A Census of 163 Large-scale (>10 pc), Velocity-coherent Filaments in the Inner Galactic Plane: Physical Properties, Dense-gas Fraction, and Association with Spiral Arms

Yifei Ge1,2 and Ke Wang1

1 Kavli Institute for Astronomy and Astrophysics, Peking University, 5 Yiheyuan Road, Haidian District, Beijing 100871, People’s Republic of China
2 Department of Astronomy, School of Physics, Peking University, 5 Yiheyuan Road, Haidian District, Beijing 100871, People’s Republic of China

Abstract

The interstellar medium has a highly filamentary and hierarchical structure that may play a significant role in star formation. A systematical study of the large-scale filaments toward their physical parameters, distribution, structures, and kinematics will inform us about which types of filaments have the potential to form stars, how the material feeds protostars through filaments, and the connection between star formation and Galactic spiral arms. Unlike the traditional by-eye searches, we use a customized minimum spanning tree algorithm to identify filaments by linking Galactic clumps from the APEX Telescope Large Area Survey of the Galaxy catalog. In the inner Galactic plane (|l| < 60°), we identify 163 large-scale filaments with physical properties derived, including the dense-gas mass fraction, and we compare them with an updated spiral arm model in position–position–velocity space. The dense-gas mass fraction is found not to differ significantly in various Galactic positions or in different spiral arms. We also find that most filaments are interarm filaments after adding a distance constraint, and filaments in arms differ a little with those not in arms. One surprising result is that clumps on and off filaments have no significant distinction in their mass at the same size.

Unified Astronomy Thesaurus concepts: Catalogs (205); Interstellar filaments (842); Galaxy structure (622); Star formation (1569); Interstellar clouds (834)

Supporting material: extended figures, machine-readable table

1. Introduction

Conspicuous filamentary structures are ubiquitous in the interstellar medium (ISM). It is not clearly understood how the filaments form, however. Numerical simulations of the multi-phase atomic ISM in the Galactic disk show that cold, dense gas prefers to organize itself naturally in the filamentary network (Tasker & Tan 2009). The widths derived from observed filaments support the argument that the filaments may form as a result of the dissipation of large-scale turbulence (Arzoumanian et al. 2011). Filaments are constituents of molecular clouds and may play a pivotal role in the process of star formation (e.g., Schneider & Elmegreen 1979; Arzoumanian et al. 2011; Zhang et al. 2019). According to simulations, massive star-forming clumps are filamentary structures in their early stages (Smith et al. 2009). Andre et al. (2010) observed filaments in both quiescent and star-forming clouds, where a large number of prestellar cores resides.

Filamentary structures have been observed by different tracers. Because they are infrared-dark, filaments can be found in extinction maps at optical (Schneider & Elmegreen 1979) and infrared wavelengths (Jackson et al. 2010; Kainulainen et al. 2013; Wang et al. 2014). Far-infrared or submillimeter dust emission maps are also employed (Andre et al. 2010; Hacar et al. 2013; Wang et al. 2015; Li et al. 2016). Researches on filaments have been focused on nearby star-forming regions such as Taurus and Orion (Palmeirim et al. 2013; Takahashi et al. 2013; Kainulainen et al. 2017), or on substructures such as fibers (Hacar et al. 2018).

A variety of filaments have been found, with diverse lengths and morphology. For instance, the infrared-dark cloud Nessie (Jackson et al. 2010) was found to be as long as 430 pc in the CO map and runs along the Scutum–Centaurus Arm in position–position–velocity (PPV) space (Goodman et al. 2014). Careful studies of the individual filaments provide important hints on their properties, the role they play in the star formation, and so on. However, on one hand, individuality is inevitable. On the other hand, more distant filaments that also contain high-mass star formation have not been studied comprehensively. An unbiased sample of filaments in the Milky Way is therefore needed.

Large-scale (>10 pc) filaments may be connected to the large-scale, Galactic spiral structures. From a position–velocity analysis, Goodman et al. (2014) suggest that Nessie, with a length of hundreds of parsecs, forms a bone-like feature that closely follows the center of the Scutum–Centaurus Arm of the Milky Way. The inspection of this linkage between large-scale filaments and spiral arms is extended to other Galactic positions in the inner Galactic plane (e.g., Ragan et al. 2014; Wang et al. 2015; Zucker et al. 2015; Abreu-Vicente et al. 2016). Unlike small-scale filaments (e.g., Schisano et al. 2014, 2020; Koch & Rosolowsky 2015; Li et al. 2016; Mattern et al. 2018), of which nearby ones could be found, large-scale filaments in the Milky Way are typically far from us, so researches toward them have not been made until recent years, when modern multi-wavelength surveys that cover the Galactic plane at high resolution and sensitivity. Several powerful algorithms have been developed to identify filaments through intensity or column-density maps, such as DisPerSE, FilFinder, a local...
Hessian matrix, getFilaments, and related algorithms based on wavelet transforms (Molinari et al. 2010a; Sousbie 2011; Sousbie et al. 2011; Men’shchikov 2013; Schisano et al. 2014; Koch & Rosolowsky 2015; Ossenkopf-Okada & Stepanov 2018). The algorithms above perform well in finding parsec-scale filaments. However, large-scale filaments are rather different. One important feature is that they may not be continuous in millimeter or submillimeter continuum images. Therefore, although several catalogs of large-scale filaments have been produced in the past few years (Ragan et al. 2014; Wang et al. 2015, 2016; Abreu-Vicente et al. 2016; Zucker et al. 2018) and different search methods have different selection criteria to identify filaments, most of which are an artificial inspection of dust characteristics, except for Wang et al. (2016).

Wang et al. (2016) adopted a minimum spanning tree (MST) algorithm to automatically identify filaments by connecting velocity-coherent clumps. The MST, a construct from graph theory, is the unique set of straight lines (edges) connecting a given set of points (vortices) without closed loops, such that the sum of the edge lengths is a minimum. Two algorithms have been developed independently by Kruskal (1956) and Prim (1957). MSTs have been associated for the first time with cluster analysis by Gower & Ross (1969). In astrophysics, MSTs have so far mainly been used to analyze the large-scale distribution of galaxies or galaxy clusters, and filamentary features have been found (e.g., Barrow et al. 1985; Adami & Mazure 1999; Doroshekevich et al. 2004; Colberg 2007; Park & Lee 2009; Alpaslan et al. 2014; Naidoo et al. 2020; Pereyra et al. 2020). MSTs have also been employed to identify star clusters (e.g., Cartwright & Whitworth 2004; Schmeja & Klessen 2006; Gutermuth et al. 2009; Wu et al. 2017). In high-energy astrophysics, MSTs have been used for γ-ray source detection (e.g., Di Gesù & Sacco 1983; Campa et al. 2007, 2013). Recently, MSTs have also been introduced to quantify core separations and mass segregation (e.g., Dib & Henning 2019; Sanhueza et al. 2019). However, MSTs have also been criticized for giving unreliable cluster catalogs because they chain unrelated structure when noise is present (Getman et al. 2018). Therefore, in addition to the kinematic coherence examination carried out by Wang et al. (2016), we also performed checks on the 2D (Galactic longitude and latitude) MSTs. The clumps Wang et al. (2016) employ are from the Bolocam Galactic Plane Survey (BGPS, Rosolowsky et al. 2010), which only covers half of the Galactic plane ($7.5^\circ \lesssim l \lesssim 195^\circ$). Now, ATLASGAL gives a better estimation of distance, temperature, and velocity toward a complete sample of Galactic clumps (Urquhart et al. 2018). High-resolution column-density maps derived from the Herschel Hi-GAL survey with PPMAP (Marsh et al. 2017) also make it possible to obtain a more accurate estimation of the filament mass in the entire Galactic plane.

The physical properties of large-scale filaments associated with spiral structure have been inspected recently (e.g., Ragan et al. 2014; Wang et al. 2015, 2016; Abreu-Vicente et al. 2016; Zucker et al. 2018). However, most of the studies lack a large sample size, or use spiral arm models that do not have a high-accuracy distance estimation. Now we are able to obtain a large sample of large-scale filaments and an updated spiral arm model fitted from trigonometric parallaxes of high-mass star-forming regions (Reid et al. 2019), which assists us better in studying the relation between filaments and spiral arms. With the advance of the radial velocity measurements of ATLASGAL clumps (Urquhart et al. 2014) and the updated spiral arm model (Reid et al. 2019), we extend the work of Wang et al. (2016) in the northern Galactic plane to the entire (except Galactic center) inner Galactic plane covered by ATLASGAL.

The paper is structured as follows. We describe the data and methods we use in Section 2. Section 3 describes our results and contains physical properties of our large-scale filaments and some statistics as well as their dense-gas mass fraction. Then in Section 4, we examine the robustness of MST method and investigate the Galactic distribution of the filaments. Fragmentation of large-scale filaments is also inspected. Finally, our conclusions are summarized in Section 5.

