

THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Position Estimation and Tracking in Colloidal Particle Microscopy

MATS KVARNSTRÖM

CHALMERS | GÖTEBORG UNIVERSITY



Department of Mathematical Statistics
Chalmers University of Technology
and Göteborg University
Göteborg, Sweden 2005

Position Estimation and Tracking in Colloidal Particle Microscopy
MATS KVARNSTRÖM
ISBN 91-7291-574-9

© MATS KVARNSTRÖM, 2005.

Doktorsavhandlingar vid Chalmers Tekniska Högskola
Ny serie nr 2256
ISSN-0346-718X

Department of Mathematical Statistics
Chalmers University of Technology and Göteborg University
SE-412 96 Göteborg
Sweden
Telephone + 46 (0)31-772 1000

Position Estimation and Tracking in Colloidal Particle Microscopy
MATS KVARNSTRÖM
Department of Mathematical Statistics
Chalmers University of Technology and Göteborg University

Abstract

This thesis presents methods for estimating the locations (including depth) of spherical colloidal particles in images recorded in video microscopy. Understanding the behavior of colloidal interactions and diffusion is of crucial importance in a vast number of areas. However, since the theory fails to predict the behavior of several important colloidal suspensions, observations and measurements on the microscopic level are needed. Examples of common, everyday colloids are milk, paint and pharmaceuticals. The positioning methods developed here can be used for tracking of particles in three dimensions observed in video microscopy. We make several suggestions on how the positioning method should be modified and implemented to be used for this purpose.

Paper I introduces a method based on rotational symmetry to estimate the center of circular objects in images. Standard errors are also estimated. The accuracy of the estimates goes well beyond sub-pixel accuracy, which is validated in a simulation study. A modification of the local polynomial kernel estimator for censored data is also suggested. In Paper II we estimate the intensity profiles of particles at different known depths. These intensity profiles are then used for depth estimation in a template matching approach. The matching criterion takes into account both different background levels and censoring of pixel values. Paper III deals with the estimation of the diffusion coefficient from particle trajectories observed with measurement noise. The model includes two types of particles, fixed and diffusing. This is appropriate since this is the typical situation for particles in the images considered.

Key words: censored regression, colloidal chemistry, depth estimation, diffusion coefficient, nonparametric function estimation, position estimation, rotational symmetry, tracking, template matching, video microscopy

Acknowledgements

I would like to thank my scientific advisers Mats Rudemo and Aila Särkkä. Without your advise, encouragement, and patience, this thesis would not have been finished in time. The amount of hours you have spent reading half-finished drafts must have been enormous.

Special thanks go to Chris Glasbey, both for his hospitality during my visits in Edinburgh and the discussions we have had during the years. You have kept me going and you are a great source of inspiration to me. Thank you for everything!

Thanks also to Lennart Lindfors, who, besides having produced the images, often supported me with good and encouraging comments.

Thanks go to my friends and colleagues who have given me great laughs and good discussions of various subjects over the years. Among the people in the senior staff, I would in particular like to thank Olle Hågström, Jeff Steif, Torgny Lindvall, and Patrik Albin for being great teachers and sources of inspiration. I am also grateful for the ever-present friendliness and support from the administrative staff.

I am fortunately blessed with lots of good friends. Thank you all for suitably distracting me with social events such as guitar playing, concerts, dinners and drinks.

I would like to thank my family for always being supportive, whatever my plans were and whatever I have attempted. Self-confidence is born out of trust, and you have always trusted me. You know I love you. You are my heroes.

Most of all however, I would like to thank my Anna. I love you. *This is your song.*

Mats Kvarnström, January 2005

Contents

1 Introduction	1
1.1 Guide for the Thesis	5
2 Data	9
2.1 Instrument setup	9
2.2 Sequence images	10
2.3 Z-scans	15
2.4 Use of the z-scans in Paper I and II	19
3 Position Estimation	21
3.1 Model of the appearance of pixels in the images	21
3.2 Estimating particle positions in 2-D	22
3.3 Estimating the depth	28
3.4 2-D template surface for sub-index estimation	32

4 Tracking	37
4.1 Linking positions to trajectories	38
4.2 Handling partial occlusion	39
4.3 Results for tracking two particles	43
4.4 Modifications	45
5 Supplementary studies	47
5.1 Dependence between the size of \mathcal{N}_t and the standard error in the 2-D estimation	47
5.2 Estimated intensity profiles for different bandwidths	51
6 Conclusions and Future Research	57
6.1 Depth estimation using templates	57
6.2 Measurements of diffusion coefficient and interaction	58
6.3 Automated depth calibration	60

Chapter 1

Introduction

The understanding of the behavior of colloidal suspensions is of crucial importance in a vast number of different areas. The standard theory for the interactions of colloidal particles, the DLVO-theory (see for example Evans and Wennerström (1999)), is merely an approximation, and experiments have shown that it fails to predict the behavior of several important suspensions, see for example Crocker and Grier (1994) and Grier (1998). Therefore, observations and measurements on the microscopic level are needed if we are to fully understand the behavior of colloidal systems.

Examples of common, everyday life colloidal systems are milk and paint. In milk, various interactions between the small (100 nm to 1 μm in diameter) colloidal milk fat particles and proteins suspended in the fluid, decide whether it coagulates into cheese or yoghurt. These interactions depend on how the milk was treated before the coagulation. For the second example, the pigments in the paint must stay suspended in the liquid in a can for years, yet, as they are spread on a wall, be able to coagulate fast. Another example of important everyday colloids are pharmaceuticals.

This thesis presents techniques developed for the quantitative study of diffusing particles in a colloidal system using video microscopy. The prin-

cipal application is to pharmacy, where properties such as diffusion coefficients and interaction between particles are important factors when formulating drugs. Here, possible modifications of for example the surface characteristics of the colloids are believed to have a large impact on modern therapies such as oral vaccines. However, the effects of such modifications need to be quantitatively measured and verified.

The idea is to make inference on properties of a colloidal particle system, such as diffusion coefficient of the particle, from a series of light microscope images of moving latex spheres. Figure 1.1 illustrates an example of what an image from such a sequence may look like. The particles in these image are spherical, made of latex (polystyrene), and have all a diameter of 494 nm. Each image consists of 512 times 512 square pixels with a side-length of 0.18 μm . The reason for studying latex spheres, and in particular, of a single size, are that suspension of these kind of polymer colloids can be used to simulate many features of a colloidal system by varying the solvent and salt concentration. Therefore they are widely used for studying the behavior of colloidal suspensions, see Evans and Wennerström (1999, chap. 9). For a recent review of different kinds of microscopy used for colloids, see Elliot and Poon (2001).

The apparent differences in size and brightness variations of the particles are due to different depths relative to the focal plane. Particles in the focal plane are depicted as small, distinct, black spots, while particles above or below the focal plane, are either light or dark in middle, respectively. Also, the further away from the focal plane a particle is, the larger and more blurred it appears. This out-of-focus effect will give us a method to estimate the depth of a particle. It should be mentioned that the light is practically considered as coherent, which is the reason for this optical effect. If light would have been incoherent, particles off-focus would simply be blurred.

This thesis introduces precise methods for estimating the locations (including depth) of particles in images like the one in Figure 1.1. A highly precise method for estimating particle centers in the image plane (that is, the horizontal and vertical coordinates) is presented in Paper I. The stan-

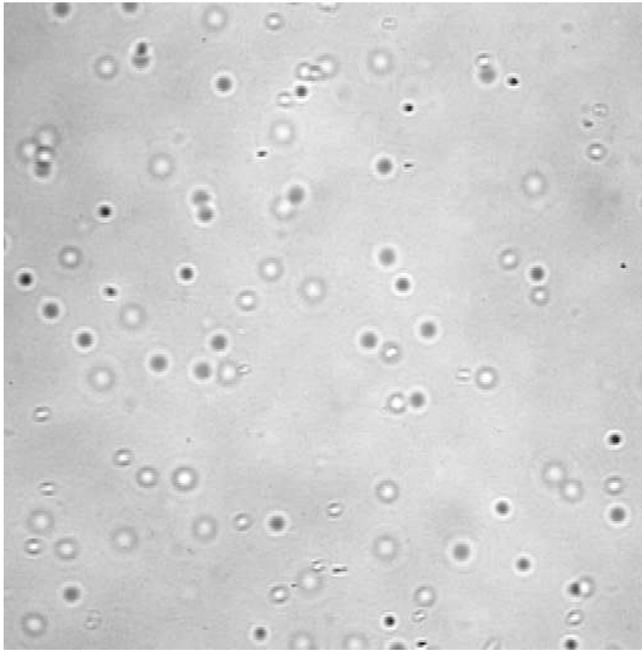


Figure 1.1: A single microscope image in a sequence of images. The particles are all equal in size and the difference in the appearances of the particles in the image is an optical effect of particles being at different depths relative to the focal plane.

Standard errors of the estimates are between 0.02 and 0.10 pixels, depending on the appearance of the particle, with lower values for particles closer to the focal plane. The method makes use of the rotational symmetry of the appearances of the particles in the image and the accuracy is well less than the sub-pixel level, by which we mean that the standard deviation of the

positioning error is less than one pixel. Standard errors for the position estimates are also estimated for each particle. Furthermore, we present a method for nonparametric function estimation when the response values are censored. This is needed since pixel values in the images are censored above a certain level. The depth is estimated using a template matching approach, covered in Paper II. The templates are empirically constructed using images of particles at known relative depth to the focal plane. Each template represents the appearance of a particle at a certain depth z and the corresponding depth between the templates is $0.2\mu\text{m}$. In Paper III we consider the estimation of the diffusion coefficient given a set of particle trajectories observed with measurement noise. However, since some particles seem to be fixed, a model with two kinds of particles, fixed and diffusing, is introduced. This is the typical situation for trajectories of particles in images like the one in Figure 1.1. Instead of manually discarding the particles which are fixed, the model permits them to be used in the estimation, which is the appropriate approach since they contain information on the measurement noise.

The latex particles in the sequence images like the one in Figure 1.1 have been treated in such a way that the suspended particles can be assumed to perform Brownian motion. This is the case at least for particles not too close to the confining glasses of the specimen, since the glass surfaces act attractively on the particles. These sequence images were solely constructed so that estimated properties can be easily verified, since the true properties of the suspension are known.

Methods commonly used for measurements on colloidal suspensions make collective measurements of properties for the entire sample, by which we mean that they cannot measure properties of each single colloid particle. Rather, the properties related to the entire population of colloids in the suspension is measured. Examples of such methods are various light-scattering techniques, for example dynamic light scattering (DLS). Nuclear magnetic resonance (NMR) and neutron scattering, are also used, see Evans and Wennerström (1999) and the references therein.

Previous work using digital video microscopy for quantitative colloidal

studies, have been made by Crocker and Grier (1996, 1998). There however, the depth of focus was ± 500 nm, which makes the particles similar in appearance and therefore easier to find in the images, since the sought-after objects in each image are similar. In their study, each particle is a bright spherical set of pixels and they achieve sub-pixel accuracy by calculating the geometric center of the brightness-weighted centroid, achieving standard errors of about 0.1 pixels (where each pixel is 85 nm). The shallow focal depth also restricts their methods to measurements in colloids confined to a crystallized structure. In our case, the depth of focus is $\pm 15\mu\text{m}$, resulting in a much wider variety of appearances of the particles in the image. The strength of the methods for particle position estimation developed here, is that our method is not restricted to particles similar in appearance. This opens up for position estimation for particles in a much wider range of depths. What we need however, is rotational symmetry of the appearance of particles in the images.

1.1 Guide for the Thesis

Preferred reading order

In Chapter 2, various aspects of the images used in this thesis are explained. Chapter 3 covers the main ideas as well as some extensions to the methods of estimating the particle locations developed in Paper I and II. In Chapter 4 we propose how the positioning methods can be used for the tracking of particles in three dimensions. Problems are also highlighted and possible modifications are discussed. Chapter 5 consists of two supplementary simulation studies related to Paper I.

If a quick start is preferred, Paper I and II should be read first and then Chapters 2 through 5. There is no preferred order of when to read Paper III.

Paper I: Estimating centers and intensity profiles of spherical particles in microscopy

This paper deals with estimation of particle position in the image plane. The underlying assumption for the estimation is the rotational symmetry of the appearance of particles. For a particle with true center at $x \in \mathbb{R}^2$, the main idea is to use the minimizer of

$$S(y) = \min_{f \in \mathcal{C}^2} \sum_{i \in \mathcal{N}_x} \{I_i - f(r_i(y))\}^2 = \sum_{i \in \mathcal{N}_x} \{I_i - \hat{f}(r_i(y))\}^2 \quad (1.1)$$

for $y \in \mathbb{R}^2$, where I_i for $i \in \mathcal{N}_x$ are the pixel values in a neighborhood of pixel locations close to x and $r_i(y)$ the distance between the candidate center y and pixel location i . Furthermore, \mathcal{C}^2 is the set of functions $f : \mathbb{R} \rightarrow \mathbb{R}$ with second order continuous derivative and symmetric in r . We calculate \hat{f} using a local quadratic kernel estimator with (appropriately chosen) bandwidth h . The idea behind minimizing the equation above is to find the position of (local) maximum rotational symmetry.

We also present a method for estimating the standard error of each particle center estimate, by using a sandwich estimator, see for example Owen (2001). A simulation study shows that these standard error estimates are consistent with the observed root-mean-square errors of the position estimates. The standard errors of the particle position estimates depend on the depth of the particle, and are in the range of 0.02 to 0.10 pixels, with larger values for depths further away from the focal plane. This is comparable to the results of Crocker and Grier (1996), however our method is not confined to particles of similar appearance.

A method for nonparametric function estimation when the response values are censored at a fixed level, is also introduced in this paper. We modify the local quadratic kernel estimate to take care of the censored values under the assumption of normally distributed observation errors. We have not found this anywhere else in the literature.

Supplementary studies for this paper can be found in Chapter 5 where we investigate the bias and the mean squared error when estimating the

intensity profiles, in particular when the bandwidth is varied. Furthermore in Chapter 5, the way the standard errors depend on the size of the neighborhood \mathcal{N}_x , for particles at different depths, is investigated.

Paper II: Depth estimation of colloidal particles in microscopy

Having found the particle center, the depth is estimated by comparing the pixel values I_i and distances $r_i(y)$, with templates (intensity profiles) of the appearance of particles at different depths. The templates are constructed by estimating the intensity profiles of a particle at known distances relative to the focal plane. The distance in depth between each template is $0.20\mu\text{m}$. The matching criterion we propose here takes care of different background intensities and possible censoring of pixel values, both of which are important features of the images considered.

This empirical approach to template construction was chosen since theoretical derivation of the appearance of the latex particles at different depths seems difficult. For example, a ray-tracing methods such as Fourier optics, used by for example Young et al. (1998) for constructing templates in DIC microscopy, does not work here since the wavelength of light is in the same order of magnitude as the size of the particles. An alternative and more advanced approach to Fourier optics is to use Mie-theory, which was used by Ovryn and Izen (2000) to predict the appearance of a polystyrene sphere of diameter $7\mu\text{m}$. However this is 14 times bigger than the particles considered here and it is unclear to what extent this approach can be applied to our particles. In fact, the imaging of spherical objects is still a topic of large theoretical interest in the optics research society.

Other particles at known depths were used to validate the estimation procedure. The root-mean-square error is concluded to be at least in the order of the distance in depth between the templates, that is $0.2\mu\text{m}$, at least for particles not too far away from the focal plane.

Paper III: Estimation of the diffusion coefficient in a mixture model

In Paper III we estimate the diffusion coefficient given a set of particle trajectories performing Brownian motion, observed under measurement noise. However, since some particles seems to be fixed, a model is introduced with two kinds of particles, fixed and diffusing. We regard the problem as an incomplete data problem since we do not know *a priori* which particles are really diffusing. The maximum likelihood estimator is computed via the EM algorithm, see Dempster et al. (1977), and it is shown to be strongly consistent and asymptotically normal, as the number of particles approaches infinity, under a reasonable restriction on the parameter space. A simulation study shows that the method is robust even for large measurement errors, and that the estimated parameters are approximately normally distributed even for small sample sizes.

The position estimates of the particles used in this paper are integer valued. They were estimated using a filtering technique called the rotational Hough Transform which is a common tool in image processing used for circle detection in images. See for example Gonzales and Woods (2002) or Kerbyson and Atherton (1995). For the full details on estimating these trajectories, the reader is referred to Kvarnström (2002). The reason for not using the methods developed in Paper I and II for the trajectories in this paper, is that Paper III was written first.

Chapter 2

Data

There are two kinds of images considered in this thesis, *sequence images* and *z-scans*. Sequence images are the ones used for inference on properties of the colloidal particles. An example of such an image is given in Figure 1.1. Z-scans are images of particles at known relative depths from the focal plane and they are mainly used for constructing the templates used in the depth estimation. In this chapter, we will explain the various aspects of the images together with the general setup of how the images were constructed.

2.1 Instrument setup

Latex particles made of polystyrene with a diameter of 494nm were placed between an objective and a cover glass and sealed. The illumination consisted of coherent light. The sample was studied in a Zeiss Axiovert 135 TV microscope equipped with a Newicon video camera. The video signal was then digitized and stored as TIF files.

Pixel values are stored as unsigned integers in 8 bits. The pixels assumes integers between 0 and 255 which are interpreted as gray scale intensity

values. This means that zero means black while 255 means white. For pixel values in between, the larger the magnitude, the brighter the shade of gray.

An important effect of this truncation to integer values is that we get censoring of pixel values above 255. This is dealt with both in the estimation of particle centers in Paper I, and the template matching procedure of finding the depth in Paper II. In Figure 2.1 we illustrate censoring by zooming in on two particles from Figure 1.1. Below each image, the pixel values surrounding the estimated particle center (using the method of finding the position of maximal rotational symmetry from Paper I) are plotted versus their corresponding distances to the center. We do not have censoring for pixel values below 0, however there seems to be a lower limit of pixel values around 30-35. If this is censoring or not is not known.

2.2 Sequence images

The image in Figure 1.1 is an example of what a sequence image looks like. These are the kind of images that will be used to make inference on the properties of the colloidal system of particles. On the left in Figure 2.2 we have zoomed in on the middle region of size 256 times 256 pixels of the sequence image in Figure 1.1. To the right of this, the same region in the next consecutive image in the sequence is shown. The movement in the image plane of the particles between two consecutive images are on the scale of a few pixels. In Figure 2.3 we display the difference between the two images and if it was not apparent from Figure 2.2, we see here that most of the particles have moved.

Each image in the video sequence consists of 512 times 512 pixels. Each pixel has a side length of 180 nm. The focal plane is set at a depth approximately between the cover and the specimen glass of the sample specimen. The maximal difference in depth (relative to the focal plane) is believed to be $15\mu\text{m}$. Therefore, the domain in which the particles are confined, and are available for our inspection through the images,

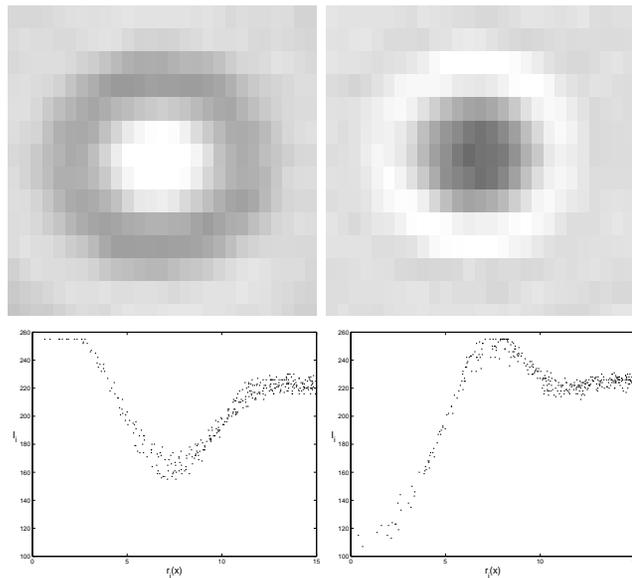


Figure 2.1: Zooming in on two particles in Figure 1.1 to illustrate censoring of pixel values at 255. Below each image the pixel values surrounding the estimated particle center are plotted versus the distances to the estimated center. This figure also demonstrates that the rotational symmetry assumption of pixel values surrounding a particle center is reasonable.

is a box with equal length of the sides of about $90\mu\text{m}$, and a depth of (approximately) $30\mu\text{m}$.

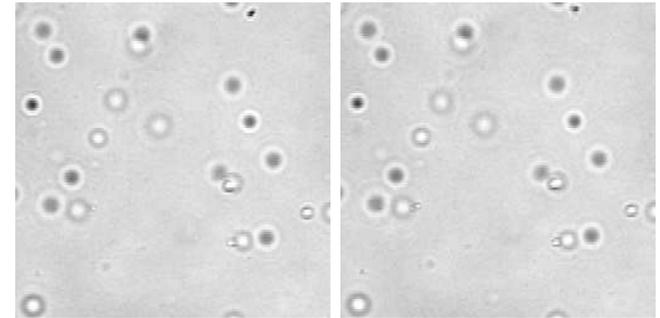


Figure 2.2: Magnified part of two consecutive sequence images. The displacement in the image plane of the particles between two consecutive images is on the scale of a few pixels. In Figure 2.3 we display the difference between the two images.

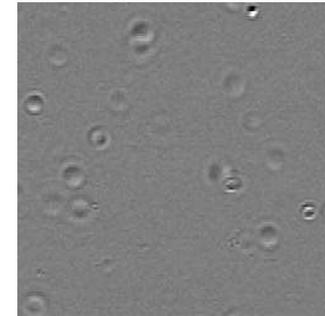


Figure 2.3: The difference between the two images in Figure 2.2. Mid-gray represents zero. Notice that movement in depth is also apparent for some particles.

Even and odd frames

The images are recorded at video rate, which is 50 images (or frames) per second. In practice however, only half of the rows in each image contain

new information. The camera records only half of the rows at each scan, alternating between the even and odd rows (also called the even and the odd fields) and duplicates this information to the rows which were not scanned. The images are called even and odd frames, respectively, depending on whether the even or the odd rows were updated. The images of Figure 2.4 show the same region zoomed in, for two consecutive frames, one even and one odd. Note that the pixels look rectangular, this is because of the duplication of pixel values from the rows that was updated to the ones that were not.

There are three possibilities on how to deal with even and odd frames in the image sequence; interlacing, interpolating, or raw images. Interlaced means that two consecutive frames (one even and one odd) are interlaced into a single image, using the even rows from the even frame and the odd from the odd frame. However, since there is an interval of a $1/50$ of a second between the even and the odd frame, this will cause problems when we are observing moving particles, which is the case here. Figure 2.4 illustrates this problem with interlacing two consecutive frames. Clearly, interlacing is not a good idea when observing moving particles. An alternative is to interpolate the pixel values. We use the updated rows (that is, even rows for even frames) and interpolate these to the pixel values on the non-updated rows (that is, odd rows, for even frames). In the bottom left image of Figure 2.4 we have displayed the linearly interpolated version of the even frame (located straight above in the figure). However, what interpolating does, is just making the image look nicer to the eye; we do not add information, rather, if anything, we distort the information.

For sequence images we will therefore use the *raw* images, by which we mean that we use each frame separately. Practically each frame is an image consisting of 256 times 512 pixel values. The coordinates (pixel locations) to these pixel values then alternates between the two pixel locations. See the upper images of Figure 2.4. The important thing is therefore to keep track of the location to which each pixel value corresponds to; the upper pixel for the even images and the lower for the odd images. This is important since it is these locations that correspond to the physical locations which we are interested in measuring.

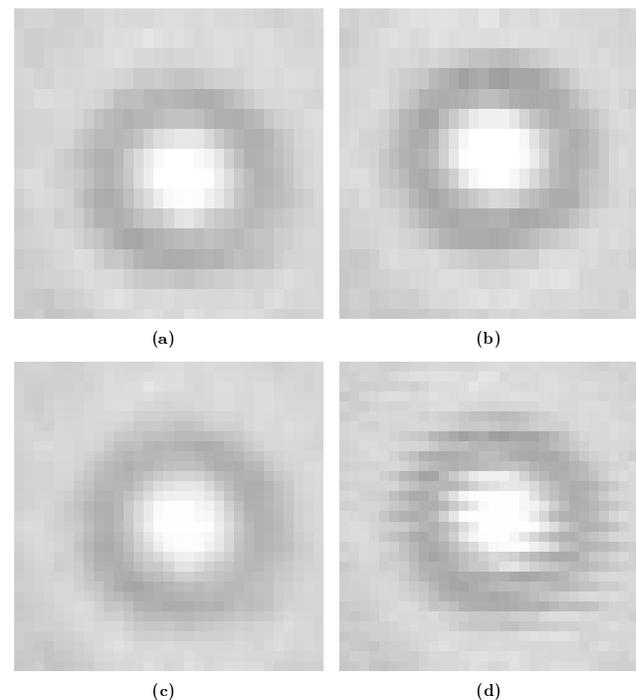


Figure 2.4: Illustration of even and odd frames (images) and the problem with interlacing two consecutive frames. The two upper images are the same area zoomed in for two consecutive frames, a) is an even image and b) is an odd image. Note the duplication of pixel values between rows, making the pixels to look rectangular. The interpolated version of a) is shown in c) and the resulting interlaced version, using the even frame for even rows and the odd frame for the odd rows, is shown in d).

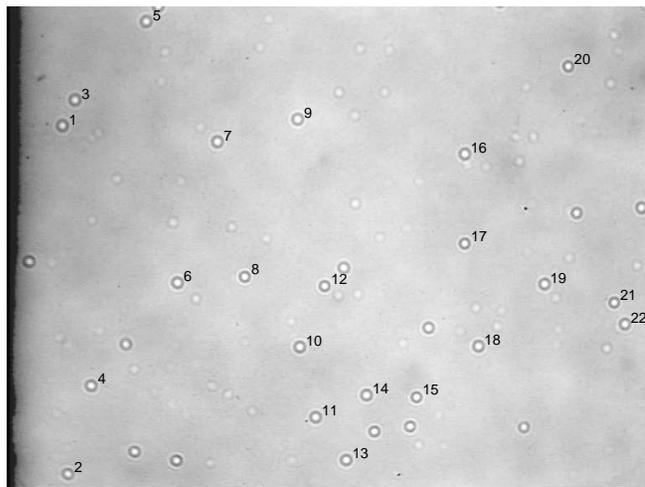


Figure 2.5: Z-scan of particles $3.0\mu\text{m}$ below the focal plane. The 22 labelled particles are the ones which are adsorbed on one of the confining glasses of the specimen and present in all 73 z-scans. The other particles in the image is either moving or not at the same depth relative to the focal plane as the labelled particles.

2.3 Z-scans

In order to know what particles look like at various depths from the focal plane, *z-scans* were constructed. In Figure 2.5 we display the *z-scan* at depth approximately $3.0\mu\text{m}$ below the focal plane. We have 73 *z-scans* at our disposal ranging from $7.2\mu\text{m}$ below to $7.2\mu\text{m}$ above the focal plane. The distance in depth between two consecutive *z-scans* is $0.2\mu\text{m}$.

The *z-scans* have been constructed by letting particles adsorb on the cover

glass surface of the specimen, and then the specimen was moved relative to the optics of the microscope. However, there are other particles than the adsorbed ones present in the *z-scans*. In Figure 2.5 the particles labelled with numbers are adsorbed on the glass. These 22 particles are the ones that were practically fixed in position through all *z-scans*. (We write "practically" since they are moving slightly, about 2 pixels throughout the entire sequence of *z-scans*.) As seen, there are several other particles present, some of which are moving but also some which are fixed in position but not adsorbed on the cover glass. Since the *z-scans* are used to depict fixed particles, they are presented in interlaced format; we will see that the interlacing effect is visible for moving particles.

In Figure 2.6 we have zoomed in on the region containing the particles labelled 6, 8, 10, 11, and 12 for the *z-scan* in Figure 2.5 and for three other *z-scans*. Note that particles below the focal plane are bright in the middle and particles above the focal plane are dark in the middle.

Here it is clear that other particles are present in the *z-scans*. We also see the same kind of interlacing effect for moving particles as we saw in Figure 2.4; look at the particle to the right above particle number 12 in image a). Furthermore, there are particles that seem to be fixed, but at another depth; these are the more vague particles, for example to the right below particles 6 and 12. These are particles adsorbed somewhere else in the specimen, possibly on the outer surface of the glass. Note also the black dot to the left of particle number 11, which looks the same throughout all *z-scans* and is probably a stain or defect in the optics. In c) we also see that mobile particles in the specimen sometimes occlude the 22 fixed particles; see particles number 8 and 11 in c).

Most importantly however, the 22 labelled particles in Figure 2.5 seem to be at a slightly different depth relative to each other. In Figure 2.6 this is clearly visible in the *z-scan* b), corresponding to particles at the focal plane, and in *z-scan* d). Particles 8 and 12 seem to be slightly more above the focal plane (since they are larger and dark in the middle than the others). The same is true for all *z-scans* and this also applies to the particles labelled 15, 17, 20 and 21 in Figure 2.5. This was also validated

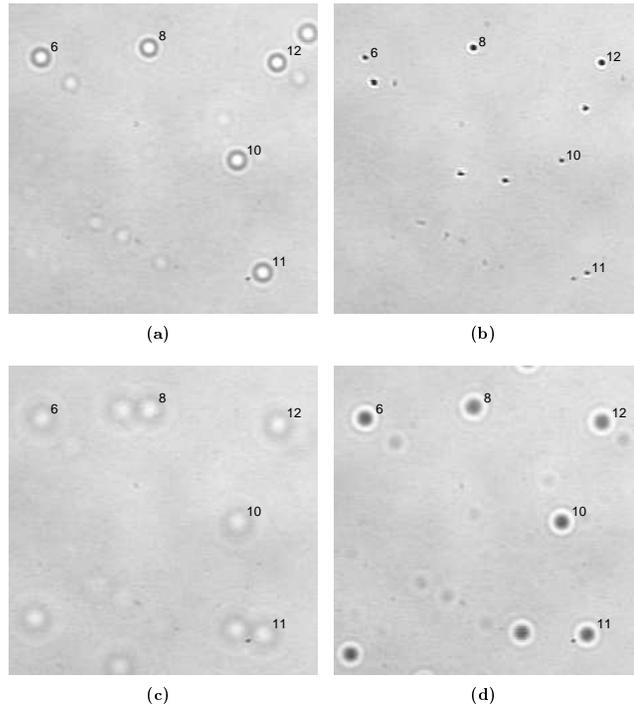


Figure 2.6: Zooming in on four z-scans. a) is the same z-scan as in Figure 2.5 and b) is the z-scan corresponding to particles at the focal plane. c) and d) are z-scans corresponding to $7.2\mu\text{m}$ below the focal plane and $3.0\mu\text{m}$ above the focal plane. Note that there are other particles present and that the labelled particles are (partially) occluded for some z-scans.

when estimating the depth of the particles in the z-scan in Paper II, see below.

