



A CENSUS OF LARGE-SCALE (≥ 10 PC), VELOCITY-COHERENT, DENSE FILAMENTS IN THE NORTHERN GALACTIC PLANE: AUTOMATED IDENTIFICATION USING MINIMUM SPANNING TREE

KE WANG (王科)¹, LEONARDO TESTI^{1,2,3}, ANDREAS BURKERT^{4,5}, C. MALCOLM WALMSLEY^{6,7}, HENRIK BEUTHER⁸, AND THOMAS HENNING⁸

¹ European Southern Observatory (ESO) Headquarters, Karl-Schwarzschild-Str. 2, D-85748 Garching bei München, Germany; kwang@eso.org

² Excellence Cluster universe, Boltzmannstr. 2, D-85748 Garching bei München, Germany

³ INAF—Osservatorio astrofisico di Arcetri, Largo E. Fermi 5, I-50125 Firenze, Italy

⁴ University Observatory Munich, Scheinerstrasse 1, D-81679 Munich, Germany

⁵ Max-Planck-Institute for Extraterrestrial Physics, Giessenbachstrasse 1, D-85758 Garching, Germany

⁶ INAF—Osservatorio astrofisico di Arcetri, Largo E. Fermi 5, I-50125 Firenze, Italy

⁷ Dublin Institute of Advanced Studies, Fitzwilliam Place 31, Dublin 2, Ireland

⁸ Max-Planck Institute für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany

Received 2016 May 8; revised 2016 July 5; accepted 2016 July 8; published 2016 September 8

ABSTRACT

Large-scale gaseous filaments with lengths up to the order of 100 pc are on the upper end of the filamentary hierarchy of the Galactic interstellar medium (ISM). Their association with respect to the Galactic structure and their role in Galactic star formation are of great interest from both an observational and theoretical point of view. Previous “by-eye” searches, combined together, have started to uncover the Galactic distribution of large filaments, yet inherent bias and small sample size limit conclusive statistical results from being drawn. Here, we present (1) a new, automated method for identifying large-scale velocity-coherent dense filaments, and (2) the first statistics and the Galactic distribution of these filaments. We use a customized minimum spanning tree algorithm to identify filaments by connecting voxels in the position–position–velocity space, using the Bolocam Galactic Plane Survey spectroscopic catalog. In the range of $7.5^\circ \leq l \leq 194^\circ$, we have identified 54 large-scale filaments and derived mass ($\sim 10^3$ – $10^5 M_\odot$), length (10–276 pc), linear mass density (54 – $8625 M_\odot \text{ pc}^{-1}$), aspect ratio, linearity, velocity gradient, temperature, fragmentation, Galactic location, and orientation angle. The filaments concentrate along major spiral arms. They are widely distributed across the Galactic disk, with 50% located within ± 20 pc from the Galactic mid-plane and 27% run in the center of spiral arms. An order of 1% of the molecular ISM is confined in large filaments. Massive star formation is more favorable in large filaments compared to elsewhere. This is the first comprehensive catalog of large filaments that can be useful for a quantitative comparison with spiral structures and numerical simulations.

Key words: catalogs – Galaxy: structure – ISM: clouds – ISM: structure – stars: formation

1. INTRODUCTION

The interstellar medium (ISM) has a highly filamentary and hierarchical structure. On the upper end of this filamentary hierarchy are large-scale gaseous filaments with lengths up to the order of 100 pc. What is their distribution in our Galaxy and what role do they play in the context of Galactic star formation? The answers to these questions are important for a critical comparison with theoretical studies and numerical simulations of galaxy formation and filamentary cloud formation. The observational key to answering these questions is a homogeneous sample of large filaments across the Galaxy identified in a uniform way.

Studies in the past years have revealed a number of large filaments with a wide range of aspect ratios and morphologies, from linear filaments to a collection of cloud complexes. Goodman et al. (2014) find that the 80 pc-long infrared dark cloud (IRDC) “Nessie” (Jackson et al. 2010), in the southern sky, can be traced up to 430 pc in the position–position–velocity (PPV) space in ^{12}CO (1–0), guided by connecting the IR-dark patches presumably caused by high column density regions extinguishing the otherwise smooth IR background emission from the Galactic plane. They argue that Nessie runs in the center of the Scutum–Centaurus spiral arm in the PPV space, referred to as a “bone” of the Milky Way. In a follow-up study, Zucker et al. (2015) searched the region covered by the MIPS GAL (*Spitzer*/MIPS Galactic Plane Survey, $|l| < 62^\circ$, $|b| < 1^\circ$;

Carey et al. 2009), focusing on the PPV loci of arms expected by various spiral arm models and finding 10 bone candidates with lengths of 13–52 pc and aspect ratios of 25–150. Ragan et al. (2014) and Abreu-Vicente et al. (2016) extend this “mid-IR extinction” method to a blind search, i.e., not restricted to arm loci but the full extent of the observed PPV space. They find 7 and 9 filaments with lengths of 38–234 pc in part of the first and fourth Galactic quadrants covered by the Galactic Ring Survey (GRS) (Jackson et al. 2006), and the Three-mm Ultimate Mopra Milky Way Survey (ThrUMMS; Barnes et al. 2015), respectively. The aspect ratios of those filaments are not well defined due to the complex morphology, but inferring from the figures in the papers, the typical aspect ratio is much less than 10.

In contrast to the indirect⁹ “mid-IR extinction” method, Wang et al. (2015) identify large filaments directly based on emission at far-IR wavelengths near the spectral energy distribution (SED) peak of cold filaments. They develop a Fourier Transform filter to separate high-contrast filaments from the low-contrast background/foreground emission. Fitting the SED built up from the multi-wavelength *Herschel* data from the Hi-GAL survey (Molinari et al. 2010), they derive temperature and column density maps, and have used those maps to select the “largest, coldest, and densest” filaments.

⁹ Indirect because “IR-dark” does not necessarily correspond to a dense cloud; it can also be caused by a real “hole in the sky” (Jackson et al. 2008; Wilcock et al. 2012).

They present 9 filaments with lengths of 37–99 pc and aspect ratios of 19–80, identified primarily from the GRS field.

These systematic searches have started to uncover the spatial distribution of large filaments in our Galaxy, revealing filaments within and outside major spiral arms. However, with different searching methods and selection criteria, in addition to inherent bias from manual inspection, it is difficult to cross-compare the results from these studies. The small sample size also limits the robustness of statistical attempts (e.g., see discussion in Wang et al. 2015). All of the above mentioned searches start from a “by-eye” inspection of dust features (either mid-IR extinction or far-IR emission), identify candidate filaments, and then verify the coherence in radial velocity using gas tracers-spectral line data.

We automate the identification process by applying a customized minimum spanning tree algorithm to the PPV space. We present the first homogeneous sample of 54 large-scale velocity-coherent filaments in the range of $7.5 \leq l \leq 194^\circ$ (see exact coverage in Section 2). We derive mass, length, linearity, aspect ratio, velocity gradient and dispersion, temperature, column/volume density, fragmentation, Galactic location, and orientation angle. For the first time, we are able to investigate the Galactic distribution of their physical properties, and to estimate the fraction of the ISM confined in large filaments and star formation therein.

We describe the data set in Section 2 and present our identification method in Section 3. The identified sample of filaments and their physical properties and statistics are presented in Section 4, followed by a discussion of the nature and implications of the filaments in Section 5. Our main conclusions are summarized in Section 6. Following the spirit of Wang et al. (2015), we focus on the densest filaments traced by millimeter dust continuum emission, and not the more diffused CO filaments.

2. DATA: A COMPLETE SPECTROSCOPIC CATALOG

The Bolocam Galactic Plane Survey (BGPS) is a blind mapping of the northern Galactic plane at 1.1 mm using the Caltech Submillimeter Observatory 10 m telescope with an effective resolution of $33''$, revealing over 8400 continuum sources (Rosolowsky et al. 2010; Aguirre et al. 2011). Spectroscopic follow-ups carried out by Schlingman et al. (2011) and Shirley et al. (2013) have observed all the 6194 BGPS sources in the longitude range of $7.5 \leq l \leq 194^\circ$ in dense gas tracers HCO^+ (3–2) and N_2H^+ (3–2), using the 10 m Heinrich Hertz Submillimeter telescope, with a FWHM beamwidth of $15''$. The detection rate is about 50%, and about 99% of the detections show a unique velocity component.

From these observations Shirley et al. (2013) compiled a complete spectroscopic catalog of 3126 sources with a single velocity component resolved in HCO^+ (3–2) and/or N_2H^+ (3–2). In a typical temperature range of 10–20 K of the BGPS sources (Dunham et al. 2011), these two lines have a critical density of $\sim 10^6 \text{ cm}^{-3}$ and an effective excitation density of the order of 10^4 – 10^5 cm^{-3} (Shirley 2015), thus they trace very dense gas. In the following, we refer to these sources as “dense BGPS sources.” At a typical distance of few kiloparsecs, a detection of the lines toward a BGPS 1.1 mm continuum peak marks the presence of parsec-scale dense gas, and a chain of such clumps connected in PPV means a rather prominent structure. Therefore, this catalog is an excellent data set for searching for velocity-coherent filaments.

Note that the coverage of the BGPS spectroscopic catalog is contiguous in the range of $7.5 \leq l \leq 90.5$, $|b| \leq 0.5$, with latitude coverage flaring up to $|b| \leq 1.5$ in several longitude cuts. In the outer Galaxy, four selected regions were observed (l range in [98.85, 100], [110, 112], [132.5, 138.5], and [187.5, 193.5] degrees). The BGPS spectroscopic survey (Shirley et al. 2013) used the version 1 BGPS continuum source catalog (Aguirre et al. 2011). In the version 2 catalog, Ginsburg et al. (2013) resolved an offset in flux scale: $S_{v2} = 1.5S_{v1}$. In this study, we use the flux from v2 (Section 4).

3. METHOD: CONNECTING DOTS USING CUSTOMIZED MST

The minimum spanning tree (MST) was first introduced by Borůvka (1926a, 1926b) to optimize the cost of electrical grids by minimizing its total length. It is now widely used in general optimization of a variety of networks (see the review by Nešetřil et al. 2001). We adopt the MST algorithm to isolate coherent filaments out of a PPV catalog, bearing in mind that “coherence” means “close proximity” in position *and* velocity. Our method is demonstrated in Figure 1. An MST connects all the nodes (BGPS clumps) in a graph with the cost of a minimum sum of edge lengths, where “edge” refers to the separation between two nodes (Figure 1(a)). We customize the MST such that a graph is connected not in one MST, but in a collection of MSTs with the following criteria:

- (1) The accepted MST must contain at least five BGPS clumps: $N_{cl} \geq 5$.
- (2) Only edges shorter than a maximum length can be connected ($\Delta L < 0.1$, Figure 1(b)).
- (3) For any two clumps to be connected, the difference in line-of-sight velocity (Δv) must be less than 2 km s^{-1} (Figure 1(c)).

The maximum edge length and velocity difference (criteria 2–3) are chosen based on characteristics of previously known filaments (Wang et al. 2015). As in Wang et al. (2015), we have used the “Snake” nebula, one of the first identified IRDCs (Carey et al. 1998), as a primary guide to test criteria (2–3). In addition to its prominent IR extinction feature, the Snake is known to have velocity-coherent structures at multiple spatial scales (Carey et al. 2000; Wang et al. 2014; Wang 2015). Together, these properties make the Snake a good test case. During the tests, we relax criteria (2–3) by increasing ΔL and Δv until the MST starts to (incorrectly) connect unrelated BGPS sources into the Snake. We also require the algorithm to be able to identify as many as possible previously known filaments from other studies (see Section 5.1), but not to connect unrelated sources. Criterion (1) is arbitrary, but we note that there is no difference if we set the minimum number of clumps N_{cl} as 5 or 6. It is clear in Table 1 which filaments would be picked up if increased the required N_{cl} .

