Fast, flexible algorithm for calculating photon correlations

Ted A. Laurence

Physical Biosciences Institute, Chemistry and Materials Science Directorate, Lawrence Livermore National Laboratory, Livermore, California 94550

Samantha Fore and Thomas Huser

NSF Center for Biophotonics Science and Technology, University of California, Davis, 2700 Stockton Boulevard, Suite 1400, Sacramento, California 95817

Received October 24, 2005; accepted November 11, 2005; posted December 1, 2005 (Doc. ID 65533)

We introduce a new algorithm for computing correlations of photon arrival time data acquired in single-molecule fluorescence spectroscopy and fluorescence correlation spectroscopy (FCS). The algorithm is based on rewriting the correlation as a counting operation on photon pairs and can be used with arbitrary bin widths and spacing. The flexibility of the algorithm is demonstrated by use of FCS simulations and single-molecule photon antibunching experiments. Execution speed is comparable to the commonly used multiple-tau correlation technique. Wide bin spacings are possible that allow for real-time software calculation of correlations, even for high count rates. © 2006 Optical Society of America

OCIS codes: 170.6280, 180.1790, 300.2530.

In single-molecule fluorescence spectroscopy and fluorescence correlation spectroscopy (FCS), the most flexible data-acquisition methods record the arrival time of each detected photon and perform analysis of the resultant photon streams.[1–4] One of the most common data analysis operations performed is that of the temporal correlation function used in FCS.[5] Although they are potentially time consuming, appropriate algorithms such as the multiple-tau correlation technique[6] allow this operation to be performed quickly.

We demonstrate a new correlation algorithm that, in contrast to the multiple-tau correlation technique, allows for bins to be spaced in any way desired. We show two examples of data analysis that use this flexibility. First, by using a wider spacing of histogram bins, one may perform online calculations of correlations at higher count rates than were previously possible. Second, we calculate correlations for a photon antibunching experiment that uses pulsed excitation. Our new algorithm allows the nanosecond-scale details of the correlation to be calculated, even at large time lags. This algorithm may be adapted to extensions of the correlation such as the PAID (photon arrival-time interval distribution) histogram.[4]

The temporal correlation function is $C_{AB}(\tau) = \langle I_A(t) I_B(t+\tau)\rangle / \langle I_A(t)\rangle \langle I_B(t+\tau)\rangle$, where $I_A(t)$ and $I_B(t)$ are detected intensities and t and τ are continuous time and time-lag variables.

The data recorded for photon-timing single-molecule fluorescence spectroscopy–FCS experiments are series of photon time stamps with time resolution Δt. t_i is the arrival time of the ith photon from channel A, and u_j is the arrival time of the jth photon from channel B. Assuming stationarity, the ensemble averages in the expression for $C_{AB}(\tau)$ are converted to averages over all time. Averaging over a finite experimental time T with N_A and N_B photons detected in the respective channels gives the correlogram $\hat{C}_{AB}(\tau)$.

In terms of discrete photon time stamps t, $I_A(t)$ is the number of photons i such that $t=t_i$, or $I_A(t) = n(\{i \ni t_i = t\})/\Delta t$. $\{i \ni t_i = t\}$ is the set of all photons i such that $t_i = t$, and the operator n counts the number of elements in the set. Similarly, $I_B(t) = n(\{j \ni u_j = t\})/\Delta t$. In this notation, using discrete time lag τ yields

$$\hat{C}_{AB}(\tau) = \frac{n(\{(i,j) \ni t_i = u_j - \tau\})(T-\tau)}{n(\{i \ni t_i \leqslant T-\tau\})n(\{j \ni u_j \geqslant \tau\})}. \qquad (1)$$

The restrictions on the average intensities in the denominator are for symmetric normalization.[7]

Equation (1) changes the correlogram calculation to a counting task. To compute $\hat{C}_{AB}(\tau)$ at full resolution Δt, one can iterate through each pair of photons (i,j), calculating the time lag $\tau = u_j - t_i$ and adding 1 to the correlogram $\hat{C}_{AB}(\tau)$. This procedure is slow, requiring $N_A N_B$ operations.

The full time resolution is not needed or wanted in FCS measurements for the entire range of time lags (for $\Delta t = 10$ ns, 10^8 elements are needed for calculating time lags up to $\tau = 1$ s). In the multiple-tau algorithm,[6] only the first 16 time lags are calculated at the full resolution. The time resolution is then halved by addition of the number of photons in adjacent bins, and eight more time lags are calculated. This process is repeated until the desired dynamic range is reached. This algorithm, which was originally developed for hardware correlators, may also use photon timing data.[8,9] The process of successively reducing the time resolution increases the amount of triangular averaging with increasing time lags.[8,10] The bins are chosen to keep this effect below 0.1%, preventing the use of wider bin spacings. The photon-mode correlation of Ref. 8 has triangular averaging only over time scale Δt, but the algorithm is performed efficiently only for short time lags.

