

Spatial statistics and image analysis. Lecture 11

Mats Rudemo

May 11, 2020

Today's lecture will cover:

RICS (Raster Image Correlation Spectroscopy) and SPRIA (Single Particle Raster Image Analysis) of diffusion

Identification of 3D gel microstructures from TEM (Transmission Electron Micrographs) images

Structural characterization from scanning TEM images

Single particle raster image analysis of diffusion

Previously we have seen several powerful techniques for estimating diffusion coefficients and particle concentrations from series of images:

- (i) Single particle tracking
- (ii) FRAP (Fluorescence Recovery After Photobleaching)
- (iii) Estimation of particle concentration from single particle tracking
- (iv) Estimation of particle concentration from count time series

In this section we shall study a method that in principle allows us to estimate diffusion coefficients from single images.

The method is based on using a clever raster scan pattern (described by Digman et al. 2005) collected with a confocal laser scanning microscope

During the scanning, adjacent pixels in the x -direction are visited within a short time interval whereas the time between visits to adjacent pixels in the y -direction is much longer, see Figure 1.

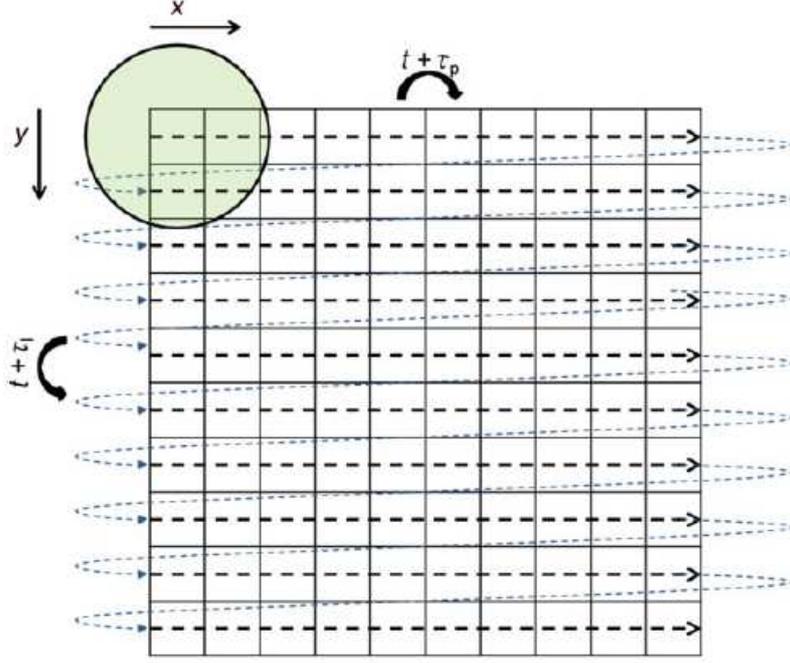


Figure 1: Movement of scanning beam according to the raster scan pattern used in RICS and SPRIA. The scanning time between adjacent pixels in the x - and y -directions are τ_p and τ_l , respectively, and $\tau_p \ll \tau_l$.

The method RICS (Raster Image Correlation Spectroscopy) is suggested in Digman et al. (2005) to analyse raster images by use of correlation function estimation typically computed by the fast Fourier transform method.

The theoretical correlation function $G(\xi, \psi)$ for the scanned image corresponding to two points (x, y) and $(x + \xi, y + \psi)$ is

$$G(\xi, \psi) = \frac{1}{\langle N \rangle} \left(1 + \frac{4D\tau(\xi, \psi)}{\omega_0^2} \right)^{-1} \left(1 + \frac{4D\tau(\xi, \psi)}{\omega_z^2} \right)^{-1/2} \exp \left[-\frac{(S\xi)^2 + (S\psi)^2}{\omega_0^2 + 4D\tau(\xi, \psi)} \right] \quad (1)$$

Here $\langle N \rangle$ is the average number of particles in the observation volume, S is pixel size, and

$$\tau(\xi, \psi) = |\tau_p \xi + \tau_l \psi| \quad (2)$$

corresponding to the time it takes to move between the points (x, y) and $(x + \xi, y + \psi)$.

The parameters ω_0 and ω_z in (1) correspond to the decay rate of the point spread function in the lateral and vertical directions. They can be estimated as standard deviations of a Gaussian distribution fitted to the laser point spread function.

In Longfils et al. (2017) an alternative analysis method SPRIA, Single Particle Raster Image Analysis, is introduced.

Each particle is analyzed separately which opens for analysis of systems of particle mixtures with varying diffusion coefficients and of heterogenous materials with diffusion properties varying with location.

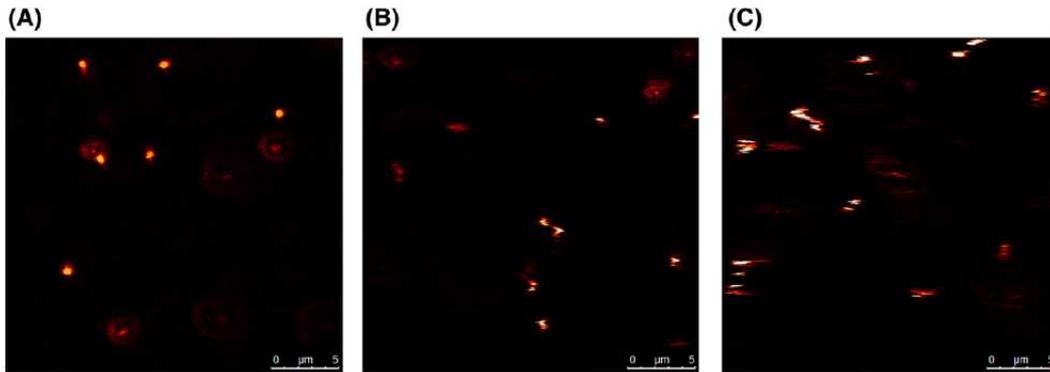


Figure 2: Raster scan images of 175 nm beads at a pixel size of 48.1 nm and varying scan rate: (A) 8000 Hz, (B) 400 Hz and (C) 100 Hz.

The sampling time structure provides information on the dynamics of the particles inside each image. In Figure 2 typical raster scan images are shown for different time scales demonstrating the effect of varying scan rates visually on experimental data with 175-nm beads scanned at decreasing speed.

In (A) at scan rate 8000 Hz almost immobile round-shaped particles are observed. In (B) particles scanned at 400 Hz move slowly and the horizontal particle lines look like sequences of shifted bright lines. In (C) the scan rate is decreased to 100 Hz. The sequence of bright lines corresponding to one particle become even more shifted as particles are moving significantly between aquisition of consecutive lines.

For application of the single particle method, individual particles have to be extracted from an image as shown in Figure 3. The procedure is based on two threshold levels $T_1 > T_2$. First find local photon count maxima above the level T_1 . Then find around each chosen maximum the smallest axis-parallel rectangle such that all observed photon count levels just outside the rectangle border are below T_2 . It turns out that the choice of levels is not critical, which seems quite plausible from a look at Figure 3.

