Cosmic Web Detection and Inference Through Directional Density Ridge

Yikun Zhang† (yikun@uw.edu)
Joint work with Yen-Chi Chen†, Rafael S. de Souza*, and Alexander Giessing†
†*Department of Statistics, University of Washington* * *Department of Physics, Astronomy and Mathematics, University of Hertfordshire*

Research Problem

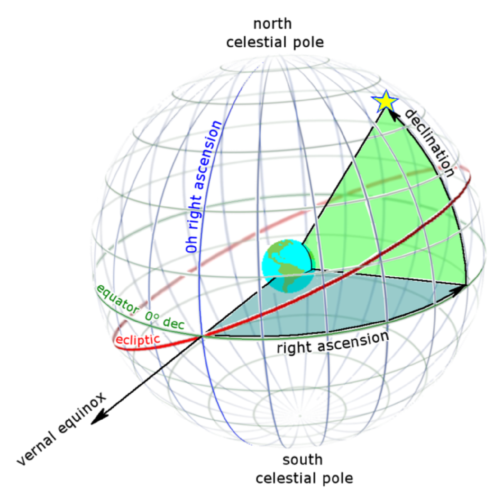
Our research objective is to detect the cosmic web structure based on the distribution of observed galaxies by the Sloan Digital Sky Survey (SDSS) and study its effects on the stellar properties of nearby galaxies. In particular, we mainly focus on the one-dimensional cosmic filaments.

Background: Previous simulation and observational studies have shown that on megaparsec scales, matter in the Universe is not uniformly distributed but rather forms a complicated large-scale network structure called **Cosmic Web**. It has four main components:

- 0D massive and overdense **galaxy clusters** (or **nodes**),
- 1D interconnected **cosmic filaments**,
- 2D tenuous **cosmic sheets** (or **cosmic walls**),
- 3D vast and near-empty **cosmic voids**.

Observational Galaxy Data: $\{(\alpha_i, \delta_i, z_i)\}_{i=1}^n \subset \mathbb{S}^2 \times \mathbb{R}^+$:

- $\alpha_i \in [0, 360°)$ is the *right ascension* (RA; celestial longitude),
- $\delta_i \in [-90°, 90°]$ is the *declination* (DEC; celestial latitude),
- $z_i \in [0, \infty)$ is the *redshift* value that measures the distance from an observer to the galaxy.



