

GASTON: Galactic Star Formation with NIKA2 – evidence for the mass growth of star-forming clumps

A. J. Rigby¹,¹★ N. Peretto¹,¹★ R. Adam,² P. Ade,¹ M. Anderson¹,¹ P. André,³ A. Andrianasolo,⁴ H. Aussel,³ A. Bacmann⁴,⁴ A. Beelen,⁵ A. Benoît,⁶ S. Berta,⁷ O. Bourrion,⁸ A. Bracco,⁹ M. Calvo,⁶ A. Catalano,⁸ M. De Petris,¹⁰ F.-X. Désert,⁴ S. Doyle,¹ E. F. C. Driessen,⁷ P. García,^{11,12} A. Gomez,¹³ J. Goupy,⁶ F. Kéruzoré⁸,⁸ C. Kramer,^{7,14} B. Ladjelate,¹⁴ G. Lagache,¹⁵ S. Leclercq,⁷ J.-F. Lestrade,¹⁶ J. F. Macías-Pérez,⁸ P. Mauskopf,^{1,17} F. Mayet,⁸ A. Monfardini,⁶ L. Perotto,⁸ G. Pisano,¹ N. Ponthieu,⁴ V. Révère,³ I. Ristorcelli,¹⁸ A. Ritacco,^{5,19} C. Romero,²⁰ H. Roussel,²¹ F. Ruppin¹²,²² K. Schuster,⁷ S. Shu,⁷ A. Sievers,¹⁴ C. Tucker¹ and E. J. Watkins^{1,23}

Affiliations are listed at the end of the paper

Accepted 2021 January 20. Received 2021 January 20; in original form 2020 July 21

ABSTRACT

Determining the mechanism by which high-mass stars are formed is essential for our understanding of the energy budget and chemical evolution of galaxies. By using the New IRAM KIDs Array 2 (NIKA2) camera on the Institut de Radio Astronomie Millimétrique (IRAM) 30-m telescope, we have conducted high-sensitivity and large-scale mapping of a fraction of the Galactic plane (GP) in order to search for signatures of the transition between the high- and low-mass star-forming modes. Here, we present the first results from the Galactic Star Formation with NIKA2 (GASTON) project, a Large Programme at the IRAM 30-m telescope that is mapping $\approx 2 \text{ deg}^2$ of the inner GP, centred on $\ell = 23^\circ.9$, $b = 0^\circ.05$, as well as targets in Taurus and Ophiuchus in 1.15- and 2.00-mm continuum wavebands. In this paper, we present the first of the GASTON GP data taken, and present initial science results. We conduct an extraction of structures from the 1.15-mm maps using a dendrogram analysis and, by comparison to the compact source catalogues from *Herschel* survey data, we identify a population of 321 previously undetected clumps. Approximately 80 per cent of these new clumps are 70- μm -quiet, and may be considered as starless candidates. We find that this new population of clumps are less massive and cooler, on average, than clumps that have already been identified. Further, by classifying the full sample of clumps based upon their infrared-bright fraction – an indicator of evolutionary stage – we find evidence for clump mass growth, supporting models of clump-fed high-mass star formation.

Key words: surveys – stars: formation – stars: massive – ISM: evolution – ISM: structure – Galaxy: disc.

1 INTRODUCTION

The stellar initial mass function (IMF) is an essential ingredient in cosmological simulations of galaxy formation and evolution, and its origin remains one of the most fundamental and important questions in the field of astronomy. In the post-*Herschel* era, a link between filamentary structures in the interstellar medium (ISM) and star formation has been firmly established (Molinari et al. 2010; André et al. 2010, 2014). At the low-mass end of the IMF, most stars with masses in the range $m_* \sim 0.1$ to a few M_\odot are found to form within dense filaments (Könyves et al. 2015, 2020; Marsh et al. 2016). If the mass-per-unit-length exceeds a critical value set by the local sound speed, gravitational instabilities can develop, leading to the fragmentation of the filament, and the formation of stellar systems from this mass reservoir (Inutsuka & Miyama 1997; Clarke et al. 2017). This *core-fed* scheme of fragmentation of trans- or supercritical virialized filaments may explain the origin

of the IMF from ~ 0.1 to $\sim 5 M_\odot$ (Lee, Hennebelle & Chabrier 2017; André et al. 2019). However, outside this range, the proposed formation mechanisms for brown dwarfs (Whitworth et al. 2007; Luhman 2012), and the formation of high-mass ($m_* \gtrsim 8 M_\odot$) stars (Tan et al. 2014; Motte, Bontemps & Louvet 2018), remain controversial.

It is well established that the Jeans-like fragmentation of molecular clouds is unable to produce dense cores that are sufficiently massive to be the progenitors of high-mass stars (Bontemps et al. 2010; Sanhueza et al. 2017), and so the formation pathway must incorporate additional mechanisms. There are broadly two families of models for high-mass star formation that remain actively debated within the field: ones in which the formation of high-mass stars resembles a scaled-up version of the low-mass star formation models (e.g. McKee & Tan 2003), where the high-mass protostellar object accretes material from a compact ($\leq 0.1 \text{ pc}$) fixed mass reservoir – i.e. a core supported by turbulence – and ones in which protostars grow in mass as a result of the large-scale ($\geq 1 \text{ pc}$) and hierarchical gravitational collapse of its parent molecular clump (e.g. Bonnell & Bate 2006; Vázquez-Semadeni et al. 2009, 2019). These two families of high-

* E-mail: rigbya@cardiff.ac.uk (AJR); peretton@cardiff.ac.uk (NP)

mass star formation models may be classified as *core-fed* and *clump-fed* scenarios, respectively (Wang et al. 2010).

A growing weight of observational evidence supports a picture in which large-scale gravitational collapse, resulting in large infall rates, plays a key role in the formation of high-mass stars. Systems with the so-called ‘hub-filament’ configuration (Myers 2009) are also routinely found in high-spatial resolution observations of high-mass clumps, with stellar protoclusters (Schneider et al. 2012; Liu et al. 2016), or the most massive dense cores found at the filament intersections (e.g. Peretto et al. 2013, 2014). The filaments within these systems have also been seen to exhibit longitudinal velocity gradients that could indicate large-scale gas flows (Kirk et al. 2013; Peretto et al. 2014; Lu et al. 2018; Williams et al. 2018; Chen 2019) that supply additional kinetic support at the hub centre, whilst funnelling sufficient mass to the high-mass cores on short time-scales. However, it is neither clear whether longitudinal filamentary accretion and global collapse in high-mass star-forming hub-filament systems is universal, nor at which masses the transition from the solar-type star-forming mode to the high-mass mode occurs.

While the execution of such high-spatial resolution studies across the entire Galaxy remain unfeasible, large and unbiased samples of dense clumps now exist from single-dish Galactic plane (GP) surveys (Schuller et al. 2009; Aguirre et al. 2011; Moore et al. 2015; Molinari et al. 2016a). The Apex Telescope Large-Area Survey of the Galaxy (ATLASGAL; Schuller et al. 2009) at 870 μm identified all of the high-mass clumps ($M > 1000 M_{\odot}$) within the Galaxy (Urquhart et al. 2018), and Jackson et al. (2019) showed that the majority of high-mass clumps are undergoing gravitational collapse. However, limitations in sensitivity or angular resolution prevent these surveys from being able to detect the precursors of intermediate-mass star-forming clumps, where any transition between the core-fed and clump-fed regimes is expected to occur. Studies of clump morphology with higher angular resolution single-dish facilities may allow progress on this front (e.g. Rigby et al. 2018), but large samples must still be acquired.

In this paper, we present the first results from GASTON, a new 200-h Large Programme (PI: N. Peretto) being undertaken at the IRAM 30-m telescope using the NIKA2 camera (Bourrion et al. 2016; Calvo et al. 2016; Adam et al. 2018; Perotto et al. 2020). The GASTON project is comprised of three parts aimed at investigating the enigmatic origin of the IMF from different angles: (i) an $\sim 2 \text{ deg}^2$ survey of a section of the inner GP searching for the transition between the solar-type and high-mass star-forming scenarios; (ii) a search for pre-brown dwarf cores, (iii) a study of dust properties in nearby resolved pre-stellar cores. NIKA2 is a state-of-the-art focal plane array at the Institut de Radio Astronomie Millimétrique (IRAM) 30-m telescope at Pico Veleta in Spain, which completed its commissioning campaign in April 2017. NIKA2 observes in 260- and 150-GHz (hereafter 1.15 and 2.00 mm) continuum wavebands simultaneously by means of a dichroic mirror, with angular resolutions of 11.1 and 17.6 arcsec, respectively. There are two 1.15-mm arrays, each made up of 1140 kinetic inductance detectors (KIDs), and there are 616 detectors on the 2.00-mm array, filling the 6.5-diameter instantaneous field of view (FoV) in each case. The noise-equivalent flux densities (NEFDs) are approximately $30 \text{ mJy s}^{1/2}$ at 1.15 mm and $9 \text{ mJy s}^{1/2}$ at 2.00 mm that, when combined with the large instantaneous FoV, result in mapping speeds that are an order of magnitude larger than the previous generation of instrumentation in operation at the IRAM 30-m telescope.

In this paper, we describe the observations and present the first results from the GASTON GP project, in which we explore the

potential for the NIKA2 observations to identify signatures of the core-fed or clump-fed scenarios of high-mass star formation. This paper is structured in the following way. In Section 2, we describe the observations and data reduction, while in Section 3, we describe our analysis of the 1.15-mm data, and present the results in Section 4. We summarize our results in Section 5

2 OBSERVATIONS

2.1 Observing strategy

The GASTON GP field, centred on $\ell = 23^{\circ}9$, $b = 0^{\circ}05$, was selected for the project due to its extremely high density of compact dust continuum sources, as identified by the Herschel infrared Galactic Plane Survey (Hi-GAL; Molinari et al. 2016a), as well as a large number of infrared-dark clouds (IRDCs; Peretto & Fuller 2009). The GP part of GASTON has been allocated a total of 73.6 h of telescope time. The approximately 2 deg^2 field was divided into three sub-maps, centred on $\ell = 23^{\circ}3$, $23^{\circ}9$, and $24^{\circ}5$, and $b = 0^{\circ}05$, which were observed using an on-the-fly mode, and combined to form the survey area. Pairs of observations with alternating position angles at these three positions are repeated in order to obtain the total integration time, and in such a manner that an approximately uniform sensitivity should be achieved over a large fraction of the map.

The pairs of rectangular raster-scanned maps are performed with alternating position angles in an effort to reduce striping in the reduced maps and are scanned across the plane, as illustrated in Fig. 1, to ensure that the start point of each scan is in a position of low emission. Initially, we adopted a strategy in which scans were alternated by $\pm 15^{\circ}$ with respect to the normal to the GP, as used by ATLASGAL (Schuller et al. 2009), but a $\pm 30^{\circ}$ strategy was finally adopted for most scans, representing a compromise between reducing striping artefacts from data reduction, and maintaining a good mapping efficiency. For the first of the observing runs (N2R12 and N2R14), scans were made with a length of 78 arcmin, with 195-arcsec row spacing (i.e. one half of the 6.5 FoV) between a total of 11 adjacent subscans, and with a scanning speed of 60 arcsec s^{-1} . Following the second observing run (N2R14), some minor adjustments were made: (i) The scanning speed was increased to 70 arcsec s^{-1} , in order to assist with the recovery of extended emission; (ii) the subscan spacing reduced to 180 arcsec to achieve more consistent coverage between subscans; and (iii) the subscan length was changed to 87 arcmin to achieve the same latitude coverage as the earlier scans.

The maps presented in this study are the result of a combined 30 h of integration time, and consist of a total of 108 individual observations taken over five observing pools: N2R12 (2017 October 27–30), N2R14 (2018 January 17–23), N2R15 (2018 February 17–18), N2R26 (2019 January 15–22), and N2R28 (2019 February 12–19). The observations were carried out with sky opacities ranging from $\tau_{225 \text{ GHz}} = 0.08$ to 0.44, and a mean value of 0.19, as measured by the IRAM 30-m taumeter. The mean source elevation was $39^{\circ}9$, and the majority ($\gtrsim 80$ per cent) of the observations were obtained between the hours of 08:00 and 14:00, indicating that the primary beams are likely to be relatively stable around their nominal values of 11.1 (1.15 mm) and 17.6 (2.00 mm) arcsec (Perotto et al. 2020, see their fig. 12). The remainder were observed between 15:00–19:00, during which time the primary beams tend to degrade slightly to ~ 12.5 and 18.0 arcsec, respectively. A detailed description of the operation, calibration and data reduction methods for NIKA2 can be found in Adam et al. (2018) and Perotto et al. (2020).

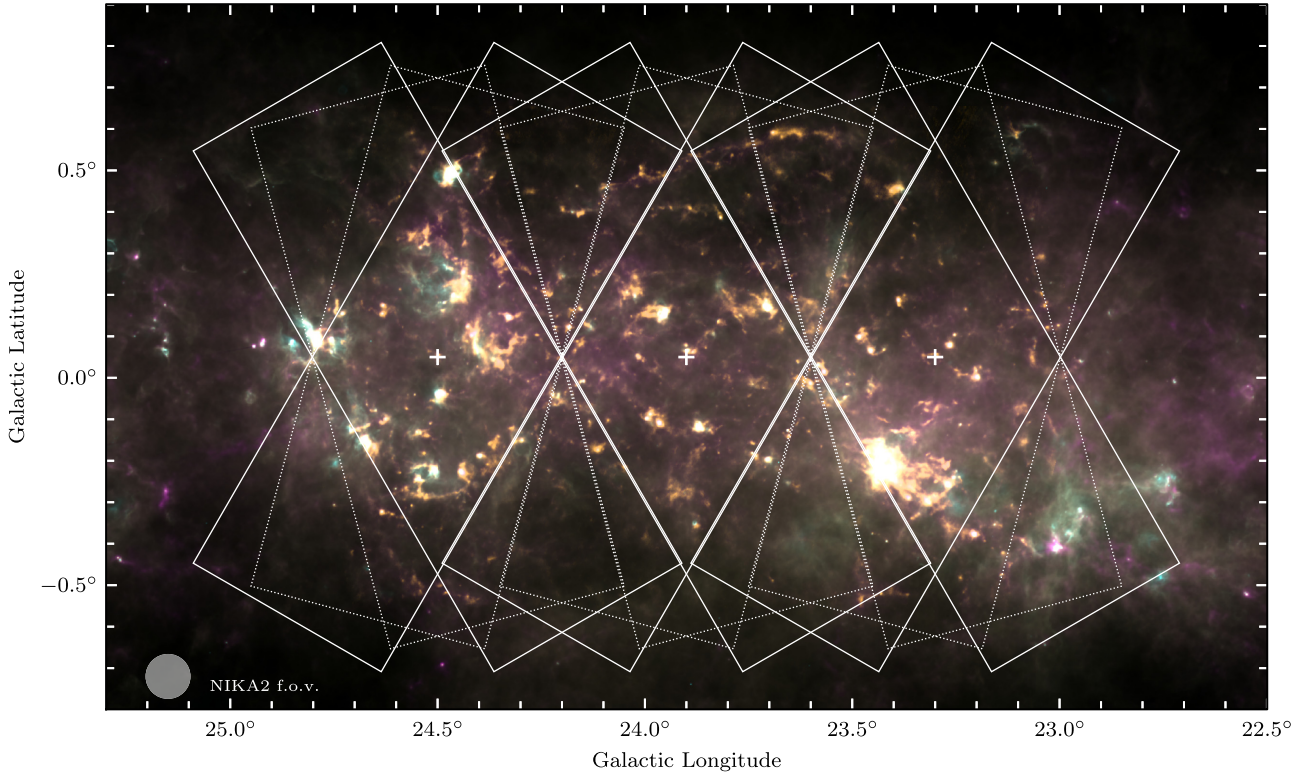


Figure 1. Scanning strategy for the GASTON GP field. Rectangles show the area covered by the central pixel of the arrays in the various scans centred on the white crosses. Dotted white rectangles show the initial strategy, and solid white rectangles show the final adopted strategy. The background image is a four-colour composite¹ consisting of *Herschel* Hi-GAL (Molinari et al. 2016a) 70-, 160-, and 350- μ m images in cyan, yellow, and magenta, and NIKA2 1.15-mm signal-to-noise ratio (S/N, to exclude noisy edges) in orange (this study). The circle in the lower left corner shows the extent of the NIKA2 FoV.

2.2 Data reduction

The data were reduced using the IDL pipeline developed by the NIKA Core Team, which converts the raw time-series data into calibrated maps. While a detailed description of this pipeline will be presented in Ponthieu et al. (in preparation), we summarize the main steps here.

First, individual KIDs that do not meet the performance criteria, are masked from the timelines, as is also done for samples associated with detected cosmic ray impacts. Next, the pointing information from the telescope control system is used in conjunction with a record of the positional offsets of each KID with respect to a reference position to determine the position of each KID on the sky as a function of time. The KIDs are then inter-calibrated, using coefficients determined from their relative gains, before applying a correction for the instantaneous atmospheric opacity, and finally applying an absolute calibration determined for the specific observing run (see Section 2.3).

