Membrane heterogeneity : from lipid domains to curvature effects

Semrau, S.

Chapter 7

Counting autofluorescent proteins *in vivo*

The formation of protein complexes or clusters in the plasma membrane is crucial for cell signaling and other biological processes. Therefore, it is valuable to follow complex formation in vivo. *Typically, autofluorescent proteins are genetically fused to the proteins of interest to make their interaction visible. With highly sensitive single-molecule microscopy it is possible to detect single proteins and protein clusters despite autofluorescent cellular background. By measuring the fluorescence intensity of a diffraction limited spot, the number of proteins in that spot can be determined. However, the poor photophysical stability of fluorescent proteins hampers this straightforward approach. Also the presence of noise and autofluorescent background – unavoidable in live cell recordings – influence the measurement. Here we present solutions to these problems and show that the number of proteins in a diffraction limited spot can be determined* in vivo *in an accurate and robust way. By quantification of the number of YFP molecules in diffraction limited spots we confirm that the membrane anchor of human H-ras heterogeneously distributes in the plasma membrane. Our description of the single YFP intensity distribution therefore provides an accurate approach for quantitative* in vivo *investigations of protein cluster formations.*

7.1 Introduction

A prominent example for the importance of protein complex formation is found in plasma membrane located signaling cascades. Here receptor molecules such as the Toll-like-receptor 3 (212) or the epidermal growth factor receptor (213) form dimers upon binding to their respective ligands. In addition to the formation of true dimers or oligomers, also protein clusters and lipid domains lead to heterogeneities in the spatial distribution of specific proteins in the plasma membrane and are known to have an effect on the dynamics of the reactions they are involved in. The small GTPase H-ras has been shown to confine to such small membrane domains upon its activation (131). Clearly, the demand for the determination of cluster formations with sufficient temporal resolution in the context of living cells is high. Quite a number of techniques have been addressed to this problem, namely bioluminescence resonance energy transfer (214), bimolecular fluorescence complementation (215), number and brightness mapping (216), image correlation spectroscopy (217), confocal laser scanning microscopy with fast scanning (218) and photon counting histograms (219). These techniques all measure ensembles of molecules, making assumptions about their collective behavior including thermal equilibrium and spatial homogeneity. The validity of such assumptions, which are potentially violated in the context of a living cell, is difficult to prove. Furthermore, many of these techniques require exact knowledge about experimental parameters, like e.g. the point spread function of the microscope, which the results might sensitively depend on.

Single-molecule experiments, on the other hand, mostly depend on universal properties of fluorescent tags and inherently possess the aspired sensitivity. It was shown *in vitro* that the number of molecules can be determined from the intensity of attached fluorophores (121). Both, the use of fluorescent proteins, which exhibit complex photophysics (109), and the presence of high noise make this type of analysis more challenging in the *in vivo* context. In a recent study Ulbrich et al. demonstrated that molecule numbers can be assessed *in vivo* from the bleaching steps of autofluorescent proteins (106). This method requires the selection of intensity trajectories that show the expected step-wise decrease of fluorescence. Another approach, introduced by Cognet et al. (107), is to collect all emitted photons until photobleaching, such that fluorescence intermittency is averaged out. This method essentially uses photon counting histograms (141) to determine the underlying distribution of molecule numbers. Here we report

on the adaptation of this single-molecule method, which was demonstrated *in vitro* (107, 121), to autofluorescent proteins in living cells. We put the approach taken intuitively by Cognet et al. on firm theoretical grounds using semi-classical Mandel theory (220). Our theoretical result allows us to choose experimental parameters for a faithful measurement of single-molecule intensity distributions. Furthermore, we address the problems arising from an autofluorescent background, which are especially severe in living cells. In particular, the probability to detect a fluorophore in a noisy background depends on the intensity of the fluorophore and therefore modulates measured intensity distributions. We quantify this detection probability and its influence on intensity measurements. We verified our theoretical predictions by measuring the integrated intensities of single YFP molecules in living cells under experimental conditions where the molecules bleach within the illumination time. We show that the resulting distributions can be described by a simple one-parameter model, which allows for the quantification of molecule numbers by established methods (121).

Finally, we apply our method to the membrane distribution of the H-Ras membrane anchor. While measurements at low spatial densities of the protein yield a strictly monomeric distribution, only slight increases in H-Ras density cause evident increases of the dimeric fraction. Assuming a random spatial distribution of H-Ras, such a density dependent effect would be expected only at much higher concentrations. We therefore have to assume a non-random distribution of H-ras. Hence, we are able to confirm the results of an earlier study (150) that this membrane anchor clusters on length scales below the width of the point spread function ($\approx 200nm$), exclusively by using information from measurements of single molecule intensities.

Taken together, our method to accurately and quantitatively describe the intensity distribution of YFP-fusion proteins *in vivo* is suited, to characterize the spatial distribution of membrane proteins based on membrane heterogeneities or domain formations on the length scale of or below the diffraction limit. While inherently the method is not able to report the exact size of such domains, it makes up for this lack of spatial resolution by providing temporal resolution. With our method, the status of domain presence can in principle be monitored many times in the course of a biological reaction, such as a signaling event. In the same way, the presence or formation of true protein complexes (i.e. dimers, trimers or

higher multimers) would lead to intensity distributions, from which our method could extract the stoichiometry and its changes. However, as our results show, membrane heterogeneity has to be taken into account when true complex formation is to be measured.

7.2 Materials and Methods

7.2.1 DNA constructs

The protocol for the preparation of the DNA constructs was previously described in detail in (150). The DNA sequence encoding the 10 C-terminal amino acids of human H-Ras (GCMSCKCVLS), which includes the CAAX motif, was inserted in frame at the C-terminus of the enhanced yellow-fluorescent protein (EYFP, S65G/S72A/T203Y) coding sequence using two complementary synthetic oligonucleotides (Isogen Bioscience, Maarssen, The Netherlands). The integrity of the reading frame of the resulting EYFP-C10HRas construct was verified by sequence analysis. For expression in mammalian cells, the complete coding sequence of EYFP-C10HRas was cloned into the pcDNA3.1 vector (Invitrogen, Groningen, The Netherlands).

7.2.2 Cell culture

For all experiments a Chinese Hamster Ovary (CHO) cell line (clone D3) was used. Cells were cultured in DMEM:F12 1:1 medium supplemented with streptomycin ($100\mu g/ml$), penicillin ($100U/ml$) and 10% new born calf serum in a 7% CO_2 humidified atmosphere at $37°C$. Cells were used for $25-30$ passages and were transferred every 4 days. For microscopy cells were cultured on cover glass slides (Assistent, Karl Hecht KG, Sondheim Germany) and transfected with $250ng$ DNA and $3\mu l$ FUGENE HD (Roche Molecular Biochemicals, Indianapolis, USA) per glass slide (1h incubation time). For a convenient expression level cells were used $3-4$ days after transfection.

7.2.3 Single-molecule microscopy

The experimental setup for single-molecule imaging has been described in detail previously (150). Briefly, the microscope (Axiovert 100; Zeiss, Oberkochen Germany) was equipped with a 100x oil-immersion objective

(NA=1.4,Zeiss, Oberkochen, Germany). The samples were illuminated for $T = 50ms$ by an Ar^+ laser (Spectra Physics, Mountain View, CA, USA) at a wavelength of $514nm$. The illumination intensity was set to $3 \pm 0.3\,\mathrm{kW/cm^2}$. A circular diaphragm was introduced in the back focal plane of the tube lens to confine the illumination area. This results in a flat laser illumination profile. An appropriate filter combination (DCLP530, ET550/50m, Chroma Technology, Brattleboro,USA) permitted the detection of individual fluorophores by a liquid nitrogen cooled slow-scan CCD camera system (Princeton Instruments, Trenton, NY, USA). The total detection efficiency of the imaging optics was $\eta_o = 0.12$. The time between consecutive images (time lag, Δt) was set to $254ms$. Typically, 4000-8000 images were obtained per cell. For the observation of the intensity of individual EYFP-CAAX molecules, CHO cells adhered to glass slides were mounted onto the microscope and kept in phosphate buffered saline (PBS: 150mM NaCl, 10mM Na_2HPO_4/NaH_2PO_4, pH 7.4) at $37°C$. The focus of the microscope was set to the ventral surface membrane of individual cells (depth of focus $\approx 1\mu m$). The density of fluorescent proteins on the plasma membrane of selected transfected cells was less than $1\mu m^{-2}$ to permit imaging of individual fluorophores. According to (110) the bleaching time τ_{bl} for the used laser intensity $I_{\mathrm{ill}} = 3kW/cm^2$ is $10.4ms$. The probability that a single EYFP bleaches within the illumination/integration time $T = 50ms$ is therefore $p_{\mathrm{bl}} > 99\%$ (110). In other words, a single EYFP is bleached within the illumination time. The bleaching rate $k_{\mathrm{bl}} = 1/\tau_{\mathrm{bl}}$ is well separated from $1/T$ ($k_{\mathrm{bl}} \approx 0.2T^{-1}$) and therefore the simplified model described below (see Eq. 7.1) is applicable. The expected photon emission rate expected from results in (110) is $F = 775\,\mathrm{photons/ms}$. Therefore, $N = \eta_o \tau_{\mathrm{bl}} F = 967$ photons are expected to be detected during the average lifetime τ_{bl} of the fluorophore, where the detection efficiency is $\eta_o = 0.12$.

7.2.4 Image analysis

At first, the autofluorescent background is subtracted from the raw images. The background subtracted images are subsequently filtered with a Gaussian whose width corresponds to the width of the point spread function (PSF) of the microscope. This procedure optimizes the signal to noise ratio. The positions of the pixels whose value after filtering exceeds a certain multiple of the noise are used as initial values for the fitting of a 2D Gaussian in the unfiltered image. From this fit, position, width and

integrated intensity of the single molecule signal are determined. More details can be found in the supplements (Sec. 7.B).

7.3 Results and Discussion

If several fluorescent molecules are colocalized on a length scale of $\lesssim 200nm$, their fluorescence signal will be a single diffraction limited spot in a widefield microscope. In the following we will refer to a certain number of molecules in a diffraction limited spot as monomer, dimer, trimer, ... irrespective of the origin of colocalization: the molecules might e.g. be part of a stable complex or transiently reside in the same nanoscopic domain. Although the molecules cannot be resolved, it is possible to infer their number from the integrated fluorescence signal. Since the number of molecules cannot be calculated from a single signal (due to noise), it is is necessary to analyze distributions of fluorescence intensities of many diffraction limited spots. To first approximation, the total fluorescence signal integrated over the diffraction limited spot should be linearly proportional to the number of molecules. In experiments this simple relationship does not hold due to the complex photophysics of fluorescent tags and the data analysis process, as detailed in the subsequent sections.