Drawbacks of Previous Works

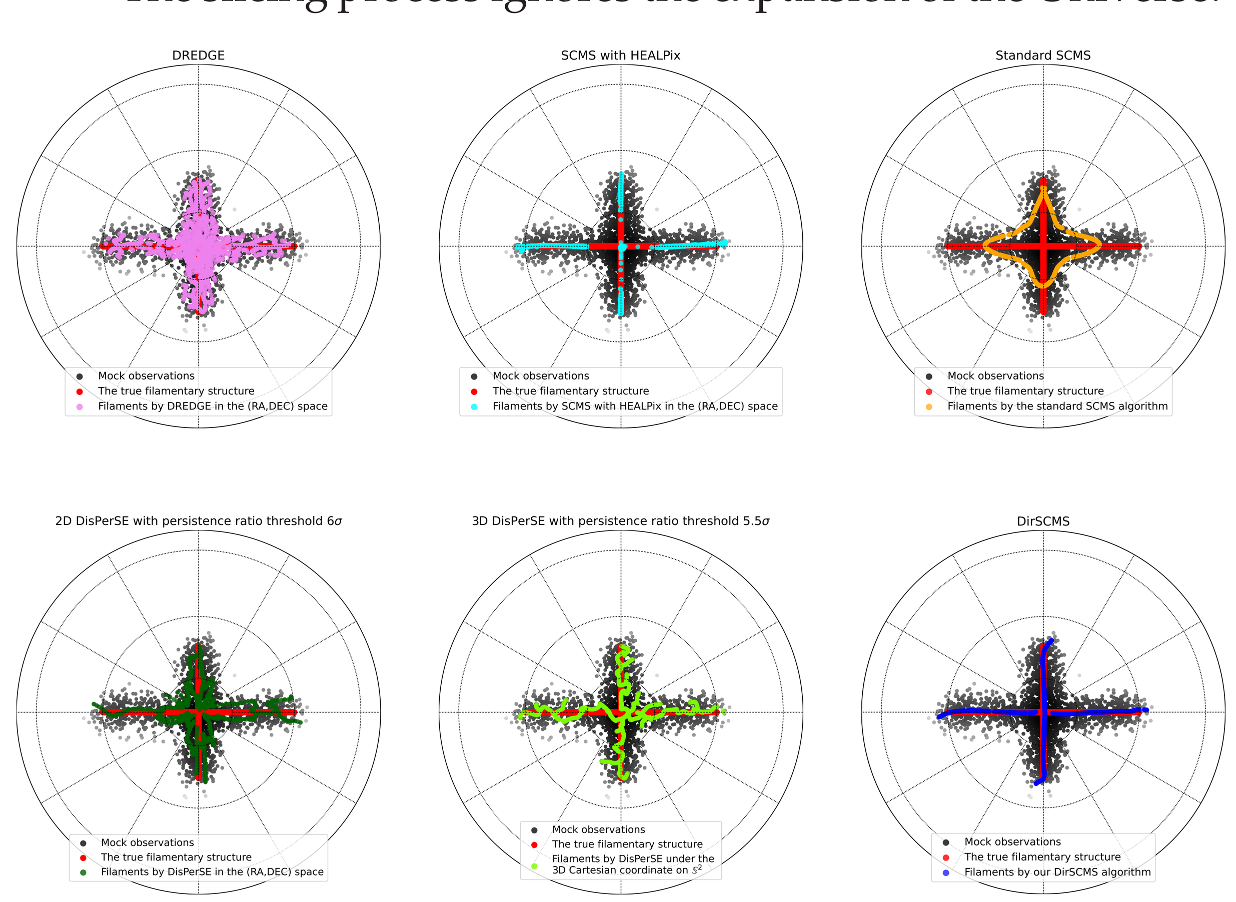
The existing filament detection methods lie in two categories:

3D Method: Convert redshift into (comoving) distances and recover the filaments in the 3D space.

- The distance transformation function depends on complicated cosmological models.
- The detected filaments could be spurious due to the finger-of-god effects.

2D Method: Partition the Universe into thin redshift slices and estimate the filaments in each slice.

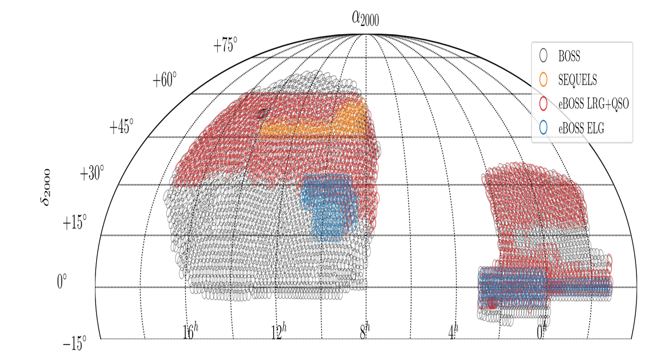
- The existing 2D slicing approaches assume a flat-sky approximation.
- The slicing process ignores the expansion of the Universe.



► **Indeed, our proposed method can easily switch between the above two categories and is adaptive to the underlying spherical geometry!**

Proposed Methodology

Step 1 (Slice the Universe): We partition the redshift range of the observed galaxies into 325 non-overlapping thin slices based on the comoving distance $\Delta L = 25$ Mpc (i.e., tomographic analysis).



Step 2 (Restrict to the Region of Interest): Within each slice, we subset the galaxies in some specific regions of the *North Galactic Cap* and *South Galactic Cap*.

