Applications of point processes/image analysis to nerve fiber data

Konstantinos Konstantinou

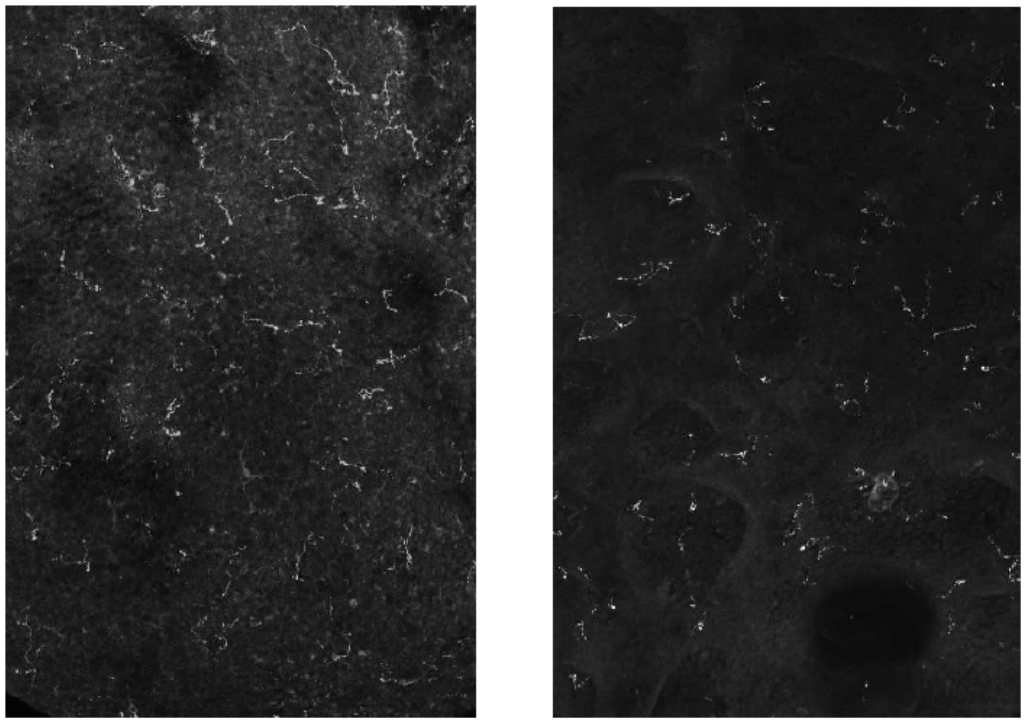
Mathematical sciences
Chalmers University of Technology and University of Gothenburg
Gothenburg, Sweden

What are epidermal nerve fibers (ENFs) ?

- Thin sensory fibers in the epidermis.
- Responsible for transferring signals to the brain (e.g heat and pain).
- Their existence was theorized for many years before it was conclusively established in 1993 in a series of confocal microscopy studies.
- Diagnostic tool for the assessment of the degree of peripheral neuropathy, a condition associated with poor nerve functionality.
- Peripheral neuropathy:
  1. Decrease of the ENF counts.
  2. Changes to the structure of the ENFs.

Confocal microscope images

- Confocal microscope images of the skin of healthy and patients with neuropathy
- Spatial distribution of nerve fibers (white)



Left: A non-diabetic subject

Right: A subject with moderate diabetic neuropathy

Objectives of the study

Goals:

- Identify and extract the locations of some points of interest(i.e termination points of the nerve, branching points of the nerves) from the images.
- Reconstruct the spatial structure of ENFs using statistics obtained from confocal microscope images for healthy subjects.

Image processing

Some image processing steps are required to obtain a skeletonized image of the nerve fibers:

1. Noise reduction : A 3×3 Wiener filter was used.
2. Background trend : Morphological operations were used in this step.
3. Obtain a more uniform nerve structure : A 3×3 max filter was used.

