Quantitative image analysis for fluorescence microscopy

Staszowska, Adela

Awarding institution:
King's College London

KING'S COLLEGE LONDON:

Randall Division of Cell and Molecular Biophysics

Quantitative image analysis for fluorescence microscopy

Adela Danuta Staszowska

March 2017

Abstract

Localisation microscopy enables imaging beyond the Abbé resolution limit and is based on detecting randomly activated single molecules in a sequence of images. A super-resolution image is then reconstructed using these localisations. Analysis of these points can provide quantitative information about number, shape, and size of features. This thesis presents novel approaches to clustering analysis, identification and characterisation of specific features with a known shape, and confirming the presence of clusters in the sample.

Clustering analysis can be used to both detect clusters, and to measure their size and density. Analysis of localisation microscopy images of biological samples is challenging because clusters are usually small and surrounded by relatively high noise. Here, the Rényi divergence, which can be adjusted to the properties of the data was used. This method provides a more precise measurement of the cluster size, than the commonly used Ripley's K function.

Biological samples are often highly structured, and the distribution of proteins within these structures is of great interest. In this work a type of adhesive structure called podosomes which consist of an f-actin core surrounded by a protein ring was investigated. Custom written software identified podosome rings in images using a circular model and calculated the relative positions of different ring proteins. This information was used to build a model of podosome ring composition.

The appearance of podosome rings imaged with localisation microscopy depends on the sample preparation and image analysis techniques used, sometimes appearing strongly clustered and sometimes continuous. To attempt to distinguish whether the apparent

clusters were due to fluorophore reappearance, a microscopy system was developed to measure fluorescence resonance energy transfer with anisotropy.

The different analysis methods presented in this thesis illustrate the ways in which data analysis and experimental methods can provide a better understanding of a biological system.

Acknowledgements

First and foremost I would like to thank my supervisor Dr Susan Cox. Susan helped me to grow as a scientist and develop numerous skills while working on my PhD project. She has encouraged me to test new ideas and challenge what I thought I already knew. Also during my writing-up period, she offered me guidance to improve my writing skills.

Special thanks are also due to my second supervisor Prof. Gareth Jones. Gareth helped me to learn new things about biology and I have enjoyed our discussions very much.

I would like to also thank my colleagues: Dr Patrick Fox-Roberts for teaching me about importance of code optimisation and avoidance of brute-force approaches to programming. Dr Sandeep Mavadia for his help with optical system design. Dr Richard Marsh for discussions about DNA-origami sample preparation. Elizabeth Foxall for tirelessly preparing podosome samples for me. Dr Rafal Szepietowski for his help with statistics and algorithm design.

I also thank my examiners, Dr Ricardo Henriques and Prof. John Girkin for the time they spend reading my thesis and for their insightful comments.

I am also grateful to Dr Simon Ameer-Beg, who encouraged me to pursue further studies, and for his help while I was applying for this PhD project.

I would like to acknowledge Helen Rudkin and Sarah Meredith, who both helped me with the bureaucratic side of being a student and patiently answered my numerous questions.

Last, but not least, I would like to thank of my parents and family for their loving support.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1   Light microscopy and Fluorescence

The microscope is an optical system used to investigate objects smaller than easily observable with, for example human eye. The emergence of microscopes was made possible by improvements in lens production in the 17th century. Robert Hook presented the first known etching of the microscope and extensive research report in *Micrographia* (1665). Many of the first microscopes (including Hooke's) were constructed using only two lenses: a magnifying lens and an eye piece lens [Hooke, 1665; Lipson et al., 2011]. The microscopes we use today still have these two basic components, with addition of other optics elements for illumination and imaging control.

Microscopes enable imaging of small structures and samples such as cells. However, many biological materials are largely transparent to light from the visible spectrum. There are two most commonly used types of methods used to overcome this problem are dark field microscopes (phase contrast and DIC) and marking the area of interest with stains or fluorophores. The phase contrast microscopy is based on a contrast enhancement by detecting phase change and displaying it as a variation in brightness. In phase-contrast microscopy the phase change is caused by the difference in the refractive index in the sample. Structures with high refractive index bend light to greater angles than objects with lower refractive index. Higher refractive index can also introduce delay in the light passage

introducing the phase shift. Thus, in wide-field microscope, light which passed through the high refractive index area in the sample will be pass trough the collecting lens further from its centre than light passing through the area with low refractive index [Zernike, 1942a,b]. The differential interference contrast microscopy (or DIC microscopy) uses a shearing principle to image difference in light paths of the light rays passing through the sample. This microscope detect the difference in the light path of two perpendicularly polarised beams by combining them with a second prism and analysed using a polariser [Oldfield, 2001]. The second method of imaging transparent samples is to introduce non-transparent substances to the sample to mark areas or structures of interest and different types substances can be used to mark different proteins or organelles in the samples. These substances can either absorb light or absorb and emit light with different wavelength, in process called fluorescence. Substances absorbing light are called stains and are visible as a coloration of the portion of the sample, for example, haematoxylin and eosin stains are used in histology to mark nuclei and cytoplasm in tissue sections. Staining is mainly used in transmission light microscopy and electron microscopy.

Substances can emit light using energy of their excited electrons, in process called luminescence. Luminescent light emission can occur from different excited states: fluorescence, occurring from a singlet excited state and phosphorescence from a triplet state. Singlet state has a pair of electron of opposite spin in excited and ground states. Thus, the return of the electron from the excited state to the ground state is allowed and it occurs very fast. An average lifetime of fluorophores is 10 ns. Conversion from the triplet state is spin forbidden, because the excited electron has the same spin orientation as its pair electron in the ground state. Consequentially, the phosphorescence occurs slowly (from milliseconds to seconds, but it can have even longer lifetime) [Lakowicz, 2006]. Energetic states of fluorescent materials can be described visually using a scheme called Jablonski diagram. A simplified version of this diagram is shown in Figure 1.1.

Molecules exhibiting fluorescence are called fluorophores. Fluorescent molecules are typically organic and contain aromatic amino acids. The first identified fluorescent protein, observed in 1845, was quinine. It can be excited with UV light and emits light with wavelength 450 nm. Similarly other naturally occurring fluorophores, tryptophan
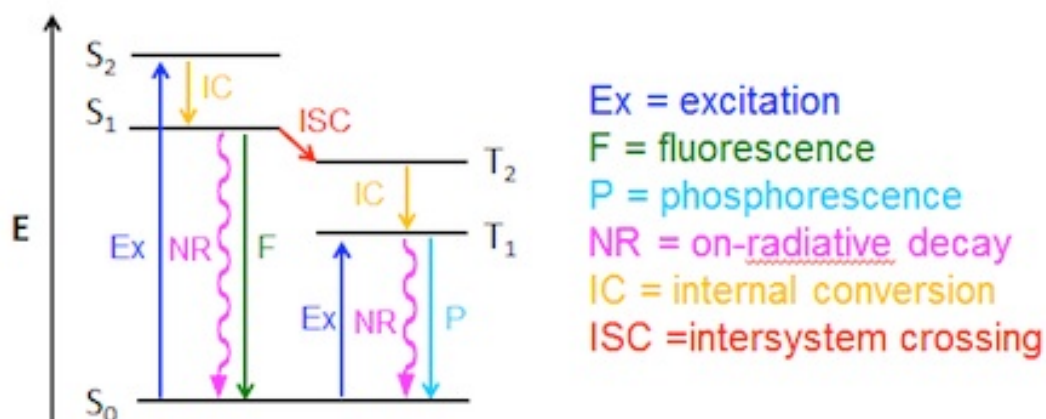
Figure 1.1: The Jablonski diagram. An electron in the ground state can be excited to the excited singlet state S_2 or triplet state T_1 (excitation marked with navy). Excited electron looses its energy to vibrational modes of the molecule in process of internal conversion changing excited levels from S_2 to S_1 or T_2 to T_1 (also called radiationless de-excitation – marked with yellow). Excited electron on the state S_1 or T_1 can return to the ground state by emitting light or radiative decay (heat – marked with magenta arrows). Light emitted by electron in singlet state is called fluorescence (marked with green) or from triplet state phosphorescence (marked with light blue). Image adapted from: [Janson, 2016].

and tyrosine can excited by the UV light and their emission spectra are similar (the difference being in the spectrum spread, the spectrum of tyrosine is slightly more narrow than spectrum of tryptophan). Tryptophan is very sensitive to changes in its environment and spectral shift was observed as a result of for example ligand binding, protein-protein association or protein unfolding [Lakowicz, 2006].

However, most substances do not exhibit natural fluorescence (for example DNA or lipids). To image such materials they have to be labelled with different kinds of fluorescent markers. The mode of imaging (fixed or live cells) or available methods of introducing fluorescent markers (for example immunostaining or method of introducing fluorescent molecule) limit the possible number of fluorescent labels that can be used to mark a given sample. The first major division of fluorophores is based on the way they attach to the sample through antibody labelling or are directly encoded by the protein DNA. The first

group uses antibody labelling, where a fluorophore is attached to an antibody, either to a secondary (which later attaches to a primary antibody) or directly to the primary antibody marking the structure of interest [Hyatt and Wise, 2001]. Organic dyes such as fluorescein and rhodamine are popular labels, and can be conjugated to antibodies and used for sample labelling and in immunoassays [Matveeva et al., 2004]. Other examples of widely used dyes are the Alexa Fluor, ATTO, or Cy dyes, their advantage is photostability compared to fluorescein and rhodamine [Hyatt and Wise, 2001]. More specialistic probes are: probes used for DNA binding (for example ethidium bromide, or EB Hedley et al. [1983]), membrane probes (for example insoluble in water, DPH, 1,6-diphenyl-1,3,5-hexatriene [Kaiser and London, 1998]), or chemical sensing probes (for example probe Fura-2 can be used to monitor levels of Ca^{2+} [Williams et al., 1985]).

The other group of fluorescent markers are the fluorophores directly expressed by the protein of interest. Probes such as are green fluorescent protein (GFP) can generally be found in jellyfish and in some species of coral [Dickson et al., 1997; Lakowicz, 2006; Tsien, 1998]. GFP is built with a β-barrel structure surrounding a chromophore in its centre. The most important feature of GFP is that the chromophore forms during folding of the polipeptide chain, which means that genes coding the expression of GFP can be inserted directly into the sample DNA [Tsien, 1998]. GFP will be then synthesized directly in the cells (it is even possible to create whole organism expressing GFP, for example GFP transgenic zebra fish [Gong et al., 2001]). GFP is known for its good photostability and relatively high quantum yield (lower than the quantum yield of organic dyes, but high for the fluorophores). Now, mutants of GFP are more commonly used, they have better stability and longer excitation and emission wavelengths [Kremers et al., 2010]. Also, a broader range of different fluorescent proteins is now available. Their emission maxima ranges from 448 to 600 nm (see Figure 1.2).
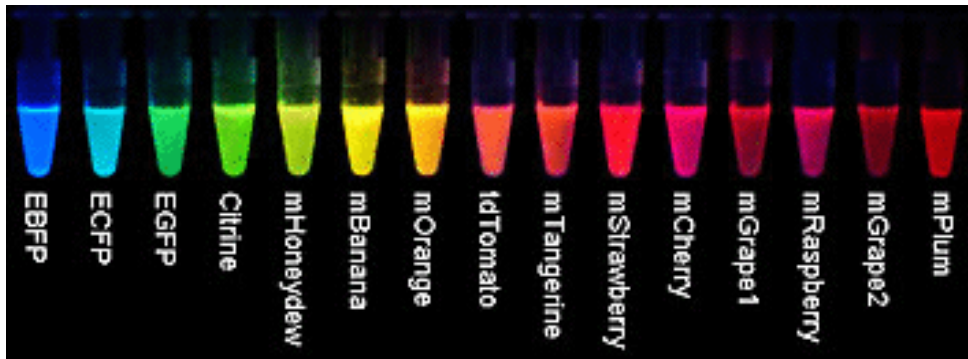
Figure 1.2: Different types of fluorescent proteins. Image retrieved online from: [Tsien, 2016].

1.2 Fluorescence microscopy

Fluorescence microscopy is a type of optical microscopy using fluorescent molecules naturally present or introduced to the sample and detecting only fluorescent light. The most basic fluorescence microscope system is composed of the light source (for example xenon arc lamp or mercury lamp), the excitation filter (which only transmits the required wavelengths of light), the beam splitter, the emission filter (blocking non-fluorescent light) [Lakowicz, 2006; Lipson et al., 2011].

Fluorescence microscopy is an extremely relevant tool to study transparent biological systems. Many kinds of biological and non-organic specimens remain nearly transparent or display very few details in light microscopy despite any contrast method used. Therefore, a way to make the structures of interest visible is to introduce a fluorophore to the sample (or use an endogenous fluorophore) and induce fluorescence. Use of more than one type of fluorophore to mark different components of the sample reveals more information about the relative positions of proteins building the imaged structure.

Ability to label specific protein in the sample is the main advantage of the fluorescence microscopy. However, the fluorescence microscopy has also limitations connected to the properties of fluorophores such as their spectral overlap, especially important when using two types of fluorescent probes at the same time due to possible cross talk, or quantum yield, or the efficiency of the fluorophore (the ratio between the number of exciting and

emitted photons, usually smaller than 100%). Another example of a problem is photo-bleaching. Electrons excited during imaging can damage the fluorophore, for example due to non-radiative decay. Photobleaching is the main factor limiting the time for which sample remains fluorescent post excitation. Photobleaching can be reduced by using photoprotective substances such as p-phenylenediamine, n-propyl gallate [Longin et al., 1993], or as in this study the ProLong® Gold embedding medium. Living samples can be affected by phototoxicity caused by heating or build-up of to toxic chemicals during imaging. Additionally, attaching or introducing fluorophore to the sample may disrupt the structure of the sample. This effect can be observed both in immunolabelling (due to weight of the antibody linker) and using transfected fluorophores, for example, Alexa Fluor 647 weights 1.2 kDa, an antibody can weight between 150-900 kDa [Sigma-Aldrich, 2017], and mCherry weights 28 kDa. These weights are comparable or bigger with the protein weight, for example the proteins discussed later in Chapter 3, such as vinculin (117 kDa), paxillin (69 kDa), and talin (270 kDa). Lastly, fluorescence microscopy allows only observation of the fluorescently labelled structures, others areas will be invisible.

1.3   Diffraction limit and point spread function

The image of any object imaged with optical system is broadened by diffraction. The diffractional broadening can be described by measuring the light intensity distribution on the image. The measured intensity distribution of a point source is also called point spread function (PSF). In the process of imaging, the object function is convolved with the PSF giving the image. The PSF depends on the imaging system and determines size of features distinguishable on the image.

There are three different methods of defining the diffraction limit such as Abbé limit, the Rayleigh criterion and the Sparrow condition. The resolution of any light based imaging device is related to the wavelength of light used in imaging process. To be able to separate two objects the distance between them needs to be greater than the resolution limit (also known as Abbé limit). The maximum possible in plane (lateral) resolution which can be obtained using wide-field light based techniques is given by equation 1.1

and the axial (z-) resolution by equation 1.2 [Abbe, 1873; Gould et al., 2009]:

$$R_{xy} = \frac{\lambda_{illumination}}{2N.A.} \qquad (1.1)$$

and

$$R_z = \frac{2\lambda_{illumination}}{N.A.^2}, \qquad (1.2)$$

where R is resolution, λ is the wavelength of the illuminating light and N.A. is the numerical aperture of the optical device. The relation between the theoretical resolution and the emitted wavelength of light used for imaging is shown in Figure 1.3. Structures with
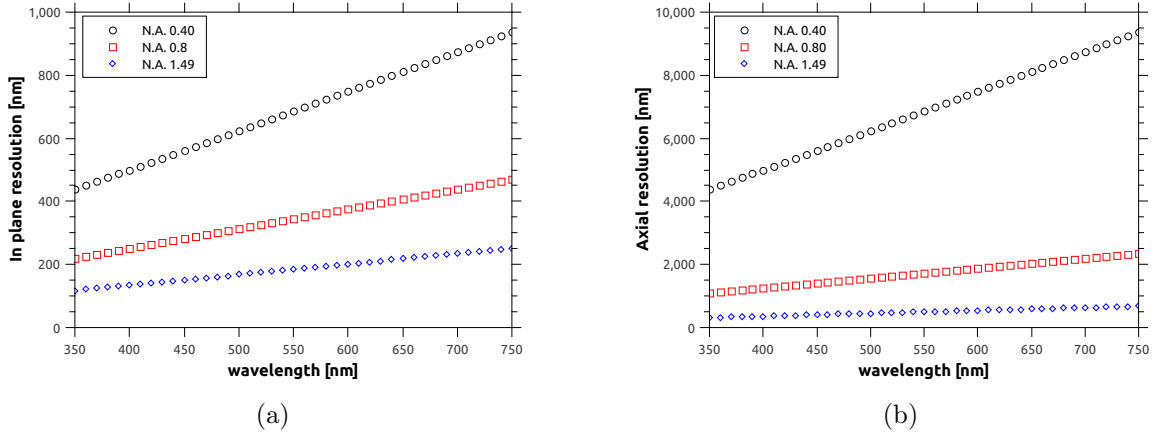
Figure 1.3: Theoretical in-plane and axial resolution limit calculated for three different objectives used in fluorescence anisotropy imaging system in our lab (see Chapter 4). Numerical aperture is marked on the graphs with different colours: N.A. 0.4 with black, N.A. 0.8 with red, and N.A. 1.49 with blue. a) Theoretical in plane resolution for different wavelengths of light for three N.A values. b) Theoretical axial resolution for imaging with different wavelength for three N.A. values.

sizes smaller than the diffraction limit will not be distinguishable in the final image [Lipson et al., 2011]. For a typical light microscope with an objective 1.3 N.A, the in plane resolution will be around 250 nm and axial resolution around 600 nm. The main limitation of this resolution limit definition is that it was defined for an infinite object and it

requires that full diffraction image is collected. Thus this definition of diffraction cannot be used directly for scanning techniques or structured illumination microscopy [Lipson et al., 2011].

The resolution limit can also be understood using the definition proposed by Rayleigh. The Rayleigh resolution limit definition uses angular displacement of two or more objects to define resolution and it is formulated for incoherently illuminated objects. Consider two points separated by distance x or angular displacement θ (see Figure 1.4).

The normalized intensity of the diffraction pattern for the aperture diameter D is given as:

$$I(\theta) = \left[2J_1 \left(\frac{1}{2} k_0 D \sin\theta \right) \Big/ \left(\frac{1}{2} k_0 D \sin\theta \right) \right]^2, \tag{1.3}$$

where J_1 is the first order Bessel function. According to the Rayleigh definition, two equally bright points are distinguishable when they are separated by at least a distance for which the central maximum of the diffraction pattern of the first object does not coincide with the first minimum of the diffraction of the second object [Rayleigh, 1879]. The Bessel function $J_1(x)$ has it first zero at $x = 3.83$:

$$\frac{1}{2} k_0 D \sin\theta_1 = \pi D \sin\theta_1 / \lambda \tag{1.4}$$

The angle θ_1 is the minimal possible angle between two light sources in order for them to be distinguishable. Also, the $\theta_1 \ll 1$ thus:

$$\theta_1 = 3.83\lambda/\pi D = 1.22\lambda/D \tag{1.5}$$

The Rayleigh criterion has one major weakness: it is only defined for diffraction patterns with no-well defined zero points (or the zero points are positioned far from the central maximum).

The Sparrow condition does not depend on the zeros of the point spread function but on the relative intensities of the imaged points [Lipson et al., 2011]. This condition is very close to the method human eye uses to distinguish two object. According to the Sparrow
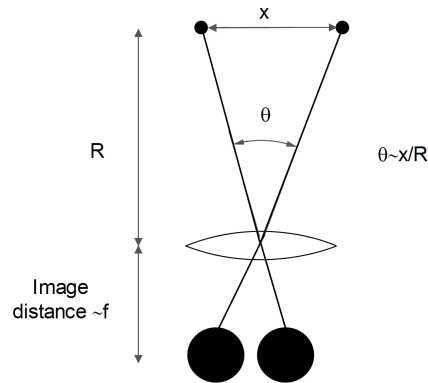
Figure 1.4: The Rayleigh resolution criterion. Top: Two point light sources distanced by x imaged with a lens. The light sources are in range R from the lens. Bottom: Image of two point light sources.

definition two objects are seen as separate when the sum of their point spread functions has a central minimum. For two points with equal intensity values the θ_{min} is given as:

$$\left(\frac{d^2 I}{d\theta^2}\right)_{\theta=\theta_{min}/2} = 0. \tag{1.6}$$

What after differentiation of Bessel function gives:

$$\theta_{min} = \frac{0.95\lambda}{D} \tag{1.7}$$

In wide-field fluorescence microscopy λ is the wavelength of emitted light.

The Sparrow criterion limits the resolution for very similar values to the theoretical Abbé limit. It is however, about 20% smaller than the Rayleigh resolution limit (see Figure 1.5). Additionally both Rayleigh and Sparrow limits do not provide a definition of resolution. They also do no take into account the number of photons detected from two adjacent sources. Both of those definitions also do not include a finite limit to the resolution (due to diffraction of light).
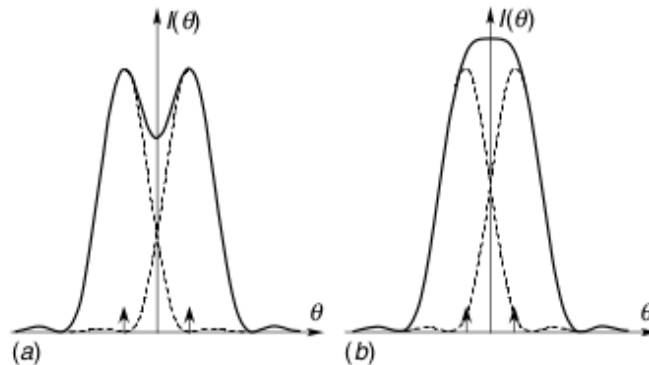
Figure 1.5: The Rayleigh and Sparrow resolution criterion for incoherently illuminated objects. Individual intensities of the light sources are shown with dash lines. The resulting intensity (marked with black line) for a)the Rayleigh criterion and b) the Sparrow criterion. Image adapted from: [Lipson et al., 2011].

1.4 Higher resolution microscopy

Obtaining higher resolution is crucial while imaging biological samples. Most animal and plant cells have dimensions from 1 to 100 μm, thus they can be imaged with practically any microscope. However, the intracellular and molecular structures smaller than 200 nm cannot be imaged without introducing new imaging methods improving the resolution below the diffraction limit. For example microtubules have a diameter of 25 nm [Lampe et al., 2012]).

Achieving better resolution in an optical system is possible by improving its quality, using shorter wavelength or increasing the N.A. of the objective lens. The Abbé theoretical limit does not take into account, for example, out-of-focus light in thick samples, thus the actual measurable resolution is lower than the theoretical limit. This can be improved using optical sectioning method (see Section 1.4.1). Similarly, the quality of the images can be improved using objective lens with higher N.A. The N.A. of an objective depends on the refractive index of the medium in which the lens is working and this limits the maximum possible numerical aperture. The mismatch between the immersion media and sample refractive indexes causes aberrations. In practice, objectives with N.A. around

1.45 are usually used for fluorescence microscopy. Lastly, a shorter wavelength of light can be used for imaging. Resolution of acquired image is proportional to the wavelength of light used for imaging, as the shorter the wavelength used the higher the resolution according to the Abbé's equations (see Section 1.4.2).

1.4.1 Confocal microscopy and deconvolution

The first approaches for increasing the resolution of optical instruments were based on improving the design and manufacturing the lenses and optical system design. The resolution (calculated with equations 1.1 and 1.2) is the maximal possible resolution. In practice the achievable resolution is affected by system or sample imperfections such as for example optical aberrations, inhomogeneous refractive index, low signal-to-noise ratio [Schermelleh et al., 2010].

Two examples of techniques commonly used to improve the quality of images are confocal microscopy and deconvolution [Murray, 2011; Schermelleh et al., 2010]. Confocal microscopy is a scanning technique. A laser beam is scanned through the sample spot by spot and the fluorescent light from every illuminated spot is recorded by a light detector. The image of the whole sample is not readily available but has to be recorded point by point and rendered using a software. There are two types of scanning modes: laser scanning, using two mirror system to shift beam in the x and y directions and sample scanning, where sample is placed on a moving platform [Minsky, 1961, 1988].

Additionally, confocal microscopy systems use a pinhole, placed in such a way that only a certain amount (a majority) of light from the in-focus plane passes through the pinhole. Generally, the pinhole is placed at a distance of a focal length from lens collecting light from in-focus plane. The out-of-focus light focused by the same lens will have its focal point in different place than in-focus light. Consequentially, out-of-focus light will be blocked, what enables optical sectioning and 3D imaging [Minsky, 1961, 1988]. The sectioning power depends on the pinhole size, for pinholes smaller than a single Airy disk it is constant and for pinholes bigger the sectioning power is weaker [Wilson and Carlini, 1987; Wilson, 1995]. However, the major advantage of using the pinhole is

resolution improvement. By restricting the size of the pinhole to block the higher orders of the diffraction pattern [Lipson et al., 2011]. Thus, the in-plane resolution of a confocal microscope is approximately given as:

$$R_{xy} = \frac{1.22\lambda_{emitted}}{2N.A.},$$ (1.8)

where the $\lambda_{emitted}$ is the wavelength of the emitted light. On the other hand introducing a small size pinhole means that the number of photons detected from the sample is going to be lowered. This can lead to a poor image contrast [Murray, 2011]. Other limitations of confocal microscopy are long image acquisition time, which affects its applications to image live cell samples [Schermelleh et al., 2010].

Deconvolution uses information about the diffraction broadening to reconstruct better resolution images. In this method standard fluorescence microscopy images are computationally processed to improve their quality using a guess or estimate of the imaged structure shape and PSF of the optical system. This process can be designed to model other factors affecting the quality of the image for example out-of-focus light. The computation usually is performed until the estimate of imaged object does not change between computational iterations. This technique can be used to improve images of both fixed and living cells, as it is performed after the images were collected. The limitations of deconvolution are mainly caused by using a single model of the PSF for whole image area. This can introduce bias or amplify noise. Deconvolution methods also require a certain signal-to-noise ratio in order to correctly process an image (usually more than S/N 20) [Murray, 2011]. Lastly, both the confocal microscopy and deconvolution methods are able to significantly improve resolution and quality of images, but they cannot provide better than the resolution given by the Abbé limit [Lipson et al., 2011].

1.4.2 Short wavelength microscopy

Achieving higher resolution is also possible by using shorter wavelength light, e.g. ultraviolet, soft x-rays and electrons as the illumination. Ultra-violet microscopy is mainly used as a non-destructive way of imaging in semiconductors or in photolithography. X-rays

are great tool for crystallography (the diffraction pattern is used to examine the crystal structure of material). X-ray crystallography can be used on bio-molecules, which can be crystallised. However, using x-rays for microscope built with glass lenses is not possible. There are no available x-ray lenses to built a conventional microscope system[1] [Lipson et al., 2011].

The most prominent example of short wavelength of microscopy is electron microscopy, where imaging is done using an electron beam. The wavelength of the electron beam can be controlled by adjusting the energy of electrons (according to the de Broglie formula $\lambda = \frac{h}{mv}$, where h is Planck constant, m is electron mass and v is the electron velocity). Since electron wavelengths are very short (for example: 2.5 pm at 200 keV), electron microscopy can image with resolution of 50 pm [Erni et al., 2009]. There are two main types of electron microscopy: scanning electron microscopy also called SEM (creating surface images) and transmission electron microscopy or TEM (creating 3D images).

The main drawback of electron microscopy is its high cost and demanding sample preparation which can change the sample properties. Samples for transmission electron microscopy (TEM) need to be fixed (no live cell imaging). Fixation is performed either with chemicals or cryotechniques. It appears that the rapid freezing of the sample can better preserve sample integrity and cause less artefacts in the samples. However, use of frozen samples can lower the achievable resolution. Scanning electron microscopy (SEM) requires a vacuum environment in the imaging chamber, thus samples need to be dehydrated and fixed before imaging and imaging of live cells is not possible. Dehydration of samples can further change their internal structure. Also, transmission electron microscopy is not able to collect true 3D images of the sample – most of the microscopes can rotate the sample only by 60-70° [Pawley, 2008; Wilson and Bacic, 2012].

---

[1]Refractive index for wavelengths of x-rays is smaller than 1 for every material used to construct standard lenses. Collecting x-rays is possible using Fresnel plates and recently a microscope was build using Fresnel plates acting as lenses.

1.5   Super-resolution fluorescence microscopy

The term super-resolution microscopy refers to the techniques which can image (and spatially resolve) structures smaller than the diffraction limit. In the last 20 years super-resolution microscopy techniques using wavelengths of light from the visible region of the spectrum were developed. These techniques overcome the diffraction limit by taking advantage of a non-linear response from the sample. Super-resolution techniques can be divided into two groups: far-field and near-field methods. Generally, far-field methods use waves which can propagate indefinitely (if they do not get absorbed, scattered or refracted by their environment). Near-field methods use waves, also called evanescent waves, which are limited to very short distances from their source. Therefore, the main difference between far-field and near-field is the propagation distance of the light. In reality, near-field methods usually also collect some portion of far-field light [Betzig et al., 1986; Lipson et al., 2011; Schermelleh et al., 2010]. One of the drawbacks of near-field microscopy is their limitation to be used on flat, 2D samples (because of its depth of field around 30-200 nm). In comparison far-field techniques have broader applications because, for example, they can be used for 3D imaging. The most prominent examples of these techniques are:

1. Stimulated Emission Depletion (STED) – two light beams of different shape which are used for imaging and de-excitation of the sample (see section 1.5.2),

2. Structured Illumination Microscopy (SIM), which uses moiré fringes to extract high resolution information from the images (see section 1.5.3), and

3. Localisation Microscopy which uses randomly excited single molecules, detects them and reconstructs a super-resolution image (see section 1.5.4).

1.5.1   Near-field microscopy

There are two near-field approaches which can provide super-resolution quality for optical imaging of biological material surfaces. The most basic near-field technique is called total internal reflection fluorescence microscope (TIRF). This technique uses an evanescent

wave to illuminate and excite only a thin (up to 300 nm in depth) portion of the sample located on the coverslip. This helps to improve the signal-to-noise ratio of the image (since only signal from the surface of the sample is detected). TIRF microscopy enhances only the axial resolution to 60-100 nm which makes it a perfect tool for examining, for example the cell surface [Axelrod, 1981; Lakowicz, 2006]. A more complex approach is used by the second technique, near-field scanning optical microscopy (SNOM/NSOM). It works without an objective lens, using a small aperture for scanning purposes (the aperture can be, for example, be positioned inside a glass tip – see Figure 1.6). The axial resolution is improved by use of an evanescent wave for illumination. The in-plane resolutions is improved by recording only small portions of the evanescent wave and depends on the proximity to the sample and diameter of the tip [Betzig et al., 1986; Schermelleh et al., 2010]. This approach can achieve a resolution of 20 nm when imaging surface structures [Betzig et al., 1986; Lipson et al., 2011; Schermelleh et al., 2010]. The main drawbacks of SNOM are long scanning times, the fact that scanning, with a tip or objective, has to be done practically on the surface of the sample (a very small working distance around 3-10 nm for objective and depth of field limited to 200 nm for the tip).
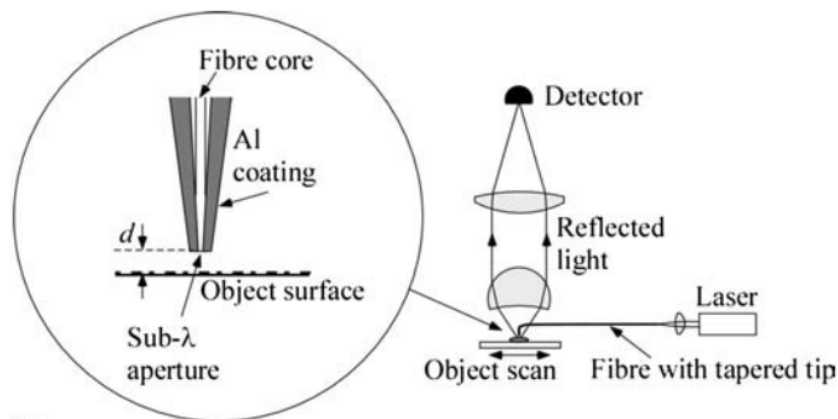
Figure 1.6: Near-field Scanning Optical Microscopy operation principle. Left: Large image of the aluminium coated glass tip used for imaging. Right: Microscope system. NSOM microscope has a standard scanning microscope set up. Image adapted from: [Lipson et al., 2011].

1.5.2   STED

Stimulated Emission Depletion (STED) microscopy is a far-field fluorescence scanning technique. It uses two beams, one to excite and the other to de-excite regions of the sample. The first beam excites and images fluorophores and has a Gaussian profile. The second, de-exciting, beam is depleting fluorophores from the parts of the region activated by the first beam. The de-exciting beam has a doughnut shape (the intensity in the middle of this beam is close to zero – see Figure 1.8). Both beams are aligned so that they have the same center point (that is the exciting beam is placed in the middle of the de-exciting beam) [Hell and Wichmann, 1994]. This composite beam scans across the sample. The structures which are in the center of the beam are imaged and all the other (off-centred) are de-excited by the doughnut beam.

The composite beam for STED imaging can be built using pulse and continuous lasers. In pulse-laser systems the light emission of the two laser sources is coupled [Hell and Wichmann, 1994]. A fluorophore has a number of different energy levels: L_0, L_1, L_2, and L_3 levels (see Figure 1.7). The vibrational relaxations lifetimes ($L_1 \rightarrow L_2$ and $L_3 \rightarrow L_0$) are significantly shorter than the lifetime of spontaneous emission ($L_2 \rightarrow L_3$) (around three times the magnitude). However, the L_2 is a very dynamic process thus it is advantageous to use a shorter, picosecond light pulse to stimulate emission. Additionally, the delay between the two pulses should also be selected to enable a temporal discrimination of the two pulses. It is the most advantageous to emit the depletion pulse immediately after the excitation pulse, because then the level L_3 would not be populated [Hell and Wichmann, 1994]. The exact desired length of the exciting and depletion beams can be calculated using the methodology presented, for example, in [Hell and Wichmann, 1994]. In practice, the exciting pulse is usually around 80 ps and is followed by 300 ps de-exciting pulse. This causes the main drawback of this approach, the laser pulses have to be synchronised (additionally pulse lasers are more expensive) [Willig et al., 2007]. The second approach uses continuous wave lasers and is much more cheaper and easy to align than the pulse laser system. The continuous depletion beam has a high intensity which influence the rate of fluorescence decay. The fluorescent molecules in the depletion zone
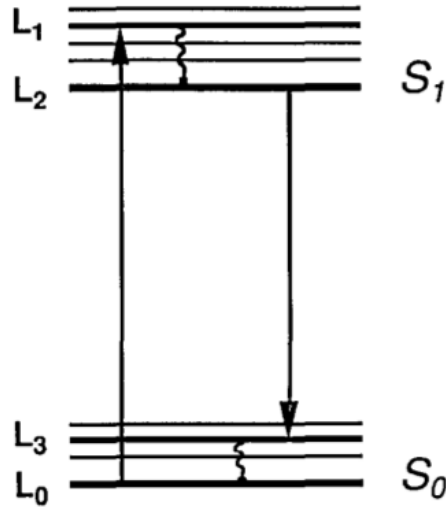
Figure 1.7: Four singlet energy states of an excited fluorophore. Image adapted from: [Hell and Wichmann, 1994].

have a shorter lifetime than those in the excited region [Willig et al., 2007]. The main drawback of this method is the fact that the continuous depletion beam requires a higher power than the pulse laser, which increases the bleaching rate.

The resolution improvement in STED microscopy is achieved due to the use of a beam with a smaller effective point spread function (see effective beam in Figure 1.8). The original STED design only allows the lateral resolution improvement and the axial resolution remains diffraction limited. Typically, STED microscopy can achieve in-plane resolution of 30-80 nm in biological samples [Hell and Wichmann, 1994; Lipson et al., 2011; Schermelleh et al., 2010; Willig et al., 2007]. Recently, z resolution was improved by creating the depletion beam using xy and z components combined with a beamsplitter [Schmidt et al., 2008]. They had also reported collection of optical sections distanced by 22 nm in the z direction [Schmidt et al., 2008]. Another method of 3D resolution improvement is use of a 2-photon laser as a excitation beam and standard STED depletion beam (2-photon excitation limits the excitation spot z dimension) [Ding et al., 2010].

The disadvantages of STED microscopy arise directly from its set-up. The composite beam illuminates, with high power, a larger sample area than is actually imaged (since only central beam of the illuminated area is imaged). This leads to high rates of bleaching
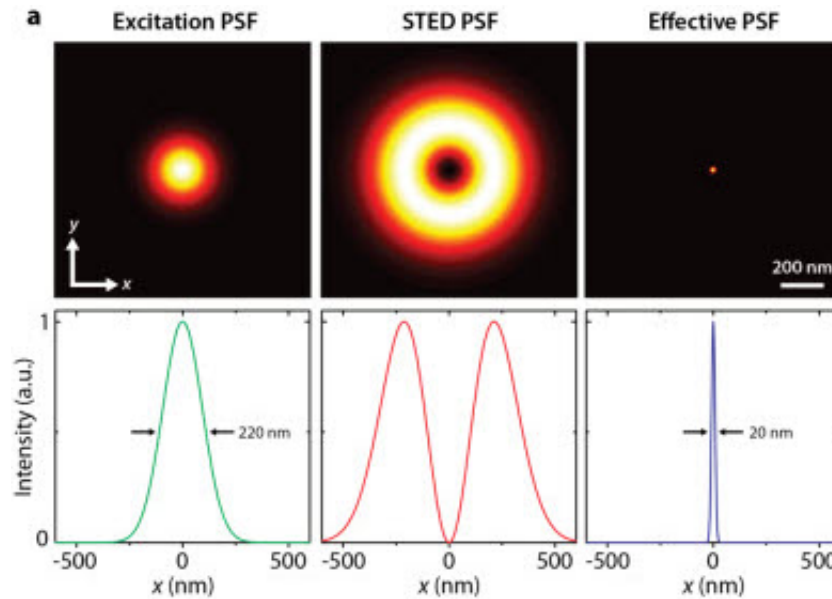
Figure 1.8: Stimulated Emission Depletion Microscope operation principle. Left: Exciting beam and its intensity profile. Similarly to standard scanning microscopes, exciting beam induces fluorescence in the area of the sample it illuminates. Middle: Beam stimulation depletion and its profile - this beam deactivates all the molecules apart from the centre point. There will be no fluorescence registered from area illuminated by this beam. Right: Effective exciting beam – only fluorescent light from the the effective exciting beam is going to be registered. Image adapted from: [Bio-nanoscopy Lab, 2015].

and phototoxicity because the same areas are excited more than once. Also, building and maintenance of a composite laser system can be costly [Gitai, 2009]. Lastly, STED is a scanning microscope thus taking an image requires time, which is an disadvantage when imaging live samples [Schermelleh et al., 2010].

1.5.3   SIM

Structured Illumination Microscopy (SIM) is based on projecting a light pattern into the sample (in the form of fringes). A SIM microscope is usually built using a wide-field microscope with a grating pattern inserted into the illuminating light path (before the sample). The grating pattern is usually a plate with a sinusoidal high frequency spatial

pattern (it can also be created using interference of two or more beams). Light going through the grating forms an illuminating pattern on the sample. The image captured by the camera contains the grid pattern and the sample images (see Figure 1.9). Then the illumination pattern displayed on the sample is translated or rotated (at least three times [Neil et al., 1997]) and followed by image acquisition. Using the data collected in the series a final image with enhanced resolution can be reconstructed using a methodology presented, for example, in [Neil et al., 1997]. This method increases the resolution in both the lateral and axial dimensions, and allows the creation of 3D images [Gustafsson et al., 2008; Neil et al., 1997].

For further resolution enhancement saturation of the fluorophores can be used [Gustafsson, 2005; Heintzmann et al., 2002]. In a typical fluorescence event the intensity of emitted light is linearly proportional to the intensity of emitted light. However, in case of saturation the intensity of emitted light will also depend on the local concentration of fluorophore [Heintzmann et al., 2002]. This fact is used in SSIM, where the sample is imaged with sinusoidal illuminating pattern with intensity equal or greater than the saturation threshold, then the emission rate (per average fluorophore) is not sinusoidal and contains higher harmonics. Use of higher harmonics enables acquiring higher resolutions. Saturated structured illumination (SSIM) can achieve resolution about 50 nm [Gustafsson, 2005; Li et al., 2015; Lipson et al., 2011; Schermelleh et al., 2010].

The disadvantages of SIM microscopy are mostly caused by out-of-focus light. The shot noise caused by the out-of-focus background is increased. The signal registered for the illumination frequencies close to the highest recordable frequency is significantly reduced [Mertz, 2011], which can lead to loss in image quality. Lastly, the optical sectioning power for imaging thick samples is much smaller for SIM than for example for confocal systems [Chasles et al., 2007; Fu et al., 2013]
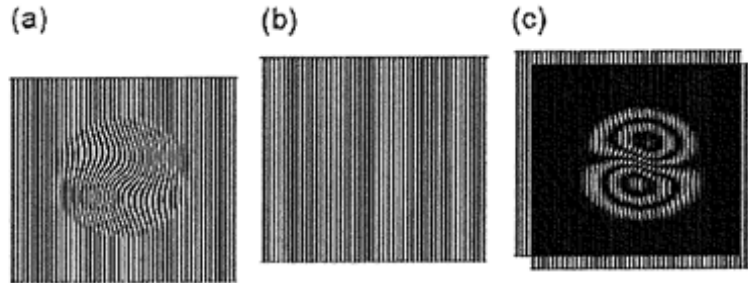
Figure 1.9: Structured Illumination operation principle. a) Object image with high frequency information. b) Vertical grid pattern with known spatial frequency. c) Image of an object and overlaid grid pattern with visible moiré fringes. This beat pattern occur at the spatial differences between spatial frequency of the pattern and spatial frequency components of the sample. The resulting frequency can be low enough to be seen by eye or imaging system, even if it was invisible before. High frequency components can be restored from the moiré pattern. Image adapted from: [Zeiss, 2015].

1.5.4   Localisation Microscopy

Localisation microscopy is a super-resolution technique which uses information about positions of single molecules in the sample. Localisation microscopy is usually referred to using one of the acronyms: STORM (Stochastic Optical Reconstruction Microscopy) [Betzig et al., 2006] or PALM (PhotoActivated Localization Microscopy) [Hess et al., 2006]. Localisation microscopy is based on the principle of single molecule detection [Betzig et al., 2006; Hess et al., 2006; Rust et al., 2006]. A single molecule can be localised with a higher precision than the resolution of an imaging system. Usually, fluorescent molecules in a sample have too high a density to localise individual molecules when they are emitting light at the same time. Thus, photoswitchable/photoactivable fluorescent molecules in the sample are randomly switched between bright and dark state with laser light [Hess et al., 2006; Rust et al., 2006]. Consequently, only a limited, random number of molecules are emitting light in each step, and to localise a high proportion of fluorescent molecules

present in the sample a long sequence of images is required (as shown in Figure 1.10). The final image is reconstructed using information about localisations from each image in the series.
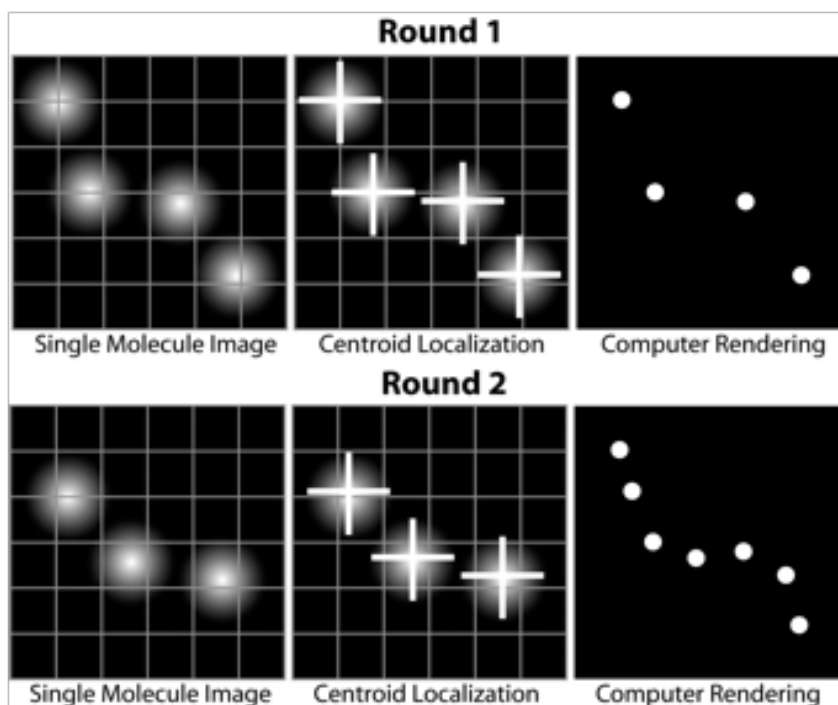
Figure 1.10: The localisation microscopy imaging principle using randomly activated dyes. In each round a random set of fluorescent molecules is imaged. Then sequence of images is analysed to recover positions of single molecules. The super-resolved image is then reconstructed using list of single molecule localisations. Image adapted from: [Photometrics, 2014].

Localisation techniques, such as PALM and STORM, require fluorescent molecules which can be switched between emitting and non-emitting states with laser light. Theoretically any photoswitchable/ photoactivable molecule is suitable for localisation imaging. However, in practice the fluorophore needs to fulfil three conditions:

1. The photobleaching rate needs to be equal or higher than the activation rate (the activation rate can be controlled by adjusting the laser power).

2. The typical number of detected photons from one molecule in one acquisition needs

to be high enough to provide good localisation precision.

3. The background and inactive state fluorescence needs to be lower than the active state fluorescence [Shroff et al., 2007].

Even if these conditions are met, the resolution of the reconstructed images is limited by the ability to localise each fluorescent molecule in the sample [Rust et al., 2006]. The accuracy of an localisation is practically limited by the sample and raw images quality. Samples for localisation microscopy has to have an optimal labelling density. Such density would allow to accurately image the sample structure, and at the same time would not be too high to produce images with multiple fluorophore appearances in close proximity. Additionally, fluorescent labelling can cause additional artefacts called false positives and negatives, where the fluorescent marker is attached to a wrong structure or is missing from the structure of interest. Lastly, the the noise in the raw images is the factor limiting the accuracy of localisation. The noise is usually caused by the shot noise of the camera, high density of labelling (fluorescent labels attached to wrong structures or free fluorophores in the sample or in the buffer).

The super-resolution image is reconstructed using a software package. Positions of imaged molecules can be calculated by fitting a function (Gaussian or theoretical PSF) to the observable PSFs of each of the molecules or centroiding algorithms. The centroiding algorithms use a weighting method: the centre is found by finding the "centre of mass" of the intensity of pixels. The calculation used for this process are exactly the same as, for example, for finding a centre of mass of a physical object. These methods suffer from slightly longer running time than the centroiding methods (for example analysis of 20000 frames with ThunderSTORM centroding algorithm takes around 7 minutes, when analysis of the same data set using Gaussian fitting takes between 15 to 24 minutes [Hirvonen et al., 2015]). However, they are suppose to provide better accuracy of results, because are less influenced by background [Small and Parthasarathy, 2014].

Thus, the accuracy of a single localisation is really limited by the number of photons emitted by a fluorophore. Each molecule is localised with uncertainty dependent on number of photons detected and can be found using an approximated equation 1.9 [Thompson

et al., 2002].

$$\sigma_x^2 = \frac{r_0^2 + q^2/12}{N} + \frac{8\pi r_0^4 b^2}{q^2 N^2}, \tag{1.9}$$

where r_0 is the standard deviation of the PSF, N is the total volume of photons collected during the acquisition, q is the size of the pixel in the image, and b is the background noise per pixel. According to the equation 1.9, better localisation of a molecule is a result of higher photon count or lower background. However, the main pitfall of this approach is that it relies on the least-squares criterion, which is best suited for a Gaussian distributed processes [Ober et al., 2004]. Another method of calculating the localisation accuracy had been proposed by Ober et al. [2004]. This method was defined using a statistical measure of a Fisher information matrix to measure the amount of information that single molecule localisations provide about the unknown structure of the sample. The Ober localisation accuracy limit is defined as:

$$\delta = \frac{\lambda_{em}}{2\pi n_a \sqrt{\gamma A t}}, \tag{1.10}$$

where A is the photon emission rate of a single molecule, γ is the optical efficiency of the detection system ($\gamma \in [0, 1]$), which quantifies the probability that an emitted photon is going to be detected by the detector in acquisition time t. This accuracy testing method is called fundamental because it does not take into account any effects hindering the resolution such as pixelation or noise occuring in experimental systems [Ober et al., 2004].