Step 3 (Estimate the Galaxy Distribution): We convert the angular coordinates $(\alpha_i, \delta_i) \in [0, 360°) \times [-90°, 90°]$, $i = 1, ..., n$ of galaxies in each slice to their Cartesian ones on $\mathbb{S}^2$ as:

$$\boldsymbol{X}_i = (\cos\delta_i \cos\alpha_i, \cos\delta_i \sin\alpha_i, \sin\delta_i) \quad \text{for} \quad i = 1, ..., n,$$

and estimate the galaxy density field on $\mathbb{S}^2$ by the directional kernel density estimator (DirKDE):

$$f(\boldsymbol{x}) = \frac{C(b)}{n} \sum_{i=1}^n \exp\left(\frac{\boldsymbol{x}^T \boldsymbol{X}_i}{b^2}\right), \tag{1}$$

where $b > 0$ is the smoothing bandwidth parameter and $C(b) > 0$ is a normalizing constant that depends on b.

Step 4 (Directional Density Ridge and SCONCE Algorithm): Given the (estimated) density field f defined in (1), we model the 1D cosmic filament on $\mathbb{S}^2$ through directional density ridge:

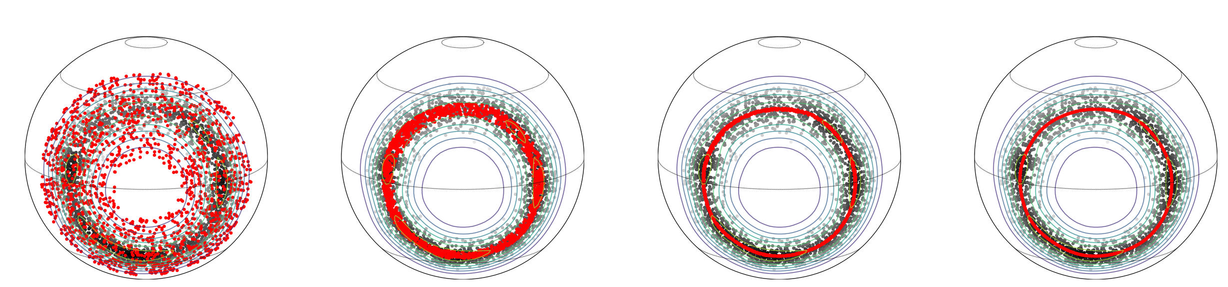
$$R_1(f) = \left\{\boldsymbol{x} \in \mathbb{S}^2 : \boldsymbol{v}_2(\boldsymbol{x})^T \mathtt{grad}\, f(\boldsymbol{x}) = \boldsymbol{0}, \lambda_2(\boldsymbol{x}) < 0\right\}, \tag{2}$$

where

- $\mathtt{grad}\, f(\boldsymbol{x})$ is the *Riemannian gradient* of f at $\boldsymbol{x} \in \mathbb{S}^2$,
- $\mathcal{H} f(\boldsymbol{x})$ is the *Riemannian Hessian* matrix of f at $\boldsymbol{x} \in \mathbb{S}^2$,
- $\boldsymbol{v}_1(\boldsymbol{x}), \boldsymbol{v}_2(\boldsymbol{x})$ are unit eigenvectors of $\mathcal{H} f(\boldsymbol{x})$ that lie within the tangent space $T_{\boldsymbol{x}}$ at $\boldsymbol{x} \in \mathbb{S}^2$ with associated eigenvalues $\lambda_1(\boldsymbol{x}) \geq \lambda_2(\boldsymbol{x})$.

Practically, we estimate the cosmic filament (2) by directional subspace constrained mean shift (DirSCMS) algorithm from our filament finder `SCONCE` with iterations ($t = 0, 1, ...$):

$$\boldsymbol{x}^{(t+1)} \leftarrow \boldsymbol{x}^{(t)} + \boldsymbol{v}_2(\boldsymbol{x}^{(t)}) \boldsymbol{v}_2(\boldsymbol{x}^{(t)})^T \frac{\sum_{i=1}^n \boldsymbol{X}_i \exp\left(\frac{\boldsymbol{X}_i^T \boldsymbol{x}^{(t)}}{b^2}\right)}{\left\|\sum_{i=1}^n \boldsymbol{X}_i \exp\left(\frac{\boldsymbol{X}_i^T \boldsymbol{x}^{(t)}}{b^2}\right)\right\|_2}.$$

