Universidade de São Paulo
Instituto de Astronomia, Geofísica e Ciências Atmosféricas
Departamento de Astronomia

Felipe Donizeti Teston Navarete

The Formation of High-Mass Stars: from High-Mass Clumps to Accretion Discs and Molecular Outflows

(A Formação de Estrelas de Alta Massa: dos Glóbulos de Alta Massa aos Discos de Acreção e Jatos Moleculares)

São Paulo
2018

Felipe Donizeti Teston Navarete

The Formation of High-Mass Stars: from High-Mass Clumps to Accretion Discs and Molecular Outflows

(A Formação de Estrelas de Alta Massa: dos Glóbulos de Alta Massa aos Discos de Acreção e Jatos Moleculares)

Tese apresentada ao Departamento de Astronomia do Instituto de Astronomia, Geofísica e Ciências Atmosféricas da Universidade de São Paulo como requisito parcial para a obtenção do título de Doutor em Ciências.

Versão corrigida. O original encontra-se disponível na Unidade.

Área de Concentração: Astronomia
Orientador: Prof. Dr. Augusto Damineli Neto

São Paulo

2018

To my beloved parents. / Aos meus amados pais.

Agradecimentos/Acknowledgements

A elaboração deste trabalho não teria sido possível sem a colaboração, estímulo e empenho de diversas pessoas. O presente volume representa grande parte do que foi a minha vida nos últimos cinco anos, tanto nos aspectos positivos quanto nos negativos, nas alegrias e nas tristezas, nas noites em claro sem poder dormir por conta de crises de ansiedade ou hipoglicêmicas e nas horas que deixei de ficar com as pessoas que eu amo para me dedicar ao meu Amor pela Ciência e Astronomia. Gostaria de expressar minha gratidão a todos os amigos e colegas que, direta ou indiretamente, contribuíram para que esta tarefa se tornasse uma realidade. A todos vocês, manifesto os meus sinceros agradecimentos.

Aos meus pais, Donizeti e Lidia, pelo amor, carinho e pelas lições de vida, trabalho e dignidade que recebi ao longo de toda a minha existência. Palavras não são suficientes para expressar a minha eterna gratidão e, a eles dedico este trabalho. Agradeço também meus familiares, por todo o suporte e carinho recebido ao longo de toda a minha vida.

Ao meu orientador, Prof. Dr. Augusto Damineli, pela oportunidade de trabalhar com um tema encantador, por sua paciência, amizade e imprescindível orientação acadêmica nesses quase dez anos de trabalho em equipe.

Ao meu relator, Prof. Dr. Marcos Diaz, pelas suas críticas construtivas e sugestões. A todo o corpo docente do Depto. de Astronomia do IAG, pelos ensinamentos dentro e fora de uma sala de aula, discussões e oportunidades oferecidas ao longo do meu Doutorado.

Aos amigos do IAG e da Astronomia, Marcus, Diana, Márcio, Iara, Quint, Paulo, Fellipy, Mota, Leonardo, Moser, o Ti e a Ju; pelas conversas, risadas, os cafés (principalmente os cafés!) e aventuras que enfrentamos juntos ao longo desses anos. Em especial, ao Pedro Beaklini que me ajudou muito durante minha guinada na direção da Radio Astronomia. Um agradecimento especial às equipes técnica e administrativa do Depto. de Astronomia,

que contribuíram ativamente para o bom andamento deste trabalho.

Aos companheiros Marco Antônio, Paulo Bianchi, Jonatan Sena e Natália Monteiro; pela amizade em todos esses anos, pelas risadas e brincadeiras, e por todo o apoio e suporte nas horas mais difíceis. Aos amigos da Física e da vida: Teruya, Adriano e Mario. Aos amigos da ETE: Vitor, Rafael, Paulo, Jaime, Caio, São e Fernando. Ao pessoal do CASP, pelos ensinamentos e pela excelente companhia nas noites de atendimento ao público no IAG. Agradeço à Annee, pelo companheirismo, carinho e paciência nesses últimos meses do Doutorado.

I am grateful for having worked under the supervision of Dr Silvia Leurini and Prof Dr Karl Menten during my one-year experience as a Guest Student at the Max-Planck Institute for Radio Astronomy, in Bonn, Germany. I especially thank Dr Leurini for her assistance and patience during my entire stay at MPIfR and along my process of learning on how to process and analyse sub-millimetre data. I greatly appreciate all the discussion and possibilities she provided during our collaboration. I also appreciate the disposition of Dr Menten for always opening up possibilities of collaborations in different fields and using different techniques. A special thank to the ATLASGAL group, the MPIfR staff and the friends that made my stay in Bonn so pleasant and productive. I would also like to express my gratitude to my flatmate, Bastian Piltz, for sharing his apartment in the lovely Endenich, and with all the help on my lack of knowledge of the German language! Vielen Dank für Alles!

Agradeço à FAPESP pelo financiamento da minha pesquisa sob os projetos 2013/11680-2 e 2014/20522-4. *This work is based on data acquired with the Atacama Pathfinder Experiment (APEX). APEX is a collaboration between the Max-Planck-Institut fur Radioastronomie, the European Southern Observatory, and the Onsala Space Observatory; and based on observations obtained at the Gemini Observatory, which is operated by the Association of Universities for Research in Astronomy, Inc., under a cooperative agreement with the NSF on behalf of the Gemini partnership: the National Science Foundation (United States), the National Research Council (Canada), CONICYT (Chile), Ministerio de Ciencia, Tecnologa e Innovacin Productiva (Argentina), and Ministrio da Cincia, Tecnologia e Inovao (Brazil).*

This thesis was written in LaTeX with the class IAGTESE, developed to write thesis and dissertations from IAG.

" *Look again at that dot. That's here. That's home. That's us. On it everyone you love, everyone you know, everyone you ever heard of, every human being who ever was, lived out their lives. The aggregate of our joy and suffering, thousands of confident religions, ideologies, and economic doctrines, every hunter and forager, every hero and coward, every creator and destroyer of civilization, every king and peasant, every young couple in love, every mother and father, hopeful child, inventor and explorer, every teacher of morals, every corrupt politician, every 'superstar', every 'supreme leader', every saint and sinner in the history of our species lived there-on a mote of dust suspended in a sunbeam.*"

Carl Sagan (1934-1996)

"*May it be a light to you in dark places, when all other lights go out.*"

J. R. R. Tolkien (1892-1973)

Resumo

Estrelas de alta massa têm grande impacto na evolução do Universo e o processo de formação destes objetos ainda é um problema em aberto na Astrofísica. Os detalhes das estruturas associadas às regiões mais próximas dos objetos centrais, tais como os discos circunstelares e a morfologia dos jatos próximos à base de lançamento, ainda não foram estudados em detalhe e carecem de evidências observacionais. Esta tese apresenta um estudo da formação de estrelas de alta massa em termos da evolução de glóbulos de alta massa (*clumps*), selecionados a partir do levantamento ATLASGAL, a partir de observações da molécula ^{12}CO na faixa espectral do sub-milimétrico. Enquanto observações *single-dish* no sub-milimétrico possibilitam o estudo em larga escala do processo de formação de estrelas de alta massa, observações com maior resolução angular são necessárias para investigar os detalhes das protoestrelas no interior dos glóbulos. Para isso, espectroscopia tri-dimensional no infra-vermelho próximo foi obtida para um grupo de fontes RMS para caracterizar o meio circunstelar de objetos estelares jovens e de alta massa (HMYSOs) em escalas lineares de ∼100-1000 UA.

A amostra TOP100 oferece uma oportunidade ímpar de analisar um conjunto estatisticamente completo de glóbulos de alta massa em diversas fases evolutivas. Observações realizadas com o radiotelescópio APEX de três transições rotacionais da molécula do CO (CO (4–3), CO (6–5) e CO (7–6)) foram utilizadas para estudar as propriedades do gás morno ($\lesssim$155 K) associado aos glóbulos, e obter as relações entre a emissão do CO e as propriedades físicas dos glóbulos. A luminosidade das diferentes transições do CO ($L_{\rm CO}$) foi obtida e sua análise mostrou que a emissão do gás aumenta em função do estágio evolutivo dos glóbulos (de glóbulos com emissão fraca no infravermelho longínquo a regiões H II) e em função da luminosidade bolométrica ($L_{\rm bol}$) e massa dos glóbulos ($M_{\rm clump}$). A

comparação entre os glóbulos de alta massa presentes na amostra TOP100 com fontes de menor massa observadas nas transições do CO (6–5) e CO (7–6), juntamente com a análise de uma amostra complementar de fontes observadas na transição do CO (10–9) mostrou que a dependência da luminosidade do CO ($L_{\rm CO}$) com $L_{\rm bol}$ aumenta em função do número quântico J associado à transição do CO. Este estudo também mostrou que as relações entre $L_{\rm CO}$ e $L_{\rm bol}$ são dominadas pelas fontes de alta luminosidade presentes na amostra analisada. a análise individual de fontes de baixa e alta luminosidade sugerem que a dependência entre $L_{\rm CO}$ e $L_{\rm bol}$ é a mesma em ambos os regimes de luminosidade, embora os valores de $L_{\rm CO}$ sejam sistematicamente maiores para os glóbulos de alta massa. Por fim, a análise da emissão do CO em altas-velocidades mostrou que ∼85% dos glóbulos presentes na amostra TOP100 apresentam jatos moleculares.

A seleção de objetos de alta massa isolados em estágio de acreção ativa é crucial para decidir se ela ocorre através de um disco de acreção e/ou via fusão de YSOs de menor massa. Para isso, observações no infra-vermelho próximo são ideais para se investigar o conteúdo dos glóbulos sub-milimétricos e resolver seus membros individuais. Devido a alta resolução espacial na banda K e a extinção interestelar moderada nesta faixa espectral, um conjunto de oito (8) HMYSOs associados a jatos em H_2 em larga-escala foram selecionados para observações espectroscópicas na banda K utilizando o espectrômetro NIFS no Gemini Norte.

Todos os objetos investigados com o NIFS apresentam emissão extendida no contínuo, bem como nas linhas espectrais típicas de fontes jovens, tais como o Brγ, transições do H_2 e a emissão nas bandas moleculares do CO. A emissão em H_2 está associada aos jatos moleculares em escalas de ∼100 UA em cinco das oito fontes (63%). A indentificação de jatos moleculares em escalas tão próximas ao objeto central indica que o processo de acreção de massa ainda está ativo nestes objetos. A emissão do Brγ provém do gás ionizado nas regiões mais próximas das fontes centrais ou regiões de choque próximas aos jatos. A espectro-astrometria da linha do Brγ em escalas de sub-píxeis, indica que a emissão do gás ocorre nas cavidades dos jatos moleculares ou delineiam estruturas alinhadas perpendicularmente aos jatos, tais como os discos de acreção. Cinco fontes também apresentam emissão nas bandas do CO (63%), e três HMYSOs apresentam linhas do CO em absorção (38%), indicando que estes objetos apresentam discos de acreção. A massa total do sistema "disco e protoestrela" foi determinada a partir do estudo da cinemática das linhas de

absorção do CO, detectadas em três objetos. A partir de modelos de rotação Kepleriana, as massas das fontes foram estimadas em 5±3, 8±5 e 30±10 $M_{\odot}$.

Os resultados obtidos a partir da espectroscopia tri-dimensional no infravermelho corroboram a hipótese de que estrelas de alta massa são formadas a partir de acreção por discos, de maneira similar ao observado para estrelas de baixa massa. A comparação entre a morfologia dos jatos moleculares identificados nos campos do NIFS e das correspondentes contrapartidas em escalas maiores indicam que os jatos apresentam diferentes graus de colimação ao longo de suas estruturas, explicadas pela multiplicidade de fontes nas proximidades da base de lançamento dos jatos ou efeitos de precessão no objeto central.

Abstract

High-mass stars play a significant role in the evolution of the Universe and the process that leads to the formation of such objects is still an open question in Astrophysics. The details of the structures connected to the central sources, such as the circumstellar discs and the morphology of the jets at their launching points, still lack of observational evidence. In this thesis, the high-mass star forming process is investigated in terms of the evolution of high-mass clumps selected from the ATLASGAL survey based on their ^{12}CO emission in the sub-millimetre. While single-dish sub-millimetre observations provide a large-scale view of the high-mass star formation process, higher angular resolution observations are required to disentangle the details of the protostars within the clumps. For this, three-dimensional infrared spectroscopy was obtained for a group of RMS sources to characterise the circumstellar environment of high-mass YSOs in linear scales of ∼100-1000 AU.

The ATLASGAL TOP100 sample offers a unique opportunity to analyse a statistically complete sample of high-mass clumps at different evolutionary stages. APEX data of three rotational J transitions of the CO (the CO (4–3), CO (6–5) and CO (7–6)) were used to characterise the properties of their warm gas ($\lesssim$155 K) content and to derive the relations between the CO and the clump properties. The CO line luminosities were derived and the analysis indicated that the CO emission increases as a function of the evolutionary stage of the clumps (from infrared-weak to H II regions) and as a function of the bolometric luminosity ($L_{\rm bol}$) and mass of the sources ($M_{\rm clump}$). The comparison of the TOP100 with low-mass objects observed in the CO (6–5) and CO (7–6), together with CO (10–9) data observed for a complementary sample of objects indicated that the dependency of the CO luminosity ($L_{\rm CO}$) with the bolometric luminosity of the sources gets steeper towards higher-J transitions. Although the CO luminosity of more luminous clumps are systematically

larger than the values obtained for the less luminous sources, the individual analysis of each subsample suggests a similar dependency of $L_{\rm CO}$ versus $L_{\rm bol}$ for each luminosity regime. Finally, the presence of high-velocity CO emission observed for the TOP100 suggests that $\sim$85% of the sources are driving molecular outflows.

The selection of isolated high-mass objects undergoing mass accretion is fundamental to investigate if these objects are formed through an accretion disc or if they are formed by merging of low-mass YSOs. The near-infrared window provides one of the best opportunities to investigate the interior of the sub-mm clumps and study in details their individual members. Thanks to the relatively high-resolution obtained in the *K*-band and the moderate reddening effectsin the *K*-band, a sample of eight (8) HMYSOs exhibiting large-scale H_2 outflows were selected to follow-up *K*-band spectroscopic observations using the NIFS spectrometer (Gemini North).

All sources exhibit extended continuum emission and exhibit atomic and molecular transitions typical of embedded objects, such as Brγ, H_2 and the CO lines. The H_2 lines are tracing the launching point of the large-scale jets in scales of $\sim 100\,$AU in five of eight sources (63%). The identification of jets at such small scales indicates that these objects are still undergoing mass accretion. The Brγ emission probes the ionised gas around the HMYSOs. The analysis of the Brγ spectro-astrometry at sub-pixel scales suggests that the line arises from the cavity of the outflows or from rotating structures perpendicular to the H_2 jets (*i.e.*, disc). Five sources also exhibit CO emission features (63%), and three HMYSOs display CO absorption features (38%), indicating that they are likely associated with circumstellar discs. By further investigating the kinematics of the spatially resolved CO absorption features, the Keplerian mass of three sources was estimated in 5$\pm$3, 8$\pm$5 and 30$\pm$10 $M_\odot$.

These results support that high-mass stars are formed through discs, similarly as observed towards low-mass stars. The comparison between the collimation degree of the molecular jets or outflows detected in the NIFS data with their large-scale counterparts indicate that these structures present a relatively wide range of collimation degrees.

List of Figures

List of Tables

Contents

List of Acronyms/Abbreviations

2MASS: Two-Micron All Sky Survey

AGB: asymptotic giant branch

ALMA: Atacama Large Millimeter Array

ALTAIR: Altitude conjugate Adaptive optics for the InfraRed

APEX: Atacama Pathfinder EXperiment (telescope)

AR: atmospheric refraction

ATLASGAL: The APEX Telescope Large Area Survey of the GALaxy

CHAMP$^+$: The Carbon Heterodyne Array of the MPIfR

CLASS: Continuum and Line Analysis Single Dish Software

CORNISH: Co-Ordinated Radio 'N' Infrared Survey for High-mass star formation

E: east (direction)

EGO: extremely green object

FIR: far-infrared

FLASH$^+$: First Light APEX Submillimeter Heterodyne receiver

FOV: field-of-view

FWHM: full width at half maximum

FWTM: full width at tenth maximum

GILDAS: Grenoble Image and Line Data Analysis Software

GLIMPSE: the (Spitzer) Galactic Legacy Infrared Mid-Plane Survey Extraordinaire

GMC: giant molecular cloud

HC: hot core

HFA: high-frequency array

Hi-GAL: Herschel infrared Galactic Plane Survey

HMSF: high-mass star-forming

HMYSO: high-mass young stellar object

IFU: integral-field unit

IMF: initial mass function

IR: infrared

IRAC: the (Spitzer) InfraRed Array Camera

IRDC: infrared dark cloud

ISM: interstellar medium

LABOCA: Large APEX Bolometer Camera

LFA: low-frequency array

LGS: laser guide star

LTE: local thermodynamic equilibrium

KS: Kolmogorov-smirnov (test)

MIPSGAL: Multiband Imaging Photometer Galactic Plane Survey

MIR: mid-infrared

MS: main-sequence

MSX: Midcourse Space Experiment

N: north (direction)

NE: north-east (direction)

NIFS: Near-Infrared Integral Field Spectrometer

NIR: near-infrared

NOAO: National Optical Astronomy Observatory

NW: north-west (direction)

OTF: on-the-fly

p2p: pixel-to-pixel (analysis)

PMS: pre-main sequence

PNe: planetary nebulae

PSC: pre-stellar core

PSF: point spread function

RGB: red, green and blue (image)

RMS: Red MSX Source (survey)

S: south (direction)

SE: south-east (direction)

SFR: star formation rate

SMA: Sub-Millimeter Array

sub-mm: sub-millimetre

SW: south-west (direction)

UCH II: ultra-compact H II region

UT: universal time

UV: ultra-violet

YSO: (low-mass) young stellar object

W: west (direction)

ZAMS: zero-age main-sequence

List of Symbols, Physical Constants and Units

$''$: arcsecond

$'$: arcminute

[$H_2/^{12}CO$]: abundance of H_2 over the number of ^{12}CO molecules

A: airmass

$A_{\rm pix}$: area of a spaxel, projected at the distance of the source (cm^{-2})

A_θ: area of the convolution beam ($arcsec^2$)

$A_{\upsilon,\rm J}$: Einstein coefficient of a given H_2 transition (s^{-1})

A_λ: interstellar reddening at a given photometric band (mag)

$A_{\rm ul}$: Einstein coefficient between the two energy levels u and l (in s^{-1})

AU: astronomical unit ($1.496{\cdot}10^{13}$ cm)

Å: Angstrom (10^{-10} m)

α: power-law index of the IMF, the luminosity-to-mass relation or the reddening law.

α, β, ϵ: intercept, slope and intrinsic scatter of the linear regression fit, respectively

$B_{\rm eff}$: beam efficiency

B_λ: black-body flux at a given wavelength ($erg\,s^{-1}\,cm^{-2}\,\mu m^{-1}$)

b: Galactic latitude (deg)

$\beta_{\rm N}$: power-law exponent of the radial distribution of the disc column density

$\beta_{\rm T}$: power-law exponent of the radial distribution of the disc temperature

$\beta_{\rm r}$: ratio between the outer and inner disc radii

c: light velocity ($2.998 \cdot 10^{10}\,cm\,s^{-1}$)

C_λ: calibration factor function (in logarithm units of ADU $[erg\,s^{-1}\,cm^{-2}\,\mu m^{-1}]^{-1}$)

d: distance of the source (pc or kpc)

$\Delta\lambda$: spectral width (Å or μm)

$\Delta\lambda_{\rm disp}$: linear spectral dispersion ($\mathrm{\AA\,pixel^{-1}}$)

Δp: projected pixel scale at the distance of the source ($\mathrm{AU\,pix^{-1}}$)

Δs: linear size of the CO emission (pc)

$\Delta s_{70\,\mu\mathrm{m}}$: linear size of the far-infrared emission at $70\,\mu$m (pc)

E_λ: atmospheric extinction curve ($\mathrm{mag\,airmass^{-1}}$)

$E_{\rm out}$: entrained mass of the outflow ($\mathrm{M_\odot\,km^2\,s^{-2}}$)

$E_{\rm u}$: upper-level energy of a given state u (erg or K)

$F_{\rm eff}$: forward efficiency

$F_{\rm Ly}$: Lyman continuum flux ($\mathrm{s^{-1}}$)

$F_{\rm norm}$: normalised spectral flux

F_λ: spectral flux at a given wavelength ($\mathrm{erg\,s^{-1}\,cm^{-2}\,sr^{-1}}$)

$F_{\rm CO}$: synthetic CO flux ($\mathrm{erg\,s^{-1}\,cm^{-2}\,\mu m^{-1}}$ units)

$F_{Ks,0}$: zero-point flux at the $K_{\rm s}$-band ($\mathrm{erg\,s^{-1}\,cm^{-2}\,\mu m^{-1}}$)

f_{H2}: ratio between the intensity of the $\mathrm{H_2}$ (1–0) S(1) transition at $2.1218\,\mu$m and the total intensity of the $\mathrm{H_2}$ emission

$F_{\rm out}$: mechanical force of the outflow ($\mathrm{M_\odot\,km\,s^{-1}\,yr^{-1}}$)

$\phi_{\rm beam}$: beam filling factor

G: gravitational constant ($6.67{\cdot}10^{-8}\,\mathrm{cm^3\,g^{-1}\,s^{-2}}$)

GHz: gigaherts (10^9 Hz)

$g_{\rm u}$: statistical weight of a given state u

i: inclination angle (degrees)

$I_{v,\rm J}$: intensity of a given $\mathrm{H_2}$ transition ($\mathrm{erg\,cm^{-2}\,s^{-1}}$)

$I_{2.12}$: intensity of the $\mathrm{H_2}$ (1–0) S(1) transition at $2.1218\,\mu$m ($\mathrm{erg\,s^{-1}\,cm^{-2}}$)

I_{H2}: total intensity of the $\mathrm{H_2}$ emission ($\mathrm{erg\,s^{-1}\,cm^{-2}}$)

$J_{\rm min}$: minimum rotational J CO level

$J_{\rm max}$: maximum rotational J CO level

h: Planck constant ($6.626 \cdot 10^{-27}\,\mathrm{erg\,s^{-1}}$)

$\mathrm{L_{bol}}$: bolometric luminosity ($\mathrm{L_\odot}$)

$L(H_2)$ or L_{H2}: luminosity of the $\mathrm{H_2}$ emission ($\mathrm{L_\odot}$)

$L_{\rm out}$: luminosity of the outflow ($\mathrm{M_\odot\,km^2\,s^{-2}\,yr^{-1}}$)

$\mathrm{L_\odot}$: Solar luminosity unit ($3.6{\cdot}10^{33}\,\mathrm{erg\,s^{-1}}$)

$\mathrm{L_*}$: stellar luminosity ($\mathrm{L_\odot}$)

ℓ: Galactic longitude (deg)

$\ell_{\rm proj}$: projected length (pc or AU)

λ: Wavelength unit (Å or μm)

λ_c: central wavelength (μm)

λ_0: central wavelength (Å or μm)

$\lambda/\Delta\lambda_{\rm disp}$: spectral resolving power

kpc: kiloparsec (10^3 pc)

KS: Kolmogorov-smirnov (rank factor)

$M_{\rm out}$: entrained mass of the outflow ($M_\odot$)

m_λ: magnitude at a given photometric band (mag)

$\dot{M}$: mass accretion rate ($M_\odot\,{\rm yr}^{-1}$)

$\dot{M}_{\rm out}$: mass-loss rate of the outflow ($M_\odot\,{\rm yr}^{-1}$)

m_H: mass of the hydrogen atom ($1.6733{\cdot}10^{-24}$ g)

$M(H_2)$ or M_{H2}: mass of the H_2 emission ($M_\odot$)

$M_\odot$: Solar mass unit ($1.99{\cdot}10^{33}$ g)

M_*: Stellar mass ($M_\odot$)

mas: milli-arcsecond

μm: micron (10^{-6} m)

μ_{H_2}: molecular weight per hydrogen molecule

$N_{\rm u}$: column density of the state u (cm^{-2})

$N_{\upsilon,{\rm J}}$: column density of a given H_2 transition (cm^{-2})

n: density of molecules (cm^3)

$n_{\rm u}$: density of molecules in a given state u (cm^3)

n_θ: number of azimuthal elements of the disc

$n_{\rm r}$: number of radial elements of the disc

N_*: number of stars per unit of stellar mass (M_*^{-1})

$N_{\rm tot}$: total column density (of the CO or H_2, cm^{-2})

$N(r)$: total column density of the disc at a given radius r (cm$^-2$)

$n_{\rm CO}$: total number of CO molecules

O_λ: observed flux per wavelength (ADU)

Ω: solid angle symbol (sr)

$\Omega_{\rm beam}$: solid angle of the beam (sr)

p: associated probability of the Spearman or the KS test

$P_{\rm out}$: momentum of the outflow ($\mathrm{M}_\odot\,\mathrm{km\,s^{-1}}$)

PA: position angle (degrees, from N to E direction)

pc: parsec ($3.086{\cdot}10^{18}\,\mathrm{cm}$)

PWV: precipitable water vapour (mm)

$Q(T)$: partition function of a given molecule at the temperature T

r: Pearson correlation coefficient

$r_{\rm in}$: inner radius of the disc (AU)

$r_{\rm out}$: outer radius of the disc (AU)

r_i: radial element of the disc model

$\mathrm{R}_\odot$: Solar radius ($6.96{\cdot}10^{10}\,\mathrm{cm}$)

R_*: Stellar radius ($\mathrm{R}_\odot$)

rms: root-mean square (K or NIR unit)

ρ: Spearman correlation coefficient

$\rho_{\rm p}$: partial Spearman correlation coefficient

S/N: signal-to-noise ratio

σ: same definition as rms

$\sigma_{\rm f}$: uncertainty in the flux ($\mathrm{erg\,s^{-1}\,cm^{-2}\,sr^{-1}}$)

$\sigma_{\rm fit}$: standard deviation of the residuals between the observed data and the fit

$\sigma_{\rm rms}$: uncertainty in the flux at the continuum level ($\mathrm{erg\,s^{-1}\,cm^{-2}}$)

$V_{\rm peak}$: peak velocity ($\mathrm{km\,s^{-1}}$)

$T_{\rm A}^*$: antenna temperature (K)

$T_{\rm bg}$: background temperature (K)

T_ν: brightness temperature at the frequency ν (K)

$T(r)$: disc temperature at a given radius r (K)

$t_{\rm dyn}$: dynamical timescale of the outflow (yr)

$T_{\rm eff}$: effective temperature (K)

$T_{\rm ex}$: excitation temperature (K)

$t_{\rm exp}$: exposure time (seconds)

$T_{\rm mb}$: main-beam temperature (K)

$T_{\rm sys}$: system temperature (K)

$T_{\rm u}$: temperature of a given state u (K, see also $E_{\rm u}$)

$\tau_{\rm KH}$: Kelvin-Helmholtz timescale (yr)

θ_i: azimuthal element of the disc model

τ_ν: optical depth at the frequency ν (K)

$\theta_{\rm beam}$: beam size (in $''$)

$\theta_{\rm conv}$: convolved beam size (in $''$)

$\theta_{\rm obs}$: observed beam size (in $''$)

$\theta_{\rm source}$: source size (in $''$)

$V_{\rm lsr}$: local standard rest velocity (in $\mathrm{km\,s^{-1}}$)

$V_{\rm max}$: maximum velocity of the outflow ($\mathrm{km\,s^{-1}}$)

$v(r)$: radial velocity at a given radius r ($\mathrm{km\,s^{-1}}$)

yr: year unit ($3.156{\cdot}10^7$ s)

Chapter 1

Introduction

High-mass stars (HMS, $M_* > 8\,M_\odot$ or spectral type earlier than B3 in the Main Sequence - MS) are the main responsible objects that rule the evolution of the galaxies throughout their whole life. These objects impact the interstellar medium (ISM) by emitting ionising radiation, injecting mechanical energy and momentum into the ISM through jets and outflows at early stages of evolution and, later, through their intense stellar winds. At the end of their evolution, they explode as supernovae, enhancing the chemistry of the ISM with heavy elements, and their shocks can trigger new episodes of star formation in their vicinity.

Despite their importance, the mechanism that leads to the formation of high-mass stars is still not completely understood and lacks observational constraints. Consequently, the detailed process of high-mass star formation is a significant gap in Astrophysics.

1.1 The star formation process: linking low- and high-mass protostars

The natural hypothesis is to expect that high-mass stars are formed through a similar process as their low-mass counterparts. Indeed, the formation of low-mass stars is well described by the disc-mediated accretion scenario proposed by Shu et al. (1987). In that scenario, the infalling material from the collapsing molecular cloud is funnelled towards the protostar, forming an equatorial disc. Part of the infalling material that does not reach the central source is accelerated through the magnetic field in the polar region (Pudritz et al., 2006), creating a collimated jet that extends up to scales of 10-100 AU. The jet drives an outflow, dragging out the surrounding dense gas and creating density-enhanced cavity walls (Visser et al., 2012). For low-mass protostars, the structures predicted by

the accretion scenario (see left panel of Fig. 1.1) are supported by recent high-resolution observations of low-mass young stellar objects (YSOs) at GHz frequencies by Wang et al. (2014). As illustrated in Fig. 1.1, these observations provided a detailed view and a clear separation between the high- from the low-velocity components of the ejected material: the high-velocity component (with velocities above $\sim 30\,\mathrm{km\,s^{-1}}$) probes the collimated jet while the low-velocity component (up to a few $\mathrm{km\,s^{-1}}$) traces the cavity walls between the jet and the passively heated envelope.

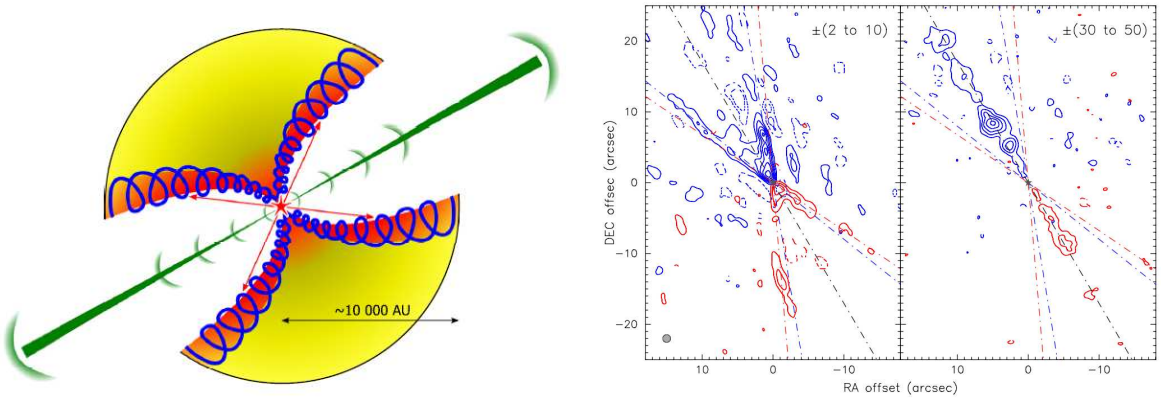


Figure 1.1: Left panel: Schematic view of the large-scale physical components in a low-mass YSO: a passively heated envelope (yellow), a hot core (extended red), a bipolar jet (green), UV-heated outflow cavity walls (red), and small-scale shocks within the cavity walls (blue). Extracted from Visser et al. (2012). Right: Integrated CO (3–2) maps of the jet driven by the low-mass YSO IRAS 04166+2706. The low-velocity emission (left panel) traces the cavity of the jet/outflow while the high-velocity emission (right) probes the collimated jet. Extracted from Wang et al. (2014).

An extensive effort was made to extend the disc-mediated accretion scenario towards HMS due to the notorious agreement between the theory and observations of the low-mass star formation process. Indeed, Palla and Stahler (1993) have demonstrated that the isotropic accretion scenario can explain the formation of stars with masses up to $\sim 6\,\mathrm{M_\odot}$. However, the intense radiative flux associated with more massive objects (the stellar luminosity increases with the mass by following an exponential law, $L_* \propto M_*^{\alpha}$, with power-law indices varying between 3.0 and 3.5) may dissipate the circumstellar disc, interrupting the accretion flow.

More recently, three-dimensional numeric simulations have found that a disc-like structure is still efficient to shield the inflowing material from the intense radiation pressure, directing the radiation towards the polar region of protostars with masses up to 30-35 $\mathrm{M_\odot}$ (Kuiper et al., 2011b,a) – the so-called *flash-light effect* (Yorke and Sonnhalter, 2002). As observed towards low-mass YSOs, the collimated radiation jet carries away the material

from the molecular envelope through outflows, playing an essential role also in the high-mass regime and allowing the accretion process to proceed up to stellar masses of $\sim 140\,M_\odot$ (Kuiper et al., 2011b, 2015).

In recent years, evidence of discs around a few B-type (proto)stars ($L_{\rm bol} \gtrsim 10^4\,L_\odot$) were found (*e.g.*, Cepheus A HW2 from Patel et al. 2005, and W33A from Davies et al. 2010). Table 1.1 summarises the recent determination of the mass of high-mass protostars through the kinematics of their disc-like structure. Despite that, there is still a lack of detections of discs around O-type objects ($L_{\rm bol} \gtrsim 10^5\,L_\odot$).

Table 1.1 - Known HMYSOs with direct determination of their stellar masses.

Source	MSX Name	$\log(L_{\rm bol}/L_\odot)$	M_* ($M_\odot$)	Ref.
Cepheus A HW2	G109.8715+02.1156	4.18	19	P05
W33A	G012.9090−00.2607	4.20	10	D10
IRAS 13481−6124	G310.0135+00.3892	3.20	18	K10
AFGL 2136	G017.6380+00.1566	4.72	20	M13
IRAS 18151−1208	G018.3412+01.7681	4.34	30	M13
S106	G076.3829−00.6210	4.60	22	M13
V645 Cyg	G094.6028−01.7966	4.63	20	M13
IRAS 20126+4104	G078.1224+03.6320	3.60	7–10	C14
AFGL 4176	G308.9176+00.1231	4.95	15	J15

Notes: The columns are: (1) Name of the source; (2) MSX name from Lumsden et al. (2013); (3) logarithm of the bolometric luminosity of the MSX source (in $L_\odot$); (4) derived mass of the protostar (in $M_\odot$). (5) References: P05 - Patel et al. (2005); D10 - Davies et al. (2010); K10 - Kraus et al. (2010); M13 - Murakawa et al. (2013); C14 - Cesaroni et al. (2014); J15 - Johnston et al. (2015).

Accordingly to Cesaroni et al. (2017), discs might be detectable towards high-mass protostars only during the intermediate stages of the pre-Main Sequence (PMS). Such hypothesis arises from the fact that evidence for actively accreting discs might be weak due to confusion with the surrounding envelope at the youngest stages of the high-mass star-forming (HMSF) process. Also, at the latest evolutionary stages of the HMSF process, most of the accretion process is already ceased and, so, the discs are likely to be already dissipated or passively evolving (Blum et al., 2004).

Another possible scenario that explains the formation of high-mass stars in denser environments is the competitive accretion, from Bonnell et al. (2001). Those authors propose that high-mass stars are formed through coalescence and merging of low-mass stars. As a consequence, the mass of the high-mass star would be regulated by the mass of the pre-existing protostars in their vicinity. Recent three-dimensional simulations of the HMSF process suggest that a combination of both scenarios – disc-mediated and competitive ac-

cretion – occurs during the formation of stars with masses greater than 20 $M_\odot$ (Krumholz et al., 2009; Krumholz, 2014).

Although the disc-mediated scenario seems to explain reasonably well the formation of high-mass stars, some theoretical problems still need to be addressed from the observational point of view. They are presented as follows.

1.2 Chasing high-mass YSOs: Observational challenges and theoretical problems

From the observational point-of-view, high-mass (proto)stars are statistically scarce mainly due to two reasons: the Initial Mass Function (IMF) and our location in the Galaxy. The IMF states that the number of stars with a given mass M_* decreases exponentially with M_*, that is, $N_* \propto M_*^{-\alpha}$, with a constant slope (*e.g.*, α=2.35, Salpeter, 1955) or changing slopes as a function of the mass (*e.g.*, α=0.3, 1.3 and 2.30 for $M_* < 0.08\,M_\odot$, $0.08 < M_* < 0.5\,M_\odot$, and $M_* < 0.5\,M_\odot$, respectively, Kroupa, 2007). Assuming the Salpeter 1955 law and integrating the IMF over the mass range between 8 $M_\odot$ (B3 V) and 150 $M_\odot$ (early O-type stars), the number of high-mass stars present in the Galaxy is about $\sim 8{\cdot}10^3$.

The second reason that hampers the identification of high-mass protostars is related to the low star formation rate (SFR) in the vicinity of the Sun. When compared to the global Galactic star formation rate ($(0.68\text{-}1.45){\cdot}10^6\,M_\odot\,\mathrm{Myr}^{-1}$, Robitaille and Whitney, 2010), the solar neighbourhood (defined as the region closer than 1 kpc from the Sun) is associated with relatively lower SFR values ($(2.0\text{-}3.0){\cdot}10^3\,M_\odot\,\mathrm{Myr}^{-1}$, Bonatto and Bica, 2011). Such low SFR suggests that no high-mass stars are likely to be formed in the solar neighbourhood. Indeed, HMSF regions are often more distant ($\gtrsim$1 kpc), making difficult the identification of individual high-mass YSO (HMYSO) candidates. In addition, the high-mass star-forming regions often exhibit high stellar density ($\lesssim 10^5\,\mathrm{pc}^{-3}$, Megeath et al. 2003). As a consequence of their relatively farther location in the Galactic plane, these regions appear as unresolved structures and require very high resolutions to avoid issues of confusion of their individual members.

A theoretical problem of the high-mass star formation process is related to the short PMS evolutionary timescale of high-mass stars when compared to the timescale of their low-mass counterparts. While low-mass stars display a relatively slow PMS evolution (*e.g.*,

a $1\,M_\odot$ star spends ∼30 million of years in the PMS phase, see Eq. (1.1)), high-mass stars are formed in a shorter timescale (about a few tens of thousands of years), exhibiting a rapid and intense PMS evolution. The time spent on the PMS phase is related to the Kelvin-Helmholtz timescale ($\tau_{\rm KH}$):

$$\tau_{KH} = \frac{GM_*^2}{R_* L_*} \approx 3.085 \cdot 10^7 \left(\frac{M_*}{M_\odot}\right)^2 \left(\frac{R_*}{R_\odot}\right)^{-1} \left(\frac{L_*}{L_\odot}\right)^{-1} \text{yr} \tag{1.1}$$

where G corresponds to the gravitational constant, M_* is the stellar mass, R_* is the stellar radii and L_* is the stellar luminosity. For example, the Kelvin-Helmholtz scale for a B0 V star (with the stellar parameters taken from Vacca et al. 1996: $M_* = 18.4\,M_\odot$, $\log(L_{\rm bol}/L_\odot) = 4.88$ and $R_* = 8.3\,R_\odot$) corresponds to $\tau_{\rm KH} \approx 16{\cdot}10^3$ yrs. Therefore, the formation of a very massive star in a small timescale requires large mass accretion rates through a process of high efficiency.

Assuming that the mass accretion rate ($\dot{M}$) is constant over the Kelvin-Helmholtz timescale, it can be expressed as:

$$\dot{M} = \frac{M_*}{\tau_{KH}} \approx 10^{-7} \left(\frac{M_*}{M_\odot}\right) \left(\frac{\tau_{KH}}{10^7\,\text{yr}}\right)^{-1} M_\odot\,\text{yr}^{-1} \tag{1.2}$$

Equation (1.2) leads to accretion rates of $\dot{M} \approx 10^{-8}\,M_\odot$ yr^{-1} for a solar-like star, and $\dot{M} \approx 10^{-3}\,M_\odot$ yr^{-1} for a B0 V star. This example indicates that the formation of high-mass stars requires a highly efficient mechanism to transfer a substantial amount of material from the parental cloud into the protostar in a short timescale.

Another theoretical problem on the HMSF process is related to the intense radiative pressure driven by the stellar winds of those objects, even at the early stages of their formation. As the high-mass protostar contracts, its temperature increases and, then, its radiative field becomes strong enough to interrupt the accreting flow. The intense radiation pressure from the protostar itself would carry away a considerable fraction of the infalling material. In addition, the stellar winds associated with high-mass stars can exhibit mechanical momentum of 10 to 100 times greater than the stellar radiative momentum (Lada, 1985; Richer et al., 2000).

Three-dimensional simulations of the HMSF process from Krumholz et al. (2009) and Kuiper et al. (2011b) suggest that the presence of a circumstellar disc around the protostar

allows the accretion to proceed even in the existence of intense radiative flux conditions. Moreover, the disc collimates the radiation towards the polar region of the protostar, creating an accretion flow in the equatorial plane. A fraction of the material in the accretion flow that cannot reach the central source is ejected through bipolar jets in the perpendicular direction of the disc, carrying away the excess of angular momentum and radiation. Such simulations indicate that the disc-mediated accretion scenario can produce stars with masses up to $\sim$140 $M_\odot$ (Kuiper et al., 2011a).

The direct observation of the "key" structures of the disc-mediated accretion scenario (*e.g.*, the discs, the outflow cavity walls and even the launching point of the jets at scales of $\sim 10^2$-10^3 AU) associated to the HMSF process is difficult because of the large distances of the sources and the consequent poor resolution for small scales structures. As an example, Fig. 1.2 compares the observations of a nearby and low-mass YSO (located at 140 pc, Burrows et al., 1996) with a distant high-mass YSO (4.9 kpc, Navarete et al., 2015), taken at similar angular resolutions. While the details of the innermost structures of the low-mass YSO (that is, the disc and the jet at scales of $\sim 10^2$ AU) are clearly detected towards the low-mass YSO (see left panel of Fig. 1.2), only the large-scale structure of the jets are accessible for more distant HMYSOs (as illustrated in the right panel of Fig. 1.2).

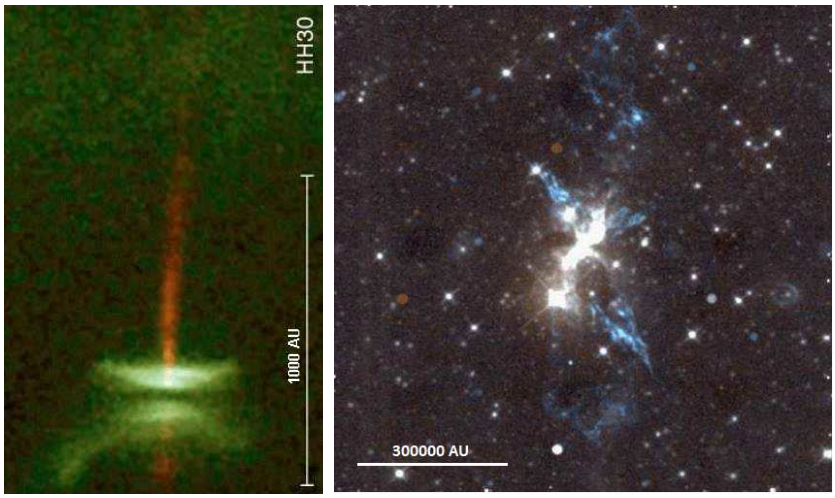


Figure 1.2: Comparison between the structures identified towards the closer and low-mass YSO HH 30 (left panel, extracted from Burrows et al. 1996) and the high-mass YSO G094.4637−00.8043 (right, extracted from Navarete et al. 2015). Due to the angular resolution of the observations and consequently different linear scales of the images (indicated by the white lines), only the large-scale structure of the jets (in blue) is accessible for high-mass (and more distant) objects, while even the structure of the disc can be identified for the (closer) low-mass YSO.

The short formation time of HMS combined with the lack of direct identification of the "key" components associated with the HMSF process highlight the difficulty in comparing

the theory and observations of the early stages of these objects. Therefore, most of the assumptions of the HMSF process relies on the identification of the large-scale structures, such as the molecular envelopes and outflows. The scaling between the large-scale properties observed towards low- and high-mass objects would indicate that similar processes are likely to be taking place at smaller (and unresolved) scales within the circumstellar region of the HMYSOs, such as the observed towards their low-mass counterparts.

1.3 The high-mass star-forming regions and their evolutionary phases

The star-forming process starts with the gravitational instability and collapse of giant molecular clouds (GMCs). Churchwell (2002) proposed an hierarchical scenario to explain the evolution sequence of the high-mass star formation based on the evolution of individual protostars within the GMCs. In the hierarchical scenario, the formation of a high-mass protostar starts with a pre-stellar core (PSC), evolving into a hot core (HC) as the gravitational collapse proceeds. The protostar reaches the high-mass young stellar object phase at the peak of the accretion process and, finally, it forms an H II region as soon as the Lyman continuum flux ($F_{\rm Ly}$) is strong enough to ionise the neutral Hydrogen in the vicinity of the protostar ($\log[F_{\rm Ly}\,({\rm s}^{-1})] \gtrsim 45.5$ dex, Urquhart et al., 2013).

Later on, Chambers et al. (2009) and Battersby et al. (2010) suggested an extension of the evolutionary sequence of the HMSF process that encompasses the formation of multiple sources, as illustrated in Fig. 1.3. The proposed evolutionary sequence is based on observational signatures towards a sample of 190 clumps within infrared dark clouds (IRDC). The sequence starts with a quiescent clump (Stage 1a) with no clear star-forming activity. In their definition, a quiescent clump has no emission in the mid-IR Spitzer maps (no 4.5 μm or 24 μm point-like source). As soon as the clump exhibits collapse signatures (a discussion on the collapse signatures in molecular line profiles is presented in Sect. 6.2 from Schneider et al. 2015), it becomes brighter in the far-IR (Stage 1b). Then, as its temperature rises, the source gets brighter in the mid-IR (*e.g.*, at 24 μm), and the active accretion phase starts (Stage 2). In Stage 2, the object exhibits a clear mid-IR point-like emission associated with extended emission at 4.5 μm, identified in the Spitzer/IRAC maps (Fazio et al., 2004). The 4.5 μm emission has a substantial contribution of several H_2 transitions tracing shocked emission and is commonly attributed to the outflow activity

within the individual protostellar members of the clump (Cyganowski et al., 2008). The propagation of the outflow within the dense circumstellar environment ignites the emission of maser lines, such as water, OH and class II methanol masers, which are often used as tracers of high-density environments and HMSF regions (Pestalozzi et al., 2005; Edris et al., 2007; Urquhart et al., 2009b). The IR-bright phase also coincides with the *high-mass young stellar object* phase (HMYSO[1], Lumsden et al. 2002). According to Lumsden et al. (2002), HMYSOs are young objects with bolometric luminosities $\gtrsim 10^4\,L_\odot$, still embedded into their parental envelopes of gas and dust, and associated with a weak Lyman continuum flux ($\log[F_{Ly}\,(s^{-1})] \lesssim 45$ dex, Urquhart et al., 2013), unable to ionise the material in their surroundings to form an H II region. The protostar starts to clean up their surroundings as soon as its continuum Lyman flux increases, appearing as a radio-loud object associated with ultra-compact H II regions (UCH II, stage 3). Then, the object evolves into a "red" clump (Stage 4), with enhanced 8 μm emission and strong cm-emission due to the expansion of the H II shell. The final evolutionary stage is characterised by the formation of a young cluster.

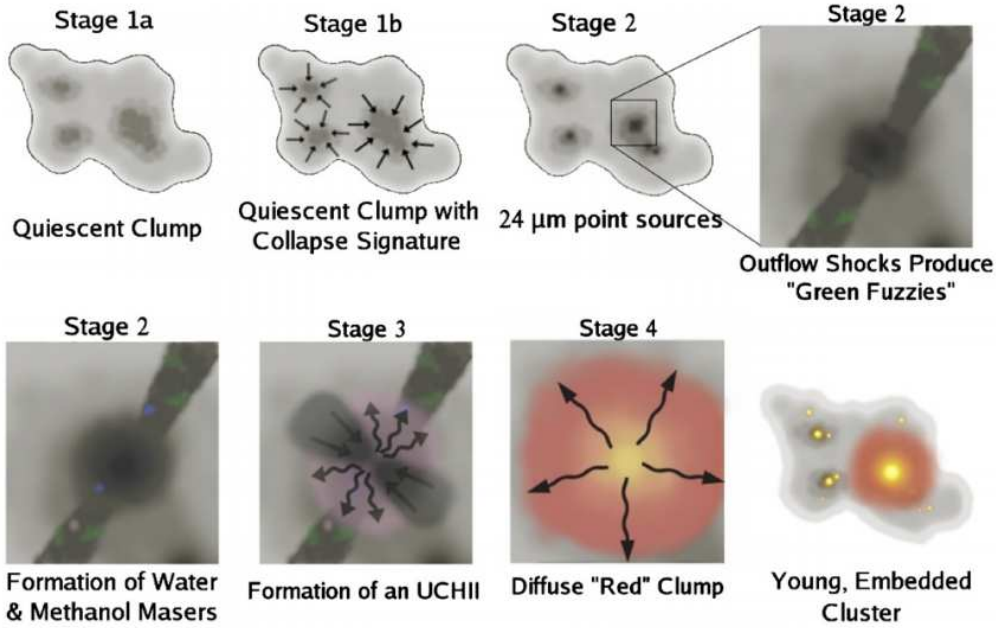


Figure 1.3: Schematic evolution sequence of the high-mass star formation (see details in the text). Extracted from Battersby et al. (2010).

The classification illustrated in Fig. 1.3 relies on empirical features observed towards the high-mass IRDCs before 2010's. In the last years, the advance of terahertz (THz) and sub-millimetre (sub-mm) Astronomy offered a new window for investigating the details of the

[1] the authors adopt the *massive young stellar object* (MYSO) nomenclature.

PMS evolution of the HMYSOs and their surroundings. The combination between ground-based (*e.g.*, SMA, APEX and, more recently, ALMA) and stratospheric observations (*e.g.*, SOFIA) provided an unprecedented view of the structure and chemistry of the first stages of the high-mass star formation process. Indeed, a more recent classification of high-mass clumps, selected from the 870 μm ATLASGAL survey (Schuller et al., 2009), has defined four main phases of the pre-MS evolution of these objects: *i*) the 70 μm-weak clumps, *ii*) the mid-IR weak clumps, *iii*) the mid-IR bright clumps and *iv*) the more evolved clumps associated with H II regions. These classes were defined by König et al. (2017) and are illustrated in Fig. 1.4.

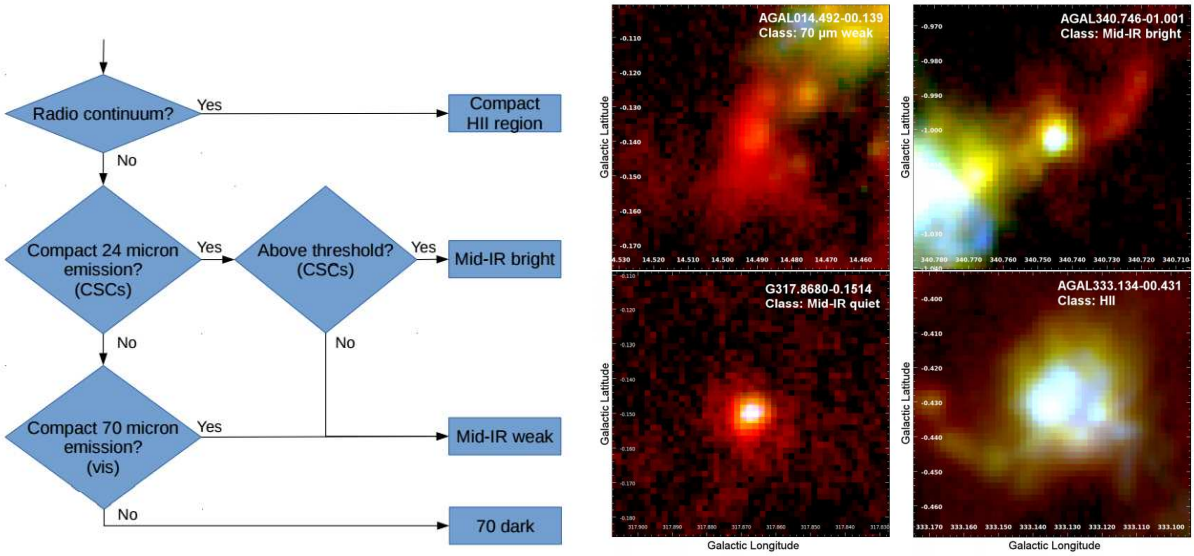


Figure 1.4: Left panel: Classification of ATLASGAL clumps based on their infrared to centimetre continuum emission. Right: False-colour RGB maps of ATLASGAL clumps at different evolutionary classes from the youngest (top left) to the most evolved (bottom right). In each map, the channels are: the *Herschel*/PACS 70 μm map (blue), the *Herschel*/PACS 160 μm (green); and the ATLASGAL 870 μm continuum emission (red). Extracted from König et al. (2017).

1.4 The importance of jets/outflows

The protostellar jets and outflows are elongated structures with linear scales of $\sim 10^4$ AU (Bally et al., 2007). Compared to the typical sizes of the discs associated with HMYSOs ($\lesssim 10^3$ AU), the jets have linear sizes extending up to a few parsecs ($\sim 10^5$ AU, Varricatt et al., 2010a), making them the ideal probes of star-forming activity across the whole Galaxy.

The protostellar jets are often identified through narrow-band imaging of forbidden

lines (*e.g.*, [Fe II] at 1.64 μm, Reiter et al. 2016), atomic lines (*e.g.*, Hα at 656.28 nm, Bally et al. 2007) and molecular lines (e.g., H_2 at 2.12 μm, Caratti o Garatti et al., 2006; Varricatt et al., 2010a). In addition, sources driving molecular jets and outflows are often associated with an excess of emission at 4.5 μm in the IRAC/Spitzer maps. The 4.5 μm emission has an important contribution from three transitions of the molecular hydrogen at that wavelength range (H_2 (0–0) S(9), (0–0) S(10) and (0–0) S(11) at 4.695, 4.410 and 4.181 μm, respectively). Indeed, most of the so-called "Extremely Green Objects" (EGOs[2]) from the GLIMPSE catalogue (Cyganowski et al., 2008) are known YSOs driving molecular outflows.

The protostellar jets are produced by the ejection of material from the central object. Their propagation process seems to be related to the mass of the central object. In the case of low-mass stars with intense magnetic activity, the infalling material that is not accreted onto the protostar is carried away from the surface of the accretion disc through the magnetic field lines, forming high-collimated jets. In such cases, the jets are expected to exhibit rotation, carrying away the excess of angular momentum of the accretion disc, allowing the material to accrete onto the central protostar. However, a few protostars associated with rotating jets were detected yet (*e.g.*, HH 26 and HH 72 from Chrysostomou et al. 2008; HH 212 from Lee et al. 2017).

The magnetic effects are weaker in high-mass protostars, and the intensity of the magnetic field is insufficient to collimate the jets as observed towards low-mass protostars. Therefore, the jets associated with HMYSOs are expected to be less collimated than the low-mass counterparts Bachiller (1996). The launching mechanism of these less collimated structures is still on debate.

1.5 Formation and evolution of the bipolar jets

The classical theory of the formation of low mass stars by Shu et al. (1987) also discusses the origin and formation of the jets driven by these objects. Later on, Goodson et al. (1997) presented a model based on numeric simulations of the formation of protostellar jets. Those authors suggest that the ejection of material as bipolar jets is due to the

[2] by following the definition of Cyganowski et al. (2008), EGOs are bright and extended objects in the green channel of the false-colour RGB IRAC maps, where the blue, green and red channels correspond to the maps at 3.6, 4.5 and 8.0 μm, respectively.

interaction between the stellar magnetosphere and the circumstellar disc. Goodson et al. (1999) identified two different components in the bipolar structures: a fast, high-collimated and less massive inner component (the "jet" itself), and a slower, wider and more massive external component (the "disc wind" component). These two components from Goodson et al. (1999) model are illustrated in Fig. 1.5. For low-mass YSOs, the *Hubble Space Telescope* (HST) observations of DG Tau from (Bacciotti et al., 2000) corroborate the theoretical predictions from Goodson et al. (1997) and indicates that the jet propagates with velocities up to 200 km s^{-1} and its innermost region is associated with densities larger than 10^4 cm^{-3}.

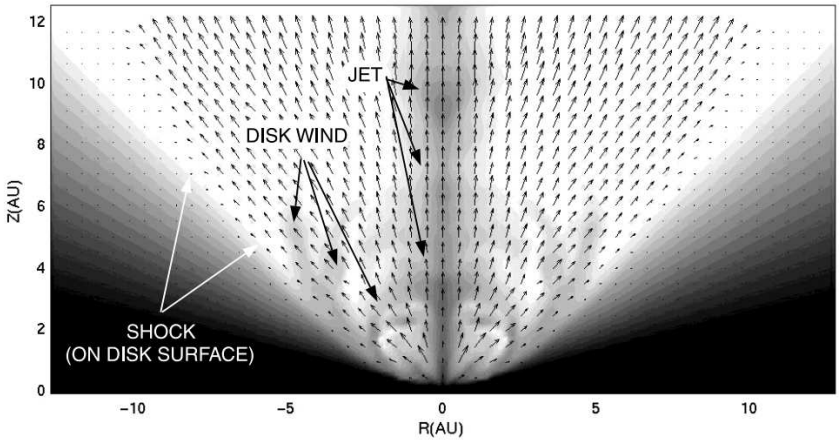


Figure 1.5: Edge-on view of the velocity profile of a simulated outflow driven by a young stellar object. The arrows indicate the direction and intensity of the velocity. The grey scale presents the density profile. Extracted from Goodson et al. (1999).

1.6 Motivation and Goals

This thesis presents the study of the large-scale properties of HMSF clumps at sub-millimetre wavelengths and the details of the circumstellar environment associated with a sample of individual high-mass YSOs in the near-infrared.

The study of the large-scale CO emission and the outflow content in a selected sample of 99 clumps, representative of the high-mass clump population in the whole Galaxy, will answer a few questions on the scaling between the properties of the gas from low- to high-mass regimes.

First, the comparison between the integrated properties of the CO and the physical properties of the clumps (that is, their bolometric luminosity or mass) will evidence whether the large-scale distribution and properties of the gas are correlated (or not) with the properties of their hosting pre- or protostars in a wide range of pre-main sequence evolutionary stages. Then, the comparison between the CO emission of low- and high-mass clumps will show if the large-scale CO gas is scaled or not over a broader range of luminosity (and mass) regimes.

The CO data will also provide an unprecedented study on the characterisation of the outflow content of these clumps. The data will reveal how common is the outflow phenomena towards high-mass clumps and if the outflow activity is more frequent in a given evolutionary stage of their driving sources. The derivation of the physical properties of the outflows (*e.g.*, their masses, mechanical force and luminosity) and further comparison with the clump properties will show whether they are scaled or not. In addition, the comparison between the properties of the outflows in a wider range of luminosity (and mass) regimes will help to answer if the high-mass outflows are scaled up versions of their low-mass counterparts.

In the near-infrared, three-dimensional spectroscopic observations at high-angular resolution provide the ideal instrumental setup to investigate the circumstellar environment of relatively close and high-luminosity objects (*e.g.*, Davies et al., 2010; Murakawa et al., 2013).

Seven HMYSO candidates associated with large-scale H_2 jets from Navarete et al. (2015) were selected to follow-up three-dimensional spectroscopic observations at the Gemini North telescope to characterise their circumstellar environment at linear scales of $\sim 10^2$-10^3 AU. The identification and analysis of the discrete spectral features (in particular, the H_2 transitions) are expected to probe the H_2 jets at relatively closer distances from their launching points. In addition, the identification of CO features towards the central sources are expected to be originated from circumstellar discs. The study of the kinematics of the disc can provide an estimate for the Keplerian mass of the HMYSOs.

These data will help to answer if the HMYSO candidates are really high-mass objects. The morphology and physical parameters of the H_2 jets closer to their launching points will be compared to their large-scale counterpart to answer whether the jets are well-collimated

at their launching point, such as observed towards low-mass YSOs, or if they are wide-angle wind-disc driven outflows.

Ultimately, one of the aims of this thesis is to considerably increase the number of HMYSOs associated with circumstellar discs to decide whether high-mass stars are formed through a similar process as their low-mass counterparts.

1.7 Thesis outline

The basic concepts of the high-mass star formation process were introduced in this Chapter, together with the current status of the research on this field.

Chapter 2 describes the ATLASGAL and RMS surveys, used for selecting the sources analysed in this volume.

Chapter 3 presents the characterisation of the large-scale CO content of a sample of high-mass clumps in the inner Galaxy from the ATLASGAL survey (TOP100) and the large-scale CO outflow content towards the TOP100 sample.

Chapter 4 presents the analysis of the circumstellar environment of eight HMYSOs exhibiting large-scale H_2 emission and the derivation of the parameters of the jets and their discs.

Finally, the most important conclusions of this thesis are outlined in Chapter 5, together with the perspectives of the projects planned for the future.

Chapter 2

Infrared and sub-millimetre Galactic surveys of high-mass clumps and HMYSOs

In the last decade, the RMS (Lumsden et al., 2013) and the ATLASGAL (Schuller et al., 2009) surveys were designed to identify a complete and unbiased list of high-mass protostellar objects and massive proto-clusters distributed along the Galactic plane. The RMS and ATLASGAL are complementary surveys and provides a ubiquitous view of the galactic plane at wavelengths ranging from the infrared (from $\sim$10 μm) to the sub-millimetre (870 μm). A concise description of both surveys regarding their importance, their observational strategy and synergies is provided by Urquhart et al. (2014). For completeness, a brief explanation of both surveys is presented as follows.

2.1 The ATLASGAL survey

In the sub-millimeter range, a considerable fraction of the continuum emission in the Galactic plane arises from the dust which is one of the best tracers to identify dense star-forming regions at their earliest stages (Chandler and Sargent, 1997).

Aware of the correlation between star-forming regions and the dust emission, a systematic and large-scale survey of the inner region of the Galactic plane was designed to search for the densest regions in the interstellar medium and possible hosts of high-mass protostellar objects: the APEX Telescope Large Area Survey of the Galaxy (ATLASGAL). The observations of the ATLASGAL survey were carried at the *Atacama Pathfinder EXperiment* (APEX) telescope, using the *Large APEX Bolometer Camera* (LABOCA Siringo et al., 2009), an array of 295 bolometers observing at 870 μm (345 GHz).

The ATLASGAL comprises observation of the inner $\sim$95 deg^2 of the Galactic Plane

($\ell \leq \pm 60°$, $b \leq \pm 1.5°$) at 870 μm with a beam size of 19.″2 and a typical noise level of 5070 mJy/beam. Figure 2.1 illustrates a typical region of the Galactic plane covered by the ATLASGAL observations.

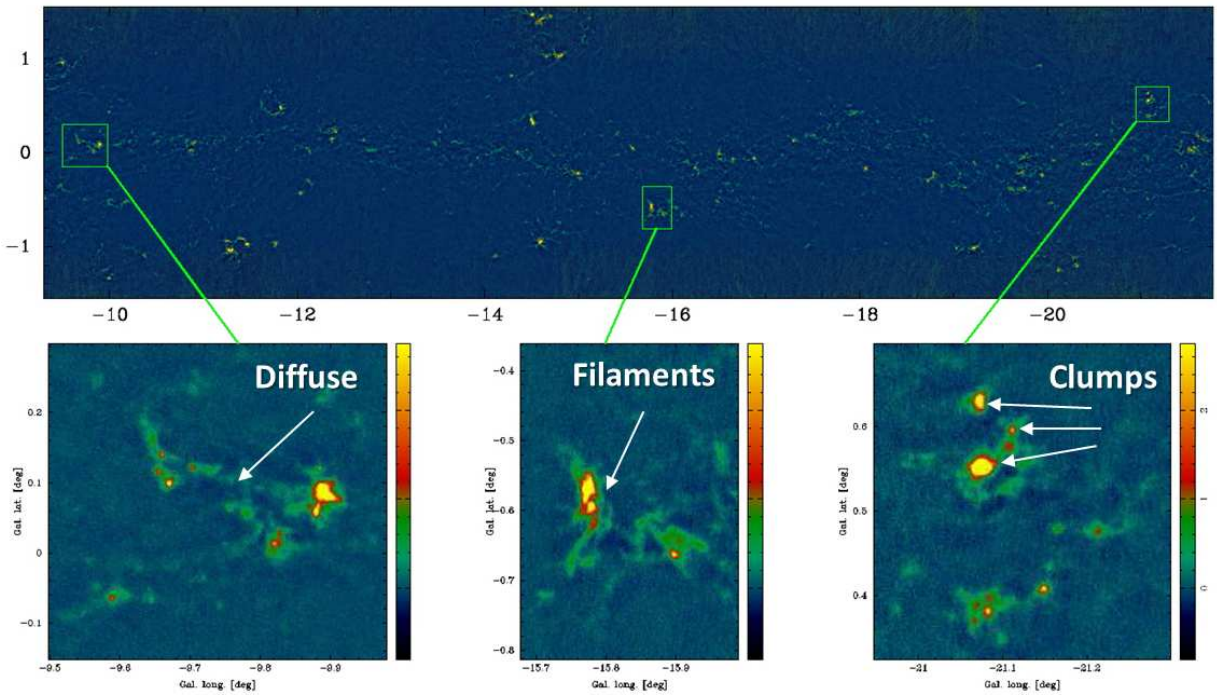


Figure 2.1: Large-scale view of the ATLASGAL 870 μm emission map in the Galactic plane between the Galactic longitudes $-10°$ and $-21°$ and zooms into selected regions with distinct morphologies of the sub-mm emission: diffuse (left inset), filamentary (middle) and compact structures (right). Adapted from Schuller et al. (2009).

By analysing the whole area covered by the ATLASGAL survey, Schuller et al. (2009) have found ∼6000 compact sources brighter than 0.25 Jy plus a variety of diffuse and filamentary structures. Those authors also found that most of the compact sources are clumps with masses ranging from a few 10 to a few 100 $M_\odot$ that might harbour high-mass protostellar objects and proto-clusters in a broad range of evolutionary phases.

The left panel of Fig. 2.2 presents the two-dimensional distribution of the ATLASGAL clumps on a face-on view of the Galactic plane, indicating a strong correlation between the position of the clumps and the position of the spiral arms.

2.2 The Red MSX Source survey

Between 1996 and 1997, the Midcourse Space eXperiment (MSX, Price et al. 2001) satellite observed the whole Galactic plane ($|b| \leq 5°$) in four mid-infrared bands centred at

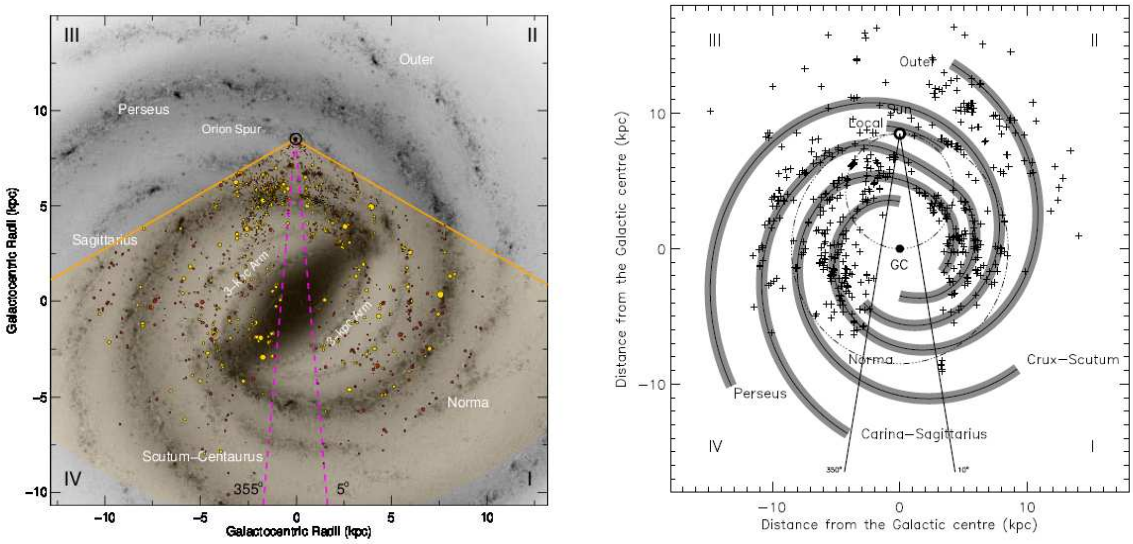


Figure 2.2: Galactic distribution of the ATLASGAL clumps (left panel) and the RMS sources classified as either YSO or HII (right). The position of the Galactic spiral arms is presented by the grey-scale image or the grey contours, using comparable Galactic models (mode information is available on the original works). The numbers in the corners refer to the Galactic quadrants and the lines (purple and orange in the left, black in the right) originate from the location of the Sun and encloses the regions excluded from the surveys. Extracted from Urquhart et al. (2008) and Urquhart et al. (2018).

8.28 μm, 12.13 μm, 14.65 μm and 21.34 μm, at a angular resolution of 18.″3 in all bands. The MSX observations were used as the starting point for the compilation of the Red MSX Source survey, presented as follows.

The Red MSX Source survey (RMS Lumsden et al., 2013) presents a refined catalogue of high-mass YSO and ultra-compact H II region candidates, selected through a photometric criteria based in near- (NIR) and mid-infrared (MIR) observations. This section presents a summary of the selection criteria adopted by the RMS survey, and the additional follow-up multi-wavelength observations, from the NIR to the centimetre window, to further characterise and derive the properties of the individual RMS sources.

Initially, Lumsden et al. (2002) have selected a sample of ≈5000 objects from the MSX point-like sources catalogue (Egan et al., 1999, 2003) based on the following infrared photometric criteria:

- $F_{8\mu m} < F_{14\mu m} < F_{21\mu m}$;
- $F_{21\mu m}/F_{8\mu m} > 2$;
- $F_{8\mu m}/F_{K} > 5$;
- $F_{K}/F_{J} > 2$.

where the fluxes at 8 μm, 14 μm and 21 μm were extracted from the MSX catalogue and the J and K-band fluxes were converted from the 2MASS catalogue (Skrutskie et al., 2006). The cross-match between objects in different catalogues was performed by assuming a maximum positional error of 10″. The relatively large positional error allows that a neighbour but separate near-infrared counterpart is attributed to a near-infrared faint MSX source. In according to Lumsden et al. (2002), a wrong association between the 2MASS and the MSX catalogues is found in less than one percent ($\lesssim$1%) of the sources. In addition, the inclusion of the NIR information on the adopted photometric criteria was necessary to exclude about two thirds of planetary nebulae (PNe) and evolved stars (*e.g.*, asymptotic giant branch stars – AGB) contained in the initial sample that were not filtered using only the MIR photometric criteria.

The infrared colours of the HMYSOs are similar to those of low-mass YSOs, ultra-compact H II regions and evolved stars (AGB and PNe) also present in the sample. Additional follow-up observations were performed to *i*) further identify contaminations and refine the sample of HMYSO candidates, and to *ii*) determine the physical properties of the objects.

The bolometric fluxes of the RMS sources were derived by Mottram et al. (2011) by combining the fluxes of the MSX and 2MASS bands with additional mid-IR observations at 10 μm (Mottram et al., 2007), far-IR (Mottram et al., 2010), 6 cm observations (Urquhart et al., 2007, 2009a) and sub-mm observations at 450 and 850 μm (Di Francesco et al., 2008). The kinematic distances of ≈1230 RMS sources (591 and 638 for the Southern and Northern hemispheres, respectively) were derived from ^{13}CO observations from Urquhart et al. (2007, 2008). The distances for the other RMS sources were derived from maser parallax or spectrophotometric analysis, and from other works in the literature (more information are provided in Sect. 2.4 from Mottram et al. 2011).

The mentioned follow-up observations, combined with additional molecular line and maser observations (*e.g.*, NH_3 and H_2O from Urquhart et al. 2009b and Urquhart et al. 2011) were used to further classify the RMS sources, separating the HMYSO candidates from more evolved objects presented in the initial catalogue. The final classification of the ∼5000 RMS sources contains: 803 YSO candidates (YSO), 59 YSOs associated with H II regions (YSO/H II), 900 YSOs driving H II regions (H II), 872 evolved stars (AGB or PNe), 1590 rejections and ≈270 RMS sources with doubtful classifications. The right

panel of Fig. 2.2 presents the distribution of the RMS sources classified as YSO and HII in the Galactic plane. Compared with the distribution of the ATLASGAL clumps, the RMS provides a complementary infrared view of their content at similar angular resolutions. Besides, the RMS is the only survey with reasonable coverage of the anti-centre location of the Galaxy.

A preliminary version of the final RMS catalogue was presented by Mottram et al. (2011). Those authors identified a list of ∼400 HMYSO candidates (YSO and YSO/H II objects), from which 353 well-classified YSO sources were selected to follow-up observations in the H_2 transition at 2.12 μm, presented by Navarete et al. (2015).

Chapter 3

Single-dish sub-millimetre observations of high-mass clumps in the inner Galaxy

This Chapter presents the study of the ^{12}CO emission of a sample of 99 clumps, selected from the ATLASGAL survey and discussed in terms of its general importance and the highlight results from Navarete et al. (submitted), presented in Appendix C.1. Every mention to particular figures or tables from that work are ***emphasized*** (*e.g.*, "see the maps presented in ***Fig. B.1***").

This Chapter is structured as follows. The motivations of this work are presented in Section 3.1. Section 3.2 describes the selection criteria of the ATLASGAL TOP100 sample. The observations are described in Sect. 3.3. The data processing is presented in Sect. 3.4. Section 3.5 presents the analysis of the APEX data, while the results are presented and discussed in Sect. 3.6. The conclusions of this Chapter are presented together with those from Chapter 4 in Chapter 5.

3.1 Motivation

This Chapter focuses on the characterisation of a sample of 99 clumps (TOP100), selected from the ATLASGAL survey and representative of the Galactic proto-clump population, in terms of their ^{12}CO emission through the observation of three rotational J transitions: CO (4–3), CO (6–5) and CO (7–6). These transitions are easily excited at relatively low temperatures (≤ 155 K) and are powerful probes of the densest layers of the star-forming regions (Bayet et al., 2009).

Besides their importance as tracers of high-density environments, they are accessible from vast range of ground-based millimeter and sub-millimeter facilities, *e.g.*, the *Atacama*

Pathfinder EXperiment (APEX) telescope, the *Atacama Large millimeter Array* (ALMA) and, in the future, the *Large Latin American Millimeter Array* (LLAMA) in the Southern hemisphere; the *Sub-millimeter Array* (SMA) and the NOEMA interferometer (the extended version of the *Plateau de Bure* interferometer) in the Northern hemisphere.

3.2 *The ATLASGAL* TOP100 *sample*

The TOP100 is a flux-limited sample of clumps selected from the ATLASGAL survey, through additional photometric criteria based on available data from the literature: the 24 μm maps from the MIPSGAL survey (Carey et al., 2009), the 70 μm maps from the Hi-GAL survey (Molinari et al., 2010) and radio continuum emission information from the CORNISH (Purcell et al., 2013) or the RMS surveys (Lumsden et al., 2013).

The TOP100 sample was initially defined by Giannetti et al. (2014) and, later, refined by König et al. (2017). The sample is divided into four evolutionary classes, defined in terms of the infrared and radio continuum emission, as:

- 70w: 14 clumps that display weak, diffuse emission or are not associated with any point-like emission in the Hi-GAL 70 μm images;
- IRw: 31 clumps that display point-like emission weaker than 2.6 Jy in the MIPSGAL 24 μm images;
- IRb: 33 clumps that are bright at 24 μm images, but are still not associated with radio continuum emission;
- H II: 21 clumps that are bright in the IR and are also associated with radio continuum emission due to the presence of an H II region.

The physical properties of the TOP100 sample were analysed in terms of distance, mass and luminosity by König et al. (2017). Those authors showed that the TOP100 sample represents a statistically significant catalogue of high-mass star-forming clumps covering a broad range of evolutionary phases, from the earliest quiescent 70 μm-weak clumps to the most evolved clumps hosting H II regions, with no bias in terms of distance, luminosity and mass among each class. The bolometric luminosities and masses of the TOP100 clumps range from $\sim$60 to 3.6×10^6 $L_\odot$ and $\sim$20 to 5.2×10^5 $M_\odot$, respectively. The distance of the

clumps ranges between 0.86 and 12.6 kpc, indicating that observations of the TOP100 at the same angular resolution sample different linear scales.

3.3 APEX Observations

This section describes the single-pointing CO (4–3) observations and on-the-fly mapping of the CO (6–5) and CO (7–6) emission towards the TOP100 clumps, observed with two heterodyne receivers installed on the 12-m APEX telescope.

The *First Light APEX Submillimeter Heterodyne receiver* (FLASH$^+$, Klein et al., 2014) is a dual-frequency receiver that operates in the frequency range between 262 and 516 GHz. The FLASH$^+$ instrument was used to observe the TOP100 clumps in the CO (4–3) transition (461.05 GHz) in the following dates of 2011: June 15 and 24, August 11 and 12. The beam size of the antenna ($\theta_{\rm beam}$) corresponds to 13.″4 at the CO (4–3) frequency. Details of the instrument setup and the properties of the CO (4–3) and the mid-J CO transitions are summarised in Table 3.1. The single-pointing CO (4–3) spectra were processed by using the same procedure adopted for calibrating the individual scans of the mid-J CO observations, presented as follows.

Table 3.1 - Instrumental setup of the APEX observations.

Transition	CO (4–3)	CO (6–5)	CO (7–6)
$E_{\rm u}$ (K)	55	116	155
Frequency (GHz)	461.04	691.47	806.65
Instrument	FLASH$^+$	CHAMP$^+$	CHAMP$^+$
$B_{\rm eff}$ [a]	0.60	0.41	0.34
$F_{\rm eff}$ [a]	0.95	0.95	0.95
$\theta_{\rm beam}$ ($''$)	13.4	9.0	7.7
ΔV (km s^{-1})	0.95	0.32	0.27
$T_{\rm sys}$ (10^3 K)	1.4 ± 0.8	1.3 ± 0.3	5.0 ± 1.5
rms (K)	0.4 ± 0.2	0.2 ± 0.2	0.8 ± 0.3
Map size	–	$80'' \times 80''$	$80'' \times 80''$

Notes: The lines are: (1) observed transition; (2) upper-level energy of the line; (3) rest frequency; (4) instrument; (5) main-beam efficiency; (6) forward efficiency; (7) beam size at the rest frequency; (8) spectral resolution of the original data; (9) mean systemic temperature of the observations; (10) mean rms of the spectra convolved to 13.″4 (see details in Sect. 3.5.1); (11) size of the maps. [a] extracted from the MPIfR website: http://www3.mpifr-bonn.mpg.de/div/submmtech/

The *Carbon Heterodyne Array of the MPIfR* heterodyne receiver (CHAMP$^+$, Kasemann et al., 2006; Güsten et al., 2008) was used to map the TOP100 sources simultaneously in the CO (6–5) and CO (7–6) transitions at 691.47 and 806.65 GHz, respectively. The beam

sizes ($\theta_{\rm beam}$) are about 9.″0 and 7.″7 at the given frequencies. As illustrated in Fig. 3.1, the CHAMP$^+$ consists in a 2×7 pixels that operate simultaneously in the radio frequency tuning ranges 620-720 GHz in the low-frequency array (LFA) and 780-950 GHz in the high-frequency array (HFA), respectively.

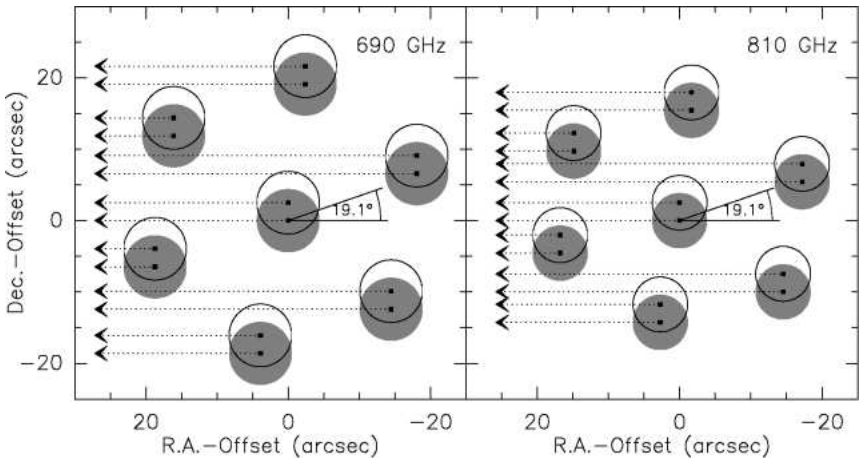


Figure 3.1: The footprints of the two CHAMP$^+$ sub-arrays, based on the positions of the pixels as actually measured on the sky. The position of each pixel is indicated by the filled grey and opened black circles, centred on the black dots. Two successive scans are indicated by filled grey and the opened black circles. The horizontally dashed arrows indicate the scanning direction in the plane of the sky. The beam sizes are about 9.″0 and 7.″7 at 691 and 806 GHz, respectively. For on-the-fly observations, the instrument is rotated by 19.1°, resulting in maps sampled at $0.7 \times \theta_{\rm beam}$ (see text for details). Extracted from the MPIfR website.

The APEX data were obtained by following a typical observing run. Each observing block consists of pointing and focus calibrations (repeated once per hour), followed by a sequence of antenna temperature calibration using a cold and a hot source (repeated every ∼20 minutes). Then, the science observations are performed. The pointing and focus were performed by observing a planet (Mars or Saturn) or nearby hot-cores: G10.47+0.03 B1, G34.26, G327.3−0.6 and NGC6334I. The same sources were used as flux calibrators. The temperature calibration of the instrument setup is relatively constant for intervals of ∼20 minutes, which defines a limit of two science targets per observing block using the same calibration conditions.

The CHAMP$^+$ observations were performed in continuous on-the-fly (OTF) mode and maps with a field-of-view (FOV) of about $80'' \times 80''$ size were obtained for each source. The OTF mapping procedure performs a continuous scan of the sky along one direction in a rectangular map while recording out the data after a given time interval. Ideally, the spacings along the scanning axis should not exceed about one-third of the beam to avoid beam broadening. A schematic view of the CHAMP$^+$ footprints is presented in Fig. 3.1.

For OTF observations, the array is rotated by 19.°1, resulting in maps sampled at the resolution of $0.7 \times \theta_{\rm beam}$.

For OTF observations, the CHAMP$^+$ array is rotated by 19.°1 to increase the resolution of the final maps. This strategy is sufficient for obtaining final data sampled at angular resolution of $0.7 \times \theta_{\rm beam}$. Then, the OTF mapping was performed by moving the antenna over a square region of $36'' \times 36''$ around the central position of each source. The pointings were spaced by $3''$ (corresponding to one-third of the CO (6–5) beam size) along the scanning axis (*e.g.*, along the right ascension direction, as indicated by the dashed horizontal arrows in Fig. 3.1) and, then, the antenna is offset by the same spacing in the perpendicular direction. The data was observed using an integration time of 1.5 seconds per pointing.

The OTF scanning was executed in two steps. First, the scan was performed by moving the antenna and integrating the observations along the x-direction (corresponding to the right ascension direction) and, then, the same procedure was performed along the y-direction (corresponding to the declination direction). The individual scans are recorded in a single FITS file, together with a header containing, among other observational information, the corresponding offset from the central position and the identification of the instrumental pixel. This information is required for regridding the data into a final datacube, as explained as follows.

The total integration time was ≈ 500 seconds (≈ 8.5 minutes) per source. The sky subtraction was performed by observing a blank sky field, offset by $600''$ from the central positions of the sources. The average precipitable water vapour (PWV) of the observations ranged between 0.28 and 0.68 mm per day, with a median value of 0.50 mm. The median system temperatures ($T_{\rm sys}$) of the observations are 1300 ± 250 and 5000 ± 1500 K, at 691 and 807 GHz, respectively. The CO (4–3) observations were performed with similar observing conditions (see Table 3.1).

3.4 APEX data processing

The individual spectra observed with the FLASH$^+$ and CHAMP$^+$ receivers were processed using the *"Continuum and Line Analysis Single Dish Software"* (`CLASS`), which is part of the `GILDAS` software[1]. The spectra were rest-frequency corrected and binned to a

[1] http://www.iram.fr/IRAMFR/GILDAS

final spectral resolution of 2.0 km s^{-1} to smooth any noisy fluctuations in the data. The baseline subtraction was performed using a first-order fit to the channels out of the main CO emission, in a window centred at $V_{\rm lsr}$ ± 100 km s^{-1} wide. A broader window was used for sources displaying CO emission broader than 80 km s^{-1} (see ***Table A.6***).

Antenna temperatures ($T_{\rm A}^{*}$) were converted to main-beam temperatures ($T_{\rm mb}$) using Eq. (3.1) (see Bensch et al. 2001), assuming beam efficiencies ($B_{\rm eff}$) and forward efficiencies reported in Table 3.1.

$$T_{\rm mb} = T_{\rm A}^{*} \frac{B_{\rm eff}}{F_{\rm eff}} \tag{3.1}$$

Finally, the CHAMP^{+} datacubes were constructed by convolving the gridded data with a Gaussian beam of one-third of the beam telescope size, using the gridding routine `XY_MAP` in `CLASS`. The procedure leads to a slightly enlarged final angular resolution (9.″6 and 8.″2 for CO (6–5) and CO (7–6), respectively) than the original beam size.

The uncertainty on the integrated flux was estimated based on the observed flux of the hot cores used as calibrators. Figure 4.5 presents the ratio between the daily integrated CO (6–5) flux of the flux calibrators with their average value as a function of the observing conditions and instrumental parameters: the days and UT time of the observations, the system temperature and the elevation angle of the telescope. The dispersion of the points is consistent with a ±20% limit and, thus, this value was adopted as the upper limit of the flux uncertainty for all the CO transitions. Two observing runs were associated with more significant deviances on the integrated flux of their calibrators and, therefore, a larger uncertainty was adopted for the sources observed immediately after the hot-core G327.3−0.6. Thus, a 30% error was adopted for the clumps AGAL320.881−00.397, AGAL326.661+00.519 and AGAL327.119+00.509, and a 50% uncertainty for sources AGAL329.066−00.307 and AGAL342.484+00.182. The upper limit value of 20% was also adopted for the error on the integrated CO (4–3) and CO (7–6) fluxes.

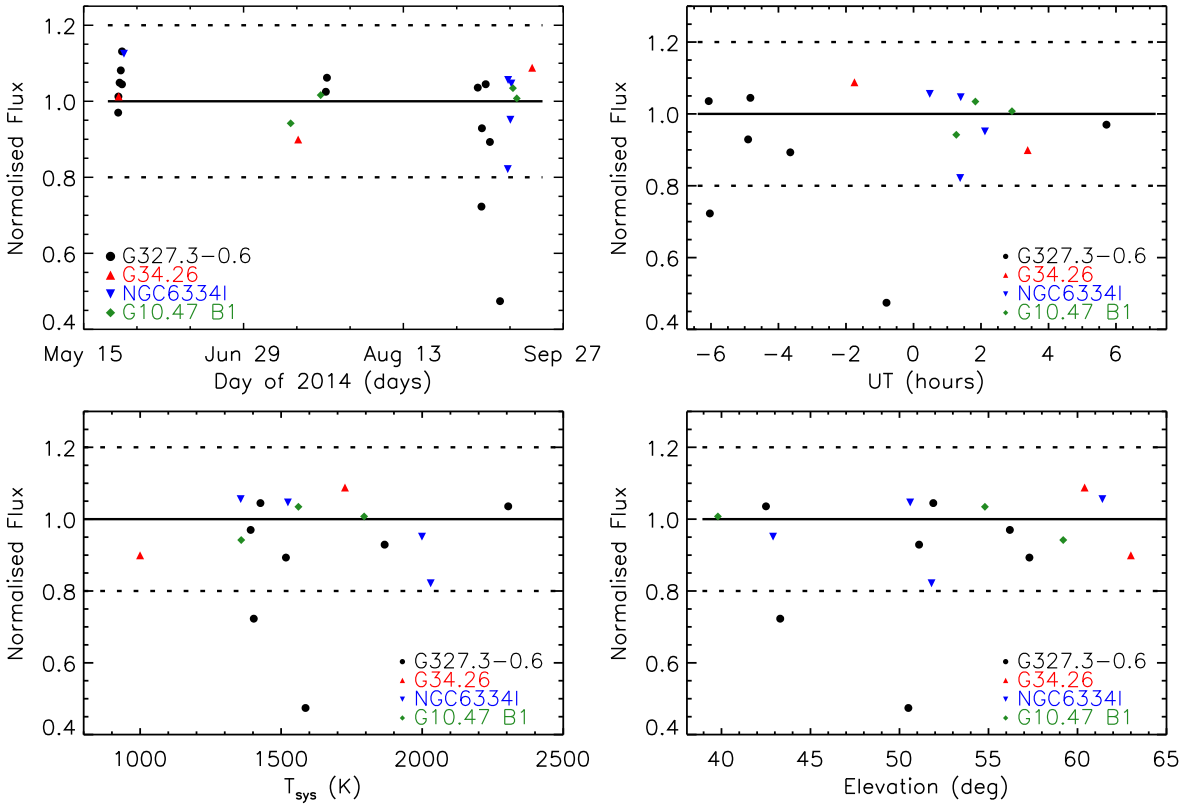


Figure 3.2: The ratio between the daily integrated flux and the average flux of the CO (6–5) transition for the four calibration sources observed during the observational campaign. The ratio is presented as a function of the days (top left panel), the universal time of the observations (top right), the system temperature (bottom left) and the elevation of the telescope (bottom right). A solid horizontal line is placed at 1.0 and the dashed lines indicate a deviation of ±20% from the unity.

3.5 Analysis of the CO spectra and maps

3.5.1 The convolution of the mid-J CO data

By taking advantage of the spatial information available on the CHAMP$^+$ datacubes, the mid-J CO spectra can be extracted at any angular resolution larger than the original beam size (see Table 3.1). Indeed, the mid-J CO lines were convolved to the same angular resolution of the single-pointing CO (4–3) observations (13.″4). The convolution of the mid-J CO data was performed by assuming a Gaussian beam centred at the coordinates given in Table A.1. The convolution method is explained as follows.

The following convention was adopted for describing the spatial and spectral elements of the datacubes and the quantities measured or evaluated directly from them. The main-beam temperature of each element of the datacube is defined as $T_{ij}(v)$, where the indices i and j indicate the position of the spatial pixel (spaxels) in the x- and y-direction (cor-

responding to the right ascension and declination axis, respectively), and v is the corresponding velocity channel.

In case of Gaussian beams and following Eq. 4.29 from Marr et al. (2015), the observed beam ($\theta_{\rm obs}$) is the result of a convolution between the original beam ($\theta_{\rm beam}$) and the source size ($\theta_{\rm source}$):

$$\theta^2_{\rm obs} = \theta^2_{\rm beam} + \theta^2_{\rm source} \tag{3.2}$$

According to Eq. (3.2), the convolution beam ($\theta_{\rm conv}$) corresponds to an arbitrary source size ($\theta_{\rm conv} = \theta_{\rm source}$), expressed in terms of the beam size (9.″6 and 8.″2 for CO (6–5) and CO (7–6), respectively) and the final angular resolution of the data, $\theta_{\rm obs}$:

$$\theta^2_{\rm conv} = \theta^2_{\rm obs} - \theta^2_{\rm beam} \tag{3.3}$$

The area of the convolution beam, A_θ, is defined as:

$$A_\theta = \sum_{ij} exp\left[-4\ln(2)\frac{O^2_{ij}}{\theta^2_{\rm conv}}\right] \tag{3.4}$$

where $O^2_{ij} = (i - x_0)^2 + (j - y_0)^2$ corresponds to the square-offset of the spaxel ij from the central position of the beam, (x_0,y_0), expressed in the same units as $\theta_{\rm conv}$. The flux of each spaxel in the datacube is multiplied by a weight-factor (w_{ij}) before computing the final spectrum. The weight w_{ij} is defined as:

$$w_{ij} = \frac{exp\left[-4\ln(2)\frac{O^2_{ij}}{\theta^2_{\rm conv}}\right]}{\frac{1}{N}\sum_{ij} exp\left[-4\ln(2)\frac{O^2_{ij}}{\theta^2_{\rm conv}}\right]} = exp\left[-4\ln(2)\frac{O^2_{ij}}{\theta^2_{\rm conv}}\right]\frac{N}{A_\theta} \tag{3.5}$$

where N corresponds to the total number of spaxels per velocity channel.

