pubs.acs.org/JPCC

A Maximum Likelihood Method for Power Law Distributions That Does Not Break Down When the Slope Is Close to Unity

Zhaoyan Zhu and R. A. Marcus*

Noyes Laboratory of Chemical Physics, California Institute of Technology, Pasadena, California 91125, United States

ABSTRACT: A general maximum likelihood estimation (MLE) method is given to analyze $\ln \hat{L} = \sum_{i} \ln p(x_i, \theta)$ experimental data with a power law form with any power exponent which does not break down for a power close to -1. It contrasts thereby with a standard procedure that does. It can be extended to a power law with an exponential tail and more generally to other distribution forms. Inasmuch as the theoretical value of the power for dye-sensitized charge recombination in semiconductors systems, and for certain charge injection, is -1 (Chen, W.; Marcus, R. A., J. Phys. Chem. C, accepted), the present correction to the current MLE method has immediate application to the data in these systems, but it is equally applicable to other systems, regardless of whether the power is -1.

INTRODUCTION

Single-molecule spectroscopy (SMS) has been an important tool in physics, chemistry, and biology.²⁻⁶ It allows for photophysical measurements of individual luminophores, revealing behavior undetectable in ensemble measurements. With recent advances in technique, 7 room-temperature singleemitter experiments have provided observations of a pronounced blinking behavior that is defined as the random switching on and off of the fluorescence or luminescence intensity of a single emitter under continuous or pulsed excitation.⁸⁻¹³ A well-known and intensively studied example is quantum dots blinking. 14-23 A unique feature of this blinking behavior is that the distribution of "on" and "off" duration times is a power law with broad range of decades duration instead of an exponential.^{8,12–16,24,25}

Although the power law phenomenon is less intensively studied in the blinking of single dye molecules than of inorganic quantum dots, there are a few observations of power law behavior of organic dye molecules embedded in polymer or on a glass or inorganic crystal surface, 24-26 in contrast with the numerous studies of quantum dots. However, because of the relatively low fluorescence efficiency and tendency to bleach, only limited data sets can be collected from experiments with the dyes. Recently, the maximum likelihood estimator (MLE) method has been adopted to analyze these power-law distributed data, especially for the limited data sets. 27-31

In a recent analysis of experimental results, the fitting obtained by MLE did not pass through or near many data points.³⁰ After a study of the method, we are able to understand the difficulty with the literature method and provide a solution.

■ MLE METHOD AND ALGORITHM

We recall that MLE is a method of estimating the parameters of a statistical model. Given a statistical model and a set of data (observations), MLE provides estimates of the model's parameters.³² Suppose one makes N independent and identically distributed (iid) observations (measurements) x_1 , x_2 , ..., x_N from a distribution with an unknown probability density function (pdf), $p_0(x)$. With certain experience, it is then

conjectured that $p_0(x)$ belongs to a certain family of distributions, $p_0(x,\theta),\theta \in \Theta$. Here we denote by θ the parameters of the pdf and by Θ the range of θ . $p_0(x) = p_0(x,\theta_0)$ with θ_0 as the unknown "true value" of the parameters in this

The problem of obtaining the parameter closest to the "true value" becomes equivalent to selecting the value that produces a distribution that gives the greatest probability (likelihood) for

For an iid sample, the joint density function is 32

$$p(x_1, x_2, ..., x_n, \theta) = p(x_1, \theta)p(x_2, \theta)...p(x_N, \theta)$$
 (1)

where x_1 , x_2 ..., x_N are the measurements in the following discussion. The above joint density function can also be taken as the likelihood when we consider θ as the variable³²

$$\hat{l}(\theta, x_1, x_2, ..., x_n) = p(x_1, x_2, ..., x_n, \theta)$$
(2)

In the method of maximum likelihood, one finds a value of θ , θ_{0} , that maximizes the likelihood. The likelihood is a multiplication, and if we use a monotonic transformation of it, a logarithm, then we will have a summation, log-likelihood, which is easier to work with and maximize than the product

$$\ln \hat{l} = \sum_{i=1}^{N} \ln p(x_i, \theta)$$
(3)

■ MLE FOR POWER LAW DISTRIBUTION

Method. We indicate the model values as \hat{X} in comparison with the experimental data X. In a perylene bisimide dye molecule fluorescence blinking case, which prompted our interest, 30 the measurements are the "on" or "off" durations t_i s, corresponding to the above x_i s. The pdf of power law distribution is