2. Data and Method

2.1. ATLASGAL Galactic Clumps

The APEX Telescope Large Area Survey of the Galaxy (ATLASGAL) is an unbiased 870 μm submillimeter survey of the inner Galactic plane ($|l| < 60^\circ$ with $|b| < 1.5^\circ$). They identified 10,163 compact sources with a resolution of 19′ and generated a comprehensive and unbiased catalog of massive, dense clumps that are located in the inner Galaxy (Urquhart et al. 2014). Among the total 10,163 sources, ∼8000 dense clumps located away from the Galactic center region ($300^\circ < l < 355^\circ$ and $5^\circ < l < 60^\circ$) were studied in physical properties, such as velocity, by Urquhart et al. (2018). The velocity information of the clumps makes it possible to investigate their coherence in PPV space.

2.2. Herschel Column Density Map

The Herschel satellite (Pilbratt et al. 2010) provides a powerful tool for studying the structure of high-mass molecular clouds on different scales (e.g., Beuther et al. 2010; Schneider et al. 2012). Hi-GAL, the Herschel infrared Galactic Plane Survey, makes an unbiased photometric survey of the inner Galactic plane by mapping a 2° wide strip in the entire Galactic Plane in five wave bands between 70 and 500 μm (Molinari et al. 2010b). High-resolution temperature-differential column-density maps with an angular resolution of 12″ have been derived from Herschel Hi-GAL data using the PPMAP tool, which represents an improvement over those obtained with standard analysis techniques (Marsh et al. 2017). We employ the temperature-integrated column-density maps for the following mass calculation of our filaments.

2.3. Filament Identification

The filaments are identified from ATLASGAL (Urquhart et al. 2018) clumps using the MST method developed by Wang et al. (2016). The customized MST algorithm is adopted to isolate filaments in PPV space, meaning that clumps only cluster as a filament when they are close to each other with similar velocities. The clumps (circles in Figure 2(a) and filled circles in Figure 1) are connected with the cost of a minimum sum of edge lengths, where edges (rectangles in Figure 2(a) and white lines in Figure 1) mean straight lines linking clumps, and edge lengths are the separation between each pair of clumps. The criteria for MST matching and filament selection follow Wang et al. (2016):
1. The accepted MST must contain at least five ATLASGAL clumps: $N_{cl} \geq 5$.
2. Only edges shorter than a maximum length (cutoff length) can be connected ($\Delta L < 0.1$).
3. For any two clumps to be connected, the difference in line-of-sight velocity (matching velocity $\Delta v$) must be less than 2 km s$^{-1}$.
4. Linearity $f_L > 1.5$. Here linearity is defined to quantify the degree of similarity between targets with a linear shape.
5. Projected length $L_{sum} \geq 10$ pc. We only focus on large-scale filaments in this work.

The initial values of the cutoff length $\Delta L$ in item (2) and matching velocity $\Delta v$ in item (3) are chosen based on characteristics of previously known filaments. A variety of values around the initial values are tested, and the best combination is chosen to identify as many known filaments as possible, but not connect unrelated sources; see Appendix A.2 for details. $N_{cl}$ in item (1), the minimum number of clumps,
refers to the pruning level of MSTs in Park & Lee (2009) and Pereyra et al. (2020). In their work, when a branch of an MST has fewer than five nodes, it is thought to be a minor branch and should be removed from the tree. Linearity in item (4) is the ratio of the spread (standard deviation) of clumps along the major axis and the spread along the minor axis. To define the major axis of a filament, we plot all the clumps in the projected sky (Galactic longitude as x and Galactic latitude as y) and fit a line as the major axis of this filament. Instead of ordinary least squares, however, we obtain this line with the help of principle component analysis (PCA), which is detailed in Appendix B.1. The minor axis is perpendicular to the major axis and passes through the mean position \((\bar{x}, \bar{y})\). The larger \(f_L\) means that the shape is more likely to be linear. For a straight line, \(f_L \rightarrow \infty\). The distances of clumps in a filament are thought to be the same. If the distances of clumps in one filament are different, a unified distance will be used (detailed in Section 3.1). The identified filaments are shown in Figure 1 and Section 3.
2.4. Mass Calculation

The mass of the filaments is calculated by summing the mass within the boundary of each filament in the Hi-GAL-based column-density maps. The boundary of a filament is a combination of a series of circles and rectangles, as shown in Figure 2. A circle represents the position of a clump with a diameter of four times the major axis of a clump. The diameter is chosen referring to the aperture radius ATLASGAL used to measure the clump flux density from Herschel, which is twice the major axis (Urquhart et al. 2018). Our mass is derived from the Herschel column-density map. This diameter is large enough for most of the source emission to be within the boundary, while still being small enough to avoid including the background. Rectangles consist of edges linking each of the two clumps in a filament. The length of a rectangle is the length of the edge (i.e., the separation between two clumps at both ends), and the width is the same as the smaller circle diameter of two clumps at both ends. Then the mass of filament is obtained by summing the contributions from the pixels of the column-density map in the boundary. Owing to gaps in coverage in the column-density map, 29 filaments are affected (filaments with DGMF of “...” in Table 1). For a filament with an incomplete column-density map, we can only obtain a total clump mass, which is the sum of the ATLASGAL clump mass in this filament. So for these filaments, instead of integrating the column density within the boundary, we first calculate the mean ratio of the total clump mass to filament mass obtained from the column-density map of other 134 filaments with the complete column-density map. Then we scale the total clump mass for these filaments by the mean ratio.

For the dense-gas mass calculation, several approaches have been employed in previous studies. They can be classified into two categories. One category is calculating mass from a denser gas tracer within the same boundary as that of the whole filament mass (e.g., Ragan et al. 2014). The other category is to take a contour with a higher level in the extinction map as the boundary to calculate the dense gas, and the contour with the lower level as the boundary of the total mass (e.g., Zhang et al. 2019). To avoid systematic bias, we adopt the latter category. That is, our dense-gas mass is also acquired from the Herschel Hi-GAL column-density map. The dense-gas mass of a filament is derived from the sum of the clump mass in this filament derived by an integral of the column density in circle regions (shown as blue circles in Figure 2(b)) with the diameter twice the major axis of the clumps. We also tried to take the clump mass directly from the ATLASGAL catalog (Urquhart et al. 2018), where the mass is calculated with the integrated 870 μm flux density as the dense gas. We compared the clump mass from ATLASGAL and that from the Herschel column-density map with the same radius. We find that some clump masses from ATLASGAL are dramatically higher than the mass from the Herschel column-density map. Considering the consistency of the calculation method between the dense mass and the total mass, we did not use the mass from the ATLASGAL catalog. Another reason is that the column-density map obtained with PPMAP considered the temperature variance pixel by pixel, while ATLASGAL takes a single temperature from a spectral graybody fitting for each clump to calculate the mass.

Unlike employing a fixed extinction or column-density contour to define the boundaries of filaments or the dense-gas region, our boundaries vary in different Galactic longitudes. As we know, the column density in the inner Galactic plane is not the same in various Galactic longitudes, it is higher in the center and decreases outward. A fixed extinction or column-density contour, on one hand, in Galactic longitude where the column density is high might add additional mass that should not be in a filament in that filament. On the other hand, in less dense Galactic longitudes, the contour may contain nothing. Our polygon boundaries guided by filaments avoid this problem.

3. Results

We identify 163 large-scale filaments in the inner Galactic plane (|l| < 60° with |b| < 1°5) from ATLASGAL clumps, and the filaments are shown in Figure 1. The color-coded circles in Figure 1 denote clumps with different velocities. For backgrounds, cyan represents intermediate-infrared 24 μm emission on the logarithmic scale from MIPS GAL (Carey et al. 2009), and red shows submillimeter 870 μm emission on a linear scale from an APEX + Planck combined image (Csengeri et al. 2016). Other two-color views of filaments and their column-density maps are displayed in Appendix D.

