Hidden Markov model analysis for fluorescent time series of quantum dots

Tatsuhiro Furuta,* Keisuke Hamada, Masaru Oda, and Kazuma Nakamura†
Graduate School of Engineering, Kyushu Institute of Technology, Kitakyushu 804-8550, Japan

(Received 21 June 2022; accepted 1 September 2022; published 19 September 2022)

We present a hidden Markov model analysis for fluorescent time series of colloidal quantum dots. A fundamental quantity to measure optical performance of the quantum dots is a distribution function for the light-emission duration. So far, to estimate it, a threshold value for the fluorescent intensity was introduced, and the light-emission state was evaluated as a state above the threshold. With this definition, the light-emission duration was estimated, and its distribution function was derived as a blinking plot. Due to the noise in the fluorescent data, however, this treatment generates a large number of artificially short-lived emission states, thus leading to an erroneous blinking plot. In the present paper, we propose a hidden Markov model to eliminate these artifacts. The hidden Markov model introduces a hidden variable specifying the light-emission and quenching states behind the observed fluorescence. We found that it is possible to avoid the above artifacts by identifying the state from the hidden-variable time series. We found that, from the analysis of experimental and theoretical benchmark data, the accuracy of our hidden Markov model is beyond human cognitive ability.

DOI: 10.1103/PhysRevB.106.104305

I. INTRODUCTION

Recently, studies on analyzing various material data by machine-learning techniques to propose new materials and new synthetic methods have become active [1–9]. The development of databases including material structures and their physical properties is progressing [10–13], and the development of software for facilitating machine-learning analyses is also making great progress [14,15]. Most of the research so far has focused on the correlation analysis of static data of materials, but recently, there has been growing interest in analyses for the dynamic data (time-dependent data) [16–26].

In the present paper, we present an analysis for time-series data on optical properties of single colloidal quantum dots (QDs). The QD is a fine particle of several nanometers and is a system that can confine electrons and holes in all dimensions. Due to its unique electronic and fluorescent properties, there are wide applications including single-electron transistors [27], quantum teleportation [28], QD lasers [29], QD solar cells [30], and quantum computers [31]. From the viewpoint of the performance of QDs as emitters, it is desirable that the light emission continue as long as possible under continuous excitation, leading to a high performance on luminescence quantum yield. However, it is known that the single colloidal QDs show blinking behaviors [32] in most cases. As a fundamental property for this measure, a distribution function on the light-emission states of the single QD is widely employed and called a blinking plot [33].

In the experiment, we irradiate the QD with the light continuously and observe a situation where the QD emits or quenches in real time. Figure 1 shows an example of the observed fluorescent intensities of a single QD as a function of time. As can be seen from the figure, the time series of the fluorescent intensity is noisy due to the experimental equipment and the measurement environment. It is difficult to quantitatively distinguish the light-emission and quenching states especially in such case that the noises are larger than or comparable to the fluorescent signals. Conventionally, in order to define the light-emission states, an artificial threshold is introduced by hand in the intensity analysis, and all the states beyond this threshold are evaluated as the light-emission state. With this definition, the light-emission duration is estimated, and a distribution function on the duration is evaluated. With this method, however, the result will include a large amount of short light-emission duration that does not actually exist, and then the obtained distribution function would clearly be erroneous.

In this study, we present a time-series data analysis using a hidden Markov model (HMM) [34,35] to solve this problem. The HMM is a typical machine-learning method to handle time-series data, and is applied to various dynamic data analyses such as stock price prediction [36] and anomaly detection [37]. In materials science, it is also beginning to be used for several applications including random telegraph noise [16–19], nitrogen-vacancy centers in diamond [20,21], electron holograms [22], atom quantum jump dynamics [23], fluorescence resonance energy transfer [24], and crack propagation [25]. The HMM introduces hidden variables for specifying states behind the observed real time series. We will show that the above mentioned noises in the fluorescence data can be eliminated with the HMM method; the time series of the hidden variable are noise-suppressed, and as a result, the duration evaluation becomes stable.

The present paper is organized as follows: In Sec. II, we describe a basic idea for treating the present problem, details

*furuta.tatsuhiro704@mail.kyutech.jp
†kazuma@mns.kyutech.ac.jp

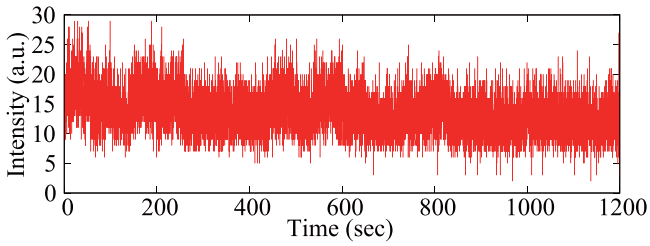

FIG. 1. Experimental fluorescence time-series data of quantum dots. See Sec. III A for experimental details and conditions.

of the HMM, and how to analyze the time series data with the HMM. In Sec. III, we present results of the HMM analysis for experimental fluorescence data. On top of that, to verify accuracy of the present HMM analysis, we show performance for the theoretical benchmark data. Section IV summarizes conclusions. We also in the Appendix describe details for the blocking Gibbs sampling to solve the HMM.

II. METHOD

A. Basic idea on state identification

Before presenting details, we first show the basic idea for studying blinking phenomena. As an example, we consider a time series of a fluorescent intensity $I(t)$ in Fig. 2(a). In this time series, the intensity is high at the regions of $t_1 \leqslant t \leqslant t_2$ and low in the range of $0 \leqslant t \leqslant t_1$ and $t_2 \leqslant t$. Here, we define the former region as an ON (bright) state and the latter region as an OFF (dark) state. The time series contains a noise η, which makes it difficult to determine the ON or OFF state; if the level difference Δ between the two states competes with the η, the state assignment becomes difficult.

In the conventional method, a threshold intensity I_{th} in panel (a) is introduced to define the ON and OFF states; if the intensity $I(t)$ is larger than the I_{th}, the state at the time t is specified as the ON states. On the other hand, if the intensity $I(t)$ is smaller than the I_{th}, the state is classified as the OFF state. Figure 2(b) shows the statistics on the duration τ of the ON state, based on the conventional method (see Sec. II C for details). Since the intensity $I(t)$ largely fluctuates due to the noise, many short-duration states of $\tau \sim 0.1$–0.5 s are generated. In the normal sense, however, in the case of this time series, the ON duration would be $t_2 - t_1 = 62.6$ s and there are no shorter ON states. From this discussion, the evaluation of the duration with the conventional method would bring about an erroneous result for the noisy time series.

Next, we consider the state assignment by the HMM. The HMM introduces a time series of a hidden variable $s(t)$ that governs the generation of the $I(t)$ (see Sec. II B). When $s(t) = 1$, the state at the time t is ON associated with the high fluorescent intensity, while when $s(t) = 0$, the state at the time t is OFF with the low intensity. The $s(t)$ time series can be determined with the Bayesian inference for the observed time series $I(t)$. Note that the noise is suppressed in the $s(t)$ time series [Fig. 2(c)]. Since the $s(t)$ has basically a binary feature, by introducing a threshold s_{th} in panel (c), we can perform a stable state assignment.

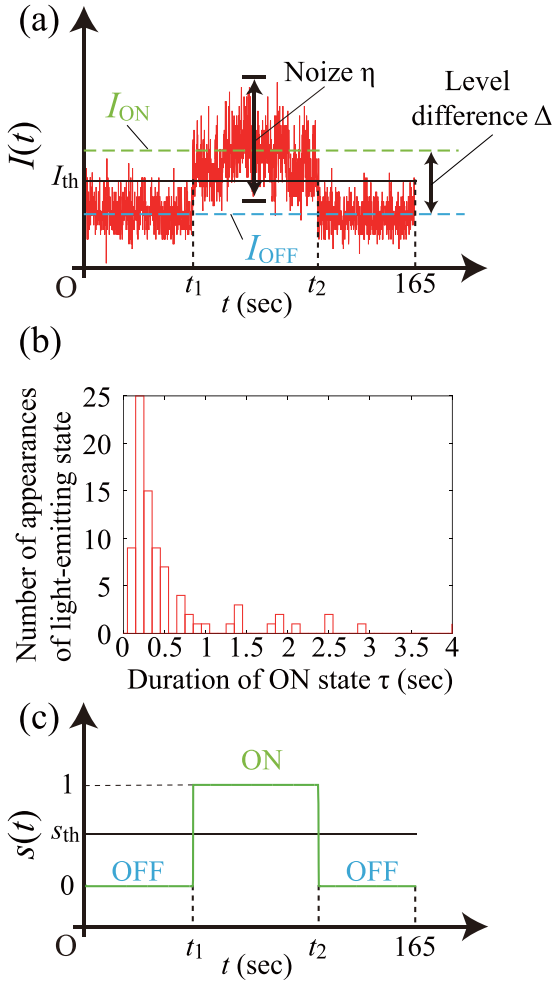

FIG. 2. Summary of an analysis for fluorescent time series. (a) Time series of the fluorescent intensity $I(t)$, where I_{ON} and I_{OFF} are the guides representing baselines of intensity for the ON and OFF states. Δ and η are the level difference between the two states and noise, respectively. I_{th} is a threshold to distinguish the ON and OFF states, which is represented by a solid line (see the text). (b) Number of appearances of the ON (bright) state, where we count the appearances by duration for the time series in panel (a). (c) Time series of a hidden variable $s(t)$ obtained from HMM simulations for the fluorescent time series in panel (a) (see Sec. II B for details), where $s(t) = 1$ and $s(t) = 0$ indicate the ON and OFF states, respectively. Note that the noise is totally suppressed in the $s(t)$ time series.

B. Hidden Markov model

The HMM is in general expressed with a combination of various probability distributions of stochastic variables [34,35]. The form of the HMM employed in the present study is shown below:

$$p(\boldsymbol{I}, \mathbf{S}, \boldsymbol{\Theta}, \boldsymbol{\pi}, \mathbf{A}) = p(\boldsymbol{I}|\mathbf{S}, \boldsymbol{\Theta})p(\mathbf{S}|\boldsymbol{\pi}, \mathbf{A})p(\boldsymbol{\Theta})p(\boldsymbol{\pi})p(\mathbf{A}). \quad (1)$$

Here, $\boldsymbol{I}$ is observed data, $\mathbf{S}$ is hidden variables, $\boldsymbol{\Theta}$ describes parameters characterizing the distribution function of $\boldsymbol{I}$, $\boldsymbol{\pi}$ is parameters to set the initial step of the hidden variable, and $\mathbf{A}$ describes a transition matrix for the time evolution of hidden variables. The $p(\mathbf{S}|\boldsymbol{\pi}, \mathbf{A})$ in Eq. (1) describes the conditional

probability distribution of $\mathbf{S}$ with $\mathbf{A}$ and $\boldsymbol{\pi}$ fixed, and similarly, $p(\boldsymbol{I}|\mathbf{S}, \boldsymbol{\Theta})$ is the conditional probability distribution of $\boldsymbol{I}$ after $\mathbf{S}$ and $\boldsymbol{\Theta}$ are determined. We note that the form in Eq. (1) is a standard type; it is not a special one which is optimized or modified for the present problem, and is basically the same as that used in the usual time-series data analysis such as stock price and seismic wave data.

$\boldsymbol{I}$ in Eq. (1) is discrete time-series data written as

$$\boldsymbol{I} = (I_1, I_2, \ldots, I_N)^T, \tag{2}$$

where I_n is an observed value at the nth time grid, and the total number of grids is N. In the present study, the $\boldsymbol{I}$ corresponds to the time series of the fluorescent intensity of a single QD.