This observation is important when the z-scans are used for constructing templates for depth estimation, and also when validating the performance of the depth estimation. The templates for depth estimation in Paper II, were constructed using mainly the particle labelled number 6. In z-scans where particle 6 was occluded by a moving particle, particle 14 was used. This way, we constructed templates of the appearance of particles at depth indexed by the z-scans, enabling us to estimate the depth of the rest of the particles in the z-scans by comparing the appearance of the particles to the templates. Of course, since we only had templates for particle appearance indexed by the z-scans, the precision in the depth estimation will be limited by the distance in depth between the z-scans, at least if no other assumptions are made. Note that since we only have a finite number of templates, estimating the depth this way is a kind of a classification problem.

When estimating the depths in the z-scans, we raised doubts about if all adsorbed particles in fact were at the same depth. The particles labelled 8, 10, 15, 17, 20, and 21 were easily recognized as having an offset in depth relative to the particles 6 and 14, which were the ones used for template construction. However, there were also some disturbing depth estimates for other particles. The grounds for these doubts were that when the depth of supposedly adsorbed particles were estimated, there seemed to be a systematic error in their estimated depths in the order of one z-scan above or below the depth of the template. (See also the next section.) One should keep in mind that the particles are approximately $0.5\mu\text{m}$ in diameter and the distance between z-scans are $0.2\mu\text{m}$. Hence, the distance between two consecutive z-scans is smaller than the radius of the particles. Therefore, some fluctuations in estimated depth may be accounted to the uncertainties in the true actual depth of the particles in the z-scans.

2.4 Use of the z-scans in Paper I and II

In Paper I we restricted the number of depths considered in the particle position estimation. We used the z-scan of particles at the focal plane together with every third z-scan below and above, up to a maximal displacement in depth of $4.2\mu\text{m}$. We indexed these from -7 to 7. Particle number 6 was used when constructing the true intensity profiles used in the Simulation Study (it is also this particle that is displayed in Figure 2 in Paper I) and in the Result section, particles 6, 7, 13, 14, 19 and 19 were used when estimating the standard errors for real data.

In Paper II we constructed the templates using the particles labelled 6 and 14. We used 61 z-scans (of the total amount of 73), from $6\mu\text{m}$ below, to $6\mu\text{m}$ above the focal plane, and they were indexed from -30 to +30. In a pre-study, we estimated the depth for all non-occluded particles in all z-scans and by looking at the median of the offset in estimated depth, relative to the template particles and calculated over all depths, three categories of particles stood out; the six particle mentioned above, which were 2 z-indices, that is $0.4\mu\text{m}$, above the template; particle 1, 3, 5 and 9 which were $0.2\mu\text{m}$ below the template; the remaining 12 particles had median offset equal to zero, relative to the template particle. It was this latter category that was used in the Results section of Paper II, however, particles 2 and 4 were not part of the study.

Chapter 3

Position Estimation

3.1 Model of the appearance of pixels in the images

We denote an image by I . This is effectively a matrix of pixel values I_i for pixel locations $i = (i_1, i_2) \in \mathcal{D}_I \subset \mathbb{Z}^2$, where \mathcal{D}_I is the set of pixel locations i for which the image is defined. We will use the terms pixel and pixel location interchangeable to mean the same thing when no risk of ambiguity exists.

A particle center is denoted by $x = (x_1, x_2) \in \mathbb{R}^2$, and to each particle, we associate a set $\mathcal{N}_x \subset \mathcal{D}_I$ of pixels in the image I called the *neighborhood of the particle at x* . Typically, we let

$$\mathcal{N}_x = \{i \in \mathcal{D}_I : r_i(x) \leq r_{max}\}, \quad (3.1)$$

where $r_i(x)$ is the Euclidean distance from the particle center x to the pixel location i , and r_{max} is an appropriately chosen distance.

The main assumption is that, for a particle center at $x \in \mathbb{R}^2$ and at depth

$z \in \mathbb{R}$, we have

$$I_i = f_z(r_i(x)) + \alpha + \epsilon_i \quad \text{for } i \in \mathcal{N}_x, \quad (3.2)$$

where f_z is called the *intensity profile* for depth z . The image noise, ϵ_i for $i \in \mathcal{D}_I$, is assumed to be normally distributed with isotropically correlated pixel values. The intensity profile is furthermore assumed to be a smooth function $f : \mathbb{R} \mapsto \mathbb{R}$ with at least two continuous derivatives and symmetric in r . Furthermore, $\alpha \in \mathbb{R}$ corresponds to the background intensity in the image and this is generally different for each particle. This is an important factor to take into account when estimating the depth.

In Figure 3.1 we display zoomed-in sub-images of what the particles look like at different depths. The true depth between to consecutive indices is $0.2\mu\text{m}$ and index 0 represents the focal plane. These sub-images are from the z -scans. The rotational symmetry assumption seems reasonable, at least for particles not too close to the focal plane.

3.2 Estimating particle positions in 2-D

The idea for estimating the particle center in 2-D goes as follows. For a particle located at $x \in \mathbb{R}^2$, we use the minimizer of equation (1.1) repeated here for convenience

$$S(y) = \min_{f \in \mathcal{C}^2} \sum_{i \in \mathcal{N}_x} \{I_i - f(r_i(y))\}^2 = \sum_{i \in \mathcal{N}_x} \{I_i - \hat{f}(r_i(y))\}^2$$

for $y \in \mathbb{R}^2$, as an estimate of x . The idea behind minimizing equation (1.1), is to find the position of (local) maximal rotational symmetry. This method of estimating the particle center in the image plane to sub-pixel accuracy is dealt with in Paper I.

In practice, we calculate \hat{f} using a local quadratic kernel estimator with a Gaussian kernel with (appropriately chosen) bandwidth h . References on nonparametric function estimation are, for example Hastie and Tibshirani (1990), Fan and Gijbels (1996), or Györfi et al. (2002). In Paper I, we

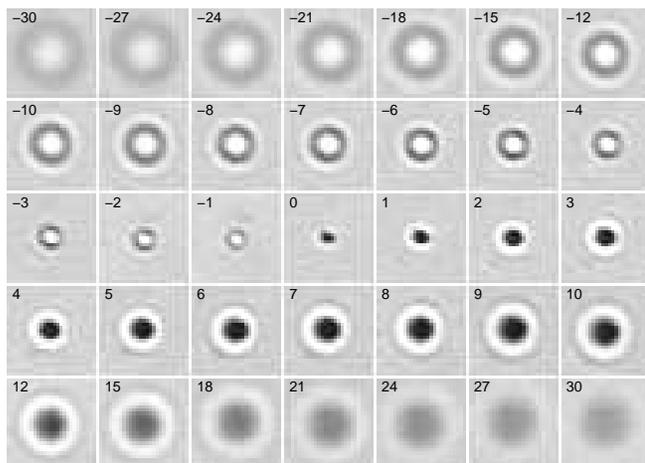


Figure 3.1: The appearances of particles at different depths relative to the focal plane. Sub-image 0 corresponds to a particle at the focal plane, and sub-images with negative and positive labels are below and above the focal plane, respectively. The distance in depth between two consecutive label numbers is $0.2\mu\text{m}$.

introduced a modification of the local quadratic kernel estimates when the response variables are censored above a certain (known) value. We call this censored regression. If nothing else is said however, when talking about the local quadratic kernel estimate, we mean the standard local quadratic kernel estimates, without taking care of censored pixel values. A simulation study, the results of which are presented in Paper I, showed that the difference in estimating the particle center is very small between using the censored and standard local quadratic method.

The reason for choosing the local quadratic kernel estimate as nonparametric method, instead of for example a spline smoother, is that the

estimate of the derivative of f is practically given to us for free using a local quadratic, and the derivative is needed when estimating the standard errors (see Paper I for details). The Gaussian kernel was chosen since it is smooth and has unbounded support, which makes the function S differentiable. Regarding the choice of bandwidth, simulation studies have shown that the choice of bandwidth is not very important when estimating the particle centers (see Chapter 5). This is good, since one of the main ideas with this method (that is, minimizing equation (1.1)) of estimating particle centers, is that it should be applicable to particles with different appearances, as long as they look rotational symmetric in the image. The idea presented in Paper I was to use a *pilot bandwidth* h_{pilot} of 0.7 to find a first, preliminary particle center and then update this bandwidth with a cross-validation study. Then the final center estimate is calculated using the updated bandwidth.

Regarding the neighborhoods, we will in Chapter 5 conduct a study of how the standard errors in the 2-D estimation depends on r_{max} . Also, when two or more particles are close together however, the assumption in (3.2) these circular neighborhoods with a fixed r_{max} does not apply, if the neighborhoods of the particles intersect. In Chapter 4, we will present a way to circumvent this, by adaptively selecting the shape of the neighborhood according to nearby particles. Another aspect concerning the choice of neighborhood \mathcal{N}_x by (3.1), is that it depends on the unknown center x . Nevertheless, if we are given an approximate particle center y_0 , we let $\mathcal{N}_x = \mathcal{N}_{y_0}$ where \mathcal{N}_{y_0} is defined by equation (3.1) for $x = y_0$. Approximate particle centers can be given either manually, or by some automatic image analysis method. Below, we will present one such automatic method.

Candidate particle centers

Before we can estimate the particle center by minimizing the criterion (1.1), dealt with in Paper I, we need to have a first approximate position. We call these approximate positions candidate particle centers. Candidate particle center are usually integer valued positions if they are the results

from an image analysis stage using some filtering technique (with appropriate post-processing). For tracking in sequence images however, we will use the position estimates in the previous image as candidate particle centers and this is dealt with in Chapter 4. Below, we will present one idea to a filtering technique for getting candidate particle centers.

Local maximal rotational symmetry

We will here introduce a (non-linear) filter that could be used to find positions of local maximal rotational symmetry. It is also presented to give the flavor of the difficulties one run into when trying to automatically find the objects of interest in an image, particularly when the sought-after objects are different in appearance.

The main computational effort in the minimization of (1.1), is spent on calculating the nonparametric estimate of the intensity profile f at each candidate position y . In particular, much of the effort is spent on calculating the distances $r_i(y)$ and the inter-distances between these, which are needed when calculating the weights in the *equivalent kernels*. Each estimate $\hat{f}(r_i(y))$ in a local quadratic kernel estimate can be written as a linear combination of response values (pixel values):

$$\hat{f}(r_i(y)) = \sum_{j \in \mathcal{N}_y} W_{ij} I_j$$

for all i . The i th row of the matrix W is the equivalent kernel for the estimated value at point $r_i(y)$. The elements in the matrix W only depend on $r_i(y)$ and the bandwidth h . (This is however not true if we use the modified version of the local quadratic kernel estimate, that takes care of censored pixel values.)

If $y \in \mathcal{D}_I$, that is, if the candidate center is an (integer valued) pixel location, and we use \mathcal{N}_y as neighborhood, the set of distances $r_i(y)$ for $i \in \mathcal{N}_y$ are the same for all $y \in \mathcal{D}_I$ (apart of course from pixels near the boundary of the image). Therefore, since the matrix of equivalent kernels depend only on $r_i(y)$, the matrix W only has to be calculated once. This

speeds up things considerable. Assuming there are n pixels in \mathcal{N}_y , the calculation of W takes $\mathcal{O}(n^2)$ multiplications into account, which for the original minimization of equation (1.1) has to be done for each y since the distances $r_i(y)$ are different for each $y \in \mathbb{R}^2$. Given the matrix W , the calculation of $S(y)$ takes n^2 multiplications.

Note that this approach is different from the method of minimizing (1.1), in the way that \mathcal{N}_y here changes with y . In (1.1) we first fix the neighborhood \mathcal{N}_x and then search for the minimizer of $S(y)$. Therefore, in order to separate the two sums, we denote by S_{pix} , the pixel-wise calculation of S introduced above

$$S_{pix}(y) = \sum_{i \in \mathcal{N}_y} \{I_i - \hat{f}(r_i(y))\}^2 = \sum_{i \in \mathcal{N}_y} \left\{ I_i - \sum_{j \in \mathcal{N}_y} W_{ij} I_j \right\}^2 \quad (3.3)$$

for $y \in \mathcal{D}_I$. The matrix W does not depend on either y or the pixel values I_i which is the important fact about S_{pix} . It depends only on the bandwidth h and the radius of the neighborhood r_{max} .

In Figure 3.2a we display a sub-region of a sequence image. In b) we display S_{pix} computed for this image with $r_{max} = 6$ and bandwidth $h = 0.7$. Sub-figure c) is the same as b) but displayed as a contour plot. The idea is to use the local minima of S_{pix} as candidate particle centers.

Now we can use morphological operations to find the local minima. The classical reference of morphology in image analysis is Serra (1982). To extract the minima, we will preform an operation called *bottom hat*. It is defined as follows. First we define a *structuring element* B , which we here let be a ball in the image plane of radius r . For an image J , the *closing* J^B of J using structure element B is defined as

$$J^B(i) = \min_{k \in B_i} \{ \max_{m \in B_k} J(m) \}$$

where B_i is the structure element centered at i . Finally, the bottom hat of J is defined as the difference

$$J_{bottom\ hat} = J^B - J$$

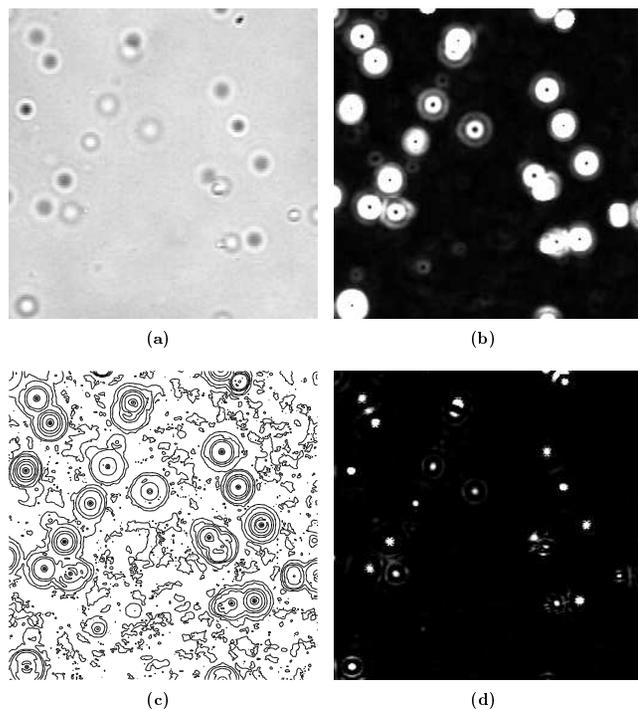


Figure 3.2: A sub-region of an image sequence a), together with its S_{pix} in b). In c) we display the logarithm of S_{pix} as a contour plot instead, making it somewhat easier to localize the local minima in S_{pix} , which are to be used as candidate particle centers. In d) we show the result of a morphological operation called bottom hat to S_{pix} . In Figure 3.3, the result after thresholding the image in d) can be seen.

between the closing J^B and the original image J . In Figure 3.2d, the bottom hat of S_{pix} in Figure 3.2b is shown, using a ball with radius 3 as structuring element (that is, all pixels within radius 3 from origo).

All values in the bottom-hat filtered S_{pix} that are below a certain threshold t are set to zero, and the largest element in each connected component in this thresholded image is denoted a candidate particle center. In Figure 3.3, we have plotted the original sub-image from Figure 3.2, together with the candidate particle centers using threshold $t = 10$. As seen from the figure, the operation with finding the positions local rotational symmetry works fairly well, most of the true particles have been found and only a few false particles were found. The true particles that were missed, were all particles close to other particles. The reason for this is of course that particles close together, disrupt the rotational symmetry.

One problem with using a filtering method like the one illustrated above, is the large number of parameters, which were here chosen more or less *ad hoc*. The parameters here are four: the bandwidth h , the size of the neighborhood r_{max} , the radius r of the structuring element, and finally, the threshold t .

Even though a filtering step like this is not good enough for finding all particles present in an image, it could be used as a "watch-dog" to look for "intruder particles" that comes into the image domain as we are tracking particles.

3.3 Estimating the depth

The idea for estimating the depth, as presented in Paper II, goes as follows. After a particle center has been estimated in the image plane, we estimate the depth by comparing the pixel values I_i and their distances $r_i(\hat{x})$ from the estimated center \hat{x} with a set of templates intensity profiles of what particles look like at different depths. The corresponding depth of the one that gives the best correspondence is the estimate of the depth of the particle. Consequently, what it comes down to, is to construct

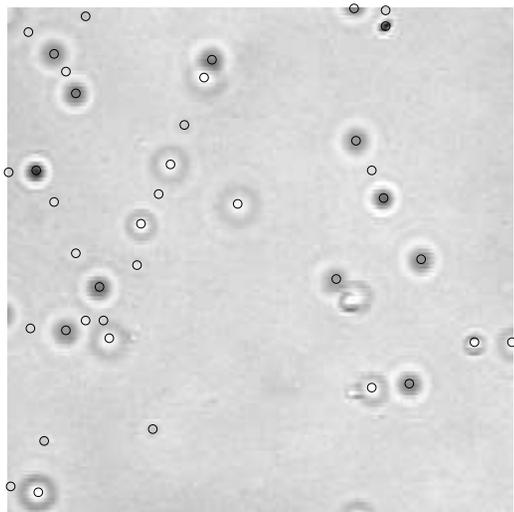


Figure 3.3: The resulting candidate particle centers after thresholding the bottom-hat filtered S_{pix} and taking the maxima in each connected component. The result is fairly good, the only true particles missed, are the ones that are close to other particles.

templates and to find a suitable measure on what best correspondence means.

The templates were constructed by estimating the intensity function of the adsorbed particles at different depths in the z-scans, see Section 2.3. Since the exact true depth of the particles in the z-scans were hard to verify (see the discussion in Section 2.4), we focused on using as few particles possible when constructing the templates. We used mainly the particle labelled number 6 in Figure 2.5 but for the z-scans in which this was occluded, particle 14 was used instead.

In Figure 3.1 we have displayed the appearance of particles at a subset of depths used for template construction. The template intensity function for these depths are plotted in Figure 3.4. These profiles were estimated using the modified version of the local quadratic kernel estimate, taking care of censored pixel values above 255. That this is the case, is evident from the observation in Figure 3.4 that the profiles are assigned values greater than the censoring limit 255. The template profile for depth z is denoted f_z .

To measure best correspondence between template and pixel values surrounding a particle that we want to estimate the depth of, we use the criterion function

$$M(z) = \frac{1}{\sigma^2} \sum_{i \in A_T} \{I_i - \hat{\alpha}_z - f_z(r_i(x))\}^2 - \sum_{i \in A_T^c} \log \left\{ \Phi \left(\frac{f_z(r_i(x)) + \hat{\alpha}_z - T}{\sigma} \right) \right\} \quad (3.4)$$

where $\hat{\alpha}_z$ is the minimizer of the expression (3.4) viewed as a function of both z and α , but where we keep z fixed. As seen, this criterion function takes care of both censored pixel values for the particle that we want to estimate the depth for, and, in fact more importantly, the different background intensities α (see the assumption regarding the appearances of particles in the images in equation (3.2)).

The main problem with this approach to depth estimation is first that we only have a finite set of template profiles, indexed from -30 to -30, thus limiting the precision by the corresponding depth between the indices of $0.2\mu\text{m}$. Also, since it is the cover glass that is moved relative to the optics of microscope when constructing the z-scans, it is important that the particles from which we estimate the templates, are at the same relative depth to the cover glass in all z-scans. Otherwise the distance in depth between each template will not be the same between the templates.

In Paper II, a simulation study showed a good precision in depth estimation, at least for particles within $3.0\mu\text{m}$ from the focal plane (corresponding to index -15 to 15 in Figures 3.1 and 3.4). For these depths, the depth was only misclassified for 14 simulations out of 25000. One objection to this highly optimistic result should be that the image noise seems to have

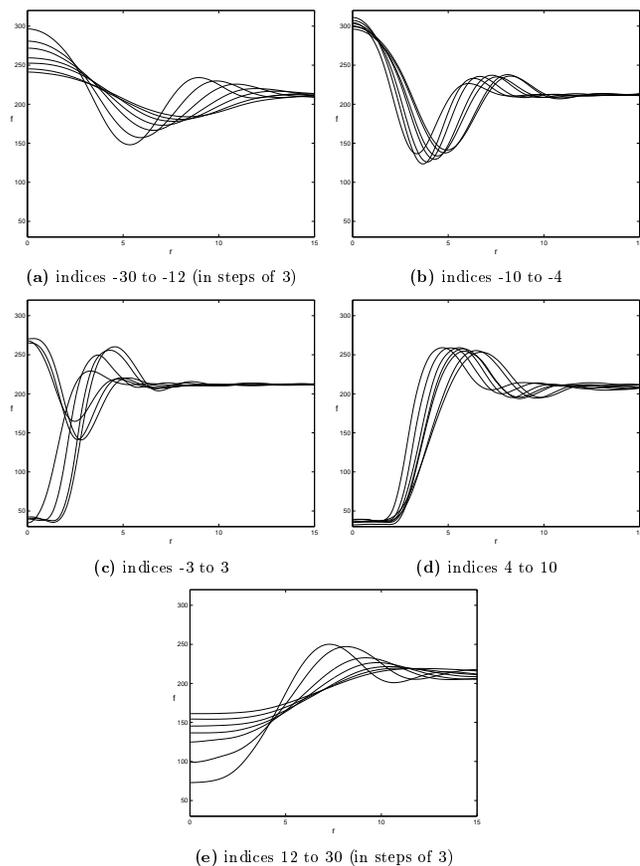


Figure 3.4: A subset of the template profiles used for template matching in the depth estimation. The distance between two consecutive indices is $0.2\mu\text{m}$. The sub-figures correspond to the rows in Figure 3.1.

larger variance for particles close to the focal plane, as reported in Paper I. Even so, compared with the result for the real images for these depths, the conclusion is that the template matching approach works well and the precision is at least within one z -index of the z -scans, corresponding to $\pm 0.2\mu\text{m}$. Compared with the standard errors for estimating the center in the plane, which were between 0.02 and 0.10 pixels ($3.6\text{nm} - 18\text{nm}$) is is of course much worse.

3.4 2-D template surface for sub-index estimation

We will here present an idea on how to construct a bivariate regression surface $f_z(r)$, as a function of both depth z and distance from center r simultaneously. This will lead us in to a discussion on how to measure the distances between two intensity profiles, which obviously is also a crucial matter in depth estimation. The matching criterion (3.4) presented above, is basically the L_2 -norm, but it is quite unsatisfactory from a theoretical point of view since it does not take special consideration of the functional features of the intensity profiles, such as for example the first stationary point of the intensity profile.

We record the pixel values I_i and corresponding distances $r_i(x)$ surrounding an adsorbed particle in the z -scans for each depth z we want to estimate the template for. To illustrate what a template surface might look like, we have in Figure 3.5 displayed the bivariate regression surface calculated via a local bilinear kernel estimate. No special care have here been taken to censored pixel values. In Figure 3.6 we present the template surface as an image instead.

The bandwidths for the bivariate regression are two; the first, h_r , is for the r -direction (in the same way as before), and the second, h_z is for the depth z . For the template surface in Figure 3.5, both bandwidths were a function of z . For the bandwidth in the r -direction, h_r , this is the same as we did in Paper I and II when estimating the intensity profiles for each

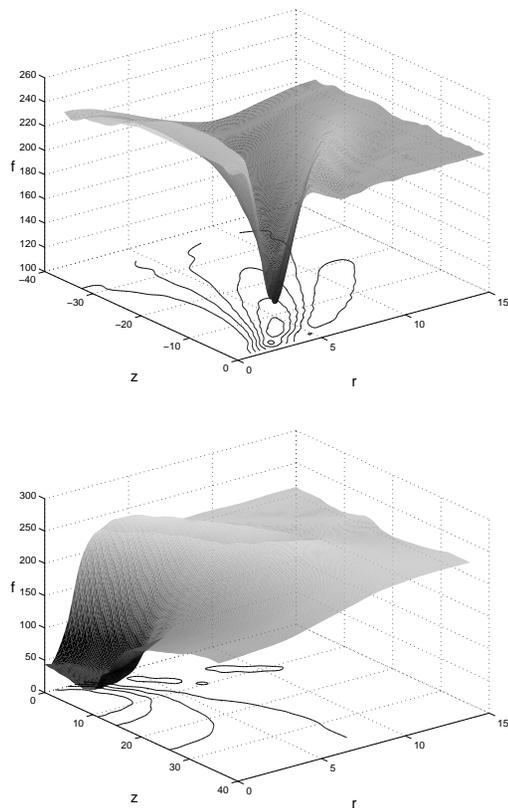


Figure 3.5: The two-dimensional template surface of intensity profiles constructed via a local bilinear kernel estimate. The estimation, as well as the presentation here, was split in two parts, one each for particles below and above the focal plane, respectively. The reason for this is the apparent phase shift at the focal plane. No consideration to censoring has been done here.

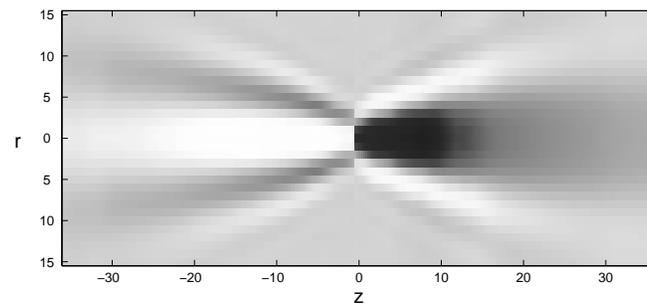


Figure 3.6: The same two-dimensional template surface as in Figure 3.5 presented as an image and reflected around $r = 0$.

depth; particles close to the focal plane need a small bandwidth h_r and vice versa for particles further away from the focal plane. The amount of smoothing in the z -direction is changed since the resemblance in appearance between particles at consecutive z -scans varies (see Figure 3.1 and Figure 3.4) with depth. For particles further away from the focal plane, smoothing between the depths can be quite large, whereas for particles close to focal plane, one must be careful not to smooth too much. This is an important point and has to do with how we measure distance in the functional space of intensity profiles, which we will get back to below. The most extreme occurrence of this, is obviously the difference in appearance between a particle at the focal plane compared to a particle just below (see Figure 3.4). Because of this, the estimation of the template surface is in two parts, one for particles below and one for particles above the focal plane. In effect, this means that we do not allow for any smoothing in the z -direction at $z = 0$.

The main benefit of a template surface instead of a set of template profiles for a fixed number of depth, is that we can use the template surface to estimate what a particle would look like at an arbitrary depth. Hence, at least in theory, we can use this surface to estimate the depth for a

continuously varying z . Another benefit is that we can get estimates of the standard error of a depth estimate, using the same approach as we used for the 2-D estimation. The reason for this, is that we can estimate the derivative of $f_z(r)$ in the z direction. Compare with the standard error derivation in Section 3 in Paper I.

Possibly the main problem with using a 2-D template is the smoothing in the z direction. This problem is related to the notion of distance between the intensity profile for different depths. For two particles at different depths, it not really clear how to smooth (or interpolate for that matter) the corresponding pixel values from the two particles, if the objective is to estimate the appearance (that is, the intensity profile) of a particle at a depth between. The way one usually does it, is to, for fixed r , smooth across the z direction. Let us for arguments sake, say that the intensity profile of the first particle has a peak at this r , and the intensity profile of the second particle "almost" has a dip at this r (this is almost the case for particles close to the focal plane). Then the resulting estimate for the depth between the two particles becomes something in-between. In these cases, the smoothing in the z -direction must be quite small, which were the case in the computation of the template surface in Figure 3.5.

Chapter 4

Tracking

The methods from the previous chapter will here be combined to illustrate the possibilities the positioning methods in three dimensions give to tracking. Possibly more, however, various problems for tracking colloidal particles in a dilute suspension using a large focal depth, as is the case for the sequence images considered here, will be highlighted.

In Figure 4.1 we demonstrate what occlusions might look like. The particles depicted in the two images are the two big particles in the upper part of Figure 3.2a. The image on the right, b), is recorded 20 time steps (0.4 sec) after the image on the left. As can be observed in the images, the two particles (optically) interfere with each other, causing a partial occlusion of the bright particle on the left. In b), the bright particle is almost fully occluded by the dark one. Note however that the actual particle centers in 3-D are far apart.

Remember that 50 sequence images are recorded each second and that the video camera alternates between updating the even and the odd rows in the image. See Section 2.2 for further details. The sequence starts with an even frame and then alternates between even and odd.

We will present a method that manages, to some extent, to handle dis-

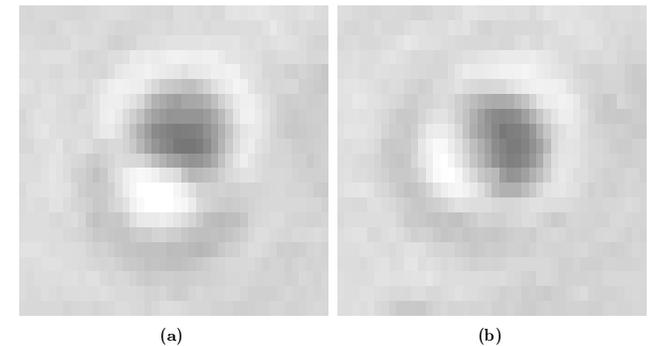


Figure 4.1: The two images demonstrate occlusion of particles. The image region is the same in both images, however the image b) is recorded 20 time steps (0.4 sec) after a).

tortion and partial occlusion. First we will however discuss two different approaches to linking position estimates of particles to trajectories.

4.1 Linking positions to trajectories

Assume first that our sequence of images only consists of two images. Now we want to track the particles present in the two images. There are basically two methods to do this. Either we first find all the particles in the two subsequent images and then find the correspondences between the two sets of particle positions, or we find the particles in the first image and then, for each particle in the first image, search in the vicinity of this particle in the second image for the corresponding particle.

There are pros and cons with both methods. For the first method, the drawbacks compared to the second are two: We do not use our knowledge of the positions of the particles in the previous image, and also, we have

to find the correspondences between the particles in the two images *after* we have estimated their positions. The drawback of the second method, is, at least unmodified, that it only allows for tracking of particles found in the first image.

As we saw in Section 3.2, it is very hard to find a method of finding all the particles in an image without a large amount of false particles, that is, candidate particle positions that do not correspond to true particles. Hence, we have to allow for a lot of false positives since we do not want to fail to hit the true particles, and consequently, the linking procedure of finding correspondences will involve a lot of candidate particles and inevitably becomes trickier. Examples in the literature of finding correspondences between sets of point patterns are Lund and Rudemo (2000), where correspondences between estimated tree positions from aerial photographs and the true positions were linked, or Cross and Hancock (1998), where the two sets of points were assumed to be the same up to an affine transformation plus a Gaussian error in the positions and where the false points were modelled by a Poisson process.

One simple approach to tracking would be: Manually assign the candidate particle positions in the first image, refine these positions, and then update the positions for each new image using the information contained in the previous particle configuration.