Then, the main-beam temperature of each spaxel is multiplied by its weight w_{ij} and the final temperature per velocity channel, $T_{\rm conv}(v)$, is obtained by integrating $T_{ij}(v)$ over the spaxels:

$$T_{\rm conv}(v) = \sum_{ij} w_{ij} T_{ij}(v) \tag{3.6}$$

The result of Eq. (3.6) is a one-dimensional spectrum of main-beam temperatures as a function of the velocity.

3.5.2 Samples at different angular resolutions

Three datasets were created based on the following criteria:

1. the sample referred as "13.″4 Sample" was defined to study of the CO emission as a function of distinct J transitions. Thus, the mid-J CO data were convolved to the same angular size of the CO (4–3) (13.″4);

2. the "Distance-limited Sample" was based on the convolution of the mid-J CO spectra to the same projected linear scale ($\ell_{\rm proj} = 0.24\,{\rm pc}$) to investigate the CO emission probed by the same volume of gas. The projected linear scale of the beam is defined as:

$$\theta_{\rm obs} = \tan^{-1}\left(\frac{\ell_{\rm proj}}{d}\right) \tag{3.7}$$

   The adopted $\ell_{\rm proj}$ value of 0.24 pc corresponds to the maximum linear size probed by the CHAMP$^+$ maps towards the nearest TOP100 clump, AGAL353.066+00.452, at d = 0.86 kpc. The maximum distance of 5 kpc is defined by the original angular resolution of the data ($\sim$10″, corresponding to $\ell_{\rm proj} \approx$0.24 pc at 5 *kpc*); and

3. the "Full-Map Sample" was defined by the integration of the mid-J CO emission over the full FOV of the datacubes. This sample was aimed to investigate any bias introduced by beam dilution effects presented in the data.

As described above, the definition of each dataset was based in a distance-selection criteria. Thus, the same TOP100 source can be present in more than one dataset depending on its distance.

3.5.3 Gaussian decomposition of the CO profiles

The 13.″4 CO spectra were fitted using Gaussian profiles to investigate the complexity of the CO emission as a function of the different J transitions and, also, the evolutionary stages of the clumps. The fits were performed interactively using the MINIMIZE task in `CLASS/GILDAS`. Initially, each spectrum was fitted assuming a single Gaussian component and the residuals between the observed and the fitted profile were used as a proxy of the goodness of the fit: if the residuals had substructures larger than 3-σ, a second or even a third component were included. The number of adopted components are related to their

physical meaning. A typical CO profile exhibits a narrow and a broad component, with line-widths up to a few $\mathrm{km\,s^{-1}}$ and tens of $\mathrm{km\,s^{-1}}$, respectively. These components are often attributed to different phases of the molecular envelope associated to these sources (*e.g.*, San José-García et al., 2013). Sources exhibiting more complex structures (*e.g.*, molecular outflows) often exhibit non-Gaussian high-velocity wings with line-widths up to $\sim$100 $\mathrm{km\,s^{-1}}$ and, therefore, require a third component. All the Gaussian components were simultaneously fitted in each spectrum. In the particular case of the CO (4–3) data, absorption features shifted in velocity from the main line are detected towards several sources. These features are likely due to emission in the line-of-sight of the reference position and, thus, were masked before fitting the data.

Figure. 3.3 presents examples of CO profiles with satisfactory Gaussian decomposition. Examples of bad fitting are presented in Fig. 3.4. As illustrated in these figures, most of the CO profiles associated with the TOP100 exhibit self-absorption features at channels closer the $V_{\rm lsr}$ position. In such cases, the affected channels were masked out before performing the fit and such condition was sufficient to fit most of the CO profiles exhibiting self-absorption features. However, in other cases, the self-absorption is so strong that the narrowest Gaussian component cannot be adequately fitted. For the CO profiles that were not fit properly, the results of the Gaussian fit and the corresponding integrated properties of the CO profile were excluded from further analysis (see flagged sources in ***Table A.6***).

The general statistics of the Gaussian decomposition are summarised in Table 3.2. The results of the Gaussian fit of each CO spectra convolved to 13.″4 are listed in ***Tables A.2*** to ***A.4***.

Table 3.2 - Results of the Gaussian decomposition of the CO spectra convolved to a common angular resolution of 13.″4.

Transition	Number of Gaussian components		
	One comp.	Two comp.	Three comp.
CO (4–3)	28 (11,11, 3, 3)	67 (3,20,28,16)	3 (0, 0, 2, 1)
CO (6–5)	12 (6, 4, 1, 1)	58 (8,20,20,10)	29 (0, 7,12,10)
CO (7–6)	35 (14,14, 6, 1)	53 (0,16,24,13)	11 (0, 1, 3, 7)

Notes: The columns are as follows: (1) CO transition, (2)–(4) number of sources fitted using (2) a single Gaussian component, (3) two and (4) three components. In each column, the values in parenthesis indicate the corresponding number of 70w, IRw, IRb and H II regions.

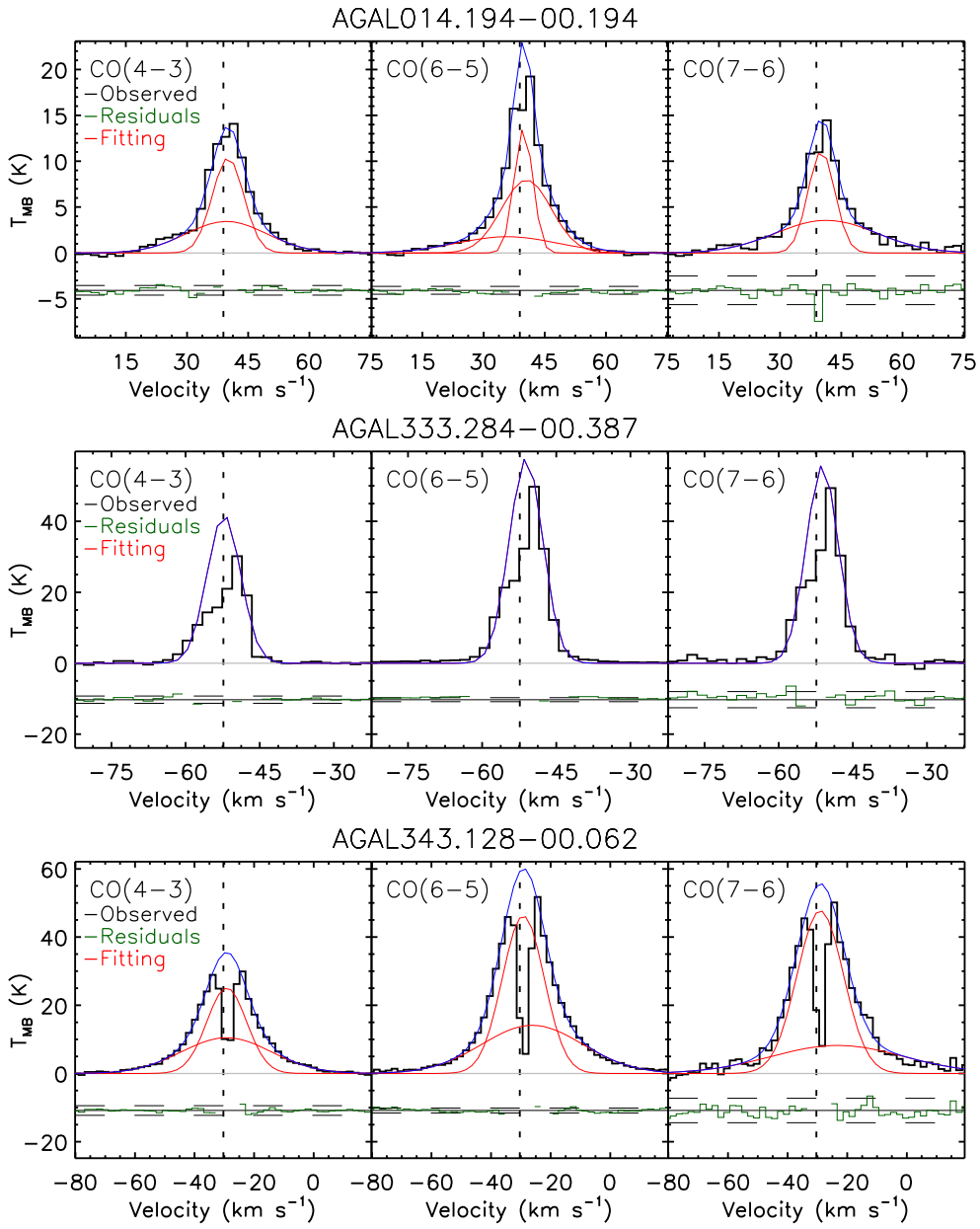


Figure 3.3: Examples of the Gaussian decomposition of the CO profiles towards the TOP100 clumps. For each source, the CO (4–3), CO (6–5) and CO (7–6) profiles are shown in the left, middle and right panel, respectively. The Gaussian components are overlaid as red curves and the total Gaussian fit is indicated by the blue curve. The residuals between the observed and the fitted data are indicated in green. The grey lines indicate the baseline. The dashed horizontal lines, placed on top of the residuals, correspond to the 3-σ level. Self-absorption or absorption features larger than 5-σ were masked out from the residuals.

3.5.4 The integrated properties of the CO emission

The characterisation and analysis of the integrated properties of the CO emission were performed as follows. For each J transition, the full-width at zero power (FWZP), the integrated intensity ($S_{\rm int}$) and the CO line luminosity ($L_{\rm CO}$) were evaluated for the three datasets defined in Sect. 3.5.2.

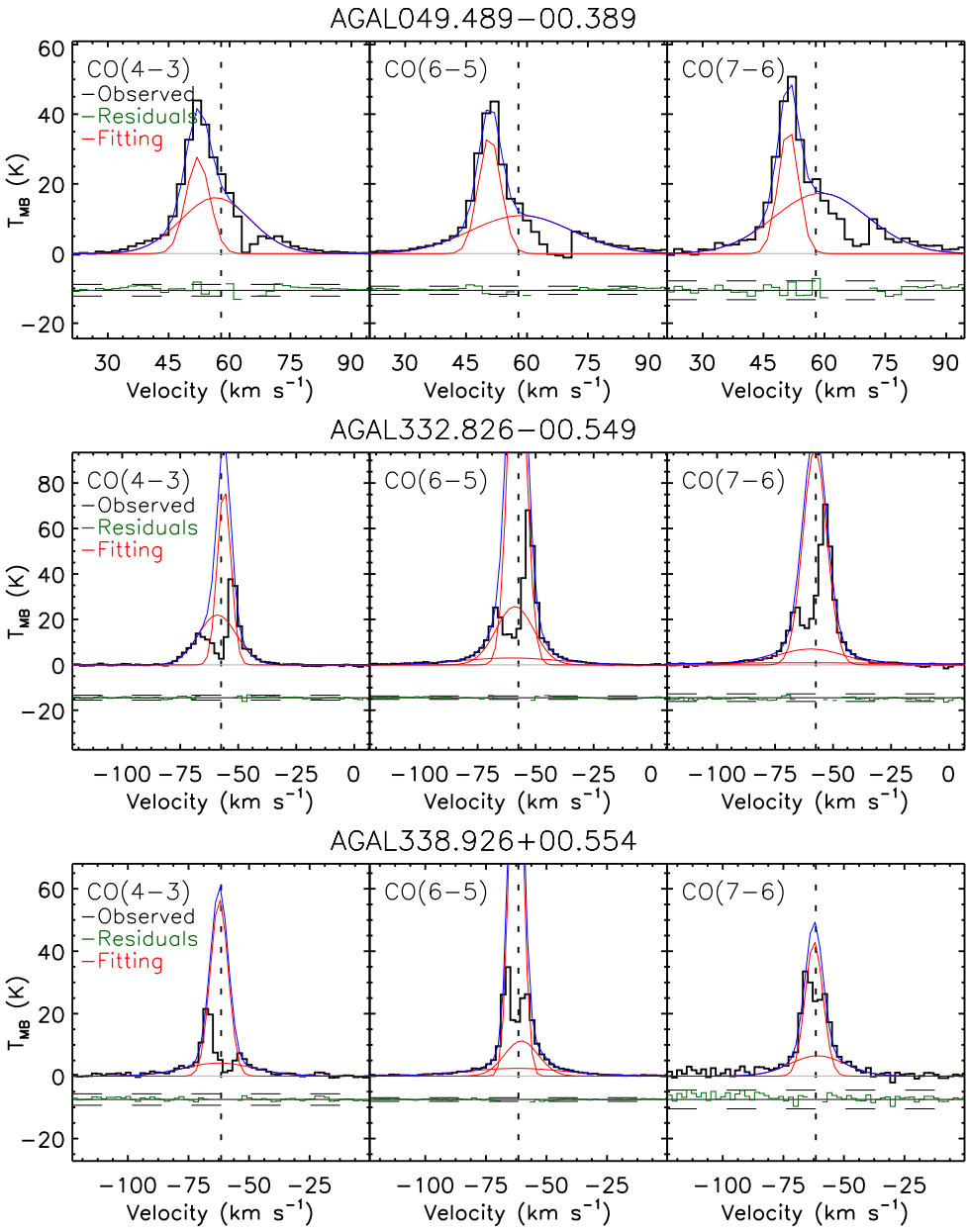


Figure 3.4: Examples of unreliable Gaussian decomposition of the CO profiles towards the TOP100 clumps. In these cases, the Gaussian decomposition was not properly determined and, thus, these profiles were excluded from further analysis.

The FWZP is a non-parametric measurement of the line-width of the CO profiles, defined as the velocity interval where $T_{\rm mb}(v)$ is equal or above a 1-σ threshold. The integrated intensity of the line ($S_{\rm int}$, in K km s^1 units) corresponds to the integration of $T_{\rm mb}$ over the velocity channels within the FWZP:

$$S_{\rm int} = \int_{\rm FWZP} T_{\rm mb}(v)\, dv = \sum_{\rm FWZP} T_{\rm mb}(v)\, \Delta v \tag{3.8}$$

where Δv corresponds to the resolution of data (in km s^{-1}).

The integrated intensity of the CO emission was used to evaluate the CO line luminosity ($L_{\rm CO}$, in K km s^{-1} pc^2 units) by following Eq. 2 from Wu et al. (2004) and derived as follows. The CO line luminosity is defined as:

$$L_{\rm CO} = \int_{\Omega} \left(\int_{v} T_{\rm mb}\, dv \right) (1 + \phi_{\rm beam})\, d^2\, d\Omega_{\rm beam} \tag{3.9}$$

where $\phi_{\rm beam}$ is the filling factor of the beam, d is the distance of the source (in pc) and $\Omega_{\rm beam}$ is the solid angle of the beam (in sr units), defined as:

$$\Omega_{\rm beam} = \int \exp\left[-4\ln(2) \frac{\theta^2}{\theta^2_{\rm beam}} \right] 2\pi\theta\, d\theta = 2.35 \cdot 10^{-11} \frac{\pi \theta^2_{\rm beam}}{4\ln(2)} \tag{3.10}$$

The numeric factor in Eq. (3.10) corresponds to the conversion from radians^2 to arcseconds^2. The filling factor of the beam is defined as the ratio between the area of the source and the beam:

$$\phi_{\rm beam} = \frac{\Omega_{\rm source}}{\Omega_{\rm beam}} = \frac{\theta^2_{\rm source}}{\theta^2_{\rm beam}} \tag{3.11}$$

From Eq. (3.11), the factor $(1 + \phi_{\rm beam})$ in Eq. 3.9 can be rewritten in terms of the sizes of the beam and the source:

$$(1 + \phi_{\rm beam}) = \left(1 + \frac{\theta^2_{\rm source}}{\theta^2_{\rm beam}} \right) = \left(\frac{\theta^2_{\rm beam} + \theta^2_{\rm source}}{\theta^2_{\rm beam}} \right) \tag{3.12}$$

Substituting Eqs. (3.10) and (3.12) in Eq. (3.9):

$$L_{\rm CO} = 23.5 \times 10^{-6} \frac{\pi \theta^2_{\rm source}}{4\ln(2)} \left(\frac{\theta^2_{\rm source} + \theta^2_{\rm beam}}{\theta^2_{\rm source}} \right) S_{\rm int} d^2 \tag{3.13}$$

which express $L_{\rm CO}$ in terms of the observed quantities: $S_{\rm int}$ (in K km s^{-1}), $\theta_{\rm source}$, $\theta_{\rm beam}$ (in arcseconds) and the distance of the source (in kpc).

The FWZP, $S_{\rm int}$ and $L_{\rm CO}$ values for the 13.″4 sample are listed in ***Table A.6***.

3.5.5 The excitation temperature of the CO gas

In local thermodynamic equilibrium (LTE), the population of two distinct energy levels of a given molecule (*e.g.*, the CO), characterised by the indices u and l, are defined by the

Boltzmann distribution:

$$\frac{n_{\rm u}}{n_{\rm l}} = \frac{g_{\rm u}}{g_{\rm l}} \exp\left(-\frac{T_{\rm u} - T_{\rm l}}{T_{\rm ex}}\right) \tag{3.14}$$

where $n_{\rm u}$ is the density of molecules in a given state (in $\rm cm^{-3}$), $g_{\rm u}$ is the statistical weight of the state, $T_{\rm u} = E_{\rm u}/k_{\rm b}$, $E_{\rm u}$ is the upper-level energy of the transition (in erg), $k_{\rm b}$ is the Boltzmann constant (in $\rm erg\,K^{-1}$) and $T_{\rm ex}$ is the excitation temperature (in K).

Ideally, the intensity ratio of different CO transitions (*e.g.*, CO (4–3) and CO (6–5)) provides an accurate estimative of the excitation temperature of the gas. However, in most of the TOP100 sources, a considerable fraction of the CO flux is underestimated due to the presence of self-absorption features at channels closer to the systemic velocity (see examples in Fig. 3.4), leading to unreliable ratios between J transitions.

Alternatively, the excitation temperature of the gas can be inferred from the peak intensity of the CO (6–5)[2] profiles as follows. The main-beam temperature can be expressed in terms of a specific solution of the radiative transfer equation, given by:

$$T_{\rm mb} = T_\nu \left[J_\nu(T_{\rm ex}) - J_\nu(T_{\rm bg})\right](1 - e^{-\tau_\nu}) \tag{3.15}$$

where $T_{\rm bg}$ is the background temperature (2.73 K), $T_\nu = h\nu/k_{\rm b}$, $J_\nu(T) = [\exp(T_\nu/T) - 1]^{-1}$ and τ_ν is the optical depth of the line. Assuming optically thick emission ($\tau_\nu \gg 1$; $e^{-\tau_\nu} \approx 0$), the excitation temperature can be directly evaluated from Eq. (3.15) as:

$$T_{\rm ex} = \frac{T_\nu}{\ln\left[1 + \frac{T_\nu}{T_{\rm mb} + T_\nu J_\nu(T_{\rm bg})}\right]} \tag{3.16}$$

The $T_{\rm ex}$ of the CO (6–5) line was computed for the whole TOP100 sample by using the observed peak $T_{\rm mb}$ of the 13.″4 spectra presented in ***Fig. B.1*** (hereafter, the 'observed' excitation temperature), and the peak temperature of the Gaussian fit (hereafter, the 'Gaussian' excitation temperature; see Sect. 3.5.3 and the spectra presented in ***Fig. B.1***). Both values (*i.e.*, the 'observed' and the 'Gaussian' excitation temperatures) are reported in ***Table A.7***. The uncertainty of the 'observed' $T_{\rm ex}$ values was estimated based on the residual between the $T_{\rm ex}$ value obtained assuming the peak $T_{\rm mb}$ and the one calculated assuming the peak $T_{\rm mb}$ plus/minus the rms of the spectrum. Similarly, the uncertainty of

[2] the CO (6–5) data were chosen since they are associated with the lowest rms values among the transitions investigated in this work (see Table 3.1).

the 'Gaussian' $T_{\rm ex}$ values was evaluated based on the errors of the Gaussian fit parameters from ***Table A.3***.

3.5.6 Linear regression fitting procedure and non-parametric statistical tests

This section describes the linear regression fitting procedure adopted in this work and the two statistical tests used on the comparison between the quantities investigated in the forthcoming sections.

The relation between the CO and the clump properties were performed using a Bayesian approach, using an `IDL` code based on the LINMIX_ERR algorithm. The code performs a Bayesian linear regression fit on the data considering errors in both x- and y-axis. The points are fitted by adjusting the linear model with three free parameters, $y = \alpha + \beta \cdot x \pm \epsilon$, where α, corresponds to the intercept, β is the slope and ϵ is the intrinsic scatter. The fitting can be performed on the linear-linear, log-linear, or log-log scales. In order to obtain a statistically reliable solution, about 100 000 iterations were computed per fit. A FITS file is created at the end of the execution, containing a two-dimensional map of (α,β) values, and all the other information of the fit is available on its header (see Fig. A.2 in Appendix A.1). Figure 3.5 illustrates an example of the linear regression fit of the CO (6–5) line luminosity against the bolometric luminosity of the TOP100 clumps. The parameters of the best fit are determined as the median values of the α and β parameters from their distribution map, and their uncertainty are obtained from the maximum elongation of the region populated with counts above the 68% of the peak level, indicated as the white contour in the right panel of Fig. 3.5.

The two-side Kolmogorov-Smirnov (KS) test was adopted to check whether the derived properties are distinct between different evolutionary stages of the clumps (*e.g.*, the distribution of the CO line luminosity for 70w and IRb clumps). The KS rank factor and its associated confidence value, p-value, are present for each comparison. The null hypothesis is that both subsamples are drawn from the same population. However, if the p-value is smaller than 3σ (*i.e.*, $p < 0.001$), the null hypothesis can be rejected, allowing to conclude that there is sufficient evidence to consider the samples to be statistically different. Weaker but considerable significances are obtained for p-values up to 0.05, representing the $\sim 2\,\sigma$ confidence level. The KS test is sensitive to differences in the location and the shape of the cumulative distributions of the samples. A KS rank factor closer to 1 indicates more

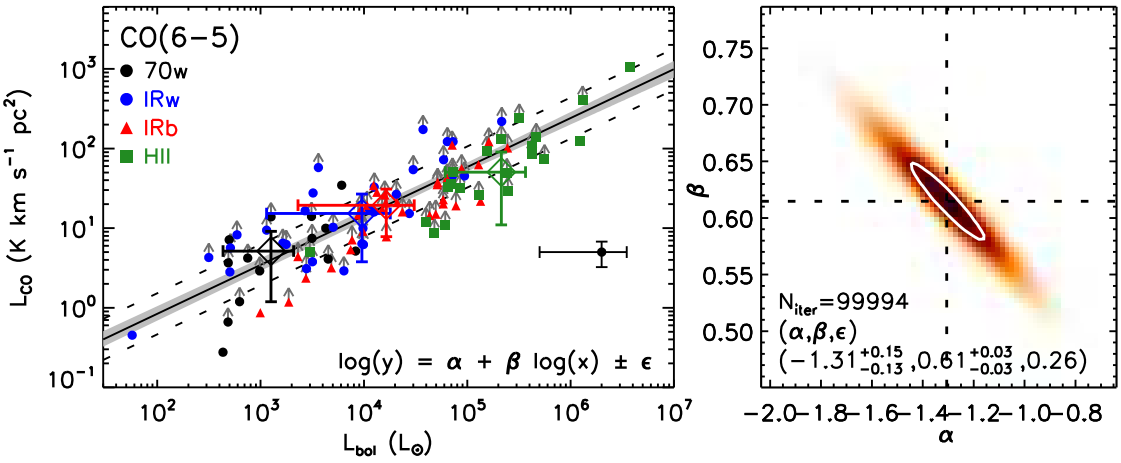


Figure 3.5: Left panel: Line luminosity of the CO (6–5) transitions as a function of the bolometric luminosity of the TOP100 sources. The median values for each class are shown as open diamonds and their error bars correspond to the mean absolute deviation of the data from their median value. Points having an upward arrow indicate a self-absorption feature in the spectrum convolved to 13.″4 and correspond to a lower limit. The typical error bars are shown on the bottom right side of the plots. The black solid line is the best fit, the light grey shaded area indicates the 68% uncertainty, and the dashed lines show the intrinsic scatter (ϵ) of the relation. Right: Distribution of (α,β) coefficients based on the Bayesian fit. The uncertainties on the coefficients were estimated in the region where the distribution equals to 68% of the peak (shown as the white contour). The number of iterations and the values of each coefficient are shown in the left bottom side of each plot.

significant differences between the cumulative distribution of two subsamples, while KS factors closer to 0 indicates weak differences.

The level of correlation between two quantities was checked by computing their Spearman correlation (ρ) rank and the associated p-value. The ρ values vary between -1 and $+1$, ρ values closer to 1 indicates that both quantities are strongly correlated, while negative ρ values indicate anti-correlation instead. ρ values around 0 indicate weak correlation level between the two quantities. For quantities that are mutually dependent on a third parameter (*e.g.*, the CO line luminosity and the bolometric luminosity depend equally on the distance of the sources), a partial Spearman correlation test was performed. In these cases, the partial Spearman coefficient, $\rho_{\rm p}$ is also provided.

3.6 Results

3.6.1 The CO line profiles from 70 μm-weak clumps to H II regions

The overall analysis of the Gaussian fitting from Sect. 3.5.3 indicates that most of the CO profiles are well fitted with two Gaussian components: one for the envelope (narrow

component) and another for the high-velocity emission (broad component). In the particular case of the CO (6–5) spectra, exhibiting the higher signal-to-noise ratio of the CO transitions, a third component is often required for fitting the profile towards one-third of the sample (29 from 99 profiles, see Table 3.2).The need of two or even three Gaussian components suggests that the high-velocity components are likely tracing molecular outflows in most of the CO (6–5) profiles (∼88%, see Table 3.2).The detailed discussion of the outflow content in the TOP100 sample is presented in Sect. 3.7.4.

Differences between the CO profile as a function of the evolutionary stages of the TOP100 clumps were investigated by computing the average spectrum of each class. The average spectra of the full TOP100 sample using the spectra convolved to 13.″4 are presented in Fig. 3.6. The results on the distance-limited subsample are presented in Fig. 3.7. The average spectra were constructed by shifting the individual spectra to 0 km s^{-1} using the corresponding $V_{\rm lsr}$ from Table A.1. Then, the average was obtained by scaling each individual spectrum by the square of the ratio between the median distance of the sample (d = 3.8 kpc for the full sample and 3.26 kpc for the distance-limited subsample) and the distance of the source.

The most significant results are found towards the average spectrum of the distance-limited subsample, presented in Fig. 3.7. The results for the distance-limited subsample indicates an increase of the integrated intensity and line width (expressed in terms of the full-width at zero power (FWZP, see description in Sect 3.5.4) of the CO profiles as a function of the evolutionary class of the clumps, from 70w clumps towards H II regions. These findings are less evident in Fig. 3.6 due to the presence of clumps located at larger distances ($\gtrsim$ 12 kpc), which are sampling a larger amount of gas in the 13.″4 beam than the corresponding volume sampled towards closer objects of the TOP100 sample. Indeed, only after including the IRw sources at distances larger than 12 kpc, as illustrated in Fig. 3.8, the IRw and IRb classes are less distinguishable.

The increase of the line width and the intensity of the CO emission as a function of the evolution of the clumps also suggests that the emission arises from different structures within the young clumps. For example, Leurini et al. (2013) have shown that the mid-J and high-J CO emission towards the hot core G327.30.6 (corresponding to the ATLASGAL clump AGAL327.293−00.579) arise from photon-dissociation regions (PDR) associated with an H II region rather than from the younger sources in the field.

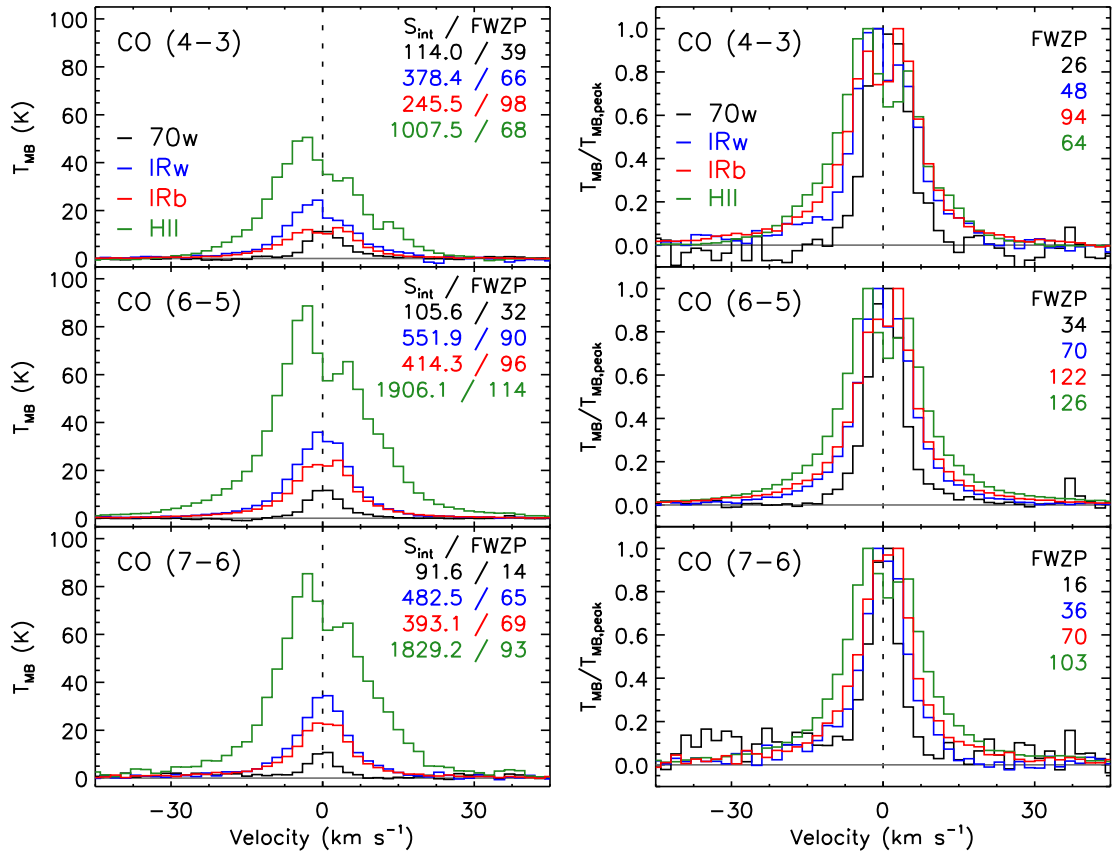


Figure 3.6: Left panel: Average spectra per evolutionary class of the CO (4–3) (top panels), CO (6–5) (middle) and CO (7–6) (bottom) lines using the spectra convolved to 13.″4. Right: Same plot, but the average CO spectra are normalised by their peak intensity. The baseline level is indicated by the solid grey line and the black dashed line is placed at 0 km s^{-1}. The FWZP of the profiles is shown in the upper right side of the panels (in km s^{-1} units), together with the integrated intensities ($S_{\rm int}$, in K km s^{-1} units) of the CO profiles shown in the left panels.

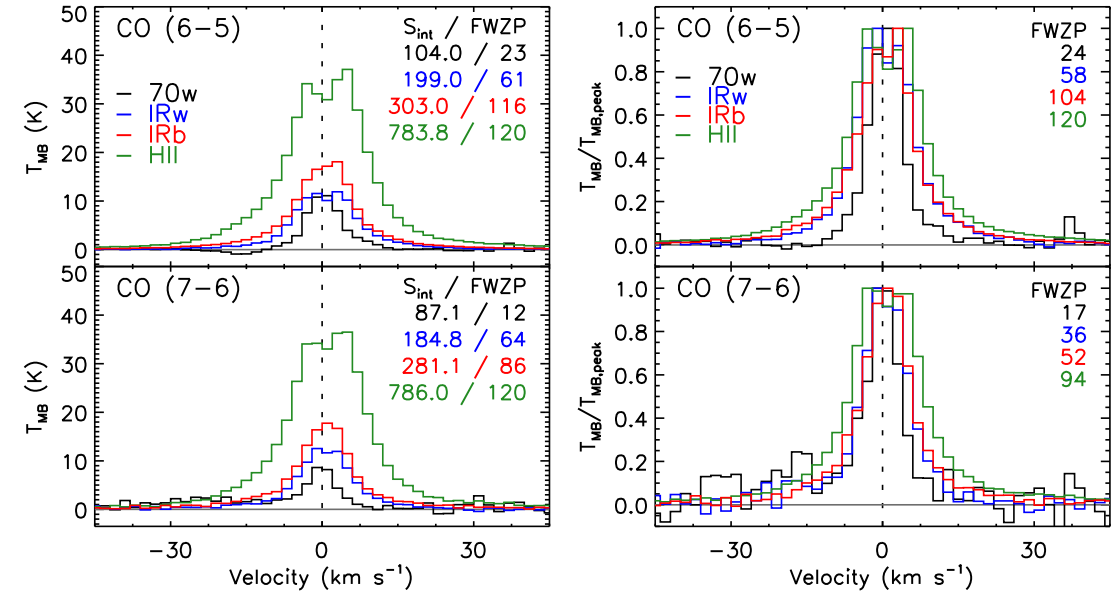


Figure 3.7: Similar as Fig. 3.6 but using the distance-limited subsample of the TOP100 clumps.

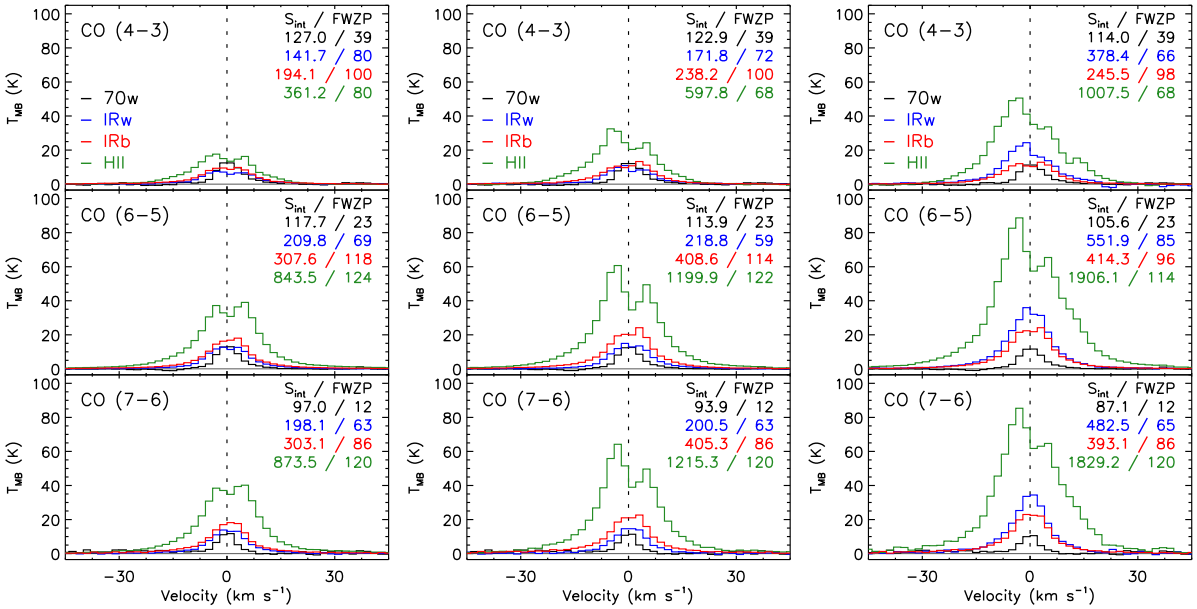


Figure 3.8: Average CO profiles per evolutionary class of the TOP100 clumps. The averaging was performed for the clumps located at distances closer than 5.0 (left panel), 10.0 (middle) and 15.0 kpc (right panel), indicating that the presence of outliers IRw clumps at distances larger than 10 kpc leads to an overestimation of the CO fluxes at early evolutionary stages.

3.6.2 The extension of the CO emission

The spatial extension of the CO (6–5) emission was investigated for the TOP100 clumps. The CO maps were created by integrating the CHAMP$^+$ datacubes over the velocity channels within the FWZP of each clump. Fig. 3.9 presents examples of the *Herschel*-PACS 70 μm maps at similar angular resolution ($\sim$10″) of the CO maps, overlaid by the CO (left panels) and 70 μm contours (right) for two TOP100 clumps. The FIR maps and CO contours indicate that the molecular and dust emission are not only tracing the central region of the clumps but also exhibit extend structures in most of the clumps. The CO (6–5) maps of the whole TOP100 clumps are presented in ***Fig. B.1***.

The linear size of the CO emission, Δs, was estimated based on the mean value of the maximum and minimum elongation of the 50% peak intensity contour level of the CO maps. The same procedure was applied on the *Herschel*-PACS 70 μm maps of the TOP100 sources to estimate the extension of FIR emission, $\Delta s_{70\mu m}$. The linear sizes of the CO and the 70 μm emission are reported in ***Table A.5***.

The extent of the mid-J CO emission is plotted against the clump properties in Fig. 3.10. The analysis of the plots together with the Spearman rank tests comparing Δs and the clump properties indicate that the CO extension is strongly correlated with $M_{\rm clump}$ while it exhibits a weaker correlation with $L_{\rm bol}$ (ρ=0.72 and 0.42, respectively;

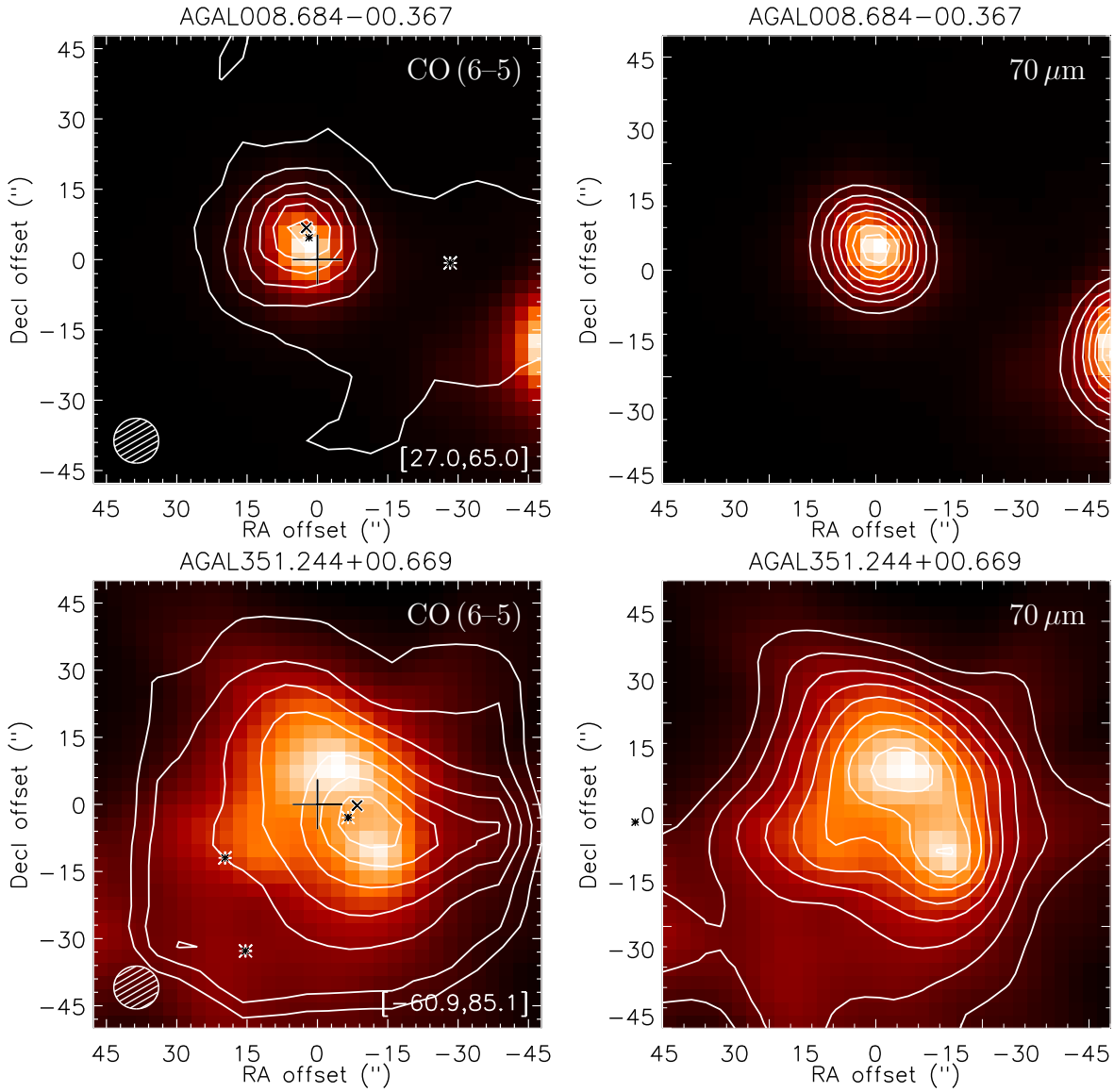


Figure 3.9: Herschel/PACS images at 70 μm overlaid by the CO (6–5) emission (left panels) and the FIR emission contours (right panels) towards two TOP100 clumps. The CO contours correspond to the emission integrated over the full-width at zero power (FWZP) of the CO (6–5) profile, the velocity range shown in the bottom right side of the image. In both panels, the contour levels are shown from 20% to 90% of the peak emission of each map, in steps of 10%. The (0,0) position of the map is shown as a + symbol, the position of the CSC source from Contreras et al. (2013) is shown as a × symbol and the dust continuum emission peaks from Csengeri et al. (2014) are shown as asterisks. The beam size is indicated in the bottom left region of the map.

$p < 0.001$ for both comparisons) or L/M ($\rho = -0.02$, $p = 0.82$). These results suggest that the size of the CO emission is likely dependent on the available mass of the clumps but not on their radiative flux or their evolutionary phase.

The bottom right panel of Fig. 3.10 presents the extension of the CO versus the size of the FIR emission. The plot indicates that both quantities are well correlated, confirmed by their Spearman rank coefficient ($\rho = 0.67$, $p < 0.001$). Most of the clumps are located

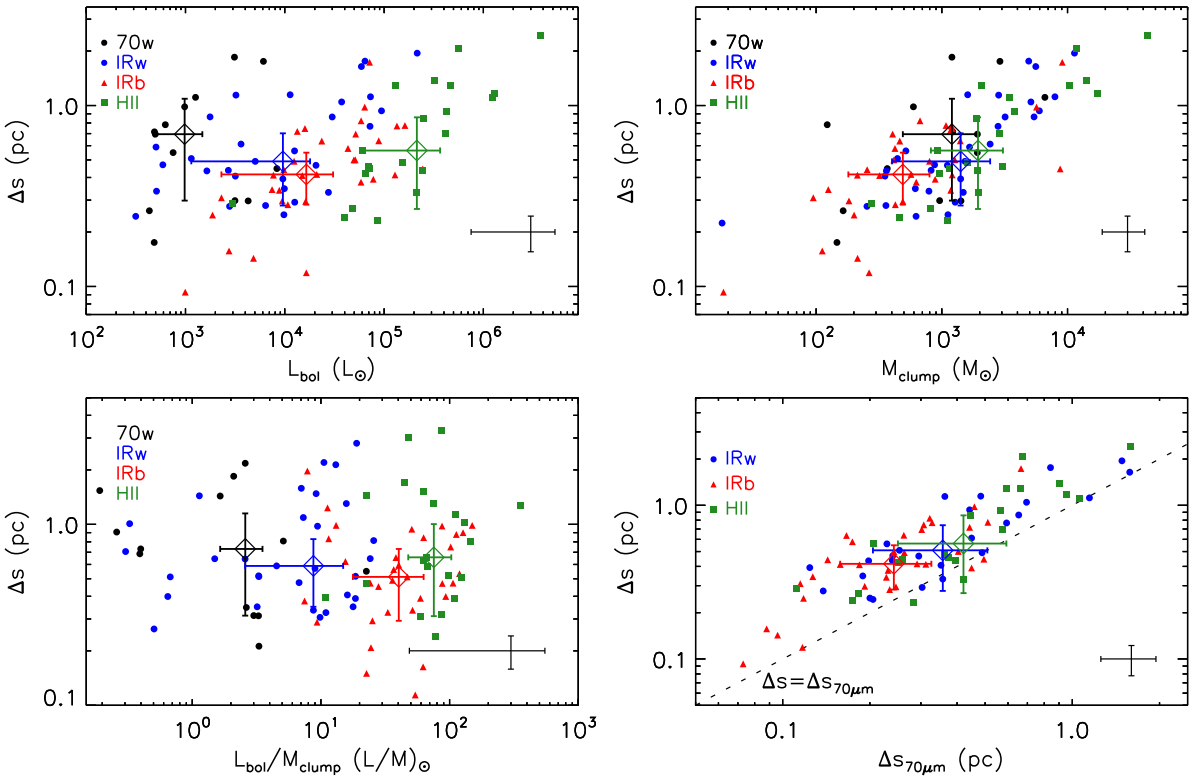


Figure 3.10: Size of the CO (6–5) emission towards the TOP100 sample versus the bolometric luminosity (top left panel), the mass (top right) and the luminosity-to-mass ratio (bottom left) of the whole TOP100 sample. Bottom right: the mid-J CO extension against the size of the 70 μm continuum emission for the IRw, IRb and H II classes. For each class, the median and the dispersion from the median values are shown in their corresponding colours. The typical uncertainties are indicated by the error bars on the bottom right of each plot.

above the $y = x$ line, suggesting that the component of the gas probed by the CO (6–5) is more extended than the dust emission at 70 μm.

3.6.3 *Evolution of the CO properties and comparison with the clump properties*

The left panel of Fig. 3.11 presents the cumulative distribution function (CDF) of the CO line luminosities of the CO (4–3), CO (6–5) and CO (7–6) transitions as a function of the evolutionary class of the TOP100 clumps. The CDFs indicate a progressive increase in the L_{CO} values as a function of the evolutionary stage of the clumps in every J transition, from 70w clumps towards H II regions. Each evolutionary class of the TOP100 clumps was tested against the others (*e.g.*, 70w versus IRw; IRb versus H II) by computing their two-sided Kolmogorov-Smirnov coefficient. The statistics are reported in Table 3.3 and the KS coefficients indicate that the most significant differences are found when comparing the earlier (*e.g.*, 70w and IRw) against the most evolved class (H II). Weaker correlations (KS $\lesssim 0.5$) are found when comparing consecutive classes, indicating that L_{CO} presents a

mild increase from class to class.

Table 3.3 - Kolmogorov-Smirnov statistics of the CO line luminosity as a function of the evolutionary class of the TOP100 clumps.

Classes	CO (4–3)	CO (6–5)	CO (7–6)
70w-IRw	0.32, $p = 0.24$	0.44, $p = 0.03$	0.46, $p = 0.02$
70w-IRb	0.34, $p = 0.16$	0.66, $p < 0.001$	0.66, $p < 0.001$
70w-H II	0.55, $p = 0.007$	0.74, $p < 0.001$	0.83, $p < 0.001$
IRw-IRb	0.21, $p = 0.46$	0.22, $p = 0.38$	0.24, $p = 0.26$
IRw-H II	0.36, $p = 0.06$	0.44, $p = 0.01$	0.45, $p = 0.007$
IRb-H II	0.42, $p = 0.02$	0.48, $p = 0.002$	0.51, $p = 0.002$

Notes: The rank KS and its corresponding probability (p) are shown for each comparison. A p-value of < 0.001 indicate a correlation at 0.001 significance level. p-values of 0.05, 0.002 and < 0.001 represent the ~ 2, 3 and $> 3\,\sigma$ confidence levels.

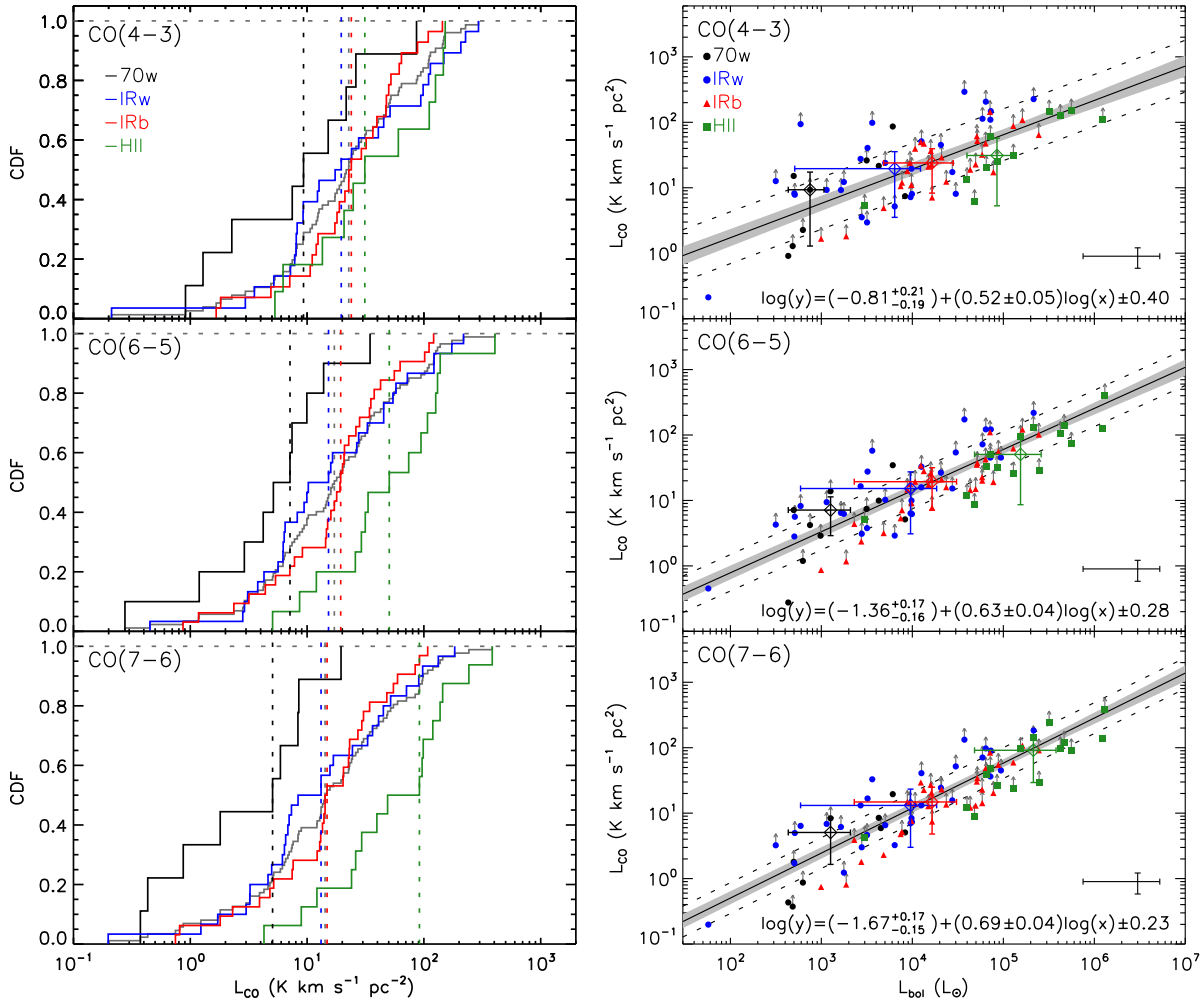


Figure 3.11: Left panels: Cumulative distribution function of the line luminosity of the CO (4–3) (top panel), CO (6–5) (middle) and CO (7–6) transitions (bottom) towards the TOP100 sample. The median values per class are shown as vertical dashed lines in their corresponding colours. Right panels: Line luminosity of the same CO J-transitions against the bolometric luminosity of the clumps. Points exhibiting an upward arrow indicate a self-absorption feature in the spectrum and correspond to a lower limit. For a complete description of the elements in the plots, see Fig. 3.5.

The CO line luminosity of the CO (4–3), CO (6–5) and CO (7–6) transitions were investigated against the clump properties and the results are presented as follows. The Spearman rank correlation coefficient for comparisons between $L_{\rm CO}$ and the clump properties are listed in Table 3.4, together with the partial Spearman coefficients for $L_{\rm CO}$ versus $L_{\rm bol}$ and $M_{\rm clump}$. The parameters of the best fits between the quantities are summarised in Table 3.5.

Table 3.4 - Spearman rank correlation statistics for the CO line luminosity as a function of the clump properties towards the TOP100 sample.

Property	CO (4–3)	CO (6–5)	CO (7–6)
$L_{\rm bol}$	0.67, $p<0.001$; $\rho_p=0.68$	0.83, $p<0.001$; $\rho_p=0.87$	0.87, $p<0.001$; $\rho_p=0.89$
$M_{\rm clump}$	0.72, $p<0.001$; $\rho_p=0.32$	0.74, $p<0.001$; $\rho_p=0.47$	0.75, $p<0.001$; $\rho_p=0.52$
L/M	0.20, $p=0.09$	0.39, $p<0.001$	0.43, $p<0.001$

Notes: The rank ρ and its corresponding probability (p) are shown for each comparison. A p-value of <0.001 indicate a correlation at 0.001 significance level. p-values of 0.05, 0.002 and <0.001 represent the ~ 2, 3 and $>3\,\sigma$ confidence levels. For $L_{\rm bol}$ and $M_{\rm clump}$, the partial correlation coefficient, ρ_p, is also shown.

Table 3.5 - Parameters of the fits of $L_{\rm CO}$ as a function of the clump properties.

Transition	Property	α	β	ϵ
CO (4–3)	$L_{\rm bol}$	$-0.81^{+0.21}_{-0.19}$	0.52±0.05	0.40
	$M_{\rm clump}$	$-1.24^{+0.20}_{-0.18}$	0.87±0.06	0.33
	L/M	$1.03^{+0.13}_{-0.14}$	0.21±0.10	0.60
CO (6–5)	$L_{\rm bol}$	$-1.36^{+0.17}_{-0.16}$	0.63±0.04	0.28
	$M_{\rm clump}$	-1.49±0.21	0.90±0.07	0.36
	L/M	$0.73^{+0.10}_{-0.12}$	0.41±0.10	0.57
CO (7–6)	$L_{\rm bol}$	$-1.67^{+0.17}_{-0.15}$	0.69±0.04	0.23
	$M_{\rm clump}$	-1.53±0.23	0.89±0.08	0.43
	L/M	0.56±0.11	0.54±0.10	0.56

Notes: The fits were performed by adjusting a Bayesian model with three free parameters in the form of $\log(y) = \alpha + \beta \log(x) \pm \epsilon$, where α, β and ϵ correspond to the intercept, the slope and the intrinsic scatter, respectively.

The right panel of Fig. 3.11 presents the $L_{\rm CO}$ values as a function of $L_{\rm bol}$ of the TOP100 clumps for the CO (4–3), CO (6–5) and CO (7–6) transitions. The Spearman rank correlation indicates that both quantities are well correlated ($\rho \gtrsim 0.7$, $p<0.001$), even when accounting for their dependence on the distance ($\rho_p \gtrsim 0.7$). The best fit of the data indicates that the dependency of $L_{\rm CO}$ with the bolometric luminosity of the clumps gets steeper as a function of J: the slope increases from 0.52±0.05 to 0.69±0.04 for the CO (4–3) towards the CO (7–6) transition. Similar results are also found towards the mid-J CO lines when

considering the distance-limited subsample (0.53±0.03 and 0.59±0.03 for the CO (6–5) and CO (7–6), see left panels from Fig. A.3).

The CO line luminosities are plotted against the mass of the clumps in the left panel of Fig. 3.12. The data display similar correlation factors ($\rho \geq 0.72$, $p < 0.001$, see Table 3.4) as those found between $L_{\rm CO}$ and $L_{\rm bol}$. However, the partial Spearman rank factor indicates that the distance plays a significant role in the observed correlation between $L_{\rm CO}$ and $M_{\rm clump}$ ($\rho_p \lesssim 0.5$). No significant differences on the slope of the relations were found when comparing distinct J levels (all values are consistent within 1-σ, see Table 3.5).

The stronger correlation between $L_{\rm CO}$ and $L_{\rm bol}$ and $M_{\rm clump}$ indicates that the CO emission is more dependent on the bolometric flux than the available mass in the clumps.

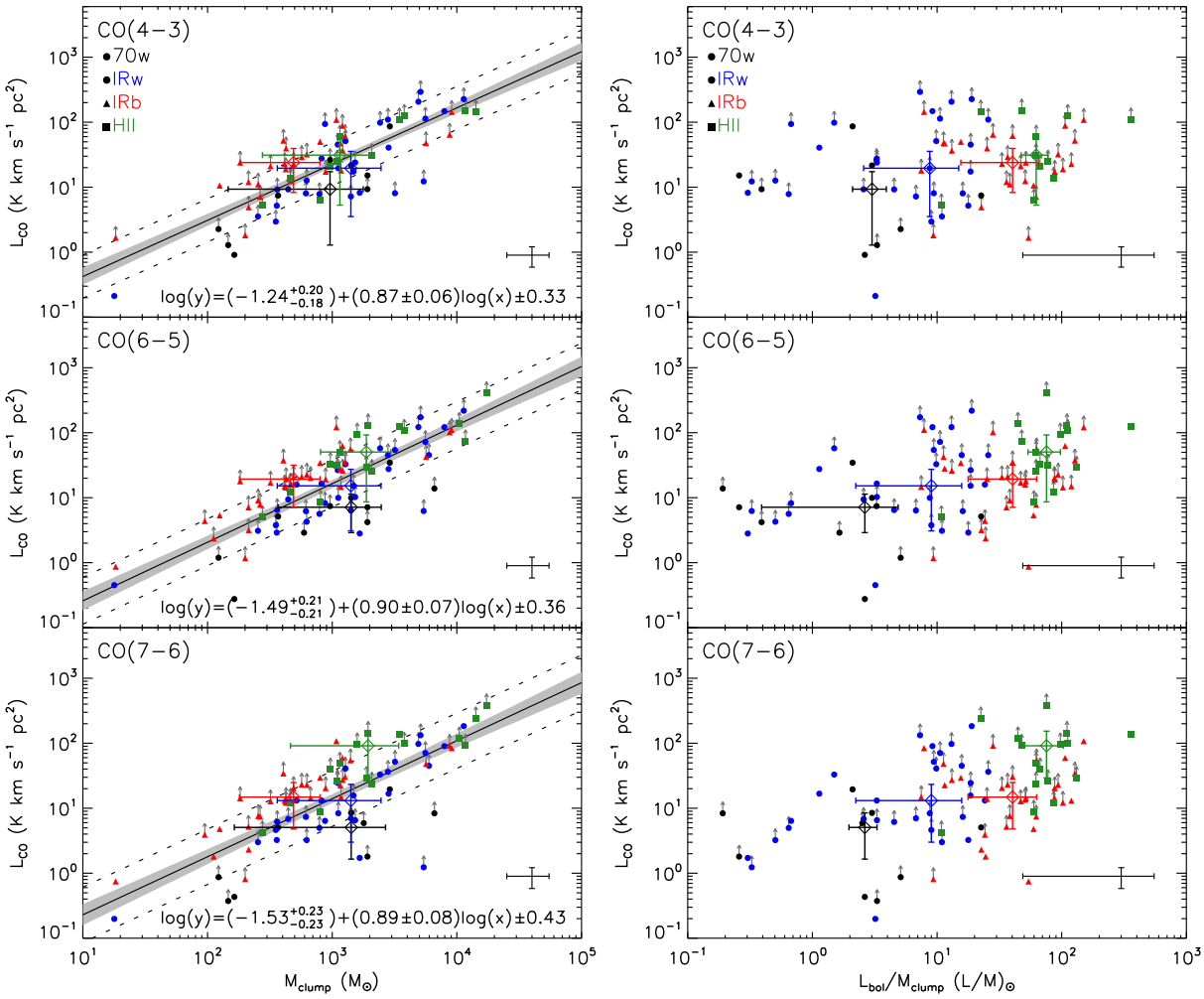


Figure 3.12: The distribution of the CO (4–3) (top panels), CO (6–5) (middle) and CO (7–6) (bottom) line luminosities as a function of the mass ($M_{\rm clump}$, left panels) and the luminosity-to-mass ratio of the clumps (L/M, right panel). For a complete description of the elements in the plots, see Fig. 3.5.

Finally, the right panel of Fig. 3.12 presents the $L_{\rm CO}$ *versus* the luminosity-to-mass ratio

of the clumps (L/M). The L/M ratio has been used as a rough estimator of evolution in the pre-main-sequence stages for both low- (Saraceno et al., 1996) and high-mass sources (e.g., Molinari et al., 2008): embedded and younger regions are often associated with small L/M values while higher L/M values are found in sources with stronger radiative flux and that have accreted most of the mass (Molinari et al., 2016; Giannetti et al., 2017).

Even though the median $L_{\rm CO}$ values per class increase towards larger L/M values, no strong correlation between these quantities was found ($\rho \lesssim 0.4$, see Table 3.4).

3.7 Discussion

3.7.1 The scaling of the CO emission from low- to high-luminosity clumps

The relations between $L_{\rm CO}$ and $L_{\rm bol}$, presented in Sect. 3.6.3 were further investigated in terms of the different linear size probed by the observations. The identification of possible bias that might be introduced in the analysis of sources distributed in a broad range of luminosities and distances (*i.e.*, the different linear sizes probed by the same beam size) is crucial for comparing galactic (*e.g.*, San José-García et al., 2013, 2016) and extragalactic studies (*e.g.*, Weiß et al., 2007; Decarli et al., 2016).

The mid-J CO data of the TOP100 sample was combined with observations towards a low-luminosity sample of eight (8) low-mass YSOs from van Kempen et al. (2009, hereafter, vK09), also observed in the CO (6–5) and CO (7–6) transitions. In addition, results for the CO (10–9) line from San José-García et al. (2013, hereafter, SJG13) were also included in the analysis. The sample from SJG13 covers a broad range of luminosities ranging from <1 to $\sim 10^5\,{\rm L}_\odot$ at different evolutionary stages, defining a complementary sample to the TOP100 and vK09 samples, observed at lower-J transitions.

The sources were divided into two luminosity regimes: low-luminosity sources with $L_{\rm bol} \leq 50\,{\rm L}_\odot$; and high-luminosity objects with $L_{\rm bol} > 50\,{\rm L}_\odot$. The limit $L_{\rm bol} = 50\,{\rm L}_\odot$ was chosen to separate low- and high-luminosity YSOs by following the definition from SJG13. Using Eq. (3.13), the $L_{\rm CO}$ values for the TOP100 clumps were recomputed from the central position of the CHAMP$^+$ maps at the original resolution of the data (see Table 3.1), probing linear scales ranging from ~ 0.04 to $0.6\,{\rm pc}$. Similarly, the line luminosities of the low-mass YSOs from vK09 were computed using the integrated intensities reported in their Table 3 at the original resolution of the CHAMP$^+$ maps, probing scales of $\sim 0.01\,{\rm pc}$. Finally, the

CO (10–9) line luminosities from SJG13 were computed using the integrated intensities reported in their Table A.1 obtained at a resolution of $\sim 20''$ and probing linear scales ranging from ~ 0.01 to 0.77 pc.

Figure 3.13 presents the CO line luminosity as a function of the bolometric luminosity for the CO (6–5), CO (7–6) and CO (10–9) transitions. For each transition, the low- and the high-luminosity subsamples were fitted individually using a simple linear regression model, $\log(y) = \alpha + \beta \cdot \log(x)$. The power-law indices (β) derived for the CO (6–5) data are 0.59 ± 0.25 and 0.59 ± 0.04 for the low- and high-luminosity subsamples, respectively. Although the power-law indices are consistent within 1-σ, the difference on the intercept (α) of each fit is roughly one order of magnitude, from -2.75 to -1.54 dex. Such large offset indicates that the $L_{\rm CO}$ values are systematically larger towards more luminous sources. Indeed, the observed offset explains reasonably well the steeper power-law index found when combining both subsamples, $\beta = 0.74 \pm 0.03$. Similar results are found for the CO (7–6) and the CO (10–9) data. The offset between the low- and high-luminosity subsamples seems to decrease towards higher J transitions (~ 1.0, ~ 0.8, and 0.3 dex for the CO (6–5), CO (7–6) and CO (10–9) transitions, respectively). In every case, the observed offset explains the steeper slope found when fitting the subsamples together.

The offsets observed between the CO line luminosity as a function of $L_{\rm bol}$ for the low- and high-luminosity subsamples are interpreted as caused by the different linear scales and, therefore, different volumes of gas sampled by the same angular resolution of the data across the x-axis in Fig. 3.13. Since more luminous clumps are often located at larger distances than less luminous objects, the volume of gas increases with $L_{\rm bol}$ and so the observed CO luminosity. In addition, as presented in the maps of ***Fig. B.1***, the mid-J CO line traces extended gas, probably as a consequence of clustered star-forming process. Thus, for sources located at closer distances (and, in general, less luminous objects), only a fraction of the total CO emission is probed by the beam size. In the case of the CO (10–9) emission, low- and high-mass sources are differently affected by dilution effects. For low-mass YSOs, the line emission is extended and mainly arises from UV-heated outflow cavities (van Kempen et al., 2010). In the case of high-mass YSOs, most of the emission comes from the inner region of the passively heated envelope (Karska et al., 2014), and could also suffer from beam dilution effects. The distinct origin of the high-J CO emission for each luminosity regime, suffering from similar beam dilution effects, could explain the smaller

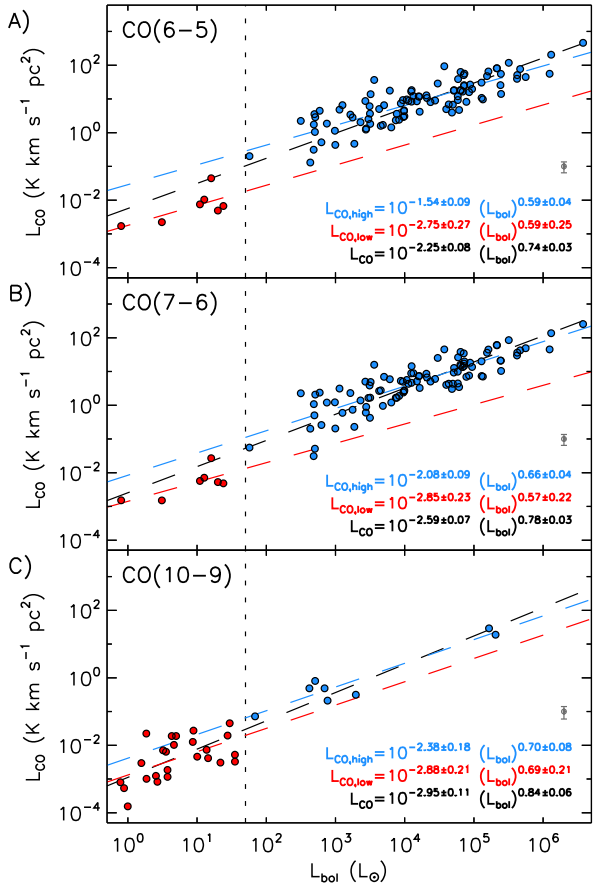


Figure 3.13: CO line luminosity against the bolometric luminosity for the CO (6–5) (top panel), CO (7–6) (middle panel) and CO (10–9) (bottom panel) lines. The fits were performed on the low- and high-luminosity subsamples (indicated as red and blue symbols, respectively) and combining both subsamples. The best fits are indicated by the red, blue and black dashed lines, respectively. The dashed vertical line at $L_{\rm bol} = 50\,{\rm L}_{\odot}$ marks the transition from low- to high-luminosity. The typical error bars are shown on the bottom right side of the plots. See the text for more information about the plots.

offset observed in the CO (10–9) data.

The complexity on extending Galactic relations to extragalactic objects was recently pointed out by SJG13 and San José-García et al. (2016), who found a similar increase on the power-law indices of the high-J CO line luminosity as a function of the bolometric luminosity when including extragalactic sources in the fit (*e.g.*, see Fig. 14 from SJG13 and Fig. 9 from San José-García et al. 2016).

3.7.2 Opacity effects and correlation between CO and the clump properties

A systematic increase in the slope of the relations between $L_{\rm CO}$ and $L_{\rm bol}$ as a function of the J transition (see right panel of Fig. 3.11) indicates that higher-J CO lines are more dependent on the $L_{\rm bol}$ than lower-J transitions. Such increase is even more evident

when including the results towards higher-J transitions, such as the study of the CO (10–9) emission towards a complementary sample of the TOP100, investigated by San José-García et al. (2013). The best fit to their data, $\log(L_{\rm CO}) = (-2.9\pm0.2)+(0.84\pm0.06)\log(L_{\rm bol})$, indicates a even steeper relation than those found towards lower-J transitions from Fig. 3.11. However, as discussed in Sect. 3.7.1, the fit of the data from San José-García et al. (2013) is dominated by the high-mass sources in their sample. Thus, the best fit for their high-mass sources, $\log(L_{\rm CO}) = (-2.4 \pm 0.2) + (0.70 \pm 0.08)\log(L_{\rm bol})$, was used to compare with the power-law indices derived towards the TOP100 sample. The power-law indices, β, of the $L_{\rm CO}$ versus $L_{\rm bol}$ are plotted against the upper-level J number in Fig. 3.14, confirming that β gets steeper towards higher-J transitions.

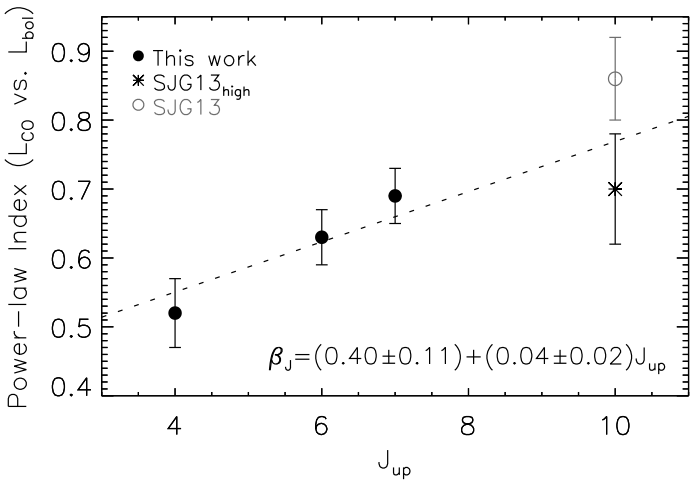


Figure 3.14: Power-law indices of the $L_{\rm CO}$ versus $L_{\rm bol}$ relations for different J transitions as a function of the upper-level J number. The β indices from Table 3.5 (filled black circles) are plotted together with the CO (10–9) slope from San José-García et al. (2013) (open grey circle, excluded from the fit) and the exponent derived for their high-luminosity subsample ($*$ symbol, see Fig. 3.13). The best fit is indicated by the dashed black line.

The CO lines are certainly optically thick and their opacity is likely to decrease as a function of J. Recently, Benz et al. (2016) discussed the effects of the opacity and the steepness of the relations between the line luminosity and the clump properties. Those authors propose that the value of the power-law exponent of the line luminosity of particular molecules (*e.g.*, HCO) and transitions depends on the radius where the line gets optically thick and, therefore, depends on the temperature of the gas. Thus, the increasing of the power-law index as a function of J suggests that higher J CO transitions are tracing a warmer component of the gas, closer to the protostar. This is also confirmed by the intercept value of the fits presented in Fig. 3.11, decreasing from -0.81 ± 0.21 to -1.67 ± 0.17 from CO (4–3) to CO (7–6) transitions (see Table 3.5), in agreement with Benz et al. (2016).

The weak correlation observed between L_{CO} and L/M (see right panel of Fig. 3.12) is quite unexpected, considering that the median L/M values per class are well separated among the TOP100 clumps (2.6, 7.2, 37.6 and 75.5 for 70w, IRw, IRb and H II regions, respectively, König et al., 2017). Previous works have found a similar behaviour towards SiO observations of sources with L/M values as those of the TOP100 sample (*e.g.*, Leurini et al., 2014; Csengeri et al., 2016). While the line luminosity of low-excitation SiO transitions does not increase as a function of L/M, higher energy SiO lines display a more pronounced increase with time. This behaviour is interpreted in terms of changes in the excitation conditions with time, which is not probed by low excitation transitions. This is likely the case for the low- to mid-J CO lines, associated with relatively low energies (≤ 155 K), while higher-J transitions are likely to be more sensitive to changes in the excitation due to their energies higher than 300 K (*e.g.*, the CO (10–9) or even higher-J transitions). Indeed, the Spearman rank factor increases from CO (4–3) ($\rho = 0.20$, see Table 3.4) towards the CO (7–6) ($\rho = 0.43$), indicating that higher-J CO line luminosities are stronger correlated with L/M. This scenario can be further investigated with observations of even higher-J CO transitions (with upper J levels ranging from 11 to 16), which are now available using the SOFIA telescope.

3.7.3 The warming-up of the CO gas as a function of the evolutionary stages

The excitation temperature of the TOP100 clumps was derived in Sect. 3.5.5, based on the peak intensity of the CO (6–5) profiles convolved to 13.″4 ('observed' temperatures) and the peak of the Gaussian fit to the observed profile ('Gaussian' temperatures). The Gaussian T_{ex} values are expected to overcome the self-absorbed peaks present in most of the CO profiles towards the TOP100 clumps. Thus, in those cases where the Gaussian decomposition of the CO profiles was reasonably good at velocity channels far from the peak, it could provide a more realistic estimate of the temperature of the gas.

The Gaussian T_{ex} values range from 14-158 K, with a median value of 37 K. Similar results are found towards the T_{ex} values derived from the peak of the CO profiles (T_{ex} values ranging between 14 and 147 K, with a median value of 34 K).

The CO (6–5) temperature was compared with temperatures from other tracers: ammonia (NH_3) from Wienen et al. (2012); $C^{17}O$ (3–2) from Giannetti et al. (2014); methanol (CH_3OH), methyl acetylene (CH_3CCH), acetonitrile (CH_3CN) and the dust temperature

from Giannetti et al. (2017). The plots are presented in Fig. 3.15.

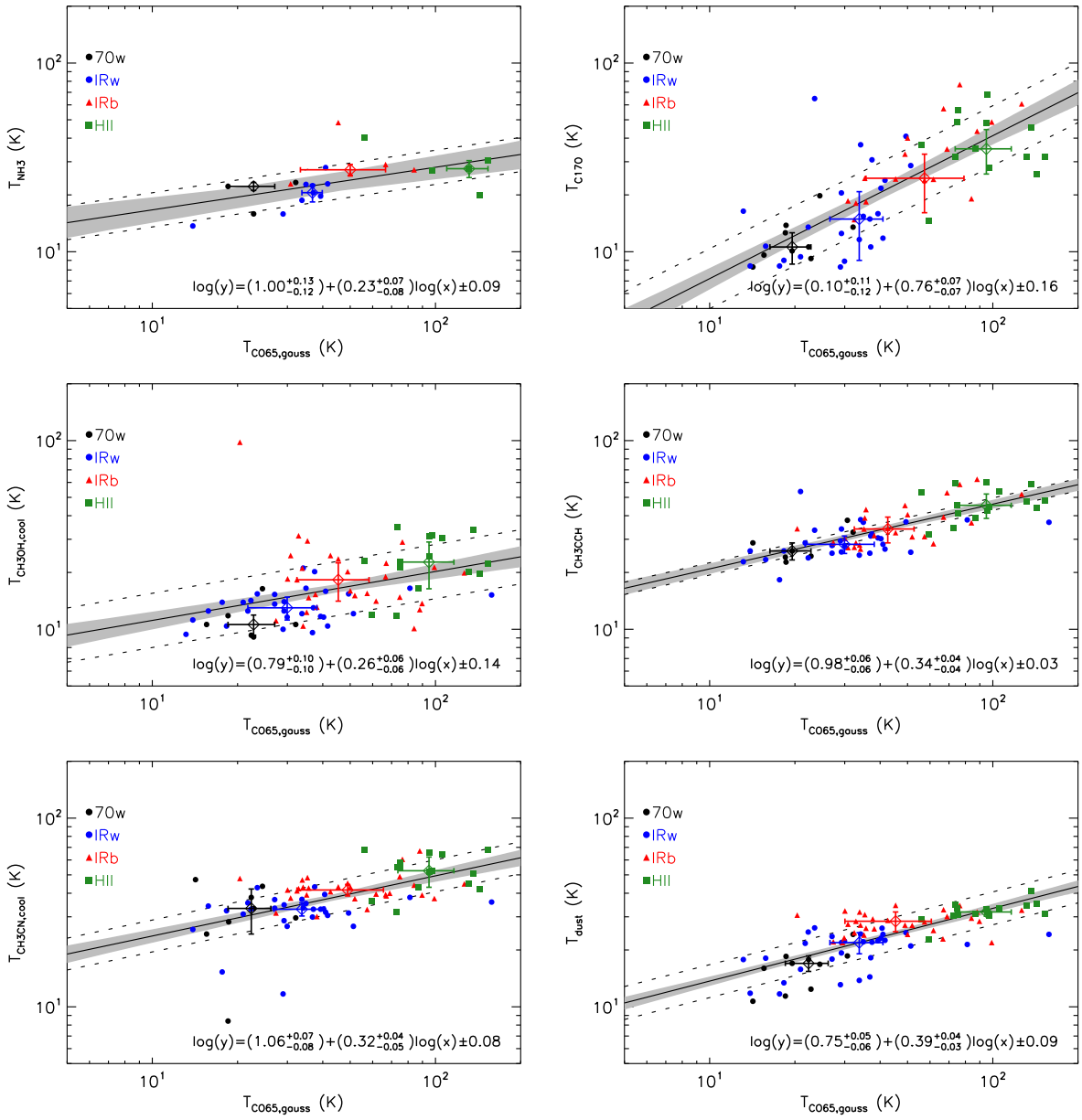


Figure 3.15: Correlation between the CO (6–5) excitation temperature versus the temperature of other tracers: from the top left to the bottom right panels: ammonia (NH_3), $C^{17}O$ (3–2), methanol (CH_3OH), methyl acetylene (CH_3CCH), the acetonitrile (CH_3CN) and dust. See text for references. For a complete description of the elements of the plots, see Fig. 3.5.

Table 3.6 summarises the Spearman rank factor comparing T_{ex} with the temperatures derived from the other tracers. In general, the temperatures are all correlated ($\rho \geq 0.44$), indicating a progressive warming up of distinct phases of the gas probed by different molecules.

Finally, the excitation temperature of the CO gas was plotted against the clump properties. Fig. 3.16 presents the T_{ex} values derived from the CO (6–5) line as a function of L_{bol}

Table 3.6 - Spearman rank correlation statistics between the excitation temperature of the CO towards the TOP100 clumps and the temperatures derived from other molecules.

	Observed	Gaussian
$T_{\rm NH3}$	0.67, $p<0.001$	0.68, $p<0.001$
$T_{\rm C17O}$	0.79, $p<0.001$	0.76, $p<0.001$
$T_{\rm CH3OH}$	0.46, $p<0.001$	0.44, $p<0.001$
$T_{\rm CH3CCH}$	0.75, $p<0.001$	0.70, $p<0.001$
$T_{\rm CH3CN}$	0.59, $p<0.001$	0.54, $p<0.001$
$T_{\rm dust}$	0.72, $p<0.001$	0.68, $p<0.001$

Notes: The comparisons are made for the excitation temperature of the CO (6–5) derived from the peak of the observed profiles and the peak of the Gaussian fit to the profiles. The rank ρ and its corresponding probability (p) are shown for each comparison. A p-value of <0.001 indicate a correlation at 0.001 significance level. p-values of 0.05, 0.002 and <0.001 represent the $\sim$2, 3 and $>$3-σ confidence levels.

and L/M. The distribution of $T_{\rm ex}$ versus $M_{\rm clump}$ is presented in Fig. 3.17. The distribution of the points and the associated Spearman coefficients indicate that $T_{\rm ex}$ is strongly correlated with the luminosity and the L/M ratio of the clumps ($\rho \geq 0.65$, $p<0.001$). However, $T_{\rm ex}$ exhibits no significant correlation with $M_{\rm clump}$ ($\rho = 0.07$, $p = 0.52$). When compared with the weak correlation found between the total CO luminosity with L/M (see right panel of Fig. 3.12), the stronger correlation observed between $T_{\rm ex}$ and L/M is an interesting result, indicating a progressively warming up of the CO gas as a function of the evolutionary stage of the clumps.

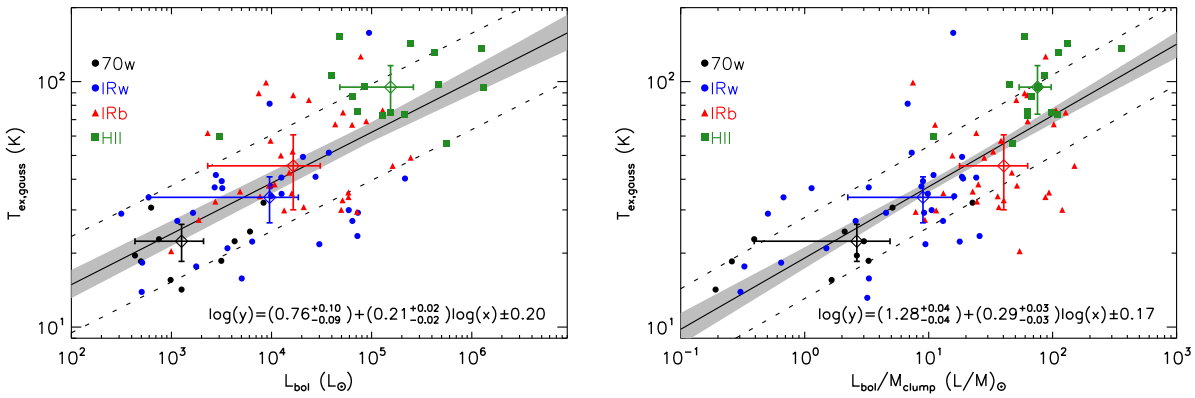


Figure 3.16: Excitation temperature of the CO (6–5) transition as a function of the bolometric luminosity (left panel) and the luminosity-to-mass ratio of the TOP100 clumps (right). The excitation temperature was derived using the peak of the Gaussian fit of the CO profiles. The median values for each class are shown as open diamonds and their error bars correspond to the median absolute deviation of the data from their median value.

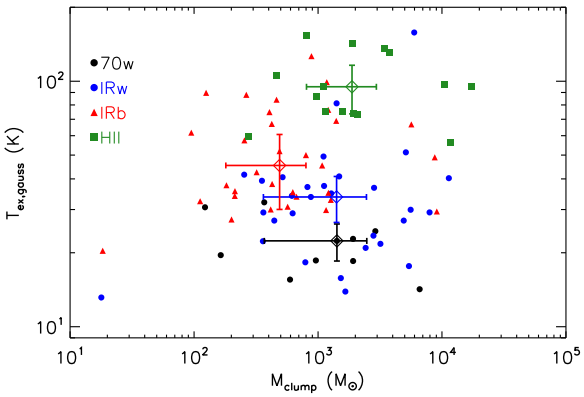


Figure 3.17: Excitation temperature of the CO (6–5) transition as a function of the mass of the TOP100 clumps (right). The full description of the elements of the plot is presented in Fig. 3.16.

3.7.4 Large-scale molecular outflows traced by the mid-J CO emission

The analysis of the mid-J CO emission presented in this Chapter indicates that high-velocity components are found in most of the spectra towards the TOP100 clumps (see Sect. 3.6.1. The high-velocity CO emission is often associated with the presence of molecular outflows that can be further exploited using the available data and high-angular interferometric observations from ALMA.

The spatial distribution of the outflows was obtained by integrating the high-velocity emission of the CO (6–5) line and, as illustrated in Fig. 3.18, integrated maps and CO spectra of the blue and red lobe were obtained. The maximum and minimum velocity of each lobe were defined by making sure that the observed integral map of each lobe probes a different structure than the one from the integrated CO (6–5) maps, presented in Fig. 3.9 and in ***Fig. B.1***.

The analysis of the high-velocity emission of the CO (6–5) datacubes has shown that about 85% (84 of 99 sources) of the sample exhibits outflow signatures, from 70w clumps to H II regions, indicating that the outflows correspond to a common phenomenon in every stage of the high-mass star formation process.

Figure 3.19 illustrates a few examples of the mid-J CO outflows identified towards the TOP100 clumps, overlaid on the Spitzer/IRAC mid-infrared maps. The maps indicate that the outflows can be either extended (*e.g.*, AGAL019.882−00.534) or compact/unresolved structures (*e.g.*, AGAL337.916−00.477). In most of the cases, the mid-IR counterpart also exhibits an excess of emission at 4.5 μm, appearing as a bright and extended green object in the false-colour RGB Spitzer/IRAC images, similarly as the Extremely Green Objects

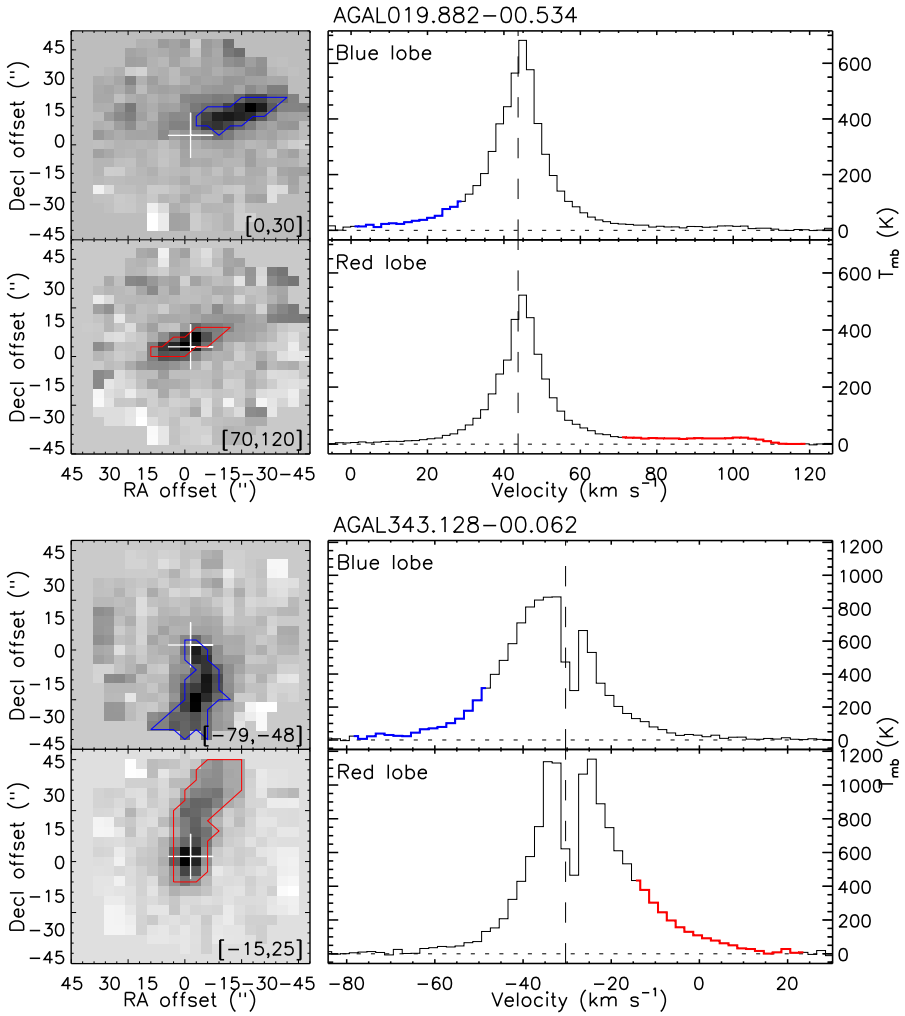


Figure 3.18: Derivation of the spatial structure of the CO (6–5) outflows associated with two TOP100 clumps. The integrated intensity map of the blue and red lobe emission are presented in the top and bottom left panels, respectively. The blue and red contours indicate the regions where the integrated fluxes of the lobes were extracted. The (0,0) position is marked as a white + symbol. The corresponding CO spectra, integrated within the spaxels delimited by the contours shown on the maps, are presented in the right panels. The velocity range of the blue and red lobes are indicated over the maps and shown in colour over the spectra. The vertical dashed black line is placed at the rest velocity ($V_{\rm lsr}$) of each source. The horizontal dashed line indicates the baseline of the data.

(EGOs) from the GLIMPSE catalogue (Cyganowski et al., 2008).

The physical parameters of the outflows were calculated in according to Beuther et al. (2002) and Yıldız et al. (2015). The high-velocity CO (6–5) flux was used to derive a series of physical parameters of the outflows, *i.e.*, the entrained mass ($M_{\rm out}$), the momentum ($P_{\rm out}$) and energy ($E_{\rm out}$), the mass-loss rate ($\dot{M}_{\rm out}$), the mechanical force ($F_{\rm out}$) and luminosity ($L_{\rm out}$).

Assuming that the emission at the wings of the CO profiles is optically thin ($\tau \ll 1$), the integrated flux of each lobe was used to compute its upper-level column density per

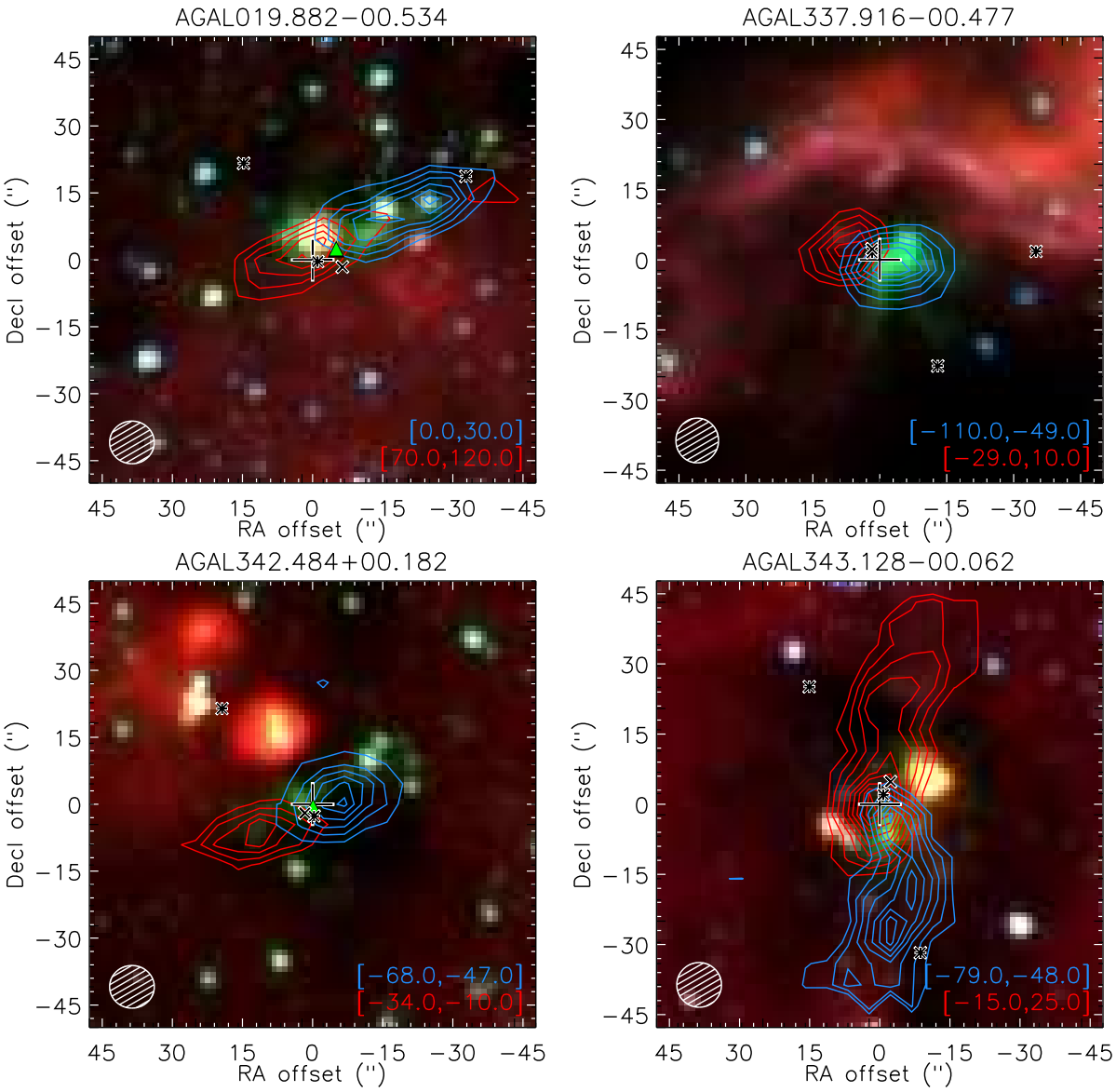


Figure 3.19: Examples of molecular outflows probed by the CO (6–5) emission towards the TOP100 sample. The false-colour RGB Spitzer/IRAC maps (blue: 3.6 μm; green: 4.5 μm; and red: 8.0 μm) are overlaid by the CO (6–5) contours of the blue- and red-shifted CO emission. The CO maps were integrated over the velocity range indicated in the bottom right region of the maps. The (0,0) position of the map is marked by the dashed horizontal and vertical lines, the position of the CSC source from Contreras et al. (2013) is shown as a × symbol and the dust continuum emission peaks from Csengeri et al. (2014) are shown as asterisks. The peak of methanol maser emission from Urquhart et al. (2013) is indicated by a green upward triangle symbol. The beam size of the CO (6–5) observations is indicated in the bottom left region of the maps.

statistical weight, $N_\mathrm{u}/g_\mathrm{u}$ (in cm^{-2}):

$$\frac{N_\mathrm{u}}{g_\mathrm{u}} = \frac{\beta\nu^2}{A_\mathrm{ul}g_\mathrm{u}} \int T_\mathrm{mb}\, dV \tag{3.17}$$

where $\beta = 8\pi k_\mathrm{b}/hc^3$, ν is the rest frequency of the transition, $\int T_\mathrm{mb}\, dV$ is the integrated intensity of the lobe (S_int). The total CO column density of each lobe (N_total, in cm^{-2}) is

obtained as:

$$N_{\rm total} = \frac{N_{\rm u}}{g_{\rm u}} Q(T_{\rm ex}) \exp\left(\frac{E_{\rm u}}{k_{\rm b} T_{\rm ex}}\right) \tag{3.18}$$

where $Q(T)$ corresponds to the partition function of the CO at a given temperature T, and $T_{\rm ex}$ is the excitation temperature of the gas (derived in Sect. 3.5.5). The uncertainty on each parameter was estimated by propagating the error of the integrated CO intensity (20% for most of the clumps, see details in Sect. 3.4) and the error of 20% in the distance.

