Astronomy & Astrophysics

The coordinated radio and infrared survey for high-mass star formation

III. A catalogue of northern ultra-compact H II regions⋆

I. E. Kalcheva[1], M. G. Hoare[1], J. S. Urquhart[2], S. Kurtz[3], S. L. Lumsden[1], C. R. Purcell[4,5], and A. A. Zijlstra[6]

[1] School of Physics and Astronomy, University of Leeds, Leeds LS2 9JT, UK
   e-mail: pyiek@leeds.ac.uk
[2] Centre for Astrophysics and Planetary Science, University of Kent, Canterbury CT2 7NH, UK
[3] Instituto de Radioastronomía y Astrofísica, Universidad Nacional Autónoma de México, 58089 Morelia, Michoacán, Mexico
[4] Research Centre for Astronomy, Astrophysics, and Astrophotonics, Macquarie University, NSW 2109, Australia
[5] Sydney Institute for Astronomy, School of Physics, The University of Sydney, NSW 2006, Australia
[6] Jodrell Bank Centre for Astrophysics, Alan Turing Building, School of Physics and Astronomy, The University of Manchester, Oxford Road, Manchester M13 9PL, UK

ABSTRACT

A catalogue of 239 ultra-compact H II regions (UCHIIs) found in the CORNISH survey at 5 GHz and 1.5″ resolution in the region $10° < l < 65°$, $|b| < 1°$ is presented. This is the largest complete and well-selected sample of UCHIIs to date and provides the opportunity to explore the global and individual properties of this key state in massive star formation at multiple wavelengths. The nature of the candidates was validated, based on observational properties and calculated spectral indices, and the analysis is presented in this work. The physical sizes, luminosities and other physical properties were computed by utilising literature distances or calculating the distances whenever a value was not available. The near- and mid-infrared extended source fluxes were measured and the extinctions towards the UCHIIs were computed. The new results were combined with available data at longer wavelengths and the spectral energy distributions (SEDs) were reconstructed for 177 UCHIIs. The bolometric luminosities obtained from SED fitting are presented. By comparing the radio flux densities to previous observational epochs, we find about 5% of the sources appear to be time variable. This first high-resolution area survey of the Galactic plane shows that the total number of UCHIIs in the Galaxy is ∼750 – a factor of 3–4 fewer than found in previous large area radio surveys. It will form the basis for future tests of models of massive star formation.

Key words. stars: formation – HII regions – radio continuum: ISM

1. Introduction

The puzzle of the birth and early life of stars exceeding 8 $M_\odot$ is not yet fully assembled. Some of the obstacles towards building a complete evolutionary sequence for these massive stars include their rarity due to their brief lifetime and the rapid evolution of each observable stage. The main sequence is reached while the young star is still embedded within a dense core and as a result the early phases of its development are hidden behind a heavy veil of dust. A well-founded distinction between global and individual properties of sources in each evolutionary stage is hampered by the strong influence of other objects within the multiple systems where massive stars typically form.

It is a vital task for modern astronomy to overcome these challenges. Massive stars affect not only their immediate surroundings, but also shape their parent galaxy. Their formation controls phase changes in the interstellar medium (ISM) via the profuse emission of ionising UV photons (Molinari et al. 2014). Processes associated with their evolution, such as winds, outflows, expanding H II regions and supernovae, stir the ISM

and enrich it with heavy elements (Zinnecker & Yorke 2007). This makes their understanding a stepping stone towards a more detailed picture of the Milky Way, as well as the extent to which galaxy formation and evolution in general is driven by stellar populations.

After a massive star has formed, it ionises a pocket of hydrogen gas which remains confined in its vicinity while expanding – thus forming an H II region. H II regions are highly convenient tracers of massive star formation, as they are clearly visible across the Galactic plane in the cm-regime (Anderson & Bania 2009). Distinguished as a separate observational class by Wood & Churchwell (1989a), ultra-compact H II (UCHII) regions link the accretion phase when a massive protostar is formed, and the development of a more diffuse and less obscured H II region. UCHII regions are defined as embedded photoionised regions ≲0.1 pc in diameter, with emission measures ≳10^7 pc cm^{-6} and electron densities $n_e \gtrsim 10^4$ cm^{-3} (Wood & Churchwell 1989a). They are the most luminous objects in the Milky Way in the far-IR, and are observable in the radio part of the spectrum if their luminosities are equivalent or higher than a B0.5 main-sequence star. The Lyman continuum ionising flux corresponding to zero-age main-sequence stars with spectral class from B2 to O5 is in the range 10^{44}–10^{49} photons s^{-1}. Estimating the distances to UCHII regions, together with their density distributions,

luminosities, morphologies, kinematics, and relationship to the parent molecular clouds is essential. These properties can be used to help understand not only the effect of UCHII regions on their environment, but also test the existing evolutionary models of massive star formation and the structure of the Milky Way (Hoare et al. 2007).

The bounds of current understanding of massive star formation are widened by the modern family of Galactic plane surveys, covering the dust (from hot to cold), the molecular and the ionised gas. These include the GLIMPSE programme (Churchwell et al. 2009; Benjamin et al. 2003) and its companion MIPSGAL survey (mid-IR; Carey et al. 2009), the UKIDSS GPS survey (near-IR; Lucas et al. 2008), the BU-FCRAO Galactic Ring Survey (CO; Jackson et al. 2006), the ATLASGAL survey (sub-mm; Schuller et al. 2009), the VGPS survey (H I; Stil et al. 2006), the CORNISH survey (radio; Hoare et al. 2012, Purcell et al. 2013). These legacy surveys provide resolution and sensitivity apposite to the detection and discerning of sources occupying angular scales down to ∼1″. At the same time, they cover wide areas on the sky and overcome the high extinction of the plane. In this way, a multi-wavelength treasure trove of unbiased, high-resolution and statistically representative data are available to aid the studies of the earliest phases of massive star formation.

The CORNISH survey[1], the first Galactic plane survey that is comparable in resolution and coverage with the GLIMPSE data, maps the compact ionised gas within the ISM. At present, the CORNISH catalogue of the northern Galactic plane, imaged with the VLA, is the most uniformly sensitive, homogeneous and complete list of northern compact radio sources at 5 GHz. The CORNISH team identified 240 ultra-compact H II region candidates. The sample provides the largest unbiased and uniform collection of these objects to date.

Previous radio UCHII samples comprise predominantly IR-targeted surveys based on IRAS point sources with far-IR colours similar to well-known UCHIIs. Wood & Churchwell (1989a,b) selected a sample of 75 UCHIIs (out of ∼1600 candidates in the Galaxy) to observe at 6 cm and 2 cm with the VLA using this method and classified them morphologically. Similarly, Miralles et al. (1994) selected and observed 12 sources at 6 and 2 cm with the VLA. Garay et al. (1993) also based their selection on strong IRAS point sources associated with compact H II regions and produced multi-frequency observations with the VLA (resolved and morphologically classified). Kurtz et al. (1994) performed radio-continuum observations on 59 UCHIIs, again IRAS-selected. De Pree et al. (2005) located and resolved a hundred objects within the massive star forming regions W49A and Sgr B2 from VLA radio continuum and radio recombination line emission observations (and revisited the Wood & Churchwell 1989a morphological classification).

The RMS survey (Urquhart et al. 2007, 2009; Lumsden et al. 2013) marked a new era of massive star formation studies. Colour-selected sources from MSX (Price et al. 2001) and 2MASS (Skrutskie et al. 2006) were followed up by arcsecond-resolution IR, as well as mm and radio observations. These, together with archival data, were used to identify for the first time a Galaxy-wide sample of ∼2000 candidate massive young stellar objects (MYSOs) and H II regions in approximately equal numbers (Urquhart et al. 2012).

The biggest disadvantage of IR-selection in UCHII studies is the discrimination against the most deeply-embedded sources.

The issue is resolved by unbiased radio surveys. The first larger-scale unbiased survey at 1.4 GHz (inner Galaxy, VLA B and A/B configuration) was conducted by Zoonematkermani et al. (1990) and was followed by (VLA C configuration) 5 GHz observations (Becker et al. 1994) covering about a fourth of the GLIMPSE region. The survey (now contained within a larger collection of re-reduced archival radio data known as MAGPIS, see Helfand et al. 2006) is useful for the study of extended thermal sources such as evolved H II regions, bubbles, etc. However, UCHII regions are unresolved or marginally resolved, and even in some instances missed altogether due to insufficient sensitivity. The CORNISH survey covers the entire GLIMPSE region and its noise level of 0.4 mJy ensures the detection of virtually all UCHIIs around a B0.5V star or earlier within the covered area (Hoare et al. 2012).

This work explores the sample of northern ultra-compact H II regions from the CORNISH survey, the majority of which are also conveniently available within the related surveys, to study the properties of this deeply embedded phase. The sample selection procedure is presented (Sect. 2) and the nature of the identified sources is verified through their observational properties (Sect. 3) and spectral indices (Sect. 3.2). Candidate short-timescale variable sources are presented in Sect. 3.3. The methodology of obtaining the distance information and the computed distances are presented in Sect. 4. The derived physical properties are discussed in Sect. 5. Results from performing automated polygon-based aperture photometry on UKIDSS and GLIMPSE infrared associations are discussed in Sect. 6 and presented in an extended source catalogue table. The spectral energy distributions (SEDs) of the sample sources from near-IR to sub-mm wavelengths were explored and utilized via SED fitting to obtain the UCHII bolometric luminosities (Sect. 7). In Sect. 8, different UCHII search methods in blind surveys are compared. The present work is summarised in Sect. 9.

2. Identification of the CORNISH UCHII sample

The CORNISH catalogue comprises 3062 sources above a 7σ detection limit. Above this limit, less than one spurious source is expected (Purcell et al. 2013). The 240 UCHII regions were selected from this high-reliability catalogue. All UCHIIs were visually identified, following criteria similar to the RMS survey, where millimetre, infrared and radio data were used for source classification (see Lumsden et al. 2013). It should be noted that the CORNISH team also identified 48 diffuse H II regions (as judged by comparison to the MAGPIS and GLIMPSE surveys), which are a part of the larger sample of CORNISH H II regions.

The full UCHII radio sample has counterparts in GLIMPSE, in all four bands (namely, IRAC 3.6, 4.5, 5.8, and 8.0 μm), with excellent positional accuracy in both surveys. This was utilised for the source identification. In the case of UCHIIs, there is overall a good agreement between the mid-IR and the 5 GHz source morphology, which ensures that the same source was captured by both surveys. A particularly good check for this are the 8 μm images. They show the morphology produced by a combination of warm Lyman-α heated dust inside the ionised zone (Hoare et al. 1991) and polyaromatic hydrocarbon (PAH) emission from just outside the ionisation front (Watson et al. 2008). This can be seen in Fig. 1. Comparison between both wavelengths is therefore useful for the distinction of adjacent unrelated sources
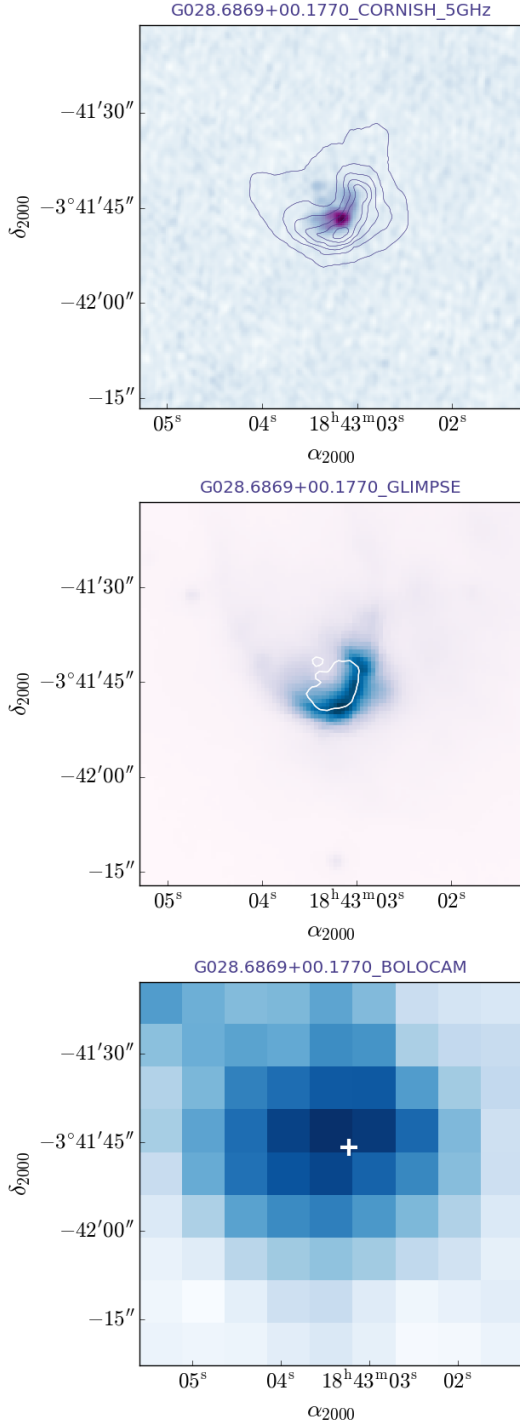
---

[1] http://cornish.leeds.ac.uk/public/index.php

Fig. 1. Images of the cometary UCHII G028.6869+00.1770. *Top panel*: CORNISH 6 cm image with overplotted GLIMPSE 8 micron contours, with the corresponding GLIMPSE image (*middle panel*), with overplotted radio contours. The morphology at both wavelengths is in very good agreement, and there is excellent positional coincidence between the two catalogues, allowing reliable source identification. The BOLOCAM image (*bottom panel*) shows a bright unresolved source at 1.1 mm coincident with G028.6869+00.1770 (the source position is marked by a white cross).

Fig. 2. CORNISH (*top panel*) and GLIMPSE (*bottom panel*) images of G010.6223−00.3788, presented as in Fig. 1, with overplotted 8 micron and 6 cm contours, respectively. A dust lane is hiding the source in the 8 micron image.

MYSOs, unlike UCHIIs, do not have strong 8.0 μm PAH emission, which is consistent with the lack of a strong UV continuum (e.g. Gibb et al. 2004). They are also generally undetected at 5 GHz, even though there are a few known MYSOs observed at radio wavelengths, with radio luminosities ($S_\nu D^2$) always below ~30 mJy kpc^2 (discussed in Hoare et al. 2007, Lumsden et al. 2013, and seen from the recent sample by Purser et al. 2016). Sources above this limit are thus H II regions or planetary nebulae (PNe).

