

### +/- Message

```
## Loading required package: knitr
```

### +/- R Code

### +/- R Code

# Quantitative Big Imaging

author: Kevin Mader date: 3 April 2014 width: 1440 height: 900 css: ../template.css transition: rotate

ETHZ: 227-0966-00L

## Many Objects and Distributions

## Course Outline

- 20th February - Introductory Lecture
- 27th February - Filtering and Image Enhancement (A. Kaestner)
- 6th March - Basic Segmentation, Discrete Binary Structures
- 13th March - Advanced Segmentation
- 20th March - Analyzing Single Objects
- 27th March - Analyzing Complex Objects
- 3rd April - **Many Objects and Distributions**
- 10th April - Statistics and Reproducibility
- 17th April - Dynamic Experiments
- 8th May - Big Data
- 15th May - Guest Lecture - In-Operando Imaging of Batteries (V. Wood)
- 22th May - Project Presentations

## Literature / Useful References

### Books

- Jean Claude, Morphometry with R
    - Online (<http://link.springer.com/book/10.1007%2F978-0-387-77789-4>) through ETHZ
    - Buy it (<http://www.amazon.com/Morphometrics-R-Use-Julien-Claude/dp/038777789X>)
  - John C. Russ, “The Image Processing Handbook”, (Boca Raton, CRC Press)
    - Available online (<http://dx.doi.org/10.1201/9780203881095>) within domain ethz.ch (or proxy.ethz.ch / public VPN)
  - J. Weickert, Visualization and Processing of Tensor Fields
    - Online (<http://books.google.ch/books?id=ScLxPORMob4C&lpg=PA220&ots=mYleQbaVXP&dq=&pg=PA220#v=onepage&q&f=false>)
- 

## Papers / Sites

- Voronoi Tesselations
  - Ghosh, S. (1997). Tessellation-based computational methods for the characterization and analysis of heterogeneous microstructures. Composites Science and Technology, 57(9-10), 1187–1210
  - Wolfram Explanation (<http://mathworld.wolfram.com/VoronoiDiagram.html>)
- Self-Avoiding / Nearest Neighbor
  - Schwarz, H., & Exner, H. E. (1983). The characterization of the arrangement of feature centroids in planes and volumes. Journal of Microscopy, 129(2), 155–169.
  - Kubitscheck, U. et al. (1996). Single nuclear pores visualized by confocal microscopy and image processing. Biophysical Journal, 70(5), 2067–77.
- Alignment / Distribution Tensor
  - Mader, K. et al (2013). A quantitative framework for the 3D characterization of the osteocyte lacunar system. Bone, 57(1), 142–154
- Two point correlation
  - Dinis, L., et. al. (2007). Analysis of 3D solids using the natural neighbour radial point interpolation method. Computer Methods in Applied Mechanics and Engineering, 196(13-16)

## Previously on QBI ...

- Image Enhancement
  - Highlighting the contrast of interest in images
  - Minimizing Noise
- Understanding image histograms
- Automatic Methods
- Component Labeling
- Single Shape Analysis
- Complicated Shapes

# Outline

- Motivation (Why and How?)
- Local Environment
  - Neighbors
  - Voronoi Tesselation
  - Distribution Tensor
- Global Environment
  - Alignment
  - Self-Avoidance
  - Two Point Correlation Function

## What do we start with?

Going back to our original cell image

1. We have been able to get rid of the noise in the image and find all the cells (lecture 2-4)
2. We have analyzed the shape of the cells using the shape tensor (lecture 5)
3. We even separated cells joined together using Watershed (lecture 6)

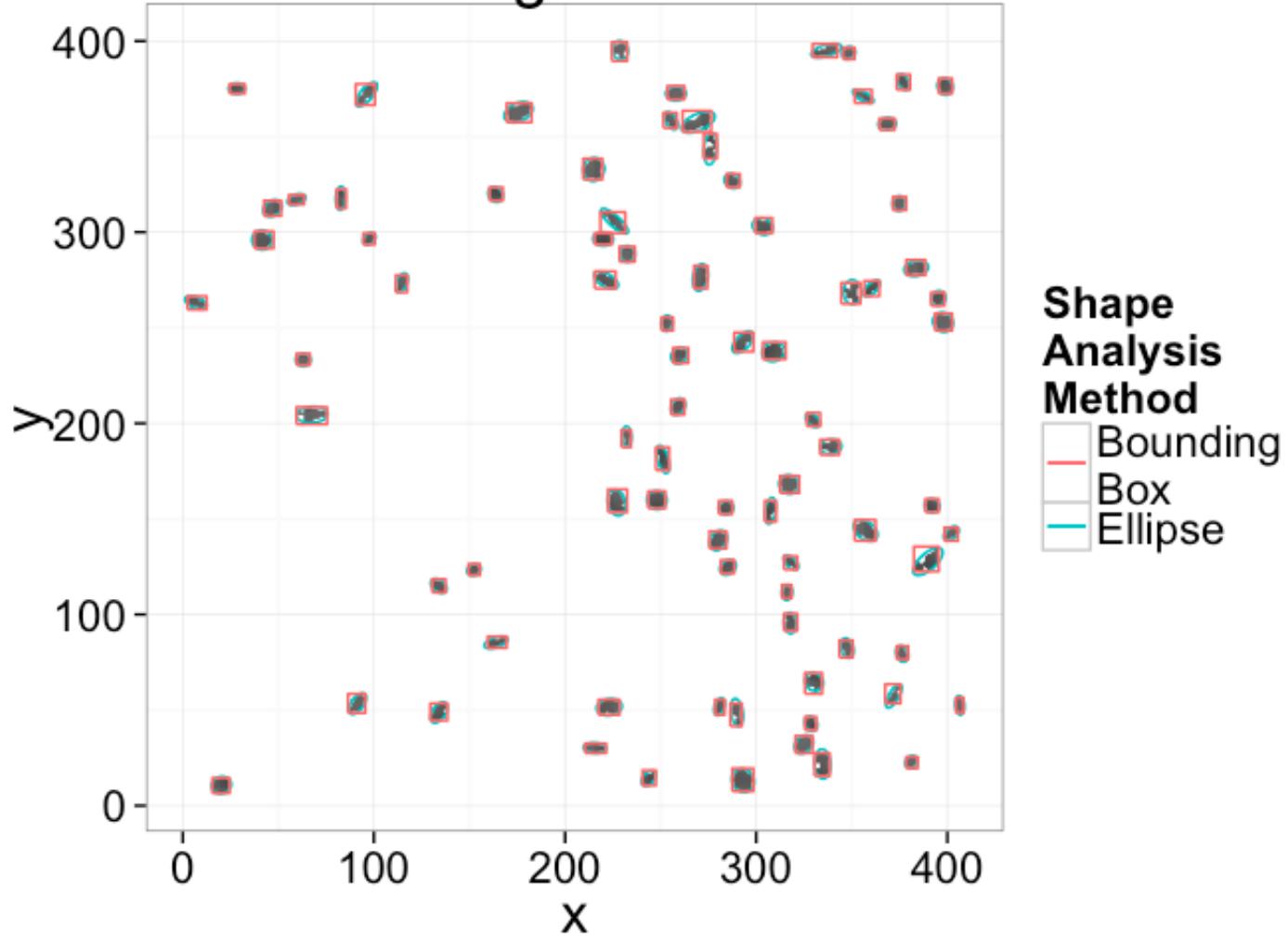
We can characterize the sample and the average and standard deviations of volume, orientation, surface area, and other metrics

## Motivation (Why and How?)

With all of these images, the first step is always to understand exactly what we are trying to learn from our images.

+/- R Code

## Single Cell



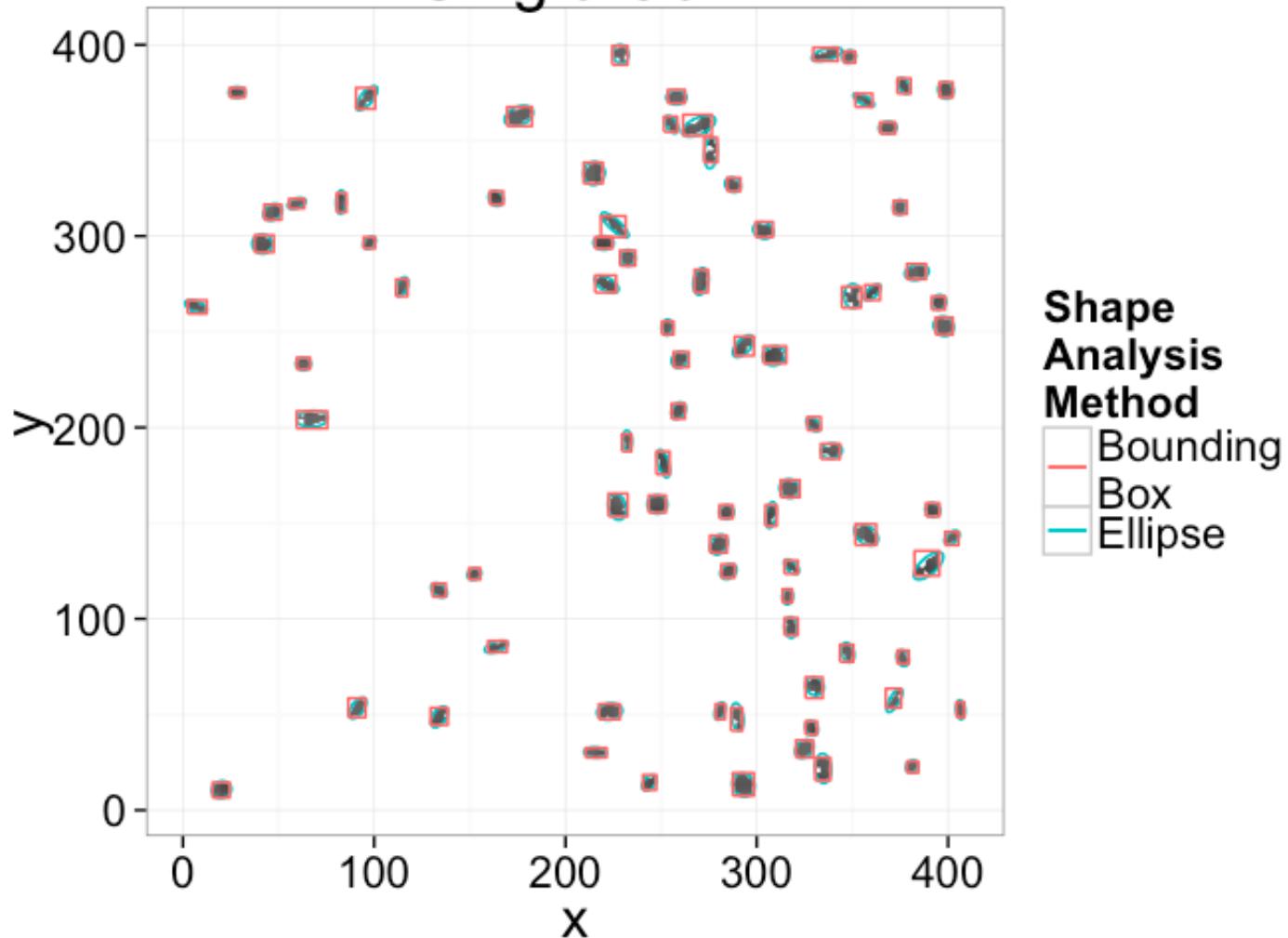
1. We want to know how many cells are alive
  - Maybe small cells are dead and larger cells are alive → examine the volume distribution
  - Maybe living cells are round and dead cells are really spiky and pointy → examine anisotropy
2. We want to know where the cells are alive or most densely packed

- We can visually inspect the sample (maybe even color by volume)
- We can examine the raw positions (x,y,z) but what does that really tell us?
- We can make boxes and count the cells inside each one
- How do we compare two regions in the same sample or even two samples?

# Motivation (continued)

+/- R Code

## Single Cell



1. We want to know how the cells are communicating

- Maybe physically connected cells (touching) are communicating → watershed
- Maybe cells oriented the same direction are communicating → *average?* orientation
- Maybe cells which are close **enough** are communicating → ?

- Maybe cells form hub and spoke networks → ?
2. We want to know how the cells are nourished
- Maybe closely packed cells are better nourished → count cells in a box?
  - Maybe cells are oriented around canals which supply them → ?

## So what do we still need

1. A way for counting cells in a region and estimating density without creating arbitrary boxes
  2. A way for finding out how many cells are *near* a given cell, it's nearest neighbors
  3. A way for quantifying how far apart cells are and then comparing different regions within a sample
  4. A way for quantifying and comparing orientations
    - the mean of  $0^\circ$  and  $180^\circ$  =  $90^\circ$
    - the distance between  $-180^\circ$  and  $179^\circ$  is  $359^\circ$
    - since we have not defined a tip or head,  $0^\circ$  and  $180^\circ$  are actually the same
- 

## What would be really great?

A tool which could be adapted to answering a large variety of problems

- multiple types of structures
- multiple phases

## Destructive Measurements

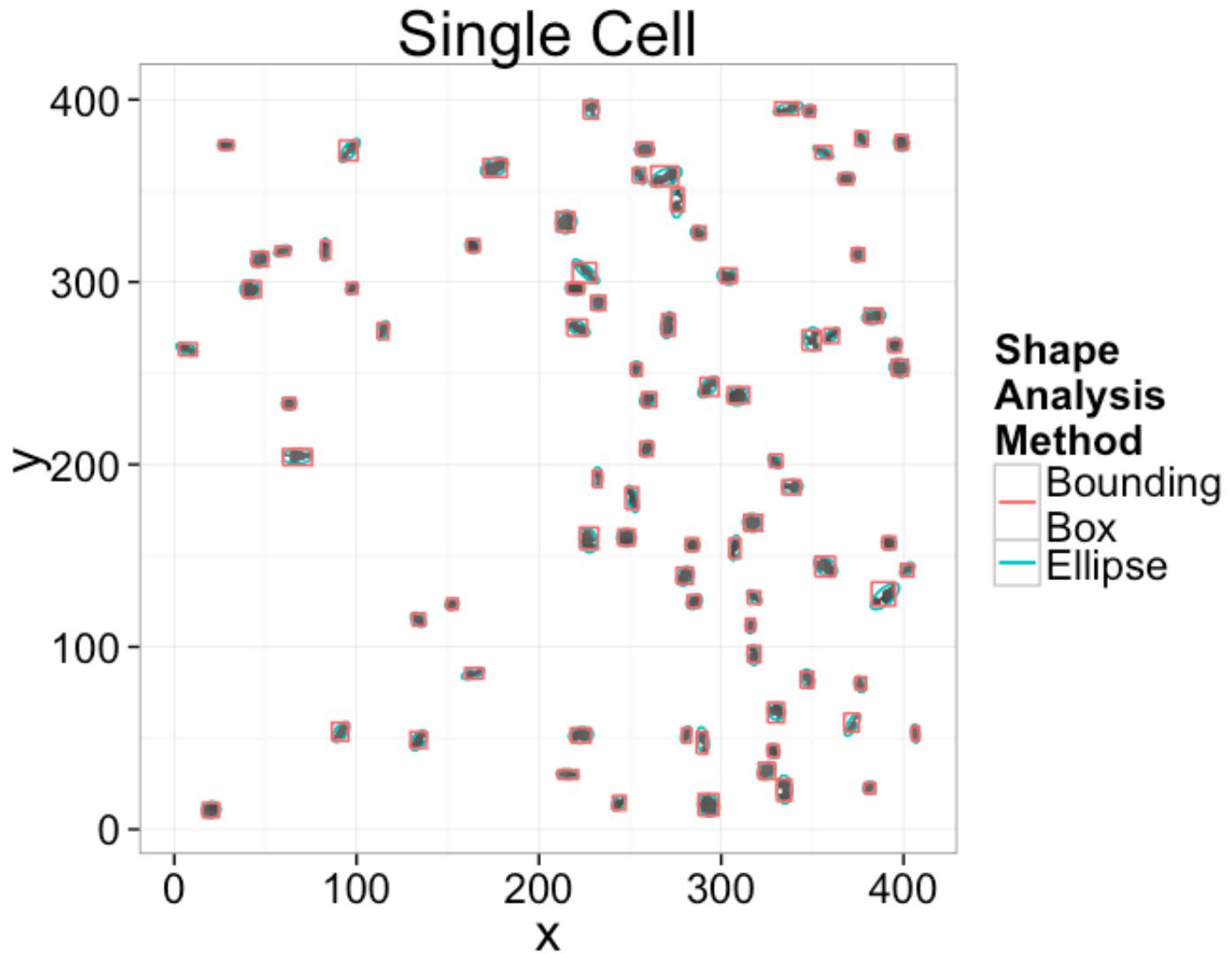
With most imaging techniques and sample types, the task of measurement itself impacts the sample.

- Even techniques like X-ray tomography which *claim* to be non-destructive still impart significant to lethal doses of X-ray radiation for high resolution imaging
  - Electron microscopy, auto-tome-based methods, histology are all markedly more destructive and make longitudinal studies impossible
  - Even when such measurements are possible
    - Registration can be a difficult task and introduce artifacts
- 

## Why is this important?

- techniques which allow us to compare different samples of the same type.
- are sensitive to common transformations
  - Sample B after the treatment looks like Sample A stretched to be 2x larger
  - The volume fraction at the center is higher than the edges but organization remains the same

# Ok, so now what?



+/- R Code

<b>x</b>	<b>y</b>	<b>vx</b>	<b>vy</b>
20.19	10.69	-0.95	-0.30
20.19	10.69	0.30	-0.95
293.08	13.18	-0.50	0.86
293.08	13.18	-0.86	-0.50
243.81	14.23	0.68	0.74
243.81	14.23	-0.74	0.68
...			

So if we want to know the the mean or standard deviations of the position or orientations we can analyze them easily.

### +/- R Code

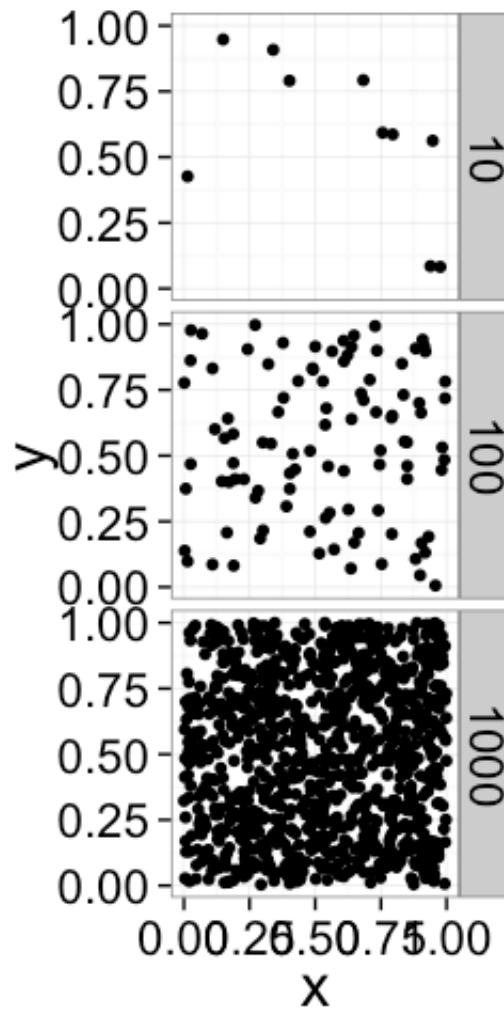
<b>id</b>	<b>Min.</b>	<b>1st Qu.</b>	<b>Median</b>	<b>Mean</b>	<b>3rd Qu.</b>	<b>Max.</b>
x	6.90	216.00	281.00	258.00	339.00	406.00
y	10.70	112.00	221.00	209.00	313.00	395.00
Length	1.06	1.57	1.95	2.08	2.41	4.32
vx	-1.00	-0.94	-0.70	-0.42	0.07	0.71
vy	-1.00	-0.70	0.02	0.04	0.71	1.00
Theta	-180.00	-134.00	-0.50	-4.67	131.00	178.00

- But what if we want more or other information?

## Calculating Density

One of the first metrics to examine with distribution is density → how many objects in a given region or volume.

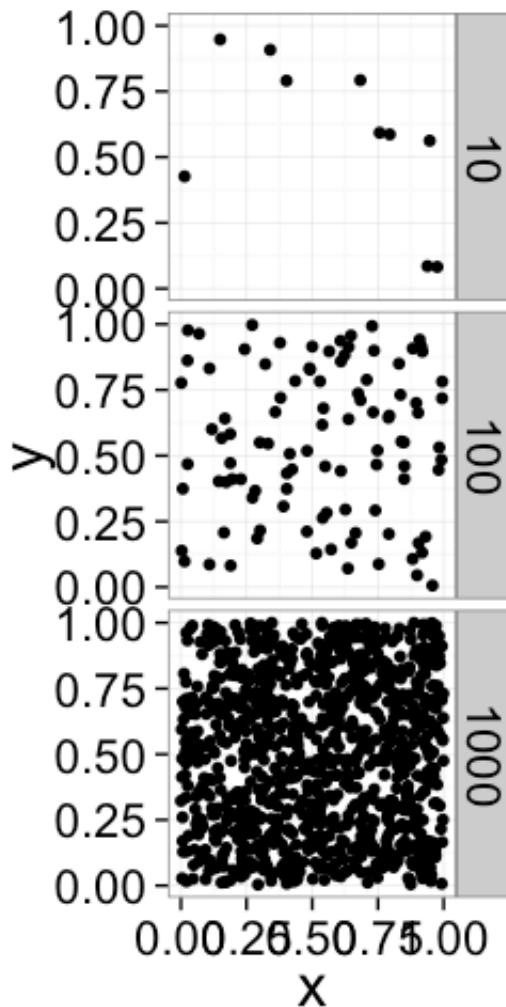
It is deceptively easy to calculate involving the ratio of the number of objects divided by the volume.



+/- R Code

It doesn't tell us much, many very different systems with the *same* density and what if we want the density of a single point? Does that even make sense?

+/- R Code



# Neighbors

## Definition

Oxford American → be situated next to or very near to (another)

- Does not sound very scientific
- How close?
  - Touching, closer than anything else?

## Nearest Neighbor (distance)

+/- R Code

Given a set of objects with centroids at

$$\mathbf{P} = [\vec{x}_0, \vec{x}_1, \dots, \vec{x}_i]$$

We can define the nearest neighbor as the position of the object in our set which is closest

$$\vec{\text{NN}}(\vec{y}) = \operatorname{argmin}(\|\vec{y} - \vec{x}\| \forall \vec{x} \in \mathbf{P} - \vec{y})$$

We define the distance as the Euclidean distance from the current point to that point, and the angle as the

$$\text{NND}(\vec{y}) = \min(\|\vec{y} - \vec{x}\| \forall \vec{x} \in \mathbf{P} - \vec{y})$$

$$\text{NN}\theta(\vec{y}) = \tan^{-1} \frac{(\vec{\text{NN}} - \vec{y}) \cdot \vec{j}}{(\vec{\text{NN}} - \vec{y}) \cdot \vec{i}}$$

## Nearest Neighbor Definition

So examining a simple starting system like a grid, we already start running into issues.

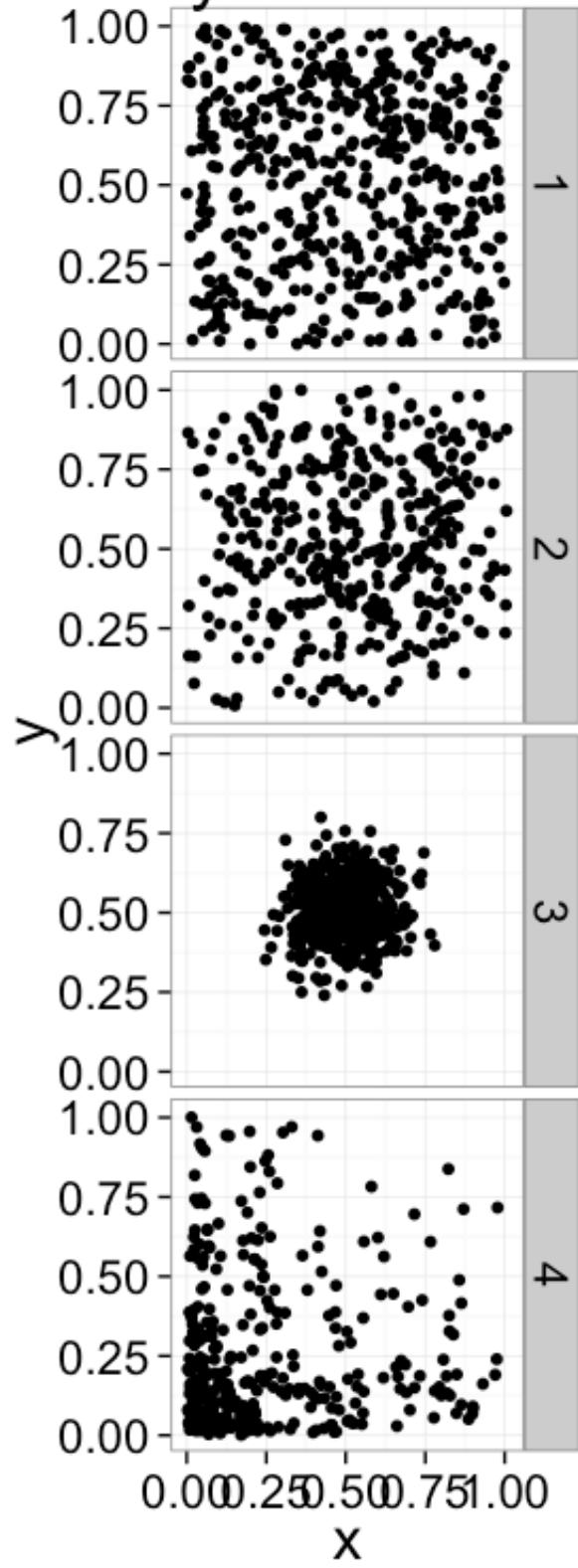
- In a perfect grid like structure each object has 4 equidistant neighbors (6 in 3D)
- Which one is closest?

We thus add an additional clause (only relevant for simulated data) where if there are multiple equidistant neighbors, a *nearest* is chosen randomly

This ensures when we examine the orientation distribution ( $\text{NN}\theta$ ) of the neighbors it is evenly distributed

+/- R Code

Density = 500 N/mm<sup>2</sup>



# In-Silico Systems

For the rest of these sections we will repeatedly use several simple in-silico systems to test our methods and try to better understand the kind of results we obtain from them.

- Compression

- The most simple system simply involves a scaling in every direction by  $\alpha$
- $\alpha < 1$  the system is compressed
- $\alpha > 1$  the system is expanded

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \alpha \begin{bmatrix} x \\ y \end{bmatrix}$$


---

- Shearing

- Slightly more complicated system where objects are shifted based on their location using a slope of  $\alpha$

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} 1 & \alpha \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

# In-Silico Systems (Continued)

- Stretch

- A non-evenly distributed system with a parameter  $\alpha$  controlling if objects bunch near the edges or the center. A maximum distance  $m$  is defined as the magnitude of the largest offset in x or y
- + Same total volume just arranged differently

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \text{sign}(x) \left( \frac{|x|}{m} \right)^\alpha m \\ \text{sign}(y) \left( \frac{|y|}{m} \right)^\alpha m \end{bmatrix}$$


---

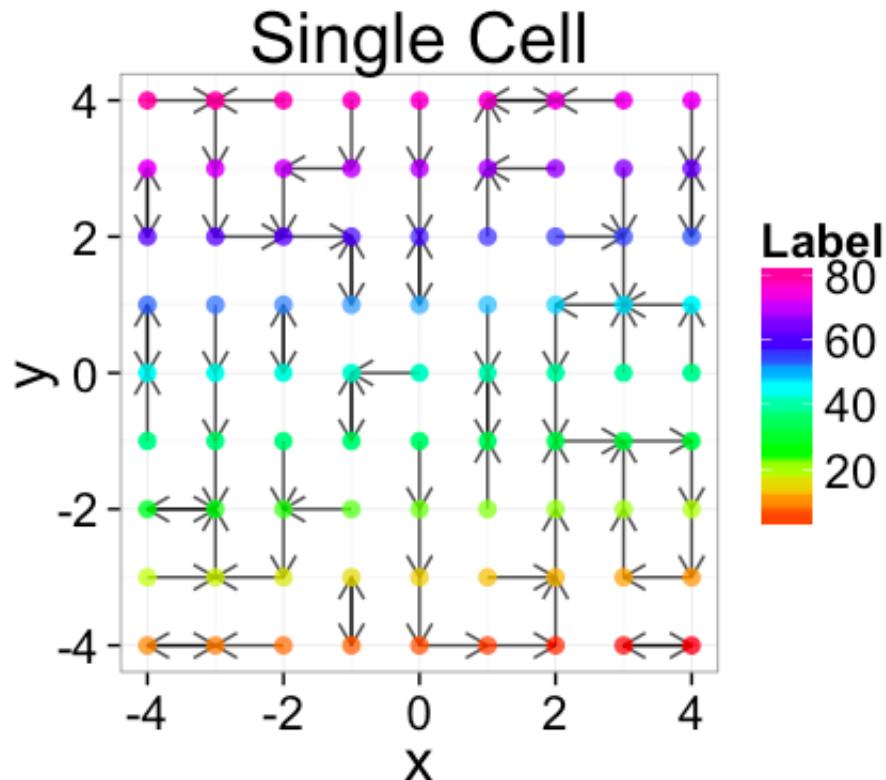
- Swirl

- A transformation where the points are rotated more based on how far away they are from the center and the slope of the swirl ( $\alpha$ ),

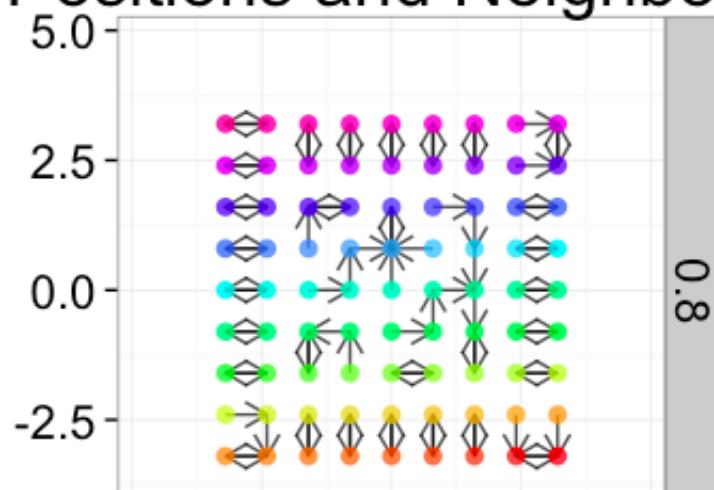
$$\theta(x, y) = \alpha \sqrt{x^2 + y^2}$$

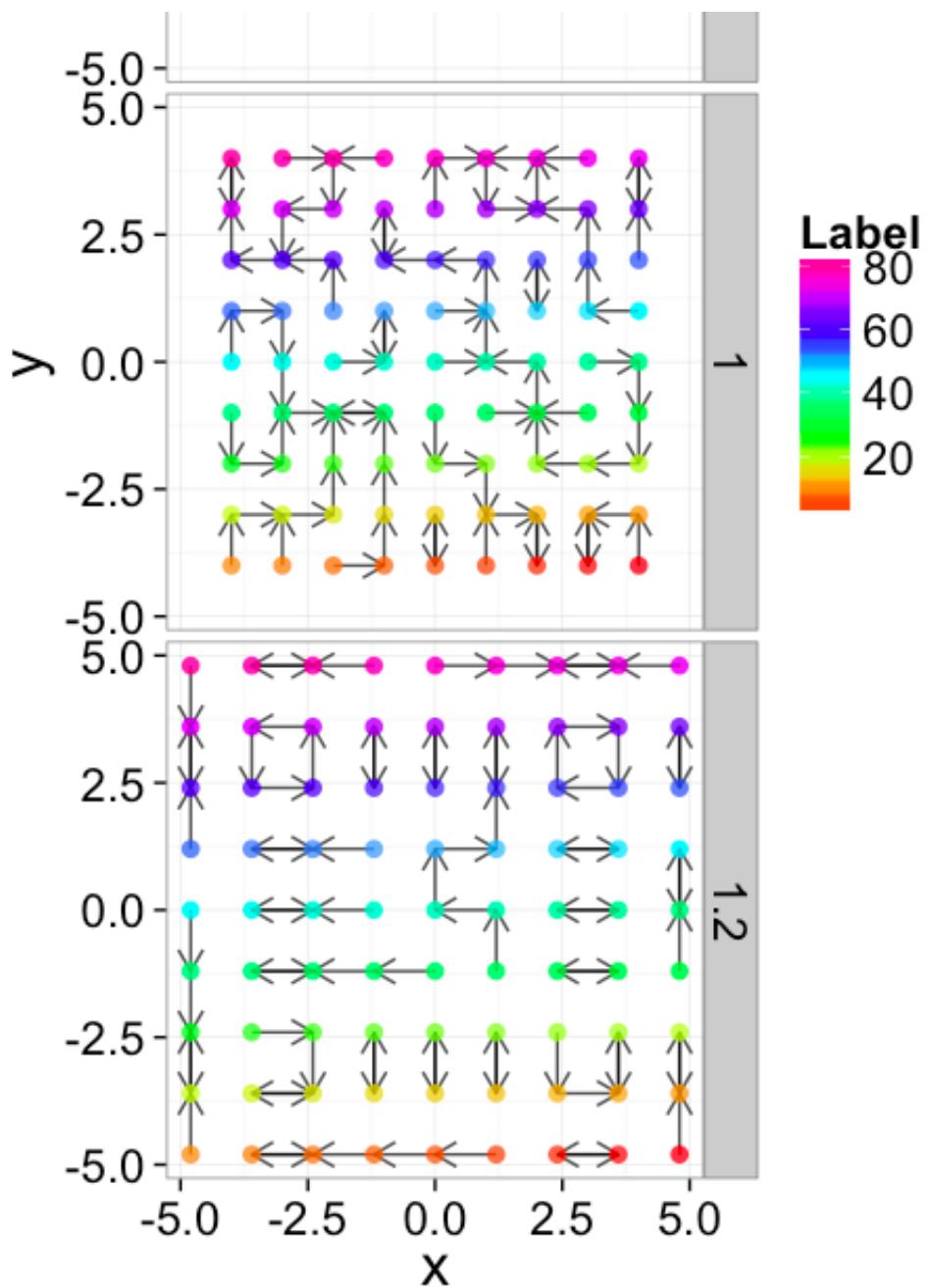
$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos \theta(x, y) & -\sin \theta(x, y) \\ \sin \theta(x, y) & \cos \theta(x, y) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

