Fluctuations as a source of information in fluorescence microscopy

Zdeněk Petrašek* and Petra Schwille

Biophysics group, Biotechnologisches Zentrum, Technische Universität Dresden, Tatzberg 47-51, 01307 Dresden, Germany

Fluctuations in fluorescence spectroscopy and microscopy have traditionally been regarded as noise—they lower the resolution and contrast and do not permit high acquisition rates. However, fluctuations can also be used to gain additional information about a system. This fact has been exploited in single-point microscopic techniques, such as fluorescence correlation spectroscopy and analysis of single molecule trajectories, and also in the imaging field, e.g. in spatio-temporal image correlation spectroscopy. Here, we discuss how fluctuations are used to obtain more quantitative information from the data than that given by average values, while minimizing the effects of noise due to stochastic photon detection.

Keywords: fluorescence spectroscopy; imaging; single molecule

1. INTRODUCTION

Fluorescence spectroscopy and microscopy techniques use the sensitivity of fluorescence to the environment of the probe to gain important information about the investigated system. The temporal variation of fluorescence parameters, such as intensity, lifetime, anisotropy or spectral properties, reflects the dynamic behaviour, while the spatial variation of these quantities provides contrast in imaging. The effects of fluorescence resonance energy transfer (FRET) or electron transfer on fluorescence allow spatial resolution far above the optical diffraction limit.

When studying samples with low concentration of fluorescent species or individual molecules, or when operating at high temporal resolution, the measured signal is typically weak and fluctuating. Often only the mean value of the signal is considered relevant, while the fluctuations are regarded as noise. This apparent noise does not, however, have its origin only in the imperfections of the experimental procedure, but the observable fluctuations are an inherent property of the thermodynamic system due to its small size. The presence of equilibrium fluctuations of an observable physical parameter has been exploited in various non-fluorescence fluctuation techniques, such as quasi-elastic light scattering (Berne & Pecora 1976), conductivity fluctuations in solutions (Feher & Weissman 1973) or through channels in a membrane (Neher & Stevens 1977), etc. (Weissman 1981). The high sensitivity combined with high temporal and spatial resolution that can be achieved experimentally by detecting fluorescence is the reason why the techniques where the fluctuations of physical quantities of a system are reflected by fluctuations in fluorescence have become the most widespread. Fluorescence fluctuation techniques have been extensively applied in chemistry, biophysics, biology and related fields in studies of photochemistry, diffusion properties, chemical kinetics, binding equilibria, aggregation, conformational fluctuations, to name a few (Rigler & Elson 2001; Krichevsky & Bonnet 2002; Petrov & Schwille 2008), both in vitro (Kahya & Schwille 2006) and in living cells (Bacia & Schwille 2003; Bacia et al. 2006), on small molecular ensembles as well as on single molecules (Edman & Rigler 2000).

The observed fluorescence fluctuations have usually several different sources. First, the signal fluctuations originate from varying number of emitters in the probed volume, or from variations of the emission properties of the emitting species due to interactions (lifetime and spectrum), reorientation (anisotropy) or internal dynamics. These fluctuations reflect an underlying physical evolution of the system, and are therefore of high interest. A second source of fluctuations, particularly relevant at low light levels or short integration times, is the variation of the number of detected photons due to the stochastic nature of light emission and detection. Also referred to as shot noise, these fluctuations do not carry useful information, but rather obscure the investigated physical process. Other fluctuation sources are those linked to background fluorescence, impurities or scattered light, and noise associated with the detection device (read-out noise, etc.).

While the relevance of the latter fluctuation sources can be reduced by eliminating background signals,
The division of the techniques into single-point measurements and imaging is based on the differences in the information content of the data typically obtained with these two techniques, calling for different analysis methods and therefore providing somewhat different information. In comparison with single-point measurements, imaging has usually lower temporal resolution and suffers from stronger shot noise contributions. On the other hand, it has the advantage of parallelism and offers the possibility of spatial correlation between different locations.

2. SINGLE-POINT MEASUREMENTS

With single-point measurements, we understand measurements on a small sample volume containing on average few or only one molecule. A small number of molecules is necessary for the amplitude of fluctuations to be large relative to the mean value, making fluctuation analysis experimentally possible (figure 1).