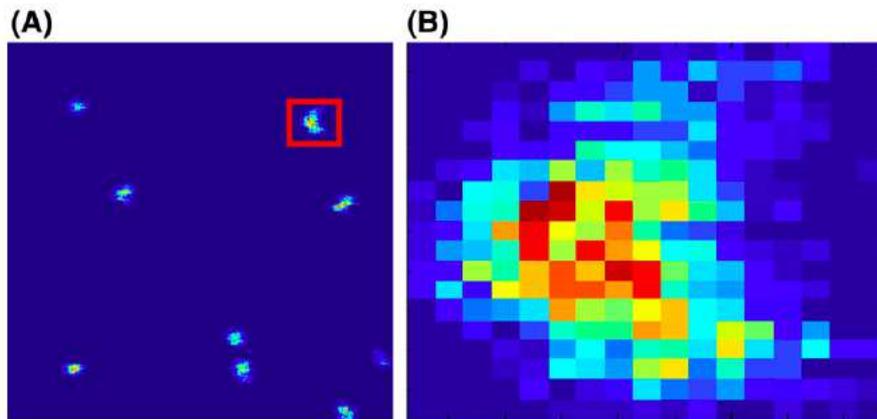


Figure 3: Freely diffusing fluorescent beads with 175 nm diameter. (A) A 256×256 pixel raster scan image. (B) The particle highlighted with a red square in (A) extracted from the image.

Figure 4 shows an axis-parallel rectangle chosen with $T_1 = 10$ and $T_2 = 5$.

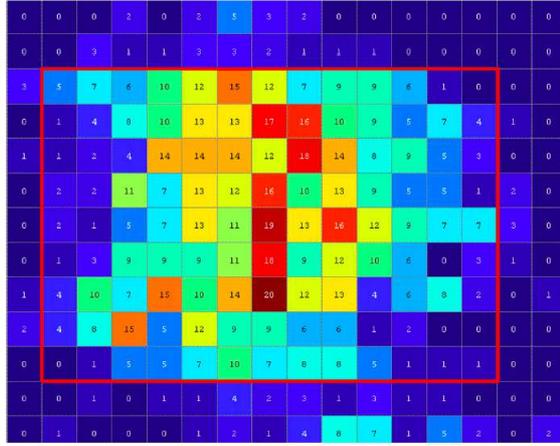


Figure 4: Magnified neighbourhood of a 175-nm bead in a 2048×2048 raster scan image. The numbers correspond to photon counts in each pixel and the local maximum is 20. The levels chosen are $T_1 = 10$ and $T_2 = 5$, and the rectangle defining the particle is found within the contour defined by the red lines. The corresponding border consists of the pixels immediately outside the red contour.

Let a particle P be defined by the axis-parallel rectangle

$$P = \{(x, y) : a < x < a + L, b < y < b + K\} \quad (3)$$

around a local maximum of photon counts. The trajectory of the particle can be estimated by use of the extracted image and used to estimate the diffusion coefficient D of the particle.

In Longfils et al. (2017) a maximum likelihood method for estimating the trajectory and the diffusion coefficient D is described.

Here describe a more direct way to estimate the trajectory and D . Let $t(y)$ denote the time at which we scan the horizontal line at y , and let $N(x, y, t_k)$ denote the measured number of photons for a given particle at the pixel with centre (x, y) at time $t(y) = t_k$, where $k = 0, \dots, K$. The x position at time $t(y) = t_k$ is estimated by the centroid

$$\psi_k = \frac{\sum_{\{(x,y) \in P: t(y)=t_k\}} N(x, y, t_k) \cdot x}{\sum_{\{(x,y) \in P: t(y)=t_k\}} N(x, y, t_k)}. \quad (4)$$

Then

$$\hat{D} = \frac{1}{2\Delta t K} \sum_{k=1}^K (\psi_k - \psi_{k-1})^2, \quad (5)$$

where $\Delta t = \tau_l$ denotes the time between two consecutive line scans, is a suitable estimator of D .

From the simulated image shown in Figure 5, where we know the true trajectory, it can be seen that both the maximum likelihood method (green) and the centroid method (black) give accurate estimates of the true trajectory (red).

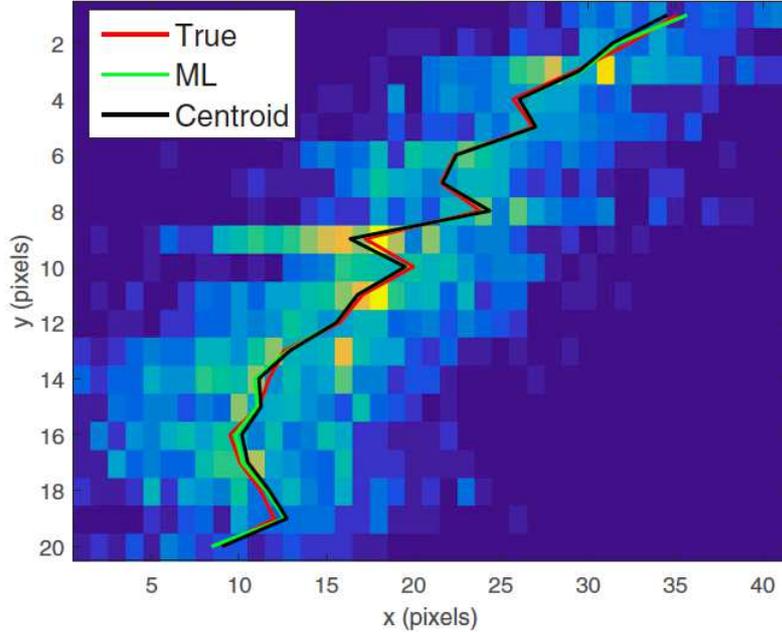


Figure 5: A simulated raster scan image of a particle with true trajectory (red), the corresponding estimated trajectory computed with the maximum likelihood method (green) and with the centroid method (black) see (4).

By use of analyses of images from both simulations and experiments the SPRIA method is in Longfils et al. (2017) evaluated and also compared with the traditional RICS method.

In the simulation study, Gaussian random walk (discrete time Brownian motion) of spheres was generated in a box with periodic boundary conditions. The sphere diameter ranged from 15 to 1000 nm. A minimum image size of 256×256 was chosen. The number of images simulated was chosen so that at least 300 particles and at least 100 images were collected.

The point spread function parameters were $\omega_0 = 248$ and $\omega_z = 1270$ nm corresponding to Gaussian fit to an average measured z -scan of immobile 175-nm fluorescent beads. The pixel dwell time was $\tau_p = 1.71 \times 10^{-7}$ s, the line time $\tau_l = 1.4 \times 10^{-3}$ s and the pixel size $S = 0.03 \mu\text{m}$. The results for 10 D -values between 0.0625 and $64 \mu\text{m}^2\text{s}^{-1}$ are shown in Figure 6 including bootstrap standard deviation estimates obtained by $B = 40$ repetitions.

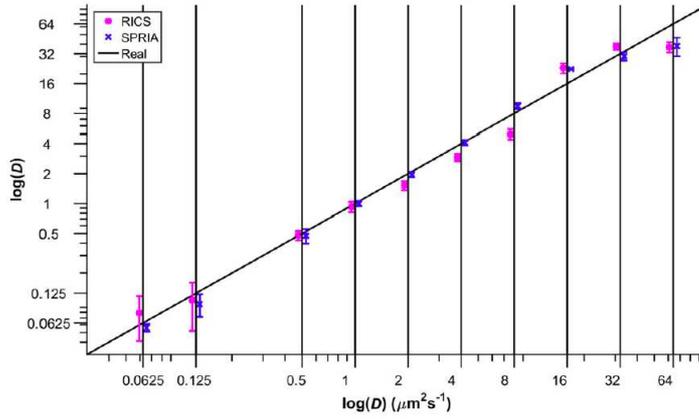


Figure 6: Illustration of simulation results with logarithmic scales on both axes. Vertical black lines correspond to the expected diffusion coefficient according to Stoke-Einstein's equation, and ideally the estimated D -values should be located at the crossings of the black identity line and the vertical black lines. Blue markers refer to SPRIA and magenta to RICS, and both estimates are presented as estimates \pm standard errors. Markers have been moved horizontally relative to each true value of the diffusion coefficient to make the figure more legible.

