Spatial statistics and image analysis (TMS016/MSA301)

Estimating diffusion coefficient based on raster images

2022-05-11

Diffusion

Diffusion or Brownian motion can be interpreted/modelled

- as a Gaussian random walk with normally distributed increments
- in terms of mean squared displacement
- by using the diffusion equation (Fick's second law).

Data: Fluorescence particles imaged by using a raster scan pattern collected with a confocal laser scanning microscope.

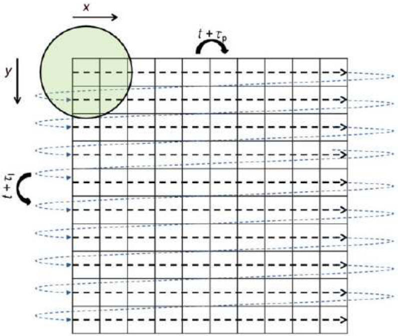
Diffusion coefficient can be estimated by

- raster image correlation spectroscopy (RICS)
- single particle raster image analysis (SPRIA)

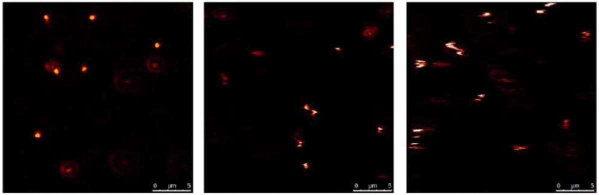
Raster scanning

The scanning is started in the top left corner of the sample and

- the scanning time between adjacent pixels in the x direction is τ_p
- the scanning time between adjacent pixels in the y direction is τ_l
- $\tau_p \ll \tau_l$
- introduces time information within the image.



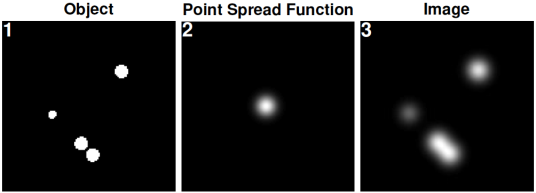
Raster scan images with varying scan rate



- Left: Scanned at rate 8000 Hz, the particles look almost immobile.
- Middle: Scanned at rate 400 Hz, the particles look like sequences of shifted bright line segments.
- Right: Scanned at rate and 100 Hz, the particles are moving significantly between two consecutive lines.

Raster image correlation spectroscopy (RICS): Point spread function

- The intensity of a point fluorescent source will be spread out upon detection due to the diffraction of light.
- The diffraction pattern is described by the point spread function, which is assumed to be a three dimensional Gaussian function with (possibly) different standard deviations in the z direction and the xy plane.



- Spatial correlation will be introduced between adjacent pixels of the image.

Raster image correlation spectroscopy (RICS)

Diffusion coefficient is estimated using the correlation function estimated from the image.

In case of pure diffusion, the theoretical correlation function $G(\xi, \psi)$ for the scanned image corresponding to two points (x, y) and $(x+\xi, y+\psi)$ (ξ and ψ are the spatial increments in number of pixels) is

$$\begin{aligned} G(\xi, \psi) &= \frac{1}{\bar{N}} \left(1 + \frac{4D|\tau_p \xi + \tau_l \psi|}{\omega_0^2}\right)^{-1} \left(1 + \frac{4D|\tau_p \xi + \tau_l \psi|}{\omega_z^2}\right)^{-1/2} \\ &\quad \times \exp\left[-\frac{(S\xi)^2 + (S\psi)^2}{\omega_0^2 + 4D|\tau_p \xi + \tau_l \psi|}\right]. \end{aligned}$$

Remark: This is not really a correlation function but a normalized covariance function.