Depending on the type of fluorescent molecule used one can divide localisation microscopy methods into specific categories. PALM uses photoactivable dyes (for example paGFP, pa-mCherry etc.) [Betzig et al., 2006] where STORM uses photoswitchable dye pairs (for example Cy3-Cy5) [Rust et al., 2006]. Spectral Precision Distance Measurement (SPDM) relies on high precision localisation usually of two fluorescent dyes (with different colour) and measuring their fluorescent lifetime [Lemmer et al., 2008]. PALM with Independently Running Acquisition or PALMIRA, in which due to use a rapid blinking

fluorophores (e.g. rsFastLime, derived from Dronpa) there is no need for an activation step [Egner et al., 2007]. Ground State Depletion and Individual Molecule rerun (GSDIM) uses fluorophores which are more likely to remain in dark, non-emitting state after activation (for example cross to the triplet state) [Fölling et al., 2008]. directSTORM, usually shortened to dSTORM, uses similar principle: single fluorescent dyes, for example Alexa Fluor 647, are directly activated (in contrast to STORM which uses dye pairs). Upon activation small amount of the fluorescent molecules are going to emit light, when a large fraction of fluorophores will remain in long lived dark state [Heilemann et al., 2008].

However, localisation microscopy was originally established as a 2D technique, a number of approaches enable 3D imaging. All of the available 3D localisations methods take an advantage of the z-dependency of the detected PSF of a single molecule. In biplane illumination PALM the sample is illuminated at two z-positions, above and below the plane in-focus. The image detected from the separate z-planes can be used to reconstruct the actual z-position of the single molecule [Juette et al., 2008]. This approach achieves an axial resolution of 75 nm and has imaging depth of 1 μm. Other 3D method uses an astigmatic lens to distort the PSF of a single molecule. The detected PSF is elliptical and by analysing its x and y dimensions, the exact z position can be then recovered computationally. The astigmatic method provides an axial resolution about 50 nm for depth of 600 nm [Huang et al., 2008]. The third approach modifies the shape of PSF to a double-helix using a phase mask (placed in the detection path) [Pavani et al., 2009], called double-helix PSF imaging or DH-PSF, or using a glass wedge in technique known as phase ramp imaging localization microscopy (PRILM) [Baddeley et al., 2011]. The PSF shape in both of cases has a similar appearance with a lobe sections forming a double-helix. The relative position and orientation of the sections of the PSF can be used to compute the axial position of a single molecule, with resolution of 40 nm (depth of field up to 2 μm) [Baddeley et al., 2011; Pavani et al., 2009].

The disadvantages of localisation microscopy are mostly connected to the image acquisition. A super-resolution image in localisation microscopy is reconstructed using long sequences of images. Acquiring an image sequence takes time, which is a major challenge when imaging live samples. Long acquisition times can also lead to drift artefacts in

the reconstructed image. Another drawback of collecting long sequences of data is their size. Also, super-resolution image reconstruction requires specialised programs and takes time [Soeller and Baddeley, 2013]. Imaging samples stained with carbocyanine dyes (for example Cy5, Alexa Fluor 647), generally requires use of an oxygen scavenging buffer, which has an acidic pH that damages biological structures over time [Endesfelder et al., 2011].

Lastly, localisation microscopy can be used along with other super-resolution techniques to collect more complete information about the imaged sample. Generally, localisation microscopy provides the highest resolution improvement. However, it performs worse than SIM or STED in other areas. For example imaging of live cell samples is relatively easy and fast with SIM microscopy, possible using high density activation approaches for localisation microscopy, but nearly impossible using STED. Additionally, localisation microscopy and STED require dedicated dyes with optimal properties (e.g. long lived dark state for localisation methods or activation spectrum enabling high depletion rate for STED), while SIM can be performed with any fluorescent labels. Another major difference in performance of these techniques is an ability to image more than one component of the sample. SIM microscopy can image more than two components and is only limited by time and available light sources. In localisation microscopy and STED usually imaging up to two different tagged structures is possible, but beyond two it is challenging because it is hard to find fluorophores which would have good performance characteristics. Lastly, 3D imaging is possible using all three techniques. STED and localisation microscopy can provide similar resolution improvement (around 30 nm axial resolution [Ding et al., 2010; Schermelleh et al., 2010; Schmidt et al., 2008]). 3D-SIM usually has resolution of 250-300 nm [Schermelleh et al., 2010].

1.6   Summary

Optical microscopy is a vital tool for examining small features in organic and inorganic samples and is a essential tool for imaging of biological samples. However, many biological structures are transparent in visible light. Thus, the invention of fluorescence microscopy was important for biological imaging. Today, transparent structures are tagged with fluorescent dyes and imaged with light from the visible spectrum. This enables imaging of only the areas of interest, what with use of different types of fluorescent tags helps to study physiology and dynamics of biological processes. The next important milestone was overcoming the diffraction limit with methods called super-resolution microscopy. Super-resolution microscopy enables imaging in sub-cellular level of live and fixed cells and provides data for further qualitative and quantitative analysis.

Chapter 2

Clustering analysis for localisation microscopy

2.1 Clustering: Introduction

Cluster analysis is a mathematical analysis method looking for groups of points in the data. Depending on the chosen criteria the points can be grouped in a cluster due to be similar position in space, sharing the same species, visual characteristics or function. In localisation microscopy the structure of interest is reconstructed using single molecule localisations and usually the features of interest appear as high density areas. The shape of such aggregation can be different, for example clusters formed on the cell membrane can be circular or elongated. Additionally, the distribution of points in clusters can also differ depending on the type of structure or sample preparation. For example, clusters can have a uniform distribution with sharp cut-off at the edge, as for example, for discussed later in this chapter DNA origami plates, Gaussian [Sengupta et al., 2011], or Lorentzian/Cauchy [Pertsinidis et al., 2013]. Thus, the performance of our clustering method will be tested on uniform and Gaussian circular clusters, and uniform rectangular clusters.

Clustering analysis was also used for data analysis in other fields, for example, physiology, biology, statistics, pattern recognition, information retrieval, and data mining [Stein-

bach et al., 2003]. One of the the first examples of the use of clustering analysis is creation of taxonomy of animals by Carl Linnaeus[1]. More recently cluster analysis has been famously used to analyse gene data in order to look for clusters of genes with similar functions [Eisen et al., 1998; MacKay, 2003].

Analysis of localisation microscopy data was previously attempted using a number of approaches. This chapter will discuss the basic operation of four most prominent examples of clustering analysis used for localisation microscopy data. The most broadly used method for this type of data is Ripley's functions (see section 2.2.1), which can detect clustering by comparing the data distribution with the random distribution. The other three clustering methods, which are going to be discussed, are Density-Based Spatial Clustering ANalysis (DBSCAN), which is based on detecting differences in density (see section 2.2.2), k-means, where the data is partitioned into k clusters (see section 2.2.3), and hierarchical clustering, which groups the data using a definition of similarity (see section 2.2.3). Lastly, the strength and weaknesses of these four methods will be discussed and compared with our method.

The main part of this chapter is dedicated to our method of clustering analysis based on the Rényi divergence. The Rényi divergence uses a similar principle to Ripley's functions (see section 2.4). It compares the data distribution to the reference distribution (for this study the reference distribution was selected to be a clustered distribution). However, its main advantage arises from it robustness to noise due to a scaling coefficient, called α. The scaling coefficient promotes clustered areas (with high density) over the noise (low density areas, see section 2.6.2). Thus, selection of an α coefficient is vitally important for accuracy and precision of the cluster size measurement and will be discussed extensively later (see section 2.6.4).

The last part of this chapter presents results of cluster analysis performed with our method and Ripley's H function of simulated and microscopy datasets. The simulated datasets were created to be similar to clusters present in the cell membrane (see for example, [Owen et al., 2013; Williamson et al., 2011]). Because the exact shape and

---

[1]Linnaeus divided all living animals into six classes - clusters - of mammals, birds, amphibians, fish, insects, and worms.

distribution of points in these clusters is still unclear, two types of clusters were simulated uniform and Gaussian distributed circles (with radius 8 or 16 pixels, see section 2.6.4). Noise was also added to simulated data sets with varying density. For purpose of this study the noise was modelled to originate from the false positive localisations (usually the false positive localisations can arise from the free floating dye, non-specific binging or cross-talk) and had uniform random distribution. Lastly, localisation microscopy data of rectangular DNA origami plates were analysed (see section 2.6.5). The DNA origami plates were selected because the information about their size is readily available. However, the Rényi divergence and Ripley's functions are not optimal tools to be used for non-circular clusters (because they use a circular shape to measure clustering). Thus, both methods were also tested on simulated data to account for their behaviour with different shaped clusters.

It should also be noted that a number of images of simulated data in this chapter was blurred to improve the visibility. This process was performed using a GIMP tool, called Gaussian blur (symmetrical 5 pixel). Also, the most of the discussion of the clustering in this chapter is performed for one type of cluster (namely disc clusters with radius of 8 pixels). Similarly, the Rényi divergence was calculated using a scaling coefficient $\alpha=70$. The same values were used intentionally for consistency and simplicity. Moreover, $\alpha=70$ is a value high enough to provide a good filtering tool for the most of the noise levels discussed here (it fails only when the density of noise is almost equal to the density of the data distribution).

2.2 Clustering analysis for localisation microscopy

2.2.1 Ripley's functions and the pair correlation function (CPF)

Ripley's K function is an analysis method used to study point patterns by comparing them to a uniform distribution. The function was defined by D. B. Ripley as the ratio of the number of points inside a circle surrounding a central point and the average density

of points on the image (see Figure 2.1), which can be written as:

$$K(r) = \lambda^{-1} \left[\sum_{i=1}^{N} \sum_{j=1}^{N} I(d_{ij} < r) \right] / N \qquad (2.1)$$

where $\lambda = N/A$ and is the average density of points on an image with area A. I is the indicator function, which equal to 1 when the jth point is inside radius r around a central point [Ripley, 1976]. The K function is usually normalised to a more useful form called

(a)   (b)
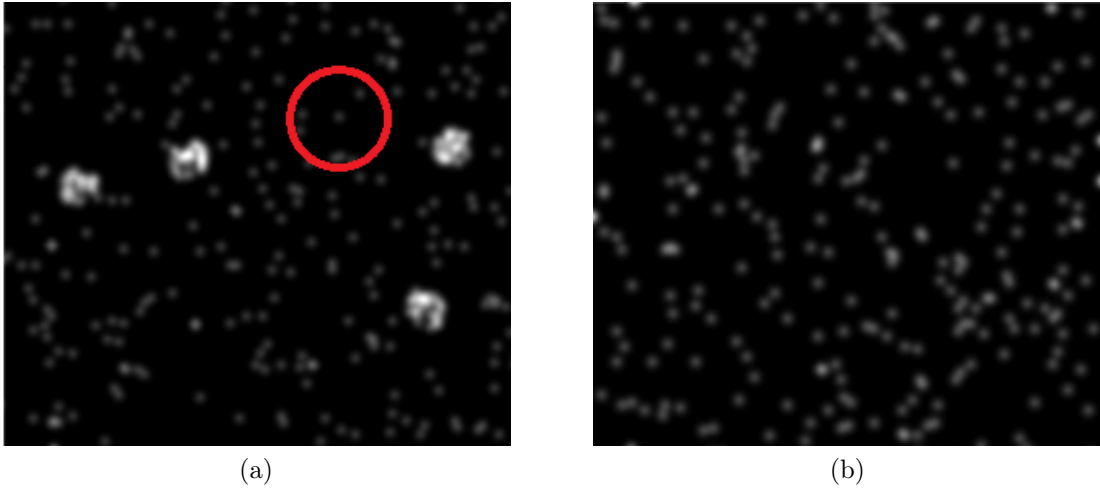
Figure 2.1: The principle of Ripley's K function. Simulated images of a) a clustered distribution with noise and b) a uniform distribution. Ripley's K function performs calculations for every point in the data set. Every point in the data set will be treated as a centre point of circular area with radius r (see red circle on image a)) and all the points positioned inside of the circle will be counted. The final result of this calculation is a total number of instances when two points of the data were positioned in or closer than a certain proximity. This number is then normalised by the number of points in the data and compared to uniform reference distribution.

L(r) (which has a linear expected value)

$$L(r) = \sqrt{K(r)/\pi}. \qquad (2.2)$$

The presence or absence of clustering can be deduced by comparing the value of the L(r) function to the linear expected value. For a randomly distributed sample the L(r) function is equal to the expected value. If the value of L(r) is higher than the expected value the function is detecting clustering in the data (see Figure 2.2).

The L(r) function can be normalised to a function called H(r) which has an expected value equal to 0 for randomly distributed points:

$$H(r) = L(r) - r. \tag{2.3}$$
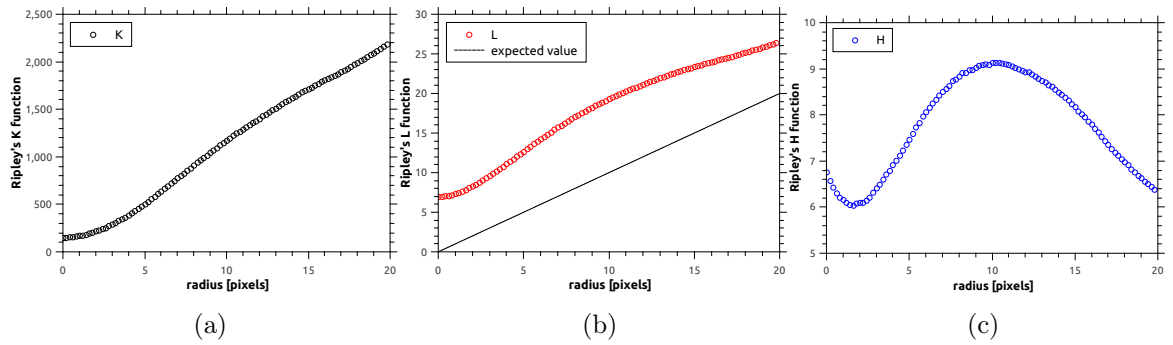
Figure 2.2: Ripley's K, L, and H functions calculated for simulated data with 10 circular clusters (radius 8 pixels) and background noise. a) Ripley's K function (black), exhibits a small change in gradient of the curve around r=10 pixels, but no quantitative measurement is possible. b) For Ripley's L function (red), a linear function can be used to detect clustering by comparing behaviour of the experimental L function with its the expected value for a random distribution (black line). If the L function has values higher than the expected value the data set is clustered. The L function values equal the expected value are an indicator of a random data distribution. c) Ripley's H function (blue) can be used for measuring cluster radius using H_{MAX}, or by detecting the sign change from positive to negative of the first derivative of the H function.

The H(r) function can also provide information about the presence or absence of clustering. A positive value of the H-function can indicate the presence of clustering (a negative value means a random distribution) [Annibale, 2012; Deschout et al., 2014;

Hess et al., 2005; Kiskowski et al., 2009; Scarselli et al., 2012]. A number of authors also have used the radius for $H(r)_{MAX}$ to deduce the average number of molecules in a cluster [Deschout et al., 2014; Sengupta et al., 2011; Zhang et al., 2006]. Information about the degree of clustering can be deduced from the magnitude of the H(r) function and used for a comparison between two or more datasets with different conditions [Deschout et al., 2014; Hess et al., 2007].

There are two different methods of calculating the size of clusters using Ripley's H function. The maximum of the H(r) function can be used as a means to provide the cluster domain size (the function has a maximum for the radius for which there is maximal point density in the cluster) [Owen et al., 2013; Parton and Hancock, 2004; Zhang et al., 2006] (see Figure 2.2). However, it has been noted that there could be some difference between the radius of the biggest point density and the radius for which there is a maximum of H(r) function [Kiskowski et al., 2009; Stoyan and Penttinen, 2000; Zhang et al., 2006]. In their paper Kiskowski et al. (2009) found that the maximum of the H(r) is a measure of cluster separation rather than a measure of cluster size. Another study discussed the issue of Ripley's H function not providing an accurate measurement of the cluster size [Shivanandan et al., 2015a]. Others have also suggested that there is a need for a pre-processing step to account for limited localisation precision [Deschout et al., 2014; Shivanandan et al., 2015b].

In their study, Kiskowski et al. (2009) claimed that the difference between the actual cluster radius and the radius giving the maximal value of H(r) is caused by contributions from points at shorter distances than the size of the cluster. This effect can be eliminated by using the first derivative of the H(r) function to measure cluster radius, which is one version of the pair correlation function (PCF). The radius is measured by detecting the first negative value of the derivative (see Figure 2.3). However, for datasets with noise the point for which the first derivative changes sign is strongly influenced by the noise (see Figure 2.3b).

The Pair Correlation Function is another function derived from Ripley's K function and it measures clustering in rings instead of circles. In this method a smaller and a bigger threshold are used to measure inter-point distances. The main difference between the PCF
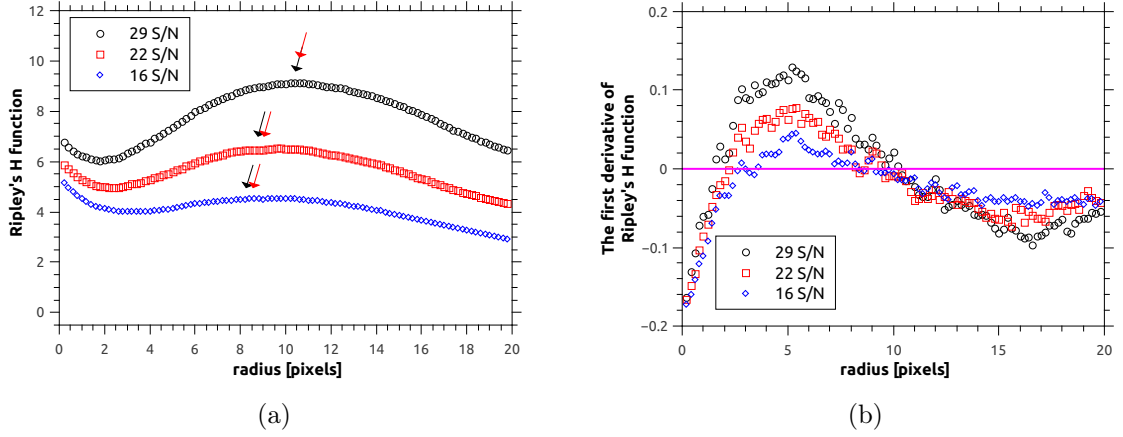
Figure 2.3: Cluster radius measurement using Ripley's H function and artefacts caused by noise. Ripley's H function was calculated for three data sets with ten round clusters (radius 8 pixels) with noise. Noise was added to the signal with three different densities to create data with a S/N 29 (marked with black dots), 22 (red squares), and 16 (blue diamonds). a) Ripley's H function was calculated for data sets with increasing noise levels. For the higher noise level, the function has smaller values and its maximum is less pronounced. According to the methodology presented in [Kiskowski et al., 2009] the radius was found by looking for the sign change of the first derivative of Ripley's H function (from positive to negative). Radius values measured with this method are marked with black arrows. The maximum of the H function is marked with red arrows. b) The first derivative of Ripley's H function calculated numerically using symmetrical equation (see section 2.6.1). For higher noise levels the noise in the first derivative calculation starts to play significant role in the crossing value and affects the measured cluster radius.

and Ripley's function is that the PCF only considers a range of distances, looking at the points inside of a ring instead of a circle (see Figure 2.4). The pair correlation function is related to the Ripley's K function by:

$$g(r) = \frac{d}{dr}K(r)/(2\pi r) \ \text{ for } r \geq 0. \tag{2.4}$$

The pair correlation function can be calculated for any value of radius from zero to infinity. The value of the pair correlation function $g(r)$ for a completely random data set is equal
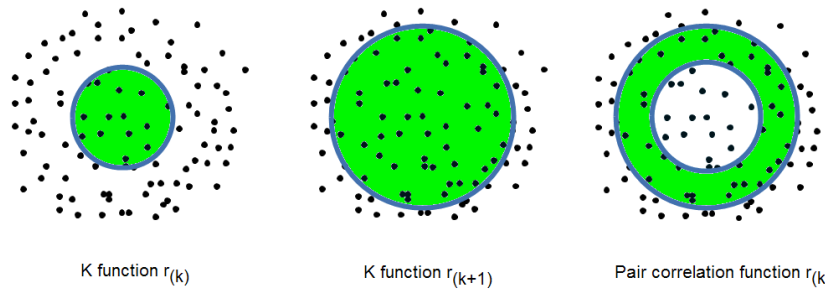
Figure 2.4: The pair correlation function (PCF) operation mechanism. The PCF (right) is calculated as a difference between two Ripley's K functions values calculated for inner and outer radius (left and middle).

to 1. If the value of the pair correlation function is larger than 1 then there is clustering in the data. For large values of radius $g(r) \rightarrow 1$ [Stoyan and Penttinen, 2000].

The H(r) and PCF functions can be modified so they account for artefacts due to the labelling and imaging process. For example, pair correlation PALM (PC-PALM) combines the PCF function with an algorithm which attempts to distinguishes actual clusters from multiple appearances of the same fluorophore. This approach assumes that a fluorescent molecule can be activated repeatedly before is bleached permanently. Thus using an experimentally obtained time parameter, t_d, after which molecule is likely to be bleached, all molecule appearances in close proximity of a previous localisation within time t_d are counted as reappearance of the same molecule. After time t_d all appearances close to the previous localisation are counted as new molecules [Annibale et al., 2011; Annibale, 2012; Sengupta et al., 2011]. Other modifications were made to account for curvature of 2D samples in the z direction by creating a topographical image of the sample surface [Owen et al., 2013]. In this work Owen (2013) also uses the Ripley's function to find clusters in 3D images.

2.2.2 DBSCAN

DBSCAN is a density based clustering algorithm which finds regions of high density surrounded by regions with low density. The first step of the algorithm is to divide all points in the image into three categories:

- core points – these points are positioned in the centre of the clusters; a point can be called a core point if the density of points inside a set radius is higher than threshold and the density outside that radius is smaller,

- border points – these points are separated by less than set radius from the core point.

- noise points – is a point which is not a core point and is further away from the closest core point than a set radius.

Then all noise points are eliminated from further examination. All core points which are positioned closer to each other than a distance within the set radius are separated by a border line (creating two separate clusters). The points forming the cluster are called border points (these points are positioned across the whole area of the cluster not only on the edges) are then assigned to the closest core point [Steinbach et al., 2003].

The most important strengths of DBSCAN are that it does not define cluster shape and the number of clusters does not have be known in advance. DBSCAN is not highly influenced by noise, because noise is removed in the beginning of the algorithm operation. The performance of the DBSCAN algorithm is shown in Figure 2.5 (image adapted from the original DBSCAN paper [Ester et al., 1996]). The original paper [Ester et al., 1996] also presented a test to asses DBSCAN performance compared with a k-mean type algorithm, called CLARANS[2], to underline its strengths: no cluster shape restriction and improved speed (DBSCAN is more than 200x faster than CLARANS). The main drawback of this method is that the all of the clusters in the data set have to have similar density in order to be found. On a data set with two clusters, of which one has low density and

---

[2]A Clustering Algorithm based on Randomized Search (CLARANS) operation is based on looking for clusters by dynamically selecting a portion of data to search and applying data partitioning around means. K-means method is going to be discussed in section 2.2.3

the second high, DBSCAN will only find one cluster and the second will be categorised as noise. This issue is addressed by for example the OPTICS algorithm (Ordering Points To Identify the Clustering Structure). OPTICS is a density based algorithm which does not use a global definition of density[3]. The major disadvantage of the DBSCAN algorithm, especially when imaging big data sets such as localisation microscopy data, is the fact that it is hard to define density. Thus, analysis of big data sets may be expensive computationally [Ankers et al., 1999; Steinbach et al., 2003].
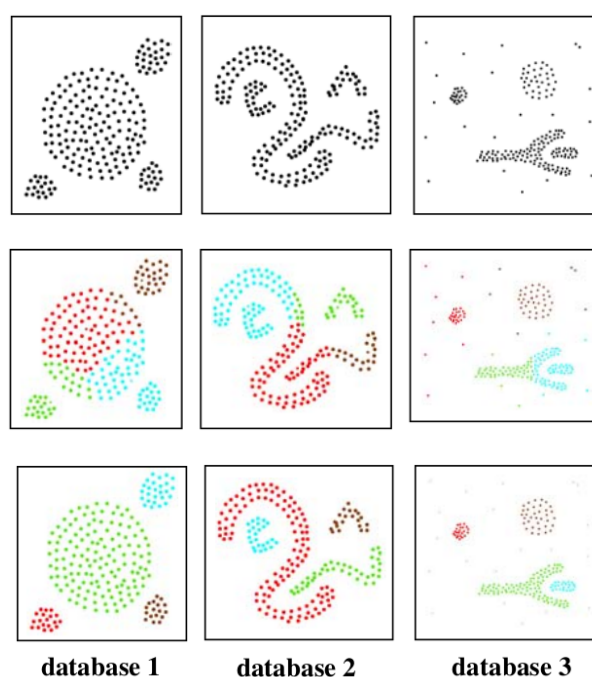
Figure 2.5: DBSCAN performance test. Top row: Different data sets with clusters. The clusters can be easily spotted by human eye. Middle: Clusters found in the data by an algorithm based on k-means method (CLARANS). Bottom: Clusters found with DBSCAN algorithm. Separate clusters found by algorithms are marked with different colours. The noise is not included in any cluster and is marked with light grey dots in bottom right figure. Image adapted from: [Ester et al., 1996].

The clustering analysis methods based on the DBSCAN principle have been used for data analysis of localisation microscopy images, for example for analysis of β2-receptors

---

[3]All points in the data set are ordered by their distance to another point. These distances are stored and are used for setting local density by creating histograms or by use of the hierarchical clustering methods [Ankers et al., 1999].

in HeLa cells [Annibale, 2012] or RNA polymerase organisation in Escherichia Coli [Endesfelder et al., 2011]. However, in most cases a modification of DBSCAN is used where the DBSCAN algorithm is parallelised using hierarchical clustering, providing faster analysis [Li and Xi, 2011]. The OPTICS algorithm has also been used to quantify the degree of clustering and co-localisation of two types of synaptic molecules [Guo et al., 2013; Pertsinidis et al., 2013].

2.2.3 Other methods of clustering analysis

K-means clustering

The k-means algorithm is an example of a iterative algorithm. One of the parameters of the algorithm is the initial number of centroids k which needs to be set by the user. The initial positions of centroids are usually selected randomly from the data points. All points are assigned to the closest centroid position forming k clusters. Next, new centroid positions are calculated. The last two steps are repeated until all centroid positions remain in the same place. Sometimes a weaker stopping condition is defined, e.g. the algorithm stops if less than 1% of points change their cluster affiliation in each iteration [MacKay, 2003; MacQueen, 1967; Steinbach et al., 2003]. An example of the k-means iteration cycle

(a) Iteration 1.   (b) Iteration 2.   (c) Iteration 3.   (d) Iteration 4.
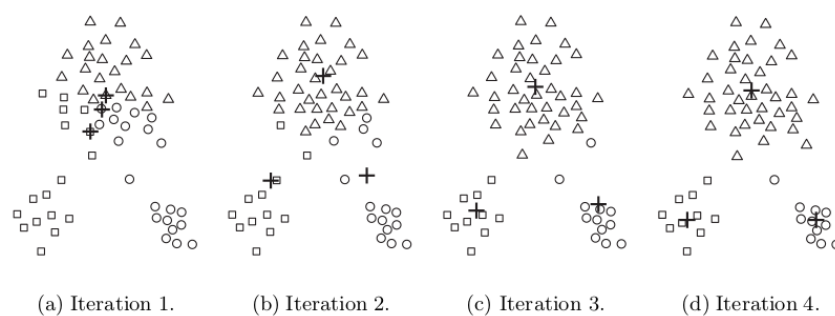
Figure 2.6: K-means algorithm operation on three-cluster data set. Points belonging to a different centroid are marked with squares, circles and triangles. Four iterations of the algorithm for $k = 3$ are shown. The centroid positions are indicated with black crosses. Image adapted from: [Steinbach et al., 2003].

is shown in Figure 2.6. The first iteration in Figure 2.6a shows all the points are assigned to the centroids placed around the middle point of the image. The new centroid positions are calculated for all points in the clusters. In the second iteration all the data points are reassigned to the new centroids shown in Figure 2.6 b). Iterations 3 and 4 are shown in Figure 2.6 c) and d) are done in the same way.

The k-means algorithm requires a definition of the proximity function which quantifies which points are the closest to the centroid. A standard measure for Euclidean data is Euclidean distance (L_2). For points p and q in n-dimensional Euclidean space this is given by

$$L_2 = \sqrt{\sum_{i=1}^{n} (p_i - q_i)^2}.$$

K-means is relatively simple and easy to use algorithm, and can be used for various types of data e.g. to find pore complexes in nuclear envelopes of oocytes [Löschberger et al., 2014]. K-means can find clusters in data with varying density and is time efficient. The main issues of this algorithm are that it can only detect circular clusters (see middle row of Figure 2.5) and the number of clusters in the data must be specified by the user. Usually the latter problem is resolved by setting the number of clusters higher than necessary [Purwar et al., 2011]. This results in localisation of actual clusters and a number of false positives. False clusters can be removed manually or by the clustering analysis software itself. Lastly the k-means algorithm can only be used for data which has a defined mean value. Usually k-menoid[4] does not have the same problem but it is more costly computationally [Steinbach et al., 2003].

Agglomerative Hierarchical Clustering

Hierarchical clustering methods are relatively old compared to Ripley's K function. However, they are still widely used for building visual clustering maps or trees underlying similarities in data sets. There are two basic types of hierarchical clustering:

- agglomerative - algorithm starts by treating every data point as a single cluster. Then it merges close points to one cluster. The algorithm requires an input param-

---

[4]K-menoid searches for the most representative point of a group of points. It can be applied to both discrete and continuous data sets.

eter defining cluster size (also called proximity), which is used to group points into
clusters;

- divisive - the reverse of the agglomerative clustering method. The analysis starts
  with one big cluster containing all points. In each step the algorithm decides which
  clusters to split. Analysis ends when all clusters contain a single point [Mooi and
  Sarstedt, 2011; Steinbach et al., 2003].

Generally, the agglomerative hierarchical clustering methods are more commonly used to
analyse information.

The hierarchical clustering outcome depends on selecting two functions defining dis-
tance and proximity measures. Similarly to the k-means algorithm, hierarchical clustering
can use a range of distance measures. Then the proximity function is selected. The prox-
imity function is defined as the distance between two clusters and it is usually set taking
into account the type of the data. For example, the MIN function, shown in Figure 2.7,
finds a shortest link between two clusters and points in these clusters are equally im-
portant. Another example of proximity function is centroid link, where the proximity
is defined as a link between centres of mass of two clusters. The final outcome of the
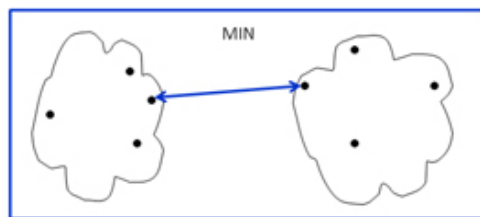
Figure 2.7: Graph interpretation of MIN proximity function. The MIN proximity is also
called single link. Image adapted from: [Verma, 2015].

analysis depends on the selected proximity definition.

The particular features of hierarchical clustering are its lack of one global function
defining clustering and its ability to detect different sizes of clusters. The main feature
of this method is that there is no global clustering function, since all decisions about
cluster merging are taken on a local level (given by the proximity function). This allows

the algorithm to avoid large scale optimization problems or selecting starting points. Another issue with the algorithm is that once merged, clusters are merged permanently (the information about previous steps is not saved). For some types of proximity function the clustering properties of the algorithm are not affected by the cluster size. This is true, for example, for Ward's proximity, which is defined as the increase in the square error caused by merging two clusters together [Steinbach et al., 2003].

The main applications of hierarchical clustering are connected to creating taxonomies, for example in evolutionary biology. It has also been used as a secondary analysis tool for characterising HIV infected HeLa cells [Lelek et al., 2012]. Clusters in localisation microscopy images were identified by a home-written software[5] and the morphology of the identified clusters was further studied using the hierarchical clustering algorithm. Similarly, DBSCAN has been aided by hierarchical clustering to analyse localisation microscopy images to identify insulin dependent clustering in RBL-2H3 cell line [Křížek et al., 2014].

2.3   Clustering algorithms comparison

Clustering algorithms discussed in this chapter are best suited, depending on their strengths and weaknesses, for analysis of different types of data. Clustering algorithms can be compared using a number of criteria, for example, the running time, required input information or preferred type of clusters they can successfully analyse (see Table 2.1). The required input information varies a great deal between algorithms, for example DBSCAN requires a density definition to successfully separate data into clustered and noisy points, the k-means requires the number of clusters to be known in advance, where Ripley's H function and hierarchical clustering require no prior knowledge about the data. The required input information is a major drawback, due to the fact that it is usually hard to provide accurate information about the cluster number or their density for localisation microscopy data, which can contain a great number of small clusters, for example, see

---

[5]A random sample of cluster centre points was selected. For each centre point the algorithm calculated the number of points which were closer to the central point than a set radius. Then the radius was increased, and the analysis stopped when the number of points close to centre point was stabilised.

clustering of adrenergic receptors in [Scarselli et al., 2012]. One of the most used approach to overcome this limitation is to run DBSCAN or k-means by running tests with different starting parameters (density or number of clusters), at a cost of longer total computation time.

The cluster type and size also plays a role in the algorithm selection. Ripley's H function and k-means require circular clusters with similar size and density to provide correct cluster detection. For example, analysis of non-circular clusters with k-means algorithm can lead to incorrect cluster identifications (see Figure 2.5). DBSCAN can analyse any shape and size, however their density needs to be similar. Usually in this situation the lower density cluster will be classified as noise by DBSCAN. Similarly, the noise sensitivity of these methods varies a great deal. DBSCAN has the highest noise tolerance, because it removes the noise points from further consideration. Ripley's H function has a moderate noise resistance – generally low and medium noise causes this method to produce non-precise cluster radius measurements. With high noise Ripley's H function is both imprecise and inaccurate (for example measuring cluster size smaller than it actually is, see section 2.6.2). K-means and hierarchical clustering, both have a low tolerance to noise. Noisy points will be treated as cluster border points and can strongly influence the answer provided by the algorithms.

Lastly, speed of the algorithms is also an important factor to consider. The localisation microscopy data usually contains thousands molecule localisations and the analysis of the clustering present in the data needs to be fast for this amount of data. The fastest from discussed here algorithms is Ripley's H function, which takes seconds to run on a typical microscopy data set (for example for data with 32,000 points it takes 16 s). DBSCAN is a slow algorithm, its computation time is usually around minutes to hours for localisation microscopy data. However, it is possible to better DBSCAN computation time by presorting the data with hierarchical clustering. An example of such faster DBSCAN based algorithm is OPTICS. K-means is usually two magnitudes slower than DBSCAN (see section 2.2.2). Hierarchical clustering algorithms are often slow, however their speed is limited by the number of decisions it needs to make to connect data into clusters.

Table 2.1: Selected properties of clustering algorithms discussed in this chapter.

	Ripley's H function	DBSCAN	k-means	hierarchical
Input information	none	cluster density	number of clusters	none
Cluster size	similar	any	similar	any
Cluster shape	circular	any	circular	any
Cluster density	similar	similar	similar	any
Noise sensitivity	moderate	high	low	low
Speed	fast (seconds to minutes)	slow (minutes to hours)	slow (hours)	slow

2.4 The Rényi divergence: General equation and properties

Divergence is a measure of difference between two distributions and is a type of statistical distance measurement. Rényi proposed a divergence measure linked to his definition of entropy, and as an extension of the Kullback–Leibler divergence (which was defined for Shannon entropy). The Rényi entropy is defined as:

$$H_\alpha = -\frac{1}{1-\alpha} \log \left[\sum_x p(x)^\alpha \right]. \tag{2.5}$$

The Rényi divergence is defined using two probability distributions, p(x) - the data distribution and q(x), a reference distribution to which the data distribution is compared:

$$D_\alpha(p(x)||q(x)) = \frac{1}{\alpha-1} \ln \left[\int_x p(x) \left(\frac{p(x)}{q(x)} \right)^{\alpha-1} dx \right]. \tag{2.6}$$

It has following properties:

1. The Rényi divergence is non-negative for all distributions $p(x_i)$, $q(x_i)$, and for α bigger than 0: $D_\alpha(p||q) \geq 0, \forall_{p,q,\alpha>0}$

2. The Rényi divergence is equal to 0 only when distributions $p(x_i)$ and $q(x_i)$ are equal: $D_\alpha(p||q) = 0$ if $p(x) = q(x) \forall_{x\in\mathbb{R}}$

3. The Rényi divergence for $\alpha = 1$ is equal to the Kullback-Liebler divergence (D_{KL}): $\lim_{\alpha\to1} D_\alpha(p||q) = D_{KL}(p||q)$.

The Rényi divergence value is strongly dependent on the selected value of the coefficient α. If we consider the Rényi divergence calculated as with equation 2.17 we can see that α acts as a scaling or weighting coefficient. That property of α can be explained by considering how the calculations were made. For each central point the number of points were inside of a given radius (radius is the variable) was counted. The number of points inside a radius is then raised to the power $(\alpha - 1)$. This means that the higher density

areas are going to be promoted over low density areas. With higher α values more dense areas are going to be promoted even more.

2.5 Clustering analysis

Cluster analysis for localisation microscopy is based on continuous variables. The single molecule localisation datasets calculated for localisation microscopy images are not pixelated and have a continuous distribution. The Rényi divergence for continuous variables is given as:

$$D_\alpha(p(X)||q(X)) = \frac{1}{\alpha - 1} \ln \left[\int p(x) \left(\frac{p(X)}{(q(X)} \right)^{\alpha-1} \mathrm{d}X \right], \tag{2.7}$$

where the data set is defined using variables $X = \{x_1, x_2, ..., x_N\}$. For the purpose of clustering analysis we use a sampling approximation on equation 2.7

$$D_\alpha(p(x_i)||q(x_i)) \approx \frac{1}{\alpha - 1} \ln \left[\frac{1}{N} \sum_{i=1}^{N} \left(\frac{p(x_i)}{q(x_i)} \right)^{\alpha-1} \right] \tag{2.8}$$

where the probability distribution $p(x_i)$ is defined using a type of kernel estimation techniques, called the Parzen window estimation [Parzen, 1962] as:

$$p(x_i) = \frac{1}{N} \sum_{j=1}^{N} \frac{I(d_{ij} < r)}{\pi r^2}, \tag{2.9}$$

where N is the number of all data points, d_{ij} is the distance between points i and j, r is the radius around the central point, and $I(d_{ij} < r)$ is an indicator function. The probability $q(x_i)$, for the purpose of clustering analysis, is defined using a Parzen window estimation for all points being in one cluster:

$$q(x_i) = \frac{1}{\pi r^2 N_c}, \tag{2.10}$$

where r is the radius of a cluster and N_c is the number of clusters. The Rényi divergence has a general equation for $\alpha \in (1, +\infty)$ and three special cases for $\alpha = 0, 1$ and ∞. The Rényi divergence for $\alpha = 0$ returns the number of clusters in the dataset (provided that there is no noise in the data). For $\alpha = 1$ the term $\frac{1}{\alpha - 1}$ has a zero in the numerator, however it is equal to the Kullback-Liebler divergence in the limit. Lastly, for $\alpha = \infty$ the divergence provides a maximum value of difference between two compared distributions.

2.5.1 The Rényi divergence for $\alpha \in (0, +\infty) \wedge \alpha \neq 1$

Using the equation 2.8 and probability density functions as defined in equations 2.9 and 2.10:

$$D_\alpha = \frac{1}{\alpha - 1} \ln \left[\frac{1}{N} \sum_{i=1}^{N} \left(\frac{\frac{1}{N} \sum_{j=1}^{N} \frac{1}{\pi r^2} I(d_{ij} < r)}{\frac{1}{\pi r^2 N_c}} \right)^{\alpha - 1} \right], \tag{2.11}$$

which can be written as:

$$D_\alpha - \frac{1}{\alpha - 1} \ln \left(N_c^{\alpha - 1} \right) = \frac{1}{\alpha - 1} \ln \left[\frac{1}{N} \sum_{i=1}^{N} \left(\frac{1}{N} \sum_{j=1}^{N} I(d_{ij} < r) \right)^{\alpha - 1} \right]. \tag{2.12}$$

Equation 2.12 can be rewritten by split exponentiation base $((ab)^c = a^c b^c)$ as:

$$D_\alpha - \frac{1}{\alpha - 1} \ln \left(N_c^{\alpha - 1} \right) = \frac{1}{\alpha - 1} \ln \left[\frac{1}{N} \sum_{i=1}^{N} \left(\frac{1}{N} \right)^{\alpha - 1} \left(\sum_{j=1}^{N} I(d_{ij} < r) \right)^{\alpha - 1} \right] \tag{2.13}$$

Then because number of points on the image N is constant and independent on sums over i and j the equation 2.13 can be written as:

$$D_\alpha - \frac{1}{\alpha - 1} \ln \left(N_c^{\alpha - 1} \right) = \frac{1}{\alpha - 1} \ln \left[\left(\frac{1}{N} \right)^\alpha \sum_{i=1}^{N} \left(\sum_{j=1}^{N} I(d_{ij} < r) \right)^{\alpha - 1} \right], \tag{2.14}$$

which is an equivalent of:

$$D_\alpha - \frac{1}{\alpha - 1} \ln \left(N_c^{\alpha-1} \right) = \frac{1}{\alpha - 1} \ln \left[\frac{\sum_{i=1}^{N} \left(\sum_{j=1}^{N} I(d_{ij} < r) \right)^{\alpha-1}}{N^\alpha} \right].$$ (2.15)

Equation 2.15 can be simplified:

$$D_\alpha - \frac{1}{\alpha - 1} \ln \left(N_c^{\alpha-1} \right) = \frac{1}{\alpha - 1} \left[\ln \left(\sum_{i=1}^{N} \left(\sum_{j=1}^{N} I(d_{ij} < r) \right)^{\alpha-1} \right) - \ln \left(N^\alpha \right) \right],$$ (2.16)

which can be written using $\ln(a^b) = b \ln(a)$ as:

$$D_\alpha - \frac{1}{\alpha - 1} ln \left(N_c^{\alpha-1} \right) = \frac{1}{\alpha - 1} \left[\ln \left(\sum_{i=1}^{N} \left(\sum_{j=1}^{N} I(d_{ij} < r) \right)^{\alpha-1} \right) - \alpha \ln \left(N \right) \right].$$ (2.17)

The main purpose of transforming equation 2.12 to the form in 2.17 is to split the exponen-tiation base. While operating a C++ program one needs to consider memory availability allocated to a type of variable used to store the calculated divergence value (here variable type double was used). Calculating the divergence with equation 2.17 returns smaller values than calculation with a standard equation. Therefore, this form of the Rényi di-vergence can still be used for calculation of datasets with 100 thousands of points and with α values smaller than 200 with the memory allocated to the double-type variable. Calculation with higher values of α is possible after using gmp library, a C++ library for calculation on large numbers.

2.5.2   The Rényi divergence for $\alpha = 0$

Using equation 2.8 and for $\alpha = 0$ and $p(x_i)$ and $q(x_i)$ defined by 2.9 and 2.10:

$$D_0 = \frac{1}{0 - 1} \ln \left[\frac{1}{N} \sum_{i=1}^{N} \left(\frac{\frac{1}{N} \sum_{j=1}^{N} \frac{1}{\pi r^2} I(d_{ij} < r)}{\frac{1}{\pi r^2 N_c}} \right)^{-1} \right],$$ (2.18)

the equation can be simplified as:

$$D_0 - \ln\left(N_c\right) = -\ln\left[\sum_{i=1}^{N} \frac{1}{\sum_{j=1}^{N} I(d_{ij} < r)}\right], \tag{2.19}$$

and taking the exponential:

$$e^{-D_0} - N_c = \sum_{i=1}^{N} \frac{1}{\sum_{j=1}^{N}(d_{ij} < r)}. \tag{2.20}$$

This equation returns the number of clusters in the data set (N_c).

2.5.3 The Rényi divergence for $\alpha = 1$

For $\alpha = 1$ the Rényi divergence is equal to the Kullback-Leibler divergence for Shannon entropy. Using the sampling approximation and probability functions $p(x_i)$ and $q(x_i)$ the same as defined in section 2.5.2 the divergence is:

$$D_{\alpha \to 1} = \frac{1}{N} \sum_{i=1}^{N} \ln\left[\frac{\frac{1}{N}\sum_{j=1}^{N}\frac{1}{\pi r^2}I(d_{ij} < r)}{\frac{1}{\pi r^2 N_c}}\right], \tag{2.21}$$

which can be simplified to:

$$D_{\alpha \to 1} - \ln\left(N_c\right) = \frac{1}{N} \sum_{i=1}^{N} \ln\left[\frac{1}{N}\sum_{j=1}^{N} I(d_{ij} < r)\right]. \tag{2.22}$$

2.5.4 The Rényi divergence for $\alpha = \infty$

This case maximises the difference between two compared distributions. The divergence is given by the natural logarithm of the maximal possible ratio of probability distributions $p(x_i)$ and $q(x_i)$ and is given by equation:

$$D_\infty = \ln\left(\operatorname{ess\,sup} \frac{p(r)}{q(r)}\right), \tag{2.23}$$

The essential supremum is defined for measuring theory or for functional analysis, which often are used to consider statements which apply to only a subgroup of the analysed data. The essential supremum, for the Rényi divergence should be understood as a maximal possible difference between the two distributions. It can be written using equations 2.9 and 2.10 as:

$$D_\infty = \ln \left[\max \left(\frac{1}{N} \frac{\sum_{j=1}^{N} I(d_{ij} < r)}{\pi r^2} \pi r^2 N_c \right) \right] \tag{2.24}$$

This can be simplified by cancelling the πr^2 term and moving N_c (number of clusters) outside of the sum and to the left side of equation giving:

$$D_\infty - \ln (N_c) = \ln \left[\max \left(\frac{1}{N} \sum_{j=1}^{N} I(d_{ij} < r) \right) \right]. \tag{2.25}$$

2.5.5   The Rényi divergence for a uniform reference distribution and its special case: Ripley's K function

Ripley's K function is a special case of the Rényi divergence for $\alpha = 2$ and a uniform reference distribution. The uniform reference distribution, indexed "U" for uniform, is given as:

$$q^U(x_i) = \frac{1}{A}, \tag{2.26}$$

where A is the image area. The Rényi divergence equation 2.8, for $\alpha = 2$, a data distribution given by equation 2.9, and a uniform reference distribution (equation 2.26) is equal to:

$$D_\alpha^U = \frac{1}{\alpha - 1} \ln \left[\frac{1}{N} \sum_{i=1}^{N} \left(\frac{\frac{1}{N} \sum_{j=1}^{N} \frac{1}{\pi r^2} I(d_{ij} < r)}{\frac{1}{A}} \right)^{\alpha - 1} \right]. \tag{2.27}$$

Both the uniform and previously discussed clustered reference distributions are point independent (their value does not depend on the point they are calculated for). The

uniform distribution is constant for the whole image and the clustered distribution depends on the number of clusters in the whole image. Thus, one can separate equation 2.27 into a point-specific part and a point-independent part. This means that all of the parameters which have no dependency on i and/or j can excluded from summation. The Rényi divergence calculated for clustered reference distribution using equation 2.12 (for $\alpha \in [0, \infty) \wedge \alpha \neq 1$) can also be arranged into data constant and clustering counting parts:

$$D_\alpha = \frac{1}{\alpha - 1} \ln \left(N_c^{\alpha-1} \right) + \frac{1}{\alpha - 1} \ln \left[\left(\frac{1}{N} \right)^\alpha \sum_{i=1}^{N} \left(\sum_{j=1}^{N} I(d_{ij} < r) \right)^{\alpha-1} \right], \qquad (2.28)$$

where N is number of all points in the data set, $I(d_{ij} < r)$ is an indicator function, πr^2 is area of a cluster, and N_c is the number of clusters. The term $\frac{1}{\alpha-1} \ln \left(N_c^{\alpha-1} \right)$ is constant and provides an offset to the divergence function. The Rényi divergence defined for an uniform distribution differs only by a constant value $\frac{1}{\alpha-1} \ln \left[\left(\frac{A}{\pi r^2} \right)^{\alpha-1} \right]$. Thus, the Rényi divergence for a uniform reference distribution can be written as proportional to:

$$D_\alpha^U \propto \frac{1}{\alpha - 1} \ln \left[\left(\frac{1}{N} \right)^\alpha \sum_{i=1}^{N} \left(\sum_{j=1}^{N} I(d_{ij} < r) \right)^{\alpha-1} \right], \qquad (2.29)$$

Therefore, from the perspective of clustering calculations there is no difference between using the Rényi divergence with a uniform or clustered distribution.

The Rényi divergence for $\alpha = 2$, $q(x_i) = \frac{1}{A}$, and $p(x_i)$ as defined in 2.9 gives Ripley's K function:

$$D_2 = \frac{1}{2 - 1} \ln \left[\frac{1}{N} \sum_{i=1}^{N} \left(\frac{\frac{1}{N} \sum_{j=1}^{N} \frac{I(d_{ij}<r)}{\pi r^2}}{\frac{1}{A}} \right)^{2-1} \right], \qquad (2.30)$$

what can be written as:

$$\pi r^2 \exp(D_2) = \lambda^{-1} \left[\sum_{i=1}^{N} \sum_{j=1}^{N} I(d_{ij} < r) \right] / N, \qquad (2.31)$$

which is Ripley's K function.