3.1. Basic Physical Properties

The physical properties of the 163 filaments are shown in Table 1. The meaning of the columns is listed in Table 2. The determination of ATLASGAL clump distances can be roughly divided into four steps (Urquhart et al. 2018). First, if maser parallax and spectroscopic distances are available, the literature solution is adopted. Second, kinematic distances are determined for clumps with radial velocity. Third, clumps located within the solar circle face distance ambiguities, which are resolved using archival H1 data. Fourth, they group the clumps to assign or correct their distances. The typical uncertainty for kinematic distances of ATLASGAL clumps is 0.3 kpc. As mentioned at the end of Section 2.3, we consider that clumps in a velocity-coherent filament have the same distance. This could be regarded as a correction for a distance measurement of ATLASGAL clumps. The physical properties of the clumps that we directly take from the ATLASGAL catalog related with distance are radius and galactocentric radius. They are both linearly correlated with distance. So we could correct them by simply multiplying them by a factor $d_0/d_{c1}$, where $d_{c1}$ is the clump distance directly from the ATLASGAL catalog, and $d_0$ is the distance of the filament. In most of our filaments (95%), the clumps have the same distances. One probable reason for this is that most of the velocity-coherent structures found by the MST are physically real structures, so clumps in the same structure naturally have the same distances. A small portion of MSTs may be a collection of clumps without a physical connection. Another probable reason is that ATLASGAL applied the friends-of-friends method to group the clumps (Urquhart et al. 2018). For some of clumps (∼10%) without well-defined distances, the distance of the cluster they belong is assigned to them. For some well-studied complexes with reliable distances, their associated clumps adopted the reliable distances.

We include standard deviations of Galactic longitude, Galactic latitude, velocity, velocity gradient, and height above the Galactic midplane in Table 1. For a fractal turbulent cloud with a mass of 1000 $M_\odot$, the standard deviation of the mass derived from PPMAP is 612 $M_\odot$ (Marsh et al. 2015). The
| ID | \( l_\text{gal} \) | \( b_\text{gal} \) | \( r_\text{gal} \) | \( d \) | \( N_\text{fil} \) | \( L_\text{fil} \) | \( T_\text{fil} \) | \( T_\text{max} \) | \( T_\text{lim} \) | \( T_\text{min} \) | \( M_\text{fil} \) | \( M_\text{los} \) | \( N_{\text{H}_2} \) | \( n_{\text{H}_2} \) | \( \zeta \) | \( \sigma_\text{v} \) | \( f_1 \) | \( f_2 \) | \( L_\text{fil} \) | \( R_{\text{fil}} \) | \( \gamma \) | \( \theta \) | Mor. | DGMF |
|----|---------|---------|---------|-----|--------|-------|------|-------|------|-----|-------|-------|------|-------|-------|---|-------|-------|-----|---------|-------|---|---|
| 1  | 5.37 ± 0.02 | 0.13 ± 0.06 | 11.0 ± 1.1 | 2.9 | 6 | 12.8 | 9.3 | 0.41 ± 0.06 | 10.8 | 18.0 | 4.6 | 364.1 | 1.8 | 5.8 | 10.0 | 7.03 | 3.02 | 11.1 | 5.4 | 25.1 ± 3.3 | 82.9 ± 2.7 | L,C | 0.20 |
| 2  | 6.20 ± 0.06 | −0.12 ± 0.01 | 11.3 ± 2.3 | 3.0 | 6 | 11.2 | 8.6 | 0.91 ± 0.28 | 16.1 | 26.0 | 5.8 | 521.4 | 2.3 | 6.6 | 7.8 | 7.06 | 5.73 | 9.1 | 5.4 | 12.0 ± 0.7 | 8.0 ± 2.5 | X | 0.32 |
| 3  | 6.16 ± 0.05 | −0.61 ± 0.02 | 17.4 ± 1.2 | 3.0 | 8 | 11.8 | 7.2 | 1.02 ± 0.10 | 14.6 | 25.0 | 9.7 | 824.8 | 2.2 | 3.7 | 5.0 | 4.43 | 3.11 | 8.5 | 5.4 | −13.7 ± 1.3 | 18.8 ± 2.7 | S | 0.38 |
| 4  | 8.58 ± 0.17 | −0.33 ± 0.05 | 37.2 ± 1.2 | 4.4 | 38 | 91.1 | 51.2 | 0.41 ± 0.02 | 9.9 | 37.6 | 135.9 | 1491.3 | 3.6 | 5.6 | 34.9 | 8.07 | 5.96 | 57.3 | 4.0 | −10.2 ± 3.7 | 7.6 ± 0.4 | S.H | ... |
| 5  | 8.16 ± 0.07 | 0.22 ± 0.03 | 13.6 ± 1.0 | 10.2 | 5.0 | 13.6 | 10.2 | 0.54 ± 0.02 | 11.2 | 29.5 | 11.1 | 820.0 | 3.1 | 7.9 | 8.3 | 10.15 | 5.27 | 11.0 | 5.4 | 29.8 ± 1.6 | 23.4 ± 1.8 | L,H | 0.46 |
| 6  | 8.51 ± 0.04 | −0.64 ± 0.04 | 17.0 ± 0.6 | 3.0 | 9 | 12.8 | 8.4 | 0.49 ± 0.02 | 11.9 | 26.8 | 2.6 | 200.4 | 1.1 | 4.3 | 11.7 | 5.21 | 1.86 | 12.3 | 5.4 | −15.1 ± 0.2 | 41.3 ± 0.6 | X | 0.26 |
| 7  | 10.22 ± 0.11 | −0.27 ± 0.10 | 11.7 ± 1.3 | 3.5 | 48 | 73.8 | 27.5 | 0.62 ± 0.10 | 11.3 | 47.7 | 102.3 | 1386.5 | 3.0 | 4.2 | 25.3 | 6.61 | 2.42 | 35.7 | 4.9 | 0.9 ± 0.1 | 41.2 ± 0.5 | S,X | ... |
| 8  | 11.09 ± 0.10 | −0.01 ± 0.03 | 29.7 ± 1.4 | 2.9 | 12 | 25.0 | 16.9 | 0.65 ± 0.04 | 10.5 | 15.8 | 2.0 | 801.2 | 2.6 | 5.7 | 13.2 | 5.99 | 3.84 | 19.1 | 5.5 | 13.8 ± 1.5 | 6.7 ± 1.1 | S | 0.32 |
| 9  | 12.49 ± 0.04 | −0.21 ± 0.08 | 34.8 ± 0.9 | 2.6 | 7 | 16.8 | 14.1 | 0.24 ± 0.01 | 10.0 | 18.5 | 8.3 | 496.1 | 2.2 | 6.7 | 12.2 | 9.75 | 5.04 | 15.2 | 5.9 | 9.9 ± 0.1 | 65.3 ± 1.5 | S | 0.20 |
| 10 | 12.87 ± 0.16 | −0.21 ± 0.08 | 35.5 ± 1.5 | 2.6 | 11 | 26.7 | 26.8 | 0.86 ± 0.05 | 9.6 | 26.6 | 8.66 | 1328.3 | 4.5 | 9.8 | 32.8 | 6.21 | 2.36 | 35.2 | 5.9 | 9.9 ± 3.8 | 17.8 ± 0.5 | C,H | - |

Note. Meaning of columns is detailed in Table 2. Column (1) ID of large-scale filaments. Columns (2)–(4): flux-weighted longitude, latitude (degrees), and LSR velocity (km s\(^{-1}\)). Column (5): distance (kiloparsecs). Column (6): number of clumps in filaments. Column (7): sum of the edge lengths (parsecs). Column (8): end-to-end length (parsecs). Column (9): velocity gradient (km s\(^{-1}\) pc\(^{-1}\)). Columns (10) and (11): minimum and maximum temperatures (kelvin) of clumps. Column (12): mass of the filament (10\(^{14}\)M\(_\odot\)). Column (13): line mass (M\(_\odot\) pc\(^{-1}\)). Columns (14) and (15): \(H_2\) column density (10\(^{22}\) cm\(^{-2}\)) and volume density (10\(^{23}\) cm\(^{-3}\)). Column (16): sum of the edge lengths over the filament width. Column (17): aspect ratio. Column (18): linearity. Column (19): linearity-weighted length. Column (20): galactocentric radius (kiloparsecs). Column (21): height above the Galactic midplane (parsecs). Column (22): orientation angle between the filament major axes and the Galactic midplane in the projected sky (degrees). Column (23): morphology class. Column (24): dense-gas mass fraction. "..." means no DGMF measurement. (This table is available in its entirety in machine-readable form.)
The normal distribution has positive skewness, which indicates that the tail is on the right. Kurtosis means that the tail is on the left side of the distribution, and skewness measures the asymmetry of probability distribution. For instance, the normal distribution has skewness 

\[ \mu, \sigma, \kappa \]

of clumps, and then obtain one filament. We repeat this process 1000 times and take the standard deviation of the 1000 values of \( \theta \) as the uncertainty of \( \theta \).

The statistics of the parameters are shown in the last few rows of Table 1, including minimum, maximum, median, mean, standard deviation, skewness (\( S \)), and kurtosis (\( K \)). Skewness measures the asymmetry of probability distribution. For instance, normal distribution has \( S = 0 \). Negative skewness means that the tail is on the left side of the distribution, and positive skewness indicates that the tail is on the right. Kurtosis characterizes how the distribution is compared to a normal distribution. The normal distribution has \( K = 0 \). A distribution with \( K > 0 \) is more centrally peaked than normal distribution while for \( K < 0 \) is flatter.

