Statistical analysis of data from single molecule experiment

Tomasz Burzykowski[a], Jacek P. Szubiakowski[*b], Tobias Rydén[c]
[a]Center for Statistics, Limburgs Universitair Centrum, Belgium;
[b]Department of Physics and Computer Methods, Warmia and Masuria University in Olsztyn, Poland;
[c]Centre for Mathematical Sciences, Department of Mathematical Statistics, Lund University, Sweden

ABSTRACT

The maximum-likelihood estimation method is considered. It is based on three-level Dehmelt scheme of isolated quantum system and it makes use of the theory of Markov-modulated Poison process (MMPP).
This method is applied to the analysis of the photon-counting distribution resulting from a single molecule fluorescence experiment. It is assumed that the data available for the estimation contain only the number of collected photons per time interval.

Keywords: Markov-modulated Poisson process, hyperexponential distribution, single molecule spectroscopy

1. INTRODUCTION

Dehmelt[1] suggested a scheme of experiment to detect the quantum jumps in a single double-resonance atom involving weak and strong transition. According to the scheme, quantum jumps at the weak transition cause the fluorescence of the strong transition to turn on and off abruptly. The fluorescence is off when the weak transition is excited and on when it is not. Consequently, the atomic fluorescence should have the form of a random telegraph signal, with alternating periods of zero ("off" periods) and non-zero ("on" periods) intensity. Thus, the quantum jumps could be directly monitored by observation of the random telegraph signal radiated by the strong transition.. Cook and Kimble[2] followed Dehmelt's idea and developed a mathematical description of his two parameter scheme. Their main conclusion was that the durations of "off" and "on" intervals should follow exponential distributions, with different rates.

Quantum jumps between excited electronic levels have been also observed as a fluorescence "blinking" in experiments involving a single molecule of different organic dyes. These molecules exhibit intersystem crossing to the long-lived triplet state T_1. The single molecule fluorescence experiments may be divided into two types. In a common single molecule experiment arriving photons of the fluorescence are collected consecutively during the defined intervals (dwell periods). This technique produces the fluorescence trace of the form of the telegraph noise[3]. In such a type of experiment arrival time of fluorescence photon is lost. The other type of single molecule fluorescence experiments that provide more information on the photophysical processes has been developed in recent years. In this kind of experiments the fluorescence photon arriving time is also registered[4]. In this paper we restrict our consideration to the first type of single molecule experiments.

The parameters of the stochastic process commonly estimated on the basis of experimental data[3] are "off" and "on" - time duration and the number of "on"-state photon counts. In the approaches proposed by Ha et al.[3] and Yip et al.[5], each dwell period of the trace is classified as belonging to "on" or "off" period, depending on the number of emitted photons collected in it. The classification rule can use, e.g., a threshold derived from the distribution of observed counts[5]. Subsequently, the histograms of the durations of "on" and "off" are analyzed using separate exponential distributions.

The aforementioned methods require arbitrary choices for the construction of the classification rule. Recently, assuming a kinetics model for the excitation-relaxation process in the molecule, a theoretical form of the histogram of the total

*jpszub@matman.uwm.edu.pl; phone: +48 89 527 60 33 internal 332; fax: +48 89 533 49 84; Department of Physics and Computer Methods, Warmia and Masuria University in Olsztyn, ul. Żołnierska 14, 10-561 Olsztyn, Poland.

Proceedings of SPIE Vol. 5258 IV Workshop on Atomic and Molecular Physics,
edited by Józef Heldt (SPIE, Bellingham, WA, 2003) · 0277-786X/03/$15.00

number of collected photons per dwell period has been proposed[6]. In principle, this method allows to estimate rates of the photophysical process providing the knowledge of the number of photons collected in dwell periods.

In this paper we consider an alternative formal approach to the estimation of the parameters of a single-molecule fluorescence experiment, involving three electronic levels. In particular, we consider maximum-likelihood estimation, using the theory of Markov-modulated Poisson processes (MMPP). Though the theory is well-known, only recently Haflidadóttir[7] has addressed the problem of parameter estimation if the available data contain only counts of events over a certain number of time-intervals that is relevant to case considered in this paper.