$\mathbf{S}$ in Eq. (1) is hidden variables written as

$$\mathbf{S} = (\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_N), \tag{3}$$

which is also a time series. The component $\mathbf{s}_n$ is a vector and consists of K components as

$$\mathbf{s}_n = (s_{1n}, s_{2n}, \ldots, s_{Kn})^T. \tag{4}$$

The $\mathbf{s}_n$ describes an internal state behind the fluorescence data at the time t_n. In the present study, the number of the internal states is 2 ($K = 2$), and thus, the $\mathbf{s}_n$ is described with the binary as

$$\mathbf{s}_n = \begin{cases} \begin{pmatrix} 1 \\ 0 \end{pmatrix} & \text{(ON state)}, \\ \begin{pmatrix} 0 \\ 1 \end{pmatrix} & \text{(OFF state)}. \end{cases} \tag{5}$$

Note that this variable satisfies a sum rule as

$$\sum_{k=1}^{K} s_{kn} = 1. \tag{6}$$

$\boldsymbol{\Theta}$ in Eq. (1) expresses variables to characterize a probability distribution function of the time series $\boldsymbol{I}$ in Eq. (2). Since the total number of the internal states is K, we consider K distribution functions and prepare parameters for each distribution function. In the present study, we assume that the distribution function of each state follows the Gaussian distribution, which is based on the observation that experimental fluorescence intensities exhibiting blinking often follow the Gaussian distribution [38]. This is also true in our present experimental results (see Sec. III A). This assumption on the distribution function of the fluorescent intensity might not be true in the case where the photon antibunching tendency is strong. In that case, however, one first examines what kind of distribution of the fluorescence intensities there is, and changes the distribution form from the Gaussian type to the other one.

In the Gaussian function, it is characterized by mean μ and precision λ, where λ^{-1} represents a variance. Thus, we rewrite $\boldsymbol{\Theta}$ as $\{\boldsymbol{\mu}, \boldsymbol{\lambda}\}$, where $\boldsymbol{\mu}$ contains the mean of each state as

$$\boldsymbol{\mu} = (\mu_1, \mu_2, \ldots, \mu_K)^T, \tag{7}$$

and $\boldsymbol{\lambda}$ contains the precision of each state as

$$\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_K)^T. \tag{8}$$

Here, μ_k and λ_k are the mean and precision of the kth Gaussian distribution.

$\boldsymbol{\pi}$ in Eq. (1) describes parameters to set an initial-step hidden variable $\mathbf{s}_1$, and this is also a K-component vector as

$$\boldsymbol{\pi} = (\pi_1, \pi_2, \ldots, \pi_K)^T \tag{9}$$

with π_k being the probability of taking the kth state at the initial step. Note a normalization of $\sum_{k=1}^{K} \pi_k = 1$.

$\mathbf{A}$ in Eq. (1) is variables for describing the transition between successive hidden variables, which are expressed as

$$\mathbf{A} = (\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_K) \tag{10}$$

with

$$\mathbf{A}_{k'} = (A_{1k'}, A_{2k'}, \ldots, A_{Kk'})^T \tag{11}$$

representing a transition probability from the k'th state to other states. Thus, $\mathbf{A}$ is a $K \times K$ matrix, and the matrix element $A_{kk'}$ expresses a transition probability from the k'th state to the kth state. Note the normalization of $\sum_{k=1}^{K} A_{kk'} = 1$.

We now rewrite the right-hand side of Eq. (1) by considering independence of the stochastic variables as [34,35]

$$\begin{aligned} &p(\boldsymbol{I}|\mathbf{S}, \boldsymbol{\mu}, \boldsymbol{\lambda}) p(\mathbf{S}|\boldsymbol{\pi}, \mathbf{A}) p(\boldsymbol{\mu}, \boldsymbol{\lambda}) p(\boldsymbol{\pi}) p(\mathbf{A}) \\ &= \left(\prod_{n=1}^{N} \prod_{k=1}^{K} p(I_n|\mu_k, \lambda_k)^{s_{kn}} \right) \\ &\quad \times \left(p(\mathbf{s}_1|\boldsymbol{\pi}) \prod_{n=2}^{N} \prod_{k'=1}^{K} p(\mathbf{s}_n|\mathbf{A}_{k'})^{s_{k'n-1}} \right) \\ &\quad \times \left(\prod_{k=1}^{K} p(\mu_k, \lambda_k) \right) p(\boldsymbol{\pi}) \left(\prod_{k'=1}^{K} p(\mathbf{A}_{k'}) \right). \end{aligned} \tag{12}$$

Here, we note that the forms of $\prod_{k=1}^{K} p(I_n|\mu_k, \lambda_k)^{s_{kn}}$ and $\prod_{k'=1}^{K} p(\mathbf{s}_n|\mathbf{A}_{k'})^{s_{k'n-1}}$ in Eq. (12) select only one of the K states due to the binary feature of the hidden variable $\mathbf{s}_n$ [Eqs. (5) and (6)]. The variable dependency in this model is represented as a graph in Fig. 3. Arrows represent a dependency between variables; for example, in this model, I_n depends on $\mathbf{s}_n$, $\boldsymbol{\mu}$, and $\boldsymbol{\lambda}$, which corresponds to $\prod_{k=1}^{K} p(I_n|\mu_k, \lambda_k)^{s_{kn}}$ of the right-hand side in Eq. (12).

Concrete forms of each probability distribution in Eq. (12) are summarized as follows:

(1) $p(I_n|\mu_k, \lambda_k)$ is the Gaussian distribution as

$$\begin{aligned} p(I_n|\mu_k, \lambda_k) &= \mathcal{N}(I_n|\mu_k, \lambda_k^{-1}) \\ &= \sqrt{\frac{\lambda_k}{2\pi}} \exp\left(-\frac{\lambda_k}{2} (I_n - \mu_k)^2 \right). \end{aligned} \tag{13}$$

(2) $p(\mathbf{s}_1|\boldsymbol{\pi})$ and $p(\mathbf{s}_n|\mathbf{A}_{k'})$ are the categorical distribution as

$$p(\mathbf{s}_1|\boldsymbol{\pi}) = \mathrm{Cat}(\mathbf{s}_1|\boldsymbol{\pi}) = \prod_{k=1}^{K} \pi_k^{s_{k1}} \tag{14}$$

and

$$p(\mathbf{s}_n|\mathbf{A}_{k'}) = \mathrm{Cat}(\mathbf{s}_n|\mathbf{A}_{k'}) = \prod_{k=1}^{K} A_{kk'}^{s_{kn}}, \tag{15}$$

respectively.

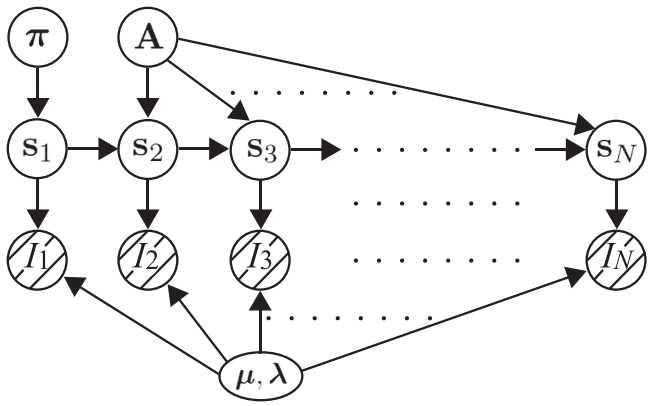

FIG. 3. A graphical model for HMM defined as Eq. (12). $\{I_n\}$ are observed time series [Eq. (2)] and the observed data are drawn with a shadow. $\{\mathbf{s}_n\}$ is a time series of hidden variables [Eq. (3)]. $\boldsymbol{\mu}$ and $\boldsymbol{\lambda}$ in Eqs. (7) and (8) are parameters characterizing distribution functions that time series $\boldsymbol{I}$ follows. An initial-step hidden variable $\mathbf{s}_1$ is generated based on the parameter $\boldsymbol{\pi}$ in Eq. (9), while $\mathbf{s}_n$ is generated by the previous step $\mathbf{s}_{n-1}$ and the transition matrix $\mathbf{A}$ in Eq. (10). Stochastic variables are represented with nodes and the dependency between the variables is represented by arrows.

(3) $p(\mu_k, \lambda_k)$ is the Gaussian-Gamma distribution as

$$\begin{aligned} p(\mu_k, \lambda_k) &= \mathrm{NG}(\mu_k, \lambda_k | m, \nu, a, b) \\ &= \mathcal{N}(\mu_k | m, (\nu\lambda_k)^{-1}) \mathrm{Gam}(\lambda_k | a, b) \\ &= \left\{ \sqrt{\frac{\nu\lambda_k}{2\pi}} \exp\left(-\frac{\nu\lambda_k}{2}(\mu_k - m)^2 \right) \right\} \\ &\quad \times \left\{ C_G(a,b) \lambda_k^{a-1} e^{-b\lambda_k} \right\}, \end{aligned} \tag{16}$$

where m, a, b, and ν are hyperparameters in the Gaussian-Gamma distribution, and $C_G(a,b)$ is the normalization constant as

$$C_G(a,b) = \frac{b^a}{\Gamma(a)} \tag{17}$$

with $\Gamma(a)$ being the Gamma function.

(4) $p(\boldsymbol{\pi})$ and $p(\mathbf{A}_{k'})$ are the Dirichlet distribution as

$$p(\boldsymbol{\pi}) = \mathrm{Dir}(\boldsymbol{\pi}|\boldsymbol{\alpha}) = C_D(\boldsymbol{\alpha}) \prod_{k=1}^{K} \pi_k^{\alpha_k - 1} \tag{18}$$

and

$$p(\mathbf{A}_{k'}) = \mathrm{Dir}(\mathbf{A}_{k'}|\boldsymbol{\beta}_{k'}) = C_D(\boldsymbol{\beta}_{k'}) \prod_{k=1}^{K} A_{kk'}^{\beta_{kk'}-1}, \tag{19}$$

respectively, and $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are hyperparameters in the Dirichlet distribution. Also, $C_D(\mathbf{x})$ is the normalization constant written as

$$C_D(\mathbf{x}) = \frac{\Gamma\left(\sum_{k=1}^{K} x_k\right)}{\prod_{k=1}^{K} \Gamma(x_k)}. \tag{20}$$

In the practical simulation, we need to calculate the conditional distribution function under the observed time-series data $\boldsymbol{I}$ in Eq. (2) as

$$p(\mathbf{S}, \boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\pi}, \mathbf{A}|\boldsymbol{I}) = \frac{p(\boldsymbol{I}, \mathbf{S}, \boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\pi}, \mathbf{A})}{p(\boldsymbol{I})}, \tag{21}$$

where $p(\boldsymbol{I})$ is a marginal distribution written as

$$p(\boldsymbol{I}) = \sum_{\mathbf{S}} \iiiint p(\boldsymbol{I}, \mathbf{S}, \boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\pi}, \mathbf{A}) d\boldsymbol{\mu} d\boldsymbol{\lambda} d\boldsymbol{\pi} d\mathbf{A}. \tag{22}$$

In the present study, we used the blocking Gibbs sampling [39] based on the Bayesian inference to optimize the conditional distribution function in Eq. (21). Details can be found in the Appendix. Anyway, by evaluating the conditional distribution function, the time-series data of the hidden variable $\mathbf{S}$ in Eq. (3) can be obtained, which is used for the following analysis of the ON/OFF duration.

