ALMA TDA Project Meeting

1/2/2017

Outline

- Introductions and project overview
- Discussion of Radio Astronomy Data Analysis Challenges

  <Break>
- Technical Description of our work

  <Lunch>
- Live Demo using datasets provided

  <Break>
- Discussion
  - Improvements to software, additional features, etc.
  - Scientific challenges/opportunities using this approach
  - Spreading the word about the technique
  - Revisit data challenges facing Radio Astronomers

Personnel Status

- USF
  - Paul Rosen
  - Junyi Tu
- NRAO
  - Jeff Kern
- San Jose State
  - Betsy Mills
- Utah
  - Bei Wang
  - Chris Johnson
  - Ayla Khan
  - Dan White
- Collaborators

Project Overview

Project history—ALMA Development Study

Project objectives—Investigate applications of Topological Data Analysis (TDA) to ALMA data

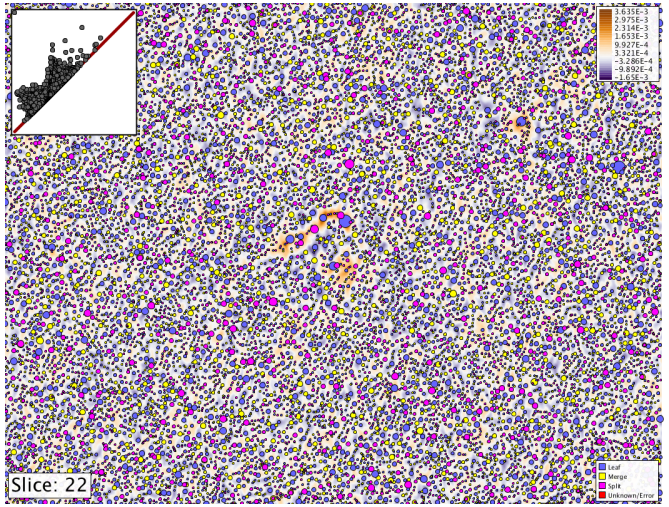
TOPOLOGICAL DATA ANALYSIS (TDA)

study of approaches to EXTRACT structure from NOISY or COMPLEX data and REPRESENT that data in an actionable form

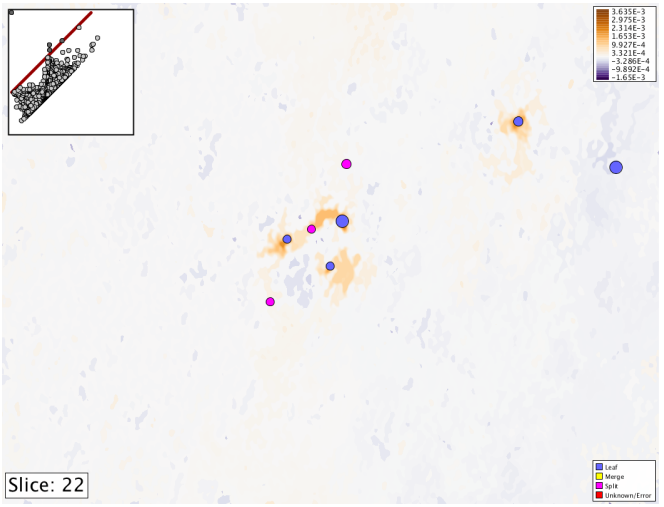
TDA represents a DIVERSE toolbox capable of addressing analysis needs in many contexts

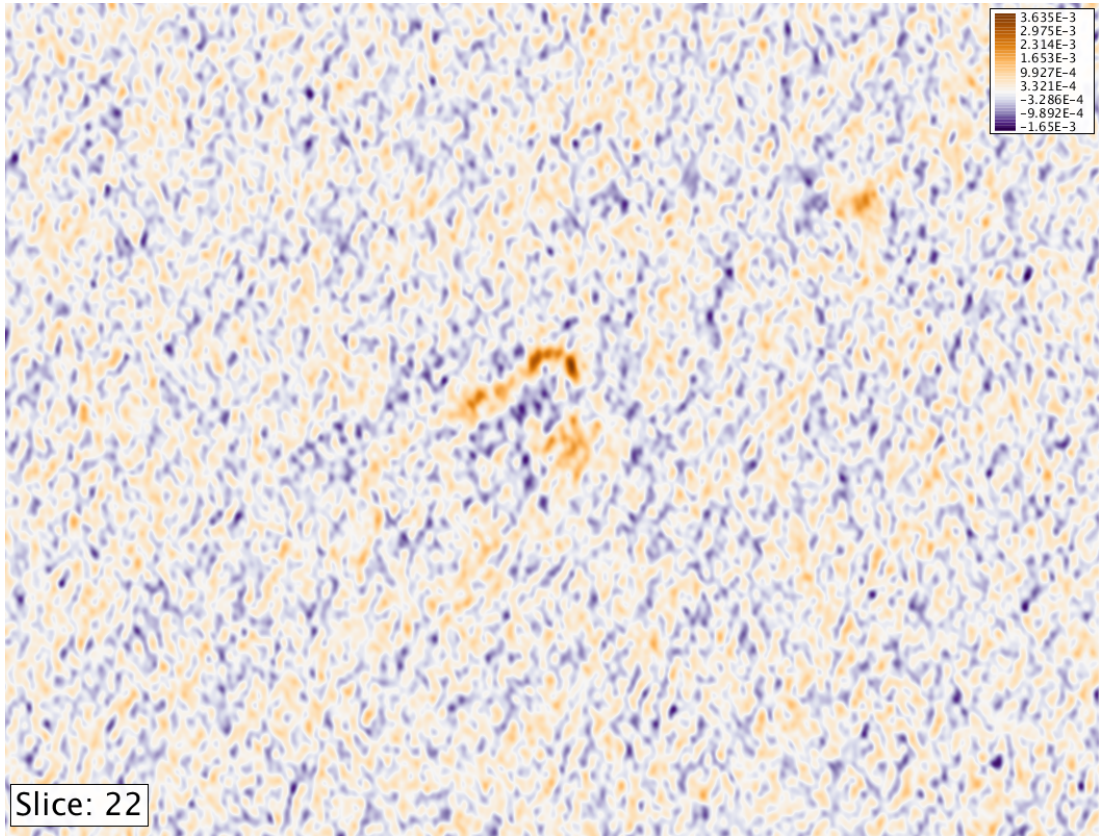
Our current work uses the CONTOUR TREE for data cube denoising

original



simplified





Discussion of Radio Astronomy Data Analysis Challenges

BREAK

Technical Description of our work

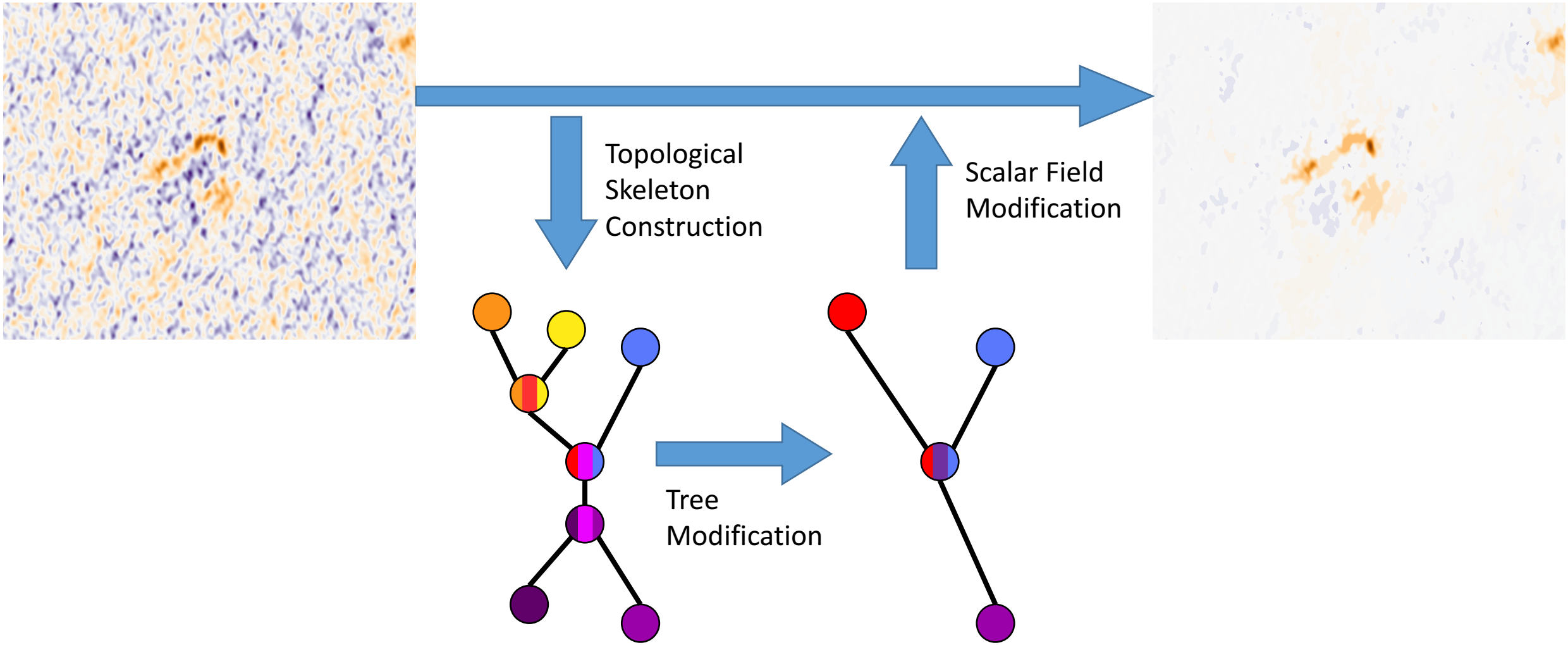
TOPOLOGICAL DATA ANALYSIS (TDA)

study of approaches to EXTRACT structure from NOISY or COMPLEX data and REPRESENT that data in an actionable form

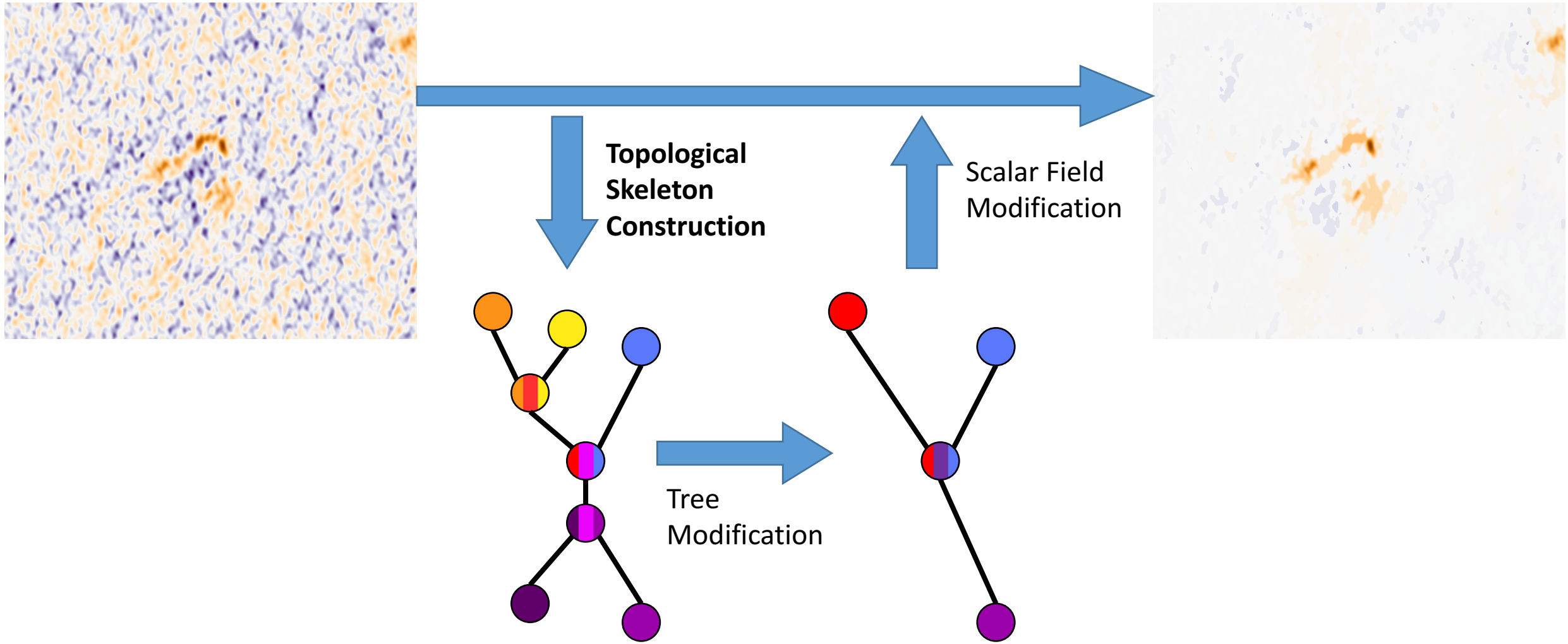
PERSISTENT HOMOLOGY (PH)

a method for computing topological FEATURES of a space at DIFFERENT spatial RESOLUTIONS simultaneously