### 3.2. Statistics

Some representative physical properties of the filaments are shown in Figure 3. Panels (a)–(h) are the normalized distribution of the Galactic longitude, the height to the Galactic midplane, the Galactocentric radius, the distance, and the radial velocity for the filaments. Panels (g)–(j) are the mass, the linearity-weighted length, the line mass, and the orientation angle of the filaments.

No filaments are found within \( |l| < 5^\circ \) due to the limitation of the catalog we use. The distribution of the filaments with Galactic longitude has two peaks around \( l = -22^\circ \) and \( l = 23^\circ \) that are shown in Figure 3(a). Mattern et al. (2018) also find the first peak, but not the second peak in their small-scale ATLASGAL filaments, because their Galactic longitude ranged from \( l = -60^\circ \) to \( l = 18^\circ \). They attribute the decrease in filament count toward the center to the difficulty of identifying filaments in crowded structures. They think the outward suppression means that it is unlikely to find filaments toward the outer Galaxy. In our sample, the position of the filaments shows a similarity with the positions of the ATLASGAL filaments, because their Galactic longitude ranged from \( l = -60^\circ \) to \( l = 18^\circ \). They attribute the decrease in filament count toward the center to the difficulty of identifying filaments in crowded structures. They think the outward suppression means that it is unlikely to find filaments toward the outer Galaxy. In our sample, the position of the filaments shows a similarity with the positions of the ATLASGAL filaments, implying that filaments tend to be found in the region in which clumps are crowded, and this can be directly perceived in Figure 6(c). So the peaks in the distribution of the filaments with Galactic longitude are merely due to peaks in the distribution of the ATLASGAL clumps. This result also implies that large-scale filaments are ubiquitous in the inner Galactic plane and are therefore an unavoidable issue in the study of star formation. But we also note that this similarity in the distribution of the filaments and clumps does not always exist (e.g., the galactocentric radius in Figure 3(d)). The distribution of the Galactic latitude of the filaments shown in Figure 3(b) has a peak and mean \( b = -0^\circ 01 \) close to the Galactic midplane, which is consistent with Mattern et al. (2018).

Of our filaments, F120 stands out, whose linearity-weighted length is 121 pc, and the mass is \( 4.36 \times 10^5 M_\odot \). Except for this extreme filament, the linearity-weighted lengths of the filaments range from 7–74 pc, with a mass of about \( 10^3–10^5 M_\odot \). We also find no obvious correlation between length and distance. The vertical distances from the filaments to the Galactic midplane (Figure 3(c)) are not symmetrical about the Galactic midplane (\( S = -0.6 \)), which agrees with the result for the northern sky (Wang et al. 2016).

### Table 2

| Column     | Explanation                                                                 |
|------------|-----------------------------------------------------------------------------|
| (1) ID     | Assigned filament ID                                                        |
| (2) \( l_{\text{ext}} \) (deg) | Flux-weighted Galactic longitude                                             |
| (3) \( b_{\text{ext}} \) (deg) | Flux-weighted Galactic latitude                                              |
| (4) \( v_{\text{ext}} \) (km s\(^{-1}\)) | Flux-weighted local standard of rest (LSR) velocity                           |
| (5) \( d \) (kpc) | Distance, the median of all clump distances in this filament                 |
| (6) \( N_c \) | The number of clumps in the filament                                          |
| (7) \( l_{\text{sum}} \) (pc) | The sum of the edges (separation of two connected clumps) of this filament  |
| (8) \( l_{\text{end}} \) (pc) | End-to-end length along the major axis                                        |
| (9) \( v_{\text{grad}} \) (km s\(^{-1}\) pc\(^{-1}\)) | Mean velocity gradient at the edges                                          |
| (10) \( T_{\text{min}} \) (K) | Minimum temperature of the clumps                                            |
| (11) \( T_{\text{max}} \) (K) | Maximum temperature of the clumps                                            |
| (12) Mass (\( 10^3 M_\odot \)) | Mass of the filament, derived from the Herschel Hi-GAL column-density map (detailed in Section 2.4) |
| (13) \( M_{\text{line}} \) (\( 10^2 M_\odot \) pc\(^{-1}\)) | Line mass, mass per unit length                                               |
| (14) \( N_{\text{H}_2} \) (\( 10^{22} \) cm\(^{-2}\)) | \( \text{H}_2 \) column density                                              |
| (15) \( n_{\text{H}_2} \) (\( 10^3 \) cm\(^{-3}\)) | \( \text{H}_2 \) volume density                                              |
| (16) \( \frac{\text{sum}}{\text{end}} \) | The ratio of the sum of the edge lengths and width. The width is the mean of the clump diameters |
| (17) \( f_k \) | Aspect ratio, the ratio of the area between the circle enclosing the filament and the concave hull (detailed in Appendix B.2). For an approximately elliptical distribution, \( f_k \) is very similar to the aspect ratio of the ellipse |
| (18) \( f_L \) | Linearity, the ratio of the spread (standard deviation) of the clumps along the major axis and the spread along the minor axis |
| (19) \( L_{\text{ext}} \) (pc) | Linearity-weighted length, \( L_{\text{ext}} = L_{\text{end}} \cdot \sqrt{\frac{1 + 4 \kappa^2}{\kappa^2}} \) |
| (20) \( R_{\text{gc}} \) (kpc) | Galactocentric radius, median of the clump galactocentric radius              |
| (21) \( z \) (pc) | Height above the Galactic midplane                                           |
| (22) \( \theta \) (deg) | Orientation angle between the filament major axes and the Galactic midplane in the projected sky (the calculation of \( \theta \) is from a PCA, detailed in Appendix B.1) |
| (23) Mor. | Morphology class following Wang et al. (2015), L represents the linear straight L-shape, C is the bent C-shape, S is the quasi-sinusoidal shape, X is the crossing of multiple filaments, and H is the head-tail or hub-filament system |
| (24) DGMF | Dense-gas mass fraction, the ratio of the dense-gas mass and the total mass of the filament. It is illustrated in Section 3.3 |
The relation between mass and lengths (sum of the edge lengths) of the large-scale filaments is shown in Figure 4. The black curve is the best-fit line using the least-squares method, whose expression is \( \log L = 0.46 \log M - 0.40 \). This implies a tight power-law relation between length and mass, \( M \sim L^{2.17} \). The power-law index is similar to that of the filaments from BGPS sources (Wang et al. 2016), indicating a fractal dimension of 2.17. For comparison, the relation between the mean cylinder radii and the mass of large-scale filaments is also examined (Figure 4(b)), which gives a relation of \( \log L = 0.56 \log M - 2.14 \). The corresponding dimension is 1.77, beyond the range 2–3 from Roman-Duval et al. (2010), implying that the mean cylinder radii are a less essential property of large-scale filaments.

Do filaments in the proximity of the Galactic midplane prefer to align with the Galactic midplane? The relation between orientation angle and height to the Galactic midplane of the large-scale filaments is plotted in Figure 5. There is no obvious
correlation between the two, which is identical to the result for the northern sky (Wang et al. 2016).

3.3. Dense-gas Mass Fraction

The dense-gas mass fraction (DGMF) of a filament is the ratio of the dense-gas mass to the total mass of the whole filament, which is an important quantity related to the star formation rate (SFR) and star formation efficiency (SFE) of molecular clouds (Heiderman et al. 2010; Lada et al. 2010, 2012), although Kainulainen et al. (2013) suggest that DGMF and SFE may not be closely linked. To avoid deviation engendered by distinct methods to derive the mass, both dense-gas mass and filament mass are derived from the Herschel Hi-GAL column-density map, which is illustrated in Section 2.2. After discarding filaments with incomplete column-density information, 130 of the 163 large-scale filaments have a DGMF measurement. Their DGMFs range from 14.7% to 62.4%, with a mean value of 35.6%. The result is larger than that from Ragan et al. (2014) and Abreu-Vicente et al. (2016), who take the ratio of the mass from ATLASGAL 870 μm dust emission and the mass from 13CO emission as DGMF. Our result is consistent with a value of 50% from Nessie (Goodman et al. 2014), however, where the DGMF is from the ratio of the mass in an envelope traced by HNC observation to the mass in cylinders with a fixed diameter and above a column-density threshold. We conjecture that diverse definitions of the dense-gas mass account for the difference. We also employ a similar dense-gas mass calculation as Abreu-Vicente et al. (2016), where the dense-gas mass is acquired from the ATLASGAL flux. The DGMFs become smaller.
4. Discussion

4.1. Comparison of Our MST Filaments to MST Filaments from BGPS Sources and Other Previously Known Filaments

To see whether the MST method is robust on sources from different catalogs, we compare our filaments from ATLASGAL Galactic clumps to those identified from BGPS by Wang et al. (2016) in the common longitude range in which the two surveys reside (7°5 < l < 60°). In the common region, 42 filaments are identified from BGPS sources, while 67 are from ATLASGAL. The detailed comparison is presented in Appendix A.1. On the whole, in the common Galactic region of the two catalogs (7°5 < l < 60°), most (70%) of the filaments identified from BGPS sources are also found from ATLASGAL clumps, despite different methods they use to extract sources, the distinct lower limit of the source luminosities, and the various surveys referred to obtain the radial velocities of the sources (Shirley et al. 2013; Urquhart et al. 2018).