C. Duration estimate based on HMM

We next describe an evaluation of the duration of the ON or OFF state of QDs. We first consider a sampling average for the time-series data of the hidden variable as

$$\bar{\mathbf{S}} = \frac{1}{N_{itr}} \sum_{i=1}^{N_{itr}} \mathbf{S}^{(i)}, \tag{23}$$

where $\mathbf{S}^{(i)}$ is the time-series data of the hidden variable in the ith Gibbs sampling step, and N_{itr} is the total number of the Gibbs sampling. The $\bar{\mathbf{S}}$ is given in the same form as in Eq. (3):

$$\bar{\mathbf{S}} = (\bar{\mathbf{s}}_1, \bar{\mathbf{s}}_2, \ldots, \bar{\mathbf{s}}_N). \tag{24}$$

In the present study, the $\mathbf{s}_n$ has two components: the ON and OFF states. Therefore, the $\bar{\mathbf{s}}_n$ is written as

$$\bar{\mathbf{s}}_n = \begin{pmatrix} s_n^{\mathrm{ON}} \\ s_n^{\mathrm{OFF}} \end{pmatrix} = \begin{pmatrix} s_n^{\mathrm{ON}} \\ 1 - s_n^{\mathrm{ON}} \end{pmatrix} \tag{25}$$

with $s_n^{\mathrm{ON}} = (1/N_{itr}) \sum_i^{N_{itr}} s_{1n}^{(i)}$ and $s_n^{\mathrm{OFF}} = (1/N_{itr}) \sum_i^{N_{itr}} s_{2n}^{(i)}$. Also, the sum rule in Eq. (6) was used in the transformation from the middle to right-side equations. We note that the component s_n^{ON} is a real number, not a binary, after the ensemble average. Now, using s_n^{ON}, we define a time series $\mathbf{s}^{\mathrm{ON}}$ as

$$\mathbf{s}^{\mathrm{ON}} = \left(s_1^{\mathrm{ON}}, s_2^{\mathrm{ON}}, \ldots, s_N^{\mathrm{ON}}\right)^T. \tag{26}$$

Figure 4 shows a schematic diagram of $\mathbf{s}^{\mathrm{ON}}$ denoted by a thick green solid curve. Generally, in the $\mathbf{s}^{\mathrm{ON}}$, statistical noise is suppressed and smooth behavior is obtained (see Sec. III). We thus estimate the ON duration τ^{ON} and the OFF duration τ^{OFF} from the $\mathbf{s}^{\mathrm{ON}}$ data, where a threshold s_{th} denoted by the thin red solid line is introduced. Since the $\mathbf{s}^{\mathrm{ON}}$ takes a value from 0 to 1, $s_{th} = 0.5$ is adopted. It is defined as entering the ON state when the $\mathbf{s}^{\mathrm{ON}}$ curve exceeds the s_{th}. We set this time to t_α^{ON} with a suffix α to specify the number of the ON event. On the other hand, when the $\mathbf{s}^{\mathrm{ON}}$ curve becomes smaller than the s_{th}, it is defined as entering the OFF state. We set this time to t_α^{OFF}. From these two times, the duration of the ON state is defined as

$$\tau_\alpha^{\mathrm{ON}} = t_\alpha^{\mathrm{OFF}} - t_\alpha^{\mathrm{ON}}. \tag{27}$$

Similarly, we define the duration of the OFF state as

$$\tau_\alpha^{\mathrm{OFF}} = t_{\alpha+1}^{\mathrm{ON}} - t_\alpha^{\mathrm{OFF}}. \tag{28}$$

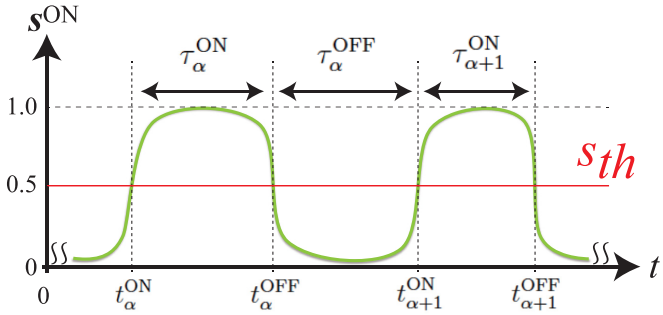

FIG. 4. A schematic diagram of a time series of $\mathbf{s}^{\mathrm{ON}}$ [Eq. (26)] which is described by a thick green solid curve. s_{th} is a threshold to distinguish the ON and OFF states, which is represented by a thin solid red line. Since the $\mathbf{s}^{\mathrm{ON}}$ takes a value from 0 to 1, $s_{th} = 0.5$ is adopted. It is defined as entering the ON state when the $\mathbf{s}^{\mathrm{ON}}$ curve exceeds the value of s_{th}. We set this time to t_α^{ON} with a suffix α specifying the number of the ON event. Similarly, when the $\mathbf{s}^{\mathrm{ON}}$ curve becomes smaller than s_{th}, it is defined as entering the OFF state. We set this time to t_α^{OFF}. From these two times, the ON and OFF duration, $\tau_\alpha^{\mathrm{ON}}$ and $\tau_\alpha^{\mathrm{OFF}}$, are estimated with using Eqs. (27) and (28), respectively.

With the obtained $\{\tau_\alpha^{\mathrm{ON}}\}$ and $\{\tau_\alpha^{\mathrm{OFF}}\}$ data, the probability distributions of the ON and OFF durations are calculated by

$$P_{\mathrm{ON}}(\tau) = \frac{1}{N_{\mathrm{ON}}} \sum_{\alpha'=1}^{N'_{\mathrm{ON}}} f_{\alpha'}^{\mathrm{ON}} w_{\alpha'}^{\mathrm{ON}} \delta\left(\tau - \tau_{\alpha'}^{\mathrm{ON}}\right) \tag{29}$$

and

$$P_{\mathrm{OFF}}(\tau) = \frac{1}{N_{\mathrm{OFF}}} \sum_{\alpha'=1}^{N'_{\mathrm{OFF}}} f_{\alpha'}^{\mathrm{OFF}} w_{\alpha'}^{\mathrm{OFF}} \delta\left(\tau - \tau_{\alpha'}^{\mathrm{OFF}}\right), \tag{30}$$

respectively, where N'_{ON} and N'_{OFF} are the total numbers of independent ON and OFF events, respectively. We note that, in this calculation, the two durations are considered equal within the grid spacing $\Delta\tau$. $f_{\alpha'}^{\mathrm{ON}}$ and $f_{\alpha'}^{\mathrm{OFF}}$ are the numbers of events with the same ON and OFF durations, respectively. Also, $w_{\alpha'}^{\mathrm{ON}}$ and $w_{\alpha'}^{\mathrm{OFF}}$ are weights given as [33]

$$w_{\alpha'}^{\mathrm{ON}} = \frac{2\Delta\tau}{\tau_{\alpha'+1}^{\mathrm{ON}} - \tau_{\alpha'-1}^{\mathrm{ON}}} \tag{31}$$

and

$$w_{\alpha'}^{\mathrm{OFF}} = \frac{2\Delta\tau}{\tau_{\alpha'+1}^{\mathrm{OFF}} - \tau_{\alpha'-1}^{\mathrm{OFF}}}, \tag{32}$$

respectively. Lastly, the denominators N_{ON} and N_{OFF} in Eqs. (29) and (30) are the total numbers of the ON and OFF events, which satisfy the following sum rules:

$$N_{\mathrm{ON}} = \sum_{\alpha'=1}^{N'_{\mathrm{ON}}} f_{\alpha'}^{\mathrm{ON}} \frac{w_{\alpha'}^{\mathrm{ON}}}{S^{\mathrm{ON}}} \tag{33}$$

and

$$N_{\mathrm{OFF}} = \sum_{\alpha'=1}^{N'_{\mathrm{OFF}}} f_{\alpha'}^{\mathrm{OFF}} \frac{w_{\alpha'}^{\mathrm{OFF}}}{S^{\mathrm{OFF}}}, \tag{34}$$

TABLE I. Initial hyperparameter setting for the present HMM simulations. K is the number of the hidden states, and ν, m, a, and b are the hyperparameters for the Gaussian-Gamma distribution of $p(\mu_k, \lambda_k)$ in Eq. (16). $\boldsymbol{\alpha}$ is the hyperparameter for the distribution function $p(\boldsymbol{\pi})$ in Eq. (18), which is a K-component vector. Also, $\boldsymbol{\beta}$ is the hyperparameter for the distribution function $p(\mathbf{A}_{k'})$ in Eq. (19), which is represented by a $K \times K$ matrix. The matrix elements are characterized by N, D, and γ, where N is the total length of the time series, and D and γ were set to 1 and 0.01 in the present simulation, respectively. N_{itr} is the total number of the iteration steps of the Gibbs sampling, which depends on the simulations.

K in Eq. (4)	2
(ν, m, a, b) in Eq. (16)	$(1, 0, 1, 1)$
$\boldsymbol{\alpha}$ in Eq. (18)	$(100, 100)^T$
$\boldsymbol{\beta}$ in Eq. (19)	$\begin{pmatrix} \frac{(1-2\gamma)N+4(1-\gamma)D}{4\gamma} & D \\ D & \frac{(1-2\gamma)N+4(1-\gamma)D}{4\gamma} \end{pmatrix}$
N_{itr} in Eq. (23)	1000 [Expt. (Figs. 6, 7)] 200 [TS-I and TS-II (Figs. 9(a), 9(b))] 1000 [TS-III (Fig. 9(c))]

respectively, where we introduced correction factors S^{ON} and S^{OFF} which are determined from the normalization conditions of $P_{\mathrm{ON}}(\tau)$ and $P_{\mathrm{OFF}}(\tau)$.

III. RESULTS AND DISCUSSION

We implemented the method described in Sec. II into the program code written by PYTHON. With this program, we performed HMM simulations for experimental and theoretical benchmark data. Table I summarizes the present setting of our HMM simulation and the employed hyperparameters. In the simulations, the observed fluorescence data $\boldsymbol{I}$ in Eq. (2) are standardized as

$$\tilde{I}_n = \frac{I_n - \bar{I}}{\sigma_I} \tag{35}$$

with a mean of

$$\bar{I} = \frac{1}{N} \sum_{n=1}^{N} I_n \tag{36}$$

and a standard deviation of

$$\sigma_I = \sqrt{\frac{1}{N} \sum_{n=1}^{N} (I_n - \bar{I})^2}. \tag{37}$$

The standardized time series $\tilde{\boldsymbol{I}}$ has a mean value of 0 and a standard deviation of 1, which is used as input as the HMM simulations.

A. HMM analysis for experimental data

1. *Experimental fluorescence data*

Here, we apply the above method to analysis of the experimental data. Before presenting the analysis results, we first describe experimental details. The QDs used in the experiment are CdSe/ZnS core-shell QDs capped with

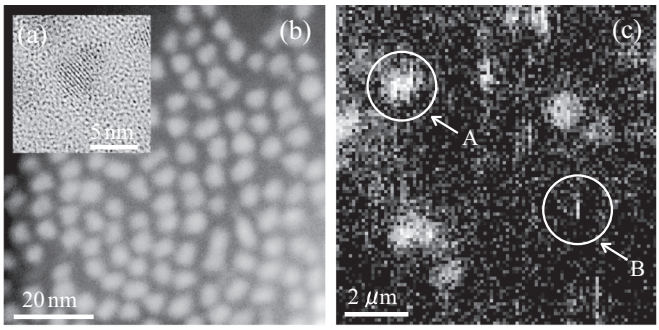

FIG. 5. (a) Transmission electron microscope (TEM) and (b) scanning TEM images. These images were taken with a FEI Tecnai 20 TEM. (c) Scanning fluorescence image of the QDs. The circled areas A and B in the panel show examples of ON (bright) and OFF (dark) states in the QDs, respectively.

trioctylphosphine oxide and hexadecyl amine, synthesized via the pyrolytic decomposition of organometallic compounds [40]. The QDs have a spheroid shape (minor axis: 4.1 ± 1.2 nm; major axis: 5.3 ± 1.3 nm), as confirmed in the transmission electron microscope (TEM) [Fig. 5(a)] and the scanning TEM [Fig. 5(b)] images. The thickness of the ZnS shell is 1.7 monolayers. The QDs were spin-cast from a toluene solution at 3000 rpm onto silica glass plates with a hydrophobic surface. The spin-cast samples were set in a sealed chamber with optical windows.

We measured scanning fluorescence images and time series of a fluorescent intensity of single QDs using a scanning laser microscope [41] equipped with an objective lens (100×; 0.9 NA) and avalanche photodetector. The laser-light wavelength and excitation intensity were 488 nm and 133 W/cm^2, respectively. The fluorescence images were acquired by raster-scanning a laser spot on the sample surface. The laser spot size and scanning step length were the diffraction-limited size (less than 1 μm) and 0.1 μm, respectively. An exposure time for each step was 0.1 s. The time series data of the single QDs were recorded at fixed positions on the sample surface with an exposure time of 0.1 s under continuous laser illumination. We confirmed that the noise due to the dead time of the detector, i.e., the blind time after each photon detection, on the experimental time series to be analyzed is sufficiently small [42]. The fluorescence images and the time series were measured for two different environments: vacuum and wet-nitrogen atmospheres. The wet-nitrogen atmosphere was prepared by bubbling nitrogen gas through distilled water. We took 12 samples with a length of 1200 s for each of the vacuum and wet-nitrogen atmospheres.

Figure 5(c) shows a typical fluorescence image of the spin-cast sample in vacuum. There are several circular spots like A with diffraction-limited size and several line segments like B along the scanning direction (vertical direction). All the circular spots and the line segments correspond to the fluorescence images of individual single QDs; the circular spots indicate that the QDs were in the ON state during the raster scanning, while the line segments indicate that the QDs were mainly in the OFF state and temporally in the ON state. All time series data of single QDs shown below were recorded at the center positions of the bright spots in scanning fluorescence images.