So far, these MSTs are *coherent* structures in the (l, b, v) space, but not necessarily *filamentary* structures. To further select filaments we introduce a linearity check. For a given filament, we fit a straight line to the (l, b) data points. The fitted line represents the filament’s major axis. Linearity is defined as the ratio between the spread (standard deviation) of data points measured along the filament’s major axis to the spread along minor axis: $f_L = \sigma_{\text{major}}/\sigma_{\text{minor}}$. After visual inspection we accept structures with

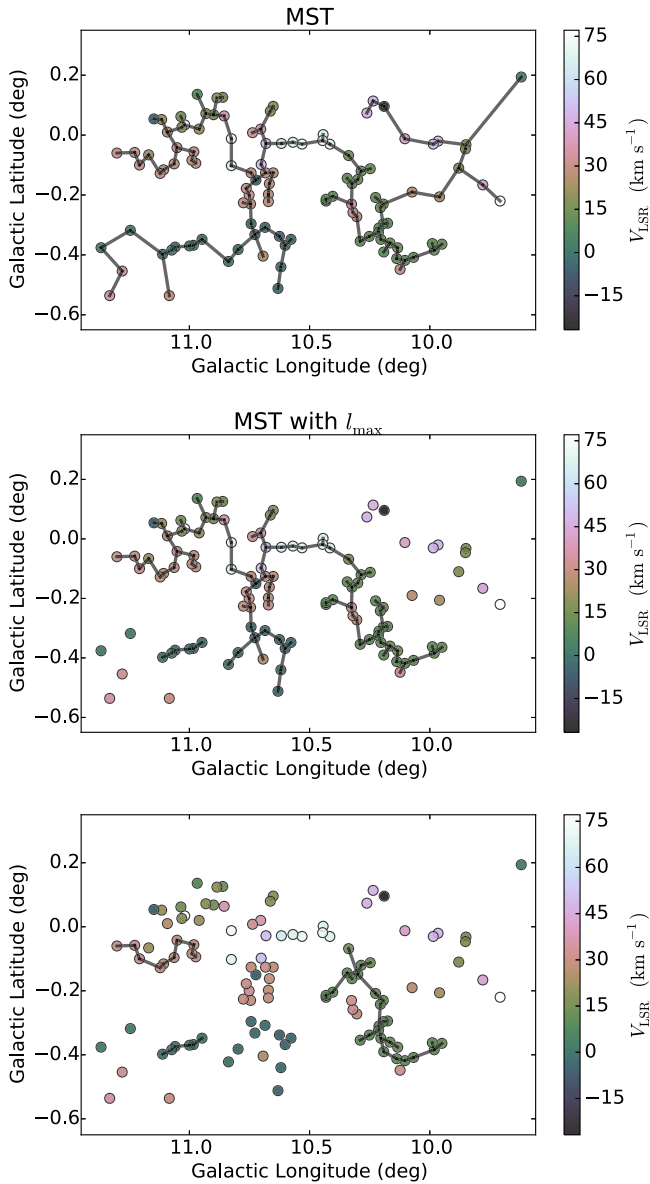


Figure 1. Demonstration of the customized MST applied to a subset of the Galactic longitude. Upper panel: an MST. Middle panel: MSTs with a maximum edge length and minimum number of clumps (criteria 1–2 in Section 3). Lower panel: MSTs with edge limited to $\Delta v < 2 \text{ km s}^{-1}$ (criteria 1–3). Circles represent BGPS sources color-coded by radial velocity as shown in the color bar.

(4) Linearity $f_L > 1.5$

as filaments. Note that the linearity is defined to quantify how far the structure is away from a linear structure. For a straight line, $f_L \rightarrow \infty$. For bent or wiggling filaments, which is often the case, the linearity is much smaller than the aspect ratio. For instance, the famous IRDC “Snake” (F7 in Table 1) has an aspect ratio of 43, while its linearity is only 4 because of its sine wiggling in the (l, b) space (Wang et al. 2014, 2015; Wang 2015). Finally, after determining the distance (Section 4.1) we accept long filaments with

(5) Projected length $\geq 10 \text{ pc}$

as large-scale filaments for the interest of this study.

Applying the methodology to the BGPS spectroscopic catalog, we have identified 91 velocity-coherent structures

(satisfying criteria 1–3), 70 of which are linear filaments (satisfying criteria 1–4). Among these, 54 are large-scale filaments (satisfying criteria 1–5), including 48 in the first Galactic quadrant where the BGPS coverage is contiguous, and 5/1 in the second/third quadrants, respectively, where the BGPS coverage is targeted to known star formation regions. Of the 54 filaments, only 9 are previously known (F7, F13, F25, F32–F33, F35–F37, and F41; Section 5.1). Clearly, our filaments identification method depends on free parameters (like many other methods),¹⁰ hence the identified filaments and their properties depend on the chosen parameters of criteria (1–5). Here we have chosen reasonable criteria to select the *representative* large-scale velocity-coherent filaments based on previously known filaments.

Each of the 54 filaments are plotted in Figure 6 in a two-color view, where the mid-IR 24 or 22 μm emission (Carey et al. 2009; Wright et al. 2010) is shown in cyan and the (sub) millimeter 0.87 or 1.1 mm emission (Ginsburg et al. 2013; Csengeri et al. 2016) is shown in red. The MST edges are drawn to outline the filaments. Most (40 out of the 54) filaments are IRDCs, while 14 are IR-bright filaments. The filaments show a wide range of filamentary morphologies (see the next section).

4. PHYSICAL PARAMETERS

Table 1 lists the physical parameters of the 54 identified filaments. Column (1) assigns identification numbers running from F1 to F54. Columns (2)–(4) feature flux-weighted longitude, latitude (in degree), and local standard of rest (LSR) velocity (km s^{-1}). For instance, for a filament containing n clumps, its flux-weighted longitude is $l_{\text{wt}} = \frac{\sum_{i=1}^n F_i \times l_i}{\sum_{i=1}^n F_i}$, where F_i and l_i are the BGPS flux and the longitude of the i th clump, respectively. Columns (5) and (6) are distance (kpc) and its type (see Section 4.1). Column (7) is the number of clumps in the filament. Columns (8)–(9) show the length of the filament in degrees and parsecs by summing all the edges in the filament. Column (10) shows the velocity gradient ($\text{km s}^{-1} \text{pc}^{-1}$): the mean of all the edges. We caution that both velocity and length are subject to projection effect. Column (11) is the dispersion of the central velocity of all the clumps in the filament (km s^{-1}). Columns (12)–(13) show the minimum and maximum temperatures of the clumps (see Section 4.2). Columns (14)–(15) show the filament mass (in unit $10^3 M_\odot$) and linear mass density ($M_\odot \text{pc}^{-1}$). Mass is computed from the integrated BGPS $\nu 2.1 \text{ mm}$ dust emission flux measured in a polygon encompassing the filament guided by the MST, adopting the Ossenkopf & Henning (1994) dust opacity law and $\beta = 1.5$, and accounting for the different temperatures in the clumps (Section 4.2). Columns (16)–(17) feature the molecular hydrogen column density (10^{22} cm^{-2}) and volume density (10^3 cm^{-3}) of the filament. These are estimated by simplifying the filament as a cylinder with a length of the filament length and a diameter of the mean major axes of the clumps. Column (18) contains the aspect ratio f_A , estimated by dividing filament length with the averaged major axes of the clumps. Column (19) features linearity f_L (see the definition in

¹⁰ Other filament-finding methods include: getfilaments (Men’shchikov 2013), DisPerSE (Sousbie 2011; Panopoulou et al. 2014), FilFinder (Koch & Rosolowsky 2015), Hessian matrix (Schisano et al. 2014; Salji et al. 2015), and Bisous model (Tempel et al. 2016). Our customized MST and the Bisous model work on discrete points, and others work on continuous images.

Table 1
(Continued)

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)	(27)
ID	l_{wt}	b_{wt}	v_{wt}	d	d_{type}	N_{cl}	L_{deg}	L_{pc}	$\left(\frac{\Delta v_i}{\Delta L_i}\right)$	σ_v	T_{min}	T_{max}	Mass	M/L	N_{H_2}	n_{H_2}	f_{Λ}	f_L	R_{gc}	z	θ	Morph.	l_{min}	l_{max}	b_{min}	b_{max}
F53	134.21	0.75	-48.4	4.1	ML	8	0.31	22.1	0.31	1.50	7.0	15.0	49.6	2242.2	4.91	7.11	27.3	2.8	11.6	90.5	85.0	C	134.20	134.28	0.64	0.86
F54	192.62	-0.03	8.0	1.6	ML	10	0.50	13.8	0.93	1.08	9.2	42.9	2.2	156.0	1.18	5.85	55.8	3.2	9.9	30.1	57.1	C	192.58	192.82	-0.16	0.13
Min	8.01	-1.17	-56.8	1.2	...	6	0.17	10.1	0.06	0.30	4.8	15.0	0.7	53.5	0.23	0.38	18.1	1.5	3.4	-120.8	-80.2
Max	192.62	0.98	112.9	16.4	...	35	1.50	276.2	1.67	2.35	15.0	48.7	640.4	8624.7	16.94	22.0	176.1	16.9	11.6	90.5	87.2
Med	25.04	-0.10	40.6	4.4	...	8	0.35	27.5	0.33	1.26	15.0	23.2	10.5	360.5	0.94	1.77	41.4	3.0	5.4	10.9	9.1
Mean	40.95	-0.02	41.8	5.1	...	10	0.43	35.7	0.43	1.27	13.3	23.9	31.2	644.1	1.57	3.53	50.0	3.7	6.1	11.5	9.8
Std	38.95	0.38	45.2	3.1	...	6	0.27	37.3	0.31	0.43	2.8	8.5	90.5	1224.6	2.38	4.8	29.7	2.7	2.1	35.5	42.0
S	1.78	0.32	-0.5	1.7	...	1	1.91	5.1	1.79	0.35	-1.7	1.1	5.9	5.4	0.53	0.26	2.0	2.9	1.0	-0.7	-0.2
K	3.11	1.11	-0.3	3.0	...	3	3.73	30.3	3.69	-0.01	1.8	0.7	36.7	32.0	3.08	0.62	4.7	10.4	0.0	2.8	-0.3

Note. A brief description of the columns (for a detailed description including uncertainties see Section 4): Column (1) assigned ID; Columns (2)–(4) flux-weighted longitude, latitude (in degrees), and LSR velocity (km s^{-1}); Columns (5) and (6) distance (kpc) and its type; Column (7) number of clumps in the filament; Columns (8) and (9) length of the filament in degree and parsecs; Column (10) velocity gradient ($\text{km s}^{-1} \text{pc}^{-1}$); Column (11) dispersion of the central velocity (km s^{-1}) of all the clumps in the filament; Columns (12) and (13) minimum and maximum temperature of the clumps in the filament; Columns (14) and (15) filament mass (in unit $10^3 M_{\odot}$) and linear mass density (M_{\odot}/pc); Column (16) and (17) estimated H_2 column density (10^{22}cm^{-2}) and volume density (10^3cm^{-3}) of the filament; Columns (18) and (19) aspect ratio and linearity; Column (20) Galactocentric radius (kpc); Column (21) vertical distance (pc) to the physical Galactic mid-plane; Column (22) orientation angle (degree) between the filament's long axis and the physical Galactic mid-plane; positive/negative angle means Galactic latitude increases/decreases with increasing longitude; Column (23) Morphology class; Columns (24)–(27) Galactic coordinate boundary of the filament (degree).

Section 3). Column (20) R_{gc} is the Galactocentric radius (kpc). Column (21) z is the vertical distance (pc) to the physical Galactic mid-plane after correction for the Sun's displacement of 25 pc above the mid-plane and the true position of the Galactic Center (Goodman et al. 2014; Wang et al. 2015). Column (22) θ is the orientation angle (degree) between the filament's long axis and the physical Galactic mid-plane. A positive/negative angle means that Galactic latitude increases/decreases with increasing longitude (Figure 1(c)). Column (23) shows the morphology class as defined in Wang et al. (2015): L: linear straight or L-shape; C: bent C-shape; S: quasi-sinusoidal shape; X: crossing of multiple filaments; H: head-tail or hub-filament system. Some filaments are characterized by more than one class. Different morphologies may have resulted from different filament formation processes. For example, an expansion of bubbles can produce C-shaped filaments, a collision of bubbles can make S-shaped filaments, a gravitational contraction of a clump embedded in a sheet can make H-type filaments, while turbulence and converging flows can make filaments of any shape. Columns (24)–(27) feature the Galactic coordinate boundary of the filament (degrees).

The last rows of Table 1 list statistics of these physical parameters. These include minimum, maximum, median, mean, standard deviation, skewness (K), and kurtosis (K). Skewness is a measure of symmetry. A symmetric distribution has $S = 0$, while a negative/positive value of S means asymmetric tails with lower/higher values around the mean, respectively. Kurtosis is a measure of how the distribution is compared to a normal distribution (e.g., Gaussian): a Gaussian distribution has $K = 0$, while a negative or positive value K means the distribution is a flatter or more centrally peaked distribution than Gaussian, respectively. These statistics provide a description of each parameter, which is discussed in Section 4.3.