4.2 Handling partial occlusion

Let us start by considering the image on the left in Figure 4.1. Denote the true centers of the two particles in this image by x_1 and x_2 , where x_1 is the true center of the brighter particle on the left. Assume we are given initial candidate centers y_1 and y_2 , for example from the filtering method in Chapter 3 of finding the local maximal rotational symmetry. Using all pixels within let us say $r_{max} = 15$ from a candidate center as neighborhood, the two neighborhoods \mathcal{N}_{x_1} and \mathcal{N}_{x_2} would intersect, with the result that they would use the same pixels for estimating the particle

centers. This is illustrated in Figure 4.2a, where the candidate centers are indicated by plus signs and where the two neighborhoods are we have plotted a circle of radius $r_{max} = 15$ around each candidate center.

The corresponding results from the two optimizations of S in equation (1.1), using these neighborhoods for the two particles, would affect the particle center estimation in the plane, most probably with a bias directed away from the other particle. The reasoning behind this, is that the position of local maximal rotational symmetry would be pushed away from the interfering particle. In Figure 4.2b we have plotted the scatter plot of pixel values within 15 pixels away from the candidate center y_1 for the bright particle in Figure 4.2a. We clearly see the interfering pixel values resulting from the fact that there is another particle nearby.

An easy way to get around this, at least to some extent, is to allow pixel locations to be part of \mathcal{N}_{x_k} to the k th particle, only if the corresponding candidate center y_k is closest to the pixel among the other candidate centers. Let K denote the set of candidate centers in the image. Then we let the neighborhood of the k th particle be

$$\mathcal{N}_{x_k} = \{i \in \mathcal{D}_I : r_i(y_k) \leq r_{max} \text{ and } r_i(y_k) = \min_{m \in K} r_i(y_m)\}. \quad (4.1)$$

Compare this with equation (3.1). This definition of neighborhood thus requires that we are given a set of candidate centers. In Figures 4.2c and 4.2d, we display what the two neighborhoods look like by letting the pixel values be black for the pixel locations that are in the neighborhood of the other particle.

In Figure 4.3 we plot all pixel values within distance 15 from the two candidate particle centers. The pixel values in the two neighborhoods illustrated in Figure 4.2c and 4.2c, respectively, are however displayed with dots and the others are displayed with plus signs. It is obvious that the pixels outside the corresponding neighborhoods are interfering.

Furthermore, we will iteratively update the neighborhoods according to (4.1) as we are minimizing the criterion for maximal rotational symmetry (1.1). After the particle center in the plane has been found, the depth

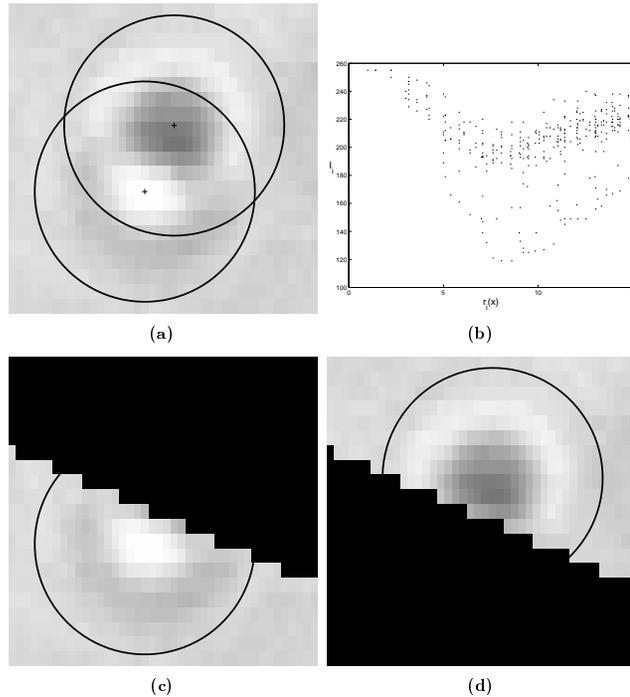


Figure 4.2: In a), the two plus signs are the two candidate particle centers. The two big circles have radii 15 and are centered at the candidate centers, representing possible neighborhoods of the particles. In b) there is a scatter plot of the pixel values up to distance 15 from the candidate center on the left in a). In c) and d) we demonstrate what the modified neighborhoods, defined by equation (4.1), look like, where a black pixel represents that the pixel is closer to the other particle.

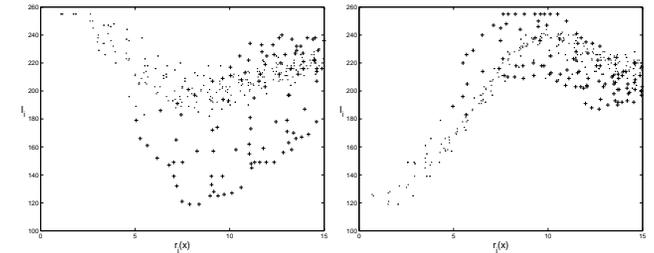


Figure 4.3: Scatter plots of the pixel values surrounding the estimated centers of the two particles in Figure 4.2a. The pixel values not in the corresponding neighborhood but within distance 15 from the candidate center are displayed with plus signs.

is estimated as in Paper II. Note that the precision in the position estimates will usually be worse when using smaller neighborhoods (compare with the simulation study in Section 5.1). Furthermore, the precision of the particle center estimate in the plane, will in general differ for the two coordinates.

The proposed algorithm for tracking in 3-D is as follows:

1. Calculate the distances between the candidate particle centers and construct the neighborhoods to each particle according to equation (4.1).
2. Minimize the local rotational symmetry function S for each particle using the corresponding neighborhood from Step 1.
3. Use the minimizers from Step 2 as candidate center and go to Step 1. Continue this until there are no changes in the position estimates.
4. Estimate the depth for each particle using template matching.
5. Load a new image from the sequence and let the estimated position from above be the candidate centers for the new image. Goto Step 1.

4.3 Results for tracking two particles

The two particles in Figure 4.1 were tracked using the proposed algorithm above. However, for the 16:th image the estimation broke down. This was indicated by that the estimated positions of the particles practically coincided.

In Figure 4.4a we show the same region as before, but for the 15:th image in the sequence together with the center estimates of the two particles in this image, after having tracked them for 14 images. A plus sign surrounded by a small circle will indicate a resulting particle center estimate. We see that the two particles have moved slightly towards each other, compared to the initial image in the sequence. Here, one could possibly argue that the estimated center of the particle on the left is somewhat biased downwards to the left.

The estimated centers in Figure 4.4a are used as candidate centers for the next image in the sequence, which is shown in Figure 4.4b together with the candidates. This next image comes from an odd frame, see Section 2.2. It seems as if both particles have moved upwards in the image plane. However, it is easy to be deceived by the eye since a transition from an even to an odd frame has the effect that everything seems to have moved upwards in the image.

As shown in Figure 4.4c, the estimation of the center of the particle on the left does not work. There are probably several reasons for this breakdown, but the main underlying cause is of course that one of the particles is heavily occluded by the other. Probably a more restrictive choice of neighborhood would be able to track both particles past this image. A few modifications in this direction will be discussed below. Nevertheless, it is hard to believe that it would be possible to track both particles past the image on the right in Figure 4.1, which is the 21st image in the sequence. There, the dark particle almost totally occludes the bright one.

When the position of a particle is wrongly estimated, we say that we

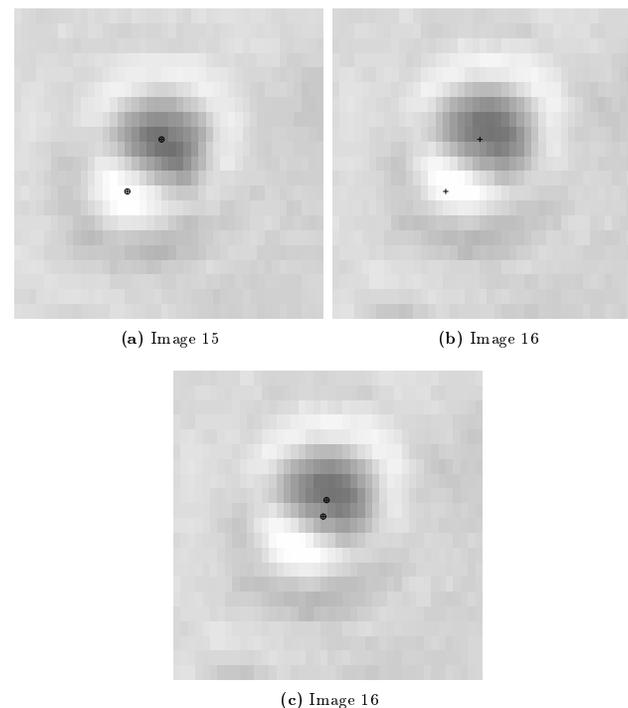


Figure 4.4: The resulting estimates of the particle centers for the 15:th image is shown in a). In b) the estimates from a) are plotted on the next image (which is an odd frame). The estimated positions for image b) is displayed in c) and clearly the tracking of the particle on the left has broke down.

lose the particle and call it a *lost particle*. Detecting the loss of a particle should not be hard. By looking at the increments of the particle trajectory

and at the distance to its nearest particles, it should be easy to detect the loss of a particle. A lost particle should probably not be discarded. The main benefit of the approach presented above for tracking, is that we get estimates of where the other particles in the image are located. One could say that these estimates tell us when to be cautious about which pixels to choose when estimating the center. A lost particle should therefore be kept since it signals that there could be another particle present in this part of the image.

4.4 Modifications

There are of course numerous ways to modify the presented tracking algorithm above. We here present a few ideas. Much of the issues have a strong algorithmic nature, by which we mean that much is associated with various tricks used in the implementation of tracking, however inspired by statistical and probabilistic reasoning.

First of all, the presented method of tracking does not make use of the depths of the particles in the previous image in the sequence. As is the case for the positions in the image plane, the depths of the particles cannot change so much in the time interval between two consecutive sequence images. A modification would be to incorporate the information of the appearance of the particles in the previous image when constructing the neighborhoods of the particles in the present image, since this tells us how big the domain of interfering pixels is.

More generally, this brings up the subject of using different shapes of neighborhoods to particles in the image. The method we presented is based on dividing the image plane into Voronoi cells. More elaborate methods could of course be used. One possibility would be to modify equation (4.1) such that the distance $r_i(y_k)$ must be smaller than a constant $c < 1$ times the minimum distance to the other candidate centers $\min_{m \in K \setminus \{k\}} r_i(y_m)$. The reason why this is believed to perform better, is that pixels in the region in-between particles are affected by both of the

nearby particles, and hence should not be used for positioning at all.

One could also incorporate detection of outliers in the scatter plot of pixels in the neighborhood compared to the fitted intensity profile. It is important then to relate the possible outliers, to where in the image they are located, in other words, we have to incorporate the spatial nature of the data. Drawing conclusions on outliers simply from pixel values plotted versus distance, is highly unsafe. Only if a possible group of outliers in the one-dimensional scatter plot can be spatially related in some way, can we allow to discard them. (By spatially related pixels, we mean pixels that are connected in the image plane.) This is the main reason why a robust method such as the LOWESS (locally weighted scatter plot smoother), Cleveland (1979), was not used for the nonparametric estimation of f in the estimation of particle centers in the plane. One way to spatially relate outliers could be to discard them only if they are sufficiently close to the boundary of the neighborhood used. This way, we could let the neighborhood shrink, as outliers near the boundary are discarded.

Chapter 5

Supplementary studies

We will present some supplementary studies based on simulations which might be of interest when reading Paper I and to some extent Paper II. Therefore it is advisable to have read at least Paper I before reading this chapter.

5.1 Dependence between the size of \mathcal{N}_x and the standard error in the 2-D estimation

We will investigate how the estimated standard error of the particle center estimate varies as a function of r_{max} , when circular neighborhoods \mathcal{N}_x , as in equation (3.1), around the particle center x , are used.

As in Paper I, we let g denote the \mathbb{R}^n valued function of $x = (x_1, x_2)$ with k th element $g_k(x) = \hat{f}(r_{ik}(x))$. Also, J denotes the Jacobian of g , the n times 2 matrix with element $(J(x))_{k,m} = \frac{\partial g_k}{\partial x_m}(y)$ for $k = 1, \dots, n$ and $m = 1, 2$. The sandwich estimator of the variance matrix $\mathbf{Var}\{\hat{x}\}$ of the estimated position error at the true center x is

$$V = (J^T J)^{-1} J^T \Sigma J (J^T J)^{-1}. \quad (5.1)$$

where Σ is the variance matrix of the residuals between the pixel values and the corresponding estimates of f .

In this simulation study, the intention is to focus on the dependence of r_{max} on the diagonal elements of \hat{V} . Therefore we will use the true (but random) value of the particle center x . Furthermore, instead of estimating the image noise parameters as we did in the simulation study of Paper I, the true values of the image noise parameters σ^2 and c , will be used. The elements of J are approximated in the same way as in Paper I, that is, by the estimated derivatives of the intensity profile.

For each simulated image I with random particle center x , we calculated V for r_{max} between 3 and 15 from the true center x . The parameters of the image noise were $\sigma^2 = 25$ and $c = 0.6$.

Let $\sigma_m(r_{max})$ be the square-root of the maximal element of the two diagonal elements of V for the m th simulated image with r_{max} as radius of the neighborhood \mathcal{N}_x . In Figure 5.1 the result after $M = 100$ simulations each for particles $3\mu\text{m}$ below and above the focal plane is shown. These depths correspond to indices ± 5 in Paper I. For each r_{max} , the mean $\bar{\sigma}(r_{max}) = M^{-1} \sum_{m=1}^M \sigma_m(r_{max})$ over the M simulated images, is displayed. Below this, we have plotted the intensity curve f_z used in the simulation. See also Figures 5.2 and 5.3, where the same kind of dependence is illustrated, this time for particles $\pm 1.8\mu\text{m}$ and $\pm 4.2\mu\text{m}$ from the focal plane, corresponding to depths indexed by ± 3 and ± 7 in Paper I.

The observation made from the plots, is that the way $\bar{\sigma}$ varies with r_{max} , clearly seems to depend on f_z . After each point r where f_z has zero derivative (that is, after each stationary value of f_z), the standard error seems to drop. Furthermore, the magnitude of this dip depends on the distance from the particle center, r . Then, for a sufficiently big r_{max} , the increase in r_{max} does not seem to affect the standard error. Note also the non-symmetry around the focal plane in each figure; the plots on the left and right represent particles at the same distance from the focal plane, but the shapes of the standard error curves are different. This is consistent with that the location of the first stationary value of f_z for $r > 0$ is closer to $r = 0$ for a particle below the focal plane than for a

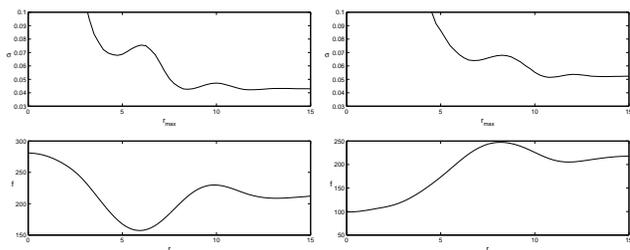


Figure 5.1: The upper curves show how the estimated standard error from the sandwich estimator varies with r_{max} for a particle $3\mu\text{m}$ below (left) and above (right) the focal plane, respectively. Below the curve, the corresponding intensity profile for the particle is shown. The way the standard error varies with r_{max} clearly depends on the underlying intensity curve. These depths correspond to depths -5 and 7 in Paper I.

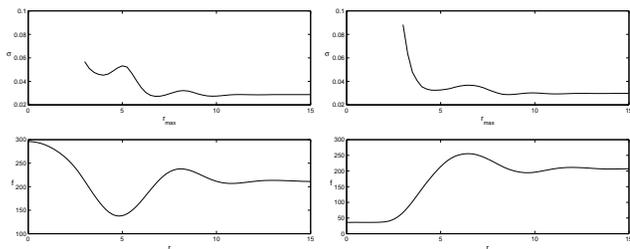


Figure 5.2: The same as Figure 5.1 but here for particles $1.8\mu\text{m}$ below (left) and above (right) the focal plane. These depths correspond to -3 and 3 in Paper I.

particle at the same distance, but above the focal plane; compare the two intensity profiles to the left and right in each figure.

The computational effort of calculating the estimate of the intensity profile at a candidate center y from pixels in a neighborhood \mathcal{N}_x with radius

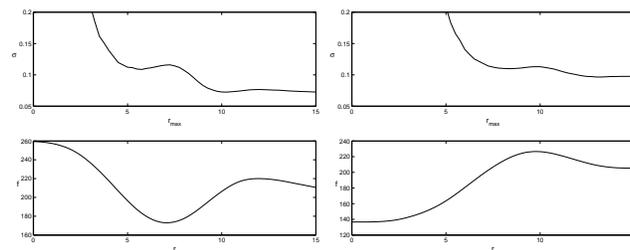


Figure 5.3: The same as Figure 5.1 but here for particles $4.2\mu\text{m}$ below (left) and above (right) the focal plane. These depths correspond to -7 and 7 in Paper I.

r_{max} , increases quadratically in r_{max} . Therefore there is a trade-off between increased precision in the estimated particle center and the time it takes to estimate the center. As seen from Figure 5.1, 5.2, and 5.3, after a certain value of r_{max} , depending on depth of the particle, there is no decrease in standard error as we increase the neighborhood. Therefore, there is no need to use a larger neighborhood than necessary. On the other hand, since the depth of the particle is unknown, we want to use the same r_{max} for all particles. $r_{max} = 15$ seems to be a good choice since this includes the major fluctuations of the intensity profiles for the majority of the depths considered.

There are also times when a small r_{max} is preferable (when the shape of \mathcal{N}_x is fixed). This is the case when particles are close together since then the assumption of rotational symmetry might not be fulfilled if r_{max} is taken too large and consequently we probably get a bias in the particle center estimation directed from the interfering particle (because the value of (1.1) is probably smaller away from the interfering particle). However, in those cases, we have hopefully already detected the interfering particle and adjusted the shape of the neighborhood accordingly, as discussed in Chapter 4. However, for automatically finding candidate centers as proposed in Section 3.2, we use the same shape of the neighborhood for all pixel locations and then a small r_{max} might perform better.

A note of warning should also be said about over-interpreting the dependence of the standard error estimates on the size of the neighborhood. The sandwich estimator is sort of an estimate of the local curvature of $S(y)$ at the true center x . It does not say anything about *how to get to this true center*. In other words it says nothing about the consistency of the particle center estimator, only what the standard error will be *if* the estimated \hat{x} gets sufficiently close to x .

5.2 Estimated intensity profiles for different bandwidths

In Paper I a cross-validation study was performed on the data corresponding to the sub-images of a particle at 15 different depths. The conclusion was first of all that the local quadratic kernel estimate was relatively insensitive to the choice of bandwidth and secondly, that bandwidth $h = 0.7$ worked sufficiently well to be used as a pilot bandwidth when finding the particle centers. Before the actual estimation of the intensity profiles however, a cross-validation study was made for each particle separately. Here we will first investigate the bias in the nonparametric estimation of the intensity profile at different depths for different bandwidths. Then, we will investigate the accuracy of cross-validation for the kind of data considered here.

By an estimate of an intensity profile f given pixel values I_i and corresponding distances $r_i(x)$, we will below mean the local quadratic kernel estimate of f with $(I_i, r_i(x))$ as data, with no censoring.

Here we will focus on the estimation of intensity profiles, and instead of using the data-driven method of cross-validation, we will conduct a simulation study to illustrate how the estimated intensity profiles depend on the choice of bandwidth. We simulate images of particles with added image noise (in the same way as in Paper I and II), and then the intensity profiles will be estimated using the true center as the particle center. This will be done for different bandwidths for each image.

The mean of the estimated intensity profiles from 100 simulated images of a particle at the focal plane, using bandwidths $h = 0.5, 0.7$, and 1.0 , respectively, from left to right, is shown in solid in the three sub-figures in the upper row of Figure 5.4. The true intensity profile used in the simulations, is displayed in dashed and the dotted lines are the pointwise maximum and minimum of the estimated values of the intensity profiles. The bias in the estimation is very high for the largest bandwidth and almost zero for the smallest bandwidth. Interestingly, the pointwise distance between the minimum and maximum value of the intensity functions is not increased much for the lowest bandwidth, compared to the highest.

The story for the two other depths of Figure 5.4, $1.8\mu\text{m}$ and $4.2\mu\text{m}$ above the focal plane, is basically the same as the story for the upper three sub-figures. Furthermore, this is true for all other depths; the pointwise bias is lower for smaller bandwidths. However, looking at functional characteristics of the estimated functions, such as the position of the first stationary value of the estimated function, we get another story. Using a smaller bandwidth than necessary, makes the intensity profile estimates too irregular with too much fluctuation. Therefore, the plots in Figure 5.4 are a bit misleading, as one might think that we should always choose a small bandwidth. Anyhow, it is of course clear that the bandwidth $h = 1.0$ is too large for estimating the intensity profiles in the two upper rows.

In Figure 5.4 we calculated the estimate for uniformly spaced values of r . The covariates, the distances $r_i(x)$ from particle center x to the pixel locations i , are however not uniformly spaced. More precisely, they are randomly distributed (since the particle center x is random), with linearly increasing density of covariates with distance, since the number of pixels within distance r from a particle center increases quadratically. Since there are few pixels for small r values, the variation is fairly large for all bandwidths at small r values, as seen by the larger span between the pointwise minimum and maximum of estimated functions in Figure 5.4. Remember also that the image noise in the simulations is correlated.

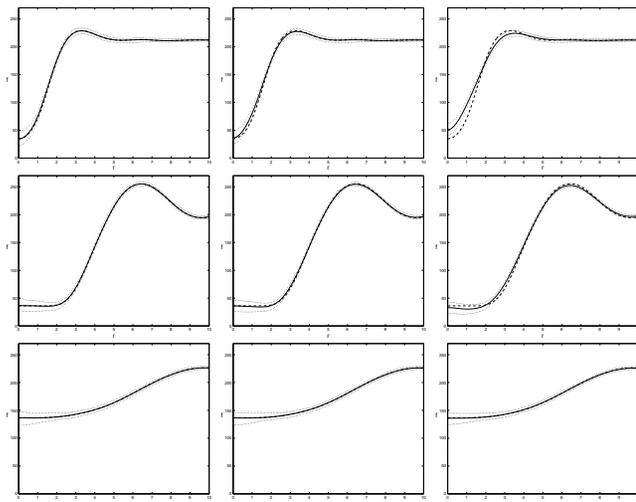


Figure 5.4: Mean of the estimated intensity profiles (solid) from 100 simulations corresponding to a particle at the focal plane in the upper row, together with particles $1.8\mu\text{m}$ and $4.2\mu\text{m}$ above in the middle and lower row, respectively, when using three different bandwidths. The bandwidths 0.5, 0.7, and 1.0, are shown from left to right, respectively. The true intensity profile is displayed in dashed and the dotted lines are the pointwise maximum and minimum values of the 100 estimated intensity profiles. Clearly the function estimates have a large bias when using too large a bandwidth; see for example the plot in the upper right corner, where the mean of the estimated functions (solid) is far from the true function (dashed).

Validating Cross-validation

Let us denote by \hat{f}_h^{-i} the local quadratic kernel estimate of f when leaving out the i :th data point, and using h as bandwidth. For each h the cross-

validation score is defined as

$$CV(h) = \frac{1}{n} \sum_{i \in \mathcal{N}_x} \{I_i - \hat{f}_h^{-i}(r_i(x))\}^2. \quad (5.2)$$

The cross-validation score is calculated at a finite number of bandwidths. The idea behind the cross-validation score, is that it is an estimate of the expected value of the squared difference between the estimated and the true regression curve (intensity profile) under the bivariate distribution of covariates and response values, $(r_i(x), I_i)$,

$$\mathbf{E}\left\{ \int (\hat{f}(\tau) - f(\tau))^2 d\mu(\tau) \right\} \quad (5.3)$$

where μ is the (marginal) distribution of covariates. In a simulation study, we know the true intensity profile, and therefore we can estimate expression (5.3) by

$$E(h) = \frac{1}{M} \sum_{m=1}^M \frac{1}{n_m} \sum_{k=1}^{n_m} (\hat{f}_h(r_{ik}(x_m)) - f(r_{ik}(x_m)))^2 \quad (5.4)$$

where \hat{f}^h is the intensity profile estimate using the data from the m th simulated image and i^1, \dots, i^n is an arbitrary enumeration of the n_m pixel values within distance r_{max} of the (random) particle center x_m . Note that for each simulation, both the pixel values I_i as well as the distances to the pixels from the particle center, $r_{ik}(x_m)$, are random. The pixel values are implicitly present in the summation (5.4) above, in the estimate of the intensity profile \hat{f}^h .

We conducted a new simulation study of $M = 100$ images for each of the 15 depths from $-4.2\mu\text{m}$ to $4.2\mu\text{m}$ with $0.6\mu\text{m}$ in between. These are the same depths considered in Paper I. For each image, we estimated the intensity profile f up to distance $r_{max} = 15$ using 17 different bandwidths $h = 0.4, 0.45, 0.5, \dots, 1.2$. Then E in equation (5.4) was calculated for each depth. Also, for each image, the cross-validation score was calculated, using the same set bandwidths. In Figure 5.5 we present the h minimizing E for each depth in solid together with the mean of the bandwidths minimizing the corresponding cross-validation score for each

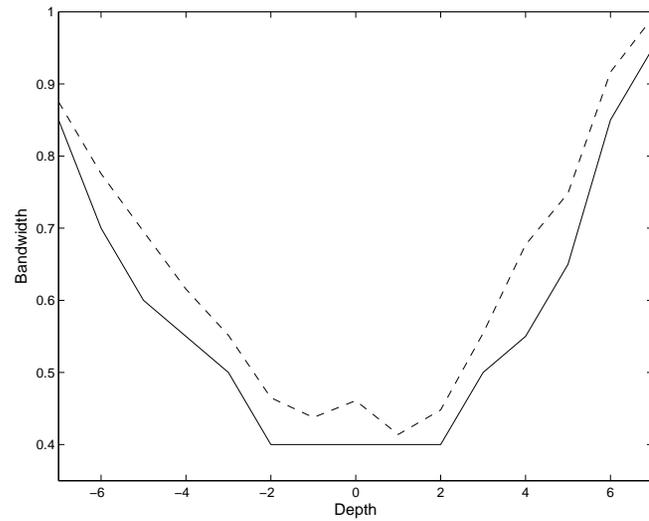


Figure 5.5: For each depth, the minimizer of $E(h)$ in equation (5.3) using 100 simulated images of particles at this depth is shown in solid. The dashed plot is the mean over each depth of the minimizers of the cross-validation score for each image.

image. The shape of the two plots is the same, but there is a discrepancy of roughly 0.06 between the two plots. It seems as if the cross-validation minimizer is biased. In practice however, this discrepancy in bandwidth does not affect the estimation of f severely.

Chapter 6

Conclusions and Future Research

6.1 Depth estimation using templates

As discussed in Section 3.4, there related to the estimation of a 2-D surface of intensity profiles, a measure of distance between the intensity profiles should take special consideration of the functional features unique for each profile. One example of such a feature is the first dip or peak of the intensity profile, depending on if the particle is below or above the focal plane. Looking at the intensity profiles, it seems as if the profiles are related to each other with a scaling parameter depending on depth, that acts on the argument r and the amplitude of the variation, stretching out the profile with increasing distance to the focal plane. An improved measure of distance between intensity profiles should take this into consideration. In short, it would be nice to have a measure of distance between shapes that exploits the important functional features like the first dip, in a more explicit manner than the L_2 -norm, which what the matching criterion (3.4) basically is.

Templates that allow for continuous depth estimation are also desirable. In Section 3.4 we presented one method for construction a two dimensional regression surface of intensity profiles. From a theoretical point of view, mathematically constructed templates would of course be preferable, since this would also allow for construction of templates for particles of different sizes and shape. However, as mentioned in the Introduction, predicting the appearance of spherical objects in microscopy of this size still seems to be an intriguing theoretical challenge, see Ovrjn and Izen (2000).

The 2-D template approach described briefly in Section 3.4 could be developed further. However, this requires better data of the appearances of particles at different depths. Several images at each depth would probably improve the estimates of the intensity profiles in general. It should also be possible to record the z-scans without having censored pixel values. Even so, this approach would still confine us to the study of particles of only the sizes and shapes we have estimated templates for.

6.2 Measurements of diffusion coefficient and interaction

As illustrated in Chapter 4, it is hard to track particles unsupervised for a longer sequence of images considered here. The main cause is that the focal depth of the microscope is very large, causing particles to occlude each other rather frequently, at least compared to the ratio of the total number of particles present in the specimen and the volume of the domain in which they are confined.

Loosing track of particles is not very important if we are only concerned with the estimation of the diffusion coefficient of the particles. Then broken trajectories of particles do not affect the estimation (more than that the sample size of observed increments decreases); when a particle is lost, we do not have to bother were it goes and when we possibly start tracking this particle again, we could consider it as a new particle (at least if all particles are of the same size). If interaction between particles is

believed to be present, one could furthermore restrict tracking to regions in the image where there are no other particles present within a reasonable range of particle interaction.

However, when estimating a possible interaction between the particles, it is crucial to estimate the positions of all particles. In particular, it is important to estimate the positions of those particles which are close to each other. But this is exactly when unsupervised tracking is hard! Supervised estimation of particles centers is of course possible, but for a larger sequence of images, this is a very tedious job. The interaction between particles can however in principle be estimated from the estimated particle positions in three dimensions in a single image using methods of statistical inference for spatial point processes, see for example Møller and Waagepetersen (2004).

Since unsupervised tracking is hard in dilute suspensions, one alternative is to use an optical trapping device called an optical tweezer. An optical tweezer is tightly focused laser beam that creates a local minimum in the optical energy strong enough to overcome both radiation forces and thermal forces. Thereby it is possible to attract a particle and move it to a specified location. A dual optical tweezer could be used to attract two particles, bring them close together, and then turn off the laser. The particles would then diffuse freely and we can track the two particles for a sequence of images. The whole thing can be repeated until enough data has been collected. Optical tweezers have been used in studies of the interaction potential in Crocker and Grier (1996) but also in for example video microscopy studies of DNA, see Perkins et al. (1994). Given the high accuracy of position estimation of spherical latex particles developed here, the possibilities for high precision estimation of the interaction between particles using this more advanced microscope setup should be large.

6.3 Automated depth calibration

An interesting problem would be to automatically estimate the depth, simply by observing trajectories of particles performing Brownian motion (like the sequence images we have used). Since Brownian motion is isotropic, a particle performing Brownian motion has the same diffusion coefficient in all three dimensions. Since tracking the position of a particle works fairly well in the image plane, we can estimate the diffusion coefficient. Now the idea is as follows. As we track the particle in two dimensions, we record some kind of feature that relates to the depth of the particle. This could be the estimated intensity profile, but it could also be a simpler attribute, for example the distance from the particle center to the first stationary value of the estimated intensity profile, that is, the distance to the first peak or dip in the profile. Then, if it is possible to order these measured features of the depth, and assuming that there is a function that relates this feature with the depth of the particle, we could associate the track of this depth feature with the diffusion coefficient already estimated from the measured diffusion in the plane.