TABLE II. A summary of the fitting of $f(\tau)$ [Eq. (38)] to the data of Fig. 8, where we list the inverse exponent m and its standard errors.

	Vacuum		Wet	
	ON	OFF	ON	OFF
Conventional	1.99 ± 0.05	2.14 ± 0.04	2.09 ± 0.06	1.83 ± 0.04
HMM	1.10 ± 0.05	1.20 ± 0.04	0.88 ± 0.06	1.08 ± 0.05

Figures 6(a)–6(c) and 7(a)–7(c) show typical time series of the experimental fluorescent intensity in vacuum and wet-nitrogen atmospheres, respectively, which are described by red solid lines. The single QDs in both environments exhibited blinking phenomena [32], i.e., an irregular change between the ON and OFF states. The blinking occurrence rate appeared to be reduced in the wet-nitrogen atmosphere, as demonstrated qualitatively in Ref. [41]. It is known that the OFF state of blinking corresponds to the trapping of photogenerated electrons (or holes) in trap site(s) located on or near QD surfaces [32,33]. Thus, we can qualitatively infer that suppression in the wet-nitrogen atmosphere is due to inactivation of the trap site(s) by photoadsorption of water molecules in the environment [41]. However, it is difficult to quantify blinking properties with commonly used conventional analysis, assuming an artificial threshold between the ON and OFF states (Sec. II A), due to the low signal-to-noise ratio of the data in Figs. 6 and 7. As such, we can apply an HMM analysis described in Secs. II B and II C to quantify blinking within the noisy data.

2. HMM analysis

The green dashed lines in Figs. 6 and 7 show time series of the hidden variable $\mathbf{s}^{\mathrm{ON}}$ in Eq. (26), which are obtained from the HMM simulations for the time series of the experimental fluorescent intensity $\boldsymbol{I}$ in Eq. (2). We see from the figure that the behavior of the $\mathbf{s}^{\mathrm{ON}}$ is in good agreement with the variation in the experimental intensities (red solid lines). Also, the $\mathbf{s}^{\mathrm{ON}}$ basically takes 1 or 0 discontinuously and is hardly affected by noise. For such data, the duration can be estimated stably by the method described in Sec. II C. The blue dotted line I_{th} is a threshold to distinguish the ON and OFF states for the $\boldsymbol{I}$ data, while s_{th} denoted by the black solid line is a threshold for the $\mathbf{s}^{\mathrm{ON}}$ data. When estimating the duration in the conventional way with using I_{th} (Sec. II A), it becomes erroneous. For example, in the case of Fig. 6(c), the difference in the experimental intensity of the ON and OFF states is clearly smaller than the noise. In this case, it would be difficult to distinguish the two states from fluorescent time series with I_{th}. If the way is forcibly performed for estimating the duration, a large amount of short duration due to noise will occur. In contrast, in the analysis for $\mathbf{s}^{\mathrm{ON}}$ with using s_{th}, since the noise is suppressed in $\mathbf{s}^{\mathrm{ON}}$, the ON/OFF assignment can be stably performed.

Figure 8 is our calculated probability distributions for the ON/OFF duration, $P_{\mathrm{ON}}(\tau)$ in Eq. (29) and $P_{\mathrm{OFF}}(\tau)$ in Eq. (30),

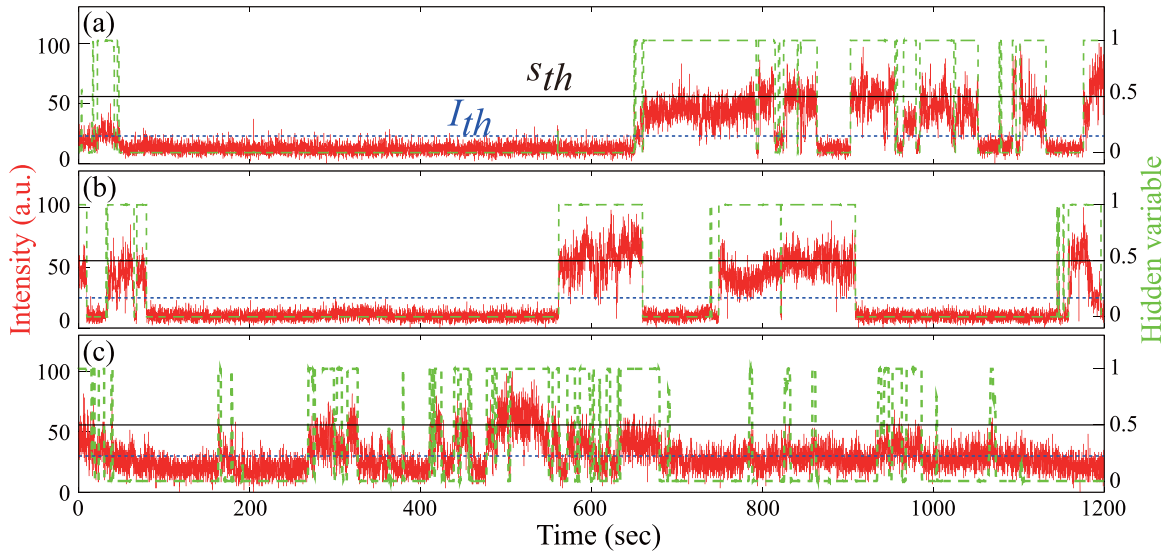

FIG. 6. Three typical time series of the fluorescent intensity measured in a vacuum atmosphere, denoted by red solid lines. These are used in the HMM simulations as the observed time series $\boldsymbol{I}$ in Eq. (2). The grid spacing of the time series is 0.1 s. The green dashed lines describe the time series of the hidden variable $\mathbf{s}^{\mathrm{ON}}$ in Eq. (26) obtained from the HMM simulations. The left scale is for the fluorescent intensity and the right scale is for the hidden variable. The I_{th} denoted by blue dotted lines is a threshold to distinguish the ON (bright) and OFF (dark) states for the $\boldsymbol{I}$ time series, and is set by hand artificially. The s_{th} denoted by black solid line is a threshold for the $\mathbf{s}^{\mathrm{ON}}$ time series, and is set to 0.5.

denoted by red dots. Left-side panels (a), (c), (e), and (g) are the results of the ON duration, and the right-side panels (b), (d), (f), and (h) are the results of the OFF duration. Also, the upper four panels [(a), (b), (c), and (d)] show the results for the vacuum condition, and the lower ones [(e), (f), (g), and (h)] show the results under the wet conditions. The (a), (b), (e), and (f) panels contain the results based on the conventional method, while the (c), (d), (g), and (h) ones represent the results based on the HMM. To analyze trends of each data, we performed a fitting of the following function [43] to the data as

$$f(\tau) = A\tau^{-m}, \tag{38}$$

where A is a coefficient and m is an inverse exponent. The fitted function is described by a black solid line and the obtained m values are summarized in Table II. A small m indicates that long-duration states tend to be formed; for example, the plot of panel (g) exhibits the smallest $m = 0.884$ and therefore

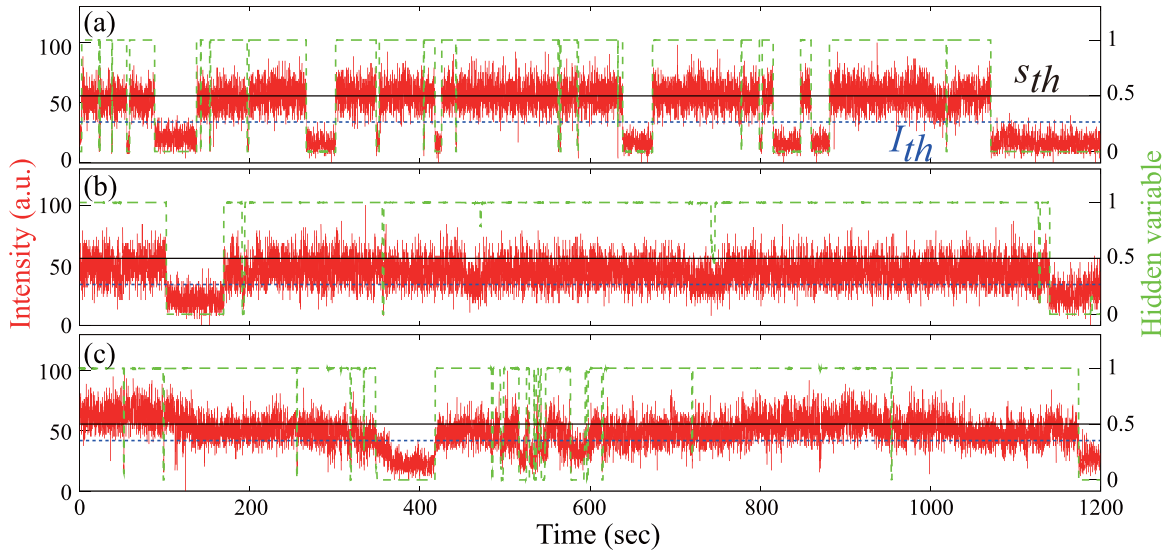

FIG. 7. Three typical time series of the fluorescent intensity measured in a wet-nitrogen atmosphere, denoted by red solid lines. These are used in the HMM simulations as the observed time series $\boldsymbol{I}$ in Eq. (2). The grid spacing of the time series is 0.1 s. The green dashed lines describe the time series of the hidden variable $\mathbf{s}^{\mathrm{ON}}$ in Eq. (26) obtained from the HMM simulations. The left scale is for the fluorescent intensity and the right scale is for the hidden variable. The I_{th} denoted by blue dotted lines is a threshold to distinguish the ON (bright) and OFF (dark) states for the $\boldsymbol{I}$ time series, and is set by hand artificially. The s_{th} denoted by black solid line is a threshold for the $\mathbf{s}^{\mathrm{ON}}$ time series, and is set to 0.5.

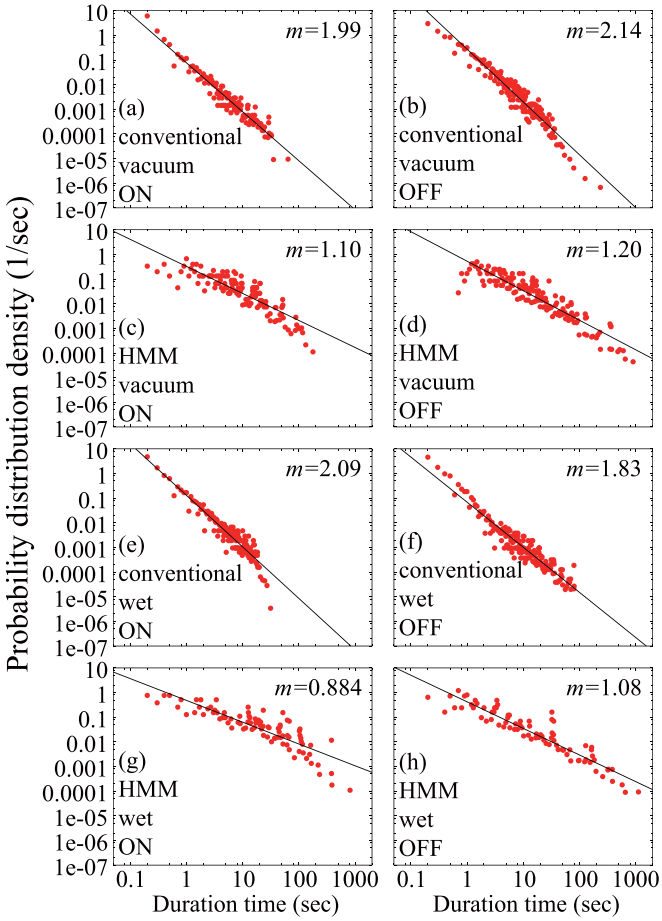

FIG. 8. Our calculated probability distributions for the ON/OFF duration in the experimental fluorescent time series, $P_{\rm ON}(\tau)$ in Eq. (29) and $P_{\rm OFF}(\tau)$ in Eq. (30), denoted by red dots. Black solid lines are $f(\tau)$ in Eq. (38) and the inverse exponent m represents a slope of the line, whose result is summarized in Table II. Also, the small slope corresponds to the high frequency of long duration. Left-side panels (a), (c), (e), and (g) are the results of the ON duration, and the right-side panels (b), (d), (f), and (h) are the results of the OFF duration. Also, the upper four panels [(a), (b), (c), and (d)] show the results for the vacuum condition, and the lower ones [(e), (f), (g), and (h)] show the results under the wet conditions. Panels (a), (b), (e), and (f) contain the results based on the conventional method, while (c), (d), (g), and (h) represent the results based on the HMM.

has many long-duration data compared with the other plots. In contrast, in the largest m plot of panel (b), the data concentrate in the short-duration regime. By comparison of m of each plot, we found that the probability distributions based on the conventional method are approximated by τ^{-2}, while the probability distributions based on the HMM decay as τ^{-1}. Thus, the plots based on the conventional method tend clearly to reflect many artificial short-duration states, and the HMM corrects the long-duration data. We note that, on the results based on the HMM, the m values with the wet condition are basically small compared to the m with the vacuum condition, indicating that single QDs emit long and quench long in the wet atmospheres, while in the vacuum atmosphere, the single QDs blink with moderate length.