The major sources of uncertainty of the parameters originate from the uncertainties in distance, dust opacity, and projection. Because projection is unknown, we do not correct for it in Table 1, but we discuss its effect on different parameters here. The dust opacity tabulated in Ossenkopf & Henning (1994) is subject to a factor of 2 uncertainty. The typical distance uncertainty is 10%–30%, depending on different distance types (see Section 4.1). Distance uncertainty propagates to other parameters. In the following estimation we use a conservative 30% distance uncertainty. Length includes distance uncertainty of 30% and projection. Due to projection, the length is a lower limit. For a random inclination angle (ϕ , defined as the angle between the line of sight and the filament's long axis) of 1 radian, $L_{\text{obs}} = L \times \sin(\phi) = 0.84L$. Velocity gradient includes distance uncertainty of 30% and projection. Projection affects both velocity and length, in the form of $\tan(\phi)$, thus projection can lead to an over- or underestimation of the velocity gradient. For $\phi = 30^\circ, 57.3^\circ, 75^\circ$, the observed velocity gradient is the true value times $\cos(\phi)/\sin(\phi) = 1.73, 0.64, 0.27$, respectively. Mass: includes uncertainties in dust opacity, distance (scaling as d^2), and flux (10%, Ginsburg et al. 2013). Together, these uncertainties propagate to <50% uncertainty in mass. Note that we have accounted for the temperature variance across the filament (similar to Wang et al. 2015), an improvement over other studies in which a uniform temperature is assumed for all filaments (e.g., Ragan et al. 2014; Abreu-Vicente et al. 2016). Mass per unit length: Compared to mass, M/L has a smaller dependence on distance uncertainty (scaling as d^1), thus the error is less than that of mass. But M/L is affected by

projection in the form of $1/L$. Column density: defined as $\text{mass}/(\text{length} \times \text{width})$, the distance error is eliminated, thus the uncertainty in column density is less than that of mass. Projection affects length but not width, so projection is in the form of $1/L$. Additionally, simplifying filaments into cylinders can cause uncertainties for some filaments that deviate from a linear structure (with a relatively small linearity f_L). Volume density is defined as $\text{mass}/(\text{length} \times \text{width} \times \text{depth})$, where depth is assumed to be the same as width. Compared to mass, volume density has a smaller dependence on distance uncertainty (scaling as $1/d$), thus the uncertainty is less than that of mass. But volume density is affected by projection in the form of $1/L$. Similar to column density, simplifying filaments into cylinders can cause additional uncertainties for some filaments that deviate from a linear structure.

4.1. Distance Estimation

Distance is determined by three methods, listed in decreasing order of robustness: (1) type P: trigonometric parallax measurements of associated masers from the BeSSeL project (Brunthaler et al. 2011, 2009; Moscadelli et al. 2009; Kurayama et al. 2011; Xu et al. 2011; Immer et al. 2013; Wu et al. 2014; A. Sanna 2016, private communication); (2) type ML: maximum likelihood distance from Bayesian evaluation of kinematic distance using external data to place priors (Ellsworth-Bowers et al. 2013, 2015); and (3) type KN, KF: near or far kinematic distance computed using the procedure of Reid et al. (2009), with updated Galactic parameters from Reid et al. (2014).

Wherever available, we use parallax distance, then Bayesian distance, then kinematic distance. Of the 54 filaments, 9 are assigned for type P, 38 for ML, 4 for KN, and 3 for KF distances. When more than one clump in a given filament has an ML distance, the median of the clump distances is used, computed after excluding extreme values. This occurs, for example, when a filament is generally IR-dark (therefore more likely to be located at near distance), but a minority of its clumps are IR-bright, so the far distance is assigned by the ML evaluation. Kinematic distance is used in seven cases. The distance ambiguity is resolved with the IR emission/extinction (see Figure 6) and the fact that all clumps in a given filament should have a consistent distance computed by the Reid et al. (2009) code.

4.2. Temperature Estimation

We evaluate the temperature of every clump from three resources, listed in order of decreasing priority to search for a match: (1) gas kinetic temperature determined from NH_3 for a subset of the BGPS sample (Dunham et al. 2011), (2) same but for a subset of the Atacama Pathfinder EXperiment Telescope Large Area Survey of the Galaxy (ATLASGAL) sample (Wienen et al. 2012), and (3) dust color temperature determined by comparing 350 μm and 1.1 mm fluxes for a subset of the BGPS sample (Merello et al. 2015). We exclude color temperatures with large uncertainties (>100%).

When querying for NH_3 temperature the clump must match in position and velocity, and for color temperature only position is available for matching. If a clump does not have a match, we assume a temperature of 15 K based on the average value of BGPS clumps (Dunham et al. 2011). Among the 54 filaments, 46 have at least 1 temperature match. The minimum and maximum temperature values are listed in Table 1.

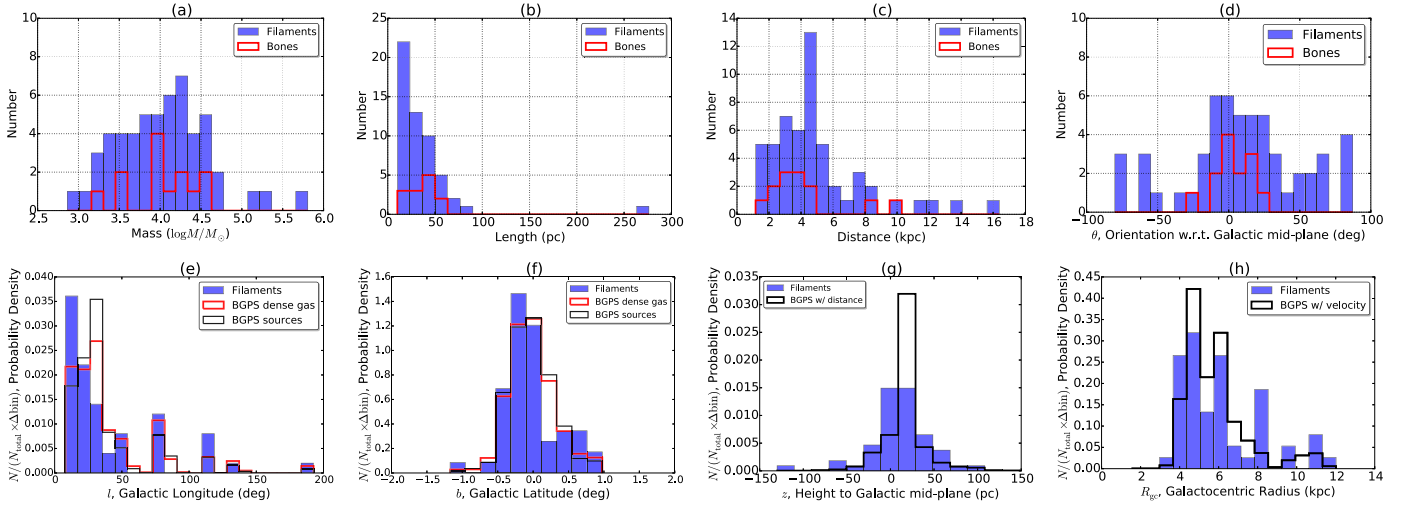


Figure 2. Panels (a)–(d) show histograms of filament mass, length, distance, and orientation angle. The 54 filaments and 13 bones are labeled. Panels (e)–(h) plot normalized distributions of l , b , z , and R_{gc} , comparing filaments with BGPS sources. “BGPS sources” are the BGPS clumps covered in the searching field. “BGPS dense gas” means the 3126 BGPS clumps with dense gas detection in Schlingman et al. (2011) and Shirley et al. (2013). The 1710 “BGPS w/distance” source and 3508 “BGPS w/velocity” sources are from Ellsworth-Bowers et al. (2015).

4.3. Statistics of the Parameters

Figure 2 presents histograms of selected parameters. Panels (a)–(d) show mass, length, distance, and orientation angle for the filaments. Panels (e)–(h) present normalized histograms of l , b , z , R_{gc} where the distribution of filaments is compared to an appropriate BGPS sub-sample when data are available.

Among the 54 identified filaments, F5 stands out as an extreme: 276 pc and $6.4 \times 10^5 M_\odot$. Excluding this extreme, the other 53 filaments extend 10–93 pc over the projected sky, with a mass of the order of 10^3 – $10^5 M_\odot$. For all 54 filaments, the angular length lies in the range of $0^\circ.17$ – $1^\circ.50$, with a median of $0^\circ.35$. The column density N_{H_2} lies in the range of $(0.2$ – $16.9) \times 10^{22} \text{ cm}^{-2}$, with a median value of $0.9 \times 10^{22} \text{ cm}^{-2}$, and volume density n_{H_2} is in the range $(0.4$ – $22) \times 10^3 \text{ cm}^{-3}$, with a median value of $1.8 \times 10^3 \text{ cm}^{-3}$. The filaments are 1–2 orders of magnitude denser than the filaments identified by the “mid-IR extinction” (Ragan et al. 2014; Zucker et al. 2015; Abreu-Vicente et al. 2016), while as dense as the filaments identified from *Herschel* far-IR emission by Wang et al. (2015).

The mass–length relation is best-fitted as $\lg L = 0.41 \times \lg M - 0.19$. This suggests that the filaments are not N small unrelaxed filaments that appear to be accidentally connected. Suppose the sub-units have a typical length l and a typical mass m . The total length of the filament would then be $L = N \times l$, with mass $M = N \times m$, resulting in $L \sim M$ and not $L \sim M^{0.41}$ as observed here. The mass–length relation implies $M \sim L^{2.4}$, i.e., a fractal dimension of 2.4. This is comparable with the observed GRS molecular cloud catalog (although of whole clouds) in the Milky Way (Roman-Duval et al. 2010) and with numerical simulations of supersonic gas turbulence (Federrath et al. 2009). The mass–length scaling of filaments, from large to small-scales, will be discussed in detail in a forthcoming paper (A. Burkert et al. 2016, in preparation).

The orientation of the filaments is clearly not random. The distribution of the orientation angles θ is close to Gaussian ($K = -0.3$), showing a weak but significant concentration toward 0° . The vertical distribution z of the filaments is surprisingly not symmetric with respect to the Galactic plane ($S = -0.7$). It skews toward the negative values, with a mean at

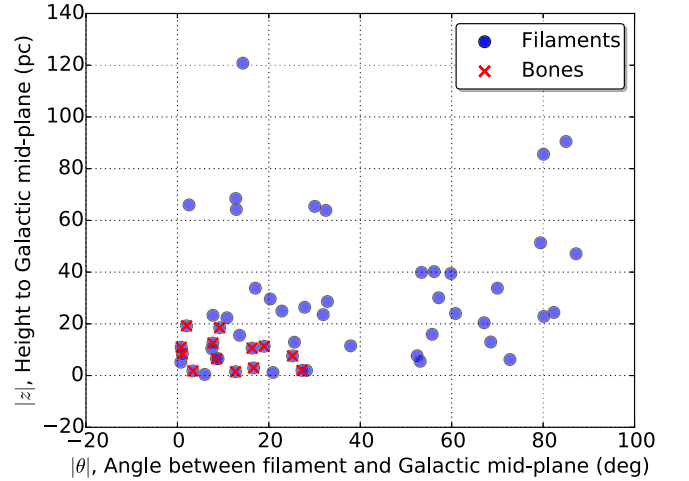


Figure 3. Filament orientation angle $|\theta|$ plotted vs. the vertical height to the physical Galactic mid-plane $|z|$. The 54 filaments and 13 bones are labeled. For reference, the height of the Sun is $z_\odot = +25$ pc.

a positive value of $z = 11.5$ pc. This is likely an observational bias (see Section 5.2). Nevertheless, z concentrates toward small values, with 50% of the (27/54) filaments located within $|z| \leq 20$ pc and 70% (38/54) within $|z| \leq 30$ pc. Toward higher vertical positions the number of filaments decreases much faster than a Gaussian function ($K = 2.8$). Three parameters have a Gaussian-like distribution: the Galactocentric radius R_{gc} ($K = 0.0$), the clump velocity dispersion $\sigma(v)$ ($K = -0.01$), and the flux-weighted LSR velocity v_{wt} ($K = -0.3$). Notably, the mean velocity gradient along the filament is small ($0.43 \pm 0.31 \text{ km s}^{-1} \text{ pc}^{-1}$), but in broad agreement with simulations (Duarte-Cabral & Dobbs 2016).

Across the Galaxy, in general, the distribution of filaments follows the number density of BGPS sources, i.e., it is more likely to find a filament where the BGPS sources are crowded. Specifically, the probability density¹¹ of filaments (P_{Filament})

¹¹ The probability density is defined as $P = N_{\text{bin}}/(N_{\text{total}} \times \Delta_{\text{bin}})$, so that the integral of the histogram will sum to 1.