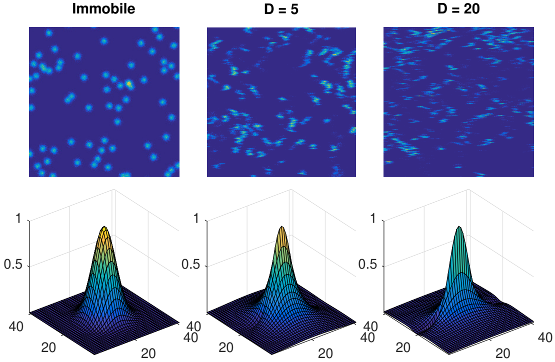
RICS correlation function

$$G(\xi,\psi) = \frac{1}{\bar{N}}\left(1+\frac{4D|\tau_p\xi+\tau_l\psi|}{\omega_0^2}\right)^{-1}\left(1+\frac{4D|\tau_p\xi+\tau_l\psi|}{\omega_z^2}\right)^{-1/2}$$
$$\times \exp\left[-\frac{(S\xi)^2+(S\psi)^2}{\omega_0^2+4D|\tau_p\xi+\tau_l\psi|}\right].$$

where

- $\bar{N}$ is the average number of particles in the observation volume
- D is the diffusion coefficient
- S is the pixel size
- $|\tau_p\xi+\tau_l\psi|$ is the time it takes to move between the points (x,y) and $(x+\xi,y+\psi)$.
- ω_0 and ω_z correspond to the decay rate of the point spread function (standard deviation of a Gaussian distribution) in the lateral and vertical directions, respectively.

Raster image correlation spectroscopy (RICS)



RICS: estimation algorithm

Let us have n images with resolution $M \times M$ pixels.

Let $G_E(\xi, \psi, j)$ be the empirical correlation function relative to a shift of ξ pixels in the x direction and ψ pixels in the y direction, $1 \leq \xi, \psi \leq M$, of the jth image, $j = 1, ..., n$.

Then, we estimate $G_E(\cdot, \cdot, j)$ for all $1 \leq j \leq n$ via the fast Fourier transform algorithm and compute the average of the empirical correlation functions

$$\hat{G}(\xi, \psi) = \frac{1}{n} \sum_{j=1}^{n} G_E(\xi, \psi, j).$$

RICS: estimation algorithm

Estimate the parameter vector θ (including D) by

$$\hat{\theta} = \text{argmin}_\theta \sum_{\xi,\psi} w(\xi,\psi)\left(G(\xi,\psi,\theta) - \hat{G}(\xi,\psi)\right)^2,$$

where the weights $w(\xi,\psi) = \left(\text{Var}(\hat{G}(\xi,\psi))\right)^{-1}$ are computed from the set of independent images.

RICS versus single particle analysis

RICS

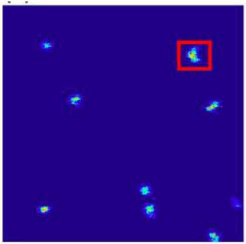
- estimates the diffusion coefficient by averaging over the observed patterns of all particles in several images
- does not give straightforward standard error estimates
- can be sensitive to the choice of the scanning rate.

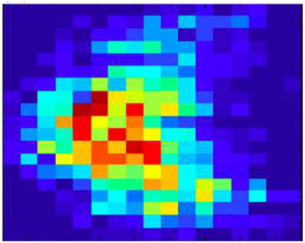
Analyzing each particle separately

- gives us a diffusion coefficient estimate for each particle and a straightforward way to estimate standard errors
- allows us to analyze systems of particle mixtures with varying diffusion coefficients and heterogenous materials with diffusion properties varying with location.

Single particle raster image analysis (SPRIA): extracting the particles

To be able to use the single particle method, individual particles have to be extracted from an image:





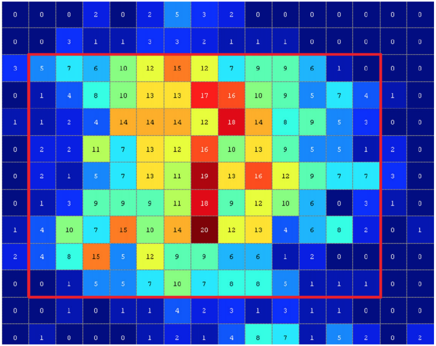
The more red the pixel is, the larger the photon count.