At this stage, a model of the correlated low-frequency noise components, which consist of electronic noise that is coupled to the readout electronics of subsets of KIDs, and atmospheric noise, which is common to all KIDs, is created; this is the so-called ‘common-mode’. We use the ‘Most Correlated Pixels’ method (Perotto et al. 2020), which works by first discarding any samples in each timeline that, when combined with the telescope pointing information and a source mask, are thought to coincide with an astrophysical source.

Next, cross-correlation coefficients between all KIDs are determined and, for each KID, a model of the low-frequency noise is built by co-adding the timelines of that KID and the 15 most correlated KIDs. A linear fit of this common mode is made and subtracted from each timeline, leaving, in principle, the true astrophysical signal.

Since the GASTON GP field is known to contain a large number of bright and diffuse structures with complex morphologies, we adopt an iterative method to define the source mask used in the calculation of the common mode. For the initial iteration, we use a source mask created² from a Hi-GAL map of the region at 500 μ m (Molinari et al. 2016a). On subsequent iterations, the source mask is determined from the astrophysical signal from the previous iteration, masking out all samples where the S/N is greater than 2, and the process is repeated to improve the estimate of the common mode, and thereby the astrophysical signal. We used a total of 55 iterations, at which point the change in the flux is negligible compared to the map produced by the penultimate iteration.

Finally, the timelines, when combined with the telescope pointing information, are projected (tangential projection) on to a pixel grid with 2.5-arcsec pixels. Inverse-variance noise weighting is used to assign the data samples on to the pixel grid, using a nearest grid point

¹Image generated using the PYTHON package MULTICOLORFITS: <https://github.com/pjcgan/multicolorfits>.

²The source mask is created by first rescaling the 500- μ m Hi-GAL map to 1.15 mm, assuming a greybody spectral energy distribution with a dust emissivity spectral index of $\beta = 1.8$ and a dust temperature of 12 K. We then subtract a version of those data smoothed to 6.5 arcmin to approximate the NIKA2 spatial filtering, and convert the resulting map into a binary source mask by identifying those pixels which exceed half of the expected sensitivity of 1.55 mJy beam⁻¹.

method, and to combine the maps associated with each individual scan into a mosaic of the GP field.

The correlated noise removal is the origin of the spatial filtering within the reduced maps, which is common to all submillimetre and millimetre-continuum imaging from ground-based telescopes. The level of spatial filtering applied here is typically on the order of the instantaneous FoV which, for NIKA2, is approximately 6.5. We do not calculate the pipeline transfer function that describes the ratio of the power spectrum within the reduced image to an estimate of the absolute source power spectrum over the sky region in this paper. An example of the 2.00-mm NIKA2 transfer function towards a galaxy cluster may be found in Ruppén et al. (2018, see their fig. 3), who found that around 95 per cent of emission at angular scales between the 2.00-mm beam size and the FoV is recovered, falling off rapidly at larger scales, and we expect a similar level of recovery here.

2.3 Calibration

Uranus and Neptune are used as the primary flux calibrators, and so colour corrections must be applied to our photometric maps to account for the difference between the spectral energy distributions (SEDs) of those planets, and the approximate SEDs of the sources is the NIKA2 imaging. We apply colour correction factors of 1.02 and 0.97 to the 1.15- and 2.00-mm bands, respectively, which were interpolated from table 12 of Perotto et al. (2020), assuming an SED of the form $I_\nu \propto (\nu/\nu_0)^{2+\beta}$, with the mean GP value for the dust emissivity spectral index of $\beta = 1.8$ (Planck Collaboration XXV 2011). The total calibration uncertainties are determined from the quadrature sum of the 5 per cent planet model uncertainties, the 5.7 and 3.0 per cent point-source rms calibration uncertainty, the 0.6 and 0.3 per cent systematic calibration uncertainties, and the 3 and 2 per cent uncertainties in the reference beam efficiencies, giving a total of 8 and 6 per cent calibration uncertainty on the 1.15- and 2.00-mm integrated flux densities, respectively. A detailed description of the absolute calibration procedure can be found in Perotto et al. (2020).

2.4 Ancillary data

We make use of two spectral line data sets in order to determine velocities to continuum sources in Section 3.3. First, we use ^{13}CO (1–0) data (110.201 GHz) from the FOREST unbiased GP imaging survey with the Nobeyama 45-m telescope (FUGIN; Umemoto et al. 2017), which have an angular resolution of 21 arcsec, and an rms of $\sigma(T_A^*) = 0.87$ K per 0.65 km s^{-1} velocity channel. The spatial extent of the FUGIN data covers the entire GASTON GP field, and the velocity range of $-100 < v_{\text{LSR}} < 200 \text{ km s}^{-1}$ covers all of the Galactic spiral arms along this sightline (Dame, Hartmann & Thaddeus 2001).

We also make use of data from the pilot study for the Radio Ammonia Mid-Plane Survey (RAMPS; Hogge et al. 2018), a GP survey of ammonia and water maser emission carried out with the Green Bank Telescope. The L23 and L24 regions of the pilot study have a considerable overlap with the GASTON GP field, spanning $22^\circ.5 \gtrsim \ell \lesssim 24^\circ.5$, with $|b| \lesssim 0.4$, and we make use of the moment 1 maps of NH_3 (1,1) inversion emission at 23.694 GHz. These data have an angular resolution of 32 arcsec, and 0.018 km s^{-1} spectral channels.

Data from Hi-GAL (Molinari et al. 2016a) at 160 and 250 μm are used in Section 3.5 for colour temperature determinations. The 160- μm PACS (Poglitsch et al. 2010) data have an angular resolution of 12 arcsec, and the 250- μm SPIRE (Griffin et al. 2010) data have an angular resolution of 18 arcsec. In both cases the uncertainties

are dominated by calibration, which we take as 7 per cent for PACS³ and 6.5 per cent for SPIRE,⁴ including a 1 per cent contribution for extended sources due to uncertainty in the SPIRE beam.

We also use 8- μm data from the Galactic Legacy Infrared Mid-Plane Survey Extraordinaire (GLIMPSE; Benjamin et al. 2003; Churchwell et al. 2009) in Section 4.3. These *Spitzer* data have an angular resolution of 2.0 arcsec.

3 ANALYSIS

3.1 Photometric maps

In Fig. 2, we present the NIKA2 photometric maps, after cropping to the largest contiguous rectangular field where the sensitivity is approximately uniform. The total integration time obtained so far is ~ 30 h in 108 scans, corresponding to roughly 40 per cent of what will be acquired upon the completion of GASTON.

Although we do not use the 2.00-mm map in the analysis presented in this paper, we present it in Fig. 2 for the sake of completeness, and to display the data quality. We will make use of these data in future GASTON studies and, in particular, the 1.15- and 2.00-mm data will be used in combination to study the dust emissivity spectral index, β , as in Rigby et al. (2018).

3.2 Source extraction

A source extraction was carried out on the 1.15-mm maps using the PYTHON ASTRODENDRO⁵ package, which uses a dendrogram (Rosolowsky et al. 2008) scheme to segment the emission while maintaining the ability to track its hierarchical structure. The dendrogram separates emission features when they are considered to be significant given their size, and difference in their intensity with respect to the background and to overlapping emission features.

The 1.15-mm map was first cropped to $22^\circ.95 \leq \ell \leq 24^\circ.85$ and $-0^\circ.55 \leq b \leq 0^\circ.65$ (as in Fig. 2) in order to exclude the noisier edge regions. The map was then smoothed to 13-arcsec resolution in order to increase the S/N. Measuring the sensitivity in such deep maps of the inner GP is problematic because of source confusion, and so several methods were used to measure the rms noise level. The preferred method was to perform a Gaussian fit to the distribution of pixel values taken from the null maps within a sample of four octagonal sky apertures with radii of ~ 1.9 arcmin (displayed in Appendix A) located across the main survey region. The four sky apertures were carefully placed across the map in order to both sample the range of integration depth in a representative manner, and to avoid regions with residual emission. The null maps are produced by the data reduction pipeline, and are made by alternately adding and subtracting the maps generated by individual scans of the full observation set. In principle, this produces a map free of emission structures whilst preserving the noise characteristics of the map. This method yields the most robust noise estimate in a line-of-sight with emission structures at all positions. The rms determined in this way was $4.18 \text{ mJy } 13\text{-arcsec beam}^{-1}$.

Based on the estimated rms noise level, we set the minimum value (min.value) above which emission is considered to be real, and the minimum difference between substructures (min.delta) to three times the global rms, while the minimum structure size

³<https://www.cosmos.esa.int/web/herschel/pacs-overview>.

⁴<https://www.cosmos.esa.int/web/herschel/spire-overview>.

⁵<http://www.dendrograms.org/>.

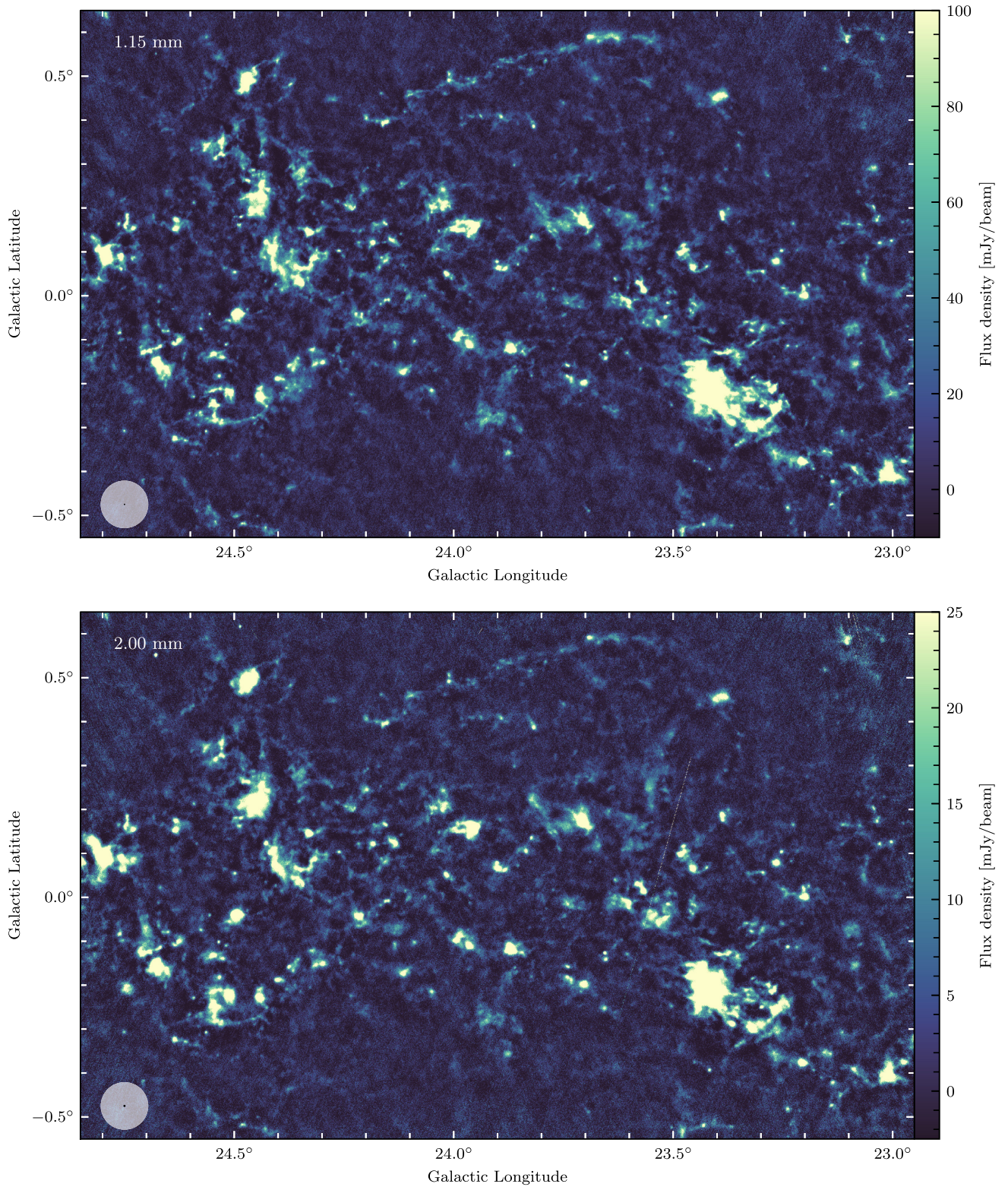


Figure 2. GASTON GP field at 1.15 (top panel) and 2.00 mm (bottom panel), which have FWHM resolutions of 11.1 and 17.6 arcsec, respectively. The FoV and primary beam sizes are shown in the bottom left-hand corners as opaque white, and black circles, respectively.

(`min_npix`) to be 16 pixels – equivalent to half of the 13-arcsec beam area. The latter choice was made for the minimum size criterion because pure point sources will be reported as smaller than the beamsize due to the clipping of the wings that extend below the minimum intensity level (as in the ‘bijection’ scheme described in Rosolowsky et al. 2008), and so would be filtered from the data if the full beam area was used. This is especially true for low S/N sources.

The application of a single average rms noise level over the entire area results in a small number of spurious detections in the noisier-than-average regions and, conversely, some undetected compact sources are visible in the residuals in the regions where the sensitivity is greater than this average value. To enable us to eliminate spurious sources, we calculate the local S/N of each source in each waveband. As mentioned earlier, the null maps contain some residual positive or negative emission, and are unsuitable for use as noise maps for all sources. We therefore produce a noise map in each waveband, that is calculated from the weight maps produced by the IDL pipeline. The per-pixel weights are proportional to the inverse variance of the flux densities measured at each position on the pixel grid from the contributing observations. The noise at each pixel position, σ_i , is therefore calculated as $\sigma_i = \sigma_0 \sqrt{w_0/w_i}$, where σ_0 is the rms measured from the pixels within the four sky apertures, w_0 is the mean value of the pixel weights within the four sky apertures, and w_i is the per-pixel weight. We apply a ‘pruning’ scheme in the dendrogram construction such that any objects that do not have a local S/N that exceeds 6 – in keeping with the dendrogram’s input parameters – in at least one of their constituent pixels are removed.

In the dendrogram nomenclature, emission features may be described as ‘trunks’, ‘branches’, or ‘leaves’, depending on their position within the hierarchy. The trunks are the lowest lying structures in the hierarchy, at the minimum detection threshold and thus describe the maximum extent of any isolated emission region, while branches are at a higher contour level within a trunk, and may contain multiple leaves. In the following analysis, we refer to the dendrogram leaves – those emission features that contain no further discernible substructure – as ‘clumps’, whereas we use the term ‘source’ to refer to any emission feature extracted using the dendrogram, regardless of its position within the hierarchy. That is to say that a source may contain multiple clumps. The dendrogram extraction found a total of 2446 sources of all kinds, consisting of 488 isolated emission structures (trunks) with 1467 clumps (leaves). We summarize the basic properties – flux densities and source sizes – in Appendix B.

3.3 Source velocity assignments

Determining the systemic velocity, and thereby a kinematic distance, for each extracted source is a necessary step in calculating many of its physical properties. However, due to both the high sensitivity of the GASTON GP maps at 1.15 mm, and the particularly complicated line-of-sight for the survey region, source confusion presents a significant challenge in their determination. We therefore adopted a procedure with several stages to establish the most reliable possible velocities, which are summarized in this Section, and further details may be found in Appendix C.

To determine a velocity for each GASTON source, first, an integrated ^{13}CO (1–0) spectrum is extracted from the FUGIN data of all pixels that lie within the source footprint. Many of the spectra exhibit multiple emission features, and so the velocity with the greatest intensity is assigned as the centroid velocity, v_{cen} , of each source. In cases where the line of sight contains multiple velocity components, this velocity is assumed to trace the highest-column density material that is most likely to be associated with the dust

continuum source. A quality flag was assigned to each velocity assignment, with a value of 3 given to the most robust assignments, and 1 to the poorest, and a value of 2 is given for intermediate-quality assignments.

The main advantage of this method is its simplicity, but it may obviously provide wrong velocity assignments in some cases, primarily due to the difference in densities traced by the molecular and continuum data. In principle, the dust continuum emission traces the full column density of molecular hydrogen, but the spatial filtering applied by the data reduction pipeline removes much of the emission from diffuse gas, leaving the more compact, and higher column-density gas. Observations of dense gas tracers would provide more suitable velocities for the GASTON GP sources than ^{13}CO (1–0), and so, as a second stage, we use data from the RAMPS pilot study (Hogge et al. 2018) of NH_3 (1,1), which has a critical density of $n_{\text{crit}} \sim 3 \times 10^3 \text{ cm}^{-3}$ to determine dense-gas velocities. We adopt velocities for 517 (21 per cent) of the sources from the first component from the line-fitting procedure, and a further 513 (21 per cent) from the first-moment maps of Hogge et al. (2018), assigning a quality flag of 3 for all ammonia-derived velocities.