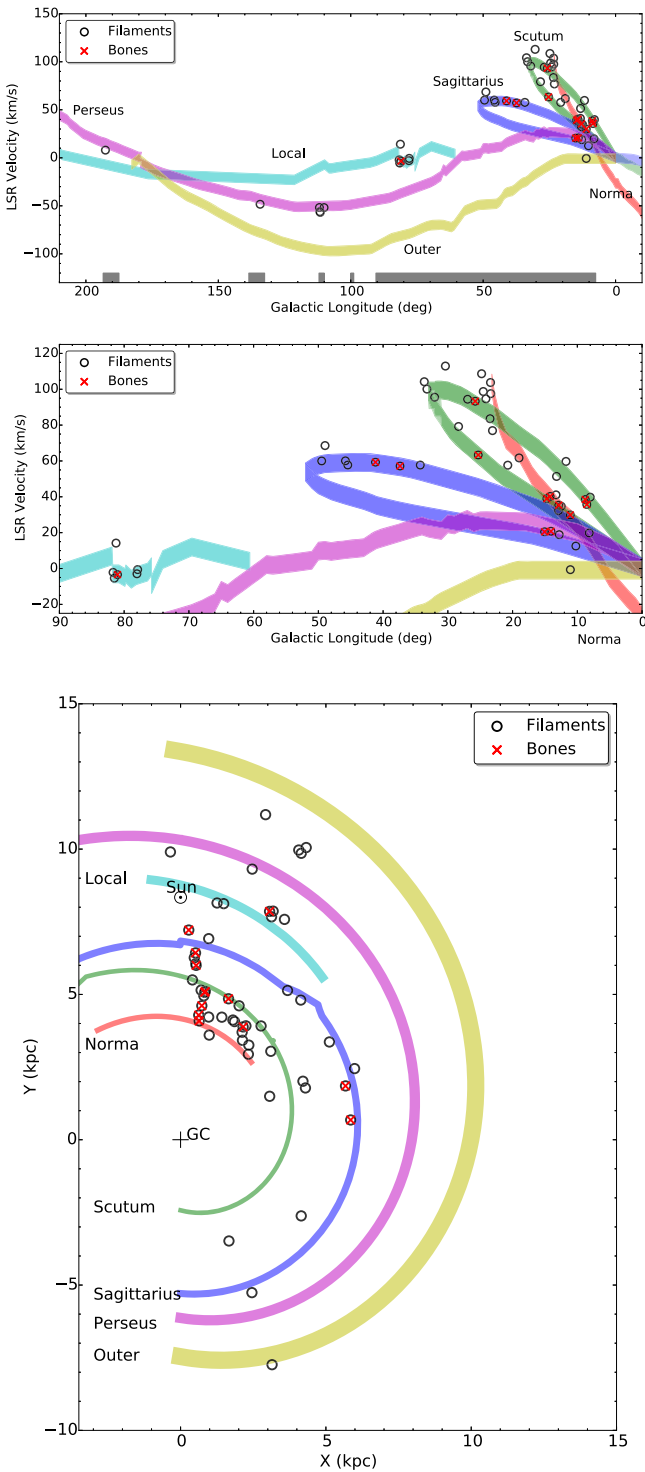


Figure 4. Galactic distribution of the filaments. Upper panel: the longitude-velocity plot showing the spiral arm segments derived from maser parallaxes (Reid et al. 2014; M. Reid et al. 2015, private communication). For simplicity, only related arms (Scutum, Sagittarius, Norma, Local, Perseus, and Outer) are plotted. The color shaded segments are of $\pm 5 \text{ km s}^{-1}$ width with respect to the arm centers. The 54 filaments and 13 bones are labeled. The gray shaded horizontal strips along the x -axis depict the searched longitude ranges. Middle panel: a zoom-in of the upper panel for clarity. Lower panel: a “face-on” view from the northern Galactic pole. The arm widths (170–630 pc) are from Reid et al. (2014), except for Norma, whose width is not available, and we plot 200 pc width for reference. The solar symbol \odot is plotted at (0, 8.34) kpc.

largely agrees with the probability density of BGPS sources: $P_{\text{Filament}} \sim P_{\text{BGPS}}$, as binned in l , b , z , R_{gc} (Figures 2(a)–(d)). This is expected because we used the BGPS spectroscopic catalog as input, which is a homogeneous sub-sample of the BGPS catalog. However, there are interesting exceptions. First, a significantly lower filament probability ($P_{\text{Filament}} \lesssim 0.5 \times P_{\text{BGPS}}$) is seen toward $l \sim 30^\circ$, the Scutum tangent. The same discrepancy is seen toward $b \sim 0.2$, $z \sim 20 \text{ pc}$, and $R_{\text{gc}} \sim 5 \text{ kpc}$. Second, on the other hand, a much higher P_{Filament} is found in the inner Galaxy of $b \sim 10^\circ$, toward a Galactocentric radius of $R_{\text{gc}} \sim 4$ and 8 kpc, and most significantly, toward a zero Galactic vertical scale height, z . Third, the averaged b and z for filaments are closer to zero than those of the BGPS sources. In another word, the filaments show a more symmetric distribution (than BGPS sources) with respect to the physical Galactic mid-plane.

Are filaments closer to the mid-plane more likely to align with the mid-plane, i.e., θ approaches 0° ? The data do not indicate so (Figure 3): the Pearson’s product-moment correlation coefficient between $|z|$ and $|\theta|$ is $C_{\text{Pearson}} = +0.26$ or -0.35 for the filaments and bones, respectively, far from a linear correlation ($+1$ or -1).

The statistical trends observed in these filaments provide excellent targets for quantitative tests with future theoretical calculations and numerical simulations.

5. DISCUSSION

5.1. Comparison to Previously Known Filaments

Our MST algorithm finds some filaments previously identified by other methods. Our search field (Section 2) partially overlaps with previous searches by Ragan et al. (2014), Wang et al. (2015), Zucker et al. (2015), and Li et al. (2016).

Of the 9 cold and dense prominent filaments presented by Wang et al. (2015), 7 are in our search field, 3 of which are identified by MST (F7, F25, and F41 correspond to G11, G24, and G49, respectively). Others are not identified because of a lack of dense BGPS sources (G29, G47, and G64) or too large of a disruption in velocity space due to active star formation (G28).

Among the 10 “bone” candidates presented by Zucker et al. (2015), 6 are in our search field, 2 of which have dense BGPS sources: BC011.13–0.12 and BC024.95–0.17. The former corresponds to F7 (the Snake), and the latter is not identified by MST because of too large of a velocity disruption. In addition, our MST filament F28 is visible in their Figure 13 and seems to fulfill all their criteria, but was not identified by Zucker et al. (2015).

Among the 7 giant molecular filaments presented by Ragan et al. (2014), 6 are partially covered in our field (“partially” because most of those filaments extend beyond our coverage in $|b|$). F36 is a small dense part of GMF38.1–32.4a, but note that Ragan et al. (2014) used a kinematic distance of 3.3–3.7 kpc, a factor of 2 larger than the parallax distance of 1.56 kpc. F19 and F38 fall in the positional coverage of GMF20.0–17.9 and GMF41.0–41.3, respectively, but outside the velocity ranges.

Li et al. (2016) identified 20 filaments longer than 10 pc in our searching field using ATLASGAL continuum data, 6 of which partially match with our filaments. F7, F25, and F32 are part of G011.046–0.069, G023.985+0.479, and

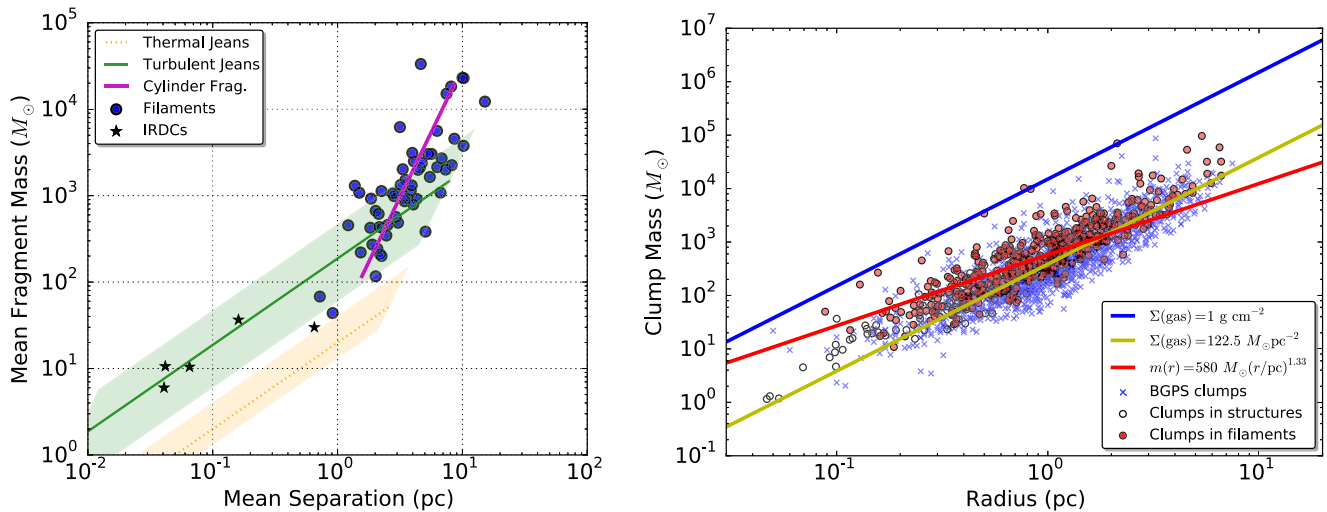


Figure 5. Left: mean clump mass vs. mean edge length for the 54 filaments, as compared to IRDCs and theoretical predictions. The magenta line is *not* a fit to the data points; instead, it depicts cylindrical fragmentation appropriate for the filaments (see the text). In comparison, the resolved fragmentation of IRDCs (data from Wang et al. 2011, 2012, 2014; Zhang & Wang 2011) is consistent with turbulent Jeans fragmentation. The green line, orange line, and associated shaded regions correspond to a range of density and temperature appropriate for IRDCs (Wang et al. 2014). Right: clump mass vs. radius for BGPS clumps, clumps in large filaments, and clumps in other velocity-coherent structures. The color lines depict various empirical criteria of star formation: blue is the Krumholz & McKee (2008) criterion of massive star formation; yellow is the average of Heiderman et al. (2010) and Lada et al. (2010) for “efficient” star formation; red is the Kauffmann & Pillai (2010) criterion of massive star formation with a correction of the adopted dust opacity law as in Dunham et al. (2011).

G030.315-0.154, respectively. The IR-dark part of F35 matches with G033.685-0.020. F33 and F37 match with G032.060+0.082 and G037.410-0.070, respectively. In the cases of F7, F25, and F32, Li et al. (2016) cataloged a longer angular extent than our filaments. However, the extra extent is not coherent in velocity.

In addition, F33 is the dense part of the “massive molecular filament” G32.02+0.06 presented by Battersby & Bally (2014) in a case study. F13 is part of the IRDC G14.225-0.506 (Busquet et al. 2013). F36 contains IRDC G034.43+00.24 (Kurayama et al. 2011; Sakai et al. 2015), and the full extent of the filament has also been reported recently by Xu et al. (2016).

F31 runs across a well-studied IRDC G28.34+0.06, also known as the “Dragon” nebula (Wang et al. 2011, 2012; Wang 2015). The IRDC is the IR-dark and submillimeter-bright arc, bent toward the bottom of the panel in Figure 6. The MST filament F31 is a new filament that runs across the IRDC at P1, where a protocluster is forming (Wang et al. 2011, 2012; Zhang et al. 2015). Interestingly, the clump-scale magnetic fields (Wang et al. 2012) are aligned with F31. At scales of the order of 10 pc, magnetic fields may be shaped by gravity, while on smaller scales (within 1 pc, or clump-scale), the magnetic fields control the formation of a secondary filament, as interpreted in Wang et al. (2012). The secondary filament is a small part of F31. Dust polarization observations of these filaments are needed to further investigate the role of magnetic fields on the formation and evolution of these filaments.

In summary, our MST method successfully finds previously known filaments where the criteria are satisfied. An important difference between the MST identified filaments and others is that the former contains dense clumps over the *whole* extent, while this is not the case for previously “by-eye” identified large filaments (“gaps” in velocity space are allowed). The MST method also finds filaments embedded in a crowded PPV space, which are difficult to isolate with human eyes (e.g., F31).

It is noteworthy that, because of the filamentary and hierarchical nature of the ISM, one can find an arbitrary number of filaments in the same data set using different criteria. Therefore, when presenting a filament sample, it is *equally important* to explicitly list the criteria used to define filaments. For the same reason, when comparing different samples of filaments one has to note the difference in criteria, otherwise the comparison is misleading.

5.2. Completeness and Bias

The 54 filaments form the first comprehensive sample of large-scale velocity-coherent gas structures in the northern Galactic plane covered by the BGPS spectroscopic survey. The homogeneous sample allows us to investigate statistical trends (Section 4.3) for the first time. With length in the range of 10–276 pc and average column density above 10^{21} cm^{-2} , these filaments are among the densest and largest structures observed in the Galaxy, and provide excellent tracers for Galactic structure and kinematics (e.g., Englmaier & Gerhard 1999; Dame et al. 2001; Dobbs et al. 2012; Reid et al. 2014; Smith et al. 2016; Vallée 2016).