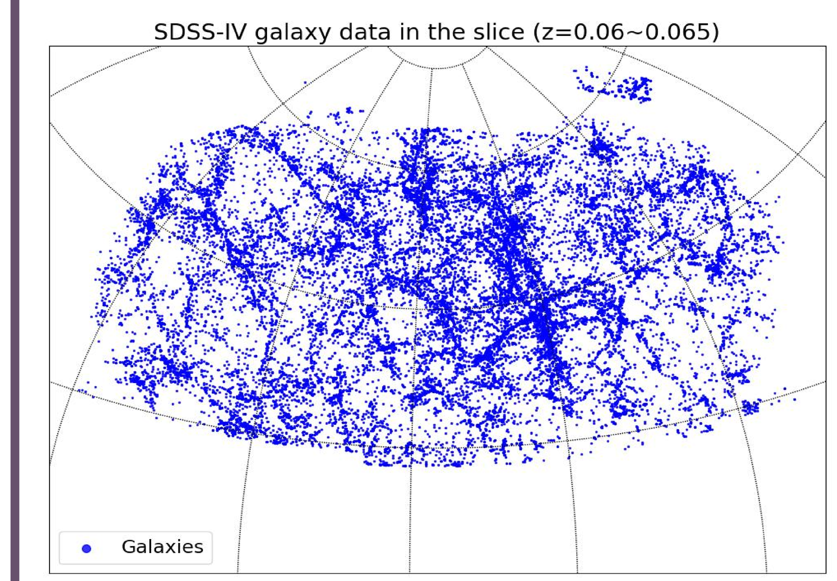


The panels from left to right present the DirSCMS algorithm applied to the mesh points (red dots) at iterations 0, 1, 2, and 8, where the contour lines indicate DirKDE (1) based on the input observations (gray dots).

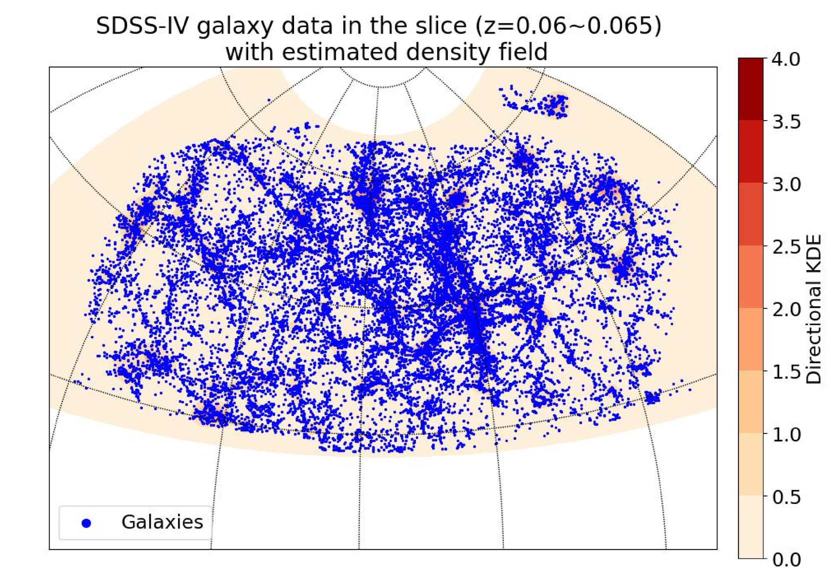
Step 5 (Bootstrap Inference): We quantify the local uncertainty of each filament points via nonparametric bootstrap.

Step 6 (Identify Local Modes and Filament Knots): We identify local modes of f via directional mean shift algorithm and intersection points of detected filaments via metric graph reconstruction algorithm.