The hierarchical structure of the 1.15-mm emission is recorded within the dendrogram catalogue, and since each source now has an individual velocity assignment, along with a quality flag, the structure can be used to inform us about likelihood of the more dubious velocity assignments being the correct ones. For example, if a velocity assignment of 50.5 km s^{-1} is given for a source with a poor-quality flag, but its parent structure has a robust velocity assignment of 50.9 km s^{-1} , then it is considered to be part of that same velocity group. If a poor velocity assignment is made to a source at a significantly different velocity to its neighbours within the hierarchy that have robust assignments, then we adopt the median velocity of those neighbours for that source. In this way, the velocity assignments for all sources (where possible) with a quality flag of either 1 or 2 were refined, and we adopt the quality flag of the most robust velocity within that group for all its constituent sources. We refer to the velocities refined in this way as v_{group} . In Table 1, we present the distribution of quality flags assigned to the velocities determined and refined in this Section.

Finally, in order to ensure that velocity-coherent structures are retained, we perform a friends-of-friends grouping. Clusters of sources were identified by recursively linking groups that lie within a tolerance of 0.06 in the ℓ and b axes, and within 2.5 km s^{-1} in line-of-sight velocities, v_{group} , which approximates the characteristics of small (10 pc-diameter) molecular clouds (e.g. Roman-Duval et al. 2010) at a distance of 5 kpc – the median near-kinematic distance for the revised source velocities along this sightline. This allows, for example, the prominent large-scale filament that extends from $[\ell, b, v_{\text{group}}] = [24^\circ.16, 0^\circ.40, 94]$ to $[23^\circ.57, 0^\circ.59, 98 \text{ km s}^{-1}]$ to be identified as a single cluster. Of all GASTON sources, 2249 (92 per cent) are connected to one of the 211 clusters identified in this way, which contain up to 233 individual sources, with a median size of 25 sources. For each cluster, we also calculate the median group velocity, v_{cluster} , and the associated standard deviation, σ_{cluster} , using only the sources that have the most reliable velocity flags from that cluster.

The usage of 21-arcsec-resolution spectral line data with our 13-arcsec resolution element means that there is an element of beam-averaging that is likely to blend together velocity components that would be distinct at matching resolution. We have not made use of available survey data in ^{12}CO (3–2) (Dempsey, Thomas & Currie 2013) at 16.6-arcsec resolution, due to optical depth considerations. While new facilities (e.g. Klaassen et al. 2019; Stanke et al. 2019) and instrumentation (e.g. Frayer et al. 2020) hold much promise,

Table 1. Distribution of quality flags for each stage in the velocity determination process.

Stage	Description	Flag value					
		<i>N</i>	3 (per cent)	<i>N</i>	2 (per cent)	<i>N</i>	1 (per cent)
i)	FUGIN ^{13}CO (1–0)	1685	68.9	298	12.2	463	18.9
ii)	RAMPS NH_3 (1,1)	1916	78.3	165	6.7	365	14.9
iii)	Dendrogram refinement	2169	88.7	75	3.1	202	8.3

Note. For each flag value, the total number of sources with that value (*N*) and the percentage of the sample made up by those sources (per cent) is given.

unbiased GP surveys in dense gas tracers such as N_2H^+ , HCN, NH_3 , or even C^{18}O at sufficiently high sensitivity and angular resolution are not currently available, and the FUGIN and RAMPS data present the best data for the moment. We note that CHIMPS2 (Eden et al. 2020) aims to cover the GASTON GP field in ^{13}CO (3–2) that, at an angular resolution of 15 arcsec and with a higher critical density molecular gas tracer than that of the FUGIN, should complement GASTON upon its completion, though it is limited to $|b| < 0.5$. We note that we do not use data from the Galactic Ring Survey (GRS; Jackson et al. 2006), which have velocity resolution and sensitivity superior to the FUGIN data, because the higher angular resolution of the latter are more important in this case.

3.4 Distance determination

Distances to sources lying within the GP are typically determined by measuring their radial velocity, and by assuming that the source is following a circular orbit described by a model of the Galactic rotation curve. For sources lying within the Solar Circle, each radial velocity corresponds to two kinematic distance solutions. In order to determine whether a source lies at the near- or far-kinematic distance, one either requires additional information, or some assumptions must be made. We determined kinematic distance solutions to each source using the Galactic rotation model of Reid et al. (2019) based on their ℓ , b , v_{group} coordinates. As a first estimate, we use version 2.4.1 of the Bayesian distance calculator⁶ (Reid et al. 2016), adopting the recommended weightings for the priors based upon the spiral arm model, GP scale height, and proximity to sources with a measured parallax. This gives an independent distance estimate to each source (regardless of position within the hierarchy), but we note that this preferentially concentrates sources into the spiral arms. However, since the spiral structure of the Milky Way is very much an ongoing matter of research, we do not adopt these distance solutions directly. Rather, we use the distances calculated with these various priors to distinguish between the near and far kinematic distances.

Of the 2446 sources, 2274 (93 per cent) are assigned to the near kinematic distance in this way, with only 120 (5 per cent) assigned to the far distance, and the remainder are located at the tangent point ($d \approx 7.5$ kpc, assuming the Reid et al. 2019 distance to the Galactic centre of 8.15 kpc), where the near and far solutions are equal. By comparison, 77 per cent of the ATLASGAL clumps within the catalogue of Urquhart et al. (2018) have distances less than the tangent distance, after reducing the sample to those within the GASTON field, and adjusting for the difference in tangent distance associated with the adopted Galactic rotation model, and a similar figure of 79 per cent is found for 1.1-mm sources from the Bolocam Galactic Plane Survey (BGPS; Ellsworth-Bowers et al.

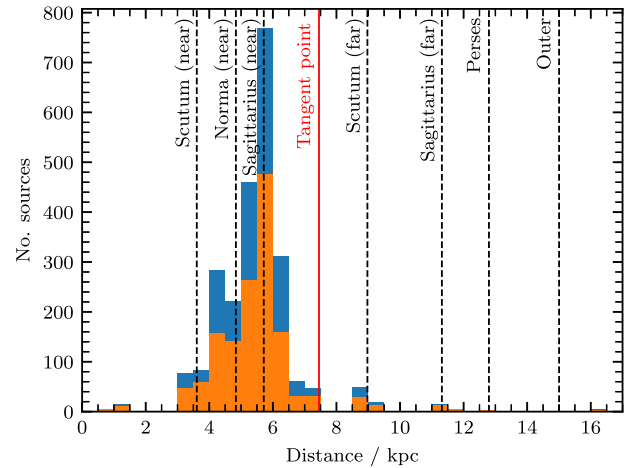


Figure 3. Distribution of kinematic distances in 0.5-kpc-wide bins, as determined for all sources (blue), and the subset of clumps (orange), in Section 3.4. The approximate locations of the intersection of the spiral arms with the line of sight, as in the Reid et al. (2019) model, are shown as vertical lines.

2015). Only 35 per cent of clumps within the Hi-GAL catalogue of Elia et al. (2017) were assigned distances on the near-side of the tangent point. However, in the latter case, there is a bias towards far distances, as objects for which data (such as extinction data) that may allow a near determination to be made were unavailable are assumed to be located at the far distance. After removing clumps with this default assumption from the Hi-GAL catalogue, the number rises sharply to 89 per cent. We show the resulting distribution of distances in Fig. 3, and compare the values to their counterparts in other catalogues in Appendix E. We find that two-thirds of our distance estimates match with those of the other catalogues to within 1 kpc and can be considered to be in general agreement.

It is not surprising that a large fraction of sources are found to be located at the near distance as a result of the combination of Malmquist bias, and the fact that the $\ell = 24^\circ$ sightline covers nearby sections of the Sagittarius, Scutum, and Norma spiral arms (Reid et al. 2019). However, we caution that the friends-of-friends analysis used here will inevitably lead to cases where sources at the far distance appear in clusters of sources for which our methodology favours a near distance.

In a manner similar to the determination of the cluster velocities, the median distance to each cluster is determined from the set of distances stemming from its constituent sources with the most robust velocity measurements, and then assigned to all sources within the cluster. Finally, we account for clusters that host maser sources with a known parallax distance. A total of 11 masers with parallax distances are located in the field, and after identifying the

⁶<http://bessel.vlbi-astrometry.org/>.

cluster associated with the l , b , v position of each maser source, we assign the parallax distance to all sources within that cluster. There are a total of nine clusters that contain maser sources and are assigned distances in this way. Two of the eleven maser sources, G023.00–0.41 ($v = 79 \pm 5 \text{ km s}^{-1}$, $d = 4.88 \pm 0.36 \text{ kpc}$) and G023.20–0.37 ($v = 82 \pm 10 \text{ km s}^{-1}$, $d = 4.18 \pm 0.61 \text{ kpc}$), are associated with objects within the same cluster, and since the associated distances are significantly different, a choice must be made. For this cluster, we adopt the parallax distance associated with G023.00–0.41, which is in stronger agreement with both the mean of the predetermined kinematic distances ($d = 5.20 \text{ kpc}$) and systemic velocities ($77.5 \pm 0.2 \text{ km s}^{-1}$). A further two of the maser sources, G023.65–0.12 and G23.6, are located within the same cluster, but their distances and velocities are identical, and so present no conflict. The median discrepancy between the kinematic and maser parallax distances to these seven clusters is 0.41 kpc, though the largest difference is 1.73 kpc.

At this point, we have derived two distances for each object within the dendrogram, d_{group} and d_{cluster} , both of which are derived principally from the sources' line-of-sight velocities, and the assumption that they simply follow circular motions about the Galactic Centre. However, it is well known that line-of-sight streaming motions of up to $\sim 10 \text{ km s}^{-1}$ can be present for molecular clouds, and peculiar motions within molecular cloud complexes further complicate this picture. To mitigate these effects in the subsequent analysis, the cluster distances were, therefore, used to determine any distance-dependent properties, and we note that the largest internal velocity dispersion of the clusters identified using the friends-of-friends grouping is 2.3 km s^{-1} . The difference between the two sets of distances is minimal and, when considering the population of clumps, 98 per cent have a difference between d_{group} and d_{cluster} of less than 0.2 kpc.

3.5 Physical properties

Estimating dust temperatures (T_d) and column densities (N_{H_2}) of molecular hydrogen from dust continuum imaging is typically done by fitting a modified blackbody to the SED:

$$I_\nu = \mu_{\text{H}_2} m_{\text{H}} N_{\text{H}_2} \kappa_\nu B_\nu(T_d), \quad (1)$$

where I_ν is the specific intensity, μ_{H_2} is the mean molecular weight per hydrogen molecular (with a value of 2.8), m_{H} is the mass of a hydrogen atom, κ_ν is the specific dust mass absorption coefficient, and B_ν is the Planck function evaluated at the frequency ν for the dust temperature. Three or more photometry points taken from imaging at submillimetre or millimetre wavelengths, convolved to the same resolution, are usually required (assuming both μ and κ_ν are fixed), providing a fit at the resolution of the longest wavelength being used – typically 36 arcsec for *Herschel* 500- μm data. However, the usage of dust temperatures derived at 36-arcsec resolution for the 13-arcsec-resolution scales of the 1.15-mm data would mean that local minima and maxima would be beam-diluted. The different spatial frequencies probed by the ground- and space-based observatories also presents a further limitation.

To overcome these limitations, we adopt colour temperatures derived from the ratio of Hi-GAL 160- to 250- μm flux densities (e.g. Peretto et al. 2016; Rigby et al. 2018), which have an effective resolution of 18 arcsec. The colour temperature is related to the ratio of flux densities in the following way:

$$\frac{S_{160}}{S_{250}} = \frac{B_{160}(T_{\text{col}})}{B_{250}(T_{\text{col}})} \left(\frac{250 \mu\text{m}}{160 \mu\text{m}} \right)^\beta. \quad (2)$$

where S_ν is the source-integrated flux density $S_\nu = \int I_\nu d\Omega$. By adopting a fixed value of $\beta = 1.8$, and using 160- and 250- μm photometry from Hi-GAL imaging, we sampled T_{col} from a grid of values ranging from 2.7 to 50 K, with 0.1 K intervals. Uncertainties on the *Herschel* photometry result in uncertainties on T_{col} of $\sim 1.5 \text{ K}$. We first performed a convolution to effectively match the PSF of the 160 μm image to that of the 250- μm image following Aniano et al. (2011), and both images were high-pass filtered to remove spatial frequencies larger than 6.5 arcmin using the NEBULISER application of the Cambridge Astronomy Survey Unit Tools software package⁷ to approximate the spatial filtering present in the NIKA2 data.

Many of the sources in the GP field consist of a mixture of compact and extended structure, and the flux densities of compact sources, therefore, usually contain contributions both of the compact object itself, and the extended emission structure on which they reside. We have therefore adopted the ‘clipping’ scheme (Rosolowsky et al. 2008) of flux estimation for S_ν (and thereby the mass) for each source, in which the local background level of flux is subtracted from the total, or put another way, only flux above the contour level defining the structure is counted in the integrated flux. Although Rosolowsky et al. (2008) prefer the ‘bijection’ scheme, in which the entirety of the flux density integrated across the source’s sky position is used for the mass calculation within molecular clouds, the clipping scheme is used here, since we are most interested in tracing the mass of the smallest structures within the hierarchy.

We determine the appropriate background level for each source by first performing a binary dilation on each of the source masks, which expands the perimeter by one pixel, and allows the encompassing perimeter pixels to be isolated (i.e. pixels just below the 1.15-mm contour level at which the source was identified). The background flux density at 1.15 mm is the contour level at which the source was separated from its parent structure, and was determined as the mean of the minimum pixel value within the source, and the maximum pixel value within the perimeter pixels. A slightly different approach must be taken to determine the background flux densities for the 160 and 250 μm because the morphology of emission at those wavelengths may differ from the morphology at 1.15 mm at which the source was defined. In contrast to the 1.15-mm case, some of the pixels within the boundary region may be brighter than some of those within the source area. At these wavelengths, the background value is therefore taken as half of the sum of the mean value of pixels within the perimeter, and the minimum pixel value within the source, thus approximating the 1.15-mm contour.

The masses of the sources identified in the dendrogram are calculated from S_ν , the monochromatic flux densities integrated over the solid angle subtended by the source at 1.15 mm using

$$M = \frac{S_\nu d^2}{\kappa_\nu B_\nu(T_{\text{col}})}, \quad (3)$$

where the dust absorption coefficient is given by $\kappa_\nu = 0.1(\nu/1\text{THz})^\beta$ (Beckwith et al. 1990). We use a value evaluated at 1.15 mm (260 GHz) of $\kappa_{260} = 0.009 \text{ cm}^2 \text{ g}^{-1}$, assuming $\beta = 1.8$, and incorporating a gas-to-dust mass ratio of 100. This value is close to the value of $0.008 \text{ cm}^2 \text{ g}^{-1}$ used in our previous NIKA study (Rigby et al. 2018), and that adopted by Hi-GAL in Elia et al. (2017). We discuss the impact of different values of β upon our results in Section 4.4. For all sources, we use their cluster distance, d_{cluster} , as determined

⁷<http://casu.ast.cam.ac.uk/surveys-projects/software-release/background-filtering>.

in Section 3.4, and we decrease the integrated flux densities by 4 per cent to account for the median level of CO (2–1) contamination (determined in Appendix D), adding a 4 per cent contribution to the corresponding uncertainties. Uncertainties are calculated using Monte Carlo methods for the primary quantities for each source (i.e. distance, absolute flux calibration uncertainty), and propagated to all derived quantities, such as mass and colour temperature. We note that the contribution of the factor of ~ 2 uncertainty on κ_v is not included in our mass uncertainties since its effects are likely to be mostly systematic across the region, and will consequently not have an impact on the analyses presented here.

We note that, as a result of the 18-arcsec resolution of the colour temperature estimates, smaller scale temperature variations will not be detected. This may result in the overestimation of source masses where compact sources warmer than the beam-smeared average are present, and vice-versa for colder compact sources. At large radii, the temperature of protostellar sources is of the form $T(r) \propto r^{2/(4+\beta)}$, where β is the dust emissivity spectral index (Terebey, Chandler & Andre 1993) and, by adopting $\beta = 1.8$, we can expect to underestimate dust temperatures of point sources at 18 arcsec by up to ~ 11 per cent compared to those at 13 arcsec. For resolved sources, this effect is unlikely to play a significant role, but we conservatively increase the uncertainties on T_{col} for all sources by 11 per cent to account for this effect.

We discuss the resulting distributions of mass and temperature in Sections 4.2 and 4.3, and present the overall distribution in Fig. 6.

4 RESULTS

4.1 Comparison to other ground-based (sub-)millimetre data

This particular region of the GP is well covered in many recent surveys of (sub-)millimetre continuum emission. The most similar ground-based surveys to our NIKA2 observations are the 1.1-mm BGPS at 33-arcsec resolution (Aguirre et al. 2011; Ginsburg et al. 2013) and, at 870 μm and with 19.2-arcsec resolution, ATLASGAL (Schuller et al. 2009). While the data from these surveys are very well matched to GASTON in terms of wavelength, they were designed to cover far wider regions of the GP (covering 192 and 360 deg^2 , respectively), and so their sensitivities are lower than those of the data presented here.

