Exploring biological shape analysis through topology, geometry and statistics

Ph. D. summer school: Biomedical image analysis, 2024/08/13

Jon Sporring,
Department of Computer Science

UNIVERSITY OF COPENHAGEN

Imaging research and pipeline

Research loop

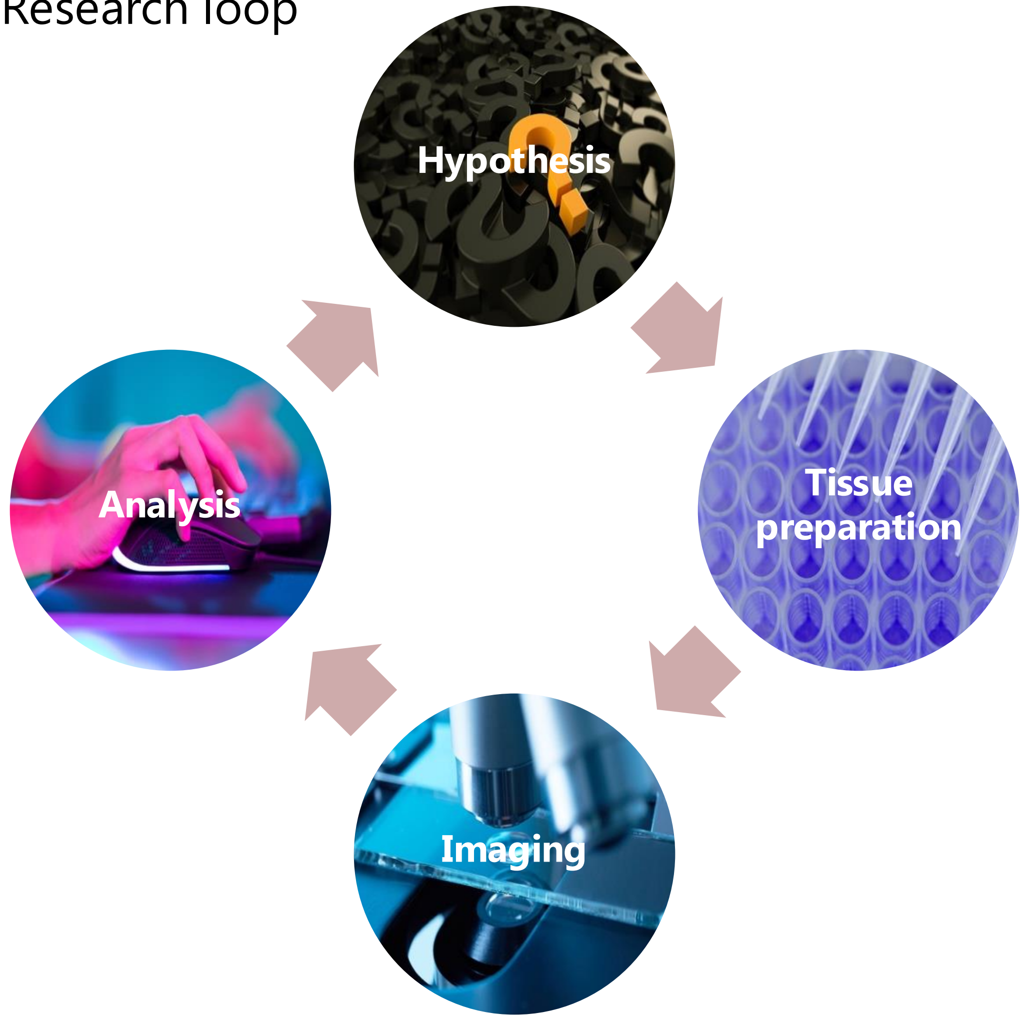
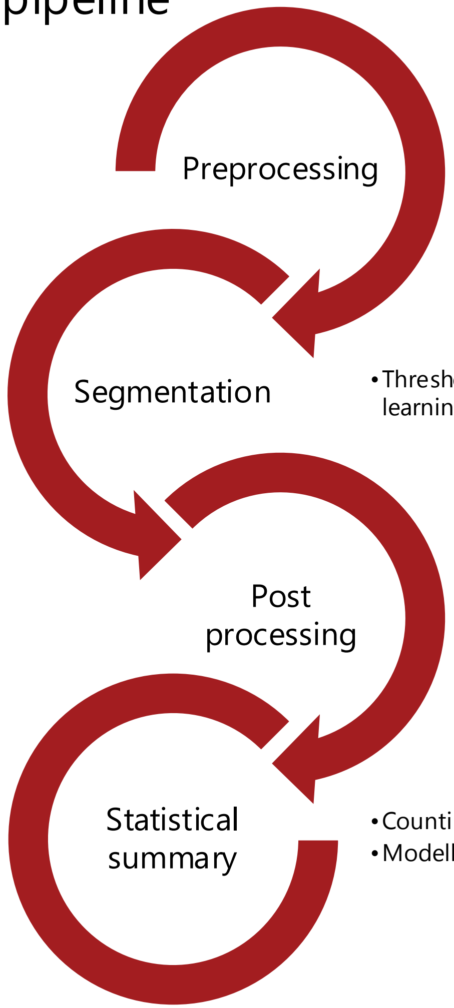


Image analysis pipeline



- Bias correction
- Cropping

- Thresholding/deep learning

- Mathematical morphology
- Shape selection

- Counting
- Modelling

AI is pretty good at segmenting stuff, what's next?

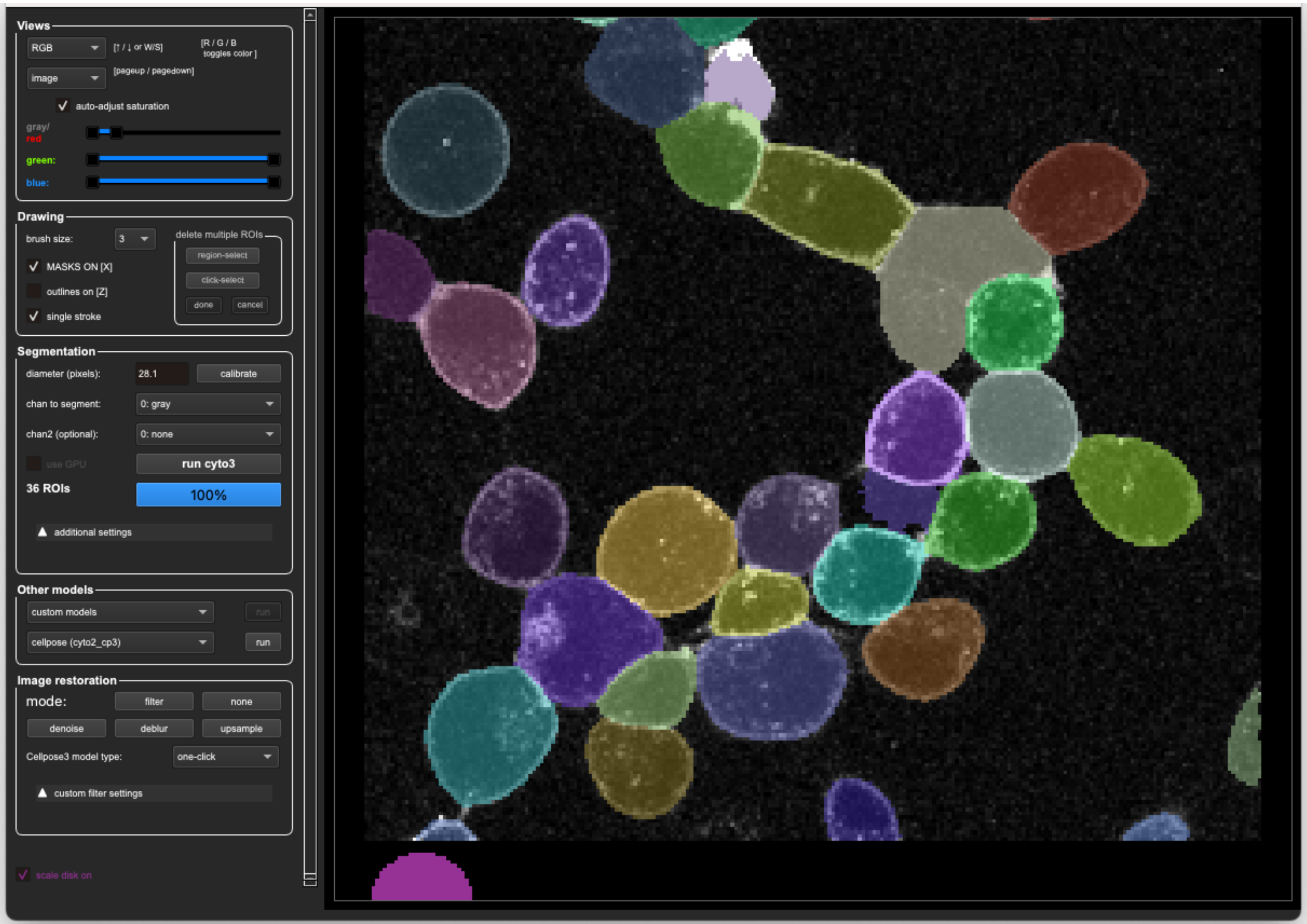
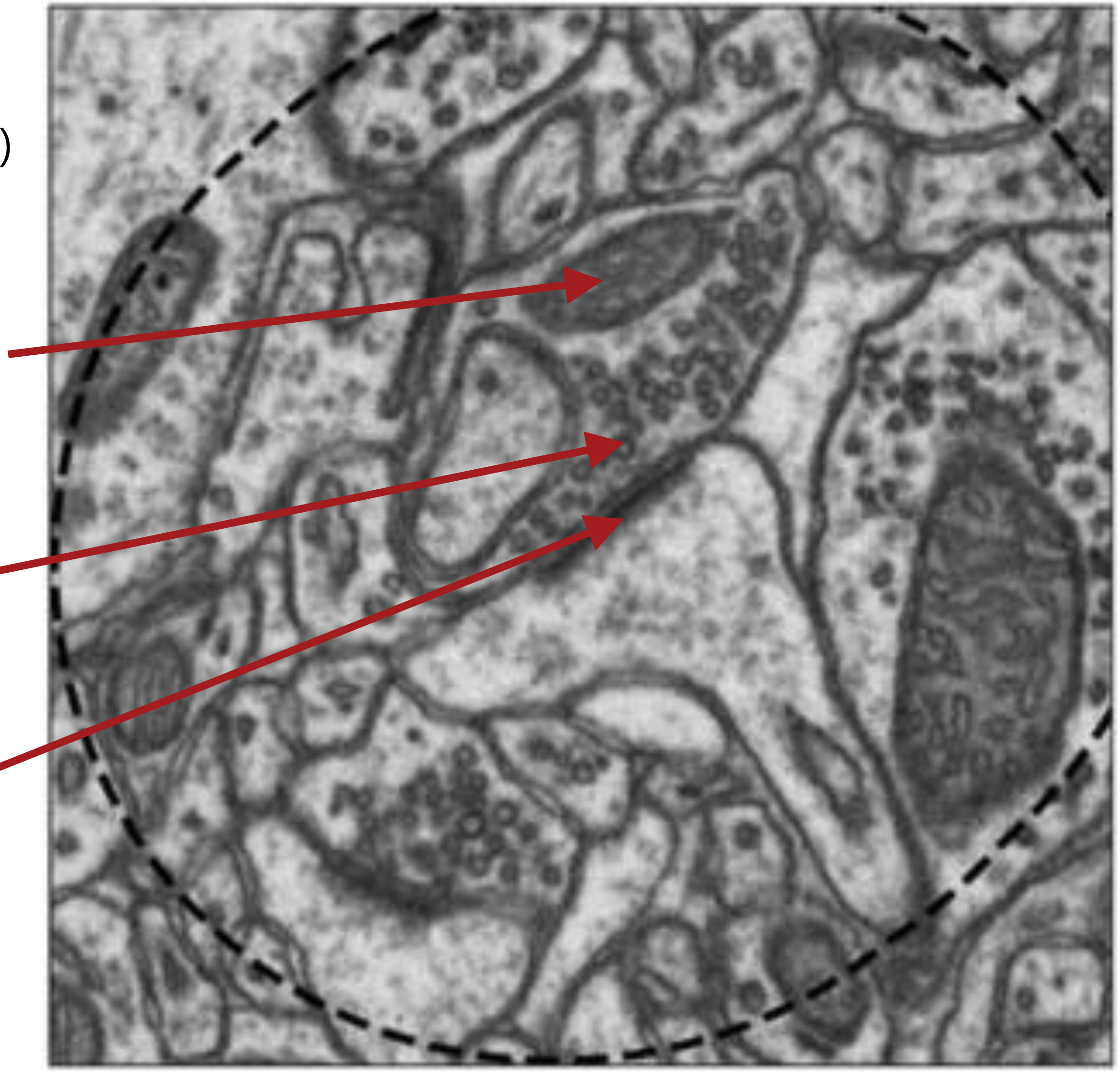


Image courtesy: Karen Martinez & Gabriella von Scheel von Rosing; AI: http://www.cellpose.org/

What to do next: Shape analysis

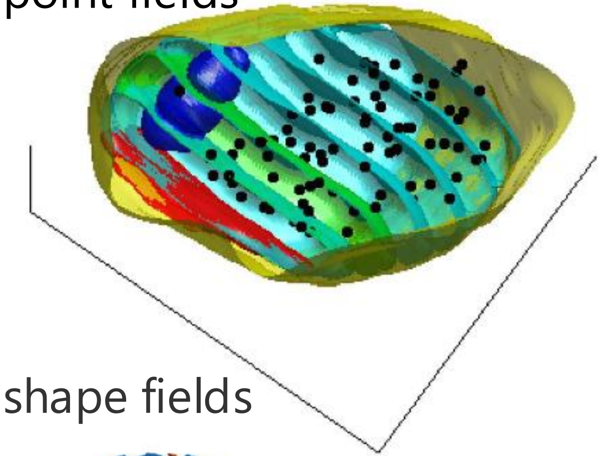
Focused ion-beam scanning electron microsopy (FIB-SEM)

Voxel size: (5 nm)^3

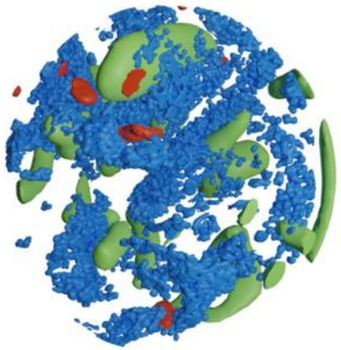


Graham Knott and Marco Cantoni. Electron microscopy dataset. https://cvlab.epfl.ch/data/ data-em/