The distinction between filaments from the two catalogs mainly arises from the number of sources used to identify filaments. BGPS sources contain over 8400 continuum sources, but only 3126 of them have a velocity measurement from HCO+(3 − 2) and/or N2H+(3 − 2) spectral. Since we aim to identify velocity-coherent filaments, the velocity is indispensable for the MST method. For ATLASGAL sources, 7809 of them have radial velocities obtained from 21 archival molecular line surveys (Urquhart et al. 2018). In the overlaid longitude we compare (7°5 < l < 60°), 2201 BGPS sources have a velocity measurement, while this holds for 3379 ATLASGAL sources. Therefore, it is natural that the number of filaments identified from ATLASGAL sources is larger than that from BGPS sources. More sources with velocity information also inform us that some of the filaments identified from MST may actually be part of a larger structure (e.g., BGPS filaments F2 and F3 are the eastern and western part of ATLASGAL filament F4, respectively). The missing velocity information of a portion of sources may lead to a failure to identify some filaments.

Our MST method identified some previously known large-scale filaments in the southern sky. Of the three bone candidates in the southern sky from Zucker et al. (2015), filaments BC_355.31-0.29 and BC_332.21-0.04 are found by the MST. Filament BC_355.31-0.29 is our F117. Filament BC_332.21-0.04 is a complex consisting of F128, F127, and the southern part of F126. Filament G350.54+0.69 found by Liu et al. (2018) is a part of our F80. Filament G316.75-0.1 (Watkins et al. 2019) spatially overlaps our F156.

4.2. Large-scale Filaments in the Milky Way

To examine the Galactic distribution of large-scale filaments, we exhibit the locations of the filaments in the longitude–velocity (PV) space and draw the Galactic spiral arms as reference (Figure 6(a)). The spiral arm model is from a fitting result of Galactic high-mass star-forming regions with trigonometric parallaxes (Reid et al. 2019). According to the previous criterion with which we judged whether a large-scale filament is associated with spiral arms (Ragan et al. 2014; Abreu-Vicente et al. 2016), if a filament has an LSR velocity within ±5 km s⁻¹ (Wang et al. 2016) of a spiral arm in the same Galactic longitude, it is thought to be an arm filament.

Figure 6. (a) Longitude–velocity view of filaments and spiral arms. Spiral arms from Reid et al. (2019) are plotted as belts in a variety of colors. Our data do not support a grand design spiral pattern, but neither are they opposed to this pattern. We just simply show this optimistic presumption of a spiral structure. Gray arms are arm segments that may join together to form arms. Filaments are shown as black circles, and if a filament is a bone, a red cross is overlaid. Gray rectangles at the bottom mark the longitude range of our filaments. (b) Longitude–distance view of filaments and spiral arms. If a bone is also within an arm in this longitude–distance diagram, a blue plus is overlaid. (c) Face-on view of filaments and spiral arms. Spiral arms are shown with their kink widths (taken from Table 2 of Reid et al. 2019), which can be regarded as characteristic widths of arms. Red dots in the background mark the clumps used to identify filaments from Urquhart et al. (2018).
This criterion is illustrated in Figure 6(a). Belts show spiral arms in PV space with widths of 10 km s\(^{-1}\). If a filament is located in any of the belts, it is thought to be an arm filament. Of our 163 large-scale filaments, 87 (53\%) are in spiral arms or spurs according to this criterion. But this value is only for comparison to previous work. We will update the criteria with which we judged which filaments could be thought of as arm filaments later. Ragan et al. (2014) find that most of their filaments are interarm filaments, while Abreu-Vicente et al. (2016) and Wang et al. (2016) find that the arm filament percentages are 67\% and 80\%, respectively.

Some filaments lie in the center of the spiral arms and so sketch out the bones of the Milky Way (Goodman et al. 2014; Zucker et al. 2015). Bones are also found in our large-scale filaments if we add the following three additional criteria (Wang et al. 2016):

1. Lie very close to the Galactic midplane, \(|z| \leq 20 \text{ pc}\);
2. Run roughly parallel to the arms in the projected sky, \(\theta \leq 30\);
3. Flux-weighted LSR velocity is within \(\pm 5 \text{ km s}^{-1}\) of a spiral arm in the same Galactic longitude.

Their Galactic distribution is shown as crosses in Figure 6, and the physical properties are denoted as red bars in Figure 3. Figure 6(c) shows filaments and clumps (filled red circles) overlaid on Galactic spiral arms as viewed from the Northern Galactic Pole. The exhibited widths of the spiral arms are intrinsic (Gaussian 1\(\sigma\)) arm widths at the Galactic radius of the kinks (Reid et al. 2019). There are 23, 12, 20, and 6 large-scale filaments in the Norma-Outer, Scutum–Centaurus-OSC, Sagittarius–Carina, and Perseus arm, respectively, while for the bones, these values are 10, 8, 12, and 5, respectively (summarized in Table 3). Here, to judge whether a filament is associated with an arm, we still follow the criterion from previous work. The differences between arm filaments and bones are that an arm filament need only to satisfy criterion (8), while bones should fulfill not only criterion (8), but also criteria (6) and (7). Surprisingly, most filaments in the Scutum–Centaurus-OSC arm and Perseus arm are bones (8/12 and 5/6), and these two arms have been thought to be two dominant spiral arms (Drimmel 2000; Churchwell et al. 2009). However, we note that the bone fractions in the four arms have no significant differences, and we lack a statistic sample.

From the face-on view of the filaments in spiral arms, we note that some bones that were thought to be arm filaments are obviously not in any spiral arm or spur (see the red crosses between the arms in Figure 6(c)). That is because in the previous criteria that were used to judge whether a filament is in a spiral arm, we only require that the filament is velocity coherent with a spiral arm in the same Galactic longitude. However, the distance of the filament to that arm was not considered. So this is a 2D position–velocity (PV) match. Researches of the Galactic location of filaments have also been confined to a 2D match PV (Ragan et al. 2014; Abreu-Vicente et al. 2016; Wang et al. 2016) or to a position–position match (PP; Wang et al. 2015; Matern et al. 2018). Now that we obtain better a distance estimation of both filaments and spiral arms, a PPV match becomes possible.

Our new definition of a bone adds a distance constraint compared with the previous definition when judging whether a filament is in a spiral arm. So it becomes a PPV match. This constraint is another criterion:

1. The difference between its distance to us and the distance of its PV-related spiral arm to us is less than 1 kpc in the same Galactic longitude.

That is, for a bone that met the previous definition criteria (6), (7), and (8), we further examine whether it is close enough to its velocity-associated arm. If a previous bone also meets criterion (9), it will pass our distance constraint and be thought of as a bone according to the new definition. Otherwise, it will be excluded from bones, because in this situation, the filament may just look near an arm on the line of sight. The 1 kpc tolerance is from a combination of the uncertainties of the filament distances and 1\(\sigma\) widths of arms. Distances from the Sun versus Galactic longitudes of filaments and spiral arms are plotted in Figure 6(b). Bones satisfying our new definition are denoted with blue pluses. As we can see, previous bones (denoted as red crosses) far from spiral arms are eliminated in our new definition. This is also clear in the face-on map. However, some of the previous bones that seem to be in one arm in Figure 6(b) are also excluded. This situation occurs when a filament has a similar velocity as a spiral arm at a Galactic longitude, but far from this arm and near another arm in distance.

The number of filaments in the spiral arms decreases significantly after adding the distance constraint. Unlike for bones, when a filament matches a spiral arm in PPV space (at a certain Galactic longitude, they have a similar distance from the Sun and a similar velocity), it is thought to be an arm filament. Of the 163 large-scale filaments, 138 are not in any spiral arm or spur, but 8, 1, 4, and 1 are in Norma-Outer, Scutum–
Centaurus-OSC, Sagittarius–Carina, and the Perseus arm (satisfying criteria (8) and (9)), respectively. The values for bones (satisfying criteria (6) ~ (9)) are 5, 1, 3, and 1. The fact that 85% of the large-scale filaments are not in any spiral arm or structure seems to indicate that the filament may not be associated so tightly with the spiral arms as thought before. But we also note that due to the existence of the distance ambiguity of filaments and the uncertainty in the spiral arm distance, especially in the fourth quadrant, some filaments that should be in spiral arms may be ruled out by the distance constraint, resulting in an underestimation of the arm filaments count.