We have identified a total of 1467 GASTON clumps (dendrogram leaves), compared to 346 ATLASGAL compact sources (Urquhart et al. 2018) and 164 from the BGPS (Ellsworth-Bowers et al. 2015) within the same sky area. Of the 164 sources from the BGPS, only 2 sources do not have centroid coordinates lying within three convolved beam radii⁸ of a GASTON clump centroid. In both cases, the BGPS source is located within a large-scale diffuse region in which the GASTON image has partially resolved and detected clumps around the periphery, and have centroid coordinates located far away from the particular region of emission extracted from the BGPS data. All of the ATLASGAL clumps lie within three convolved beam radii of a GASTON compact source, and they all fall within the boundaries of emission extracted by the dendrogram analysis in this work. The centroid coordinates of 1237 (84 per cent) and 1114 (75 per cent) out of the 1467 GASTON clumps are located further than three convolved beam radii of any BGPS or ATLASGAL source centroid,

respectively. Although this positional matching based on centroid coordinates will be less effective for sources that are not centrally concentrated, ~ 95 per cent of GASTON clumps have angular radii of less than three beam radii (i.e. 39 arcsec), and so it is clear that the majority of GASTON clumps are new identifications.

In Fig. 4, we compare the GASTON 1.15-mm imaging in a region centred on $\ell = 24^\circ 465$, $b = -0^\circ 205$ to observations from ATLASGAL and the BGPS, and 8- μm emission from the GLIMPSE. The greater angular resolution of the GASTON GP data compared to BGPS and ATLASGAL clearly allows new substructures to be resolved within previously detected emission regions, in addition to the detection of new faint compact sources. By comparison to the *Spitzer*/GLIMPSE 8- μm image, it can also be seen that the 1.15-mm continuum is effective at tracing structures seen in absorption as infrared-dark features as well as compact regions of bright emission at 8 μm . Not all features from the 8 μm image have discernible 1.15-mm counterparts, and although most of the faint and extended 8 μm emission (presumably arising from hot dust grains or PAH emission at ionization fronts) is not seen in the GASTON image, there are visible counterparts to some of these features, such as the column extending from $[24^\circ 62, -0^\circ 10]$ to $[24^\circ 60, -0^\circ 20]$ and, surprisingly, the bow shock at $[24^\circ 32, -0^\circ 15]$.

A more sensitive study of this region was recently carried out with the Large Millimetre Telescope at 1.1 mm using AzTEC (Heyer et al. 2018). Those observations were centred on $\ell = 24^\circ 5$, $b = 0^\circ 0$, with a radius of $0^\circ 55$ – partially overlapping with the GASTON GP field – inside which they reached a median sensitivity of 9 mJy beam⁻¹ with a resolution of 8.5 arcsec. After rescaling to 1.15 mm, assuming a spectral energy distribution of the form $I_\nu \propto \nu^{3.8}$, to match our NIKA2 observations, the point-source sensitivity would correspond to ~ 7.6 mJy beam⁻¹, making them around half as sensitive as those presented here. There are a total of 993 AzTEC compact sources within the region that overlaps with GASTON, after restricting the source catalogue to those considered by Heyer et al. (2018) to be robust, i.e. probability of a false-positive detection of < 2 per cent, compared to 693 compact GASTON clumps, and 1175 GASTON sources of all kinds.

Of the AzTEC source centroids, 698 (70 per cent) fall upon a pixel associated with an extracted GASTON source, and the remaining 30 per cent are likely to be a result of (i) the greater resolution of the 32-m LMT Alfonso Serrano at 1.1 mm (8.5 arcsec) compared to the 30-m IRAM observations at 1.15 mm (11.1 arcsec) and (ii) the more pronounced spatial filtering in the AzTEC observations (removing spatial scales > 50 arcsec) compared to that of NIKA2 (~ 6.5); (iii) the area of maximum sensitivity in the AzTEC map extends into regions of higher noise in the GASTON map, where our detection rate is lower. Furthermore, we have adopted dendrogram parameters in this study optimized to detect emission structures on all scales, and a treatment using harsher spatial filtering like that applied in the AzTEC data, or those of the ArTéMiS observations of Peretto et al. (2020) would be likely to increase the overall number of compact sources detected.

The statistics discussed in this section are summarized in Table 2.

4.2 A new population of clumps

NIKA2 is expected to be more sensitive to cold and compact dust sources within the Milky Way than *Herschel*, owing to its greater angular resolution at longer wavelengths. To determine whether we have identified compact structures within the GASTON GP data that were not identified using the various extraction techniques applied to the *Herschel* data, we compared our catalogue to two sets of Hi-GAL

⁸Our search parameters adopt convolved beam radii, that are the quadrature sum of the (smoothed) 13-arcsec GASTON beam and the beam size of the data set in question.

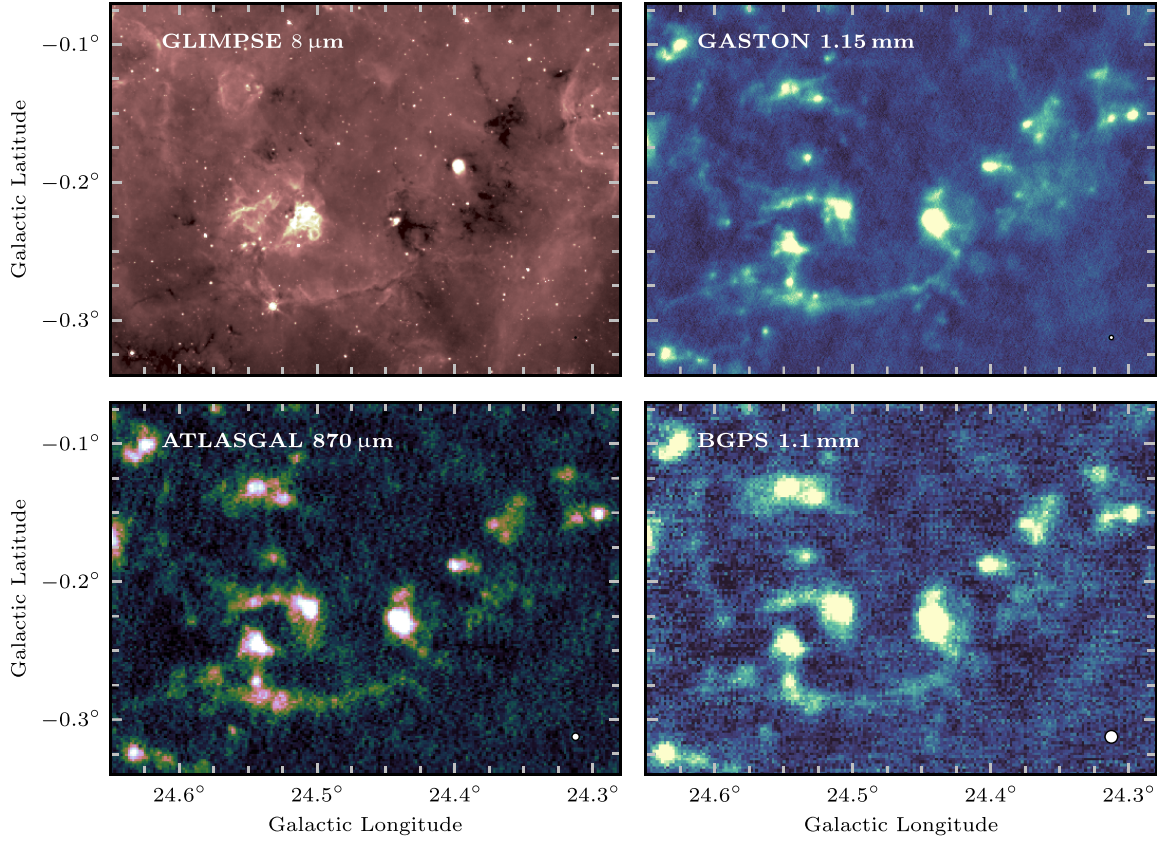


Figure 4. Comparison of the GASTON 1.15-mm imaging (upper right-hand panel) to imaging from *Spitzer*/GLIMPSE 8- μ m (upper left-hand panel; Benjamin et al. 2003; Churchwell et al. 2009), APEX/ATLASGAL 870- μ m (lower left-hand panel; Schuller et al. 2009), and CSO/BGPS (lower right-hand panel; Ginsburg et al. 2013) imaging at 1.1 mm. FWHM beam sizes are shown in the lower-right corners of each image. All images are shown at their native resolution and pixel sizes.

Table 2. Summary of the results of catalogue cross-matching procedure used to compare the GASTON source statistics with other surveys in Sections 4.1 and 4.2.

Data (i)	Reference (ii)	Wavelength (iii)	Resolution (arcsec) (iv)	N (v)	N_{source} (vi)	N_{clump} (vii)	N_{matched} (viii)	N_{unique} (ix)
BGPS	Ellsworth-Bowers et al. (2015)	1.1 mm	33	164	162	162	230	1237
ATLASGAL	Urquhart et al. (2018)	870 μ m	19.2	346	346	345	365	1102
AzTEC	Heyer et al. (2018)	1.1 mm	8.5	993	698	420	336	357
Hi-GAL	Elia et al. (2017)	BM ^c	36 ^d	857	731	734	702	765
Hi-GAL	Molinari et al. (2016a)	70 μ m	6	2807	1698	814	693	774
Hi-GAL	Molinari et al. (2016a)	160 μ m	12	4533	2334	1516	1145	322
Hi-GAL	Molinari et al. (2016a)	250 μ m	18	2591	1701	1466	1146	321
Hi-GAL	Molinari et al. (2016a)	350 μ m	24	1655	1247	1194	1075	392
Hi-GAL	Molinari et al. (2016a)	500 μ m	35	938	771	798	838	629

Notes. For each data set (i), we identify (ii) the catalogue used; (iii) the wavelength; (iv) the angular resolution; (v) N , the total number of sources in the catalogue falling within the cropped GASTON GP field; (vi) N_{source} , the number of source centroids that fall upon a pixel belonging to an identified GASTON source; (vii) N_{clump} , the number of sources coincident^a with a GASTON clump centroid; (viii) N_{source} , the number of GASTON clump centroids coincident^b with a catalogued source centroid; (ix) N_{unique} , the number of GASTON clump centroids that are not coincident with a catalogued source centroid. ^aSources are regarded as coincident when the separation of their centroids is less than or equal to three convolved beam radii (Section 4.1).

^bGASTON sources are regarded as coincident with a catalogued source as above with the exception of the Hi-GAL sources for which the circular or elliptical footprints are checked for overlapping GASTON sources (Section 4.2).

^cSources in this catalogue require at least three consecutive detections in the 160-, 250-, 350-, and 500- μ m wavebands.

^dWe adopt the most conservative beam size for the band-merged (BM) Hi-GAL catalogue.

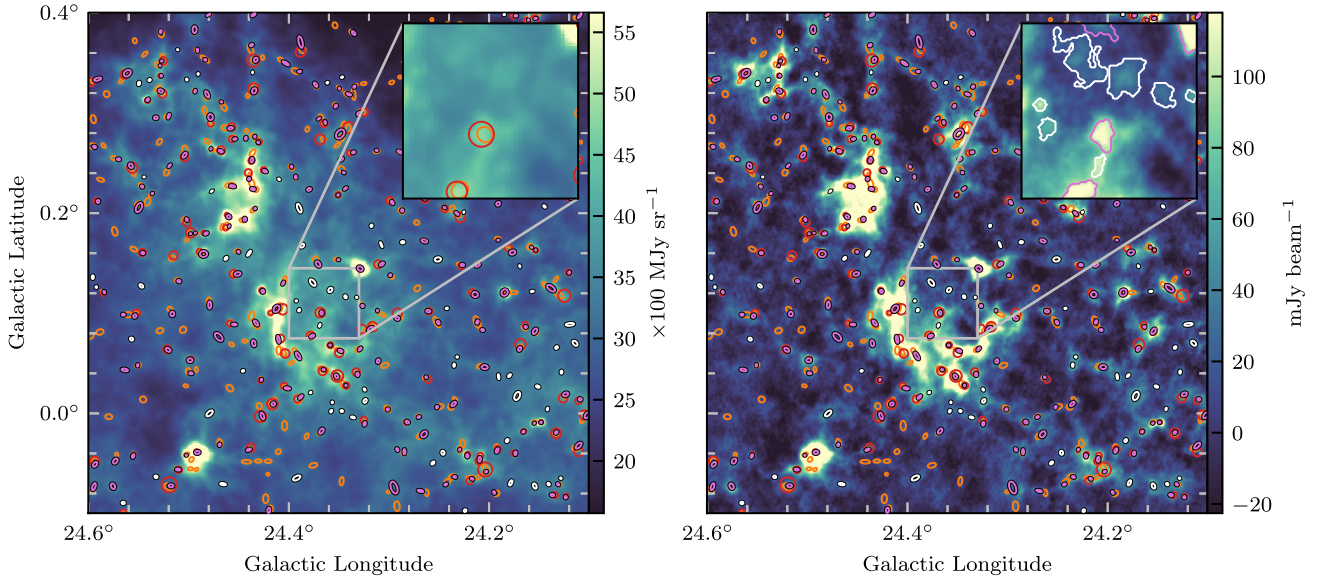


Figure 5. *Herschel*/Hi-GAL (Molinari et al. 2016a) 250- μm (left-hand panel) and NIKA2/GASTON 1.15-mm (right-hand panel) continuum emission maps for a $0.5^\circ \times 0.5^\circ$ region centred on $\ell = 24^\circ 35'$, $b = 0^\circ 15'$. The positions and sizes of the compact 250- μm sources from the Molinari et al. (2016a) Hi-GAL catalogue are shown as orange ellipses, and sources identified in the BM Hi-GAL catalogue of Elia et al. (2017) are shown as red circles. The positions of the 1.15-mm GASTON clumps are shown by markers with black borders: Those that have no cospatial 250- μm counterpart are given in white, while 1.15-mm sources with counterparts are shown as pink ellipses. Both panels contain an inlaid zoomed region, with *Herschel* markers only shown on the left-hand panel, while in the right-hand panel – for the sake of clarity – we show the boundaries of the GASTON clumps, adopting the same colour scheme as for the ellipses in the main image.

compact source catalogues. We performed a series of cross-matching exercises to identify any 1.15-mm GASTON clumps that have not been identified by either the five monochromatic source catalogues of Molinari et al. (2016a), or the ‘high-reliability’ BM catalogue of Elia et al. (2017).

To carry out the source matching, we adopted the elliptical source sizes from the monochromatic catalogues, and the circular source sizes for the BM catalogue – after convolving with a Gaussian kernel representative of the 1.15-mm NIKA2 beam – to look for overlaps with the footprint of each dendrogram structure, and record the number of matches in each case (see Table 2). By comparison with the monochromatic catalogues, we found the lowest fraction of GASTON clumps that do not match to any Hi-GAL compact source in the 250- μm band, with 321 GASTON clumps (22 per cent) without a 250- μm counterpart, and 765 (40 per cent) with no BM counterpart. We compare the *Herschel* 250- μm data and the GASTON GP 1.15-mm data for a $0.5^\circ \times 0.5^\circ$ region centred on $\ell = 24^\circ 35'$, $b = 0^\circ 15'$ in Fig. 5, overlaying markers representing catalogued sources from the 1.15-mm dendrogram extraction and the two Hi-GAL catalogues. The highest match fraction with sources in the 250- μm was expected, as these particular *Herschel* observations represent a compromise between the greater angular resolution available at the shorter wavelengths, and the greater sensitivity to the coldest structures at the longer wavelengths. Although the 500- μm catalogue represents the closest match in wavelength, the large difference in angular resolution results in a substantially different segmentation of the emission under the different source extraction methods.

In Fig. 6, we compare the distributions of mass and colour temperature, T_{col} , for the full sample of sources. Their statistical properties are summarized in Table 3. The new clumps identified by our NIKA2 observations with no counterpart in the Hi-GAL catalogues fall, on average, towards lower mass and tempera-

ture than those associated with *Herschel*-catalogued sources. We compare the $\log_{10}(M/M_\odot)$ and T_{col} distributions for the clumps that do have 250- μm Hi-GAL matches and those that do not, using two-sample Anderson–Darling. The results show that the distributions of both properties in the two populations are inconsistent at a high confidence level, with p -values of <0.02 in both cases.

There are 14 clumps with masses in excess of $100 M_\odot$ that are new detections. We note that to compare these values to other survey data, such as ATLASGAL, one must consider the masses measured in the same way, and the ATLASGAL clump masses of Urquhart et al. (2018) adopt a method more similar to the ‘bijection’ scheme (as discussed in Section 3.5). The approximate conversion between the ‘clipped’ masses reported in this study, and the equivalent bjected masses, calculated from combined *Herschel* and NIKA2 column density maps (following the method of Rigby et al. 2018) is $M_{\text{bijection}} \approx 41 \times M^{0.67}$, with a simple linear least-squares fit in log-space. In this metric, 33 of the 321 new clumps have $M_{\text{bijection}} > 1000 M_\odot$, and may be considered as high-mass. Of the 321 new clumps, 251 (80 per cent) have no compact 70- μm counterpart found by matching to the Molinari et al. (2016a) catalogue, and so may be considered as candidates for starless sources, though we caution that this is an upper limit, and a lack of a 70- μm match does not guarantee a lack of embedded star formation (e.g. Traficante et al. 2017).

We point out here that the various catalogues used in the cross-matching exercises in this Section and in the previous one, were generated using different source extraction techniques. We caution that we expect the total number of sources in the different data sets to vary with the methodology, but it is clear that we detect a substantial number of previously undetected sources, with the 321 sources that do not coincide with 250- μm compact sources representing the most conservative estimate.

