Voronoi and Delaunay Tessellations for Bayesian Wombling on LHC Data

Presenter: Alex Roman

University of Florida
Department of Physics

May 4th, 2020

K. Matchev, A. Roman, P. Shyamsundar

Detecting New Physics

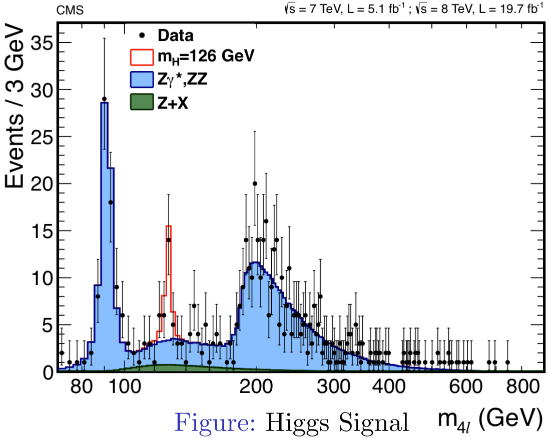


Figure: Higgs Signal

The process of detecting new physics can be viewed as the search for the overdensity of events in a suitable phase space.

Simulated Signals

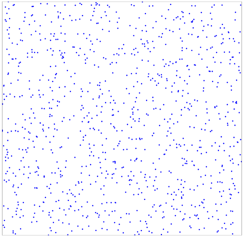


Figure: No Signal

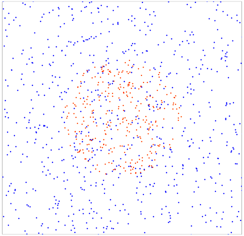


Figure: Circular Signal

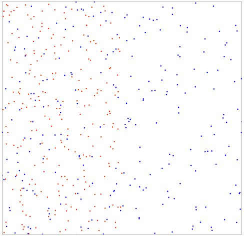


Figure: Step Signal

We generate two generic types of signals, shown in orange above.

Simulated Signals

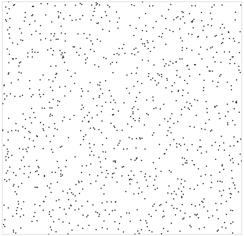


Figure: No Signal

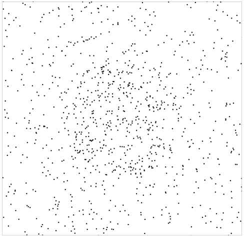


Figure: Circular Signal

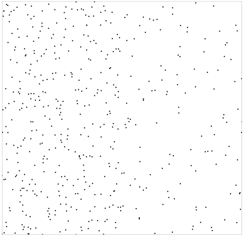


Figure: Step Signal

We discard the labels and all points appear black, as in real data.

Boundary Detection

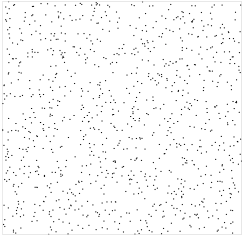


Figure: No Signal

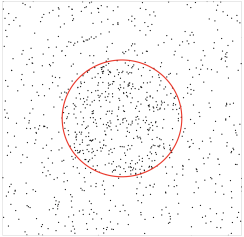


Figure: Circular Signal

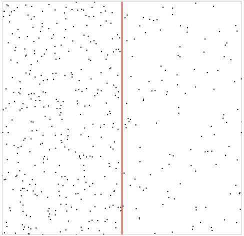
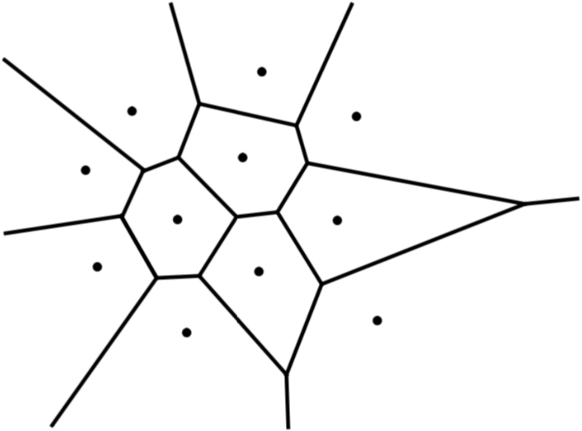