Single particle raster image analysis (SPRIA): extracting the particles

Particles are defined by using two threshold levels T_1 and T_2, where $T_1 > T_2$ as follows:

- First, find the 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.
- The choice of the levels is not critical.

$T_1 = 10$ and $T_2 = 5$:



The rectangle defined by the red lines defines the particle.

Single particle raster image analysis (SPRIA): definition of particles

As shown above, the particle P is defined as the axis-parallel rectangle

$$P = \{(x, y) : a < x < a + L, b < y < b + K\} \tag{1}$$

around the corresponding local maximum of photon counts. Here,

- (a, b) is the position of the top-left pixel in the corner of the rectangle P
- L and K are the lengths of its sides.

The trajectory and the diffusion coefficient D of the particle can be estimated based on the extracted particles.

Single particle raster image analysis (SPRIA): estimating the trajectory and D

Diffusion can be modelled in terms of mean squared replacement. In one dimension,

$$\mathbb{E}[(X(t+\Delta t) - X(t))^2] = 2D\Delta t.$$

If we knew the x coordinate X_k of the particle position at time t_k (the time to scan the line k), $k = 0, \ldots, K$, then

$$\tilde{D} = \frac{1}{2\Delta t K} \sum_{k=1}^{K} (X_k - X_{k-1})^2,$$

where $\Delta t = \tau_l$ is the time needed to scan a line, would be an unbiased estimator for the diffusion coefficient D.

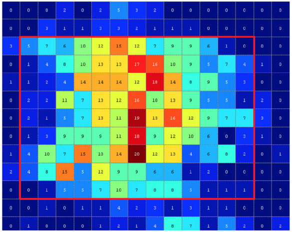
However, we do not know X_k's.

Single particle raster image analysis (SPRIA): estimating the trajectory and D

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_k, where $t_k = t_k(y)$, $k = 0, \ldots, K$, is the time at which we observe the horizontal line at y.

X_k can be estimated as a weighted mean of the x values on the same line, where the weights are the photon counts (centroid method).

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



Single particle raster image analysis (SPRIA): estimating the trajectory and D

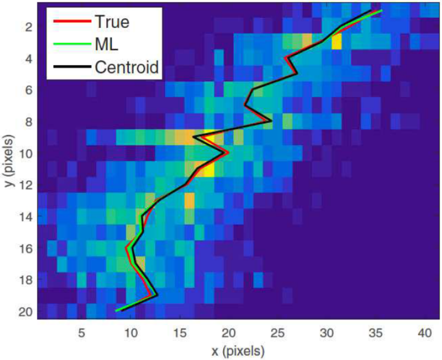
Finally, the diffusion coefficient D can be estimated by

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

where $\Delta t = \tau_l$ is the time between two consecutive line scans.

Estimated trajectory

The true (red) and estimated (black and green) trajectories in a simulated image:



Remark: The green trajectory is estimated by using maximum likelihood as described in Longfils et al. (2017).

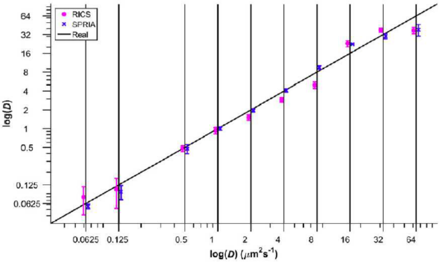
Performance of RICS and SPRIA

A simulation study: Diffusing particles were generated by using a Gaussian random walk (discrete time Brownian motion) of spheres in a box with periodic boundary conditions.

- Spheres with 10 different diameters varying between 0.015μm 1μm were simulated. The larger the particle, the slower the movement and smaller the diffusion coefficient.
- The pixel size was 0.03μm.
- More details of the simulation experiment can be found in the lecture notes by Mats Rudemo.

