

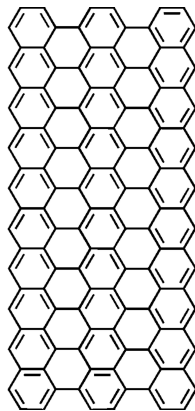
Big Data Algorithms in Nanoelectronics

1. Introduction to Nanoelectronics
2. The Design of Clusterplot
3. Test of Clusterplot with Generic Data
4. Back to Physics
5. Conclusion

Electric Transport in Graphene Nanoribbons

Graphene:

- Single layer carbon structure
- High electron mobility ($>15000 \text{ cm}^2\text{V}^{-1}\text{s}^{-1}$)
- Zero-gap semiconductor



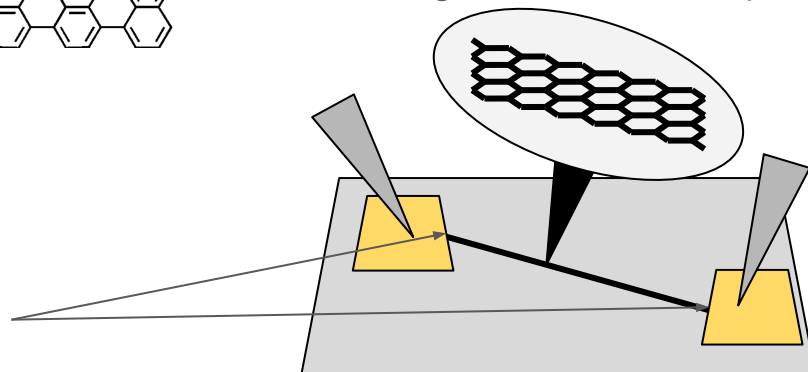
Graphene Nanoribbons:

- Stripes of graphene
- Small ($<50 \text{ nm}$ wide)
- Metallic/semiconductive depending on the edge formation

My Project:

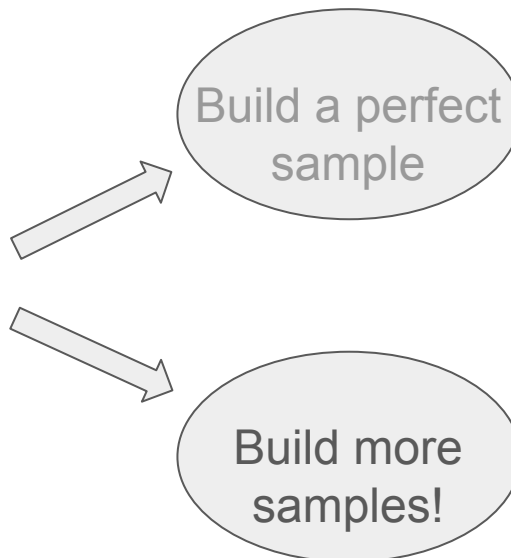
- GNRs assembled directly on surface of hexagonal Boron Nitride
- Current I as a function of applied Voltage V is measured between gold contacts on top

6-7 μm



Measurement Problems in Nanoelectronics ...