2. THEORY

2.1 MARKOV-MODULATED POISSON PROCESS

MMPP is a Poisson process with arrival rate that varies according to a continuous-time m-state irreducible Markov chain[8]. We restrict the consideration of MMPP to the case of $m = 2$, that fits to the problem of single molecule fluorescence "blinking". Depending on the state, the arrival rate equals λ_1 and λ_2, respectively. The MMPP is parametrized by the generator Q and the arrival rates λ_1 and λ_2 represented through the matrix Λ:

$$\mathbf{Q} = \begin{bmatrix} -\omega_1 & \omega_1 \\ \omega_2 & -\omega_2 \end{bmatrix}, \qquad \Lambda = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}.$$

The off diagonal elements $\{\mathbf{Q}\}_{ij}$ $(i \neq j)$ of $\mathbf{Q}$ represent the parameters of the exponential density for the transition to state j, given that the Markov chain is in state i. The diagonal elements $\{\mathbf{Q}\}_{ii}$ stand for the parameters of the exponential density for leaving the current state i; they equal the negative of the sums of the off diagonal elements on the respective rows. The diagonal elements $\{\Lambda\}_{ii}$ represent the parameters of the exponential density for time between arrivals from the Poisson process, given that the Markov chain is in state i.

To specify an MMPP completely, one needs to define the state of Markov chain at time $t = 0$. Let J_0 denote the state of the Markov chain at time t=0. If the initial moment is chosen to correspond to one of arrivals, then we obtain the *event-stationary* version of MMPP[8]. In this case, $\{P(J_0 = 1), P(J_0 = 2)\}^T$ is chosen to equal

$$v = \begin{pmatrix} \dfrac{\lambda_1 \omega_2}{\lambda_1 \omega_2 + \lambda_2 \omega_1} \\ \dfrac{\lambda_2 \omega_1}{\lambda_1 \omega_2 + \lambda_2 \omega_1} \end{pmatrix} \tag{1}$$

The components of vector v in Eq. **(1)** can be interpreted as the proportions of arrivals generated in the corresponding state if the Markov chain is observed over a long period of time.

In what follows, probabilities for counts of events over time intervals will be needed. Let N_t be the number of arrivals in (0, t] period and J_t be the state of the Markov chain at time t. Define $P_{ij}(n,t) = P(N_t = n, J_t = j \mid N_0 = 0, J_0 = i)$, where n = 0,1,2,3,...and i,j = 1,2, and put

$$P(n,t) = \begin{bmatrix} P_{11}(n,t) & P_{12}(n,t) \\ P_{21}(n,t) & P_{22}(n,t) \end{bmatrix} \tag{2}$$

The matrices $\boldsymbol{P}(n\ t)$ satisfy the forward Chapman-Kolmogorov equations[9]

$$P(0,0) = I,$$

$$\frac{\partial}{\partial t} P(0,t) = P(0,t)(\mathbf{Q} - \Lambda),$$

$$\frac{\partial}{\partial t} P(n,t) = P(n,t)(\mathbf{Q} - \Lambda) + P(n-1,t)\Lambda.$$

The solution can be formally given by matrix exponential. The closed-form expression for the matrix may be provided only in simple cases. In general, the elements of the matrix can be computed numerically by applying the uniformization method[9].

If $\lambda_2 = 0$,the MMPP becomes a Markov renewal process known as the interrupted Poisson process (IPP)[8]. It can be instantly seen, that "on" and "off" times in Dehmelt's scheme can be considered as states of Markov chain, with exponentially distributed durations. The photon emissions, on the other hand, can be assumed to occur according to a Poisson process only during the "on" time (chain state).

