11-2022

Deep Near-Infrared Survey Towards the W40 and Serpens South Region in the Aquila Rift: A Comprehensive Catalogue of Young Stellar Objects

Min Long
Boise State University

Publication Information

Sun, Jia; Gutermuth, Robert A.; Wang, Hongchi; Zhang, Miaomiao; Zhang, Shuinai; Ma, Yuehui; . . . and Long, Min. (2022). "Deep Near-Infrared Survey Towards the W40 and Serpens South Region in the Aquila Rift: A Comprehensive Catalogue of Young Stellar Objects". *Monthly Notices of the Royal Astronomical Society, 516*(4), 5244-5257. https://doi.org/10.1093/mnras/stac2191

Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY
MNRAS **516,** 5244–5257 (2022) https://doi.org/10.1093/mnras/stac2191

Deep near-infrared survey towards the W40 and Serpens South region in the Aquila Rift: A comprehensive catalogue of young stellar objects

Jia Sun (孙佳)[1,2★] Robert A. Gutermuth[3] Hongchi Wang[1] Miaomiao Zhang[1] Shuinai Zhang (张水乃)[1,4] Yuehui Ma[1] Xinyu Du[1] and Min Long[5]

[1]*Purple Mountain Observatory, Chinese Academy of Sciences, No. 10 Yuanhua Rd, Qixia District, Nanjing 210033, China*
[2]*University of Chinese Academy of Sciences, No. 19(A) Yuquan Rd, Shijingshan District, Beijing 100049, China*
[3]*Department of Astronomy, University of Massachusetts, Amherst, MA 01003, USA*
[4]*Key Laboratory of Dark Matter and Space Astronomy, Chinese Academy of Sciences, Nanjing 210033, China*
[5]*Department of Computer Science, Boise State University, Boise, ID 83725, USA*

Accepted 2022 July 17. Received 2022 July 17; in original form 2022 March 10

ABSTRACT

Active star-forming regions are excellent laboratories for studying the origins and evolution of young stellar object (YSO) clustering. The W40–Serpens South region is such a region, and we compile a large near- and mid-infrared catalogue of point sources in it, based on deep near-infrared observations of Canada-France-Hawaii Telescope (CFHT) in combination with Two Micron All Sky Survey (2MASS), UKIRT Infrared Deep Sky Survey (UKIDSS), and *Spitzer* catalogues. From this catalogue, we identify 832 YSOs, and classify 15, 135, 647, and 35 of them to be deeply embedded sources, Class I YSOs, Class II YSOs, and transition disc sources, respectively. In general, these YSOs are well correlated with the filamentary structures of molecular clouds, especially the deeply embedded sources and the Class I YSOs. The W40 central region is dominated by Class II YSOs, but in the Serpens South region, half of the YSOs are Class I. We further generate a minimum spanning tree (MST) for all the YSOs. Around the W40 cluster, there are eight prominent MST branches that may trace the vestigial molecular gas filaments that once fed gas to the central natal gas clump. Of the eight, only two now include detectable filamentary gas in *Herschel* data and corresponding Class I YSOs, while the other six are populated exclusively with Class II YSOs. Four MST branches overlap with the Serpens South main filament, and where they intersect, molecular gas 'hubs' and more Class I YSOs are found. Our results imply a mixture of YSO distributions composed of both primordial and somewhat evolved YSOs in this star-forming region.

Key words: stars: formation – stars: pre-main sequence – stars: protostars – infrared: stars.

1 INTRODUCTION

Stars are the fundamental building blocks of galaxies. They originate from the collapse of molecular clouds, the coldest and densest parts of the interstellar medium (ISM), and form mainly in overdense groupings and clusters within these clouds. The forming stars can provide significant feedback to the surroundings and ultimately affect the formation and evolution of subsequent generations of stars. The deaths of high-mass stars in supernovae can have profound effects on nearby ISM conditions. Thus, stellar evolution from birth to death influences the evolution of multiple objects: stellar systems, molecular clouds, and host galaxies. An understanding of current star formation is critical to our growing knowledge of galaxy formation and evolution, especially in those regions containing young stellar objects (YSOs) at different evolutionary stages.

YSOs usually reside in dusty molecular cloud cores, which make them dim in the optical band or even undetectable. Nevertheless, they are visible in the near-, mid-, and far-infrared (NIR, MIR, and FIR) bands, not only because of their photospheric blackbody spectrum but also because of the heavy dust emission in their circumstellar material. Therefore, at different evolutionary stages, YSOs have distinct infrared spectral energy signatures (Lada 1987; André 2002). Meanwhile, they also emit occasionally in the radio band owing to free–free emission from ionized plasma (e.g. Shang et al. 2004), in the UV band owing to accretion shocks or strong stellar winds (e.g. Calvet et al. 2004), and in the X-ray band owing to flares (Preibisch & Zinnecker 1999; Tsuboi et al. 2001; Winston et al. 2018).

During the cloud evolution and star formation, there are a few observable stages of YSOs after the starless core stage, as follows. (1) Class 0 sources are deeply embedded protostellar condensations that can be most effectively identified in the submillimetre, millimetre, and FIR bands owing to the dust blackbody emission (Hurt & Barsony 1996; Testi & Sargent 1998; Bontemps et al. 2010). (2) The canonical category of Class I YSOs, including flat-spectrum objects, describes protostars that have accreted enough mass and started to form circumstellar discs, but still possess infalling dusty envelopes (Caratti o Garatti et al. 2017; Greene et al. 1994). (3) For a Class II YSO in the 'discy' pre-main-sequence star stage, the envelope is largely cleared and an optically thick protoplanetary disc remains behind (Fang et al. 2009; Liu et al. 2019). (4) Class III YSOs have started hydrogen nuclear fusion,

★ E-mail: sun.jiaaaa@gmail.com

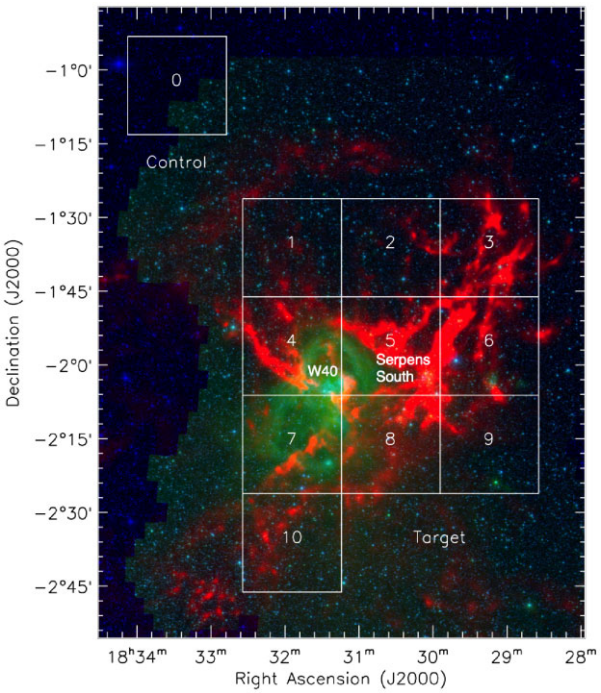

Figure 1. Composite IR image of the W40 and Serpens South regions. The white boxes denote the 'Control' (no. 0) and 'Target' (nos 1–10) fields in WIRCam. The underlying image is composed of 2MASS K_s (blue), *Spitzer* 4.5-μm (green) and *Herschel* hydrogen column density (red) data.

and their discs have only debris or some planets left (Su et al. 2006). Finally, with increasing temperature, the YSOs enter the main sequence.

Sensitive large-scale sky surveys at infrared wavelengths can provide a powerful tool with which to generate comprehensive and high-confidence catalogues of YSOs, especially in nearby (within 1 kpc) star-forming regions. Such surveys include 'From Molecular Cores to Planet' (c2d; Evans et al. 2009) and the 'Gould Belt Survey' (GB Survey; Dunham et al. 2013) conducted by the *Spitzer* Space Telescope, and the 'Two Micron All Sky Survey' (2MASS) (Skrutskie et al. 2006). Based on these surveys, thousands of YSOs have been identified and classified in many nearby star-forming regions (e.g. Gutermuth et al. 2009; Dunham et al. 2015).

This work studies two associated star-forming regions, namely W40 and Serpens South (Fig. 1), both of which belong to a larger complex of molecular clouds collectively known as the Aquila Rift. The W40 region was first observed in the NIR in 1978 (Zeilik & Lada 1978), but it was another 30 years before the Serpens South region was discovered and named (Gutermuth et al. 2008a). The W40 region is famous for its hourglass-shaped H II region illuminated by a few massive OB stars (Shuping et al. 2012), and its cloud structure, magnetic field, and star-formation activities, such as outflows and YSO populations, have been studied in detail through observations from NIR to millimetre bands (Gutermuth et al. 2008a; Nakamura et al. 2011; Sugitani et al. 2011; Zhang et al. 2015; Winston et al. 2018; Shimoikura et al. 2019). The Serpens South region is embedded in a dark cloud filament and contains a large number of extremely young objects, which have also received a lot of attention. The two regions are separated by only ∼20 arcmin (or ∼3 pc) from each other, and both of them are at a distance of ∼436 pc (Ortiz-León et al. 2017). As a consequence, the W40–Serpens South region can be considered as one complex with YSOs at different states, presenting an excellent laboratory in which study feedback influences.

In this region, Bontemps et al. (2010) used the *Herschel* GB Survey data to identify several tens of Class 0/I sources in very early evolutionary stages. Dunham et al. (2015) used the c2d and *Spitzer* GB survey data to identify YSOs, with NIR data based on 2MASS, which is not deep enough for those objects embedded in thick molecular filaments. In this work, we analyse NIR data provided by the Canada–France–Hawaii Telescope (CFHT) for the W40–Serpens South region, complemented by catalogues such as 2MASS and UKIRT Infrared Deep Sky Survey (UKIDSS) . The deeper NIR observations can penetrate better towards the filaments around the W40 and Serpens South cores, and allow us to identify more embedded YSOs and depict the detailed structures of young clusters. Although we do not include a specific filament analysis here, we want briefly to check the spatial association between YSOs and filaments in the overall W40–Serpens South region. Such YSO structures can also be compared with a more delicate structure of molecular clouds, for example as detected by Atacama Large Millimeter Array (ALMA) (Plunkett et al. 2018).

The paper is organized as follows. The used observations and data are introduced in Section 2. Methods of image-processing and point-source catalogue production are explained in Section 3. We present the identification of YSOs and discuss their relationships with molecular clouds and the entire star-forming region in Section 4, and give a summary in Section 5.

2 OBSERVATIONS AND DATA

2.1 CFHT observations

We used the Wide-Field Infrared Camera (WIRCam) on the CFHT to take NIR images of the W40 and Serpens South regions in the J, H, and K_s bands. WIRCam has a field of view of 20 arcmin × 20 arcmin, and a resolution of 0.3 arcsec per pixel. The 'Target' field has a size of ∼60 arcmin × 80 arcmin, or ∼ 7.6 pc × 10.1 pc, and covers most of the main molecular complex. Fig. 1 shows the entire region, consisting of 10 subfields of the 'Target', labelled from no. 1 to no. 10, during the actual observations, and one 'Control' field (labelled no. 0) of 20 arcmin × 20 arcmin as reference. The 'Control' field was selected based on the distance to the 'Target', and has a very low density of molecular clouds, so there are no young stars enclosed. 2MASS (blue), *Spitzer* 4.5-μm (green) and *Herschel* hydrogen column density (red) data are superimposed for illustrations. It can be seen that the 2MASS data cover the entire area; *Spitzer* data cover only the 'Target' and ∼1/4 of the 'Control' field; and *Herschel* data cover only the 'Target' field.