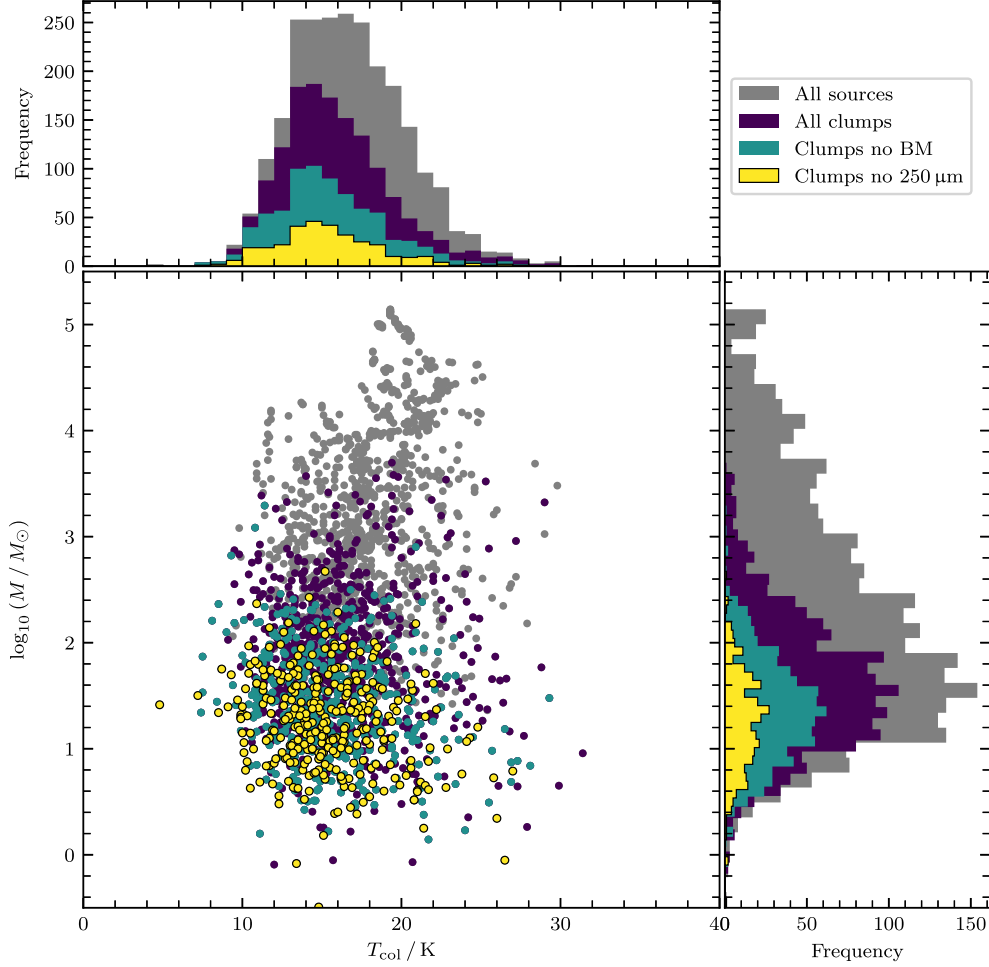


Figure 6. Mass and colour temperature distributions for all sources (grey), clumps (purple), clumps with no BM Hi-GAL counterpart (green), and clumps with no Hi-GAL 250- μ m counterpart (yellow).

Table 3. Statistical results – mean, standard deviation, sample size, standard error on the mean – of the masses and colour temperatures for the various samples presented in Fig. 6.

Sample	Mean	σ	N^a	σ/\sqrt{N}
$\log_{10}(M/M_{\odot})$				
All sources	2.24	1.08	2443	0.02
Clumps	1.60	0.63	1464	0.02
Clumps no BM	1.35	0.49	763	0.02
Clumps no 250 μ m	1.26	0.46	319	0.03
T_{col}/K				
All sources	16.66	3.59	2443	0.07
Clumps	15.84	3.56	1464	0.09
Clumps no BM	15.45	3.53	763	0.13
Clumps no 250 μ m	15.35	3.42	319	0.19

^aNote that three clumps have 160/250- μ m flux ratios that correspond to colour temperatures that outside of the range $2.7 \geq T_{\text{col}} \leq 50$ K, and so have no colour temperatures or masses derived here.

4.3 The evolution of star-forming clumps

To characterize our sample, we have computed the infrared-bright fraction for each source, f_{IRB} , based on GLIMPSE 8- μ m imaging (Benjamin et al. 2003; Churchwell et al. 2009). The infrared-bright

fraction is defined as the fraction of pixels within the source boundary that are brighter than the median value of all pixels within a 4.8-arcmin box centred on each pixel (see Watkins et al., in preparation for a full description). The filter size was selected so that the most extended sources within the 8- μ m image – typically bubbles associated with HII regions – could be compared to their local background. Although this filter scale represents a different spatial scale for the nearest and farthest sources within the region, it offers robust measurements for sources with angular sizes less than half of the filter scale, which is true for all GASTON clumps. We note that a very similar the classification scheme was used by Battersby et al. (2011), who used GLIMPSE 8- μ m images to classify a sample of Hi-GAL sources.

In Section 4.2, each clump was also matched to the 70- μ m catalogue of Molinari et al. (2016a) and we record the sum of the positionally matched 70- μ m source integrated flux densities. The 70- μ m luminosity is known to be a good tracer of the bolometric luminosity of any embedded sources (e.g. Dunham et al. 2008; Ragan et al. 2012). We calculate the bolometric luminosity by using the Elia et al. (2017) relationships:

$$\frac{L_{\text{bol}}}{L_{\odot}} = \begin{cases} 2.56 \left(\frac{S_{70}}{\text{Jy}} \right)^{1.00} \left(\frac{d}{\text{kpc}} \right)^2 & \text{if } S_{70} \geq 50 \text{ Jy,} \\ 3.29 \left(\frac{S_{70}}{\text{Jy}} \right)^{0.79} \left(\frac{d}{\text{kpc}} \right)^2 & \text{if } S_{70} < 50 \text{ Jy,} \end{cases} \quad (4)$$

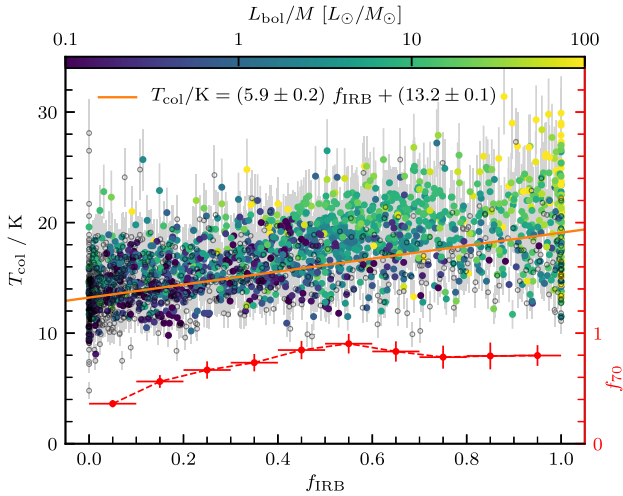


Figure 7. Colour temperature as a function of the infrared-bright fraction, f_{IRB} . The coloured points show the relation for all GASTON sources that have at least one matched Hi-GAL 70- μm compact source from Molinari et al. (2016a), and are colour-coded by the ratio of their bolometric luminosity to 1.15-mm mass. The empty circles show the relationship for unmatched GASTON sources. The trend determined by a linear least-squares fit is shown as the solid orange line. The red points show, on the secondary y-axis, the fraction of GASTON sources that are associated with compact 70- μm sources as a function of f_{IRB} in 0.1-wide bins.

where S_{70} is the sum of the integrated flux densities of any compact sources from the catalogue of Molinari et al. (2016a) that lie within one beam-radius of a GASTON clump boundary.

The infrared-bright fraction, f_{IRB} , may serve as an indicator for the evolutionary stage of a source: sources that are completely infrared-dark ($f_{\text{IRB}} = 0$) are likely to be at the earliest stages of evolution, containing no observable indications of star formation in terms of their 8- μm flux and, conversely, sources that are completely infrared-bright ($f_{\text{IRB}} = 1$) are likely to be sources exhibiting signs of advanced star formation, most commonly in the form of emission from young stellar objects, H II regions, or heated polycyclic aromatic hydrocarbon (PAH) molecules. Although a more detailed analysis will be presented in Watkins et al. (in preparation), we demonstrate the basic utility of f_{IRB} as an evolutionary indicator in several ways in Fig. 7. First, there is a positive correlation (Pearson correlation coefficient $\rho = 0.57$, $p\text{-value} \ll 0.1$) between f_{IRB} and the source temperature as traced by T_{col} . Secondly, there is also a positive correlation ($\rho = 0.57$, $p\text{-value} \ll 0.1$) between f_{IRB} and $\log_{10}(L_{\text{bol}}/M)$. Temperature and L_{bol}/M are both generally regarded as robust tracers of the time evolution of clumps (e.g. Urquhart et al. 2014, 2018; Molinari et al. 2016b; Svoboda et al. 2016; Elia et al. 2017). Finally, we show that the fraction of GASTON sources that are associated with compact 70- μm sources from Hi-GAL (Molinari et al. 2016a) also increases with f_{IRB} , and reaches a plateau of ~ 85 per cent for sources with $f_{\text{IRB}} \gtrsim 0.5$.

It is important to be aware of the potential for distance biases using the f_{IRB} parameter, since darkness at 8 μm usually requires a bright background against which the radiation can be absorbed. We therefore reduce our sample in the following analysis to include only those clumps that lie at distances between 3.5 and 7.0 kpc. This criterion serves a dual function; first, it limits the effect of the distance bias in f_{IRB} since infrared-dark clouds are generally located at the near kinematic distance (e.g. Ellsworth-Bowers et al. 2013), and the tangent for kinematic distances at this longitude range is ~ 7.5 kpc.

Secondly, this limits the role of a varying spatial resolution element in the source extraction process, and these particular limits mean that the most distant sources have only a factor of 2 lower spatial resolution than the nearest ones. Conveniently, these limits remove only ~ 10 per cent of the sample.

We define four categories through which we divide the sample of clumps based upon their infrared-bright fractions that can be used to broadly trace evolutionary phase: (i) $f_{\text{IRB}} < 0.040$ for predominantly infrared-dark clumps; (ii) $0.040 \geq f_{\text{IRB}} < 0.229$ for clumps that are mostly infrared-dark; (iii) $0.229 \geq f_{\text{IRB}} < 0.663$ for clumps that are partially infrared-bright; (iv) $f_{\text{IRB}} \geq 0.663$ for clumps that are predominantly infrared-bright. The specific values of f_{IRB} were chosen such that the four subsamples are of equal size, each containing a total of 321 clumps. The fraction of clumps in each category that are associated with compact 70- μm counterparts, f_{70} , also increases monotonically with the proposed evolutionary sequence, with 21, 33, 59, and 69 per cent of the clumps having 70 μm counterparts in samples (i)–(iv), respectively.

Tracing the precise time evolution of star-forming regions is not a trivial matter, but although we caution that f_{IRB} does not give an *absolute* value of ‘evolutionary stage’ or age for individual sources, it may be regarded as *relative* age indicator and, therefore, a suitable measurement when dealing with large samples of sources. Using a similar methodology to that used in this paper, Battersby, Bally & Svoboda (2017) estimated the fraction of a sample of dense molecular regions above a column density threshold associated with high-mass star formation that were starless or star-forming. By comparing the relative fractions of these categories to those associated with methanol masers and ultracompact H II regions – which have established lifetimes – the authors estimated the lifetime of the starless and star-forming phases as 0.2–1.7 and 0.1–0.7 Myr, respectively. These two time-scales may broadly map on to stages (i)–(ii) (predominantly starless) and (iii)–(iv) (predominantly star-forming), respectively, indicating a total duration of $\lesssim 2.5$ Myr for the evolution from $f_{\text{IRB}} = 0$ to $f_{\text{IRB}} = 1$.

In Fig. 8, we display these four samples in (logarithmic) mass and colour temperature space. By plotting the mean position of the four subsamples in the two axes, a progressive movement of the distributions is seen as the clumps get brighter at 8 μm . These mean values suggest that the clumps, on average, gain a moderate amount of mass – increasing by ~ 0.2 dex between stages (i) and (iii), reflecting a 60 per cent increase – before losing the same amount of mass, all the while increasing in temperature. The errors on the mean values indicate that although these changes in mass are subtle, they are significant at the 3.1σ level between stages (i) and (ii), at the 4.3σ level between stages (i) and (iii), and at the 4.0σ level of significance between stages (iii) and (iv). The increase in source counts as the GASTON GP project progresses will allow these values to be revisited with greater statistical power. The mean values and their associated standard errors are given in Table 4. A further striking feature of Fig. 8 is the increasing spread in temperatures along the evolutionary sequence. This behaviour is expected as a consequence of the more rapid evolution of the highest mass sources, in addition to the mass function. The more numerous low-mass sources are expected to evolve relatively slowly, only increasing in temperature, and losing mass at later stages as the clump is dispersed. The rarer high-mass clumps are expected to rapidly gain in mass and temperature at early stages, before losing mass at later stages.

This reflects the behaviour that might be expected from a clumped model of star formation, whereby the mass gain of the clump is the result of accretion from its direct environment, and the clump is

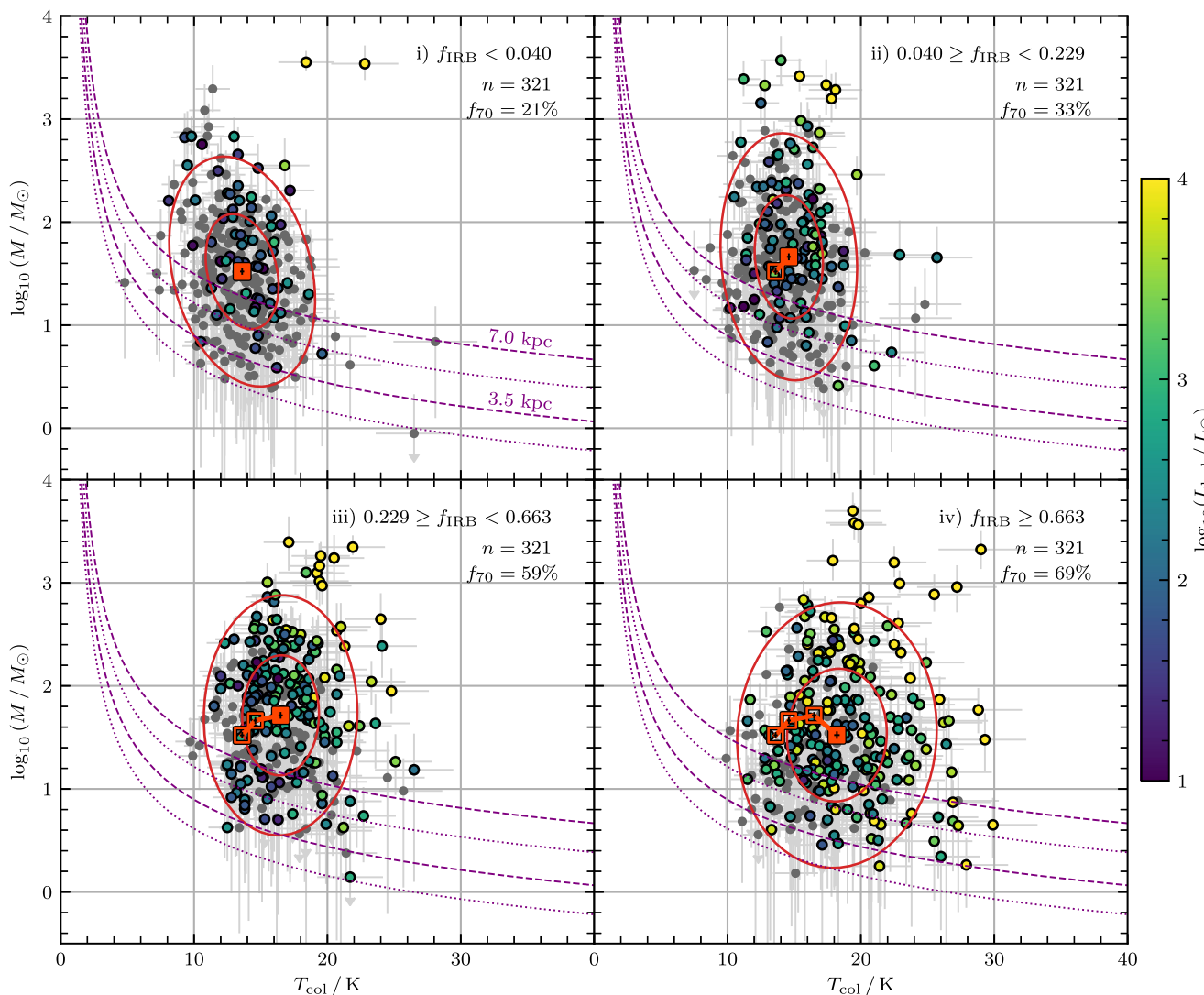


Figure 8. Mass–colour temperature relationships for the four subsamples of GASTON clumps, classified by their infrared-bright fraction, f_{IRB} . Masses are calculated using the ‘clipping’ method. Filled orange squares show the mean value of the quantities plotted in each axis, while the open squares show the mean values for the previous stages in the proposed evolutionary sequence. In each case, the standard error on the mean in both axes is shown in black within the marker shapes. The colour of the points shows the total bolometric luminosity from matched 70- μm compact sources, while the grey points have no compact 70- μm counterparts from Molinari et al. (2016a). Red ellipses show the approximate 1σ and 2σ confidence intervals for the distribution, containing ~ 68 and 95 per cent of the data points, respectively. The dashed and dotted purple curves correspond to the 1.15-mm point-like and extended source sensitivities for the combination of the data and source extraction parameters at the near- and far-distance limits of the sample.