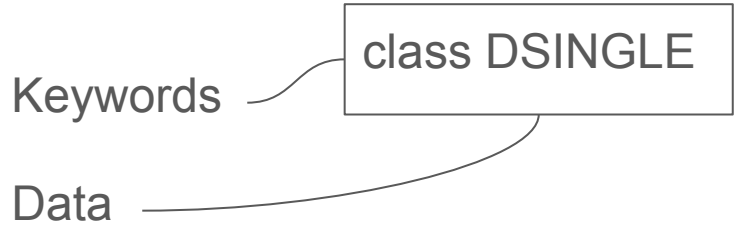
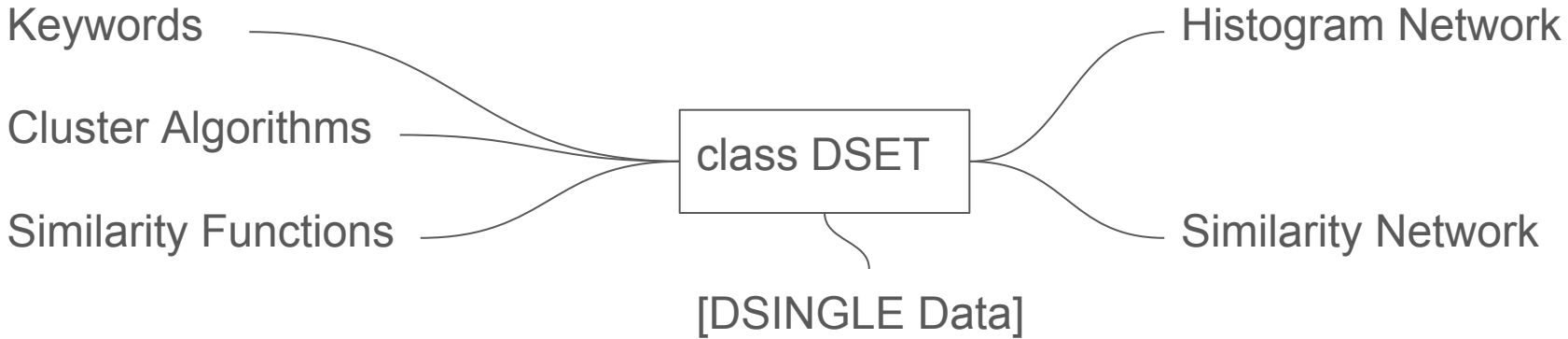
- Contamination
- Surface irregularities
- Structure breaks after measurements
- Errors and irregularities in building procedure



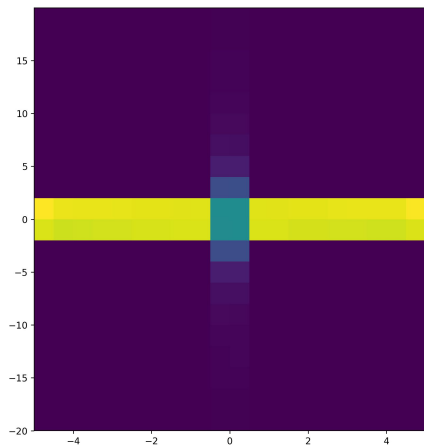
... Led to the Idea for Clusterplot!

- Provide a framework to handle big amounts of measurement files (currently used for up to 4000)
- Find similar curves in a data set and group them together
- Filter the data set easily for specific attributes

The Design of Clusterplot

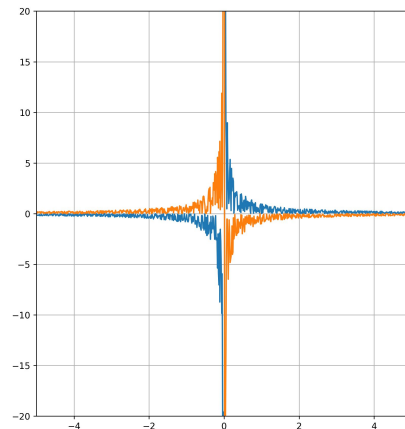


Histogram vs. Histogram Network

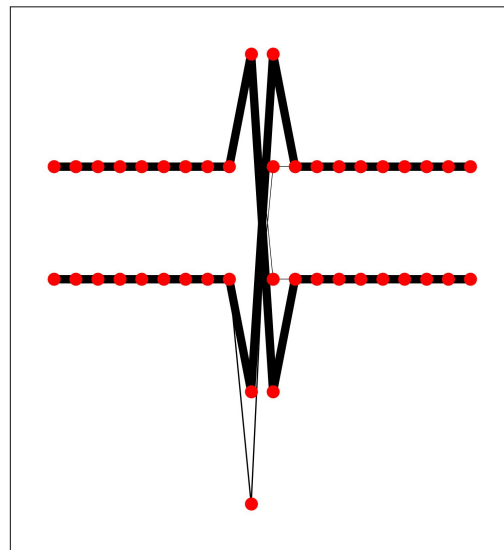


- Feature Space: $n \times m$
- Good for representing single features

- 400 hyperbolas with random noise
- 200 have a positive coefficient
- 200 have a negative coefficient

