Photon Counts Simulation in Fluorescence Fluctuation Spectroscopy

Shingaryov I. P.[1)*], Skakun V. V.[1)], Apanasovich V. V.[1)]
1) Belarusian State University, Nezavisimosty, 4, Minsk, 220050, Belarus
*Author to whom correspondence should be addressed. E-mail: shynharou@bsu.by

Abstract: *Developing of new data analysis models and methods requires comprehensive testing of their validity, accuracy and robustness. This can be done by means of simulation of an analyzed characteristic. In this work, we propose a simulation model of photon counts detection process in fluorescence fluctuation spectroscopy (FFS) experiments. This model can be used to obtain data to test data analysis models and methods in FFS.*

***Keywords*:** Fluorescence Fluctuation Spectroscopy, Photon Counting Histogram, Photon Counts.

1. INTRODUCTION

Fluorescence Fluctuation Spectroscopy (FFS) methods are widely used in modern biophysical and biochemical research [1]. In FFS, the information about dynamics, interactions, and structure of fluorescently labeled macromolecules is extracted from detected fluorescence intensity fluctuations. These fluctuations are produced by kinetic processes that alter the number and intrinsic fluorescence of molecules in the observation volume defined by focusing optics of the experimental setup. The detected fluorescence signal thus contains information about molecular diffusion, photophysical and chemical dynamics [2, 3, 4, 5].

There are various FFS methods for extracting this information. The commonly used methods are fluorescence correlation spectroscopy (FCS) [6], the photon counting histogram (PCH) analysis [7], the fluorescence intensity distribution analysis (FIDA) [8], and their modifications. FCS provides information about diffusion coefficients, chemical kinetics, exited-state molecular dynamics, picomolar concentrations and dynamics of the interaction of fluorescent molecules. It allows to resolve species by their molecular weight. PCH and FIDA are used to obtain information about the concentration and the specific brightness of molecules. In contrast to FCS, these methods allow to resolve species by their specific brightness.

Developing of new data analysis models and methods requires comprehensive testing of its validity, accuracy, resolvability, and robustness. This can be done by simulation of analyzed characteristic with given parameter values and signal-to-noise (S/N) ratio and subsequent application of an analysis method to the simulated data.

In this work, we establish an approach to simulate the process of photon detection in the FFS system. This allows to get a photon counting distribution (PCD) with given parameters and S/N ratio.

2. THEORY

In thermodynamic equilibrium, the main processes affecting fluorescence intensity fluctuations are diffusion of molecules in the observation volume with the inhomogeneous spatial distribution of excitation energy and changes in the molecule quantum yield. Although molecules can be excited in each point of the laser beam propagation path in the sample, the major number of photons is detected from the molecules situated in the focus of the registration optics.

Fluorescence intensity fluctuations are defined as

$$\delta F(t) = F(t) - < F(t) >, \tag{1}$$

where $F(t)$ and $<F(t)>$ is the measured and the temporal average fluorescence intensity respectively [9].

Assuming that the intensity of the excitation light is constant and all fluctuations arise only from changes in the local concentration $\delta C(r,t)$ the fluorescence intensity fluctuations can be expressed as

$$\delta F(t) = \kappa\sigma\eta_f \int_V I(r)S(r)\delta C(r,t)dr, \tag{2}$$

where κ – the overall detection efficiency, σ – the molecular absorption cross-section, η_f – the fluorescence quantum efficiency, $I(r)$ – the spatial distribution of the excitation energy with the maximum amplitude I_0, $S(r)$ – the optical transfer function, which describes the spatial collection efficiency of the setup [9].

The functions $I(r)$ and $S(r)$ can be combined into a single function $W(r) = I(r)S(r)/I_0$ called a brightness profile. It describes the spatial distribution of the emitted light. Often, it is approximated by a three-dimensional Gaussian, which is decayed to $1/e^2$ at ω_0 in lateral and z_0 in axial direction [10]:

$$B(r) = B_0 \exp(-2\frac{(x^2+y^2)}{\omega_0^2} - 2\frac{z^2}{z_0^2}), \tag{3}$$

where $B_0 = B(0)$.

Parameters κ, σ, η_f can be combined with I_0 to give the specific brightness $q = I_0\sigma_a\eta_f\kappa$ that determines the photon count rate detected from a molecule per second.