Bibliography

- W. Cleveland. Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association*, 74:829–836, 1979.
- J. Crocker and D. Grier. Microscopic measurements of the pair interaction potential of charge-stabilized colloid. *Physical Review Letters*, 73:352–355, 1994.
- J. Crocker and D. Grier. Methods of digital video microscopy for colloidal studies. *Journal of Colloid and Interface Science*, 179:298–310, 1996.
- J. Crocker and D. Grier. Interactions and dynamics in charge-stabilized colloid. *MRS Bulletin*, 23:24–31, 1998.
- A. Cross and E. Hancock. Graph matching with a dual-step EM algorithm. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20:1263–1253, 1998.
- A. Dempster, N. Laird, and D. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *J. Roy. Statist. Soc. Ser. B*, 39:1–38, 1977.
- M. S. Elliot and W. Poon. Conventional optical microscopy of colloidal suspensions. *Advances in Colloid and Interface Science*, 92:133–194, 2001.
- D. F. Evans and H. Wennerström. *The Colloidal Domain. Where Physics, Chemistry, Biology, and Technology meet*. Wiley-VCH, New York, second edition, 1999.
- J. Fan and I. Gijbels. *Local Polynomial Modelling and Its Applications*. Chapman and Hall, London, first edition, 1996.

BIBLIOGRAPHY

- R. Gonzales and R. Woods. *Digital Image Processing*. Prentice-Hall, Upper Saddle River, New Jersey, 2002.
- D. Grier. Colloids: A surprisingly attractive couple. *Nature*, 393:621, 1998.
- L. Györfi, M. Kohler, A. Krzyzak, and H. Walk. *A Distribution-Free Theory of Nonparametric Regression*. Springer-Verlag, New York, first edition, 2002.
- T. Hastie and R. Tibshirani. *Generalized Additive Models*. Chapman and Hall, London, first edition, 1990.
- D. Kerbyson and T. Atherton. Circle detection using Hough Transform Filters. In *Fifth International Conference on Image Processing and its applications*. IEEE Conference Publications, 4-6 July 1995.
- M. Kvarnström. Estimating diffusion coefficients in colloidal particle systems. Technical report, Mathematical Sciences, Chalmers University of Technology, 2002. Licentiate Thesis.
- J. Lund and M. Rudemo. Models for point processes observed with noise. *Biometrika*, 87:235–249, 2000.
- J. Møller and R. Waagepetersen. *Statistical Inference and Simulation for Spatial Point Processes*. Chapman and Hall/CRC, Boca Raton, Florida, first edition, 2004.
- B. Ovryn and S. Izen. Imaging of transparent spheres through a planar interface using a high-numerical-aperture optical microscope. *Journal of the Optical Society of America A. Optics, Image Science and Vision*, 17:1202–1213, 2000.
- A. Owen. *Empirical Likelihood*. Chapman and Hall/CRC, London, first edition, 2001.
- T. Perkins, S. Quake, D. Smith, and S. Chu. Relaxation of a single dna molecule observed in optical microscopy. *Science*, 264:822–826, 1994.
- J. Serra. *Image Analysis and Mathematical Morphology*. Academic Press, London, first edition, 1982.
- D. Young, C. Glasbey, A. Gray, and M. N.J. Towards automatic cell identification in dic microscopy. *J. Microsc.*, 192:186–193, 1998.

Estimating centers and intensity profiles of spherical particles in microscopy

Mats Kvarnström and Chris Glasbey

January 24, 2005

Abstract

We present a method for estimation of particle center in digitized microscope images, based on an assumption of rotational symmetry of pixel values surrounding a true particle center. The functional form of how the pixel values vary with distance from a particle center is called the intensity profile and depends on the depth of the particle. In addition to estimating the center, the intensity profile is also estimated using a nonparametric estimator. However, pixel values are censored above a certain known value. We propose a modification of the local quadratic kernel estimate for non-parametric function estimation using censored response values.

Furthermore, for each center estimate, we also estimate the standard error of the estimate using a sandwich estimator. A simulation study shows that these standard errors are consistent with the observed RMS errors. The standard errors depend on distance to the focal plane and are in the range of 0.02 to 0.10 pixels, depending on depth of the particles, with lower values for particles closer to the focal plane.

1 Introduction

In order to track and subsequently estimate the diffusion coefficients of diffusing colloidal particles observed in video microscopy, a highly precise, automatic method for estimating particle centers is needed. It is also of interest to have reliable estimates on the standard errors, since errors in the position estimates affect the subsequent estimation of the diffusion coefficient. Furthermore, sub-pixel accuracy is needed since the particles typically only diffuse distances in the order of one or two pixels between two subsequent images.

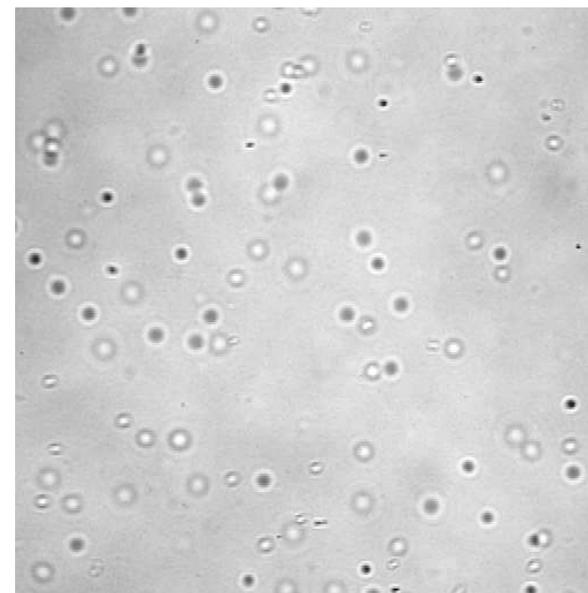


Figure 1: An image from a video sequence of diffusing particles. The particles are all equal in size and the difference in the appearances of the particles in the image is due to that the particles are at different depth relative to the focal plane.

Figure 1 shows one of the images in a typical sequence of images recorded and digitized by the video microscope. The images are recorded at a frame rate of 50 images per second. Each image consists of 512 times 512 square pixels with a side-length of 180 nm. The particles are all of the same size, 494 nm in diameter; the difference in appearance is due to different depth relative to the focal plane.

Our method of estimating the particle centers in the image relies on an assumption of rotational symmetry of the pixel values in the vicinity of a true particle center. We will estimate the particle center by the position with largest rotational symmetry, where the symmetry is measured as the sum of

squares between nearby pixel values and a fitted nonparametric estimate of how pixel values vary with distance from a (candidate) particle center. The pixel values in the images considered are censored (saturated) at an upper limit of $T = 255$ and are instead set to T . We introduce a method for nonparametric estimation of a regression function when the response variables (the pixel values) are censored above an upper limit under the assumption of normally distributed homoscedastic observation errors (the image noise).

Censored (saturated) pixel values are common in microarrays where the estimated expression of genes get biased when no adjustment is done due to the censoring, see Glasbey et al. (2005) and Ekström et al. (2004). A nonparametric approach was pursued by Glasbey et al. (2005) using a principal components model. Ekström et al. (2004) used a parametric approach where several parametric forms were tested for the expression of genes in the microarrays.

Standard error on each center estimate will be based on the sandwich estimator, see for example Owen (2001). These standard errors are shown to be consistent with root-mean-square (RMS) errors in a simulation study in Section 5. In this simulation study, we also examine the precision when only half of the rows in the image is used in the estimation. This is done since for image like the one in Figure 1, in fact only half of the rows are updated each $1/50$ second, alternating between the even and the odd rows. Since only half as many pixels are used, one might have expected $\sqrt{2}$ times larger RMS errors and furthermore, since we are losing information in the vertical direction, the errors would be greater in the vertical than in the horizontal direction. The study shows however, that the loss in precision is not as high as expected and especially that the precision is still equal in both coordinates, except for the particles closest to the focal plane.

In Figure 2, we have extracted and zoomed in on a particle like the ones in Figure 1 at different distances from the focal plane. Each sub-image consists of 27 times 27 pixels. Sub-images with negative and positive labels correspond to particles below and above the focal plane, respectively. The distance in depth between two consecutive sub-images is approximately $0.6\mu\text{m}$ and sub-image 0 corresponds to a particle at the focal plane. These images are called *z-scans* and have been constructed by letting particles adsorb on the glass surface of the specimen, and then moving the specimen relative to the optics of the microscope. This way, we were able to record the appearances of particles at different depths of focus.

Previous work using digital video microscopy for colloidal studies, have been made by Crocker and Grier (1996, 1998), but there, the depth of focus was approximately $\pm 0.5\mu\text{m}$, which makes the particles similar in appearance and therefore easier to find in the images, since the sought-after objects in each image

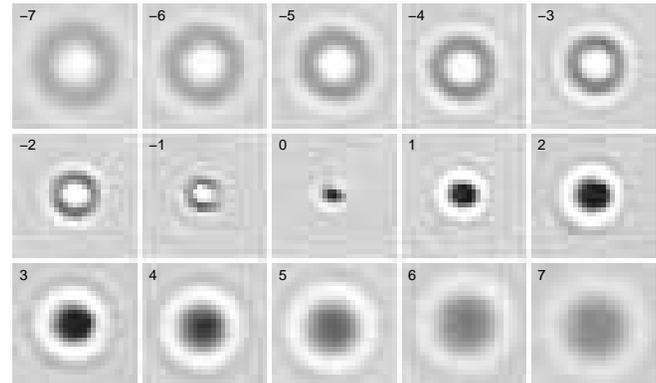


Figure 2: The appearance of a particle at 15 different distances to the focal plane. Sub-image 0 corresponds to a particle at the focal plane, and sub-images with negative and positive labels are below and above the focal plane, respectively. The distance in depth between two consecutive sub-images is approximately $0.6\mu\text{m}$.

are similar. In their study, each particle is a bright spherical set of pixels and they achieve sub-pixel accuracy by calculating the geometric center of the brightness-weighted centroid, obtaining standard errors of about 0.1 pixels (where each pixel is 85 nm). Also, the particles in their colloidal suspensions are effectively confined to a crystallized structure. In our case, the depth of focus is $\pm 15\mu\text{m}$, resulting in a much wider variety of appearances of the particles in the image.

Assumptions and main idea

We denote an image by I . This is effectively a matrix of pixel values I_i for pixel locations $i = (i_1, i_2) \in \mathcal{D}_I \subset \mathbb{Z}^2$, where \mathcal{D}_I is the set of indices i for which the image is defined. We will use the terms pixel and pixel location interchangeable to mean the same thing when no risk of ambiguity exists.

A particle center is denoted by $x = (x_1, x_2) \in \mathbb{R}^2$, and to each particle, we associate a set $\mathcal{N}_x \subset \mathcal{D}_I$ of pixels in the image I called the *neighborhood of the particle at x* . For the particles in Figure 2, we can take all pixels in the

corresponding sub-image as the neighborhood for the particle, but typically, we let

$$\mathcal{N}_x = \{i \in \mathcal{D}_I : r_i(x) \leq r_{max}\}, \quad (1)$$

where $r_i(x)$ is the Euclidean distance from the particle center x to pixel i , and r_{max} is an appropriately chosen distance. In this paper, we will let r_{max} be 15. The main assumption is that, for a particle center at $x \in \mathbb{R}^2$ we have

$$I_i = f(r_i(x)) + \epsilon_i \quad \text{for } i \in \mathcal{N}_x, \quad (2)$$

where f is called the *intensity profile*. The image noise, ϵ_i for $i \in \mathcal{D}_I$, is assumed to be normally distributed with isotropically correlated pixel values. The intensity profile $f : \mathbb{R} \rightarrow \mathbb{R}$ is assumed to be a smooth function with at least two continuous derivatives and symmetric in its argument r , here denoted by \mathcal{C}^2 .

The basic idea for estimating the particle center is as follows. For a particle located at x , we use the minimizer of

$$S(y) = \min_{f \in \mathcal{C}^2} \sum_{i \in \mathcal{N}_x} \{I_i - f(r_i(y))\}^2 = \sum_{i \in \mathcal{N}_x} \{I_i - \hat{f}(r_i(y))\}^2 \quad (3)$$

for $y \in \mathbb{R}^2$ as estimate of x . We calculate \hat{f} using a local quadratic kernel estimator with (appropriately chosen) bandwidth h . The idea behind minimizing the equation above is to find the position of (local) maximum rotational symmetry. We denote the estimate of x by \hat{x} . At \hat{x} , the corresponding \hat{f} will be the estimate of the intensity profile for this particle. The idea behind minimizing equation (3), is to find the position of (local) maximal rotational symmetry.

Pixel values are censored above an upper limit T . Here, T equals 255. In Figure 2, we have censoring for all particles except at the focal plane and at depths 6 and 7. For particles below the focal plane, censoring occurs for pixels close to the center of the particle, and for particles above the focal plane, on the first fringe. The relative amount of censoring is however not very large; between 0.5 and 4.5 percent of the $27 \cdot 27 = 729$ pixels are censored, with the largest amount for the particles just above the focal plane (depth 1, 2, and 3). For estimating the particle center, that is, finding the position of maximal rotational symmetry, censoring does not affect much, since censoring occurs at an annulus around the true particle center. Nevertheless, for estimating the intensity profile, it is important.

In Figure 3 and 4 we illustrate some of the assumptions and methods presented so far. In Figure 3 we have zoomed in on the particle at depth -5 and made a surface plot of S for a few values surrounding the minimizer. In Figure 4, we see that it is reasonable to assume rotational symmetry of the pixel values

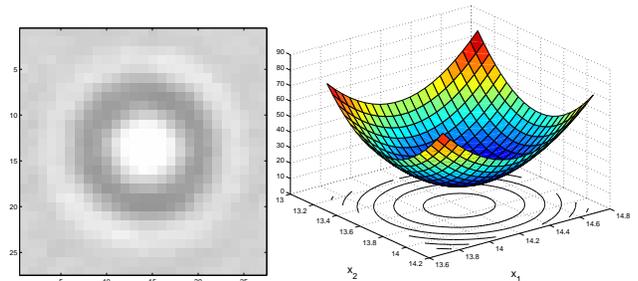


Figure 3: The particle at depth -5 together with a surface plot of S on the right. The center pixel in the image is $(x_1, x_2) = (14, 14)$. The estimated center is at $(14.21, 13.67)$. The estimated standard error of this position estimate is 0.035 (see Section 3). In Figure 4 a comparison is made between the scatter plots of pixel values surrounding $(14, 14)$ and the estimated center.

surrounding a true center. Furthermore, the two plots indicate that it should be possible to estimate the particle center at sub-pixel accuracy; in the scatter plot on the left, there is a clear “shift” in the scatter plot, which is not present in the scatter plot on the right, corresponding to the estimated center. The figure also demonstrates censoring of pixel values at the upper limit of $T = 255$.

2 Intensity profile estimation

For the case when there are no censored pixel values in \mathcal{N}_x , we will use the local quadratic kernel estimator with a Gaussian kernel function. Common references for nonparametric local polynomial kernel estimation are Hastie and Tibshirani (1990) and Fan and Gijbels (1996). The reason for choosing the local quadratic kernel estimate as nonparametric method, instead of for example a spline smoother, is that the estimate of the derivative of f is practically given to us for free using a local quadratic, and the derivative is needed when estimating the standard errors. Compared to a local *linear* kernel estimate, the quadratic is much less sensitive to the choice of bandwidth. The Gaussian kernel was chosen since it is smooth and has unbounded support, which makes the function S differentiable. Regarding the choice of bandwidth, simulation studies have shown that the bandwidth is not very important when estimating the particle centers.

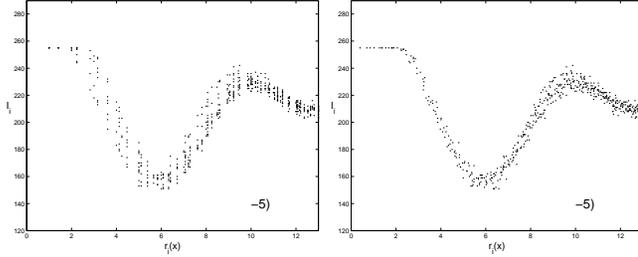


Figure 4: Two scatter plots for the particle at depth -5 in Figure 2. On the left, the scatter plot of pixel values surrounding the center pixel of the sub-image (14, 14), and on the right, the same for the estimated particle center at (14.21, 13.67). Note the censoring of pixel values above $T = 255$. See Figure 3 for an image of the particle and the plot of S .

We will however use cross-validation to automatically choose bandwidth.

Local polynomial kernel estimate

We will explain briefly how a local polynomial kernel estimate is calculated. The value of \hat{f} at each point of evaluation r_0 is a solution to a local weighted least-squares problem, where each observation I_i is given a weight from the Gaussian kernel function according to distance between r_0 and $r_i(y)$. For estimation when there is censoring, however, we will modify this weighted least-squares problem. For easier notation, we will henceforth write r_i instead of $r_i(y)$ for $i \in \mathcal{N}_x$ when there is no possibility of misunderstanding.

Let K denote the Gaussian kernel, that is $K(x) = \exp(-x^2/2)$. The local polynomial kernel estimate is based on the assumption that the unknown regression function f can be approximated locally by a polynomial of degree p ,

$$f(r) \simeq \sum_{k=0}^p \frac{f^{(k)}(r_0)}{k!} (r - r_0)^k$$

for r close to r_0 . To calculate the value of the estimate at an arbitrary point r_0 , we locally fit a polynomial of degree p by solving the weighted least-square

problem of finding the minimizer to

$$\sum_{i \in \mathcal{N}_x} \left\{ I_i - \sum_{k=0}^p \beta_k (r_i - r_0)^k \right\}^2 K_h(r_i - r_0) \quad (4)$$

with respect to $\beta = (\beta_0, \dots, \beta_p)$ and denote the minimizer by $\hat{\beta}$. Here, $K_h(x) = h^{-1}K(x/h)$ is the kernel with bandwidth $h > 0$. The estimates of $f^{(k)}$ at r_0 , for $k = 0, 1, \dots, p$, are

$$\hat{f}^{(k)}(r_0) = k! \hat{\beta}_k.$$

Notice that since $\hat{f}(r_0)$ is the solution to a weighted least-squares, $\hat{f}(r_0)$ is a linear combination of the pixel values I_j for $j \in \mathcal{N}_x$.

Local regression with censored response values

Denote the set of uncensored pixels by $A_T = \{i \in \mathcal{N}_x : I_i < T\}$, where $T = 255$. Then the complement, A_T^c , is the set of censored pixels. With pixel values in \mathcal{N}_x censored, we minimize

$$\begin{aligned} & \frac{1}{2} \log \sigma^2 \sum_{i \in A_T} K_h(r_i - r_0) + \frac{1}{2\sigma^2} \sum_{i \in A_T} \left\{ I_i - \sum_{k=0}^p \beta_k (r_i - r_0)^k \right\}^2 K_h(r_i - r_0) \\ & - \sum_{i \in A_T^c} \log \left\{ \Phi \left(\frac{\sum_{k=0}^p \beta_k (r_i - r_0)^k - T}{\sigma} \right) \right\} K_h(r_i - r_0) \end{aligned} \quad (5)$$

instead of the sum of squares (4), where Φ is the distribution function of $N(0,1)$, that is, a zero mean, normally distributed random variable with unit variance. The reason behind minimizing (5) is that it corresponds to maximizing the local likelihood

$$\prod_{i \in A_T} \left\{ \frac{1}{\sigma} \phi \left(\frac{I_i - \sum_{k=0}^p \beta_k (r_i - r_0)^k}{\sigma} \right) \right\}^{w_i} \prod_{i \in A_T^c} \left\{ \Phi \left(\frac{\sum_{k=0}^p \beta_k (r_i - r_0)^k - T}{\sigma} \right) \right\}^{w_i}$$

of observations I_i , that are possibly censored above T . The weights $w_i = K_h(r_i - r_0)$, could be interpreted as the proportions of observations from a larger sample being at r_i . We here implicitly assume the image noise ϵ_i to be i.i.d. and normal. Notice however that we generally, in the main assumption (2), assume the image noise to be correlated.

When referring to the two kinds of local quadratic kernel estimates dealt with above, we will sometimes call the two versions "the standard" and "the censored" version, respectively. We will also use "censored regression" for the latter.

Remark 1: In order to have the estimated intensity profile symmetric, as assumed in the model, we use the reflected data $\{-r_i(y), I_i\}$ together with $\{r_i(y), I_i\}$ as data when calculating the regression. However, we only evaluate \hat{f} at the points $r_i(y)$.

Remark 2: Note that the original local quadratic kernel estimate, defined by the minimizer of (4), is linear in the pixel values. This means that we can write

$$\hat{f}(r_i) = \sum_{j \in \mathcal{N}_x} W_{ij} I_j \quad (6)$$

for each $i \in \mathcal{N}_x$ for some matrix W . This matrix is called the *equivalent kernel matrix*. The trace of W , is a common estimate of the degrees of freedom for fitting of f to the data $\{r_i, I_i\}$, see Hastie and Tibshirani (1990, chap. 3). We will use this below when estimating the variance σ^2 . Note that W only depends on the bandwidth and the distances $r_i = r_i(y)$.

In contrast to the local quadratic kernel estimate, the censored version, defined by (5), is not linear in the pixel values since it contains the non-linear term involving Φ . It should also be noted that \hat{f} is never linear in candidate particle center y .

Algorithm for censored regression

The problem with equation (5), is that the parameter σ^2 , the variance of the image noise ϵ_i , is unknown. Hence, this needs to be estimated as well. However, instead of minimizing (5) with respect to both β and σ^2 locally for each r_0 , we will use an iterative scheme, alternating between estimation σ^2 globally and updating the regression estimates.

Let f_i^0 denote the estimate from a standard local quadratic kernel estimate at point r_i . The first, pilot, estimate of σ^2 , is based on the residuals between f^0 and the pixel values:

$$\hat{\sigma}_0^2 = \frac{1}{|A_T| - \text{tr}\{W\}} \sum_{i \in A_T} \{I_i - f_i^0\}^2 \quad (7)$$

where $|A_T|$ is the number of elements in A_T , and W the equivalent kernel matrix defined in equation (6) above for the standard local quadratic estimate.

Next, minimize equation (5) at each point r_i using the pilot estimate $\hat{\sigma}_0^2$ and denote this intensity profile estimate f_i^1 for $i \in \mathcal{N}_x$. Then, estimate σ^2 by (7) again (using the same W as before), however this time using f^1 instead of f^0 .

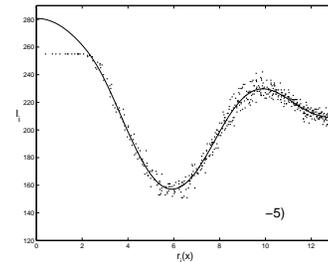


Figure 5: The resulting estimate of the intensity profile for the particle at depth -5 from Figure 2. For particles with censoring for small r -values it is essential to use reflected data when calculating the estimate (see Remark 1 after equation 5). Otherwise the estimated profile would have had a large negative derivative.

This above is repeated until the relative difference $|\sigma_k^2 - \sigma_{k+1}^2|/\sigma_k^2$ between the two consecutive estimates of σ^2 is smaller than a certain limit δ , chosen such that an update gives a negligible effect on the minimization of (5). Here, we let $\delta = 0.01$ and the iterative scheme typically converges after two or three iterations.

In Figure 5 we have plotted the resulting intensity profile to the particle at depth -5 in Figure 2. Without reflecting the data around $r = 0$ (see Remark 1 above), the estimated f would have been quite different, with a large negative derivative of the estimated profile at zero.

Bandwidth selection: cross-validation

Denote by \hat{f}_h^{-i} the (standard) local quadratic kernel estimate of f when leaving out the i :th data point, and using h as bandwidth. For each h we construct the cross-validation score

$$CV(h) = \frac{1}{n} \sum_{i \in \mathcal{N}_x} \{I_i - \hat{f}_h^{-i}(r_i(x))\}^2. \quad (8)$$

In practice, CV is computed for a finite number of suitably chosen bandwidths. For the data considered in this paper, bandwidth between 0.4 and 1.2 have turned out to be suitable choices. The bandwidth minimizing the cross-validation

is denoted by h_{opt} . When pixel values are censored, the summation in (8) is restricted to A_T , the uncensored pixels.

When estimating the centers, a pilot bandwidth h_{pilot} will first be used to find the minimizer of S . Then, a cross-validation study will be performed and the resulting h_{opt} will be used to re-estimate the center, if needed. As pilot bandwidth $h = 0.7$ has turned out to work well for estimating the center for all appearances of particles considered. In fact, the center estimation procedure is highly insensitive to the choice of bandwidth. Simulation studies have also shown that the difference between updating and not updating the bandwidth for a second re-estimation of the center is practically negligible. For the estimation of f , it is more important however. Choosing a too large bandwidth causes a large bias in the estimation and picking a too small bandwidth increases the variance.

3 Particle Center Estimation

Given a neighborhood \mathcal{N}_x to a particle with center in x , we estimate the center by the minimizer \hat{x} of

$$S(y) = \sum_{i \in A_T} \{I_i - \hat{f}(r_i(y))\}^2 \quad (9)$$

for $y \in \mathcal{N}_x$, where \hat{f} is the local quadratic kernel estimate using $\{r_i(y), I_i\}$ for $i \in \mathcal{N}_x$ as data, as discussed in the previous section. Furthermore, A_T is the set of uncensored pixels, that is, pixels below the censoring limit T .

Note that in (9), it is the covariates $r_i(y)$ that change with y . The response variables, the pixel values I_i for $i \in \mathcal{N}_x$, are the same for each y .

Choice of \mathcal{N}_x

Before the minimization of (9), we need a neighborhood \mathcal{N}_x of pixel locations to x . However, when x is unknown, so is \mathcal{N}_x . Nevertheless, if we are given an approximate particle center y_0 , manually or by some automatic image analysis method, we can let $\mathcal{N}_x = \mathcal{N}_{y_0}$ where \mathcal{N}_{y_0} is defined by equation (1) for $x = y_0$ for a suitable choice of r_{max} .

The choice of r_{max} is a balance between wanted precision of the estimated center on one hand, and computational effort and distance to nearby particles on the other hand. Picking a large r_{max} increases the number of pixels in the sum (9) and we therefore expect the estimate to get better. This vague argument is of course only true to a certain extent, since assumption (2) is in practice only

valid up to a certain size of the neighborhood, since there are other interfering particles present in the image. This is realized after looking at the sequence image of Figure 1. In the simulation study, we will use $r_{max} = 15$ and for the particles in the sub-images of Figure 2, we will take the entire sub-image as neighborhood.

Algorithm for estimating the center

We assume that a first, candidate position y_0 is given to us. We also assume that a suitable neighborhood \mathcal{N}_x is picked according to the discussion in the previous section. Furthermore, a pilot bandwidth h_{pilot} should also be chosen. For the data considered here, we use $h_{pilot} = 0.7$.

To calculate $S(y)$ at a candidate center y , do as follows:

1. Calculate the distances $r_i(y)$ for the candidate center y to all pixel locations $i \in \mathcal{N}_x$.
2. Compute $\hat{f}(r_i(y))$ for each point of evaluation $r_i(y)$ using bandwidth h_{pilot} .
3. Calculate $S(y)$ according to equation (9).

The minimizer \hat{x} of S using h_{pilot} as bandwidth is a preliminary estimate of the particle center x . A cross-validation study is then conducted using the data $\{r_i(\hat{x}), I_i\}$ for $i \in \mathcal{N}_x$ and the minimizing bandwidth is called h_{opt} . After this, the center is re-estimated as above using h_{opt} as bandwidth.

The actual optimization scheme can be chosen according to personal taste. We used a greedy search: compute the value of $S(y)$ for values y in a 5 times 5 uniform mesh with distance dy to nearest horizontal and vertical neighbour. Find the minimizer y_1 among these, and re-calculate $S(y)$ for a similar mesh, but this time centered above y_1 and with inter-distance $dy/2$.

Remark: In Step 2, $\hat{f}(r_i(y))$ is calculated either using the algorithm for censored regression in Section 2, or using the standard local quadratic kernel estimation. However, note that finding the minimizer of equation (5) is much more computationally costly than finding the minimizer to equation (4). Therefore, one way of reducing the computational cost when estimating the center, would be to first use the standard version to find a preliminary estimate as above. After the preliminary estimate has been found, an update in position is done, this time using the censored version of the local quadratic kernel estimate. For the degrees of censoring present in the data considered here, simulation studies has shown that

the difference in precision of the particle center estimate is practically negligible between the two methods of local quadratic kernel estimates.

Standard Error Estimates

Estimates on the standard error on particle center estimation will be derived using a sandwich estimator, which allows for the residuals to be correlated. See for example Owen (2001).

Recall that the estimate of the particle center x was the y minimizing

$$S(y) = \sum_{i \in \mathcal{A}_\tau} \left\{ I_i - \hat{f}(r_i(y)) \right\}^2$$

This expression is differentiable with respect to y since the Gaussian kernel is an infinitely differentiable kernel with unbounded support. If we had used a kernel with bounded support, a small perturbation in y can change the covariates $r_i(y)$ within the supported region centered at the point of evaluation r_0 in equation (4), and hence change the estimated value.

Assume that there are n pixels in \mathcal{N}_x and enumerate the pixels (arbitrarily) as i^1, i^2, \dots, i^n . (We use super-indices since sub-indices denotes the two coordinates of a pixel location.) We will now let g denote the \mathbb{R}^n valued function of $x = (x_1, x_2)$ with k th element $g_k(x) = \hat{f}(r_{i^k}(x))$. The reason is that we want to emphasize that we here consider the estimated intensity profile \hat{f} as a function of x .

Let J denote the Jacobian of g , that is the n times 2 matrix with element $(J(x))_{k,m} = \frac{\partial g_k}{\partial x_m}(y)$ for $k = 1, \dots, n$ and $m = 1, 2$. Furthermore, we let R be the vector of residuals at x :

$$R = R(x) = \begin{pmatrix} I_1 - g_1(x) \\ I_2 - g_2(x) \\ \vdots \\ I_n - g_n(x) \end{pmatrix},$$

where we have used the less awkward notation I_k instead of I_{i^k} , for $k = 1, \dots, n$. Moreover, let $\Sigma = \Sigma(x) = \mathbf{Var}\{R(x)\}$ denote the variance matrix of the residuals.

The sandwich estimator of the variance matrix $\mathbf{Var}\{\hat{x}\}$ is

$$V = (J^T J)^{-1} J^T \Sigma J (J^T J)^{-1}. \quad (10)$$

Since we do not know $J = J(x)$ and $\Sigma = \Sigma(x)$, we have to estimate or approximate them.

