Discrete graph extraction from simulated scalar fields

Hari Prasad, T.K.M College of Engineering, India
Mentors : **Oskar Elek, Farhanul Hasan**

Google Summer of Code

UNIVERSITY OF CALIFORNIA SANTA CRUZ
CROSS | CENTER FOR RESEARCH IN OPEN SOURCE SOFTWARE

Introduction

The PolyPhy Graph Extraction project aims at generating graph structures from simulated scalar fields. The project is part of PolyPhy, a python implementation of PolyPhorm which uses the Monte Carlo Physarum Machine (MCPM) algorithm inspired by the foraging behavior of Physarum polycephalum 'slime mold', in order to analyze intergalactic gas and dark matter filaments together known as the cosmic web. The project focuses on two methodologies, one using a custom graph extraction algorithm based on an agent-based approach. The second method utilizes the TTK-Paraiew Library to develop a pipeline for graph extraction and visualization using the Morse Smale Complex.

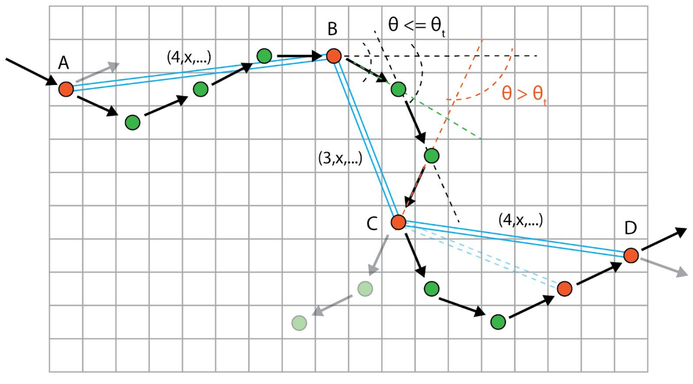
Existing solution and drawbacks

DisPerSe is an established method in astronomy and people in the field have good intuition about the topological structures. But the main limitations of DisPerSe are the setup process and the legacy code. Also, the library is written in C++ which limits the users from easily extending its capabilities and also adds difficulty in prototyping with the library. The solution we arrived at is a pipeline that is written in python, which provides a much lesser learning curve to the users and the ability to connect to several of the topology and analytics libraries available with python.

Proposed solution 1

Custom graph extraction algorithm

The method is to apply the concept already used in the MCPM algorithm, which is an agent-based approach, but on the scalar trace field output by the PolyPhorm. The algorithm will use a number of randomly placed probes in the scalar trace field which will navigate through the field based on the density while adding nodes and edges in the path traced by the probes, based on various criteria such as the curvature in the path, the density of the field at a given point, number of intermediate points, etc.



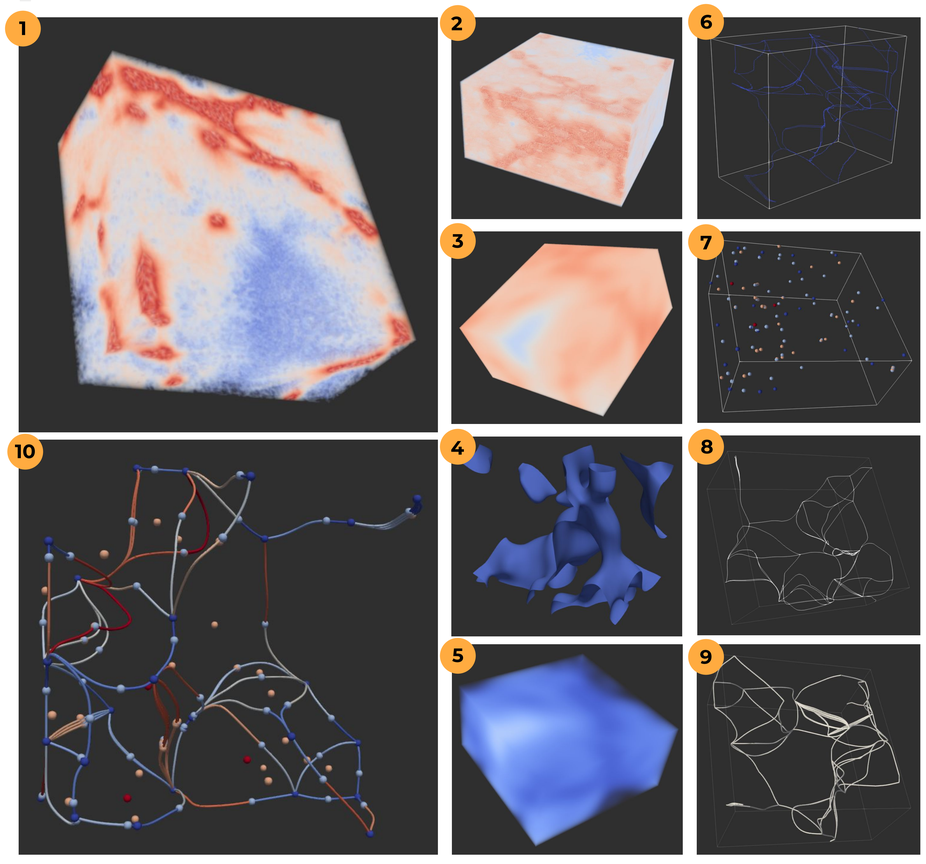
The challenges faced while developing the algorithm, which still need to be worked on are

- The probes when traversing the field over several iterations, tend to produce spurious connections.
- A criterion is required to terminate the path of a probe
- Connecting two intersecting paths and identifying nearest probes or scalar points already traversed by other probes.
- Incorporating the thickness of the filament structures while using a probe-based algorithm.