# Examining Compression



## Positions and Neighbors

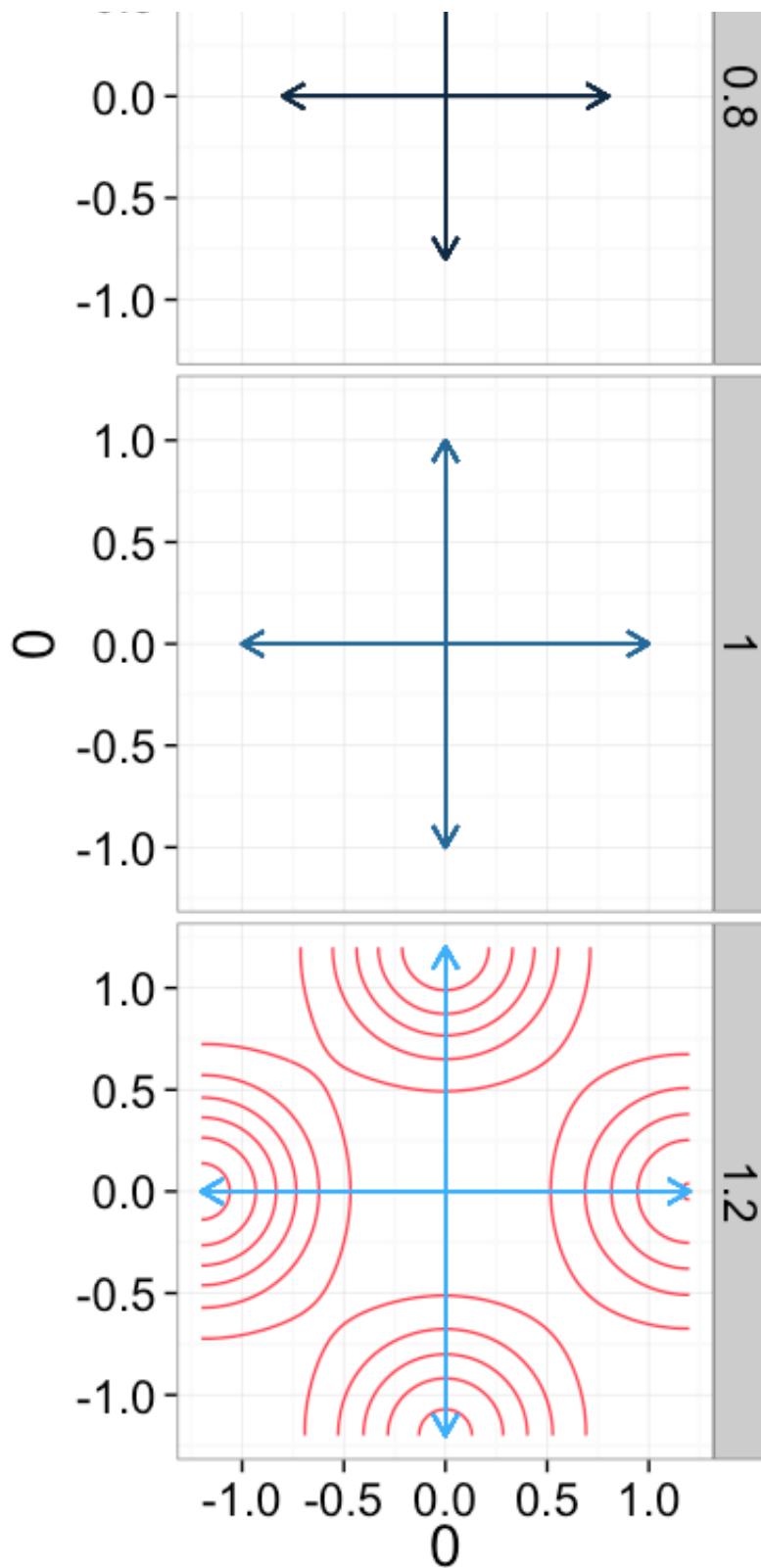




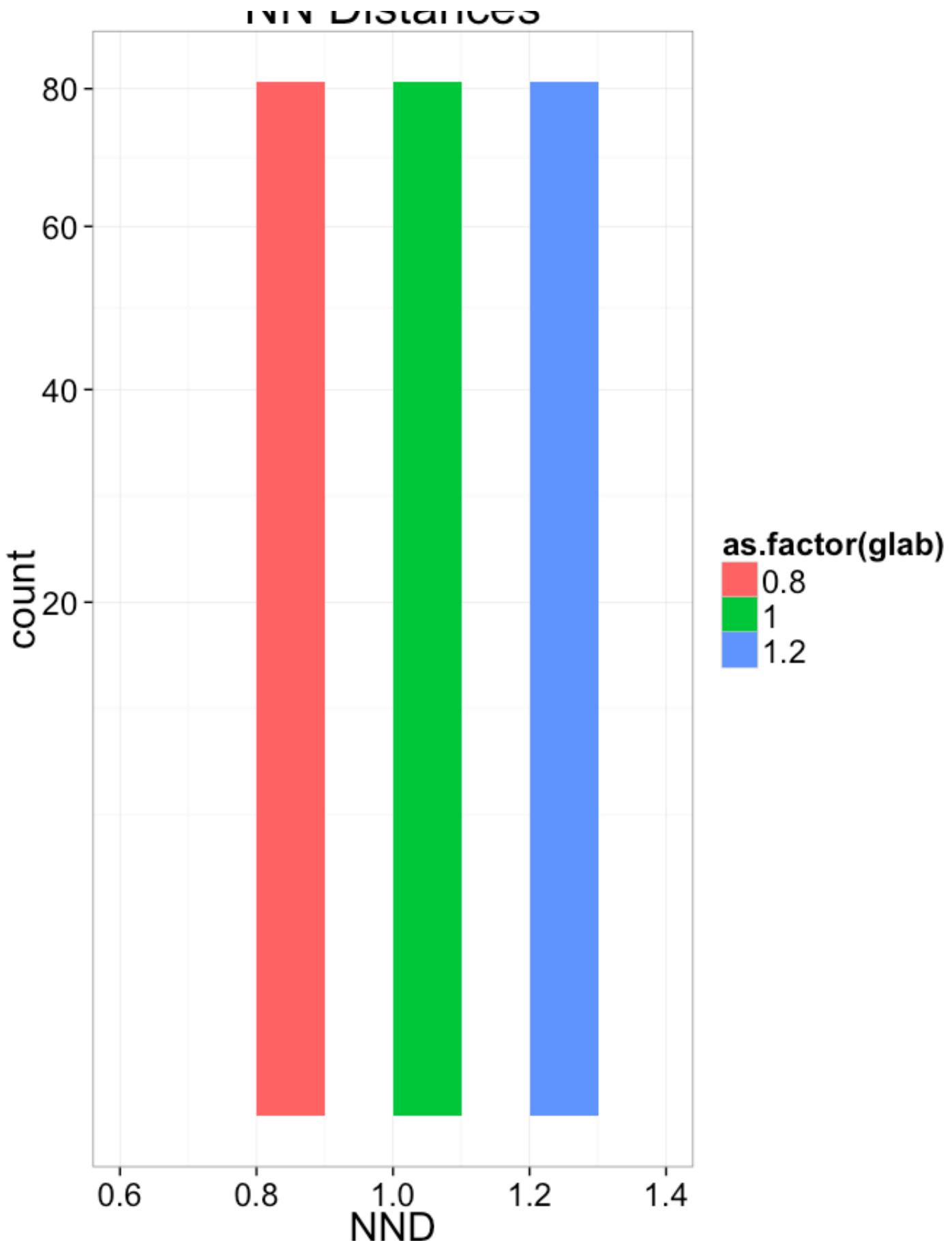
## Compression Distributions

### Nearest Neighbor Orientations



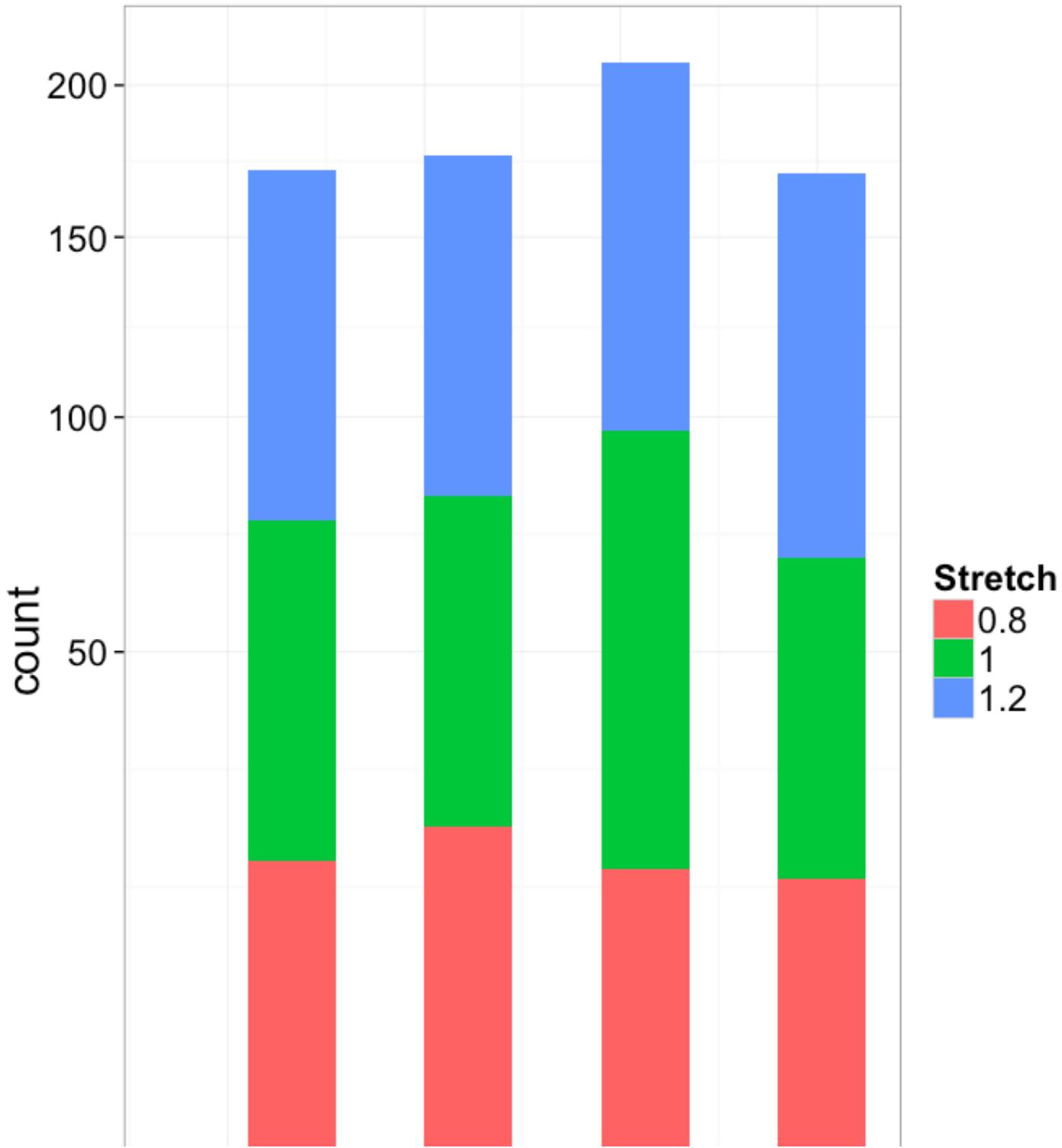


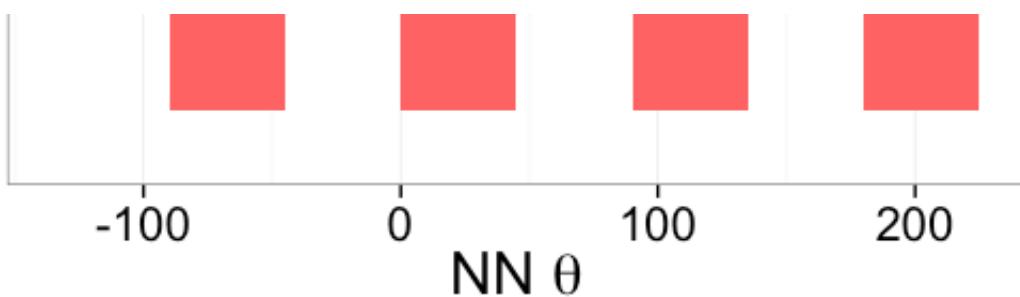
## NN Distances



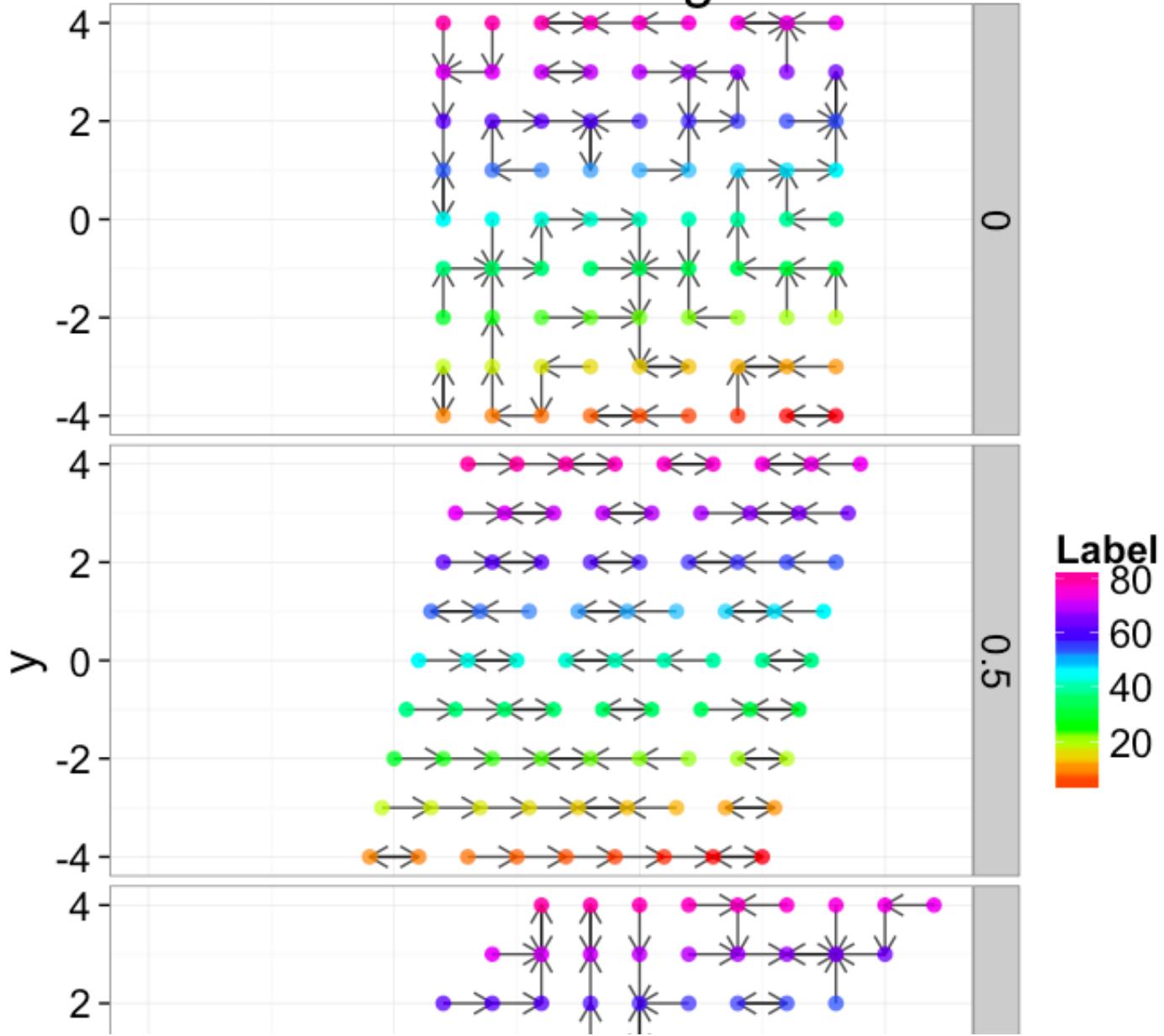
# Examining Different Shears

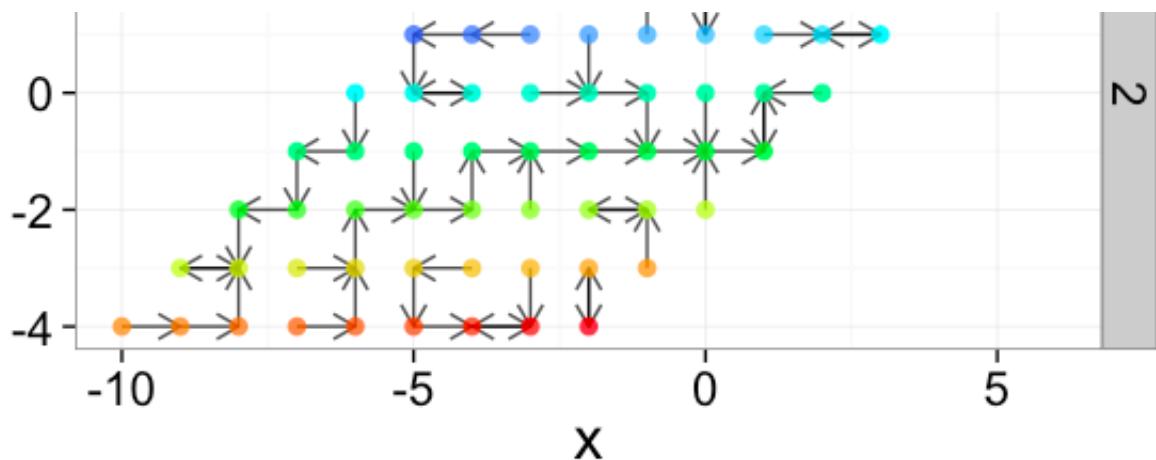
NN Orientation





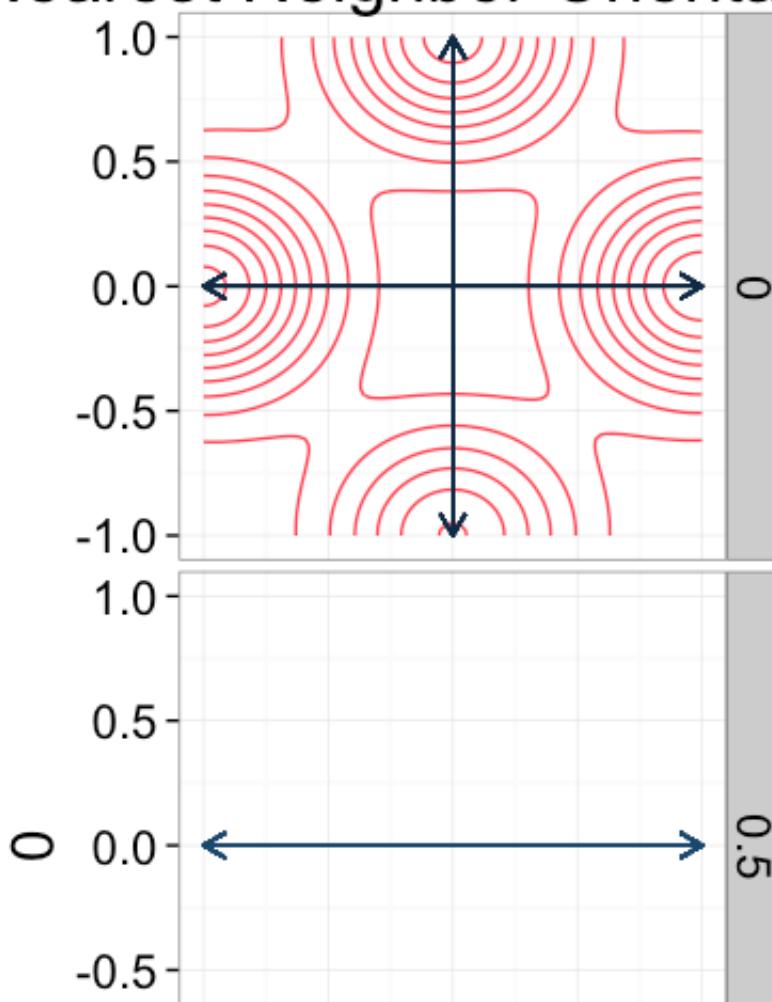
## Positions and Neighbors

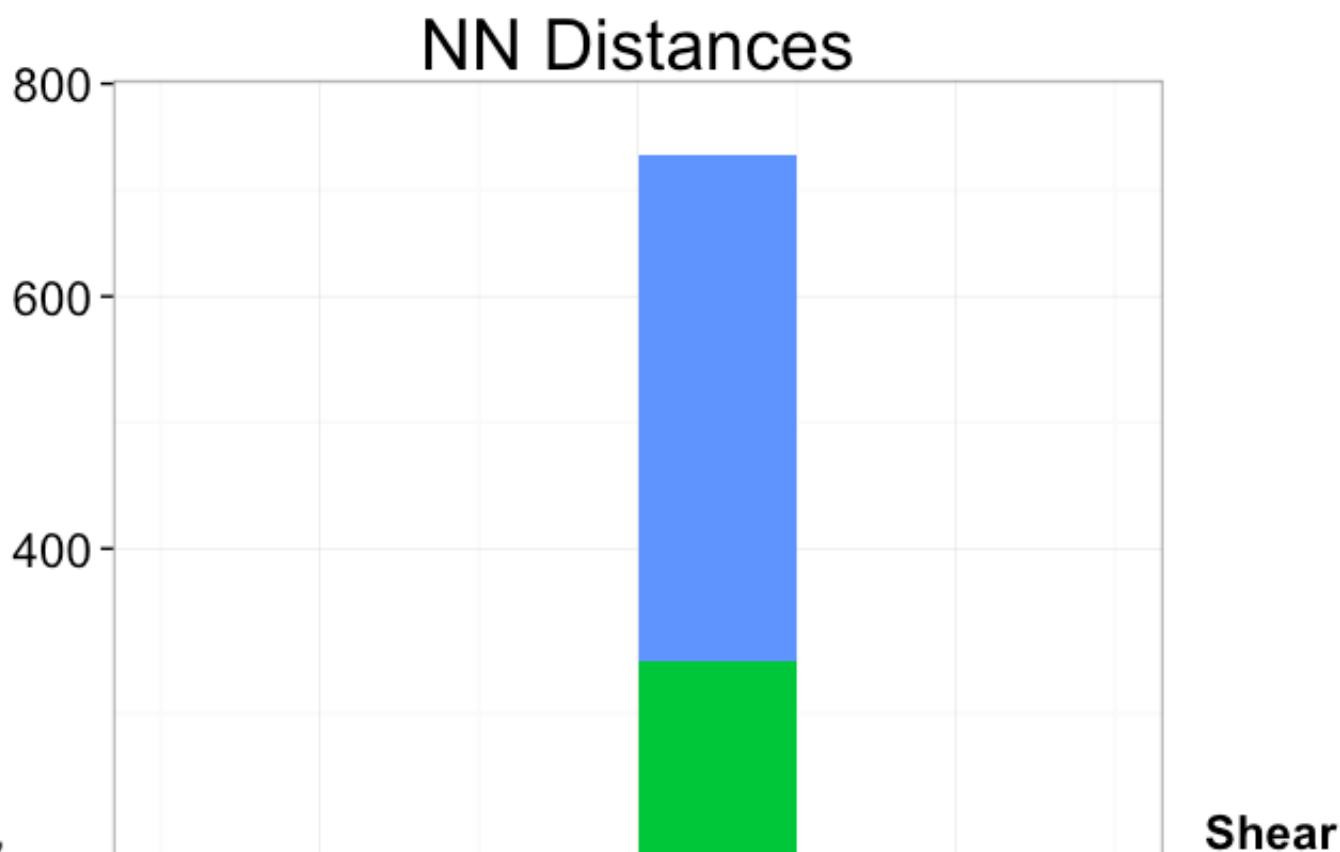
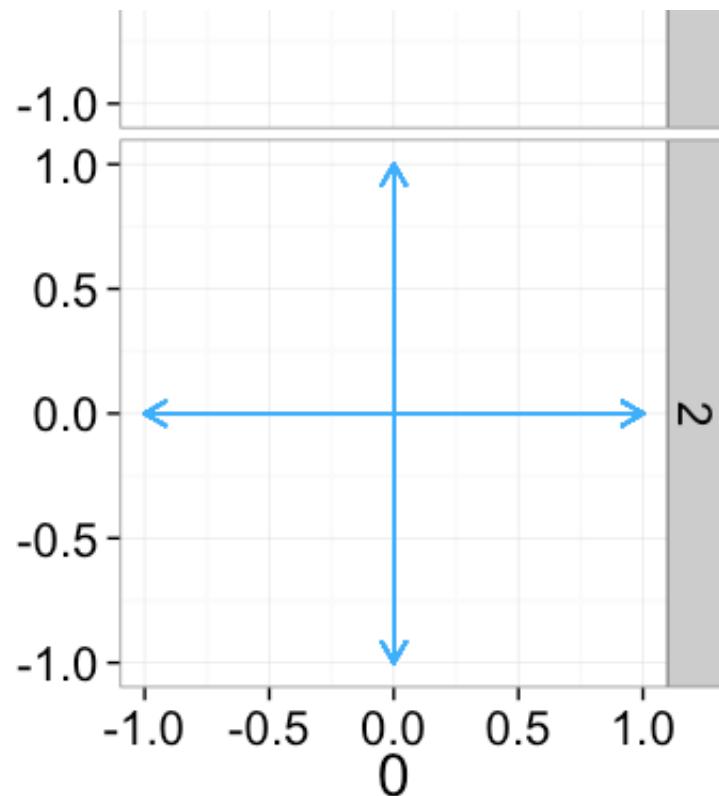