The standard way to approximate $J(x)$ would be by $J(\hat{x})$, but this involves differentiating each g_k with respect to x , which is very cumbersome. Instead, we approximate the derivatives of g_k in the following way. First we use the chain rule to get

$$\frac{\partial g_k}{\partial x_1} = \frac{\partial \hat{f}(r_{i^k}(x))}{\partial x_1} = \frac{\partial \hat{f}}{\partial r} \frac{\partial r_{i^k}(x)}{\partial x_1}.$$

Then we approximate the derivative of \hat{f} with respect to r with the estimate of the derivative of the intensity profile, namely $\hat{\beta}_1$ at the point of evaluation $r_{i^k}(x)$, that is, the estimate of the derivative of the intensity profile. Therefore we get:

$$\frac{\partial g_k}{\partial x_1} \simeq \beta_1(r_{i^k}(x)) \frac{x_1 - i_1^k}{r_{i^k}(x)}$$

since the derivative of the distance $r_{i^k}(x)$ with respect to x_1 is $(x_1 - i_1^k)/r_{i^k}(x)$. The analogous is done for the derivative of g_k with respect to x_2 . Finally, since x is unknown, we approximate x by \hat{x} . We denote the approximate J at point \hat{x} by \hat{J} .

Image noise model

Here, we assume that the image noise, and therefore the residuals, are spatially correlated with covariance function $\mathbf{Cov}\{\epsilon_i, \epsilon_j\} = \sigma^2 \exp(-c||i - j||)$ between pixels $i, j \in \mathcal{N}_x \subset \mathcal{D}_I$, for some constants σ^2 and $c > 0$, see Section 4. This has been observed to be a reasonable model when the noise structure of background images (that is, images where no particle is present) has been examined. Consequently, the variance matrix $\Sigma(x)$ is replaced by the estimate $\hat{\Sigma}$, with elements

$$\hat{\Sigma}_{k,\ell} = \hat{\sigma}^2 \exp(-\hat{c}||i^k - i^\ell||)$$

where $||i - j||$ is the Euclidean distance between the two pixel locations i and $j \in \mathcal{N}_x$, and $\hat{\sigma}^2$ and \hat{c} are the estimated parameters to the model $\gamma(\tau) = \sigma^2 \exp(-c\tau)$, using the observed residuals. See Section 4 for how to estimate σ^2 and c .

For a given estimate of a particle center, we let \hat{V} denote the estimated covariance matrix

$$\hat{V} = (\hat{J}^T \hat{J})^{-1} \hat{J}^T \hat{\Sigma} \hat{J} (\hat{J}^T \hat{J})^{-1}$$

when the estimated model parameters of the noise, $\hat{\sigma}^2$ and \hat{c} from the residuals are used in the expression for Σ .

The square-root of the diagonal elements of the estimated variance matrix \hat{V} , are called the *estimated standard errors*.

4 Results

In Figure 6 we display the reconstructed versions of the sub-images of Figure 2 after the centers and the intensity functions had been estimated. Below the reconstructed sub-images, we have displayed the spatial residuals between the original image and the reconstructed version.

The fit is clearly better for particles with more smoothly varying intensity profile, that is, for particles further away from the focal plane. This could be interpreted as the assumption of rotational symmetry does not fit very well for particles near the focal plane. However, small perturbations in the center estimate make large difference in the residuals when the underlying intensity profile is varying quickly. See Figure 7, where the pixel values surrounding the particle at the focal plane are plotted together with the corresponding estimate of the intensity profile. The residuals are in general larger for particles closer to the focal plane than for particles further away. This is probably due to that the model of rotational symmetry is not entirely correct for these depths, due to some unknown optical effect. Also, it may be so that the image noise is multiplicative rather than additive. This would partly explain why the noise seems to have larger variance and for particles with large fluctuations in the intensity profile.

Spatially correlated residuals

A close-up view of the residuals reveals that there seems to be some correlation between the residuals. On the left in Figure 8 we have displayed the sample autocovariance scores from the residuals between the original and the reconstructed version of the particle at depth -5. The sample autocovariance C_{km} for lags k and m , in the vertical and the horizontal direction, respectively, is defined as

$$C_{km} = N_{km}^{-1} \sum \{(R_{(i_1, i_2)} - \bar{R})(R_{(i_1+k, i_2+m)} - \bar{R})\}$$

where the summation is made over all pixel locations $i = (i_1, i_2) \in A_T$ at uncensored pixels, such that both i and the shifted version (i_1+k, i_2+m) is in A_T . Moreover, N_{km} is the number of elements in the summation, and \bar{R} is the sample mean of the residuals. Note that pixels that are censored are not included in the sample covariance. In Figure 8, the largest value is represented as white and the smallest as black. It is immediate that the residual are correlated.

On the right of Figure 8, we have plotted C_{km} as a function of $\tau = \sqrt{k^2 + m^2}$ instead. We have also plotted the fitted exponential correlation function $\gamma(\tau) = \sigma^2 \exp(-c\tau)$ in solid and in dashed, the nonparametric smooth of C_{km} using a

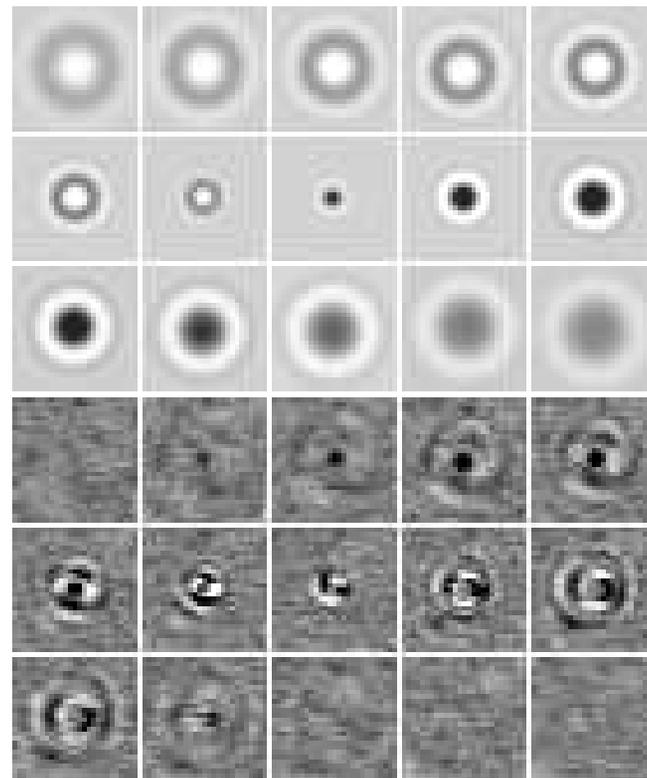


Figure 6: Above: The reconstructed versions of the sub-images in Figure 2. Below, we display the differences (magnified 5 times) between the original and the reconstructed versions. Clearly the fit is better for particles with less rapid changes in the intensity function.

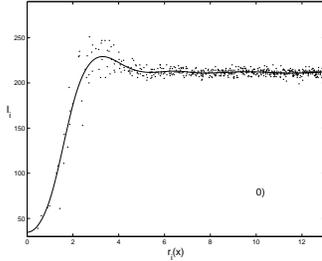


Figure 7: The resulting estimate of the intensity profile for the particle at the focal plane from Figure 2. Comparing f for this depth with the scatter plot, the rotational symmetry seems reasonable. However the residuals are larger than for particles further away from the focal plane, see Figure 4. This is reflected in the spatial residuals in Figure 6.

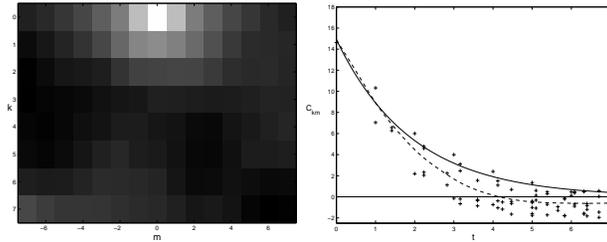


Figure 8: Sample autocovariance for the spatial residuals to fit of the particle at depth -5 in Figure 2. On the left they are displayed as image for horizontal m and vertical k lags separately, and on the right as a function of $\tau = \sqrt{k^2 + m^2}$, where we also have plotted the fitted exponential correlation function $\sigma^2 \exp(-c\tau)$ with $\sigma^2 = 14.9$ and $c = 0.51$ in solid. The dashed line is the nonparametric smooth of C_{km} using a local linear kernel estimator with bandwidth 0.5. As seen, the exponential correlation function is a bit pessimistic regarding the decay of the correlations.

local linear kernel estimator with bandwidth 0.5. The exponential correlation

function was fitted by estimating σ^2 and c by

$$\hat{\sigma}^2 = C_{00}$$

and

$$\hat{c} = -\log(\bar{C}_1/C_{00}),$$

where \bar{C}_1 is the mean of C_{01} and C_{10} , that is, the mean of the two sample autocovariance scores for lags of length 1.

Standard errors and estimated noise parameters

We estimated the particle center for another 5 particles present in the same kind of images (z-scans) as the particle in Figure 2, where we know the depth of the particles. For each particle, the noise parameters σ^2 and c were estimated as above, and the standard error for the center estimate, was estimated as described in Section 3.

In Figure 9 we have plotted the mean of the resulting standard errors for the 15 depths (the off-diagonal terms were negligible for all particles). Not surprisingly, the precision in the estimation depends on the appearance of the particle, and hence, on the depth of the particle. On the right in Figure 9, the mean of the corresponding estimates of σ^2 and c are displayed for each depth. The noise parameters also depend on the depth of the particles, probably due to the worse fit of the rotational symmetry model near the focal plane, as observed in Figure 6. As mentioned above in connection with Figure 7, multiplicative image noise could also be a contributing factor to the larger residuals near the focal plane.

5 Simulation Study

We conducted a simulation study using artificially produced images of particles to examine if the root-mean-square (RMS) error of the estimated particle centers vary with depth in the same way as the estimated standard errors of the real data displayed in Figure 9. Furthermore, we wanted to see if the estimated standard errors from the simulated images were consistent with the RMS errors. The third objective was to examine how much the RMS errors were affected when pixel values were censored but if we used the standard local quadratic kernel estimate in the calculation of S , instead of the censored version, see Section 3.

We created an artificial image of size 33 times 33 for a particle at depth z as follows. The true center for the particle was chosen as the middle pixel of the

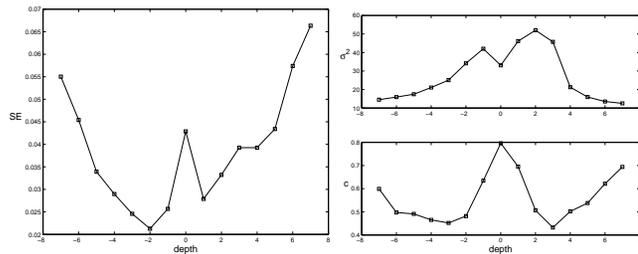


Figure 9: We estimated the particle center for 6 particles at each of the depths in Figure 2. On the left, the mean of the estimated standard errors of the particle center estimation is displayed and on the right the estimated noise parameters σ^2 and c to the exponentially decaying correlation function.

image (17,17) but with an offset chosen uniformly from $[-0.5,0.5]$ independently for both coordinates. We choose the particle center at random in order to get randomness in the distances $r_i(x)$. Then the pixel values for all pixel locations are given the value predicted by the corresponding intensity profile f to depth z . As true intensity profiles, we used the ones estimated for the particle in 2. To this we add zero mean Gaussian image noise with covariance function $\sigma^2 \exp(-c||i-j||)$ between pixels i and j in the image. The realization of this noise is done via a Cholesky factorization of the covariance matrix with the elements $\sigma^2 \exp(-c||i-j||)$. We used noise parameters $\sigma^2 = 25$ and $c = 0.6$, which was considered a suitable choice of parameters, both from looking at the estimated parameters in Figure 9 but also after an examination of the noise structure in background images (that is images with no particles present) where the noise variance were generally lower, around 13-15, and the correlation factor c around 0.6. Finally, the pixel values are rounded to the nearest integer. Regarding censoring, this was only done in one of the studies, see below.

The middle pixel in the artificial image was used as approximate starting center and the pixels within distance $r_{max} = 15$ from this center pixel was used as neighborhood \mathcal{N}_x . Bandwidth $h_{pilot} = 0.7$ was used as pilot bandwidth for all depths, see Section 3 for details on estimating the center.

We conducted three different simulation studies, each of which consisted of 500 simulated images with subsequent particle center estimation for each depth. In Figure 10 we have displayed the resulting errors in the center estimation as a boxplot. The results in the horizontal and vertical coordinates are presented

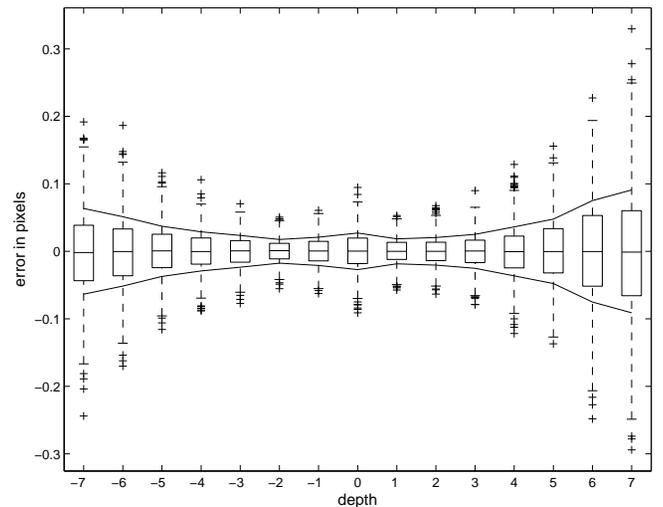


Figure 10: Boxplot of the resulting particle center estimation errors for 500 simulated images at each depth. The errors in both coordinates are presented simultaneously so each boxplot consists of 1000 values. The solid line is the calculated RMS error from the same simulation study for each depth. See Figure 11 for the consistency check of the estimated standard errors in this simulation.

together since there was no correlation between the coordinate position errors. The box has lines at the lower quartile, median, and upper quartile values and the length of the whiskers are 1.5 times the interquartile range, which is defined as the difference between the 75th percentile and the 25th percentile of the data. Outliers are displayed using the symbol '+'. The solid lines are the RMS errors plotted symmetrically on both sides of the x-axis.

Regarding the consistency of the estimated standard errors with the RMS errors, we have in Figure 11 plotted the RMS errors (boxes), and the mean of the 500 estimated standard errors for each depth (stars). It seems as if the estimated standard errors slightly over-estimates the standard errors. The qualitative shape of the two curves is however the same. In Figure 11, we have also

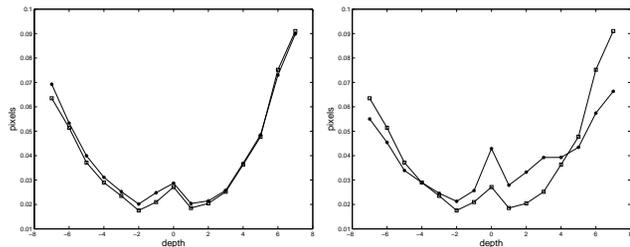


Figure 11: Left: Comparison between the RMS errors (boxes) of the particle center estimation errors (in both coordinates simultaneously) and the mean of the estimated standard errors (stars) for the simulation study of 500 simulated images presented in Figure 10. The standard errors slightly over-estimates the RMS errors, which could be due to the rather pessimistic noise correlation model, see Figure 8. Right: Comparison between the RMS errors (boxes), also plotted in the left plot, and the mean of the standard errors estimated from the real data (stars), which are also plotted in Figure 9.

plotted the estimated standard errors from the real data compared with the RMS errors from the simulation study. The most probable reason why the standard errors from the real data, is due to larger estimates of the variance of the noise for particle in the real data close to the focal plane, and vice versa for particles further away.

Two more simulation studies were conducted. In the first of these we wanted to compare the RMS errors if we censored pixel values in the images at 255 but did not take this into consideration when estimating the center, that is, we used the standard local quadratic kernel estimator when calculating S . The numbers of simulations were again 500 images per depth. The difference between the resulting RMS errors here and in the previous study, were not statistically significant on a 95 percent confidence level (based on a normal distribution assumption on the errors, which by a chi-square plot was seen to be highly plausible).

In the last simulation study the objective was to investigate how the precision in estimating the center was affected if only the even rows in the images were used for the estimation. The reason for this investigation is that for sequence images like the one in Figure 1, only half of the rows contain information. These kind of images are recorded in video format and only half of the rows are updated for each image. Therefore, we would perhaps expect a lower precision in the vertical coordinate.

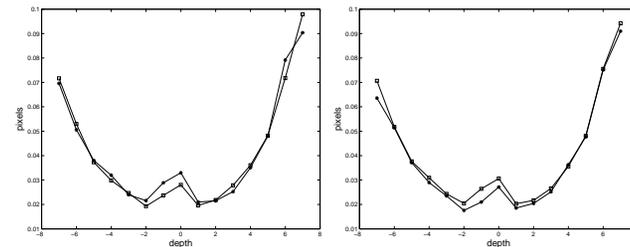


Figure 12: In the plot on the left, we have plotted the RMS errors for the horizontal (boxes) and the vertical coordinates (stars) separately for the simulation study where we only use the pixels on the even rows. The plot on the right is a comparison between the magnitude of the absolute values of the RMS errors in the two simulation studies, using only even rows (boxes) and using all rows (stars).

In Figure 12, the resulting RMS errors for the two coordinates separately are presented and quite surprisingly, the difference between the errors in the two coordinates is small. The depths for which the difference between the RMS errors are statistically significant on a 95 percent confidence level, are for depths -1, 0, and 1. Furthermore, considering both coordinates simultaneously, in the plot on the right, the RMS errors are hardly affected even though only half as many pixels are used in the estimation.

The reason for the first observation that the errors are almost the same in both coordinates, is probably that the rotational symmetry imposes such a strong condition on the estimation procedure so that all pixels contribute equally, disregarded of the fact that we actually have half as many pixels in one direction. For the particles closest to the focal plane, the particles are smaller, and hence there are fewer pixels involved. Possibly more surprising is that the magnitude of the errors hardly increases. This is probably due to that the image noise is correlated; the relative increase in effective sample size is not 1/2 but much smaller since the observations are correlated.

6 Conclusions and Discussion

Our method of estimating the centers of spherical particles in images is possible up to precision well beyond sub-pixel accuracy. Since it only relies on the

assumption of rotational symmetry of the appearance of the particles in the images, it can be used for tracking even when the appearance of the particles changes (just as long as the rotational symmetry assumption still holds). We therefore believe it to be a highly versatile tool for automatic measurements in video microscopy of for example colloidal suspensions.

The simulation study showed that the estimated standard errors are consistent with the RMS errors. It is however necessary that the assumptions on rotational symmetry and the isotropy of the image noise are correct. For the real data in Figure 2, the spatial residuals showed an increasingly worse fit for particles near the focal plane (see Figure 6). One reason could be that the image model in equation (2) with additive image noise is not correct. In Figure 9, the estimated noise parameters from the real images show signs of some kind of dependence of depth, and hence of the appearance of the particles. It could be so that the image noise is multiplicative, since we seem to have larger variance and shorter correlation length (larger c) for particles near the focal plane.

The estimator for the noise parameters is clearly biased. An improvement in this estimation might give better results in the estimation of the standard error. However, the standard error estimator, predicts the magnitude of the errors fairly well, see Figure 11, and it is only here that the noise estimates are needed.

An alternative measure of the local rotational symmetry to use when pixels are censored, is

$$S_{att}(y) = \frac{1}{\sigma^2} \sum_{i \in A_T} \{I_i - \hat{f}(r_i(y))\}^2 - \sum_{i \in A_c^*} \log \left\{ \Phi \left(\frac{\hat{f}(r_i(y)) - T}{\sigma} \right) \right\}$$

where, as usual, A_T is the set of uncensored pixels in \mathcal{N}_x and \hat{f} the censored version of the local quadratic kernel estimate using the data $\{r_i(y), I_i\}$. Here, σ^2 is the estimated variance from equation (7). This measure of rotational symmetry takes care of the censored pixels in an appropriate manner. Note that the equation above equals (3) when there is no censoring. However, since the proportion of censored pixels is so small here, it would probably not change the result.

The result of the simulation study where we only used the even rows in the images, is interesting; see Figure 12. First since the difference between the particle estimation errors in the vertical and the horizontal coordinates were very small. Second because the magnitude of the RMS error did not decrease with a factor $\sqrt{2}$ as one might have expected considering that we used half as many pixels in the particle center estimation. Both these observations are good news for estimating particle centers in images recorded in video rate, where only half of the rows in the images are updated at each recording instant.

There is no known parametric form for the intensity profiles. In fact, imaging of spherical objects is still a topic of large theoretical interest in optics research, see for example Ovrn and Izen (2000). Ray-tracing using Fourier optics (see for example a standard textbook on optics as Hecht (1998)) does not work here since the size of particles is in the scale of the wavelength of light used. An alternative and more advanced approach to Fourier optics is to use Mie-theory, which was used by Ovrn and Izen (2000) to predict the appearance of a polystyrene sphere of diameter $7\mu\text{m}$. However, no closed functional form of the intensity profiles is known. For a review of colloidal suspensions in microscopy see Elliot and Poon (2001).

References

- J. Crocker and D. Grier. Methods of digital video microscopy for colloidal studies. *Journal of Colloid and Interface Science*, 179:298–310, 1996.
- J. Crocker and D. Grier. Interactions and dynamics in charge-stabilized colloid. *MRS Bulletin*, 23:24–31, 1998.
- C. Ekström, S. Bak, C. Kristensen, and M. Rudemo. Spot shape modelling and data transformations for microarrays. *Bioinformatics*, 20:2270–2278, 2004.
- M. S. Elliot and W. Poon. Conventional optical microscopy of colloidal suspensions. *Advances in Colloid and Interface Science*, 92:133–194, 2001.
- J. Fan and I. Gijbels. *Local Polynomial Modelling and Its Applications*. Chapman and Hall, London, first edition, 1996.
- C. Glasbey, T. Forster, and P. Ghazal. Estimation of expression levels in spotted microarrays with saturated pixels. *submitted*, 2005.
- T. Hastie and R. Tibshirani. *Generalized Additive Models*. Chapman and Hall, London, first edition, 1990.
- E. Hecht. *Optics*. Addison-Wesley, Reading, Massachusetts, third edition, 1998.
- B. Ovrn and S. Izen. Imaging of transparent spheres through a planar interface using a high-numerical-aperture optical microscope. *Journal of the Optical Society of America A. Optics, Image Science and Vision*, 17:1202–1213, 2000.
- A. Owen. *Empirical Likelihood*. Chapman and Hall/CRC, London, first edition, 2001.

Depth estimation of colloidal particles in microscopy

Mats Kvarnström

January 24, 2005

Abstract

Estimates of the locations of particles are crucial for subsequent estimation of interaction of particles. It is also of interest for tracking particles in three dimensions. Here we focus on the estimation of depth of colloidal latex particles observed in bright-field video microscopy. Precise estimation of the particle centers in 2-D of spherical particles in microscopy is covered in Kvarnström and Glasbey (2005). Here, we introduce a method for estimating the depth of spherical colloidal particles, using empirically derived templates. The criterion function used for finding the best correspondence between the template and the image takes care of possible censoring of pixel values in the image and different levels of background intensity. From both real data and a simulation study, the conclusion is that the depth estimation has a standard error below at least $0.2\mu\text{m}$, which is the corresponding distance in depth between the templates used.

1 Introduction

In colloidal chemistry, it is crucial to quantitatively be able to measure the stability of the colloidal system of particles. Digital microscopy offers vast opportunities for automated measurements of interaction and diffusivity of particle systems, both of which are important factors when determining the stability. However, estimating interaction of particles in a three dimensional domain demands position estimates of all three coordinates. In Figure 1 we give an example of an image that could be used for measurements of the properties of a colloidal suspension. The particles in the images are diffusing in a dilute solution and the image is taken from a larger sequence of images, recorded at a frame rate of 50 images per second. The particles are made of latex and are all of the same size,

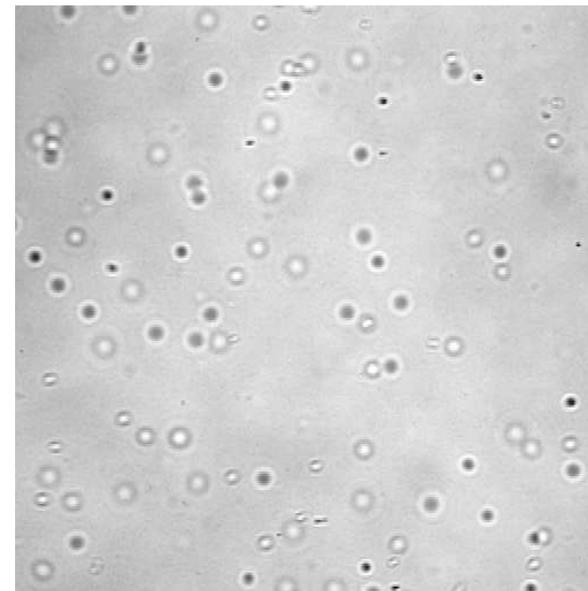


Figure 1: An image of latex particles diffusing in a dilute solution. The difference in appearance is due to an out-of-focus effect.

494 nm in diameter. Difference in appearance of particles in a given image is due to an out-of-focus effect caused by difference in depth relative to the focal plane. It is this difference in appearance that will be used when estimating the depth of the particles.

The image in Figure 1 consists of 512 times 512 square pixels of side-length 180 nm, so the distance is roughly $90\mu\text{m}$ across the horizontal and vertical domain of the image. The exact maximum deviation of the particles from the focal plane in this experimental setup is unknown but it is believed to be at least $\pm 10\mu\text{m}$. In Figure 2, the sub-images show the appearance of particles at known depths relative to the focal plane. Images like these will be used to construct templates

of what the particles look like, at a set of different depths. The templates could then be used for depth estimation of particles in images like the sequence images of Figure 1. Previous studies of measurements of colloidal particles using digital microscopy include Crocker and Grier (1996, 1998), however their studies concerned systems of particles efficiently confined at a single depth. For an overview of colloidal chemistry in general, see Evans and Wennerström (1999) and for an overview of microscopy used for observing colloidal suspensions, see Elliot and Poon (2001).

Examples of depth estimation in vision and image processing are several. The problem formulation is however usually rather different, as is the solution. One application is reconstruction of 3-D scenery from video sequences. For example, in Chowdhury and Chellappa (2001), this is done by tracking correspondences of various landmarks of geometrical objects (such as corners of a building) and relate these to the known movements of the camera. Another application is to determine the distance from the camera to various objects in an image scenery, see for example Gil et al. (2004) and Ahn et al. (1997). The underlying assumption is that the images of objects not in focus, are convolutions of the original, true image, and a linear (known) low-pass filter. The depths of the objects in the image are then estimated by means of estimating the amount of smoothing in the observed image, either by a deconvolution or a measure of the fraction of high-pass versus low-pass components. None of these techniques are applicable here since there are obviously no suitable landmarks on the latex particles, and the different appearances of the de-focused particles cannot be represented by a linear operation.

Precise estimation of the centers in the image plane (2-D position estimation), is covered in Kvarnström and Glasbey (2003). This estimation was based on the assumption of rotational symmetry of the pixel values surrounding a particle center. The standard errors of the estimates were both estimated and verified in a simulation study to be between 0.02 and 0.10 pixels, depending on the distance in depth from the focal plane, with larger deviation for particles further away from the focal plane. This corresponds to standard errors between 3.6 and 18 nm. Furthermore, a method was developed to estimate the functional form of how pixel values change with distance (in the image domain) from the particle center. In the present paper, the *intensity profiles*, for particles at different (known) depths will first be estimated and then used as *template profiles* to determine the depths of particles for which we do not know the true depth. This kind of approach is called *template matching*. The intensity profiles (or templates) will be estimated nonparametrically using a local quadratic kernel estimate. From now on, we will by particle center mean the 2-D center, that the location of the particle center in the image plane. The third coordinate will be

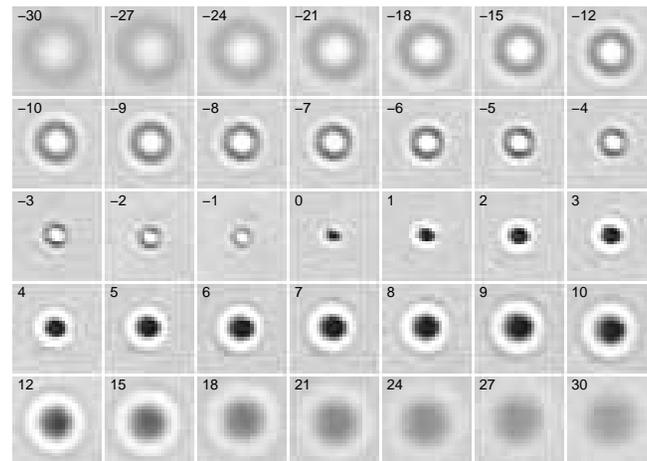


Figure 2: The appearances of the particles at 35 different depths relative to the focal plane. Sub-image 0 corresponds to a particle at the focal plane, and sub-images with negative and positive labels are below and above the focal plane, respectively. The distance in depth between two consecutive label numbers is $0.2 \mu\text{m}$.

called depth and denoted by z .

Template matching is a fairly common method in image processing where the best correspondence between an image and a set of templates is sought. Examples where the method has been used successfully, is in Young et al. (1998) to automatically identify and measure yeast cells in DIC microscopy, and Dralle and Rudemo (1997), where positions of trees are estimated from areal photographs. Here, our set of templates should mimic the appearance of particles at different depth of focus. Given an image of a particle at an unknown depth, the corresponding depth of the template that gives the best correspondence, would then be the estimate of the depth of the particle in the image. We therefore need both to construct the set of template profiles and to define a suitable criterion function to find the “best correspondence”.

The templates will be constructed by estimating the intensity profiles in images of

particles at known depths, like the sub-images of Figure 2. Images like these were recorded by first letting particles adsorb on one of the confining glasses of the specimen, and then move the specimen relative to the optics of the microscope. We have 61 *z-scan* images like these to our disposal, with distance to the focal plane ranging from $-6 \mu\text{m}$ to $+6 \mu\text{m}$ with $0.2 \mu\text{m}$ between each. In Figure 2, we display a subset of the total number of depths. For each *z-scan*, there are 10 particles present and believed to be totally adsorbed on the glass, and hence, at the same relative depth. In each sub-image in Figure 2 we have extracted the neighborhood around a particle in the *z-scan* for the depth specified by the corresponding label. We have used the same particle in all sub-images except for depths -2, -1, 8, and 12. In the *z-scans* for these depths, the original particle was occluded by other, mobile particles in the specimen. Therefore, for sub-images -2, -1, 8, and 12, another particle was used. We are not able to present the appearance of the same particle at each depth since all of the 10 adsorbed particles are occluded in at least one of the 61 *z-scans*.

The criterion function used here is based on least squared distance between the template profile and the pixel values surrounding the particle center. Care must however be taken since censoring in the images occurs for pixel values larger than an upper limit T . Also, particles in the images are generally at different background pixel levels. This affects the relative level of censoring for particles at different background levels; a particle at a location with large background intensity has more pixels censored than a particle with low background intensity. Hence, the criterion function should also take this into account. We will compare the performance between five different criterion functions and find out that taking care of different background levels is far more important than taking care of censored pixel values.