The experimental techniques can be divided into two groups, based on the number of molecules observed. In measurements performed on small numbers of molecules, without attempting to separate the signal from individual molecules, the fluctuations reflect the motion of molecules and intermolecular interactions, as well as the intramolecular dynamics. These techniques are suitable for the investigation of diffusion properties, association/dissociation reactions, aggregation, etc. Measurements performed strictly on one molecule are particularly suited for studying the intramolecular dynamics. There are two main reasons for that: the fluctuations due to changing number of molecules observed simultaneously are not present, and consequently, the data complexity is reduced because one does not have to deal with molecules in different phases on their internal state trajectory.

How can we extract meaningful physical information from a fluctuating signal? The simplest analysis involves the calculation of the mean of the observables from the measurement, e.g. fluorescence anisotropy. In this analysis, however, no use is made of the fluctuations.

In one of the most widely used fluctuation techniques, fluorescence correlation spectroscopy (FCS; Elson & Magde 1974; Kriechesky & Bonnet 2002; Petrov & Schwille 2008), temporal correlations $g(\tau) = \langle F(t + \tau)F(t) \rangle$ of the signal $F(t)$ between different times $\tau$ are calculated in addition to the mean $\langle F(t) \rangle$ (figure 2). This procedure allows us to decide whether the fluctuating signal originates from a few diffusing molecules or continuous fluorescent mass. The correlation values and their decay with time are a rich source of information on concentration, diffusion properties, chemical reactions, internal dynamics, photophysics, etc., on temporal scales from below nanoseconds to seconds and longer. Correlation analysis between different detection channels (spatial, spectral and polarization) widens the range of accessible system parameters even further (Bacia et al. 2006). FCS has been applied not only to in vitro or model systems, improvement of the optical design, development in detection technology and the use of photon-counting detectors, the first two sources will always be present. The aim of fluctuation-oriented techniques is to use the fluctuations of an observable fluorescence parameter to gain better insight into the physics of the system. Shot noise fluctuations present an obstacle on the way to this goal. Proper data treatment is required to eliminate the shot noise effects, while preserving the relevant information during the data reduction process, without introducing biases and at the same time keeping the precision as high as possible. The data analysis, usually involving averaging, should ideally return enough information to allow discrimination between different models.

The great importance of equilibrium fluctuations in physics stems from their connection to the macroscopic properties of the system. This is expressed by the fluctuation–dissipation theorem, which relates the fluctuations of a system in thermal equilibrium to the system response to external forces when brought out of equilibrium (Kubo 1966). Since the microscopic origins of equilibrium fluctuations and the macroscopic response are the same, analysis of fluctuations provides important parameters describing the macroscopic behaviour, and vice versa. For example, the fluctuation–dissipation theorem provides a relation between frictional coefficient (force dissipation) and the fluctuations in particle velocity in the case of Brownian motion; and the relation between thermal noise in a resistor and impedance in the electric circuit. The applications of fluorescence fluctuation methods as described here aim not only at obtaining the macroscopic response parameters (e.g. diffusion coefficient) but also at elucidating directly the microscopic properties of the system (e.g. rate and equilibrium constants).

The majority of the techniques described in this paper study fluctuations in a steady state, as opposed to relaxation techniques, where the system is perturbed and the relaxation back to the steady state is monitored. The exceptions are the correlation analysis of singlet–triplet transitions and antibunching (Widengren et al. 1995), where the processes of interest are the excitation and subsequent relaxation back to the ground state.

We would like to discuss here how current fluorescence microscopic techniques use fluorescence fluctuations to obtain more information about the investigated system than by considering only simple mean fluorescence values, and how knowledge about the photon statistics is incorporated into the analysis, in order to reduce or eliminate bias and extract maximum information. By fluorescence microscopy we mean not only imaging, but also techniques where measurement is performed at one microscopic location: single-point measurements. It is the size of the volume, in combination with low concentration, which results in a relatively low number of emitting species, and therefore gives the relevance to fluctuations. In an extreme case, measurement on a single molecule can be performed to investigate the internal dynamics. Imaging can be viewed as many point measurements in (semi-)parallel, and many considerations for point measurements can also be applied there.
Figure 1. (a) A schematic of a generic experimental set-up for single-point and scanning fluorescence measurements. The parallel excitation beam (ex) is deflected by the scanning optics (sc) and focused by a high numerical aperture objective (ob) into a desired diffraction-limited focal spot, where the fluorescent molecules are excited. The fluorescence is collected and collimated by the same objective, descanned, and separated from the excitation path by a wavelength-selective beam splitter (bs1). The following beam splitter(s) (bs2) divide the fluorescence into two or more channels, based on the differences in wavelength or polarization. The separated fluorescence is optionally spatially filtered by a confocal pinhole to ensure a well-confined measurement volume with small effective number of molecules (not shown), and may be further filtered by additional emission filters to suppress unwanted background, before being detected by point detectors, such as avalanche photodiodes or photomultiplier tubes (d1, d2). (b) Freely diffusing fluorescent molecules enter the measurement volume randomly, where they are excited and from which the fluorescence is detected. The number of molecules in the volume fluctuates; the relative amplitude and duration of fluctuations can be evaluated to provide concentration and diffusion coefficients. Independently diffusing molecules with different emission characteristics (light red circles, light green circles) can be distinguished from the bound species (joined light green and light red circles) by cross-correlation analysis. In addition to number fluctuations, fluctuations in intensity or other fluorescence parameter (light red circles ↔ dark red circles, light green circles ↔ dark green circles) can be detected, reflecting the intramolecular dynamics or intermolecular interactions.

