equivalent to the value adopted in ref. 20. However, the correctness is lower (<0.1) for $L = 0$ (without a frame average). These results indicate that our method shows better performance when each trajectory is long enough to be frame averaged. This can be achieved under the conditions of relatively small Δt or long diffusion-state lifetime.[11,14] On the other hand, vbSPT[23] shows better performance for short trajectories as explained in ref. 20 and 23.

4.2 Performance for trajectories having two diffusive states without localization error

4.2.1 Typical example of estimation. We investigate the detailed performance of our method based on the analysis of the trajectories consisting of two diffusive states. Hereafter, we consider the trajectories without the localization error of the particles. Since the localization error will increase the apparent diffusion coefficient of the particles and obscure the true value, we compare the estimated diffusion coefficient with the exact one without considering the localization error. The trajectories are generated under the conditions of $D_{R,1} = 1.0$, $D_{R,2} = 4.0$, and $\Delta t = 4.0 \times 10^{-2}$. The particles transfer between the two diffusive states with the transition probability of 0.05 or stay in the same states with the probability of 0.95 during the time interval of Δt. We prepare 1000 trajectories with consecutive 30 frames (time steps). These conditions are similar to those discussed in the previous section and the literature[20,22,23] in which the conditions were determined as a model of membrane or protein kinetics. We extract the number of the diffusive states, their most probable sequence, and the transition probability from the trajectories using our method. The squared displacements, $d(t)$, defined by eqn (4) are calculated from the generated trajectory data. Then, the number of diffusive states and each diffusion coefficient are estimated from the probability distribution of the squared displacement using the GMM. We consider a frame average of the squared displacements defined as eqn (22) to reduce the scattering of the squared displacements and improve the estimation by the GMM. The choice of a certain value of L means that the probability of the state transition is sufficiently small in the frames of $2L + 1$, and L should be less than 10 in the experimental conditions reported in the literature.[20,22,23] By substituting $\bar{d}(t_j)$ for x_j, $\gamma(z_{nk})$ is calculated by eqn (12) as the E step. Then, as the M step, $\boldsymbol{\alpha}^{new}$, $\boldsymbol{\beta}^{new}$, and $\boldsymbol{\pi}^{new}$ are calculated from eqn (20), (19) and (16), respectively. In the GMM, the maximum likelihood parameters are obtained by the iterative calculation of the E and M steps. Fig. 2a shows the typical result of the GMM for the trajectories, where consecutive 5 ($L = 2$) steps of the squared displacements are averaged. The histogram of the squared displacement is also shown in the figure for the comparison. Note that the histogram is not used in the GMM calculation and the bin width does not affect the GMM result. Two peaks corresponding to the two diffusive states have been detected using the GMM. The ratios of diffusion coefficients are estimated as $D_{R,1} = 1.03$ and $D_{R,2} = 3.68$ in this example. In the HMM calculation, the initial estimation of transfer probability between each state is determined based on the number of frames for averaging (5 in this example). The most likely path for each trajectory is calculated by the Viterbi algorithm. Fig. 2b shows a typical example of an estimation for the most likely path. The correctness of the estimation is defined as

$$C_{HMM} = \frac{\text{number of correctly estimated states}}{\text{number of the squared displacements}}. \quad (23)$$

In this example, $C_{HMM} = 0.966$, and the reliability of the estimation is reduced near the transition from state 1 to state 2 (time step of 14 in Fig. 2b). The transition probability is estimated as 0.033 from state 1 to 2, where the exact probability is 0.05. This example shows that our proposed method is effective to extract the diffusive states and estimate the transfer between the states.