The mass entrained by the outflow ($M_{\rm out}$) is calculated as

$$M_{\rm out} = \mu_{H_2} m_H A_{\rm pix} \left[\frac{H_2}{^{12}CO}\right] N_{\rm total} \tag{3.19}$$

where the molecular weight per hydrogen molecule, $\mu_{H_2} = 2.8$, includes the contribution of helium (Kauffmann et al., 2008), m_H is the mass of the hydrogen atom and $A_{\rm pix}$ is the area of a spaxel projected at the distance of the source (in cm^{-2}), and $[H_2/^{12}CO] = 1.2{\cdot}10^4$ corresponds to the abundance of H_2 over the number of ^{12}CO molecules (Yıldız et al., 2012).

The momentum and energy of the outflows ($P_{\rm out}$ and $E_{\rm out}$, respectively) are estimated by Eqs.(3.20) and (3.21):

$$P_{\rm out} = M_{\rm out} V_{\rm max} \tag{3.20}$$

$$E_{\rm out} = \frac{1}{2} M_{\rm out} V_{\rm max}^2 \tag{3.21}$$

Assuming the entrained gas moves with a constant velocity along the extension of the outflow, the dynamical timescale ($t_{\rm dyn}$) is given by:

$$t_{\rm dyn} = \frac{\ell_{\rm proj}}{< V_{\rm max} >} \tag{3.22}$$

where $\ell_{\rm proj}$ corresponds to the average size of the blue and red lobes; and, $< V_{\rm max} >$, the average between the relative maximum velocity of the blue and the red lobe in respect of the rest velocity of the source ($V_{\rm lsr}$).

The time-dependent parameters, the mass-loss rate ($\dot{M}_{\rm out}$), the outflow force ($F_{\rm out}$) and

the outflow luminosity ($L_{\rm out}$) are calculated using Eqs. (3.23), (3.24) and (3.25):

$$\dot{M}_{\rm out} = \frac{M_{\rm out}}{t_{\rm dyn}} \tag{3.23}$$

$$F_{\rm out} = \frac{p_{\rm out,b} + p_{\rm out,r}}{t_{\rm dyn}} \tag{3.24}$$

$$L_{\rm out} = \frac{E_{\rm out,b} + E_{\rm out,r}}{t_{\rm dyn}} \tag{3.25}$$

The preliminary results indicate that the outflow parameters are correlated with the bolometric luminosity and the mass of the clumps, while weaker correlations are found when comparing the outflow parameters with the luminosity-to-mass ratio of the clumps. The stronger correlations with $L_{\rm bol}$ and $M_{\rm clump}$ but not with L/M are similar to those observed when comparing the total CO line luminosity with the clump properties (see Sect. 3.6.3).

Figure 3.20 presents $M_{\rm out}$ as a function of the clump properties. The plots indicate that $M_{\rm out}$ increases as a function of the evolutionary class of the clumps (from 70 μm-weak clumps to H II regions) and as a function of the $L_{\rm bol}$ and $M_{\rm clump}$ values. The relations found between $M_{\rm out}$ *versus* the bolometric luminosity or the mass of the clumps (see top and middle panels of Fig. 3.20) display similar power-law distributions to those derived towards low-J CO observations from Beuther et al. (2002), Wu et al. (2004) and Maud et al. (2015). In addition, the mid-J CO analysis also extends the validity of these relations towards more luminous objects with $L_{\rm bol} \gtrsim 10^6\,{\rm L}_\odot$. Despite the agreement between the slope of the trends obtained for different J transitions, they are offset by roughly two orders of magnitude, indicating that the low-J CO is probing a larger amount of gas when compared to the CO (6–5) transition.

Similar results are also found among the other parameters of the outflows *versus* the clump properties (see Fig. 3.21 for the distribution of the outflows parameters versus $L_{\rm bol}$). The most discrepant results are found for the mechanical force of the outflows. For the TOP100 sample, the derived power-law index (0.74±0.10) is shallower than the one reported by Yıldız et al. (2015, 1.13 ± 0.37, indicated by the green line in Fig. 3.21), who investigated a sample of low-luminosity YSOs in the CO (6–5) transition. The steeper relation found towards less luminous YSOs suggests that the dependency between the mechanical force of the outflows and the properties of their driven source decreases from low-

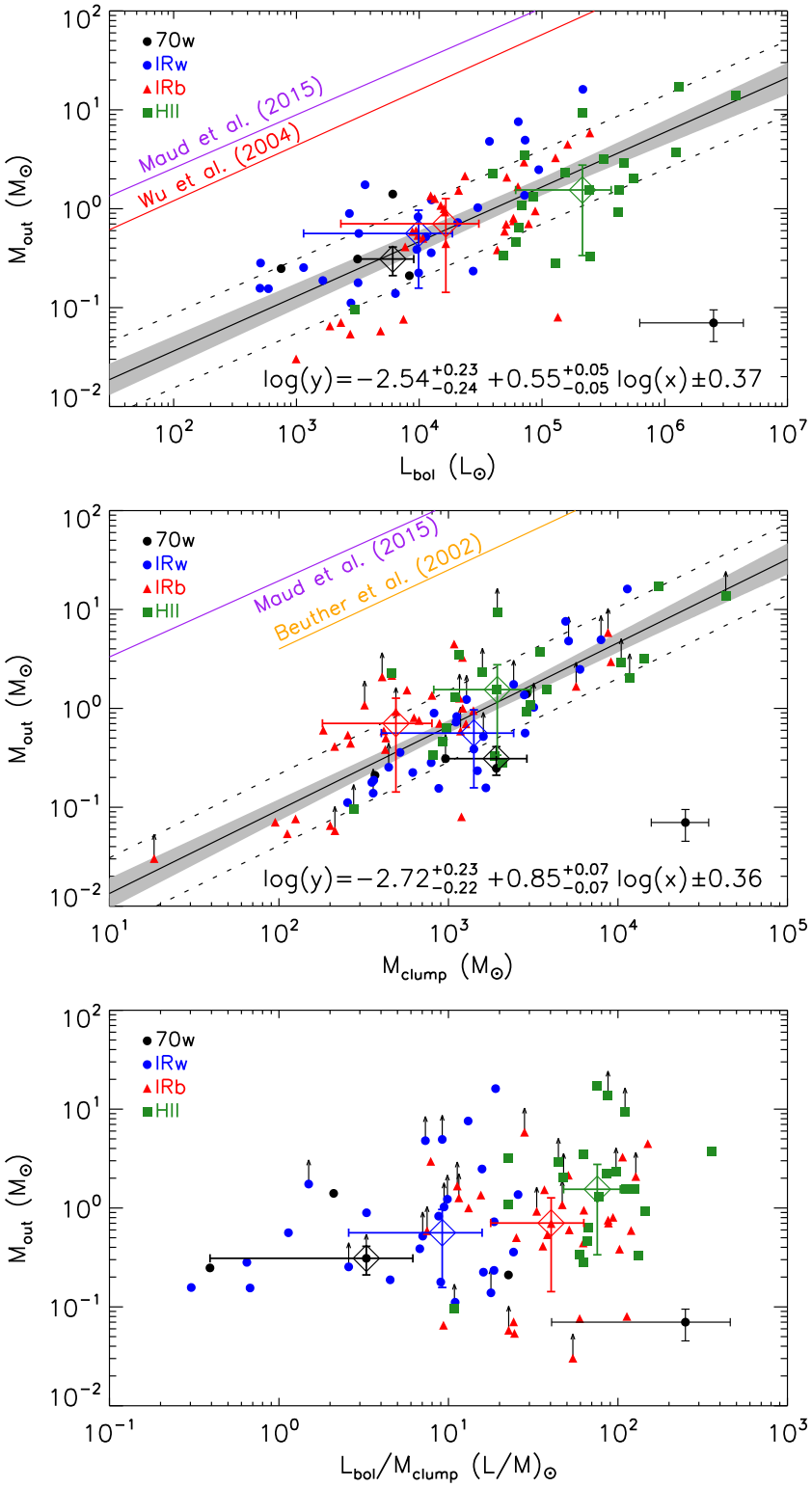


Figure 3.20: The dependency of the mass entrained by the outflows as a function of the bolometric luminosity (top panel), the mass (middle) ad the luminosity-to-mass ratio of the TOP100 clumps (bottom). The filled yellow, red and purple lines correspond to fits of low-J CO observations from Beuther et al. (2002), Wu et al. (2004) and Maud et al. (2015), respectively. For a complete description of the elements in the plots, see Fig. 3.5.

to high-luminosity regimes.

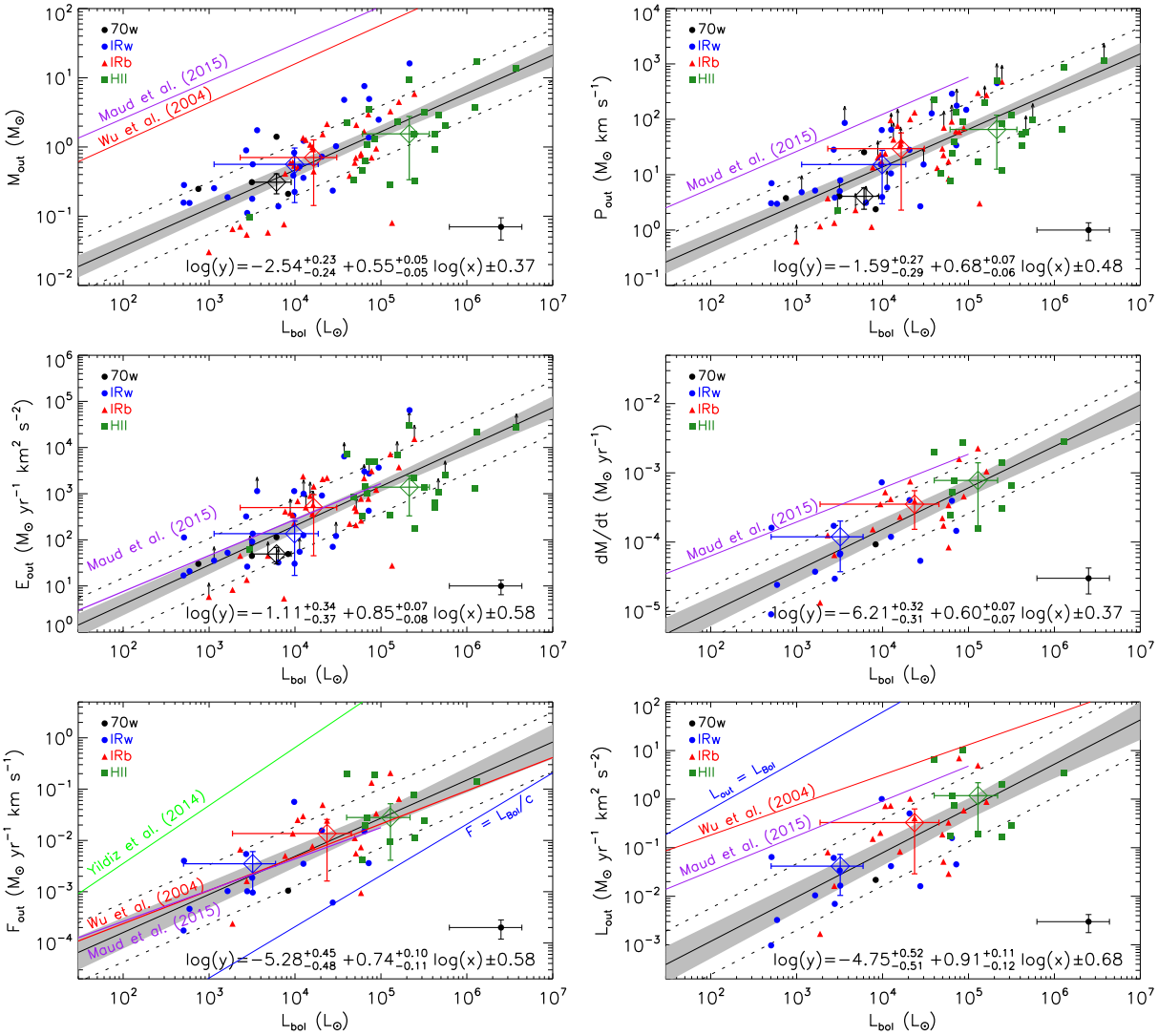


Figure 3.21: The dependency of the inclination-dependent outflow parameters as a function of the bolometric luminosity of the TOP100 clumps. The filled green, red and purple lines correspond to fits of low-J CO observations from Yıldız et al. (2015), Wu et al. (2004) and Maud et al. (2015), respectively. The blue lines correspond to the radiative force and $L_{\rm bol}$ =$L_{\rm out}$ lines in the bottom panels. For a complete description of the elements in the plots, see Fig. 3.5.

The increment on the values of the parameters of the outflows as a function of the clump properties can be interpreted as the result of the significant star-forming process from the whole stellar population being formed within the clumps, which are likely to form multiple-stellar system.

The properties of the outflows associated with the TOP100 clumps will also be compared to those observed towards less luminous sources, such as presented by Yıldız et al. (2015). This analysis will help to identify whether the large-scale properties of the outflows are scaled or not in a broader range of bolometric luminosities.

The identification of the outflows in the TOP100 sample delivered a subsample of high-luminosity clumps with clear outflow activity to be further investigated using, *e.g.*, high angular resolution interferometric data from ALMA.

Chapter 4

Integral field spectroscopy of HMYSOs

This chapter presents the *K*-band integral field spectroscopy of seven high-mass young stellar objects (HMYSOs) associated with large-scale H_2 outflows. Section 4.1 describes the selection criteria of the sample, the observations are presented in Sect. 4.2 and Sect. 4.3 presents the observational results, together with the derivation of the physical parameters of the central sources and their extended emission. The results are discussed in Sect.4.4.

4.1 The sample

The sources analysed in this Chapter were initially selected from a sample of HMYSO candidates from the Red MSX Source survey (RMS, Lumsden et al., 2013) catalogue, exhibiting large-scale H_2 jets identified by Navarete et al. (2015, hereafter, Nav15). Those authors found that 21% (74 sources) HMYSO candidates from a sample of 353 objects are associated with polar-like H_2 emission as expected for bipolar molecular jets. These 74 objects define a sample of good candidates to investigate in details the structure and physical parameters of the molecular jets closer to their origin, together with the characterisation of the circumstellar environment of these HMYSOs. The following additional selection criteria were used to select seven HMYSOs exhibiting large-scale H_2 jets from Nav15 to follow-up high-angular resolution spectroscopic using the Gemini North Telescope:

1. Bolometric luminosity greater than $10^3\,L_\odot$, corresponding to a B3 V_0 star or earlier (Panagia, 1973);

2. Association with large-scale polar-like H_2 emission from Nav15;

3. Accessible from the Gemini North telescope ($-25^\circ \geq \delta \geq +65^\circ$)

4. Sufficiently closer ($\leq 15''$) to a bright ($8.5 < \mathrm{R} < 15$ mag) tip-tilt star for a high Strehl Correction (Adaptive Optics system requirement).

The properties of the sources, extracted from the RMS database[1] are summarised in Table 4.1. The well-known source W33A, investigated by Davies et al. (2010), was also included in the analysis presented in this Chapter to validate the methodology and the results obtained for the other HMYSOs.

Table 4.1 - Properties of the HMYSOs.

ID	MSX Name	L_{bol} ($\mathrm{L}_{\odot}$)	Distance (kpc)	V_{lsr} ($\mathrm{km\,s^{-1}}$)	Ref.
W33A	G012.9090−00.2607	3.2×10^4	2.4	$+35.3^a$	D10
N33	G079.1272+02.2782	1.6×10^3	1.4	$+1.7^b$	Nav15
N58	G094.4637−00.8043	2.1×10^4	4.9	-44.6^a	Nav15
N35	G100.3779−03.5784	1.7×10^4	3.7	-37.6^a	Nav15
N16	G108.7575−00.9863	1.4×10^4	4.3	-51.5^a	Nav15
V49	G109.0974−00.3458	2.3×10^4	3.8	-46.7^b	Nav15/V10
N47	G124.0144−00.0267	1.8×10^3	3.1	-43.4^b	Nav15
N21	G173.6328+02.8064	3.1×10^3	2.0	-20.0^a	Nav15

Notes. The columns are as follows: (1) Short name for each source; (2) Designation of the source based on the MSX catalogue; (3) bolometric luminosity; (4) kinematic distance; (5) the rest velocity, the footnote indicate where the given velocity was obtained from: *a*) NH_3 (1,1) or *b*) CS (2–1) observations (for more information see Lumsden et al., 2013); (6) reference: D10 - Davies et al. (2010), V10 - Varricatt et al. (2010b), Nav15- Navarete et al. (2015). Columns (2)-(5) are extracted from the RMS database (Lumsden et al., 2013).

The large-scale continuum-subtracted H_2 maps of the sources are presented in Fig. 4.1. A summary of the H_2 structures identified in each FOV is given in Table 4.2.

4.2 Observations and Data processing

The *K*-band ($\lambda \approx 2.2\,\mu$m) spectra were obtained with the *Near-Infrared Integral Field Spectrometer* (NIFS, McGregor et al., 2003) during the second semester of 2011. The data were taken in queue-scheduled mode (as part of Gemini observing program GN-2011B-Q25). Data for the source W33A, part of the project GN-2008A-Q-44, were downloaded from the Gemini Science Archive and processed using the same methodology presented in this section. NIFS has a full wavelength coverage it the *K*-band of about 4200 Å, with a linear dispersion of $\Delta\lambda_{\mathrm{disp}} = 2.13$ Å pixel^{-1} and a nominal spectral resolving power of $\lambda/\Delta\lambda_{\mathrm{disp}} = 5160$ (Blum and McGregor, 2008), or $\approx 58\,\mathrm{km\,s^{-1}}$ in the velocity scale.

[1] rms.leeds.ac.uk

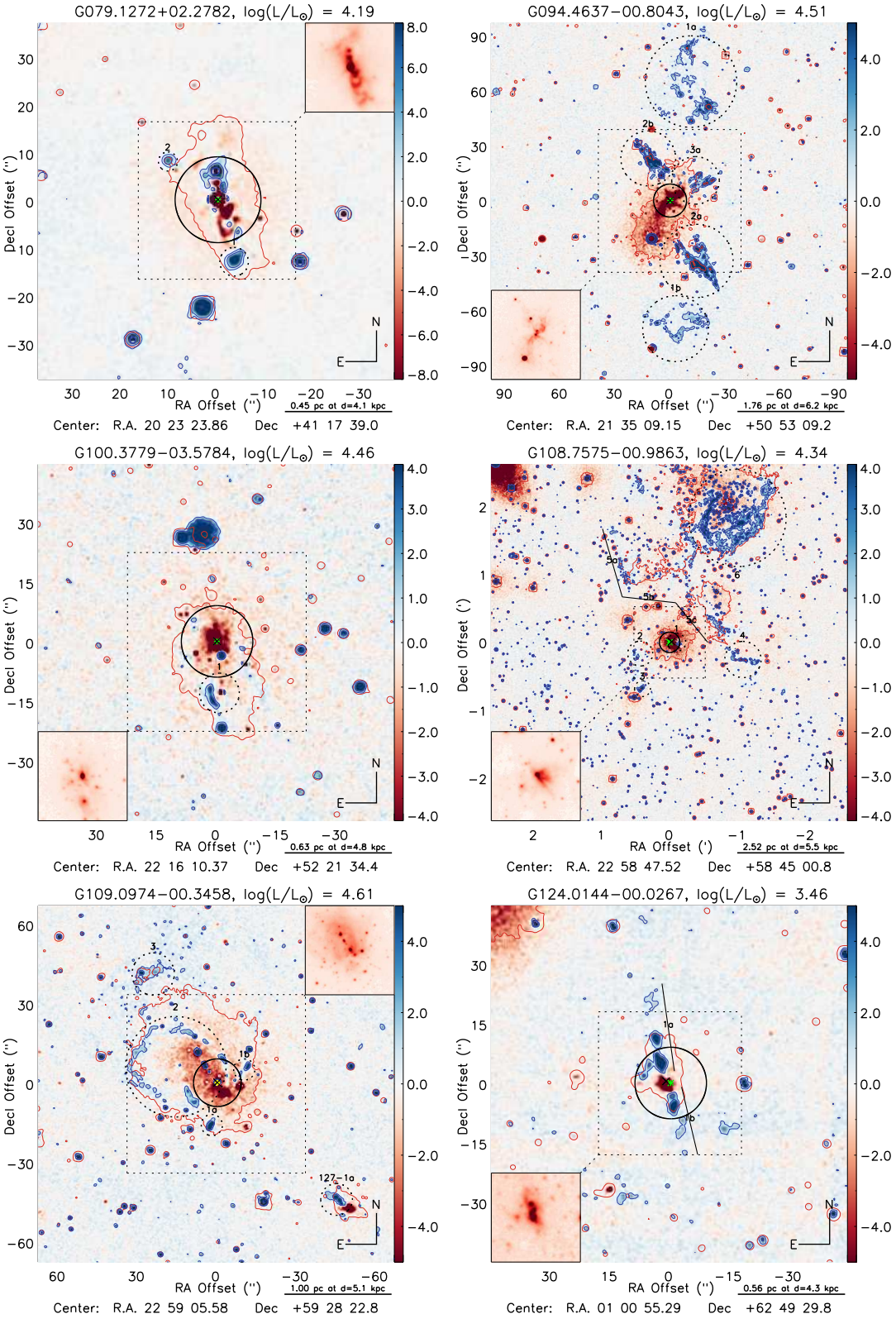


Figure 4.1: Continuum-subtracted H_2 maps of the sources. Each image is scaled by the rms of the continuum-subtracted map. Excess H_2 emission is shown in blue while excess in the continuum filter is shown in red. Radio-quiet/loud RMS sources are represented by a green $*$ or a yellow $\diamond$ symbol, respectively. The inset frame displays the details in the continuum filter of the region delimited by the dashed box in the main map. Extracted from Navarete et al. (2015).

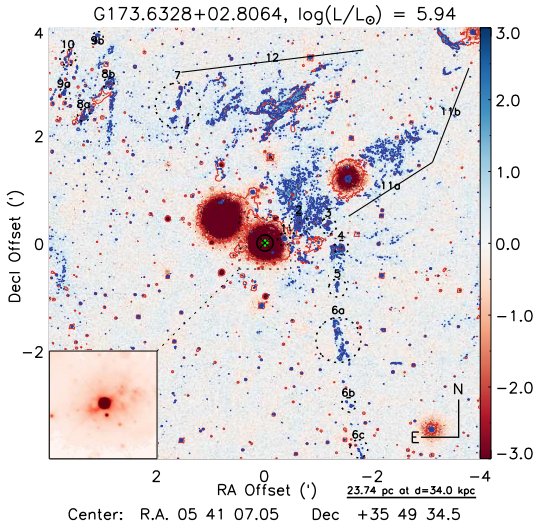


Figure 4.1: Continued.

Table 4.2 - Properties of the large-scale H_2 emission associated with the HMYSOs.

ID	Class	ℓ_{proj} (pc)	PA (°)	R	Comments
N16	BP1	0.32	270	1.1	–
N21	K	3.30	275	–	H_2 emission is not likely associated with the source
N33	BP2	0.28	195	3.4	bright knot located to the SW direction of the MSX source
N35	K	0.44	170	–	–
N47	BP2	0.48/0.38	10/200	8.0/6.3	–
N58	BP5	2.70/2.38	355/180	3.5/3.3	Outer outflow oriented in the S-N direction
		1.66/1.24	210/30	3.2/3.0	Intermediary outflow oriented in the SW-NE direction
		0.78	300	2.5	Inner jet oriented in the NW direction (no red counterpart)
V49	BP2	0.39/0.39	150/320	5.2/7.8	–
W33A	BP1	0.56	135	2.4	Elongated continuum structure in the SE direction of the source as reported by D10.

Notes. The columns are as follows: (1) ID of the source; (2) the classification of the H_2 emission: BPn for a n-polar outflow or K for knot; (3) projected length of the structure. In case of bipolar structure, the values of each lobe are shown; (4) the position angle of the structure, from N to E; (5) the collimation factor of the outflow (R=length/width); (6) any comments on the large-scale H_2 structure. These information were extracted from Navarete et al. (2015).

NIFS was used with the *Altitude conjugate Adaptive optics for the InfraRed* (ALTAIR) module in laser guide star (LGS) mode, providing spatial resolutions closer to the diffraction limit of the telescope (about $\approx 0\farcs068$ at the *K*-band for an 8-m telescope). The ALTAIR LGS mode consists on the propagation of a beam of coherent radiation at 589 nm

into the sky. The laser beam excites sodium atoms at ∼90 km altitude. The excited atoms re-emit and produce an artificial star on which ALTAIR can guide for high-order corrections. A natural guide star is still required for guidance, focus and tip/tilt correction. Natural stars with $R < 15$ mag and located within 15″ from the target may be used for high-Strehl LGS correction, while natural stars with $R < 18.5$ mag and within 25″ still provides a low-Strehl LGS correction. Figure 4.2 presents the K_s-band ($\lambda_0 = 2.159\,\mu$m, $\Delta\lambda = 0.262\,\mu$m Cohen et al., 2003) maps from the *Two-Micron All Sky Survey* (2MASS, Skrutskie et al., 2006) overlaid by the reference sources and positions used for the observations.

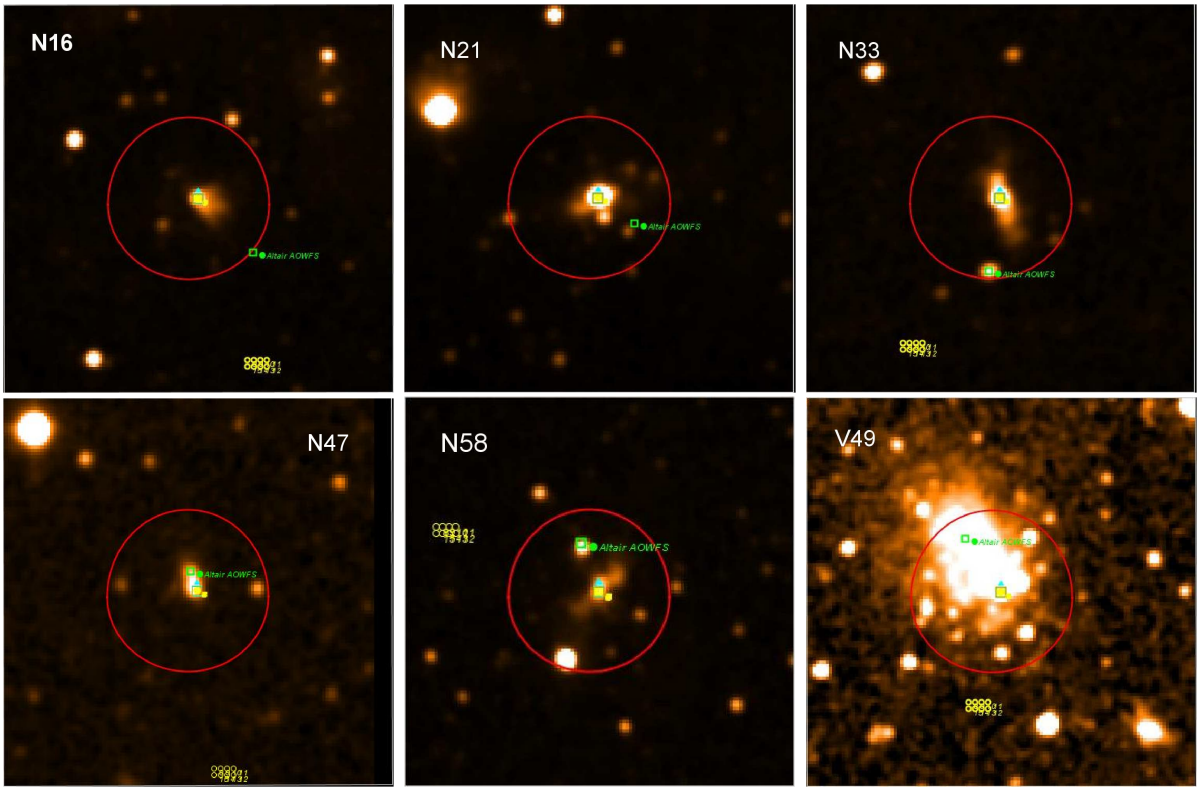


Figure 4.2: 2MASS K_s-band maps of the sources observed with NIFS. From the left top to the right bottom: N16, N21, N33, N47, N58 and V49. The position of the tip/tilt star is indicated by the green circle, the pointings for the sky observation are shown as yellow circles (numbered from 1 to 8), and the NIFS FOV (3″ × 3″) is indicated by the grey box in the centre of the field. The red circle indicates the 15″ limit for the ALTAIR natural guide star.

The on-source and off-source observations were performed in an ABAABA[2] pattern. The off-source positions were set a nearby sky field (∼30″, see Fig. 4.2) with the same exposure time (t_{exp}) of the on-source positions. Each source was observed using an 8-points dithering pattern, distributed in a 4 × 2 box layout. The on-source frames were obtained by dithering the telescope by 0.″064 along the slits and 0.″126 across them (Fig. 4.3). The

[2] A and B correspond to on- and off-source positions, respectively.

adopted dithering pattern maximises the point spread function (PSF) sampling of the datacubes. This is a fundamental step to improve the angular resolution of the final data using the a Richardson-Lucy deconvolution procedure (Steiner, priv. conv.).

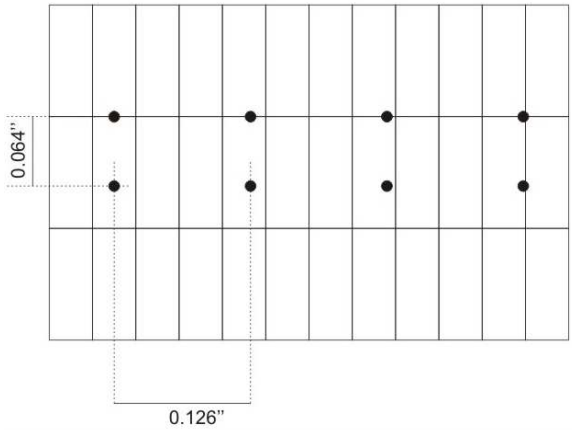


Figure 4.3: Dithering pattern used for the NIFS observations. The NIFS pixel scale corresponds to $0.103 \times 0.043\,\mathrm{arcsec\,pixel^{-1}}$ across and along the directions of the slits, respectively. The central position of each pointing is indicated by a filled circle, spaced by $0\farcs126$ and $0\farcs064$ along the x- and y-directions, respectively.

The properties of the sources and the stars used as telluric standards are listed in Tables 4.3 and 4.4, respectively. Due to problems in the processing, the telluric standard observed for V49 was not used and HIP 7291, observed at similar airmass conditions, was adopted as its telluric standard instead. The typical seeing of the observations, estimated based on the full width at half maximum (FWHM) of the telluric standard stars, is about $0.15''$. The signal-to-noise ratio (S/N) values of the science observations are reported in Table 4.3. The S/N ratio was estimated in a line-free region of the continuum between 2.20 and $2.275\,\mu$m, exhibiting values close to 100 in most of the observations. The relatively high S/N ratio values ensure that weak lines are detectable along the entire spectral range, such as the low-J CO absorption features at $\approx 2.35\,\mu$m.

As part of the calibration procedure, the following additional set of images were taken for processing the IFU data:

1. a set of flat-field images taken with Quartz lamps on (flat-on) and off (flat-off);

2. flat-field images with lamps on and a Ronchi mask (Ronchi flats), used to correct the spatial distortion along the spectral dimension;

3. a set of dark images, taken with the same t_{exp} of the science and telluric sources.

Table 4.3 - Observational properties of the HMYSO sample observed with the Gemini North telescope.

ID	RA (J2000)	Decl. (J2000)	K_s (mag)	t_{exp} (s)	$t_{exp,total}$ (s)	Date	Air Mass	S/N	Δp (AU/pix)
N16	22:58:47.38	58:45:01.9	10.0	23	184	2011-09-13	1.32	130	215
N21	05:41:07.04	35:49:34.8	8.0	8	64	2011-10-16	1.05	80	100
N33	20:23:23.83	41:17:39.3	9.6	23	184	2011-09-13	1.21	100	70
N35	22:16:10.37	52:21:34.4	10.5	200	1600	2011-12-09	1.30	100	185
N47	01:00:55.27	62:49:29.4	11.5	150	1200	2011-10-11	1.44	90	155
N58	21:35:09.11	50:53:09.6	11.1	150	1200	2011-09-13	1.19	150	245
V49	22:59:05.12	59:28:19.3	14.2	450	3600	2011-12-06	1.38	150	190
W33A[a]	18:14:39.51	−17:52:00.1	10.6	180	1980	2008-05-25	1.29		120

Notes. The columns are: (1) Short name of the source from Table 4.1; (2)-(3) central coordinates of the observed field; (4) 2MASS K_s-band magnitude; (5) individual exposure time (in seconds); (6) total exposure time (in seconds); (7) date of the observation (aaaa-mm-dd); (8) mean airmass of the observations; (9) signal-to-noise ratio of the brightest pixel of the point-like object in each FOV, measured in the continuum between 2.200 and 2.275 μm; (10) projected pixel scale at the distance of the source (in AU/pixel). *a* Observations from the project GN-2008A-Q-44 using the data from Davies et al. (2010).

Table 4.4 - Properties of the telluric standards observed as part of the observations.

ID	Telluric Standard	Spectral Type	K_s (mag)	t_{exp} (s)	Seeing ($''$)	Mean Airmass	Airmass difference[a]
N16	HIP 114332	A0 V	7.90	120.0	0.149	1.37	−0.05
N21	HIP 32268	A2 V	6.04	30.0	0.156	1.10	−0.05
N33	HIP 102074	A0 V	8.00	135.0	0.156	1.22	−0.02
N35	HIP 114714	A0 V	6.32	39.0	0.157	1.25	+0.05
N47	HIP 7291	A2 V	7.52	75.0	0.158	1.43	+0.01
N58	HIP 109079	A0 V	6.56	39.0	0.153	1.13	+0.06
V49	HIP 117450	A0 V	6.33	30.0	0.157	1.39	−0.01
W33A	HR 6798	A2 V	5.87	10.0	0.150	1.45	−0.16

Notes. The columns are: (1) Short name of the science source from Table 4.1; (2) Name of the telluric standard; (3) Spectral type of the telluric star; (4) 2MASS K_s-band magnitude; (5) total exposure time (in seconds); (6) average seeing (in arcseconds), measured as the FWHM of the telluric star; (7) mean airmass of the telluric observations; (8) difference between the average airmass of the science observations and the corresponding telluric standard.

4. Argon and Neon lamp spectra (Ar-Ne) for wavelength calibration of the data.

The flat-field and dark images were taken at the beginning of the night while the Ar-Ne spectra were taken right after the science observations.

4.2.1 NIFS data processing

The standard data processing of the NIFS observations was performed using the Gemini/NIFS `IRAF`[3] package. Calibration data were prepared with Gemini-`IRAF` `NFPREPARE` task which adds the Data Quality and Variance extensions to each FITS file. A flat-field template was created using the set of flat-on and flat-off images. The wavelength transformation was performed using Ar-Ne spectrum, obtained after each observation. The rms of the wavelength solution was ≤ 0.15 Å for each slit. The spatial rectification model was made using a Ronchi mask image (for more information, see Blum and McGregor, 2008).

After the creation of the calibration templates, the telluric and science observations were prepared with `NFPREPARE`, were flat-fielded, wavelength corrected, and sky subtracted. The spectral template of the telluric lines was created by the extraction of a one-dimensional spectrum from a 0.5″ radius area centred on the telluric standard star. Then, the Brγ feature was removed by fitting a Voigt profile between two continuum points in the telluric spectrum. Next, the telluric correction was applied to both the science and the telluric star itself using the `NFTELLURIC` task. Finally, the datacube was created by resampling the original rectangular spatial pixels to squared pixels of 0.″05 in both x- and y-directions. At the end of the processing, final datacubes with dimensions of $\sim 60 \times 60$ spatial pixels and 2040 spectral pixels were delivered.

4.2.2 Post-processing of the datacubes

The datacubes created by the standard `IRAF` reduction pipeline were further processed using a set of routines written in `IDL`. The purpose of these additional procedures includes *i*) the atmospheric refraction correction of the data, *ii*) the improvement of the signal-to-noise of the data, and *iii*) the flux-calibration of the final datacubes.

4.2.2.1 Atmospheric refraction correction

The atmospheric refraction (AR) is an optical effect caused by the change of the refraction index across the passage of the light of the source from the top of the atmosphere to the focal plane of the instrument. Although the AR is neglected on long-slit spectroscopic

[3] IRAF is distributed by the National Optical Astronomy Observatory, which is operated by the Association of Universities for Research in Astronomy, Inc., under cooperative agreement with the National Science Foundation.

data, it appears as a significant effect on IFU observations as it introduces a shift of the source in the (x, y) plane as a function of wavelength. Figure 4.4 illustrates the AR effects in the datacube of the telluric standard HIP 102074 before and after the AR-correction. The top panel indicates the existence of a systematic variation of the position in the plane of the sky with a position angle (PA) of $\sim 335°$ from longer wavelengths towards shorter ones. The offsets in the x- and y-directions range from -1.0 to +1.5 mas and -0.8 to +0.8 mas, respectively. In general, such offsets correspond to a small fraction of the pixel size (*e.g.*, 2.5 mas corresponds to a fraction of 0.05 of the NIFS pixel size) and can be attenuated to scales smaller than ~ 1 mas, as presented by the lower panel of Fig. 4.4.

The AR-correction was performed using the `IDL` routine `REFRACTION_CORRECTOR_PLUS`, which is part of the package of post-processing routines from Steiner et al. (2009) and available in `http://www.astro.iag.usp.br/~pcatomography/`. The code performs a n-th order polynomial fit to the position of the source on x- and y-direction as a function of λ. For NIFS datacube, a n=3 polynomial fit was applied to the data ($x(\lambda) = \Sigma_{n=0}^{n=3} A_n x^n$, where A_n is the n-th coefficient of the polynomial). The coefficients of the fit were obtained by measuring the position of the central source in four spectral regions, centred at $\lambda \approx 2.040$, 2.145, 2.253 and 2.236 μm, respectively. Then, the coefficients for both x- and y-directions were provided as input for the `IDL` code. At the end of the processing, the routine delivered a set of AR-corrected datacubes that were combined as described as follows.

4.2.2.2 Datacube combining

The AR-corrected datacubes were combined using the `IDL` routine `CUBE_COMBINE_PLUS`, also part of the routines from Steiner et al. (2009). The code computes and delivers either the sum, the average or the median of a list of input datacubes. The combination of the input AR-corrected datacubes by their median was chosen to suppress any bad pixels or possible cosmic rays present in the original data.

4.2.2.3 Flux calibration of the datacubes

The last procedure applied to the datacubes consisted in the flux calibration of the observed counts. For this, the `IDL` routine `IFU_fluxcal` was written to derive *i*)the instrumental sensitivity function based on the telluric standard data; *ii*)the flux calibration using a black-body approximation for the telluric spectra; and *iii*)the atmospheric extinc-

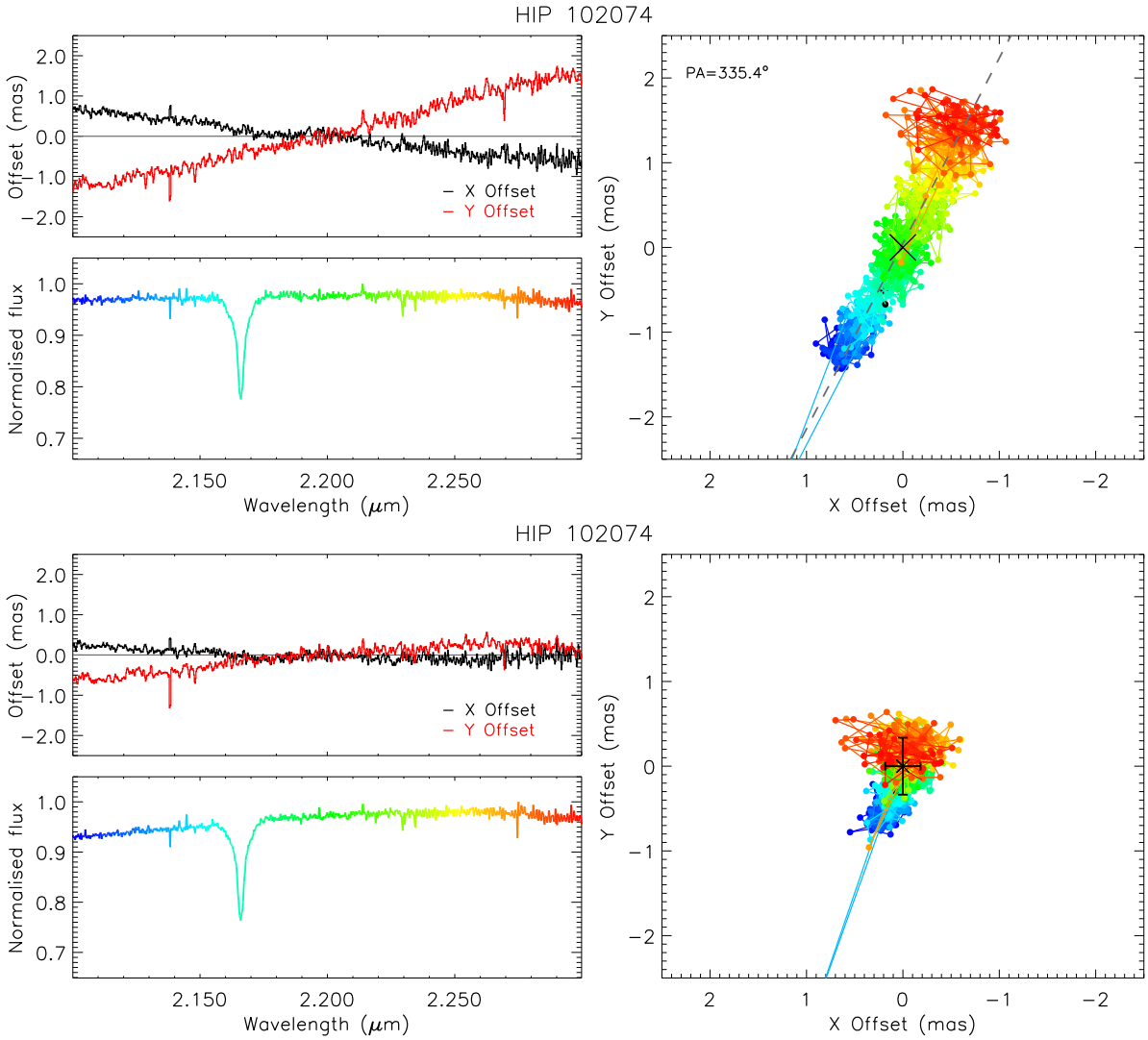


Figure 4.4: Spatial variation of the centroid position towards the telluric standard HIP 102074 before (top panel) and after the atmospheric refraction correction (bottom). Top left panel: the variation of the centroid of the emission in the x- (in black) and y-axis (in red) as a function of the wavelength. The baseline is indicated by the horizontal grey line. Bottom left: Normalised average spectrum of the source. The colour scale indicates the spectral counterpart of the spatial position shown in the map on the left. Right panel: Spatial variation of the centroid as a function of the wavelength (λ). The (0,0) position corresponds to the median position of the source along the whole spectral range. In the top panel, the dashed grey line indicates the direction of the maximum variation of the centroid position as a function of λ. In the bottom panel, the error bars placed at (0,0) indicates the standard deviation of the centroid position along λ after the AR-correction.

tion correction of the telluric standard. The code relies in the procedures of the `STANDARD`, `SENSFUNC` and `CALIBRATE` routines from the `IRAF`-NOAO package. The three corrections performed by each routine are applied simultaneously by computing the 'calibration factor' function (C_λ), given by:

$$C_\lambda = 2.5 \log \left(\frac{O_\lambda}{t_{\rm exp}\ B_\lambda(T_{\rm eff}, m_{Ks})} \right) + A\, E_\lambda \qquad (4.1)$$

where, O_λ corresponds to the observed counts in a bandpass of the observation (in ADU), $t_{\rm exp}$ is the exposure time of the observation (in seconds), $B_\lambda(T_{\rm eff}, m_{Ks})$ is the black-body flux (in erg s^{-1} cm^{-2} μm^{-1}) at a given effective temperature $T_{\rm eff}$, scaled by the $K_{\rm s}$ magnitude of the telluric standard, A is the average airmass of the observation and E_λ is the atmospheric extinction curve (in mag airmass^{-1} units).

The construction of O_λ was performed by computing the integral one-dimensional spectrum of the telluric standard star over the spaxels with average flux above 3-σ. Then, the continuum emission was modelled by fitting a polynomial function of order $n \leq 4$, using an outlier-resistant algorithm[4], as illustrated in the top panel of Fig. 4.5. Next, $B_\lambda(T_{\rm eff}, m_{Ks})$ was constructed by following a two-steps procedure. First, the shape of the black-body flux was obtained by adopting the effective temperature corresponding to the spectral type of the telluric standard (from Table 4.4). The list of effective temperatures as a function of spectral types from Tokunaga (2000) was adopted. Then, the black-body emission was scaled using the $K_{\rm s}$-band flux of the telluric star. The 2MASS $K_{\rm s}$-band magnitudes ($m_{\rm Ks}$) were converted to fluxes ($F_{\rm Ks}$) using Eq 4.2:

$$m_{Ks} = -2.5 \log \left(\frac{F_{Ks}}{F_{Ks,0}} \right) \tag{4.2}$$

where $F_{Ks,0} = 4.28(81) \cdot 10^{-7}$ erg s^{-1} cm^{-2} μm $^{-1}$ is the zero-point flux at the $K_{\rm s}$-band (Cohen et al., 2003), and the number in parentheses corresponds to the uncertainty on its last two digits.

Finally, the additive term of Eq. (4.1) requires a model for the atmospheric extinction as a function of the wavelength. Thus, the mean extinction through the broad K-band filter for Mauna Kea, $E_\lambda = 0.033$ mag/airmass (Tokunaga et al., 2002), was adopted.

After constructing C_λ (see middle panel of Fig. 4.5), each spatial pixel of the datacube (here, represented as $D_{xy\lambda}$) was divided by the exposure time of the observations, and was extinction-corrected and flux-calibrated as indicated by Eq. (4.3):

$$D_{fcal,xy\lambda} = 10^{0.4 \cdot C_\lambda} \cdot \left(\frac{D_{xy\lambda}}{t_{\rm exp}} \right) \tag{4.3}$$

The resulting calibrated spectrum is presented in the bottom panel of Fig. 4.5, where the telluric standard HIP 109079 is overlaid on top of the corresponding black-body profile

[4] the `IDL` routine `robust_poly_fit` was used.

for the corresponding effective temperature of the star, $T_{\rm eff} = 9480\,{\rm K}$, and scaled by the observed $K_{\rm s}$-band flux of the telluric standard star.

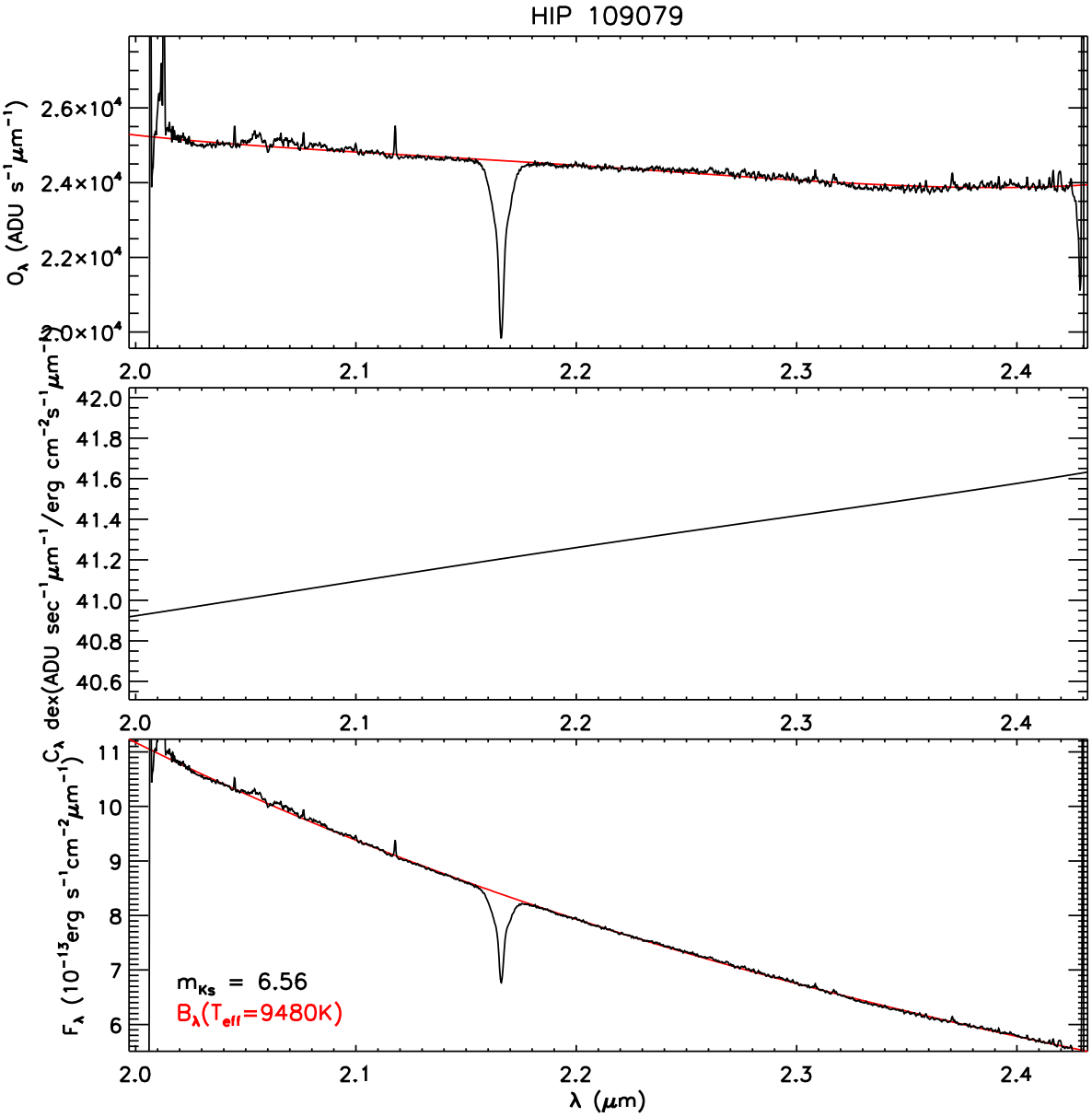


Figure 4.5: Derivation of the calibration factor for the telluric standard star HIP 109079. Top panel: the integrated spectrum of the source overlaid by the polynomial fit of the stellar continuum. Middle panel: the logarithm of the calibration factor function for the telluric standard. Bottom panel: The flux-calibrated spectrum overlaid on top of the black-body curve calculated for the corresponding effective temperature of the star and scaled by the observed $K_{\rm s}$-band flux of HIP 109079 (the $T_{\rm eff}$and $m_{\rm Ks}$ values are indicated in the bottom left corner of the plot).

The results of the flux calibration were checked using the telluric standard data as follows. First, the total flux of the telluric standards were integrated over the bandwidth of the 2MASS $K_{\rm s}$-band filter ($\lambda_c = 2.159\,\mu{\rm m}$, $\Delta\lambda = 0.262\,\mu{\rm m}$, Cohen et al., 2003). Then, the corresponding $K_{\rm s}$-band magnitude was derived using Eq. (4.2) to be compared with

the 2MASS K_s-band magnitude from Table 4.4. Figure 4.6 presents the residuals between the calibrated magnitudes and the 2MASS values. The absolute values are systematically shifted towards negative values, displaying a median value of $\Delta K_s = -0.03$ mag, corresponding to a K_s-band flux ratio of ≈1.028 between the NIFS and 2MASS fluxes. However, the associated errors indicate that the magnitudes derived from the NIFS data and those from 2MASS are compatible within a 2-σ threshold, indicating that the flux calibration procedure was successfully applied to the datacubes.

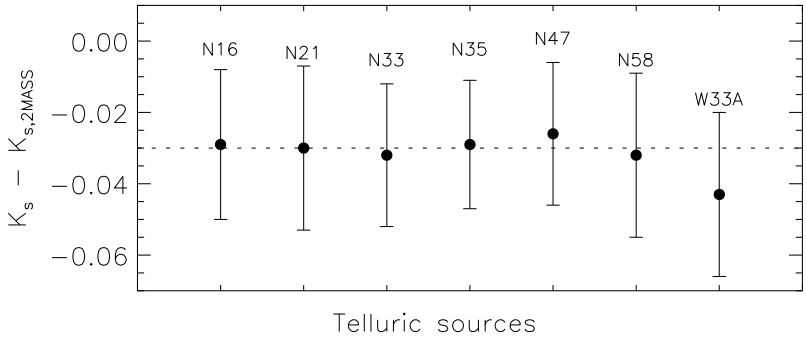


Figure 4.6: Residuals between the K_s-band magnitudes computed from the flux-calibrated NIFS data and the 2MASS magnitudes of the telluric standards. The error bars correspond to the error on the 2MASS K_s-band magnitudes, and the dashed horizontal line indicates the median value of all objects.

4.3 Results

This section presents the identification of the K-band spectral features associated with the HMYSOs and their surrounding nebulae, followed by the analysis of how they vary across the FOV of the NIFS datacubes. The detailed analysis of the NIFS data is presented in the next subsections. Sect. 4.3.1 discusses the spatial distribution of the K-band emission and the slope of the continuum towards the HMYSOs. In Sect. 4.3.2, the spectral features of the central sources are analysed. Section 4.3.3 presents the spatial distribution of discrete spectral features. The analysis and determination of the physics properties of the extended H_2 emission are presented in Sect. 4.3.4. Sect. 4.3.5 presents the examination of the integrated emission and kinematics of the Brγ emission. Finally, the study of the CO features, both in emission and in absorption are presented in Sect. 4.3.6).

4.3.1 Spatial distribution of the continuum emission

The spatial distribution of the continuum emission associated with the HMYSOs was investigated by comparing the average spatial distribution of the continuum emission of the HMYSOs with their telluric standard stars (see Table 4.4). The datacubes were integrated along the spectral direction and the average fluxes of the maps, in the x- and y-direction, were obtained. Figure 4.7 presents the distribution of the average fluxes as a function of the offset from the central position for all sources. The extension of the continuum emission was estimated in terms of the full width at tenth maximum (FWTM), defined as the width of the profile at its 10% level. The results are reported in Table 4.5. The FWTM values for the telluric standards range between 0.2-0.6″, while the science sources display FWTM values greater than 1.3″. The ratio between the science and the telluric FWTM values are greater than 3.0, indicating that all HMYSOs are clearly associated with extended continuum emission.

Table 4.5 - Full width at the tenth maximum of the spatial distribution of the continuum emission towards the telluric standards and the science objects.

ID	$FWTM_{tel}$ (″)	$FWTM_{sci}$ (″)	$FWTM_{sci}/FWTM_{tel}$
N16	0.42(07)	1.25(21)	3.0±0.7
N21	0.63(14)	>2.40(19)	>3.8±0.9
N33	0.22(26)	2.08(35)	5.4±1.0
N35	0.38(07)	1.28(40)	3.3±1.2
N47	0.52(12)	>2.22(21)	>4.3±1.1
N58	0.45(07)	>2.53(09)	>5.6±0.9
V49[a]	0.52(12)	1.64(01)	3.2±0.7
W33A[b]	0.31(—)	1.04(09)	3.3±0.3

Notes: The columns are: (1) ID of the source; (2) average and standard deviation of the full width at tenth maximum (FWTM) in the x- and y direction of the telluric standard; (3) same as column (2), but for the science object. A > symbol indicates a lower-limit; (4) ratio between column (3) and column (2). (a) due to problems on the telluric standard data of V49, the telluric standard of source N47 was used for comparison. (b) the telluric standard HR 6798 is a double star and, therefore, only the profile in the y-direction was used to estimate the FWTM of the telluric standard.

The distribution of the continuum emission was investigated as a function of the spectral range by obtaining integrated maps of the datacubes in three spectral regions centred at 2.05, 2.20 and 2.35 μm. The maps are presented as false-colour RGB images in Fig. 4.8. In order to improve the contrast and evidence any extended or faint emission in the observed fields, the maps are shown in square-root scaling of their intensity. The minimum and maximum intensity of each map were set to 0 and 30-σ, where σ corresponds to the noise level of the map. The adopted dynamical range provided the best contrast between

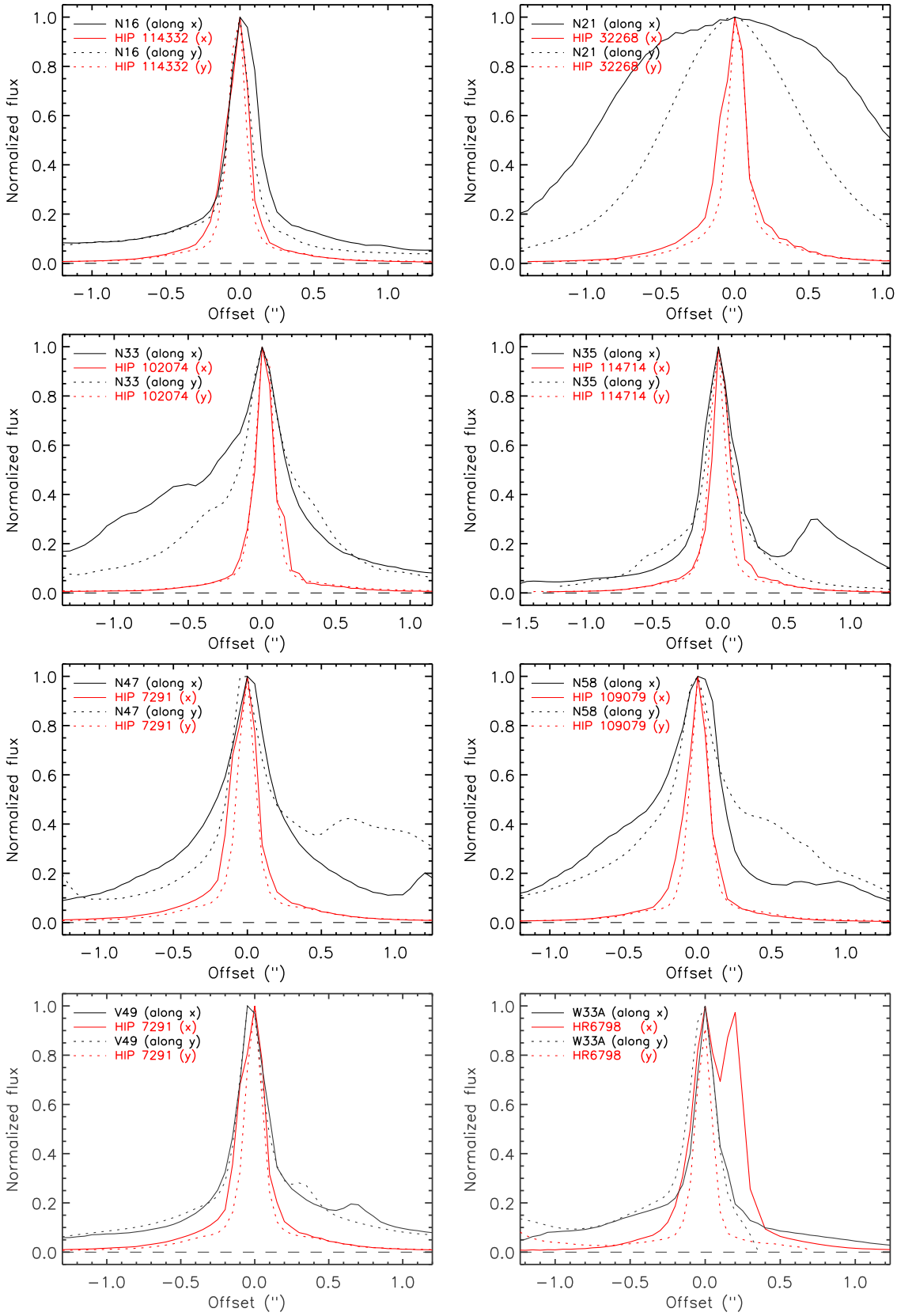


Figure 4.7: Average flux in the x- and y-directions of the continuum emission associated with the HMYSOs (in black) and their telluric standard stars (in red). The peak of each profile was set as the reference position. The horizontal dashed line indicates the baseline of the data.

structures brighter at longer wavelengths (*e.g.*, embedded point-like objects) or at shorter ones (*e.g.*, diffuse nebulae).

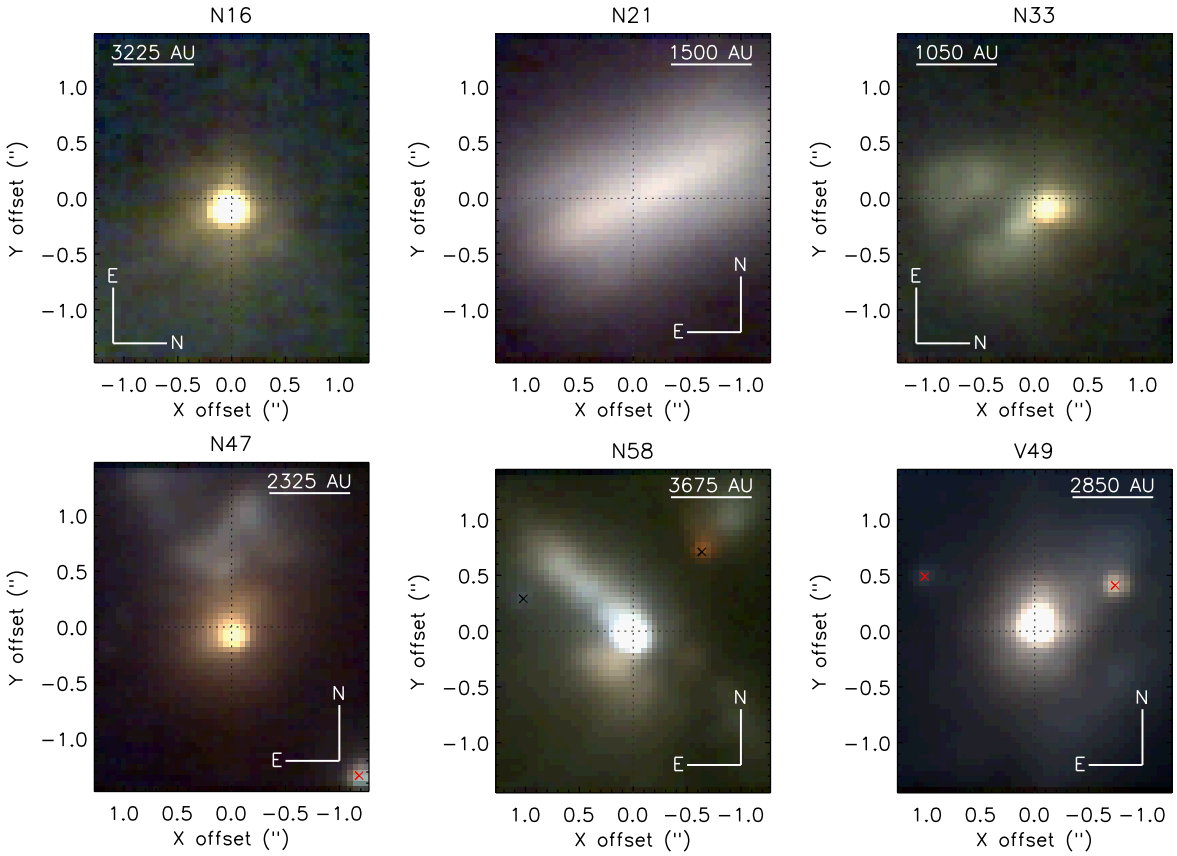


Figure 4.8: False-colour RGB images of the datacubes towards the HMYSOs. Each channel corresponds to the integrated emission in three spectral regions of the data (blue: 2.05 μm, green: 2.20 μm, red: 2.35 μm; using a bandwidth of 0.01 μm per map). The scale in the upper right indicates the projected length at the distance of the source. The orientation of the FOV is indicated in the bottom left or bottom right corner. The dashed black lines indicate the central position of the FOV. Additional point-like sources in the FOV are indicated by $\times$ symbols.

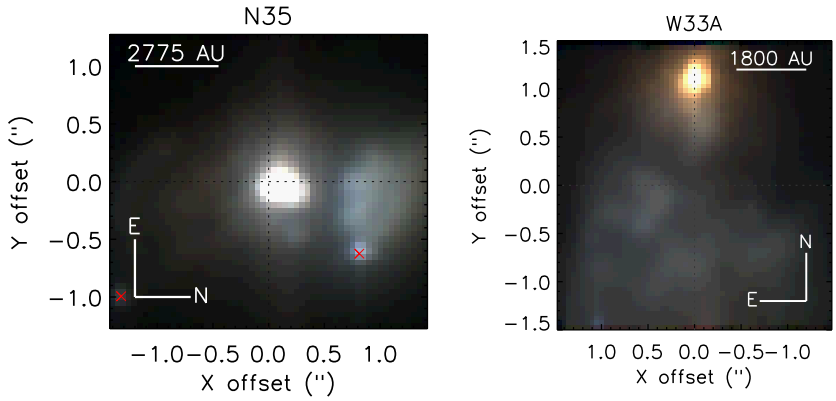


Figure 4.8: Continued.

The RGB maps reveal the presence of extended continuum emission towards all the

sources, as expected for embedded young objects and in agreement with the spatial profiles presented in Fig. 4.7. In half of the fields, the maps indicate that the flux at larger wavelengths is dominated by the embedded HMYSO rather than its associated extended nebula (see N33, N47, V49, W33A). In the case of N58, the nebulosity located at the SE direction is redder than the point-like object itself, which could be an indication of dust directly heated by the radiation from the central source.Finally, source N21 exhibits no point-like emission but an elongated structure in the SE-NW direction that runs across most of the FOV. Table 4.6 lists the properties of the structures identified in the FOV towards the sources, including the direction of the extended nebulae, the number of point-like sources identified in each field and a visual comparison of the contrast between the nebular emission and the central source.

Table 4.6 - Summary of the structures identified in the FOV of the datacubes.

Source	Direction of the Extend emission	Multiplicity	Comments
N16	towards W	1	point-like source (PS) is mainly red, extended emission is red (e.g. scattered by dust).
N21	SE-NW	1	No PS, ext. emission is blue
N33	towards S	1	Red PS
N35	towards N	3	Red PS, ext. emission is blue, blue companion
N47	towards N	2	Red PS, ext. emission is blue, comp. at SW
N58	NE-SW, SE-NW	3	Blue PS, red comp. (NW) and blue comp. (E), blue ext. emission in the NE-SW direction, red ext. emission in SW direction.
V49	towards NW	3	Red PS, ext. emission is blue, two comp. at NE and NW directions.
W33A	towards S	1	Red PS, ext. emission is blue

Notes: The columns are: (1) ID of the source; (2) direction of the extended emission based on the position of the main source; (3) multiplicity of point-like objects identified in the fields.

Fainter companion point-like sources were identified in the FOV of four objects (N35, N47, N58 and V49). The position of each point-like object was measured from the continuum maps and, assuming they are located at the same distance of the HMYSOs, the projected separation between the brightest source and each companion star was computed and listed in Table 4.7. These high-angular resolution observations confirm the existence of multiple, fainter and less massive companions in the vicinity of the HMYSOs, in according

with the theoretical prediction from three-dimensional simulations of the high-mass star formation process by Krumholz et al. (2009).

Table 4.7 - Relative distances of the companion stars to the main HMYSOs.

Source	N35	N47	N58	V49
X_c ($''$)	[+0.12,−0.02]	[+0.03,−0.04]	[−0.00,−0.00]	[+0.05,+0.08]
X_1 ($''$)	[+0.86,−0.57]	[+1.23,−1.30]	[−0.68,+0.77]	[+0.75,+0.44]
X_2 ($''$)	[−1.29,−1.01]	–	[+1.02,+0.29]	[−0.98,+0.51]
Scale (10^3 AU/$''$)	3.70(74)	3.10(62)	4.90(98)	3.80(76)
$\Delta\ell_1$ (10^3 AU)	3.38(68)	5.4(11)	5.0(10)	3.01(60)
$\Delta\ell_2$ (10^3 AU)	6.4(13)	–	5.2(10)	4.26(85)

Notes: The rows are: (1) ID of the source; (2) offset of the centre of the brightest source from the (0,0) position of the FOV (in arcseconds); (3) offset of the centre of the brightest companion object in the FOV; (4) when available, same as (3) for the third and faintest object in the FOV; (5) linear scale of the FOV, in AU arcsec^{-1}, computed using the distance of each source from Table 4.1; (6)-(7) projected separation between the point-like companions and the brightest source in the field. The numbers in parenthesis corresponds to the uncertainty on the last two digits, based on the error assumed for the distance (20%).

4.3.2 The central sources

This section presents the analysis of the spatial distribution of the continuum emission and the characterisation of the spectral features of the brightest source in each field presented in Fig. 4.8.

Figure 4.9 presents the integrated maps of the continuum emission of the HMYSOs, extracted in an emission-line-free region between 2.275 μm and 2.285 μm. As found in Fig. 4.8, the continuum maps are not only tracing point-like emission but often probe more extended structures around the HMYSOs. Such extended structures are likely reflection nebulae due to scattering of the continuum emission from the point-like sources and are detected along the entire spectral range of the NIFS data.

A two-dimensional Gaussian fit was performed to constrain the position and the extension of the point-like sources presented in each map presented in Fig 4.9. The parameters of the fits are summarised in Table 4.8. The comparison between the full width at half maximum (FWHM) values of the central sources and those from the telluric standards also confirms that the continuum emission of the HMYSOs is more extended than a point-like object. In addition, the Gaussian model of the telluric standards is a good approximation for estimating the Point Spread Function (PSF) of the NIFS observations. The FWHM of the HMYSOs was converted to linear scales (presented in the last columns of Table 4.8),

and the values indicate that the emission associated with the HMYSOs trace scales in the range of $\sim$(3-13) $\cdot 10^2$ AU (excluding the source N21, which has a completely distinct morphology than the rest of the sample).

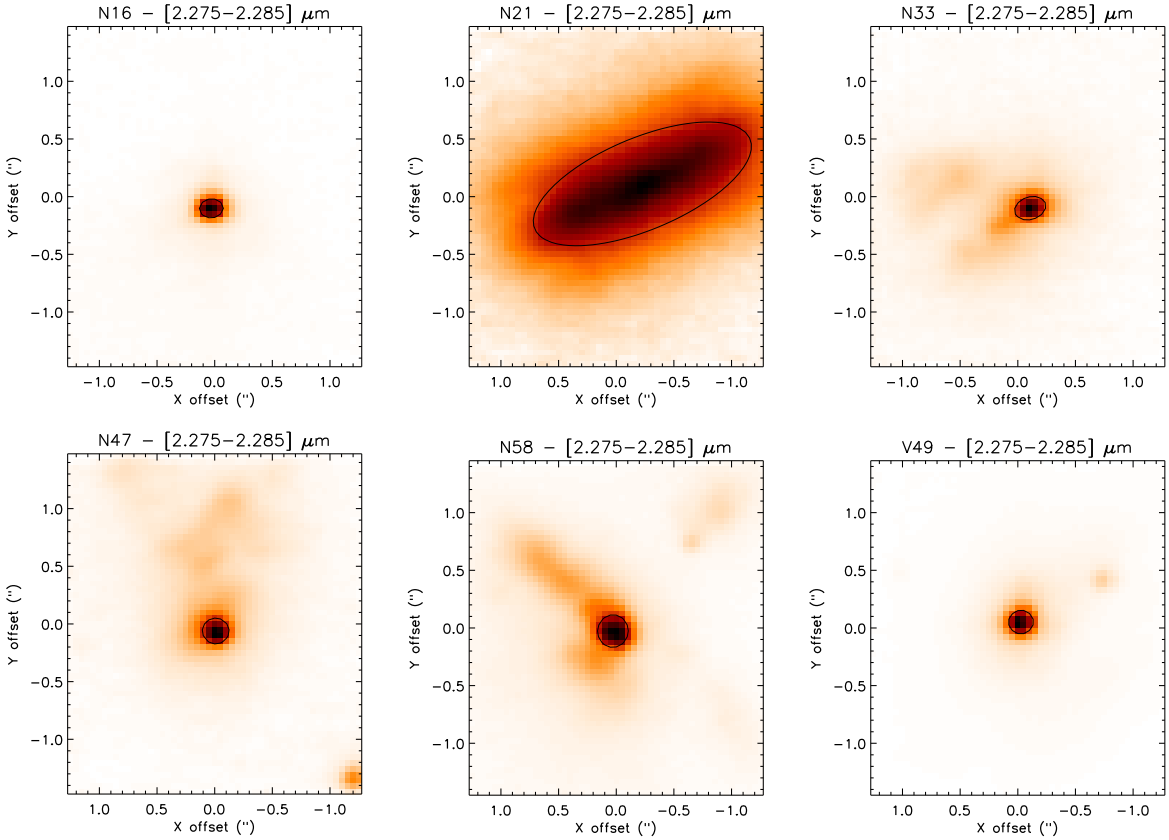


Figure 4.9: Continuum maps of the sources, integrated between the spectral range 2.275-2.285 μm. The black curve indicates the full width at half maximum of the Gaussian model (see Table 4.8) fitted for the point-like object. The orientation of the maps is indicated in Fig. 4.8.

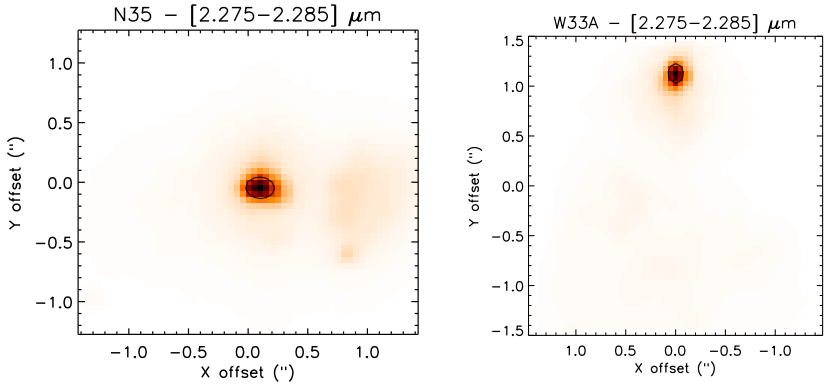


Figure 4.9: Continued.

The analysis of the extended emission was performed by analysing the residual maps presented in Fig. 4.10. These maps were obtained by subtracting the Gaussian model of

Table 4.8 - Parameters of the Gaussian fit model of the point-like continuum emission at 2.28 μm.

Source	x_c (pixel)	y_c (pixel)	FWHM$_1$ (pixel)	FWHM$_2$ (pixel)	PA (°)	Ω (10^{-13} sr)	<FWHM> (10^2 AU)	<FWHM>$_{tel}$ (pixel)
N16	25.43(1)	28.00(1)	3.94(3)	3.13(24)	94.6(3)	5.7	7.6(1)	3.02(5)
N21	30.44(7)	32.17(4)	39.69(25)	15.67(86)	112.8(1)	287.0	33.2(1)	3.13(5)
N33	28.11(4)	28.03(3)	5.45(11)	3.84(8)	17.8(7)	9.7	3.3(9)	2.73(5)
N35	30.98(2)	25.10(2)	4.72(5)	3.51(4)	176.4(4)	7.6	7.6(1)	2.91(5)
N47	26.16(3)	28.81(3)	4.51(8)	4.37(8)	90.3(7)	9.1	6.9(2)	3.10(5)
N58	25.44(5)	28.96(5)	5.13(12)	5.47(1)	-6.4(9)	13.0	13.0(4)	3.08(4)
V49	26.54(2)	30.51(2)	3.97(4)	3.76(4)	7.5(4)	6.9	7.4(1)	3.10(5)
W33A	30.04(1)	52.68(2)	2.91(3)	3.77(4)	1.8(5)	5.1	4.0(1)	2.97(4)

Notes: The columns are: (1) ID of the source; (2)-(3) coordinates of the centre on the x- and y-direction; (4)-(5) full width at half maximum of the principal axes; (6) position angle of the first principal axis (from N to E); (7) solid angle of the source projected on the sky (in sr); (8) mean projected FWHM at the distance of the source (in AU); (9) mean FWHM of the corresponding telluric standard star (in pixel); The numbers in parenthesis corresponds to the uncertainty on the last one or two digits.

the central sources from the maps presented in Fig. 4.9. They exhibit a detailed view of the extended continuum emission associated with the sources, as well as the presence of secondary and fainter point-like objects as already mentioned in Sect. 4.3.1. As expected, the maps indicate that all sources exhibit extended emission in the continuum. The most complex morphology is associated with the source N58, exhibiting extended structures in the NE-SW and NW-SE directions, respectively. Excluding the sources N21 and V49, the continuum emission associated with the other HMYSOs exhibit polar-like structures, suggesting the presence of heated dust along the propagation axis of molecular outflows.

Figure 4.11 presents the K-band spectra of the central sources, computed as the average spectra of the spaxels within the FWHM of the two-dimensional Gaussian models (see Table 4.8). The visual inspection of the spectra of all sources indicates that the continuum emission increases towards larger wavelengths, as expected for young and embedded protostellar objects.

The discrete spectral features associated with the HMYSOs were identified on the normalised spectra presented in Fig. 4.12. The K-band spectra were continuum-normalised by fitting a Spline function of order $n \leq 3$. A summary of the spectral lines identified in the spectra is given in Table 4.9.

The analysis of the K-band spectra of the sources indicates no evidence of photospheric features of OB stars (e.g., He I or He II lines, Brγ feature in absorption, ionised metal lines such as the C IV triplet and N III, Hanson et al., 2005), indicating that the photosphere

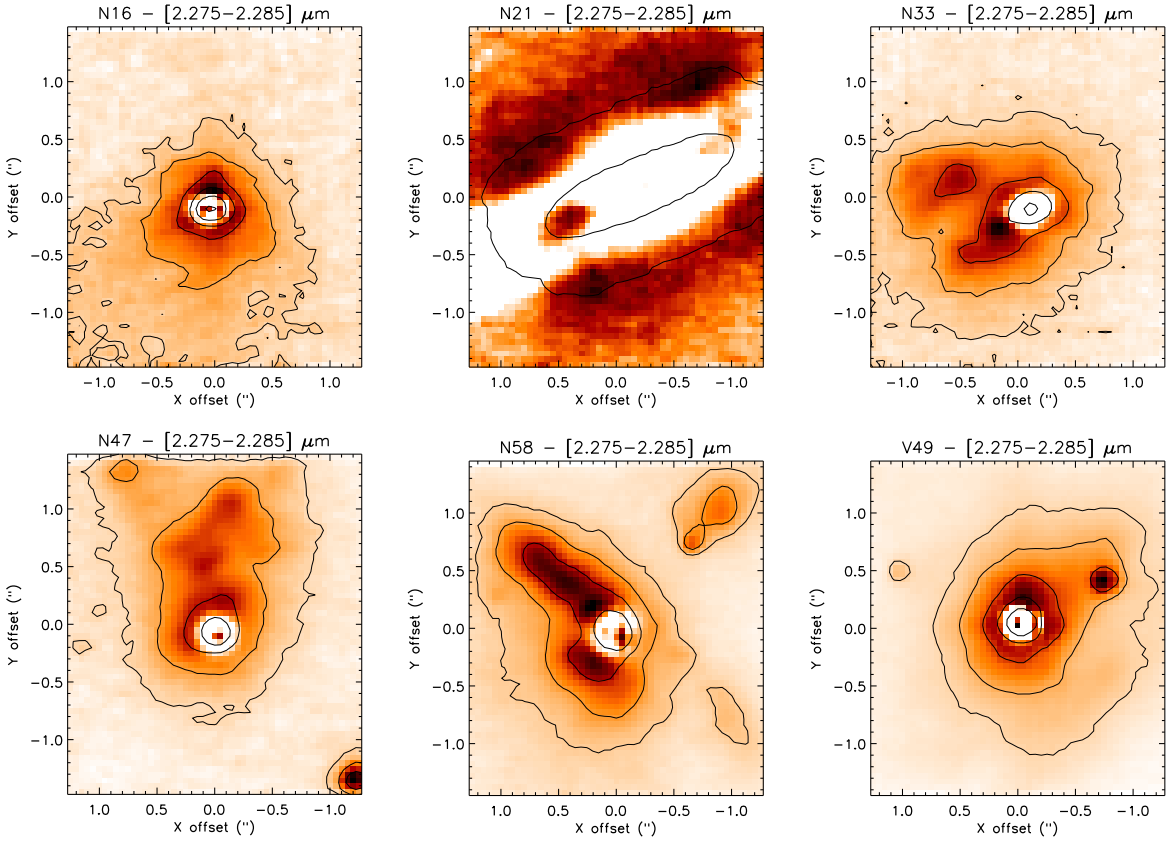


Figure 4.10: Residuals between the maps shown in Fig. 4.9 and the Gaussian model fitted for the point-like source. The black contours are placed at 3^n-σ levels (with $n = 0, 1, 2, ...$) of the continuum maps from Fig. 4.9, tracing both the extended regions with continuum fluxes above a 1-σ threshold, and the position of additional point-like objects in the FOV.

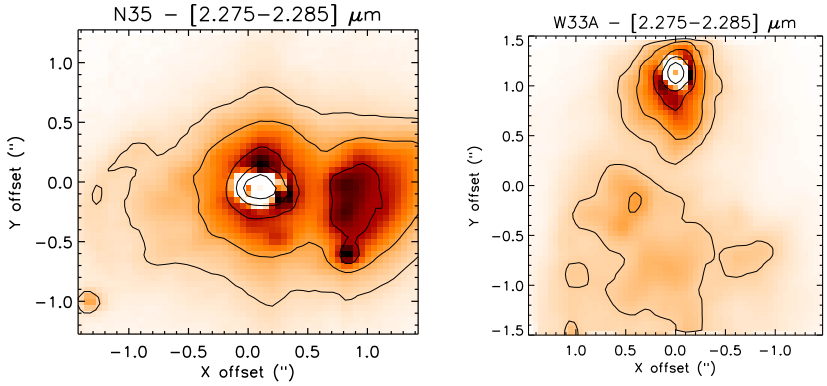


Figure 4.10: Continued.

of the underlying (proto)stars is completely obscured by their circumstellar material. The overall analysis of the spectra is presented in terms of the spectral features identified in the K-band range as follows.

The Brγ feature: The spectra of sources N16 and N21 are dominated by the dust thermal emission (see Fig. 4.11) and only a tentative detection of Brγ emission at 2.1655 μm was

identified in their normalised spectra, shown in Fig. 4.12. All the other sources exhibit a relatively broad Brγ emission feature.

The H_2 emission: Several transitions of the H_2 are identified towards the spectra of sources N35, N47, N58 and V49, indicating that the large-scale H_2 emission from Nav15 is likely originated by the main object in the FOV of these objects.

The CO emission features: Sources N47, N58, V49 and W33A exhibit the CO overtone emission features starting with the υ=(2–0) rotational-vibrational bandhead at 2.2935 μm and going up to the υ=(5–3) or even the υ=(6–4) bandhead, appearing as a blended line with the H_2 (1–0) Q(2) at 2.4134 μm. The CO bandhead features are likely to be originated at an unresolved inner disc region ($\sim$ 3 AU), where the physical conditions (temperatures of 2500-5000 K and densities of $n > 10^{11}$ cm^3, Kraus et al., 2000; Blum et al., 2004) are sufficient to maintain the CO molecules coupled.

The CO absorption features: Sources N33, N35, N47, V49 and W33A exhibit the low-J CO $\nu = 2$–0 lines as absorption features. These lines are likely originated in a colder ($\sim$30 K) and far ($\sim 10^3$ AU) region of the disc (Davies et al., 2010).

Metal features: Neutral sodium lines at 2.206 and 2.209μm are detected in emission towards sources N47, N58 and W33A (as previously reported by D10). The Na I lines are fluorescent and pumped by UV photons (Thompson and Boroson, 1977; Scoville et al., 1983). The lines are observed in emission, indicating that the sodium is under-ionised and hence shielded from the direct stellar radiation field. Accordingly to Porter et al. (1998), the Na I lines can also arise from a circumstellar disc, where the sodium atoms are shielded from the intense stellar UV-radiation field. Indeed, Table 4.9 indicates that the detection of Na I lines is correlated with the presence of the CO bandhead emission for sources N47, N58 and W33A, suggesting that both features are likely arising from circumstellar discs.

The spatial distribution of the individual spectral features identified in the spectra of the HMYSOs is presented in the next subsection.

Table 4.9 - Spectral features identified on the spectra of the sources.

Line	λ_0 (μm)	Source							
		N16	N21	N33	N35	N47	N58	V49	W33A
H_2 3–2 S(6)	2.0130	–	–	–	Y	–	Y	Y	–
H_2 1–0 S(2)	2.0338	–	–	–	Y	–	Y	Y[a]	–
H_2 1–0 S(1)	2.1218	–	–	–	Y	Y	Y	Y	–
H_2 2–1 S(2)	2.1542	–	–	–	Y	–	–	Y	–
Brγ	2.1655	(?)	(?)	Y	Y	Y	Y	Y	Y
Na I	2.2062	–	–	–	–	Y	Y	–	Y
Na I	2.2090	–	–	–	–	Y	Y	–	Y
H_2 1–0 S(0)	2.2235	–	–	–	Y	Y	Y	–	–
H_2 2–1 S(1)	2.2477	–	–	–	Y	Y	Y	–	–
CO 2–0	2.2935	–	–	–	–	Y	Y	Y	Y
CO (low-J)	2.33–2.35	–	Y	–	Y	Y	–	Y	Y
CO 3–1	2.3227	–	–	–	–	Y	Y	Y	Y
CO 4–2	2.3535	–	–	–	–	Y	Y	Y	Y
CO 5–3	2.3829	–	–	–	–	Y	Y	Y	Y
H_2 1–0 Q(1)	2.4066	–	–	–	Y	Y	Y	Y[a]	–
H_2 1–0 Q(2)	2.4134	–	–	–	Y	Y	Y[a]	–	–
H_2 1–0 Q(3)	2.4237	–	–	–	Y	Y	Y	–	–

Notes: The columns are: (1) Identification of the transition; (2) central wavelength of the line; (3)-(10) detected lines are indicated as 'Y', tentative detections are indicated as (?), no detections are shown as '–'. [a] blended line.

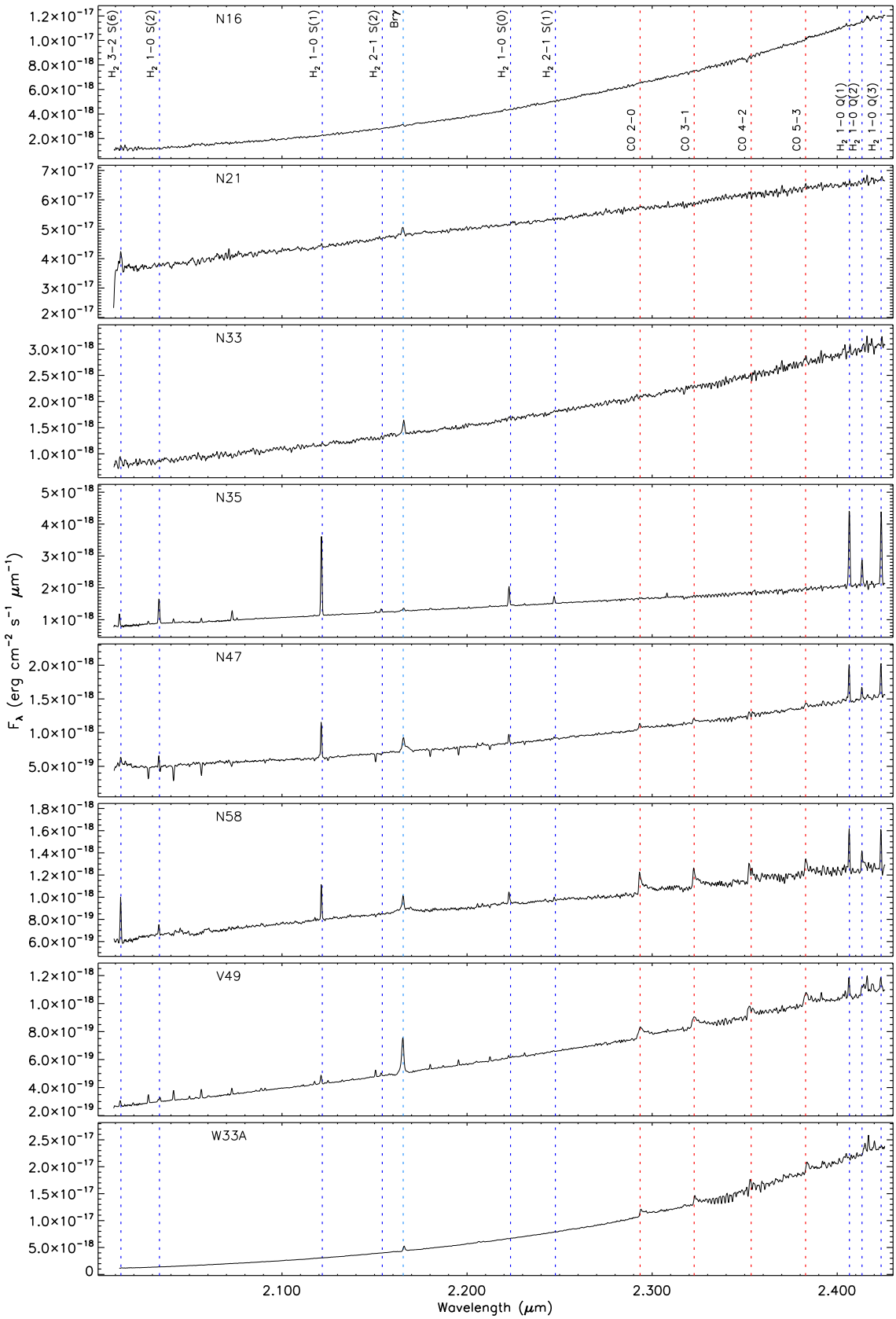


Figure 4.11: *K*-band spectra of the HMYSOs. The typical emission features are indicated by the vertical dashed lines (H_2 transitions are shown in blue, CO emission features in red, Bracket-γ in green, neutral Sodium doublet in black).