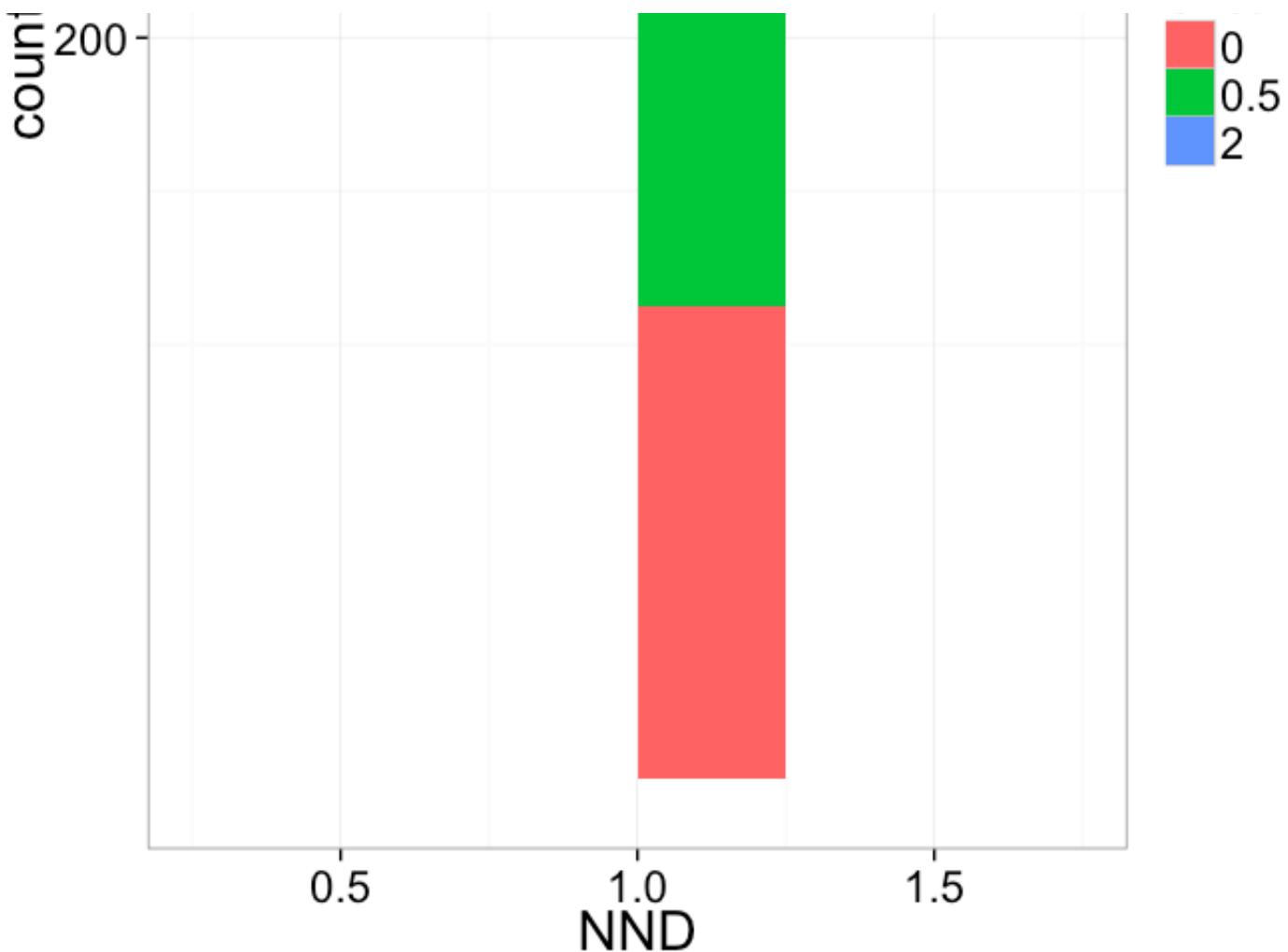


## Shear Distributions

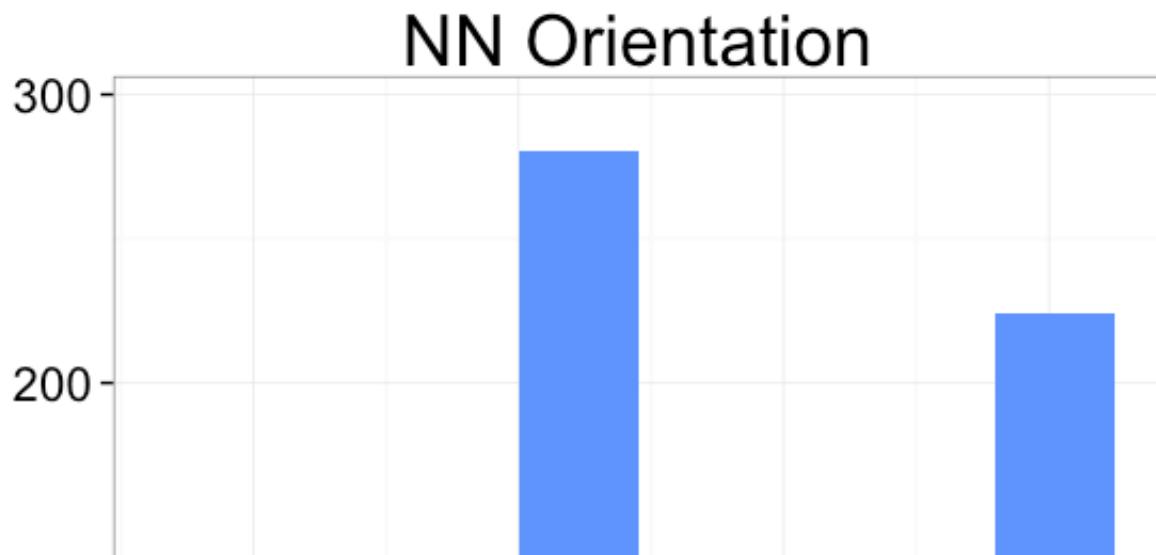
Nearest Neighbor Orientations

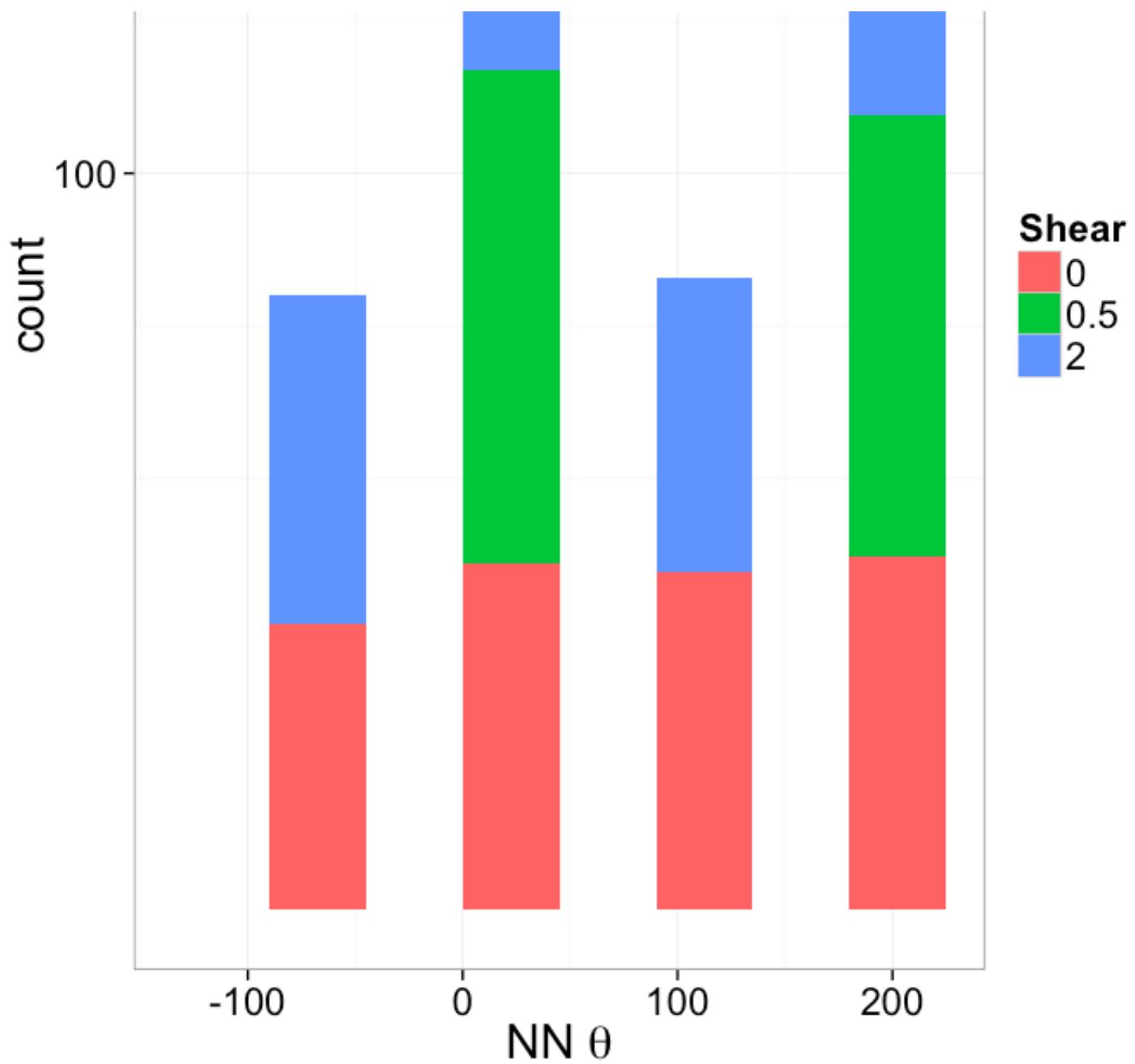




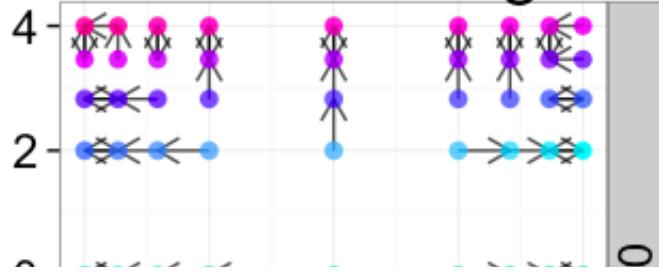


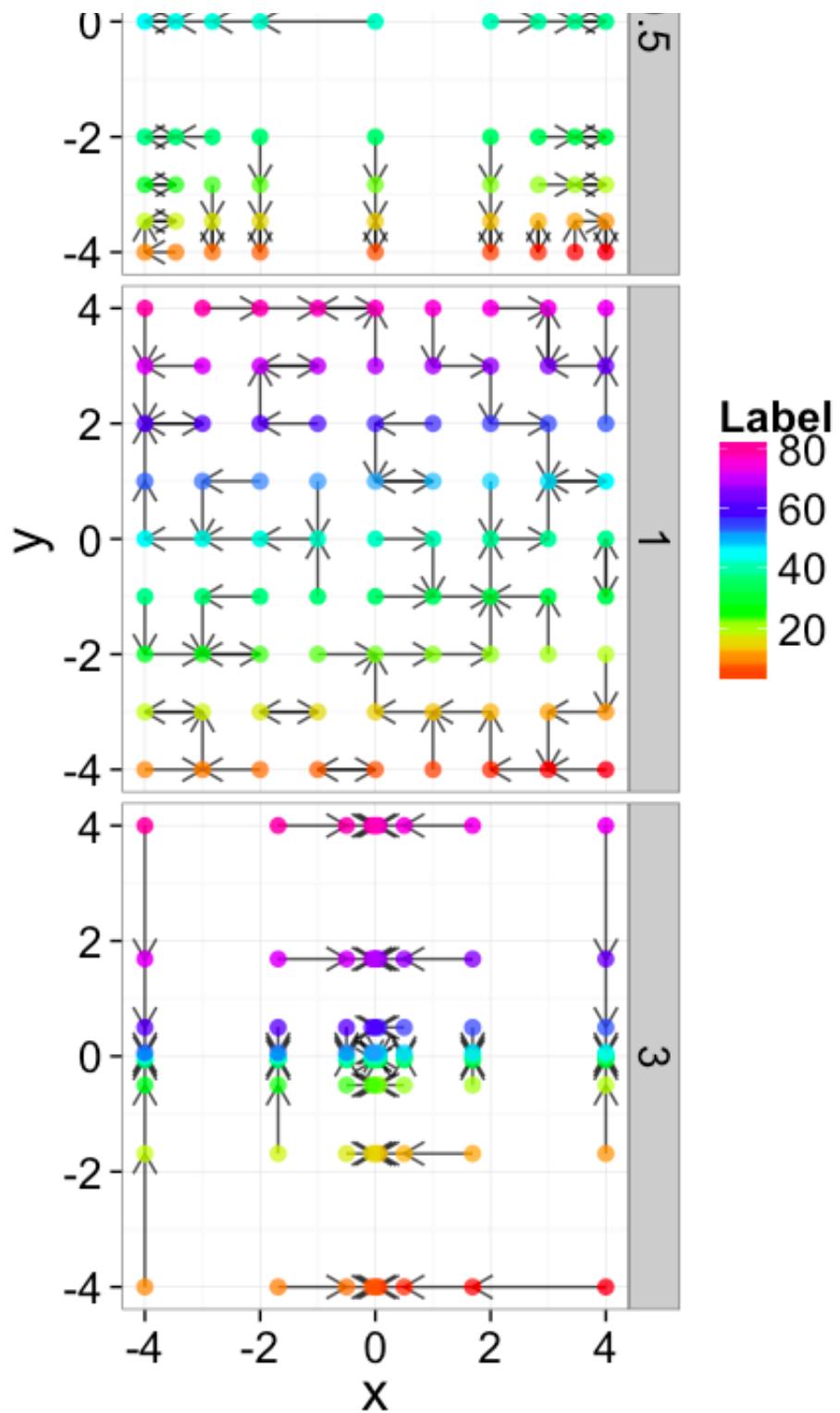
## Examining Different Stretches





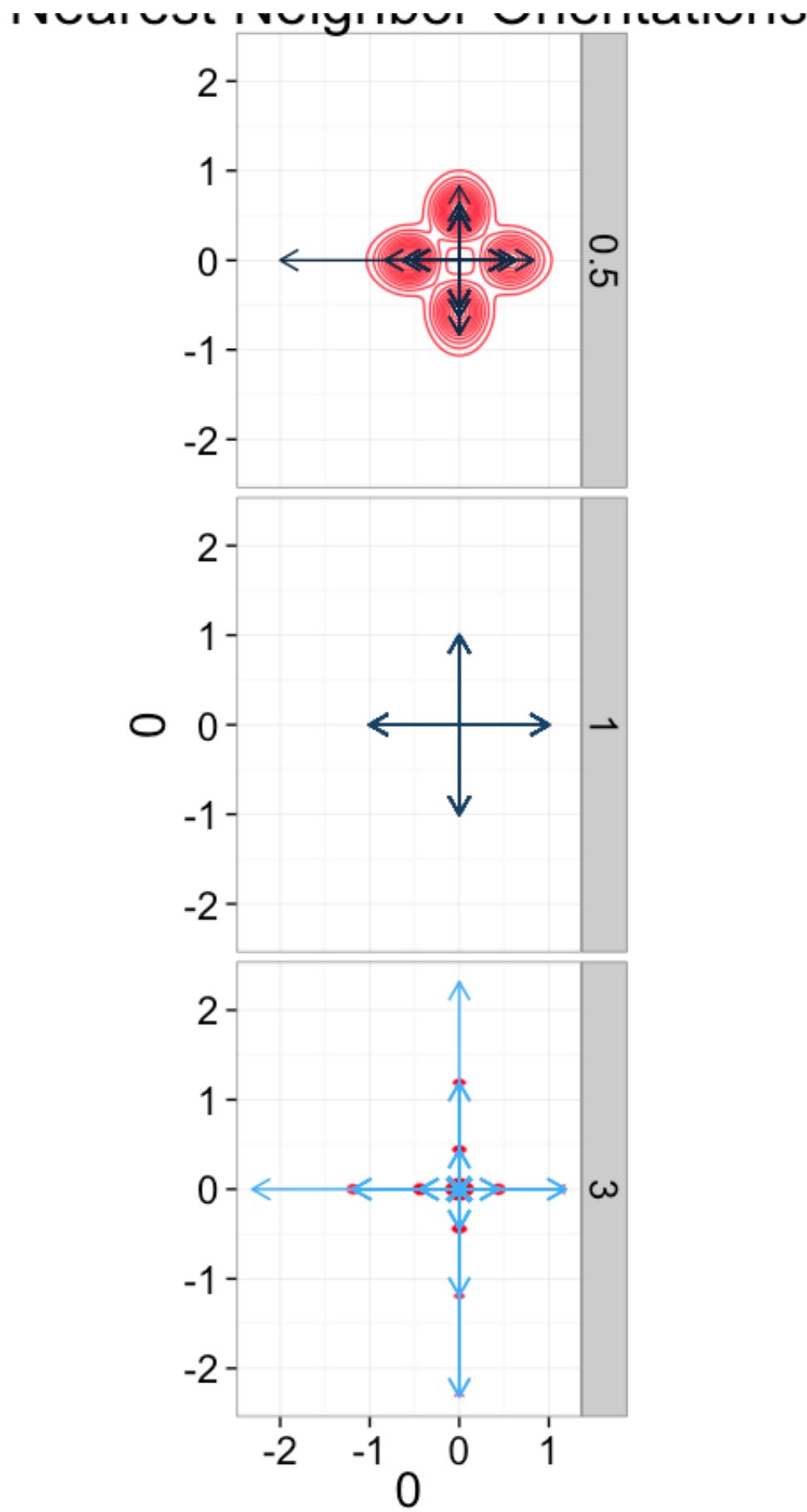
## Positions and Neighbors



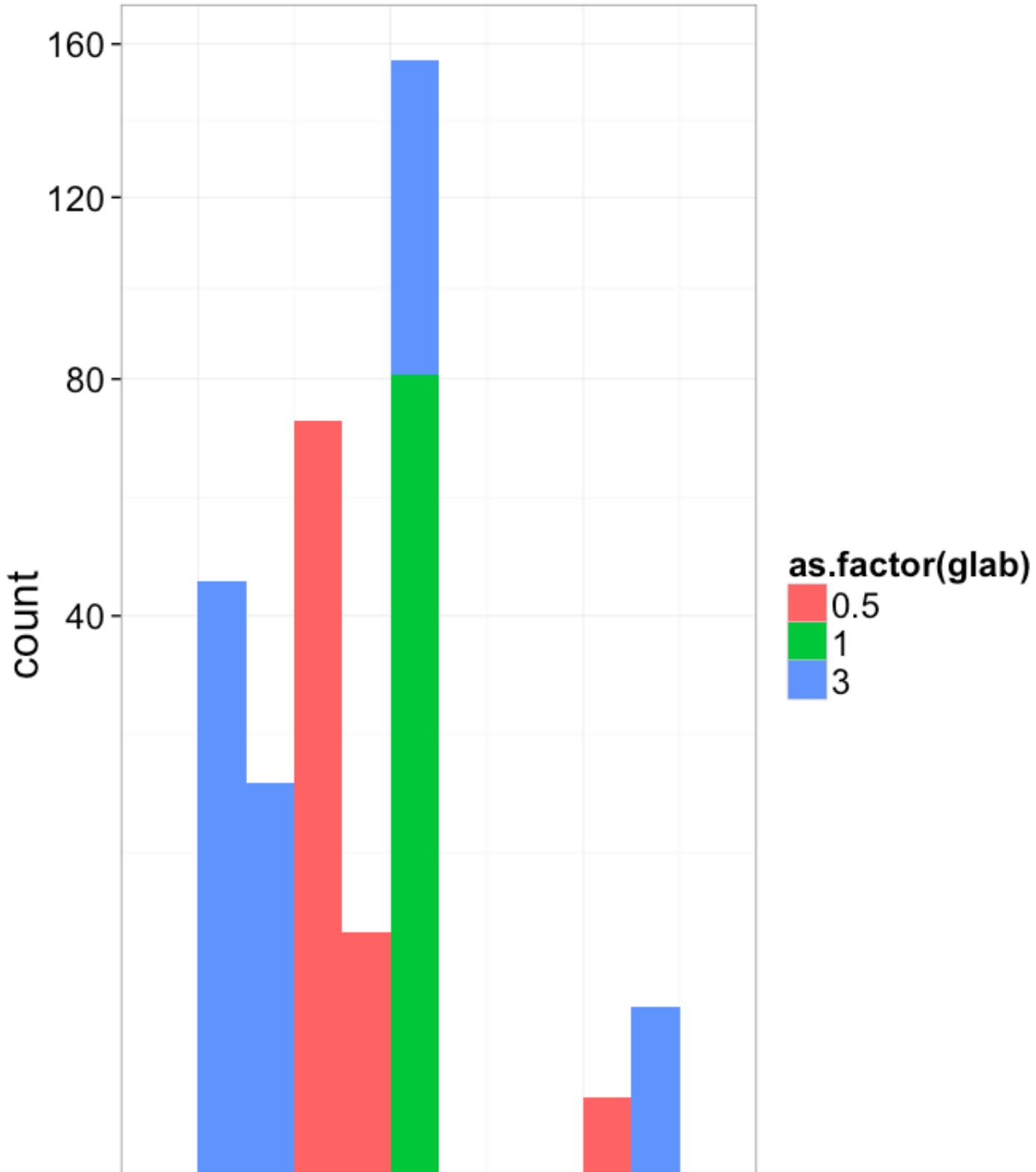


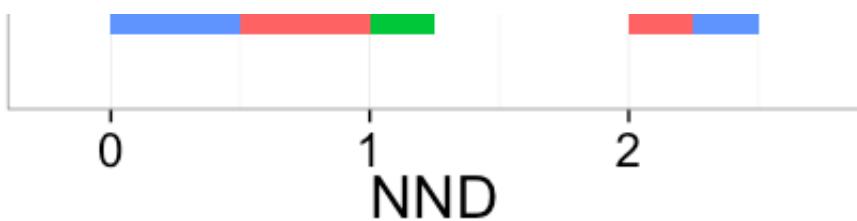
# Stretch Distributions

Nearest Neighbor Orientations



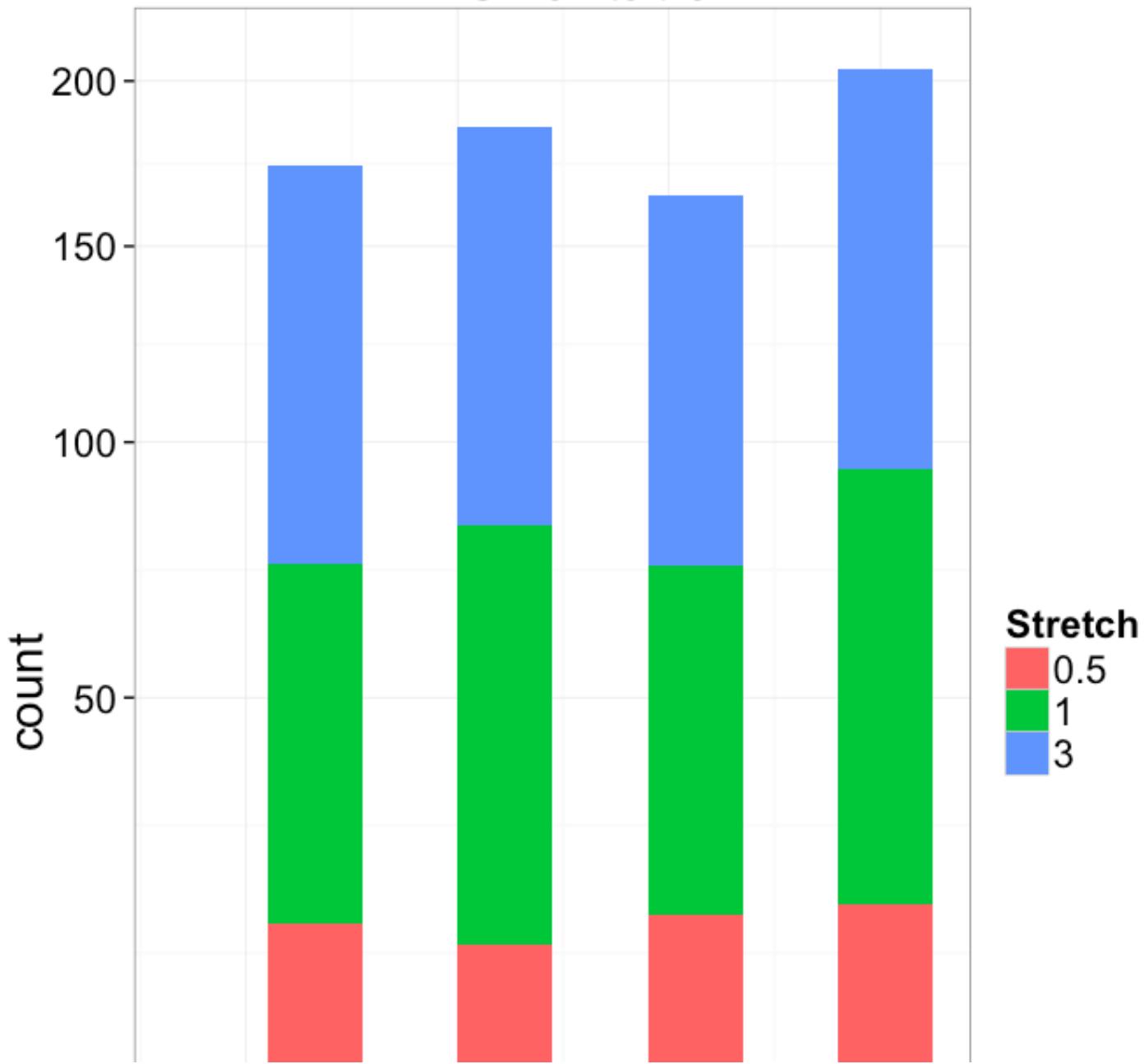
## NN Distances

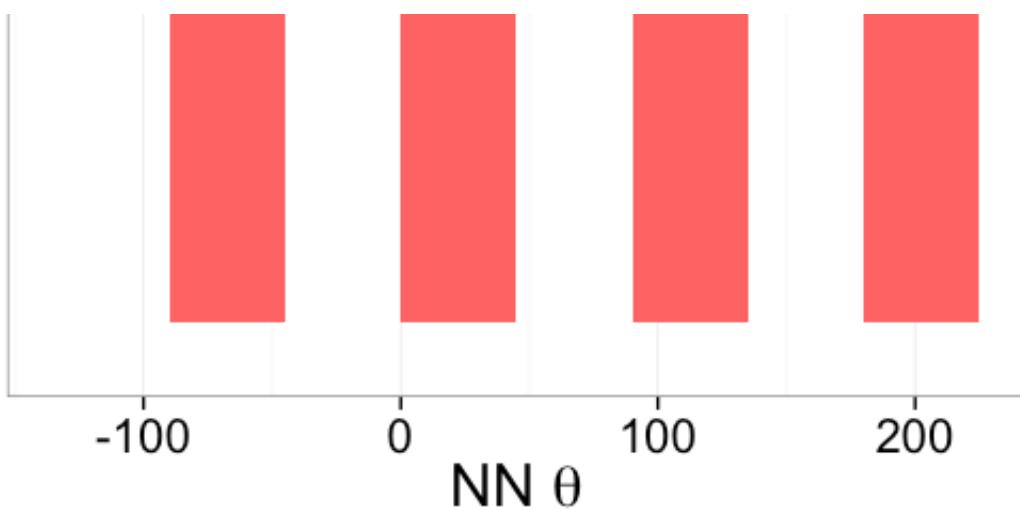




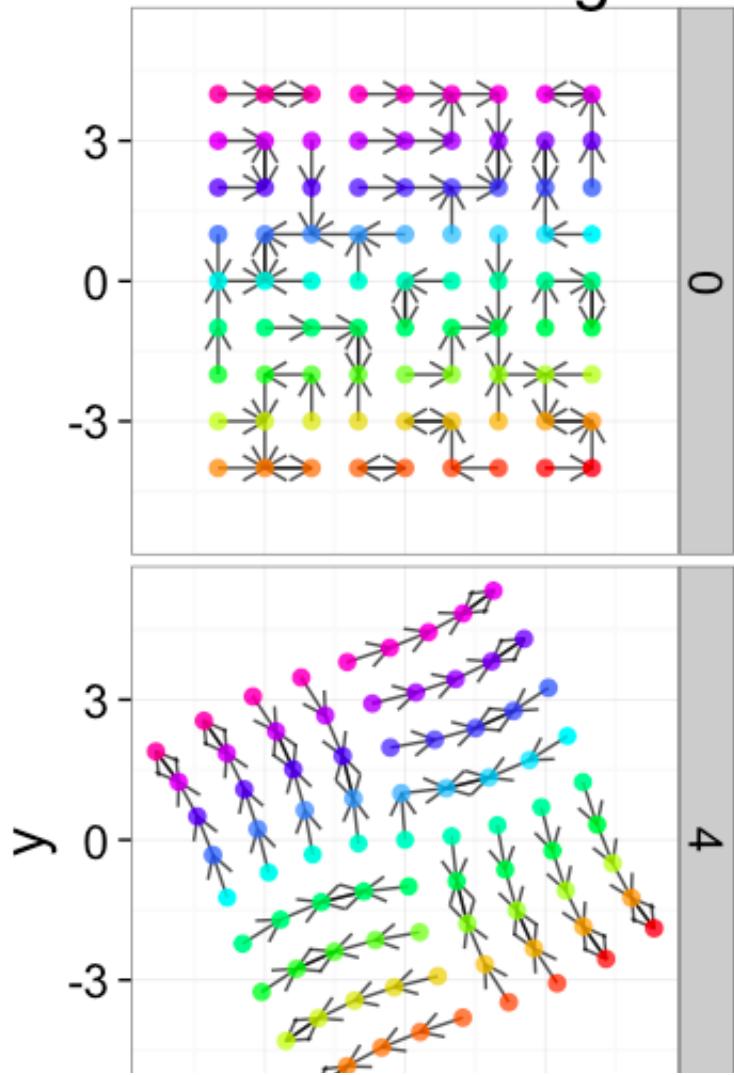
## Examining Swirl Systems

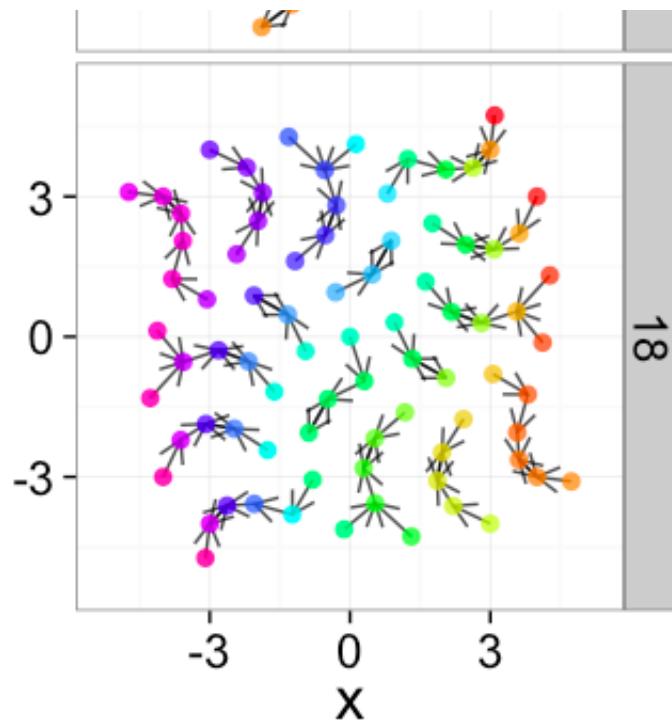
NN Orientation





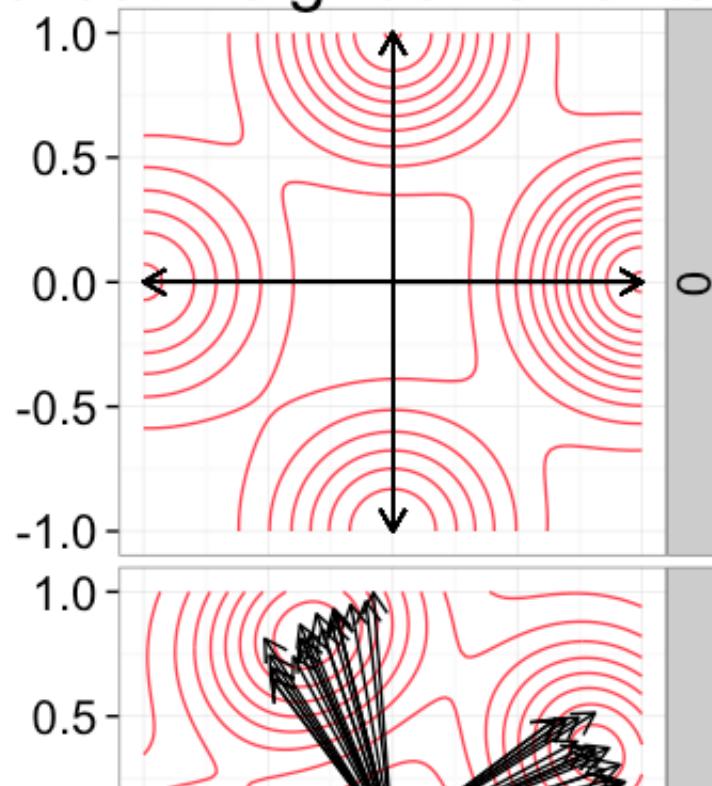
## Positions and Neighbors

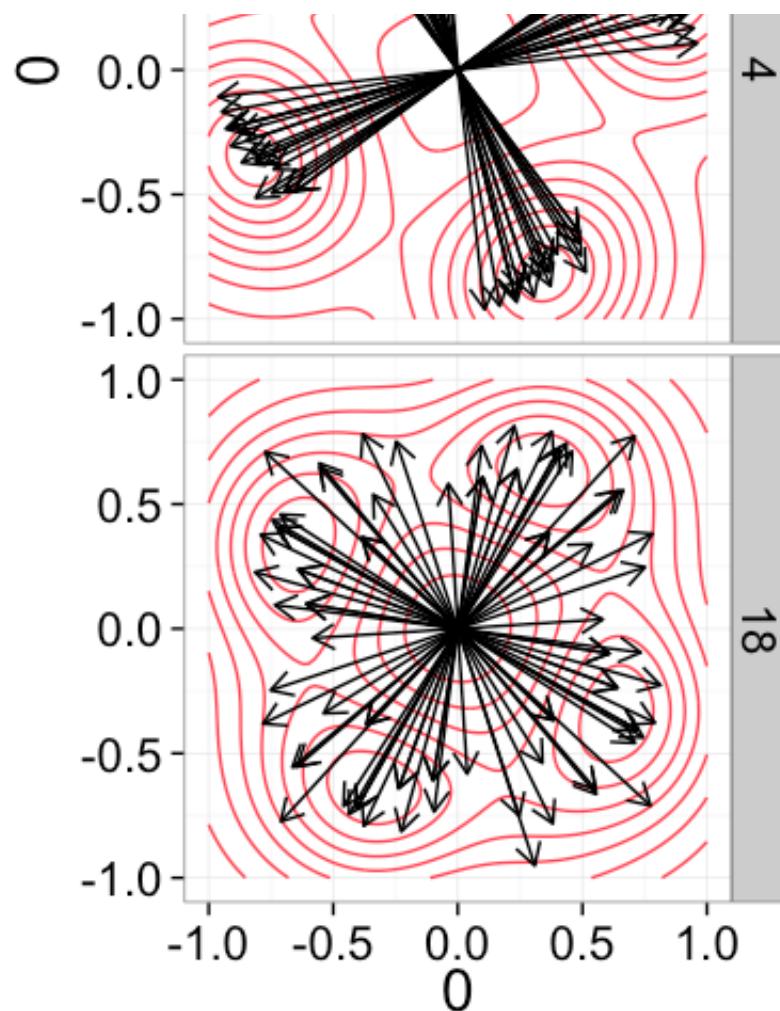




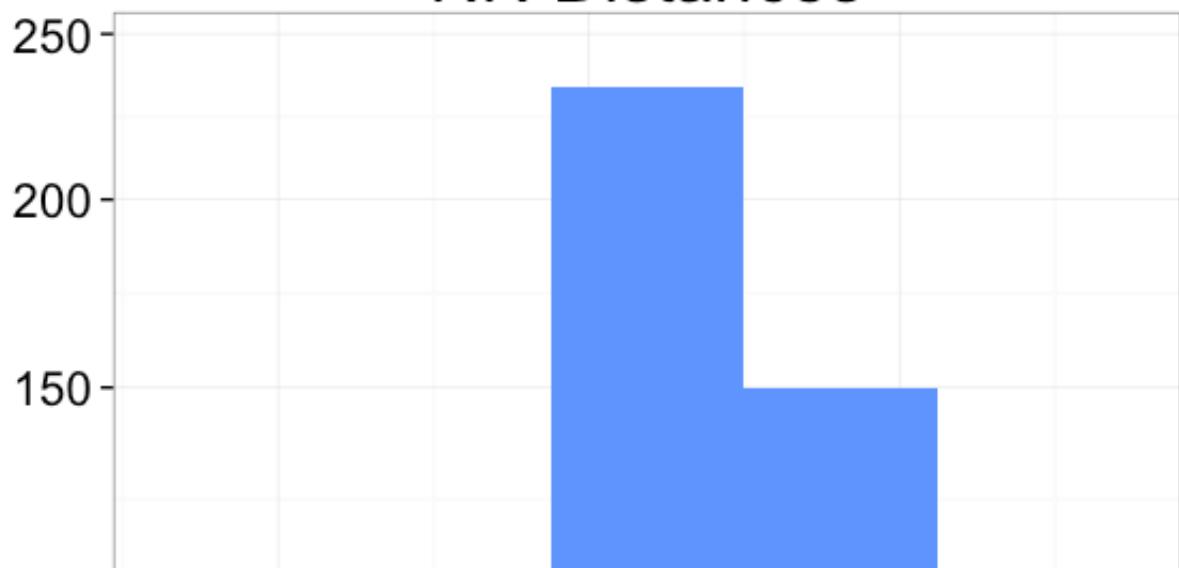
## Swirl NN Distributions

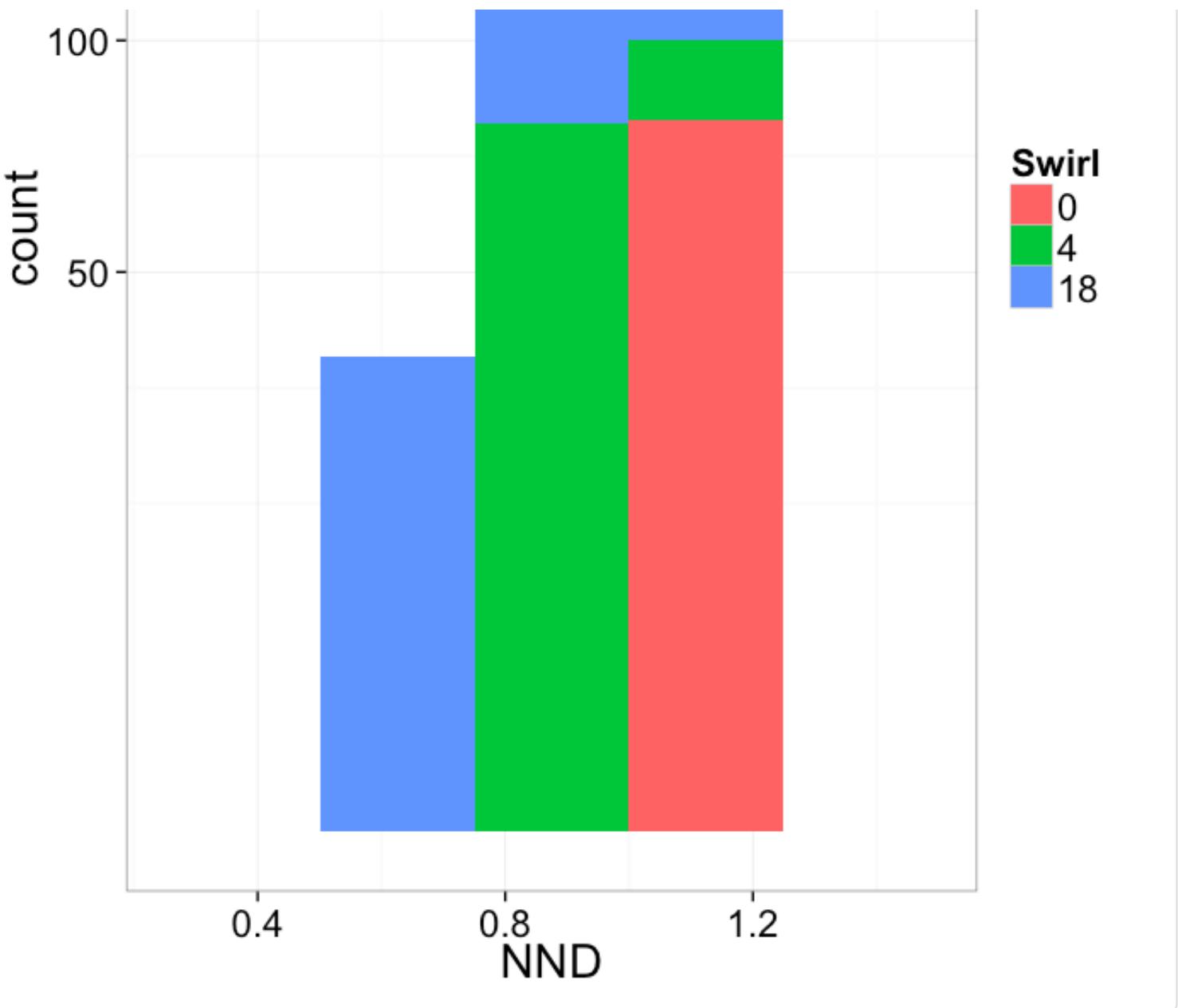
Nearest Neighbor Orientations





## NN Distances





## What we notice

We notice there are several fairly significant short-comings of these metrics (particularly with in-silico systems)

1. Orientation appears to be useful but random
  - Why should it matter if one side is 0.01% closer?
2. Single outlier objects skew results
3. We only extract one piece of information
4. Difficult to create metrics
  - Fit a peak to the angle distribution and measure the width as the "angle variability"?

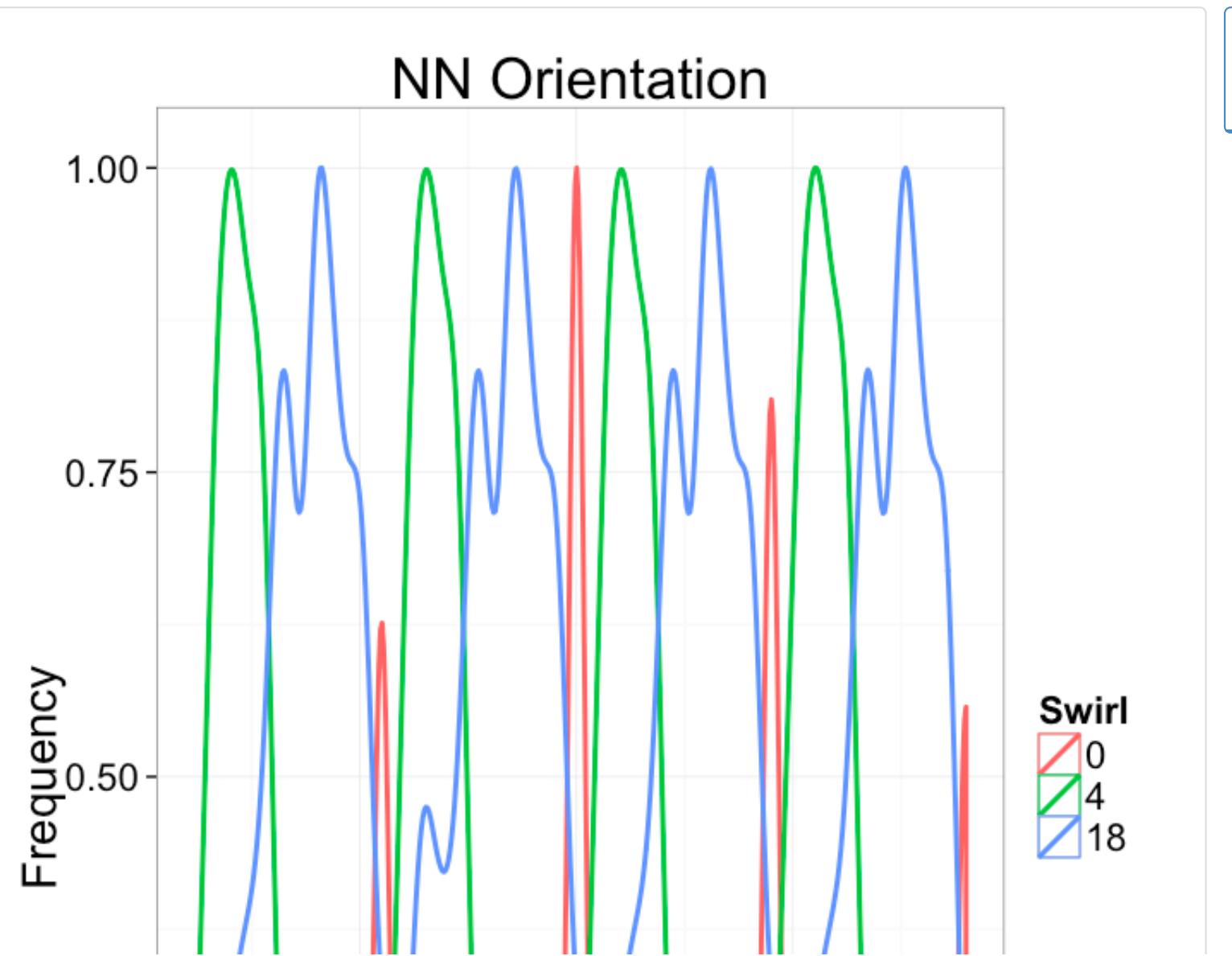
Luckily we are not the first people to address this issue

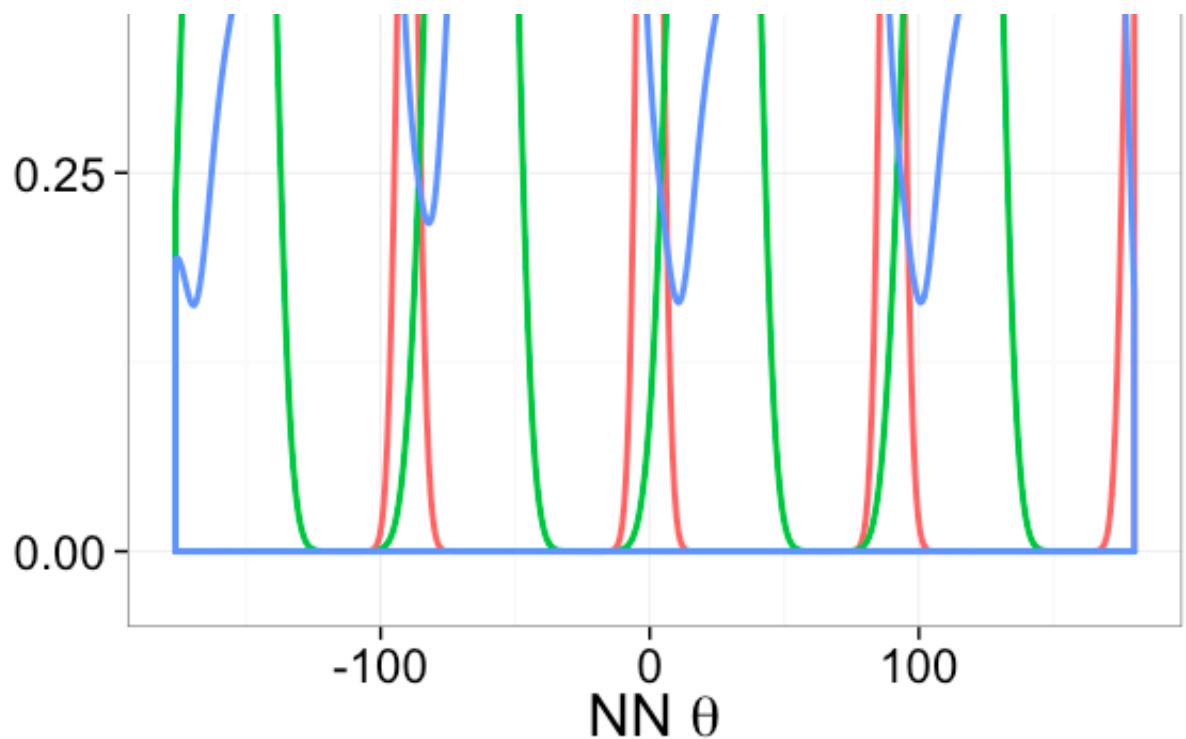
# Random Systems

Using a uniform grid of points as a starting point has a strong influence on the results. A better approach is to use a randomly distributed series of points

- resembles real data much better
- avoids these symmetry problems
  - $\epsilon$  sized edges or overlaps
  - identical distances to nearby objects