The CFHT observations were carried out in the J, H, and K_s bands between 2012 July 27 and 29. Seeing was good, and the average full width at half-maximum (FWHM) of the subfields varied from 0.45 to 0.96 arcsec. The single-image exposure times of J, H, and K_s were 30, 10.3, and 15 s, respectively. We selected the dithering mode so that bad pixels and cosmic rays could be effectively removed. Each subfield was covered by 4–24 exposures, and about 70 per cent of the final integrated images contain more than 16 exposures. The total exposure times reached 240, 123.6, and 120 s, on average, in the J, H, and K_s bands, respectively.

The image of each individual exposure was processed by CFHT's IDL Interpreter of the WIRCam Images, namely the `'I'wii`

pipeline.[1] In the pipeline, the following processes are performed automatically: chipbias and dark subtraction, flat-fielding, non-linearity correction, cross-talk removal, and sky subtraction. In addition, pixels that are bad or saturated are marked.

2.2 2MASS and UKIDSS data

Bright stars in our long-exposure-time images could be saturated and need to be resolved in our CFHT observations. This is also because of the wide field and high sensitivity of the WIRCam. For this reason, we combined 2MASS and UKIDSS data for the calibration of the bright stars.

First, we used the 2MASS Point Source Catalog (PSC) (Skrutskie et al. 2006). 2MASS data corresponding to the 'Target' and the 'Control' fields are available from the VizieR II/246 catalogue.[2] We used stars with a high-quality flag (`qflag` = 'AAA'), which means that the source has a signal-to-noise (S/N) ratio greater than 10 and an error less than 0.1. We also used a low-contamination flag (`cflag` = '000'), which means that the source is unaffected by any known artefacts, or simply has no detection in some bands.

Second, we introduced UKIDSS Galactic Plane Survey (GPS) data (Lucas et al. 2008) for calibration and used them as the medium-bright part of our catalogue. UKIDSS has integration times of 80, 80, and 40 s for the *J*, *H*, and *K* bands respectively. The UKIDSS data we used are available from the WFCAM Science UKIDSS Archive release DR11.[3] Quality flags were selected as follows: `priOrSec` = 0 or `frameSetID` for selecting out a seamless, best catalogue; `ppErrbits` < 256 to exclude all *K*-band sources with any error quality condition; $-2 \leq$ `Class` ≤ -1 to exclude noises, saturated sources, and galaxies. The last flag is strict and cuts off ∼32 per cent of the total number of UKIDSS stars. Among them, sources marked as 'galaxy' are mostly pairs of stars (Lucas et al. 2008); therefore we used CFHT data instead.

2.3 *Spitzer* data

MIR data is a good supplement to NIR data for both contamination removal and YSO identification. Aquila data from the *Spitzer* GB Legacy Survey were therefore included in this work. This section is a summary of work performed by Gutermuth et al. (in preparation).

The 'Target' field is fully covered, and the 'Control' field is 1/4 covered by the *Spitzer* survey. The *Spitzer* photometry is derived from a draft catalogue of the *Spitzer* Extended Solar Neighborhood Archive (SESNA; Gutermuth et al., in preparation), and we summarize below the creation process via the Cluster Grinder IDL software collection for *Spitzer* data treatment (Gutermuth et al. 2009; Winston et al. 2018; Li et al. 2019; Pokhrel et al. 2020).

The IRAC imaging has four bands, at 3.6, 4.5, 5.8, 8.0 μm, and the MIPS imaging is at 24 μm (Gutermuth et al. 2008a). Images are treated for bright-source artefacts using custom IDL scripts. Cosmic-ray hits are identified and masked out, and all images are merged into mosaics using the World Coordinate System (WCS) mosaic module, with 0.87 arcsec per pixel resolution for IRAC and 1.8 arcsec per pixel for the MIPS.

Sources are identified using the `PhotVis` module (Gutermuth et al. 2008b), and aperture photometry of those sources is performed using `aper.pro` from the IDL Astronomy User's Library (Landsman 1993). The adopted aperture radius for IRAC is 2.4 arcsec, and the background is estimated from an annulus field with radii of 2.4 and 7.2 arcsec. Corresponding MIPS aperture and inner and outer annulus radii are 6.35, 7.62, and 17.78 arcsec, respectively. Instrumental fluxes are converted to Vega-standard magnitudes using magnitude zero points (e.g. magnitude for 1 DN/s net flux in aperture) of 20.21, 19.45, 17.25, and 17.64 mag for IRAC channels 1 through 4, and 15.35 mag for MIPS 24 μm. Magnitude limits for fields averaging 90 per cent differential completeness are 14.5, 14.2, 13.2, 12.0, and 8.4 mag at 3.6, 4.5, 5.8, 8.0 and 24 μm, respectively.

2.4 *Herschel* data

We also included the publicly available *Herschel* hydrogen column density map from the *Herschel* GB Survey for a higher resolution of the dust extinction estimation and better demonstration of the Molecular Clouds, as underlying images in Figs 7 and 8 show. The map is derived from greybody fitting of the 160-, 250-, 350-, and 500-μm FIR images, and resampled and smoothed to match the latter image's 36-arcsec FWHM beam size (Könyves et al. 2015).

3 IMAGE REDUCTION, APERTURE PHOTOMETRY AND THE LARGE INFRARED CATALOGUE

After 352, 528, and 384 WIRCam small images in the *J*, *H* and K_s bands are output by the `'I'wii` pipeline, many corrections are performed before the final catalogue is generated. This section describes the reduction and analysis of the NIR data only.

3.1 Follow-up reduction of images

Original images have many void pixels that affect both faint and bright stars, and combined images can effectively remove them and also give a larger S/N ratio for star detection. Meanwhile, the resampling process can correct the distortions of the original images. To co-add individual images together, we chose a suite of astronomical pipeline software tools, AstrOmatic (Bertin 2012), which includes the functions SExtractor (Source Extractor), scamp (Software for Calibrating AstroMetry and Photometry), and SWarp (Source Warp) (Bertin & Arnouts 2010; Bertin 2010a, b). These tools were developed at the Institut d'Astrophysique de Paris (IAP) in France and are designed to be run in batch mode on large quantities of data.

First, SExtractor extracts a point-source catalogue from images and outputs a rough catalogue. Second, scamp compares it with the standard 2MASS catalogue and calculates an astrometric solution. For our data, scamp corrects the catalogue distortion with polynomial functions under the TPV[4] WCS in FITS (Calabretta & Greisen 2002). With these equation coefficients, scamp outputs the corresponding external header to each image. Eventually, SWarp performs image resampling and co-addition using polynomial distortion correction terms in external headers. We observe that this suite of tools performs better than the Simple Imaging and Mosaicking Pipeline (simple) (Wang 2010). In the simple, astrometry is done with

[1] www.cfht.hawaii.edu/Instruments/Imaging/WIRCam/IiwiVersion1Doc.html
[2] http://vizier.u-strasbg.fr/viz-bin/VizieR?-source=II/246&-to = 3
[3] http://wsa.roe.ac.uk:8080/wsa/SQL_form.jsp
[4] https://fits.gsfc.nasa.gov/registry/tpvwcs/tpv.html

more resampling, during which extra spatially correlated noise is introduced (Zhang & Kainulainen 2019) and the FWHM of sources increases by ∼10 per cent.

Detrended images output by the CFHT `'I'iwi` pipeline are input to the above pipeline. For SExtractor and SCAMP, we used all the default settings. We changed three parameters instead of using the default in SWarp: combine with 'AVERAGE', resampled pixel scale to be '0.15 arcsec', and background 'NOT' subtracted. For our images, 'AVERAGE' is better than the default 'MEDIAN', although the latter can effectively remove cosmic rays. Although the FWHMs of stars increase on the final images, the sources are shaped well for photometry, as 'AVERAGE' can better preserve the shape of a source, especially for the peak. A 0.15-arcsec pixel scale gives a better image quality where pixels are not integrally overlapped and defines a better location for a star. The background is not subtracted because it contains nebular emission. After SWarp, the 'Target' field is stitched to 10 pieces of images instead of one, which is required by the astigmatism correction in Section 3.4. Therefore, the final catalogue is also stitched from 10 separated catalogues. After distortion correction, the typical astrometry difference with 2MASS for a star is 0.06 arcsec. In the final derived maps, the mean FWHM varies from 0.65∼1.0 arcsec, as images are more stretched after being combined.

3.2 Aperture photometry

On the final stacked images, the point spread functions (PSFs) are not uniform, and their sizes in each field are not the same. So we chose to perform aperture photometry instead of PSF photometry.

We used PhotVis (Gutermuth et al. 2008b) to identify stars and perform aperture photometry. The radius of the aperture, inner and outer radii of the background annulus were chosen to be 0.75, 1.5, and 3 arcsec, which are relatively small to avoid nearby starlight pollution when the typical nearest-neighbour distance is ∼1.8 arcsec for a star. We identified sources with statistics better than a S/N ratio of 6. The numbers of sources from the initial photometry were 422 442 for J, 543 212 for H, and 813 307 for K_s. The number of detections in the K_s band is nearly double that in the J band, mainly because the longer wavelength emitted by blocked background Milky Way stars can better penetrate dark clouds. We used TOPCAT to perform the cross-match, between the K_s and H bands first, and then between the K_s/H and J bands, because stars in the first two bands have lower FWHM values and a larger number of detections. In total 870 471 sources were identified, and 369 420 sources appeared in all three bands. The position tolerance of cross-matches of the two catalogues was 1 arcsec.

There are two methods to estimate detection completeness. The first is to successively add synthetic stars with known magnitude and count their detection ratio, as described in Gutermuth et al. (2005). The second is to fit a linear relationship between the magnitude histogram of log(N) and magnitude, and then assume that the decline at faint magnitudes is entirely due to limited sensitivity (Lucas et al. 2008). The latter method is easier in our case, as synthetic stars change in those corrections from the first photometry to the final magnitude catalogue. The typical 90 per cent completeness limits in uncrowded fields are $J = 19.8$, $H = 18.6$, and $K = 18.0$, with uncertainties of ∼0.2 mag. The non-uniform number of exposures in the various subfields also leads to different limiting magnitudes within the subfields and to a non-homogeneous photometric catalogue. We checked that the mixed magnitude for the three bands is ∼0.12 mag less than the theoretical magnitude, which is consistent with our derived uncertainties.

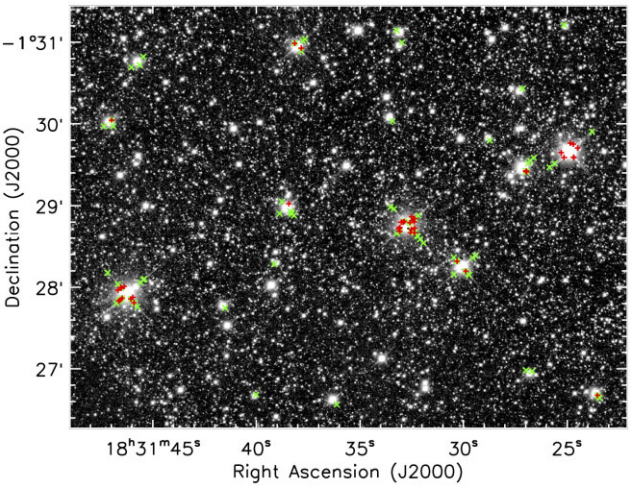

Figure 2. Effects of bright stars on faint stars owing to telescope spikes and long exposure on J-band images. Green cross signs are faint stars that are marked as 'affected by a spike'. Red plus signs are stars that are marked as 'affected by bright field'.

