

Infall Profiles for Supercluster-scale Filaments

Mary Crone Odekon¹ , Michael G. Jones² , Lucas Graham³ , Jessica Kelley-Derzon⁴ , and Evan Halstead¹ 

¹ Department of Physics, Skidmore College, Saratoga Springs, NY 12866, USA; mcrone@skidmore.edu, ehalste@skidmore.edu

² Steward Observatory, University of Arizona, 933 North Cherry Avenue, Rm. N204, Tucson, AZ 85721-0065, USA

³ Department of Physics, Washington University in St. Louis, St. Louis, MO 63105, USA; graham.l@wustl.edu

⁴ Department of Physics, University of Florida, Gainesville, FL 32603, USA; jkellyderzon@ufl.edu

Received 2022 June 2; revised 2022 July 6; accepted 2022 July 13; published 2022 August 23

Abstract

We present theoretical expectations for infall toward supercluster-scale cosmological filaments, motivated by the Arecibo Pisces–Perseus Supercluster Survey (APPSS) to map the velocity field around the Pisces–Perseus Supercluster (PPS) filament. We use a minimum spanning tree applied to dark matter halos the size of galaxy clusters to identify 236 large filaments within the Millennium simulation. Stacking the filaments along their principal axes, we determine a well-defined, sharp-peaked velocity profile function that can be expressed in terms of the maximum infall rate V_{\max} and the distance ρ_{\max} between the location of maximum infall and the principal axis of the filament. This simple, two-parameter functional form is surprisingly universal across a wide range of linear mass densities. V_{\max} is positively correlated with the halo mass per length along the filament, and ρ_{\max} is negatively correlated with the degree to which the halos are concentrated along the principal axis. We also assess an alternative, single-parameter method using V_{25} , the infall rate at a distance of 25 Mpc from the axis of the filament. Filaments similar to the PPS have $V_{\max} = 612 \pm 116 \text{ km s}^{-1}$, $\rho_{\max} = 8.9 \pm 2.1 \text{ Mpc}$, and $V_{25} = 329 \pm 68 \text{ km s}^{-1}$. We create mock observations to model uncertainties associated with viewing angle, lack of three-dimensional velocity information, limited sample size, and distance uncertainties. Our results suggest that it would be especially useful to measure infall for a larger sample of filaments to test our predictions for the shape of the infall profile and the relationships among infall rates and filament properties.

Unified Astronomy Thesaurus concepts: [Large-scale structure of the universe \(902\)](#); [Dark matter distribution \(356\)](#); [Redshift surveys \(1378\)](#); [Galaxies \(573\)](#)

1. Introduction

As the initial density perturbations of the early universe grow, they form large-scale structures such as sheets, filaments, and clusters. Observations of these various structures can provide tests of cosmological models over different spatial scales and in different dynamical environments. For example, the observed infall of galaxies toward large-scale filaments can provide a test of cosmological models on scales intermediate between clusters of galaxies (e.g., Sorce & Tempel 2018) and larger bulk flows (e.g., Howlett et al. 2022). The ongoing Arecibo Pisces–Perseus Supercluster Survey (APPSS; O’Dognoghe et al. 2019), in particular, is designed to provide observational constraints on the infall rate to the main filament of the Pisces–Perseus Supercluster (PPS). Here we provide a theoretical comparison for observations of infall to the PPS and similar filaments.

The PPS was chosen as the target for APPSS in part because of its simple orientation geometry: it is oriented across the sky, approximately perpendicular to the line of sight. It is about 70 Mpc away, on the far side of the local void that separates the Virgo supercluster region from the PPS region (see Haynes & Giovanelli 1986).

The APPSS survey will determine galaxy distances using the Baryonic Tully–Fisher Relation (BTFR; McGaugh et al. 2000; Papastergis et al. 2016; Lelli et al. 2016), an observed correlation between galaxy rotation speed and baryonic mass

that provides distance estimates to a precision of about 30%. By including all the baryonic mass in the galaxy—gas mass as well as stellar mass—the BTFR extends the traditional Tully–Fisher Relation (Tully & Fisher 1977) down to low-mass galaxies, allowing us to adequately sample the volume surrounding the PPS.

The APPSS survey is just one of a growing body of peculiar velocity catalogs used to test cosmological models. Especially notable are the recent Sloan Digital Sky Survey Peculiar Velocity Catalog (34,059 objects; Howlett et al. 2022), the Cosmicflows-4 Tully–Fisher catalog (9800 objects; Kourkchi et al. 2020), and the 6° Field Galaxy Survey peculiar velocity sample (8800 objects; Springob et al. 2014). Large surveys like these, combined with smaller targeted surveys like APPSS that focus on specific regions, will provide data that can be compared with theoretical expectations using a variety of different techniques. For example, Shaya et al. (2022) use numerical reconstructions on objects from the Cosmicflows-3 catalog (Tully et al. 2016) and groups from the Two Micron All Sky Survey (Tully 2015) to trace trajectories of galaxy systems in the local universe and constrain the value of the Hubble constant.

An analytic treatment of the velocity field near filaments is more complicated than treatments of the velocity field on either smaller or larger scales. The regions around filaments are in an intermediate stage of evolution, neither dynamically relaxed like small, dense clusters nor in the linear regime. In addition, unlike infall toward clusters of galaxies, which can be approximated as spherical, the gravitational attraction of external sources along one direction does not cancel out for a cylindrical geometry. Therefore, rather than use an analytic



Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](#). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

approximation, we determine the expected infall using cosmological simulations. In particular, we use the Millennium simulation (Springel et al. 2005) because it is large enough to provide a statistically useful sample of filaments. The Millennium volume is a periodic cube of length $500 h^{-1}$ Mpc, or about 714 Mpc for $h = 0.7$. As described below, we are able to obtain a sample of 236 large filaments in the Millennium volume, including 40 with properties similar to the PPS. For comparison, the Illustris TNG300 (Pillepich et al. 2018) simulation volume is a cube of length 300 Mpc, nearly 15 times smaller.

The Millennium database also includes galaxy catalogs based on semi-analytic models. To mimic the observational methods of APPSS, we trace infall using gas-rich galaxies from the catalog of Lacey et al. (2016), which applies the GALFORM semi-analytic model to the updated “MR7” Millennium run.

Several previous studies have investigated the expected structure of velocity fields in the region of filaments. One particular approach is to measure the dispersion of galaxy redshifts within a filament and relate this to the linear mass density of the filament (Eisenstein et al. 1997; Pereyra et al. 2020). This method has the significant advantage that it can be used for galaxies without redshift-independent distance measurements. For a filament aligned across the sky, like the PPS filament, it can be based entirely on the width of the filament in redshift space. This is somewhat analogous to determining the virial mass of a cluster of galaxies using the velocity dispersion. However, it does not directly measure the mass in the same way as it would for a relaxed, virialized object. Within a filament, the velocity dispersion may be dominated by orbital motions of galaxies within embedded dense groups and clusters. In addition, the width of a filament in redshift space likely includes a large number of infalling galaxies as well as orbiting galaxies. Nonetheless, it is not surprising that the velocity dispersion of filaments may be *correlated* to their underlying mass density, because larger filaments are likely to have larger clusters and greater infall, and might therefore be used to estimate the underlying mass.

Another reason to consider the velocity field of simulated filaments is its effect on galaxy evolution. For example, Laigle et al. (2015) and Dubois et al. (2014) find that the spins of halos embedded in filaments are aligned with the vorticity of the filament. Indeed, Xia et al. (2021) conclude that rotation is an important filament property, and Wang et al. (2021) find possible observational evidence for filament spin. Another example is presented by Kraljic et al. (2019), who show that the properties of simulated galaxies at different redshifts support the picture of motion along filaments, away from saddle points between large nodes.

Velocity fields can also be used as part of the criteria for defining filaments. While it is typical for filaments to be identified using the density field or some tracer thereof, simulations can take advantage of the full six-dimensional phase space information to define what is special about a filament. Buehlmann & Hahn (2019), for example, use the eigenvalues of the velocity dispersion tensor to identify filament-like environments where two dimensions are collapsed.

In a study of filaments more closely related to the project we discuss in this paper, Pereyra et al. (2020) show that the average velocity field surrounding simulated filaments is

perpendicular to the filament spine, while the motion within the filament (that is, close to the filament spine) is primarily along the filament.

While these studies illustrate the importance of filament velocity fields in understanding cosmology and galaxy evolution, they are not designed to provide a precise and observationally practical comparison for projects like APPSS that measure the large-scale infall surrounding a large (supercluster-scale) filament. Our goal here is to provide expected infall rates as a function of observable properties of the filament and to find an observationally feasible way to make these measurements.

In this context, we choose a methodology that is straightforward to apply to observations like those available for the PSS and that focuses on measurable, large-scale features of the filament structure and surrounding velocity field. In particular, to compensate for observational uncertainties in distance, we average over many galaxies within a large region surrounding each filament. This means we are not sensitive to variations along the filaments, but only to the overall bulk flow toward or away from them. Our approach is to identify the primary, straight-line axis of each filament using principal component analysis and then find the average motion toward or away from this axis for galaxies in a large volume surrounding it. This differs from studies that trace the curves of filaments on smaller scales and consider variations along them.

More generally, given the multiscale nature of cosmological filaments, it is important to keep in mind that our study is specifically designed to address supercluster-scale filaments. The multiscale nature of filaments can be seen in large-scale cosmological simulations (e.g., Kraljic et al. 2019; Galárraga-Espinosa et al. 2020) and observations (e.g., Carrón Duque et al. 2022; Pereyra et al. 2020; Martínez et al. 2016), and also in studies that focus on filaments that drain into particular clusters and superclusters, whether in simulations (e.g., Kuchner et al. 2020; Gouin et al. 2020) or observations (e.g., Castignani et al. 2022; Einasto et al. 2020; Mahajan et al. 2018).

This paper is organized as follows. In Section 2 we describe our method for defining filaments. In Section 3 we present the results for infall profiles for the simulated filaments, including two different ways to parameterize infall, and show how the parameters relate to the observable structure of the filaments and the underlying dark matter distribution. In Section 4 we present results for mock observations, including measurement uncertainties, and in Section 5 we summarize our results and discuss the implications.

Note that in Sections 2 and 3, we present distances and masses in terms of the dimensionless Hubble constant $h = H_0/100 \text{ km s}^{-1} \text{ Mpc}^{-1}$, the same way distances and masses are presented in the Millennium simulation database. For the mock observations in Section 4, which include h -dependent magnitude limits, we use $h = 0.7$.

2. Defining Filaments

Cosmological filaments are defined in a variety of ways, depending on the information available and the purpose of a particular study. If the information is from a cosmological simulation, for example, filament identification can be based on the dark matter distribution itself rather than the galaxy distribution. Methods for finding filaments are reviewed in Libeskind et al. (2018), Rost et al. (2020), and Carrón Duque et al. (2022).

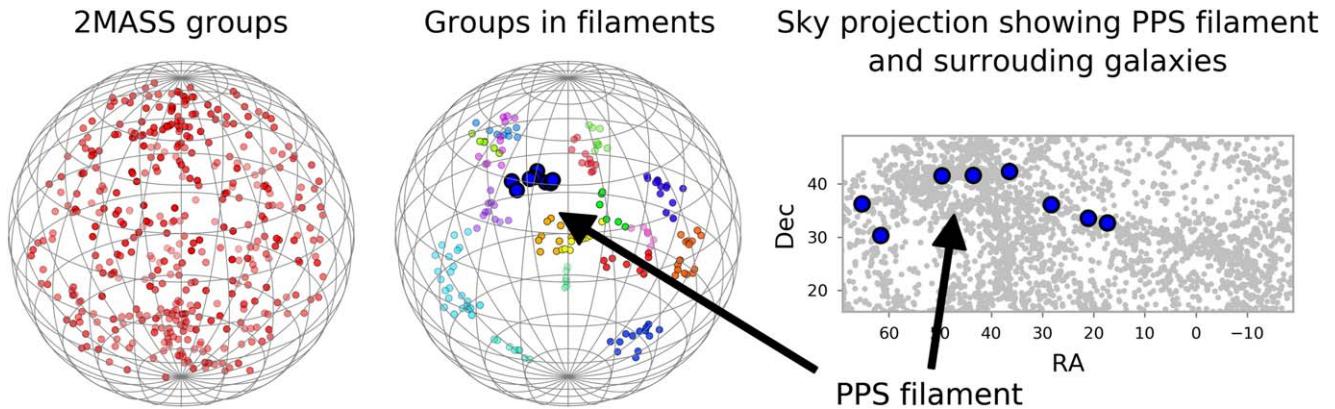


Figure 1. Results of our filament identification technique applied to the 2MASS groups catalog. The left panel shows all groups with $M_{vir} > 10^{14} h^{-1} M_{\odot}$. The center panel shows only groups identified to be in filaments, with different filaments indicated by color. The right panel is a sky view showing the groups in the Pisces–Perseus Supercluster filament (large blue dots) in the context of the 2MASS galaxy distribution (small gray dots).

Our method for defining filaments is designed to be used in the context of galaxy surveys as well as simulations, and is similar to those of Alpaslan et al. (2014), Crone Odekon et al. (2018), Pereyra et al. (2020), Bonnaire et al. (2020) and other graph-based methods (see Section 2 of Libeskind et al. 2018). We start with groups of galaxies with virial masses above $10^{14} h^{-1} M_{\odot}$, roughly the mass range often referred to as galaxy clusters. We connect these large groups with a minimum spanning tree and break the tree where connecting edges are greater than $17 h^{-1} \text{Mpc}$. The remaining structures that have at least six large groups are considered filaments. Breaking the tree at this particular edge length maximizes the number of filaments with at least six connected groups in the simulation.

Figure 1 illustrates this process applied to observations of the local universe, specifically the Two Micron All Sky Survey (2MASS) groups catalog of Tully (2015). This catalog has the advantage of covering nearly the entire sky, because it is based on infrared wavelengths that can penetrate through dust in the Milky Way down to low galactic latitudes. We apply the filament finder in three dimensions with each group placed at distance cz/H_o . The left panel in Figure 1 shows all the groups above a virial mass of $10^{14} h^{-1} M_{\odot}$ out to $cz = 10,000 \text{ km s}^{-1}$, within which the sample is approximately complete (see Section 6 in Tully 2015). The center panel in Figure 1 shows the resulting filaments with at least six connected groups and with the PPS filament indicated by large blue dots. The right panel shows the positions on the sky of these large groups in the PPS filament in the context of the 2MASS galaxy distribution.

Using our filament identification process, the PPS filament is defined by eight large groups that have a combined virial mass of $2.9 \times 10^{15} h^{-1} M_{\odot}$. If we define the filament length as the straight-line distance between the two most distant groups, the PPS filament has a length of $38 h^{-1} \text{Mpc}$. It is important to keep in mind, however, that different filament detection algorithms would result in different properties for the PPS and other filaments. Our approach is to use a method that works with existing observational data and that succeeds in identifying the main body of the PPS filament.