2.6 Clustering analysis of simulated and localisation microscopy data

2.6.1 Data simulation and clustering analysis software

The simulated data for cluster analysis was created using custom software written in C++ using the libCVD library. The code created clusters and the noise separately. Its operation started by creating a list of random cluster centre positions (each cluster was placed at a distance from the image border so the whole cluster was visible on the image). Then clusters were created around the centre points with a set density. Two types of clusters were generated, but each dataset only had one of these types. A sharp edge cut-off, disk-shaped clusters (which were created using random number generator) and Gaussian shaped clusters (which were created using a Matlab mvnrnd() function to create a random sample of points with a Gaussian distribution). Noise was randomly added with a predefined density. The noise level was defined as a percentage of noisy pixels and also as Signal-to-Noise (S/N), and calculated as $S/N = \frac{I_{signal}}{I_{noise}}$, where I_{signal} is the intensity of signal and I_{noise} is the intensity of noise from the image patches with the same area. S/N can be calculated in this manner for simulated data, because the true number of points in the signal is known and noise is added to the data after simulating the signal. For the microscopy data S/N was calculated as: $S/N = \frac{I_{signal} - I_{noise}}{I_{noise}}$. The intensities of the signal and noise were estimated as a number of points in signal (here clustered points) and a number of noise points. Examples of datasets of a range of noise levels created for cluster analysis are shown in Figure 2.8.

The Rényi divergence analysis was performed using equation 2.17 which can be used within the memory range of double variables in C++. The Rényi divergence was calculated for α equal to 0, 2 and 10-120 (with increments of 10). The flowchart presenting the optimised software for the Rényi divergence and Ripley's H function calculation is presented in Figure 2.9. These results can be used to estimate the optimal α value for use with the data. The first and second numerical derivatives of the Rényi divergence were calculated using the symmetrical numerical derivative: $f'(x) = \frac{f(x+h) - f(x-h)}{2h}$. For each

(a) (b) (c)

(d) (e) (f)
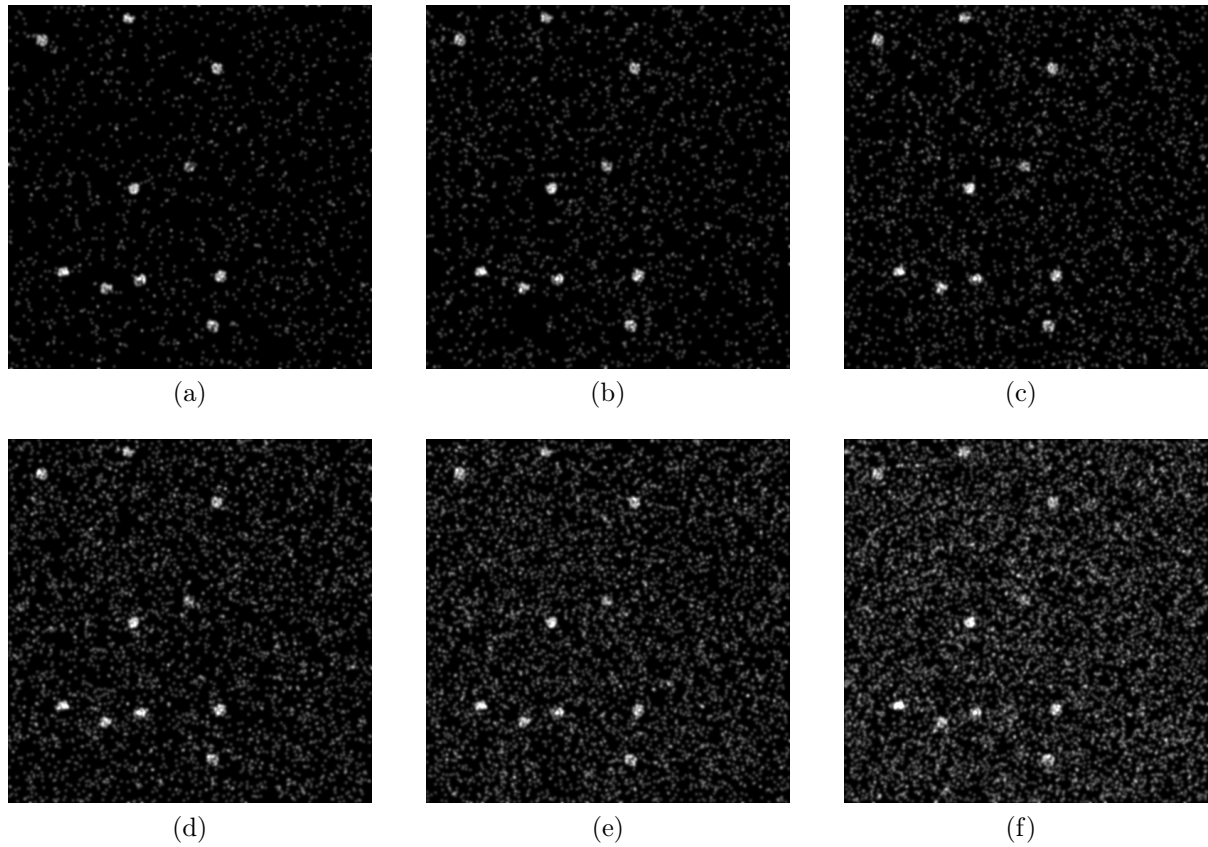
Figure 2.8: Examples of simulated data sets for cluster analysis. Data sets with 10 clusters with a radius of 8 pixels. Noise was added to each data set (a) 0.5% of pixels are noise, which corresponds to a S/N of 29, (b) 0.74% – S/N 22, (c) 1.0% – S/N 16 , (d) 1.5% – S/N 10, (e) 2.0% – S/N 7, and (f) 3.0% – S/N 6. For improved visibility the image was blurred.

data set Ripley's H function and its first derivative was also calculated. The derivatives of the Rényi divergence and Ripley's H function are going to be used for the cluster radius measurement (see section 2.6.2).
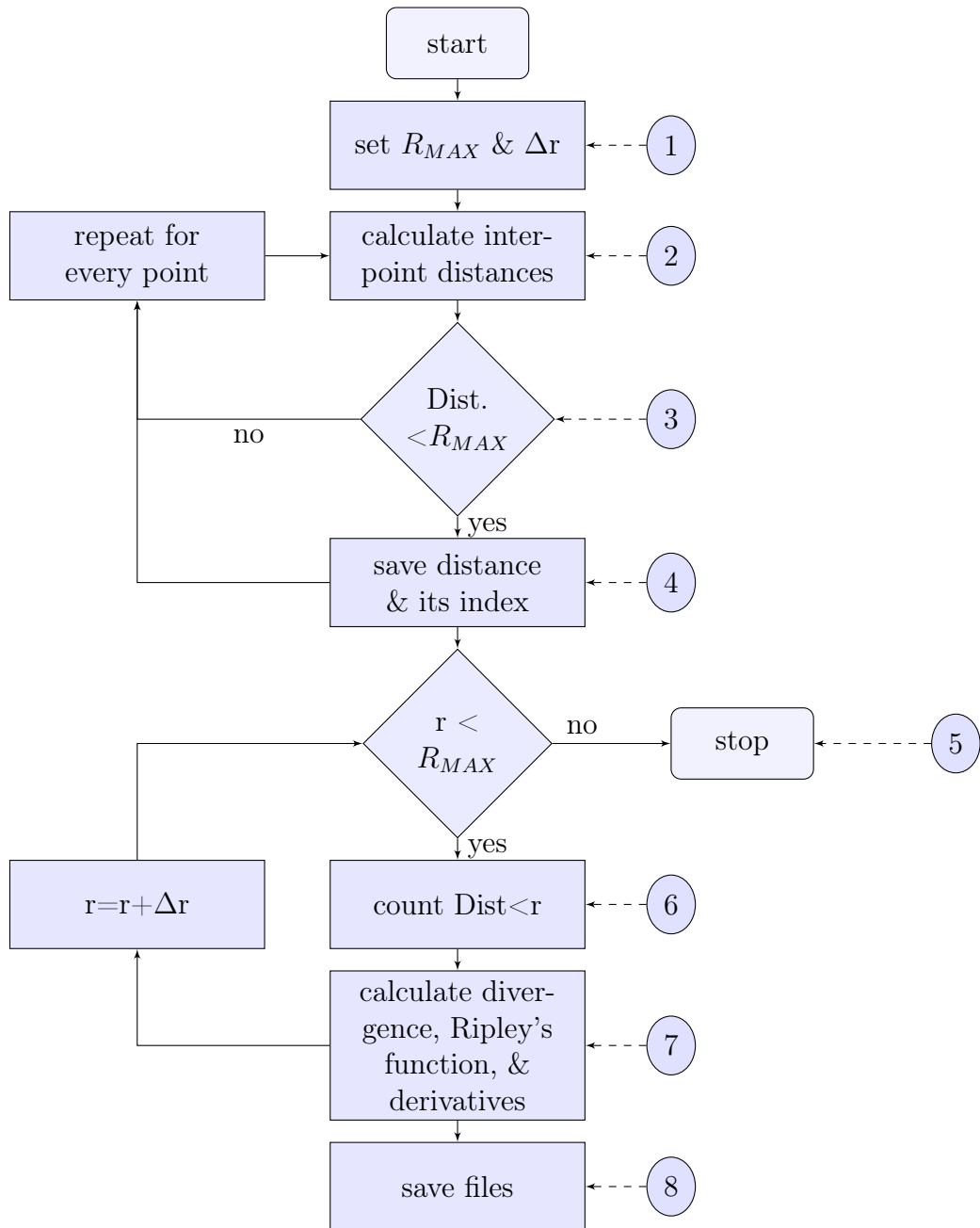
Figure 2.9: The cluster analysis software operation. The software will analyse the data in a range of radius values defined with a maximal radius (R_{MAX}) and radius increment (Δr, step 1). Then the inter-point distances are calculated for every pair of points in the data set (step 2) and code stores only those which are smaller than the maximal radius R_{MAX} (step 3 and 4). Then the Rényi divergence and Ripley's H function are calculated for a range of radius values using previously calculated inter-point distances (steps 5-7) and saved to files for further analysis (step 8).

2.6.2   Cluster radius measurements

The cluster radius was found by identifying a plateau in the Rényi divergence function. If the Rényi divergence is calculated as a function of a radius then the function will have plateau for radius equal (or approximately equal for noisy data) to the actual cluster size. The Rényi divergence function counts all points inside the area defined by a given radius (see Figure 2.10 where areas in which points will be counted are marked with red circles). The Rényi divergence function has different behaviour when calculated for three main groups of radius values. For radius values smaller than the cluster radius the Rényi divergence will increase its value very rapidly (see Figure 2.11 divergence for $\alpha = 120$, radius values 0 to 8.2). Even a small change of the calculation radius includes more point into the calculation (see Figure 2.10a). For radius values similar to cluster radius the Rényi divergence function has a plateau (see Figure 2.11 divergence for $\alpha = 120$ for radius values 8.2 to 10.4). The Rényi divergence value does not change with increasing radius as only a small number of new points will be included in the calculation (see Figure 2.10b). For radius values larger than the cluster radius the Rényi divergence value increases slowly as the calculation radius includes more points from outside of the cluster (see Figure 2.10c).

This means that the radius can be found by looking for a saddle point (first and second derivatives of the function need to be equal to 0). Figure 2.11 shows the divergence calculated for a simulated data set with 10 clusters (radius 8 pixels) with 0.5% (S/N 29) noise. The radius measurement identifies the first point in the plateau as the point at which the first derivative is equal to zero within a set tolerance interval.

Radius measurements using Ripley's H function were performed to compare the results with the divergence method. The cluster radius was calculated using the methodology presented in [Kiskowski et al., 2009] and identified by looking for the radius value for which the first derivative changed sign from positive to negative. However, this method is strongly affected by noise (see Figure 2.12b). Three possible methods for noise reduction were tested: Gaussian kernel smoothing of the first derivative of Ripley's H function, consistency in negative values of the first derivative (where more than one point in a row needs to have a negative value to define a cluster radius measurement), and looking for

<div style="text-align:center">(a)   (b)   (c)</div>
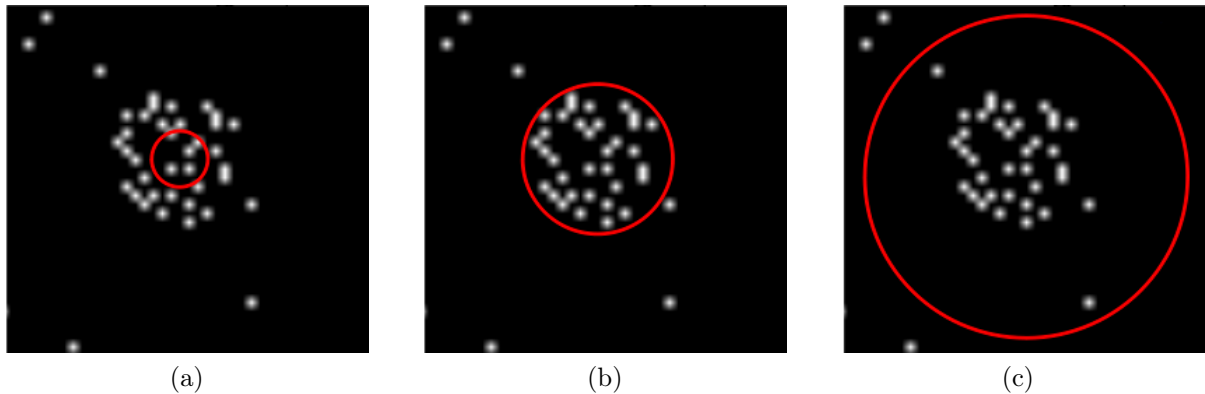
Figure 2.10: The origin of the plateaus in Rényi divergence function. The Rényi divergence function counts all points inside a radius – here marked with red circles. a) The radius for the calculation is smaller than the cluster size. Increasing the radius size will greatly increase a number of points inside the new radius, which causes a sharp increase in the value of the Rényi divergence function. b) Radius is similar to the actual cluster radius. Only a limited number of new points are added with the radius increase. The Rényi divergence function values does not change, forming a plateau. c) Radius is larger than cluster radius. New points (from the background) are included inside the radius. The Rényi divergence function has a small gradient.

the first negative point of the first derivative after the maximum of Ripley's H function. An example of Gaussian kernel smoothing is shown in Figure 2.12b. It should be noted that the smoothing process introduced a new error into the measurements.

To compare the performance of different cluster radius measurement methods for Ripley's H function was compared, it was calculated for data sets created for Monte Carlo testing (see section 2.6.3). These data sets had 10 clusters with 8 pixels radius with varying noise levels (no noise, 0.5%, 0.74%, 1%, and 1.5%). The cluster radius was then measured using four methods: standard, smoothing, consistency (at least three point in a row), and maximum (detecting the first derivative sign change after maximum). The average measured radius was calculated as a mean of the resultant distribution of the cluster radius measurement. The standard deviation was also calculated to measure the distribution spread.

The maximum and consistency methods provided an improvement of noise resistance
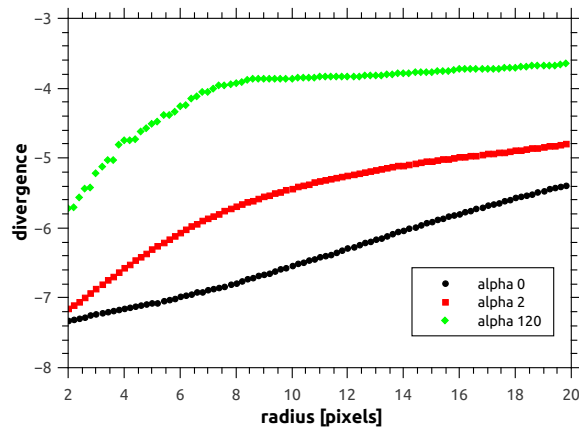
Figure 2.11: The Rényi divergence function calculated for three values of α 0, 2 and 120. The data set was created by the software described in section 2.6.1 with 10 clusters (8 pixels radius) and 0.5% noise (S/N 29). With a higher α value the function plateaus appear for radius values approximately equal to the actual cluster radius. For example, here for α =120 (marked with green) the plateau begins at radius 8.2).
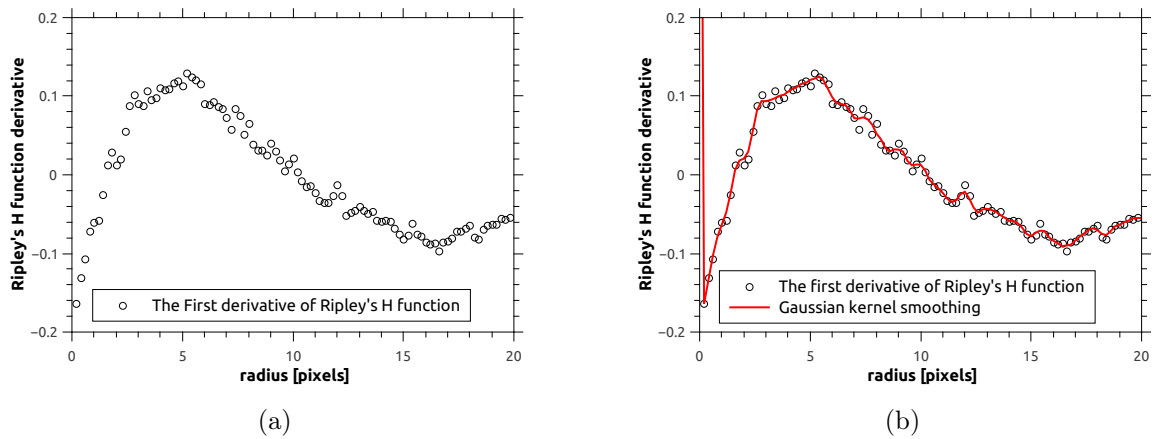
Figure 2.12: The first derivative of Ripley's H function calculated for a data set of 10 clusters (radius of 8 pixels) with 0.5% noise (S/N 29). a) The first derivative of Ripley's H function (marked with black circles). b) The first derivative of Ripley's H function (marked with black circles) and Gaussian kernel smoothing curve (red solid line).

for cluster radius measurement. The most accurate cluster radius measurements were achieved by detecting the first negative value of the first derivative after maximum in

Ripley's H function (see Figure 2.13). In further studies the standard methodology as presented in [Kiskowski et al., 2009] was used, because it is a well established method and the comparison between our cluster analysis method based on the Rényi divergence and the Ripley's H function should be made using a widely used and published methodology.
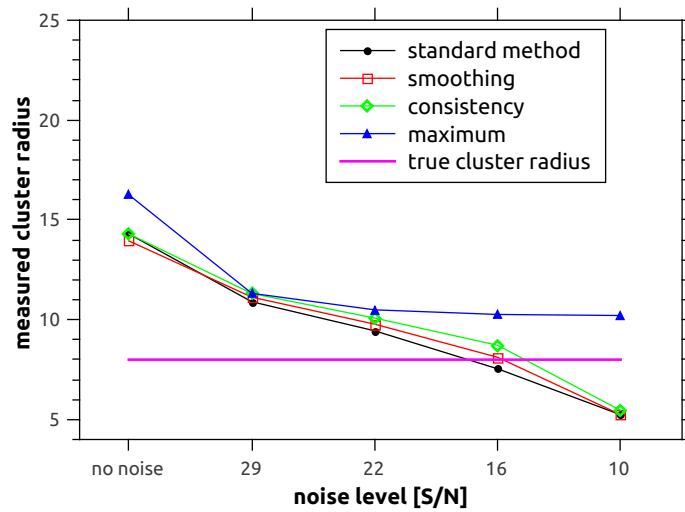
Figure 2.13: Comparison of noise influence on different methods of cluster radius measurement for Ripley's H function. Monte Carlo simulations were performed on data sets with 10 clusters (8 pixels radius). The cluster radius was measured for different noise levels using: a standard method (black) as presented in [Kiskowski et al., 2009] by looking for the point in which the first derivative changes sign from positive to negative, smoothing (red) when the first derivative values were smoothed with a Gaussian kernel and used to detect the sign change in the first derivative, the consistency method (green), which detected the point for which the first derivative remains negative, and the maximum (blue), which detects the first negative value of the first derivative after the maximum of Ripley's H function. The true radius of simulated clusters is marked with magenta line.
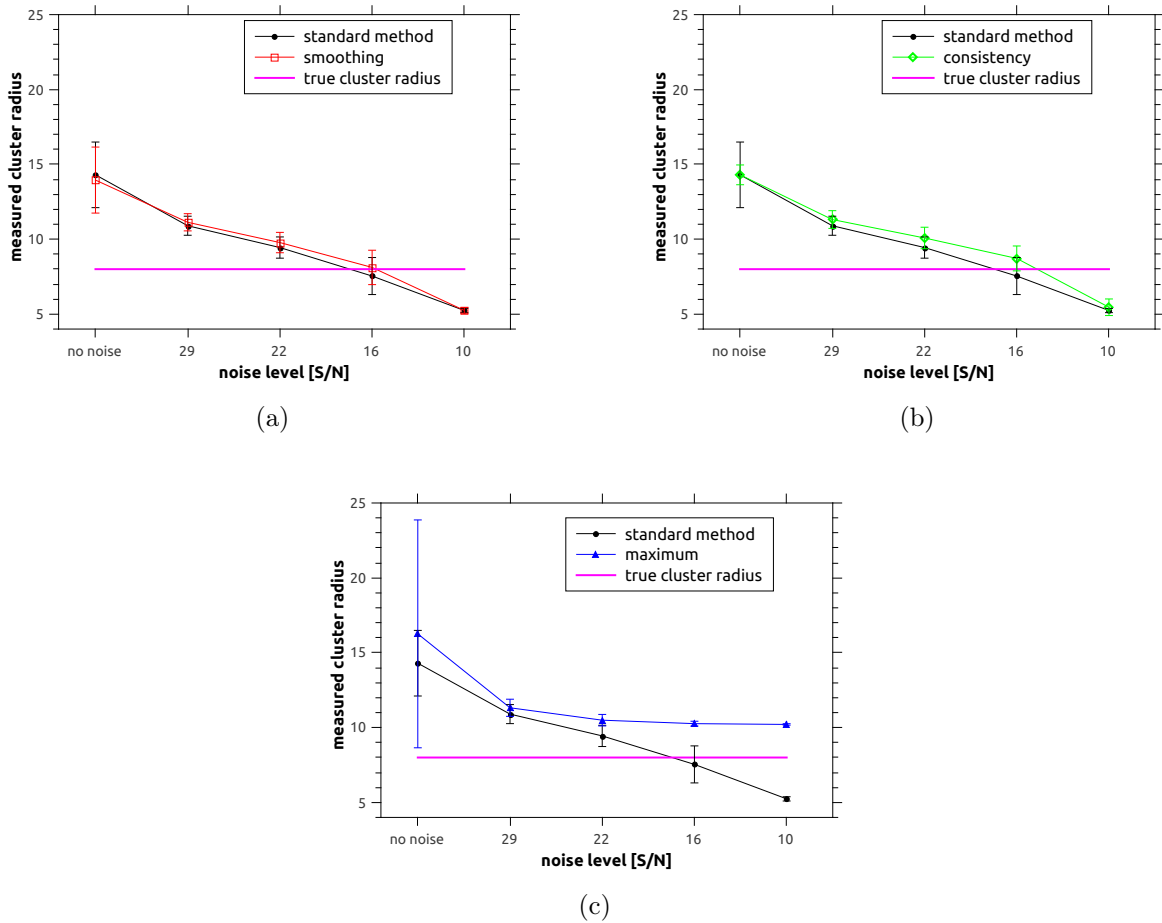
Figure 2.14: Comparison of errors for testing the noise influence on different methods of cluster radius measurement for Ripley's H function. Errors were calculated as standard deviation for different cluster radius measuring methods a) smoothing, b) consistency, c) maximum and compared to the errors of the standard method.

2.6.3 Statistical analysis of cluster radius measurement

Monte Carlo testing

Monte Carlo methods can be used for testing the accuracy and stability of a given measurement method. In this testing method the signal (here clusters) and noise are generated using the same general parameters but with different, random positions. Usually the number of datasets simulated for the testing is selected either by a "pre-run" on the data [Davidson and MacKinnon, 2000] or by adaptive choice during computation [Guo

and Peddada, 2008]. For cluster analysis testing the number of repetitions were selected using a "pre-run" method: the results distribution was evaluated and its mean and standard deviation calculated. The final number of repetitions was then set at a number for which the distribution of results did not change its properties significantly (here defined as a 1% change). An example of mean and standard deviations calculated for Monte Carlo testing of distribution with 10 clusters (8 pixel radius with S/N 29) is presented in Figure 2.15.

(a)   (b)
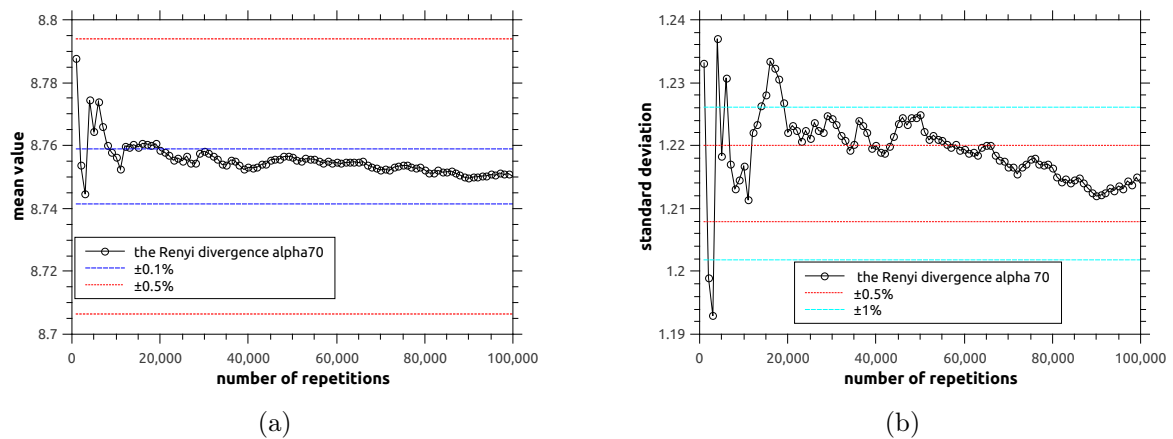
Figure 2.15: Test accounting for the error introduced by the finite number of repetitions of Monte Carlo simulations. The radius was measured with the Rényi divergence with $\alpha = 70$ for a dataset with 10 clusters (8 pixels radius with S/N 29). a) mean and b) standard deviation values measured for the results of the cluster size measurement distributions created from a limited number of repetitions. For this type of data set around 30000 repetitions provided a less than 1% change in measured value compared with the last measurement for 100000 repetitions.

The values of mean and standard deviation were strongly affected by the number of repetitions used to create the results distribution. With a higher number of repetitions, for this data set around 30,000, the mean and standard deviation have a relatively stable value. The change measured as a percentage varies by less than 0.1% for mean and less than 1% for standard deviation. The percentage change was calculated relative to the last measured values. This testing was performed to find a suitable number of Monte Carlo

repetitions, which would provide the most accurate and stable answer for a finite number of repetitions. This also enabled us to perform the Monte Carlo testing with the smallest possible number of repetitions shortening the computation time significantly. Here, only an example of the stability testing for a finite number of repetition was discussed, however, for each tested data type the same procedure was performed for both the Rényi divergence and Ripley's H function. Any change in either the number of clusters, size of clusters, or noise level would require separate assessment of the stability of the Monte Carlo testing and the error introduced to the measurement due to a finite number of repetitions.

The Monte Carlo simulations were used in this work to establish the mean, median, standard deviation, and quartiles of the radius measurements results for simulated datasets. An example of the results for datasets generated with Monte Carlo methods containing 10 clusters of 8 pixel radius and S/N 29, is shown in Figure 2.16. Monte Carlo simulations were previously used by [Kiskowski et al., 2009], to simulate datasets for performance testing of Ripley's K function.

The results distributions for the Rényi divergence and Ripley's H function have different shapes. The results distribution for the Réni divergence is asymmetric and has long tails, where the distribution for Ripley's H has a Gaussian like appearance and is symmetric. Examples of these results distributions are shown in Figure 2.17. Thus to investigate if the Ripley's function results were Gaussian distributed the skewness and kurtosis were calculated and are presented in Table 2.2 and compared with the the values of a Gaussian distribution. The kurtosis and skewness calculated for the Ripley's H function results were very close to the table values for a Gaussian distribution. The Rényi divergence skewness and kurtosis have values were completely different. This also means that the mean and median for the Rényi divergence distribution are going to have different values - the median will indicate where the peak of the distribution is located. The position of the mean will be influenced by higher values from the tails. The statistical parameters of distributions presented in Figure 2.16 are collected in Table 2.3.

Three probability density functions (pdf), Gaussian, Logistic, and t Location-scale, were fitted to the cluster radius distributions acquired with the Rényi divergence. The Logistic and t Location-scale distributions were chosen because they have heavy tails.
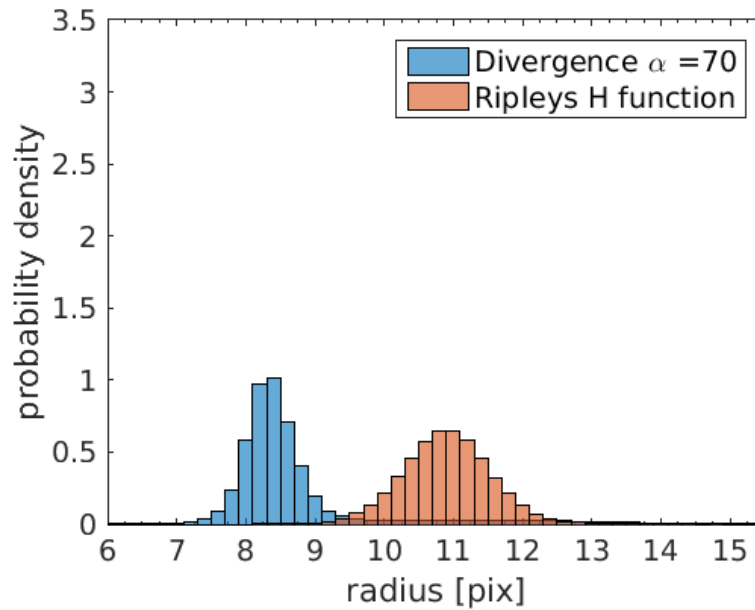
Figure 2.16: Monte Carlo statistical testing of the Rényi divergence cluster analysis method compared with Ripley's H function. Data sets were simulated with 10 round clusters (radius 8 pixels) with S/N 29 (0.5% points in the image were noise). Radius measurement using the Rényi divergence, α=70 are presented with blue bars, and Ripley's H function results are presented with orange bars.

Table 2.2: Skewness and Kurtosis calculated for the result distributions of cluster radius measurement acquired using the Rényi divergence and Ripley's H function. The results distributions were acquired for Monte Carlo data sets (10 circular clusters, radius 8 pixels, S/N 29). The kurtosis and skewness values of the results distribution acquired with Ripley's function are very close to the skewness and kurtosis values of a Gaussian distribution. The skewness and kurtosis values for the results distributions acquired with the Rényi divergence are suggesting that this result distribution is not Gaussian distributed.

	Rényi divergence $(\alpha = 70)$	Ripley's H function	Gaussian distribution
Kurtosis	11	3	1
Skewness	2.73	0.062	0
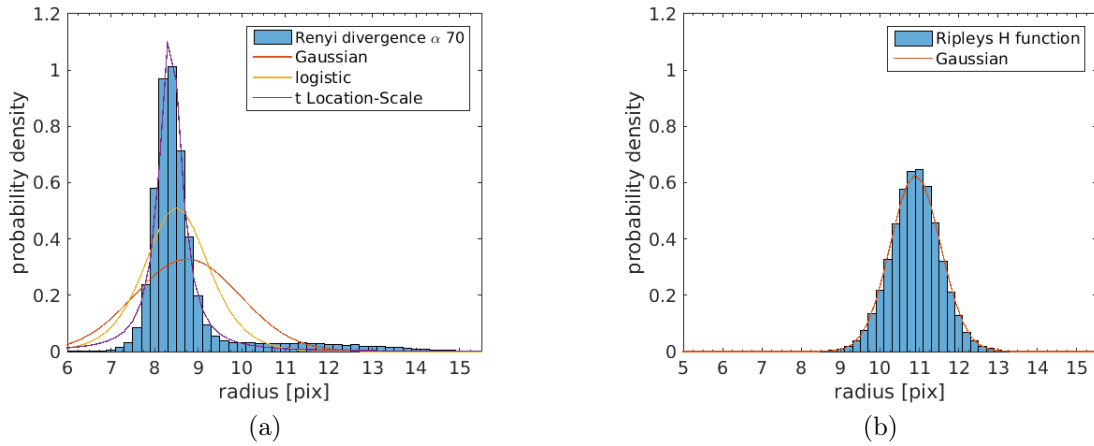
Figure 2.17: The cluster radius measurement results distributions acquired using the Rényi divergence and Ripley's H function with fitted probability distributions. a) Gaussian, Logistic, and t Location-scale distributions were fitted to the distribution of results of cluster radius measurement acquired with the Rényi divergence (for distributions parameters see Table B.1. b) Gaussian function was fitted to results distribution of cluster radius measurement acquired using Ripley's H function (see Table B.1).

Additionally, the t Location-scale distribution gives a Lorentzian (Cauchy) distribution with a very sharp peak and heavy tails. However, the pdfs fitted to the Rényi divergence results did not correspond to the data distributions optimally (see Figure 2.17a). A Gaussian pdf was fitted to radius values acquired with Ripley's H function and it provides a good estimation of the results (see Table 2.2). The equations and properties of the pdf fitted can be found in Appendix B.

Bootstrapping and noise resistance testing

Bootstrapping is a re-sampling test of accuracy of the parameter or hypothesis. It starts with the creation of new sets (or subsets) using the original data. The new sets are created by random selection of elements with replacement from the original set (each element can be selected more than once) and the number of elements in a subset is equal to the number of elements of original data. Bootstrapping can be used for testing and creating confidence intervals for different parameters [Efron and Tibshirani, 1994]. Here

bootstrapping was used to establish the mean, median, standard deviation, and quartiles of the radius measurement distributions and as a validation for Monte Carlo simulations results.

Noise resistance testing was performed to check how the answer the method provides changes with noise. Data sets were simulated to include the same cluster number and position. Background noise was added randomly with the same density to each data set. For each data realisation the cluster radius was calculated according to procedures described above in sections 2.6.1 and 2.6.2. The histograms presenting the results of the radius measurement are shown in Figure 2.20.

Bootstrapping and noise resistance testing were performed for six data sets from those generated for Monte Carlo testing. Data sets were selected to include these for which the Rényi divergence provided an overestimated, an underestimated and four accurate cluster size measurements (see Appendix A). For example the first dataset had a two groups of "super-clusters" where the small clusters were grouped together to create bigger clusters (see Figure A.1a). Similarly dataset four has uneven density of points through the clusters (high density of points inside the clusters and a smaller density at the edge) leading to underestimated measurement of the cluster size with the Rényi divergence (see Figure A.1d). The resulting distributions of the bootstrapping testing are presented in Figure 2.18.

The bootstrapping results confirm that the Rényi divergence is able to provide more precise and accurate measurement of the cluster size than the Ripley's H function. The results distributions for the Rényi divergence are more compact (see standard deviation values in Table A.1) and the mean and median of the Rényi divergence results distribution are closer to the actual cluster size of 8 pixels, than the values measured with the Ripley's H function.

Noise resistance testing was also performed to asses the noise influence on the final cluster size measurement. For this purpose data sets were generated with the same clusters (signal) and the new background noise was added to each dataset. The same signal data used for bootstrapping testing were used (see Figure A.1). Data sets were generated to acquire S/N 29 ratio (the same noise density but with different random positions). The
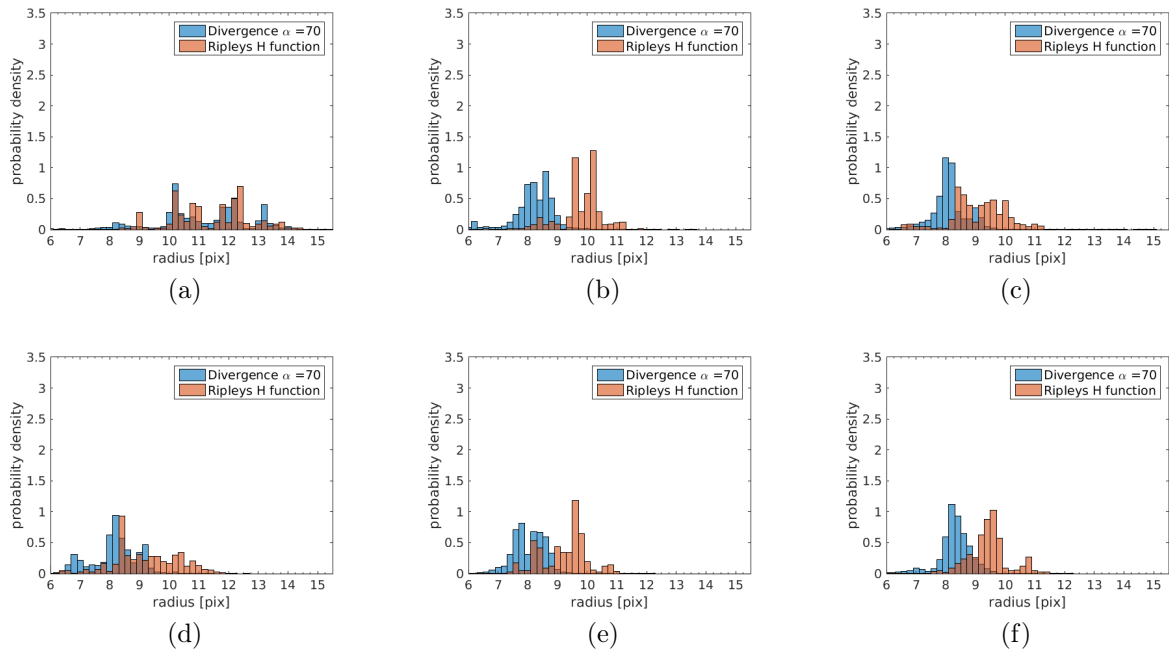
Figure 2.18: Results of bootstrapping tests of the stability of cluster radius measurements for six different datasets. The Rényi divergence results are marked with blue bars and Ripley's H function with orange. The Rényi divergence distributions provide better estimation of the true cluster size (actual cluster size 8 pixels) than the distributions for Ripley's H function. The statistical parameters of the results are collected in Table A.1. The actual cluster size is 8 pixels, S/N 29.

results distributions for the six data sets tested are presented in Figure 2.19.

The result distributions had different shapes for the Monte Carlo and bootstrapping testing. They consisted of a single (for the Rényi divergence) and multiple (for Ripley's H function) sharp peaks surrounded by relatively low tails. The appearance of sharp peaks means that the method provided the same cluster radius measurement for different realisations of background noise. The Ripley's H function appears to be more affected by changes in noise and thus is likely to provide a less accurate cluster radius measurement for the same noise level (see standard deviation in Table A.2). For this testing method the Rényi divergence provides more accurate measurement of the cluster size than the Ripley's H function (see mean and median values in Table A.2).

Bootstrapping and noise resistance result distributions were summed to create bulk
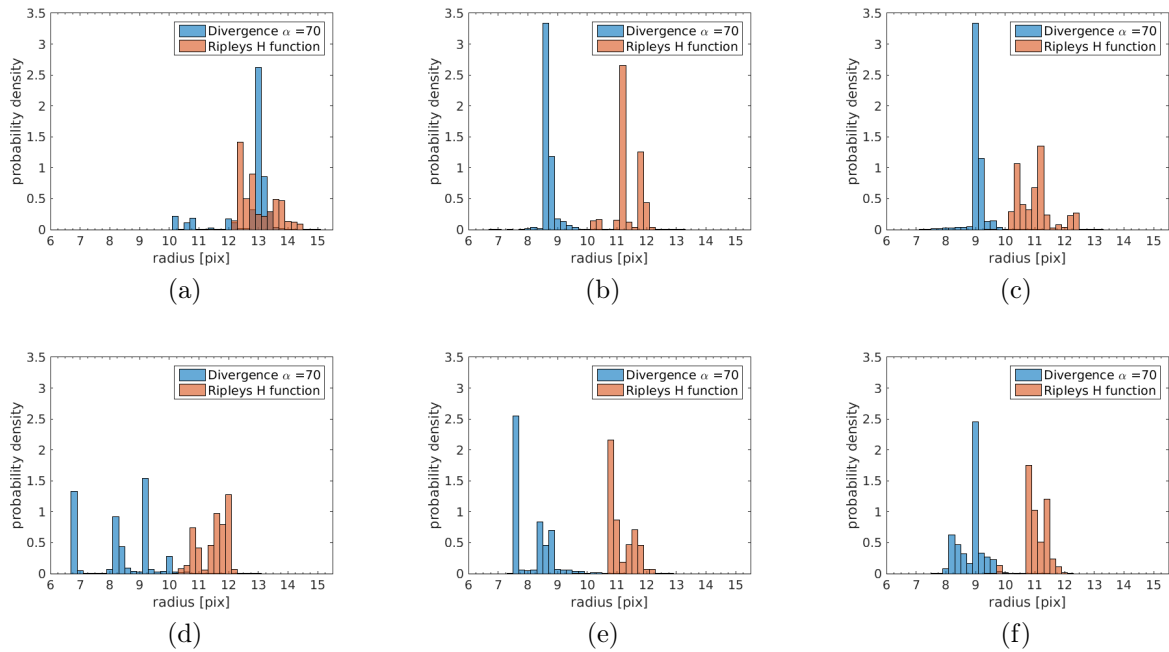
Figure 2.19: Results distributions for noise resistance tests of the stability of cluster radius measurements for six different datasets. The Rényi divergence distributions are marked with blue and Ripley's H function with orange. The Rényi divergence results distributions generally have a single sharp peak and comparably weak tails. Ripley's H function results have a broader distributions.

results distributions for the six analysed data sets. To compare the results of these tests with the Monte Carlo results bulk distributions were created. These distributions were built by adding together distributions acquired with bootstrapping and separately in noise resistance. These bulk distributions were then compared with the results distribution acquired for Monte Carlo testing (see Figure 2.20). The parameters of the final distributions (mean, median, standard deviation, and quartiles) for each test performed are collected in Table 2.3.

The three tests performed (Monte Carlo, bootstrapping, and noise resistance) confirmed that the Rényi divergence provides more accurate and precise cluster radius measurements than Ripley's H function. For each of the statistical tests performed the value of mean and median was closer to the real cluster radius when measured with the Rényi divergence than with the Ripley's H function. Also, interquartile distance (measured as
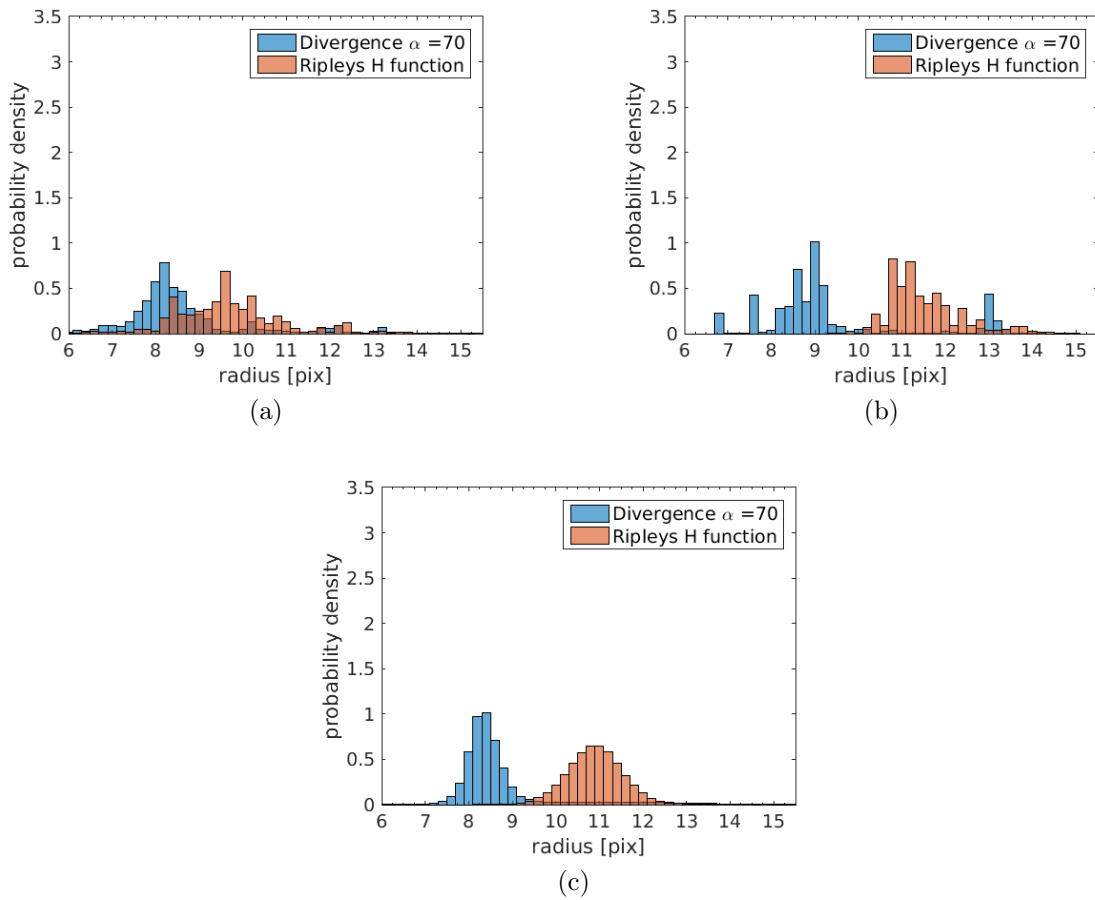
Figure 2.20: Comparison of results distributions for stability testing (Monte Carlo, boot-strapping, noise resistance). Bulk distributions of six data sets were created by adding results distributions for each data set for a) bootstrapping, b) noise resistance, c) Monte Carlo results distribution. These cumulative distributions are broader than the Monte Carlo distribution results, however, their mean and median were very similar these parameter values for Monte Carlo testing. Rényi divergence results are in blue bars and Ripley's H in orange.

distance between the 1st and the 3rd quartile) values were smaller for the distributions measured with the Rényi divergence. Standard deviations calculated for distributions in bootstrapping and noise resistance were bigger for the Rényi divergence. However, it should be noted that the distributions for the Rényi divergence are asymmetric and the right tail affects the standard deviation values.

Table 2.3: Statistical parameters of results distributions for stability testing with bootstrapping, Monte Carlo and noise resistance. Mean, median, standard deviation and quartiles (the 1st and the 3rd respectively) of the distributions of cluster radius measurement with the Rényi divergence and Ripley's H function on simulated data sets with 10 clusters, with S/N 29 (noise level 0.5%) generated using bootstrapping, Monte Carlo and noise resistance testing.

		Monte Carlo	Bootstrapping	Noise resistance
Rényi div.	mean	7.80	8.63	9.30
	median	7.80	8.20	9.00
	st. dev	0.61	1.45	1.69
	quartiles	7.60, 8.20	8.00, 8.80	8.40, 9.20
Ripley's H	mean	10.40	9.64	11.51
	median	10.40	9.60	11.20
	st. dev	0.46	1.33	0.83
	quartiles	10.40, 11.00	8.80, 10.20	11.00, 12.00

The cluster measurement provided by the Rényi divergence is strongly dependent on the quality of the signal. For example, cluster radius measured for data sets with overlapping clusters is bigger and this leads to overestimation of cluster size. When testing the influence of noise on the measurement in noise resistance testing the Rényi divergence provides a stable radius measurement. This means that the quality of signal is the limiting factor for the precise and accurate cluster radius measurement with the Rényi divergence. Ripley's H function is less affected by changes in signal and in noise providing a broader and thus less specific results distributions during testing.

Simulated datasets with only background noise

As a validation of the quantitative accuracy the Rényi divergence was tested on data sets with only background noise. The data sets for this testing were generated with approximately 1800 points. The number of points to create these data was selected to correspond to an average number of points (clusters + noise) in data sets with 10 clusters (8 pixels radius) with S/N 29. The average density of points for this data set was

0.006 points per pixel. An example of a data set created for this testing is presented in Figure 2.21.
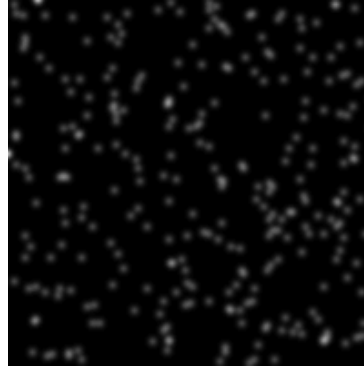
Figure 2.21: An example of a data set containing only background noise. Because noise was generated randomly small aggregations of points were present. For improved visibility the image was blurred.