We also note that, in the HMM analysis, the evaluation of long-duration events is rather accurate, while the evaluation for short-duration events is partially limited. For example, in the case of the present experimental data, the time-series data are taken at an interval Δt of 0.1 s, and in this case, the reliable lower bound of the evaluated duration would be about 0.5 s or more (i.e., $\sim 5 \times \Delta t$). We will discuss this point in Sec. III B in more detail.

B. HMM analysis for theoretical benchmark data

In this part, we check the quantitative accuracy of the present HMM simulation. In Sec. III A, we have shown that there is a discernible difference between the results based on the conventional method and the HMM simulation, but it does not mean the quantitative accuracy for the HMM. In order to verify the accuracy of the HMM, it is necessary to perform HMM analyses for a time series with correct answers and check whether the HMM can reproduce the correct results.

1. Generation of model time series

For this purpose, we consider the following model distribution functions for the ON and OFF duration [38] as

$$p_{\rm ON}(\tau) = A^{\rm ON}\tau^{-q}e^{-\tau/\xi}\theta(\tau-\tau_{\rm min})\theta(\tau_{\rm max}-\tau) \quad (39)$$

and

$$p_{\rm OFF}(\tau) = A^{\rm OFF}\tau^{-l}\theta(\tau-\tau_{\rm min})\theta(\tau_{\rm max}-\tau), \quad (40)$$

respectively, where $A^{\rm ON}$ and $A^{\rm OFF}$ are normalization constants. $\theta(x)$ is a step function, and q, ξ, and l are parameters of the model functions. $\tau_{\rm min}$ and $\tau_{\rm max}$ are respectively lower and upper cutoffs of the duration considered in the present model. These functions have been widely used in analyses for the blinking phenomena with the power law of QDs [38,44]. The exponential term in Eq. (39) is valid in the case of the truncated power law. In the simulation, we first generate a time series with the model distribution functions. Next, we perform an HMM analysis for the generated time series, and evaluate the ON and OFF duration. Finally, we check whether the distribution function calculated with the simulation [Eqs. (29) and (30)] reproduces the original model distribution functions [Eqs. (39) and (40)].

The model time series is generated as follows:

(1) We first define baselines for the ON and OFF intensities ($I_{\rm ON}$ and $I_{\rm OFF}$) and their difference

$$\Delta = I_{\rm ON} - I_{\rm OFF}. \quad (41)$$

(2) We next sample $\{\tau_\alpha^{\rm ON}\}$ from $p_{\rm ON}(\tau)$ in Eq. (39) and $\{\tau_\alpha^{\rm OFF}\}$ from $p_{\rm OFF}(\tau)$ in Eq. (40) with α specifying a sampling number. The $\tau_\alpha^{\rm ON}$ and $\tau_\alpha^{\rm OFF}$ represent the duration of the $I_{\rm ON}$ and $I_{\rm OFF}$ intensities, respectively. With these data, we make a sequence $\boldsymbol{\tau}$ as

$$\boldsymbol{\tau} = \left(\tau_1^{\rm ON}, \tau_1^{\rm OFF}, \tau_2^{\rm ON}, \ldots, \tau_{N_e}^{\rm ON}, \tau_{N_e}^{\rm OFF}\right), \quad (42)$$

where N_e is the total number of the ON or OFF events.

(3) Then, based on the above $\boldsymbol{\tau}$, $I_{\rm ON}$, and $I_{\rm OFF}$, we construct a bare time series

$$\boldsymbol{I}_0 = \left(I_1^0, I_2^0, \ldots, I_N^0\right)^T \quad (43)$$

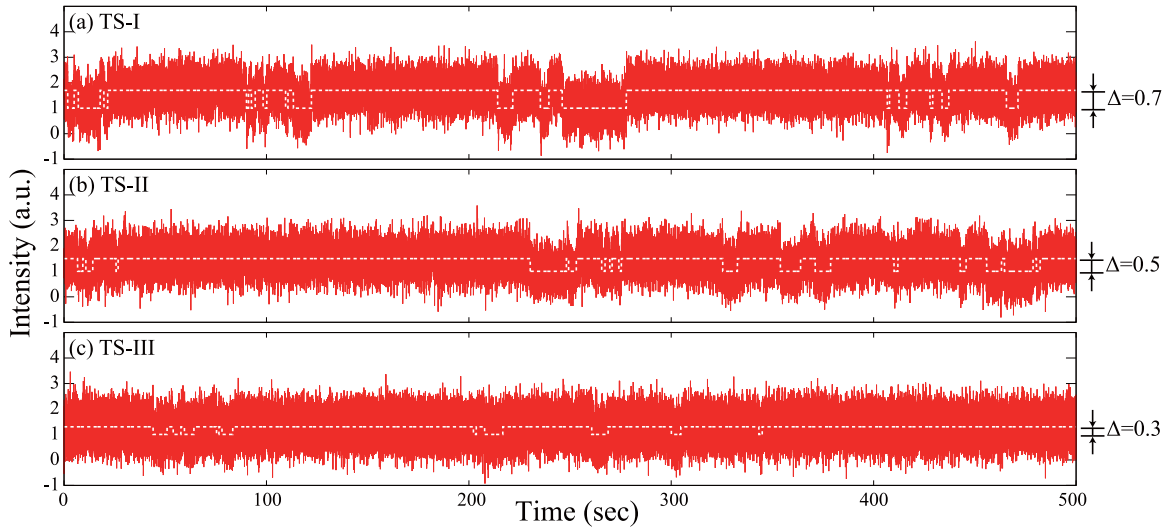

FIG. 9. Model time series $\boldsymbol{I}$ [Eq. (2)] generated in the manner of Sec. III B, denoted by red solid lines. White dashed lines are the bare time series $\boldsymbol{I}_0$ in Eq. (43). $\Delta = I_{\rm ON} - I_{\rm OFF}$ in Eq. (41) represents the bare intensity difference. (a) TS-I ($\Delta = 0.7$), (b) TS-II ($\Delta = 0.5$), and (c) TS-III ($\Delta = 0.3$). The noise η is 0.5. The time-grid interval Δt was set to 0.01 s for TS-I and TS-II and 0.005 s for TS-III and a single time series contains 20 ON and 20 OFF events ($N_e = 20$).

which are discretized at interval Δt and in a similar form to Eq. (2).

(4) Finally, we add an artificial noise η to the above time series $\boldsymbol{I}_0$, where η is sampled from the Gaussian functions as

$$\eta \sim \mathcal{N}(\eta|0,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{\eta^2}{2\sigma^2}\right) \tag{44}$$

with σ being a standard deviation of the Gaussian function. Thus, the final model time series $\boldsymbol{I}$ in Eq. (2) is obtained.

Figure 9 shows the generated time series $\boldsymbol{I}$ in Eq. (2) (red solid lines). The bare time series $\boldsymbol{I}_0$ in Eq. (43) is also shown with white dashed lines. Parameters in generating the time series and characterizing the model functions $p_{\rm ON}(\tau)$ in Eq. (39) and $p_{\rm OFF}(\tau)$ in Eq. (40) are summarized in Table III. In the present study, we consider three time series with different Δ (0.7, 0.5, and 0.3) in Eq. (41) with the noise amplitude η in Eq. (44) kept at 0.5. We call the time series with $\Delta = 0.7$, 0.5, and 0.3 TS-I, TS-II, and TS-III, respectively. The time-grid interval Δt of the time series was set to 0.01 s for TS-I and TS-II and 0.005 s for TS-III, and one time series contains 20 ON and 20 OFF events ($N_e = 20$). We generated 1000 samples for each time series and took an ensemble average of them in the analysis. Panels (a), (b), and (c) in Fig. 9 compare TS-I, TS-II, and TS-III. As the Δ becomes smaller ($\Delta = 0.7 \to 0.5 \to 0.3$), it becomes difficult to distinguish between the ON and OFF intensities. At $\Delta = 0.3$ (TS-III), it would be no longer possible for the human eye to distinguish between the ON and OFF states correctly.

We note that, to ensure numerical accuracy of the HMM simulation, a proper choice of a grid spacing Δt of the time series is important; Δt must ideally be small enough compared to a grid spacing $\Delta\tau$ of the duration grid which is introduced in the calculation of $P_{\rm ON}(\tau)$ in Eq. (29) and $P_{\rm OFF}$ in Eq. (30). Then,

$$\Delta t \ll \Delta\tau. \tag{45}$$

By definition, $\Delta\tau$ represents the minimum duration. In order to evaluate it accurately, the Δt must be small enough. In the present study, Δt and $\Delta\tau$ were respectively set to 0.01 s and 0.1 s for TS-I and TS-II and 0.005 s and 0.1 s for TS-III.

TABLE III. Parameters for generating model time series of Fig. 9, based on the model functions $p_{\rm ON}(\tau)$ in Eq. (39) and $p_{\rm OFF}(\tau)$ in Eq. (40). ξ, q, l, $\tau_{\rm max}$, and $\tau_{\rm min}$ are parameters characterizing $p_{\rm ON}(\tau)$ and $p_{\rm OFF}(\tau)$. $I_{\rm ON}$ and $I_{\rm OFF}$ are baselines for the ON and OFF intensities, respectively, and Δ is their difference [Eq. (41)]. According to the Δ, we named the time series with $\Delta = 0.7$, 0.5, and 0.3 TS-I, TS-II, and TS-III, respectively. σ is a standard deviation of the Gaussian function to generate noise η [Eq. (44)]. N_e is the total number of the ON or OFF events [Eq. (42)] in the time series. Note that the total number of the ON and OFF events in the time series is the same. Δt is a grid spacing of time series, and $\Delta\tau$ is a grid spacing of the duration grid, which is introduced in the calculation of $P_{\rm ON}(\tau)$ in Eq. (29) and $P_{\rm OFF}(\tau)$ in Eq. (30). The $\tau_{\rm max}$, $\tau_{\rm min}$, $\Delta\tau$, and Δt are given in units of s.

