Cosmic Filament Detection Under the Survey Geometry

Yikun Zhang[*]
(Joint work with *Yen-Chi Chen*[*]
and *Rafael S. de Souza*[†])

[*]Department of Statistics,
University of Washington
[†] Shanghai Astronomical Observatory

Venue: March 11 at Cosmic Cartography 2022

What Is Cosmic Web?

Cosmic Web is a large-scale network structure revealing that the matter in our Universe is not uniformly distributed (Zel'Dovich, 1970; Shandarin and Zeldovich, 1989; Bond et al., 1996).

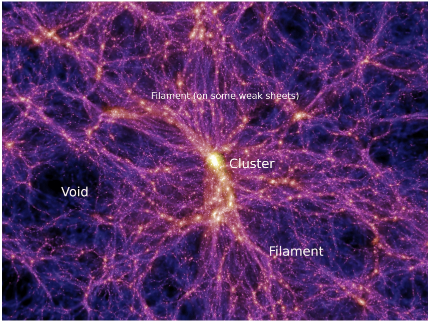


Figure 1: Characteristics of *Cosmic Web* (credited to the millennium simulation project (Springel et al., 2005)).

Focus of Our Research

In this talk, we present a novel methodology to detect cosmic filaments as well as the nodes on the filaments based on the newly released SDSS-IV galaxy observations.

- Our algorithm is adaptive to the survey geometry.
- Our filament model is statistically consistent.

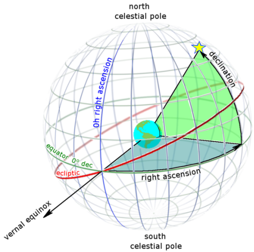


Figure 2: Illustration of right ascension (RA) and declination (DEC) (Image Courtesy of Wikipedia).

∗ Notice that each astronomical object has a coordinate (RA,DEC,Redshift) in the survey data, where (RA,DEC) encodes its position on the celestial sphere.

Significance of Cosmic Filaments

- They connect complexes of super-clusters (Lynden-Bell et al., 1988).
- They contain information about the global cosmology and the nature of dark matter (Zhang et al., 2009; Tempel et al., 2014).
- The trajectory of cosmic microwave background light can be distorted due to cosmic filaments, creating the weak lensing effect.

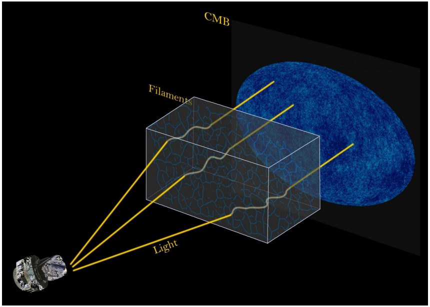


Figure 3: Illustration of the bending trajectory of CMB lights (credit to Siyu He, Shadab Alam, Wei Chen, and Planck/ESA; see He et al. (2018) for details).

Previous Works of Filament Detection on Survey Data

In astronomical survey data, such as SDSS or the Dark Energy Survey, the positions of observed objects are recorded as

$$\{(\alpha_1, \delta_1, Z_1), ..., (\alpha_n, \delta_n, Z_n)\},$$

where, for $i = 1, ..., n$,

- $\alpha_i \in [0, 360°)$ is the *right ascension* (RA), i.e., celestial longitude,
- $\eta_i \in [-90°, 90°]$ is the *declination* (DEC), i.e., celestial latitude,
- $Z_i \in (0, \infty)$ is the *redshift* value.

The existing filament detection methods applied to survey data come from two different categories:

- **3D method**: Convert redshifts into (comoving) distances (Tempel et al., 2014; Sousbie et al., 2011).
- **2D method**: Slice the Universe into thin redshift slices (Chen et al., 2015b; Duque et al., 2021).

Our method can easily switch between the above two categories.

Slicing the Universe (Tomographic Analysis)

The tomographic filament detection has its own advantages over 3D methods:

- It controls the redshift distortions along the line-of-sight direction (i.e., the *finger-of-god* effect).
- The measurement error in one slice won't propagate to other slices.
- It helps reduce computational cost...

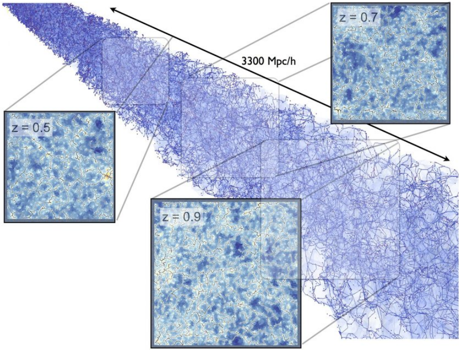


Figure 4: Illustration of slicing the Universe (credit to Laigle et al. 2018)

Previous 2D Methods on Survey Data

With each slice, says $z = 0.470$–0.475,

- the redshift values of observed objects are thought to be identical.
- the locations of these objects, encoded by their (RAs, DECs), are considered in a flat Euclidean space.

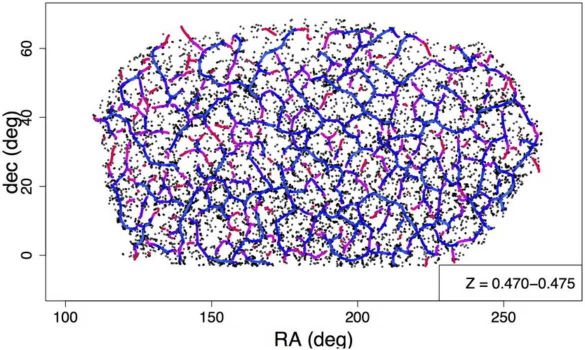
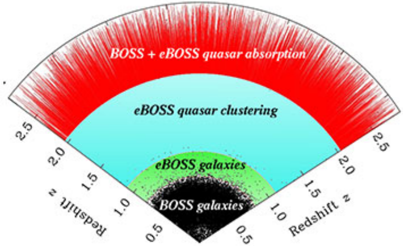


Figure 5: Cosmic filaments via density ridges on a 2D slice (Chen et al., 2015b, 2016)

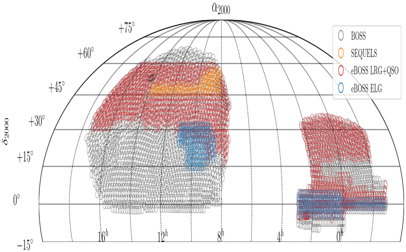
Problems of Existing Methods on 2D Slices

The slices ($\Delta z = 0.005$) in the survey data are not some flat 2D planes, but some **spherical shells**, which have a *nonlinear* curvature!

- Recall that the locations of astronomical objects in a slice are recorded by $\{(\alpha_i, \delta_i)\}_{i=1}^{n}$ on a celestial sphere.



(a) Planned eBOSS coverage of the Universe (credit to M. Blanton and SDSS)



(b) BOSS/eBOSS Spectroscopic Footprint as of DR16 (credit to SDSS)

Why can't we ignore the spherical geometry? (I)

Setup: Suppose that we want to recover the true ring/filament structure across the North and South pole of a unit sphere given some noisy data points from it.