The data generated for this testing had a small degree of clustering (as seen in Figure 2.21) which means that the final result of the testing will depend on the selected value of the increment of radius in calculations. Two values of increment were tested, 0.2 and 1 (sampling was set with increment of the radius for which the Rényi divergence was calculated - meaning that for this data set radius increased by 1 or 0.2 pixel). For an increment 1 there were no plateaus present in the Rényi divergence function and no maximum in Ripley's H function. Thus, there was no cluster radius measured.

Both the Rényi divergence calculated with $\alpha = 10$ and $\alpha = 70$ for sampling 0.2 provided a cluster identification and cluster radius measurement. The distribution for smaller α values provided a range of possible cluster radius measurements (see Figure 2.22 for $\alpha = 10$). The distribution was very broad (standard deviation 4.72) and without a clear peak maximum (mean 9.13). For higher alpha the distribution was less broad (standard deviation 0.79) and positioned around smaller values of radius (mean 1.42). Additionally for higher α values the distribution had a bimodal distribution originating from small groups of points in random distribution (see Figure 2.21 and 2.22). The results provided by the Rényi divergence were completely different for data sets with only background
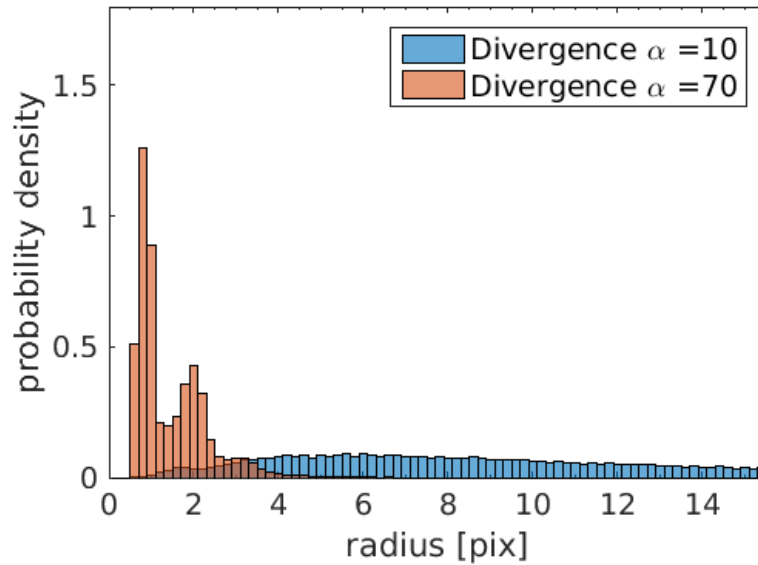
Figure 2.22: Distribution of cluster radius measurements with the Rényi divergence with $\alpha = 10$ and $\alpha = 70$ for datasets with only background noise.

noise from the data with clusters present. The Ripley's H function measured sub-pixel cluster sizes for sampling 0.2.

The sampling 0.2 was selected for noise testing because this sampling rate was used for cluster radius measurement of clustered data sets (see section 2.6.4). The reason clustered data required a higher sampling rate was caused by a inhomogeneous data density in these datasets. The areas with clusters had a magnitude higher point density than the areas with noise or background noise data. Thus to accurately detect cluster radius a higher sampling was used.

2.6.4 Cluster radius measurement and α adjustment for different sample properties

The properties of localisation microscopy data sets can vary a great deal. Usually a similar noise level is present for similar samples and with an identical preparation and imaging method. However, changing the sample preparation or using, for example, a new batch of dye or antibody can influence the level of background noise. Thus, noise resistance is

a requirement for a good cluster size measurement method for localisation microscopy. Here, the accuracy and precision of the Rényi divergence and Ripley's H function was compared on simulated Monte Carlo data sets with the same number of clusters with eight different noise levels (no noise, and S/N 29, 22, 16, 10, 7, 6, 3). For each noise level Monte Carlo simulations were performed, by creating the same number of data sets. Cluster radius measurements with the Rényi divergence and Ripley's H function provided distributions of results similarly to already discussed distributions in section 2.6.3.

The Rényi divergence behaviour is strongly dependent on the value of α. Generally, the higher the noise level (or smaller the signal-to-noise ratio) the higher the α value needs to be to precisely measure the cluster size. The easiest way to select the optimal α value is to calculate the Rényi divergence for a rage of α values and select an α value for which the measured cluster radius value and confidence levels start to have similar or the same value. The cluster radius calculation was performed with the Rényi divergence for α values in range of 10-120. This was performed on simulated datasets containing ten round clusters with a uniform point distribution and a sharp edge (and a radius of either 8, see Figure 2.24, or 16 pixels, see Figure 2.26) and Gaussian clusters (points inside the clusters had a Gaussian distribution, with $\sigma = 8$, see Figure 2.28). A number of different noise levels was added to those datasets to create images similar to those observed for localisation microscopy. Depending on the properties of the cluster (size, type) and background noise level, a different value of α was needed to accurately measure cluster size. For example, the cluster radius was measured with increasingly higher precision for α=10-30 (see Figure 2.23) and for α values higher than 30 there is no change in measurement precision.

The distribution for measurements with the Rényi divergence is asymmetric and the median provides a better measurement for the maximum of the distribution. Similarly, the 1st and 2rd quartiles carry more information about the distribution behaviour. However, because the distribution of Ripley's H function results are symmetrical and other studies cited mostly the mean of the distribution (for example, [Kiskowski et al., 2009]) the mean values for both methods were also calculated. Here, the mean and median of the Rényi divergence and Ripley's H function were compared to provide a broader understanding of

the properties and shape of their results distribution.

As discussed above the α value determines the shape of the Rényi divergence function. With higher α the influence of noise is minimised and cluster size is measured more accurately. To illustrate that behaviour the cluster radius was measured using a range of α values for datasets containing 10 round uniform clusters (radius 8 pixels and S/N 29). The measured cluster size and α values used for the measurements are presented in Figure 2.23.

(a)	(b)
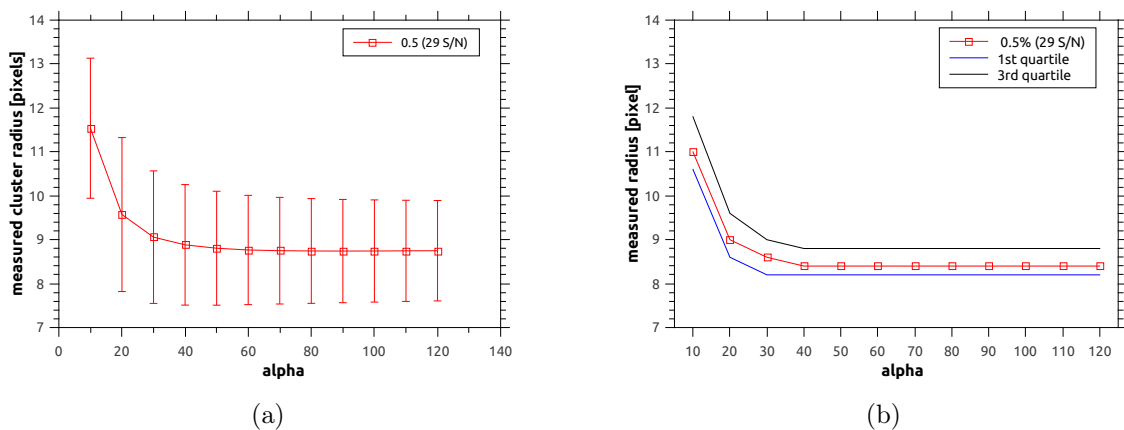
Figure 2.23: Cluster radius measurement using different values of α on datasets with S/N 29. Datasets were simulated using Monte Carlo methods (see Section 2.6.3) with 10 clusters (8 pixels radius and S/N 29). Here, for comparison, the average radius is presented as a) mean with standard deviation (marked in red) and b) median (red) with confidence levels established with the 1st and 3rd quartiles marked with blue and black lines respectively. For this type of dataset a stable and accurate cluster radius measurement was provided for α higher than 30.

The same investigation of the minimal α value for the cluster radius measurement was performed for a Monte Carlo generated series of datasets. Each series of data had added background noise (starting from no noise to S/N 3). An average cluster radius measured for these noise levels using different α values is presented in Figure 2.24.

With increasing noise the minimal α value needed to provide an accurate answer increases. Generally, the higher the noise the higher the α value needed to be. This rule holds well up to the point when noise has a very similar density to the density of points
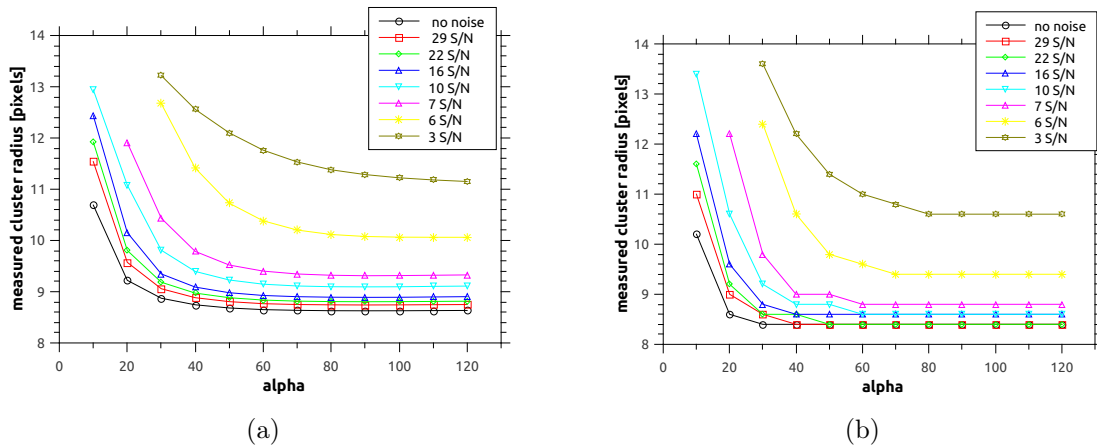
Figure 2.24: Cluster radius measurement using different values of α on data sets with increasing noise levels. Results of cluster radius measurements with the Rényi divergence for Monte Carlo simulated datasets with 10 clusters (radius of 8 pixels) with various noise levels. The results distributions were asymmetric thus a) mean and b) median were calculated for the results distribution. The cluster radius was calculated for measurements with different α values for data sets with increasing levels of noise (noise levels were marked with: no noise – black, S/N 29 (0.5%) – red, S/N 22 (0.74%) – green, S/N 16 (1.0%) – blue, S/N 10 (1.5%) – turquoise, S/N 7 (2.0%) – magenta, S/N 6 (3%) – yellow, and S/N 3 (4%) – dark yellow). The Rényi divergence for this type of data set provides stable radius measurement for noise levels smaller than 3.0% (S/N 6) and α values higher than 50.

in clusters (noise level S/N 3). For this type of data sets α value needs to be higher than 50 to provide accurate and precise measurements of cluster radius.

Results of cluster size measurements with the Rényi divergence were compared with cluster sizes measured with Ripley's H function using Monte Carlo simulations. The Rényi divergence provided similar values of cluster radius measurement for noise levels up to S/N 6, where Ripley's H function provided a cluster radius measurement for noise levels S/N 16 (1.0%) and higher. Cluster radius measurements with the Rényi divergence for $\alpha = 70$ and Ripley's H function for simulated data sets with different noise levels are shown in Figure 2.25. Comparison of the Rényi divergence with Ripley's function results suggested that the Rényi divergence provides both more accurate and more precise measurement. Error bars shown in Figure 2.25 were calculated as standard deviations

for the mean and interquartile distance for the median. Monte Carlo simulations were performed according to the protocol presented in section 2.6.3.

(a)   (b)
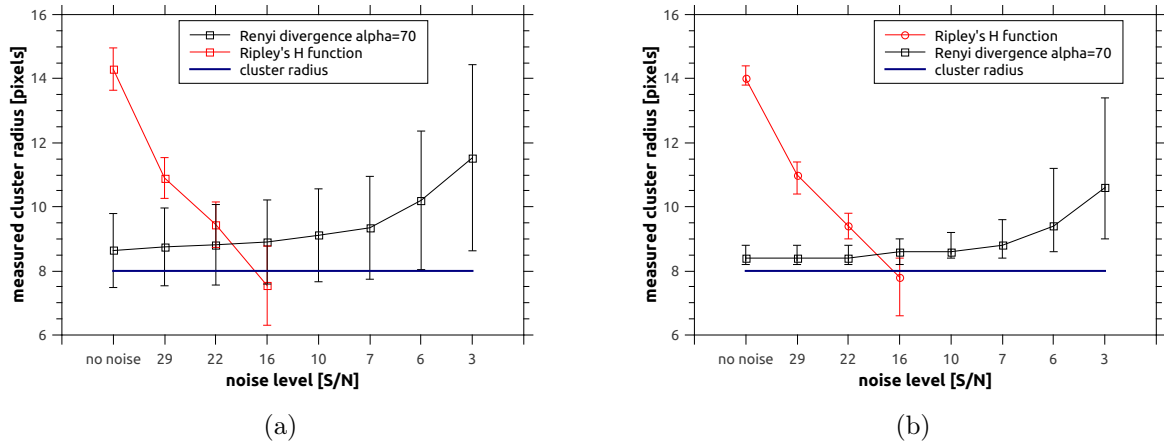
Figure 2.25: Comparison of cluster radius values measured with the Rényi divergence and Ripley's H function for simulated datasets with 10 clusters (8 pixel radius) and increasing noise levels. The average cluster radius was measured as a) mean (error bars are standard deviation) and b) median (error bars interquartile distance). The cluster radius size was measured with the Rényi divergence (marked with black) and Ripley's H function (red). Ripley's H function radius measurement was only possible for noise levels smaller than 1.0% or higher than S/N 16. The Rényi divergence provided radius measurement up to 3.0% noise (S/N 6) and provided both more precise and accurate measurements.

Similarly, the performance of the Rényi divergence and Ripley's H function was assessed for data with bigger clusters sizes (10 clusters, radius 16 pixels, with a uniform distribution of points in the cluster with a sharp cut off). The cluster radius was measured for data sets simulated with Monte Carlo methods (see section 2.6.3) and α values in the range 10-120. The cluster radius was measured as the mean (error measured as standard deviation) and the median (error measured as interquartile distance) was used to provide a more broad understanding of the distribution shape (see Figure 2.26) (the same methodology was used as discussed above for data sets with 8 pixels radius).
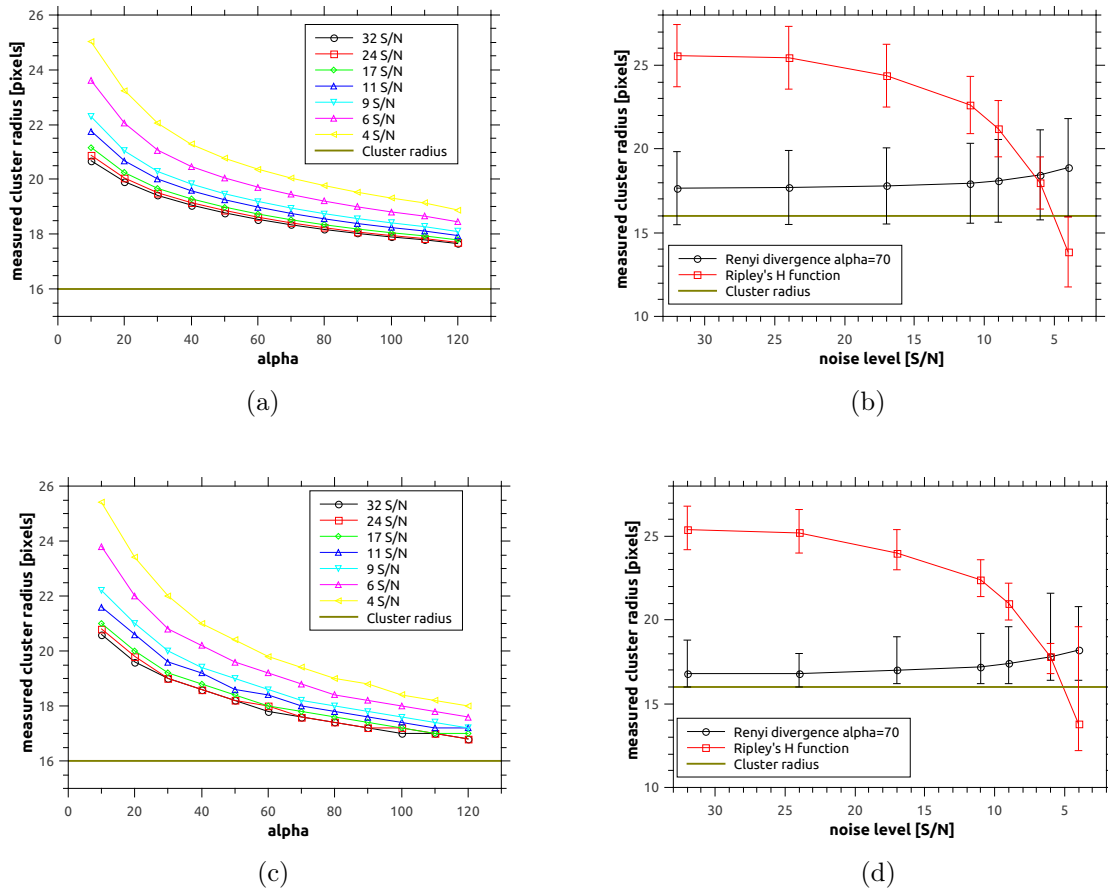
Figure 2.26: Cluster radius measurements with the Rényi divergence and Ripley's H function for simulated clusters with 16 pixels radius with different noise levels. The results distribution is asymmetric for the Rényi divergence, thus the average radius was presented as a) & b) mean, and c) & d) median. Error bars are standard deviation for mean and interquartile distance for median. The cluster size measured with the Rényi divergence was closer to the real cluster size than the cluster radius measured with Ripley's H function.

The results of cluster radius measurements on data with 10 clusters with 16 pixel radius also suggest that the Rényi divergence $\alpha = 70$ provided a more accurate and precise measurement of cluster size. The mean and median of the results distribution of cluster radius measurements were closer to the real cluster radius for the Rényi divergence ($\alpha = 70$) (see Figure 2.26b and d). While comparing the noise resistance the Rényi divergence also performs better. The divergence measure provided a stable cluster radius

measurement for different noise levels. The change in the value of cluster radius measured was less than 10%, compared with 45% decrease in value measured with Ripley's H function. It should be also noted that Ripley's H for the S/N lower than 6 measured cluster size which was smaller than the actual cluster radius. This indicated that Ripley's function cannot separate clusters (or signal) from the background noise for this noise level.

Analysis of clusters with sharp defined edges provides some insight into the performance of the Rényi divergence and Ripley's H function. Unsurprisingly, the behaviour of both methods was affected by the data quality and properties, for example size of clusters. As discussed before the accuracy of the cluster radius measured using the Rényi divergence is higher than when measured with Ripley's H function based on analysis of clusters with radius 8 and 16 pixels. For both of the analysed data set types the Rényi divergence provides more accurate measurement of cluster size compared with Ripley's H function which overestimates the cluster size.

When considering the noise resistance the behaviour of the two methods was different. The cluster size measured with the Rényi divergence increased slightly for higher noise levels (see Figures 2.25 and 2.26). This was caused mostly by a very small difference between the density of background and clusters, leading to overestimation of the cluster size. In contrast, the value of cluster radius measured with Ripley's H function decreased to the point of detecting clusters smaller than the actual cluster size, as the noise level increased (see Figures 2.25 and 2.26). This was caused by lack of any density filtering in Ripley's function. For larger clusters both methods were able to provide cluster radius measurements even for higher noise levels. With smaller cluster sizes and higher noise levels the measurement of size was less accurate for the Rényi divergence when Ripley's H function measured cluster size smaller than the actual cluster size.

The clusters discussed up to this point were simulated with a sharp cut-off at the edge. However, the clusters in biological samples have been reported to have diffuse and less pronounced edges or non-circular, elongated shape. A number of studies identified clusters present in the localisation microscopy images to have shape of a Gaussian (for example [Owen et al., 2013; Sengupta et al., 2011; Williamson et al., 2011]) or Cauchy-Lorentz [Pertsinidis et al., 2013]. To test the performance of the Rényi divergence and

Ripley's H function against these types of clusters, we simulated data sets with 10 symmet-rical Gaussian clusters with $\sigma = 8$ pixels. To account for varying conditions in localisation microscopy background noise was added with increasing density. An example of a dataset with Gaussian clusters and S/N 29 is shown in Figure 2.27.

Figure 2.27: An example of a dataset with Gaussian clusters. Datasets contains 10 Gaussian ($\sigma = 8$ pixels) shaped clusters with S/N 29. The radius of Gaussian clusters was modelled as half of the Gaussian FWHM and for this data it was equal to 9.4 pixels. For improved visibility the image was blurred.

A symmetrical Gaussian function is defined by two parameters, mean and standard deviation (σ), and has an infinite area. The size of a Gaussian it is usually estimated using standard deviation (σ) or full width at half maximum (FWHM), given as $FWHM = 2\sqrt{2\ln 2}\sigma$. For clusters simulated for this testing the FWHM was equal to 9.4 pixels.

To measure the cluster radius a bigger increment of radius for calculation was used in both the Rényi divergence and Ripley's H function. Monte Carlo simulations were used to simulate data sets with the same cluster properties and different noise levels. Here, the results distributions were characterised using the mean, median and standard deviation values (see Figure 2.28).

The comparison of accuracy of the cluster radius measurements for the Rényi divergence and Ripley's H function did not show a big difference in performance of these two
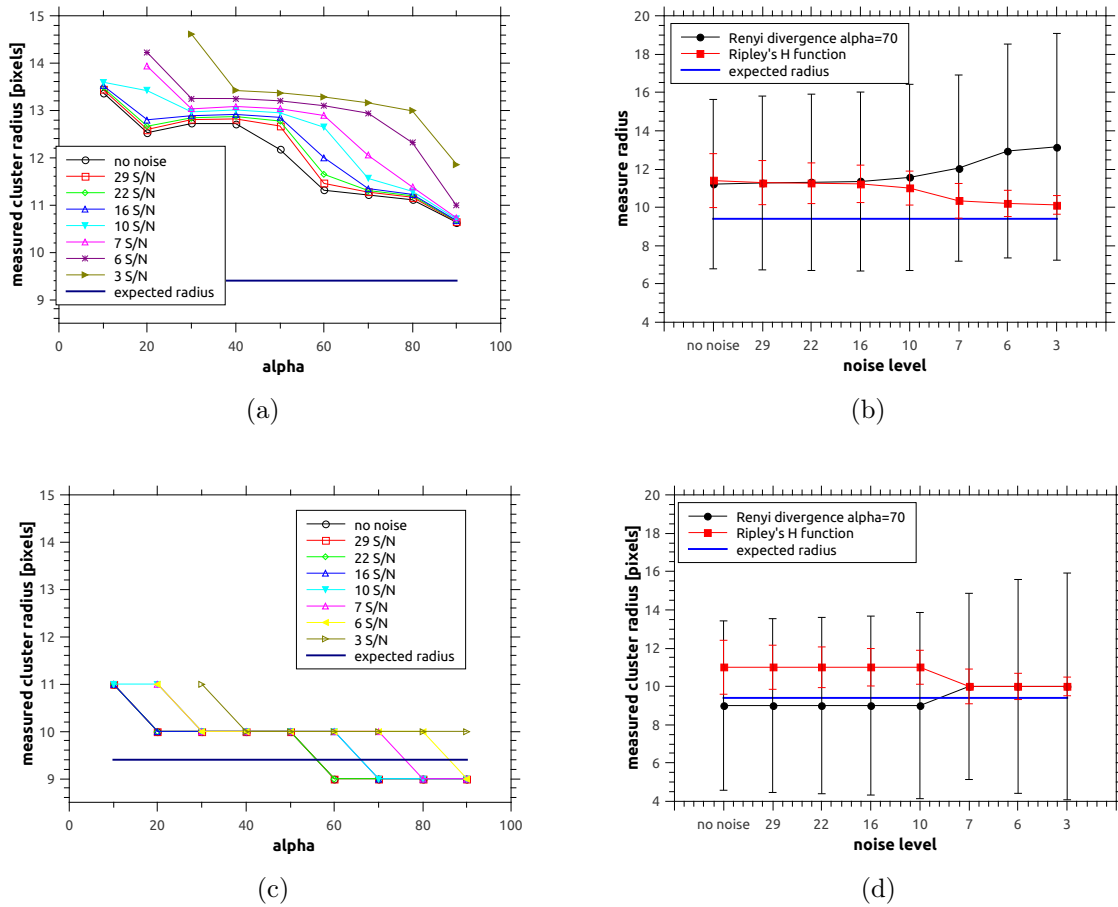
Figure 2.28: Results of the cluster radius measurement for the Gaussian clusters ($\sigma = 8$) with different noise levels using the Rényi divergence and Ripley's H function. a) The mean and c) median cluster radius measured using the Rényi divergence with different α values. Different noise levels are marked with: no noise – black, S/N 29 – red, S/N 22 – green, S/N 16 – blue, S/N 10 –turquoise, S/N 7 – magenta, S/N 6 – purple, and S/N 3 with dark yellow. Comparison of b) mean and d) median cluster radius measured using the Rényi divergence (marked with black) and Ripley's H function (marked in red). The expected cluster radius is marked with a navy line. Error bars are the standard deviations.

methods. Both of the methods slightly overestimated the expected cluster size of 9.4 pixels. For the higher noise levels the measured cluster size started to be affected. Similarly to already observed results for sharp-cut off clusters, the Rényi divergence measured cluster radius to be bigger and Ripley's H smaller than those measured with both methods for

lower noise levels. The main difference in the results presented here, for Gaussian clusters, is that the Rényi divergence measurements were less precise than the Ripley's H function, which was different from results discussed above for sharp-cut off clusters. This effect can originate from the cluster morphology itself, it is harder for the Rényi divergence to identify the point density change between the cluster and the noise especially for higher noise levels. It should be also noted that because of higher increment of radius used for the Gaussian clusters the results provided were less precise. To improve the quality of the results the method of cluster radius measurement will have to be changed. The current method of radius detection designed for high sampling rates, simply detects plateaus in the function. For low sampling rates a two step approach can be used, the plateau can be detected in low sampling data, and used to probe the data with higher sampling density to detect the starting point of the plateau in the function. These higher precision calculations are currently in progress.

2.6.5   Analysis of localisation microscopy images of DNA origami samples

DNA origami are two or three dimensional structures formed by artificial folding of DNA strands. These structures were first reported by [Rothemund, 2006] and various shapes can be achieved. DNA origami structures (supplied by GATTAquant) were DNA origami plates with known sizes and an approximately constant number of Alexa Fluor 647 molecules attached (around 20 molecules per plate, see a model of the plate in Figure 2.29). The sizes of the ordered DNA origami plates were 30x30, 60x60 and 60x90 nm. Knowledge about the actual size of imaged structure was particularly necessary to be able to compare the accuracy and precision of the Rényi divergence and Ripley's H function.
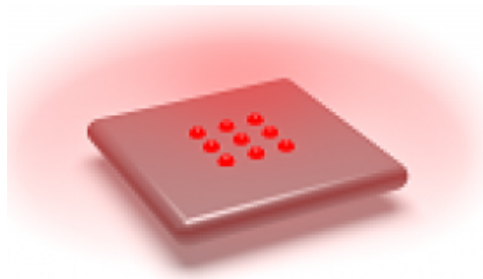
Figure 2.29: An example of DNA origami plate image supplied by GATTAquant. Here only nine fluorescent molecules are attached to the plate. For the results shown in this section DNA origami plates used to gather clustering information were labelled uniformly with approximately twenty Alexa Fluor 647 molecules. Image adapted from: [GATTAquant, 2016].

The DNA plates were delivered in liquid form and samples for localisation microscopy were made according to protocol supplied by GATTAquant. This protocol immobilised the DNA origami on BSA-biotin-neutravidin surface. The first dish was washed three times with PBS. Then the dishes were incubated with 200 μl of BSA-biotin solution (0.5 $\frac{mg}{ml}$ in PBS) for 5 minutes. After incubation, the BSA-biotin solution was removed and dish was washed three times with PBS. Next, the dish was incubated with the neutravidin solution (0.5 $\frac{mg}{ml}$ in PBS). After 5 minutes the neutravidin solution was removed and the dish was

washed three times with PBS. Then, DNA origami was diluted between 10 and 100 times (0.5 μl of DNA origami solution + 5 to 50 ml of PBS) with PBS. The DNA plates (in liquid) were placed in the dish and incubated for 5 minutes. The optimal dilution factor was selected using guidelines provided in the sample preparation protocol, so the average density of DNA plates on the sample surface has been $\sim 1/\mu m^2$. Lastly, the dish was washed with PBS three times.

Imaging was performed at the Nikon N-STORM system. The exposure was set to 10 ms, and around 10000 images per imaged region were collected. The 647 nm laser power was adjusted to provide the same number of single molecules appearances in each frame (as far as possible). The 405 nm laser was used towards the end of the imaging to shorten the subpopulation of molecules in a long lived dark state in Alexa Fluor molecules. Post-processing and super-resolution image reconstruction was then performed in quickPALM [Henriques et al., 2010].

Simulated DNA origami data sets

DNA origami plates are rectangular and the Rényi divergence and Ripley's H function provide optimal results for circularly shaped clusters. The DNA origami plates ordered from GATTAquant had rectangular or square shapes (30x30, 60x60, 60x90 nm). Both the Rényi divergence and Ripley's H function use a circle to measure the local clustering patterns, which means that the cluster size measured by these methods will be a radius of a circle. Thus, the measured cluster radius will have a relation to the size of the structure itself but it will not be its actual size. The potential values of the measured radius can be estimated: the measured cluster will not be smaller than a circle fitting inside the rectangle and not bigger than the circumcircle (see Figure 2.30). The lower limit, or minimal expected radius was underestimating the cluster size, as a big part of the cluster with high density of molecules is not included into a cluster. The minimal radius was given as half of the length of the shorter side of the rectangle. The higher limit and maximal expected radius for the measured cluster radius was a circumcircle (see Figure 2.30) with radius given as half of the diagonal of the rectangle: $r = \frac{\sqrt{a^2+b^2}}{2}$.

To account for the difference between the measured cluster radius and actual plate

size, the rectangular plates were simulated. The rectangular DNA origami were created to measure 16x24 pixels, this preserved the 2:3 length and width ratio of real 60x90 nm plates. The square plates size of 16x16 pixels corresponded to 60x60 nm plates. Each data set had 10 clusters with S/N 29.
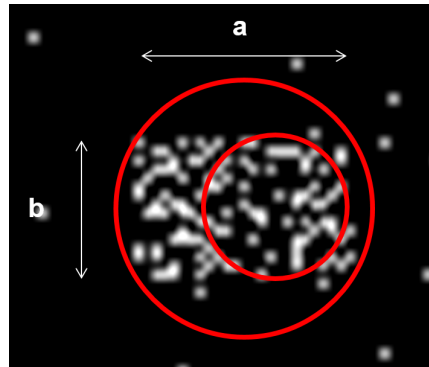
Figure 2.30: Simulated DNA plate and cluster radius estimation (plate is a fragment of the image Figure 2.32). Circle passing through DNA origami corners, radius of that circle can be calculated as half of the diagonal and inner-circle radius is equal to half of the length of the shorter side of the plate.

The simulations of the square and rectangular DNA origami allowed us to account for the difference in the geometry of the sample. The square DNA origami plate area size is more close to the circumcircle area than the rectangular plate (see Figure 2.30). Thus it was expected that the percentage difference between the value of the measured cluster radius and the maximal expected radius can be different for the two considered shapes of plates. This hypothesis was checked by simulating square and rectangular plates and comparing the measured cluster size with the maximal expected radius. The simulated DNA origami data sets were analysed using the Rényi divergence and Ripley's H function.

The comparison of the cluster radius measured by the Rényi divergence with the maximal expected radius suggested that there is a small difference in measured cluster radius for the two plate shapes. Figure 2.31 a and b present the measured cluster radius (as mean and median respectively). The absolute difference between the measured radius and maximal expected radius is bigger for the rectangular plate. However, the percentage difference in measurements for both plates are similar (mean and median measurements
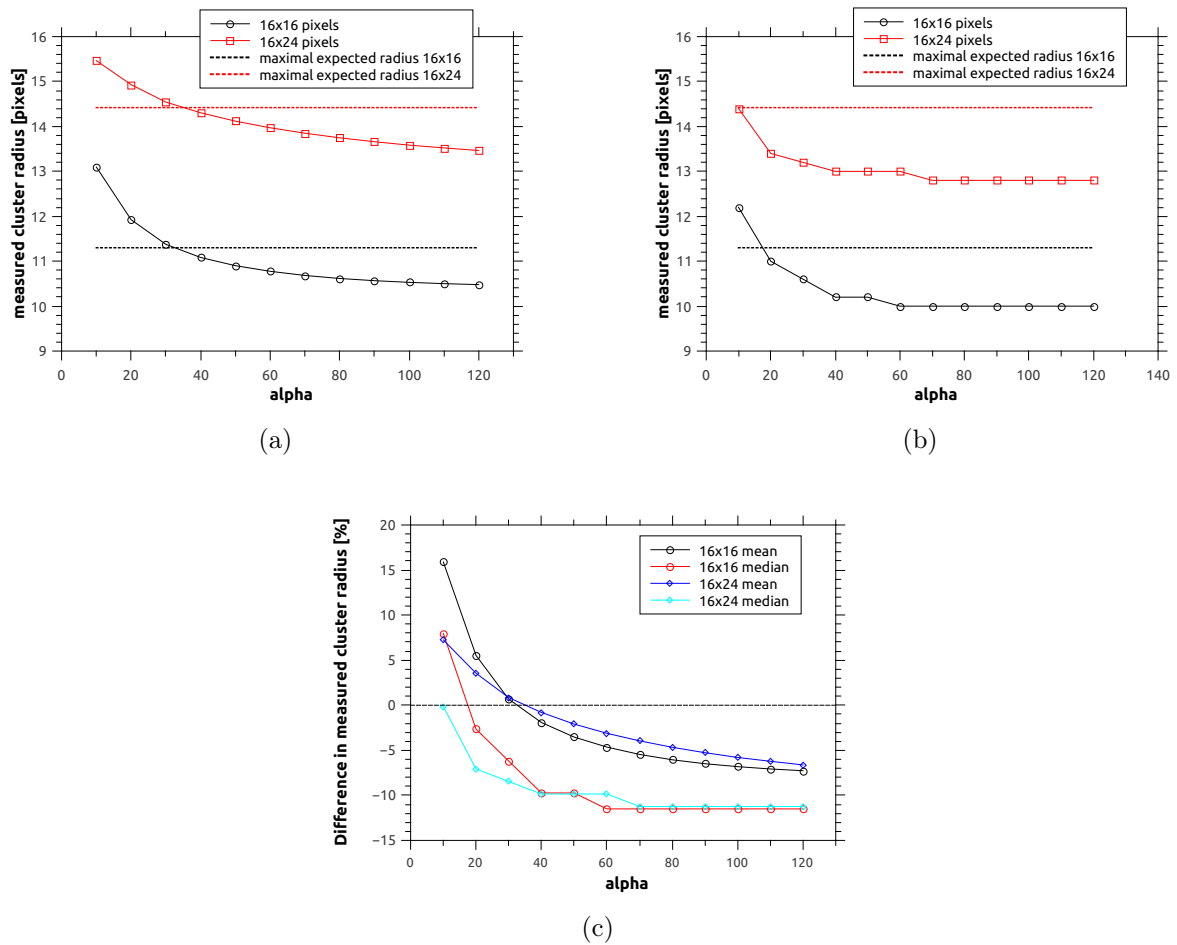
Figure 2.31: The average cluster radius measured for the two analysed shapes of simulated DNA origami plates. The average radius measured with the Rényi divergence for a range of α values with a) mean and b) median for data sets with square and rectangular clusters. c) Percentage difference between the measured cluster size and the maximal expected radius for both plate shapes and measurement with mean and median.

were analysed separately, see Figure 2.31c). The percentage difference for measurements with the Rényi divergence and Ripley's H function was presented in Table A.3.

The Rényi divergence measure of cluster size is more closely related to the actual size of the objects than the Ripley's H function. The absolute value of measured radius and the percentage difference indicate that Ripley's H function overestimated the maximal expected radius by 30%.

Additionally, the noise resistance was evaluated for rectangular DNA plates in order

to provide a guideline for localisation microscopy imaging. For this testing datasets were simulated with 10 clusters, 16x24 pixel plates with S/N 29, 16 and 7. An example of simulated DNA origami data set is shown in Figure 2.32. For the simulated plates the minimal expected radius was 8 pixels and the maximal radius 14.4 pixels.
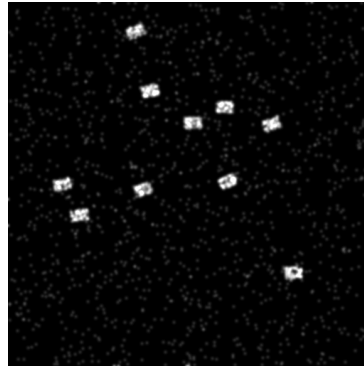
Figure 2.32: Simulated DNA origami plate dataset with 10 clusters and S/N 29. The length of the sizes of the simulated plates corresponded to the ratio of the width and length of the actual plates (DNA origami plates dimensions were 60x90 nm and simulated 16x24 pixels).

Results of the cluster radius measurements suggest that the Rényi divergence provides a more accurate cluster radius measurement than the Ripley's H function (see Figure 2.33). The measured cluster radius distribution for the Rényi divergence has an asymmetrical shape (see Section 2.6.3), thus the mean and median were used to provide information about the results. The cluster radius measured with the Rényi divergence was closer to the cluster size set by the limits discussed above and the measured cluster size was smaller than the radius of the circle passing through the rectangle corners. The Ripley's H function overestimated the cluster size and the measured cluster size was bigger than the cluster itself. The cluster size measured with Ripley's H function for higher noise levels is closer to the estimated upper limit of circumcircle – this is caused by the noise influence.
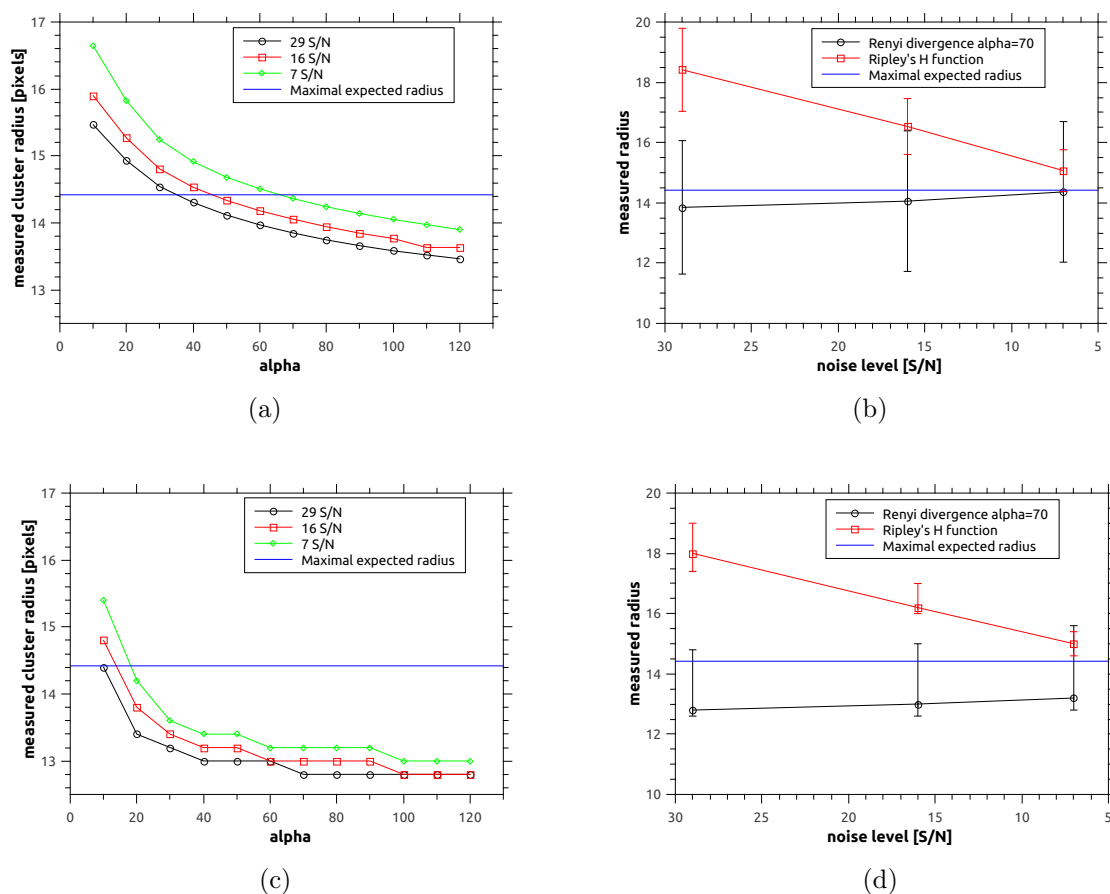
Figure 2.33: Cluster radius measurements with the Rényi divergence and Ripley's H function for simulated DNA origami datasets with increasing noise levels. Radius measurement for simulations of DNA origami data sets a & b) radius measured as mean, and c & d) median. Error bars are standard deviation for mean and 1st and 3rd quartiles for median. The cluster size provided by the Rényi divergence measured as mean and median was closer to the real cluster size, than one provided by Ripley's H function.

Localisation microscopy images of DNA origami

The DNA origami localisation microscopy data sets of 30x30, 60x60, and 60x90 nm plates samples were prepared, imaged and analysed. An example of localisation microscopy image of 60x90 nm plates is shown in Figure 2.34. There was some variation in the number of attached dye molecules per plate. This results in a variation in brightness and shape of the images of the plates. It should be also mentioned that the three sizes of

plates mentioned here only correspond to the area where dyes were attached. The size of the folded DNA was the same for the three types and equal to 60x90 nm.
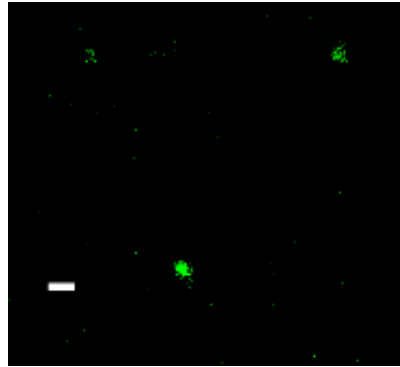
Figure 2.34: Localisation microscopy image of three DNA origami rectangular plates, size 60x90 nm. Different numbers of Alexa Fluor 647 molecules attached to each plate resulting in small variations in brightness of the clusters and accounts for their non-square shape. Scale bar 300 nm.

The DNA origami data were analysed using the Rényi divergence and Ripley's H function to measure the average cluster size for these structures. As discussed above, for the simulated data the cluster radius is not going to be directly equal to the plate dimension, but will be limited by the maximal and minimal expected radius (circumcircle and circle fitting inside the plate). The maximal and minimal expected radius values for imaged plates are collected in Table 2.4. The measured cluster radius for three sizes of plates analysed are presented in Figures 2.35 and 2.36.

The Rényi divergence provided very similar values of the cluster size for all the three types of plates. This result was partially caused by the quality variation in the samples. The 30x30 nm plates were very bright and hard to bleach during imaging. The 60x60 plates were dim and bleached quickly. It was very challenging during imaging to capture the same number of frames as for the 30x30 or 60x90 plates. The 60x90 nm plates were bright and had an intermediate bleaching rate. Another possible reason is that the dyes were attached to the folded DNA origami via covalent bonding to a designated sites. We conclude that either the labelling of these plates has incorrect or the stability of the sample was compromised. This will be subject of future investigation.
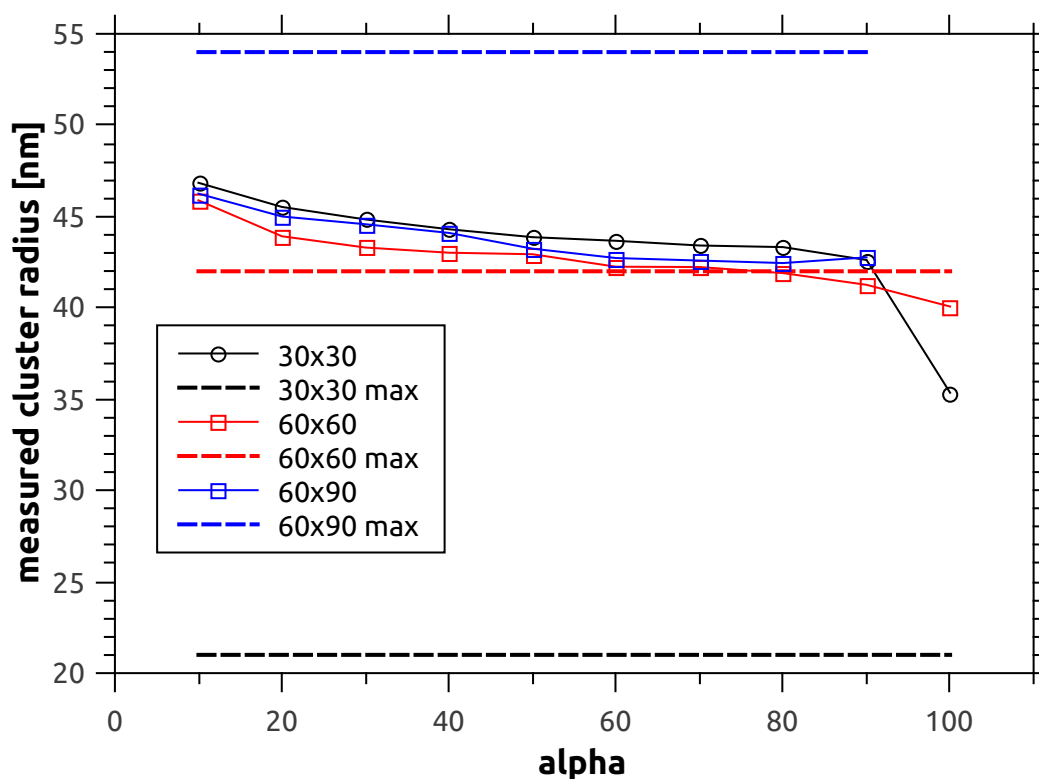
Figure 2.35: Cluster radius measured for localisation microscopy datasets of DNA origami plates. Different values of α were use to measure the cluster size in three types of plates 30x30, 60x60, and 60x90 nm (marked in black, red and blue respectively). For each plate the maximal expected radius was marked with a dash line of the same colour as of the results. Values presented here are an average of at least 10 different regions in three or more samples.

The 30x30 plates were brighter than the other imaged plates producing reconstructions with features bigger than these for 60x60 and 60x90 nm plates. A variation in structural stability of different batches of DNA origami structures was also observed. Thus, the 30x30 nm images should be treated with caution.