Image processing

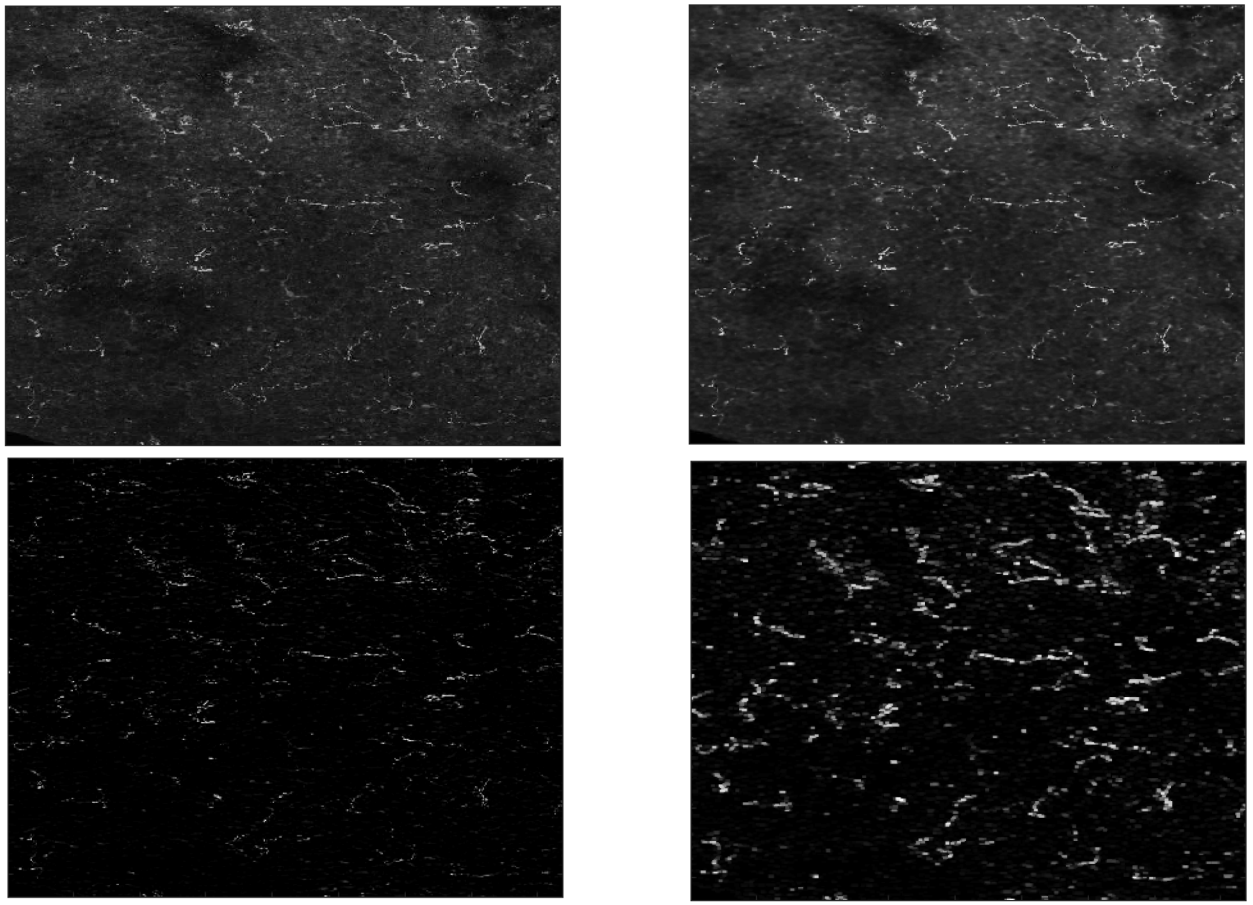


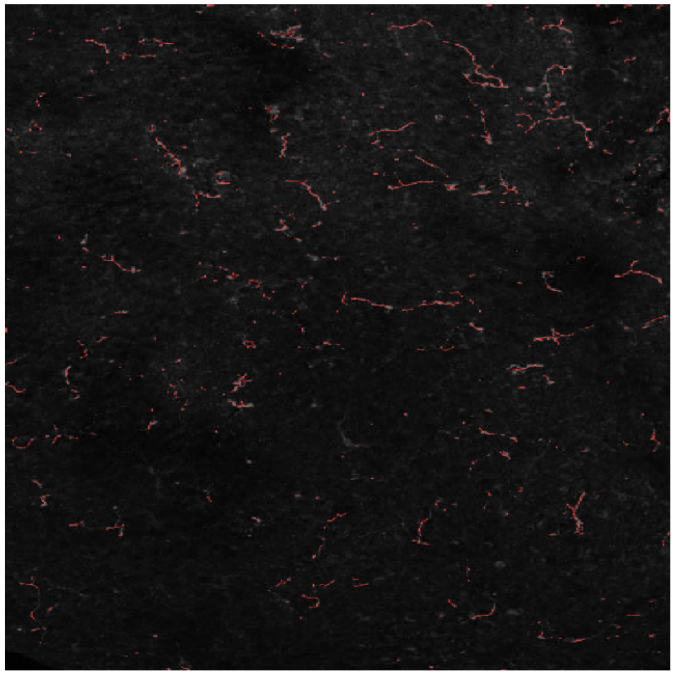
Figure: Original image(top right), smoothed image(top left), the image after removing the background trend(bottom left) and the image after the max filter is applied

Image segmentation

- At this stage, the nerve trees (white) are easier to be separated from the other parts of the epidermis that are not of interest to us.
- To segment the image into two classes (nerve and not nerve) we need to define a threshold t.
- The threshold selection method proposed by Otsu was used.

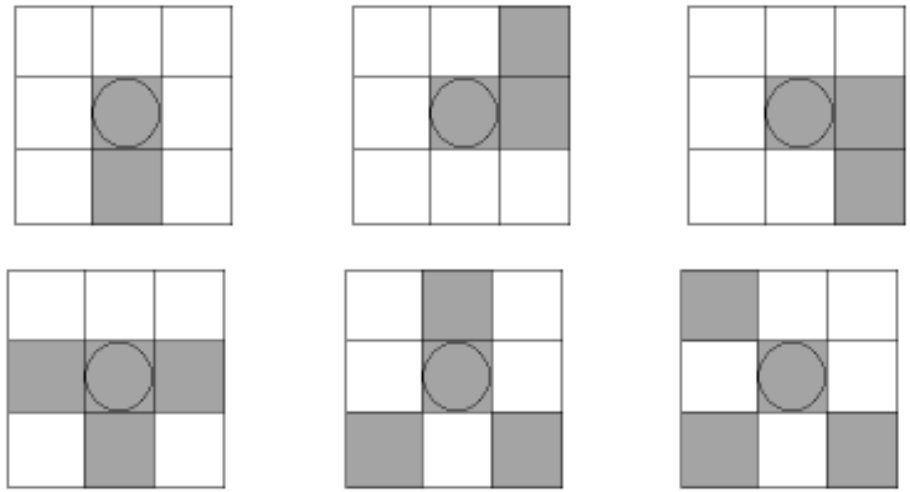
Detecting the nerves from the image

- The skeletonized image superimposed on the original image.
- Nerve fibers are detected from the images. We now need to identify points of interest.