7.3.1 Blinking and bleaching of fluorescent proteins

YFP and other autofluorescent proteins are popular tags for biomolecules *in vivo* because of their ease of use and the guaranteed 1:1 labeling ratio. Unfortunately, fluorescent proteins exhibit complex photophysics: they are known to blink, i.e. switch transiently between fluorescent and non-fluorescent states, and bleach fairly quickly. This poor photostability can make it difficult to infer molecule numbers from the fluorescence signal. We illustrate the photophysics of a fluorescent protein with a 3-state model derived in Sec. 7.C.1, see inset to Fig. 7.1A. In this model the fluorophore switches between 'on' and 'off' with a rate k and bleaches with a rate k_{bl} from the 'on' states. Only in the 'on' state the protein emits photons with a mean rate $\bar{I}$. It cannot return to a fluorescent state once it is bleached. Fig. 7.1 shows the number of photons emitted by a single fluorophore during illumination time T calculated from the 3-state model. In Fig. 7.1A the influence of blinking is illustrated. Since the photon emission rate in the 'on' state is set to $\bar{I} = 100/T$ the mean of the distribution is

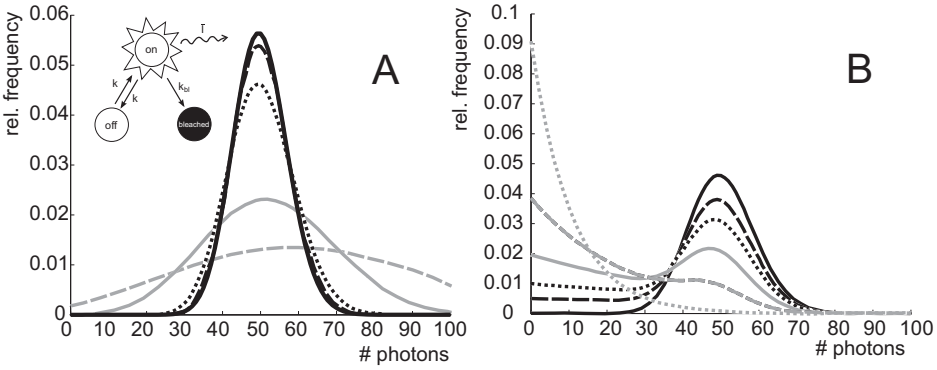

Figure 7.1
Photophysical model for blinking behavior of YFP. Inset panel A: Schematic representation of the model. The fluorophore switches between 'on' and 'off' with a rate k and bleaches with a rate k_{bl} from the 'on' states. Only in the 'on' state the protein emits photons with a mean rate $\bar{I}$. Once bleached it cannot return to a fluorescent state. A) Influence of blinking for negligible bleaching. Relative frequency of numbers of photons emitted by a single fluorophore during illumination time T. The bleaching rate k_{bl} is in all cases $k_{bl} = (10000T)^{-1}$, the rate of photon emission in the 'on' state is $f = 100/T$. The blinking rate k is $k \to \infty$ (solid black line, limit given by Poisson distribution), $k = (0.3333T)^{-1}$ (dashed black line) ,$k = (0.1T)^{-1}$ (dotted black line), $k = (0.01T)^{-1}$ (solid grey line) , $k = (0.002T)^{-1}$ (dashed grey line). B) Influence of bleaching for fixed blinking rate. Relative frequency of numbers of photons emitted by a single fluorophore during illumination time T. The blinking rate k is in all cases $k = (0.01T)^{-1}$, the rate of photon emission in the 'on' state is $f = 100/T$. The bleaching rate k_{bl} is $k_{bl} = (10^4T)^{-1}$ (solid black line),$k_{bl} = (2T)^{-1}$ (dashed black line) ,$k_{bl} = T^{-1}$ (dotted black line) , $k_{bl} = (0.5T)^{-1}$ (solid grey line), $k_{bl} = (0.25T)^{-1}$ (dashed grey line), $k_{bl} = (0.1T)^{-1}$ (dotted grey line).

approximately 50, at least for high blinking rates. Clearly, the width of the distribution increases with decreasing blinking rate. Note that even for infinitely fast blinking the distribution has a finite width. This minimal width is due to the fact that photon emission is a stochastic process. The variance of the Poisson distribution, which describes this process, is equal to the mean (here: 50), so the minimal width (standard deviation) is $\sqrt{50}$. For very slow blinking, the mean shifts to the right (dashed grey line in Fig. 7.1), if the fluorophore is initially 'on'. Both effects make it difficult to distinguish the monomer with mean intensity 50 from the dimer, which would have a mean intensity of 100. As with blinking, bleaching strongly distorts the intensity distributions, (Fig. 7.1B). While for small bleaching rates the distribution shows a clear local maximum (black solid line in

Fig. 7.1B), the distribution follows an exponential decay for fast bleaching (black dotted grey line in Fig. 7.1B).

7.3.2 Robust intensity distributions

For both blinking and bleaching the shape of the distributions changes the most with varying k and k_{bl} when the time scales for blinking $(1/k)$ and bleaching $(1/k_{\mathrm{bl}})$ are comparable to the illumination time T.

Since both k and k_{bl} sensitively depend on many experimental parameters (illumination intensity, local pH, ...) the observed variability in the intensity distributions prevents a robust assessment of molecule numbers. Along the lines of ideas developed by Cognet et al. (107), we therefore propose to use long illumination times T such that $T \gg 1/k \gg 1/k_{\mathrm{bl}}$.

In that case the intensity distribution assumes a very simple form, which is independent of the value of k. In Sec. 7.C.1 we show that for large T the distribution of the number n of photons detected during time T is

$$p(n;N) = \frac{1}{N}\left(1+\frac{1}{N}\right)^{-(n+1)} \tag{7.1}$$

where N is the mean number of photons detected. $N = \eta_d \eta_o \bar{I} k_{\mathrm{bl}}^{-1}$ where η_d and η_o are the quantum yield of the detector and the detection efficiency of the imaging optics, respectively. This intensity distribution is not influenced by blinking and depends on k_{bl} in a defined way, making it a good starting point for the measurement of molecule intensities. Below we show that the intensities of single YFP molecules follow this distribution if T is sufficiently large.

The distribution of the intensity of a dimer (i.e. two fluorophores in a diffraction limited spot) $p_2(n;N)$ is obtained from the convolution of Eq. 7.1 with itself (121):

$$p_2(n;N) = \sum_{n'=0}^{\infty} p(n-n')p(n') = \frac{n+1}{(1+N)^2}\left(1+\frac{1}{N}\right)^{-n} \tag{7.2}$$

Continued convolution with $p(n;N))$ (Eq. 7.1) gives the distribution for higher multimers, like e.g. a trimer

$$p_3(n;N) = \sum_{n'=0}^{\infty} p(n-n')p_2(n') = \frac{(n+1)\left((n/2)+1\right)}{(1+N)^3}\left(1+\frac{1}{N}\right)^{-n} \tag{7.3}$$

In Fig. 7.2 the intensity distributions for a monomer, dimer and trimer with the same mean number of detected photons N (per fluorophore) are compared.

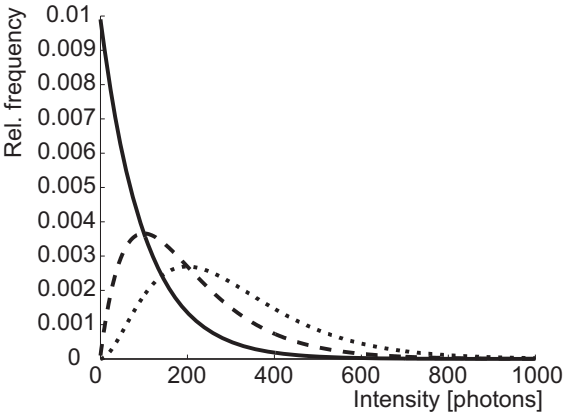

Figure 7.2
Intensity distribution $p(n;N)$) for the monomer (solid line) given by Eq. 7.1 and for the dimer (dashed line) and trimer (dotted line) obtained from convolution of $p(n;N)$) with itself. The mean number of detected photons is $N = 100$ in all cases.

In principle one could use the distributions derived so far to fit a measured intensity distribution and determine the fraction of monomers, dimers, etc. However, experimental factors modulate the measured intensity distributions as detailed in the subsequent section.

7.3.3 Detection probability

In the preceding section we showed that the intensity distribution of a single fluorophore follows an exponential decay for long illumination times T (Eq. 7.1), which means that a significant fraction of the molecules has a very small intensity. However, such dim molecules cannot be detected due to the presence of noise. This experimental noise, which originates from photon counting statistics, the detection apparatus and background autofluorescence of the cell, is unavoidable. To increase the signal to noise ratio (SNR) the acquired image is filtered with a Gaussian of width w, which should equal the width of the present signals, as prescribed by optimal filtering theory (see Sec. 7.B). Still, even after filtering, a threshold has to be defined to distinguish noise from a real single-molecule signal: only those pixels that exceed the noise by a threshold factor t (i.e. SNR $> t$) are considered to be part of potential single-molecule signals. If the threshold factor t is chosen too large, few single molecules will be detected, if it is too small, however, noise will be falsely identified as

single-molecule signals.

To quantify the influence of thresholding on intensity distributions we derived in Sec. 7.B the probability to detect a single molecule signal of width w and integrated intensity A at a noise level σ and threshold t

$$p_{\text{det}}^{\max}(A;\sigma,w,t) = \frac{1}{2}\left(1+\text{erf}\left(\frac{A}{\sqrt{8\pi}\sigma w}-\frac{t}{\sqrt{2}}\right)\right) \tag{7.4}$$

where t is the threshold imposed on the SNR after filtering the image with a Gaussian filter of width w. Fig. 7.3A and Fig. 7.3B show very good agreement between this theoretical result and simulated data. The slight systematic underestimation of the simulated detection probability is probably due to the fact that only the maximum (i.e. brightest pixel) of a single-molecule signal is considered in the model (Sec. 7.B). This pixel has the highest chance to be detected (i.e. to exceed the threshold) in the presence of noise. Other, adjacent pixels, which belong to the same single-molecule signal and by definition are less bright, also slightly contribute to the detection probability of the whole signal. Their contribution has been neglected in the derived model.

With the help of Eq. 7.4 we can now theoretically determine the shape of *measured* intensity distributions. Measured intensity distributions are products of the distribution of the emitted intensities (Eq. 7.1) and the detection probability (Eq. 7.4). An example for such a distribution is given in Fig. 7.3C. Since the detection probability goes to 0 for small intensities, measured intensity distributions always have a peak at finite intensities, despite the fact that the underlying distribution of emitted intensities is *maximal* for small intensities. In Sec. 7.3.4 we will show that experimentally determined intensity distributions indeed have the predicted shape shown in Fig. 7.3C.