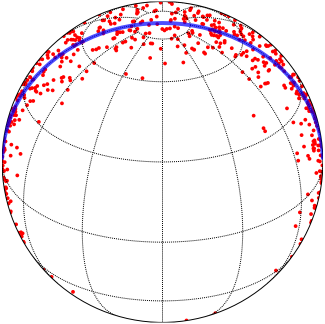
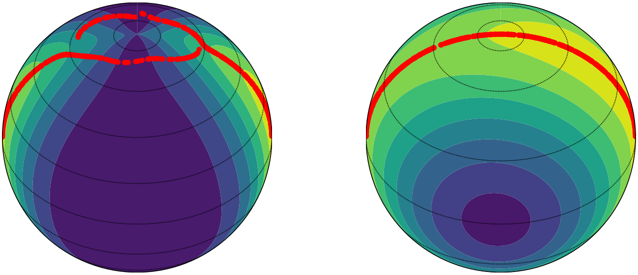


Figure 7: Noisy observations (red points) and the underlying true ring/filament structure (blue line)

Why can't we ignore the spherical geometry? (III)

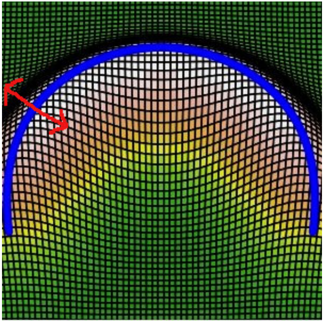
The background contour plots are kernel density estimators on the flat plane $[-90°, 90°] \times [0°, 360°)$ and unit sphere $\Omega_2 = \{\boldsymbol{x} \in \mathbb{R}^3 : ||\boldsymbol{x}||_2 = 1\}$, respectively.



(a) Euclidean SCMS Method. (b) Directional SCMS Method.

$*$ SCMS: subspace constrained mean shift (Ozertem and Erdogmus, 2011).

(Directional) density ridges are generalized local maxima (within some subspaces) of the underlying density function (on Ω_q).



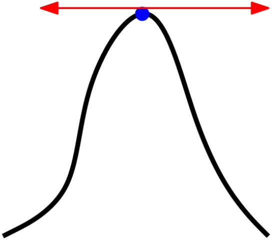


Figure 9: Density ridge (lifted onto the underlying density function; Chen et al. 2015a)

Formal Definitions of Directional Density Ridges

Under our scenario of detecting cosmic filaments within a spherical (redshift) slice, $q = 2$ and $d = 1$.

- A smooth density function $f : \Omega_q \to \mathbb{R}$. ($q = 2$ in a spherical slice.)
- Riemannian gradient $\texttt{grad} f(\boldsymbol{x})$ and Riemannian Hessian $\mathcal{H}f(\boldsymbol{x})$.
- Denote $V_d(\boldsymbol{x}) = \left[\boldsymbol{v}_{d+1}(\boldsymbol{x}), ..., \boldsymbol{v}_q(\boldsymbol{x})\right] \in \mathbb{R}^{(q+1)\times(q-d)}$ with columns as the second to the last eigenvectors of $\mathcal{H}f(\boldsymbol{x})$ lying within the tangent space $T_{\boldsymbol{x}}$ at $\boldsymbol{x} \in \Omega_q$.

$\Longrightarrow$

Local modes of f on Ω_q:

$$\mathcal{M} \equiv \texttt{Mode}(f) = \left\{\boldsymbol{x} \in \Omega_q : \texttt{grad} f(\boldsymbol{x}) = \mathbf{0}, \lambda_1(\boldsymbol{x}) < 0\right\}$$

Order-d density ridge on Ω_q (or directional density ridge) of f:

$$\mathcal{R}_d \equiv \texttt{Ridge}(f) = \left\{\boldsymbol{x} \in \Omega_q : V_d(\boldsymbol{x}) V_d(\boldsymbol{x})^T \texttt{grad} f(\boldsymbol{x}) = \mathbf{0}, \lambda_{d+1}(\boldsymbol{x}) < 0\right\}.$$

$*$ Note that the Riemannian Hessian $\mathcal{H}f(\boldsymbol{x})$ has a unit eigenvector $\boldsymbol{x}$ that is orthogonal to $T_{\boldsymbol{x}}$ and corresponds to eigenvalue 0.

Question: How can we recover the directional density ridges (or equivalently, cosmic filaments) from some discrete observations?

1. Density Estimation: We estimate the galaxy distribution via the *directional* kernel density estimator (KDE; Hall et al. 1987; Bai et al. 1988).

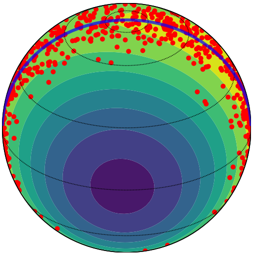


Figure 10: Counter plot of directional KDE

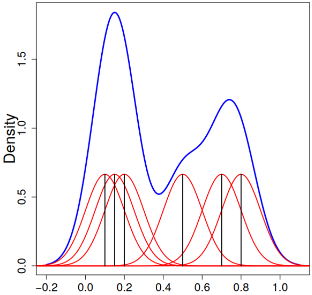


Figure 11: Illustration of one-dimensional KDE (Chen, 2017)

DirSCMS Algorithm

2. Filament Estimation: We propose the directional subspace constrained mean shift (DirSCMS) algorithm (Zhang and Chen, 2021c), which iterates a point on Ω_2 along the (subspace constrained) *gradient* of directional KDE.

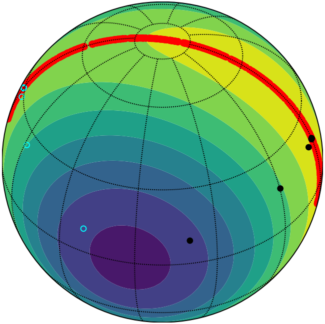
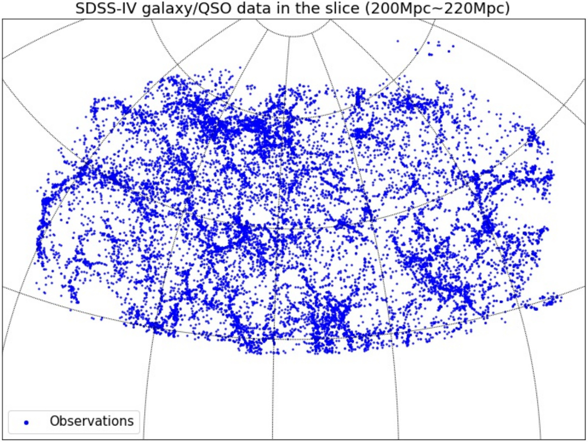


Figure 12: Two DirSCMS iterative paths

Cosmic Filament Detection on SDSS-IV Galaxy Data

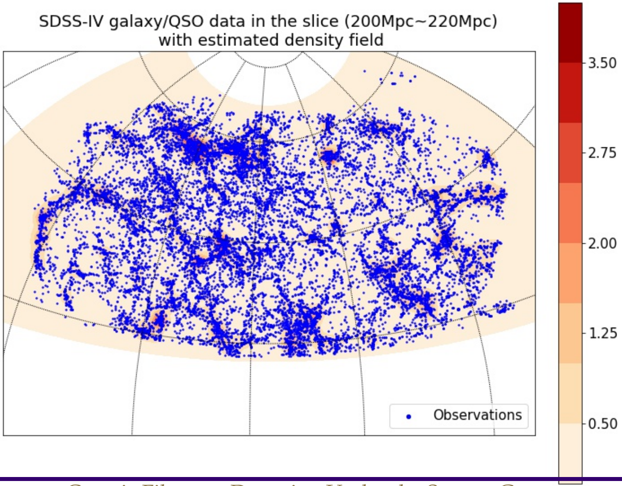
Step 1 (Slicing the Universe): Partition the redshift range into 325 spherical slices based on the comoving distance $\Delta L = 20\,\text{Mpc}$.