Identifying points of interest

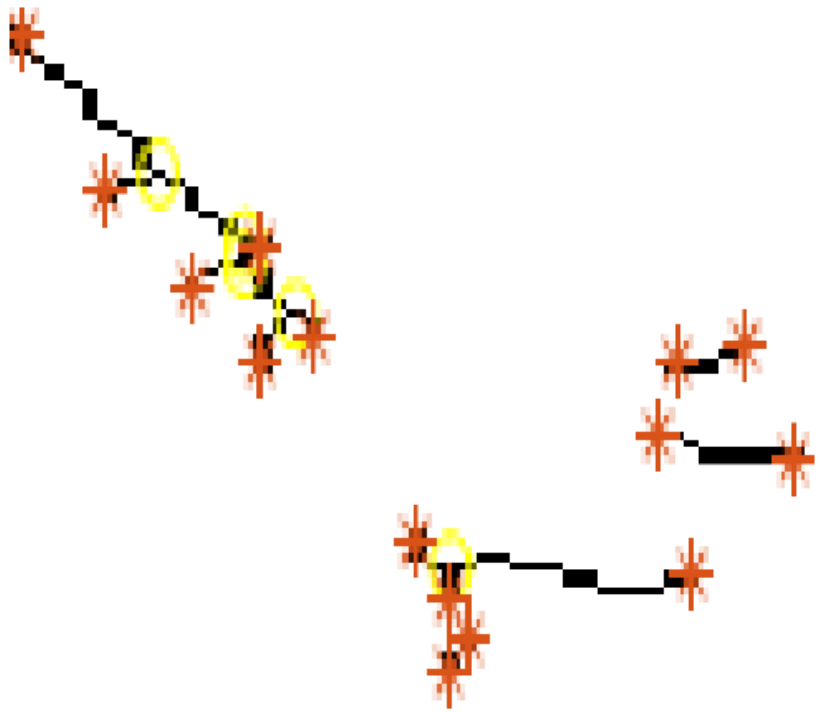
- Need to find all points where branches intersect and all points where branches end.
- Need to define all possible 3×3 neighbourhoods representing end points (20 such neighborhoods) or intersection points (60 such neighborhoods).



Top: Example of three neighborhoods of an endpoint
Bottom: Example of three neighborhoods of a branching point.

Identifying points of interest

- A hit-and-miss transform operation is applied to identify branching and end points.
- Skeleton of the nerve trees (black) with branching points (yellow circles) and end points (red marks)



Obtain mean statistics from the points

Statistics considered are:

1. Mean numbers of segments at each branching points $n_0 = 3.07$
2. Mean angles between the segments connected $a_0 = 116.7$
3. Mean branch lengths $d_0 = 13.7$

Reconstruct the spatial structure of the epidermal nerve fibers using the obtained statistics.

Reconstruction

- Idea: Use Markov chain Monte Carlo (MCMC) methods to sample from a given distribution π.
- Unfortunately, in our application the probability distribution π is unknown.
- However, we can use a Boltzmann distribution that puts most of its probability mass on states that have some desired properties.

Boltzmann distribution

The **Boltzmann distribution** $\pi_{f,T}$ on a finite set S, with an **energy function** $f : S \to \mathbb{R}$ and a **temperature parameter** $T > 0$, is the probability distribution on S which to each element s assigns probability

$$\pi_{f,T}(s) = \frac{1}{Z_{f,T}} exp\left(\frac{-f(s)}{T}\right) \tag{1}$$

where

$$Z_{f,T} = \sum_{s \in S} exp\left(\frac{-f(s)}{T}\right) \tag{2}$$

is a normalizing constant to ensure that $\pi_{f,T}$ is a probability function with total probability of one.

Energy function

- A simple energy function that favours configurations that are similar to the calculated mean statistics is used.
- We define our energy function as

$$f(s) = f_1(s) + f_2(s) + f_3(s) \tag{3}$$

where

$$f_1(s) = c_1 \sum_{\substack{u,v \\ u \sim v}} (d_{uv} - \hat{d}_0)^2 e_{uv} \tag{4}$$

$$f_2(s) = c_2 \sum_{u} (n_u - n_0)^2 \tag{5}$$

$$f_3(s) = c_3 \sum_{u} \sum_{a_u} (a_u - a_0)^2 \tag{6}$$

Energy function

- c_1, c_2, c_3 are constants weighting each part of the energy function.
- u, v are branching points
- d_{uv} is the distance between the points u, v
- $u \sim v$ denotes that u and v are neighbors.
- a_0, d_0 and n_0 denote the mean statistics obtained for the angles, the mean branch length and number of intersections per branching point.
- a_u denotes the angles less than 180 degrees between branches connected to the branching point
- n_u is the number of connected branches for the branching point u
- e_{uv} gives 1 if the branching points u and v are connected, and 0 otherwise.

Simulation

- The 3D structure of the nerve fibers is modelled as a random graph $G = (V, E)$, where V denotes the set of nodes v_j and E denotes the set of undirected edges between nodes.
- An edge $< u, v > \in E$ if and only if there is a branch connecting the two nodes.
- For this graph we assign a Boltzmann distribution with the energy function defined above for the states g of the graph.
- Our simulations are performed in a box of size $320 \times 432 \times 50$ (periodic boundary conditions).
- Our starting configuration of the nodes were constructed by taking the starting nodes to be completely randomly distributed points in the box and no connections between the nodes i.e $E_0 = \varnothing$.
- Our algorithm creates a sequence of graphs G_n where at each iteration the graph is updated using an MCMC algorithm.

Algorithm: Add/remove connections(Step 1)

Step 1: For $n = 1$ to N_1 do

1. We randomly choose a pair of nodes $u, v \in V_n$ that are neighbors.
2. If $< u, v > \in E_n$ we first remove it.
3. To obtain E_{n+1} we either add $< u, v >$ or leave the set of edges unchanged according to the conditional probability$\pi_{f,T}$ given V_n and all other edges of E_n.
4. The probability to change from state g to state g' is given by

$$P_{g,g'} = \frac{1}{1 + exp(-\frac{(f(g)-f(g')}{T})} \tag{7}$$

5. Accept this change if $u < P_{g,g'}$, where $u \sim U(0, 1)$.

end do

Algorithm: Move points (Step 2)