To find the optimal value for the threshold t we have to balance the number of rejected single-molecule signals, which increases with t, with the number of false positives (i.e. noise accepted as signal), which decreases with t. To predict the number of false positives we calculated the probability $p_{\text{false}}(t)$ to falsely detect a single molecule in a pixel with only noise at threshold t (see also Sec. 7.B).

$$p_{\text{false}}(t) = \frac{1}{2}\left(1+\text{erf}\left(-\frac{t}{\sqrt{2}}\right)\right) \tag{7.5}$$

In an image with M pixels roughly $M \cdot p_{\text{false}}(t)$ noise peaks are falsely detected as single molecules signals, if the pixels can be considered in-

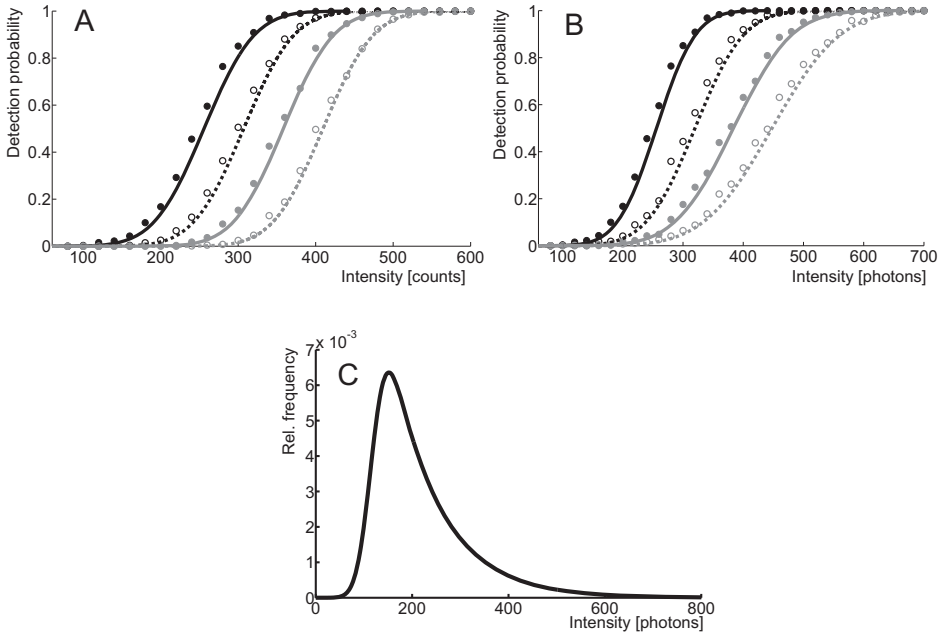

Figure 7.3
A) Detection probability determined from simulations for a threshold t of 5 (black solid circles),6 (black open circles), 7 (grey solid circles) and 8 (grey open circles) respectively at noise level $\sigma = 20$ and signal width $w = 0.7pxl$. The lines give the detection probability predicted by Eq. 7.4 for a threshold t of 5 (black solid line),6 (black dashed line), 7 (grey solid line) and 8 (grey dashed line). B) Detection probability determined from simulations for noise levels σ of 20 (black solid circles), 25 (black open circles), 30 (grey solid circles) and 35 (grey open circles) respectively at a threshold t of 5 and signal width $w = 0.7pxl$. The lines give the detection probability predicted by Eq. 7.4 for a noise level σ of 20 (black solid line), 25 (black dashed line),30 (grey solid line) and 35 (grey dashed line) C) Complete intensity distribution for a single fluorophore (monomer) given by Eq. 7.46. This distribution is calculated as the product of the monomer distribution (Eq. 7.1) and the detection probability (Eq. 7.4) and subsequent normalization to 1. The assumed parameters are: number of detected photons $N = 100$, noise level $\sigma = 10$, threshold $t = 5$, signal width $w = 0.7pxl$.

dependent. We define ϵ as the maximal allowed ratio of false positives $(M \cdot p_{\text{false}}(t))$ to all detected signals N_{signals} : $Mp_{\text{false}}(t) < \epsilon N_{\text{signals}}$. This definition leads to an upper limit for t: $t > \sqrt{2}\,\text{erf}^{-1}\,(1 - 2\epsilon(N_{\text{signals}}/M))$ For example with $\epsilon = 0.01$, $M = 50^2$ and typically $N_{\text{signals}} = 10$ we get $t \gtrsim 4$. Since the total number of detected signals depends on N_{signals} and will decrease with t, t should not be chosen too large to avoid loss of data.

7.3.4 Experimental validation

To verify our theoretical derivations we performed single-molecule fluorescence experiments on EYFP molecules in living CHO cells. EYFP was tagged to the membrane anchor of H-Ras, which resulted in a membrane localization of the EYFP and thereby greatly facilitated the measurements. We measured the intensities of ≈ 23000 single-molecule signals for an illumination time of $T = 50ms$. We estimated from earlier experiments (see Sec. 7.2) that an illumination time of $50ms$ is several times bigger than the bleaching time expected at the illumination intensities used. To ensure that there was only one EYFP in each diffraction limited spot the signal density was kept very low ($\rho < 0.2\mu m^{-2}$). Fig. 7.4A shows the obtained intensity distributions. As expected from theory (Eq. 7.1) the distribution approximately follows an exponential decay. The measured mean number of photons of $N = 837 \pm 3$ is roughly in agreement with the value expected from earlier results on EYFP (see Sec. 7.2), which was 967. The single molecule signals were obtained with a threshold factor of $t = 3$ which minimizes the influence of the detection probability on the intensity distribution. However, only signal intensities above $1000\,photons$ were used, which roughly corresponds to a threshold of $t = 15$. Hence, the number of false positives due to noise is minimized. To show that thresholding influences intensity distributions in the way we derived above, we present in Fig. 7.4B several intensity distributions based on the same raw data, which differ by the threshold factor t. In principle, these distributions should be the product of the distribution of emitted photons (Eq. 7.1), with the mean number of photons N determined above, see Fig. 7.4A), and the detection probability at a certain threshold factor t and noise level σ. This product is given in Eq. 7.46. The noise level σ is not constant in our experiments but varies between images and, more strongly, between different cells. Therefore the detection probability has to be determined as the average over all acquired images using the noise levels in those images

$$\bar{p}_{\mathrm{det}}^{\max}(A; w, t) = \frac{1}{N_{\mathrm{images}}} \sum_{i}^{N_{\mathrm{images}}} p_{\mathrm{det}}^{\max}(A; \sigma_i, w, t) \tag{7.6}$$

where σ_i is the noise level in image i and N_{images} is the number of acquired images.

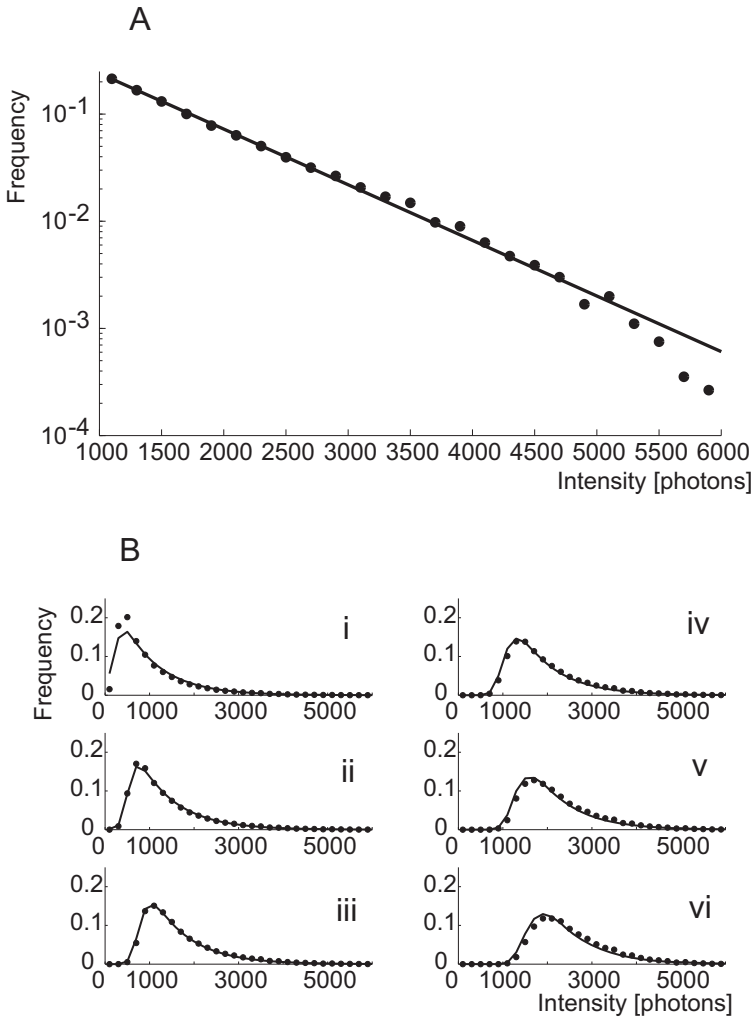

Figure 7.4
Intensity distribution of single EYFPs on the membrane of living CHO cells. A) Experimental intensity distribution of the intensities of 22615 single-molecule fluorescence signals (solid circles). A one-parameter fit of Eq. 7.1 (solid line) to the experimental intensity distribution for single EYFPs gives $N = 837 \pm 3$ photons. The width w of the single molecule signals was restricted to the interval $0.64 - 0.81\,pixel$, the width of the Gaussian filter was $r = 0.72\,pxl$, $pxl = 220nm$. The threshold factor was $t = 3$ and only signals with an intensity bigger than $1000\,photons$ were used. B) Influence of thresholding on the shape of the intensity distributions. The raw data is the same as in panel A, but the threshold factor t is varied. $t = 5, 10, 15, 20, 25, 30$ from i) to vi) respectively. The experimental data (solid circles) is compared to the full theoretical distribution Eq. 7.46 (solid line). The mean number of detected photons N was determined above from a restricted data set (see panel A), the detection probability was calculated by integration over the noise levels found in the raw images.)

The noise levels were estimated as described in Sec. 7.B. The resulting detection probability $\bar{p}_{\mathrm{det}}^{\mathrm{max}}(A; w, t)$ is then multiplied with the distribution of emitted intensities (Eq. 7.1) to get the full theoretical description of the measured intensity distributions (solid lines in Fig. 7.4B). The measured distributions nicely follow the theoretical expressions determined in this way.