0146-9592/06/060829-3/$15.00

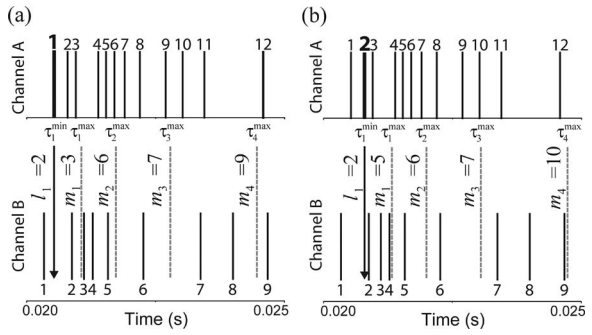

Fig. 1. Calculating the cross correlation of photon sequences in two channels, A and B (photon times shown as vertical solid lines), by using our new algorithm. The maximum and minimum limits for M bins in time lag τ $([\tau_1^{\min},\tau_1^{\max}),[\tau_2^{\min},\tau_2^{\max}),\ldots,[\tau_M^{\min},\tau_M^{\max}))$ are added to the arrival time of each photon in channel A. (a) These limits (arrows for $\tau_1^{\min}$; lighter dotted lines for other limits) for photon 1 from channel A; (b) these limits for photon 2. In this example, $M=4$, $\tau_1^{\min}=0$, $\tau_2^{\min}=\tau_1^{\max}$, $\tau_3^{\min}=\tau_2^{\max}$, and $\tau_4^{\min}=\tau_3^{\max}$. To the left of each limit in (a) is index j of the photon in channel B, such that $u_j>t_1+\tau_{\text{curr}}$, where, in turn, $\tau_{\text{curr}}=\tau_1^{\min}$, $\tau_{\text{curr}}=\tau_1^{\max}$, etc. Similar indices j are shown in (b). The contribution to correlogram Y_k for each time-lag bin k is m_k-l_k. In going from photon 1 to 2 in channel A [comparing (a) and (b)] one makes only small adjustments in the values of l_k and m_k.

As we now show, it is not necessary to use the rigid bin spacings of the multiple-tau algorithm to calculate the correlogram efficiently. We specify M bins with arbitrary limits in time lag τ_k, say, $[\tau_1^{\min},\tau_1^{\max}),[\tau_2^{\min},\tau_2^{\max}),\ldots,[\tau_M^{\min},\tau_M^{\max})$. The algorithm is implemented as follows:

(1) Initialize a correlogram Y_k with M bins to 0.
(2) Starting with the first photon, t_1, find l_k such that $u_{l_k-1}<t_1+\tau_k^{\min}\leqslant u_{l_k}$ and m_k such that $u_{m_k-1}<t_1+\tau_k^{\max}\leqslant u_{m_k}$ for all $k=1\ldots M$ (a linear search is adequate). l_k and m_k are the photon indices in channel B that correspond, respectively, to the minimum and maximum time-lag limits of bin k.
(3) For every k, add m_k-l_k to Y_k.
(4) Repeat steps (2) and (3) for $t_2\ldots t_{N_A}$. The important time-saving tactic lies in this step. Bin limits l_k and m_k are only slightly adjusted when one goes from t_1 to t_2 in step (2) (Fig. 1). Hence it is possible to use the previously found l_k and m_k as starting points to find new values that satisfy the conditions in step (2) for t_2.
(5) Normalize according to Eq. (1).

In step (4) the boundaries of the M time-lag bins are in general shifted only by a small amount (the time between two successive photons in channel A) when one goes from the first to the second photon in channel A (Fig. 1). The number of operations then scales as N_AM rather than N_AN_B operations, dramatically reducing the computational cost. Triangular averaging[8,10] occurs over time resolution Δt; rectangular averaging occurs over width $\tau_k^{\max}-\tau_k^{\min}$ of bin k.

We compare the performance of our new algorithm with the multiple-tau algorithm, using 10 s simulations of fluorescent molecules diffusing through a confocal detection volume.[4] Each molecule's average count rate in the volume is 35.4 kHz. The average molecular occupancy c in the volume is 0.1, 1.0, and 10.0 in three simulations, leading to total average count rates of 3.54, 35.4, and 354 kHz, respectively. Figure 2 shows example correlations and curve fits. Table 1 compares execution times for our new algorithm and the multiple-tau algorithm (on a 2 GHz Pentium M computer). In the second and third columns, the bins are dictated by the multiple-tau algorithm, with τ from 1 μs to 2 s. At low count rates the execution times are similar. At high count rates (354 kHz) the multiple-tau algorithm is 40% faster.