Figure 2 illustrates the filament identification process applied to the Millennium simulation. The left panel shows all halos with $M_{200} > 0.72 \times 10^{14} h^{-1} M_{\odot}$, where M_{200} is the mass within a region where the density is 200 times the average density of the universe. This mass cutoff includes a correction factor of

0.72 to relate the M_{200} halo masses in the Millennium database to the virial masses in the 2MASS group catalog (see Duffy et al. 2008; Coe 2010; Saez et al. 2016). There are 5481 halos above this cutoff. The center panel in Figure 2 shows the resulting 236 filaments that have at least six halos, and the right panel shows one filament for which the elongated shape is clear in the y – z projection.

As illustrated in Figure 3, the lengths and combined group masses of the simulated filaments bracket those of the PPS filament, suggesting that these filaments are indeed an appropriate sample to compare with the PPS.

Figure 4 shows a histogram of the lengths of the simulated filaments, in a format similar to that used in several recent papers that describe filament catalogs. Note that different catalogs define filaments in different ways, resulting in different typical filament lengths. Our filaments tend to be larger, for example, than those in the three catalogs compared by Rost et al. (2020) and in the lowest redshift block of Carrón Duque et al. (2022). Our choice to define filaments using large groups is designed to detect supercluster-scale filaments like the PPS filament.

3. Determining Infall Profiles

In this section we describe our method for selecting galaxies around the simulated filaments and measuring their infall rate as a function of distance from the filament axis.

The Millennium simulation (Springel et al. 2005) uses a TreePM code to model gravity for collisionless matter and then applies semi-analytical models to trace the formation and evolution of galaxies within merging dark matter halos. The Millennium database provides several simulated galaxy catalogs, some based on the original Millennium Run (MR), which used cosmological parameters consistent with 2dFGRS and 1 yr WMAP data, and others based on the updated MR7 (Guo et al. 2013), which used cosmological parameters consistent with the 7 yr WMAP data. There are also two different semi-analytic codes used to create the Millennium galaxy catalogs. Some catalogs use the L-Galaxies code developed at the Max Planck Institute (Croton et al. 2006; De Lucia et al. 2006), and some use the GALFORM code developed at the Durham University Institute for Computational Cosmology (Bower et al. 2006). In this paper, we focus on the Lacey et al. (2016) catalog, as it uses the updated MR7 cosmological parameters and the most recent semi-analytic

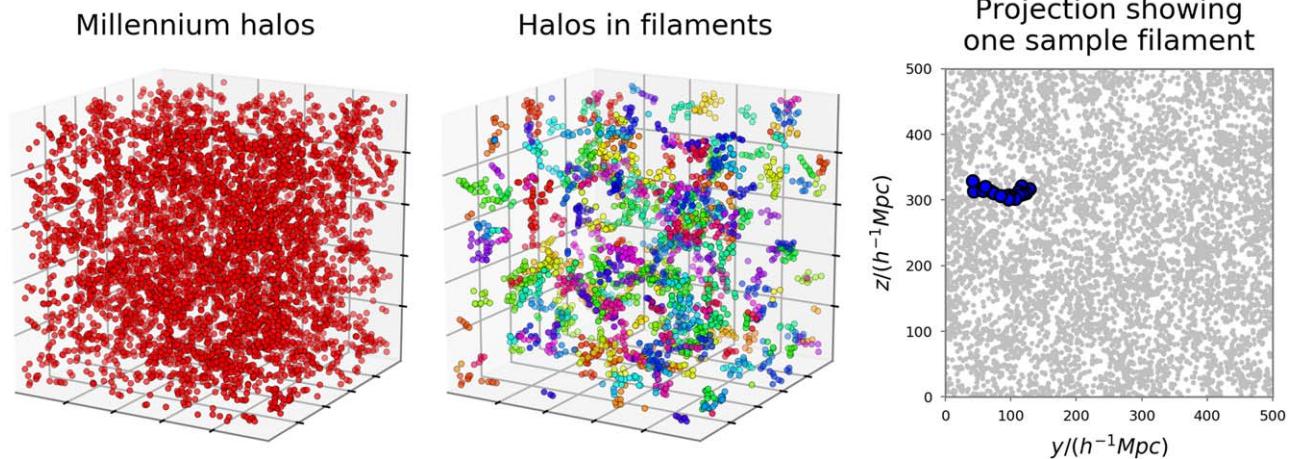


Figure 2. Results of our filament identification technique applied to the Millennium simulation. The left panel shows all halos with $M_{200} > 0.72 \times 10^{14} h^{-1} M_{\odot}$, approximately equivalent to the cutoff of $M_{\text{vir}} > 10^{14} h^{-1} M_{\odot}$ used in Figure 1 for the 2MASS groups. The center panel shows only halos identified to be in filaments, with different filaments indicated by color. The right panel is a rotated view showing the elongated geometry of a typical filament. Note that the Millennium simulation has periodic boundaries, so that some filaments are cut off on one side and continue on the opposite side.

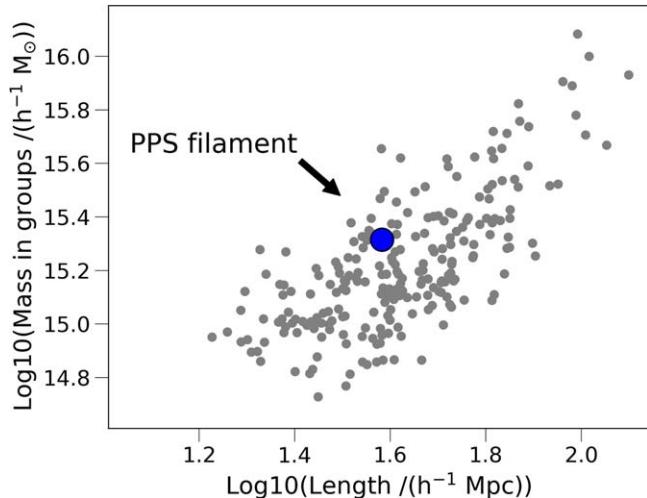


Figure 3. Comparison of the length and mass of the Pisces–Perseus Supercluster filament (large blue dot) with the simulated filaments (small gray dots). The horizontal axis shows the distance between the two most distant groups that define the filament. The vertical axis shows the sum of the virial masses of the groups or halos that define the filament. The range of values for the simulated filaments brackets the values for the PPS filament.

models (which are based on GALFORM). The APPSS survey is based on gas-rich late-type galaxies, so we make our sample selection based on cold-gas mass, selecting galaxies with a minimum cold-gas mass of $M_{\text{min}} = 10^8 h^{-1} M_{\odot}$.

For our purposes, galaxies serve mostly as tracers of the flow of matter toward filaments, so we do not expect our results to be very sensitive to the details of semi-analytic modeling. However, we expect the selection criteria for galaxies to make *some* difference, if only because early-type galaxies are more concentrated in groups and clusters. Therefore, their velocities will be more strongly affected by local orbits, potentially making it more difficult to measure the larger-scale flow toward filaments.

We also make use of the galaxy catalog of Lagos et al. (2012), which applies the GALFORM semi-analytic model to the earlier MR run. This is partly to check the dependence of our results on small changes to cosmological parameters. It also allows us to compare the galaxy velocity field to the large-scale

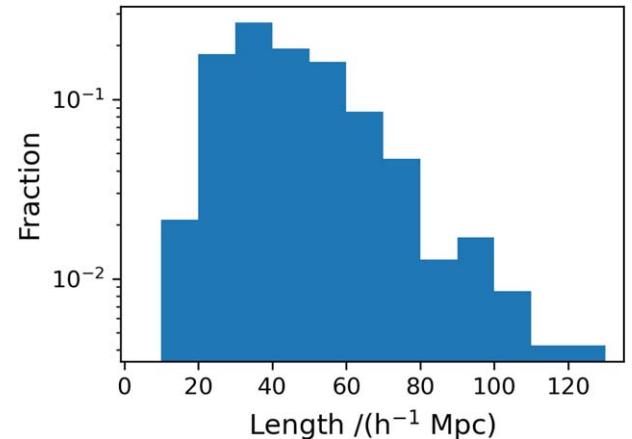


Figure 4. Histogram of filament lengths, in a format similar to Figure 19 of Carrón Duque et al. (2022). Our filament detection algorithm is tuned to select larger filaments than some other filament catalogs (e.g., Colberg et al. 2005; Rost et al. 2020).

underlying matter distribution, which is provided in the Millennium database for MR (but not MR7) in the form of a 256^3 density grid. Note that the MR run is based on a different random realization of the initial conditions than the MR7, so it produces a totally different set of filaments.

As described in the introduction, our goal is to determine an overall, large-scale infall profile for each filament. To do this, we select galaxies in a large cylindrical-shaped region centered on the main axis of each filament. We identify the axis by applying principal component analysis (PCA) to the large halos that were used to identify the filament, weighting the halos in proportion to their masses. The principal axis is defined as the straight line that minimizes the weighted distances of the halos from the axis. We then select the galaxies in a cylinder along the principal axis of each filament, with length $30 h^{-1} \text{ Mpc}$ along the principal axis of each filament and radius that extends to $100 h^{-1} \text{ Mpc}$. The length of $30 h^{-1} \text{ Mpc}$ is large enough to measure a representative infall profile but small enough to be dominated by the filament. The radius of $100 h^{-1} \text{ Mpc}$ extends far enough out to see the large-scale gravitational effect of the filament.

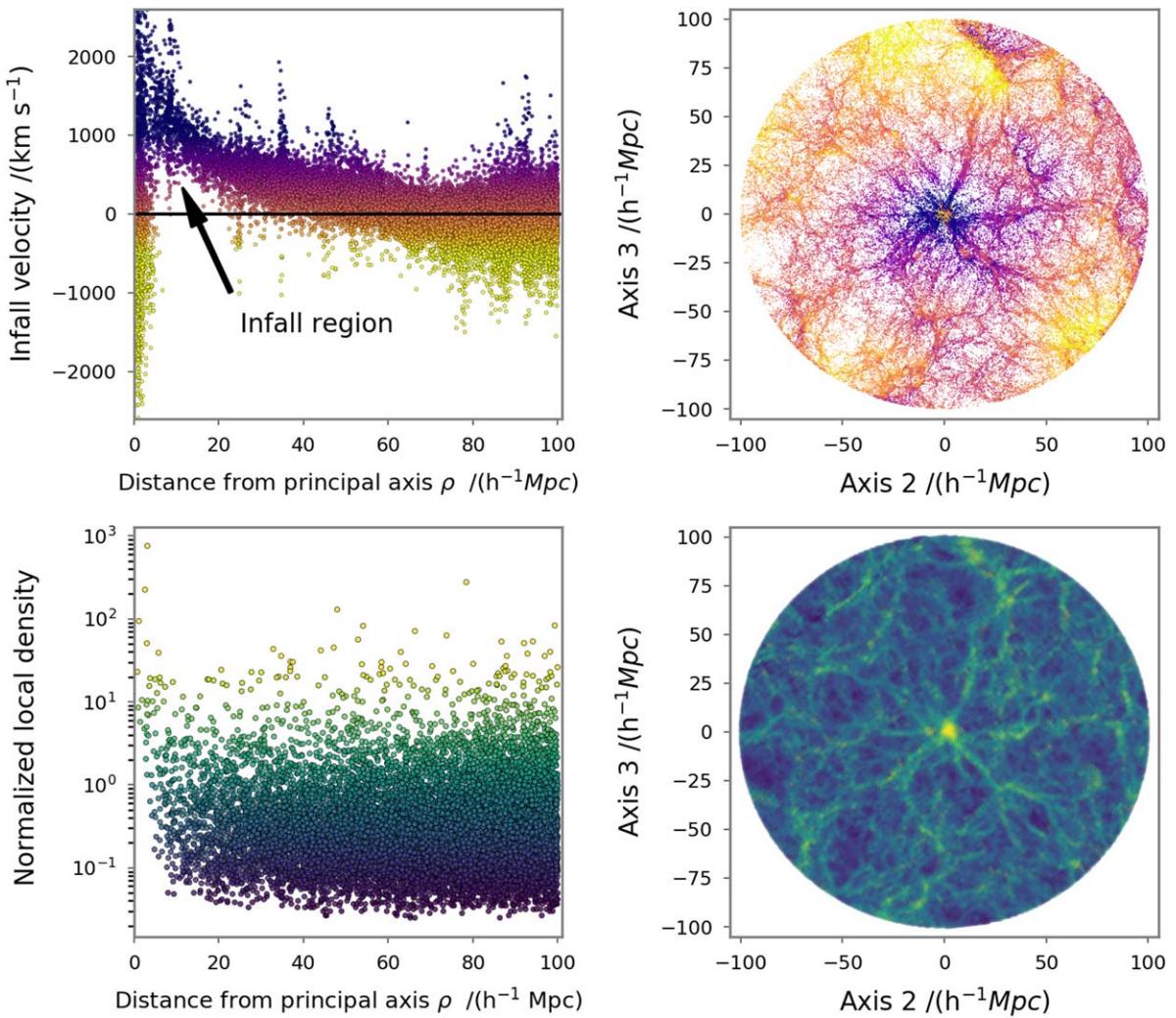


Figure 5. Velocity and density structure in a cylinder-shaped region of thickness $30 h^{-1}$ Mpc surrounding one sample simulated filament. The top left panel shows the component of velocity toward the principal axis of the filament as a function of the distance ρ from the principal axis. The top right panel shows the cylinder-shaped region projected along the principal axis, with the same color coding for infall velocity as on the left. Galaxies in the very center are primarily in virialized structures where orbital motions produce both positive and negative infall velocities, while at about $10 h^{-1}$ Mpc, nearly all galaxies are falling toward the filament (positive infall velocity). At high distances from the filament, the infall speed approaches zero. (Note that the values for velocity are in comoving coordinates, which do not include the expansion of the universe.) The bottom left panel shows density as a function of distance from the principal axis, and the bottom right panel shows the projected cylinder with the same color coding for density as on the left. The filament appears as a light-yellow dense region in the center surrounded by smaller filamentary structures.

We express infall in terms of cylindrical coordinates, with the axis of the coordinate system defined by the principal axis of each filament. We use the phrase “infall profile” to mean $V_{\text{infall}}(\rho)$, where cylindrical coordinate ρ is the perpendicular distance to the principal axis, and the “infall velocity” $V_{\text{infall}} = -\mathbf{v} \cdot \hat{\rho}$ is the component of the velocity directed toward the principal axis (in the direction of negative ρ).

Figure 5 illustrates the density and velocity structure for the cylinder-shaped region selected in this way surrounding a sample filament from the MR run. The top left panel shows the infall profile. Galaxies in the very center are primarily in virialized structures where orbital motions produce both positive and negative values for V_ρ . At about $10 h^{-1}$ Mpc, nearly all galaxies are falling toward the filament, with a positive infall velocity. At large distances, galaxies are not on average affected by that particular filament, so the average infall speed approaches zero. Note that the values for velocity are in comoving coordinates, which do not include the expansion of the universe.

The top right panel of Figure 5 uses the same color coding based on the infall velocity but shows galaxy positions within the projected cylinder, with the principal axis directed perpendicular to the page. The dark blue region surrounding the center is the infall region.

The bottom row of Figure 5 is similar to the top row, but instead of showing the infall it shows the density. On the bottom left is the density for each grid point in the cylinder as a function of distance from the principal axis, and on the bottom right is an image of the projected cylinder, using the same color coding for density as on the bottom left. The region near the filament axis appears as a light-yellow overdense region in the center of the cylinder surrounded by a complicated web of smaller filamentary structures.