7.3.5 Clustering due to membrane heterogeneity

The measurements presented in the previous section were performed at low signal densities ($< 0.2\mu m^{-2}$) to ensure that there was only one single YFP molecule per diffraction limited spot. In subsequent measurements at increased signal densities we observed that the intensity distributions are shifted to higher intensities, see Fig. 7.5. This shift is due to the presence of several molecules in one diffraction limited spot (multimers). To quantify the amounts of monomers, dimers and higher multimers we compare the intensity distributions at various densities of single molecule signals to the intensity distribution of the monomer(121). For simplicity we assume here that only monomers and dimers are present and describe the measured distributions as a weighted sum of the intensity distribution of a monomer $p(n; N)$ and a dimer $p_2(n; N)$

$$p_{\mathrm{total}}(n) = \alpha p(n; N) + (1 - \alpha) p_2(n; N). \tag{7.7}$$

$p(n; N)$ and $p_2(n; N)$ were presented in Sec. 7.3.2 (see Sec. 7.C.1 for the derivation). N, the average number of detected photons, is determined from the monomer distribution as described in the previous section which leaves the fraction of monomers α as a free fit parameter. Fig. 7.5 shows an example for a fit of this model to experimental data. We find that α decreases quickly with increasing signal density, see inset of Fig. 7.5. Strikingly, there are many more dimers than expected for a uniform distribution of molecules at such low densities ($< 1\mu m^{-2}$, see the model derived in Sec. 7.C.1). In fact, this is in agreement with earlier results on the used construct (150) where a certain fraction of the molecules were shown to exhibit confined diffusion in $\approx 200nm$ domains. Consequently, one would expect colocalization of molecules even at low densities. This result has important implications. On the one hand intensity can be used as a readout for membrane heterogeneity. On the other hand, this result shows that membrane heterogeneity has to be taken into account when

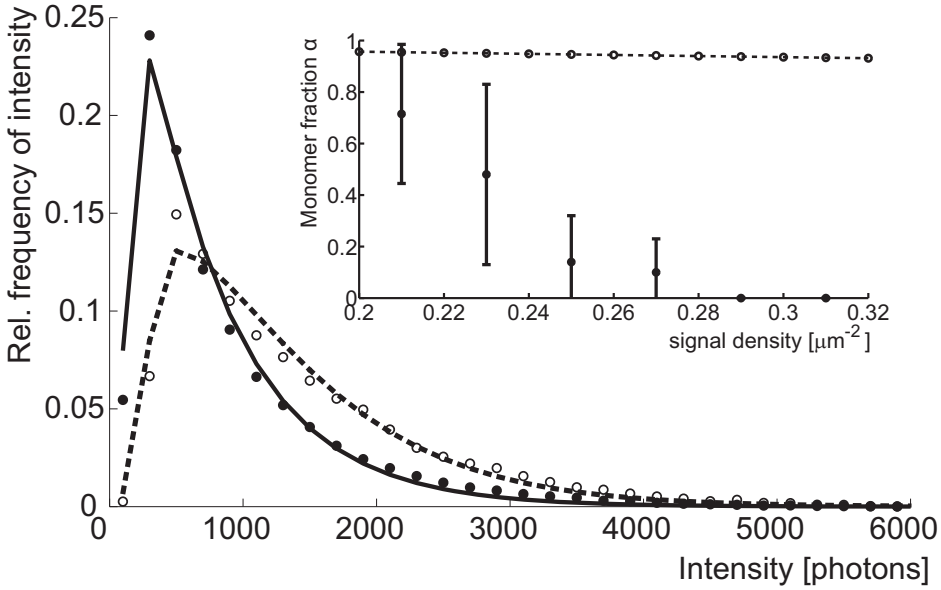

Figure 7.5
Influence of membrane heterogeneity on the intensity distribution of EYFP in the membrane of living CHO cells. The solid circles give the intensity distribution at low signal densities ($< 0.2\mu m^{-2}$), already shown in Fig. 7.4. The solid line is the corresponding theoretically expected distribution obtained as detailed above. This distribution is compared to one taken at a signal density of $0.25\mu m^{-2}$ (open circles). The threshold factor is $t = 3$ for both distributions. The visible shift to higher intensities with increased signal densities is due to the presence of multimers (i.e. several molecules colocalized in a diffraction limited spot). A fit to Eq. 7.7 (dashed line) gives that at this density the fraction of monomers is $\alpha = 0.14$. Inset: Fraction of monomers versus density of single molecule signals (solid circles). The error was determined as standard deviation calculated from all data sets used in a certain bin. The open circles show the theoretically expected monomer fraction for a uniform distribution of molecules, see Eq. 7.49.

true complex formation (in contrast to mere colocalization) is to be measured.

7.3.6 Limitations and errors

The method presented here requires long illumination times T. Mobile molecules in living cells can be observed only for a limited time, which sets an upper boundary for T. In our experiments this did not present a severe limitation but it might be e.g. for molecules which diffuse more

quickly.

Secondly, the diffusion of molecules during illumination broadens the observed single-molecule signals and might render fitting the signals difficult, especially if signal density is high. To quantify the influence of diffusion we estimate the expected signal width and the broadening caused by diffusion. For a single YFP with emission maximum at $\lambda = 527nm$ and a numerical aperture of the microscope objective of $NA = 1.4$ we would expect the full width half maximum of a signal to be $\approx 2 \cdot \lambda/(2\mathrm{NA}) = 376nm$ which corresponds to a signal width of $w_{\mathrm{PSF}} = 160nm$. The signal width is defined as the width of a gaussian which is fit the point spread function (PSF) of the microscope, see Sec. 7.B. Additionally, the movement of the molecule in the period before bleaching leads to a broadening of the peak. With a bleaching time of $\tau_{\mathrm{bl}} = 10.4ms$ (see Sec. 7.2) and a diffusion coefficient of roughly $D = 0.8\mu m^2/s$ (150), the total signal width is, according to (221), $w = \sqrt{(w_{\mathrm{PSF}})^2 + D\tau_{\mathrm{bl}}} = 178nm$. This broadening effect is small and does not hamper the applicability of the method in our experiments. In general, the signal width puts an upper limit on the density of signals: two molecules which come closer to each other than w_{PSF} cannot be resolved.

In addition to long illumination times T our method also requires that the bleaching time scale is much longer than time scale of any blinking (see Sec. 7.C.1). In our experiments blinking is obviously fast enough because we observe the exponential decay predicted for well separated time scales. Different fluorophores might have blinking rates which are comparable to the bleaching rate. Such molecules cannot be used with our method.

We estimated the error for the measurement of the monomer fraction in Sec. 7.3.5 with the help of simulations. In particular, we assume that the mean number of emitted photons N is known and randomly generate N_{signals} signals with intensities drawn from the distribution Eq. 7.7. The randomly drawn intensities are binned in equally sized bins and the resulting distribution is normalized. This distribution is fit with Eq. 7.7 with α as the only free fit parameter. The whole procedure is repeated 100 times for each set of parameters and the error is determined as the standard deviation $\Delta\alpha$ over the 100 values obtained for α. Fig. 7.6 compares the influence of the experimental parameters on the relative error of α determined by these simulations. Fig. 7.6A shows that $\Delta\alpha$ approximately scales like $\propto 1/sqrtN_{\mathrm{signals}}$, which means that the accuracy can always be increased by measuring more signals. As expected the relative

error of α increases with decreasing α. In Fig. 7.6B we show that $\Delta\alpha$ scales approximately like $\propto 1/\alpha$. If the bin size is sufficiently small, $\Delta\alpha$ is independent of the mean number of emitted photons N, see Fig. 7.6C. If the the bin size is too large, all signals will fall in one (or a few bin), which makes fitting of the distribution impossible. Conversely, $\Delta\alpha$ is independent of the bin size, see Fig. 7.6C, unless the mean number of emitted photons N is small. In that case the bin size has to be chosen sufficiently small. The relative error we expect for the parameters of the experiment

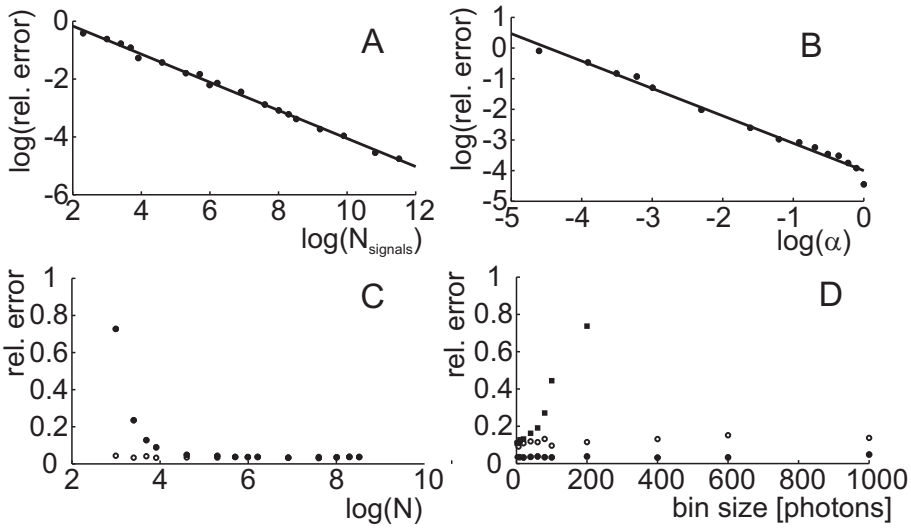

Figure 7.6
Accuracy of the measurement of the monomer fraction α determined from simulations. A) Dependence of the relative error of α (solid circles) on the number of detected signals $N_{signals}$. A linear fit to the data in the log-log graph (solid line) has a slope of -0.49. The bin size is 200, $N = 500$, $\alpha = 0.5$. B) Dependence of the relative error of α (solid circles) on α. A linear fit to the data in the log-log graph (solid line) has a slope of -0.89. The bin size is 200, $N = 500$, $N_{signals} = 5000$. C) Dependence of the relative error of α on the mean number of emitted photons N. Solid circles correspond to simulations with a bin size of 200, the open circles to simulations with a bin size of 20. $N_{signals} = 5000$, $\alpha = 0.5$ D) Dependence of the relative error of α on the bin size b. The solid circles correspond to simulations with $N_{signals} = 5000$ and $N = 500$, the open circles to simulations with $N_{signals} = 500$ and $N = 500$ and the solid squares to simulations with $N_{signals} = 500$ and $N = 20$

presented above ($N = 837$, bin size 200, $N_{\text{signals}} \approx 20000$) ranges from

0.08 for $\alpha = 0.1$ to 0.01 for $\alpha = 1$. The errors we observe in experiments are much larger, see inset of Fig. 7.5. The reason for that is twofold: Our simulations do not account for biological variability which is probably considerable. Secondly, we combined data sets with different signal densities in bins with a bin size of $0.02\mu m^{-2}$, see inset of Fig. 7.5. Since the monomer fraction is decreasing quickly with signal density, heterogeneity within a bin contributes significantly to the reported error.