	Eq.	TS-I	TS-II	TS-III
ξ	(39)	1800	1800	1800
q	(39)	1.3	1.3	1.3
l	(40)	1.7	1.7	1.7
$\tau_{\rm max}$	(39), (40)	1000	1000	500
$\tau_{\rm min}$	(39), (40)	1	1	1
$\Delta\tau$	(45)	0.1	0.1	0.1
$I_{\rm ON}$	(41)	1.7	1.5	1.3
$I_{\rm OFF}$	(41)	1.0	1.0	1.0
Δ	(41)	0.7	0.5	0.3
σ	(44)	0.5	0.5	0.5
N_e	(42)	20	20	20
N_{itr}	(23)	200	200	1000
Δt	(45)	0.01	0.01	0.005
$N_{\rm sample}$		1000	1000	1000

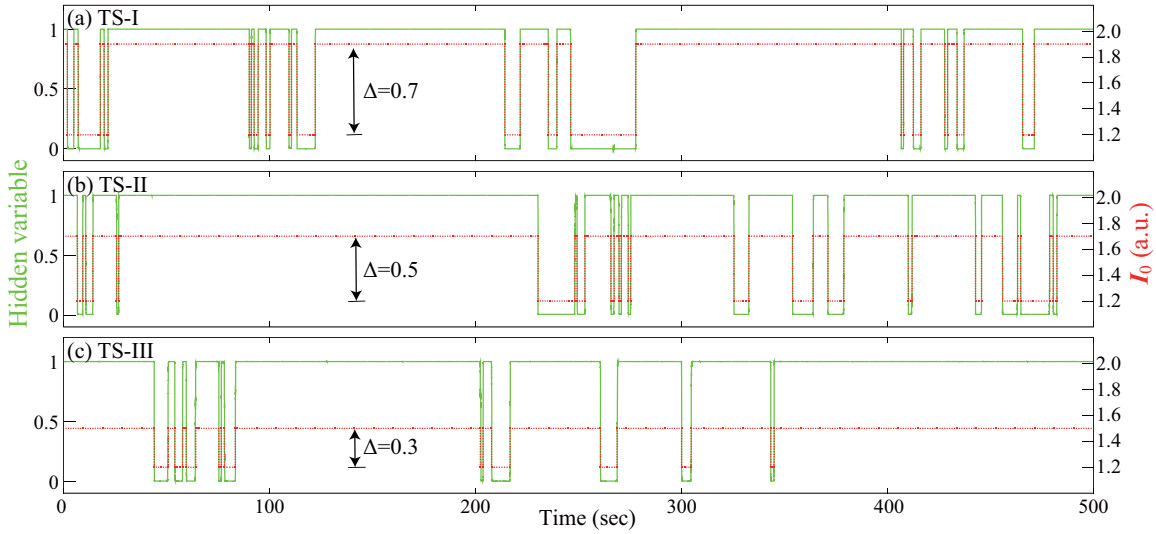

FIG. 10. Calculated time series of hidden variable $\mathbf{s}^{\mathrm{ON}}$ [Eq. (26)] with the HMM simulations, denoted by green solid lines. Panel (a) shows the $\mathbf{s}^{\mathrm{ON}}$ obtained for TS-I ($\Delta = 0.7$) in Fig. 9(a). Similarly, panels (b) and (c) are the results for TS-II ($\Delta = 0.5$) in Fig. 9(b) and TS-III ($\Delta = 0.3$) in Fig. 9(c), respectively. Bare time series $\boldsymbol{I}_0$ denoted by red dotted lines are also shown for comparison. If the HMM simulation is successful, the $\mathbf{s}^{\mathrm{ON}}$ should match $\boldsymbol{I}_0$.

We also comment on the computational time on the present HMM simulation. If the time-series length is long, we can understand the longer-duration behavior of the blinking phenomena rather accurately. On the other hand, in this case, the calculation time will become long. So, in an actual treatment, rather than analyzing a single long time series, we divide it and analyze multiple samples of medium-size time series. For the present case, to analyze the data of one time series, it takes about 7 days with a single-core machine. Also, we considered 1000 samples for the statistical average. Since we used 20 computer nodes and each node consists of 28 cores, the total computational time was $1000/(28\times20)\times 7 \sim 12.5$ days.

2. *The time series of hidden variables*

We show in Fig. 10 results of our HMM analysis for the time series in Fig. 9. Panels (a), (b), and (c) correspond to the results for TS-I [Fig. 9(a)], TS-II [Fig. 9(b)], and TS-III [Fig. 9(c)], respectively. The solid green line describes the time series of the hidden variables $\mathbf{s}^{\mathrm{ON}}$ in Eq. (26), which is compared with the bare time series $\boldsymbol{I}_0$ in Eq. (43) (red dotted line). We see that the behavior of $\mathbf{s}^{\mathrm{ON}}$ rather well reproduces the $\boldsymbol{I}_0$. There are almost no misjudgments even for the time series with $\Delta = 0.3$ (TS-III), and thus our HMM has very high accuracy.

3. *Probability density of duration time*

From the analysis to the $\mathbf{s}^{\mathrm{ON}}$ time series as shown in Fig. 10, we evaluated the durations $\{\tau_\alpha^{\mathrm{ON}}\}$ and $\{\tau_\alpha^{\mathrm{OFF}}\}$ using the method described in Sec. II C, and calculated their probability distributions $P_{\mathrm{ON}}(\tau)$ in Eq. (29) and $P_{\mathrm{OFF}}(\tau)$ in Eq. (30). The resulting blinking plots are shown in Fig. 11 by red dots. The left-side panels represent $P_{\mathrm{ON}}(\tau)$, and the right-side panels are $P_{\mathrm{OFF}}(\tau)$. Also, the upper (a) and (b) panels describe the results for the model time series TS-I [Fig. 9(a)]. Similarly, the middle (c) and (d) and lower (e) and (f) panels describe the results for the time series TS-II [Fig. 9(b)] and TS-III [Fig. 9(c)], respectively. We also give the original model distributions

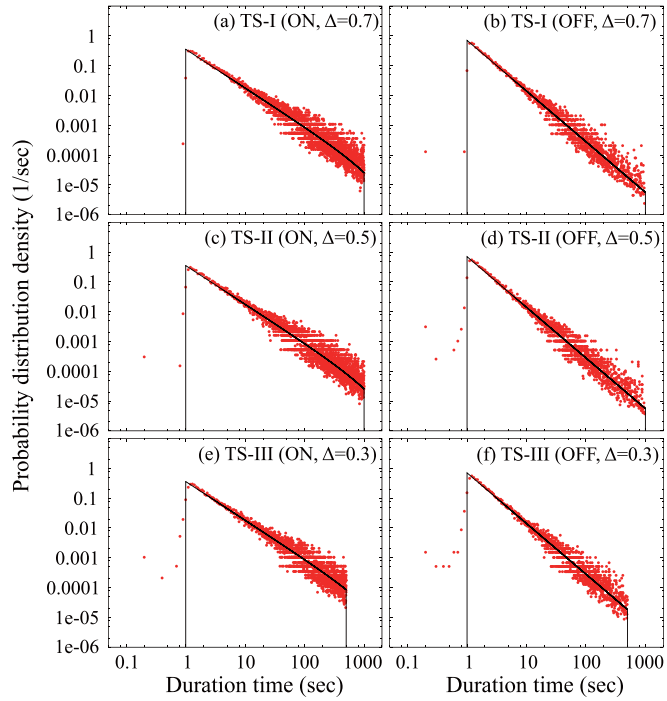

FIG. 11. Our calculated probability distributions for the ON/OFF duration in the fluorescent time series, $P_{\mathrm{ON}}(\tau)$ in Eq. (29) and $P_{\mathrm{OFF}}(\tau)$ in Eq. (30), denoted by red dots. Black solid lines represent model probability distributions $p_{\mathrm{ON}}(\tau)$ in Eq. (39) and $p_{\mathrm{OFF}}(\tau)$ in Eq. (40). The upper (a) and (b), middle (c) and (d), and lower (e) and (f) panels correspond to the results for the time series with $\Delta = 0.7$, 0.5, and 0.3, respectively. Note that each result is obtained for N_{sample} time series. Also, left (a), (c), and (e) panels are the results for the ON duration, while right (b), (d), and (f) panels are the results of the OFF state.

TABLE IV. Analysis of the HMM prediction accuracy. Upper, middle, and lower blocks describe the analysis results for the time series with $\Delta = 0.7$, 0.5, and 0.3, respectively. Note that each result is obtained for $N_{\rm sample}$ time series. In the table, we present the duration results digit by digit. $N_{\rm ans}$ is the total number of actual events in the ON or OFF states, which is evaluated from $\boldsymbol{I}_0$ in Eq. (43). $N_{\rm sim}$ is the total number of events in the ON or OFF states, which is estimated from the time series $\mathbf{s}^{\rm ON}$ in Eq. (26), based on the HMM simulation. $N_{\rm suc}$ is the total number of correctly predicted events. Recall and precision are $N_{\rm suc}/N_{\rm ans}$ in Eq. (46) and $N_{\rm suc}/N_{\rm sim}$ in Eq. (47), respectively. These results are for $N_{\rm sample}$ time series in Table III. The unit of τ is sec.

		ON					OFF				
	τ	$N_{\rm ans}$	$N_{\rm sim}$	$N_{\rm suc}$	Recall	Precision	$N_{\rm ans}$	$N_{\rm sim}$	$N_{\rm suc}$	Recall	Precision
	0–1	0	358	0	–	0	0	363	0	–	0
	1–10	11832	11804	11625	0.983	0.985	16058	16025	15682	0.977	0.979
TS-I	10–100	5728	5728	5728	1.0	1.0	3325	3325	3325	1.0	1.0
	100–1000	2440	2440	2440	1.0	1.0	617	617	617	1.0	1.0
	1000–	0	0	0	–	–	0	0	0	–	–
	0–1	0	1396	0	–	0	0	1459	0	–	0
	1–10	11702	11620	10747	0.918	0.925	16123	15978	14501	0.899	0.908
TS-II	10–100	5844	5844	5844	1.0	1.0	3247	3247	3247	1.0	1.0
	100–1000	2454	2454	2454	1.0	1.0	630	630	630	1.0	1.0
	1000–	0	0	0	–	–	0	0	0	–	–
	0–1	0	2031	0	–	0	0	2163	0	–	0
	1–10	12085	11926	10398	0.86	0.872	16236	15945	13458	0.829	0.844
TS-III	10–100	5933	5934	5931	0.999	0.999	3210	3212	3208	0.999	0.999
	100–500	1982	1982	1981	0.999	0.999	554	554	554	1.0	1.0
	500–	0	0	0	–	–	0	0	0	–	–

$p_{\rm ON}(\tau)$ in Eq. (39) and $p_{\rm OFF}(\tau)$ in Eq. (40) by black solid lines. We see that the calculated blinking plots well reproduce the original distribution functions. As emphasized in Fig. 9(c), it is difficult for humans to identify the ON or OFF state in the $\boldsymbol{I}$ time series TS-III ($\Delta = 0.3$, $\eta = 0.5$), but the present HMM simulation is able to discriminate between the ON and OFF states with a fairly high accuracy.

4. Quantitative accuracy of HMM analysis

To check the HMM-simulation accuracy quantitatively, we evaluated recall and precision, which are defined as

$$\text{Recall} = \frac{N_{\rm suc}}{N_{\rm ans}} \tag{46}$$

and

$$\text{Precision} = \frac{N_{\rm suc}}{N_{\rm sim}}, \tag{47}$$

respectively. Here, $N_{\rm ans}$, $N_{\rm sim}$, and $N_{\rm suc}$ are the number of actual events with a certain duration, the number of events with a certain duration obtained by the HMM simulation, and the number of correctly predicted events, respectively. In the present model time series, we introduce cutoffs $\tau_{\rm min}$ and $\tau_{\rm max}$ [Eqs. (39) and (40)], and then the actual events exist only in 1–1000 s. In this analysis, we tolerate a 10% error. For example, in the case of data with the duration of $\tau = 1$ s, if the simulation evaluates it within 0.9–1.1 s, it is regarded as a success event. Table IV shows analysis results for the HMM simulation, where we accumulated the duration data from the $N_{\rm sample}$ time series, where $N_{\rm sample}$ is given in Table III. In the table, we present the duration results digit by digit. From the table, we see that, for all the time series, accuracy of the recall and precision is satisfactory. Certainly the accuracy of the short duration is less good; especially, simulations slightly observe a duration of 0–1 s that does not actually exist. Overall, however, the accuracy is very high and satisfactory, and the ability of the reproduction of the ON/OFF states based on the present HMM is impressive.

IV. SUMMARY

We have presented an HMM analysis for experimental and theoretical fluorescent time series of QDs. In this simulation, we have calculated the time series of hidden variables to evaluate the ON/OFF duration of the fluorescence. With the resulting duration data, we have calculated blinking plots. The blinking plots are well-known fundamental data for understanding the optical property and performance of the QDs and microscopic modeling for the emission and quenching processes are done based on this plot. So, the quantitative reliability of the blinking plot would be important for real understanding of the QD physics.

Through the comparison between the results based on the conventional and HMM analyses, we have found discernible quantitative differences; in the conventional method, the ON or OFF state is directly evaluated for the noisy fluorescent time series, thus leading to a large amount of artificial short duration data. On the other hand, in the HMM analysis, the hidden-variable time series which is noise-suppressed is calculated for the ON/OFF assignment, so we can accumulate reliable duration data; the artificial short-duration data are suppressed and long-duration data are properly evaluated. It was found that these differences in the evaluation methods have a great influence on the analysis for the noisy time-series data; in the case of the experimental fluorescence data of the single colloidal QDs, it had a significant effect on the power

of the duration probability distribution. Also, in order to show the accuracy of the present HMM analysis, we have analyzed the theoretical benchmark time series generated from model distribution functions for the duration. The distribution functions obtained from the HMM simulations well reproduce the model distribution functions, and we have found that the ON/OFF assignments can be performed rather accurately even for the low signal-to-noise time series.