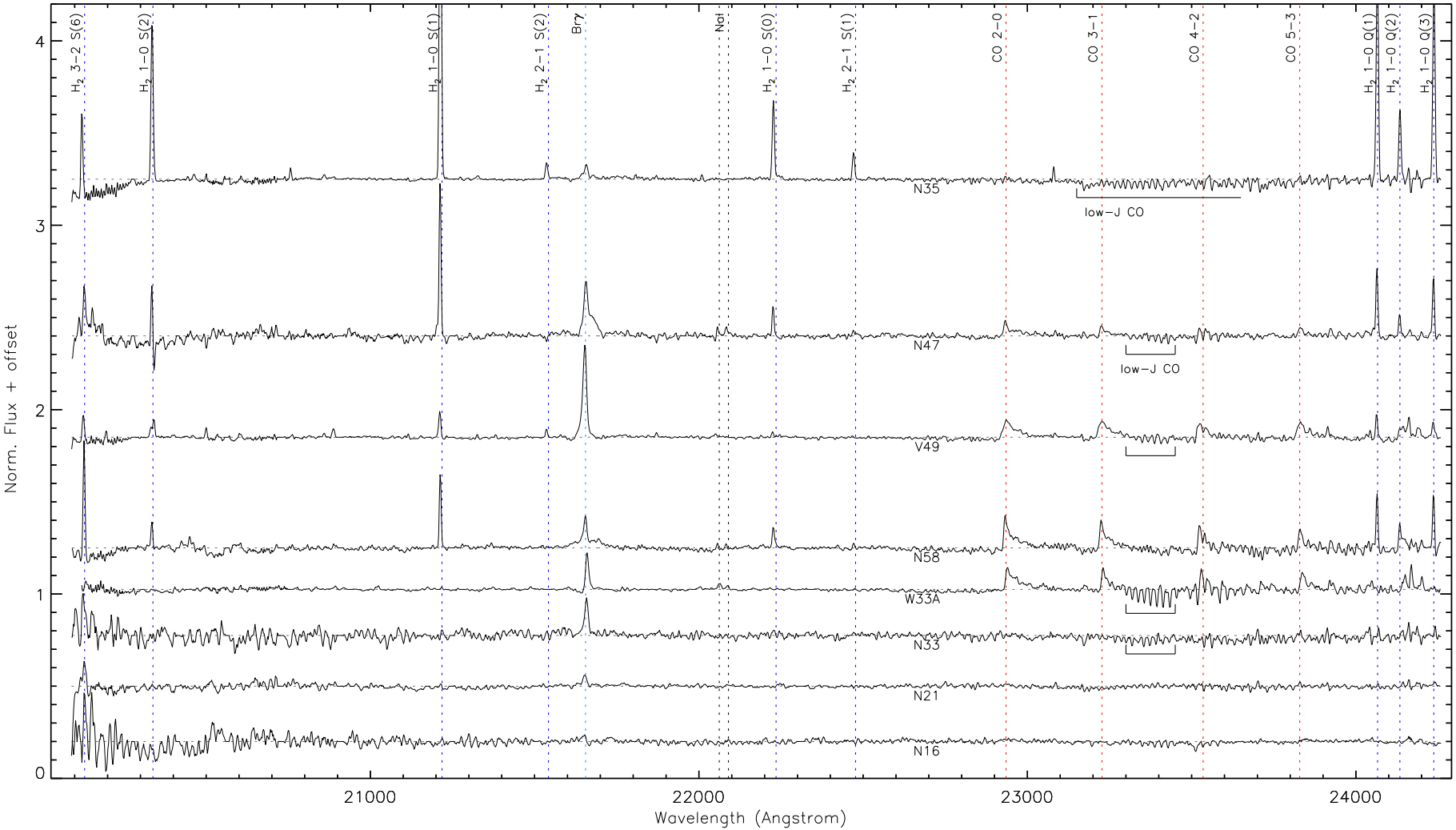


Figure 4.12: Continuum-normalised *K*-band spectra towards the HMYSOs. The typical spectral features of HMYSOs are indicated by the vertical dashed lines (H_2 transitions are shown in blue, CO emission features are shown in red, Brγ in green; other spectral lines are shown in black) and low-*J* CO in absorption are identified around 2.3 μm. The baseline of each spectrum is indicated by a dashed grey line. The peak of the emission features of source N35 was suppressed for better visualisation.

4.3.3 Spatial distribution of discrete spectral features

The analysis of the *K*-band spectra of the central sources has shown that Brγ and CO bandhead emission are present in most of the objects (N33, N35, N47, N58, V49 and W33A). In addition, four sources also exhibit molecular H_2 transitions towards their central spectra (N35, N47, N58 and V49). In order to study the spatial distribution of these emission lines, continuum-subtracted line-maps were constructed using an `IDL` routine that extracts a section of the datacube in a given wavelength interval, fits and subtracts the continuum emission from the input datacube using an outlier-resistant third-order polynomial fit[5]. Then, channel-map plots of the continuum-subtracted datacubes were created to highlight the distribution of the emission as a function of the wavelength (or velocity) along the spectral dimension.

Channel-map plots of three spectral lines (the H_2 (1–0) S(1) transition at 2.1218 μm, the Brγ line at 2.1655 μm and the CO (2–1) bandhead emission at 2.2935 μm) were generated for each source by using the procedure described above. Each channel-map is scaled in units of the standard deviation (σ) of each line-cube, estimated over the continuum in a spectral region outside of the line profile. To improve the contrast between emitting and absorbing regions, the maps are displayed in square-root of the intensity with respect to the baseline level[6].

Figure 4.13 presents the channel-map plots of the Brγ emission at 2.1655 μm and the H_2 (1–0) S(1) line at 2.1218 μm and the CO (2–0) bandhead feature at 2.35 for source N58. The Brγ emission is found between $v \sim [-195, +71]\,\mathrm{km\,s^{-1}}$. The analysis of the spectral distribution of the Brγ feature indicates that the line is not only associated with the main point-like source (indicated by the red ellipsoid on the maps), but traces an extended component towards the NE direction with a PA $\approx$50°.

The channel-map plot of the H_2 (1–0) S(1) transition indicates that the emission exhibits a more complex structure at velocities ranging from -174 to $+6.9\,\mathrm{km\,s^{-1}}$, suggesting a pair of bipolar structures. The brightest H_2 emission peaks at $v = -83.5\,\mathrm{km\,s^{-1}}$ and extends to the SE direction from the main HMYSO. The corresponding counter-jet was identified as a fainter structure, oriented in the opposite direction (NW) of the brightest nebular

[5] ROBUST_POLY_FIT

[6] that is, $\mathrm{displ}(X_{ij}) = \mathrm{sign}(X_{ij}) \cdot X_{ij}^{0.5}$, where X_{ij} corresponds to the observed intensity in a given spaxel ij, displ() is the scaled value of the intensity and sign() corresponds to ± 1 based on the sign of the intensity

emission, at the same velocity. The brightest lobe of the second bipolar structure identified in the FOV corresponds to the elongated structure oriented to the NE direction at $v = -83.5\,km\,s^{-1}$, similarly as the extended structure detected in Brγ. Its counter-jet-like emission was detected at $v = -53.4\,km\,s^{-1}$, and is oriented to the SW direction.

Finally, the CO (2–0) channel-map plot indicates that the CO emission peaks at $v = 42\,km\,s^{-1}$ and is detected at velocities ranging from -97.8 to $+236.8\,km\,s^{-1}$. The CO emission is also tracing extended emission to the NE direction (PA $\approx 50°$), similarly as the one probed by the Brγ and H_2 lines.

The comparison between the structures probed by different tracers around source N58 indicates that most of the H_2 emission is spatially different than the one from the Brγ or the CO lines. Besides, in every case, the emission is partially extended and not only associated with the central object.

The channel-map plots obtained for the other sources are presented in Appendix B.1. Table 4.10 summarises the velocity ranges of the emission probed by each spectral line in the FOV of the NIFS datacubes.

Table 4.10 - Velocity range of the Brγ, H_2 (1–0) S(1) and CO (2–0) emission towards the HMYSOs.

Source	V_{lsr} ($km\,s^{-1}$)	Brγ (2.1655 μm)	H_2 (1–0) S(1) (2.1218 μm)	CO (2–0) (2.2935 μm)
N21	−20.0	[−131.9,+45.4]	–	–
N33	+1.7	[−106.4,+159.5]	[−53.9,+96.8]	–
N35	−37.6	[−204.5,+150.5]	[−214.9,+86.6]	[−49.7,+33.9]
N47	−43.4	[−170.1,+273.0]	[−269.2,+2.1]	[−103.4,+119.7]
N58	−44.6	[−194.6,+130.3]	[−174.0,+6.9]	[−97.8,+236.8]
V49	−46.7	[−356.3,+205.0]	[−279.6,+22.0]	[−108.7,+198.1]
W33A	+35.3	[−80.7,+303.3]	[−88.5,+32.1]	[−44.4,+262.4]

Notes: The columns are: (1) ID of the source; (2) V_{lsr} from Table 4.1 (in $km\,s^{-1}$ units); (3)-(6) the velocity range of the Brγ, the H_2 (1–0) S(1), the H_2 (1–0) S(0) line and the CO (2–0) bandhead lines, respectively. The velocity ranges were not corrected by the V_{lsr} of the sources.

The line map of each spectral feature was obtained by integrating the continuum-subtracted datacubes within the spectral interval of the velocity ranges from Table 4.10. The line maps of the three spectral features are presented in Fig. 4.14. The analysis of the integrated line maps, combined with the velocity information from the channel maps (see Appendix B.1), led to the identification of distinct structures traced by different spectral lines. For example, Brγ and CO emission are tracing a compact structure associated with the point-like objects (*e.g.*, see V49 and W33A), but also exhibit a spatially extended

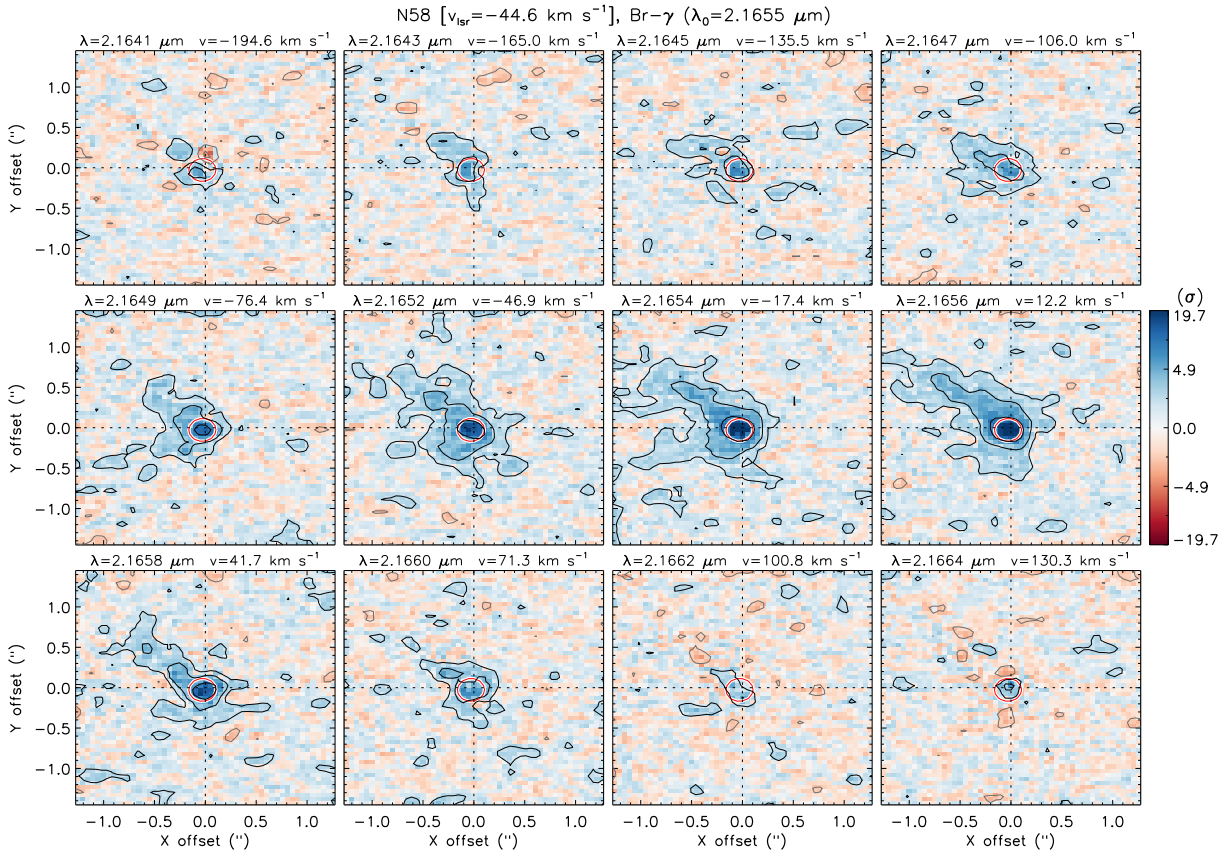


Figure 4.13: Channel-maps of the Brγ transition towards the source N58. The corresponding wavelength and its relative velocity with respect to λ_0 are indicated in the top of each channel. The intensity is scaled in σ units and is shown in a red-to-blue divergent scale, as indicated by the colour bar at the right of each panel. The maps are shown in linear or square-root scaling to improve the contrast between the structures. The Gaussian fit of the point-like continuum source and its position are indicated by the red ellipsoid. The black contours are placed at 3^n-σ, where $n = 0, 1, 2, ...$, while the grey contours are placed at -3^n-σ levels.

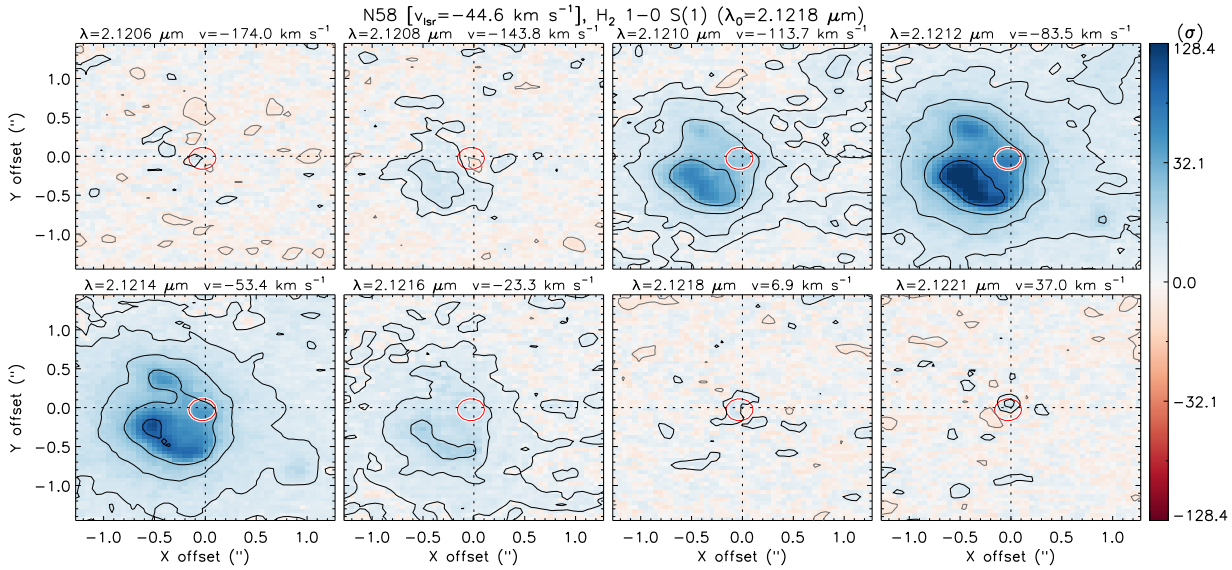


Figure 4.13: Continued. Channel-maps of the H_2 (1–0) S(1) emission towards the source N58.

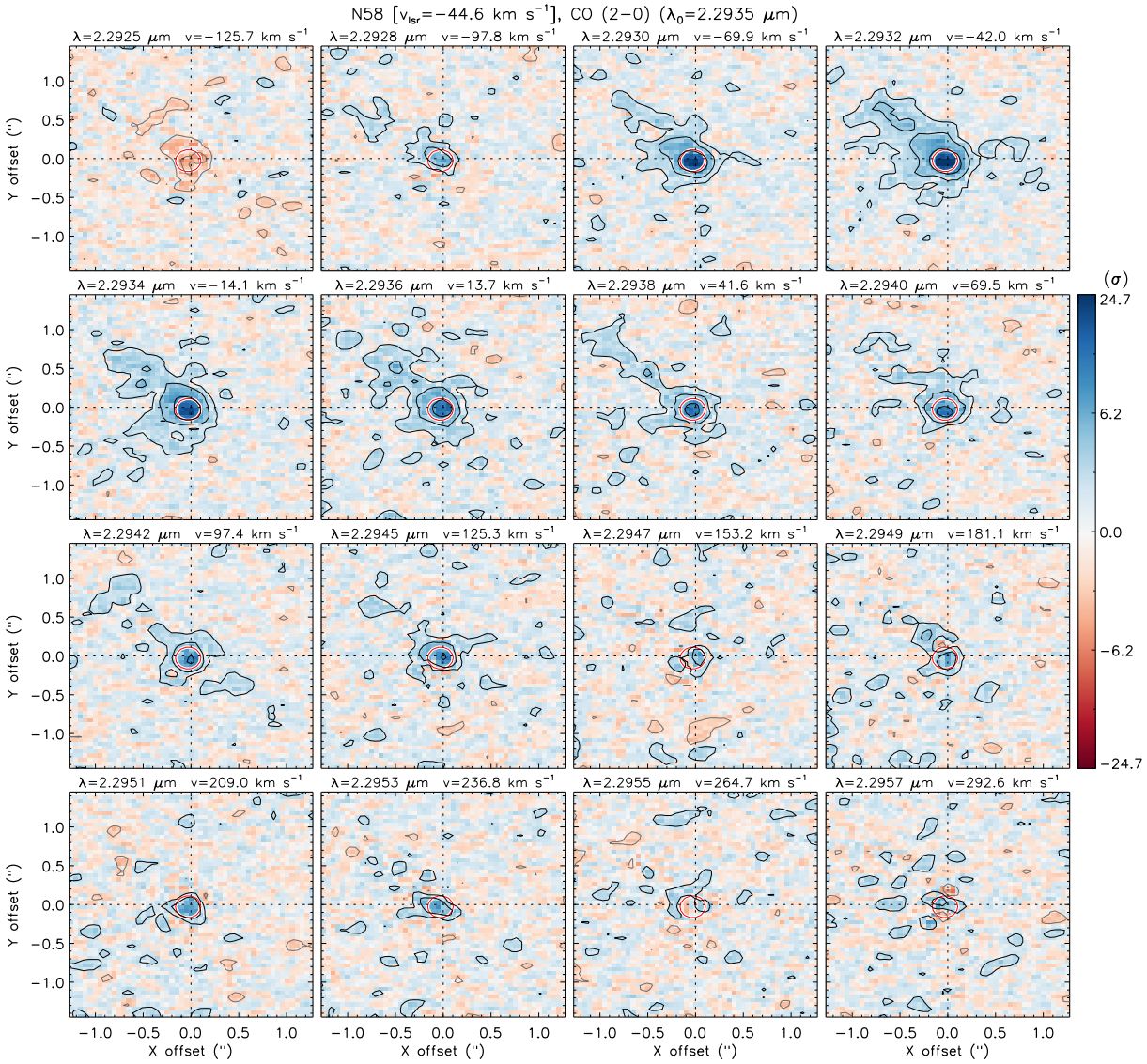


Figure 4.13: Continued. Channel-maps of the CO (2–0) bandhead feature towards the source N58.

component. The extended Brγ and CO emission are usually more compact than the extension of the H_2 emission (*e.g.*, N47 and N58).

The line-maps were used to derive the geometric parameters of the structures traced by each line. The peak velocity of the structures traced by each line, their extension and position angles are reported in Table 4.12. In the particular case of the H_2 (1–0) S(1) transition, jet-like structures are found among six of the eight sources (excluding sources N16 and N21) studied in this thesis. The geometrical parameters of these jet-like structures, their direction, length, aspect ratio (R=length/width) and the angular area of the brightest component ($\Delta\Omega$, in steradians), are summarised in Table 4.11.

The physical parameters of the individual spectral features are derived and analysed

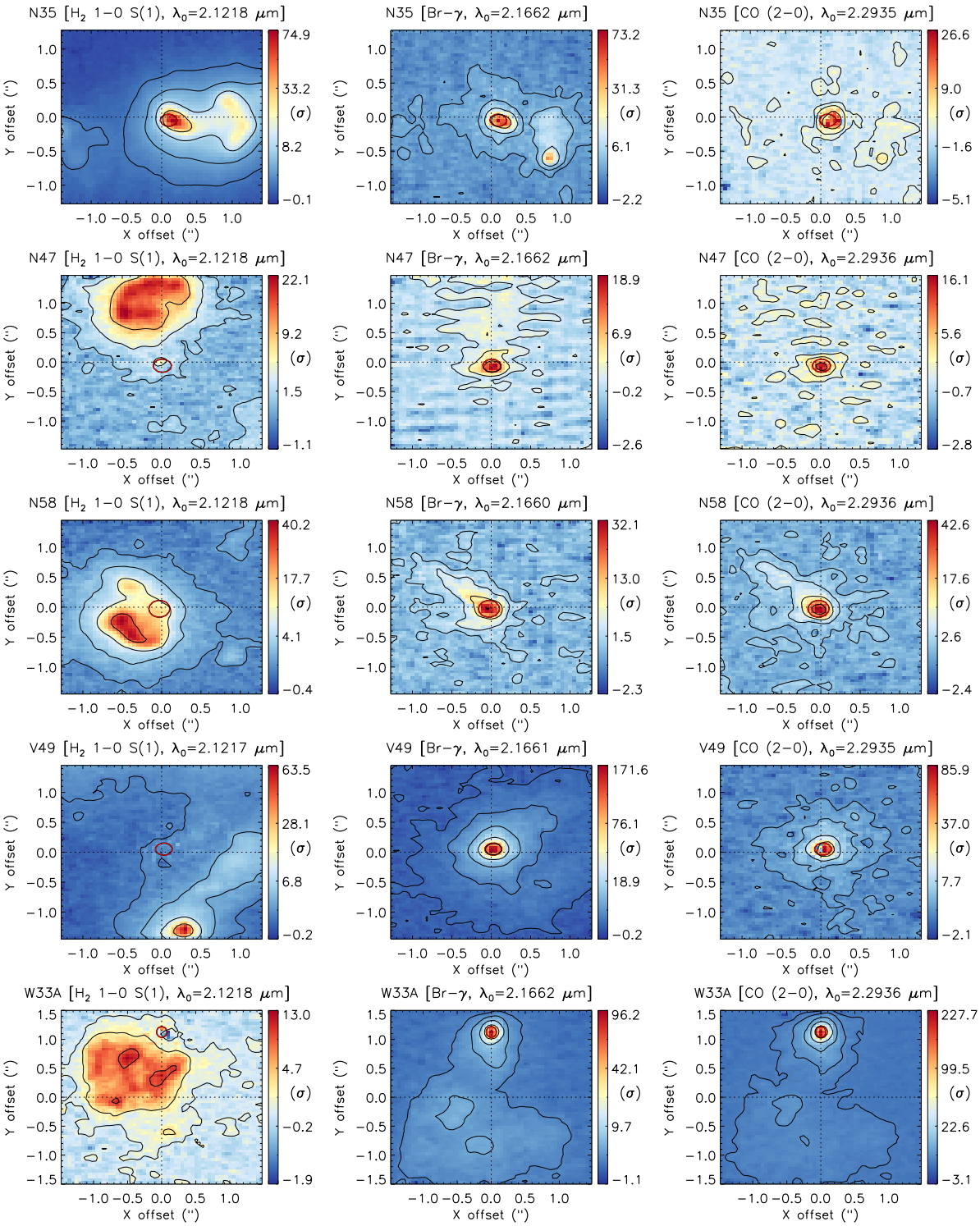


Figure 4.14: Integrated maps of the H_2 (1–0) S(1) (left), Brγ (middle) and CO (2–1) (right) features towards the HMYSOs. The corresponding rest velocity is indicated in the top of each image. The intensity is scaled in σ units and is shown in a red-to-blue divergent scale, as indicated by the colour bar at the right of each panel. The Gaussian fit of the point-like continuum source and its position are indicated by the red ellipsoid. The black contours are placed at 3^n-σ, where $n = 0, 1, 2, ...$

in the next subsections.

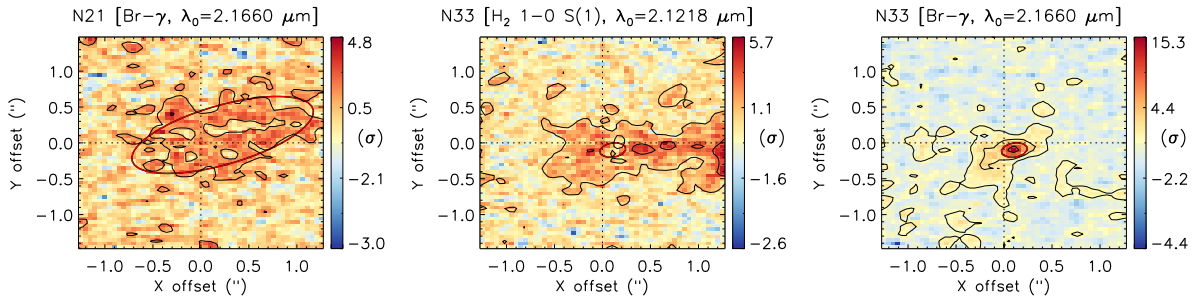


Figure 4.14: Continued. The integrated Brγ emission map towards N21 and the Brγ and H_2 (1–0) S(1) line-maps of source N33.

Table 4.11 - Extension of the H_2 nebula associated with the HMYSOs.

Source	Direction	$\Delta\ell$ (10^3 AU)	R	$\Delta\Omega$ (10^{-11} sr)
N33	N	1.55	1.6	1.70
N33	S	0.98	1.1	–
N35	N	4.14	1.4	3.16
N35	S	5.51	2.0	–
N47	N	4.68	1.1	2.61
N58	SE	3.09	1.1	3.11
N58	NE	3.14	2.7	–
N58	SW	5.59	3.5	–
N58	NW	7.40	4.3	–
V49	NW	6.42	3.0	3.30
W33A	SE	3.35	1.0	5.03

Notes: The columns are: (1) ID of the source; (2) direction of the emission; (3) projected length of the structures based on the angular length from Table 4.12; (3) aspect ratio of the H_2 emission (R=length/width); (4) Approximate area in the plane of the sky (in steradians) covered by the brightest lobe of the jets identified in each FOV.

Table 4.12 - Spatial structures traced by different lines towards the HMYSOs.

Source	Line	v_{peak}	$\Delta\ell$ ($''$)	PA ($^\circ$)	Comment
N21	Brγ	-126.3	2.80	123	Diffuse and weak emission (see Fig. B.1).
N33	H_2	$+66.7$	1.11	3	Brightest jet (red-shifted)
	H_2	-53.9	0.70	175	Counter-jet (blue-shifted), identified in Fig. B.2.
	Brγ	-41.8	0.43	217	Elongated structure to the SW direction.
N35	H_2	-136.2	1.12	3	Brightest structure
	H_2	-21.1	1.49	190	Red counterpart of the H_2 jet, see Fig. B.4.
	Brγ	-51.2	0.85	344	Extended emission between the source and the bow-shock of the H_2 emission
	CO	-21.8	0.99	355	Extended emission within the H_2 bow-shock
N47	H_2	-239.1,-88.3	1.51	12	Main H_2 jet, double-peaked emission, see Fig. B.6.
	Brγ	-75.9	0.80	6	Extended and point-like emission
	CO	-47.7	–	–	Point-like emission.
N58	H_2	-83.5	0.63	125	Brightest emission
	H_2	-83.5	0.64	45	Same orientation of the Brγ and CO features.
	H_2	-53.4	1.14	238	Red-shifted counterpart of the structure at 56°
	H_2	-83.5	1.51	313	Red-shifted counterpart of the structure at 125°
	Brγ	-70.9	0.88	45	Elongated structure to NE direction
	CO	-42.0	0.91	45	Elongated structure to NE direction
V49	H_2	-98.6	1.69	325	Main jet-like structure, origin at ($-0.25''$,$-1.30''$).
	Brγ	-114.3	0.86	315	Point-like emission and elongated nebula towards NW direction (see Fig. B.9)
	CO	$+2.9$	–	–	Point-like emission.
W33A	H_2	-28.2	1.34	146	H_2 jet
	Brγ	-16.1	2.13	172	point-like and extended emission to the S direction
	CO	$+39.3$	2.13	172	point-like and extended emission to the S direction, similar to the Brγ emission.

Notes: The columns are: (1) ID of the source; (2) the spectral line: H_2 (1–0) S(1) at 2.1218 μm, Brγ at 2.1661 μm; and CO (2–0) at 2.3529 μm; (3) the velocity of the brightest emission, $v = 0\,km\,s^{-1}$ for the vacuum wavelength of each transition/ (4) the maximum projected length of the structure (in arcseconds); (5) the position angle of the emission (in degrees, from N to E); (6) comments on each structure.

4.3.4 *Extended H_2 emission as a probe of molecular jets*

This section presents the derivation of the physical parameters of the gas traced by the H_2 emission, identified through the line-map of the H_2 (1–0) S(1) feature at 2.12 μm (see Fig. 4.14). Extended H_2 emission is present in six of the eight HMYSOs investigated in this work (N33, N35, N47, N58, V49 and W33A). The morphology of the H_2 emission associated to these sources suggests that the H_2 is probing the molecular jets at closer distances to their launching points, in scales of $\sim 10^3$ AU, (see Table 4.11).

The spectral characterisation of the H_2 emission was performed by obtaining the average K-band spectrum of the spaxels with fluxes above a 1-σ threshold in the H_2 (1–0) S(1) line-map (see Fig. 4.14). The resulting spectra are presented in Fig 4.15, together with the identification of the spectral lines associated with the jets. Several residual telluric lines are present in most of the spectra, shown as grey features in the spectra to avoid confusion with intrinsic nebular features. At least ten H_2 transitions coming from the rotational levels J=2 to 1 and 1 to 0, from the vibrational levels ν=0 to 4 are identified in most of the nebular spectra. In addition, the Brγ emission feature is also present in every spectrum displayed in Fig. 4.15, while CO bandhead emission is also found towards the nebular spectra of sources N58 and W33A.

In order to check if the emission features are intrinsically originated by the extended nebula or if they are reflected radiation of the central source, the ratio between the average nebular (Fig. 4.15) and the central source spectra (Fig. 4.11) was computed ($F_{\rm nebula}/F_{\rm central}$). The resulting spectra are presented in Fig. 4.16. The comparison between the features present in the original spectra and those identified in Fig. 4.16 indicates that the Brγ and CO features, previously detected on the individual spectra, appear as weak or even absent features in the ratio of the central to the nebular spectra. Their absence in Fig. 4.16 suggests that such lines are nor produced in the nebula, but are mainly reflected radiation from the central source. On the other hand, the H_2 transitions are detected as strong features, indicating that the H_2 emission is intrinsically originated by the nebulae.

Except for source N58, the continuum level of the $F_{\rm nebula}/F_{\rm central}$ ratio decreases towards larger wavelengths for all the other five HMYSOs, suggesting that the extended emission is brighter in the blue than the central source.

The next subsection presents the derivation of the physical parameters of the H_2 emis-

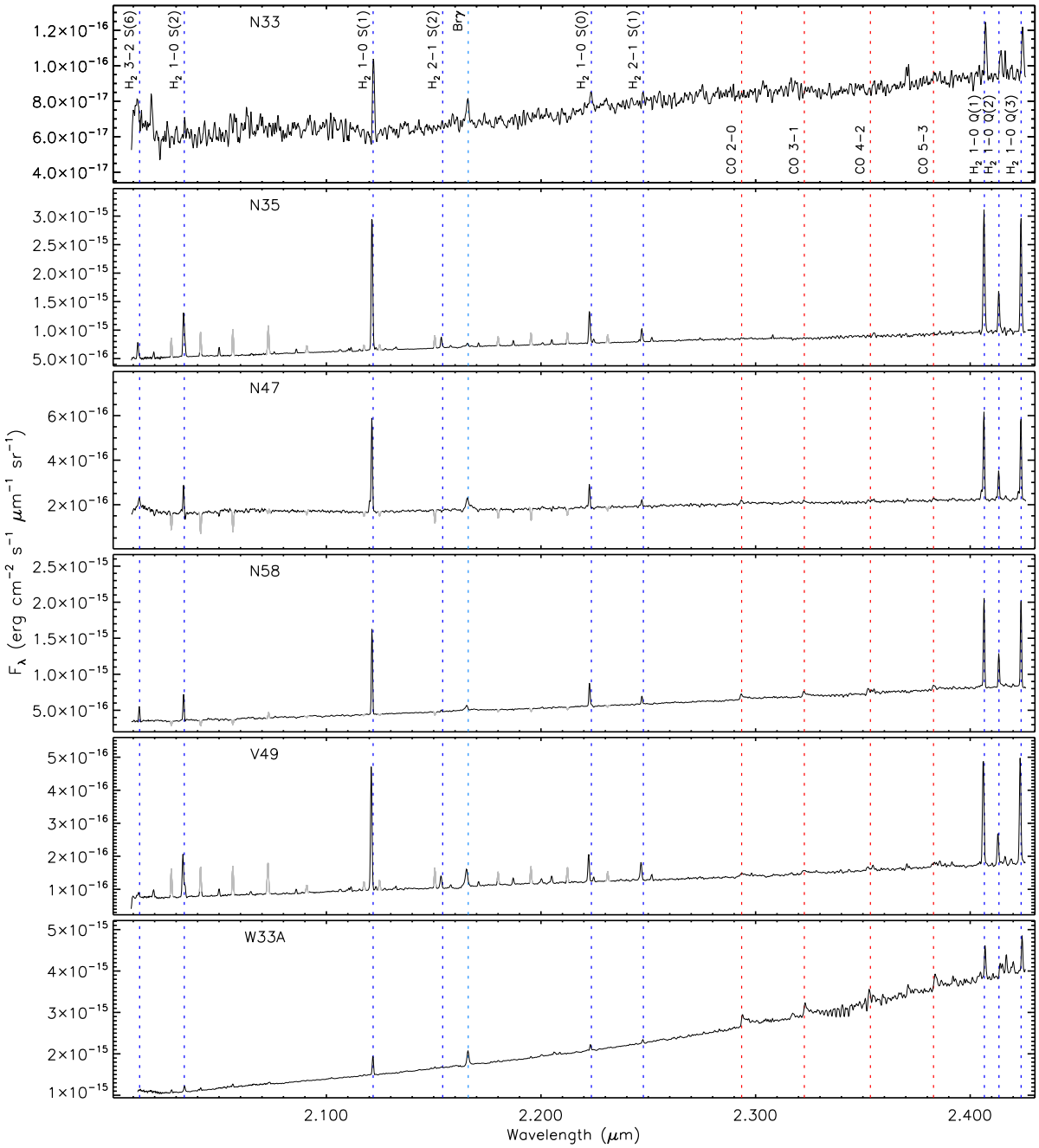


Figure 4.15: Average *K*-band spectra towards the H_2 jets associated with the HMYSOs. The typical spectral features identified in the spectra are indicated by the vertical dashed lines: H_2 transitions are shown in blue, CO emission features are shown in red, Brγ in green. Residual telluric lines are indicated in grey.

sion through the analysis of the flux of different H_2 transitions present in their spectra.

4.3.4.1 *The average physical parameters of the jets*

The physical parameters of the H_2 jets (*e.g.*, their excitation temperature, column density and reddening) were derived based on the analysis of the integrated flux of about 16 H_2 transitions present in the *K*-band spectra from Fig. 4.15, with upper-level energies

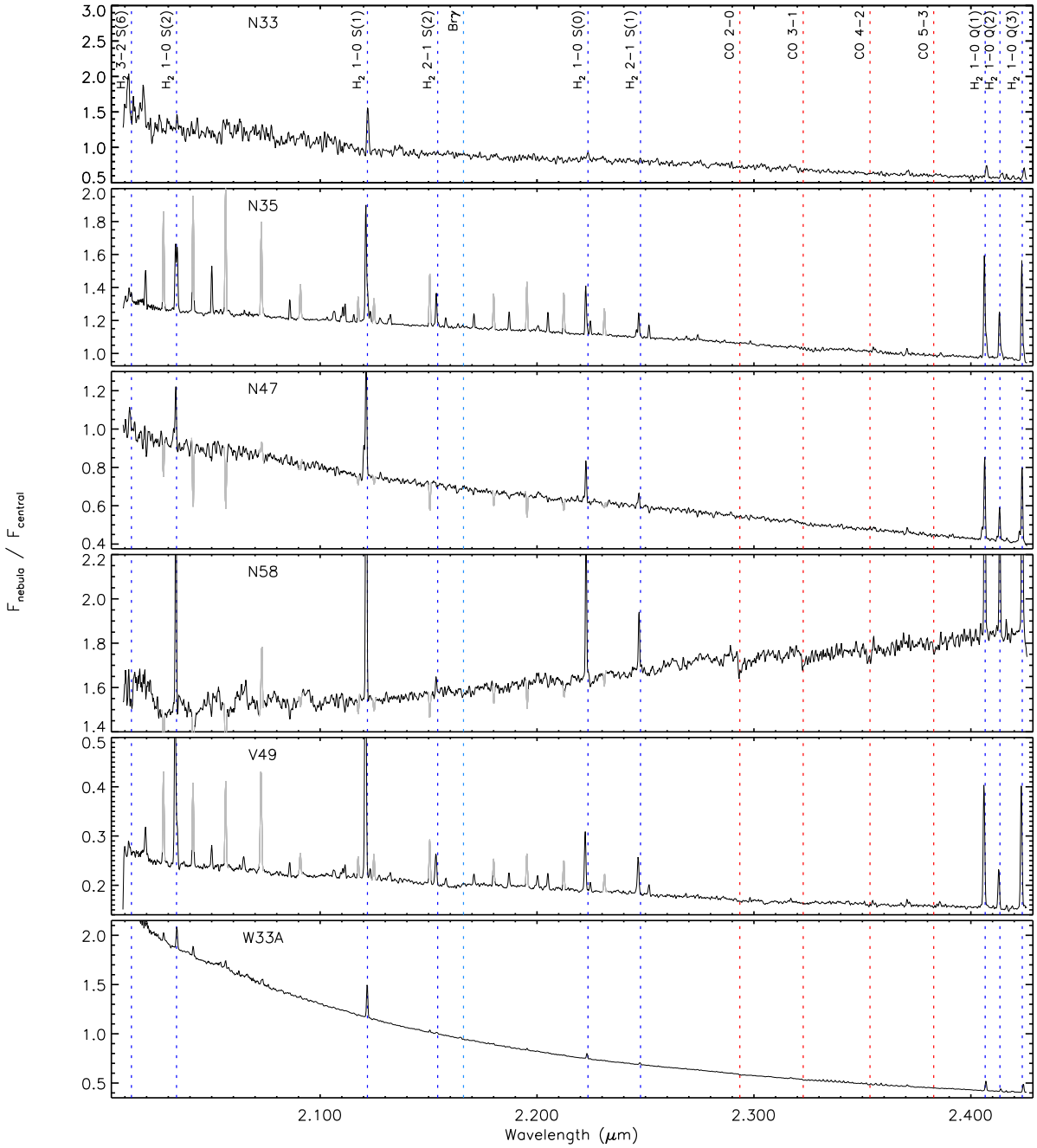


Figure 4.16: Ratio between the nebular and the central source spectra (presented in Figs. 4.15 and 4.11, respectively). The typical spectral features identified in the spectra are indicated by the vertical dashed lines: H_2 transitions are shown in blue, CO emission features are shown in red, Brγ in green. Residual telluric lines are indicated in grey.

ranging from 6149 to 21911 K. A list of all the H_2 transitions available in the *K*-band and their molecular properties is presented in Table 4.13.

The IDL routine LINE_FLUX was written to automatically extract the flux and compute the width of the lines detected above a given threshold. The integrated flux is obtained by summing the flux of the spectral pixels with fluxes above a 1-σ threshold. The uncertainty

Table 4.13 - Properties of the H_2 transitions present in the *K*-band data.

Transition	$\lambda_{J,v}$ (Å)	$g_{J,v}$	E_{up} (K)	$A_{J,v}$ (10^{-7} s^{-1})
H_2 (3–2) S(6)	20130	17	21911	3.57
H_2 (1–0) S(2)	20338	9	7584	3.98
H_2 (3–2) S(5)	20656	45	20856	4.50
H_2 (2–1) S(3)	20735	33	13890	5.77
H_2 (1–0) S(1)	21218	21	6956	3.47
H_2 (3–2) S(4)	21280	13	19912	5.22
H_2 (2–1) S(2)	21542	9	13150	5.60
H_2 (3–2) S(3)	22014	33	19086	5.63
H_2 (1–0) S(0)	22235	5	6471	2.53
H_2 (2–1) S(1)	22477	21	12550	4.98
H_2 (3–2) S(2)	22870	9	18386	5.63
H_2 (2–1) S(0)	23556	5	12095	3.68
H_2 (3–2) S(1)	23864	21	17818	5.14
H_2 (1–0) Q(1)	24066	9	6149	4.29
H_2 (1–0) Q(2)	24134	5	6471	3.03
H_2 (1–0) Q(3)	24237	21	6956	2.78

Notes: The columns are: (1) Identification of the transition based on the rotational and vibrational levels, J and v, respectively; (2) the rest wavelength; (3) the statistical weight; (4) the upper-level energy; and (5) the Einstein coefficient of the transition. Data from Turner et al. (1977) and Dabrowski (1984).

of the flux is proportional to the rms of the nearby continuum multiplied by the square root of the number of spectral channels within the spectral feature ($\sigma_f = \sigma_{rms}\sqrt{n_{pix}}$).

As an example, Table 4.14 summarises the fluxes of the H_2 transitions towards the average spectrum of the H_2 jet associated with the source N35 (see Fig. 4.15), which exhibits the most extensive number of H_2 lines in the sample. The tables containing the line fluxes measured towards the other sources are presented in Tables B.1 to B.5, in Appendix B.2.

The physical parameters of the H_2 emission were derived based on to the methodology presented by Caratti o Garatti et al. (2015, hereafter, C15), who investigated the properties of large-scale H_2 jets based on long-slit spectroscopic data. As defined in C15, the column density of each H_2 transition ($N_{v,J}$, characterised by its upper-rotational and vibrational levels, J and v, respectively) is given by:

$$N_{v,J} = 4\pi \frac{\lambda_{v,J}}{hc} \frac{I_{v,J}}{A_{v,J}} \tag{4.4}$$

where $\lambda_{v,J}$ is the rest wavelength of the transition, $I_{v,J} = F_{v,J}\,\Omega$ is the line intensity (in

Table 4.14 - Parameters of the emission lines of the flux-calibrated average H_2 nebular spectrum of the source N35.

Transition	λ_{peak}	FWHM	F_λ	σ_F	$F_\lambda/F_{2.12\mu m}$	$\sigma_{F\lambda/2.12}$
H_2 3-2 S(6)	20120.9	2.13	1.19	0.16	0.098	0.015
H_2 1-0 S(2)	20334.3	4.27	3.88	0.14	0.319	0.023
H_2 2-1 S(3)	20650.1	4.27	0.21	0.10	0.017	0.008
H_2 1-0 S(1)	21213.5	4.27	12.19	0.75	1.000	0.087
H_2 3-2 S(4)	21273.2	2.13	0.06	0.02	0.005	0.001
H_2 2-1 S(2)	21535.7	4.27	0.54	0.04	0.044	0.004
H_2 3-2 S(3)	22007.2	4.27	0.19	0.02	0.015	0.002
H_2 1-0 S(0)	22227.0	4.27	2.63	0.33	0.216	0.030
H_2 2-1 S(1)	22470.3	4.27	1.12	0.09	0.092	0.010
H_2 3-2 S(1)	23859.4	4.27	0.30	0.06	0.025	0.005
H_2 1-0 Q(1)	24064.3	4.27	12.19	0.26	1.000	0.065
H_2 1-0 Q(2)	24132.5	4.27	3.88	0.15	0.318	0.023
H_2 1-0 Q(3)	24237.1	4.27	11.52	0.19	0.945	0.060

Notes: The columns are: (1) Identification of the spectral feature; (2) the peak wavelength of the emission (in Å); (3) the full width at half maximum (FWHM) of the emission (in Å); (4)-(5) the observed flux and its error (in $10^{-18}\,erg\,cm^{-2}\,s^{-1}$ units); (6)-(7) the ratio between the flux at a given transition and the flux at 2.12 μm, and its error.

erg cm^{-2} s^{-1}), Ω is the solid angle of the spaxel[7] (in sr), and $A_{v,J}$ is the Einstein coefficient (in s^{-1}). The $A_{v,J}$ values are taken from Turner et al. (1977) and Dabrowski (1984).

An interstellar reddening law must be adopted for obtaining the real column density of each H_2 transition from the observed $N_{v,J}$ values (from Eq. (4.4)). The extinction correction should be applied to the observed intensity of the lines, as given by:

$$I_{\lambda,cor} = I_{\lambda,obs} 10^{A_\lambda/2.5} \tag{4.5}$$

where A_λ is the extinction in a given wavelength (in mag), $I_{\lambda,obs}$ is the observed intensity and $I_{\lambda,cor}$ is the de-reddened intensity. In the near-infrared, the ISM extinction law follows a power-law distribution given by:

$$\frac{A_\lambda}{A_{\lambda 0}} = \left(\frac{\lambda}{\lambda_0}\right)^{-\alpha} \tag{4.6}$$

where A_λ is the extinction in a given wavelength (in mag), α is slope of the extinction law and λ_0 is the reference wavelength. Using Eqs. (4.5) and (4.6) in Eq. (4.4), the extinction-

[7] the corresponding solid angle of a 0.″05 spaxel is $\Omega \approx (\Delta\theta)^2 \approx 5.88{\cdot}10^{-14}$ sr.

corrected column density of each line can be evaluated as:

$$N_{v,\mathrm{J}} = 4\pi \frac{\lambda_{v,\mathrm{J}}}{hc} \frac{I_{v,\mathrm{J}}}{A_{v,\mathrm{J}}} 10^{\wedge} \left[\frac{A_{\mathrm{K}}}{2.5} \left(\frac{\lambda_{v,\mathrm{J}}}{\lambda_{\mathrm{K}}} \right)^{-\alpha} \right] \tag{4.7}$$

Recently, Damineli et al. (2016, hereafter D16) derived an ISM reddening law towards the inner Galaxy with a power-law index of $\alpha = 2.13 \pm 0.08$ in the 0.4-4.8μm range. The ratio between A_{V} and A_{K} derived by those authors, $A_{\mathrm{V}}/A_{\mathrm{K}} = 14.95$, is compatible with other recent determinations ($A_{\mathrm{V}}/A_{\mathrm{K}} = 16.13$ Nishiyama et al., 2008) and is systematically larger than previous IR-reddening laws ($A_{\mathrm{V}}/A_{\mathrm{K}} = 8.77$-$8.92$ Rieke and Lebofsky, 1985; Cardelli et al., 1989). The following results were obtained by adopting the ISM reddening-law from D16 and defining the reference wavelength for the K-band as $\lambda_0 = 2.159\,\mu$m.

Assuming that the H_2 emission is in local thermodynamic equilibrium (LTE), the population of each transition of the molecule follows the Boltzmann distribution, given by:

$$\left(\frac{N}{g} \right)_{v,\mathrm{J}} \propto \exp\left(-\frac{E_{v,\mathrm{J}}}{k_{\mathrm{B}} T_{\mathrm{ex}}} \right) \tag{4.8}$$

where $(N/g)_{v,J}$ is the column density of the transition divided by its statistical weight, $g_{v,\mathrm{J}}$; $E_{v,J}$ is the upper-level energy of the transition, k_{B} is the Boltzmann constant and T_{ex} is the excitation temperature of the gas. By following the formulae from Zhang et al. (1998), the column density of a given transition is related to T_{ex} and the total column density of the molecule, N_{tot}, as:

$$\log\left(\frac{N}{g} \right)_{v,\mathrm{J}} = \log \frac{N_{\mathrm{tot}}}{Q(T)} - \frac{E_{v,\mathrm{J}}}{k_{\mathrm{B}} T_{\mathrm{ex}}} \tag{4.9}$$

where $Q(T)$ is the partition function given by:

$$Q(T) = \sum_{v,\mathrm{J}} g_{v,\mathrm{J}} \exp\left(-\frac{E_{v,\mathrm{J}}}{k_{\mathrm{B}} T_{\mathrm{ex}}} \right) \tag{4.10}$$

The *Boltzmann diagram* of the H_2 emission was constructed by plotting the logarithm of the column density divided by its statistical weight against the corresponding upper-level energy of each H_2 transition. A linear fit to the data in the Boltzmann diagram, $\log(y) = \alpha\, x + \beta$ provides an estimate of the extinction (characterised by the A_K value), the excitation temperature ($T_{\mathrm{ex}} = -1/\alpha$) and the total column density [$N_{\mathrm{tot}} = Q(T_{\mathrm{ex}}) \exp(\beta)$]

of the H_2 gas.

By following the procedure described above, an `IDL` routine was written to construct the Boltzmann diagrams of the H_2 lines and to compute the corresponding $N_{\rm tot}$, $T_{\rm ex}$ and A_K values. Pairs of H_2 transitions originating from the same energy level should lie in the sample position in the diagram. Thus, the correlation between the column density and the energy is expected to be the highest so the extinction and the temperature can be evaluated simultaneously (see Caratti o Garatti et al., 2015, and references therein).

Fig. 4.17 presents the analysis of the Boltzmann diagram towards the H_2 jet associated with the source N35 (see Fig. 4.15 and Table 4.14). Each plot of the top panel exhibits the Boltzmann diagram of the H_2 emission constructed by varying A_K between 0 and $\sim$5.0 mag in steps of 0.05 mag (to improve the visualisation, the top panel of Fig. 4.17 displays A_K varying between 0 and 2.4 mag, in steps of 0.10 mag). The best fit of the data is indicated in each plot and the goodness of the fit was tested by investigating the variation of two independent statistical coefficients: the linear Pearson correlation coefficient of the data (r[8]) and the "standard deviation" of the residuals between the observed data and the fit ($\sigma_{\rm fit}$). The correlation between the quantities should be maximum as $\sigma_{\rm fit}$ tends to its minimum value. In addition, the correlation is also expected to be maximum as the absolute value of r (*i.e.*, $r := |r|$) tends to 1. Therefore, a hybrid combination between these two parameters in the form of $(1-r)\sigma_{\rm fit}$[9] should be minimum at the highest correlation of the data in the Boltzmann diagram. A discussion on the choice of the best criterion to determining the A_K value is presented in Appendix B.3.

The left bottom panel of Fig. 4.17 presents the distribution of the $(1-r)\sigma_{\rm fit}$ parameter as a function of the A_K values. For the case of N35, the minimum value of $(1-r)\sigma_{\rm fit}$ occurs for $A_K = 1.9$ mag and the corresponding de-reddened Boltzmann diagram is presented in the right bottom panel of the same figure. Three fits were performed in the data of the de-reddened Boltzmann diagram, first by considering the $\upsilon = 1$, and $\upsilon = 2$ plus $\upsilon = 3$ vibrational states and, then, by mutually fitting all the observed transitions. This last fit provides a measure of the mean physical parameters of the H_2 emission, which is

[8] The linear Pearson coefficient varies between -1 and 1. r values close to 1 indicate positive correlation, r values closer to 1 indicate anti-correlation and r values around 0 indicates weak correlation between the quantities

[9] Different approaches for determining the best parameter were attempted (*e.g.*, by normalising both r and $\sigma_{\rm fit}$ values before combining them or summing instead of multiplying the factors) but all the methods led to very similar results.

reported in Table 4.15 together with the results of the other sources. When considering only the lowest vibrational transitions ($v = 1$), the fit indicates that the H_2 emission is likely associated with a colder and denser phase of the gas (T_{ex} is about $(2.1 \pm 0.2)\dot{1}0^3$ K and the logarithm[10] of N_{tot} is 18.46 ± 0.11 dex). The same analysis considering only the higher v transitions ($v = 2$ and $v = 3$) leads to a relatively warmer and less dense gas (T_{ex} $= (2.5 \pm 0.2)\dot{1}0^3$ K and $\log(N_{tot}) = 18.20 \pm 0.10$ dex), which is expected for exciting higher energy transitions. Indeed, the $v = 3$ transitions display excitation energies higher than $16{\cdot}10^3$ K. These high vibrational states are more sensitive to higher temperatures and most of these lines are located at spectral ranges outside the K-band window (*e.g.*, in the J-band). Similar physical conditions are obtained when considering the whole dataset (T_{ex} $= 2432 \pm 57$ K and $\log(N_{tot}) = 20.39 \pm 0.03$ dex).

Table 4.15 - Physical parameters of the H_2 jet towards the HMYSOs.

ID	A_K (mag)	A_V (mag)	T_{ex} (10^3 K)	$\log(N_{tot})$ dex(cm^{-2})
N33	4.30	64.3	4.07±0.22	18.40±0.09
N35	1.45	21.7	2.33±0.17	18.40±0.14
N47	1.25	18.7	2.24±0.23	17.51±0.30
N58	1.45	21.7	2.34±0.13	18.19±0.16
V49	0.00	0.0	2.67±0.20	16.48±0.09
W33A	2.35	35.1	3.06±0.15	17.96±0.10
Median	1.45	21.7	2502	18.08
Mean	2.16	32.3	2784	17.82
Sigma	1.27	19.0	700	0.74

Notes: The parameters were estimated through the Boltzmann diagram fitting and considering all the available H_2 transitions with fluxes above a 1-σ threshold. The columns are: (1) ID of the source; (2) the extinction in the K-band; (3) the extinction in the V-band; (4) the excitation temperature; (5) the logarithm of the total column density. The uncertainty in (2) and (3) are ±0.05 and ±0.75 mag, respectively.

Figure 4.18 presents the Boltzmann diagrams derived for the other spectra from Fig. 4.15. Due to the questionable fitting of the Boltzmann diagram of the extended emission associated with source V49, it was removed from the further analysis. The computed A_K values ranged between 1.25 and 4.30 mag, with a median value of 1.45 mag, and the extinction at the V-band ranged between 18.7 and 64.3 mag, with a median value of 21.7 mag. C15 derived the mean visual extinction of a sample of large-scale H_2 jets ranging from 1 to 50 mag, with a median of 15 mag. The corresponding A_K values range from 0.1 to 5.7 mag, with a median value of 1.7 mag by considering the IR-reddening law from Cardelli et al.

[10] $\log(x)$ corresponds to the logarithm of x to the base 10.

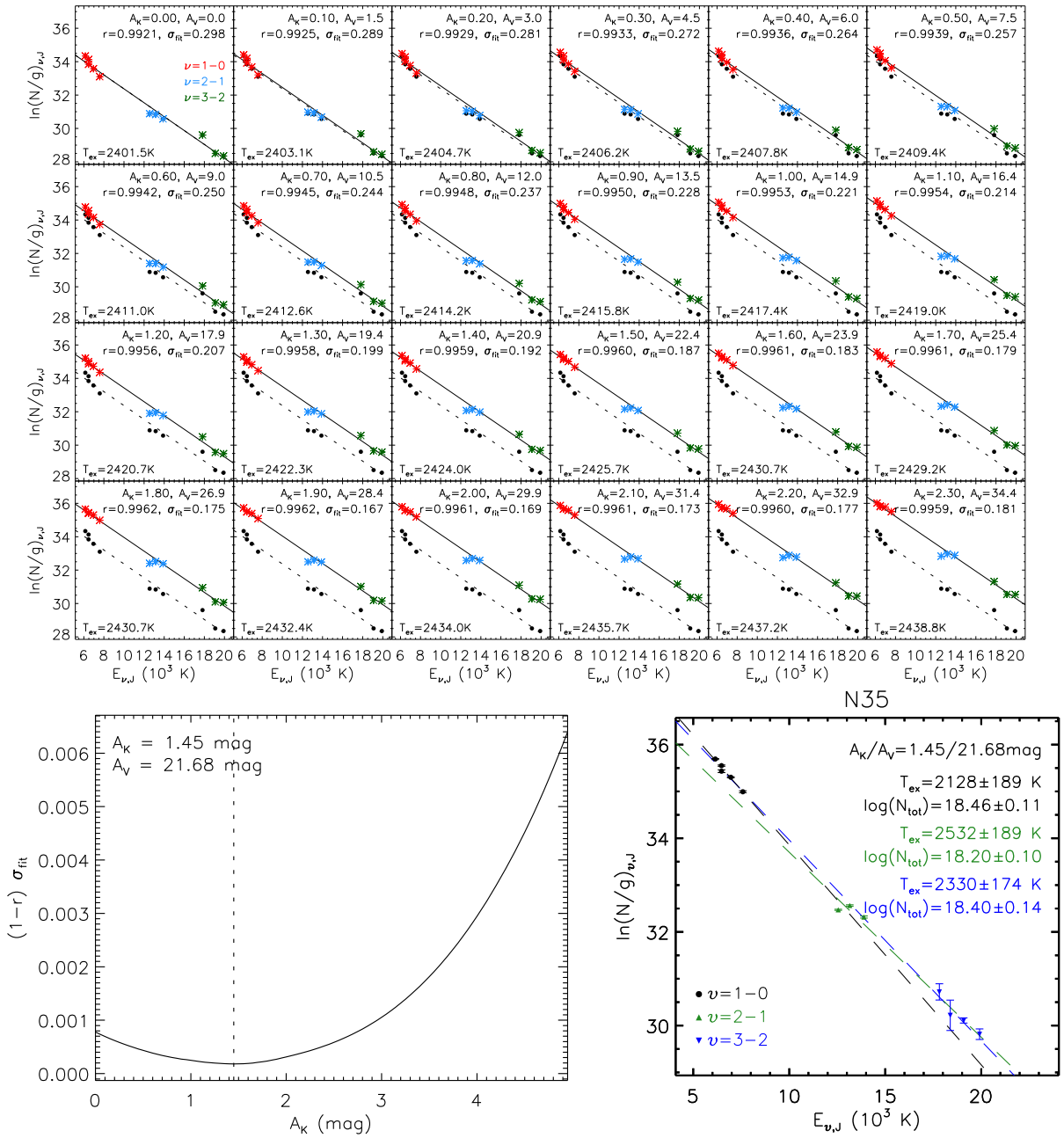


Figure 4.17: Boltzmann diagram of the H_2 gas associated with the extended emission detected towards the source N35. The diagram is constructed based on the intensity of different H_2 transitions present in the spectrum. Top panel: Variation of the A_K value from 0 to 2.4 mag, in steps of 0.1 mag. Points from different vibrational levels are shown in different colours (red: $\epsilon = 1$–0; blue: $\epsilon = 2$–1; green: $\epsilon = 3$–2;). The best fit to the data is indicated by the filled black line. The original dataset ($A_K = 0$ mag) is shown as black dots in each panel, together with its best fit (indicated by the dashed black line). Values of A_K, A_V, the linear Pearson coefficient of the data (r), the "standard deviation" of the residuals ($\sigma_{\rm fit}$) and the corresponding $T_{\rm ex}$ using the de-reddened points are indicated in each plot. Bottom left panel: the combination between the linear Pearson coefficient and the standard deviation of the residuals of the fit in the form of $(1-r)\sigma_{\rm fit}$ versus the A_K value. The best A_K value, indicated by the dashed black line, occurs at the minimum value of $(1-r)\sigma_{\rm fit}$. Bottom right panel: De-reddened Boltzmann diagram, assuming A_K =1.90 mag. The dashed lines indicate the best fit to the data using the $\epsilon = 1$–0 transitions only (black), the higher $\epsilon = 2$–1 and 3–2 transitions (green) and the whole dataset (blue). The corresponding temperatures and total column density are indicated in the upper-right corner of each panel.

(1989, , that is, $A_V/A_K = 8.77$-8.92) adopted in their work. Although the HMYSOs studied in this work and those investigated by C15 are distinct, the A_K mean value and the limits and median values derived in both works are equivalent, indicating no significant variation of the reddening of the jets at the vicinity of their launching point and at larger scales.

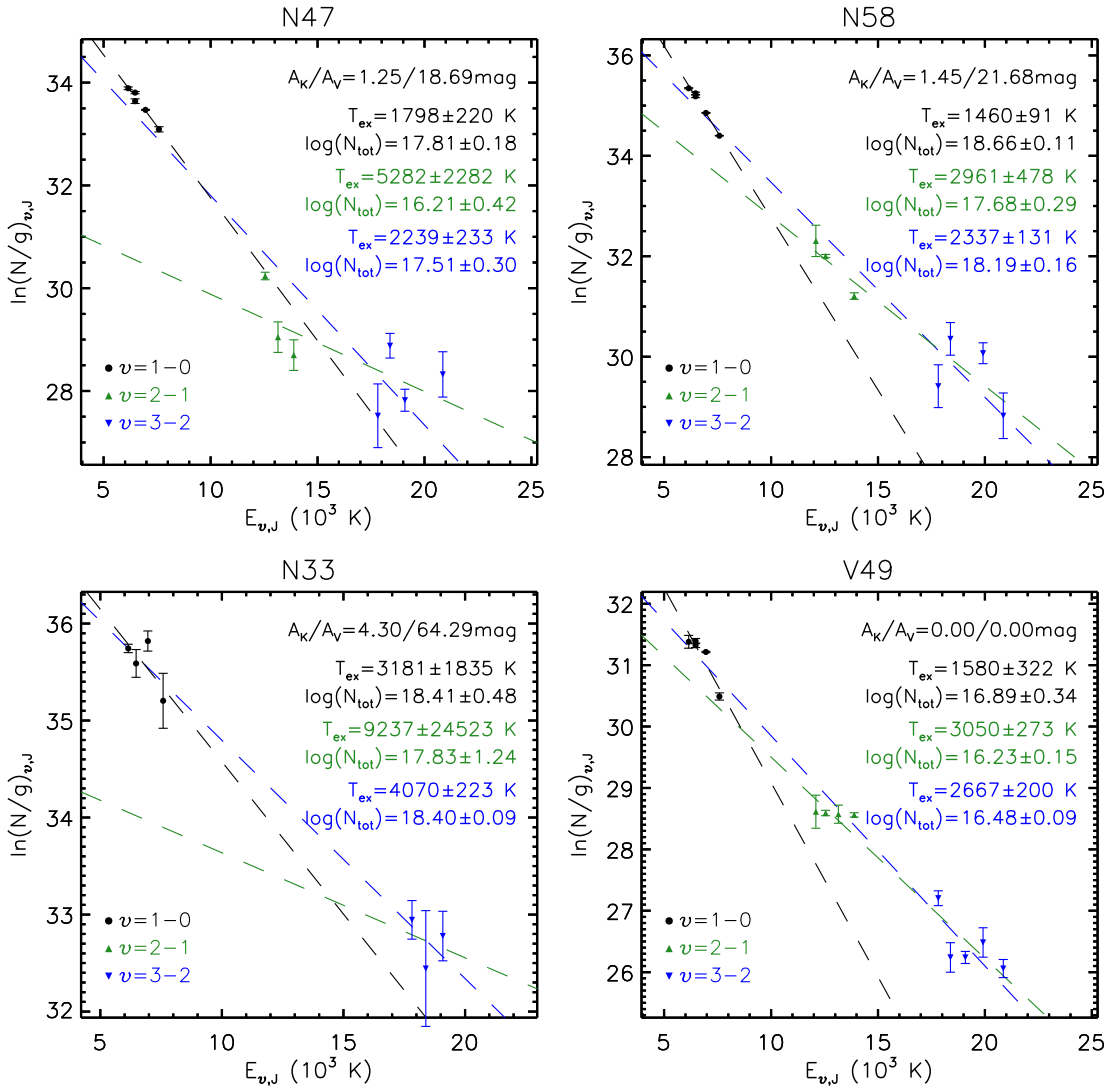


Figure 4.18: Boltzmann diagrams of the H_2 gas based on the average spectra of the extended H_2 emission associated with the HMYSOs. For a complete description of the plots, see Fig. 4.17.

The analysis of the physical parameters of the extended H_2 emission indicates that a colder and less dense gas is producing the low-J H_2 transitions ($J = 1$–0, black points). When including both low- and high-J H_2 transitions in the fitting of the Boltzmann diagrams, the median excitation temperature increases, followed by a decrease in the total column density. In every Boltzmann diagram (see Figs. 4.17 and 4.18) the low-J low-J H_2 transitions (in black) are often tracing colder and denser fraction of the extended emis-

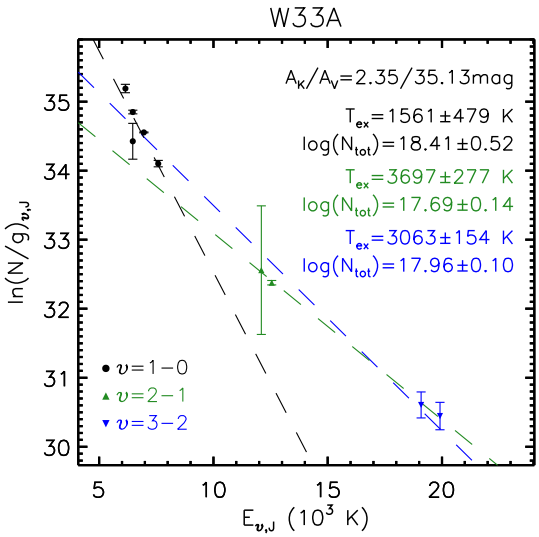


Figure 4.18: Continued.

sion, while higher-J transitions (green and blue points) are usually tracing warmer and less dense gas. The median temperature of the H_2 emission is about $2.5{\cdot}10^3$ K, ranging between $2.2{\cdot}10^3$ and $4.1{\cdot}10^3$ K. The excitation temperatures derived at the vicinity of the driven sources are relatively higher than those computed at larger scales (1200–3300 K, see C15), indicating a slight decrease of the temperature of the jets as a function of the distance from the source.

The median value of the total column density of the H_2 emission regions is about 10^{18} cm^{-2}. The densest H_2 gas are associated with sources N33 and N35, with total column density estimated in $N_{\rm tot} \approx 10^{18.4}$ cm^{-2}, while the more diffuse emission is found towards N47 and W33A ($N_{\rm tot} \sim 10^{17.5}$-$10^{18.0}$ cm^{-2}). The logarithm of the total column density of the large-scale jets from C15 ranges between 17.9 and 20.0 dex, with a median value of 18.7 dex. The apparent increase of $N_{\rm tot}$, from smaller towards larger scales, suggests that a larger fraction of the surrounding ISM is excited as the jet propagates farther away from its launching point. In addition, a decrease of $T_{\rm ex}$ is also observed, similarly as observed when analysing the NIFS data alone.

Finally, the Spearman coefficient factor (ρ) and its associated probability (p) between the physical parameters of the H_2 emission and the bolometric luminosity of the main HMYSO (see Table 4.1) was computed. No strong correlation between the temperature or the total column density and $L_{\rm bol}$ was found ($\rho = -0.41$ and $p = 0.42$ for $T_{\rm ex}$ against $L_{\rm bol}$; and $\rho = +0.18$ and $p = 0.74$ for $N_{\rm tot}$ *versus* $L_{\rm bol}$), suggesting that the physical parameters

of the jets are not likely directly linked to their driven sources. On the other hand, these results might be biased by the low number of sources in the present study. Therefore, a more statistically significant study is required to constrain any relation between the physical parameters of the jets with the properties of their driven sources.

4.3.4.2 Pixel-to-pixel determination of the physical parameters of the jets

By taking advantage of the spatial information available in the NIFS data, the same procedure adopted for evaluating the physical parameters of the H_2 emission from the average spectra from Fig. 4.15 was used to compute the physical parameters from every spaxel of the datacubes. Spatial maps of A_K, $T_{\rm ex}$ and $N_{\rm tot}$ were obtained for every source by measuring the flux of the H_2 transitions present in the K-band spectra and fitting the Boltzmann diagram for each spaxel in the datacubes.

The mass and luminosity of the H_2 gas were evaluated as presented below. First, the mass of the H_2 gas (M_{H2}) was computed based on the total column density of the H_2 molecules as:

$$M_{H2}(x,y) = 2\mu m_H A_{\rm pix} d^2 N_{\rm tot}(x,y) \tag{4.11}$$

where the (x,y) indices correspond to the position of a given spaxel on the datacube, $\mu = 1.24$ is the mean atomic weight of the H_2, $m_{\rm H}$ is the proton mass (in g) and $A_{\rm pix}$ corresponds to the projected area of the spaxel[11] (in cm^2). Then, the total H_2 luminosity ($L(H_2)$) was computed by following the procedure presented in Caratti o Garatti et al. (2006) and briefly described below. The total flux of the H_2 (I_{H2}) was estimated using their LTE calculations, based the temperature of the gas ($T_{\rm ex}$) and the intensity of the H_2 (1–0) S(1) transition at 2.12 μm ($I_{2.12}$). Figure 4.19 presents the distribution of the ratio between $I_{2.12}$ and I_{H2} (hereafter, f_{H2}), as a function of $T_{\rm ex}$. By computing the pixel-to-pixel intensity of the 2.12 μm line, a map of the total luminosity of the H_2 was obtained as:

$$L_{\rm H2}(x,y) = A_{\rm pix} \cdot \frac{I_{2.12}(x,y)}{f_{\rm H2}(T_{\rm ex}(x,y))} \tag{4.12}$$

The maps of all the physical parameters of the H_2 gas in the FOV of the source N35 are presented in Fig. 4.20. In each map, the contours of the H_2 (1–0) S(1) flux are placed at 3^n-σ ($n = 0, 1, 2, ...$). The range of each parameter was chosen in order to improve the

[11] For NIFS observations, $A_{\rm pix} = 5.596 \cdot 10^{29} (d/{\rm kpc})^2$ cm^2

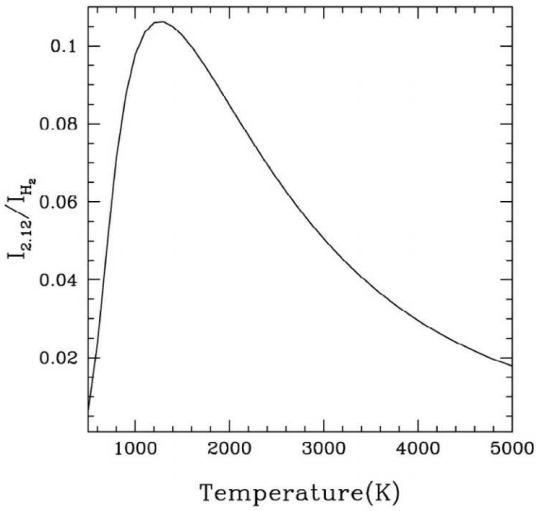


Figure 4.19: Ratio between the intensity of the H_2 (1–0) S(1) line at 2.12 μm and the total intensity of the H_2 emission as a function of the excitation temperature of the gas in LTE conditions. Extracted from Caratti o Garatti et al. (2006).

visualisation of the maps.

The top left panel of Fig. 4.20 presents the distribution of the *K*-band extinction in the FOV of the source N35. The median *K*-band extinction, computed over the spaxels within the 3-σ contour, is $A_K = 0.9 \pm 0.5$ mag. The map exhibits relatively high reddening values ($\gtrsim 1.5$ mag) closer to the position of the HMYSO and the bright bow-shock at $\sim 1''$ in the E direction of the source. The excitation temperature of the H_2 emission is presented in the top middle panel of the same figure. The median $T_{\rm ex}$ value is $(2.5 \pm 0.2) \cdot 10^3$ K over the same 3-σ region. The total column density of the H_2 gas is presented in the top right panel of the figure. The distribution of $N_{\rm tot}$ is similar to the flux of the H_2 (1–0) S(1) line, indicated by the contours, indicating that the region closer to the main HMYSO and the bow-shock structure located to the E are denser ($N_{\rm tot} \gtrsim 10^{20}$ cm^{-2}) than the median $N_{\rm tot}$ value of the H_2 gas in the FOV ($N_{\rm tot} \approx 10^{18.7}$ cm^{-2}). The comparison between the temperature and density of the H_2 gas suggests that both quantities are anti-correlated. Indeed, the region around the central source exhibits temperatures below $2.5{\cdot}10^3$ K and is associated with the highest total column density values in the FOV ($N_{\rm tot} \gtrsim 10^{20}$ cm^{-2}). Similar physical conditions are also found towards the front region of the bow-shock structure located to the E direction. On the other hand, the H_2 emission situated between the main HMYSO and the bow-shock is associated with relatively larger temperatures ($T_{\rm ex} \gtrsim 3.5{\cdot}10^3$ K) and less dense gas ($N_{\rm tot} \lesssim 10^{19}$ cm^{-2}).

The last two panels in Fig 4.20 exhibit the mass and the luminosity of the H_2 gas, com-

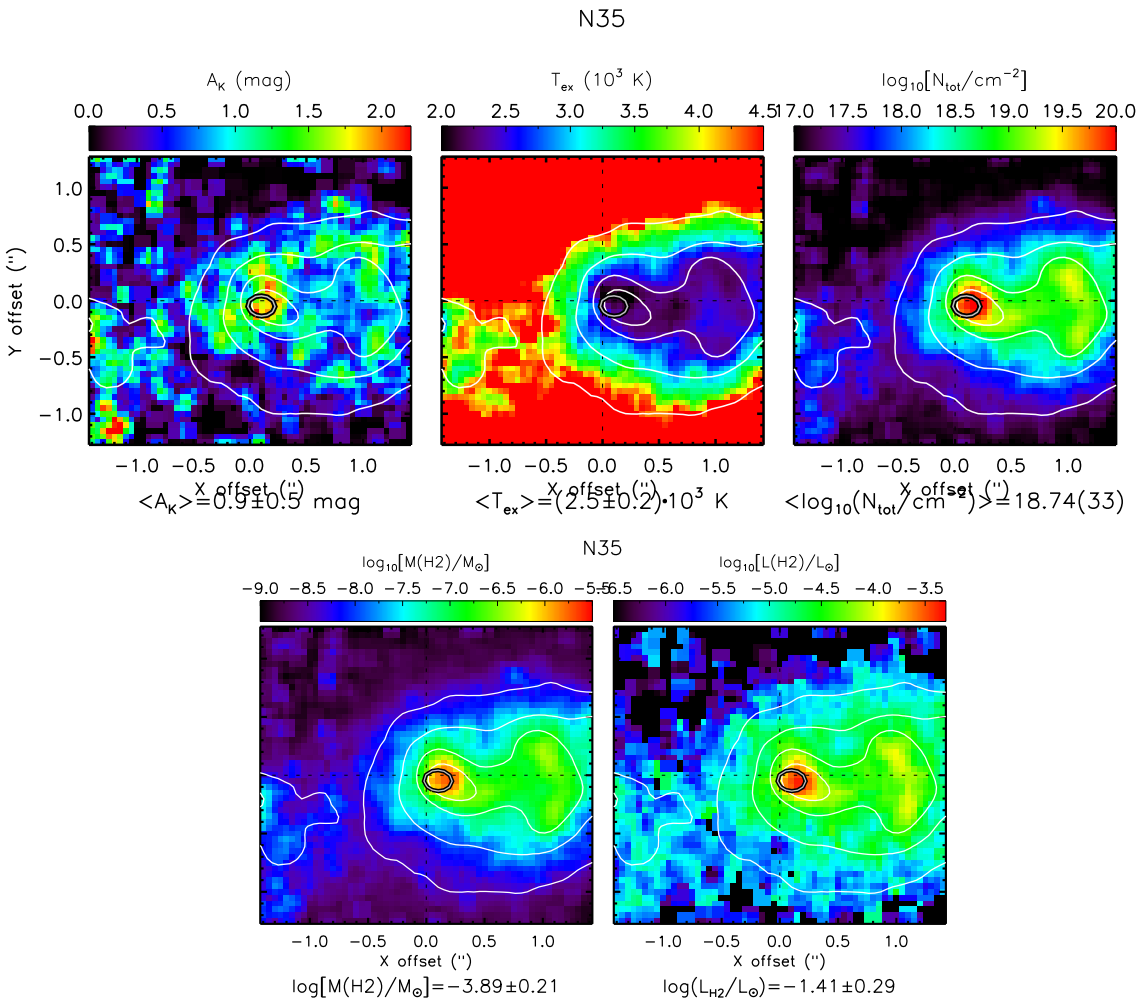


Figure 4.20: Top panels: Spatial distribution of the K-band reddening (left), the excitation temperature (middle), the total column density (right) of the H_2 emission associated with the source N35. Bottom panels: spatial distribution of the mass (left) and the luminosity (right) of the H_2 emission. The position of the main source is indicated by the black ellipsoid. To improve the visualisation, a boxcar average filter with a width of 2 pixels was applied to each map. The dynamical range of the values used for displaying each map is indicated by the corresponding colour-bar. The white contours are placed at 3^n-σ (with $n = 0, 1, 2, ...$) levels of the H_2 1–0 S(1) emission at 2.12 μm. For A_K, T_{ex} and N_{tot}, the median values of the spaxels within the 3-σ contour are indicated below each map. For $M(H_2)$ and $L(H_2)$, the values presented below each map are integrated over the 1-σ contour. The error on each value is calculated as the standard deviation of the values on the corresponding map.

puted using Eqs. (4.11) and (4.12). The values presented below each of these maps present the integrated mass and luminosity over the 1-σ contour. The distribution of $M(H_2)$ and $L(L_2)$ are similar to the map of the total column density and, therefore, does not add new information to the analysis. However, the integrated values the mass and luminosity of the H_2 are important quantities that can be further compared with the properties of the main sources. The maps for the other sources are presented in Appendix B.5 and the results obtained from the pixel-to-pixel analysis (p2p) are summarised in Table 4.16.

The physical parameters derived from the average spectra of the H_2 emission (Table 4.13) and the mean physical parameters from the p2p analysis (Table 4.16) are compatible within 1- or 2-σ, indicating that the results obtained from both methods are, in general, consistent.

Figure 4.21 compares the physical parameters of the H_2 gas, derived from the p2p analysis (listed in Table 4.16), against the bolometric luminosity of the HMYSOs. The linear Spearman factor is presented for each comparison. Except for the excitation temperature, all the other parameters of the H_2 gas exhibit a moderate correlation with $L_{\rm bol}$ ($\rho \gtrsim 0.6$), suggesting that more luminous YSOs are driving denser, more massive and brighter jets at the linear scales probed by the NIFS data ($\sim 10^3$ AU). However, due to the small number of points in the plots, these relations should be further explored by including more sources in the analysis.

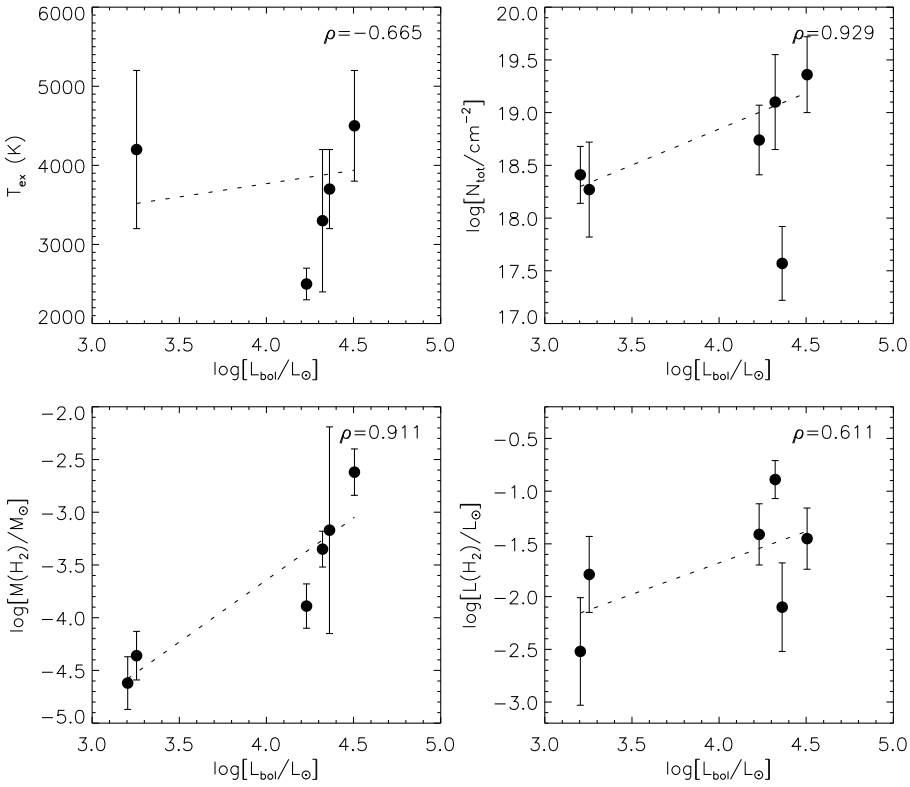


Figure 4.21: Distribution of the physical parameters of the H_2 gas as a function of the bolometric luminosity of the driven source. The linear Spearman factor is presented on the top right corner of each panel. The black dashed line indicates the best fit to the data.

The relations between each pair of the physical parameters of the gas (*e.g.*, $N_{\rm tot}$ against $T_{\rm ex}$) were obtained and analysed as follows. The individual values of each spaxel of one physical parameter were plotted against the other parameters, and the distribution of the

Table 4.16 - Physical parameters of the H_2 emission based on the p2p calculation.

	A_K (mag)	T_{ex} (10^3 K)	$\log(N_{tot}/cm^{-2})$	$\log[M(H_2)/M_\odot]$	$\log[L(H_2)/L_\odot]$
N33	0.6 ± 0.6	9.0 ± 1.7	18.41 ± 0.27	-4.62 ± 0.25	-2.52 ± 0.51
N35	0.9 ± 0.5	2.5 ± 0.2	18.74 ± 0.33	-3.89 ± 0.21	-1.41 ± 0.29
N47	1.4 ± 1.0	4.2 ± 1.0	18.27 ± 0.45	-4.36 ± 0.23	-1.79 ± 0.36
N58	2.1 ± 0.8	3.3 ± 0.9	19.10 ± 0.45	-3.35 ± 0.17	-0.89 ± 0.18
V49	1.3 ± 0.8	3.7 ± 0.5	17.57 ± 0.35	-3.17 ± 0.98	-2.10 ± 0.42
W33A	4.1 ± 1.0	4.5 ± 0.7	19.36 ± 0.36	-2.62 ± 0.22	-1.45 ± 0.29
Median	1.35	3.70^a	18.58	-3.62	-1.62
Mean	1.73	3.64^a	18.58	-3.67	-1.69
Sigma	1.27	0.79^a	0.64	0.76	0.57

Notes: The columns are: (1) ID of the source; (2) the *K*-band extinction and its associated error (in mag); (3) the excitation temperature and its associated error (in K); (4) the total column density and its associated error (in dex units of cm^{-2}); (5) the total mass of H_2, integrated over the 1-σ contour shown on each map (in units of $M_\odot$); (6) the total luminosity of H_2, integrated over the 1-σ contour shown on each map (in units of $L_\odot$). (^{a}) the results for source N33 were excluded from the statistics computed for the T_{ex} parameter.

points for the full map was obtained as presented in Fig. 4.22 for the N35 source. To quantify the strength of the correlation between the pairs of parameters, the Spearman correlation rank coefficients (ρ) were obtained for all comparisons.

Fig. 4.22 indicates that some quantities are reasonably well-correlated (*e.g.*, T_{ex} vs. $\log(N_{tot})$, T_{ex} vs. $\log[M(H_2)]$) exhibiting $|\rho| \geq 0.9$. Others comparisons are less correlated, with $|\rho| \lesssim 0.6$ (*e.g.*, T_{ex} vs. A_K, $\log(N_{tot})$ vs. A_K).

The comparison between pairs of the primary physical parameters (A_K, T_{ex} and N_{tot}) indicates that T_{ex} and N_{tot} are anti-correlated while both quantities exhibit weak correlations with A_K. As previously described on the fitting the Boltzmann diagrams, the A_K parameter is the first parameter set before fitting the data and, therefore, it is uncorrelated with T_{ex} and N_{tot} values. On the other hand, the strong dependency observed between these last quantities ($\rho = -0.941$) originates from the fact that they are simultaneously evaluated from the fit (see Fig. 4.17). As previously observed in the maps from Fig. 4.20, T_{ex} is anti-correlated with N_{tot} (that is, the temperature increases as the gas get less dense). In general, comparisons between the properties of the gas with the corresponding A_K value display the weakest correlations in the analysis ($|\rho| \lesssim 0.7$), while comparisons between the other four parameters (N_{tot}, T_{ex}, $M(H_2)$ and $L(H_2)$) are stronger correlated ($|\rho| \gtrsim 0.85$). The plots for the other sources are presented in Appendix B.5. Similar results were found for the other sources.

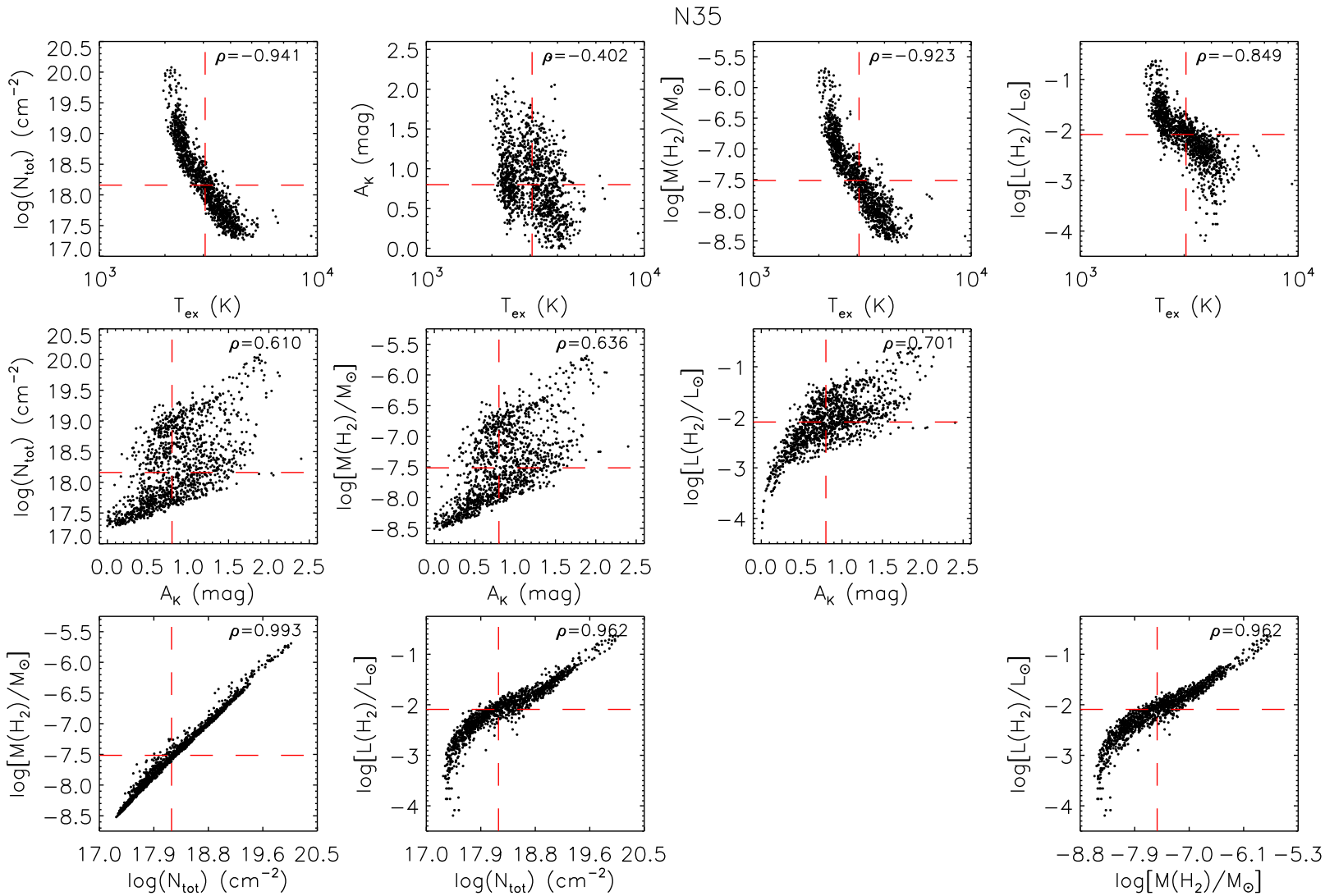


Figure 4.22: Correlation between the physical parameters of the H_2 gas for source N35, derived for each spaxel of the NIFS datacube. The Spearman test (ρ) is indicated on the top right corner of each panel. The mean values on the x- and y-axis are indicated as red dashed lines.

4.3.5 Brγ *emission towards the HMYSOs*

The Brγ feature is a typical hydrogen recombination line (H I, $n = 7$–4 at 2.166 μm) emitted from ionised regions. This line is often used as a tracer of stellar or disc winds towards a wide range of stellar masses (e.g., for Herbig Ae/Be objects, see Corcoran and Ray 1998; for HMYSOs, seeDrew et al. 1998; Davies et al. 2010; Murakawa et al. 2013), exhibiting velocities ranging from $\sim$200-700km s^{-1} (Lumsden et al., 2012). In the case of HMYSOs, previous studies have suggested that the Brγ emission traces the cavity of the ionised jet closer to the main source (*e.g.*, W33A, from D10) or arises from a disc-driven wind (*e.g.*, S106 and I18151, from M13). In the first scenario, the spatial distribution of the Brγ emission is aligned with the axis of the molecular jet, while in the case of disc-driven winds, the Brγ emission is likely to be perpendicularly aligned to the jet. This section presents the analysis of the Brγ feature to test which of these scenarios explains the origin of the line observed towards the HMYSOs investigated in this thesis.

As previously presented on the middle column of Fig. 4.14 (see Sect. 4.3.3), the integrated maps of the Brγ emission are likely tracing the main point-like objects in each FOV, but are also outlining some extended structures. In some cases, the extended Brγ emission is similar to those traced by the H_2 (1–0) S(1) line (*e.g.*, see N35, N47, N58, and W33A), while other sources exhibit a different morphology of the extended Brγ emission when compared to those probed by the molecular line (*e.g.*, N33 and V49). Moreover, the Brγ maps appear as similar as the continuum maps (see Fig. 4.9). Accordingly to Cooper et al. (2013), the temperature of the gas that originates the Brγ emission lies between 7500 and 30000 K. Such limits correspond to the typical temperature of an UCH II region and an OB star, respectively. Therefore, the correlation between the Brγ and the continuum emission indicates that a warmer component of the gas is producing the H I emission, while a colder gas phase is responsible for the more extended H_2 emission detected at 2.12 μm. Indeed, the median excitation temperatures of the H_2 emission found in Sects. 4.3.4 ($(2.5 \pm 0.7)\cdot10^3$ K) and 4.3.4.2 ($(3.7 \pm 0.8)\cdot10^3$ K) are below 7500 K.