Figure: Step Signal

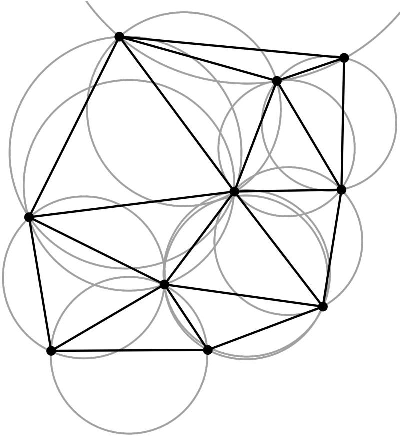
The goal is to reconstruct the true boundaries used to generated the data.

Voronoi Tessellations



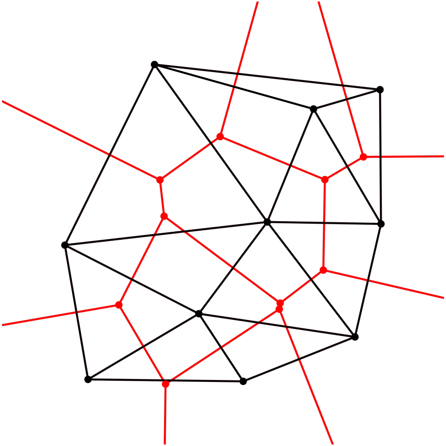
The Voronoi Tessellation assigns a polygon to each data point, whose edges bisect the lines between each point and it's neighbors.

Delaunay Triangulation



The Delaunay triangulation tiles the space with triangles. Three points share a triangle if their circumcircle contains no other points.

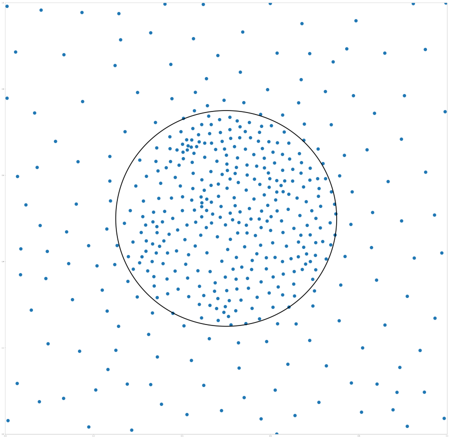
Delaunay Triangulation



It turns out that the two graphs formed by the Voronoi tessellation and the Delaunay triangulation are dual to each other.

Voronoi Tessellation on our Data

We consider some simulated data.

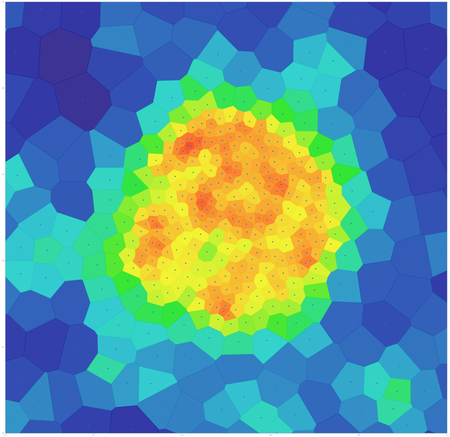


Voronoi Tessellation on our Data

We compute the Voronoi tessellation.
We estimate the functional density by the inverse area.