If we do not care whether large-scale filaments are bones of the Milky Way and just wish to know the influence of spiral arms on filaments, we could take the filaments meeting criteria (1) ~ (5), (8), and (9) as arm filaments. The others are interarm filaments. Between the two, we find no significant distinctions in some of their physical properties such as mean temperatures, nonthermal velocity dispersion, and column densities. But they also have differences in properties such as Galactocentric radius, height above the Galactic midplane, mass, and length. The detailed discussion for arm filaments and interarm filaments are given in Appendix C.

4.3 Dense-gas Mass Fraction

The dense-gas mass fraction is found to be associated with the Galactic position of the filaments. From the analysis of seven filaments identified in the Galactic Ring Survey (GRS; Jackson et al. 2006), Ragan et al. (2014) suggested that the DGMF decreases as the galactocentric radius increases. They also found that filaments near the Galactic midplane tend to have a larger DGMF. In contrast, Abreu-Vicente et al. (2016) find no obvious correlation between height above the Galactic midplane and DGMF from an inspection of the 10 filaments they identified in the fourth Galactic quadrant.

To investigate this issue in a larger statistical sample, we examined the DGMF and Galactic position of our 130 filaments with DGMF measurements. The results are shown in Figure 7. Surprisingly, the DGMF is neither related to galactocentric radius nor to the height above the Galactic midplane. We changed the size of the boundaries to derive the filament mass and dense-gas mass, and obtained the similar result. We also tried to use the clump mass from ATLASGAL as the dense-gas mass, and the DGMF was still not correlated with galactocentric radius or with the height above the Galactic midplane.

Figure 7. Dense gas mass fraction and Galactic position of the filaments.

4.4 Dense Clumps in Filaments

Dense clumps are thought to be birthplaces of massive stars (e.g., Motte et al. 2018). Their typical mass is $10^3 M_\odot$, and the typical molecular hydrogen column density is $10^{22}$ cm$^{-2}$ with a typical radius of 1 pc. According to Wang et al. (2016), dense clumps in large-scale filaments are slightly denser than those outside of filaments, indicating that filaments prefer to gather material and assist the formation of massive stars. To examine whether this preference can be extended to the whole inner Galactic plane, we compare the mass of clumps with various effective radii between those in and outside of filaments. The differences between clumps in and outside of filaments are shown in Figure 8(a), where the lines are the fit results. Surprisingly, the two have no remarkable distinctions. We also bin the logarithmic radius for greater clarity (Figure 8(c)). The bins range from about −1.6 to 1.2 with a step of 0.1. In each bin, we calculate the mean logarithmic mass of the clumps in and outside of filaments, and then plot them versus the median logarithmic radius of each bin. The black curve shows clumps in filaments, and the red curve shows clumps outside of filaments. Error bars denote the standard deviations of the logarithmic clump mass in each bin. As we can see, the two curves almost coincide, except at both ends, where the number of clumps is small. But even at both ends, the differences are within the error bars. Therefore, the mass of clumps in filaments or outside of filaments has no significant distinction on the same scale.

We then make a histogram to show the molecular hydrogen column density of the clumps in filaments or outside of filaments (Figure 8(b)). The densest clump has a column density of about $25 \times 10^{22}$ cm$^{-2}$, but the number of extremely dense clumps is small. So we only show the density ranging from 0 to $6 \times 10^{22}$ cm$^{-2}$. The medians of the molecular hydrogen column density for 2628 clumps in large-scale filaments and 4213 outside of filaments are $1.03 \times 10^{22}$ cm$^{-2}$ and $1.28 \times 10^{22}$ cm$^{-2}$, respectively. Therefore, for our filaments, clumps outside of filaments are slightly denser than those in filaments. This result may explain why the SFE and SFR surface density in the filaments are similar to those in other star-forming regions discovered by Zhang et al. (2019). To figure out whether this result is caused by not demanding enough for filaments, we increase the critical linearity for filaments and repeated the same experiment. We also use the bones to examine this issue. However, the results do not change much. In some cases, filamentary structures have been thought to fragment through cylinder fragmentation.
Clumps from cylinder fragmentation will be denser than those from Jeans fragmentation (Wang et al. 2014). Therefore, our results give us a hint that the fragmentation process of filamentary structures may not be cylinder fragmentation.

To figure out why we find no distinction between clumps on and outside of filaments, while Wang et al. (2016) obtained different results, we restore Figure 5(b) of Wang et al. (2016) and overlay the ATLASGAL clumps in the northern sky in our Figure 8(d). We only plot ATLASGAL clumps in the northern sky in this panel because the BGPS clumps we compare are almost in the northern sky. Brown triangles are ATLASGAL clumps in the northern sky. Filled red circles are 496 clumps in the BGPS filaments in Wang et al. (2016). Blue crosses show the BGPS clumps from Ellsworth-Bowers et al. (2015). The lines are the fit results.

4.5. Fragmentation of Large-scale Filaments

Linear filaments were approximately regarded as gas cylinders when their fragmentation was investigated (e.g.,
Fischera & Martin 2012; Wang et al. 2014, 2016). According to Chandrasekhar & Fermi (1953) and Ostriker (1964), an isothermal gas cylinder will fragment due to gravitational instability when the mass per unit length exceeds a critical line mass. Under this condition, the fragments will be equally spaced with a separation $\lambda_{cl}$, and the fragment mass will be

$$M_{cl} = 301.7M_\odot \left( \frac{\lambda_{cl}}{\text{pc}} \right)^3.$$  \hspace{1cm} (1)

Alternatively, if the fragmentation is governed by the Jeans instability, the relation between fragment mass and separation is different. For a thermal Jeans fragmentation, the fragment mass is

$$M_{\text{th}} = 13.29M_\odot \left( \frac{T}{10 \text{ K}} \right) \left( \frac{\lambda_{\text{th}}}{\text{pc}} \right).$$ \hspace{1cm} (2)

While for turbulent fragmentation, the fragment mass is

$$M_{\text{turb}} = 381.7M_\odot \left( \frac{\sigma}{\text{km s}^{-1}} \right)^2 \left( \frac{\lambda_{\text{turb}}}{\text{pc}} \right).$$ \hspace{1cm} (3)

where $\sigma$ is the turbulent width, which is well approximated by the velocity dispersion measured from dense-gas tracers such as NH$_3$. The above equations are deduced from Wang et al. (2014). As can be seen, the power-law index for the relation between fragment mass and separation is 3 for cylinder fragmentation and 1 for Jeans fragmentation. So we show the mean clump mass versus mean clump separation in each filament in Figure 9(a). The slope of the fitted line is 1.46, representing the power-law index between the mean clump mass and the mean separation. If we take the projection effect into consideration, the slope will be even shallower. This power-law index indicates that large-scale filaments are more likely to have Jeans fragmentation than cylinder fragmentation, which is in contrast to Wang et al. (2016), who suggest a cylinder fragmentation. We note that our sample is larger (163 filaments) than that of Wang et al. (2016, 54 filaments). However, it is too early to draw a conclusion because Equation (1) for cylinder fragmentation is based on an isothermal linear gas cylinder with infinite length. In practice, a filament could not be strictly isothermal, and it also has a finite length with some bends.

To inspect whether straight filaments would prefer to have cylinder fragmentation, we derive the power-law index for filaments with a linearity larger than a set of values. For instance, the mean clump mass versus mean separation of filaments with a linearity larger than 3 are plotted. We fit a line for mass and separation, and take the slope of this line as a power-law index for this linearity. For filaments with a linearity larger than 5, we repeat the procedure and obtain another power-law index. Similarly, we choose a critical linearity from 1.5 to 8.5 and obtain a series of slopes (power-law index). Then we plot the power-law index versus linearity in Figure 9(b). As can be seen, the power-law index has a rising trend in linearity between 4 and 6. This result gives us a hint that linearity might play an important role in the fragmentation mechanism of filaments. The larger the linearity, the more likely is cylinder fragmentation. For linearity larger than 6, the trend of the curve is uncertain due to the large error caused by the small sample size.

5. Conclusion

We employ the MST method (Wang et al. 2016) to search for large-scale filaments in PPV space in the inner Galactic plane. The algorithm is applied to 7809 clumps with a velocity measurement in the ATLASGAL Galactic clumps catalog in the range $|l| < 60^\circ$ with $|b| < 1^\circ.5$ (Urquhart et al. 2018). We produce a sample of large-scale filaments in the inner Galactic plane consisting of 163 filaments. We derive their physical properties and examine their DGMFs with the help of a Herschel Hi-GAL column-density map derived from PPMAP (Marsh et al. 2017). We inspect the Galactic distribution of the filaments and compare them with the latest spiral arm model (Reid et al. 2019). Fragmentation of large-scale filaments is also investigated. The main results are listed below.
1. The robustness of the MST method is verified by a comparison between the results from ATLASGAL Galactic clumps and BGPS dust clumps. Most (70%) of the filaments identified from BGPS dust clumps are found from ATLASGAL clumps in the common Galactic region of the two catalogs, although different methods were used to extract them and although a distinct lower limit of the clump luminosities and various surveys was referred to to obtain the radial velocities of the clumps.