+/- R Code

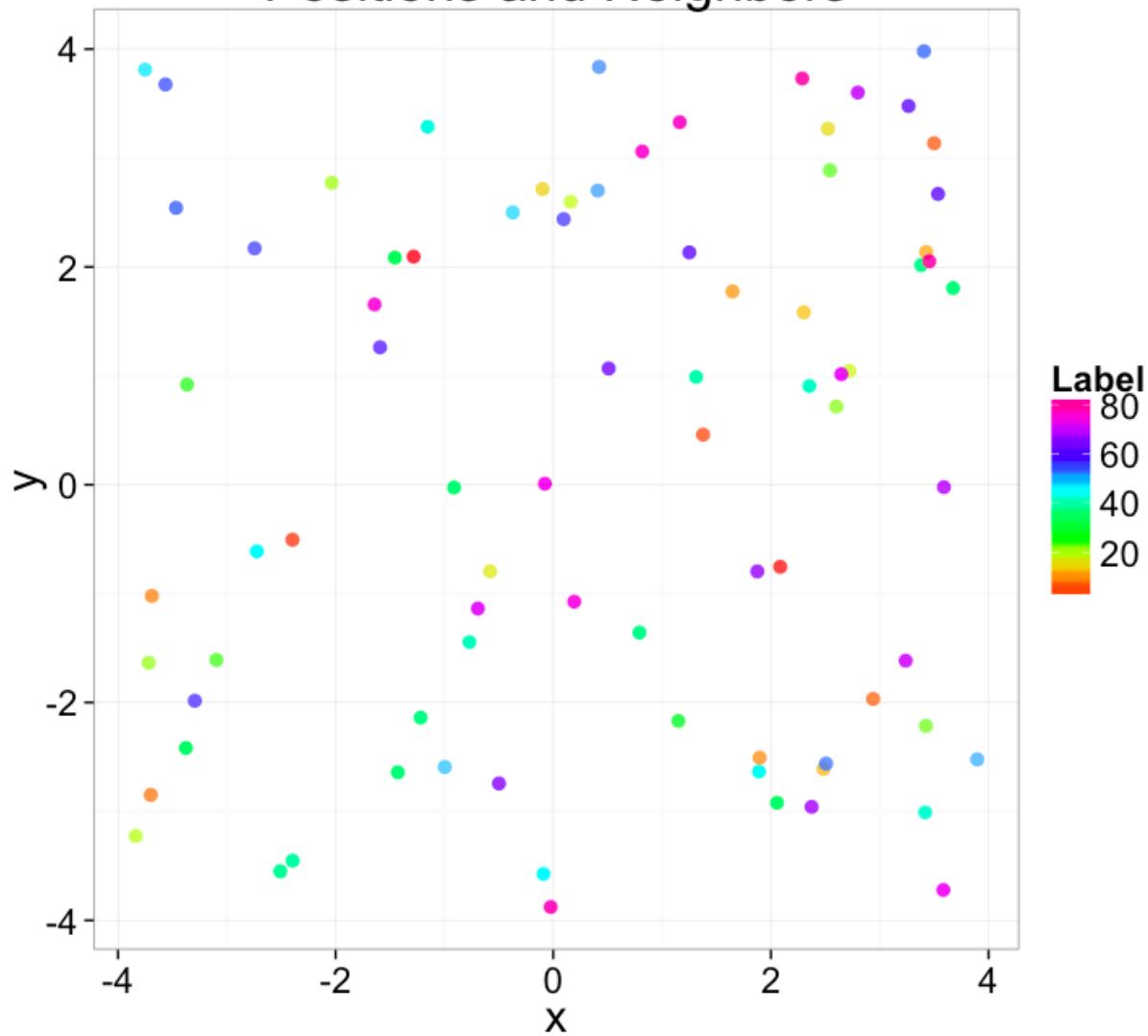




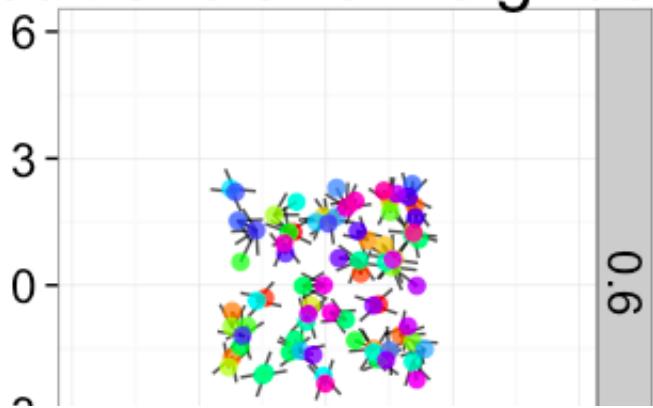
## Examining Compression

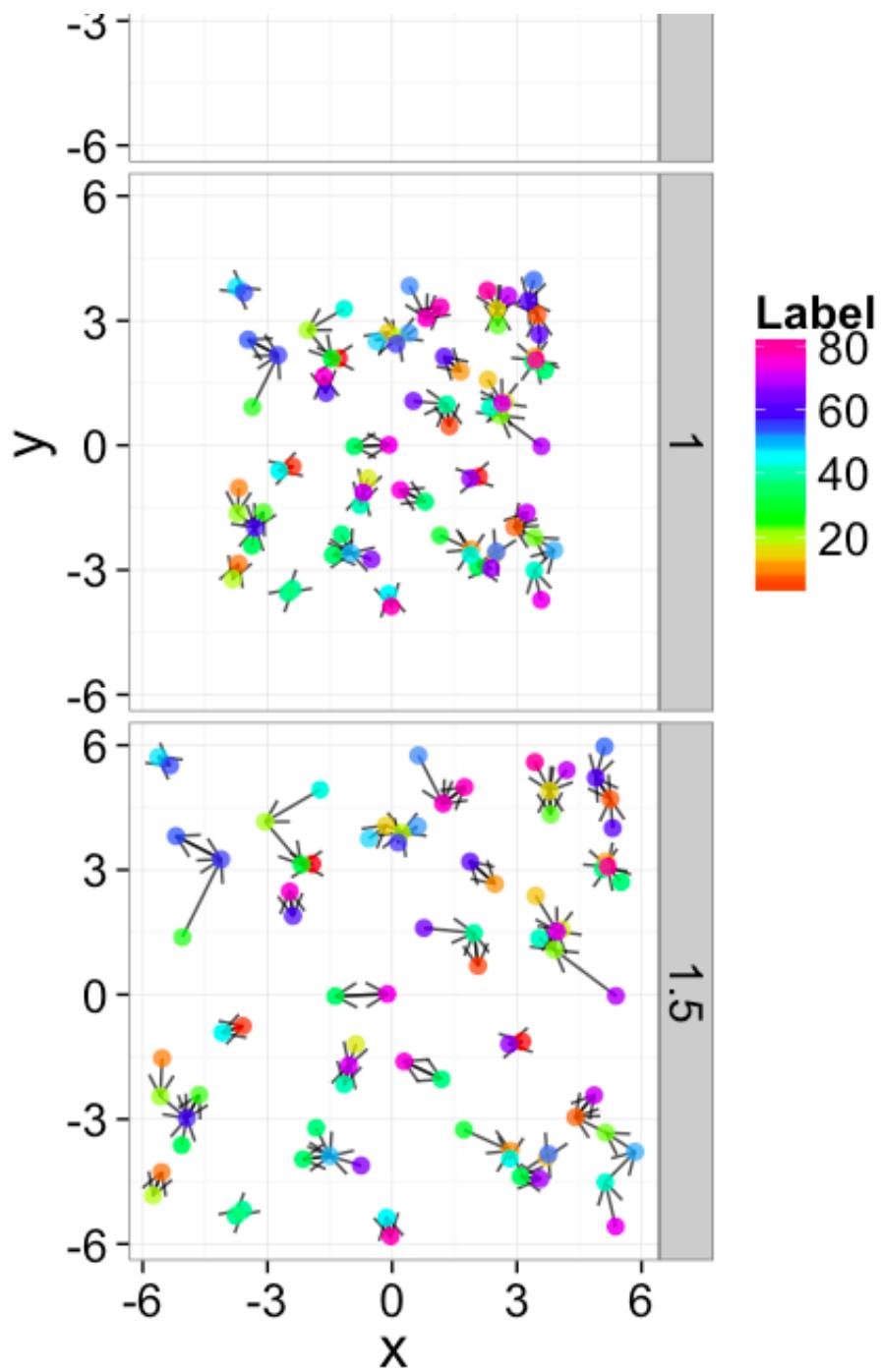
+/- R Code

## Positions and Neighbors



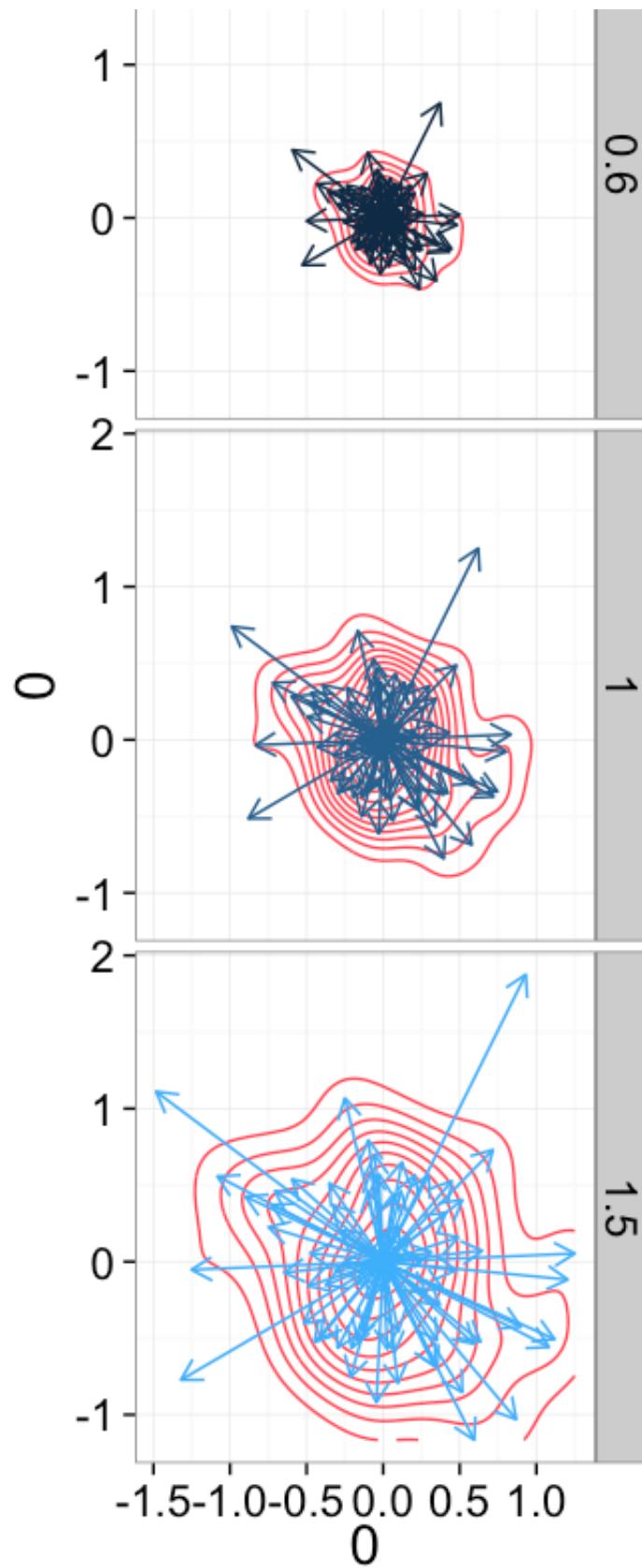
## Positions and Neighbors



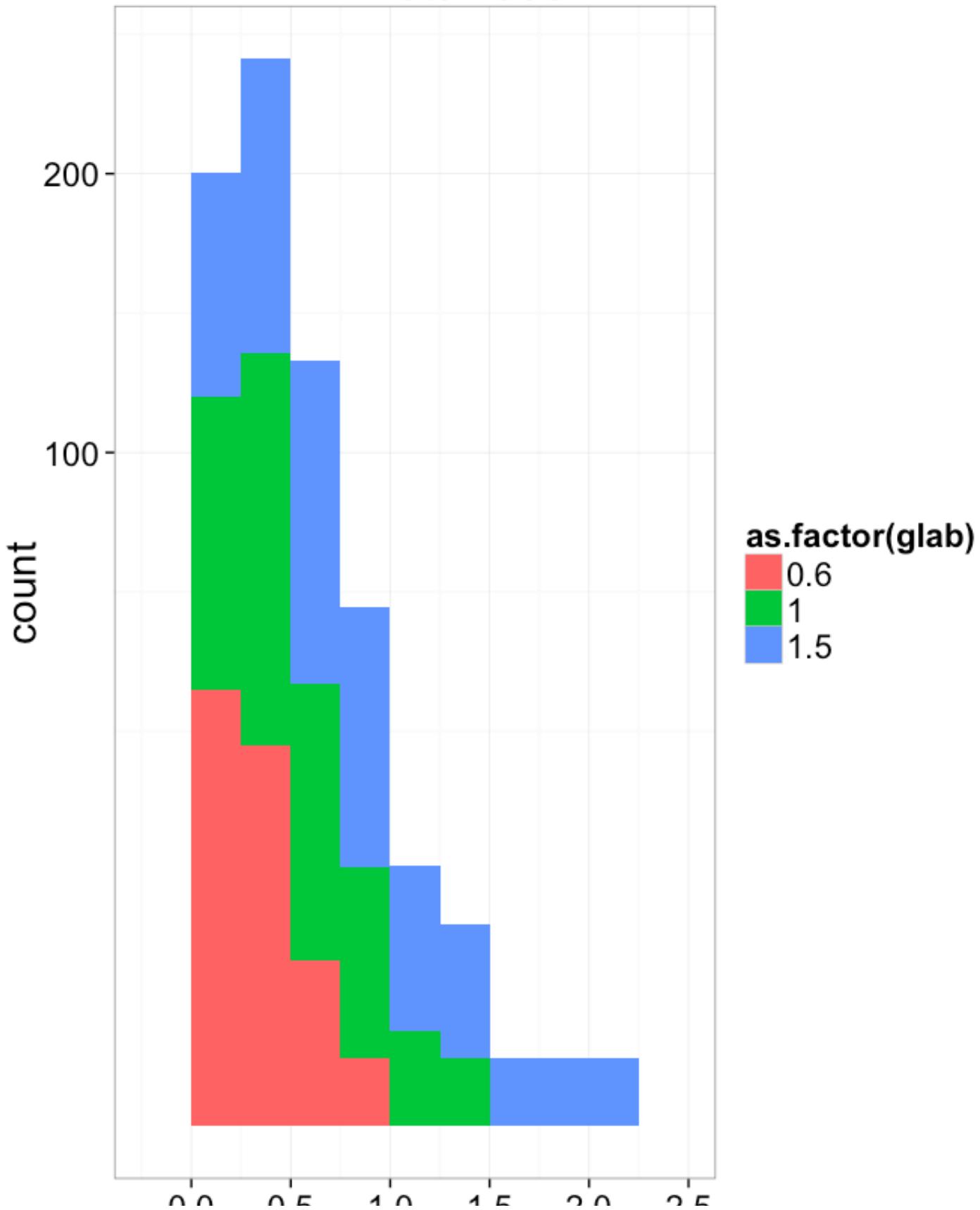


## Compression Distributions

Nearest Neighbor Orientations

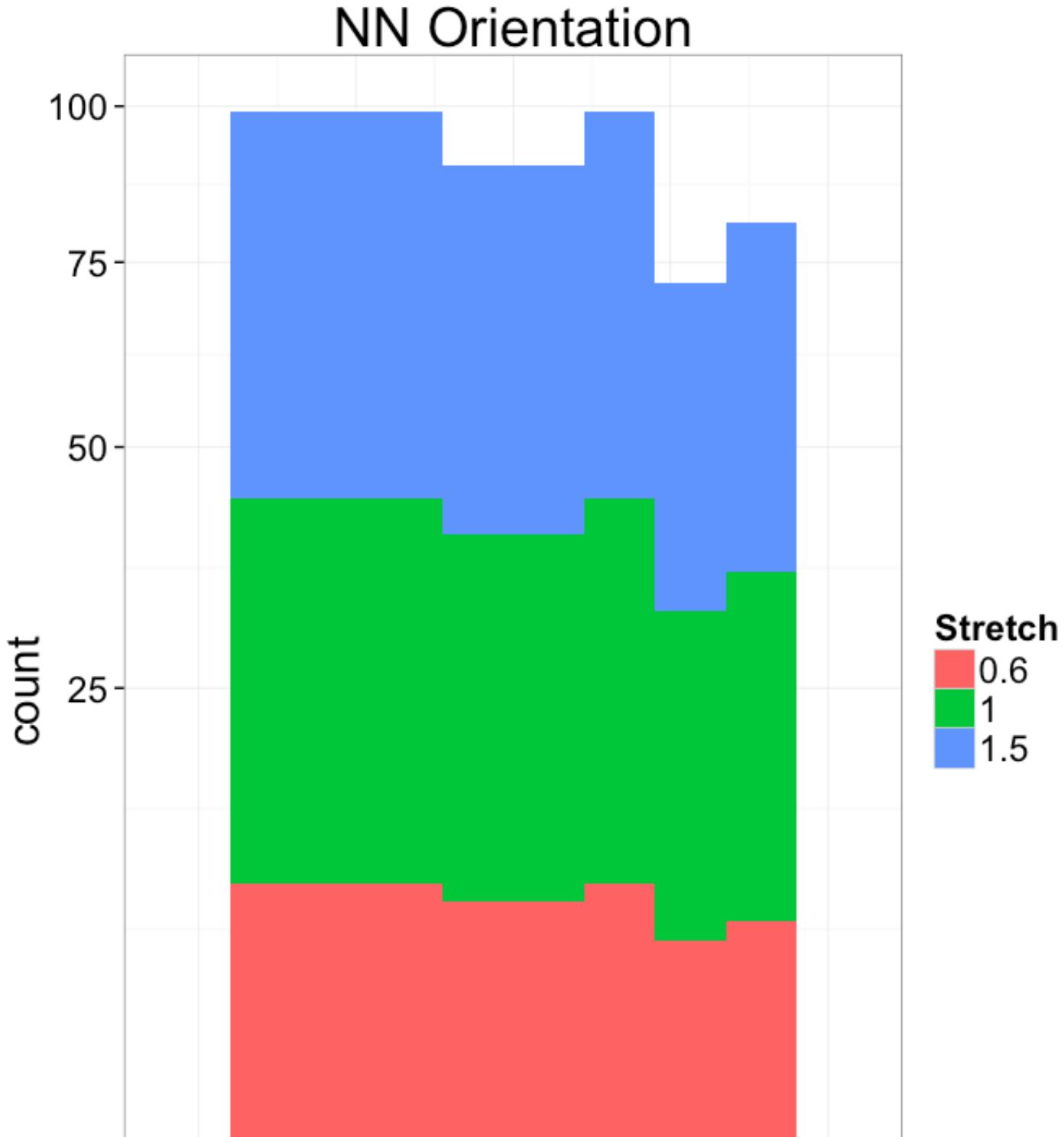


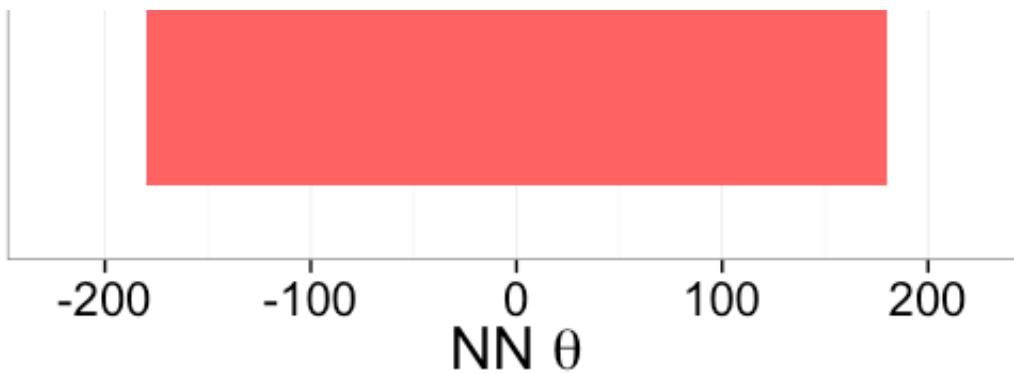
# NN Distances



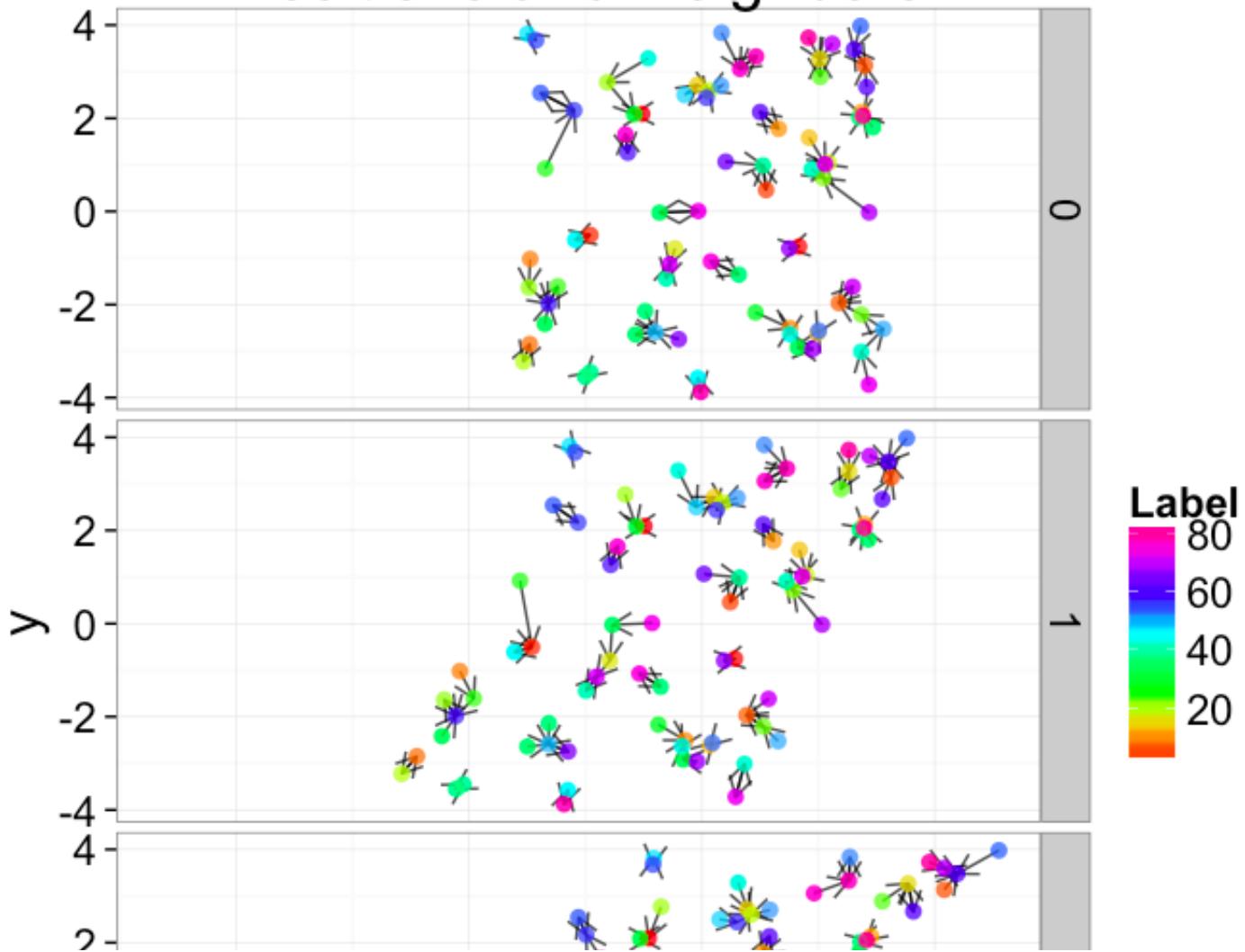


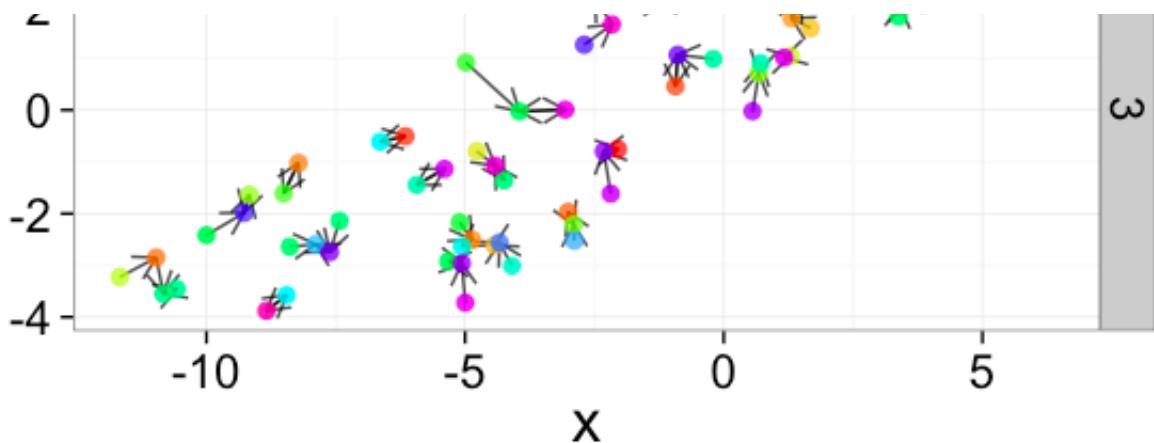
## Examining Different Shears





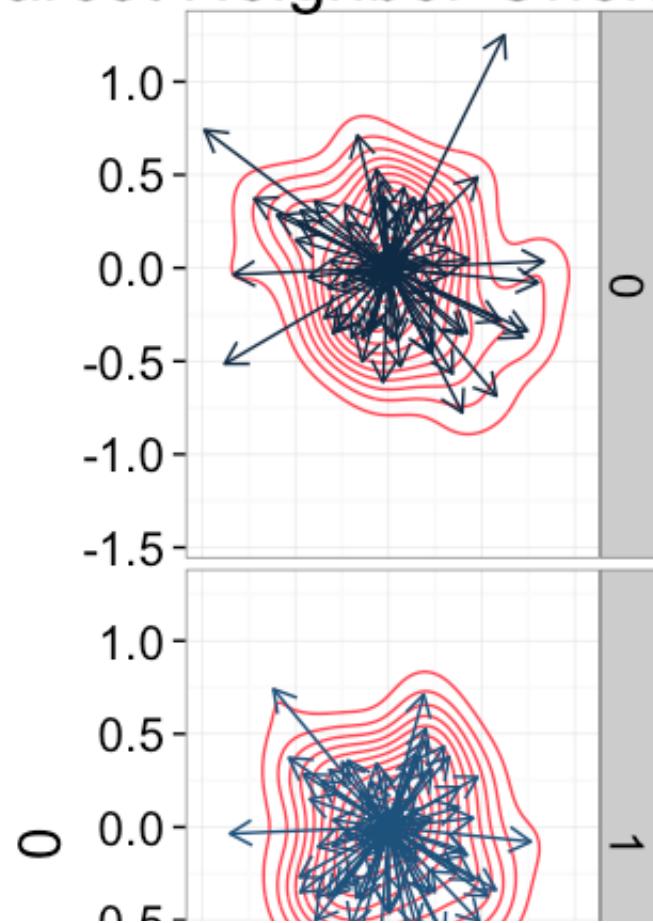
## Positions and Neighbors

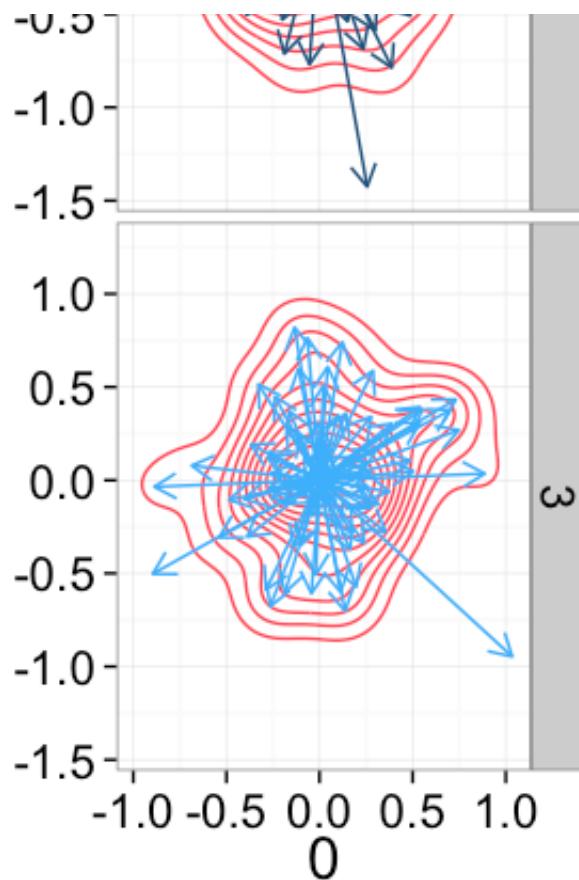




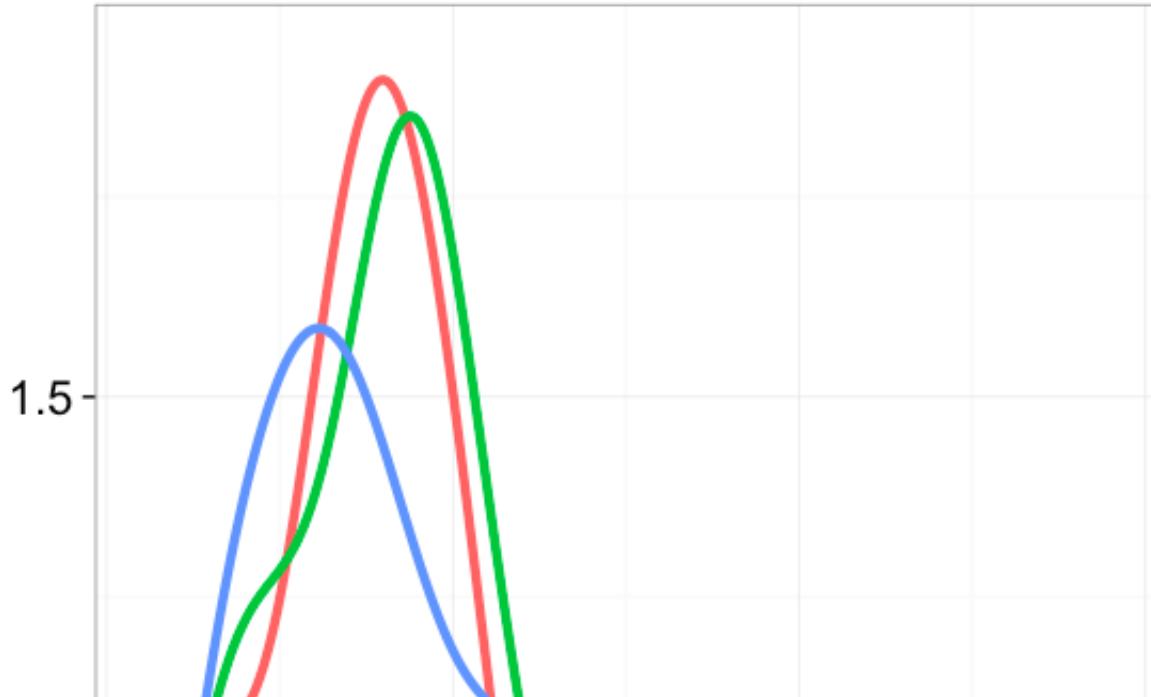
## Shear Distributions

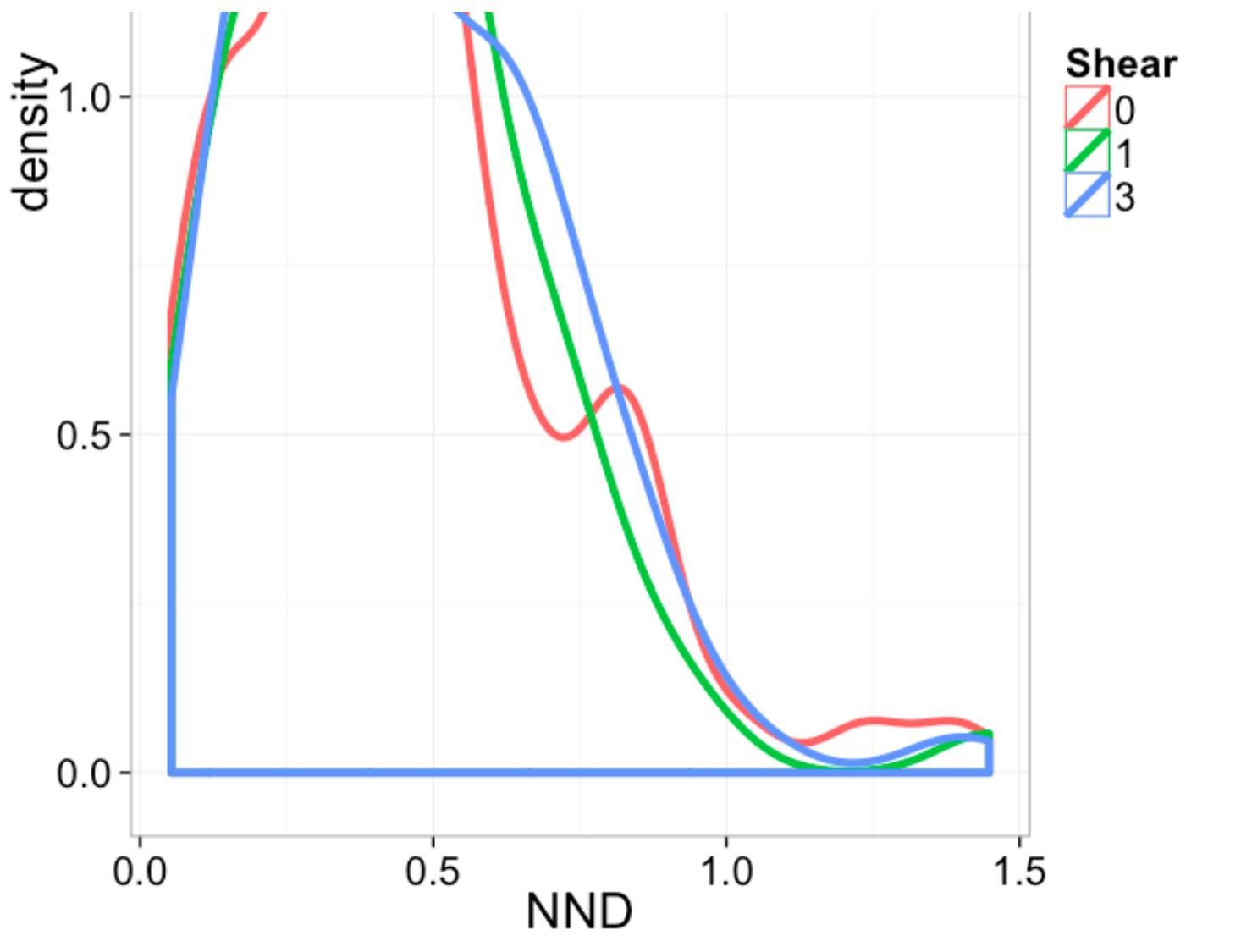
### Nearest Neighbor Orientations





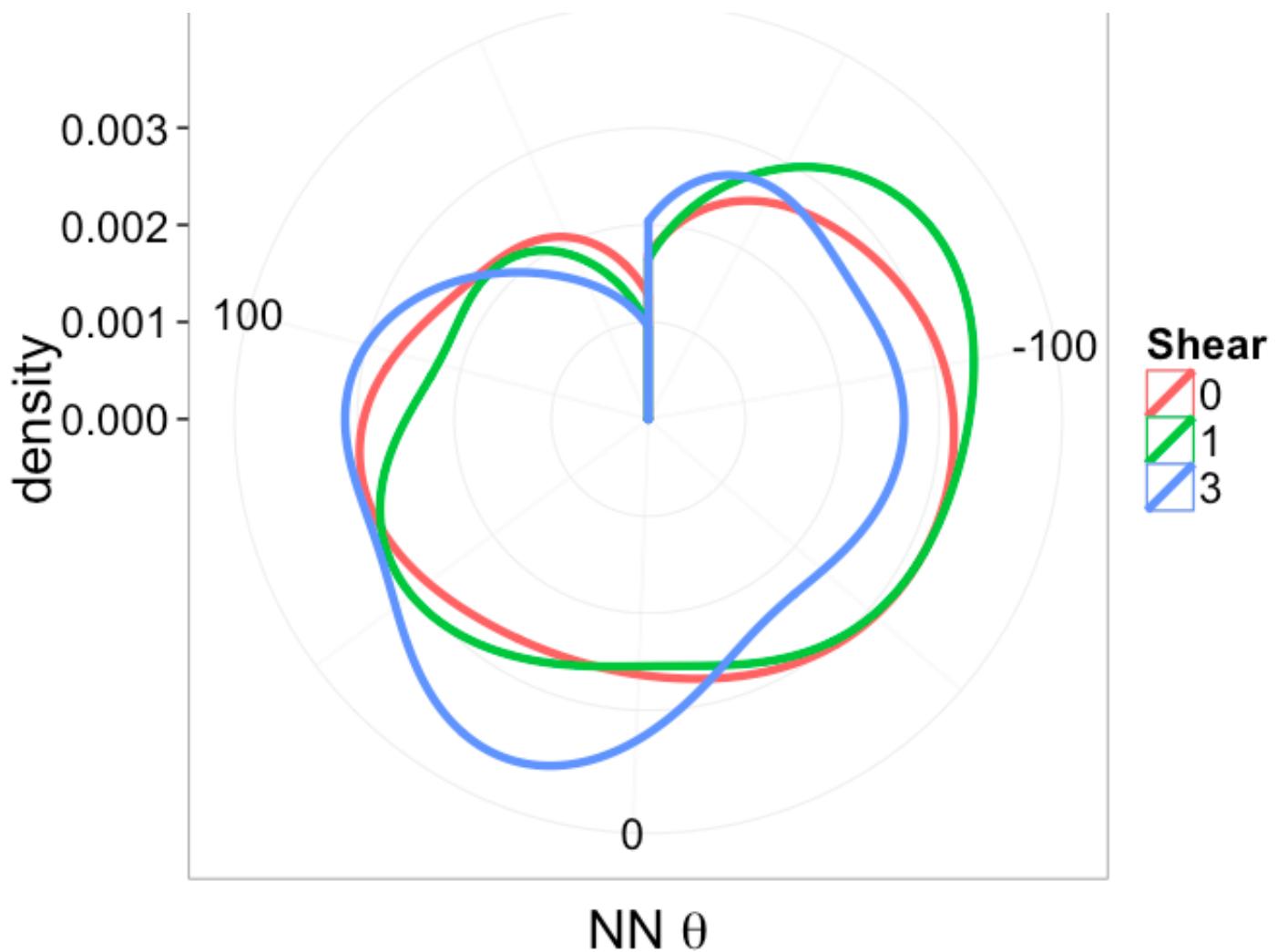
## NN Distances



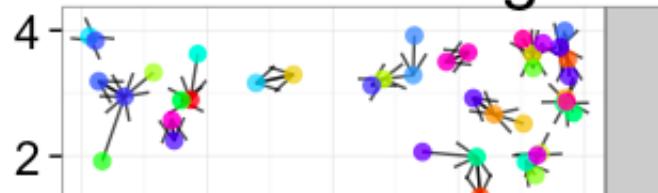


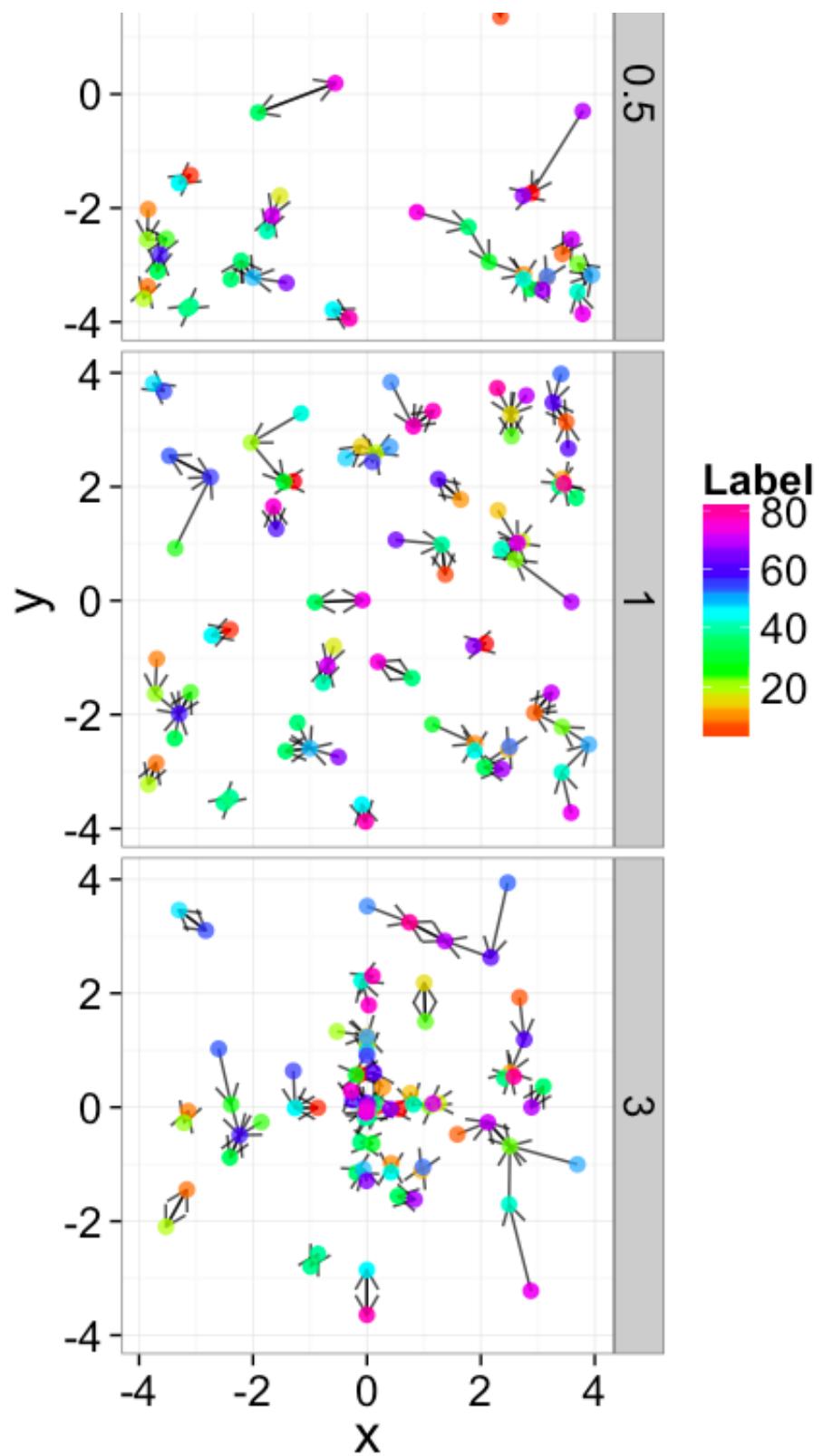
## Examining Different Stretches

NN Orientation



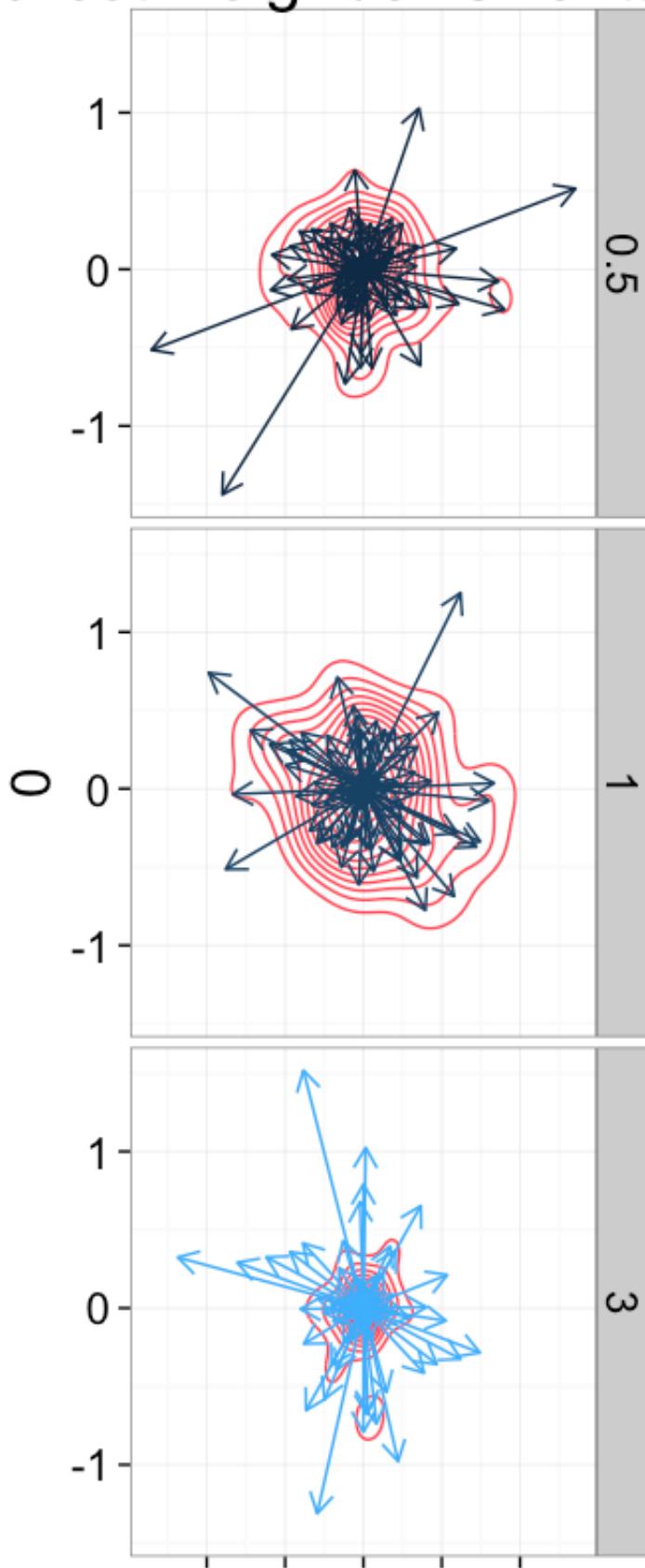
## Positions and Neighbors





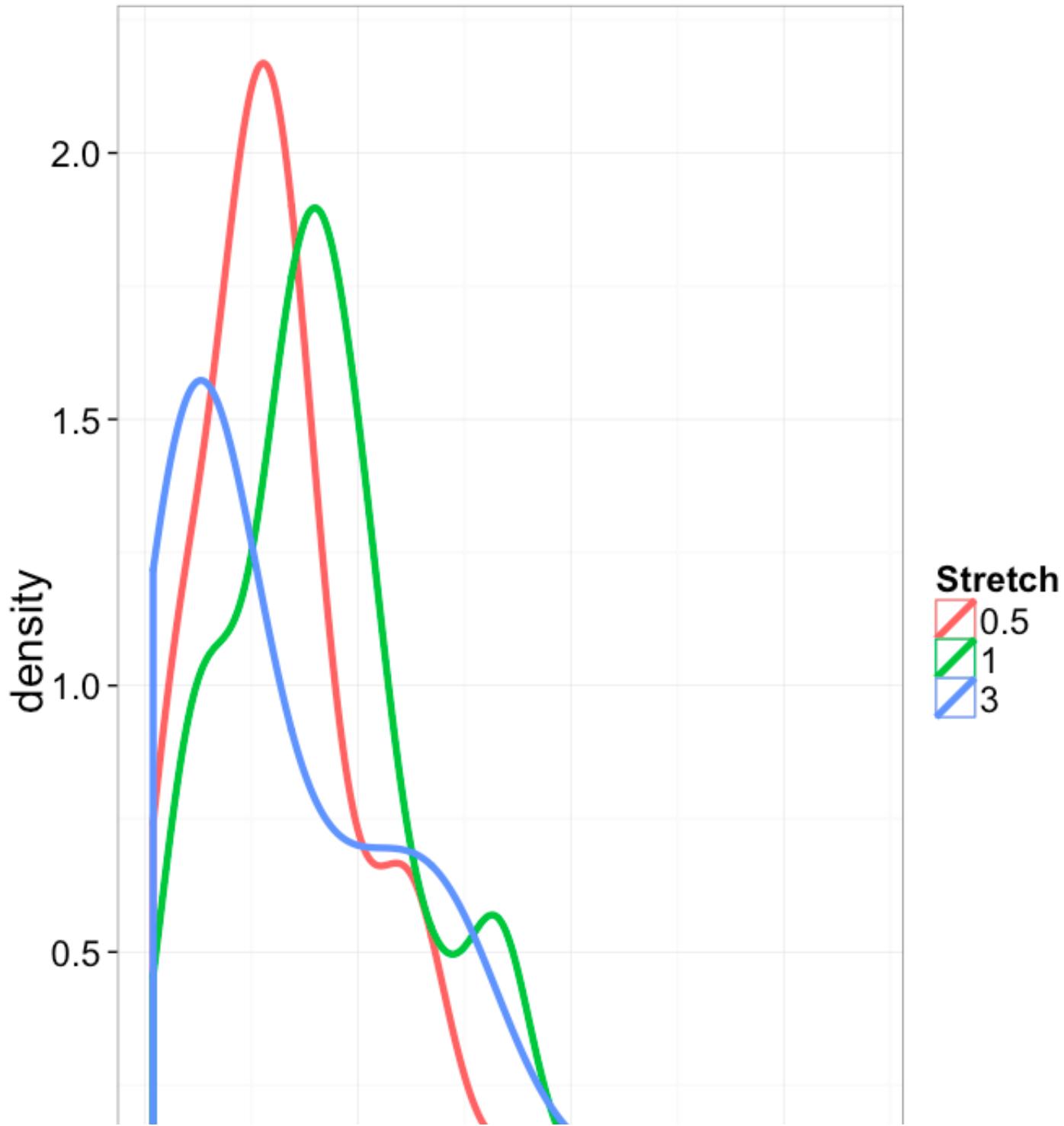
# Stretch Distributions

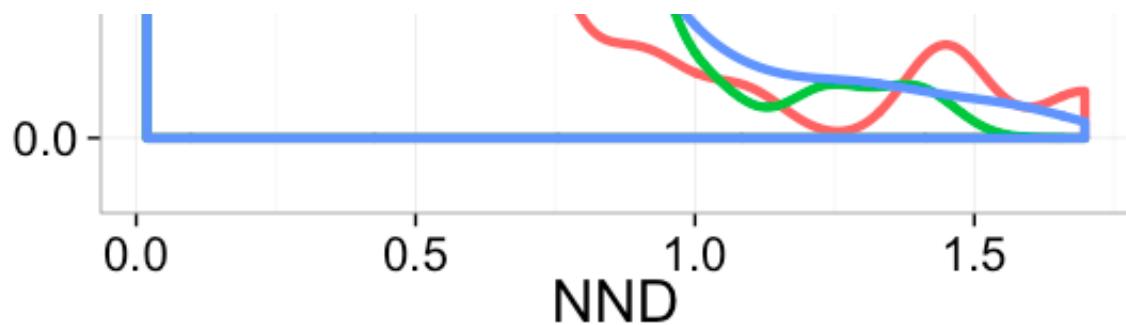
# Nearest Neighbor Orientations



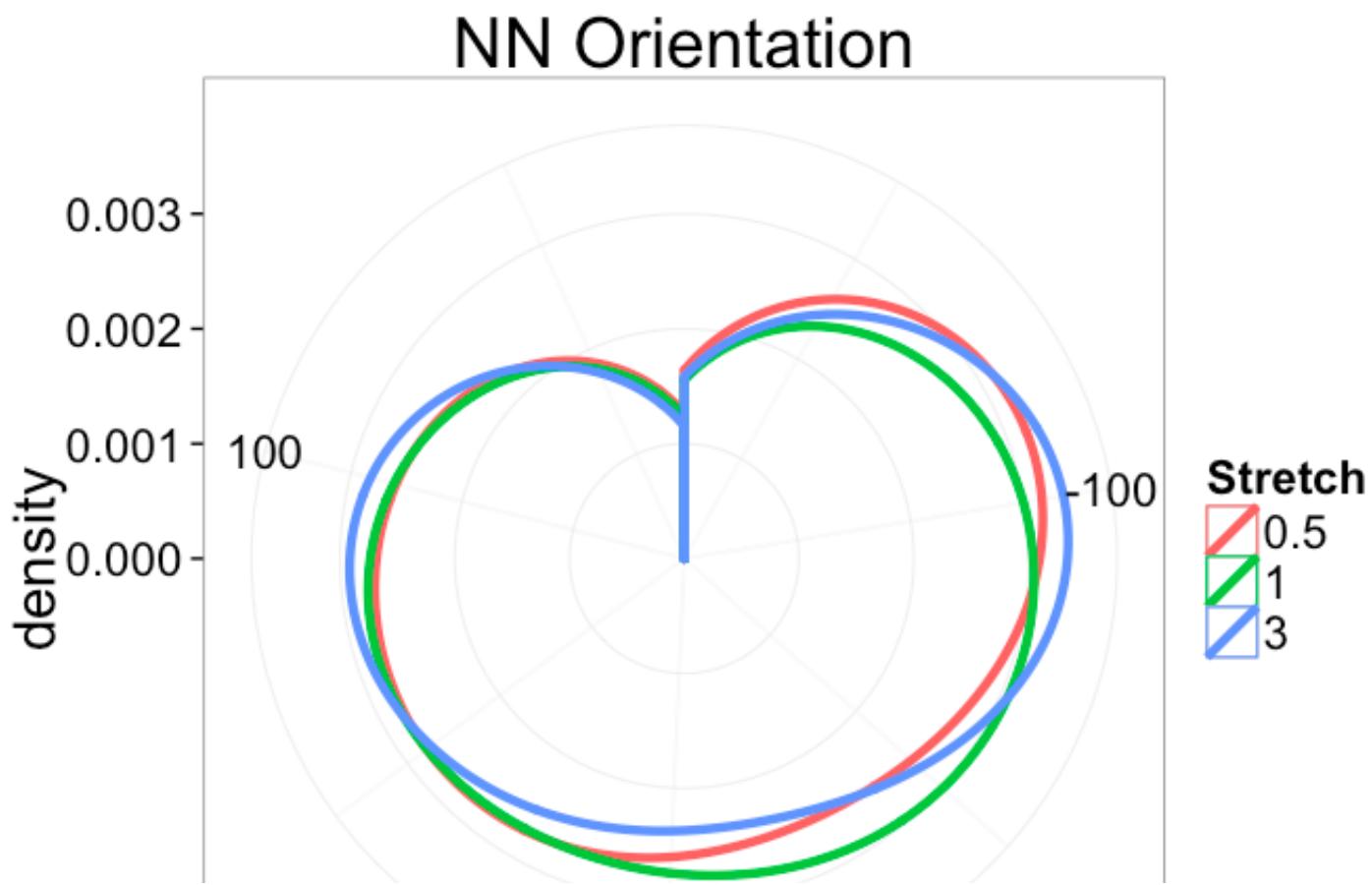
-1.0 -0.5 0.0 0.5 1.0  
0

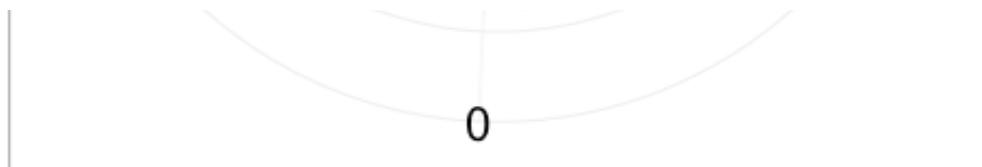
## NN Distances



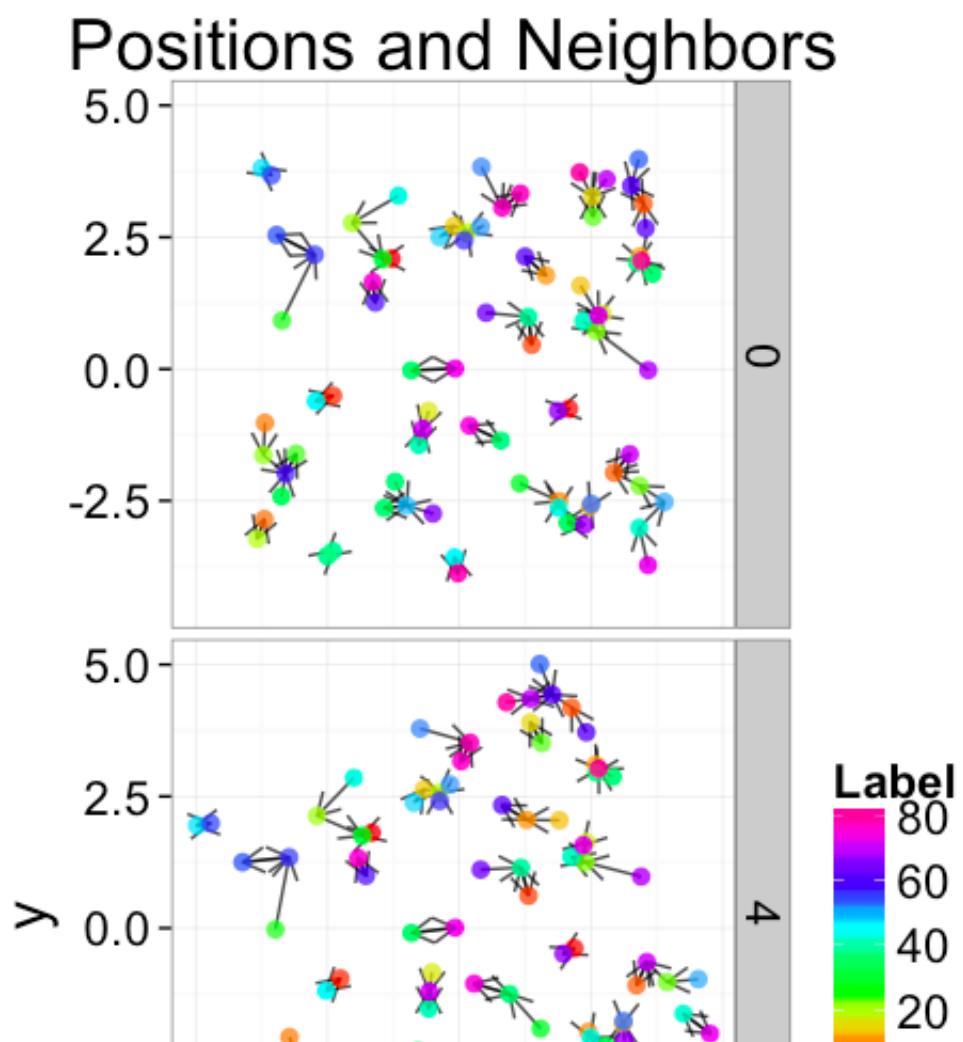


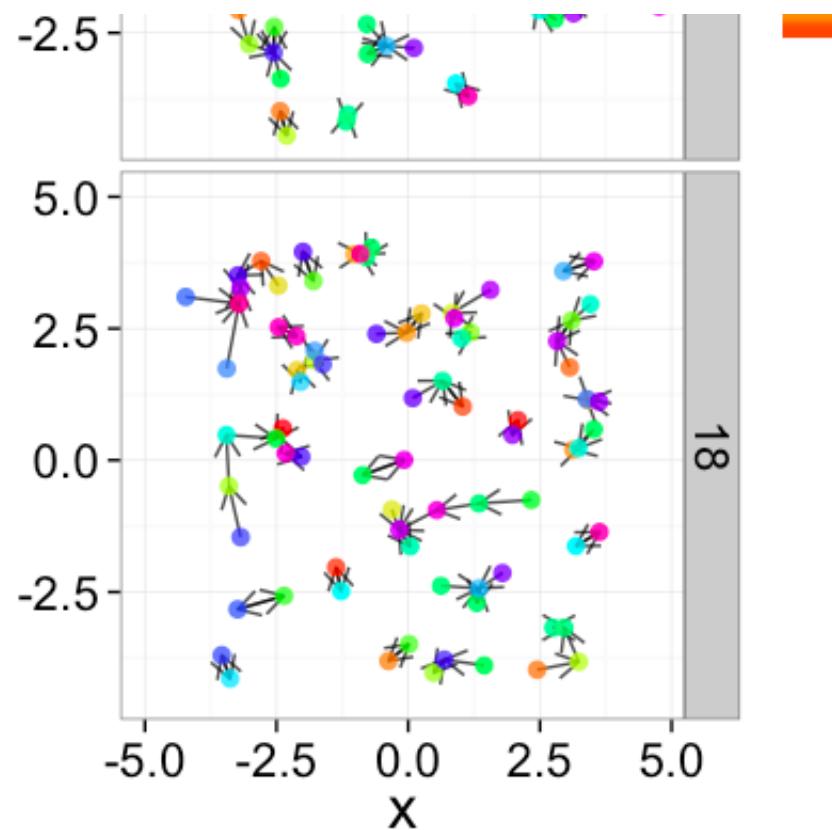
## Examining Swirl Systems





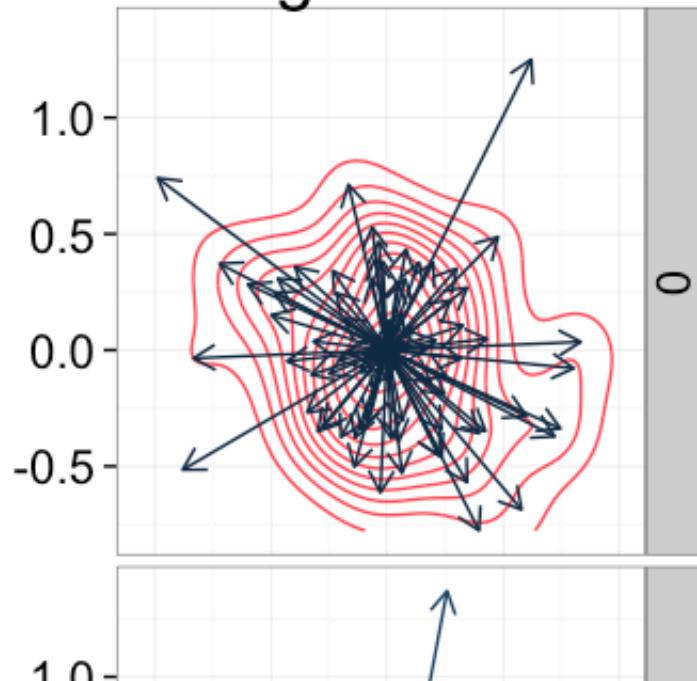
$NN \theta$

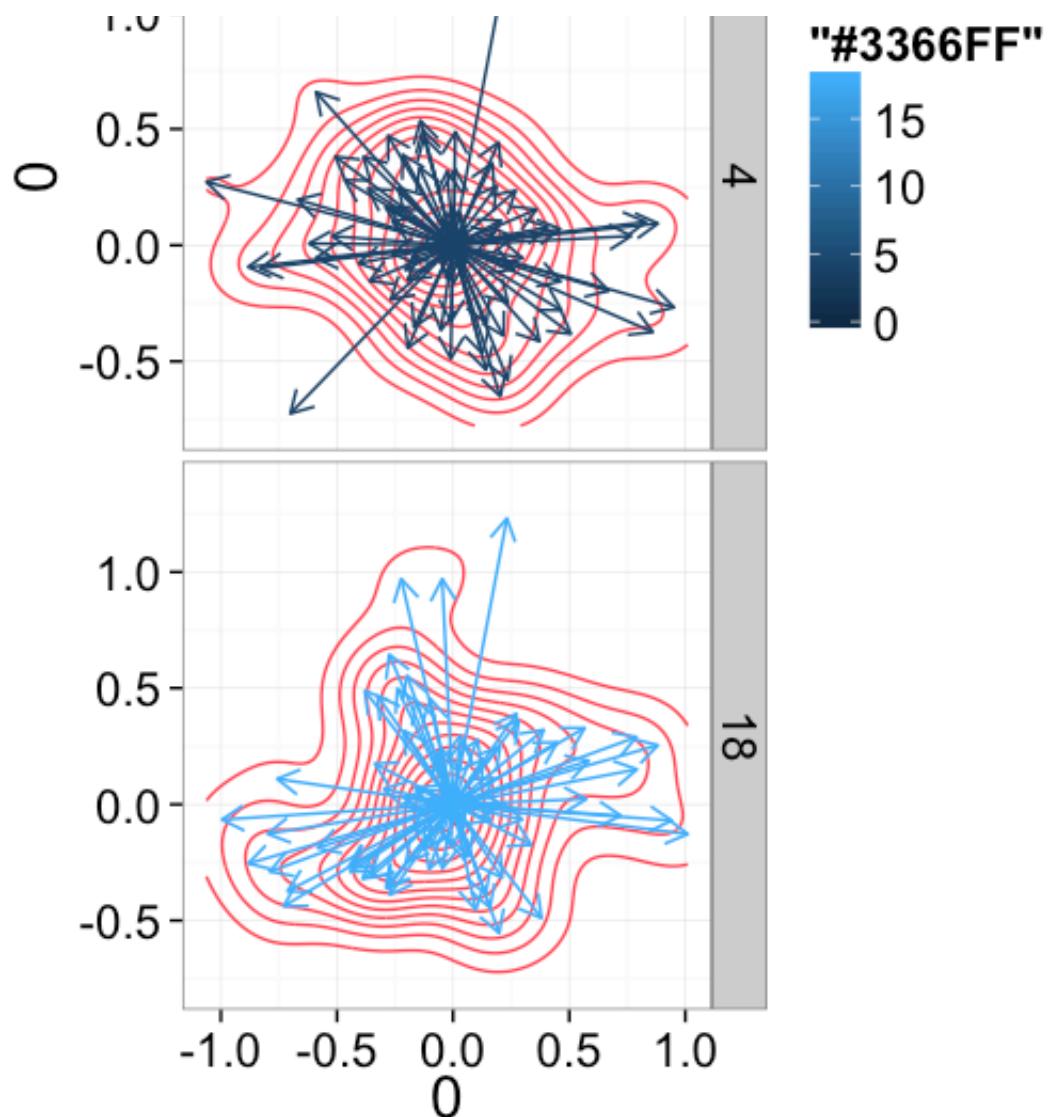




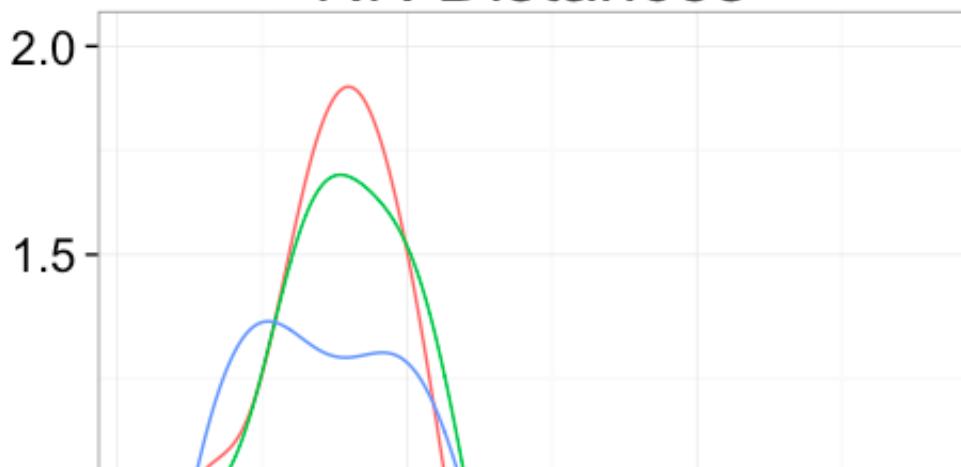
## Swirl NN Distributions

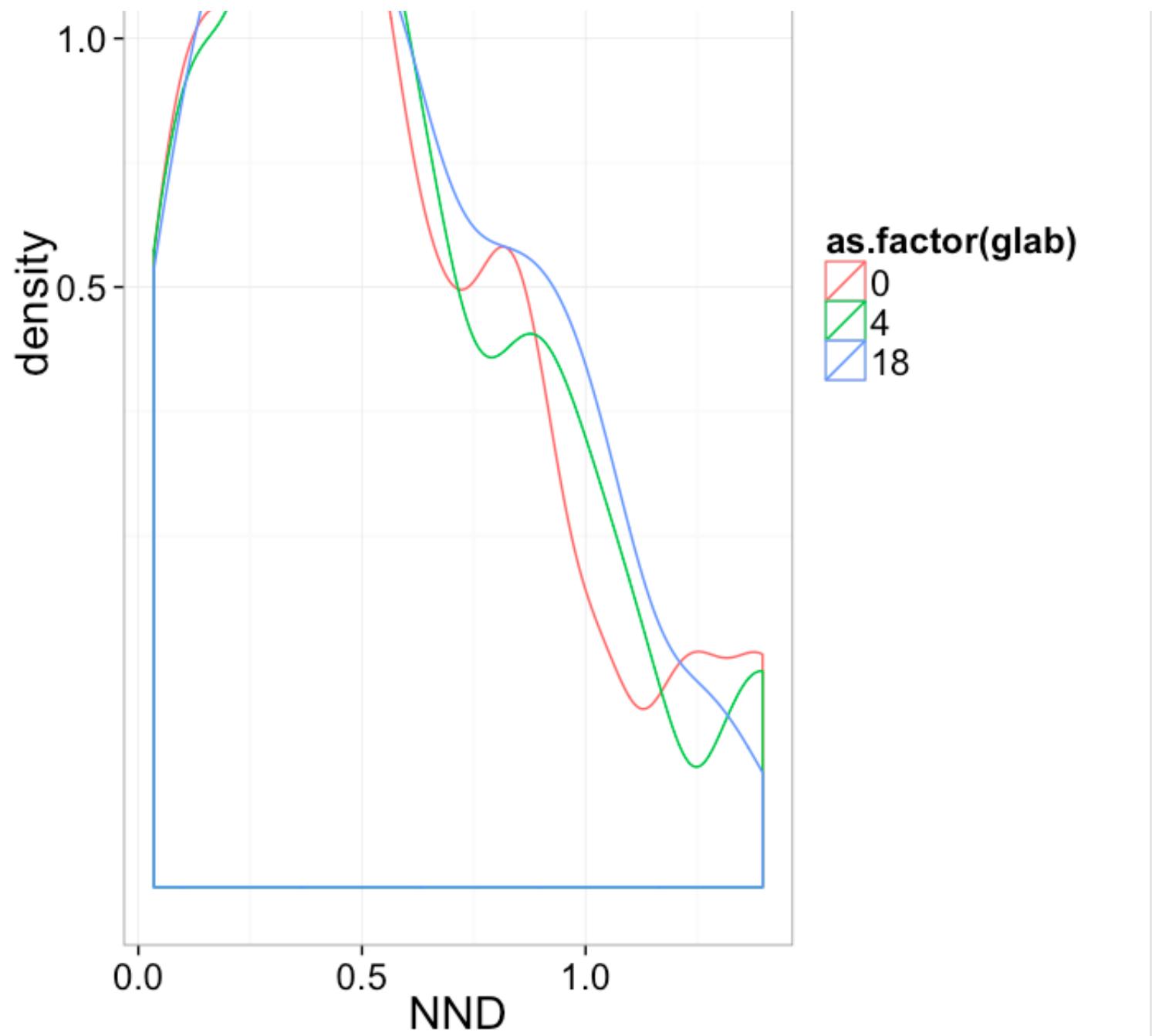
### Nearest Neighbor Orientations





## NN Distances



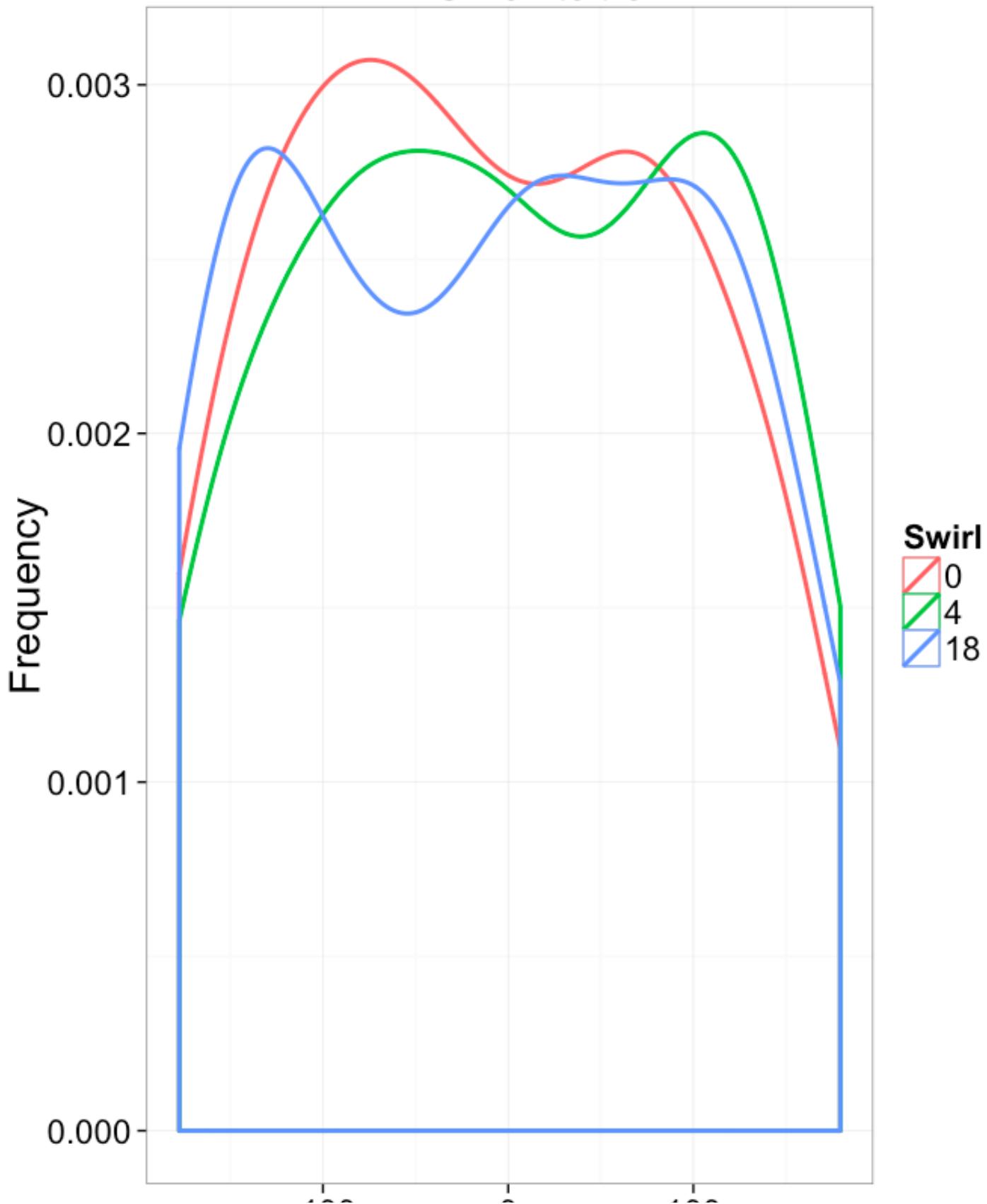


## Voronoi Tesselation

+/- R Code