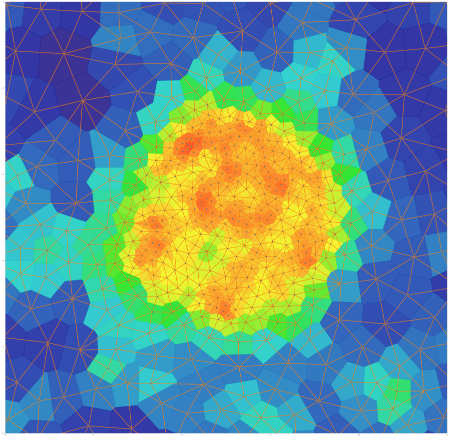
$$Value \sim \frac{1}{Area}$$

(The cells are colored by the value)



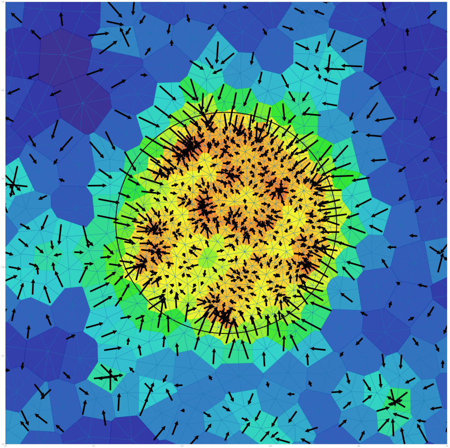
Voronoi Tessellation and Delaunay Triangulation

Superimposing the Delaunay triangulation we can see a natural way to associate a gradient. Since each triangle's vertices have associated values, we can simply fit a plane.



Voronoi Tessellation and Delaunay Triangulation

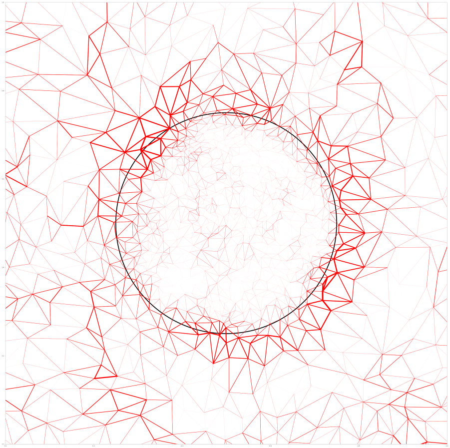
Superimposing the Delaunay triangulation we can see a natural way to associate a gradient. Since each triangle's vertices have associated values, we can simply fit a plane.



Dot Product Correlation

Note that the gradients on the true boundary are large and aligned.

We calculate the dot product between neighboring gradients, and impose a cut on the value.

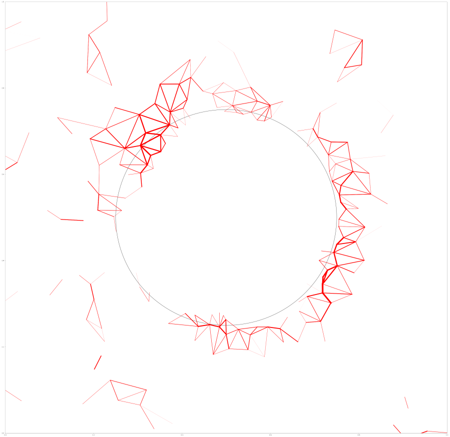


Dot Product Correlation (Truncated)

Note that the gradients on the true boundary are large and aligned.

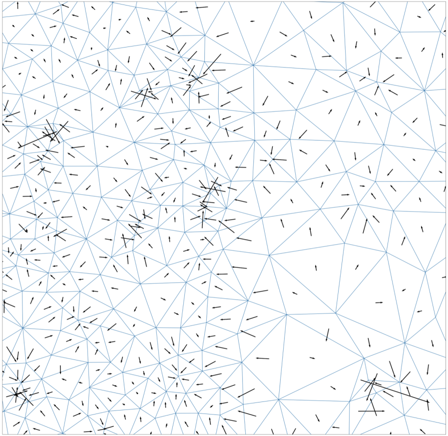
We calculate the dot product between neighboring gradients, and impose a cut on the value.