For IPP the distribution of time between successive arrivals from the Poission process is hyperexponential, with probability density function given by[10]

$$f(t) = p\gamma_1 \exp(-\gamma_1 t) + (1-p)\gamma_2 \exp(-\gamma_2 t). \tag{3}$$

Formally the hyperexponential density is a probability density of mixture of two exponential distributions with a Bernoulli mixing distribution with parameter p. It can be shown that, contrary to an ordinary Poisson process, an IPP does not possess the "lack of memory" property[11].

The following relationships between the parameters of an IPP (λ_1, ω_1, ω_2) and the hyperexponential distribution (p,γ_1, γ_2) hold[11]:

$$\lambda_1 = p\gamma_1 + (1-p)\gamma_2, \quad \omega_1 = \frac{p(1-p)(\gamma_1-\gamma_2)^2}{\lambda_1}, \quad \omega_2 = \frac{\gamma_1\gamma_2}{\lambda_1}. \tag{4}$$

2.2 MODEL OF MOLECULAR KINETICS

The fluorescence trace obtained in the single molecule spectroscopy experiment can be described by the analogy to the double resonance scheme by Dehmelt[1] by distinguishing among three energy-levels of the molecule: ground state - S_0, singlet excited state - S_1 and triplet state - T_1. The ground state S_0 is common for both direct $S_1 \rightarrow S_0$ relaxation and for deactivation via the triplet state $S_1 \rightarrow T_1 \rightarrow S_0$. For the sake of simplicity, let us assume that each $S_1 \rightarrow S_0$ relaxation results in the emission of a photon. Thus, if two direct relaxation of excited singlet state take place immediately one after the other, two photons will be registered. In principle the time that elapses between the emission of these photons can be considered as associated with electron passage through the following three stages: excitation $S_0 \rightarrow S_1$, residence at level S_1 and relaxation $S_1 \rightarrow S_0$. An alternative way consists of the sequence of the following stages: excitation $S_0 \rightarrow S_1$, residence at level S_1, intersystem crossing and relaxation $T_1 \rightarrow S_0$. Note that $S_1 \rightarrow T_1 \rightarrow S_0$ transition is not directly observable. Consequently, during the time that elapses between the observation of two emitted photons, it is possible that a sequence of $i = 0,1,2,\ldots$ decays *via* triplet state, followed by one direct singlet relaxation, have taken place.

In a typical single molecule spectroscopy experiment the molecule is irradiated with cw laser light in the resonance with the $S_0 \rightarrow S_1$ zero phonon transition. The excited molecule may fluoresce and relax to the ground state or with a very low probability it may undergo the intersystem crossing to T_1. Due to the spin-forbiden character of singlet-triplet transitions, T_1 is a metastable state with the lifetime much longer then that for S_1. The eventual relaxation of T_1 is predominantly a nonradiative process, which occurs between the isoenergetic levels of the singlet and triplet systems. The excitation-relaxation process may be modelled as a continuous time stochastic process.[12] The kinetics of this process is described by the set of first-order differential equations for probabilities $p_i(t)$ that the molecule is in state i = 0, 1, 2 for ground (S_0), singlet excited state (S_1), or triplet state (T_1), respectively:

$$\begin{aligned}
\frac{dp_0(t)}{dt} &= k_{01}\, p_1(t) + k_{02}\, p_2(t), \\
\frac{dp_1(t)}{dt} &= -(k_{01} + k_{21})\, p_1(t), \\
\frac{dp_2(t)}{dt} &= k_{21}\, p_1(t) - k_{02}\, p_2(t),
\end{aligned} \tag{5}$$