Voronoi tessellation is a method for partitioning a space based on points. The basic idea is that each point  $\vec{p}$  is assigned a region **R** consisting of points which are closer to  $\vec{p}$  than any of the other points. Below the diagram is shown in a dashed line for the points shown as small circles.

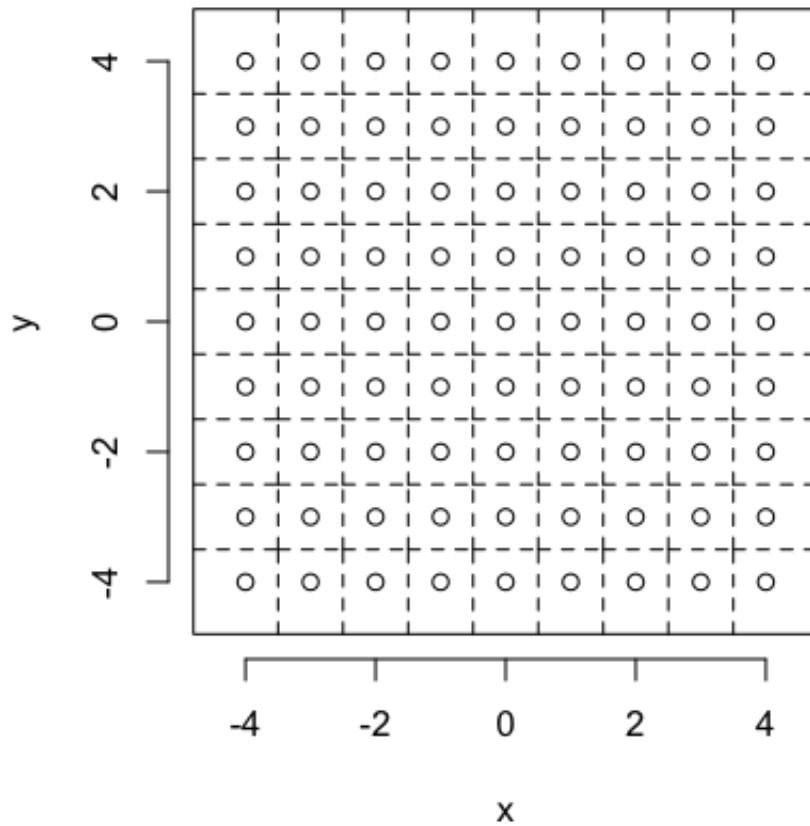
## NN Orientation



-100                    0                    100  
**NN**  $\theta$

We call the area of a region ( $R$ ) around point  $\vec{p}$  its territory.

The grid on the random system, shows much more diversity in territory area.



# Calculating Density

Back to our original density problem of having just one number to broadly describe the system.

- Can a voronoi tessellation help us with this?
- **YES**

With density we calculated

$$\text{Density} = \frac{\text{Number of Objects}}{\text{Total Volume}}$$

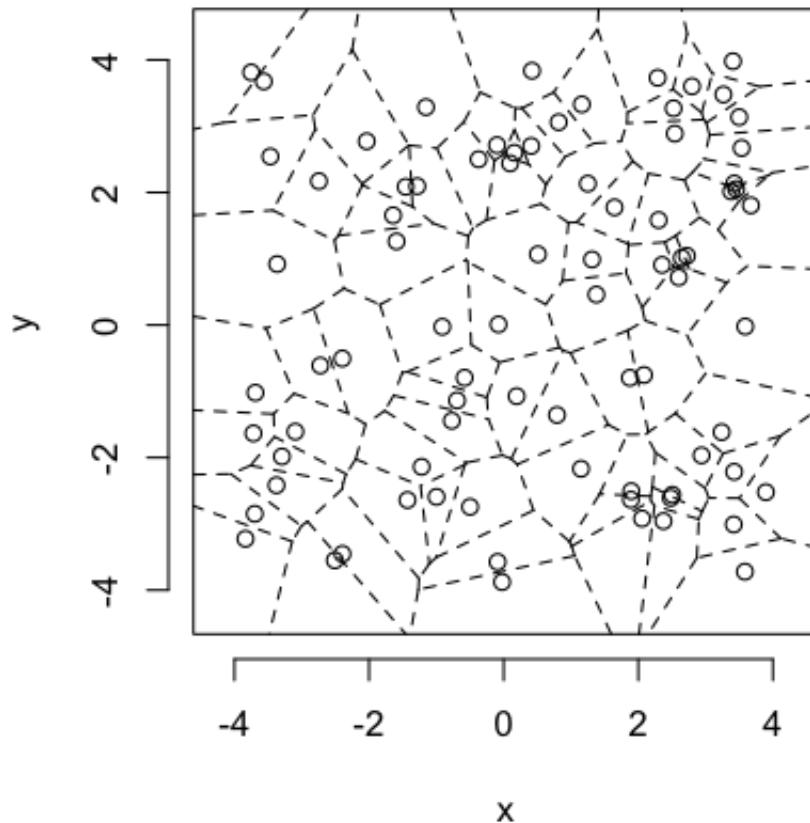
with the regions we have a territory (volume) per object so the average territory is

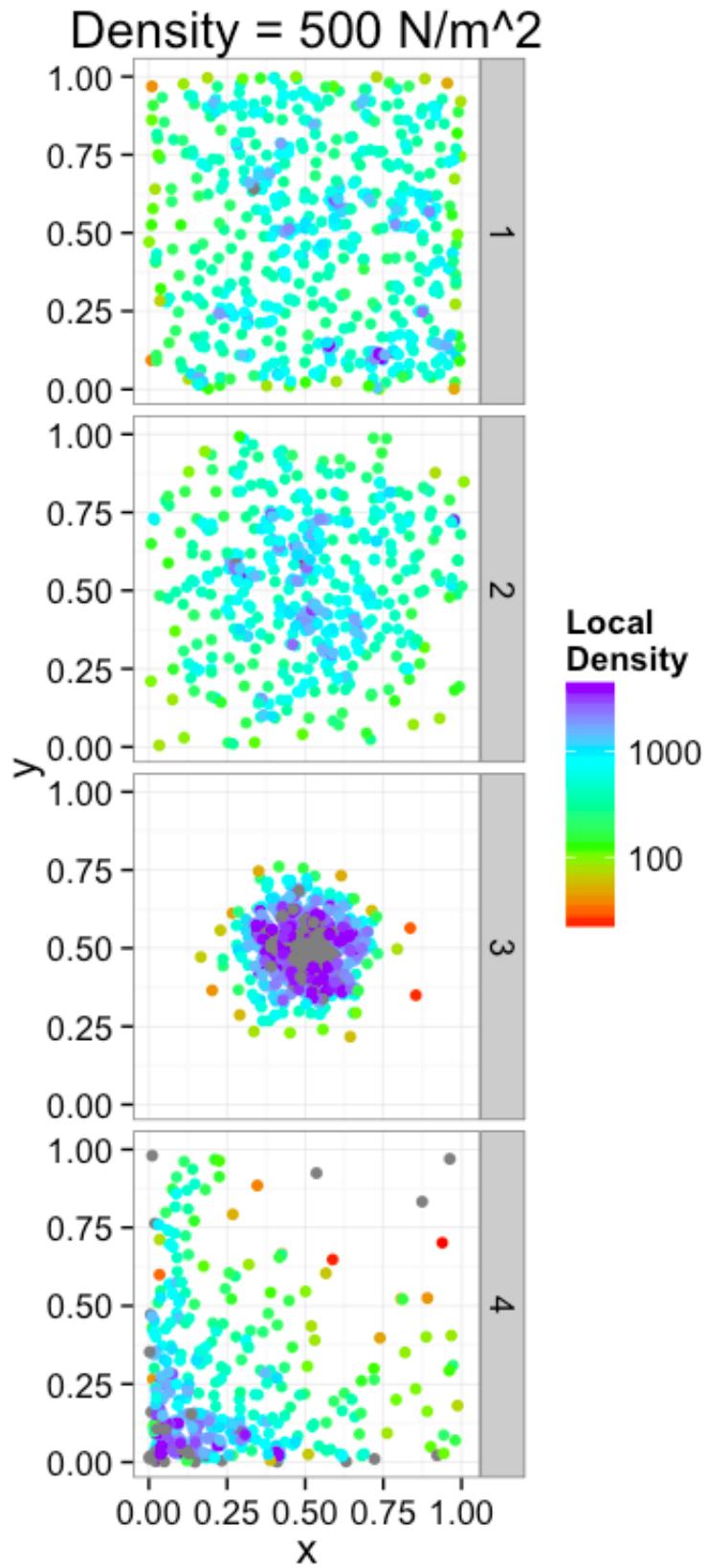
$$\bar{\text{Territory}} = \frac{\sum \text{Territory}_i}{\text{Number of Objects}} = \frac{\text{Total Volume}}{\text{Number of Objects}} = \frac{1}{\text{Density}}$$

So the same, but we now have a density definition for a single point!

$$\text{Density}_i = \frac{1}{\text{Territory}_i}$$

## Density Examples



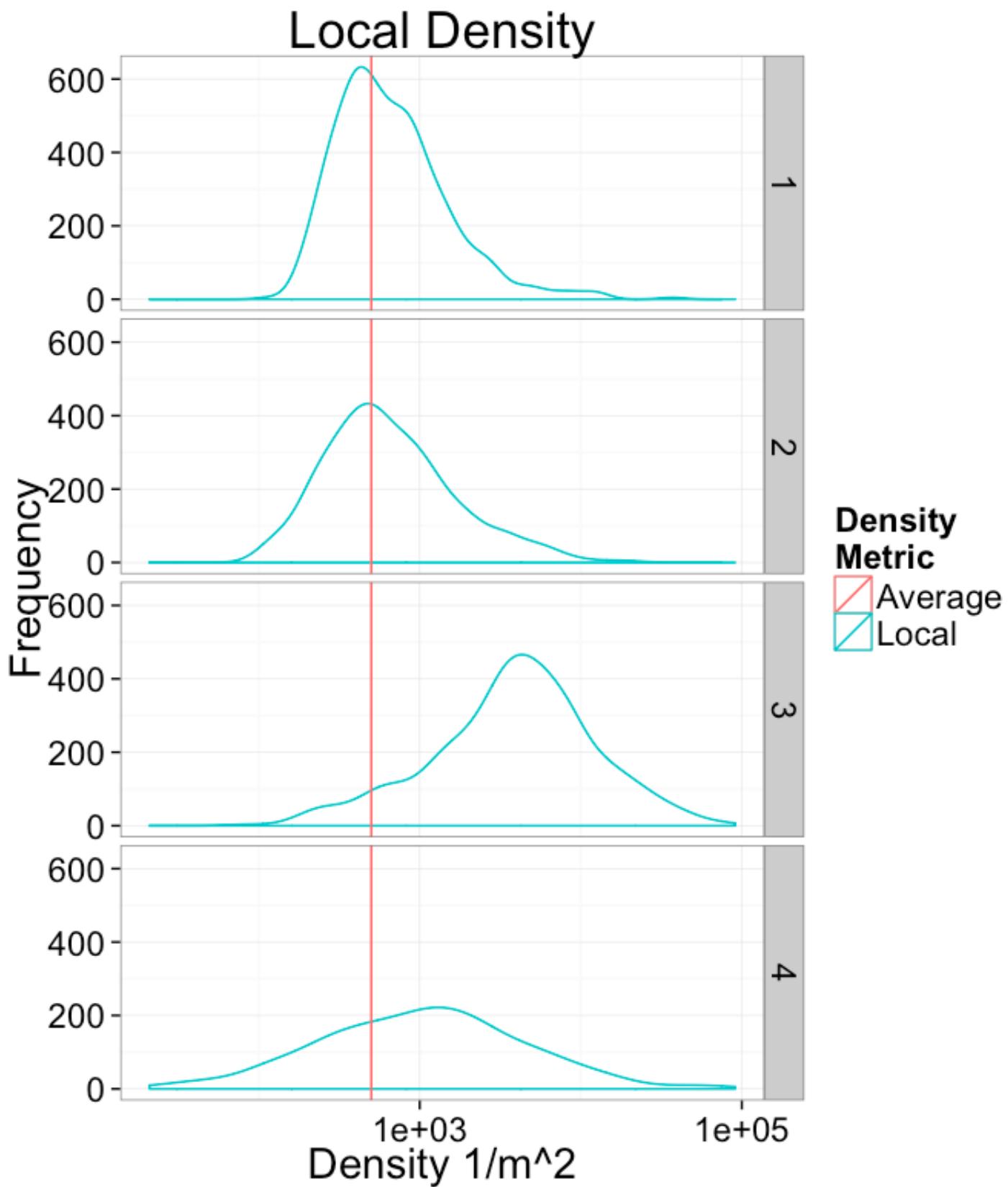


+/- R Code

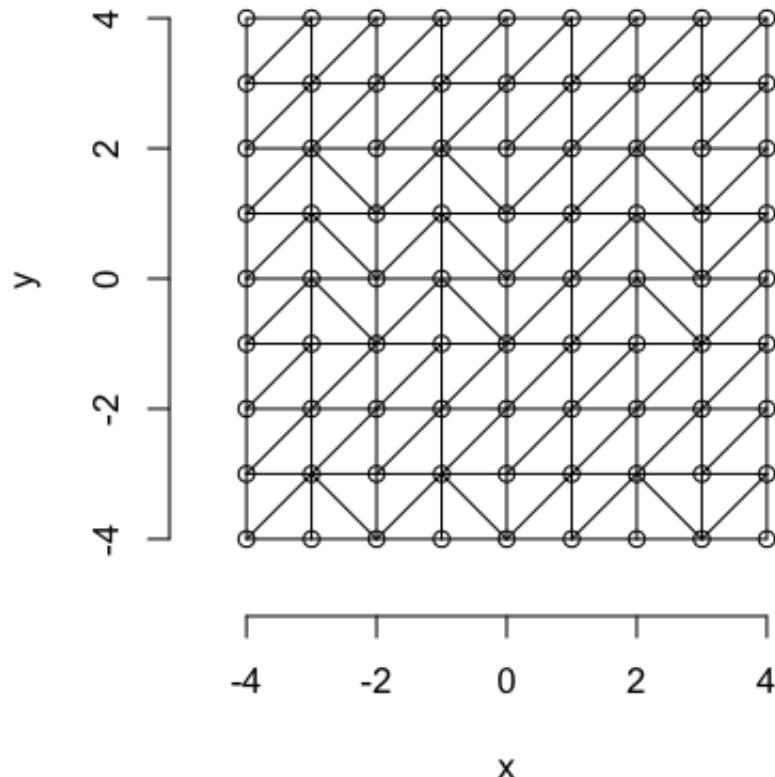
# Delaunay Triangulation

A parallel or *dual* idea (<http://mathworld.wolfram.com/DelaunayTriangulation.html>) where triangles are used and each triangle is created such that the circle which encloses it contains no other points. The triangulation makes the *neighbors* explicit since connected points in the triangulation correspond to points in our tesselation which share an edge (or face in 3D)