We can see the boundary is partially reconstructed.



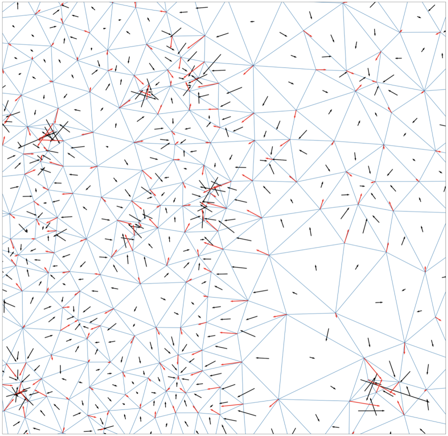
Gradient Averaging

- The **Black** vectors are the gradients of the triangles



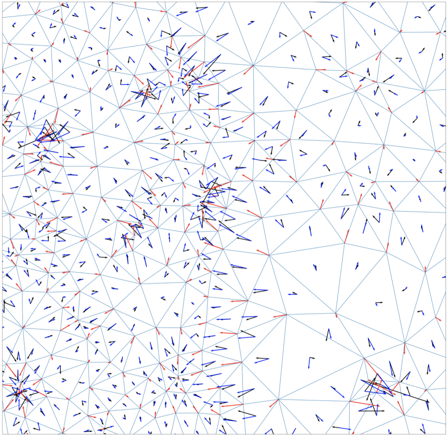
Gradient Averaging

- The **Black** vectors are the gradients of the triangles
- The **Red** vectors are the averages of the gradients neighboring each Delaunay vertex.



Gradient Averaging

- The **Black** vectors are the gradients of the triangles
- The **Red** vectors are the averages of the gradients neighboring each Delaunay vertex.
- The **Blue** vectors are the averages of values on the vertices of each triangle.



Wombling

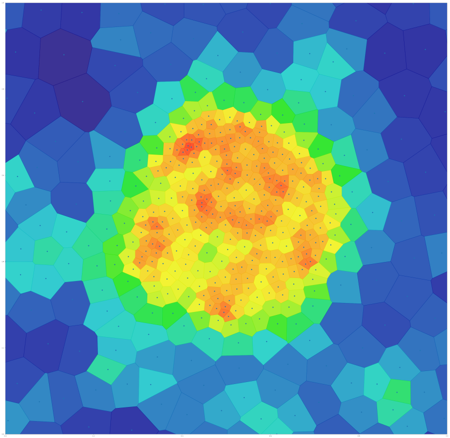
What is Wombling?

Define Wombling:

In statistics, **Wombling** is any of a number of techniques used for identifying zones of rapid change, typically in some quantity as it varies across some geographical or Euclidean space. It is named for statistician William H. Womble.

Wombling

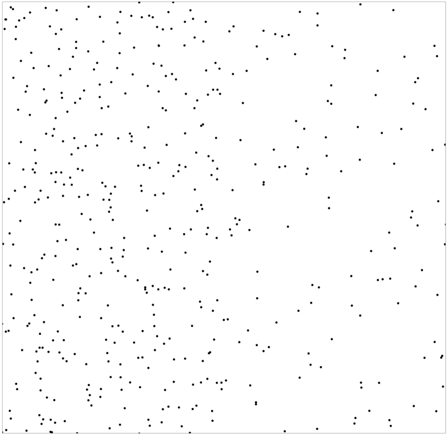
We are looking for regions of rapid change in the functional value.



Wombling is Finding Cliff Faces

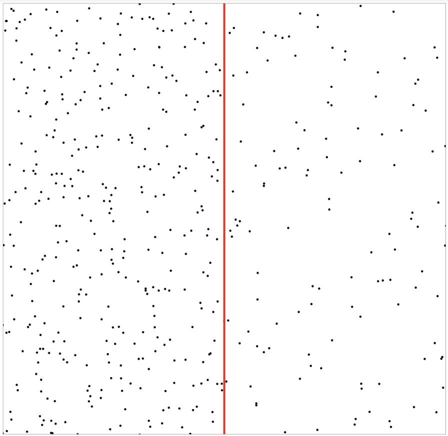
Finding the boundary

Consider a step distribution.