For the highest count rate (bottom row in Table 1), neither algorithm can compute the correlation within 10 s, the duration of the simulated experiment, preventing real-time correlation calculations. However, with wider bins (two per decade from 1 μs to 1 s; fourth column), our algorithm can keep up with a count rate of 354 kHz. Triangular averaging occurs only over time resolution Δt, permitting the use of widely spaced bins that are problematic with the multiple-tau method.[8,10]

We are currently performing measurements that attempt to quantify the number of independent fluorescent quantum emitters by using fluorescence photon antibunching[11–13]; a detailed description of the setup is given in Ref. 13. We measure fluorescence from single oligonucleotides labeled with the dye Atto

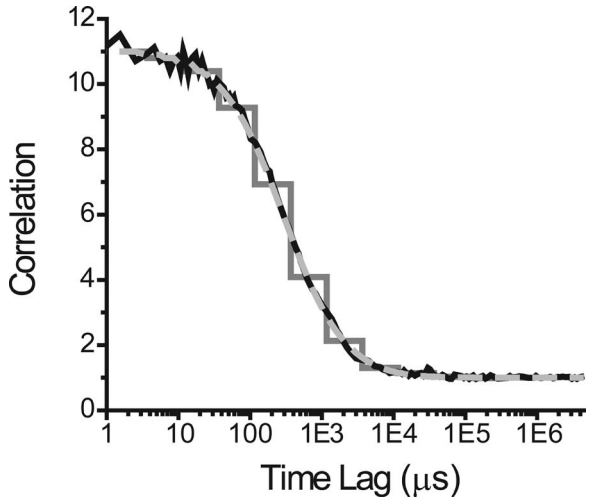

Fig. 2. Correlations and fits for simulated data of fluorescent molecules diffusing through a Gaussian detection volume. In this case, $c=0.1$, and diffusion time τ_D through the detection volume is 300 μs. The correlation is fitted to $C(\tau)=1+1/[c(1+\tau/\tau_D)\sqrt{1+\tau/(25\tau_D)}]$. The correlation calculated with our new algorithm with quasi-logarithmic bin spacings from the multiple-tau correlation algorithm (8 bins per octave, 160 total bins) is shown in black. The dotted light gray curve is a fit with $c=0.099\pm0.002$ and $\tau_D=294\pm3$ μs. The dark gray curve is the correlation with wide bin spacings (2 bins per decade, 12 total bins). A fit (not shown) recovered the values $N=0.10\pm0.01$ and 300 ± 100 μs. The wider spacing sacrifices some accuracy but increases speed of computation.

Table 1. Comparison of Correlation Execution Times

Molecular Occupancy c	Multiple Tau (s)	New Algorithm (s)	Wide Bin (s)
0.1	0.21	0.21	0.07
1.0	1.9	2.1	0.44
10.0	16	23	4.0

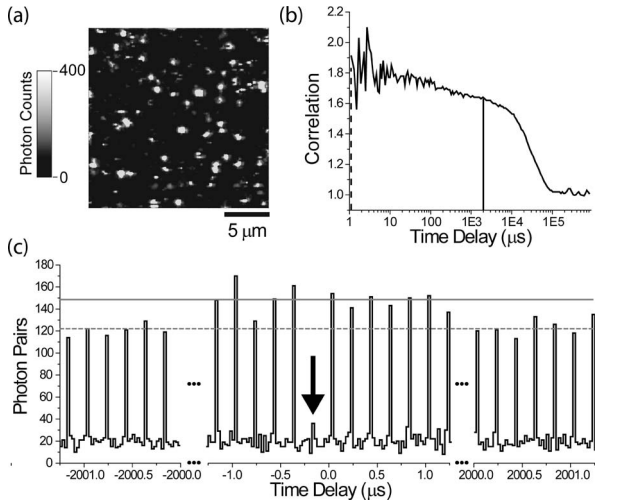

Fig. 3. Photon antibunching experiments performed on fluorescently labeled DNA oligomers attached to a glass surface. (a) Image of single surface-immobilized DNA oligomers; image size, 20 μm $\times$ 20 μm. (b) Long-time-scale correlation over the entire image (bin spacings similar to those in Fig. 2). (c) Short-time scale correlations with linearly spaced bins (25 ns bins). The nonnormalized cross correlation for the region with time lags from −1.25 to 1.25 μs is shown [central plot in (c), dotted curve in (b)], along with the regions within 1.25 μs of the ±2 ms time lags [side plot in (c), solid line in (b)]. The spikes in (c) correspond to pulses from the excitation laser. The solid horizontal line is the average peak height in the region −1.25 to 1.25 μs, and the dotted horizontal line is the average spike height in the regions within 1.25 μs of ±2 ms time-lag regions. The peak at −0.2 μs with low photon pair counts is due to photon antibunching (dark arrow). It is not at 0 time lag owing to a cable delay (150 ns) and a software-adjusted digital delay (37.5 ns).