Table 1. Fitting parameters for the spike length and radius of affection.

Band	a_0	a_1	b_0	b_1
J	3.91	−0.206	3.30	−0.182
H	3.74	−0.209	3.19	−0.187
K_s	3.51	−0.167	2.84	−0.145

3.3 Effects of telescope spike and long time exposure

Bright stars can affect the identification of faint stars owing to the effects of telescope spikes and long exposure. Modern reflecting telescopes usually need spider vanes to support their second mirror. When incident light reaches the diffraction limit, telescope spikes can appear near these bright stars. In our CFHT images, bright stars have four main spikes. In addition, after a long exposure, the bright stars can look bigger and more square in shape, and thus affect their surrounding field. Fig. 2 shows both effects on J-band images. We examined bright stars with J/H/K mag < 13.5 and marked affected faint stars.

To quantify these effects and mark faint stars, we introduced two values: the spike length $L_{\rm spike}$ and the affected radius $R_{\rm affect}$. First, we chose bright stars from the 2MASS catalogue, and their magnitudes are evenly distributed. Second, we roughly recorded the pixel length of their spikes and their affected radius. By assuming an exponential relationship, we fitted both $\log(L_{\rm spike})$ and $\log(R_{\rm affect})$ as a linear function of magnitude M. The resultant formulas are as follows:

$$L_{\rm spike}({\rm pixels}) = 10^{a_0+a_1 M}, \quad (1)$$

$$R_{\rm affect}({\rm pixels}) = 10^{b_0+b_1 M}. \quad (2)$$

The fitting parameters are shown in Table 1.

These functions were applied to our field to mark any star that is located near the bright sources. There are about 12 000 sources marked as 'affected by a spike' and 4700 sources marked as 'affected by a bright field'. Two of our final YSOs are marked as 'affected', so we re-classified them with the 2MASS data instead. Coincidentally, their evolutionary stages remain unchanged.

3.4 Astigmatism correction

We cross-matched CFHT with UKIDSS to process corrections in the CFHT catalogue. The cross-match was carried out in each band separately, and there are about 289 000, 451 000, and 458 000 stars cross-matched in the J, H, and K_s bands respectively. When examining CFHT and UKIDSS magnitude-difference diagrams, we found that the points at the brighter end approached two distinct values rather than one value and thus formed two branches in the diagram. The second branch is wider and fainter than the first one. The top middle panel of Fig. 3 shows this effect.

The right-hand panel of Fig. 3 shows that relatively faint stars are distributed mostly on the edges or corners, where the astigmatism is worse than around the centre of the image. Astigmatism occurs because the telescope has slightly different refractive powers on different meridians. After the parallel light enters the optic mirrors, it cannot be concentrated at an accurate point but at a slightly deflected point. Distorted stars are stretched with images by SWarp during the previous step discussed in Section 3.1, and they can occupy a bigger area than the undistorted stars. Under the same aperture, these bigger blobs lose some starlight counts.

As fields 1–10 of the 'Target' repeat in the same observed pattern as the 'Control' field, their astigmatism patterns are also the same. With no extinction from clouds, the 'Control' field has enough stars that are randomly distributed. Therefore the 'Control' field can be used to draw a compensation map that can be applied to both itself and fields 1–10 of the 'Target' field. For each point of the compensation map, we found 40 nearest-neighbour stars in the 'Control' field that are detected by both UKIDSS and CFHT. The mean magnitude difference of them was calculated and taken to be the value of magnitude compensation at the point. The standard deviation of these differences was also calculated and placed on a 'compensation deviation' map. Finally, each star in field 0–10 has a compensated magnitude added, and magnitude uncertainties also have an error propagation in the same way as for the compensation deviation map. The J-band aperture 'compensation' map that we made for the astigmatism correction is shown in the left-hand panel of Fig. 3. After this correction, the bright end of the magnitude difference between CFHT and UKIDSS converges, as the bottom-middle panel of Fig. 3 shows.

3.5 Colour-term fitting correction

Each telescope has its own unique filter transmission curve, and therefore we need a colour term to correlate observations of the same source through CFHT and UKIDSS. Furthermore, UKIDSS uses the K band, while CFHT uses the K_s band, which is ∼25 per cent narrower. Given that UKIDSS is a large-scale NIR survey covering a range of areas and depths, it is easier for us to shift the CFHT catalogue to the UKIDSS colour system. The 2MASS catalogue is not shifted to the UKIDSS colour, because only bright stars from 2MASS are combined. They are usually saturated in UKIDSS, and lack detections to finish the fittings.

By fitting a linear term with the median value of each segment, the transformation equations are as follows. First, for stars detected in at least two bands of the CFHT, we use the equations:

$$J_{\rm CFHT} - J_{\rm UKIDSS} = a_J + b_J(J_{\rm CFHT} - H_{\rm CFHT}), \quad (3)$$

$$H_{\rm CFHT} - H_{\rm UKIDSS} = a_H + b_H(H_{\rm CFHT} - K_{\rm sCFHT}), \quad (4)$$

$$K_{\rm sCFHT} - K_{\rm UKIDSS} = a_K + b_K(H_{\rm CFHT} - K_{\rm sCFHT}). \quad (5)$$

In the first case, we used CFHT magnitudes instead of UKDISS on the right side of the equals sign, because stars have to be detected in two bands to use these equations. The number of qualifying stars in UKIDSS is about half of that in CFHT (174 000 versus 376 000 for both J and H detected; 248 000 versus 510 000 for both H and K). The relationships between $J_{\rm CFHT} - H_{\rm CFHT}$ and $J_{\rm UKDISS} - H_{\rm UKIDSS}$, $H_{\rm CFHT} - K_{\rm CFHT}$ and $H_{\rm UKIDSS} - K_{\rm UKIDSS}$ were checked and found to be linear, and therefore these equations are compliant with those using UKIDSS on the right side of the equals sign.

Second, for stars detected only in one band, or stars lacking a key band in the first case, we used $A_{\rm K}$ as the colour term for fitting instead:

$$J_{\rm CFHT} - J_{\rm UKIDSS} = a_J + b_J A_K, \quad (6)$$

$$H_{\rm CFHT} - H_{\rm UKIDSS} = a_H + b_H A_K, \quad (7)$$

$$K_{\rm sCFHT} - K_{\rm UKIDSS} = a_K + b_K A_K. \quad (8)$$

The second case is for the stars that are significantly different from UKIDSS, lack detections, and are located in a high-extinction place. Their magnitude difference will be reduced through fitting with the $A_{\rm K}$ value. $A_{\rm K}$ data are derived from the *Herschel* column density map and converted using $N_{\rm H}/A_{\rm V} = 1.8 \times 10^{21}{\rm cm}^{-2}/{\rm mag}$ (Predehl & Schmitt 1995) and $A_{\rm K}/A_{\rm V} = 0.112$ (Rieke & Lebofsky 1985). Here $A_{\rm K}$ is roughly a total extinction by all the molecular clouds along the line of sight, and is used as $A_{\rm K}$ TOTAL in Section 4.2.2. The relationships between $J_{\rm CFHT} - H_{\rm CFHT}$, $H_{\rm CFHT} - K_{\rm CFHT}$ and A_K are also checked to be approaching linear, which makes the second case compliant with the first.

After the colour term is applied, a deeply reddened star has less colour difference between the CFHT and UKIDSS catalogues. Fig. 4 shows the magnitude difference of the K_s band before and after colour-term fitting. Stars that have low magnitude errors (< 0.08), a nearest-neighbour distance > 10 pixels, and a de-reddened $K_s < 16$ mag were used to perform the fitting, and the results are applied to all other stars, as the bottom two panels of Fig. 4 show. For those stars with $H - K_s > 2.7$ mag, K_s can receive a colour correction of about 0.05 mag; for those stars with $A_{\rm K} > 4$ mag, K_s can receive a colour correction of over 0.2 mag.

3.6 Zero-magnitude correction

The observations were carried out over three days, and therefore the seeing turned out to be a variable of time. Fig. 5 shows the zero-magnitude correction we adopted in our work. The top panel shows the FWHM as a function of field number in those three days. Fields 1, 2, 3, 4 were observed on 2012 July 27; fields 8, 9, 0 on 2012 July 28; and fields 5, 6, 7, 10 on 2012 July 29. As a result, zero magnitude has different levels among the 10 subfields inside the 'Target' field. Therefore, we calculated a zero magnitude for each field so that the CFHT stars can be as close to the UKIDSS as possible. The zero magnitudes were calculated by fitting. First, we chose bright stars inside a window of 2 mag that started from 14.5, 14, and 13.5 mag for J, H, and K_s, respectively. Second, their UKIDSS – CFHT magnitudes were fitted with a Gaussian distribution. Third, the peak and the error of fitting were taken as the zero magnitude and its uncertainty in this field. The bottom panel of Fig. 5 shows the zero magnitudes for each band, which will be added back to our data accordingly.

Uncertainty during fitting will also be added to all stars' uncertainties through error propagation. The magnitude difference between two catalogues should be consistent with their uncertainties. 2MASS uncertainty is normal, but uncertainty of UKIDSS

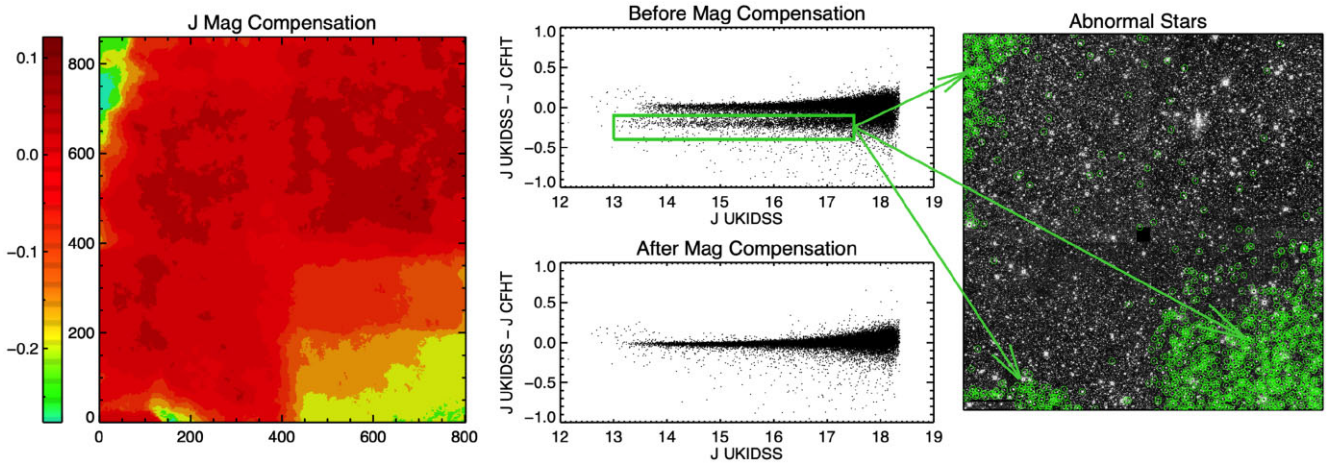

Figure 3. *Left-hand panel*: *J*-band magnitude 'aperture compensation' map. Most parts, mainly located in the centre, are nearly zero. Those located on the edges have a bad performance and need correction. *Top-middle panel*: *J*-band magnitude difference between the UKIDSS and CFHT catalogues for the 'Control' field. The lower branch at the bright end is marked in the green box. *Bottom-middle panel*: Same stars as in the top right-hand panel after the compensation map is applied. *Right-hand panel*: Stars in the green box of the deviated branch are placed on the 'Control' field map with green circles. The compensation maps for *H* and K_s are almost same as for *J*. For *J*, the compensation magnitude varies from –0.23 to 0.17 mag, with a mean value of ∼0.038 mag; for *H*, the compensation magnitude varies from –0.27 to 0.058 mag, with a mean value of ∼–0.057 mag; for K_s, the compensation magnitude varies from –0.18 to 0.094 mag, with a mean value of ∼0.025 mag.