Figure 2. Fluctuation analysis in single-point measurements. The (a) photon sequences recorded in multiple detection channels are (b) typically binned before (c) the correlation analysis or various types of intensity distribution analysis: (d) FIDA/PCH, (e) fluorescence intensity multiple distributions analysis (FIMDA), (f) two-dimensional FIDA, and (g) fluorescence intensity and lifetime distribution analysis (FILDA). The binned data from different channels can be combined to yield another parameter, for example, (h) the proximity ratio \( r = \frac{F_1}{F_2} \). The histograms of proximity ratios are analysed by (i) probability distribution analysis. The states can be identified from the trace of \( r \) revealing (j) the state trajectory. Alternatively, the raw photon sequences can be (k) directly correlated without binning, analysed by means of (k) photon arrival-time interval distribution (PAID), or by other methods operating directly on the photon-by-photon basis.
such as molecules diffusing and interacting in phospholipid membranes (Kahya & Schwille 2006), but also directly in living cells, for example, to monitor endocytosis of cholera toxin followed by separation of its subunits (Bacia et al. 2002), and to study binding of calmodulin to protein kinase II in cytoplasm (Kim et al. 2004).

Since the correlation is calculated between different time channels, and photon detection events are uncorrelated, the contribution due to photon-counting statistics to the correlation value is eliminated. The photon-counting noise affects only the variance of the correlation value.

It is worth noting that in FCS we are still performing ensemble averaging, assumed equivalent to temporal averaging, but extract additional observables (correlations) from the data in addition to the mean. Even in a true single molecule experiment we seek to perform an ensemble average, in a sense that the extracted parameters should be representative of a class of molecules (ensemble), not only of one particular molecule on which the experiment has been performed. The crucial question is how to perform the data reduction (‘averaging’ in a sense) in order to retain the maximum amount of relevant information and to eliminate the superfluous fluctuations due to the photon-counting noise without biasing the results.

With all its benefits, classical FCS does not exploit all the information contained in the data. A natural question is: what other observables can we calculate from the data to learn even more about the system?

One way is to consider correlations of higher order, such as \( g_m(\tau) = \langle F^m(t + \tau)F^n(t) \rangle \). Analysis of higher order correlations of the fluctuating signal, or higher order moments, allows characterization of distribution of aggregate sizes (Palmer & Thompson 1987). Considering multiple time correlations \( g(\tau_1, \ldots, \tau_{n-1}) = \langle F(t)F(t + \tau_1)\ldots F(t + \tau_{n-1}) \rangle \) with \( \tau_i = 0 \) instead of higher order moments \( \langle F(t)^m \rangle \) eliminate the contributions of shot noise (Qian & Elson 1990), because the shot noise between different channels is uncorrelated.

Higher order correlation functions, together with dual-colour cross-correlations, were shown to allow differentiation between systems in equilibrium and nonequilibrium steady state (Qian & Elson 2004). The non-equilibrium steady state of an open system is characterized by the presence of externally controlled fluxes (exchange of material), non-zero gradients in chemical potential and time irreversibility of fluctuations. It may exhibit concentration oscillations, but the average concentrations of reacting species may also be constant, as in a system in thermodynamic equilibrium (Qian et al. 2002). Living systems, or their parts, such as distinct reaction networks, are often in a non-equilibrium steady state, rather than in equilibrium. It is, therefore, important to be able to recognize and characterize the non-equilibrium steady state, for example, by determining the fluxes. Experimental approaches employing fluctuation analysis have been proposed for this task, and are expected to play an important role in the studies of reaction networks in living cells (Qian 2006, 2007).