2 Assumptions

The underlying assumption of the appearances of the particles in the images we are considering here, is the rotational symmetry of the pixel values close to the true particle center. More precisely, for a particle located at $x = (x_1, x_2) \in \mathbb{R}^2$, we assume that there exists a neighborhood \mathcal{N}_x of (integer valued) pixel centers $i = (i_1, i_2)$ such that

$$I_i = f_z(r_i(x)) + \alpha + \epsilon_i \quad (1)$$

for $i \in \mathcal{N}_x$, where I_i is the pixel value at i and $r_i(x)$ is the Euclidean distance from the particle center x to the pixel center i . We denote the domain for which the image I is defined by \mathcal{D}_I . The function f_z is called the *intensity profile at depth* z of a particle. It is assumed to be a smooth function with at least second

order continuous derivative. Also, we assume that f_z is an even function of r (even though r is never negative), so that the derivative of f_z at $r = 0$ is equal to zero, that is $f_z'(0) = 0$. The image noise, ϵ_i for $i \in \mathcal{D}_I$, is assumed to be normally distributed, with (isotropically) correlated pixel values, which is what have been observed in the images considered here. The background level α is generally different for each particle.

Regarding the neighborhood \mathcal{N}_x , there are several choices. For the particles in Figure 2, we can take all pixels in the corresponding sub-image as the neighborhood \mathcal{N}_x . More typically however, we let

$$\mathcal{N}_x = \{i \in \mathcal{D}_I : r_i(x) \leq r_{max}\} \quad (2)$$

for an appropriately chosen maximum distance r_{max} . For images like Figure 1, when the distance between two particles is closer than $2r_{max}$, we might however need to exclude some pixels from the two neighborhoods. This will however not be dealt with in this paper. Throughout this paper, we will use a neighborhood like the one given in (2), with r_{max} equal to 15.

As mentioned earlier, the pixel values in the images are censored above an upper limit T . Here, T equals 255. Censoring has to be dealt with properly, both when estimating the template profiles and in the subsequent matching. In Figure 3, the scatter plots illustrate what censoring typically looks like. Each scatter plot consists of the pixel values I_i and the distances $r_i(y)$ from a candidate particle center y of the particle. The scatter plots on the right are for distances $r_i(\hat{x})$ from the corresponding estimates of the particle centers. Henceforth, when we refer to a scatter plot for a particle, we mean the plot of pixel values I_i to the distances $r_i(y)$ for a candidate or estimated particle center y .

We assume the noise in the images to be zero-mean additive Gaussian noise with isotropic covariance $\sigma^2 \exp(-c\|i - j\|)$ between pixel centers i and $j \in \mathcal{D}_I$. The constants σ^2 and c , are the variance and the inverse correlation length, respectively and $\|\cdot\|$ denotes the Euclidean distance. This was found in Kvarnström and Glasbey (2005) to be a reasonable approximation of the correlation structure of the noise in the considered images. Even though the noise assumed to be correlated, when estimating the intensity profiles, we proceed in a quasi-likelihood approach, implicitly assuming independent errors ϵ_i . In fact, the correlation structure will in this paper only be used when simulating image noise in the simulation study. The noise parameters $\sigma^2 = 25$ and $c = 0.6$ are thought to be sufficiently conservative estimates.

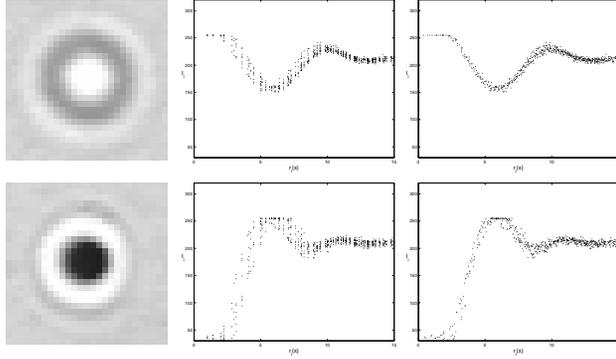


Figure 3: Here we have zoomed in on depth -15 (top row) and 8 (bottom row) of the particle in Figure 2. The middle column show the scatter plots of pixel values and distances from the center pixel in the sub-images, and the right columns the scatter plots after refining the center (separately for each). Notice the censoring of pixel values above 255.

3 Template construction

Figure 2 shows what a particle looks like at 35 different depths. In sub-image 0, the particle is at the focal plane. Negative labels correspond to particles below, and positive labels to particles above the focal plane. The distance in depth between two consecutive labels is $0.2 \mu\text{m}$.

The scatter plots in the middle and the right columns of Figure 3 for the two particles, illustrate the need to refine the centers of the particles at greater accuracy than to the (integer-valued) pixel centers.

For a particle at known depth z , the template profile is constructed as follows:

1. Refine the position estimate by finding the position with maximal rotational symmetry by minimizing equation (3) below.
2. Find the bandwidth h_{opt} that minimizes the cross-validation score for the local quadratic kernel estimate used on the pixel values I_i and the distances to the estimated center.

3. Estimate the intensity profile with the bandwidth h_{opt} . If there are censored pixels, use censored regression as described below.

The resulting intensity profile estimate for this depth z is denoted by f_z .

Much of the material found in this section can be found in greater detail in Kvarnström and Glasbey (2005).

3.1 Estimating centers in 2-D

We estimate the particle center by minimizing

$$S(y) = \sum_{i \in \mathcal{N}_z} \{I_i - g(y, r_i(y))\}^2 \quad (3)$$

for the particle located at x (see Kvarnström and Glasbey (2005)). Here, $g(y, \cdot)$ is a scatter smooth of the data points $\{r_i(y), I_i\}$ calculated by means of a local quadratic kernel estimate with a Gaussian kernel function (see below). We let \hat{x} denote the minimizer of S and call it the estimate of x . The idea behind minimizing equation (3), is to find the position of (local) maximal rotational symmetry.

For easier notation, we will henceforth write r_i instead of $r_i(y)$ when there is no possibility of misunderstanding. We will now explain briefly how the local quadratic kernel estimate is calculated. Standard references for nonparametric local polynomial kernel estimation are Hastie and Tibshirani (1990) and Fan and Gijbels (1996). Let K denote the Gaussian kernel, that is $K(x) = \exp(-x^2/2)$. The local quadratic kernel estimate is based on the assumption that the unknown regression function f can be approximated locally by a polynomial of degree two,

$$f(r) \simeq \sum_{k=0}^2 \frac{f^{(k)}(r_0)}{k!} (r - r_0)^k$$

for r close to r_0 . To calculate the value of the estimate at an arbitrary point r_0 , we locally fit a quadratic polynomial using weighted least-squares by minimizing

$$\sum_{i \in \mathcal{N}_z} \left\{ I_i - \sum_{k=0}^2 \beta_k (r_i - r_0)^k \right\}^2 K_h(r_i - r_0) \quad (4)$$

with respect to $\beta = (\beta_0, \beta_1, \beta_2)$ and denote the minimizer by $\hat{\beta}$. Here, $K_h(x) = h^{-1}K(x/h)$ is the kernel with bandwidth $h > 0$. The estimates of $f^{(k)}$ at r_0 , for $k = 0, 1$, and 2 are

$$\hat{f}^{(k)}(r_0) = k! \hat{\beta}_k$$

The value of g at the point $r_0 = r_i(y)$ for the candidate center y , is then $g(y, r_i(y)) = \hat{f}^{(0)}(r_i(y))$. For each candidate center y in the evaluation of $S(y)$ in expression (3), this is done for all distances $r_i(y)$ to the pixel center $i \in \mathcal{N}_x$. Notice also that since g is the solution to a weighted least-squares, $g(y, r_i(y))$ is a linear combination of the pixel values I_j for $j \in \mathcal{N}_x$.

For the data considered here, the local quadratic model is not very sensitive to choice of bandwidth. This was shown in a cross-validation study in Kvarnström and Glasbey (2005), where $h = 0.7$ turned out to be a good choice of a *pilot bandwidth* when estimating the particle center, for the depths considered here. The bandwidth is then updated, via a cross-validation study, as we get closer to the minimizer of S . How to conduct a cross-validation is explained below.

3.2 Estimating the intensity profiles

When the particle center has been found, we estimate the intensity function f_z .

First of all, the bandwidth has to be updated from the pilot bandwidth $h = 0.7$ in the center estimation. The bandwidth minimizing the cross-validation score CV , defined as

$$CV(h) = \sum_{i \in \mathcal{N}_x} \{I_i - \hat{f}_h^{-i}(r_i(\hat{x}))\}^2,$$

is denoted by h_{opt} , where \hat{f}_h^{-i} is the local quadratic kernel estimate using all pixels except the i th. In practice, the set over which CV is calculated, is finite. Here, we calculate CV for $h = 0.5, 0.6, \dots, 1.5$.

When none of the pixel values I_i are censored, f_z is re-estimated with the local quadratic kernel estimate, this time with h_{opt} as bandwidth. If there is censoring however, we will use a slightly modified estimation scheme.

Censored regression

Let A_T denote the set of uncensored pixels, that is $A_T = \{i \in \mathcal{N}_x : I_i < T\}$. Hence, the complement A_T^c is the set of censored pixels. For an arbitrary point r_0 , denote the minimizer of

$$\begin{aligned} & \frac{1}{2} \log \sigma^2 \sum_{i \in A_T} K_h(r_i - r_0) + \frac{1}{2\sigma^2} \sum_{i \in A_T} \left\{ I_i - \sum_{k=0}^2 \beta_k(r_i - r_0)^k \right\}^2 K_h(r_i - r_0) \\ & - \sum_{i \in A_T^c} \log \left\{ \Phi \left(\frac{\sum_{k=0}^2 \beta_k(r_i - r_0)^k - T}{\sigma} \right) \right\} K_h(r_i - r_0) \end{aligned} \quad (5)$$

by $\hat{\beta}$. Here, Φ is the distribution function of a standard normal random variable. The problem with equation (5), is that the parameter σ^2 , the variance of the image noise ϵ_i , is unknown. Hence, this needs to be estimated as well. However, instead of minimizing (5) with respect to both β and σ^2 locally for each r_0 , we will use an iterative scheme, alternating between estimation σ^2 globally and updating the regression estimates.

Let \hat{f}_i^0 denote the estimate from a standard local quadratic kernel estimate at point r_i . The first, pilot, estimate of σ^2 , is based on the residuals between \hat{f}_i^0 and the pixel values:

$$\hat{\sigma}_0^2 = \frac{1}{|A_T|} \sum_{i \in A_T} \{I_i - \hat{f}_i^0\}^2 \quad (6)$$

where $|A_T|$ is the number of elements in A_T . Next, minimize equation (5) at each point r_i using the pilot estimate $\hat{\sigma}_0^2$ and denote this intensity profile estimate \hat{f}_i^1 for $i \in \mathcal{N}_x$. Then, estimate σ^2 by (6) again, this time using \hat{f}_i^1 . Repeat this until the relative change $|\sigma_k^2 - \sigma_{k+1}^2|/\sigma_k^2$ between two consecutive estimates of σ^2 is smaller than a certain limit δ , chosen such that an update gives a negligible effect on the minimization of (5). Here, we let $\delta = 0.01$ and the iterative scheme typically converges after two or three iterations.

Remark: The estimator of σ^2 in equation (6) is biased. In Kvarnström and Glasbey (2005) it was proposed to use $|A_T| - \text{tr}\{W\}$ in the denominator instead of $|A_T|$. Here, W is the *equivalent kernel matrix*, for which $\hat{f}(r_i(y)) = \sum_{j \in \mathcal{N}_x} W_{ij} I_j$, where \hat{f} is the standard local quadratic kernel estimate. (Remember that this was linear in the pixel values since it is the solution to a weighted least-squares problem.) Thus, $|A_T| - \text{tr}\{W\}$ is an estimate of the degrees of freedom in the local quadratic kernel estimate. An better alternative is therefore to use this in the denominator.

Resulting intensity profiles

The set of final estimates of the profiles are called template profiles and denoted by $\{f_z\}$, thus dropping the hat-sign for easier notation. They are calculated and stored for 200 equally spaced points between 0 and 15. In Figure 4, the resulting profiles for the 35 depths corresponding to the sub-images of Figure 2 are shown.

Intensity profiles of particles closer to the focal plane have larger fluctuations than the corresponding intensity profiles for particles further away. Furthermore, for particles below the focal plane, the intensity profile assumes large values close to $r = 0$, whereas the opposite is valid for particles above the focal plane (compare with Figure 2). Notice that the estimated intensity profiles can

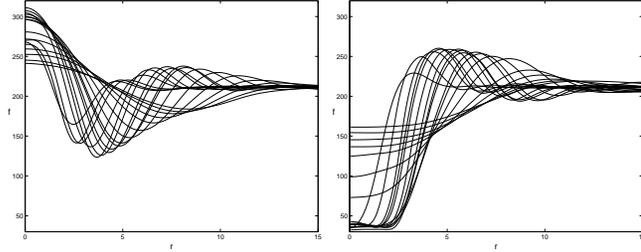


Figure 4: The resulting template profiles for the sub-images of Figure 2. The intensity profiles corresponding to sub-images with negative labels are plotted on the left and the ones with positive labels, including the particle at the focal plane, on the right.

assume larger values than the censoring level $T = 255$. For particles below the focal plane, where censoring occurs for small r -values, this was possible since we assume the true profile to be an even function of r , which is natural considering the assumptions of rotational symmetry and the smoothness of the intensity profile.

4 Template matching

Let us assume that the center x of a particle is given. Then our data consist of $\{r_i, I_i\} = \{r_i(x), I_i\}$, of pixel values and distances to x . Having constructed the template profiles in Section 3, we are to search for the template profile f_z that gives the best correspondence to the data $\{r_i, I_i\}$. This correspondence can be measured in a number of ways.

Arguably the most popular criterion is the square difference between the template and the data, which in our case would be to let the minimizer of

$$\sum_{i \in \mathcal{N}_z} \{I_i - f_z(r_i(x))\}^2$$

with respect to the profiles f_z , be the estimate of the depth. This criterion function does not however take into account either censored pixels or different background levels α . Since both these issues are typical for our data, a modification is called for.

Assume first that the image noise variance σ^2 and background level α are both known. Then the minimizer of

$$\frac{1}{\sigma^2} \sum_{i \in \mathcal{A}_T} \{I_i - \alpha - f_z(r_i(x))\}^2 - \sum_{i \in \mathcal{A}_T^c} \log \left\{ \Phi \left(\frac{f_z(r_i(x)) + \alpha - T}{\sigma} \right) \right\} \quad (7)$$

would be an estimate of the depth z taking into account both censoring and different background levels. As before, T is the censoring level, and Φ the cumulative distribution function of a standard normal random variable. In fact, this estimate of z is the maximum likelihood estimate of z under the assumption that the pixel values I_i are

$$I_i = f_z(r_i(x)) + \alpha + \epsilon_i$$

for some f_z in our set of templates, the true 2-D center is at x , and that the image noise variables ϵ_i are i.i.d. and $N(0, \sigma^2)$ -distributed. Since the templates f_z are stored only for 200 equally spaced values over the interval $[0, 15]$, the values of f_z at the points $r_i(x)$ are evaluated by linear interpolation from these stored values of the template profiles.

Regarding σ^2 , we will use the residuals between the pixel values I_i and the local regression smooth $g_i = g(x, r_i(x))$ from the particle center estimation procedure of minimizing (3) using only the uncensored pixel values, in exactly the same that we did in equation (6).

For α , we will use a profile likelihood approach, estimating α separately for each candidate template profile f_z . Thus, the minimizer of

$$M(z) = \frac{1}{\sigma^2} \sum_{i \in \mathcal{A}_T} \{I_i - \hat{\alpha}_z - f_z(r_i(x))\}^2 - \sum_{i \in \mathcal{A}_T^c} \log \left\{ \Phi \left(\frac{f_z(r_i(x)) + \hat{\alpha}_z - T}{\sigma} \right) \right\} \quad (8)$$

will be used as an estimate of z , where $\hat{\alpha}_z$ is the minimizer of equation (7) with respect to α , given f_z .

For comparison, we will use four other criterion functions, namely:

$$M_1(z) = \sum_{i \in \mathcal{N}_z} \{I_i - f_z(r_i(x))\}^2$$

$$M_2(z) = \sum_{i \in \mathcal{N}_z} \{I_i - \bar{I} - (f_z(r_i(x)) - \bar{f}_z)\}^2$$

$$M_3(z) = \sum_{i \in \mathcal{N}_z} \{I_i - \bar{\alpha}_z - f_z(r_i(x))\}^2$$

$$M_4(z) = \sum_{i \in \mathcal{A}_T} \{I_i - \bar{\alpha}_z - f_z(r_i(x))\}^2$$

where $\bar{\alpha}_z$ is computed for each z separately as the minimizer of

$$\sum_{i \in \mathcal{A}_x} \{I_i - \alpha - f_z(r_i(x))\}^2$$

while keeping z fixed, and where \bar{I} is the mean of the pixel values in \mathcal{N}_x and \bar{f}_z the mean of $f_z(r_i(x))$ for $i \in \mathcal{N}_x$.

The criterion function M_1 is not believed to perform very well if we have either censoring or different background levels. The last three criteria all take into account the background level. The difference between them is subtle. M_2 simply subtracts the corresponding mean from each term, without bothering about the possibility of censoring; one could say that M_2 estimates α by $\bar{I} - \bar{f}_z$. On the other hand M_3 and M_4 estimate α from the uncensored pixel values only. Also, whereas M_1 , M_2 , and M_3 all sum over all pixels in \mathcal{N}_x , M_4 only sums over the uncensored pixels. Furthermore, note that when there is no censoring, $M = M_2 = M_3 = M_4$.

For particles for which we want to estimate the depth, we first estimate the particle center by minimizing equation (3). The pixel values and distances from the estimated particle center, $\{r_i, I_i\} = \{r_i(\hat{x}), I_i\}$, are then used in template matching.

Matching example

In Figure 5 we have displayed the scatter plot from a particle from the same z -scan as sub-image -15 in Figure 2 after the center has been refined. We have also plotted the template profile f_z that gave the best fit, the template for depth $z = -15$. The logarithm of the matching criteria M , M_1 , M_3 , and M_4 are plotted in the right panel of Figure 5. To be comparable, we divided the matching criteria M_k , $k = 1, 3, 4$, by σ^2 . M is the solid line, M_1 the dotted, M_3 the dashed, and M_4 the solid-dotted. We see that all four have a dip towards $z = -15$. M_2 is not displayed since it was indistinguishable from M_3 . The difference between the criteria is clearly visible for templates corresponding to particles above the focal plane (the positive labels), since they all have low pixel values for d close to zero, and this is where censoring occurs for this particle. Compare also this result with the template profiles in Figure 4.

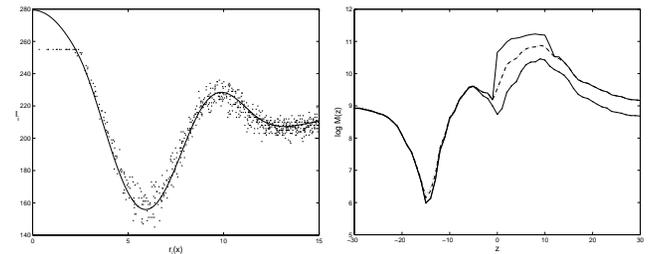


Figure 5: The scatter plot in the left panel is from a particle, presumed to be at depth -15. The matching criteria M (solid), M_1 (dotted), M_3 (dashed), and M_4 (solid-dotted) are plotted on the right and they all have a dip towards -15. M_1 is practically indistinguishable from M_3 ; the difference is only visible around the dip and for z smaller than -20.

5 Results

We conducted a small performance study of the depth estimation procedure developed in Sections 3 and 4 using real data consisting of the z -scans. In the 61 z -scans used in the construction of the templates, we have in total 10 particles believed to be at the same relative depth as the corresponding particles in the sub-images of Figure 2. All 10 particles are however not present in every z -scan, since even though our particles are adsorbed on one of the glasses of the specimen, not all other particles in the specimen are. As a consequence, the adsorbed particles are sometimes occluded by the mobile ones. For each particle, we do not use the z -scans where it is occluded. Therefore, the number of particles for each depth used varies between 8 and 10. (There are 10 particles in 43 z -scans, 9 in 14 z -scans, and 8 in 4 of the z -scans.)

The only criterion function used here, was M . Also, we used the neighborhoods \mathcal{N}_x of equation (2) with $r_{max} = 15$, that is, all pixels within distance 15 from the corresponding estimate of the particle center are taken into account.

In the left panel of Figure 6, we have plotted the difference between the estimated depth using our constructed templates, and the (believed) true depth for the depths between -30 and 30. A dot indicates that at least one of the particles had this error for that depth, therefore multiple errors at a specific depth are not visible in this plot. We also calculated the sample mean error (bias) and sample standard deviation of the errors for each z . The solid line in Figure 6

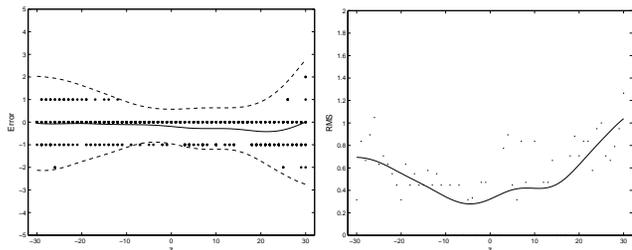


Figure 6: On the left, the dots represents the errors in depth versus the (believed) true depth for all ten particles. The solid line is the smoothed sample bias. The dashed-dotted lines are the smoothed bias plus and minus three times the smoothed sample standard deviation for each depth. The dots on the right are the pointwise root-mean-square errors for each depth and the solid line is the smoothed RMS.

is the smoothed mean of the error (bias) and the two dashed-dotted lines are the smoothed sample bias plus and minus three times the smoothed sample standard deviation, respectively. The pointwise root-mean-square (RMS) errors are plotted in the right panel of Figure 6, together with its smoothed version. All smoothing was conducted using a local linear kernel estimate with bandwidth 5. The two figures tell us that the accuracy in depth estimation is better for depths closer to the focal plane.

The RMS error calculated over all depths and particles is 0.60 z -units and the mean bias is -0.23 z -units. Even though the number of replicates are few, the investigation indicates that depth estimation is possible at least with a level of accuracy corresponding to a standard deviation of one z -unit, which corresponds to 0.2 μm . This should be compared with the accuracy of the position estimation in 2-D reported in Kvarnström and Glasbey (2005), namely the accuracy varying from 0.02 to 0.10 pixels in RMS error, which corresponds to 3.6 nm and 20 nm, respectively.

The major contributing factor to the RMS errors for this study however, seems to be the systematic offset in depth relative to the depth of the two particles used in the template construction. Furthermore, for some particles this offset seems to vary for the z -scans. For example, in Figure 7 we have plotted the errors in depth estimation as a function of depth (that is, for different z -scans) for two particles. There clearly seem to be a negative and positive trend, respectively,

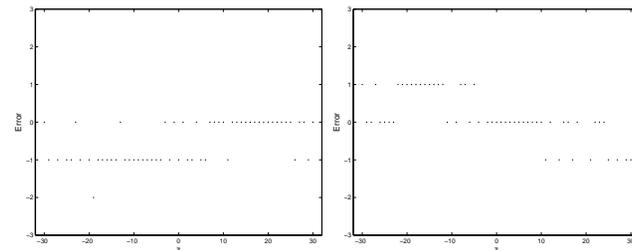


Figure 7: The errors in the depth estimates versus the (believed) true depth, for two different particles. There seems to be a trend upwards and downwards, respectively for the two particles, as the depth goes from negative to positive.

as the depth goes from negative to positive. The reason for this behavior is not known. Naturally, systematic offsets from the believed true depth makes it hard to draw conclusions on the accuracy of the depth estimation.

6 Simulation study

Since the number of particles in the z -scans is so small, and since the true depth of them did not seem to be same as the corresponding depths for the particles used in the template construction, a simulation study was conducted.

We create an artificial image of size 33 times 33 for a particle at depth z as follows. First the center for the particle is chosen at the middle pixel of the image (17,17) but with an offset chosen uniformly from $[-0.5,0.5]$ independently for each coordinate. We choose the particle center at random in order to get randomness in the distances $r_i(x)$. Then the pixel values for all pixel centers are given the value predicted by the template profile f_z for depth z plus a random background level α , chosen uniformly among the real values between -40 and 10. This variability is typically what is observed in images like the ones in Figure 1. To this we add zero mean Gaussian image noise with covariance function $\sigma^2 \exp(-c\|i - j\|)$ between pixels i and j in the image. The realization of this noise is done via a Cholesky factorization of the covariance matrix with the elements $\sigma^2 \exp(-c\|i - j\|)$. Finally, the pixel values are rounded to the nearest integer between 0 and 255.

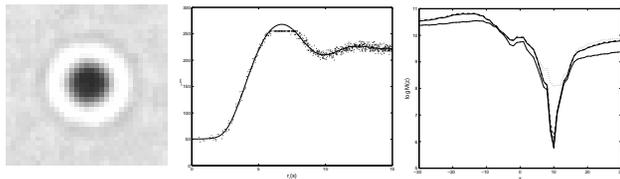


Figure 8: Example of a simulated image (left) of a particle at depth 10. The middle panel shows the scatter plot together with the template profile corresponding to the correctly estimated depth. The criterion functions M (solid), M_1 (dotted), M_3 (dashed), and M_4 (solid-dotted) are plotted in the right panel.

In Figure 8, we have, from left to right, an example of a simulated image at depth $z = 10$, its scatter plot, and the resulting matching criteria. With the scatter plot, we have also plotted the fitted template profile corresponding to the (correctly) estimated depth. The background level α used here is 15 which explains the fairly large amount of censoring. The image noise parameters σ^2 and c , were 25 and 0.6, respectively.

For each of the 35 depths in Figure 2, we simulated 1000 independent images and estimated the depth for each using the four template matching criteria of Section 4. The matching for each image was done using all 61 template profiles, that is, we use all templates from depths -30 to 30, not just the 35 depths for which we simulate images for. We did not estimate the particle centers x for the distances $r_i(x)$, but used instead the known (but random) centers. We used the same neighborhoods \mathcal{N}_x as in the previous section, that is pixels up to distance $r_{max} = 15$ from the particle center.

The overall RMS in the simulation study for all 35 depths was 0.24 z -units or 49 nm. However, the errors were practically zero for depths between 15 z -units from the focal plane. For these depths there were only 14 misclassifications. Considering that the total number of simulations for these depths were 25 000, we can conclude that the precision is very high for depths close to the focal plane. Therefore, the main contribution to the overall RMS of 0.24 z -units comes from the depths further away from 15 z -units from the focal plane. The RMS goes up almost linearly from zero at depth ± 15 , to 0.6 and 0.8, for depths -30 and +30, respectively. Looking at the estimated intensity profiles in Figure 4 or the appearance of the particles in Figure 2, the general tendency of precision is maybe not surprising. The particles are more similar looking for depths further away from the focal plane.

Compared to the other matching criteria, all but M_1 were almost identical in performance. The overall RMS errors for M_4 , M_3 , and M_2 were 0.24, 0.25 and 0.25 z -units, respectively. M_1 however, the matching criterion that did not take into account either censoring or different background levels, had an overall RMS error of 5.0 z -units. The conclusion is therefore that taking into account different background levels is far more important than taking care of the censored pixels.

7 Conclusions and discussion

As discussed in Section 5, it is hard to draw conclusions on the accuracy of the depth estimation presented here by using real data as the z -scans. We simply cannot infer whether the errors are due to our depth estimation or an effect of deviances in depth among our ten particles. The simulation study however shows that the depth estimation is highly accurate with an overall RMS of 0.25 z units, or 50 nm. However the RMS varies heavily with the true depth, with RMS errors up to 0.8 z -units for particles 6 μm from the focal plane, and practically zero for particles within 2 μm from the focal plane. This should be compared with the simulation study for the 2-D estimation in Kvarnström and Glasbey (2005), where the RMS errors varied from 0.020 to 0.10 pixels (3.6 nm to 18 nm) depending on depth.

In the simulation study, we did not estimate the particle center in the image plane, but used the known true centers. One might think that this may contribute to the optimistic result of Section 6, since errors in particle centers affect the distances $r_i(x)$ used in the template matching. However, a simulation study (not reported here) indicated that this is not the case at all. Probably a more important objection to the highly optimistic result of the simulation study, is that the image noise in real images seems to have larger variance for particles close to the focal plane, as reported in Kvarnström and Glasbey (2005). Even so, compared with the result for the real images for these depths, the conclusion is that the template matching approach works well and the precision (measured as standard error) is at least within one z -index of the z -scans, corresponding to $\pm 0.2\mu\text{m}$.

Regarding the choice of matching criterion, even though M scored best among the five in the simulation study, the three others that at least took care of the different background levels (M_4 , M_3 , and M_2), were only slightly worse than M . Therefore, the conclusion is that taking into account different background levels is far more important than taking care of the censored pixels.

additive

Our template construction was based on particles adsorbed on the specimen glass. If the appearance of particles is affected by this adsorption, our estimated profiles might differ from what particles look like in when they are diffusing freely in a solution, as is the case for the particles in Figure 1. Consequently, the depth estimation of free particles might be affected. In relation to this, it should also be mentioned that the templates can most likely only be used for depth estimation of particles of the same size. If there are particles of different sizes in the solution the template matching might be of limited use, unless we construct a set of templates for particles for a number of different sizes.

Regarding the construction of templates, an alternative method to using empirical data, would be to construct the templates mathematically, using theory from optics for the tracing of rays in the image formation. This was done by Larsen and Rudemo (1998) for creating templates for the appearances of tree canopies observed from the air under different lighting conditions. However, since the wavelength of light is in the same order of magnitude as the diameter of the particles, the two standard approximations of optics, geometric and Fourier optics (ray-tracing), cannot be used. Therefore, this approach has not been pursued. An alternative and more advanced approach to Fourier optics could be to use Mie-theory, which was used by Ovrn and Izen (2000) to predict the appearance of a polystyrene sphere of diameter $7\ \mu\text{m}$, that is 14 times bigger than the particles considered here.

References

- S. Ahn, S. Lee, A. Meyyappan, and P. Schenker. Experiments on depth from magnification and blurring. Proceedings of the 1997 IEEE/RSJ International Conference on Intelligent Robots and Systems, pages 733 – 739, 7-11 Sept 1997.
- A. Chowdhury and R. Chellappa. Robust estimation of depth and motion using stochastic approximation. 2001 International Conference on Image Processing, pages 642–645, 7-10 Oct 2001.
- J. Crocker and D. Grier. Methods of digital video microscopy for colloidal studies. *Journal of Colloid and Interface Science*, 179:298–310, 1996.
- J. Crocker and D. Grier. Interactions and dynamics in charge-stabilized colloid. *MRS Bulletin*, 23:24–31, 1998.
- K. Dralle and M. Rudemo. Automatic estimation of individual tree positions from aerial photos. *Canadian Journal of Forest Research*, 27:1728–1736, 1997.