Let us assume that only the singlet excited state is populated immediately after excitation, i.e., $p_1(0) = 1$. From this condition it follows that $p_0(0) = p_2(0) = 0$. The transitions between states are specified by the rate coefficients $k_{01} = k_f + k_{nr}$, where k_f is the rate constant for fluorescence and k_{nr} for nonradiative deactivation, k_{21} is the rate of inter-system crossing, k_{02} the triplet deactivation. The corresponding eigenvalues are: $\gamma_0 = 0$, $\gamma_1 = -k_{01} + k_{21}$, $\gamma_2 = -k_{02}$. The solution for the probability of ground state occupation from the instant $t = 0$ to t has the form:

$$p_0(t) = 1 - \frac{k_{01} - k_{02}}{k_{01} + k_{21} - k_{02}} \exp\left[-(k_{01} + k_{21})\, t\right] - \frac{k_{21}}{k_{01} + k_{21} - k_{02}} \exp\left[-k_{02}\, t\right], \tag{6}$$

The probability density function for the relaxation will take place at instant t can be obtained by differentiation of Eq. (6). On the introduction of the Bernouli parameter

$$p = \frac{k_{01} - k_{02}}{k_{01} + k_{21} - k_{02}}, \tag{7}$$

it is evident that the probability density function of times elapsed between successive photon arrivals is hyperexponential (see Eq. . (3)) and the kinetic scheme fits nicely to the IPP framework. Moreover, given estimates of the parameters in (3) or (4), one can easily obtain estimates of the parameters k_{01}, k_{02} and k_{21} involved in (6). For instance,

$$\begin{aligned} k_{21} &= (1-p)(\gamma_1 - \gamma_2) \\ k_{02} &= \gamma_2 \\ k_{01} &= p\gamma_1 + (1-p)\gamma_2 \end{aligned} \tag{8}$$

2.3 THE PARAMETER ESTIMATION PROCEDURES FOR A PHOTON-COUNTING DISTRIBUTION OF SINGLE MOLECULE FLUORESCENCE

In an experiment a single molecule is observed during time-interval $(0, \tau]$ and the number of emitted photons in each of the K sub-intervals $((k\tau/K,\ (k+1)\tau/K]$, with $k = 0,1, \ldots, K-1$, is registered. Note that such a scenario differs from K independent repetition of the elementary experiment. That is, if dwell period is much shorter than "off time", then the "off time" results in a series of consecutive dwell periods with 0 counts, while the sequence of independent repetitions yields just a single dwell period with 0 counts.

The likelihood function for the observed photon counts $(n_1, \ldots, n_K)$, assuming that the registration of photons begins when the first photon is registered can be written[7] as

$$l(\lambda_1, \omega_1, \omega_2) = v^T \left\{ \prod_{i=1}^{K} P(n_i, t) \right\} u, \tag{9}$$

where $u=(1,1)^T$, v^T is the steady-state probability vector and $\boldsymbol{P}(n_i, t)$ is defined in Eq. (2).

3. EXPERIMENTAL

The sample that consists of separated immobilised molecules was prepared by spin coating a droplet of polystyrene (M_W 44000, Aldrich) in toluene solution containing 5×10^{-10} M of DiI (1,1'-didocyl-3,3,3',3'-tetramethyl-indocarbocyanine perchlorate) on a cover glass. The resulting sample contains less than 0.2 molecules per 1μm^2. The 543.5 nm laser line provided by a CW He-Ne laser (Melles Griot 05-LPG-193), was used to excite the molecules. The excitation light was focused on the sample using 1.4 N.A. oil immersion objective. The power density on the diffraction limited light spot with diameter 300nm was approximately 140 W/cm^2. Fluorescence from a single molecule was collected with the same lens used for excitation and was imaged confocally on a single photon counting avalanche photodiode (SPCM AQ, EG&G) as a detector. The scattered laser light was removed by a combination of dichroic mirror and a notch plus filter.[4]

4. RESULTS

Figure 1 presents the observed counts of registered photons in consecutive dwell periods of duration 10ms. Across the whole time-scale a consistent pattern of many dwell periods with a very low (0-3) photon-counts can be observed. This background noise may result from a scattering of the light and a noise in the electronic equipment.