Results from experiments with particles of four different sizes: 100 nm, 175 nm, 490 nm and 1000 nm, are shown in Figure 7.

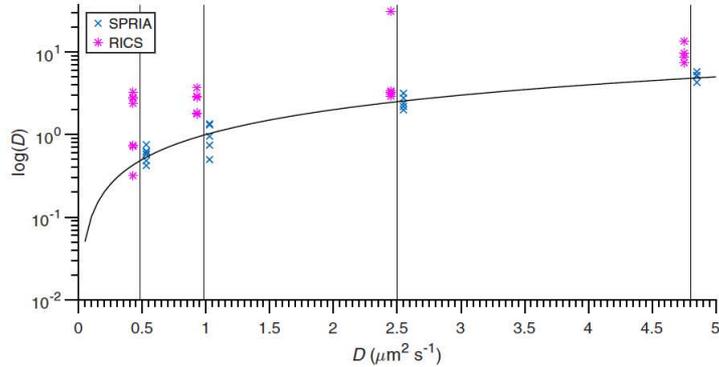


Figure 7: Illustration of experimental results for the SPRIA (blue markers) and the RICS (magenta markers) diffusion-coefficient estimations methods with fluorescent beads. A logarithmic scale is used on the y -axis. Vertical black lines correspond to the expected diffusion coefficient according to Stoke-Einsteins's equation, and ideally the measured D -values should be located at the crossings of the black logarithmic curve and the vertical black lines. Markers have been moved horizontally to make the figure more legible.

Both Figure 6 and Figure 7 show that the SPRIA method performs well for the settings chosen.

To indicate that the SPRIA method may be used to analyse diffusion in heterogenous media a simple simulation example was performed in Longfils et al. (2017) with results shown in Figure 8.

The dynamics of particles is supposed to vary such that the diffusion coefficient is $0.8 \mu\text{m}^2\text{s}^{-1}$ inside a circle and $0.4 \mu\text{m}^2\text{s}^{-1}$ outside the circle. In the simulation 2142 particles in 300 images were found. The pixel size was $0.03 \mu\text{m}$ and the image resolution was 256×256 pixels.

The mobility map shown in the right part of Figure 8 was obtained by smoothing with a Gaussian kernel with bandwidth $\sigma = 15$ pixels. The mobility map indicates that the SPRIA method allows recovery of medium heterogeneity to a certain extent.

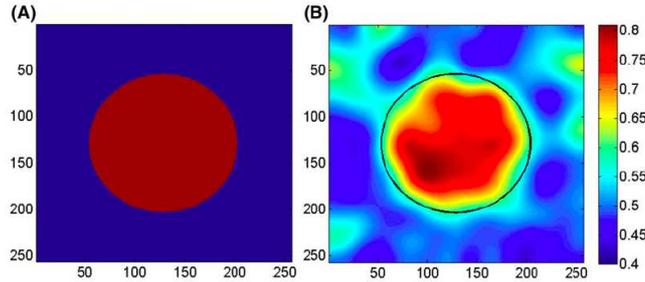


Figure 8: (A) Image of a heterogenous sample used in a simulation study with pixel colours corresponding to the expected diffusion coefficient in $\mu\text{m}^2\text{s}^{-1}$. (B) Mobility map, smoothed with a Gaussian kernel with bandwidth $\sigma = 15$ pixels, based on 2142 particles in 300 images with the boundary circle between the two media in black. The pixel size is $0.03 \mu\text{m}$ and the resolution of the images is 256×256 pixels.

What are the conditions for using the SPRIA method? Roughly two conditions should be satisfied:

- (i) The density of particles must not be too large if individual particles should be identifiable.
- (ii) Sampling time between lines should be such that adjacent horizontal particle lines should differ to some extent, but not too much as particles then become split into several parts.

The second condition is illustrated in Figure 9. Particularly the lower right part (D) shows a situation where the detection algorithm will split the bead into several (presumably five) particles.

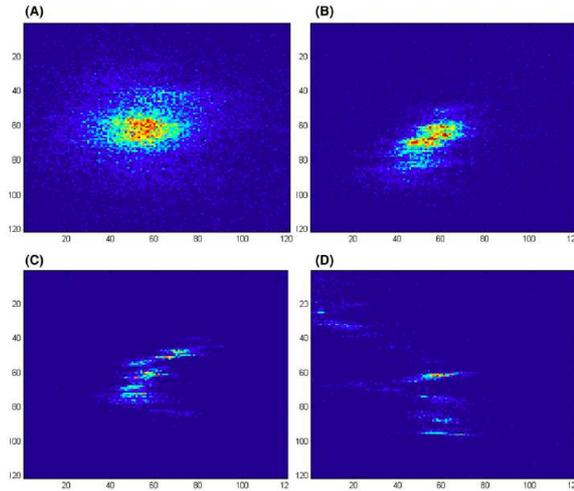


Figure 9: Typical particle images with settings: $S_x = 0.03\mu\text{m}$, $\tau_p = 7.5 \cdot 10^{-7}$ s, $\tau_l = 2.5 \cdot 10^{-3}$ s, and scan rate 400 Hz. (A) 1000-nm bead; (B) 490-nm bead; (C) 175-nm bead; (D) 100-nm bead. The extraction step will identify exactly one bead in (A) and in (B). In (C) the bead may be split into two particles, as the last four to six lines are not connected to the previous. In (D) the algorithm will (depending on the thresholds used) split the bead into presumably five particles (corresponding to the following maxima: one in the top left, and four in the middle lower part of the image separated by lines with low photon counts).

Image analysis of transmission electron micrographs

Identification of three-dimensional gel microstructures

Mass transport in gels depends crucially on local properties of the gel network. In Nisslert et al. (2007) a method for identifying the three-dimensional gel microstructure from statistical information in transmission electron micrographs (TEM) is suggested.

The gel-strand network is modelled as a random graph with nodes and edges (branches). The distribution of edge length, the number of edges at nodes and the angles between edges at a node are estimated from transmission electron micrographs.

The 3D gel network is simulated by Markov chain Monte Carlo (MCMC) methods based on statistical information found from the micrographs. The micrographs can be viewed as projections of stained gel-strands in slices, and a formula is derived for estimating the thickness of the stained gel slice based on the total projected gel-strand length and the number of times that gel-strands enter or exit the slice.

To find relevant features of a gel-strand network we aim at separating the gel-strands from the background, and to create a skeleton showing the network as thin lines.

In Figure 10 we see in the left part a TEM micrograph of a Sepharose gel to be analysed. The image corresponds to an area of about 1700×1700 nm and shows the microstructure of the gel network at a magnification of 10 000 times. Images are saved as 8-bit grey-scale images of size 1024×1024 pixels. One could hope for directly finding the gel-strand network by thresholding, but the histogram in Figure 11 shows that that seems difficult.