- M. S. Elliot and W. Poon. Conventional optical microscopy of colloidal suspensions. *Advances in Colloid and Interface Science*, 92:133–194, 2001.
- D. F. Evans and H. Wennerström. *The Colloidal Domain. Where Physics, Chemistry, Biology, and Technology meet*. Wiley-VCH, New York, second edition, 1999.
- J. Fan and I. Gijbels. *Local Polynomial Modelling and Its Applications*. Chapman and Hall, London, first edition, 1996.
- P. Gil, S. Lafuente, S. Maldonado, and F. Acevedo. Distance estimation from image defocus for video surveillance systems. *Electronics Letters*, 40:1047 – 1049, 2004.
- T. Hastie and R. Tibshirani. *Generalized Additive Models*. Chapman and Hall, London, first edition, 1990.
- M. Kvarnström and C. Glasbey. Estimating particle centres and intensity profiles in microscope images. *Submitted*, 2005.
- M. Larsen and M. Rudemo. Optimizing templates for finding trees in aerial photographs. *Pattern Recognition Letters*, 19:1153–1162, 1998.
- B. Ovrn and S. Izen. Imaging of transparent spheres through a planar interface using a high-numerical-aperature optical microscope. *Journal of the Optical Society of America A. Optics, Image Science and Vision*, 17:1202–1213, 2000.
- D. Young, C. Glasbey, A. Gray, and M. N.J. Towards automatic cell identification in dic microscopy. *J. Microsc.*, 192:186–193, 1998.

Estimation of the diffusion coefficient in a mixture model

Mats Kvarnström

Mathematical Statistics
Chalmers University of Technology
412 96 Göteborg, Sweden
matskv@math.chalmers.se

Abstract

The positions of particles assumed to perform Brownian motion have been observed in a series of images. Since some of them seem to be fixed, a model with two kinds of particles, diffusing and fixed, is introduced. For each particle position observation we also assume an additive normal measurement error. We regard the problem as an incomplete data problem since we do not know *a priori* which particles are really diffusing. The complete data is of curved exponential type and the observed data is a mixture of two normal components. The maximum likelihood estimator is computed via the EM algorithm and it is shown to be strongly consistent and asymptotically normal, as the number of particles approaches infinity, under a reasonable restriction on the parameter space. A simulation study shows that the method is robust even for large measurement errors, and that the estimates are close to normal even for small sample sizes.

Key Words and Phrases: discretely observed diffusion, measurement error, mixture distribution, EM algorithm, asymptotic normality, strong consistency, curved exponential family

1 Introduction

We investigate the estimation of the diffusion variance (or equivalently, the diffusion coefficient) of colloidal particles which have been observed in a series of images recorded with a video microscope. The positions of the particles have been estimated using image processing algorithms and tools. The moving particles are assumed to perform Brownian motion in three dimensions. Furthermore, we assume the observed positions to be imperfect, i.e. we assume some measurement error. A further complicating fact is that some of the observed particles are not moving but are instead particles adsorbed on the objective or cover glass of the specimen. Also, some observed particle positions are considered to be from “false” particles, which do not correspond

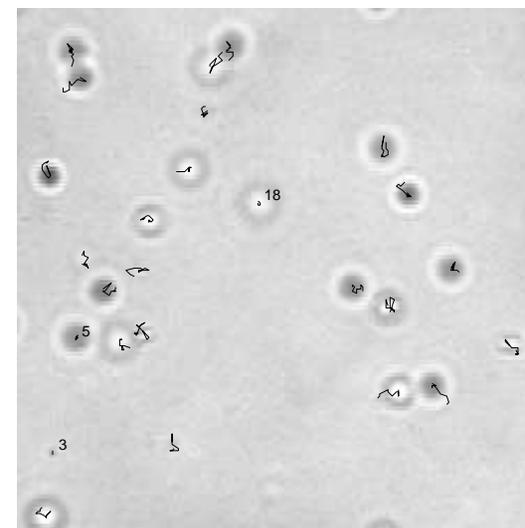


Figure 1: The 26 trajectories estimated in a sequence of 12 images together with the first image of the sequence. The three numbered particles seem to be fixed.

to real particles, but instead, defects in the optics of the microscope. These fixed particles could be removed manually but we decided not to do that, first, because this should be possible to do automatically using statistical methods, and second, because the observed positions of the fixed particles actually give us information on the measurement error. Instead, we introduce a model which is a mixture of diffusing and fixed particles, and for both types, the positions are observed under additive Gaussian measurement noise. The parameters of the model are the diffusion variance, σ^2 , the measurement error variance, σ_e^2 , and the proportion of diffusing particles, p . The problem can be considered an incomplete data problem since we do not know *a priori* which particles are diffusing.

An example of what the situation may look like, can be seen in Figure 1. The figure shows the initial image in a sequence of 12 images, together with the positions of the particles in the subsequent 11 images, together forming the estimated trajectories of the particles. Here, the positions of the particles have only been estimated in two dimensions. For details regarding the estimation of the particle positions, see KVARNSTRÖM (2002). By manual inspection, we made sure that no change of identities of the particles occurred in the process of converting the observed positions in the images into trajectories. The time interval between two images is 4 milliseconds. The particles are spherical, made of polystyrene and are all equal in size, 494 nm in diameter. The apparent differences in size and brightness are due to an out-of-focus effect and depend on placement in depth of the particles relative to the focal plane. Particles above

the focal plane are bright in the middle and dark on the circumference and vice versa for the particles below the focal plane. Also, the depicted size of a particle increases with its distance to the focal plane. In the sequence displayed Figure 1, three particles seem to be fixed. These are the particles with numbers next to them; particles 5 and 18 are adsorbed on the cover and the objective glass, respectively, and particle 3 probably corresponds to a defect in the optics.

Several papers in the mathematical statistics literature deal with the estimation of the diffusion coefficient of a diffusion. Usually however, only the case where data consist of a single observed trajectory, is considered. Furthermore, the diffusion coefficient is usually a space-dependent function which is either estimated non-parametrically or parametrically. The asymptotic properties are then studied either as the sampling interval goes to zero with a fixed total observation length (see DOHNAL (1987) and FLORENS-ZMIROU (1993) for parametric, and HOFFMANN (1999), JACOD (2000), and HOFFMANN (2001) for non-parametric estimation), or as the total observation time goes to infinity while the sampling interval is kept constant (see BIBBY and SØRENSEN (1995) and KESSLER and SØRENSEN (1999)). GENON-CATALOT and JACOD (1994) also pursues the latter approach, but with a random sampling scheme.

In the situation covered here, we have several observed particle trajectories, each with a fixed number of samples. Furthermore, the observed positions of the trajectories are subject to measurement error and not all observed particles are diffusing. Our data becomes a finite mixture of diffusing and non-diffusing particles. The asymptotic properties of the maximum likelihood estimator of the model parameters, of which the (constant) diffusion coefficient is one, is then analysed. We show that the estimator is strongly consistent and asymptotically normal, as the number of particles approaches infinity, under a reasonable restriction on the parameter space. To the author's knowledge, this is not covered elsewhere in the literature.

As an application, we use the trajectories from the video sequence in Figure 1 as data. The maximum likelihood estimator of the parameter is computed via the EM algorithm, which gives us, in addition to the parameter estimate, the posteriori estimates of whether a particle is moving or fixed, called the classification variables. The estimated diffusion coefficient is in good agreement with that predicted by Stoke-Einstein's relation. Moreover, the classification variable estimates correspond very well to our manual classification.

The paper is organized as follows. In Section 2 we introduce the model with two kinds of particles, diffusing and fixed, both observed with additive measurement error on the position estimates. Various properties of the likelihood and the maximum likelihood estimator are discussed in Section 3. We introduce a simple and reasonable restriction on the parameter space and prove that there always exists a maximum likelihood estimator under this restriction. Furthermore, we show how to implement the EM algorithm in this particular setup.

In Section 4 we study the asymptotic properties of the estimator when we keep the observation length fixed and let the number of particles go to infinity. The estimator of the triple $\theta = (\sigma^2, \sigma_e^2, p)$ using the observed data, i.e. the observed increments, is verified to be strongly consistent and asymptotically normally distributed under the previously mentioned restriction on the parameter space. The same asymptotic result is also shown to be true if instead the complete data is used.

In Section 5 we use the model assumption and estimate the diffusion variance for the data corresponding to the trajectories in Figure 1. In a simulation study, the EM algorithm is shown to be

a practical alternative when computing the maximum likelihood estimate, at least for signal-to-noise ratios σ^2/σ_e^2 above 1/2, and as long as the observation length, N , is larger than 10. Another simulation study shows that normal approximation of the distribution of θ seems reasonable for as small population sizes n as 10. Finally, in Section 6, we discuss the results and possible future developments.

2 Model assumptions and notation

Let us start by regarding one-dimensional observations of a particle and denote the true and observed position of a generic particle at time $k = 0, \dots, N$, by R_k and S_k , respectively.

For a diffusing particle we assume the following state-space model:

$$\begin{aligned} R_k &= R_{k-1} + w_k \\ S_k &= R_k + e_k \end{aligned} \tag{1}$$

for $k = 1, \dots, N$. The initial position R_0 is assumed to be a non-random constant. Furthermore $S_0 = R_0 + e_0$. Since each moving particle is assumed to perform a Brownian motion, the increments $\{w_k\}_{k=1}^N$ are zero mean i.i.d. normally distributed random variables with variance $\sigma^2 = 2D\tau$, where τ is the time interval between images and D the diffusion coefficient. The measurement errors, $\{e_k\}_{k=0}^N$ are assumed to be i.i.d. zero mean normal variables with variance σ_e^2 , independent of the increments $\{w_k\}$. For a fixed particle, we use the same state-space model but with $w_k = 0$ for all k .

Let n be the number of observed particles and let each particle $i = 1, \dots, n$ be diffusing with probability p independently of each other. Define the classification variables Z_i as

$$Z_i = \begin{cases} 1 & \text{if the } i\text{:th particle is diffusing} \\ 0 & \text{if the } i\text{:th particle is fixed} \end{cases}$$

for $i = 1, \dots, n$. We assume that a particle is either diffusing or fixed for the entire sequence of $N + 1$ observations.

The model can now easily be extended to noisy observations of a Brownian motion in d dimensions if we assume the measurement error in each dimension to be distributed as $\{e_k\}$ above and independent of each other. Then, the observations of a particle follow the state-space model (1) in each dimension independently of each other, and each particle will be assigned independent coordinate processes. Henceforth, we will however for ease of notation, write as $d = 1$ unless otherwise stated.

The index $i = 1, \dots, n$, is used to distinguish between the n particles. By a subindex i to an entity, as in Z_i , we mean that the entity belongs to the i :th particle. If the subindex i is neglected we mean a generic particle. The index $k = 0, \dots, N$, is used for a generic particle only, and corresponds to the discrete time k in the state-space model.

We denote the observed increments for a particle by $Y_k = S_k - S_{k-1}$, $k = 1, \dots, N$. The covariance matrix of the increment vector, $Y = [Y_1, \dots, Y_N]^T$, becomes

$$\Sigma_1 = \sigma^2 I + \sigma_e^2 T \tag{2}$$

for a diffusing particle and

$$\Sigma_0 = \sigma_e^2 T$$

for a fixed particle, where I is the $N \times N$ identity matrix and T is the tri-diagonal matrix

$$T = \begin{bmatrix} 2 & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & \cdots & 0 \\ 0 & -1 & 2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 2 \end{bmatrix}.$$

We see from the covariance matrix above that the measurement error on the observed positions induces a dependence between the observed increments, which originally, by definition of Brownian motion, were independent.

2.1 Transformation of the increment vector

To make our formulas look nicer in the subsequent sections, we use some basic linear algebra to transform the increment vector so that the elements of the transformed vectors become uncorrelated.

In (2), Σ_1 has the same eigenvectors as T since every vector is an eigenvector to I . If we denote the eigenvalues of T by λ_k , $k = 1, \dots, N$, then the eigenvalues of Σ_1 are

$$\gamma_k = \sigma^2 + \sigma_e^2 \lambda_k, \quad k = 1, \dots, N.$$

Let U be the matrix with the eigenvectors of T as columns. Then we can write, by the spectral decomposition theorem, $T = U \Lambda U^T$, where $\Lambda = \text{diag}\{\lambda_1, \dots, \lambda_N\}$. If

$$\tilde{Y} = U^T Y \quad (3)$$

is the transformed increment vector, its covariance matrix will be diagonal:

$$\begin{aligned} \text{Var}\{\tilde{Y}\} &= U^T \text{Var}\{Y\} U = U^T (\sigma^2 I + \sigma_e^2 U \Lambda U^T) U = \\ &= \sigma^2 I + \sigma_e^2 \Lambda = \text{diag}\{\gamma_1, \dots, \gamma_N\} \end{aligned} \quad (4)$$

The analogous is valid for a fixed particle, but with $\sigma^2 = 0$. The dependence between the increments is now "hidden" in U and Λ , which do not depend on σ^2 or σ_e^2 , but only on the length of the increment vector N , which of course is known.

2.2 Observed and complete data

We classify data into two categories, observed and unobserved. The observed data consist of the noise-corrupted increment vectors Y_i , $i = 1, \dots, n$, while the classification variables Z_i are unobserved. Together, they constitute the complete data, denoted by $X_i = (Y_i, Z_i)$, $i = 1, \dots, n$.

The probability density function of the complete data X is

$$g_C(x; \sigma^2, \sigma_e^2, p) = [p f_1(y; \sigma^2, \sigma_e^2)]^z [(1-p) f_0(y; \sigma_e^2)]^{1-z} \quad (5)$$

for a single generic particle, where f_1 and f_0 are the pdfs of a zero mean N -variate normal distributed random vector with covariance matrices $\Sigma_1 = \Sigma_1(\sigma^2, \sigma_e^2)$ and $\Sigma_0 = \Sigma_0(\sigma_e^2)$, respectively.

In the d dimensional case, f_i will be a dN -variate normal density with d independent particles one for each dimension, since, by assumption, the coordinate processes of a particle are independent.

The complete data belongs to an exponential family of distributions (see for example LINDSEY (1996)). However, if $N > 1$, which typically is the case, the distribution is non-regular, curved, since the parameter space is 3-dimensional and the dimension of the minimal sufficient statistic is $N+2$ (see the Appendix for a derivation of this). This holds irrespectively of how many dimensions we observe. The case $N = 1$ is non-typical since we think of our problem as studying a video sequence of images of particles and as such we usually observe more than one increment.

The probability density of the observed data for a generic particle, Y , is obtained by integrating (5) over the distribution of Z

$$g(y; \sigma^2, \sigma_e^2, p) = p f_1(y; \sigma^2, \sigma_e^2) + (1-p) f_0(y; \sigma_e^2), \quad (6)$$

and we see that our observed data is a finite mixture of two normal components. Note however that one of the parameters, σ_e^2 , is present in both components, which is typically not the case for finite mixtures. For a thorough account on finite mixture models and their applications, we refer to MCLACHLAN and PEEL (2000).

3 Likelihood Estimation

We denote the parameter vector by $\theta = (\sigma^2, \sigma_e^2, p)$. Let Ω be the parameter space consisting of those θ defining valid finite mixture densities (6). In other words, $\Omega = \{\theta = (\sigma^2, \sigma_e^2, p) \in \mathbb{R}^3 \mid p \in [0, 1], \sigma^2 > 0, \sigma_e^2 > 0\}$. The true parameter point θ_0 is always assumed to lie in the interior of Ω , i.e. $\theta_0 \in \text{int}(\Omega)$.

The complete likelihood L_C induced by the complete data (observed increments and classification variables) from n observed particles is

$$L_C(\theta) = \prod_{i=1}^n [p f_1(y_i; \sigma^2, \sigma_e^2)]^{z_i} [(1-p) f_0(y_i; \sigma_e^2)]^{1-z_i} \quad (7)$$

However, since our observed data consist of the increments only, the observed likelihood becomes

$$L(\theta) = \prod_{i=1}^n \{p f_1(y_i; \sigma^2, \sigma_e^2) + (1-p) f_0(y_i; \sigma_e^2)\} \quad (8)$$

Often with finite mixtures, there is a problem of identifiability, i.e. that a permutation of the parameters in the model yields the same distribution, see DAY (1969), SUNDBERG (1974), and REDNER and WALKER (1984). In our model, as long as the true parameter θ_0 lies in the interior of Ω , we do not have this problem since the two distributions in the mixture are not interchangeable due to the presence of σ_e^2 in both mixture components.

3.1 Existence of a maximum likelihood estimator

First, we should address the important question of the existence of a global maximizer of L for a given set of observations $\{Y_1, \dots, Y_n\}$. DAY (1969) pointed out that for univariate normal mixtures, the likelihood is not bounded if both the expected values and the variances of the mixture components are considered as unknown parameters. Hence, with no extra conditions, a global maximizer does not generally need to exist for normal mixtures. Even though our present setup of the model does not involve the estimation of a drift term in the diffusion, or systematic position measurement errors, we will propose a restriction on the parameter space that will guarantee the existence of a maximum likelihood estimator even under the natural extension of the model to non-zero drift components of the diffusions.

For $c \in (0, 1)$, the restricted parameter space Ω_c is defined to be a subset of Ω such that

$$0 < c \leq \frac{\sigma^2}{\sigma_e^2} \leq c^{-1} < \infty \tag{9}$$

holds. We will assume that this holds for some $c \in (0, 1)$, i.e. $\theta_0 \in \Omega_c$ for some $c \in (0, 1)$. The restriction is reasonable in the sense that we do not allow the signal-to-noise ratio to be too small, neither too big.

This kind of restriction on the parameter space, was originally proposed by HATHAWAY (1985). Here, it will guarantee, that for any $c \in (0, 1)$, the likelihood function will have a global maximizer for each n . In Section 4, the restriction will be used in a sufficient condition for strong consistency as n approaches infinity (see Theorem 1), where the condition is such that we assume that $\theta_0 \in \Omega_c$ holds for some $c \in (0, 1)$.

Lemma 1. *Let $\{Y_1, \dots, Y_n\}$ be a set of observations from the finite mixture specified by the density (6) and let $c \in (0, 1)$. Then, with probability one, there exists a global constrained maximizer for $L(\theta)$ in Ω_c .*

Proof. The idea is to show that

$$\sup_{\theta \in \Omega_c} L(\theta) = \sup_{\theta \in K} L(\theta)$$

for some appropriate, compact $K \subset \Omega$.

With probability one, the increment vectors will all be different from the zero vector. Therefore all the terms in the likelihood will stay bounded. Also, it will go to zero if both σ^2 and σ_e^2 either go to zero or to infinity. By condition (9) above however, it is enough to show that one of the two variances goes to zero or infinity; the other variance will also be forced to go to zero or infinity, respectively.

So, there exist constants a_i and b_i such that $K = \{\theta \in \Omega_c : a_1 \leq \sigma_e^2 \leq a_2, b_1 \leq \sigma^2 \leq b_2\}$ which gives the desired result.

Remark 1: A maximum hence exists, but it does not necessarily have to be unique for fixed n . For $p = 0$, σ^2 becomes a free variable. Also, if $p = 1$ and $N = 1$, all values of σ^2 and σ_e^2 satisfying $\sigma^2 + 2\sigma_e^2 = c$ for some constant c , are maximum likelihood estimators. Note however that the restriction is not a necessary condition for a maximizer to exist, neither is it a sufficient condition for a *unique* global maximizer to exist. It is a sufficient condition for the likelihood to be bounded for all $\theta \in \Omega_c$, and hence for a maximum likelihood estimator to exist.

Remark 2: If the number of observations n is larger than three (one more than the number of mixture components), the restriction (9) of the parameter space also gives us the conclusion of Lemma 1 even under an expanded model with a drift term in the diffusion together with systematic position measurement errors, that is, if the mixture components have non-zero expected values, μ_1 and μ_2 , and we need to estimate these as well; just let K be as before concerning the parameters σ^2 and σ_e^2 , and with $|\mu_1|, |\mu_2| \leq \max_i \{\|Y_i\|\} < \infty$ (see HATHAWAY (1985)) where $\|\cdot\|$ is the L_2 -norm of a vector. As mentioned previously, this might turn out to be useful if it is necessary to determine whether systematic effects are present or not.

3.2 The EM algorithm

When computing maximum likelihood estimates for finite mixtures, the EM algorithm is an appealing method to use. In fact, although the algorithm got its name and was generalized by DEMPSTER et al. (1977), it was actually introduced and used, for the special case of finite mixtures of exponential distributions as early as in HASSELBLAD (1969). For mixture distribution the method takes full advantage of the simple structure of the complete likelihood together with easily calculated estimates of the unobserved data, the classification variables. For an overview of the theory and contemporary applications of the EM algorithm, we refer to MCLACHLAN and KRISHNAN (1997).

The algorithm is an iterative scheme consisting of two steps; the expectation and the maximization step, accordingly called the E-step and the M-step. The E-step consists of estimating the unobserved data, which in the case of a mixture are the classification variables. In the M-step we maximize the complete likelihood (7) using the estimated classification variables, \hat{Z}_i , from the E-step together with our data Y_i . The algorithm hence becomes:

Assume that $\theta^{(m)}$ is the estimate from the m :th iteration of the EM algorithm.

- E-step: For each $i = 1, \dots, n$, compute

$$\hat{Z}_i = \mathbf{E}_{\theta^{(m)}}\{Z_i|Y_i\} = \frac{p^{(m)} f_1(Y_i; \Sigma_1^{(m)})}{p^{(m)} f_1(Y_i; \Sigma_1^{(m)}) + (1-p^{(m)}) f_0(Y_i; \Sigma_0^{(m)})}$$

- M-step: Maximize

$$\mathbf{E}_{\theta^{(m)}}\{\log L_c(\theta)|Y\} = \sum_{i=1}^n \left\{ \hat{Z}_i \log\{p f_1(Y_i; \sigma^2, \sigma_e^2)\} + (1 - \hat{Z}_i) \log\{(1-p) f_0(Y_i; \sigma_e^2)\} \right\}$$

with respect to $\theta = (\sigma^2, \sigma_e^2, p)$, denote the result by $\theta^{(m+1)}$ and go to the E-step.

In this application of the EM algorithm, each of the two steps has a probabilistic meaning; in the E-step we classify each particle using a quadratic discriminant rule, and in the M-step we maximize the complete likelihood using the classifications from the E-step as if classifications from the previous E-step in fact were the observed data. Note, however, that the estimated classification variables are not confined to zero or one, but could be any real number in-between.

As is the case for most numerical algorithms for maximizing a function, there is no way of guaranteeing that we actually end up in the global maximum. The EM algorithm can guarantee however, that

$$L(\theta^{(m+1)}) \geq L(\theta^{(m)})$$

and since in our case the likelihood is bounded and continuous, every sequence of iterates $\{\theta^{(m)}\}$ of the EM algorithm will have a stationary value of $L(\theta)$ as a limit point (Theorem 2 of WU (1983)). This stationary value does of course not have to be the global maximum. Nothing more than this, can in general be said about the $\theta^{(m)}$ -sequence. For a thorough discussion regarding the convergence of the EM algorithm, we refer to the already mentioned work by WU (1983), which is covered also in the review article by REDNER and WALKER (1984), and in the general reference on the EM algorithm, MCLACHLAN and KRISHNAN (1997).

For the typical application of the mixture model discussed in this article, with N usually larger than 10 and a ratio σ^2/σ_e^2 larger than 1/2, simulations show that the EM algorithm works sufficiently well, both concerning the number of steps of the iteration until convergence to a stationary point of the likelihood function, and with respect to its ability to converge to the same estimate θ regardless of starting value θ^0 ; see the simulation study in Section 5.

4 Asymptotics

In this section we study the asymptotic properties of the maximum likelihood estimator as the number of particles n grows large. It turns out that the estimator is both strongly consistent and asymptotically normal. This case is not covered in the literature probably since one of the parameters is present in both mixture components. In addition, many authors compactify the parameter space in order to get asymptotic results, see e.g. CHENG and LIU (2001). We do not want to do this, because we feel that it imposes an unnatural restriction on the parameter space.

For $N = 1$, SUNDBERG (1974) gives the consistency and asymptotic normality of the maximum likelihood estimator $\hat{\theta}_n$, under the single condition that the information matrix $\mathcal{I}(\theta)$ is positive definite at the true parameter point θ_0 . Since Lemma 2 below says that this is true for all $\theta_0 \in \text{int}(\Omega)$, we are actually done for $N = 1$, both with consistency and asymptotic normality.

Hence, what is left to prove is consistency and asymptotic normality for $N > 1$. REDNER and WALKER (1984) discuss this issue for mixtures of distributions but only for regular exponential distributions and when the parameters of separate mixture components are mutually independent. Some modifications are thus needed.

4.1 Consistency

To prove consistency of the maximum likelihood estimator for general N , we verify that the classical conditions for consistency of WALD (1949) are satisfied for the mixture density in (6) when the true parameter is in Ω_c . In the process, we use results from REDNER (1981).

Theorem 1 (Strong consistency). *Let the true parameter point θ_0 be in Ω_c for some $c \in (0, 1)$ and let $\hat{\theta}_n$ be the global maximizer of $L(\theta)$ over Ω_c for each n . Then*

$$\mathbf{P}\{\hat{\theta}_n \rightarrow \theta_0 \text{ as } n \rightarrow \infty\} = 1$$

Proof. Wald's conditions are enumerated as in REDNER (1981) to 1 through 6. For the specification of these conditions, we refer the reader to that article.

Conditions 1,2,4' and 5 are satisfied for Ω and the mixture component densities f_1 and f_2 . Inspection of the proof of Redner's Theorem 5 shows that Conditions 2 and 4 also are satisfied for the mixture density (6). If we restrict Ω to Ω_c as defined in (9), then also Conditions 3 and 6 are satisfied, giving us the result by applying Theorems 1 and 2 from WALD (1949). \square

Remark 1: The extra condition (9) is used, first to prove that a maximum likelihood estimator exists for all n , and second, to prove that Condition 3 of REDNER (1981) is satisfied, i.e. that $L(\theta_i) \rightarrow 0$ when $d(\theta_0, \theta_i) \rightarrow \infty$, where d means Euclidean distance.

Remark 2: Under the restricted parameter space, Ω_c , Wald's conditions 1-6 are satisfied even under the previously mentioned expanded model with a drift term in the diffusion and systematic position measurement errors, that is, if the mixture components have non-zero expected values and we need to estimate them as well.

4.2 Asymptotic normality

Sufficient conditions for the asymptotic normality of the maximum likelihood estimator $\hat{\theta}_n$ can be found in for example Theorem 5.23 of VAN DER VAART (1999). Since we have consistency and $\log g(y; \theta)$ is smooth, what remains to be proved is that the mapping $\theta \mapsto \mathbf{E}_{\theta_0} \log g(Y; \theta)$ admits a second order Taylor expansion around $\theta_0 \in \text{int}(\Omega)$ with non-singular second derivative matrix. In other words, we have to prove that the information matrix $\mathcal{I}(\theta_0)$ is positive definite. This is the result of Lemma 2 below, so Theorem 2 is a consequence of that.

Theorem 2 (Asymptotic normality). *Assume the true parameter $\theta_0 \in \text{int}(\Omega_c)$ for some $c \in (0, 1)$. Then the maximum likelihood estimator $\hat{\theta}_n$ is asymptotically normal, i.e.*

$$n^{1/2}(\hat{\theta}_n - \theta_0) \xrightarrow{\mathcal{D}} N(0, \mathcal{I}(\theta_0)^{-1}) \tag{10}$$

as $n \rightarrow \infty$.

The restriction to Ω_c is needed only to guarantee consistency; the positive definiteness of $\mathcal{I}(\theta)$ is in fact valid for all $\theta \in \text{int}(\Omega)$ as can be seen from the next lemma.

Below, we denote the gradient of a function f with respect to θ by $\partial_\theta f$. We also refrain from writing down the arguments of the functions when no risk of ambiguity exists.

Lemma 2. *The information matrix $\mathcal{I}(\theta)$ is positive definite for all $\theta \in \text{int}(\Omega)$.*

Proof. Positive definiteness of $\mathcal{I}(\theta)$ means that $a^T \mathcal{I}(\theta) a > 0$, for all $a \in \mathbb{R}^3 \setminus \{0\}$. Since $\mathcal{I}(\theta)$ is the variance of the score function $\partial_\theta \log g(Y; \theta)$, we interpret $a^T \mathcal{I}(\theta) a$ as the variance of the linear combination $a^T \partial_\theta \log g(Y; \theta)$. Hence, what we have to prove is that

$$\text{Var}\{a^T \partial_\theta \log g(Y; \theta)\} > 0$$

for all $a \in \mathbb{R}^3 \setminus \{0\}$.

Assume the opposite. Then we have, with probability one, that

$$a^T \partial_\theta \log g(Y; \theta) = 0 \tag{11}$$

for some $a \in \mathbb{R}^3 \setminus \{0\}$ since the mean of the score is zero. Writing out the components of the score function $\partial_\theta \log g(Y; \theta)$, we have

$$\begin{aligned} \partial_{\sigma^2} \log g &= \frac{p \partial_{\sigma^2} f_1}{p f_1 + (1-p) f_0} \\ \partial_{\sigma_e^2} \log g &= \frac{p \partial_{\sigma_e^2} f_1 + (1-p) \partial_{\sigma_e^2} f_0}{p f_1 + (1-p) f_0} \\ \partial_p \log g &= \frac{f_1 - f_0}{p f_1 + (1-p) f_0} \end{aligned}$$

where

$$\begin{aligned} \partial_{\sigma^2} f_1 &= \frac{1}{2} \sum_{k=1}^N \left(\frac{\tilde{y}_k^2}{(\sigma^2 + \lambda_k \sigma_e^2)^2} - \frac{1}{\sigma^2 + \lambda_k \sigma_e^2} \right) f_1(y; \sigma^2, \sigma_e^2) = k_1(y) f_1(y; \sigma^2, \sigma_e^2) \\ \partial_{\sigma_e^2} f_1 &= \frac{1}{2} \sum_{k=1}^N \left(\frac{\lambda_k \tilde{y}_k^2}{(\sigma^2 + \lambda_k \sigma_e^2)^2} - \frac{\lambda_k}{\sigma^2 + \lambda_k \sigma_e^2} \right) f_1(y; \sigma^2, \sigma_e^2) = k_2(y) f_1(y; \sigma^2, \sigma_e^2) \\ \partial_{\sigma_e^2} f_0 &= \left(\frac{1}{2(\sigma_e^2)^2} \sum_{k=1}^N \frac{\tilde{y}_k^2}{\lambda_k} - \frac{N}{2} \frac{1}{\sigma_e^2} \right) f_0(y; \sigma_e^2) = k_3(y) f_0(y; \sigma_e^2) \end{aligned}$$

Equation (11) can be written as

$$a_1 p \partial_{\sigma^2} f_1 + a_2 [p \partial_{\sigma_e^2} f_1 + (1-p) \partial_{\sigma_e^2} f_0] + a_3 [f_1 - f_0] = 0.$$

After re-arranging in terms of f_1 and f_0 and noticing that $f_1(Y) \neq f_0(Y) > 0$ with probability one for all $\theta \in \text{int}(\Omega)$, we see that this is equivalent to

$$\begin{cases} a_1 p k_1(Y) + a_2 p k_2(Y) + a_3 = 0 \\ a_2 (1-p) k_3(Y) - a_3 = 0 \end{cases} \tag{12}$$

For $N > 1$, since $k_1(Y)$, $k_2(Y)$, and $k_3(Y)$ are linearly independent and non-zero with probability one for all $\theta \in \text{int}(\Omega)$, equation (12), and hence equation (11), is satisfied only if a is zero, and we have a contradiction. For $N = 1$, even though $k_2 = \lambda_1 k_1$, we still have that $k_1(Y)$ and $k_3(Y)$ are linearly independent, which suffices to arrive at the same conclusion. \square

Remark: Notice that (12) is satisfied for non-zero a if $p = 0$. This is also what we would expect since then we have no information on σ^2 . Also, if $N = 1$, then $k_2(Y) = \lambda_1 k_1(Y)$, so if $p = 1$ (12) is satisfied as long as $a_1 + \lambda_1 a_2 = 0$ and $a_3 = 0$.