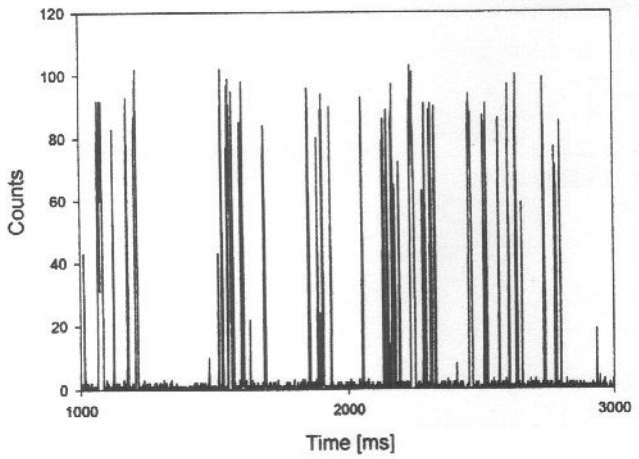

Figure 1: The fluorescence of a single DiI molecule measured at 10 ms temporal resolution.

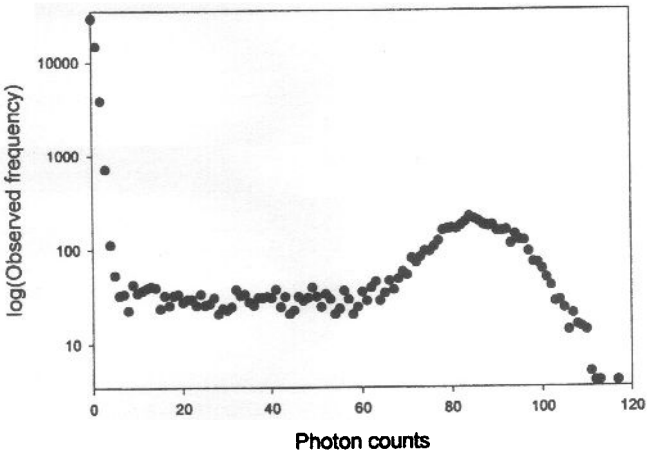

Figure 2: Photon-counting histogram obtained from the data in Figure 1.

We analyze the data from the [0,530s] interval using several different methods:

1. the method by Yip et al., considering consecutive "dwell periods" with the number of observed photon-counts smaller or equal (larger) than a particular threshold value as belonging to one "off" ("on") time; for comparison purposes, two threshold values were used: 20 and 40 photons;
2. assuming an underlying two-stage IPP;
3. using the method by Yip et al. with threshold values of 20 and 40 photons and subtracting 1, 2 or 3 counts from all observed photon-counts (putting 0 if the result of the subtraction is negative);
4. assuming an underlying two-stage IPP, subtracting 1, 2 or 3 counts from all observed photon-counts (putting 0 if the result of the subtraction is negative);
5. assuming an underlying two-stage MMPP;
6. using the theoretical form of the photon-count histogram proposed by Molski[6].

The first two scenarios ignore the presence of the "background noise". In scenarios 3-4, the analysis is informally adjusted for the "noise" by subtracting a small number of photons from the observed counts. Scenarios 5-6 allow for a formal correction for the "noise". This is because, as it has been mentioned earlier, a two-stage MMPP can be viewed as a sum of an IPP and an independent Poisson process.

Table 1 presents results of the analysis. In what follows we will assume the length of one dwell period (10ms) as the time unit.

In the analysis assuming an underlying MMPP scenario, the expected number of emitted photons during "off" time (λ_2) was estimated to equal 0.51 (standard error SE=0.003) time units. This value of λ_2 corresponds exactly to the estimate obtained in the analysis of the sub-period without fluorescence assuming a simple Poisson process (0.51, SE=0.003). Thus the value can be interpreted as an estimate of the background noise level. The expected number of emitted photons per dwell period during "on" time (λ_1) was estimated to equal 86.51 (SE=0.15). It follows that the expected number of emitted photons per dwell period during "on" time, in the absence of the "noise", can be estimated to equal 86 photons (SE=0.15). The expected durations of "on" and "off" times ($1/\omega_1$ and $1/\omega_2$, respectively) were estimated to equal 4.06 (SE=0.11) and 38.4 (SE=1.09) time units, respectively.