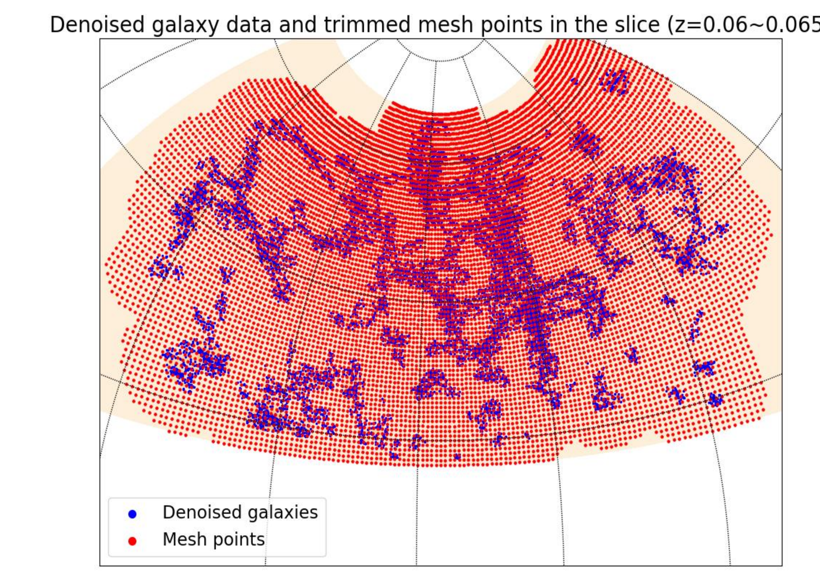
Cosmic Web Detection From SDSS Galaxies on an Example Redshift Slice



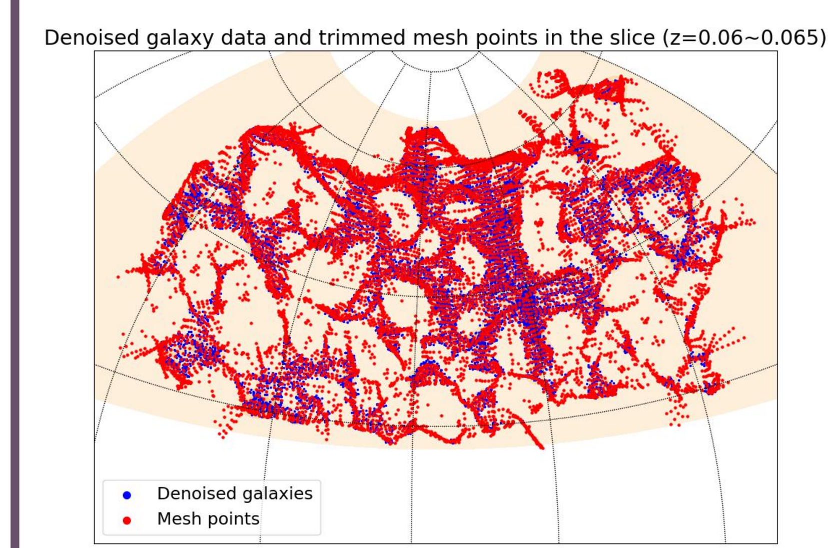
Steps 1 and 2: Slice the Universe and Restrict to the North Galactic Cap.



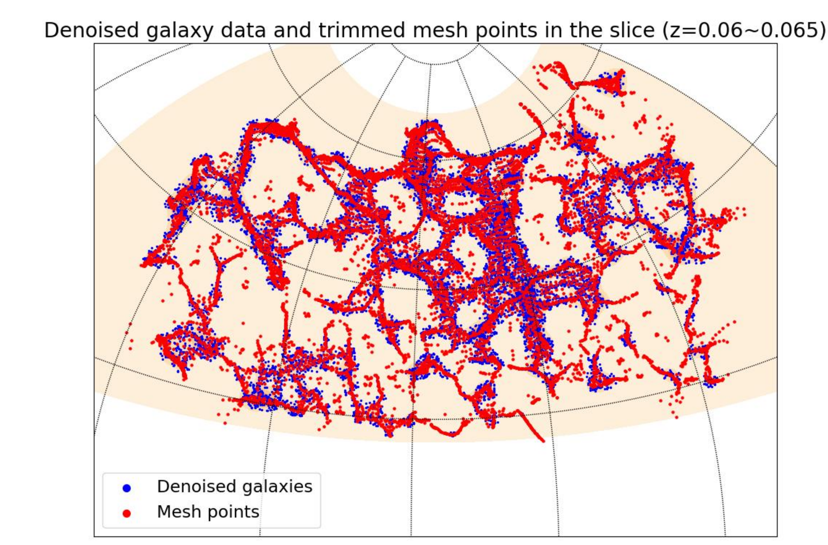
Step 3: Estimate the Galaxy Density Field.