Going one step further, one can sort the detected photons in bins of a fixed temporal width shorter than the diffusion time (diffusion can then be neglected), and analyse the full distribution of the number of photons per bin instead of calculating only the moments of the distribution (fluorescence intensity distribution analysis (FIDA: Kask et al. 1999), photon counting histogram (PCH; Chen et al. 1999)). The distribution is determined by both the fluctuation of the number of molecules in the observed volume and the photon-counting statistics. Assuming essentially Poisson statistics of these two sources of fluctuations allows one to determine the mean number of molecules (concentration) and the molecular brightness (mean number of photons detected from one molecule per unit time). Contrary to FCS, this method is sensitive to the presence of more species of different brightness, regardless of their diffusion properties (Müller et al. 2000). FCS can resolve two species only if their diffusion times differ significantly, implying large difference in size, and even then provides only product of concentration and square of the brightness of each species. FIDA was applied to study hybridization and cleavage of labelled oligonucleotides, with resolution of all cleavage products based on molecular brightness (Kask et al. 1999).

Sorting photons into bins shorter than the diffusion time and analysing the histogram of photon counts means that all dynamic information due to diffusion contained in the correlations between different bins is lost. However, if a series of histograms is constructed where the bin width increases from one histogram to another, in the range of bin widths from shorter than to longer than the diffusion time, the motion of molecules will affect the distributions (Palo et al. 2000; Gopich & Szabo 2005b). Then, in addition to brightness and concentration, the temporal characteristics of fluctuations can be extracted, as done by evaluating correlations in FCS. The technique has been demonstrated by determining the binding constant of a protein–ligand interaction (Palo et al. 2000). A similar way to extract information about molecular diffusion from photon count histograms with a range of bin widths is to analyse the cumulants of the resulting photon count distributions (Wu & Müller 2005).

As in FCS, dividing the fluorescence signal into two channels on the basis of different polarization, emission wavelength, etc., with selectively higher sensitivity for one or the other species to be resolved, leads to significantly higher accuracy compared with the one-channel technique. Both photon streams are binned and a joint photon count number distribution (two-dimensional histogram) is constructed and fitted to a model (Kask et al. 2000). This approach was shown to be suitable for studying the antibody–antigen binding and ligand–receptor interactions (Kask et al. 2000). Adding nanosecond resolution and pulsed excitation introduces the ability to distinguish fluorescence species by the differences in their fluorescence lifetime, in addition to brightness. The raw data are reduced and analysed by constructing two-dimensional histograms, where the second axis is determined by the sum of the delay times from the excitation pulse of all photons in the bin. Binding of peptide to calmodulin could be monitored with this method, owing to the change of fluorescence lifetime upon complex formation (Palo et al. 2002).
Another analysis approach, bringing together the benefits of FCS and photon count distribution analysis, avoids binning the photon counts and involves construction of a two-dimensional histogram of photon pairs instead (photon arrival-time interval distribution (PAID; Laurence et al. 2004)). One histogram dimension is defined by the temporal separation of the two photons (analogous to FCS correlation) and the other dimension corresponds to the number of photons detected within the time interval defined by the detection of the two photons. The analysis of the histogram allows simultaneous determination of the mean number of molecules, molecular brightness and characterization of diffusion. The technique was shown to resolve and characterize different species involved in the RNA–polymerase–DNA interaction (Laurence et al. 2004). As with the other methods, PAID can be extended by combining photon streams from different detection channels.

The methods discussed so far use correlation or photon count distribution analysis to analyse signal from a small number of molecules. When the interest is in the intramolecular dynamics, e.g. conformation fluctuations and the related fluctuations in enzymatic activity (Xie 2002), observation of one molecule at a time allows even more detailed data analysis. This can be realized by immobilizing the molecule, confining it into a volume smaller than the probed volume, or diluting the sample so that statistically at most one molecule is present in the observation volume.