Table 4. Mean values and associated standard errors of the distributions of the four evolutionary stages presented in Fig. 8.

Stage	$\log_{10}(M/M_{\odot})$		$T_{\text{col}} / \text{K}$	
	Mean	Error	Mean	Error
(i)	1.52	0.03	13.61	0.15
(ii)	1.66	0.03	14.60	0.14
(iii)	1.71	0.03	16.49	0.16
(iv)	1.52	0.04	18.20	0.21

subsequently dispersed by feedback as star formation progresses. By contrast, in a core-fed scenario, such clumps do not exhibit an initial increase in mass, since all of the mass is in situ before star formation begins, which seems incompatible with Fig. 8. Furthermore, the

movement of the mean points in Fig. 8 bears a striking resemblance to the mass–temperature tracks described by the clump-fed models of Peretto et al. (2020, see their fig. 7), supporting the coeval mass growth of clumps and embedded protostars within. However, it is important to note that the sizes of the GASTON sources presented here are, on average, larger than those presented in Peretto et al. (2020). This might suggest that the same accretion process from large to small scales occur over a relatively wide range of scales. A qualitatively similar behaviour as that seen in Fig. 8 is observed by Elia et al. (2017, see their fig. 22). In that study, clumps were categorized as starless or star-forming based upon the presence of 70- μm emission, with the star-forming sample further divided according to the detection or non-detection of 21–24 μm counterparts, and coincidence with H II regions.

We highlight that each of the four subsamples is likely to contain a mix of both core-fed and clump-fed sources, and so the signal

present here in the progressive movement of the mean position will be moderated by the core-fed sources toward the low-mass end of the mass distribution. If the clump-fed scenario is more likely to be important for the highest mass clumps, then this behaviour could be reflected in the upper envelope of the distributions. The 95th percentile in mass changes by +0.08, −0.03, and +0.06 dex, while the 95th percentile in temperature increases by 0.7, 2.7, and 4.4 K between stages (i)–(ii), (ii)–(iii), and (iii)–(iv), respectively. The lack of a reduction in the 95th percentile in logarithmic mass in the final stages is more difficult to interpret, as it suggests that the clump dispersal phase of the most massive clumps is either not being sampled well here, or lasts a relatively long time. The former explanation could easily be true, as the number of sources in the 95th percentile is only 16 in each stage, and thus the measurement suffers from small numbers. Urquhart et al. (2014) argued that, although the accretion phase for high-mass protostars is probably very rapid, they may then take a relatively long time to begin dispersing the enveloping clump, resulting in the clustering of sources at the apex of their mass–luminosity evolution. Qualitatively, this would result in the same behaviour seen here with these highest mass clumps spending much of their time in the high-mass, high-temperature part of this diagram. This is also supported by the clump-fed models of Peretto et al. (2020), whose evolutionary tracks for the highest mass stars shows an accretion phase that is relatively short-lived compared to the lifetime approaching their maximum mass.

It is unlikely that this sequence arises as a result of observational effects such as sensitivity or completeness. The point-source sensitivity limits are displayed in each panel of Fig. 8, and do not display any signs of correlation with the movement of the mean position of the distributions in stages (i)–(iii); lower mass clumps are more likely to be observable at later stages when temperatures are higher, and so this effect may in fact be suppressing mean value of the mass in the middle stages, and the average mass gain in high-mass sources is therefore likely to be greater than the value of 60 per cent deduced from this study.

The source extraction, which was carried out on the 1.15-mm map, is completely independent of the f_{IRB} parameter, and so we do not expect any systematic bias in terms of completeness as a function of evolutionary stage. The effects of mass completeness manifest themselves in this parameter space by the position of the peak of the distribution. The real distribution of clump masses is expected to consist of a power law $N \propto M^\alpha$, where $\alpha < 0$, and so the turnover close to the mean value reflects the completeness that is a function of the sensitivity and source extraction methodology. Since both of these factors are consistent across all samples, it is difficult to conceive a scenario in which differences in completeness could be dominating the position of the mean.

4.4 Caveats on the proposed evolutionary scenario

In the previous Section we have provided evidence, acquired from the catalogue of clumps identified at 1.15 mm, that the overall population is most massive at intermediate evolutionary stages, as traced by the f_{IRB} proxy for relative evolution. In this section, we explore various caveats to this analysis, and test the robustness of the observed trend.

First, our methods for acquiring clump distances in Section 3.4 have resulted in a large fraction – 93 per cent – of sources that are located at the near kinematic distance. We expect a high fraction along this line of sight, and it is, perhaps, not surprising that a higher fraction is recovered than for clumps in ATLASGAL (77 per cent), BGPS (79 per cent), or Hi-GAL (89 per cent of the ‘good-quality’ assignments), due simply to the greater mass sensitivity

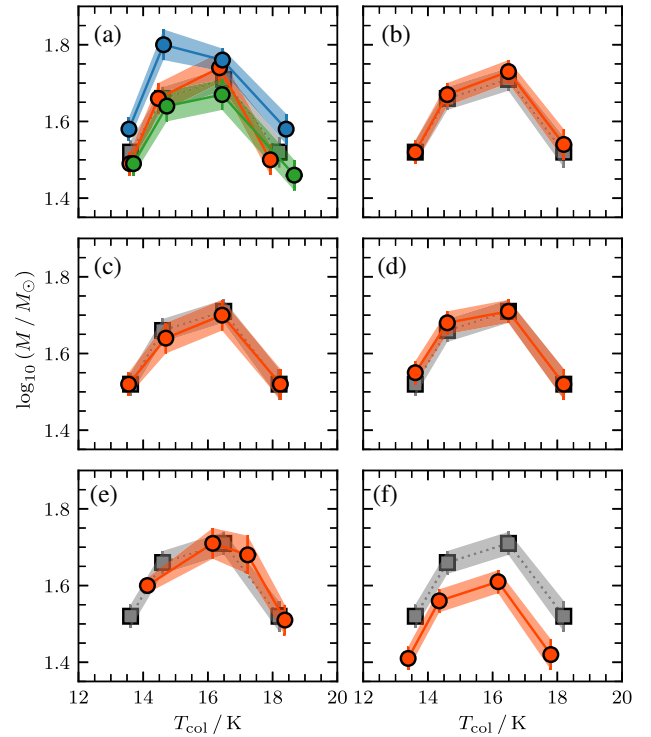


Figure 9. Mean evolutionary trend after altering the samples to (panel a) displace a randomly selected third of clumps to the far kinematic distance (using three sets); (panel b) adopt d_{group} as the distance estimate; (panel c) adjust the distance limits to $4 < d_{\text{cluster}} < 6$ kpc; (panel d) normalize the clump masses to the mean distance of 5.35 kpc; (panel e) adopt equally spaced f_{IRB} limits for the subsamples; and (panel f) use alternative values of β . In each case, the original trend is shown as grey squares, with mean positions of the altered samples overlaid in coloured circles. The shaded regions show the $\pm 1\sigma$ confidence limits.

of GASTON, and the greater prevalence of low-mass compared to high-mass clumps. However, we have tested the robustness of the proposed evolutionary trend by repeating the analysis three times after first randomly assigning 1/3 of the population of clumps to their far kinematic distance solutions. The resulting trends are illustrated in Fig. 9, and the overall effect is, in two of the three cases, to slightly reduce the significance of the differences between stages (i)–(ii), (i)–(iii), and (iii)–(iv) compared to the original distance assignments, while the significance is increased in the third case. This tendency to slightly reduce the significance can be mainly ascribed to the resulting smaller sample sizes, as the clumps now assigned to the far-kinematic distance fall outside the distance limits of the sample, consequently increasing the error on the mean.

Another choice affecting the results was the decision to adopt the cluster-averaged distances, d_{cluster} , as opposed to the dendrogram group-averaged distances, d_{group} . If we adopt the latter in place of the former, and repeat the analysis, we see almost no change in the results, since the average difference in the two distances is on the order of 100 pc. In fact, as can be seen in panel b) of Fig. 9, the evolutionary trend is slightly strengthened in this scenario with the largest change coming between stages (i) and (iii), now at the 4.5σ level of significance, compared to 4.3σ in the original analysis.

The choice of distance limits used could also have an effect upon the recovered trend, and we repeated the experiment by adopting more conservative distance limits, such that a maximum factor of

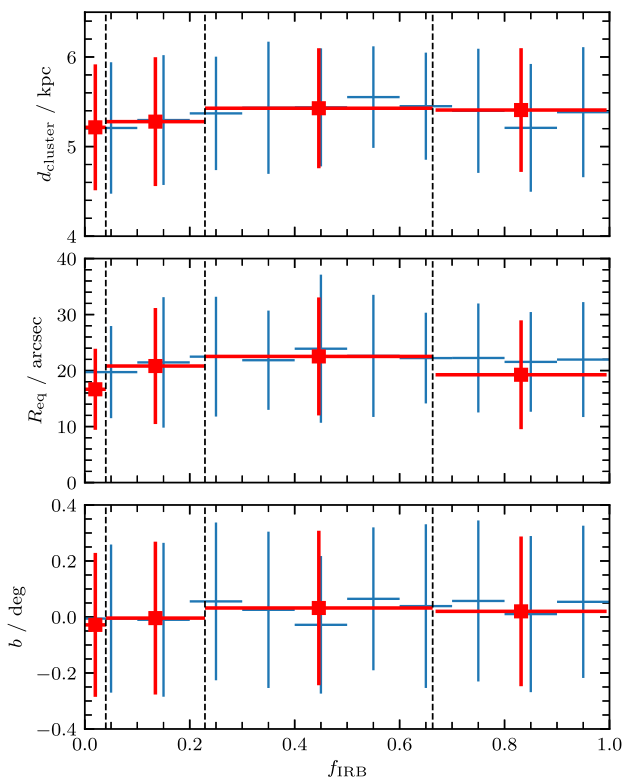


Figure 10. Average values of distance (top panel), angular size (middle panel), and centroid latitude (bottom panel) of GASTON clumps as a function of the infrared-bright fraction. The red data points show the mean values for samples (i)–(iv), and the blue data points show the mean values for clumps sampled in 0.1-wide bins of f_{IRB} . The vertical error bars indicate the standard deviation of the populations, and the horizontal error bars indicate the f_{IRB} range of the sample. Dashed vertical lines indicate the boundaries used for the samples (i)–(iv).

1.5 difference in the spatial resolution at the near and far cut-off was used, i.e. $4 < d_{\text{cluster}} < 6$ kpc. The effect here is, again, to reduce the sample size, now with 259 or 260 clumps in each stage compared to the original 321, and thereby reduce the significance of the trend. In this case, the largest change in $\log_{10}(M/M_{\odot})$ between stages (i) and (iii) is still significant at the 3.7σ level (panel c) of Fig. 9).

We also note the presence of potential biases in both distance and angular size as a function of f_{IRB} . Relative to one another, the average values of both quantities show the same behaviour as the average mass throughout the four proposed evolutionary stages. A bias towards nearer distances for the more infrared-dark sources can be understood by their requirement for a bright infrared background against which they can be seen, though it is not clear why there should be a bias for the brightest sources to also be located closer to the observer. The bias towards larger sources at intermediate values of f_{IRB} can be understood by the nature of the median filter, by which sources should tend towards $f_{\text{IRB}} = 0.5$ as their size approaches the filter size, by construction. We illustrate these biases in Fig. 10, indicating the mean values and standard deviations. One way to explore the effect of the distance bias is to normalize all of the clump masses to the same distance, and we do this by rescaling the masses as if they were located at the mean distance within the sample, multiplying them by a factor of $(5.35 \text{ kpc}/d_{\text{cluster}})^2$, and we include the modified evolutionary trend in panel d) of Fig. 9. The bias in distance is extremely mild, and has little effect upon our conclusions. However, the effect of the bias in angular size is more difficult to test,

and so we caution that there may be an effect at play here though, again, the effect is probably very mild, considering the overlapping distributions. In Fig. 10, we also show the lack of any significant bias in f_{IRB} as result of the latitude of the clumps, indicating that this particular line of sight has a sufficiently bright infrared-background so as to make any latitude-dependent bias negligible.

In Section 4.3, the limits to f_{IRB} were selected to equalize the size of the four samples, and thereby eliminate the statistical effects of varying sample sizes. Since f_{IRB} is expected to trace the relative evolution of each clump (however non-linearly), a secondary effect of this choice of samples is that a similar fraction of the clumps' total lifetimes may be covered in each stage. We can, however, investigate whether these particular boundaries of f_{IRB} that define the evolutionary stages, are responsible for the apparent sequence. Therefore, we conducted a further test, by defining the subsamples with equally spaced f_{IRB} limits at $f_{\text{IRB}} = 0.25, 0.50, 0.75$. The four samples (i)–(iv) now contain 656, 202, 148, and 248 clumps, and, again, the same general trend is present (panel e) of Fig. 9), though at a reduced significance, with 2.35σ , 1.57σ and 2.73σ changes in the value of $\log_{10}(M)$ between stages (i)–(ii), (i)–(iii), and (iii)–(iv), respectively. That equal time is spent – on average – in each stage. It is conceivable that other f_{IRB} limits could be adopted to either amplify or diminish the significance of the proposed evolutionary trend, but with a lack of physical or statistical justification, we explore these no further here.

Our calculation of temperatures and masses in Section 3.5, could also be having an effect upon the evolutionary trend. We adopted a value of $\beta = 1.8$, based upon the GP average value, as measured by *Planck* (Planck Collaboration XXV 2011), and thereby calculated a value of κ_{260} . However, the value of β has been shown to vary as a function of frequency (e.g. Planck Collaboration XIV 2014), and so we test the impact of a two-valued β upon the reported evolutionary trend. Following Planck Collaboration XIV (2014), we adopt a far-infrared value of $\beta_{\text{FIR}} = 1.88 \pm 0.08$ for our calculation of colour temperatures (as per equation 2), and a millimetre value of $\beta_{\text{mm}} = 1.60 \pm 0.06$, for the determination of κ_{260} used in the mass calculation (equation 3). In panel (f) of Fig. 9, it can be seen that the relative evolutionary trends are unaffected by these changes, with the exception that the value of $\log_{10}(M/M_{\odot})$ is systematically shifted downwards by ~ 0.1 , accompanied by a small reduction in temperature of $0.2\text{--}0.4$ K.

5 SUMMARY AND CONCLUSIONS

In this paper, we present the Galactic Star Formation with NIKA2 (GASTON) project, a guaranteed-time large programme using the IRAM 30 m telescope's new millimetre continuum camera, NIKA2. The GASTON large-programme consists of two projects aimed at probing the origin of different regions of the stellar initial mass function: (i) a survey of high- to intermediate-mass star-forming clumps within a 2 deg^2 region of the inner GP centred on $\ell = 23^{\circ}9$, $b = 0^{\circ}05$; (ii) a search for pre-brown dwarf cores; and a third project (iii) consisting of a study of dust property variations in nearby well-resolved cores.

We have described the observing strategy for the GASTON GP project, and presented the first results obtained primarily through analysis of the 1.15-mm photometric maps after ~ 40 per cent of the total integration time. A dendrogram extraction technique has allowed us to isolate resolved and unresolved compact sources as well as more extended structures, whilst maintaining a description of the hierarchy of emission complexes. By extracting ^{13}CO (1–0) spectra for each source using FUGIN (Umemoto et al. 2017) and

RAMPS (Hogge et al. 2018) data in combination with the latest models of the Galactic rotation curve (Reid et al. 2019), we have assigned kinematic distances to each structure and, in concert with Hi-GAL (Molinari et al. 2016a) data, have been able to calculate their masses.

Of the 2446 dendrogram structures that we have identified, a total of 1467 structures are either compact or unstructured (within the constraints of the observations) and 22 per cent of these appear to be a new population of clumps with no counterpart in the *Herschel* Hi-GAL catalogues (Molinari et al. 2016a; Elia et al. 2017). These new clumps are, on average, less massive, and colder than their 250- μm -detected counterparts, and represent a population of sources for which the *Herschel* resolution element at wavelengths in excess of 250 μm is not sensitive to. The majority of these sources are candidates for starless clumps, with 80 per cent having no compact 70- μm counterparts. By the end of the GASTON GP observing campaign, we expect this number to increase significantly.