Usage Scenario

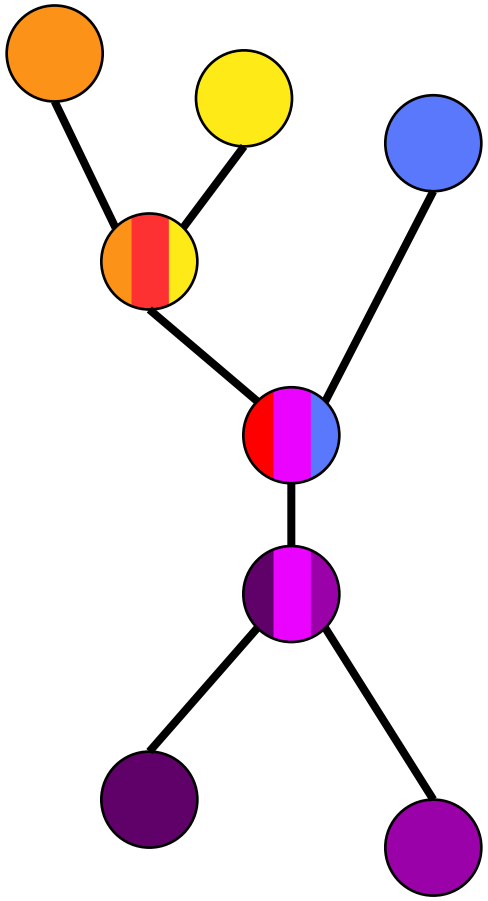


Usage Scenario

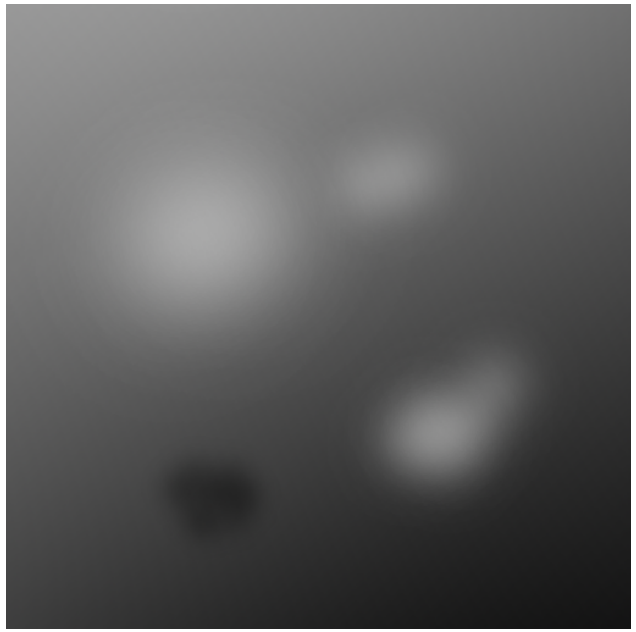


Topological Skeleton Construction

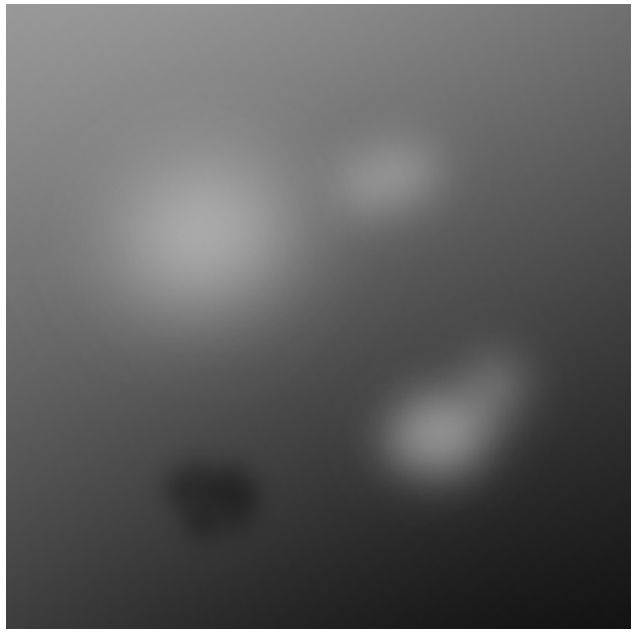
The Contour tree

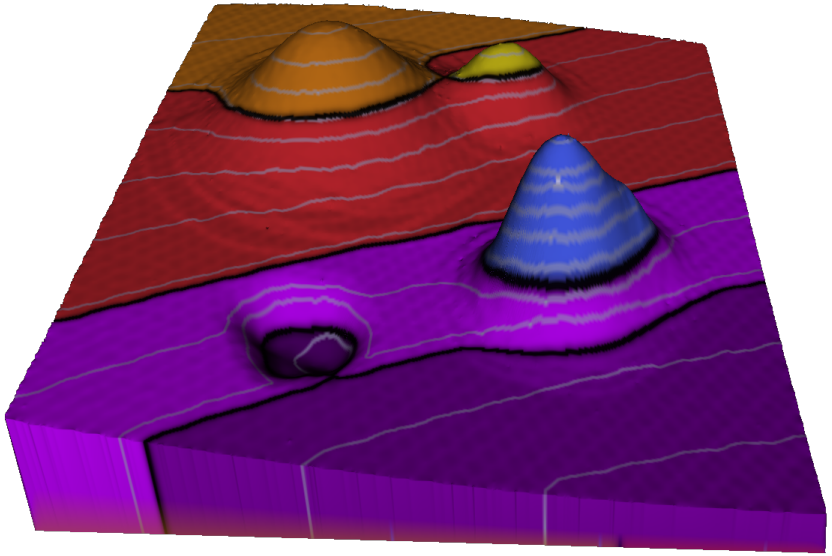


CONTOUR TREES

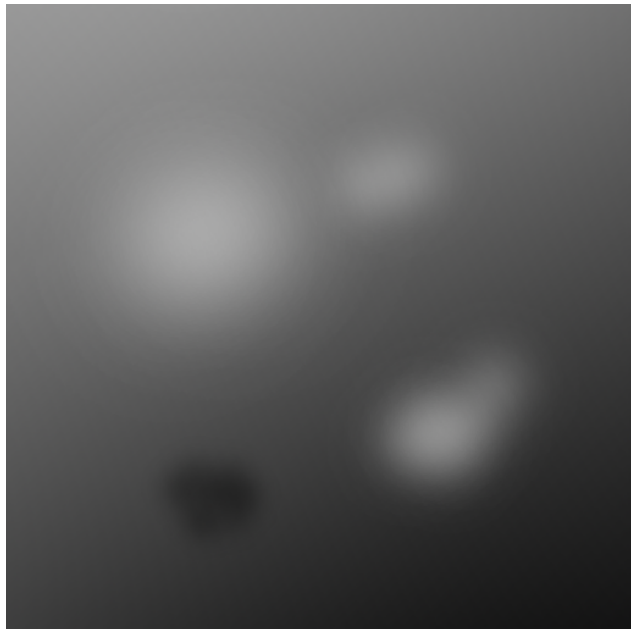


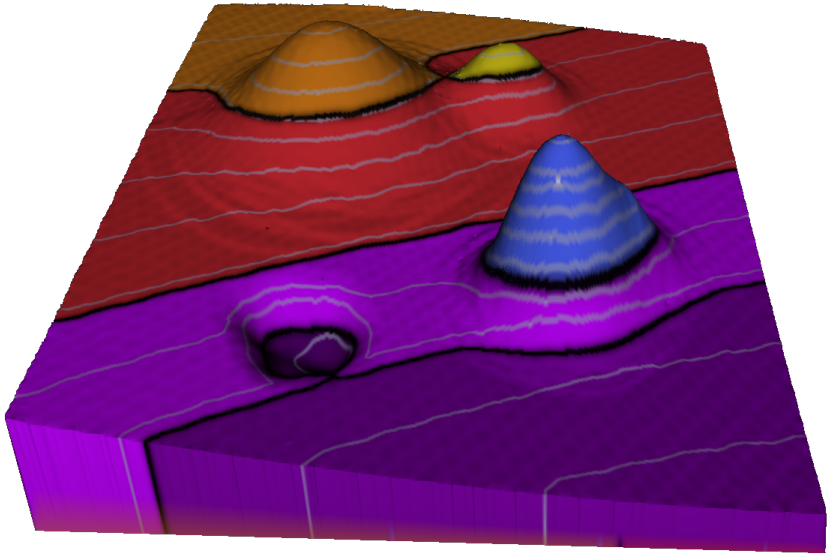
Contour trees

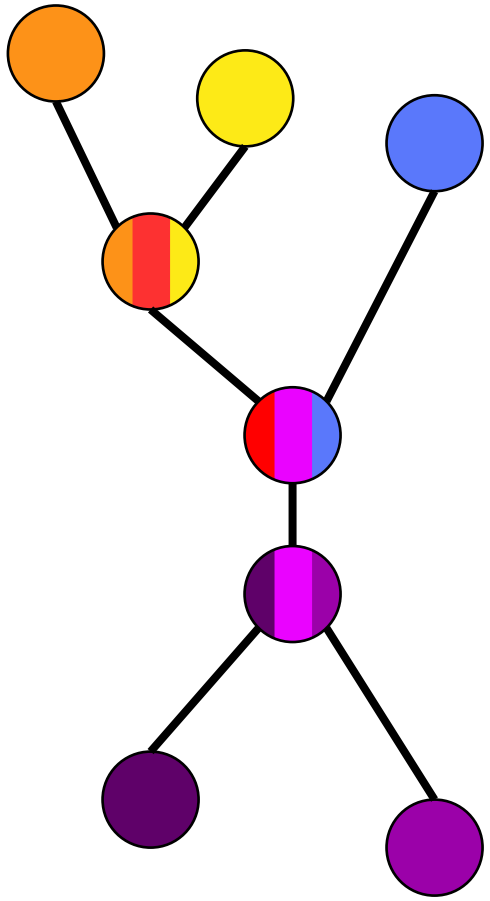




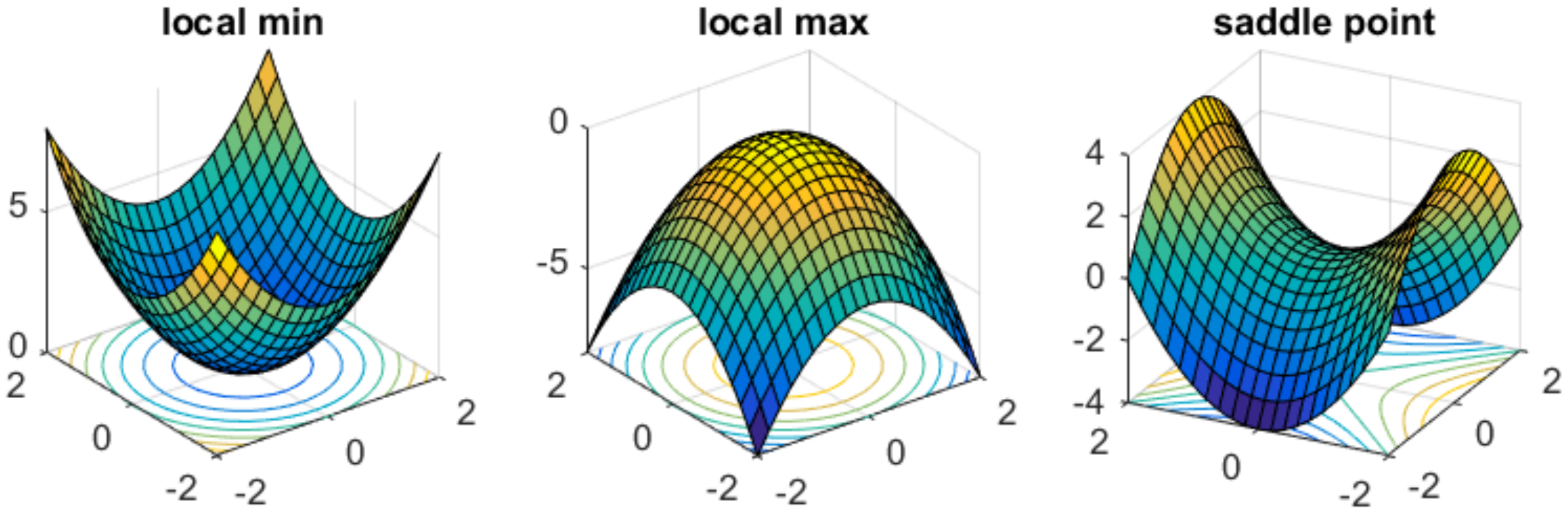
Contour trees



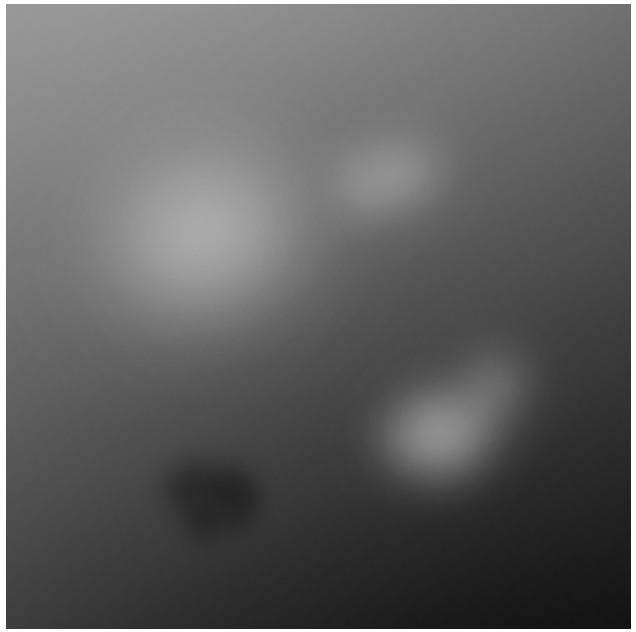


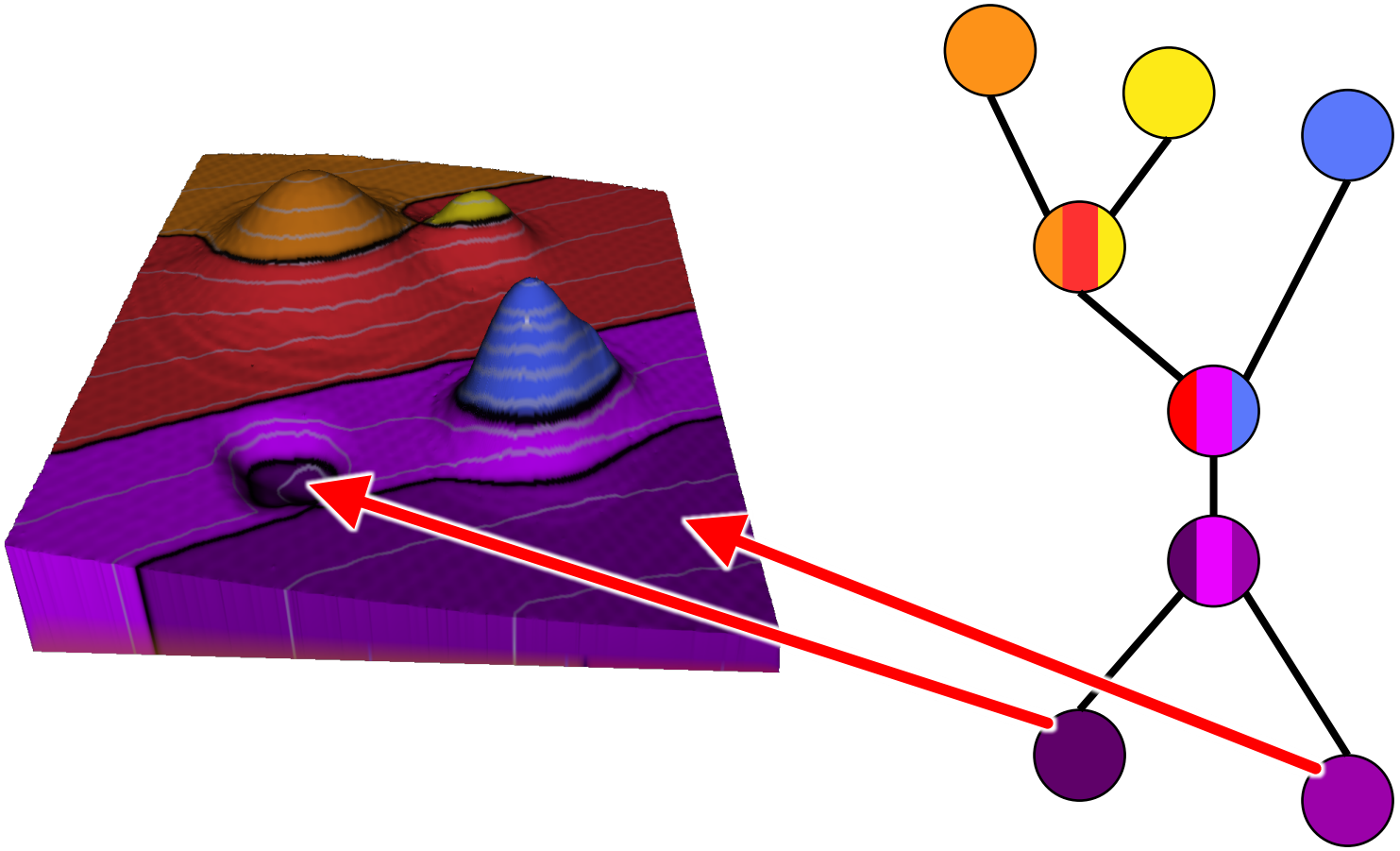


Critical Point Types

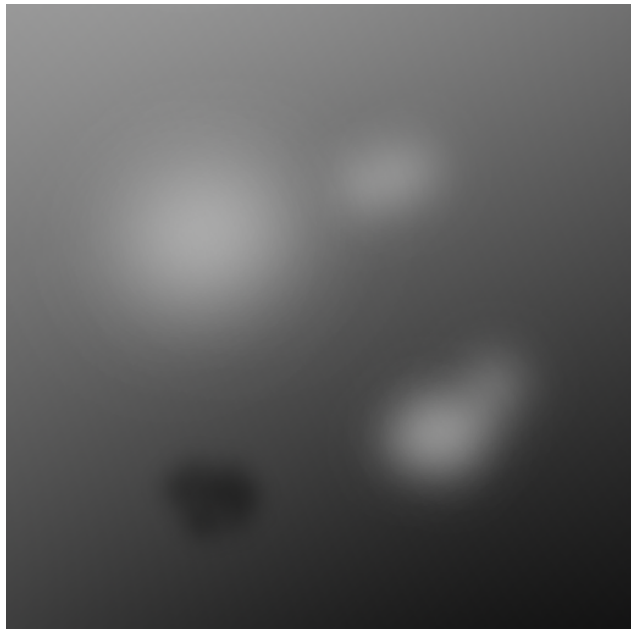


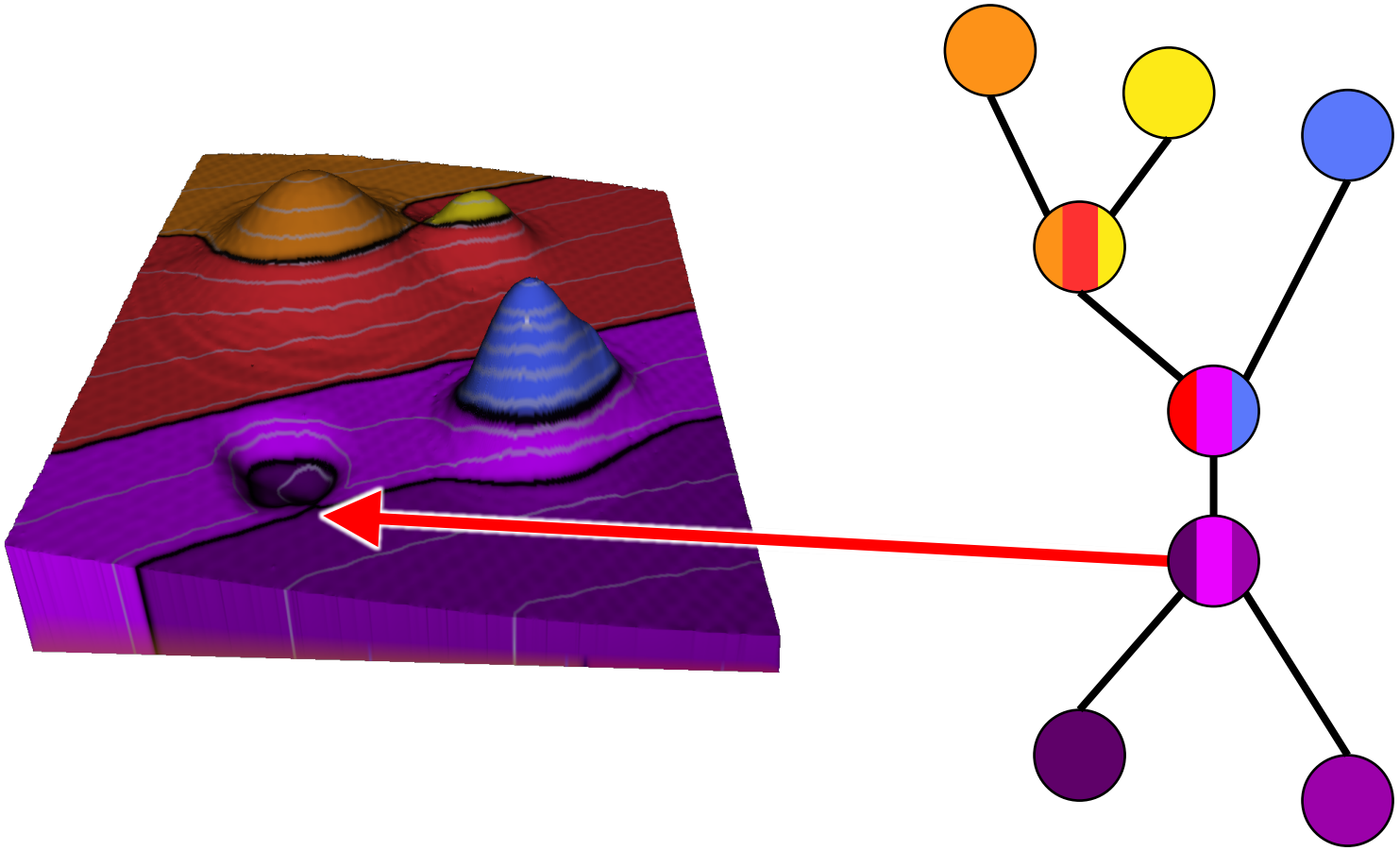
Contour trees



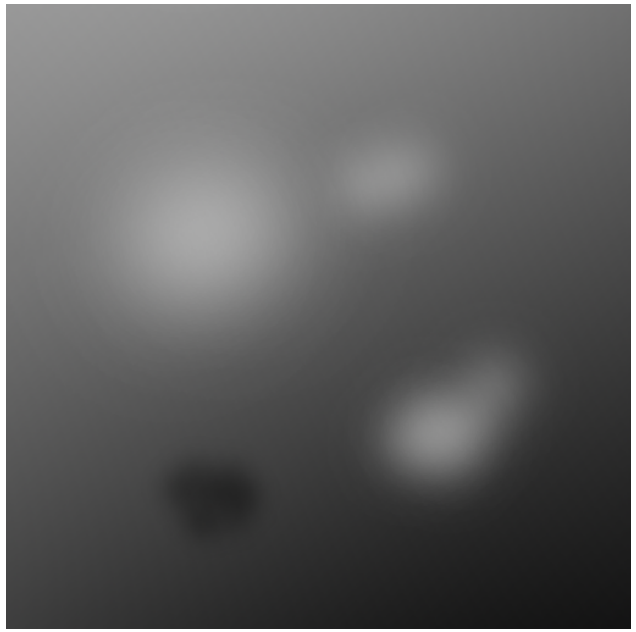


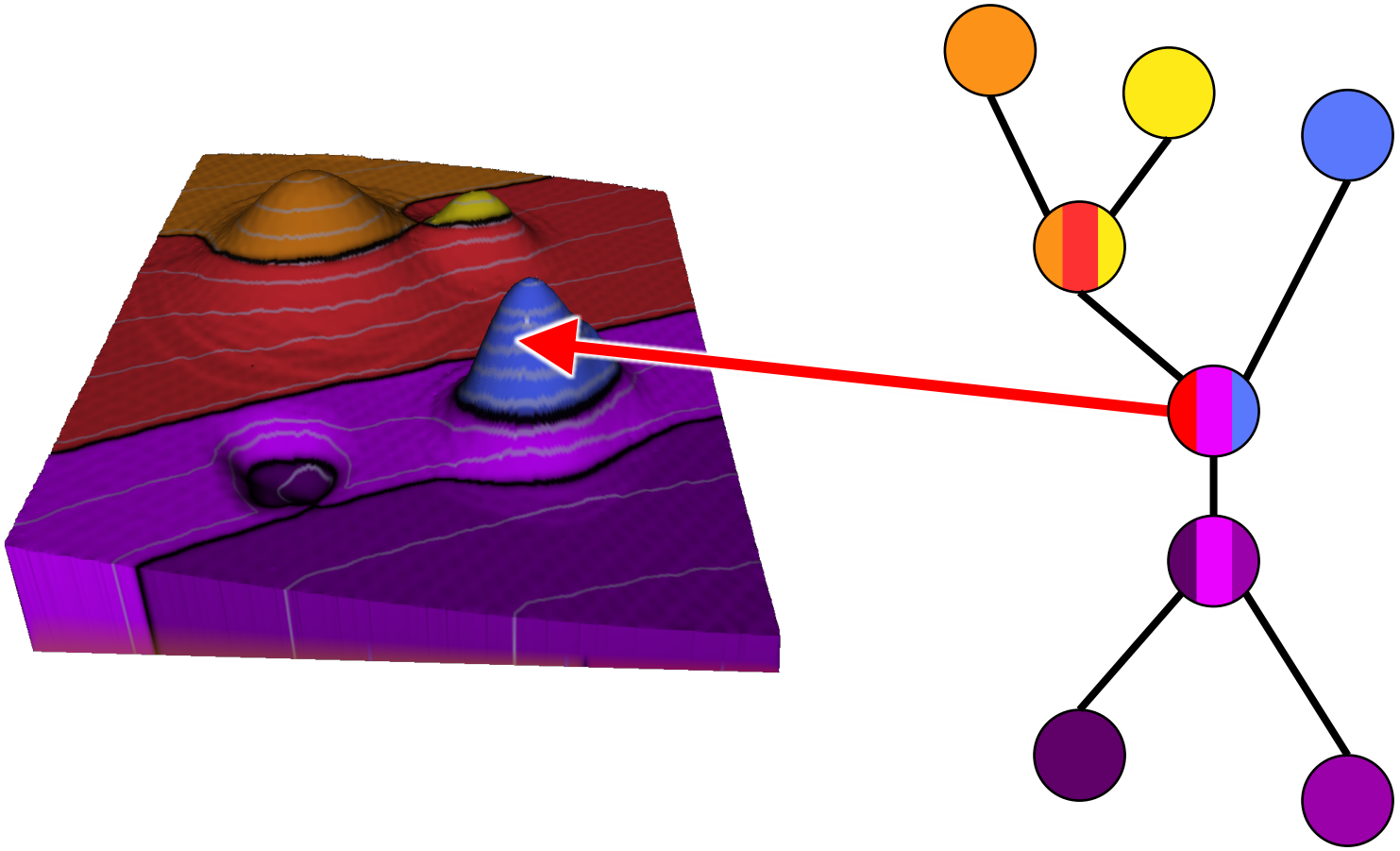
Contour trees



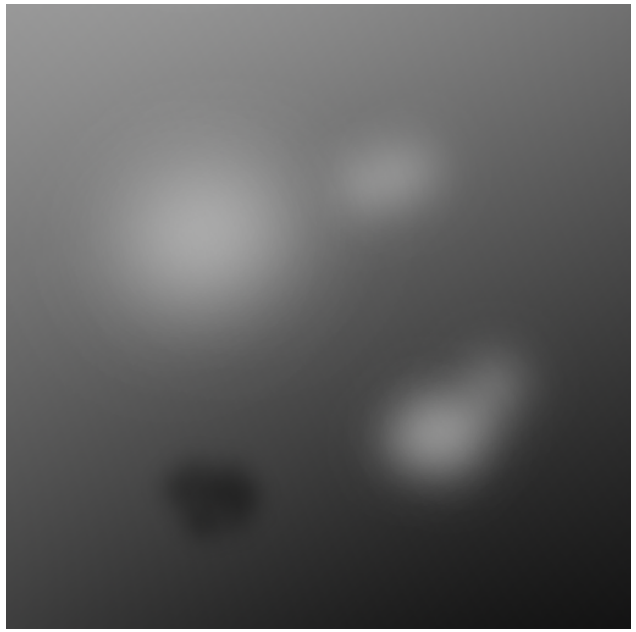


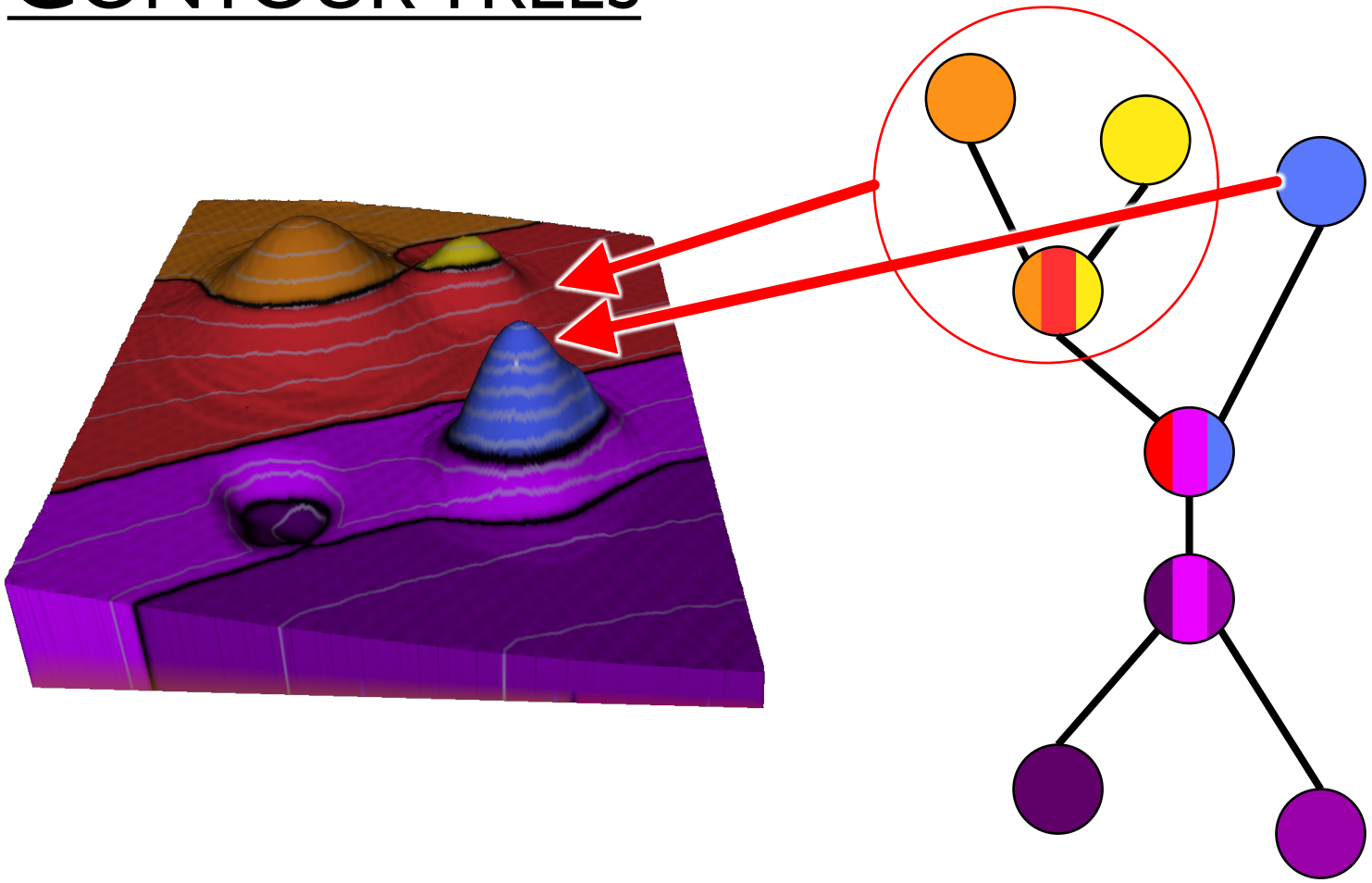
Contour trees



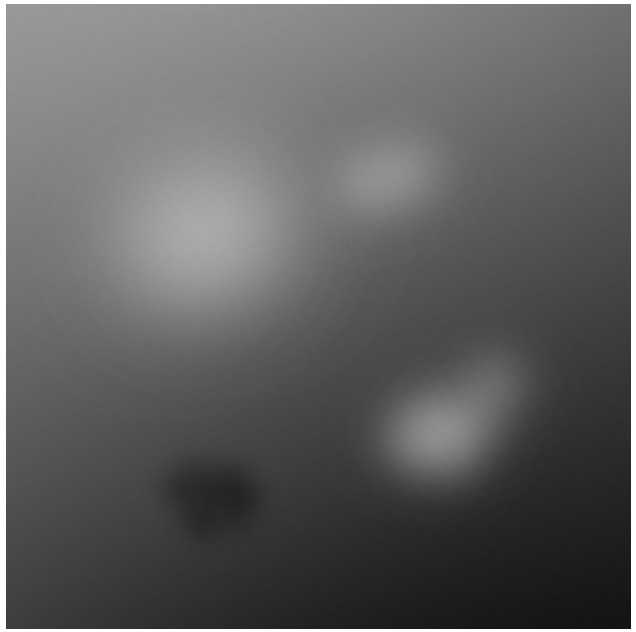


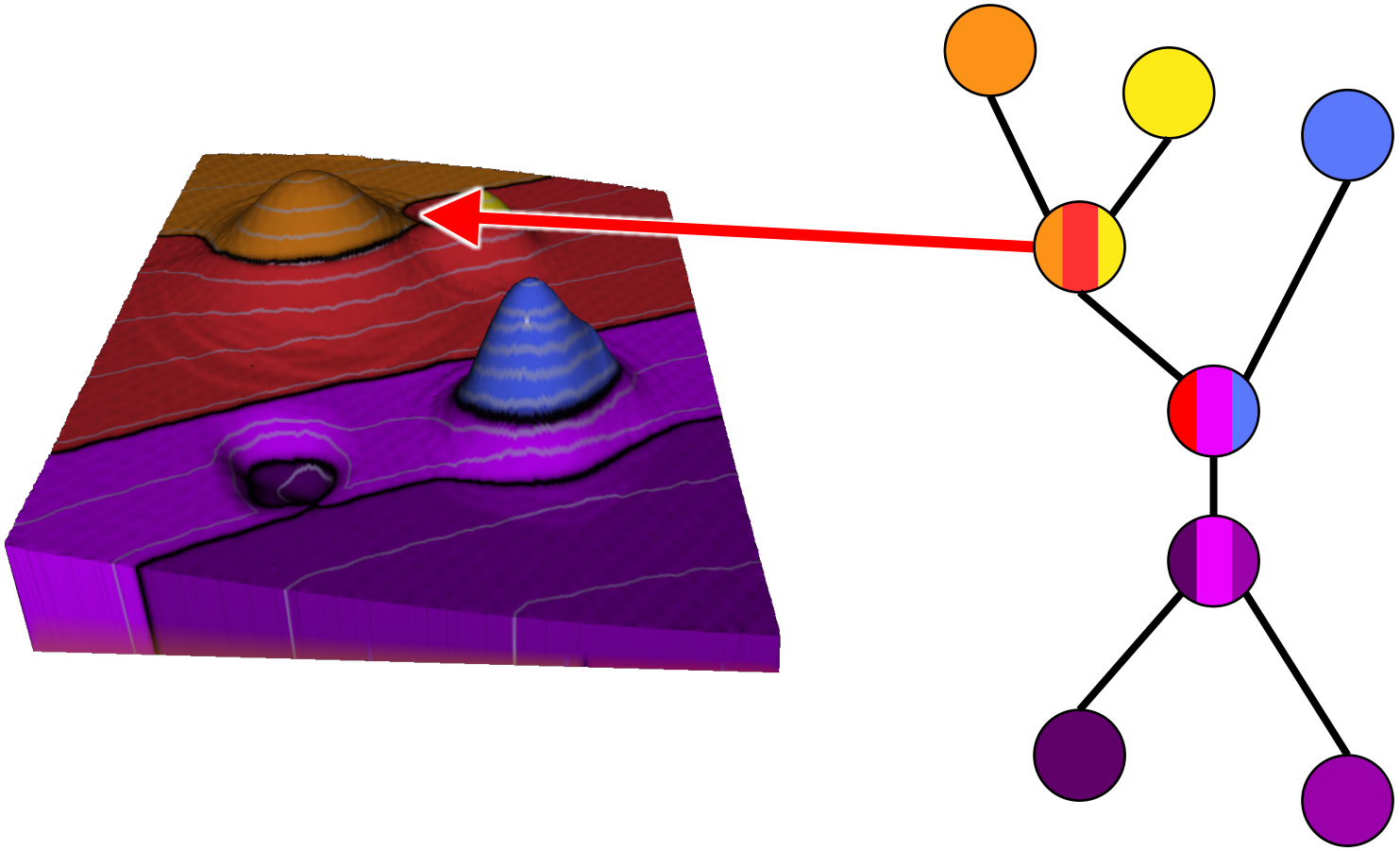
Contour trees



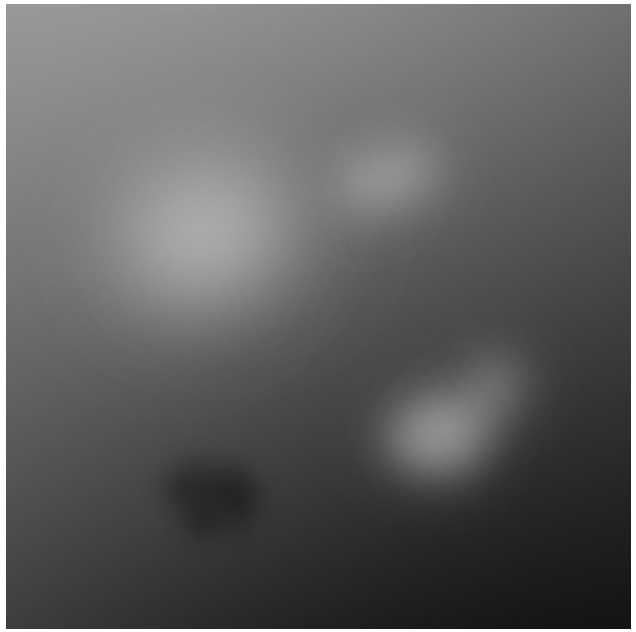


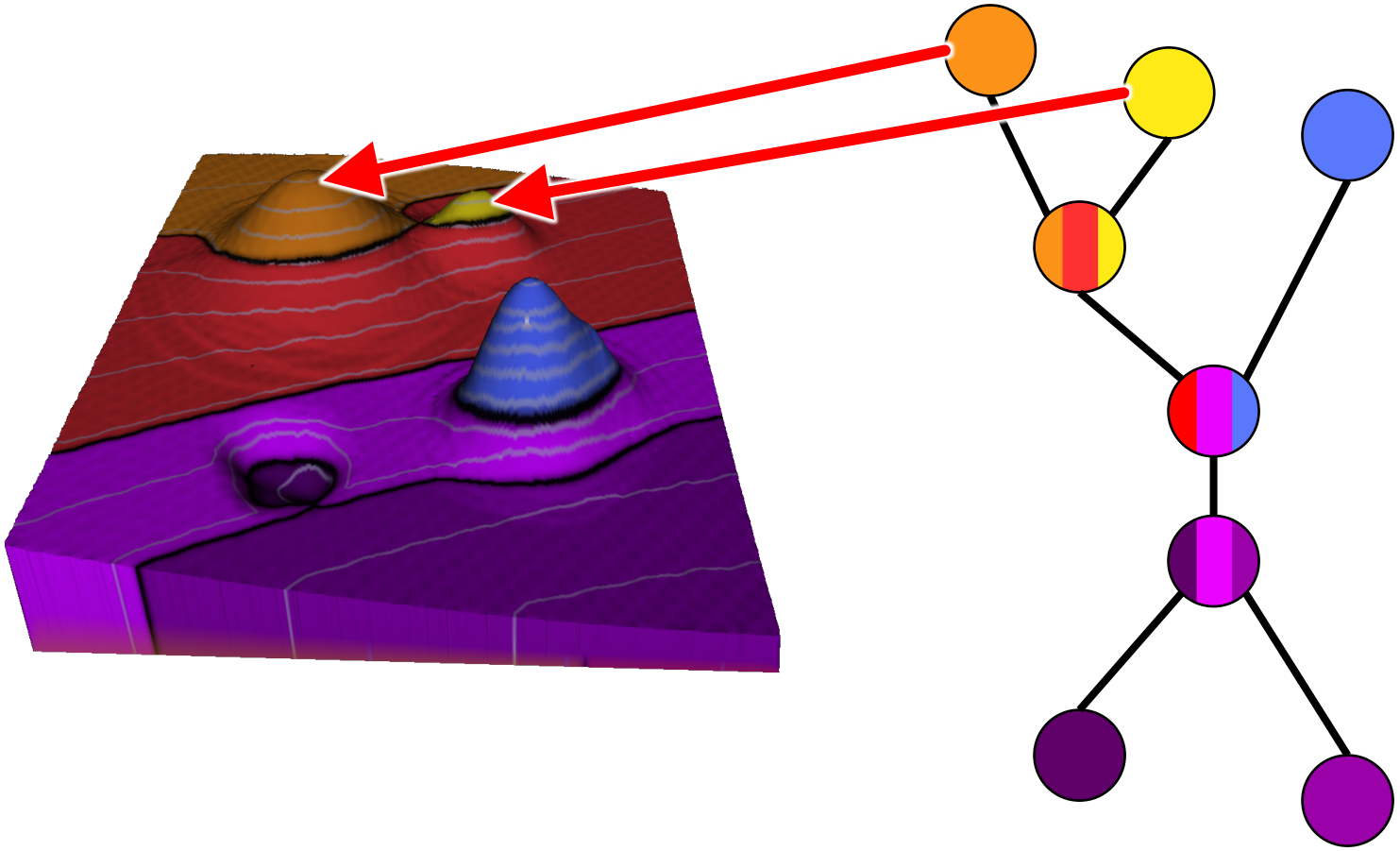
Contour trees



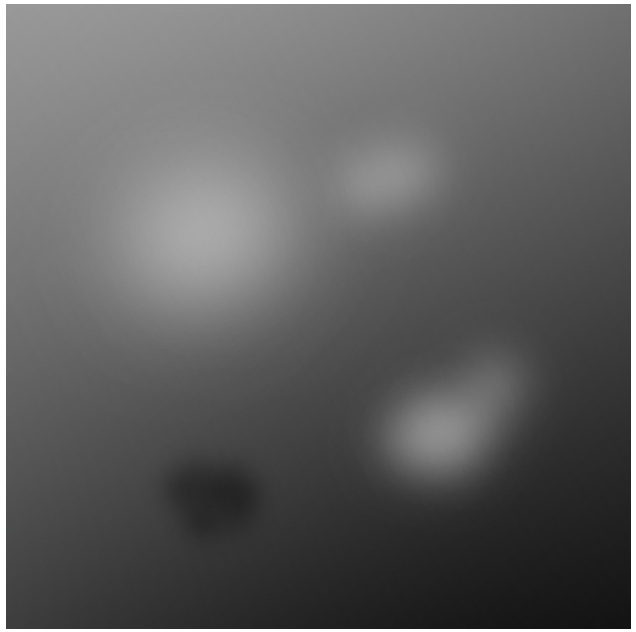


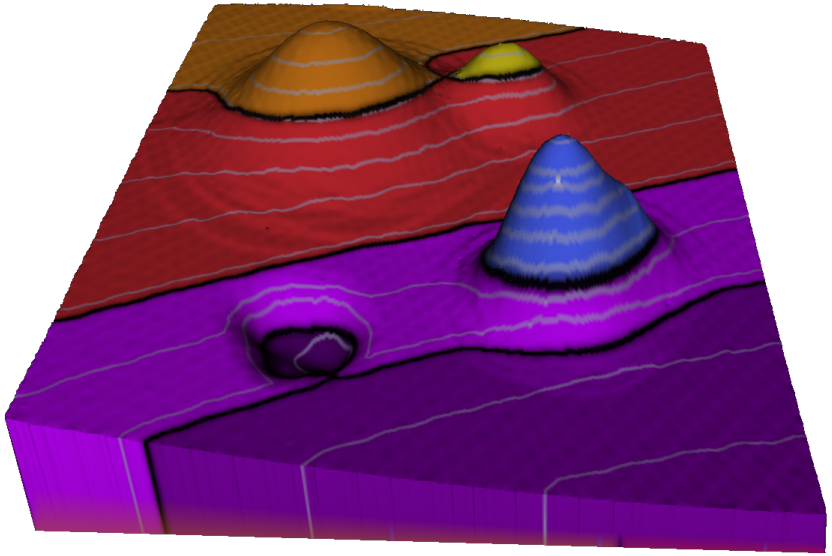
Contour trees

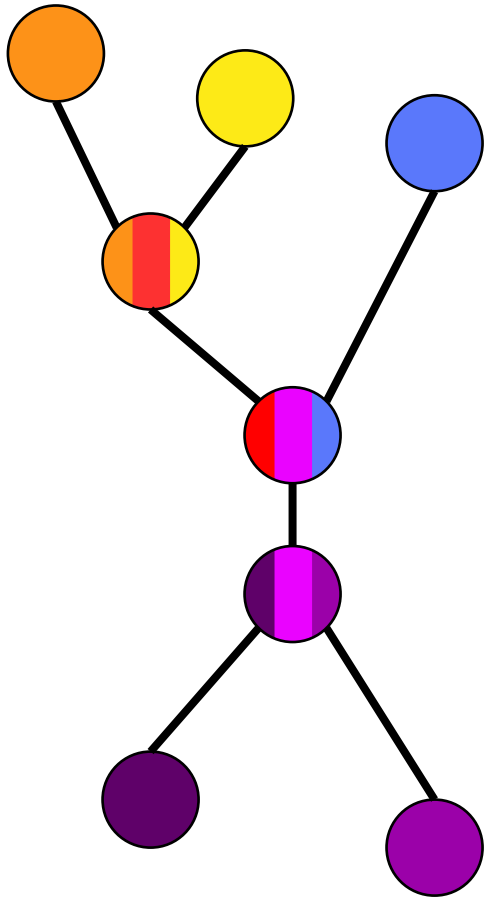




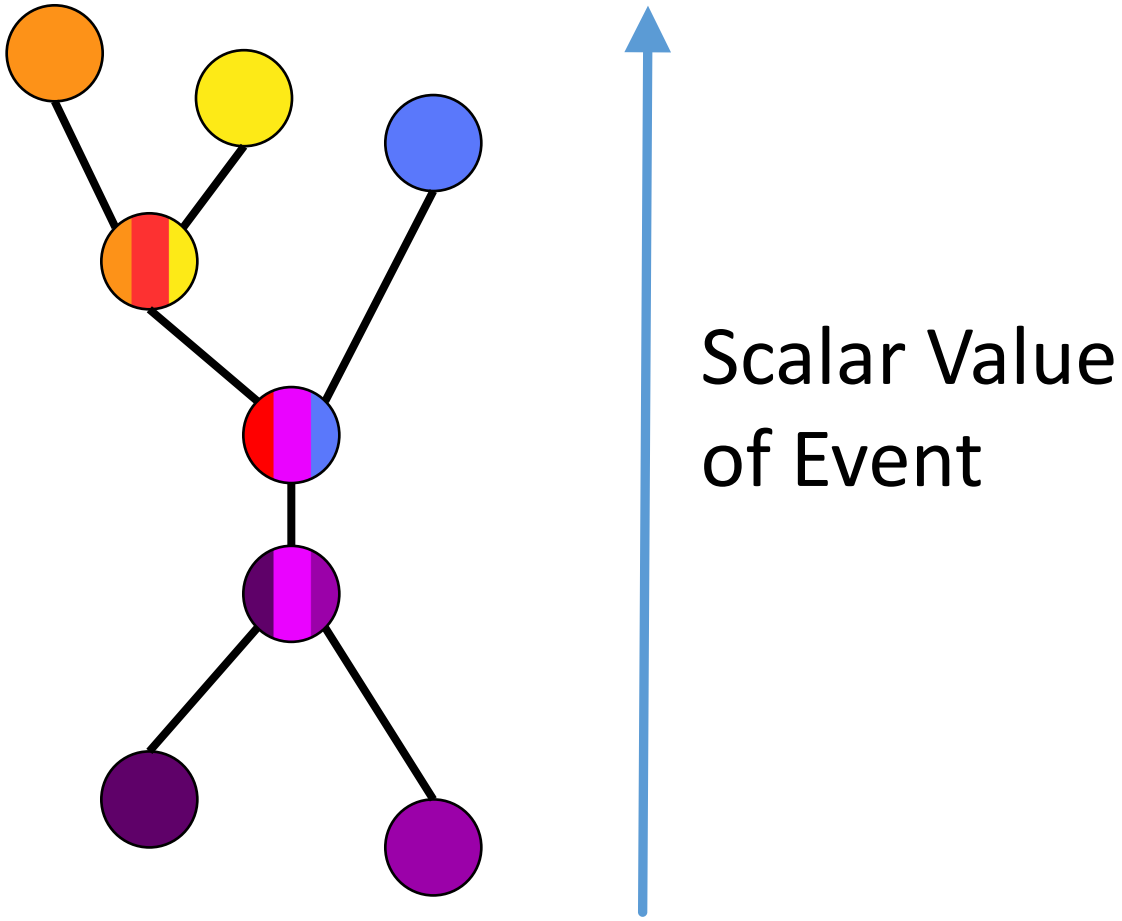
Contour trees



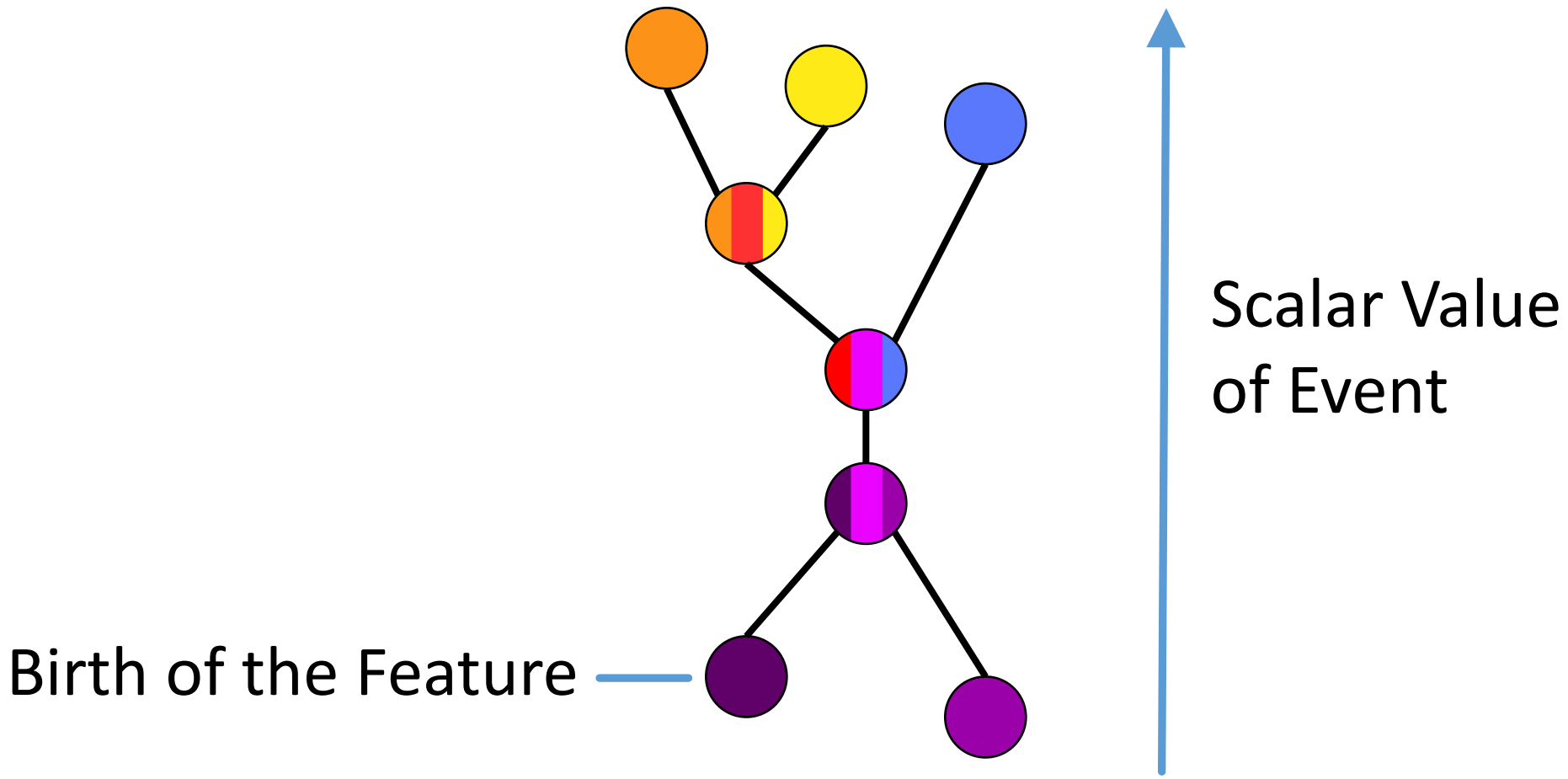