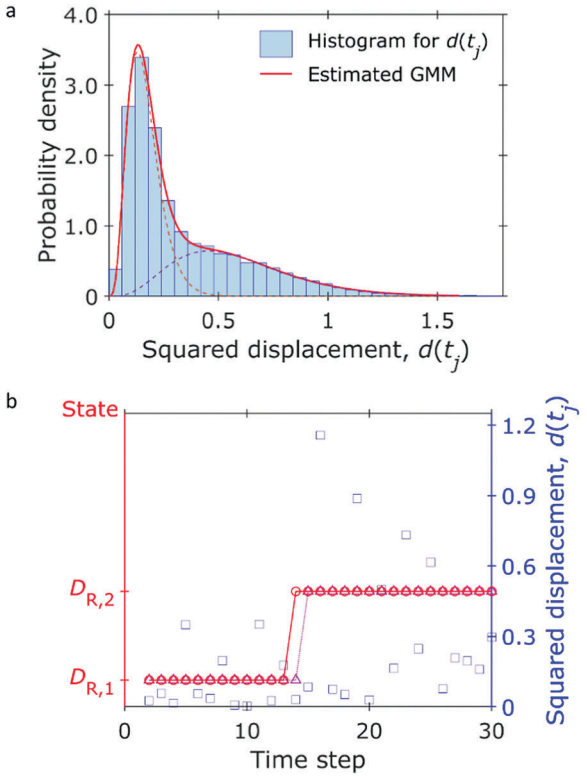

Fig. 2 Typical results of the estimation using our proposed method. The trajectory was generated under the conditions of $\Delta t = 4.0 \times 10^{-2}$, $D_{R,1} = 1.0$, and $D_{R,2} = 4.0$. The number of trajectories is 1000, and each trajectory consists of consecutive 30 time steps. (a) Estimation for the number of diffusive states using the GMM. (b) Most likely sequential path estimated by the Viterbi algorithm is shown as circles and the actual state path given in the trajectory generation is shown as triangles (left y axis). The squared displacements are also shown in squares (right y axis).

4.2.2 Effect of number of averaging frames and trajectories. We investigate the dependence of the classifying performance on the number of averaging frames ($2L + 1$). One thousand trajectories are prepared under the condition of $D_{R,1} = 1.0$, $D_{R,2} = 4.0$, $\Delta t = 4.0 \times 10^{-2}$, and each trajectory consists of 15 to 60 time steps. Fig. 3a shows the correctness of the estimated the number of diffusive states using the GMM, where the correctness of estimation by the GMM, C_{GMM}, is defined as following equation,

$$C_{GMM} = \frac{\text{number of correct estimations}}{\text{number of independent trials}}. \quad (24)$$

In this study, we calculate C_{GMM} from 50 independent trials. The tested numbers of averaging frames are from 3, 5, and 7 ($L = 1$, 2, 3). The correctness, C_{GMM}, drastically increases with increasing number of averaging frames or time steps as shown in Fig. 3a. The correctness, C_{GMM}, is 0 for $L = 0$, when the time steps of each trajectory is from 15 to 100. This result shows that

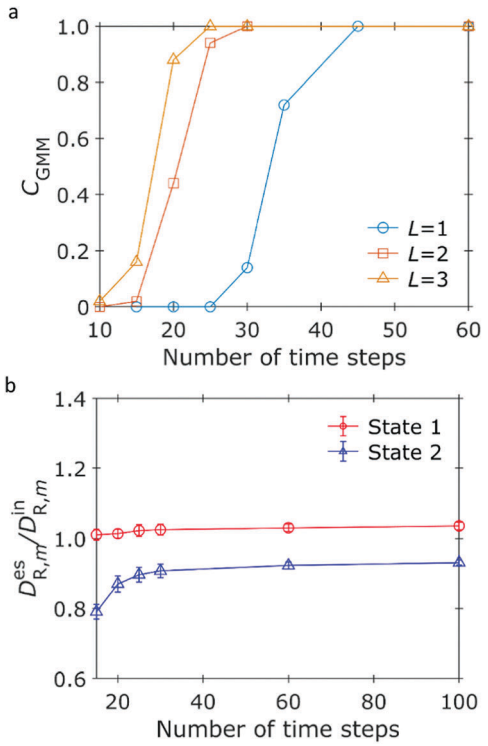

Fig. 3 Performance of the proposed method. Trajectories are generated under the conditions of $\Delta t = 4.0 \times 10^{-2}$, $D_{R,1} = 1.0$, and $D_{R,2} = 4.0$. (a) Correctness of estimation for number of diffusive states using the GMM. The number of averaging frames are $L = 1, 2, 3$. (b) Ratio of the estimated diffusion coefficients (superscript "es") to the input (superscript "in") coefficients for $L = 2$.