The completeness of our filament sample largely depends on the data we use. The BGPS continuum catalog is 98% complete at the 0.4 Jy level (Rosolowsky et al. 2010). The spectroscopic catalog (Shirley et al. 2013) contains 50% of the sources in the survey coverage with dense gas lines detected. As our identification used the spectroscopic catalog, the results are biased to dense clumps. This is evident in the high averaged column density and high linear mass density. However, we emphasize that this is indeed our goal—we are interested in the most prominent dense filaments.

On the other hand, given the location of our Sun in the Galaxy, even homogeneous surveys like BGPS or ATLAS-GAL are biased to structures (a) closer to the Sun and (b) on the same side with respect to the mid-plane as the Sun (Figure 2; Schuller et al. 2009, but see discussion in Rosolowsky et al.

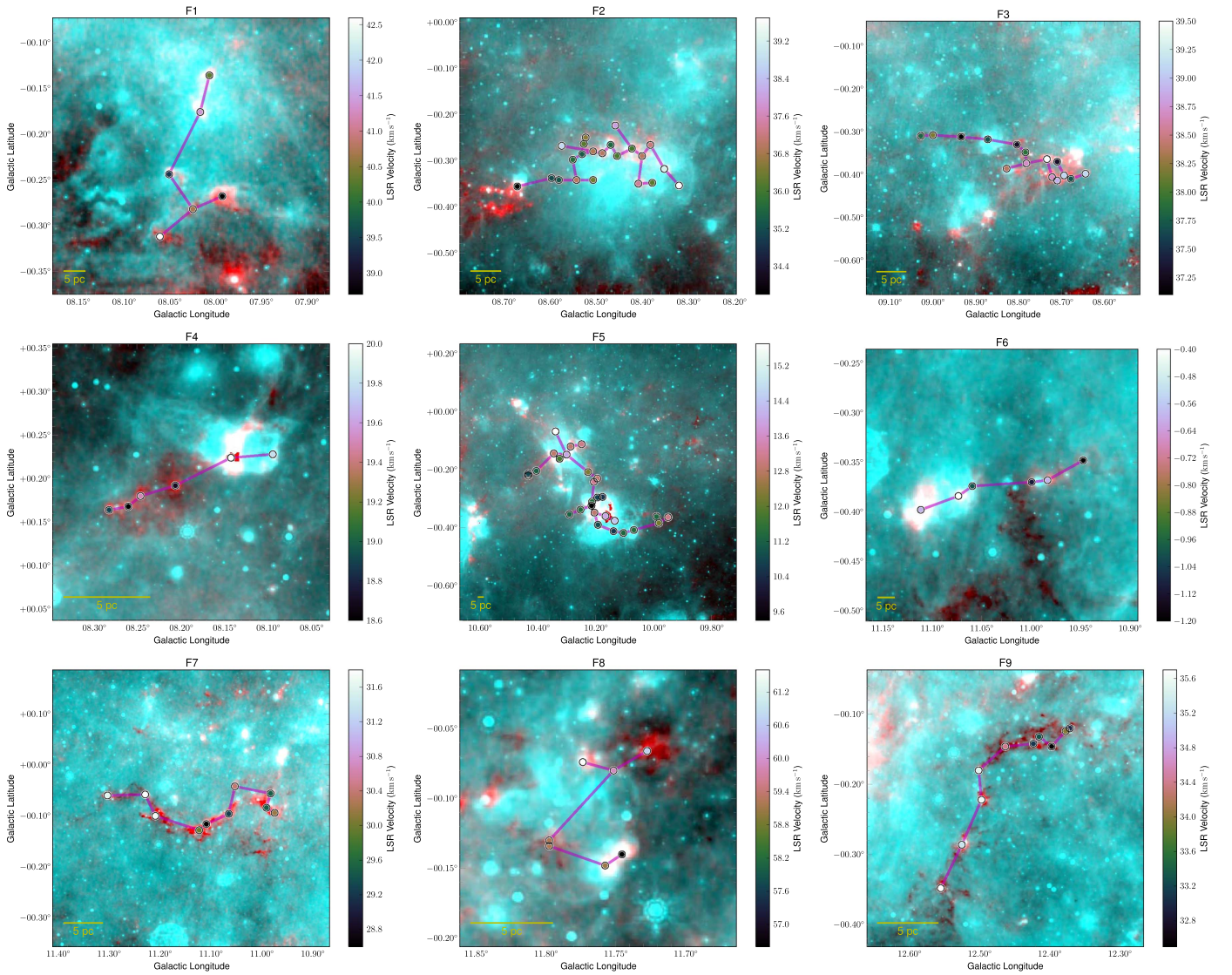


Figure 6. Two-color view of the filaments. For F1–F42, cyan shows the *Spitzer* 24 μm emission in logarithmic scale (from the MIPS GAL survey, Carey et al. 2009), and red shows the APEX+*Planck* 0.87 mm emission in linear scale (from the ATLASGAL survey, Csengeri et al. 2016). For F43–F54, cyan is *WISE* 22 μm emission (Wright et al. 2010) and red is the BGPS 1.1 mm emission (Ginsburg et al. 2013). The filled circles represent dense BGPS sources with color-coded velocity as shown in the color bar. The sources are connected by MST edges (see the text). A scale bar of 5 pc is shown for reference.

2010). As mentioned in Section 4.3, the large filaments show less bias than BGPS sources in z , but the distribution of z is indeed not perfectly symmetric. Should we mirror the distribution of $z > 0$ to $z < 0$, the total number of filaments would increase to 74. That is, 27% of the filaments may be missed due to this effect.

Although it has a much more improved completeness compared to previous methods, the MST approach cannot find all large filaments in our Galaxy. One of the main strengths of this method is the repeatability compared to manual approaches.

5.3. Galactic Distribution and Number of Large Filaments and “Bones” in the Galaxy

Most of the filaments are associated with major spiral arms (Figure 4), consistent with the observations by Wang et al. (2015). Many of them concentrate along the longitude–velocity tracks of the Scutum, Sagittarius, and Norma arms, and a few are associated with the Local arm, the Perseus arm, and one is

associated with the Outer arm.¹² Only a small fraction (11/54, or 20%) of the filaments are not within $\pm 5 \text{ km s}^{-1}$ of any arm structure, and are analogs of “spurs” observed in other galaxies.

How many large filaments exist in our Galaxy? Using our method, we have identified 48 filaments in the contiguous coverage of $7^{\circ}.5 \leq l \leq 90^{\circ}.5$. It is reasonable to estimate a similar number of filaments in the fourth quadrant. In the outer Galaxy, the BGPS survey is targeted to several star formation regions, therefore the 6 identified filaments provide an extremely lower limit. Taking all these into account, and correcting for the bias as discussed in Section 5.2, we estimate about 200 velocity-coherent filaments longer than 10 pc and with a global column density above 10^{21} cm^{-2} , as the filaments presented in this study.

¹² In the $l - v$ view (Figure 4), most filaments follow the spiral arms, while the association is less evident in the face-on view. This originates from the difference in the distance determination methods for the arm segments (parallax measurements) and the filaments (mainly kinematic distance). The same is seen in, e.g., Abreu-Vicente et al. (2016).

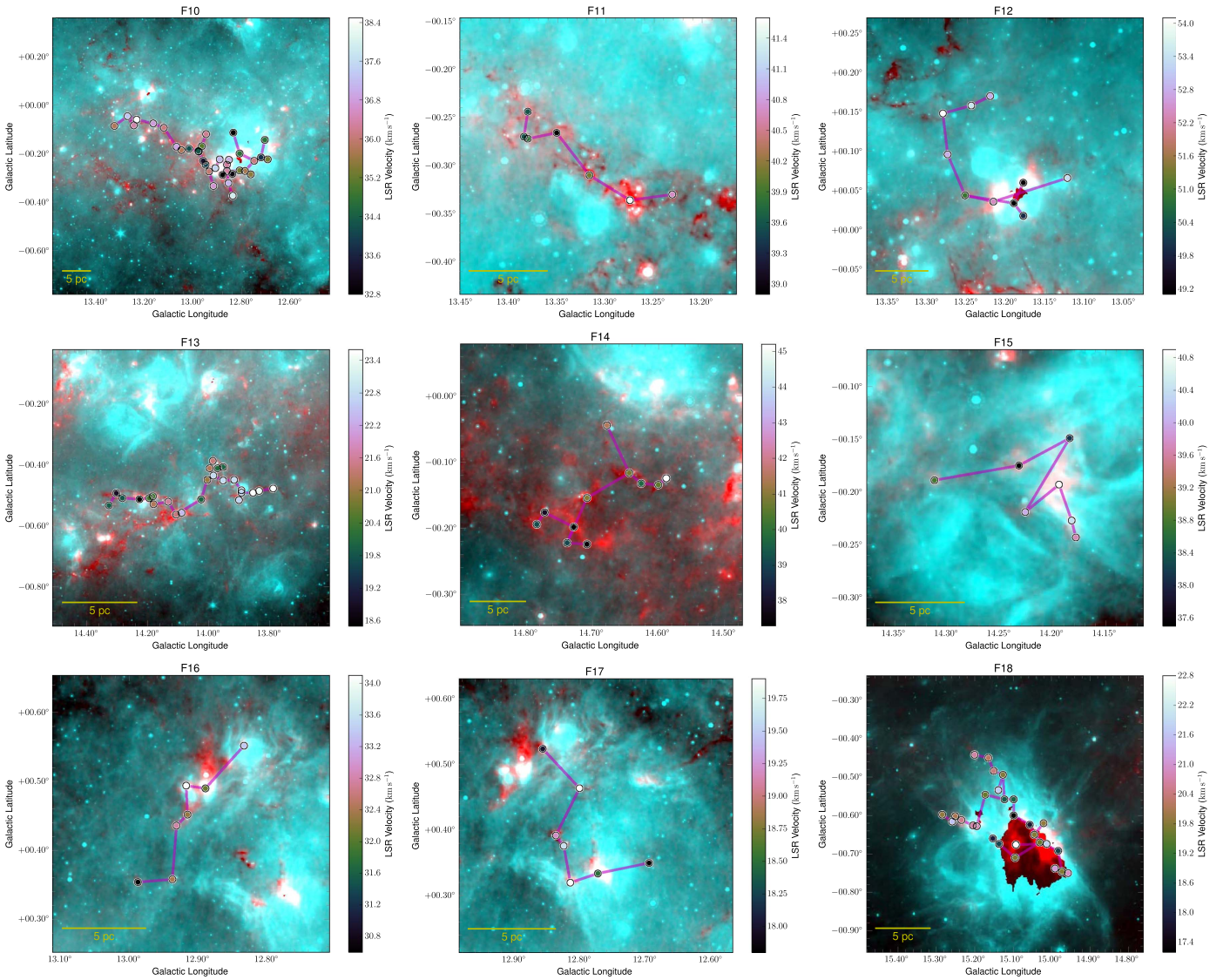


Figure 6. (Continued.)

How many filaments lie in the center of spiral arms and thus sketch out the “bones” of the Milky Way? Goodman et al. (2014) argued that the long and skinny IRDC “Nessie” lies in the center of the Scutum–Centaurus spiral arm in the (l, v) space, and within $z = \pm 20$ pc from the physical mid-plane. Following Goodman et al. (2014) and Zucker et al. (2015), our criteria for a “bone,” on top of our large-filaments criteria (1–5) (Section 3), are:

- (6) Lies in the very center of the physical Galactic mid-plane, with $|z| \leq 20$ pc.
- (7) Runs almost parallel to arms in the projected sky, with $|\theta| \leq 30^\circ$.
- (8) The flux-weighted LSR velocity v_{wt} is within ± 5 km s⁻¹ from spiral arms.

However, the exact structure and position of the spiral arms in our Galaxy are not well-established. Diverse models have been derived from a variety of data ranging from atomic, molecular, ionized gas to stars and pulsars (e.g., Hou & Han 2014; Reid & Honma 2014; Vallée 2015, 2016). Here we have adopted the spiral segments derived from maser parallaxes (Reid et al. 2014; M. Reid et al. 2015, private communication), which have

constrained distances. In Figure 4 we superpose the filaments on the spiral segments. Among the 54 filaments, 27 fulfill criteria (1–6), 21 fulfill criteria (1–7), and 13 of them also fulfill criterion (8). These 13 filaments (F2, F3, F7, F10, F13–F15, F18, F28, F29, F37, F38, and F48) are “bones” according to our definition. Our criteria for a bone are more strict than Zucker et al. (2015) in terms of velocity coherence and mean column density. When compared to other filaments, bones do not stand out in mass, length (Figures 2(a) and (b)), column/volume density, or temperature. All 13 bones are located in the first quadrant (which is not surprising given our search field; see Section 2), making 27% of the 48 filaments in the same region of blind BGPS survey. Obviously, owing to disagreement among the many spiral arm models, adopting a different model will lead to different “bones.” But the fraction of bones in filaments should not change dramatically in a reasonable spiral model.