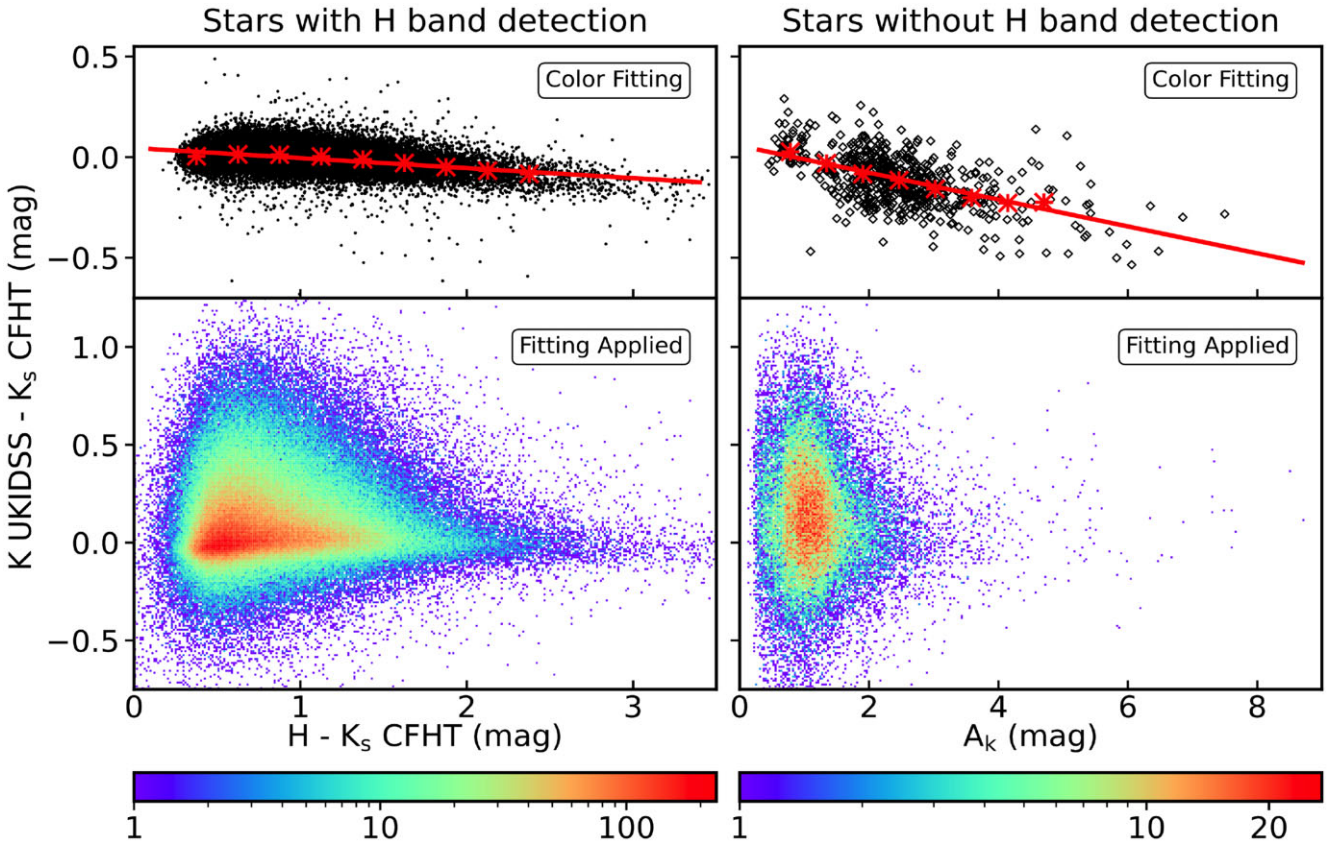

Figure 4. *Left-hand panel*: Colour-term fitting of the K_s-band magnitude with $H - K_s$ (top), and after it is applied to the CFHT catalogue (bottom). *Right-hand panel*: Colour-term fitting with A_K (top), and after it is applied to all qualified stars (bottom). The median values of each magnitude bin, namely the red asterisks, are the actual fitted points. The colour bars for the bottom panels are counts of a 2D histogram divided by 300 × 200 pixel grids. Considering that the colour correction is aimed mainly at stars with larger A_K, the red blob in the bottom middle is allowed to be higher than zero here. These differences with UKIDSS are better corrected in the section on zero-magnitude correction.

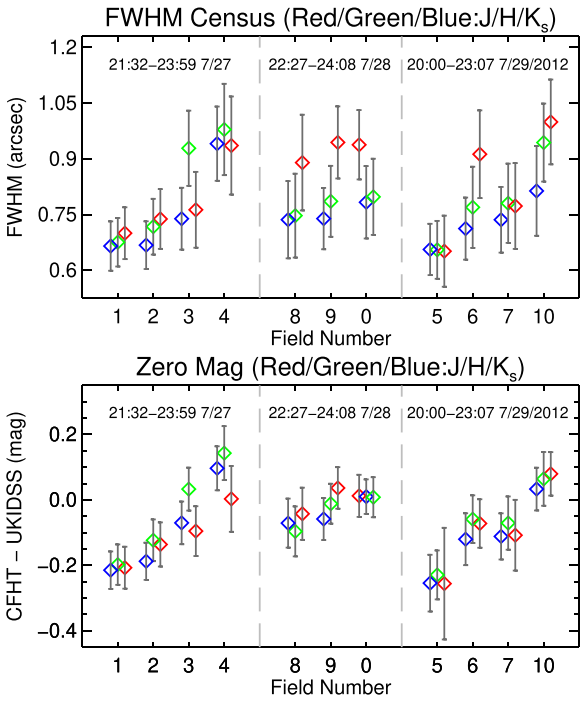

Figure 5. *Top panel*: FWHM level and variation census of each band in the subfield images. *Bottom panel*: Zero points of magnitude and magnitude difference variation level of each band in the subfield catalogue. We finally added zero magnitude accordingly.

is underestimated because it contains only a synthetic error term, as Lucas et al. (2008) issued. . Therefore an extra constant term, $C_{\rm Unc.UKIDSS}$, is required for expanding the UKIDSS uncertainty. We correct it by calculating the compromised magnitude difference: $(Mag_{\rm 2MASS} - Mag_{\rm UKIDSS})(Unc^2_{\rm 2MASS} + Unc^2_{\rm UKIDSS})^{-1/2}$. For bright stars, the standard deviation should be mostly less than 1. In this calculation, only those stars that are isolated, round-shaped and not very sharp in shape were used. We calculated $C_{\rm Unc.UKIDSS}$ to be 0.0265, 0.0347, and 0.0337 in the J, H, and K_s bands. Then this constant term is added back, and the bigger $(Unc^2_{\rm UKIDSS} + C^2_{\rm Unc.UKIDSS})^{1/2}$ replaces the smaller $Unc_{\rm UKIDSS}$. After this step, the two catalogues, 2MASS and UKIDSS, are compatible within their uncertainties. The CFHT uncertainty had already increased gradually in previous processes, such as astigmatism correction, colour-term fitting, and zero-magnitude adding. It is compatible with 2MASS and revised UKIDSS uncertainties.

3.7 The large near-mid-infrared catalog

After the above steps, a large NIR catalogue is generated. Selection rules for NIR photometry in multiple catalogues are as follows: 1, if a star has $J < 12.5$, or $H < 12.5$, or $K_s < 12$ then its data come from 2MASS; 2, if a star has any band of J, H, K_s missing then its data come from CFHT. Then We cross-matched it with *Spitzer* MIR data and obtained a large near-mid-infrared catalogue. A census of this large near-mid-infrared catalogue is shown in Table 2, and the point-source catalogue has a format as in Table 3. It contains the following information: YSO ID; coordinates; magnitude and magnitude uncertainty of $J/H/K_s$; IRAC Ch1/Ch2/Ch3/Ch4; MIPS 24 μm; whether this source is detected in the 2MASS/UKIDSS/CFHT surveys; front extinction (see Section 4.1); and total extinction (see Section 3.5) .

4 IDENTIFICATION AND DISTRIBUTION OF YSOS

4.1 Classification of YSOs

Based on the large near-mid-infrared catalogue, we can identify IR-excess-bearing dusty YSOs in the 'Target' field and generate a comprehensive YSO catalogue. Fig. 6 shows our classification diagrams and processing steps, using the method described in Gutermuth et al. (2009) (G09 hereafter).

First, we removed contamination from non-stellar objects, which may have abnormally high fluxes in some specific narrow band channels owing to PAH line features, AGN torus emission, or shock-heated gas emission, and are confusing for YSO classifications. Figs 6(a)–(d) show how the distinct colour gradient lines are used to separate stars and contamination sources: the exact expressions of these lines are detailed in G09's appendix A. With reference to the diagrams, we excluded 28 unresolved star-forming galaxies, 46 AGNs, 12 shock-emission knots, and 49 PAH aperture-contaminated sources in the 'Target' field. Fig. 7(a) shows the distribution of these contamination sources. The shock-emission knots tend to appear in places where the molecular clouds have higher column densities, while the other three have roughly random distributions.

Second, we identified and classified Class I/II YSOs according to their infrared excess, mainly by using *Spitzer* data if they were reliably detected by all IRAC Ch1–4 bands (Figs 6e–f). However, for those that lack IRAC Ch3/4 detections, which turns out to be ∼31 per cent YSOs, we finally identified them using the deep K_s-band data (Fig. 6i), with their reddening measured with another NIR-band detection. The dereddening magnitude $A_{\rm K}$ (also called the 'FRONT' extinction in Section 4.2.2) was calculated for all those stars with good NIR detections ($\sigma < 0.1$) in Figs 6(g)–(h). The results are listed in the point-source catalogue in Section 3.7. Visually, the extinction of a star is proportional to the distance of its locus in Figs 6(g)–(h) from the dashed line in the arrow direction of $A_{\rm K} = 2$.

Third, we searched for deeply embedded sources and transition disc sources according to the *Spitzer* MIPS 24-μm data (Fig. 6 j–k). A deeply embedded source with a lower bolometric temperature or extreme reddening is dim in the MIR bands, and thus frequently lacks some vital MIR detections or has been misclassified as an AGN contaminant, but is relatively quite bright at MIPS 24 μm. These sources can be either Class 0 or Class I, but it is not possible to distinguish between them with current data. Transition disc sources have significant dust clearing within their inner discs, leading to an impression of no disc based on the MIR, but the relatively bright MIPS 24 μm reveals the presence of the disc beyond the large inferred inner gaps. These sources are added to the Class II

Table 2. Census of the final large near-mid-infrared catalogue composition.