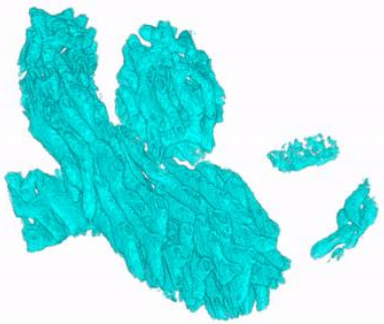
Spatial point fields



Spatial shape fields

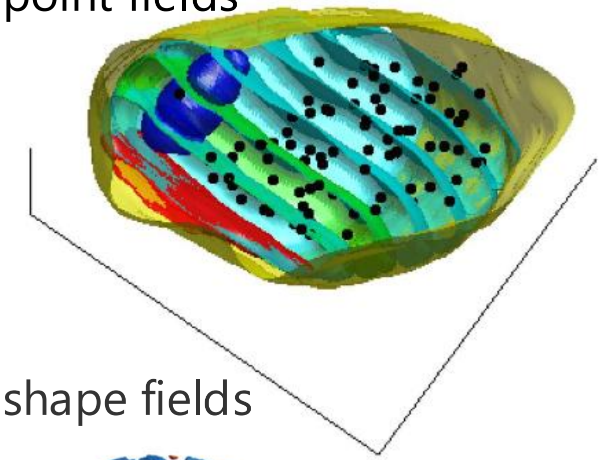


Topological data analysis

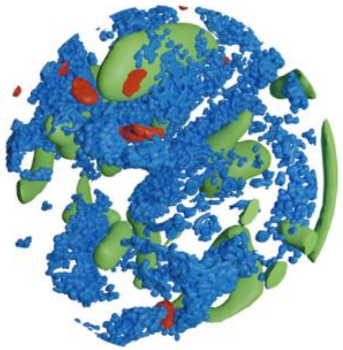


Literature

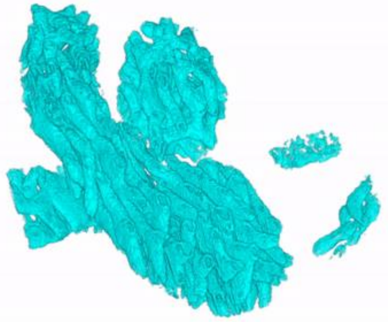
Spatial point fields



Spatial shape fields

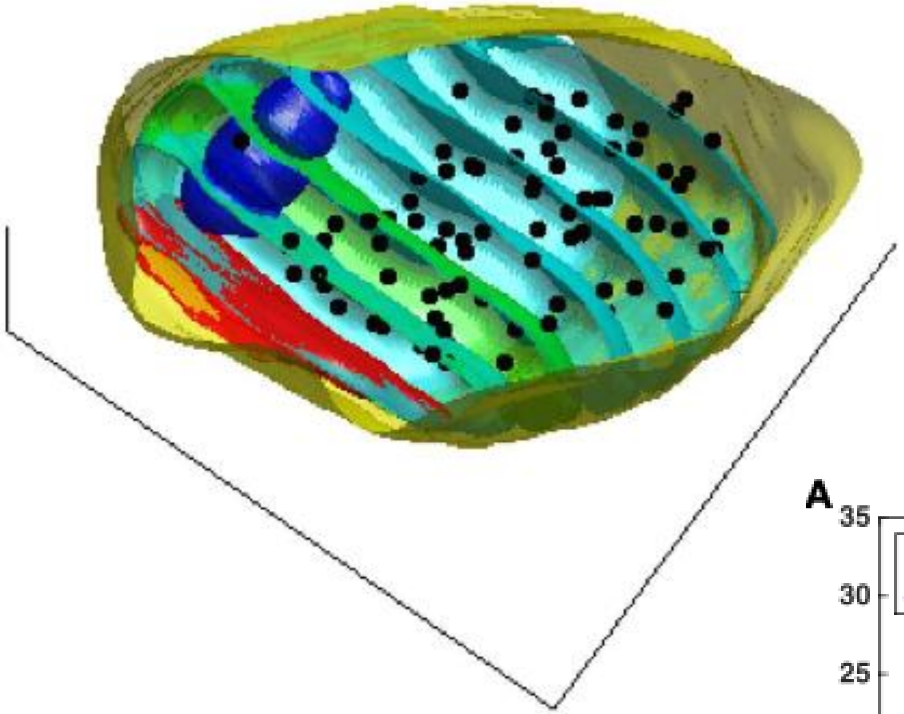


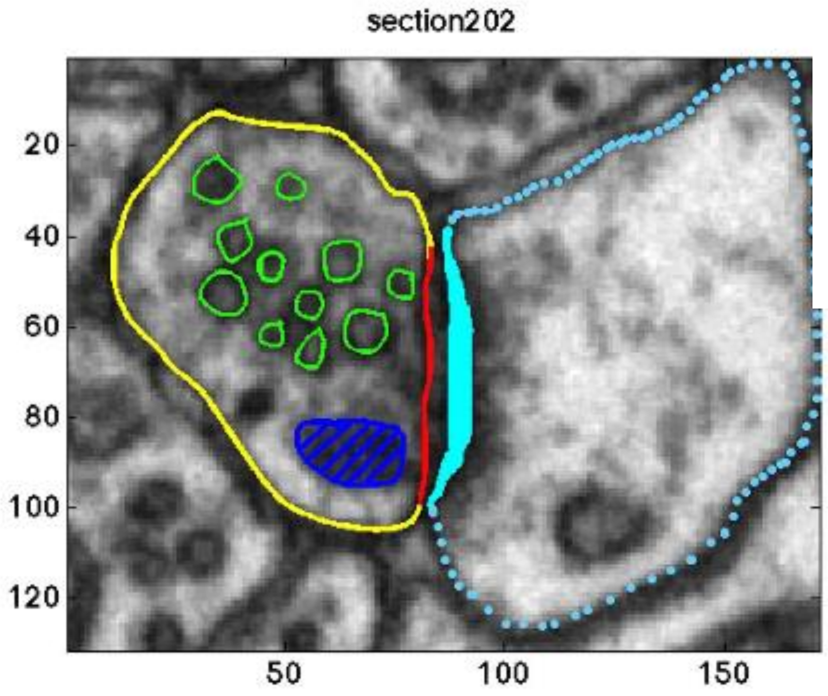
Topological data analysis

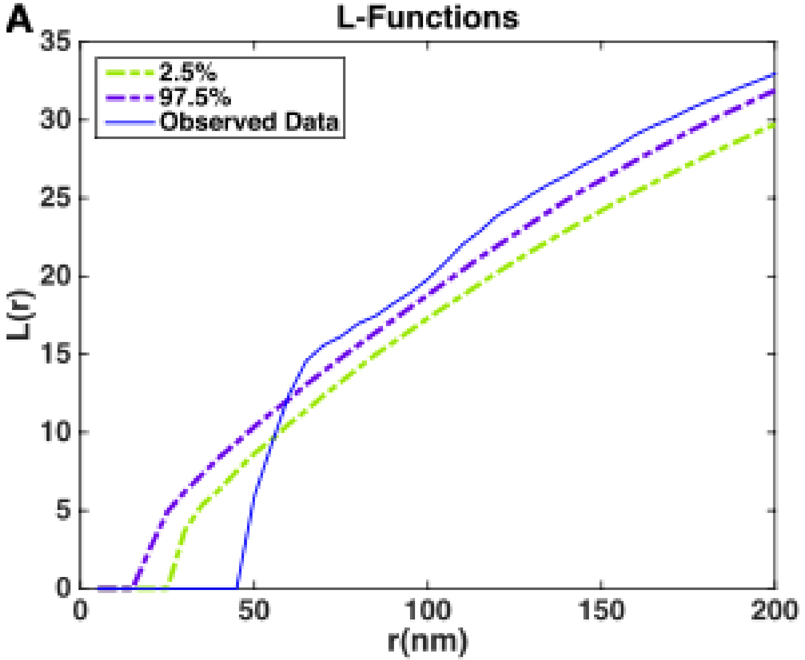


- Analysis of shape and spatial interaction of synaptic vesicles using data from focused ion beam scanning electron microscopy (FIB-SEM); M Khanmohammadi, RP Waagepetersen & J Sporring, Frontiers in Neuroanatomy, 2015
- Stoyan, D. (2006). Fundamentals of Point Process Statistics. In: Baddeley, A. et al. (eds) Case Studies in Spatial Point Process Modeling. Lecture Notes in Statistics, vol 185. Springer
- Mrkvička, Tomáš, et al. "A one-way ANOVA test for functional data with graphical interpretation." *Kybernetika* 56.3 (2020): 432-458.
- Stephensen, H.J.T., Svane, A.M., Villanueva, C.B. et al. Measuring Shape Relations Using r-Parallel Sets. J Math Imaging Vis, vol 63, 2021.
- Chazal F., Michel B., An Introduction to Topological Data Analysis: Fundamental and Practical Aspects for Data Scientists, In: Frontiers in Artificial Intelligence, vol 4, 2021

1. Spatial point fields





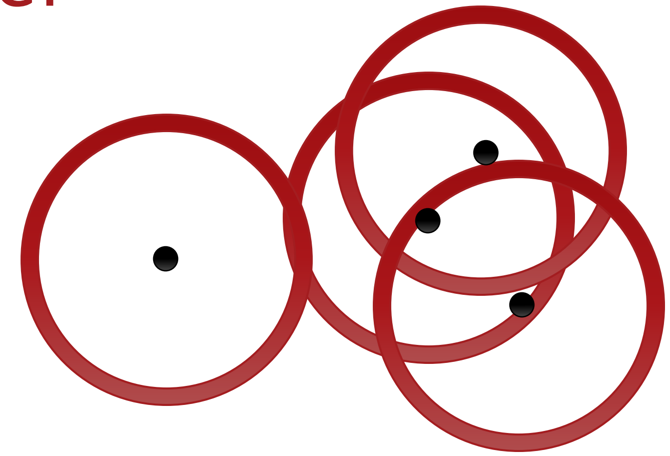


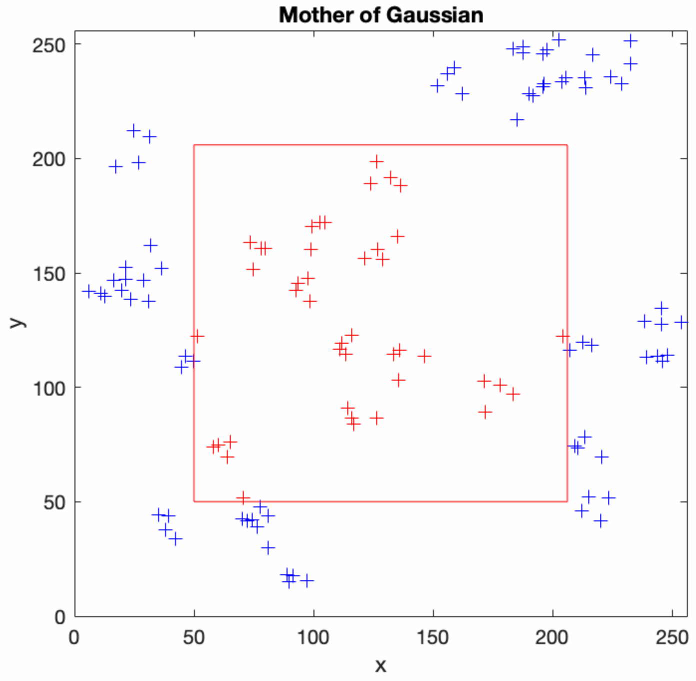
Analysis of shape and spatial interaction of synaptic vesicles using data from focused ion beam scanning electron microscopy (FIB-SEM); M Khanmohammadi, RP Waagepetersen & J Sporring, Frontiers in Neuroanatomy, 2015

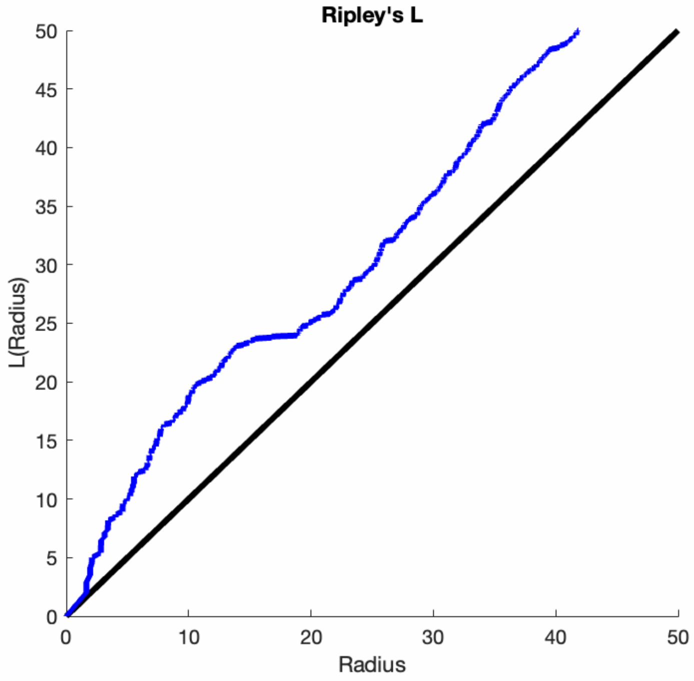
Ripley's K- and L-functions: expected number of neighboring points by radius

$$K(r) = \frac{1}{\lambda}\mathbb{E}\left[I\left(d_{ij} < r\right)\right]$$

$$L = \sqrt{\frac{K}{\pi}}$$







R and rpy2 demo:
https://sporring.github.io/bia2024/talk.pdf
https://sporring.github.io/bia2024/spatstat_bia2024.zip

https://cran.r-project.org/
https://spatstat.org/
https://cran.r-project.org/web/packages/GET/vignettes/pointpatterns.pdf

demoRpy2.py: Installation instructions for R, R-packages, and python packages:

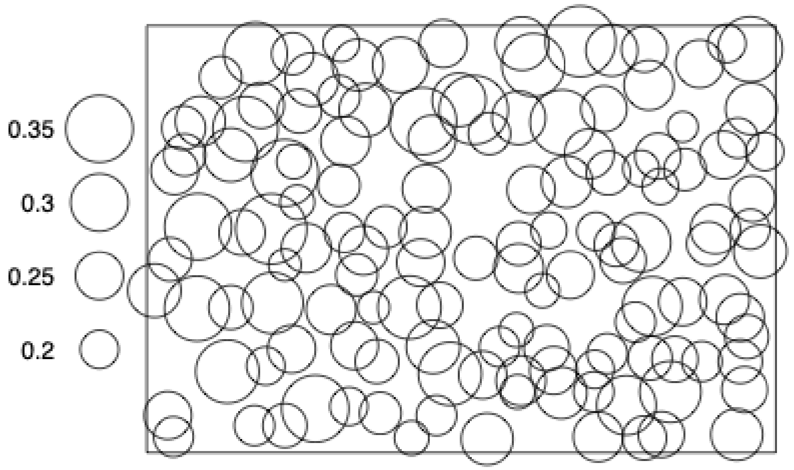
```
# 1. Install R, which to my experience works best directly from  https://cran.r-project.org/
# then start R and install some packes:
#  install.packages("spatstat")
#  install.packages("lazyeval")
#  install.packages("GET")
…
```