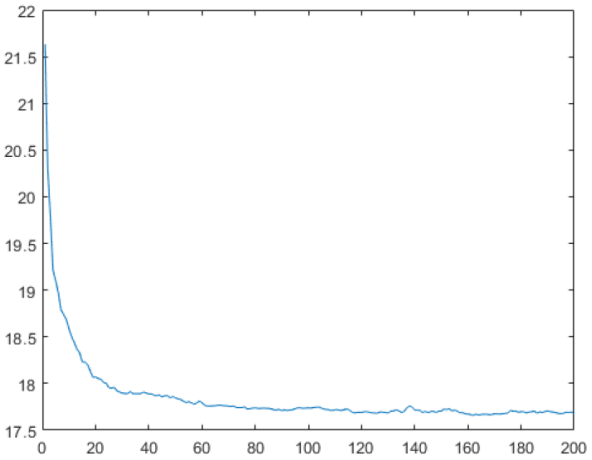
Step 2: For $n = 1$ to N_2 do

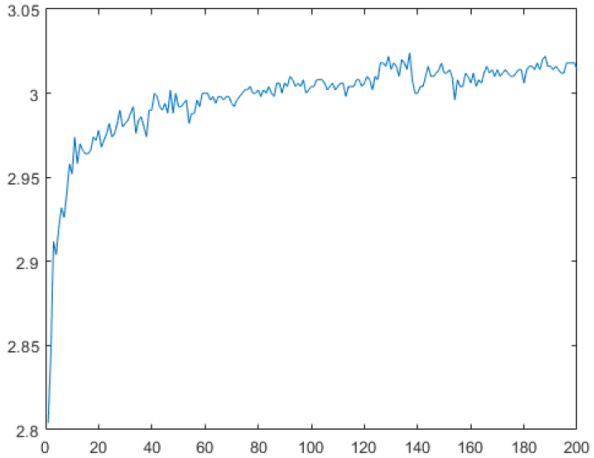
1. We randomly pick a node $u \in V_n$.
2. Sample a random movement ΔR from a uniform distribution in a $3D$ sphere with radius 1.
3. To obtain V_{n+1} we either move node u by ΔR or leave the set of nodes unchanged according to the conditional probability $\pi_{f,T}$ given all other nodes of V_n and the edges of E_{n+1}.
4. The probability to change from state g to state g' is given by equation (7)
5. Accept the translation of node u by ΔR if $u < P_{g,g'}$, where $u \sim U(0, 1)$.

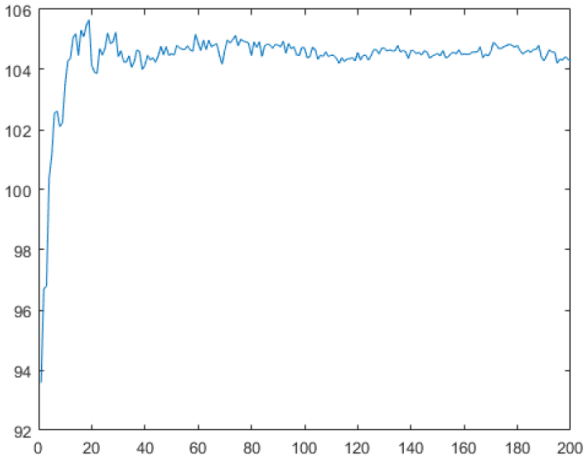
end do

Results

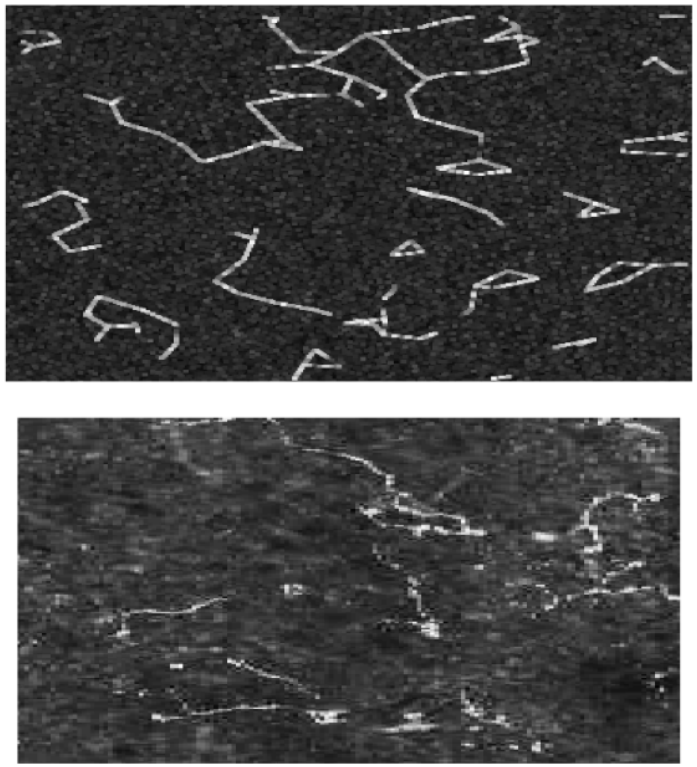
- Mean statistics of the simulated structure mimic converge to the desired mean statistics.