Received: April 17, 2012 Revised: June 1, 2012 Published: June 7, 2012

$$\hat{p}(t) = At^{-m} \tag{4}$$

The probability of an observation at t_i is $\hat{P}(t_i) \equiv \int_{t_i}^{t_i + \Delta t_i} \hat{p}(t) \, dt$. In experiments, $P(t_i) = (n_{t_i})/\Sigma_{i=1}^k \ n_{t_i}$ and $P(t_i) \equiv P(t_i)/\Delta t_i$. Each n_i is the number of observations obtained in the time interval Δt_i around t_i and $\sum_{i=1}^k n_i = N$ and k is the number of intervals. There are several ways of choosing Δt_i . We adopt the commonly used one $\Delta t_i = (t_{i+1} - t_{i-1})/2$ for 1 < i < k, and for Δt_1 and Δt_k we use the experimental data points at the cited times t_i . In the model, we use the approximation $\hat{p}(t_i) \simeq \hat{P}(t_i)/\Delta t_i$, and then we have $A(m) = (1)/(\sum t_i^{-m} \Delta t_i)$. We now have the model with parameter m

$$\hat{p}(t) = \frac{1}{\sum_{i=1}^{k} t_i^{-m} \Delta t_i} t^{-m}$$
(5)

The likelihood function in terms of the model parameter m is $N!\prod_{i=1}^k \hat{P}(t_i)n_i/n_i!$. We obtain an estimation of m closest to m_0 by maximizing the log-likelihood function

$$\ln \hat{l} = \ln N! + \sum \left[n_i \ln \hat{P}(t_i) - \ln n_i! \right]$$
(6)

More generally, if the functional form is p(t,m), where we now denote by m the collective parameters, m_1 , m_2 , ..., then for $\hat{p}(t_i)$ we would have instead of eq 5

$$\hat{p}(t) = \frac{p(t, m)}{\sum p(t_i, m)\Delta t_i}$$
(7)

One can obtain an estimate of the error variance³³ σ^2 of $\ln p(t_i)$ from the residue (the difference between the observation $\ln p(t_i)$ and the estimation $\ln \hat{p}_i(t)$)

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^k (\ln \hat{p}_i(t) - \ln p_i(t))^2}{k-2}$$
 (8)

Here in the fitting there is one parameter, m, and one constraint $(\sum p(t_i)\Delta t_i = 1)$, and thus the residue sum of squares has k-2 degrees of freedom. We may also estimate the variance of m^{33}

$$Var(m) = \frac{\hat{\sigma}^2}{\sum \left(\frac{\partial \ln \hat{P}_i(t)}{\partial m}\right)^2}$$
(9)

and the standard error is

$$se(m) = \sqrt{Var(m)}$$
 (10)

Comparison with an Earlier MLE Method. To normalize the PDF and obtain an expression of A as a function of m, we avoid the integration of the previous method.²⁷ Instead, we use the summation of the probability of all measurements. Unlike the requirement for the previous method that $m_0 \neq 1$, with the present method, even when the true value is $m_0 = 1$, one can still fit the experimental data well and get an m close to m_0 . The method is applicable regardless of whether m is 1.

RESULTS AND DISCUSSION

We next apply the method to the intermittent fluorescence experimental data of the single perylene bisimide dye molecules on the ${\rm Al_2O_3}$ system.³⁴ Processing the fluorescence data in the algorithm described by eqs 4–6, we have obtained a power law distribution with slope around 1 for both off and on times, as in Figures 1 and 2. These results agree well with the diffusion-based model of Chen and Marcus.¹

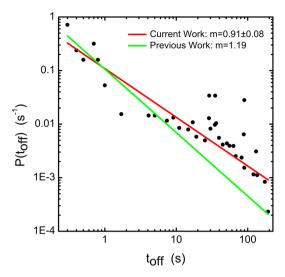


Figure 1. "Off" probability density for 25 molecules observed on sapphire (0001) using original experimental data in ref 30 from Prof. Monti. The green line is from the original MLE fitting with $P = At^{-m}$, where m = 1.19 in ref 30 and the red line is obtained from the present modified MLE fitting with $P = At^{-m}$ and where m = 0.91 and se = 0.08.

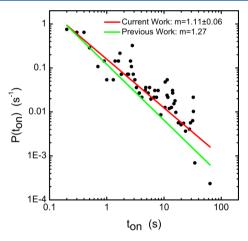


Figure 2. "On" probability density for 51 molecules observed on sapphire (0001) using original experimental data in ref 30 from Prof. Monti. The green line is from the original MLE fitting with $P = At^{-m}$, where m = 1.27 in ref 30. The red line is obtained from the present modified MLE fitting with $P = At^{-m}$ and where m = 1.11 and se = 0.06.