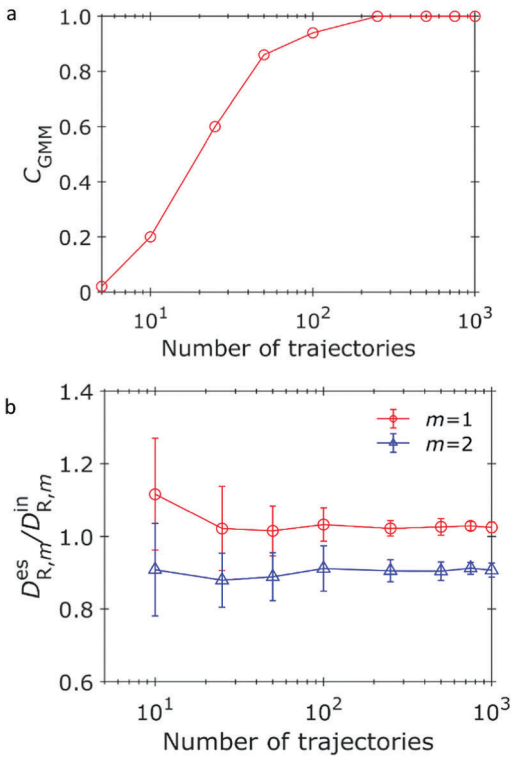

Fig. 4 Dependence of estimation on the number of trajectories. Trajectories of 30 time steps are generated under the conditions of $\Delta t = 4.0 \times 10^{-2}$, $D_{R,1} = 1.0$, and $D_{R,2} = 4.0$. (a) Correctness of estimation for number of diffusive states using the GMM. The number of averaging frames is $L = 2$. (b) Ratio of the estimated diffusion coefficients (superscript "es") to the input (superscript "in") coefficients for $L = 2$.

our method works well when each trajectory is long enough to calculate a frame average (>15). Fig. 3b shows the estimated $D_{R,1}$ and $D_{R,2}$ calculated from the correct estimations of the number of diffusion states for $L = 2$. The estimated value of $D^{es}_{R,2}$ approaches the assigned value for large numbers of time steps. Though $D^{es}_{R,1}$ is larger than the assigned value, the deviation of $D^{es}_{R,1}$ from the assigned value is small ($\sim 3\%$).

In the above discussion, the number of trajectories is fixed at 1000. In the following, we fix the time steps of each trajectory of 30 and vary the number of trajectories. The other conditions remain the same (*i.e.*, $D_{R,1} = 1.0$, $D_{R,2} = 4.0$, $\Delta t = 4.0 \times 10^{-2}$, $L = 2$). Fig. 4a shows the correctness of estimation by the GMM, C_{GMM}. The correctness C_{GMM} monotonically increases and is larger than 0.9 when the number of trajectories is larger than 50. The estimated $D_{R,1}$ and $D_{R,2}$ shown in Fig. 4b indicate that reliable estimation is achieved by our method even when the number of trajectories is small. For example, the $D_{R,1}$ and $D_{R,2}$ can be estimated with 10% when the number of trajectories larger than 25.

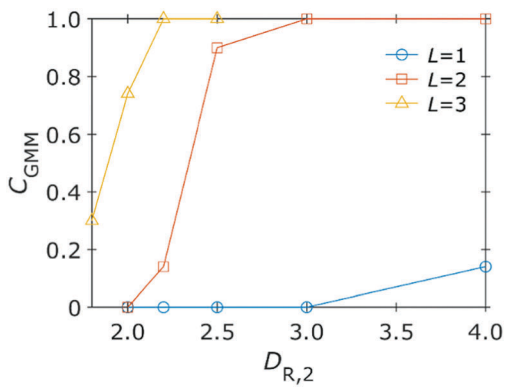

Fig. 5 Correctness of estimation for number of diffusive states using the GMM, C_{GMM}. The number of averaging frames are $L = 1, 2, 3$. The trajectory was generated under the conditions of $\Delta t = 4.0 \times 10^{-2}$ and $D_{R,1} = 1.0$ (fixed). The number of trajectories is 1000, and each trajectory consists of consecutive 30 time steps.

4.2.3 Effect of ratio of diffusion coefficient. The classifying performance would depend on the ratio of diffusion coefficient, $D_{R,2}$. We prepare 1000 trajectories consisting of 30 time steps under the conditions of $D_{R,1} = 1.0$, $1.8 \leq D_{R,2} \leq 4.0$, $\Delta t = 4.0 \times 10^{-2}$. Fig. 5 shows the correctness C_{GMM}, where $L = 1, 2, 3$. It is considered that the correct estimation will be difficult for small $D_{R,2}$. The correctness also drastically increases with increasing $D_{R,2}$ for $L = 2$, and the reliable estimation is achieved for $L = 3$ even when $D_{R,2}$ is small ($C_{GMM} = 0.74$ for $D_{R,2} = 2.0$). The correctness for $L = 1$ is low even when $D_{R,2} \geq 4$. Thus, a larger number of trajectories or time steps of each trajectory is required for reliable estimation when $L = 1$ (see also Fig. 3).