The 60x60 and 60x90 nm plates behaviour during imaging was comparable of typical behaviour of a labelled sample and other DNA origami structures imaged in our lab. The localisation microscopy data for these plates were analysed using the Rényi divergence
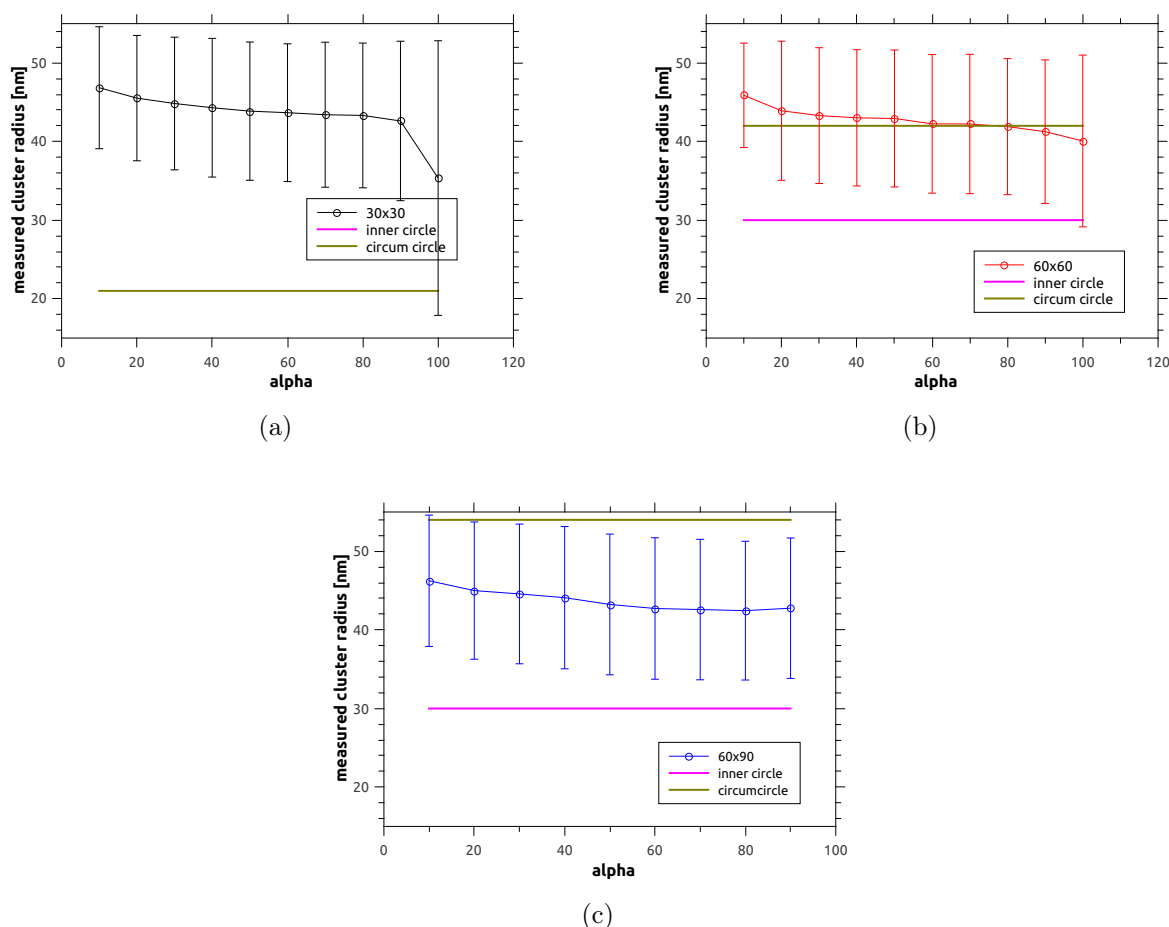
Figure 2.36: The average cluster radius measured for the three analysed DNA origami plates. Cluster radius measured with the Rényi divergence (using different α values) for three types of plates: a) 30x30 nm, b) 60x60 nm and 60x90 nm. For each imaged plate type at least 10 regions from three or more samples were analysed. The error bars are the standard deviations.

and Ripley's H function (see Table 2.4 and Figure 2.36). The measured cluster radius had very similar values for these structures. However, for both of those plates it was positioned in the range set by the minimal and maximal expected radius (as discussed for simulated DNA origami). The Ripley's H function measured the cluster radius to be on average three times bigger than the maximal expected radius.

Table 2.4: The cluster radius measured with the Rényi divergence and Ripley's H function for 30x30, 60x60, and 60x90 nm plates. The minimal and maximal expected radius for each plate and the number of data sets used to create the average is also given.

	plate	30x30	60x60	60x90
expected radius	min	15	30	30
	max	21	42	54
measured radius	Rényi divergence	43±9	42±9	43±9
	Ripley's H function	148±21	142±19	151±21
number of data sets analysed		58	30	45

Analysis of 60x90 nm DNA origami plates with added background noise

To test the influence of noise on the data sets 60x90 nm plates were imaged with added solution of Alexa Fluor 647 to the buffer. This test was only performed for the 60x90 nm plates because this structure was the most stable (see discussion above). The 60x90 nm plates were prepared according to the standard protocol and imaged without background noise added to the sample. Then 60 μl of 30 nM solution of Alexa Fluor 647 in PBS was added to the sample and left at room temperature for 30 minutes. Before imaging the sample was washed with 800 μl of PBS. After collecting images from 10 randomly selected areas on the sample, it was again washed with PBS. The washing was repeated after each imaging step to create background noise levels with different density (see Figure 2.37).

The cluster radius measured for sample after adding the liquid solution of Alexa Fluor 647 changed with the number of washes performed. The change can be quantified for both for the Rényi divergence and Ripley's H function by comparing the measured value with the value of measured radius for a sample without added background (see Table 2.5). The difference in measured cluster size is biggest for two washes following the addition of Alexa Fluor 647 solution to the sample.

The change in the cluster radius measured for the same sample with different stages of washing out floating dye can be explained by considering the 3D architecture of the sample. The DNA origami was immobilised on biotin-neutravidin coated glass slides. Thus after addition of the free floating dye, it may have congregated around the DNA origami
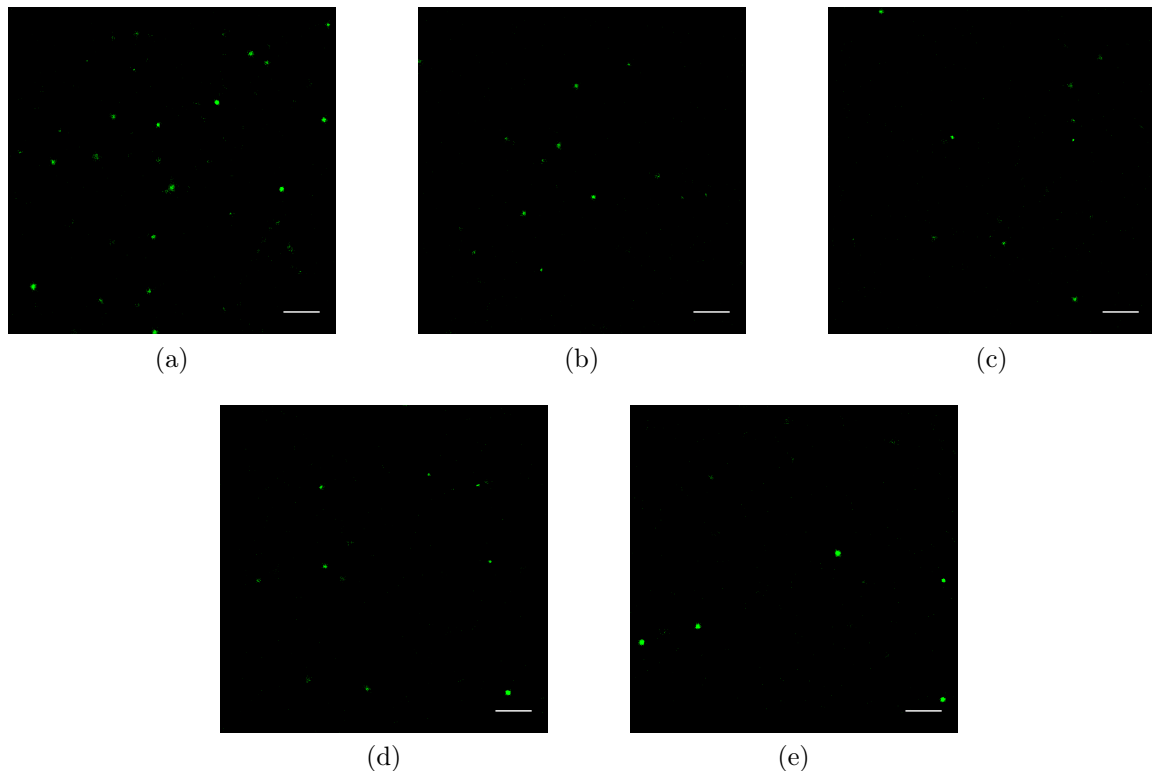
Figure 2.37: Fragments of localisation microscopy images of DNA origami plates 60x90. a) Imaged without adding any background. After adding 60 μl of 30 nM solution of Alexa Fluor 647 sample was washed b) one, c) two, d) three, and e)four times. Small clusters are especially visible in c and d). Scale bar: 1 μm.

creating slightly bigger clusters. Then, dye molecules, aided by washing steps, could have crossed to the biotin layer. Biotin may have created a mesh network where dye molecules were trapped and moving in a limited space. This could lead to appearance of artificial clustering visible on the localisation images (see Figure 2.37b-d). The cluster radius measured for different noise levels in the sample was lowered by the artificial clusters (see Figure 2.38).

The cluster radius measured with the Ripley's H function was less precise and accurate. However, it should be noted that it also displayed the same behaviour with the addition of free dyes in the sample (slightly higher radius measured after the first wash, lower values after the second and the third, and after the fourth the cluster radius had the same value as in the beginning).

Table 2.5: Percentage change in cluster radius measured with the Rényi divergence and Ripley's H function. The change in measured cluster size was calculated in respect to the measured cluster radius for DNA origami sample without background solution added. For "+" values the measured cluster size was bigger and for "−" smaller than the original measurement.

Number of washes	the Rényi divergence $\alpha = 70$	Ripley's H function
1	+5.05%	+1.13%
2	-6.36%	-3.33%
3	-2.64%	-4.30%
4	+0.33%	+0.38%

(a)

(b)
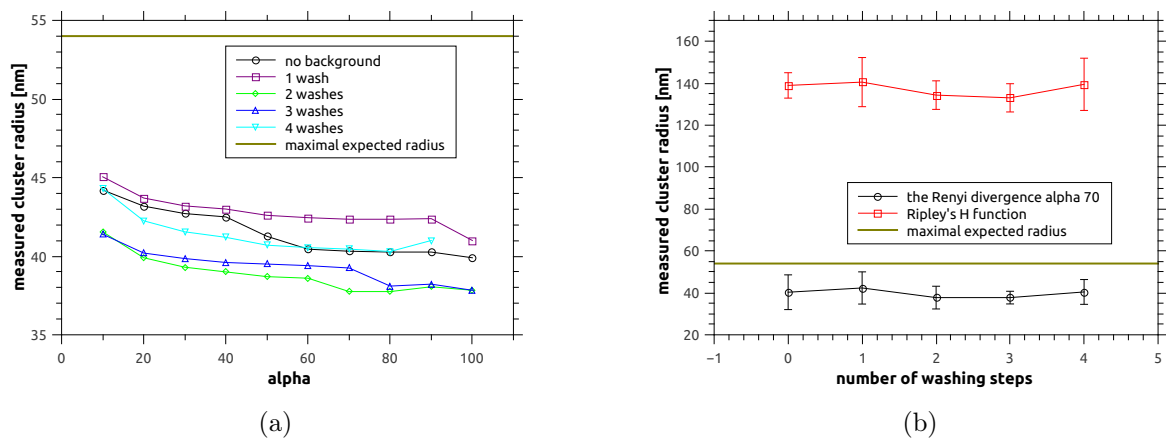
Figure 2.38: The cluster radius of localisation microscopy images of 60x90 nm DNA origami plates measured for samples with added background noise. a) the cluster radius measured with the Rényi divergence for different α values. The measured cluster size for DNA origami sample without added background is marked with black. After adding solution with Alexa Fluor 647 the cluster size increase by a small amount (marked with purple). After two and three washes the measured cluster size decreased due to presence of false clusters (marked with green and blue). Imaged after four washes the measured cluster radius value has risen to the value before the background was added. b) Comparison of the cluster radius measured with the Rényi divergence $\alpha = 70$ and Ripley's H function. The maximal expected radius for this data set is marked with olive (green) line.

2.7 Discussion and Outlook

Clustering analysis for localisation microscopy data can provide a lot of very important information about imaged structure. The structure presence, density or size can be evaluated and used for further study. Here, only the cluster size measurement was discussed. The main goal of this work was to evaluate the precision and accuracy of a new method of cluster analysis developed using the Rényi divergence. The results acquired with our method were compared with the well established Ripley's H function, and provided better accuracy in measurement of actual cluster size for both Gaussian and disk clusters. However, the precision of the Rényi divergence was worse than precision of the Ripley's H function for Gaussian clusters. Some properties of the Rényi divergence and Ripley's H function are collected in Table 2.6.

Some consideration should be also given to the selection of the scaling coefficient α. In this work the α was selected to be 70 for all of the analysed datasets mainly for consistency. This value was selected because it provided a stable radius measurement for different data types and noise levels analysed here. Generally, the α value should be selected using a pre-run on the data (similarly as in section 2.6.4) and set to a value for which the measured cluster radius does not change between different datasets which are going to be analysed.

One of the future aims is to use the Rényi divergence for analysis and study of localisation microscopy data of biological structures. This will require a prior knowledge about the actual size of the imaged structure. This condition is usually hard to meet. However, new advances have enabled imaging of biological structures with both fluorescence and electron microscopy. Correlative microscopy can provide information about the size of clusters with nearly unbiased electron microscopy precision and a localisation microscopy image of the sample structure.

Additionally, the high precision cluster radius measurement provided by the Rényi divergence could be used to develop analysis tools quantifying different properties of clusters present in the images. The already existing cluster analysis software can be changed to facilitate cluster counting or cluster density calculation. This can provide

even more complete set of information about the analysed structure.

Lastly, it is possible to change the cluster model shape and data estimation to accommodate different clusters shapes. This may be important especially for elongated, rod like clusters when the round shape of model cluster will cause underestimation of cluster size.

Table 2.6: Comparison of properties of the Rényi divergence and Ripley's H function. The calculation of both, the Rényi divergence and Ripley's H function was performed simultaneously by a single software. Both of these cluster analysis methods are the most suitable for analysis of circular clusters. Next, the Rényi divergence provides better accuracy of the cluster measurement than Ripley's H function. The precision of the cluster size measurement is different and depends on the cluster type: the Rényi divergence performs better for disk clusters and Ripley's function for Gaussian ones. Lastly, the Rényi divergence is much more noise resistant than Ripley's H function.

	the Rényi divergence	Ripley's H function
running time	fast(16 s for 32,000 point data set)	fast(16 s for 32,000 point data set)
cluster shape	circular	circular
accuracy	high	low
precision	high to medium	medium to low
noise resistance	high	low

Chapter 3

Modelling protein arrangements in the podosome ring

3.1 Introduction

Podosomes are discreet structures formed at the cell surface and form adhesive links to the extracellular matrix. They were first identified on the surface of bone cells (osteoclasts) in 1983 [Zambonin-Zallone et al., 1983]. In the first images of podosomes, they appear as bright dots surrounded by a microfilament web. These bright dots were identified, at the time, as actin clusters [Zambonin-Zallone et al., 1983]. These actin rich structures were finally named 'podosomes' after repeated analysis of the samples prepared with the same cell line, reviewed in [Murphy and Courtneidge, 2011]. Podosome presence was later confirmed in a range of different cell types: monocytes [Calle et al., 2004, 2006; Linder et al., 1999], dendritic cells [Calle et al., 2004, 2006; Meddens et al., 2014], stimulated endothelial cells [Rottiers et al., 2009; Seano et al., 2014], smooth muscle cells [Gimona et al., 2008], Src-transformed fibroblasts [Courtneidge et al., 2005], megakaryocytes [Schachtner et al., 2013a], microglia [Siddiqui et al., 2012], trabecular mesh work cells [Aga et al., 2008; Han et al., 2013], trophoblasts [Patel and Dash, 2012], and neural crest cells [Murphy and Courtneidge, 2011; Santiago-Medina et al., 2015]. The presence of podosomes or their analogues invadopodia, in a range of cell types indicates that these protein struc-

tures play an extremely important role in many physiological processes such as tissue invasion [Foxall et al., 2016].

Podosomes are build with an actin-rich core and a protein ring, which is the main focus of this chapter. The ring consists of integrins and integrins-associated proteins, such as vinculin, paxillin, and talin. The integrins-associated proteins are thought to form concentric circles around the podosome core [Meddens et al., 2013] and occupy different zones in the ring [Cox et al., 2012; Meddens et al., 2013]. However, previous investigations of the protein arrangement in the podosome ring were either based visual assessment of podosome rings in super-resolution images Cox et al. [2012] or on confocal studies [Meddens et al., 2013] (see section 3.1.1). Therefore, to investigate the protein arrangement further, we designed a pipeline for podosome identification and protein position measurement. The exact operation of this pipeline and proof of concept results are going to be discussed in sections 3.4, 3.4.1, and 3.5.1. The work presented in this chapter was published in [Staszowska et al., 2016].

3.1.1   Structure

Podosomes consist of three main components: a core surrounded by a ring structure and a cap placed on top of the core. The core is rich in F-actin and actin associated proteins, for example Wiskott-Aldrich Syndrome Protein (WASP) [Schachtner et al., 2013a]. WASP plays an important role in podosome formation through activation of Arp2/3, a major actin filament generator [Schachtner et al., 2013b]. The podosome ring consists of integrins and integrin associated proteins including vinculin, paxillin, and talin. The ring moderates and maintains the adhesion to the extracellular matrix (discussed in section 3.1.2) [Calle et al., 2006; Murphy and Courtneidge, 2011]. Recently, a number of studies reported a cap-like structure positioned on top of the podosome core [Bhuwania et al., 2012; Wiesner et al., 2014] (see Figure 3.1). This cap was reported to interlock with the core structure and only a limited number of proteins building it were identified for example supervillin, (see Figure 3.1) [Wiesner et al., 2014]. It was confirmed that the actin filaments connecting individual podosomes are branching from the cap [Bhuwania et al., 2012; Linder and

Wiesner, 2015]. However, the exact role of this structure is not yet understood [Linder and Wiesner, 2015].
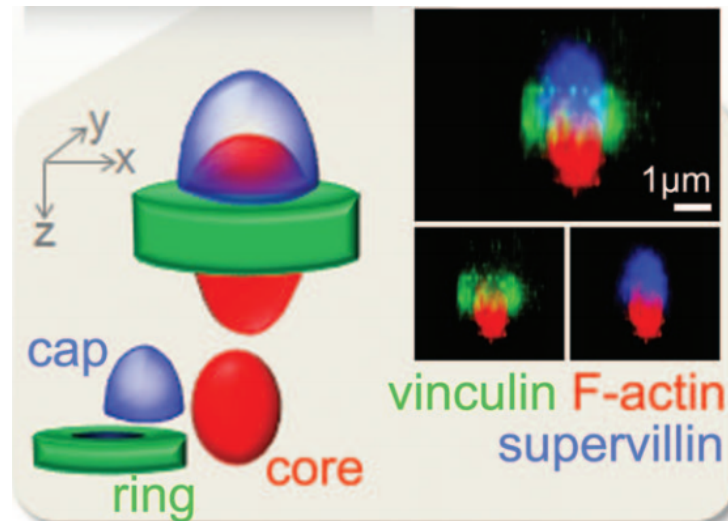
Figure 3.1: 3D structure of podosome. The core (red) is surrounded by the ring structure (green). The cap (blue) is positioned on top of the cove above the ring. Right: A model of podosome reconstructed using confocal images of podosomes. Vinculin from the ring is seen surrounding F-actin core (red). F-actin core is topped with the supervillin built cap. Image adapted from: [Wiesner et al., 2014].

The size of podosomes varies from 0.5 μm to 2 μm in diameter and 0.6 μm to 1 μm in depth (the core usually has a diameter of 0.5 μm and height of 0.5 μm [Meddens et al., 2014]). Their lifetime is limited to a few minutes [Linder and Kopp, 2005; Veillat et al., 2015]. Until recently, the shape of the podosome ring was believed to be round [Monypenny et al., 2011]. Recent studies of the podosome ring based on analysis of super-resolution microscopy images, present two conflicting models of the ring structure. The ring was reported to have a hexagonal shape when imaged with high density localisation methods using live and fixed cells expressing fluorescent proteins [Cox et al., 2012]. An opposing model was presented in a study using primary/secondary antibody labelling with low density activation localisation microscopy. Here, the rings were seen as sparse collections of protein clusters [van den Dries et al., 2013; Meddens et al., 2014].

Because of the conflicting observations, the arrangement of proteins in the podosome ring is still unknown. It was, however, confirmed that the different proteins in the podosome ring occupy discrete spaces in the ring [Cox et al., 2012; Meddens et al., 2014]. This led to speculation about possible arrangements of proteins in the podosome rings. One possible configuration of proteins in the ring is shown in Figure 3.2. The protein arrangement was also assessed visually using localisation microscopy images, finding talin to be closer to the podosome core than vinculin [Cox et al., 2012]. A complete model of podosome ring structure was proposed in Meddens et al. [2013], however it was based on basic confocal images of podosomes. Development and preliminary results of a quantitative pipeline for analysis protein arrangement in the ring will be discussed in section 3.3. Development of a model for protein arrangement of the proteins in the podosome ring could provide an insight into podosome dynamics and functions.

The main difference between podosomes and other types of adhesions is that podosomes have an actin core which plays a part in the matrix degradation process. Podosomes also contain WASP protein. Detection of WASP helps to distinguish podosomes from focal adhesions [Calle et al., 2004; Linder et al., 1999] (see Figure 3.2).

Podosomes associate as groups inside cells. They are interconnected by myosin cables and form structures such as clusters or belts [Veillat et al., 2015]. Veillat et al (2015) also present several other types of podosome organisation in cells shown in Figure 3.3. This organisation behaviour usually depends on the type of cell forming podosomes and their physiological function. For example, an array of podosomes is formed in moving cells.
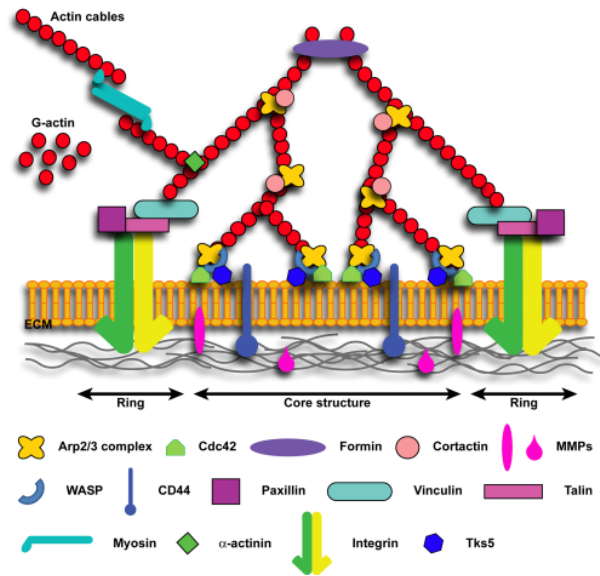
Figure 3.2: Podosome structure: a schematic representation of podosome structure. Integrin proteins control binding to the extracellular matrix (ECM). The actin cables link the ring structure with the core. Talin and paxillin are linked to the integrins and vinculin provides a link to the actin cables branching from the core. Note that this is only a model of one possible arrangement which has been proposed for the proteins in the ring. The actin core is attached to the cell membrane by Cdc42-activated WASP and Arp2/3 complex. Image adapted from: [Schachtner et al., 2013b].
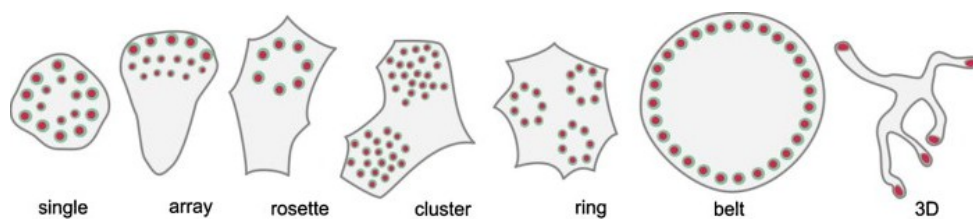
Figure 3.3: Examples of podosome organisation in different cell types and life states. Starting from the left, single podosomes are present in stationary macrophage cells and arrays in moving cells. Podosome rosettes are present in TGFβ-induced endothelial cells. Clusters can be found in immature dendritic cells. Podosome rings and belts are present in osteoclasts. 3D podosomes organisation present in cells embedded in gelled collagen matrix. Image adapted from: [Veillat et al., 2015].

3.1.2   Main functions

Adhesion

Adhesion is one of the essential functions of podosomes. Podosomes form when there is contact between a cell and the substratum. It was recently confirmed that podosome rings are in a very close proximity to the extracellular matrix material [Schachtner et al., 2013a]. Theoretically, any of the extracellular matrix proteins can be used to form an adhesion site, however some proteins are used more than others [van den Dries et al., 2012]. For example, fibronectin is a preferred substratum for macrophage podosomes and vitronectin for osteoclast podosomes [Veillat et al., 2015]. Although the actual signalling pathway for podosome assembly is not yet known, a number of studies have reported on the order in which proteins are recruited to the forming podosome. It was confirmed that the podosome core appears before the podosome ring and later induces its formation [Luxenburg et al., 2012]. The first protein recruited to the ring in osteoclasts is paxillin (one of the podosome ring proteins) [Veillat et al., 2015].

Once podosomes are made they then play a role in podosome mediated tissue remodelling. In bone remodelling, podosomes form a tight adhesion junction what enables bone resorption by osteoclasts (followed by remodelling the bone tissue by osteoblasts). Podosomes are also responsible for creating new capillary vessels from existing blood vessels, they induce the outgrowth and creation of new capillaries (this process is known as angiogenesis) [Seano et al., 2014].

Mechanosensing

Podosomes can sense and adjust to the geometry and stiffness of the extracellular matrix (ECM). The mechanosensory ability of podosomes was discovered by observing how the stiffness of the matrix influences podosomes [Veillat et al., 2015], and was recently confirmed by detecting myosin II and actin fibres in podosomes [Burgstaller and Gimona, 2004; Meddens et al., 2014]. The mechanosensing mechanism is driven by actin polymerisation, moving the actomyosin fibres connecting the extracellular matrix to the integrins in the ring [van den Dries et al., 2013]. An other example of this function is a mechanosens-

ing mediated relation between the topography and composition of the extracellular matrix and the lifetime and shape of the podosome. Another example of mechanosensing in podosomes is probing the 3D architecture of the substrate present in immature dendritic cells (see Figure 3.3: showing a cell displaying podosomes in 3D environment) [van den Dries et al., 2012]. The same probing property is used by the podosomes to form protrusion sites. The protrusion formed to aid tissue transmigration, forms in places offering the least resistance. In cultured cells this was seen on porous hard substrate filled with softer substance, for example gelatin (see Figure 3.4) [Veillat et al., 2015].
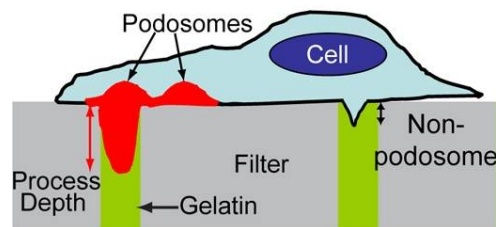
Figure 3.4: Model of a podosome forming a protrusion to the softer environment. Cell was planted on a porous substrate. Podosome forms on the softer substrate (here gelatin) to form a protrusion site. For comparison protrusion formed by the cell (on the right). Image adapted from: [Gawden-Bone et al., 2010]

The ability of podosomes to form protrusions is used in two different processes to find sites for the cell to cross a barrier in processes called extravasation and cell invasion. Extravasation is the process of crossing between capillaries and tissue. Immune response to inflammation is based on leukocyte movement between infected tissues and lymphatic system. The process of extravasation is enabled by podosomes, which form protrusions in and between endothelial cells to find suitable crossing site [Carman, 2009; Veillat et al., 2015].

3.1.3 Pathology

Any disturbance in podosome formation and motility can cause malfunction in cell function leading to disease. The most well-known disease connected to abnormalities in podosome formation is Wiskott-Aldrich syndrome [Calle et al., 2004]. It is an immunodeficiency syndrome caused by mutation of the WASP gene [Thrasher and Burns1, 2010]. Macrophages and immature dendritic cells, collected from the patients suffering from this condition, are unable to form podosomes [Jones et al., 2002; Zicha et al., 1998]. Lack of podosomes affect the ability macrophages to move in the body [Blundell et al., 2008]. Other example of podosome related disease is an autoimmunitive desease known as PAPA syndrome (pyogenic sterile arthritis, pyoderma gangrenosum, and acne). In this condition podosomes formed by macrophages are malformed and have much stronger adhesive and matrix degrading abilities [Veillat et al., 2015].

Metastatic cancer is an another example of disease state in which podosome-like structures, called invadopodia, play a crucial role. The cancer cells in HNSCC (head and neck squamous cell carcinoma), breast carcinoma and melanoma cells all assemble invadopodia (as opposed to healthy cells). Invadopodia are believed to facilitate cancer metastasis [Murphy and Courtneidge, 2011]. Despite their similarity to podosomes, invadopodia differ in few aspects. Firstly, invadopodia have higher protrusive and matrix degradation abilities. They are also bigger, have a longer life and are located in the centre of a cell [Foxall et al., 2016]. Another difference is that there are no connections linking individual invadopodia in the same cell and as a result, they seem not to associate with one another.

3.2 Imaging podosomes at super-resolution

3.2.1 Podosome sample preparation

Live and fixed samples of macrophage podosomes were prepared by Elizabeth Foxall, a PhD student in Gareth Jones group, according to the protocol presented in [Vijayakumar et al., 2014]. For a complete protocol see Appendix C.

3.2.2   Tandem dyes for two colour imaging

Tandem dye pairs were used for two colour localisation microscopy imaging. These pairs of dyes are composed of two fluorescent dyes covalently bonded to an antibody. One of these dyes acts as an activator and the other as an acceptor. A tandem dye pair behaves as a dye which has the excitation characteristics of the donor dye and emission properties of the acceptor. This arises because of Förster resonance energy transfer (FRET), a contactless process of energy transfer (see section 4.2 for more details) [Bates et al., 2007].

For the purpose of two-colour imaging three tandem dye pairs were used: Alexa Fluor 405-Alexa Fluor 647, Cy2-Alexa Fluor 647, and Cy3-Alexa Fluor 647. During imaging the activator dye (Alexa Fluor 405, Cy2, and Cy3) was excited using relevant wavelength of laser light (405, 488, 561 nm respectively). The read-out was performed using by exciting Alexa Fluor 647 dye, photoswitching of which was facilitated using the activator dye [Bates et al., 2007]. To label sample with tandem dyes, they were conjugated to the secondary antibody. The dye was handled and conjugated by Oleg Glebov, using a protocol provided by Nikon (see Appendix D).

The main disadvantage of tandem dyes is cross-talk when using them for two colour imaging [Dani et al., 2010]. Some portion of activator dyes are activated by laser light with wavelengths different than the wavelength of maximum absorption. For example Cy2 will have maximum efficiency when activated with 488 nm wavelength light, and will be less sensitive to activation with 405 or 561 nm laser light. However, it can be still excited by 405 or 561 nm light [Bates et al., 2007]. This means that dyes can be activated by laser light used to image a second dye pair leading to miss-identification of tandem pairs attached to different proteins.

The amount of cross-talk can be measured in cross-talk experiments. In this test samples are labelled with only one different tandem dye pair. Then the imaging is performed as for sample labelled with two different dye pairs. The amount of cross-talk is measured as a ratio between the number of localisations measured for the optimum activation and number of localisations for the second channel (used in two colour imaging to activate the second dye pair) [Bates et al., 2007; Dani et al., 2010]. Another problem is smaller than

100% efficiency of FRET, which depends on the distance between donor-acceptor pair (it is $\propto \frac{1}{r^6}$) and is reduced by acceptor quenching [Fiori and Meller, 2010]. The amount of quenching depends on the tandem dye pair used [Fiori and Meller, 2010]. Lastly, tandem dye pairs, as any other fluorophores, can be damaged by photobleaching and oxygen radicals.

3.2.3   Imaging

The localisation microscopy imaging was performed using the Nikon N-STORM system at the Nikon Imaging Centre at King's College London. The Nikon system is built with an Eclipse Ti-E Inverted Nikon Microscope with TIRF 100x objective, N.A. 1.49, an Andor DU-897 camera, laser and LED light sources (laser wavelengths and powers: 405 nm, 30 mW; 488 nm, 90 mW; 514 nm, 50 mW; 561 nm, 90 mW, and 647 nm, 170 mW) and operated with NIS Elements software. The laser power was adjusted during the acquisition to compensate for the bleaching rate so a similar number of counts was acquired in every frame (as far as possible). During acquisition the 405 nm laser was used induce population of fluorophores in the dark state to re-enter the emitting state. This was important towards the end of the acquisition when the sample started to become bleached. The 405 nm laser power was used at the end of the acquisition was usually smaller than 12 mW (40% of the maximal laser power). To improve the signal-to-noise ratio of the acquired images, the imaging was performed in TIRF (or near-TIRF). This restricted excitation to the sample surface, so only fluorophores on the sample surface (up to 200 nm) emitted light. This was particularly advantageous when imaging structures like podosomes, which form on the sample surface. In this mode fluorescent molecules attached to different parts of the cell were not illuminated.

During imaging samples were placed in an imaging buffer to control the density of active fluorophores by inducing a long lived dark state. The basic imaging buffer contains an oxygen scavenging system and reducing agent. The oxygen scavenging is performed by two buffer ingredients: glucose oxidase and glucose. Glucose is oxidised by glucose oxidase and the oxygen from air is reduced to hydrogen peroxide (H_2O_2) in this reaction.

This creates an relatively oxygen free environment around the sample. Hydrogen peroxide is a strong oxidising and corrosive compound. In order to control its concentration the imaging buffer contains catalase enzyme. Catalase breaks down the hydrogen peroxide into to water and oxygen. The third active ingredient is a reducing agent, usually a thiol, for example β-Mercaptoethylamine (MEA). The role of reducing agent is to quench the triplet state [Dave et al., 2009; Heilemann et al., 2005; Rasnik et al., 2006]. The basic imaging buffer recipe was based on the protocol supplied by Nikon [Nikon, 2015]. Recently a number of studies reported resolution improvement through improving buffer recipes [Dave et al., 2009; Olivier et al., 2013]. According with the methodology presented in [Dave et al., 2009] Cyclooctatetraene (COT) was added to the methanolamine (MEA) buffer to final 2mM concentration. The recipe for the imaging buffer can be found in Appendix E.

3.3 Quantitative model of the podosome ring structure

Localisation microscopy images are reconstructed from a collection of points, thus the structures of interest must be identified manually or by a software. It is possible to try to detect recurring shapes without a model of what the shape should be but this is computationally expensive. Thus it is easier to identify features using a model of the structure. In order to identify podosome rings, a model of the ring structure was created. Podosomes rings were modelled as circular (see localisation microscopy images of the podosome rings in fixed cells in Figure 3.5). The circle centre and radius can be found analytically using coordinates of three points positioned on that circle. The circular model is a good method to identify potential podosome centres and radius values even if the actual podosome structure is hexagonal (a circle can be passed through corners of a hexagon an it will have the same centre point and radius of a half of the hexagon diagonal) or elongated (when a bigger range of possible podosome centres is returned, for example, see podosome centre identifications in Figure 3.8c). The protein positions were later investigated using the podosome centre identifications. After podosome identification the relative protein positions were calculated. Since the podosome size varies between

different podosomes and some of the podosomes have an elongated shape and are not actually circular the protein positions were calculated relative to each other (the average position of one protein was subtracted from the average position of the second).

<div align="center">(a)   (b)</div>
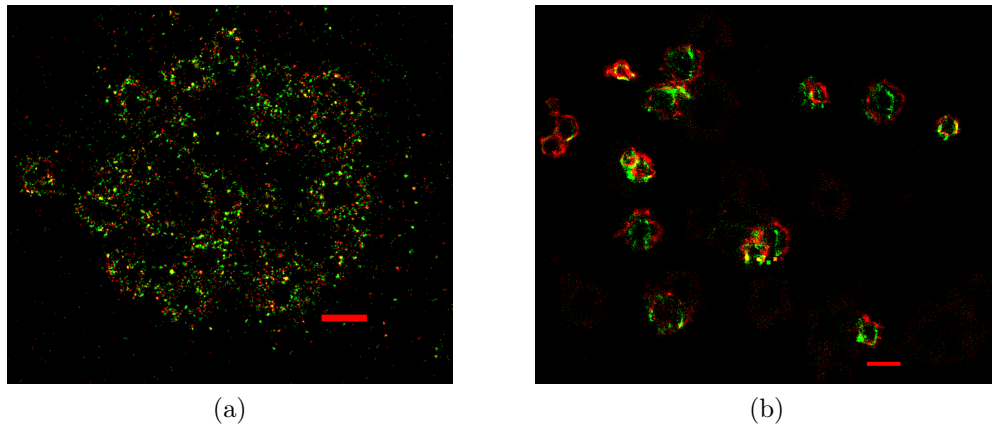
Figure 3.5: Example of images of podosome rings imaged with localisation microscopy and 3B. a) Podosomes imaged with localisation microscopy. Podosome ring proteins were stained with tandem dye pairs: vinculin (green) with Cy2-Alexa Fluor 647 and paxillin (red) with Cy3-Alexa Fluor 647. b) Podosome rings imaged with 3B method. Talin was transfected with mCherry-talin construct (green) and vinculin labelled with Alexa Fluor 488 (red). Scale bar: 1 μm.

The circular model of podosome ring can be easily implemented computationally. A circle passing through three non-collinear points (x_1, y_1), (x_2, y_2), and (x_3, y_3) can be found either geometrically, by finding the crossing point between lines perpendicular to the segments connecting the points (see Figure 3.6), or by solving the equation of the circle:

$$(x - x_0)^2 + (y - y_0)^2 = r^2, \tag{3.1}$$

where the (x_0, y_0) is the centre point of the circle and r is the radius. The equation 3.1 was solved for three points positioned in the same circle and the centre point was found. The radius was calculated as the distance between the centre point and any of the three points positioned on that circle.
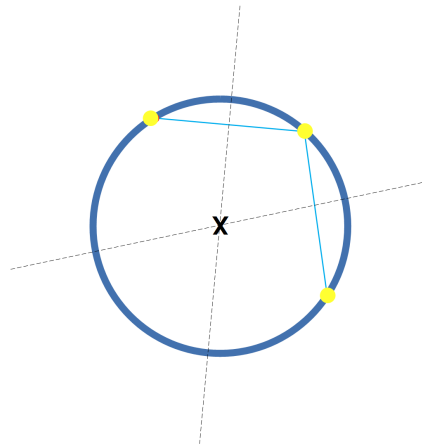
Figure 3.6: A geometrical construction of a circle passing through three points. If we draw two perpendicular lines crossing the segments connecting the points, then their crossing point will be the centre of the circle.

3.4   Podosome identification algorithm

The potential locations of podosome rings in localisation microscopy images were found by random sampling of the localisation microscopy dataset and fitting these random samples a circular model of the podosome ring structure. The equation of the circle fitted to the three randomly selected points was found using equation 3.1. To ensure that the points selected were suitable, the software randomly selected three points separated by small enough distances that they could belong to the same podosome (step 1 in Figure 3.7) [Fischler and Bolles, 1981]. Then the circle centre and radius were calculated (step 2 in Figure 3.7). The fitted circle parameters had to have biologically sensible values: the circle radius had to have a similar value to an actual podosome ring radius (500–1000 nm) and the inside of the circle had to have a very small number of protein localisations (less than 10 localisations in the centre, because there should be no fluorophores present in the podosome cores). Lastly, the overlap between the fitted circle and actual podosome structure was evaluated by examining the localisations surrounding the fitted circle in the proximity of 400 nm which corresponds to an actual podosome ring thickness. The values used for this step were based on experimental values acquired for a typical podosome ring

from a localisation microscopy image (usually for a good overlap the density of points should be approximately uniform, step 3 in Figure 3.7). If the criteria were met the parameters of the fitted circle were saved (step 4 in Figure 3.7), and if not they were discarded. Next the new set of three points was selected and fitting steps were repeated (steps 1-3 in Figure 3.7). Usually, circles were found across the whole image, however the density of identification in the area with podosomes was much higher.

The software ran for a predefined number of repetitions, set to be 100x higher than the number of points in the data. This was smaller than the number of all possible combinations of selecting three point from the data, for typical podosome data it provided around 3000 circle fittings for an image of a cell with around 20 podosomes, and multiple circles were fitted to the same podosome (after the filtering step). Additionally, the lower number of repetition shortened time of the computation significantly. The filtered circle fittings were saved and plotted in an image for reference (see Figure 3.8b). The podosome centres were then found by convolving a Gaussian function with the centre points of the fitted circles and finding the local maxima (images created in convolution process and identified podosome centres were also saved, see Figure 3.8). The maxima are localised around the true podosome centres, as the fitted circle centres were more densely localised in these areas (step 5 in Figure 3.7). Compared with the visual assessment of the images, the software had an 80% success rate (the remaining 20% were false positives and negatives). This success rate is comparable with other methods of podosome identification reporting 75% success rate [Meddens et al., 2013].
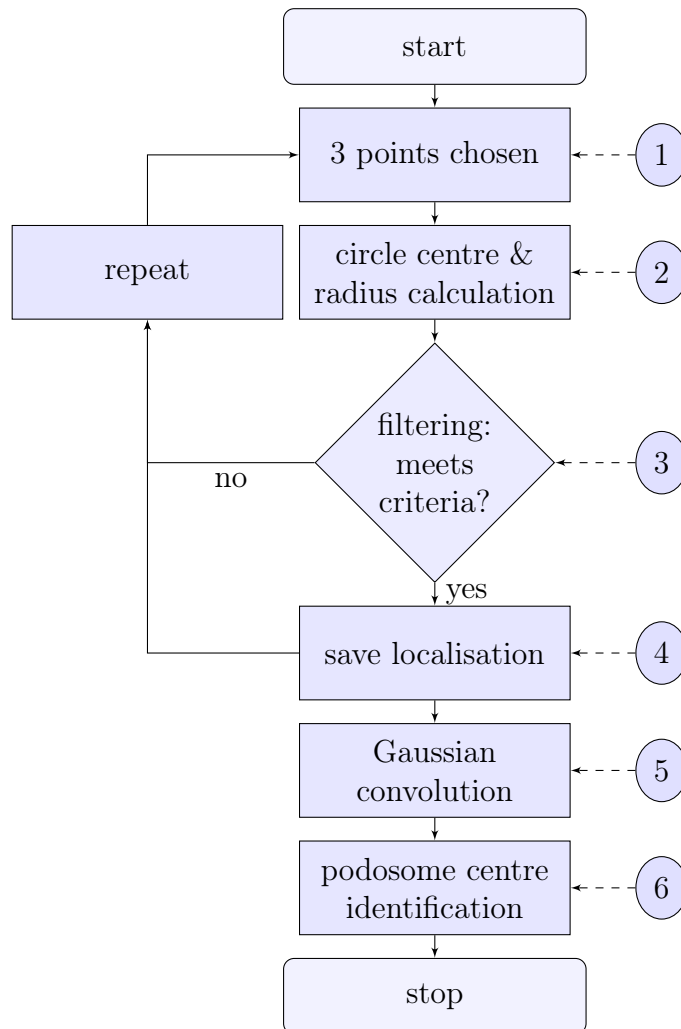
Figure 3.7: The podosome localisation software operation. 1) The three points are selected randomly from the localisation microscopy data. 2) The radius and a center of circle passing by three points are calculated. 3) Filtering step checking if the circle fit is close to the biological properties of the podosome ring. If the filtering criteria are met the circle centre is saved (step 4) the steps 1-3 are repeated. Otherwise a new set of points is chosen (back to step 1). 5) The circle center positions are convolved with the Gaussian and saved as image. 6) The podosome centres are find by looking for a local maxima on the Gaussian image from step 5.
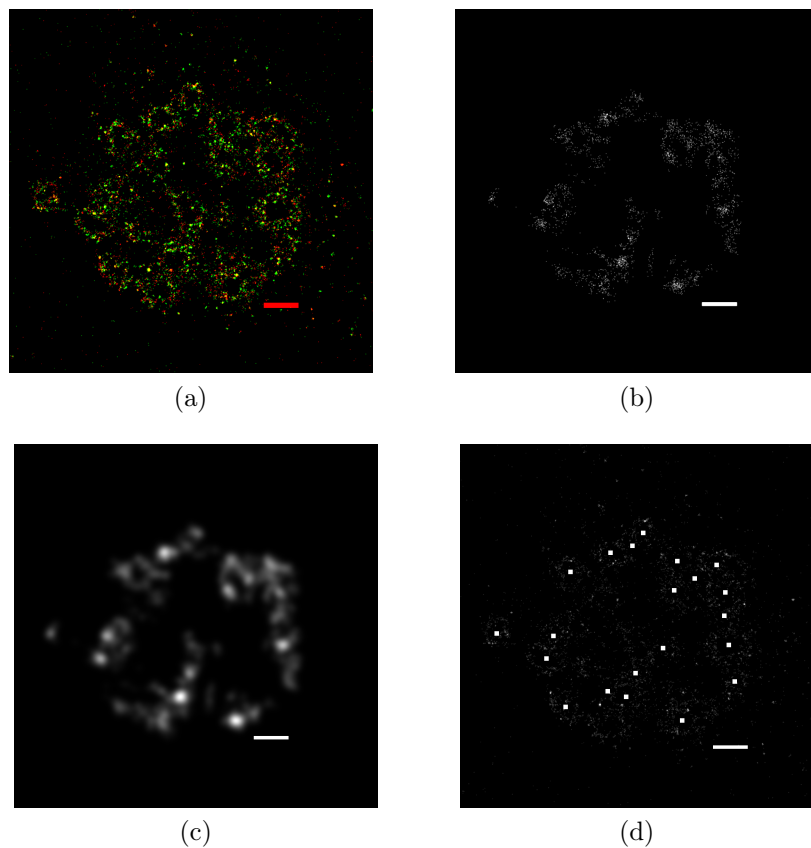
Figure 3.8: Podosome identification in localisation microscopy images. a) Image of po-
dosome rings with vincullin (green) stained with Cy2-Alexa Fluor 647 and paxillin (red)
with Cy3-Alexa Fluor 647. b) Filtered circle centres found using the algorithm (image
was blurred for improved visibility). c) All possible centre points were convolved with the
Gaussian function. The actual centres of podosomes were found as local intensity maxima.
d) Identified podosome centres displayed as white squares displayed over the localisation
microscopy image of a cell displaying podosomes. Scale bar 1 μm.

3.4.1   Protein distance calculation

The ring protein positions were calculated by subtracting average position of one protein from average position of the other protein. This method of calculation helped to minimise bias introduced by different podosome size and the ring elongation (see Figure 3.10). The podosome centres found in the previous step (see section 3.4) were used for the protein distance calculations. The boundary of the podosome ring was hard to define computationally due to close proximity of other podosomes and background noise. However, the ring boundaries can be easily distinguished by human eye. Thus to define the boundaries, a black-and-white mask image was created showing the areas with podosomes in white and background in black (step 1 in Figure 3.9). Creating the mask images had two additional advantages of excluding areas where the relative positions of the proteins in the ring might be different (for example parts of the ring shared two podosomes) and removing false positive podosome identifications.

The podosome ring shape can sometimes be elongated, thus the proteins in the ring may not be equally spaced from the podosome centre. The calculations of the relative positions were performed for small sections of the rings. The angular increment for creating these sections was set to 0.1 radian – this value was selected to provide on average about ten localisation of each protein per segment. For easier computation the coordinate system was changed to polar coordinates (where the angular increment is easy to define, step 2 in Figure 3.9). The calculated average difference in positions of the two protein in each segment was then weighted by the number of points in given segment – so the regions with small numbers of localisations would not influence strongly the end result (step 3 in Figure 3.9).
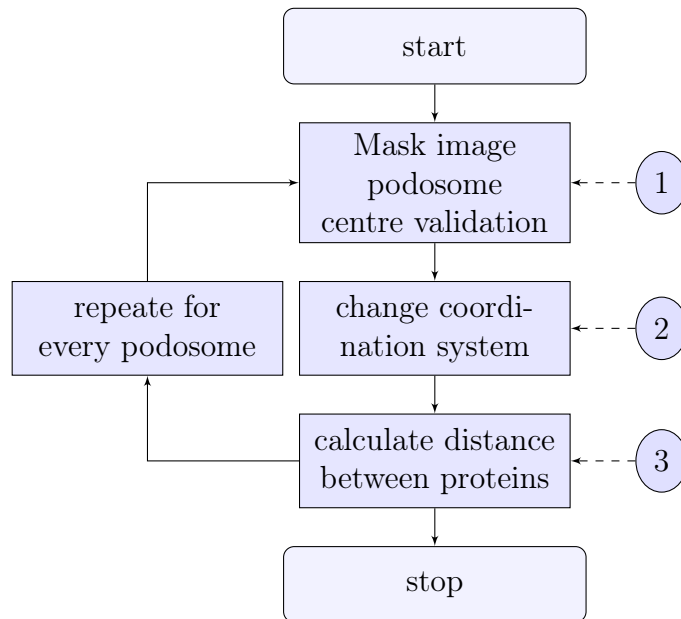
Figure 3.9: The operation of the program for calculating positions of proteins in the podosome rings. The program iterates through the list of podosome centres provided by the podosome identification software and for every podosome centre the points which belonged to the same podosome are identified using the mask image (step 1). Then the coordinate system is changed from Cartesian to polar coordinates (step 2), with the current podosome centre becoming the centre of the polar coordinate system. Each point belonging to the podosome is described by the distance to the centre and its angular position. For details of the coordinate conversion see appendix F. The relative distance between two proteins was calculated and saved (step 3).
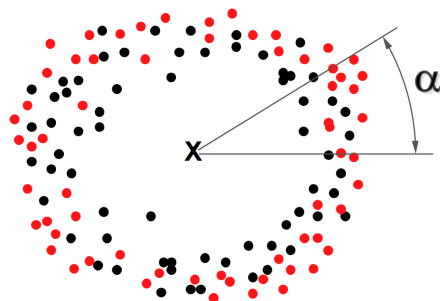
Figure 3.10: The relative protein position calculation. The average position of each ring protein is calculated in a segment (α). The average position of the first protein (marked with black dots) is subtracted from the average position of the second (red dots). The positive value of the relative distance means that the first protein is closer to the podosome centre than the second, the negative means that the second protein is closer. This method of protein distance measurement provides values unbiased by difference in podosome size and ring elongation.