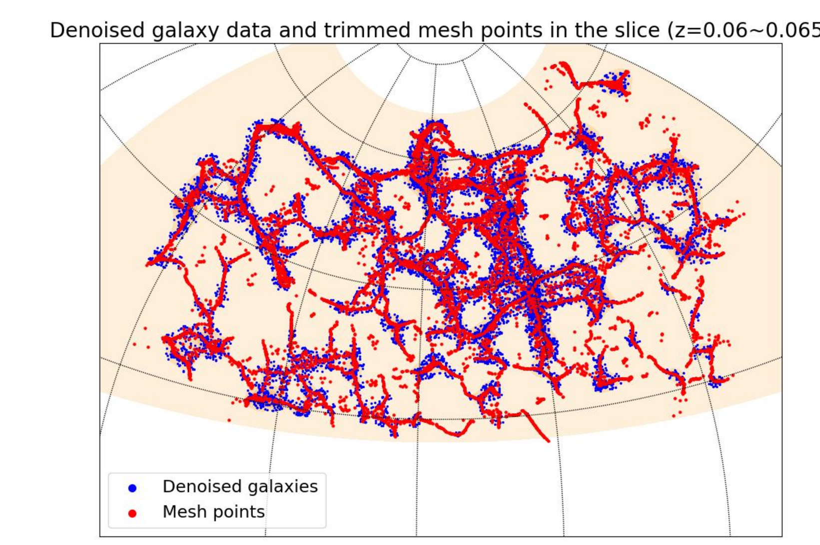
Step 4-1: Denoise the Density Field and Lay Down Mesh Points.



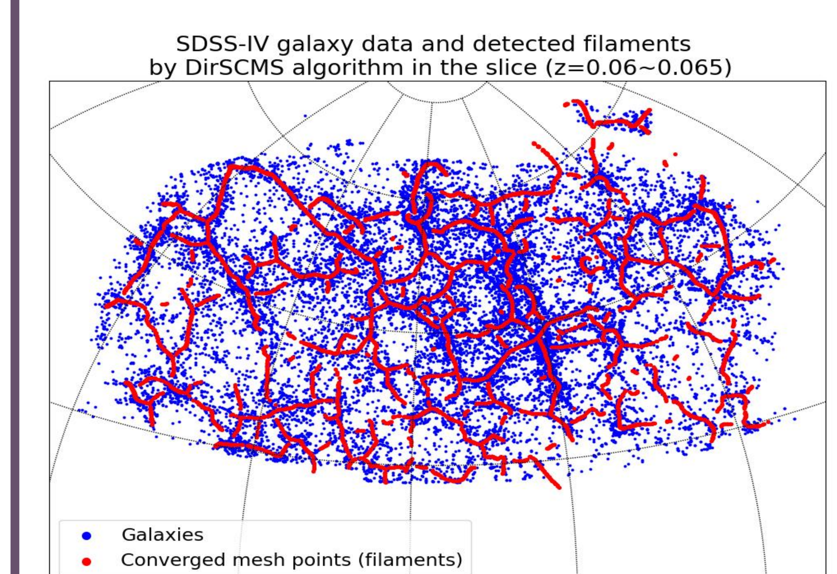
Step 4-2: Apply Our DirSCMS Algorithm (Iteration 1).