This leaves PNe as possible contaminants of the selected sample. Unlike PNe, UCHIIs are found within molecular clouds, often in close proximity to IR clusters and dust lanes, which aids the visual classification. A lower but significant fraction of sources are found near other radio sources. About 33% of the CORNISH UCHIIs are situated in a radio cluster (within 12″ of another source), with 30% in a sky region containing more than seven detections of 7σ sources. The outlines of 24% of the UCHIIs overlap one or more 7σ sources (see Sect. 3). H II regions are expected to be strong sources in 1 mm continuum (which maps the cool dust), whereas PNe are not. BOLOCAM 1.1 mm images (see Rosolowsky et al. 2010) centred at the radio source position were visually inspected in conjunction with the IR images to verify that the UCHII sample is not contaminated by PNe.

and over-resolved emission (see Purcell et al. 2013). It can also reveal the most heavily obscured objects (those deeply embedded in infrared-dark clouds (IRDCs) or hidden behind dust lanes), as shown in Fig. 2.
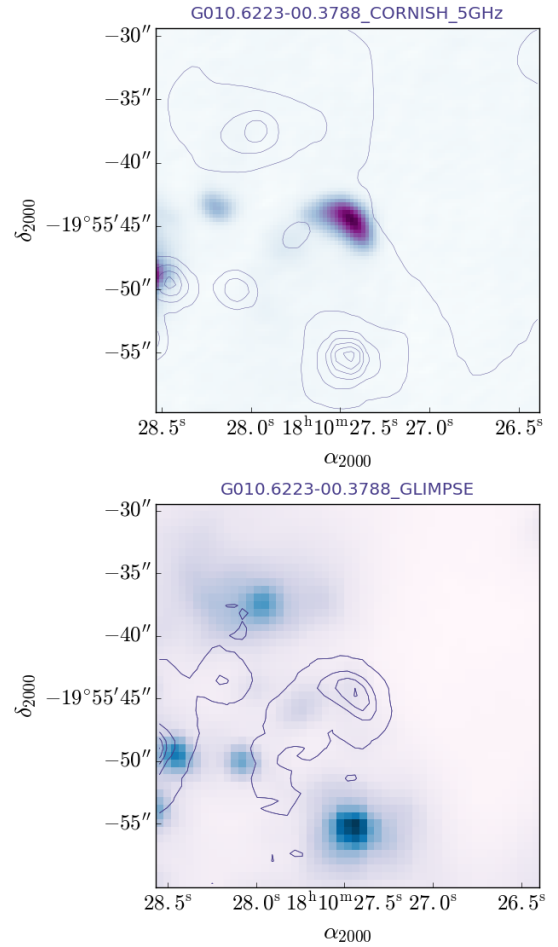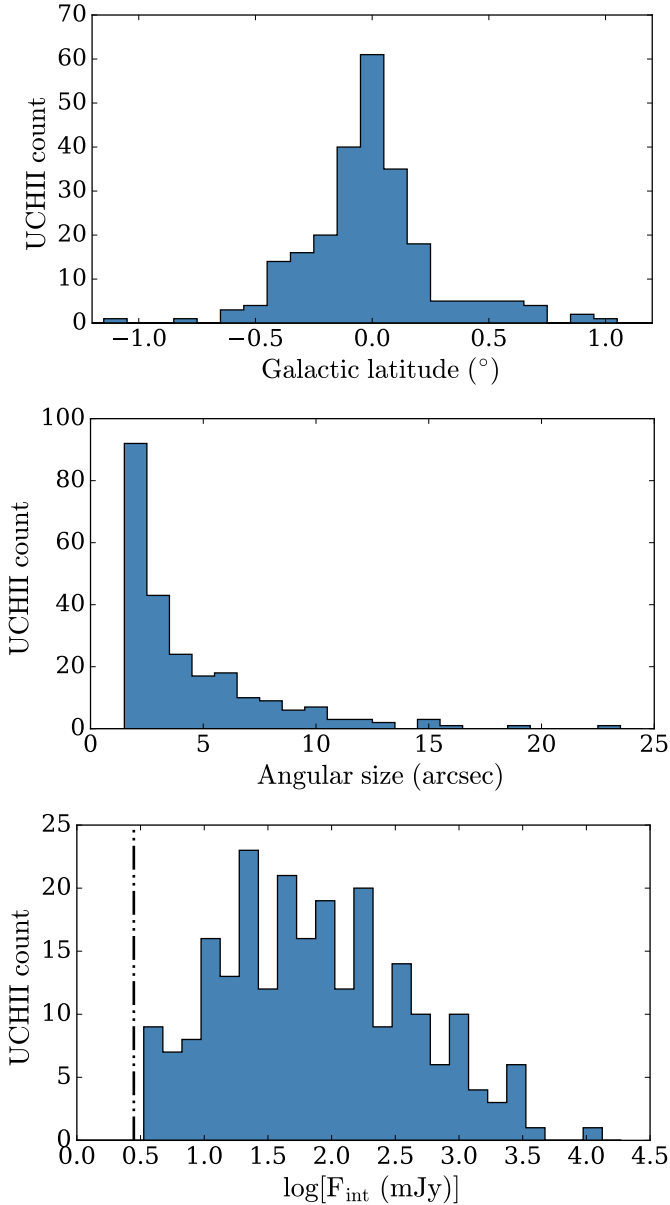
Fig. 4. Ratio of integrated to peak flux vs. angular size of the 239 COR-NISH UCHIIs. Fluxes estimated from Gaussian fits and from polygon apertures are shown as black circles and grey diamonds, respectively.

sources. As discussed in Purcell et al. (2013), sources begin to suffer from over-resolution above angular sizes of 14″. Most sources have integrated fluxes above 10 mJy, and the brightest source has a flux density of 12.6 Jy. The integrated flux histogram shows a clear downturn towards the lowest values, that is, the UCHII sample is close to complete.

The CORNISH beam size is 1.5″ and all sources with $\theta < 1.8″$ are marked as unresolved in the catalogue table (detailed checks for the entire CORNISH catalogue are discussed in Purcell et al. 2013). Within the CORNISH UCHII candidate sample, the flux was measured by fitting a Gaussian in 90 out of the 239 cases (angular size range 1.5″ to 6.2″), and for the remaining 149 sources, a hand-drawn polygon was used instead (angular size range from 1.8″ to 23.4″). The integrated and peak fluxes were compared as a function of angular size of each source, as shown in Fig. 4. Naturally, those sources whose fluxes were measured from a Gaussian fit show a clear trend of the flux ratio with increasing angular size, whereas the remaining sources with manually drawn contours show more variation.

3.1. Lower-resolution radio counterparts

The Multi-Array Galactic Plane Imaging Survey (MAGPIS; Helfand et al. 2006) is useful for the study of evolved H II regions and other extended, optically thin thermal emitters. However, it is not well-suited to explore dense, thermal sources, as those are unresolved or only marginally resolved. Catalogues at 20 cm (VLA B,C,D configuration) and at 6 cm (VLA C configuration) are available (White et al. 2005). The 6 cm catalogue covers ∼23% of the northern-GLIMPSE region, and at the survey resolution most of the detected CORNISH counterparts are unresolved – no morphological information is available.

The benefits of a comparison between CORNISH and MAGPIS are explained in detail in Purcell et al. (2013; see Figs. 20, 21). In brief, for extended sources in the CORNISH sample, the measured flux density could be a lower limit in cases where some of the extended emission was filtered out due to gaps in uv coverage, and the measured angular sizes could be underestimated by as much as 50% for non-Gaussian sources in interferometric measurements (see e.g. Panagia & Walmsley 1978). To test for such instances, a comparison to the MAGPIS

Fig. 3. Observational properties of the candidate sample – the confinement to the Galactic plane (*top panel*), small angular sizes (*middle panel*), and total radio fluxes (*bottom panel*) are consistent with UCHII regions. The dot-dashed line in the bottom panel marks the 7σ (2.8 mJy) sensitivity limit of CORNISH.

It is easier to sift out other classes of sources such as radio stars and radio galaxies. Radio stars can be distinguished by their lack of mid- and far-IR emission, whereas radio galaxies have no infrared counterparts.

3. Radio properties of the CORNISH UCHIIs

The distribution of the Galactic latitudes, angular sizes and integrated fluxes of the sample of candidate UCHIIs are shown in Fig. 3. The sources are closely confined to the Galactic plane, as expected for very young massive star forming regions. Using the CORNISH survey, Urquhart et al. (2013) fit a scale-height of 20.7 ± 1.7 pc for compact and ultra-compact H II regions.

The majority of the ultra-compact H II regions have angular sizes below 5″, with the histogram peaking towards unresolved
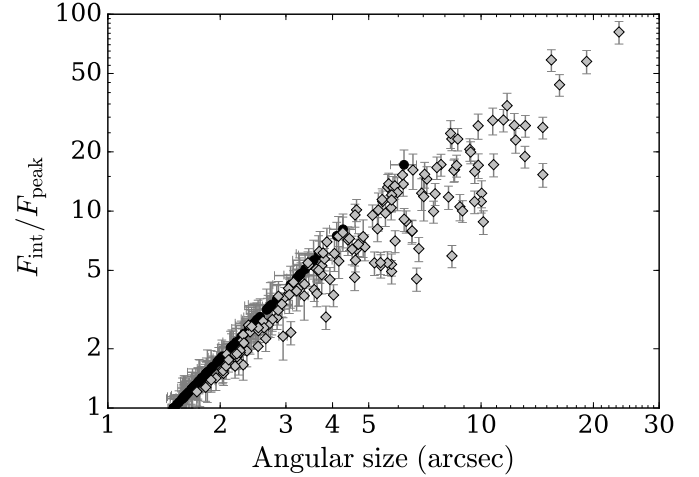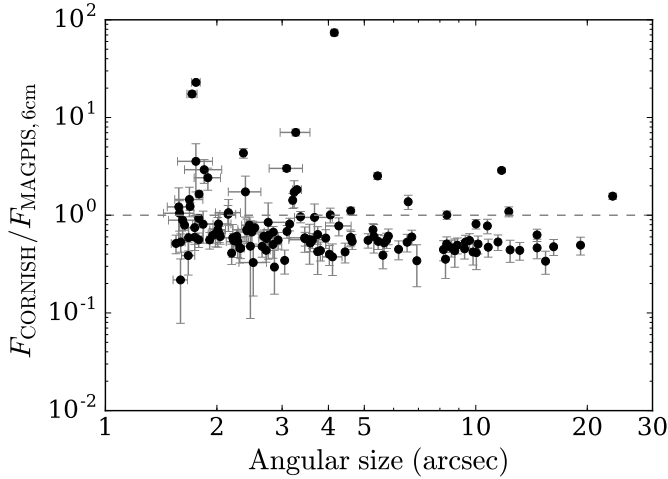
Fig. 5. Comparison of the MAGPIS 6 cm and CORNISH integrated fluxes (black circles) with plotted line of equality, plotted against the corresponding CORNISH angular size.

Fig. 6. Spectral indices between 1.4 GHz and 5 GHz against brightness temperature of the corresponding CORNISH source at 5 GHz. The model lines signify different T_e. All CORNISH sources with an overlapping 5σ or 7σ source were excluded, as they are not resolved in MAGPIS.

6 cm data is useful, as the survey configuration allowed the recovery of more diffuse emission at the cost of lower resolution.

It is important to note that there are instances of repeats in the combined 6–20 cm MAGPIS catalogue. This is due to the fact that sources matching more than once are listed multiple times, for example a 6 cm source listed with each 20 cm counterpart and vice versa (White et al. 2005). Since the structure of the MAGPIS catalogue leads to repeats even within a small cross-matching radius, to obtain a sufficiently reliable cross-match with the MAGPIS catalogue table, the matched sample had to be limited to only 47 associations. Visual identification of the associations was therefore preferred. The CORNISH team visually identified 21 620 cm matches (out of which 162 have both a 6 cm and 20 cm MAGPIS detection).

The fluxes of the 20 cm and 6 cm MAGPIS UCHII associations were measured independently of the catalogue table values via automated aperture photometry scripts. Radio outlines from the CORNISH database − hand-drawn polygons in the case of extended radio sources, and Gaussian outlines in the case of compact radio sources, were utilised. In order to use these outlines as apertures for the lower-resolution MAPGIS data, they were expanded accordingly. The necessary "padding" value (i.e. the required radial expansion) was determined after multiple runs with different aperture sizes. Based on the curve of growth, a padding value of 4″ (i.e. total expansion of 8″) was chosen for the flux measurement.

Figure 5 shows that the majority of the flux ratio values occupy the range between ∼0.2 and 1.1, which indicates that some flux was not recovered at the higher resolution (around 65% of the flux was detected for these sources). The slightly negative slope of the flux ratio is due to worsening over-resolution with increasing angular size. Sources overlapping with one or more 5σ or 7σ CORNISH neighbours were excluded, as these sources are unresolved and merged in the MAGPIS 6 cm images and their flux measurements are unreliable. The outliers above the equality line, that is, with CORNISH fluxes significantly higher than their MAGPIS 6 cm counterpart, are investigated for potential short-timescale variability in Sect. 3.3.

The reliability of the measured 21 620-cm and 1626-cm flux values was judged on the basis of their median brightness level of the sky. For all sources with abnormal (i.e. outside the range of the majority of sources) median level in the sky-annuli, upper
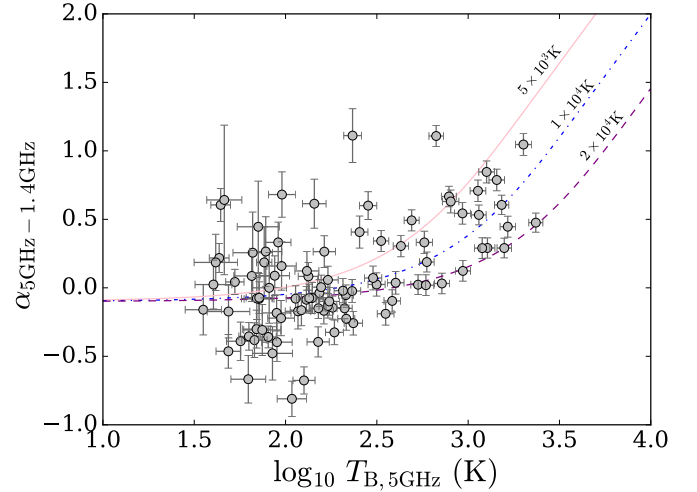
flux limits are included in the flux table (an example is given in Table B.1, the full table is available at the CDS).

Sources with consistent sky values that are also not overlapping with a 5σ or 7σ source are here considered to be the highly-reliable MAGPIS flux measurement subset. These were used to compute the spectral indices of the CORNISH UCHIIs – see Sect. 3.2.

A few MAGPIS UCHII associations were picked at random out of the highly-reliable subset and their fluxes were also measured with CASA, by Gaussian fits in unresolved cases or computing the flux for the extended source region otherwise[2]. Table 1 shows a comparison of the flux values in the published MAGPIS catalogue table with our remeasured fluxes for these sources. Clearly, neither the photometry performed with the automated script or with CASA reproduced the catalogued 6 cm MAGPIS values to a reasonable degree, with a much better agreement between our two photometric measurements. This also appears to be true when comparing the results for extended 20 cm sources. This is why the fluxes of all available lower-resolution counterparts were remeasured in this work. Our results obtained for unresolved 20 cm sources appear to be overall in better agreement with the MAGPIS catalogue.

3.2. 5 GHz–1.4 GHz spectral indices

The spectral indices (α) were computed from the relation

$$\alpha = \frac{\ln\left(S_{5\,\text{GHz}}/S_{1.4\,\text{GHz}}\right)}{\ln\left(5/1.4\right)}\,,\tag{1}$$

where $S_{5\,\text{GHz}}$ and $S_{1.4\,\text{GHz}}$ are the integrated fluxes at 6 cm and 20 cm, respectively. The uncertainty was found from

$$\Delta\alpha = \frac{\sqrt{(\sigma_{5\,\text{GHz}}/S_{5\,\text{GHz}})^2 + (\sigma_{1.4\,\text{GHz}}/S_{1.4\,\text{GHz}})^2}}{\ln\left(5/1.4\right)}.\tag{2}$$

The indices were calculated for the CORNISH 6 cm flux and using the newly measured 20 cm fluxes. As discussed in the

---

[2] The same method was tested on the CORNISH images, in good agreement with the catalogued values.