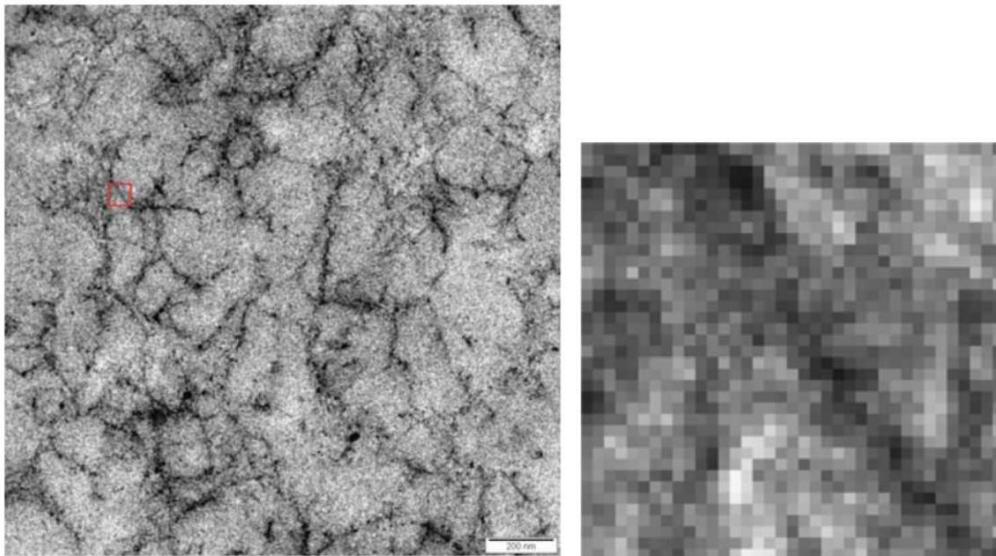


Figure 10: Left: one of the 1024×1024 pixel TEM micrographs showing a Sepharose gel at 10 000 times magnification. Right: magnification of the area within the red box in the upper left part of the left image.

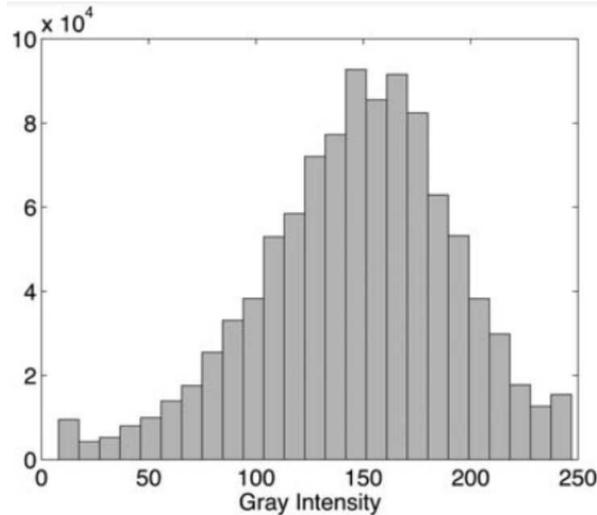


Figure 11: Histogram of the grey level intensity of the micrograph in the left part of Figure 10.

To find the gel-strand network a series of image operations are performed as shown in Figure 12. We will here give a brief description of the different steps illustrated in this figure.

The upper left image in Figure 12 shows the starting image. To decrease the level of noise a Wiener filter is used to smooth the image. The Wiener filter is a low-pass filter that uses prior information about the noise in the image to optimize the noise reduction.

As the intensities of pixels that represent the gel-strands vary considerably, a minimum filter is then applied. This takes the minimum pixel value from a 3×3 neighbourhood, resulting in an overall darker image with a more uniform grey-level polymer structure as shown in the right part in the upper figure row.

The minimum filter also produces some new dark spots in the background, and to suppress these a 5×5 median filter is applied.

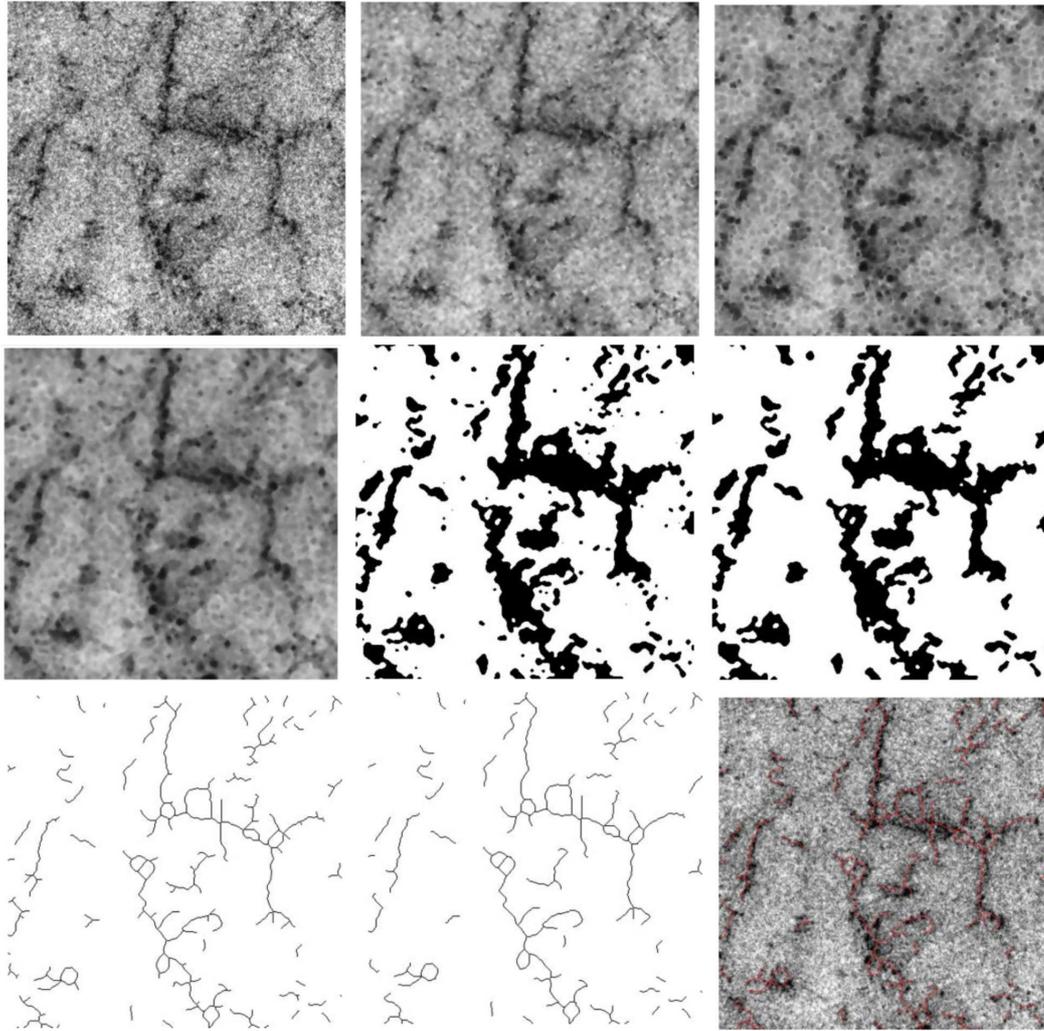


Figure 12: Illustration of the sequence of image analysis steps used to find the gel-strand network. In the three upper images we see (from left to right): a 400×400 pixel part of the image to left in Figure 10, the same image after application of a Wiener filter, and the result of applying a minimum filter. In the middle row we see similarly (from left to right): the result after application of a 5×5 median filter, the binary image after thresholding, and the result after removal of small black spots. The lower row shows (from left to right): the skeleton created from the binary image, a skeleton with branches shorter than 10 pixels removed, and finally to the lower right the resulting skeleton displayed in red onto the original image.