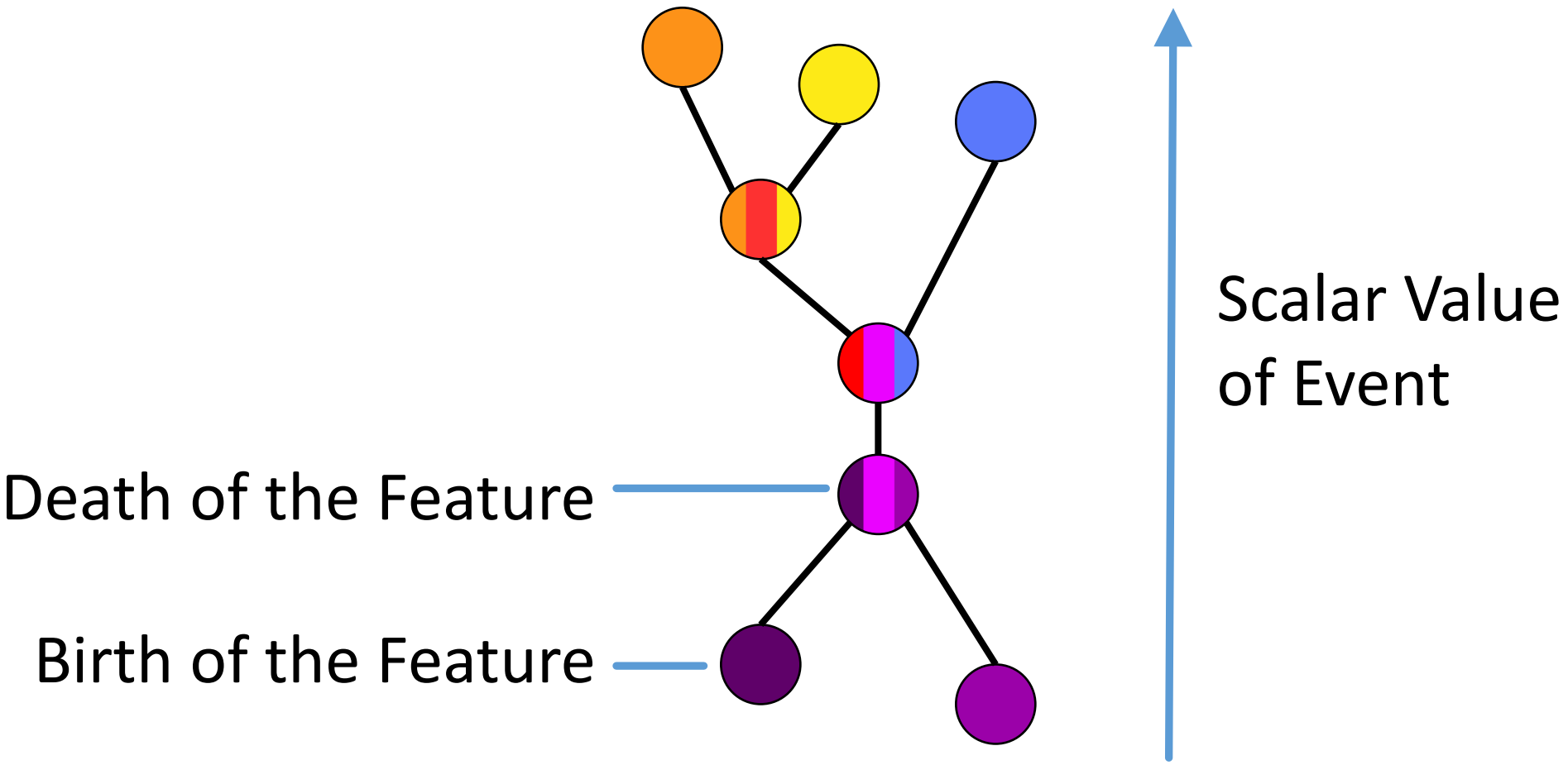
A Closer Look at the Contour Tree



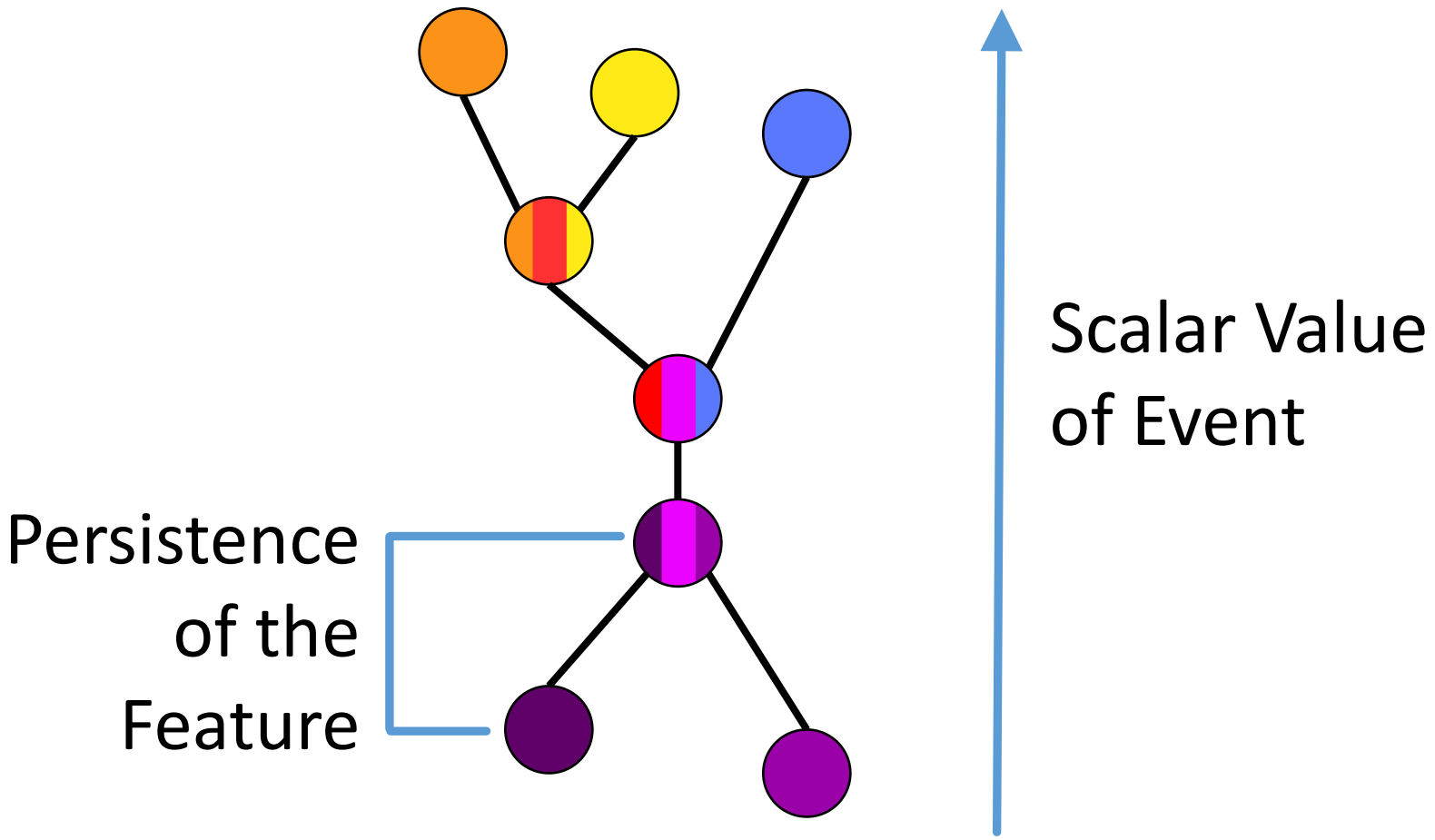
A Closer Look at the Contour Tree



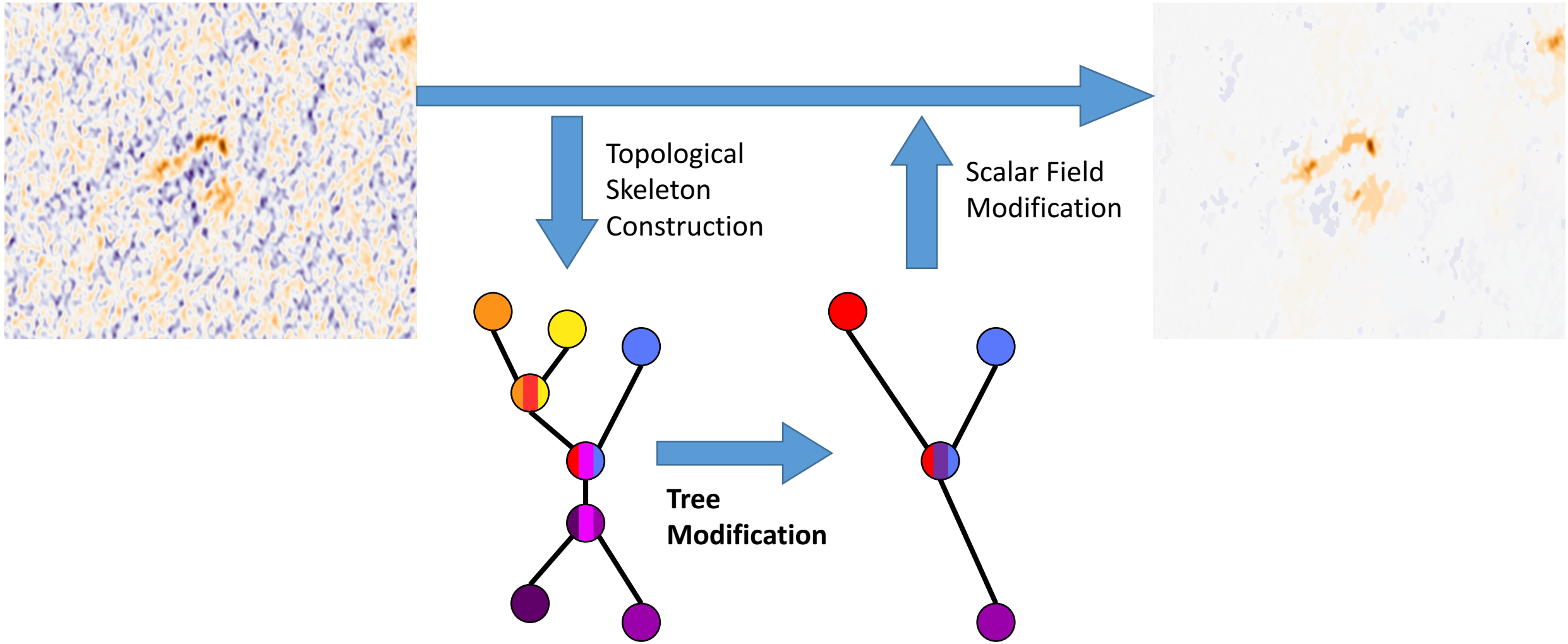
A Closer Look at the Contour Tree



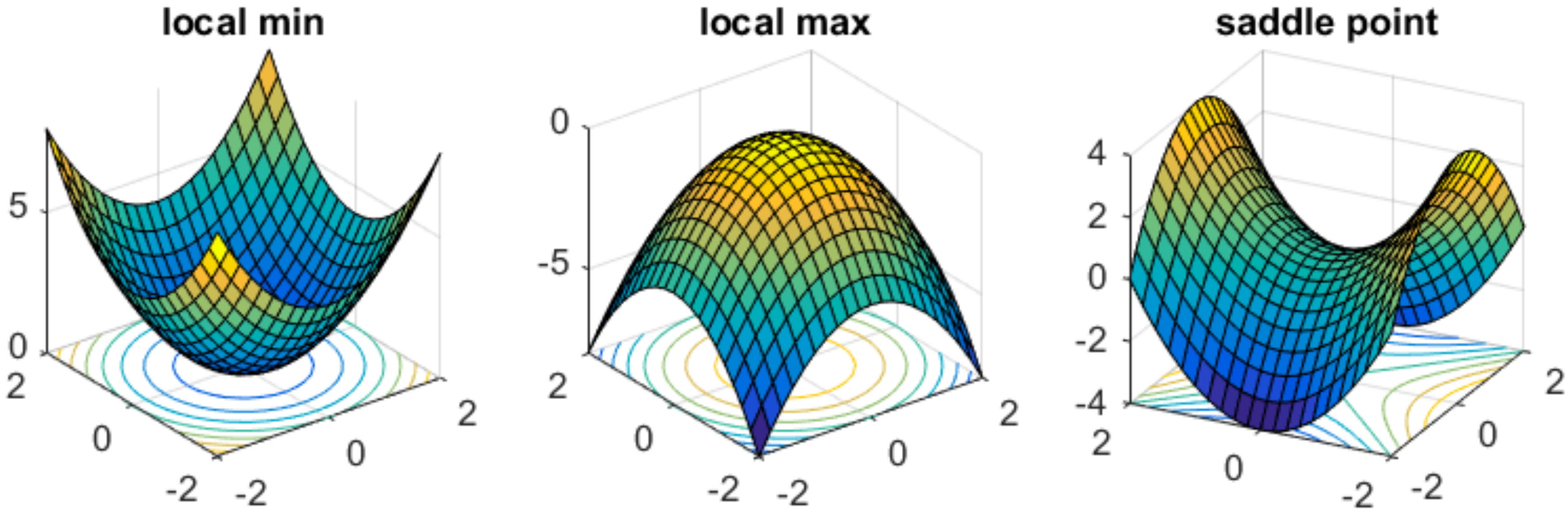
A Closer Look at the Contour Tree



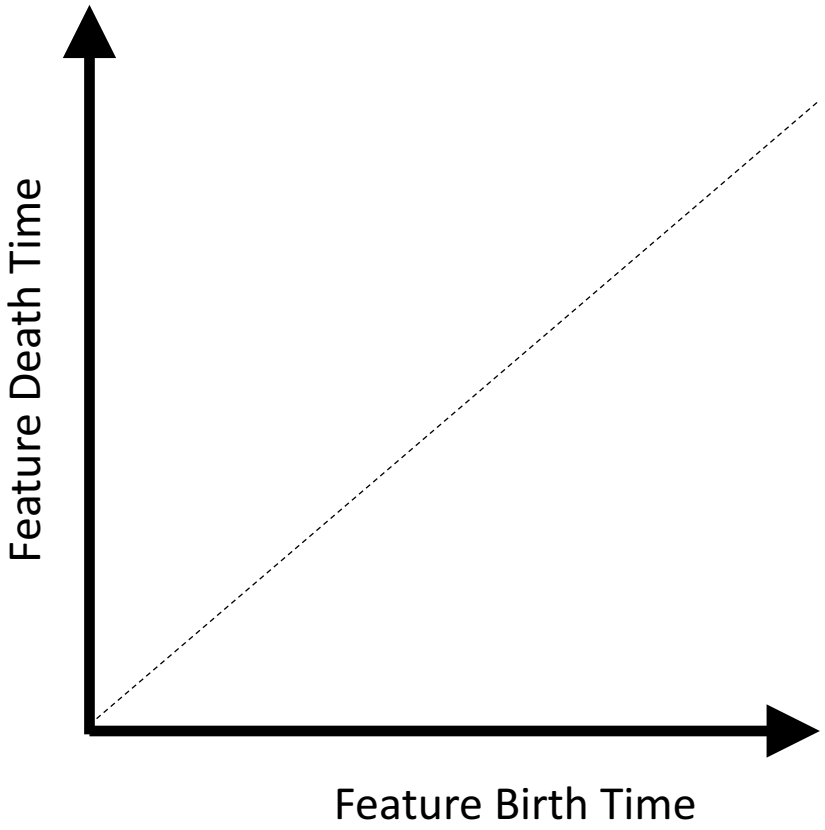
Usage Scenario

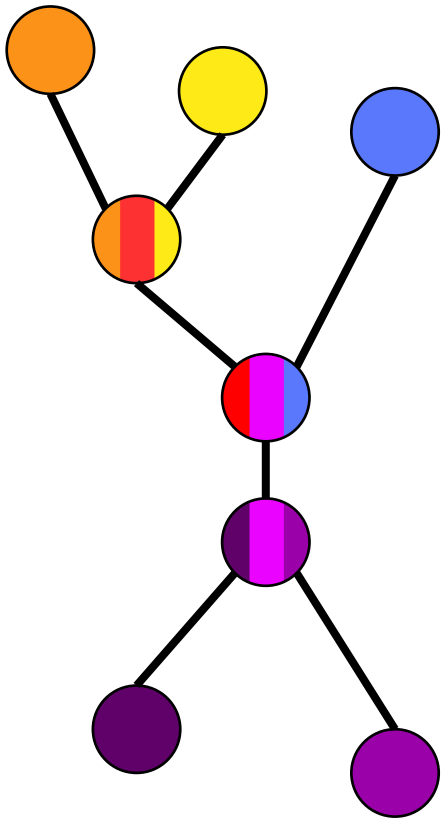


Critical Point Cancellation

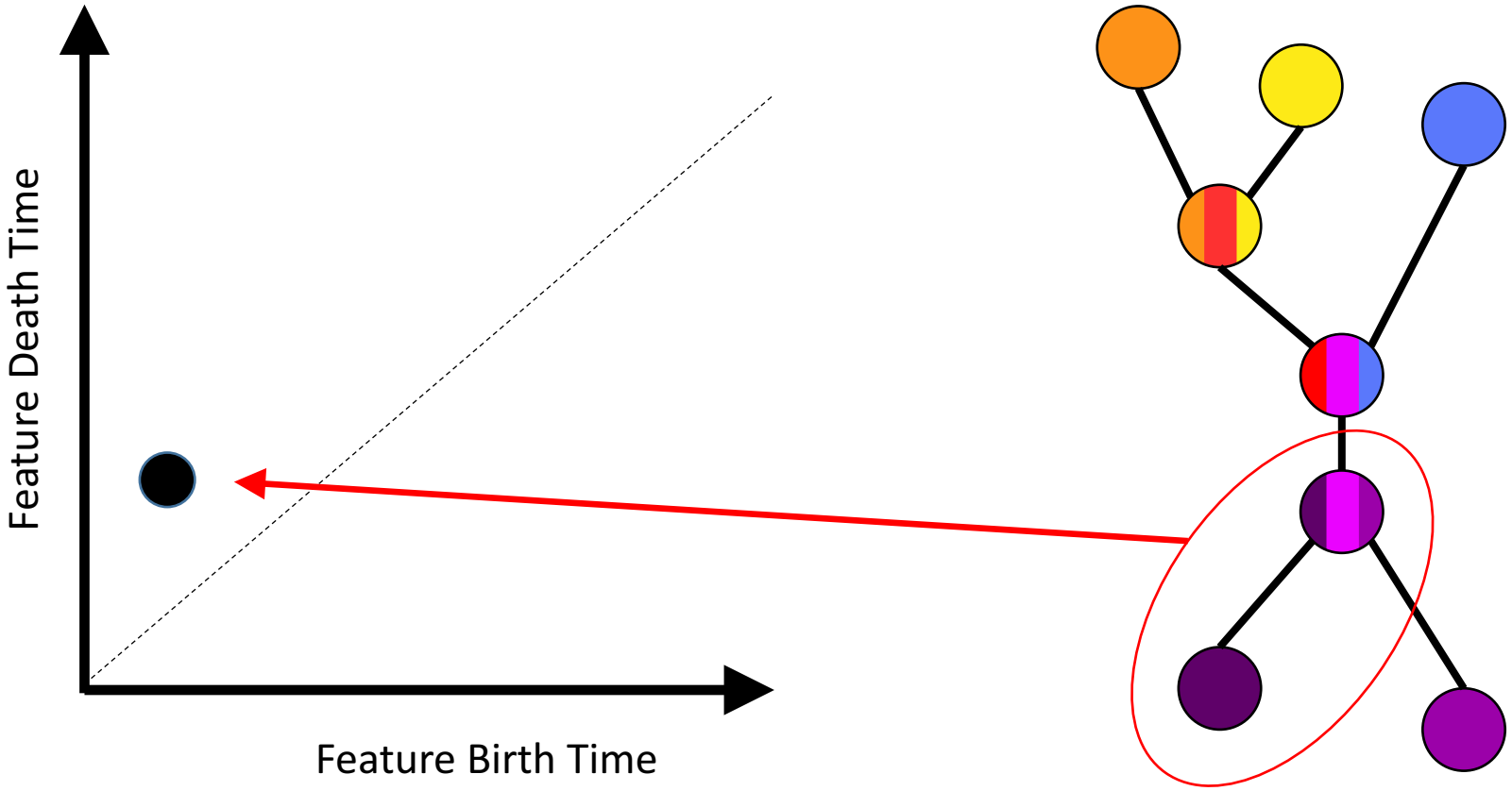


Controlling Simplification
The Persistence Diagram

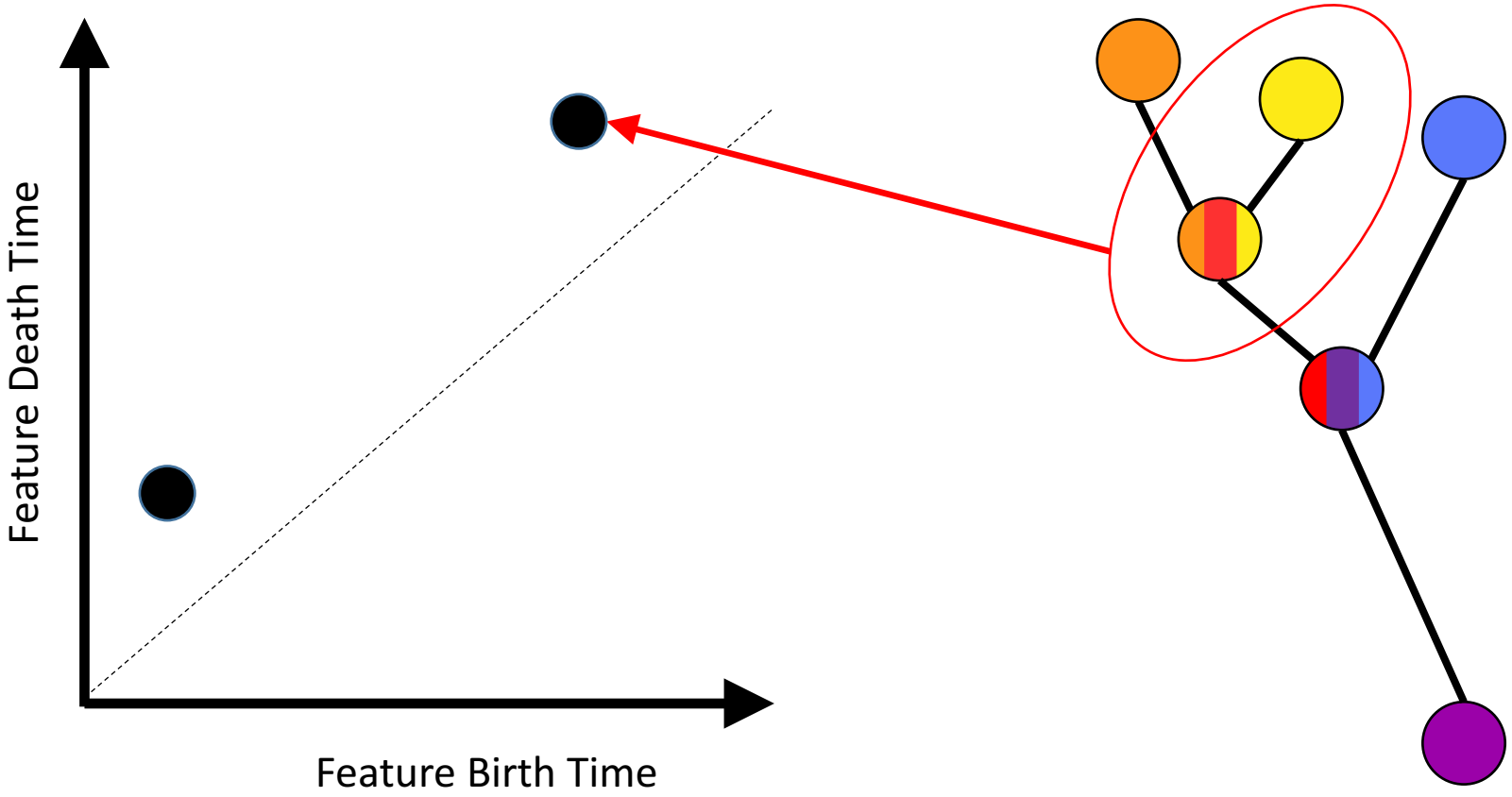




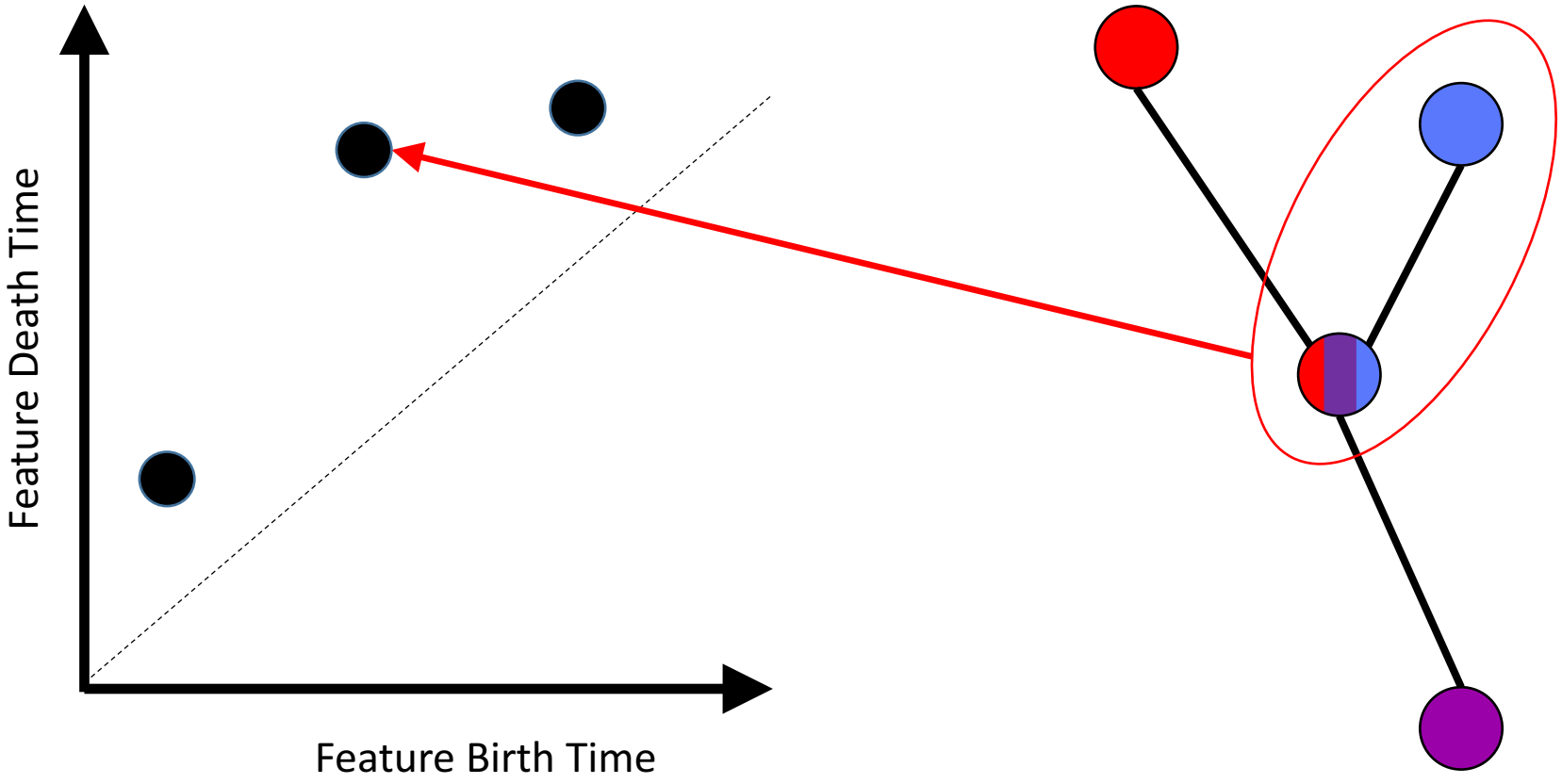
Controlling Simplification
The Persistence Diagram



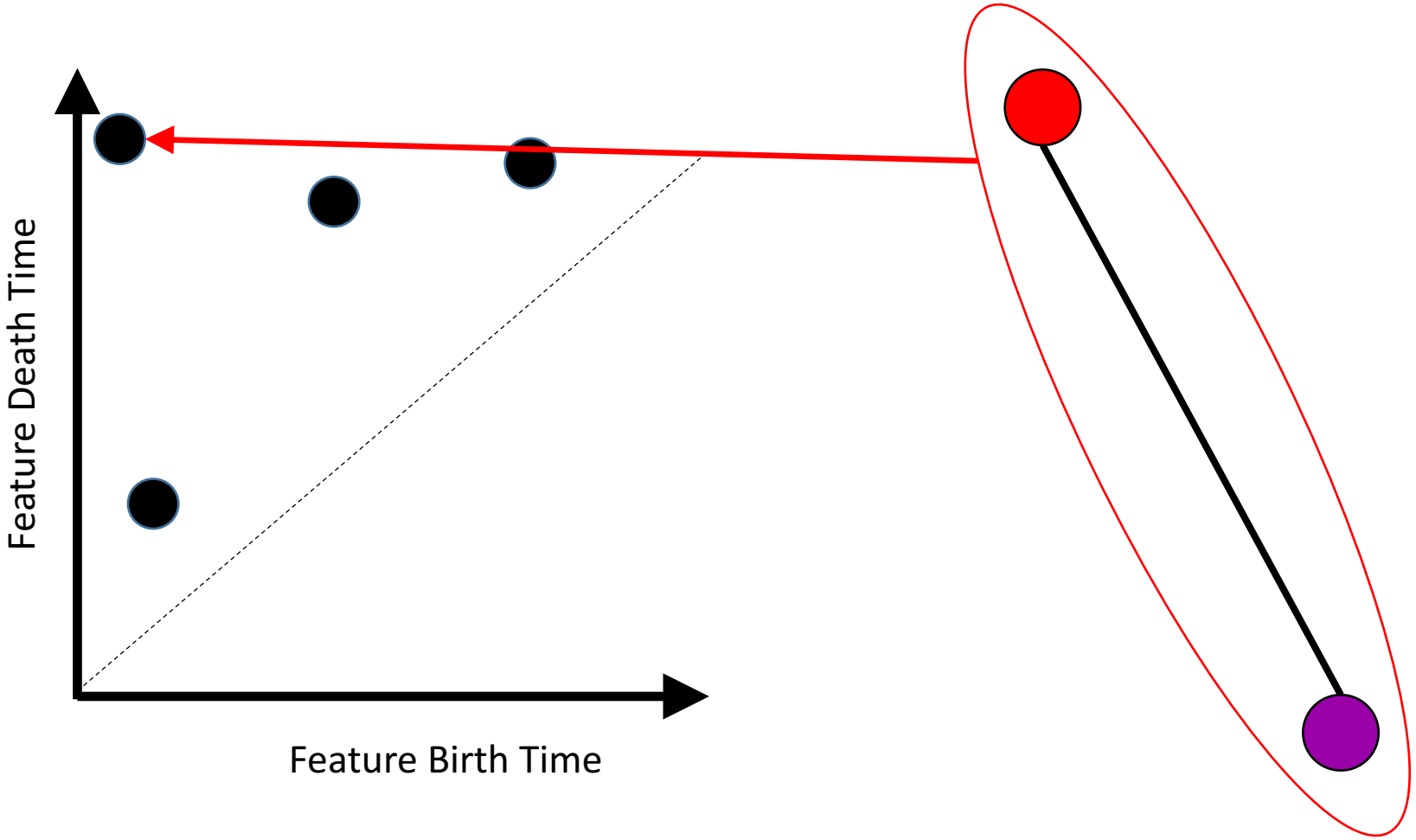
Controlling Simplification
The Persistence Diagram



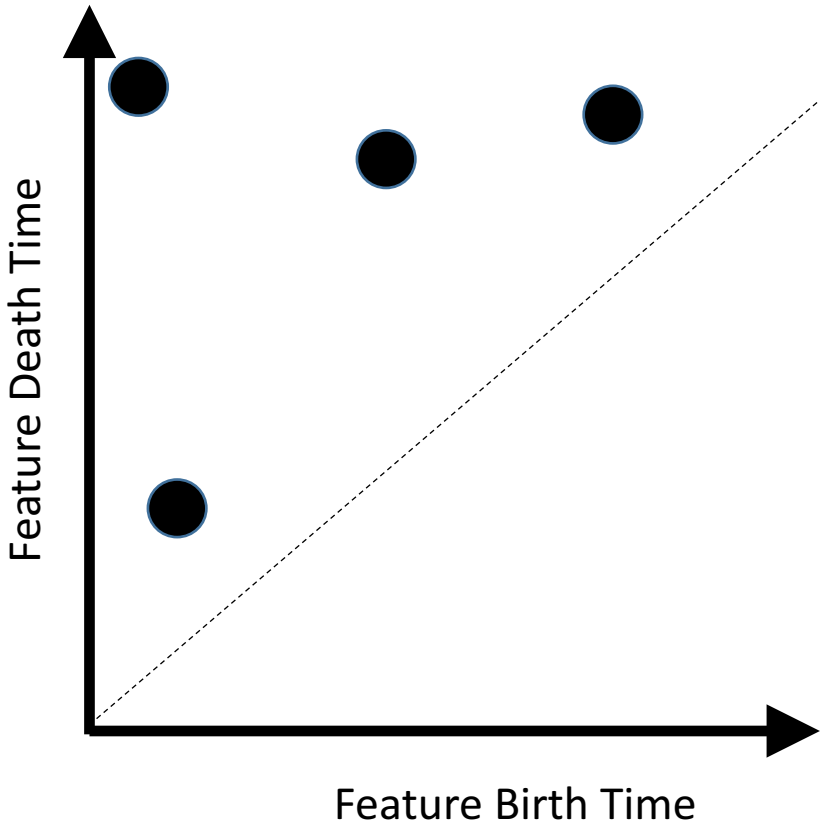
Controlling Simplification
The Persistence Diagram

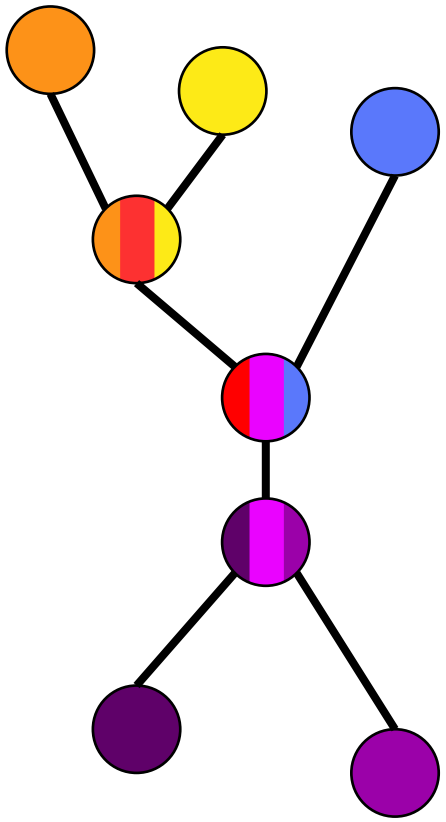


Controlling Simplification
The Persistence Diagram

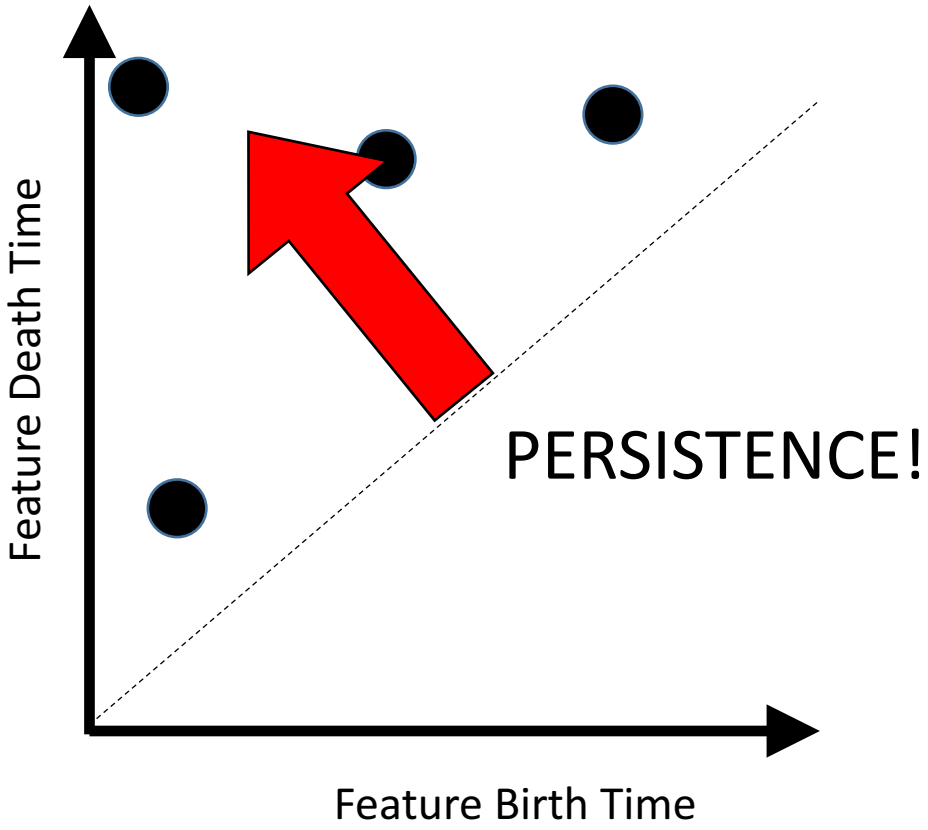


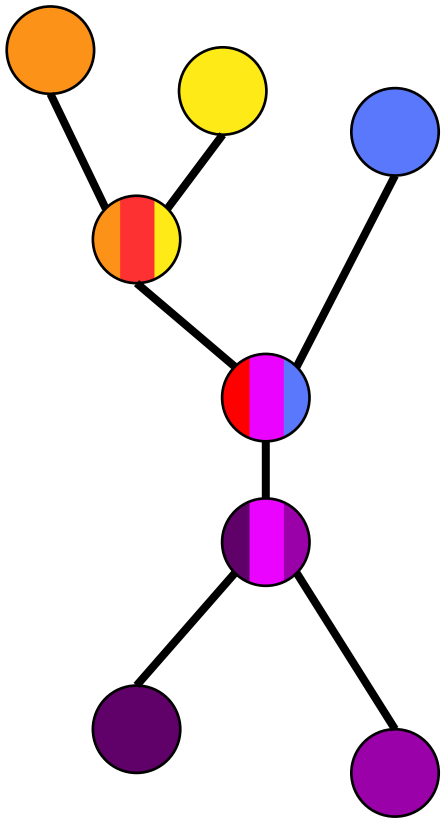
Controlling Simplification
The Persistence Diagram



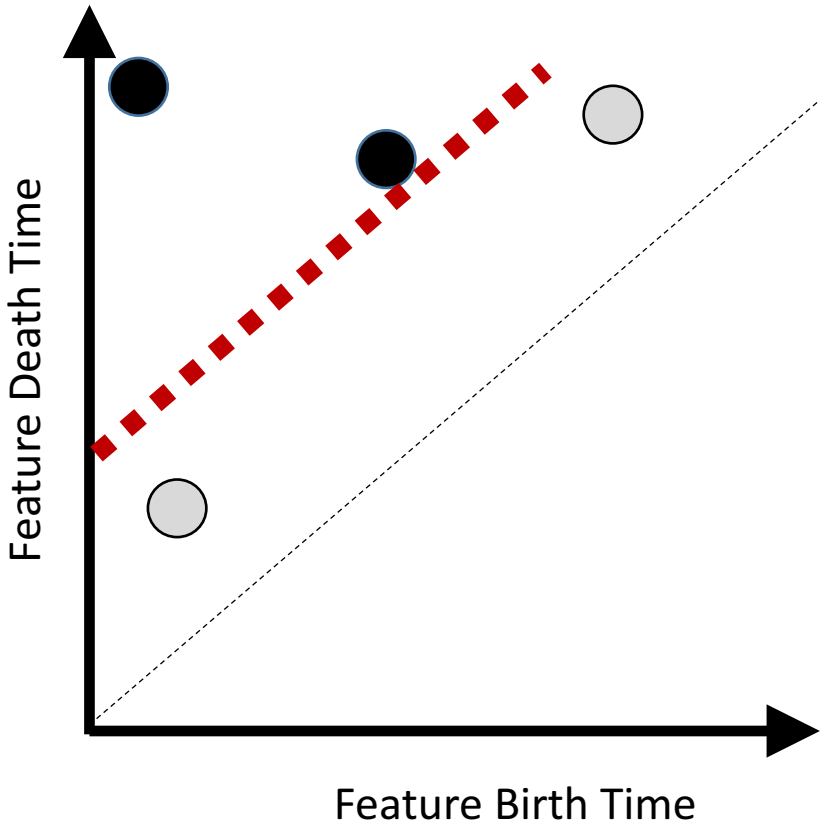


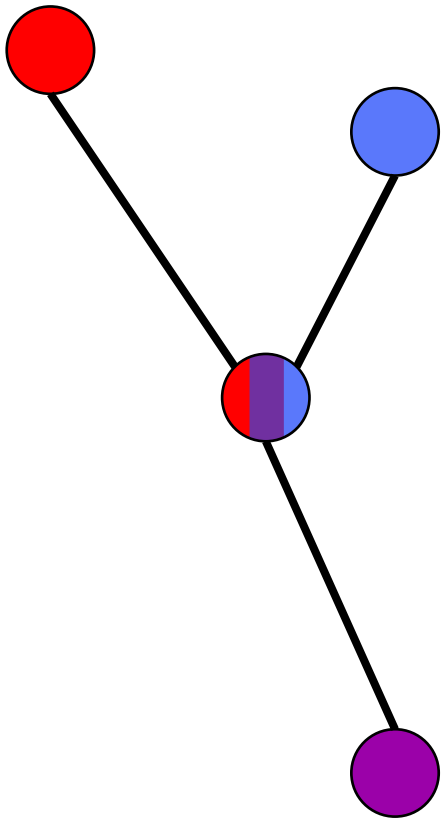
Controlling Simplification
The Persistence Diagram



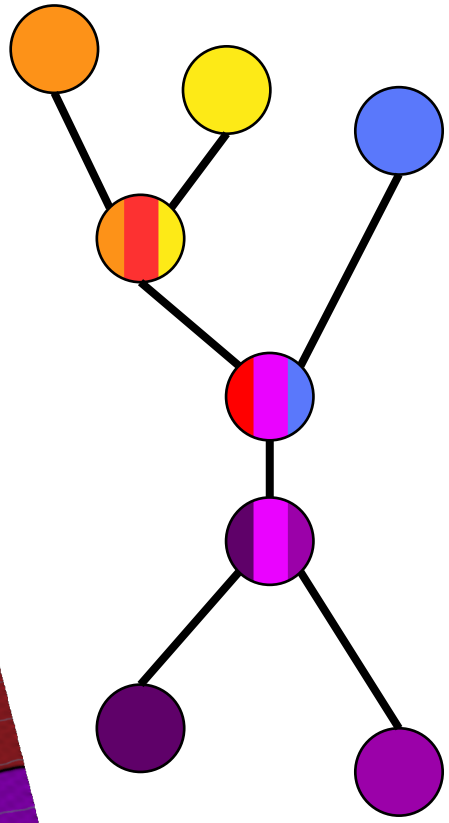


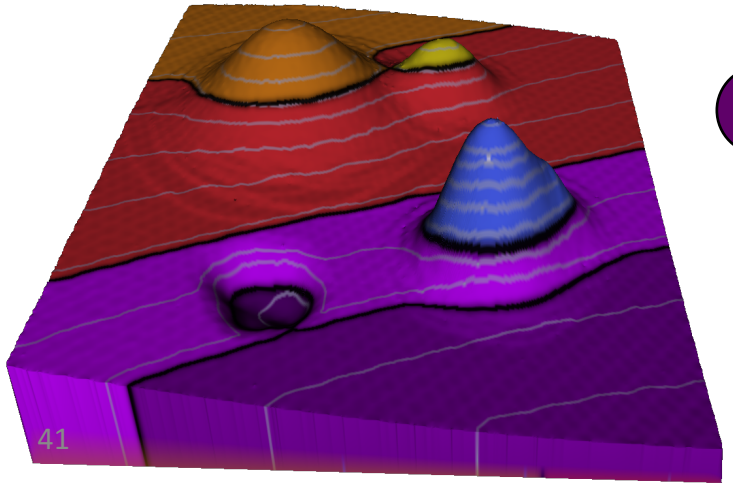
Controlling Simplification
The Persistence Diagram