We next compare with a previous method of implementing the MLE. In comparison with the normalization method used above to obtain the function A(m), one can compare with the method of approximation used previously in the literature. The observation data points are within a range between the experimental resolution time t_{\min} and the time window of the experiment t_{\max} . So $\int_{t_{\min}}^{t_{\max}} \hat{p}(t_i) dt = 1$, and under the condition that $m_0 \neq 1$, $A(m) = (m-1)/(t_{\min}^{1-m} - t_{\max}^{1-m})$ and we have

$$\hat{p}(t) = \frac{m-1}{t_{\min}^{1-m} - t_{\max}^{1-m}} t^{-m} \tag{11}$$

so that the integral of $\hat{p}(t)$ from t_{\min} to t_{\max} is unity. However, when m approaches unity, $\hat{p}(t)$ approaches 0/0 and so is indeterminate.

Taking $t_{\rm max}\gg t_{\rm min}$ because the experimental time window spans several orders of magnitude and presuming $m_0>1$, eq 11 becomes

$$\hat{p}(t) \cong \frac{m-1}{t_{\min}^{1-m}} t^{-m} \tag{12}$$

This method was previously adopted to analyze the fluorescence of tetraphenoxy-perylene diimide dye. ²⁸ In that case, the diffusion process was a "spectral diffusion" rather than a particle diffusion, and one observed a power exponent $m_0 \cong 1.5$. ^{35–37} However, in the special case $m_0 \cong 1$, which applies when a particle diffusion on a 2-D surface is involved, ³⁰ the condition for the validity of eq 11 is no longer satisfied, and so we cannot retain the expression in eq 12. Indeed, eqs 11 and 12 do not give a close fit to the data (Figures 1and 2). For the single-molecule study of the electron injection from the dye onto a semiconductor surface, $m_0 \cong 1$ is the theoretically expected value, ¹ and a treatment of this particular case is particularly necessary.

In analyzing data with forms other than power law, for example, power law with an exponential tail, the normalization demonstrated above can also be applied.

In the "on" plot, the last two points deviate from the power law. One could fit these with an exponential tail. However, the two points occur where the error is large (small signal intensity) and should be given little or no weight. More data in that region are needed. Because there are only two data points, one can not decide definitively whether this occurrence is an exponential cutoff, bleaching of the dye molecule or other. There is, however, another photoexcited dye-on-surface result that does show an exponential cutoff.³⁸

CONCLUDING REMARKS

A literature method for treating the experimental data with MLE has a singularity when the data obey a power law with a slope equal to -1. We have given instead a general MLE method to analyze experimental data with a distribution of a power law form having any power exponent, including -1. It can be extended to many other functional forms such as a power law with an exponential tail.

AUTHOR INFORMATION

Corresponding Author

*E-mail: ram@caltech.edu.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We acknowledge the support of this research by NSF, ARO, and ONR and thank Prof. Monti for providing the original experimental data on which Figures 1 and 2 are based. We use data given in ref 30 here, but to be more precise we have the original data from Prof. Monti to verify.

REFERENCES

- (1) Chen, W.-C.; Marcus, R. A. J. Phys. Chem. C, DOI: http://dx.doi.org/10.1021/jp303837g.
- (2) Moerner, W. E.; Orrit, M. Science 1999, 283, 1670-1676.
- (3) Kultzer, F.; Orrit, M. Annu. Rev. Phys. Chem. 2004, 55, 585-611.
- (4) Tinnefeld, P.; Sauer, M. Angew. Chem., Int. Ed. 2005, 44, 2642–2671.
- (5) Michalet, X.; Kapanidis, A. N.; Laurence, T.; Pinaud, F.; Doose, S.; Pflughoefft, M.; Weiss, S. Annu. Rev. Biophys. Biomol. Struct. 2003, 32, 161–182.
- (6) Barkai, E.; Jung, Y. J.; Silbey, R. Annu. Rev. Phys. Chem. 2004, 55, 457–507.