In the present simulation, we have focused on two-state analysis, but the HMM can also be applied to analysis of three or more states. If such an analysis is performed successfully, it will be possible to construct a more detailed model on fluorescence of QDs, especially on the relaxation from photoexcited to ground states [45], which is left as a future task.

ACKNOWLEDGMENTS

We thank Yoshihide Yoshimoto for helpful discussions. We acknowledge the financial support of JSPS KAKENHI Grants No. 16H06345, No. 16K05452, No. 17H03393, No. 17H03379, No. 19K03673, and No. 22H01183. A part of the computation was done at Supercomputer Center, Institute for Solid State Physics, University of Tokyo.

APPENDIX: BLOCKING GIBBS SAMPLING

In this Appendix, we describe how to calculate the conditional distribution function $p(\mathbf{S}, \boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\pi}, \mathbf{A}|\boldsymbol{I})$ in Eq. (21). For this purpose, we use the Gibbs sampling [35,46] which gives the approximate solution of $p(\mathbf{S}, \boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\pi}, \mathbf{A}|\boldsymbol{I})$. The calculation proceeds as follows:

(1) We first calculate a conditional probability distribution of $p(\mathbf{S}^{(i)}|\boldsymbol{\mu}^{(i-1)}, \boldsymbol{\lambda}^{(i-1)}, \boldsymbol{\pi}^{(i-1)}, \mathbf{A}^{(i-1)}, \boldsymbol{I})$ fixed by $\boldsymbol{\mu}^{(i-1)}$, $\boldsymbol{\lambda}^{(i-1)}$, $\boldsymbol{\pi}^{(i-1)}$, $\mathbf{A}^{(i-1)}$, and $\boldsymbol{I}$, and sample $\mathbf{S}^{(i)}$ from the resulting conditional probability distribution as

$$\mathbf{S}^{(i)} \sim p(\mathbf{S}^{(i)}|\boldsymbol{\mu}^{(i-1)}, \boldsymbol{\lambda}^{(i-1)}, \boldsymbol{\pi}^{(i-1)}, \mathbf{A}^{(i-1)}, \boldsymbol{I}), \tag{A1}$$

where the upper suffix i specifies the iteration step of the Gibbs sampling, and the variables of the initial step, $\boldsymbol{\mu}^{(0)}$, $\boldsymbol{\lambda}^{(0)}$, $\boldsymbol{\pi}^{(0)}$, $\mathbf{A}^{(0)}$, are evaluated with the hyperparameters in Table I.

(2) We next calculate a conditional probability distribution $p(\boldsymbol{\mu}^{(i)}, \boldsymbol{\lambda}^{(i)}|\mathbf{S}^{(i)}, \boldsymbol{\pi}^{(i-1)}, \mathbf{A}^{(i-1)}, \boldsymbol{I})$ with fixed $\mathbf{S}^{(i)}$, $\boldsymbol{\pi}^{(i-1)}$, $\mathbf{A}^{(i-1)}$, and $\boldsymbol{I}$, and sample $\boldsymbol{\mu}^{(i)}$, $\boldsymbol{\lambda}^{(i)}$ from the obtained conditional distribution function as

$$\boldsymbol{\mu}^{(i)}, \boldsymbol{\lambda}^{(i)} \sim p(\boldsymbol{\mu}^{(i)}, \boldsymbol{\lambda}^{(i)}|\mathbf{S}^{(i)}, \boldsymbol{\pi}^{(i-1)}, \mathbf{A}^{(i-1)}, \boldsymbol{I}). \tag{A2}$$

(3) The same treatment applies to $\boldsymbol{\pi}^{(i)}$; we evaluate a conditional distribution function $p(\boldsymbol{\pi}^{(i)}|\mathbf{S}^{(i)}, \boldsymbol{\mu}^{(i)}, \boldsymbol{\lambda}^{(i)}, \mathbf{A}^{(i-1)}, \boldsymbol{I})$ and sample $\boldsymbol{\pi}^{(i)}$ as follows:

$$\boldsymbol{\pi}^{(i)} \sim p(\boldsymbol{\pi}^{(i)}|\mathbf{S}^{(i)}, \boldsymbol{\mu}^{(i)}, \boldsymbol{\lambda}^{(i)}, \mathbf{A}^{(i-1)}, \boldsymbol{I}). \tag{A3}$$

(4) Finally, we sample $\mathbf{A}^{(i)}$ from a conditional distribution function with fixed $\mathbf{S}^{(i)}$, $\boldsymbol{\mu}^{(i)}$, $\boldsymbol{\lambda}^{(i)}$, $\boldsymbol{\pi}^{(i)}$, and $\boldsymbol{I}$ as follows:

$$\mathbf{A}^{(i)} \sim p(\mathbf{A}^{(i)}|\mathbf{S}^{(i)}, \boldsymbol{\mu}^{(i)}, \boldsymbol{\lambda}^{(i)}, \boldsymbol{\pi}^{(i)}, \boldsymbol{I}). \tag{A4}$$

One iteration loop of the Gibbs sampling consists of Eq. (A1) → Eq. (A2) → Eq. (A3) → Eq. (A4) → Eq. (A1). The $\mathbf{S}^{(i)}$, $\boldsymbol{\mu}^{(i)}$, $\boldsymbol{\lambda}^{(i)}$, $\boldsymbol{\pi}^{(i)}$, $\mathbf{A}^{(i)}$ are optimized by repetition of this iteration. In the present HMM, since each conditional probability distribution is analytically obtained (see below), the Gibbs sampling is especially effective.

We next describe the conditional distribution for Eqs. (A1), (A2), (A3), and (A4). The calculation is a little complicated, but the derivation itself is straightforward, so only the results are shown.

(1) $p(\mathbf{S}^{(i)}|\boldsymbol{\mu}^{(i-1)}, \boldsymbol{\lambda}^{(i-1)}, \boldsymbol{\pi}^{(i-1)}, \mathbf{A}^{(i-1)}, \boldsymbol{I})$ [Eq. (A1)]:

We used a blocking Gibbs sampling based on a forward-backward algorithm [34,35,47]. In this treatment, we utilize a marginalization as

$$p(\mathbf{s}_n|\boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\pi}, \mathbf{A}, \boldsymbol{I}) = \sum_{\mathbf{S}_{\backslash n}} p(\mathbf{S}|\boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\pi}, \mathbf{A}, \boldsymbol{I}), \tag{A5}$$

where we drop the upper suffix (i) or $(i-1)$ on the iteration step for simplicity. Also, $\mathbf{S}_{\backslash n}$ is a subset of $\mathbf{S}$ minus $\mathbf{s}_n$. $p(\mathbf{s}_n|\boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\pi}, \mathbf{A}, \boldsymbol{I})$ is described with the following category distribution:

$$p(\mathbf{s}_n|\boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\pi}, \mathbf{A}, \boldsymbol{I}) = \mathrm{Cat}(\mathbf{s}_n|\boldsymbol{\eta}_n) = \prod_{k=1}^{K} \eta_{kn}^{s_{kn}} \tag{A6}$$

with

$$\eta_{kn} = \frac{\tilde{\eta}_{kn}}{\sum_{k'=1}^{K} \tilde{\eta}_{k'n}} \tag{A7}$$

and

$$\tilde{\eta}_{kn} = f_{kn} b_{kn}. \tag{A8}$$

Here, f_{kn} and b_{kn} on the right-hand side are given as follows:

$$f_{kn} = \frac{\hat{f}_{kn}}{\sum_{k'=1}^{K} \hat{f}_{k'n}} \tag{A9}$$

with

$$\hat{f}_{kn} = \begin{cases} p(I_1|\mu_k, \lambda_k)\pi_k, & (n=1), \\ p(I_n|\mu_k, \lambda_k) \sum_{k'=1}^{K} A_{kk'} f_{k'n-1}, & (n \neq 1), \end{cases} \tag{A10}$$

and

$$b_{kn} = \frac{\hat{b}_{kn}}{\sum_{k'=1}^{K} \hat{b}_{k'n}} \tag{A11}$$

with

$$\hat{b}_{kn} = \begin{cases} \sum_{k'=1}^{K} p(I_{n+1}|\mu_{k'}, \lambda_{k'}) A_{k'k} b_{k'n+1}, & (n \neq N), \\ 1, & (n = N). \end{cases} \tag{A12}$$

In this calculation, $\boldsymbol{\mu}^{(0)}$, $\boldsymbol{\lambda}^{(0)}$, $\boldsymbol{\pi}^{(0)}$, $\mathbf{A}^{(0)}$ are required to find $p(\mathbf{s}_n|\boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\pi}, \mathbf{A}, \boldsymbol{I})$ [Eq. (A6)] in the initial step. These are evaluated with the hyperparameters in Table I.

(2) $p(\boldsymbol{\mu}, \boldsymbol{\lambda}|\mathbf{S}, \boldsymbol{\pi}, \mathbf{A}, \boldsymbol{I})$ [Eq. (A2)]:

This conditional probability distribution is described as the Gaussian-Gamma distribution:

$$\begin{aligned} &p(\boldsymbol{\mu}, \boldsymbol{\lambda}|\mathbf{S}, \boldsymbol{\pi}, \mathbf{A}, \boldsymbol{I}) \\ &= p(\boldsymbol{\mu}, \boldsymbol{\lambda}|\mathbf{S}, \boldsymbol{I}) \\ &= \prod_{k=1}^{K} \mathrm{NG}(\mu_k, \lambda_k|\hat{m}_k, \hat{\nu}_k, \hat{a}_k, \hat{b}_k) \end{aligned} \tag{A13}$$

with

$$\hat{\nu}_k = \sum_{n=1}^{N} s_{kn} + \nu, \tag{A14}$$

$$\hat{m}_k = \frac{\sum_{n=1}^{N} s_{kn} I_n + \nu m}{\hat{\nu}_k}, \tag{A15}$$

$$\hat{a}_k = \frac{1}{2}\sum_{n=1}^{N} s_{kn} + a, \tag{A16}$$

$$\hat{b}_k = \frac{1}{2}\left(\sum_{n=1}^{N} s_{kn} I_n^2 + \nu m^2 - \hat{\nu}_k \hat{m}_k^2\right) + b, \tag{A17}$$

where the original hyperparameters (ν, m, a, b) of the Gaussian-Gamma distribution are renormalized in as $(\hat{\nu}_k, \hat{m}_k, \hat{a}_k, \hat{b}_k)$.

(3) $p(\boldsymbol{\pi}|\mathbf{S}, \boldsymbol{\mu}, \boldsymbol{\lambda}, \mathbf{A}, \boldsymbol{I})$ [Eq. (A3)]:

This is the Dirichlet distribution:

$$p(\boldsymbol{\pi}|\mathbf{S}, \boldsymbol{\mu}, \boldsymbol{\lambda}, \mathbf{A}, \boldsymbol{I}) = p(\boldsymbol{\pi}|\mathbf{s}_1) = \mathrm{Dir}(\boldsymbol{\pi}|\hat{\boldsymbol{\alpha}}), \tag{A18}$$

where

$$\hat{\alpha}_k = s_{k1} + \alpha_k \tag{A19}$$

with $\hat{\alpha}_k$ being the renormalized hyperparameter of the Dirichlet distribution.

(4) $p(\mathbf{A}|\mathbf{S}, \boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\pi}, \boldsymbol{I})$ [Eq. (A4)]:

This is also the Dirichlet distribution as

$$p(\mathbf{A}|\mathbf{S}, \boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\pi}, \boldsymbol{I}) = p(\mathbf{A}|\mathbf{S}) = \prod_{k'=1}^{K} \mathrm{Dir}(\mathbf{A}_{k'}|\hat{\boldsymbol{\beta}}_{k'}), \tag{A20}$$

where $\hat{\beta}_{kk'}$ is the renormalized hyperparameter from a bare parameter $\beta_{kk'}$ as

$$\hat{\beta}_{kk'} = \sum_{n=2}^{N} s_{k'n-1} s_{kn} + \beta_{kk'}. \tag{A21}$$

[1] A. Seko, T. Maekawa, K. Tsuda, and I. Tanaka, Machine learning with systematic density-functional theory calculations: Application to melting temperatures of single- and binary-component solids, Phys. Rev. B **89**, 054303 (2014).