5.4. Fraction of ISM Confined in Large Filaments

Given the importance of filamentary geometry for enhancing massive clustered star formation, it is of great interest to

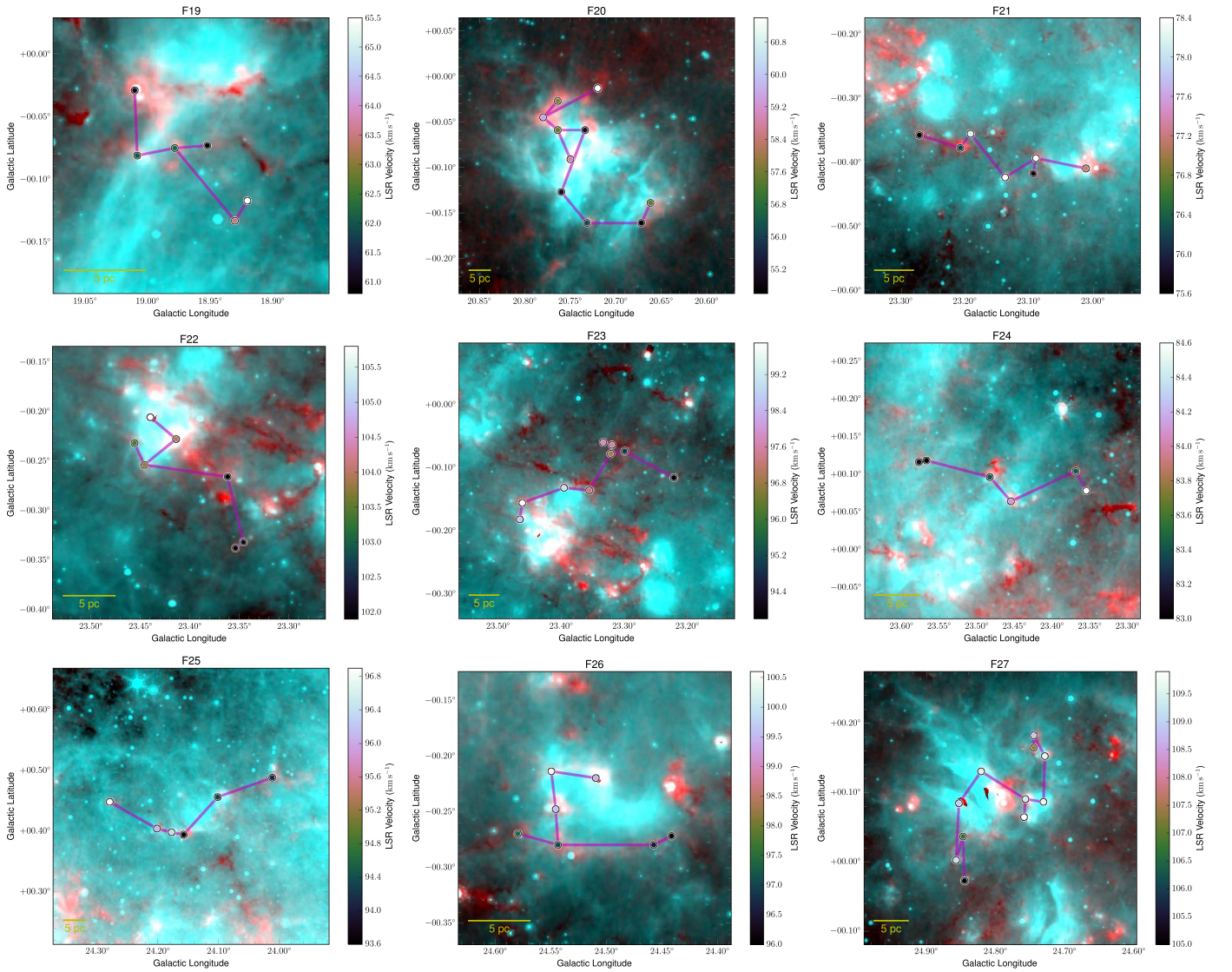


Figure 6. (Continued.)

quantify the fraction of the ISM contained in large filaments and to evaluate the star formation activities therein.

To address this question we consider only the range of $7.5 \leq l \leq 90.5$ where the BGPS and its spectroscopic follow-up are contiguous. In this field there are 5841 BGPS v1 sources, 2893 having HCO⁺ or N₂H⁺(3–2) detection, which we refer to as “dense BGPS sources.” We identified 48 filaments in this field, which are comprised of 512 BGPS sources. That means that 17.7% (512/2893) of dense BGPS sources, or 8.8% (512/5841) of all BGPS sources, are confined in large filaments. If we count BGPS clumps in the bones only, 6.8% of dense BGPS sources, or 3.4% of all BGPS sources, are confined in bones.

Compact 1.1 mm continuum emission of the BGPS sources outline the dense, inner part of much larger and less dense envelopes of molecular clouds (Dunham et al. 2011). Assuming a dense gas mass fraction of 10%–20% (Ragan et al. 2014; Ginsburg et al. 2015), we infer an order of 1% of the molecular ISM is confined in large filaments, and about 1/3 of this amount is confined in bones—marking spiral arm centers.

5.5. An Apparent Length Limit of 100 pc and the Longest Filaments beyond this Limit

Wang et al. (2015) pointed out an apparent upper limit of 100 pc projected length for the longest filaments in their study designed to find cold and dense filaments based on *Herschel* far-IR emission. In this study, we use a different approach without limiting the temperature. Except for the extremely long filament F5, all other filaments are indeed shorter than 100 pc. Interestingly, this limit is also seen in Zucker et al. (2015) despite the different search method. The 100 pc limit seems to be present in filaments with a global column density above the order of 10^{21} – 10^{22} cm⁻². Relaxing the column density to a lower cut of $<10^{20}$ cm⁻², longer filaments start to be picked up in ¹²CO/¹³CO (1–0): the 430 pc “optimistic Nessie” (Goodman et al. 2014); the 500 pc “wisp” (Li et al. 2013), and a few filaments by Ragan et al. (2014) and Abreu-Vicente et al. (2016). However, those CO filaments have much smaller aspect ratios (typically $\ll 10$, see figures in their papers), and the low-*J* CO gas outlines the relatively diffuse envelopes of denser

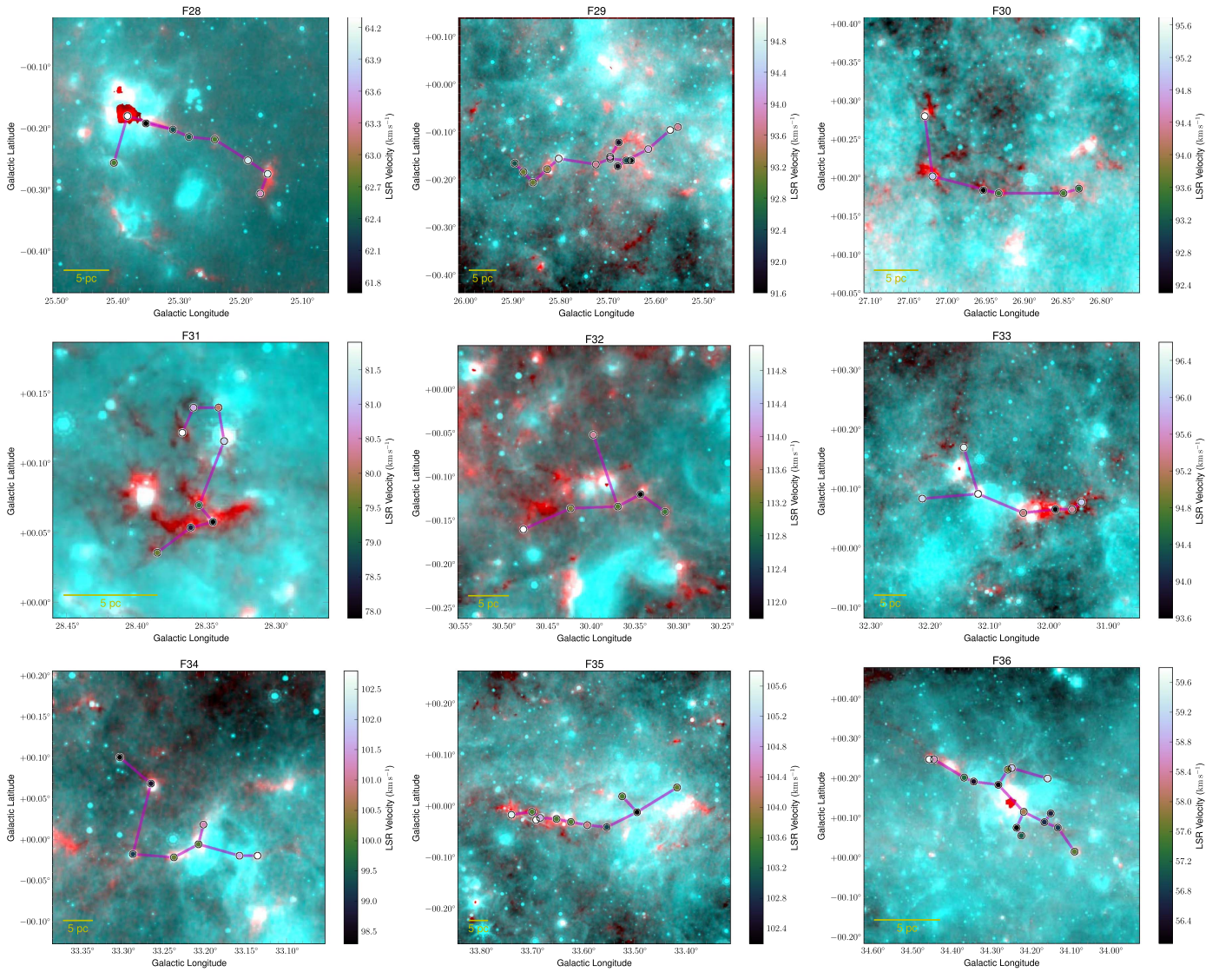


Figure 6. (Continued.)

structures traced by MIR extinction. For example, one filament reported by Ragan et al. (2014) contains F36 as a small part (Section 5.1). Whether 100 pc is a true limit for dense filaments warrants further study. This provides a quantitative test case for numerical simulations (e.g., Falceta-Gonçalves et al. 2015).

So far, the 276 pc-long filament F5 (Figure 6) is the only exception of dense ($>10^{21}$ cm⁻²) filaments longer than 100 pc. Compared to the above mentioned extremely long CO filaments, F5 is at least 10 times denser in average column density, and it may also have less dense extensions, similar to the 80 pc “classic Nessie.” More systematic searches and comparison to numerical simulations can resolve the true length limit of the longest filaments. We emphasize that the average column density is a crucial parameter for defining the boundary of filaments and thus the length and aspect ratio. It is also worthwhile to note that our filaments, as defined by a collection of dense BGPS clumps, form the center of larger and less dense structures.

The origin of large-scale velocity-coherent filaments is still a mystery. Numerical simulations of the multiphase ISM in galactic disks demonstrate that the cold, dense gas component

tends to organize itself naturally into a filamentary network (e.g., Tasker & Tan 2009; Smith et al. 2014). Spiral arms can sweep up and compress gas, generating bones (Goodman et al. 2014). Gravitationally unstable disk regions condense into gaseous rings and arcs (Behrendt et al. 2015). In differentially rotating disks, structures like molecular cloud complexes could be sheared into elongated filaments.

Our calculations (A. Burkert et al. 2016, in preparation) show that tidal effects of the Milky Way are too weak to affect the maximum length of filaments. This is consistent with our observations, where we find filament lengths do not correlate with Galactocentric radii ($C_{\text{Pearson}} = -0.07$). The maximum filament length of the order of 100 pc might be related to the timescale of $\tau_{\text{SF}} = 10^7$ years (e.g., Burkert & Hartmann 2013) on which stars form in dense molecular gas and destroy their environments. Typical turbulent velocities on large scales in galactic disks are of the order of $\sigma = 10$ km s⁻¹ (Dib et al. 2006). If σ is the maximum velocity with which coherent filaments can grow and if τ_{SF} denotes the timescale on which they are destroyed again, their length is limited by $l = \sigma \times \tau_{\text{SF}} = 100$ pc, in agreement with the observations.