We have proposed a categorization that describes the relative evolutionary stage of the clumps in terms of their infrared-bright fraction, f_{IRB} . The mean temperature of the clumps, along with the fraction of 70- μm -bright sources, and the ratio bolometric luminosity to mass in each stage increases monotonically with the proposed sequence, supporting the use of f_{IRB} as an evolutionary indicator. In mass–temperature space, the mean position of the distribution of clumps follows a path that agrees well with clump-fed scenarios for high-mass star formation, in which high-mass star-forming cores accrete mass from their parsec-scale environment, before finally losing mass at later stages as the dense gas initially associated with the star-forming region is dispersed and accreted.

Upon completion of the GASTON GP field, we will publish the full catalogue of sources, and revisit the measurements we have made in this study.

ACKNOWLEDGEMENTS

We wish to thank the anonymous referee for their helpful comments and suggestions that have improved the quality of this paper. AJR would like to thank his newborn son, Joseph, whose tardy arrival assisted with the completion of this work. AJR and NP would like to thank the STFC for financial support under the consolidated grant number ST/N000706/1. We would like to thank the IRAM staff for their support during the campaigns. The NIKA dilution cryostat has been designed and built at the Institut Néel. In particular, we acknowledge the crucial contribution of the Cryogenics Group, and in particular Gregory Garde, Henri Rodenas, Jean Paul Leggeri, Philippe Camus. This work has been partially funded by the Foundation Nanoscience Grenoble and the LabEx FOCUS ANR-11-LABX-0013. This work is supported by the French National Research Agency under the contracts ‘MKIDS’, ‘NIKA’, and ANR-15-CE31-0017 and in the framework of the ‘Investissements d’avenir’ program (ANR-15-IDEX-02). This work has benefited from the support of the European Research Council Advanced Grant ORISTARS under the European Union’s Seventh Framework Programme (Grant Agreement no. 291294). FR acknowledges financial supports provided by NASA through SAO Award Number SV2-82023 issued by the Chandra X-Ray Observatory Center, which is operated by the Smithsonian Astrophysical Observatory for and on behalf of NASA under contract NAS8-03060. A.B. acknowledges the support of the European Union’s Horizon 2020 research and innovation program under the Marie Skłodowska-Curie Grant agreement No. 843008 (MUSICA). This research made use of the PYTHON pack-

ages ASTROPY⁹ (The Astropy Collaboration et al. 2013, 2018), IPYTHON¹⁰ (Perez & Granger 2007), NUMPY¹¹, SCIPY¹² (Virtanen et al. 2020), MATPLOTLIB¹³ (Hunter 2007). This research has also made use of NASA’s Astrophysics Data System Bibliographic Services, TOPCAT¹⁴ (Taylor 2005), and SAOIMAGEDS¹⁵ (Joye & Mandel 2003).

DATA AVAILABILITY

In keeping with the IRAM Large Program policy,¹⁶ the GASTON data – including raw and processed products – will be made publicly available 18 months after the last scheduling semester in which these are observed.

REFERENCES

- Adam R. et al., 2018, *A&A*, 609, A115
 Aguirre J. E. et al., 2011, *ApJS*, 192, 4
 André P. et al., 2010, *A&A*, 518, L102
 André P., Di Francesco J., Ward-Thompson D., Inutsuka S.-I., Pudritz R. E., Pineda J. E., 2014, *Protostars and Planets VI*. University of Arizona Press, Tucson, AZ, p. 27
 André P., Arzoumanian D., Könyves V., Shimajiri Y., Palmeirim P., 2019, *A&A*, 629, L4
 Aniano G., Draine B. T., Gordon K. D., Sandstrom K., 2011, *PASP*, 123, 1218
 Battersby C. et al., 2011, *A&A*, 535, A128
 Battersby C., Bally J., Svoboda B., 2017, *ApJ*, 835, 263
 Beckwith S. V. W., Sargent A. I., Chini R. S., Guesten R., 1990, *AJ*, 99, 924
 Benjamin R. A. et al., 2003, *PASP*, 115, 953
 Bonnell I. A., Bate M. R., 2006, *MNRAS*, 370, 488
 Bontemps S., Motte F., Csengeri T., Schneider N., 2010, *A&A*, 524, A18
 Bourrion O. et al., 2016, *J. Instrum.*, 11, P11001
 Calvo M. et al., 2016, *J. Low Temp. Phys.*, 184, 816
 Chen H.-R. V., 2019, *ApJ*, 875, 24
 Churchwell E. et al., 2009, *PASP*, 121, 213
 Clarke S. D., Whitworth A. P., Duarte-Cabral A., Hubber D. A., 2017, *MNRAS*, 468, 2489
 Dame T. M., Hartmann D., Thaddeus P., 2001, *ApJ*, 547, 792
 Dempsey J. T., Thomas H. S., Currie M. J., 2013, *ApJS*, 209, 8
 Drabek E. et al., 2012, *MNRAS*, 426, 23
 Dunham M. M., Crapsi A., Evans N. J., II, Bourke T. L., Huard T. L., Myers P. C., Kauffmann J., 2008, *ApJS*, 179, 249
 Eden D. J. et al., 2020, *MNRAS*, 498, 5936
 Elia D. et al., 2017, *MNRAS*, 471, 100
 Ellsworth-Bowers T. P. et al., 2013, *ApJ*, 770, 39
 Ellsworth-Bowers T. P., Rosolowsky E., Glenn J., Ginsburg A., Evans N. J., II, Battersby C., Shirley Y. L., Svoboda B., 2015, *ApJ*, 799, 29
 Frayer D. et al., 2020, *BAAS*, 52, 272.17
 Ginsburg A. et al., 2013, *ApJS*, 208, 14
 Griffin M. J. et al., 2010, *A&A*, 518, L3
 Heyer M. et al., 2018, *MNRAS*, 473, 2222
 Hogge T. et al., 2018, *ApJS*, 237, 27
 Hunter J. D., 2007, *Comput. Sci. Eng.*, 9, 90
 Inutsuka S.-i., Miyama S. M., 1997, *ApJ*, 480, 681

⁹<https://astropy.org/>.

¹⁰<https://ipython.org/>.

¹¹<https://numpy.org/>.

¹²<https://scipy.org/>.

¹³<https://matplotlib.org/>.

¹⁴<http://www.star.bris.ac.uk/~mbt/topcat/>.

¹⁵<http://ds9.si.edu/>.

¹⁶<https://www.iram-institute.org/EN/content-page-243-7-158-243-0-0.html>.

- Jackson J. M. et al., 2006, *ApJS*, 163, 145
- Jackson J. M. et al., 2019, *ApJ*, 870, 5
- Joye W. A., Mandel E., 2003, in Payne H. E., Jedrzejewski R. I., Hook R. N., eds, ASP Conf. Ser. Vol. 295, *Astronomical Data Analysis Software and Systems XII*. Astron. Soc. Pac., San Francisco, p. 489
- Kirk H., Myers P. C., Bourke T. L., Gutermuth R. A., Hedden A., Wilson G. W., 2013, *ApJ*, 766, 115
- Klaassen P. et al., 2019, *BAAS*, 51, 58
- Könyves V. et al., 2015, *A&A*, 584, A91
- Könyves V. et al., 2020, *A&A*, 635, A34
- Lee Y.-N., Hennebelle P., Chabrier G., 2017, *ApJ*, 847, 114
- Liu T. et al., 2016, *ApJ*, 829, 59
- Lu X. et al., 2018, *ApJ*, 855, 9
- Luhman K. L., 2012, *ARA&A*, 50, 65
- Marsh K. A. et al., 2016, *MNEAS*, 459, 342
- McKee C. F., Tan J. C., 2003, *ApJ*, 585, 850
- Molinari S. et al., 2010, *A&A*, 518, L100
- Molinari S. et al., 2016a, *A&A*, 591, A149
- Molinari S., Merello M., Elia D., Cesaroni R., Testi L., Robitaille T., 2016b, *ApJ*, 826, L8
- Moore T. J. T. et al., 2015, *MNEAS*, 453, 4265
- Motte F., Bontemps S., Louvet F., 2018, *ARA&A*, 56, 41
- Myers P. C., 2009, *ApJ*, 700, 1609
- Peñaloza C. H., Clark P. C., Glover S. C. O., Shetty R., Klessen R. S., 2017, *MNEAS*, 465, 2277
- Peretto N., Fuller G. A., 2009, *A&A*, 505, 405
- Peretto N. et al., 2013, *A&A*, 555, A112
- Peretto N. et al., 2014, *A&A*, 561, A83
- Peretto N., Lenfestey C., Fuller G. A., Traficante A., Molinari S., Thompson M. A., Ward-Thompson D., 2016, *A&A*, 590, A72
- Peretto N. et al., 2020, *MNRAS*, 496, 3482
- Perez F., Granger B. E., 2007, *Comput. Sci. Eng.*, 9, 21
- Perotto L. et al., 2020, *A&A*, 637, A71
- Planck Collaboration XXV, 2011, *A&A*, 536, A25
- Planck Collaboration XIV, 2014, *A&A*, 564, A45
- Poglitsch A. et al., 2010, *A&A*, 518, L2
- Ragan S. et al., 2012, *A&A*, 547, A49
- Reid M. J., Dame T. M., Menten K. M., Brunthaler A., 2016, *ApJ*, 823, 77
- Reid M. J. et al., 2019, *ApJ*, 885, 131
- Rigby A. J. et al., 2018, *A&A*, 615, A18
- Roman-Duval J., Jackson J. M., Heyer M., Rathborne J., Simon R., 2010, *ApJ*, 723, 492
- Rosolowsky E. W., Pineda J. E., Kauffmann J., Goodman A. A., 2008, *ApJ*, 679, 1338
- Ruppin F. et al., 2018, *A&A*, 615, A112
- Sanhueza P., Jackson J. M., Zhang Q., Guzmán A. E., Lu X., Stephens I. W., Wang K., Tatematsu K., 2017, *ApJ*, 841, 97
- Schneider N. et al., 2012, *A&A*, 540, L11
- Schuller F. et al., 2009, *A&A*, 504, 415
- Stanke T. et al., 2019, *BAAS*, 51, 542
- Svoboda B. E. et al., 2016, *ApJ*, 822, 59
- Tan J. C., Beltrán M. T., Caselli P., Fontani F., Fuente A., Krumholz M. R., McKee C. F., Stolte A., 2014, *Protostars and Planets VI*. University of Arizona Press, Tucson, AZ, p. 149
- Taylor M. B., 2005, in Shopbell P., Britton M., Ebert R., eds, ASP Conf. Ser. Vol. 347, *Astronomical Data Analysis Software and Systems XIV*. Astron. Soc. Pac., San Francisco, p. 29
- Terebey S., Chandler C. J., Andre P., 1993, *ApJ*, 414, 759
- The Astropy Collaboration et al., 2013, *A&A*, 558, A33
- The Astropy Collaboration et al., 2018, *AJ*, 156, 123
- Traficante A., Fuller G. A., Billot N., Duarte-Cabral A., Merello M., Molinari S., Peretto N., Schisano E., 2017, *MNRAS*, 470, 3882
- Umemoto T. et al., 2017, *PASJ*, 69, 78
- Urquhart J. S. et al., 2014, *A&A*, 568, A41
- Urquhart J. S. et al., 2018, *MNRAS*, 473, 1059
- Vázquez-Semadeni E., Gómez G. C., Jappsen A.-K., Ballesteros-Paredes J., Klessen R. S., 2009, *ApJ*, 707, 1023
- Vázquez-Semadeni E., Palau A., Ballesteros-Paredes J., Gómez G. C., Zamora-Avilés M., 2019, *MNRAS*, 490, 3061
- Virtanen P. et al., 2020, *Nat. Methods*, 17, 261
- Wang P., Li Z.-Y., Abel T., Nakamura F., 2010, *ApJ*, 709, 27
- Whitworth A., Bate M. R., Nordlund Å., Reipurth B., Zinnecker H., 2007, *Protostars and Planets V*. University of Arizona Press, Tucson, AZ, p. 459
- Williams G. M., Peretto N., Avison A., Duarte-Cabral A., Fuller G. A., 2018, *A&A*, 613, A11

APPENDIX A: NULL MAPS

We display the 1.15-mm null map, used in Section 3.2, in Fig. A1.

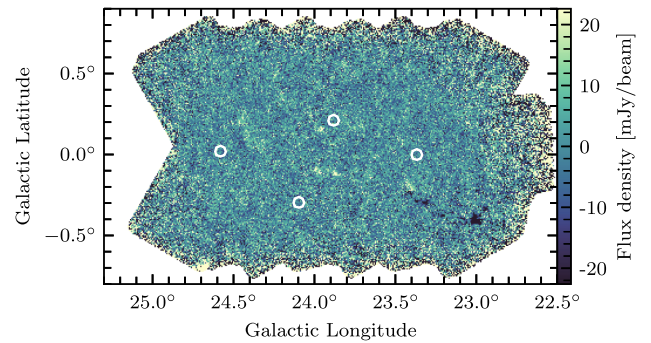


Figure A1. Null map for the GASTON GP field at 1.15 mm. The positions of the four sky apertures are shown as white polygons.

APPENDIX B: BASIC SOURCE PROPERTIES

In Section 3.2, we described the dendrogram-based procedure used to identify discrete sources within the 1.15-mm data, and here we describe their basic properties. We have measured two types of integrated flux density, termed ‘bijected’ and ‘clipped’, as in the scheme of Rosolowsky et al. (2008). The bijected integrated flux density of a source is calculated from the sum of pixel values within the source. In contrast, the clipped integrated flux density is a background-subtracted measurement, in which an estimate of the local background (as described in Section 3.5) is first subtracted from each pixel value before calculating their sum. We further define a contrast parameter for the clipped flux densities, which is the ratio of the peak to the background flux density value, and has a value of unity for faint sources at the detection threshold of our dendrogram extraction. None of the sources in our dendrogram extraction have a contrast value of less than unity, although, in principle, point sources located in regions with a bright background could result in values of less than 1.

Fig. B1 illustrates the relationship between peak flux density, and the two types of integrated flux density. The relationship between peak and bijected flux densities shows that the majority of sources lie above the line of equality, which isolated point sources at very high S/N would be expected to follow. Sources lying above this line

are extended sources, as can be seen by the colour scale indicating the sources’ equivalent radii, R_{eq} (the radius of a circle with the equivalent area). In the case of the clipped flux densities, sources lying in regions with significant backgrounds may lie below the line of equality, and indeed the most significant departures occur for sources located in the region around $\ell = 23^\circ 4$, $b = -0^\circ 2$. The minimum expected clipped integrated flux density is 4.1 mJy, for point sources at the detection threshold, which could occur at any value of peak intensity if the local background is sufficiently bright. Some sources lie just below this line, which may occur where sources are either not point-like, or have local noise levels slightly higher than the adopted average rms in the extraction parameters. The relationship between clipped and bijected integrated intensities is also shown, in which it can be seen that sources with the highest contrast, i.e. sources in regions with relatively low backgrounds, lie closer to the 1:1 relation. A linear least-squares fit in log-space finds an average relationship between the clipped and bijected flux densities of $S_{\text{bijected}} = 9.6 S_{\text{clipped}}^{0.8}$.

In Fig. B2, we display histograms of the peak and integrated flux densities, and equivalent radii of all sources identified by the dendrogram. The distributions of further subsamples are shown, and we show the population of clumps (sources containing no further observed substructures), and complexes (dendrogram ‘trunks’ containing more than one clump). Note that in the case of the peak flux densities, the peak flux within a complex is the same value as

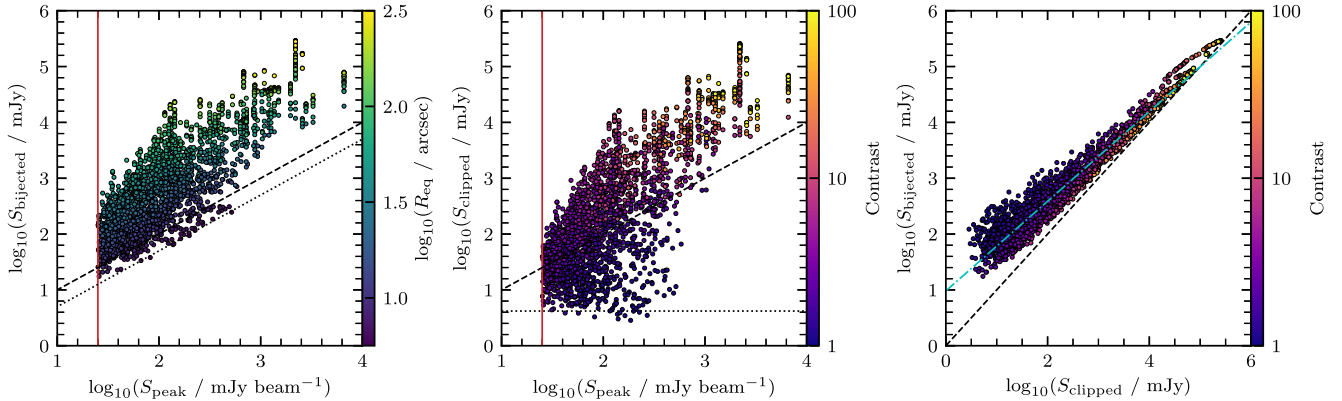


Figure B1. Left-hand panel: peak versus integrated (bijected) flux density for all sources, colour-coded according to sources’ equivalent radii. The dashed 1:1 line shows the expected relationship for isolated point sources in all panels. The dotted line shows the lower bound of the data, where the integrated flux density is equal to half of the expected value for a point source. Middle panel: peak versus integrated (clipped) flux density for all sources, colour-coded according to the contrast parameter (see the text). The dashed black line again shows the expected relationship for isolated point sources. The dotted line shows the minimum integrated flux density of a point source at the detection threshold. Right-hand panel: clipped versus bijected integrated flux density, coloured by contrast parameter.