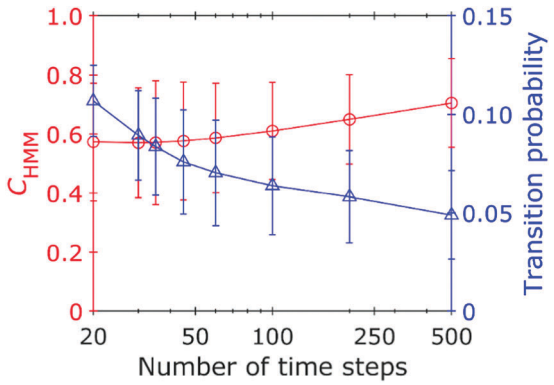

Fig. 6 Correctness of estimated state path and transition probability using the HMM, C_{HMM}. Trajectories are generated under the conditions of $\Delta t = 4.0 \times 10^{-2}$, $D_{R,1} = 1.0$, and $D_{R,2} = 4.0$. The transition probabilities from state 1 to 2 and *vice versa* are set at 0.05. The number of trajectories is 1000 and the time steps of each trajectory is shown as horizontal axis.

We also investigate the effect of the time step Δt for $L = 2$ and confirm that Δt does not affect the result because the mean ratio of the squared displacements does not change in different Δt.

4.2.4 Performance of HMM. The most probable path of diffusive sates is estimated for each trajectory. The correctness C_{HMM} and transfer probability depend on the number of time steps. Note that the frame average is not used in the HMM calculation. The trajectories are generated under the conditions of $D_{R,1} = 1.0$, $D_{R,2} = 4.0$, $\Delta t = 4.0 \times 10^{-2}$, and the transition probabilities from state 1 to 2 and *vice versa* are set at 0.05. These conditions are the same as those of Section 4.2.1. We prepare 1000 trajectories varying the time steps of each trajectory and calculate the most probable path for each trajectory by the HMM. Fig. 6 shows the correctness, C_{HMM}, and the estimated transfer probability from state 1 to 2. The correctness, C_{HMM}, gradually increases with increasing the time steps of each trajectory. However, C_{HMM} is smaller than C_{GMM}. Though there are a lot of trajectories without state transitions in the given time steps due to small transition probability and number of time steps, the HMM always classifies the data into two states; thus, C_{HMM} is small and the transition probability is over estimated for the small numbers of time steps.

4.3 Three diffusive states

We discuss the application of our method to the analysis of the trajectories for particles traveling in a medium having three diffusive states. We generate the trajectories under the following conditions: three diffusive states with $D_{R,1} = 1.0$ (state 1), $D_{R,2} = 4.0$ (state 2), and $D_{R,3} = 12.0$ (state 3) and the discretized time $\Delta t = 3.0 \times 10^{-3}$. The characteristic time and length are $\tau = 1$ s and 0.5 μm, respectively. The particles transfer each state during the time interval of Δt under the following probabilities: the transition probability from state 1 to state 2 is 0.031 and to state 3 is 0.1. The transition probability from state 2 to state 1 is 0.017 and to state 3 is 0.01. The transition probability from state 3 to state 1 is 0.055 and to state 3 is 0.055. These conditions correspond to those discussed in Persson *et al.*,[23]