To estimate the concentration of molecules the effective volume is defined as [6]:

$$V_{eff} = \frac{\psi_1^2}{\psi_2}, \tag{4}$$

where

$$\psi_k = \int_V W^k(r)dr. \tag{5}$$

Here V is the total volume of the sample. For practical usage, the integration limits in (5) can be changed to infinity because the size of the illuminated region is much smaller than the size of a reservoir containing a sample [7].

For the Gaussian approximation (3):

$$\psi_{Gk} = \int_V B^k(r)dr = B_0^k\left(\frac{\pi}{2k}\right)^{3/2}\omega_0^2 z_0 \tag{6}$$

and

$$V_{Geff} = \pi^{3/2}\omega_0^2 z_0 . \quad (7)$$

Analysis of the Photon Counting Distribution

The photon counting distribution (PCD) defines the probability *P(n)* of detecting *n* photons during a sampling time *T*. The generating function (GF) of the probabilities *P(n)* can be written as [8]:

$$\theta(\xi) = \exp\left\{\sum_{i=1}^{M} \overline{C_i} \int_V \left(\exp\left((\xi-1)q_i TW(r)\right)-1\right)dr\right\}, \quad (8)$$

where $\overline{C_i}$ and q_i is the average concentration and the specific brightness of the *i-th* species correspondingly, *M* is the number of species in a sample.

The relation between the PCD and the GF is [11]:

$$\theta(\xi) = \sum_{n=0}^{\infty} P(n)\xi^n , \quad (9)$$

where ξ is the auxiliary variable and

$$P(n) = \frac{1}{n!}\frac{d^n\theta(\xi)}{d\xi^n}\bigg|_{\xi=0} . \quad (10)$$

In the PCH analysis, different molecules are assumed to have independent coordinates, and thus contribute independently into the fluorescence signal. The duration of the photon-counting interval is assumed to be short compared to the typical diffusion time of the fluorescent molecules through the observation volume. The light intensity emitted by a molecule at a certain position $\boldsymbol{r}$ is expressed as a product of its specific brightness *q* and the brightness profile *W(r)* only [8]. Therefore, the number of photons detected from a molecule per sampling time *T* is

$$n = qTW(r) . \quad (11)$$

It was demonstrated that deviation of the Gaussian approximation from the true brightness profile causes the PCH model to fail under certain conditions, which could lead to a complete misinterpretation of the data [12]. In the analytical expression for the PCD, there are no parameters that can be adjusted to compensate this deviation [7]. To overcome this, correction factors were proposed [13]:

$$F_k = \frac{\psi_k - \psi_{Gk}}{\psi_{Gk}}, k = 1,2,... \quad (12)$$

These correction factors characterize the contribution of detected photons from the non-Gaussian part of the brightness profile into the fluorescence signal.

The photon counting GF with correction factors was obtained by expanding the exponent function in (8) into a Taylor series [14]:

$$\theta(\xi) = \exp\left\{\sum_{i=1}^{M} \overline{C_i} \sum_{k=1}^{\infty} \frac{(\xi-1)^k q_i^k T^k}{k!}\psi_k\right\} . \quad (13)$$

Using (12) and (13) the expression for the GF with correction factors can be written as a product of two GFs:

$$\theta(\xi) = \theta_G(\xi)\theta_c(\xi) , \quad (14)$$

where

$$\theta_G(\xi) = \exp\left\{\sum_{i=1}^{M} \overline{C_i} \int_V \left(\exp\left((\xi-1)q_i TB(r)\right)-1\right)dr\right\} \quad (15)$$

and

$$\theta_c(\xi) = \exp\left\{\sum_{i=1}^{M} \overline{C_i} \sum_{k=1}^{\infty} \frac{(\xi-1)^k q_i^k T^k F_k}{k!}\psi_{Gk}\right\} . \quad (16)$$

Therefore, the additional photons detected from the non-Gaussian part of *W(r)* are taken into account by $\theta_c(\xi)$.

Statistical noise of the PCD can be approximated by the Binomial distribution [7]

$$P_b(k) = \frac{m!}{k!(m-k)!}P_n^k(1-P_n)^{m-k}, k = 0,1,...m , \quad (17)$$

where P_n is the probability of detecting *n* photons during the sampling time *T*, *m* is the number of sampling intervals. Thus, for each point *n* of the PCD the mean μ_n and the variance σ_n^2 are calculated as $\mu_n = mP_n$ and $\sigma_n^2 = mP_n(1-P_n)$ respectively.