[2] A. Seko, A. Togo, H. Hayashi, K. Tsuda, L. Chaput, and I. Tanaka, Prediction of Low-Thermal-Conductivity Compounds with First-Principles Anharmonic Lattice-Dynamics Calculations and Bayesian Optimization, Phys. Rev. Lett. **115**, 205901 (2015).

[3] J. Lee, A. Seko, K. Shitara, K. Nakayama, and I. Tanaka, Prediction model of band gap for inorganic compounds by combination of density functional theory calculations and machine learning techniques, Phys. Rev. B **93**, 115104 (2016).

[4] K. Osada, K. Kutsukake, J. Yamamoto, S. Yamashita, T. Kodera, Y. Nagai, T. Horikawa, K. Matsui, I. Takeuchi, and T. Ujihara, Adaptive Bayesian optimization for epitaxial growth of Si thin films under various constraints, Mater. Today Commun. **25**, 101538 (2020).

[5] K. A. Severson, P. M. Attia, N. Jin, N. Perkins, B. Jiang, Z. Yang, M. H. Chen, M. Aykol, P. K. Herring, D. Fraggedakis, M. Z. Bazant, S. J. Harris, W. C. Chueh, and R. D. Braatz, Data-driven prediction of battery cycle life before capacity degradation, Nat. Energy **4**, 383 (2019).

[6] W. Jiang, J. Ahn, F.-L. Xu, C.-Y. Liaw, Y.-C. Chan, Y. Zhou, and Y.-L. Lam, Fabrication of a low-operating voltage diamond thin film metal semiconductor metal photodetector by laser writing lithography, Appl. Phys. Lett. **72**, 1131 (1998).

[7] V. Stanev, K. Choudhary, A. G. Kusne, J. Paglione, and I. Takeuchi, Artificial intelligence for search and discovery of quantum materials, Commun. Mater. **2**, 105 (2021).

[8] Q. Tao, P. Xu, M. Li, and W. Lu, Machine learning for perovskite materials design and discovery, npj Comput. Mater. **7**, 23 (2021).

[9] G. G. C. Peterson and J. Brgoch, Materials discovery through machine learning formation energy, J. Phys.: Energy **3**, 022002 (2021).

[10] A. Jain, S. P. Ong, G. Hautier, W. Chen, W. D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, and K. A. Persson, The Materials Project: A materials genome approach to accelerating materials innovation, APL Mater. **1**, 011002 (2013).

[11] S. P. Ong, W. D. Richards, A. Jain, G. Hautier, M. Kocher, S. Cholia, D. Gunter, V. L. Chevrier, K. A. Persson, and G. Ceder, Python Materials Genomics (pymatgen): A robust, open-source python library for materials analysis, Comput. Mater. Sci. **68**, 314 (2013).

[12] S. P. Ong, S. Cholia, A. Jain, M. Brafman, D. Gunter, G. Ceder, and K. A. Persson, The materials application programming interface (API): A simple, flexible and efficient API for materials data based on REpresentational state transfer (REST) principles, Comput. Mater. Sci. **97**, 209 (2015).

[13] M. de Jong, W. Chen, T. Angsten, A. Jain, R. Notestine, A. Gamst, M. Sluiter, C. Krishna Ande, S. van der Zwaag, J. J. Plata, C. Toher, S. Curtarolo, G. Ceder, K. A. Persson, and M. Asta, Charting the complete elastic properties of inorganic crystalline compounds, Sci. Data **2**, 150009 (2015).

[14] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg *et al.*, Scikit-learn: Machine learning in python, J. Mach. Learn. Res. **12**, 2825 (2011).

[15] C.-C. Chang and C.-J. Lin, Libsvm: A library for support vector machines, ACM Trans. Intell. Syst. Technol. **2**, 1 (2011).

[16] F. M. Puglisi, P. Pavan, A. Padovani, L. Larcher, and G. Bersuker, RTS noise characterization of HfO_x RRAM in high resistive state, Solid-State Electron. **84**, 160 (2013).

[17] F. M. Puglisi, P. Pavan, A. Padovani, and L. Larcher, A study on HfO_2 RRAM in HRS based on *I-V* and RTN analysis, Solid-State Electron. **102**, 69 (2014).

[18] F. Puglisi and P. Pavan, Factorial hidden Markov model analysis of random telegraph noise in resistive random access memories, ECTI Trans. Elect. Eng. Electr. Commun. **12**, 24 (2014).

[19] B. Stampfer, F. Zhang, Y. Y. Illarionov, T. Knobloch, P. Wu, M. Waltl, A. Grill, J. Appenzeller, and T. Grasser, Characterization of single defects in ultrascaled MoS_2 field-effect transistors, ACS Nano **12**, 5368 (2018).

[20] T. Murai, T. Makino, H. Kato, M. Shimizu, T. Murooka, E. D. Herbschleb, Y. Doi, H. Morishita, M. Fujiwara, M. Hatano, S. Yamasaki, and N. Mizuochi, Engineering of Fermi level by *nin* diamond junction for control of charge states of NV centers, Appl. Phys. Lett. **112**, 111903 (2018).

[21] Y. Doi, T. Fukui, H. Kato, T. Makino, S. Yamasaki, T. Tashima, H. Morishita, S. Miwa, F. Jelezko, Y. Suzuki, and N. Mizuochi, Pure negatively charged state of the NV center in *n*-type diamond, Phys. Rev. B **93**, 081203(R) (2016).

[22] T. Tamaoka, Y. Midoh, K. Yamamoto, S. Aritomi, T. Tanigaki, M. Nakamura, K. Nakamae, M. Kawasaki, and Y. Murakami, Denoising electron holograms using the wavelet hidden Markov model for phase retrieval-applications to the phase-shifting method, AIP Adv. **11**, 025135 (2021).

[23] S. Gammelmark, K. Mølmer, W. Alt, T. Kampschulte, and D. Meschede, Hidden Markov model of atomic quantum jump dynamics in an optically probed cavity, Phys. Rev. A **89**, 043839 (2014).

[24] M. Pirchi, R. Tsukanov, R. Khamis, T. E. Tomov, Y. Berger, D. C. Khara, H. Volkov, G. Haran, and E. Nir, Photon-by-photon hidden Markov model analysis for microsecond single-molecule FRET kinetics, J. Phys. Chem. B **120**, 13065 (2016).

[25] D. H. Nguyen-Le, Q. Tao, V.-H. Nguyen, M. Abdel-Wahab, and H. Nguyen-Xuan, A data-driven approach based on long short-term memory and hidden Markov model for crack propagation prediction, Eng. Fract. Mech. **235**, 107085 (2020).

[26] C. C. Strelioff, J. P. Crutchfield, and A. W. Hübler, Inferring Markov chains: Bayesian estimation, model comparison, entropy rate, and out-of-class modeling, Phys. Rev. E **76**, 011106 (2007).

[27] L. Zhuang, L. Guo, and S. Y. Chou, Silicon single-electron quantum-dot transistor switch operating at room temperature, Appl. Phys. Lett. **72**, 1205 (1998).

[28] S. Pirandola, J. Eisert, C. Weedbrook, A. Furusawa, and S. L. Braunstein, Advances in quantum teleportation, Nat. Photon. **9**, 641 (2015).

[29] D. L. Huffaker, G. Park, Z. Zou, O. B. Shchekin, and D. G. Deppe, 1.3 μm room-temperature GaAs-based quantum-dot laser, Appl. Phys. Lett. **73**, 2564 (1998).

[30] A. Nozik, Quantum dot solar cells, Phys. E **14**, 115 (2002).

[31] A. Imamoğlu, D. D. Awschalom, G. Burkard, D. P. DiVincenzo, D. Loss, M. Sherwin, and A. Small, Quantum Information Processing Using Quantum Dot Spins and Cavity QED, Phys. Rev. Lett. **83**, 4204 (1999).

[32] M. Nirmal, B. O. Dabbousi, M. G. Bawendi, J. J. Macklin, J. K. Trautman, T. D. Harris, and L. E. Brus, Fluorescence intermittency in single cadmium selenide nanocrystals, Nature (London) **383**, 802 (1996).

[33] M. Kuno, D. P. Fromm, H. F. Hamann, A. Gallagher, and D. J. Nesbitt, "On"/"off" fluorescence intermittency of single semiconductor quantum dots, J. Chem. Phys. **115**, 1028 (2001).

[34] L. Rabiner, A tutorial on hidden Markov models and selected applications in speech recognition, Proc. IEEE **77**, 257 (1989).

[35] Z. Ghahramani, An introduction to hidden Markov models and Bayesian networks, in *Hidden Markov Models* (World Scientific, 2001), pp. 9–41.

[36] M. R. Hassan, A combination of hidden Markov model and fuzzy model for stock market forecasting, Neurocomputing **72**, 3439 (2009).

[37] K.-C. Kwon and J.-H. Kim, Accident identification in nuclear power plants using hidden Markov models, Eng. Appl. Artif. Intell. **12**, 491 (1999).

[38] B. Bruhn, F. Qejvanaj, I. Sychugov, and J. Linnros, Blinking statistics and excitation-dependent luminescence yield in Si and CdSe nanocrystals, J. Phys. Chem. C **118**, 2202 (2014).

[39] C. S. Jensen, U. Kjærulff, and A. Kong, Blocking Gibbs sampling in very large probabilistic expert systems, Int. J. Hum. Comput. Stud. **42**, 647 (1995).

[40] K. Hashizume, M. Matsubayashi, M. Vacha, and T. Tani, Individual mesoscopic structures studied with sub-micrometer optical detection techniques: CdSe nanocrystals capped with TOPO and ZnS-overcoated system, J. Lumin. **98**, 49 (2002).

[41] M. Oda, A. Hasegawa, N. Iwami, K. Nishiura, N. Ando, A. Nishiyama, H. Horiuchi, and T. Tani, Photoluminescence behaviors of single CdSe/ZnS/TOPO nanocrystals: Adsorption effects of water molecules onto nanocrystal surfaces, J. Lumin. **127**, 198 (2007).

[42] We estimated the noise size due to the dead time of the detector quantitatively as follows: First, the dead time of the detector is estimated as 50 ns. Next, the exposure time (= one division of the time grid) Δt is 50 ms, the photoluminescence intensity per exposure time in the light-emission state I_{ON} is ∼3000 counts, and the fluorescence intensity per exposure time in the quenching state I_{OFF} is ∼2800 counts. Here, we note that one photon is recorded as one count. Thus, the total dead times per exposure time of the light-emission and quenching states are 0.15 ms (= 3000×50 ns) and 0.14 ms (= 2800×50 ns), respectively. Then, the numbers of counting loss during the exposure time in the light-emission and quenching states are estimated as 9[= 3000×(0.15 ms/50 ms)] ± 3 counts and 7.8[= 2800×(0.14 ms/50 ms)] ± 2.8 counts, respectively, where the deviation is evaluated as a square root of the number of the counting loss. Now, in the present experimental time series, relatively large intensity fluctuations were observed for the light-emission and quenching states; the range of the fluctuations is ±150 counts, and therefore $I_{\mathrm{ON}} = 3000 \pm 150$ counts and $I_{\mathrm{OFF}} = 2800 \pm 150$ counts. These fluctuations are much larger than those due to the dead time (about ±3 estimated above). Thus, we think that the noise effect due to the dead time of the detector is negligible in the results with the present HMM analysis.

[43] K. T. Shimizu, R. G. Neuhauser, C. A. Leatherdale, S. A. Empedocles, W. K. Woo, and M. G. Bawendi, Blinking statistics in single semiconductor nanocrystal quantum dots, Phys. Rev. B **63**, 205316 (2001).

[44] K. T. Early and D. J. Nesbitt, Size-dependent photoionization in single CdSe/ZnS nanocrystals, Nano Lett. **13**, 4844 (2013).

[45] X. Hou, J. Kang, H. Qin, X. Chen, J. Ma, J. Zhou, L. Chen, L. Wang, L.-W. Wang, and X. Peng, Engineering Auger

recombination in colloidal quantum dots via dielectric screening, Nat. Commun. **10**, 1750 (2019).

[46] S. Geman and D. Geman, Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images, IEEE Trans. Pattern Anal. Mach. Intell. **PAMI-6**, 721 (1984).

[47] L. E. Baum, An inequality and associated maximization technique in statistical estimation for probabilistic functions of Markov processes, in *Inequalities III: Proceedings of the Third Symposium on Inequalities* (Academic Press, 1972), Vol. 3, pp. 1–8.