Survey	2MASS	UKIDSS	CFHT	Ch1	Ch2	Ch3	Ch4	MIPS	H & K_s & 1Ch
Total	4247	159 980	737 683	122 704	119 859	43 083	32 547	1494	111 481

Table 3. The near-mid-infrared catalogue of identified YSOs in the 'Target' field. Two default values, –100 and 10, are used as placeholders for non-detections in magnitude and uncertainty. The full table is available online.

ID	RA (J2000)	Dec. (J2000)	J	Unc. J	H	Unc. H	K	Unc. K	NIR catalogue	3.6	Unc. 3.6	4.5	Unc. 4.5	5.8	Unc. 5.8	8.0	Unc. 8.0	24	Unc. 24	A_K FRONT	A_K TOTAL	Class
	(hh:mm:ss.sss)	(dd:mm:ss.sss)	(mag)	(mag)	(mag)	(mag)	(mag)	(mag)	(mag)	(mag)	(mag)	(mag)	(mag)	(mag)	(mag)	(mag)	(mag)	(mag)	(mag)	(mag)	(mag)	(I∗,I,II,II∗)
Deeply Embedded Sources																						
1	18:30:3.115	−1:36:32.770	−100.000	10.000	17.840	0.029	16.194	0.019	CFHT	13.653	0.019	12.388	0.011	11.842	0.024	11.521	0.060	4.500	0.005	0.000	2.223	I∗
2	18:28:47.771	−1:38:7.869	−100.000	10.000	−100.000	10.000	17.021	10.000	CFHT	13.509	0.011	11.524	0.010	11.160	0.009	10.346	0.017	3.062	0.002	−100.000	2.452	I∗
3	18:31:57.414	−2:1:57.435	17.005	0.030	16.227	0.036	16.168	0.047	UKIDSS	16.134	0.254	15.026	0.137	13.227	0.357	11.461	0.346	6.587	0.191	0.187	1.536	I∗
4	18:31:52.180	−2:1:26.148	22.972	0.973	−100.000	10.000	17.782	10.000	CFHT	13.632	0.017	11.876	0.007	11.196	0.019	10.728	0.080	3.685	0.024	−100.000	3.358	I∗
5	18:30:3.259	−2:3:26.575	−100.000	10.000	−100.000	10.000	−100.000	10.000	NONE	−100.000	10.000	−100.000	10.000	−100.000	10.000	10.414	0.040	4.400	0.040	−100.000	12.796	I∗
...																						
14	18:30:15.645	−2:7:19.774	−100.000	10.000	18.288	0.032	15.600	0.022	CFHT	13.207	0.016	12.388	0.012	12.074	0.057	11.647	0.066	4.141	0.006	3.518	5.800	I∗
15	18:29:43.972	−2:12:55.408	−100.000	10.000	−100.000	10.000	17.493	10.000	CFHT	14.203	0.025	12.612	0.015	12.256	0.042	12.279	0.079	4.003	0.002	−100.000	2.396	I∗
Class I YSOs																						
16	18:31:51.010	−1:33:54.766	19.490	0.127	17.431	0.046	15.625	0.039	UKIDSS	13.503	0.013	12.641	0.008	11.777	0.017	10.852	0.014	7.286	0.032	0.781	0.723	I
17	18:31:58.086	−1:27:53.997	19.945	0.065	18.232	0.036	16.694	0.022	CFHT	14.562	0.030	13.764	0.025	12.964	0.041	11.901	0.034	7.852	0.031	0.477	0.748	I
18	18:28:55.834	−1:43:13.634	20.755	0.069	17.533	0.015	15.204	0.019	CFHT	12.330	0.005	11.330	0.004	10.503	0.007	9.767	0.005	5.734	0.008	2.141	1.905	I
19	18:29:8.907	−1:42:35.671	−100.000	10.000	−100.000	10.000	17.357	10.000	CFHT	15.117	0.063	13.684	0.047	12.524	0.058	12.113	0.140	−100.000	10.000	−100.000	2.878	I
20	18:29:8.696	−1:42:33.115	−100.000	10.000	−100.000	10.000	−100.000	10.000	NONE	15.090	0.049	13.537	0.035	11.952	0.031	11.233	0.044	−100.000	10.000	−100.000	2.949	I
...																						
149	18:32:15.980	−2:34:43.280	18.352	0.018	15.503	0.009	13.003	0.019	CFHT	11.795	0.004	10.181	0.004	9.334	0.003	8.780	0.010	3.310	0.001	1.394	3.395	I
150	18:32:9.134	−2:27:3.158	−100.000	10.000	17.652	0.021	16.355	0.016	CFHT	14.781	0.099	13.945	0.065	−100.000	10.000	−100.000	10.000	−100.000	10.000	0.130	1.023	I
Class II YSOs																						
151	18:31:24.861	−1:46:1.704	10.697	0.023	8.870	0.053	7.756	0.026	2MASS	6.809	0.002	6.592	0.002	6.313	0.001	6.018	0.001	5.121	0.005	1.047	0.739	II
152	18:31:24.681	−1:45:20.551	12.394	0.023	10.488	0.022	9.652	0.019	2MASS	8.841	0.002	8.527	0.002	7.930	0.002	6.893	0.002	4.713	0.006	1.375	0.714	II
153	18:31:16.688	−1:44:47.220	16.455	0.028	15.130	0.035	14.475	0.035	UKIDSS	13.577	0.013	13.293	0.013	13.382	0.050	−100.000	10.000	−100.000	10.000	0.672	0.671	II
154	18:31:28.334	−1:44:43.855	17.563	0.036	16.231	0.036	15.590	0.040	UKIDSS	14.600	0.037	14.218	0.027	−100.000	10.000	−100.000	10.000	−100.000	10.000	0.698	0.677	II
155	18:31:24.678	−1:42:51.629	12.763	0.026	11.033	0.022	10.172	0.021	2MASS	9.298	0.002	9.137	0.003	8.795	0.003	8.588	0.002	8.007	0.033	1.126	0.721	II
...																						
796	18:32:20.591	−2:28:45.903	18.827	0.020	17.110	0.014	16.170	0.015	CFHT	15.074	0.061	14.424	0.038	−100.000	10.000	−100.000	10.000	−100.000	10.000	1.033	1.013	II
797	18:32:21.005	−2:28:22.727	18.549	0.020	16.638	0.012	15.700	0.013	CFHT	14.761	0.038	14.199	0.032	−100.000	10.000	−100.000	10.000	−100.000	10.000	1.342	0.920	II
Transition Disc Sources																						
798	18:32:0.791	−1:45:6.379	17.958	0.038	15.224	0.035	13.746	0.034	UKIDSS	12.846	0.007	12.697	0.008	12.464	0.025	12.406	0.060	9.739	0.191	2.150	1.336	II∗
799	18:32:1.153	−1:39:31.191	−100.000	10.000	17.681	0.017	16.401	0.016	CFHT	14.573	0.034	14.334	0.037	−100.000	10.000	12.199	0.040	9.206	0.073	1.963	0.893	II∗
800	18:31:23.138	−1:27:44.323	18.997	0.034	17.436	0.021	16.197	0.022	CFHT	14.666	0.057	14.336	0.036	−100.000	10.000	11.317	0.031	8.776	0.070	0.511	0.530	II∗
801	18:32:13.131	−1:28:27.408	19.285	0.025	17.887	0.020	16.779	0.019	CFHT	15.313	0.060	15.104	0.084	−100.000	10.000	12.355	0.056	9.538	0.090	0.373	0.573	II∗
802	18:31:6.791	−1:34:26.820	17.577	0.036	16.516	0.038	16.082	0.049	UKIDSS	15.263	0.066	15.119	0.076	−100.000	10.000	13.243	0.073	9.997	0.167	0.457	0.537	II∗
...																						
831	18:31:21.688	−2:30:24.095	11.869	0.028	9.429	0.024	7.951	0.027	2MASS	6.400	0.002	6.293	0.002	5.777	0.001	5.527	0.001	3.664	0.003	1.684	0.854	II∗
832	18:32:26.095	−2:30:20.597	17.363	0.015	15.061	0.007	13.852	0.015	CFHT	13.096	0.008	12.896	0.011	12.712	0.034	12.792	0.103	7.987	0.109	1.713	1.116	II∗

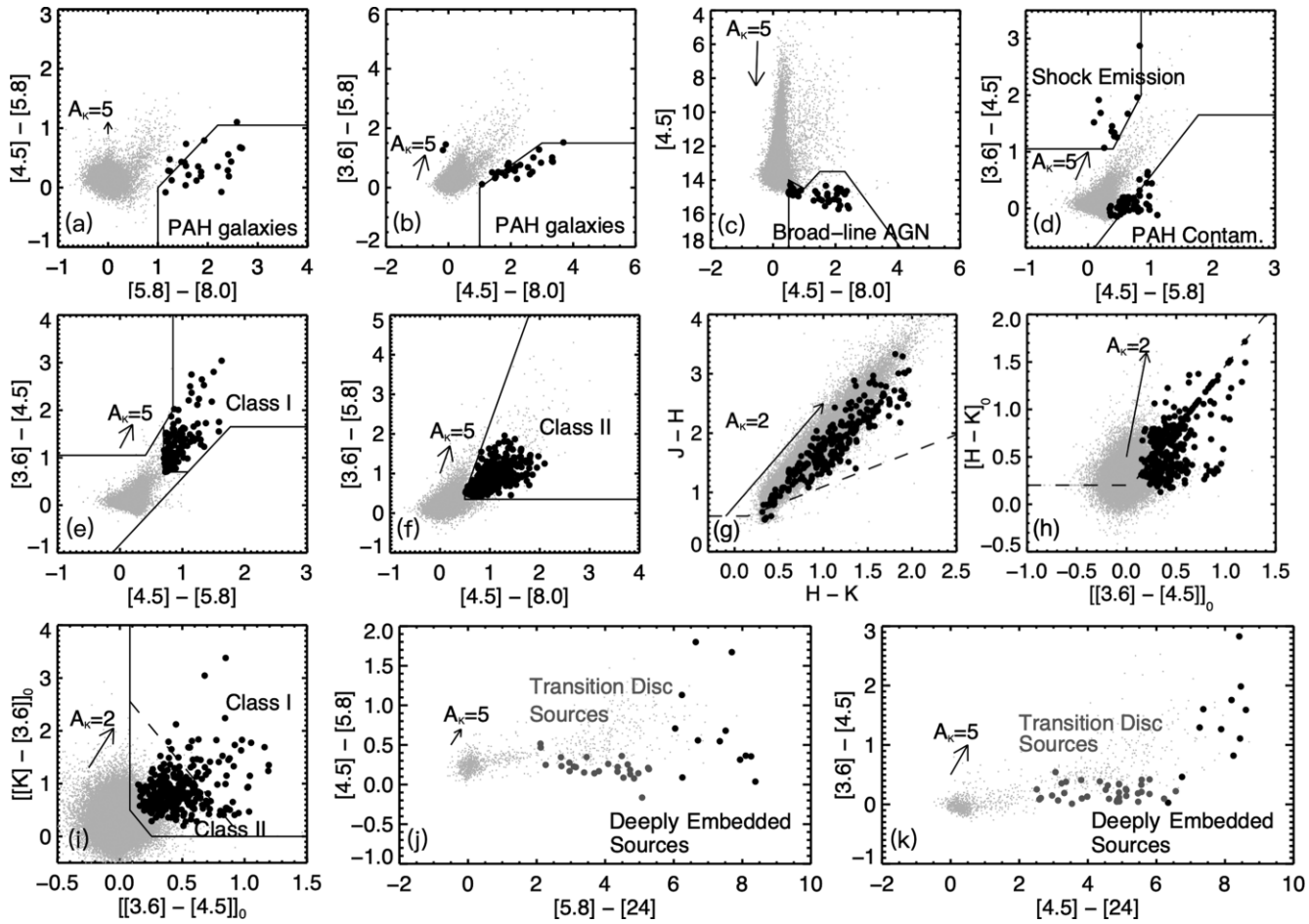

Figure 6. *First Row*: (a–d) Removal of non-YSO sources of star-forming galaxies (28), AGNs (47), shock emission (12), and PAH aperture-contaminated sources (53). *Second Row*: (e, f) Classification of Class I and Class II YSOs with *Spitzer* Ch1–4 bands. (g, h) Measuring extinction for dereddening of all the points and dots in (i). (g) Star loci before dereddening: the black dots will be on the two dashed lines after the dereddening, like those in (h). Stars not located on the two dashed lines in (h) have already been de-reddened in (g). *Bottom Row*: (i) Classification of Class I and Class II YSOs with *K* and *Spitzer* Ch1–2 bands, and (j, k) classification of deeply embedded sources and transition disc sources with MIPS. All of the colour excess ratios in this figure adopt the extinction law reported by Flaherty et al. (2007).

tallies. The distributions of these classified YSOs are shown in Figs 7(b)–(c).