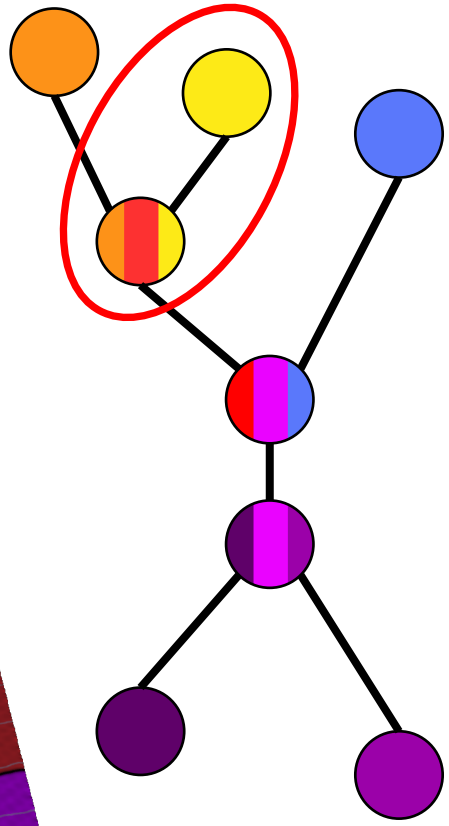


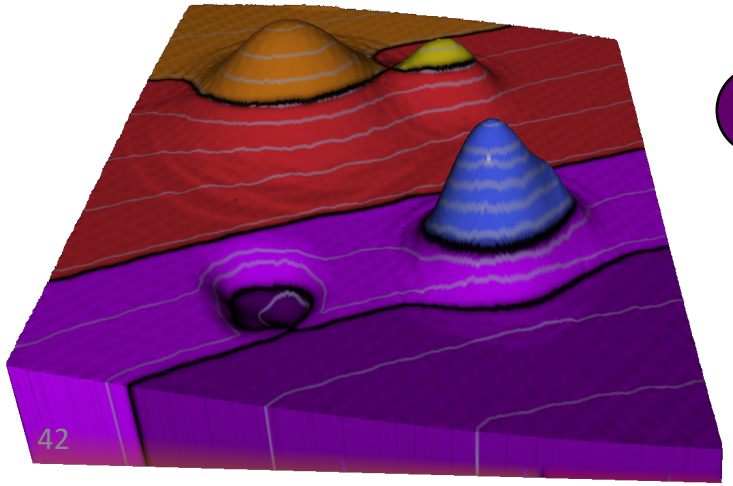
Feature Removal



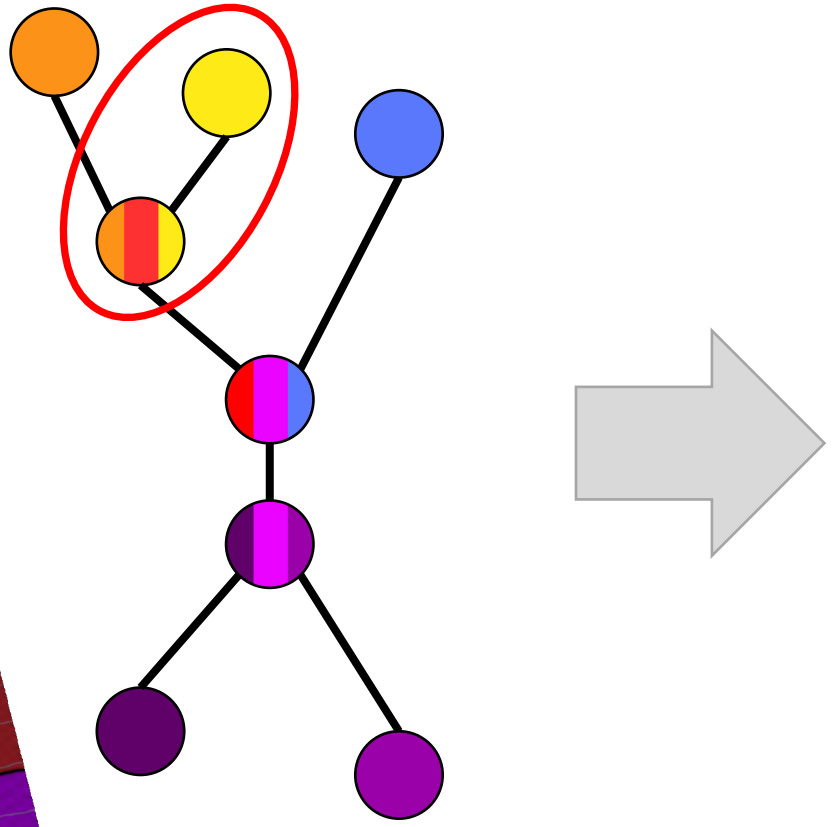


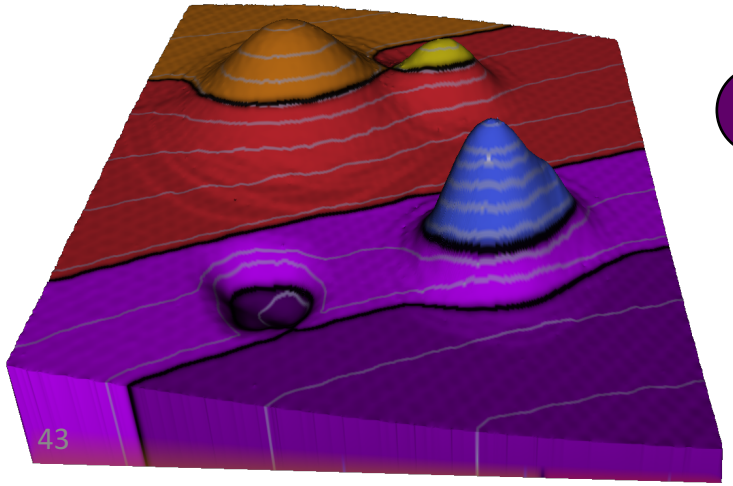
Feature Removal



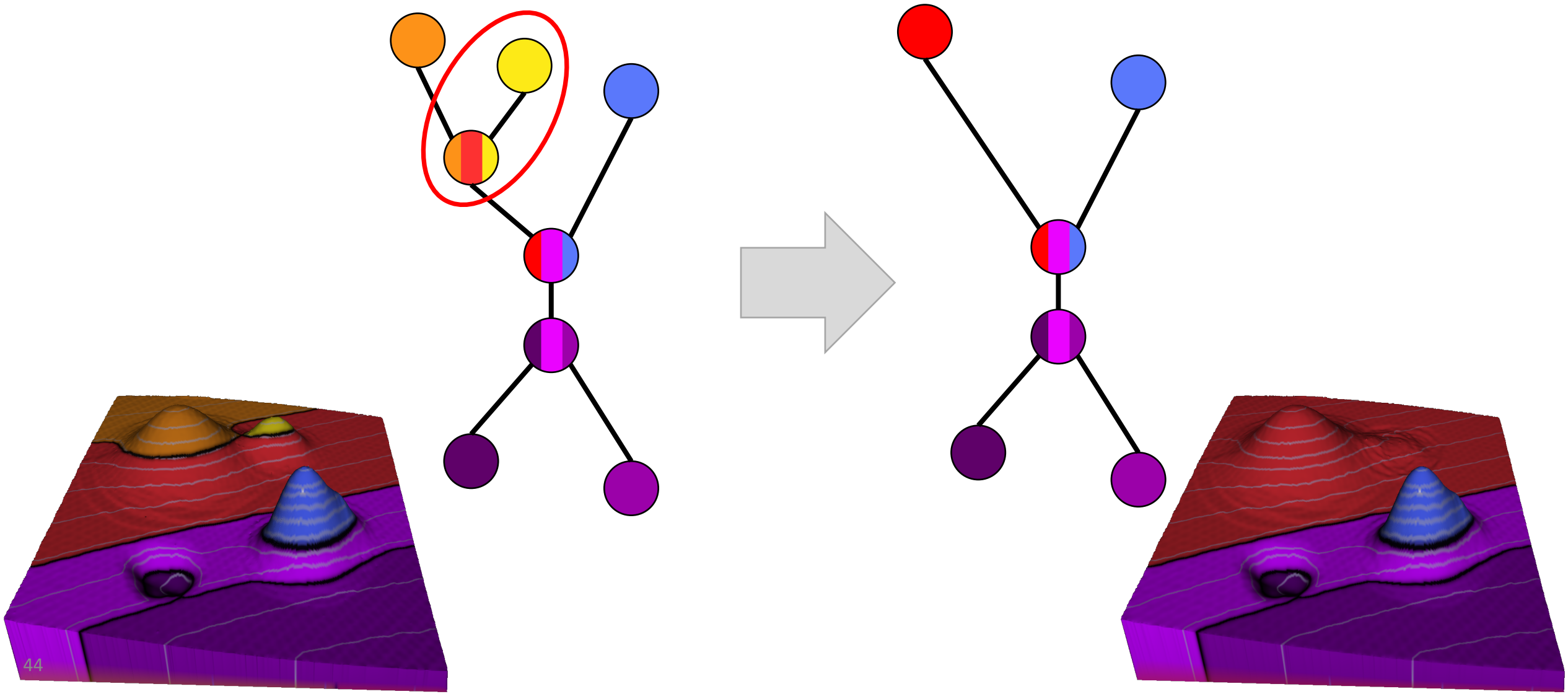


Feature Removal

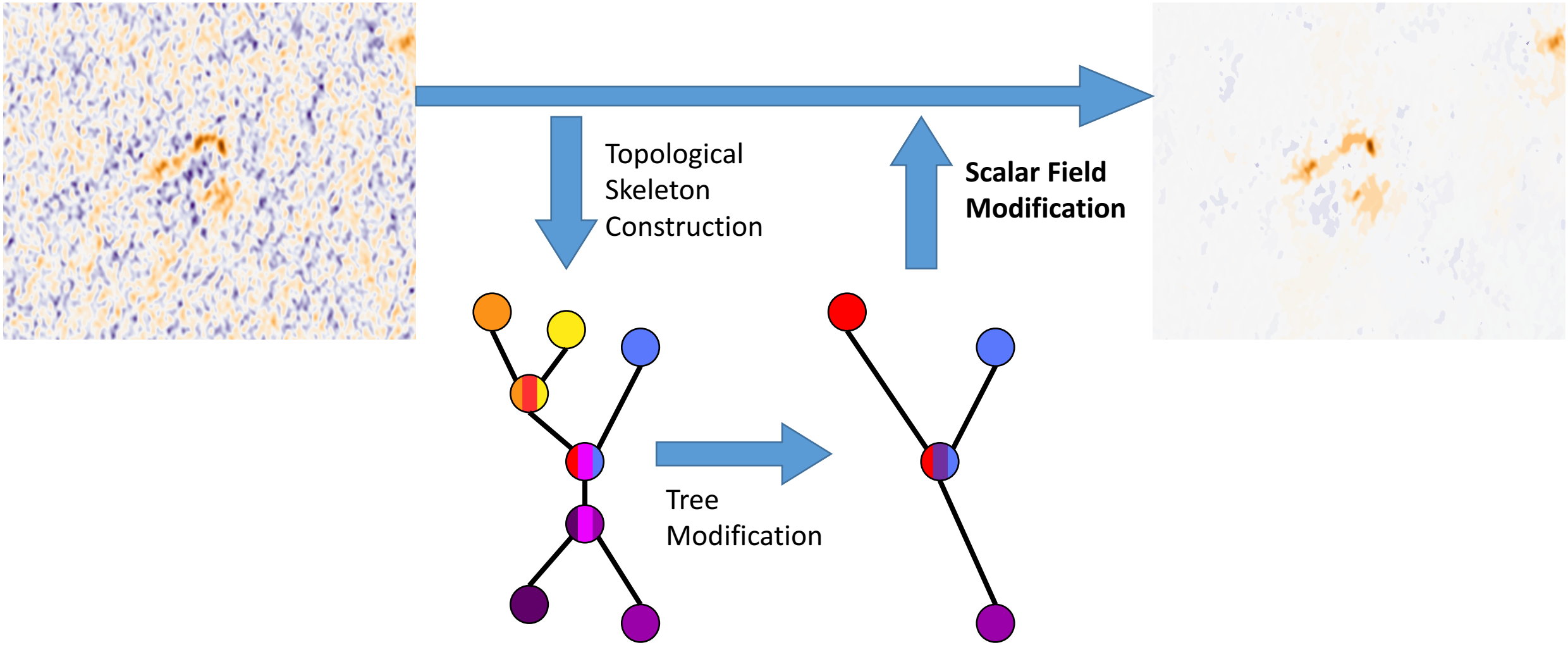




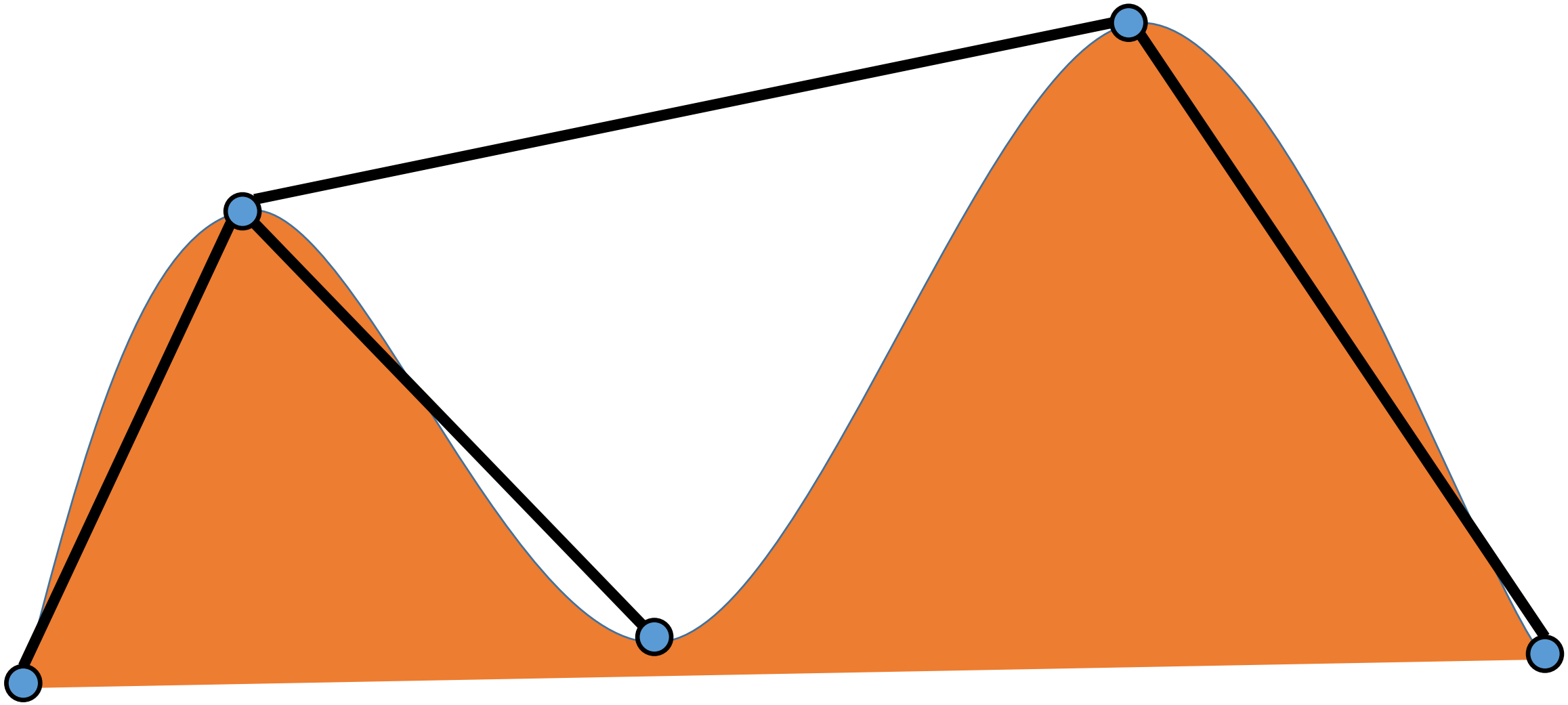
Feature Removal



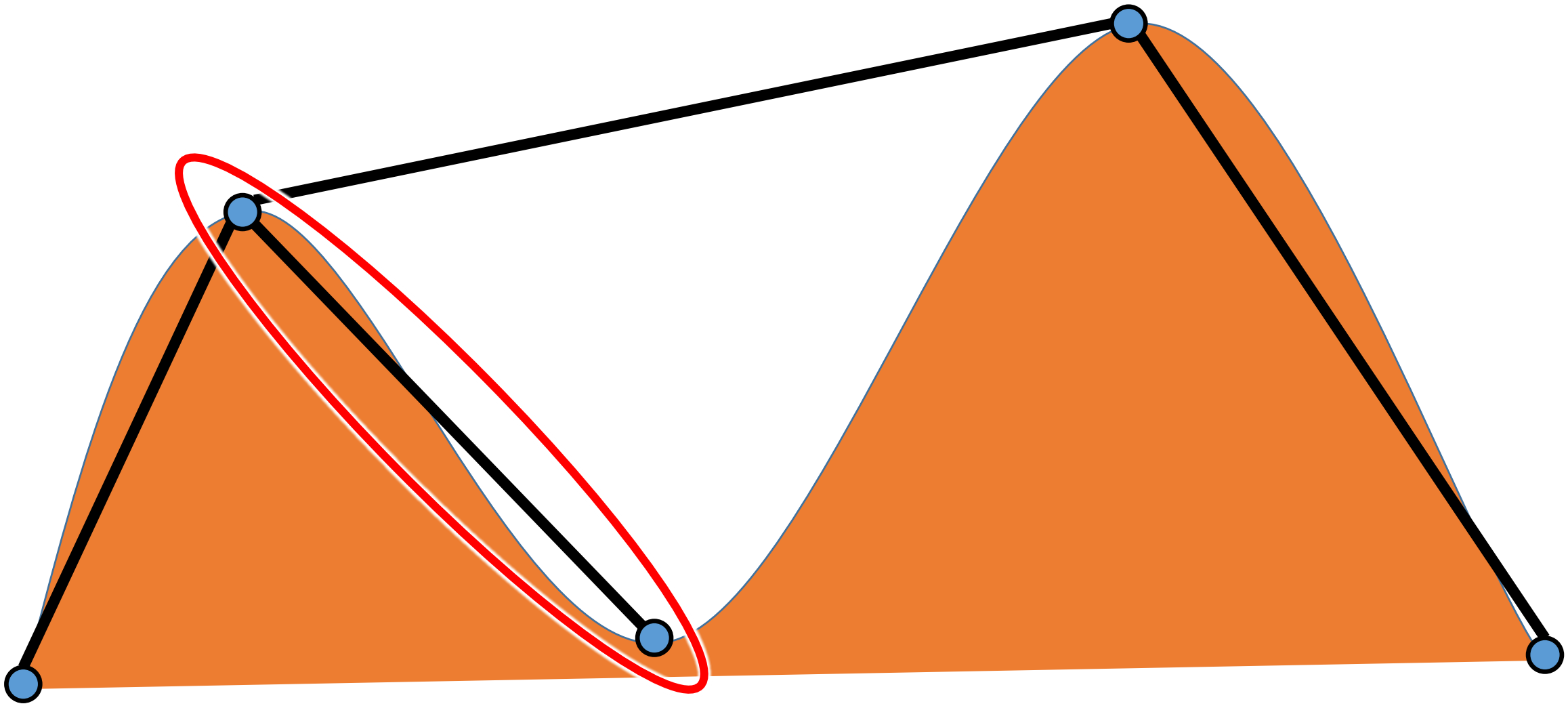
Usage Scenario



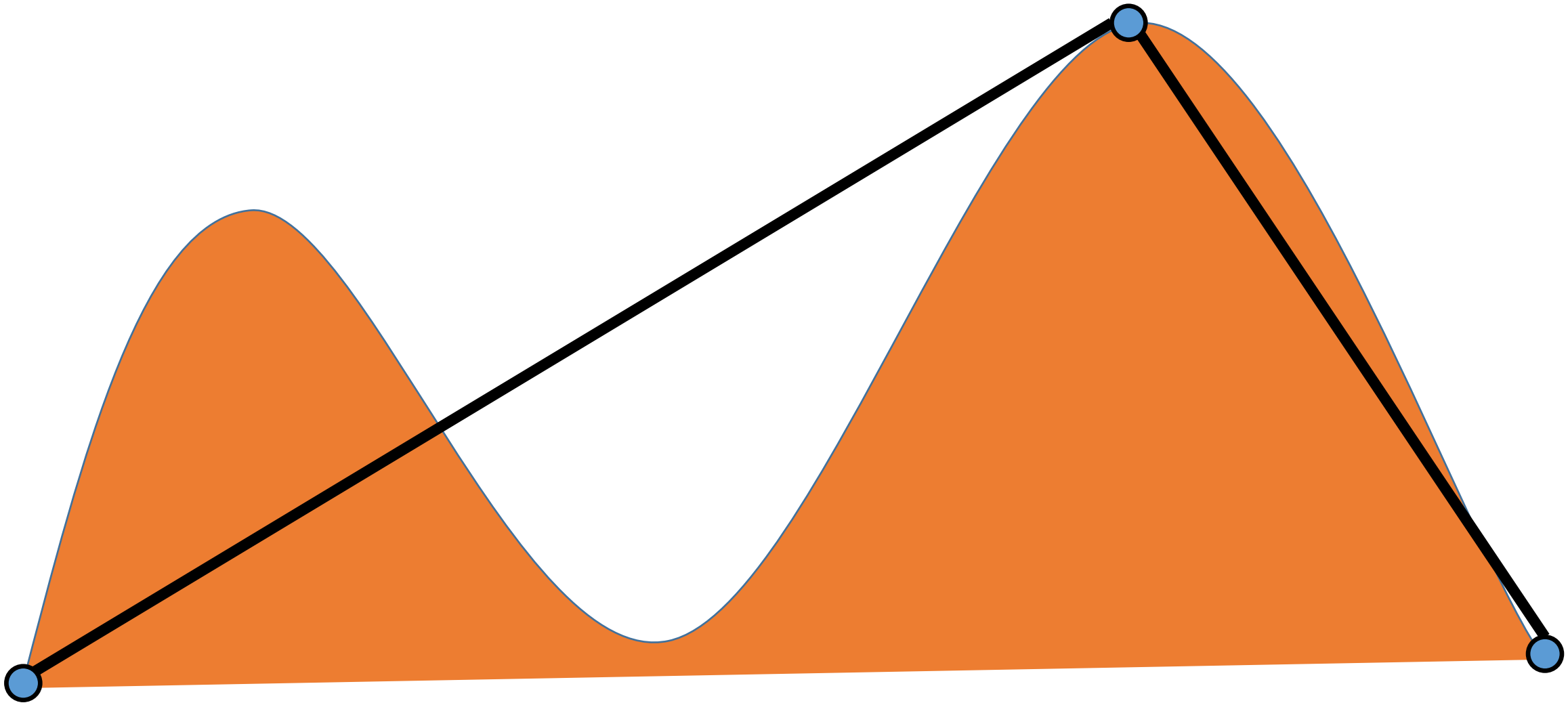
Scalar Field Modification



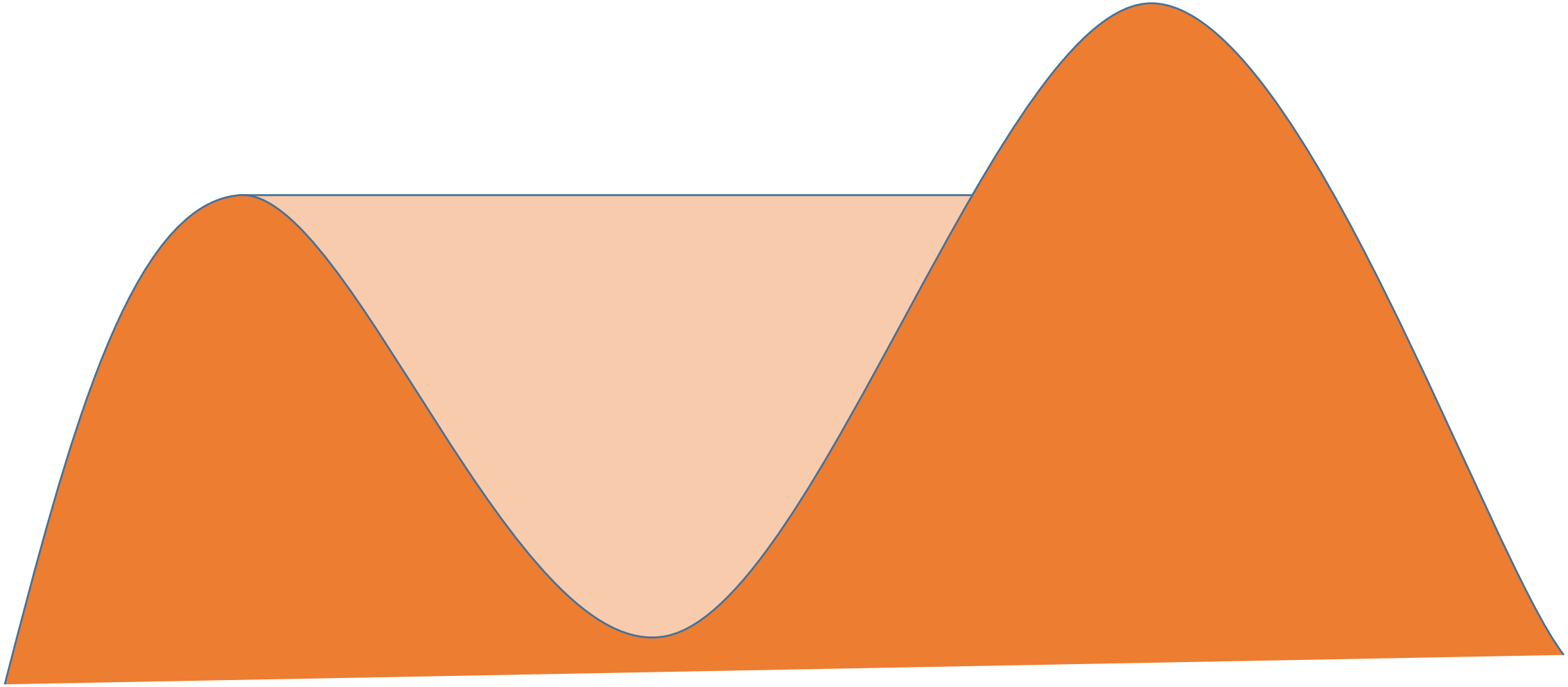
Scalar Field Modification



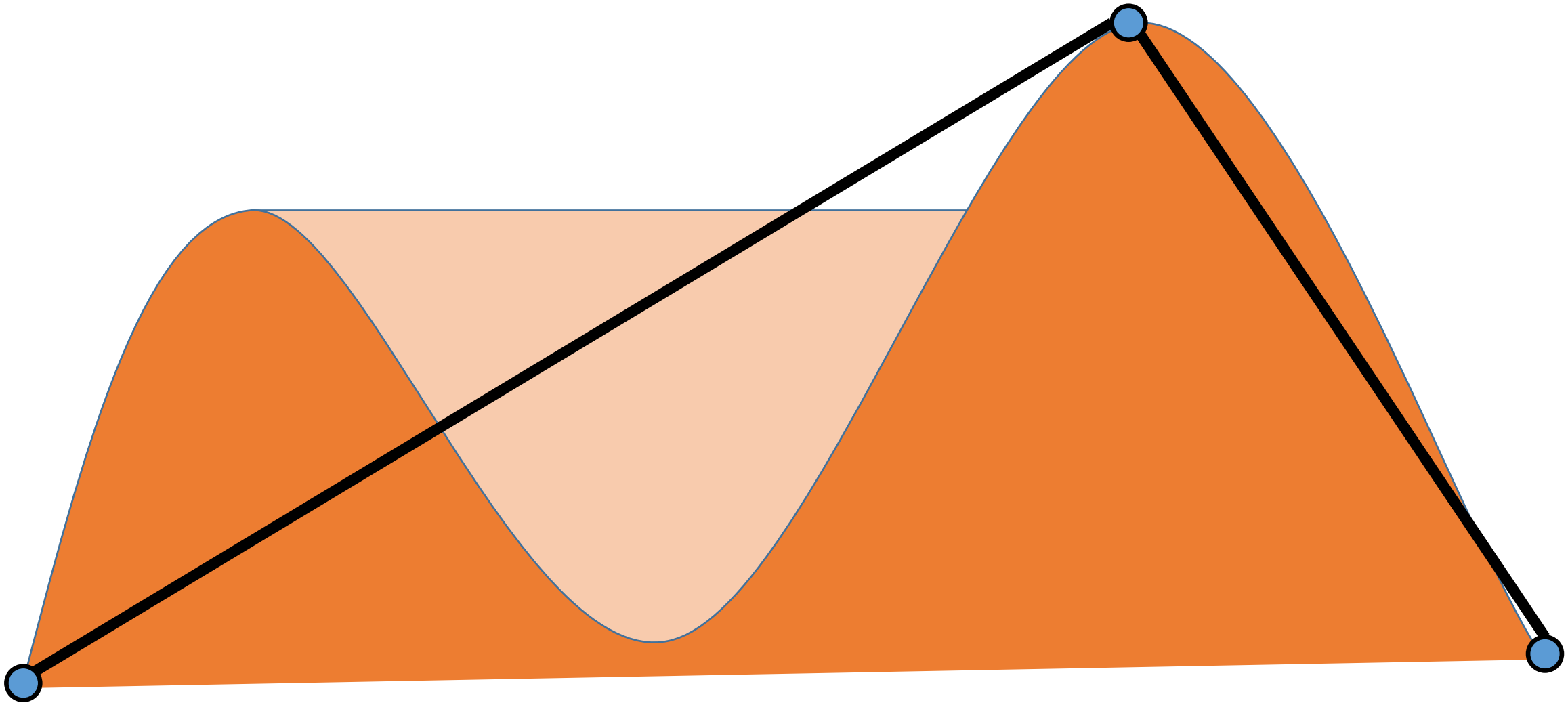
Scalar Field Modification



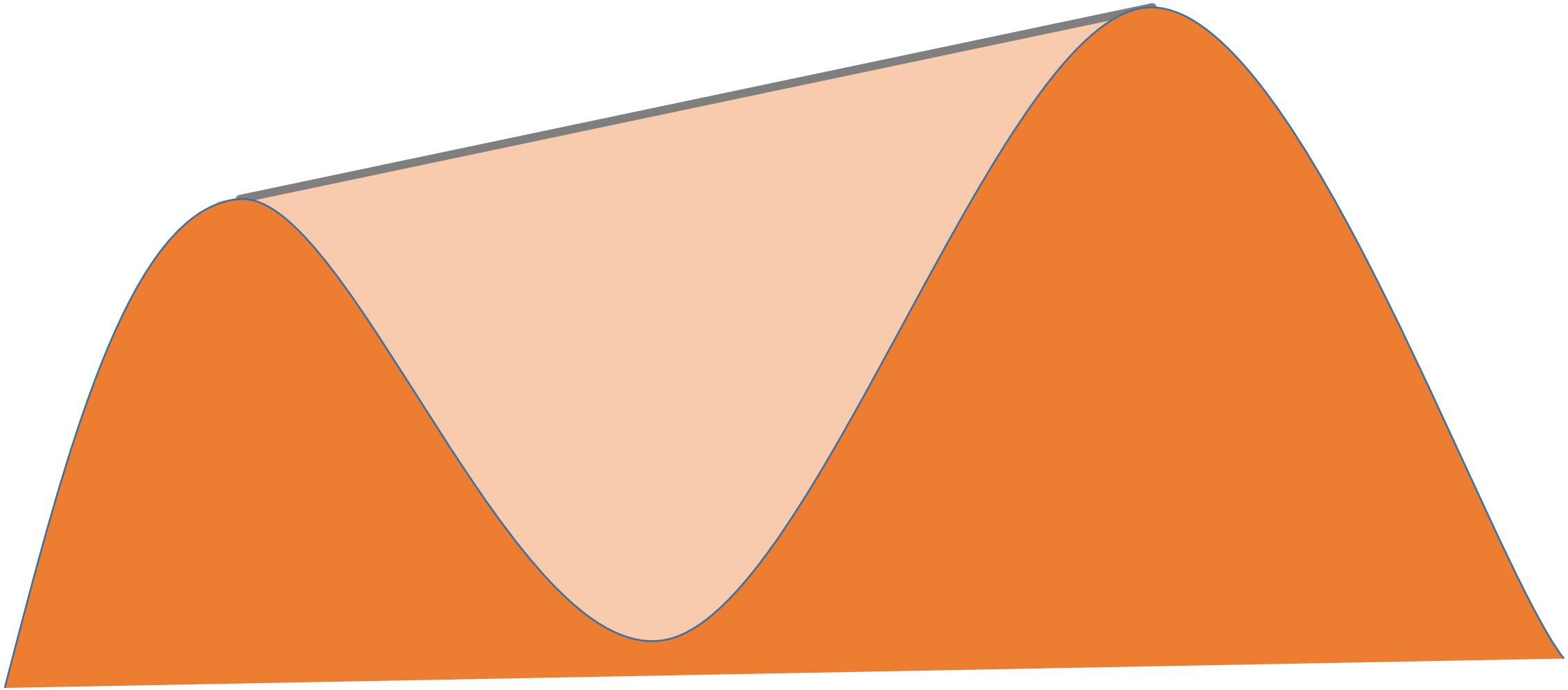
Scalar Field Modification



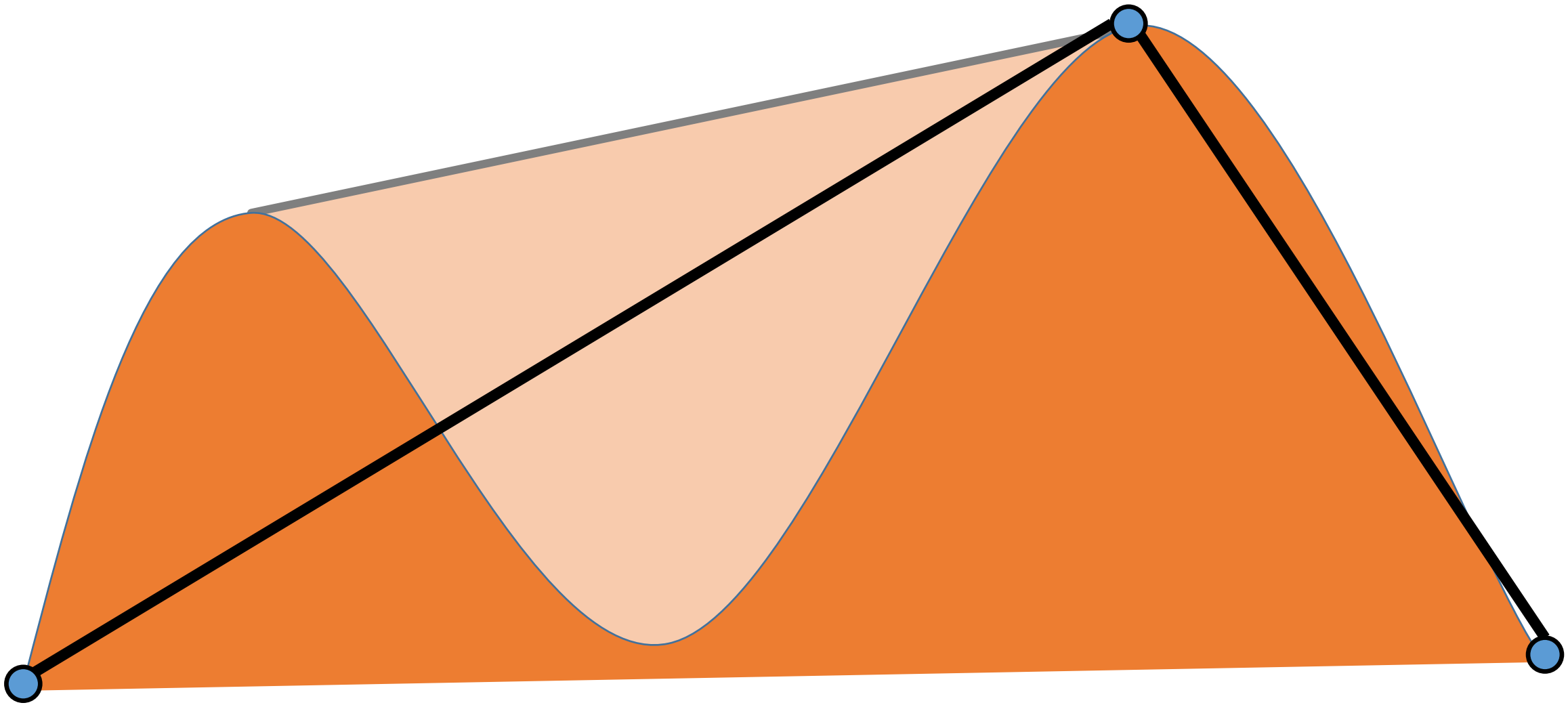
Scalar Field Modification



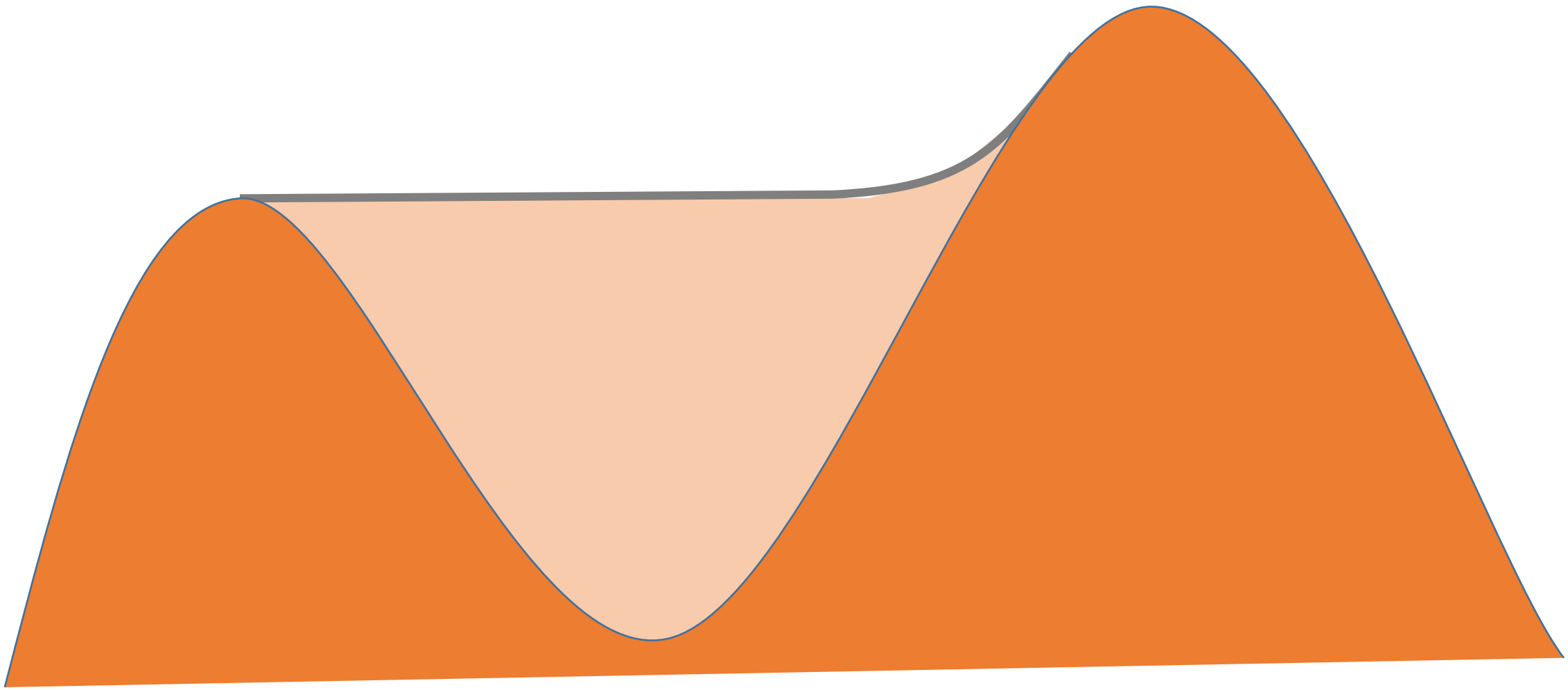
Alternative Simplifications



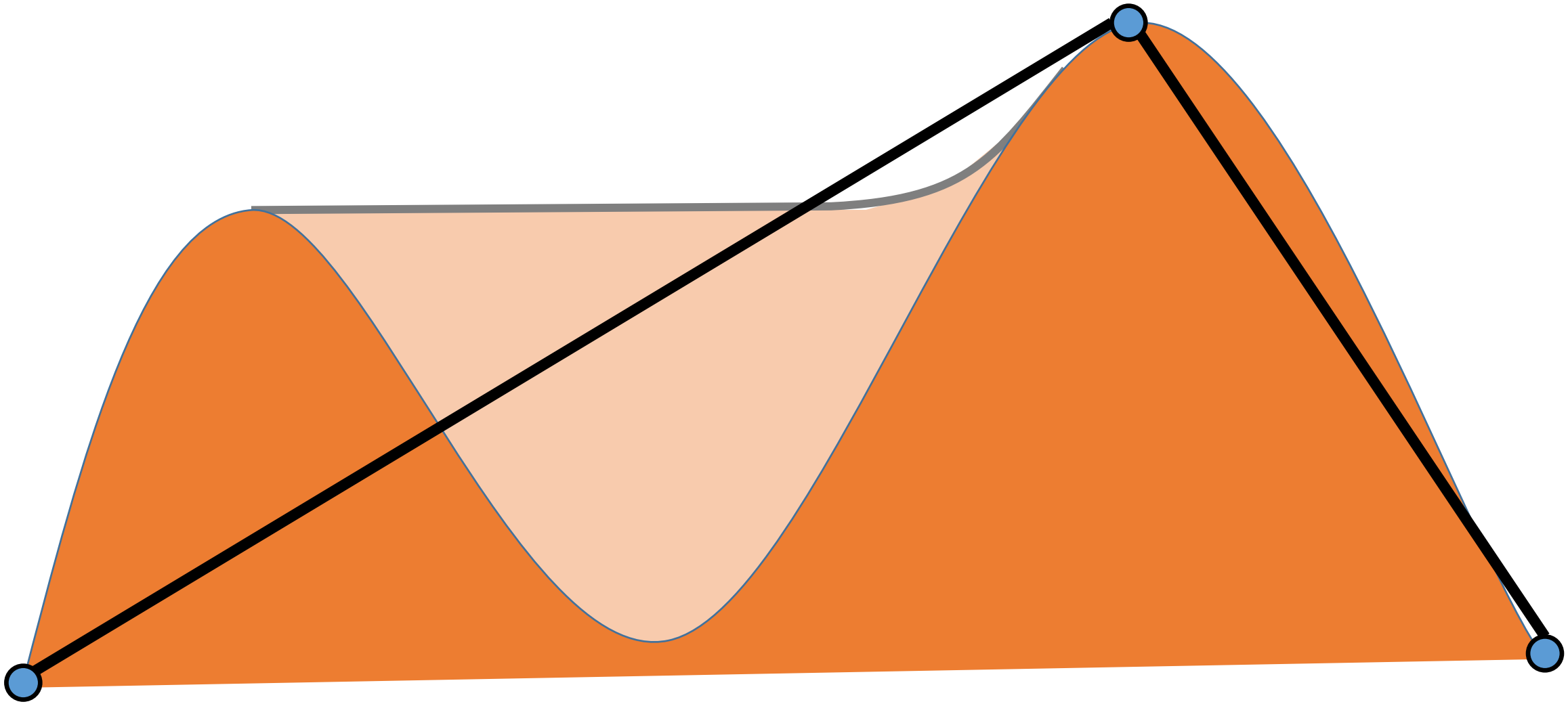
Alternative Simplifications



Alternative Simplifications



Alternative Simplifications



EXAMPLE

Simple Spinning Disk

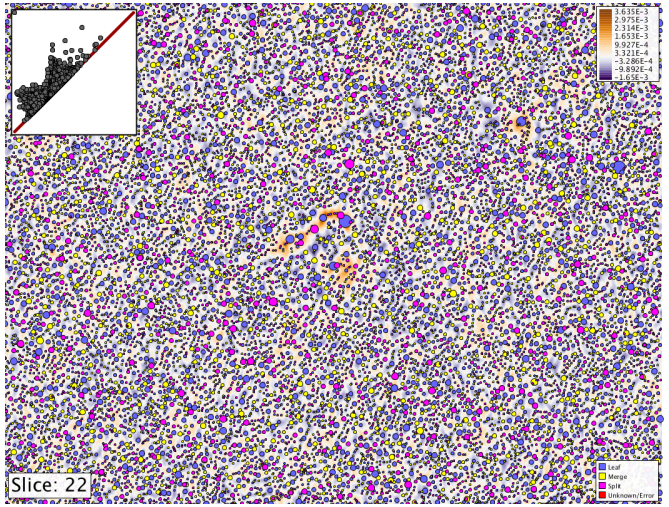
from Anil Seth

Phys. & Astro.