2. The Galactic positions of the filaments are asymmetric about the Galactic midplane, which may be an observational bias.

3. The dense-gas mass fractions of the filaments have no significant distinctions in different Galactic radii and vertical height above the Galactic midplane, in contrast to previous studies.

4. The filaments are compared with an updated spiral arm model, and a new PPV match is employed to judge whether a filament is in a spiral arm. With this matching method, a number of filaments that are thought to be in an arm from a PP or PV match are eliminated. So the number of filaments that is actually associated with arms decreases greatly.

5. The bone fraction and dense-gas mass fraction do not vary too much in different spiral arms.

6. Dense clumps in filaments have no obvious distinction in mass compared with those outside of filaments on the same scale.

We are grateful to Chao Wang, Wenyu Jiao, K. A. Marsh, Nannan Yue, Siju Zhang, Fengwei Xu, and the anonymous referee for helpful discussions and suggestions that improved the scientific content as well as clarity of the manuscript. We acknowledge support by the National Key Research and Development Program of China (2017YFA0402702), the National Science Foundation of China (12041305, 11973013, 11721303), and the Weiming-I High-Performance Computing Platform of Peking University through the instrumental analysis fund of Peking University (0000057511). We acknowledge the science research grants from the China Manned Space Project with Nos. CMS-CSST-2021-B06 and CMS-CSST-2021-A09. This research has made use of SAOImageDS9, developed by the Smithsonian Astrophysical Observatory. This research made use of Montage. It is funded by the National Science Foundation under grant No. ACI-1440620, and was previously funded by the National Science Foundation under grant No. ACI-2017YFA0402702.

Appendix A
Robustness of the Algorithm

A.1. Comparison between Filaments from BGPS and ATLASGAL Sources

To figure out whether the MST method is sensitive to data sets, we compare filaments identified from BGPS and ATLASGAL Galactic clumps. The result is summarized in Table A1. The common part of the two Galactic clump catalogs reaches from Galactic longitude 7°.5 to 60°, corresponding to 42 filaments (F1 to F42) found from BGPS sources by Wang et al. (2016) and 67 filaments identified from ATLASGAL sources by us. We list the 42 filaments from BGPS sources and scrutinize whether they are found in ATLASGAL sources. In Table A1, if a filament from ATLASGAL is similar to a filament from BGPS, the sequence number of that ATLASGAL filament will be listed below the ID of the BGPS filament. If a BGPS filament is a part of an ATLASGAL filament, “E,” “W,” “N,” or “S” will be added to the sequence number of the ATLASGAL filament for the east, west, north, or south part. On the other hand, if an ATLASGAL filament is a part of a BGPS filament, “P” will be added before the sequence number of the ATLASGAL filament. If both a BGPS filament and an ATLASGAL filament are at the same position but look different, a “D” will be added. If a BGPS filament is not found in ATLASGAL sources, an “NF” will be filled below.

A.2. Choice of Parameters in MST

The choice of parameters starts from a physical consideration. Since the scale we are interested in is tens of parsec and the observed angle for a filament with a length of 10 pc at a distance of 5 kpc is about 0°.1, we choose 0°.1 as the initial cutoff length. The velocity difference between the two ends of that filament has an order of 1 km s$^{-1}$ if we treat the velocity gradient as 0.1 km s$^{-1}$ pc$^{-1}$. This order of velocity gradient is estimated from the velocity difference divided by the total length of the filaments from several large-scale filament catalogs (Ragan et al. 2014; Wang et al. 2015; Zucker et al. 2015; Abreu-Vicente et al. 2016).

After initial cutoff length and matching velocity are chosen, we test other values around them. For the cutoff length, we give 0°.05, 0°.10, 0°.20, and 0°.40. Meanwhile, 1, 2, 3, and 5 km s$^{-1}$ are given to the matching velocity. For illustration, tests in a small part of the sky are shown in Figure A1. Figure A1 exhibits MSTs with different parameters in the region around the Snake nebula, one of the first identified IRDCs (Carey et al. 1998) located at a Galactic longitude 11°.11 and latitude −0°.12. The panels in each row show the MSTs obtained with the same matching velocity, and the cutoff lengths range from 0°.05 to 0°.40. Different rows of panels have different matching velocities, ranging from 1 to 5 km s$^{-1}$. Dots are ATLASGAL clumps color-coded by LSR velocities. Black line segments are the edges of the MSTs. As we can see, the Snake, the MST at Galactic longitude ∼11°, has been successfully found with cutoff lengths 0°.10, 0°.20, and 0°.40. When the cutoff length is 0°.05, the Snake cannot be found. When the cutoff length is set to 0°.20 or 0°.40, too many unrelated structures are linked. So the best cutoff length for the test of the Snake is 0°.10. Since MSTs do not vary much when we alter the matching velocity in this region, tests in other parts of the sky show that a very strict matching velocity will cause a loss of real structures. Results obtained with a matching velocity of 2, 3, and 5 km s$^{-1}$ do not vary too much. This is reasonable considering the possibility that a physically isolated source is observed exactly within a structure. Then we choose the relatively strict matching velocity, 2 km s$^{-1}$.

Since the results of the MST are relatively sensitive to cutoff length, we apply a widely used approach in star clusters (e.g., Gutermuth et al. 2009) to examine the chosen cutoff length. First, we set the cutoff length to maximum and cluster all clumps in a single MST. Next we plot the cumulative distribution function (CDF) of the MST edge lengths in Figure A2 as brown steps. As we can see, the CDF is steep in...
Figure A1. MSTs with different parameters in the region around the Snake. Panels in each row show MSTs obtained with the same matching velocity, and the cutoff lengths range from 0°.05 to 0°.40. Different rows of panels have different matching velocities, ranging from 1 to 5 km s\(^{-1}\). Dots are ATLASGAL clumps color-coded by LSR velocity. Black line segments are MST edges.

Table A1

Comparison between Filaments from BGPS and ATLASGAL Sources

| BGPS Filaments | F1  | F2  | F3  | F4  | F5  | F6  | F7  | F8  | F9  | F10 | F11 | F12 | F13 | F14 |
|---------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| ATLASGAL filaments | NF  | F4-E | F4-W | F5  | F7  | NF  | F8  | NF  | F9  | F10 | P-F13 | D-F11 | D-F17 | F16 |
| BGPS filaments | F15 | F16 | F17 | F18 | F19 | F20 | F21 | F22 | F23 | F24 | F25 | F26 | F27 | F28 |
| ATLASGAL filaments | D-F15 | NF  | NF  | F18 | D-F26 | F29 | F33 | F35-N | F35-S | NF  | D-F40 | NF  | F41 | NF  |
| BGPS filaments | F29 | F30 | F31 | F32 | F33 | F34 | F35 | F36 | F37 | F38 | F39 | F40 | F41 | F42 |
| ATLASGAL filaments | F44 | NF  | NF  | P-F54 | F56 | F57 | F58 | F59 | NF  | F62 | NF  | NF  | F65 | P-F66 |
the short edge part and shallow in the long edge part. So the CDF can be approximated by three parts: a steep-sloped segment, a shallow-sloped segment, and a transition segment. We fit a line for the steep-sloped segment and shallow-sloped segment, respectively. Then the intersection of the two lines gives us the cutoff length. The cutoff length acquired from this approach is 0.95°, which is very close to our criteria. If we alter the fitting range with reason, the result changes by less than 10%.

A.3. Comparison of the Methodological Approach

We take the well-studied filament of the Snake nebula as an example. We employ FilFinder (Koch & Rosolowsky 2015) to find filamentary structures in this region through the column density. In Figure A3, the red curves in panel (a) are the skeleton found by FilFinder. The background is the flattened (a preprocess for the map of FilFinder that suppresses objects that are significantly brighter) column-density map. Panel (b) is the filament found by the MST as a comparison. As we can see, FilFinder is powerful in finding minor structures. But if we would like to extract large-scale structure, by-eye inspection has to be used on the result from FilFinder.