+/- R Code



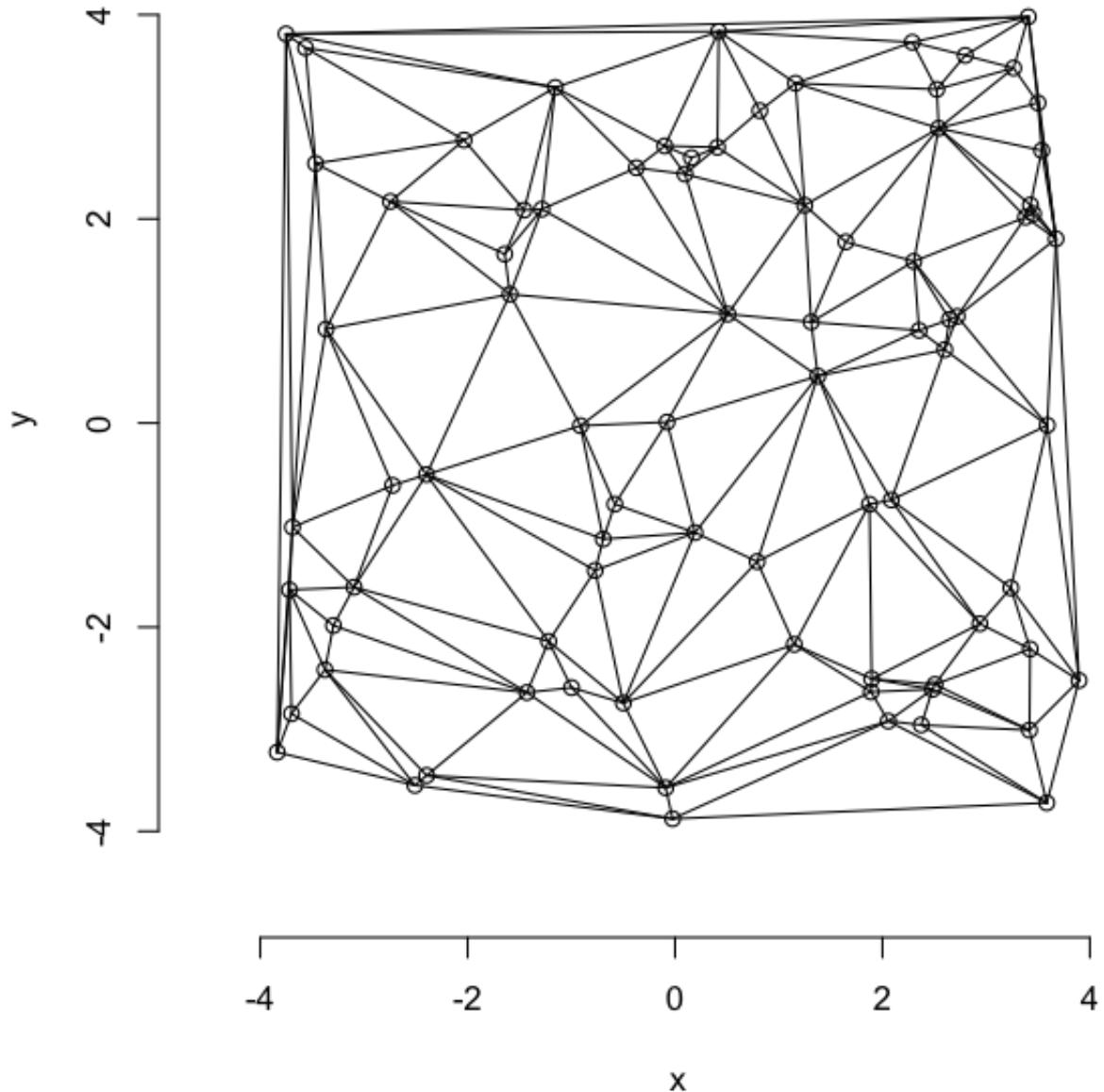
We define the number of connections each point  $\vec{p}$  has the Neighbor Count or Delaunay Neighbor Count.  
The triangulation on a random system has a much higher diversity in neighbor count