Table 1. Comparison of a few MAGPIS flux values in the published 6–20 cm catalogue table (5th and 8th columns) with the measured flux values presented in this work – as from the automated script (3rd and 6th columns) and individual measurements with CASA (4th and 6th columns).

CORNISH name	F_C (cat.)	$F_{M6\,cm}$ (ap.)	$F_{M6\,cm}$ (CASA)	$F_{M6\,cm}$ (cat.)	$F_{M20\,cm}$ (ap.)	$F_{M20\,cm}$ (CASA)	$F_{M20\,cm}$ (cat.)
G011.1104−00.3985	305.37 ± 28.55	303.63 ± 9.12	320.3	112.37	326.73 ± 3.86	370.91	187.21
G012.1988−00.0345	62.71 ± 5.92	102.95 ± 1.47	110.18	59.82	42.52 ± 1.34	48.00	62.05
G016.3913−00.1383‡	124.27 ± 15.43	43.18 ± 3.61	42.89	20.34	57.52 ± 5.18	66.71	58.23
G018.7106+00.0002	107.46 ± 10.62	177.74 ± 1.71	188.78	102.96	36.60 ± 0.94	41.97	35.39
G021.8751+00.0075	566.73 ± 54.14	1064.73 ± 8.85	1073.00	264.32	685.90 ± 7.84	719.53	599.63
G037.8731−00.3996	2561.21 ± 234.04	5189.35 ±23.38	5219.50	1518.5	1769.56 ± 6.57	1800.10	1279.40

Notes. Only highly-reliable values (based on median sky values and no overlapping sources) are compared in the table. All values are in mJy. The second column includes the corresponding CORNISH fluxes. The source denoted by ‡ could be intrinsically variable at 6 cm.

previous section, for the purpose of obtaining reliable spectral indices, all overlapping sources were excluded, thus making it possible to compute 93 spectral indices. This should be sufficiently representative of the sample as a whole, as the selection is in this way based on image quality alone.

The brightness temperatures of the sources in the UCHII sample at 6 cm were also calculated. This was done for all CORNISH UCHII regions (see Fig. 7e), using the equation

$$T_b = \frac{1.36\lambda^2 S_\nu}{\theta^2} \,, \tag{3}$$

where the flux density S_ν is in mJy, the brightness temperature T_b is in K, the wavelength λ is in cm and the HPBW (half-power beam width) θ for a Gaussian beam is in arcseconds. For all unresolved sources, this provides a lower limit. T_b serves as a measure of optical thickness – increasing with larger optical depth and reaching the electron temperature (T_e) of the ionised region in the optically thick limit (Siódmiak & Tylenda 2001). The optical depths τ_ν at 5 GHz were found from

$$T_b = T_e \left(1 - e^{-\tau_\nu}\right) \,, \tag{4}$$

for $T_e = 10^4$ K (Eq. (4) in Siódmiak & Tylenda 2001). We note that the optical depth results (Fig. 7f) are lower limits in the case of unresolved sources (due to the angular size dependence of T_b), as well as for the most extended sources (due to the possible underestimation of the 5 GHz flux, see Fig. 5).

The spectral index analysis was performed by plotting the UCHII spectral indices (Fig. 6) as a function of their corresponding 6 cm brightness temperatures and comparing the result to the theoretical model developed by Bojičić et al. (2011; see Eq. (6) in their paper), who followed up on the work by Siódmiak & Tylenda (2001). The focus of these works were PNe, however the model is general and should be applicable to all ionised nebulae. It assumes a uniform nebula characterized by its electron temperature and optical thickness at a corresponding reference frequency. The individual spectral index values (along with the measured fluxes at 6 cm and 20 cm) can be found in Table B.1.

The majority of UCHII candidates were found to be within the expected theoretical limits for free-free emission ($-0.1 \lesssim \alpha \lesssim 2$) − that is, between sources with optically thin or optically thick radio continuum emission at both frequencies. None of the sources have a spectral index above 1.5. However, ~18% of the sources are found below the lower limit (taking errors into account), which is inconsistent with thermal emission. The reason for the apparently non-thermal spectral indices is likely to be

the difference in the VLA configuration at 1.4 GHz and 5 GHz (i.e. the uv coverage at higher resolution resulting in filtering out of some of the flux, as seen in Fig. 5). If the 6 cm fluxes were ~1.6 times higher (as high as the measured 6 cm MAPGIS fluxes, Fig. 5), all but two sources (G010.6297−00.3380 and G023.8985+00.0647) would be shifted within the thermal bounds. Despite this, we choose to use the CORNISH 6 cm data over the MAGPIS 6 cm images, as the latter show significantly greater image-to-image variations in median sky levels and thus provide fewer 6 cm frames that are viable for the spectral index calculation (only 71/162 MAGPIS 6 cm sources when overlapping sources are excluded). The 20 cm images appear overall of higher quality than the MAGPIS 6 cm data. Therefore, by utilising the CORNISH 6 cm images and the (reduced sample of 93) MAGPIS 20 cm images, we have limited our spectral index analysis to the best available data. Due to the time difference between the 20 cm observations and (both sets of) the 6 cm observations, source variability, as discussed in Sect. 3.3, is another possible explanation for some instances of non-thermal (appearing) UCHIIs. Unfortunately, it is not possible at this time to quantify this effect without more 20 cm observations.

The location of each UCHII in Fig. 6 is determined by both its optical thickness and its electron temperature. The model lines in the diagram correspond to applying Eq. (6) of Bojičić et al. (2011) for electron temperatures $T_e = 5 \times 10^3$ K, $T_e = 1 \times 10^4$ K and $T_e = 2 \times 10^4$ K. For reference, the mean value for the electron temperature of Galactic H II regions has been estimated to be 8000 K (Quireza et al. 2006). At lower brightness temperatures, the UCHIIs scatter around the optically thin limit, where we find the majority of the sample. The much larger associated uncertainties of the sources with lower T_b (i.e. of lower opacity at 6 cm) should be noted in this case. With increasing T_b, α increases in agreement with the model, corresponding to 5×10^3 K $\lesssim T_e \lesssim 2 \times 10^4$ K, revealing sources that appear optically thick. Lower electron temperatures would be a better fit to the sample if we take into account the possible under-estimation of the 6 cm flux.

The presented spectral index results are a validation of the UCHII region nature of our sample.

3.3. Evidence for short-timescale UCHII variability?

All sources with available MAGPIS and CORNISH 6 cm images were revisited to look for significant flux changes between the two epochs that cannot be attributed to image quality or other individual reasons. Such sources could be variable over short timescales comparable with the difference in time between the

Table 2. Ratio of CORNISH to MAGPIS 5 GHz fluxes for candidate variable UCHII regions (observed 15 years apart).

CORNISH name	$F_C/F_{M,6\,cm}$
G011.0328+00.0274	2.42 ± 0.61
G011.9786−00.0973	3.56 ± 1.83
G014.5988+00.0198	2.92 ± 0.80
G016.3913−00.1383	2.88 ± 0.15
G023.4553−00.2010	17.42 ± 1.28
G025.7157+00.0487	4.33 ± 0.49
G030.7579+00.2042	73.85 ± 4.49
G030.7661−00.0348	7.04 ± 0.56
G037.7347−00.1128	22.90 ± 1.09

Notes. We note that in some cases, the source was not detected in MAGPIS, so the upper flux limit was used instead for the comparison.

MAGPIS 6 cm survey (obs. 1989–1991) and the CORNISH survey (obs. 2006–2008). The flux ratio lower limits were taken into account to quantify this. The higher flux at the earlier epoch (due to the detection of more extended emission) prevents a reliable investigation of instances of intrinsic flux decrease, particularly for sources larger than 5″. A further hindrance is that any hypothetical UCHIIs that are completely invisible at the later epoch due to a significant flux decrease over time cannot be reliably differentiated from more extended H II region phases due to the lower resolution at the earlier epoch.

Due to these limitations imposed by the different VLA configurations used at the two epochs, in the context of this work, we use the term variability to refer to increase in flux in the ∼15 years separating the two surveys. The sources which appear to be intrinsically variable all have a flux increase greater than ∼50%. These are listed in Table 2.

The 6-cm variables appear to have several properties in common:

1. They exist in relative isolation (no overlapping sources, no other radio sources within ≳1′, no busy complexes). An exception to this is G030.7661−00.0348, which is likely experiencing substantially different effects than the rest in the bustling environment of W43;
2. They are all near-IR dark (apart from G030.7661−00.0348);
3. Most are particularly compact (≲0.1 pc), with the exception of G016.3913−00.1383.
4. Their 5 GHz–1.4 GHz spectral indices are not anomalous with respect to the rest of the sample.

The potentially variable sources comprise ∼5% of the CORNISH UCHII sample. In four cases (G011.9786−00.0973, G014.5988+00.0198, G023.4553−00.2010, and G030.7661−00.0348), the CORNISH source was not detected at both 6 and 20 cm at the earlier observational epoch. Three sources (G025.7157+00.0487, G030.7579+00.2042, and G037.7347−00.1128) all have only a 20 cm detection at the earlier epoch (and there is a dim 20 cm counterpart for G011.0328+00.0274). The extended G016.3913−00.1383 has 6 and 20 cm counterparts, but these appear smaller and dimmer at the earlier epoch. Although there are some suggestive correlations, this group of candidate variable sources is clearly statistically insufficient to establish any common pattern linking the (presence or lack of) emission at 6 and 20 cm at the same epoch. Unfortunately, the 6 cm variability cannot be linked in any way to variability at 20 cm, due to the lack of other available 20 cm data of comparable quality for the 6-cm variables.

Time-variable radio flux densities of several ultra-compact and hyper-compact (HC) H II regions have been reported previously (Acord et al. 1998, Franco-Hernández & Rodríguez 2004, Rodríguez et al. 2007, Gómez et al. 2008, Galván-Madrid et al. 2008). The flux changes have been associated with morphological changes across observational epochs, on timescales of a few years. The UC and HC H II regions have been caught expanding (Acord et al. 1998). Galván-Madrid et al. (2008) discuss contracting UC and HC H II regions; however their sources are unresolved. Variability over observable timescales could be caused by factors that are either external or internal to the forming star. The former could be the result of chaotic motions of the material surrounding the ionising star. In this scenario, optically thick gas (e.g. clumps in the stellar wind) would occasionally block the outgoing radiation, shielding the outer ionised gas layers and thus neutralising them (Peters et al. 2010a). In the latter scenario, the forming star itself is undergoing changes (Franco-Hernández & Rodríguez 2004) − surface temperature fluctuations affect the UV flux and thus the H II region size.

Theoretical studies have reproduced this behaviour, referred to as flickering. The three-dimensional collapse simulations of massive star formation of Peters et al. (2010a,b,c) and Klassen et al. (2012) include feedback by ionising radiation and show time variability leading to changes in H II region appearance and flux comparable to observations. The Peters et al. (2010a) model produces flickering on scales of ∼10 years. In this model, accretion has not ceased prior to the UCHII stage. The infalling neutral flow becomes ionised when in close proximity to the star. The H II region is gravitationally trapped early on, which is followed by a fluctuation between trapped and extended states, and thus changes in flux, size and morphology are seen over time. Galván-Madrid et al. (2011) performed statistical analysis of simulated radio-continuum observations separated by 10 year steps, using the Peters et al. (2010a,b,c) models to form H II regions. They found that 7% of the simulated HC and UC H II regions have a detectable flux increase (larger than 10%) and 3% have a detectable flux decrease, but expect only ∼0.3% of their H II regions to have a flux increase of over 50%. The observations discussed in this work show that ∼5% of the CORNISH UCHII regions have become brighter by 50% or more over a comparable time scale (∼15 years), based on the 6 cm data. In practice, any similar statistic of observed sources with flux increase ≲10% would be unreliable, given the associated flux uncertainties. In any case, assuming that all of our candidate variable sources truly undergo intrinsic changes, variable HC and UC H II regions could be significantly more common and their brightness could fluctuate more than predicted. If this is the case, invoking ongoing accretion cannot account for the observed dramatic change in flux. The Peters et al. (2010a) model does not include radiation pressure, magnetic fields, winds and outflows, all of which are components of the star formation process and might be related to variability.

4. Distances

A crucial step towards characterising UCHII regions is to determine the distance to each source. One can then convert measured parameters (e.g. fluxes and angular sizes) into physical quantities (e.g. luminosities and physical sizes). Accurately derived distances to UCHII regions are used to test the current models of the face-on Galactic structure (see e.g. Urquhart et al. 2013). As the heavy obscuration hinders the use of any optical distance determination techniques, the distances to most UCHII regions are kinematically derived. The kinematic distance is found by

fitting the radial velocity of the source to a Galactic rotation curve (e.g. Brand & Blitz 1993; Reid et al. 2009), using radio or mm spectral line data. Errors in the calculated distance arise when the source radial velocity differs from the one assumed by the model (e.g. velocity errors of about 10 km s^{-1} due to the velocities departing from circular rotation as a result of streaming motions; Anderson & Bania 2009).

Distance estimates for the outer Galaxy are relatively straightforward. However, a major obstacle arises when one seeks the kinematic distance for objects within the Solar circle. At Galactic radii smaller than that of the Sun, two possible distances exist for each radial velocity. These distances, known as near and far, are situated at equal intervals from the tangent point distance. The kinematic distance ambiguity (KDA) is not present only for the tangent point velocity, which is the maximum radial velocity. Different methods exist to assign the correct kinematic distance to the sources of interest – for example, H I emission/absorption (H I E/A), H I self-absorption (H I SA), or using absorption lines from other molecules, for example formaldehyde (H$_2$CO) (as discussed by Anderson & Bania 2009).

These methods are particularly effective for bright UCHII regions, whose free-free continuum emission is substantially stronger than the Galactic H I emission, thus resulting in unambiguous absorption spectra. As the maximum radial velocity along the line of sight is always the tangent velocity, lack of absorption between the source and tangent velocity reveals that the source is located at the near distance. Otherwise, the far distance is assigned (or the tangent distance, in the cases when the source velocity equals the tangent velocity). In this work, the H I E/A method was adopted (see e.g. Urquhart et al. 2012). This method makes use of CO emission line data and H I absorption to obtain the near and far kinematic distances and to attempt to resolve the KDA.

4.1. Distances from ATLASGAL and RMS

The work by Urquhart et al. (2013) presents an unbiased and complete sample of 170 molecular clumps with 213 embedded compact and ultra-compact H II regions over the common GLIMPSE, ATLASGAL and CORNISH survey region. Table 3 from Urquhart et al. (2013) contains distances to all CHII- and UCHII-hosting clumps and Table 4 contains Lyman continuum fluxes (F_{Ly}) and source physical sizes. Out of the associated 213 H II regions, 203 also belong to the CORNISH UCHII sample (ten were classified as more extended). Of the remaining 36 UCHII candidates without ATLASGAL distances, eight were found to have a distance estimate in the Red MSX source survey (RMS) database[3] (Lumsden et al. 2013). These are the distances for G010.6297−00.3380, G030.6881−00.0718, G032.0297+00.0491, G035.0524−00.5177, G038.5493+00.1646, G048.6099+00.0270, G060.8842−00.1286, and G061.7207+00.8630.