7.4 Conclusion

We have shown that, despite the many challenges presented by living cells, single-molecule intensities can be measured in a robust way. We have put such measurements on firm mathematical and analytical grounds. In particular, we have analyzed and minimized the influence of noise and the noise related thresholding procedure. We have shown experimentally that using long illumination times overcomes the problems arising from single-molecule blinking and bleaching. The resulting intensity distributions are well described by a simple, quantitative model. Finally, we have quantified the amount of clustering of a certain membrane protein due to membrane heterogeneity. We hope that our work paves the way for accurate *in vivo* measurements of protein complex and cluster formation. In particular, since our method is based on imaging, it should allow for the construction of cell-wide stoichiometry maps.

7.5 Acknowledgement

The authors thank Dr. S. Olthuis-Meunier for preparing the cell lines, Joe Brzostowski and Susan Pierce for providing microscopy and other lab equipment. This work was supported by funds from the Netherlands Organization for Scientific Research (NWO-FOM) within the program on Material Properties of Biological Assemblies (FOM-L1707M)

7.A Image processing and analysis

7.A.1 Definitions

The image taken by the CCD camera $s(x, y)$ is composed of the following contributions

$$s(x, y) = \sum_i G_i(x, y) + B(x, y) + n_0(x, y) \tag{7.8}$$

where G_i are the single-molecule signals at positions (x_i, y_i), B is the autofluorescent backgroundand, $n_0(x, y)$ is a constant Gaussian noise (e.g. read-out noise).

Since the signal is obtained by counting single-photons, the probability for a certain signal $G_i(x, y)$ coming from a single molecule is given by a Poisson distribution with mean $g_i(x, y)$

$$p(G_i(x, y)) = \frac{g_i(x, y)^{G_i(x,y)}}{G_i(x, y)!} \cdot \exp\left[-g_i(x, y)\right] \tag{7.9}$$

where $g_i(x, y)$ is the point spread function (PSF) of the microscope centered at molecule position (x_i, y_i). The PSF, which is ideally given by a 2D Airy-function, is most commonly approximated by a 2D Gaussian

$$g_i(x, y) = \frac{A}{2\pi w^2} \cdot \exp\left[\frac{(x - x_i)^2 + (y - y_i)^2}{2w^2}\right] \tag{7.10}$$

Correspondingly the contribution of the background is given by

$$p(B(x, y)) = \frac{b(x, y)^{B(x,y)}}{B(x, y)!} \cdot \exp\left[-b(x, y)\right] \tag{7.11}$$

where $b(x, y)$ is given by the product of the laser profile and the autofluorescence coming from the cell.

The constant noise contribution $n_0(x, y)$ is assumed to be a normal distribution with mean 0 and standard deviation σ_0.

7.A.2 Autofluorescent background subtraction

We assume that $b(x, y)$ only fluctuates weakly on the length scale of the signal width w at the time scale of the time lag between frames Δt. Therefore, a low pass Fourier filtering of single images or a sliding minimum of an

image stack gives approximately $b(x, y)$ (see below). After subtraction of $b(x, y)$ the background is approximately flat. We further assume that the influence of the background is so weak that the remaining shot-noise can be described by a normal distribution with standard deviation $\sqrt{b(x, y)}$. These simplifications lead to the following estimation of the signal

$$s(x, y) = \sum_i G_i(x, y) + n(x, y) \tag{7.12}$$

where $n(x, y)$ is distributed normally with mean 0 and standard deviation $\sigma(x, y) = \sqrt{\sigma_0^2 + b(x, y)}$. Since we assumed that $b(x, y)$ is approximately constant on the length scale w, we can also write $\sigma = \sqrt{\sigma_0^2 + \overline{b}}$ where $\overline{b}$ is the background averaged over the whole image.

i) Sliding minimum filter Since experimental parameters are chosen such that single molecule signals only appear in a single frame (single molecules are bleached during integration time T) and the autofluorescent background bleaches slowly, a sliding minimum filter can be applied. In the time trace of each pixel the value of a pixel at a certain time t is exchanged for the local minimum (in time). The local minimum is defined as the minimal pixel value in a time window of width τ around t. By using a sliding minimum instead of a sliding mean, the intensity of single-molecule signals remains unchanged. The remaining noise, however, is biased towards positive values.

ii) High pass FFT filter Since the width of single molecule signals is smaller than typical length scales of autofluorescent structures, the background can be diminished by high pas Fourier filtering. As described in the following, the filter does not change positions or intensities of single molecule signals, if applied correctly. It is less effective than the sliding minimum filter but causes no bias in the remaining noise.

For the FFT filter we can estimate the influence of the filtering on the signal intensities. As already mentioned, the PSF of the microscope is approximated by a Gaussian

$$g_i(x, y) = \frac{A}{2\pi w^2} \cdot \exp\left[\frac{(x - \hat{x}_i)^2 + (y - \hat{y}_i)^2}{2w^2}\right] \tag{7.13}$$

The Fourier transform results in a modified gaussian

$$g_i(k_x, k_y) = e^{-\frac{(2\pi w)^2 (k_x^2 + k_y^2)}{2}} e^{-i2\pi(\hat{x}_i k_x + \hat{y}_i k_y)} \tag{7.14}$$

Note that the information about the width of the signal is reflected in the amplitude, and the information about the position is reflected in the phase. By changing the amplitude only, the positions are preserved. Therefore

$$h(k_x, k_y) = e^{-\frac{(2\pi v)^2(k_x^2+k_y^2)}{2}} \Rightarrow \tilde{g}(k) = g_i(k_x, k_y) \cdot h(k_x, k_y) \tag{7.15}$$

Inverse Fourier transform yields

$$\hat{g}_i(x, y) = \frac{A}{2\pi(v^2 + w^2)} \cdot \exp\left[\frac{(x - \hat{x}_i)^2 + (y - \hat{y}_i)^2}{2(v^2 + w^2)}\right] \tag{7.16}$$

The filtered image is subsequently subtracted from the original image. $f(x, y) = g_i(x, y) - \hat{g}_i(x, y)$ If $v \gg w$ then the form of the signal is hardly distorted. Since the width of the PSF depends approximately quadratically on the z-position ($w(z) = w_0(1 + cz^2)$), filtering is more effective for out of focus light For $z \neq 0$ the PSF and $\hat{g}_i(x, y)$ (the filtered PSF) differ less (since $w(z) > w(0)$), which results in a reduction of out of focus signal.

As an estimate for the change in intensity due to filtering the relative signal height after filtering is calculated

$$\frac{\text{height of filtered signal}}{\text{original signal height}} = \frac{\frac{1}{w^2} - \frac{1}{w^2+v^2}}{\frac{1}{w^2}} = 1 - \left(1 + \left(\frac{v}{w}\right)^2\right)^{-1} \tag{7.17}$$

If v is too small, signal height and intensity decrease too much. For $v = 5w$ the height is still 96% of the original signal height, so $v > 5w$ is a reasonable limit. If v is too big, however, the filtering effect will be small.

The different filters are compared in Fig. 7.7.

It is evident from Fig. 7.7 that the sliding minimum filter increases the signal to background ratio most. However, background estimation methods should be compared as well in terms of noise, offset and signal width.

Fig. 7.8 shows that the FFT filter results in an offset that is symmetrically distributed around 0, the sliding minimum filter, however, causes a systematic shift in offset.

Also in terms of noise the sliding minimum filter is superior, which is shown in Fig. 7.9. Images filtered with the FFT method show a systematically higher noise level.

Furthermore, a difference in width of the found signals is evident in Fig. 7.10. The maximum at $1.7 pixel$ corresponds to the signal from single

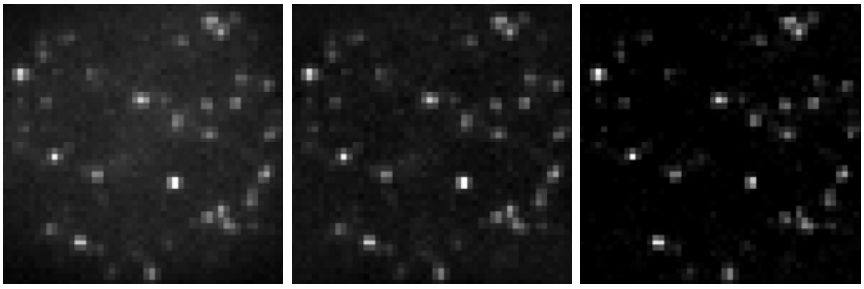

Figure 7.7
From left to right: raw image, image filtered with high pass FFT filter ($v = 5\,\mathrm{pxl}$), image stack filtered with sliding minimum filter (window size: 2 frames)

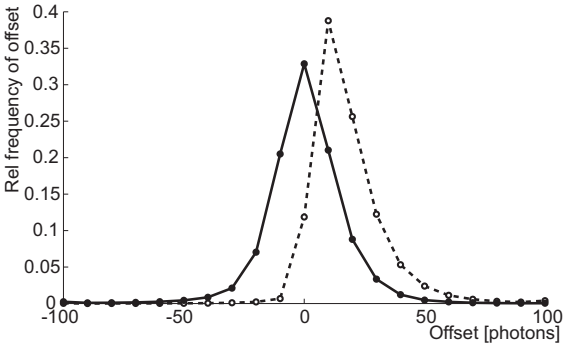

Figure 7.8
Offset distribution. Solid line: FFT filter, Dashed line: sliding minimum, threshold $t = 5$

molecules, see Sec. 7.3.6 the additional maximum around $1pxl$ is due to noise that is falsely identified as a single molecule (see 7.B). The FFT filter clearly results in less false positive detections.

Although the sliding minimum filter gives a better signal to background ratio and less noise we chose to employ in the following only the FFT filter. It outperforms the other method in terms of offset and false positive detections.