A detailed view of the Brγ features, extracted from the normalised spectra from Fig. 4.12 (see Sect. 4.3.2), is presented on a velocity scale in Fig. 4.23. The left panel of Fig. 4.23 presents the Brγ profiles in the rest frame of the sources, using the $V_{\rm lsr}$ values reported in Table 4.1. The analysis of the profiles presented in both panels indicates that

most of the peaks of the Brγ emission are offset from the rest position (that is, 0 km s^{-1} corresponds to $\lambda_0 = 2.1655\,\mu$m), suggesting that the real $V_{\rm lsr}$ of the HMYSOs disagrees from the $V_{\rm lsr}$ value derived from observations at higher angular resolutions ($\sim 30''$ for NH_3 (1,1), Urquhart et al., 2011). Thus, the $V_{\rm lsr}$ value was corrected in order to match the position of the peak intensity of the Brγ profile with the $V_{\rm lsr}$ of the line. The best results are presented in the middle panel of Fig. 4.23. The strongest Brγ emission is found towards the source V49, while the faintest emission is associated with N16. The strength of the Brγ feature is often correlated with the accretion rate in low-mass stars (Muzerolle et al., 1998; Beck et al., 2010). Later, Cooper et al. (2013) found a good agreement between the Brγ intensity and the accretion luminosity for a wide interval of $L_{\rm bol}$, even including more luminous YSOs.

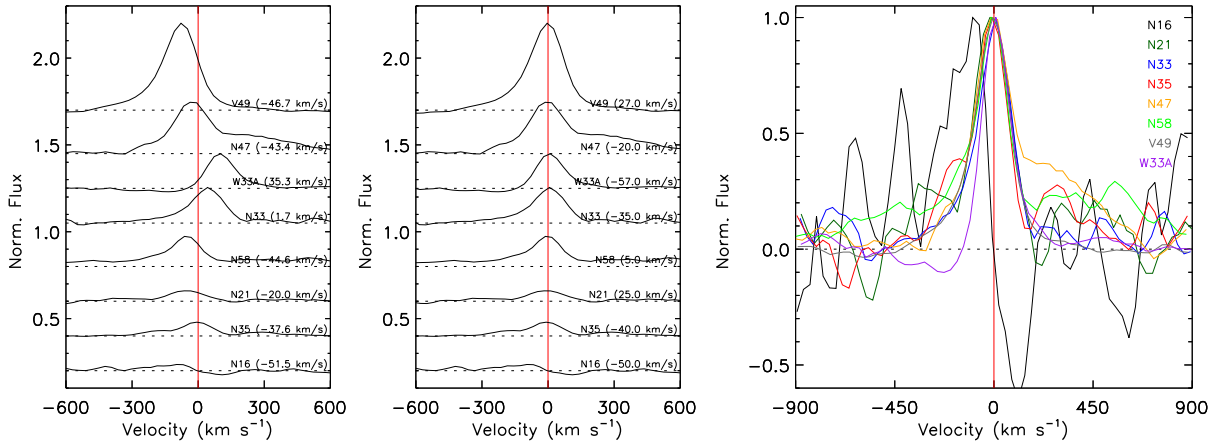


Figure 4.23: Left panel: A zoom into the Brγ feature of the normalised spectra from Fig. 4.12, presented in the velocity scale. Middle panel: Same data, but after realignment of the Brγ peak position with respect to the rest position of the line. The best $V_{\rm lsr}$ value is indicated to the right of each profile. Right: the Brγ profile normalised by its peak intensity for all the sources. The profiles were velocity-corrected by using the $V_{\rm lsr}$ value from Table 4.1. The red vertical line is placed at 0 kms and the baseline of each spectrum is indicated by the horizontal black dashed lines.

The right panel of Fig. 4.23 presents the Brγ feature normalised by its peak intensity on the same velocity scale for all the sources. The plot indicates that the Brγ profiles are broad, covering a velocity interval ranging from ~ -800 to $\sim +800$ kms. The broad distribution of the Brγ emission suggests that the feature is tracing the quiescent gas at velocities closer to the peak of the line, but it also exhibits a high-velocity component, probably associated with rotating discs (*e.g.*, S106, see M13), expanding H II shells (*e.g.*, AFGL 2136, see M13) or outflow cavities (*e.g.*, W33A, see D10). As previously discussed by D10, the Brγ feature of W33A displays a P-Cygni absorption in a velocity range between

-500 and $-100\,\mathrm{km\,s^{-1}}$, and an extended red wing. Although the majority of HMYSOs in the sample is associated with broad Brγ wings, only source N33 exhibits an inverse P-Cygni Brγ profile. As opposed to W33A, from which the P-Cygni profile is expected to be originated by winds (Murakawa et al., 2013), inverse P-Cygni profiles are often associated with edge-on systems (at high inclination angles), where the temperature in the accretion flow decreases toward larger distances from the central region (Sicilia-Aguilar et al., 2017).

A spectro-astrometric analysis was conducted to investigate the spatial variation of the Brγ emission as a function of its velocity scale. The spectro-astrometry consists on the measurement of the positional centroid of the emission along the wavelength range, with superior accuracy than the diffraction limit. This technique has been extensively used on NIFS data, providing unprecedented results on the structures associated with HMYSOs (*e.g.*, see Davies et al., 2010; Murakawa et al., 2013, and references therein). By following the procedure described in D10, an `IDL` routine was written to perform the spectro-astrometry of the Brγ feature associated with each source investigated in this thesis.

The routine `IFU_centroid_fitting.pro` performs the two-dimensional centroid fit of the spatial distribution of the emission along the spectral direction of the datacube. The validation of the procedure was performed by comparing the results of W33A with those obtained by D10 (cf. Fig. 12 from D10). Any large-scale gradient on the variation of the centroid position was removed by fitting a robust 3rd-order polynomial to the data using the ROBUST_POLY_FIT function. Several centroid fitting `IDL` routines (*e.g.*, CNTRD, CENTROID, GCENTROID, FWCENTROID and GAUSS2DFIT) were tested to check if the results are robust or biased by choosing a different fitting algorithm. The most similar results as those from D10 were obtained when using the FWCENTROID algorithm and, thus, it was used to perform the spectro-astrometry analysis for the other sources. A note of caution must be given due to the contribution of the continuum on the total flux of the line emission. Indeed, such high accuracy can only be achieved when considering the 'continuum+line' datacube and, therefore, the measured positional offsets obtained through the spectro-astrometry method lie nearer to the continuum than a pure line spectrum would provide.

The spectro-astrometry of the Brγ feature towards the source W33A is presented in Fig. 4.24. The top left panel presents the offset of the centroid in both right ascension and declination axes as a function of the wavelength. As similarly found by D10, the

centroid exhibit a significant shift from south to north, together with a shift from east to west directions. The bottom left panel presents the Brγ profile at the same wavelength scale. The coloured scale indicates the effective spectral region of the Brγ feature. Finally, the right panel presents the spatial variation of the centroid position as a function of the velocity of the emission, indicated by the colour bar and matching the colour scale of the Brγ profile (see bottom left panel).

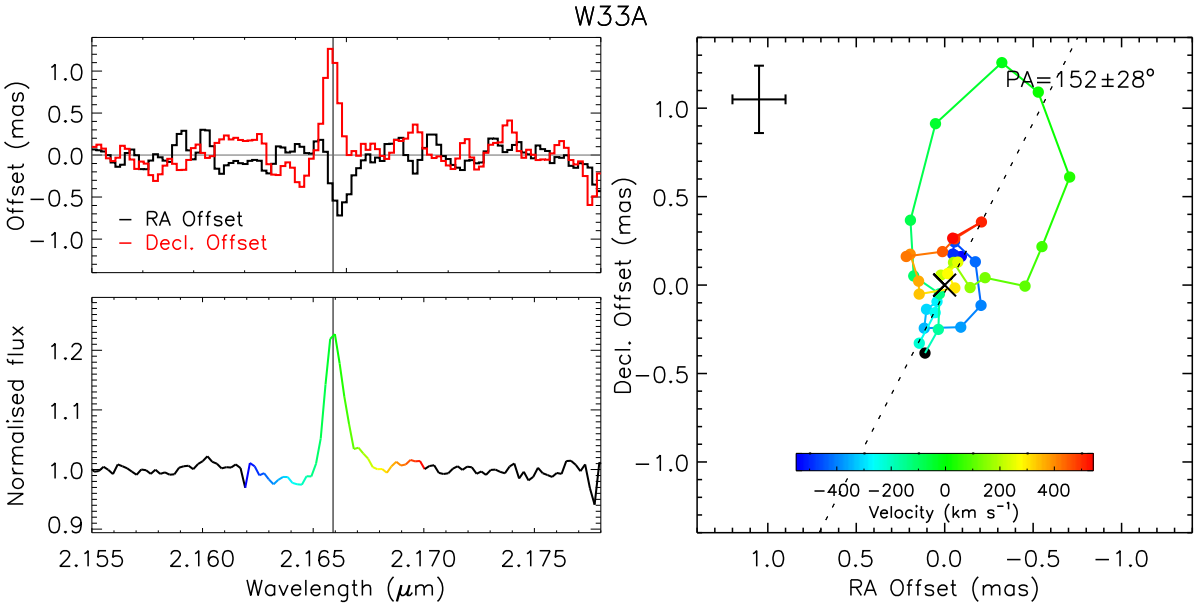


Figure 4.24: Spectro-astrometry of the Brγ feature for the source W33A. Left: the top panel shows the variation of the centroid of the emission in the right ascension (RA, in black) and declination axis (Decl., in red) as a function of the wavelength. The baseline is indicated by a horizontal grey line. The peak of the Brγ feature (at 0 km s^{-1}) is indicated by a vertical line. Bottom left: the normalised flux of the Brγ emission. The coloured region indicates the spectral interval shown on the spatial variation map of the Brγ emission. Right panel: Spatial variation of the Brγ emission. The velocity with respect to the peak of the emission is indicated by the horizontal colour bar shown in the bottom region of the plot. The dispersion of the data, estimated using the corresponding telluric standard, is shown in the left top corner of the plot. The dashed line indicates the position-angle (PA, in degrees from N to E), measured from the reddest to the bluest high-velocity positions.

The blue and red high-velocity components (corresponding to velocities of -300 to $+400$ km s^{-1}) are offset by roughly $\sim$1 mas, from the south-east to the north-west direction. Such angular scale corresponds to $\approx$2.4 AU at the distance of 2.4 kpc, indicating that the centroid of the Brγ feature is tracing coherent variation of the emission at sub-pixel scales towards the central object. The displacement between the blue and red high-velocity components roughly follows a linear distribution with a position angle (PA, measured from N to E) of $150 \pm 28°$, which is compatible with the PA $= 113 \pm 25°$ derived by D10. By following the discussion in D10, the spectro-astrometry of the Brγ feature is probing the base of the bipolar wind/jet traced by the H_2 emission (see Fig. 4.14), which is aligned

with the large-scale outflow identified towards the source (see Fig. 1 from D10).

The accuracy of the technique was estimated based on the analysis of the centroid offset of the Brγ feature of the telluric standard star, as presented in Fig. 4.25. In the case of the telluric standard stars, the Brγ feature is originated at the unresolved stellar photosphere and appears in absorption. Therefore, the Brγ feature is not expected to exhibit any coherent spatial variation across the spectral direction. Indeed, Fig. 4.25 shows no systematic spatial variation of the Brγ feature towards HR 6798. The standard deviation of the centroid position of the Brγ emission in the 'RA' and 'Decl' directions, were estimated as 0.15 and 0.19 mas respectively. These values were adopted as the error on the centroid position of the results presented in Fig. 4.24.

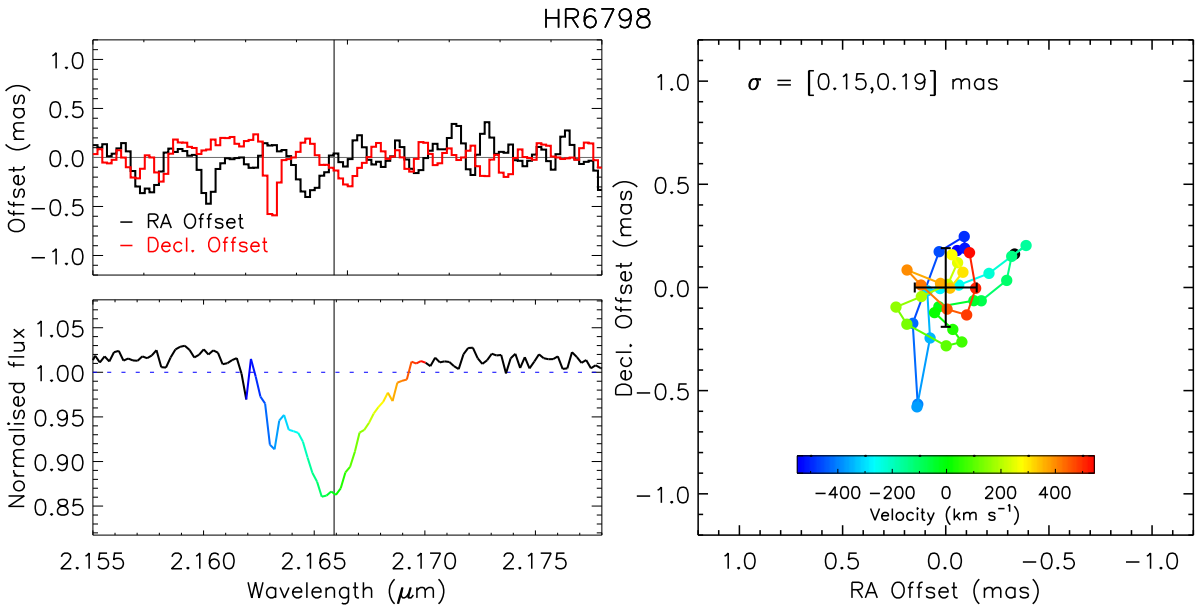


Figure 4.25: Same as Fig. 4.24, but displaying the spectro-astrometry of the Brγ feature of the HR 6798 source, used as the telluric standard star for W33A observations. The error bars placed at the (0,0) position indicate the standard deviation on each axis, whose values are indicated in the top left region of the map.

Figure 4.26 presents the spectro-astrometry of the Brγ feature towards the other HMYSOs from the sample. The corresponding plots of the telluric standard stars are presented in Fig. B.33. The individual analysis of the spectro-astrometry of the Brγ feature towards the HMYSOs is discussed below. The main results are summarised in Table 4.17.

N47: The spectro-astrometry of the Brγ feature associated with the source N47 (see Fig. 4.26) indicates a wide-angle structured detected at velocities ranging from -450 to $-250\,\mathrm{km\,s^{-1}}$. A red-shifted small-angle counterpart was detected at velocities between $+250$ to $-450\,\mathrm{km\,s^{-1}}$. The angular separation between the median position of the blue

and red structures is about 3.2 mas, corresponding to a linear scale of 9.8 AU at the distance of 3.1 kpc. The structure is roughly oriented in the north-south direction, with a PA = 341±32°, and is aligned with the large-scale H_2 emission detected by Nav15 (PA = 10°, see Fig. 4.1). The wide-angle blue-shifted structure traced by the Brγ is perpendicular to the bright H_2 lobe identified in Fig. 4.14, suggesting that the Brγ is probing the cavity/ionised region around the launching point of the H_2 jet. No water maser emission was found towards the RMS object (Urquhart et al., 2011).

N58: The positional offsets between the blue- and red-shifted emission are well separated by ≈1.5 mas in the east-west direction (PA = 67±14°), corresponding to a linear scale of 7.4 AU at 4.9 kpc. The structure traced by the Brγ feature is perpendicular to the inner H_2 jet identified in the large-scale maps from Nav15 (see Fig. 4.1) and also with the brightest H_2 jet identified in the H_2 map from Fig. 4.14. This configuration suggests that the Brγ emission is tracing a circumstellar disc around the HMYSO, similarly as found for source S106 by Murakawa et al. (2013), who assumed a Keplerian rotation for the gas to estimate the mass of the source. The Keplerian mass is given by:

$$M = \frac{rv(r)^2}{G} \approx 2.3 \left(\frac{r}{\mathrm{AU}}\right) \left(\frac{v(r)}{\mathrm{km\,s^{-1}}}\right)^2 \mathrm{M_\odot} \tag{4.13}$$

where r corresponds to the distance to the central position (in AU), $v(r)$ is the velocity at the position r and G is the gravitational constant. Thus, by assuming a Keplerian rotation model for the gas with an edge-on orientation and radial velocities of ±300-400 km s^{-1} at 7.4 AU, the total mass was estimated as out to be $\sim (1.0\pm0.4)\cdot10^3$ M$_\odot$. The RMS source has a bolometric luminosity of $2.1\cdot10^4$ L$_\odot$, which translates into a 17-28 M$_\odot$ object assuming a mass-luminosity relation with a power-law index between 3.0 and 3.5. The masses obtained from the Keplerian analysis and the L-M relation are inconsistent, indicating that the observed velocities of the Brγ feature are not likely following a Keplerian distribution.

V49: The positional offset between the blue- and red-shifted centroids corresponds to an angular offset of 0.64 mas (∼2 AU at 3.1 kpc), roughly oriented in the east to west direction (PA = 104±24°). The orientation of the Brγ is roughly perpendicular to the H_2 jet identified in the FOV (PA = 325°, see Table 4.12), which is associated with a nearby source located to the South direction of the main object in the field. The observed distribution

of the Brγ feature suggests that it is tracing a disc-like structure, with maximum blue- and red-shifted velocities around -400 and $+300\,\mathrm{km\,s^{-1}}$. Assuming a Keplerian model for these velocities, the mass of the central source is estimated in $\sim(2.8\pm1.1)\cdot10^2\,\mathrm{M_\odot}$. The bolometric luminosity of the RMS source associated with V49 is $1.8\cdot10^3\,\mathrm{L_\odot}$, which is consistent with a 8-12 $\mathrm{M_\odot}$ object. As similarly found for source N58, the Keplerian mass of source V49, derived from spectro-astrometry of the Brγ feature, is inconsistent with the mass obtained from the mass-luminosity relation.

N33: The spectro-astrometry of the Brγ feature associated with source N33 exhibits larger offsets ($\sim$6 mas) when compared to most of the sources due to its relatively proximity. Indeed, N33 is the closest object of the sample and, therefore, larger offsets are expected for the same region probed towards more distant objects. The observed offset between the blue- and red-shifted components corresponds to a linear size of $\sim$10 AU at the distance of 1.4 kpc. This source exhibits a positional offset between the blue- and the red-shifted centroids of $\sim$7 mas in the N to S direction, with a PA estimated in $148\pm10°$. The kinematic structure probed by the spectro-astrometry is roughly aligned with the direction of the H_2 jets identified in Fig. 4.14, suggesting that the Brγ feature is tracing the cavity of the outflow at scales of $\sim$10 AU.

N35: The plot of this source is rotated in 90° counter-clockwise, properly matching the RA and Decl. axes with the x- and y-axis, respectively. This source exhibits a clear coherent positional offset between the blue- and the red-shifted centroids of $\sim$0.45 mas in the east to west direction. The structure probed by the Brγ feature is perpendicular to the H_2 jets identified in Fig. 4.14, suggesting that it arises from a circumstellar disc around the HMYSO. The angular offset corresponds to a linear scale of $\sim$1.7 AU at the distance of 3.7 kpc. The velocities of the blue and red wing are about $\pm$350-400 $\mathrm{km\,s^{-1}}$, corresponding to a Keplerian mass of $\sim(2.7\pm0.5)\cdot10^2\,\mathrm{M_\odot}$. The bolometric luminosity of the RMS source is $1.7\cdot10^4\,\mathrm{L_\odot}$, which leads to a stellar mass of 16-26 $\mathrm{M_\odot}$. As previously found towards sources N58 and V49, the Keplerian mass of the source N35, derived from the Brγ analysis, is inconsistent with the value obtained from the L-M relation.

N16 and N21:the results of the Brγ spectro-astrometry towards N16 and N21 sources correspond to the less significant among the sample due to the weak Brγ features (and, therefore, detected at lower signal-to-noise ratios) identified in their spectra. The positional offsets detected towards source N16 seem to exhibit random patterns, covering the whole

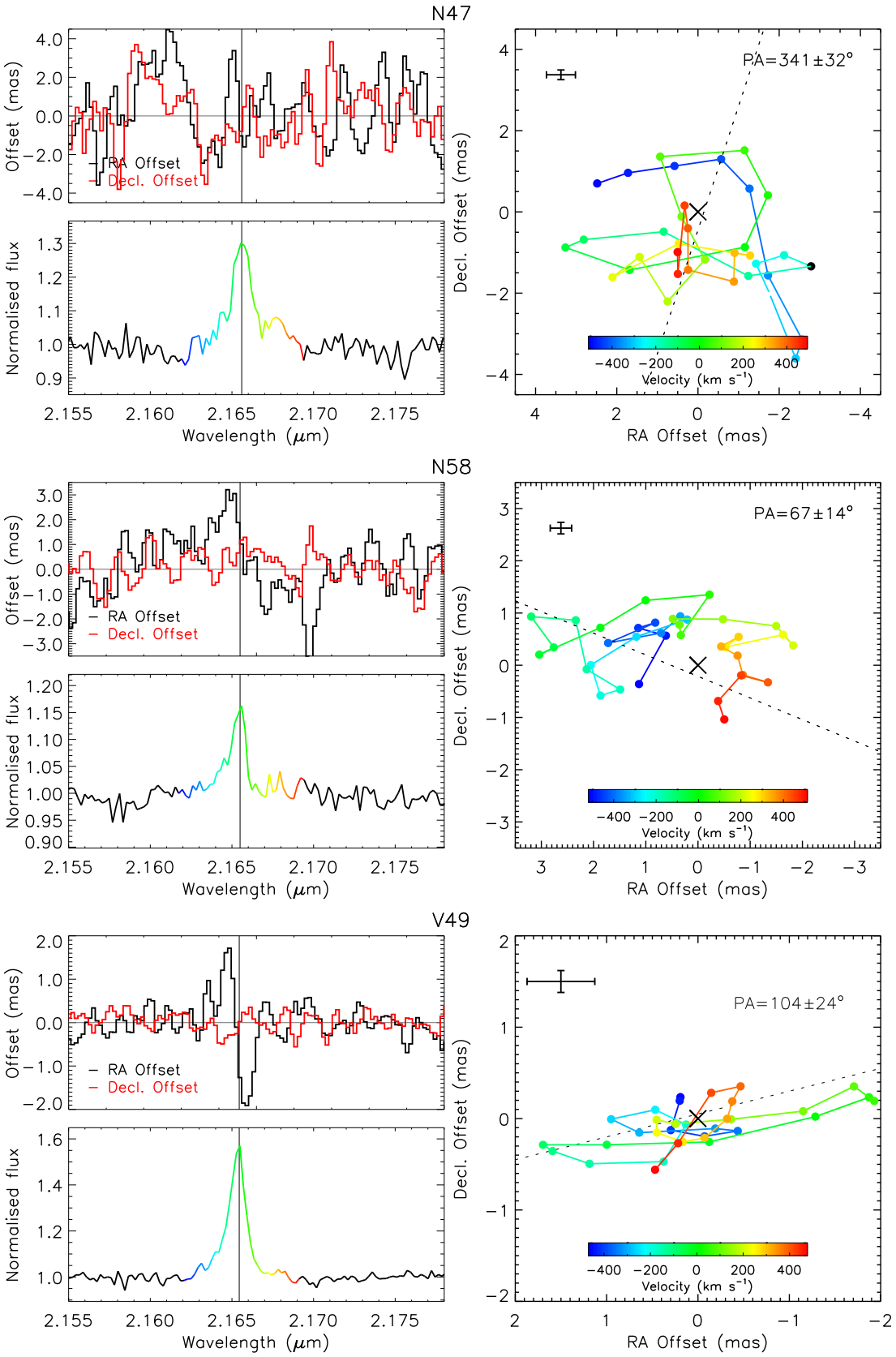


Figure 4.26: Spectro-astrometry of the Brγ feature for sources N47 (top), N58 (middle) and V49 (bottom). The full description of the plots is presented in Fig. 4.24.

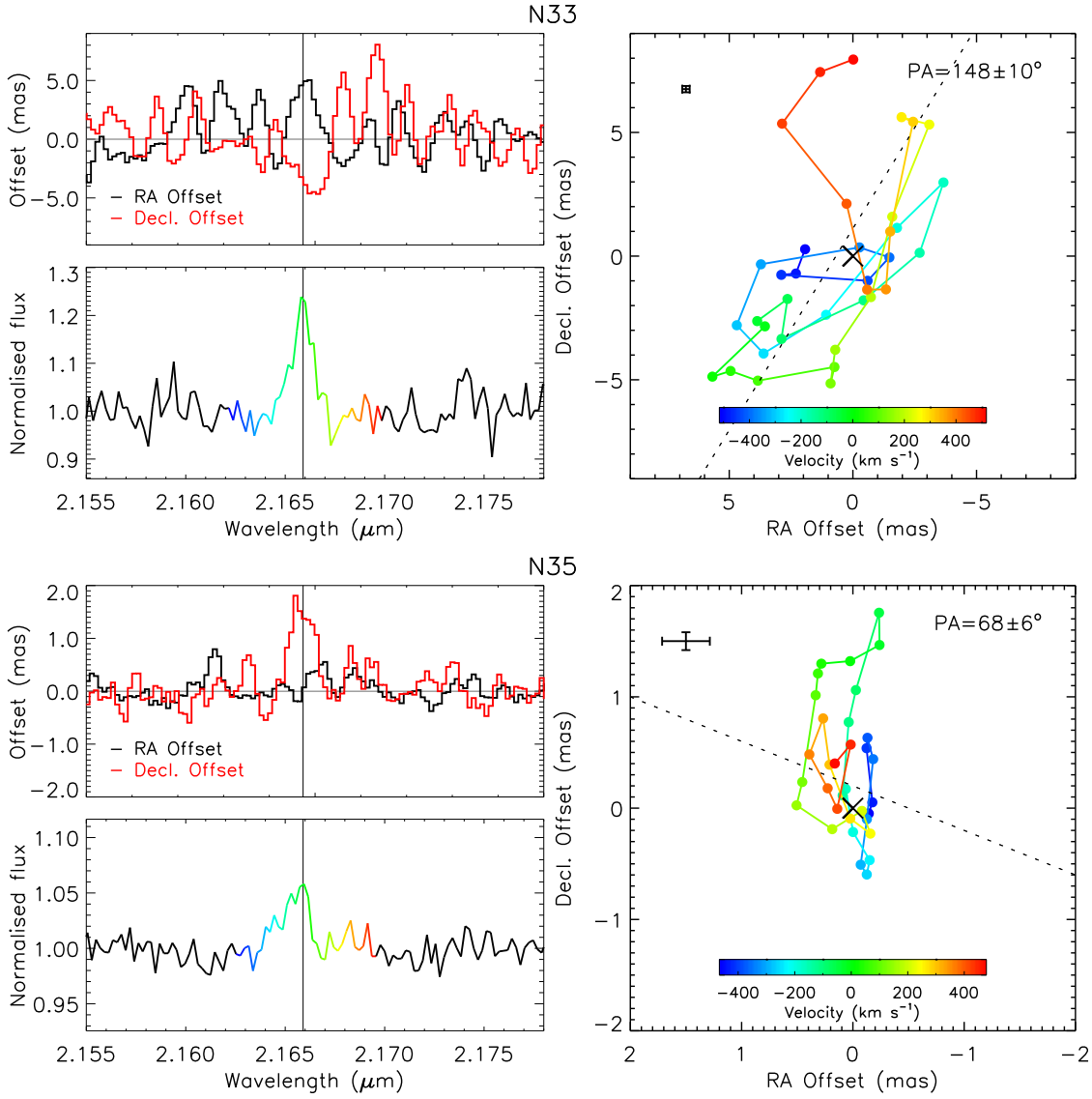


Figure 4.26: Continued. Sources N33 (top) and N35 (bottom).

velocity range from −500 to +500 km s^{-1}. The random patterns trace a region of ∼2 mas, corresponding to a linear scale of 8.6 AU around the main continuum position. This is consistent with the absence of discrete spectral features in the mean spectrum of this object (see Fig. 4.11) and, therefore, not much information can be extracted from the spectro-astrometry analysis. Finally, source N21 exhibits the largest positional offsets of the Brγ centroid of the sample (50 mas, equivalent to the NIFS pixel size). The blue- and red-shifted positions seem to lie in the same region, 25 mas to the east direction of the region probed at velocities near the systemic velocity. As presented on the continuum maps, this source is elongated on the SW-NE direction (see Figs. 4.8 and 4.9), with a position angle of 112.8±0.1°, the positional offsets of the Brγ emission has a PA = 296±21°, which is

consistent with the orientation of the continuum emission ($PA_{Br\gamma} = PA_{Cont} + 180°$). One can also speculate that multiple sub-structures at distinct velocities are present within the larger structure detected in the NIFS FOV.

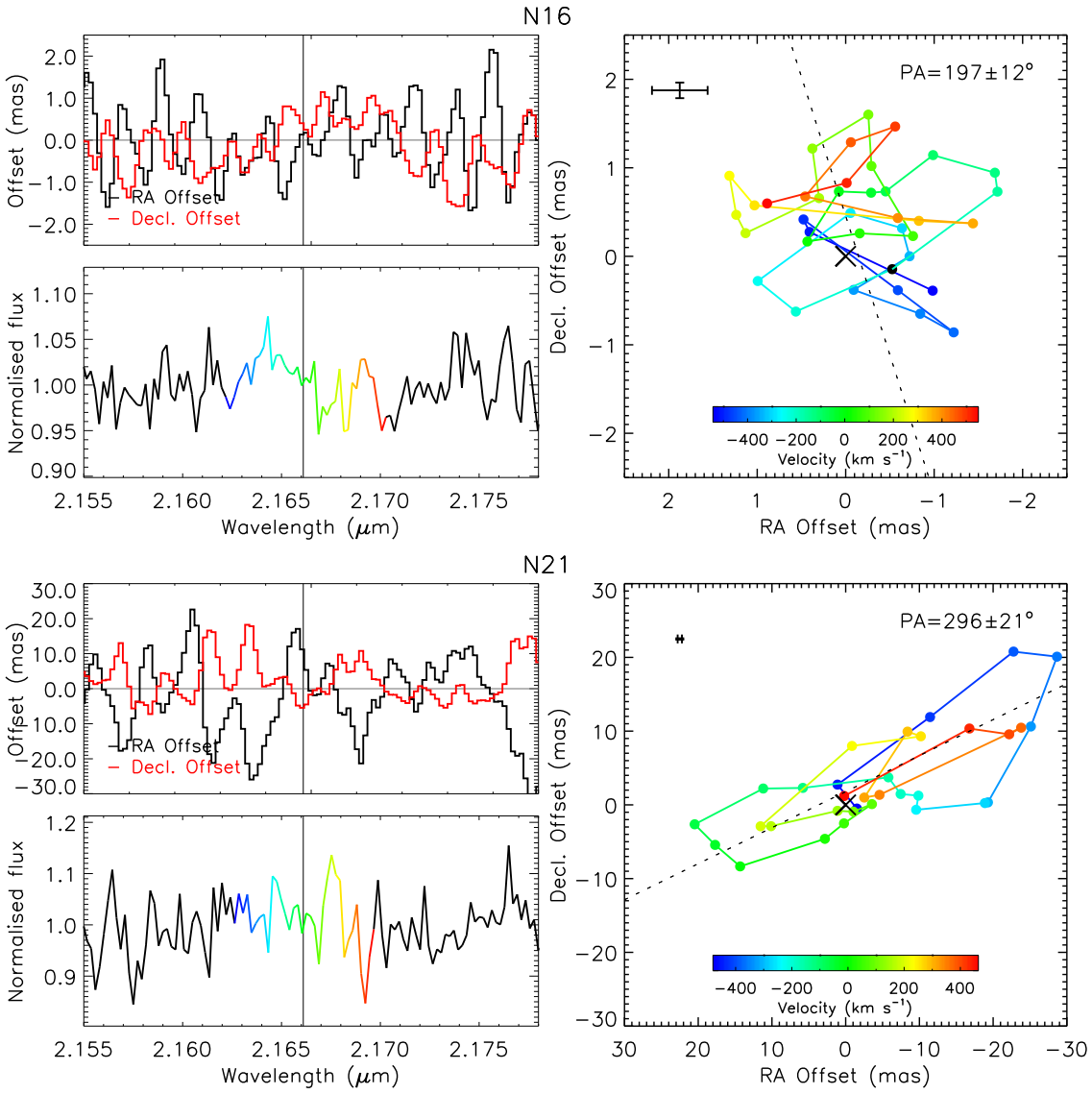


Figure 4.26: Continued. Sources N16 (top) and N21 (bottom).

4.3.6 *The CO features at 2.3 μm*

The analysis of the average spectra of the central sources (see Sect. 4.3.2) indicates that half of the sample (four of eight sources) exhibit CO bandhead emission features at ∼2.3 μm (N47, N58, V49 and W33A). In addition, the low-J transitions of the CO vibrational $v = 2 - 0$ state are observed in absorption towards the spectra of five sources (N33, N35, N47, V49 and W33A). Figure 4.27 presents a detailed view of the normalised spectra

Table 4.17 - Parameters derived from the spectro-astrometry of the Brγ feature towards the HMYSOs.

Source	Positional Accuracy (mas)	PA (°)	$\Delta\theta_{\rm blue-red}$ (mas)	$\Delta\ell_{\rm blue-red}$ (AU)	Origin
N21	[0.29,0.09]	296(21)	38.1	76.1	Multiple sub-structures (?)
N33	[0.15,0.13]	148(10)	7.4	10.3	Outflow cavity
N35	[0.21,0.08]	68(6)	0.45	1.7	Circumstellar disc
N47	[0.36,0.12]	341(32)	3.2	9.8	Outflow cavity
N58	[0.21,0.11]	67(14)	1.5	7.4	Circumstellar disc
V49	[0.36,0.12]	104(24)	0.64	2.0	Circumstellar disc (?)
W33A	[0.15,0.19]	152(28)	1.0	2.4	Outflow cavity

Notes: The columns are as follows: (1) ID of the source; (2) positional accuracy based on the telluric standards (see text); (3) position angle (from N to E) measured from the red to blue position; (4) angular distance between the red and blue extreme positions; (5) linear distance between the red and blue extreme positions; (6) most likely origin of the Brγ feature.

from Fig. 4.12 in the spectral region between 2.287 and 2.380 μm. Three CO bandhead emission features – CO (2–0), CO (3–1) and CO (4–2) – are observed together with the low-J CO absorption features associated with the HMYSOs. The identification of both CO bandhead emission and CO absorption features indicate that both warm (>1000 K) and cold ($\lesssim 100$ K) gas are present along the line-of-sight to the sources W33A, N47 and V49. In this section, the analysis of the CO emission and absorption features are discussed in terms of their origin and the physical parameters of the disc and the central sources.

4.3.6.1 The CO bandhead emission

The CO bandhead emission features were first identified in the spectra of YSOs by Carr (1989), who proposed two scenarios for explaining their origin: a circumstellar disc or a neutral stellar wind. The author found a good agreement between the observed CO profiles with the prediction obtained by modelling a dense ($\geq 10^{10}$ cm^{-3}) and warm (~ 3000 K) gaseous Keplerian disc. In the past years, some works have estimated proto-stellar masses by modelling the observed CO bandhead emission features by assuming the Keplerian disc hypothesis (*e.g.*, Kraus et al., 2000; Bik and Thi, 2004; Blum et al., 2004; Wheelwright et al., 2010, see Table 1.1). Indeed, the ultimate goal of studying the disc structure associated with the protostar is to assess the mass of the central source.

Assuming that the warm component of the disc shows Keplerian rotation at radii between 0.1 and 1.0 AU, the disc around a 10 $M_\odot$-like star would exhibit radial velocities in the range of $\sim$100-300 km s^{-1}, which can be analysed with the resolving power of the NIFS data for edge-on sources. The linear scales of the innermost region of the discs

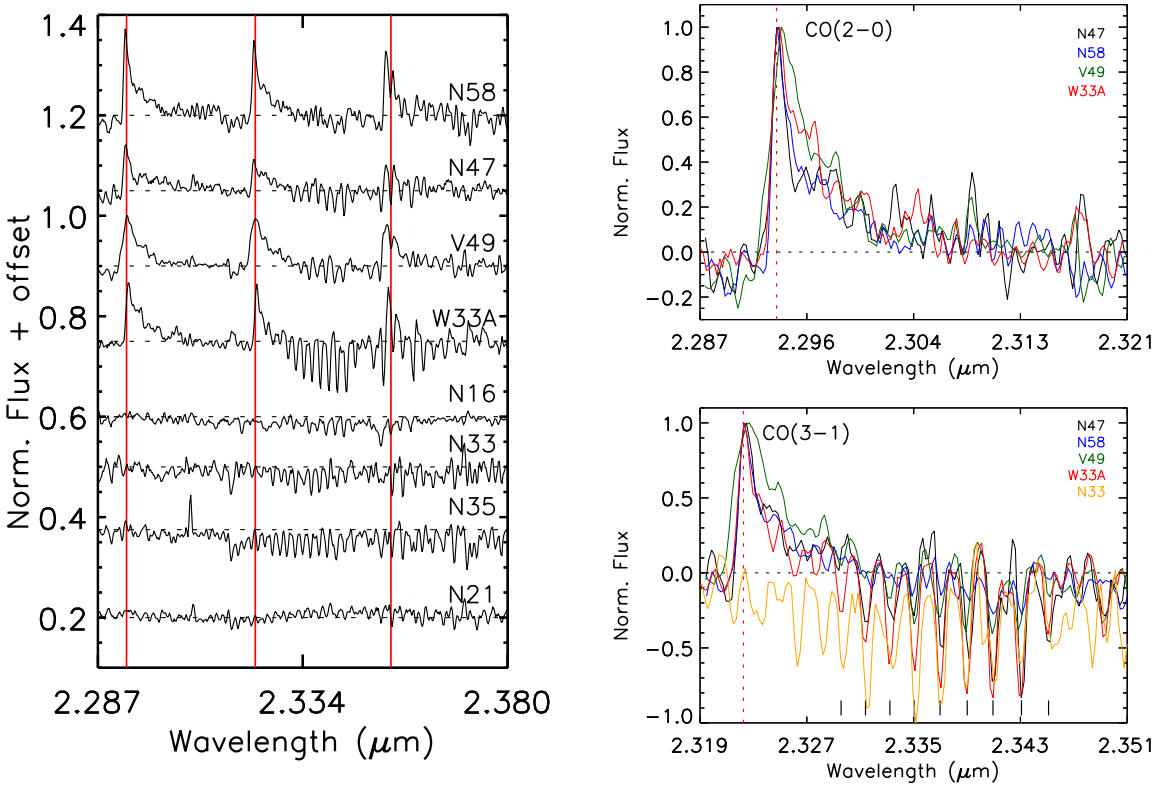


Figure 4.27: Left: Details of the CO features on the normalised spectra presented in Fig. 4.12. The rest wavelength of the CO (2–0), CO (3–1) and CO (4–2) bandhead features is indicated by the red vertical lines. The baseline of each spectrum is indicated by the horizontal black dashed lines. Right: Normalised emission around the CO (2–0) (top) and the CO (3–1) (bottom) bandhead emission features for sources N47, N58, V49 and W33A (source N33 is also indicated in the bottom panel). The data was velocity-corrected using the $V_{\rm lsr}$ values from Table 4.1. The baseline is indicated by the horizontal black dashed line. The rest wavelength of the CO bandhead feature is indicated as the red dashed line, and the low-J CO absorption features are indicated by the black filled lines.

are unresolved and, therefore, no kinematic information can be inferred from the CO emission features. However, the observed CO profile can be modelled in terms of the general physical parameters and the geometry of the discs (*e.g.*, Blum et al., 2004; Bik and Thi, 2004; Ilee et al., 2013). Bik and Thi (2004) found good correspondence between the observed and modelled CO bandhead emission features for CO surface densities in the range 10^{20}–10^{21} cm^{-2} and excitation temperatures between 1700 and 4500 K.

Sources N58 and W33A exhibit spatially extended CO emission (see integrated maps presented in Fig. 4.8). One-dimensional spectra were extracted at different positions along a pseudo-slit, oriented in the direction of the extended CO emission to investigate any variation of the CO bandhead profile along the spatial distribution of the emission. The results for source N58 are presented in Fig. 4.28. Spectra were extracted in steps of 0.″2 along the direction of the extended CO emission, as indicated by the grey arrow placed at

PA = 45°. The spectra were continuum-normalised and were over-plotted, as presented in the left panel of the figure. The data indicate that the CO (2–0) overtone is broader closer to the central position and gets narrower towards the extended CO emission. Such broadening of the CO bandhead overtone was already observed by D10 in the source W33A, when they compared the profile of the CO emission features in the nebular spectrum with the one detected in the spectrum of the central source (see Fig. 11 from D10). Those authors interpreted the broadening of the CO bandhead emission as caused by a rotating disc associated with the main source. Therefore, the broadening of the CO bandhead observed towards source N58 suggests the presence of a rotating disc as similarly as observed towards W33A. For the other sources that do not display extended CO emission (N47 and V49), the spectra of the central sources were analysed by assuming that the CO emission arises from a circumstellar disc as well.

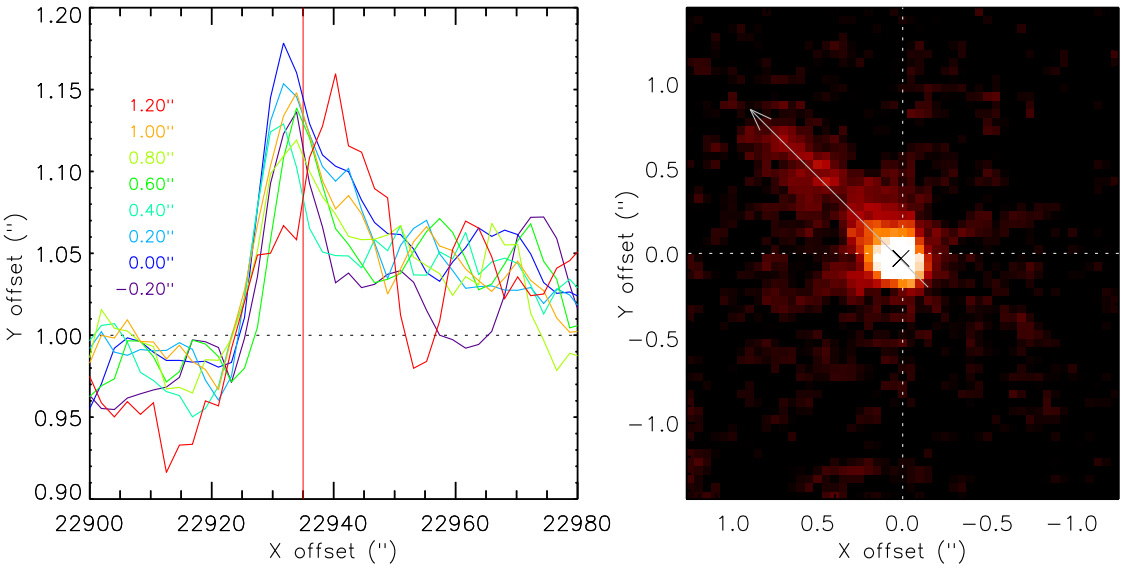


Figure 4.28: Spatial variation of the CO (2–0) bandhead overtone towards the extended CO emission associated with source N58. The map shown on the right presents the integrated CO emission at 2.29 μm and the red arrow indicates the direction of the extended CO emission. The central position (0.″0,0.″0) is indicated by the black × symbol.

As discussed above, the top right panel of Fig. 4.27 indicates that the CO bandhead emission features exhibit variations in the shapes and broadening of their profiles. The shape of the CO emission observed at the position of the main source is mostly affected by inclination effects, the density and temperature of the gas. The physical parameters and inclination of the system "disc+HMYSO" can be constrained by modelling the observed CO profile using the methodology from Kraus et al. 2000 and, later, adapted by Blum

et al. 2004.

Based on the methodology presented by Blum et al. (2004), a set of `IDL` routines were written to model a synthetic CO spectrum based on the given physical and geometric parameters (constrained by six free-parameters), and to compare it with the observations. The free-parameters of the model are: the surface density at the inner radius for emission ($N(r_{\rm in})$); the power-law indices $\beta_{\rm T}$ and $\beta_{\rm N}$ for the temperature and density profiles, respectively; the radial velocity at the inner radii ($v(r_{\rm in})$); the inclination of the system (i); and the systemic velocity of the system ($V_{\rm lsr}$). The modelled CO spectrum uses power-laws for the disc temperature and the column density distribution, given by:

$$T(r) = T(r_{\rm in}) \left(\frac{r}{r_{\rm in}} \right)^{-\beta_{\rm T}} \tag{4.14}$$

$$N(r) = N(r_{\rm in}) \left(\frac{r}{r_{\rm in}} \right)^{-\beta_{\rm N}} \tag{4.15}$$

where $T(r_{\rm in})$ corresponds to the temperature at the inner radius ($r_{\rm in}$).The inner radius of the disc ($r_{\rm in}$) is defined in terms of the radial velocity at that position and the mass of the central source as:

$$r_{\rm in} = \frac{GM}{v(r_{\rm in})^2} \tag{4.16}$$

It worth mentioning that the method itself does not provide a direct estimate of the mass of the source. Instead, it requires an estimate of the mass to derive the physical extension of the disc.

The optical depth and CO flux are calculated in a radial plus azimuthal coordinate grid (r,θ). The total CO flux is obtained by adding the emission from each (r,θ) cell, corrected by its radial velocity ($v(r)$) and inclination (i). The line flux is calculated assuming LTE, and adding the individual contribution of rotational CO transitions with J between $J_{\rm min}$ and $J_{\rm max}$. The synthetic spectrum is, then, convolved using a Gaussian or Lorentzian line profile. An intrinsic Lorentzian line profile was chosen by following the same assumptions made by Blum et al. (2004). The parametrisation of the disc geometry is fixed for every model. The number of radial elements is equal to $n_{\rm r} = 75$, as well as the number azimuthal elements (n_θ).

The maximum temperature at the inner radius of the CO disc is fixed at 5000 K, above which CO should be dissociated. The maximum surface density is a free-parameter which

depends on the observed CO profile (larger N values broaden the red wing of the CO bandhead profile; *e.g.*, see Fig. 12 from Kraus et al. 2000). The outer radius of the disc ($r_{\rm out}$) is defined at the radial location where the temperature drops to 1000 K. The ratio between the outer and inner radius of the disc ($\beta_{\rm r}$) is defined in terms of the ratio between the temperatures at these radii as:

$$\beta_{\rm r} = \left(\frac{T_{\rm min}}{T_{\rm max}}\right)^{-\frac{1}{\beta_{\rm T}}} \tag{4.17}$$

The radial element (r_i) is parametrised as:

$$r_i = r_{\rm in} + \beta_{\rm r}\left(\frac{i-1}{n_{\rm r}-1}\right) \tag{4.18}$$

where $\beta_{\rm r}$ corresponds to the ratio between the outer and inner disc radii, and i varies between 1 and $n_{\rm r}$. The azimuthal element, θ_i varies between 0 and 2π, in steps of $2\pi n_\theta^{-1}$. For each cell, the CO flux is calculated in terms of the temperature and density at the given position.

The total CO flux ($F_{\rm CO}$) is obtained as the sum of the flux evaluated at each cell, $F_{\rm CO}(r,\theta)$, corrected by the inclination angle of the system and considering a velocity-correction based on the radial velocity of the individual cells. The total flux is given by:

$$F_{\rm CO}(r,\theta) = \sum_{r,\theta}\left[F_{\rm CO}(T,N,r,\theta)\cos(i)\cos(\theta)\left(\frac{v_r}{c}\frac{\lambda_c}{\Delta\lambda}\right)\right] r\ dr\ d\theta \tag{4.19}$$

Finally, the modelled spectrum is continuum-normalised and scaled by the peak intensity of the observed spectrum. The χ^2 and $\chi^2_{\rm red}$ parameters are computed considering only the first CO bandhead feature, between 2.291 and 2.296 μm.

The procedure described above was tested on source W33A, using the same physical and geometrical parameters adopted by D10: the total mass of the system, $M = 15\,{\rm M}_\odot$, the inclination of the system, $i = 60°$, and the radial velocity at the inner radius of the disc, $v(r_{\rm in}) = 50\,{\rm km\,s^{-1}}$. The results obtained with the methodology presented in this section are shown in Fig. 4.29. The best fit was obtained by assuming the surface density at the inner disc, $N(r_{\rm in}) = 10^{19}\,{\rm cm^{-2}}$, and the exponents $\beta_{\rm N} = 4.5$, and $\beta_{\rm T} = 0.5$.

Figure 4.30 presents the results of the CO bandhead emission fitting for sources N47 and N58. The CO emission modelling for source V49 is presented in the left panel of

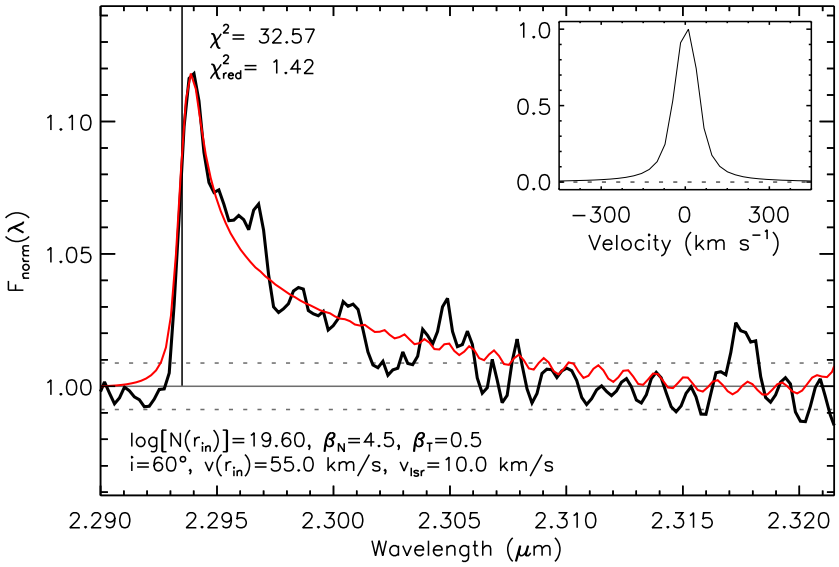


Figure 4.29: Normalised CO (2–0) bandhead emission for source W33A. The observed and synthetic spectra are indicated by the black and red curves, respectively. The parameters adopted for computing the synthetic spectrum are presented in the bottom of the plot. Both χ^2 and reduced χ^2, evaluated in the spectral region between 2.291 and 2.296 μm, are presented in the top left region of the panel. The vertical filled line is placed at the rest-wavelength of the CO bandhead (2.2935 μm), the horizontal filled line indicates the baseline and the dashed lines correspond to the $\pm$1-σ limit. The inset presents a single CO line profile as a function of the velocity, normalised in respect to the peak intensity.

Fig. 4.31. The results are summarised in Table 4.18.

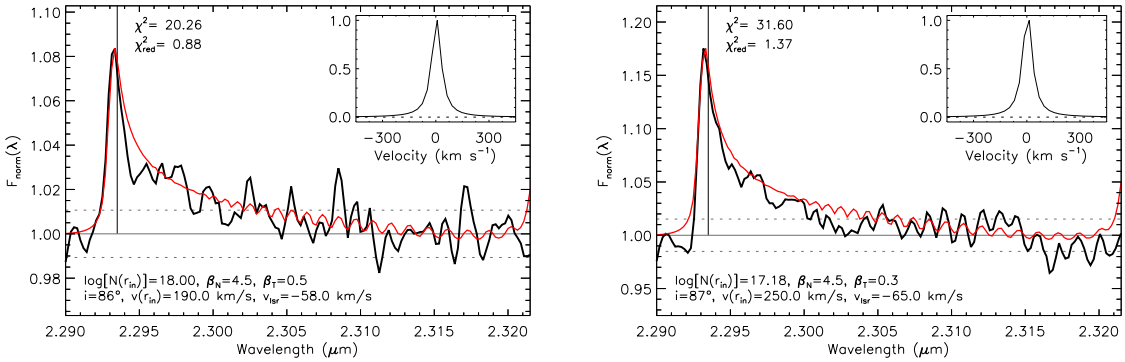


Figure 4.30: Same as Fig. 4.29 but displaying the modelling of the CO bandhead emission in the normalised spectra of sources N47 (left) and N58 (right).

Among the four sources, V49 is the only source that exhibits a pronounced blue shoulder and red-shifted peak. In according to Blum et al. (2004), such profile is well fitted by a rotating disc. In the observed spectrum, the velocity separation between the blue shoulder and the red peak is roughly twice the projected rotational velocity at the outer radius of the disc, while the terminal velocity of the blue shoulder corresponds to the projected velocity of the CO at the inner radius of the disc. Figure 4.31 presents the CO (2–0) bandhead

Table 4.18 - Parameters derived from the fitting of the CO bandhead emission features in the spectra of the HMYSOs.

Source	i (°)	$\log[N_{\rm CO}(r_{\rm in})]$ ($\rm cm^{-2}$)	$\beta_{\rm N}$	$\beta_{\rm T}$	$v(r_{\rm in})$ ($\rm km\,s^{-1}$)	$V_{\rm lsr}$ ($\rm km\,s^{-1}$)	M_* ($\rm M_\odot$)	$\log(M_{\rm disc}/M_\odot)$
N47	86	18.00	4.5	0.5	190.0	−58.0	29-59	–7.89,–7.28
N58	87	17.18	2.5	0.5	50.0	−75.0	12-23	–6.71,–6.14
V49	65	20.48	4.5	0.5	270.0	−95.0	8-80	–7.17,–5.17
W33A	60	19.60	4.5	0.5	55.0	+15.0	4-8	–5.88,–5.28

Notes: The columns are as follows: (1) ID of the source; (2) the inclination of the "disc+HMYSO" system (90° corresponds to edge-on view); (3) logarithm of the surface density of the CO at the inner radius of the disc (in $\rm cm^{-2}$); (4) the power-law index of the radial density distribution; (5) the power-law index of the radial temperature distribution; (6) the radial velocity at the inner radius of the disc (in $\rm km\,s^{-1}$); (7) the systemic velocity of the source that best fits the CO emission (in $\rm km\,s^{-1}$); (8) the range of mass of the central source used to derive the physical parameters of the disc (in $\rm M_\odot$); (9) the range of disc masses (in $\rm M_\odot$).

feature of source V49 in the velocity scale. For source V49, the velocity of the outer and inner radius of the disc were estimated in 95 and 175 $\rm km\,s^{-1}$, respectively.

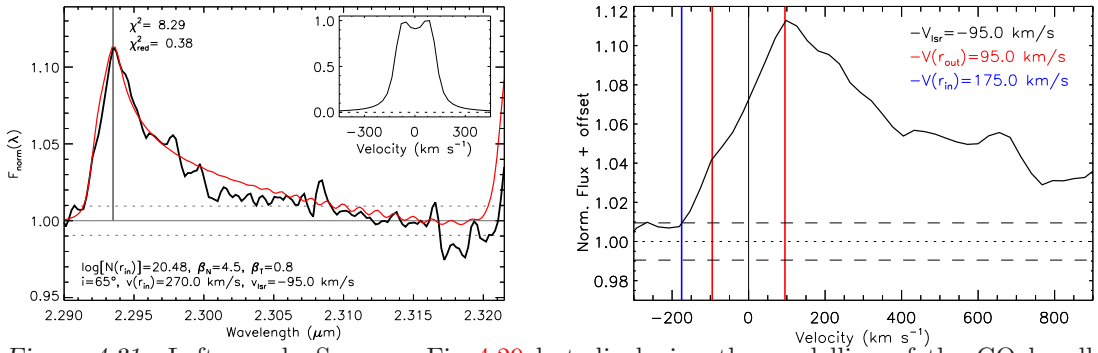


Figure 4.31: Left panel: Same as Fig. 4.29 but displaying the modelling of the CO bandhead emission in the normalised spectra of source V49. Right panel: detailed view of the CO (2–0) bandhead emission feature on the velocity scale, exhibiting a shoulder feature on the blue side of the CO feature. A smooth factor of 2 was applied to the spectrum. The rest velocity is placed at 0 $\rm km\,s^{-1}$, indicated by the black vertical line. The terminal velocity at the blue side defines the velocity of the outer disc (indicated by the blue vertical line) while the velocity of the blue shoulder and the peak of the CO feature define the velocity of the inner disc (indicated by the red vertical lines).

The CO emission arises from the innermost area of the disc, and the mass of that specific region can be evaluated using the results from the CO modelling. The mass of the disc is obtained by integrating the surface density over its extension as presented as follows. The number of molecules of CO present in the disc can be evaluated as:

$$n_{\rm CO} = \sum_{r=r_{\rm in}}^{r_{\rm out}} 2\pi r \Delta r N_{\rm CO}(r) \tag{4.20}$$

Using Eq. (4.15),

$$n_{\rm CO} = 2\pi \frac{\Delta r N_{\rm CO}(r_{\rm in})}{r_{\rm in}^{-\beta_{\rm N}}} \sum_{r=r_{\rm in}}^{r_{\rm out}} r^{-\beta_{\rm N}+1} \tag{4.21}$$

The total mass of the disc, in grams, is obtained as similarly as the mass of the H_2 jets (Eq. (4.11)):

$$M_{\rm disc} = \mu_{\rm H2} m_{\rm H} [H_2/CO] n_{\rm CO} \tag{4.22}$$

where $[H_2/CO] = 1.2{\cdot}10^4$ is the typical abundance of the H_2 in terms of the CO molecules (Yıldız et al., 2012), $\mu_{H2} = 2.8$ is the mean particle weight per hydrogen molecule (Kauffmann et al., 2008), and m_H is the mass of the Hydrogen atom. Due to similarity between the physical conditions of the hot and dense gas phase of the circumstellar discs with those from dense and warm molecular outflows, the relatively larger $[H_2/CO]$ ratio from Yıldız et al. (2012) was adopted instead of the typical ratio found towards cold, dark clouds ($(6.4{\pm}0.5){\cdot}10^3$, Kulesa, 2002).

As indicated by Eq. (4.16), the physical extension of the disc (that is, the limits $r_{\rm in}$ and $r_{\rm out}$) depends on the mass of the central source. Therefore, the mass of the CO disc is proportional to the mass of the central source. Another point that must be taken into account is that the derived CO mass corresponds to a small fraction of the total mass of the disc. In fact, the CO bandhead emission originates in the inner region of the disc, between 1-2 AU, while the total extension of a disc associated with a HMYSO is expected to be 10^2-10^3 AU.

4.3.7 Correlation between the mass of the discs and the jets

Previous works have successfully modelled the CO bandhead emission spectra of YSOs and other astrophysical objects (*e.g.*, Ae/Be stars). A summary of the density profile and physical extension of the discs from the literature is presented in Table 4.19, as well as the estimate of the disc masses, evaluated using Eq. (4.22).

The disc masses are presented as a function of the bolometric luminosity of their driven sources in Fig. 4.32. The plot indicates a large scatter of the data and a relatively weak correlation between the mass of the disc and the bolometric luminosity of the sources ($\rho = 0.41$, $p = 0.05$). The best fit to the data indicates that $M_{\rm disc}$ scales with $L_{\rm bol}$ by following a power-law distribution with index $0.71{\pm}0.30$.

Table 4.19 - CO bandhead modelling results for other YSOs from the literature.

Source	$\log(L_{\rm bol}/L_{\odot})$	$r_{\rm in}$ (AU)	$r_{\rm out}$ (AU)	$\log[N_{\rm CO}(r_{\rm in})]$	$\beta_{\rm N}$	$\log(M_{\rm disc}/M_{\odot})$	Ref.
IRAS 16164	5.11±0.12	–	–	–	–	–8.70±0.17	Bik04
IRAS 08576	2.86±0.12	–	–	–	–	–6.08±0.17	Bik04
NGC 3576	4.00±0.44	24.08	50.56	22.13	1.0	–0.46±0.17	B04
M17-268	3.25±0.35	0.13	0.77	22.35	3.8	–5.73±0.17	B04
M17-275	3.82±0.42	1.10	7.61	21.54	4.1	–4.44±0.17	B04
G333	2.27±0.25	0.38	2.47	21.11	3.6	–5.33±0.17	B04
IRAS08576	2.55±0.28	0.09	0.78	21.90	1.5	–5.00±0.17	W10
M8E	3.67±0.40	0.31	2.61	23.04	1.5	–2.79±0.17	W10
G033	4.11±0.12	0.24	2.05	21.97	1.5	–4.08±0.17	W10
VG308	5.59±0.12	0.94	8.00	22.69	1.5	–2.17±0.17	W10
G332	5.11±0.12	0.59	5.08	21.34	1.5	–3.92±0.17	W10
G347	4.63±0.12	0.45	3.81	22.23	1.5	–3.27±0.17	W10
IRAS 18151	4.34±0.12	2.40	28.60	21.40	1.5	–2.54±0.17	M13
S106	4.60±0.12	0.30	4.00	22.70	1.5	–3.01±0.17	M13
V645 Cyg	4.63±0.12	3.20	44.00	20.00	1.5	–3.65±0.17	M13
HD 369	2.20±0.12	0.10	1.50	20.78	1.8	–6.05±0.17	I14
HD 259	3.19±0.12	0.89	4.30	19.20	3.3	–6.45±0.17	I14
HD 58	2.95±0.12	0.18	0.62	19.79	1.6	–7.20±0.17	I14
HD 101	1.58±0.12	1.00	1.10	22.15	0.5	–6.13±0.17	I14
PDS 37	3.27±0.12	1.60	3.80	19.00	1.7	–5.99±0.17	I14

Notes: The columns are as follows: (1) ID of the source; (2) the bolometric luminosity of the source; (3)-(4) the inner and outer radius of the disc; (5) the surface density of the CO at the inner radius of the disc; (6) the power-law index of the radial density distribution; (7) the CO disc mass derived from Eq. (4.22); (8) References: Bik04 – Bik and Thi (2004); B04 – Blum et al. (2004); W10 – Wheelwright et al. (2010); M13 – Murakawa et al. (2013); I14 – Ilee et al. (2014).

4.3.8 Kinematics of the CO absorption features: outer disc signatures?

An ultimate goal of the investigation of the star formation process is to identify a connection between the inflow and outflow processes during the active accretion phase of the protostars. One might expect that the available mass in the discs and the expelled mass through the molecular jets and outflows are correlated. The CO modelling only provides a small fraction of the total mass of the discs, responsible for producing the observed CO emission. Therefore, a tracer of the outer and colder regions of the circumstellar environment of the HMYSOs is required to estimate the total mass of the disc. Davies et al. (2010) and Murakawa et al. (2013) presented the analysis of the kinematics of the low-J CO absorption features at $\sim 2.33\,\mu$m, and their spatial distribution for W33A and two other objects (I18151 and AFGL 2136). According to those authors, the absorption CO features probes an outer and colder region ($\gtrsim 10^2$ AU, with temperatures of $\sim$30 K)

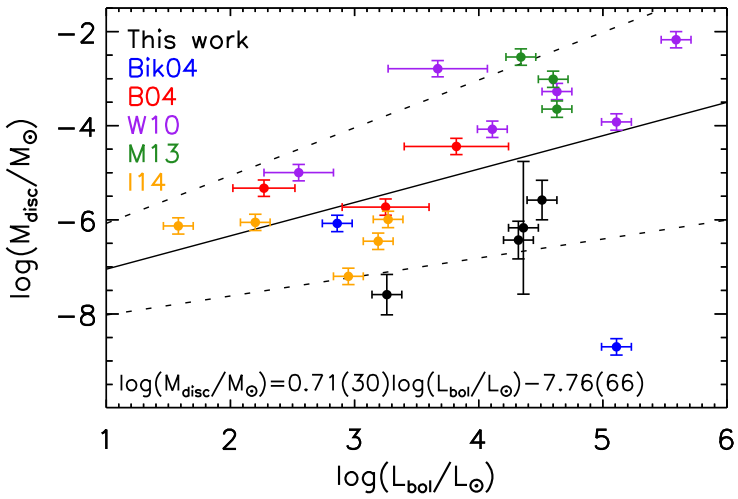


Figure 4.32: The mass of the discs as a function of the bolometric luminosity of HMYSOs. Data from other works in the literature are indicated using the following colour-scheme: Bik04 (blue) – Bik and Thi (2004); B04 (red) – Blum et al. (2004); W10 (purple) – Wheelwright et al. (2010); M13 (green) – Murakawa et al. (2013); I14 (yellow) – Ilee et al. (2014). The best fit to the data is indicated by the filled black line. The dashed black lines are placed at the 1-σ level of the fit.

of the circumstellar discs. The mass of the 'disc+protostar' system can be determined by fitting a Keplerian rotation curve on the position-velocity diagram of these transitions (see Fig. 17 from D10).

By following their procedure, IDL routines were written to resample the input datacubes into a ten times finer spectral resolution data and to derive the kinematics of the CO absorption features. A section of the datacubes was extracted in the spectral direction (λ-axis) between 2.32 and 2.38 μm. These data were resampled in the λ-axis by a factor of ten (that is, the pixel scale changed from $\sim$2.13 Å pixel^{-1} to $\sim$0.213 Å pixel^{-1}) using a linear interpolation algorithm[12]. As an example, the original and the resampled spectra of source W33A in the region of the CO absorption features are presented in Fig. 4.33.

Then, two approaches were followed to analyse these data: *a*) differential images between the blue- and red-shifted emission were obtained to check for disc-like candidates traced by the CO absorption features; and *b*) position-velocity diagram was obtained for each source. These results are discussed as follows.

[12] see the IDL CONGRID() routine.

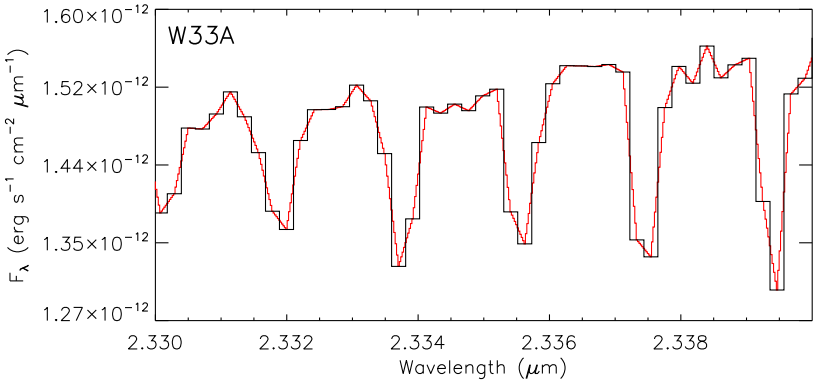


Figure 4.33: Zoom into the low-J CO absorption features identified in the spectrum of W33A at the original spectral pixel-scale (in black) and the resampled spectrum into a 10 times smaller spectral dispersion (in red).

4.3.8.1 *Differential CO maps*

A final datacube was created by adding up each CO feature together to improve the signal-to-noise ratio of the data. Then, two integrated intensity maps were created by summing the velocity channels (*i.e.,* the spectral channels) corresponding to velocities bluer than the central velocity ($X_{\rm blue}(x,y)$) and those redder than the central velocity ($X_{\rm red}(x,y)$). The subtraction between the blue- and the red-shifted maps delivered the residual maps presented in Fig. 4.34.

All maps indicate sufficient contrast between the blue- and the red-shifted emission as expected from a rotating structure. In the case of W33A, the position angle of the structure, PA = 130° is roughly consistent with the orientation derived by D10 (PA = 150±10°). By following the convention adopted for W33A, the position angles of the discs were measured from North to East direction, with the red-shifted emission oriented to the left and the blue-shifted emission to the right of the disc axis.

For source N35, the orientation axis of the disc (PA = −90°) is perpendicularly aligned with the propagation of the H_2 jet (PA = 3°; see Fig. 4.14 and Table 4.12). As opposed to W33A, the disc associated with the source N47 is counter-aligned with the brightest H_2 jet (with position angles estimated in 160° and 12°, respectively). V49 has the less obvious presence of a disc, from which the axis of the disc candidate is offset by 0.″1−0.″2 to the W direction of the peak of the continuum emission. It worth mentioning that, as previously found in the H_2 map presented in Fig. 4.14, the source at the centre of the FOV is not the driving source of the H_2 jet that propagates towards the NW direction in the map.

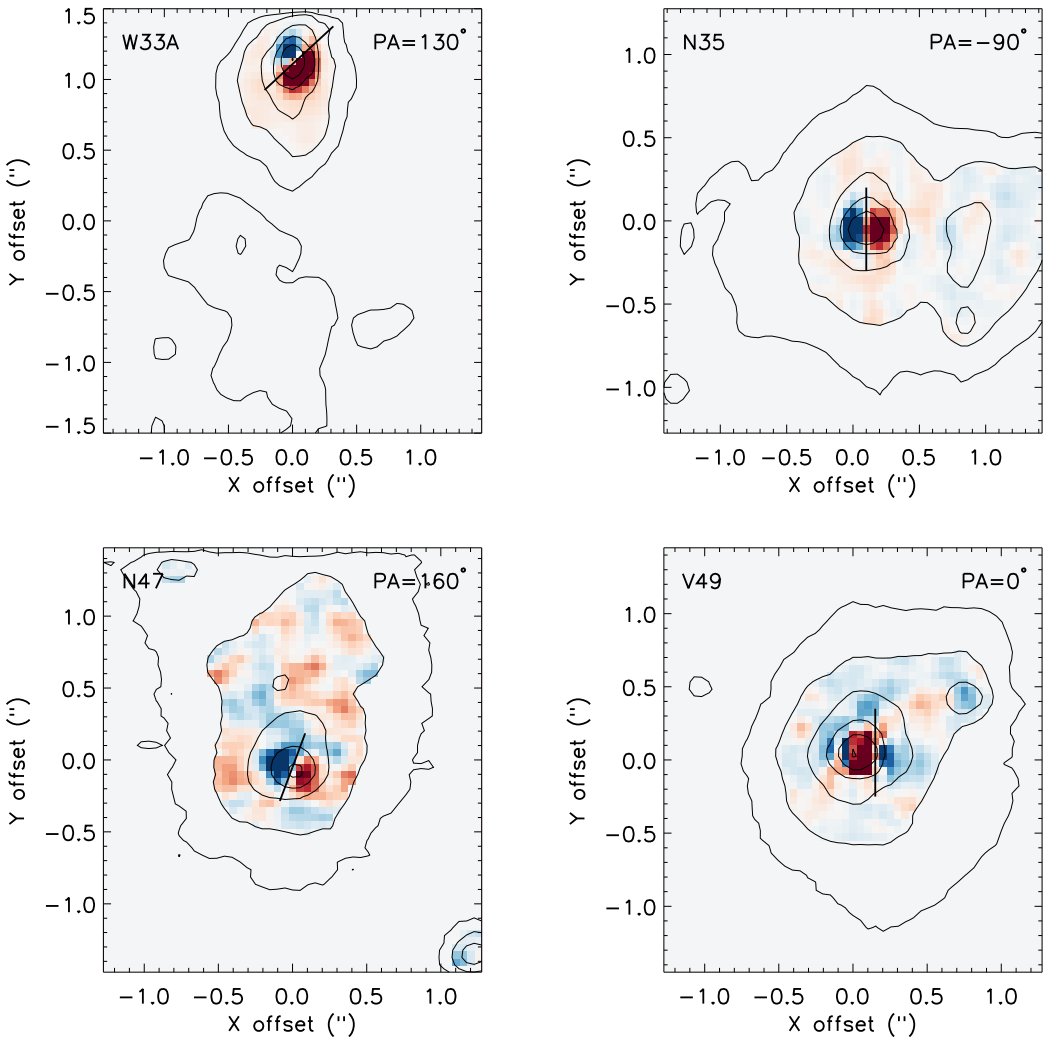


Figure 4.34: Residual images between the blue-shifted and red-shifted components of the CO absorption features. The strength of the CO feature in the blue- and in the red-shifted components is indicated by following the blue-to-red colour scale. The black contours are placed at the $3^n - \sigma$ ($n = 0, 1, 2, ...$) level of the integrated emission over the entire spectral range of the CO datacubes. The filled line indicates the polar axis of the disc candidates, measured from N to E (see the convention in the text).

4.3.8.2 Kinematics of the CO absorption features

The analysis of the CO absorption features kinematics was further investigated by fitting the velocity offset of each spaxel in the datacube. This procedure was performed using the *Penalized Pixel-Fitting* (pPXF) algorithm from Cappellari and Emsellem (2004). The pPXF models an input spectrum based on a given template, and, among a series of parameters (see Cappellari and Emsellem (2004) for details), delivers the radial velocity of the observed data. For this work, the brightest spaxel of the continuum emission at $\sim$2.3 μm (generally at the central region of the point-source object) was used as the tem-

plate spectrum (at $0\,\mathrm{km\,s^{-1}}$). At the end of the modelling, a two-dimensional map of radial velocities is delivered for each object. The results for W33A is presented in the left panel of Fig. 4.35, together with the map obtained by D10 (see right panel).

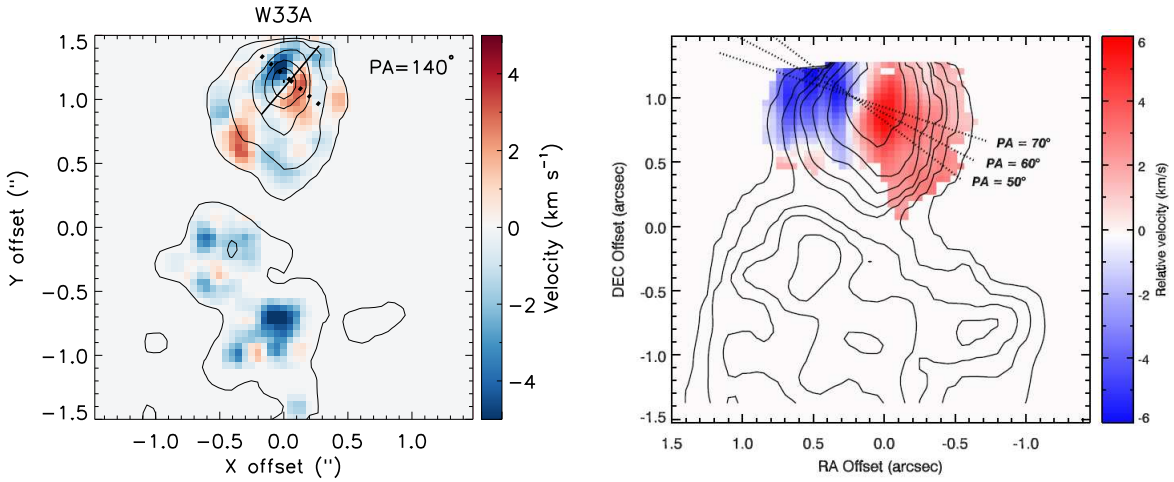


Figure 4.35: Left panel: Radial velocity map of the CO absorption features towards W33A, evaluated from the pPXF analysis. The black contours are placed at the $3^n - \sigma$ level ($n = 0, 1, 2, ...$) of the integrated continuum emission at $\sim$2.3 μm. The velocities were estimated for spaxels with fluxes above 2-σ. The velocity scale is presented on the right side of the map. The filled black line indicates the position angle (from N to E) of the rotation axis of the disc. The dashed black line is perpendicularly aligned to the disc axis. Right panel: similar velocity map for the same source, obtained by Davies et al. (2010).

The rotation direction and position angle of the structure obtained in this work is consistent with the parameters derived in D10. However, some differences in the methodology adopted in each work and their results should be discussed. First, in according to D10, a datacube similar to the obtained in Sect. 4.3.8.1 was smoothed with a boxcar filter of 5-spaxel width (0.″25) before evaluating the radial velocity map presented in the right panel of Fig. 4.35. As a consequence, the kinematic structure probed by their map extends towards larger offsets from the central source position than the corresponding map presented in the left panel of Fig. 4.35. On the contrary, no smoothing was applied to the input data used for obtaining the radial velocity map presented in the left panel of Fig. 4.35. As a consequence, the spatial distribution of the kinematics of the CO features is less extended than the one obtained by D10 and presented in the right panel of Fig. 4.35.

Although the orientation of the disc candidate around W33A is well defined on the residual map presented in Fig. 4.35, no kinematic information is available. Complementary maps obtained from the pPXF analysis were used to evaluate the position-velocity diagram (P-V diagram) across the disc-like structure associated with W33A, as presented

in Fig. 4.36. For comparison, the results obtained from D10 are presented in the same figure. The P-V diagram was constructed by extracting a slice of the velocity map along the dashed black line indicated in Fig. 4.35. Three models of Keplerian rotation curves are overlaid on the data points: the best fit to the data and the corresponding upper and lower limit considering 1-σ deviation. The best model corresponds to the one with the lowest χ^2 parameter, based on the points between $\pm 0\farcs2$ from the central position. Finally, the maximum radius of the disc is set by fixing the position where the radial velocity goes to zero at positional offsets far from the central position.

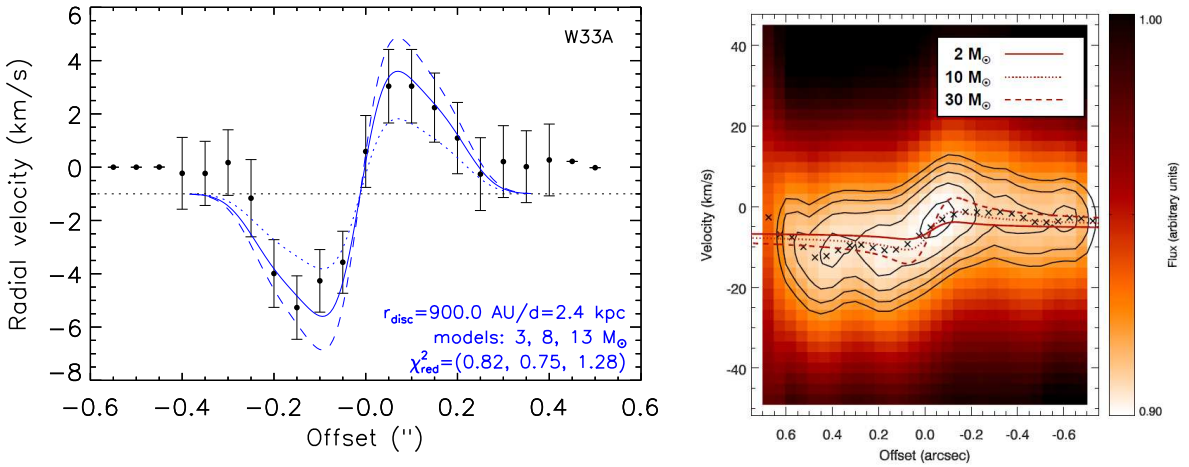


Figure 4.36: Left panel: Position-velocity diagram across the disc-like structures identified towards W33A. The data points were extracted across the dashed lines indicated in Fig. 4.35. The black dots and their errors were obtained from the pPXF analysis of the CO absorption features. Blue curves indicate the Keplerian rotation models for three different central masses: the best fit is shown as a filled line, the lower-and upper-limit as dotted and dashed lines, respectively. The horizontal dotted line indicates the velocity offset that better fit the models with the observed data. The truncation radii of the disc and distance of the sources are indicated in the bottom right corner of the plot. Right panel: The P-V diagram for W33A, obtained by Davies et al. (2010). The map is scaled in terms of the intensity of the CO absorption features below the continuum. The crosses indicate the velocity of the minimum intensity for the absorption as a function of position along the disc direction. The contours are placed at intensity levels of 0.895 to 0.92 at intervals of 0.005. The lines indicate three models of Keplerian rotation for different central masses at the distance of 3.8 kpc.

The P-V diagram for W33A indicates a clear separation between the blue- and red-shifted components (with peak velocities around $+4$ and $-4\,\mathrm{km\,s^{-1}}$, respectively) within positional offsets ranging between $-0\farcs2 \leq r \leq +0\farcs2$. The best Keplerian model, shown as the filled blue curve, indicates that the total mass of the system "protostar+disc" is about $8 \pm 5\,\mathrm{M_\odot}$. The uncertainty on the mass corresponds to the deviation between the mass assumed for the best fit and those used for the upper and lower limit models, shown as the dashed and dotted blue curves in Fig. 4.36. Assuming the inclination angle of 60° from the CO emission modelling (see Table 4.18 in Sect. 4.3.6.1), the inclination-corrected mass

of W33A is $9.2 \pm 5.8\,M_{\odot}$. At last, the outer radius of the disc associated with W33A was estimated as $r_{out} \sim 900 \pm 100$ AU. The adopted r_{out} value is consistent with the observed sizes of disc candidates associated with high-mass YSOs (∼500-15000 AU, see Table 2 from Reipurth et al., 2007).

The Keplerian mass obtained from the pPXF analysis, presented in the left panel of Fig. 4.36, can be compared with the results from D10, who estimated the total mass of W33A as $13 \pm 2\,M_{\odot}$, assuming a distance of 3.8 kpc. Both results can be compared by scaling the mass obtained by D10 to the same distance of 2.4 kpc (from Table 4.1). The scaled mass is roughly $8 \pm 3\,M_{\odot}$, which is consistent with the mass obtained from the pPXF analysis. The agreement between the Keplerian mass of W33A – reported by D10 after scaling to the distance of 2.4 kpc and the mass derived through the pPXF method – confirms that the methodology adopted in this work is suitable for investigating the kinematics of discs associated with HMYSOs. In addition, the pPXF method also provides an estimation of the errors and delivers reasonable results without requiring highly-smoothed data as used in the results presented by D10 (*e.g.,* see Fig. 4.35), where the errors cannot be clearly assessed.

Results for sources N35, N47 and V49 are presented in Fig. 4.37. The results of all sources are summarised in Table 4.20.

Table 4.20 - Parameters derived from the modelling of the CO absorption features in the spectra of the HMYSOs.

Source	PA (°)	i (°)	$V_{r,out}$ (km s^{-1})	d (kpc)	$M \sin(i)$ ($M_{\odot}$)	r_{out} (AU)
N47	-5	86	±4	3.1	5±3	700±100
V49	165	65	±9	3.8	40±10	700±100
W33A	140	60	±4	2.4	8±3	900±100

Notes: The columns are as follows: (1) ID of the source; (2) the position angle of the rotating axis of the disc (in degrees, from N to E and from the blue- to the red-shifted component); (3) the inclination of the "HMYSO+disc" system (90° corresponds to edge-on view); (4) peak velocity of the blue- and- red-shifted components; (5) distance of the sources; (6) Keplerian mass of the sources; (7) outer radius of the disc.

The top panels of Fig. 4.37 present the radial velocity map of the CO absorption features and the P-V diagram for source V49. As similarly as the residual map from Fig. 4.34, the velocity map indicates a disc-like structure well-separated in velocity and offset by ∼0.2″ from the peak of the continuum emission. The P-V diagram, extracted across the direction of the disc structure, indicates a clear separation between blue- and red-

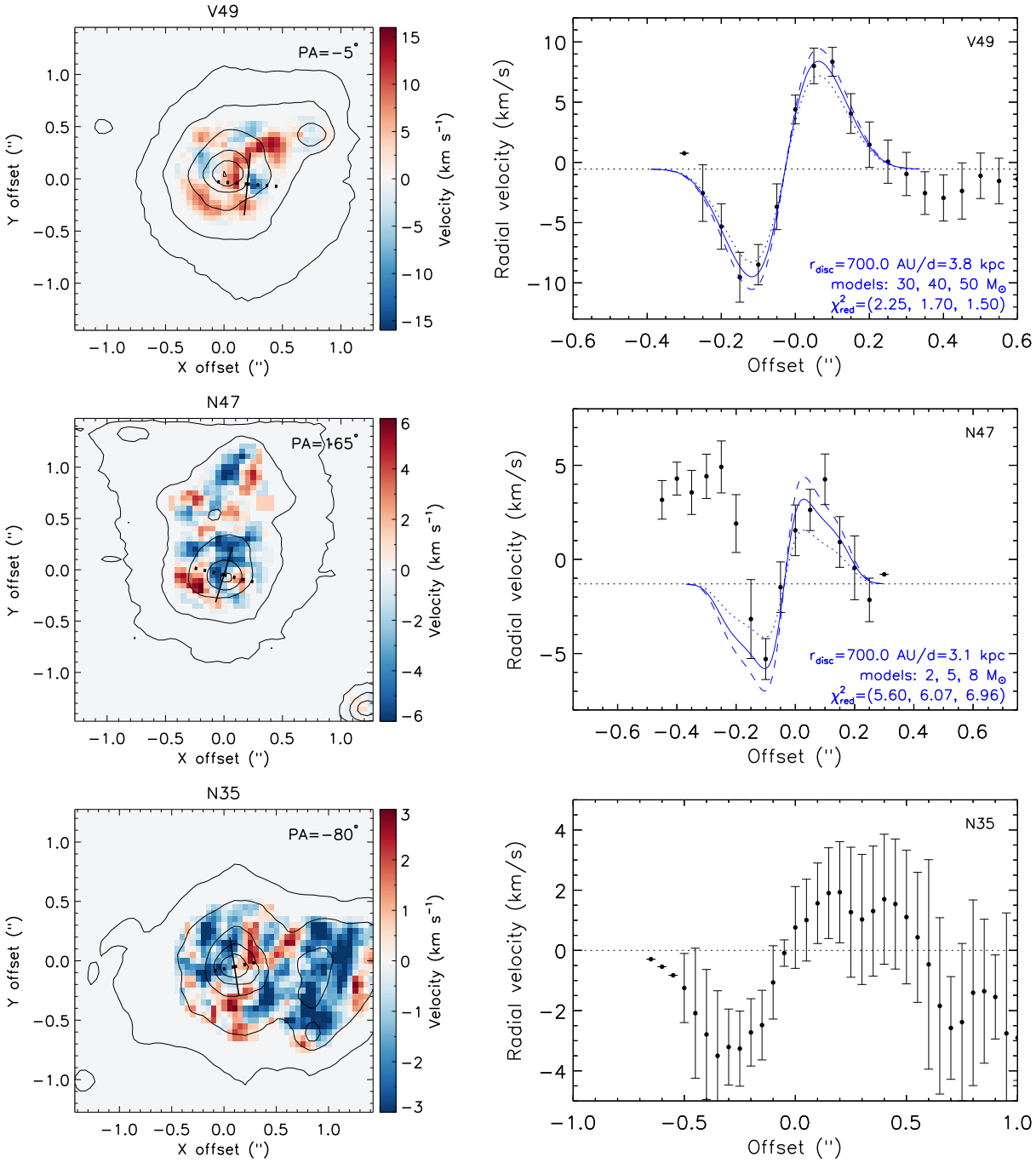


Figure 4.37: Left panels: Radial velocity maps of the CO absorption features, evaluated from the pPXF analysis for sources V49 (top), N47 (middle) and N35 (bottom). A full description of the maps is presented in the left panel of Fig. 4.35. The results towards N35 lack of clear evidence of a rotating structure associated with the HMYSO, while more evident disc signatures are found towards sources N47 and V49. Right panels: Position-velocity diagrams across the dashed line indicated in the residual maps. For a complete description of the figure, see Fig. 4.36. No fit was performed for source N35 (see discussion in the text).

shifted components, with maximum velocities estimated as $\pm 9\,\mathrm{km\,s^{-1}}$. A Keplerian mass of $40 \pm 10\,\mathrm{M_\odot}$ and a disc with a maximum extension of $700 \pm 100\,\mathrm{AU}$ provides the best fit to the observed velocities. Assuming an inclination of 65° (see Table 4.18), the inclination-

corrected Keplerian mass of the source is about $44 \pm 11\,M_{\odot}$.

When compared to the previous two objects, the velocity map of source N47 exhibits a weaker disc signature aligned with a PA = 165°. The P-V diagram was extracted across the disc structure and exhibits spurious red-shifted features at $r \lesssim -0.2''$, closer to the blue-shifted component of the disc. Despite of that, the data display a clear separation between the blue- and red-shifted peaks between $-0.2 \leq r \leq +0.2''$. The velocity offset of $\pm$4-5 km s^{-1} is consistent with a $5 \pm 3\,M_{\odot}$ object at the distance of 3.1 kpc. The inclination of 86° derived in Sect. 4.3.6.1) indicates that the N47 is an almost edge-on object ($\sin(i) \approx 1$).

Finally, the blue- and red-shifted components associated with source N35 are roughly oriented in the same direction of the H_2 jet. Such alignment with respect to the H_2 jets indicates that the kinematics of the CO absorption features are unlikely to be associated with a rotating disc and, therefore, no Keplerian fit was performed. The kinematics of the CO absorption features obtained from the pPXF analysis are relatively less obvious than the residual map presented in Fig. 4.34, which clearly indicates a red-shifted component oriented in the same direction of the blue-shifted H_2 jet. An unresolved system of multiple protostars could explain the observed orientation of the disc-like structure and the observed H_2 jets. In this case, the driving source of the H_2 jet would be a different object than the one associated with the disc candidate. The absence of extended H_2 emission in the direction of the rotating axis of the disc might indicate that the accreting phase of this object has already ceased and the disc is passively evolving, such as those identified by Blum et al. (2004). Alternatively, the absence of H_2 emission aligned with the rotating axis of the CO disc candidate also suggests that the observed CO absorption features have a different origin than a circumstellar disc.

4.4 Discussion of the K-band data

4.4.1 Extended structures associated with the HMYSOs and traced by different lines

The line maps of the H_2, the Brγ and the CO features were obtained in Sect. 4.3.3 and their morphologies and physical parameters were derived. These lines are originated in regions with different conditions of temperature and density. The CO emission is often originated in dense and warm regions, with temperatures ranging between 1000 and 5000 K (Kraus et al., 2000). The Brγ arises from gas with temperatures in the range between 7500

and 30000 K Cooper et al. (2013). The H_2 emission, originates in a colder and more diffuse medium, with temperatures of 2000 K and 4000 K (see Table 4.15).

The analysis of the line maps indicates that the Brγ and the CO features are often tracing similar structures, closer to the main HMYSOs in the FOV, while the H_2 probes more extended and diffuse structures. The extended emission probed by the CO and Brγ lines is correlated with the continuum emission (*e.g.*, see the continuum maps in Fig. 4.10 and the line maps in Fig. 4.14).

By investigating a sample of low-mass YSOs using a similar instrumental setup, Beck et al. (2010) found that extended Brγ emission is present in most of their sources. In some sources, the Brγ emission extends over the direction of the reflection nebulae probed by the continuum maps as similarly as found towards the HMYSOs presented in this Chapter (*e.g.*, see N33, N35, N58 and V49). Those authors also found that, for other low-mass YSOs, the extended Brγ emission is aligned with the direction of the outflows, traced by the [Fe II] line. In the high-mass regime, source N47 exhibits a similar configuration between the Brγ and the molecular outflow probed by the H_2 emission, similarly as reported by Beck et al. (2010) towards low-mass YSOs.

The spectro-astrometry of the Brγ feature is presented in Sect. 4.3.5. The analysis indicated that coherent sub-pixel offsets between the blue- and red-shifted components were found in six of the eight sources investigated in this Chapter. The study of the spectro-astrometry of the Brγ feature towards W33A presented by Davies et al. (2010) suggested that the emission is tracing the cavity of the bipolar jet detected in the H_2 transitions. Indeed, the results presented in Sect. 4.3.5 corroborate the scenario proposed by D10 towards W33A. Those authors measured the position angle of the Brγ feature as 157±25° from N to E direction. The spectro-astrometric analysis presented in this work delivered a PA of 152±28°, which is consistent with the orientation reported by D10. In addition, the projected length of the coherent structure traced by the Brγ emission probes a scale of $\sim$2 AU (or 0.81 mas, see Table 4.17). Such scale corresponds to a fraction 50× smaller than the spaxel-size of the observations (50 mas), evidencing that the spectro-astrometry is a powerful technique to investigate the information present in the NIFS datacubes at sub-pixel scales. The analysis of the spectro-astrometry of the Brγ feature indicated that the sub-pixel variation of the emission arises from a circumstellar structure (*i.e.*, a disc; see N35, N58 and V49) or it probes the ionised gas within the cavity of

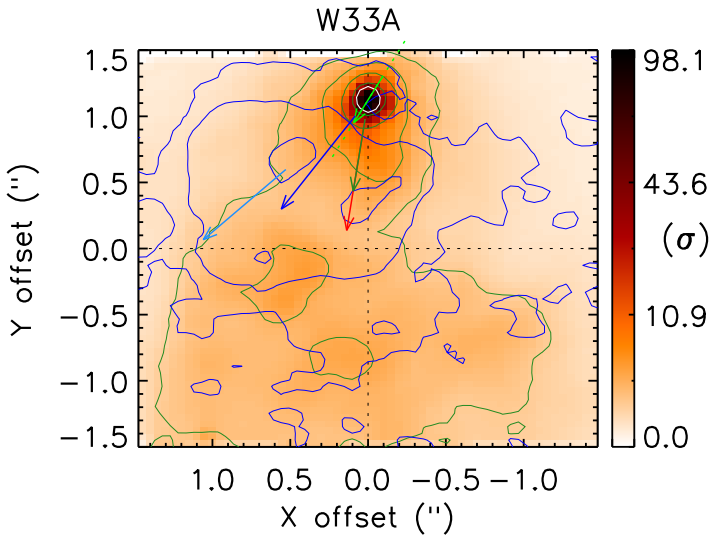


Figure 4.38: Direction of the extended line emission identified in the FOV of W33A source. North is up and East is to the right. The continuum map at 2.2 μm is overlaid by the H_2 (1–0) S(1) (in blue) and Brγ line-map contours (in green). The arrows indicate the position angle (PA) of the following features: the large-scale H_2 emission from Nav15 (in light blue), the H_2 (1–0) S(1) (in blue), Brγ (in green), the spectro-astrometry of the Brγ (in light green) and the CO (2–0) bandhead emission (in red).

a molecular outflow driven by the central source (see N33, N47 and W33A). No clear sub-pixel structures were identified in the analysis of the other two sources (N16 and N21).