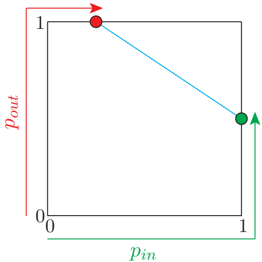


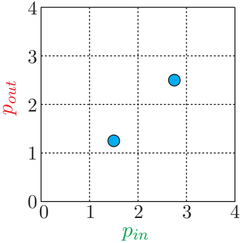
Finding the boundary

We confine ourselves to searching only for straight line boundaries.



Parameterization of Lines

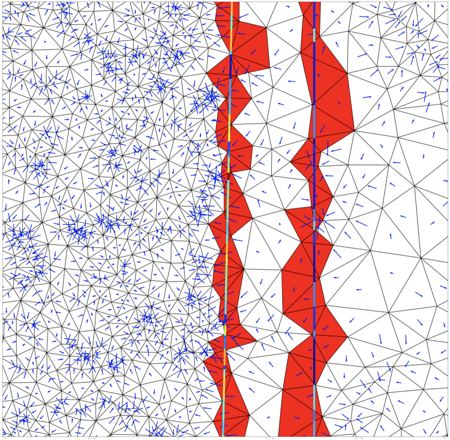




We parameterize all possible lines intersecting the unit square in terms of coordinates as measured along the perimeter of the entry point p_{in} and the exit point p_{out}.

Straight Line Wombling

Each line is evaluated by the total flux of averaged gradients. The segments are colored by their flux density. We can see that the left (true) boundary has a higher flux than the right one.



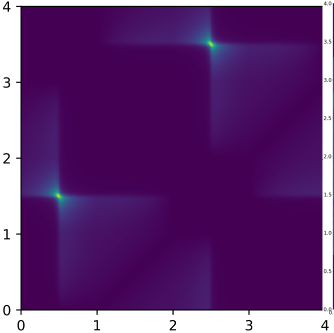


Figure: Theoretical

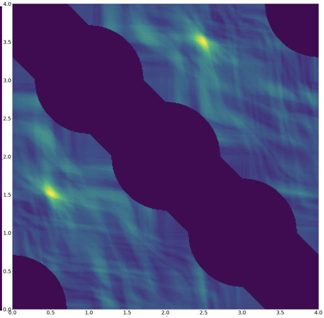
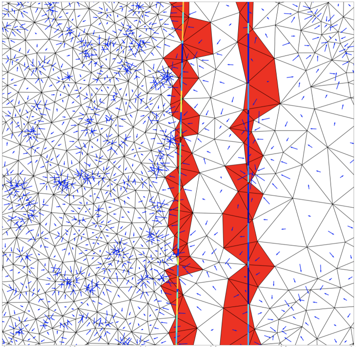


Figure: Actual

Heat-map of flux on the space of all lines.
The winning lines correspond to the true boundary.

Results



We assign each segment of each line a value:

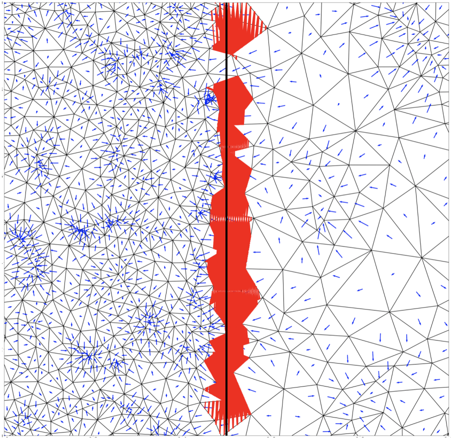
$$\text{SegmentValue} = \frac{\text{SegmentFlux}}{\text{SegmentLength}} \cdot \left(\frac{\text{TotalFlux}}{\text{TotalLegth}}\right)^p$$

p is a free parameter.