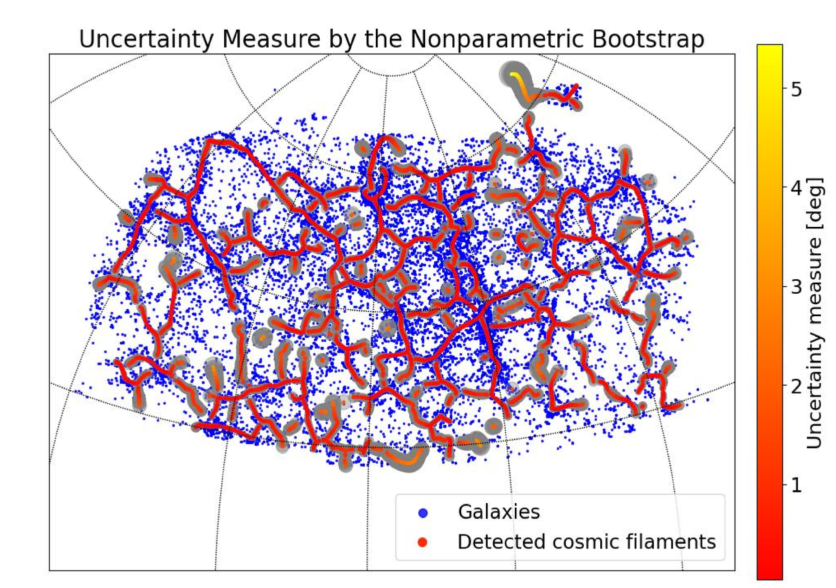
Step 4-2: Apply Our DirSCMS Algorithm (Iteration 2).



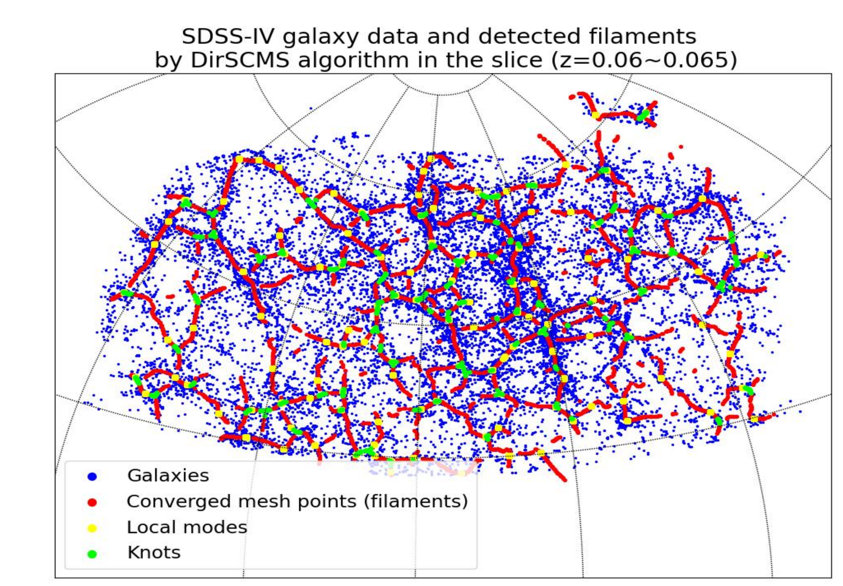
Step 4-2: Apply Our DirSCMS Algorithm (Iteration 3).



Step 4-2: Apply Our DirSCMS Algorithm (Final Convergence).