In the rest of this section, the orientation of the different structures traced by the line-maps of the H_2, Brγ and the CO bandhead emission (Sect. 4.3.3), and their alignment with the extended continuum emission (Sect. 4.3.1) are discussed. In addition, the geometrical properties of the structures identified in the high-angular resolution NIFS datacubes are compared with those derived for their large-scale counterparts, reported by Navarete et al. (2015) and summarised in Table 4.2. The individual discussion of each source is presented as follows.

W33A (G012.9090−00.2607): Figure 4.38 presents the 2.2 μm continuum map of W33A overlaid by the integrated flux contours of its extended H_2 (in blue) and Brγ emission (in green) maps. The arrows indicate the direction of the extended emission traced by different lines: H_2 (2–1) S(1) at 2.12 μm (in blue), the large-scale H_2 emission from D10 (in light blue), Brγ (in green), the spectro-astrometry of the Brγ feature (in light green) and, when available, the CO (2–0) bandhead emission at 2.29 μm (in red).

N33 (G079.1272+02.2782): This source displays a relatively weak H_2 emission in the N-S direction. The blue-shifted lobe (identified as '1b' in Fig. 4.39) matches the direction of the large-scale H_2 emission from Nav15, but no large-scale H_2 counterpart was found in

the direction of the red-shifted emission ('1r') detected in the NIFS FOV. The H_2 emission is found within a velocity range of $\sim$120 km s^{-1} (from $-$53.9 to +66.7 km s^{-1}). The average excitation temperature of the H_2 emission is about 4.0×10^3 K with a total column density of $\approx 2.5\cdot10^{18}$ cm^{-3} (from Table 4.15). The Brγ emission is oriented to the SW direction roughly following the direction of the continuum emission also observed in the NIFS FOV, suggesting that the extended component of the Brγ is scattered radiation from the central source by the nebula. The structure traced by the spectro-astrometry of the Brγ feature exhibits a PA of 148$\pm$10° and is almost perpendicular to the large-scale Brγ emission (PA = 217°). The linear projected size probed by the spectro-astrometry is about $\sim$10 AU and is tilted by $\sim$25° with respect to the blue-shifted H_2 emission. The orientation of the structure traced by the spectro-astrometry of the Brγ suggests that the line is tracing the cavity of the H_2 jet.

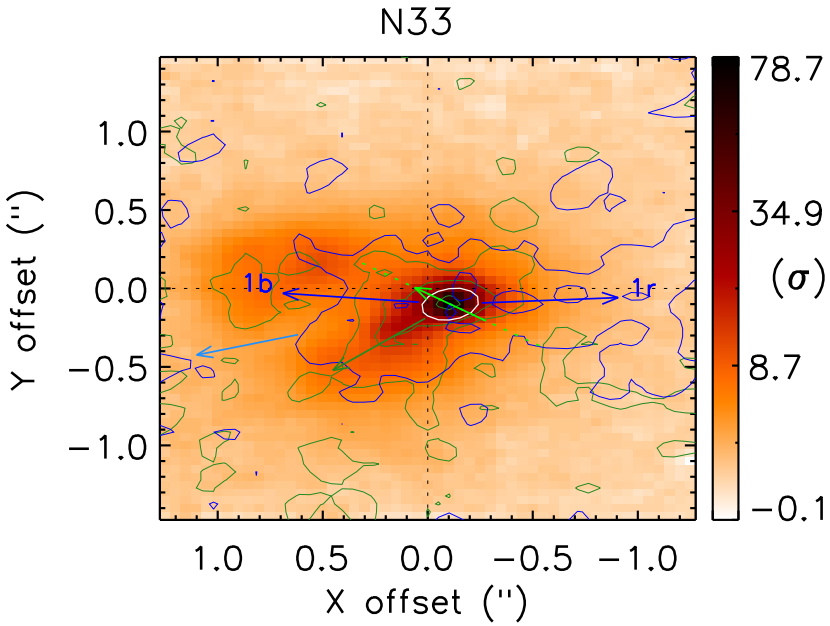


Figure 4.39: Direction of the extended line emission identified in the FOV of source N33. The North direction is oriented to the right and the East is upward. A full description of the elements in the map is given in Fig. 4.38.

N58 (G094.4637−00.8043): Figure 4.40 presents the direction of the structures identified towards the source N58. Extended Brγ and CO emission were identified towards the NE direction of the main object, which also coincide with the extended continuum emission and the orientation of the blue-shifted lobe of the bipolar H_2 jet aligned in the NE-SW direction. The H_2 emission associated with the bipolar structure is detected over a velocity range of $\sim$90 km s^{-1} (from $-$143.8 to $-$53.4 km s^{-1}) and is compatible with the

orientation of the large-scale H_2 outflow from Nav15 (labelled as 2A-2B). The large-scale H_2 map also exhibits another set of bipolar emission in the N-S direction with no counterpart in the observed NIFS FOV. Finally, a third jet-like component is found towards the NW direction (3A), matching with the orientation of the red-shifted component of the jet-like emission found in the SE-NW direction (1b,1r). Although the blue-shifted jet is the brightest in the NIFS FOV, no large-scale counterpart was identified in the map presented by Nav15. The emission of the blue-shifted lobe spans between -143.8 and $-23.3\,\mathrm{km\,s^{-1}}$ ($\Delta v \sim 120\,\mathrm{km\,s^{-1}}$) while the red-shifted lobe is only detected within -113.7 and $-83.5\,\mathrm{km\,s^{-1}}$ ($\Delta v \sim 30\,\mathrm{km\,s^{-1}}$). This either suggests that the structure is aligned to the plane of the sky or that the emission is not likely tracing a bipolar outflow. The main H_2 jet, detected in the NIFS FOV and labelled as '1b' has no large-scale H_2 emission identified in the map presented by Nav15. The orientation of the red-shifted counterpart, labelled as '1r', is aligned with the large H_2 jet '3A'. A second bipolar structure found within the NIFS FOV, labelled as '2b' and '2r' are aligned in the same direction of the large-scale H_2 bipolar emission in the NE-SW direction (labelled as '2A' and '2B'). The NIFS FOV exhibits no clear counterpart of the more distant large-scale bipolar H_2 structure detected in the map presented by Nav15. The observed large-scale distribution of the H_2 emission can be explained by a periodic accretion process combined with the precession of the propagation axis of the outflows. The origin of the precession motion can be caused by gravitational interaction between the main HMYSO presented in the region with companion objects present in the same region. Alternatively, the formation of a multiple stellar system can explain the variety of jet-like structure identified in the NIFS FOV. Indeed, two companion point-like objects separated by $\sim 5{\cdot}10^3$ AU from the main HMYSO were identified in the *K*-band data, supporting that this source is located in a relatively crowded region. Finally, the structure traced by the spectro-astrometry of the Brγ feature is roughly aligned with the other tracers of dense gas (Brγ and CO) and is perpendicular to the orientation of the brightest H_2 jet. The observed configuration suggests that Brγ is tracing a circumstellar disc at scales of $\sim$15 AU. The circumstellar disc was further investigated by the analysis of the CO bandhead emission features detected in the spectrum of the central source. The modelling of the CO (2–0) at $\sim 2.29\,\mu$m is consistent with an almost edge-on system (inclination of 87°), with a stellar mass ranging between 12 and 23 $\mathrm{M_\odot}$.

N35 (G100.3779−03.5784): This source exhibits a bipolar H_2 emission in the N-S di-

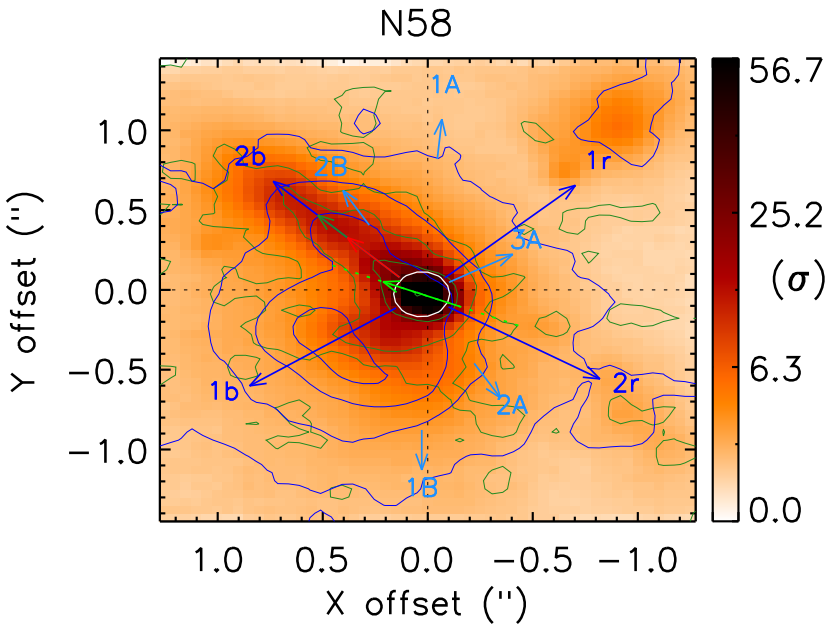


Figure 4.40: Direction of the extended line emission identified in the FOV of source N58. A full description of the elements in the map is given in Fig. 4.38.

rection with velocities spanning from -214.9 to $+86.6\,\mathrm{km\,s^{-1}}$ ($\Delta v \sim 300\,\mathrm{km\,s^{-1}}$). The extended Br$\gamma$ and CO emission coincide with the distribution of the continuum emission and the direction of propagation of the blue-shifted H_2 lobe (which peaks at $v = -94.3\,\mathrm{km\,s^{-1}}$, see Fig. B.4). The large-scale H_2 emission is oriented towards the same direction of the red-shifted H_2 jet identified in the NIFS FOV ($v_{\mathrm{peak}} = +26.3\,\mathrm{km\,s^{-1}}$). No large-scale counterpart of the blue-shifted H_2 jet labelled as '1b' was detected in the map presented by Nav15. The spectro-astrometry of the Brγ feature revealed a structure perpendicular to the orientation of the extended emission traced by the other lines and the continuum. Such configuration suggests that the spectro-astrometry is tracing a circumstellar disc associated with the HMYSO. The projected length separating the blue- and the red-shifted components is about 3.6 AU, with velocities ranging from -300 to $+300\,\mathrm{km\,s^{-1}}$. Assuming a Keplerian motion, the mass enclosed by the structure traced by the Brγ feature is about $M = rv^2G^{-1} \sim 50\,\mathrm{M_\odot}$. The presence of a disc is also confirmed by the identification of CO features in the spectrum of the central source (see Fig. 4.12). However, as presented in Sect. 4.3.8, the kinematics of the CO features is not consistent with a circumstellar disc scenario. Two compact companions were also detected in the NIFS FOV, with linear separations between 3.5-6.5·10^3 AU from the main HMYSO.

V49 (G109.0974−00.3458): The central object exhibits an extended Brγ emission dis-

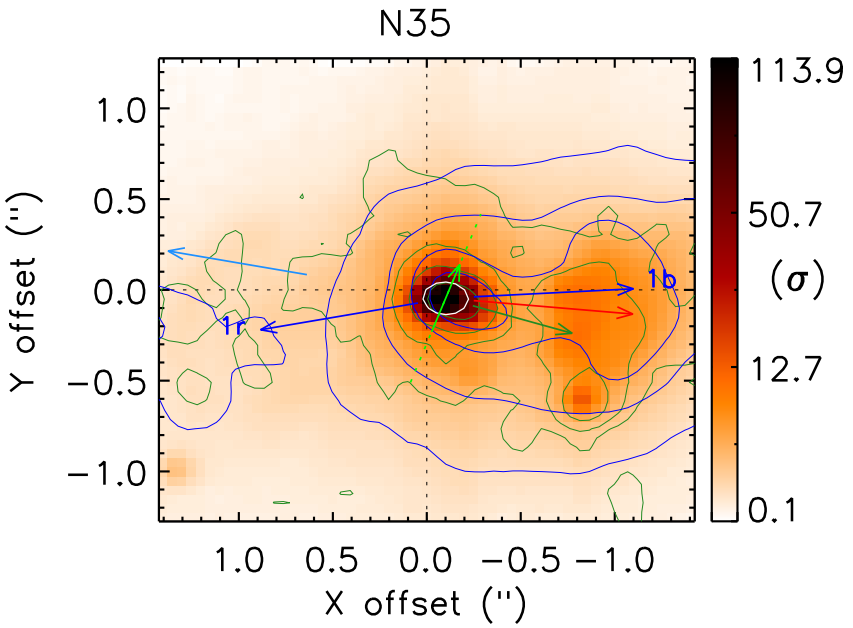


Figure 4.41: Direction of the extended line emission identified in the FOV of the source N35. The North direction is oriented to the right and the East is upward. A full description of the elements in the map is given in Fig. 4.38.

tributed towards the NW direction, following the distribution of the extended continuum emission presented in the map from Fig. 4.42. The spectro-astrometry of the Brγ feature suggests that the line is probing a circumstellar disc around the main source in the field. The Brγ feature exhibits a coherent offset between blue- and red-shifted component, offset by a linear scale of $\sim$2 AU and displaying a velocity separation between -400 and $+300\,\mathrm{km\,s^{-1}}$. The position angle of the structure traced by the Brγ feature was estimated as PA $= 14\pm24°$ (measured in the direction of its rotation axis), matching the orientation of the disc-like structure obtained from the analysis of the CO absorption features ($-5°$, see Fig. 4.37). The mass of the central source was estimated as 8-80 $M_\odot$ through the modelling of its CO emission, while the analysis of the kinematics of the CO absorption feature led to a total mass of the system of $40\pm10\,M_\odot$. The lower-limit of these estimates are roughly consistent with the total mass expected from the bolometric luminosity of the RMS source, $M_* \sim$ 8-12 $M_\odot$. Although disc signatures where found towards the main object in the NIFS FOV, the H_2 emission seems to arise from a source offset by $\approx1\farcs5$ to the S direction of the brightest object in the field, identified as the compact and point-like H_2 emission at the given position. The direction of the H_2 emission (peak velocity of $-98.6\,\mathrm{km\,s^{-1}}$, see Fig. B.11) is consistent with the large-scale H_2 jet labelled as '1B'. In addition to the compact source traced by the H_2 line, two secondary compact objects were also identified in

the continuum maps. They have linear separations of ~ 3 and $4.3{\cdot}10^3$ AU from the brightest object in the centre of the FOV. The presence of these companions indicates a high multiplicity of objects in the vicinity of the RMS source.

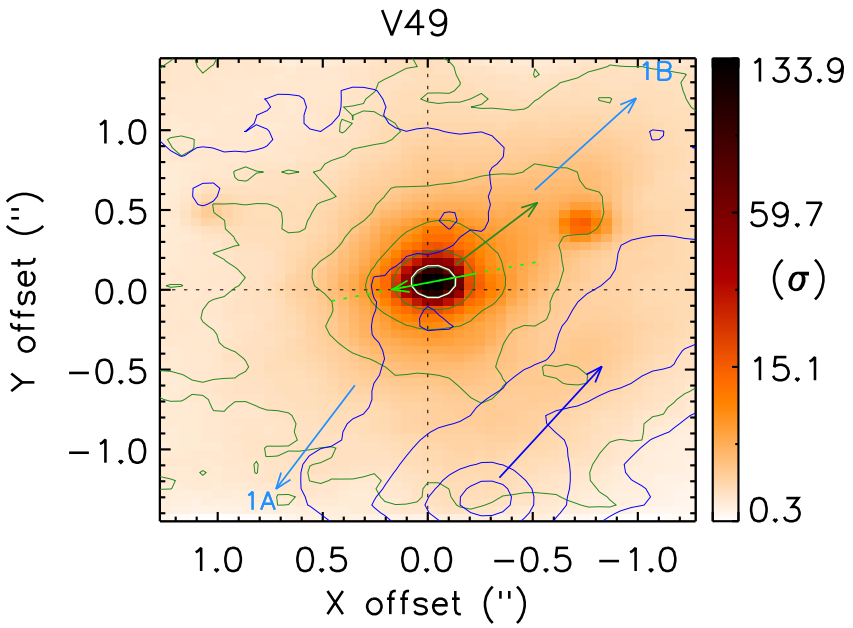


Figure 4.42: Direction of the extended line emission identified in the FOV of source V49. A full description of the elements in the map is given in Fig. 4.38.

N47 (G124.0144−00.0267): The H_2 emission is detected in a wide range of velocities (from -269.2 to $-28.0\,\mathrm{km\,s^{-1}}$, $\Delta v \sim 240\,\mathrm{km\,s^{-1}}$), and its maximum intensity peaks at $-88.3\,\mathrm{km\,s^{-1}}$. The orientation of the H_2 structure matches the direction of the extended Brγ emission and the large-scale H_2 emission '1A' from Nav15. Except for a weak knot to the S from the main source, no red-shifted counterpart of the H_2 emission is detected in the FOV. The spectro-astrometry of the Brγ feature probes a structure that extends up to 9.8 AU, with a position angle of 341±32°, roughly aligned with the H_2 emission. The shape of the Brγ kinematics indicates a wide angle opening towards velocities ranging from -400 to $-300\,\mathrm{km\,s^{-1}}$ and a narrow red-shifted region with velocities ranging from $+300$ to $+400\,\mathrm{km\,s^{-1}}$. Such structure is likely probing the cavity of the observed H_2 jet and, together with the absence of a red-shifted H_2 emission, is consistent with a source displaying a high-inclination angle with respect to the plane of the sky. However, the analysis of the CO emission suggests the presence of a disc in an almost edge-on orientation ($i = 86°$) and the shape of the CO bandhead profile is consistent with a source of mass ranging between 29 and 59 $M_\odot$. The analysis of the CO absorption features indicates disc-like structure

with a PA estimated in −5° and a Keplerian mass of the system estimated in 5±3 $M_\odot$. This source has a point-like companion towards the SW direction, with a projected linear distance of ∼5.4·10^3 AU from the brightest object in the centre of the FOV.

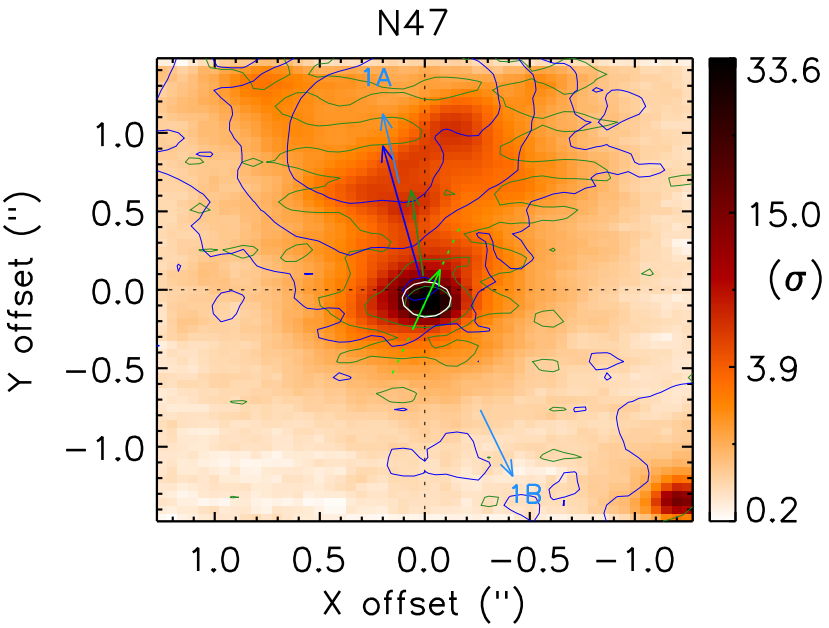


Figure 4.43: Direction of the extended line emission identified in the FOV of source N47. A full description of the elements in the map is given in Fig. 4.38.

N16 (G108.7575−00.9863): The map indicates that the direction of the large-scale H_2 emission is aligned with the maximum elongation of the continuum emission, which seems to trace the cavity of an outflow. However, this source does not exhibit any discrete spectral feature towards its central spectrum neither any defined morphology traced by the most common lines detected in the other sources investigated in this thesis (e.g. H_2 and Brγ). The absence of discrete spectral features combined with the steeper spectrum of the central source suggests that the thermal emission from the circumstellar dust is producing a veiling-like effect on the observed spectrum. Martin (1996) suggested that the NIR veiling arises from the heated accretion flow in T Tauri stars, with temperatures between 3000 and 6000 K. Indeed, the observed *J*-band veiling measured towards these objects is roughly proportional to the mass accretion rate, with values ranging between 10^{-9} and 10^{-5} $M_\odot yr^{-1}$ (Folha, 2003). Despite the agreement between these effects in low-mass YSOs, the extension and validity of the veiling mechanism in high-mass YSOs is still debatable. In the high-mass regime, the strength of the veiling effect may be correlated with high accretion rates or outbursts (Connelley and Greene, 2010).

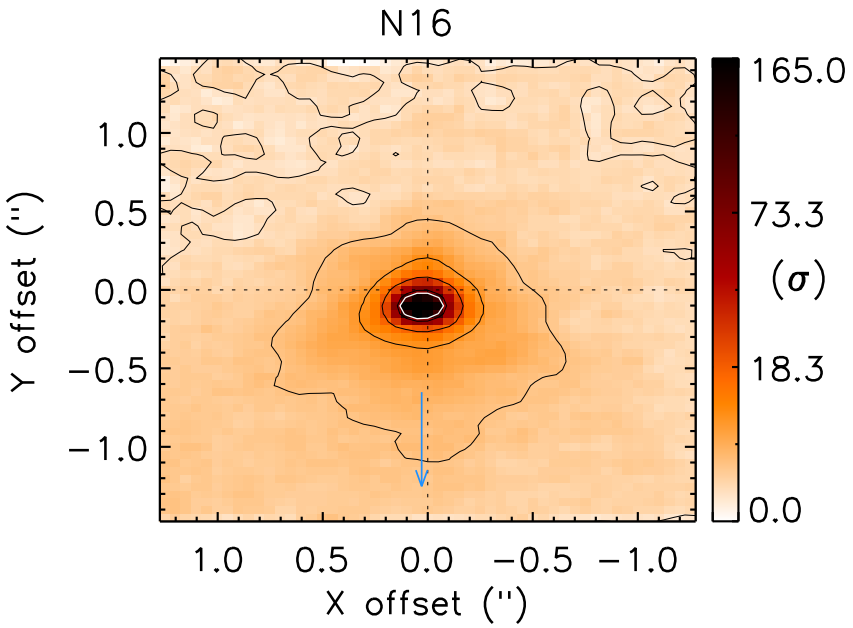


Figure 4.44: Direction of the extended line emission identified in the FOV of source N16. The black contours corresponds to the 3^n-σ levels ($n = 0, 1, 2, ...$) of the continuum emission at 2.20 μm. The North direction is oriented to the right and the East is upward. A full description of the elements in the map is given in Fig. 4.38.

N21 (G173.6328+02.8064): The analysis of source N21 does not provide much information on its nature. The source exhibits an elongated continuum emission with no clear discrete spectral feature presented on its *K*-band data, except for a weak Brγ emission. The spatial distribution of the Brγ is relatively diffuse and faint (see Fig. 4.45), leading to a doubtful estimation of its origin. The analysis of the spectro-astrometry of the Brγ feature, however, exhibits the largest offsets observed in this work (around 50 mas and corresponding to a linear scale of ∼76 AU) oriented in the same direction of the continuum emission. No clear conclusions were found for this object and it is unlikely that the source observed in the NIFS FOV corresponds to the driving source of the large-scale H_2 emission identified in Nav15.

4.4.2 Collimation degree of the jets along their propagation axis

One of the primary objectives of this work was to estimate the collimation of the H_2 jets at large- and small-scales, comparing the high-angular resolution NIFS observations with the large-scale H_2 maps presented by Nav15.

Figure 4.46 presents the aspect ratio (defined as the ratio between the length and the width of the structures) of the H_2 jets identified in the NIFS datacubes (from Table 4.11)

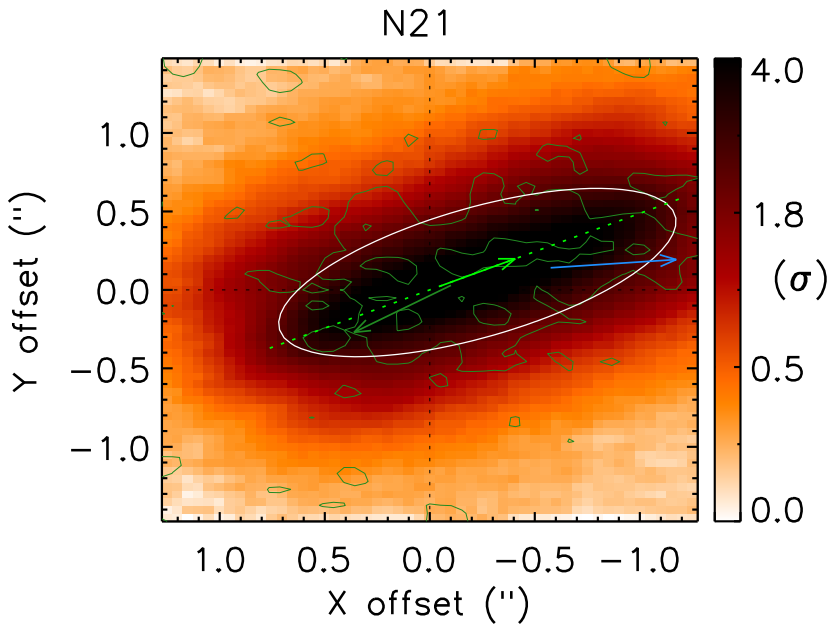


Figure 4.45: Direction of the extended line emission identified in the FOV of source N21. A full description of the elements in the map is given in Fig. 4.38

as a function of the aspect ratio of their large-scale H_2 counterpart (reported in Table 4.2). The comparison between the collimation degree of the jets closer to the launching point and their large-scale counterpart indicates that the observed aspect ratio of the structures can either increase or decrease along their propagation axis. The plot indicates that the H_2 jets driven by the sources N33, N47, V49 and W33A are located above the $y{=}x$ line, suggesting that the H_2 jets are more collimated at larger distances from their launching points. Source N58, however, exhibits a different behaviour: the jets seem to de-collimate along their propagation axis.

Magneto-hydrodynamical (MHD) simulations and high-resolution observations of a wide range of outflow configurations were used to investigate the effects of collimation or de-collimation of the molecular outflows in a wide range of (proto)stellar masses and configurations (*e.g.,* multiplicity, precession motions). The principal results are listed as follows. (Pudritz et al., 2006) have shown that the highly collimated and wide-angle jets and outflows can be produced by a disc-wind driven mechanism, depending on the amount of the so-called mass load (that is, the observed mass-loss rate of the outflow per unit of magnetic flux). According to those authors, low mass load values lead to episodic and rapid behaviour while higher values correspond to stationary outflows. More recently, Vaidya et al. (2011) performed MHD simulations including the effects of the radiative field

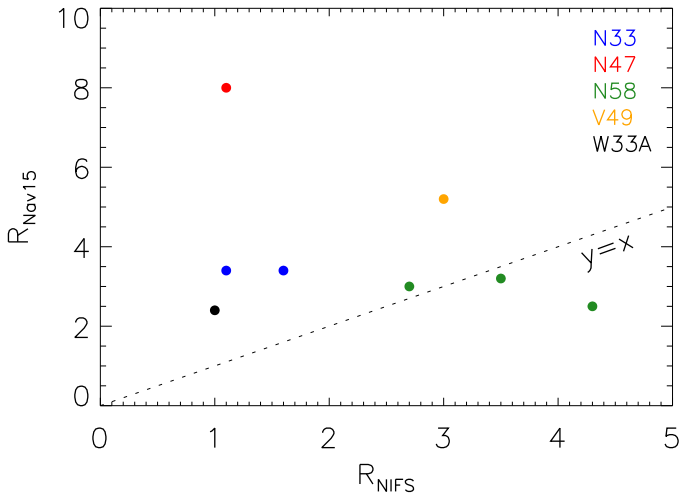


Figure 4.46: Aspect ratio of the H_2 jets identified in the NIFS datacubes against the aspect ratio of their large-scale counterparts from Nav15. Data from different sources are shown in different colours. The dashed black line indicates the $y{=}x$.

of the protostar onto the jet and outflow dynamics. They found that the outflows become less collimated as a function of the mass of the protostar, with opening angles going from 20° to 32° as the stellar mass increased from 20 to 60 $M_\odot$. They also found that changes in the magnetic flux do not significantly affect the collimation of the jets. Therefore, high-collimated jets would be expected at the early stages of high-mass star formation while poorly collimated jets would be observed in more evolved HMYSOs. Cesaroni et al. (2017) reported the detection of poorly collimated bipolar SiO outflows in two (G29.96 and G31.41) of six HMYSOs in their sample, observed with the ALMA interferometer at high-angular resolution (∼0.″2). Based on the simulations of Peters et al. (2014), Cesaroni et al. 2017 speculate that the observed wide-angle SiO emission could be a consequence of a precessing jet due to the presence of interacting companions. Indeed, Peters et al. (2014) found that wide-angle outflows are observed when multiple stars form in the same accretion flow.

The apparent de-collimation observed in the H_2 structures associated with source N58 (see Fig. 4.46) can be interpreted as an effect of precession of the launching point of the collimated jet (*i.e.*, the circumstellar disc), which results in well-collimated structures closer to their launching point, producing a large-scale wide-angle structure as a result of the combined effect of the interaction between the flow and the quiescent interstellar medium. Poorly collimated outflows can be produced as a result of extreme precession of the jet (*e.g.*, Shepherd et al., 2000). The presence of multi-polar H_2 jets located at different

projected distances from the that seems to be driven by the RMS source is an indicator that precession motions and episodic accretion process were frequent during the whole accretion period of this source, suggesting that the large-scale outflows are wide-angle structures due to these phenomena.

Chapter 5

Conclusions

This thesis presented the analysis of the large-scale CO emission from high-mass clumps selected from the ATLASGAL survey, and the characterisation of the circumstellar environment of high-mass YSOs from the RMS survey, exhibiting large-scale molecular H_2 jets.

In Chapter 3, the ATLASGAL TOP100 sample was characterised in terms of their CO (4–3), CO (6–5) and CO (7–6) emission. All the TOP100 clumps were detected in the three J CO transitions. The analysis of the CO (6–5) maps indicated that the mid-J CO emission probes linear scales between 0.1 and 2.4 pc. The extension of the mid-J CO emission is correlated with the mass of the clumps, suggesting that the extension of the CO gas is linked to the available amount of the total gas in the region. In addition, the mid-J CO extension of the 70 μm-bright TOP100 clumps is also correlated with the far-infrared emission probed by the *Herschel*-PACS 70 μm maps.

The CO emission was investigated in terms of different spatial resolutions on the mid-J CO. First, the CO (6–5) and CO (7–6) data were convolved to a common angular resolution of $13\farcs4$, including the single-pointing CO (4–3) data. Then, the mid-J CO data were convolved to a common linear size of $\sim$0.24 pc using a distance limited sub-sample of clumps and the mid-J CO emission was integrated over the entire field-of-view of the CHAMP$^+$ maps for another subsample of the TOP100 sample. The analysis of the CO emission integrated over different angular and linear scales indicated that the relations between the CO and the clump properties usually are not dependent on the spatial resolution of the data, at least for typical distances within the Galaxy. These results are encouraging for studies of large samples of star-forming regions across the Galaxy based on single-pointing observations.

The CO luminosity ($L_{\rm CO}$) of the three transitions exhibit a progressively increase as a function of the evolutionary class of the clumps (from 70w clumps to H II regions). The CO luminosity is correlated with $L_{\rm bol}$, suggesting that the integrated flux of the CO is dependent on the radiative flux of the objects within the clumps. In addition, the $L_{\rm CO}$ is correlated with the mass of the clumps, but not with their luminosity-to-mass ratio. The dependency of the $L_{\rm CO}$ with $L_{\rm bol}$ is steeper at higher-J CO transitions, suggesting that higher J CO transitions are more sensitive to the temperature of the gas and likely arise from an inner envelope, as recently discussed by Benz et al. (2016).

The excitation temperature ($T_{\rm ex}$) of the clumps was derived from the peak intensity of the Gaussian fit on the CO (6–5) spectra. The comparison with the clump properties indicated that the temperature of the CO gas increases as a function of $L_{\rm bol}$ and L/M, as expected for a warming up of the gas as a function of time.

Finally, the CO luminosity was investigated as a function of the bolometric luminosity of low- and high-luminosity sources for the CO (6–5), CO (7–6) and CO (10–9) lines. For every J transition, the CO luminosity shows similar slopes as a function of the$L_{\rm bol}$ of the low- and high-luminosity sources. However, the relations between $L_{\rm CO}$ and $L_{\rm bol}$ get steeper when combining both subsamples. The steeper relations are interpreted in terms of an existent offset between the CO luminosity from low- and high-luminosity clumps. The CO emission observed in more luminous and distant objects is likely originated from multiple sources within the linear scale probed by the size of beam (up to 0.84 pc) and are systematically larger than the emission from resolved and nearby less luminous objects, from which the CO emission is integrated over smaller linear scales ($\sim$0.01 pc).

The line-width of the CO lines increases as a function of the evolutionary stages of the TOP100 clumps (from 70w to H II regions, suggesting that high-velocity CO components are likely to trace molecular outflows even at the latest evolutionary classes. This suggests that accretion, resulting in outflows, is still undergoing in the more evolved clumps of the TOP100.

In Chapter 4, the circumstellar environment of eight high-mass YSOs exhibiting large-scale H_2 jets was investigated through high-angular resolution K-band integral field spectroscopic observations. At linear resolutions of $\sim10^2$ AU, all sources exhibit a compact point-like object associated with extended emission in both the continuum and in discrete

spectral features (Brγ, H_2 and CO). Most of the central sources exhibit typical features of embedded objects, including a steeper continuum emission and discrete spectral features such as the Brγ, H_2 and CO emission and absorption lines.

The H_2 emission traces extended emission closer to the launching point of the observed large-scale H_2 jets. At scales of hundreds of AUs, the extended H_2 exhibits typical temperatures of $\sim(3.7\pm0.8)\cdot10^3\,K$ and column densities around $10^{18-19}\,cm^{-2}$. The comparison between the aspect ratio of the H_2 jets closer to their launching point and their large-scale counterparts indicated that most of the structures are more collimated at larger distances from their origin. Source N58 exhibits multiple polar structures that are less collimated at larger distances from their launching point. The decollimation of the jets along their propagation axis can be explained by multiplicity of sources and precession motions associated with the launching point of the jets. These predictions are supported by recent MHD simulations.

The Brγ emission probes the ionized gas in the vicinity of the central HMYSOs. At sub-pixel resolutions, the spectro-astrometry of the Brγ feature indicated that the line either traces the cavity of the outflows or probes rotating structures perpendicular to the H_2 jets (*i.e.*, the inner region of the circumstellar discs).

The CO features are originated in high-density environments, such as the circumstellar discs. The combined analysis of the CO emission modelling and the kinematics of the spatially resolved CO absorption features, the Keplerian masses of three HMYSOs were estimated in 5 ± 3, 8 ± 5 and $30\pm10\,M_\odot$.

These results support that the key-structures of the disc-mediated accretion scenario – the circumstellar discs and molecular jets – are present in intermediary and high-mass protostars. These evidences support that stars more massive than $8\,M_\odot$ are likely formed through a similar process as their low-mass counterparts.

5.1 Perspectives and ongoing works on the TOP100 sample

The analysis of the mid-J CO emission presented in Chapter 3 indicates that high-velocity components are found in most of the spectra of the TOP100 clumps. Such high-velocity emission is often associated with the presence of molecular outflows that can be further exploited using the available data and high-angular interferometric observations from ALMA. In this section, the perspectives and works in progress related to the characterisation of the outflow content towards the TOP100 clumps are presented.

5.1.1 ALMA observations towards a well-selected sample of molecular outflows

The circumstellar accretion model by Shu et al. (1987) predicts that, for low-mass YSOs, the ejection of material as bipolar jets collimated by their strong magnetic field. However, observational evidence of collimated jets are small among high-mass YSOs with $L_{\rm bol} > 10^4\,{\rm L}_\odot$ (Varricatt et al., 2010a; Navarete et al., 2015), even though models predict their presence up to masses of 30-35 $\rm M_\odot$ (Kuiper et al., 2015).

The study of the mid-J CO outflows associated with the TOP100 clumps from Navarete et al. (in prep.) indicates that ∼85% of their sample is associated with molecular outflows. The TOP100 is a well-characterised and understood sample of high-mass objects in the Galaxy and, therefore, is ideally suited to investigate the structures associated to the star-forming process at scales of $\sim 10^3$ AU at GHz frequencies, that can only be probed using high-resolution interferometric data.

By taking advantage of the available information on the TOP100 sample, five clumps exhibiting clear outflow activity (see Fig. 5.1) were selected to be observed using the ALMA facility by adopting the following criteria:

1. sources with bolometric luminosity greater than $10^5\,{\rm L}_\odot$ (corresponding to a single O9 or earlier ZAMS star);

2. distances below 7 kpc (linear resolutions of ≤ 1400 AU at the proposed resolution);

3. non-Gaussian wings with high-velocity (CO (6–5) profiles with full-widths at zero power $\geq 60\,{\rm km\,s^{-1}}$).

The sample was observed using two ALMA-12m configurations: C40-6 and C40-3 (Cycle 4), with maximum baselines of 460 and 1800 m, respectively; and minimum baseline of

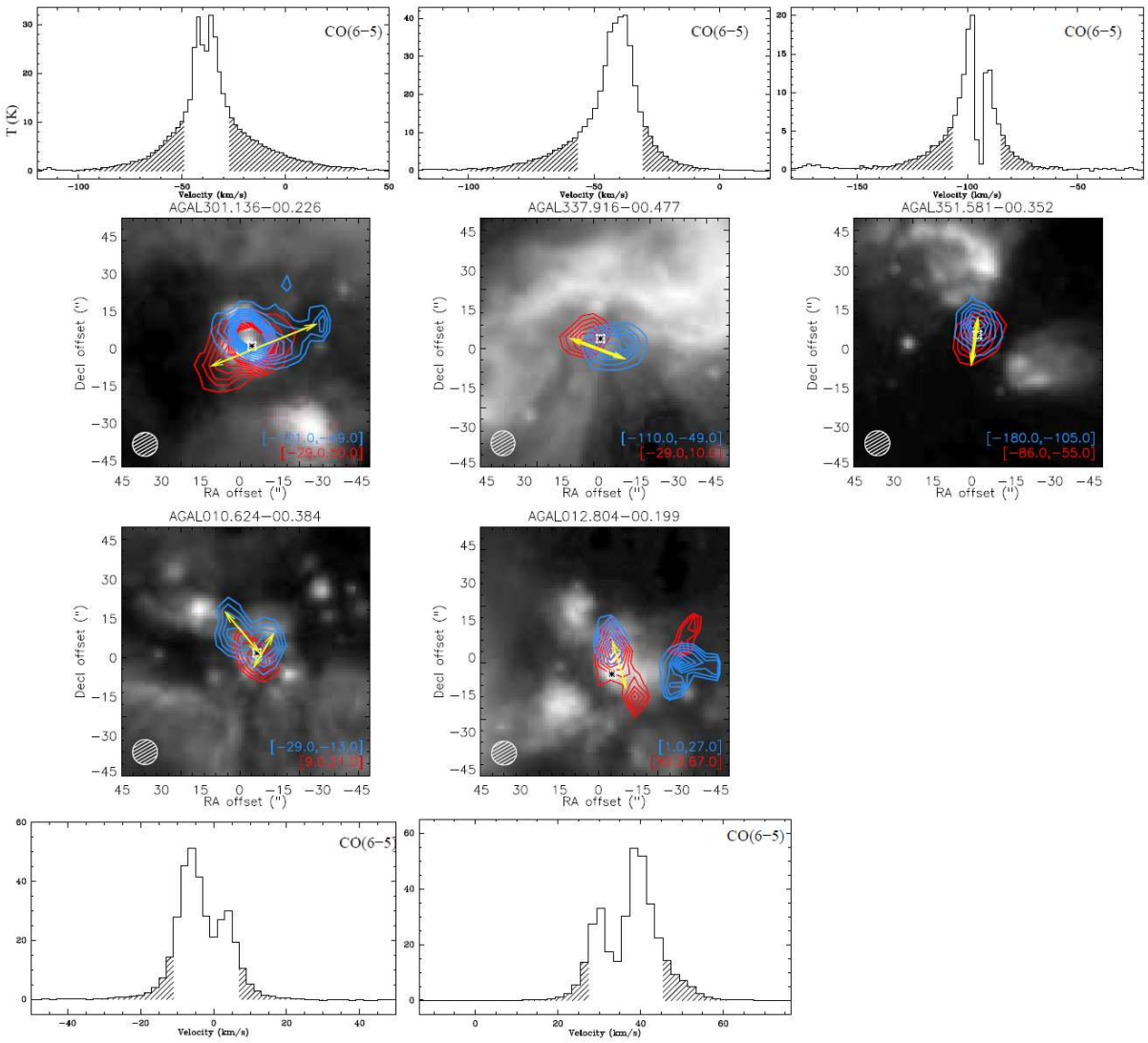


Figure 5.1: CO (6–5) spectra of the targets. The regions filled with parallel lines outline the high-velocity emission associated with the outflows. Middle panels: Spitzer 8.0 μm image and integrated CO (6–5) red- and blue-shifted emission. The velocity ranges for the integration are shown in each panel. The arrows are the main axis of the outflows covered by the mosaics. The APEX 8″ beam size is shown in the bottom left corner. From Leurini (priv. comm.).

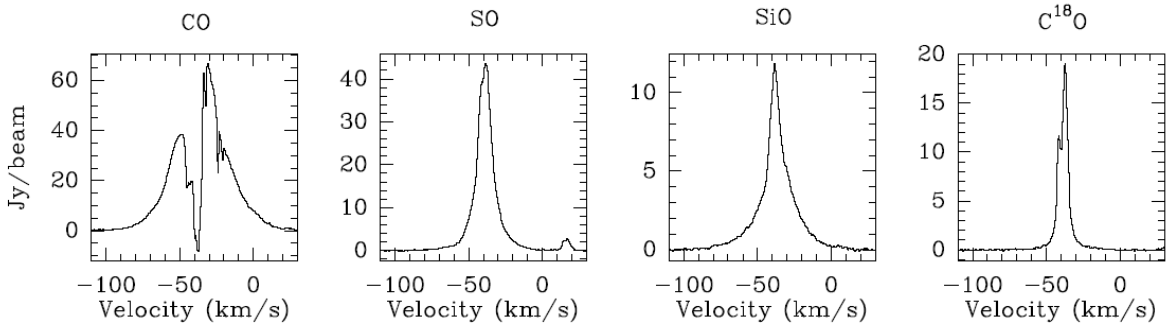


Figure 5.2: Atacama Compact Array observations of the targeted lines at the central position of source AGAL301.136−00.226. From Leurini (priv. comm.).

15 m; corresponding to angular resolutions of 1″.5 and 0″.35 at ∼230 GHz[1]. The achieved spatial resolution is sufficient to probe the structure of the objects in scales of ∼1000 AU, using the transitions listed in Table 5.1. Figure 5.2 presents the spectra of the targeted transitions taken at the central position of AGAL301.136−00.226 using the Atacama Compact Array (ACA) facility.

Table 5.1 - List of transitions and the corresponding structures probed by each line.

Transition	Frequency (GHz)	Targeted structures
SiO (5-4)	217.105	collimated jet (high velocities)
$C^{18}O$ (2-1)	219.560	outflow cavities (low to intermediate velocities)
SO(6_5-5_4)	219.949	collimated jet (high velocities)
CO (2-1)	230.538	outflow cavities and collimated jet (intermediate to high velocities)

ALMA data will provide unprecedented observational constraints and answers to two main questions of the high-mass star formation scenario: if the jet and outflow-cavity structure observed in low- and intermediate-mass YSOs (see Fig. 4 from Wang et al., 2014) is also present in high-mass YSOs; and if outflows from HMYSOs are driven by collimated jets as their low-mass counterparts.

5.2 *Principal Component Analysis of the NIFS data*

The Principal Component Analysis (PCA) is a powerful technique for extracting information from datacubes, as presented by Steiner et al. (2009) and briefly described below. The PCA transforms the observed datacubes into a system of uncorrelated coordinates

[1] informations available in https://almascience.nrao.edu/observing/prior-cycle-observing-and-configuration-schedule

ordered by their variance. These new coordinates are the eigenvectors, containing information in the spectral dimension. The projections of the data onto the eigenvectors produce two-dimensional maps (images), referred as tomograms. The eigenvectors are mutually orthogonal, and each one is associated with a tomogram. The correct association of the eigenvectors and tomograms is fundamental for the interpretation of the results.

In the past years, the PCA has been used successfully for studying the kinematics and the distribution of distinct spatial structures associated with Extragalactic objects (Ricci et al., 2014; Menezes et al., 2014b; Ricci et al., 2018) but no study in stellar objects was performed yet.

As presented in Chapter 4, the circumstellar environment of the HMYSOs are rich in terms of spectral features (*e.g.*, Brγ, H_2 and CO features) and spatial structures (*e.g.*, jets and discs), representing an ideal case for testing the PCA technique. For this reason, the NIFS datacubes will be analysed using the PCA methodology described by Steiner et al. (2009) and Menezes et al. (2014a) and the results will be compared to those presented in Chapter 4.

Bibliography

Bacciotti F., Mundt R., Ray T. P., Eislöffel J., Solf J., Camezind M., Hubble Space Telescope STIS Spectroscopy of the Optical Outflow from DG Tauri: Structure and Kinematics on Subarcsecond Scales, ApJ, 2000, vol. 537, p. L49

Bachiller R., Bipolar Molecular Outflows from Young Stars and Protostars, ARA&A, 1996, vol. 34, p. 111

Bally J., Reipurth B., Davis C. J., Observations of Jets and Outflows from Young Stars, Protostars and Planets V, 2007, pp 215–230

Battersby C., Bally J., Jackson J. M., Ginsburg A., Shirley Y. L., Schlingman W., Glenn J., An Infrared Through Radio Study of the Properties and Evolution of IRDC Clumps, ApJ, 2010, vol. 721, p. 222

Bayet E., Gerin M., Phillips T. G., Contursi A., Are ^{12}CO lines good indicators of the star formation rate in galaxies?, MNRAS, 2009, vol. 399, p. 264

Beck T. L., Bary J. S., McGregor P. J., Spatially Extended Brackett Gamma Emission in the Environments of Young Stars, ApJ, 2010, vol. 722, p. 1360

Bensch F., Stutzki J., Heithausen A., Methods and constraints for the correction of the error beam pick-up in single dish radio observations, A&A, 2001, vol. 365, p. 285

Benz A. O., Bruderer S., van Dishoeck E. F., Melchior M., Wampfler S. F., van der Tak F., Goicoechea J. R., Indriolo N., Kristensen L. E., Lis D. C., et al., WISH VI. Constraints on UV and X-ray irradiation from a survey of hydrides in low- to high-mass YSOs, ArXiv e-prints, 2016

Beuther H., Walsh A., Schilke P., Sridharan T. K., Menten K. M., Wyrowski F., CH_3OH and H_2O masers in high-mass star-forming regions, A&A, 2002, vol. 390, p. 289

Bik A., Thi W. F., Evidence for an inner molecular disk around massive Young Stellar Objects, AAp, 2004, vol. 427, p. L13

Blum R. D., Barbosa C. L., Damineli A., Conti P. S., Ridgway S., Accretion Signatures from Massive Young Stellar Objects, ApJ, 2004, vol. 617, p. 1167

Blum R. D., McGregor P. J., The Ionizing Stars of the Galactic Ultra-Compact H II Region G45.45+0.06, AJ, 2008, vol. 135, p. 1708

Bonatto C., Bica E., Constraining the star formation rate in the solar neighbourhood with star clusters, MNRAS, 2011, vol. 415, p. 2827

Bonnell I. A., Bate M. R., Clarke C. J., Pringle J. E., Competitive accretion in embedded stellar clusters, MNRAS, 2001, vol. 323, p. 785

Burrows C. J., Stapelfeldt K. R., Watson A. M., Krist J. E., Ballester G. E., Clarke J. T., Crisp D., Gallagher III J. S., Griffiths R. E., Hester J. J., Hoessel J. G., Holtzman J. A., Mould J. R., Scowen P. A., Trauger J. T., Westphal J. A., Hubble Space Telescope Observations of the Disk and Jet of HH 30, ApJ, 1996, vol. 473, p. 437

Cappellari M., Emsellem E., Parametric Recovery of Line-of-Sight Velocity Distributions from Absorption-Line Spectra of Galaxies via Penalized Likelihood, PASP, 2004, vol. 116, p. 138

Caratti o Garatti A., Giannini T., Nisini B., Lorenzetti D., H_2 active jets in the near IR as a probe of protostellar evolution, A&A, 2006, vol. 449, p. 1077

Caratti o Garatti A., Stecklum B., Linz H., Garcia Lopez R., Sanna A., A near-infrared spectroscopic survey of massive jets towards extended green objects, A&A, 2015, vol. 573, p. A82

Cardelli J. A., Clayton G. C., Mathis J. S., The relationship between infrared, optical, and ultraviolet extinction, ApJ, 1989, vol. 345, p. 245

Carey S. J., Noriega-Crespo A., Mizuno D. R., Shenoy S., Paladini R., Kraemer K. E., Price S. D., Flagey N., Ryan E., Ingalls J. G., Kuchar T. A., et al., MIPSGAL: A Survey of the Inner Galactic Plane at 24 and 70 μm, PASP, 2009, vol. 121, p. 76

Carr J. S., Near-infrared CO emission in young stellar objects, ApJ, 1989, vol. 345, p. 522

Cesaroni R., Galli D., Neri R., Walmsley C. M., Imaging the disk around IRAS 20126+4104 at subarcsecond resolution, A&A, 2014, vol. 566, p. A73

Cesaroni R., Sánchez-Monge Á., Beltrán M. T., Johnston K. G., Maud L. T., Moscadelli L., et al., Chasing discs around O-type (proto)stars: Evidence from ALMA observations, A&A, 2017, vol. 602, p. A59

Chambers E. T., Jackson J. M., Rathborne J. M., Simon R., Star Formation Activity of Cores within Infrared Dark Clouds, ApJS, 2009, vol. 181, p. 360

Chandler C. J., Sargent A. I., The Role of Dust in Star and Planet Formation: Observations. In From Stardust to Planetesimals , vol. 122 of Astronomical Society of the Pacific Conference Series, 1997, p. 25

Chrysostomou A., Bacciotti F., Nisini B., Ray T. P., Eislöffel J., Davis C. J., Takami M., Investigating the transport of angular momentum from young stellar objects. Do H2 jets from class I YSOs rotate?, A&A, 2008, vol. 482, p. 575

Churchwell E., Ultra-Compact HII Regions and Massive Star Formation, ARA&A, 2002, vol. 40, p. 27

Cohen M., Wheaton W. A., Megeath S. T., Spectral Irradiance Calibration in the Infrared. XIV. The Absolute Calibration of 2MASS, AJ, 2003, vol. 126, p. 1090

Connelley M. S., Greene T. P., A Near-infrared Spectroscopic Survey of Class I Protostars, AJ, 2010, vol. 140, p. 1214

Contreras Y., Schuller F., Urquhart J. S., Csengeri T., Wyrowski F., Beuther H., Bontemps S., Bronfman L., Henning T., Menten K. M., Schilke P., Walmsley C. M., Wienen M., Tackenberg J., Linz H., ATLASGAL - compact source catalogue, A&A, 2013, vol. 549, p. A45

Cooper H. D. B., Lumsden S. L., Oudmaijer R. D., Hoare M. G., Clarke A. J., Urquhart J. S., Mottram J. C., Moore T. J. T., Davies B., The RMS survey: near-IR spectroscopy of massive young stellar objects, MNRAS, 2013, vol. 430, p. 1125

Corcoran M., Ray T. P., Wind diagnostics and correlations with the near-infrared excess in Herbig Ae/Be stars, A&A, 1998, vol. 331, p. 147

Csengeri T., Leurini S., Wyrowski F., Urquhart J. S., Menten K. M., Walmsley M., Bontemps S., Wienen M., Beuther H., Motte F., Nguyen-Luong Q., Schilke P., Schuller F., Zavagno A., Sanna C., ATLASGAL-selected massive clumps in the inner Galaxy. II. Characterisation of different evolutionary stages and their SiO emission, A&A, 2016, vol. 586, p. A149

Csengeri T., Urquhart J. S., Schuller F., Motte F., Bontemps S., Wyrowski F., Menten K. M., Bronfman L., Beuther H., Henning T., Testi L., Zavagno A., Walmsley M., The ATLASGAL survey: a catalog of dust condensations in the Galactic plane, A&A, 2014, vol. 565, p. A75

Cyganowski C. J., Whitney B. A., Holden E., Braden E., Brogan C. L., Churchwell E., Indebetouw R., Watson D. F., Babler B. L., Benjamin R., Gomez M., Meade M. R., Povich M. S., Robitaille T. P., Watson C., A Catalog of Extended Green Objects in the GLIMPSE Survey: A New Sample of Massive Young Stellar Object Outflow Candidates, AJ, 2008, vol. 136, p. 2391

Dabrowski I., The Lyman and Werner bands of H2, Canadian Journal of Physics, 1984, vol. 62, p. 1639

Damineli A., Almeida L. A., Blum R. D., Damineli D. S. C., Navarete F., Rubinho M. S., Teodoro M., Extinction law in the range 0.4-4.8 μm and the 8620 Å DIB towards the stellar cluster Westerlund 1, MNRAS, 2016, vol. 463, p. 2653

Davies B., Lumsden S. L., Hoare M. G., Oudmaijer R. D., de Wit W.-J., The circumstellar disc, envelope and bipolar outflow of the massive young stellar object W33A, MNRAS, 2010, vol. 402, p. 1504

Decarli R., Walter F., Aravena M., Carilli C., Bouwens R., et al., The ALMA Spectroscopic Survey in the Hubble Ultra Deep Field: Molecular Gas Reservoirs in High-redshift Galaxies, ApJ, 2016, vol. 833, p. 70

Di Francesco J., Johnstone D., Kirk H., MacKenzie T., Ledwosinska E., The SCUBA Legacy Catalogues: Submillimeter-Continuum Objects Detected by SCUBA, ApJS, 2008, vol. 175, p. 277

Drew J. E., Proga D., Stone J. M., A radiation-driven disc wind model for massive young stellar objects, MNRAS, 1998, vol. 296, p. L6

Edris K. A., Fuller G. A., Cohen R. J., A survey of OH masers towards high mass protostellar objects, A&A, 2007, vol. 465, p. 865

Egan M. P., et al., MSX6C Infrared Point Source Catalog. The Midcourse Space Experiment Point Source Catalog Version 2.3 (October 2003), VizieR Online Data Catalog, 2003, vol. 5114, p. 0

Egan M. P., Price S. D., Shipman R. F., Gugliotti G. M., Tedesco E. F., Moshir M., Cohen M., The MSX Infrared Point Source Catalog, Version 1. 0. In Astrophysics with Infrared Surveys: A Prelude to SIRTF , vol. 177 of Astronomical Society of the Pacific Conference Series, 1999, p. 404

Fazio G. G., Hora J. L., Allen L. E., Ashby M. L. N., Barmby P., et al., The Infrared Array Camera (IRAC) for the Spitzer Space Telescope, ApJS, 2004, vol. 154, p. 10

Folha D. F. M., Continuum Excess Emission in Young Low Mass Stars, Ap&SS, 2003, vol. 287, p. 41

Giannetti A., Leurini S., Wyrowski F., Urquhart J., Csengeri T., Menten K. M., König C., Güsten R., ATLASGAL-selected massive clumps in the inner Galaxy: VI. Temperature structure and evolution, A&A, 2017

Giannetti A., Wyrowski F., Brand J., Csengeri T., Fontani F., Walmsley C. M., Nguyen Luong Q., Beuther H., Schuller F., Güsten R., Menten K. M., ATLASGAL-selected massive clumps in the inner Galaxy. I. CO depletion and isotopic ratios, A&A, 2014, vol. 570, p. A65

Goodson A. P., Böhm K.-H., Winglee R. M., Jets from Accreting Magnetic Young Stellar Objects. I. Comparison of Observations and High-Resolution Simulation Results, ApJ, 1999, vol. 524, p. 142

Goodson A. P., Winglee R. M., Boehm K.-H., Time-dependent Accretion by Magnetic Young Stellar Objects as a Launching Mechanism for Stellar Jets, ApJ, 1997, vol. 489, p. 199

Güsten R., Baryshev A., Bell A., Belloche A., et al., Submillimeter heterodyne arrays for APEX. In Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series , vol. 7020 of Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 2008, p. 10

Hanson M. M., Kudritzki R.-P., Kenworthy M. A., Puls J., Tokunaga A. T., A Medium Resolution Near-Infrared Spectral Atlas of O and Early-B Stars, ApJS, 2005, vol. 161, p. 154

Ilee J. D., Fairlamb J., Oudmaijer R. D., Mendigutía I., van den Ancker M. E., Kraus S., Wheelwright H. E., Investigating the inner discs of Herbig Ae/Be stars with CO bandhead and Brγ emission, MNRAS, 2014, vol. 445, p. 3723

Ilee J. D., Wheelwright H. E., Oudmaijer R. D., de Wit W. J., Maud L. T., Hoare M. G., Lumsden S. L., Moore T. J. T., Urquhart J. S., Mottram J. C., CO bandhead emission of massive young stellar objects: determining disc properties, MNRAS, 2013, vol. 429, p. 2960

Johnston K. G., Robitaille T. P., Beuther H., Linz H., Boley P., Kuiper R., Keto E., Hoare M. G., van Boekel R., A Keplerian-like Disk around the Forming O-type Star AFGL 4176, ApJ, 2015, vol. 813, p. L19

Karska A., Herpin F., Bruderer S., Goicoechea J. R., et al., Far-infrared molecular lines from low- to high-mass star forming regions observed with Herschel, A&A, 2014, vol. 562, p. A45

Kasemann C., Güsten R., Heyminck S., Klein B., Klein T., Philipp S. D., Korn A., Schneider G., Henseler A., Baryshev A., Klapwijk T. M., CHAMP $^+$: a powerful array receiver

for APEX. In Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series , vol. 6275 of Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 2006, p. 0

Kauffmann J., Bertoldi F., Bourke T. L., Evans II N. J., Lee C. W., MAMBO mapping of Spitzer c2d small clouds and cores, A&A, 2008, vol. 487, p. 993

Klein T., Ciechanowicz M., Leinz C., Heyminck S., Guesten R., C. K., et al., Flash+ -A Dual Channel Wide-Band Spectrometer for APEX, IEEE Trans. on Terahertz Science and Technology, 2014, vol. 4, p. 588

König C., Urquhart J. S., Csengeri T., Leurini S., Wyrowski F., Giannetti A., Wienen M., Pillai T., Kauffmann J., Menten K. M., Schuller F., ATLASGAL-selected massive clumps in the inner Galaxy. III. Dust continuum characterization of an evolutionary sample, A&A, 2017, vol. 599, p. A139

Kraus M., Krügel E., Thum C., Geballe T. R., CO band emission from MWC 349. I. First overtone bands from a disk or from a wind?, A&A, 2000, vol. 362, p. 158

Kraus S., Hofmann K.-H., Menten K. M., Schertl D., Weigelt G., Wyrowski F., Meilland A., Perraut K., Petrov R., Robbe-Dubois S., Schilke P., Testi L., A hot compact dust disk around a massive young stellar object, Nature, 2010, vol. 466, p. 339

Kroupa P., The initial mass function of simple and composite stellar populations, ArXiv Astrophysics e-prints, 2007

Krumholz M. R., The big problems in star formation: The star formation rate, stellar clustering, and the initial mass function, Phys. Rep., 2014, vol. 539, p. 49

Krumholz M. R., Klein R. I., McKee C. F., Offner S. S. R., Cunningham A. J., The Formation of Massive Star Systems by Accretion, Science, 2009, vol. 323, p. 754

Kuiper R., Klahr H., Beuther H., Henning T., Radiation pressure feedback in the formation of massive stars, Bulletin de la Societe Royale des Sciences de Liege, 2011a, vol. 80, p. 211

Kuiper R., Klahr H., Beuther H., Henning T., Three-dimensional simulation of massive star formation in the disk accretion scenario, ArXiv e-prints, 2011b

Kuiper R., Yorke H. W., Turner N. J., Protostellar Outflows and Radiative Feedback from Massive Stars, ApJ, 2015, vol. 800, p. 86

Kulesa C. A., Molecular hydrogen and its ions in dark interstellar clouds and star forming regions, THE UNIVERSITY OF ARIZONA, 2002, Ph.D. Thesis

Lada C. J., Cold outflows, energetic winds, and enigmatic jets around young stellar objects, ARA&A, 1985, vol. 23, p. 267

Lee C.-F., Ho P. T. P., Li Z.-Y., Hirano N., Zhang Q., Shang H., A rotating protostellar jet launched from the innermost disk of HH 212, Nature Astronomy, 2017, vol. 1, p. 0152

Leurini S., Codella C., López-Sepulcre A., Gusdorf A., Csengeri T., Anderl S., SiO excitation from dense shocks in the earliest stages of massive star formation, A&A, 2014, vol. 570, p. A49

Leurini S., Wyrowski F., Herpin F., van der Tak F., Güsten R., van Dishoeck E. F., The distribution of warm gas in the G327.3-0.6 massive star-forming region, A&A, 2013, vol. 550, p. A10

Lumsden S. L., Hoare M. G., Oudmaijer R. D., Richards D., The population of the Galactic plane as seen by MSX, MNRAS, 2002, vol. 336, p. 621

Lumsden S. L., Hoare M. G., Urquhart J. S., Oudmaijer R. D., Davies B., Mottram J. C., Cooper H. D. B., Moore T. J. T., The Red MSX Source Survey: The Massive Young Stellar Population of Our Galaxy, ApJS, 2013, vol. 208, p. 11

Lumsden S. L., Wheelwright H. E., Hoare M. G., Oudmaijer R. D., Drew J. E., Tracers of discs and winds around intermediate- and high-mass young stellar objects, MNRAS, 2012, vol. 424, p. 1088

Marr J. M., Snell R. L., Kurtz S. E., Fundamentals of Radio Astronomy: Observational Methods. CRC Press, 2015

Martin S. C., The Thermal Structure of Magnetic Accretion Funnels in Young Stellar Objects, ApJ, 1996, vol. 470, p. 537

Maud L. T., Moore T. J. T., Lumsden S. L., Mottram J. C., Urquhart J. S., Hoare M. G., A distance-limited sample of massive molecular outflows, MNRAS, 2015, vol. 453, p. 645

McGregor P. J., et al., Gemini near-infrared integral field spectrograph (NIFS). In Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series , vol. 4841 of Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 2003, p. 1581

Megeath S. T., Guthermuth R. A., Bronfman L., Pipher J. L., Wilson T. L., The Stellar Density in High Mass Star Forming Regions. In IAU Symposium , vol. 221 of IAU Symposium, 2003

Menezes R. B., Steiner J. E., Ricci T. V., A treatment procedure for Gemini North/NIFS data cubes: application to NGC 4151, MNRAS, 2014a, vol. 438, p. 2597

Menezes R. B., Steiner J. E., Ricci T. V., An Off-centered Active Galactic Nucleus in NGC 3115, ApJ, 2014b, vol. 796, p. L13

Molinari S., Pezzuto S., Cesaroni R., Brand J., Faustini F., Testi L., The evolution of the spectral energy distribution in massive young stellar objects, A&A, 2008, vol. 481, p. 345

Molinari S., Schisano E., Elia D., Pestalozzi M., et al., Hi-GAL, the Herschel infrared Galactic Plane Survey: photometric maps and compact source catalogues. First data release for the inner Milky Way: $+68^\circ \geq \ell \geq -70^\circ$, A&A, 2016, vol. 591, p. A149

Molinari S., Swinyard B., Bally J., Barlow M., Bernard J.-P., Martin P., Moore T., Noriega-Crespo A., Plume R., Testi L., et al., Hi-GAL: The Herschel Infrared Galactic Plane Survey, PASP, 2010, vol. 122, p. 314

Mottram J. C., et al., The Red MSX Source survey: the bolometric fluxes and luminosity distributions of young massive stars, A&A, 2011, vol. 525, p. A149+

Mottram J. C., Hoare M. G., Lumsden S. L., Oudmaijer R. D., Urquhart J. S., Meade M. R., Moore T. J. T., Stead J. J., The RMS survey: far-infrared photometry of young massive stars, AAp, 2010, vol. 510, p. A89+

Mottram J. C., Hoare M. G., Lumsden S. L., Oudmaijer R. D., Urquhart J. S., Sheret T. L., Clarke A. J., Allsopp J., The RMS survey: mid-infrared observations of candidate massive YSOs in the southern hemisphere, A&A, 2007, vol. 476, p. 1019

Murakawa K., Lumsden S. L., Oudmaijer R. D., Davies B., Wheelwright H. E., Hoare M. G., Ilee J. D., Near-infrared integral field spectroscopy of massive young stellar objects, MNRAS, 2013, vol. 436, p. 511

Muzerolle J., Calvet N., Hartmann L., Magnetospheric Accretion Models for the Hydrogen Emission Lines of T Tauri Stars, ApJ, 1998, vol. 492, p. 743

Navarete F., Damineli A., Barbosa C. L., Blum R. D., A survey of extended H_2 emission from massive YSOs, MNRAS, 2015, vol. 450, p. 4364

Navarete F., Leurini S., Giannetti A., Wyrowski F., Urquhart J. S., König C., Csengeri T., Güsten R., Damineli A., Menten K. M., ATLASGAL-selected high-mass clumps in the inner Galaxy.
VI. Characterisation of mid-J CO emission, submitted, submitted

Navarete F., Leurini S., Wyrowski F., Giannetti A., Urquhart J. S., König C., Csengeri T., Güsten R., Damineli A., Menten K. M., ATLASGAL-selected high-mass clumps in the inner Galaxy.
VII. Mid-J CO outflow content, A&A, forthcoming, forthcoming

Nishiyama S., Nagata T., Tamura M., Kandori R., Hatano H., Sato S., Sugitani K., The Interstellar Extinction Law toward the Galactic Center. II. V, J, H, and K_s Bands, ApJ, 2008, vol. 680, p. 1174

Palla F., Stahler S. W., The Pre-Main-Sequence Evolution of Intermediate-Mass Stars, ApJ, 1993, vol. 418, p. 414

Panagia N., Some Physical parameters of early-type stars, AJ, 1973, vol. 78, p. 929

Patel N. A., Curiel S., Sridharan T. K., Zhang Q., Hunter T. R., Ho P. T. P., Torrelles J. M., Moran J. M., Gómez J. F., Anglada G., A disk of dust and molecular gas around a high-mass protostar, Nature, 2005, vol. 437, p. 109

Pestalozzi M. R., Minier V., Booth R. S., A general catalogue of 6.7-GHz methanol masers. I. Data., A&A, 2005, vol. 432, p. 737

Peters T., Klaassen P. D., Mac Low M.-M., Schrön M., Federrath C., Smith M. D., Klessen R. S., Collective Outflow from a Small Multiple Stellar System, ApJ, 2014, vol. 788, p. 14

Porter J. M., Drew J. E., Lumsden S. L., Broad band infrared spectroscopy of massive young stellar objects, A&A, 1998, vol. 332, p. 999

Price S. D., Egan M. P., Carey S. J., Mizuno D. R., Kuchar T. A., Midcourse Space Experiment Survey of the Galactic Plane, AJ, 2001, vol. 121, p. 2819

Pudritz R. E., Rogers C. S., Ouyed R., Controlling the collimation and rotation of hydromagnetic disc winds, MNRAS, 2006, vol. 365, p. 1131

Purcell C. R., Hoare M. G., Cotton W. D., Lumsden S. L., Urquhart J. S., Chandler C., Churchwell E. B., Diamond P., Dougherty S. M., Fender R. P., Fuller G., Garrington S. T., Gledhill T. M., Goldsmith P. F., Hindson L., et al., The Coordinated Radio and Infrared Survey for High-mass Star Formation. II. Source Catalog, ApJS, 2013, vol. 205, p. 1

Reipurth B., Jewitt D., Keil K., Protostars and Planets V, Protostars and Planets V, 2007

Reiter M., Smith N., Bally J., [Fe II] jets from intermediate-mass protostars in Carina, MNRAS, 2016, vol. 463, p. 4344

Ricci T. V., Steiner J. E., May D., Garcia-Rissmann A., Menezes R. B., Optical and near-infrared IFU spectroscopy of the nuclear region of the AGN-starburst galaxy NGC 7582, MNRAS, 2018, vol. 473, p. 5334

Ricci T. V., Steiner J. E., Menezes R. B., Integral field unit spectroscopy of 10 early-type galactic nuclei - I. Principal component analysis Tomography and nuclear activity, MNRAS, 2014, vol. 440, p. 2419

Richer J. S., Shepherd D. S., Cabrit S., Bachiller R., Churchwell E., Molecular Outflows from Young Stellar Objects, Protostars and Planets IV, 2000, pp 867–+

Rieke G. H., Lebofsky M. J., The interstellar extinction law from 1 to 13 microns, ApJ, 1985, vol. 288, p. 618

Robitaille T. P., Whitney B. A., The Present-Day Star Formation Rate of the Milky Way Determined from Spitzer-Detected Young Stellar Objects, ApJ, 2010, vol. 710, p. L11

Salpeter E. E., The Luminosity Function and Stellar Evolution., ApJ, 1955, vol. 121, p. 161

San José-García I., Mottram J. C., Kristensen L. E., van Dishoeck E. F., Yıldız U. A., van der Tak F. F. S., Herpin F., Visser R., McCoey C., Wyrowski F., Braine J., Johnstone D., Herschel-HIFI observations of high-J CO and isotopologues in star-forming regions: from low to high mass, A&A, 2013, vol. 553, p. A125

San José-García I., Mottram J. C., van Dishoeck E. F., Kristensen L. E., van der Tak F. F. S., Braine J., Herpin F., Johnstone D., van Kempen T. A., Wyrowski F., Linking low- to high-mass young stellar objects with Herschel-HIFI observations of water, A&A, 2016, vol. 585, p. A103

Saraceno P., Andre P., Ceccarelli C., Griffin M., Molinari S., An evolutionary diagram for young stellar objects., A&A, 1996, vol. 309, p. 827

Schneider N., Csengeri T., Klessen R. S., Tremblin P., Ossenkopf V., Peretto N., Simon R., Bontemps S., Federrath C., Understanding star formation in molecular clouds. II. Signatures of gravitational collapse of IRDCs, A&A, 2015, vol. 578, p. A29

Schuller F., Menten K. M., Contreras Y., Wyrowski F., et al., ATLASGAL - The APEX telescope large area survey of the galaxy at 870 μm, A&A, 2009, vol. 504, p. 415

Scoville N., Kleinmann S. G., Hall D. N. B., Ridgway S. T., The circumstellar and nebular environment of the Becklin-Neugebauer object - 2-5 micron wavelength spectroscopy, ApJ, 1983, vol. 275, p. 201

Shepherd D. S., Yu K. C., Bally J., Testi L., The Molecular Outflow and Possible Precessing Jet from the Massive Young Stellar Object IRAS 20126+4104, ApJ, 2000, vol. 535, p. 833

Shu F. H., Adams F. C., Lizano S., Star formation in molecular clouds - Observation and theory, ARA&A, 1987, vol. 25, p. 23

Sicilia-Aguilar A., Oprandi A., Froebrich D., Fang M., Prieto J. L., Stanek K., Scholz A., Kochanek C. S., Henning T., Gredel R., Holoien T. W.-S., Rabus M., Shappee B. J., Billington S. J., Campbell-White J., Zegmott T. J., The 2014-2017 outburst of the young star ASASSN-13db: A time-resolved picture of a very low-mass star between EXors and FUors, ArXiv e-prints, 2017

Siringo G., Kreysa E., Kovács A., Schuller F., Weiß A., Esch W., Gemünd H.-P., Jethava N., Lundershausen G., Colin A., Güsten R., Menten K. M., Beelen A., Bertoldi F., Beeman J. W., Haller E. E., The Large APEX BOlometer CAmera LABOCA, A&A, 2009, vol. 497, p. 945

Skrutskie M. F., Cutri R. M., Stiening R., Weinberg M. D., et al., The Two Micron All Sky Survey (2MASS), AJ, 2006, vol. 131, p. 1163

Steiner J. E., Menezes R. B., Ricci T. V., Oliveira A. S., PCA Tomography: how to extract information from data cubes, MNRAS, 2009, vol. 395, p. 64

Thompson R. I., Boroson T. A., The infrared emission lines from IRC +10420., ApJ, 1977, vol. 216, p. L75

Tokunaga A. T., , 2000 Infrared Astronomy. p. 143

Tokunaga A. T., Simons D. A., Vacca W. D., The Mauna Kea Observatories Near-Infrared Filter Set. II. Specifications for a New JHKL'M' Filter Set for Infrared Astronomy, PASP, 2002, vol. 114, p. 180

Turner J., Kirby-Docken K., Dalgarno A., The Quadrupole Vibration-Rotation Transition Probabilities of Molecular Hydrogen, ApJS, 1977, vol. 35, p. 281

Urquhart J. S., Busfield A. L., Hoare M. G., Lumsden S. L., Clarke A. J., Moore T. J. T., Mottram J. C., Oudmaijer R. D., The RMS survey. Radio observations of candidate massive YSOs in the southern hemisphere, A&A, 2007, vol. 461, p. 11

Urquhart J. S., et al., The RMS survey. ^{13}CO observations of candidate massive YSOs in the southern Galactic plane, A&A, 2007, vol. 474, p. 891

Urquhart J. S., et al., The RMS survey. ^{13}CO observations of candidate massive YSOs in the northern Galactic plane, A&A, 2008, vol. 487, p. 253

Urquhart J. S., et al., The RMS survey. 6 cm continuum VLA observations towards candidate massive YSOs in the northern hemisphere, A&A, 2009a, vol. 501, p. 539

Urquhart J. S., et al., The RMS survey. H_2O masers towards a sample of southern hemisphere massive YSO candidates and ultra compact HII regions, A&A, 2009b, vol. 507, p. 795

Urquhart J. S., Hoare M. G., Lumsden S. L., Oudmaijer R. D., Moore T. J. T., The RMS Survey: A Galaxy-wide Sample of Massive Young Stellar Objects. In Massive Star Formation: Observations Confront Theory , vol. 387 of Astronomical Society of the Pacific Conference Series, 2008, p. 381

Urquhart J. S., König C., Giannetti A., Leurini S., Moore T. J. T., Eden D. J., Pillai T., Thompson M. A., Braiding C., Burton M. G., Csengeri T., Dempsey J. T., Figura C., Froebrich D., Menten K. M., Schuller F., Smith M. D., Wyrowski F., ATLASGAL - properties of a complete sample of Galactic clumps, MNRAS, 2018, vol. 473, p. 1059

Urquhart J. S., Moore T. J. T., Csengeri T., Wyrowski F., Schuller F., Hoare M. G., Lumsden S. L., Mottram J. C., Thompson M. A., Menten K. M., Walmsley C. M., Bronfman L., Pfalzner S., König C., Wienen M., ATLASGAL - towards a complete sample of massive star forming clumps, MNRAS, 2014, vol. 443, p. 1555

Urquhart J. S., Moore T. J. T., Schuller F., Wyrowski F., Menten K. M., Thompson M. A., Csengeri T., Walmsley C. M., Bronfman L., König C., ATLASGAL - environments of 6.7 GHz methanol masers, MNRAS, 2013, vol. 431, p. 1752

Urquhart J. S., Morgan L. K., Figura C. C., Moore T. J. T., Lumsden S. L., Hoare M. G., et al., The Red MSX Source survey: ammonia and water maser analysis of massive star-forming regions, MNRAS, 2011, vol. 418, p. 1689

Vacca W. D., Garmany C. D., Shull J. M., The Lyman-Continuum Fluxes and Stellar Parameters of O and Early B-Type Stars, ApJ, 1996, vol. 460, p. 914

Vaidya B., Fendt C., Beuther H., Porth O., Jet Formation from Massive Young Stars: Magnetohydrodynamics versus Radiation Pressure, ApJ, 2011, vol. 742, p. 56

van Kempen T. A., Kristensen L. E., Herczeg G. J., et al., Origin of the hot gas in low-mass protostars. Herschel-PACS spectroscopy of HH 46, A&A, 2010, vol. 518, p. L121

van Kempen T. A., van Dishoeck E. F., Güsten R., Kristensen L. E., Schilke P., Hogerheijde M. R., Boland W., Menten K. M., Wyrowski F., APEX-CHAMP$^+$ high-J CO observations of low-mass young stellar objects. II. Distribution and origin of warm molecular gas, A&A, 2009, vol. 507, p. 1425

Varricatt W. P., Davis C. J., Ramsay S., Todd S. P., A near-IR imaging survey of intermediate- and high-mass young stellar outflow candidates, MNRAS, 2010a, vol. 404, p. 661

Varricatt W. P., Davis C. J., Ramsay S., Todd S. P., A near-IR imaging survey of intermediate- and high-mass young stellar outflow candidates, MNRAS, 2010b, vol. 404, p. 661

Visser R., Kristensen L. E., Bruderer S., van Dishoeck E. F., Herczeg G. J., Brinch C., Doty S. D., Harsono D., Wolfire M. G., Modelling Herschel observations of hot molecular gas emission from embedded low-mass protostars, A&A, 2012, vol. 537, p. A55

Wang L.-Y., Shang H., Su Y.-N., Santiago-García J., Tafalla M., Zhang Q., Hirano N., Lee C.-F., Molecular Jet of IRAS 04166+2706, ApJ, 2014, vol. 780, p. 49

Weiß A., Downes D., Neri R., Walter F., Henkel C., Wilner D. J., Wagg J., Wiklind T., Highly-excited CO emission in APM 08279+5255 at z = 3.9, A&A, 2007, vol. 467, p. 955

Wheelwright H. E., Oudmaijer R. D., de Wit W. J., Hoare M. G., Lumsden S. L., Urquhart J. S., Probing discs around massive young stellar objects with CO first overtone emission, MNRAS, 2010, vol. 408, p. 1840

Wienen M., Wyrowski F., Schuller F., Menten K. M., Walmsley C. M., Bronfman L., Motte F., Ammonia from cold high-mass clumps discovered in the inner Galactic disk by the ATLASGAL survey, A&A, 2012, vol. 544, p. A146

Wu Y., Wei Y., Zhao M., Shi Y., Yu W., Qin S., Huang M., A study of high velocity molecular outflows with an up-to-date sample, A&A, 2004, vol. 426, p. 503

Yıldız U. A., Kristensen L. E., van Dishoeck E. F., Belloche A., van Kempen T. A., Hogerheijde M. R., Güsten R., van der Marel N., APEX-CHAMP$^+$ high-J CO observations of low-mass young stellar objects. III. NGC 1333 IRAS 4A/4B envelope, outflow, and ultraviolet heating, A&A, 2012, vol. 542, p. A86

Yıldız U. A., Kristensen L. E., van Dishoeck E. F., Hogerheijde M. R., Karska A., Belloche A., Endo A., et al., APEX-CHAMP$^+$ high-J CO observations of low-mass young stellar objects. IV. Mechanical and radiative feedback, A&A, 2015, vol. 576, p. A109

Yorke H. W., Sonnhalter C., On the Formation of Massive Stars, ApJ, 2002, vol. 569, p. 846

Zhang Q., Ho P. T. P., Ohashi N., Dynamical Collapse in W51 Massive Cores: CS (3-2) and CH_3CN Observations, ApJ, 1998, vol. 494, p. 636

Appendix

Appendix A

Sub-millimetric observations

A.1 Brief description of the output files from the `IDL` *routine* `bayesian_fit`

The code reads the input data and its errors, and fits the best linear models for combinations of the points, given by $(x \pm \sigma_x, y \pm \sigma_y)$. As a result, a FITS file containing the map of (α, δ) coefficients is created (see Fig. A.1) and its header (see Fig. A.2) presents a complete description of the fit. Each of the coefficients are registered as `ALPHA`, `BETA` and `EPSILON`, followed by the upper and lower errors (`SIG_A_PL` and `SIG_A_MN`; and `SIG_B_PL` and `SIG_B_MN`). The number of iterations is given in `NITER`. The construction of the α and β axes is given by:

```
ALPHA[1:NAXIS1] = CRVAL1 + CDELT1 * [1:NAXIS1]
 BETA[1:NAXIS1] = CRVAL2 + CDELT2 * [1:NAXIS1]
```

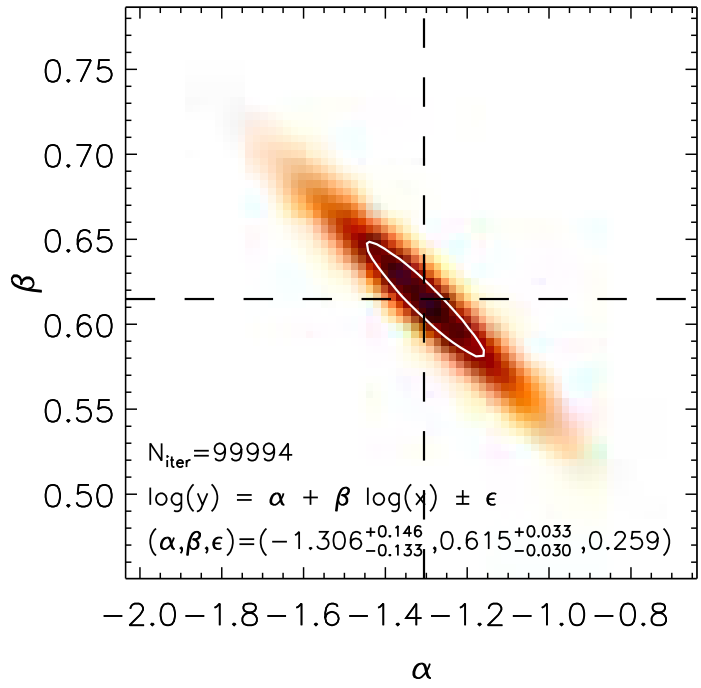


Figure A.1: Example of the distribution of (α,β) coefficients from the Bayesian fit model. The uncertainties on the coefficients were estimated in the region with population above or equal to 68% of the peak counts (shown as white contour). The number of iterations and the values of each coefficient are shown in the left bottom side of the plot.

```
SIMPLE  =                    T /image conforms to FITS standard
BITPIX  =                   16
NAXIS   =                    2
NAXIS1  =                   76
NAXIS2  =                   76
EXTEND  =                    T /file may contain extensions
ORIGIN  = 'IDL/Bayesian_fitting.pro' /FITS file originator
CTYPE1  = 'ALPHA   '           /Values for the intercept coefficient
CRVAL1  =       -2.03477333805 /
CRPIX1  =                    1 /Ref pix of axis 1
CDELT1  =      0.0186125891675 /
CTYPE2  = 'BETA    '           /Values for the slope coefficient
CRVAL2  =       0.450343174640 /
CRPIX2  =                    1 /Ref pix of axis 2
CDELT2  =     0.00448220676705 /
ALPHA   =       -1.30567972405 /Values for the intercept coefficient
DELTA   =       0.614783643740 /Values for the slope coefficient
EPSILON =       0.258848500125 /Values for the intrinsic scatter of the fit
SIG_A_PL=             0.145698 /Upper error on ALPHA
SIG_A_MN=             0.133491 /Lower error on ALPHA
SIG_B_PL=            0.0327766 /Upper error on BETA
SIG_B_MN=            0.0299743 /Lower error on BETA
NITER   =              99994.0 /Number of iterations for the fit
BSCALE  =                    1
BZERO   =                    0
END
```

Figure A.2: Example of a header of an output file from the Bayesian fit code.

A.2 Full tables

Table A.1 - Properties of the TOP100 sources. Reproduced from Navarete et al. (submitted).

ID	Source	RA(J2000) (HH:MM:SS)	DEC(J2000) (DD:MM:SS)	$V_{\rm lsr}$ (km s^{-1})	d (kpc)	$\log\left(\frac{L_{\rm bol}}{L_\odot}\right)$	$\log\left(\frac{M_{\rm clump}}{M_\odot}\right)$	Class
1	AGAL008.684−00.367	18:06:23.35	−21:37:05.2	38	4.78	4.44	3.17	IRw
2	AGAL008.706−00.414	18:06:36.72	−21:37:18.6	38.9	4.78	2.70	3.22	IRw
3	AGAL010.444−00.017	18:08:44.59	−19:54:36.4	73.1	8.55	4.05	3.21	IRw
4	AGAL010.472+00.027	18:08:38.24	−19:51:51.5	67	8.55	5.67	4.02	H II
5	AGAL010.624−00.384	18:10:28.82	−19:55:48.4	−2.0	4.95	5.63	3.58	H II
6	AGAL012.804−00.199	18:14:13.34	−17:55:45.3	35.4	2.40	5.39	3.27	H II
7	AGAL013.178+00.059	18:14:01.32	−17:28:38.6	49.2	2.40	3.92	2.57	70w
8	AGAL013.658−00.599	18:17:24.37	−17:22:08.0	47.4	4.47	4.32	2.76	IRb
9	AGAL014.114−00.574	18:18:13.14	−16:57:20.0	19.1	2.57	3.50	2.55	IRw
10	AGAL014.194−00.194	18:16:58.77	−16:42:17.5	40.4	3.90	3.43	2.91	IRw
11	AGAL014.492−00.139	18:17:22.09	−16:24:59.4	41.7	3.87	2.88	3.28	70w
12	AGAL014.632−00.577	18:19:15.22	−16:30:02.3	16.8	1.83	3.44	2.40	IRw
13	AGAL015.029−00.669	18:20:22.37	−16:11:38.8	19.4	1.98	5.13	3.08	IRb
14	AGAL018.606−00.074	18:25:08.27	−12:45:22.8	43.8	4.33	2.77	2.94	IRw
15	AGAL018.734−00.226	18:25:56.05	−12:42:50.8	39.5	12.48	4.86	3.90	IRw
16	AGAL018.888−00.474	18:27:07.69	−12:41:36.4	65.1	4.74	3.51	3.45	IRw
17	AGAL019.882−00.534	18:29:14.20	−11:50:28.4	42.9	3.66	4.09	2.90	IRb

Table A.1 - continued.

ID	Source	RA(J2000) (HH:MM:SS)	DEC(J2000) (DD:MM:SS)	$V_{\rm lsr}$ (km s^{-1})	d (kpc)	$\log\left(\frac{L_{\rm bol}}{L_\odot}\right)$	$\log\left(\frac{M_{\rm clump}}{M_\odot}\right)$	Class
18	AGAL022.376+00.447	18:30:24.26	−09:10:36.0	51.5	3.96	2.50	2.80	IRw
19	AGAL023.206−00.377	18:34:54.99	−08:49:15.5	77.4	4.59	4.10	3.11	IRw
20	AGAL024.629+00.172	18:35:35.61	−07:18:17.9	115.8	7.72	3.70	3.18	IRw
21	AGAL028.564−00.236	18:44:18.04	−03:59:40.0	87.5	5.45	3.25	3.73	IRw
22	AGAL028.861+00.066	18:43:46.06	−03:35:32.2	105.4	7.41	5.21	3.03	IRb
23	AGAL030.818−00.056	18:47:46.82	−01:54:24.6	99	4.90	4.80	3.75	IRb
24	AGAL030.848−00.081	18:47:55.45	−01:53:29.5	97.9	4.90	3.49	3.08	70w
25	AGAL030.893+00.139	18:47:13.37	−01:45:04.0	105.6	4.90	2.70	3.28	70w
26	AGAL031.412+00.307	18:47:34.33	−01:12:45.5	98	4.90	4.84	3.49	H II
27	AGAL034.258+00.154	18:53:18.52	01:15:01.5	58	1.56	4.68	2.91	H II
28	AGAL034.401+00.226	18:53:18.79	01:24:37.8	57	1.56	3.48	2.44	H II
29	AGAL034.411+00.234	18:53:18.18	01:25:23.0	57	1.56	3.68	2.33	IRb
30	AGAL034.821+00.351	18:53:38.07	01:50:28.6	57.7	1.56	3.44	2.05	IRb
31	AGAL035.197−00.742	18:58:12.79	01:40:37.6	33.9	2.19	4.37	2.67	IRb
32	AGAL037.554+00.201	18:59:09.95	04:12:16.3	84	6.71	4.71	3.10	IRb
33	AGAL043.166+00.011	19:10:13.40	09:06:11.2	7	11.11	6.58	4.64	H II
34	AGAL049.489−00.389	19:23:43.70	14:30:32.0	57.9	5.41	5.75	4.07	H II

Table A.1 - continued.

ID	Source	RA(J2000) (HH:MM:SS)	DEC(J2000) (DD:MM:SS)	$V_{\rm lsr}$ (km s^{-1})	d (kpc)	$\log\left(\frac{L_{\rm bol}}{L_\odot}\right)$	$\log\left(\frac{M_{\rm clump}}{M_\odot}\right)$	Class
35	AGAL053.141+00.069	19:29:17.80	17:56:17.9	22	1.60	3.36	1.98	IRb
36	AGAL059.782+00.066	19:43:10.83	23:44:03.3	22	2.16	3.99	2.41	IRb
37	AGAL301.136−00.226	12:35:34.89	−63:02:32.8	−37.0	4.40	5.33	3.29	H II
38	AGAL305.192−00.006	13:11:13.73	−62:47:27.1	−33.5	3.80	4.10	2.71	IRw
39	AGAL305.209+00.206	13:11:14.09	−62:34:41.5	−42.6	3.80	4.95	3.15	IRb
40	AGAL305.562+00.014	13:14:26.24	−62:44:24.9	−39.8	3.80	4.71	2.61	IRb
41	AGAL305.794−00.096	13:16:32.84	−62:49:41.3	−40.7	3.80	2.99	2.77	70w
42	AGAL309.384−00.134	13:47:24.03	−62:18:09.5	−49.5	5.34	4.20	3.08	IRb
43	AGAL310.014+00.387	13:51:38.36	−61:39:15.1	−41.6	3.61	4.70	2.62	IRb
44	AGAL313.576+00.324	14:20:08.34	−60:42:05.4	−45.5	3.78	3.97	2.26	IRb
45	AGAL316.641−00.087	14:44:18.77	−59:55:11.5	−19.1	1.19	3.00	1.26	IRb
46	AGAL317.867−00.151	14:53:16.60	−59:26:33.9	−40.3	2.95	3.22	2.56	IRw
47	AGAL318.779−00.137	14:59:33.06	−59:00:33.5	−38.5	2.78	3.81	2.56	IRw
48	AGAL320.881−00.397	15:14:33.18	−58:11:28.5	−46.0	9.97	3.78	3.46	70w
49	AGAL326.661+00.519	15:45:02.97	−54:09:10.7	−39.6	1.82	3.87	2.10	IRb
50	AGAL326.987−00.032	15:49:08.16	−54:23:06.9	−58.2	3.95	3.06	2.65	IRw
51	AGAL327.119+00.509	15:47:32.98	−53:52:45.0	−83.0	5.51	4.77	2.83	IRb

Table A.1 - continued.