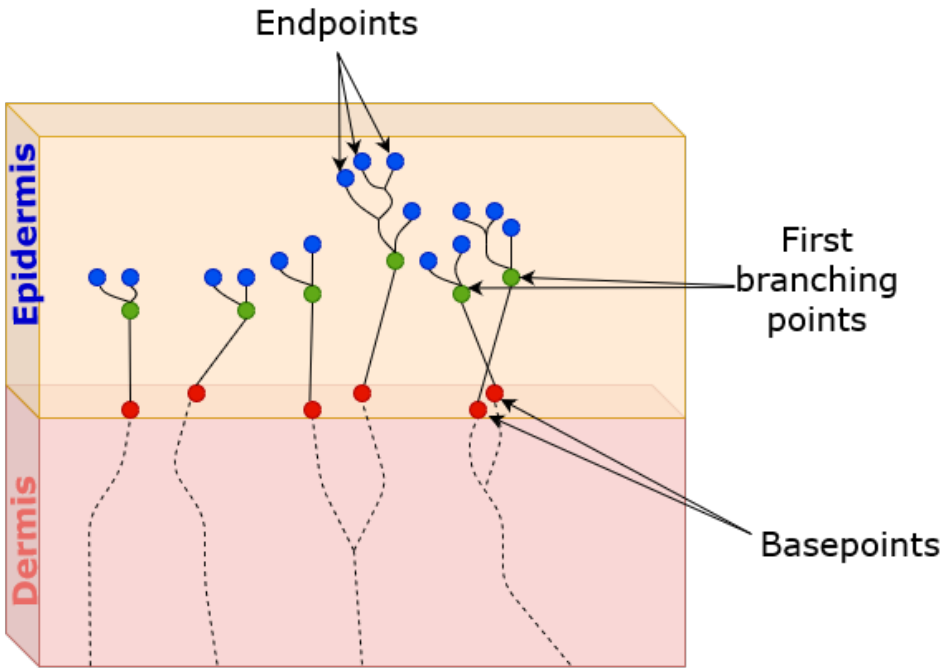


Simulated micrograph



Top: Projection of a section of the simulated image in the xy plane
Bottom: The right top section of the confocal microscope image.

Point process models for ENFs.



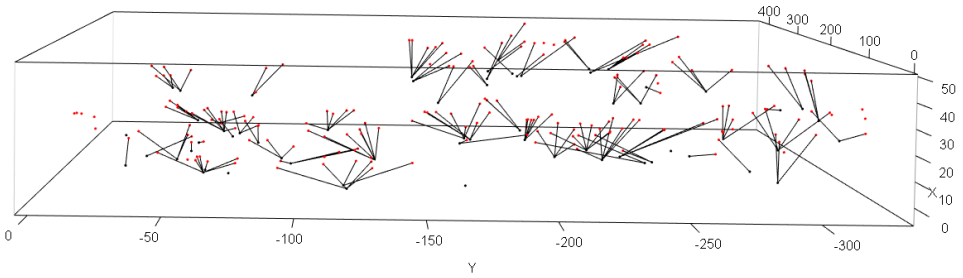
Basepoints: The locations where the nerves enter the epidermis(red)

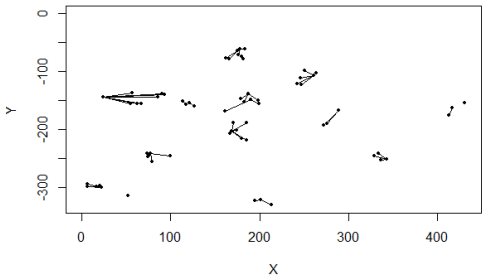
Branching points: The locations where the nerves branch(green)

End points: The locations where the nerves terminate(blue) .

ENF dataset(cont.)

- Hierarchical structure
  - Disease Groups
  - Subjects
  - Body parts (Only data from foot are considered)
  - Samples
- Point patterns in observation window $= 320 \times 432 \times z$ where $z \in (20,50)$ μm.





Objectives

Aims:

- Construct a 2D point process model able to capture the spatial structure of the endpoints.
- Extend the model in 3D.
- Compare the ENFs structure of healthy controls and subjects with mild diabetic neuropathy.

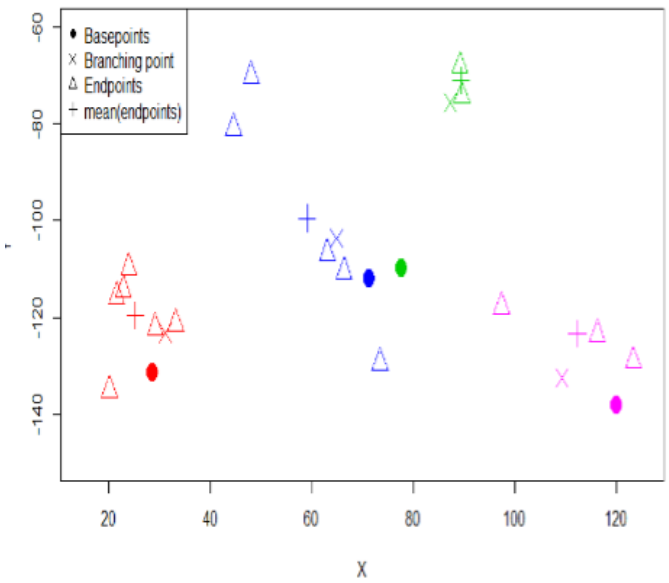
Models for $X = X_p$(conditioned on Base points)

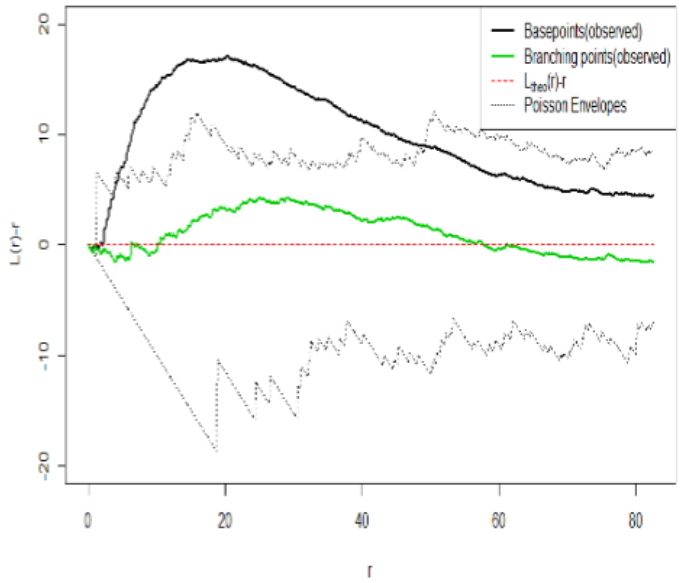
- Non Orphan Cluster Model (NOC)
  - Tree size $\sim$ Jonquiere(δ, γ)
  - Branch length $\sim$ Gamma(α, β)
  - Angles $\sim$ Von-Mises(μ, κ)

  NOC model (1) favours directions towards open space.
  I.e opposite direction from the closest other base point.
- Uniform Cluster Centre (UCC)
  - Tree size $\sim$ Negative Binomial(k, p)
  - Branch length $\sim$ Gamma(α, β)
  - Mean direction $\mu \sim$ Uniform$(0, 2\pi)$
  - Angles $\mid \mu \sim$ Von-Mises(μ, κ)

  UCC model (2) have no preferred direction.