The cylinders we examine around each filament are quite large, so we can examine the shape of the infall profile out to large distances. To illustrate the scale, as well as the geometry of the cylinder and principle axis, Figure 6 shows the cylinders surrounding three of the filaments in the context of the entire

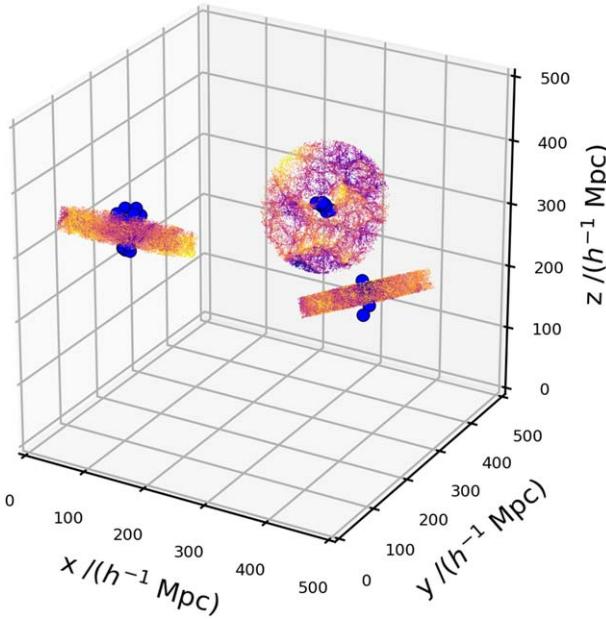


Figure 6. Cylinders perpendicular to the principal axis for three sample filaments, showing the scale of the regions from which we draw galaxies to measure the infall profile. Large blue dots are the groups that define each filament. Galaxies in the cylinders are color coded for infall velocity in the same way as the projected cylinders on the upper right of Figure 5.

Millennium volume. The groups used to define the filaments are indicated with large blue dots, and the galaxies are indicated with small points that are color coded for infall velocity as in Figure 5. The orientation of the cylinders is defined relative to the principle axis of each filament.

To find an average infall profile function, we average together the 236 simulated filaments in the MR7 run by stacking them along their principal axes. The resulting infall profile, if we average the galaxy infall velocities in bins of distance from the principal axis, is shown in the left panel of Figure 7. Each filament is equally weighted in each distance bin; the average profile exhibits considerably more scatter if we instead weight each galaxy the same, so that the profile is more heavily weighted by whichever filaments have the most galaxies at that particular distance from the principal axis. Note that the infall profiles are calculated as an average in shells (binning by the distance ρ), not as an integrated average of all galaxies within ρ .

A simple function that can be used to describe the shape in the left panel of Figure 7 is

$$v = C\rho^4 e^{-\rho^{0.2}}, \quad (1)$$

where ρ is the distance from the principal axis. The rounded peak is the result of averaging together filaments with velocity profiles that peak at different distances. To find a function that represents the typical infall profile for a *single* filament, it is more appropriate to normalize the distances by the peak infall distance before averaging. To do this, we first average the infall velocities of the galaxies for each filament in distance bins of width $1 \text{ Mpc } h^{-1}$ and find the distance bin with the highest average infall. The distances from the principal axis are then divided by this distance before stacking together with the other filaments. The right panel of Figure 7 shows the resulting

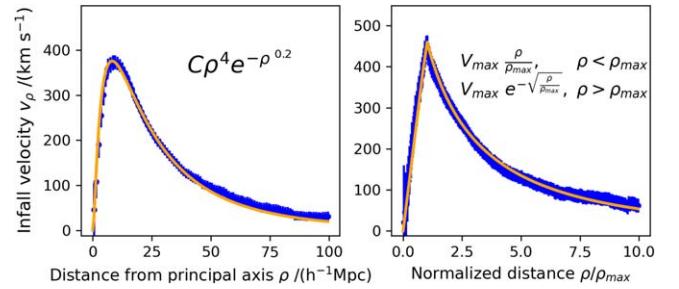


Figure 7. Average velocity profiles for all 236 simulated filaments from the MR7 run. On the left, infall velocities are stacked by distance from the principal axis. On the right, the distance for each filament is first normalized by its peak infall distance before stacking. Blue dots and error bars show binned means and the standard error in the mean. Orange lines show the functions given on the upper right of each panel. The peak in the left panel is rounded because different filaments peak at different distances. For a single filament, it is more appropriate to use the sharp-peaked function in the right panel.

average profile, which we fit to the function

$$v = V_{\max} \frac{\rho}{\rho_{\max}}, \quad \rho \leq \rho_{\max} \\ v = V_{\max} e^{-\sqrt{\frac{\rho}{\rho_{\max}}}}, \quad \rho \geq \rho_{\max}. \quad (2)$$

For the averaged profile in the right panel of Figure 7, the peak occurs by construction at $\rho/\rho_{\max} = 1$. But for an individual filament, Equation (2) can be fit to the infall profile in terms of ρ to determine the two parameters V_{\max} and ρ_{\max} .

Figure 8 shows this function fit to 15 of the 236 MR7 filaments. The infall profiles show a wide variety of shapes due to the range of internal structure and surrounding environment for each filament. Several filaments are in the vicinity of large structures that create clear gravitational signatures in the outer parts of these profiles. Filaments 3 and 6, for example, have structures at a distance of about 100 Mpc that cause the profile to be lower on the close side (smaller distance) and higher on the far side. Filament 14 has a large structure more than 100 Mpc away that pulls the entire outer part of the profile to lower infall velocities. Nonetheless, the fitting function reliably finds the peak in the infall profile.

3.1. A Two-parameter Description of Infall: V_{\max} and ρ_{\max}

In this subsection we show that the parameters V_{\max} and ρ_{\max} , as determined using Equation (2), are related to two different aspects of the internal structure of the filaments: the overall mass of the filament and the degree to which the mass is concentrated close to the principal axis. These two aspects of the structure can be seen in either the positions and masses of the groups used to identify the filaments, or in the full underlying mass distribution, including dark matter.

First, we note that V_{\max} and ρ_{\max} are independent parameters, both required to specify the infall profiles for a particular filament (Figure 9). If the situation were simpler, and infall profiles for different filaments could be differentiated by a single parameter—for example, a parameter related to the overall size of the filament—we would expect to see a correlation between V_{\max} and ρ_{\max} , indicating that they could be reduced to one parameter.

The top row of Figure 10 shows how V_{\max} and ρ_{\max} are related to two different properties of the filament: the mass per length of large groups along the filament and the degree to which the groups are concentrated toward the principal axis.

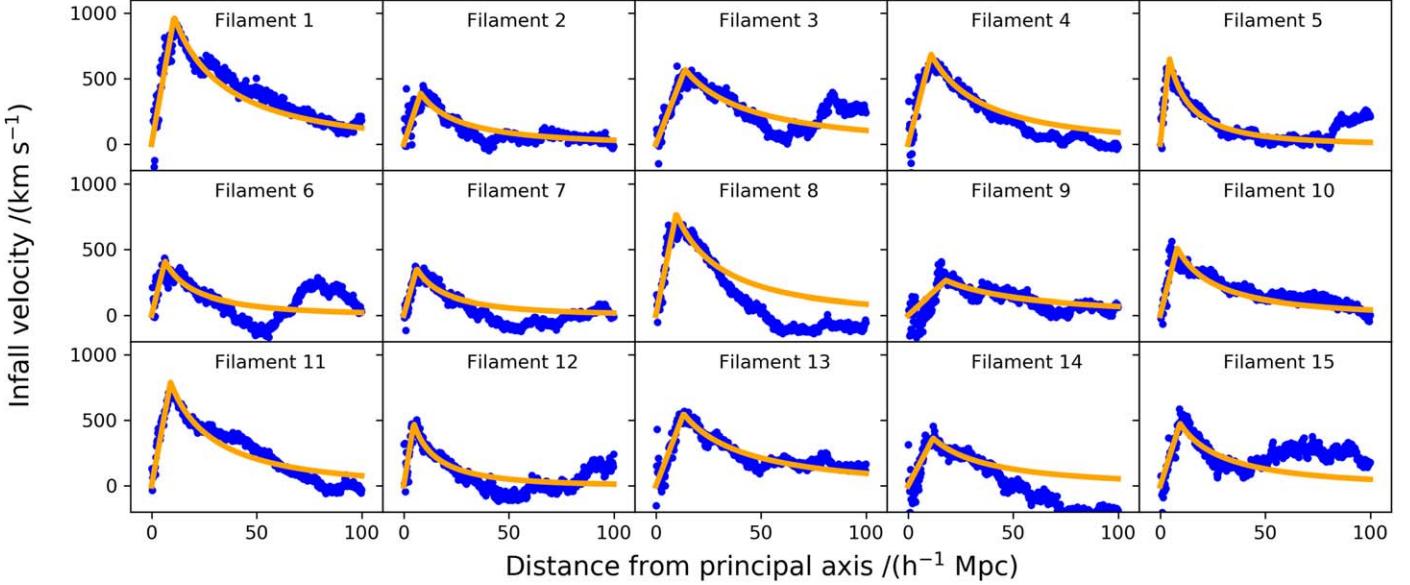


Figure 8. Velocity profiles as a function of distance from the principal axis for 15 of the 236 simulated filaments. Blue dots show averages in bins of $1 h^{-1} \text{ Mpc}$. Orange lines show the fitting function in Equation (2) applied to each filament. Filaments vary based on internal structure and surrounding environment, but the fitting function typically succeeds in identifying the maximum infall velocity and its distance from the principal axis.

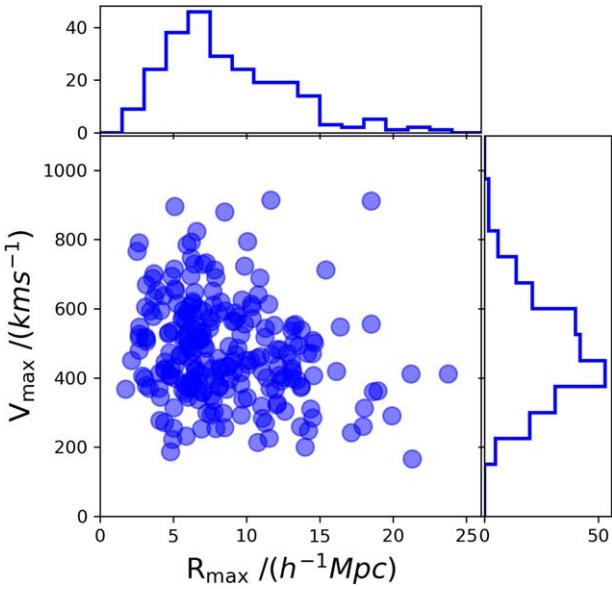


Figure 9. Distribution of values for V_{\max} and ρ_{\max} for the simulated filaments. These two parameters are independent and reflect different aspects of filament structure; they are not correlated as one would expect if the infall profile depended on a single underlying parameter.

The groups used to define these properties are the same ones used to define the filaments—those with masses greater than $M_{200} > 0.72 \times 10^{14} h^{-1} M_{\odot}$ and not separated by more than the cutoff linking length of $17 h^{-1} \text{ Mpc}$. The mass per length is defined as the sum of all the group masses in the filament divided by the straight-line distance between the two most distant groups. These are the same definitions for mass and length used in Figure 3. The group concentration is defined as the sum of the masses of groups within $5 h^{-1} \text{ Mpc}$ of the principal axis divided by the sum of the masses in groups within $15 h^{-1} \text{ Mpc}$. These groups are the same ones used to define the filaments and to calculate mass and length. This specific definition of groups concentration is somewhat

arbitrary but is useful because it produces a wide range of values when applied to the simulated filaments.

The positions and masses of groups are observable properties that can provide predictions for the expected values of V_{\max} and ρ_{\max} for a specific filament like the PPS. Figure 10 shows the values for the PPS as vertical gray lines. If we estimate the uncertainty in mass per length for the PPS to be 30% (e.g., by adding in quadrature $\sim 20\%$ uncertainty in mass to light ratio and $\sim 20\%$ in the amount of light Tully 2015), then the PPS has a group mass per length in the range $5625 \pm 1591 M_{\odot} \text{ Mpc}^{-1}$. If we estimate the group concentration to be anywhere in the range 0.6–1.0, then 40 of the simulated filaments are “like the PPS” in terms of their group mass per length and group concentration. These 40 filaments have an average $V_{\max} = 612 \text{ km s}^{-1}$ with standard deviation 116 km s^{-1} and an average $\rho_{\max} = 6.2 h^{-1} \text{ Mpc}$ with standard deviation $1.5 h^{-1} \text{ Mpc}$.

The bottom row of Figure 10 shows how V_{\max} and ρ_{\max} are related to the underlying mass distribution, including dark matter. Unlike the group structure, the underlying mass distribution is not directly observable. Instead, it can be inferred from measurements of V_{\max} and ρ_{\max} assuming the cosmological models are correct. In order to compare infall with the matter distribution, we use filaments from the MR Millennium run. (As described above, the MR has a density grid available in the Millennium database.) The overdensity in the bottom left panel is defined as the density within a cylinder of radius $5 h^{-1} \text{ Mpc}$ and length $30 h^{-1} \text{ Mpc}$ oriented along the principle minus the average density of the universe, normalized by the average density of the universe. The mass concentration in the bottom right panel is defined as the mass within a cylinder of radius $5 h^{-1} \text{ Mpc}$ and length $30 h^{-1} \text{ Mpc}$ oriented along the principle axis, divided by the mass within a cylinder of radius $15 h^{-1} \text{ Mpc}$ and length $30 h^{-1} \text{ Mpc}$. As was the case for the specific definition of group concentration, these definitions are somewhat arbitrary but are useful in that they produce a fairly wide range of values for the simulated filaments.