When freely diffusing, the molecule passing through the probed volume generates a photon burst, with its size limited by the time spent in the volume. Consequently, only a limited number of photons is available for analysis, and any observable quantity determined from the information carried by these photons, such as FRET efficiency (or alternatively, proximity ratio (Antonik et al. 2006)) or fluorescence anisotropy (Kalnin et al. 2007), is affected by the stochasticity of photon detection. With an elaborate analysis, the photon-counting statistics can be fully accounted for, and discrete probability distributions of the observables can be constructed that accurately describe the shot noise broadening of not only the experimental histograms, but also the step-like histogram structure due to the discrete nature of the photon-counting data. As a result, probability distribution analysis is very sensitive to histogram broadening due to a distribution of the values of the observable quantity, i.e. to the sample heterogeneity. When applied to FRET in donor–acceptor labelled DNA, motion inhomogeneities in the range of 0.5 nm could be resolved (Antonik et al. 2006). The method can be implemented either for the data binned into equally sized bins (Antonik et al. 2006; Kalnin et al. 2007) or for histograms of values obtained from whole bursts of varying size (Nir et al. 2006).

The probability distribution analysis reveals state heterogeneity, but not the interconversion rates between the states. These can be obtained by calculating the observable value (fluorescence lifetime, FRET efficiency, etc.) in channels of the binned signal, identifying the states and constructing histograms of dwell times, as demonstrated by identifying different conformational states of a fluctuating DNA oligonucleotide (Eggeling et al. 1998). Alternatively, temporal autocorrelation of the calculated fluctuating observable can be used.

Identification of states from the trajectory of the observable becomes difficult when the number of states, the transition rates or the photon-counting noise increase. Then, it is advantageous to view the molecular evolution through an unknown sequence of states as a Markov process, meaning that the next molecular state depends only on the previous state and not on earlier history. The state trajectory can be reconstructed and the transition rates can be determined from the fluctuating data using hidden Markov models (HMM; Talaga 2007). This procedure uses directly the raw photon sequence defined by the photon arrival times (Andrec et al. 2003). HMM were shown to be applicable to the identification of rotational steps of ATP synthase from the FRET data (Zarrabi et al. 2007), and to protein–DNA interaction studies (McKinney et al. 2006). Since the experimental observables reflecting the state of the molecule are not necessarily unique to each state, the measured sequence of the observable need not be a Markov process anymore. Non-Markovian property of the data, experimentally manifested by non-exponential relaxations, can be identified not only by HMM, but also by higher order correlation analysis (Edman & Rigler 2000).

Binning the data into channels prior to further processing may appear to have the advantage of reducing the photon-counting noise. However, binning causes loss of information contained in the precise timing of the photon. In order to avoid this information loss, photon-by-photon analysis algorithms requiring no binning have been developed. They allow calculation of correlation functions, or other type of analysis, directly from the photon sequence, without having to determine the relevant observable, such as lifetime or FRET distance, at every instance beforehand (Yang & Xie 2002; Gopich & Szabo 2005a). The benefit is a higher time resolution and a broader temporal correlation range. Photon-by-photon analysis of FRET distance trajectory using a maximum-likelihood approach additionally provides rigorous error estimates (Schröder & Grubmüller 2003). A photon-by-photon analysis using a Bayesian estimator has been developed for identification of a fluorescent molecule diffusing through a laser focus from a known set of possible species, based on the differences in diffusion coefficient or brightness (McHale et al. 2004).

Considering the various analysis methods (Lippitz et al. 2005) yielding often different types of information, an obvious question arises: is a particular method analysing the data efficiently, and what is the limit on the amount of information that can be extracted from given data with certain noise level? The answers are being sought using the concepts of information theory, such as Shannon information, entropy, mutual information or Fisher information (Watkins & Yang 2004; Talaga 2006). The information theory analysis of the data processing method can tell us, for example, whether it is possible to reconstruct a hidden molecular state trajectory from the given photon stream. The optimal methods use photon-by-photon approach, often in combination with maximum-likelihood estimators fully incorporating the Poisson statistics of photon detection.
3. BEYOND POINT MEASUREMENTS

In many cases of intermolecular interactions the fluctuations due to molecular motion are very slow. Consequently, the measurement at one location would require prohibitively long time for the system to sample sufficiently well the space of its allowed states, i.e. to approximate well enough the ensemble average. A solution to this problem is to perform the measurement on many systems in parallel.

In image correlation spectroscopy (Petersen et al. 1993), the image \( F(x, y) \) containing many copies of the investigated system (e.g. aggregating membrane receptor) is autocorrelated (figure 3). The autocorrelation amplitude \( g(0, 0) \) of the spatial image correlation \( g(\xi, \eta) \) is related to the receptor density in the image. If an assumption can be made about the distribution of the number of the fluorescent subunits of the receptor, the autocorrelation amplitude can yield information about the parameters of this distribution, such as the mean number of subunits. In this way, molecular aggregation on immobile or slowly moving complexes can be studied. As other correlation techniques, image correlation can be extended to work with more detection channels (Petersen et al. 1998), or to use higher order correlations in order to better characterize aggregation. Applications of image correlation include quantification of the aggregation state of plasma membrane receptors revealing their preclustering before ligand binding (Wiseman & Petersen 1999).