Signal-to-noise ratio is usually defined as $S/N = \mu/\sigma$ [15]. Here we will use the estimation of *S/N* as the signal-to-noise ratio at the maximum of the PCD

$$S/N = \sqrt{mP_{\max}/(1-P_{\max})} , \quad (18)$$

where $P_{\max} = \max_i(P_i)$. Therefore, the number of sampling intervals defines *S/N* ratio for a given P_{max}.

The analysis of the PCD is based on the fitting the theoretical PCD to the experimental one by the least-squares method [7]. The Marquardt method is commonly used as an optimization procedure [16]. The quality of the fit is verified by the reduced Chi-criterion value and normalized residuals of the fit [7, 16]:

$$\chi_r^2 = \frac{\sum_{k=0}^{n}\left(M\frac{P_k - E_k}{\sigma_k}\right)^2}{n-d-1} \quad (19)$$

and

$$R_k = M\frac{P_k - E_k}{\sigma_k} , \quad (20)$$

where *d* – the number of fitting parameters, P_k – the theoretical probability to detect *k* photons, E_k – the statistical frequency to detect *k* photons, *n* – the number of bins, *M* – the number of trials, σ_k – the standard deviation of the *k*-th PCD point.

3. PHOTON COUNTS SIMULATION

There are various processes affecting the fluorescent signal such as photobleaching, saturation of molecules, and triplet-state kinetics. In many experimental situations, the influence of these processes on the PCD is negligible and only translational diffusion and fluorescence of molecules must be taken into account.

From statistical physics, it is known that the probability of finding *k* non-interacting molecules in an open volume ΔV is well approximated by the Poisson distribution with the mean $\overline{N} = \overline{C}\Delta V$. The Poisson distribution with the mean $\overline{n} = qTB(r)$ also approximates the probability of detecting *n* photons from a molecule at the position $\boldsymbol{r}$ during a sampling time *T* [8].

Using these assumptions the following photon counts simulation scheme can be proposed for the case of a sample containing *i*-species of molecules. At the first

stage, it is necessary to get the number of molecules N_i of the *i*-th species by generating a Poisson random variable with the mean $\overline{N}_i$ and then uniformly distribute them in the observation volume V_0. After that, a Poisson random variable n_{ij} with the mean $\overline{n_{ij}} = q_i TB(r_{ij})$ is generated to get the number of detected photons from the *j*-th molecule of the *i*-th species, $j = 1,2...N_i$, $i = 1,2,...,M$, where *M* is the number of species. Then the numbers of generated photons are summarized to get the total number of photons S_T detected during the sampling time *T* from all molecules present in the observation volume V_0. These stages are repeated *m* times to reach a given S/*N* ratio. Consequently, the series of photon counts is obtained and the PCD can be build.

We used the normalization to the effective volume V_{Geff} commonly used in FCS [6]. In this case, the initial parameters of the model will be the average number of molecules for the *i*-th species $\overline{N}_{effi}$ in the effective volume V_{Geff}, the specific brightness q_i, the sampling time *T*, and the number of sampling intervals *m*.

The Gaussian approximation (3) can be reduced to the one-dimensional form using the coordinate transformation:

$$\begin{cases} x = r\omega_0 \cos\varphi \sin\alpha, \\ y = r\omega_0 \sin\varphi \sin\alpha, \\ z = rz_0 \cos\alpha, \end{cases} \quad (21)$$

where $r \in [0;\infty)$, $\alpha \in [0;\pi]$, $\varphi \in [0;2\pi)$.

Then the expression (3) takes form

$$B(r) = \exp(-2r^2), \quad (22)$$

where $r^2 = x^2 + y^2 + z^2$ is the length or the radius-vector ***r***. Here we set $B_0 = 1$ as it is assumed in the PCH analysis.