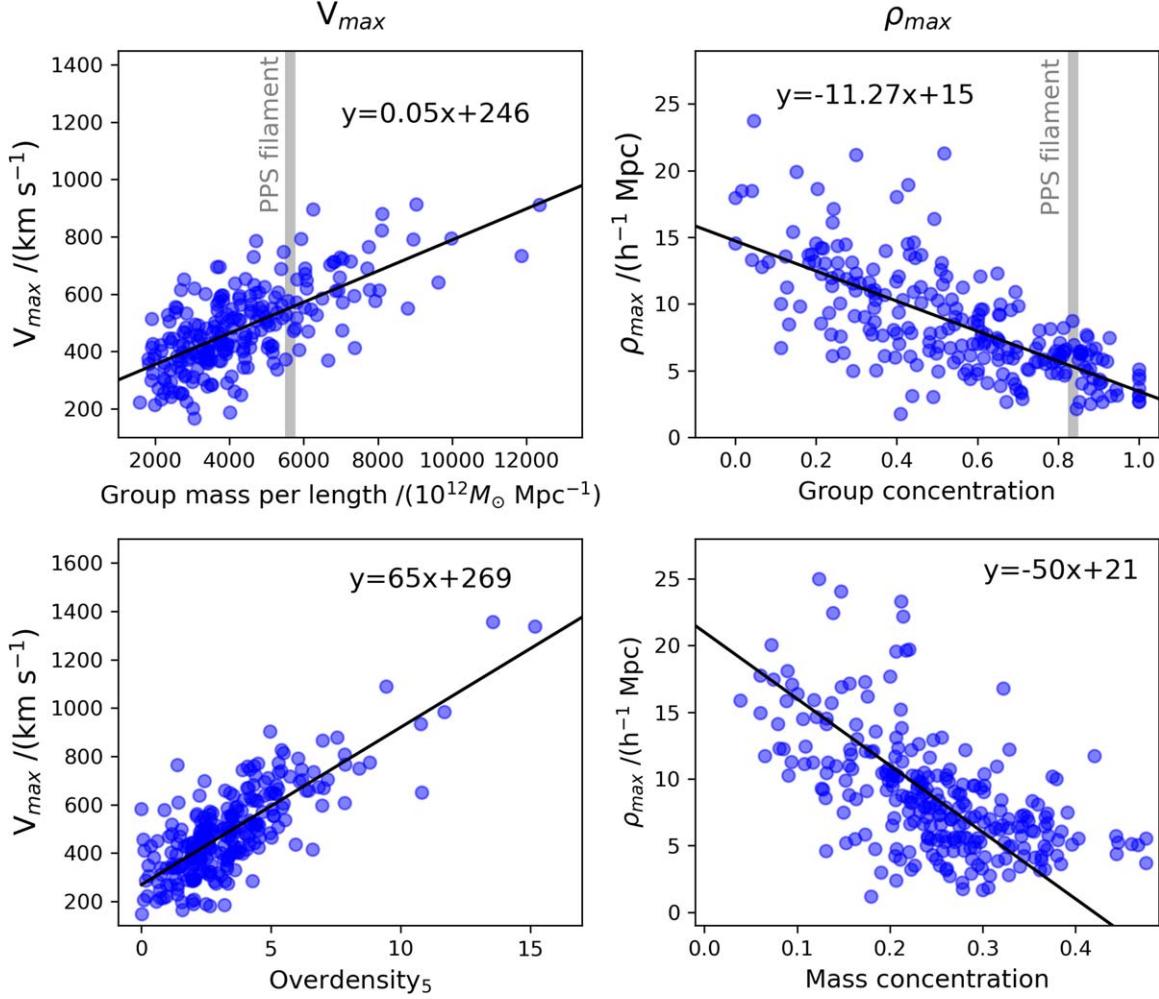


Figure 10. Relationships between V_{\max} (left) and ρ_{\max} (right) and the underlying mass distribution in each filament. The upper panels show relationships with the mass distribution of the groups used to define the filaments: V_{\max} is positively correlated with the mass of the groups in the filament divided by the length of the filament, and ρ_{\max} is negatively correlated with the degree to which groups are concentrated toward the principal axis of the filament. Vertical gray lines show the observed values of group mass per length and group concentration for the PPS filament, from which we can predict the expected values of V_{\max} and ρ_{\max} . The lower panels show relationships with the full mass distribution, including dark matter. The dark matter distribution is not directly observable but could instead be inferred from measured values of V_{\max} and ρ_{\max} . See the text for the specific definitions of group concentration, mass concentration, and overdensity.

3.2. A One-parameter Description of Infall: V_{25}

In this subsection we provide an alternative, one-parameter method using Equation (2) to predict the infall rate to filaments and to use the observed infall rate to infer the underlying dark matter distribution. This method provides less information about the filament than the two-parameter description in terms of V_{\max} and ρ_{\max} . However, it is more robust to observational limitations.

As we describe in Section 4 below, when we take into account observational limitations it becomes more difficult to reliably determine V_{\max} and ρ_{\max} separately. However, the value of the infall velocity at a fixed distance of 25 Mpc can still be recovered, and this value is also related to the structure of the filament. (Note that the value of V_{25} is based on a distance of 25 Mpc assuming $h = 0.7$, not a distance of 25 h^{-1} Mpc.)

As shown in Figure 11, V_{25} is related to both the group mass and to the overdensity of the filament. Here we define group mass as the mass of the groups used to define the filament, when limited to the central $30 h^{-1}$ Mpc along the principle axis. In the lower panel, we define overdensity in terms of the mass within a cylinder of radius $15 h^{-1}$ Mpc and length $30 h^{-1}$ Mpc.

The value of the group mass, when defined in this way, is $2.0 h^{-1} 10^{15} M_{\odot}$, or in the range 1.4 – $2.6 h^{-1} 10^{15}$ for a mass uncertainty of 30%. The subset of the 40 filaments that are “like the PPS” as defined above and that are in this range have an average $V_{25} = 329 \text{ km s}^{-1}$ with standard deviation 68 km s^{-1} .

The relationships for V_{25} in Figure 11 are generally more tightly correlated than those for the more complicated two-parameter description in Figure 10. The relationship in the upper panel of Figure 11 has an adjusted coefficient of determination $R_{\text{adj}}^2 = 0.638$ and the lower panel has $R_{\text{adj}}^2 = 0.764$. For comparison, the relationships for V_{\max} in Figure 10 have $R_{\text{adj}}^2 = 0.456$ (upper panel) and $R_{\text{adj}}^2 = 0.604$ (lower panel), and the relationships for ρ_{\max} have $R_{\text{adj}}^2 = 0.488$ (upper panel) and $R_{\text{adj}}^2 = 0.323$ (lower panel).

When relating V_{\max} , ρ_{\max} , and V_{25} to the mass distribution of the filament, we fit to a straight line $y(x) = mx + b$ rather than assume a different functional form. We investigated fitting instead to a power law $y(x) = ax^m$ but found that the straight line was generally a better fit in the sense of having a higher R_{adj}^2 . There are physical reasons one might expect a different functional form (for example, both V_{\max} and V_{25} going to zero

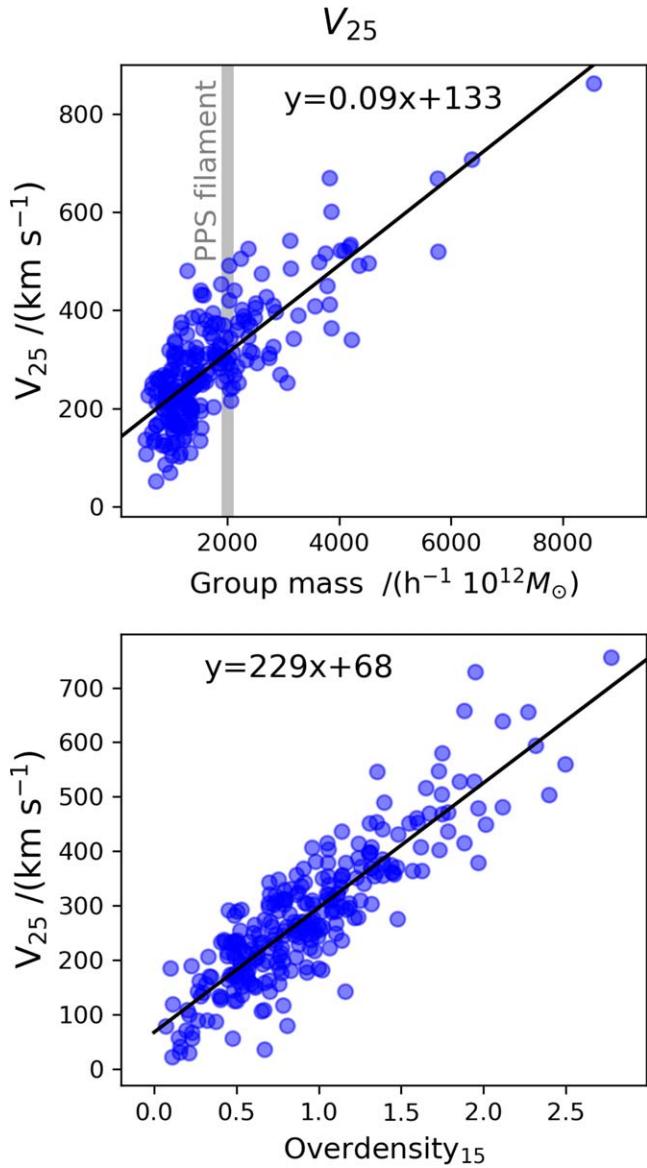


Figure 11. Similar to Figure 10, but for the description of infall in terms of the single parameter V_{25} , the value of the infall velocity at a fixed distance of 25 Mpc. V_{25} is correlated with the total mass of the groups used to define the filament, and with the underlying overdensity of the filament.

as the overdensity goes to zero), but the scatter is such that we cannot distinguish among more complicated models.

Before discussing mock observations, we pause to comment on the dependence of our results on cosmological parameters and galaxy selection criteria. We can compare the results in the upper rows of Figures 10 and 11, which are based on the MR7 run and for which galaxies were selected by cold-gas mass, with the same analysis performed on the MR run, or with galaxies selected by stellar mass. In all cases, the best-fit lines are consistent with those shown in Figures 10 and 11, so our results are robust to these differences. For example, a least-squares fit to the relationship between V_{\max} and group mass per length (upper left of Figure 10) yields a slope of 0.054 ± 0.004 and intercept of 246 ± 18 . If we instead select the galaxies by stellar mass, we find a slope of 0.053 ± 0.004 and intercept of 226 ± 18 , and if we instead use the values from the MR run (which uses a different random realization of the initial conditions, as well as slightly different cosmological

parameters), we find a slope of 0.045 ± 0.004 and intercept of 249 ± 17 . Similarly, the distribution of values of V_{\max} , ρ_{\max} , and V_{25} for the MR7, cold-gas-selected sample are not statistically distinguishable from the distributions of the same quantities obtained from the MR run or the stellar-mass-selected galaxies. For example, the p-value for the two-sided Anderson–Darling test is greater than 0.10 for each of these distributions as compared with the MR7, cold-gas-selected sample.

4. Mock Observations

The previous sections provide a description of infall toward filaments based on the full information available in simulations. In this section, we discuss the feasibility of measuring infall to filaments given observational limitations similar to those of the APPSS survey, in particular:

1. Viewing angle. Even if we assume the filament is oriented across the sky, it will not look the same from all viewing directions unless the filament and the region surrounding it are axially symmetric. Here, we quantify the variation caused by lack of axial symmetry, when the infall profile is determined from galaxies limited to a region similar to that of APPSS.
2. Lack of three-dimensional velocity information. For the simulations, we have three-dimensional information for both position and velocity. For the observations, we have three position coordinates (R.A., decl., and distance) and one coordinate that provides information about the velocity, the line-of-sight redshift. The redshift information, combined with the distance (and the Hubble constant), gives the peculiar motion along the line of sight. The observational coordinates can be converted into distance from the filament ρ and motion toward the filament v_{ρ} , but converting the line-of-sight velocity to motion toward the filament requires assumptions about the three-dimensional direction of motion.
3. Limited sample size. An observational sample will be limited to the galaxies that reside within the survey volume, and the realities of a flux-limited observational sample will mean that fewer low-mass objects are included at greater distances.
4. Measurement uncertainties. While the measurement uncertainties for redshift and sky position are small, the uncertainty for distance may be as high as $\sim 30\%$. Furthermore, with BTFR distance estimates the distance uncertainties are typically log-normal in shape, with long tails extending to higher distances.

4.1. Constructing Mock Observations

In this analysis we construct wedge-shaped volumes around each simulated filament that are designed to approximately reflect the real geometry of APPSS. The filament is placed at a distance of 70 Mpc⁵ and aligned perpendicular to the line of sight. The wedge is truncated at twice the distance to the filament, that is, at $z = 140$ Mpc. The maximum dimensions of the wedge are 80×40 Mpc, such that a section of the filament 40 Mpc long is present in the middle of the wedge, and 10 Mpc

⁵ We note that throughout this section we will use Mpc, to indicate the physical distance with $h = 0.7$.

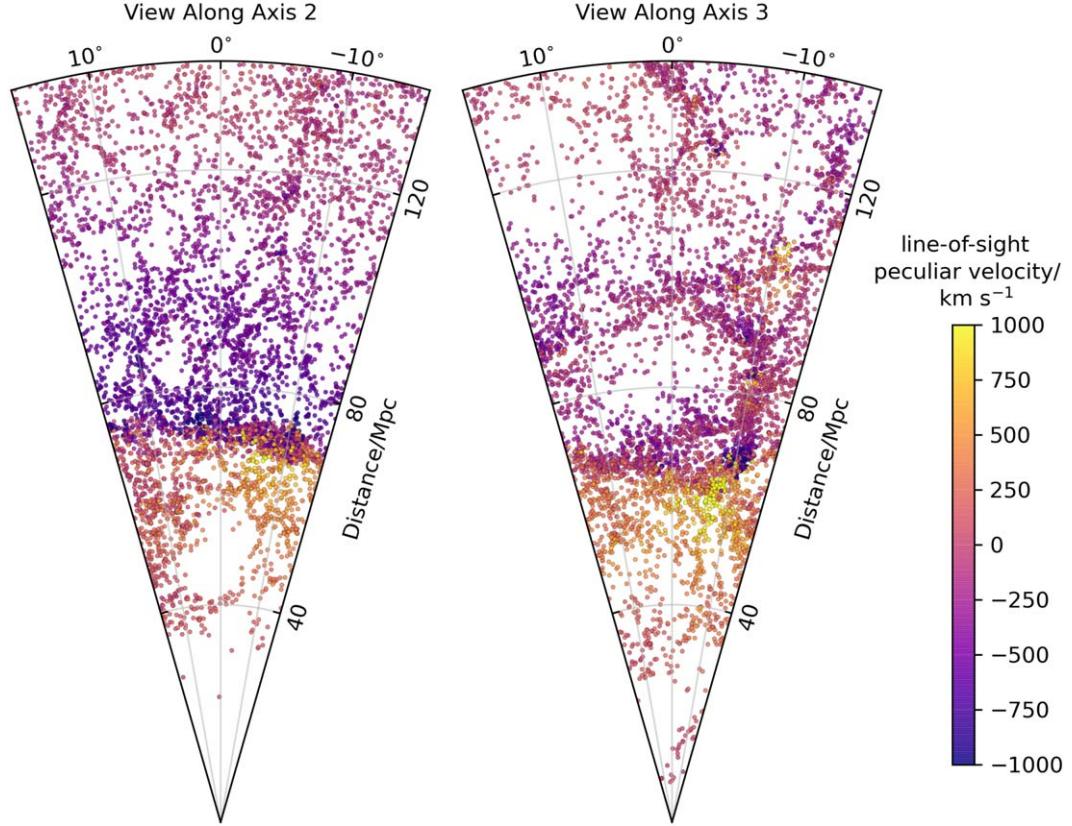


Figure 12. Mock observational wedges for a single filament as seen from two different directions. In both cases, the principal axis of the filament is oriented across the field of view at a distance of 70 Mpc. For the wedge on the left, the observer’s line of sight is along axis 2, and on the right the view is rotated by 90°, along axis 3. Each galaxy is color coded by the component of the peculiar velocity that is along the line of sight. Because the filaments and their surroundings are not axisymmetric, infall looks different from the two different perspectives.

above and below the filament (perpendicular to both the filament axis and the line of sight).