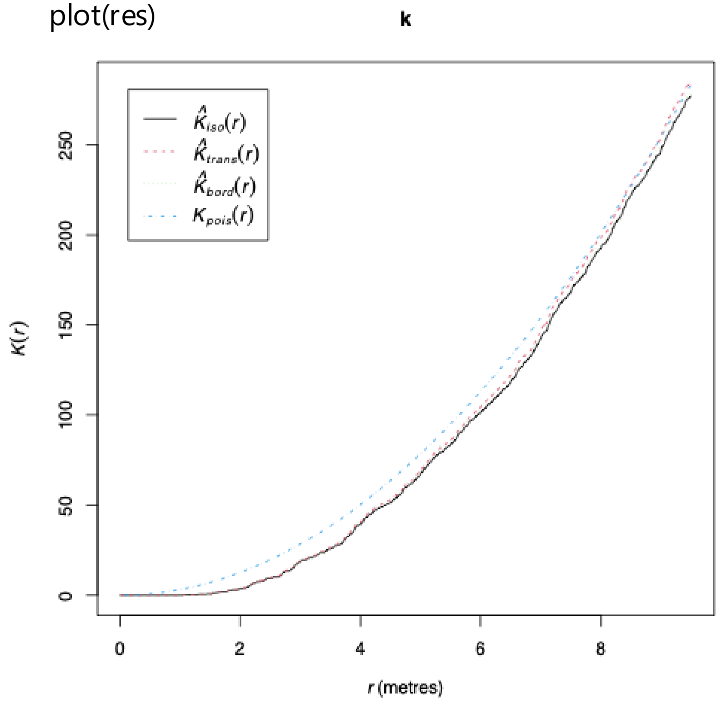
GETDemo.r

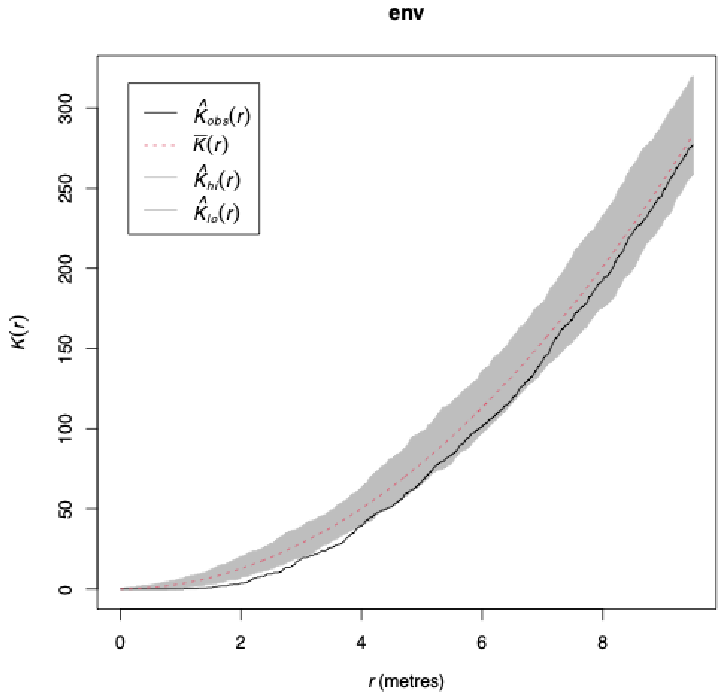
```
library("GET")
library("spatstat.model")
library("ggplot2")
X <- spruces
print(X)
plot(X)
k <- Kest(X)
print(k)
plot(k)
env <- envelope(X, nsim=1999, savefuns=TRUE, simulate=expression(runifpoint(ex=X)), verbose=FALSE)
print(env)
plot(env)
res <- global_envelope_test(env)
print(res)
plot(res)
```

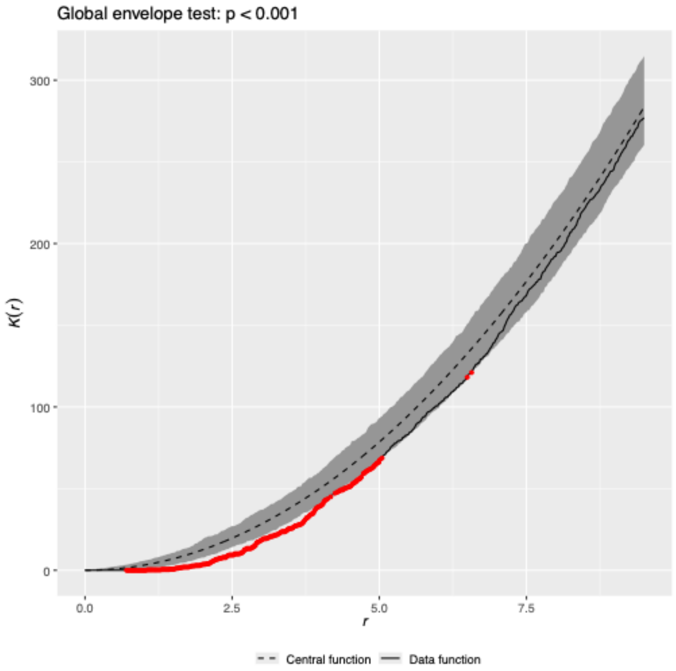
GETDemo.r

```
library("GET")
library("spatstat.model")
library("ggplot2")
X <- spruces
print(X)
plot(X)
k <- Kest(X)
print(k)
plot(k)
env <- envelope(X, nsim=1999, savefuns=TRUE, simulate=expression(runifpoint(ex=X)), verbose=FALSE)
print(env)
plot(env)
res <- global_envelope_test(env)
print(res)
plot(res)
```

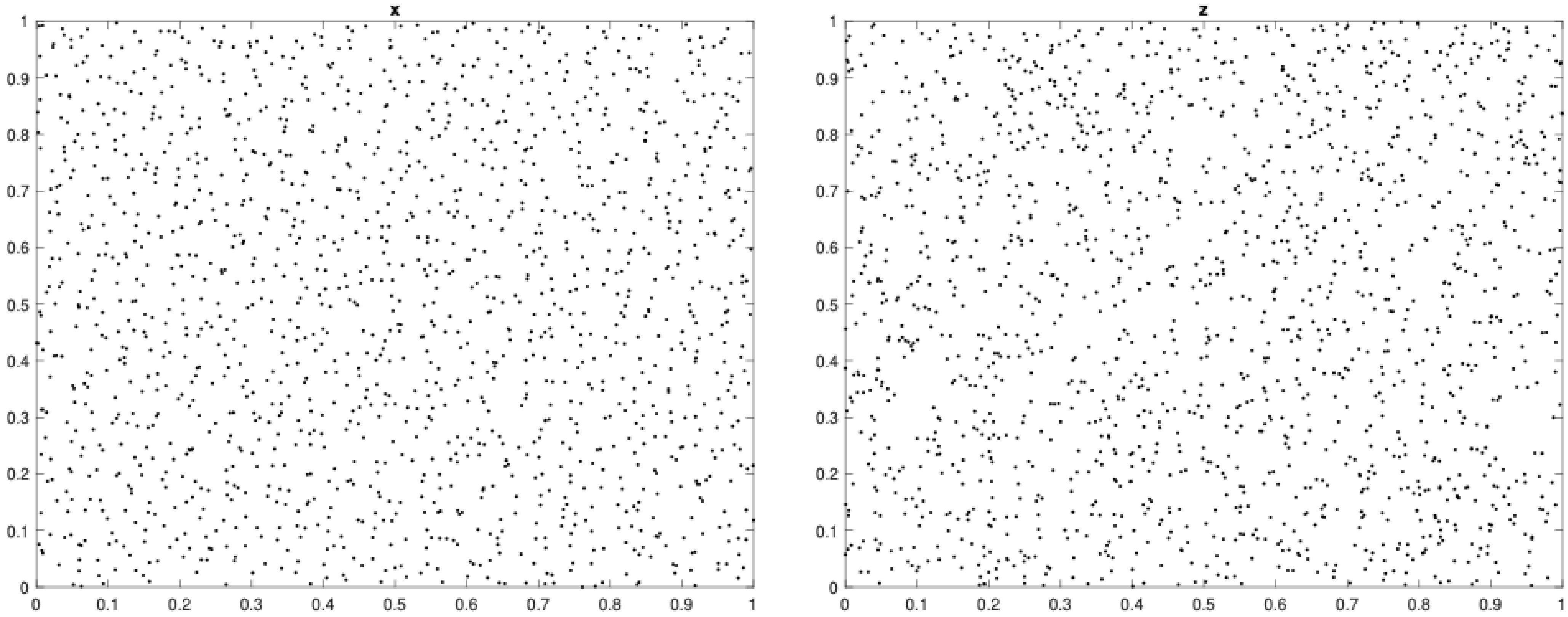




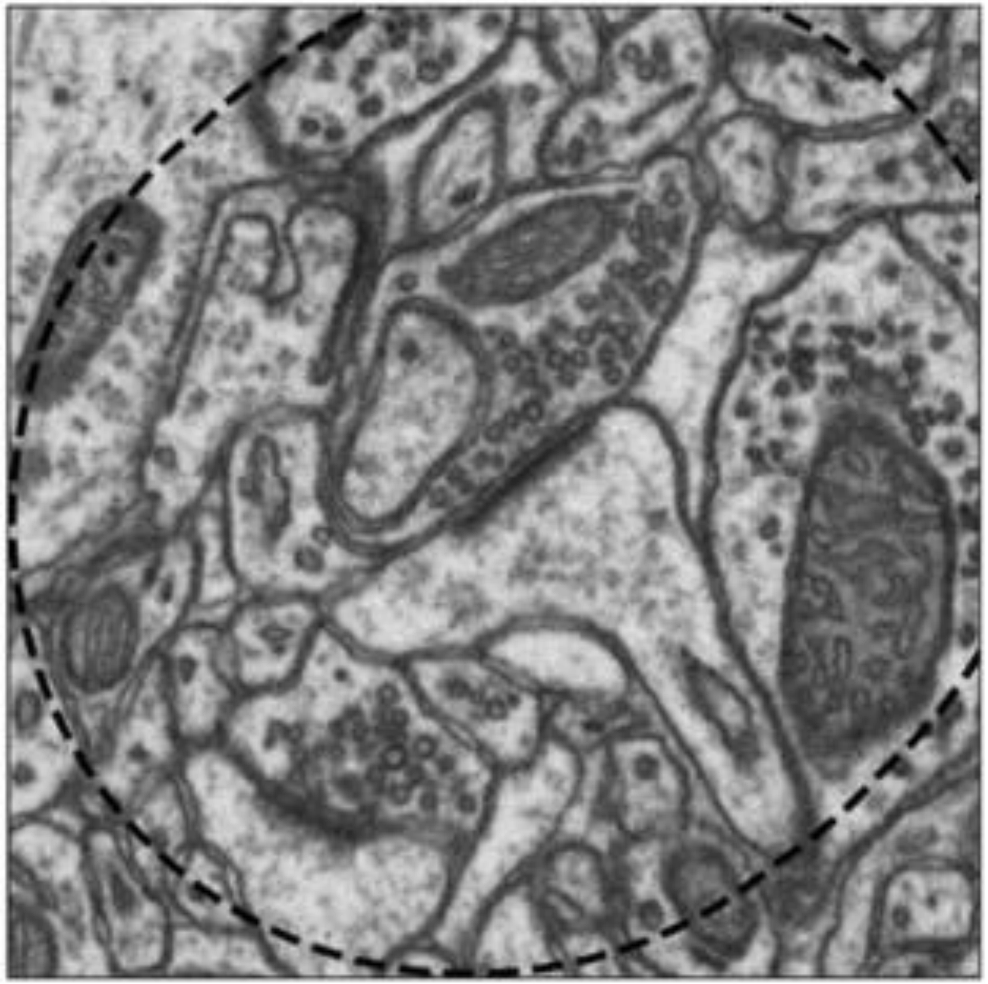




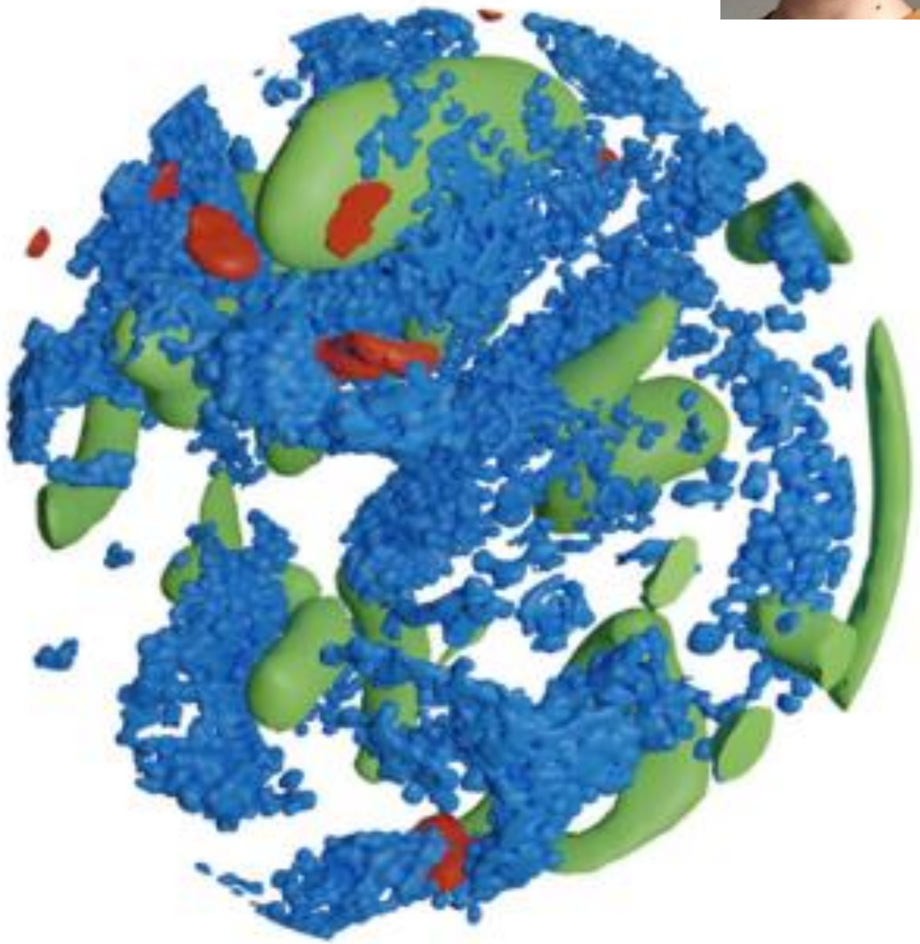
Exercise: Use GET's global_envelope_test to test whether x and/or z are likely to be random