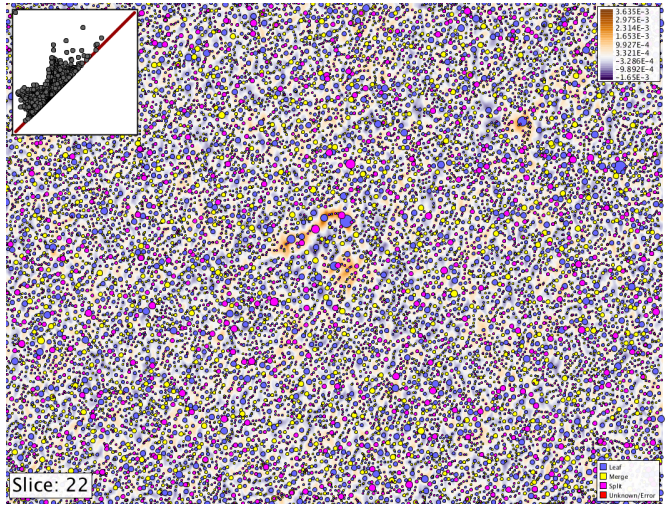
University of Utah

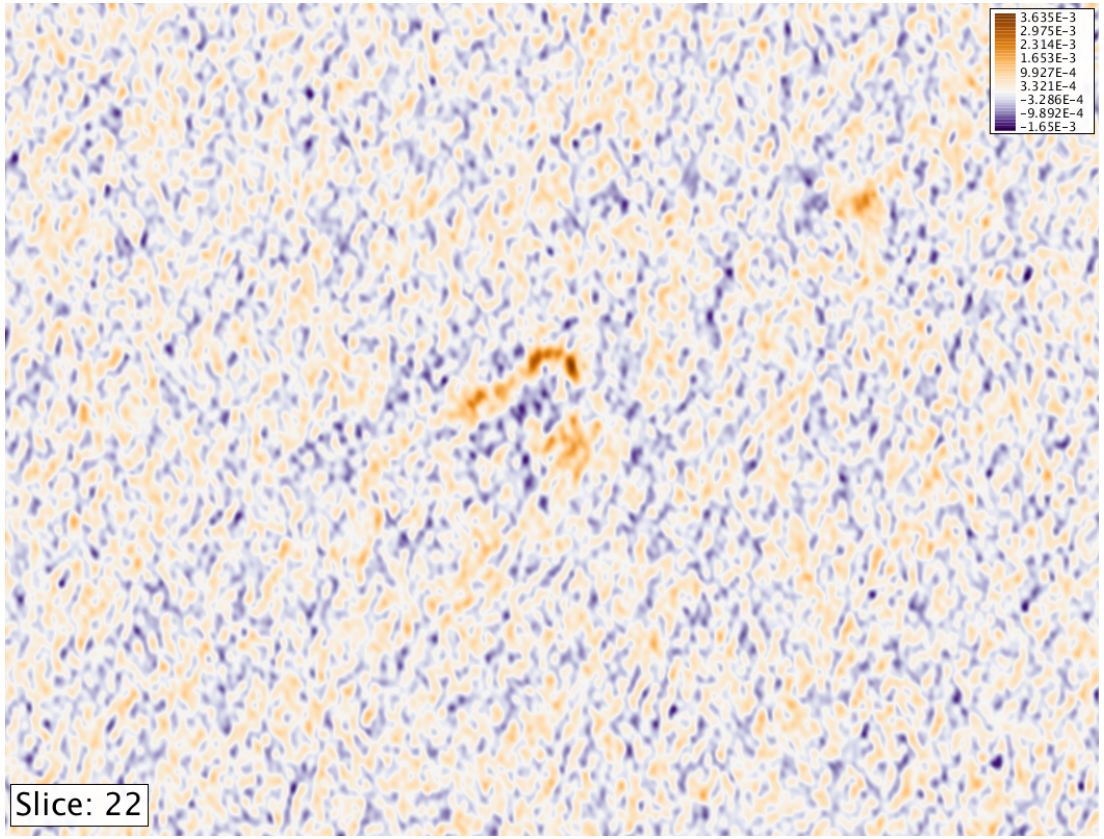
Varying Persistent Simplification

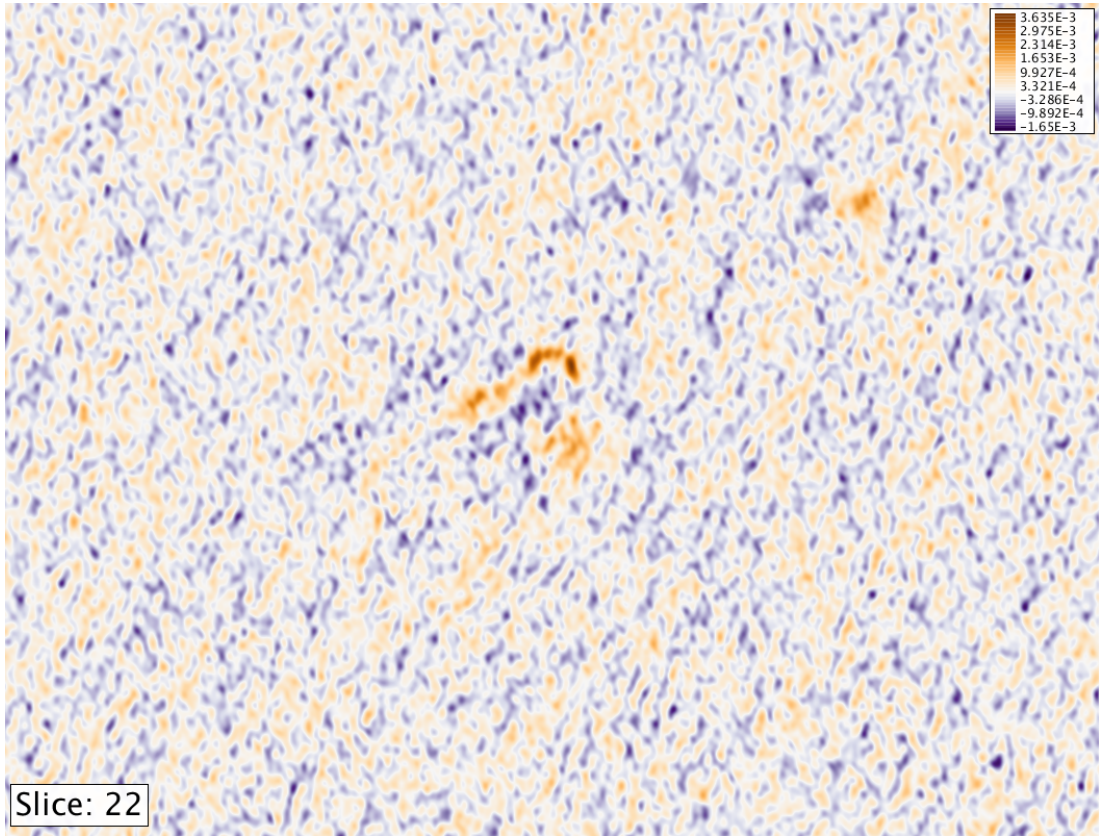
original



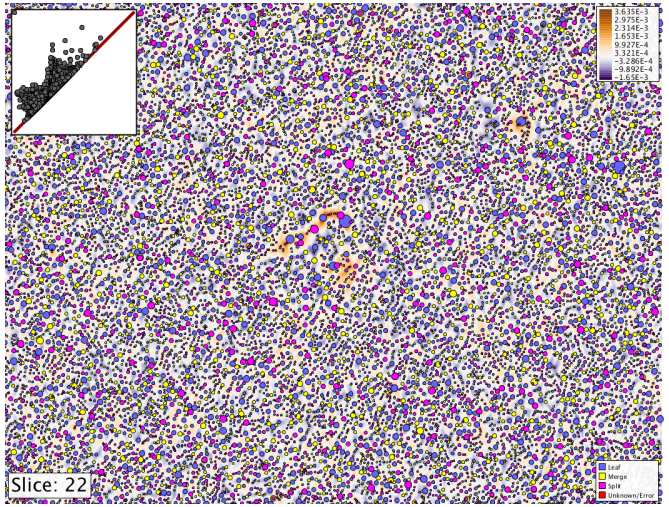
simplified



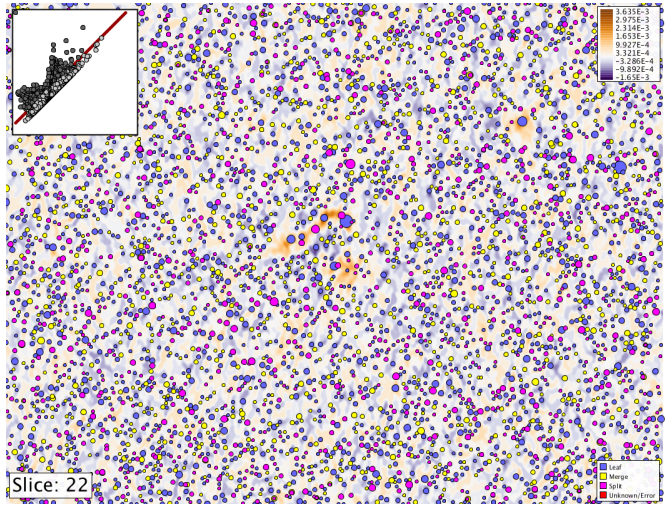


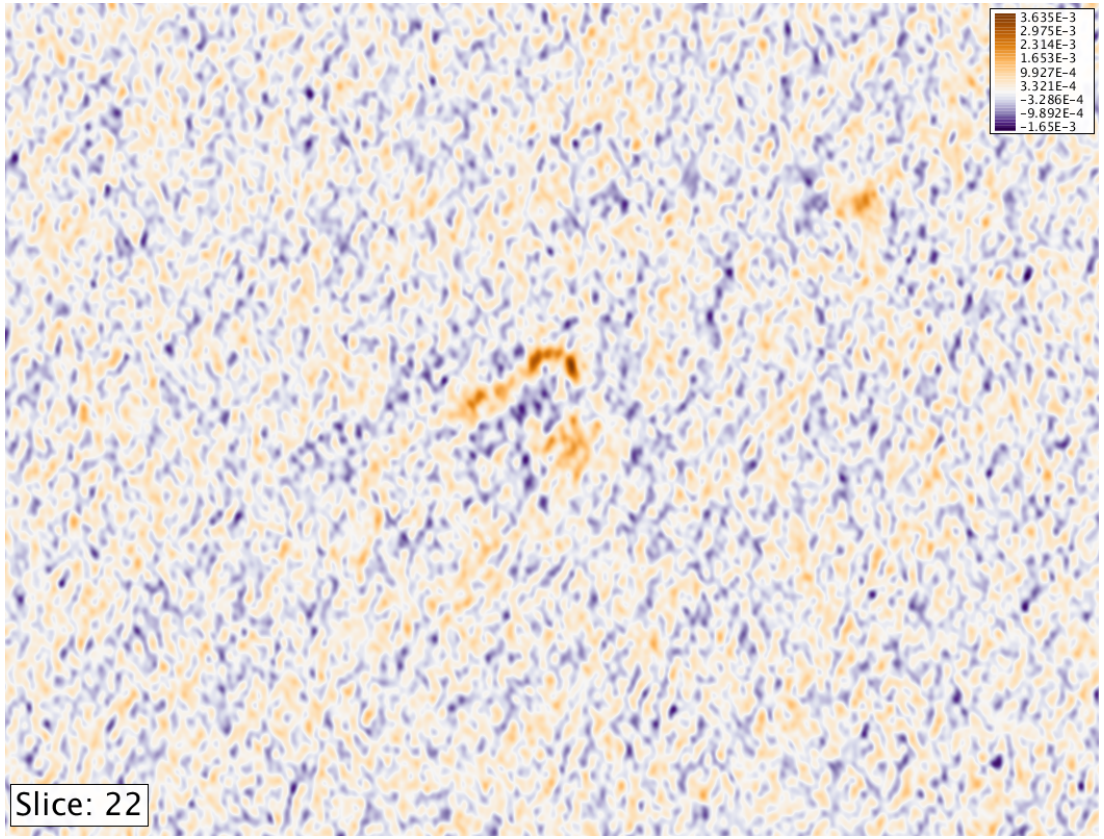


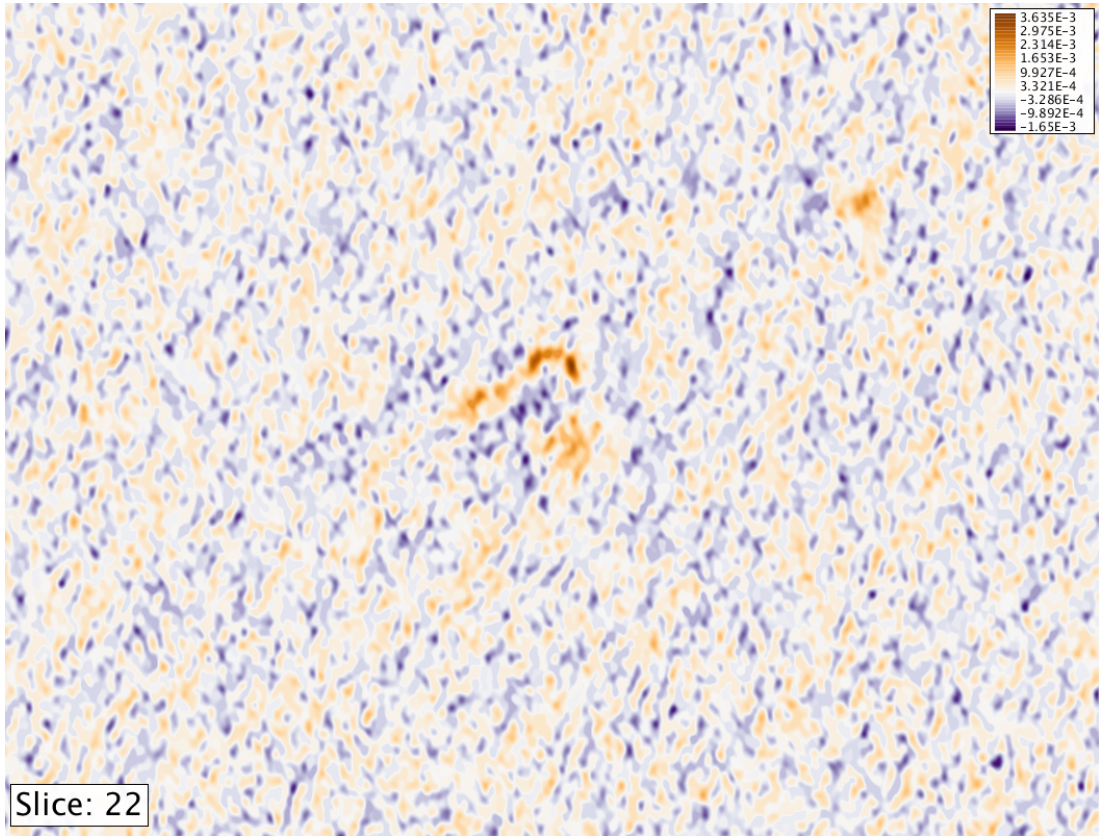
original



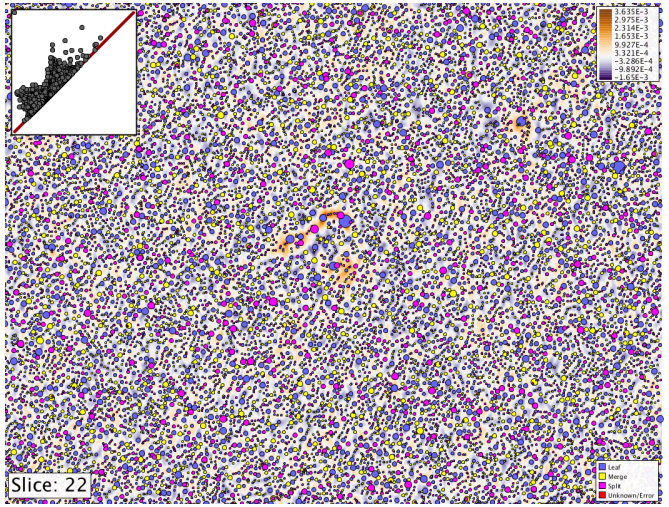
simplified



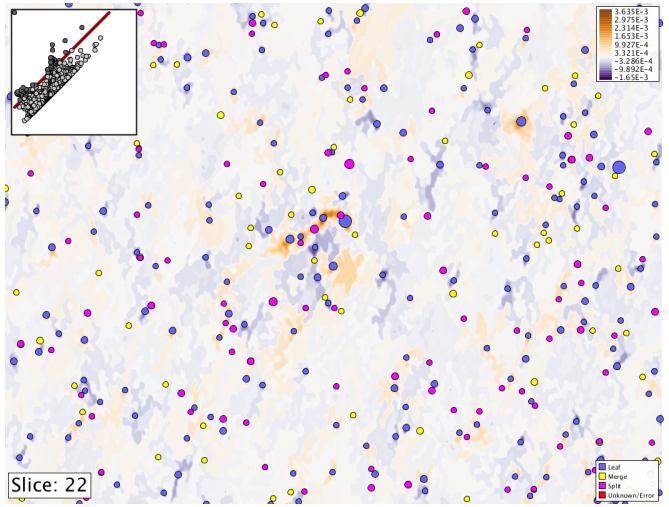


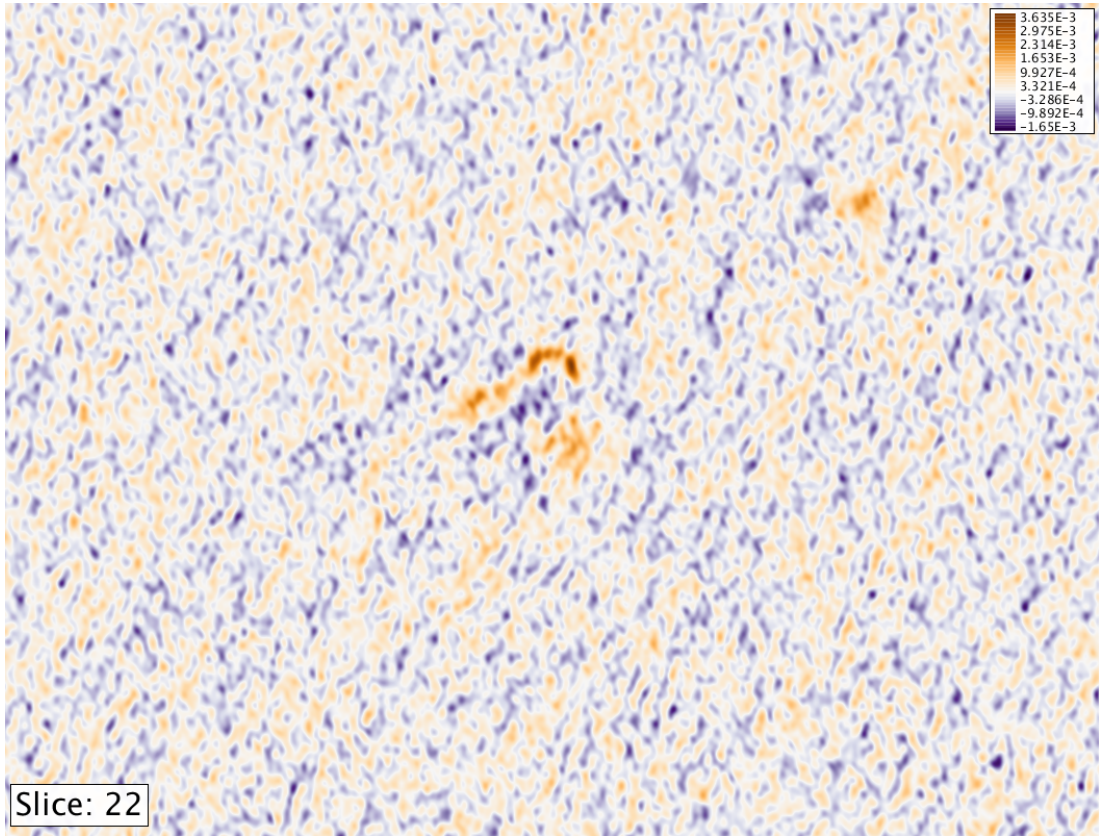


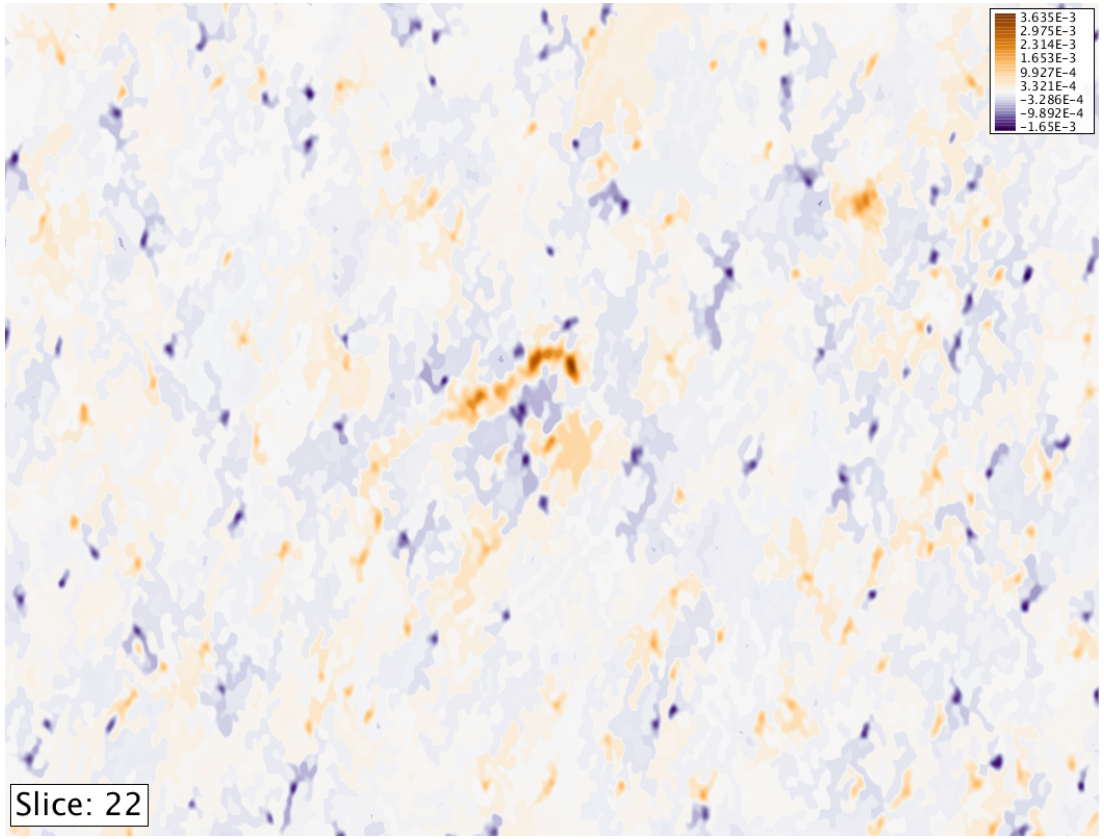
original



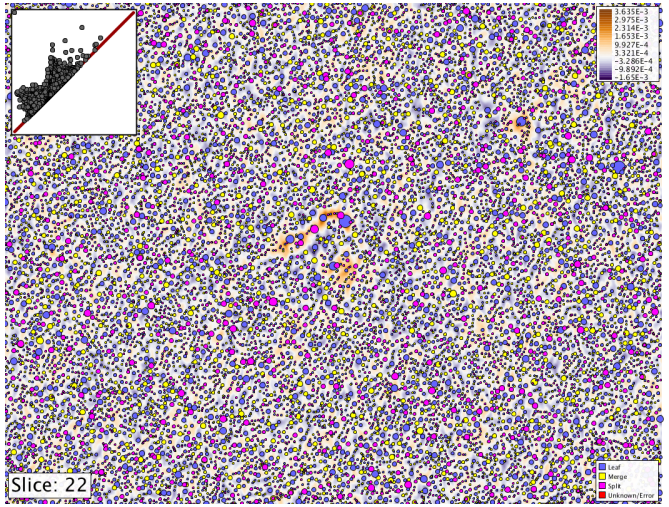
simplified



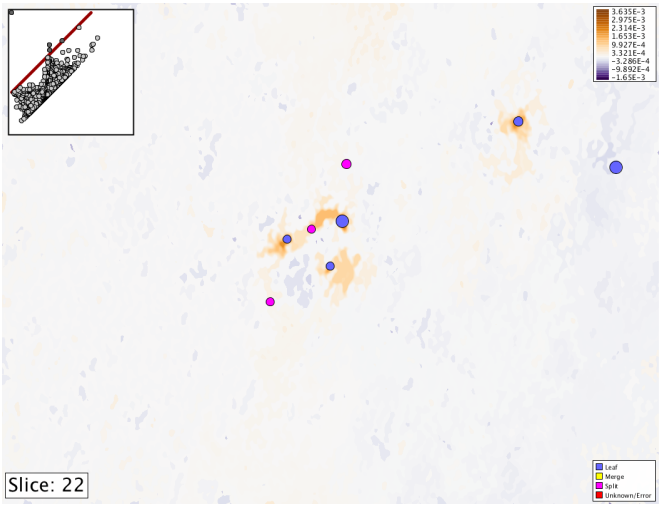


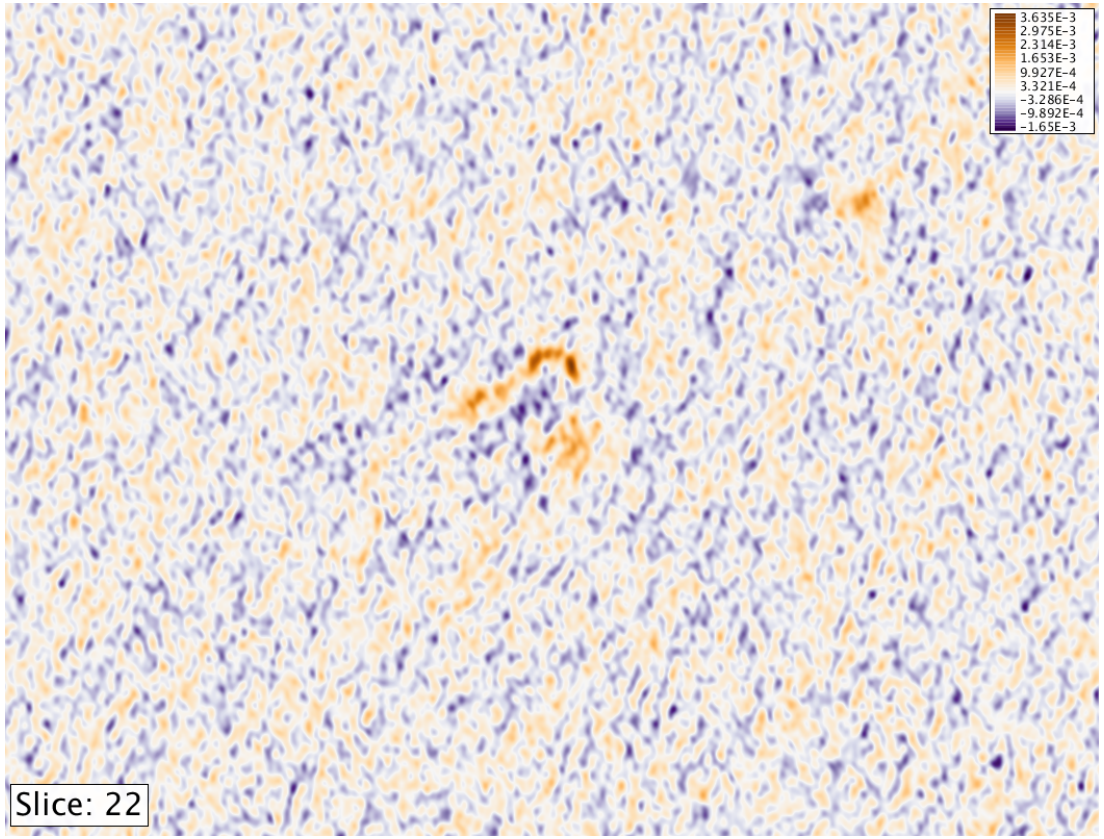


original

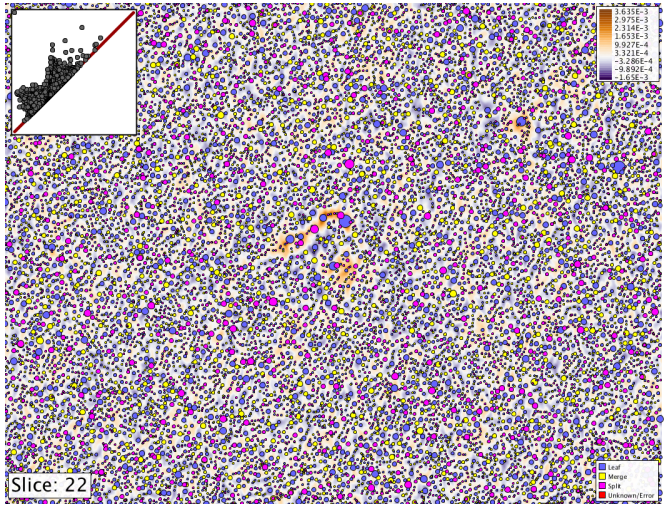


simplified

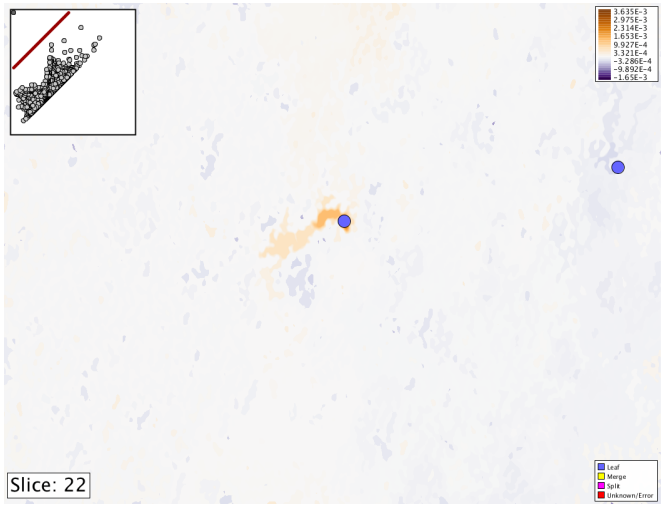


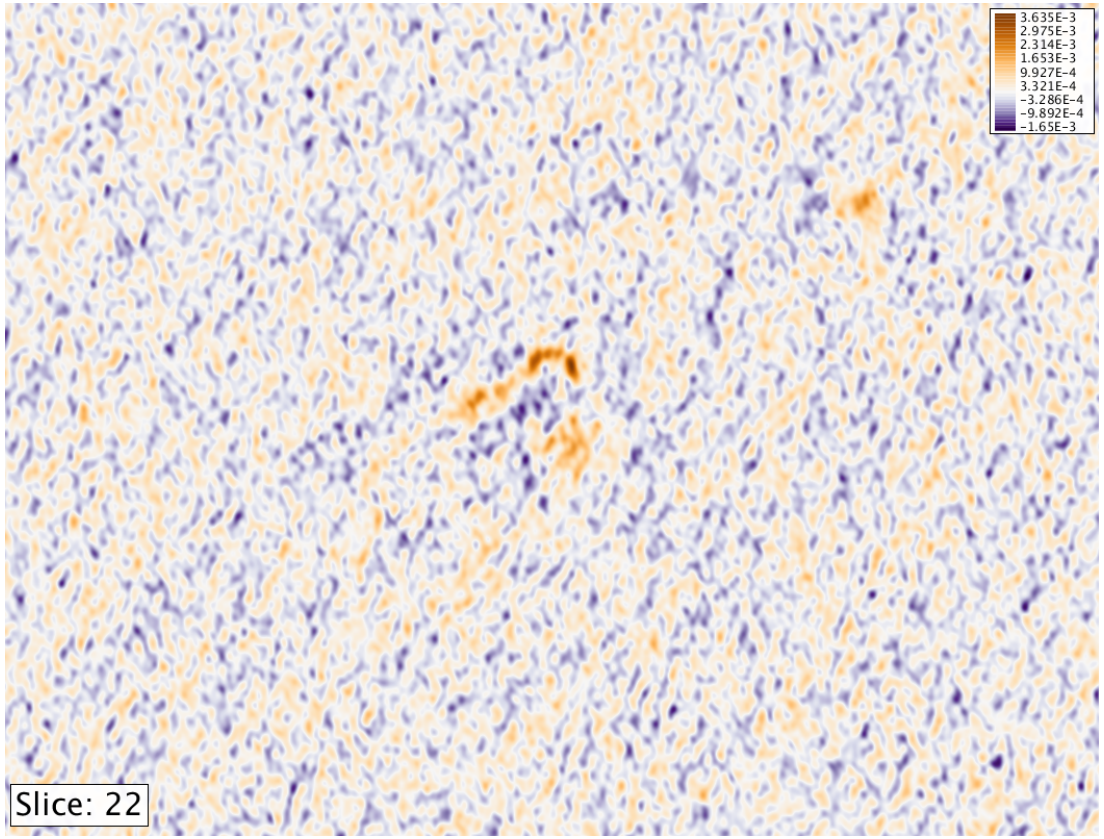


original



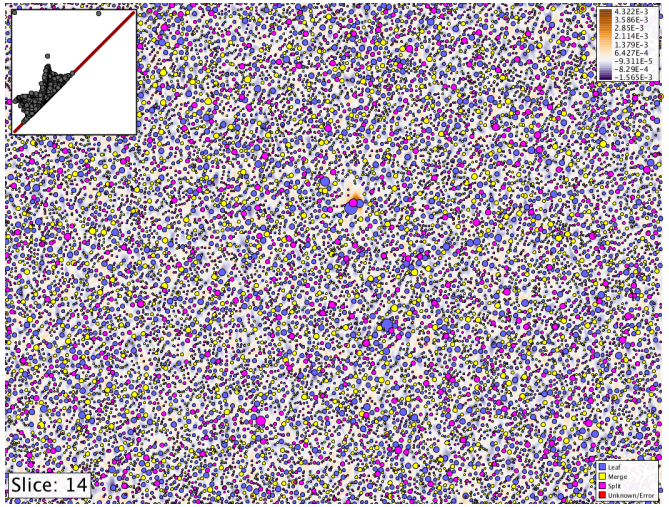
simplified



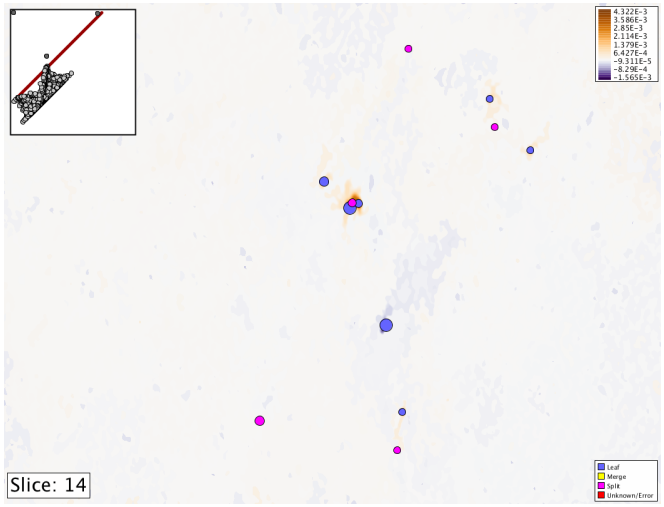


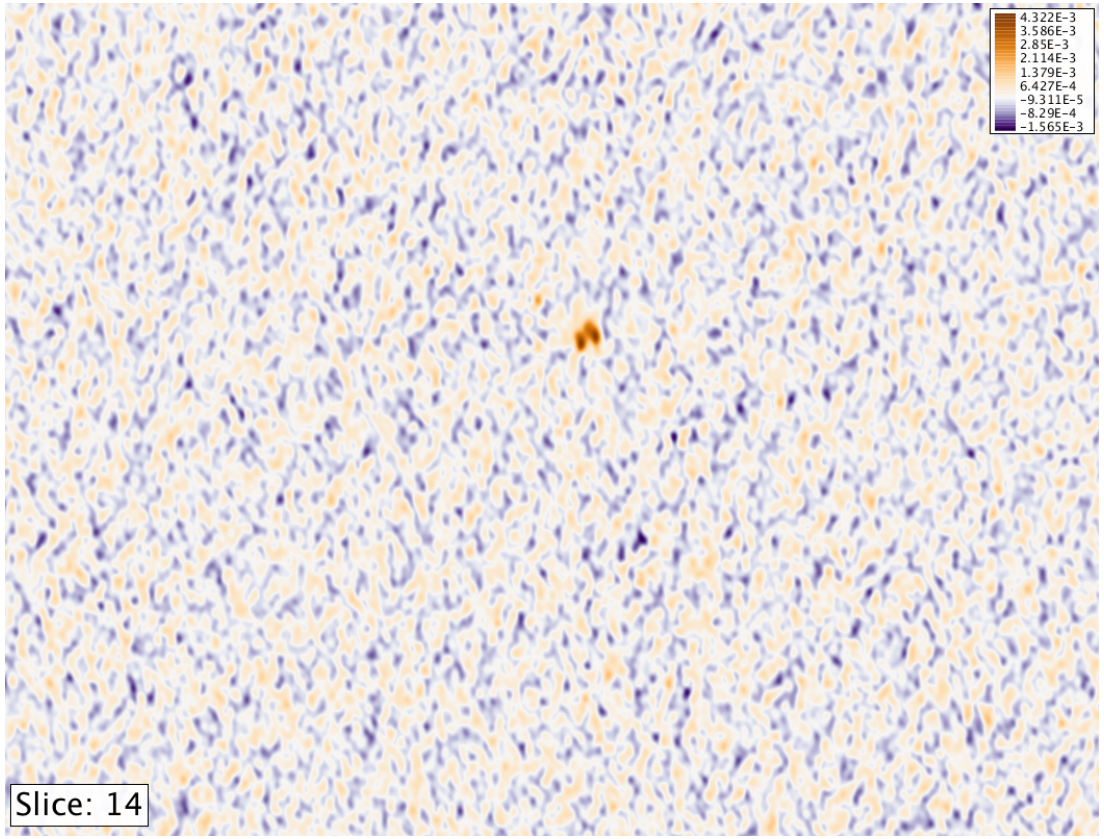