Table 1: Estimates of expected durations of "on" and "off" times and of the "on" time photon-emission, for various estimation methods.

Scenario	Estimation method	Counts	Expected duration		Expected emission "on" time (λ_1)
			"on" time ($1/\omega_1$)	"off" time ($1/\omega_2$)	
1	Yip *et al.* [20]	0	4.94 (0.14)	41.23 (1.22)	
1	Yip *et al.* [40]	0	4.77 (0.14)	44.71 (1.37)	
2	IPP	0	0.26 (0.002)	2.32 (0.02)	87.10 (0.22)
3	Yip *et al.* [20]	1	4.93 (0.14)	41.40 (1.22)	
3	Yip *et al.* [40]	1	4.76 (0.14)	44.96 (1.38)	
3	Yip *et al.* [20]	2	4.91 (0.14)	41.70 (1.24)	
3	Yip *et al.* [40]	2	4.75 (0.15)	45.35 (1.39)	
3	Yip *et al.* [20]	3	4.89 (0.14)	42.18 (1.26)	
3	Yip *et al.* [40]	3	4.73 (0.15)	45.94 (1.42)	
4	IPP	1	0.92 (0.01)	8.57 (0.12)	85.08 (0.15)
4	IPP	2	2.53 (0.06)	23.98 (0.54)	84.49 (0.14)
4	IPP	3	3.78 (0.10)	36.05 (1.00)	83.61 (0.14)
5	MMPP	0	4.06 (0.11)	38.40 (1.09)	86.00 (0.15)
6	Molski[6]	0	4.04 (0.11)	38.09 (0.96)	85.98 (0.16)

Note: "– Counts" indicates the reduction in the observed counts as adjustment for background noise. Standard errors are given in parentheses. The threshold used in the method of Yip *et al.* is indicated in squared brackets.

It is of interest to compare these results with those obtained under other scenarios. The results using the theoretical form of the photon-count histogram proposed by Molski[6] are very close to those obtained using the MMPP assumption. Note that the former in principle treat the observed histogram as arising from a series of independent dwell periods, thus ignoring the dependence between the periods induced by the observation scheme.

Under the assumption of an underlying IPP, ignoring the "background noise" gives a somewhat higher estimate of λ_1 (87.10, SE=0.22) than for MMPP. The estimates of the expected durations of "on" (0.26, SE=0.002) and "off" (2.32, SE=0.02) times are, on the other hand, dramatically smaller. The adjustment of the observed counts for background noise by subtracting from them 1, 2 or 3 photons (see Table 2) increases the estimates of the expected durations. However, the background noise adjustment has an opposite effect on λ_1 - it decreases the estimates of this parameter.

To apply the approach proposed by Yip et al.[5], 20 and 40 registered photons were chosen as threshold values for the purposes of the classification of dwell periods to "on" and "off" times respectively. As in the original paper by Yip et al., the choice was made based on the graph of the observed frequency of dwell periods with different numbers of registered photons. The analysis based on the approach proposed by Yip et al., whether without or with the adjustment for the background noise, yields larger estimates of the expected duration of "on" and "off" times than those obtained for an MMPP. The adjustment has a mild effect on the estimates. Irrespectively of the amount of the adjustment it slightly decreases the estimate of $1/\omega_1$ and increases the estimate of $1/\omega_2$.

Knowing the IPP parameters (λ_1, ω_1, ω_2) one can easily find the parameters of distribution of time between the registration of two consecutive photons by solving equations (4): $p = 0.997$, $\gamma_1 = 1/86.25$ and $\gamma_2 = 1/37.11$.

CONCLUSIONS

In this communication we proposed the formal approach to the estimation of parameters of a single molecule fluorescence photon-counting distribution. The proposed estimation procedure is based on the maximization of the likelihood function derived under the assumption of a two-stage Markov Modulated Process[7]. The physical basis for proposed approach consists of the theoretical scheme of double-resonance atom[2] and of the kinetic model for three-level dye molecule.