Proposed solution 2

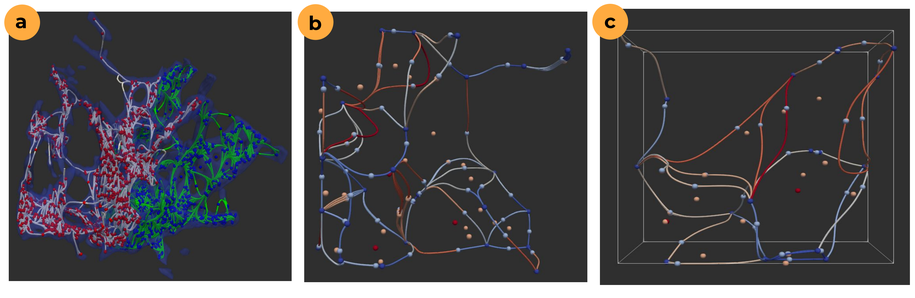
TTK-ParaView pipeline

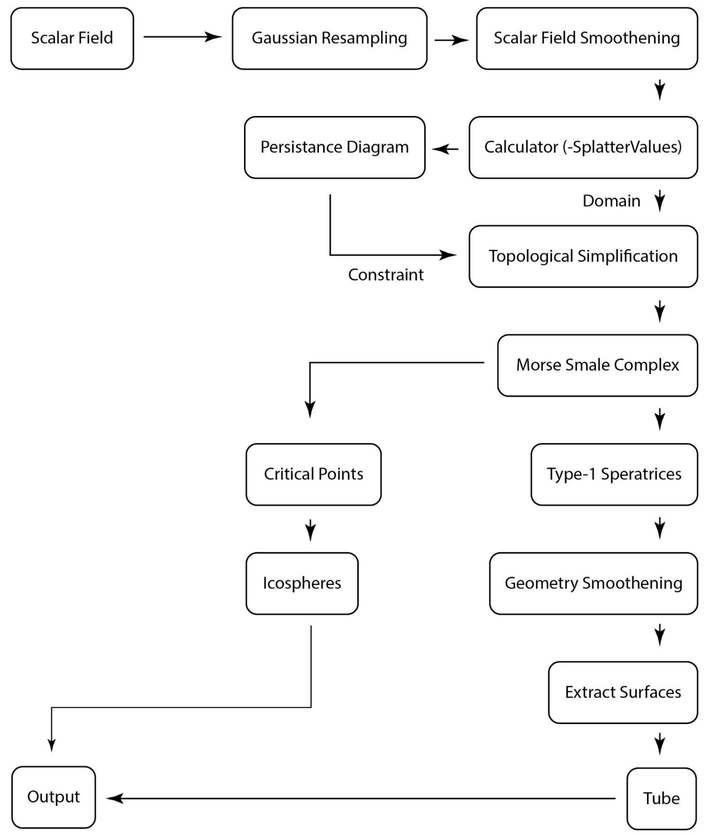
The Topology ToolKit (TTK) is an open-source library and software collection for topological data analysis and visualization. TTK can handle scalar data defined either on regular grids or triangulations, in 2D, 3D, or more, incorporating various algorithms such as critical points, persistence diagrams, persistence curves, Morse-Smale complexes, The TTK-Paraview method uses Persistence Diagram and Morse Smale complex to extract graph-like structures from the scalar field. With the user-friendly GUI of ParaView and the TTK library, we tried to develop a pipeline to perform the same in a much easier way compared to DisPerSe.



Experiments

Experimentation on the continuity of the extracted structures (a), different levels of smoothing (b) and different gaussian resampling grid size ©.





Advantages over DisPerSe

- Easy to setup and modify the pipeline.
- Ability to integrate with other python libraries.
- Extract the multiple extracted structures individually and then combine them in ParaView.
- Easy to use GUI with TTK integrated ParaView.
- Numerous visualization tools and functions available in ParaView.
- Ability to export the ParaView project as a Python script.

Future works

- Comparisons between the outputs generated by both DisPeSe and TTK-ParaView pipeline.
- Cosmological analysis using various Graph algorithms, on the extracted graphs representing the cosmic web structures, giving structural level understanding, spatial locations of structures, disentangling structures, studying the relation between star formation activity and the position of galaxies in the filaments, interactions between gas clouds and galaxies and measuring the distances and analyzing the activity within galaxies, etc.
- Developing a functional algorithm that performs comparable to the current pipeline which would provide a different solution compared to the Discrete Morse theory solution used by both DisPerSe and TTK-ParaView pipeline.