4.2. Distance estimates for remaining sources using CO data

Anderson et al. (2009) describe a large-scale study of the molecular properties of H II regions of different sizes and morphologies using fully sampled CO maps. This is the BU-FCRAO Galactic Ring Survey (Jackson et al. 2006), which uses ^{13}CO $J = 1 \longrightarrow 0$ emission. This has advantages over the commonly used ^{12}CO

[3] http://rms.leeds.ac.uk/cgi-bin/public/RMS_DATABASE.cgi

Table 3. Comparison between velocities (in km s^{-1}) measured in this work and by Anderson et al. (2009), Table 3 for the sources in common.

CORNISH	v	Anderson	v
G024.4698+00.4954	103	C24.47+0.49	102.67
G024.4721+00.4877	102.8	C24.47+0.49	102.67
G024.4736+00.4950	102.6	C24.47+0.49	102.67
G024.8497+00.0881	109.3	C24.81+0.10	108.31
G030.7661−00.0348	96.1	C30.78−0.03	94.76
G030.7661−00.0348	96.1	U30.84−0.11b	96.89
G031.2420−00.1106	21.1	U31.24−0.11a	21.07
G034.2544+00.1460	57.7	U34.26+0.15	57.1
G034.2571+00.1466	57.7	U34.26+0.15	57.1
G037.9723−00.0965	54.7	C38.05−0.04	54.1
G049.4640−00.3511	59.5	U49.49−0.37	60.08
G049.4891−00.3763	60.9	U49.49−0.37	60.08
G050.3157+00.6747	26.5	U50.32+0.68	26.31

Notes. As the peak channel was used to calculate the velocities, the associated errors are given by the channel width. This equals 0.21 km s^{-1} for each BU-FCRAO Galactic Ring Survey cube. In practice this error is outweighed by the error due to peculiar motions, which is ~10 km s^{-1}.

isotopologue, as ^{13}CO is ~50 times less abundant and thus provides a smaller optical depth, and consequently smaller line widths and better separation of velocity components along the line of sight.

CO data-cubes from the Galactic Ring Survey[4] were used to obtain radial velocities for the remaining sources without ATLASGAL or RMS distances. The positional accuracy of the GRS is ~2.3″, which is equivalent to 1/10 of the spacing between grid points on the map. Data cubes for the available sources were obtained, and the radial velocities were measured for the sources coinciding with the CORNISH coordinates. The emission line structure of the data is very complex, often with multiple emission lines. This reflects the complexity of the (l, b, v_{LSR}) structure of the molecular gas in the line of sight of the H II region. The sources were located within the 15′ data cubes by going manually through the cube channels and then mapping the cube once the source was found at or close to the precise CORNISH coordinates. A source's velocity was taken to be equal to the velocity of the most prominent emission line in the map at the exact source position. The final CO source velocity results were compared to Table 3 in Anderson et al. (2009) for the sources in common. The code by Reid et al. (2009) was used to obtain near and far kinematic distances corresponding to each radial velocity estimate (see Sect. 4.3).

4.3. Resolving the KDA

The H I Emission/Absorption (H I E/A) method was implemented as a standard way to choose between the calculated near and far distances (see e.g. Anderson & Bania 2009; Urquhart et al. 2012). The method has proven to be very successful for KDA resolution (Anderson & Bania 2009).

Out of the 28 CORNISH UCHIIs with missing CO distances, CO and VGPS H I data were found for 21 (CO data were not available within the range $10° \leq l \leq 17°$). In the VGPS spectra, the H I brightness temperature at the source location is

[4] http://www.bu.edu/galacticring/

Table 4. KDA-resolved UCHII distances.

CORNISH name	KDA	Strong H I abs.	d (kpc)
G024.4698+00.4954	n	✓	5.5 ± 0.3
G024.4721+00.4877	n	✓	5.5 ± 0.3
G024.4736+00.4950	n	✓	5.5 ± 0.3
G024.8497+00.0881	n	✓	5.8 ± 0.3
G026.0083+00.1369	f	✗	13.8 ± 0.5
G026.8304−00.2067	f	✗	11.9 ± 0.3
G029.7704+00.2189	f	✗	9.8 ± 0.3
G030.7579+00.2042	t	✗	7.2 ± 0.6
G030.7661−00.0348	t	✓	7.2 ± 0.6
G031.2420−00.1106	f	✗	12.7 ± 0.4
G034.2544+00.1460	n	✓	3.6 ± 0.4
G034.2571+00.1466	n	✓	3.6 ± 0.4
G035.4570−00.1791	f	✗	9.7 ± 0.4
G037.7562+00.5605	f	✗	12.2 ± 0.5
G037.9723−00.0965	f	✗	9.7 ± 0.4
G045.5431−00.0073	t	✗	5.9 ± 0.9
G049.4640−00.3511	t	✓	5.5 ± 2.2
G049.4891−00.3763	t	✓	5.5 ± 2.2
G050.3157+00.6747	f	✗	8.6 ± 0.5
G061.4763+00.0892	t	✓	4.0 ± 1.5
G061.4770+00.0891	t	✓	4.0 ± 1.5

Notes. The distances and errors were derived with the Reid et al. (2009) code and the H I E/A method was used to resolve the KDA. The sources that are found at the near, far, and tangent distance, are labelled with n, f, and t, correspondingly. The third column indicates whether the H I spectra used for the KDA resolution were of good quality (i.e. mostly in absorption, marked with ✓) or poor (i.e. mostly in emission, indicated with ✗).

overall ≲100 K, as expected for optically thin H I gas. Therefore, a background source of higher temperature should be seen in H I absorption. The presence of a H II region would provide a sufficiently strong continuum to detect a line in absorption. It should be noted, however, that the VGPS synthesized beam size is ∼45″ (FWHM) at 21 cm (see Stil et al. 2006 for the VGPS survey paper). The H II region may not be detected at all if the source size is not comparable to the survey beam size (Urquhart et al. 2012). The extent of the effects of this on our distance estimates are hard to quantify. Moreover, neighbouring CORNISH UCHIIs found within the beam cannot be distinguished. This is the case, for instance, for G024.4721+00.4877, G024.4736+00.4950, and G024.4698+00.4954, whose H I spectra are practically identical. The latter issue is alleviated by the fact that spatially close UCHIIs are likely to belong to the same star-forming region, and hence to be situated at the same distance.

To assign a near or a far distance, the H I data were plotted, together with a line marking the assigned CO radial velocity and its associated error (±10 km s⁻¹ due to peculiar motions) and the calculated tangent velocity was included as well (see Appendix A). All sources with CO velocity within 10 km s⁻¹ of the tangent point velocity were assigned the tangent point distance, in order to limit wrong assignments to the near distance (as suggested by Anderson & Bania 2009). Additional uncertainty arises due to half of the H I spectra being dominated by emission instead of absorption at the UCHII source position. The presence or lack of convincingly strong absorption is indicated in Table 4 for each H II region. An example of a reliable H I spectrum is that coinciding with G049.4891−00.3763, as it clearly shows very strong absorption lines. Emission-dominated spectra

appear similar to the one coincident with G029.7704+00.2189 (Appendix A).

All 21 UCHIIs with available GRS CO and VGPS H I data were assigned a distance, and four more distances were adopted from the paper by Cesaroni et al. (2015). These are the distances for G010.3204−00.2328, G011.1712−00.0662, G014.1741+00.0245, and G016.3913−00.1383. It should be noted that for these four sources, as for all sources without data to resolve the ambiguity, Cesaroni et al. (2015) assigned the far distance. Placing sources without a distance solution at the near distance is preferable. Only three sources (G010.8519−00.4407, G011.9786−00.0973, and G014.1046+00.0918) with no known distance and no available emission and absorption data remain in our sample. From the group of 21 UCHIIs, six were assigned the near distance, eight – the far, and the remaining seven sources were assigned the tangent point distance.

5. Derived physical properties

A summary of the physical properties of the sample is presented in Fig. 7. The distributions of heliocentric and galactocentric distances, physical sizes, optical depths, brightness temperatures, Lyman continuum fluxes, emission measures, and electron densities are shown. The computed values corresponding to each CORNISH UCHII can be found at the CDS (an example is shown in Appendix D).

Figure 7a shows the distribution of the UCHII regions with heliocentric distance. The five peaks are at 2, 4, 11, 13, and 17 kpc. As noted by Urquhart et al. (2013, who also find five similarly situated peaks at 2, 5, 10, 12, and 16 kpc), they likely correspond, in turn, to the near side of the Sagittarius arm, the end of the bar and Scutum-Centaurus arm, the far sides of the Sagittarius and Perseus arms, and the Norma arm. Urquhart et al. (2013) studied the associated clumps to the CORNISH compact H II regions and found that the most prominent peak in the heliocentric distance histogram at ∼11 kpc corresponds to the W49A complex (as this peak is not seen for clumps).

The galactocentric distances were also estimated (Fig. 7b). The galactocentric distance distribution depends only on the choice of Galactic rotation curve (Urquhart et al. 2011). The peaks are located at ∼4.5, 6, and 7.5 kpc, similarly to the findings of Urquhart et al. (2011, peaks at ∼4, 6, and 8 kpc) for the young massive star sample in the RMS survey. Urquhart et al. (2011) identify the peak at approximately 4 kpc to be at the intersection of the Long Bar and the Scutum-Centaurus arm, also coinciding with the W43 complex. The 6 kpc peak is coincident with the Sagittarius arm, and the 7.5 kpc peak corresponds to the Perseus arm (the bin is dominated by the W49A complex – the most active star-forming region in the Galaxy).

The distribution of the physical sizes is shown in Fig. 7c. All sources with available deconvolved angular sizes (see Purcell et al. 2013) are presented in the histogram in blue. The physical sizes of 36 UCHIIs could not be obtained because the corresponding sources were unresolved and hence their true sizes are unknown (and three more sources do not have a computed distance). Upper limits on the physical sizes were computed for these unresolved sources with available distances (grey region in the histogram). Most sources (∼66%) are larger than 0.05 pc and smaller than 0.2 pc in diameter, and the distribution peaks at 0.1 pc, which is consistent with the typical sizes of UCHII regions. About 12.5% are with sizes between 0.01 pc and 0.05 pc, and the remaining ∼21.5% sources are between 0.2 pc and 0.9 pc. The investigated properties of these sources are in accord with the rest of the sample and therefore likely
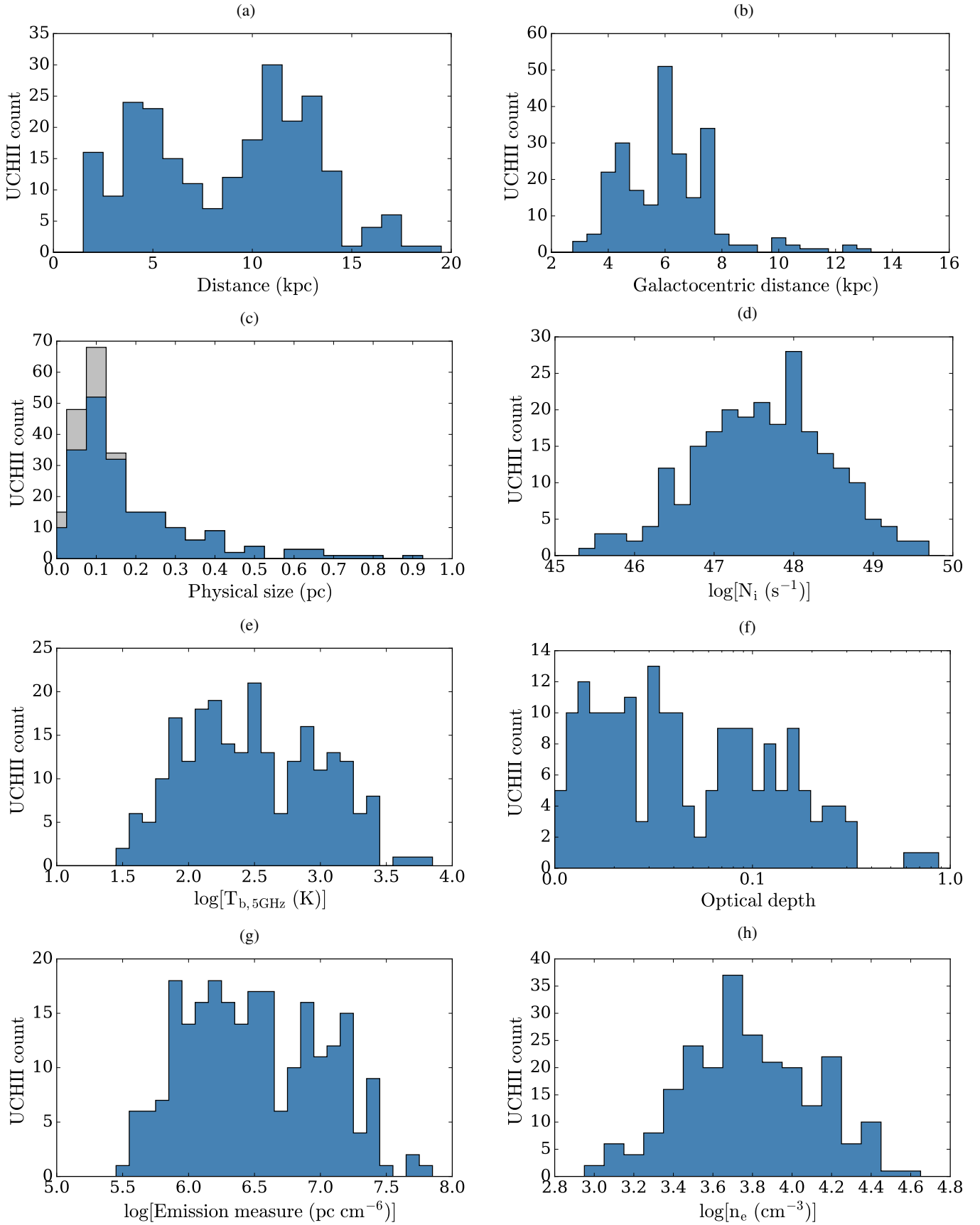
Fig. 7. Parameter distributions of the CORNISH UCHIIs. The distance histogram (*panel a*) includes values from Urquhart et al. (2012, 2013), as well as the distances computed in this work. The grey area in *panel c* shows upper limits for sizes of unresolved sources. The Lyman continuum flux histogram (*panel d*) includes results from Urquhart et al. (2013), as well as results from this work, with ±10% associated error. There are ±30% errors on the emission measures (*panel g*) and ±20% errors on the electron densities (*panel h*).

exhibit the natural variation in size one would expect for a continuum of spectral subtypes and ages, and for different ambient densities.

Lyman continuum fluxes from Urquhart et al. (2013) were available for 203 out of the 239 sources in the CORNISH sample. The Lyman continuum fluxes N_i for 33 of the remaining sources were computed from

$$\left[\frac{N_i}{\text{photons s}^{-1}}\right] = 9 \times 10^{43} \left[\frac{S_\nu}{\text{mJy}}\right]\left[\frac{d^2}{\text{kpc}}\right]\left[\frac{\nu^{0.1}}{5\,\text{GHz}}\right], \tag{5}$$

where S_ν is the flux at a frequency ν and d is the distance (Eq. (6) from Urquhart et al. 2013). The associated error on N_i is 10%, which is dominated by the distance error. As for the emission measures and electron densities, a dust-free, optically thin emitter is assumed here. The Lyman continuum flux distribution (Fig. 7d) of the CORNISH UCHII sample peaks between $47.2 < \log(N_i) < 48.5$ photons s^{-1}, which is also consistent with the Lyman continuum ionising flux from zero-age main-sequence stars with spectral class from B0 to O7. The computed Lyman continuum fluxes should be treated as lower limits.