ID	Source	RA(J2000) (HH:MM:SS)	DEC(J2000) (DD:MM:SS)	$V_{\rm lsr}$ (km s^{-1})	d (kpc)	$\log\left(\frac{L_{\rm bol}}{L_\odot}\right)$	$\log\left(\frac{M_{\rm clump}}{M_\odot}\right)$	Class
52	AGAL327.393+00.199	15:50:18.99	−53:57:04.5	−88.7	5.92	4.13	3.07	IRb
53	AGAL329.029−00.206	16:00:31.96	−53:12:54.0	−43.5	11.49	5.33	4.06	IRw
54	AGAL329.066−00.307	16:01:09.69	−53:16:01.6	−42.2	11.58	4.85	3.96	IRb
55	AGAL330.879−00.367	16:10:20.12	−52:06:06.4	−63.6	4.16	5.19	3.20	H II
56	AGAL330.954−00.182	16:09:52.67	−51:54:53.9	−90.6	9.32	6.12	4.24	H II
57	AGAL331.709+00.582	16:10:06.79	−50:50:29.4	−67.1	10.53	4.57	3.71	IRw
58	AGAL332.094−00.421	16:16:16.66	−51:18:25.3	−56.5	3.60	4.77	2.80	IRb
59	AGAL332.826−00.549	16:20:11.02	−50:53:14.5	−57.3	3.60	5.38	3.29	H II
60	AGAL333.134−00.431	16:21:02.44	−50:35:12.2	−51.9	3.60	5.62	3.46	H II
61	AGAL333.284−00.387	16:21:30.83	−50:26:58.9	−51.5	3.60	5.11	3.32	H II
62	AGAL333.314+00.106	16:19:28.56	−50:04:42.1	−45.8	3.60	4.03	2.63	IRb
63	AGAL333.604−00.212	16:22:09.35	−50:05:58.8	−47.1	3.60	6.09	3.54	H II
64	AGAL333.656+00.059	16:21:11.61	−49:52:16.3	−84.8	5.29	3.63	3.15	70w
65	AGAL335.789+00.174	16:29:47.34	−48:15:52.1	−49.5	3.67	4.31	3.04	IRw
66	AGAL336.958−00.224	16:36:17.12	−47:40:44.3	−73.1	10.91	3.56	3.38	IRw
67	AGAL337.176−00.032	16:36:18.78	−47:23:18.6	−70.0	11.00	4.77	3.75	IRw
68	AGAL337.258−00.101	16:36:56.42	−47:22:27.0	−67.6	11.00	4.48	3.50	IRw

Table A.1 - continued.

ID	Source	RA(J2000) (HH:MM:SS)	DEC(J2000) (DD:MM:SS)	$V_{\rm lsr}$ (km s^{-1})	d (kpc)	$\log\left(\frac{L_{\rm bol}}{L_\odot}\right)$	$\log\left(\frac{M_{\rm clump}}{M_\odot}\right)$	Class
69	AGAL337.286+00.007	16:36:34.76	−47:16:51.2	−105.8	9.44	3.10	3.82	70w
70	AGAL337.406−00.402	16:38:50.89	−47:27:56.2	−40.1	3.26	4.93	3.04	H II
71	AGAL337.704−00.054	16:38:29.57	−47:00:41.2	−46.5	12.26	5.50	4.15	H II
72	AGAL337.916−00.477	16:41:10.60	−47:08:02.9	−39.6	3.20	5.11	3.08	IRb
73	AGAL338.066+00.044	16:39:28.60	−46:40:34.6	−67.6	4.69	3.50	2.98	70w
74	AGAL338.786+00.476	16:40:22.55	−45:51:04.1	−64.6	4.49	2.69	3.09	70w
75	AGAL338.926+00.554	16:40:34.36	−45:41:40.6	−61.9	4.40	4.97	3.78	IRw
76	AGAL339.623−00.122	16:46:06.58	−45:36:44.3	−34.2	3.01	4.18	2.50	IRb
77	AGAL340.374−00.391	16:50:02.47	−45:12:43.0	−46.5	3.60	2.71	2.90	IRw
78	AGAL340.746−01.001	16:54:03.66	−45:18:45.1	−29.3	2.76	3.89	2.33	IRb
79	AGAL340.784−00.097	16:50:14.89	−44:42:31.6	−101.4	9.99	4.86	3.45	IRw
80	AGAL341.217−00.212	16:52:17.75	−44:26:54.5	−43.2	3.67	4.21	2.69	IRb
81	AGAL342.484+00.182	16:55:02.53	−43:13:03.1	−38.3	12.55	4.81	3.69	IRw
82	AGAL343.128−00.062	16:58:17.24	−42:52:02.0	−30.4	3.04	4.86	3.06	H II
83	AGAL343.756−00.164	17:00:50.08	−42:26:11.9	−27.9	2.90	4.00	2.79	IRw
84	AGAL344.227−00.569	17:04:07.75	−42:18:41.7	−23.0	2.52	3.99	3.05	IRw
85	AGAL345.003−00.224	17:05:11.03	−41:29:07.2	−27.1	3.02	4.81	2.99	H II

Table A.1 - continued.

ID	Source	RA(J2000) (HH:MM:SS)	DEC(J2000) (DD:MM:SS)	$V_{\rm lsr}$ (km s^{-1})	d (kpc)	$\log\left(\frac{L_{\rm bol}}{L_\odot}\right)$	$\log\left(\frac{M_{\rm clump}}{M_\odot}\right)$	Class
86	AGAL345.488+00.314	17:04:28.33	−40:46:24.6	−17.3	2.22	4.79	2.97	H II
87	AGAL345.504+00.347	17:04:23.10	−40:44:26.7	−17.1	2.25	4.64	2.63	IRb
88	AGAL345.718+00.817	17:03:06.46	−40:17:05.9	−12.0	1.56	3.27	2.30	IRb
89	AGAL351.131+00.771	17:19:34.42	−35:56:46.8	−5.6	1.82	2.80	2.09	70w
90	AGAL351.161+00.697	17:19:57.50	−35:57:50.8	−6.5	1.82	3.94	3.07	IRb
91	AGAL351.244+00.669	17:20:18.38	−35:54:43.4	−3.9	1.82	4.89	2.95	IRb
92	AGAL351.416+00.646	17:20:53.04	−35:47:02.3	−6.0	1.34	4.60	2.67	H II
93	AGAL351.444+00.659	17:20:54.60	−35:45:12.8	−5.0	1.34	3.98	3.15	IRw
94	AGAL351.571+00.762	17:20:51.00	−35:35:25.8	−3.7	1.34	2.64	2.22	70w
95	AGAL351.581−00.352	17:25:24.90	−36:12:44.0	−93.7	6.81	5.39	3.94	IRb
96	AGAL351.774−00.537	17:26:42.59	−36:09:20.6	−2.7	1.00	4.22	2.42	IRb
97	AGAL353.066+00.452	17:26:13.50	−34:31:54.0	−2.1	0.86	1.76	1.25	IRw
98	AGAL353.417−00.079	17:29:18.94	−34:32:05.9	−54.2	6.06	3.65	3.25	70w
99	AGAL354.944−00.537	17:35:11.69	−33:30:23.8	−5.5	1.91	2.68	2.17	70w

Notes: The columns are as follows: (1) ID of each source; (2) name from the ATLASGAL catalogue; (3)–(4) Equatorial coordinates (J2000) of the 870 μm continuum peak; (5) the local standard rest velocity ($V_{\rm lsr}$) from the $C^{17}O$ (3–2) data from Giannetti et al. (2014); (6)–(8) distances, bolometric luminosities and clump masses from König et al. (2017); (9) classification of the clump: 70 μm dark (70w), infrared quiet (IRw), infrared bright (IRb) or H II regions (H II) from König et al. (2017).

A.3 Possible bias on the trends between the CO and the clump properties

This Appendix presents the investigation of the possible bias that might affect the trends found in Sect. 3.6.3 using the CO spectra convolved to the same angular beam of 13.″4. Thanks to the relatively high-angular resolution data from the CHAMP$^+$ data, two major effects were further exploited towards the mid-J CO observations: the dilution effects introduced by computing the CO line luminosity at different beam sizes (detailed in Appendix A.3.1); and the underestimation of the integrated fluxes due to the presence of self-absorption features in the data (Appendix A.3.2).

A.3.1 Beam dilution effects: comparison at different beam sizes

The relations between the L_{CO} values of the data convolved to 13.″4 and the clump properties were further investigated in terms of varying the size of the convolved beam (see Eq. (3.3)). As detailed in Sects. 3.5.1 and 3.5.2, the mid-J CO spectra were convolved to the same linear scale of 0.24 pc for a distance-limited subsample of sources and were integrated over the full extension of the CHAMP$^+$ maps for 92 of 99 clumps.

Fig. A.3 presents the CO luminosity against the bolometric luminosity of the clumps for both datasets. No significant differences on the correlation between L_{CO} and L_{bol} were found when comparing the different datasets, even when including the 13.″4 in the comparison ($0.82 \leq \rho \leq 0.87$ for the CO (6–5) and CO (7–6) transitions in the three datasets, see Tables 3.4 and A.2). In addition, the change on the slopes when comparing the different datasets are roughly consistent within their errors, indicating that the distribution of the CO luminosity as a function of L_{bol} is robust against different beam sizes for the mid-J CO transitions investigated in this work. Therefore, the trends are not likely affected by beam dilution effects. Such result indicates that the bulk of the CO emission arises from a compact region within the clumps and the contribution of the extended CO emission corresponds to a small fraction of the total flux.

Figure A.4 presents the CO line luminosity as a function of the mass of the clumps. The plots indicate that the correlations for the distance-limited subsample (left panel) are weaker ($\rho \leq 0.44$, see Table A.2) than those observed towards the two other samples ($\rho \geq 0.59$, see Table 3.4). As similarly found towards the 13.″4 dataset, the slopes derived for the distance-limited subsample (0.48±0.08 and 0.47±0.09 for CO (6–5) and CO (7–6),

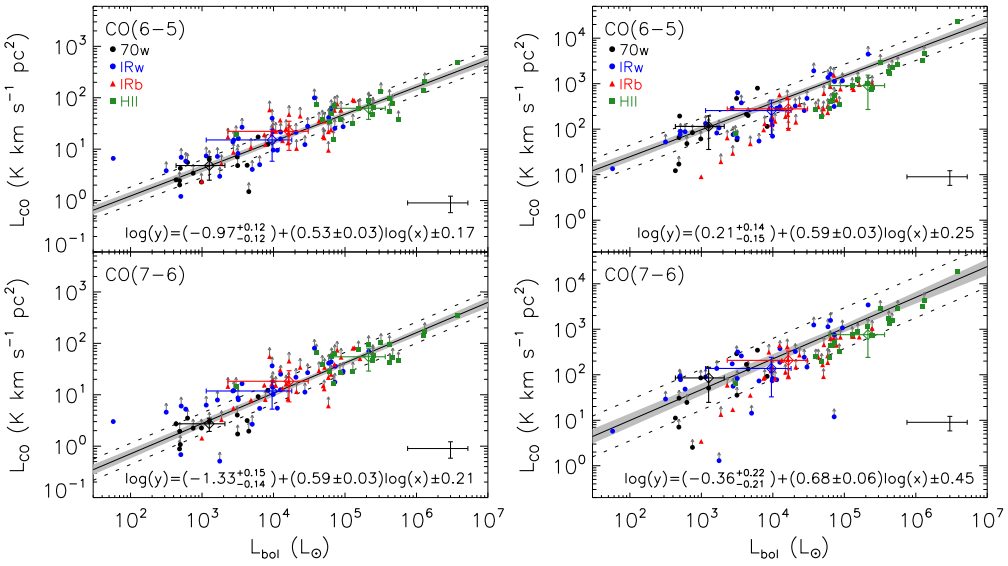


Figure A.3: The distribution of the CO (6–5) (top panels) and CO (7–6) (bottom) line luminosities as a function of the bolometric luminosity ($L_{\rm bol}$) of the TOP100 clumps. The CO luminosity was integrated over a fixed linear scale of 0.24 pc (left panels, see text for details) and over the full extension of the CHAMP$^+$ maps (right panels). For a complete description of the elements in the plots, see Fig. 3.5.

respectively, see the left panel of Fig. A.4) or even the full-map sample (0.92±0.07 and 0.89±0.10 for CO (6–5) and CO (7–6), respectively, see Fig. A.4) indicates no significant changes as a function of J. For the full-map sample (right panel of Fig. A.4), the correlation between $L_{\rm CO}$ and $M_{\rm clump}$ is stronger ($\rho \gtrsim 0.6$, see Table A.2, indicating that the aperture sizes adopted for computing $L_{\rm CO}$ and $M_{\rm clump}$ are relevant on the determination of both quantities. The weaker correlations found between $L_{\rm CO}$ derived for the common linear size and the corresponding $M_{\rm clump}$ values can be explained due to the fact that the $L_{\rm CO}$ is integrated over only a fraction of the beam used for evaluating $M_{\rm clump}$. Indeed, (König et al., 2017) evaluated the mass of the TOP100 clumps assuming a minimum aperture size of 55.″1. The $L_{\rm CO}$ values for the full-map sample, however, is integrated over a larger or equal aperture of those adopted by König et al. (2017), explaining the similar relations of those found towards the 13.″4 dataset.

Finally, the comparison between $L_{\rm CO}$ and L/M values is presented in Fig. A.5. The data display a relatively stronger correlation for the common linear scale dataset ($\rho \geq 0.64$, $p < 0.001$ for all mid-J CO lines) rather the 13.″4 spectra ($\rho \leq 0.43$, $p \leq 0.09$ for the mid-J CO lines, see Table 3.4). The observed correlation is even stronger than those values obtained towards the full-map sample ($\rho \leq 0.49$, see Table A.2).

The overall results suggests that the analysis of the mid-J CO emission is robust in

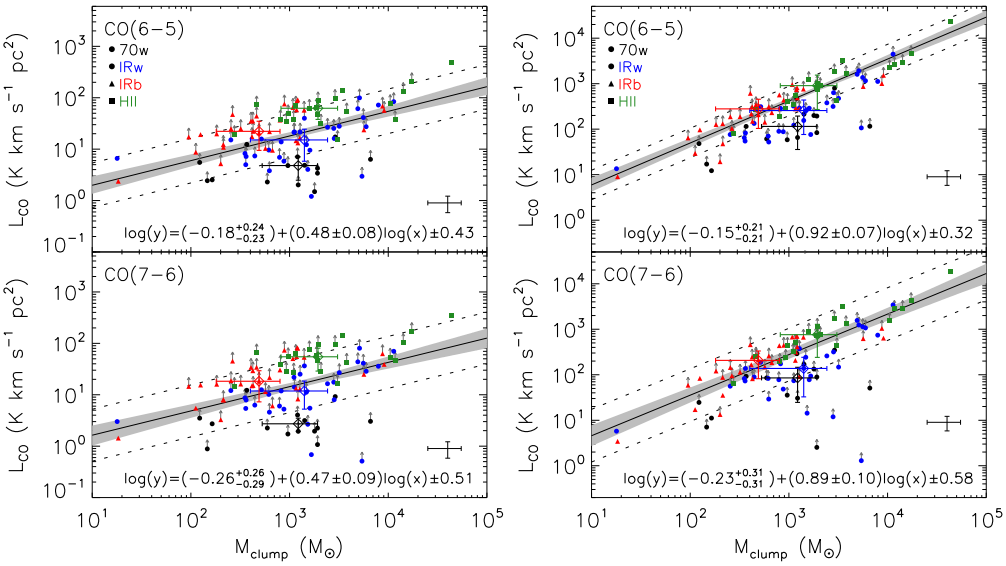


Figure A.4: Similar as Fig. A.3 but displaying the distribution of the CO line luminosities as a function of the mass ($M_{\rm clump}$) of the TOP100 clumps.

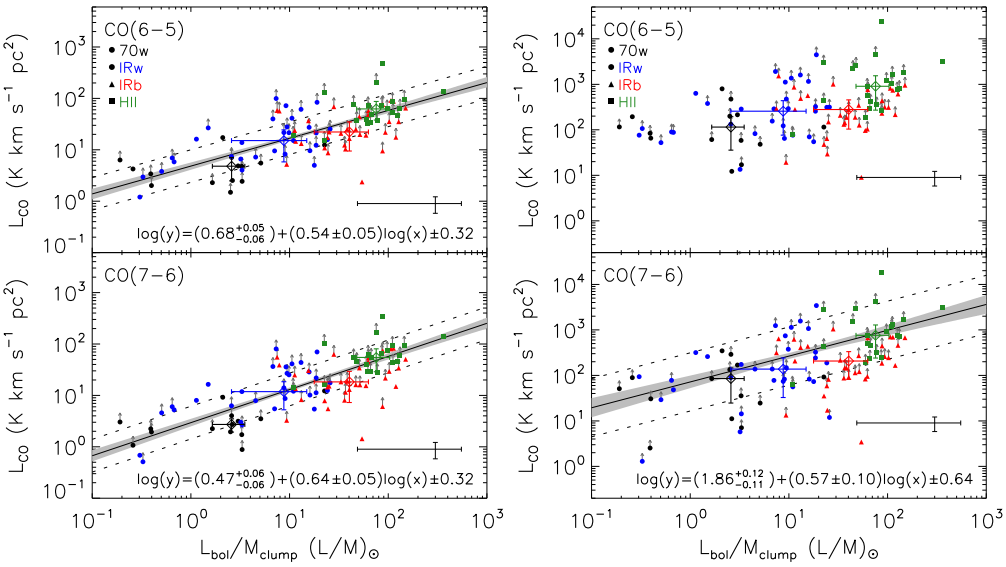


Figure A.5: Similar as Fig. A.3 but displaying the distribution of the CO line luminosities as a function of the luminosity-to-mass ratio (L/M) of the TOP100 clumps.

terms of beam dilution effects. Dilution effects play a significant role only when comparing the CO emission with quantities evaluated integrated over relatively large aperture sizes, such as the mass of the clumps.

A.3.2 The self-absorption features

This section presents the analysis on the biases that might be introduced into the trends found between the CO line luminosity and the clump properties from Sect. 3.6.3 by the self-absorption features present in the CO profiles.

Table A.2 - Spearman tests for the CO line luminosity as a function of the clump properties for the distance-limited and for the full-map samples.

	Distance-limited		Full-map	
Property	CO (6–5)	CO (7–6)	CO (6–5)	CO (7–6)
$L_{\rm bol}$	0.86, $p<0.001$; $\rho_p=0.91$	0.87, $p<0.001$; $\rho_p=0.88$	0.84, $p<0.001$; $\rho_p=0.96$	0.82, $p<0.001$; $\rho_p=0.97$
$M_{\rm clump}$	0.44, $p<0.001$; $\rho_p=0.85$	0.41, $p<0.001$; $\rho_p=0.81$	0.72, $p<0.001$; $\rho_p=0.93$	0.59, $p<0.001$; $\rho_p=0.94$
L/M	0.67, $p<0.001$	0.79, $p<0.001$	0.43, $p<0.001$	0.49, $p<0.001$

Notes: The rank ρ and its corresponding probability (p) are shown for each comparison. A p-value of <0.001 indicate a correlation at 0.001 significance level. p-values of 0.05, 0.002 and <0.001 represent the ~2, 3 and $>3\,\sigma$ confidence levels. For $L_{\rm bol}$ and $M_{\rm clump}$, the partial correlation coefficient, ρ_p, is also shown.

Figure A.6 presents the CDF of the CO line luminosity for only the sources ($\sim$15 clumps) that are likely not affected by self-absorption in their CO spectra. The CDFs indicate that the median values per evolutionary class (indicated as dashed lines in the plot) are shifted toward larger $L_{\rm CO}$ values when compared with the median values per class using the whole TOP100 sample (shown as filled lines). The observed shift towards larger $L_{\rm CO}$ values is expected since a considerable fraction of the observed flux is lost due to the self-absorption features present in most of the TOP100 clumps, suggesting that these features might play a significant role on the analysis of the integrated CO properties as a function of the clump properties.

Figure A.7 compares the CO line luminosity with the properties of the clumps that are not affected by self-absorption features on their CO spectra. The left panel presents the CO line luminosity versus $L_{\rm bol}$ for the three J transitions. The derived power-law indices are $\approx0.58\pm0.08$, 0.69 ± 0.08 and 0.70 ± 0.06, for the CO (4–3), CO (6–5) and CO (7–6), respectively. Although the indices are systematically larger than those reported in Table 3.5 (see Sect. 3.6.3), the differences between the power-law indices found for the self-absorption free subsamp and those derived in Sect. 3.6.3 decrease from CO (4–3) towards CO (7–6), suggesting that self-absorption effects are smoothed towards higher J transitions. In addition, for each J transition, the power-law indices from Fig. A.7 and those reported in Sect. 3.6.3 for the whole TOP100 sample are compatible within their errors. This result suggests that self-absorption does not play a significant role on the scaling of the CO luminosity as a function of $L_{\rm bol}$. Similar results are also found when plotting $L_{\rm CO}$ against the mass of the clumps for the same subsample (see right panel of Fig. A.7). These results

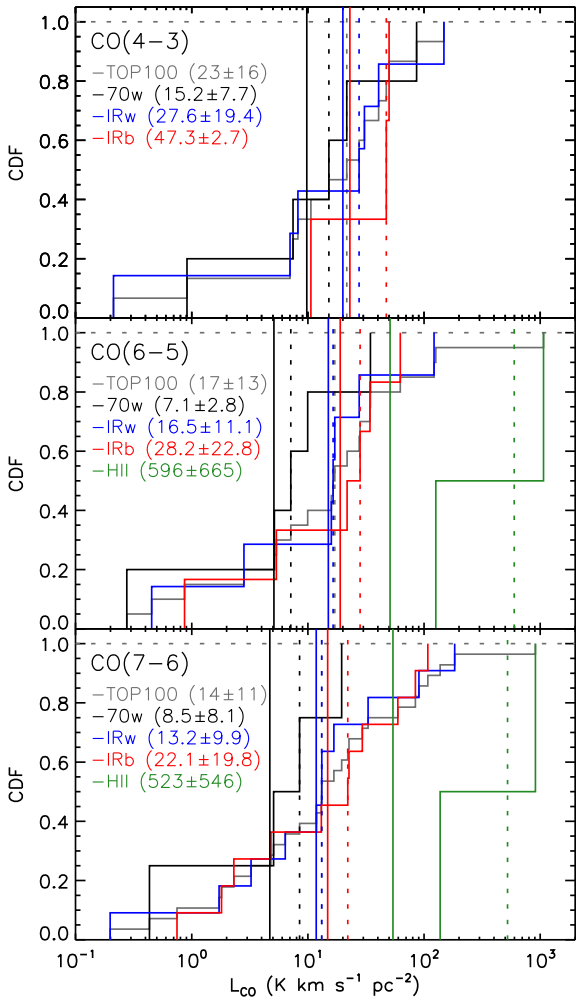


Figure A.6: Cumulative distribution function (CDF) of the CO line luminosity of the subsample of clumps that are not likely affected by self-absorption features on their CO spectra. The median values per evolutionary class and those considering the full TOP100 sample are indicated as the vertical dashed and filled lines in their corresponding colour. The median and absolute deviation from the median values per evolutionary class are presented in each panel.

reinforce the conclusion that the power-law indices of the trends found in Sect. 3.6.3 are robust against self-absorption features present in the CO profiles of the TOP100 sample, especially towards the mid-J CO transitions.

In order to further exploit the biases that might affect the relations between L_{CO} and the clump properties found in Sect. 3.6.3, the CO line luminosities were recalculated using the integrated fluxes obtained from the Gaussian fit of the CO profiles (see Sect. 3.5.3), using Eq. (3.13) and adopting the same assumptions made in Sect. 3.5.4. The observed and Gaussian CO line luminosities are plotted against the luminosity and the mass of the TOP100 clumps in Fig. A.8. For each CO transition, the slopes of the fit of L_{CO} versus L_{bol} (left panel) for the observed and Gaussian datasets are consistent within their errors,

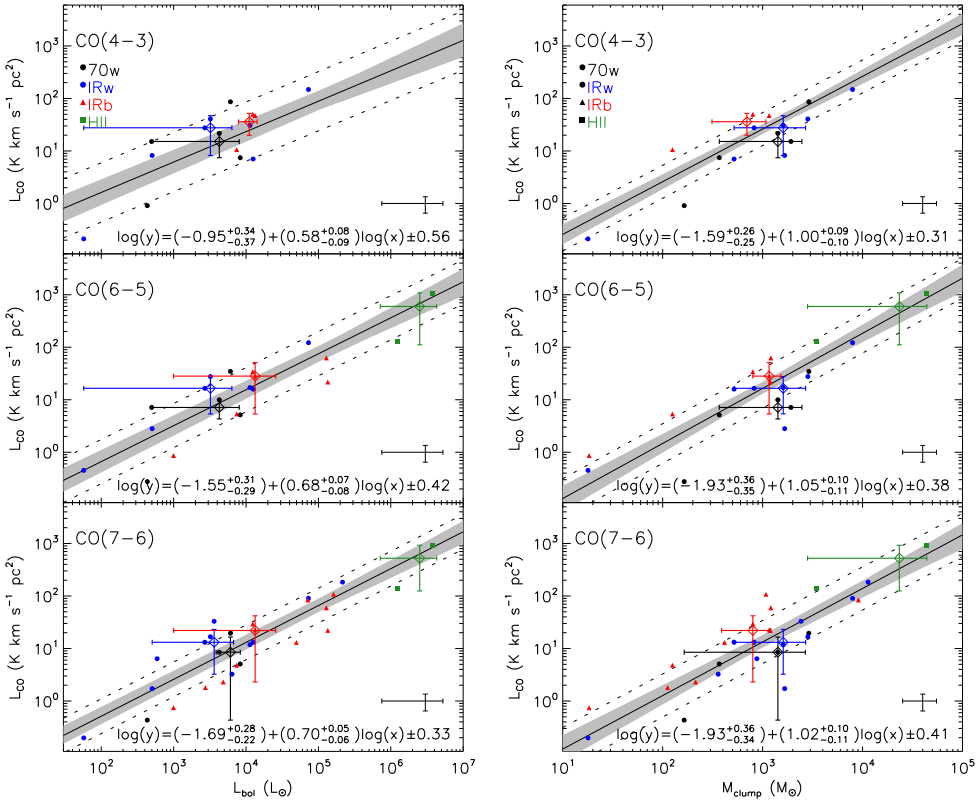


Figure A.7: CO line luminosity versus the clump properties for the sources that are not affected by self-absorption features. The L_{CO} values are plotted against the bolometric luminosity (left) and the mass (right) of the clumps (right). For a complete description of the elements in the plots, see Fig. 3.5.

indicating that the presence of self-absorption features in the data is not likely affecting the dependency of the CO emission with the bolometric luminosity of the clumps. Similar results are also found when comparing both datasets with the mass of the clumps (see right panel of Fig. A.8). These findings suggest that the slopes of the trends found in Sect. 3.6.3 are robust in terms of the presence of self-absorption features.

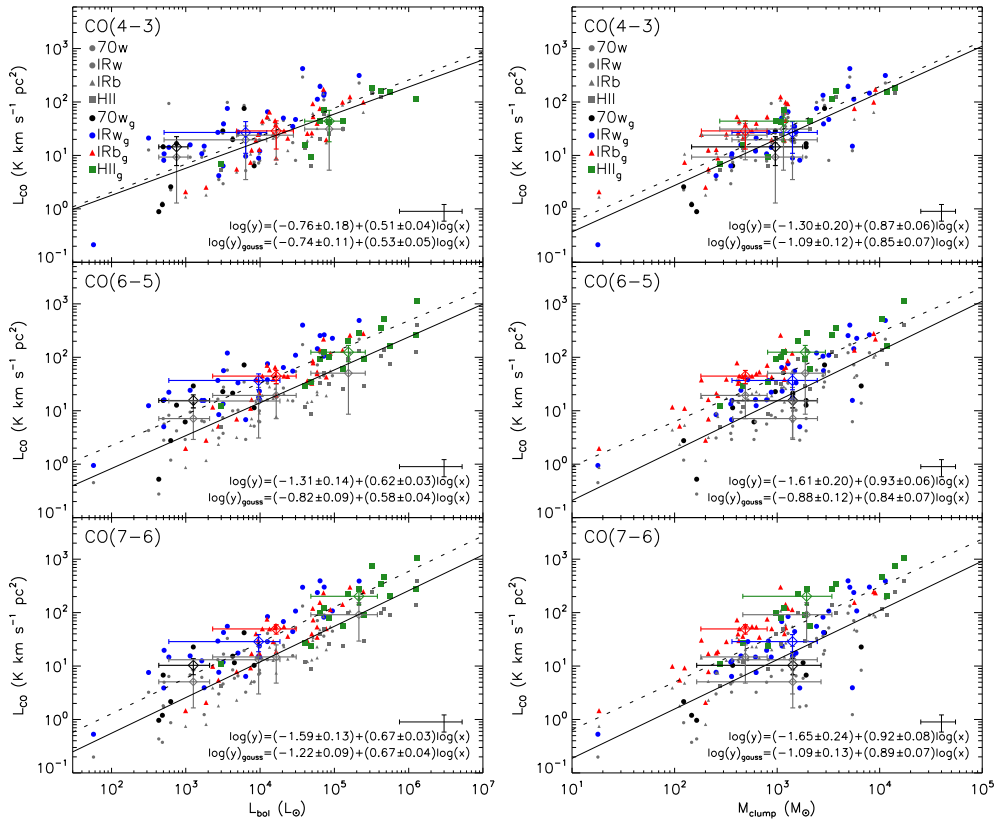


Figure A.8: Observed and Gaussian CO line luminosity versus the TOP100 clump properties. The observed L_{CO} (grey symbols) and the Gaussian CO line luminosity (coloured symbols) are plotted against the bolometric luminosity (left panel) and the mass of the clumps (right). The filled and dashed lines indicate the best fit of the observed and the Gaussian datasets. The median values for each class are shown as open diamonds and their error bars correspond to the median absolute deviation of the data from their median value. The typical error bars are shown in the bottom right side of the plots.

Appendix B

Near Infrared

B.1 Channel maps

This Appendix presents the channel maps of the Brγ feature at 2.1661 μm, the H_2 (1–0) S(1) and (1–0) S(0) at 2.1218 and 2.2235 μm, and the CO (2–1) bandhead emission at 2.2935 μm towards the HMYSOs observed with NIFS. The construction of the channel maps is described in Sect. 4.3.3. The description of the components of each plot is presented in Fig. B.1.

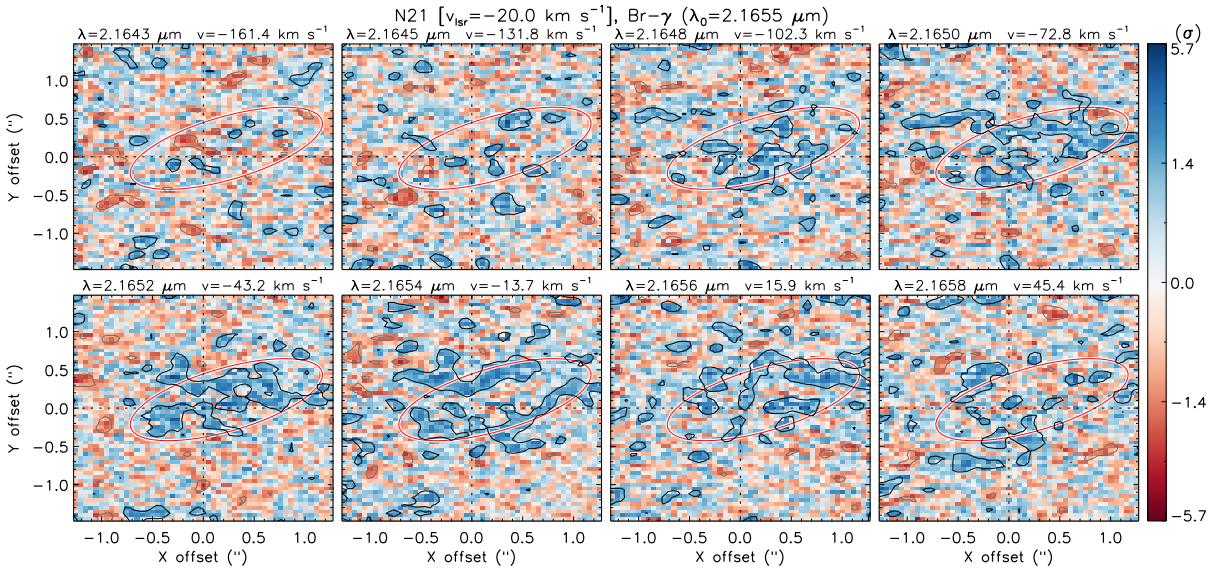


Figure B.1: Channel maps of the Brγ transition towards the source N21. The intensity is scaled in σ units and is shown in a red-to-blue divergent scale, as indicated by the bar at the right of each panel. The corresponding wavelength and the velocity ($v(\lambda_0) = 0\,\mathrm{km\,s^{-1}}$) are shown on the top of each channel. The maps are shown in linear or square-root scaling in order to improve the contrast between the structures. The full width at half maximum of the Gaussian model for the point-like continuum emission (see Sect. 4.3.2) is indicated as the red curve. The black contours are placed at 3^n-σ, where $n = 0, 1, 2, 3...$, while the grey contours are placed at -3^n-σ levels.

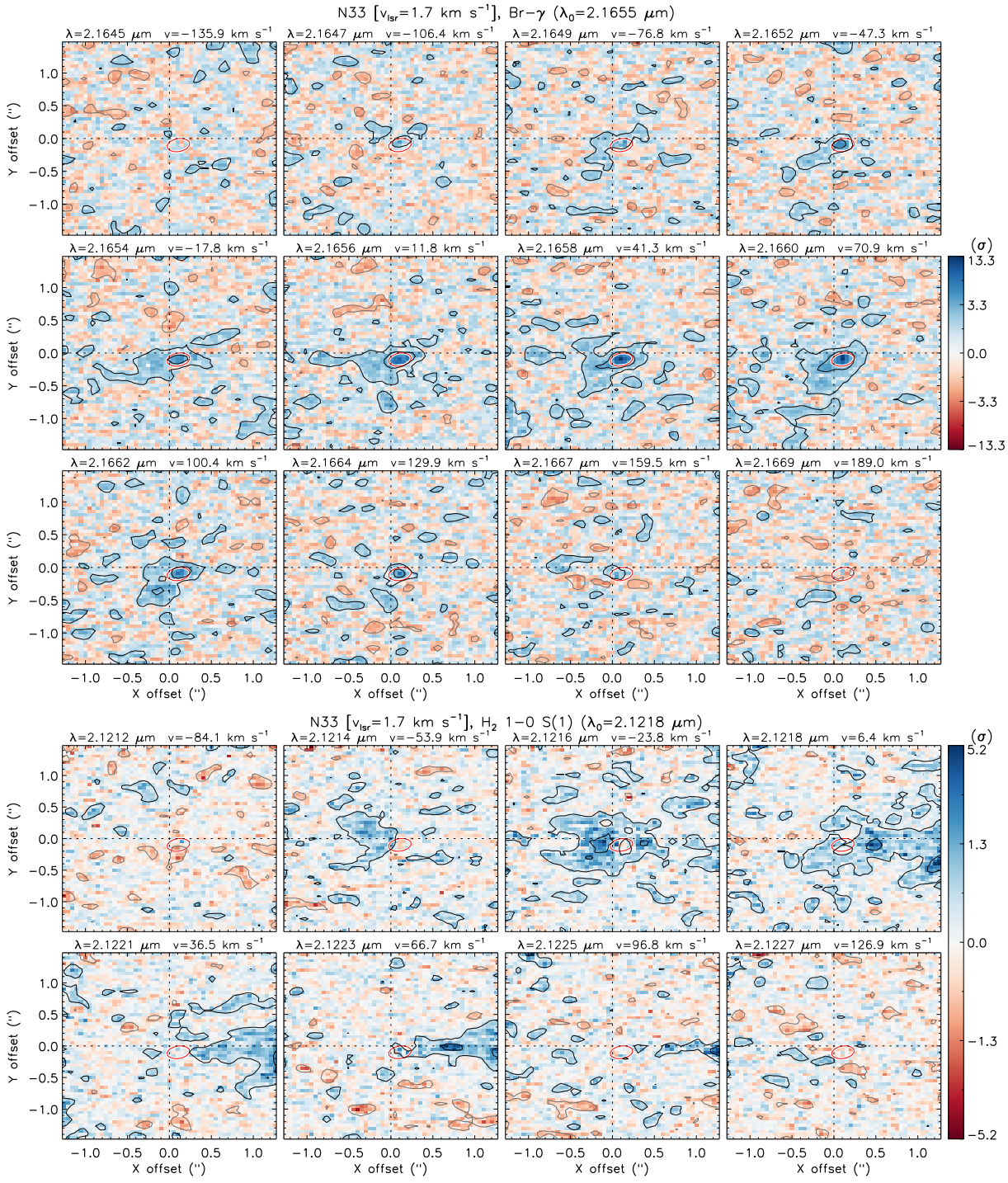


Figure B.2: Channel maps of the Brγ (top) and the H_2 (1–0) S(1) (bottom) emission towards the source N33.

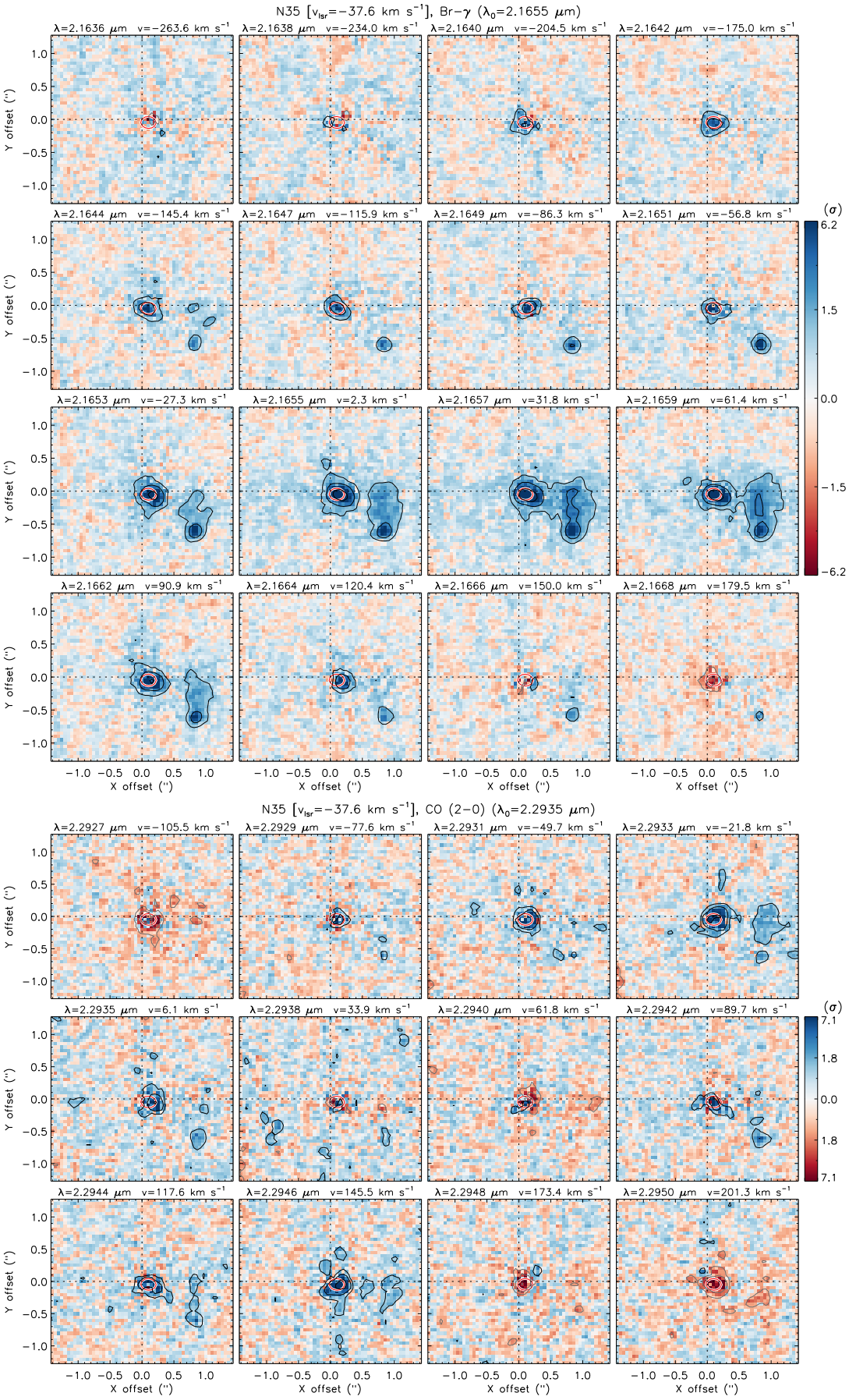


Figure B.3: Channel maps of the Brγ (top) and the CO (2–1) (bottom) emission towards the source N35.

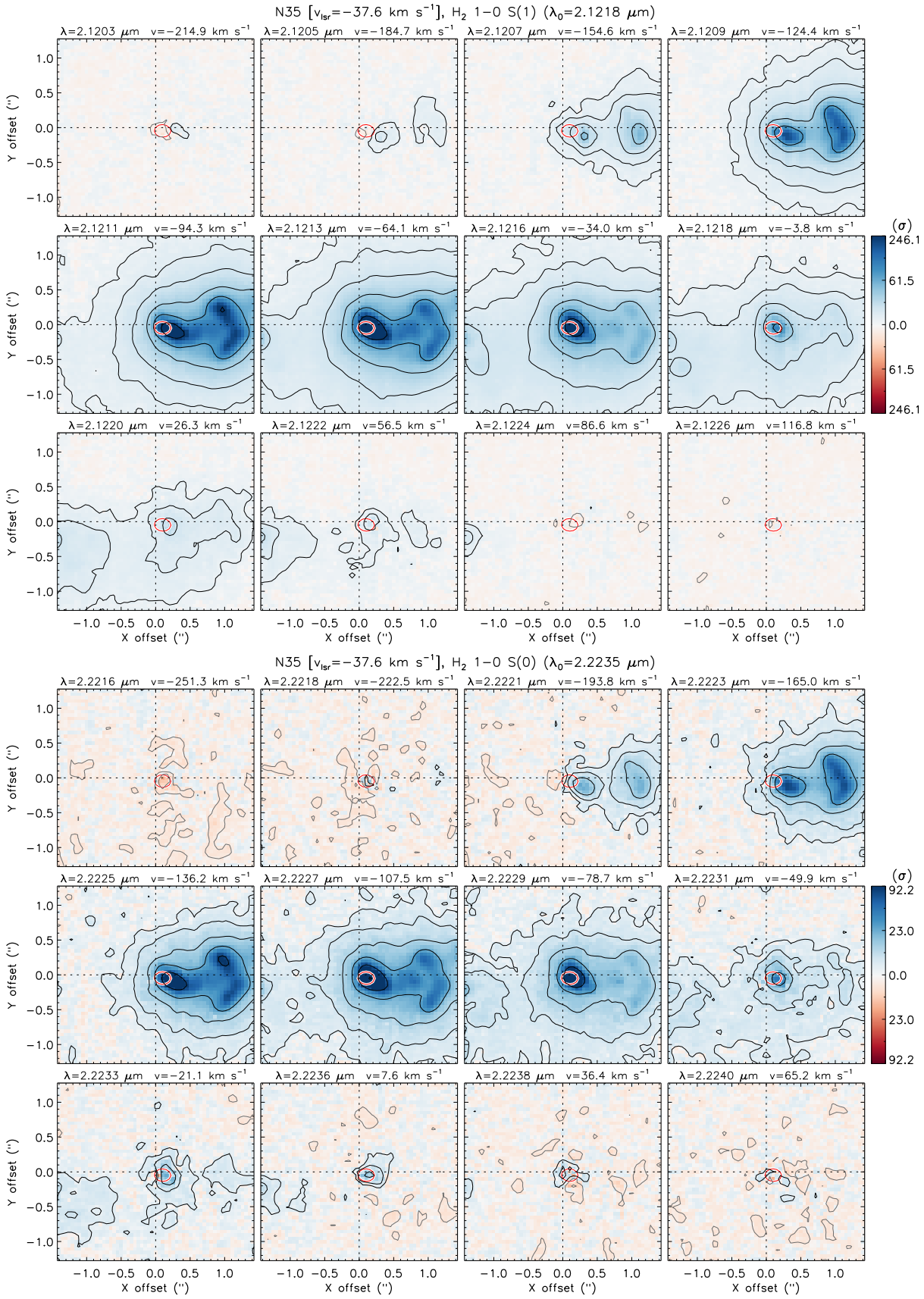


Figure B.4: Channel maps of the H_2 (1–0) S(1) (top) and the H_2 (1–0) S(0) (bottom) emission towards the source N35.

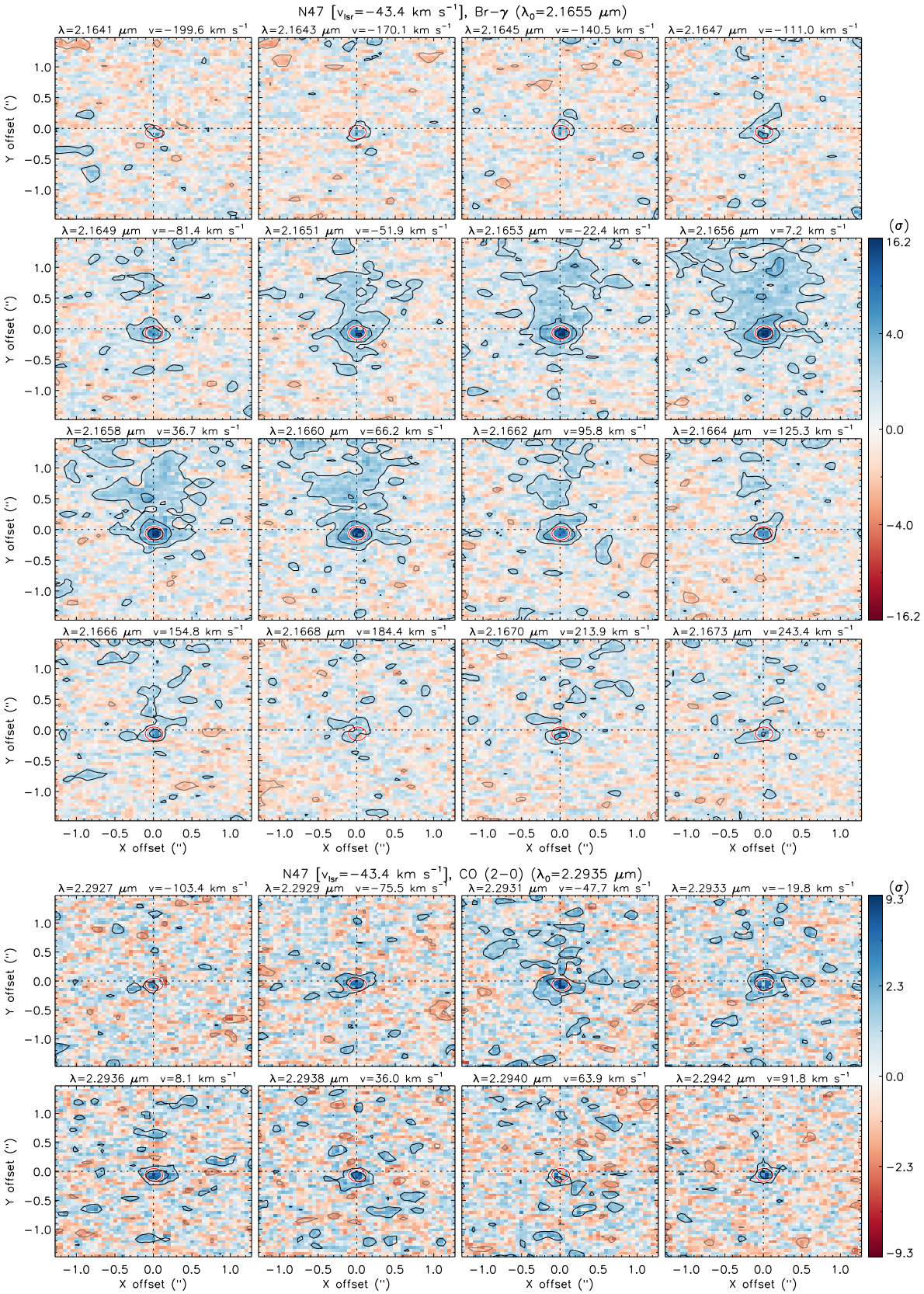


Figure B.5: Channel maps of the Brγ (top) and the CO (2–1) (bottom) emission towards the source N47.

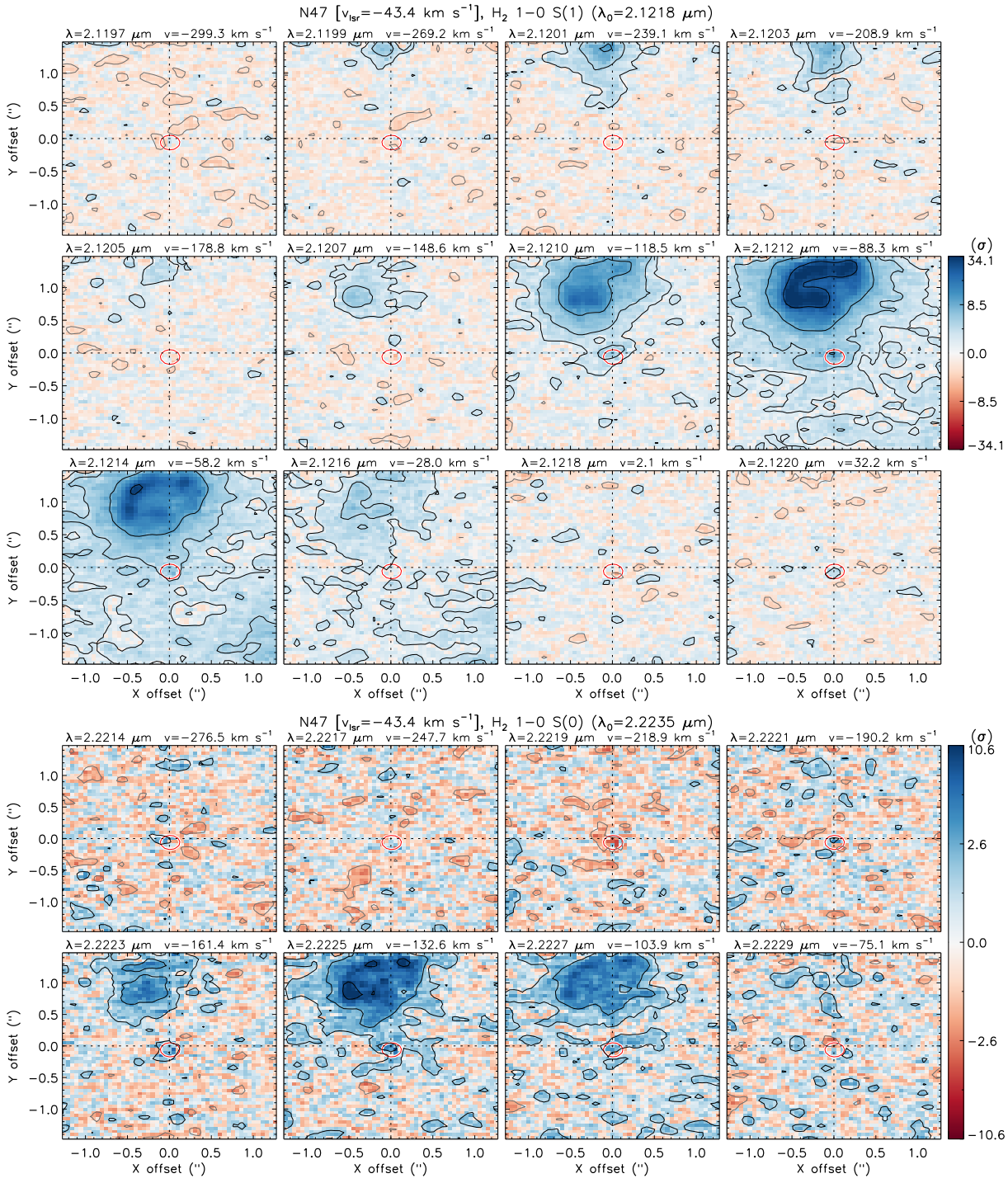


Figure B.6: Channel maps of the H_2 (1–0) S(1) (top) and the H_2 (1–0) S(0) (bottom) emission towards the source N47.

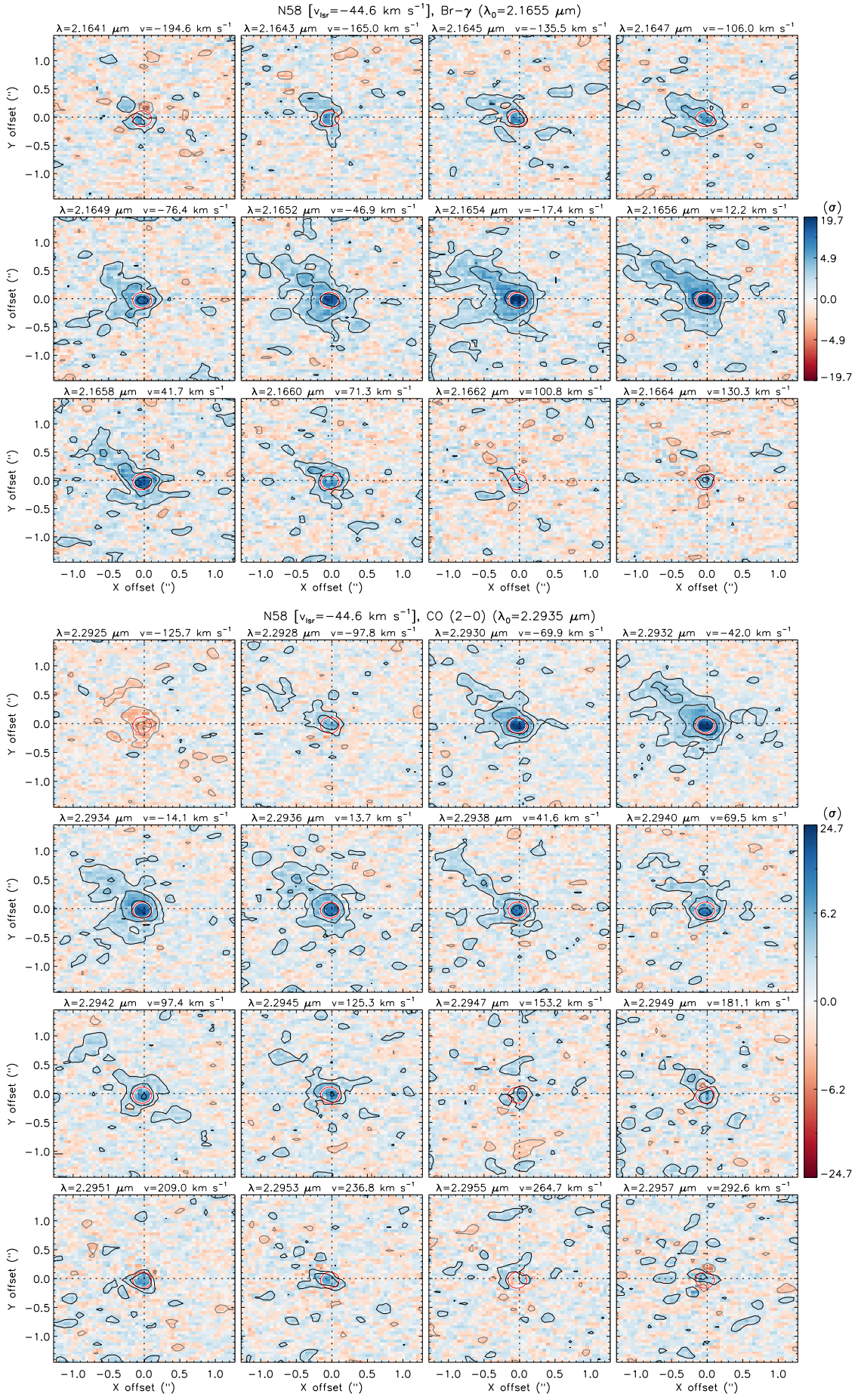


Figure B.7: Channel maps of the Brγ (top) and the CO (2–1) (bottom) emission towards the source N58.

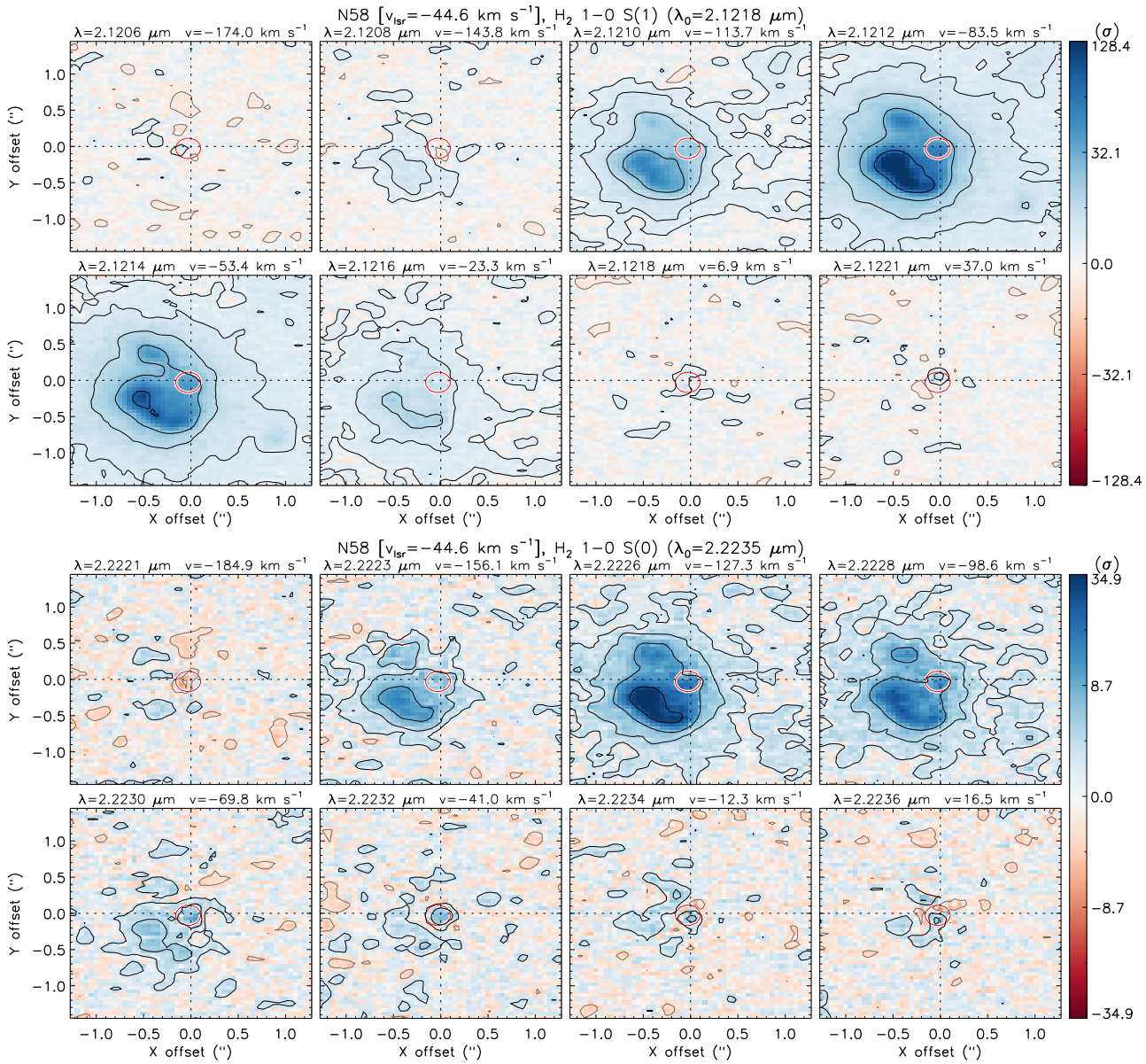


Figure B.8: Channel maps of the H_2 (1–0) S(1) (top) and the H_2 (1–0) S(0) (bottom) emission towards the source N58.

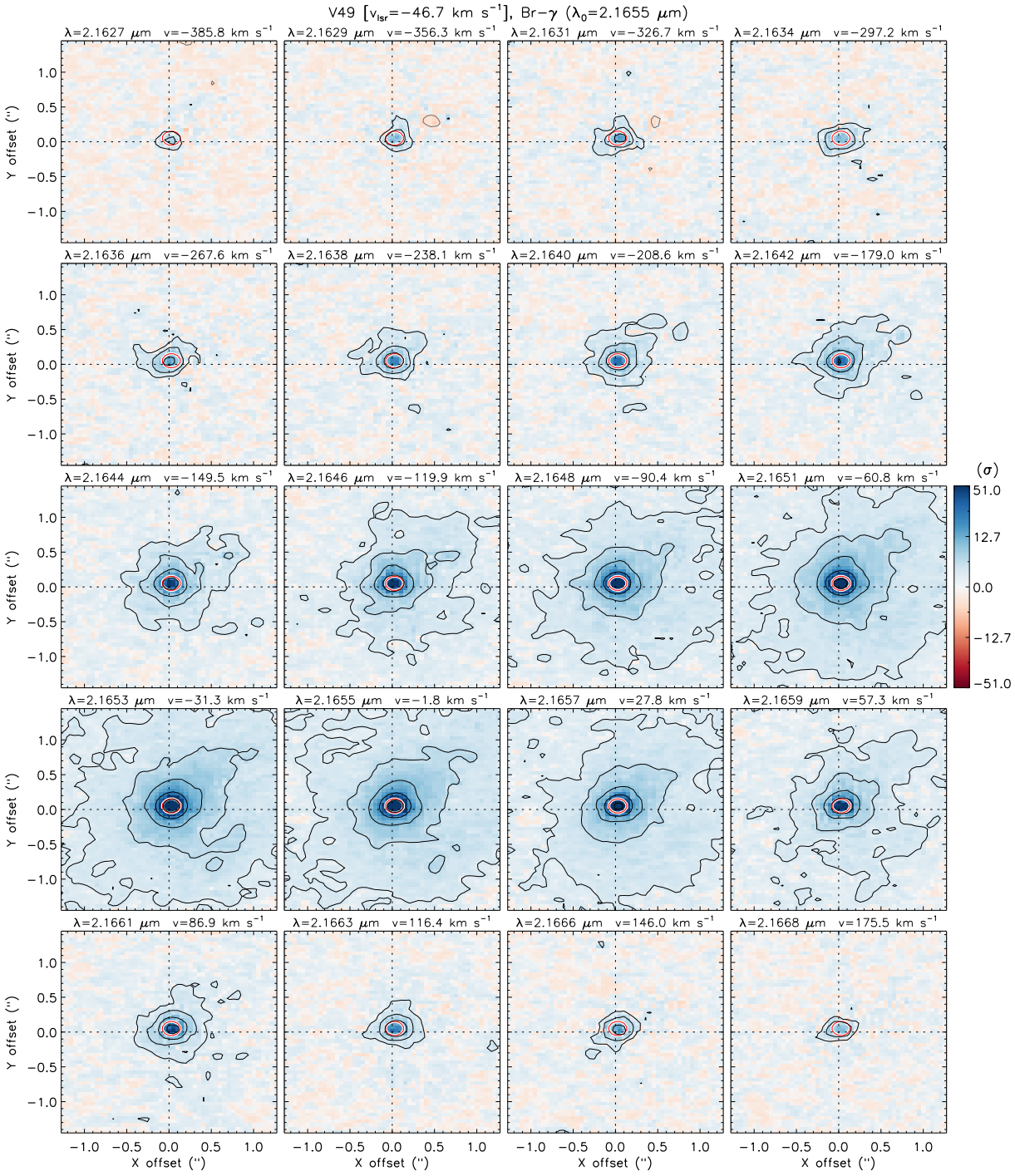


Figure B.9: Channel maps of the Brγ transition towards the source V49.

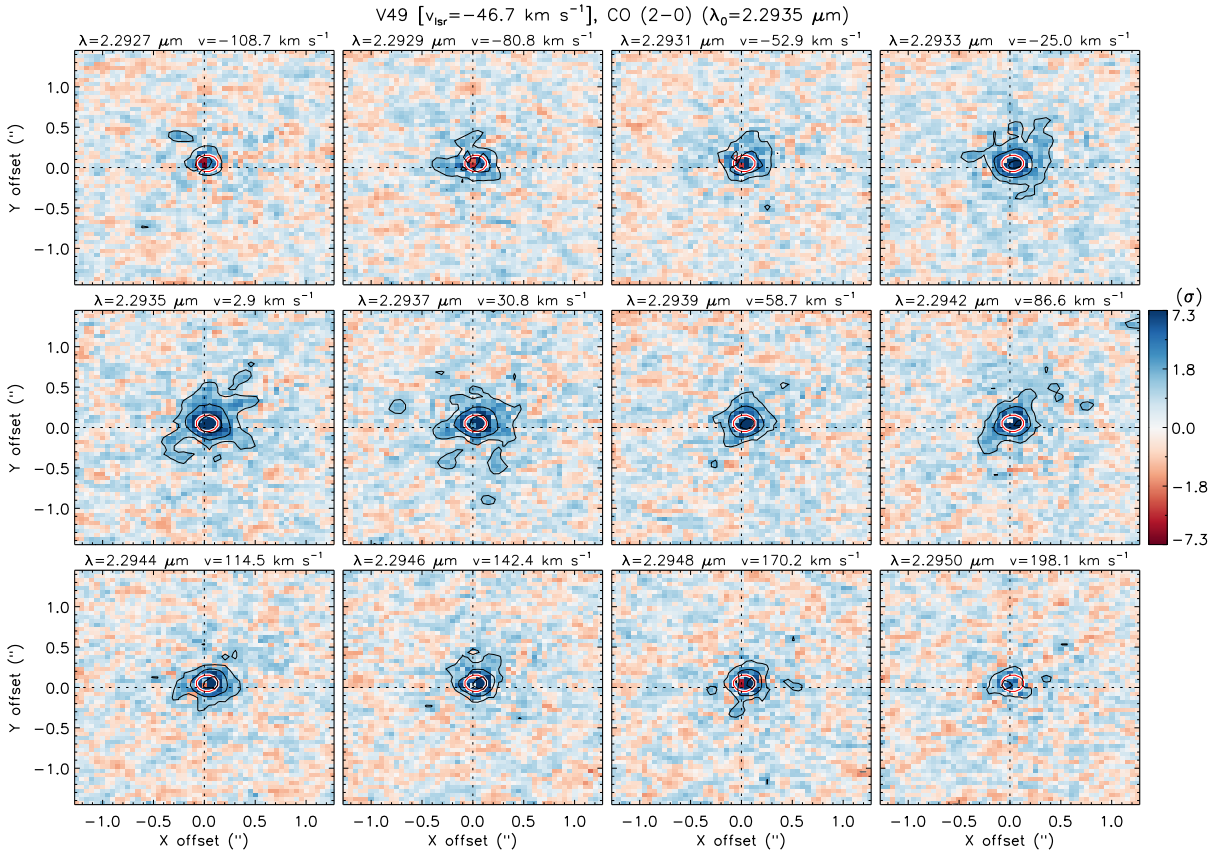


Figure B.10: Channel maps of the CO (2–1) (bottom) emission towards the source V49.

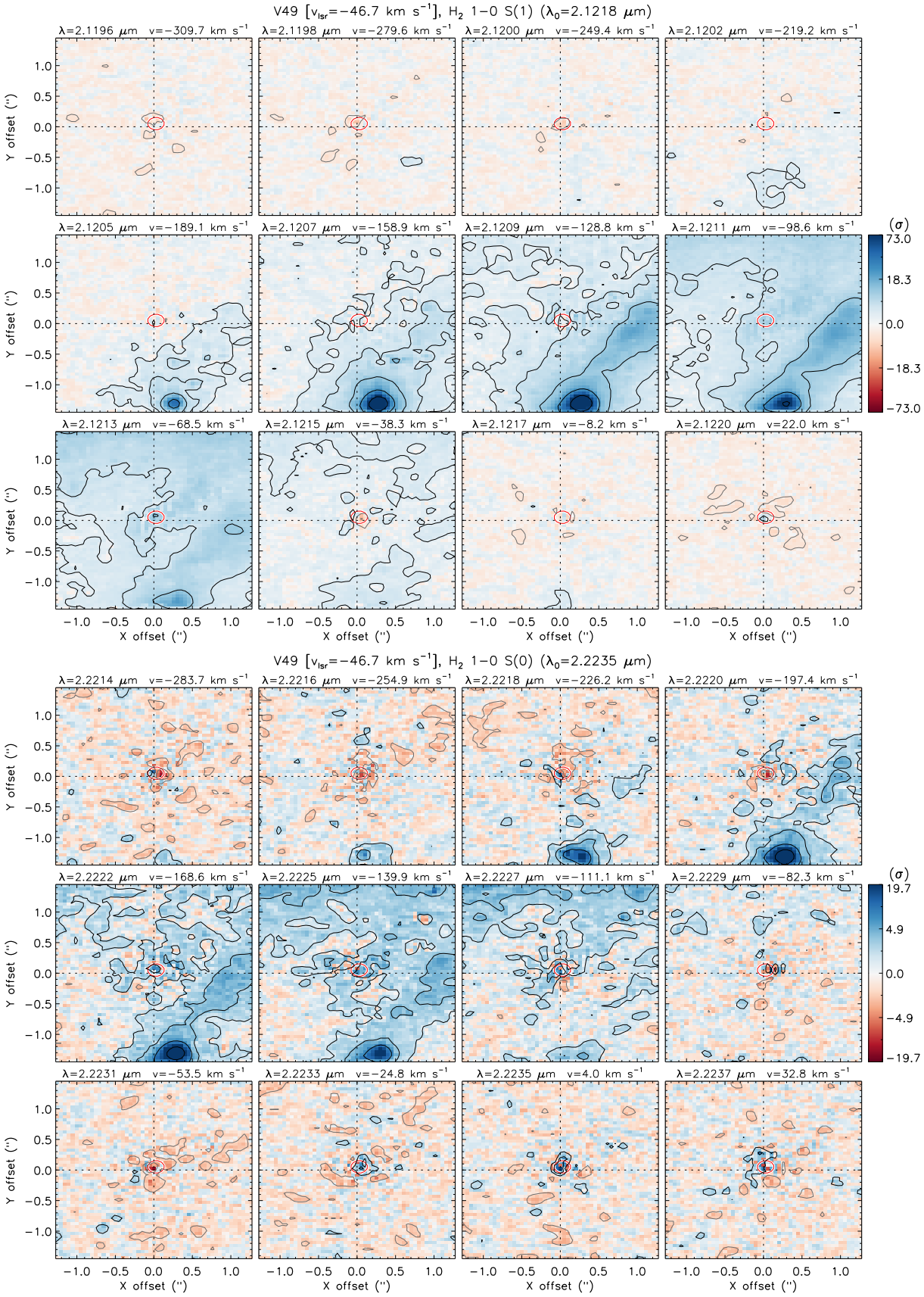


Figure B.11: Channel maps of the H_2 (1–0) S(1) (top) and the H_2 (1–0) S(0) (bottom) emission towards the source V49.

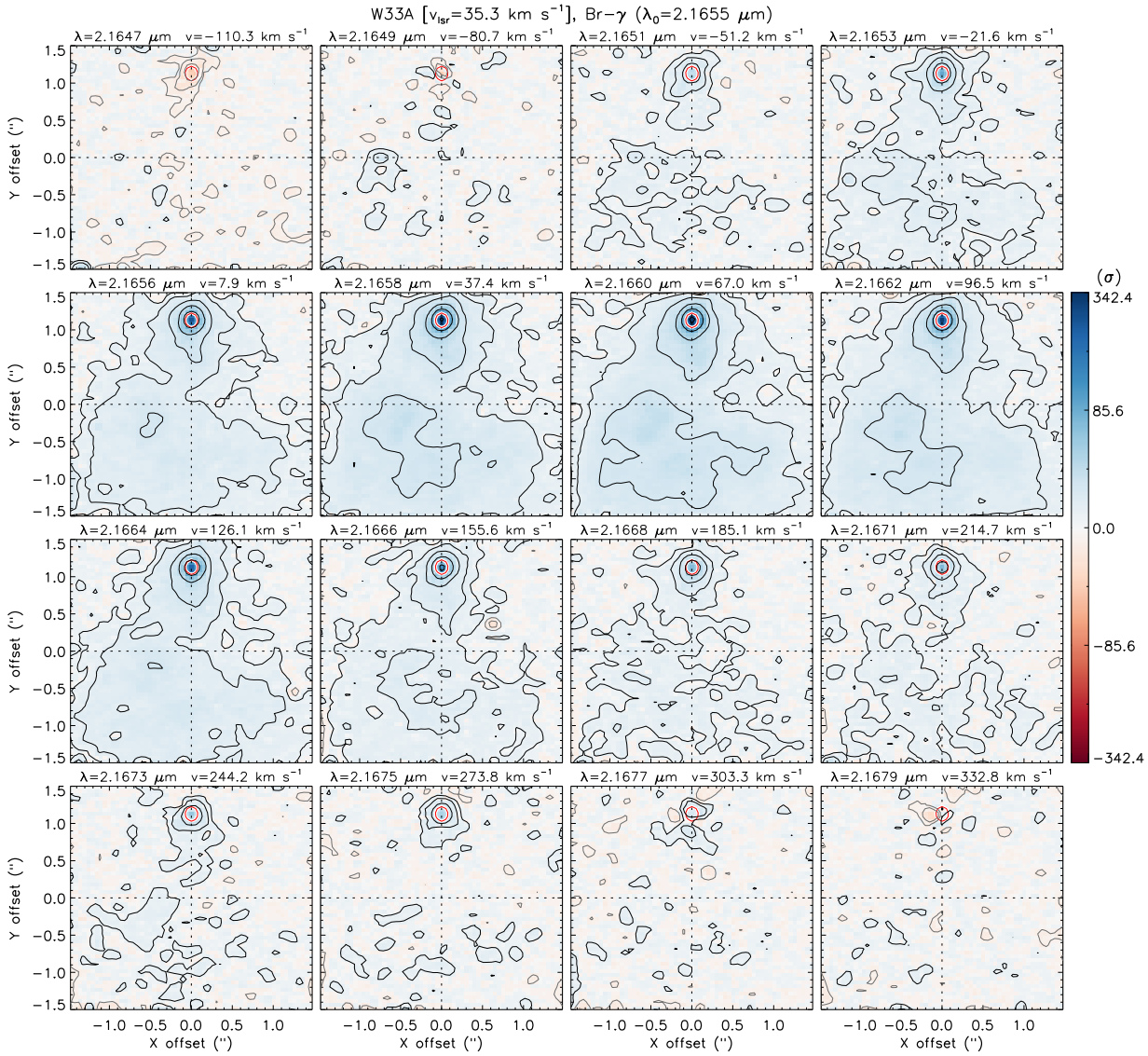


Figure B.12: Channel maps of the Brγ emission towards the source W33A.

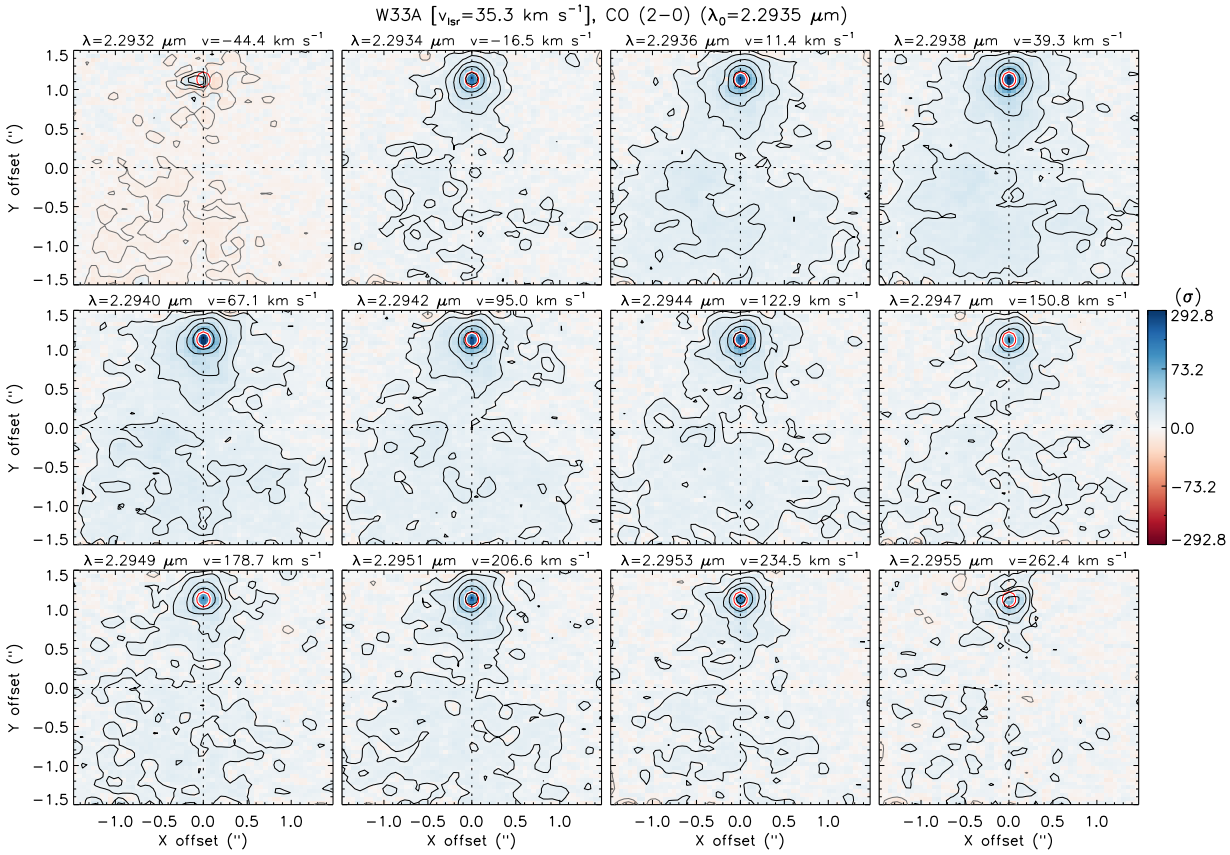


Figure B.13: Channel maps of the CO (2–1) emission towards the source W33A.

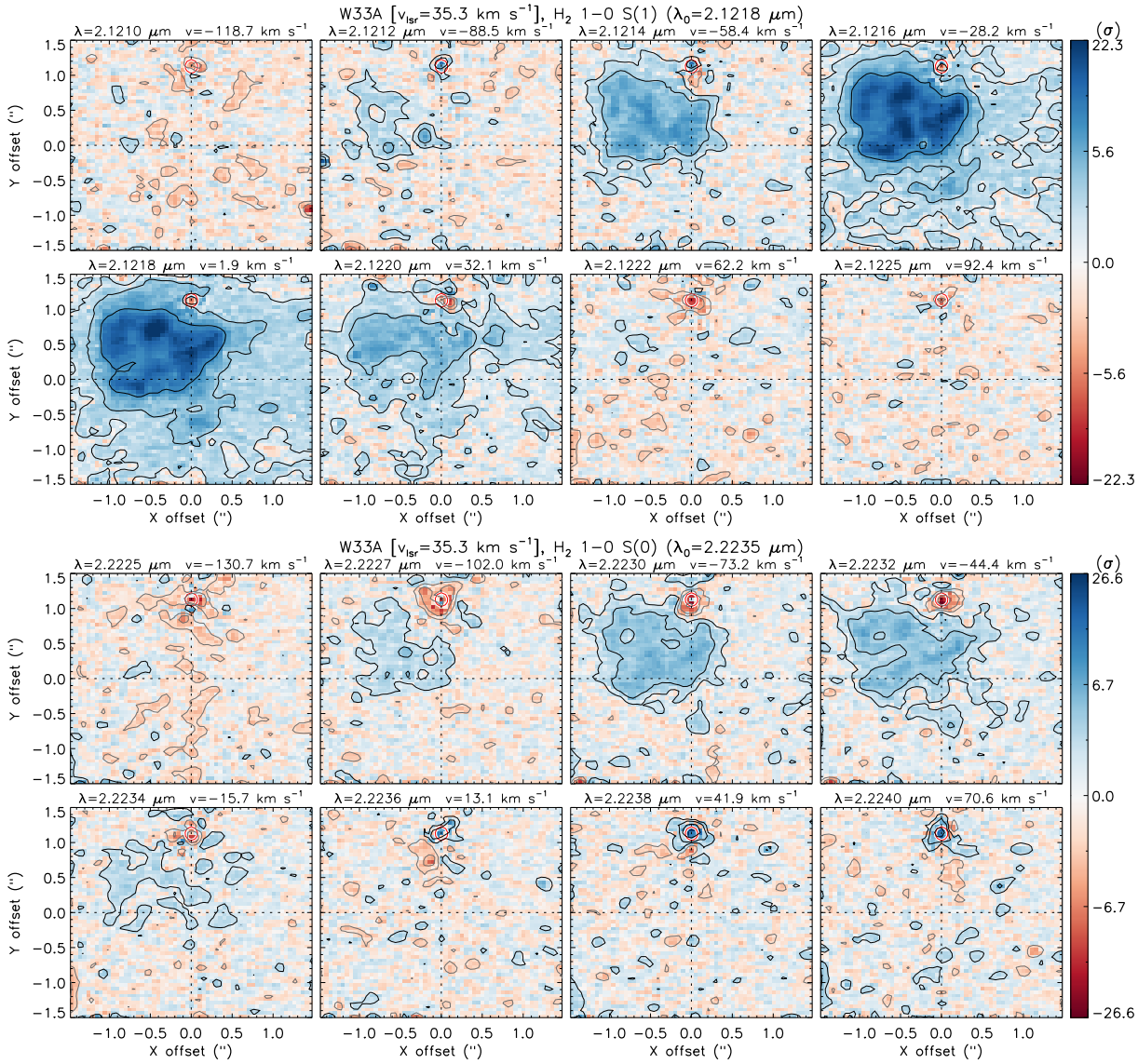


Figure B.14: Channel maps of the H_2 (1–0) S(1) (top) and the H_2 (1–0) S(0) (bottom) emission towards the source W33A.

B.2 Tables

This Section presents the tables containing the average flux of the spectra presented in Fig. 4.3.4.

Table B.1 - Parameters of the emission lines of the flux-calibrated average H_2 nebular spectrum of the source N33.

Transition	λ_0	λ_{peak}	FWHM	F_λ	σ_F	$F_\lambda/F_{2.12\mu m}$	$\sigma_{F\lambda/2.12}$
H_2 3-2 S(6)	20130	20123.9	2.13	0.58	0.25	0.207	0.090
H_2 1-0 S(1)	21218	21216.3	4.27	2.80	0.25	1.000	0.127
H_2 3-2 S(4)	21280	21273.9	4.27	0.17	0.07	0.062	0.024
H_2 3-2 S(3)	22014	21999.4	2.13	0.27	0.12	0.097	0.043
H_2 1-0 S(0)	22235	22229.8	2.13	0.49	0.10	0.175	0.039
H_2 3-2 S(1)	23864	23859.9	4.27	0.39	0.12	0.140	0.045
H_2 1-0 Q(1)	24066	24069.0	4.27	1.90	0.19	0.681	0.090
H_2 1-0 Q(3)	24237	24241.8	4.27	1.39	0.19	0.497	0.081

Notes: The columns are: (1) Identification of the spectral feature; (2) the rest wavelength of the line (in Å); (3) the peak wavelength of the emission (in Å); (4) the full width at half maximum (FWHM) of the emission (in Å); (5)-(6) the observed flux and its error (in $10^{-18}\,erg\,cm^{-2}\,s^{-1}$ units); (7)-(8) the ratio between the flux at a given transition and the flux at $2.12\,\mu m$, and its error.

Table B.2 - Parameters of the emission lines of the flux-calibrated average H_2 nebular spectrum of the source N47.

Transition	λ_0	λ_{peak}	FWHM	F_λ	σ_F	$F_\lambda/F_{2.12\mu m}$	$\sigma_{F\lambda/2.12}$
H_2 3-2 S(6)	20130	20125.8	2.13	0.34	0.07	0.131	0.028
H_2 1-0 S(2)	20338	20334.9	4.27	0.85	0.05	0.327	0.026
H_2 2-1 S(3)	20649	20629.3	2.13	0.05	0.03	0.019	0.010
H_2 1-0 S(1)	21218	21211.7	2.13	2.59	0.15	1.000	0.084
H_2 1-0 S(0)	22235	22225.2	2.13	0.60	0.06	0.233	0.028
H_2 2-1 S(1)	22477	22470.5	2.13	0.18	0.02	0.068	0.008
H_2 2-1 S(0)	23556	23552.2	4.27	0.05	0.02	0.019	0.009
H_2 1-0 Q(1)	24066	24064.2	2.13	2.54	0.08	0.983	0.067
H_2 1-0 Q(2)	24134	24132.5	4.27	0.84	0.03	0.326	0.023
H_2 1-0 Q(3)	24237	24237.1	2.13	2.47	0.08	0.954	0.065

Notes: The columns are: (1) Identification of the spectral feature; (2) the rest wavelength of the line (in Å); (3) the peak wavelength of the emission (in Å); (4) the full width at half maximum (FWHM) of the emission (in Å); (5)-(6) the observed flux and its error (in $10^{-18}\,erg\,cm^{-2}\,s^{-1}$ units); (7)-(8) the ratio between the flux at a given transition and the flux at $2.12\,\mu m$, and its error.

Table B.3 - Parameters of the emission lines of the flux-calibrated average H_2 nebular spectrum of the source N58.

Transition	λ_0	λ_{peak}	FWHM	F_λ	σ_F	$F_\lambda/F_{2.12\mu m}$	$\sigma_{F\lambda/2.12}$
H_2 3-2 S(6)	20130	20128.2	4.27	2.34	0.11	0.279	0.020
H_2 1-0 S(2)	20338	20335.2	2.13	2.61	0.06	0.311	0.018
H_2 1-0 S(1)	21218	21212.1	2.13	8.38	0.43	1.000	0.073
H_2 3-2 S(4)	21280	21280.4	2.13	0.07	0.03	0.008	0.003
H_2 2-1 S(2)	21542	21536.4	2.13	0.15	0.02	0.018	0.003
H_2 3-2 S(3)	22014	22007.9	2.13	0.13	0.03	0.015	0.004
H_2 1-0 S(0)	22235	22225.6	2.13	2.18	0.18	0.261	0.026
H_2 2-1 S(1)	22477	22470.9	2.13	0.77	0.06	0.091	0.008
H_2 3-2 S(1)	23864	23845.0	2.13	0.15	0.07	0.017	0.009
H_2 1-0 Q(1)	24066	24064.7	2.13	9.11	0.19	1.087	0.061
H_2 1-0 Q(2)	24134	24133.0	4.27	3.36	0.11	0.401	0.025
H_2 1-0 Q(3)	24237	24237.6	2.13	9.09	0.13	1.085	0.058

Notes: The columns are: (1) Identification of the spectral feature; (2) the rest wavelength of the line (in Å); (3) the peak wavelength of the emission (in Å); (4) the full width at half maximum (FWHM) of the emission (in Å); (5)-(6) the observed flux and its error (in $10^{-18}\,erg\,cm^{-2}\,s^{-1}$ units); (7)-(8) the ratio between the flux at a given transition and the flux at $2.12\,\mu m$, and its error.

Table B.4 - Parameters of the emission lines of the flux-calibrated average H_2 nebular spectrum of the source V49.

Transition	λ_0	λ_{peak}	FWHM	F_λ	σ_F	$F_\lambda/F_{2.12\mu m}$	$\sigma_{F\lambda/2.12}$
H_2 3-2 S(6)	20130	20124.9	4.27	0.03	0.01	0.039	0.013
H_2 1-0 S(2)	20338	20331.8	4.27	0.25	0.03	0.320	0.047
H_2 2-1 S(3)	20649	20647.7	4.27	0.02	0.01	0.025	0.009
H_2 3-2 S(5)	20656	20647.7	4.27	0.01	0.01	0.019	0.007
H_2 1-0 S(1)	21218	21208.9	4.27	0.78	0.08	1.000	0.140
H_2 3-2 S(4)	21280	21272.9	4.27	0.01	0.00	0.008	0.003
H_2 2-1 S(2)	21542	21535.4	4.27	0.07	0.02	0.094	0.029
H_2 3-2 S(3)	22014	22002.7	4.27	0.02	0.01	0.027	0.009
H_2 1-0 S(0)	22235	22222.5	4.27	0.13	0.02	0.167	0.033
H_2 2-1 S(1)	22477	22465.8	4.27	0.10	0.01	0.129	0.022
H_2 3-2 S(2)	22870	22860.5	2.13	0.01	0.00	0.007	0.003
H_2 2-1 S(0)	23556	23547.7	4.27	0.02	0.01	0.031	0.010
H_2 3-2 S(1)	23864	23859.2	4.27	0.03	0.01	0.039	0.014
H_2 1-0 Q(1)	24066	24059.8	4.27	0.73	0.05	0.939	0.111
H_2 1-0 Q(2)	24134	24128.1	4.27	0.21	0.02	0.275	0.034
H_2 1-0 Q(3)	24237	24232.6	4.27	0.74	0.04	0.952	0.107

Notes: The columns are: (1) Identification of the spectral feature; (2) the rest wavelength of the line (in Å); (3) the peak wavelength of the emission (in Å); (4) the full width at half maximum (FWHM) of the emission (in Å); (5)-(6) the observed flux and its error (in $10^{-18}\,erg\,cm^{-2}\,s^{-1}$ units); (7)-(8) the ratio between the flux at a given transition and the flux at $2.12\,\mu m$, and its error.

Table B.5 - Parameters of the emission lines of the flux-calibrated average H_2 nebular spectrum of the source W33A.

Transition	λ_0	$\lambda_{\rm peak}$	FWHM	F_λ	$\sigma_{\rm F}$	$F_\lambda/F_{2.12\mu{\rm m}}$	$\sigma_{{\rm F}\lambda/2.12}$
H_2 1-0 S(2)	20338	20339.0	4.27	0.77	0.03	0.164	0.014
H_2 1-0 S(1)	21218	21216.0	2.13	2.70	0.03	0.576	0.048
H_2 3-2 S(4)	21280	21265.1	2.13	0.06	0.03	0.013	0.006
H_2 2-1 S(2)	21542	21536.1	2.13	0.06	0.03	0.012	0.006
H_2 1-0 S(0)	22235	22229.6	2.13	0.75	0.05	0.161	0.017
H_2 2-1 S(1)	22477	22475.0	4.27	0.53	0.05	0.113	0.014
H_2 3-2 S(2)	22870	22878.3	2.13	0.14	0.06	0.029	0.013
H_2 3-2 S(1)	23864	23844.9	2.13	0.55	0.23	0.117	0.051
H_2 1-0 Q(1)	24066	24068.9	4.27	4.10	0.44	0.876	0.118
H_2 1-0 Q(3)	24237	24241.8	4.27	4.68	0.38	1.000	0.116

Notes: The columns are: (1) Identification of the spectral feature; (2) the rest wavelength of the line (in Å); (3) the peak wavelength of the emission (in Å); (4) the full width at half maximum (FWHM) of the emission (in Å); (5)-(6) the observed flux and its error (in $10^{-18}\,{\rm erg\,cm^{-2}\,s^{-1}}$ units); (7)-(8) the ratio between the flux at a given transition and the flux at $2.12\,\mu$m, and its error.

B.3 Construction of the Boltzmann Diagrams of the H_2 emission

This Appendix presents the comparison between three different criteria to evaluate the best A_K value of the H_2 gas associated with the extended emission towards the HMYSOs. The tests were performed towards the spectrum of the extended H_2 emission associated with W33A (Fig. 4.15) due to the larger and considerable differences found between the criteria, as discussed below.

Figure B.15 presents the distribution of the A_K value as a function of the three parameters based on the fitting of the data, they are: the absolute value of the Pearson correlation coefficient (r, left), the standard deviation of the residuals of the fit ($\sigma_{\rm fit}$, middle) and a factor that combines both parameters, $(1-r)\sigma_{\rm fit}$ (right). The middle panel of Fig. B.15 presents the standard deviation of the residuals, $\sigma_{\rm fit}$, as a function of the A_K value. The plot indicates that the best A_K value, obtained when considering the minimum value of $\sigma_{\rm fit}$, is roughly two magnitudes smaller, $A_K = 4.40$ mag. The observed distribution of $\sigma_{\rm fit}$ values is non-symmetric and rapidly decreases around the best A_K value. Finally, in order to take advantage of both independent parameters, a third parameter was proposed to guarantee that no local minimum or maximum values are introducing any bias on the results. Thus, a hybrid parameter given by $(1-r)\sigma_{\rm fit}$ should be minimum at the best A_K value, $A_K = 4.40$ mag (right), as similar as obtained when only considering the $\sigma_{\rm fit}$ value. The $(1-r)$ factor acts more effectively in the outskirts of A_K range but has smaller effects when approaching to the best A_K value.