7.B Detection probability

After background subtraction the image is filtered with an appropriate filter to optimize the signal to noise ratio (SNR). According to optimal filtering theory (142), the filter should be identical to the signal. Therefore

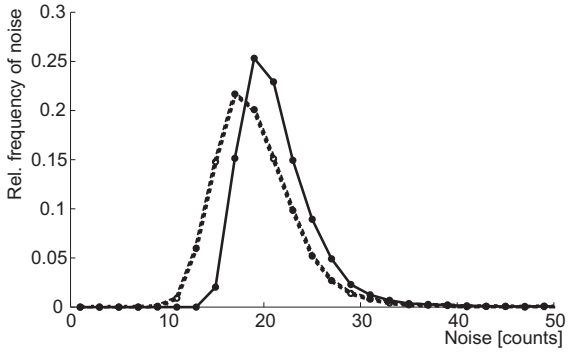

Figure 7.9
Noise distribution. Solid line: FFT filter, Dashed line: sliding minimum, threshold $t = 5$

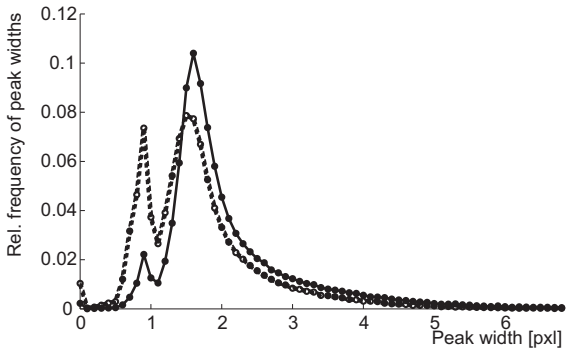

Figure 7.10
Signal width distribution. Solid line: FFT filter, Dashed line: sliding minimum, threshold $t = 5$

we filter with a Gaussian of width r. The resulting SNR is

$$\mathrm{SNR} = \frac{Ar}{\sigma\sqrt{\pi}(r^2 + w^2)} \tag{7.18}$$

where $\tilde{A} = A/(2\pi(r^2+w^2))$ is the maximum of the signal (= signal height) after filtering and $\tilde{\sigma} = (\sigma/2\sqrt{\pi}r)$ is the noise level after filtering. The noise distribution is still normal (with mean 0 and standard deviation $\tilde{\sigma}$). For the optimal choice $r = w$

$$\mathrm{SNR}_{\max} = \frac{A}{\sigma\sqrt{\pi}w} \tag{7.19}$$

. Notice that σ has the units [counts/pixel] so that the SNR is dimensionless. In order to distinguish between noise and signal, a threshold t is introduced. Only those pixels in the filtered image whose brightness

$I(x, y)$ exceeds the (filtered) noise level by a factor t are treated as potential molecule positions.

$$I(x, y) > t \cdot \frac{\sigma}{2\sqrt{\pi}r} \equiv t\tilde{\sigma} \tag{7.20}$$

Fig. 7.11 illustrates the thresholding procedure. Subimages of the unfil-

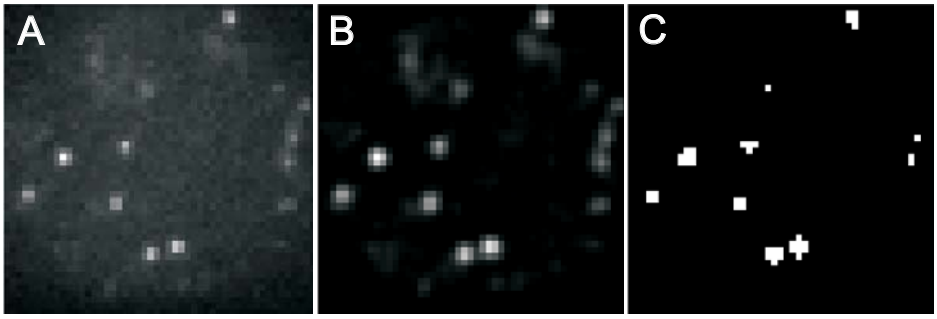

Figure 7.11
Thresholding procedure A) Raw image of single EYFP-CAAX molecules on the membrane of a living CHO cell. The noise is approximately $\sigma = 26$. The linear grey scale ranges from 0 counts (black) to 1002 counts (white). B) Image after background subtraction and filtering (i.e. correlation) with gaussian of width $0.7pxl$, $pxl = 220nm$. The linear grey scale ranges from 0 counts (black) to 1473 counts (white). C) Binary image after thresholding. White pixels correspond to pixels whose value exceeds the threshold at threshold factor $t = 10$.

tered image around the pixels which were identified as potential molecule positions are then fit with the sum of a Gaussian and a constant offset $g(x, y) + \text{off}$.

We define the detection probability $p_{\text{det}}(\sigma, A)$ as the probability that for a given noise level σ a pixel with brightness $\tilde{A}$ exceeds the threshold described, so

$$p_{\text{det}}(\sigma, A) = \int_{t\tilde{\sigma}}^{\infty} da\, \frac{1}{2\pi\tilde{\sigma}} \exp\left[-\frac{(a - (\tilde{A} + \text{off}))^2}{2\tilde{\sigma}^2}\right] \tag{7.21}$$

Integration gives

$$p_{\text{det}}(\sigma, A) = \frac{1}{2}\left(1 + \text{erf}\left(\frac{1}{\sqrt{2}}\left(\frac{\tilde{A} + \text{off}}{\tilde{\sigma}} - t\right)\right)\right) \tag{7.22}$$

$$= \frac{1}{2}\left(1 + \text{erf}\left(\frac{1}{\sqrt{2}}\left(\frac{A + \text{off}'}{\sigma} \cdot \frac{r}{\sqrt{\pi}(r^2 + w^2)} - t\right)\right)\right) \tag{7.23}$$

with the error function

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x dy\, \exp(-y^2) \tag{7.24}$$

and $\mathrm{off}' = \mathrm{off} \cdot 2\pi(r^2+w^2)$. For a given σ, A and w the detection probability is maximal for the choice $r = w$, i.e. when the SNR is maximal. If we assume $\mathrm{off} \ll A$

$$p_{\mathrm{det}}^{\mathrm{max}}(\sigma, A) = \frac{1}{2}\left(1 + \mathrm{erf}\left(\frac{A}{\sqrt{8\pi}\sigma w} - \frac{t}{\sqrt{2}}\right)\right) \tag{7.25}$$

Obviously the integrated signal intensity A should be as big as possible and noise level σ and signal width w should be as small as possible to maximize the detection probability. The threshold parameter t should be as small as possible to detect as many molecules as possible. However, with decreasing t the probability $p_{\mathrm{false}}(\sigma)$ that a noise peak is falsely detected as a signal is growing. In complete analogy to the above derivation the probability $p_{\mathrm{false}}(\sigma)$ is

$$p_{\mathrm{false}}(t) = \frac{1}{2}\left(1 + \mathrm{erf}\left(-\frac{t}{\sqrt{2}}\right)\right) \tag{7.26}$$

In an image with M pixels roughly $M \cdot p_{\mathrm{false}}(t)$ noise peaks are falsely detected as signals. As shown in Fig. 7.12 this is a good estimation for a filter width of $r = 0.7pxl$. For wider Gaussian filters the theoretical formula overestimate the number of false positives. In this case, the pixels are not independent any more, so there are fewer false positive detections than expected. Also for small thresholds t, theory and simulation differ. Here the number of false detections becomes so high, that they overlap and cannot be distinguished anymore. In any case the theoretical expression Eq. 7.26 is an upper bound for the number of false positives. It can therefore be used to safely estimate the threshold t required for a certain maximal number of false positives.

7.C Intensity distribution

7.C.1 Mandel theory

According to the semi-classical theory by Mandel (220) which is used also in Photon Counting Histograms (PCH) (141) the probability to find n

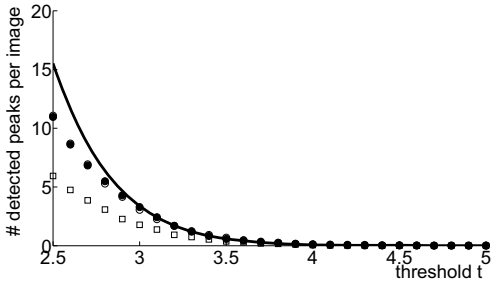

Figure 7.12
Number of signals detected in an image with only noise, image size 50x50 pixels. Filled circles: noise level $\sigma = 50$, filter width $r = 1.7pxl$; open circles: noise level $\sigma = 20$, filter width $r = 0.7pxl$; open squares: noise level $\sigma = 20$, filter width $r = 1.1pxl$, solid line: number of false positives expected from Eq. 7.26

photons at time t with an integration/illumination time T and detector quantum yield η_d is:

$$p(n,t,T) = \int_0^\infty \frac{(\eta_d W(t))^n \, e^{-\eta_d W(t)}}{n!} p_{\mathrm{inc}}(W(t),T) dW(t) \tag{7.27}$$

$$W(t) = \eta_o \int_t^{T+t} \int_A I(\mathbf{r},t) d\mathbf{A} dt \tag{7.28}$$

$W(t)$ is the number of incident photons falling on the detector given a photon emission rate $I(\mathbf{r},t)$, a detection efficiency of the imaging optics η_o and a detector area A. $p_{\mathrm{inc}}(W(t),T)$ is the probability distribution of the number of incident photons $W(t)$ at time t given an illumination time T.

Here integration is performed not over the whole detector (which would be the whole CCD chip) but over the area on the chip that is covered by a single-molecule signal so that

$$W(t) = \eta_o \int_t^{T+t} I_{\mathrm{sm}}(t) dt \tag{7.29}$$

where $I_{\mathrm{sm}}(t)$ is the photon emission rate coming from a single molecule. T is assumed to be so long that $W(t)$ is independent of the time of measurement t: $W(t) \equiv W$. Note that this is the opposite of the limit used in PCH (141) where very short integration/illumination times T are used so that the intensity fluctuations govern the counting statistics. For long enough illumination times T the probability distribution $p_{\mathrm{inc}}(W,T)$ can be obtained from the probability $p_{\mathrm{fluor}}(t_{\mathrm{on}},T)$ that a molecule is "on", i.e. emitting photons, for a period t_{on} during the integration time T. If $\bar{I}$ is

the average number of photons emitted during periods when the molecule is "on", $p_{\mathrm{inc}}(W,T)$ is

$$p_{\mathrm{inc}}(W,T) = \frac{\eta_o}{\bar{I}}\, p_{\mathrm{fluor}}\left(t_{\mathrm{on}} = \frac{W}{\bar{I}}, T\right) \tag{7.30}$$

The distribution of the "on" times, $p_{\mathrm{fluor}}(t_{\mathrm{on}},T)$, depends on the photophysics of the molecule. In the following two paragraphs a model for $p_{\mathrm{fluor}}(t_{\mathrm{on}},T)$ that includes blinking and bleaching is derived.

