IDENTIFICATION OF SHELL AND STREAM SUBSTRUCTURES IN GALAXY DEBRIS

RAHUL BISWAS, DAVID HENDEL, KATHRYN V. JOHNSTON, ROHIT K. PATRA, BODHISATTVA SEN.

Simulation studies suggest that larger galaxies are formed through accretion of smaller galaxies. The process of accretion creates debris that can be detected though present (Sloan Digital Sky Survey) and future (Large Synoptic Survey Telescope) telescopes. Even through the debris contain only ∼1% of the total mass they are spread over volume that is orders of magnitude larger and follow elegant and simple physics. Studying these structures will provide answers to fundamental problems such as (a) how often do the smaller galaxies merge into larger ones, and (b) what is the structure or locus of the debris.

Traditionally astronomers have relied on predictions from simulations of how these debris structures form and behave. In this project, we develop a method to identify the two substructures — shells and streams — from two-dimensional projected positions of a data cloud of stars. We first estimate the density of the two-dimensional position of the stars using a kernel density estimate (KDE); see Wand and Jones (1995). We crucially use features of this KDE to identify shells from streams. The feature that distinguishes streams from shells is how the density of the data cloud behaves around the 'ridge' (see the next paragraph for details): for streams, the KDE is symmetric on either side of the ridge, whereas for shells, it is asymmetric, with the density decreasing slowly as we move towards the center of the point cloud; see Figure 1.

Let us now define the notion of a *ridge* for a bivariate real-valued function; see Eberly (1996). For a smooth differentiable function $f : \mathbb{R}^2 \to \mathbb{R}$ define the 2×2 Hessian matrix $H(x) := ((H_{i,j}(x)))$, where $H_{i,j}(x) := \partial^2 f(x)/\partial x_i \partial x_j$. For each x, let $\lambda(x)$ be the smallest eigenvalue of $H(x)$ and let $p(x) \in \mathbb{R}^2$ be the corresponding eigenvector. The ridge points of f is the set of points x such that $p(x)^\top \nabla f(x) = 0$ and $p(x)^\top H(x) p(x) < 0$, i.e., the set of points of local maxima along the eigendirection of the smallest eigenvalue (sometimes also called as the 'normal direction'). Note that the normal direction is the direction along which the function has largest concavity; see Eberly (1996). We estimate the ridges of the KDE by the Subspace Constrained Mean Shift algorithm; see Chen et al. (2015), Genovese et al. (2014), Ozertem and Erdogmus (2011).

After having identified the ridges, for each point on the ridge, we test the hypothesis of symmetry (around the ridge point) of the density along its normal direction. To do this, we consider a small rectangle centered at the ridge point with sides parallel to the normal direction and project all the points in the rectangle onto a line through the ridge point parallel to the normal direction. Let x_i denote the (signed) distance of a projected point from the ridge point. We now test for the symmetry of distribution of x_i's using the two-sided Wilcoxon-signed rank test; Wilcoxon et al. (1970). P-values greater than 0.05 is indicative of streams. Adjacent stream-like

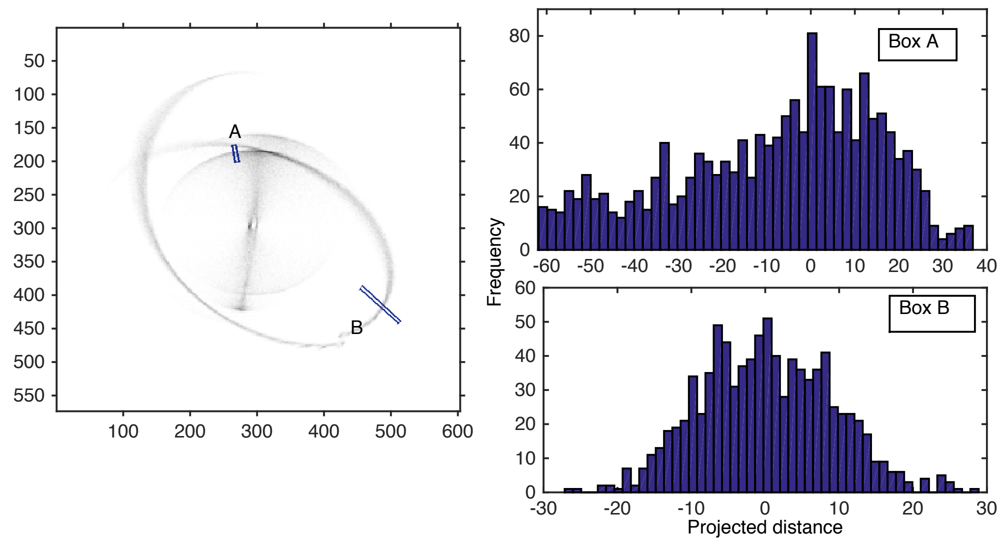

FIGURE 1. Histograms of x_i corresponding to two ridge points. "Box A" corresponds to shell and "Box B" corresponds to a stream. In each histogram the the red vertical lines represent the mode of the histograms.

points along the ridge are connected by smoothing traversing the ridge, and all particles within a few σ (as defined by the histograms) of the ridge are tagged as members of the stream. The streams, thus found, are separated out, and the shells are identified from the image obtained from the remaining points.

To identify the shells we first find the the edges in the image using standard edge detection techniques; see Canny (1986), Comaniciu and Meer (2002), and Meer and Georgescu (2001). As the shells in the image are known to lie on (nearly) concentric circles around the center of the point cloud, we use the concentric hough transform (Illingworth and Kittler, 1988) on the edges to detect significant circular arcs (and extract the corresponding radii and angles). Some post processing of the detected radii and angles are done to detect the connected component of the shells.

References

Canny, J. (1986). A computational approach to edge detection. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, (6):679–698.

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

Comaniciu, D. and Meer, P. (2002). Mean shift: A robust approach toward feature space analysis. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(5):603–619.

Eberly, D. (1996). *Ridges in image and data analysis*, volume 7. Springer Science & Business Media.

Genovese, C. R., Perone-Pacifico, M., Verdinelli, I., Wasserman, L., et al. (2014). Nonparametric ridge estimation. *The Annals of Statistics*, 42(4):1511–1545.

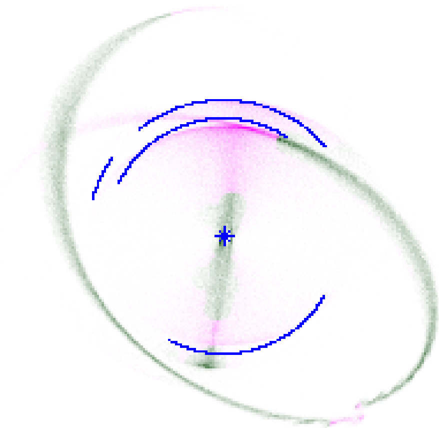

FIGURE 2. Detected streams (in green) and shells (blue arcs) overlaid upon the original image(in pink).

Illingworth, J. and Kittler, J. (1988). A survey of the hough transform. *Computer vision, graphics, and image processing*, 44(1):87–116.

Meer, P. and Georgescu, B. (2001). Edge detection with embedded confidence. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 23(12):1351–1365.

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

Wand, M. P. and Jones, M. C. (1995). *Kernel smoothing*, volume 60 of *Monographs on Statistics and Applied Probability*. Chapman and Hall, Ltd., London.

Wilcoxon, F., Katti, S., and Wilcox, R. A. (1970). Critical values and probability levels for the wilcoxon rank sum test and the wilcoxon signed rank test. *Selected tables in mathematical statistics*, 1:171–259.