The brightness temperature distribution at 6 cm and the corresponding optical depth distribution, as calculated in Sect. 3.2 (Eqs. (3) and (4)), are shown in Figs. 7e and 7f, respectively. The optical depth histogram, as well as Fig. 6, confirm that assuming optically thin sources at 5 GHz is justified for the majority of UCHIIs in our sample.

The emission measures (EM, Fig. 7g) and the electron densities (n_e, Fig. 7h) of the UCHII regions were computed from the equations

$$\left[\frac{EM}{\text{pc cm}^{-6}}\right] = 1.7 \times 10^7 \left[\frac{S_\nu}{\text{Jy}}\right]\left[\frac{\nu}{\text{GHz}}\right]^{0.1}\left[\frac{T_e}{\text{K}}\right]^{0.35}\left[\frac{\theta_S}{''}\right]^{-2} \tag{6}$$

and

$$\left[\frac{n_e}{\text{cm}^{-3}}\right] = 2.3 \times 10^6 \left[\frac{S_\nu}{\text{Jy}}\right]^{0.5}\left[\frac{\nu}{\text{GHz}}\right]^{0.05}\left[\frac{T_e}{\text{K}}\right]^{0.175}\left[\frac{\theta_S}{''}\right]^{-1.5}$$
$$\times \left[\frac{d}{\text{pc}}\right]^{-0.5}, \tag{7}$$

where $T_e = 10^4$ K is the electron temperature and θ_S is the source angular size. Both equations were adopted from Sánchez-Monge et al. (2013), who followed the formalism of Mezger & Henderson (1967) and Rubin (1968). It is assumed that the cm continuum flux is emitted from homogeneous optically thin H II regions. As can be seen from the optical depth results shown in Fig. 7f, this is a good description for the sample. The typical uncertainties on the flux density and angular diameter imply an uncertainty on the emission measure of 30% and uncertainty on the electron density of 20%. Taking the 5σ flux sensitivity of CORNISH (2 mJy, see Hoare et al. 2012) and the 1.5″ resolution, we estimate the CORNISH sensitivity to $\log(EM)$ to be $\sim$5.5 pc cm^{-6}.

The computed UCHII emission measures and electron densities are generally consistent with the results by Kurtz et al. (1994). No sources in the sample have computed electron densities and emission measures that would exceed 10^5 cm^{-3} and 10^8 pc cm^{-6}, respectively, even when taking into account the associated uncertainties. Hyper-compact regions have electron densities in excess of 10^6 cm^{-3} and emission measures in excess of 10^{10} pc cm^{-6} (Hoare et al. 2007). Thus, the CORNISH survey has identified only UCHIIs; no HCHIIs are reported. This is

not surprising, because the high density of HCHIIs implies high turnover frequencies, $\sim$30 GHz. For an optically thick free-free spectrum with $S_\nu \propto \nu^2$, the flux density at the CORNISH observing frequency of 5 GHz will be of the order of 40× lower than the flux density near the turnover frequency.

The computed optical depths at 5 GHz were used to quantify how much of the 5 GHz flux density would be missed in (the few) potential cases of optically thick H II regions (as well as how much this effect varies from source to source). In 80% of all the H II regions observed at 5 GHz, the difference between the measured and the theoretical unattenuated flux is below 10%. The bright UCHII G049.4905−00.3688 has the highest computed difference ($\sim$56.4%). The distribution of the flux difference due to attenuation tapers off above differences greater than $\sim$20%, indicating that there is most likely no significant fraction of sources that have been missed altogether. The same should be true even if the electron temperature varies from region to region (within the expected physical bounds). It should also be noted that the computed Lyman continuum flux in this case is not significantly underestimated due to optically thick free-free emission, but could still be affected by loss of ionising photons (e.g. via dust absorption), or for radio flux that was not recovered.

6. Infrared properties

6.1. Associations with mid-infrared data

As discussed by Watson et al. (2008), the emission detected in the vicinity of a hot star is dominated by a different emission process in each of the four mid-IR GLIMPSE-IRAC bands. In particular, the presence of PAH emission is based on 8 μm, 5.8 μm, and lack of 4.5 μm emission. Similar to larger IR bubbles, the mid-IR images of UCHII regions also show the presence of 8 μm shells, dominated by strong PAH features in the IRAC bands at 3.6, 5.8, and 8 μm. The inner surface of each shell is located at a distance from the ionising star that is equal to the PAH destruction radius. The 3.6 μm band is dominated by stars, with contributions from the diffuse PAH feature at 3.3 μm and perhaps from scattered starlight. The 4.5 μm band exhibits no PAH features; the brightest contributors are stars and the diffuse emission is due to the shared contribution of lines from H II regions and from shocked molecular gas. The diffuse emission in the 5.8 μm band is dominated by the 6.2 μm PAH feature, apart from the immediate vicinity of O stars (where PAHs are destroyed). The diffuse emission in the 8 μm band is dominated by the 7.7 and 8.6 μm PAH features, or by thermal emission of dust heated by the hot stars and Lyman-α photons.

The CORNISH survey was designed to cover the GLIMPSE region of the Galactic plane (Hoare et al. 2012), ensuring that all sources have mid-infrared counterparts to the radio continuum sources. Photometry of all GLIMPSE UCHII sources in the four IRAC bands was performed. Selecting the correct size of the IR source would not have been straightforward without knowledge of the position and size of the ultra-compact radio source, as the IR environment is more complex than in the radio view due to the different contributions to the emission. It is also difficult to disentangle individual sources in busy neighbourhoods.

In order to use the hand-drawn polygons (in the case of extended radio sources), and Gaussian outlines (in the case of compact radio sources) as apertures for the IR data, they were expanded accordingly. The exact padding value necessary for each of the four GLIMPSE bands was chosen after measuring
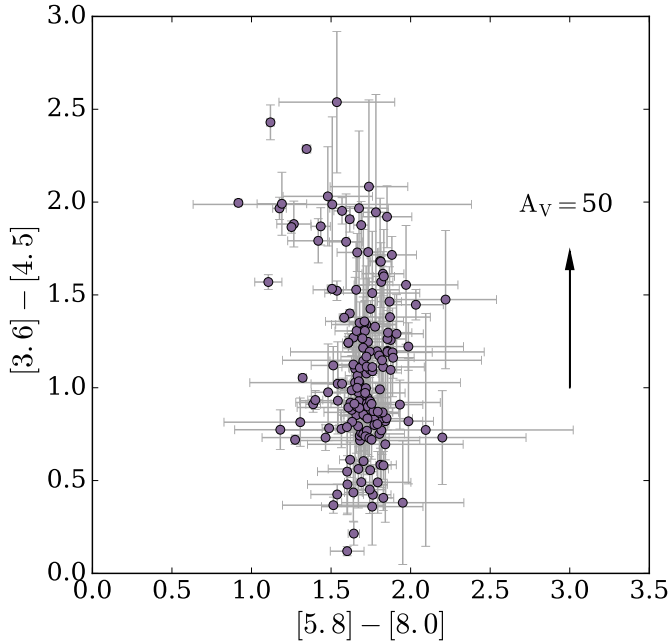
Fig. 8. Colour-colour diagram of the CORNISH UCHIIs (excluding all sources with one or more saturated GLIMPSE images or otherwise unreliable GLIMPSE fluxes). The arrow shows the reddening vector, based on the extinction law of Indebetouw et al. (2005). The plot is in the Vega magnitude system.

Fig. 9. Near-IR colour-colour diagram of near-IR nebulae associated with the CORNISH UCHIIs. The stellar contamination has been removed. Lower limits for sources visible only in H and K are also shown. The arrow shows the extinction vector, calculated using the extinction law from Stead & Hoare (2009). The predicted intrinsic colours of an ionised nebula are shown with a black triangle. The observed data and the intrinsic colours are shown in the AB magnitude system.

(for each band) the counts at different aperture sizes and examining where the curve of growth begins to plateau before starting to increase again with the inclusion of unrelated sources. This was done for a small sub-sample of sources with a range of sizes representative of the UCHII sample, in all four bands. The padding radii used for the 3.6, 4.5, 5.8, and 8 μm images were 2″, 2.4″, 3.4″, and 4.3″, respectively. Some of the measured fluxes are included in Table C.1 (the full table is available at the CDS). Median absolute deviation from the median (MADFM) background estimation was utilised (for all photometry in this work), as it is insensitive to the presence of outliers and is a reliable estimate of the noise (Purcell et al. 2013)[5]. From all 956 GLIMPSE images (all four bands), 36 8.0 μm images were found to be saturated at the source location after visual inspection, as well as two 5.8 μm images and one 4.5 μm image. These, together with non-detections, were excluded from the final results. The photometric results for G031.2801+00.0632 were also excluded, as diagnostic diagrams showed them to be dominated by a neighbouring YSO (seen in all GLIMPSE bands) rather than the UCHII region. This left 180 sources with 3.6 μm fluxes, 191 – with 4.5 μm fluxes, 190 – with 5.8 μm fluxes, and 184 – with 8.0 μm fluxes.

A colour-colour plot is shown in Fig. 8. Only sources for which it was possible to compute both the [3.6]−[4.5] and [5.8]−[8.0] colours were included (and no upper or lower limits for the remaining sources), to avoid overcrowding the plot. This was not found to affect the exhibited trends for the mid-IR colours. From the 174 sources with reliable flux values in all bands, ~85% occupy the zone 1.5 < [5.8]−[8.0] < 2. The [3.6]−[4.5] colour ranges between 0.1 and 2.1. This is similar to the results reported by de la Fuente et al. (2009) for

19 ultra-compact H II regions. They find that about 75% of the UCHIIs are grouped around [5.8]−[8.0] ≃ 1.7 and 0.5 ≲ [3.6]−[4.5] ≲2.0.

6.2. Associations with near-infrared data and extinctions

The visual extinction in the line of sight to UCHIIs can be estimated from their J, H, and K fluxes, assuming the diffuse emission is due to purely nebular gas emission. Several methods can be utilised to achieve this, such as the use of near-IR colours, or ratios of the near-IR flux to the radio flux (Willner et al. 1972). When using the colour-dependent methods, provided sufficiently reliable magnitude measurements of the embedded sources, one's choice comes down to a compromise between scattering effects and infrared excess. Dust excess (typically pronounced in the K-band) causes the H−K colours to appear redder, and scattered light (due to dust grains) at shorter wavelengths results in bluer J−H colours (Porter et al. 1998). It is useful to compare the results from the different methods and in this way weigh the severity of systematic issues while providing an extinction range for the studied sources.

The UKIDSS Galactic plane survey (Lucas et al. 2008) covered the northern and equatorial Galactic plane at |b| < 5° in the J (1.17−1.33 μm), H (1.49−1.78 μm), K (2.03−2.37 μm) bands and provides an opportunity to investigate the near-IR properties of the CORNISH UCHII sample, such as fluxes and detection statistics. Only a point-source UKIDSS catalogue is available at present. Therefore automated photometry was performed in the same manner as described in Sect. 6.1 on all CORNISH sources with available UKIDSS data. In total, 230 sources had available UKIDSS images in the J band, 228 – in H, 227 – in K. The visual

---

[5] All sources were detected at at least 9 times the median background value. Errors were computed following Masci (2009) and include prior (noise-model) and derived uncertainties.
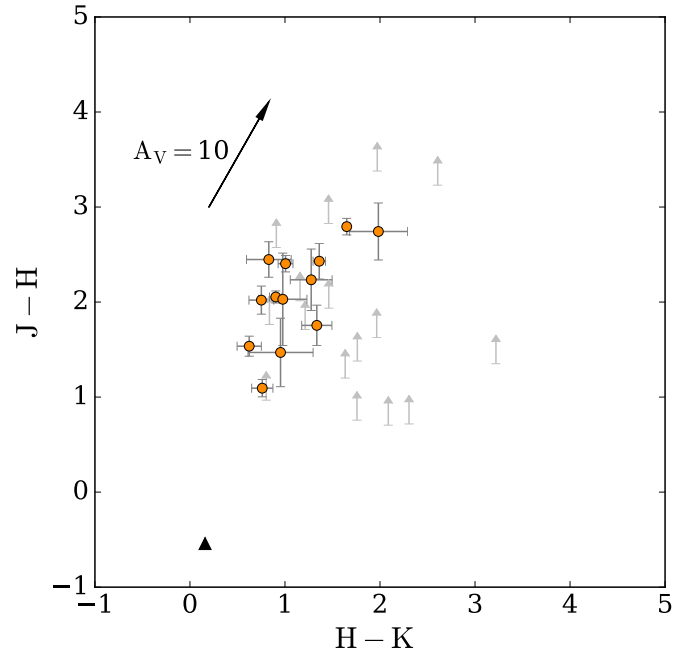
Table 5. Comparison between some extinction values in this work and in the literature, derived from comparison to 6 cm, Brackett-γ line ("Brγ"), or multi-configuration radio observations ("radio").

CORNISH name	x	A_x (mag)	Literature name	x	A_x (mag)	Reference
G029.9559−00.0168	K/6 cm	2.33 ± 0.07	G29.96−0.02	K/radio	2.14 ± 0.08	Watson et al. (1997)
				Brγ	2.20 ± 0.25	
				Brγ	2.16 ± 0.07	Moore et al. (2005)
G043.8894−00.7840	K/6 cm	2.83 ± 0.12	G43.89−0.78	Brγ	3.32 ± 0.21	Moore et al. (2005)
G045.4545+00.0591	K/6 cm	2.21 ± 0.11	G45.45+0.06	K/6 cm	2.5	Feldt et al. (1998)
G049.4905−00.3688	K/6 cm	2.33 ± 0.07	W51d	K/6 cm	2.6 ± 0.3	Goldader & Wynn-Williams (1994)
				Brγ	1.59 ± 0.07	Moore et al. (2005)

Notes. Our results and the referenced literature values are in good agreement with extinction-law fits to H II recombination-line data (Moore et al. 2005).

inspection revealed that out of all sources, 83 have a visible nebula in K, out of which 31 also have an H-band nebula. Out of these, 14 nebulae are visible in J[6]. Aperture photometry (with median background subtraction) of all contaminant bright stars found within the expanded polygon aperture used for the automated J, H, and K flux measurements was also performed. This was done for all images with a visible near-IR nebula coinciding with an UCHII. The measured stellar fluxes were subtracted from the total photometric fluxes in order to obtain the nebular fluxes. Some of the near-IR UCHII fluxes and corresponding AB magnitudes are presented in Table C.2. The full table is available at the CDS.

Figure 9 shows a diagram of the J–H against H–K nebular colours. The mean colour of the nebulae visible in all three bands is 2.1 for J–H, and 1.1 for H–K.

Figure 10 presents a comparison between the computed visual extinctions for the UCHII sample obtained from four methods: F_H/F_{5GHz} and F_K/F_{5GHz} (top panel); J–H and H–K (bottom panel). The empirically-derived R_V-dependent extinction law A_λ/A_V from Cardelli et al. (1989; Eqs. (1-3b)) was used to convert from near-IR to visual extinction. The (standard for the ISM) optical total-to-selective extinction ratio $R_V = 3.1$ was assumed. This has been found to reach values of 5-6 towards dense clouds (e.g. Cardelli et al. 1989) but using $R_V = 5$ did not affect the results within error.