RUNNING THROUGH SLICES

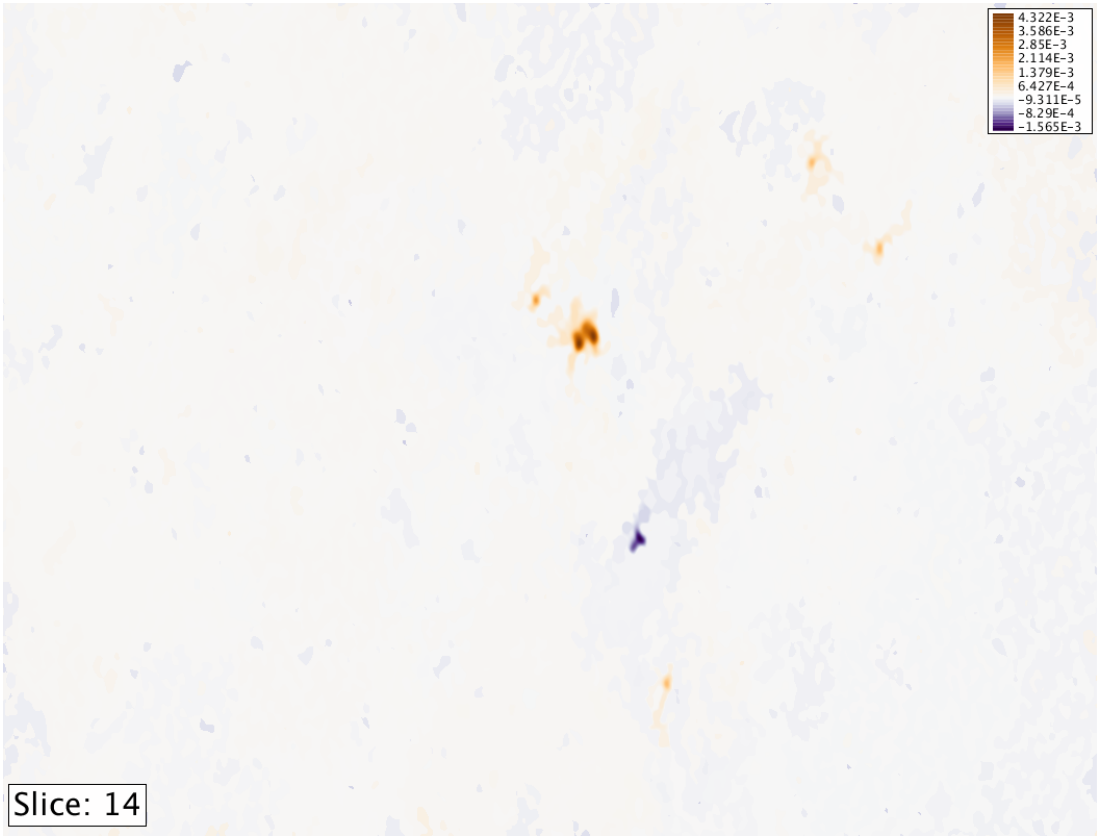
original



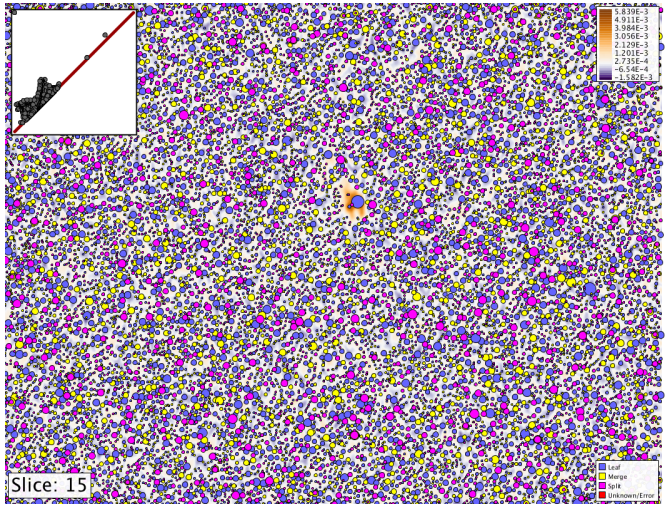
simplified



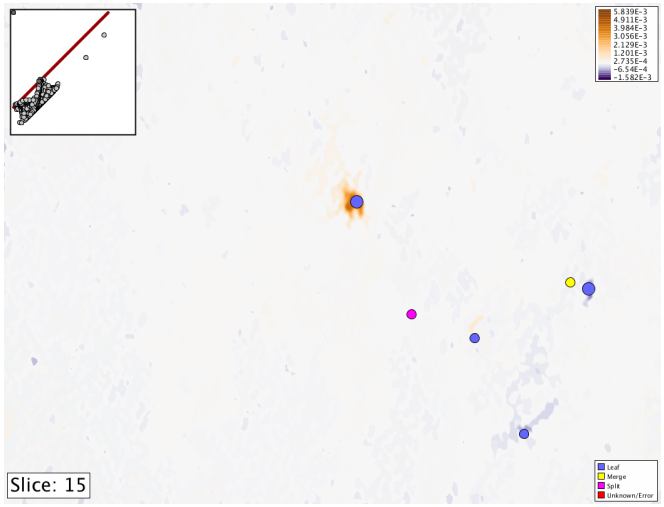


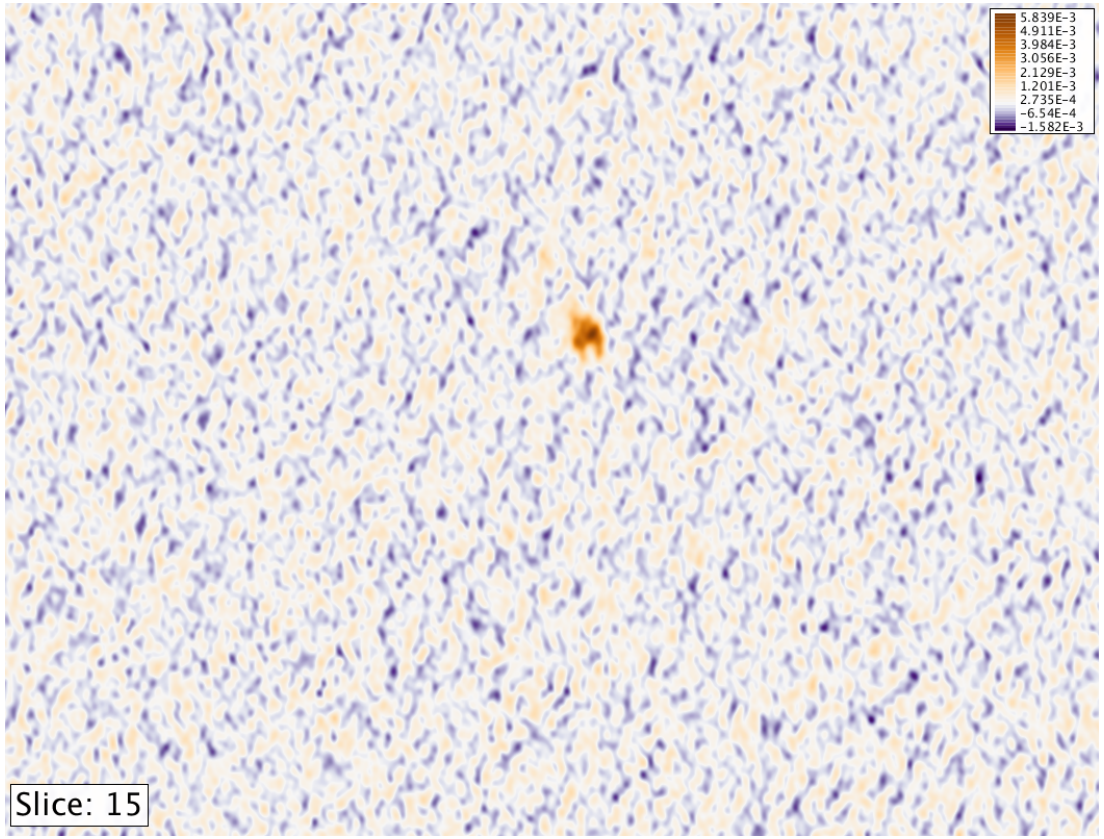


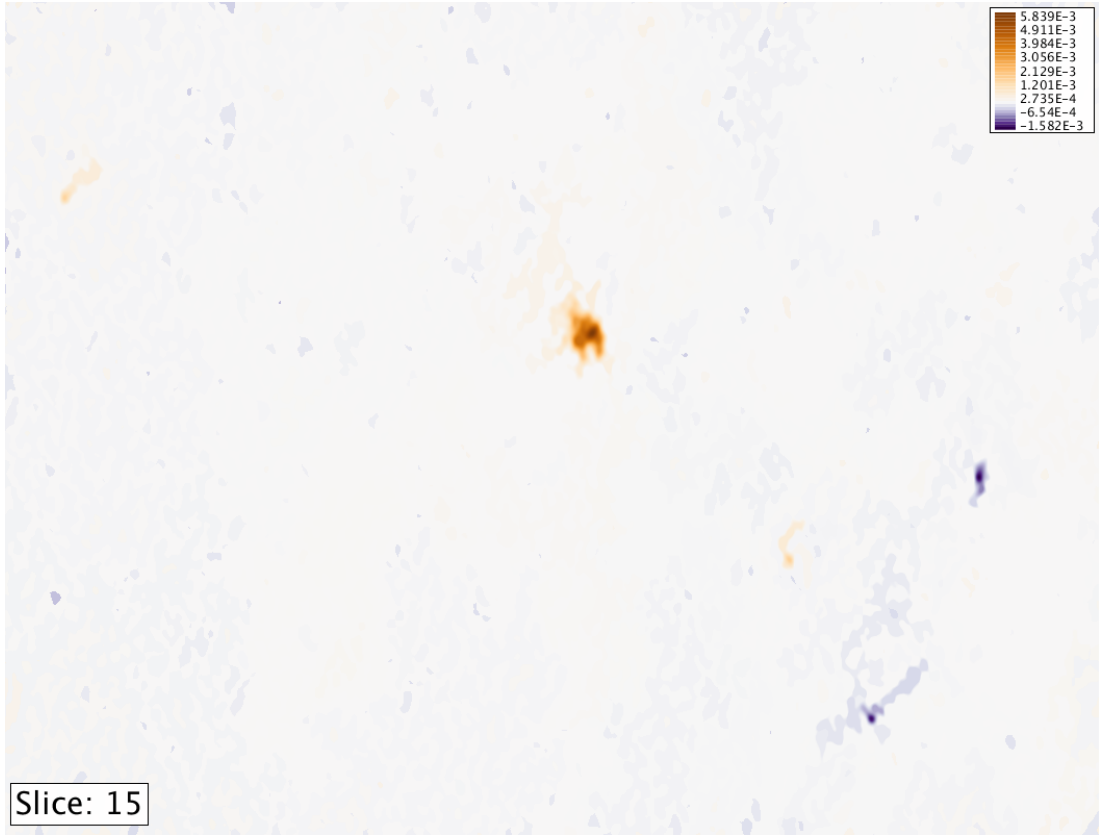
original



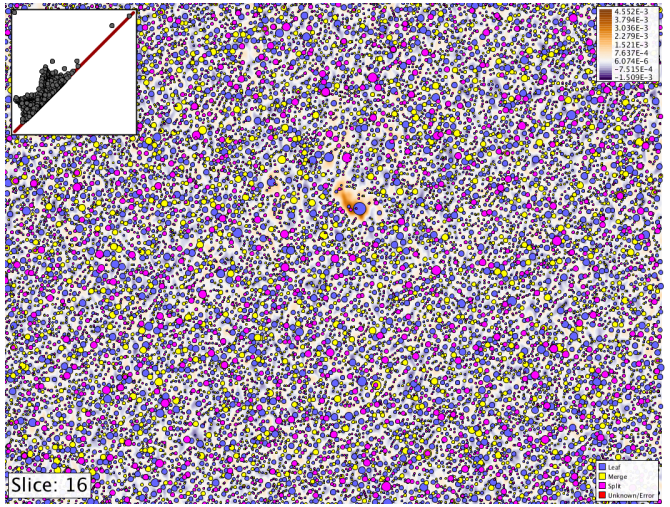
simplified



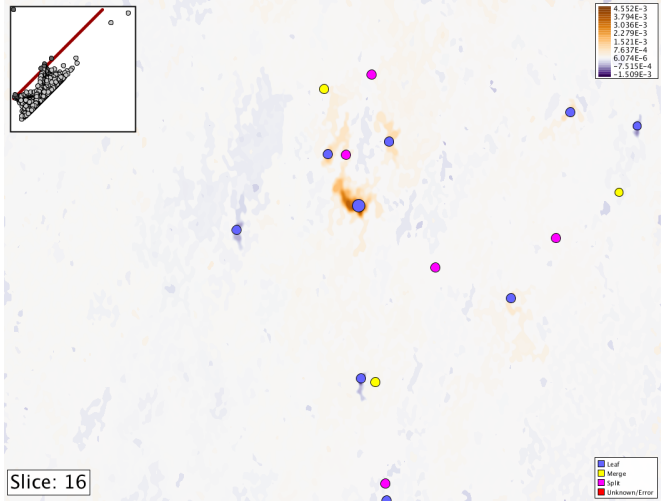


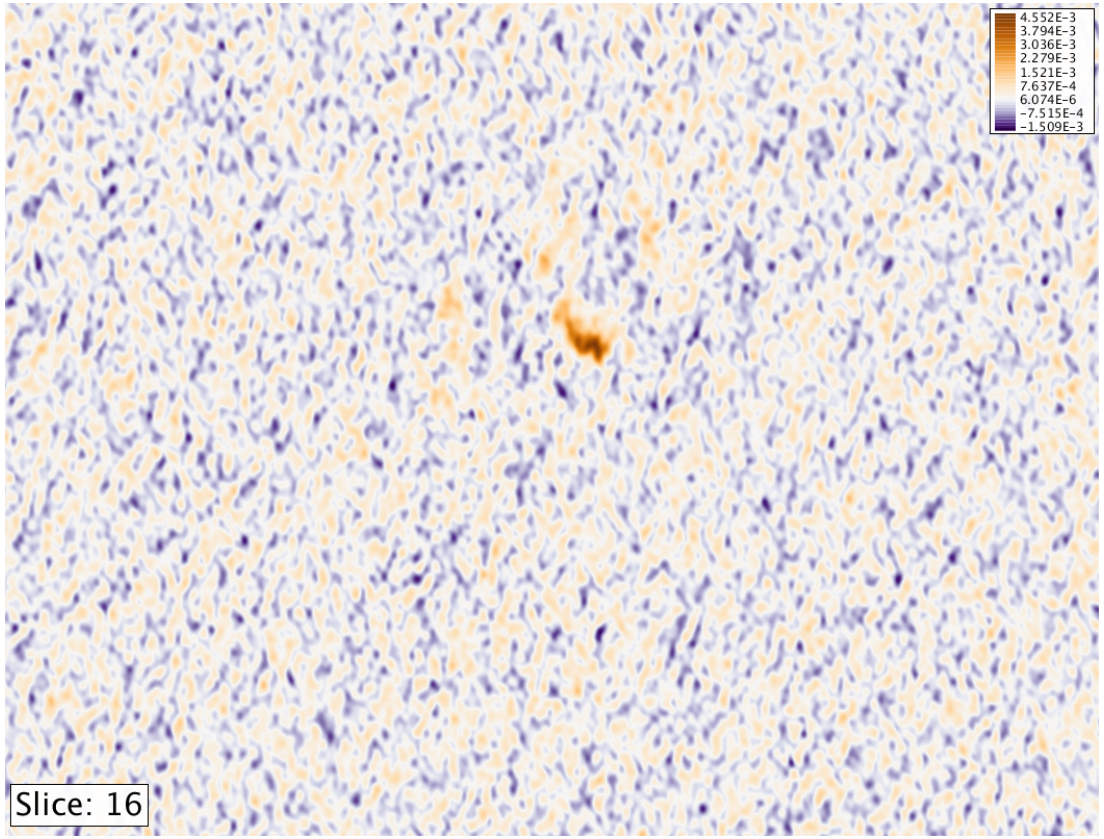


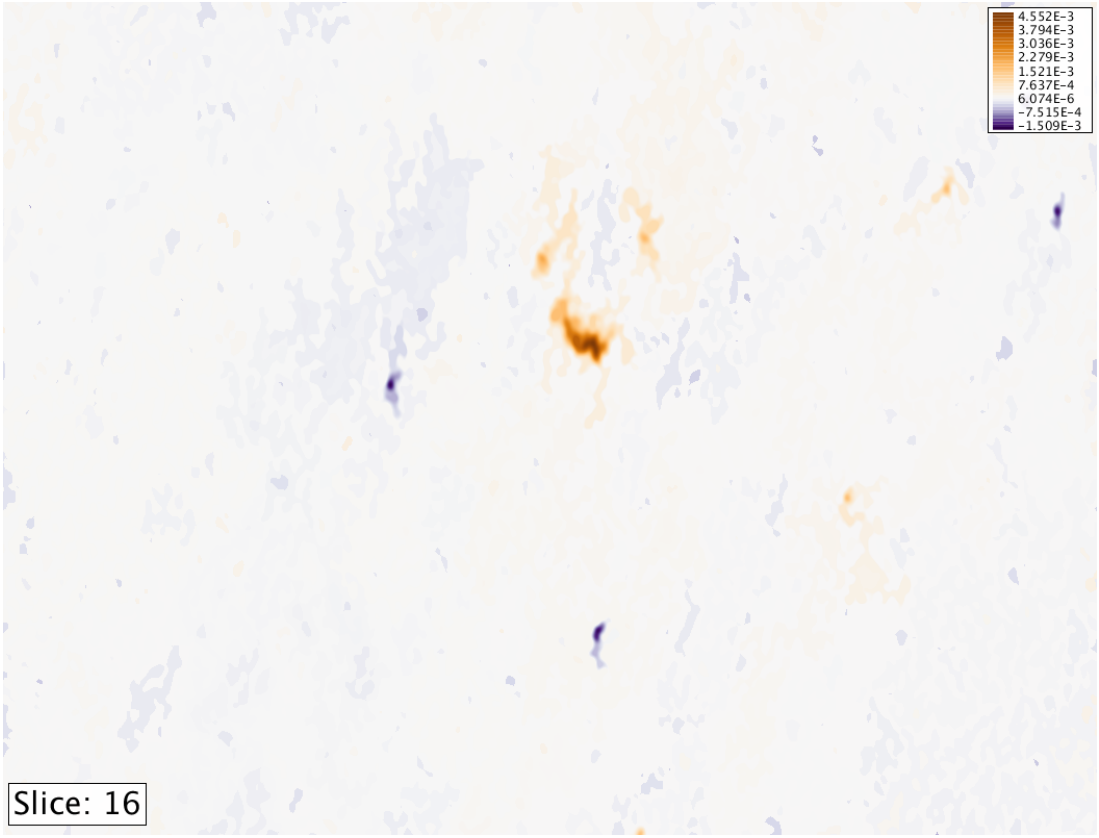
original



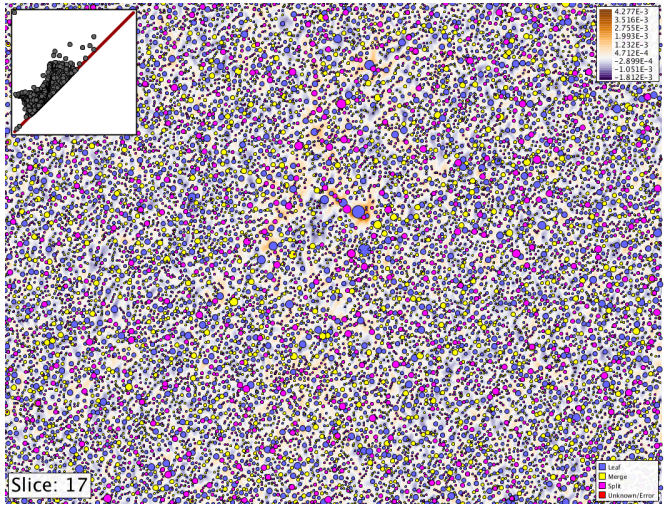
simplified



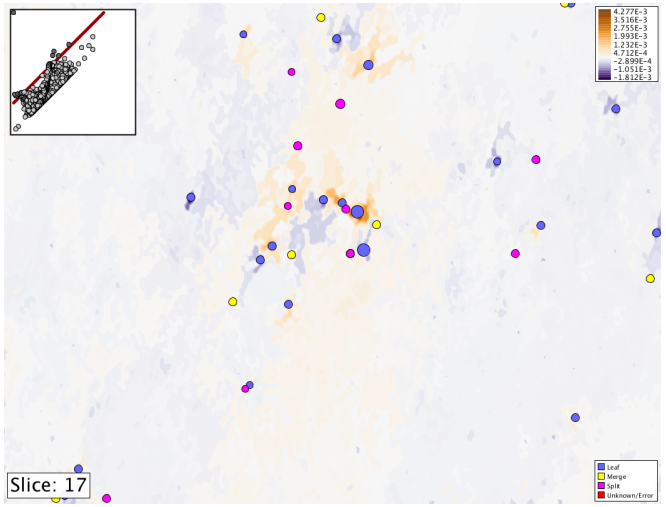


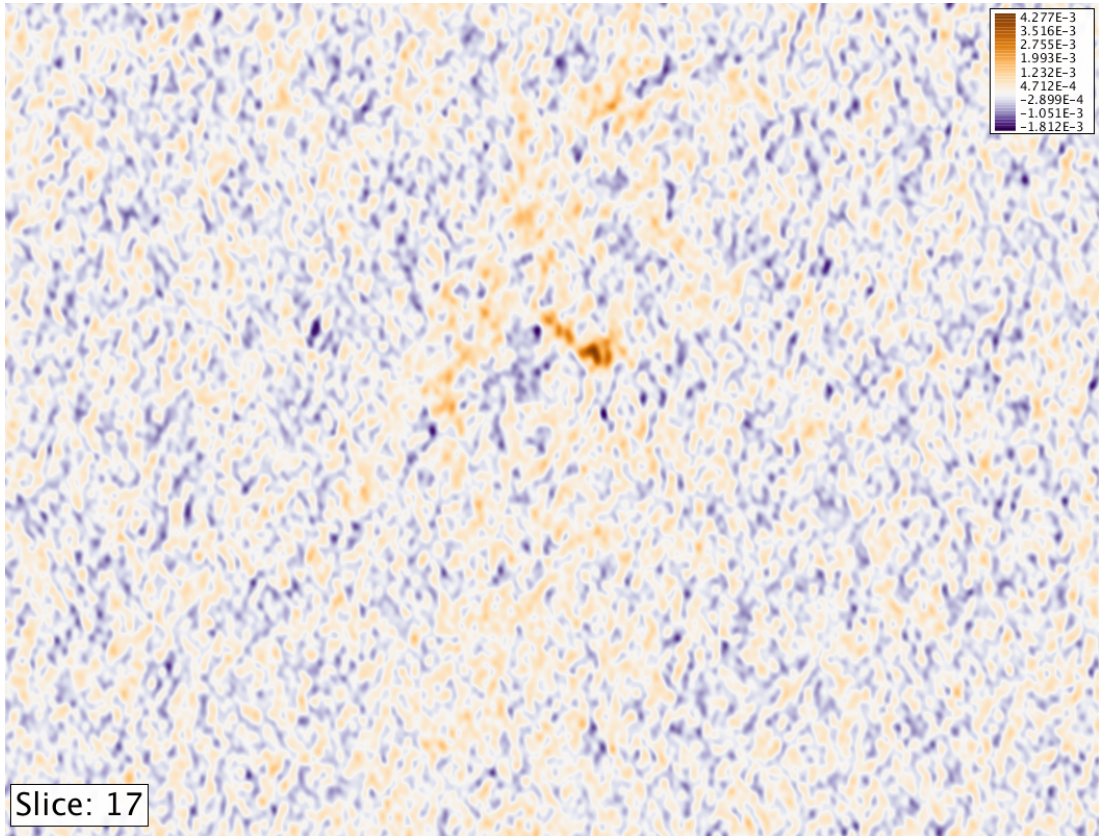


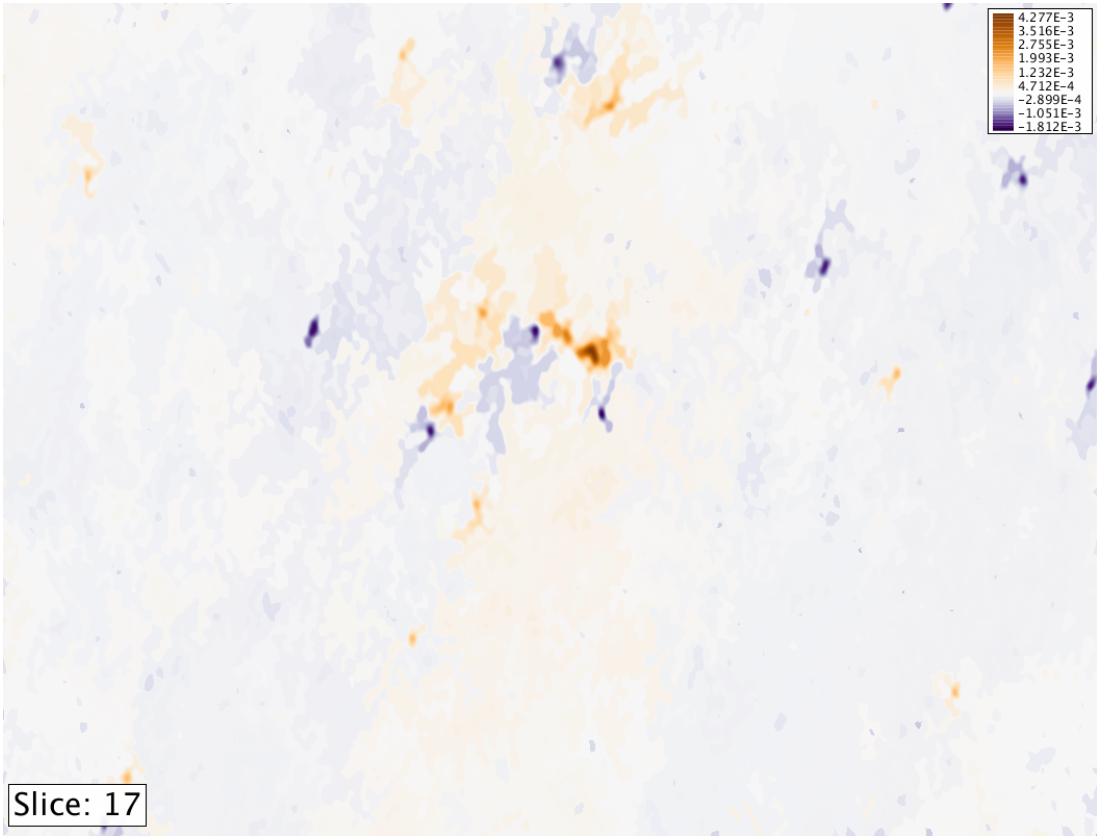
original



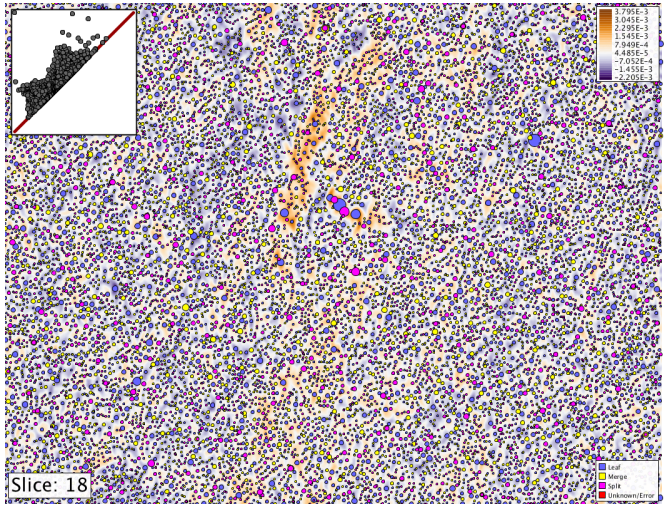
simplified



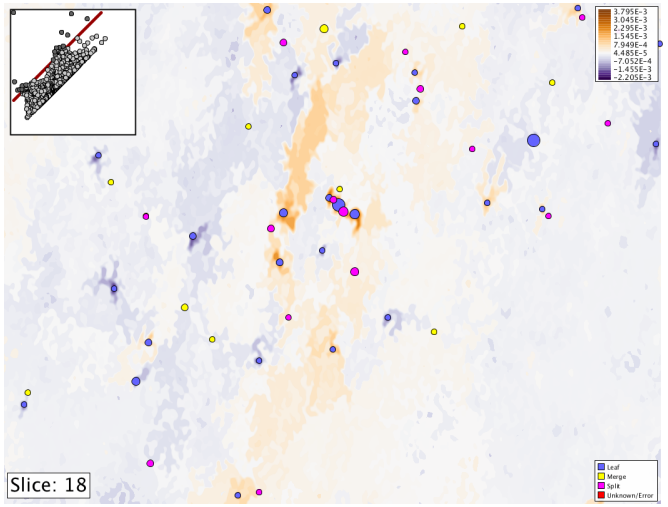


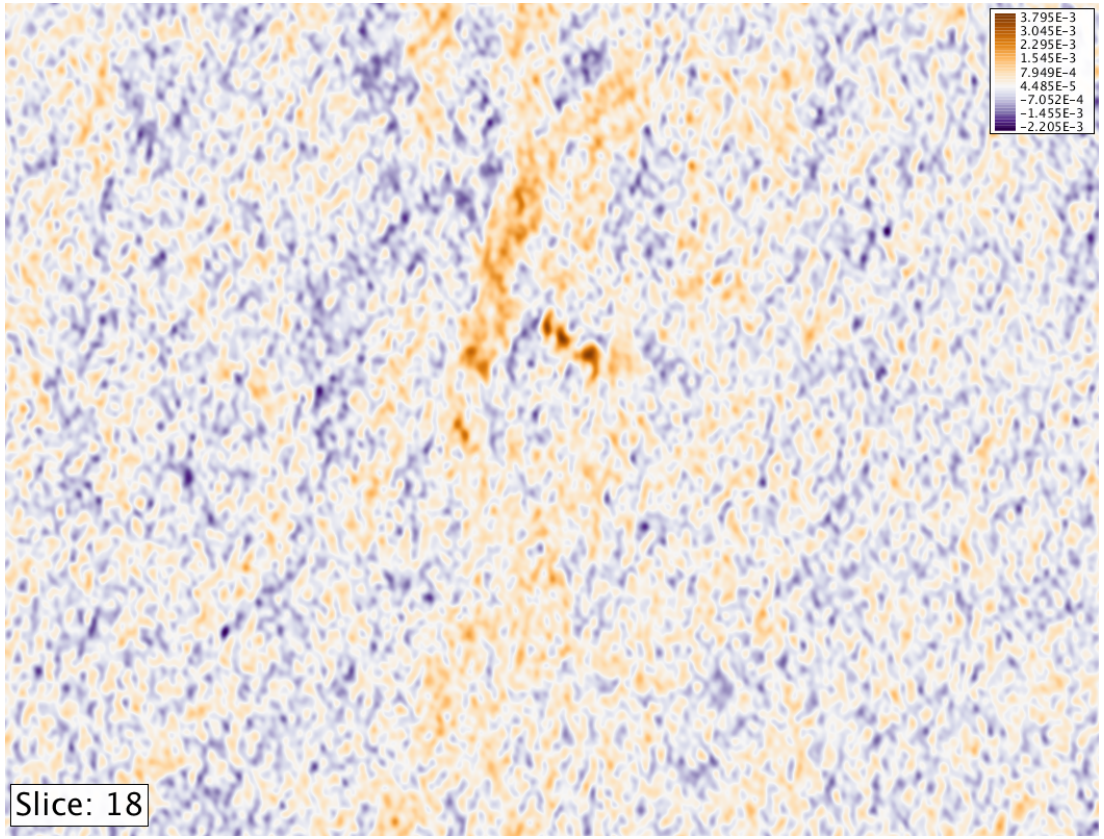


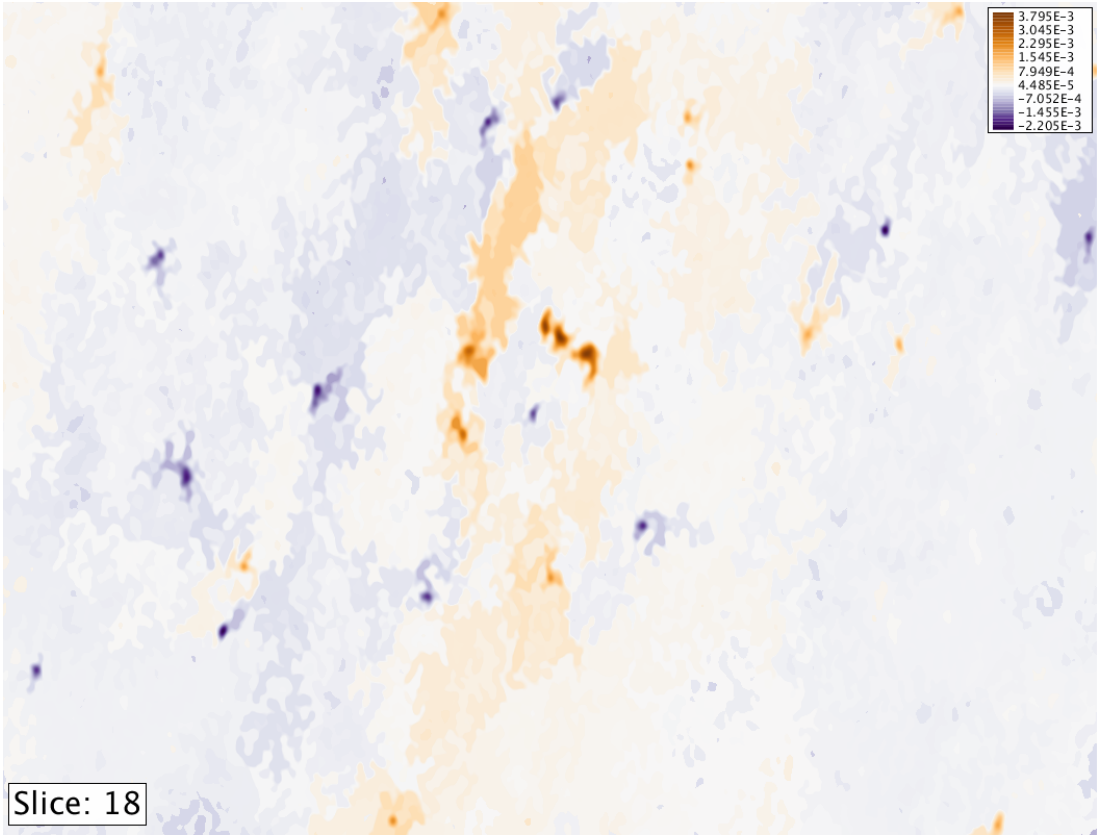
original



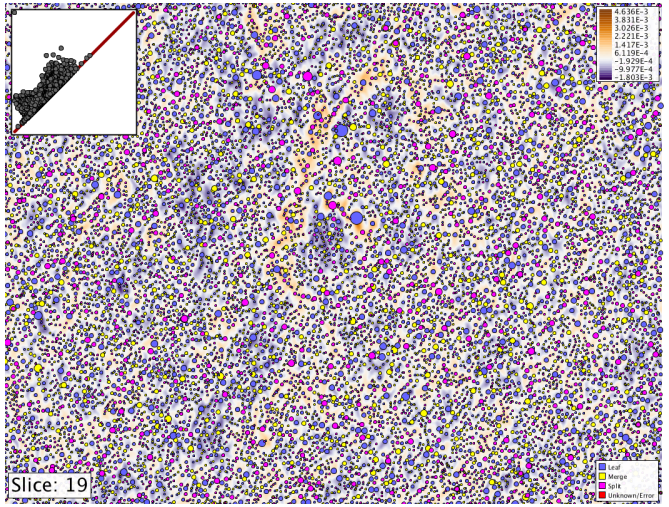
simplified



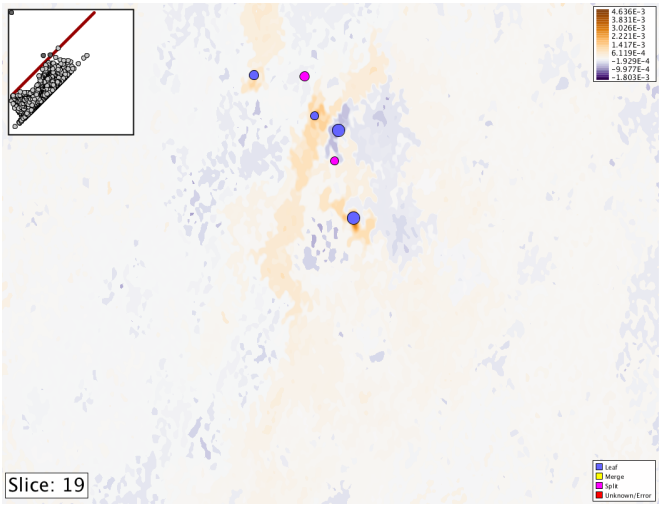


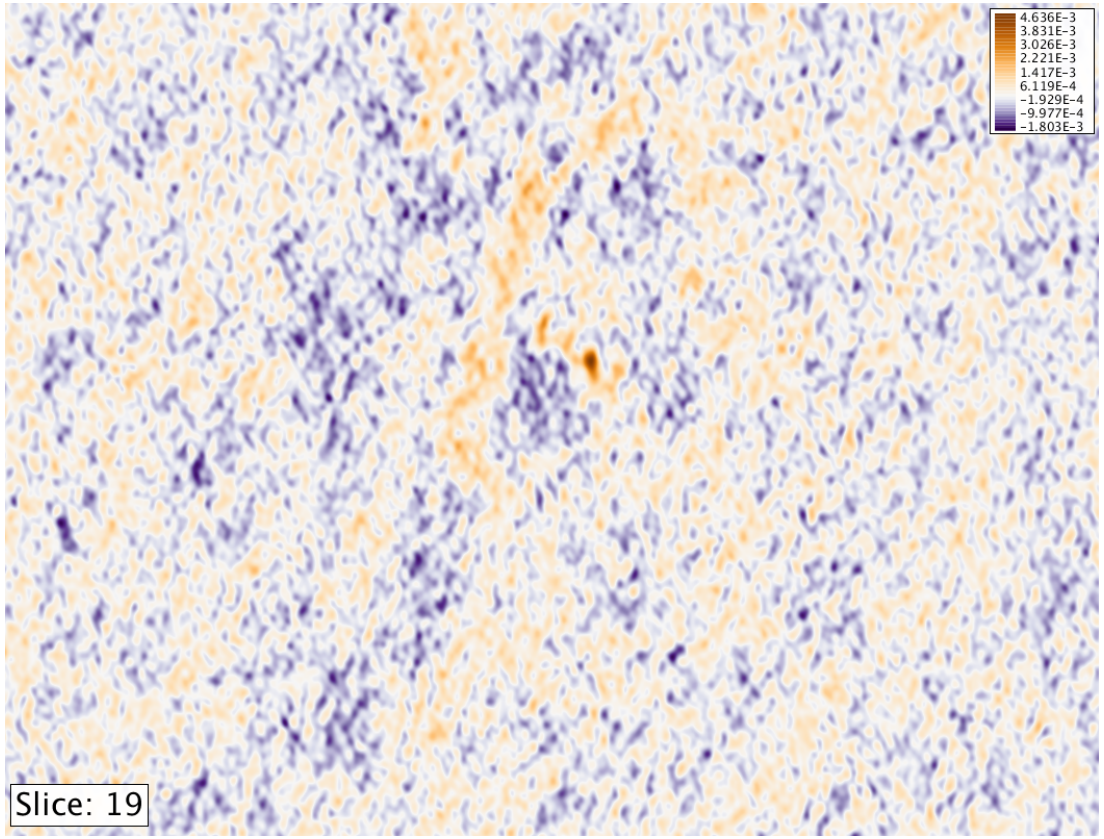