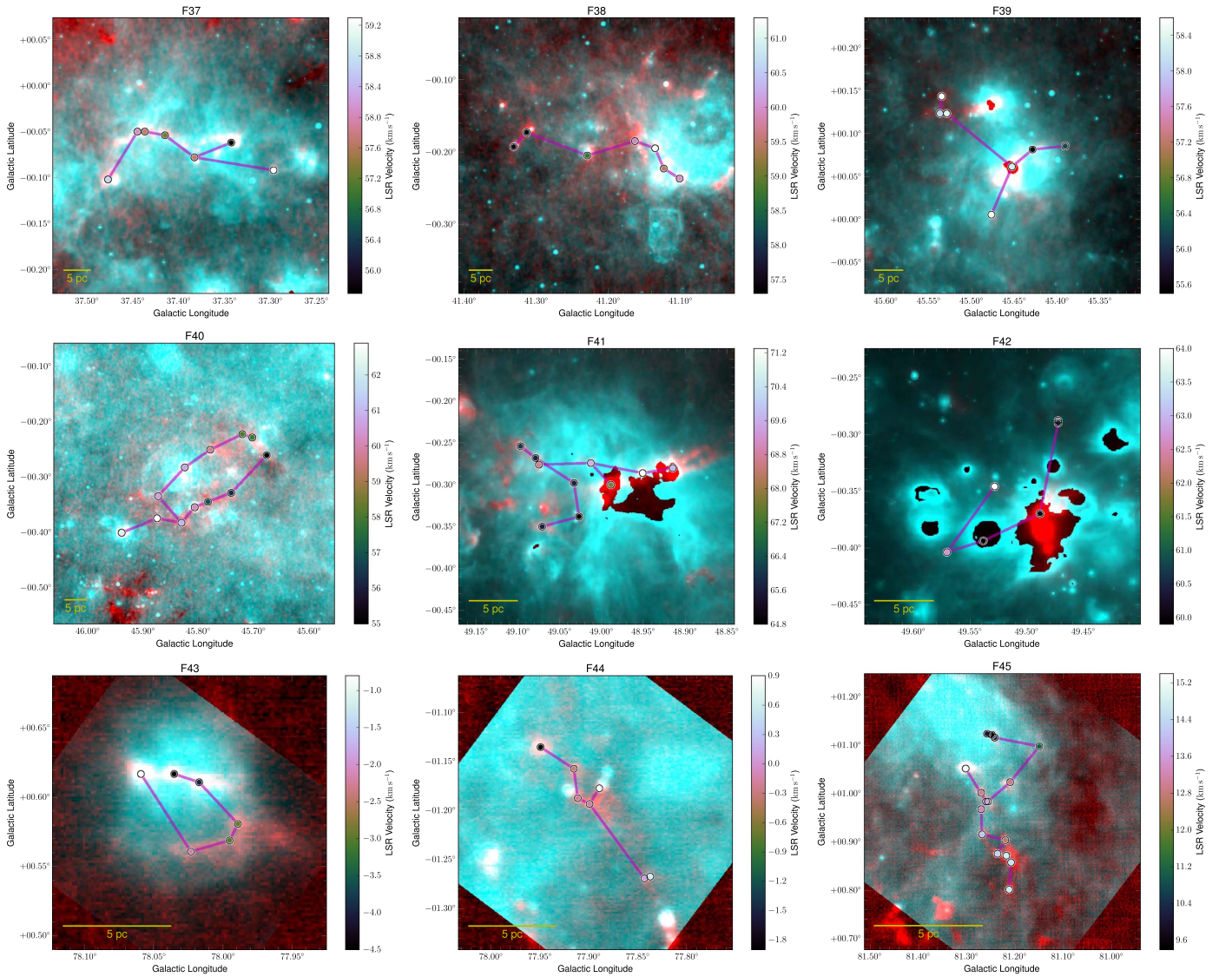


Figure 6. (Continued.)

5.6. Fragmentation of Large-scale Filaments and Subsequent Star Formation

By definition, the filaments presented in this study are in the form of a chain of dense clumps physically connected by less dense gas in between. For linear filaments, this geometry resembles a fragmented “cylinder” with regularly spaced clumps under the “sausage instability” of self-gravity (e.g., F9 in Figure 6). According to the framework of Chandrasekhar & Fermi (1953), an isothermal gas cylinder becomes supercritical when its linear mass density exceeds the critical value $(M/L)_{\text{crit}}$, and will fragment into a chain of equally spaced fragments with a spacing of λ_{cl} , with each fragment having a mass of $M_{\text{cl}} = (M/L)_{\text{crit}} \times \lambda_{\text{cl}}$. In short, the fragmentation is governed by central density and pressure (thermal plus non-thermal). This framework has been followed by many authors (e.g., Ostriker 1964; Nagasawa 1987; Bastien et al. 1991; Inutsuka & Miyama 1992; Fischera & Martin 2012). See Wang et al. (2011, 2014) for a useful deduction of the formulas. Figure 5 (left panel) plots the mean clump mass of each

filament with the mean separation between clumps (mean length of the edges in the filament). The observed fragmentation is consistent with the theoretical prediction of cylindrical fragmentation, assuming a central density of $1 \times 10^4 \text{ cm}^{-3}$ and a velocity dispersion of $0.4\text{--}2.2 \text{ km s}^{-1}$ (magenta line). The spread of the data points around the prediction line may be due to a range of central densities and imperfect cylinder geometry. Recent numerical simulations have shown that geometric bending, which is often seen in observed filaments, can change the regularity of the spacing (Gritschneider et al. 2016), indicating that more theoretical work is required in order to understand the stability and dynamics of filaments.

Dense clumps with a typical mass of $10^3 M_{\odot}$ and typical size of 1 pc (Figure 5, left panel) are, in general, capable of forming a cluster of stars. Statistically, dense clumps residing within large filaments are slightly denser than clumps elsewhere (see below). In Figure 5 (right panel), we plot clump mass versus deconvolved radius (not all BGPS clumps have a valid radius, see Section 4)

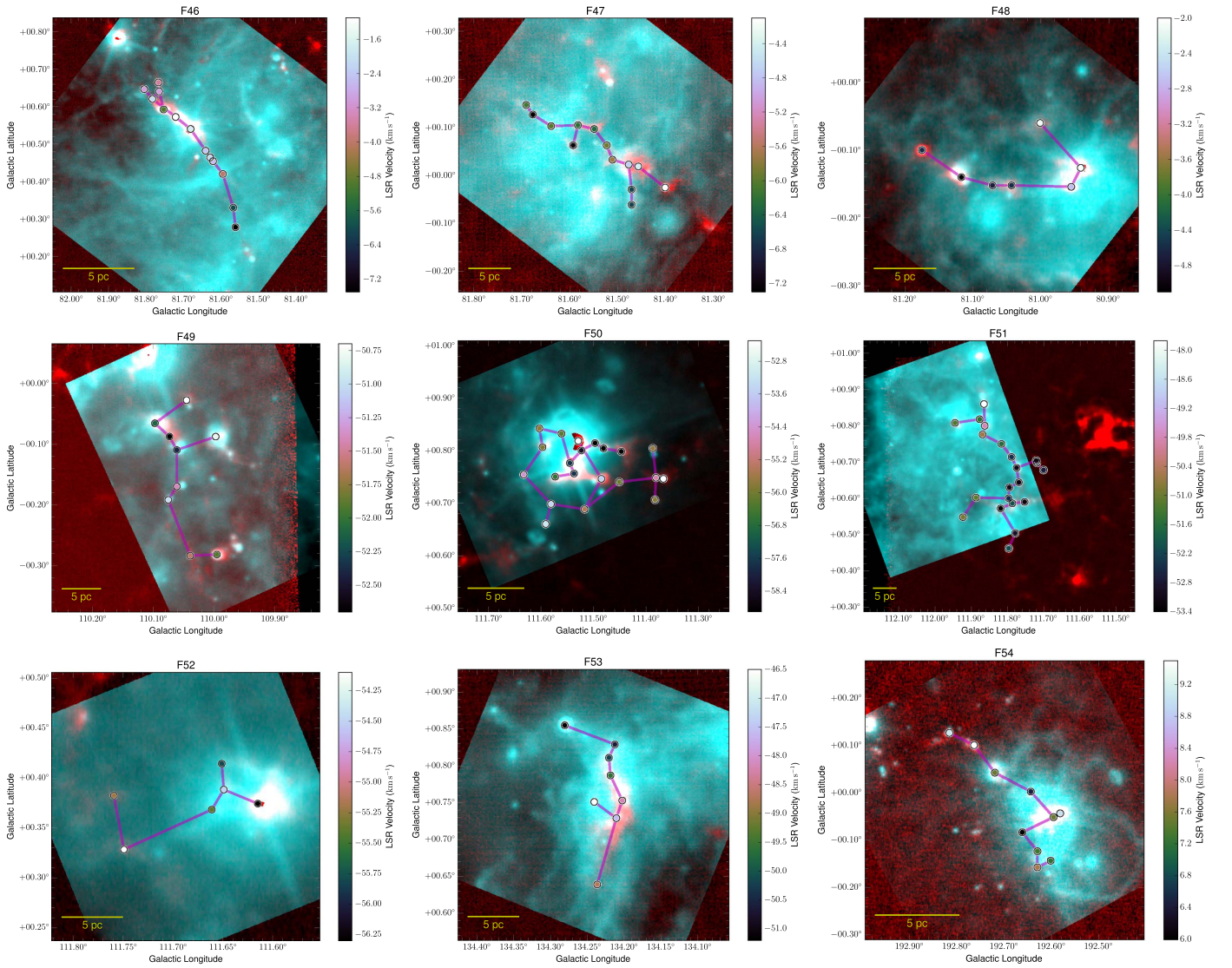


Figure 6. (Continued.)

for three categories of BGPS clumps: I—the 1710 clumps with well determined distance from Ellsworth-Bowers et al. (2015). II—the 294 clumps in velocity-coherent structures but not large filaments. III—the 469 clumps in large filaments. Categories I, II, and III have 41.0%, 39.8%, and 46.2% clumps, respectively, satisfying the Kauffmann & Pillai (2010) threshold of forming massive stars. Thus, categories I and II are indistinguishable, while in comparison, category III is slightly more favorable for massive star formation. If we count in mass instead of number of clumps, the fractions are 79.2% for BGPS sources, 86.3% for velocity-coherent structures but not large filaments, and 91.0% for large filaments. Surprisingly, bones do not show a higher fraction compared to large filaments, either counted in number or mass. This indicates that local environment such as a velocity-coherent filament plays a role in enhancing massive star formation. Filaments, in particular, provide a preferred form of geometry to channel mass flows that can inhomogeneously feed star-forming clumps (e.g., Arzoumanian et al. 2011; Peretto et al. 2014; Zhang et al. 2015; Federrath 2016; Heigl et al. 2016). On the other hand, the Galactic environment does not seem to affect local star formation across the few hundred parsecs spread of vertical

position z , consistent with previous studies (Eden et al. 2012, 2013).

6. SUMMARY

We present an automated method designed to identify large-scale velocity-coherent filaments throughout the Galaxy. The method utilizes a customized MST algorithm, which connects neighboring voxels in the PPV space. We have applied the algorithm to the BGPS spectroscopic catalog in the range $7.5 \leq l \leq 194^\circ$, $|b| < 0.5^\circ$. We have identified a comprehensive catalog of 54 large-scale filaments and derived physical parameters including mass ($\sim 10^3$ – $10^5 M_\odot$), length (10–276 pc), linear mass density (54 – $8625 M_\odot \text{pc}^{-1}$), aspect ratio (18–176), linearity, velocity gradient, temperature, fragmentation, Galactic location, and orientation angle. We investigate the Galactic distribution of these parameters and compare the filaments with an updated Galactic spiral arm model.

1. In general, the Galactic distribution of large filaments follows the dense gas traced by BGPS sources. However,

- there are interesting exceptions to be further explored by a quantitative comparison to theoretical work (Section 4.3).
2. Most of the filaments are associated with major spiral arms, including the Scutum, Sagittarius, and Norma arms, and a few associated with the Local, Perseus, and Outer arms. About 20% of the filaments are inter-arm structures, or “spurs” (Section 5.3). The filaments tend to align with Galactic plane, but the tendency does not correlate with vertical height (Figure 3).
 3. The filaments are widely distributed across the Galactic disk, with 50% located within ± 20 pc from the Galactic mid-plane and 27% runs in the center of major spiral arms, forming the “bones” of the Milky Way (Section 5.3).
 4. An order of 1% of the molecular ISM is confined in large filaments, and about 1/3 of this amount is confined in bones—marking spiral arm centers (Section 5.4).
 5. Massive star formation is more favorable in large filaments compared to elsewhere. However, Galactic environment is not observed to affect local star formation (Section 5.6).
 6. An apparent length limit of 100 pc is observed for filaments with a global column density N_{H_2} higher than 10^{21} cm^{-2} (or optical extinction $A_V \approx 1$ mag). Longer filaments are rarer, with a much lower aspect ratio, and have at least one order of magnitude lower global column density (Section 5.5).

Our method can be applied to three-dimensional PPV catalogs from observations or PPP catalogs from simulations. This study focuses on the northern Galactic plane covered by the BGPS. In the near future, the two-dimensional ATLAS-GAL (inner Galactic plane; Csengeri et al. 2014) and Hi-GAL (full Galactic plane; Molinari et al. 2016) catalogs will be complemented with velocity information from currently ongoing spectral line surveys (SEDIGISM, Schuller et al. 2016; MWISP: Milky Way Imaging Scroll Painting, Jiang & Li 2013; Sun et al. 2015; Mopra Southern Galactic Plane CO Survey, Burton et al. 2013; and ThrUMMS, Barnes et al. 2015). By then, we can complete the census of large-scale, velocity filaments in the full Galactic plane.