The near-IR extinction was computed from the difference between the measured and expected near-IR magnitudes. For the flux-ratio methods, the expected H- and K-band fluxes were obtained by utilising the intrinsic ratios between IR and radio flux found by Willner et al. (1972), $F_H/F_{5\,GHz} = 0.26$ and $F_K/F_{5\,GHz} = 0.3$. Using the Willner ratios, a value of 0.68 for H–K was computed in this work. In order to obtain the J–H extinction, $F_J/F_{5\,GHz} = 0.43$ was computed, using Eq. (1) and Table 2 in Brussaard & van de Hulst (1962), taking into account the significant Paschen-β line contribution to the J band (Hummer & Storey 1987). This resulted in a value of −0.1 for J–H. These predicted intrinsic J–H and H–K colours of ionised nebulae agree well with those from near-IR photometry of planetary nebulae (with 2MASS data) studied by Ramos-Larios & Phillips (2005): H–$K = 0.65$ and J–$H \sim -0.1$. Weidmann et al. (2013) also found comparable values, using PNe in the VVV survey[7]: H–$K = 0.62$ and J–$H = 0.0$. This comparison is in the Vega system, taking into account the 2MASS (Vega system) and UKIDSS (AB system) magnitude offsets for each band. Our

computed intrinsic AB colours, J–$H = -0.54$ and H–$K = 0.15$, were converted to the Vega system using Table 7 in Hewett et al. (2006): the AB offsets for the J, H, and K bands are 0.938, 1.379, and 1.9, respectively.

The presented results from the four different extinction methods are consistent with previous estimates of visual extinctions towards UCHIIs, which are ∼0–50 mags (see e.g. Hanson et al. 2002). Moore et al. (2005) calculated the extinction for a number of compact and ultra-compact H II regions, using observed hydrogen recombination lines. A comparison between extinction results from this work and from literature is presented in Table 5, showing good agreement.

As can be seen in Fig. 10, the offsets between the extinctions obtained from the different methods are clearly systematic. Values derived using the K band are ∼10 magnitudes lower (i.e. brighter) than those using H (and J). An addition of ∼10 magnitudes to the K band would bring the F_H/F_{5GHz} and F_K/F_{5GHz} methods to agreement and eliminate most of the unphysical negative values. Such an addition would mimic eliminating the expected boost to the K band from the contribution of very hot dust in the vicinity of the ionising star. However, a ∼10-magnitude addition to K in the $A_{V,(J-H)}$ vs. $A_{V,(H-K)}$ diagram to exclude potential hot dust contribution would actually result in a ∼20-magnitude systematic offset between these two methods. Such a discrepancy likely stems from the general nature of each pair of utilised methods and highlights the need for in-depth investigation, preferably on a case-to-case basis (beyond the scope of this work).

Two of the UCHIIs with computed extinction have coinciding *XMM-Newton* hard X-ray sources. The extinction was computed independently from X-ray spectral fitting (which also revealed that the sources have $k_B T > 2$ keV, i.e. $T > 10^7$ K). The H–K-derived visual extinction for G030.7661−00.0348 is ∼35 mags, whereas the X-ray-fitted hydrogen column density translates to $A_V \approx 67$ mags (towards the W43 star cluster as a whole). G025.3824−00.1812 has H–K − derived $A_V \approx 12$, and the value derived from the X-ray spectral fit is ∼11 mags[8].

The full table with extinction results is available at the CDS, an example can be found in Table C.3.

7. Spectral energy distributions and bolometric luminosities

The computed UKIDSS and GLIMPSE fluxes were combined with multi-wavelength data from MSX (21 μm), WISE (22 μm),

---

[6] G023.9564+00.1493 has a visible nebula in the near-IR, but photometric issues due to image quality lead to unreliable flux values, and the source is not included in the final results table.

[7] https://vvvsurvey.org/

---

[8] The X-ray-derived value is the same also for G025.3809−00.1815, as the two neighbouring sources are not resolved by *XMM-Newton*.

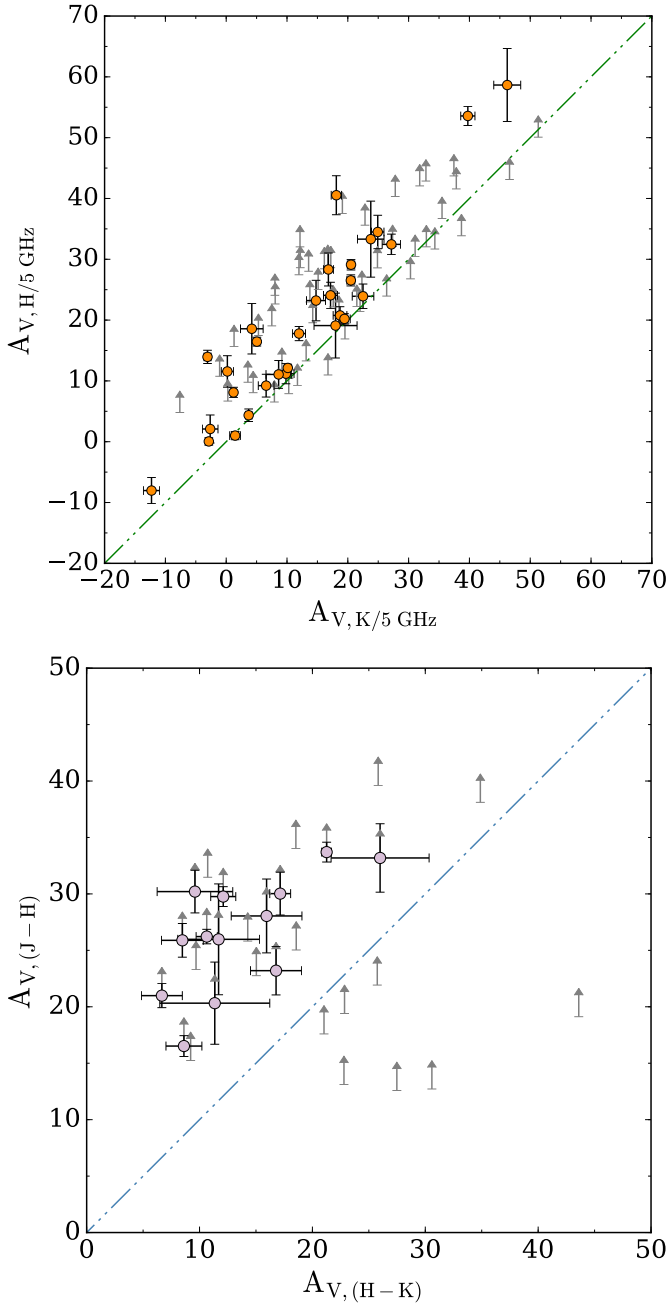**Fig. 11.** Average SED for all UCHIIs with available multi-wavelength data, normalised to the sample median distance of 9.8 kpc. The plot includes the *J*, *H*, and *K* UKIDSS fluxes and the 3.6, 4.5, 5.8, and 8.0 μm GLIMPSE fluxes from this work, together with fluxes from MSX (21 μm), WISE (22 μm), HiGAL (70, 160, 250, 350, and 500 μm), and ATLASGAL (870 μm; see Table A.1 in Cesaroni et al. 2015).

Fig. 12. Bolometric luminosity distribution for the UCHIIs with fitted SEDs.

Fig. 10. Visual extinctions of the UCHIIs, computed from their near-IR fluxes from four methods – F_H/F_{5GHz}, F_K/F_{5GHz}, *J–H*, and *H–K*. The top panel compares the *H*- and *K*-band derived A_V. In the *bottom panel*, the comparison is shown for the *J–H* against the *H–K* derived A_V. Lower limits on the extinction are shown in the cases where the UCHII is seen only in H and K (grey arrows). Lines of equality are also plotted.

HiGAL (70, 160, 250, 350, and 500 μm), and ATLASGAL (870 μm). These data were available for 177 UCHIIs and the SEDs were reconstructed for these sources. The majority of the SEDs have very reasonable shapes and exhibit the same average shape. There are a few SEDs with irregularities, typically the flux at 350 μm and 22 μm. There are many SEDs (90/177) where the 4.5 μm flux is low (comparable to the 3.6 μm flux), which is most likely caused by the gap in PAH emission at this wavelength. The average SED for the sample (when all sources are placed at the median distance of 9.8 kpc) is presented in Fig. 11.
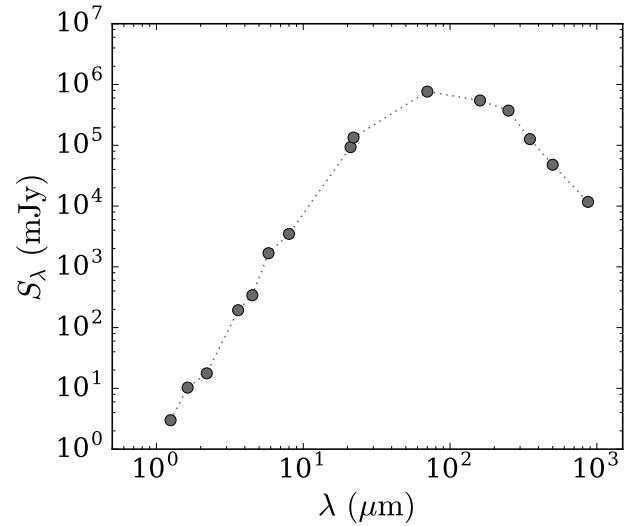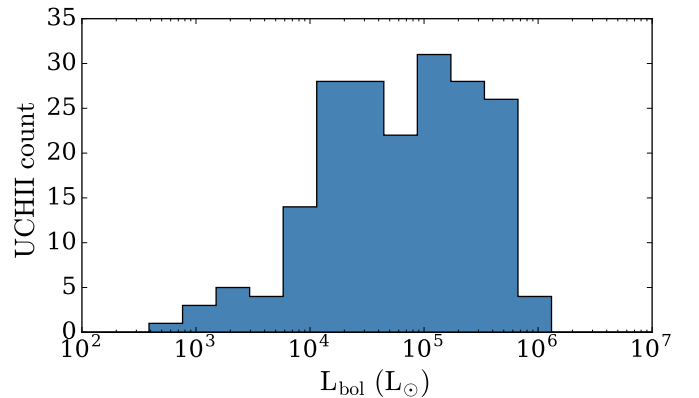
The bolometric luminosities of the CORNISH UCHIIs were calculated from fitting the SEDs of the sources with available data across multiple wavelengths. The SED Fitter from Robitaille et al. (2007) was utilised, following the procedure prescribed by Mottram et al. (2011). The complete sample was split into subsamples according to their distance (as computed in this work) to limit the distance range when running the fits. The fitted bolometric luminosities were then converted to the final bolometric luminosity values by replacing the automatically fitted distances with the distances presented in this work. The results are shown in Fig. 12. These are in good agreement with Fig. 19 in Urquhart et al. (2013), which summarises the RMS bolometric luminosities of 135 associated clumps. We note that the models used in the SED fitting are tailored to YSOs and do not take into account additional sources of dust heating in the ionised region, such as Lyman-α (Hoare et al. 1991). However, the SEDs of UCHIIs and YSOs are sufficiently similar up to radio wavelengths (which are not included in the models) and the fits are useful for computing the bolometric luminosities.

Possible contamination by MYSO jet emission was investigated with a plot of radio luminosity at 5 GHz against bolometric
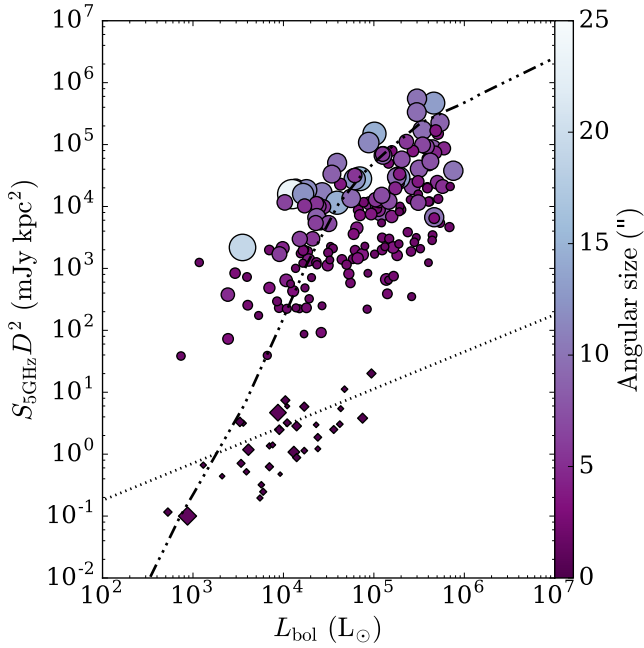
Fig. 13. 5 GHz radio luminosity against bolometric luminosity for the UCHIIs with fitted SEDs, shown with circles. A sample of confirmed ionised MYSO jets by Purser (2017) is also included (indicated by diamonds). The dash-dotted line marks the expected radio luminosity from Lyman continuum emission. The dotted line shows the empirical relation extrapolated for low-mass YSO jets (Anglada 1995). The angular size of all sources is colour-coded and emphasized through the marker sizes (the jet sample marker sizes are exaggerated, as their resolution is higher).

luminosity of the UCHII regions and comparing to a sample of confirmed MYSO jets by Purser (2017). This is presented in Fig. 13. The UCHII regions are brighter, as expected, and the two populations clearly separate. The predicted optically-thin 5 GHz radio luminosities corresponding to a range of bolometric luminosities, from the stellar models by Thompson (1984) for $L_{bol} \leq 10^3 \, L_{\odot}$ and from the models summarised in Table 1 of Davies et al. 2011 for $L_{bol} > 10^3 \, L_{\odot}$, are shown with a dash-dotted line in the plot. The radio continuum flux can be inferred from the Lyman continuum flux, as the latter determines the amount of ionised material and thus the number of free electrons participating in the thermal bremsstrahlung process.

The majority of CORNISH UCHIIs have radio luminosity that is between 1% and 100% of the theoretical value at the corresponding H II region bolometric luminosity (i.e. below the model line). A mixture of UCHII angular sizes are found at equal fractions of the model luminosity, ruling out angular size as the culprit behind the large variation in observed luminosity. It is to be expected that a large portion of UCHIIs would be dimmer than predicted, as the stellar models do not account for the portion of Lyman continuum flux that is absorbed by dust.

A significant number of sources (about a third) can be found above the line, with radio luminosities up to ten times or more than those predicted. The presence of H II regions in the "forbidden" area above the model line is attributed to a Lyman excess. There are different explanations for such H II regions, as discussed, for example, by Sánchez-Monge et al. (2013), Cesaroni et al. (2015, 2016), and in the references therein. One explanation is the assumption of spherical symmetry, whereas in reality Lyman photons could be leaking in directions tracing lower gas density, which could be away from the line of sight. It is unlikely for the contribution to be from overlapping 5σ or 7σ

UCHIIs, as no systematic trends were found in their distribution in the plot. Cesaroni et al. (2016) used molecular tracers to look for outflows and accretion shocks in the vicinity of 200 CORNISH H II regions, and found no evidence to support any outflow-related phenomenon. Instead, they found that Lyman-excess sources are more associated with infall than non-excess sources, and propose ongoing accretion and accretion shocks as an explanation, but their HCO^+ measurements are only sensitive to large-scale infall ($\lesssim 1$ pc), which is not direct evidence of accretion.