Even if we assume the filament is oriented across the sky, it will not look the same from all viewing directions unless the filament and the region surrounding it are axially symmetric. In order to investigate the degree to which this introduces uncertainty, we create wedges that view each filament from eight equally spaced angles around the principal axis.

In order to mimic the APPSS data, we use a detection limit based on the cold atomic gas mass of each simulated galaxy. Based on the preliminary APPSS data (O’Donoghue et al. 2019) we expect the typical sensitivity to be approximately three times that of ALFALFA (Haynes et al. 2018). Therefore, to construct the mock completeness limit we took the ALFALFA 50% completeness limit (Haynes et al. 2011) for a source with a velocity width of 50 km s^{-1} and decreased the flux by a factor of 3. This gave the following approximate completeness limit for an APPSS-like survey: $\log M_{\text{HI}}/M_{\odot} \geq 2 \log(D/\text{Mpc}) + 4.54$. The real sensitivity of H I observations also depends on the velocity width of each source; however, for this exercise the above approximate limit is sufficient as it requires mock galaxies to have significant cool gas reservoirs and enforces a flux-limited behavior in the sample selection. Figure 12 shows two sample mock observational wedges for a single filament, illustrating the galaxy sample and the degree to which its appearance depends on the viewing angle.

In order to construct infall profiles from mock data, we convert observational coordinates into ρ and v_{ρ} . Because the

observations measure three position coordinates (R.A., decl., and distance), it is just a matter of geometry to find the coordinate ρ (see Appendix A). However, the observer measures only one velocity component: the motion along the line of sight, as determined from the observed redshift and distance. Therefore, it is not possible to convert the velocity to cylindrical filament-based coordinates without making some assumption about the unknown velocity components. One approach is to simply use the line-of-sight peculiar velocity as a proxy for v_{ρ} , with the sign depending on whether the galaxy is on the near side or far side of the filament; this is reasonable to the extent that the galaxies are located in a thin, narrow volume that the filament cuts across. Another approach is to assume that the motion is directly along the coordinate ρ , and that the observed line-of-sight peculiar velocity is simply a component of this motion; an analogous approach is sometimes used to convert observational coordinates to spherical infall coordinates for clusters of galaxies (e.g., Karachentsev & Nasonova 2010).

Figure 13 shows the degree to which each of these approaches is able to reproduce the known, true value of v_{ρ} for one specific mock observational wedge. In this figure, galaxies are color coded by the geometrical correction factor that would be applied to convert the line-of-sight velocity to v_{ρ} under the assumption that the motion is toward the filament (see Appendix A for the derivation of this factor). Both approaches, using the line-of-sight velocity v_{los} (left panel of Figure 13) or applying the correction factor (center panel of Figure 13) work fairly well for galaxies in most of the volume,

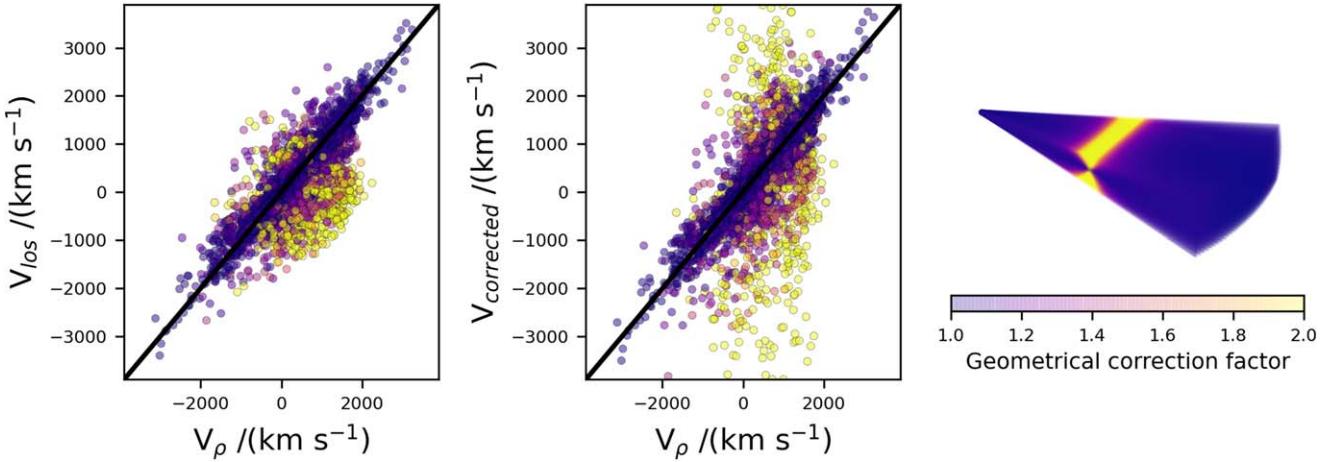


Figure 13. Impact of limited velocity information in the mock observations. The left panel shows the relationship between the measurable line-of-sight velocity component V_{los} and the infall velocity V_{ρ} for galaxies in an observational wedge around a single simulated filament. The central panel is similar, but with V_{los} modified by the geometrical correction factor that would convert it to V_{ρ} if the velocity is in the direction of the filament. The right panel shows the geometry of the observational wedge. All three panels are color coded by the geometrical correction factor. For galaxies in most parts of the observational wedge (blue), V_{los} and $V_{\text{corrected}}$ are both similar to V_{ρ} . However, for galaxies where the line of sight is nearly perpendicular to the direction of infall (yellow), it is difficult to recover the infall velocity, and applying the correction factor introduces additional scatter.

where the correction factor is less than about 1.2 (colored blue in the figure). However, for galaxies above and below the filament, in the region with a high correction factor (colored yellow), values scatter widely from the known values of v_{ρ} , making it important to either drop these galaxies entirely or carefully model their uncertainties. (We take the latter approach below). This occurs because in this region the line-of-sight velocities of the galaxies are close to perpendicular to the radial infall direction. Thus, in practice v_{los} has a minimal component from the true v_{ρ} , and the large correction factors only act to amplify unrelated components of the galaxies' motions. About 20% of the galaxies in the observational wedge are in the region where the correction factor is above 1.2.

Uncertainties are added to the distance values of the simulated galaxies similar to those expected for BTFR distance estimates. Specifically, random Gaussian errors are added to the distance in log-space (from a normal distribution with $\sigma = 0.2$ dex), meaning the uncertainties are log-normally distributed and have a long tail toward larger distances. As the BTFR is a relation between baryonic mass and rotation velocity, and observed mass depends on distance squared, this corresponds to a vertical scatter in the BTFR of 0.4 dex (total scatter, not the intrinsic scatter). This may seem unduly pessimistic; however, many of the APPSS galaxies are expected to be low-mass galaxies, for which the BTFR has increased scatter. For simplicity, no uncertainties are added to either the angular position of each object or the radial velocity, both of which are typically negligible in observations in comparison to the distance uncertainty.

The next, and principal, step is to fit the infall function to these mock data. The large distance uncertainties (± 41 Mpc at 70 Mpc) make this highly challenging. Simply binning the sample with distance would lead to a large fraction of galaxies being placed in incorrect bins and a smoothing out of the true infall pattern. Furthermore, although the positions of the galaxies can in theory be precisely converted to the filament-centric cylindrical coordinate ρ , in practice it is usually impossible to determine (independently of its observed velocity) with confidence whether any given galaxy in the vicinity of the filament is in front or behind it. These limitations rule out a conventional fitting

approach but can be appropriately addressed by the many realizations used in a Markov Chain Monte Carlo (MCMC) approach. We therefore adopt such a fitting approach using the Stan Python package `pystan`.⁶ This removes the need to bin the data as the enormous uncertainties in the distances can be incorporated into a hierarchical model that is used to fit the infall model to the full data set.

We fit each filament using a random sample of 2000 galaxies from the mock catalog produced following the steps above. This is intended to roughly match the number of tracers that the final APPSS sample is expected to have over an equivalent volume. For these computationally intensive MCMC fits, we focus on the PPS-like filaments in the MR7 simulation, performing MCMC fits to mock observations of the 40 filaments with $5625 \pm 1591 \text{ M}_\odot \text{ Mpc}^{-1}$ and group concentration between 0.6 and 1.0. We run MCMC fits for each of these filaments from four different viewing angles, separated by 90° . In addition, we run MCMC fits on a smaller sample of 20 filaments (also from four viewing angles) that span the full range of filament properties. Finally, we also fit a sample of filaments from the MR run, using the same selection criteria, in order to allow comparison to the underlying density field that is unavailable for the MR7 run of the Millennium Simulation.

Our initial tests of this fitting approach demonstrated that given the large distance uncertainties it was unlikely to be able to reliably recover the form of the infall very close to the filament spine (e.g., for $\rho < 10$ Mpc). We therefore adopted a modified version of Equation (2) where an additional parameter, ρ_0 , was added to allow the exponential decay in infall velocity to be fit independently of the slope of the inner linear rise:

$$v = V_{\text{max}} \frac{\rho}{\rho_{\text{max}}}, \quad \rho \leq \rho_{\text{max}} \\ v = V_{\text{exp}} e^{-\sqrt{\frac{\rho}{\rho_0}}}, \quad \rho > \rho_{\text{max}}, \quad (3)$$

where $V_{\text{exp}} = V_{\text{max}} \exp \sqrt{\rho_{\text{max}} / \rho_0}$.

⁶ Stan Development Team. 2021. Stan Modeling Language Users Guide and Reference Manual, 2.28. <https://mc-stan.org>.

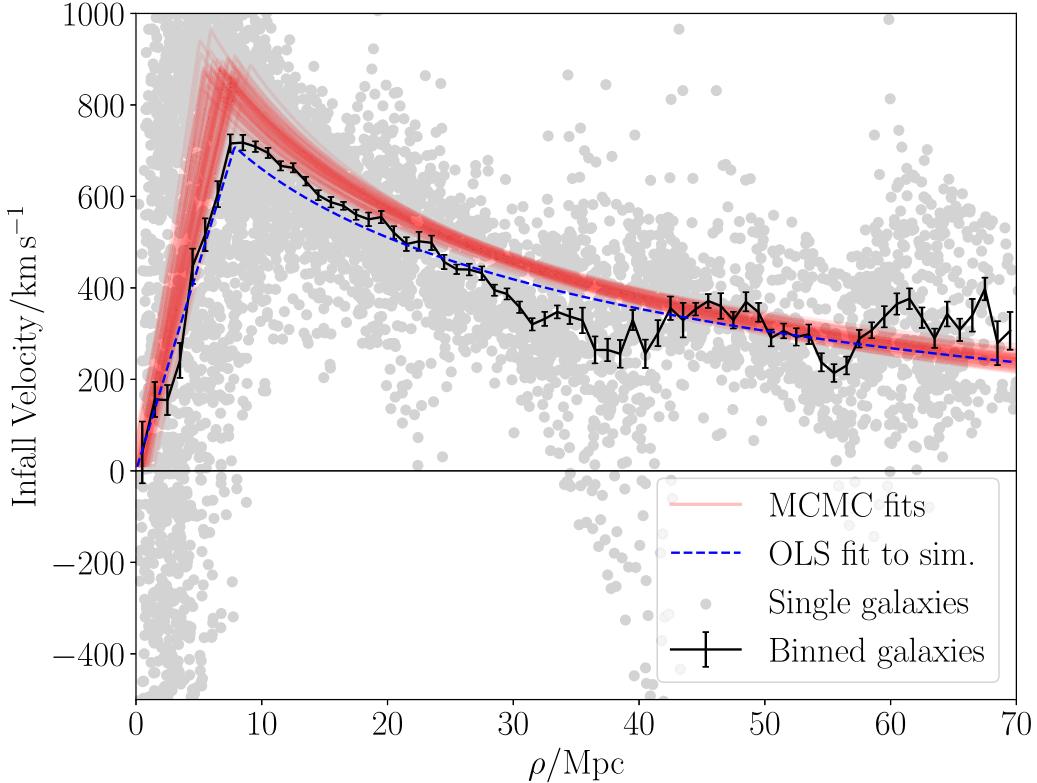


Figure 14. Example of MCMC fits to a single filament. One thousand semi-transparent fits drawn at random from the joint posterior distribution of all parameters (red lines) are overplotted on the simulation data (gray points) for individual galaxies from the observational wedge around the filament (before distance uncertainties were added). Considerable uncertainty exists near ρ_{\max} ; however, by ~ 25 Mpc all fits give consistent values for the infall velocity. The black error bars indicate the mean and standard deviations in bins of width 1 Mpc for all galaxies in the simulation wedge (with no distance uncertainties added), and the blue dashed line is the ordinary least-squares fit to these binned data.

The model parameters fit with the MCMC are those of this infall function plus the distance to the filament spine, the scale of random peculiar velocities (not associated with the filament), and the true distance to every galaxy in the sample. The last of these adds 2000 additional nuisance parameters which are marginalized at the end of the MCMC run. This greatly increases the complexity of the model, but it is necessary to address the enormous (and asymmetric) distance uncertainties that BTFR distance estimates entail. Due to this increase in parameters a normal desktop computer takes approximately 2 h to achieve convergence and fit a single filament.

We provide mostly broad priors for the model as follows: $V_{\max} \sim \text{normal}(\mu = 500 \text{ km s}^{-1}, \sigma = 200 \text{ km s}^{-1})$, $D_{\text{fil}} \sim \text{normal}(\mu = 70 \text{ Mpc}, \sigma = 0.5 \text{ Mpc})$, $\rho_{\max} \sim \text{normal}(\mu = 5 \text{ Mpc}, \sigma = 2 \text{ Mpc})$, $\rho_0 \sim \text{normal}(\mu = 20 \text{ Mpc}, \sigma = 10 \text{ Mpc})$, $\sigma_{V,\text{pec}} \sim \text{normal}(\mu = 200 \text{ km s}^{-1}, \sigma = 50 \text{ km s}^{-1})$. In addition to these priors, all parameters are constrained to be positive. The distance prior is fairly tightly constrained as there are so many galaxies in the real PPS filament that their average velocity can be used to estimate an accurate distance simply by assuming Hubble flow. The prior for the true distance to each galaxy is $p(D_{\text{true}}|D_{\max}) = 3D_{\text{true}}^2/D_{\max}^3$, that is, assuming that the galaxies are distributed randomly throughout the volume out to $D_{\max} = 140$ Mpc. This prior does not include the overdensity of the filament itself (which may be a means to marginally improve the model), but it will account for Malmquist bias by including the prior knowledge that galaxies are likely, on average, to be further away than their distance estimates would imply.

An example of a fit to a particular mock filament is provided in Figure 14, which shows a random selection of fits from the MCMC chain on top of points showing the infall velocity of each galaxy in the sample as a function of ρ . Note that these points correspond to galaxies before distance uncertainties are added, as if the uncertainties were included then no trend would be visible by eye. Appendix B provides a corner plot for this filament, showing posterior distributions of the model parameters.

The final step in the MCMC fitting process is to calculate V_{25} . This is found by evaluating the expression for the infall function at $\rho = 25$ Mpc using the mean fit parameters from the MCMC posterior. Errors are estimated by propagating the uncertainties (1σ standard deviation of posteriors). Figure 15 compares these V_{25} values to those calculated from an ordinary least-squares (OLS) fit to all galaxies in the same mock wedges (before distance uncertainties were added and the samples were reduced to 2000 galaxies), binned into 0.1 Mpc wide bins in ρ . We see that overall there is excellent agreement (within the uncertainties) in the V_{25} values from the MCMC fits and the OLS fits (with no distance uncertainties). The tight correlation between the two measurements of V_{25} demonstrates that the MCMC fitting procedure is able to accurately recover this quantity even with the considerable uncertainties of BTFR distance estimates.