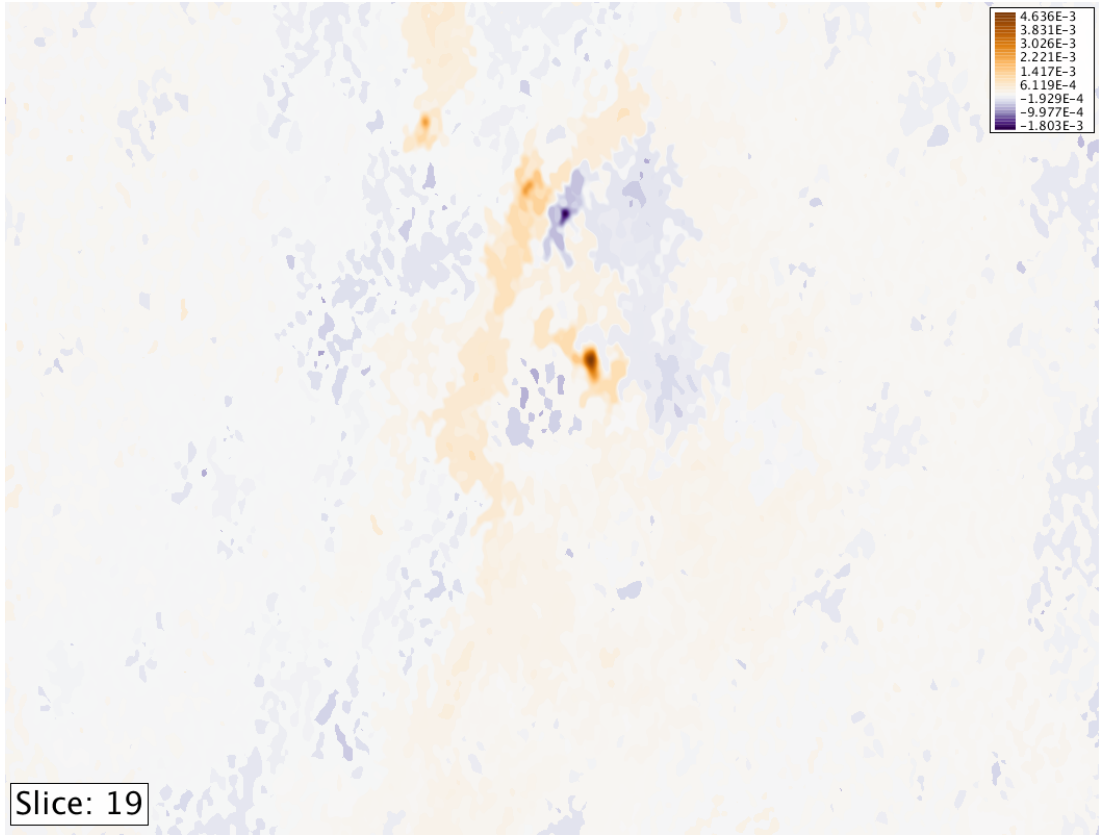
original



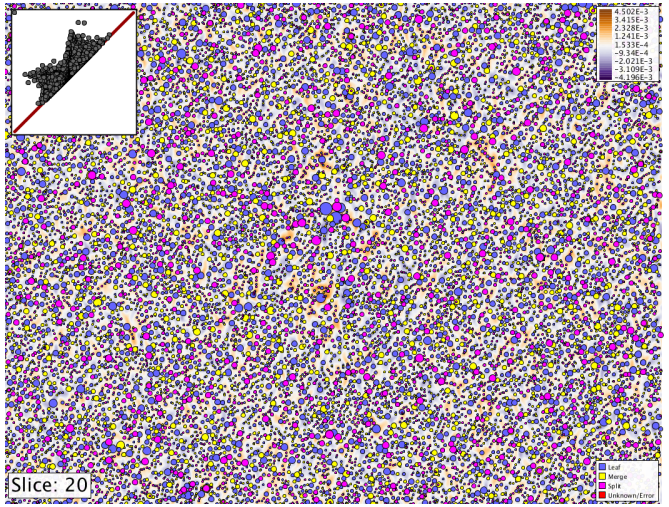
simplified



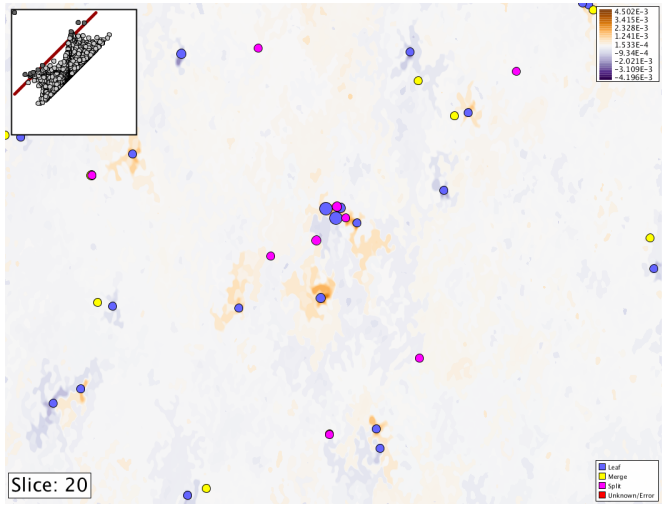


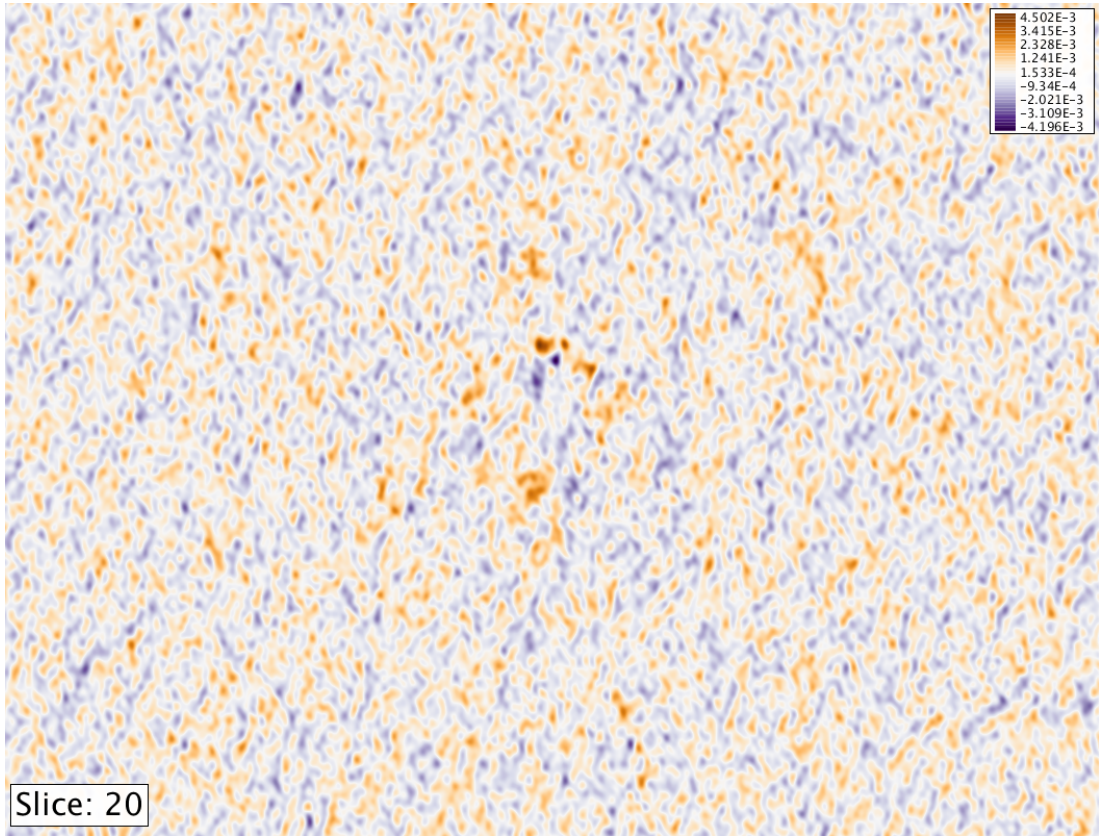


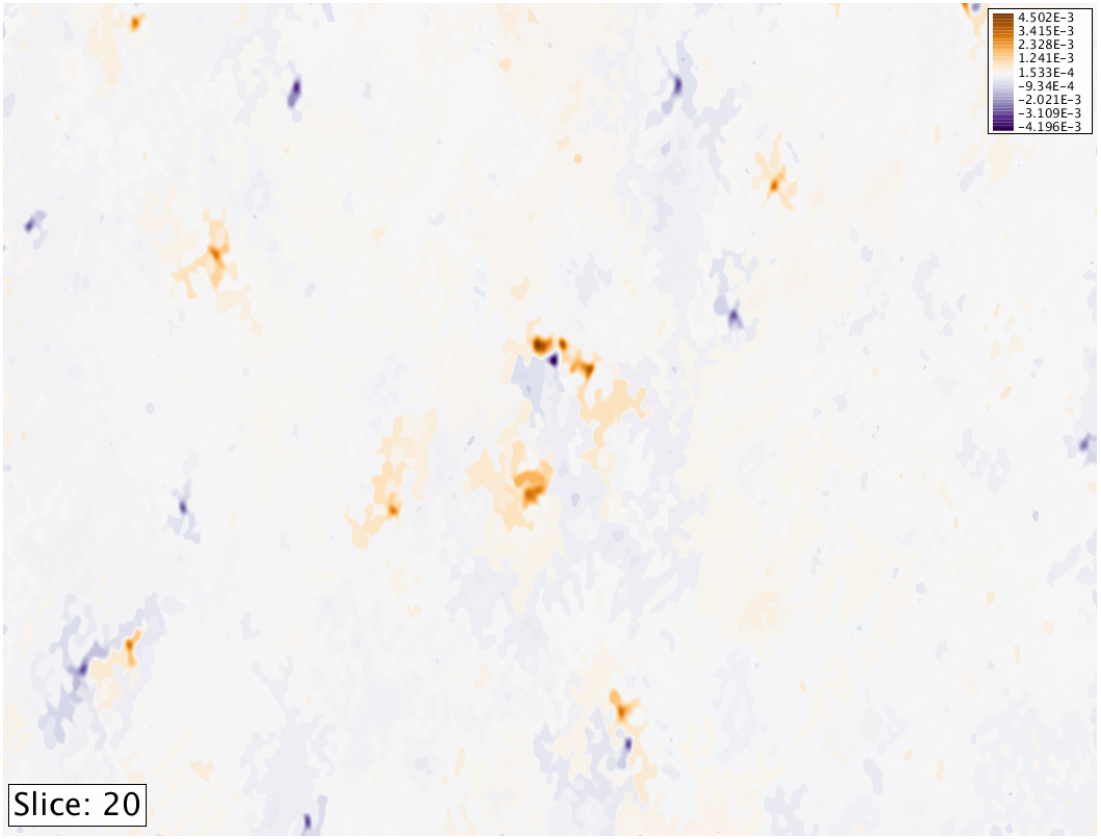
original



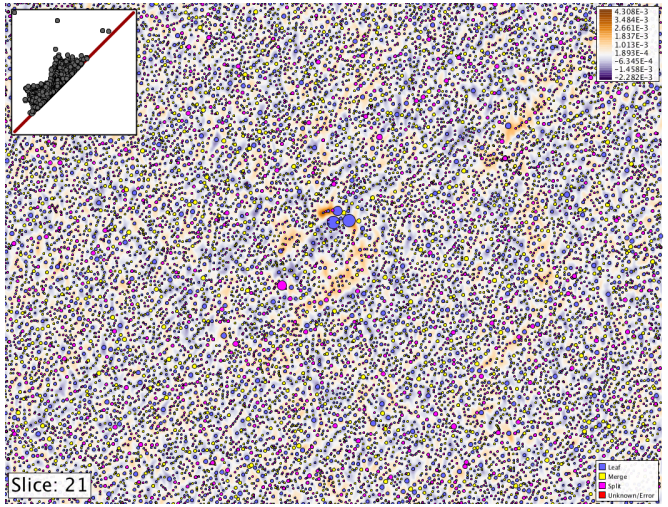
simplified



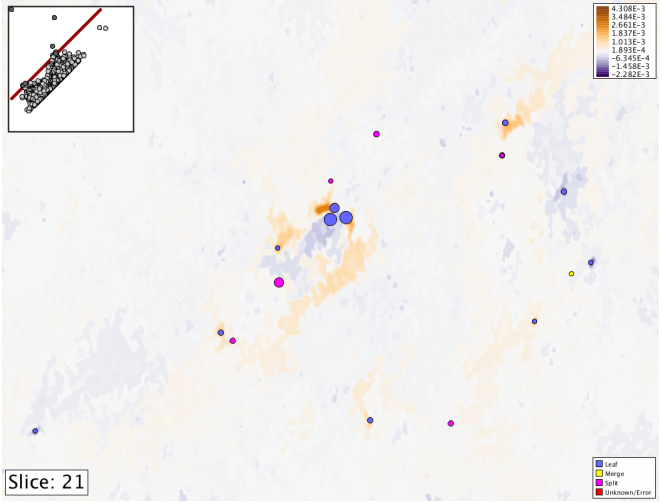


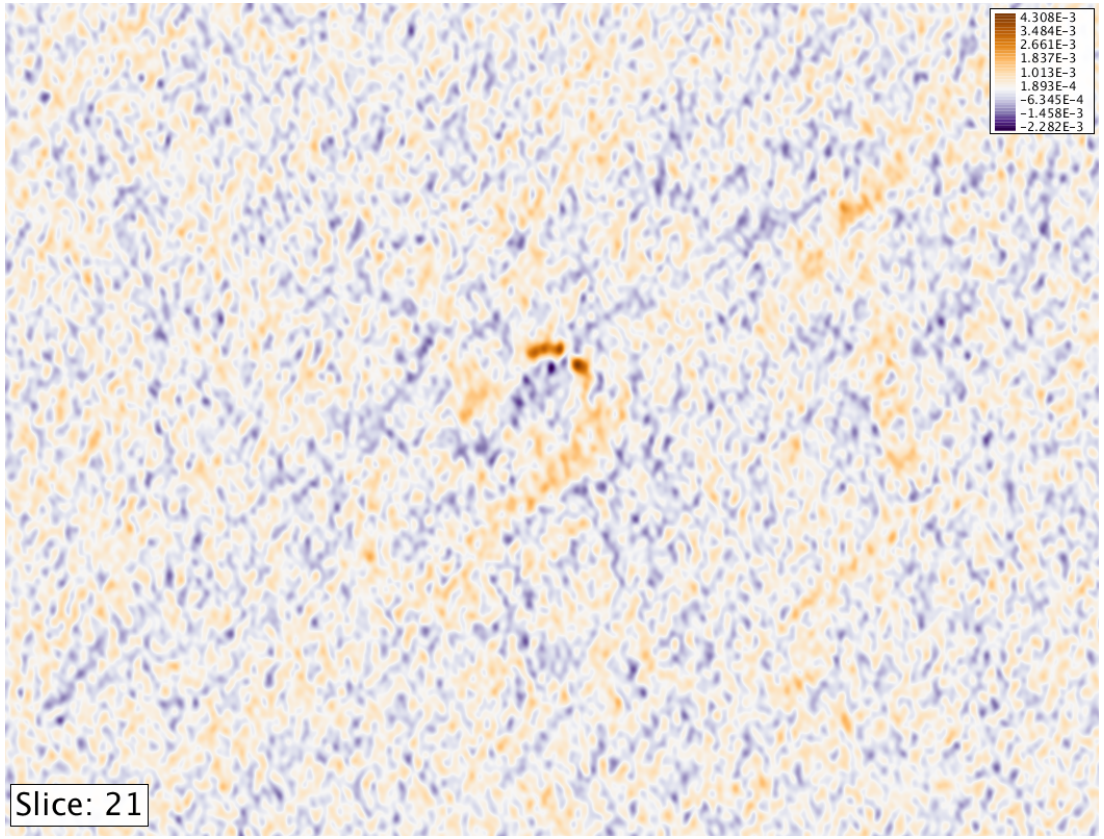


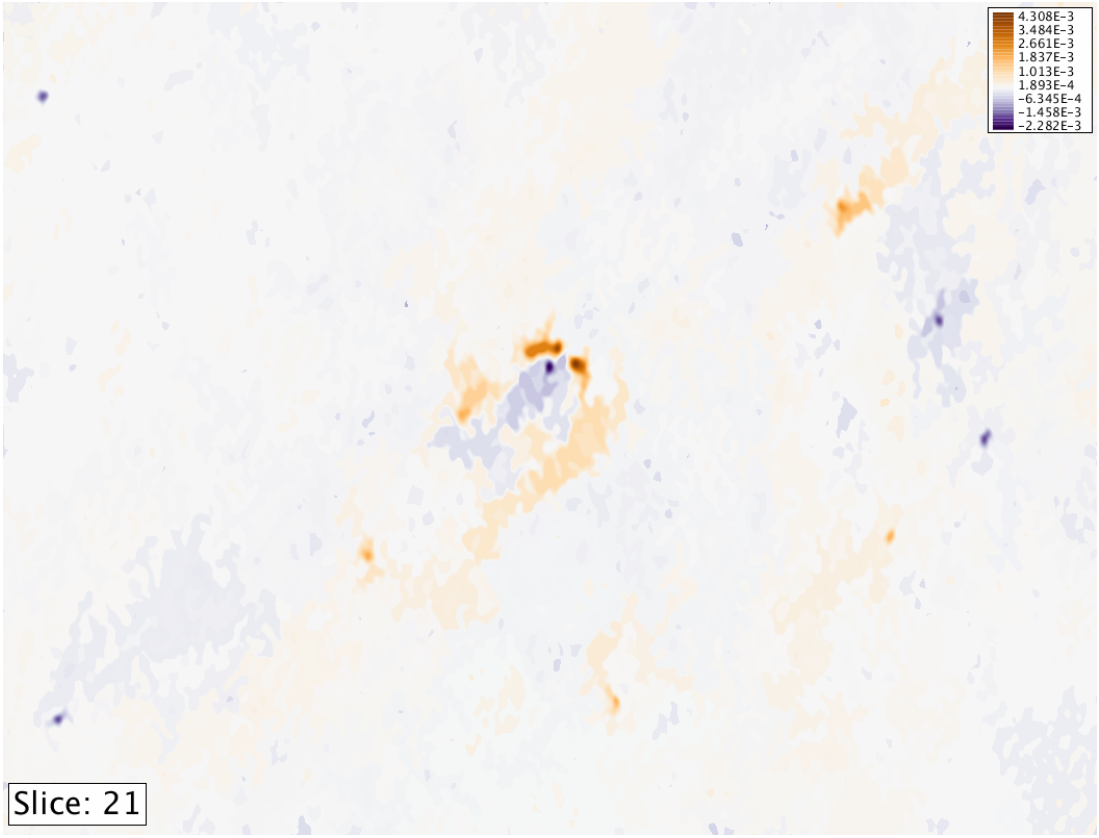
original



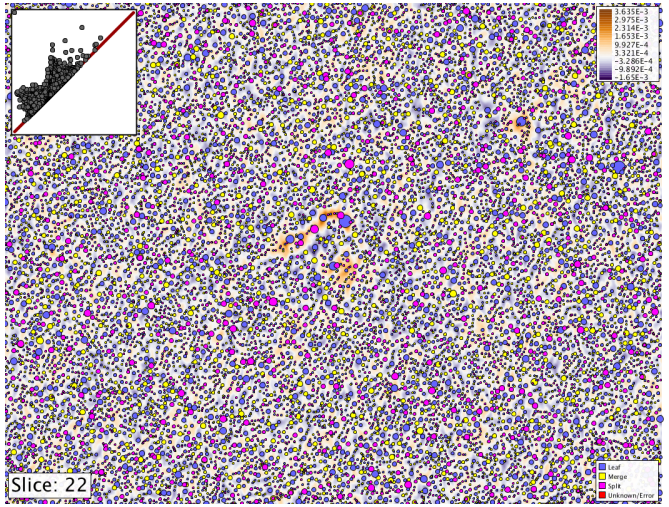
simplified



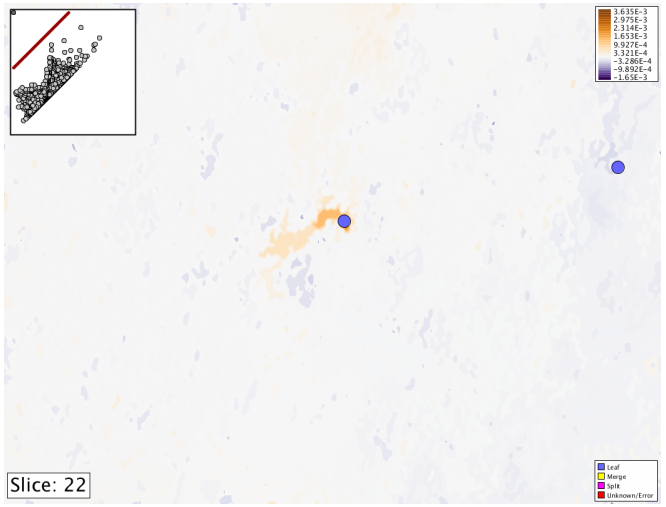


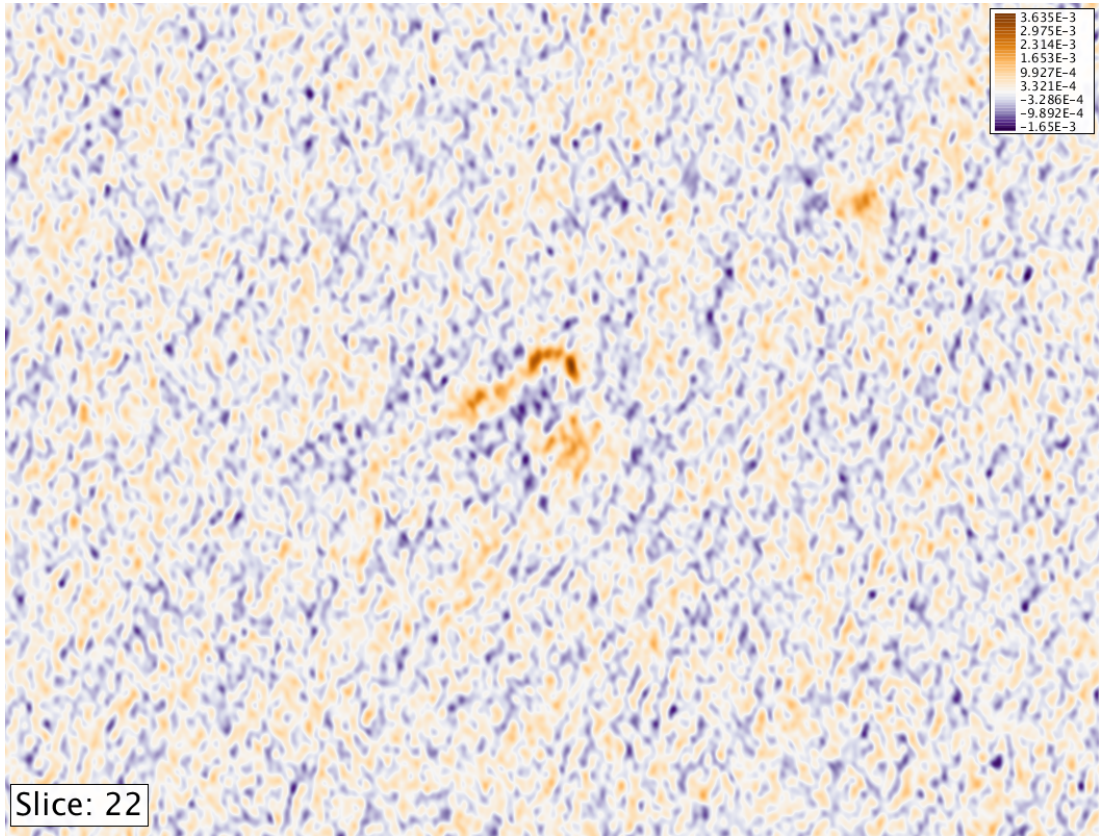


original

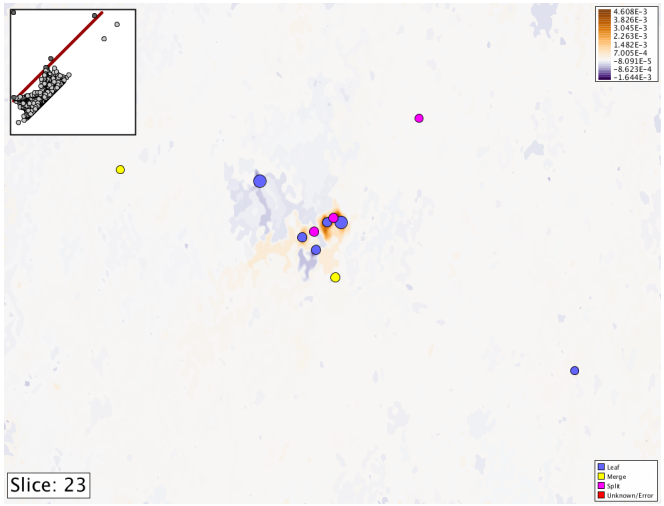


simplified

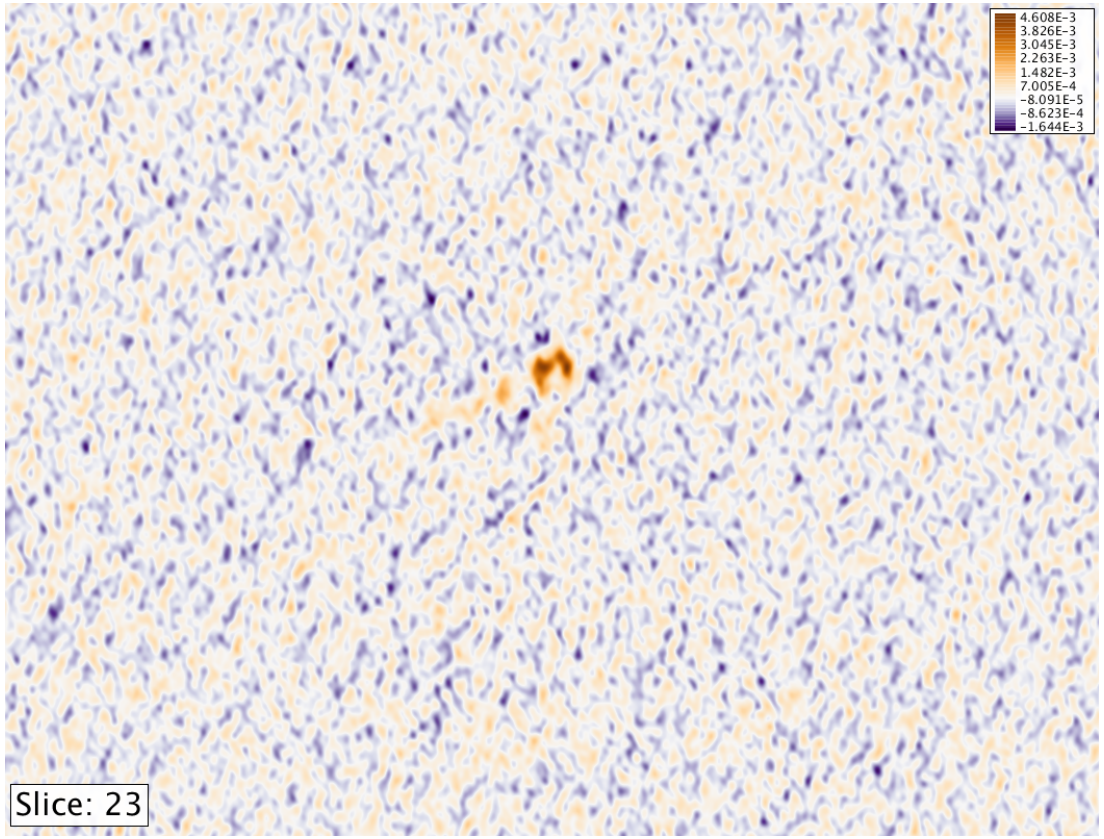


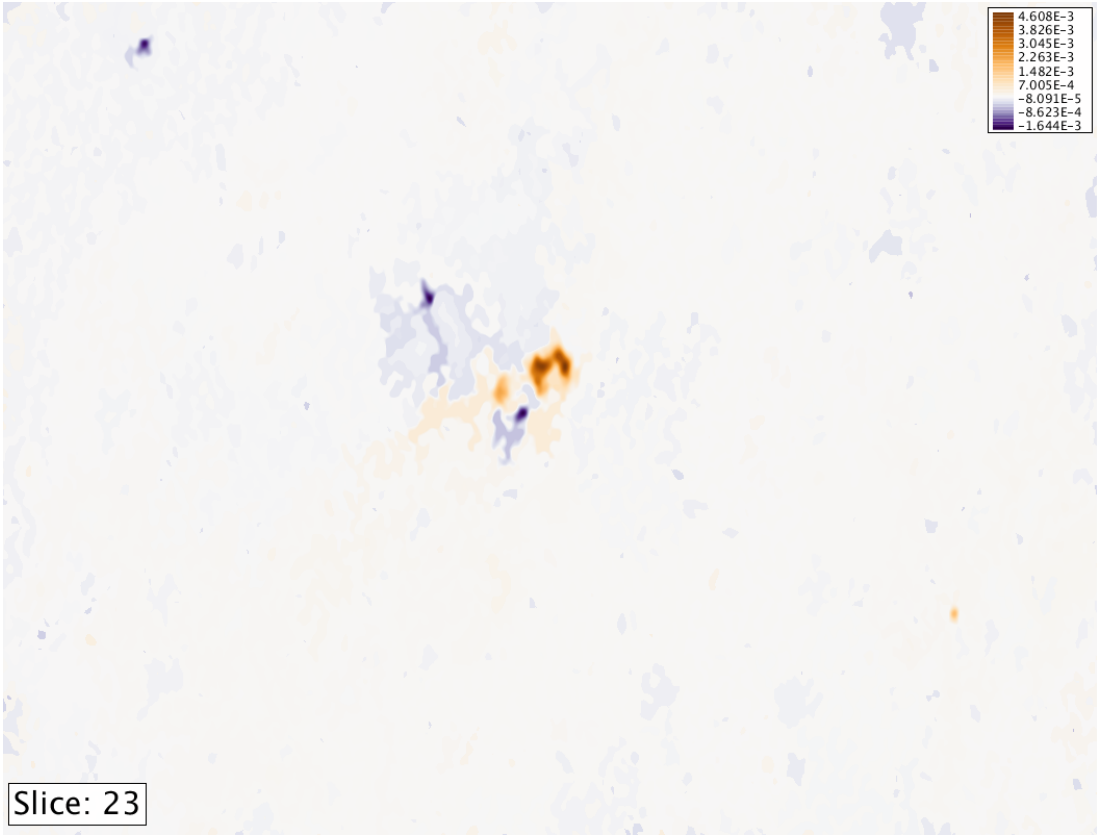


simplified

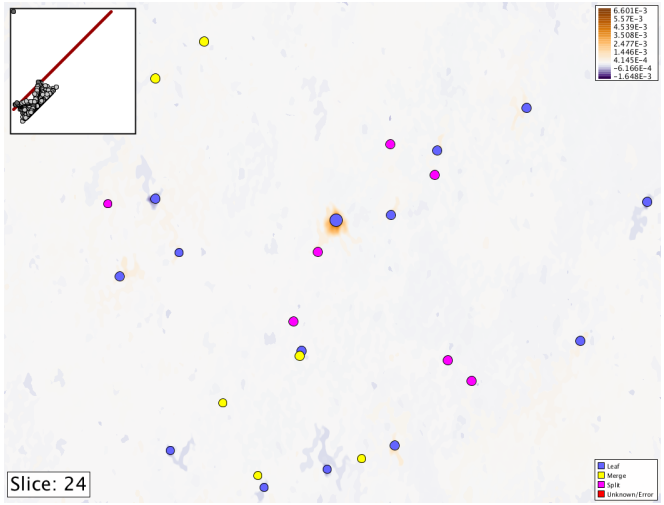


original

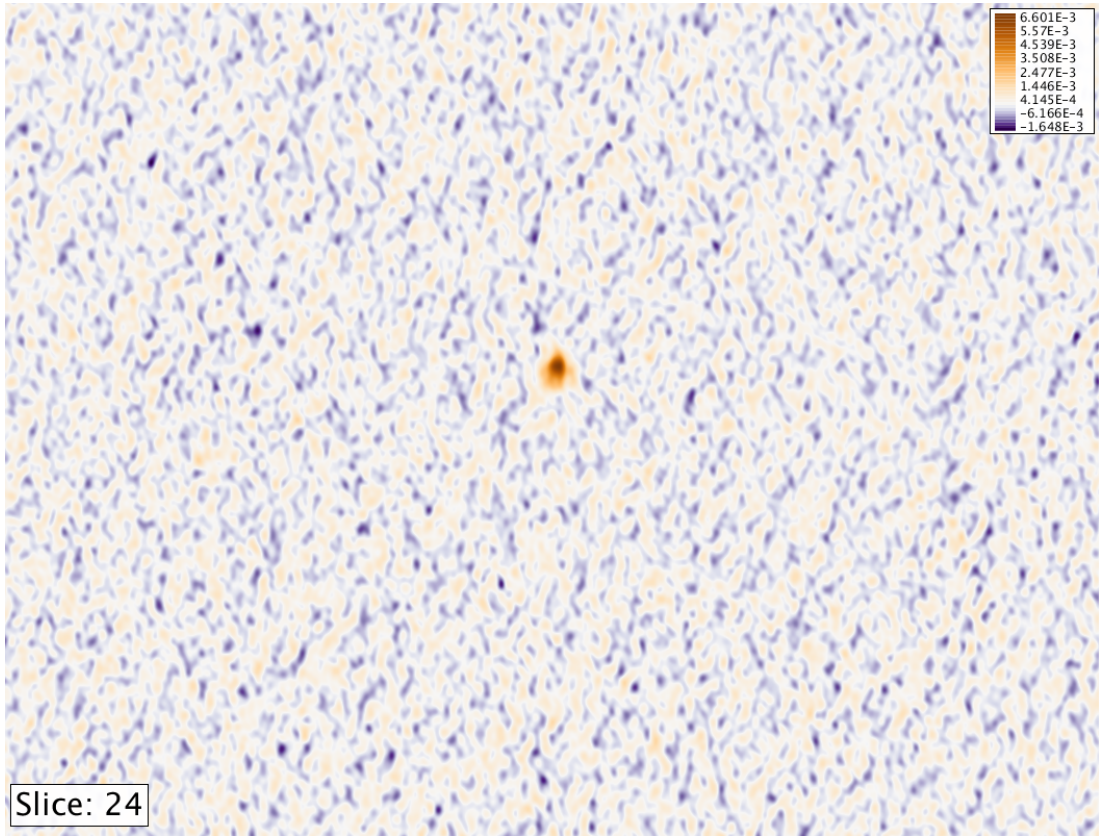


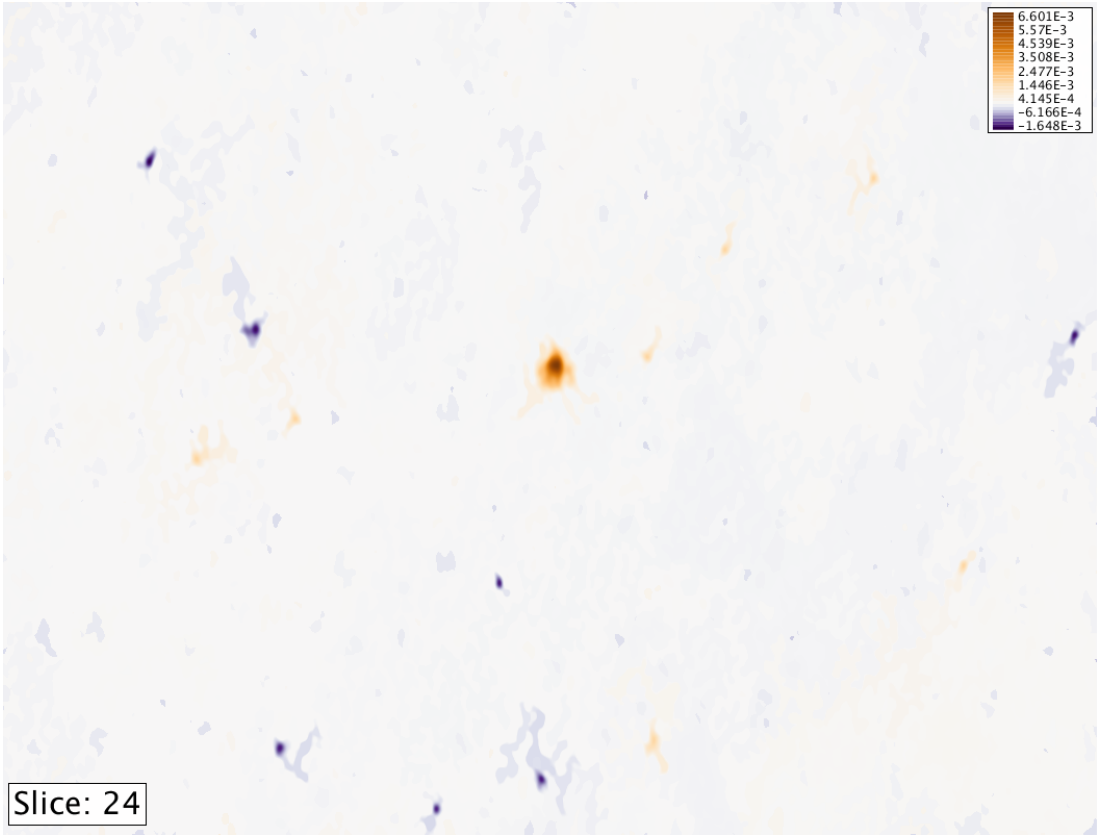


simplified



original





Lunch

Live Demo

Discussion

TOPICS

Improvements to software, additional features, etc.