Two sources, G065.2462+00.3505 and G011.0328+00.0274, were found to have computed bolometric luminosities below the lower limit for H II region formation – that is, for a B3 type star with $L_{bol} \sim 10^3 \, L_{\odot}$ (Boehm-Vitense 1981; Meynet & Maeder 2003). In the case of G011.0328+00.0274, this is likely due to an individual issue with the distance determination or the bolometric flux determination, as the radio-to-bolometric flux ratio is not extreme. G065.2462+00.3505, originally in the UCHII catalogue, was reclassified as a radio star. Although this unresolved radio source is embedded within an IR nebula $\sim 50''$ in diameter, there is only a near-IR stellar counterpart but no trace of a compact counterpart at 8 μm or 70 μm.

It should be noted that 6 sources (G011.9786−00.0973, G014.5988+00.0198, G018.6654+00.0294, G031.1590+00.0465, G036.4062+00.0221, and G043.7960−00.1286) could not be shown in the plot. No bolometric luminosity was computed for them as the available supplementary data was not sufficient to build a good portion of their SEDs for a fit. Their $S_{5GHz}D^2$ values range from ~ 34 to ~ 828 mJy kpc^2, and would thus still separate from the MYSO sample in the plot.

8. Comparison of UCHII search methods in blind surveys

We argue that high resolution blind radio surveys are the most reliable way to obtain the UCHII population census of the Milky Way. With the CORNISH sample of genuine UCHII regions, our total estimate is ~ 750 UCHIIs in the Galaxy. This was obtained by scaling up the sample size (239 sources) by a geometric correction factor of ~ 3.1 – the ratio between the detected number of RMS UCHII regions in the total RMS and CORNISH area – 500 and 160, respectively (see Urquhart et al. 2008, Lumsden et al. 2013). Colour-selected RMS sources with detectable radio emission were classified as UCHIIs due to their mid-IR morphologies. Practically the same scaling factor is obtained when taking the total RMS sample of 900 H II regions and the 297 RMS H II regions in the CORNISH area. The RMS survey encompasses $10° < l < 350°$ and $|b| < 5°$ and provides the best current map of the non-uniform distribution of massive star formation throughout the Galaxy. UCHII regions exhibit the same Galactic scale-height, $\sim 0.6°$, in CORNISH and RMS. Since the RMS UCHII counts in the total and in the CORNISH area were obtained in the same manner, their ratio should not be greatly affected by the systematic limitations of the IR-selection. We note that UCHIIs near the Galactic centre are missed in our total estimate, to avoid assumptions for the UCHII number density in this region not covered by RMS.

The sections below highlight limitations of the other UCHII search methods employed in Galactic plane surveys.

8.1. Sub-mm − ATLASGAL

Urquhart et al. (2013) used ATLASGAL-CORNISH associations to compute the surface density of UCHII regions as a function of

Galactocentric distance. They estimated that the Galactic UCHII population comprises ~400 sources around B0 and earlier type stars, out of which only ~45 around O6 or earlier type stars are detectable. However, the depth of ATLASGAL falls short of detecting all UCHII regions within the common ATLASGAL-CORNISH area – about 30 UCHII regions were missed as a result. The deeper SCUBA-2 survey should provide a higher detection certainty with the same search method.

8.2. Mid-IR – RMS

The whole CORNISH region is covered in RMS; despite this, about half of the CORNISH UCHIIs have RMS counterparts. There are ~40 further associations with the more diffuse CORNISH H II sample (48 H II regions). Many of the UCHIIs "missing" from RMS are, in fact, detected within big complexes but not listed as individual objects. Urquhart et al. (2013) discussed the larger total number of RMS versus ATLASGAL H II regions – the RMS H II sample contains extended regions identified from their mid-IR morphology, with radio continuum emission lower than the CORNISH surface brightness sensitivity. The RMS sample thus includes fewer individually listed UCHII regions, together with a number of more extended H II regions.

8.3. Previous radio surveys of the Galactic plane

Giveon et al. (2008) found 494 MSX matches to the White et al. (2005) catalogue within the area shared with CORNISH (~23%). Giveon et al. (2005) believe their sample is dominated by UCHII regions. As discussed in Sect. 3.1, through visual inspection we only found 162 CORNISH UCHII regions in common with the lower-resolution 6 cm data. A catalogue table cross-match found even fewer sources in common – 111, with 20 cross-matches to 37 CORNISH diffuse H II regions within the shared area. Thus we are finding a factor of 3–4 fewer UCHIIs than implied by Giveon et al. (2005, 2008) – the vast majority of their sample are not genuine ultra-compact H II regions but rather represent more extended H II region phases.

9. Summary

The CORNISH UCHII sample is the largest complete and unbiased high-resolution collection of ultra-compact H II regions to date. Within the mapped region ($10° < l < 65°, |b| < 1°$), 239 UCHIIs have been confirmed from 240 candidates visually identified at 5 GHz radio-continuum emission. In this work, we explored the observational properties, spectral indices, physical characteristics and SEDs of this early stage of massive star formation. In summary:

1. The selection procedure for the CORNISH UCHIIs is robust and the nature of the sample as a whole was reliably identified.

2. The majority (~82%) of UCHIIs have spectral indices that are consistent with the expected theoretical limits for thermal free-free emission. The instances of non-thermal spectral indices could be naturally resulting from the difference in VLA configuration between the higher- and lower-frequency datasets, or are the result of combining 6 cm and 20 cm fluxes of variable thermal sources from two epochs.

3. We conclude that at least 5% of UCHIIs have exhibited a significant flux increase (by ~50% or more) between two observational epochs separated by ~15 years.

4. Distances were computed for 21 UCHIIs which had no literature distance (or their KDA had not been previously resolved) prior to this work. The derived physical properties of the UCHII sample agree well with theoretical expectations.

5. We have presented results of extended source photometry of UCHII regions in the mid- and near-IR. The GLIMPSE and UKIDSS colours of the sample follow the expected trends set by results obtained from earlier, smaller samples. We expect the mid-IR results to be particularly reliable, as they combine the precise knowledge of position, radio size, and shape provided by the CORNISH survey with the good correspondence (in the vast majority of cases) to the mid-IR counterparts. The results of the extended near-IR photometry (particularly the J and H bands) should be used with much more care, due to the difficulty in accurate subtraction of the stellar contamination in the busy, diffuse environments of star forming regions seen at these wavelengths.

6. Extinctions towards the UCHII regions were computed using the intrinsic H- and K-band to radio flux ratio from Willner et al. (1972), as well as from the J–H and H–K nebular colours.

7. The average SED of the UCHII sample (from gathering available multi-wavelength data and combining them with the new near- and mid-IR results) is in excellent agreement with the expected shape (see e.g. Faison et al. 1998, Fig. 1 in Hoare et al. 2012), with a peak between 70 and 160 μm. Bolometric luminosities were computed by fitting the individual SEDs. In a plot of radio luminosity against bolometric luminosity, the CORNISH UCHII sample is clearly a separate population to confirmed MYSO jets. About a third of the UCHIIs exhibit a Lyman excess.

8. High resolution blind radio surveys are the best way to definitively find the UCHII population of the Galaxy. Radio selection provides a more reliable statistic than infrared and mm selection. We found a factor of 3–4 fewer genuine ultra-compact H II regions than in previous lower resolution radio areal surveys, which, in conjunction with up-to-date models (see Davies et al. 2011), goes towards alleviating the *lifetime problem* posed by Wood & Churchwell (1989a).

References

Acord, J. M., Churchwell, E., & Wood, D. O. S. 1998, ApJ, 495, L107
Anderson, L. D., & Bania, T. M. 2009, AJ, 690, 706
Anderson, L. D., Bania, T. M., Jackson, J. M. et al. 2009, ApJS, 181, 255
Anglada, G. 1995, Rev. Mex. Astron. Astrofís., 1, 67
Becker, R. H. et al. 1994, Bull. Am. Astron. Soc., 26, 915
Benjamin, R. A., Churchwell, E., Babler, B. L. et al. 2003, PASP, 115, 953
Boehm-Vitense, E. 1981, ARA&A, 19, 295
Bojičić, I. S., Parker, Q. A., Filipović, M. D., & Frew, D. J. 2011, MNRAS, 412, 223
Brand, J., & Blitz, L. 1993, A&A, 275, 67
Brussaard, P. J., & van de Hulst, H. C. 1962, Rev. Mod. Phys., 34, 507
Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, ApJ, 345, 245
Carey, S. J., Noriega-Crespo, A., Mizuno, D. R. et al. 2009, PASP, 121, 76
Cesaroni, R., Pestalozzi, M., Beltrán, M. T. et al. 2015, A&A, 579, A71
Cesaroni, R., Sánchez-Monge, Á., Beltrán, M. T. et al. 2016, A&A, 588, L5
Churchwell, E., Babler, B. L., Meade, M. R. et al. 2009, PASP, 121, 213
Davies, B., Hoare, M. G., Lumsden, S. L. et al. 2011, MNRAS, 416, 972

de la Fuente, E., Porras, A., Grave, J. M. C. et al. 2009, Rev. Mex. Astron. Astrofis., 37, 13

de Pree, C. G., Wilner, D. J., Deblasio, J. et al. 2005, ApJ, 624, L101

Faison, M., Churchwell, E., Hofner, P. et al. 1998, ApJ, 500, 280

Feldt, M., Stecklum, B., Henning, T. et al. 1998, A&A, 339, 759

Franco-Hernández, R., & Rodríguez, L. F. 2004, ApJ, 604, L105

Galván-Madrid, R., Rodríguez, L. F., Ho, P. T. P., & Keto, E. 2008, ApJ, 674, L33

Galván-Madrid, R., Peters, T., & Keto, E. et al. 2011, MNRAS, 416, 1033

Garay, G., Rodríguez, L. F., Moran, J. M., & Churchwell, E. 1993, ApJ, 418, 368

Gibb, E. L., Whittet, D. C. B., Boogert, A. C. A., & Tielens, A. G. G. M. 2004, ApJS, 151, 35

Giveon, U., Becker, R. H., Helfand, D. J., & White, R. L., 2005, AJ, 129, 348

Giveon, U., Becker, R. H., & White, R. L. 2008, AJ, 135, 1697

Goldader, J. D., & Wynn-Williams, C. G. 1994, ApJ, 433, 164

Gómez, L., Rodríguez, L. F., Loinard, L. et al. 2008, ApJ, 685, 333

Hanson, M. M., Luhman, K. L., & Rieke, G. H. 2002, ApJS, 138, 35

Helfand, D. J., Becker, R. H., White, R. L., Fallon, A., & Tuttle, S. 2006, AJ, 131, 2525

Hewett, P. C., Warren, S. J., Leggett, S. K., & Hodgkin, S. T. 2006, MNRAS, 367, 454

Hoare, M. G., Roche, P. F., & Glencross, W. M. 1991, MNRAS, 251, 584

Hoare, M. G., Kurtz, S. E., Lizano, S., Keto, E., & Hofner, P. 2007, Protostars and Planets, 5, 181

Hoare, M. G., Purcell, C. R., Churchwell, E. B. et al. 2012, PASP, 124, 939

Hummer, D. G., & Storey, P. J. 1987, MNRAS, 224, 801

Indebetouw, R., Mathis, J. S., Babler, B. L. et al. 2005, ApJ, 619, 931

Jackson, J. M., Rathborne, J. M., Shah, R. Y. et al. 2006, ApJS, 163, 145

Klassen, M., Pudritz, R. E., & Peters, T. 2012, MNRAS, 421, 286

Kurtz, S., Churchwell, E., & Wood, D. O. S. 1994, ApJS, 91, 659

Lucas, P. W., Hoare, M. G., Longmore, A. et al. 2008, MNRAS, 391, 136

Lumsden, S. L., Hoare, M. G., Urquhart, J. S. et al. 2013, ApJS, 208, 17

Masci, F. 2009, Aperture Photometry Uncertainties assuming Priors and Correlated Noise, version 2.0

Meynet, G., & Maeder, A. 2003, A&A, 404, 975

Mezger, P. G., & Henderson, A. P. 1967, ApJ, 147, 471

Miralles, M. P., Rodriguez, L. F., & Scalise, E. 1994, ApJS, 92, 173

Molinari, S., Bally, J., Glover, S. et al. 2014, Protostars and Planets, 6, 125

Moore, T. J. T., Lumsden, S. L., Ridge, N. A., & Puxley, P. J. 2005, MNRAS, 359, 589

Mottram, J. C., Hoare, M. G., Urquhart, J. S. et al. 2011, A&A, 525, A149

Panagia, N., & Walmsley, C. M. 1978, A&A, 70, 411

Peters, T., Banerjee, R., Klessen, R. S. et al. 2010a, ApJ, 711, 1017

Peters, T., Mac Low, M.-M., Banerjee, R., Klessen, R. S., & Dullemond, C. P. 2010b, ApJ, 719, 831

Peters, T., Klessen, R. S., Mac Low, M.-M., & Banerjee, R. 2010c, ApJ, 725, 134

Porter, J. M., Drew, J. E., & Lumsden, S. L. 1998, A&A, 332, 999

Price, S. D., Egan, M. P., Carey, S. J., Mizuno, D. R., & Kuchar, T. A. 2001, AJ, 121, 2819

Purcell, C. R., Hoare, M. G., Cotton, W. D. et al. 2013, ApJS, 205, 1

Purser, S. J. D. 2017, PhD Thesis, University of Leeds, UK

Purser, S. J. D., Lumsden, S. L., Hoare, M. G. et al. 2016, MNRAS, 460, 1039

Quireza, C., Rood, R. T., Bania, T. M., Balser, D. S., & Maciel, W. J. 2006, ApJ, 653, 1226

Ramos-Larios, G., & Phillips, J. P. 2005, MNRAS, 357, 732

Reid, M. J., Menten, K. M., Zheng, X. W. et al. 2009, ApJ, 700, 137

Robitaille, T. P., Whitney, B. A., Indebetouw, R., & Wood, K. 2007, ApJS, 169, 328

Rodríguez, L. F., Gómez, Y., & Tafoya, D. 2007, ApJ, 663, 1083

Rosolowsky, E., Dunham, M. K., Ginsburg, A. et al. 2010, ApJS, 188, 123

Rubin, R. H. 1968, ApJ, 154, 391

Sánchez-Monge, Á., Beltrán, M. T., Cesaroni, R. et al. 2013, A&A, 550, A21

Schuller, F., Menten, K. M., Contreras, Y. et al. 2009, A&A, 504, 415

Siódmiak, N., & Tylenda, R. 2001, A&A, 373, 1032

Skrutskie, M. F., Cutri, R. M., Stiening, M. D. et al. 2006, AJ, 131, 1163

Stead, J. J., & Hoare, M. G. 2009, MNRAS, 400, 731

Stil, J. M., Taylor, A. R., Dickey, J. M. et al. 2006, ApJ, 132, 1158

Thompson, R. I. 1984, ApJ, 283, 165

Urquhart, J. S., Busfield, A. L., Hoare, M. G. et al. 2007, A&A, 461, 11

Urquhart, J. S., Hoare, M. G., Lumsden, S. L., Oudmaijer, R. D., & Moore, T. J. T. 2008, ASP Conf. Ser., 387, 381

Urquhart, J. S., Hoare, M. G., Lumsden, S. L. et al. 2009, A&A, 501, 539

Urquhart, J. S., Moore, T. J. T., Hoare, M. G. et al. 2011, MNRAS, 410, 1237

Urquhart, J. S., Hoare, M. G., Lumsden, S. L. et al. 2012, MNRAS, 420, 1656

Urquhart, J. S., Thompson, M. A., Moore, J. T. et al. 2013, MNRAS, 435, 400

Watson, A. M., Coil, A. L., Shepherd, D. S., Hofner, P., & Churchwell, E. 1997, ApJ, 487, 818