The gel-strand network is now somewhat better separated from the background. Thresholding is still problematic, but possible. The method of Otsu (1979) is used with the thresholds determined by minimizing the intraclass variance between dark and light pixel values.

In the thresholding the image was further divided in 25 different parts with individual thresholds to take care of varying intensity in different parts of the image. The threshold values from Otsu's method are also decreased with 4% to get the gel-strands thinner. Further black spots smaller than 80 pixels are removed as shown in the right image in the middle row of the Figure 12.

Having now a binary image which reasonably well represents the gel structure in the micrographs, the next step is to create a skeleton representation of the network. Two methods were tried: *medial axis transformation* and *thinning*, compare Sonka et al. (2015).

Both methods create skeletons that represent the microstructure of the gel network rather well, but a lot of small branches are created without correspondence in the original micrographs. Comparing skeletons produced by the two methods, thinning was found to produce less of these artefacts and was chosen.

Branches shorter than 10 pixels are also removed from the skeletons. The resulting network is shown in the middle image of the lower row in Figure 12. Finally in the lower right part of the figure we see the resulting skeleton in red overlaid on the original image.

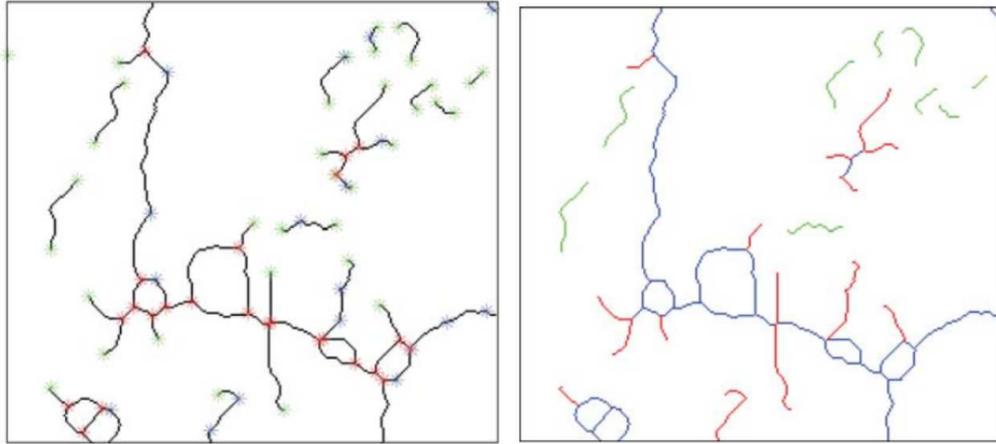


Figure 13: Left: Part of a micrograph skeleton with nodes (with three branches) marked in red, end points in green and bending points in blue. Right: Skeleton with branches between nodes marked in blue, branches between end-points in green and branches between an end point and a node in red.

By looking at the locations in the gel-strand network where branches have been removed we can find a special type of nodes called bending points, as the removed short branches were often created where the gel strands change direction.

In the left part of Figure 13 part of a skeleton is shown with nodes (with three branches) marked in red, end points in green and bending points in blue. The bending points are in the sequel treated as nodes having only two connected branches.

By using the end points and nodes we can measure branch lengths. This is done by measuring the straight line distance between two end points, two nodes or one end point and one node. The branches are also measured more accurately by looking at the distance between the pixels in the branches, where the distance between horizontally or vertically adjacent pixels is one and between diagonally adjacent pixels is $\sqrt{2}$.

In a skeleton we see, compare the right part of Figure 13, three types of branches. Note that we look at projections of a thin

section and that we assume full connectivity in the gel network. Branches between two end points correspond to gel-strands that enter and exit the section. Branches between an end point and a node correspond to gel-strands that enter or exit the section, and branches between two nodes are fully contained in the section.

We want to find the distribution of angles between branches at nodes. As we have a discrete pattern of pixels we cannot just use angles between adjacent pixels as that would give only multiples of 45° . Choosing a point too far from a node increases the probability of an incorrect branch direction. As a compromise the direction eight pixels from a node was used.

Further, to find the thickness of gel-strands a binary representation of the gel network and a distance map, see Breu et al. 1995 is constructed. The distance map gives the shortest distance for each pixel to the nearest non-gelstrand pixel. Combining this with the skeleton representation of the gel microstructure gives the radius distribution of gel-strands.

Let us now look at slice thickness estimation. To obtain contrast in the TEM images a thin slice of gel is put into a uranium and lead bath. Uranium and lead are allowed to diffuse into the gel for about one hour and are adsorbed on the gel-strands. In Nisslert et al. (2007) a method is proposed to estimate slice thickness, which corresponds to how far uranium and lead has diffused into the gel. The method as described below is based on the total projected gel-strand length and the number of gel-strand crossings into or out of the slice.

Assume that the gel-strand network consists of a collection of curves such that essentially all end points in the 2D projection correspond to a curve passing into or out of a slice of thickness D . Gel-strand curves are supposed to have locally a well-defined orientation. Figure 14 shows a short approximately linear curve segment ΔC of length ΔL that forms an angle α with a vertical line and has vertical distance y from segment midpoint to the bottom of the slice.

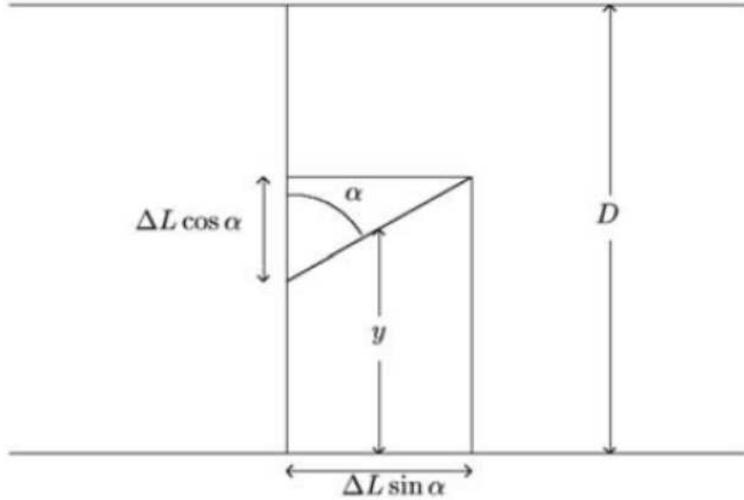


Figure 14: Short approximately linear curve segment ΔC of length ΔL with orientation α relative to a vertical line and with a vertical distance y from midpoint to bottom of slice. The slice is oriented such that it has horizontal upper and lower boundaries.

Assuming that the line segment is randomly rotated in 3D around its center one finds that the angle α has probability density

$$p(\alpha) = \sin \alpha, \quad 0 \leq \alpha \leq \pi/2, \quad (6)$$

compare Exercise 13.1. The length of the horizontal projection of the corresponding curve segment ΔC is $L_{\text{hor}}(\Delta C) = \Delta L \sin \alpha$ with expectation

$$\mathbf{E}(L_{\text{hor}}(\Delta C)) = \int_0^{\pi/2} \Delta L \sin \alpha p(\alpha) d\alpha = \frac{\pi \Delta L}{4}, \quad (7)$$

compare Baddeley & Jensen 2005).