- Within each slice, we consider the redshifts of galaxies to be the same so that the galaxies are located on a sphere.



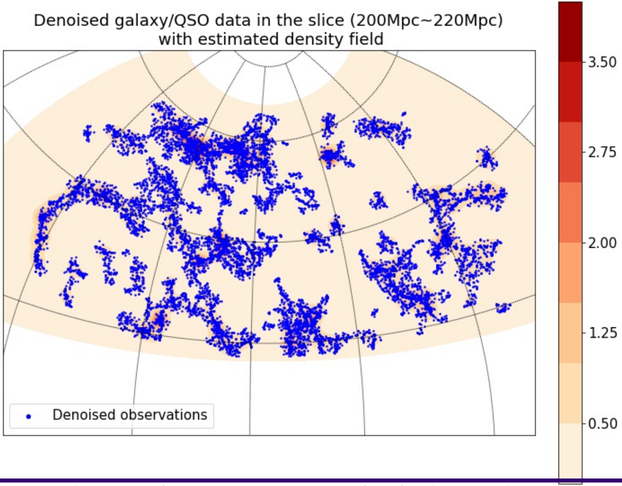
Step 2 (Density Estimation): Estimate the galaxy density field via directional KDE.

- The bandwidth parameter is selected in a data-adaptive approach.

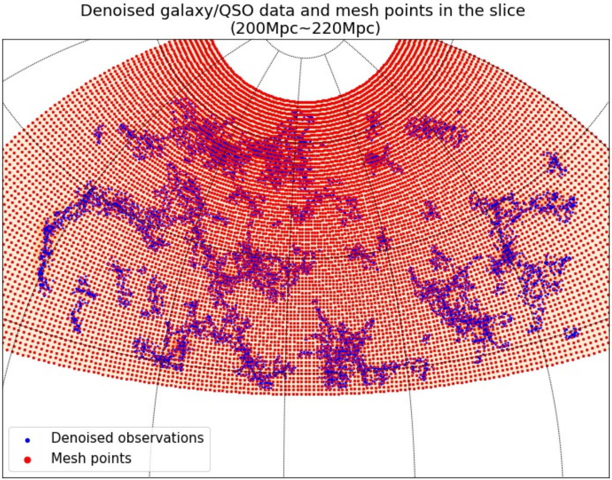


Step 3 (Denoising): Remove the observations with low-density values.

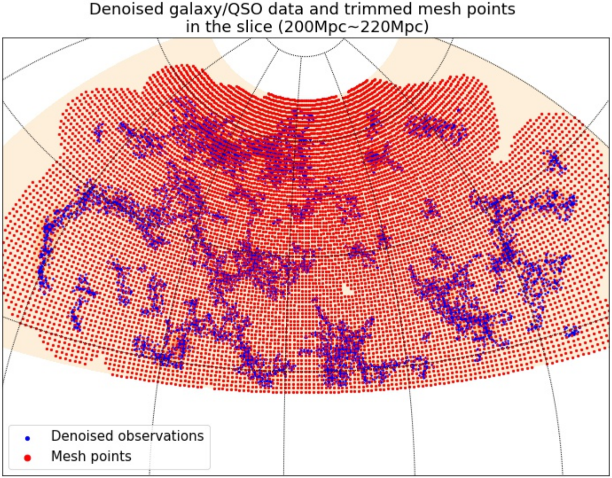
- We keep at least 80% of the original galaxy data in the slice.



Step 4 (Laying Down the Mesh Points): We place a set of dense mesh points on the interested region, which are the initial points of our DirSCMS iterations.



Step 5 (Thresholding the Mesh Points): We discard those mesh points with low-density values and keep 85% of the original mesh points.



Step 6 (DirSCMS Iterations): We iterate our DirSCMS algorithm on each remaining mesh point until convergence.

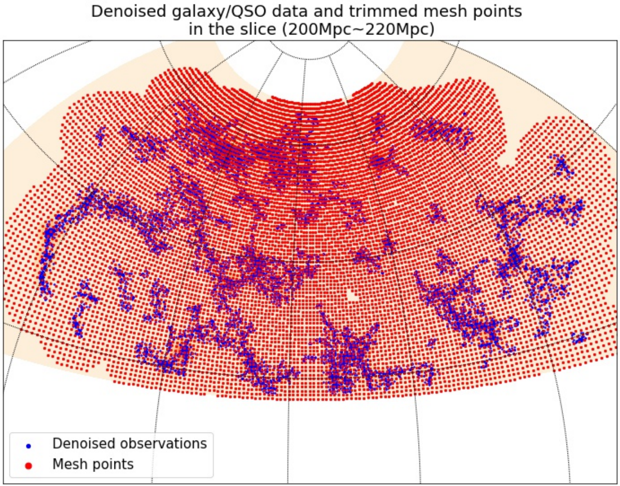


Figure 13: DirSCMS Iterations (Step 0).

Step 6 (DirSCMS Iterations): We iterate our DirSCMS algorithm on each remaining mesh point until convergence.

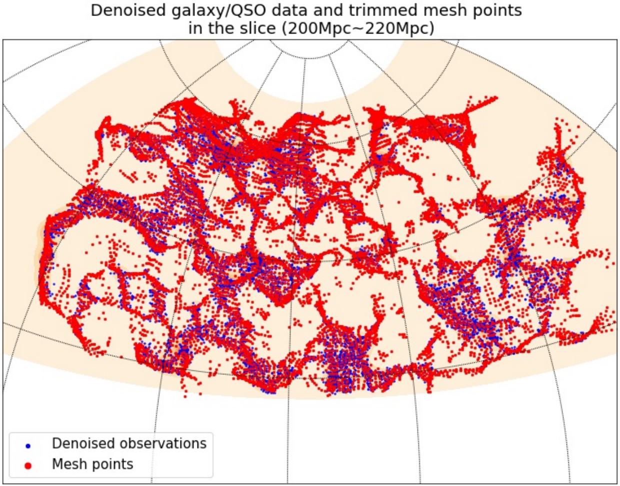


Figure 13: DirSCMS Iterations (Step 1).