Scientific challenges/opportunities using this approach

Spreading the word about the technique

Revisit data challenges facing Radio Astronomers

Outstanding Issues

- Moment analysis (simplify before or after)
- Multiple slices (3D contour tree)
- Local vs global contour trees / Scalability
- Boundary
- Scalar field simplification

Outstanding issue: Moment analysis

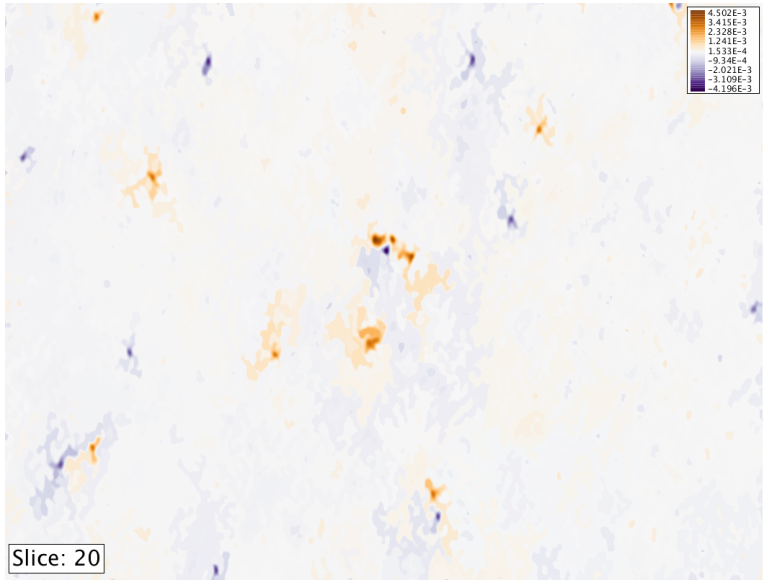
$$M_0 = PS\left(\sum Vel_i\right)$$

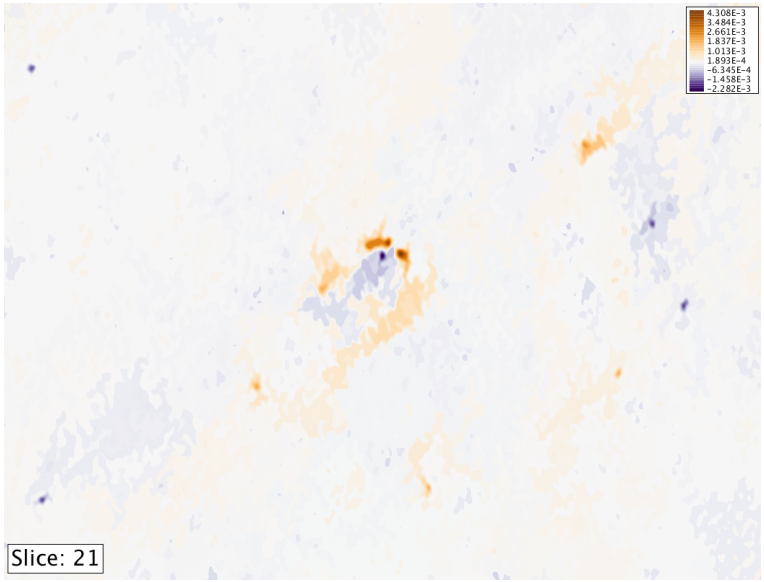
$$or$$

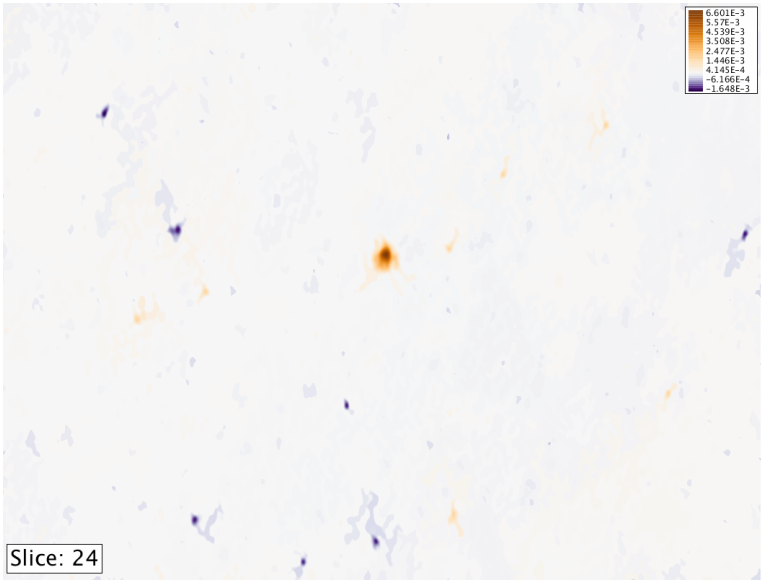
$$M_0 = \sum PS(Vel_i)$$

Outstanding issue: multiple slices

- How to co-simplify?
- Multiple 2D vs 3D contour trees?

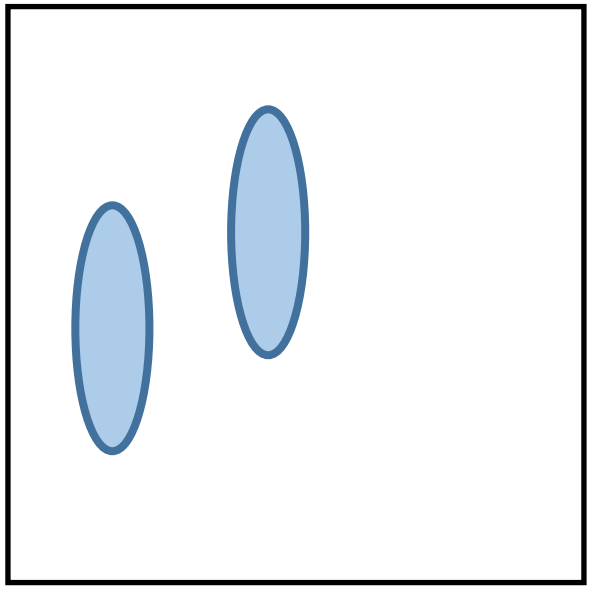




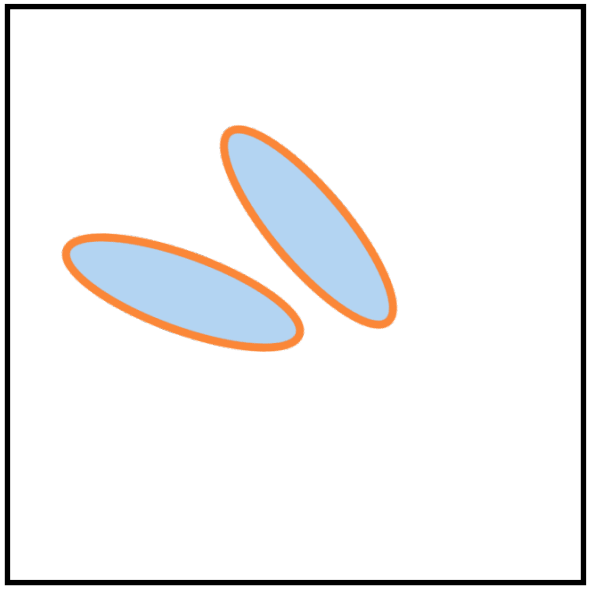


Outstanding issue: 3D contour trees (topological pants)

Slice 1



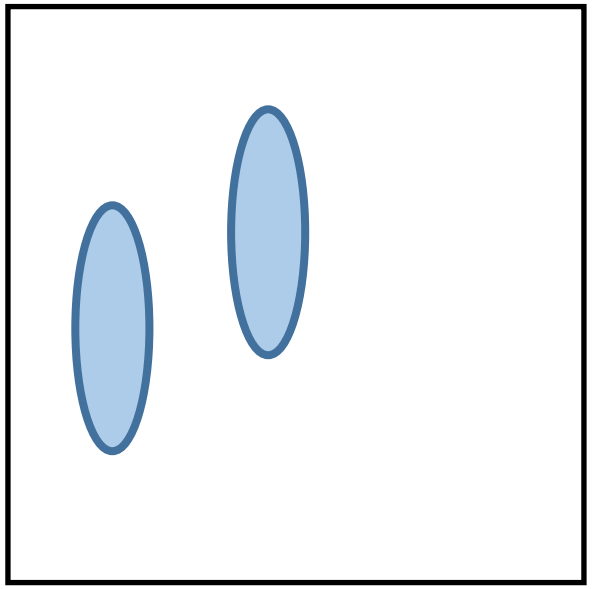
Slice 2



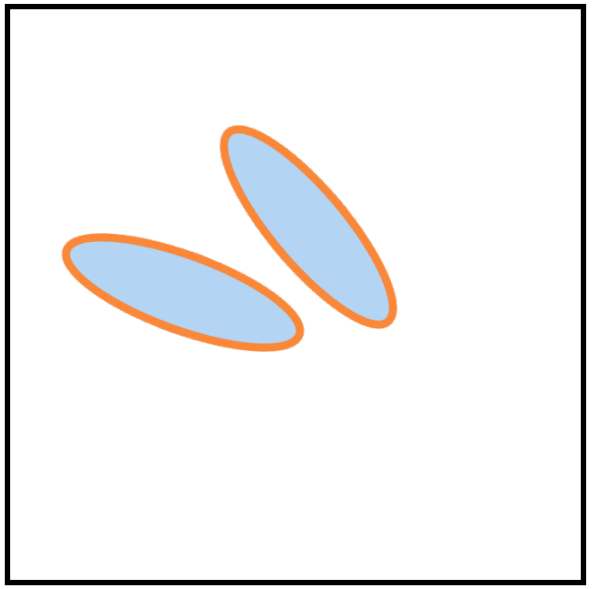
Both have the same isovalue

Outstanding issue: 3D contour trees (topological pants)

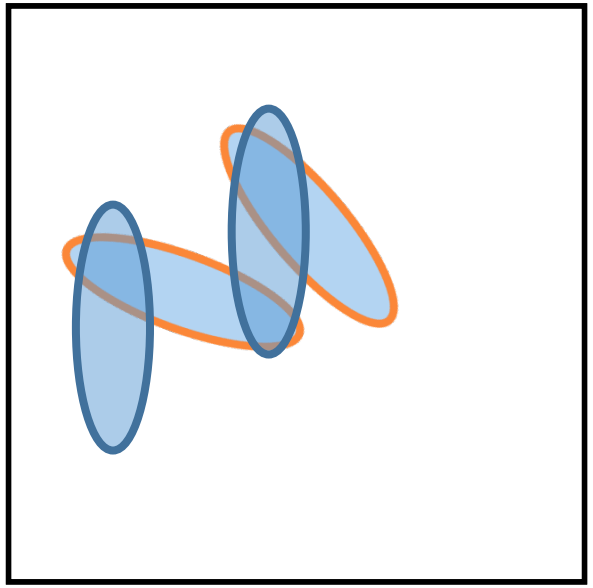
Slice 1



Slice 2

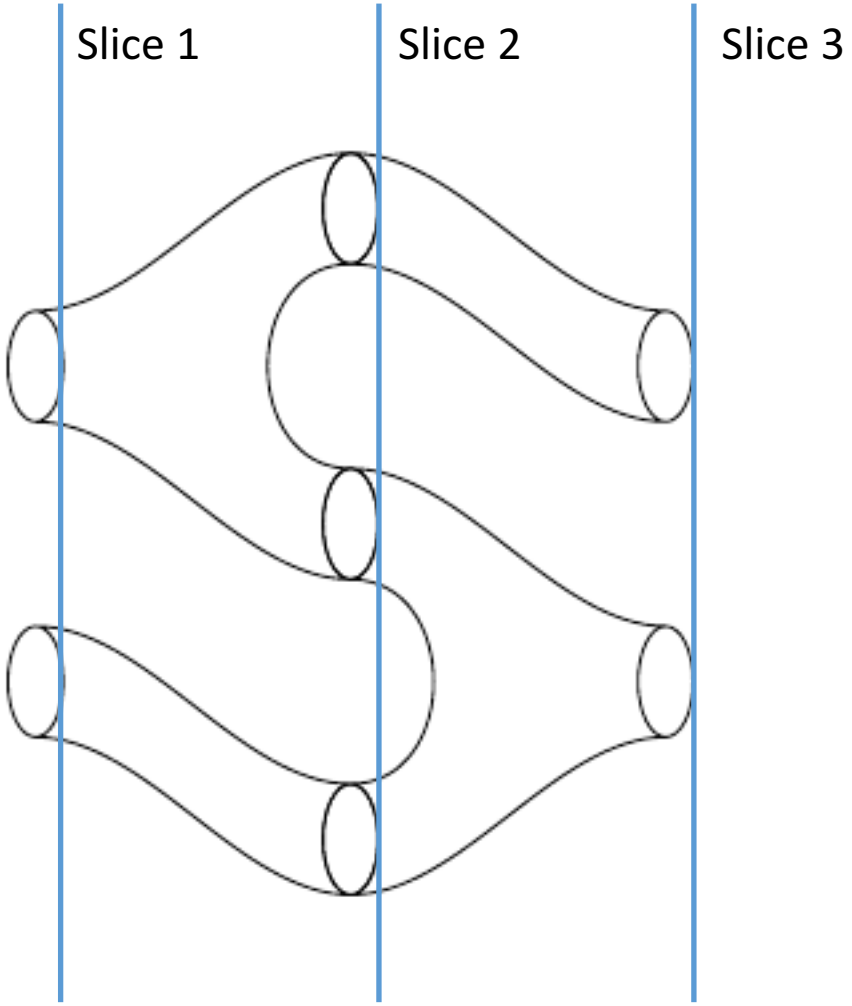


Slice 1+2

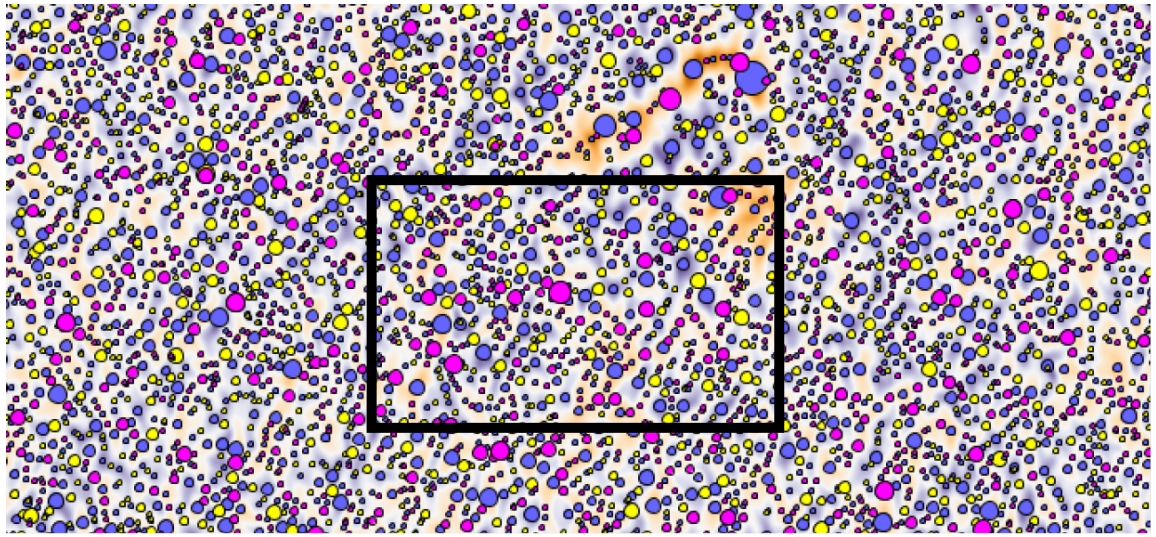


Both have the same isovalue

Outstanding issue: 3D contour trees (topological pants)



Outstanding issue: Local vs. global contour tree



Outstanding issue: Local vs. global contour tree

- 2 part question
  - Which makes sense for analysis?
  - Can we compute global?

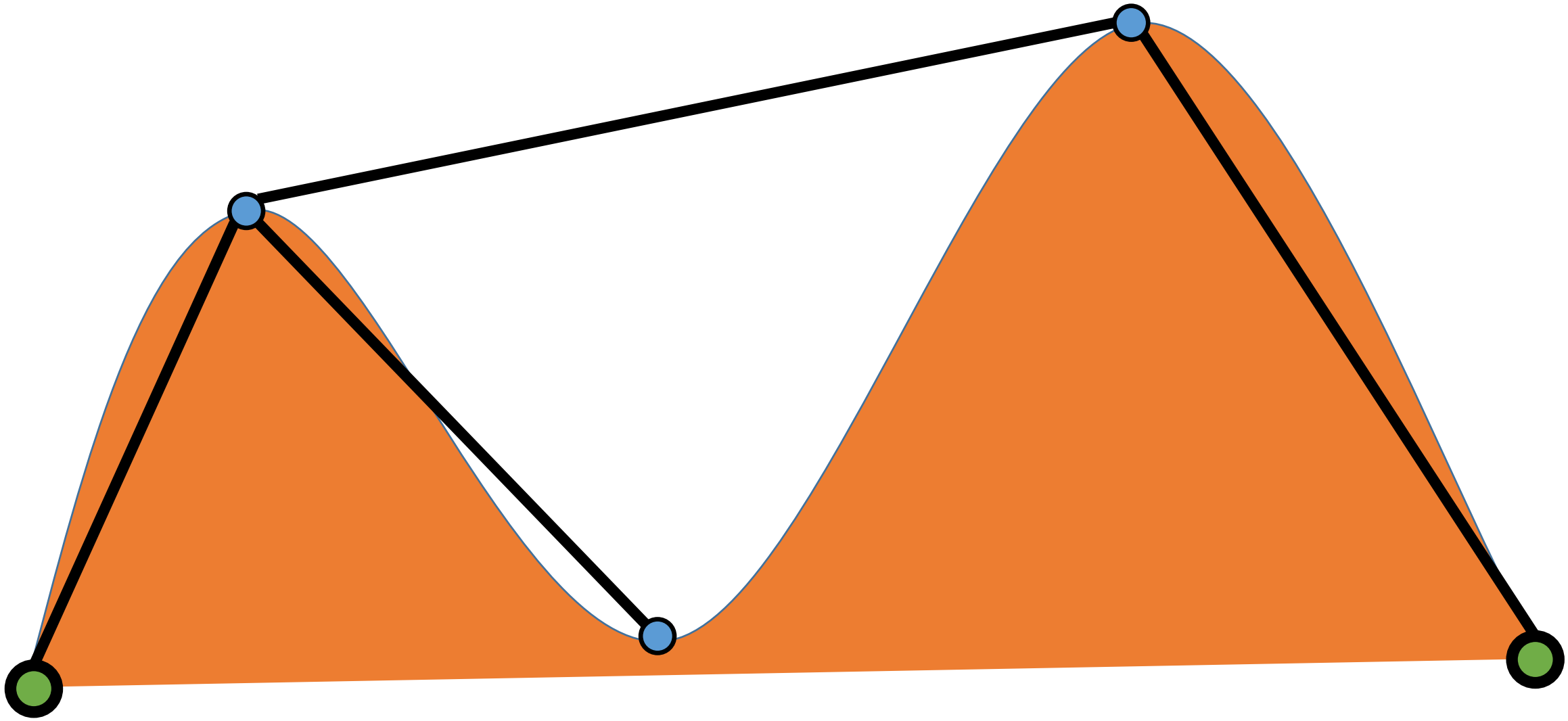
- Research questions:
  - Precomputation?
  - Data storage and query?
  - Efficient computation in parallel?

Outstanding issue: Scalability

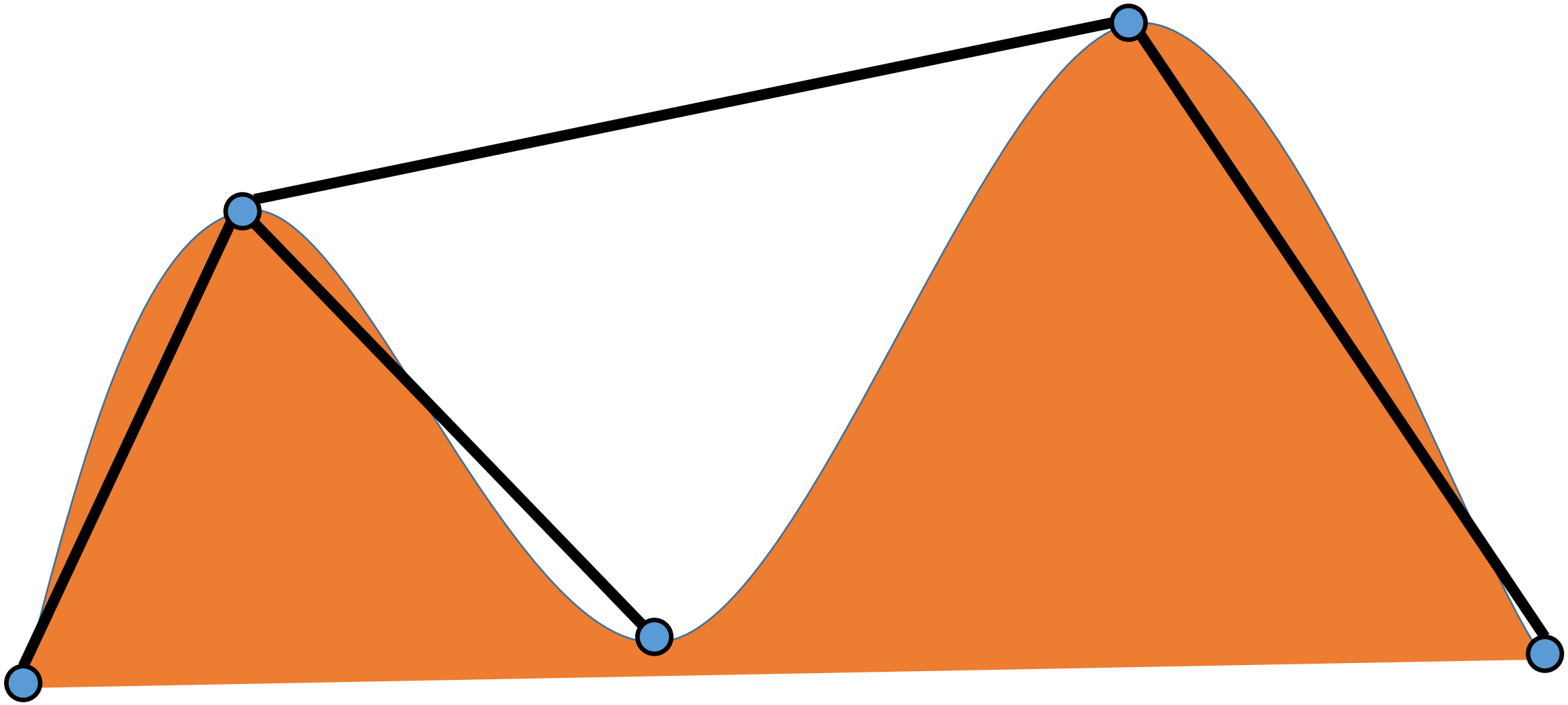
- We have a new parallel GPU implementation of the contour tree in the works (needs debugging)
- This will partially address issue from a technical standpoint
  - Implementation is limited in the size of field it can support (memory bound, not compute bound)
  - Complete solution will require slicing the field and gluing multiple trees

- Note: scalar field simplification on a large field will be costly and still needs a GPU implementation

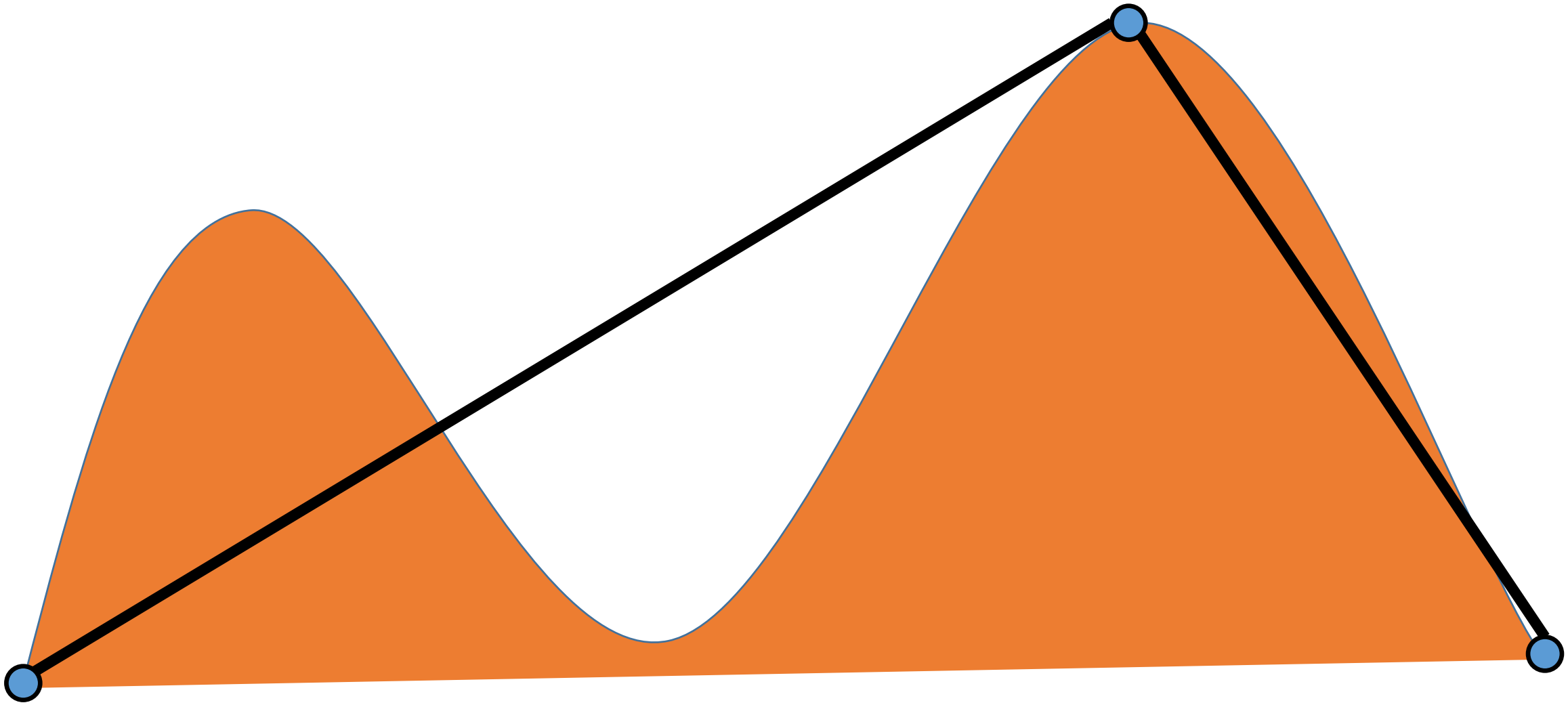
Outstanding issues: Boundary Conditions



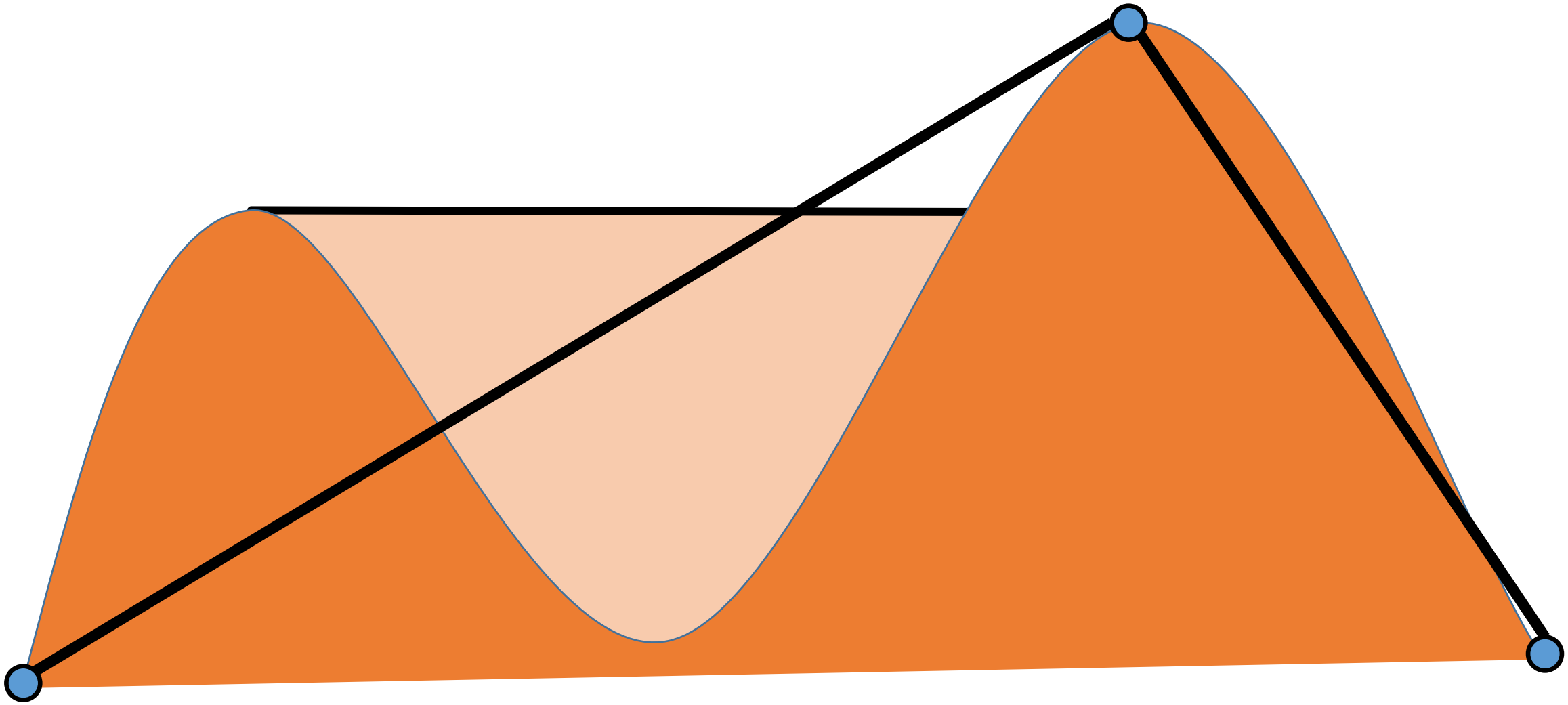
Outstanding issues: Scalar field simplification



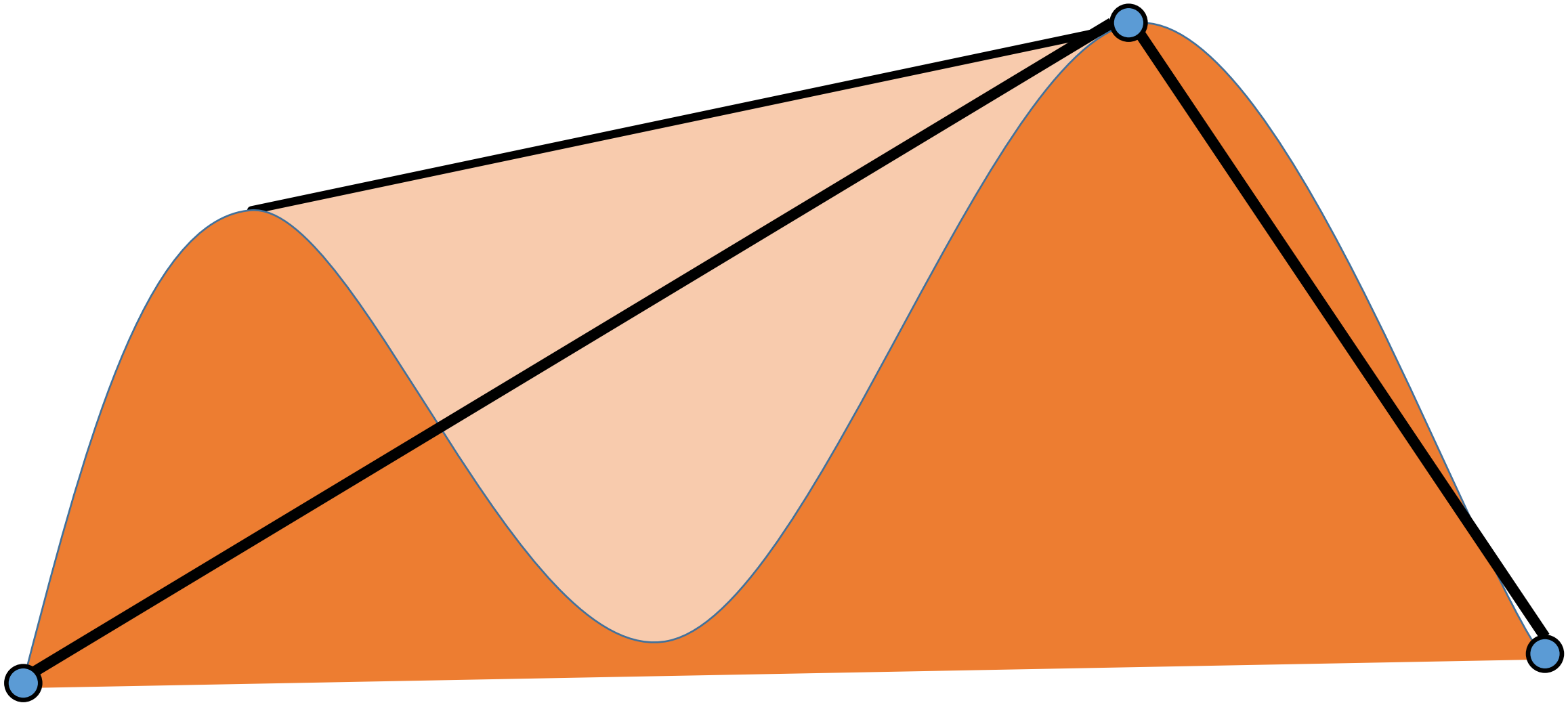
Outstanding issues: Scalar field simplification



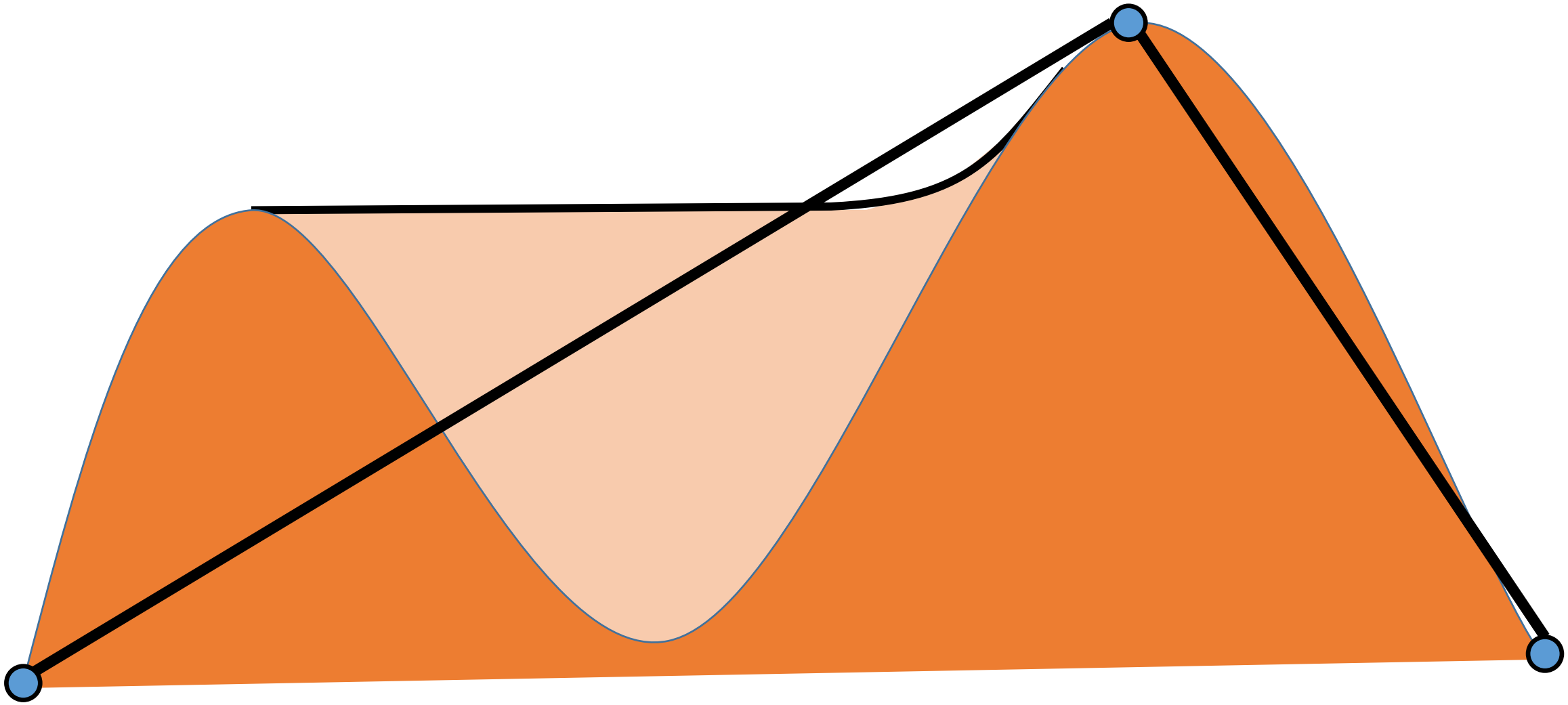
Outstanding issues: Scalar field simplification



Outstanding issues: Scalar field simplification



Outstanding issues: Scalar field simplification



Possible Implementation & Research Issues Moving Forward

- Contour tree parallelization, precomputation, storage, etc.
  - Integration of this data into fits file?
- Parallelization of scalar field simplification
- Investigate alternative scalar field simplification mechanisms
- Visualizations
  - Redesign of persistence diagram
  - Better handing of multiple cubes
  - Better handling of co-simplification interface
- Alternative topological structures (morse-smale complex, more precisely)
- Deeper dive into applications
  - Denoising / noise modeling
  - Source Finding (and prioritization)
  - Source Segmentation
- Scientific validation / verification
- Integration with CASA
- Community outreach
- Combining with other ML technologies (such as clustering or supervised learning)