We expand further on this test in Figure 16 where we plot the median deviation between these two fits (essentially the scatter about the one-to-one line in Figure 15) as a function of both the number of tracers used in the MCMC fits and the scale of the BTFR distance uncertainties that we assume. The behavior is broadly as expected, with the typical size of deviations

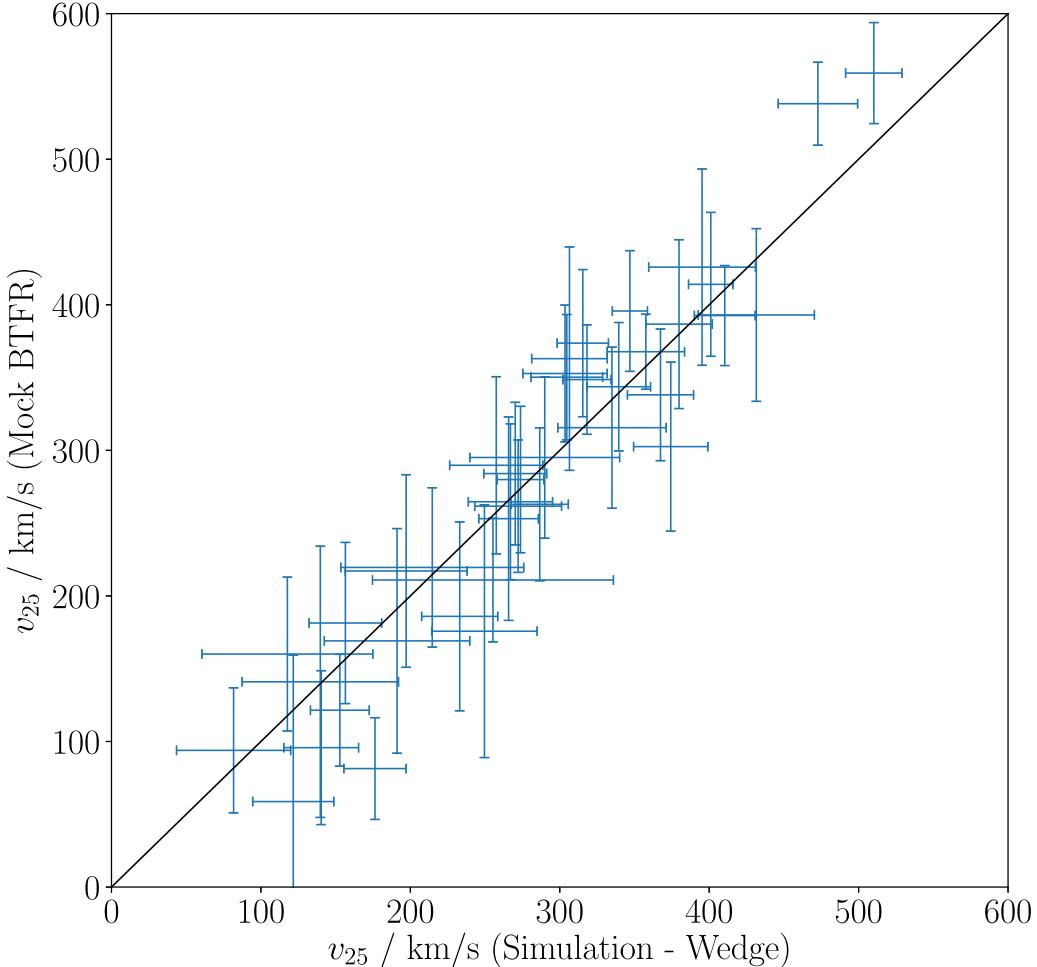


Figure 15. Infall velocity at 25 Mpc for observational wedges before (x-axis) and after (y-axis) introducing the effects of BTFR-like distance uncertainties. Each point corresponds to a single PPS-like filament and each V_{25} value (y-axis) was calculated from an MCMC fit of 2000 galaxies with 0.2 dex distance uncertainties (red lines in Figure 14). For each simulation wedge (without distance uncertainties) all galaxies were averaged in bins of width 200 kpc in radial distance from the principal axis. The infall function was then fitted via an OLS fit to the bin median values (dashed blue line in Figure 14). The uncertainties for the mock values represent propagation of the standard deviations of the posterior distributions (about the mean) of ρ_{\max} , V_{\max} , and ρ_0 . For the simulation error bars the same propagation of uncertainty was performed for the 1σ OLS fit uncertainties in these parameters.

decreasing as more tracers are used in the fits and as the assumed BTFR distance uncertainties decrease. For small sample sizes ($N < 500$) the accuracy of the fits rapidly improves with increasing sample size, but beyond $N = 2000$ there is very little benefit to expanding the sample of tracers. The improvements from reducing the scale of the distance uncertainties can be seen for all sample sizes; however, it is most pronounced for small samples.

4.2. Results from Mock Observations

Each source of observational uncertainty has the potential to introduce challenges to measuring V_{\max} , ρ_{\max} , and V_{25} . Figure 17 summarizes how well we can recover the values of these three parameters as we include progressively more sources of uncertainty. The top row shows the scatter that is introduced by observing the filament from a specific viewing angle, taking into account galaxy sample selection based on HI flux as described above. Most of the observational wedges include about 5000 galaxies above this limit. The values obtained from the wedge-shaped observational volumes are compared with the values obtained using the full, three-dimensional information from galaxies in the cylinders, as

described in Section 3. For example, the top left panel shows the average value for V_{\max} obtained from the eight observational wedges that view the filament from different angles around the principal axis, with error bars showing the standard deviation. These are plotted against the values for V_{\max} obtained using the three-dimensional cylinders. The straight line indicates the identity relation, and the inset histogram shows the scatter around the identity relation. The histogram includes each of the viewing angles for each filament as separate data points, and shows the best-fit Gaussian curve and the standard deviation. For V_{\max} , for example, the values for observational wedges have a Gaussian distribution about the identity relation with a scatter of 66 km s^{-1} .

The second row of Figure 17 includes an additional source of uncertainty: the fact that only the line-of-sight motion V_{los} is available in the observations. To do this, we consider the results from our attempts to geometrically recover V_{ρ} from V_{los} combined with position information. As shown in Figure 13, V_{los} and V_{ρ} are very similar for galaxies in most parts of the observational wedge but not for galaxies just above and below the filament, where only a small component of V_{ρ} is along the line of sight. About 20% of the galaxies in the wedge are

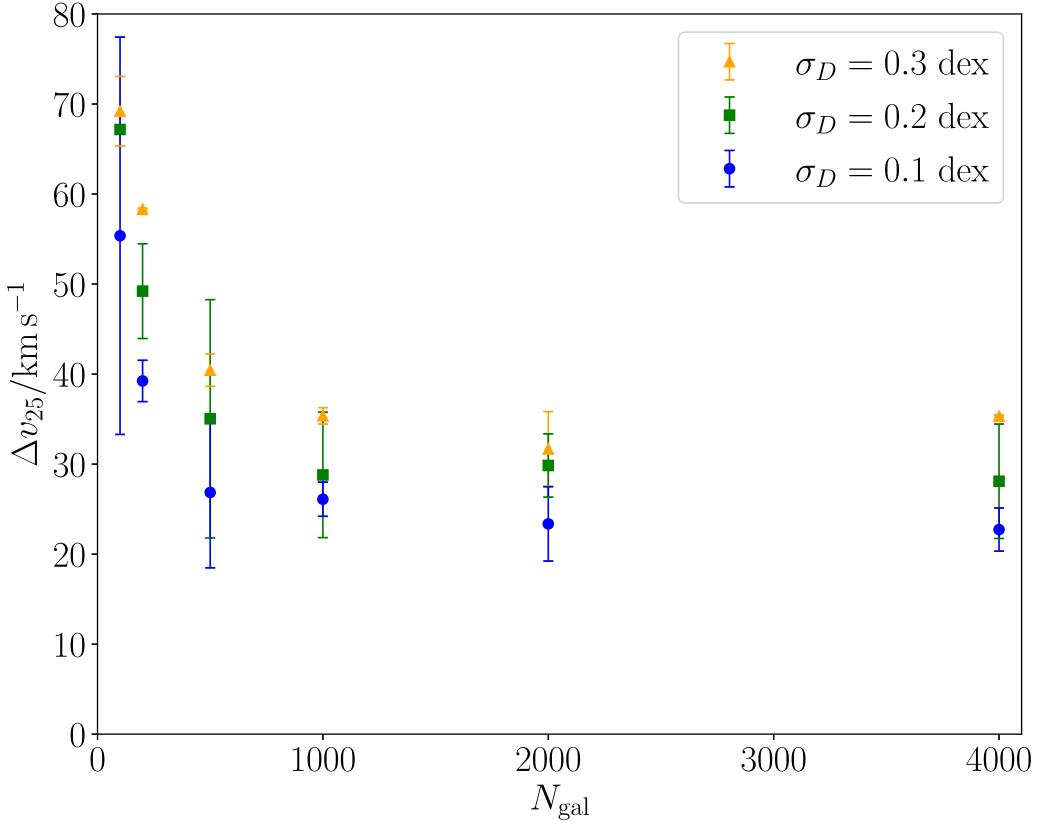


Figure 16. Median offset between the mock V_{25} fit value and that of the simulation wedge with no distance uncertainties as a function of the sample size and scale of distance uncertainties. Each point corresponds to the median value over 40 PPS-like filaments fit using MCMC. Error bars correspond to jackknife standard deviation estimates for the 40 filaments fit in each data point.

typically in the region where the geometrical correction factor is greater than 1.2. The uncertainty in V_ρ is so large for these galaxies, whether or not we attempt a correction using the geometrical correction factor in Figure 13, that we remove them from the sample. Note that the option of removing galaxies that are in this region is only feasible if their distances can be determined with precision. In the context of the second row of Figure 18, we use this approach as a way to estimate the scatter introduced specifically by the lack of three-dimensional velocity information when combined with the scatter from using observational wedges. (However, when we include distance uncertainties using MCMC fits, we include all the galaxies.) For example, the scatter in V_{max} rises to 108 km s^{-1} in the second row, compared with 66 km s^{-1} in the first row. If the uncertainties add in quadrature, this implies an additional scatter of $\sqrt{108^2 - 66^2} \text{ km s}^{-1} = 85 \text{ km s}^{-1}$ specifically from the observational limitation to one velocity component.

The third row of Figure 17 combines the limitations from the first two rows with a limited sample size of 2000 galaxies, down from about 5000 that are above the HI flux limit for most filaments. This adds only a small increase in scatter—for example, small enough that the scatter in V_{max} is still about 108 km s^{-1} .

The fourth row of Figure 17 includes all the uncertainties in our mock observations, including distance uncertainties of 0.2 dex, taken into account using MCMC fits. Here we see significant limitations to the two-parameter description of infall in terms of V_{max} and ρ_{max} . The scatter for V_{max} is both large and shifted to higher values, while the values for ρ_{max} mock observations match only for values less than about 12 Mpc;

otherwise they are considerably underestimated. Using a one-parameter description of infall in terms of V_{25} is more promising. In terms of the adjusted coefficient of determination R_{adj}^2 , the best-fit (dashed) lines for V_{max} , ρ_{max} , and V_{25} have values of 0.237, 0.278, and 0.485, respectively.

We next consider the possibility of predicting the infall rate based on the structure of the filament. As we showed in the top rows of Figures 10 and 11, V_{max} , ρ_{max} , and V_{25} are related to the group masses and group concentration of the filaments. Figure 18 shows how the scatter in these relationships increases as we introduce observational uncertainties. The top row shows the original scatter in these relationships for the full cylinders, and the following rows introduce each type of observational uncertainty, using the same format as in Figure 17. The bottom row, which includes all the uncertainties, reflects the same limitations seen in the bottom row of Figure 16. For V_{max} the scatter is high and the values are offset from the original relation, and for ρ_{max} the values are offset and the slope is flatter. As expected, a more promising approach in the context of our full sample of simulated filaments is to use V_{25} . In terms of the adjusted coefficient of determination R_{adj}^2 , the best-fit (dashed) lines for V_{max} , ρ_{max} , and V_{25} have values of 0.135, 0.213, and 0.396, respectively. Taking into account all the sources of uncertainty, filaments that are “like the PPS” in their observable group structure, as defined in Section 3, have an average $V_{\text{max}} = 791 \text{ km s}^{-1}$ with standard deviation 230 km s^{-1} , an average $\rho_{\text{max}} = 7.0 \text{ Mpc}$ with standard deviation 3.5 Mpc , and an average $V_{25} = 301 \text{ km s}^{-1}$ with standard deviation 150 km s^{-1} .