To evaluate the number of crossings we condition on the angle α . The length of the vertical projection of the curve segment is $\Delta L \cos \alpha$, and we get a crossing out of the slice if either $y < (\Delta L/2) \cos \alpha$ or $y > D - (\Delta L/2) \cos \alpha$. Thus the conditional expectation of the number of crossings of curves moving out of the slice is $(1/D)\Delta L \cos \alpha$. Since we should have equally many crossings from curves moving into the slice, the conditional expectation of the total number of crossings is $(2/D)\Delta L \cos \alpha$. Averaging over α we get the expected number of crossings

$$\mathbf{E}(N_{\text{cross}}(\Delta C)) = \frac{2}{D} \int_0^{\pi/2} \Delta L \cos \alpha \sin \alpha d\alpha = \frac{\Delta L}{D}. \quad (8)$$

Summing over all curve segments we find from (7) that the expected horizontal curve length is

$$\mathbf{E}(L_{\text{hor}}) = \frac{\pi L}{4}, \quad (9)$$

where L is the total 3D curve length, and similarly the expected number of crossings is

$$\mathbf{E}(N_{\text{cross}}) = \frac{L}{D}. \quad (10)$$

From the last two equations we find

$$\mathbf{E}(N_{\text{cross}}) = \frac{4}{\pi D} \mathbf{E}(L_{\text{hor}}). \quad (11)$$

This means that the slice thickness D can be estimated from the horizontal length L_{hor} in the projection and the number of crossings N_{cross} into or out the slice, and we get the estimate

$$\hat{D} = \frac{4L_{\text{hor}}}{\pi N_{\text{cross}}}. \quad (12)$$

Let us now see how the 3D microstructure can be modeled as a random graph $G = (V, E)$ and reconstructed by use of MCMC simulation. A random graph consists of a set $V = \{v_1, \dots, v_N\}$ of nodes (vertices) and a set E of undirected edges between nodes. An edge between two nodes u and v is denoted $\langle u, v \rangle$.

In our MCMC modelling of the gel-strand network we let nodes in the graph be either branching points or bending points. An edge $\langle u, v \rangle \in E$ if and only if there is a direct connection (that is a branch) between the nodes u and v .

Define a probability measure for the random graph $G = (V, E)$ by use of a real-valued energy function $f(g)$ such that the probability density at $G = g$ is

$$\pi_f(g) = \frac{1}{Z_f} e^{-f(g)}, \quad (13)$$

where Z_f is a normalizing constant ensuring that the probability density $\pi_f(g)$ integrates to one. The chosen energy function is

$$f(g) = f_1(g) + f_2(g) + f_3(g), \quad (14)$$

corresponding to lengths of edges, number of edges at nodes and angles between edges at nodes, respectively.

The first energy component is

$$f_1(g) = c_1 \sum_{u \sim v} (d_{uv} - d_0)^2, \quad (15)$$

where u and v are nodes of g , d_{uv} is the distance between them and $u \sim v$ means that there is an edge between u and v . The target distance d_0 is essentially the mean branch length (in pixels) from the micrographs multiplied with a factor $4/\pi$ to compensate for going from two to three dimensions, compare (7).

The second energy component corresponds to the number edges connected to nodes. In the micrographs almost all intersection points have three connected edges, but since bending points that have only two edges are also included, the target number n_0 of connections will be somewhat less than three (about 2.7). An energy component taking this into account is

$$f_2(g) = c_2 \sum_v (n_v - n_0)^2, \quad (16)$$

where we sum over all nodes v of g and n_v is the number of edges at v .

The third energy component is

$$f_3(g) = c_3 \sum_v \sum_{\alpha_v} (\alpha_v - \alpha_0)^2, \quad (17)$$

where we sum over all nodes v of g and over all angles α_v , less than π radians, between edges connected to v . With three edges at a node we expect by symmetry angles around 120° degrees, but for nodes with two edges we expect somewhat larger angles.

In the reconstruction of the gel network a sequence of graphs, $G_n = (V_n, E_n)$, $n = 0, 1, \dots, n_{\max}$, is updated with a Metropolis algorithm. The updating from G_n to G_{n+1} is performed with two types of steps. In the first type of steps we add or remove an edge to E_n (or leave E_n unchanged). In the second type we move one of the nodes of V_n a random distance (or leave V_n unchanged). The number N of nodes in V_n is kept constant. The two types of steps are performed as follows.

Type 1 step. Pick a random pair $\{u, v\}$ of nodes. If $\langle u, v \rangle \in E_n$ we first remove it. To obtain E_{n+1} either add $\langle u, v \rangle$ or leave the set of edges unchanged according to the conditional π_f distribution given V_n and all other edges of E_n .

Type 2 step. Pick a node $v \in V_n$ and sample a random movement of the position of the node v . The random movement ΔR is sampled from a uniform distribution in a 3D sphere with radius 2 pixels. Consider the conditional π_f distribution given all other nodes of V_n and the set of edges E_{n+1} , and choose according to this distribution V_{n+1} either equal to V_n or equal to V_n with the chosen node moved by ΔR .

Further details of the MCMC simulation, including choice of the weights (c_1, c_2, c_3) and the target values (d_0, n_0, α_0) in the energy function can be found in Nisslert et al. (2007). As always in this type of simulations a large number of updates is needed.

Let us now show some results obtained from the image analysis. In Figure 15 we see distributions of branch lengths, of angles between branches and of gel-strand thickness.

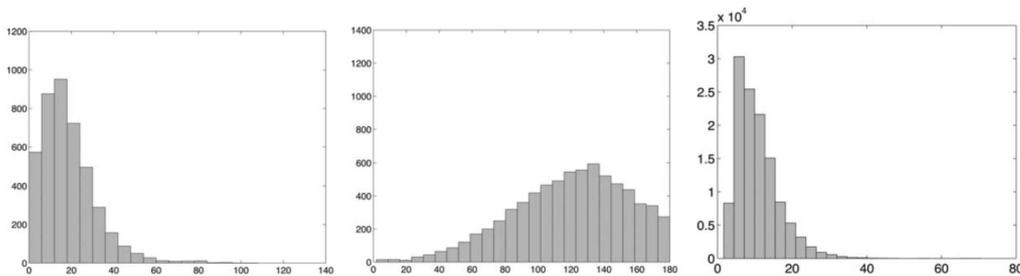


Figure 15: Left: Distribution of branch lengths (in nm) between nodes, one pixel corresponds to 1.66 nm. Middle: Distribution of angles (in degrees) between branches at nodes. Right: Distribution of gels strand thickness (gel-strand radius in nm).

The slice thickness of the section, from which the micrographs are projections, was estimated by use of (12). The total projected length was obtained as the sum of all branch lengths, and the total number of crossings was obtained as the total number of end points in the skeleton (excluding end points at borders).

From six images this resulted in a thickness estimate of 90.6 nm with a standard error of 3.5 nm. It was also concluded that uranium and lead had penetrated most of the slice.

Let us briefly look at the results from the MCMC simulations which were performed in an approximately $800 \times 800 \times 800$ nm cubic box with periodic boundary conditions. With a gel slice thickness of 90 nm, the node intensity per volume unit was found, corresponding to $N = 1604$ in the cubic box and this N value was used in the simulations.