If higher temporal resolution is required than that available with fast imaging, but the motion of fluorescent species is still too slow for measurement at one location, combination of standard FCS with scanning the measurement volume across the sample may help (Petrašek & Schwille 2008), as was shown in the measurements of slow motion of GFP-labelled proteins in the cortex of developing *C. elegans* embryos (Petrašek et al. 2008). In this way, a large total volume can be sampled allowing good averaging. The motion of the measurement volume has to be taken into account in the correlation analysis. Scanning can provide additional benefits: reduction of photobleaching (Petrašek & Schwille 2008a), elimination of slow membrane fluctuations (Ries & Schwille 2006), or simultaneous measurement of flows, diffusion and immobilization (Skinner et al. 2005).

Since scanning FCS obtains signal from different locations, autocorrelation of the detected photon stream represents a subset of the full spatio-temporal correlation \( g(\xi, \eta, \tau) \), sampled only at certain coordinates \((\xi, \tau)\) determined by the scan path, however with considerably higher temporal resolution. Owing to its relation to the probability density function \( p(\mathbf{r}, t) \) describing the particle motion, the spatio-temporal correlation provides more detailed characterization of

---

**Figure 3. Spatial correlation analysis.** (a) Spatial correlation of a fluorescence image produces a peak centred at the origin; the correlation value at \((\xi, \eta)\rightarrow (0, 0)\) is inversely proportional to the density of fluorescent species, and the width of the correlation peak is related to the instrument point spread function and the size of the fluorescent features in the image. (b) Spatio-temporal correlation spectroscopy (STICS): performing additional temporal correlation across a sequence of images enables the determination of diffusion coefficients by quantifying the broadening of the correlation peak at increasing correlation times \( \tau \). The shift of the correlation peak maximum away from the origin corresponds to the mean displacement vector (red arrows) and can be used to determine the magnitude and direction of the average flow. (c) Dividing the image into segments and performing spatio-temporal correlation analysis in each segment separately allows mapping of the flow field (green arrows).
the particle dynamics than, for example, the commonly investigated dependence of the mean square displacement on time (Petrášek et al. 2008).

One of the motivations behind scanning described so far was to obtain statistically significant data for correlation analysis by measuring many particles at different locations. A different application uses circular scanning with a feedback loop to follow the motion of one fluorescent particle in three dimensions without disturbing it (Berglund & Mabuchi 2005; Levi et al. 2005). In this way, the particle trajectory is obtained, and at the same time the fluorescence stream can be analysed for the internal dynamics on a temporal scale as long as the tracking period (McHale et al. 2007). This approach enables measurement on one particle without immobilizing or confining it, thereby minimally affecting its behaviour. The technique has been applied to the study of phagocytosis of protein-coated beads (Levi et al. 2005) and to the fluctuation analysis of freely diffusing quantum dots from nanosecond time scales, permitting monitoring of antibunching, to the time scale of seconds, where bright/dark state transitions are observed (McHale et al. 2007). The limits of tracking accuracy and maximum trackable speed are set by the bandwidth of the feedback loop and especially by the photon-counting statistics (Enderlein 2000; Berglund & Mabuchi 2006). Shot noise is, in this case, a source of two negative effects: it obscures the useful information in the detected signal, and prevents fast diffusing particles from being tracked.

4. MICROSCOPY IMAGING

Fluorescence microscopy imaging techniques are usually employed to gain different information from that obtained in single-point measurements: two- or three-dimensional spatial distribution of fluorescent molecules, particles or larger structures, or distribution of any parameter reflected by their fluorescence, and the changes of these distributions in time. However, the situation is similar regarding the useful signal and noise: the fluctuations of the sample carrying interesting information are mixed with the fluctuations due to photon-counting statistics, or the detection process in general. The relevance of detection-associated fluctuations is exacerbated by typically shorter acquisition times per pixel than in single-point measurements. For this reason, the elimination of fluctuations in imaging has been more in focus than the exploitation of fluctuations as a source of new information.