7.C.2 2-state model

The 2-state model with a fluorescent "on" and a non-fluorescent "off" state was discussed in (222). In this model the molecule can switch reversibly between the "on" state and the "off" state but it never bleaches. The corresponding probability distribution for the times t_{on} in the "on" state $p_{\mathrm{fluor}}(t_{\mathrm{on}},T)$ is the sum of contributions from 4 different kinds of fluorescence traces. 1. A molecule that starts in the "on" state can stay "on" during the whole illumination time T ($t_{\mathrm{on}} = T$). The probability for such a trace is $p_{\mathrm{on}}(0)\exp(-k_{\mathrm{off}}T)$ where $p_{\mathrm{on}}(0)$ is the probability that the molecule is initially "on" and k_{off} is the rate for switching from "on" to "off". 2. A molecule that starts in the "off" state can stay "off" during the whole illumination time T ($t_{\mathrm{on}} = 0$) . The probability for such a trace is $p_{\mathrm{off}}(0)\exp(-k_{\mathrm{on}}T)$ where $p_{\mathrm{off}}(0)$ is the probability that the molecule is initially "off" and k_{on} is the rate for switching from "off" to "on". Since a molecule can either be "on" or "off" initially, $p_{\mathrm{on}}(0) + p_{\mathrm{off}}(0) = 1$ must hold. 3. A molecule that starts in the "on" state can switch between "on" and "off". The probability density for those fluorescence traces is

$$\begin{aligned} p_{\mathrm{odd}}(t_{\mathrm{on}},T) = {} & k_{\mathrm{off}} \exp\left(-k_{\mathrm{off}}t_{\mathrm{on}} - k_{\mathrm{on}}t_{\mathrm{off}}\right) \\ & \times I_0\left(2\sqrt{k_{\mathrm{off}}k_{\mathrm{on}}t_{\mathrm{on}}t_{\mathrm{off}}}\right) \end{aligned} \tag{7.31}$$

for an odd number of switches and

$$\begin{aligned} p_{\mathrm{even}}(t_{\mathrm{on}},T) = {} & \sqrt{k_{\mathrm{off}}k_{\mathrm{on}}\frac{t_{\mathrm{on}}}{t_{\mathrm{off}}}} \exp\left(-k_{\mathrm{off}}t_{\mathrm{on}} - k_{\mathrm{on}}t_{\mathrm{off}}\right) \\ & \times I_1\left(2\sqrt{k_{\mathrm{off}}k_{\mathrm{on}}t_{\mathrm{on}}t_{\mathrm{off}}}\right) \end{aligned} \tag{7.32}$$

for an even number of switches. $t_{\mathrm{off}} = T - t_{\mathrm{on}}$ and I_0 and I_1 are modified Bessel-functions of the first kind of order 0 and 1 respectively, see (222).

The total probability for traces which start in the "on" state and switch between "on" and "off" is $p_{\mathrm{on}}(0)\,(p_{\mathrm{odd}}(t_{\mathrm{on}},T)+p_{\mathrm{even}}(t_{\mathrm{on}},T))$. 4. The probability density for a molecule which starts in the "off" state and switches between "off" and "on" are analogous: k_{on} is interchanged with k_{off} and t_{on} is interchanged with t_{off}. The total probability density for the "on" times t_{on} is the sum of all 4 contributions

$$
\begin{aligned}
&p(t_{\mathrm{on}},T) = p_{\mathrm{on}}(0)\\
&\times \exp(-k_{\mathrm{off}}T)\delta(T-t_{\mathrm{on}}) + p_{\mathrm{off}}(0)\exp(-k_{\mathrm{on}}T)\delta(t_{\mathrm{on}}) + \Theta(t_{\mathrm{on}})\Theta(T-t_{\mathrm{on}})\\
&\times \exp(-k_{\mathrm{off}}t_{\mathrm{on}}-k_{\mathrm{on}}t_{\mathrm{off}})\left[(p_{\mathrm{on}}(0)k_{\mathrm{off}}+p_{\mathrm{off}}(0)k_{\mathrm{on}})I_0\left(2\sqrt{k_{\mathrm{off}}k_{\mathrm{on}}t_{\mathrm{on}}t_{\mathrm{off}}}\right)\right.\\
&\left.+(p_{\mathrm{on}}(0)\sqrt{t_{\mathrm{on}}/t_{\mathrm{off}}}+p_{\mathrm{off}}(0)\sqrt{t_{\mathrm{off}}/t_{\mathrm{on}}})\sqrt{k_{\mathrm{off}}k_{\mathrm{on}}}I_1\left(2\sqrt{k_{\mathrm{off}}k_{\mathrm{on}}t_{\mathrm{on}}t_{\mathrm{off}}}\right)\right]
\end{aligned}
$$

7.C.3 3-state model

We generalize the above model presented in the previous section to a 3-state model in which the fluorophore can bleach from the "on" state to a bleached state with rate k_{bl}, see Inset to Fig. 7.1A. In this model the probability distribution $p_{\mathrm{fluor}}(t_{\mathrm{on}},T)$ is again given by the sum of 4 contributions. 1. A molecule starts in the "on" state and stays "on" during the whole illumination time T ($t_{\mathrm{on}}=T$). The probability for such traces is now $p_{\mathrm{on}}(0)\exp(-(k_{\mathrm{off}}+k_{\mathrm{bl}})T)$. 2. The probability that the molecule starts in the "off" state and stays "off" during T ($t_{\mathrm{on}}=0$) is as above: $p_{\mathrm{off}}(0)\exp(-k_{\mathrm{on}}T)$. 3. The contribution of traces that are switching and end in the "on" or "off" state without bleaching during the illumination time T is the same as contribution 3 of the 2-state model except for an additional factor $\exp(-k_{\mathrm{bl}}t_{\mathrm{on}})$. 4. A molecule can start in the "on" or "off" state and after several switching events the fluorescence trace is ended by a final bleaching event from the "on" state. The probability that a molecule is initially "on" and stays "on" until it bleaches at time t_{on} is $p_{\mathrm{on}}(0)k_{\mathrm{bl}}\exp(-(k_{\mathrm{bl}}+k_{\mathrm{off}})t_{\mathrm{on}})$. If the molecule switches, it has to switch an even number of times if it is initially "on" and an odd number of times if it is initially "off" to end the fluorescence trace in the "on" state before bleaching. The corresponding probabilities are $p_{\mathrm{on}}(0)p_{\mathrm{even}}(t_{\mathrm{on}},t')$ and $p_{\mathrm{off}}(0)p_{\mathrm{odd}}(t_{\mathrm{on}},t')$ where t' is the point in time when the bleaching event takes place. t' lies between t_{on} (then the molecule is continuously "on" until it bleaches) and T (then t_{on} is spread over the illumination time T), see Fig. 7.13.

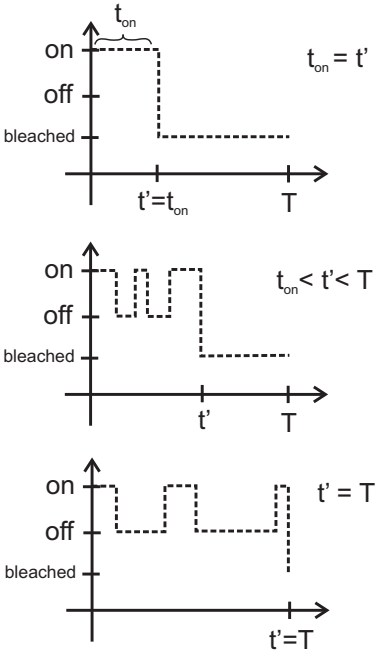

Figure 7.13
Illustration of fluorescence traces contributing to the probability distribution $p_{fluor}(t_{on}, T)$. In all three cases the "on" time is identical while the time point of bleaching t' is varied. In the top most trace the time of bleaching t' is identical to t_{on}, the molecule has to be continuously in the "on" state until bleaching. In the middle trace the "on" time is spread over the time until bleaching due to intermittent periods in the "off" state. In the bottom trace the molecule bleaches exactly at the end of the illumination period. The "on" time is spread over the whole illumination period T.

To properly account for all possible traces one has to integrate t' over all allowed values $t_{\mathrm{on}} < t' < T$. Finally, the probability for a bleaching event after an "on" time of t_{on} is $k_{\mathrm{bl}} \exp(-k_{\mathrm{bl}} t_{\mathrm{on}})$. In summary, the contribution of switching traces which are ended by a final bleaching event is:

$$p_{\mathrm{on}}(0) k_{\mathrm{bl}} \exp(-(k_{\mathrm{bl}} + k_{\mathrm{off}}) t_{\mathrm{on}}) + k_{\mathrm{bl}} \exp(-k_{\mathrm{bl}} t_{\mathrm{on}}) \int_{t_{\mathrm{on}}}^{T} dt' (p_{\mathrm{on}}(0) p_{\mathrm{even}}(t_{\mathrm{on}}, t') + p_{\mathrm{off}}(0) p_{\mathrm{odd}}(t_{\mathrm{on}}, t')) \tag{7.33}$$

Adding the contributions from the 4 different kinds of fluorescence traces

results in

$$
\begin{aligned}
&p(t_{\mathrm{on}}, T) = p_{\mathrm{on}}(0)\exp(-(k_{\mathrm{off}}+k_{\mathrm{bl}})T)\delta(T-t_{\mathrm{on}}) \\
&+ p_{\mathrm{off}}(0)\exp(-k_{\mathrm{on}}T)\delta(t_{\mathrm{on}}) + \Theta(t_{\mathrm{on}})\Theta(T-t_{\mathrm{on}})\Big[k_{\mathrm{bl}}\exp(-k_{\mathrm{bl}}t_{\mathrm{on}}) \\
&\times\left(p_{\mathrm{on}}(0)\exp(-k_{\mathrm{off}}t_{\mathrm{on}}) + \int_{t_{\mathrm{on}}}^{T} dt'(p_{\mathrm{on}}(0)p_{\mathrm{even}}(t_{\mathrm{on}};t') + p_{\mathrm{off}}(0)p_{\mathrm{odd}}(t_{\mathrm{on}};t'))\right) \\
&+ \exp(-k_{\mathrm{bl}}t_{\mathrm{on}})\exp(-k_{\mathrm{off}}t_{\mathrm{on}} - k_{\mathrm{on}}t_{\mathrm{off}}) \\
&\times\Big[(p_{\mathrm{on}}(0)k_{\mathrm{off}} + p_{\mathrm{off}}(0)k_{\mathrm{on}})I_0\left(2\sqrt{k_{\mathrm{off}}k_{\mathrm{on}}t_{\mathrm{on}}t_{\mathrm{off}}}\right) \\
&+(p_{\mathrm{on}}(0)\sqrt{t_{\mathrm{on}}/t_{\mathrm{off}}} + p_{\mathrm{off}}(0)\sqrt{t_{\mathrm{off}}/t_{\mathrm{on}}})\sqrt{k_{\mathrm{off}}k_{\mathrm{on}}}I_1\left(2\sqrt{k_{\mathrm{off}}k_{\mathrm{on}}t_{\mathrm{on}}t_{\mathrm{off}}}\right)\Big]\Big]
\end{aligned}
$$

We simplify this model by assuming that "on" and "off" rates are equal ($k_{\mathrm{on}} = k_{\mathrm{off}} = k$) and that the molecule is initially always in the "on" state ($p_{\mathrm{on}}(0) = 1, p_{\mathrm{off}}(0) = 0$).