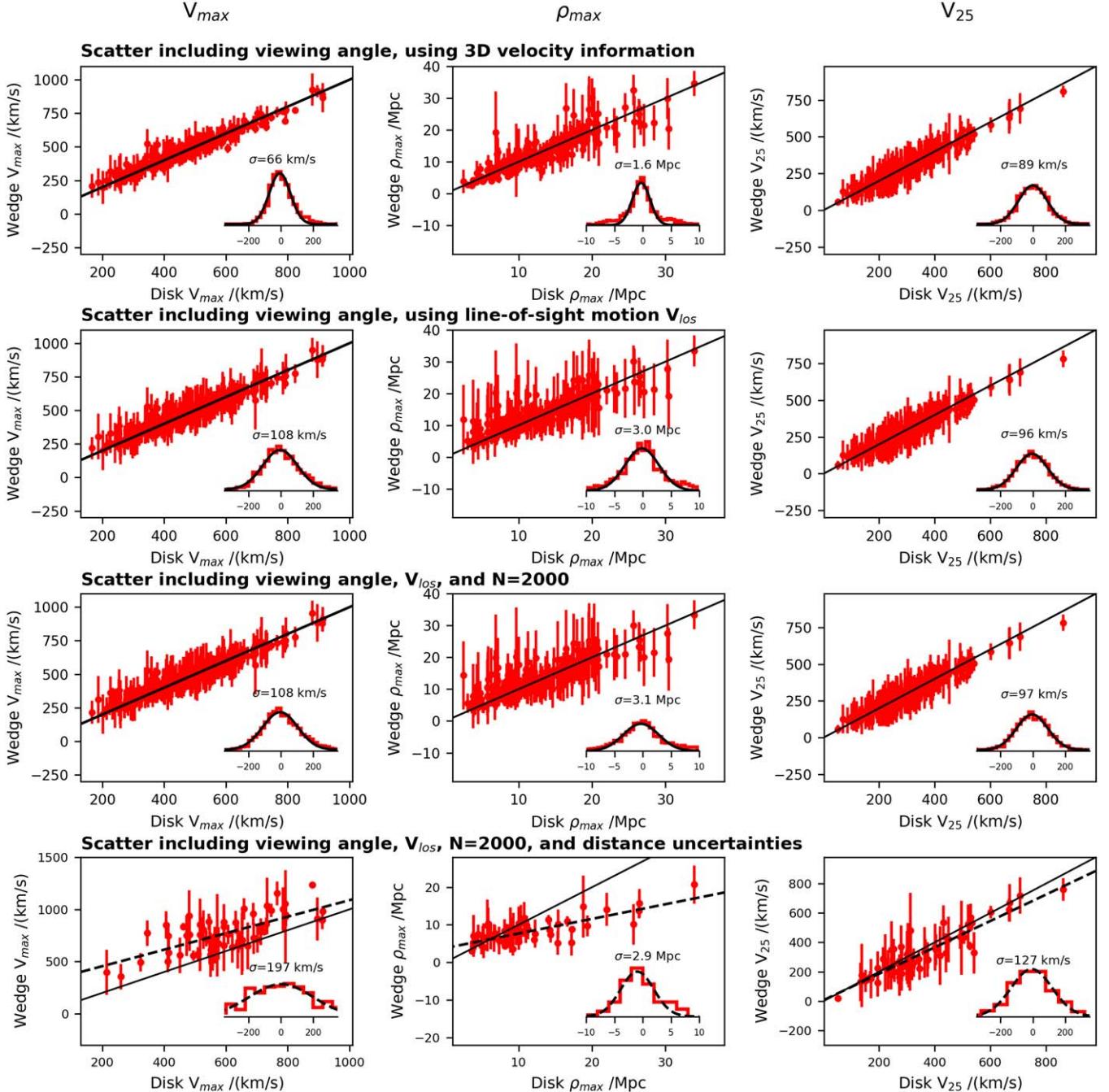


Figure 17. Scatter in V_{\max} (left column), ρ_{\max} (center column), and V_{25} (right column) introduced by adding progressively more observational limitations: viewing angle (top row), velocity limited to line-of-sight component (second row), sample size limited to 2000 galaxies (third row), and distance uncertainties (bottom row). Each panel shows the relationship between the mock observational wedges and the value obtained using a cylinder and full 3D information. Points show the average for the wedges for each filament as seen from different angles, with error bars showing the standard deviation. Solid lines show in each panel show the identity relation, and the dashed lines in the bottom row show the best fit once all uncertainties are included. Inset histograms show the deviation of the values for all the wedges from the 3D cylinder values (for the top three rows) and the deviation from the dashed best-fit line (for the bottom row). As more observational limitations are added, the deviation tends to increase.

Finally, we consider the relationships between V_{\max} , ρ_{\max} , and V_{25} and the underlying mass distribution. As we showed in the bottom rows of Figures 10 and 11, V_{\max} , ρ_{\max} , and V_{25} are related to the mass and mass concentration of the filaments, including dark matter. Figure 19 shows how the scatter in these relationships increases as we introduce observational uncertainties. In terms of the adjusted coefficient of determination R_{adj}^2 , the best-fit (dashed) lines for the fits that include all uncertainties (the bottom row) V_{\max} , ρ_{\max} , and V_{25} have values

of 0.570, 0.153, and 0.427, respectively. In this case, the relationship for V_{\max} shows the most predictive value and the relationship for ρ_{\max} shows the least predictive value.

5. Summary and Discussion

In this paper, we characterize the theoretically expected large-scale infall pattern of galaxies toward cosmological filaments, and we model the uncertainties in determining the infall profile using measurements like those in the Arecibo

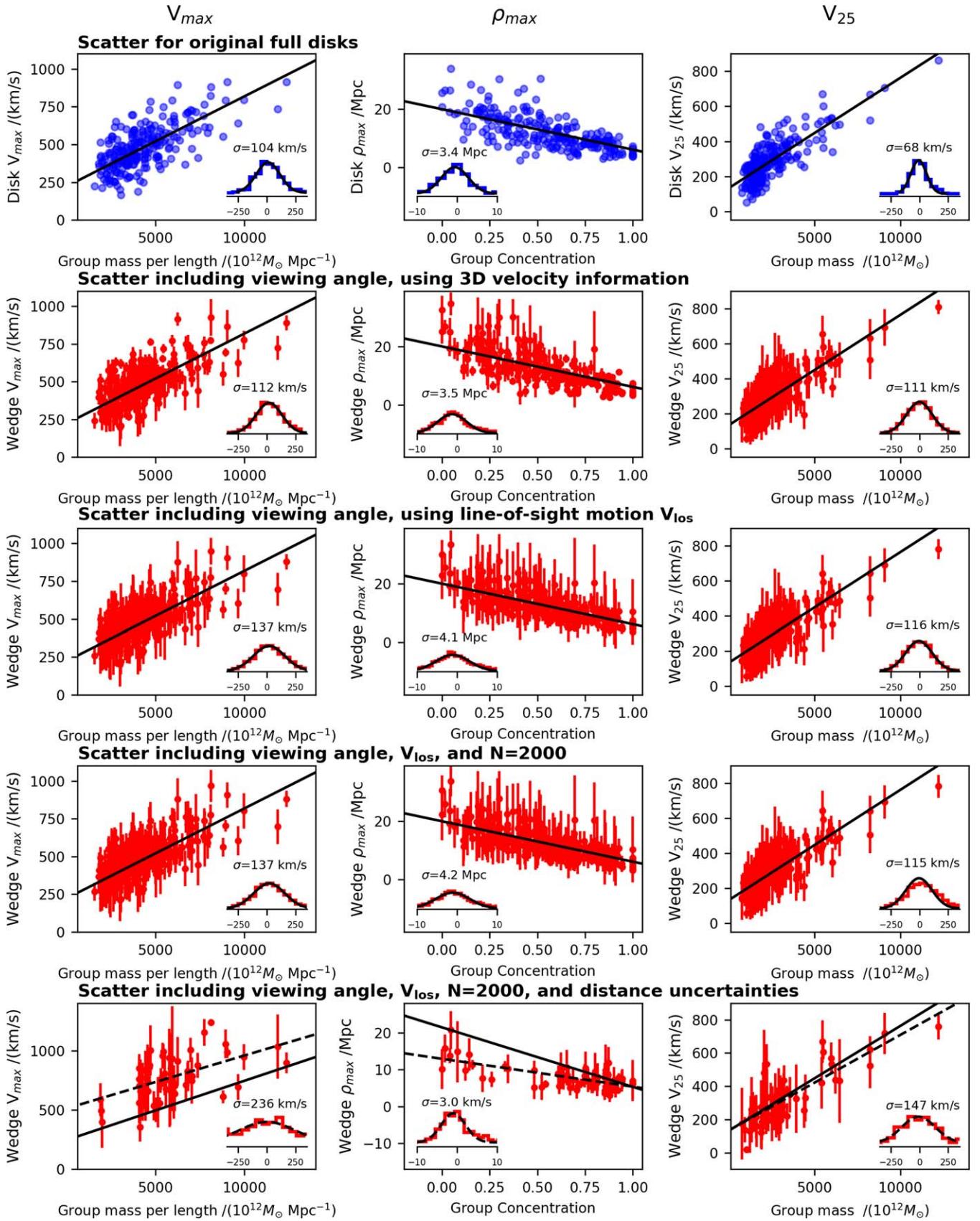


Figure 18. Scatter in relationships between V_{\max} , ρ_{\max} , and V_{25} and the group properties of each filament, including the effects of observational limitations. The layout is similar to Figure 17, except that the top row shows the scatter for the original cylinders before introducing observational limitations. Solid lines show results from the fits to the original cylinders. Dashed lines in the bottom row show results from fits to the mock observations once all uncertainties are included. With all uncertainties included, it becomes especially difficult to recover the relationships for V_{\max} , ρ_{\max} . The relationship for V_{25} is more robust.

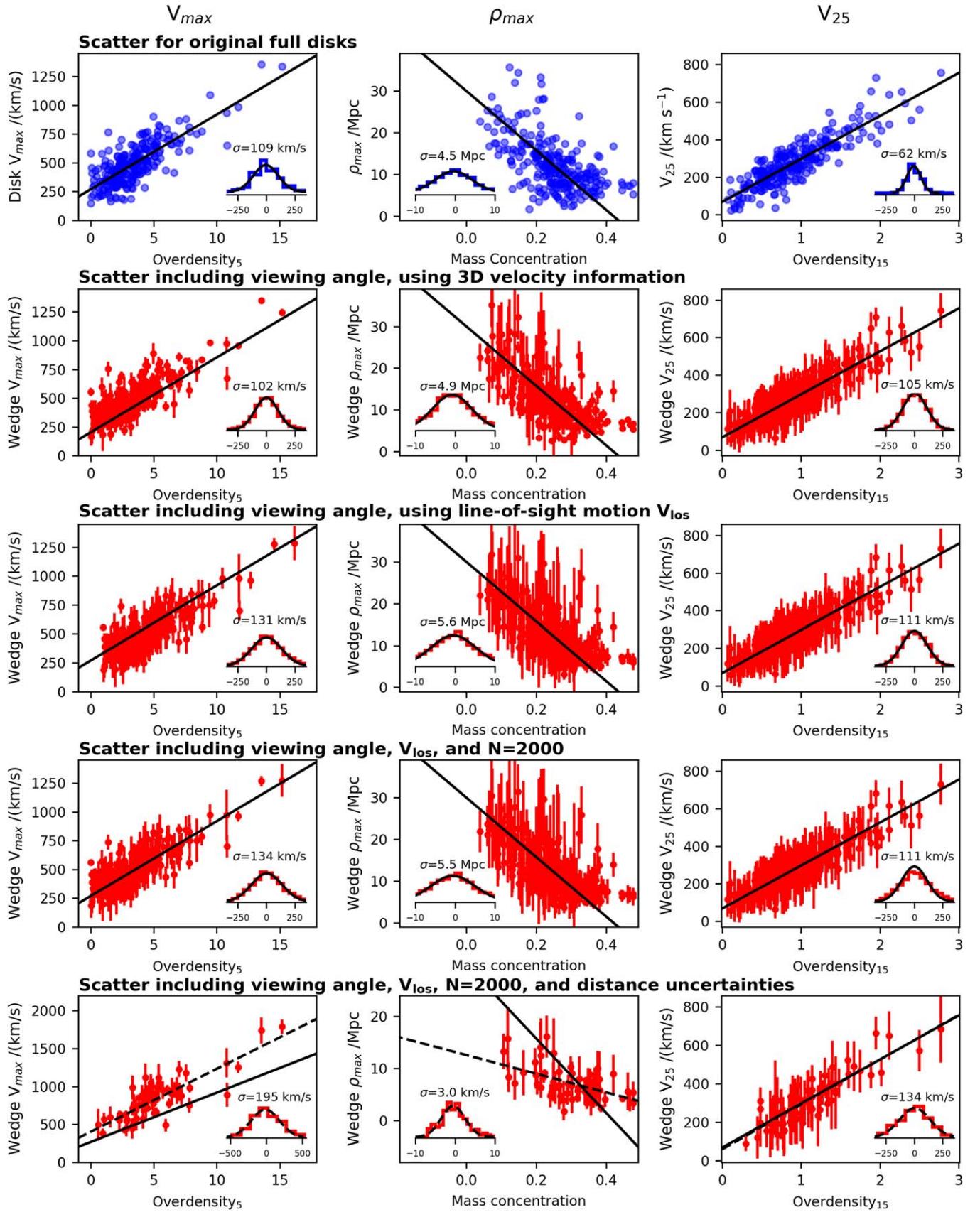


Figure 19. Scatter in relationships between V_{\max} , ρ_{\max} , and V_{25} and the underlying matter distribution of filaments, including the effects of observational limitations. The layout is the same as Figure 18.

Pisces–Perseus Supercluster Survey (APPSS). Our results can be summarized as follows:

1. The average infall profile, when stacked by distance from the principal axis, follows the equation $v_p = C\rho^4 e^{-\rho^{0.2}}$. The rounded peak in this function is the result of averaging together filaments with velocity profiles that peak at different distances. When the distance for each filament is normalized by its peak infall distance the average profile has a sharper peak, and is better fit by the piecewise function given by Equation (2). This piecewise function is the appropriate one to use in fitting the profile of an individual filament.
2. The infall for each filament can be characterized by the maximum infall velocity V_{\max} and the distance of the point of maximum infall from the principal axis of the filament ρ_{\max} . These two parameters are related to the properties of the groups in the filament: V_{\max} is positively correlated with the mass in halos per length along the filament, and ρ_{\max} is negatively correlated with the degree to which the halos are concentrated toward the principal axis. Based on these constraints, we expect the PPS infall region have $V_{\max} = 612 \pm 116 \text{ km s}^{-1}$ and $\rho_{\max} = 8.9 \pm 2.1 \text{ Mpc}$ (assuming $h = 0.7$).
3. A simpler way to characterize infall, which is easier to determine observationally, is in terms of the single parameter V_{25} , the infall at a distance of 25 Mpc from the axis of the filament. Based on the mass of the central groups in the PPS, we predict $V_{25} = 329 \pm 68 \text{ km s}^{-1}$.
4. The values for V_{\max} , ρ_{\max} , and V_{25} also reflect the underlying mass distribution in the filament, including dark matter, so they have the potential to be used to constrain the large-scale dark matter distribution. The specific relationships between these parameters and the mass distribution are given in Figures 10 and 11.
5. Our results are robust to the small differences in cosmological parameters between the MR7 and MR runs, and to the choice of selecting galaxies by cold-gas mass or stellar mass.
6. A significant source of uncertainty for observational measurements is the viewing angle of the filament, even if we assume the filament is oriented across the sky. For example, this introduces a scatter of 66 km s^{-1} in the relationship between V_{\max} as determined using a 3D cylinder versus that obtained from observational wedges for the same filament.
7. An additional source of uncertainty in determining the infall profile from observational data is the availability of only one velocity component, the motion along the line of sight. This is primarily a problem for galaxies in the region near the filament where the line of sight is nearly perpendicular to the direction of infall. For example, if we include only galaxies with velocity correction factors less than 1.2, then including this source of uncertainty along with the uncertainty in viewing angle increases the scatter in V_{\max} to about 108 km s^{-1} .
8. The galaxy sample size introduces a significant source of additional uncertainty only if the sample size is smaller than about 2000.
9. A particular challenge in measuring the infall profile is introduced by the uncertainty in galaxy distances, which can be approximately 30% using the BTFR. We model this uncertainty using MCMC sampling, and find that it is

particularly difficult to recover the expected values for V_{\max} and ρ_{\max} , while recovering the single parameter V_{25} is more promising.

10. Taking into account all sources of observational uncertainty, we find that of the three parameters V_{\max} , ρ_{\max} , and V_{25} , it is V_{25} that has the tightest correlation to the observable group structure of the filament, and V_{\max} that has the tightest correlation with the underlying mass distribution.
11. Taking into account the measurement uncertainties we expect for APPSS, we expect our measurements of the three parameters for the PPS to be in the ranges $V_{\max} = 791 \pm 230 \text{ km s}^{-1}$, $\rho_{\max} = 7.0 \pm 3.5 \text{ Mpc}$, and $V_{25} = 301 \pm 150 \text{ km s}^{-1}$.