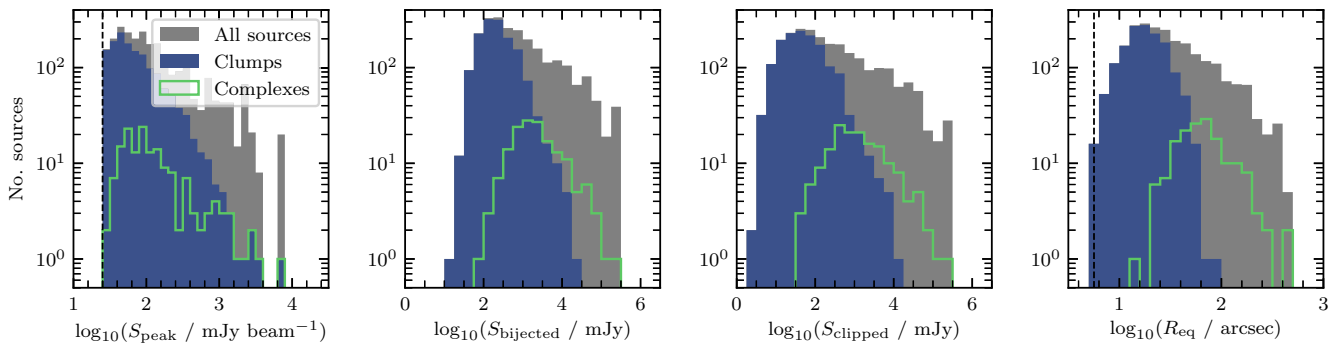


Figure B2. Distributions of various source properties, separated into all sources (grey), clumps (sources containing no discernible substructure), and complexes (sources containing two or more clumps).

the peak flux of its brightest constituent clump, and so the complex distribution is a subset of the clump distribution. The brightest clumps have peak flux densities of $\sim 10 \text{ Jy beam}^{-1}$, and the largest complexes have integrated flux densities of $\sim 100 \text{ Jy}$. Source sizes range from the minimum detectable size of $R_{\text{eq}} = 5.6 \text{ arcsec}$ to 6.5 arcmin .

APPENDIX C: VELOCITY ASSIGNMENTS

In Section 3.3, we derived an initial velocity assignment for each source based on the channel within a source area-averaged spectrum with the maximum intensity. A level of robustness for each velocity assignment was estimated by determining the standard deviation of the velocities associated with the three brightest channels within each spectrum. In cases where this standard deviation is less than 1 km s^{-1} (i.e. the three brightest channels are almost perfectly adjacent), the assignment is considered to be robust and assigned a flag value 3; where the standard deviation is more than 5 km s^{-1} , the assignment is considered to be poor and assigned a flag value of 1; cases between those limits are assigned an intermediate robustness flag value of 2. In practice, these flags allow the most robust velocities – where there is a single dominant peak – to be isolated from cases with either a low S/N or from spectra in which there are multiple velocity components of similar strength. In this initial velocity assignment, we record 1680 (69 per cent) of sources with a robustness flag of 3; 3305 (12 per cent) sources with an intermediate robustness flag, and 461 (19 per cent) with a poor robustness flag.

Following this initial velocity assignment, we make further velocity estimates by using the more appropriate – but limited in coverage – data of the NH_3 (1,1) inversion transition from the RAMPS survey. The RAMPS pilot study (Hogge et al. 2018) presented, as data products, maps of source velocities derived from a line-fitting procedure, as well maps of the first moment, with the latter having a greater extent, but the former being more accurate. For each GASTON source, and for each RAMPS velocity map, we record the mean velocity of pixels that fall within the source area. After excluding measurements where either the mean error (in the case of the fitted first components) or the standard deviation on the mean velocities (in the case of the first moments) are greater than 5 km s^{-1} , we record a total of 523 first-component velocities and 1097 first moment velocities. Where available, we replace the ^{13}CO -derived v_{cen} values with their NH_3 -derived counterparts, preferring the line-fitted measurements over the first moment-derived velocities, and adopt a robustness flag of 3. At this stage, the v_{cen} assignments are made up of 1407 (58 per cent) ^{13}CO -derived velocities, 516 (21 per cent) with NH_3 first moment-derived velocities, and the remaining 523 (21 per cent) with velocities derived from the NH_3 fitted first velocity components. There are 1913 (78 per cent), 174 (7 per cent), and 359 (15 per cent) of v_{cen} assignments with quality flags of 3, 2, and 1, respectively. Of the NH_3 -derived velocities, 888 (85 per cent) agree with their initial ^{13}CO -derived assignment to better than 5 km s^{-1} , indicating that the initial assignments are generally reliable.

In Fig. C1, we illustrate the method used in Section 3.3 for the initial velocity assignments for each of the sources identified in Sec-

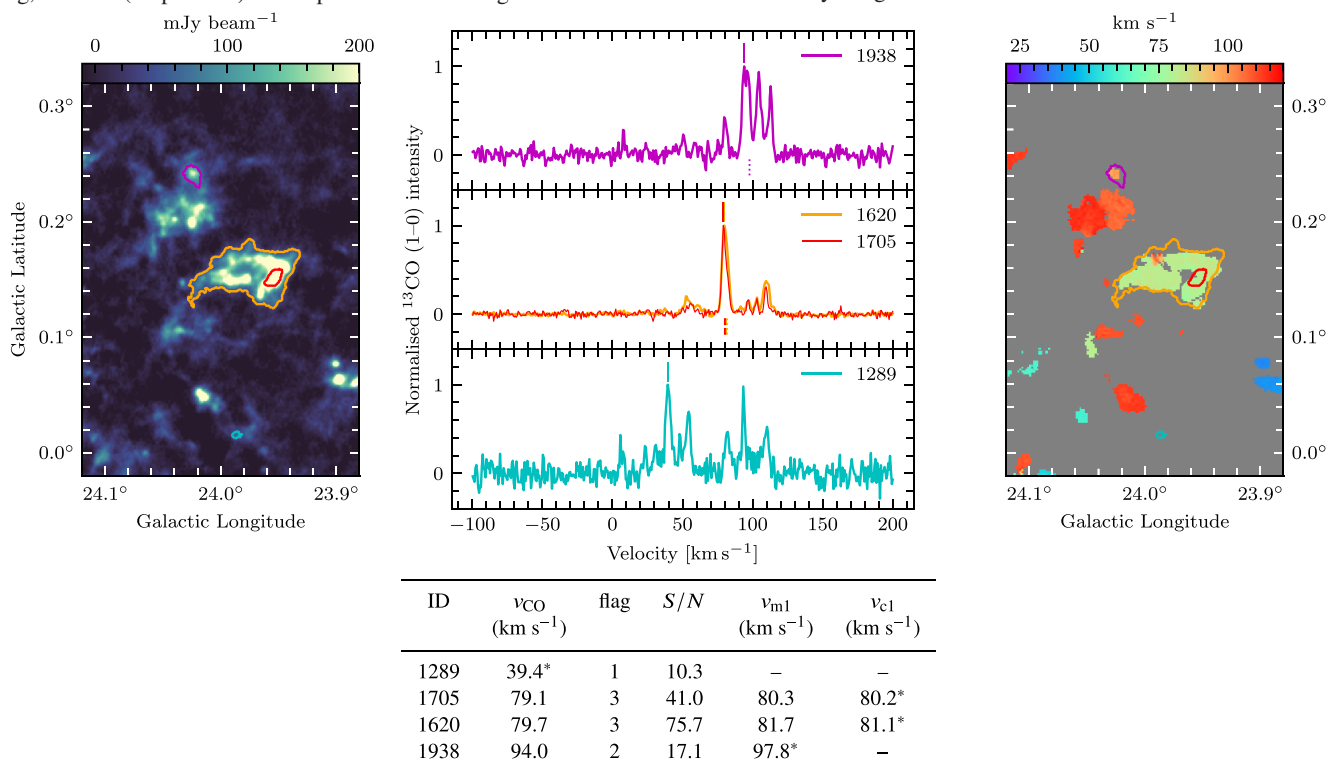


Figure C1. Example of the velocity determinations of four sources within a subset of the data. In the left-hand panel, the 1.15-mm flux density map is shown (after smoothing to 13-arcsec resolution), with contours overlaid corresponding to the footprints of catalogued sources with ID numbers 1289, 1620, 1705, and 1938. In the central panel, the corresponding ^{13}CO (1–0) spectra are presented, which were extracted over the footprint of the sources, with the velocity associated with the maximum intensity is assigned to that source indicated by a solid line above the spectra, and any velocity information determined from either the first-fitted velocity component or first moment maps of the RAMPS data are indicated by dot-dashed and dotted lines below the spectra. The right-hand image shows the RAMPS NH_3 (1,1) first moment (intensity weighted coordinate) map from Hogge et al. (2018), with the same source outlines overlaid. In the table, details of the ^{13}CO -derived velocity (v_{CO}), initial velocity flag, the S/N of the ^{13}CO (1–0) spectrum, and the source velocities derived from the RAMPS NH_3 (1,1) first moment (v_{m1}) and first velocity component (v_{c1}), are given for each source. In each case, an asterisk denotes the adopted velocity, v_{cen} .

tion 3.2. Where possible, we also display the velocity derived from the available RAMPS data. At this point, we refer the reader back to the main text of Section 3.3 for a description of how the emission structures are used to refine the individual velocity measurements.

APPENDIX D: LINE CONTAMINATION

The 1.15-mm NIKA2 band spans a range in frequency from 210–300 GHz (taking these edges as 20 per cent transmission), and as such, the continuum may contain important contributions from the CO (2–1) line at 230.538 GHz, especially for observations of the inner GP. To estimate the level of this line contamination in the 1.15-mm maps, we made use of CO (1–0) data from the FOREST unbiased GP imaging survey with the Nobeyama 45-m telescope (FUGIN; Umemoto et al. 2017). FUGIN covers the entire GASTON GP region, with an angular resolution of 20 arcsec.

The CO (1–0) data were integrated over the full velocity range of -50 to 200 km s^{-1} , and multiplied by a line ratio of $R_{2-1/1-0} = 0.7$ to provide an approximate conversion to CO (2–1) that is appropriate for cold and dense molecular gas (Peñaloza et al. 2017). The data were then converted from units of integrated intensity (K km s^{-1}) into surface brightness units (Jy beam^{-1}) following the procedure outlined in Drabek et al. (2012), which accounts for the NIKA2 filter response. Finally, we applied a 6.5-arcmin filter to the CO data to approximate the spatial filtering applied by the NIKA2 data reduction pipeline, using the NEBULISER application. Both the CO (2–1) and NIKA2 1.15-mm surface brightness maps were smoothed to a common resolution of 22 arcsec, and we measured their ratio for pixels that had a 1.15-mm S/N of more than 10.

The level of CO (2–1) line contamination in the valid pixels ranges from 0 to ~ 15 per cent, with a median value of 4 per cent. The upper limit of this measurement is most likely to be a worst-case scenario, and would apply in cases where NIKA2 is fully recovering emission up to 6.5-arcmin scales. Furthermore, the adopted $R_{2-1/1-0}$ value is the higher value from what is usually measured as a bimodal distribution, with a lower value of 0.3 often found in regions of sub-thermally excited low density CO. We do not, therefore, directly account for CO (2–1) line contamination in our analysis, but incorporate an additional factor of 4 per cent as the associated uncertainty on the flux density.

We note that these values are comparable to those estimated by (Rigby et al. 2018) for NIKA, who estimated CO contamination at the 1–3 per cent level. A slightly higher value is found here for a number of reasons: (i) The FUGIN data allow us to estimate the contamination at higher angular resolution; (ii) the FUGIN data allow us to estimate this quantity in the correct isotopologue of CO whereas the Rigby et al. (2018) study used an approximate conversion from ^{13}CO , which is less affected by optical depth effects; (iii) larger spatial frequencies are able to be recovered with NIKA2 than with NIKA, and so more large-scale CO emission survives the pipeline’s filtering.

APPENDIX E: DISTANCE COMPARISON

In Fig. E1, we compare the distances determined for GASTON clumps in Section 3.4 with the distances to their closest positionally-matched counterparts from the BGPS, ATLASGAL, and Hi-GAL surveys, whose catalogues were cross-matched in Section 4.1. We find that 64 and 69 per cent of GASTON clumps have distances that differ by less than 1 kpc when compared to the clump catalogues from the BGPS (Ellsworth-Bowers et al. 2015) and ATLASGAL (Urquhart et al. 2018), respectively. While only 25 per cent of distance determinations agree with Hi-GAL clumps from the BM catalogue of Elia et al. (2017), this value rises to 67 per cent once sources without a ‘good’ distance quality flag are removed.

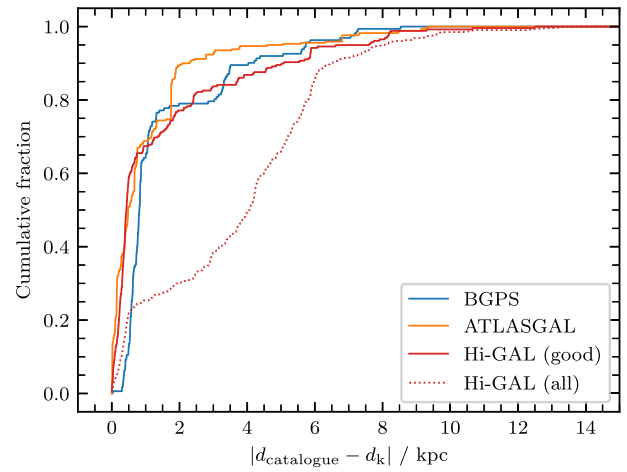


Figure E1. Cumulative distribution showing the difference in distances determined for GASTON clumps and their counterparts from the BGPS (blue), ATLASGAL (orange), and Hi-GAL (red) catalogues.

²LLR (Laboratoire Leprince-Ringuet), CNRS, École Polytechnique, Institut Polytechnique de Paris, F-91764 Palaiseau, France

³AIM, CEA, CNRS, Université Paris-Saclay, Université Paris Diderot, Sorbonne Paris Cité, F-91191 Gif-sur-Yvette, France

⁴Univ. Grenoble Alpes, CNRS, IPAG, F-38000 Grenoble, France

⁵Institut d’Astrophysique Spatiale (IAS), CNRS and Université Paris Sud, F-91400 Orsay, France

⁶Institut Néel, CNRS and Université Grenoble Alpes, F-38400 Saint-Martin-d’Hères, France

⁷Institut de RadioAstronomie Millimétrique (IRAM), Grenoble, F-38400 Saint-Martin-d’Hères, France

⁸Univ. Grenoble Alpes, CNRS, Grenoble INP, LPSC-IN2P3, 53, avenue des Martyrs, F-38000 Grenoble, France

⁹Rudjer Bošković Institute, Bijenička cesta 54, 10000 Zagreb, Croatia

¹⁰Dipartimento di Fisica, Sapienza Università di Roma, Piazzale Aldo Moro 5, I-00185 Roma, Italy

¹¹Chinese Academy of Sciences South America Center for Astronomy, National Astronomical Observatories, CAS, Beijing 100101, China

¹²Instituto de Astronomía, Universidad Católica del Norte, Av. Angamos 0610, Antofagasta 1270709, Chile

¹³Centro de Astrobiología (CSIC-INTA), Torrejón de Ardoz, E-28850 Madrid, Spain

¹⁴Instituto de Radioastronomía Milimétrica (IRAM), E-18012 Granada, Spain

¹⁵Aix Marseille Univ, CNRS, CNES, LAM (Laboratoire d’Astrophysique de Marseille), E-13284 Marseille, France

¹⁶LERMA, Observatoire de Paris, PSL Research University, CNRS, Sorbonne Universités, UPMC Univ. Paris 06, F-75014 Paris, France

¹⁷School of Earth and Space Exploration and Department of Physics, Arizona State University, Tempe, AZ 85287, USA

¹⁸Univ. Toulouse, CNRS, IRAP, 9 Av. du Colonel Roche, BP 44346, E-31028 Toulouse, France

¹⁹Département de Physique, Ecole Normale Supérieure, 24, rue Lhomond, F-75005 Paris, France

²⁰Department of Physics and Astronomy, University of Pennsylvania, 209 South 33rd Street, Philadelphia, PA 19104, USA

²¹Institut d’Astrophysique de Paris, Sorbonne Université, CNRS UMR7095, F-75014 Paris, France

²²Kavli Institute for Astrophysics and Space Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

²³Astronomisches Rechen-Institut, Zentrum für Astronomie der Universität at Heidelberg, Mönchhofstraße 12-14, D-69120 Heidelberg, Germany

This paper has been typeset from a \LaTeX file prepared by the author.