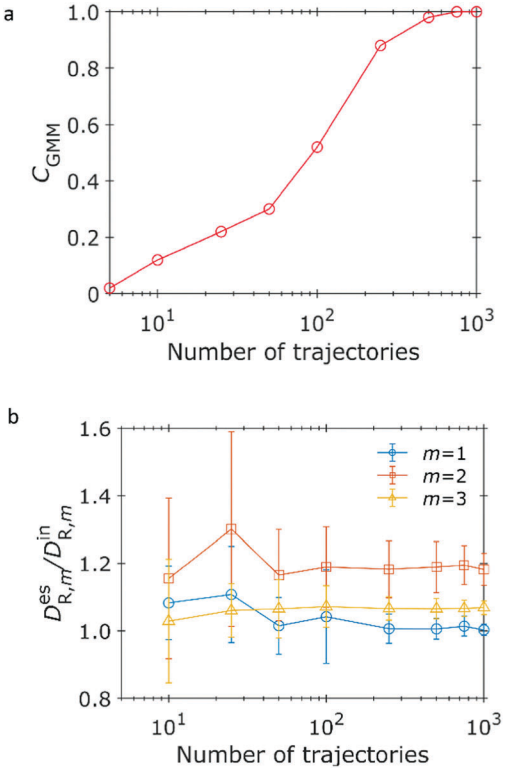

Fig. 7 Dependence of estimation on number of trajectories. Trajectories of 30 time steps are generated under the conditions of $\Delta t = 3.0 \times 10^{-3}$, $D_{R,1} = 1.0$, $D_{R,2} = 4.0$, and $D_{R,3} = 12.0$. (a) Correctness of estimation for number of diffusive states using the GMM. The number of averaging frames are $L = 2$. (b) Ratio of the estimated diffusion coefficients (superscript "es") to the input (superscript "in") coefficients for $L = 2$.

where the parameters were given in dimensional quantities: $\hat{D}_1 = 0.25$ μm^2 s^{-1}, $\hat{D}_2 = 1.0$ μm^2 s^{-1}, $\hat{D}_3 = 3.0$ μm^2 s^{-1}, and $\Delta\hat{t} = 3.0 \times 10^{-3}$ s.

We estimate the number of diffusive states, where the time steps of each trajectory are fixed at 30. The correctness for the estimation of number of diffusive states is calculated from 50 independent trials as shown in Fig. 7a. The correctness rapidly increases in 100 trajectories and reaches to unity over 750 trajectories. The correctness for this 3 states estimation is smaller than that for 2 states estimation shown in Fig. 4a because the number of data in each state is smaller by construction. Fig. 7b shows the estimated ratio of diffusion coefficients. The estimated $D_{R,1}$ converges to the input value with 1000 trajectories. On the other hand, the estimated $D_{R,2}$ and $D_{R,3}$ are 10% or 20% lager than the input values because the variances of squared displacement generated from $D_{R,2}$ and $D_{R,3}$ are larger than that from $D_{R,1}$. Our method works well in the 3-states problem.

4.4 Distinction of discrete diffusive states

In the above discussions, we have considered the particle trajectories in a medium having discrete diffusive states. The medium is assumed as consisting of several discrete diffusive states because HMM can treat only discrete hidden states.[32] This is one of the limitation of the HMM. It is sometimes convenient to regard a medium consisting of multiple discrete

diffusive states or particle motion is approximately expressed as such, when the underlying process is unclear before the analysis.[28,40–42] Therefore, it is worth clarifying how much amount of data is necessary for treating a medium consisting of discrete diffusive states. Fig. 3a, 4a, 5 and 7a can be understood as showing the required amount of data to determine whether a medium consists of discrete diffusive states or not by the GMM based on BIC.

As a comparison to a medium having several discrete diffusive states, we consider a case where a particle travels in a medium having a continuously varying diffusion coefficient. One thousand trajectories are generated by eqn (3), where $D_{R,m}$ is replaced by the time varying diffusion coefficient $D_R(t_j)$. The diffusion coefficient linearly varies from 1 to D_{max} as $D_R(t_j) = (D_{max} - 1)(j - 1)/(N - 1) + 1$, where $D_{max} = 4.0$ and the time steps $N = 30$. This condition corresponds to that discussed in Section 4.2.2 except for a continuously varying diffusion coefficient, and the correctness is unity when the number of frame averaging is over 5 ($L \geq 2$) in Fig. 3a. Fig. 8 shows the probability densities for the squared displacements of the trajectories with the continuously varying diffusion coefficient $D_R(t_j)$ and the discrete diffusion states ($D_{R,1} = 1.0$, and $D_{R,2} = 4.0$) for comparison, where the number of frame averaging is 5 ($L = 2$). The solid and dashed curves are obtained by the GMM in which the number of the mixing distribution is determined by BIC. The GMM returns a single state with the diffusion coefficient of 2.45 corresponding to the time average of $D_R(t_j)$ for the trajectory with the continuously varying diffusion coefficient (solid line). For the trajectory traveling in the discrete states, the discrete diffusive states are successfully classified by the GMM (dashed line). As shown in figures from 3 to 7, the GMM can classify the discrete diffusive states when the sufficient number of time steps is available and/or the number of averaging frames is sufficiently large for the trajectories with the discrete diffusive states. By contrast, for the trajectory in the continuously varying diffusion coefficient, the estimated number of diffusive states is determined as one by our method even when the large number of times steps and/or averaging frames are used. Our method provides a statistical validity of the assumption that a medium can be treated as consisting of discrete diffusive states for the given number of data points of trajectory.