To achieve approximate stationarity in the MCMC simulations 200 hundred iterations were used. In each of these iterations first 10 000 iterations were performed with connections between nodes created or removed (see *Type 1 step* above), and then 10 000 iterations were performed with movement of nodes (see *Type 2 step* above).

After obtaining the skeleton, the branches were thickened to obtain a volume percent of 4%. A result with a 3D rendering of the simulated gel network is shown in Figure 16.

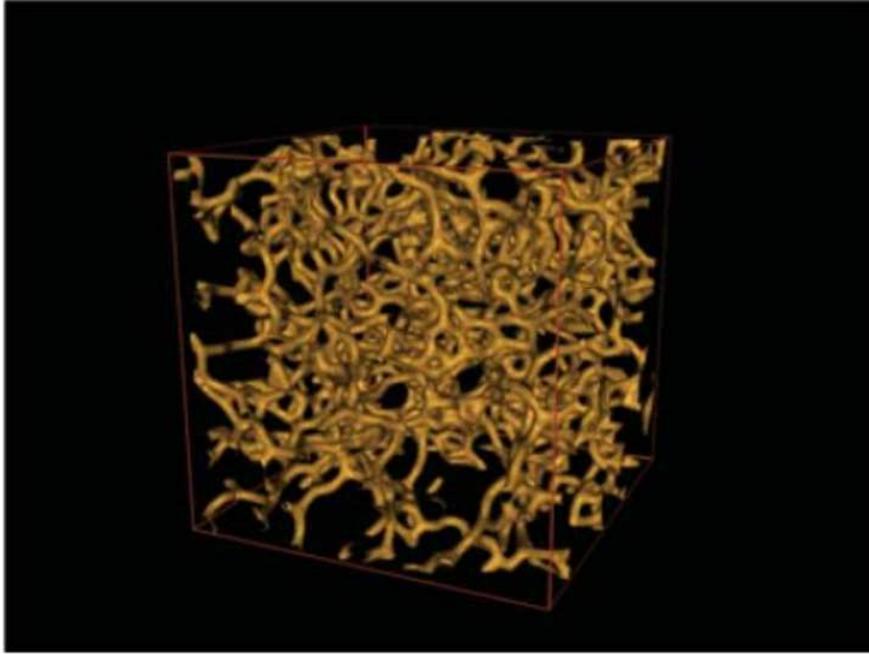


Figure 16: Screen shot of a 3D rendering of a simulated gel network with constant gel-strand radius of 9.2 nm, in a cube with 500 pixels side corresponding to 830 nm.

To evaluate the image processing algorithms and the MCMC simulation algorithms we compare in Figure 17 a micrograph with computed skeleton (red) to the left with a simulated "micrograph" also with a computed skeleton (red) to the right. Visually the left and right part look rather similar.

Finally, in Figure 18 we show (in black) the projection of an actual simulated 3D skeleton and (in red) the corresponding skeleton obtained from image processing. The skeletons look quite similar, indicating that the image processing actually succeeds in creating a skeleton representation from the micrographs.

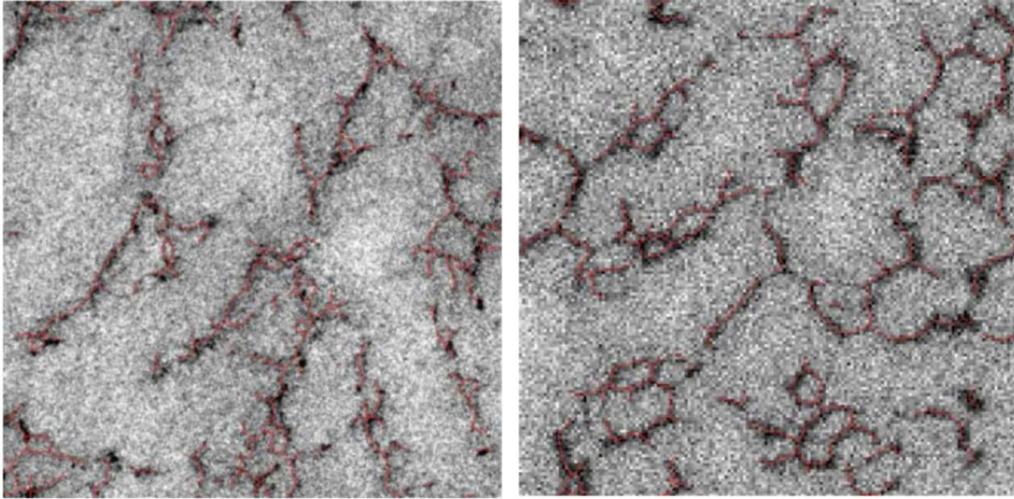


Figure 17: Left: Part of a micrograph, 500×500 pixels, with corresponding skeleton marked in red. Right: Simulated "micrograph", 500×500 pixels, with corresponding skeleton marked in red.

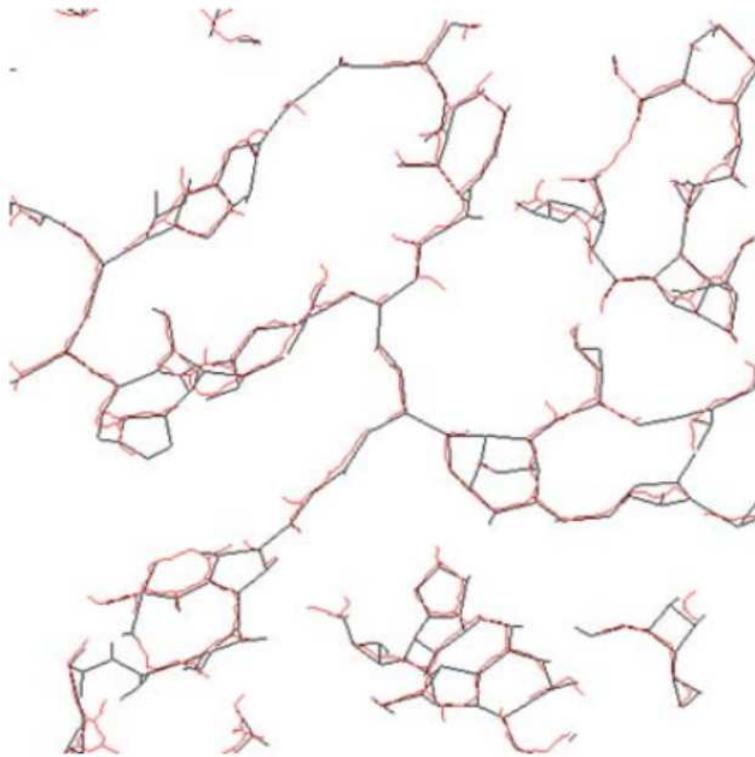


Figure 18: Projection of the actual simulated 3D skeleton (displayed in black), size 500×500 pixels, compared with the corresponding skeleton obtained from the image processing (displayed in red).