The volume V_0 defined by the radius-vector $\mathbf{r}_0$ in these coordinates equals to:

$$V_0 = 4\pi r_0^3 \omega_0^2 z_0 / 3. \quad (23)$$

As the average concentration of the *i-th* species is

$$\overline{C_i} = \frac{N_{effi}}{V_{Geff}} = \frac{\overline{N_i}}{V_0}, \quad (24)$$

the average number of the *i*-th species of molecules in the volume V_0 will be

$$\overline{N}_i = \frac{N_{effi} V_0}{V_{Geff}} = N_{effi} \frac{4r_0^3}{3\sqrt{\pi}}. \quad (25)$$

The value of r_0 can be defined from the condition

$$\int_0^{r_0}\int_0^{2\pi}\int_0^{\pi} B(r,\alpha,\varphi) J dr d\varphi d\alpha \Big/ \int_V B(r)dr \to 1. \quad (26)$$

To uniformly distribute molecules in the V_0 at the first step it is necessary to obtain the coordinates *x*, *y*, *z* of each molecule uniformly distributed in [$-r_0$; r_0]. At the second step, the fulfillment of the condition $\left(x^2 + y^2 + z^2\right)^{1/2} \le r_0$ is verified, and, if it is not true, the first step is repeated again.

Photon counts simulation with the brightness profile correction

In many real experimental conditions, it is sufficient to consider only the first-order correction F_1 ($F_j = 0$, $j = 2,3...$) [13]. In this case, the initial parameter B_0 must be taken into account because correction (12) affects the total distribution of fluorescence emission. It was shown that brightness becomes a function of the correction parameter F_1 [14]

$$q_c = B_0 q = (1 + F_1) q, \quad (27)$$

where q_c is the specific brightness with correction.

Then the expression (16) can be written as

$$\theta_c(\xi) = \exp\left\{ \sum_{i=1}^{M} 2^{-3/2} \overline{C_i} q_{ci} T F_1 \pi^{3/2} \omega_0^2 z_0 (\xi - 1) \right\}. \quad (28)$$

The GF of the Poisson distribution is [11]:

$$\theta_{pois}(\xi) = \exp\{\Lambda(T)(\xi - 1)\}. \quad (29)$$

Comparison (28) and (29) indicates that $\theta_c(\xi)$ is the GF of the Poisson distribution with the mean

$$\Lambda(T) = \sum_{i=1}^{M} 2^{-3/2} \overline{C_i} q_{ci} T F_1 \pi^{3/2} \omega_0^2 z_0. \quad (30)$$

This fact allows to propose an algorithm for photon counts simulation with the corrected brightness profile.

Taking into account (12) the volume V_{Geff} can be transformed to

$$V_{ceff} = (1 + F_1)^2 \pi^{3/2} \omega_0^2 z_0. \quad (31)$$

The expression for calculation of the average number of molecules of the *i*-th species in the observation volume V_0 takes the form:

$$\overline{N}_{ci} = N_{ceffi} \frac{4r_0^3}{3\sqrt{\pi}(1 + F_1)^2}, \quad (32)$$

where N_{ceffi} is the number of molecules in the effective volume V_{ceff}. Using (24), the expression (30) can be changed to

$$\Lambda(T) = \sum_{i=1}^{M} \frac{N_{ceffi} q_{ci} T F_1}{2\sqrt{2}(1 + F_1)^2}. \quad (33)$$

Therefore, the initial parameters of the model are N_{ceffi}, q_{ci}, *T*, *m*, and F_1. In this case, the simulation is performed according to the scheme proposed earlier. To take into account the additional number of photons emitted from the non-Gaussian part of the brightness profile a Poisson random variable with the mean $\Lambda(T)$ is generated. This additional number of photons is added to the total number of photons S_T.

4. TEST EXPERIMENTS

To confirm that our model correctly reproduces the features of real FFS experiments we made several simulations for the one- and two-component system. For each system case, we considered the cases of model with and without the brightness profile correction. The corresponding PCDs were built and analyzed using the PCH analysis with and without correction to determine if the theoretical models could fit the simulated data. That would indicate the adequacy of our model.

The sampling time in all simulations was set to $T=5*10^{-5}$ s. The asymptotic standard errors were used as an error measure of the fit parameter estimations [16].