We emphasize that in this context we use a definition of filaments that is straightforward to apply to observations like those available for the PPS, and therefore does not take into account the detailed structure that can be seen in simulations. For example, we determine the axis of the filament using principal component analysis on cluster-sized halos, resulting in a simple, straight-line estimate of the axis. This differs from studies that trace the curves of a filament using more fine-grained measures of the density. If we did this, we would be more sensitive to infall on smaller spatial scales, including the infall to individual halos, rather than the average infall on larger scales.

Although this study was motivated by the APPSS project, which focuses on one particular nearby filament, it would be especially interesting to measure the average infall profile for a sample of multiple filaments. One advantage of such a study, is that it would allow us to test the average infall profile in Equation (2), as we expect the data from one filament alone to have too much scatter from the average profile to provide an effective test. Another advantage would be that it would allow us to test the the relationships we obtain between V_{\max} , ρ_{\max} , and V_{25} and the structure of the filaments.

We are grateful for the contributions of all members of the Undergraduate ALFALFA Team, especially Trevor Viscardi, Martha Haynes, and Rebecca Koopmann. We also thank John Helly for assistance in accessing the Millennium database. This work has been supported by NSF grants AST-1211005, AST-1637339, and AST-1637271.

Appendix A

Derivation of Infall Profiles from Mock Observations

Here we show the derivation for converting positions and velocities from mock observations into infall profiles. The mock observations can be described using spherical coordinates with the observer at the origin. The observer measures three position coordinates: the distance r , and the angular positions on the sky θ and ϕ . We can express the observer's coordinates in terms of Cartesian coordinates using

$$x = r \sin \theta \cos \phi, \quad (A1)$$

$$y = r \sin \theta \sin \phi, \quad (A2)$$

$$z = r \cos \theta. \quad (A3)$$

If the filament is positioned parallel to the y -axis at $x = D_{\text{fil}}$ and $z = 0$, the distance of a given point from the principal axis of

the filament can be written as

$$\rho = \sqrt{(x - D_{\text{fil}})^2 + z^2}. \quad (\text{A4})$$

This is similar to the usual cylindrical coordinate ρ , except that it is perpendicular to the y -axis instead of the z -axis, and the x -axis is shifted by D_{fil} . Substituting expressions A1 and A3 into expression A4 gives ρ in terms of the observational coordinates:

$$\rho = \sqrt{r^2 \cos^2 \theta + (r \sin \theta \cos \phi - D_{\text{fil}})^2}. \quad (\text{A5})$$

Because the observer can measure three independent position coordinates, it is just a matter of geometry to obtain expression A5. On the other hand, the observer can measure only one velocity component: the motion along the line-of-sight v_r . Therefore, it is not possible to convert the velocity to cylindrical filament-based coordinates without making some assumption about the unknown velocity components. Here we take the approach of assuming that the motion is directly along the coordinate ρ ; an analogous approach is sometimes used to convert observational coordinates to spherical infall coordinates for clusters of galaxies (e.g., Karachentsev & Nasanova 2010). In this case, we can relate the line-of-sight component of the velocity v_r to the motion toward the filament v_ρ via the dot product

$$v_r = \mathbf{v} \cdot \hat{\mathbf{r}} = v_\rho \hat{\mathbf{r}} \cdot \hat{\mathbf{r}}, \quad (\text{A6})$$

where

$$\hat{\mathbf{r}} = \frac{x}{r} \hat{\mathbf{x}} + \frac{y}{r} \hat{\mathbf{y}} + \frac{z}{r} \hat{\mathbf{z}} \quad (\text{A7})$$

and

$$\hat{\mathbf{r}} = \frac{x - D_{\text{fil}}}{\rho} \hat{\mathbf{x}} + \frac{z}{\rho} \hat{\mathbf{z}}. \quad (\text{A8})$$

Writing A7 and A8 in terms of the observer's spherical coordinates using A1–A3, and taking the dot product,

$$\hat{\mathbf{r}} \cdot \hat{\mathbf{r}} = \frac{\sin \theta \cos \phi (r \sin \theta \cos \phi - D_{\text{fil}}) + r \cos^2 \theta}{\sqrt{(r \sin \theta \cos \phi - D_{\text{fil}})^2 + r^2 \cos^2 \theta}}. \quad (\text{A9})$$

Finally, solving for v_ρ in A6 and substituting A9 for the dot product, we obtain

$$v_\rho = v_r \frac{\sqrt{(r \sin \theta \cos \phi - D_{\text{fil}})^2 + r^2 \cos^2 \theta}}{\sin \theta \cos \phi (r \sin \theta \cos \phi - D_{\text{fil}}) + r \cos^2 \theta}. \quad (\text{A10})$$

This provides the expression for the motion toward the filament in terms of the observational coordinates. The factor by which v_r is multiplied to obtain v_ρ is the geometrical correction factor illustrated in Figure 13.

Appendix B Sample MCMC Fit

Here we provide an example of our Markov Chain Monte Carlo (MCMC) sampling method applied to a particular mock filament. Figure 20 shows a corner plot of the joint posterior distribution of the five model parameters. This mock filament is the same one for which a random selection of fits from the MCMC chain are provided in Figure 14.

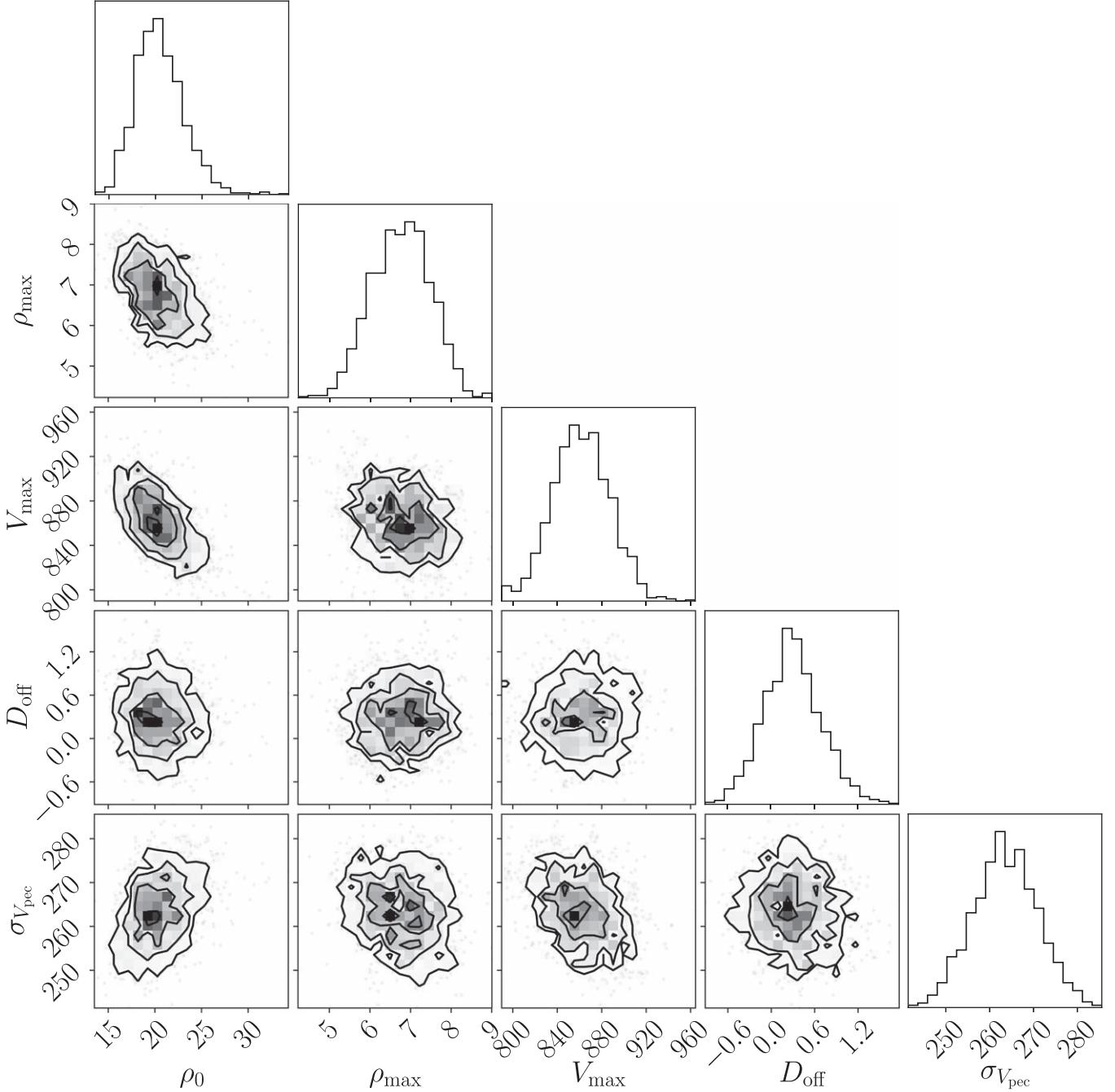


Figure 20. Example corner plot for one filament showing the posterior distributions of the model parameters fit with MCMC. In addition, to the three model parameters (ρ_0 , ρ_{\max} , and V_{\max}) the MCMC fit also fits for any offset in the distance to the filament spine (D_{off}) and fits for the typical scale of peculiar velocities (not associates with infall on to the filament), $\sigma_{V_{\text{pec}}}$.

ORCID iDs

Mary Crone Odekon <https://orcid.org/0000-0003-0162-1012>
 Michael G. Jones <https://orcid.org/0000-0002-5434-4904>
 Lucas Graham <https://orcid.org/0000-0001-8165-2323>
 Jessica Kelley-Derzon <https://orcid.org/0000-0002-7727-1824>
 Evan Halstead <https://orcid.org/0000-0001-9735-4232>

References

Alpaslan, M., Robotham, A. S. G., Obreschkow, D., et al. 2014, *MNRAS*, **440**, L106
 Bonnaire, T., Aghanim, N., Decelle, A., & Douspis, M. 2020, *A&A*, **637**, A18
 Bower, R. G., Benson, A. J., Malbon, R., et al. 2006, *MNRAS*, **370**, 645
 Buehlmann, M., & Hahn, O. 2019, *MNRAS*, **487**, 228
 Carrón Duque, J., Migliaccio, M., Marinucci, D., & Vittorio, N. 2022, *A&A*, **659**, A166
 Castignani, G., Combes, F., Jablonka, P., et al. 2022, *A&A*, **657**, A9
 Coe, D. 2010, arXiv:1005.0411
 Colberg, J. M., Krughoff, K. S., & Connolly, A. J. 2005, *MNRAS*, **359**, 272
 Crone Odekon, M., Hallenbeck, G., Haynes, M. P., et al. 2018, *ApJ*, **852**, 142
 Croton, D. J., Springel, V., White, S. D. M., et al. 2006, *MNRAS*, **365**, 11
 De Lucia, G., Springel, V., White, S. D. M., Croton, D., & Kauffmann, G. 2006, *MNRAS*, **366**, 499
 Dubois, Y., Pichon, C., Welker, C., et al. 2014, *MNRAS*, **444**, 1453
 Duffy, A. R., Schaye, J., Kay, S. T., & Dalla Vecchia, C. 2008, *MNRAS*, **390**, L64
 Einasto, M., Deshev, B., Tenjes, P., et al. 2020, *A&A*, **641**, A172
 Eisenstein, D. J., Loeb, A., & Turner, E. L. 1997, *ApJ*, **475**, 421

Galárraga-Espinosa, D., Aghanim, N., Langer, M., Gouin, C., & Malavasi, N. 2020, *A&A*, **641**, A173

Gouin, C., Aghanim, N., Bonjean, V., & Douspis, M. 2020, *A&A*, **635**, A195

Guo, Q., White, S., Angulo, R. E., et al. 2013, *MNRAS*, **428**, 1351

Haynes, M. P., & Giovanelli, R. 1986, *ApJL*, **306**, L55

Haynes, M. P., Giovanelli, R., Kent, B. R., et al. 2018, *ApJ*, **861**, 49

Haynes, M. P., Giovanelli, R., Martin, A. M., et al. 2011, *AJ*, **142**, 170

Howlett, C., Said, K., Lucey, J. R., et al. 2022, *MNRAS*, **515**, 953

Karachentsev, I. D., & Nasonova, O. G. 2010, *MNRAS*, **405**, 1075

Kourkchi, E., Tully, R. B., Eftekharzadeh, S., et al. 2020, *ApJ*, **902**, 145

Kraljic, K., Pichon, C., Dubois, Y., et al. 2019, *MNRAS*, **483**, 3227

Kuchner, U., Aragón-Salamanca, A., Pearce, F. R., et al. 2020, *MNRAS*, **494**, 5473

Lacey, C. G., Baugh, C. M., Frenk, C. S., et al. 2016, *MNRAS*, **462**, 3854

Lagos, C. D. P., Lacey, C. G., & Baugh, C. M. 2012, arXiv:1210.4974

Laigle, C., Pichon, C., Codis, S., et al. 2015, *MNRAS*, **446**, 2744

Lelli, F., McGaugh, S. S., & Schombert, J. M. 2016, *ApJL*, **816**, L14

Libeskind, N. I., van de Weygaert, R., Cautun, M., et al. 2018, *MNRAS*, **473**, 1195

Mahajan, S., Singh, A., & Shobhana, D. 2018, *MNRAS*, **478**, 4336

Martínez, H. J., Muriel, H., & Coenda, V. 2016, *MNRAS*, **455**, 127

McGaugh, S. S., Schombert, J. M., Bothun, G. D., & de Blok, W. J. G. 2000, *ApJL*, **533**, L99

O'Donoghue, A. A., Haynes, M. P., Koopmann, R. A., et al. 2019, *AJ*, **157**, 81

Papastergis, E., Adams, E. A. K., & van der Hulst, J. M. 2016, *A&A*, **593**, A39

Pereyra, L. A., Sgró, M. A., Merchán, M. E., Stasyszyn, F. A., & Paz, D. J. 2020, *MNRAS*, **499**, 4876

Pillepich, A., Springel, V., Nelson, D., et al. 2018, *MNRAS*, **473**, 4077

Rost, A., Stasyszyn, F., Pereyra, L., & Martínez, H. J. 2020, *MNRAS*, **493**, 1936

Saez, C., Campusano, L. E., Cyriano, E. S., Sodré, L., & Kneib, J. P. 2016, *MNRAS*, **460**, 4453

Shaya, E. J., Tully, R. B., Pomarède, D., & Peel, A. 2022, *ApJ*, **927**, 168

Sorce, J. G., & Tempel, E. 2018, *MNRAS*, **476**, 4362

Springel, V., White, S. D. M., Jenkins, A., et al. 2005, *Natur*, **435**, 629

Springob, C. M., Magoulas, C., Colless, M., et al. 2014, *MNRAS*, **445**, 2677

Tully, R. B. 2015, *AJ*, **149**, 171

Tully, R. B., Courtois, H. M., & Sorce, J. G. 2016, *AJ*, **152**, 50

Tully, R. B., & Fisher, J. R. 1977, *A&A*, **500**, 105

Wang, P., Libeskind, N. I., Tempel, E., Kang, X., & Guo, Q. 2021, *NatAs*, **5**, 839

Xia, Q., Neyrinck, M. C., Cai, Y.-C., & Aragón-Calvo, M. A. 2021, *MNRAS*, **506**, 1059