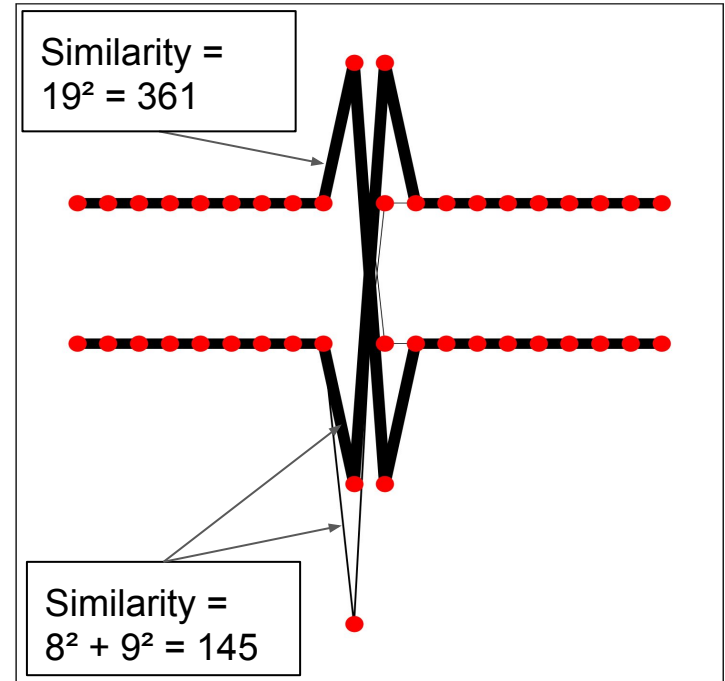


- Feature space: $n^2 \times (m-1)$
- Good for representing paths
- Stores also information about first derivative, better for overlapping curves



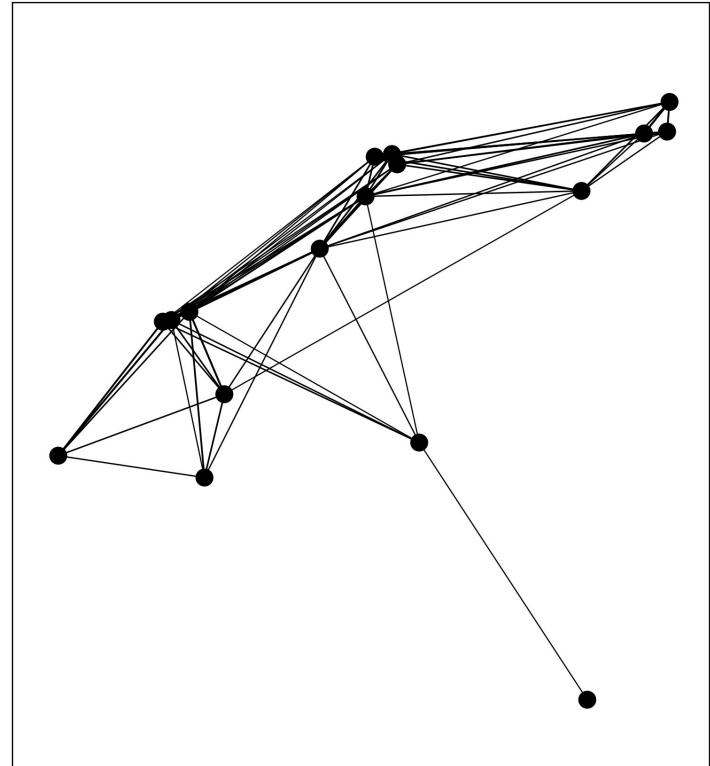
The Definition of Similarity - Squared Rows

1. Compare the edges of two single curves and find matching ones
2. Count the length of linked edges and square them
3. The similarity is the sum of the squares