In total, 832 YSOs were classified in the 'Target' field, including 15 deeply embedded sources, 135 Class I, 647 Class II, and 35 transition disc sources, as shown in Fig. 8, and their detailed measurements are listed in Table 3. In addition, two YSOs were found in the 'Control' field (one Class II, and one transition disc source). Because of the low column density of molecular clouds and the fact that these objects are unmatched with the YSOs from Dunham et al. (2015), these two are very likely contaminants and are not included in our YSO catalogue.

We repeated the same classification procedures using only 2MASS as the NIR data, and obtained 657 YSOs in the 'Target' field. With UKIDSS–CFHT data, two of the 657 YSOs are not in the final catalogue, and another two have their stages re-classified. Therefore, our UKIDSS–CFHT data finds ∼27 per cent more YSOs.

We briefly cross-matched our YSO catalogue with four earlier works (with 2-arcsec position tolerance), including: 557 YSOs with 2MASS data in all our observed field from Dunham et al. (2015); 264 Class III YSOs with *Chandra* data in the W40 subregion from Kuhn et al. (2010); 299 YSOs with *Chandra* data in the Serpens South subregion from Winston et al. (2018); and 67 Class 0/I protostars with ALMA data in the Serpens South subregion from Plunkett et al. (2018). Note that the studies by Kuhn, Winston, and Plunkett have limited fields of view, and only Dunham has the same field coverage for a full comparison. We obtained 379 YSOs matched with Dunham et al. (2015) in the 'Target' field, while obtaining 374 YSOs matched with the 657 YSOs that were identified using only 2MASS data, as described in Section 4.1. In the latter case, the other 183 YSOs that are only in the work by Dunham are randomly distributed, while the 283 only in ours trace molecular clouds well. This is due to the differences between the YSO identification methods of G09 and Harvey et al. (2007). We matched 82 YSOs with Kuhn et al. (2010) and all the 299 YSOs with Winston et al. (2018). This shows that the fraction of disc-less stars in Serpens South is not as big as in W40. We matched 33 YSOs with Plunkett et al. (2018), which suggests that Serpens South is younger than W40. In summary, when all these previously identified YSOs are deducted, 309 are unique to our catalogue.

Nevertheless, this YSO catalogue may still contain some misclassified contaminants. G09 applied these classification diagrams to the real AGNs that were identified by Stern et al. (2005) in the Bootes field. The results suggest a ubiquitous contamination such that 7.9 ± 1.0 AGNs will be classified as YSOs per square

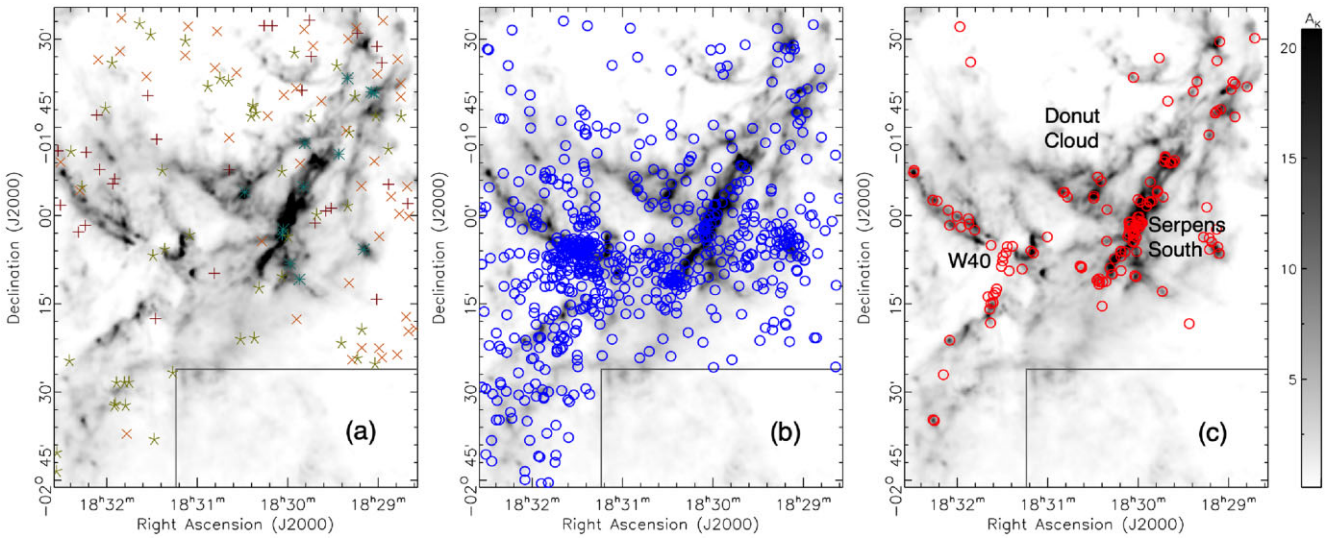

Figure 7. (a) Distributions of contaminations. (b) Class II and transition disc sources. (c) Class I and deeply embedded sources. The positions of PAH galaxies (+), AGNs (×), PAH aperture-contaminated sources (⋆), and shock-emission knots (∗) are marked in (a). The underlying images in the three panels are the same, namely the extinction map of A_K derived from the *Herschel* hydrogen column density map.

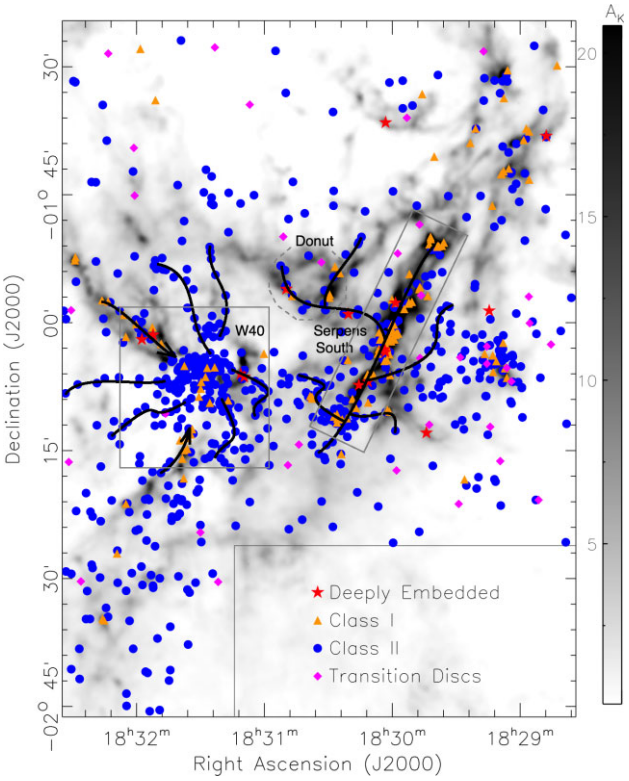

Figure 8. A more detailed YSO distribution in the 'Target' field. Black curves are the simplifications of nearby MST branches that connect surrounding YSOs to the centre cluster. The underlying image is the same as in Fig. 7. The two boxes mark the main parts of the W40 and Serpens South clusters. The dashed circle marks a 'Donut' cloud, which has a donut-shaped MST subgraph of YSOs.

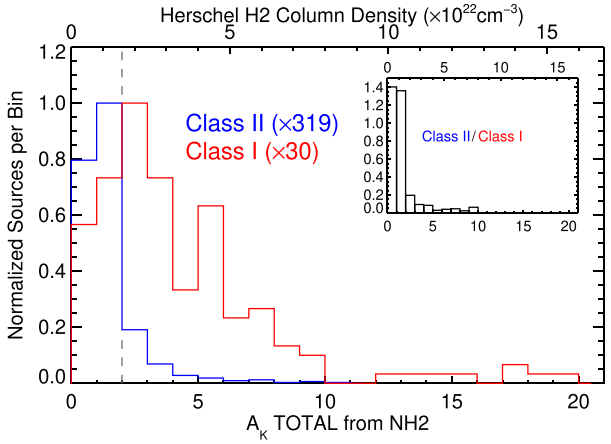

Figure 9. Histogram of the total extinction of Class I (including deeply embedded sources) and Class II (including transition disc sources) YSOs. Both are normalized by the peak bin value, which is 30 for Class I and 319 for Class II. The dashed line marks $A_K = 2$, where Class I and Class II are almost separated. The ratios of normalized sources per bin between Class II and Class I are very distinct, and therefore we show them in the subfigure.

degree. Hence this gives us about 8.8 ± 1.1 YSOs that will be AGN contaminations in the 'Target' field (area ∼ 1.11 deg^2). On the other hand, some real YSOs can be falsely flagged as PAH-emitting sources or AGNs. Gutermuth et al. (2008b) analysed a large extragalactic survey field to estimate the residual extragalactic contamination in the YSO census at 6.4 deg^{-2} for Class I sources and 3.8 deg^{-2} for Class II sources. Thus we might misclassify ∼11.3 real YSOs as contamination sources in the 'Target' field, although we have ∼8.8 contamination sources in the YSO list.

4.2 Spatial distributions of YSOs

4.2.1 Distributions of Class I and II YSOs

As shown in Figs 7(b)–(c), Class I YSOs (with deeply embedded sources included) have a stronger correlation with the molecular gas than do Class II sources (with transition disc sources included). And statistically, Fig. 9 shows the YSO histograms with the column density of molecular clouds in terms of A_K. In the entire 'Target' field, around 74 per cent of Class I YSOs fall into the range of A_K

> 2, but only 16 per cent of Class II YSOs are there. We can see that most YSOs are concentrated in the W40 region and the Serpens South region.

In W40, YSOs are mostly Class II. Based on the existence of O and B main-sequence stars in the central region of W40, Shuping et al. (2012) give an upper limit of the cluster age of about 7 Myr. Thus the age of the W40 cluster is of the order of a few Myrs, consistent with typical ages of Class II YSOs. The two molecular filaments (black curves marked with arrows in Fig. 8) that converge to W40 from the top left and bottom left contain obviously more Class I YSOs, which are likely outside the W40 H II bubble (green hourglass shape in Fig. 1).