Step 6 (DirSCMS Iterations): We iterate our DirSCMS algorithm on each remaining mesh point until convergence.

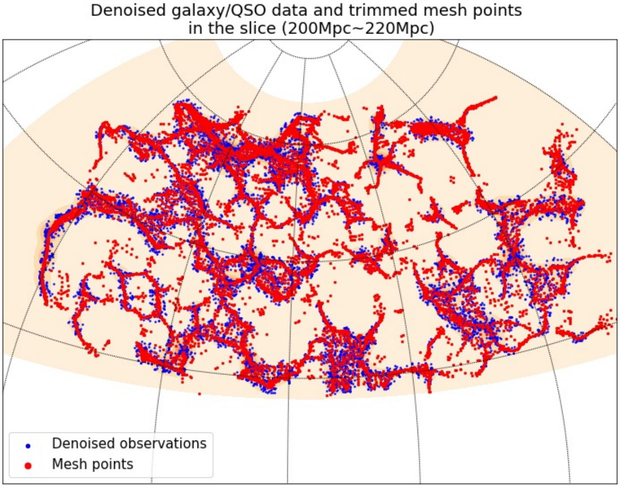


Figure 13: DirSCMS Iterations (Step 2).

Step 6 (DirSCMS Iterations): We iterate our DirSCMS algorithm on each remaining mesh point until convergence.

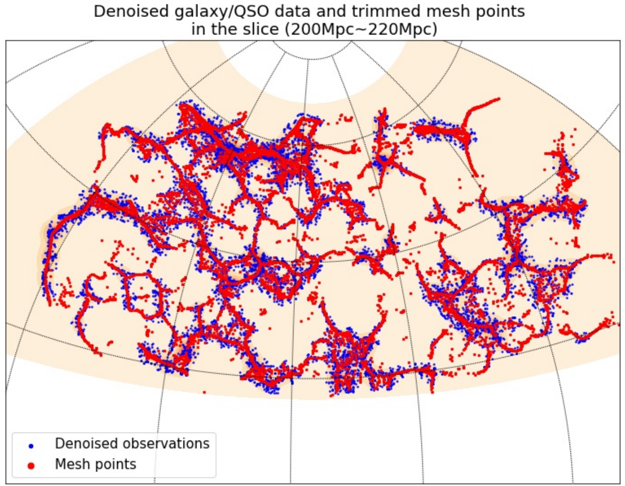


Figure 13: DirSCMS Iterations (Step 3).

Step 6 (DirSCMS Iterations): We iterate our DirSCMS algorithm on each remaining mesh point until convergence.

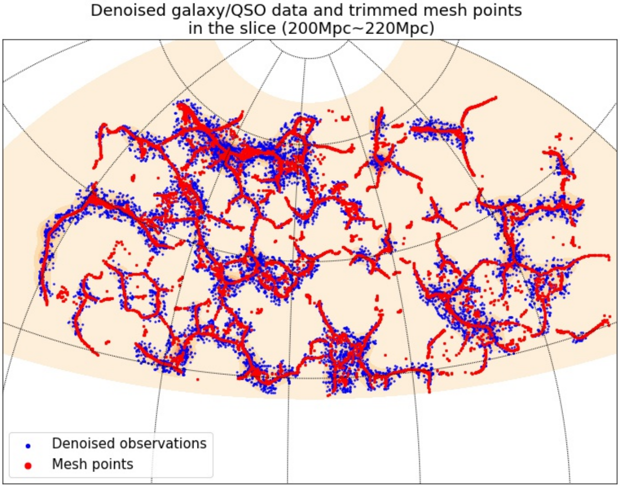


Figure 13: DirSCMS Iterations (Step 5).

Step 6 (DirSCMS Iterations): We iterate our DirSCMS algorithm on each remaining mesh point until convergence.

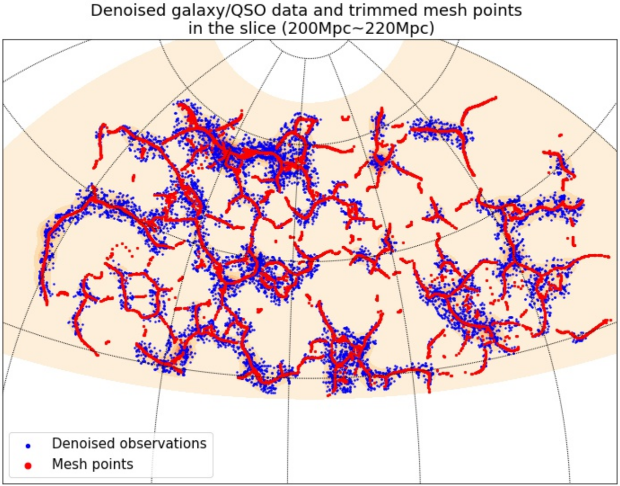


Figure 13: DirSCMS Iterations (Step 8).

Step 6 (DirSCMS Iterations): We iterate our DirSCMS algorithm on each remaining mesh point until convergence.

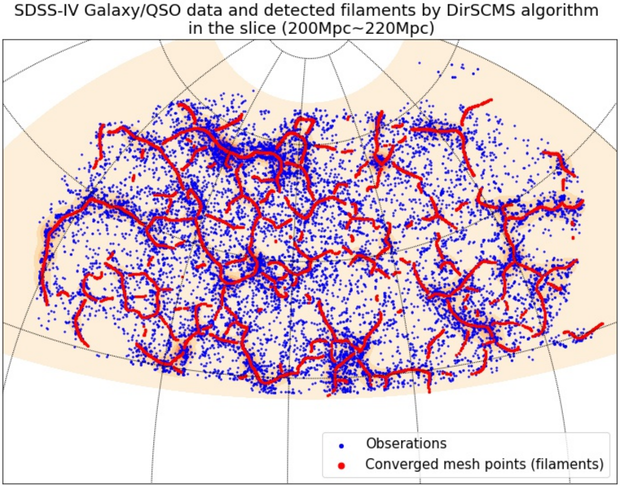


Figure 13: DirSCMS Iterations (Final).

Step 7 (Mode and Knot Estimation): We seek out the local modes and knots on the filaments as cosmic nodes.

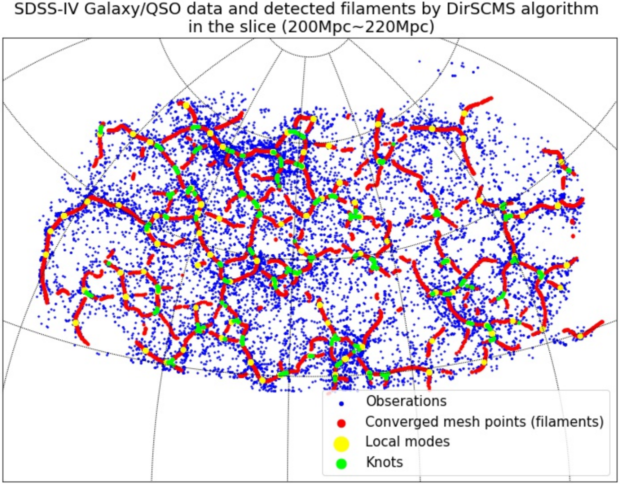


Figure 14: Nodes on the detected filaments.