Simulation results

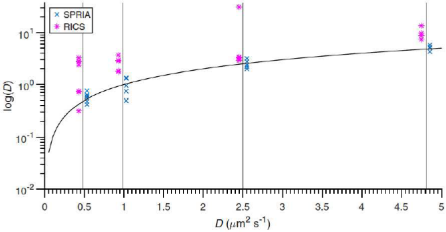
- Vertical black lines correspond to the expected diffusion coefficient according to Stoke-Einstein's equation.
- Blue markers ($\pm$ standard errors) refer to SPRIA and magenta to RICS.
- Standard deviations for the RICS estimates are estimated by bootstrap using $B = 40$ repetitions.



(Logarithmic scales on both axes.)

Performance for real data

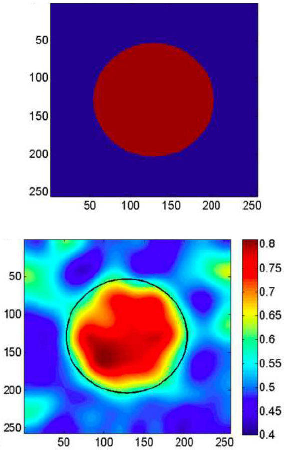
Diffusion coefficient estimates for fluorescent beads of four different sizes: 0.1 μm, 0.175 μm, 0.49 μm, and 1.0μm.



(Logarithmic scale on the y axes.)

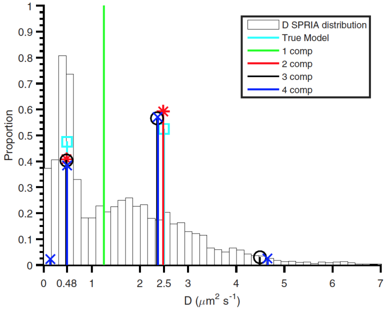
- Vertical black lines correspond to the expected diffusion coefficient according to Stoke-Einstein's equation.
- Blue markers refer to SPRIA and magenta to RICS.

- Set-up: Diffusion coefficient inside the disk is 0.8 $\mu m^2 s^{-1}$ and outside the disk 0.4 $\mu m^2 s^{-1}$. In the simulation, 2142 particles in 300 images were found.
- True diffusion map (top) and the reconstructed map by using the estimated diffusion coefficients and Gaussian smoothing (bottom).



SPRIA: Note on particle mixtures

- We can assume that the distribution of the diffusion coefficient is a mixture distribution.
- If the number of diffusion coefficients (components in the mixture distribution) is known, the diffusion coefficients can be estimated quite well.
- However, if the number of diffusion coefficients is unknown, the problem becomes more complicated.



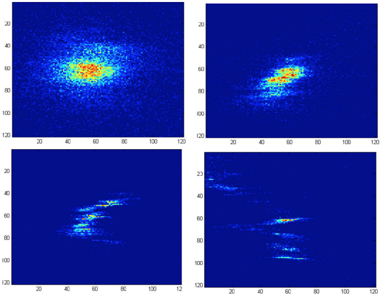
The true model has two components.

When can we use SPRIA?

1. The density of particles should not be too large since we have to be able to identify individual particles.
2. Sampling time between the lines should be long enough so that the adjacent horizontal particle lines differ from each other, but not too long so that the particles do not split into several parts.

Illustration (assumption 2)

- One particle would be identified in the images on the top row.
- Two particles would be identified on the bottom left.
- Five (or more, depending on the thresholds) identified on the bottom right.



Practical issues

Project parts 1 and 2: The deadline is today!

Project part 3: presentations

- Schedule is in Canvas.
- Each group has appr. 15 minutes to present their project (research question, data, methodology, results).
- The main purpose is to get feedback from and to give feedback to the others.
- If you have any questions, do not hesitate to contact Konstantinos or me.

Exam: both more practical and more theoretical questions may appear.