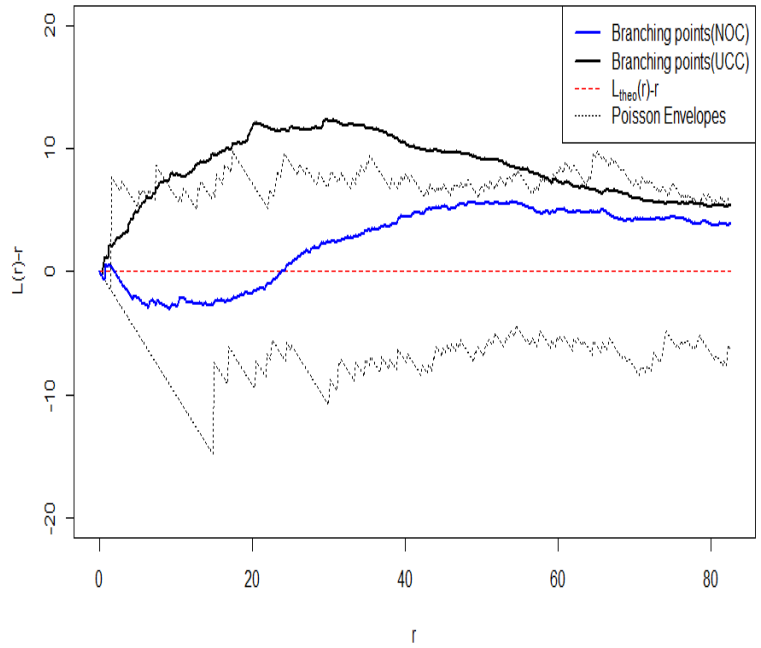
First branching points





- First branching points are better cluster centres for endpoints.
- Basepoints (clustered)
- Branchpoints (CSR)

Branching points (NOC vs UCC)



- A Non-Orhan Cluster(NOC)-like model (1) favours directions towards open space.
  I.e opposite direction from the closest other base point.
- Uniform Cluster Centre(UCC)-like model (2) have no preferred direction.

Models for $X = X_p$(Using the branching points)

- Neyman-Scott point processes (Thomas, Matern)
- Two step NOC-like model (3):
  1. Branchpoints | basepoints
     - $L_1 \sim \Gamma(\alpha_1, \beta_1)$
     - $\phi_1 \sim \mathit{VonMises}(\mu, \kappa)$, where μ is known.
  2. Endpoints | branchpoints.
     - $L_2 \sim \Gamma(\alpha_2, \beta_2)$
     - $\phi_2 \sim \mathit{Uniform}(0, 2\pi)$
     - $s \sim \mathit{NB}(r, p)$

Summary function for a subject.

- First we compute a summary function for every sample
- Then the subject wise summary function can be computed as

$$\bar{K}_i(r) = \sum_{j=1}^{m_i} w_{ij} \hat{K}_{ij}(r)$$

  where $\hat{K}_{ij}(r)$ is the estimate for sample j of subject i
- $w_{ij} = \frac{n_{ij}}{\sum_{j=1} n_{ij}}$

Goodness of fit(subject)