4.5 Analysis of experimentally obtained SPT data

Now, we apply our proposed method to the analysis of experimentally obtained SPT data. The SPT data was obtained by an optical microscope in the same manner as that of our previous study.[43] We used ZnS–$AgInS_2$ nanoparticles[44–46] as probe particles, and mixed it into a polydimethylsiloxane (PDMS) layer. An oil immersion objective lens with 100 times magnification with N.A. of 1.4 was used. The time step (frame interval) was $\Delta t = 0.2$ s. The SPT images were captured after 20 h adding a curing agent in the PDMS layer at a room temperature of 20 °C. The SPT images were analyzed using IDL-code.[47] For the completely cured PDMS layer, the diffusion coefficient was calculated as $O(10^{-4})$ μm^2 s^{-1}. This indicates that the localization error is considered as $O(10^2)$ nm. Under this condition, there were some diffusive states in the PDMS layer.[43] Fig. 9a shows the estimation result of the number of diffusive states using the GMM. Fig. 9b shows the single particle trajectory having 185 location data points with the most probable diffusive states. In the GMM analysis, it was found that there were three diffusive states of diffusion coefficients of 4×10^{-3}, 8×10^{-3}, and 12×10^{-3} μm^2 s^{-1}. These values reasonably agreed with those obtained by MSD analysis.[43] The particle

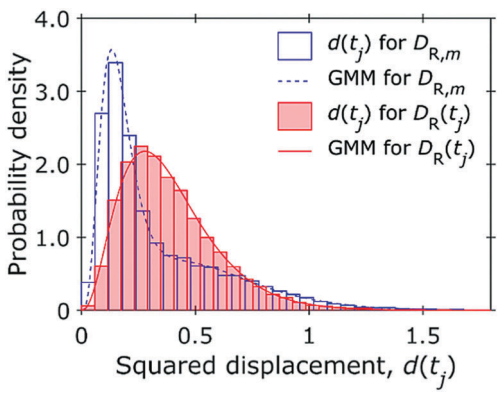

Fig. 8 Typical results of probability densities for trajectories in a medium with continuously varying diffusion coefficient and that consisting of two discrete diffusive states. The solid and dashed curves show the result of the GMM for each condition.

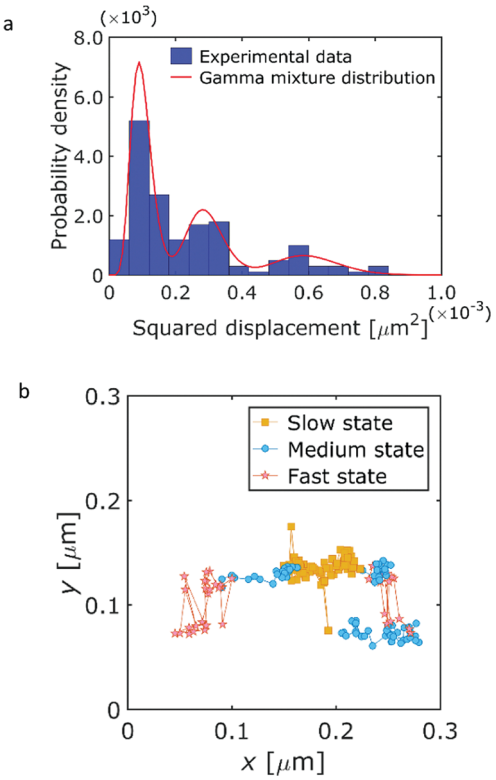

Fig. 9 Analysis results for the experimentally obtained SPT trajectory. (a) The estimation for number of diffusive states using GMM. (b) The classified result using HMM.

motions having relatively large displacement can be captured as the fast state (large diffusion coefficient). Though there is the place around $x = 0.25$, $y = 0.12$ containing two diffusive states due to the estimation error induced from the small amount of the data, each diffusive state occupies different place in the layer as shown in Fig. 9b. It is considered that the PDMS layer consists of several portions having different viscosity, when we assume discrete diffusive states. We conclude that our proposed method is applicable to the analysis of experimental SPT/SMT data.