Recall that the survey data $\{(\alpha_i, \delta_i, Z_i)\}_{i=1}^n \in \Omega_2 \times \mathbb{R}^+$ is directional-linear.

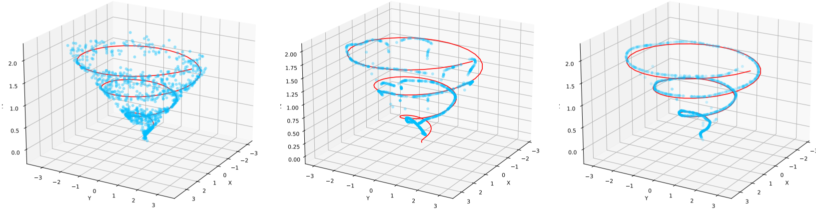
- We consider extending our DirSCMS algorithm to estimate the cosmic filaments (*i.e.*, density ridges) in a directional-linear product space (Zhang and Chen, 2021a).
- We adopt the directional-linear KDE (García-Portugués et al., 2015) with $\boldsymbol{X}_i \in \Omega_2$ being the Cartesian coordinate of (ϕ_i, η_i) for $i = 1, ..., n$:

$$\widehat{f}_{\boldsymbol{h}}(\boldsymbol{x}, \boldsymbol{z}) = \frac{C_{L,2}(h_1)}{nh_2} \sum_{i=1}^n L\left(\frac{1 - \boldsymbol{x}^T \boldsymbol{X}_i}{h_1^2}\right) K\left(\frac{z - Z_i}{h_2}\right)$$

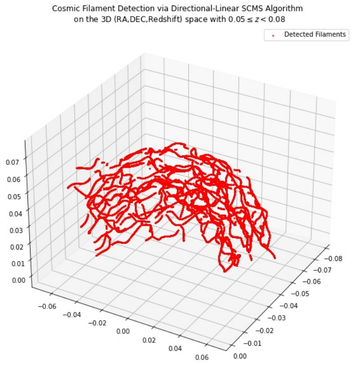
where $L(r) = e^{-r}$ and $K(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ are the kernel functions.

Filament Detection in the Directional-Linear Space (I)

Our directional-linear SCMS algorithm is stabler than its Euclidean prototype.



(a) Simulated data points. (b) Euclidean SCMS. (c) Directional-linear SCMS.



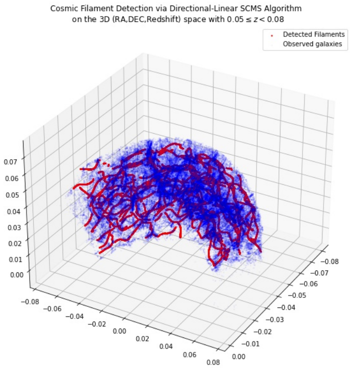


Figure 16: Cosmic filament detection in the 3D (RA,DEC,Redshift) space with our directional-linear SCMS algorithm.

Filament Effects on Galaxy Properties

1. We compute the angular distance (or equivalently, *geodesic distance*) of each observed galaxy in the redshift range $0.05 \leq z < 0.7$ to our detected filaments in the corresponding slice.

2. We obtain the galaxy properties, such as stellar mass and metallicity, from the FIREFLY value-added catalog (Wilkinson et al., 2017; Maraston and Strömbäck, 2011).

3. Our subsequent analyses focus on the following three regions:
   - **Low redshift region**: $0.05 \leq z < 0.07$.
   - **Medium redshift region**: $0.25 \leq z < 0.27$.
   - **High redshift region**: $0.55 \leq z < 0.57$.

4. We partition the galaxies within each region into several bins according to their distances to our detected filaments.

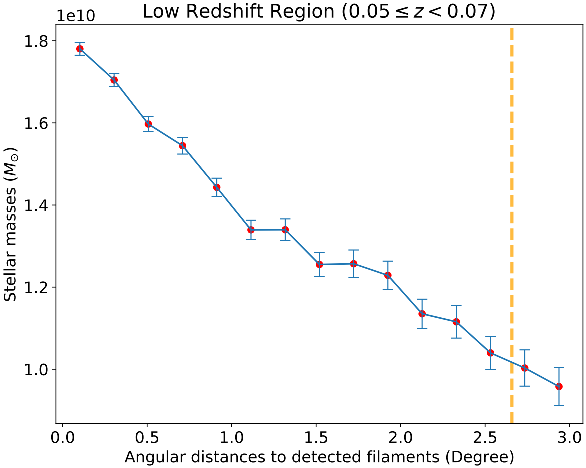


Figure 17: Comparison between stellar masses of galaxies and their distances to filaments (**Low redshift region**)

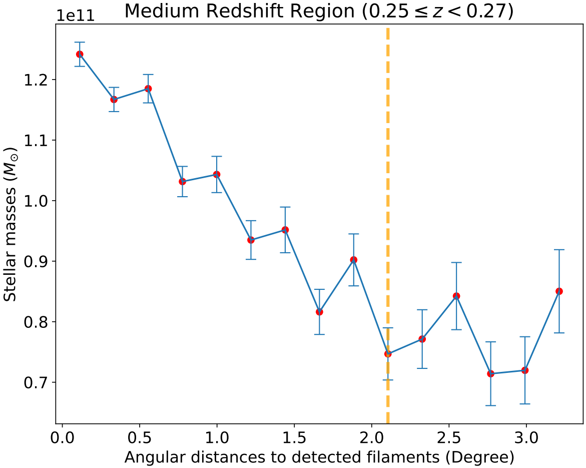


Figure 17: Comparison between stellar masses of galaxies and their distances to filaments (**Medium redshift region**)

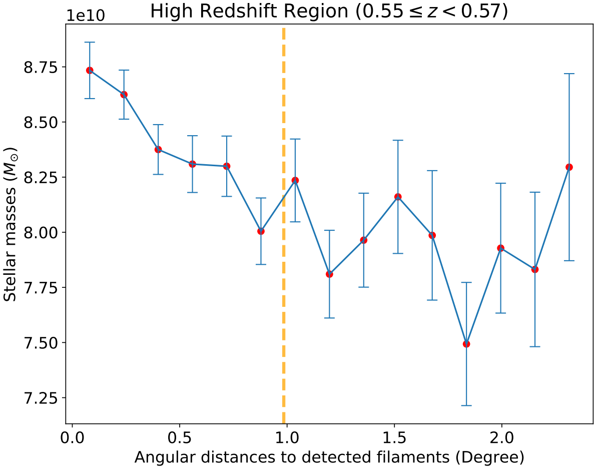


Figure 17: Comparison between stellar masses of galaxies and their distances to filaments (**High redshift region**)

Conclusions and Ongoing Works

In this talk, we discussed our methodology of recovering filament structures from some SDSS-IV galaxy data.