As increasing numbers of filaments with various morphologies are published in the literature, we urge our colleagues to explicitly list the criteria used to define their filaments before a commonly agreed, physics driven, definition of filaments can be reached in the community. The final definition will likely be scale-dependent, given the different physics behind large-scale and smaller-scale filaments in the ISM.

We are grateful to Mark Reid and Tom Dame for providing the spiral arm segments, and Alberto Sanna for the parallax distance to maser G11.11-0.11, before publication. We thank the anonymous referee for a constructive review report, and Eric Feigelson for comments on the statistical aspects of the paper. K.W. is supported by grant WA3628-1/1 of the German Research Foundation (DFG) through the priority program 1573 (“Physics of the Interstellar Medium”). The color bars in the figures utilize the CUBEHELIX color scheme introduced by Green (2011). This research made use of ASTROPY, a community-developed core Python package for Astronomy (Astropy Collaboration et al. 2013). This publication makes use of data products from the Wide-field Infrared Survey Explorer, which is a joint project of the University of California, Los Angeles, and the Jet Propulsion Laboratory/California Institute

of Technology, funded by the National Aeronautics and Space Administration.

Facilities: CSO, SMT, *Spitzer*, *WISE*, *Planck*.

Software: PYTHON, ASTROPY.

REFERENCES

- Abreu-Vicente, J., Ragan, S., Kainulainen, J., et al. 2016, *A&A*, 590, A131
- Aguirre, J. E., Ginsburg, A. G., Dunham, M. K., et al. 2011, *ApJS*, 192, 4
- Arzoumanian, D., André, P., Didelon, P., et al. 2011, *A&A*, 529, L6
- Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, *A&A*, 558, A33
- Barnes, P. J., Muller, E., Indermuehle, B., et al. 2015, *ApJ*, 812, 6
- Bastien, P., Arcoragi, J.-P., Benz, W., Bonnell, I., & Martel, H. 1991, *ApJ*, 378, 255
- Battersby, C., & Bally, J. 2014, *ASSP*, 36, 417
- Behrendt, M., Burkert, A., & Schartmann, M. 2015, *MNRAS*, 448, 1007
- Borůvka, O. 1926a, *Práce mor. přírodověd. spol. v Brně III* (in Czech), 3, 37
- Borůvka, O. 1926b, *Elektronický Obzor* (in Czech), 15, 153
- Brunthaler, A., Reid, M. J., Menten, K. M., et al. 2009, *ApJ*, 693, 424
- Brunthaler, A., Reid, M. J., Menten, K. M., et al. 2011, *AN*, 332, 461
- Burkert, A., & Hartmann, L. 2013, *ApJ*, 773, 48
- Burton, M. G., Braiding, C., Glueck, C., et al. 2013, *PASA*, 30, e044
- Busquet, G., Zhang, Q., Palau, A., et al. 2013, *ApJL*, 764, L26
- Carey, S. J., Clark, F. O., Egan, M. P., et al. 1998, *ApJ*, 508, 721
- Carey, S. J., Feldman, P. A., Redman, R. O., et al. 2000, *ApJL*, 543, L157
- Carey, S. J., Noriega-Crespo, A., Mizuno, D. R., et al. 2009, *PASP*, 121, 76
- Chandrasekhar, S., & Fermi, E. 1953, *ApJ*, 118, 116
- Csengeri, T., Urquhart, J. S., Schuller, F., et al. 2014, *A&A*, 565, A75
- Csengeri, T., Weiss, A., Wyrowski, F., et al. 2016, *A&A*, 585, A104
- Dame, T. M., Hartmann, D., & Thaddeus, P. 2001, *ApJ*, 547, 792
- Dib, S., Bell, E., & Burkert, A. 2006, *ApJ*, 638, 797
- Dobbs, C. L., Pringle, J. E., & Burkert, A. 2012, *MNRAS*, 425, 2157
- Duarte-Cabral, A., & Dobbs, C. L. 2016, *MNRAS*, 458, 3667
- Dunham, M. K., Rosolowsky, E., Evans, N. J., II, Cyganowski, C., & Urquhart, J. S. 2011, *ApJ*, 741, 110
- Eden, D. J., Moore, T. J. T., Morgan, L. K., Thompson, M. A., & Urquhart, J. S. 2013, *MNRAS*, 431, 1587
- Eden, D. J., Moore, T. J. T., Plume, R., & Morgan, L. K. 2012, *MNRAS*, 422, 3178
- Ellsworth-Bowers, T. P., Glenn, J., Rosolowsky, E., et al. 2013, *ApJ*, 770, 39
- Ellsworth-Bowers, T. P., Rosolowsky, E., Glenn, J., et al. 2015, *ApJ*, 799, 29
- Englmaier, P., & Gerhard, O. 1999, *MNRAS*, 304, 512
- Falceta-Gonçalves, D., Bonnell, I., Kowal, G., Lépine, J. R. D., & Braga, C. A. S. 2015, *MNRAS*, 446, 973
- Federrath, C. 2016, *MNRAS*, 457, 375
- Federrath, C., Klessen, R. S., & Schmidt, W. 2009, *ApJ*, 692, 364
- Fischera, J., & Martin, P. G. 2012, *A&A*, 542, A77
- Ginsburg, A., Bally, J., Battersby, C., et al. 2015, *A&A*, 573, A106
- Ginsburg, A., Glenn, J., Rosolowsky, E., et al. 2013, *ApJS*, 208, 14
- Goodman, A. A., Alves, J., Beaumont, C. N., et al. 2014, *ApJ*, 797, 53
- Green, D. A. 2011, *BASI*, 39, 289
- Gritschneider, M., Heigl, S., & Burkert, A. 2016, arXiv:1604.00378
- Heiderman, A., Evans, N. J., II, Allen, L. E., Huard, T., & Heyer, M. 2010, *ApJ*, 723, 1019
- Heigl, S., Burkert, A., & Hacar, A. 2016, arXiv:1601.02018
- Hou, L. G., & Han, J. L. 2014, *A&A*, 569, A125
- Immer, K., Reid, M. J., Menten, K. M., Brunthaler, A., & Dame, T. M. 2013, *A&A*, 553, A117
- Inutsuka, S.-I., & Miyama, S. M. 1992, *ApJ*, 388, 392
- Jackson, J. M., Finn, S. C., Chambers, E. T., Rathborne, J. M., & Simon, R. 2010, *ApJL*, 719, L185
- Jackson, J. M., Finn, S. C., Rathborne, J. M., Chambers, E. T., & Simon, R. 2008, *ApJ*, 680, 349
- Jackson, J. M., Rathborne, J. M., Shah, R. Y., et al. 2006, *ApJS*, 163, 145
- Jiang, Z., & Li, J. 2013, in *Conf. Proc., Protostars and Planets VI* (Tucson, AZ: Univ. of Arizona Press), **IB003**
- Kauffmann, J., & Pillai, T. 2010, *ApJL*, 723, L7
- Koch, E. W., & Rosolowsky, E. W. 2015, *MNRAS*, 452, 3435
- Krumholz, M. R., & McKee, C. F. 2008, *Natur*, 451, 1082
- Kurayama, T., Nakagawa, A., Sawada-Satoh, S., et al. 2011, *PASJ*, 63, 513
- Lada, C. J., Lombardi, M., & Alves, J. F. 2010, *ApJ*, 724, 687
- Li, G.-X., Urquhart, J. S., Leurini, S., et al. 2016, *A&A*, 591, A5

- Li, G.-X., Wyrowski, F., Menten, K., & Belloche, A. 2013, *A&A*, **559**, A34
- Men'shchikov, A. 2013, *A&A*, **560**, A63
- Merello, M., Evans, N. J., II, Shirley, Y. L., et al. 2015, *ApJS*, **218**, 1
- Molinari, S., Swinyard, B., Bally, J., et al. 2010, *PASP*, **122**, 314
- Molinari, S., Schisano, E., Elia, D., et al. 2016, arXiv:1604.05911
- Moscadelli, L., Reid, M. J., Menten, K. M., et al. 2009, *ApJ*, **693**, 406
- Nagasawa, M. 1987, *PThPh*, **77**, 635
- Nešetřil, J., Milková, E., & Nešetřilová, H. 2001, *Discrete Mathematics*, **233**, 3
- Ossenkopf, V., & Henning, T. 1994, *A&A*, **291**, 943
- Ostriker, J. 1964, *ApJ*, **140**, 1056
- Panopoulou, G. V., Tassis, K., Goldsmith, P. F., & Heyer, M. H. 2014, *MNRAS*, **444**, 2507
- Peretto, N., Fuller, G. A., André, P., et al. 2014, *A&A*, **561**, A83
- Ragan, S. E., Henning, T., Tackenberg, J., et al. 2014, *A&A*, **568**, A73
- Reid, M. J., & Honma, M. 2014, *ARA&A*, **52**, 339
- Reid, M. J., Menten, K. M., Brunthaler, A., et al. 2014, *ApJ*, **783**, 130
- Reid, M. J., Menten, K. M., Zheng, X. W., et al. 2009, *ApJ*, **700**, 137
- Roman-Duval, J., Jackson, J. M., Heyer, M., Rathborne, J., & Simon, R. 2010, *ApJ*, **723**, 492
- Rosolowsky, E., Dunham, M. K., Ginsburg, A., et al. 2010, *ApJS*, **188**, 123
- Sakai, T., Sakai, N., Furuya, K., et al. 2015, *ApJ*, **803**, 70
- Salji, C. J., Richer, J. S., Buckle, J. V., et al. 2015, *MNRAS*, **449**, 1782
- Schisano, E., Rygl, K. L. J., Molinari, S., et al. 2014, *ApJ*, **791**, 27
- Schlingman, W. M., Shirley, Y. L., Schenk, D. E., et al. 2011, *ApJS*, **195**, 14
- Schuller et al. 2016, *A&A*, submitted
- Schuller, F., Menten, K. M., Contreras, Y., et al. 2009, *A&A*, **504**, 415
- Shirley, Y. L. 2015, *PASP*, **127**, 299
- Shirley, Y. L., Ellsworth-Bowers, T. P., Svoboda, B., et al. 2013, *ApJS*, **209**, 2
- Smith, R. J., Glover, S. C. O., Clark, P. C., Klessen, R. S., & Springel, V. 2014, *MNRAS*, **441**, 1628
- Smith, R. J., Glover, S. C. O., Klessen, R. S., & Fuller, G. A. 2016, *MNRAS*, **455**, 3640
- Sousbie, T. 2011, *MNRAS*, **414**, 350
- Sun, Y., Xu, Y., Yang, J., et al. 2015, *ApJL*, **798**, L27
- Tasker, E. J., & Tan, J. C. 2009, *ApJ*, **700**, 358
- Tempel, E., Stoica, R. S., Kipper, R., & Saar, E. 2016, *A&C*, **16**, 17
- Vallée, J. P. 2015, *MNRAS*, **450**, 4277
- Vallée, J. P. 2016, *AJ*, **151**, 55
- Wang, K. 2015, *The Earliest Stages of Massive Clustered Star Formation: Fragmentation of Infrared Dark Clouds*, Springer Theses (Berlin: Springer)
- Wang, K., Zhang, Q., Wu, Y., Li, H.-b., & Zhang, H. 2012, *ApJL*, **745**, L30
- Wang, K., Zhang, Q., Wu, Y., & Zhang, H. 2011, *ApJ*, **735**, 64
- Wang, K., Testi, L., Ginsburg, A., et al. 2015, *MNRAS*, **450**, 4043
- Wang, K., Zhang, Q., Testi, L., et al. 2014, *MNRAS*, **439**, 3275
- Wienen, M., Wyrowski, F., Schuller, F., et al. 2012, *A&A*, **544**, A146
- Wilcock, L. A., Ward-Thompson, D., Kirk, J. M., et al. 2012, *MNRAS*, **422**, 1071
- Wright, E. L., Eisenhardt, P. R. M., Mainzer, A. K., et al. 2010, *AJ*, **140**, 1868
- Wu, Y. W., Sato, M., Reid, M. J., et al. 2014, *A&A*, **566**, A17
- Xu, J.-L., Li, D., Zhang, C.-P., et al. 2016, *ApJ*, **819**, 117
- Xu, Y., Moscadelli, L., Reid, M. J., et al. 2011, *ApJ*, **733**, 25
- Zhang, Q., & Wang, K. 2011, *ApJ*, **733**, 26
- Zhang, Q., Wang, K., Lu, X., & Jiménez-Serra, I. 2015, *ApJ*, **804**, 141
- Zucker, C., Battersby, C., & Goodman, A. 2015, *ApJ*, **815**, 23