# Compression System

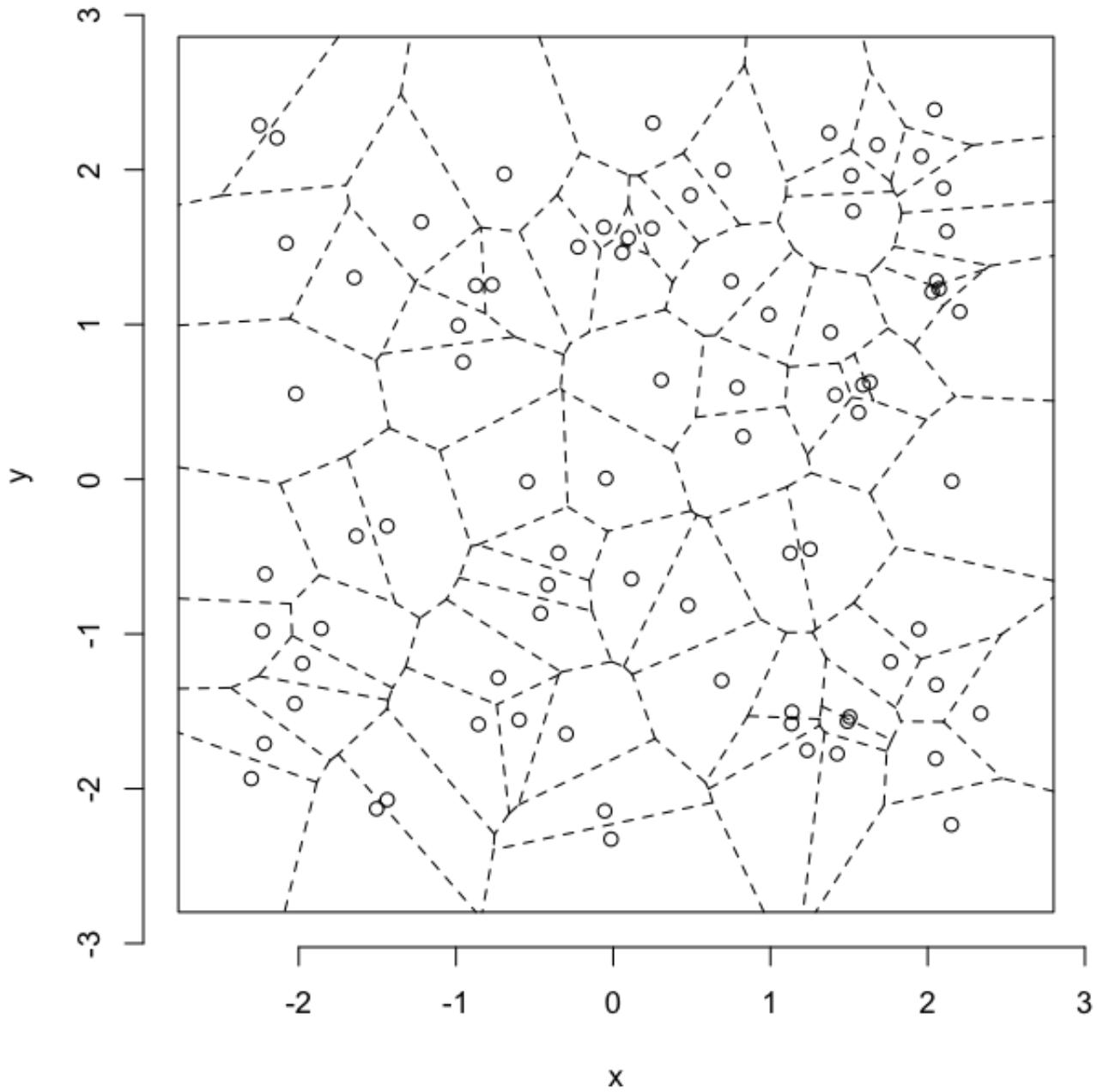
## Compression

+/- R Code



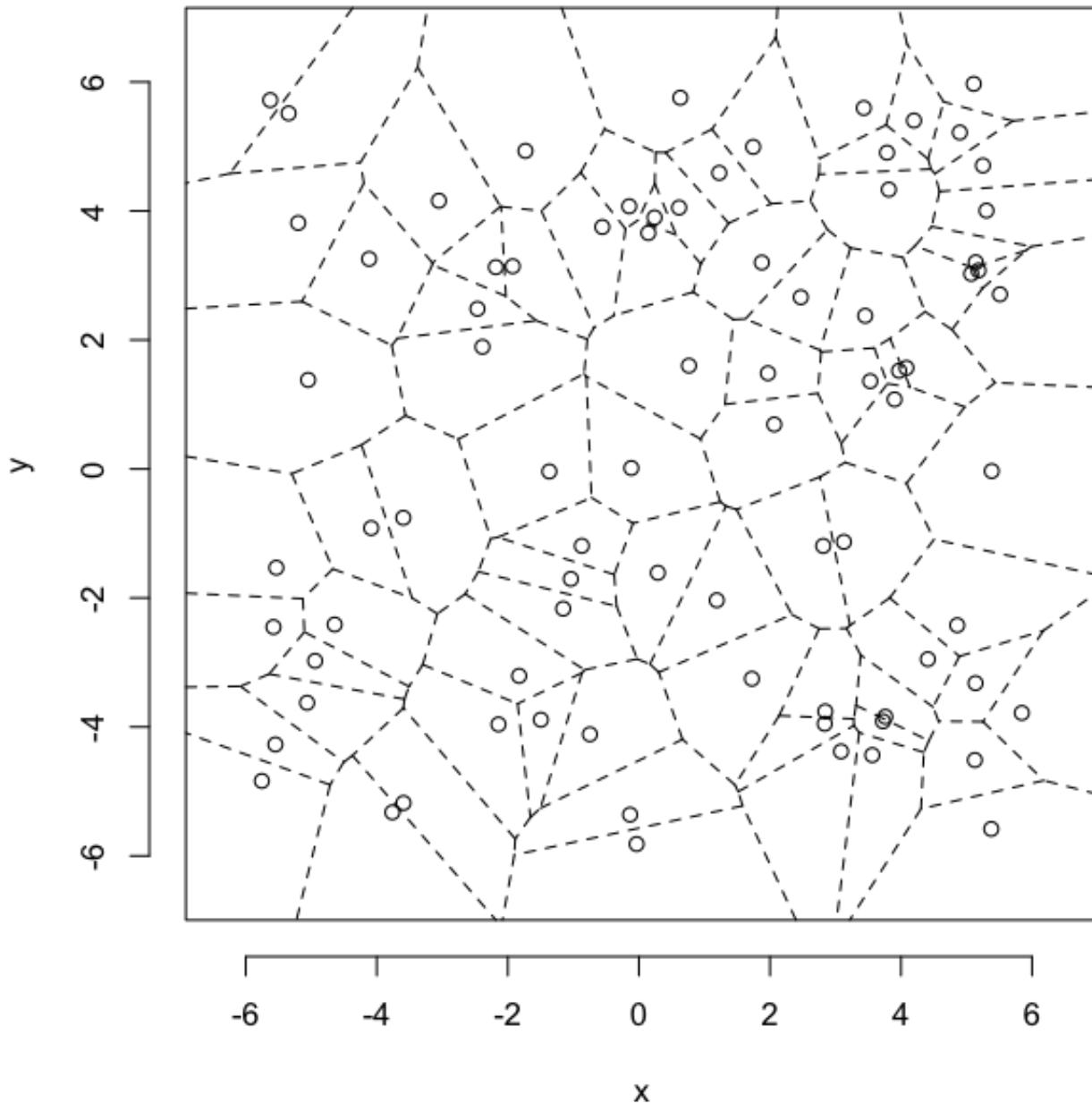
## Tension

+/- R Code



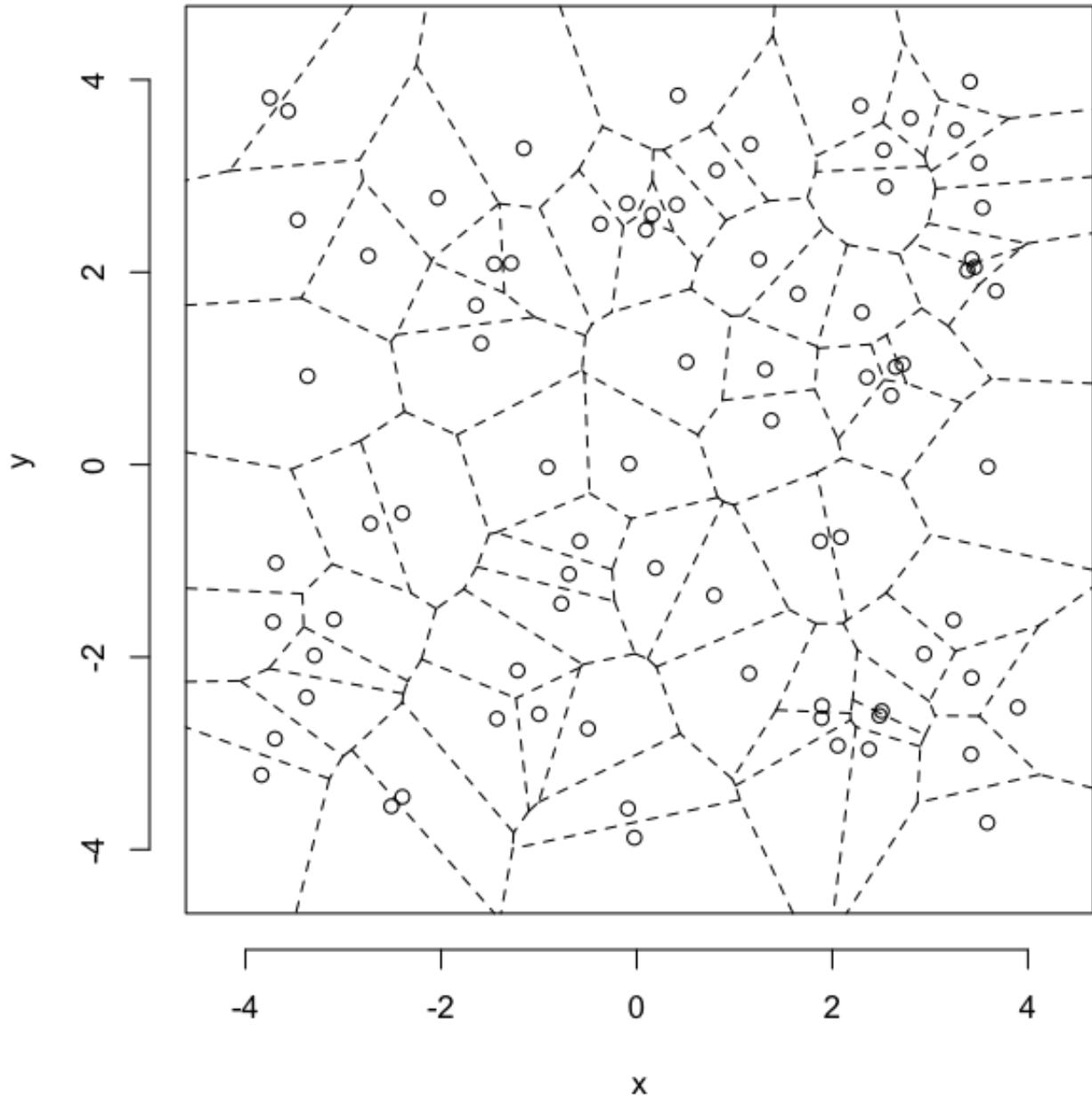
# Shear System

Low Shear



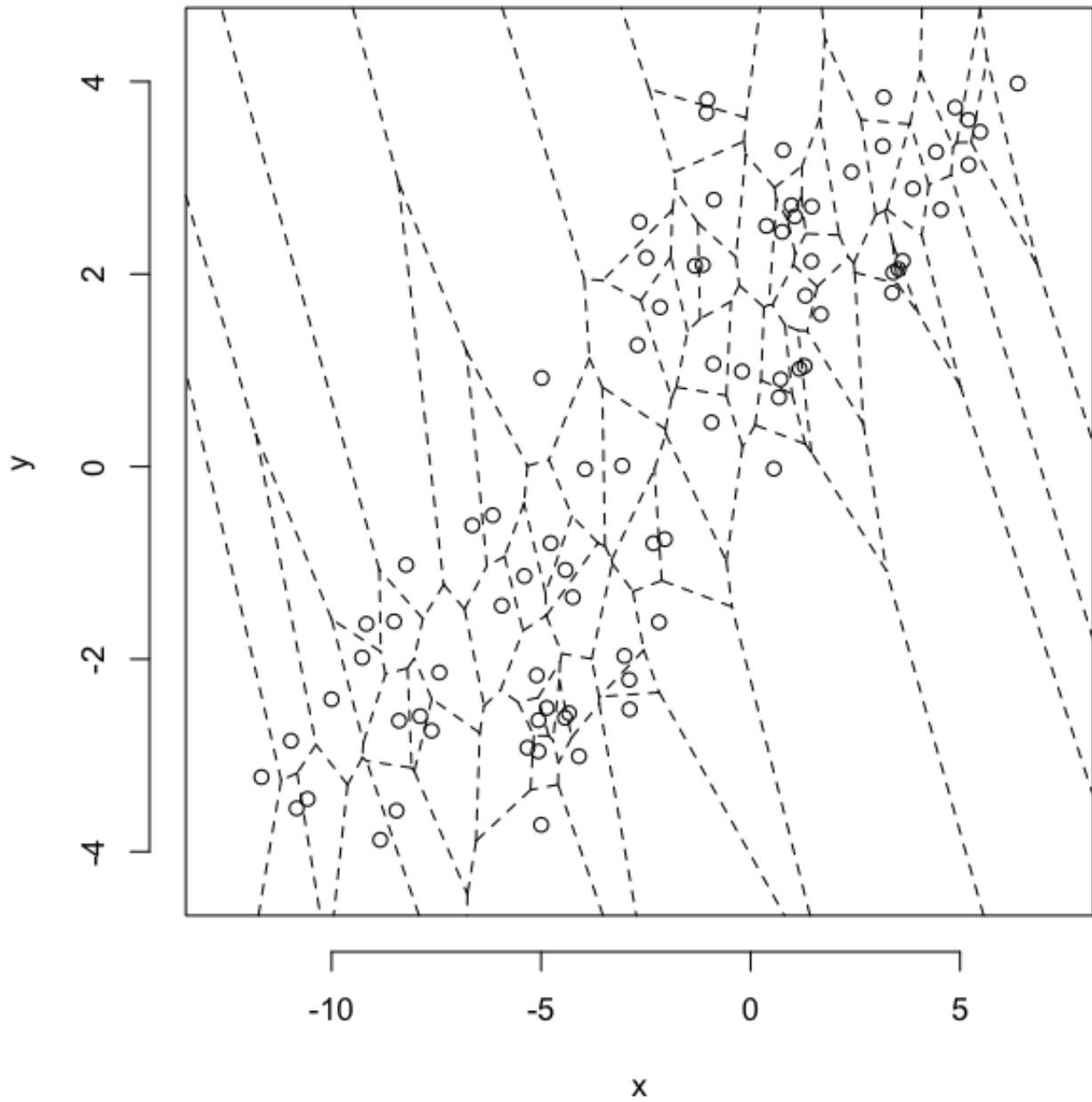
+/- R Code

## High Shear



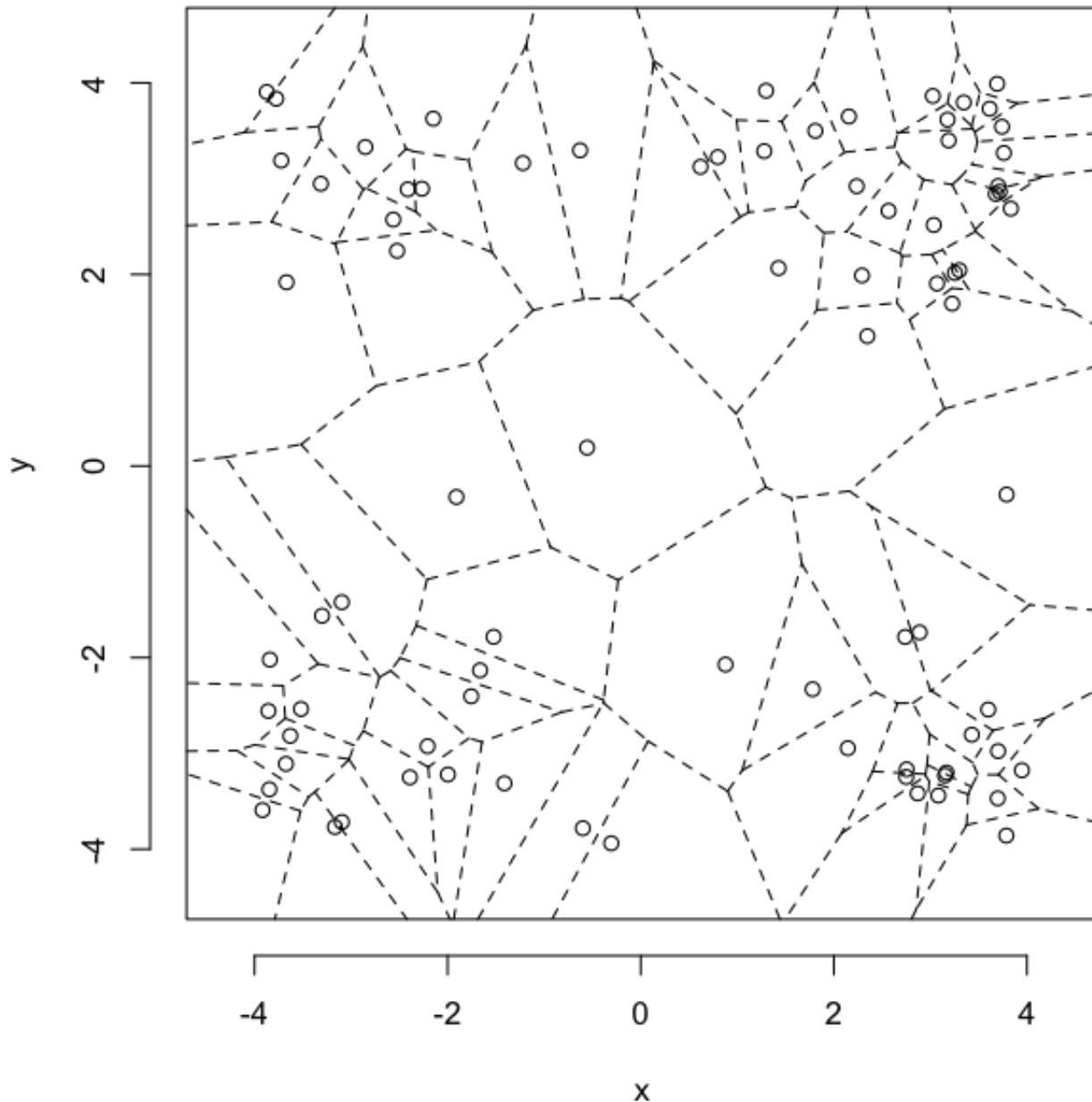
# Stretch System

Low Stretch



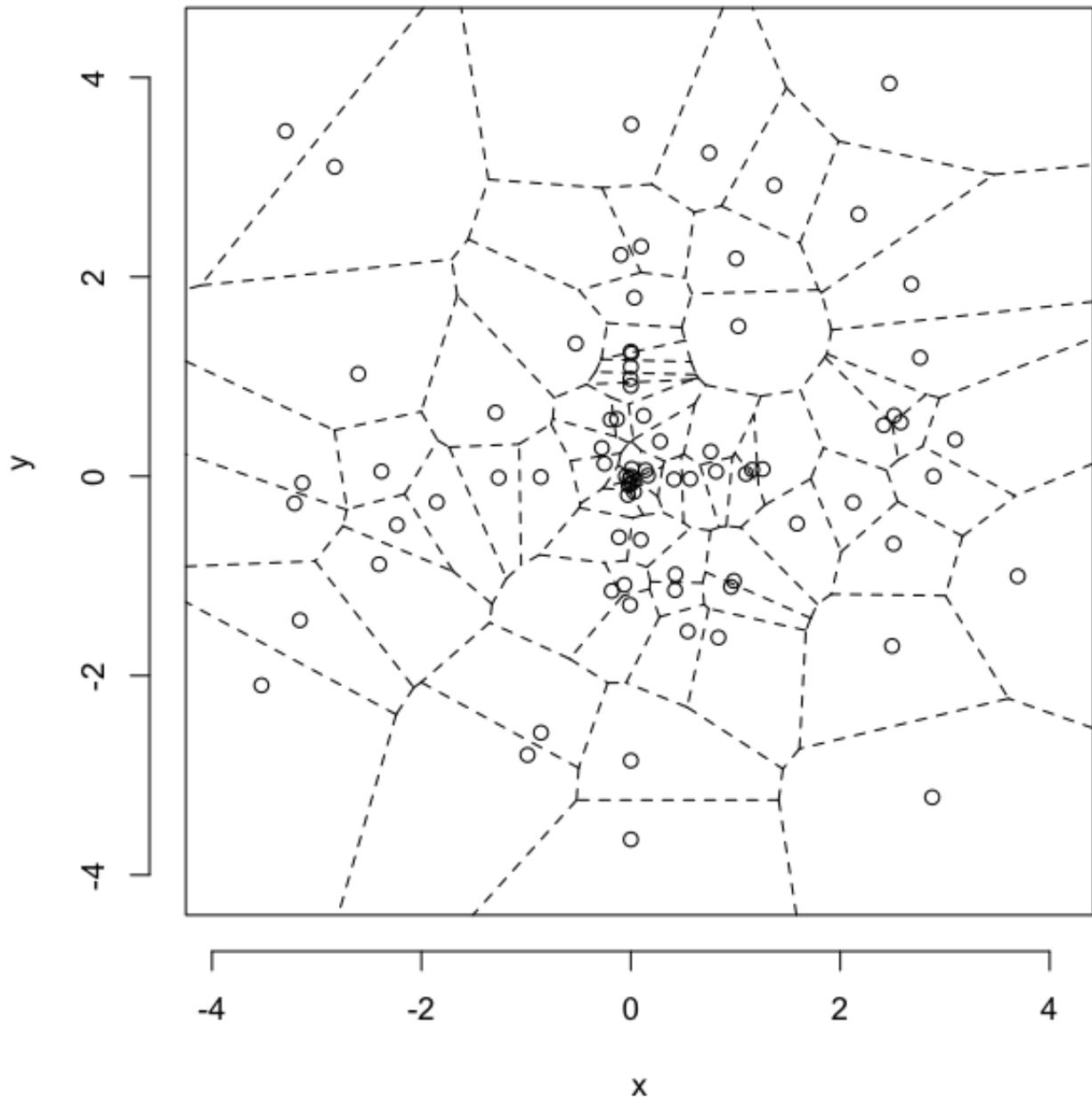
+/- R Code

## Highly Stretched System



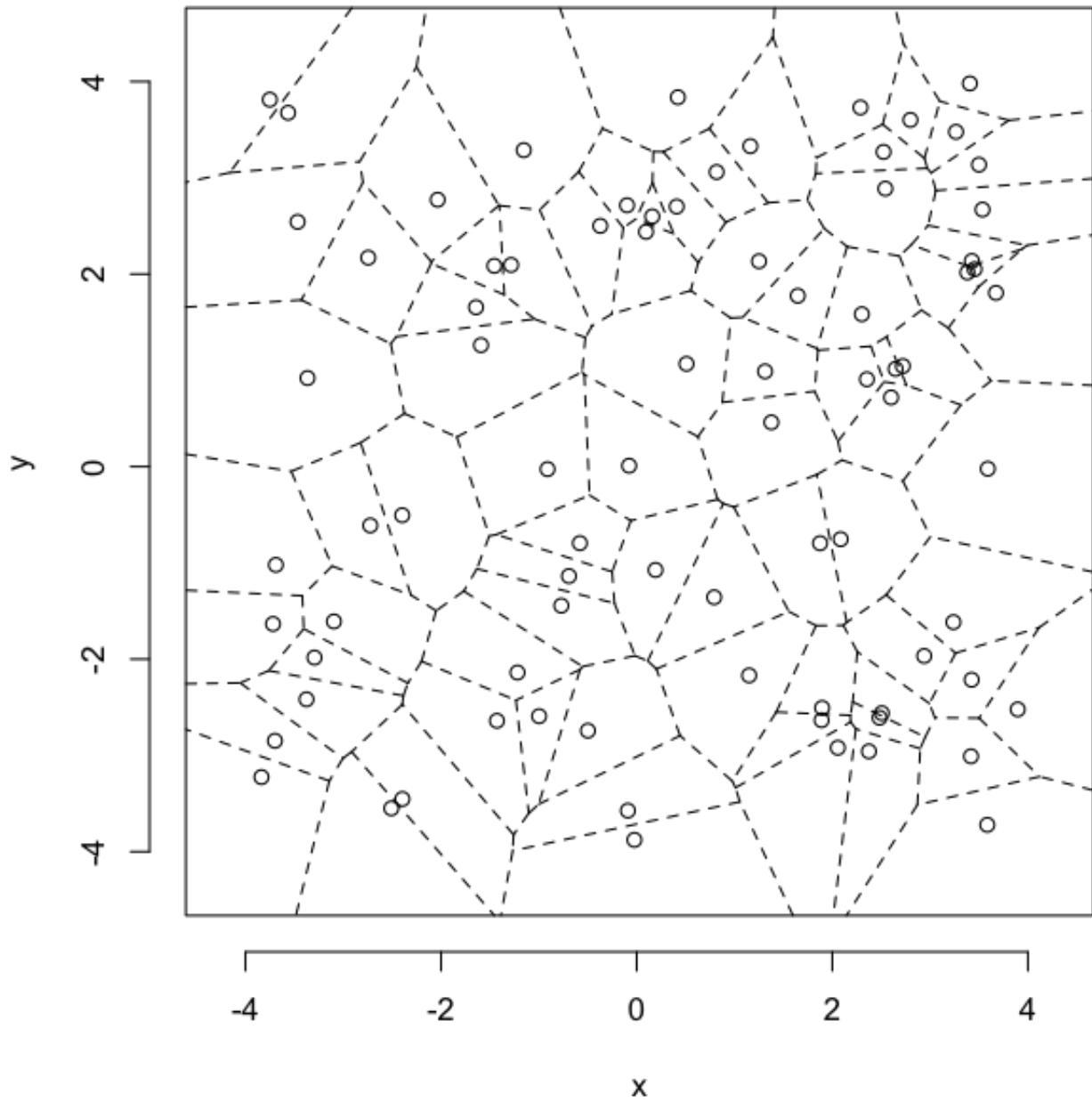
# Swirl System

## Low Swirl System



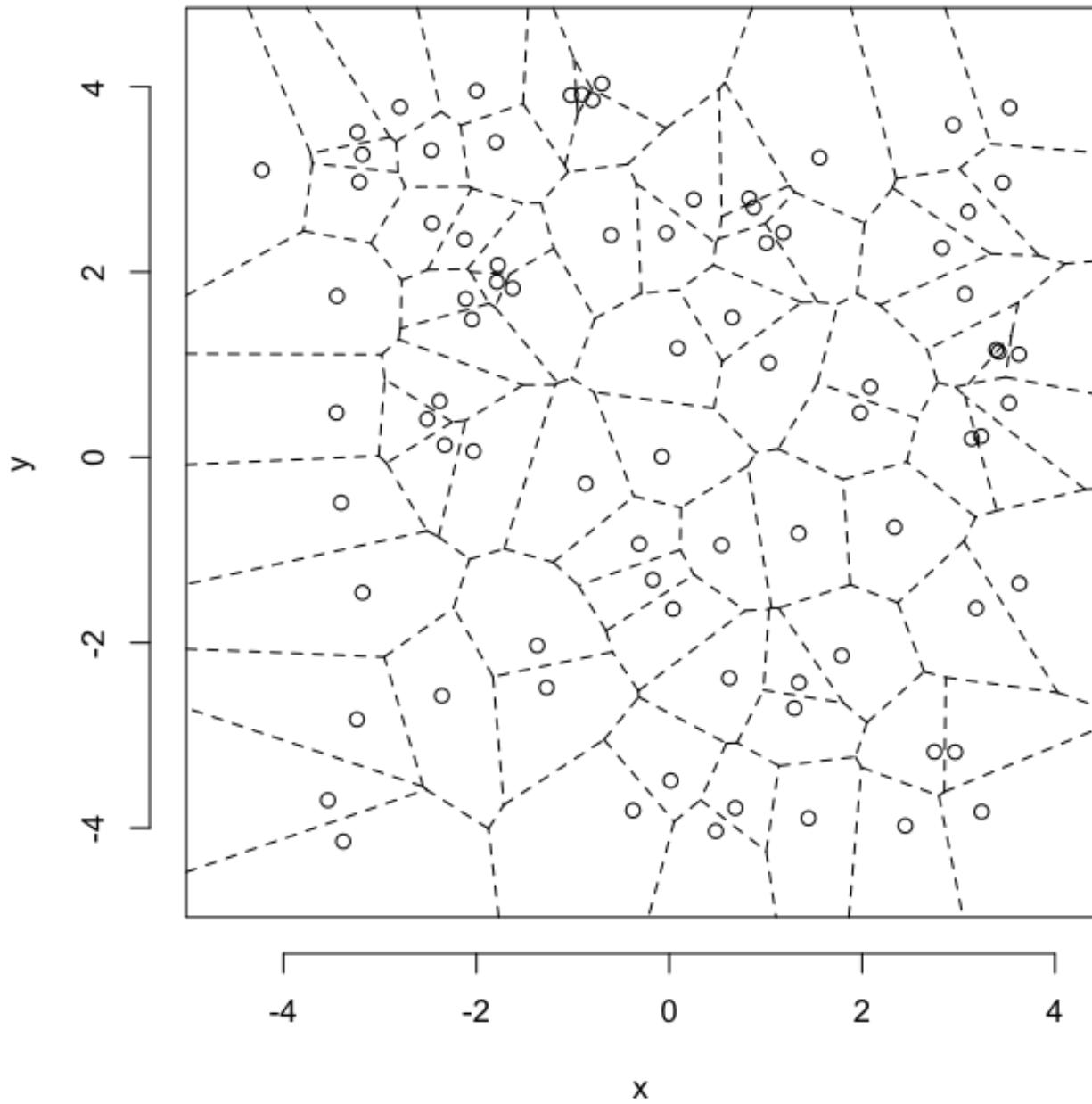
+/- R Code

## High Swirl System



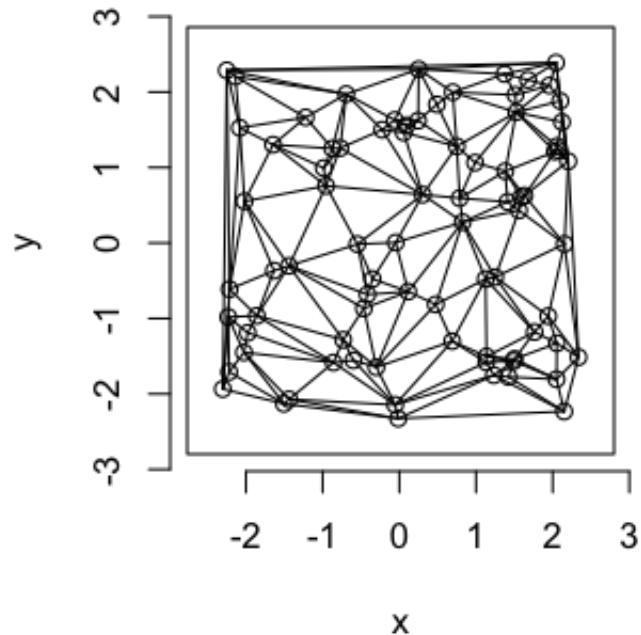
# Neighborhoods

## Compression



+/- R Code

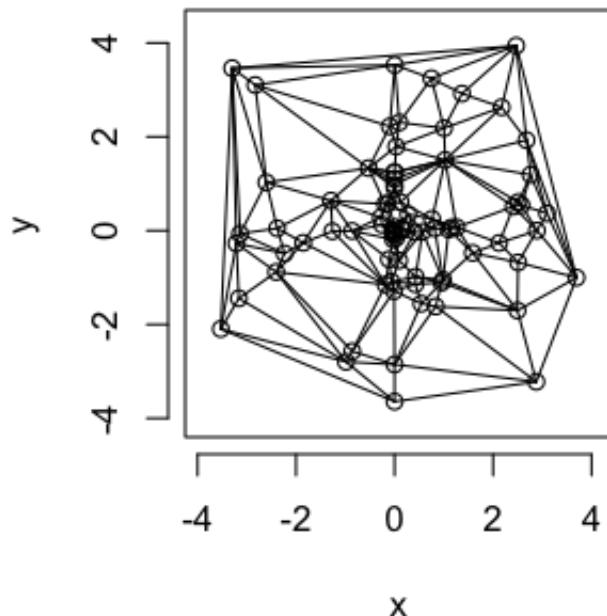
## Stretch



+/- R Code

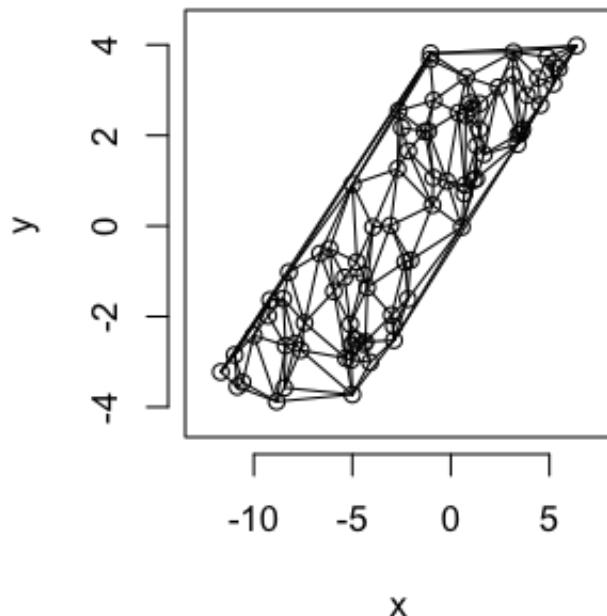
## Shear

+/- R Code



## Swirl

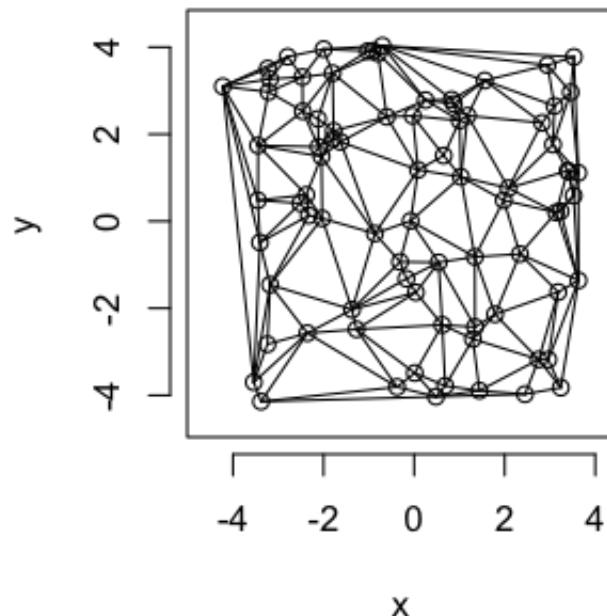
+/- R Code



# Neighbor Count

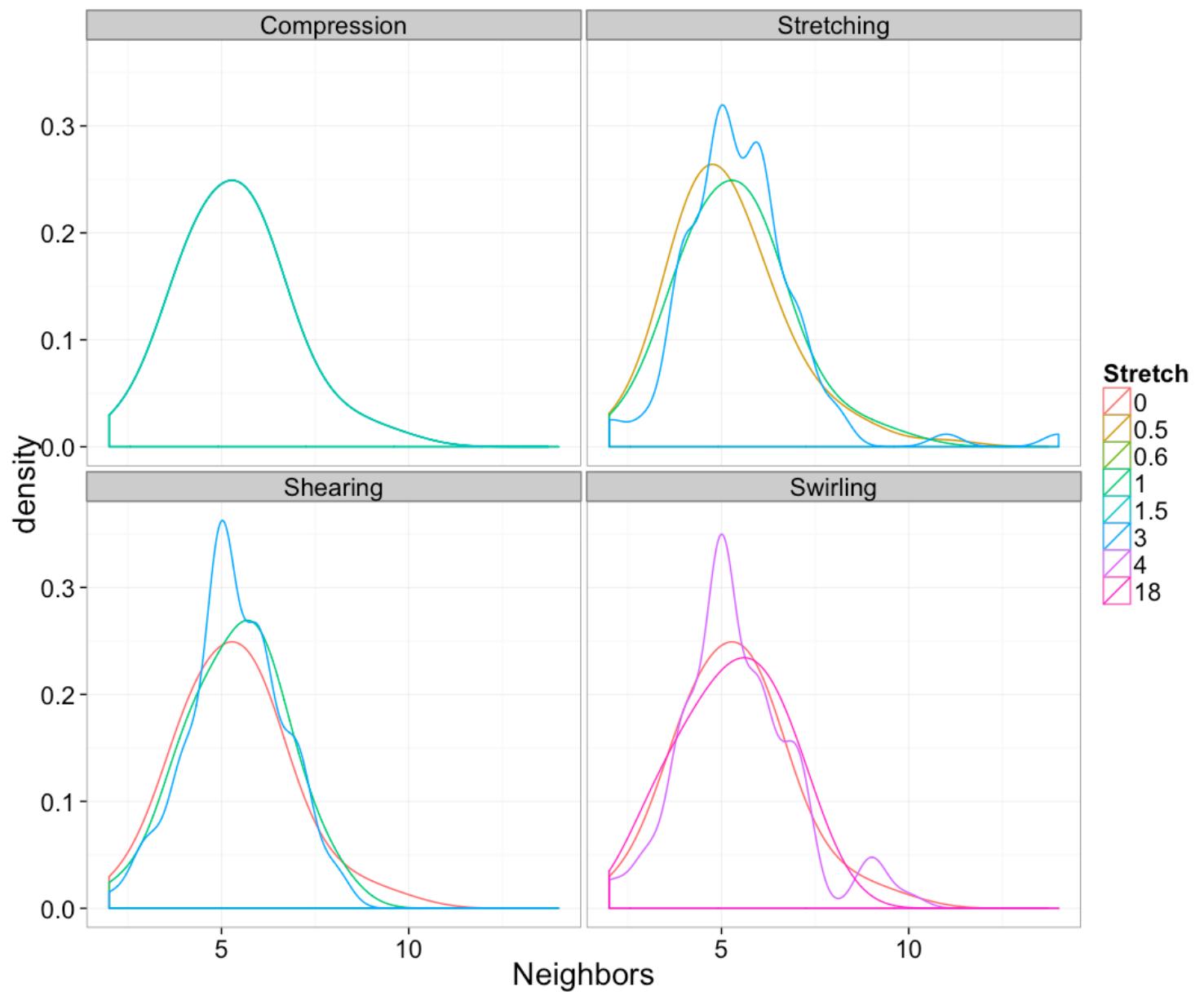
+/- R Code

+/- R Code



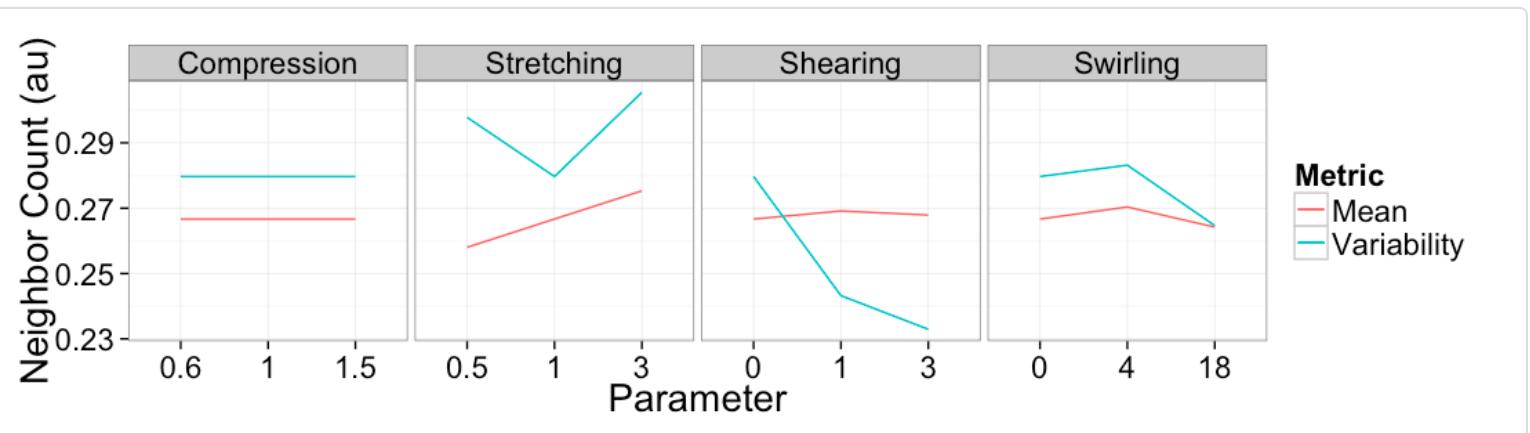
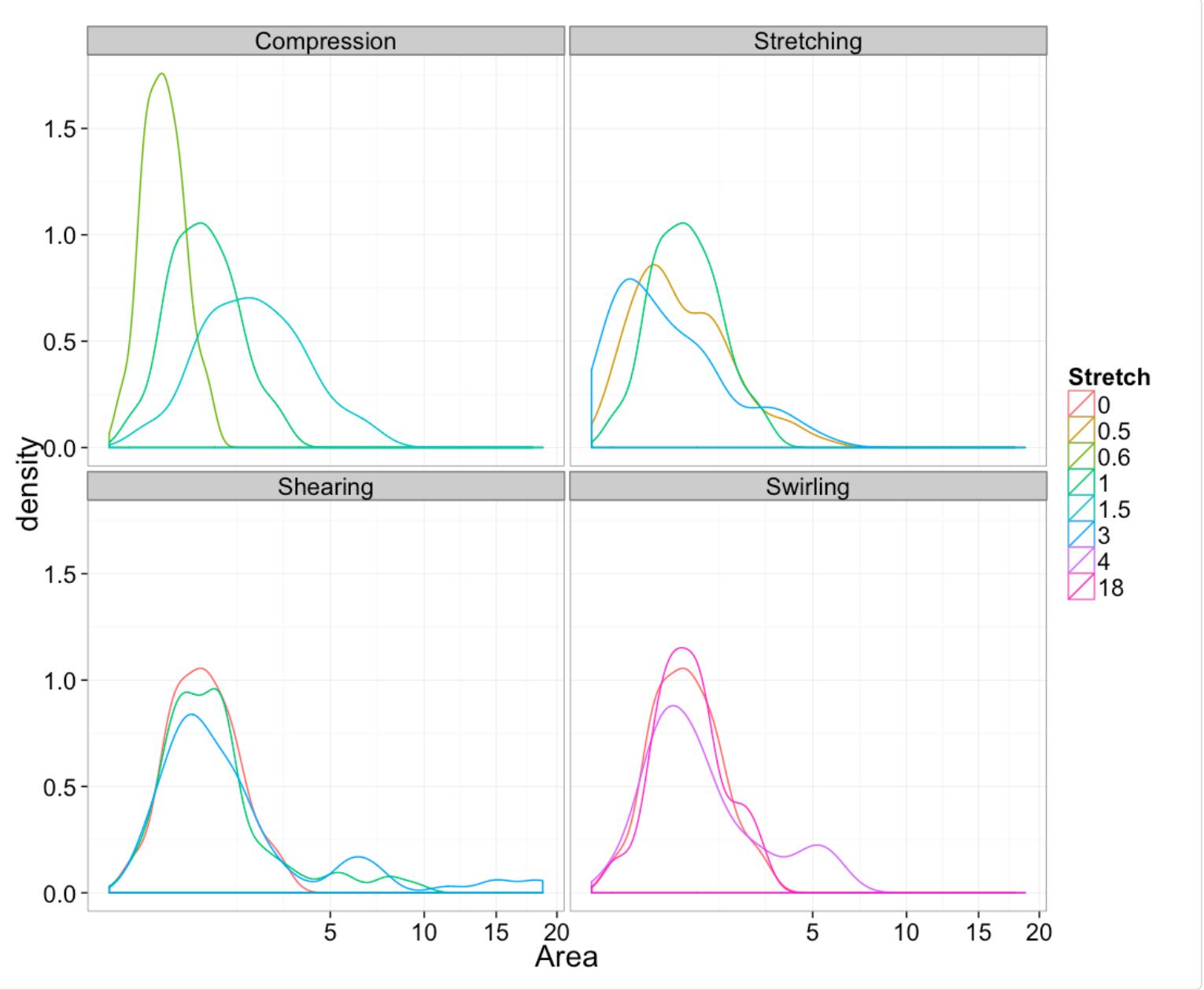
# Volume

+/- R Code



## Mean vs Variability

+/- R Code



+/- R Code

# Where are we at?

We have introduced a number of "operations" we can perform on our objects to change their positions

- compression
- stretching
- shearing
- swirling

We have introduced a number of metrics to characterize our images

- Nearest Neighbor distance
- Nearest Neighbor angle
- Delaunay Neighbor count
- Territory Area (Volume)

A single random systems is useful

- but in order to have a reasonable understanding of the behavior of a system we need to sample many of them.

## Understand metrics as a random system + a known transformation

- We take **mean** values
- Also **coefficient of variation** ([CV](http://en.wikipedia.org/wiki/Coefficient_of_variation) ([http://en.wikipedia.org/wiki/Coefficient\\_of\\_variation](http://en.wikipedia.org/wiki/Coefficient_of_variation))) values since they are "scale-free"

+/- R Code

# Understanding Metrics

In imaging science we always end up with lots of data, the tricky part is understanding the results that come out. With this simulation-based approach

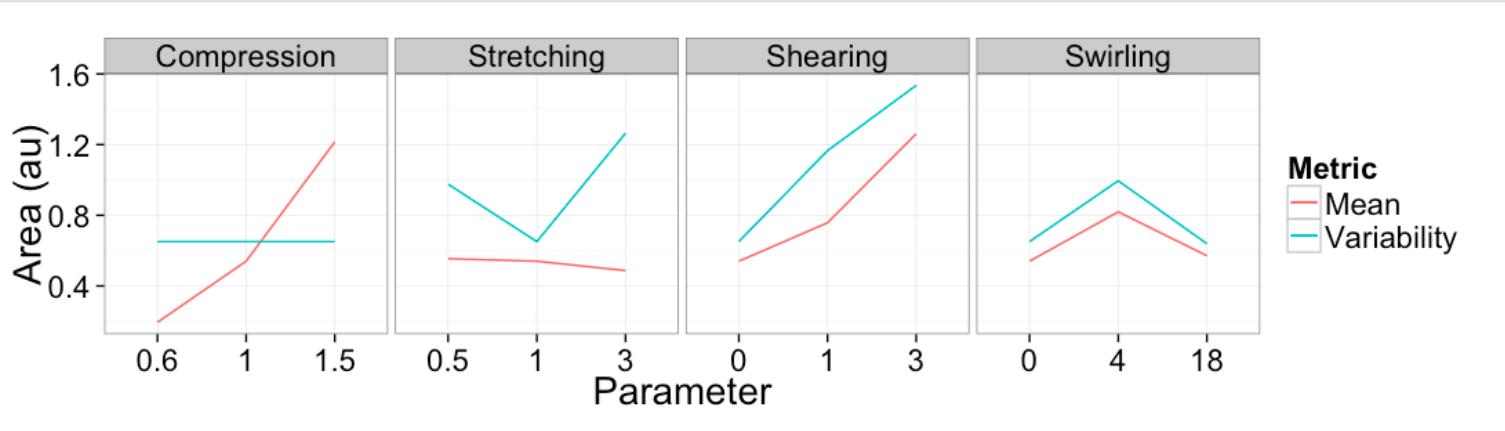
- we generate completely random data
- apply a known transformation to it  $\mathcal{F}$
- quantify the results

We can then take this knowledge and use it to interpret observed data as transformations on an initially random system. We try and find the **rules** used to produce the sample

## Examples

1. Cell distribution in bone
  - Cell position appears random
  - Metrics  $\neq$  Random Statistics
  - Cells are consistently *self-avoiding* or *stretched*
2. Egg-shell Pores
  - Pores in rock / egg shell appear random
  - Metrics  $\neq$  Random Statistics
  - Pores are also *self-avoiding*

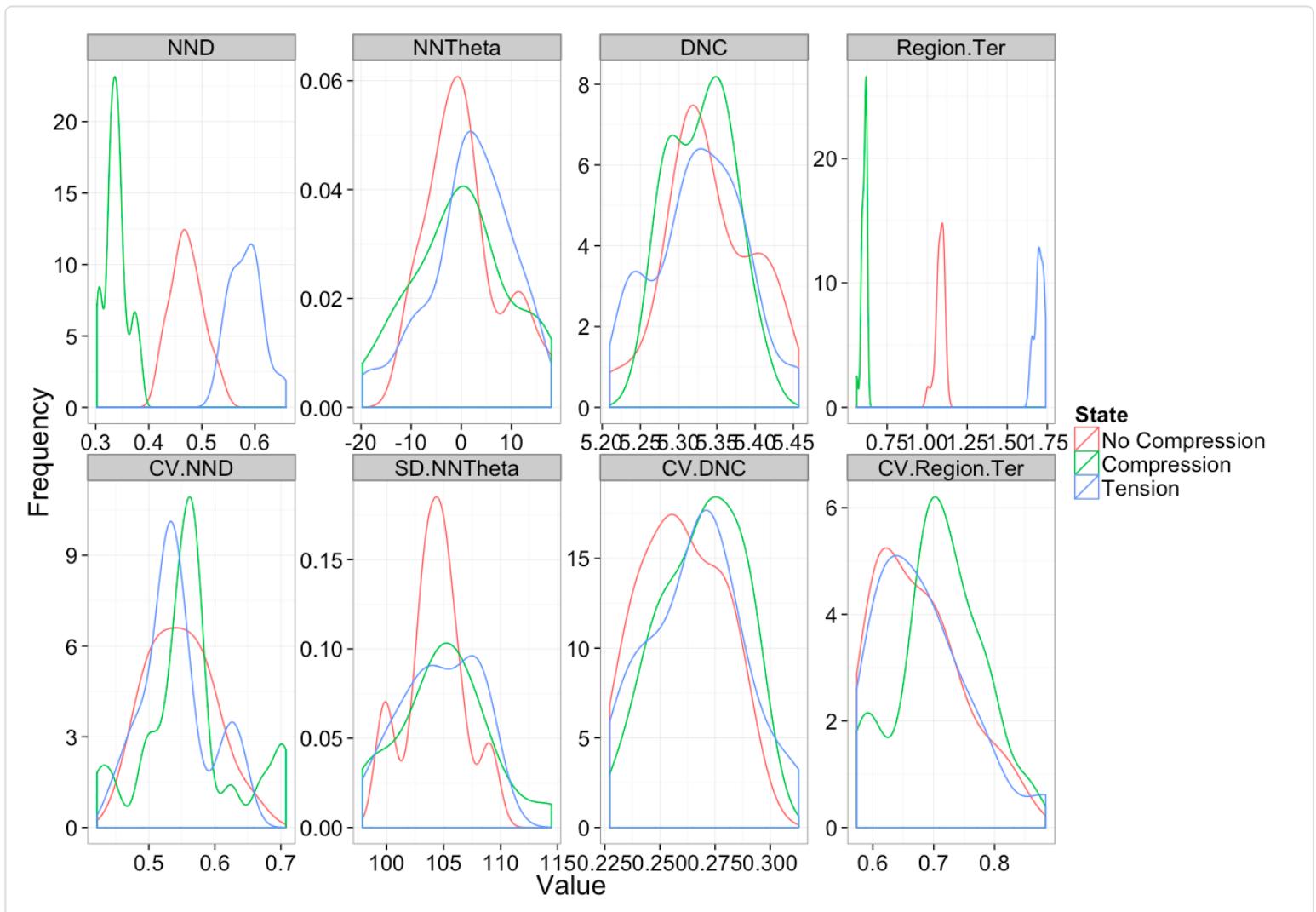
## Compression



+/- R Code

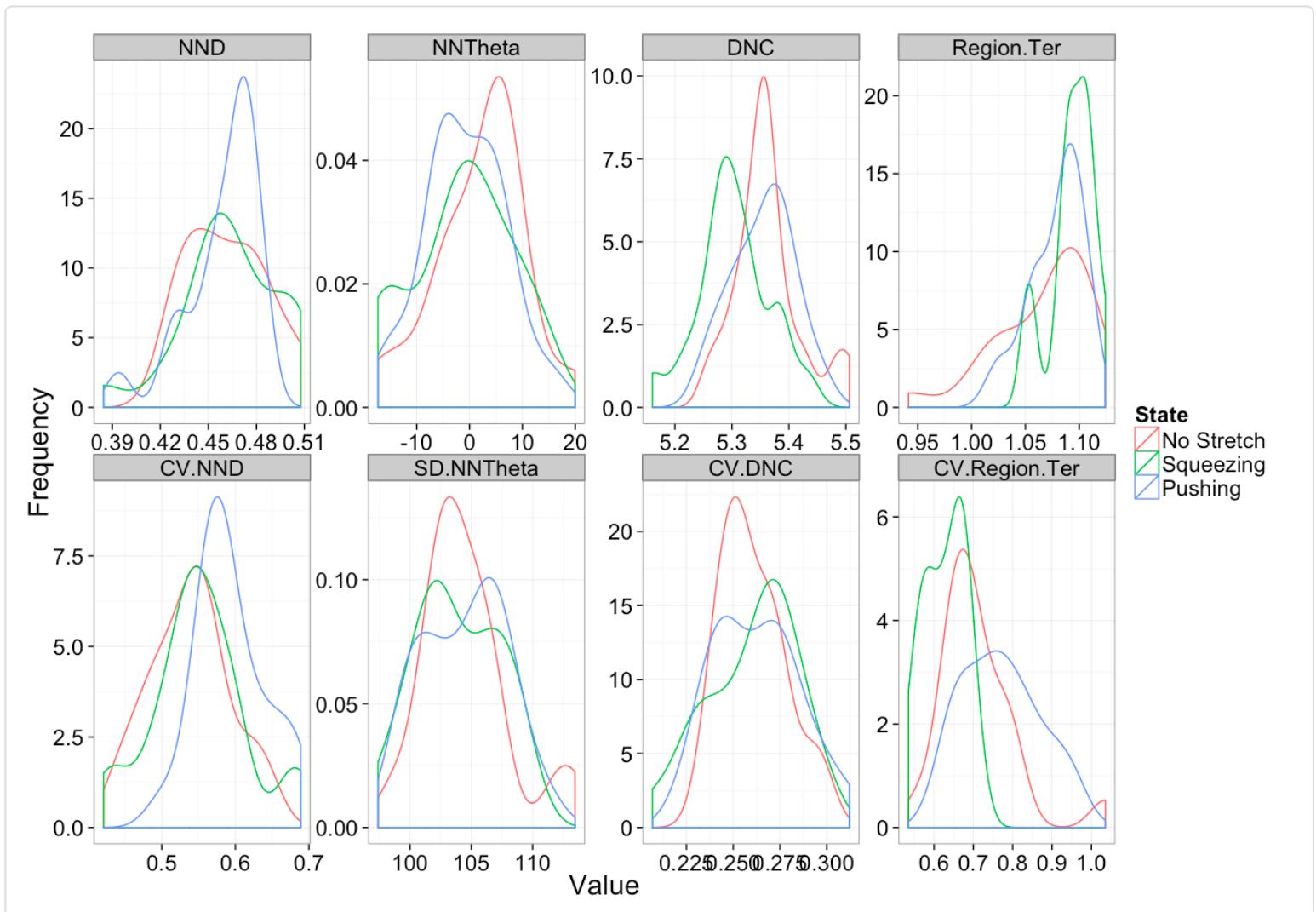
## Compression Sensitivity

+/- R Code



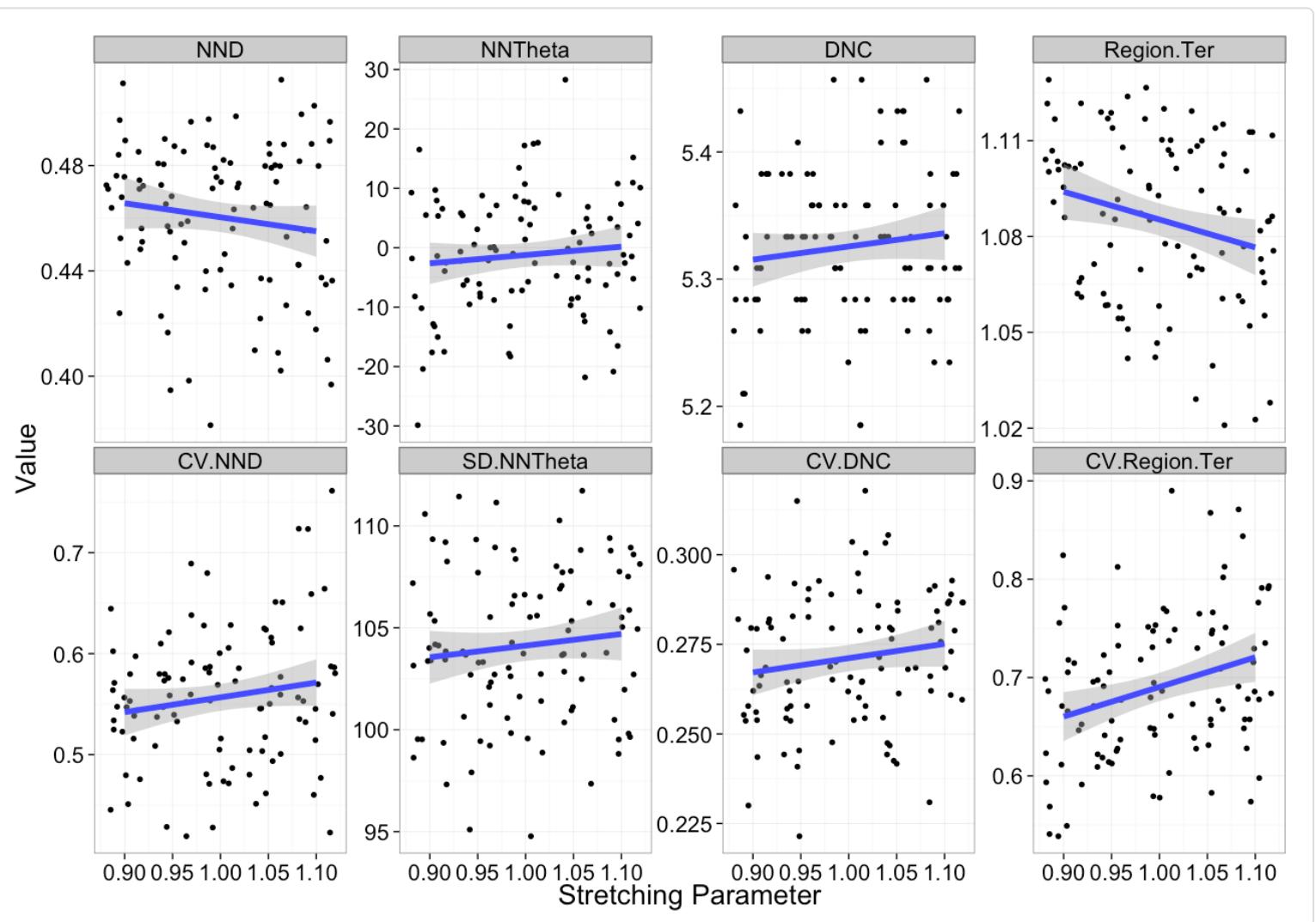
# Stretching

+/- R Code



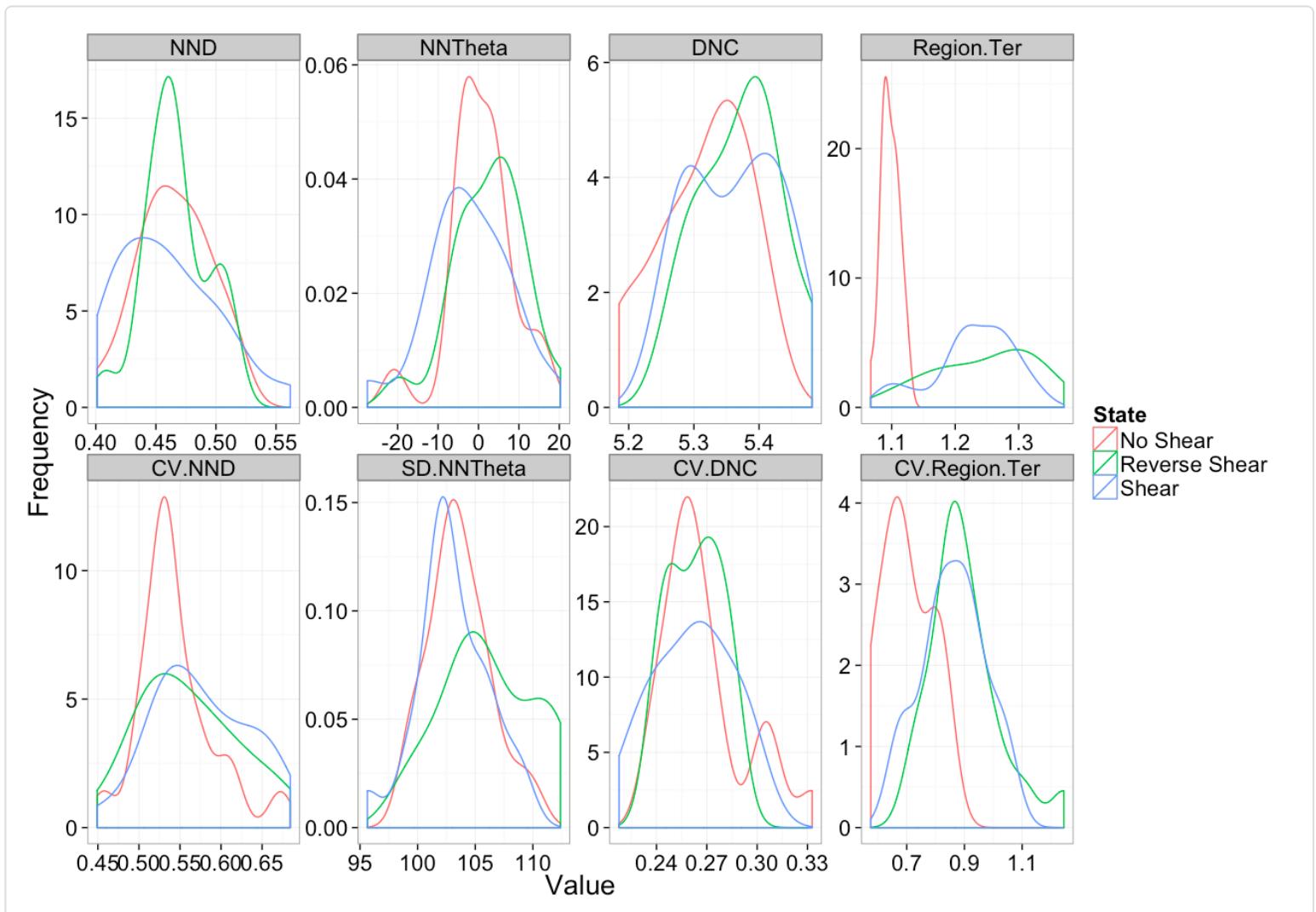
# Stretching Sensitivity

+/- R Code



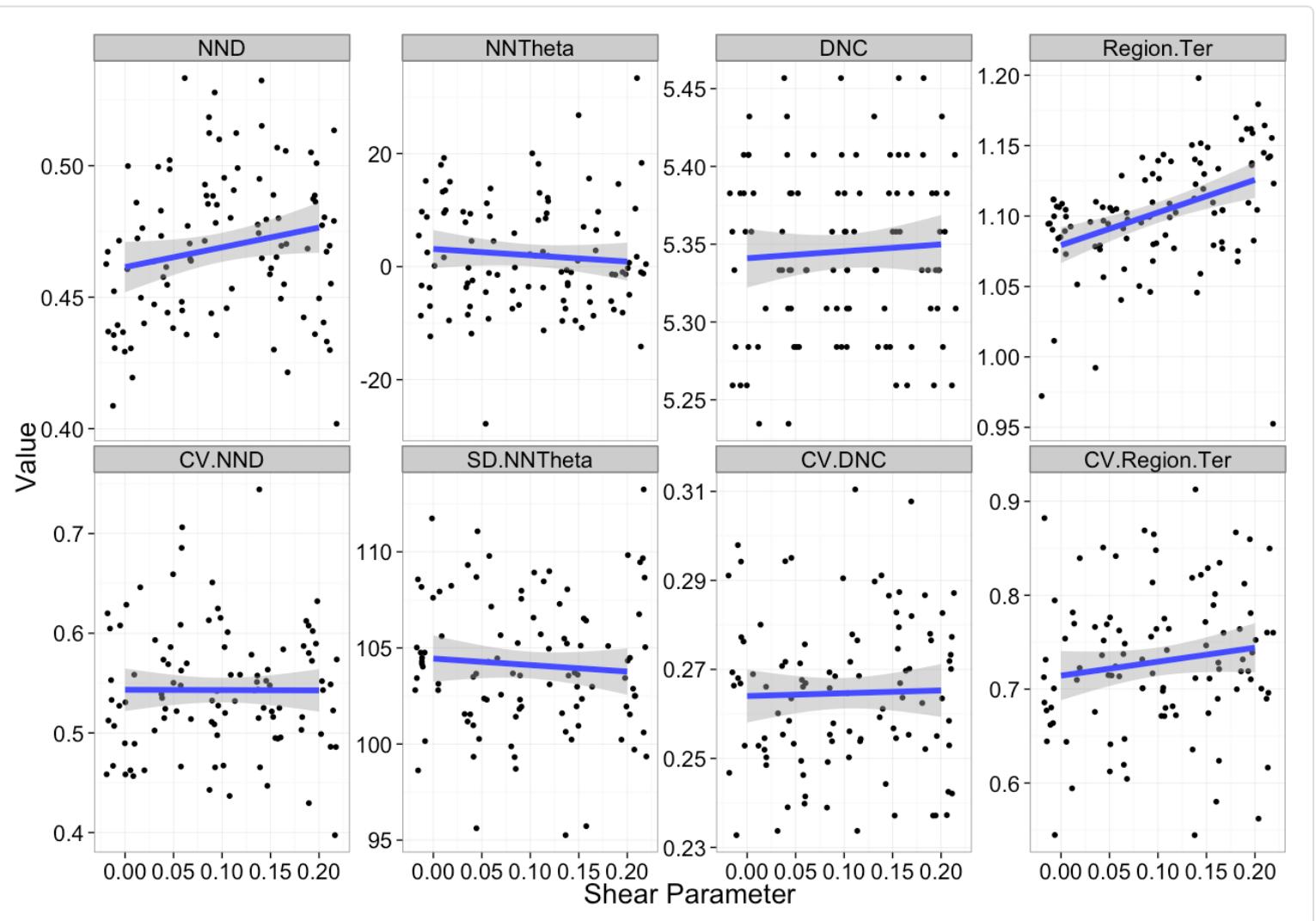
# Shearing

+/- R Code



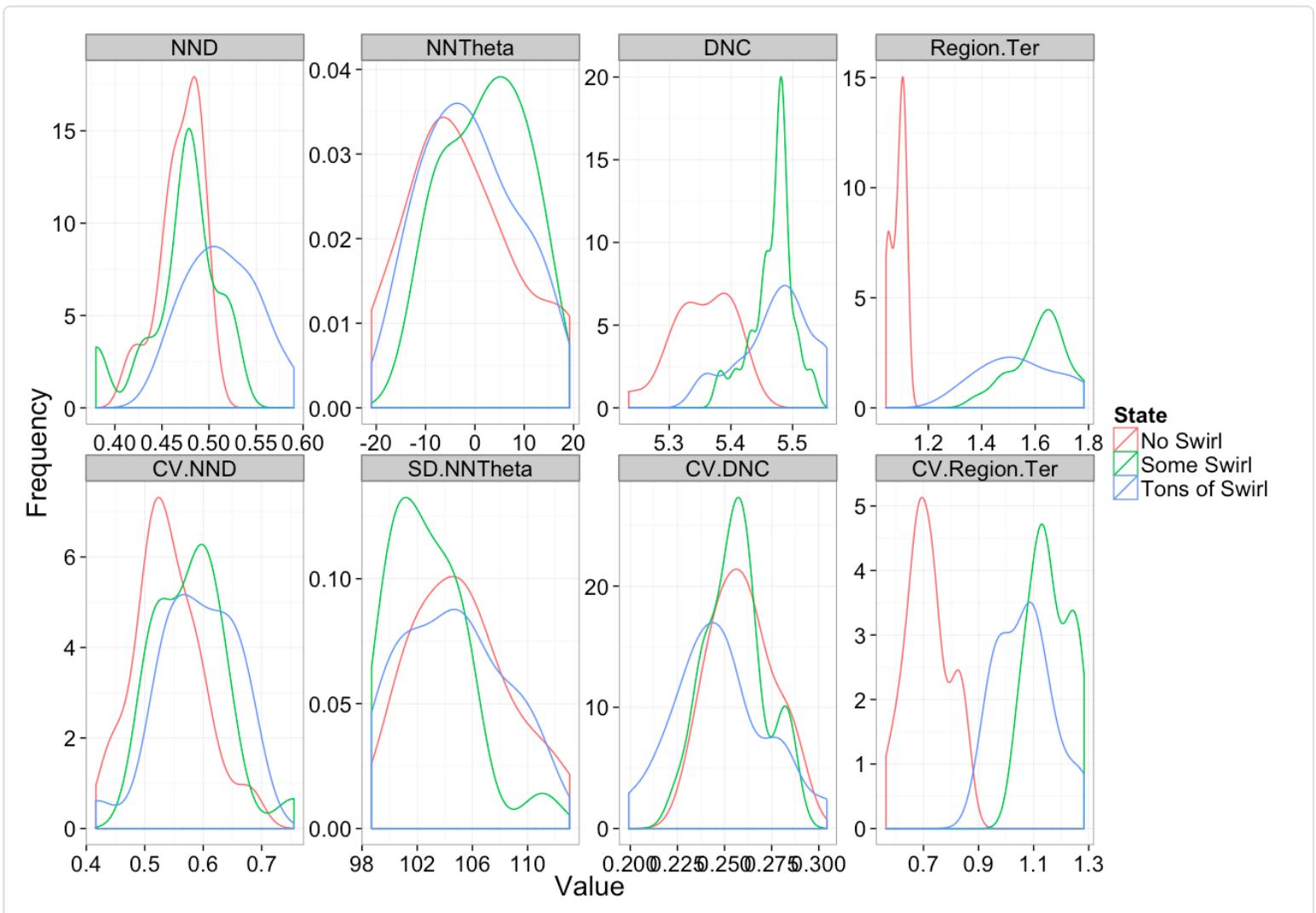
# Shearing Sensitivity

+/- R Code



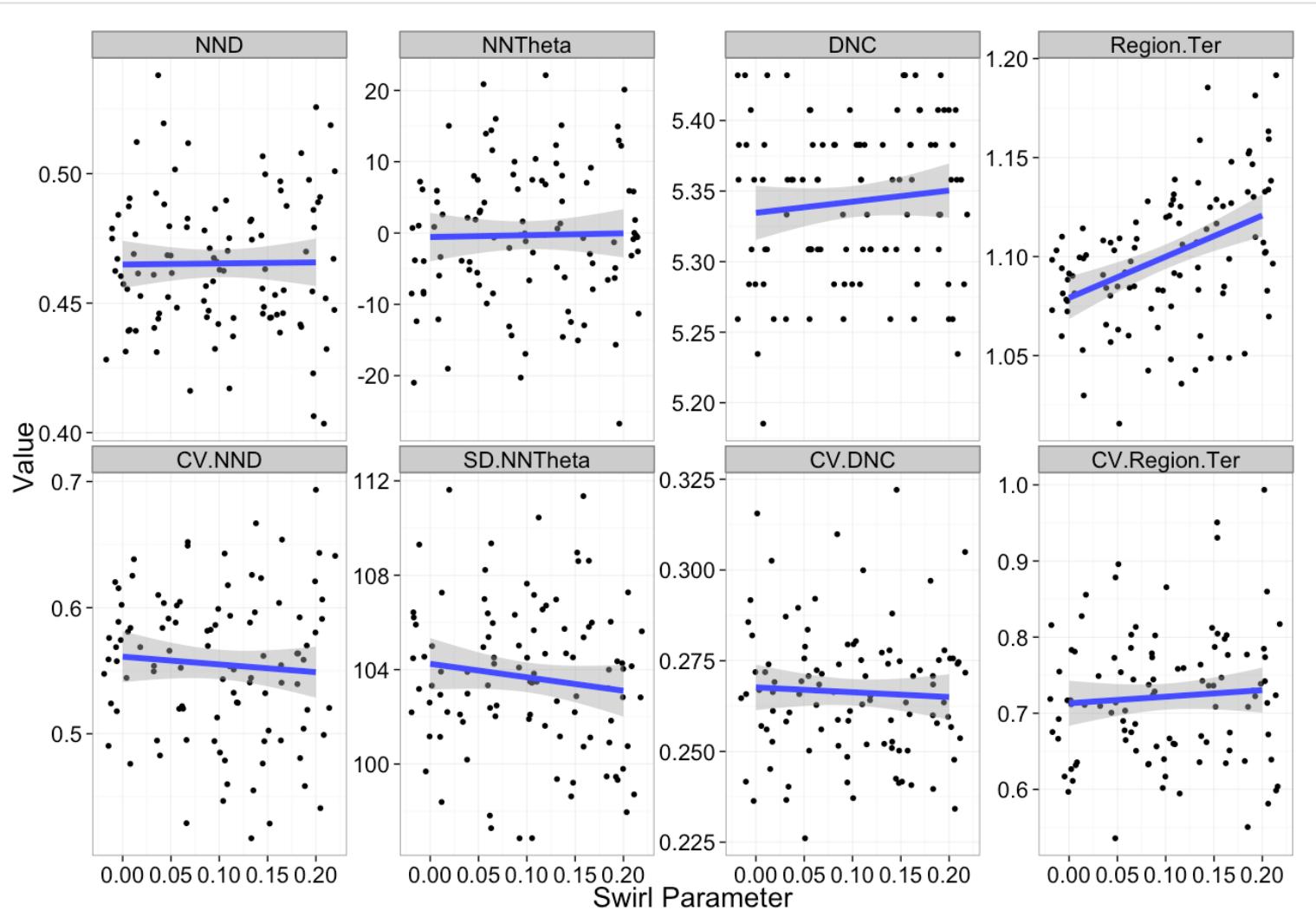
# Swirling

+/- R Code



# Swirling Sensitivity

+/- R Code



# Self-Avoiding

From the nearest neighbor distance metric, we can create a scale-free version of the metric which we call *self-avoiding coefficient* or *grouping*.

The metric is the ratio of

- observed nearest neighbor distance NND
- the expected mean nearest neighbor distance ( $r_0$ ) for a random point distribution (Poisson Point Process) with the same number of points (N.Obj) per volume (Total.Volume).

$$r_0 = \sqrt[3]{\frac{\text{Total.Volume}}{2\pi \text{N.Obj}}}$$

Using the territory we defined earlier (Region area/volume) we can simplify the definition to

$$r_0 = \sqrt[3]{\frac{\bar{T}_{\text{er}}}{2\pi}}$$

$$\text{SAC} = \frac{\text{NND}}{\sqrt[3]{\frac{\bar{T}_{\text{er}}}{2\pi}}}$$

# Distribution Tensor

So the information we have is 3D why are we taking single metrics (distance, angle, volume) to quantify it.