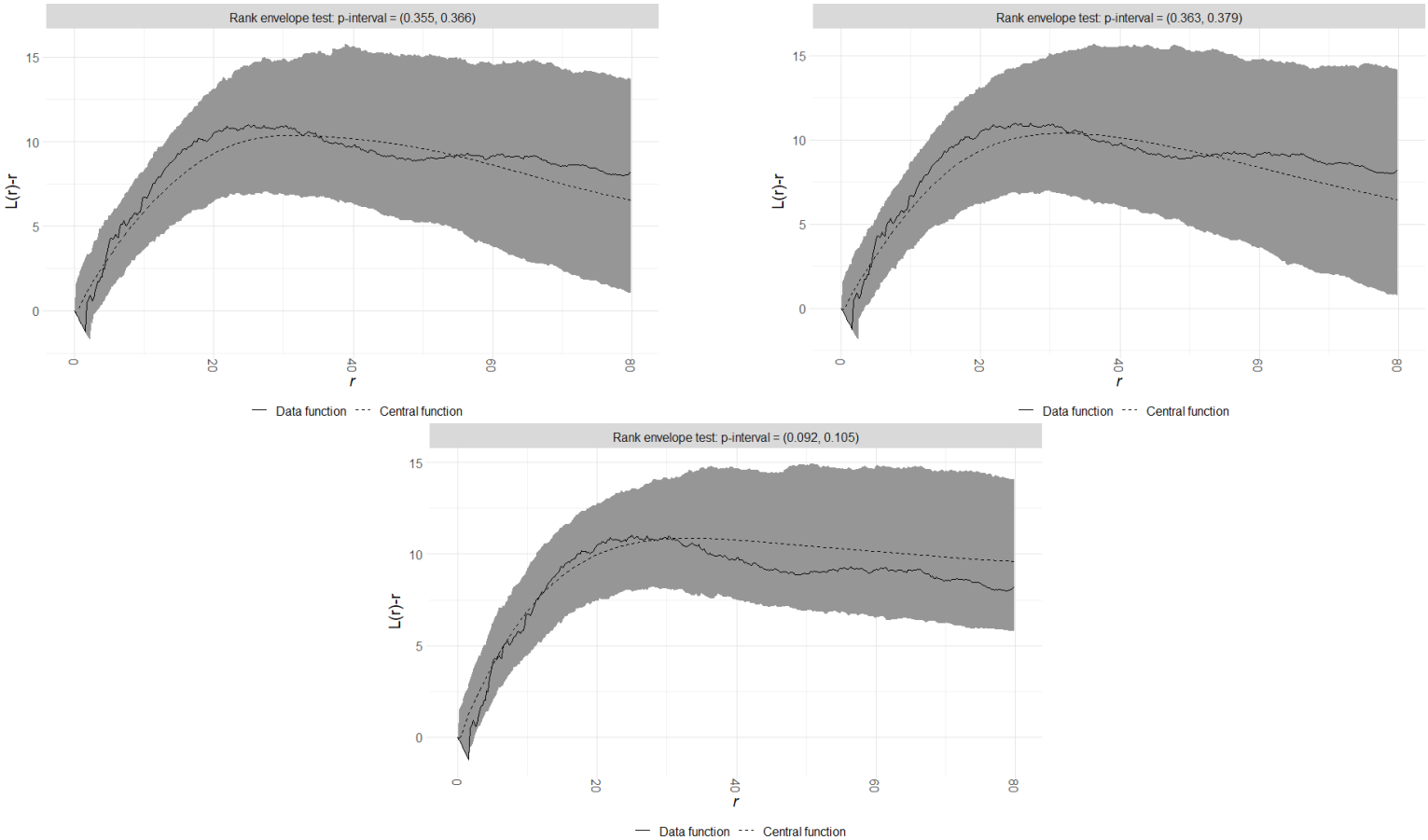


Figure: Subject wise pooled two dimensional $L(r) - r$ functions with global envelopes for the end points of a healthy subject of the two Neyman-Scott models (top) and the NOC-like model(bottom).

Pairwise interaction Markov model for $X = X_p \times X_z$

- Pairwise interaction Markov field Model for $X_z|X_p$ (4) defined by the conditional density of z_i given all other z_j and X_p given by

$$f(z_i|(x_k, y_k)_{k=1}^n, (z_j)_{j\neq i}) \propto \gamma^{s_i} 1(\|(x_i, y_i, z_i) - (x_j, y_j, z_j)\| > h \quad \text{for } j \neq i) \quad (8)$$

where,

1. s_i is the number of further points of the process in the cylinder $B(x_i, y_i, z_i; r, t)$
2. $h > 0$ is a hard core parameter.
3. γ is the interaction parameter.

Parameter estimation

Maximize the log pseudo likelihood given by

$$\begin{aligned} pl(\gamma, h, r, t) &= \sum_{i=1}^{n} log(f(z_i|(x_j, y_j)_{j=1}^{n}, (z_j)_{j\neq i}) \\ &= \sum_{i=1}^{n} log(\gamma^{s_i} 1(\|(x_i, y_i, z_i) - (x_j, y_j, z_j)\| > h \quad \text{for } j \neq i)/c_i) \end{aligned} \tag{9}$$

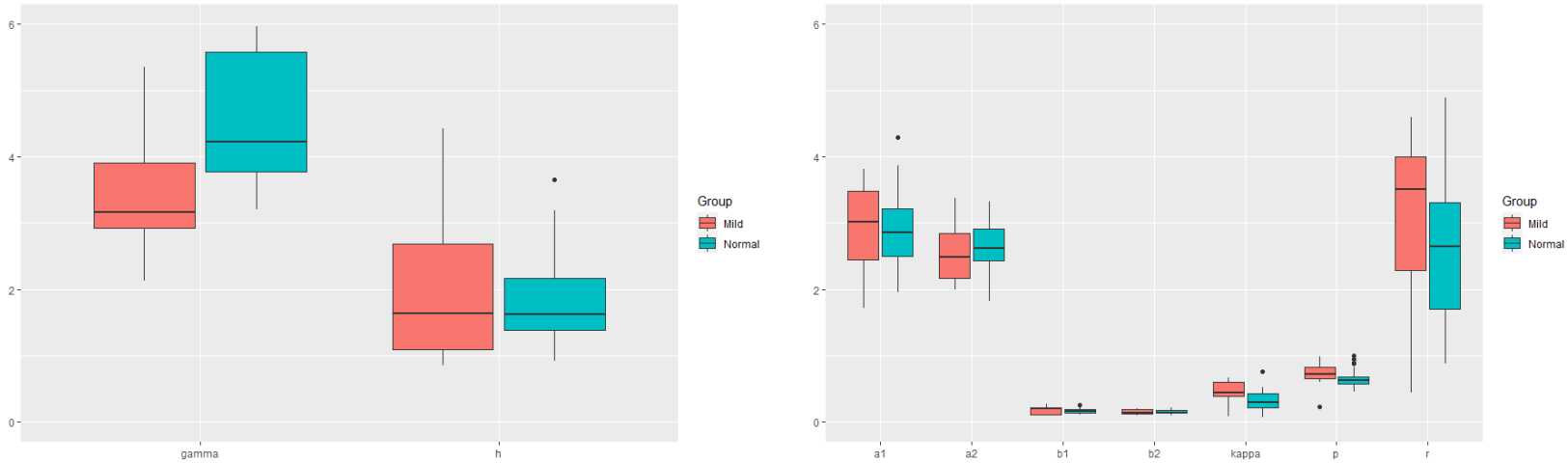
where c_i is the normalizing constant

$$\begin{aligned} c_i = \sum_{k=0}^{n-1} \gamma^k \int_{W_z} &(\|(x_i, y_i, z_i) - (x_j, y_j, z_j)\| > h \quad \text{for } j \neq i) \\ &\times 1(\sum_{j\neq i} 1((x_j, y_j, z_j) \in B(x_i, y_i, z_i; r, t)) = k) dz \end{aligned} \tag{10}$$

Parameter estimates

- $\hat{h} = (n-1)d_{min}/n$, where d_{min} is the minimum distance between two points of the process.
- $\hat{\gamma} = \underset{\gamma}{\text{argmax}}\ \text{pl}(\gamma; \hat{h}, r_g, t_g)$ over a grid of values for the parameters r_g and t_g.
- $\hat{t}$ and $\hat{r}$ are the corresponding grid values.