3.5   Results

3.5.1   Quantitative arrangement of the protein in the podosome ring

The methodology developed for podosome ring analysis was used to perform calculations of protein positions in the podosome ring. The relative positions of the vinculin-paxillin pair were calculated for samples stained with tandem dye pairs (and imaged using the Nikon N-STORM system). The talin-vinculin pair was imaged using samples prepared with mCherry-talin construct, and vinculin stained with Alexa Fluor 488 (imaging and super-resolution analysis was performed by Susan Cox, for details see [Cox et al., 2012]). These two types of sample preparation methods were chosen because they suggested a completely different podosome ring shape. Thus models built with different sample preparation model could be compared, leading to a better understanding of both structure and shape of the podosome ring. The relative positions of two proteins in each pair were calculated (see section 3.4.1) and the resulting distributions are shown in Figure 3.11. To create these distributions around 380 podosomes (data collected from four samples, see Figure 3.11a) for the vinculin-paxillin pair and 9 podosomes from a single data set for the talin-vinculin pair were analysed (see Figure 3.11b).

The distribution for the vinculin-paxillin pair has a very sharp peak and relatively heavy tails. Mean, median, standard deviation and the 1st and 3rd quartiles were calculated for the vinculin-paxillin distribution (see Table 3.1). Both mean and median of this distribution suggest a very small difference between positions of vinculin and paxillin. The mean value is equal to 4 nm and median to 1 nm, suggesting that the paxillin is located further away from the podosome centre than vinculin. These values were to small to provide a definitive answer about the arrangement of these two proteins. Both the mean and median were smaller than the proteins size (see the minimal protein size in Table 3.3).

Additionally a Gaussian and t-location scale pdfs were fitted to the vinculin-paxillin relative distance distribution. The resulting fit is shown in Figure 3.12 and parameters

(a)   (b)
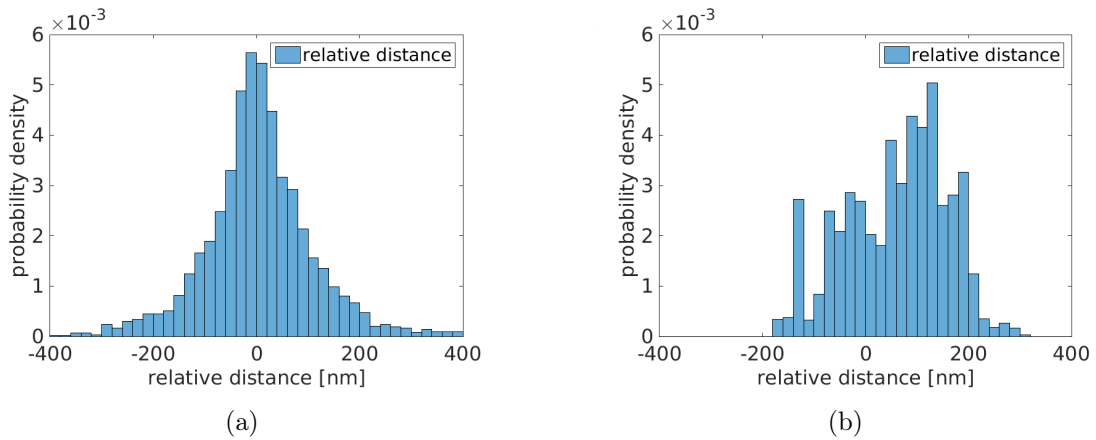
Figure 3.11: Histograms of the position difference between pairs of proteins from the po-dosome ring. a) The relative position of vinculin in respect to paxillin (results acquired with standard localisation microscopy). The relative positions were calculated by sub-tracting an average position of one protein from the second. For positive positive values paxillin is closer to the podosome centre, and for negative vinculin is closer. b) The relative positions of talin and vinculin (results for 3B analysed images). For positive values talin is closer to the centre, while for negative vinculin is closer.

Table 3.1: Statistical parameters of distributions of the relative protein distance calcula-tions for vinculin-paxillin and vinculin-talin protein pairs.

protein pair	mean relative distance [nm]	median relative distance [nm]	Standard deviation	quartiles
vinculin-paxillin	4	1	110	(-40,30)
vinculin-talin	61	75	99	(0,140)

are collected in Table 3.2. Neither of the fitted pdfs provided a good fit for the results for vinculin-paxillin pair. There was no fitting performed for the talin-vinculin pair because this results distribution has a bimodal shape and no readily available probability density function has this shape has this shape.

An analysis of nine podosome rings provided an indication of the relative positions of talin and vinculin in the podosome ring. A data set with two of the podosome ring
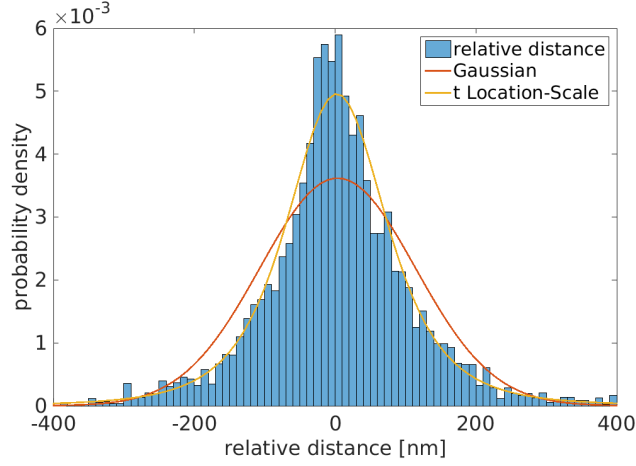
Figure 3.12: Fitting of Gaussian and t Location-scale pdf to the histogram of the relative positions of the vinculin-paxillin pair. However, neither of these functions provided a good fit for the results (see Appendix B for distribution equations and further details).

Table 3.2: Fitting parameters of Gaussian and t Location-scale probability functions fitted to the relative protein distance measurements results for vinculin-paxillin pair.

Function	μ	σ	ν
Gaussian	$3.68(0.59, 0.59)$	$110.24(0.42, 0.42)$	
t Location-Scale	$2.49(0.48, 0.47)$	$74.58(0.56, 0.57)$	$3.25(0.068, 0.070)$

proteins vinculin stained with Alexa Fluor 488 and talin encoded with mCherry-talin construct was imaged and analysed by Susan Cox [Cox et al., 2012]. The relative position measurements indicated that on average talin was 60 nm closer to to the podosome centre than vinculin (see Table 3.1). When considering the median value the relative distance is 75 nm. Similarly, the absolute distances measured for the two considered proteins from the podosome centre suggest that the talin is closer to the podosome centre than vinculin (see Figure 3.13 and Table 3.4).

Table 3.3: Minimal sizes of podosome ring proteins calculated using methodology presented in [Erickson, 2009]. The minimal volume of space occupied by protein with a certain mass can be calculated using equation $R_{min}[nm] = 0.066M^{\frac{1}{3}}$,where M is the mass of the the protein, measured in Daltons [Erickson, 2009] (for full explanations see Appendix G).

protein	mass [kDa]	R_{min} [nm]	notes
vinculin	117	3.23	circular only in inactive state
paxillin	69	2.71	
talin	270	4.27	

Table 3.4: Statistical parameters of the absolute position distributions of talin and vinculin. Mean, median, standard deviation and quartiles were calculated (the 1st and 3rd quartiles were calculated to account for spread around the median and asymmetry of the distributions).

	mean [nm]	median [nm]	standard dev.	quartiles
vinculin position	571.84	575.11	85.87	(520, 620)
talin position	511.05	466.04	151.99	(400, 500)

The absolute distances of the proteins from the podosome centre can be calculated and compared for datasets with podosomes with similar size. However, to provide a statistically significant answer a larger number of data sets would have to be analysed. This is especially important when considering the distribution of talin absolute distances which had a large standard, deviation indicating broad distribution. Analysis of the talin absolute distance provided a bimodal distribution (see Figure 3.13). This may indicate two stable conformational states of talin (this matter will be investigated further).

Two protein labelling for localisation microscopy is challenging and requires dedicated method of imaging or sample preparation. Samples must be stained with specially prepared dyes, for example tandem dye pairs or imaged and analysed by high density methods when using fluorescent proteins (for example 3B [Cox et al., 2012]). Use of conventional dyes (e.g. Alexa Fluor 488, 561, or 647) is possible only using high power laser to induce a high enough rate of random switching between emitting and dark state
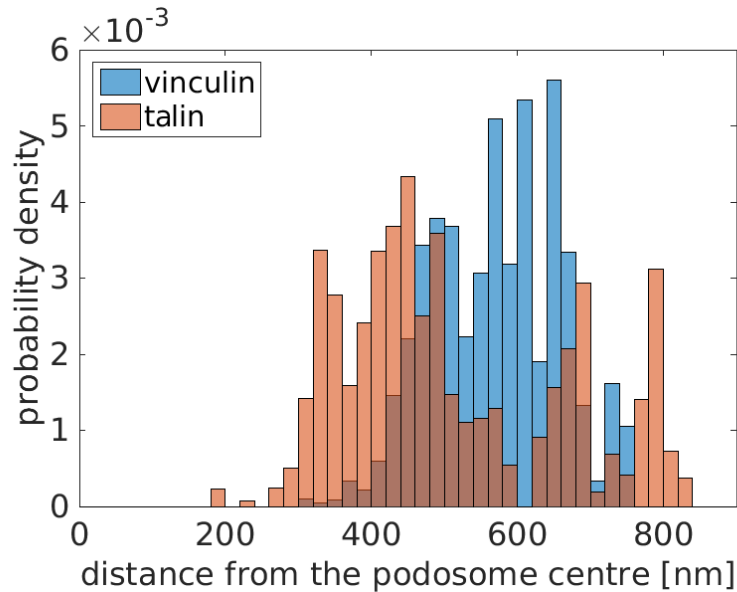
Figure 3.13: Histogram of the absolute positions of talin and vinculin in podosome rings. The absolute protein positions in each segment were calculated measuring the average distance of each protein from the podosome centre.

and requires chromatic realignment. Another possibility is imaging with a buffer with free-floating dyes, which are imaged only after they attach to a protein of interest. This is done using technique known as point accumulation for imaging in nanoscale topography (PAINT) [Sharonov and Hochstrasser, 2006]. However, there are drawbacks to each of the two-colour methods. The tandem dye pairs have to be conjugated by hand and are unstable (stability depended on the batch of dyes and varied between two weeks to a month). Also each new batch of tandem dyes has to be evaluated for cross-talk between channels (see Appendix D). The high density methods using fluorescent proteins require advanced post-processing and are time consuming. Use of Alexa Fluor dyes (other than Alexa Fluor 647) was not possible at the time because lasers available in our lab were not powerful enough for the dyes to be randomly activated. Using two organic dyes has also an another potential issue connected to an imaging buffer. Imaging buffer is usually optimised to ameliorate performance of a single dye. This means that for two dye imaging it is going to perform better for one dye than the other. Lastly, the PAINT method might suffer from a higher background than methods with stationary fluorescent

probes [Giannone et al., 2008].

The varying quality of images acquired with different super-resolution techniques underlines that the sample preparation, imaging and post-processing have a direct impact on the accuracy of the quantitative measurement and interpretation of results. In localisation microscopy an important sources of error is the labelling error caused among others by the method used to tag proteins. The primary/secondary antibody labelling separated the organic dye from the protein [Tanaka et al., 2010]. The organic dyes size is usually $\sim$0.5 nm [Resch-Genger et al., 2008] and the combined length of the two antibodies is $\sim$10 nm [Ries et al., 2012]. This means that the position of the protein can differ up to 9 nm from the localisation [Shivanandan et al., 2014]. For this study the proteins of interest were stained with tandem dye pairs in primary/secondary antibody labelling which could potentially introduce even bigger distance between the protein of interest and the probe. Another issue is that although the monoclonal primary antibody used for staining does attach to a specific location on the protein, the information about the exact attachment spot is not available. Additionally, in primary/secondary antibody labelling there can be clustering artefacts, as more than one secondary antibody can attach to each primary antibody and more than one fluorophore can be attached to a secondary antibody. Lastly, each localised molecule position is estimated using information delivered by photons coming from the molecule. Thus, the molecule position is estimated with uncertainty caused by a limited number of photons detected. This uncertainty is proportional to a square root of the number of photons detected. Assuming a Gaussian model for the uncertainty, the standard deviation of each localisation is given as: $\sigma = \frac{\sigma_{PSF}}{\sqrt{N}}$ [Thompson et al., 2002]. For Alexa Fluor 647[1] the uncertainty is $\sim$ 12 nm when using the Nikon STORM system.

Use of fluorescent proteins expressed directly by the protein of interest can remove the issue of the labelling error due to minimising the distance between the detected fluorescent marker and the protein of interest. However, the quantum yield of the fluorescent proteins is smaller than that of the organic dyes, resulting in lower intensity and in higher density

---

[1]with FWHM $= 230\,\mathrm{nm} - \sigma_{PSF} = 97\,\mathrm{nm} -$ and considering number of photons detected to be around 100

data sets. Analysis of denser data sets usually requires more time and dedicated analysis methods (for example 3B [Cox et al., 2012]). However, overcoming the drawbacks of the fluorescent proteins provides images with much smaller labelling errors and much more precise reconstruction of the actual sample structure.

Overall, the analysis method presented here can provide an accurate measurement of the absolute and relative protein positions in the ring. The accuracy of any quantitative measurement strongly depends on the quality of data. This can be achieved by ensuring that the labelling error is minimised. Here, the labelling error was decreased due to use expression of fluorophore directly to one of the protein of interest. Use of more than one organic dye can lead to even more precise measurement. The main source of localisation error, due to two antibodies to attach fluorescent tag, would be removed by directly expressing fluorophores into the desired protein for the cost of lower photon count, impaired biological function or overexpression.

3.5.2 3D SIM images of live cells displaying podosomes

Podosomes consist of three separate structures called the core, ring and cap. These three structures are thought to have a distinctive position in the podosome structure. However, little is known about 3D podosome architecture during podosome assembly and turnover. Observation of physiological processes in cells requires live cell imaging. Structured illumination microscopy (SIM) is a wide-field technique enabling fast imaging of live cells. It also provides a resolution improvement when compared with wide field methods. However, it main advantage over other super-resolution techniques is its acquisition speed, which is much faster. Additionally, SIM optically sections the image and provides a stack of images with different vertical positions (these images can later be used for 3D reconstruction of the imaged object).

This study focused on the 3D arrangement of the podosome ring and core. Podosome samples were prepared using THP-1 cells by Elizabeth Foxall, a PhD student in Gareth Jones group (see Appendix C). Two proteins were imaged. Talin from the podosome ring was transfected with mCherry-talin construct and WASP-interacting protein (WIP)

present in the podosome core was transfected with GFP-WIP. Samples were imaged using Nikon N-SIM super-resolution system with Nikon objective 60x (water, N.A. 1.27) and an EM-CCD camera (Andor Ixon) was used to collect images. Eleven sections were taken to create 1.2 μm z-stack (with step size 0.12 μm). The exposure time was set to 400 ms for GFP and 300 ms for mCherry (because of its faster bleaching rate). The super-resolution images were reconstructed using the Nikon software with raw images captured with fifteen angular grid pattern positions. An example of reconstructed z sections is shown in Figure 3.14 and 3D reconstruction is presented in Figure 3.15.

The resulting images did not provide desired resolution improvement for creating clear 3D images of podosomes. 3D reconstructions of podosomes displayed a very thick ring surrounding the core – in reality the podosome ring is a flat construct (see Figure 3.1, [Wiesner et al., 2014]). The poor quality of the resulting images was caused by a number of reasons. Firstly the expression levels of fluorescent proteins in the sample were variable – only a number of cells were bright and visible in two colour channels. The mCherry-talin transfection was performed after the WIP-GFP. The mCherry-talin construct was introduced after the WIP-GFP and not all of the cells were expressing it. Additionally, the sample was very dim, which was addressed by elongating the exposure time with smaller laser power. However, the longer acquisition time elongated total acquisition time up to 6 minutes, which introduced drift to the system. Such a long acquisition time is comparable with the acquisition time of localisation microscopy providing much better resolution. Use of higher laser power was also limited due to very fast bleaching of the cells. An additional problem was caused by instability of the system calibration. The SIM system is usually used at room temperature and when the heating chamber is activated for live cell imaging, the calibration of the system is affected (the grid pattern position was not stable). The calibration of the system affect its ability to optically section images leasing to out of focus light to be detected. All of those issues were responsible for the high noise levels visible in Figures 3.14 and 3.15.

Imaging of live cell samples is challenging because of a number of factors. Cells need to be placed in optimal conditions (37°C and buffer not affecting cells). The heating elements controlling the temperature in the immediate surroundings of the sample radiate heat,

which can affect metal elements of microscope. The heat introduced expansion affects the calibration of grid pattern in Structured Illumination Microscope so that it had to be calibrated before each imaging session and during imaging. Additionally, local heating also affects optical components of the microscope, for example, the objective. The SIM system in the Nikon Centre at King's College was recently refurbished and is more stable during imaging at 37°C and images of live cell samples displaying podosomes with a better signal-to-noise ratio were taken by members of Jones Group (King's College London). Despite the limited resolution improvement the Structure Illumination Microscopy can still provide an insight into podosome dynamics with 3D images.

(a) +600 μm

(b) +480 μm

(c) +360 μm

(d) +240 μm

(e) +120 μm

(f) plane in focus

(g) -120 μm

(h) -240 μm

(i) -360 μm

(j) -480 μm

(k) -600 μm
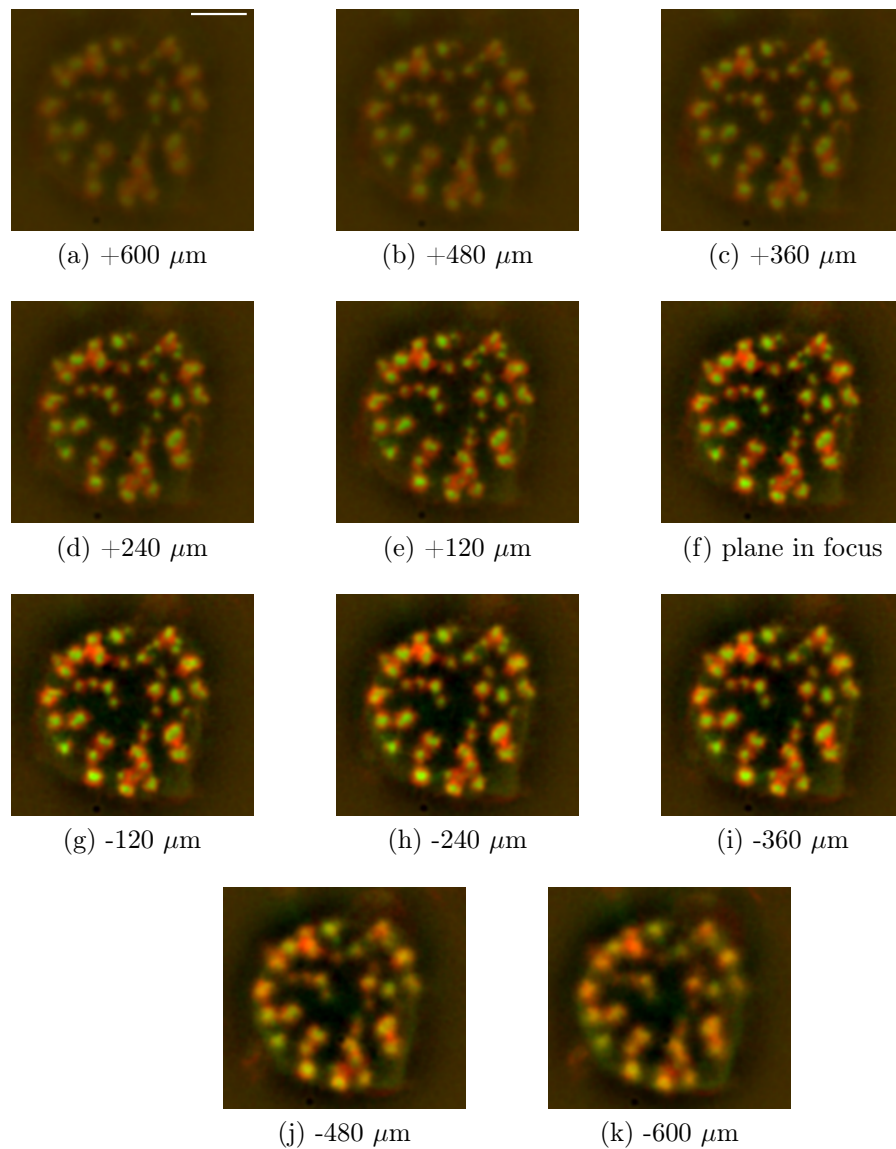
Figure 3.14: Reconstructed structured illumination images of z-sections of live cell displaying podosomes. Cells were transfected with WIP-GFP and mCherry-talin constructs. Eleven sections were recorded with interval 0.120 μm (five sections were taken above the plain of focus indicated with "+" and five below with "-"). Talin is displayed in red and WIP with green. Scale bar 6 μm.

(a)

(b)
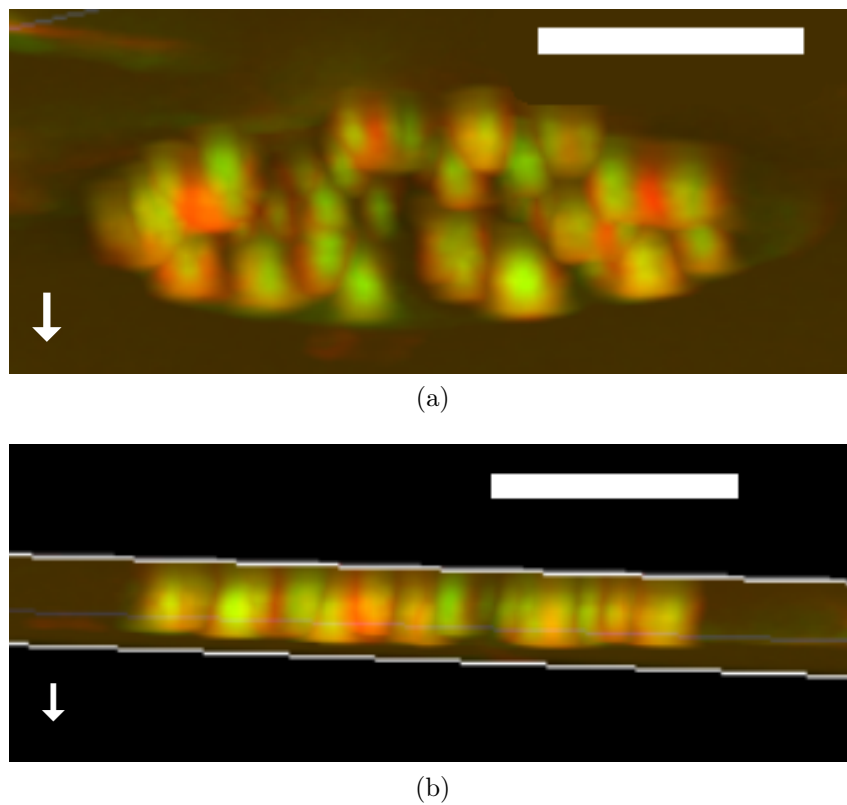
Figure 3.15: 3D reconstruction of z-plane projections from Figure 3.14. a) a xy-plane projection of reconstructed 3D image, b) an z-projection of 3D reconstruction displaying cone shaped podosome cores - marked with green and the ring structure with a shape of red cloud. The x-axis direction is marked with the white arrow. Scale bar 6 μm.

3.6 Discussion and Outlook

Podosomes are formed by a range of cells mainly to facilitate the cell motility. They are small ($\sim 1~\mu$m) structures formed on the the cell surface. Because of their small size, imaging of podosomes in detail was made possible with emergence of super-resolution microscopy. These super-resolution techniques allowed study structure and physiology of the podosomes.

The podosome ring is a complex structure which forms on the cell surface and controls the podosome function. Quantitative analysis and modelling of the podosome ring structure would lead to a better understanding of its role in podosome dynamics. The methodology presented here could be used to build such a model of protein positions. This would include collecting more data for analysis. For example, comparing position of talin with vinculin and then talin with paxillin would help to build a relative model of the ring structure. For a better quality of images using at least one fluorescent protein marking the protein of interest to minimise labelling errors. Also the number of datasets will have to be bigger than presented here and include data from at least a hundred podosomes collected from at least three samples to provide reliable statistics. It is also possible to use software presented here to identify other structured which can be approximated as circular rings (for example the nuclear pore complex). Identification of structures with a shape other than circular is also possible, however, that would include building a new model of the structure of interest.

Chapter 4

Fluorescence anisotropy

4.1 Fluorophore reappearance and clustering artefacts in localisation microscopy

Localisation microscopy provides a great improvement of resolution for imaging nano-scale biological structures. Single molecule localisations can also provide quantitative information about the sample. Usually, quantitative analysis is connected with finding the structure of the imaged sample (for example, using a model shape, see Chapter 3), measuring clustering in the data to characterise the size and/or number of features (see Chapter 2), or counting single particles for stoichiometric purposes (for example in gene expression studies [Raj and van Oudenaarden, 2008]). However, accuracy of such information is affected by two main factors: different sample preparation and analysis methods affect the fluorophore positions which will be identified, and under- or overcounting of fluorophores resulting in differences between the numbers of molecules localised and these actually present in the sample[1].

In particular, fluorophore reappearances can introduce clustering artefacts to localisation microscopy images. This has been discussed in a number of studies [Annibale et al., 2011; Annibale, 2012; Coltharp et al., 2012; Deschout et al., 2014; Sengupta et al., 2011].

---

[1]There are two main reasons for errors in molecule counting: undercounting caused by limited detection efficiency and overcounting – detecting the same fluorophore more than once [Deschout et al., 2014]

Minimising and accounting for the influence of over-counting is crucial to produce an accurate and precise measurement of imaged structures. Many fluorescent molecules have a long lived dark-state, from which they cross to the active state and emit light [Annibale et al., 2011; Annibale, 2012; Deschout et al., 2014]. It has previously been suggested that it is possible to remove multiple appearances in a post-processing step by using a time threshold, discarding localisations from subsequent frames in the area surrounding the first localisation. This threshold needs to be selected manually for each fluorescent probe used and depends on labelling density and localisation uncertainty [Deschout et al., 2014]. Thus, the final results depend on the length of the time threshold used.

Similarly, different methods of sample preparation and analysis are known to influence the resulting appearance of the imaged structure. For example, podosome rings (previously discussed in Chapter 3), the shape detected using fixed labelled and transfected samples is different. In fixed samples where the excitation density of fluorophores is low the ring appears to be a collection of sparse clusters [van den Dries et al., 2013] whereas it appears as a continuous hexagon for fixed cell imaging with encoded fluorescent proteins using high density analysis, 3B [Cox et al., 2012]. The use of antibody labelling may be responsible for clustering artefacts, for example, more than one secondary antibody can attach to the primary antibody [Tanaka et al., 2010]. This suggest that the observed clustering in podosome rings may be an artefact caused either by staining or fluorophore reappearance. However, it is still unclear if the observed clustering corresponds to clustering of the labelled proteins. This chapter discusses design and initial testing of a microscope system which can be used to investigate this phenomenon by detecting an effect of energy transfer between two fluorescent molecules of the same type, allowing areas where fluorophores are in close proximity to be identified.

4.2 Förster Resonance Energy Transfer

Flörster Resonance Energy Transfer (FRET) is a contactless process of energy transfer between two fluorophores [Gell et al., 2006]. FRET results from the interaction between dipoles of two fluorophore molecules. Energy transfer occurs when the emission spectrum

of the donor molecule overlaps with the absorption spectrum of the acceptor molecule and results in light emission from the acceptor molecule. The strength of the FRET interaction depends on the donor-acceptor pair and the distance between them (the strength is proportional to $\frac{1}{r^6}$, the molecules need to be separated by 2-10 nm for FRET to occur) [Lidke et al., 2003].

Förster resonance energy transfer is usually observed for two types of donor-acceptor pairs: these where the donor and acceptor are different (called heteroFRET) and those where the donor and acceptor are the same type of fluorophore (homoFRET or emFRET). HeteroFRET requires the presence of two different fluorescent molecules and it can be detected either by observing an increase of polarisation of the donor, decrease of the fluorescence lifetime of the donor, or the increase in emission of the donor probe after photobleaching the acceptor [Yan and Marriott, 2003]. HeteroFRET can be used to studies of colocalisation or signalling in cells [Clayton, 2009]. HomoFRET can be used for observing interactions of the same type of protein for example, the formation of dimers [Lidke et al., 2005; Yan and Marriott, 2003], or to observe protein congregation and clustering in the sample. HomoFRET does not change the intensity or the lifetime of the detected fluorescence. The fluorescence lifetime remains unchanged because in the energy transfer the excitation energy is transferred between the same types of fluorophore [Yan and Marriott, 2003]. However, it can be detected by observing polarisation changes of the emitted light. This is because an energy transfer can occur between molecules with non-parallel dipoles. Since the dipole direction controls the direction of the polarisation of the emitted light is less polarised than the excitation light (see Section 4.3). The depolarisation caused by homoFRET mechanism is shown in Figure 4.2.

4.3 Fluorescence anisotropy: principles

In general, fluorophores emit fluorescent light with the same polarisation properties as the exciting beam – the light polarization is maintained. This is observed because the fluorophores in the sample will absorb light only if their dipole moment[2] is parallel to the electric vectors of the light beam (in a process is called photoselective excitation). As a result, in an isotropic solution only fluorophores oriented along the direction of the electrical vector of the exciting light beam are excited, and thus only these molecules emit polarised light (see Figure 4.2).

In an experimental situation, some of the polarisation is lost due to, for example, rotation of the fluorescent molecule or energy transfer (this is going to be discussed is section 4.4). The polarisation loss can be accounted for using a measure called fluorescence anisotropy, which is defined as:

$$r = \frac{I_\parallel - I_\perp}{I_\parallel + 2I_\perp},$$ (4.1)

where $I_\perp$ is the intensity of the light detected with polarisation perpendicular to the polarisation of the exciting light (perpendicular polarization channel) and $I_\parallel$ is the detected intensity of light with polarisation direction parallel to the polarisation of the in the parallel polarization channel (parallel polarisation channel). In practice equation 4.1 needs to be corrected to account for any differences in camera detection efficiency between the parallel and perpendicular polarisation channels, with a calibration factor called the G factor. The G factor depends only on the camera properties and is sample independent [Tramier et al., 2000]. This factor is given by:

$$G = \sqrt{\frac{I_{\parallel 1} \cdot I_{\perp 2}}{I_{\perp 1} \cdot I_{\parallel 2}}}.$$ (4.2)

---

[2]The dipole moment (also called transition dipole moment) is an electric dipole moment connected to the transition between an initial and a final state. It is a vector property, the direction of which determines how the system will react with light of a given polarization.

Thus the anisotropy can be calculated using [Lidke et al., 2005]:

$$r = \frac{I_\parallel - GI_\perp}{I_\parallel + 2GI_\perp} \tag{4.3}$$

where G is the scaling factor due to camera properties.

The definition of fluorescence anisotropy can be derived for a single molecule. For the purpose of this definition the absorption and emission moments of this unmoving molecule will be considered as parallel. Consider a fluorescent molecule in an xyz coordinate system (see Figure 4.1). The molecule is positioned at an angle θ to z axis and ϕ to y axis. The electric field created by the molecule excited with polarised light (polarised parallel to the z axis) is given by:

$$E(\theta, \phi) = E_0 \frac{\sin \theta}{r} \hat{\theta}, \tag{4.4}$$

where E_0 is the amplitude, r is the distance between the fluorophore and the unit vector $\hat{\theta}$. The emitted light intensity is proportional to the square of the electric field created by the fluorophore and given by:

$$I(\theta, \phi) = I_0 \frac{\sin^2 \theta}{r^2} \hat{r}, \tag{4.5}$$

where I_0 is the amplitude, $\hat{r}$ is a unit vector in direction of light propagation [Lakowicz, 2006].

Using information about the electrical field $E(\theta, \phi)$ and intensity $I(\theta, \phi)$ of light emitted by a fluorescent molecule we can write equations for projections of the field and intensity to the axis of the coordination system. The z and y components of the electric field are equal to: $E_z(\theta, \phi) = E_0 \cos \theta$ and $E_y(\theta, \phi) = E_0 \sin \theta \, \sin \phi$. The I_z and I_y components of intensity are equal to the square of the corresponding components of electrical field. In this case, since the molecule was excited with light polarised parallel to the z axis, the intensity components can be also written as $I_\parallel$ for z component and $I_\perp$ for y
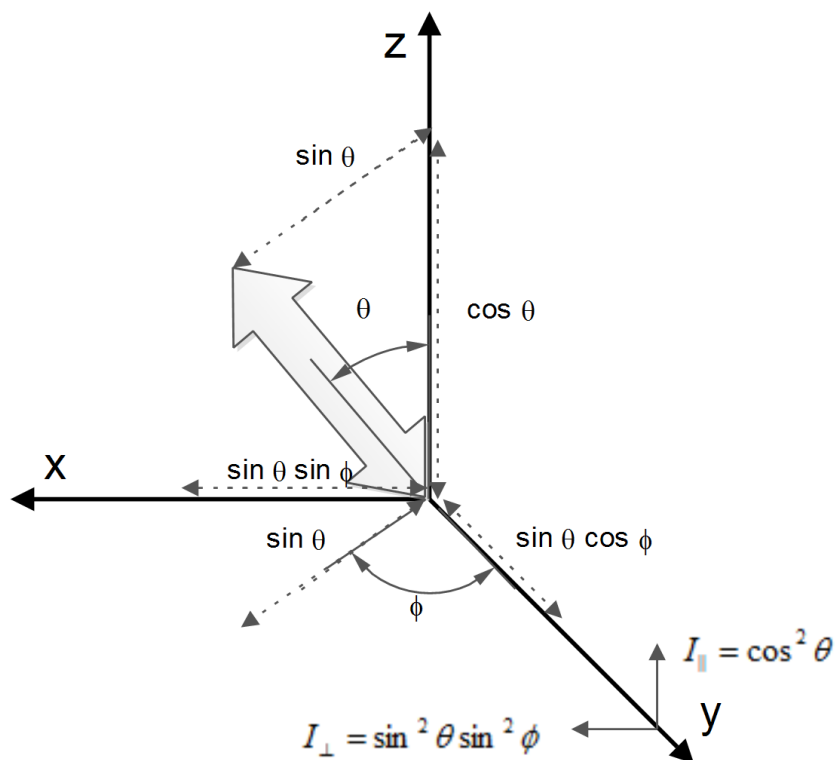
Figure 4.1: A model of a single molecule represented by an arrow in "xyz" coordinate system. The molecule is positioned to z axis on angle θ and to y axis on ϕ. Image based on image from: [Lakowicz, 2006] pp.356.

[Lakowicz, 2006]:

$$I_{\parallel}(\theta, \phi) = I_0 \cos^2\theta \tag{4.6}$$

and

$$I_{\perp}(\theta, \phi) = I_0 \sin^2\theta \, \sin^2\phi. \tag{4.7}$$

In real samples fluorophores are oriented randomly, thus the resulting intensity is an average of intensities from excited molecules. Molecules in the sample are photoselectively excited with light polarised parallel to the z axis, which results in a population of molecules positioned at angle by ϕ to the y axis which are going to be excited. For a set

of randomly oriented molecules the ϕ angle has an equally-probable value in range from 0 to 2π. Thus, we can average over the angle ϕ [Lakowicz, 2006]:

$$< \sin^2 \phi > = \frac{\int_0^{2\pi} \sin^2 \phi d\phi}{\int_0^{2\pi} d\phi} = \frac{1}{2}. \tag{4.8}$$

$I_{\parallel}$ is independent of ϕ, but $I_{\perp}$ can be written as:

$$I_{\perp}(\theta) = I_0 \frac{1}{2} \sin^2 \theta. \tag{4.9}$$

Similarly, the angle θ can be averaged because of photoselective excitation. The probability that a molecule is going to be excited depends on the $\cos^2\theta$ (θ is the angle between the molecule dipole moment and z axis, see Figure 4.1). This results in a population of excited molecules, which are symmetrically oriented around the z axis. In a random sample of fluorophores the proportion of molecules at an angle between θ and $\theta + d\theta$ is proportional to $\sin\theta d\theta$. The probability distribution of molecules excited by single photon polarised light is given as [Lakowicz, 2006]:

$$f(\theta)d\theta = \cos^2\theta \, \sin\theta \, d\theta. \tag{4.10}$$

Using a probability distribution function given by equation 4.10, one can calculate the measured fluorescence intensity [Lakowicz, 2006]:

$$I_{\parallel} = I_0 \int_0^{\pi/2} f(\theta) \cos^2\theta \, d\theta = I_0 2 \int_0^{\pi/2} \cos^2\theta \sin\theta \, d\theta = I_0 < \cos^2\theta > \tag{4.11}$$

and

$$I_{\perp} = I_0 \frac{1}{2} \int_0^{\pi/2} f(\theta) \sin^2\theta \, d\theta = I_0 \frac{1}{2} \int_0^{\pi/2} \cos^2\theta \, \sin\theta \sin^2\theta \, d\theta = \frac{I_0}{2} < \sin^2\theta >, \tag{4.12}$$

where k is constant. Using equations 4.11 and 4.12 the measured anisotropy for this

sample as can be written as [Lakowicz, 2006]:

$$r = \frac{3 < \cos^2\theta > -1}{2}. \tag{4.13}$$

Thus the anisotropy value depends on the average of $\cos^2\theta$. For a single molecule the value of anisotropy measured simply depends on the angle θ. The highest anisotropy $r = 1$ will be measured for $\theta = 0$. However, in practice the measured anisotropy is always smaller than 1, because absorption and emission dipoles are rarely parallel. However, the anisotropy measured from a sample with more than a single fluorescent molecule is usually even smaller due to random distribution of fluorophores and photoselection. To calculate the measured anisotropy for a sample containing a large number of fluorescent molecules, one needs to average $\cos^2\theta$ for angles between 0 and $\pi/2$ using equation 4.10 [Lakowicz, 2006]:

$$< \cos^2\theta >= \frac{\int_0^{\pi/2} f(\theta) \cos^2\theta \, d\theta}{\int_0^{\pi/2} f(\theta) \, d\theta} = \frac{2 \int_0^{\pi/2} \cos^2\theta \, \sin\theta d\theta}{\int_0^{\pi/2} \cos^2\theta \, \sin\theta d\theta} = \frac{3}{5} \tag{4.14}$$

The maximum possible measured fluorescence anisotropy for a sample labelled with many fluorophores can be calculated using the average value of $\cos^2\theta$ (as calculated in equation 4.14) and equation 4.13:

$$r_{MAX} = 0.4. \tag{4.15}$$

In practice observing a value of exactly 0.4 is unusual – it is only possible when there are no processes lowering the polarisation in the sample (for example Brownian motion, homoFRET, light scattering, or due to use of high N.A. optics in the microscope) [Lakowicz, 2006].

Fluorescence anisotropy and photoselection has also been used to acquire super-resolved images. The principle is similar to that of localisation microscopy, but using photoselective activation of fluorescent probes with polarised light to randomly activate fluorophores [Hafi et al., 2014]. This means that any type of fluorophore can be used for

imaging with this technique. To activate all molecules in the sample the polarisation of the exciting light beam is changed over time. The resolution achievable with this technique was reported to be ~ 50 nm. However, it is very challenging to image biological structures using polarisation nanoscopy, as any rotational diffusion will hinder the imaging [Hafi et al., 2014].

4.4 Practical measurements of fluorescence anisotropy

There are two main processes which can be detected by a decrease in the anisotropy value – rotational diffusion and homoFRET. Rotational diffusion causes spinning of the fluorescent molecules which changes the direction of their dipole moments and results in emission of light of random polarisation (light from each single molecule is still polarised but the direction of its electrical vector is random) – this behaviour is shown in Figure 4.2 [Lakowicz, 2006; Lidke et al., 2003]. The measured anisotropy value strongly depends on the rotational speed of the molecules – for higher speeds the measured anisotropy is lower [Gautier et al., 2001; Keating and Wensel, 1991; Lakowicz, 2006; Tramier et al., 2000] and can be used for measurements of rotational diffusion (see Appendix H). Time resolved fluorescence anisotropy measurements of rotational diffusion can be used to provide information about viscosity in of the sample and binding or denaturation of molecules [Siegel et al., 2003]. Rotational diffusion studies have also been used to provide an estimation of polarisation loss and subsequent lowering of anisotropy value in the microscopy system due to the use of a high N.A. objective.

Förster energy transfer between molecules decreases the anisotropy value (i.e. the presence of FRET causes decrease in anisotropy) [Yan and Marriott, 2003]. Although in FRET the donor molecule is photoselectively excited with polarised light[3] the acceptor molecule is oriented randomly[4]. Thus, light emitted by an acceptor will not have the

---

[3]The direction of molecule transition moment is parallel to the direction of electrical vector in the polarized light

[4]The direction of transition moment of the acceptor molecule can be random.
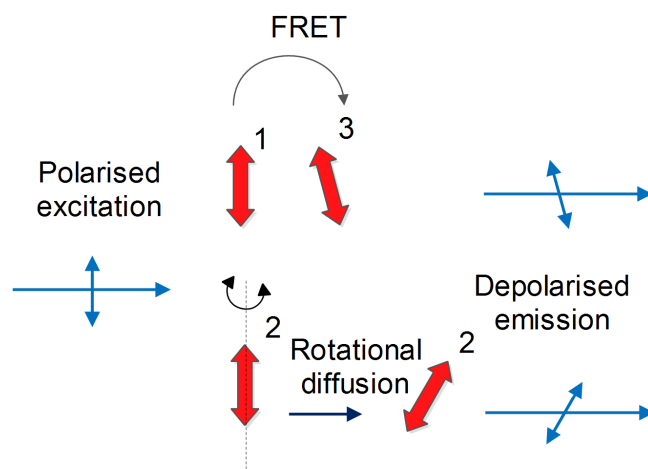
Figure 4.2: Homo FRET and rotational diffusion as factors lowering the anisotropy value. Two molecules in solution (marked 1, 2) have dipole moments oriented according with direction of polarised light (dipole moments marked with red arrows). Activated molecule 1 passes its energy to molecule 3 in the process of fluorescence resonance energy transfer (FRET). Molecule 3 is the same type of fluorophore molecule as 1, but it has different orientation of the dipole moment, which causes a depolarisation to be detected. Other process causing the depolarisation of emitted light is rotational diffusion. Molecule 2 is excited with polarised light and it is rotated before it emits light. The rotational diffusion causes rotation of the molecule dipole moment (in a random direction). This results in depolarised emission of light. Light emitted by the molecule no longer has the same direction of polarisation as of the excitation light. The activation and emission light beam and their polarisation directions are marked with blue arrows.

same polarisation as the light which can be emitted by a donor [Gautier et al., 2001; Lidke et al., 2003]. The mechanism of homoFRET transfer is shown in Figure 4.2.

4.4.1 Fluorescence anisotropy imaging

Fluorescence anisotropy microscope

The fluorescence anisotropy imaging microscopy system was build for this project. It was based on an inverted Zeiss microscope, with a 63x, 1.4 N.A. objective, with a laser light source, wavelength 473 nm, power 50 mW, and 90% light polarisation. The power of the laser light source was controlled using a natural density filter (N.D.). To achieve fully polarised light, the laser beam was transmitted through a linear polariser. A telescopic beam expanding system was used to create a beam with bigger spot size for better illumination (the magnification of the telescopic system was 1.33). Then the light passed through a $\lambda/2$ wave plate which was used to create linearly polarised light with horizontal and vertical polarisation directions. The effect of the $\lambda/2$ wave plate operation depended on the angular orientation of the plate axes to the polarisation direction of linearly polarised light. For a polarisation direction oriented along the axis of the plate the polarisation direction remained unchanged, and for an angle 45° between the polarisation direction and the plate axis, the polarisation direction of the beam was rotated by 90°.

The light was then collected by a tube lens (f=400 mm). The most important part of the set up was the OptoSplit II (Cairn research). This component acted as a polarising beam splitter dividing the light from the sample into two polarisation channels: parallel and perpendicular with respect to the excitation beam (images were spatially translated and appeared in different positions on the camera chip) (see Figure 4.3). Image acquisition was performed using emCCD camera (Cascade II: 512, Photometrics) controlled with Micro-manager 1.4 software[5] with exposure time 0.2 s. The scheme of the microscope system is shown in Figure 4.3. The linear polariser and $\lambda/2$ wave plate positions were adjusted to find the position of polariser which transmits maximal intensity of light and the axis of the half-wave plate.

The fluorescence anisotropy imaging cycle consisted of a number of steps. Firstly the image of the area of interest was captured with horizontal polarization of the exciting light. Then, to calculate the anisotropy for each pixel in the image the two sub-images

---

[5]Freeware software based on ImageJ, documentation available at http://valelab.ucsf.edu/ MM/MMwiki/
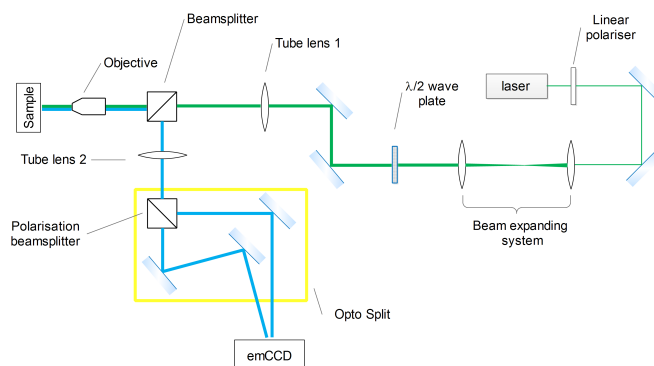
Figure 4.3: The fluorescence anisotropy imaging system. The laser light λ=473 nm (power of the laser light was adjusted by addition of a N.D. filter) was transmitted via a linear polariser to create linearly polarised light (laser was 90% polarised before this step). The light beam was expanded using a telescope two lens system (focal lengths f_1=75 mm and f_2=100 mm, resulting magnification 1.33). The expanded light beam passed through a $\lambda/2$ plate which could change the direction of polarisation by 90°. The light was passed to the microscope via the tube lens 1 (f=400 mm). An inverted Zeiss microscope with objective (63x, N.A. 1.4) was used for imaging (the tube lens 2 is a part of the microscope). Fluorescent light from the sample (marked with blue) was transmitted to OptoSplit system which consisted of a polarising beam splitter separating the fluorescent light into parallel and perpendicular polarisations. The light with different polarisations was then reflected by a system of mirrors inside of the OptoSplit and imaged by the camera chip (emCCD camera, Photometrics).

(images of the two polarization channels) were aligned. This adjustment was done using code provided by dr Fox-Roberts. The code interpolated one of the sub-images to the position of another using a bright group pixels or small object from the top right corners of the sub-images and corrected for the rotation of sub images. The alignment of the sub-images is important in fluorescent imaging as imperfect alignment can lead to detecting a incorrect anisotropy value [Siegel et al., 2003]. The G factor, which characterizes the difference in detection of different light polarisation was calculated using equation 4.2.

The G factor was calculated separately for each imaging session, since the positions of the sub-images on the camera chip change slightly over time (the positions of the sub-images was controlled via adjustment of the mirror system in the OptoSplit). Finally, the anisotropy value was calculated using equation 4.3.

Polarisation optics adjustment

The linear polariser and $\lambda/2$ wave plate were adjusted to provide light with optimal polarisation properties. The linear polariser was used to achieve linear polarisation of the laser light. The $\lambda/2$ plate was used to control the direction of the light polarisation (depending on the orientation of the plate the direction polarisation of light passing through the plate, can be either be unaffected or rotated by 90°). To check the correct operation of the system the linear polariser in the system was rotated and the light intensity was measured for $\lambda/2$ wave plate set to pass light with unchanged polarisation direction or rotated by 90°.

The measured light intensity varied sinusoidally with the input polarisation. A sine function was fitted to find the maxima points of the intensity profiles:

$$y = y_0 + A \sin\left(\pi \frac{x - \alpha}{p}\right) \tag{4.16}$$

where A is amplitude of the sine function, α is the phase shift, y_0 is offset, and p is half of the period of the function. The point for which the fitted sine function has its maximum can be calculated using the equation:

$$x_{MAX} = \alpha + \frac{p}{2}, \tag{4.17}$$

where α is the phase shift and p is the half of the periodicity value.