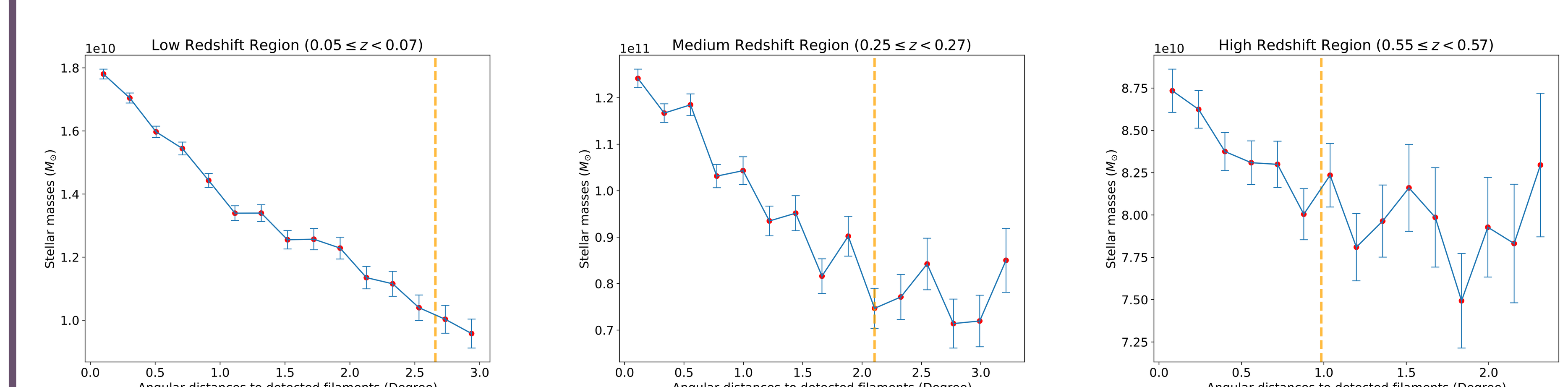
Step 5: Quantify the Uncertainties via Nonparametric Bootstrap.



Step 6: Estimate Local Modes and Filament Knots.[a]

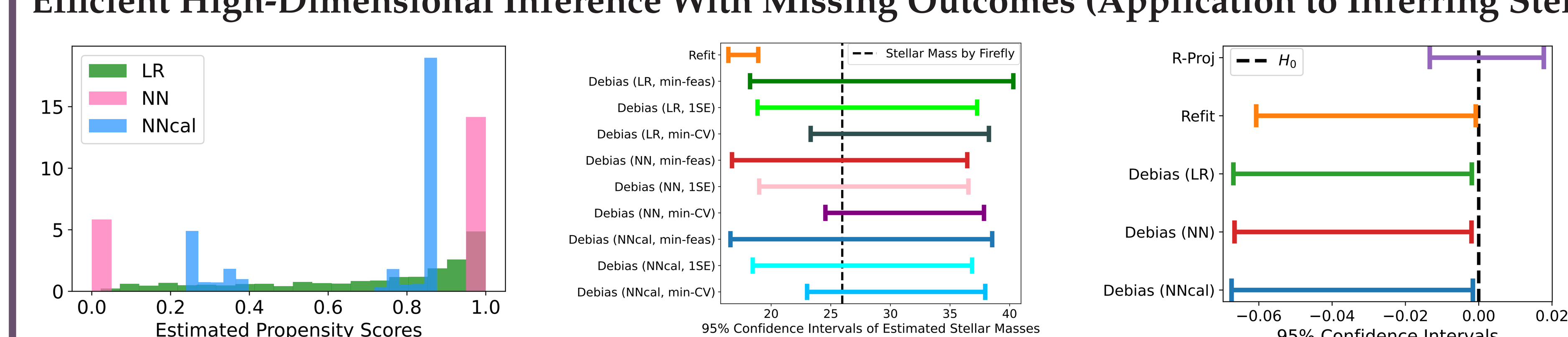
[a] The full catalog data can be downloaded at `https://doi.org/10.5281/zenodo.6244866`.

Effects of Cosmic Filaments on Galaxy Properties



These three figures present the correlations between stellar masses of galaxies in different redshift regions and their distances to our detected filaments. *It suggests that galaxies near filaments tend to be more massive than those that are further away from filaments.*

Efficient High-Dimensional Inference With Missing Outcomes (Application to Inferring Stellar Masses of SDSS Galaxies):



Middle Panel: Inference on stellar mass of a new galaxy based on its spectroscopic and photometric properties.
Right Panel: Testing the negative correlation between the stellar mass and its distance to nearby cosmic filaments.