2. Spatial shape fields: Real structures are not points, small structures are difficult to separate

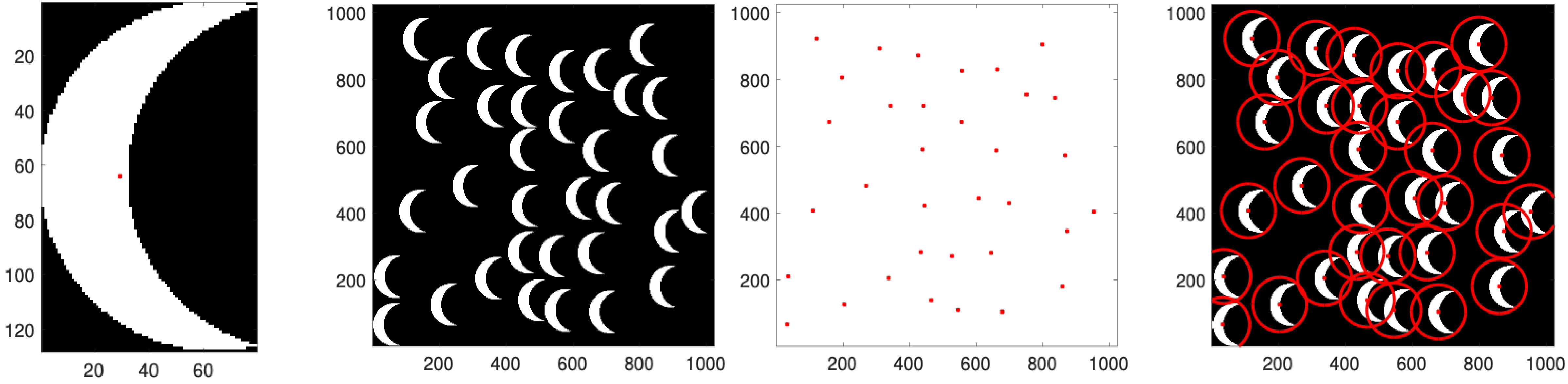


Deep Learning

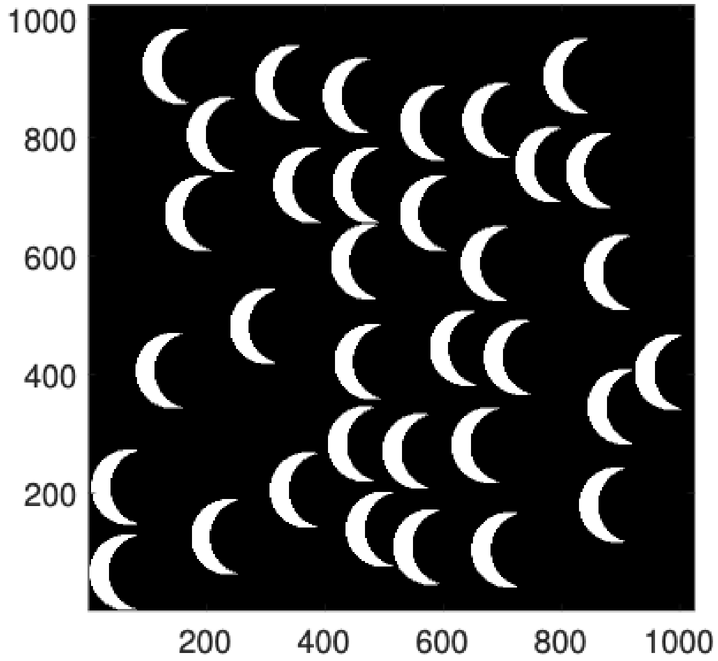


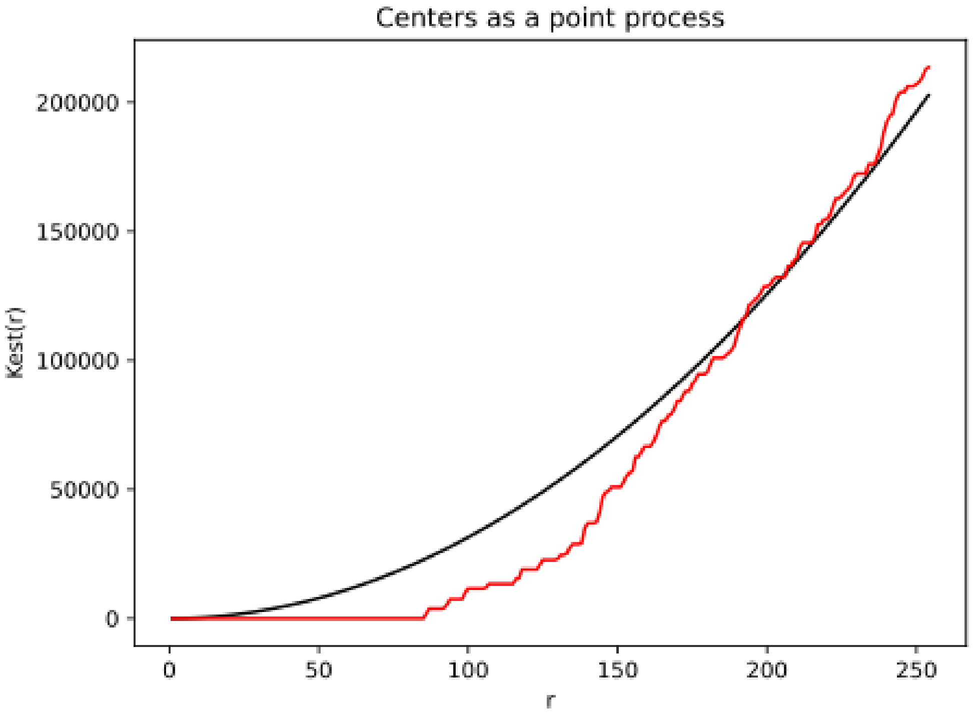
Measuring Shape Relations Using r-Parallel Sets; HJT Stephensen, AM Svane, CB Villanueva, SA Goldman, & J Sporring; Journal of mathematical imaging and vision, 2021

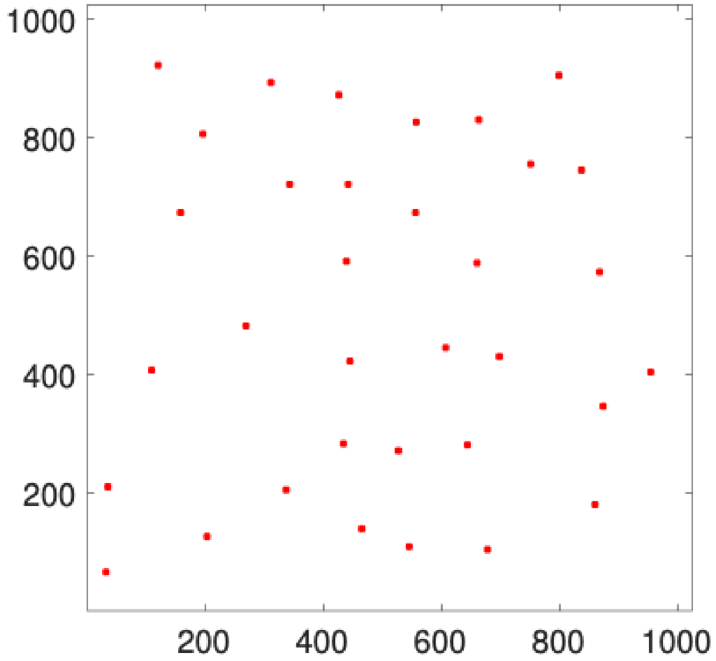
Not all shapes are well summarized as a reference point



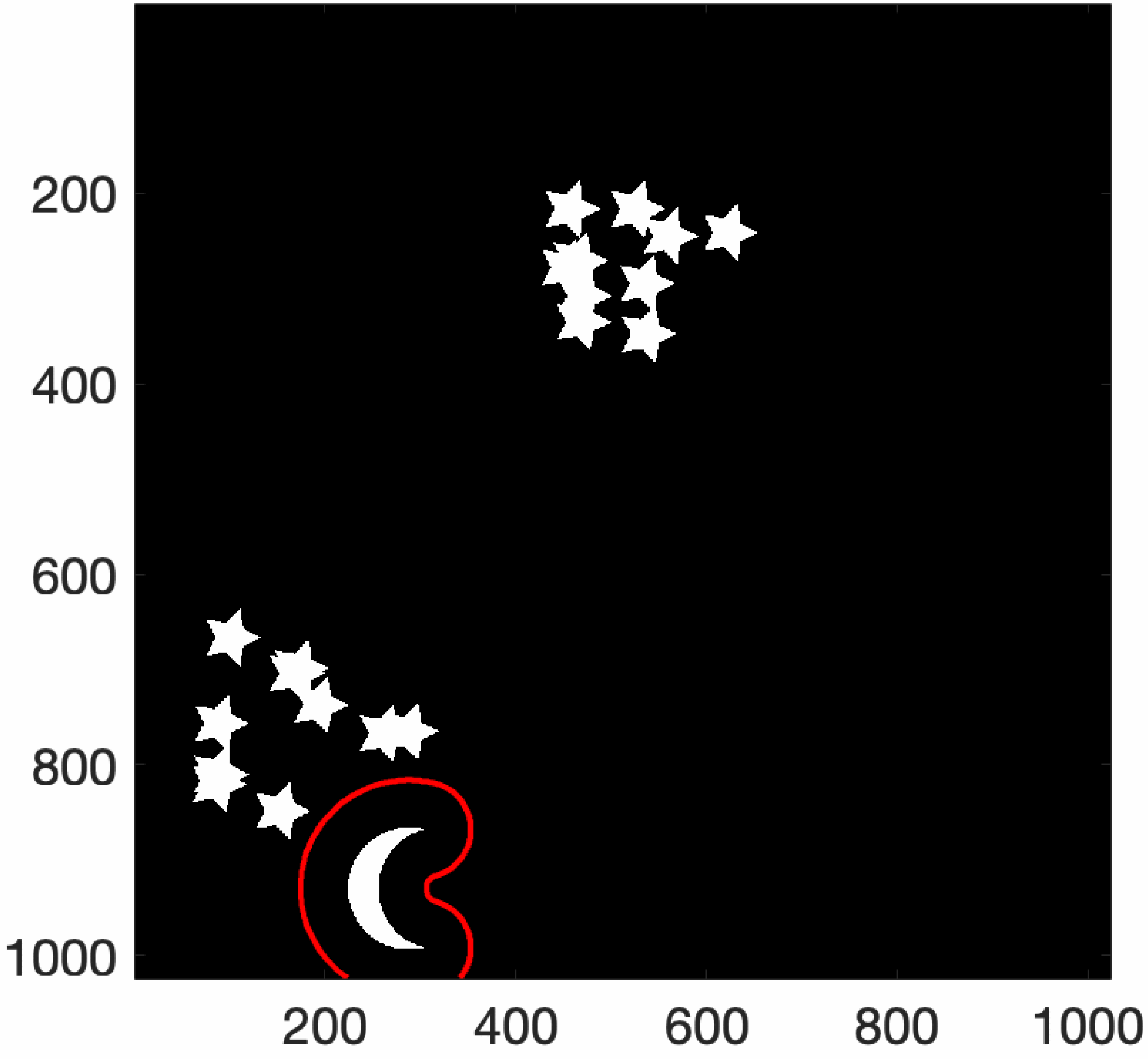
Ripley's K-function indicates strong repulsion