In the Serpens South region, which consists of multiple less-massive star-forming subregions, about half of the YSOs are classified as Class I, which is a much higher percentage than in the W40 region. Along the central ridge of the molecular filament, massive YSOs are found to have ages mostly less than ∼ 0.46–0.72 Myr (Plunkett et al. 2018), while a few flat-spectrum YSOs give an upper limit of the cluster age of about 1 Myr (Dunham et al. 2015). Furthermore, based on the observations of HC7N clumps, the chemical evolution model suggests that the star-forming age along the ridge is less than ∼0.3 Myr (Friesen et al. 2013). A subregion north of Serpens South has its YSOs distributed along the cloud in the shape of a ring, so we name it the 'Serpens South Donut'. In the Donut region, a HC7N clump is also detected around a deeply embedded source, suggesting a similar young age of about 0.2 Myr. Consistently, the high percentages of Class I YSOs identified in these two regions represent a younger cluster age of less than ∼1 Myr.

G09 calculated the number ratios of Class I/Class II in 36 star-forming cores, and the median value of them is ∼0.27. The entire W40–Serpens South region has a similar mean ratio of 0.22. However, if we take YSOs from two specific box regions in Fig. 8, the Class I/Class II ratio is 0.119 for W40, and 0.849 for Serpens South. The Serpens South region may rank among the youngest regions (Gutermuth et al. 2008a; Li et al. 2019).

4.2.2 *YSOs and molecular clouds*

Fig. 10 shows a comparison between the 'FRONT' and the 'TOTAL' extinctions of YSOs, which represents how deeply these YSOs are embedded. As stated above, A_K FRONT is measured through the NIR/MIR colour excess, and A_K TOTAL is estimated from the *Herschel* map. For the deeply embedded sources, the magnitude in the MIPS band is used, and the max–min brightness range corresponds to [2.4, 8.6] mag. For the Class I YSOs , Ch2 band is used with a range of [4.7, 15.0] mag. For the Class II YSOs, Ch1 band is used with a range of [3.2, 16.0] mag. For the transition disc sources, *H* band is used with a range of [6.7, 17.3] mag. In addition, the 'FRONT' extinction of 98 YSOs cannot be measured owing to the lack of *H*-band data, and thus they are not on the diagram.

In general, we expect all the points to be located beneath the A_K FRONT = A_K TOTAL equivalent line (the middle black dashed line in Fig. 10), because the gas in front of a star should always be less than the total. However, most YSOs in Fig. 10 are beneath the upper A_K line (with a ratio of 2). In addition to the measurement uncertainties of the front and the total extinction, this could also be caused by the different spatial resolutions of telescopes. *Herschel* images have a relatively lower resolution of ∼20 arcsec, which can attenuate the higher column densities of fibres (Hacar et al. 2018) to lower averaged values. Because NIR/MIR images have a higher

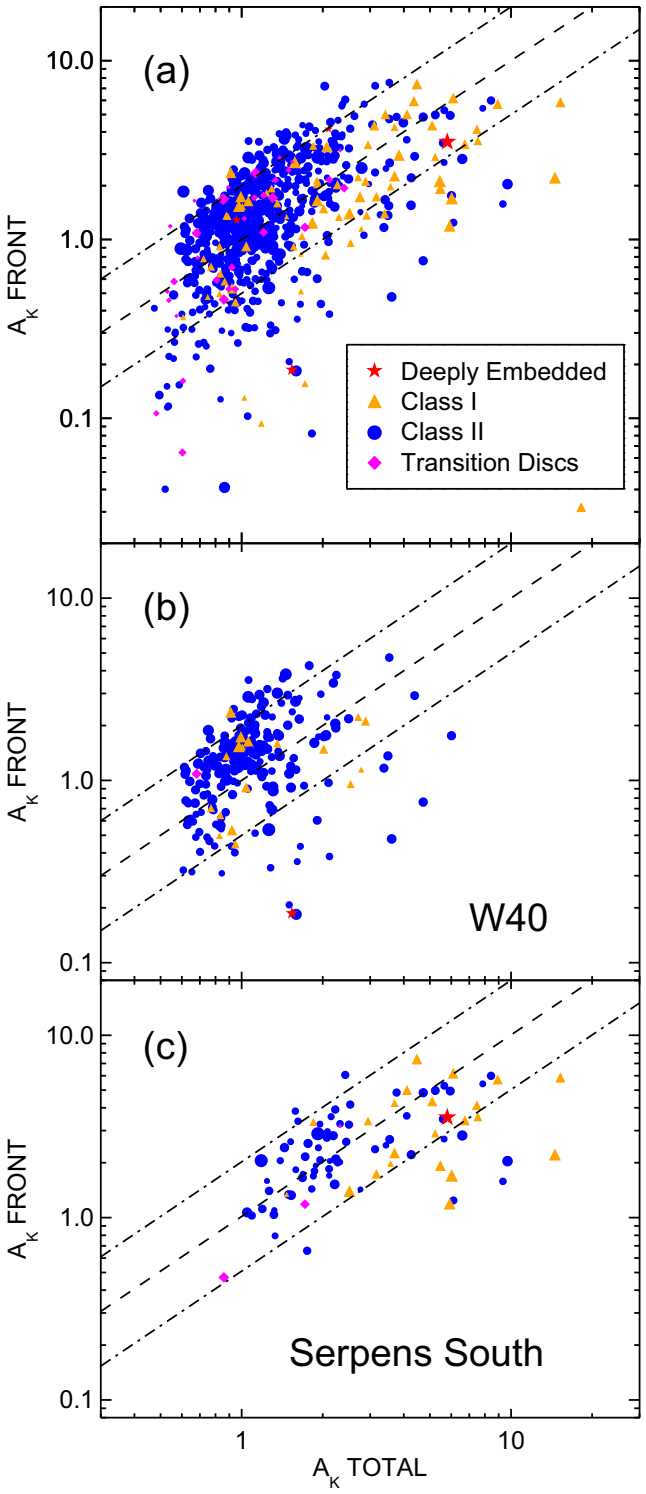

Figure 10. (a) Comparison between the total and the front extinction of all YSOs except for 98 that lack NIR detections for A_K FRONT and 13 with A_K FRONT = 0. (b) and (c) correspond to stars inside the two grey boxes in Fig. 8. The symbol sizes for each species represent the relative brightness given by the dereddened magnitude. The dotted and dashed lines are A_K FRONT = [2, 1, 0.5] A_K TOTAL, from top to bottom.

resolution of ∼1 arcsec, the front extinction to YSOs caused by dense fibres is more precisely measured. Thus the front extinction can be higher than the total value.

Nevertheless, this factor does not prevent us from noticing some interesting phenomena. (1) The Class II YSOs tend to lie in the region with A_K TOTAL < 2 mag, while the Class I YSOs appear in the region between 2 and 20 mag, as also suggested in Fig. 9. This again indicates the strong correlation between the Class I YSOs and the molecular gas. Most Class II YSOs are around the equivalent line, within a factor of 2, suggesting that they are lying inside or around the molecular gas. Meanwhile, we can also see a fraction of Class II YSOs with A_K FRONT < 0.6 mag and A_K TOTAL < 2 mag are likely having smaller symbol sizes . These sources could be slightly farther away from the molecular filaments, and in these locations it would be hard to form massive YSOs because of the difficulty in accreting enough gas. (2) The YSOs in the W40 region (the square region in Fig. 8) are more concentrated around $A_K \approx 1$ mag, both for the FRONT and for the TOTAL. These YSOs belong mainly to the compact central cluster that was born in a massive star-forming region. The cluster has been evolving for a few Myrs, and the surrounding molecular gas could be exhausted or expelled by stellar feedback. It could be the reason why the extinction of these Class II YSOs is relatively lower. The outliers at higher A_K TOTAL lie roughly on the several filamentary branches that still converge towards the W40 cluster. (3) The YSOs in the Serpens South region (the long rectangular region in Fig. 8) are obviously distributed at a higher A_K TOTAL between 1 and 20 mag. Consequently, compared with the Class II YSOs, the Class I YSOs are located at the ridge of the filament where the column densities are higher. Above A_K TOTAL >3, the A_K FRONT of YSOs does not increase linearly, but instead shows a flat trend. This suggests that some YSOs are not presented here, especially those deeply embedded in thick filaments with A_K TOTAL > 3. These YSOs are not necessarily undetected though. There are 98 YSOs not in this diagram because they lack *H*-band detection, and we find that 88 of them have A_K TOTAL >2, of which about 60 are in the Serpens South region. As a result, most YSOs in this region are likely to be detected, although some deeply embedded sources do not have the necessary *H*-band data for the front extinction calculation. In addition, while most YSOs surround the A_K FRONT $= A_K$ TOTAL equivalent line, a number of outliers are located below the lower A_K line (with a ratio of 0.5) in Fig. 10. Those outliers also have lower front extinction, although they have higher A_K TOTAL values (around 10 mag). Most of these outliers belong to a small central cluster, including the one in the lower right corner. (4) The situation for the YSOs in the remaining regions is similar, in that the Class I YSOs lie on the thicker filaments. Most Class II YSOs are around the equivalent line within a factor of 2, but a small portion of them and even a few Class I YSOs probably have a smaller front extinction, and are usually less bright. These fainter outliers may have originated in small star-forming regions, where the natal gas is quickly exhausted.

4.2.3 Minimum spanning tree of YSOs

To characterize the YSO distributions more quantitatively and explore potential connections among different regions, we constructed the minimum spanning tree (MST) for these YSOs (e.g. G09). The MST is a concept that originated in graph theory, but is frequently used to separate close stellar clusters (e.g. Battinelli 1991; Hetem & Lepine 1993; Cartwright & Whitworth 2004; Gutermuth et al. 2009; Kirk, Offner & Redmond 2014). The MST here is generated by connecting the nearest neighbour of each point first to form an 'island', and then connecting every two nearby 'islands' with the shortest path. New 'islands' gradually grow bigger, until all points are connected to form a tree structure in which the total branch length is the minimum.

Fig. 8 shows some prominent MST branches of the main body of the W40 and the Serpens South regions, while the detailed analysis of the MST is presented in Sun et al. 2022. The Donut cloud is young and its branches are well connected to the Serpens South rather than to the W40 region. Some clusters located to the west or north-west of the Serpens South region are relatively independent, and their MST branches are not marked here.

For W40, eight branches spread to all directions from the central cluster, in the shape of a spider. Six of the branches, illuminated by Class II YSOs, have no obvious corresponding filamentary molecular clouds. But the other two match the filaments well, and they contain more Class I YSOs. Therefore, the W40 central cluster is probably sitting on what was a convergence point of many filaments in the past, as suggested by many works that find that massive star-forming regions need to accrete more gas from multiple subfilaments (e.g. Schneider et al. 2012; Peretto et al. 2013; Rayner et al. 2017).

The Serpens South region also has several connected filaments, in the shape of a gecko, and its main body is on a thick and over-dense filament. The Class I YSOs (including the deeply embedded sources) lie roughly at intersection points of the MST branches and the ridge of the main filament. This again suggests that the gas is supplied from subfilaments. Although the main body is depicted by a long and straight MST branch, it consists of several smaller star-forming regions. Based on the HC7N detections, the star-forming age along the ridge of the molecular filament is less than ∼ 0.3 Myr (Friesen et al. 2013). But in the north part of this filament, where the cloud is the most overdense, the age is even younger.