1. Our DirSCMS algorithm took into account the survey (spherical) geometry when estimating the filament structures.
2. We applied our method to the latest survey data (SDSS-IV, Data Release 16).
3. Our analyses reveal some signals that galaxies near the filaments are heavier in their stellar masses.

We are planning to

- Release a comprehensive cosmic web catalog.
- Analyze if other galaxy properties are correlated by cosmic web structures.
- ...

Thank you!

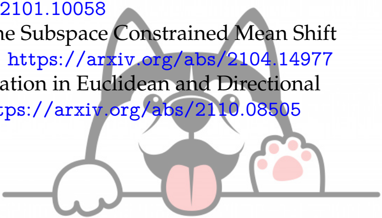
More details can be found in

[1] Y. Zhang and Y.-C. Chen. Kernel Smoothing, Mean Shift, and Their Learning Theory with Directional Data. *Journal of Machine Learning Research*, 22(154):1–92, 2021. https://arxiv.org/abs/2010.13523

[2] Y. Zhang and Y.-C. Chen. The EM Perspective of Directional Mean Shift Algorithm. 2021. https://arxiv.org/abs/2101.10058

[3] Y. Zhang and Y.-C. Chen. Linear Convergence of the Subspace Constrained Mean Shift Algorithm: From Euclidean to Directional Data. 2021. https://arxiv.org/abs/2104.14977

[4] Y. Zhang and Y.-C. Chen. Mode and Ridge Estimation in Euclidean and Directional Product Spaces: A Mean Shift Approach. 2021. https://arxiv.org/abs/2110.08505



Reference I

Z. Bai, C. Rao, and L. Zhao. Kernel estimators of density function of directional data. *Journal of Multivariate Analysis*, 27(1):24 – 39, 1988.

J. R. Bond, L. Kofman, and D. Pogosyan. How filaments of galaxies are woven into the cosmic web. *Nature*, 380(6575):603–606, 1996.

Y.-C. Chen. A tutorial on kernel density estimation and recent advances. *Biostatistics & Epidemiology*, 1 (1):161–187, 2017.

Y.-C. Chen, C. R. Genovese, and L. Wasserman. Asymptotic theory for density ridges. *The Annals of Statistics*, 43(5):1896–1928, 2015a.

Y.-C. Chen, S. Ho, P. E. Freeman, C. R. Genovese, and L. Wasserman. Cosmic web reconstruction through density ridges: method and algorithm. *Monthly Notices of the Royal Astronomical Society*, 454 (1):1140–1156, 2015b.

Y.-C. Chen, S. Ho, A. Tenneti, R. Mandelbaum, R. Croft, T. DiMatteo, P. E. Freeman, C. R. Genovese, and L. Wasserman. Investigating galaxy-filament alignments in hydrodynamic simulations using density ridges. *Monthly Notices of the Royal Astronomical Society*, 454(3):3341–3350, 2015c.

Y.-C. Chen, S. Ho, J. Brinkmann, P. E. Freeman, C. R. Genovese, D. P. Schneider, and L. Wasserman. Cosmic web reconstruction through density ridges: catalogue. *Monthly Notices of the Royal Astronomical Society*, 461(4):3896–3909, 2016.

J. C. Duque, M. Migliaccio, D. Marinucci, and N. Vittorio. A novel cosmic filament catalogue from sdss data. *arXiv preprint arXiv:2106.05253*, 2021.

E. García-Portugués. Exact risk improvement of bandwidth selectors for kernel density estimation with directional data. *Electronic Journal of Statistics*, 7:1655–1685, 2013.

E. García-Portugués, R. M. Crujeiras, and W. González-Manteiga. Central limit theorems for directional and linear random variables with applications. *Statistica Sinica*, pages 1207–1229, 2015.

Reference II

P. Hall, G. S. Watson, and J. Cabrara. Kernel density estimation with spherical data. *Biometrika*, 74(4): 751–762, 12 1987. ISSN 0006-3444. URL https://doi.org/10.1093/biomet/74.4.751.

S. He, S. Alam, S. Ferraro, Y.-C. Chen, and S. Ho. The detection of the imprint of filaments on cosmic microwave background lensing. *Nature Astronomy*, 2(5):401–406, 2018.

C. Laigle, C. Pichon, S. Arnouts, H. J. Mccracken, Y. Dubois, J. Devriendt, A. Slyz, D. Le Borgne, A. Benoit-Levy, H. S. Hwang, et al. Cosmos2015 photometric redshifts probe the impact of filaments on galaxy properties. *Monthly Notices of the Royal Astronomical Society*, 474(4):5437–5458, 2018.

D. Lynden-Bell, S. Faber, D. Burstein, R. L. Davies, A. Dressler, R. Terlevich, and G. Wegner. Spectroscopy and photometry of elliptical galaxies. v-galaxy streaming toward the new supergalactic center. *The Astrophysical Journal*, 326:19–49, 1988.

C. Maraston and G. Strömbäck. Stellar population models at high spectral resolution. *Monthly Notices of the Royal Astronomical Society*, 418(4):2785–2811, 2011.

B. Moews, M. A. Schmitz, A. J. Lawler, J. Zuntz, A. I. Malz, R. S. de Souza, R. Vilalta, A. Krone-Martins, E. E. Ishida, and C. Collaboration. Ridges in the dark energy survey for cosmic trough identification. *Monthly Notices of the Royal Astronomical Society*, 500(1):859–870, 2021.

U. Ozertem and D. Erdogmus. Locally defined principal curves and surfaces. *Journal of Machine Learning Research*, 12(34):1249–1286, 2011.

S. F. Shandarin and Y. B. Zeldovich. The large-scale structure of the universe: Turbulence, intermittency, structures in a self-gravitating medium. *Reviews of Modern Physics*, 61(2):185, 1989.

T. Sousbie, C. Pichon, and H. Kawahara. The persistent cosmic web and its filamentary structure–ii. illustrations. *Monthly Notices of the Royal Astronomical Society*, 414(1):384–403, 2011.

Reference III

V. Springel, S. D. White, A. Jenkins, C. S. Frenk, N. Yoshida, L. Gao, J. Navarro, R. Thacker, D. Croton, J. Helly, et al. Simulations of the formation, evolution and clustering of galaxies and quasars. *nature*, 435(7042):629–636, 2005.

E. Tempel, R. Stoica, V. J. Martinez, L. Liivamägi, G. Castellan, and E. Saar. Detecting filamentary pattern in the cosmic web: a catalogue of filaments for the sdss. *Monthly Notices of the Royal Astronomical Society*, 438(4):3465–3482, 2014.

D. M. Wilkinson, C. Maraston, D. Goddard, D. Thomas, and T. Parikh. Firefly (fitting iteratively for likelihood analysis): a full spectral fitting code. *Monthly Notices of the Royal Astronomical Society*, 472 (4):4297–4326, 2017.

Y. B. Zel'Dovich. Gravitational instability: An approximate theory for large density perturbations. *Astronomy and astrophysics*, 5:84–89, 1970.

Y. Zhang and Y.-C. Chen. Mode and ridge estimation in euclidean and directional product spaces: A mean shift approach. *arXiv preprint arXiv:2110.08505*, 2021a. URL https://arxiv.org/abs/2110.08505.