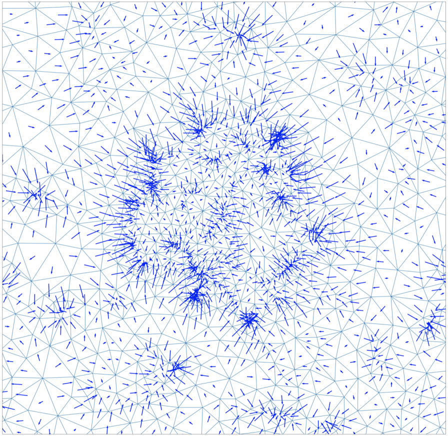
Results

Showing the top 1% of segments, we nicely reconstruct the boundary.

$(p = 1)$



Circular Signal



This method can be used to detect non-linear boundaries.

Straight Line Wombling

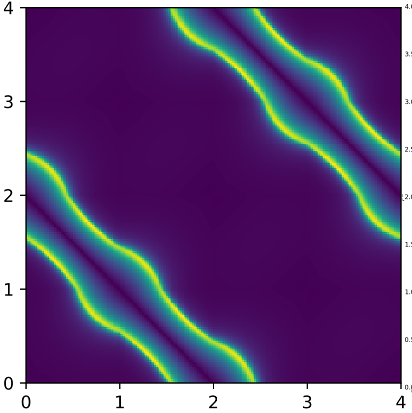


Figure: Theoretical

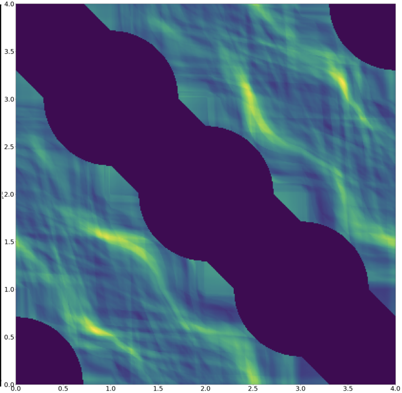
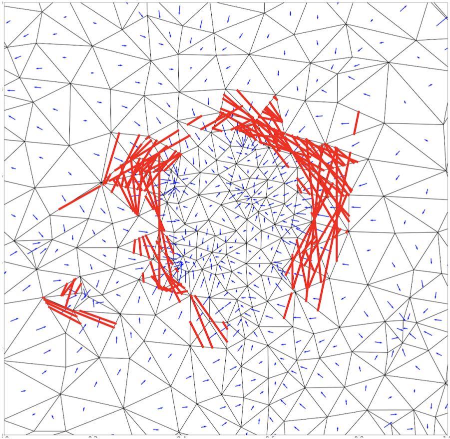


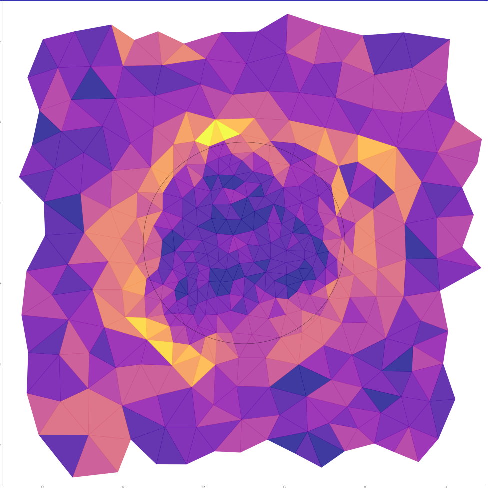
Figure: Actual

Straight Line Wombling

We are able to reconstruct this curved boundary.



The End



Questions?