Appendix B

Additional Remarks on Methods to Obtain the Physical Properties

B.1. Orientation Angles and Major Axes of Filaments

The orientation angle is the angle between the filament major axis and the Galactic midplane. That is, it is the arctan of the slope of the fitted line. But in the fitting process, we employ PCA to obtain the slope rather than using ordinary least squares (OLS). The reason to do this is that OLS is not rotation...
invariant. As can be seen in Figure B1(b), OLS acts by minimizing the sum of the squared vertical distances between the line that best represents the data and the data points. If we invert $x$ and $y$, we expect that the fitted line does not change. However, the fitted lines from OLS and OLS after inverting $x$ and $y$ are not the same, which can be seen from the black and blue lines in Figure B1(a). This is because after $x$ and $y$ are inverted, the offsets that were used to minimize in OLS become horizontal, as shown in Figure B1(c). As a result, $x$ and $y$ are not equivalent, which is accessible in data fitting, but not in the feature extraction of the image.

Therefore, we suggest to use another variant of the least-squares method, in which the sum of the squared offsets (or distances) perpendicular to a line is minimized. The PCA (Pearson 1901) uses the principle components of data on behalf of the whole data. Specifically, it reduces data from $n$-dimensional to $m$-dimensional. We take the 2D data of the red points in Figure B1 as an example. We aim to reduce them into a line (1D data) and at the same time, the loss of information should be as low as possible after the reduction. To realize this, points are projected onto the line, and the spread of the projected data should be as large as possible. Since the mean position of the points is determined, maximizing the sum of the variance along the line is equivalent to minimizing the sum of the squared distances between the points and the line. So the procedure to find this line (strictly, the direction of the line as the first base vector) with a PCA in our 2D case is the same as the least squares with perpendicular offsets shown in Figure B1(d). Then the orientation angle is obtained from the slope of the line (or from the direction of the first base vector from a PCA). It is the property of least-squares method that the fitted line passes through the mean position $(\bar{x}, \bar{y})$, so the major axis is also determined.

### B.2. Concave Hulls and Aspect Ratios

The convex hull of a set of points in two dimensions is the minimum-area polygon that contains those points such that all internal angles between adjacent edges are smaller than 180°. For example, the black polygon in Figure B2(a) is the convex hull of our filament F7. A convex hull has been used to characterize the size and aspect ratio of MSTs since Schmeja & Klessen (2006). However, in some situations such as F7 in Figure B2(a), the convex hull does not represent the boundaries of a given set of points well. So we adopt a concave hull instead. A concave hull has at least one reflex interior angle. That is, an angle with a measure that is between 180° and 360°. The blue polygon in Figure B2(a) is the concave hull for F7 based on works of Moreira & Santos (2007). As can be seen, the concave hull describes the boundary of the MST better.
We then take the ratio of the area between the circle enclosing the vertices and the concave hull as the aspect ratio of a filament, similar as Gutermuth et al. (2009) for a convex hull. As shown in Figure B2(b), the black polygon is the concave hull of the filament. We take the mean position of the hull vertices as the center to obtain a circle, that is, the gray circle in Figure B2(b). The radius of the circle is the distance of the farthest vertex to the center. Then the aspect ratio is defined as the ratio of the circle area to the concave hull area, \( a = \frac{S_{\text{circ}}}{S_{\text{con}}}. \)

For an approximately elliptical distribution, \( f_A \) is very similar to the aspect ratio of the ellipse (Schmeja & Klessen 2006).

Appendix C

Influence of Galactic Environments on Large-scale Filaments

To investigate the influence of Galactic environments on large-scale filaments, we compare the properties of our filaments in spiral arms (25 of 163 filaments) and outside of filaments (the remaining 138 filaments). We find no significant distinctions in some of their physical properties such as mean temperatures, nonthermal velocity dispersion, and column densities (Figures C1(h), (f), and (p)). Zucker et al. (2019) also find that the length of filaments in an arm or outside an arm is invariant. They find that arm filaments tend to have a higher column density than interarm filaments, but they recall that more advanced numerical simulations are required to confirm their results. The filament catalog from Schisano et al. (2020) contains both large-scale and small-scale filaments and also shows that the column densities are similar in arm filaments and interarm filaments.

In addition to similitude, large-scale filaments also have distinctions such as galactocentric radius, height above the Galactic midplane, and mass. The distribution of the galactocentric radius of the arm filaments shows a peak in 4.6 kpc (Figure C1(n)), meaning that most of the population of filaments is found in the Scutum arm, which is consistent with Zucker et al. (2019). The arm filaments, on the whole, have a higher mass and linearity-weighted length than interarm filaments (Figures C1(q) and (o)). But this might be caused by the fact that arm filaments are mostly far away from us (Figure C1(c)). The distributions of the bones in this work with distance constraint have also been overlaid in Figure C1 with blue steps. The differences between arm filaments and bones in this work are that bones in this work have two additional conditions: (6) a height above the Galactic midplane of \(|z| \leq 20\) pc, and (7) an orientation angle between the filament major axis and the physical Galactic midplane of \(|\theta| \leq 30^\circ\). We do not find obvious distinctions between the distributions of the two samples.

The other panels in Figure C1 are (a) the Galactic longitude, (b) the Galactic latitude, (c) the distance from the Sun, (d) the height above the Galactic midplane, (e) the mean radial velocity of each filament, (f) the nonthermal velocity dispersion of the clumps in each filament, (g) the mean velocity gradient, (h) the mean temperature, (i) the mean separation of the clumps in each filament, (j) the line mass, (k) the volume density, (l) the angle between the major axes and the Galactic midplane in the projected sky, and (m) the linearity. There are obviously more arm filaments in northern than in southern sky (Figure C1(a)). This might be caused by the inaccuracy of the spiral arm model, where quadrant 4 is a fit result (Figure 1 in Reid et al. 2019). The distribution of the arm filament height above the Galactic midplane (Figure C1(d)) has a doublet that peaks at about \( z = \pm 20\) pc, and this might be an observation effect considering that our Sun resides 25 pc above the Galactic midplane. The distribution of the radial velocity of the arm filaments (Figure C1(e)) is due to a combination of the Galactic latitude and the rotation of the Milky Way.

We also find that the DGMF of filaments in the spiral arms have no significant distinction from interarm filaments, neither do the bones. For the arm filaments, the DGMFs vary in different spiral arms according to Abreu-Vicente et al. (2016), but we lack a statistical sample. Our filament catalog provides a larger sample to conduct a statistical analysis. We find about 10 filaments with a DGMF measurement in each spiral arm, except for the Perseus arm, which has only one. The mean DGMFs of the filaments in each spiral arm are 35.8% ± 9.8%, 41.0% ± 5.2%, 32.5% ± 7.0%, and 44.8% for the Norma–Outer, Scutum–Centaurus-OSC, Sagittarius–Carina, and the Perseus arm, respectively. The larger DGMF of Scutum–Centaurus-OSC and Perseus arm along with the larger bone fraction give us a hint of the peculiarity of these two arms.
Figure C1. Differences of filaments in arms or outside them. (a) Galactic longitude. (b) Galactic latitude. (c) Distance from the Sun. (d) Height above the Galactic midplane. (e) Mean radial velocity of each filament. (f) Nonthermal velocity dispersion of the clumps in each filament. (g) Mean velocity gradient of the edges in each filament. (h) Mean temperature. (i) Mean separation of the clumps in each filament. (j) Line mass. (k) Volume density. (l) Angle between the major axes and the Galactic midplane in the projected sky. (m) Linearity. (n) Galactocentric radius. (o) Linearity-weighted length. (p) Column density. (q) Mass.
Appendix D
Filaments on Emission Maps and Column Density Maps

We have shown the two-color view of some filaments in Figure 1. The two-color views of other filaments are shown in Figure D1. The color-coded circles denote clumps in filaments with various velocities. For backgrounds, cyan represents intermediate-infrared 24 μm emission on a logarithmic scale from MIPSGAL (Carey et al. 2009), and red shows submillimeter 870 μm emission on a linear scale from an APEX + Planck combined image (Csengeri et al. 2016).

Filaments are also overlaid on a Herschel Hi-GAL column-density map from PPMAP (Marsh et al. 2017) in Figure D2. As in Figure D1, the color-coded circles also denote clumps in filaments with various velocities. There are some gaps in coverage on the map, such as F21 in Figure D2. The reason is that in order to cover the entire Galactic Plane in a reasonable amount of time, Marsh et al. (2017) had to truncate the field of each tile, and unfortunately, the truncation was excessive in some places. A column-density map for F147 (with Galactic latitude 1°43) is absent because it is outside the coverage of the Herschel Hi-GAL map.

Figure D1. Two-color view of filaments. The color-coded circles denote clumps in filaments with various velocities. For backgrounds, cyan represents intermediate-infrared 24 μm emission on a logarithmic scale from MIPSGAL (Carey et al. 2009), and red shows submillimeter 870 μm emission on a linear scale from an APEX + Planck combined image (Csengeri et al. 2016).

(An extended version of this figure is available.)
Figure D2. Filaments overlaid on column-density maps from PPMAP (Marsh et al. 2017). (An extended version of this figure is available.)