Group comparison



- Attraction between pairs of endpoints(stronger in the healthy group).
- Differences in the concentration parameter of the angular distribution between the two groups.
- More points per cluster in the healthy group than the mild group.

Simulation

- MCMC algorithm where $\mid X_p \mid$ is fixed
- Acceptance probability $\alpha = \frac{f(z_i^{new} | X_p, (z_j)_{j \neq i})}{f(z_i | X_p, (z_j)_{j \neq i})}$

Algorithm 1:

Result: The point pattern X= $X_p \times X_z$
Simulate X_p using a model for the planar process;
Simulate $Z_1, \ldots, Z_n \sim Uniform(\min(W_z), \max(W_z))$;
Set $X = X_p \times Z$;
for $i = 1, \ldots, M$ **do**
  for $j = 1, \ldots, n$ **do**
    Propose Z_j^* using a Uniform proposal;
    Calculate the acceptance probability α;
    Draw $U \sim Uniform(0, 1)$;
    if $U < \alpha$ **then**
      Set $Z_j = Z_j^*$;
      Set $X = X_p \times Z$;
    end
  end
end

Cylindrical K function

- Since the patterns are anisotropic we use directional summary statistics to evaluate the fit.
- An unbiased estimate for the cylindrical K function (the structuring element is a cylinder) is given by

$$K^u_{cyl}(r) = \frac{1}{\hat{\lambda}^2} \sum_{x_1, x_2 \in W}^{\neq} w(x_1, x_2) 1[x_1 - x_2 \in B^u(r, w)], \qquad r > 0 \quad (11)$$

- $\hat{\lambda}^2 = \frac{n(n-1)}{|W|^2}$
- $w(x_1, x_2) = \frac{1}{|W \cap W_{x_2 - x_1}|}$ is the translation edge correction with $W_{x_2 - x_1}$ denoting the translation of the d-dimensional window W by $x_2 - x_1$
- $B^u(r, w)$ denotes the shape created by the intersection of a cylinder with fixed half-width w and direction u with spheres of radius $r > 0$.

Model evaluation

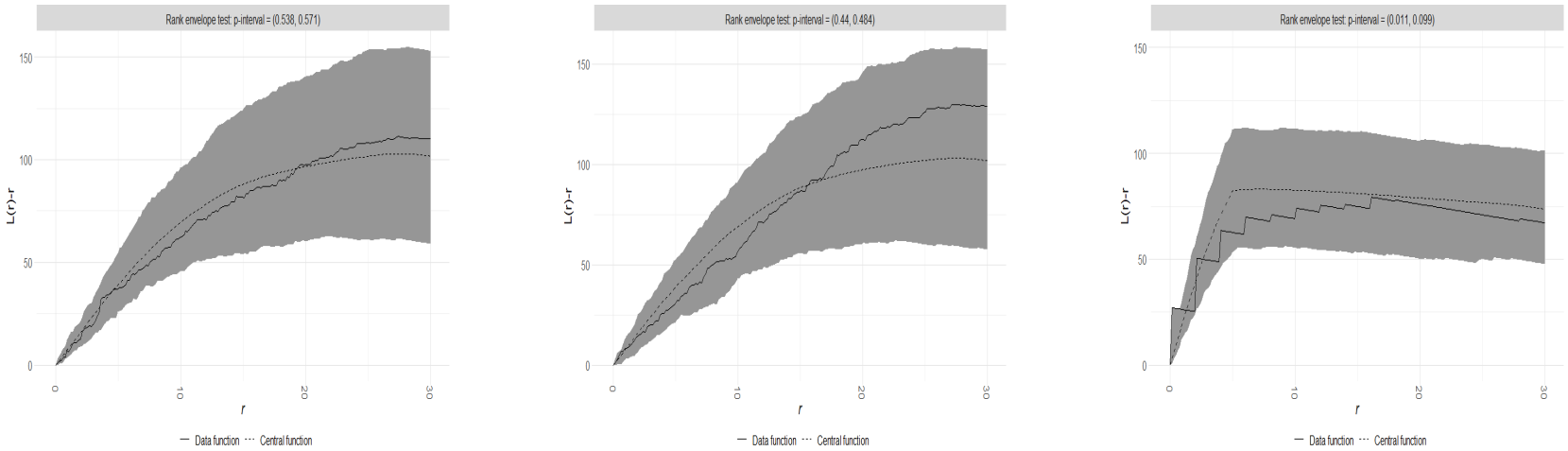


Figure: Subject wise pooled cylindrical three dimensional $L(r) - r$ functions with 95% global envelopes for the end points of a healthy subject of the NOC-like model. The left column shows the cylindrical $L(r) - r$ functions directed towards the x-axis, the middle column the results for the y-axis and the right column the respective results for the z-axis.

Summary

1. Image analysis
   - Image processing/segmentation methods to detect points of interest from an image.
   - Used those points to obtain some mean statistics.
   - Proposed an algorithm to reconstruct the structure of the fibers.
2. Point processes
   - Planar point process models for the ENFs endpoints.
   - Subject-wise K function.
   - Pairwise interaction Markov model(3D).
   - Directional K functions(cylindrical).

References

[1] V. Olsbo, M. Myllymäki, L. A. Waller, and A. Särkkä, "Development and evaluation of spatial point process models for epidermal nerve fibers," *Mathematical biosciences*, vol. 243, no. 2, pp. 178–189, 2013.

[2] C. Andersson, P. Guttorp, and A. Särkkä, "Discovering early diabetic neuropathy from epidermal nerve fiber patterns," *Statistics in medicine*, vol. 35, no. 24, pp. 4427–4442, 2016.

[3] K. Konstantinou and A. Särkkä, "3d modelling of the epidermal nerve fibers," 2021.

[4] A. D. Christoffersen, J. Møller, and H. S. Christensen, "Modelling columnarity of pyramidal cells in the human cerebral cortex," *arXiv preprint arXiv:1908.05065*, 2019.