Photon counts simulation for the one-component system

Simulations were performed with and without the brightness profile correction. In the case of simulation without correction we set the initial values of the model parameters as $N_{eff}=2$, $q=60000$, and the total number of sampling intervals as $m=10^6$ (that corresponds to the observation time in 50 s, S/N≈600). The estimations $\widetilde{N}_{eff}=2.005\pm0.009$, $\widetilde{q}=59800\pm300$ were obtained by the PCH analysis without correction. The obtained estimations are close to the initial values of the parameters. The chi-criterion value ($\chi_r^2=1.48$) and the shape of the normalized residuals (not shown) indicates a good fit quality. This confirms that our model correctly reproduces the main processes of photon detection.

In the case of simulation with the brightness profile correction we set the initial values of the model parameters as $N_{eff}=2$, $q=84000$, $F_1=0.4$. The simulated data were analyzed by the PCH analysis with and without correction. As it was expected, in the case of applying the PCH analysis without correction, the obtained estimations are not acceptable ($\chi_r^2\approx286$). The estimations obtained by the PCH analysis with correction $N_{eff}=1.99\pm0.006$, $q_c=84000\pm400$, $F_1=0.393\pm0.005$ are close to the initial values of the model parameters. The simulated and the theoretical PCDs and their normalized residuals are shown in Fig. 1.

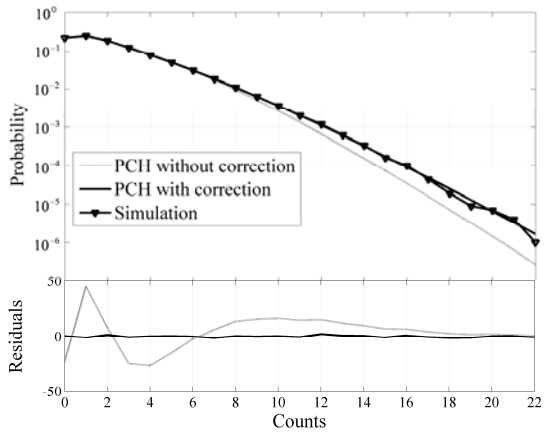

Fig. 1 – The simulated PCD for the one-component system and the result of the fit by the PCH analysis with and without correction

The chi-criterion value ($\chi_r^2=0.94$), the normalized residuals, and proximity of the obtained estimations to the initial values of the model parameters proofs the adequacy of the simulation in the case when additional photons emitted from the non-Gaussian part of the brightness profile are taken into account.

Photon counts simulation for the two-component system

The total number of sampling intervals in simulations was set to $m=6*10^6$ that corresponds to the total observation time in 300 s and S/N≈1300.

The initial values of model parameters in the case of simulation with the brightness profile without correction were $N_{eff1}=1$, $q_1=60000$, $N_{eff2}=5$, $q_2=20000$. The estimations obtained by the PCH analysis without correction $\widetilde{N}_{eff1}=1.02\pm0.09$, $\widetilde{q}_1=60000\pm1000$, $\widetilde{N}_{eff2}=4.99\pm0.03$, $\widetilde{q}_2=19900\pm700$ are close to the initial values and the quality of the fit ($\chi_r^2=0.77$) is acceptable.

In the case of simulations with the corrected brightness profile, the data were analyzed by the PCH model with and without correction (see the fit results in Table 1 and Fig. 2).

Table 1. The initial values and the estimations of the model parameters in the case of simulation for the two-component system with the corrected brightness profile

Parameter	Initial value	PCH without correction	PCH with correction
N_{ceff1}	1	0.83±0.07	1.05±0.1
q_{c1}	84000	75000±2000	83000±5000
N_{ceff2}	5	9.9±0.7	5.2±0.6
q_{c2}	28000	10000±1000	26000±6000
F_1	0.4	-	0.37±0.08

As it was expected in the first case the estimations were not in good agreement with the initial values of the model ($\chi_r^2=1.78$) as the ratio $\widetilde{N}_{ceff2}/\widetilde{N}_{ceff1}\approx12$ while the ratio of initial parameters is only 5. By the PCH analysis with correction we obtained the acceptable fit results ($\chi_r^2=1.26$) that confirmed our approach for the multicomponent systems simulation.