Y. Zhang and Y.-C. Chen. Kernel smoothing, mean shift, and their learning theory with directional data. *Journal of Machine Learning Research*, 22(154):1–92, 2021b.

Y. Zhang and Y.-C. Chen. Linear convergence of the subspace constrained mean shift algorithm: From euclidean to directional data. *arXiv preprint arXiv:2104.14977*, 2021c. URL https://arxiv.org/abs/2104.14977.

Y. Zhang, X. Yang, A. Faltenbacher, V. Springel, W. Lin, and H. Wang. The spin and orientation of dark matter halos within cosmic filaments. *The Astrophysical Journal*, 706(1):747, 2009.

Riemannian Gradient and Hessian on Ω_q

Assume tentatively that the directional function f is well-defined and smooth in $\mathbb{R}^{q+1} \setminus \{\mathbf{0}\}$ (or at least in an open neighborhood $U \supset \Omega_q$).

- *Riemannian gradient* $\texttt{grad} f(\boldsymbol{x})$ on Ω_q:

$$\texttt{grad} f(\boldsymbol{x}) = \left(\boldsymbol{I}_{q+1} - \boldsymbol{x}\boldsymbol{x}^T\right) \nabla f(\boldsymbol{x}),$$

  where $\boldsymbol{I}_{q+1}$ is the identity matrix in $\mathbb{R}^{(q+1)\times(q+1)}$.

- *Riemannian Hessian* $\mathcal{H}f(\boldsymbol{x})$ on Ω_q (Zhang and Chen, 2021b):

$$\mathcal{H}f(\boldsymbol{x}) = (\boldsymbol{I}_{q+1} - \boldsymbol{x}\boldsymbol{x}^T) \left[\nabla\nabla f(\boldsymbol{x}) - \nabla f(\boldsymbol{x})^T \boldsymbol{x} \cdot \boldsymbol{I}_{q+1}\right] (\boldsymbol{I}_{q+1} - \boldsymbol{x}\boldsymbol{x}^T).$$

Here, $\boldsymbol{I}_{q+1}$ is the identity matrix in $\mathbb{R}^{(q+1)\times(q+1)}$, while $\nabla f(\boldsymbol{x})$ and $\nabla\nabla f(\boldsymbol{x})$ are total gradient and Hessian in $\mathbb{R}^{q+1}$.

Formal Definition of Directional KDE

Directional kernel density estimator (KDE; Hall et al. 1987; Bai et al. 1988; García-Portugués 2013):

$$\widehat{f}_h(\boldsymbol{x}) = \frac{c_{L,q}(h)}{n} \sum_{i=1}^{n} L\left(\frac{1 - \boldsymbol{x}^T \boldsymbol{X}_i}{h^2}\right). \quad (1)$$

- $\boldsymbol{X}_1, ..., \boldsymbol{X}_n \in \Omega_q \subset \mathbb{R}^{q+1}$ are directional random observations.
- L is a directional kernel, *i.e.*, a rapidly decaying function with nonnegative values on $[0, \infty)$.
- $h > 0$ is the bandwidth parameter.
- $c_{L,q}(h)$ is a normalizing constant satisfying

$$c_{L,q}(h)^{-1} = \int_{\Omega_q} L\left(\frac{1 - \boldsymbol{x}^T \boldsymbol{y}}{h^2}\right) \omega_q(d\boldsymbol{y}) = h^q \lambda_{h,q}(L) \asymp h^q \lambda_q(L) \quad (2)$$

with $\lambda_q(L) = 2^{\frac{q}{2}-1} \omega_{q-1} \int_0^\infty L(r) r^{\frac{q}{2}-1} dr$.

An Example of the Directional Kernel

Under the von Mises kernel $L(r) = e^{-r}$,

- directional KDE $\widehat{f}_h(\boldsymbol{x}) = \frac{c_{L,q}(h)}{n} \sum_{i=1}^{n} L\left(\frac{1-\boldsymbol{x}^T \boldsymbol{X}_i}{h^2}\right)$

  becomes

- a mixture of von Mises-Fisher densities:

$$\begin{aligned} \widehat{f}_h(\boldsymbol{x}) &= \frac{1}{n} \sum_{i=1}^{n} f_{\text{vMF}}\left(\boldsymbol{x}; \boldsymbol{X}_i, \frac{1}{h^2}\right) \\ &= \frac{1}{n(2\pi)^{\frac{q+1}{2}} \mathcal{I}_{\frac{q-1}{2}}(1/h^2) h^{q-1}} \sum_{i=1}^{n} \exp\left(\frac{\boldsymbol{x}^T \boldsymbol{X}_i}{h^2}\right). \end{aligned}$$

Input:

- A directional data sample $\boldsymbol{X}_1, ..., \boldsymbol{X}_n \sim f(\boldsymbol{x})$ on Ω_q
- The order d of the directional ridge, smoothing bandwidth $h > 0$, and tolerance level $\epsilon > 0$.
- A suitable mesh $\mathcal{M}_D \subset \Omega_q$ of initial points.

Step 1: Compute the directional KDE $\widehat{f}_h(\boldsymbol{x}) = \frac{c_{L,q}(h)}{n} \sum_{i=1}^{n} L\left(\frac{1-\boldsymbol{x}^T\boldsymbol{X}_i}{h^2}\right)$ on the mesh $\mathcal{M}_D$.

Step 2: For each $\widehat{\boldsymbol{x}}^{(0)} \in \mathcal{M}_D$, iterate the following DirSCMS update until convergence:

while $\left|\left|\sum_{i=1}^{n} \widehat{V}_d(\widehat{\boldsymbol{x}}^{(0)})\widehat{V}_d(\widehat{\boldsymbol{x}}^{(0)})^T \boldsymbol{X}_i \cdot L'\left(\frac{1-\boldsymbol{X}_i^T\widehat{\boldsymbol{x}}^{(0)}}{h^2}\right)\right|\right|_2 > \epsilon$ **do**:

Detailed Procedures of DirSCMS Algorithm II

- **Step 2-1**: Compute the scaled version of the estimated Hessian matrix as:

$$\frac{nh^2}{c_{L,q}(h)} \mathcal{H}\widehat{f}_h(\widehat{\boldsymbol{x}}^{(t)}) = \left[\boldsymbol{I}_{q+1} - \widehat{\boldsymbol{x}}^{(t)} \left(\widehat{\boldsymbol{x}}^{(t)}\right)^T\right] \left[\frac{1}{h^2} \sum_{i=1}^{n} \boldsymbol{X}_i \boldsymbol{X}_i^T \cdot L''\left(\frac{1-\boldsymbol{X}_i^T \widehat{\boldsymbol{x}}^{(t)}}{h^2}\right) + \sum_{i=1}^{n} \boldsymbol{X}_i^T \widehat{\boldsymbol{x}}^{(t)} \boldsymbol{I}_{q+1} \cdot L'\left(\frac{1-\boldsymbol{X}_i^T \widehat{\boldsymbol{x}}^{(t)}}{h^2}\right)\right] \left[\boldsymbol{I}_{q+1} - \widehat{\boldsymbol{x}}^{(t)} \left(\widehat{\boldsymbol{x}}^{(t)}\right)^T\right].$$

- **Step 2-2**: Perform the spectral decomposition on $\frac{nh^2}{c_{L,q}(h)} \mathcal{H}\widehat{f}_h\left(\widehat{\boldsymbol{x}}^{(t)}\right)$ and compute $\widehat{V}_d(\widehat{\boldsymbol{x}}^{(t)}) = \left[\boldsymbol{v}_{d+1}(\widehat{\boldsymbol{x}}^{(t)}), ..., \boldsymbol{v}_q(\widehat{\boldsymbol{x}}^{(t)})\right]$, whose columns are orthonormal eigenvectors corresponding to the smallest $q-d$ eigenvalues inside the tangent space $T_{\widehat{\boldsymbol{x}}^{(t)}}$.