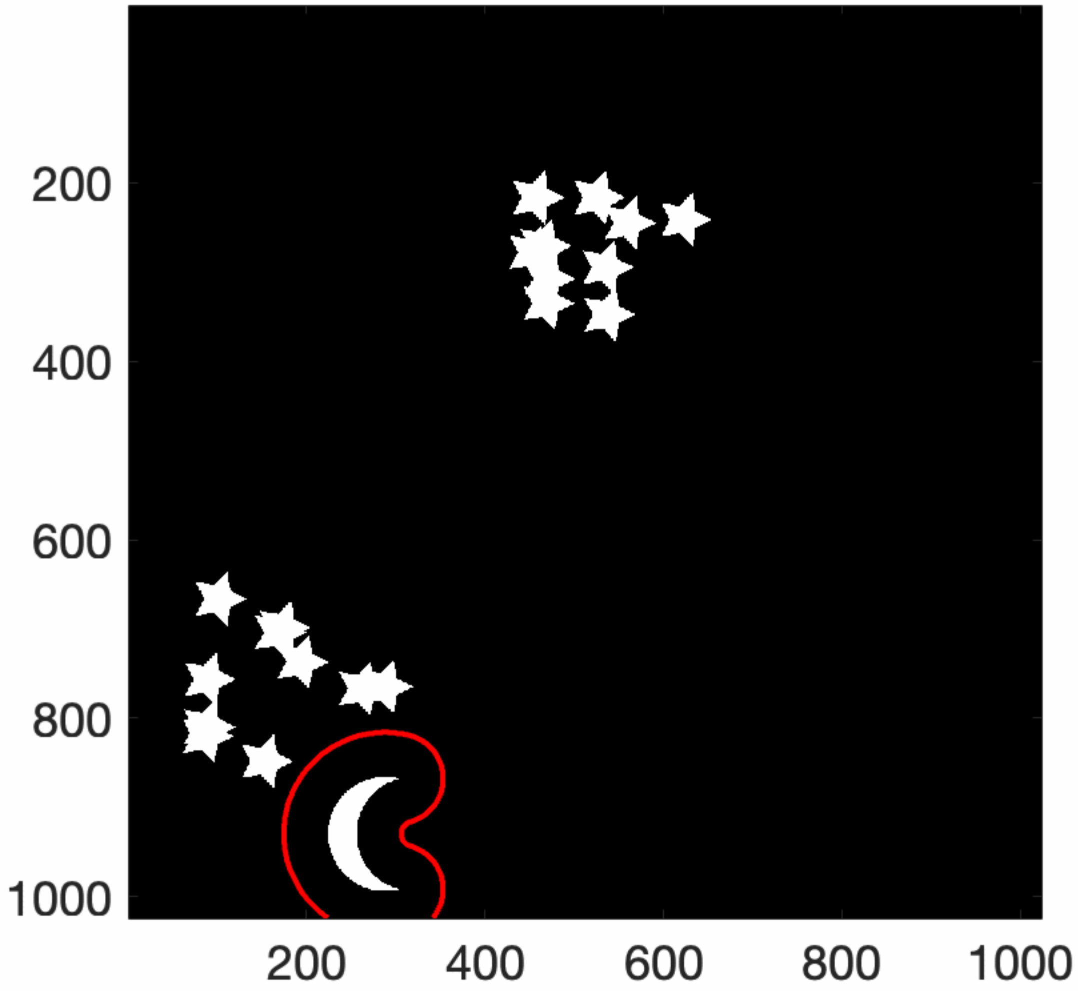


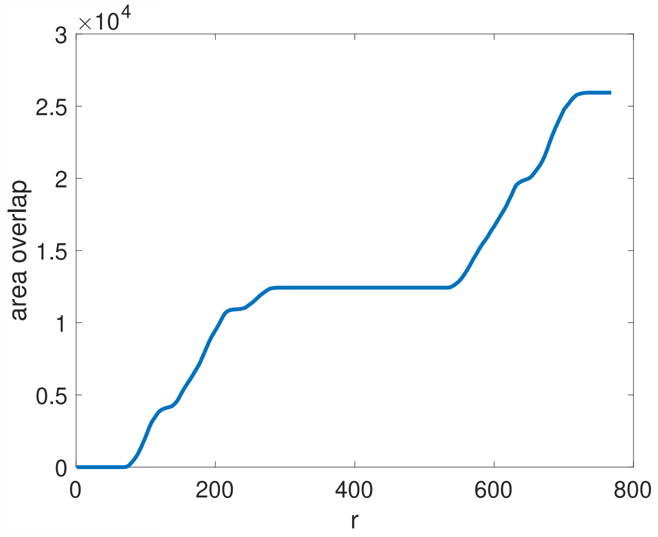


Shape relation measures = Interaction with distance fields



Shape relation measures: Some characteristic functions

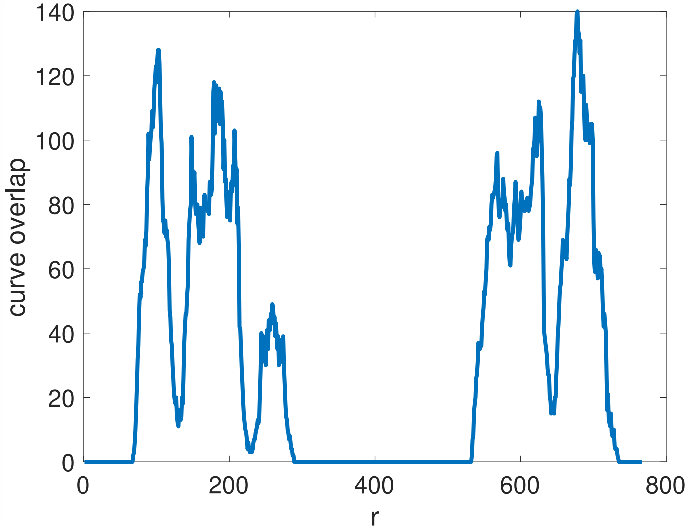


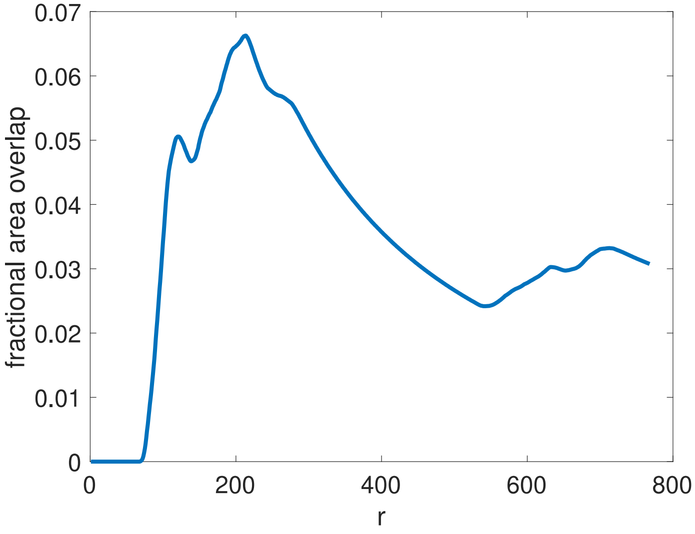


$$\mu_{00}(r) = \mathcal{H}(X \cap Y^r)$$

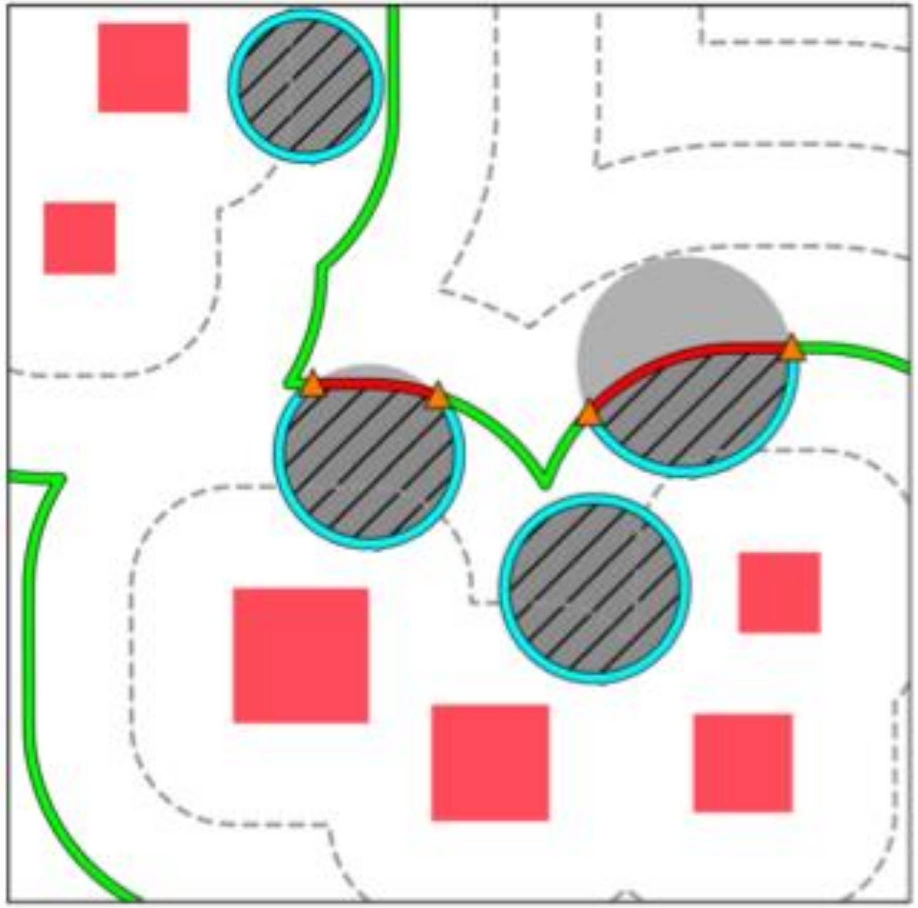
$$g_{00}(r) = \frac{d\mu_{00}(r)}{dr} = \mu_{01}(r)$$

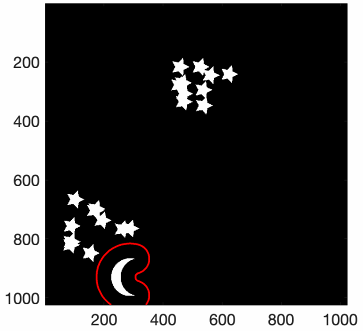
$$f_{00}(r) = \frac{\mu_{00}(r)}{\mathcal{H}(Y^r)}$$





There can be only 4





$$Y^r = \{\alpha \in \mathbb{R}^d \mid \inf_{y \in Y} d(\alpha, y) \leq r\} \tag{1}$$

$$\mu_{\varepsilon,\varepsilon'}(X, Y^r) = \mathcal{H}^{d-\varepsilon-\varepsilon'}(\partial^\varepsilon X \cap \partial^{\varepsilon'} Y^r) \ . \tag{2}$$

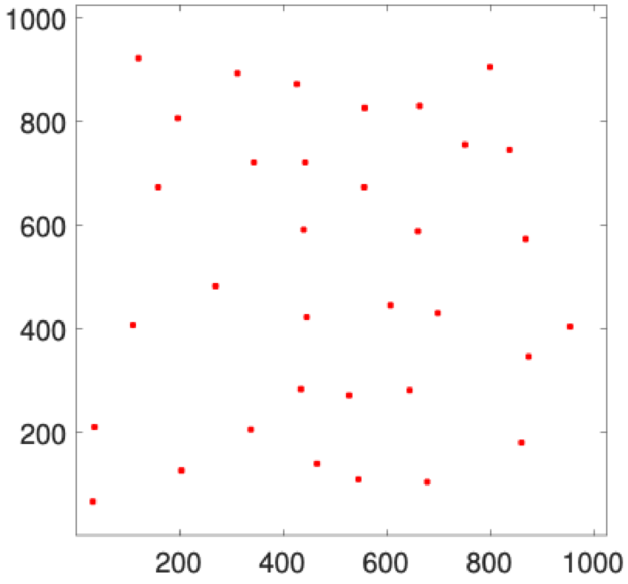
$d=2$	$(\varepsilon,\varepsilon')$	$\mathcal{H}^{d-\varepsilon-\varepsilon'}$	$\partial^\varepsilon X \cap \partial^{\varepsilon'} Y^r$	Interpretation of $\mu_{\varepsilon,\varepsilon'}(X, Y^r)$
	$(0,0)$	Area	$X \cap Y^r$	Area of intersection
	$(0,1)$	Curve length	$X \cap \partial Y^r$	Boundary length of intersection inside interior of X
	$(1,0)$	Curve length	$\partial X \cap Y^r$	Boundary length of intersection inside boundary of X
	$(1,1)$	Point counts	$\partial X \cap \partial Y^r$	Number of points in intersection of boundaries

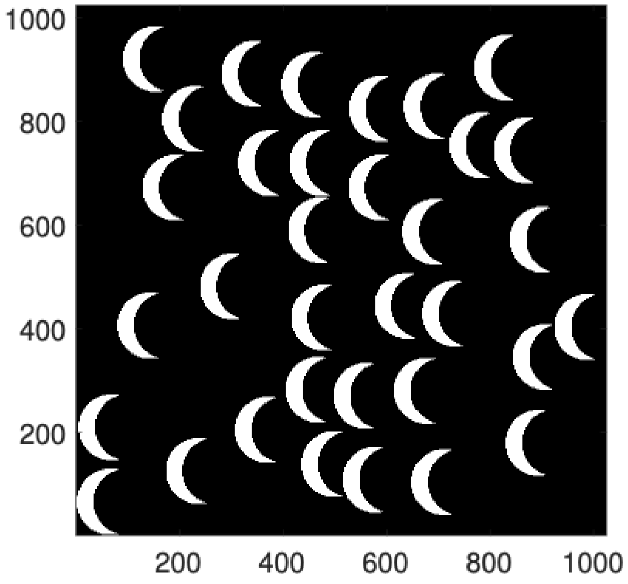
$d=3$	$(\varepsilon,\varepsilon')$	$\mathcal{H}^{d-\varepsilon-\varepsilon'}$	$\partial^\varepsilon X \cap \partial^{\varepsilon'} Y^r$	Interpretation of $\mu_{\varepsilon,\varepsilon'}(X, Y^r)$
	$(0,0)$	Volume	$X \cap Y^r$	Volume of intersection
	$(0,1)$	Surface area	$X \cap \partial Y^r$	Surface area of intersection inside interior of X
	$(1,0)$	Surface area	$\partial X \cap Y^r$	Surface area of intersection inside boundary of X
	$(1,1)$	Curve length	$\partial X \cap \partial Y^r$	Length of intersection of boundaries

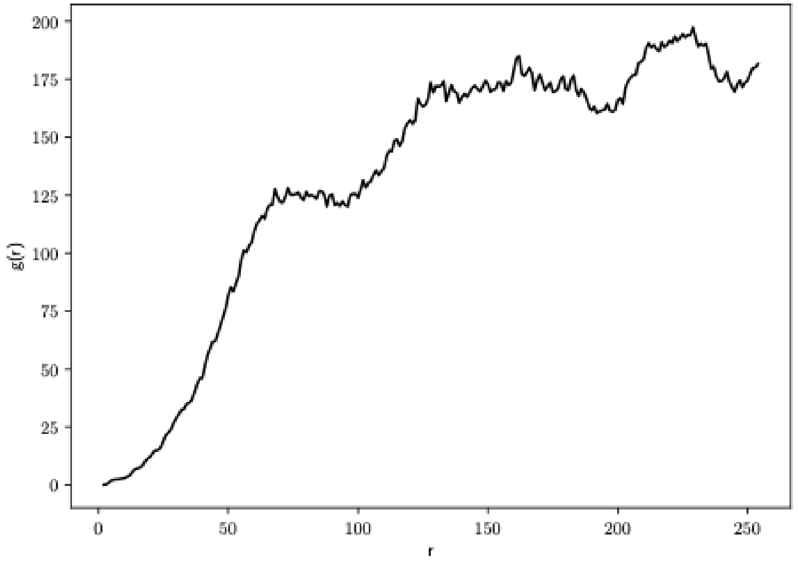
https://lab.compute.dtu.dk/QIM/tools/relationalshapemeasure

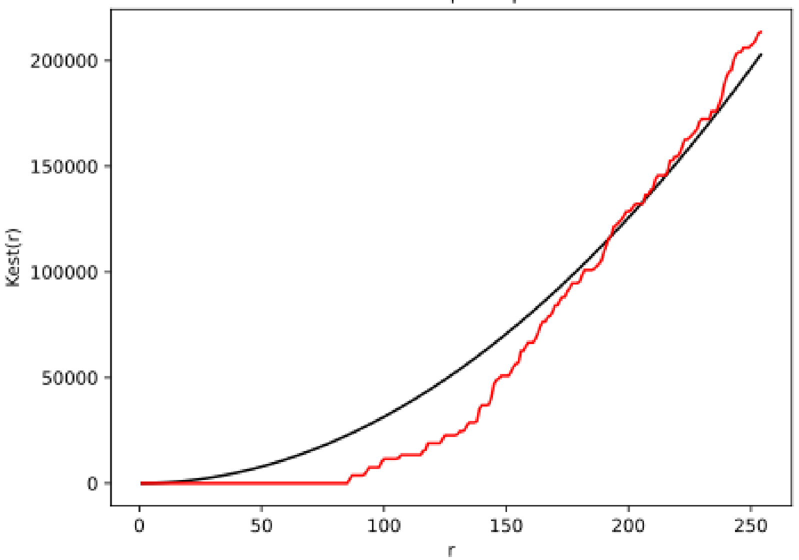
Spatial point fields vs. shape fields: shapeRelation.py

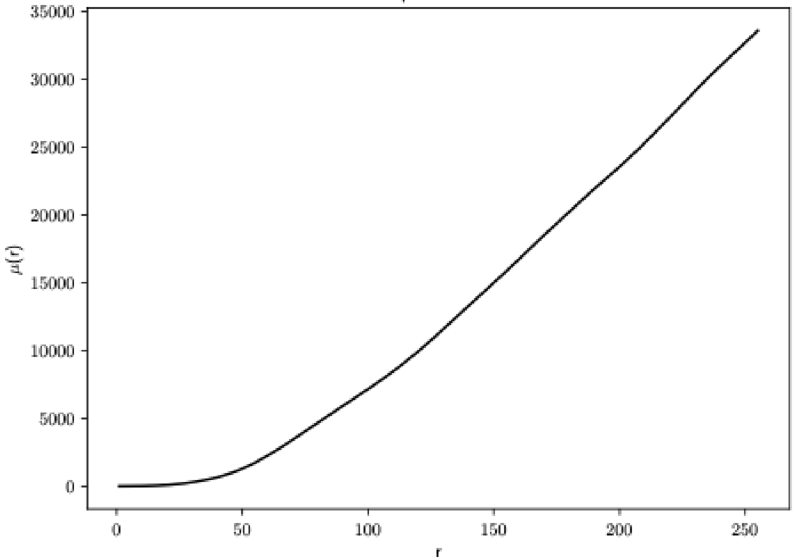
$$\mu_{00}(r) = \mathcal{H}(X \cap Y^r)$$
$$g_{00}(r) = \frac{d\mu_{00}(r)}{dr} = \mu_{01}(r)$$
$$f_{00}(r) = \frac{\mu_{00}(r)}{\mathcal{H}(Y^r)}$$