5 SUMMARY

We compiled a large NIR and MIR catalogue of point sources in the W40–Serpens South region, based on the deep NIR data of CFHT observations in combination with the 2MASS, UKIDSS, and *Spitzer* catalogues. To identify stellar sources accurately, data reduction processes including image-stacking and photometry corrections need to be performed precisely. For the CFHT data, we adopted the 'AstrOmatic' suite and optimized the reduction processes, instead of using the SIMPLE pipeline, to obtain precise images for photometry. Multiple corrections were applied to these photometry results, such as those for optics astigmatism and telescope spikes, colour term fitting, and zero-magnitude correction. To maximize the data accuracy, we used 2MASS and UKIDSS as the standard catalogues to resolve saturated bright stars, subdivide the zero magnitude, and perform the astigmatism correction. Catalogues were cross-calibrated before being integrated. In addition, we cross-matched the results with *Spitzer* data, and included the publicly available *Herschel* hydrogen column density map to reveal the total dust extinction. Finally, all the information was combined to form this large near-mid-infrared catalogue of point sources for this region.

Based on this large catalogue, we identified 832 young stellar objects (YSOs) after filtering out contamination sources, and classified 15, 135, 647, and 35 of them to be deeply embedded sources, Class I YSOs, Class II YSOs, and transition disc sources. We also compared the front extinction of these YSOs with the total column densities at their locations to check their depths in the molecular clouds, and generated the MST of the YSOs to check their spatial

correlation with the molecular clouds. In general, the YSOs are well correlated with the filamentary structures of the molecular clouds in this region, especially the Class I YSOs (including the deeply embedded sources). The W40 central region is dominated by evolved Class II YSOs, and the surrounding materials were probably expelled by previous stellar feedback. Around the W40 central cluster, there are eight prominent MST branches. Of these, six branches populated exclusively with Class II YSOs show little correlation with molecular filaments, but the other two include detectable gas and corresponding Class I YSOs. The YSOs in the Serpens South region lie mainly on the thick filamentary clouds, and the Class I YSOs emerge roughly at intersection points where the subbranches join with the main body filament. The front extinction of them is also higher compared with that in the W40 region, and the *H*-band emission of some deeply embedded ones cannot penetrate the thick molecular gas. In summary, the massive star formation in the W40–Serpens South region is commonly supported by the gas supply from multiple subfilaments. Our results imply a convoluted YSO distribution in this important star-forming region.

ACKNOWLEDGEMENTS

We sincerely appreciate the anonymous referee for their very careful reading and helpful comments, which greatly improved the manuscript. We gratefully thank all the members in the PMO star-forming group for valuable discussions. MZ is supported by the National Natural Science Foundation of China (NSFC) grant no. 12073079. HW and YM acknowledge support from NSFC grant nos 11973091 and 11973090. SNZ acknowledges support from NSFC grant no. 11573070. RAG gratefully acknowledges funding support for this work from NASA Astrophysics Data Analysis Program (ADAP) award NNX17AF24G. JS acknowledges financial support from the China Scholarship Council (CSC). JS sincerely thanks Chi-hung Yan for his valuable explanation of the CFHT 'I'wii pipeline. Observations of this work were obtained with WIRCam, a joint project of Taiwan, Korea, Canada, France and the Canada-France-Hawaii Telescope (CFHT), which is operated by the National Research Council (NRC) of Canada, the Institute National des Sciences de l'Univers of the Centre National de la Recherche Scientifique of France and the University of Hawaii. This work is based in part on observations made with the *Spitzer* Space Telescope, which was operated by the Jet Propulsion Laboratory, California Institute of Technology under a contract with NASA. This research has made use of data from the *Herschel* Gould Belt survey (HGBS) project, which is a *Herschel* Key Programme jointly carried out by SPIRE Specialist Astronomy Group 3 (SAG 3), scientists of several institutes in the PACS Consortium, and scientists of the *Herschel* Science Center (HSC). This research has made use of data products from the UKIDSS Galactic Plane Survey (GPS), which is one of the five NIR Public Legacy Surveys that are being undertaken by the UKIDSS consortium and is available online through the WFCAM Science Archive (WSA). This research has made use of the VizieR catalogue access tool, CDS, Strasbourg, France. Through VizieR, we made use of data products from the Two Micron All Sky Survey (2MASS), which is a joint project of the University of Massachusetts and the Infrared Processing and Analysis Center/California Institute of Technology, funded by NASA and the National Science Foundation. This research has made use of astronomical pipeline software from the 'AstrOmatic', which is located at the Institut d'Astrophysique de Paris (IAP), France, and benefits from support by the French PNC and PNG. This research has made use of TOPCAT, which is an interactive graphical viewer and editor for tabular data.

DATA AVAILABILITY

The YSO catalogue underlying this article is available in the Supporting Information. The 2MASS-UKIDSS-CFHT NIR catalogue has been uploaded to the Science Data Bank, at http://doi.org/10.57760/sciencedb.j00149.00001, and is publicly available. The NIR images observed or pieced together can be shared on reasonable request to the corresponding author.

REFERENCES

André P., 2002, in Bouvier J., Zahn J.-P., eds, Star Formation and the Physics of Young Stars: Summer school on Stellar Physics X, Vol. 3. EAS Publications Series, Paris, p. 1
Battinelli P., 1991, A&A, 244, 69
Bertin E., 2010a, Astrophysics Source Code Library, record ascl:1010.063
Bertin E., 2010b, Astrophysics Source Code Library, record ascl:1010.068
Bertin E., Arnouts S., 2010, Astrophysics Source Code Library, record ascl:1010.064
Bertin E., Delorme P., Bouy H., 2012, Astrophysics and Space Science Proceedings: Star Clusters in the Era of Large Surveys. Springer-Verlag, Berlin, Heidelberg, p. 71
Bontemps S. et al., 2010, A&A, 518, L85
Calabretta M. R., Greisen E. W. 2002, A&A, 395, 1077
Calvet N., Muzerolle J., Briceño C., Hernández J., Hartmann L., Saucedo J. L., Gordon K. D., 2004, AJ, 128, 1294
Caratti o Garatti A. et al., 2017, Nature Phys., 13, 276
Cartwright A., Whitworth A. P., 2004, MNRAS, 348, 589
Dunham M. M. et al., 2013, AJ, 145, 94
Dunham M. M. et al., 2015, ApJS, 220, 11
Evans Neal J. I. et al., 2009, ApJS, 181, 321
Fang M., van Boekel R., Wang W., Carmona A., Sicilia-Aguilar A., Henning T., 2009, A&A, 504, 461
Flaherty K. M., Pipher J. L., Megeath S. T., Winston E. M., Gutermuth R. A., Muzerolle J., Allen L. E., Fazio G. G., 2007, ApJ, 663, 1069
Friesen R. K., Medeiros L., Schnee S., Bourke T. L., di Francesco J., Gutermuth R., Myers P. C., 2013, MNRAS, 436, 1513
Greene T. P., Wilking B. A., Andre P., Young E. T., Lada C. J., 1994, ApJ, 434, 614
Gutermuth R. A., Megeath S. T., Pipher J. L., Williams J. P., Allen L. E., Myers P. C., Raines S. N., 2005, ApJ, 632, 397
Gutermuth R. A. et al., 2008a, ApJ, 673, L151
Gutermuth R. A. et al., 2008b, ApJ, 674, 336
Gutermuth R. A., Megeath S. T., Myers P. C., Allen L. E., Pipher J. L., Fazio G. G., 2009, ApJS, 184, 18
Hacar A., Tafalla M., Forbrich J., Alves J., Meingast S., Grossschedl J., Teixeira P. S., 2018, A&A, 610, A77
Harvey P., Merín B., Huard T. L., Rebull L. M., Chapman N., Evans Neal J. I., Myers P. C., 2007, ApJ, 663, 1149
Hetem A. J., Lepine J. R. D., 1993, A&A, 270, 451H
Hurt R. L., Barsony M., 1996, ApJ, 460, L45
Kirk H., Offner S. S. R., Redmond K. J., 2014, MNRAS, 439, 1765
Könyves V. et al., 2015, A&A, 584, A91
Kuhn M. A., Getman K. V., Feigelson E. D., Reipurth B., Rodney S. A., Garmire G. P., 2010, ApJ, 725, 2485
Lada C. J., 1987, in Peimbert M., Jugaku J., eds, Proc. IAU Symp. 115, Star Forming Regions. D. Reidel Publishing Co., Dordrecht, p. 1
Landsman W. B., 1993, in Hanisch R. J., Brissenden R. J. V., Barnes J., eds, ASP Conf. Ser. Vol. 52, Astronomical Data Analysis Software and Systems II. Kluwer, Dordrecht, p. 246
Li J., Myers P. C., Kirk H., Gutermuth R. A., Dunham M. M., Pokhrel R., 2019, ApJ, 871, 163
Liu Y. et al., 2019, A&A, 622, A75
Lucas P. W. et al., 2008, MNRAS, 391, 136
Nakamura F. et al., 2011, ApJ, 737, 56
Ortiz-León G. N. et al., 2017, ApJ, 834, 143
Peretto N. et al., 2013, A&A, 555, A112

Plunkett A. L., Fernández-López M., Arce H. G., Busquet G., Mardones D., Dunham M. M., 2018, A&A, 615, A9
Pokhrel R. et al., 2020, ApJ, 896, 60
Predehl P., Schmitt J. H. M. M., 1995, A&A, 500, 459
Preibisch T., Zinnecker H., 1999, AJ, 117, 2381
Rayner T. S. M. et al., 2017, A&A, 607, A22
Rieke G. H., Lebofsky M. J., 1985, ApJ, 288, 618
Schneider N. et al., 2012, A&A, 540, L11
Shang H., Lizano S., Glassgold A., Shu F., 2004, ApJ, 612, L69
Shimoikura T., Dobashi K., Nakamura F., Shimajiri Y., Sugitani K., 2019, PASJ, 71, S4
Shuping R. Y., Vacca W. D., Kassis M., Yu K. C., 2012, AJ, 144, 116
Skrutskie M. F. et al., 2006, AJ, 131, 1163
Stern D. et al., 2005, ApJ, 631, 163
Su K. Y. L. et al., 2006, ApJ, 653, 675
Sugitani K. et al., 2011, ApJ, 734, 63
Sun J., Gutermuth R. A., Wang H., Zhang S., Long M., 2022, preprint (arXiv:2207.09042)
Testi L., Sargent A. I., 1998, ApJ, 508, L91
Tsuboi Y., Koyama K., Hamaguchi K., Tatematsu K., Sekimoto Y., Bally J., Reipurth B., 2001, ApJ, 554, 734
Wang W. -H., 2010, in Mizumoto Y., Morita K.-I., Ohishi M., eds, ASP Conf. Ser. Vol. 434, Astronomical Data Analysis Software and Systems XIX. Astron. Soc. Pac., San Francisco, p. 87
Winston E., Wolk S. J., Gutermuth R., Bourke T. L., 2018, AJ, 155, 241
Zeilik M. I., Lada C. J., 1978, ApJ, 222, 896
Zhang M., Fang M., Wang H., Sun J., Wang M., Jiang Z., Anathipindika S., 2015, ApJS, 219, 21
Zhang M., Kainulainen J., 2019, A&A, 632, A85

SUPPORTING INFORMATION

Supplementary data are available at *MNRAS* online.

Table S1. The near-mid-infrared catalogue of identified YSOs in the 'Target' field. Two default values, –100 and 10, are used as placeholders for non-detections in magnitude and uncertainty.

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

This paper has been typeset from a TEX/LATEX file prepared by the author.