5. Conclusions

In this article, we have proposed a hybrid method of a gamma mixture model (GMM) and a hidden Markov model (HMM) to classify particle trajectory for a particle moving in a heterogeneous medium. We introduce a GMM as an extension of a Gaussian mixture model based on the expectation–maximization (EM) algorithm. The number of diffusive states is estimated by the GMM, and then the HMM is used to extract the most likely path of the diffusive states. The correct estimation of the number of diffusive states can be achieved from small amount of trajectory data by considering frame average of squared displacements. The transition path of diffusive states is estimated by the Viterbi algorithm based on the estimated number of diffusive states. We compare our method with existing methods by calculating trajectory for a particle moving in a medium having two diffusive states. It is shown that our proposed method can extract the number of diffusive states more reliably than existing methods when the number of averaging frames is large. Thus, our method is a powerful tool for the trajectories obtained with relatively large frame rate or having long diffusion-state lifetime. Furthermore, we also indicate that our method can provide an indicator whether the assumption of a medium consisting of discrete diffusive states is appropriate or not based on the amount of the given data. Our hybrid method of the GMM and HMM is promising method for analyzing single-particle/molecule tracking data when limited number of data is available.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

We thank K. Tomita for his assistance with the simulations by Python. This work was partly supported by a JSPS Grant-in-Aid for Scientific Research (B), No. 16H04277 and a research encouragement grants from the Asahi Glass Foundation.

References

1 M. J. Saxton and K. Jacobson, *Annu. Rev. Biophys. Biomol. Struct.*, 1997, **26**, 373–399.
2 M. J. Saxton, *Biophys. J.*, 1997, 72.
3 W. E. Moerner and D. P. Fromm, *Rev. Sci. Instrum.*, 2003, **74**, 3597.
4 J. R. Lakowicz, *Principles of Fluorescence Spectroscopy*, Springer, 3rd edn, 2006.
5 N. Ruthardt, D. C. Lamb and C. Brauchle, *Mol. Ther.*, 2011, **19**, 1199–1211.
6 A. Kusumi, T. A. Tsunoyama, K. M. Hirosawa, R. S. Kasai and T. K. Fujiwara, *Nat. Chem. Biol.*, 2014, **10**, 524–532.
7 C. Manzo and M. F. Garcia-Parajo, *Rep. Prog. Phys.*, 2015, **78**, 124601.
8 T. Chen and B. M. Reinhard, *Small*, 2013, **9**, 876–884.
9 J. A. Varela, C. Aberg, J. C. Simpson and K. A. Dawson, *Small*, 2015, **11**, 2026–2031.
10 D. Woll, H. Uji-i, T. Schnitzler, J. Hotta, P. Dedecker, A. Herrmann, F. C. De Schryver, K. Mullen and J. Hofkens, *Angew. Chem., Int. Ed. Engl.*, 2008, **47**, 783–787.
11 S. Ito, K. Itoh, S. Pramanik, T. Kusumi, S. Takei and H. Miyasaka, *Appl. Phys. Express*, 2009, **2**, 075004.
12 K. Paeng and L. J. Kaufman, *Macromolecules*, 2016, **49**, 2876–2885.
13 H. Qian, M. P. Sheetz and E. L. Elson, *Biophys. J.*, 1991, **60**, 910–921.
14 L. C. Elliott, M. Barhoum, J. M. Harris and P. W. Bohn, *Phys. Chem. Chem. Phys.*, 2011, **13**, 4326–4334.
15 C. Ribrault, A. Triller and K. Sekimoto, *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.*, 2007, **75**, 021112.
16 S. Burov, S. M. Tabei, T. Huynh, M. P. Murrell, L. H. Philipson, S. A. Rice, M. L. Gardel, N. F. Scherer and A. R. Dinner, *Proc. Natl. Acad. Sci. U. S. A.*, 2013, **110**, 19689–19694.
17 I. Hanasaki and Y. Isono, *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.*, 2012, **85**, 051134.
18 I. Hanasaki, S. Uehara, Y. Arai, T. Nagai and S. Kawano, *Jpn. J. Appl. Phys.*, 2015, **54**, 125601.
19 Y. Matsuda, I. Hanasaki, R. Iwao, H. Yamaguchi and T. Niimi, *Anal. Chem.*, 2016, **88**, 4502–4507.
20 P. J. Bosch, J. S. Kanger and V. Subramaniam, *Biophys. J.*, 2014, **107**, 588–598.
21 C. Metzner, C. Mark, J. Steinwachs, L. Lautscham, F. Stadler and B. Fabry, *Nat. Commun.*, 2015, **6**, 7516.
22 M. Ott, Y. Shai and G. Haran, *J. Phys. Chem. B*, 2013, **117**, 13308–13321.
23 F. Persson, M. Linden, C. Unoson and J. Elf, *Nat. Methods*, 2013, **10**, 265–269.
24 Z. Ghahramani, *Int. J. Pattern Recognit. Artif. Intell.*, 2001, **15**, 9–42.
25 J. E. Bronson, J. Fei, J. M. Hofman, R. L. Gonzalez, Jr. and C. H. Wiggins, *Biophys. J.*, 2009, **97**, 3196–3205.
26 K. Okamoto and Y. Sako, *Biophys. J.*, 2012, **103**, 1315–1324.
27 Y. Gu, X. Di, W. Sun, G. Wang and N. Fang, *Anal. Chem.*, 2012, **84**, 4111–4117.
28 S. Habuchi, S. Fujiwara, T. Yamamoto, M. Vacha and Y. Tezuka, *Anal. Chem.*, 2013, **85**, 7369–7376.
29 L. O. Mair and R. Superfine, *Soft Matter*, 2014, **10**, 4118–4125.
30 M. Kanke, E. Tahara, P. J. Huis In't Veld and T. Nishiyama, *EMBO J.*, 2016, **35**, 2686–2698.