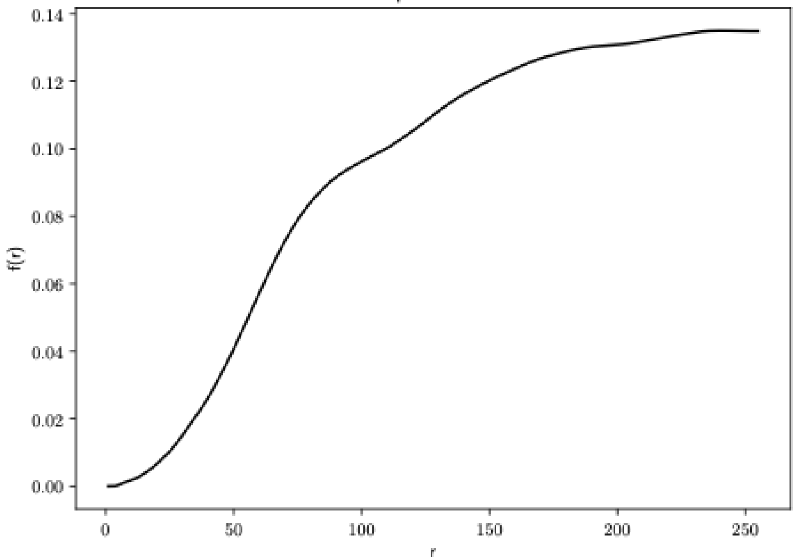




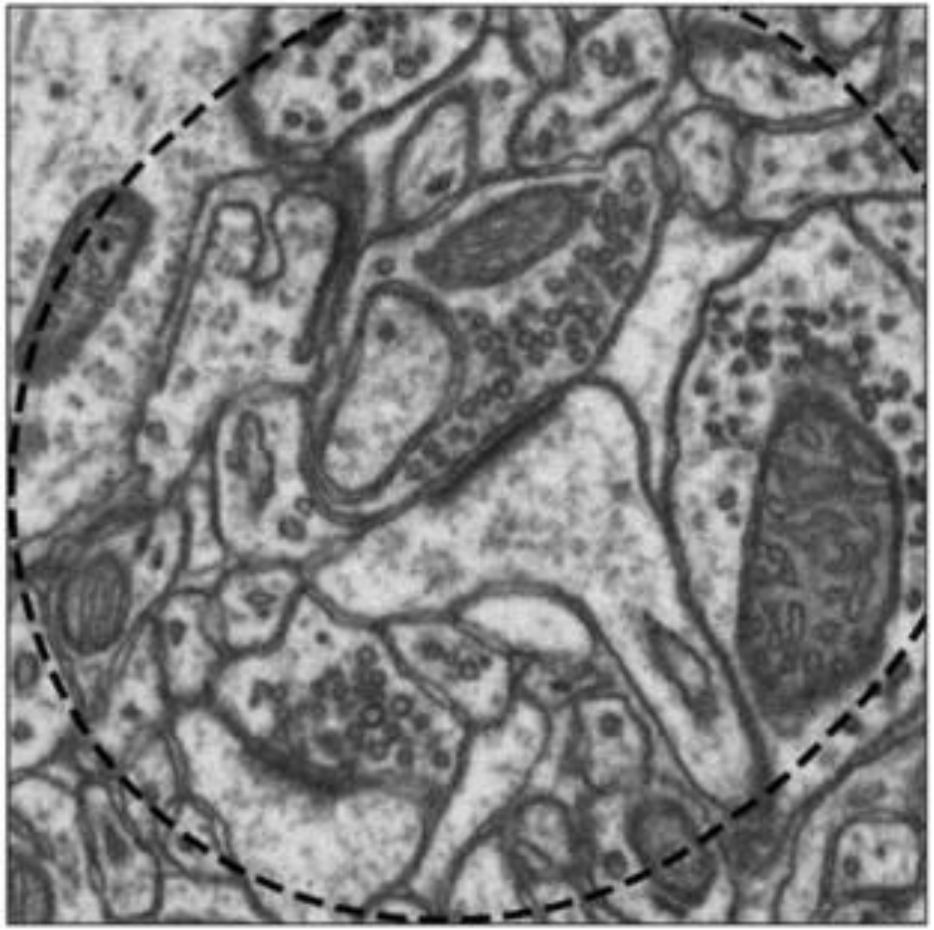


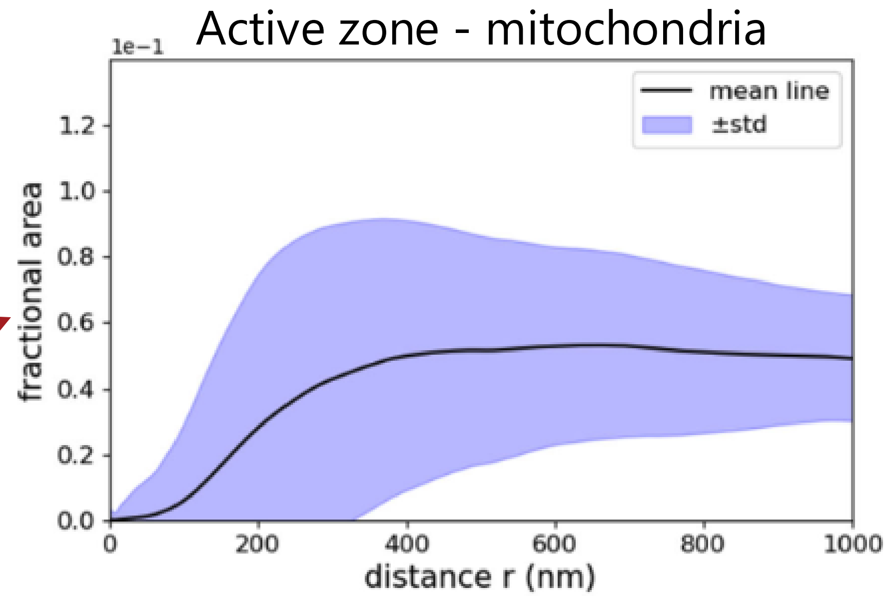


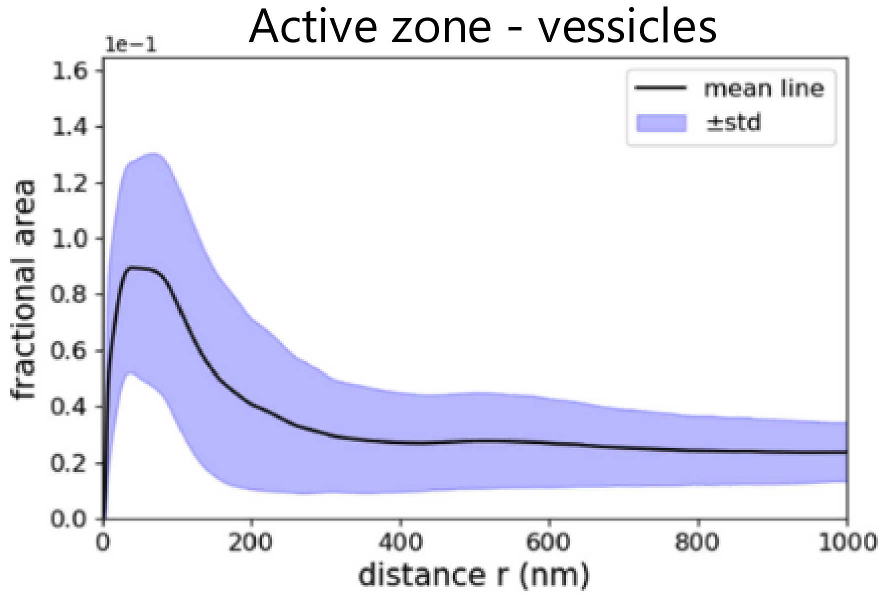




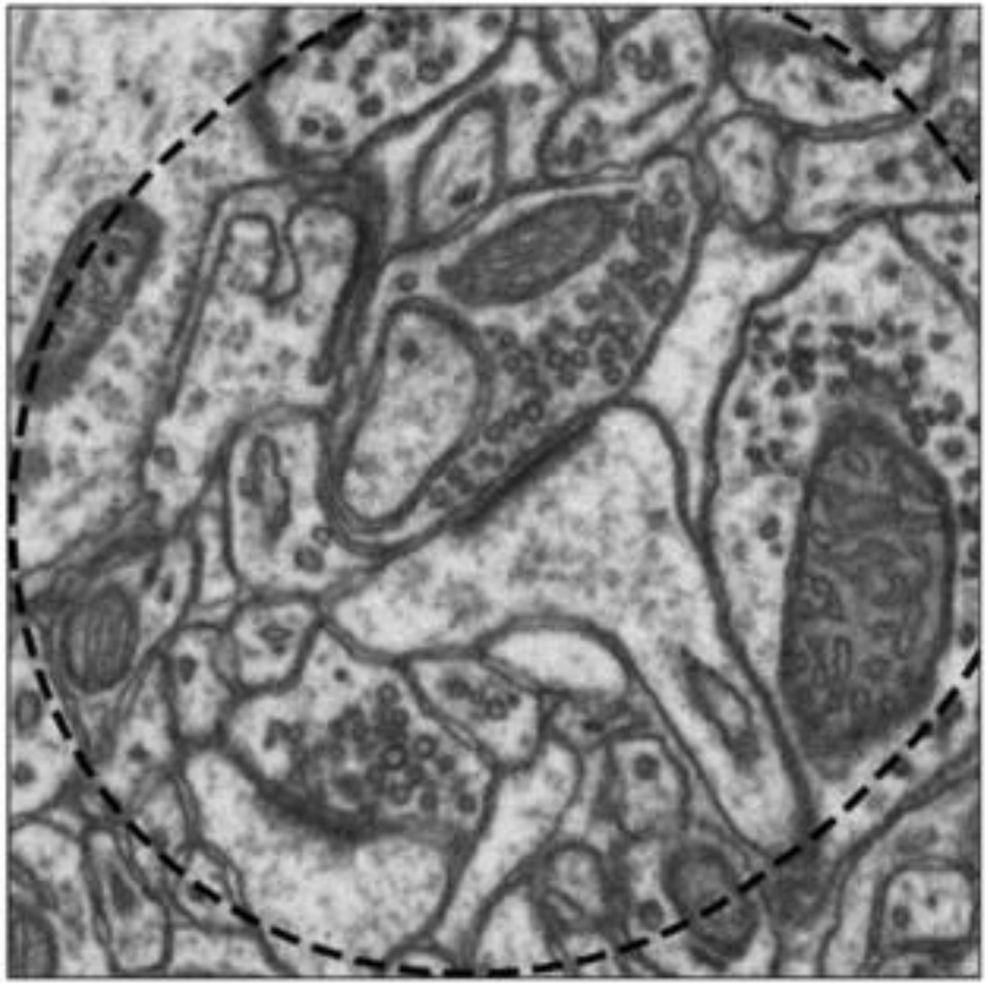
Shape relations for statistical summary of families of shapes and their relations



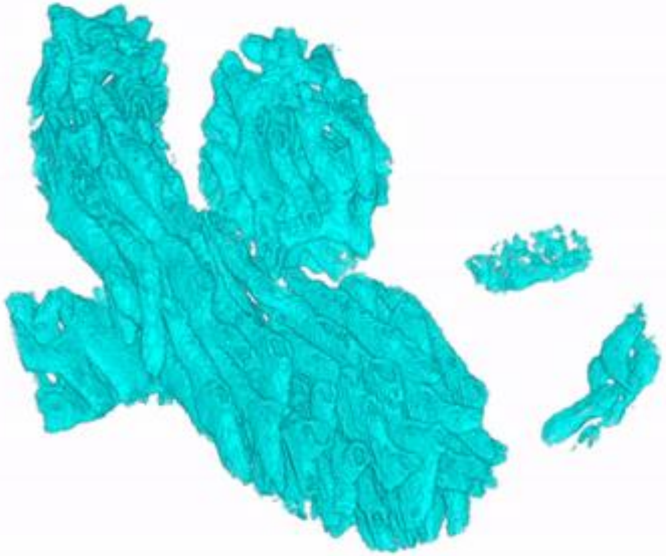




3. With a little help from topology: Cristae membranes in Mitochondria



Deep Learning

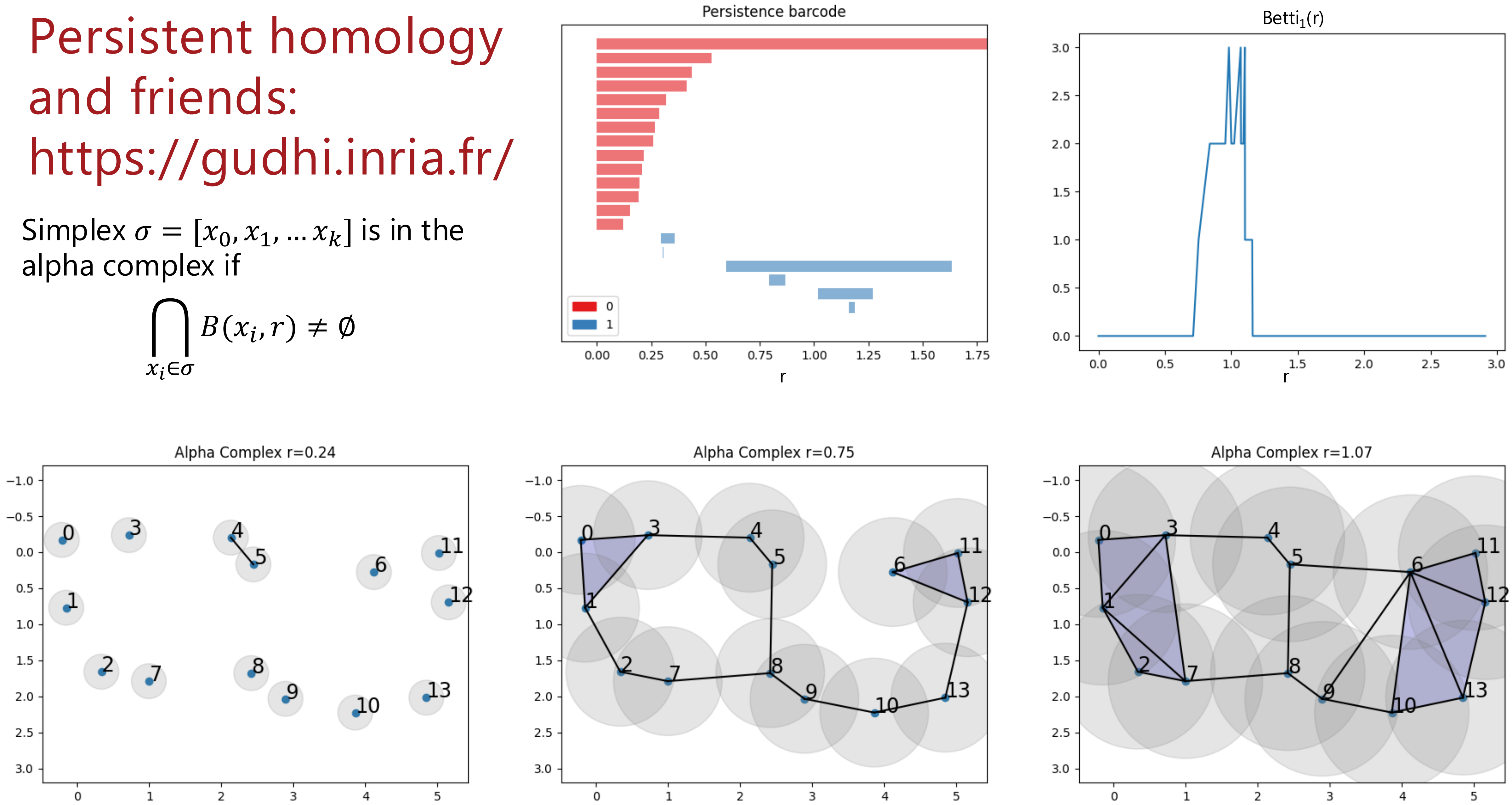


Extracting Mitochondrial Cristae Characteristics from 3D Focused Ion Beam Scanning Electron Microscopy Data, C Wang, L Østergaard, S Hasselholt, & J Sporring, to appear in Communications Biology, 2024

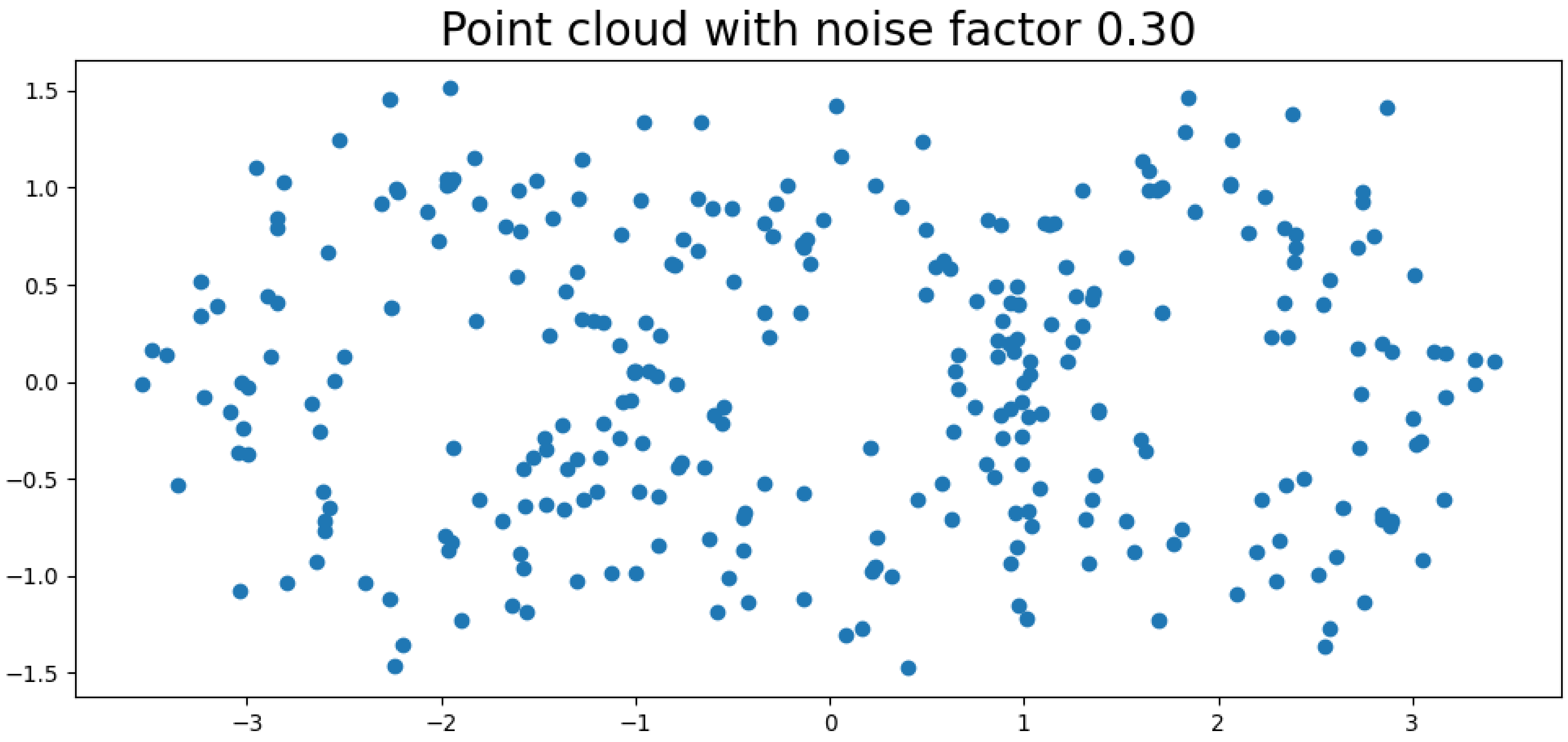
Persistent homology and friends: https://gudhi.inria.fr/