The Similarity Network

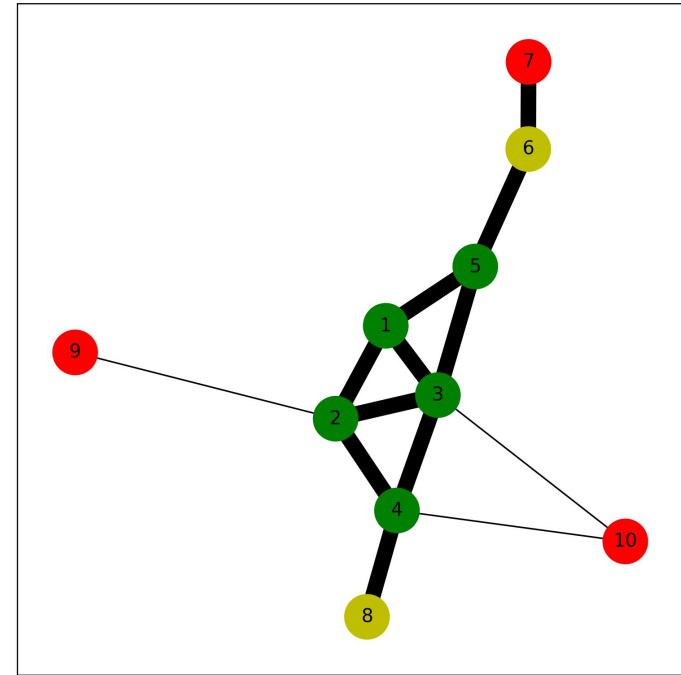
- Points are adequate to curves/files
- Edges are adequate to their similarity
- The layout is calculated by a spring algorithm, points are repelling, the stronger the edges are, the stronger is the attraction



The Similarity Network - DBSCAN Cluster Algorithm

“Density-based spatial clustering of applications with noise”

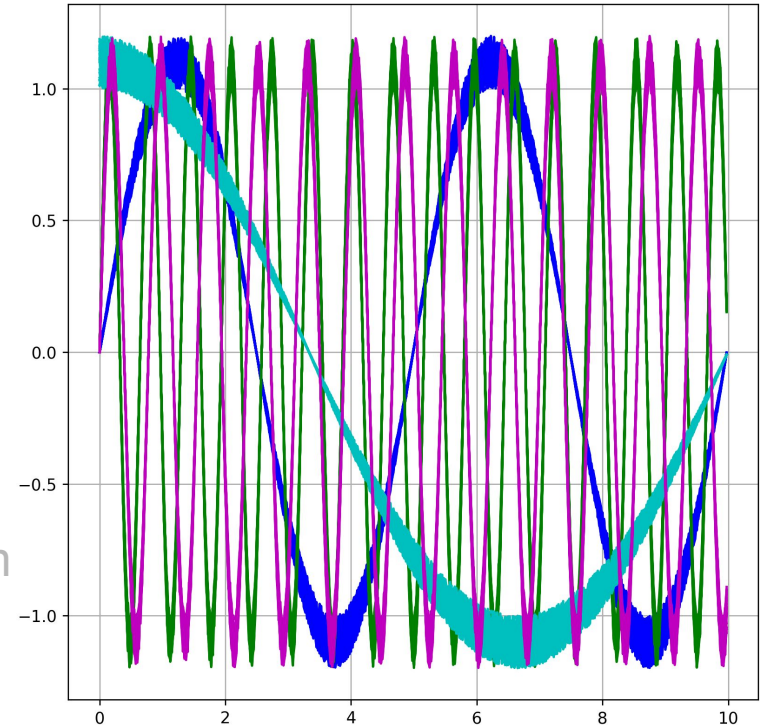
- Points are reachable from another point, when their similarity is higher than “min_sim”
- Points with at least “min_pts” points within reach are considered “Core Points”
- All core points and all points reachable from a core point are cluster members



Testing Clusterplot with Sine and Cosine Curves

Generic test data set:

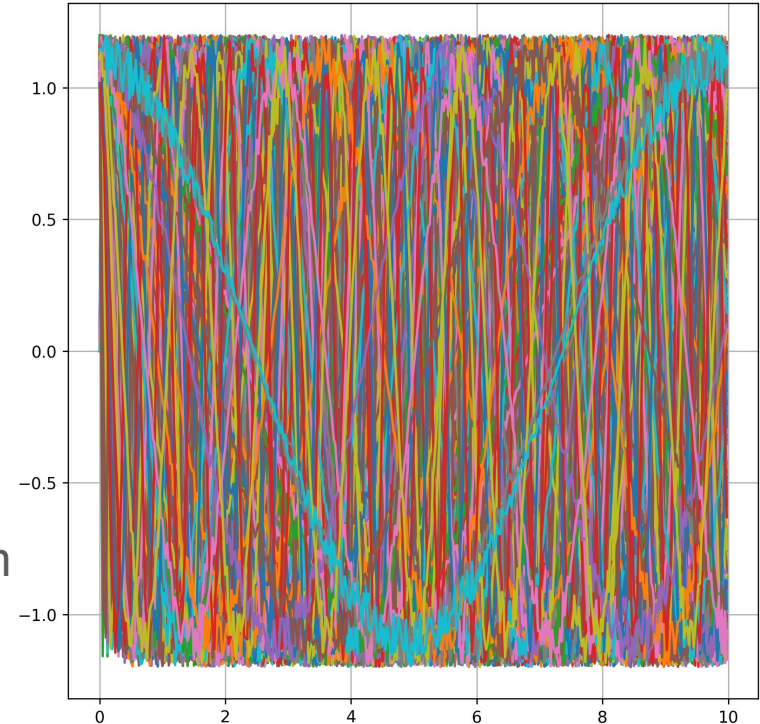
- (3x) 20 sine curves with the same random period and 20% noise
- (1x) 20 cosine curves with the same random period and 20% noise
- 300 sine curves with different random periods and 20% noise
- 300 cosine curves with different random periods and 20% noise



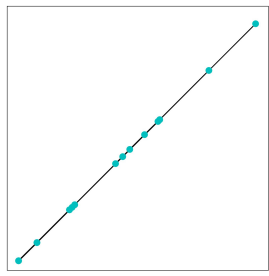
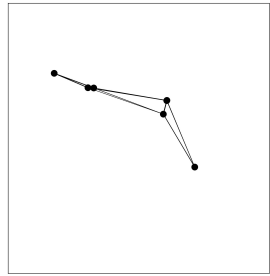
Testing Clusterplot with Sine and Cosine Curves

Generic test data set:

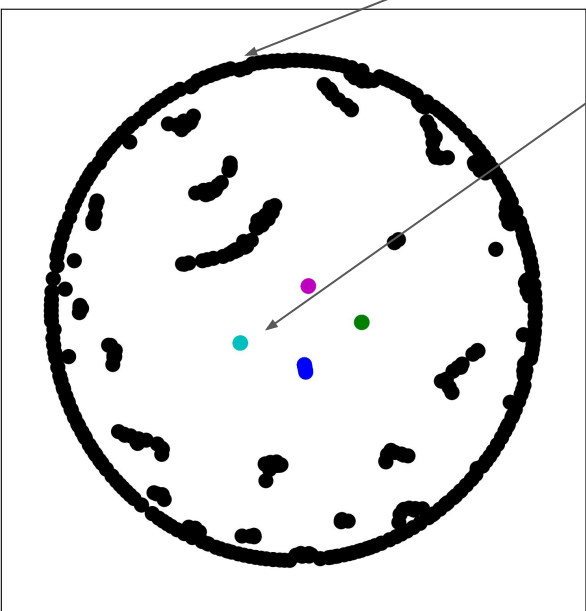
- (3x) 20 sine curves with the same random period and 20% noise
- (1x) 20 cosine curves with the same random period and 20% noise
- 300 sine curves with different random periods and 20% noise
- 300 cosine curves with different random periods and 20% noise