The general problem of noise reduction has been extensively researched in different fields reaching from astrophysics to biomedical imaging (Verveer et al. 1999; Bonnet 2004; Puettet et al. 2005). Many of these approaches are applicable also to fluorescence microscopy imaging, as can be seen in examples of three-dimensional image reconstruction and deconvolution techniques. The image restoration algorithms are able to deal with different noise distributions, and can therefore be optimized for images obtained with analogue detectors or the increasingly popular photon-counting devices.

Proper treatment of detection noise is especially important in fluorescence microscopy techniques where the contrast is not formed directly by variations in fluorescence intensity, but by variations in another parameter calculated from the data, such as in ratio imaging or lifetime imaging. Low pixel intensities imply high noise necessitating the use of unbiased estimators of these quantities.

In ratio imaging, the signal is separated into two detection channels on the basis of different polarization or emission wavelength. The signal ratio can be used to monitor rotational mobility, resonance energy transfer, spectral fluctuations, etc. The presence of noise in the two signals x and y causes not only a spread of values of calculated ratio x/y, but also deviation (bias) from its true value (Wang 2007). The estimated ratio x/y can then be falsely found to depend on the total intensity x + y. The knowledge of the noise statistics allows estimation of the bias and construction of less biased or unbiased estimators (van Kempen & van Vliet 2000). For example, the ratio estimator \( E(x/y) = x(1/y + (1 - y)/(y(y^2 - 2))) \) has almost no bias and equal or smaller variance than \( E(x/y) = x/y \) for practically useful values of x and y, if x and y are Poisson distributed with uncorrelated noise.

In fluorescence lifetime imaging, the lifetime (or decay rate) is calculated by fitting the fluorescence decay in one pixel to a model function, in the simplest case a single exponential (Suhling et al. 2005). Using the criterion of minimum \( \chi^2 \) for finding the estimate of fluorescence lifetime implicitly assumes Gaussian-distributed noise. However, in imaging applications the photon counts per pixel are often low, and the Poisson distribution of the number of photons in individual channels is no longer well approximated by the Gaussian distribution. The \( \chi^2 \) minimization then results in biased decay rate values. Maximum-likelihood estimators assuming Poisson photon-counting statistics are known to produce smaller bias that can be easily corrected for (Tellingshuisen & Wilkerson 1993).

Photon-counting noise is the fundamental factor determining the accuracy limit in recently introduced super-resolution imaging techniques (Betzig et al. 2006; Sharonov & Hochstrasser 2006; Toprak & Selvin 2007; Huang et al. 2008) and in high-accuracy particle-tracking experiments (Yildiz & Selvin 2005). Both classes of techniques rely on the calculation of the centre position of a pixelized diffraction-limited image of a much smaller emitter, built up from a limited number of photon counts. The accuracy limit has been shown to lie at approximately 1–2 nm under realistic conditions (Thompson et al. 2002). Relevant analysis methods of photon-counting data, including the centroid calculation (Suhling et al. 1999), have been previously developed in the field of astrophysics, where similar limitations and problems are encountered—weak signals combined with the need of high spatial resolution.

The previous examples illustrate how fluctuations limit the imaging techniques and how they are dealt with. But can the fluctuations of the physical properties of the sample be used to obtain additional information, as in single-point techniques?
Conceptually simpler are the methods attempting to perform the previously discussed point measurements at many locations in parallel. In addition to the abovementioned lifetime and ratio imaging, new techniques are appearing that aim at producing maps of physical parameters derived from the signal fluctuations. For example, temporal correlation of a sequence of images $g(0, 0, \tau)$ provides information about diffusion (Wiseman et al. 2000; Burkhardt & Schwille 2006), and, if performed pixel wise, can be used to generate maps of diffusion coefficients and concentrations, as demonstrated in an example of lipid molecules diffusing in a lipid bilayer, and receptor proteins diffusing on cell membranes (Kannan et al. 2007). Analysis of the mean and variance of a fluctuating fluorescence signal in individual pixels of an image sequence can be used to produce maps of average molecular brightness and average particle number (Digman et al. 2008).