Watson, C., Povich, M. S., Churchwell, E. B. et al. 2008, ApJ, 681, 1341

Weidmann, W. A., Gamen, R., van Hoof, P. A. M. et al. 2013, A&A, 552, A74

White, R. L., Becker, R. H., & Helfand, D. J. 2005, AJ, 130, 586

Willner, S. P., Becklin, E. E., & Visvanathan, N. 1972, ApJ, 175, 699

Wood, D. O. S., & Churchwell, E. 1989a, ApJS, 69, 831

Wood, D. O. S., & Churchwell, E. 1989b, ApJ, 340, 265

Zinnecker, H., & Yorke, H. W. 2007, ARA&A, 481, 45

Zoonematkermani, S., Helfand, D. J., Becker, R. H., White, R. L., & Perley, R. A. 1990, ApJS, 74, 181
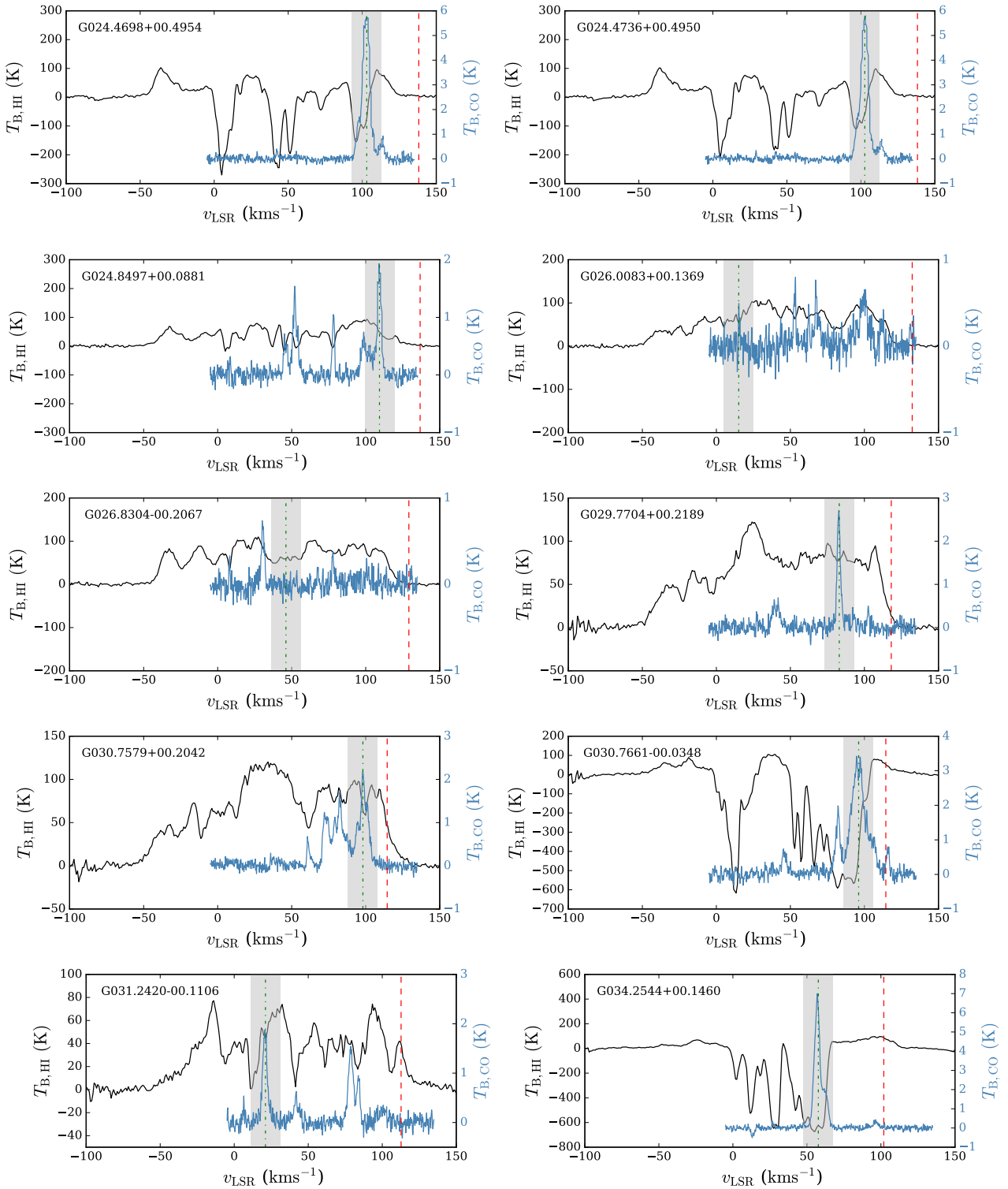
Appendix A: KDA resolution – plots

Fig. A.1. Continuum-subtracted VGPS spectra (black line) towards sources from the CORNISH UCHII region sample were used to solve their KDA. The red dashed line indicates the tangent point velocity. The blue spectrum shows the measured CO spectrum, with its corresponding *y*-axis also in blue. The green dot-dashed line represents the CO source velocity, with the region marked in grey on each side showing the expected uncertainty of ±10 km s^{-1} due to streaming motions. The H I spectrum of G024.4721+00.4877 is practically identical to G024.4698+00.4954 (*top row, left*) and therefore was omitted.
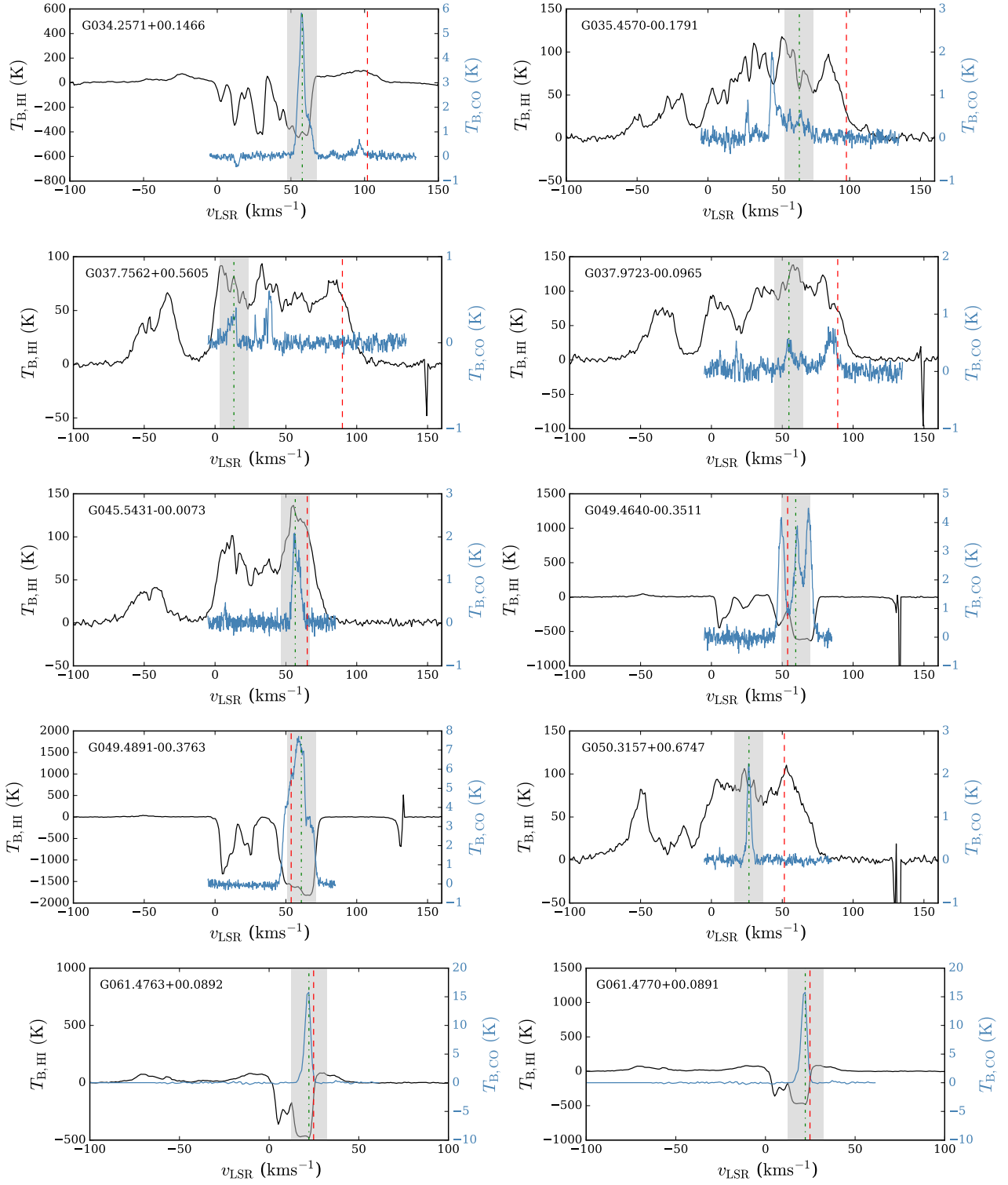
Fig. A.1. Continued.

Appendix B: Radio fluxes and spectral indices

Table B.1. Fluxes of the MAGPIS counterparts to CORNISH UCHIIs at 6 cm and 20 cm, as remeasured in this work (described in Sect. 3.1), and computed spectral indices (CORNISH 6 cm – MAGPIS 20 cm).

Name	S_C (mJy)	$S_{M,6\ cm}$ (mJy)	$S_{M,20\ cm}$ (mJy)	$\alpha_{C6\ cm-M20\ cm}$
G010.3009−00.1477	631.39 ± 59.3	1169.19 ± 9.41	456.88 ± 74.73	0.25 ± 0.15
G010.3204−00.2328	32.43 ± 5	<28.78	48.94 ± 2.56	−0.32 ± 0.13
G010.3204−00.2586	18.2 ± 2.26	47.09 ± 3.28	38.2 ± 1.97	−0.58 ± 0.11
G010.4724+00.0275	38.43 ± 4.38	92.82 ± 1.35	20.69 ± 3.36	0.49 ± 0.16
G010.4736+00.0274	19.3 ± 2.39	94.83 ± 1.01	17.68 ± 3.02	0.07 ± 0.17
...				

Notes. Missing values are due to missing MAGPIS 6 cm images. In all cases of unreliable photometric results or non-detections, the 3σ upper limits are included.

Appendix C: Extended IR source fluxes – GLIMPSE and UKIDSS

Table C.1. GLIMPSE fluxes and magnitudes of the sample of CORNISH UCHII regions, computed in this work.

Name	$S_{3.6\ \mu m}$ (mJy)	$S_{4.5\ \mu m}$ (mJy)	$S_{5.8\ \mu m}$ (mJy)	$S_{8.0\ \mu m}$ (mJy)	[3.6]	[4.5]	[5.8]	[8.0]
G010.3009−00.1477	191.81 ± 80.80	316.14 ± 99.64	1982.12 ± 712.14	5117.70 ± 2547.03	10.69 ± 0.45	10.15 ± 0.34	8.16 ± 0.39	7.13 ± 0.54
G010.3204−00.2328	107.48 ± 3.13	159.13 ± 4.36	812.79 ± 22.51	1646.41 ± 53.34	11.32 ± 0.03	10.90 ± 0.03	9.13 ± 0.03	8.36 ± 0.03
G010.3204−00.2586	8.68 ± 2.98	57.50 ± 4.64	213.39 ± 41.53	495.73 ± 136.43	14.05 ± 0.37	12.00 ± 0.09	10.58 ± 0.21	9.66 ± 0.30
G010.6218−00.3848	5.80 ± 3.48	26.34 ± 4.77	–	–	14.49 ± 0.65	12.85 ± 0.20	–	–
G010.6223−00.3788	–	53.54 ± 23.54	–	–	–	12.08 ± 0.48	–	–
...								

Notes. Missing values are due to missing data, unreliable data, and the exclusion of non-detections. The magnitudes presented here are shown in the AB system, for consistency with our UKIDSS results.

Table C.2. UKIDSS fluxes, AB magnitudes, and colours of the CORNISH UCHII regions with a visible counterpart in at least the K band.

Name	S_J (mJy)	S_H (mJy)	S_K (mJy)	J	H	K	$J–H$	$H–K$
G010.3009−00.1477	1.47	3.54	4.64 ± 1.81	15.05	13.65	12.83 ± 0.42	1.40	0.81
G010.6297−00.3380	0.89	2.18	1.38 ± 0.58	15.59	14.17	14.15 ± 0.45	1.41	0.02
G010.9584+00.0221	0.66	1.86	1.63 ± 0.77	15.92	14.35	13.97 ± 0.51	1.57	0.38
G011.1104−00.3985	2.68	4.51	26.25 ± 1.53	14.39	13.39	10.95 ± 0.06	1.01	2.43
G011.9368−00.6158	1.94	3.90	6.65 ± 1.79	14.74	13.54	12.44 ± 0.29	1.20	1.10
...								

Notes. All values without a quoted error are limits (upper flux limits and their corresponding lower magnitude limits).

Table C.3. Extinctions towards the CORNISH UCHII with visible near-IR nebulae (at least in the K band), computed from the measured near-IR fluxes using four different methods.

Name	A_H	A_K	$A_V(H)$	$A_V(K)$	$A_V(J-H)$	$A_V(H-K)$
G010.3009−00.1477	4.16	4.03 ± 0.44	36.71	35.49 ± 3.85	19.60	9.37
G010.6297−00.3380	1.24	1.9 ± 0.46	10.97	16.74 ± 4.08	19.74	−1.89
G010.9584+00.0221	3.59	3.89 ± 0.52	31.69	34.31 ± 4.6	21.31	3.17
G011.1104−00.3985	3.11	1.36 ± 0.07	27.45	11.96 ± 0.65	15.63	32.41
G011.9368−00.6158	4.72	4.29 ± 0.31	41.57	37.84 ± 2.77	17.57	13.43
...						

Notes. A_λ/A_V ratios were computed following Cardelli et al. (1989): those are ~0.28, 0.18, and 0.11, for UKIDSS $\lambda = J, H, K$, respectively, assuming $R_V = 3.1$. All values without a quoted error are lower limits.

Appendix D: Distances and physical properties

Table D.1. Heliocentric and galactocentric distances and physical properties of the 239 CORNISH UCHIIs.

Name	d_h (kpc)	d_G (kpc)	D (pc)	$\log N_i$ (s^{-1})	$\log EM$ (cm^{-6}pc)	$\log n_e$ (cm^{-3})	$\log T_b$ (K)	$\tau_{5\ GHz,10^4 K}$	$\log L_{bol}$ ($L_\odot$)
G010.3009−00.1477	2.4 (1)	6.15	0.061	47.5	7.03	4.2	3.02 ± 0.04	0.11	–
G010.3204−00.2328	12.6 (2)	4.5	0.18	47.7	6.19	3.54	2.18 ± 0.07	0.015	4.51
G010.3204−00.2586	3.5 (1)	5.1	< 0.028	46.3	6.51	4.12	2.5 ± 0.06	0.032	4.18
G010.4724+00.0275	11.0 (1)	3.06	0.091	47.6	6.58	3.84	2.57 ± 0.05	0.038	5.84
G010.4736+00.0274	11.0 (1)	3.06	0.075	47.3	6.37	3.76	2.36 ± 0.05	0.023	–
...									

References. (1) Urquhart et al. (2013); (2) Cesaroni et al. (2015).

Notes. Upper limits on the physical sizes are given for unresolved sources of known distance.