The linear polariser was placed in close proximity to the laser light source (see Figure 4.3) and mounted in Thorlabs mount with angular scale (see appendix I). The intensity of light detected in both channels changed along with the polariser rotation. The resulting intensity profiles are shown in Figure 4.4. The sample used for imaging was 2 ml

of 1 μM solution of Rhodamine 6G in Ethanol (low viscosity of ethanol, around 1 cP, provided an isotropic sample, see section 4.4). The exposure time was set to 50 ms and no N.D. filter was used. This measurements were performed without the $\lambda/2$ plate in the optical system.
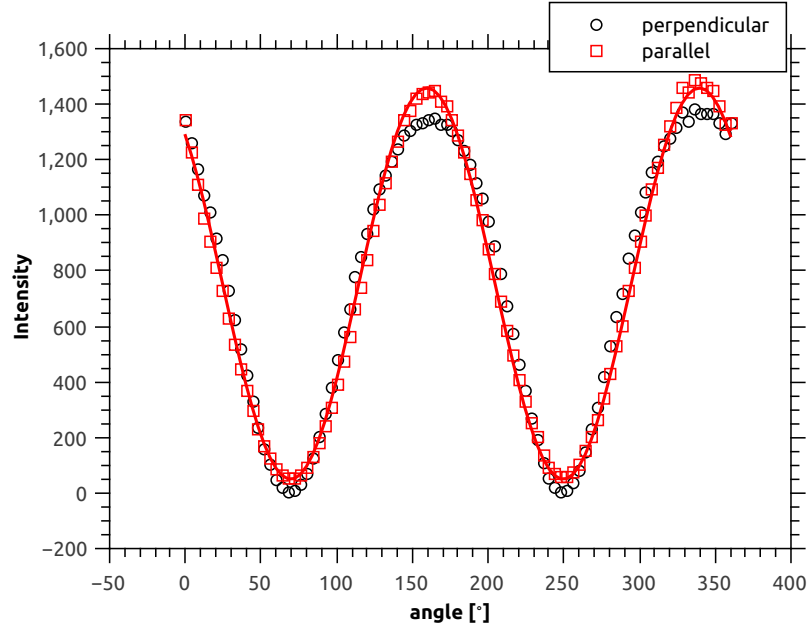
Figure 4.4: Light intensity variation for parallel and perpendicular polarisation channels measured when rotating the linear polariser (imaging with 63x, N.A. 1.4 objective, without $\lambda/2$ plate in the system). The sine function (equation 4.16) was fitted to the intensity distribution (because the shapes of the fitted functions were very similar, for simplicity only the function fitted to the parallel intensity distribution is shown, marked with red line). The fitting parameters of both fits are presented in table 4.1.

Table 4.1: Properties of the sine function (equation 4.16) fitrd to the light intensity profiles detected for linear polariser rotation.

	y_0	A	α	P	R^2	χ^2
parallel	754±3	705±4	114.55±0.17	89.99±0.14	0.9974	640
perpendicular	763±7	679±9	114.89±0.45	89.00±0.35	0.9826	4000

The optimal angular position of the linear polariser was measured as the first local maximum of the intensity profile (calculated using equation 4.17, see table 4.2). The first

Table 4.2: The first intensity maxima for parallel and perpendicular polarisation directions for linear polariser calibration. Values were manually read from the intensity distribution and calculated using equation 4.17.

	manual read [°]	calculated [°]
parallel	172	160
perpendicular	172	159

maximum calculated for the fitted sine function was used to set the linear polariser to the optimal position for providing the highest light intensity with linear polarisation.

After calibrating the linear polariser, the $\lambda/2$ wave plate was introduced to the system (see Figure 4.3 and placed in a ThorLabs mount with 360° scale, see appendix I). To find the axes of the plate it was rotated, and the light intensity profile captured in the parallel and perpendicular polarisation channels. A clean slide was used as a sample, a 2.0 N.D. filter was used to decrease the intensity of the laser light, and the exposure time was set to 10 ms.

Table 4.3: Properties of the sine function fitted to the light intensity profiles detected when rotating the $\lambda/2$ plate. The sine function fitted to the data is given by equation 4.16).

	y_0	A	α	P	R^2	χ^2
parallel	114±1	88±1	114.64±0.30	45.11±0.11	0.9732	100
perpendicular	62±1	-46±1	115.87±0.53	45.18±0.19	0.9227	90

The $\lambda/2$ plate had to be set in two positions providing an unchanged and rotated polarisation directions with respect to the polarisation direction of the laser light. To find these positions the first local maximum was estimated and calculated for the fitted sine function using equation 4.17 for the light intensity distribution (see table 4.4).

The polarisation of light does not change its direction when passed through the $\lambda/2$ plate, when the optical axis of the plate is parallel to the polarisation direction. The $\lambda/2$ plate
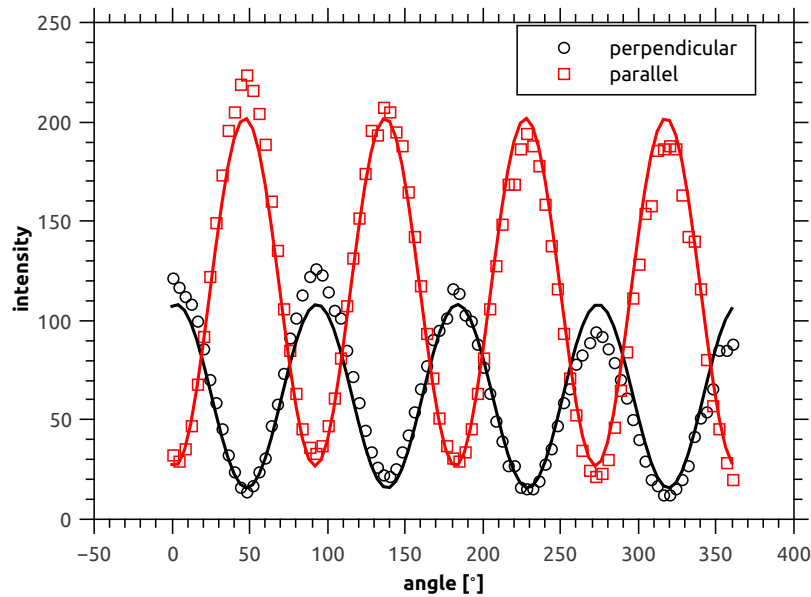
Figure 4.5: Light intensity variation for parallel and perpendicular polarisation channels measured for different values of rotation of the $\lambda/2$ plate. The sine function (equation 4.16) was fitted to the parallel and perpendicular intensity distributions – marked with red and black lines respectively. Fitting parameters are presented in table 4.3.

rotates the polarisation direction by 90° when its optical axis is positioned at a 45° angle to the polarisation direction. Thus the difference between two subsequent maxima of parallel and perpendicular channels is 45°. For the microscope system in our lab the difference between the two channels is 44° and for the fitted sine function is 46. The difference between the theoretical and measured values is probably due to a small polarisation loss. Also the angular scale on the plate mount was given with 2° precision. The calculated angular positions were then used to set the $\lambda/2$ plate to control the polarisation direction of polarised light.

Lastly, the correct operation of the system was validated by testing the light intensity detected for two polarisation directions of the excitation light (horizontal and vertical created with the $\lambda/2$ plate) when the linear polariser in the system was rotated. The $\lambda/2$ plate was positioned to rotate the polarisation direction by 90° and then to transmit light with polarisation direction unchanged (these two positions are marked as rotated

Table 4.4: The first local intensity maxima for parallel and perpendicular polarisation directions for the $\lambda/2$ plate calibration. Values were manually read from the intensity distribution and calculated using equation 4.17.

	manual read [°]	calculated [°]
parallel	48	47
perpendicular	92	93

and unchanged). For each of the polarisation directions of the excitation light the linear polariser was rotated by 360° with increments of 2°. The sample was made with 3 ml of 1 μM solution of Rhodamine 6G in ethanol and placed in a glass dish (μl-Dish, Ibidi®). Exposure time was 20 ms and no N.D. filter was used.

Table 4.5: Properties of the sine function fitted to the light intensity profiles detected in perpendicular and parallel polarisation channels. The $\lambda/2$ plate was positioned to rotate polarisation direction by 90° (denoted by R), and to transmit light with polarisation unchanged (denoted by U). A sine function was fitted to the data (given by equation 4.16).

	y_0	A	α	P	R^2	χ^2
parallel R	511±2	475±3	22.17±0.34	90.42±0.18	0.9958	470
perpendicular R	574±2	576±3	23.69±0.25	89.92±0.13	0.9977	380
parallel U	518±5	478±6	111.83±0.43	88.26±0.33	0.9837	1900
perpendicular U	530±4	489±6	111.60±0.42	88.47±0.32	0.9847	1800

The points for which the fitted sine function reaches the first local maximum were calculated using equation 4.17 and are shown in Table 4.6. The maximum intensity of light measured without and with the $\lambda/2$ plate (set to not change the polarisation direction of excitation light) had similar values. This confirms that the $\lambda/2$ plate is operating correctly in the system. The difference between the positions of first local maximums for rotated and unchanged polarisation directions is between 87 and 89°. This is very close to the theoretical value of 90°for polarisation change using the $\lambda/2$ plate.
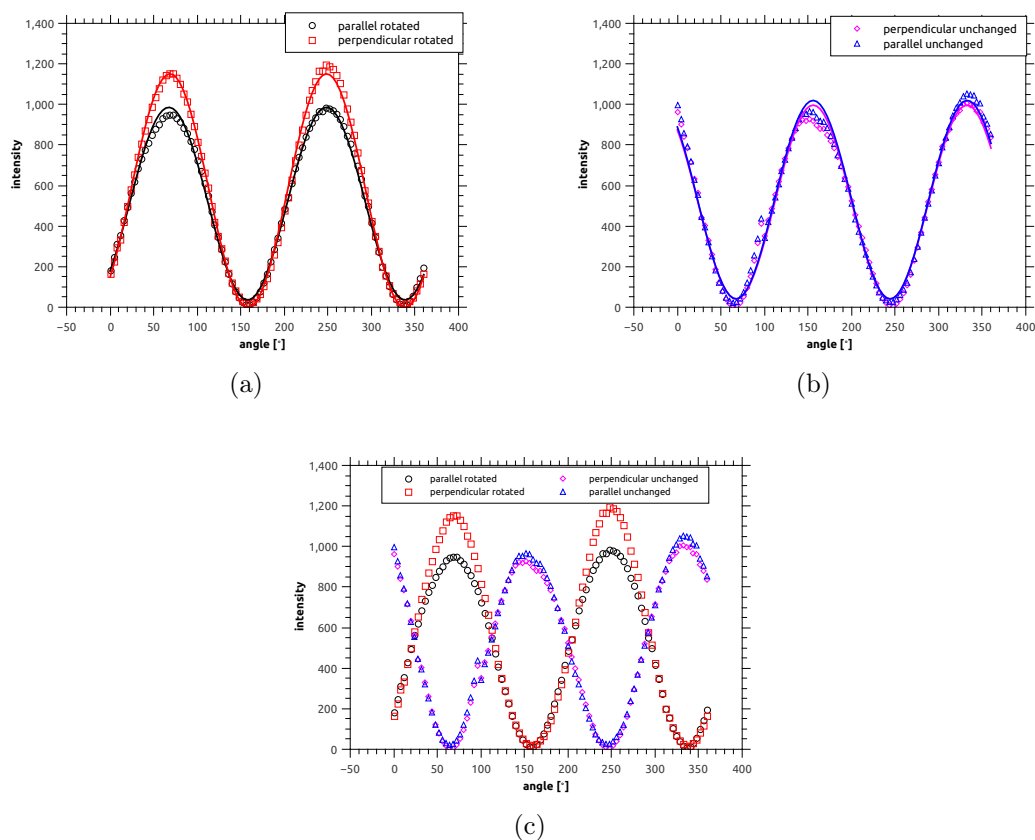
Figure 4.6: A validation test to confirm correct operation of polarisation optics. The linear polariser in the system was rotated by 360° and the $\lambda/2$ plate was set to two positions in which the polarisation direction of the activation light was a) rotated by 90° (marked as rotated) and b) unchanged (marked unchanged) polarisation direction. a) and b) Sine functions (equation 4.16) were fitted to the intensity distributions, the fitting parameters are presented in table 4.5. Fitted sine functions are displayed as: parallel rotated with black line, perpendicular rotated with red line, parallel unchanged with blue line, and perpendicular unchanged with magenta line. c) Intensity variation for linear polariser rotation for rotated and unchanged polarisation directions.

Anisotropy–viscosity calibration

Anisotropy viscosity calibration measurements were performed to estimate the polarisation loss due to the use of high N.A. optics in the system. The influence of the rotational diffusion was measured using a viscosity model sample. Thirteen water/glycerol solutions

Table 4.6: The first local intensity maximum of polarised light measured using $\lambda/2$ plate to control polarisation direction of excitation light while rotating the linear polariser by 360°. Values were manually read from the intensity distribution and calculated using equation 4.17. The position of the first local maximum of intensity corresponds to the values estimated and calculated for polariser rotation (see table 4.2), confirming correct operation of the polarisation system.

	manual read [°]	calculated [°]
parallel rotated	70	67
perpendicular rotated	70	69
parallel unchanged	150	156
perpendicular unchanged	150	156

were prepared with different concentrations of glycerol (24, 37, 49, 56, 65, 74, 83, 84, 87, 89, 92, 93, and 94%) and 1 ml of 10 μM Rhodamine 6G (in water) was added to each solution. The viscosity was found for each of the concentrations using the concentration-viscosity table from [Sheely, 1932]. Measurements of the fluorescence anisotropy were performed with the FluoroMax-4, a spectrofluorometer built without lens components (Horiba), and using the fluorescence anisotropy microscope system.

The anisotropy measurements using the FluoroMax were performed for 2 ml of glycerol solutions placed in a 3 ml cuvette (integration time was set to 0.5 s). The anisotropy calculated for the FluoroMax measurement are shown in Figure 4.7a. The anisotropy-viscosity calibration for the microscope system (using a 63x, N.A. 1.4 objective) was performed with a micro-well plate. Each glycerol Rhodamine 6G solution was placed in three wells of the plate. The mean intensity values from the images of three wells images were used to calculate the anisotropy for each solution. The G factor was also calculated. The anisotropies for each solution were then plotted against the viscosity (see Figure 4.7b). The viscosity of the solution is a factor limiting rotational diffusion of the molecules as discussed in section 4.3. Thus for higher viscosities a higher anisotropy values should be

measured. The first order exponential decays were fitted using a function:

$$y = y_0 + A \cdot \exp(-x/t), \tag{4.18}$$

to the experimental data using QtiPlot 0.9.8.8 software.

(a)   (b)
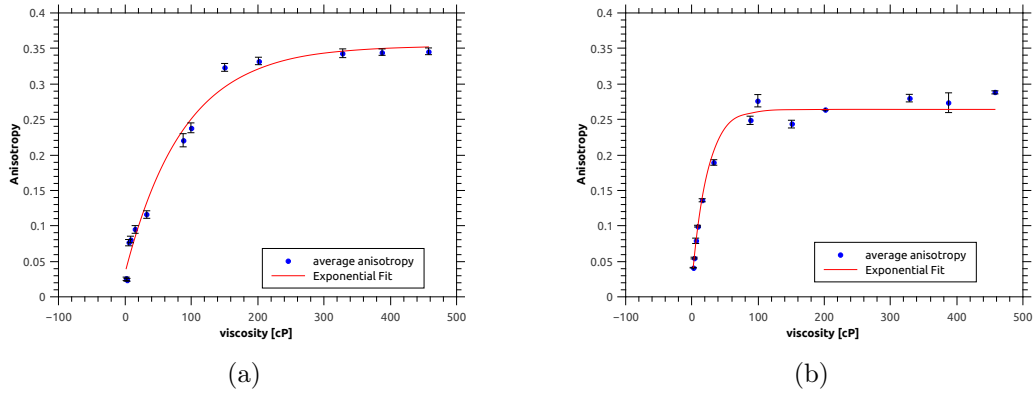
Figure 4.7: Fluorescence anisotropy measurement for Rhodamine 6G glycerol solutions with increasing viscosity. a) Fluorescence anisotropy measurement made using spectrofluorometer. b) Imaged with anisotropy microscopy system, with 63x, N.A. 1.4 objective. Fitted exponential decay function is given with equation 4.18.

Table 4.7: The fitting parameters for the viscosity-anisotropy calibration.

	FluoroMax	microscope system
y_0	0.354 ±0.012	2.6408±0.0014
A	-0.322±0.013	-2.4281±0.0034
t	90±10)	2±2
R^2	0.98422	0.99890

The anisotropy value should change for solutions with different viscosity. For higher viscosities the anisotropy value is expected to be higher (because molecules in the solution are rotating more slowly), as was discussed in section 4.3. The anisotropy-viscosity dependency was expected to be exponential [Lakowicz, 2006]. Measurement of anisotropy for a given sample can be then used to find the rotational speed of the molecules in the solution

with the Debye-Stokes-Einstein model[6]. The anisotropy values for the high viscosity Rhodamine solutions were very similar to the theoretical value of maximal anisotropy (0.4). The anisotropy measured with the microscopy system was significantly lower than the one measured with the spectrofluorometer. Quantification of polarisation loss is an important part of the calibration process and can be used to provide a guidance for the error for microscope ran anisotropy imaging.

Bleaching and its influence on measured anisotropy value

Photobleached samples exhibit an increase in the fluorescence anisotropy value compared with non bleached state. Photobleaching decreases the number of molecules which are available as homoFRET acceptors in the sample. Thus the depolarisation caused by homoFRET is decreased (homoFRET and its effect on anisotropy were discussed in section 4.2). Several studies reported similar anisotropy increases after a portion of the available homoFRET acceptors was photobleached [Cottet et al., 2013; Varma and Mayor, 1998; Yeow and Clayton, 2007].

Here, the photobleaching and simultaneous anisotropy measurements were used to confirm the correct operation of the fluorescence anisotropy imaging system. Three samples were prepared using different concentration of Rhodamine 6G in ethanol (10 μM, 100 μM and 1 mM) and left to dry on clear glass slides. The imaging was performed for 10 arbitrary selected areas on the sample surfaces (20 ms exposure). The anisotropy was calculated for every image in the time series (see Figure 4.8a-c). An average anisotropy value increase due to photobleaching is presented in Figure 4.8d.

Three samples were examined, which displayed similar increases in anisotropy values after photobleaching. For all imaged samples the anisotropy values rose by approximately 30% after bleaching. The measured increase in the anisotropy values confirmed that our optical system is sensitive enough to indicate the presence of homoFRET.

---

[6]The Debye-Stokes-Einstein model of the rotational diffusion quantifies the rotational velocity of molecules and is given by: $\theta = \frac{4\pi a^3 \eta}{3kT}$, where η is the viscosity of the solvent, a is the radius of the molecule, k is the Boltzmann constant, and T is the temperature [Siegel et al., 2003]
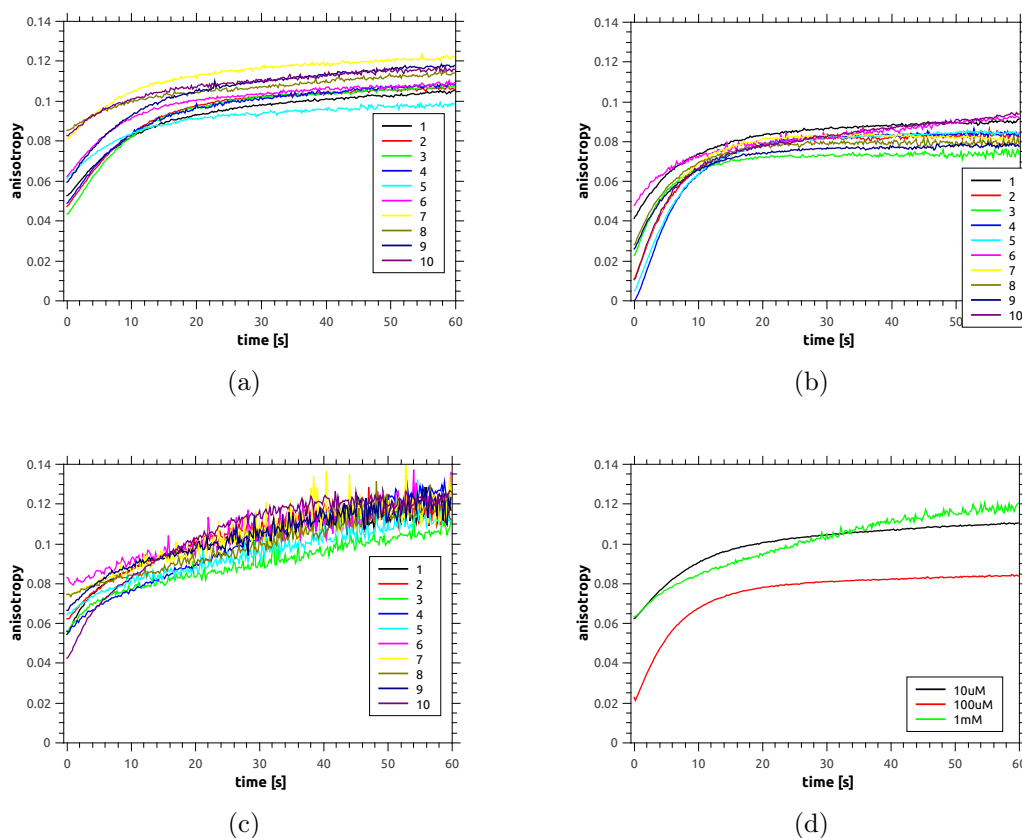
Figure 4.8: Anisotropy value increase caused by photobleaching of the homoFRET acceptors. Ten regions were with a uniform intensity were imaged. Anisotropy was calculated each region for a) 10 μM, b) 100 μM, and c) 1 mM concentration of Rhodamine 6G. d) An average anisotropy gain during the photobleaching.

4.5 Epithelial cell-cell junctions anisotropy maps

Epithelial cell-cell junctions play an important part in many biological processes, for example embryonic development, tissue homoeostasis, and inflammation. Disruption of the dynamics of cell-cell junctions has been linked to illness and cancer [Morton et al., 2013]. The cell-cell junctions are contain transmembrane proteins which form and control the junction. An example of these proteins is, the Coxsackie- and adenovirus receptor (CAR), a 46 kDa protein, which is present in cell-cell junctions in respiratory epithelial cells. The exact role of the CAR protein in maintaining the cell-cell junctions is not yet fully understood [Morton et al., 2013]. However, it was confirmed that it is important for

the correct operation of the epithelium, for example in respiratory epithelium (knockout of CAR proteins resulted in changes in lung morphology [Cohen et al., 2001; Morton et al., 2013]). Additionally, CAR expression was shown to correlate with progression of epithelial derived cancers (for example lung cancer) [Morton et al., 2013]. Thus understanding the role of the CAR protein in forming the epithelial junctions is very important. We planned to investigate the CAR-CAR interactions in the junctions by detecting homoFRET (using fluorescence anisotropy). If the CAR proteins interact in the junction then the level of homoFRET should be higher in areas of higher concentration.

4.5.1   Sample preparation

Biological samples were prepared by Dr Penny Morton (Parsons Group, Randall Division). The cell line used was Human Bronchial Endothelial Cells (HBEC). Cell were cultured so that they formed cell-cell junctions. The CAR protein which takes part in junction forming was genetically encoded with GFP and the cells were fixed. The samples were mounted with ProLong® Gold (Life Technologies) embedding medium which functioned as bleaching protectant.

4.5.2   Preliminary results – homo FRET and anisotropy decrease

The CAR protein assembly in cell-cell junctions was studied by monitoring the anisotropy value in the junctions. A higher concentration of GFP in the junction results in a higher probability of FRET (since the presence of FRET leads to depolarisation of emitted light, and lowers the anisotropy value). The images of endothelial cells were thresholded prior to the anisotropy calculation, so that only areas where the cell-cell junctions were present were analysed (the light intensity collected from the junctions was the highest in the centre of the junctions, because the most of the tagged proteins congregate in the junction area). This behaviour can be observed in anisotropy maps presented in Figure 4.9. There is a variation of anisotropy values in the junction, but generally the anisotropy value in the junction centre is smaller than value in the edge of the junctions. For each image the mean anisotropy value and standard deviation were calculated (see Table 4.8).
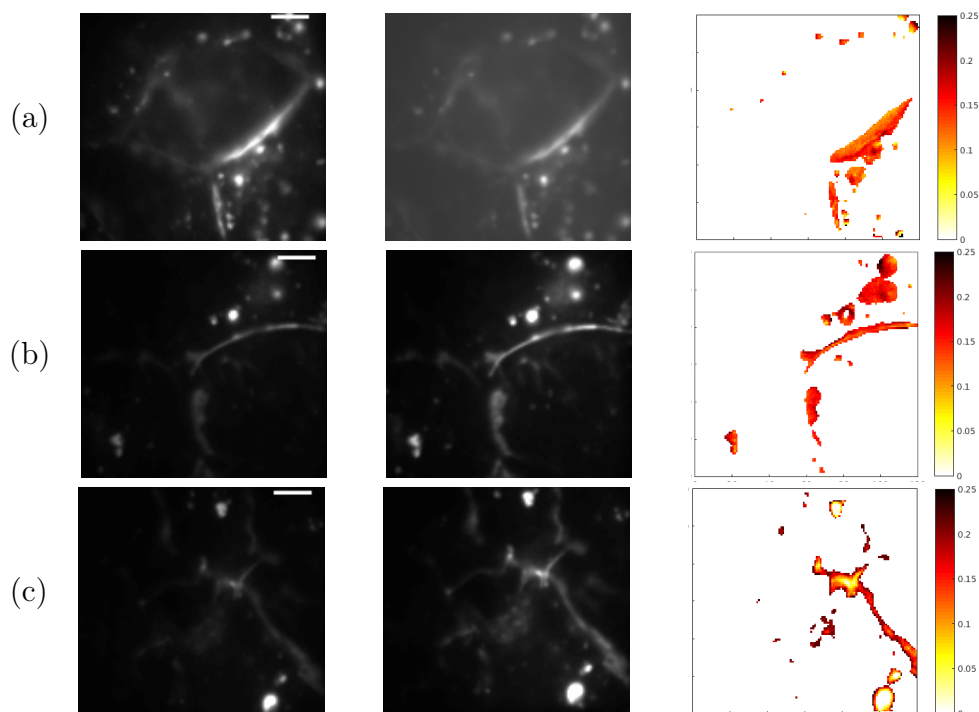
Figure 4.9: Fluorescence anisotropy imaging of the junctions between endothelial cells. Left: Perpendicular polarized light channel. Center: Parallel polarized light channel. Right: The anisotropy map. The image is displayed in false color scale. The colour scale is displayed to the right of each map. Scale bar 20 μm.

Table 4.8: The mean anisotropy values calculated for anisotropy maps which are shown in Figure 4.9.

Figure 4.9	Mean anisotropy value	Standard deviation
(a)	0.117	0.024
(b)	0.152	0.026
(c)	0.149	0.025

4.6   Discussion and outlook

The main purpose of using the anisotropy system was to be able to detect FRET and its effect on polarisation of emitted light. This behaviour was tested using samples prepared with epithelial cells. Epithelial cells form cell-cell junctions which are built with CAR proteins. Here, anisotropy maps of cell-cell junctions were presented, with lower anisotropy values towards the centre of the junctions, where the homo FRET rate is highest. To confirm the actual anisotropy lowering an additional test was planned to be performed on mutant cells not displaying cell-cell junctions. The anisotropy measured for cells without cell-cell junctions should be higher than for the cells without.

In the future the anisotropy values for epithelial cells which form cell-cell junctions will be compared with anisotropy values for mutant cells which are unable to form the junctions. In absence of tagged-protein congregations the anisotropy value would be higher than for the cells not displaying cell-cell junctions, because there is less probability of energy transfer (FRET) between fluorophores.

The anisotropy system has not yet been used with localisation microscopy. The microscopy system (see Section 4.4.1), was equipped with a single light source. 473 nm laser light can be used to excite a number of fluorophores types, however, their blinking properties and high power required make them not ideal choices for localisation microscopy imaging of biological samples. A number of data sets for samples prepared with dried Rhodamine 6G displaying random blinking were collected. However, Rhodamine is not an optimal probe to be used with biological samples of interest in our group. Thus, the microscopy system was equipped with a second light source – 647 nm laser light (see Appendix J). The 647 nm laser source has not yet been joined with polarisation optics (the work on the system expansion are still in progress). 647 nm laser light can be used with Alexa Fluor 647 dye, which was previously used for imaging of biological samples in other projects (for example podosomes and DNA origami see Chapters 2 and 3). An investigation of small scale clustering in podosome rings could help to resolve the issue of their shape: whether it is continuous or a sparse collection of clusters.

Chapter 5

Conclusions and Outlook

Localisation microscopy provides resolution beyond the diffraction limit and recovers super-resolution information using single molecule localisations. The super-resolved, reconstructed images are generally used for qualitative study of various structures. However, the single molecule localisations can be also used to provide quantitative information about the structure of interest. These analysis methods can characterise the shape and size of the imaged structure, leading to a better understanding of biological processes. This work presents two analysis methods for gaining extra information and an approach to validate accuracy of localisation microscopy data.

Clustering analysis is an example of a technique which can be used to analyse localisation microscopy data. In localisation microscopy the super-resolution image is reconstructed using single molecule localisations. This means that the structure of interest is usually preserved as a group of points with density higher than the area surrounding it. Clustering analysis helps to identify such higher density regions and can measure their properties (for example as it was discussed in Chapter 2 the size of the cluster). This work presented a new analysis method based on the Rényi divergence and compared its performance with well established Ripley's H function. In the future the analysis presented here will be extended to measure more information about clustering, for example the number of clusters or an average density of points in the cluster. The clustering analysis will also be used for the analysis of biological data sets with known size such as the nuclear pore

complex or viruses. These could be a used as a test subject because they have relatively similar and known sizes across the samples.

Localisation microscopy images can also be analysed to reveal information about protein colocalisation. As was discussed for clustering analysis, the structures of interest in localisation microscopy are visible as a collection of discrete single molecule localisations. Here, the podosome rings were imaged using two fluorescent probes to mark of different proteins in the podosome ring. Custom software identified podosomes in the images using a model of their structure and calculated the relative distance of different pairs of molecules present in the rings. In the future to create a complete podosome ring model more data sets with different proteins labelled will have to be collected and analysed. Lastly, the pipeline developed here can be used for the analysis of any ring shaped structures present in the localisation microscopy data. It can also be adjusted to include models for structures with different shapes and so could become a more widely useful tool for testing the relative distribution of different proteins.

The quality of the localisation microscopy data is the main factor limiting the accuracy of quantitative measurements. Samples labelled with organic dyes using primary and secondary antibodies often exhibit some degree of clustering, attributed to the fluorophore reappearance or clustering of the secondary antibody on the primary. It is not generally possible to distinguish the cause of the clustering, meaning it is not possible to know whether it is related to the underlying biological structure being studied. Here, a system to perform fluorescence anisotropy was constructed, which should be able to distinguish clustering due to multiple fluorophore reappearances from actual clustering of the fluorophores. Fluorescence anisotropy detects the change between the polarisation of exciting beam and that of the emitted light. One of the processes causing a detectable change in polarisation is Föster energy transfer, which occurs in areas with a higher density of dye molecules. Our preliminary results are promising, and in the future the system is going to be equipped with a more powerful laser source more suitable for localisation microscopy imaging.

Appendix A

Additional results and data sets for clustering analysis
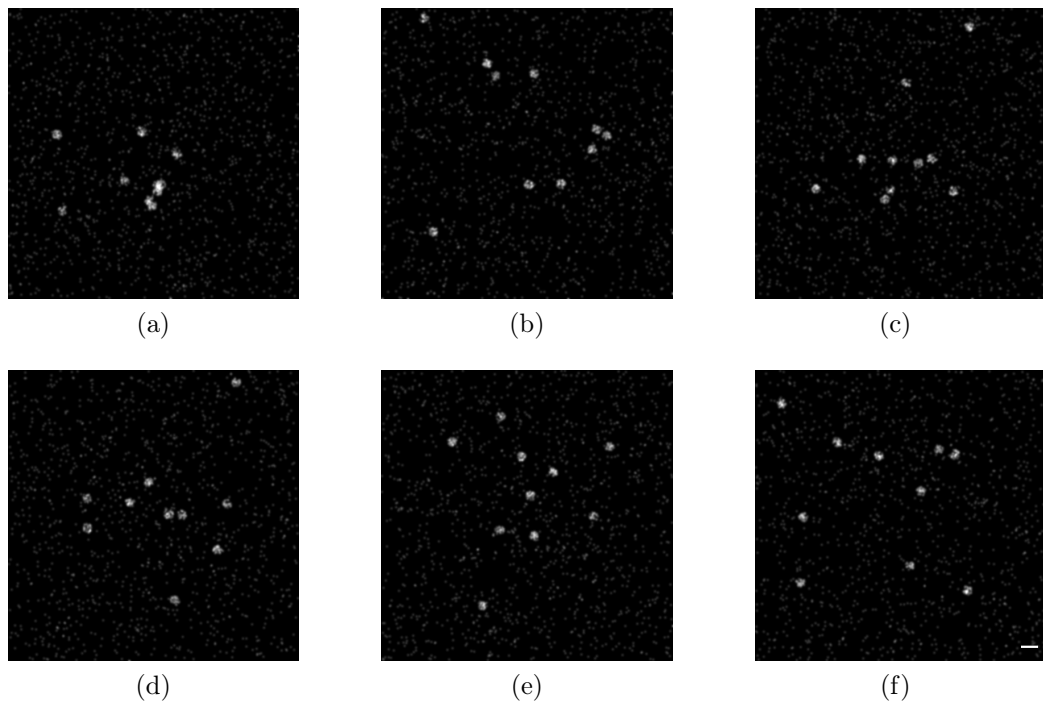
Figure A.1: Datasets used for bootstrap and noise resistance testing of the Rényi divergence and Ripley's H function. Data sets presented here were simulated for Monte Carlo testing, each simulated data set has 10 clusters with radius of 8 pixels and S/N 29 (noise was 0.5% of the number of pixels in the image). For improved visibility the images were blurred. Scale bar: 30 pixels

Table A.1: Statistical parameters for bootstrap testing of cluster radius measurements with the Rényi divergence $\alpha = 70$ for six data sets containing 10 circular clusters with a radius of 8 pixels (S/N 29). The original datasets used for this testing are presented in Figure A.1. The cluster radius value measured for this dataset with the Rényi divergence and Ripley's H function is called the 'original radius'.

No	original radius	mean	median	Standard deviation	quartiles
Rényi divergence					
a	13.20	11.14	11.20	1.56	(10.20, 12.20)
b	8.00	8.17	8.20	0.72	(8.00, 8.60)
c	8.20	8.07	8.00	0.72	(7.80, 8.40)
d	6.80	8.14	8.20	0.82	(7.80, 8.60)
e	7.80	7.99	8.00	0.67	(7.60, 8.40)
f	8.40	8.25	8.40	0.68	(8.00, 8.60)
Ripley's H function					
a	13.80	11.37	11.60	1.34	(10.40, 12.40)
b	10.20	9.75	10.00	0.88	(9.60, 10.20)
c	11.00	9.04	9.20	1.07	(8.40, 9.80)
d	11.40	9.09	9.00	1.29	(8.40, 10.20)
e	9.80	9.21	9.40	0.80	(8.40, 9.60)
f	10.80	9.40	9.40	0.79	(9.20, 9.80)

Table A.2: Statistical parameters for noise resistance testing of cluster radius measurements with the Rényi divergence $\alpha = 70$ and Ripley's H function for six data sets with 10 circular clusters with a radius of 8 pixels (S/N 29). The cluster radius value measured for this dataset with the Rényi divergence and Ripley's H function is called the 'original radius'. This 'original radius' value can be compared with the mean and median values of the results of the noise resistance testing.

No	original radius	mean	median	Standard deviation	quartiles
			Rényi divergence		
a	13.20	12.73	13.0	0.81	(13.00, 13.00)
b	8.00	8.69	8.60	0.19	(8.600, 8.80)
c	8.20	9.06	9.00	0.23	(9.00, 9.20)
d	6.80	8.33	8.40	1.08	(6.80, 9.20)
e	7.80	8.11	7.60	0.59	(7.60, 8.60)
f	8.40	8.87	9.00	0.41	(8.60, 9.00)
			Ripley's H function		
a	13.80	12.98	12.80	0.57	(12.40, 13.60)
b	10.20	11.38	11.20	0.42	(11.20, 11.80)
c	11.00	11.01	11.00	0.58	(10.40, 11.20)
d	11.40	11.50	11.60	0.48	(11.00, 12.00)
e	9.80	11.14	11.00	0.40	(10.80, 11.60)
f	10.80	11.06	11.00	0.36	(10.80, 11.40)

Table A.3: The comparison of the cluster radius measured with the Rényi divergence and Ripley's H function for rectangular and square simulated clusters. For each simulated rectangular cluster the radius was measured with both methods. The percentage difference between the measured radius and the maximal expected radius is given ("-" indicates that the measured radius was smaller than the maximal expected radius, "+" that it was bigger).

		the Rényi divergence $\alpha = 70$	Ripley's H function
16x16	max expected r.	11.31	11.31
	mean	10.68	14.81
	median	10.00	14.80
	st. dev	1.84	0.53
	difference mean	-4.64%	+31.06%
	difference median	-11.50%	+30.97%
16x24	max expected r.	14.42	14.42
	mean	13.85	9.64
	median	12.80	9.60
	st. dev	2.21	1.33
	difference mean	-3.11%	22.66%
	difference median	-11.23%	+24.82%

Appendix B

Probability density functions and their properties

The Gaussian probability density function (pdf) is given as:

$$f(x, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right),$$ (B.1)

where μ is mean and σ is standard deviation. The Logistic pdf is defined as:

$$f(x, \mu, \sigma) = \frac{\exp\left(\frac{x-\mu}{\sigma}\right)}{\sigma\left(1 + \exp\left(\frac{x-\mu}{\sigma}\right)\right)^2},$$ (B.2)

where μ is mean and σ is scale parameter. The logistic pdf is used for growth models and in logistic regression. It has longer tails and a higher kurtosis than the normal distribution pdf. The t Location-scale pdf is given as:

$$f(x, \mu, \sigma, \nu) = \frac{\Gamma(\frac{\nu+1}{2})}{\sigma\sqrt{\nu\pi}\Gamma(\frac{\nu}{2})} \left[\frac{\nu + \left(\frac{x-\mu}{\sigma}\right)}{\nu}\right]^{-\left(\frac{\nu+1}{2}\right)},$$ (B.3)

where $\Gamma(x) = (x-1)!$ is the gamma function, μ is the location parameter, σ is the scale parameter, and ν is the shape parameter. The t Location-scale pdf similarly to the Logistic pdf has heavier tails than the Gaussian. The heaviness of tails is controlled by parameter ν and the t location scale approaches the normal distribution pdf for the

$\nu = \infty$, the smaller values of ν introduce tails to the distribution.

Table B.1: The fitting parameters of the Gaussian, Logistic, and t Location-scale pdfs fitted to the Rényi divergence and Ripley's H function cluster radius measurements results. Parameters of the Gaussian pdf fitted to results of cluster radius measurements acquired with Ripley's H function.

	Function	μ	σ	ν
Rényi div.	Gaussian	8.7500 $\pm(0.0075,0.0075)$	1.2140 $\pm(0.0053,0.0053)$	
	Logistic	8.5089 $\pm(0.0049,0.0049)$	0.4891 $\pm(0.0027,0.0028)$	
	t Location-Scale	8.3673 $\pm(0.0025,0.0025)$	0.2878 $\pm(0.0026,0.0027)$	1.237 $\pm(0.014,0.014)$
Ripley's	Gaussian	10.9004 $\pm(0.0039,0.0040)$	0.6361 $\pm(0.0028,0.0028)$	

Appendix C

Fixed and live cell sample preparation

The THP1 cells (THP1 is a human monocytic cell line derived from a male leukaemia patient) were purchased from ATCC cell bank. Cells were maintained in the Roswell Park Memorial Institute medium (rpmi medium) and supplemented with 10% Fetal Calf Serum. The THP1 cells were seeded on fibronectin (fibronectin is a dimer glycoprotein and can be found in the extracellular matrix. It binds to integrin proteins and the extracellular matrix components, for example collagen or fibrin) coated coverslips (μl-Dish, glass bottom, Ibidi®). To induce the podosome growth the TGF-β1 (Transforming growth factor beta is a protein which controls the cell growth, proliferation, cell specialization (differentiation), and other cell functions) was introduced to the sample (for 16 hours).

Samples were fixed with 3.6% paraformaldehyde in PBS (for 20 minutes). The cells were washed in PBS (3x), permabilised in 0.1% Triton X-100 in PBS (for 3 minutes) and again washed in PBS (3x). After that they were blocked in 3% bovine serum albumin (BSA) in PBS (for 30 minutes), the primary antibody was added (for paxillin Paxillin antibody produced in rabbit, Novus Biologicals, #NBP1-19833 and for vinculin Monoclonal Anti-Vinculin antibody produced in mouse, clone hVIN-1, Sigma Aldrich, #V9131-.2ML, diluted in 3% BSA in PBS for 1 hour) and washed (4x for 5 minutes with PBS). Subsequently, the secondary antibody was added (for paxillin AffiniPure Donkey Anti-Rabbit IgG, #711-005-152, and for vinculin AffiniPure Donkey Anti-Mouse, #715-005-151 from Jackson ImmunoResearch Europe, diluted in 3% BSA in PBS for 30 minutes) and the

sample was washed (4x for 5 minutes). Samples were then stored in PBS in the dark at 4°C.

Live cell preparation is divided into three stages: transfecting and harvesting the virus vector from substrate cells followed by lentiviral transfection of THP-1 cells. The virus for transfecting podosome displaying cells was prepared using the HEK293T cell line (a highly transfectable derivative of human embryonic kidney 293 cells). HEK cells were seeded on a plate and trypsinised. The HEK cell culture medium was added to the plate and the cells were then counted. After spinning down at 1200 rpm for 3 minutes the cells were divided into separate wells into a 12-well plate ($3 - 6 \times 10^5$ cells per well with 1 ml of the HEK cell culture media) and were grown over night. Transfection of cells was performed using the transfection media (in an eppendorf tube, A, prepare 90 μl OptiMEM and 10 μl Polyethylenimine (PEI) and incubate for minimum 5 minutes in room temperature and in a second tube, B, add 90 μl OptiMEM to plasmid mix – see Table C.1). The contents of

Table C.1: The plasmid mix properties.

Plasmid	Stock solution	Mass	μl/well
pD8.91	0.2 μg μl	1.3 μg	6.5
pMD.G	0.2 μg/μl	0.42 μg	2.1
pLNT/Sffv	0.5 μg/μl	1.74 μg	3.5

tube A was added to tube B and incubated for 15 minutes at room temperature. 320 μl of OptiMEM was added to the mix. The HEK cells were washed with 500 μl of OptiMEM and the transfection mix was added to the cells. The cells and transfection mix were incubated at 37°C for 4 hours. After that time the transfecting mix was removed and cells were placed in 1 ml of cell culture media for 48 hours.

The virus for the transfection of the THP-1 cells was extracted from the HEK cells. The HEK cells trasfected in the previous step were transported into falcon tubes containing the fresh cell culture media. Falcon tubes were then spin down at 2000 rpm for 5 minutes. The supernatant from the spin down mixture was filtered to a 50 ml falcon with 2% solution of Virkon (one falcon per infection). Lastly, $\frac{2}{3}$ ml of the Virkon and

supernatant solutions were placed in empty falcons and used it for transfection of THP-1 cells. The remaining virus was frozen (in cryotubes at -80°C).

The THP-1 cells were transfected with the virus acquired in the previous step of the protocol. 10^5 THP-1 cells per transfection in 600-700 μl of media were used. The cells in the media were spun down at 1200 rpm for 3 minutes. The cells were then re-suspended in the new medium in separate wells of a 12-well plate. To each well 1 μl of the polybrene solution was added (solution 4 μg/ml). Around 300-400 μl of the virus was added to each well (so that the final volume in the well added to 1 ml). A well plate with the cells was then incubated at 37°C for 48-72 hours. If the infections were successful (expression of fluorescence proteins in the cells was observed), cells in solution from the well plate were placed into 15 ml falcon with 4 ml of PBS. Falcons were subsequently spin down at 2000 rpm for 5 minutes. The supernatant was removed and was placed in culture flasks (T25) with 4 ml of cell culture media.

Live cell samples were placed on fibronectin coated glass bottomed dishes (Ibidi) and imaged with colourless media to minimise background introduced by the imaging medium. The colourless media were prepared using: 10% Fetal Bovine Serum (FBS), 100 mM B-mercaptoethanol, 10 mM HEPES, and OptiMEM.

Appendix D

Tandem dye conjugation protocol

1. To aliquot the Alexa Fluor dyes, dissolve 1.0 mg in anhydrous DMSO and aliquot again into tubes for a final 0.02 mg amount of dye per tube. For the Cy dyes, dissolve one dye into sufficient amount of anhydrous DMSO to allow distribution into 10 new aliquots.

2. Using evaporator, remove all DMSO.

3. Store aliquots at -20°C.

4. For labelling with an activator-Alexa Fluor 647 pair, dissolve one activator (Alexa Fluor 405, Cy2 or Cy3) aliquot in 10 μl anhydrous DMSO and one Alexa Fluour 647 aliquot in 10 μl anhydrous DMSO.

5. Recommended labelling mixtures:
   Labelling with an activator-Alexa Fluor 647 pair: Mix 50 μl secondary antibody (1.25 mg/ml in PBS) with 6 μl of 1 M NaHCO$_3$, 1.5 μl Cy3 (or 4 μl Alexa Fluor 405 or 5 μl Cy2) and 0.6 μl Alexa Fluor 647.

6. Allow the reaction to proceed for 30 minutes at room temperature, wrapped in foil or otherwise protected from light, on a shaking platform.

7. During the reaction equilibrate NAP-5 gel filtration columns, one per labelling reaction, by running three column volumes of PBS.

8. Bring the reaction volume up to 200 μl with PBS (add $\sim$140 μl) and gently vortex.

9. Add the entire volume to the center of the column.

10. Allow the sample to enter the column and after the last drip, add 550 μl PBS to wash.

11. Add 300 μl PBS and collect the eluent into a 1.5 ml Eppendorf tube.

12. Measure absorbance of the labelled secondary antibody using the UV/Visible spectrophotometer.

Recommended Labelling Ratios

Labelling with an activator-reporter pair. The concentrations of activator dye, antibody, and reporter dye have to be in the ratio:

Activator Dye : Antibody : Reporter Dye = 2.0-3.0 : 1 : 0.6-1.0

The labelling ratios were calculated using antibody, activator dye and reporter dye concentrations. Concentrations c were calculated using the Beer-Lambert's Law:

$$A = \log_{10} \frac{I_0}{I} = \epsilon l c, \tag{D.1}$$

where A is the absorbance, I_0 is the incident intensity, I is the transmitted intensity, ϵ is the molar extinction coefficient, l is the length of solution light passes through (in cm), and c is the molar concentration. For purpose of comparing concentrations of different compounds of the same solution, the Beer-Lambert's Law can be written as:

$$A \propto \epsilon c. \tag{D.2}$$

Concentration of antibody was calculated using corrected value of antibody absorbance (measured absorbance from the antibody needs to be corrected to account absorbance by activator and reporter dyes):

$$
\begin{aligned}
A_{280,Antibody}^{actual} = A_{280,Antibody}^{measured} &- CF_{280,Activatordye} \times A_{\lambda activator}^{measured} \\
&- CF_{280,Reporterdye} \times A_{\lambda reporter}^{measured},
\end{aligned}
\tag{D.3}
$$

where $CF_{280,Activator/reporterdye}$ are correction factors at 280 nm for labelling dyes, $A^{measured}_{\lambda activator/reporter}$ is measured absorbance of the dye. The concentration of antibody is calculated using the actual absorbance of the antibody:

$$c_{Antibody} = \frac{A^{actual}_{280}}{\epsilon_{280}}. \tag{D.4}$$

Concentration of fluorescent dyes can be calculated using equation D.5.

$$c_{dye} = \frac{A_{\lambda MAX}}{\epsilon}, \tag{D.5}$$

where $A_{\lambda MAX}$ is the measured maximum of absorbance and ϵ is the extinction coefficient for relevant dye.

Measurements of absorbance were performed using NanoDrop 1000, a spectrophotometer (Thermo Scientific). For each tandem dye pair characterised three values of absorbance were measured: for activator dye (Cy2 at 488 nm, Cy3 at 560 nm, and Alexa Fluor 405 at 405 nm), reporter dye (Alexa Fluor 647 at 650 nm), and antibody (AffiniPure Donkey Anti-Rabbit and AffiniPure Donkey Anti-Mouse). Calculation of the molar extinction coefficients and correction factors (CF_{280}) for dyes used for conjugation and antibody are presented in Table D.1.

Table D.1: Molar extinction coefficients and correction factors at 280 nm (CF_{280}) for dyes and antibody used for tandem dye conjugation.

	ϵ	CF_{280}
antibody	210,000	
Cy2	150,000	0.15
Cy3	150,000	0.08
Alexa Fluor 405	34,000 μl	0.7
Alexa Fluor 647	239,000	0.03

Appendix E

Reducing-Oxidizing Imaging Buffer recipe for Localisation Microscopy Imaging

The blocking buffer recipe is shown in Table E.1 and it is given in the Nikon N-STORM system Protocol-Sample Preparation manual. Some of the ingredient are complex and were made from scratch in the lab (dilution buffer, MEA, and GLOX) and kept at 4°C – recipes are shown in Table E.1 in comments. Glox solution can be stored up to 2 weeks and MEA for 4 weeks.