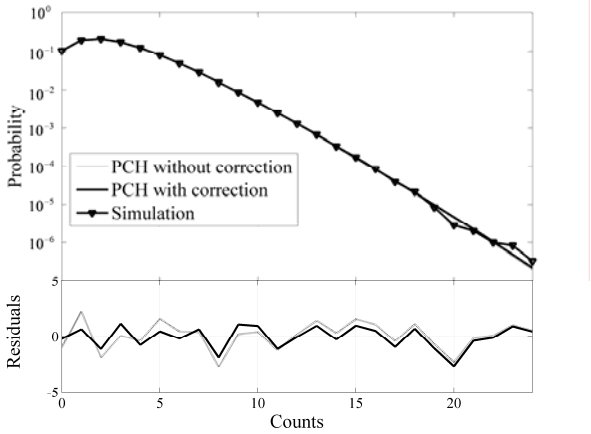

Fig. 2 – The simulated PCD for the two-component system and the result of the fit by the PCH analysis with and without correction

5. CONCLUSION

We have developed a simulation approach of the photons detection process in the FFS experiments for multicomponent molecular systems. The received estimations of the model parameters, the chi-criterion

values, and the normalized residuals have indicated the adequacy of our model in all considered cases.

6. REFERENCES

[1] S. Hess, S. Huang, A. Heikal, W. Webb. Biological and Chemical Applications of Fluorescence Correlation Spectroscopy: A Review, *Biochemistry* 41 (3) (2002) p. 697-705

[2] E. Elson, H. Qian. Interpretation of fluorescence correlation spectroscopy and photobleaching recovery in terms of molecular interactions, *Methods in Cell Biology* 30 (1989). p. 307

[3] S. Maiti, U. Haupts, W. Webb. Fluorescence correlation spectroscopy: diagnostics for sparse molecules, *Proc. Natl. Acad. Sci. U. S. A.* 94 (1997). p. 11753

[4] T. Starr, N. Thomson. Total internal reflection with fluorescence correlation spectroscopy: combined surface reaction and solution diffusion, *Biophysical Journal* 80 (2001). p. 1575

[5] E. Hom, A. Verkman. Analysis of Coupled Bimolecular Reaction Kinetics and Diffusion by Two-Color Fluorescence Correlation Spectroscopy: Enhanced Resolution of Kinetics by Resonance Energy Transfer, *Biophysical Journal* 83 (2002). p. 533

[6] O. Krichevsky, G. Bonnet. Fluorescence Correlation Spectroscopy: the Technuque and its applications, *Reports on Progress in Physics* 65 (2002). p. 251-297

[7] Y. Chen, J. Müller, P. So, E. Gratton. The Photon Counting Histogram in Fluorescence Fluctuation Spectroscopy, *Biophysical Journal* 77 (1999). p. 553-567

[8] P. Kask, K. Palo, D. Ullmann, K. Gall Fluorescence-intensity distribution analysis and its application in biomolecular detection technology, *Proc. Natl. Acad. Sci. U. S. A.* 96 (24) (1999). p. 13756-13761

[9] P. Schwille, F.-J. Meyer-Almes, R. Rigler. Dual-Color Fluorescence Cross-Correlation Spectroscopy for Multicomponent Diffusional Analysis in Solution, *Biophysical Journal*, 72 (1997). p. 1878-1886

[10] R. Rigler, Ü. Mets, J. Widengren, P. Kask. Fluorescence correlation spectroscopy with high rate and low background: analysis of translational diffusion, *European Biophysical Journal,* 22 (1993). p. 169-175

[11] I. Bolshakov, V. Rakoshits. Applied theory of stochastic flows. Moscow, 1978. (in Russian)

[12] T. Perroud, B. Huang, M. Wallace, R. Zare. Photon counting histogram for one-photon excitation, *ChemPhysChem* 4 (2003). p. 1121-1123

[13] B. Huang, T. Perroud, R. Zare. Photon Counting Histogram: One-Photon Excitation, *ChemPhysChem* 5 (2005). p. 1523-1531

[14] V. Skakun, V. Apanasovich. Photon counting distribution analysis with out-of-focus-emission correction in monomolecular fluctuation spectroscopy, *Vestnik BGU*, Series 1. Physics, Mathematics, Informatics (2) (2008). p. 31-35 (in Russian)

[15] D. Koppel. Statistical accuracy in fluorescence correlation spectroscopy, *Phys. Rev. A.* 10 (1974). p. 1938-1945.

[16] P. Bevington, D. Robinson. *Data Reduction and Error Analysis for the Physical Sciences*. New-York: McGraw-Hill, 2003.