The fluctuation analysis can, however, use also the spatial information contained in images. When investigating the dynamics of a fluorescently labelled microscopic sample, the motion of structures or even individual particles is identified in a sequence of images, and can be further evaluated to yield quantitative information. However, it is often desirable, or dictated by the biological system, that the concentration of fluorescent molecules is low. Additionally, fast dynamics necessitates short integration times. Both these requirements lead to low signal per image pixel, and therefore strong contribution of photon-counting noise, often preventing direct identification and tracking of distinct labelled structures. Still, correlation analysis of an image sequence makes it possible to identify a pattern in the fluctuating signal affected by photon-counting statistics. Images are divided into small segments $F(x, y, t)$ for which three-dimensional correlation $g(\xi, \eta, \tau)$ is calculated. The temporal decay and the location of the correlation maximum in $(\xi, \eta)$ are used to determine the average diffusion coefficient and flow vector (Hebert et al. 2005). Assembling the results from all image segments produces a map of diffusion coefficients and a flow map (figure 3). It is also possible to eliminate immobile image features and thus focus only on the dynamic (fluctuating) constituents of the system. The flow mapping via correlation analysis has been used, for example, to study the motion of adhesion-related proteins and actin in migrating cells (Brown et al. 2006).

The amplitude, and therefore detectability, of physical fluctuations in images can be enhanced by optimized labelling, as used in the technique called speckle microscopy (Danuser & Waterman-Storer 2006). Speckle microscopy is suitable for studying the assembly dynamics, subunit turnover and movement of biological structures, consisting of a large number of subunits of one or several types, such as microtubules or actin filaments. A small fraction of the subunits (0.5%) is fluorescently labelled, resulting in a speckled pattern on the otherwise invisible structure. Appearance, disappearance and motion of the pattern reflect the dynamics of the structure, without the need to identify (resolve), localize or track individual molecules (Waterman-Storer et al. 1998). Correlation analysis of image segments between subsequent frames provides an average displacement vector, describing the motion even in the presence of high noise and low speckle contrast. Consequently, flow maps can be constructed, with resolution determined by the size of correlated image segment only slightly lower than the diffraction limit (Ji & Danuser 2005). The calculated flow map can subsequently assist tracking of individual speckles in the image sequence. With sufficient contrast, the speckles can also be identified and analysed individually, without the preceding correlation step.

The spatio-temporal correlation analysis is typically performed on the data obtained from highly sensitive CCD cameras where the whole frame is exposed simultaneously, and the temporal resolution is limited by the frame transfer rate. Higher temporal resolution than in frame-based image correlation can be achieved in correlation analysis of images from a confocal laser scanning microscope (Digman et al. 2005). Since the laser beam is scanned across the sample in a well-defined fashion to create an image, correlation of the fluorescence separately along the fast and slow axis, and across the frames, gives access to three different time scales on which the molecular dynamics can be measured. This approach combines the features of imaging with scanning FCS.

5. CONCLUSION

Perhaps not surprisingly, the extraction and analysis of useful information carried by fluctuating fluorescence signal has advanced most in the field of single molecule spectroscopy. The reasons can be related to the facts that when observing only one molecule no additional fluctuations due to the occupation number or trajectory phase differences are present, the shot noise fluctuations are well defined, and background contributions to the signal can be efficiently eliminated. FCS on a small number of molecules is well established and exists in many variations, some of them optimized to determine one parameter precisely, others to measure many parameters simultaneously. In imaging applications, the focus has traditionally been on removing fluctuations regarded as noise. Nevertheless, correlation approaches can use physical fluctuations in the sample to image quantities, such as flow directions, that appear lost in noise to a human eye.

Concerning the technology development, detectors and optics have approached their limits in sensitivity, and significant further development can be expected perhaps in parallel detection (avalanche photodiode arrays). A transfer of technology is occurring from astrophysics, where position-sensitive detectors with high temporal resolution are being developed, such as quadrant anode, wedge-and-strip, or delay line multi-channel plate detectors (Michalet et al. 2007). Similarly, although fluorescent probes are improving in photostability and brightness, the limitations on photon yield due to finite lifetime and the total energy dose deposited in biological sample will persist.

Therefore, it has become increasingly important to find ways to extract the maximum possible information from the available data. Some of the experience with
the advanced analysis methods in the single molecule field is expected to be transferred to imaging. Instrumentation is available that can record the whole photon stream from which the images are built up with picosecond resolution (Emiliani et al. 2003; Becker et al. 2006; Wahl et al. 2007), in principle allowing a similar type of analysis as in single-point techniques. Implementation of multiparameter fluorescence detection (Widengren et al. 2006), i.e. separating the signal into many detection channels, is particularly important in order to resolve heterogeneities typical for complex samples typically investigated with imaging.

Additionally, new analysis methods will have to be developed that will take into account the complexity of living organisms, reflected, for example, by the fact that the investigated system is often in a non-equilibrium steady state rather than in thermodynamic equilibrium.

REFERENCES