- **Step 2-3**: Update

$$\widehat{\boldsymbol{x}}^{(t+1)} \leftarrow \widehat{\boldsymbol{x}}^{(t)} - \widehat{V}_d(\widehat{\boldsymbol{x}}^{(t)})\widehat{V}_d(\widehat{\boldsymbol{x}}^{(t)})^T \left[\frac{\sum_{i=1}^n \boldsymbol{X}_i L'\left(\frac{1-\boldsymbol{X}_i^T\widehat{\boldsymbol{x}}^{(t)}}{h^2}\right)}{\sum_{i=1}^n \boldsymbol{X}_i L'\left(\frac{1-\boldsymbol{X}_i^T\widehat{\boldsymbol{x}}^{(t)}}{h^2}\right)}\right].$$

- **Step 2-4**: Standardize $\widehat{\boldsymbol{x}}^{(t+1)}$ as $\widehat{\boldsymbol{x}}^{(t+1)} \leftarrow \frac{\widehat{\boldsymbol{x}}^{(t+1)}}{\left|\left|\widehat{\boldsymbol{x}}^{(t+1)}\right|\right|_2}$.

Output: An estimated directional d-ridge $\widehat{\mathcal{R}}_d$ represented by the collection of resulting points.

- Recall that the directional-linear KDE at $(\boldsymbol{x}, z) \in \Omega_2 \times \mathbb{R}$ is defined as:

$$\widehat{f}_h(\boldsymbol{x}, z) = \frac{C_{L,2}(h_1)}{nh_2} \sum_{i=1}^{n} L\left(\frac{1-\boldsymbol{x}^T\boldsymbol{X}_i}{h_1^2}\right) K\left(\frac{z-Z_i}{h_2}\right).$$

- Directional-linear mean shift iteration:

$$\begin{aligned}\left(\boldsymbol{x}^{(t+1)}, z^{(t+1)}\right)^T &\leftarrow \Xi(\boldsymbol{x}^{(t)}, z^{(t)}) + \left(\boldsymbol{x}^{(t)}, z^{(t)}\right)^T \\ &= \begin{pmatrix} \dfrac{\sum_{i=1}^{n} \boldsymbol{X}_i \cdot L'\left(\frac{1-\boldsymbol{X}_i^T\boldsymbol{x}^{(t)}}{h_1}\right) K\left(\frac{z^{(t)}-Z_i}{h_2}\right)}{\sum_{i=1}^{n} L'\left(\frac{1-\boldsymbol{X}_i^T\boldsymbol{x}^{(t)}}{h_1}\right) K\left(\frac{z^{(t)}-Z_i}{h_2}\right)} \\ \dfrac{\sum_{i=1}^{n} Z_i \cdot L\left(\frac{1-\boldsymbol{X}_i^T\boldsymbol{x}^{(t)}}{h_1}\right) K\left(\left|\left|\frac{z^{(t)}-Z_i}{h_2}\right|\right|_2^2\right)}{\sum_{i=1}^{n} L\left(\frac{1-\boldsymbol{X}_i^T\boldsymbol{x}^{(t)}}{h_1}\right) K\left(\left|\left|\frac{z^{(t)}-Z_i}{h_2}\right|\right|_2^2\right)} \end{pmatrix}\end{aligned}$$

with an extra standardization $\boldsymbol{x}^{(t+1)} \leftarrow \frac{\boldsymbol{x}^{(t+1)}}{\left|\left|\boldsymbol{x}^{(t+1)}\right|\right|_2}$.

- Directional-linear SCMS algorithm iteration at $\boldsymbol{y}^{(t)} = \left(\boldsymbol{x}^{(t+1)}, z^{(t+1)}\right)^T$:

$$\boldsymbol{y}^{(t)} \leftarrow \boldsymbol{y}^{(t)} + \eta \cdot \widehat{V}_d(\boldsymbol{y}^{(t)})\widehat{V}_d(\boldsymbol{y}^{(t)})^T \boldsymbol{H}^{-1}\Xi(\boldsymbol{y}^{(t)}),$$

where $\boldsymbol{H} = \mathtt{Diag}(h_1^2, h_1^2, h_2^2) \in \mathbb{R}^{3\times 3}$ is a diagonal matrix and

$$\Xi(\boldsymbol{y}^{(t)}) = \Xi(\boldsymbol{x}^{(t)}, z^{(t)}) = \begin{pmatrix} \dfrac{\sum_{i=1}^n \boldsymbol{X}_i \cdot L'\left(\frac{1-\boldsymbol{X}_i^T\boldsymbol{x}^{(t)}}{h_1}\right) K\left(\frac{z^{(t)}-Z_i}{h_2}\right)}{\sum_{i=1}^n L'\left(\frac{1-\boldsymbol{X}_i^T\boldsymbol{x}^{(t)}}{h_1}\right) K\left(\frac{z^{(t)}-Z_i}{h_2}\right)} - \boldsymbol{x}^{(t)} \\ \dfrac{\sum_{i=1}^n Z_i \cdot L\left(\frac{1-\boldsymbol{X}_i^T\boldsymbol{x}^{(t)}}{h_1}\right) K\left(\left|\left|\frac{z^{(t)}-Z_i}{h_2}\right|\right|_2^2\right)}{\sum_{i=1}^n L\left(\frac{1-\boldsymbol{X}_i^T\boldsymbol{x}^{(t)}}{h_1}\right) K\left(\left|\left|\frac{z^{(t)}-Z_i}{h_2}\right|\right|_2^2\right)} - z^{(t)} \end{pmatrix}.$$

Here, we design a theoretically motivated and empirically effective step size as $\eta = \min\{h_1 h_2, 1\}$.

$*$ Notes: A naive generalization of SCMS algorithm $\boldsymbol{y}^{(t+1)} \leftarrow \boldsymbol{y}^{(t)} + \widehat{V}_d(\boldsymbol{y}^{(t)})\widehat{V}_d(\boldsymbol{y}^{(t)})^T\Xi(\boldsymbol{y}^{(t)})$ plus standardization as with pure Euclidean/directional data does not work (Zhang and Chen, 2021a)!