- (7) Xie, X. S. Acc. Chem. Res. 1996, 29, 598-606.
- (8) Basche, T.; Kummer, S.; Brauchle, C. Nature 1995, 373, 132–134.
- (9) Kohn, F.; Hofkens, J.; Gronheil, R.; Van der Auweraer, M.; De Schryver, F. C. J. Phys. Chem. A 2002, 106, 4808-4814.
- (10) Veerman, J. A.; Garcia-Parajo, M. F.; Kuipers, L.; van Hulst, N. F. *Phys. Rev. Lett.* **1999**, 83, 2155–2158.
- (11) Kulzer, F.; Kummer, S.; Basche, T.; Brauchle, C. J. Inf. Rec. 1996, 22, 567–572.
- (12) Dickson, R. M.; Cubitt, A. B.; Tsien, R. Y.; Moerner, W. E. *Nature* **1997**, *388*, 355–358.
- (13) Vandenbout, D. A.; Yip, W. T.; Hu, D. H.; Fu, D. K.; Swager, T. M.; Barbara, P. F. *Science* **1997**, 277, 1074–1077.
- (14) Kuno, M.; Fromm, D. P.; Hamann, H. F.; Gallagher, A.; Nesbitt, D. J. *J. Chem. Phys.* **2000**, *112*, 3117–3120.
- (15) Neuhauser, R. G.; Shimizu, K. T.; Woo, W. K.; Empedocles, S. A.; Bawendi, M. G. *Phys. Rev. Lett.* **2000**, *85*, 3301–3304.
- (16) Shimizu, K. T.; Neuhauser, R. G.; Leatherdale, C. A.; Empedocles, S. A.; Woo, W. K.; Bawendi, M. G. *Phys. Rev. B* **2001**, 63, 205316–1–205316–5.
- (17) Kuno, M.; Fromm, D. P.; Hamann, H. F.; Gallagher, A.; Nesbitt, D. J. *J. Chem. Phys.* **2001**, *115*, 1028–1040.
- (18) van Sark, W.; Frederix, P.; Bol, A. A.; Gerritsen, H. C.; Meijerink, A. Chem. Phys. Chem. **2002**, *3*, 871–879.
- (19) Verberk, R.; van Oijen, A. M.; Orrit, M. *Phys. Rev. B* **2002**, *66*, 233202-1–233202-4.
- (20) Stefani, F. D.; Knoll, W.; Kreiter, M.; Zhong, X.; Han, M. Y. *Phys. Rev. B* **2005**, *72*, 125304-1–125304-7.
- (21) Verberk, R.; Chon, J. W. M.; Gu, M.; Orrit, M. *Physica E* **2005**, 26, 19–23.
- (22) Issac, A.; von Borczyskowski, C.; Cichos, F. Phys. Rev. B 2005, 71, 161302-1–161302-4.
- (23) Brokmann, X.; Hermier, J. P.; Messin, G.; Desbiolles, P.; Bouchaud, J. P.; Dahan, M. *Phys. Rev. Lett.* **2003**, *90*, 120601-1–120601-4.
- (24) Hoogenboom, J. P.; van Dijk, H. P.; Hernando, J.; van Hulst, N. F.; Garcia-Parajo, M. F. *Phys. Rev. Lett.* **2005**, *95*, 097401-1–097401-4.
- (25) Schuster, J.; Cichos, F.; von Borczyskowski, C. Appl. Phys. Lett. **2005**, 87, 051915-1–051915-3.
- (26) Yeow, E. K. L.; Melnikov, S. M.; Bell, T. D. M.; De Schryver, F. C.; Hofkens, J. *J. Phys. Chem. A* **2006**, *110*, 1726–1734.
- (27) Hoogenboom, J. P.; den Otter, W. K.; Offerhaus, H. L. J. Chem. Phys. **2006**, 125, 204713-1-204713-12.
- (28) Hoogenboom, J. P.; Hernando, J.; Garca-Paraj, M. F.; van Hulst, N. F. *J. Phys. Chem. C* **2008**, *112*, 3417–3422.
- (29) Wustholz, K. L.; Bott, E. D.; Kahr, B.; Reid, P. J. J. Phys. Chem. C 2008. 112. 7877-7885.
- (30) Schirra, L. K.; Tachett, B. S.; Blumenfeld, M. L.; Monti, O. L. A. *J. Chem. Phys.* **2009**, *131*, 124702-1–124702-7.
- (31) Watkins, L. P.; Yang, H. J. Phys. Chem. B 2005, 109, 617-628.
- (32) Pfanzagl, J. Parametric Statistical Theory; de Gruyter, New York,
- (33) Montgomery, D. C.; Peck, E. A.; Vining, G. G. Introduction to Linear Regression Analysis; John Wiley and Sons, Inc.: Hoboken, NJ, 2006.
- (34) The data are the same as those in ref 30. We thank Prof. Monti for providing us with his data points.
- (35) Tang, J.; Marcus, R. A. J. Chem. Phys. 2005, 123, 054704-1-054704-12.
- (36) Tang, J.; Marcus, R. A. J. Chem. Phys. 2005, 123, 204511-1–204511-6.
- (37) Frantsuzov, P. A.; Marcus, R. A. Phys. Rev. B 2005, 72, 155321-1-155321-10
- (38) Bian, Z.; Tachikawa, T.; Cui, S.; Tojo, S.; Fujitsuka, M.; Majima, T. Chem. Sci. **2012**, *3*, 370–379.

■ NOTE ADDED AFTER ASAP PUBLICATION

This paper was published on the Web on June 20, 2012. Corrections have been made to Reference 1. The correct version was reposted on June 25, 2012.