Simplex $\sigma = [x_0, x_1, \dots x_k]$ is in the alpha complex if

$$\bigcap_{x_i \in \sigma} B(x_i, r) \neq \emptyset$$

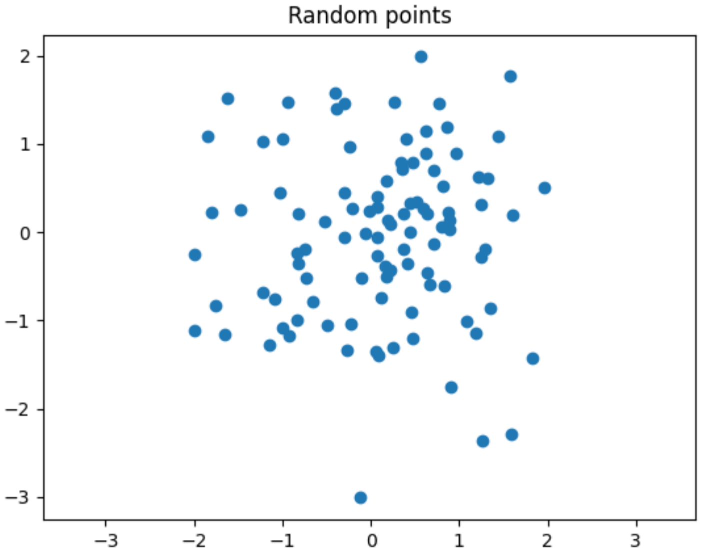


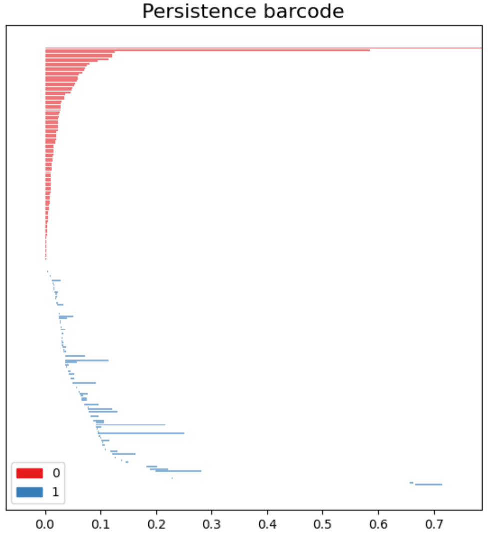
Persistence under noise: $\mathrm{dgm} = \mathrm{dgm}^{\mathrm{Signal}} \cup \mathrm{dgm}^{\mathrm{Noise}}$

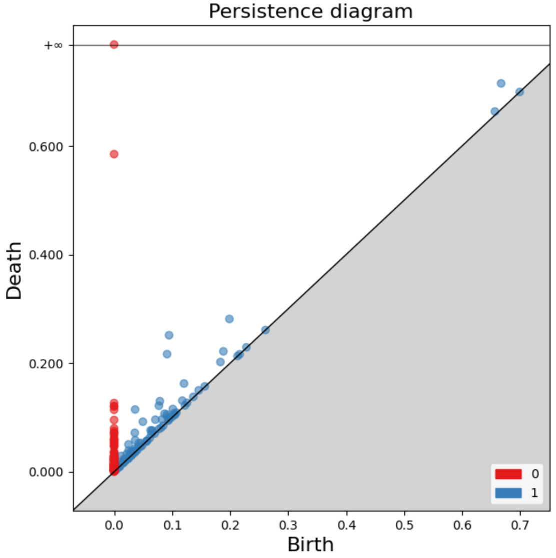


Pure noise

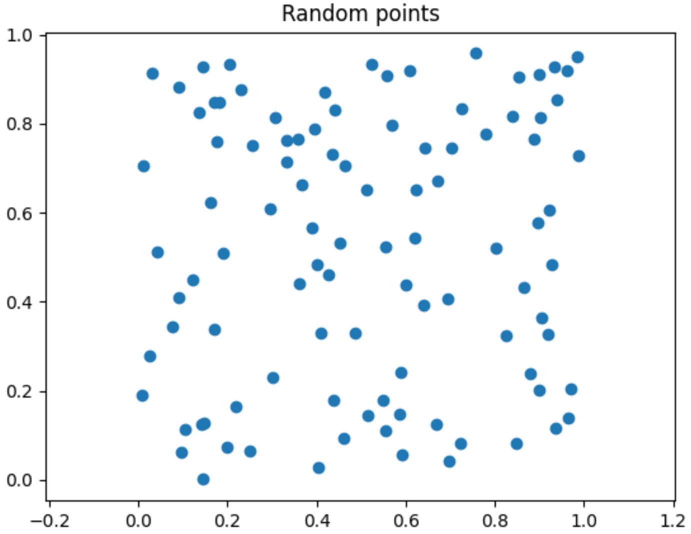
Normal

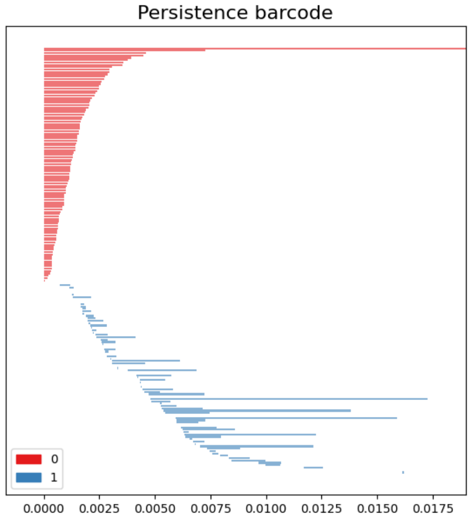


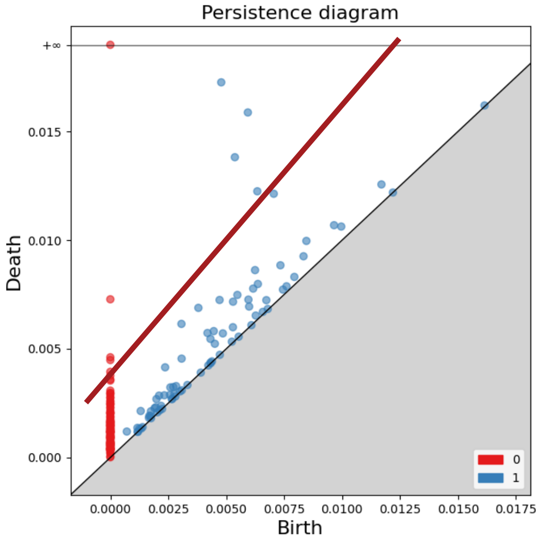




Uniform







Bobrowski & Skraba, “A universal null-distribution for topological data analysis”, Nature/Scientific Reports, 2023

Random points:

$$x \in S(d), \qquad x \sim f, \qquad p = (r_{\text{birth}}, r_{\text{death}})$$

Left-skewed Gumbel distribution:

$$F(x) = 1 - e^{-e^x}, \qquad f(x) = e^{x - e^x}, \qquad \mu = -\gamma = -0.57721, \qquad \sigma^2 = \frac{\pi^2}{6}$$

Transformation:

$$\rho = \ln \ln \frac{r_{\text{death}}}{r_{\text{birth}}}$$

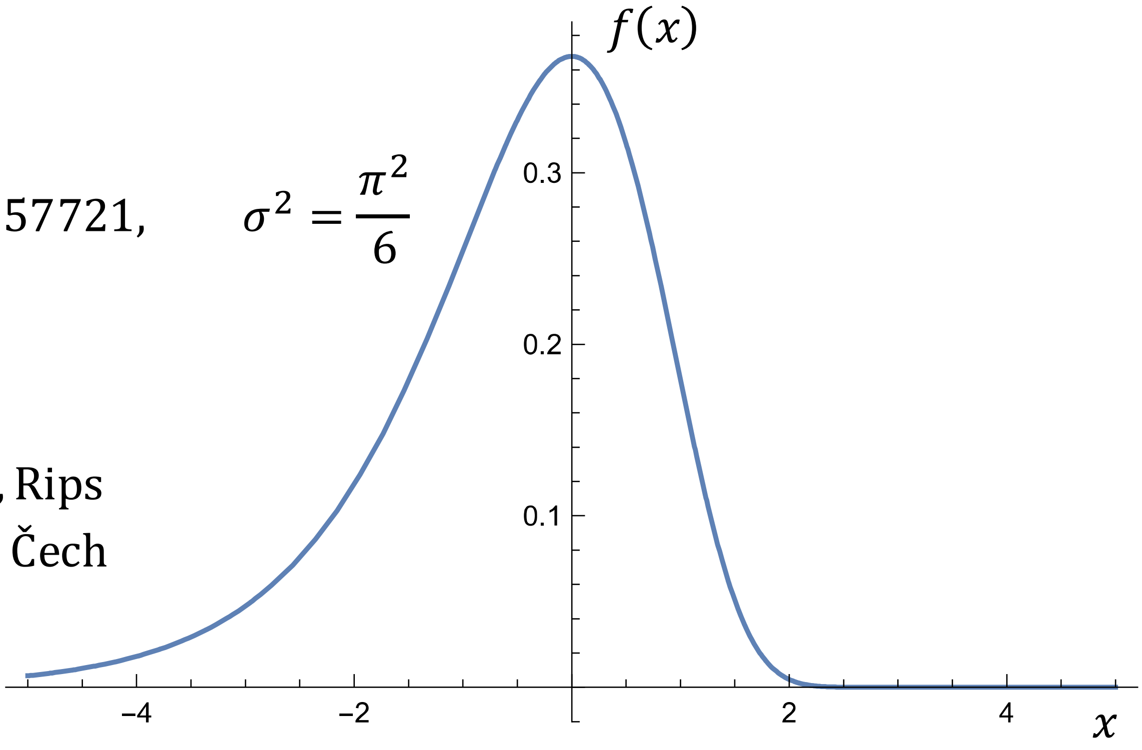
$$\mathrm{x} = \frac{(\rho - \bar{\rho})}{\beta} - \gamma, \qquad \bar{\rho} = \frac{1}{|\mathrm{dgm}_k|} \sum_{p \in \mathrm{dgm}_k} \rho\,, \qquad \beta = \begin{cases} 1, \text{Rips} \\ 2, \text{Čech} \end{cases}$$

Bonferroni testing (family-wise error rate $< \alpha$):

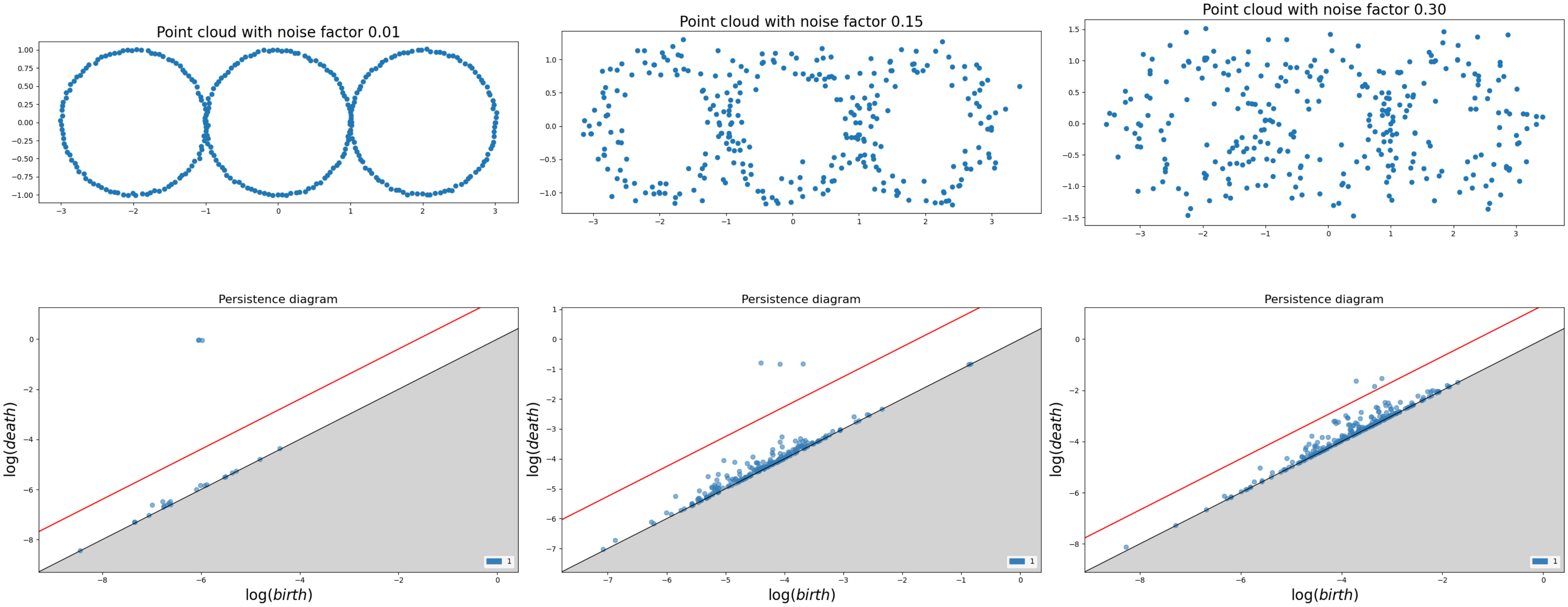
$$P(x \geq x_0 | x \text{ is noise}) = 1 - F(x) = e^{-e^{x_0}}$$

$$\mathrm{dgm}_k^{\text{Signal}}(\alpha) = \left\{ p \in \mathrm{dgm}_k\colon\ e^{-e^x} < \frac{\alpha}{|\mathrm{dgm}_k|} \right\}$$