- Shouldn't we use 3D metrics with 3D data?
  - Just like the shape tensor we covered before, we can define a *distribution* tensor to characterize the shape of the distribution.
  - The major difference instead of constituting voxels we use *edges*
    - an edge is defined from the Delaunay triangulation
    - it connects two neighboring bubbles together
- 
- We can calculate distribution for a single bubble by taking all edges that touch the object or from a region by finding all edges inside that region

We start off by calculating the covariance matrix from the list of edges  $\vec{v}_{ij}$  in a given volume  $\mathcal{V}$

$$\vec{v}_{ij} = \vec{\text{COV}}(i) - \vec{\text{COV}}(j)$$

$$\text{COV}(\mathcal{V}) = \frac{1}{N} \sum_{\forall \text{COM}(i) \in \mathcal{V}} \begin{bmatrix} \vec{v}_x \vec{v}_x & \vec{v}_x \vec{v}_y & \vec{v}_x \vec{v}_z \\ \vec{v}_y \vec{v}_x & \vec{v}_y \vec{v}_y & \vec{v}_y \vec{v}_z \\ \vec{v}_z \vec{v}_x & \vec{v}_z \vec{v}_y & \vec{v}_z \vec{v}_z \end{bmatrix}$$

# Distribution Tensor (continued)

We then take the eigentransform of this array to obtain the eigenvectors (principal components,  $\vec{\Lambda}_{1\dots 3}$ ) and eigenvalues (scores,  $\lambda_{1\dots 3}$ )

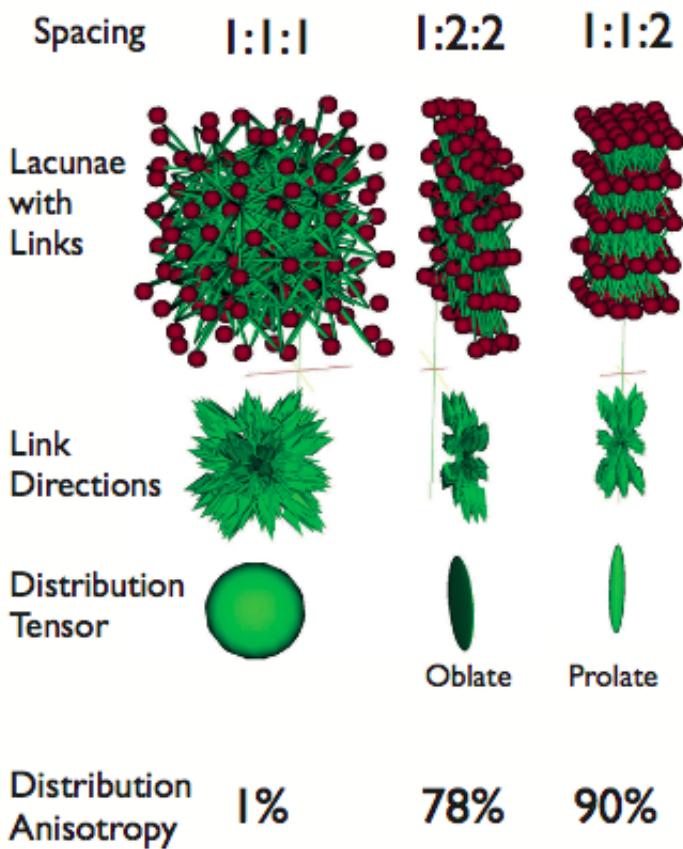
$$\text{COV}(I_{id}) \longrightarrow \underbrace{\begin{bmatrix} \vec{\Lambda}_{1x} & \vec{\Lambda}_{1y} & \vec{\Lambda}_{1z} \\ \vec{\Lambda}_{2x} & \vec{\Lambda}_{2y} & \vec{\Lambda}_{2z} \\ \vec{\Lambda}_{3x} & \vec{\Lambda}_{3y} & \vec{\Lambda}_{3z} \end{bmatrix}}_{\text{Eigenvectors}} * \underbrace{\begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}}_{\text{Eigenvalues}} * \underbrace{\begin{bmatrix} \vec{\Lambda}_{1x} & \vec{\Lambda}_{1y} & \vec{\Lambda}_{1z} \\ \vec{\Lambda}_{2x} & \vec{\Lambda}_{2y} & \vec{\Lambda}_{2z} \\ \vec{\Lambda}_{3x} & \vec{\Lambda}_{3y} & \vec{\Lambda}_{3z} \end{bmatrix}}^T$$

The principal components tell us about the orientation of the object and the scores tell us about the corresponding magnitude (or length) in that direction.

# Distribution Anisotropy

Visual example

- Tensor represents the average spacing between objects in each direction  $\approx$  thickness of background.
- Its interpretation is more difficult since it doesn't represent a *real* object



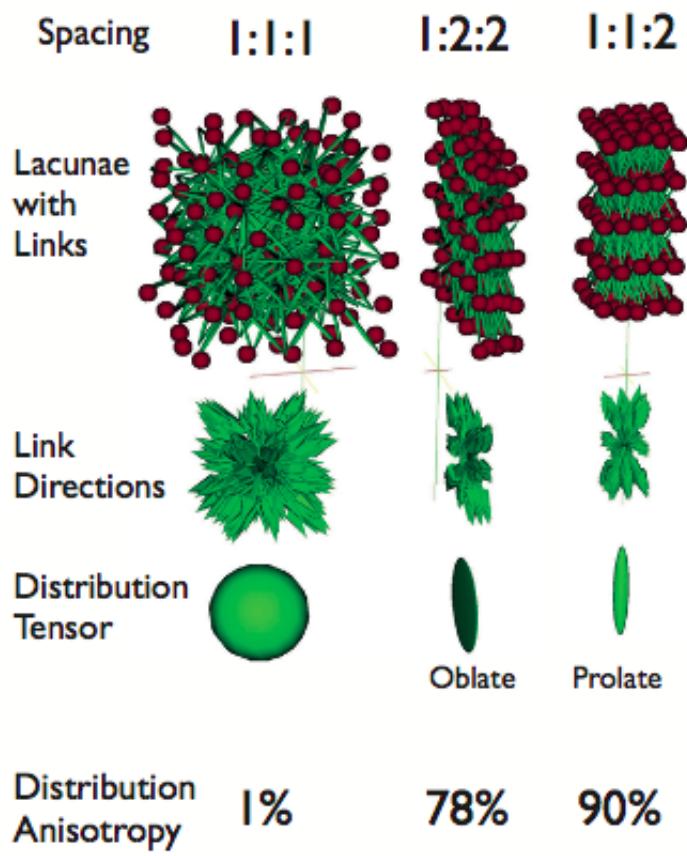
From this tensor we can define an anisotropy in the same manner as we defined for shapes. The anisotropy defined as before

$$Aiso = \frac{\text{Longest Side} - \text{Shortest Side}}{\text{Longest Side}}$$

- An isotropic distribution tensor indicates the spacing between objects is approximately the same in every direction
  - Does not mean a grid, organized, or evenly spaced!
  - Just the same in every direction
- Anisotropic distribution tensor means the spacing is smaller in one direction than the others

- Can be regular or grid-like
- Just closer on one direction

# Distribution Oblateness



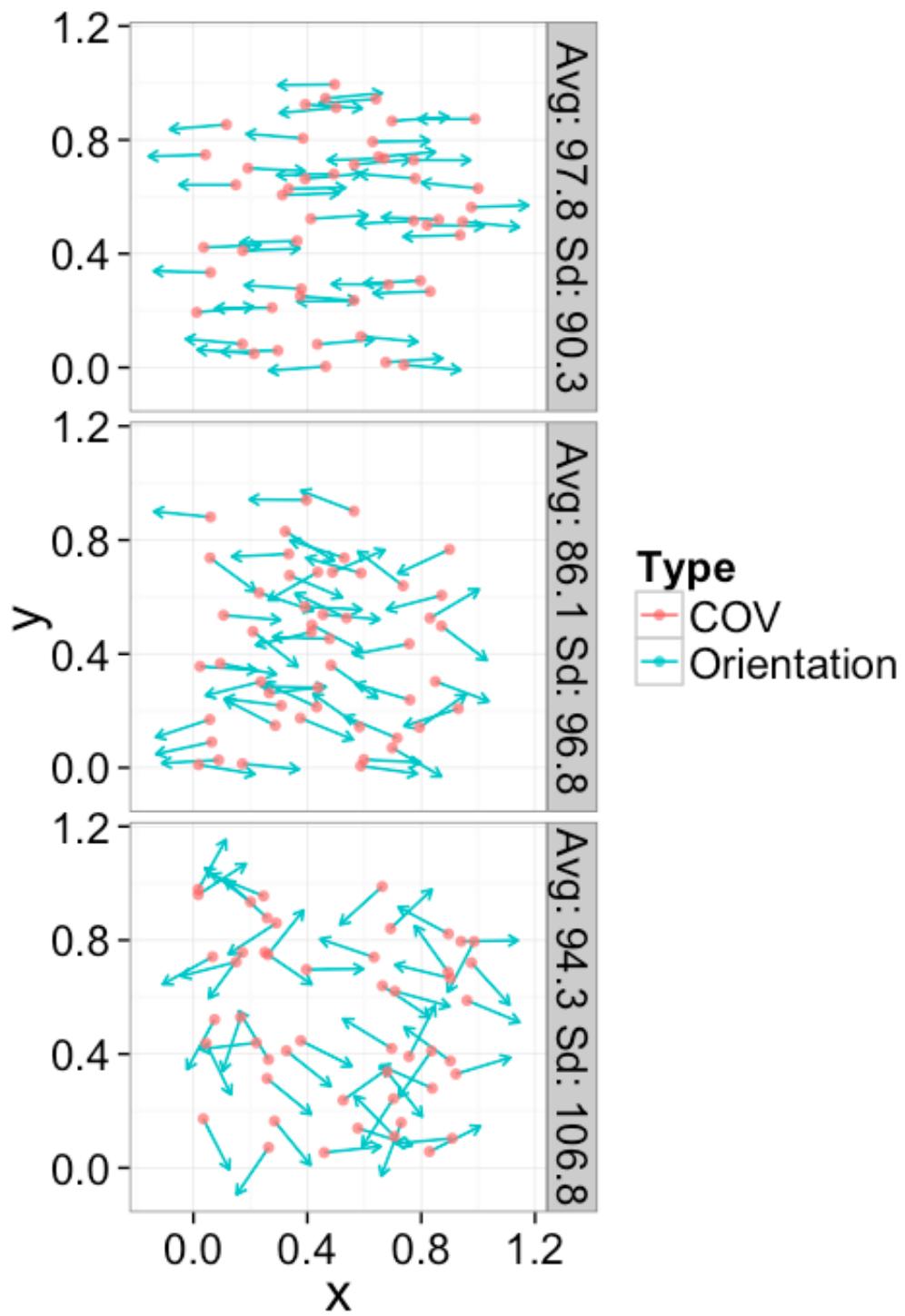
From this tensor we can also define oblateness in the same manner as we defined for shapes. The oblateness is also defined as before as a *type* of anisotropy

$$Ob = 2 \frac{\lambda_2 - \lambda_1}{\lambda_3 - \lambda_1} - 1$$

- Indicates the pancakeness of the distribution
- Oblate means it is very closely spaced in 2 directions and far in the other
  - strand-like structure
- Prolate means it is close in one direction and far in the other 2
  - sheet-like structure

# Orientation

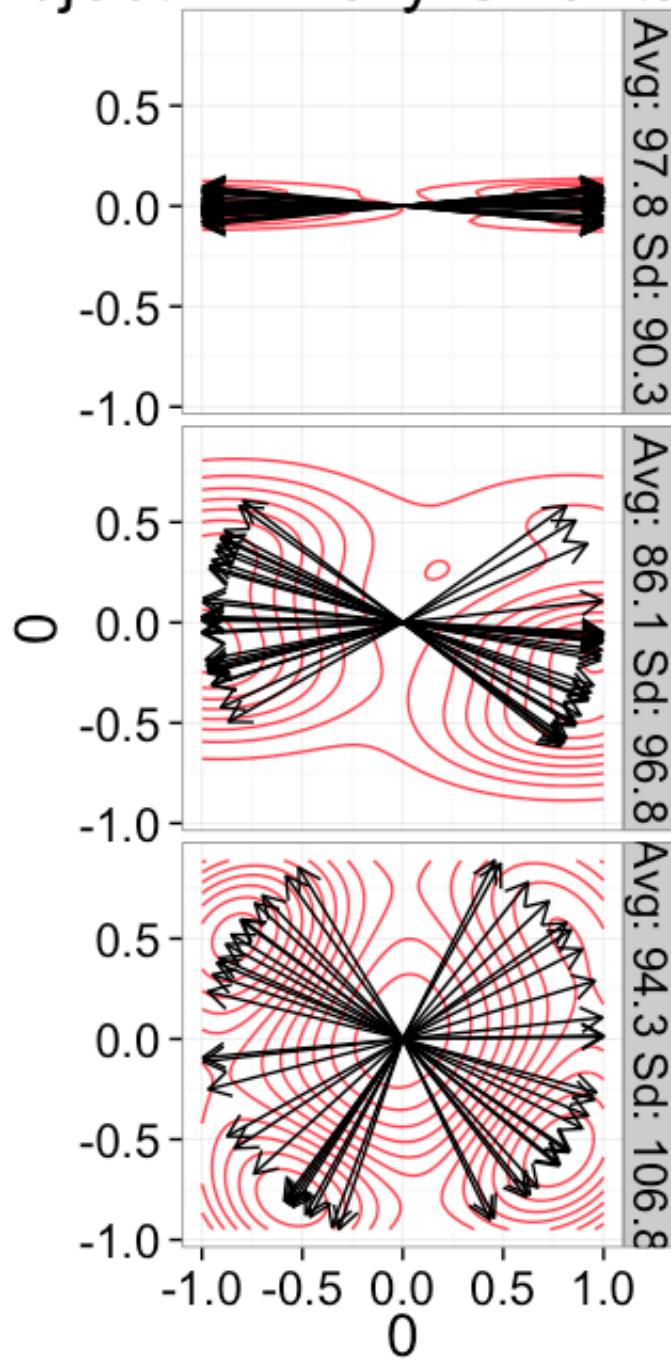
The shape tensor provides for each object 3 possible orientations (each of the eigenvectors).



As we saw in earlier plots it is difficult to express this information in point or graph format

+/- R Code

## Object Primary Orientations



## Alignment Tensor

We can again take advantage of the versatility of a tensor representation for our data and use an *alignment tensor*.

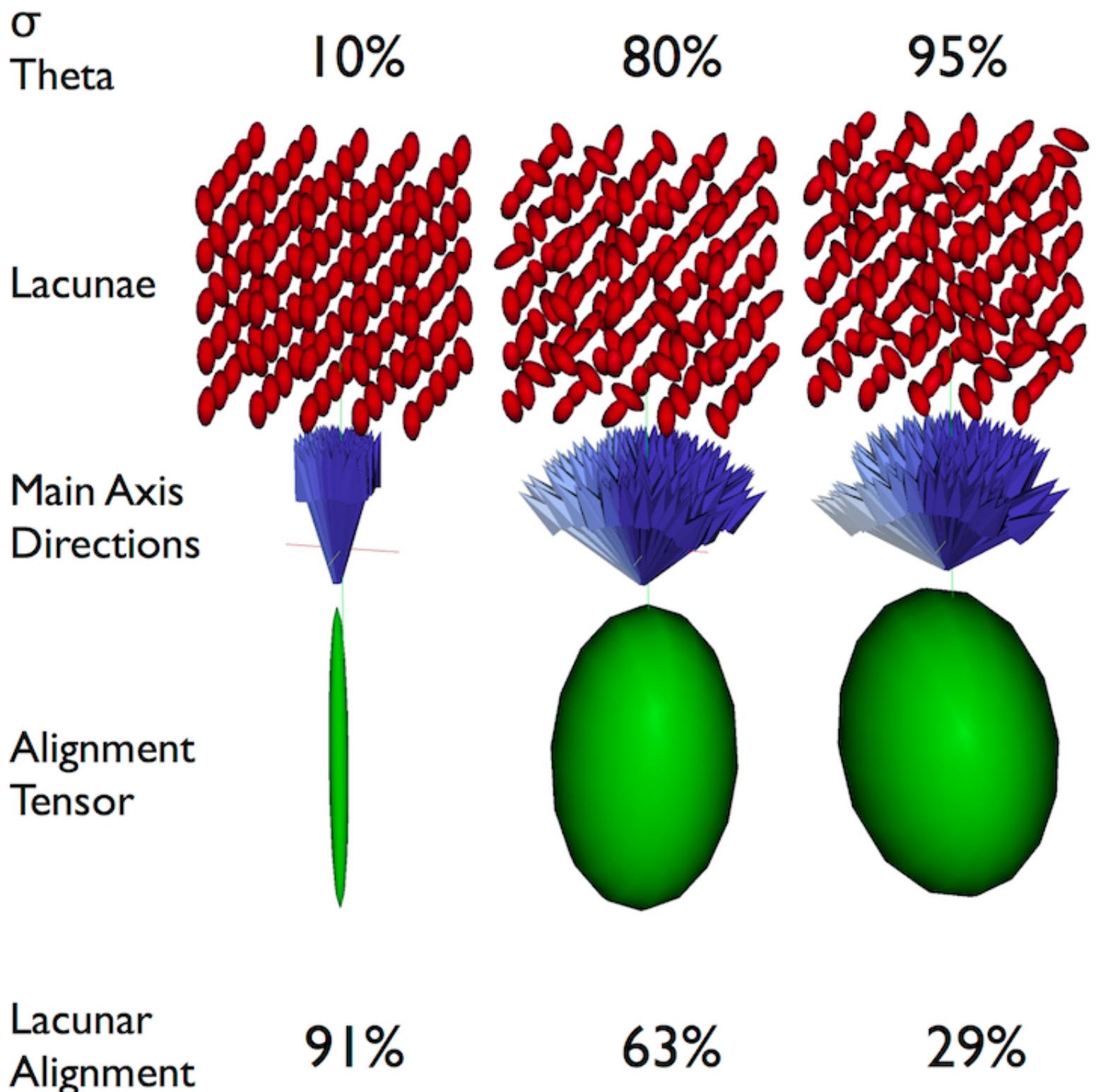
- The same as all of the tensors we have introduced
- Except we use the primary orientation instead of *voxel positions* (shape) or *edges* (distribution)

- Similar to distribution it is calculated for a volume ( $\mathcal{V}$ ) or region and is meaningless for a single object

$$\vec{v}_i = \vec{\Lambda}_1(i)$$

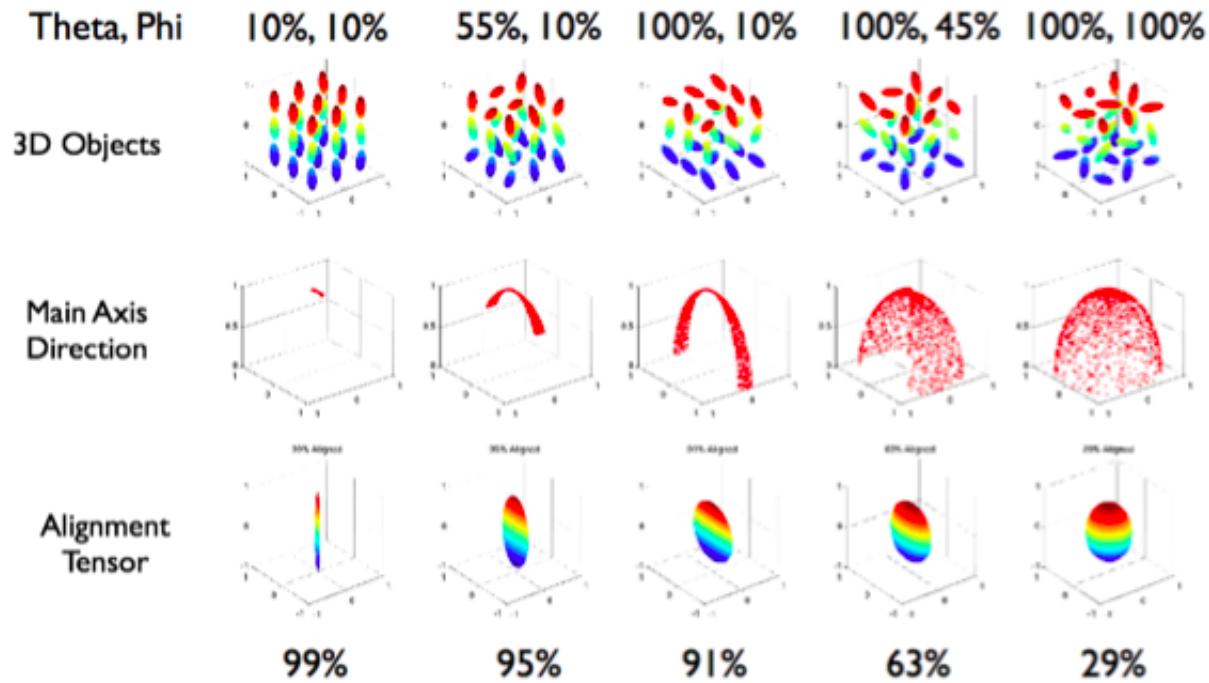
$$COV = \frac{1}{N} \sum_{\forall COM(i) \in \mathcal{V}} \begin{bmatrix} \vec{v}_x \vec{v}_x & \vec{v}_x \vec{v}_y & \vec{v}_x \vec{v}_z \\ \vec{v}_y \vec{v}_x & \vec{v}_y \vec{v}_y & \vec{v}_y \vec{v}_z \\ \vec{v}_z \vec{v}_x & \vec{v}_z \vec{v}_y & \vec{v}_z \vec{v}_z \end{bmatrix}$$


---



## Alignment Anisotropy

Anisotropy for alignment can be summarized as degree of alignment since very anisotropic distributions mean the objects are aligned well in the same direction while an isotropic distribution means the orientations are random. Oblateness can also be defined but is normally not particularly useful.



# Other Approaches

type:alert

## K-Means

K-Means can also be used to classify the point-space itself after shape analysis. It is even better suited than for images because while most images are only 2 or 3D the shape vector space can be 50-60 dimensional and is inherently much more difficult to visualize.

## 2 Point Correlation Functions

For a wider class of analysis of spatial distribution, there exist a class of functions called *N-point Correlation Functions*

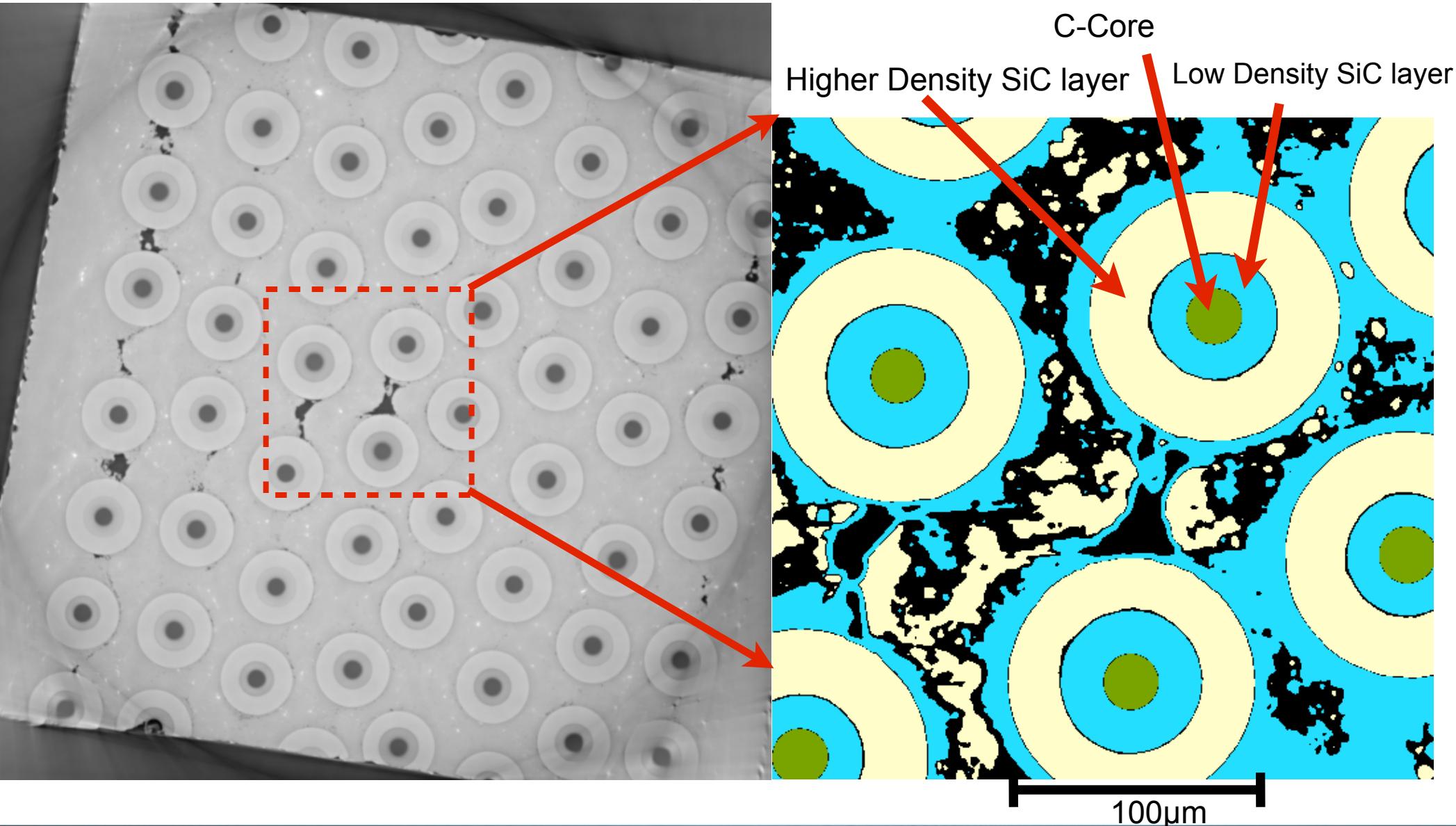
- given a point of type *A* at  $\vec{x}$
- what is the probability of *B* at  $\vec{x} + \vec{r}$
- How is it useful
  - A simple analysis can be used to see the spacing of objects (peaks in the  $F(\vec{r})$  function, *A* and *B* are the same phase).

- Where are cells located in reference to canals (peaks in  $F(\vec{r})$  function,  $A$  is a point inside a canal  $B$  is the cells)

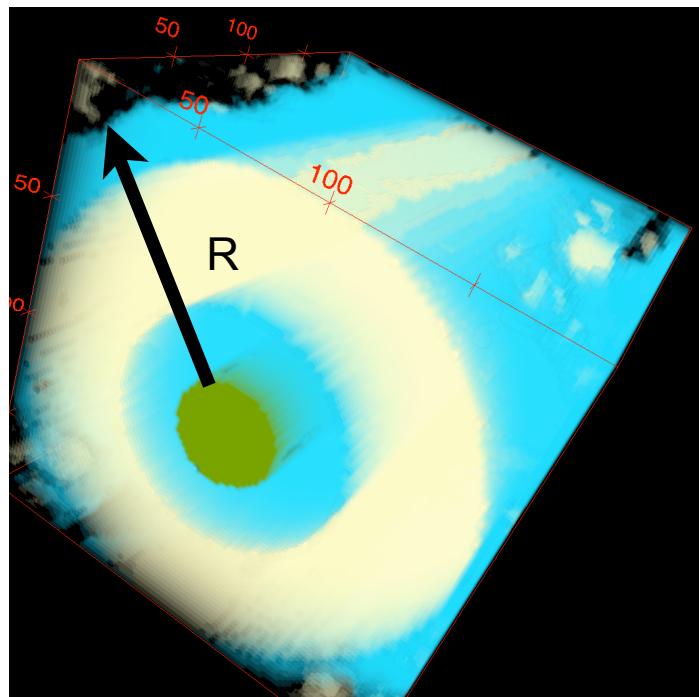
# SiC/C composite fibers in an Al Matrix

Collaborators - TOMCAT : S. Irvine, Institute for BioNanoTechnology in Medicine, Northwestern, USA :  
S. Stock

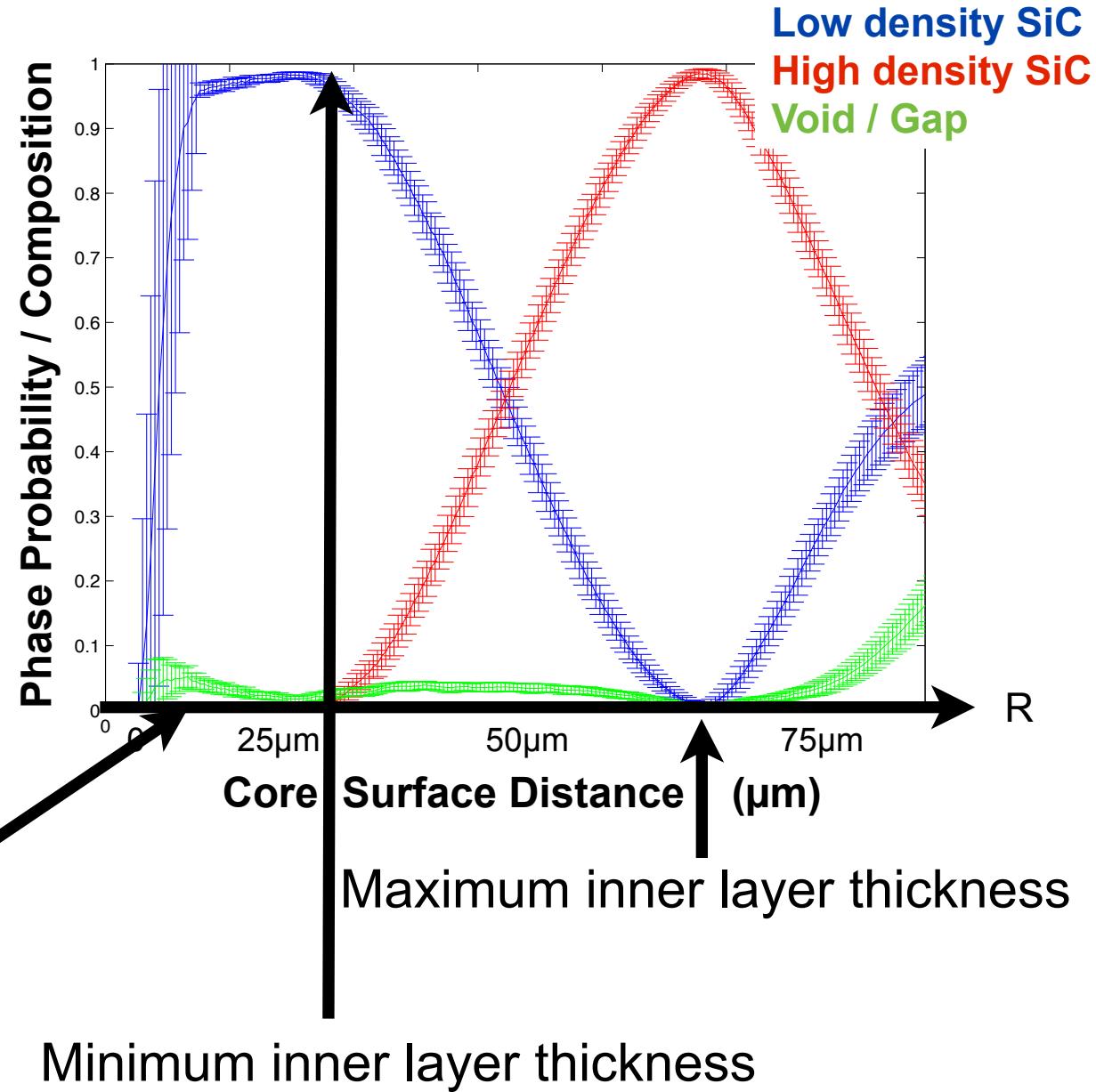
- Quantitative assessment of mass produced materials and fibers



# Quantifying layer separation and thickness



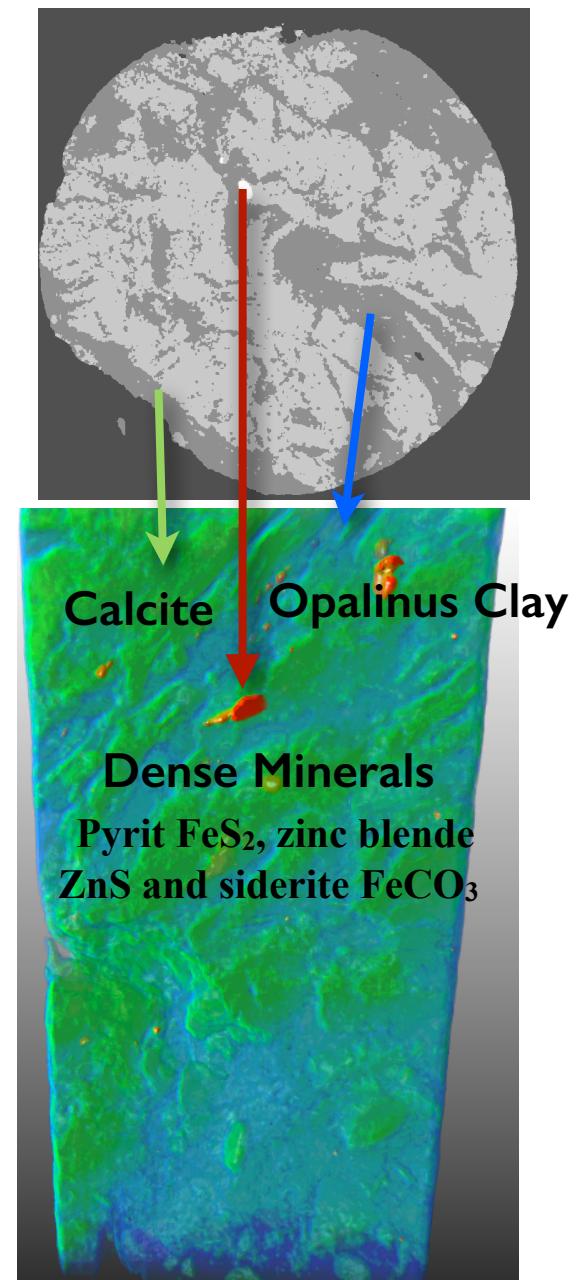
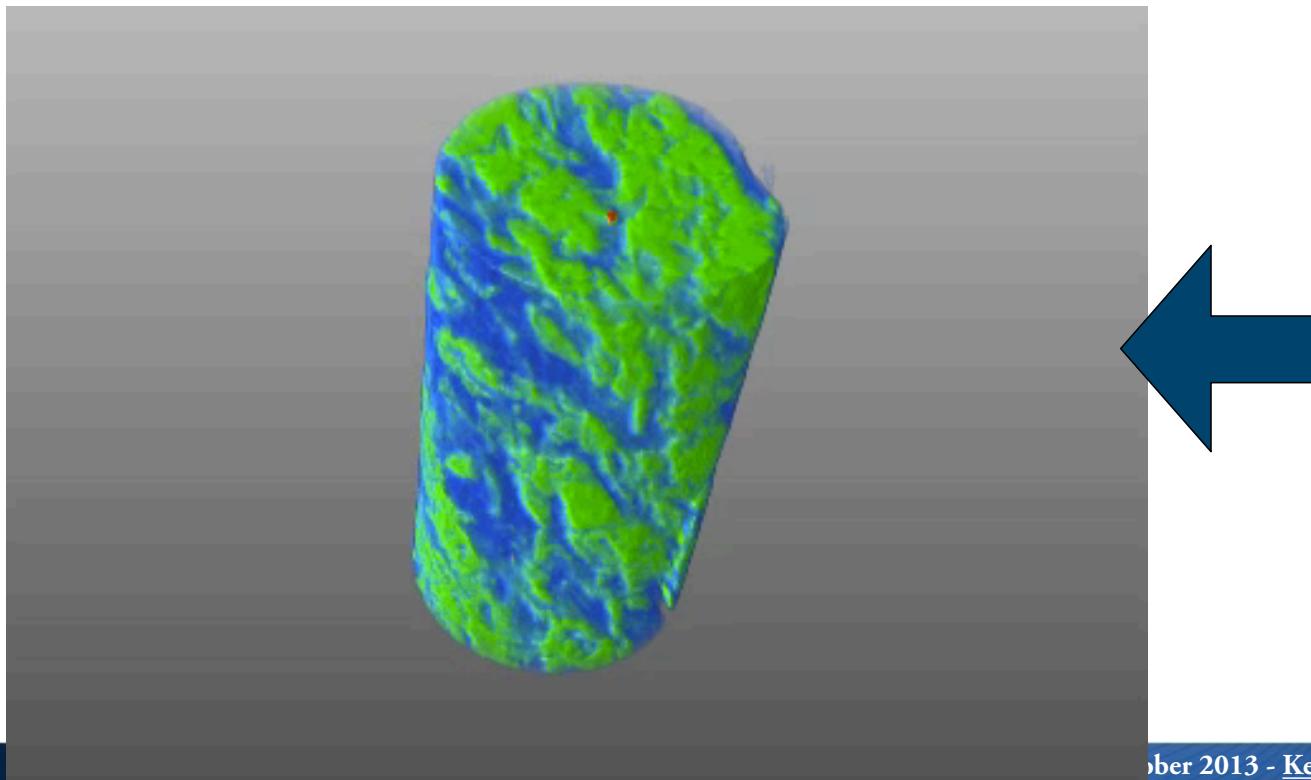
Up to 10% surface detachment between core and inner layer



# Radionuclide Diffusion in Compact Clays

Collaborators - TOMCAT : F. Marone, Laboratory for Waste Management, PSI : A. Jakob, S. Churakov

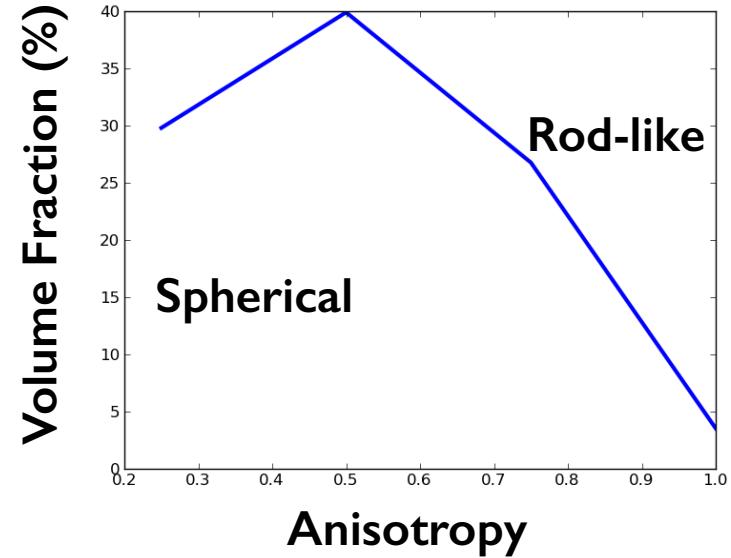
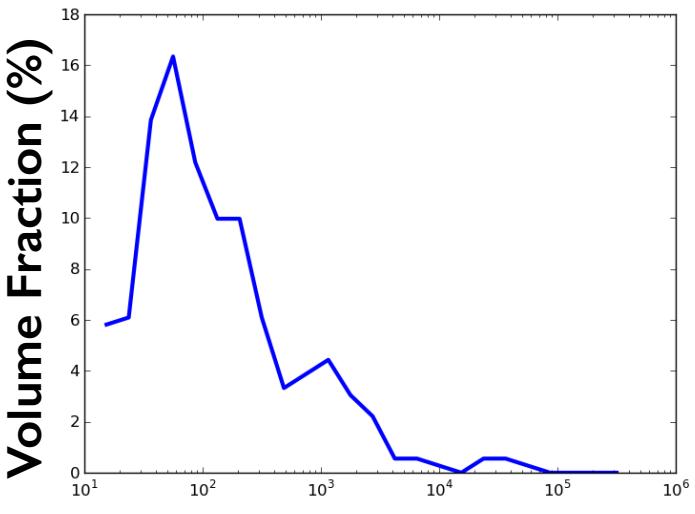
- To safely store radioactive waste migration through compact rock must be studied
- Microscopic material and structure of rock determines transport properties for tracers
- Quantify variation in structure distribution  
--> many samples needed



# Radionuclide Diffusion in Compact Clays

Collaborators - TOMCAT : F. Marone, Laboratory for Waste Management, PSI : A. Jakob, S. Churakov

- Collections of dense materials are slightly anisotropic but still form clumps



## Spatial Relationships

- Dense minerals are preferentially embedded in calcite over clay indicative of a high pH environment at time of formation