655 (Atto-tec GmbH) immobilized on a coverslip (Fig. 3). For excitation, we drive a picosecond pulsed laser (LDH 635-B, PicoQuant GmbH) with an external 5 MHz clock generated by a counter–timer board (PCI-6602, National Instruments). The emission is split to form a Hanbury–Brown–Twiss interferometer with two avalanche photodiode detectors (SPCM-AQR-14, Perkin-Elmer). Detected photons are timed on two channels by an 80 MHz clock, which also drives the 5 MHz laser trigger. Using pulsed excitation and calculating cross correlations between the two channels is an efficient way to observe photon antibunching.[11,13,14]

Figure 3(b) shows the correlation of all photon events obtained during the image scan calculated with bins similar to those in Fig. 2 (the spacing between bins is too large for the pulsed nature of the excitation to be observed). The faster fluctuations at the ~100 μs time scale are due to fluorophore triplet-state blinking.[15] The longer decay at the ~30 ms time scale is the result of scanning the sample.[16] Figure 3(c) reveals the pulsed nature of the excitation, zooming in on the correlation in Fig. 3(b) around two time lags. Our algorithm efficiently calculates the correlation with 25 ns time resolution whether the time lags are near 0 (3.7 s for 800,000 photons on a 2 GHz Pentium M computer) or ±2 ms (same execution time). The nearly missing peak at τ near −200 ns is due to photon antibunching. Comparing spikes in the correlations near 0 time lag and at long time lags allows the triplet-state effects on the photon antibunching curve to be quantified. Such effects will play a critical role in using photon antibunching to count the number of fluorophores for each molecule.

Work at LLNL was performed under the auspices of the U.S. Department of Energy by the University of California, Lawrence Livermore National Laboratory, under contract W-7405-Eng-48. The Center for Biophotonics, a National Science Foundation Science and Technology Center, is managed by the University of California, Davis, under cooperative agreement PHY 0120999. T. A. Laurence's e-mail address is laurence2@llnl.gov.

References

1. S. A. Soper, L. M. Davis, and E. B. Shera, J. Opt. Soc. Am. B **9**, 1761 (1992).
2. J. R. Fries, L. Brand, C. Eggeling, M. Kollner, and C. A. M. Seidel, J. Phys. Chem. A **102**, 6601 (1998).
3. J. S. Eid, J. D. Muller, and E. Gratton, Rev. Sci. Instrum. **V71**, 361 (2000).
4. T. A. Laurence, A. N. Kapanidis, X. X. Kong, D. S. Chemla, and S. Weiss, J. Phys. Chem. B **108**, 3051 (2004).
5. D. Magde, E. Elson, and W. W. Webb, Phys. Rev. Lett. **29**, 705 (1972).
6. K. Schatzel, Inst. Phys. Conf. Ser. **77**, session 4, 175 (1985).
7. K. Schatzel, M. Drewel, and S. Stimac, J. Mod. Opt. **35**, 711 (1988).
8. D. Magatti and F. Ferri, Rev. Sci. Instrum. **74**, 1135 (2003).
9. M. Wahl, I. Gregor, M. Patting, and J. Enderlein, Opt. Express **11**, 3583 (2003).
10. K. Schatzel and R. Peters, in Proc. SPIE **1430**, 109 (1991).
11. K. D. Weston, M. Dyck, P. Tinnefeld, C. Muller, D. P. Herten, and M. Sauer, Anal. Chem. **74**, 5342 (2002).
12. C. W. Hollars, S. M. Lane, and T. Huser, Chem. Phys. Lett. **370**, 393 (2003).
13. S. Fore, T. A. Laurence, Y. Yeh, R. Balhorn, C. W. Hollars, M. Cosman, and T. Huser, IEEE J. Sel. Top. Quantum Electron. **11**, 873 (2005).
14. B. Lounis and W. E. Moerner, Nature **407**, 491 (2000).
15. J. Widengren, U. Mets, and R. Rigler, J. Phys. Chem. **99**, 13,368 (1995).
16. Y. Xiao, V. Buschmann, and K. D. Weston, Anal. Chem. **77**, 36 (2005).