The approach proposed in this paper offers several advantages in comparison to the methods considered previously.[3,5] For instance, it allows to estimate all parameters of interest without specification of any arbitrary threshold values and avoiding a data manipulation. Moreover, due to the formal probabilistic description, the proposed technique allows to gain

more insight into the physical mechanism underlying the photon-counting distribution through the exploration of the correspondence between parameters of an IPP and of a hyperexponential distribution. It provides a natural way to adjust the analysis for the presence of background noise, which is important from a practical point of view. Finally, simulations (not reported here) indicate a very good performance of the procedure, in terms of a small bias of the estimators and of their standard errors, if a correct model (IPP in the absence and MMPP in the presence of noise) is used. Though the method is numerically involved, there exist several numerical optimization approaches that can be used for its implementation. The SAS-IML[13] program used for producing the results presented in the paper is available from the first author upon request.

ACKNOWLEDGMENTS

We thank J. Hofkens (Department of Chemistry, Katholieke Universiteit Leuven) for making the experimental data available to us and A. Haflidadóttir for providing her unpublished thesis and other valuable information.

REFERENCES

1. H. Dehmelt, "Proposed l0nu greater than nu laser fluorescence spectroscopy on T1+mono-ion oscillator II", *Bulletin of the American Physical Society*, **20**, 60-60, 1975.

2. R. J. Cook and H. J. Kimble, "Possibility of direct observation of quantum jumps" *Phys. Rev. Lett.* **54**, 1023-1026, 1985.

3. T.Ha, Th. Enderle, D.S. Chemla, P.R. Selvin and S. Weiss,"Quantum jumps of single molecules at room temperature", *Chem. Phys. Lett.* **271**, 1-5, 1997.

4. E. Novikov, J. Hofkens, M. Cotlet, M. Maus, F.C. De Schryver and Noël Boens, "A new analysis method of single molecule using series of photon arrival times: theory and experiment", *Spectrochmica Acta* Part A, 2001.

5. W. T. Yip, D. Hu, J. Yu, D. A. Vanden Bout and P.F. Barbara, "Classifying the photophysical dynamics of single- and multiple-chromophoric molecules by single molecule spectroscopy", Journal of Phys. Chem. A, **102**, 7564-7575, 1998.

6. A. Molski, "Photon-counting distribution of fluorescence from a blinking molecule", **324**, 301-306, 2000.

7. A. Haflidadóttir, "Estimation of IPP parameters based on counts" *Unpublished M.Sc. thesis*, Department of Mathematical Modeling at the Technical University of Denmark (http://www.imm.dtu.dk).

8. W. Fischer, and K. Meier-Hellstern, "The Markov-modulated Poisson process (MMPP) cookbook", *Performance Evaluation*, **18**, 149-171, 1993.

9. M. F. Neuts and J.-M. Li, "An algorithm for the $P(n,t)$ matrices of a continuous BMAP", In: *Matrix-analytic methods in stochastic models. Papers from the International Conference held in Flint*, MI. S.R. Chakravarthy and A.S. Alfa, (eds.). *Lecture Notes in Pure and Applied Mathematics*, **183**, pp. 7-19. New York: Marcel Dekker, Inc. 1997.

10. A. Kuczura, "The interrupted Poisson process as an overflow process", *Bell System Technical Journal*, **52**, 437-448, 1973.

11. J.E. Beyer, and B.F. Nielsen, "Predator foraging inpatch environments: the interrupted Poisson process (IPP) model unit", *Dana: Danish Journal of Fisheries and Marine Research*, 11, 65-130, 1996.

12. N.G. Van Kampen, Stochastic process in physics and chemistry, North Holland, 1992.

13. SAS Institute Inc., *SAS/IML Software: Changes and Enhancements trough Release 1.11*, Cary, NC: SAS Institute Inc., 1995.