4.3 Complete data asymptotics

In applications the particles may have already been labelled manually as diffusing or fixed, i.e. we have been given the complete data. If this is the case, the asymptotic properties of the estimator become easier to verify, mainly because the likelihood is a product of probability density functions.

After some elementary, but fairly cumbersome, calculations we obtain

$$\mathcal{I}_c(\theta) = \begin{bmatrix} \frac{dp}{2} \sum_{k=1}^N \frac{1}{(\sigma^2 + \lambda_k \sigma_e^2)^2} & \frac{dp}{2} \sum_{k=1}^N \frac{\lambda_k}{(\sigma^2 + \lambda_k \sigma_e^2)^2} & 0 \\ \frac{dp}{2} \sum_{k=1}^N \frac{\lambda_k}{(\sigma^2 + \lambda_k \sigma_e^2)^2} & \frac{dp}{2} \sum_{k=1}^N \frac{\lambda_k^2}{(\sigma^2 + \lambda_k \sigma_e^2)^2} + \frac{Nd(1-p)}{2(\sigma_e^2)^2} & 0 \\ 0 & 0 & \frac{1}{p(1-p)} \end{bmatrix} \tag{13}$$

for the expected information matrix to the complete data. It can be seen, by applying the Cauchy-Schwarz inequality on the upper-left 2 by 2 matrix, to be positive definite for all $\theta \in \text{int}(\Omega)$.

For $\theta_0 \in \text{int}(\Omega_c)$ we get strong consistency directly from WALD (1949) even without the restriction on the parameter space. Furthermore, since $\mathcal{I}_c(\theta_0)$ is positive definite, all conditions for asymptotic normality are satisfied.

4.4 Note on a further generalization

An interesting article with relevance to our problem, is KIEFER and WOLFOWITZ (1956). It deals with the consistency of a maximum likelihood estimator when there are infinitely many incidental parameters present. These incidental parameters could be, in a generalization of our problem, the variance of the Brownian motion σ^2 if all diffusing particles have different diffusion coefficients. This corresponds to a so-called poly-disperse solution in contrast to our present problem, which is mono-disperse (every particle has the same diffusion coefficient).

Assume that for each $i = 1, \dots, n$, we have that Y_i is N -variate normally distributed random variable with mean zero and covariance matrix $\Sigma_i = I \sigma_i^2 + T \sigma_e^2$. Then, following the language of KIEFER and WOLFOWITZ (1956), the σ_i^2 :s are the incidental parameters and σ_e^2 the parameter (even though, in our context, these names are misleading since we consider it to be the other way round). Notice that if the σ_i^2 :s are constants and different for each i we only observe one increment vector Y_i for each σ_i^2 . Obviously the estimates of the σ_i^2 :s can not be consistent. It turns out, however, that if we consider σ_i^2 , $i = 1, \dots, n$, to be independent random variables with common (but unknown) distribution function F , then, under certain assumptions on F , the maximum likelihood estimator of F converges to F at every point of continuity. Also, the maximum likelihood estimator of σ_e^2 is strongly consistent.

The model discussed in this article is of course a special case of these σ_i^2 coming from a unknown distribution function F . This distribution has only two values; zero, for the fixed

particles, and σ^2 , for the diffusing particles. In other words, F can be written as

$$F(x) = \begin{cases} 0 & \text{when } x < 0 \\ 1 - p & \text{when } 0 \leq x < \sigma^2 \\ 1 & \text{when } \sigma^2 \leq x \end{cases}$$

5 Application

As an application of the model, we estimated the diffusion coefficient of the diffusing particles in Figure 1. Here, the positions of the 26 particles were estimated in two dimensions in each image using a simple circle detection algorithm and the resulting trajectories are displayed in Figure 1. The total number of frames were 12, so $N = 11$. By manual inspection, we concluded that the three particles in Figure 1 numbered 3, 5, and 18, were fixed. The remaining 23 were classified as diffusing particles.

5.1 Results

We applied the EM algorithm to the observed data with initial value $\theta^0 = (1, 1, 0.5)$. We stopped when the change of the Z_i 's between two consecutive E-steps was smaller than 10^{-6} . This criterion was satisfied after 3 steps with the resulting maximum likelihood estimates $\hat{\theta}$ with elements

$$\begin{aligned} \hat{\sigma}^2 &= 2.2058 \\ \hat{\sigma}_e^2 &= 0.3172 \\ \hat{p} &= 0.8847 \end{aligned} \quad (14)$$

where the unit for the first two is the square of the side length of a pixel. The estimated classification variables \hat{Z}_i , defined as the conditional expectation of Z_i given Y_i at parameter point $\hat{\theta}$, were

$$\begin{aligned} \hat{Z}_3 &= 2.473 \cdot 10^{-3} \\ \hat{Z}_5 &= 1.528 \cdot 10^{-5} \\ \hat{Z}_{18} &= 1.049 \cdot 10^{-5} \\ \hat{Z}_i &= 1.000 \quad \text{otherwise} \end{aligned} \quad (15)$$

in good correspondence with our manual classification.

5.2 Observed information matrix

The observed information at the maximum likelihood estimate $\hat{\theta}$, can be written as

$$I(\hat{\theta}; y) = \mathcal{I}_C(\hat{\theta}; y) - [\mathbf{E}_\theta\{S_C(X; \theta)S_C^T(X; \theta)|y\}]_{\theta=\hat{\theta}} \quad (16)$$

by using a result from LOUIS (1982), where $\mathcal{I}_C(\theta; y) = \mathbf{E}_\theta\{I_C(\theta; x)|Y = y\}$, the conditional expectation of the complete data given $Y = y$, and $S_C(x; \theta) = \partial_\theta \log L_C(\theta)$, the score of the

complete likelihood. Intuitively, the first term in (16) corresponds to the observed information from the complete data given the estimated values of the unobserved data Z_i , and the second term corresponds to the missing information due to the fact that Z_i was in fact not observed.

Using this, the observed information matrix at the estimate becomes

$$\begin{aligned} I(\hat{\sigma}^2, \hat{\sigma}_e^2, \hat{p}; Y) &= \begin{bmatrix} 33.75 & 52.75 & 0 \\ 52.75 & 476.6 & 0 \\ 0 & 0 & 254.9 \end{bmatrix} - \begin{bmatrix} 0.034 & 0.153 & -0.090 \\ 0.153 & 0.691 & -0.405 \\ -0.090 & -0.405 & 0.240 \end{bmatrix} \\ &= \begin{bmatrix} 33.72 & 52.59 & 0.090 \\ 52.59 & 475.9 & 0.405 \\ 0.090 & 0.405 & 254.7 \end{bmatrix} \end{aligned} \quad (17)$$

with inverse

$$I^{-1}(\hat{\sigma}^2, \hat{\sigma}_e^2, \hat{p}; Y) = \begin{bmatrix} 0.0358 & -0.0040 & 0.0000 \\ -0.0040 & 0.0025 & 0.0000 \\ 0.0000 & 0.0000 & 0.0039 \end{bmatrix}. \quad (18)$$

This gives us an approximate variance of the estimate of $\hat{\sigma}^2$ equal to

$$\mathbf{Var}\{\hat{\sigma}^2\} \simeq 0.0358. \quad (19)$$

Note that some elements in the matrix in (17) are close to zero. The reason for this is that the estimated classification variables in (15) are very close to their true values of zero or one. In other words, we are very close to our manual classification, which corresponds to having complete data. Compare this with the zero elements of the complete information matrix in (13).

5.3 Comparing the estimate with the theoretical diffusion coefficient

The estimated $\hat{\sigma}^2$ above corresponds to an estimated diffusion coefficient of

$$\hat{D} = 0.893 \mu\text{m}^2/\text{s},$$

where we used the relationship between diffusion variance and diffusion coefficient, $\sigma^2 = 2D\tau$ and scaling to μm . Here, $\tau=0.040$ s is the time interval between observations, and each pixel corresponds to a square with side $M=180$ nm.

If we assume normality, motivated by the asymptotic normality result of Section 4.2 when n is sufficiently large, we get an approximate 95% confidence interval of D of

$$D = \hat{D} \pm 1.96 \cdot \frac{M^2}{2\tau} \sqrt{0.0358} = .893 \pm .150 \mu\text{m}^2/\text{s}. \quad (20)$$

The theoretical diffusion coefficient is given by Stoke-Einstein's relation (see for example EVANS and WENNERSTRÖM (1999) pages 370-372)

$$D = \frac{k_B T}{6\pi\eta R_H} \quad (21)$$

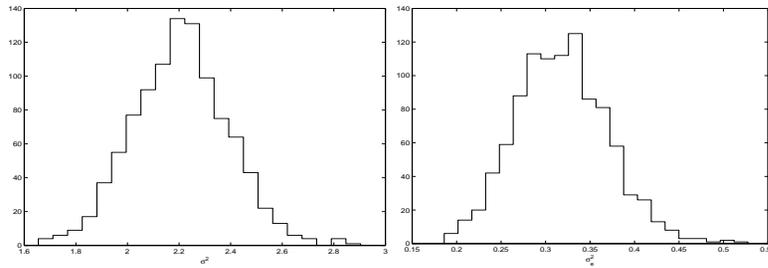


Figure 2: The histograms of the estimates of σ^2 and σ_e^2 using the EM algorithm from 1000 simulations using 2.2058 and 0.3172 as true values.

where k_B is Boltzmann's constant, η the viscosity of the solution, T the temperature and R_H the hydrological radius of the particle. The appropriate values for the viscosity and temperature are $\eta=0.9$ mPa and $T=298$ K. The geometric radius of the particles are 247 nm and this is used as the hydrological radius, even if the latter is often a bit larger than the former. Plugging all this into (21), we get

$$D = 0.982 \mu\text{m}^2/\text{s},$$

which is within the 95% confidence interval of (20).

5.4 Simulation of the approximate distribution of the estimates

We simulated 1000 realizations with 26 particles, of which 3 were fixed, over 12 frames in two dimensions, with the estimated values of $\sigma^2 = 2.2058$ and $\sigma_e^2 = 0.3172$ from (14) as the true diffusion variance and error variance. For each time series, we estimated σ^2 , σ_e^2 and p using the EM algorithm. However, since the number of fixed particles is 3 in each sequence, the estimate of p is not very interesting and will therefore not be displayed.

The histograms of the estimated values are displayed in Figure 2. The sample mean and covariance matrix of the 1000 estimates of σ^2 and σ_e^2 were

$$\begin{aligned} \bar{\sigma}^2 &= 2.2054 \\ \bar{\sigma}_e^2 &= 0.3185 \end{aligned}$$

and

$$\begin{bmatrix} .0348 & -.0040 \\ -.0040 & .0027 \end{bmatrix}$$

in good agreement with the true values of $\sigma^2 = 2.2058$ and $\sigma_e^2 = 0.3172$ and the upper-left block of the inverse of the observed information matrix in (18).

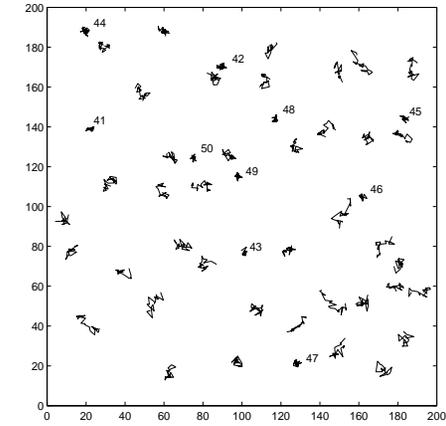


Figure 3: Simulated sequence of 50 particles of which 10 are fixed (the numbered ones) and 40 are diffusing, where $\sigma^2 = \sigma_e^2 = 1$ and $N = 20$, with the start positions taken uniformly on the square. Compare with the corresponding classification variables in Figure 4.

5.5 More simulations

In the example of Figure 1, it is easy to manually classify particles as either diffusing or fixed since the measurement error is small compared to the diffusion variance. To see what the situation looks like when the signal-to-noise ratio is significantly lower, we simulated 40 diffusing particles with $\sigma^2 = 1$ together with 10 fixed particles, all observed under measurement error with $\sigma_e^2 = 1$. The number of observed increments, N , was 20.

The resulting observed trajectories are displayed in Figure 3, where the starting positions have been chosen uniformly on the square. Clearly, here it is much harder to determine by eye, which particles are actually fixed, compared to the situation in Figure 1. However, using the analysis developed here, we are able to classify all the particles correctly, by which we mean that the estimated classification variables are within 0.5 of their true value. The estimated classification variables for this sequence are displayed in Figure 4 where the first 40 correspond to diffusing particles and the last 10 to fixed. The EM algorithm converged in 7 or 8 iterations. Also, it converged to the same estimate of θ as we varied the initial parameter θ^0 over 20 different values.

To conduct a more thorough analysis of the behaviour of the EM algorithm for this model, we simulated 100 independent data sets, each consisting of 100 particles of which 20 were fixed for different values of σ^2 , σ_e^2 , and N . Table 1 shows the result. As can be seen, the method manages to estimate the variances σ^2 and σ_e^2 appropriately, and that it is robust for all signal-to-noise ratios when the observation length is large. However, for small observation lengths, both the number of misclassifications and the number of EM iterations, become large if the signal-to-

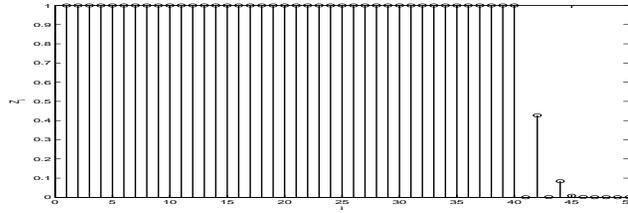


Figure 4: The estimated classification variables for the sequence in Figure 3. The first 40 correspond to diffusing particles and the last 10 to fixed particles.

noise ratio is not sufficiently high. Although not presented in Table 1, it should be mentioned that the only effect different starting values θ^0 had on the convergence of the EM algorithm, was to increase or decrease the number of iterations needed until convergence, by one or two steps.

Taking all of this into consideration, the EM algorithm is clearly an applicable method for these kind of data sets, failing only to converge within 100 iterations when $N = 10$ and the signal-to-noise ratio is $1/3$.

In Section 4 we proved that the estimator $\hat{\theta}_n$ is asymptotically normal as n goes to infinity. In applications, one might wonder for how small n it is reasonable to approximate the distribution of θ_n by the normal distribution. Looking at the histograms of Figure 2, it seems that, at least for the estimated parameter values from the example trajectories of Figure 1, this is the case for the relatively small population size of $n = 26$.

To investigate this further, we simulated 1000 independent sequences where $N = 10$, each with $n = 10$ particles, of which 2 were fixed, and where the true parameters were $\sigma^2 = 2$ and $\sigma_e^2 = 1$. For each sequence, we estimated σ^2 , σ_e^2 , and p but, as before, since the number of fixed particles was non-random, we do not include the estimate of p in the analysis. We calculated the generalized distance

$$d_m^2 = (\theta_m - \bar{\theta})S^{-1}(\theta_m - \bar{\theta})^T, \quad m = 1, \dots, 1000$$

where θ_m is the vector consisting of the final estimates of σ^2 and σ_e^2 from the m :th simulated sequence, and where $\bar{\theta}$ and S are the sample mean and the sample covariance matrix, respectively, of the 1000 estimates of σ^2 and σ_e^2 . If the estimates θ_m are normally distributed, each d_m^2 is χ^2 -distributed with 2 degrees of freedom. In Figure 5, we have plotted both the empirical cumulative distribution together with the distribution function of a χ^2_2 -distribution. The figure suggests that even for population sizes as small as $n = 10$, the estimates seem to be close to the normal distribution. The results were similar for other combinations of true parameters σ^2 and σ_e^2 .

n=100	N	$\hat{\sigma}^2$		$\hat{\sigma}_e^2$		# of EM iterations	# of misclassifications
		mean	std	mean	std		
$\sigma^2 = 3, \sigma_e^2 = 1$	10	3.01	.188	0.99	.061	11	0.9
	20	3.02	.121	1.00	.047	3.2	0.0
	40	2.99	.088	1.00	.031	1.0	0
$\sigma^2 = 2, \sigma_e^2 = 1$	10	1.98	.141	1.01	.068	13	2.5
	20	2.01	.091	1.01	.046	5.6	0.1
	40	2.00	.061	1.00	.028	1.3	0
$\sigma^2 = 1, \sigma_e^2 = 1$	10	1.01	.088	1.00	.054	19	6.4
	20	1.00	.054	0.99	.036	8.0	0.8
	40	1.00	.037	1.00	.027	2.9	0
$\sigma^2 = 1, \sigma_e^2 = 2$	10	1.00	.114	2.00	.094	33	13
	20	1.00	.068	2.00	.058	12	3.0
	40	1.00	.044	2.00	.044	5.0	0.1
$\sigma^2 = 1, \sigma_e^2 = 3$	10	1.00	.131	3.00	.128	50 ¹	17
	20	1.00	.071	3.01	.082	14	5.4
	40	1.00	.058	2.99	.059	7.8	0.4

¹Here, for 7 of the simulated sequences, the EM algorithm failed to converge within 100 iterations.

Table 1: Comparison of the estimates of σ^2 and σ_e^2 , with the corresponding standard deviation, the average number of EM iterations, and the average number of misclassifications, for 100 realizations of simulations with $n = 100$ particles, of which 20 were fixed, for different values of σ^2 , σ_e^2 , and N .

6 Discussion

When comparing the estimated classification variables to the plotted trajectories, we see that our method indeed manages to classify the particles correctly for this data set. In addition, the theoretical diffusion coefficient derived from Stoke-Einstein's relation is within the 95% confidence interval of the estimated diffusion coefficient. These two results, in combination with the simulation study in Section 5, demonstrate the efficiency of the method and are encouraging for the future analysis of larger data sets of this kind.

We should, however, keep in mind the discrete nature of the observed trajectories; the image processing used for this image sequence only positioned particles up to pixel level. Hence, the observed likelihood, here based on continuous spatial data, could be modified to take this into consideration. Alternatively, the image processing could be refined to obtain sub-pixel accuracy.

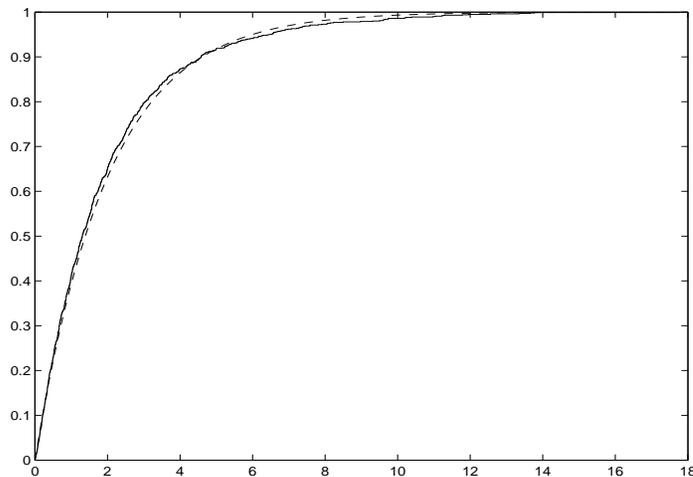


Figure 5: The χ^2 distribution function (dashed), together with the empirical cumulative distribution of the generalized distances of the estimates of σ^2 and σ_e^2 from 1000 simulated sequences of $n = 10$ particles, of which 2 were fixed, when $N = 10$ and the true parameter values were $\sigma^2 = 2$ and $\sigma_e^2 = 1$.

Nevertheless, the main result in this paper is the theoretical analysis and the data should primarily be considered as an illustrative example of an application of the method.

A natural question arises on the possibility for the particles to change states during the observation period of $N + 1$ images; that is, from diffusing to fixed or from fixed to diffusing. For colloidal particles like those in Figure 1, at least the latter possibility is believed to be highly unlikely; once a particle has become adsorbed on one of the specimen glasses, it continues to be so for the duration of the experiment. To change from diffusing to fixed is of course possible, but for the observation lengths we have studied so far, this has also been ruled out; remember that N equal to 11 corresponds to a total observation length of 440 ms. After future refinements of the image processing part however, we plan to look at far larger observation lengths and then the scenario of a particle getting stuck becomes more likely. The model must then be modified and one alternative might be a hidden Markov model.

Another way to generalize the model, would be to allow for the particles to have different sizes and hence different diffusion coefficients, as discussed in Section 4.4. If we regard the diffusion coefficients as random variables from a common (but unknown) distribution the task would then be to estimate the distribution of the diffusion coefficients rather than the actual diffusion coefficients. Another, simpler and more direct, alternative is to allow for two or three classes of particles, each with a fixed (but unknown) diffusion variance.

References

- O.E. BARNDORFF-NIELSEN and D.R. COX. *Inference and Asymptotics*. Chapman & Hall London, first edition, 1994.
- B.M. BIBBY and M. SØRENSEN. Martingale estimation functions for discretely observed diffusion processes. *Bernoulli*, 1:17–39, 1995.
- R.C.H. CHENG and W.B. LIU. The consistency of estimators in finite mixture models. *Scand. J. Statist.*, 28:603–616, 2001.
- N.E. DAY. Estimating the components of a mixture of normal distributions. *Biometrika*, 56:463–474, 1969.
- A.P. Dempster, N.M. LAIRD, and D.B. RUBIN. Maximum likelihood from incomplete data via the EM algorithm. *J. Roy. Statist. Soc. Ser. B*, 39:1–38, 1977.
- G. DOHNAL. On estimating the diffusion coefficient. *J. Appl. Probab.*, 24:105–114, 1987.
- D.F. EVANS and H. WENNERSTRÖM. *The colloidal domain. Where Physics, Chemistry, Biology, and Technology meet*. Wiley-VCH, New York, second edition, 1999.
- D. FLORENS-ZMIROU. On estimating the diffusion coefficient from discrete observations. *J. Appl. Prob.*, 30:790–804, 1993.
- V. GENON-CATALOT and J. JACOD. Estimation of the diffusion coefficient for diffusion processes: random sampling. *Scand. J. Statist.*, 21:193–221, 1994.
- V. HASSELBLAD. Estimation of finite mixtures of distributions from the exponential family. *Amer. Statist. Assoc.*, 64:1459–1471, 1969.
- R.J. HATHAWAY. A constrained formulation of maximum-likelihood estimation for normal mixture distributions. *Ann. Statist.*, 13:795–800, 1985.
- M. HOFFMANN. L_p estimation of the diffusion coefficient. *Bernoulli*, 5:447–481, 1999.
- M. HOFFMANN. On estimating the diffusion coefficient: parametric versus nonparametric. *Ann. Inst. H. Poincaré Probab. Statist.*, 37:339–372, 2001.
- J. JACOD. Non-parametric kernel estimation of the coefficient of a diffusion. *Scand. J. Statist.*, 27:83–96, 2000.
- M. KESSLER and M. SØRENSEN. Estimating equations based on eigenfunctions for a discretely observed diffusion process. *Bernoulli*, 5:299–314, 1999.
- J. KIEFER and J. WOLFOWITZ. Consistency of the maximum likelihood estimator in the presence of infinitely many incidental parameters. *Ann. Math. Statist.*, 27:887–906, 1956.

- M. KVARNSTRÖM. Estimating diffusion coefficients in colloidal particle systems. Technical report, Mathematical Sciences, Chalmers University of Technology, <http://www.math.chalmers.se/Stat/Research/Preprints/>, 2002. Licentiate Thesis.
- J.K. LINDSEY. *Parametric Statistical Inference*. Oxford University Press, Oxford, first edition, 1996.
- T.A. LOUIS. Finding the observed information matrix when using the EM algorithm. *J. Roy. Statist. Soc. Ser. B*, 44:226–233, 1982.
- G. MCLACHLAN and T. KRISHNAN. *The EM algorithm and Extensions*. Wiley-Interscience, New York, first edition, 1997.
- G. MCLACHLAN and D. PEEL. *Finite mixture models*. Wiley-Interscience, New York, first edition, 2000.
- R. REDNER. Note on the consistency of the maximum likelihood estimate for nonidentifiable distributions. *Ann. Statist.*, 9:225–228, 1981.
- R.A. REDNER and H.F. WALKER. Mixture densities, maximum likelihood and the EM algorithm. *SIAM Rev.*, 26:195–239, 1984.
- R. SUNDBERG. Maximum likelihood theory for incomplete data from an exponential family. *Scand. J. Statist.*, 1:49–58, 1974.
- A.W. VAN DER VAART. *Asymptotic Statistics*. Cambridge University Press, Cambridge, U.K., first edition, 1999.
- A. WALD. Note on the consistency of the maximum likelihood estimate. *Ann. Math. Statist.*, 20:595–601, 1949.
- C.F.J. WU. On the convergence properties of the EM algorithm. *Ann. Statist.*, 11:95–103, 1983.

Acknowledgements

The author wishes to thank Mats Rudemo for his supervision, Ib Skovgaard for interesting and helpful discussions regarding the asymptotics, and Lennart Lindfors of AstraZeneca R&D for producing the images. The author is also grateful for the comments and suggestions from the referee.

Appendix: Sufficient statistics

Consider the complete data density (5) with observed trajectories in d dimension. Take the logarithm and let \tilde{Y}_{lk} denote the k :th element of the transformed increment vector (see section 2.1) of the l :th dimension trajectory. We get

$$\begin{aligned} \log g_C &= z \log p - z \frac{d}{2} \sum_{k=1}^N \log(\sigma^2 + \lambda_k \sigma_e^2) - \frac{z}{2} \sum_{k=1}^N \frac{\sum_l \tilde{y}_{lk}^2}{\sigma^2 + \lambda_k \sigma_e^2} + \\ &+ (1-z) \log(1-p) - (1-z) \frac{d}{2} \sum_{k=1}^N \log(\lambda_k \sigma_e^2) - \frac{1-z}{2} \sum_{k=1}^N \frac{\sum_l \tilde{y}_{lk}^2}{\lambda_k \sigma_e^2} \\ &= \sum_{k=1}^N z \left(\sum_{l=1}^d \tilde{y}_{lk}^2 \right) \left(-\frac{1}{2} \frac{1}{\sigma^2 + \lambda_k \sigma_e^2} \right) - \frac{1}{2\sigma_e^2} \sum_{k=1}^N \frac{1-z}{\lambda_k} \left(\sum_{l=1}^d \tilde{y}_{lk}^2 \right) \\ &+ z \left(\log\left(\frac{p}{1-p}\right) - \frac{d}{2} \sum_{k=1}^N \log\left(\frac{\sigma^2 + \lambda_k \sigma_e^2}{\lambda_k \sigma_e^2}\right) \right) - \left(\frac{d}{2} \sum_{k=1}^N \log(\lambda_k \sigma_e^2) - \log(1-p) \right) \end{aligned}$$

and we see that a minimal sufficient statistic can be chosen to be

$$\begin{aligned} t_1 &= z \sum_{l=1}^d \tilde{y}_{l1}^2 \\ &\vdots \\ t_N &= z \sum_{l=1}^d \tilde{y}_{lN}^2 \\ t_{N+1} &= \sum_{k=1}^N \frac{1-z}{\lambda_k} \left(\sum_{l=1}^d \tilde{y}_{lk}^2 \right) \\ t_{N+2} &= z \end{aligned}$$

with the corresponding canonical parameter α

$$\begin{aligned} \alpha_1 &= -\frac{1}{2} \frac{1}{\sigma^2 + \lambda_1 \sigma_e^2} \\ &\vdots \\ \alpha_N &= -\frac{1}{2} \frac{1}{\sigma^2 + \lambda_N \sigma_e^2} \\ \alpha_{N+1} &= -\frac{1}{2\sigma_e^2} \\ \alpha_{N+2} &= \log\left(\frac{p}{1-p}\right) - \frac{d}{2} \sum_{k=1}^N \log\left(\frac{\sigma^2 + \lambda_k \sigma_e^2}{\lambda_k \sigma_e^2}\right) \end{aligned}$$

which is a function of our parameter θ . Since this is 3-dimensional and the sufficient statistics is $(N + 2)$ -dimensional, we say that the complete data belongs to a curved exponential family or, with the terminology of BARNDORFF-NIELSEN and COX (1994), a $(N + 2, 3)$ -exponential model.

Solving for p in the expression for α_{N+2} above, we get

$$p = \frac{e^{\alpha_{N+2}} \prod_{k=1}^N \left(\frac{\sigma^2 + \lambda_k \sigma_e^2}{\lambda_k \sigma_e^2} \right)^{d/2}}{1 + e^{\alpha_{N+2}} \prod_{k=1}^N \left(\frac{\sigma^2 + \lambda_k \sigma_e^2}{\lambda_k \sigma_e^2} \right)^{d/2}} = \frac{e^{\alpha_{N+2}} \prod_{k=1}^N \left(\frac{\alpha_{N+1}}{\lambda_k \alpha_k} \right)^{d/2}}{1 + e^{\alpha_{N+2}} \prod_{k=1}^N \left(\frac{\alpha_{N+1}}{\lambda_k \alpha_k} \right)^{d/2}}$$

and we can write the logarithm of the complete data density as

$$\log g_C = \alpha^T t - k(\alpha) \tag{22}$$

where $\alpha = \alpha(\theta)$ and k becomes

$$k(\alpha) = \frac{d}{2} \log(N + 1) - \frac{dN}{2} \log(-2\alpha_{N+1}) + \log \left(1 + e^{\alpha_{N+2}} \prod_{k=1}^N \left(\frac{\alpha_{N+1}}{\lambda_k \alpha_k} \right)^{d/2} \right) \tag{23}$$

From standard theory of exponential families, we get the cumulants of the sufficient statistics by differentiating $k(\alpha)$. In particular, we have $\mathbf{E}T = \frac{\partial k}{\partial \alpha}$ and $\mathbf{Var}\{T\} = \frac{\partial^2 k}{\partial \alpha \partial \alpha^T}$, which we denote μ and V , respectively.

The expectation of the sufficient statistics can be written

$$\mathbf{E}T = \begin{bmatrix} dp(\sigma^2 + \lambda_1 \sigma_e^2) \\ \vdots \\ dp(\sigma^2 + \lambda_N \sigma_e^2) \\ d(1 - p)N\sigma_e^2 \\ p \end{bmatrix}$$