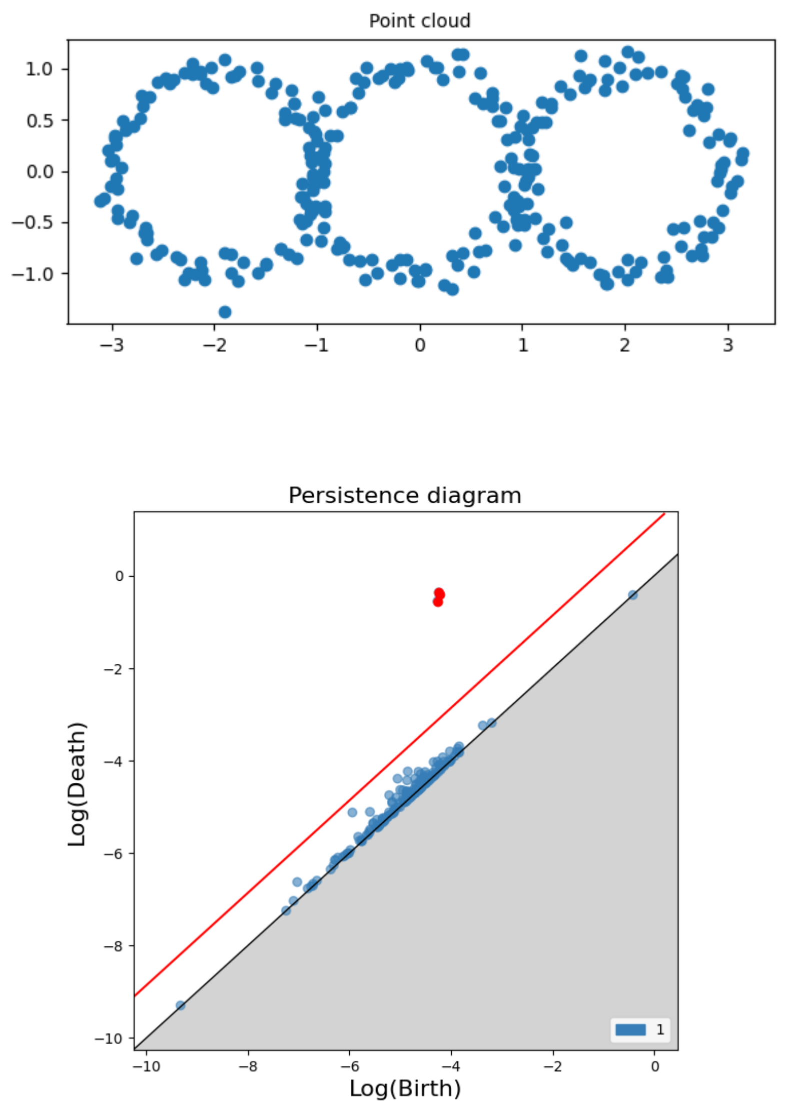
$$e^{\rho} = \ln \frac{r_{\text{death}}}{r_{\text{birth}}} = (-1)^{\beta} e^{\beta\gamma + \bar{\rho}} \left(\ln \frac{\alpha}{|\mathrm{dgm}_k|} \right)^{\beta}$$

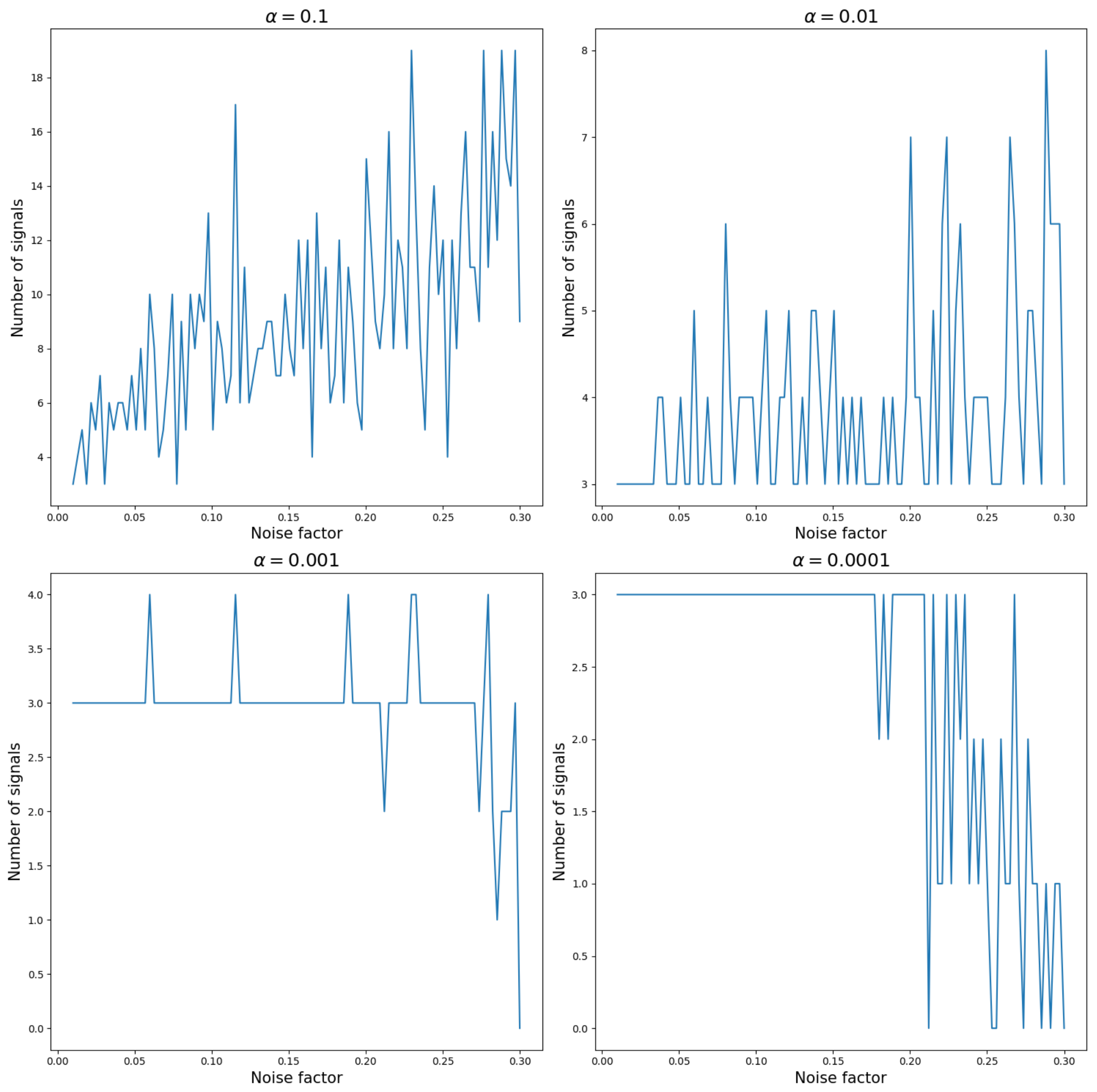


The universal distribution can separate very noisy cases

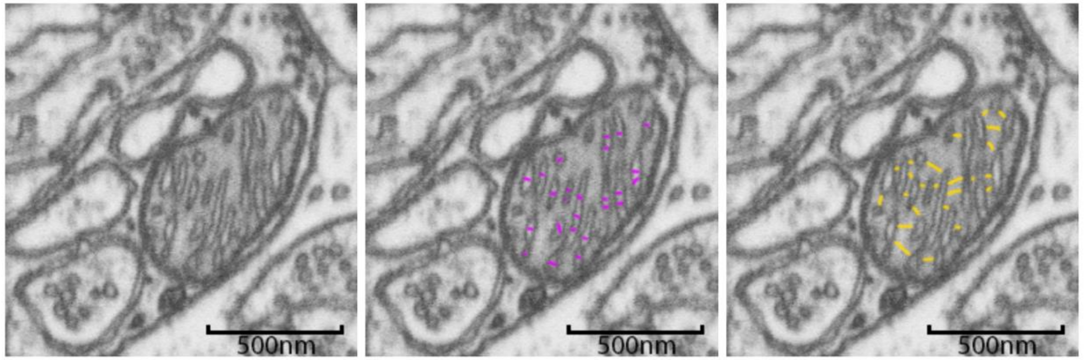


Family-wise error rate

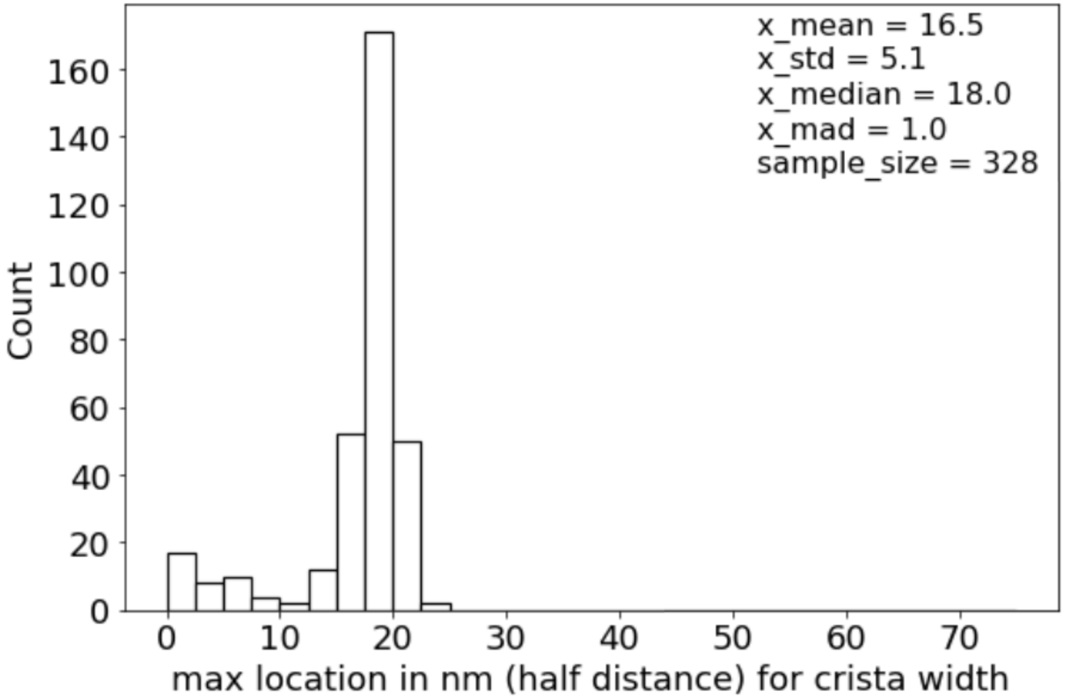




Persistent homology: Statistical measures on H_0

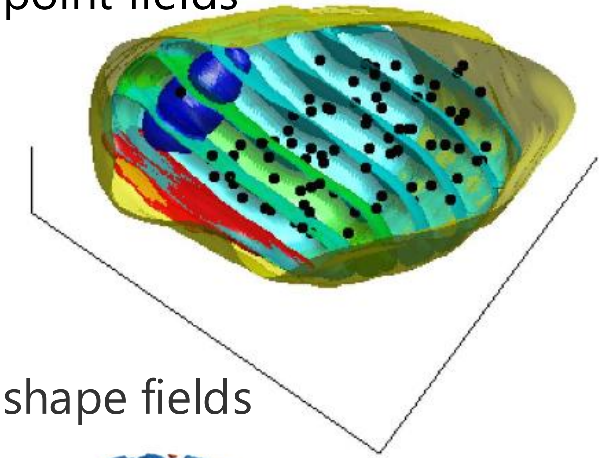


(a) (b) (c)



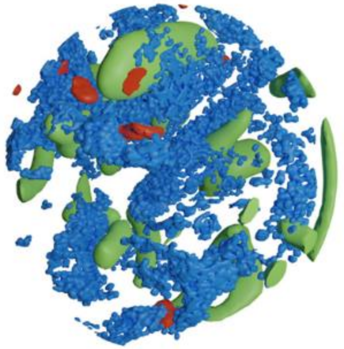
Statistical summary of object collections

Spatial point fields



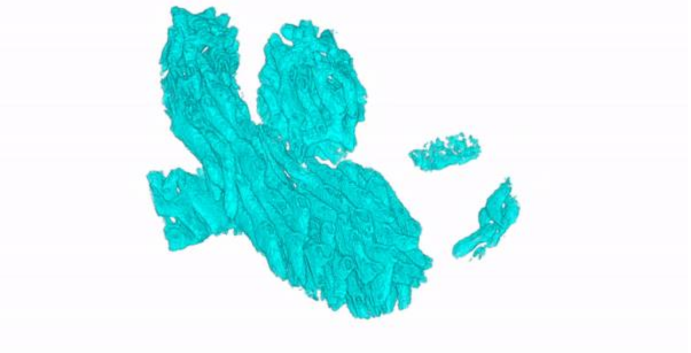
Pair correlation and Ripley's K functions summarizes 1st order point relations – e.g., do the vessicles cluster?

Spatial shape fields

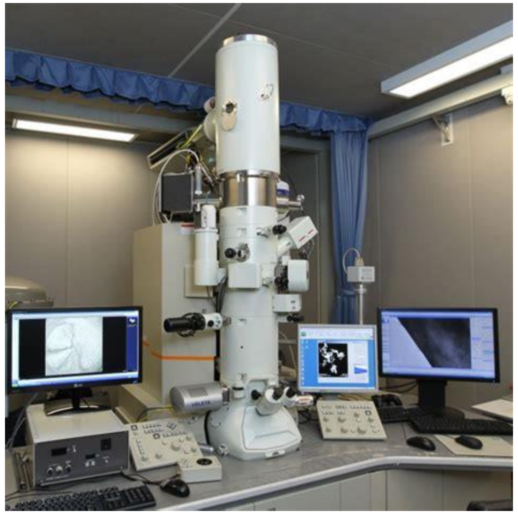


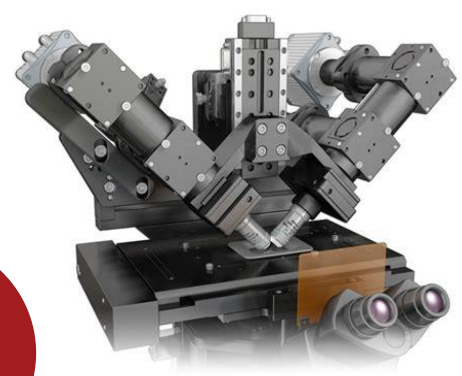
Hausdorf measures on overlaping sets extends notion of points to shapes – e.g., are mitochondria seen close to the synapse?

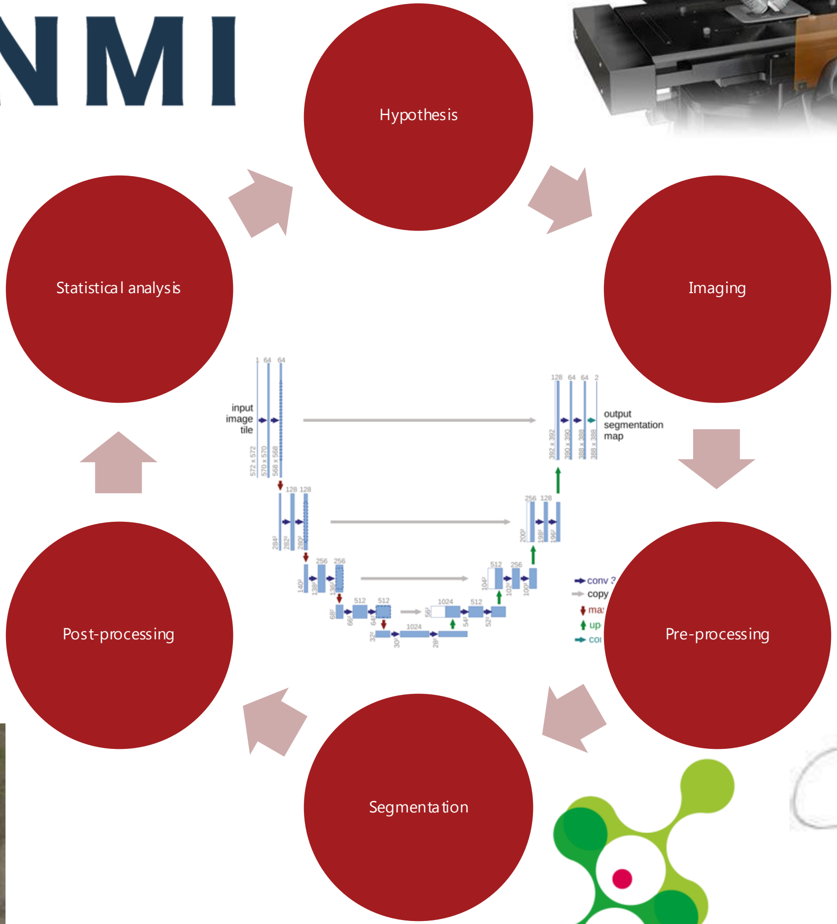
Topological data analysis



Filtrations brings topological concepts to measurements - e.g., what is the average tubular radius of complicated objects







EURO BIOIMAGING
Image life, discover the future

DANISH BIOIMAGING NETWORK

BNMI