Structural characterization from scanning transmission electron micrographs

In Nordin et al. (2014) material characterization from high angle annular dark field scanning transmission electron micrograph (HAADF-STEM) images is studied by use of maximum likelihood methods. The upper left part of Figure 19 shows the intensity response in 2D from a model material consisting of 20nm silica particles that have aggregated to form a stable particle gel. The approximately square image is obtained by scanning a gel slice of thickness about 90 nm. The mass thickness $\alpha(x, y)$ of the silica at a 2D position (x, y) can be written as fraction of the total sample thickness

$$\alpha(x, y) = \frac{1}{z_{\max}} \int_0^{z_{\max}} \rho_S(x, y, z) dz, \quad (18)$$

where $\rho_S(x, y, z)$ denotes the silica density (depending on whether there is a silica particle at the point (x, y, z) or not). A simple model for the observed intensity $I(x, y)$ is

$$I(x, y) = b + cg(\alpha(x, y)) + \epsilon(x, y), \quad (19)$$

where b and c are constants, g is a power function

$$g(\alpha) = \alpha^\beta, \quad (20)$$

and the noise $\epsilon(x, y)$ is supposed to be $N(0, \sigma^2)$ and independent for different pixels (x, y) .

Let us give a maximum likelihood function for an image such as the one shown in the upper left part of Figure 19. Suppose that in the corresponding 3D rectangular box, compare Lecture 10, there are N silica particles with 3D centres $\mathbf{x}_1, \dots, \mathbf{x}_N$. Then the

following log-likelihood function corresponds to the model (19)

$$\ell(\theta) = -|M| \log \left(\sqrt{2\pi}\sigma \right) - \frac{1}{2\sigma^2} \sum_{(x,y) \in M} [I_0(x,y) - b - cg(\alpha(x,y))]^2, \quad (21)$$

where M is the set of pixels, $|M|$ is the number of elements in M and the parameter vector is $\theta = (b, c, \beta, N, \mathbf{x}_1, \dots, \mathbf{x}_N)$. A complication in finding maximum likelihood estimates is the large number particles, see Figure 19, but it can be done (at least approximately), for instance by use of simulated annealing.

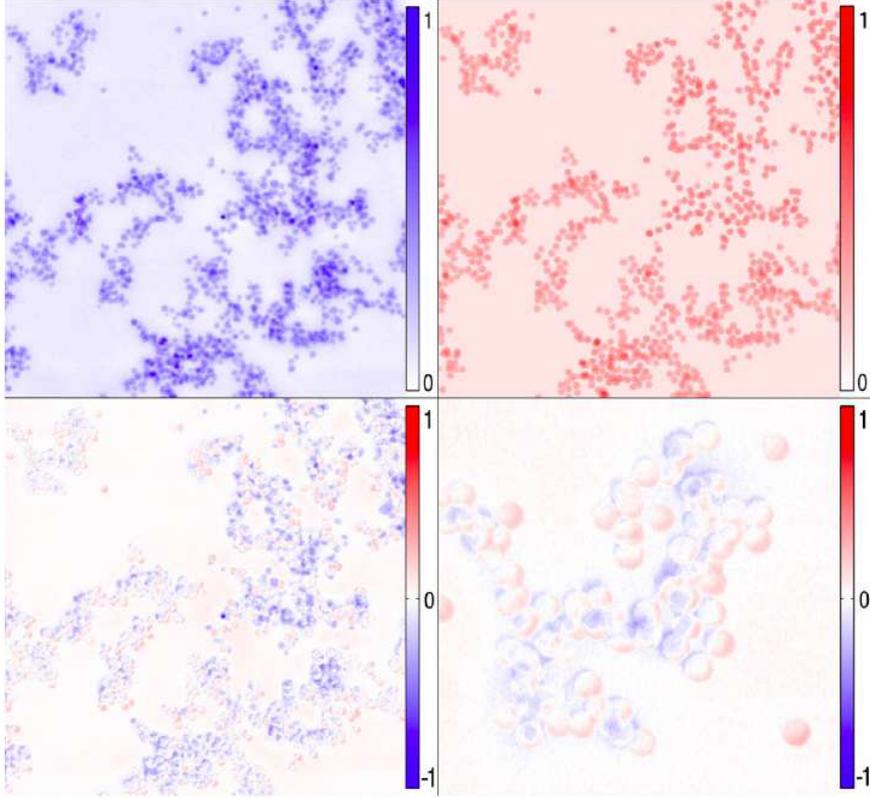


Figure 19: Upper left: Micrograph intensity image $I_0(x,y)$ of a 90 nm slice of 5 wt% aggregated nano silica obtained with HAADF-STEM. Upper right: Regenerated micrograph image $I(x,y)$ obtained by maximizing the log-likelihood function (21). Lower left: Residual image $I(x,y) - I_0(x,y)$ of the original micrograph and the regenerated micrograph. Lower right: Blow-up of the top left cluster in the residual image.

The maximum likelihood estimate of the power parameter β in (20) was $\hat{\beta} = 0.69$, which gave a considerably better fit compared to the linear response with $\beta = 1$ as shown in Figure 20. The models with $\beta = 1$ and a general β are nested and the hypothesis $\beta = 1$ can be tested for instance by use of repeated images. In the present case there was only one image available. To obtain approximately independent repetitions the image was divided into 16 subimages of equal size by three vertical and three horizontal lines. With the corresponding 16 estimates of β a t -test showed rejection of the null hypothesis $H_0 : \beta = 1$ with a p -value much less than 0.001.

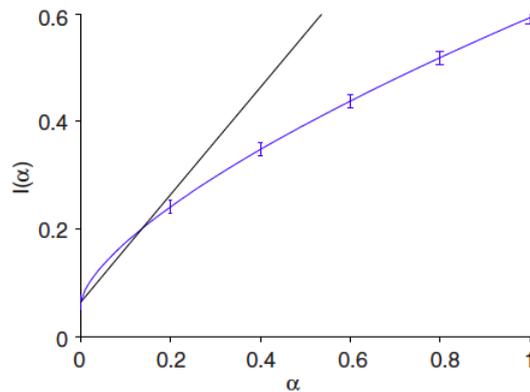


Figure 20: The intensity response I versus mass thickness α of silica as estimated by maximizing the likelihood function (21). The micrographs are STEM projections of 90 nm thick microtomed sample sections, where the individual silica spheres are 20 nm in diameter, compare Figure 19. One single sphere corresponds to $\alpha = 0.22$. The blue line shows the estimate with the estimated power $\beta = 0.69$ in (20), while the black line corresponds to $\beta = 1$. The error bars show the estimated standard deviation of the background noise.

As a check of the model (19) histograms of the pixel intensity in the observed and estimated micrograph images are shown in Figure 21. Generally there is good agreement between the observed and the estimated image intensities. But the histograms indicate that a feature not taken into account of the model is that the edges of the observed spheres are somewhat blurry.

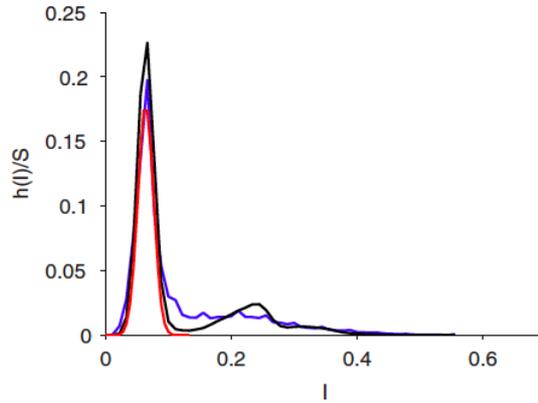


Figure 21: Histogram of the pixel intensity for the observed micrograph (blue) and the estimated micrograph intensity (black). The estimated background noise is also shown (red). Note that the observed micrograph image is blurry at the edges of the spheres, compare the lower right part of Figure 19. This is not taken into account in the model, which can explain why the estimated intensity shows an underestimate just between $I = 0.1$ and 0.2 .