Simnet

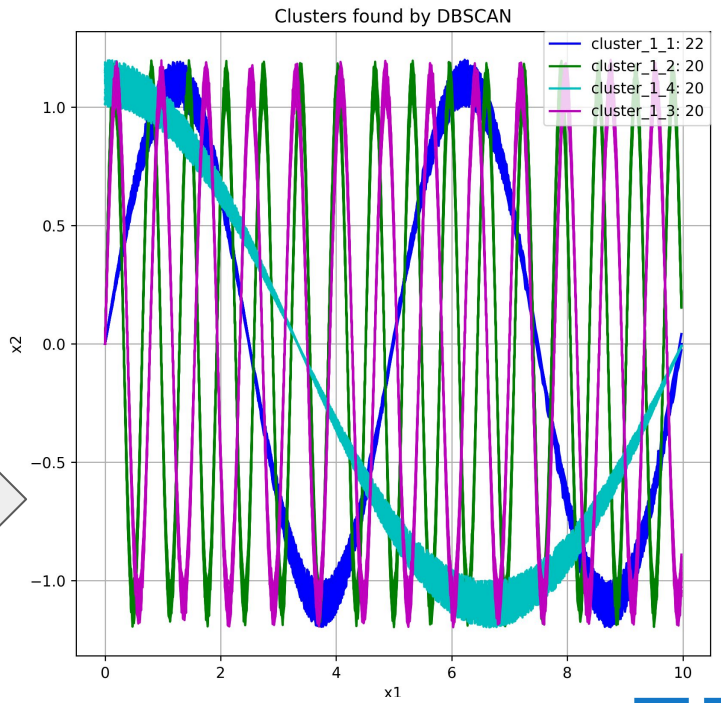


Cluster Plot

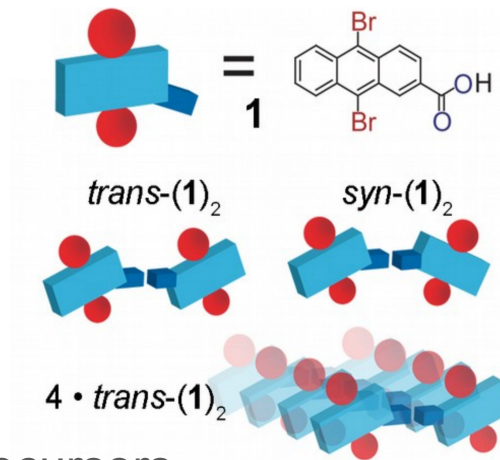


- Parameters used:
- 20 x 20 bins
 - min_sim = 50
 - min_pts = 15

Cluster Algorithm:



Back to Physics ...

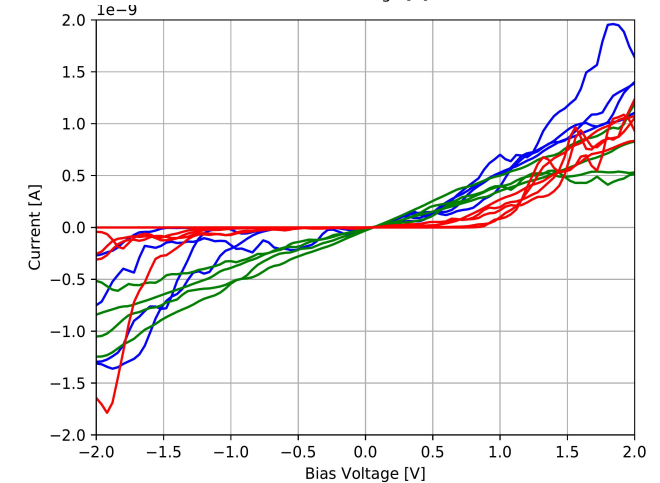
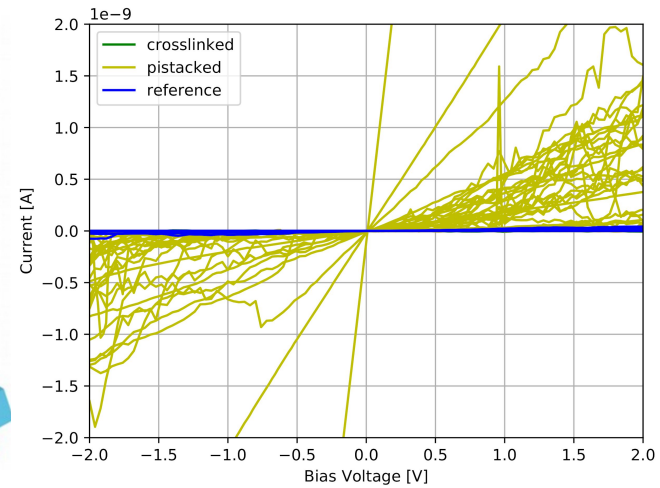


The production steps were:

1. Building the sample with precursors
2. measuring the transport
3. crosslink the precursors
4. measuring the transport

But:

- Transport behavior of the pi-stacked precursors could be shown!



Conclusion

1. Clusterplot works well and was also used in two other projects, and has been constantly developed since the end of my master's thesis and will get a graphical interface soon
2. The electric transport in Graphene Nanoribbons could not be measured
3. On the other hand, electric transport in its precursors could be shown

Thanks for Your Attention!