31 A. R. Webb, *Pattern Recogn.*, 2000, **33**, 2045–2054.
32 C. M. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)*, Springer-Verlag Inc., New York, 2006.
33 H. Shen, L. J. Tauzin, W. Wang, B. Hoener, B. Shuang, L. Kisley, A. Hoggard and C. F. Landes, *Anal. Chem.*, 2016, **88**, 9926–9933.
34 M. J. Skaug, J. N. Mabry and D. K. Schwartz, *J. Am. Chem. Soc.*, 2014, **136**, 1327–1332.
35 C. Yu, J. Guan, K. Chen, S. C. Bae and S. Granick, *ACS Nano*, 2013, **7**, 9735–9742.
36 S. Juodkazis, N. Mukai, R. Wakaki, A. Yamaguchi, S. Matsuo and H. Misawa, *Nature*, 2000, **408**, 178–181.
37 A. M. Kloxin, A. M. Kasko, C. N. Salinas and K. S. Anseth, *Science*, 2009, **324**, 59–63.
38 MathWorks, Statistics and Machine Learning Toolbox, https://jp.mathworks.com/products/statistics.html.
39 R. Weiss, S. Du, J. Grobler, D. Cournapeau, F. Pedregosa, G. Varoquaux, A. Mueller, B. Thirion, D. Nouri, G. Louppe, J. Vanderplas, J. Benediktsson, L. Buitinck, M. Korobov, R. McGibbon, S. Lattarini, V. Niculae, csytracy, A. Gramfort, S. Lebedev, D. Huppenkothen, C. Farrow and A. Yanenko, hmmlearn: Hidden Markov Models in Python, with scikit-learn like API, http://hmmlearn.readthedocs.io/en/latest/.
40 J. Kirstein, B. Platschek, C. Jung, R. Brown, T. Bein and C. Brauchle, *Nat. Mater.*, 2007, **6**, 303–310.
41 C. Hellriegel, J. Kirstein and C. Bräuchle, *New J. Phys.*, 2005, **7**, 23.
42 S. Habuchi, N. Satoh, T. Yamamoto, Y. Tezuka and M. Vacha, *Angew. Chem., Int. Ed. Engl.*, 2010, **49**, 1418–1421.
43 R. Iwao, Y. Matsuda, H. Yamaguchi and T. Niimi, *MHS*, 2015, DOI: 10.1109/MHS.2015.7438256.
44 T. Torimoto, T. Adachi, K. Okazaki, M. Sakuraoka, T. Shibayama, B. Ohtani, A. Kudo and S. Kuwabata, *J. Am. Chem. Soc.*, 2007, **129**, 12388–12389.
45 T. Torimoto, S. Ogawa, T. Adachi, T. Kameyama, K. Okazaki, T. Shibayama, A. Kudo and S. Kuwabata, *Chem. Commun.*, 2010, **46**, 2082–2084.
46 Y. Matsuda, T. Torimoto, T. Kameya, T. Kameyama, S. Kuwabata, H. Yamaguchi and T. Niimi, *Sens. Actuators, B*, 2013, **176**, 505–508.
47 J. C. Crocker and D. G. Grier, *J. Colloid Interface Sci.*, 1996, **179**, 298.