According to findings presented in [Dave et al., 2009], the final concentration of COT in the imaging buffer was 2 mM. The recipe for imaging buffer with added COT is shown in Table E.2.

Table E.1: The basic imaging buffer for localisation microscopy imaging recipe.

ingredient	volume	comments
distilled water	800 μl	
dilution buffer	200 μl	recipe: 10mM Tris + 50 mM NaCl
MEA	200 μl	recipe for 1 ml: 77 g Cysteamine (MEA) + 1.0 ml 1M HCl
GLOX	20 μl	recipe for 250 μl: 14 mg Glucose Oxidase + 50 μl Catalase (17mg/ml), + 200 ml Bufer A (10mM Tris + 50 mM NaCl)

Table E.2: The imaging buffer for localisation microscopy imaging.

ingredient	volume	comments
basic buffer	1 ml	made according to recipe from Table E.1
COT solution	250 μl	recipe for $\sim$1 ml: 1ml DMSO + 1.4 μl COT

Appendix F

Cartesian and polar coordinate systems

The polar coordinates are related to Cartesian coordinates as:

$$x = r \cos \phi$$
$$y = r \sin \phi.$$

$$\text{(F.1)}$$

The Cartesian coordinates can be calculated from polar coordinates as:

$$r = \sqrt{x^2 + y^2}$$
$$\phi = \text{atan2}(y, x),$$

$$\text{(F.2)}$$

where

$$\text{atan2}(y, x) = \begin{cases} \arctan(\frac{y}{x}), & \text{if } x > 0 \\ \arctan(\frac{y}{x}) + \pi, & \text{if } x < 0 \text{ and } y \geq 0, \\ \arctan(\frac{y}{x}) - \pi, & \text{if } x < 0 \text{ and } y < 0, \\ \frac{\pi}{2}, & \text{if } x = 0 \text{ and } y > 0, \\ -\frac{\pi}{2}, & \text{if } x = 0 \text{ and } y < 0, \\ \text{undefined} & \text{if } x = 0 \text{ and } y = 0, \end{cases}$$

$$\text{(F.3)}$$

Appendix G

The minimal radius of the protein molecule

Proteins are rigid structures mostly impenetrable to water, with elasticity that can be compared to hard-plastic [Erickson, 2009]. For the purpose of estimating the molecule size one can use an estimation of the protein density as 1.37 $\frac{g}{cm^3}$ [Erickson, 2009]. The inverse value of the protein density, called the partial specific volume or v_2 is usually used for the protein size calculations. The partial specific volume has values between 0.70 and 0.76 $\frac{cm^3}{g}$ (an average value of 0.73 $\frac{cm^3}{g}$ was used for further protein size calculation). Using the partial specific volume v_2, the volume of the molecule can be approximated as:

$$V[nm^3] = \frac{v_2 \times 10_{21}}{N_A} \times M, \tag{G.1}$$

where N_A is the Avogadro constant ($6.023 \times 10^{23} \frac{Da}{g}$) and M is the mass of the the protein, measured in Daltons. The equation G.1 can be written as:

$$V[nm^3] = 1.212 \times 10^{-3} \times M. \tag{G.2}$$

Approximating protein shape to be spherical, the minimal radius of the protein (for minimal volume containing the mass of the protein) can be calculated as:

$$R_{min}[nm] = \left(\frac{3V}{4\pi}\right)^{\frac{1}{3}} = 0.066 M^{\frac{1}{3}}, \tag{G.3}$$

where M is the mass of the the protein, measured in Daltons [Erickson, 2009].

Appendix H

Perrin equations

The anisotropy measurements can be used to calculate the average rotational diffraction speeds using Perrin equations. The anisotropy values indirectly depends on the viscosity of the the sample and is higher for the high viscosity liquids (molecules are rotating slower) than for the low viscosity liquids. The anisotropy loss due to the rotational diffusion is given by the Perrin equations H.1 and H.2 [Lakowicz, 2006].

$$r = \frac{r_0}{1 + (\tau/\theta)} \tag{H.1}$$

or it can be written as displaying direct link to the rotational diffusion [Clayton et al., 2002]:

$$\frac{r_0}{r} = 1 + \tau/\theta = 1 + 6D\tau \tag{H.2}$$

where r_0 is the anisotropy that would be measured in the absence of rotational diffusion, θ is the rotational correlation time for the diffusion process, D is the rotational diffusion coefficient, and τ is the fluorescence lifetime [Lakowicz, 2006; Siegel et al., 2003].

Appendix I

List of optical components used to build Fluorescence Anisotropy System

Following optical components were used for building the fluorescence anisotropy system:

1. ThorLabs

   - Four aluminium mirrors $\varnothing$1", PF10-03-G01, with clear edge mounts KM1CE,

   - Plano-Convex lens f = 75 mm, with no coating, LA1608-A, $\varnothing$1",

   - Plano-Convex lens f = 100 nm, with no coating, LA1509-A, $\varnothing$1",

   - Tube lens f = 400.0 mm, $\varnothing$1", Unmounted, Visible Achromat, AntiReflective-Coating: 400-700 nm, AC254-400-A1,

   - Threaded Manual Beam Shutter, SM1SH1,.

   - Unmounted linear polariser for 480 – 550 nm, $\varnothing$25.0 mm, LPVISA100,

   - Mounted Zero Order, 1/2 Waveplate 473nm, WPH05M-473,

   - Two Rotation Mount for $\varnothing$1" Optics, RSP1C/M,

   - Reflective filters, N.D. 1.0 and 2.0, ND510A and ND520A (respectively),

2. Zeiss

   - Plan-Neofluor, 10x, N.A. 0.3,

- EC Plan-Neofluor, 20x, N.A. 0.5

- Plan-Apochromat 63x, N.A. 1.4, Oil,

- Axiovert 200, inverted microscope,

3. Photometrics

- Cascade II, 512x512, emCCD camera,

4. Cairn Research

- OptSplit II LS, Image Splitter

Appendix J

Fluorescence anisotropy imaging system with two light sources

The fluorescence anisotropy imaging system with two light sources was based on a single light source system (see Figure 4.3). The second light source was added to the system, a 647 nm laser diode (700 mW, Optoelectronics). The diode emitted light with an astigmatic profile, what was adjusted using by characterising beam profile with Beam view (Coherent Inc.) and two astigmatic lenses (see Figure J.1). Because the spot size of the 647 nm beam was similar to size of 473 nm laser, no beam expansion system was used. Polarisation optics (linear polariser and $\lambda/2$ plate) are going to be for this light source to control the polarisation properties of the laser light (the linear polarisation and polarisation direction). Polarisation optics would need to be calibrated to find optimal angular positions of the polarisation optics (see Section 4.4.1). The 647 nm light would be passed directly to a dichroic mirror, reflecting 647 nm light (at angle 45°) and transmitted 473 nm light, thus enabling two colour imaging if needed.

1. Components of 473 nm light system were discussed in Section 4.4.1 and Appendix I

2. Thorlabs

   - Molded Glass Aspheric lens (mounted)f = 4.03 mm, N.A. = 0.64, C340TMD-B,

   - Plano-Convex Cylindrical lens f = 200 mm, H = 30 mm, L = 32 mm, N-BK7,

- TE-Cooled Mount for Dia=5.6 mm and Dia=9 mm Lasers, TCLDM9

- General-Purpose Plate Holder, FP01

- Benchtop Laser Diode/TEC Controller, 1 A / 96 W, ITC4001

3. Optoelectronics

- Laser diode 647 nm, 700 mW, HL63193MG

4. Comar

- Dichroic mirror 40 x 25 mm, 495 IB 125

5. Chroma

- Filter set for Alexa Fluor 647, 49913
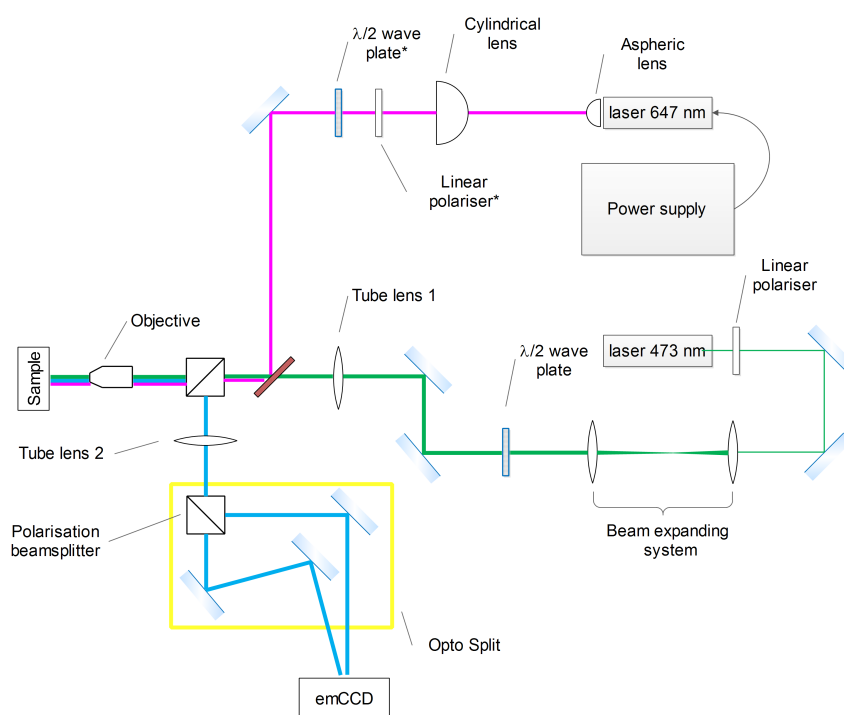
Figure J.1: Fluorescence anisotropy imaging system with two light sources. The design of light source 473 nm was discussed in Section 4.4.1. A 647 nm laser light source has an aspheric shape and was collimated using to astigmatic lenses. Similarly as in 473 nm light source set up, linear polariser and a $\lambda/2$ plate will be added to the system to control the polarisation properties of the laser beam. Then light was reflected by a dichroic plate and passes to an inverted Zeiss microscope with objective (63x, N.A. 1.4). After imaging the sample fluorescent light was divided into separate polarisation channels (parallel and perpendicular in regards to polarisation of exciting beam) and imaged with an CCD camera (Photometrics).

Bibliography

Abbe, E.
  1873. "Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung".
  (German) [Contributions to the Theory of the Microscope and of Microscopic Percep-
  tion]. 9(1):413–468.

Aga, M., J. M. Bradley, K. E. Keller, M. J. Kelley, and T. S. Acott
  2008. Specialized podosome - or invadopodia-like structures (PILS) for focal trabecular
  meshwork extracellular matrix turnover. *Investigative Ophthalmology Visual Science*,
  49(12):5353–5365.

Ankers, M., M. M. Breunig, H.-P. Kriegel, and J. Sander
  1999. OPTICS: Ordering Points To Identify the Clustering Structure. ACM SIGMOD
  international conference on management of data. *ACM Press*, 28(2):49–60.

Annibale, P.
  2012. *Investigating the impact of single molecule fluorescence dynamics on photo ac-
  tivated localization microscopy experiments*. PhD thesis, École Polytechnique Fédérale
  de Lausanne, Switzerland.

Annibale, P., S. Vanni, M. Scarselli, U. Rothlisberger, and A. Radenovic
  2011. Identification of clustering artifacts in in photoactivated localization microscopy.
  Nature Methods, 8(7):527–528.

Axelrod, D.
  1981. Cell-Substrate Contacts Illuminated by Total Internal Reflection Fluorescence.
  The Journal of Cell Biology, 89(1):141–145.

Baddeley, D., M. B. Cannell, and C. Soeller
  2011. Three-dimensional sub-100 nm super-resolution imaging of biological samples
  using a phase ramp in the objective pupil. *Nano Research*, 4(6):589–598.

Bates, M., B. Huang, G. T. Dempsey, and X. Zhuang
  2007. Multicolor Super-Resolution Imaging with Photo-Switchable Fluorescent Probes.
  Science, 317(5845):1749–1753.

Betzig, E., A. Lewis, A. Harootunian, E. Isaacson, and E. Kratschmer
  1986. Near-field scanning optical microscopy (NSOM). *Biophysical Journal*, 49(1):269–
  279.

Betzig, E., G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino,
  M. W. Davidson, J. Lippincott-Schwartz, and H. F. Hess
  2006. Imaging intracellular fluorescent proteins at nanometer resolution. *Science*,
  313(5793):1642–1645.

Bhuwania, R., S. Cornfine, Z. Fang, M. Krüger, E. J. Luna, and S. Linder
2012. Supervillin couples myosin-dependent contractility to podosomes and enables their turnover. *Journal of Cell Science*, 125(9):2300–2314.

Bio-nanoscopy Lab
2015. STED principle [Image]. [Online; retrieved on 17th of November 2015].

Blundell, M. P., G. Bouma, Y. Calle, G. E. Jones, C. Kinnon, and A. J. Thrasher
2008. Improvement of migratory defects in a murine model of wiskott–aldrich syndrome gene therapy. *Molecular Therapy*, 16(5):836–844.

Burgstaller, G. and M. Gimona
2004. Actin cytoskeleton remodelling via local inhibition of contractility at discrete microdomains. *Journal of Cell Science*, 117(2):223–231.

Calle, Y., S. Burns, A. J. Thrasher, and G. E. Jones
2006. The leukocyte podosome. *European Journal of Cell Biology*, 85(3–4):151–157.

Calle, Y., H. chuan Chou, A. J. Thrasher, and G. E. Jones
2004. Wiskott-Aldrich syndrome protein and the cytoskeletal dynamics of dendritic cells. *Journal of Pathology*, 204(4):460–469.

Carman, C. V.
2009. Mechanisms for transcellular diapedesis: probing and pathfinding by invadosome-like protrusions. *Journal of Cell Science*, 122(17):3025–3035.

Chasles, F., B. Dubertret, and A. C. Boccara
2007. Optimization and characterization of a structured illumination microscope. *Optics Express*, 15(24):16130–16140.

Clayton, A.
2009. Fluorescence resonance energy transfer and applications. In *Fluorescence Applications in Biotechnology and Life Sciences*, E. M. Goldys, ed., chapter 8, Pp. 259–266. New Jersey, USA: Wiley-Blackwell.

Clayton, A. H., Q. S. Hanley, D. J. Arndt-Jovin, V. Subramaniam, and T. M. Jovin
2002. Dynamic Fluorescence Anisotropy Imaging Microscopy in the Frequency domain (rFLIM). *Biophysical Journal*, 83(3):1631–1649.

Cohen, C. J., J. T. C. Shieh, R. J. Pickles, T. Okegawa, J.-T. Hsieh, and J. M. Bergelson
2001. The coxsackievirus and adenovirus receptor is a transmembrane component of the tight junction. *PNAS*, 98(26):15191–15196.

Coltharp, C., R. P. Kessler, and J. Xiao
2012. Accurate Construction of Photoactivated Localization Microscopy (PALM) Images for Quantitative Measurements. *PLoS One*, 7(12):e51725.

Cottet, M., O. Faklaris, E. Trinquet, J.-P. Pin, and T. Durroux
2013. New Fluorescent Strategies Shine Light on the Evolving Concept of GPCR Oligomerization. In *Springer Series on Fluorescence 13: Fluorescent Methods to Study Biological Membranes*, Y. Mély and G. Duportail, eds., Pp. 389–415. New York, USA: Springer-Verlag.

Courtneidge, S. A., E. F. A. Jr., I. Pass, D. F. Seals, and L. Tesfay
2005. The Src Substrate Tks5, Podosomes (Invadopodia), and Cancer Cell Invasion. *Cold Spring Harbor Symposia on Quantitative Biology*, 70:167–171.

Cox, S., E. Rosten, J. Monypenny, T. Jovanovic-Talisman, D. T. Burnette, J. Lippincott-Schwartz, G. E. Jones, and R. Heintzmann
2012. Bayesian localization microscopy reveals nanoscale podosome dynamics. *Nature Methods*, 9(2):195–200.

Dani, A., B. Huang, J. Bergan, C. Dulac, and X. Zhuang
2010. Superresolution imaging of chemical synapses in the brain. *Neuron*, 68(5):843–856.

Dave, R., D. S. Terry, J. B. Munro, and S. C. Blanchard
2009. Mitigating unwanted photophysical processes for improved single-molecule fluorescence imaging. *Biophysical Journal*, 96(6):2371–2381.

Davidson, R. and J. G. MacKinnon
2000. Bootstrap tests: How many bootstraps? *Econometric Reviews*, 19(1):55–68.

Deschout, H., A. Shivanandan, P. Annibale, M. Scarselli, and A. Radenovic
2014. Progress in quantitative single-molecule localization microscopy. *Histochemistry and Cell Biology*, 142(1):5–17.

Dickson, R. M., A. B. Cubitt, R. Y. Tsien, and W. E. Moerner
1997. On/off blinking and switching behaviour of single molecules of green fluorescent protein. *Nature*, 388(6640):355–358.

Ding, J., K. T. Takasaki, and B. L. Sabatini
2010. Supraresolution imaging in brain slices using stimulated-emission depletion 2-photon laser scanning microscopy. *Neuron*, 63(4):429–437.

Efron, B. and R. J. Tibshirani
1994. *An Introduction to the Bootstrap*. New York, USA: Springer.

Egner, A., C. Geisler, C. von Middendorff, H. Bock, D. Wenzel, R. Medda, M. Andresen, A. C. Stiel, S. Jakobs, C. Eggeling, and A. S. S. W. Hell
2007. Fluorescence nanoscopy in whole cells by asynchronous localization of photoswitching emitters. *Biophysical Journal*, 93(9):3285–3290.

Eisen, M. B., P. T. Spellman, P. O. Brown, and D. Botstein
1998. Cluster analysis and display of genome-wide expression patterns. *PNAS*, 95(25):14863–14868.

Endesfelder, U., S. Malkusch, B. Flottmann, J. Mondry, P. Liguzinski, P. J. Verveer, and M. Heilemann
2011. Chemically induced photoswitching of fluorescent probes – a general concept for super-resolution microscopy. *Molecules*, 16(4):3106–3118.

Erickson, H. P.
2009. Size and shape of protein molecules at the nanometer level determined by sedimentation, gel filtration, and electron microscopy. *Biological Procedures Online*, 11(1):32–51.

Erni, R., M. D. Rossell, C. Kisielowski, and U. Dahmen
2009. Atomic-resolution imaging with a sub-50-pm electron probe. *Physical Review Letters*, 102(9):096101.

Ester, M., H.-P. Kriegel, J. Sander, and X. Xu
1996. A density-based algorithm for discovering clusters in large spatial databases with noise. In *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD-96)*, Pp. 226–231. AAAI Press.

Fiori, N. D. and A. Meller
2010. The effect of dye-dye interactions on the spatial resolution of single-molecule FRET measurements in nucleic acids. *Biophysical Journal*, 98(10):2265–2272.

Fischler, M. A. and R. C. Bolles
1981. Random sample consensus: A praradigm for model fitting with applications to image analysis and automated cartography. *Comm. of the ACM.*, 24(6):381–395.

Fölling, J., M. Bossi, H. Bock, R. Medda, C. A. Wurm, B. Hein, S. Jakobs, C. Eggeling, and S. W. Hell
2008. Fluorescence nanoscopy by ground-state depletion and single-molecule return. *Nature Methods*, 5(11):943–945.

Foxall, E., A. Pipili, G. E. Jones, and C. M. Wells
2016. Significance of kinase activity in the dynamic invadosome. *European Journal of Cell Biology*.

Fu, H. L., J. L. Mueller, M. P. Javid, J. K. Mito, D. G. Kirsch, N. Ramanujam, and J. Q. Brown
2013. Optimization of a widefield structured illumination microscope for non-destructive assessment and quantification of nuclear features in tumor margins of a primary mouse model of sarcoma. *PLoS One*, 8(7):e68868.

GATTAquant
2016. DNA origami plate [Image]. [Online; retrieved on 23rd of February 2016].

Gautier, I., M. Tramier, C. Durieux, J. Coppey, R. Pansu, J.-C. Nicolas, K. Kemnitz, and M. Coppey-Moisan
2001. Homo-FRET microscopy in living cells to measure monomer-dimer transition of GFP-tagged proteins. *Biophysical Journal*, 80(6):3000–3008.

Gawden-Bone, C., Z. Zhou, E. King, A. Prescott, C. Watts, and J. Lucocq
2010. Dendritic cell podosomes are protrusive and invade the extracellular matrix using metalloproteinase MMP-14. *Journal of Cell Science*, 123(9):1427–1437.

Gell, C., D. Brockwell, and A. Smith
2006. *Handbook of Single Molecule Fluorescence Spectroscopy*. Oxford, UK: Oxford University Press.

Giannone, G., E. Hosy, F. Levet, A. Constals, K. Schulze, A. I. Sobolevsky, M. P. Rosconi, E. Gouaux, R. Tampé, D. Choquet, and L. Cognet
2008. Dynamic superresolution imaging of endogenous proteins on living cells at ultra-high density. *Current Opinion in Cell Biology*, 99(4):1303–1310.

Gimona, M., R. Buccione, S. A. Courtneidge, and S. Linder
2008. Assembly and biological role of podosomes and invadopodia. *Current Opinion in Cell Biology*, 20(2):235–241.

Gitai, Z.
2009. New fluorescence microscopy methods for microbiology: Sharper, faster, and quantitative. *Current Opinion in Microbiology*, 12(3):341–346.

Gong, Z., B. Ju, and H. Wan
2001. Green fluorescent protein (GFP) transgenic fish and their applications. *Genetica*, 111(1):213–225.

Gould, T. J., V. V. Verkhusha, and S. T. Hess
2009. Imaging biological structures with fluorescence photoactivation localization microscopy. *Nature Protocols*, 4(3):291–308.

Guo, D., H. Zhou, Y. Zou, W. Yin, H. Yu, Y. Si, J. Li, Y. Zhou, X. Zhou, and R. J. Soares-Magalhāes
2013. Geographical Analysis of the Distribution and Spread of Human Rabies in China from 2005 to 2011. *PLoS ONE*, 8(8):e72352.

Guo, W. and S. Peddada
2008. Adaptive choice of the number of bootstrap samples in large scale multiple testing. *Statistical Applications in Genetics and Molecular Biolog*, 7(1). Article13.

Gustafsson, M. G. L.
2005. Nonlinear structured-illumination microscopy: Wide-field fluorescence imaging with theoretically unlimited resolution. *PNAS*, 102(37):13081–13086.

Gustafsson, M. G. L., L. Shao, I. N. G. Peter M. Carlton, C. J. Rachel Wang, W. Z. Cande, D. A. Agard, and J. W. Sedat
2008. Three-dimensional resolution doubling in wide-field fluorescence. *Biophysical Journal*, 94(12):4957–4970.

Hafi, N., M. Grunwald, L. S. van den Heuvel, T. Aspelmeier, J.-H. Chen, M. Zagrebelsky, O. M. Schütte, C. Steinem, M. Korte, A. Munk, and P. J. Walla
2014. Fluorescence nanoscopy by polarization modulation and polarization angle narrowing. *Nature Methods*, 11(5):579–584.

Han, H., D. Kampik, F. Grehn, and G. Schlunck
2013. Tgf-β2-induced invadosomes in human trabecular meshwork cells. *PLoS One*, 8(8):e70595.

Hedley, D. W., M. L. Friedlander, I. W. Taylor, C. A. Rugg, and E. A. Musgrove
1983. Method for analysis of cellular dna content of paraffin-embedded pathological material using flow cytometry. *Journal of Histochemistry & Cytochemistry*, 31(11):1333–1335.

Heilemann, M., E. Margeat, R. Kasper, M. Sauer, and P. Tinnefeld
2005. Carbocyanine dyes as efficient reversible single-molecule optical switch. *Journal of the American Chemical Society*, 127(11):3801–3806.

Heilemann, M., S. van de Linde, M. Schüttpelz, R. Kasper, B. Seefeldt, A. Mukherjee, P. Tinnefeld, and M. Sauer
2008. Subdiffraction-resolution fluorescence imaging with conventional fluorescent probes. *Angewandte Chemie*, 47(33):6172–6176.

Heintzmann, R., T. M. Jovin, and C. Cremer
2002. Saturated patterned excitation microscopy—a concept for optical resolution improvement. *Journal of the Optical Society of America A*, 19(8):1599–1609.

Hell, S. W. and J. Wichmann
1994. Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy. *Optics Letters*, 19(11):780–782.

Henriques, R., M. Lelek, E. F. Fornasiero, F. Valtorta, C. Zimmer, and M. M. Mhlanga
2010. QuickPALM: 3D real-time photoactivation nanoscopy image processing in ImageJ. *Nature Methods*, 7(5):339–340.

Hess, S. T., T. P. K. Girirajan, and M. D. Mason
2006. Ultra-high resolution imaging by fluorescence photoactivation localization microscopy. *Biophysical Journal*, 91(11):4258–4272.

Hess, S. T., T. J. Gould, M. V. Gudheti, S. A. Maas, K. D. Mills, and J. Zimmerberg
2007. Dynamic clustered distribution of hemagglutinin resolved at 40 nm in living cell membranes discriminates between raft theories. *PNAS*, 104(44):17370–17375.

Hess, S. T., M. Kumar, A. Verma, J. Farrington, A. Kenworthy, and J. Zimmerberg
2005. Quantitative electron microscopy and fluorescence spectroscopy of the membrane distribution of influenza hemagglutinin. *Journal of Cell Biology*, 169(6):965–976.

Hirvonen, L. M., T. Kilfeather, and K. Suhling
2015. Single-molecule localization software applied to photon counting imaging. *Applied Optics*, 54(16):5074–5082.

Hooke, R.
1665. *Micrographia*. Project Guttenberg. Ebook file.

Huang, B., W. Wang, M. Bates, and X. Zhuang
2008. Three-dimensional super-resolution imaging by stochastic optical reconstruction microscopy. *Science*, 319(5864):810–813.

Hyatt, A. D. and T. G. Wise
2001. *Immunolabeling*, Pp. "73–107. Boston, MA: Birkhäuser Boston.

Janson, K.
2016. Jablonski diagram [Image]. [Online; retrieved on 27th of November 2016].

Jones, G. E., D. Zicha, G. A. Dunn, M. Blundell, and A. Thrasher
2002. Restoration of podosomes and chemotaxis in Wiskott-Aldrich syndrome macrophages following induced expression of WASp. *The International Journal of Biochemistry & Cell Biology*, 34(7):806–815.

Juette, M. F., T. J. Gould, M. D. Lessard, M. J. Mlodzianoski, B. S. Nagpure, B. T. Bennett, S. T. Hess, and J. Bewersdorf
2008. Three-dimensional sub–100 nm resolution fluorescence microscopy of thick samples. *Nature Methods*, 5(6):527–529.

Kaiser, R. D. and E. London
1998. Location of diphenylhexatriene (DPH) and its derivatives within membranes: Comparison of different fluorescence quenching analyses of membrane depth. *Biochemistry*, 37(22):8180–8190.

Keating, S. and T. Wensel
1991. Nanosecond fluorescence microscopy. emission kinetics of fura-2 in single cells. *Biophysical Journal*, 59(1):186–202.

Kiskowski, M. A., J. F. Hancock, and A. K. Kenworthy
2009. On the use of Ripley's K-Function and its derivatives to analyze domain size. *Biophysical Journal*, 97(4):1095–1103.

Kremers, G.-J., S. G. Gilbert, P. J. Cranfill, M. W. Davidson, and D. W. Piston
2010. Fluorescent proteins at a glance. *Journal of Cell Science*, 124(2):157–160.

Křížek, P., P. W. Winter, Z. Švindrych, J. Borkovec, M. Ovesný, D. A. Roess, B. G. Barisas, and G. M. Hagen
2014. Imaging of insulin receptors in the plasma membrane of cells using super-resolution single molecule localization microscopy. In *Microscopy: advances in scientific research and education*, A. Méndez-Vilas, ed., chapter 2, Pp. 259–266. Formatex Research Center.

Lakowicz, J. R.
2006. *Principles of Fluorescence Spectroscopy*. New York, USA: Springer.

Lampe, A., V. Haucke, S. J. Sigrist, M. Heilemann, and J. Schmoranzer
2012. Multi-colour direct STORM with red emitting carbocyanines. *Biology of the Cell*, 104(4):229–237.

Lelek, M., F. D. Nunzio, R. Henriques, P. Charneau, N. Arhel, and C. Zimmer
2012. Superresolution imaging of HIV in infected cells with Flash-PALM. *PNAS*, 109(22):8564–8569.

Lemmer, P., M. Gunkel, D. Baddeley, R. Kaufmann, A. Urich, Y. Weiland, J. Reymann, P. Müller, M. Hausmann, and C. Cremer
2008. SPDM: light microscopy with single-molecule resolution at the nanoscale. *Applied Physics B*, 93(1):1–12.

Li, D., L. Shao, B.-C. Chen, X. Zhang, M. Zhang, B. Moses, D. E. Milkie, J. R. Beach, J. A. H. III, M. Pasham, T. Kirchhausen, M. A. Baird, M. W. Davidson, P. Xu, and E. Betzig
2015. Extended-resolution structured illumination imaging of endocytic and cytoskeletal dynamics. *Science*, 349(6251):aab3500–1–aab3500–10.

Li, L. and Y. Xi
2011. Research on Clustering Algorithm and Its Parallelization Strategy. In *Procedings of International Conference on Computational and Information Sciences*, Pp. 325–328.

Lidke, D., P. Nagy, B. Barisas, R. Heintzmann, J. Post, K. Lidke, A. Clayton, D. Arndt-Jovin, and T. Jovin
2003. Imaging molecular interactions in cells by dynamic and static fluorescence anisotropy (rflim and emfret). *Biochemical Society Transactions*, 31(5):1020–1027.

Lidke, K. A., B. Rieger, D. S. Lidke, and T. M. Jovin
2005. The role of photos statistic in fluorescence anisotropy imaging. In *IEEE Transactions on Image Processing: a Publication of the IEEE Signal Processing Society*, volume 14, Pp. 1237–1245.

Linder, S. and P. Kopp
2005. Podosomes at a glance. *Journal of Cell Science*, 118(10):2079–2082.

Linder, S., D. Nelson, M. Weiss, and M. Aepfelbacher
1999. Wiskott-aldrich syndrome protein regulates podosomes in primary human macrophages. *PNAS*, 96(17):9648–9653.

Linder, S. and C. Wiesner
2015. Tools of the trade: podosomes as multipurpose organelles of monocytic cells. *Cellular and Molecular Life Sciences*, 72(1):121–135.

Lipson, A., S. G. Lipson, and H. Lipson
2011. *Optical Physics*, 4 edition. Cambridge, UK: Cambridge University Press.

Longin, A., C. Souchier, M. Ffrench, and P.-A. Bryon
1993. Comparison of anti-fading agents used in fluorescence microscopy: image analysis and laser confocal microscopy study. *Journal of Histochemistry & Cytochemistry*, 41(12):1833–1840.

Löschberger, A., C. Franke, G. Krohne, S. van de Linde, and M. Sauer
2014. Correlative super-resolution fluorescence and electron microscopy of the nuclear pore complex with molecular resolution. *Journal of Cell Science*, 127(20):4351–4355.

Luxenburg, C., S. Winograd-Katz, L. Addadi, and B. Geiger
2012. Involvement of actin polymerization in podosome dynamics. *Journal of Cell Science*, 125(7):1666–1672.

MacKay, D. J. C.
2003. *Information Theory, Inference, and Learning Algorithms*. Cambridge, UK: Cambridge University Press.

MacQueen, J.
1967. Some Methods for Classification and Analysis of Multivariate Observations. In *Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability*, volume 1, Pp. 281–297. University of California Press.

Matveeva, E., Z. Gryczynski, I. Gryczynski, and J. R. Lakowicz
2004. Immunoassays based on directional surface plasmon-coupled emission. *Journal of Immunological Methods*, 286(1–2):133–140.

Meddens, M. B., B. Rieger, C. G. Figdora, A. Cambi, and K. van den Dries
2013. Automated podosome identification and characterization in fluorescence microscopy. *Microscopy and Microanalysis*, 19(1):180–189.

Meddens, M. B., K. van den Dries, and A. Cambi
2014. Podosomes revealed by advanced bioimaging: What did we learn? *European Journal of Cell Biology*, 93(10):380–387.

Mertz, J.
2011. Optical sectioning microscopy with planar or structured illumination. *Nature Methods*, 8(10):811–819.

Minsky, M.
1961. Patent no. us3013467 (a) – 1961-12-19.

Minsky, M.
1988. Memoir on Inventing the Confocal Scanning Microscope. *Scanning*, 10(4):128–138.

Monypenny, J., H.-C. Chou, I. Bañón-Rodríguez, A. J. Thrasher, G. E. J. Inés M. Antónb, and Y. Calle
2011. Role of wasp in cell polarity and podosome dynamics of myeloid cells. *European Journal of Cell Biology*, 90(2-3):198–204.

Mooi, E. and M. Sarstedt
2011. *A Concise Guide to Market Research: The Process, Data and Methods using IBM SPSS Statistics*. New York, USA: Springer.

Morton, P. E., A. Hicks, T. Nastos, G. Santis, and M. Parsons
2013. CAR regulates epithelial cell junction stability through control of E-cadherin trafficking. *Scientific Reports*, 3:2889.

Murphy, D. A. and S. A. Courtneidge
2011. The 'ins' and 'outs' of podosomes and invadopodia: characteristics, formation and function. *Nature Reviews*, 12(7):413–426.

Murray, J. M.
2011. Methods for imaging thick specimens: Confocal microscopy, deconvolution and structured illumination. *Cold Spring Harbor Protocols*.

Neil, M. A. A., R. Juskaitis, and T. Wilson
1997. Method of obtaining optical sectioning by using structures light in conventional microscope. *Optics Letters*, 22(24):1905–1907.

Nikon
2015. Super Resolution Microscope N-STORM. STORM Protocol-Sample Preparation. [Online; retrieved on 10th of November 2015].

Ober, R. J., S. Ram, and E. S. Ward
2004. Localization accuracy in single-molecule micorscopy. *Biophysical Journal*, 86(2):1185–1200.

Oldfield, R. J.
2001. *Differential Interference Contrast Light Microscopy*. John Wiley & Sons, Ltd.

Olivier, N., D. Keller, P. Gönczy, and S. Manley
  2013. Resolution doubling in 3D-STORM imaging through improved buffers. *PLOS
  ONE*, 8(7):e69004.

Owen, D. M., D. J. Williamson, L. Boelen, A. Magenau, J. Rossy, and K. Gaus
  2013. Quantitative analysis of three-dimensional fluorescence localization microscopy
  data. *Biophysical Journal*, 105(2):L05–L07.

Parton, R. G. and J. F. Hancock
  2004. Lipid rafts and plasma membrane microorganisation: insights from Ras. *Trends
  in Cell Biology*, 14(3):141–147.

Parzen, E.
  1962. On estimation of a probability density function and mode. *Annals of Mathematical
  Statistics*, 33(3):1065–1076.

Patel, A. and P. R. Dash
  2012. Formation of atypical podosomes in extravillous trophoblasts regulates extracel-
  lular matrix degradation. *European Journal of Cell Biology*, 91(3):171–179.

Pavani, S. R. P., M. A. Thompson, J. S. Biteen, S. J. Lord, N. Liu, R. J. Twieg, R. Piestun,
  and W. E. Moerner
  2009. Three-dimensional, single-molecule fluorescence imaging beyond the diffraction
  limit by using a double-helix point spread function. *PNAS*, 106(9):2995–2999.

Pawley, J. B.
  2008. LVSEM for Biology. In *Biological Low-Voltage Scanning Electron Microscopy*,
  H. Schatten and J. B. Pawley, eds., chapter 2, Pp. 27–106. New York, USA: Springer-
  Verlag.

Pertsinidis, A., K. Mukherjee, M. Sharma, Z. P. Pang, S. R. Park, Y. Zhang, A. T.
  Brunger, T. C. Südhof, and S. Chu
  2013. Ultrahigh-resolution imaging reveals formation of neuronal SNARE/Munc18 com-
  plexes in situ. *PNAS*, 110(2):E2812–E2820.

Photometrics
  2014. Localisation microscopy principle [Image]. [Online; retrieved on 14th of May
  2014].

Purwar, Y., S. L. Shah, G. Clarke, A. Almugairi, and A. Muehlenbachs
  2011. Automated and unsupervised detection of malarial parasites in microscopic im-
  ages. *Malaria Journal*, 10(364):364–374.

Raj, A. and A. van Oudenaarden
  2008. Nature, Nurture, or Chance: Stochastic Gene Expression and Its Consequences.
  Cell, 135(2):216–226.

Rasnik, I., S. A. McKinney, and T. Ha
  2006. Nonblinking and long-lasting single-molecule fluorescence imaging. *Nature Meth-
  hods*, 2(11):891–893.

Rayleigh, L.
  1879. Xxxi. investigations in optics, with special reference to the spectroscope. *Philosophical Magazine Series 5*, 8(49):261–274.

Resch-Genger, U., M. Grabolle, S. Cavaliere-Jaricot, R. Nitschke, and T. Nann
  2008. Quantum dots versus organic dyes as fluorescent labels. *Nature Methods*, 5(9):763–775.

Ries, J., C. Kaplan, E. Platonova, H. Eghlidi, and H. Ewers
  2012. A simple, versatile method for gfp-based super-resolution microscopy via nanobodies. *Nature Methods*, 9(6):582–584.

Ripley, B.
  1976. The second-order analysis of stationary point processes. *Journal of Applied Probability*, 13(2):255–266.

Rothemund, P. W. K.
  2006. Folding DNA to create nanoscale shapes and patterns. *Nature*, 440(7082):297–302.

Rottiers, P., F. Saltel, T. Daubon, B. Chaigne-Delalande, V. Tridon, C. Billottet, E. Reuzeau, and E. Génot
  2009. TGF-induced endothelial podosomes mediate basement membrane collagen degradation in arterial vessels. *Journal of Cell Science*, 122(23):4311–4318.

Rust, M. J., M. Bates, and X. Zhuang
  2006. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nature Methods*, 3(10):793–795.

Santiago-Medina, M., K. A. Gregus, R. H. Nichol, S. M. O'Toole, and T. M. Gomez
  2015. Regulation of ecm degradation and axon guidance by growth cone invadosomes. *Development*, 142:486–496.

Scarselli, M., P. Annibale, and A. Radenovic
  2012. Cell Type-specific β2-Adrenergic Receptor Clusters Identified Using Photoactivated Localization Microscopy Are Not Lipid Raft Related, but Depend on Actin Cytoskeleton Integrity. *Journal of Biological Chemistry*, 287:16768–16780.

Schachtner, H., S. D. J. Calaminus, A. Sinclair, J. Monypenny, M. P. Blundell, C. Leon, T. L. Holyoake, A. J. Thrasher, A. M. Michie, M. Vukovic, C. Gachet, G. E. Jones, S. G. Thomas, S. P. Watson, , and L. M. Machesky
  2013a. Megakaryocytes assemble podosomes that degrade matrix and protrude through basement membrane. *Blood*, 121:2542–2552.

Schachtner, H., S. D. J. Calaminus, S. G. Thomas, and L. M. Machesky
  2013b. Podosomes in adhesion, migration, mechanosensing and matrix remodeling. *Cytoskeleton*, 70(10):572–589.

Schermelleh, L., R. Heintzmann, and H. Leonhardt
  2010. A guide to super-resolution fluorescence microscopy. *Journal of Cell Biology*, 190(2):165–175.

Schmidt, R., C. A. Wurm, S. Jakobs, J. Engelhardt, A. Egner, and S. W. Hell
2008. Spherical nanosized focal spot unravels the interior of cells. *Nature Methods*, 5(6):539–544.

Seano, G., G. Chiaverina, P. A. Gagliardi, L. di Blasio, A. Puliafito, C. Bouvard, R. Sessa, G. Tarone, L. Sorokin, D. Helley, R. K. Jain, G. Serini, F. Bussolino, and L. Primo
2014. Endothelial podosome rosettes regulate vascular branching in tumour angiogenesis. *Nature Cell Biology*, 16(10):931–941.

Sengupta, P., T. Jovanovic-Talisman, D. Skoko, M. Renz, S. L. Veatch, and J. Lippincott-Schwartz
2011. Probing protein heterogeneity in the plasma membrane using PALM and pair correlation analysis. *Nature Methods*, 8(11):969–975.

Sharonov, A. and R. M. Hochstrasser
2006. A mathematical theory of communication. *The Bell System Technical Journal*, 103(50):18911–18916.

Sheely, M. L.
1932. Glycerol viscosity tables. *Industrial and Engineering Chemistry*, 24(9):1060–1064.

Shivanandan, A., H. Deschout, M. Scarselli, and A. Radenovic
2014. Challenges in quantitative single molecule localization microscopy. *FEBS Letters*, 588(19):3595–3602.

Shivanandan, A., J. Unnikrishnan, and A. Radenovic
2015a. Accounting for limited detection efficiency and localization precision in cluster analysis in single molecule localization microscopy. *PLoS one*, 10(3):1–15.

Shivanandan, A., J. Unnikrishnan, and A. Radenovic
2015b. On characterizing membrane protein clusters with model-free spatial correlation approaches. *PLoS one*.

Shroff, H., C. G. Galbraith, J. A. Galbraith, H. White, J. Gillette, S. Olenych, M. W. Davidson, and E. Betzig
2007. Dual-color superresolution imaging of genetically expressed probes within individual adhesion complexes. *PNAS*, 104(57):20308–20313.

Siddiqui, T. A., S. Lively, C. Vincent, and L. C. Schlichter
2012. Regulation of podosome formation, microglial migration and invasion by Ca^{2+} -signaling molecules expressed in podosomes. *Journal of Neuroinflammation*, 9(250):250–266.

Siegel, J., K. Suhling, S. Lévêque-Fort, S. E. D. Webb, D. M. Davis, D. Phillips, Y. Sabharwal, and P. M. W. French
2003. Wide-field time-resolved fluorescence anisotropy imaging (TR-FAIM): Imaging the rotational mobility of a fluorophore. *Review of Scientific Instruments*, 74(1):182–192.

Sigma-Aldrich
2017. Antibody classes and subclasses. [Online; retrieved on 23th of January 2017].

Small, A. R. and R. Parthasarathy
2014. Superresolution localization methods. *The Annual Review of Physical Chemistry*, 65(1):107–125.

Soeller, C. and D. Baddeley
2013. Super-resolution imaging of EC coupling protein distribution in the heart. *Journal of Molecular and Cellular Cardiology*, 58:32–40.

Staszowska, A. D., P. Fox-Roberts, E. Foxall, G. E. Jones, and S. Cox
2016. Investigation of podosome ring protein arrangement using localization microscopy images. *Methods*, Pp. –.

Steinbach, M., V. Kumar, and P.-N. Tan
2003. *Introduction to Data Mining*. New York, USA: Springer.

Stoyan, D. and A. Penttinen
2000. Recent applications of spatial point process methods in forestry statistics. *Statistical Science*, 15(1):61–78.

Tanaka, K. A. K., K. G. N. Suzuki, Y. M. Shirai, S. T. Shibutani, M. S. H. Miyahara, H. Tsuboi, M. Yahara, A. Yoshimura, S. Mayor, T. K. Fujiwara, and A. Kusumi
2010. Membrane molecules mobile even after chemical fixation. *Nature Methods*, 7(11):865–866.

Thompson, R. E., D. R. Larson, and W. W. Webb
2002. Precise nanometer localization analysis for individual fluorescent probes. *Biophysical Journal*, 82(3):182–192.

Thrasher, A. J. and S. O. Burns1
2010. WASP: a key immunological multitasker. *Nature Reviews Immunology*, 10(5):2775–2783.

Tramier, M., K. Kemnitz, C. Durieux, J. Coppey, P. Denjean, R. B. Pansu, and M. Coppey-Moisan
2000. Restrained Torsional Dynamics of Nuclear DNA in Living Proliferative Mammalian Cells. *Biophysical Journal*, 78(5):2614–2627.

Tsien, R. Y.
1998. The Green Fluorescent Protein. *Annual Reviews of Biochemistry*, 67:509–544.

Tsien, R. Y.
2016. Fluorescent proteins [Image]. [Online; retrieved on 27th of January 2016].

van den Dries, K., S. L. Schwartz, J. Byars, M. Meddens, M. Bolomini-Vittoria, D. S. Lidke, C. G. Figdor, K. A. Lidke, and A. Cambi
2013. Dual color super-resolution microscopy reveals nanoscale organization of mechanosensory podosomes. *Molecular Biology of the Cell*, 24(13):2112–2123.

van den Dries, K., S. F. G. van Helden, J. te Riet, R. Diez-Ahedo, C. Manzo, M. M. Oud, F. N. van Leeuwen, R. Brock, M. F. Garcia-Parajo, A. Cambi, and C. G. Figdor
2012. Geometry sensing by dendritic cells dictates spatial organization and PGE(2)-induced dissolution of podosomes. *Cellular and Molecular Life Sciences*, 69(11):1889–1901.

Varma, R. and S. Mayor
1998. GPI-anchored proteins are organized in submicron domains at the cell surface. *Nature*, 394(6695):798–801.

Veillat, V., P. Spuul, T. Daubon, I. Egaña, I. Kramer, and E. Génot
2015. Podosomes: Multipurpose organelles? *The International Journal of Biochemistry and Cell Biology*, 65(1):52–60.

Verma, R.
2015. The MIN proximity function [Image]. [Online; retrieved on 22nd of January 2015].

Vijayakumar, V., J. Monypenny, X. J. Chen, L. Machesky, S. Lilla, A. J. Thrasher, I. M. Antón, Y. Calle, and G. E. Jones
2014. Tyrosine phosphorylation of WIP releases bound WASP and impairs podosome assembly in macrophages. *Journal of Cell Science*.

Wiesner, C., V. Le-Cabec, K. E. Azzouzi, I. Maridonneau-Parini, and S. Linder
2014. Podosomes in space. Macrophage migration and matrix degradation in 2D and 3D settings. *Cell Adhesion & Migration*, 8(3):179–191.

Williams, D. A., K. E. Fogarty, R. Y. Tsien, and F. S. Fay
1985. Calcium gradients in single smooth muscle cells revealed by the digital imaging microscope using Fura-2. *Nature*, 318(6046):558–561.

Williamson, D. J., D. M. Owen, J. Rossy, A. Magenau, M. Wehrmann, J. J. Gooding, and K. Gaus
2011. Pre-existing clusters of the adaptor Lat do not participate in early T cell signaling events. *Nature Immunology*, 12(7):655–662.

Willig, K. I., B. Harke, R. Medda, and S. W. Hell
2007. STED microscopy with continuous wave beams. *Nature Methods*, 4(11):915–918.

Wilson, S. M. and A. Bacic
2012. Preparation of plant cells for transmission electron microscopy to optimize immunogold labeling of carbohydrate and protein epitopes. *Nature Protocols*, 7(9):1716–1727.

Wilson, T.
1995. The role of the pinhole in confocal imaging system. In *Handbook of biological confocal micorscopy*, J. B. Pawley, ed., chapter 11, Pp. 167–182. New York, USA: Springer Science.

Wilson, T. and A. R. Carlini
1987. Size of the detector in confocal imaging systems. *Opt. Lett.*, 12(4):227–229.

Yan, Y. and G. Marriott
2003. Analysis of protein interactions using fluorescence technologies. *Current Opinion in Chemical Biology*, 7(5):635–640.

Yeow, E. K. L. and A. H. A. Clayton
2007. Enumeration of Oligomerization States of Membrane Proteins in Living Cells by Homo-FRET Spectroscopy and Microscopy: Theory and Application. *Biophysical Journal*, 92(9):3098–9104.

Zambonin-Zallone, A., A. Teti, M. V. Primavera, L. Naldini, and P. C. Marchisio
1983. Osteoclasts and monocytes have similar cytoskeletal structures and adhesion property in vitro. *Journal of Anatomy*, 137(1):57–70.

Zeiss
2015. SIM principle [Image]. [Online; retrieved on 17th of November 2015].

Zernike, F.
1942a. Phase contrast, a new method for the microscopic observation of transparent objects part I. *Physica*, 9(7):686–698.

Zernike, F.
1942b. Phase contrast, a new method for the microscopic observation of transparent objects part II. *Physica*, 9(10):974–980.

Zhang, J., K. Leiderman, J. R. Pfeiffer, B. S. Wilson, J. M. Oliver, and S. L. Steinberg
2006. Characterizing the topography of membrane receptors and signaling molecules from spatial patterns obtained using nanometer-scale electron dense probes and electron microscopy. *Micron*, 37(1):14–34.

Zicha, D., W. Allen, P. Brickell, C. Kinnon, G. Dunn, G. Jones, and A. Thrasher
1998. Chemotaxis of macrophages is abolished in the wiskott-aldrich syndrome. *British Journal of Haematology*, 101(4):659–665.