If we use the probability distribution $p_{\mathrm{fluor}}(t_{\mathrm{on}}, T)$ for the 3-state model (Eq. 7.34) to calculate the distribution of incident photons $p_{\mathrm{inc}}(W, T)$ (Eq. 7.30) and insert $p_{\mathrm{inc}}(W, T)$ in Mandel's theory (Eq. 7.27) we obtain the probability distribution $p(n, T)$ for the number of photons detected during illumination time T. This distribution was used to illustrate the influence of blinking and bleaching in Sec. 7.3.1.

7.C.4 Robust intensity distributions

If the illumination time T is so big that bleaching is fast on the time scale set by T ($k_{\mathrm{bl}} > 1/T$), the molecule bleaches within the illumination time T. If, additionally, blinking is faster than bleaching $k > k_{\mathrm{bl}}$, the molecule switches frequently between the "on" and "off" state during $1/k_{\mathrm{bl}}$ and k_{bl} is the only relevant rate. Under these conditions we can significantly simplify Eq. 7.34 to

$$
p_{\mathrm{fluor}}(t_{\mathrm{on}}) = k_{\mathrm{bl}}\exp\left(-k_{\mathrm{bl}}t_{\mathrm{on}}\right) \tag{7.34}
$$

$$
\Rightarrow p_{\mathrm{inc}}(W) = \frac{k_{\mathrm{bl}}}{\eta_o \bar{I}}\exp\left(-\frac{k_{\mathrm{bl}}W}{\eta_o \bar{I}}\right) \tag{7.35}
$$

If we insert this distribution into Mandel's formula Eq. 7.27 we get:

$$p(n,T) \equiv p(n) = \int_0^\infty \frac{(\eta_d W)^n\, e^{-\eta_d W}}{n!} \cdot \frac{k_{\mathrm{bl}}}{\eta_o \bar{I}} \exp\left(-\frac{k_{\mathrm{bl}} W}{\eta_o \bar{I}}\right) dW \tag{7.36}$$

$$= \frac{k_{\mathrm{bl}}}{\eta_o \bar{I}} \frac{\eta_d^n}{n!} \int_0^\infty W^n e^{-\eta W} dW \quad \text{with} \quad \eta = \eta_d + \frac{k_{\mathrm{bl}}}{\eta_o \bar{I}} \tag{7.37}$$

$$= \frac{k_{\mathrm{bl}}}{\eta_o \bar{I}} \eta_d^n \eta^{-(n+1)} = \frac{1}{1 + \eta_d \eta_o \bar{I} k_{\mathrm{bl}}^{-1}} \left(1 + \frac{1}{\eta_d \eta_o \bar{I}} k_{\mathrm{bl}}^{-1}\right)^{-n} \tag{7.38}$$

By introducing $N = \eta_d \eta_o \bar{I} k_{\mathrm{bl}}^{-1}$ we can write this distribution for the number of detected photons n as

$$p(n;N) = \frac{1}{N}\left(1 + \frac{1}{N}\right)^{-(n+1)} \tag{7.39}$$

where N is the average number of photons detected $N = \sum_{n=0}^{\infty} n\, p(n)$.

In order to estimate the range of parameters in which the simplified model is valid, we compare it to the full model Eq. 7.34. Fig. 7.14 shows that the model works best, if k_{bl} is well separated from the blinking rates k and $1/T$. As mentioned in Sec. 7.2 and Sec. 7.3.6, this is indeed the case for the used experimental parameters.

The distribution of the intensity of a dimer (i.e. two fluorophores in a diffraction limited spot) $p_2(n;N)$ is obtained from the convolution of Eq. 7.39 with itself (121):

$$p_2(n;N) = \sum_{n'=0}^{\infty} p(n-n')p(n') \tag{7.40}$$

$$= \frac{1}{(1+N)^2} \sum_{n'=0}^{\infty} \left(1 + \frac{1}{N}\right)^{-(n-n')} \Theta(n-n') \left(1 + \frac{1}{N}\right)^{-n'} \tag{7.41}$$

$$= \frac{1}{(1+N)^2} \left(1 + \frac{1}{N}\right)^{-n} \sum_{n'=0}^{n} = \frac{n+1}{(1+N)^2} \left(1 + \frac{1}{N}\right)^{-n} \tag{7.42}$$

Continued convolution with $p(n;N)$) (Eq. 7.39) gives the distribution for

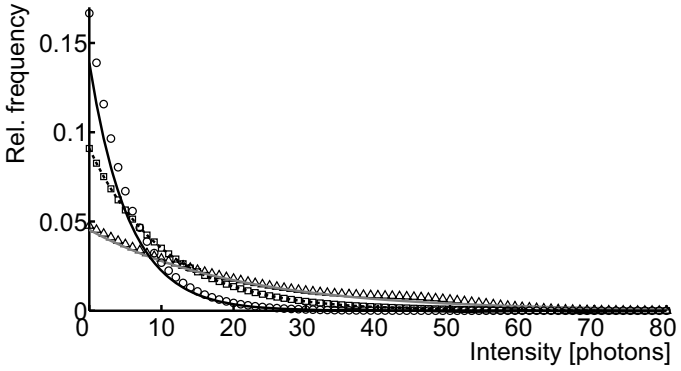

Figure 7.14
Comparison of the general model based on Eq. 7.34 to the simplified model Eq. 7.39 for long illumination times T. The open symbols correspond to the intensities calculated using the general model based on Eq. 7.34 with bleaching rate $k_{bl} = (0.05T)^{-1}$ (open circles), $k_{bl} = (0.1T)^{-1}$ (open squares), $k_{bl} = (0.2T)^{-1}$ (open circles). The switching rate is $k = (0.01T)^{-1}$ for all distributions. The lines correspond to intensities given by Eq. 7.39 with the same values for k_{bl}: $k_{bl} = (0.05T)^{-1}$ (black solid line), $k_{bl} = (0.1T)^{-1}$ (black dashed line), $k_{bl} = (0.2T)^{-1}$ (grey solid line)

higher multimers, like e.g. a trimer

$$p_3(n;N) = \sum_{n'=0}^{\infty} p(n-n')p_2(n') \tag{7.43}$$

$$= \frac{1}{(1+N)^3}\left(1+\frac{1}{N}\right)^{-n}\sum_{n'=0}^{n}(n'+1) \tag{7.44}$$

$$= \frac{(n+1)\,((n/2)+1)}{(1+N)^3}\left(1+\frac{1}{N}\right)^{-n} \tag{7.45}$$

In Fig. 7.2 the intensity distributions for a monomer, dimer and trimer with the same average number of detected photons N (per fluorophore) are compared.

7.C.5 Complete intensity distribution

By combination of the simplified model for the number of emitted photons Eq. 7.39 with the detection probability Eq. 7.25 a complete description of

experimental intensity distributions is obtained:

$$p(n,N) = \frac{C}{2}\left(1+\operatorname{erf}\left(\frac{1}{\sqrt{2}}\left(\frac{n+\mathrm{off}'}{2\sqrt{\pi}\sigma w}-t\right)\right)\right)\frac{1}{N}\left(1+\frac{1}{N}\right)^{-(n+1)} \tag{7.46}$$

Since the signal width w is a property of the optics and the fluorophore and σ, the noise level, can be estimated from the data, N, the mean number of photons detected during the integration time, is the only free parameter. C is determined by normalization $\sum_n p(n,N) = 1$. C also gives the ratio between the number of actually detected single-molecule signals and the total number of single-molecule signals present: C = #(detected signals)/#(all signals). Fig. 7.3C shows an example of the distribution for typical experimental values. Note that single-molecule intensity distributions have an intrinsic asymmetry due to the influence of the detection probability. As suggested in (223) there might be other experimental factors that can lead to asymmetric intensity distributions.

7.D Overlapping single-molecule signals

When two single-molecule signals are closer together than the diffraction limit $2 \cdot w$ (= 2 × signal width), they cannot be resolved anymore. Consequently, a dimer is observed. For a homogeneous distribution of molecules with surface density ρ, on average $\rho\pi w^2$ molecules are found in a circle of radius w. Those molecules would be observed as a single signal. Assuming a Poisson process for the positions of the molecules, the probability to observe an n-mer $p_{\mathrm{cluster}}(n,w)$ at a signal width w is therefore:

$$p_{\mathrm{cluster}}(n,w) = C\frac{(\rho\pi w^2)^n e^{-\rho\pi w^2}}{n!} \tag{7.47}$$

where C is determined by normalization: $\sum_{n=1}^{\infty} p_{\mathrm{cluster}}(n,w) = 1$, so

$$p_{\mathrm{cluster}}(n,w) = \frac{(\rho\pi w^2)^n e^{-\rho\pi w^2}}{\left(1-e^{-\rho\pi w^2}\right) n!} \tag{7.48}$$

ρ is given by $N_{\mathrm{mol}}/A_{\mathrm{ROI}}$ where N_{mol} is the number of molecules in the region of interest (ROI) with area A_{ROI}. Note that this model is only valid if the density is far below the percolation limit.

The relation between the number of molecules N_{mol} and the number N_{signals} of observed signals is given by

$$N_{\mathrm{signals}} = N_{\mathrm{mol}} \left(\sum_{n=1}^{\infty} n p_{\mathrm{cluster}}(n, w) \right)^{-1} \tag{7.49}$$

Fig. 7.15 shows the ratio of observed signals to the number of molecules $N_{\mathrm{signals}}/N_{\mathrm{mol}}$ for a typical experimental value of $w = 0.7\,\mathrm{pxl}$, $pxl = 220nm$. Here, n-mers up to $n = 50$ are considered. For comparison, if only monomers and dimers are admitted (dashed line in Fig. 7.15), the amount of overlap is underestimated. For $\rho < 0.25/\mathrm{pxl}^2$ both curves coincide, which means that for low densities, signals consist exclusively of monomers and dimers. In that regime, the ratio $N_{\mathrm{signals}}/N_{\mathrm{mol}}$ decreases linearly with density ρ. The monomer fraction α used in Sec. 7.3.5 is defined as $p_{\mathrm{cluster}}(1, w)/N_{\mathrm{signals}}$.

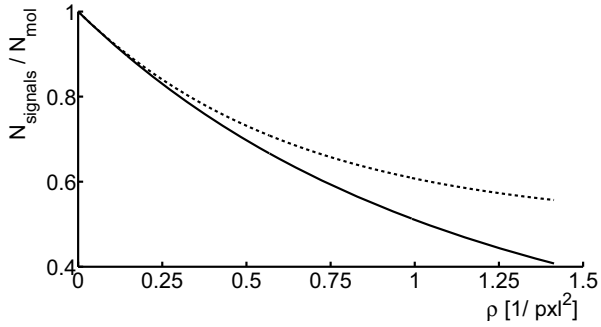

Figure 7.15
$N_{signals}/N_{mol}$ for $w = 0.7pxl$, $pxl = 220nm$. Solid line: n-mers up to n=50 are considered , dashed line: only monomers and dimers are considered