Figure B.16 presents the H_2 diagrams obtained when considering the best A_K value derived from the method using only the Pearson correlation coefficient (in the left) and when using the hybrid parameter, $(1-r)\sigma_{\rm fit}$ (right). The left panel indicates that the adoption of the maximum value of the Pearson correlation coefficient lead to an A_K value of 6.15 mag. The corresponding Boltzmann diagram is presented in the left panel of Fig. B.16. The diagram exhibits divergent physical parameters derived from each set of vibrational transitions (e.g., excitation temperatures ranging from $\sim$ 3500 to $\sim$ 15000 K) and, therefore, indicates that the A_K value is not physically meaningful. The right panel presents the Boltzmann diagram obtained when considering $A_K = 4.40$ mag. Although the data from $v = 2$–1 is underestimated, the plot clearly shows that the determination of the physical parameters converges when using the the lower-v dataset and the whole dataset ($T_{\rm ex}$ $\sim$3700 K and

$\log(N_{\rm tot}) \approx 20.8$ dex). Therefore, either the usage of the hybrid parameter or simply the $\sigma_{\rm fit}$ value leads to a better determination of the physical parameters of the H_2 gas rather than the usage of the linear Pearson coefficient alone. The usage of the hybrid parameter as the criterion to determine the best A_K value led to more robust determinations of the pixel-to-pixel calculations presented in this thesis (see Sect. 4.3.4.2).

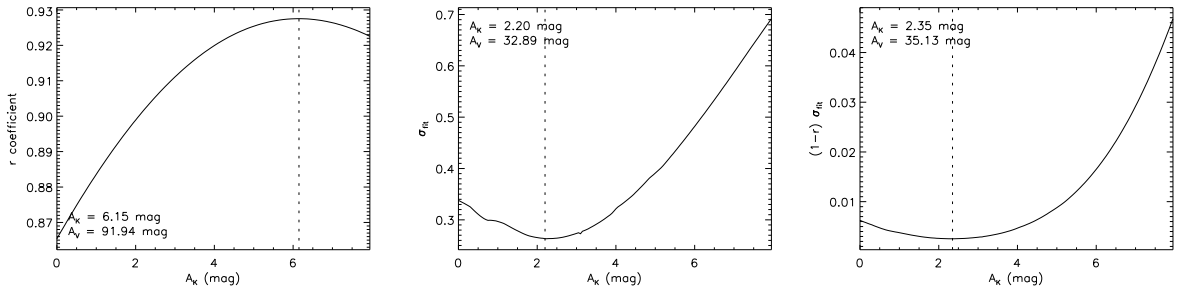


Figure B.15: Comparison between the distribution of the three parameters used for deriving the best A_K value for the H_2 emission of W33A. See definition of each parameter in the text.

Figure B.16 presents the H_2 diagrams obtained when considering the best A_K value derived from the method using only the Pearson correlation coefficient (in the left) and when using the hybrid parameter, $(1-r)\sigma_{\rm fit}$ (right). The major difference is found when computing the physical parameters of the lower υ transitions (e.g., υ=1–0). The fit diverges in the left panel while it displays reasonable values in the right panel. Therefore, the usage of the hybrid parameter leads to more robust results when fitting the H_2 diagrams, avoiding misleading determinations on the pixel-to-pixel calculations presented in this Thesis.

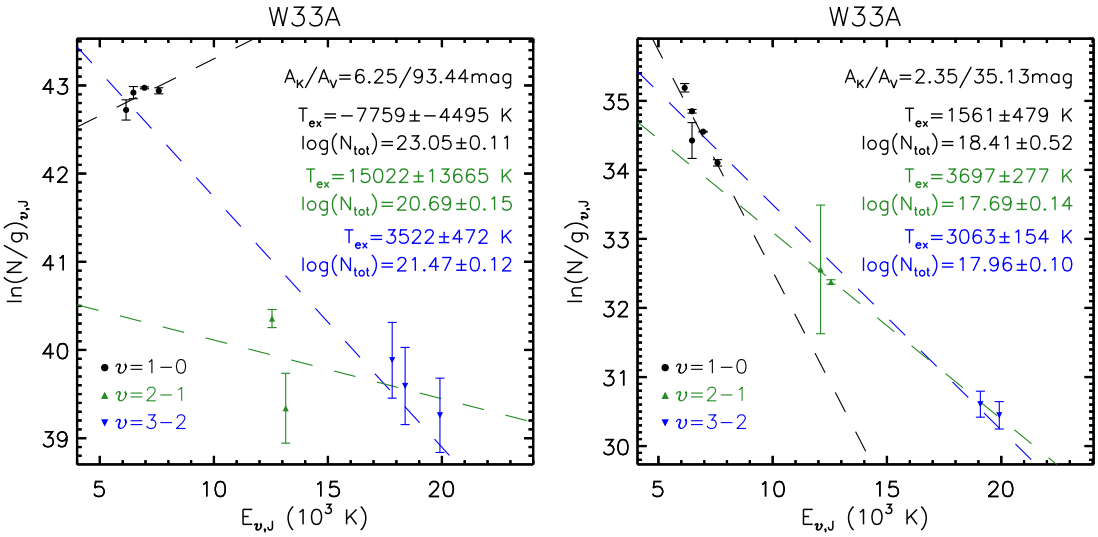


Figure B.16: Boltzmann diagram of the H_2 gas associated with the extended emission detected towards the source W33A, assuming distinct values for the K-band reddening. In the left panel, the extinction in the K-band was adopted as 6.25 mag, closer to the best A_K value derived assuming the minimum value of the linear Pearson coefficient (r, see left panel of Fig. B.15). In the right panel, $A_K = 2.35$ mag, derived in the right panel of Fig. B.15, was assumed instead. The dashed lines indicate the best fit to the data using the $\epsilon = 1$–0 transitions only (black), the higher $\epsilon = 2$–1 and 3–2 transitions (green) and the whole dataset (blue). The corresponding temperatures and total column density are indicated in the upper-right corner of each panel.

B.4 Boltzmann Diagrams of the H_2 emission

Analysis of the excitation diagram of the H_2 molecule based on the intensity of different transitions found towards the brightest H_2 jet associated with the N35 source.

Top panel: Variation of the A_K value from 0 to 2.3 mag, in steps of 0.1 mag. Points from different vibrational levels are shown in distinct colours (red: $\epsilon = 1$–0; blue: $\epsilon = 2$–1; green: $\epsilon = 3$–2;). The best fit to the data is indicated by the filled black line. The original dataset ($A_K = 0$ mag) is shown as black dots in each panel, together with its best fit (indicated by the dashed black line). Values of A_K, A_V, the linear Pearson coefficient of the data (r) and the corresponding $T_{\rm ex}$ using the de-reddened points are indicated in each plot. Bottom left panel: linear Pearson coefficient versus the A_K value. The best A_K value, indicated by the dashed black line, occurs at the maximum value of r. Bottom right panel: De-reddened excitation diagram, assuming A_K =1.30 mag. The dashed lines indicate the best fit to the data using the $\epsilon = 1$–0 transitions only (black), the $\epsilon = 2$–1 and 3–2 transitions (green) and the whole dataset (blue). The corresponding temperatures and total column density are indicated in the upper-right corner of each panel.

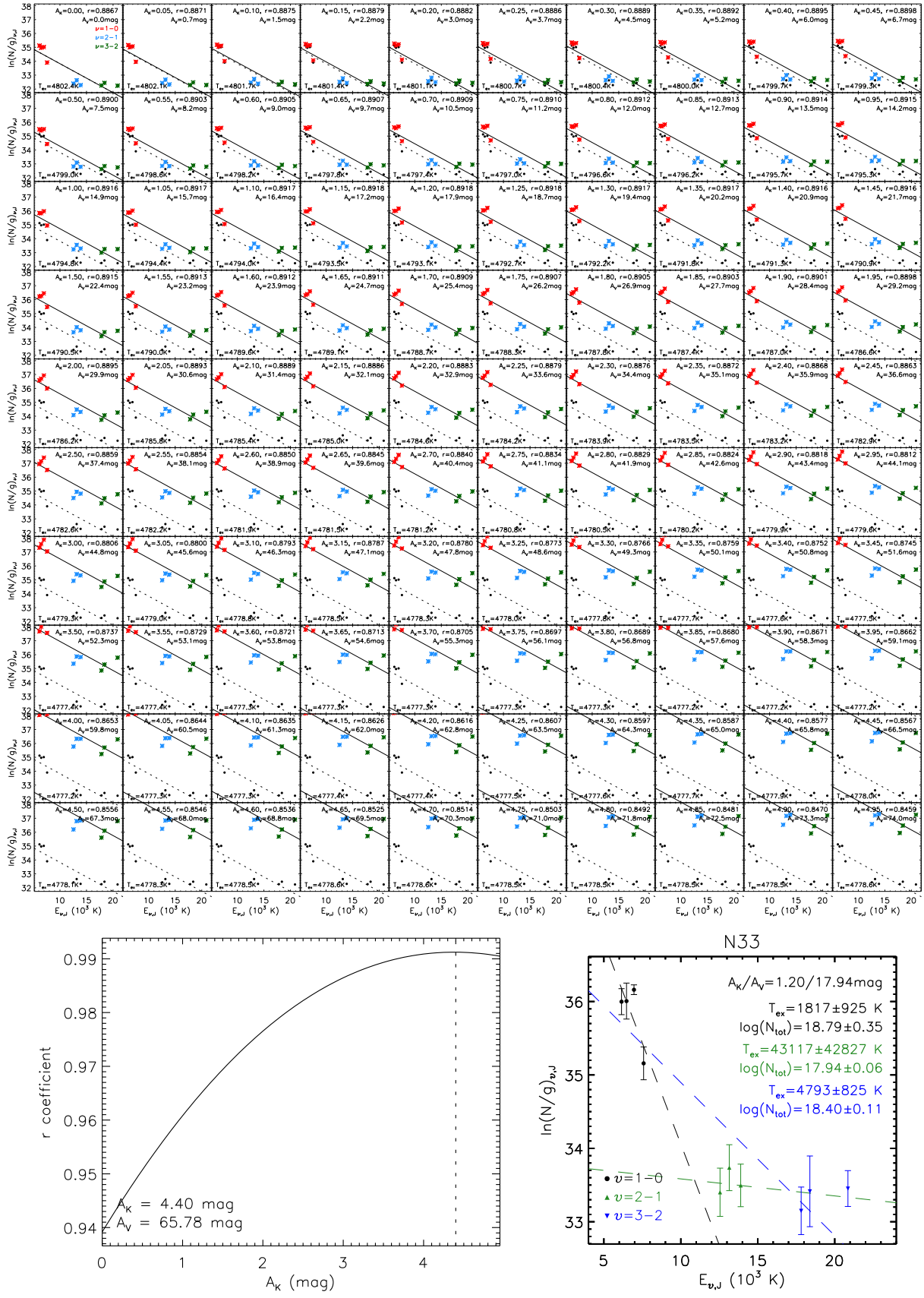


Figure B.17: Excitation diagram of the H_2 emission towards the brightest jet associated with the N33 source.

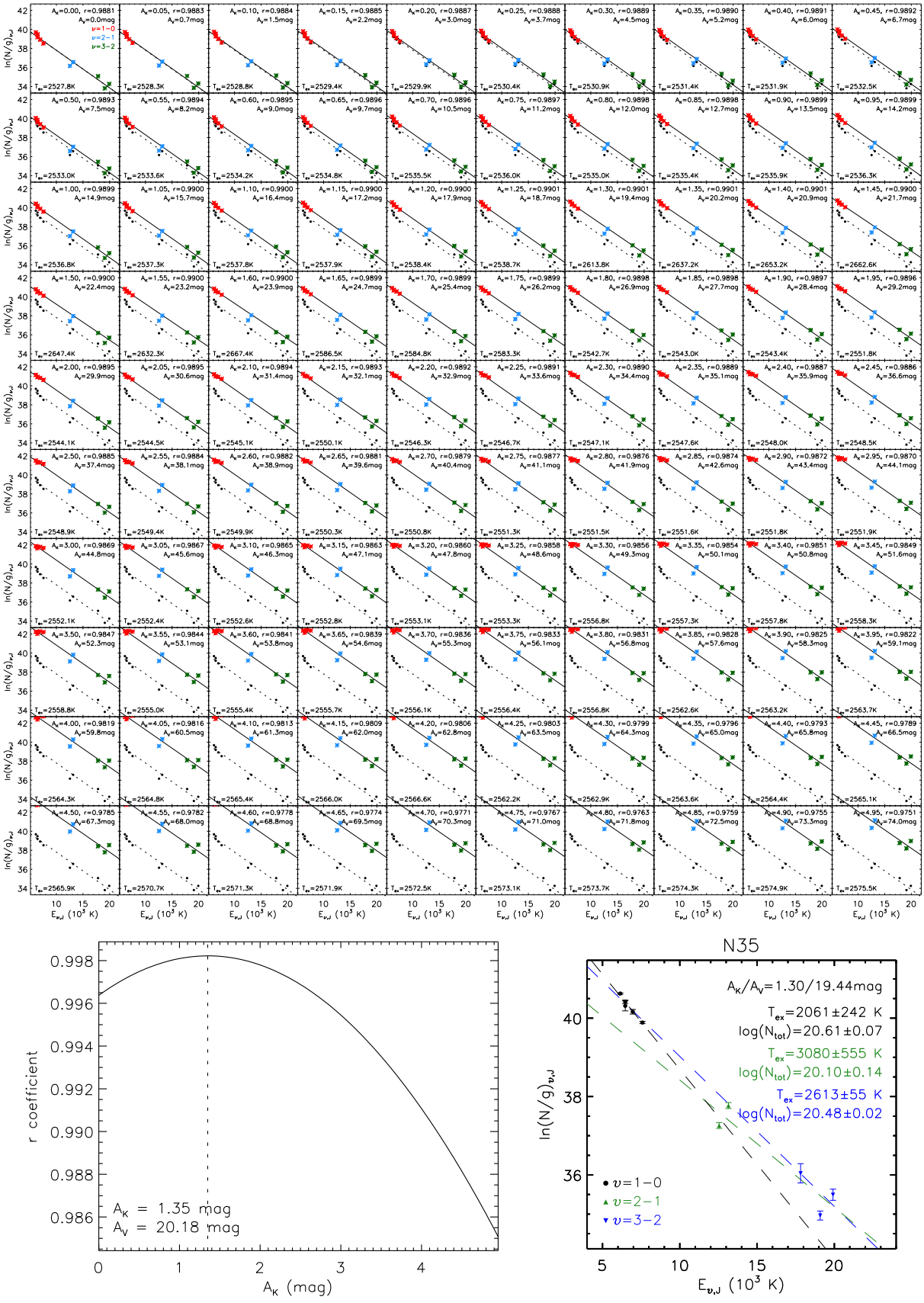


Figure B.18: Excitation diagram of the H_2 emission towards the brightest jet associated with the N35 source.

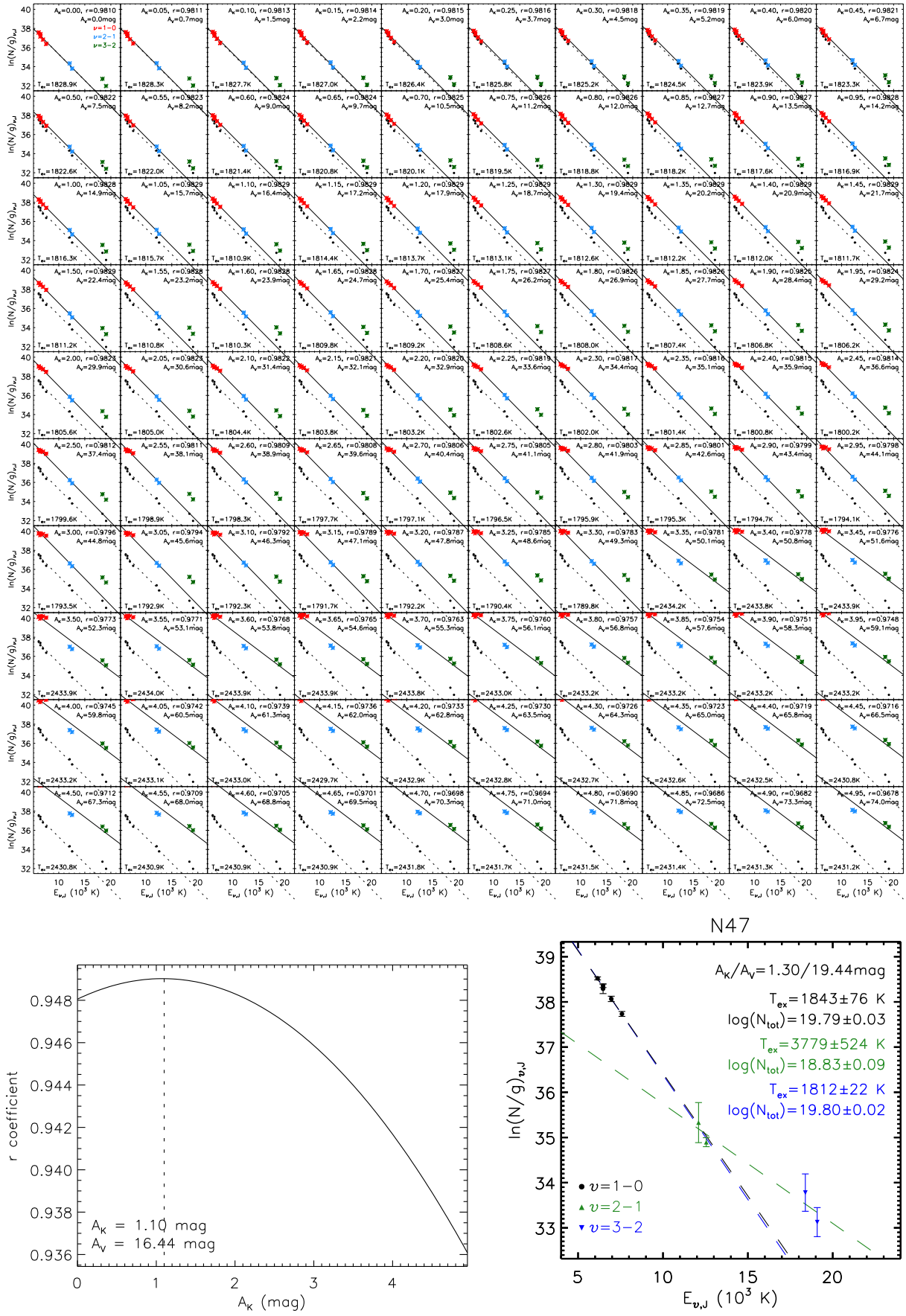


Figure B.19: Excitation diagram of the H_2 emission towards the brightest jet associated with the N47 source.

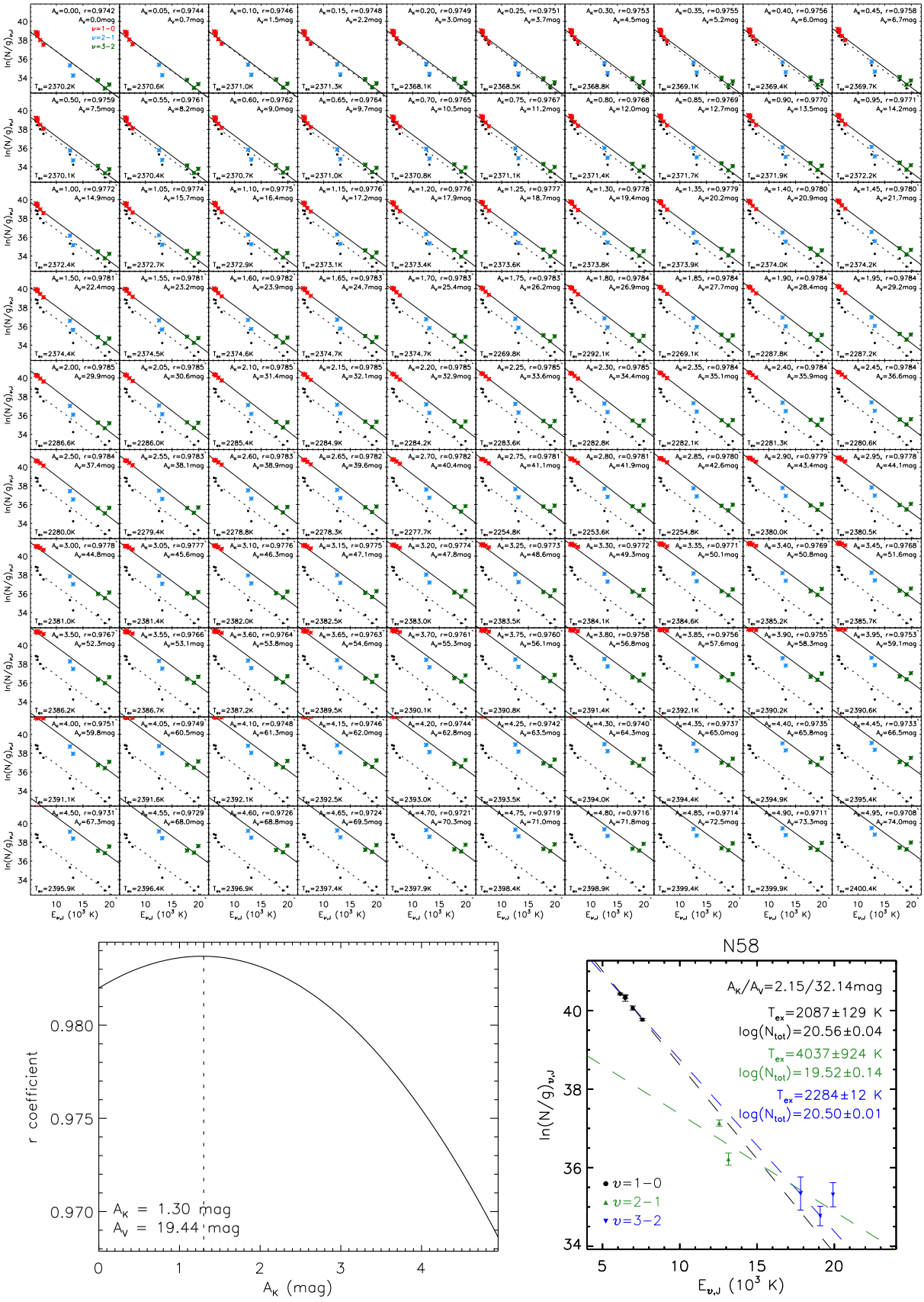


Figure B.20: Excitation diagram of the H_2 emission towards the brightest jet associated with the N58 source.

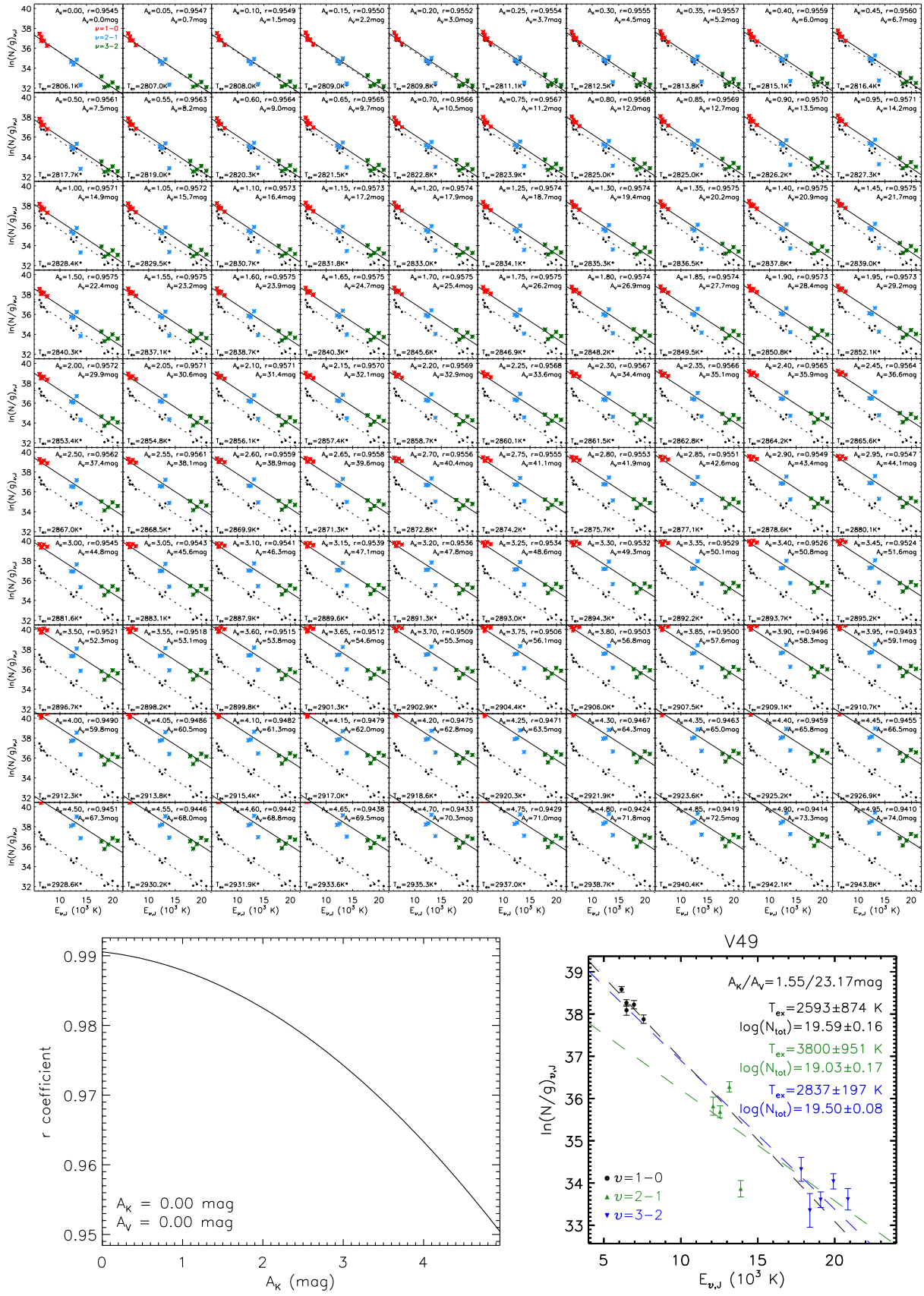


Figure B.21: Excitation diagram of the H_2 emission towards the brightest jet associated with the V49 source.

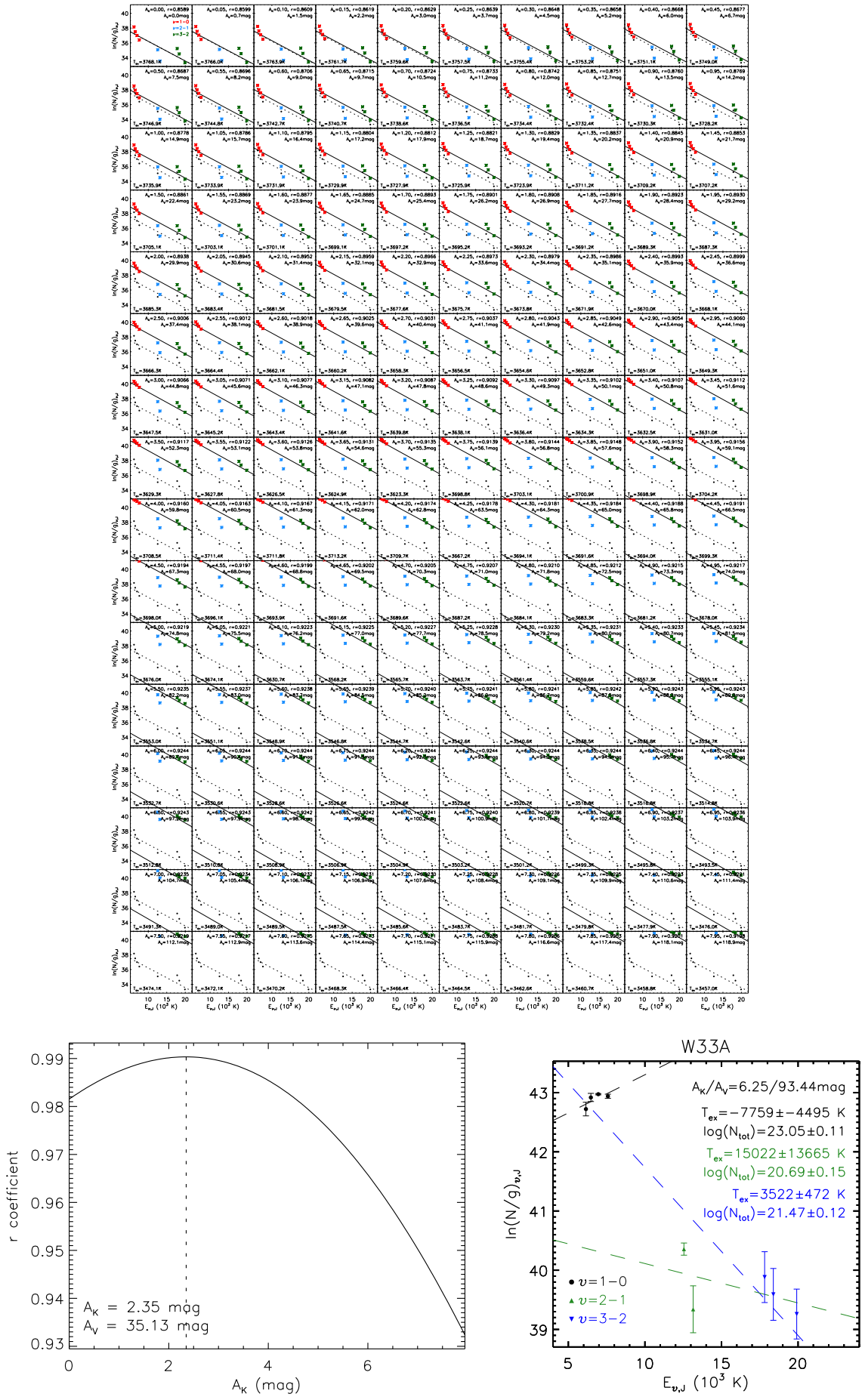


Figure B.22: Channel maps of the CO (2–1) emission towards the source W33A.

B.5 Pixel-to-pixel analysis of the H_2 emission

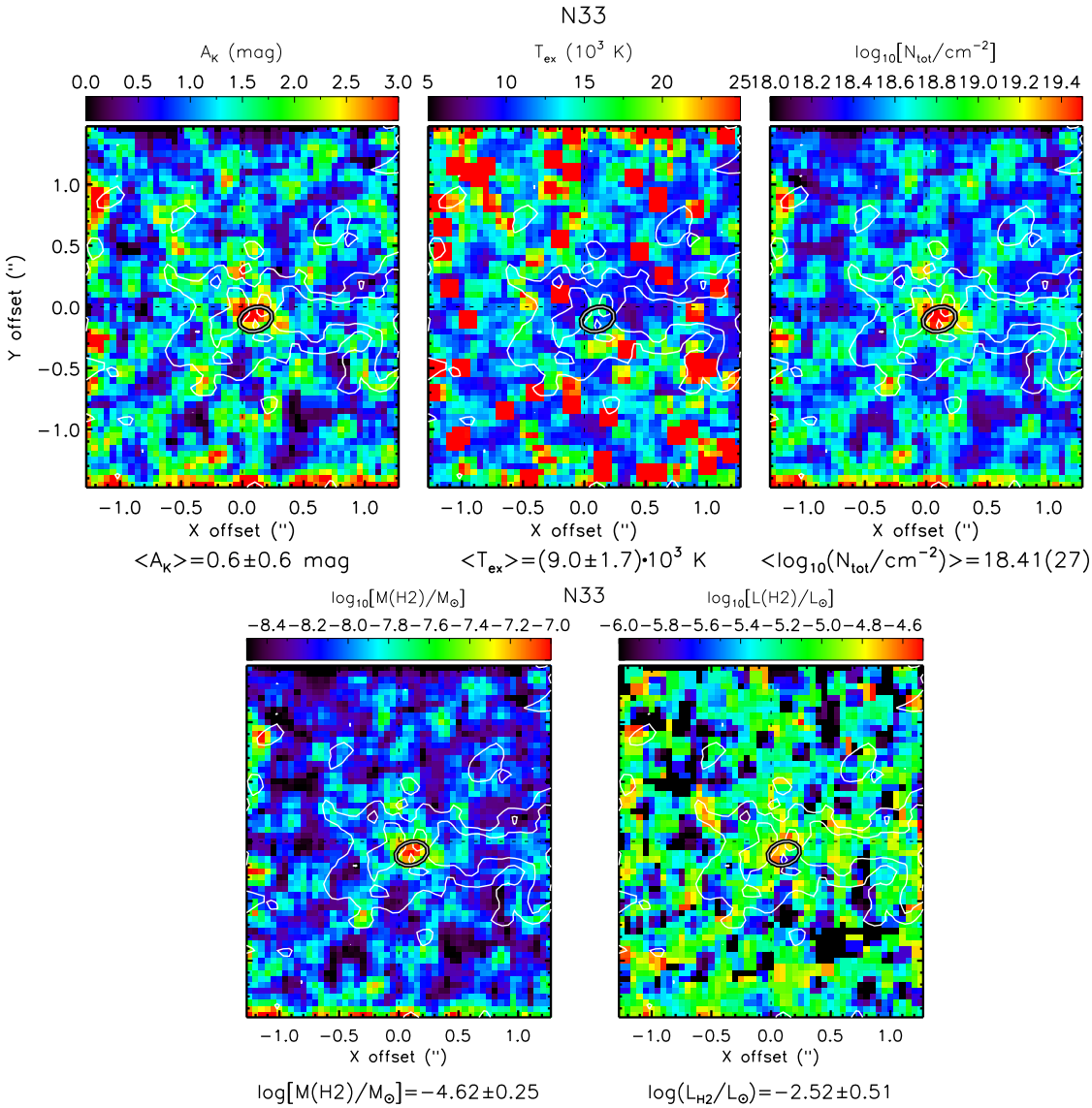


Figure B.23: Upper panels: Spatial distribution of the K-band reddening (left), the excitation temperature (middle), the total column density (right) of the H_2 emission associated with the source N33. Bottom panels: spatial distribution of the mass (left) and the luminosity (right) of the H_2 emission. The position of the main source is indicated by the black ellipsoid. To improve the visualisation, a boxcar average filter with a width of 2 pixels was applied to each map. The dynamical range of the values used for displaying each map is indicated by the corresponding colour-bar. The white contours are placed at 3^n-σ (with $n = 0, 1, 2, ...$) levels of the H_2 1–0 S(1) emission at 2.12 μm. For A_K, $T_{\rm ex}$ and $N_{\rm tot}$, the median values of the spaxels within the 3-σ contour are indicated below each map. For $M(H_2)$ and $L(H_2)$, the values presented below each map are integrated over the 1-σ contour. The error on each value is calculated as the standard deviation of the maps.

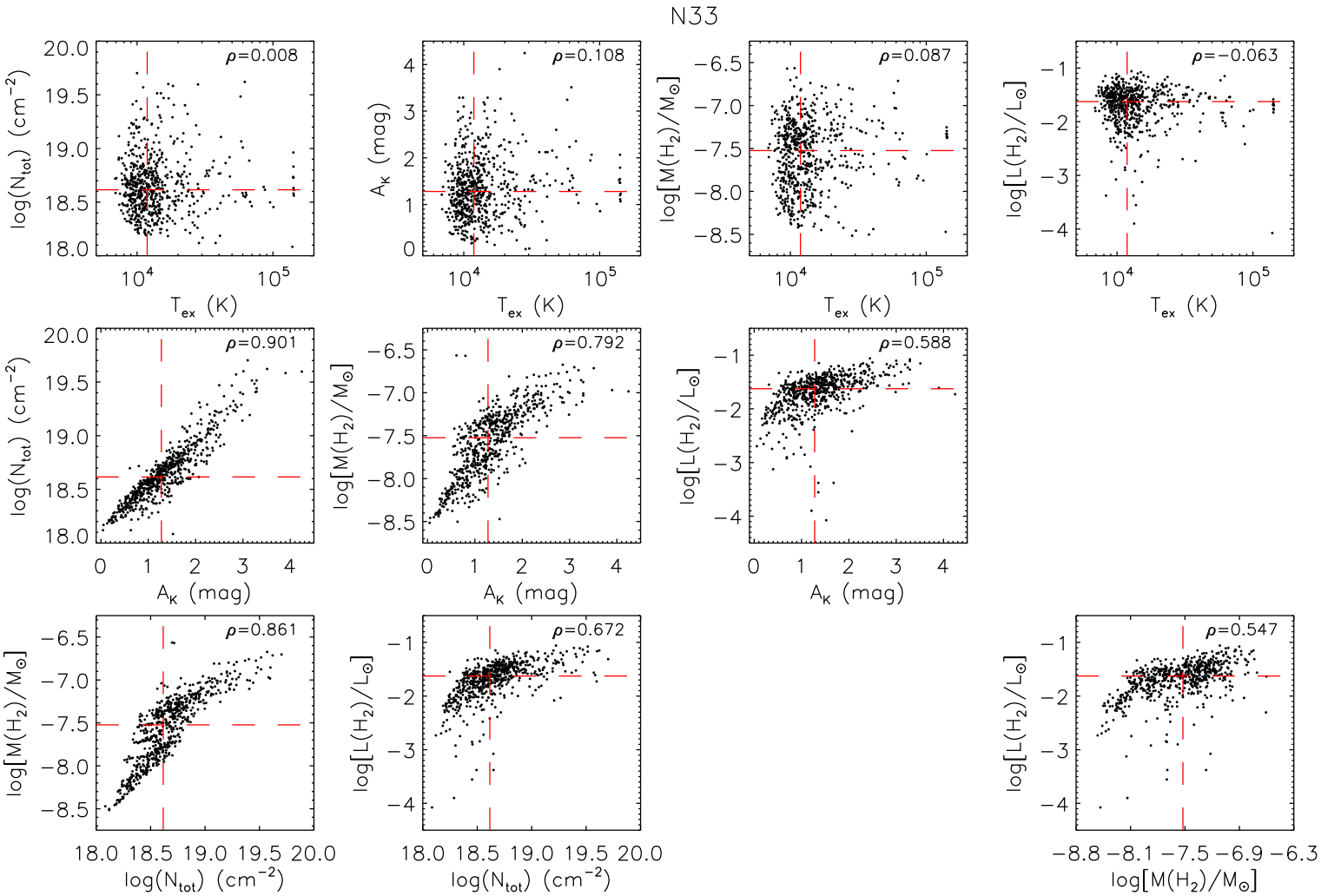


Figure B.24: Correlation between the physical parameters of the H_2 gas for source N33, derived for each spaxel of the NIFS datacube. The Spearman test (ρ) is indicated on the top right corner of each panel. The mean values on the x- and y-axis are indicated as red dashed lines.

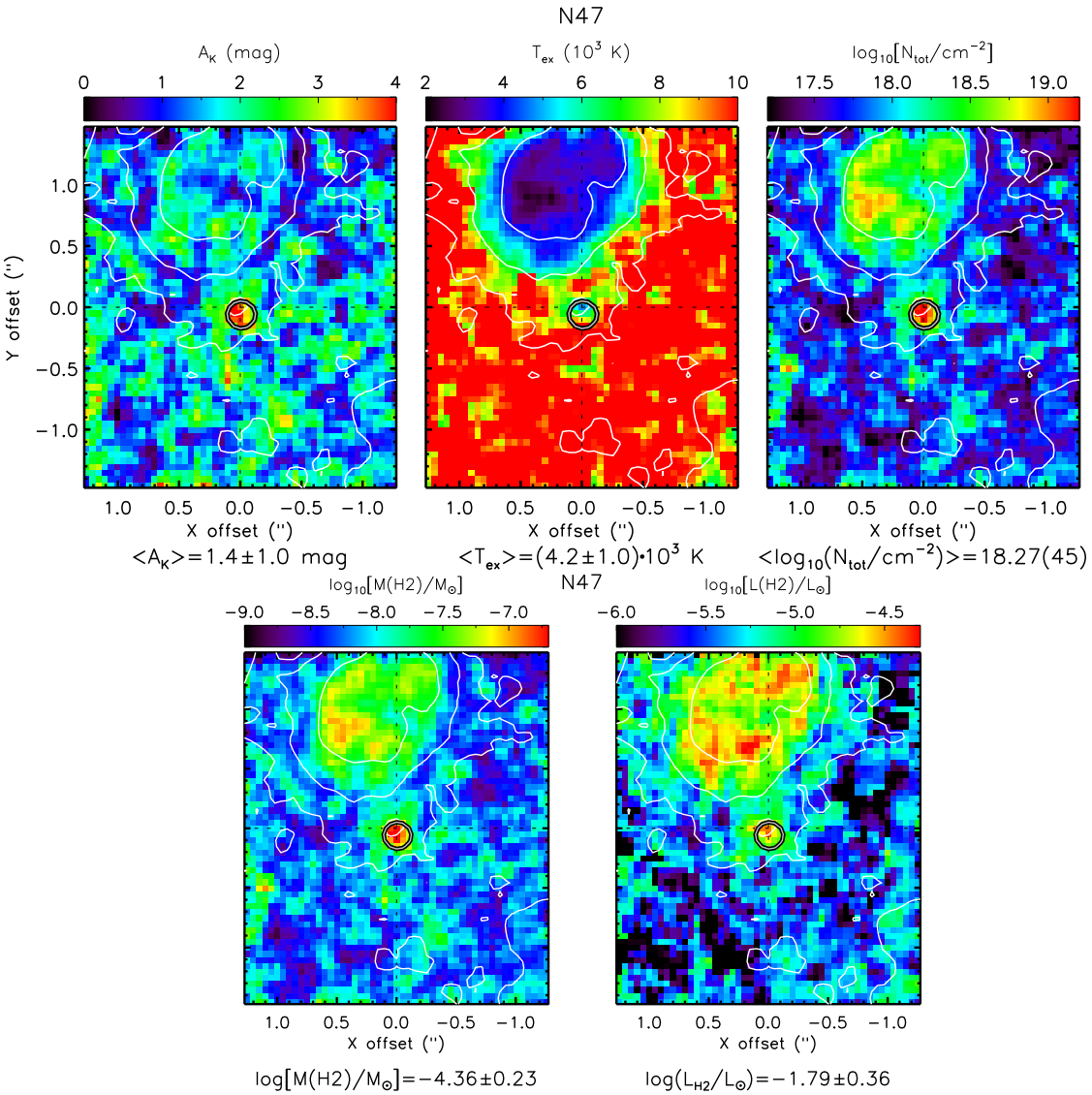


Figure B.25: Same as Fig. B.23 but for source N47.

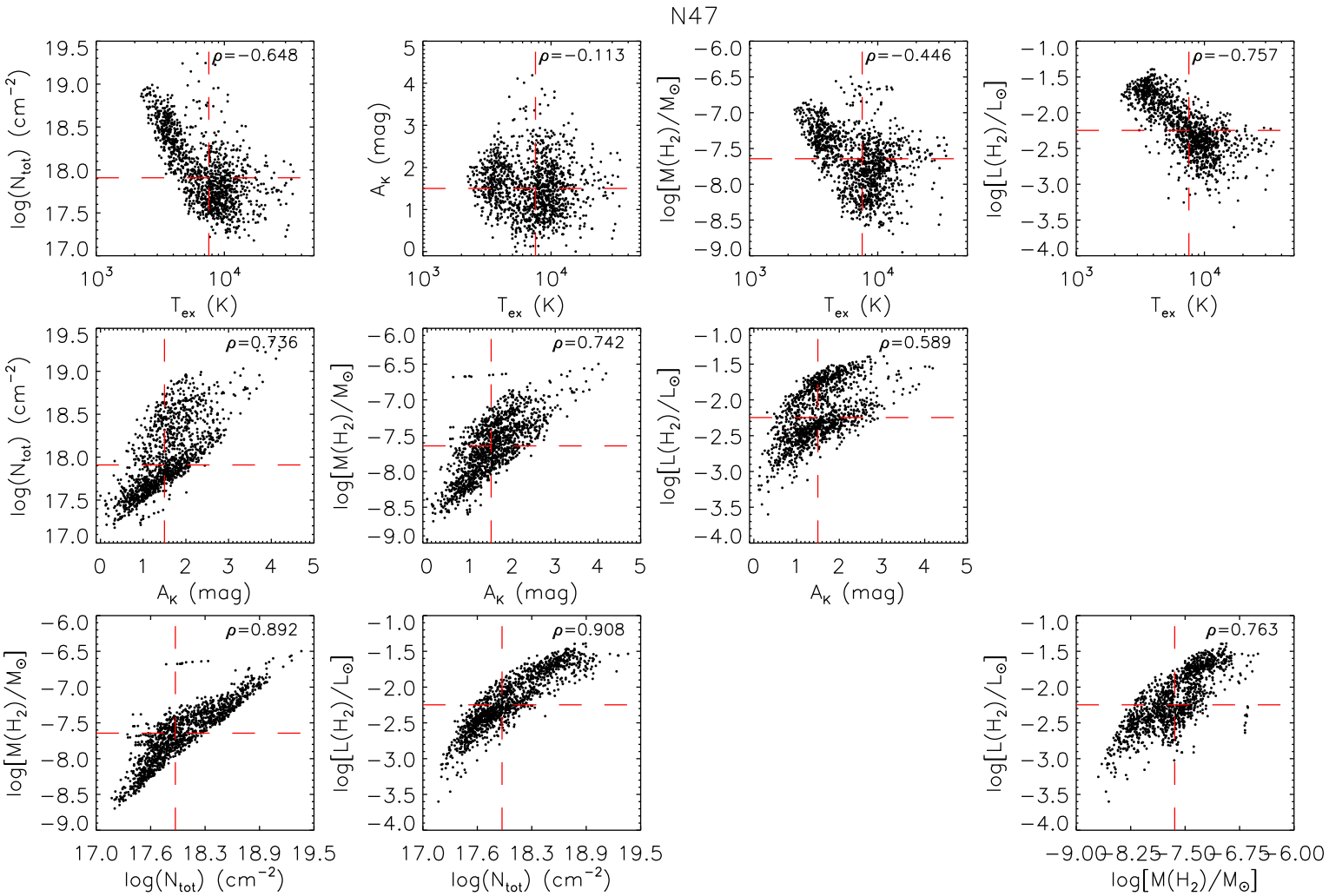


Figure B.26: Correlation between the physical parameters of the H_2 gas for source N47, derived for each spaxel of the NIFS datacube. For a full description of the image, see Fig. B.24.

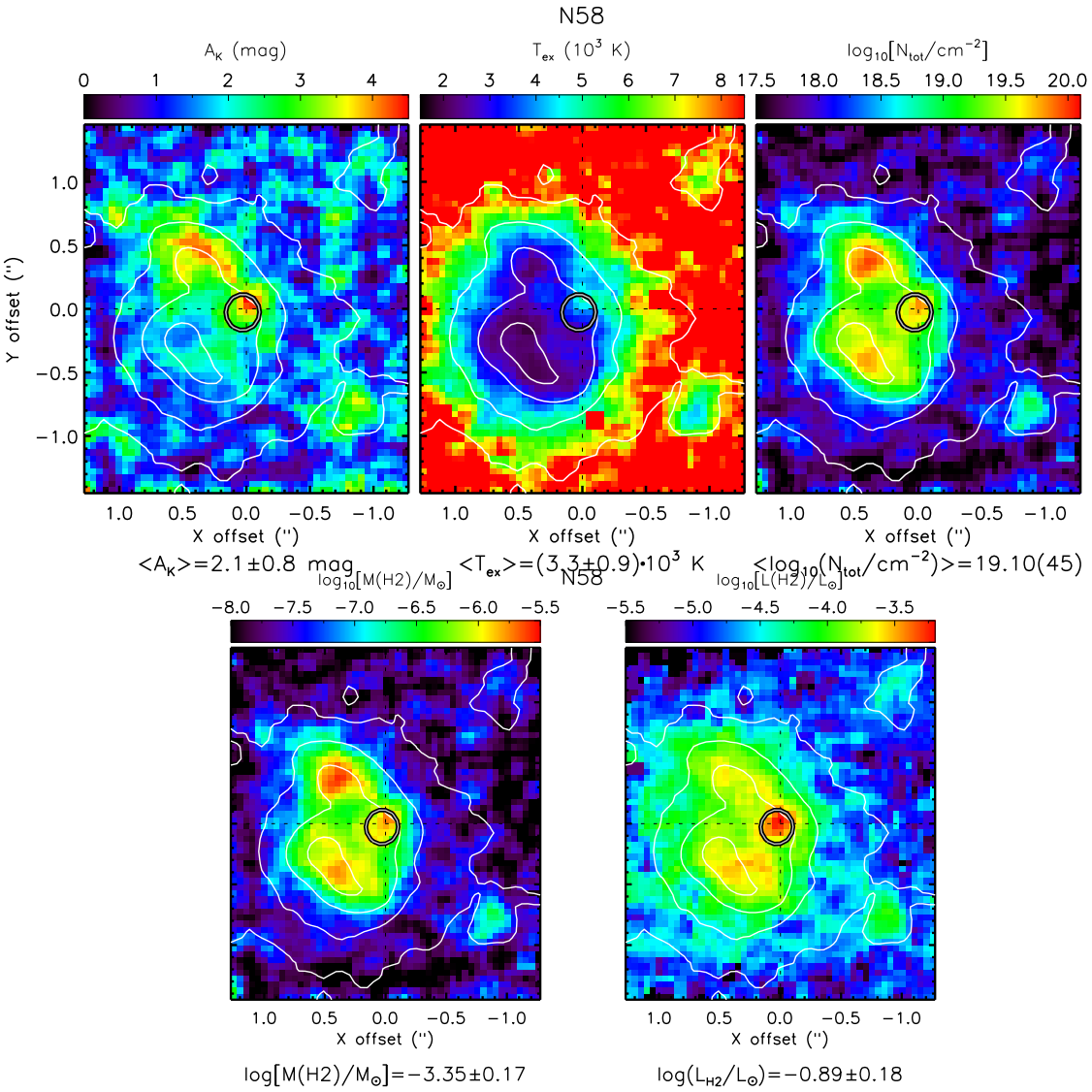


Figure B.27: Same as Fig. B.23 but for source N47.

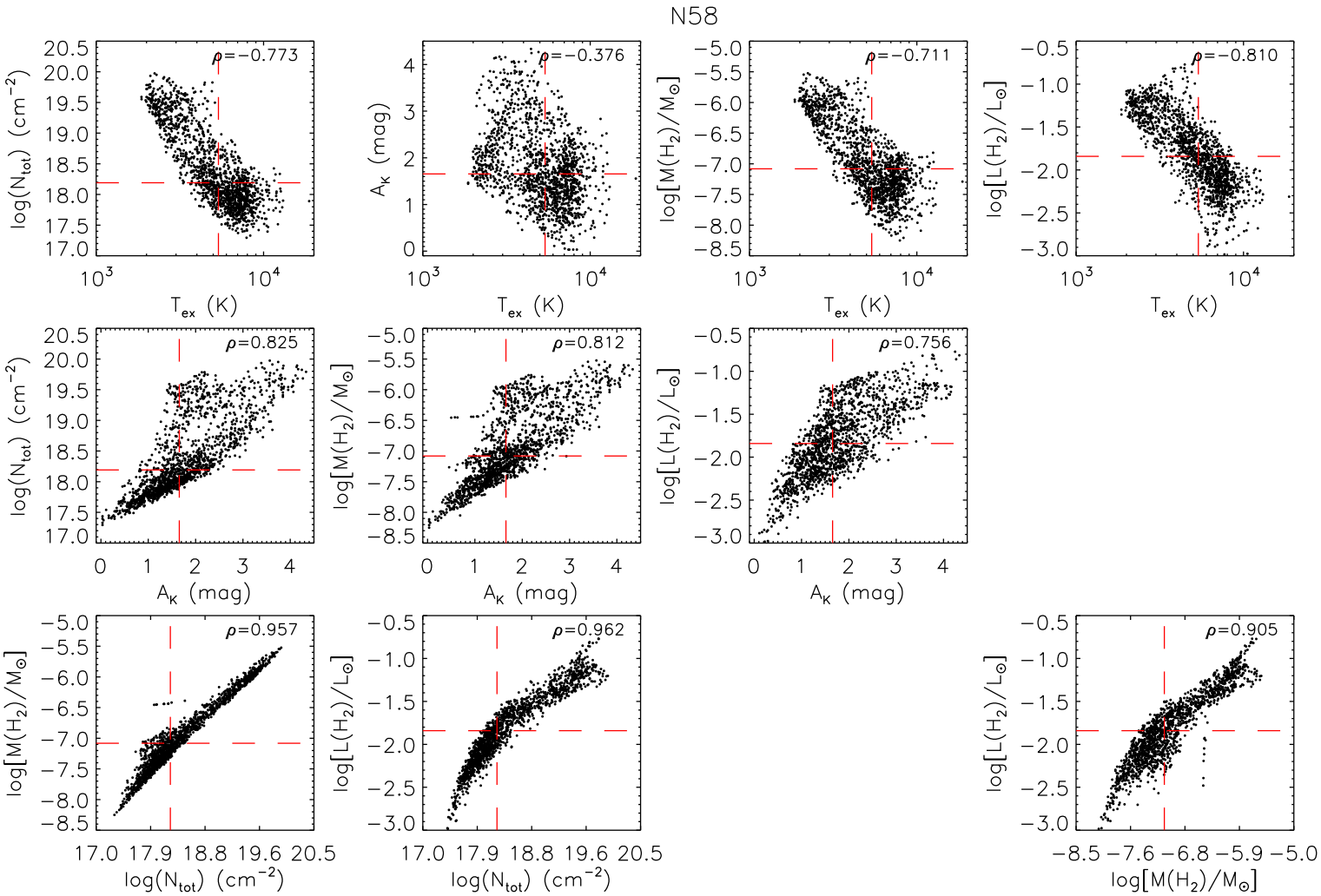


Figure B.28: Correlation between the physical parameters of the H_2 gas for source N58, derived for each spaxel of the NIFS datacube. For a full description of the image, see Fig. B.24.

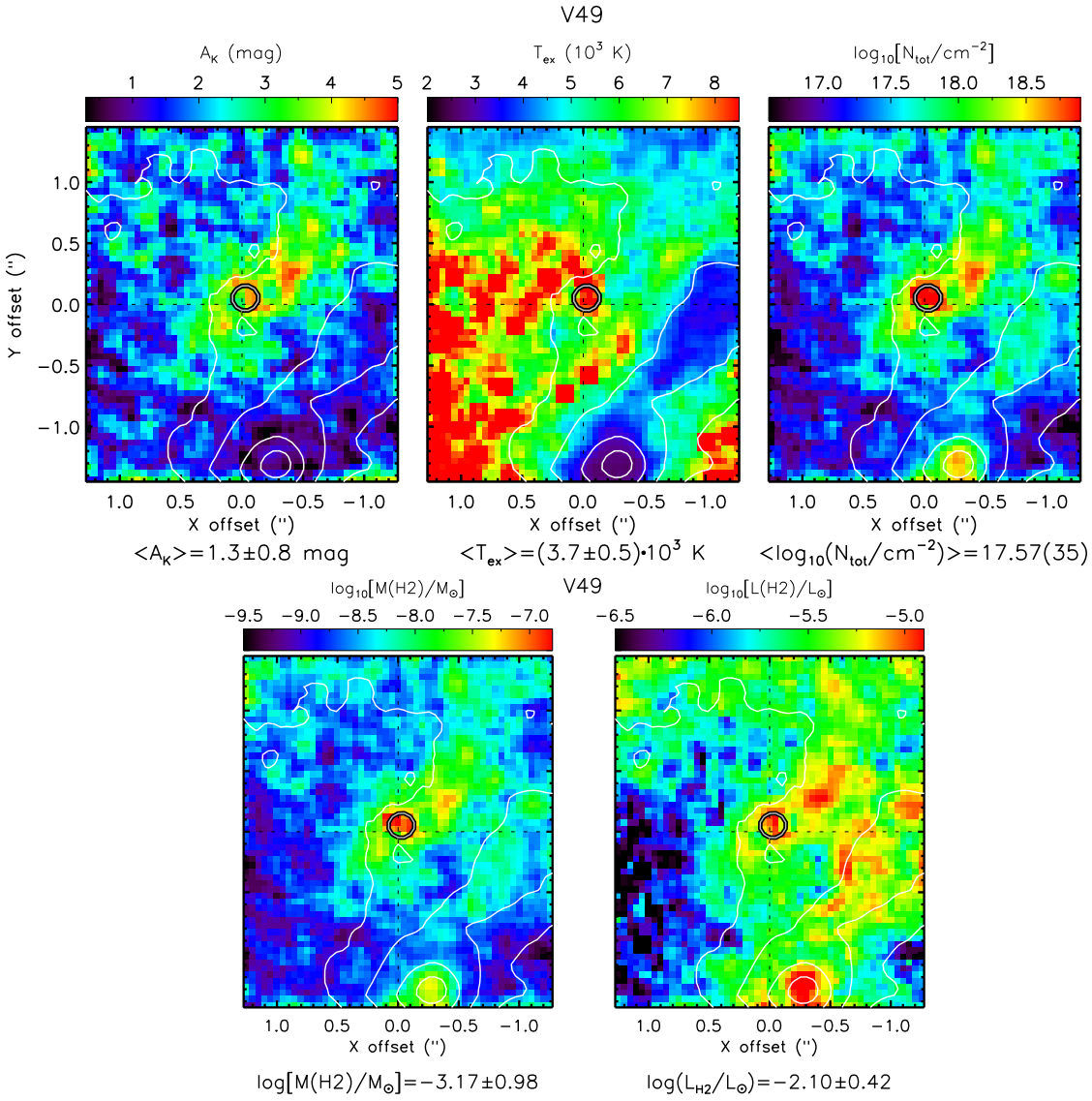


Figure B.29: Same as Fig. B.23 but for source V49.

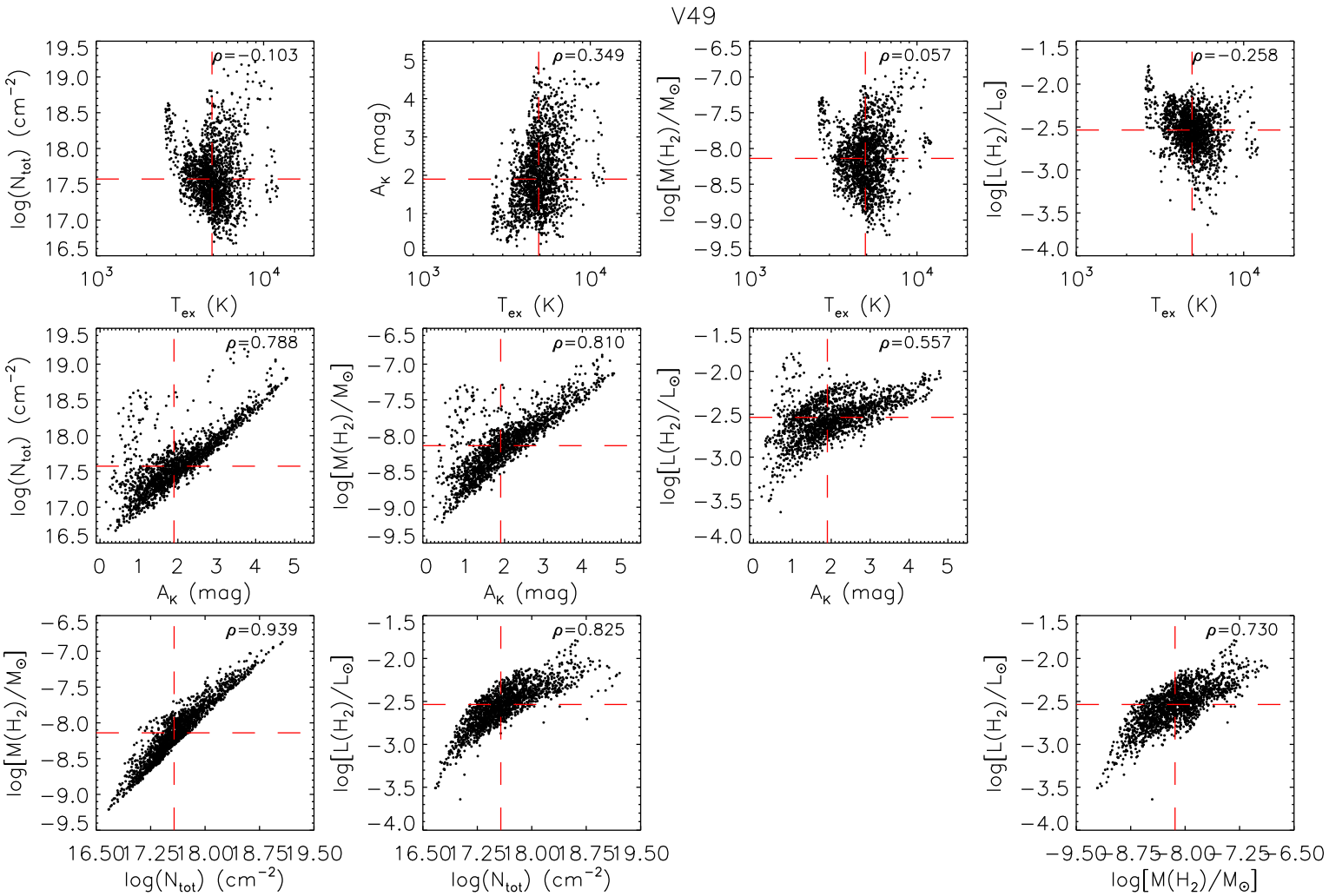


Figure B.30: Correlation between the physical parameters of the H_2 gas for source V49, derived for each spaxel of the NIFS datacube. For a full description of the image, see Fig. B.24.

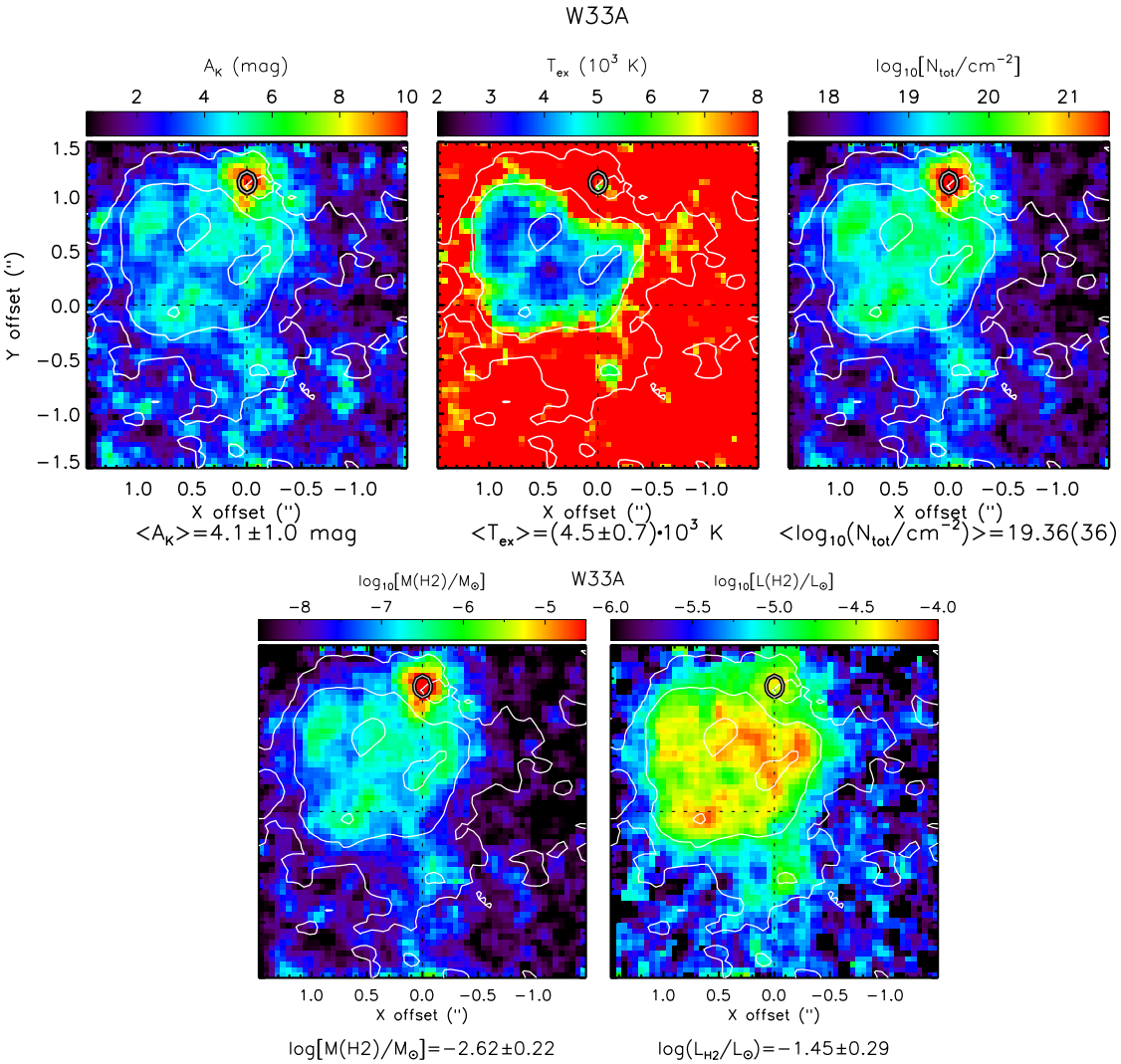


Figure B.31: Same as Fig. B.23 but for source V49.

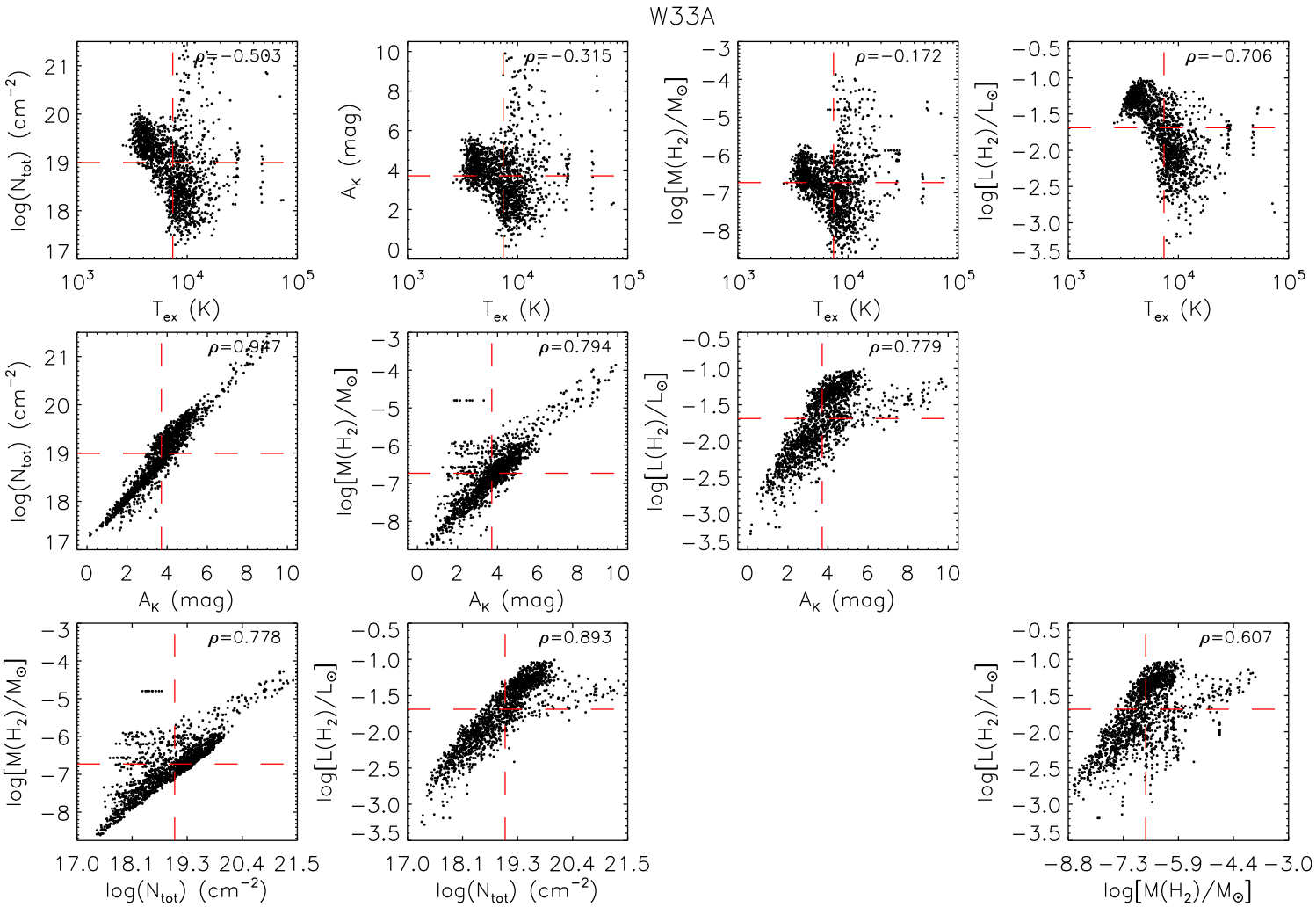


Figure B.32: Correlation between the physical parameters of the H_2 gas for source W33A, derived for each spaxel of the NIFS datacube. For a full description of the image, see Fig. B.24.

B.6 Spectro-astrometry of the Brγ feature of the telluric standard stars

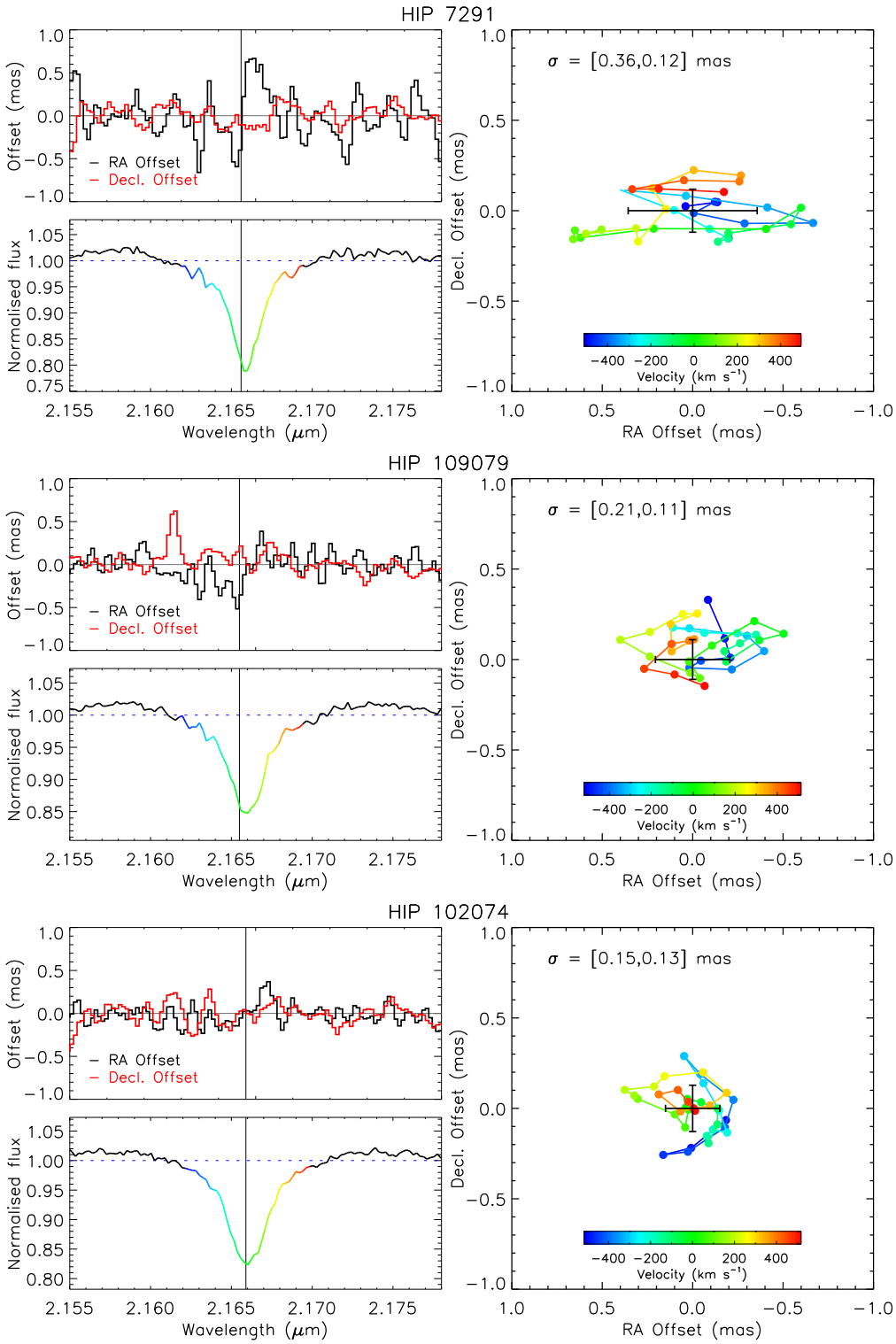


Figure B.33: Similar as Fig. 4.25, but for the telluric standard stars of sources N47 (top), N58 (middle) and V49 (bottom). The error bars placed at the (0,0) position indicate the standard deviation on each axis, whose values are indicated in the top left region of the map.

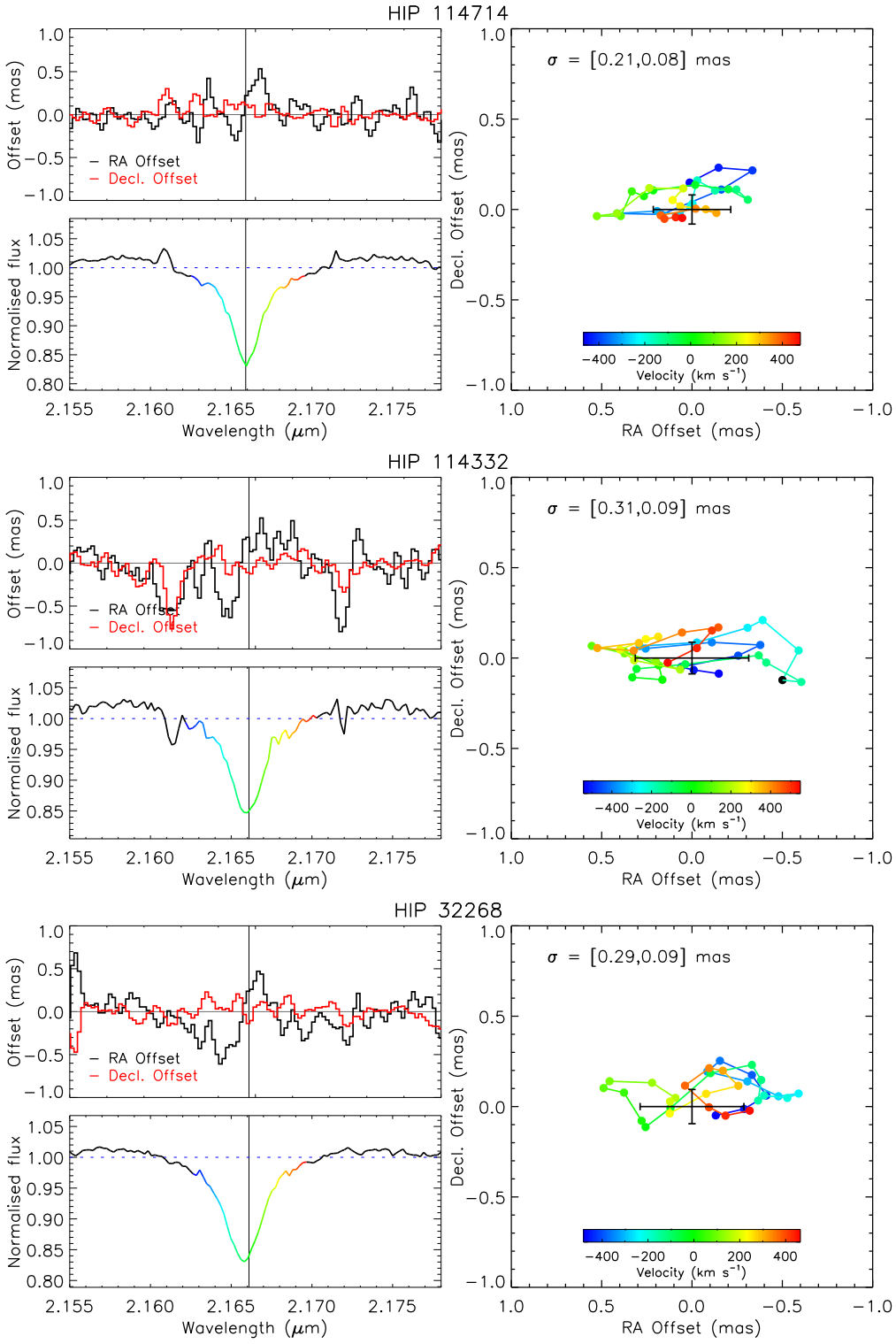


Figure B.33: Continued. Sources N35 (top), N16 (middle) and N21 (bottom).

Appendix C

Submitted and Published Articles

C.1 ATLASGAL Article (submitted to A&A)

Astronomy & Astrophysics manuscript no. champ_top100_paper_one
March 15, 2018

ATLASGAL-selected high-mass clumps in the inner Galaxy. VII. Characterisation of mid-*J* CO emission

F. Navarete[1,2], S. Leurini[3,1], A. Giannetti[4,1], F. Wyrowski[1], J. S. Urquhart[5,1], C. König[1], T. Csengeri[1], R. Güsten[1], A. Damineli[2], and K. M. Menten[1]

[1] Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, 53121 Bonn, Germany
e-mail: fnavarete@mpifr-bonn.mpg.de
[2] Universidade de São Paulo, Instituto de Astronomia, Geofísica e Ciências Atmosféricas, Departamento de Astronomia, Rua do Matão 1226, Cidade Universitária São Paulo-SP, 05508-090, Brazil
[3] INAF - Osservatorio Astronomico di Cagliari, Via della Scienza 5, I-09047, Selargius (CA)
[4] INAF - Istituto di Radioastronomia & Italian ALMA Regional Centre, Via P. Gobetti 101, I-40129 Bologna, Italy
[5] Centre for Astrophysics and Planetary Science, The University of Kent, Canterbury, Kent CT2 7NH, UK

ABSTRACT

Context. High-mass stars are **formed within massive molecular clumps**, where a large number of stars form close together. The evolution of the clumps with different masses and luminosities is mainly regulated by its high-mass stellar content and the formation of such objects is still not well understood.
Aims. **In this work, we characterise the mid-*J* CO emission in a statistical sample of 99 clumps** (TOP100) **selected from the ATLASGAL survey that are representative of the Galactic proto-cluster population.**
Methods. High-spatial resolution APEX-CHAMP$^+$ maps of the CO (6–5) and CO (7–6) transitions were obtained and combined with additional single-pointing APEX-FLASH$^+$ spectra of the CO (4–3) line. The data were convolved to a common angular resolution of 13.″4. We analysed the line profiles by fitting the spectra with up to three Gaussian components, classified as narrow or broad, and computed CO line luminosities for each transition. **Additionally, we defined a distance-limited sample of 72 sources within 5 kpc to check the robustness of our analysis against beam dilution effects. We study the correlations of the line luminosities and profiles for the three CO transitions with the clump properties and investigate if and how they change as a function of the evolution.**
Results. **All sources were detected above 3-σ in all three CO transitions and most of the sources exhibit broad CO emission likely associated with molecular outflows. We found that the extension of the mid-*J* CO emission is correlated with the size of the dust emission traced by the Herschel-PACS 70 μm maps. The CO line luminosity ($L_{\rm CO}$) is correlated with the luminosity and mass of the clumps. However, it does not correlate with the luminosity-to-mass ratio.**
Conclusions. **The dependency of the CO luminosity with the properties of the clumps is steeper for higher-*J* transitions. Our data seem to exclude that this trend is biased by self-absorption features in the CO emission, but rather suggest that different *J* transitions arise from different regions of the inner envelope. Moreover, high-mass clumps show similar trends in CO luminosity as lower mass clumps, but are systematically offset towards larger values, suggesting that higher column density and/or temperature (of unresolved) CO emitters are found inside high-mass clumps.**

Key words. stars: formation – stars: protostars – ISM: molecules – ISM: kinematics and dynamics – line: profiles

1. Introduction

High-mass stars are responsible for the dynamical and chemical evolution of the interstellar medium and of their host galaxies by injecting heavier elements and energy in their surrounding environment by means of their strong UV emission and winds. Despite their importance, the processes that lead to the formation of high-mass stars are still not well understood (Zinnecker & Yorke 2007).

Observations at high-angular resolution have confirmed a high **degree of multiplicity for** high-mass stars, suggesting these objects are not formed in isolated systems (Grellmann et al. 2013). The same scenario is **supported** by three-dimensional simulations of high-mass star formation (Krumholz et al. 2009; Rosen et al. 2016). These objects are formed **on** a relatively short timescale ($\sim 10^5$ yr), requiring large accretion rates **($\sim 10^{-4}$ M$_\odot$ yr^{-1}, Hosokawa & Omukai 2009)**. **Such** conditions can only be achieved in the densest clumps in molecular clouds, with sizes of $\lesssim 1$ pc and masses of **order** 100-1000 M$_\odot$ **(Bergin & Tafalla 2007)**. **These** clumps are associated with large visual extinctions, thus observations at long wavelengths are required to study their properties and the star formation **process**.

After molecular hydrogen (H_2), which is difficult to observe directly in dense cold gas, carbon monoxide (CO) is the most abundant molecular species. Thus, rotational transitions of CO are commonly used to investigate the physics and kinematics of star-forming regions **(SFRs)**. Traditionally, observations of CO transitions with low angular momentum quantum number J from $J = 1$–0 to 4–3 (here defined as low-J transitions) have been used for this purpose **(e.g., see Schulz et al. 1995, Zhang et al. 2001 and Beuther et al. 2002)**. **These lines have upper level energies, $E_{\rm u}$, lower than 55 K** and are easily excited at relatively low temperatures and moderate densities. **Therefore, low-*J* CO lines are not selective tracers of the**

densest regions of SFRs, but are contaminated by emission from the ambient molecular cloud. On the other hand, **higher-***J* CO transitions are less contaminated by ambient gas emission and likely probe the warm gas directly associated with embedded young stellar objects (YSOs). In this paper we make use of the ***J* = 6–5 and 7–6 lines of CO, with $E_u \sim 116$ K and 155 K, respectively**, and in the following we refer to them simply as mid-*J* CO transitions. **Over the past** decade the Atacama Pathfinder Experiment telescope (APEX[1], Güsten et al. 2006) has enabled routine observations of mid-*J* CO lines, **while** *Herschel* **and SOFIA have opened the possibility of spectroscopically resolved observations of even higher-*J* transitions (*J*=10–9 and higher, e.g., Gómez-Ruiz et al. 2012; San José-García et al. 2013, hereafter, SJG13; Leurini et al. 2015; Mottram et al. 2017)**. van Kempen et al. (2009a, **hereafter, vK09**) and van Kempen et al. (2009b) have shown the importance of mid-*J* CO transitions in tracing warm gas in the envelopes and outflows of low-mass protostars. More recently, **SJG13** used the High Frequency Instrument for the Far Infrared **(HIFI, de Graauw et al. 2010)** on board of Herschel to study a sample of low- and high-mass star-forming regions in high-*J* transitions of several CO isotopologues **(e.g., CO, ^{13}CO and C^{18}O $J = 10 - 9$), finding that the link between entrained outflowing gas and envelope motions is independent of the source mass.**

In this paper, we present CO (6–5) and CO (7–6) maps towards a sample of 99 high-mass clumps selected from the APEX Telescope Large Area Survey of the Galaxy (ATLASGAL), which **has provided an unbiased coverage** of the plane of the inner Milky Way in the continuum emission at 870 μm (Schuller et al. 2009). Complementary single-pointing observations of the CO (4–3) line are also included in the analysis in order to characterise the CO emission towards the clumps. Section 2 describes the sample and **Sect.** 3 presents the observations and data reduction. **In Sect. 4 we present the distribution and extent of the mid-*J* CO lines and their line profiles, compute the CO line luminosities and the excitation temperature of the gas, and compare them with the clump properties. In Sect. 5 we discuss our results in the context of previous works. Finally, the conclusions are summarised in Sect. 6.**

2. Sample

ATLASGAL detected the vast majority of all current and future high-mass star forming clumps ($M_{clump} > 1000 M_\odot$) in the inner Galaxy. Recently, Urquhart et al. (2018) completed the distance assignment for ∼ 97 per cent of the ATLASGAL sources and analysed their masses, luminosities and temperatures based on thermal dust emission, and discussed how these properties evolve. Despite the statistical relevance of the ATLASGAL sample, detailed spectroscopic observations are not feasible on the whole sample. Therefore, we defined the ATLASGAL Top 100 (hereafter, TOP100**, Giannetti et al. 2014; König et al. 2017), a flux-limited sample of clumps selected from this survey with additional infrared (IR) selection criteria to ensure it encompasses a full range of luminosities and evolutionary stages (from 70 μm-weak quiescent clumps to H II regions). The 99 sources analysed in this paper are a subsample of the original** TOP100 **(König et al. 2017) and are classified as follows:**

- Clumps which either do not display any point-like emission in the Hi-GAL Survey (Molinari et al. 2010) 70 μm images **and/or** only show weak**, diffuse emission at this wavelength** (hereafter, 70w, 14 sources);
- Mid-**IR** weak sources that are either not associated with any point-like counterparts or the associated compact emission is weaker than 2.6 Jy in the MIPSGAL survey (Carey et al. 2009) 24 μm images (hereafter IRw, 31 sources);
- Mid-IR bright sources in an active phase of **the** high-mass **star formation process**, with strong compact emission seen in 8 μm and 24 μm images, but still not associated with radio continuum emission (hereafter IRb, 33 sources);
- Sources in a later phase of **the** high-mass **star formation process that are** still deeply embedded in their envelope, **but are** bright in the mid-IR and **associated** with radio continuum emission (H II regions, 21 sources).

König et al. (2017) analysed the physical properties of the TOP100 sample in terms of distance, mass and luminosity. They found that **at least 85%** of the sources have the ability to form **high-mass** stars and that most of them are likely gravitationally unstable and would collapse without the presence of a significant magnetic field. These authors showed that the TOP100 **represents a statistically significant sample of high-mass** star-forming clumps covering a range of evolutionary phases, from the coldest and quiescent 70 μm-weak to the most evolved clumps hosting H II regions, with no bias in terms of distance, luminosity and mass among the different classes. The masses and bolometric luminosities of the clumps range from ∼20 to 5.2×10^5 $M_\odot$ and from ∼60 to 3.6×10^6 $L_\odot$, respectively. The distance of the clumps ranges between 0.86 and 12.6 kpc, **and 72 of the 99 clumps have distances below 5 kpc. This implies that observations of the** TOP100 **at the same angular resolution sample quite different linear scales.** In Appendix A, Table A.1, we list the main properties of the observed sources. We adopted the Compact Source Catalogue (CSC) names from Contreras et al. (2013) for the TOP100 sample although the centre of the maps **may** not exactly coincide with those positions (the average offset is ∼5″.4, with values ranging between ∼0″.5–25″.8 **, see Table A.1)**.

In this paper, we investigate the properties of mid-*J* CO lines for a subsample of the original TOP100 **as part of our effort to observationally establish a solid evolutionary sequence for high-mass star formation. In addition to the dust continuum analysis of König et al. (2017), we further characterised the** TOP100 **in terms of the content of the shocked gas in outflows traced by SiO emission (Csengeri et al. 2016) and the ionised gas content (Kim et al. 2017), the CO depletion (Giannetti et al. 2014), and the progressive heating of gas due to feedback of the central objects (Giannetti et al. 2017; Tang et al. 2017). These studies confirm an evolution of the targeted properties with the original selection criteria and strengthen our initial idea that the** TOP100 **sample constitutes a valuable inventory of high-mass clumps in different evolutionary stages.**

3. Observations and data reduction

3.1. CHAMP$^+$ observations

Observations of the TOP100 sample were performed with the APEX 12-m telescope on the following dates of 2014 May 17-20, July 10, 15-19, September 9-11 and 20. The CHAMP$^+$ (Kasemann et al. 2006; Güsten et al. 2008) **multi-beam** heterodyne receiver was used to map the sources simultaneously in the

[1] This publication is based on data acquired with the Atacama Pathfinder Experiment (APEX). APEX is a collaboration between the Max-Planck-Institut für Radioastronomie, the European Southern Observatory, and the Onsala Space Observatory.

CO (6–5) and CO (7–6) transitions. Information about the instrument setup configuration is given in Table 1.

The CHAMP$^+$ array has 2×7 pixels that operate simultaneously in the **radio frequency** tuning ranges 620-720 GHz in the low **frequency array** (LFA) and **the other half in the range** 780-950 GHz in the high frequency **array** (HFA), respectively. The half-power beam widths ($\theta_{\rm mb}$) are 9.″0 (at 691 GHz) and 7.″7 (807 GHz), and the beam-spacing is ∼2.15 $\theta_{\rm mb}$ for both sub-arrays. The observations were performed in continuous on-the-fly (OTF) mode and maps of **80″ × 80″** size, **centred on the coordinates given in Table A.1**, were obtained for each source. The sky subtraction was performed by observing a blank sky field, offset from the central positions of the sources by 600″ in right ascension. The average precipitable water vapour (PWV) of the observations varied from 0.28 to 0.68 mm per day, having a median value of 0.50 mm. The average system temperatures ($T_{\rm sys}$) **ranged from 1050 to 1550 K and 3500 to 6500 K**, at 691 and 807 GHz, respectively. Pointing and focus were checked on planets at the beginning of each observing session. The pointing was also checked every hour on Saturn and Mars, and on hot cores (G10.47+0.03 B1, G34.26, G327.3−0.6, and NGC6334I) during the observations.

Each spectrum was rest-frequency corrected and baseline subtracted using the "Continuum and Line Analysis Single Dish Software" (CLASS), which is part of the GILDAS software[2]. The data were binned to a final spectral resolution of 2.0 km s^{-1} in order to improve the signal-to-noise ratio of the spectra. The baseline subtraction was performed using a first-order fit to the **line-free** channels outside a window **of ± 100** km s^{-1} **wide, centred on the systemic velocity, $V_{\rm lsr}$, of each source**. We used a broader window for sources exhibiting wings broader than ∼ 80 km s^{-1} (AGAL034.2572+00.1535, AGAL301.136−00.226, AGAL327.393+00.199, AGAL337.406−00.402, AGAL351.244+00.669 and AGAL351.774−00.537, see Table A.6). Antenna temperatures ($T_{\rm A}^*$) were converted to main-beam temperatures ($T_{\rm mb}$) using beam efficiencies of 0.41 at 691 GHz and 0.34 at 809 GHz[3]. Forward efficiencies are 0.95 in all observations. **The gridding routine XY_MAP in CLASS was used to construct the final datacubes. This routine convolves the gridded data with a Gaussian of one third of the beam telescope size, yielding a final angular resolution slightly coarser (9.″6 for CO (6–5) and 8.″2 for CO (7–6)) than the original beam size (9.″0 and 7.″7, respectively). The final spectra at the central position of the maps have an average rms noise of 0.20 and 0.87 K for CO (6–5) and CO (7–6) data, respectively.**

Figure 1 presents the ratio between the daily integrated flux and the corresponding average flux for the CO (6–5) transition of each hot core used as calibrator as a function of the observing day. The deviation of the majority of the points with respect to their average value is consistent **within** a ±20% limit; thus, this value was adopted as the uncertainty on the integrated flux for both mid-J CO transitions. On September 10, the observations of G327.3−0.6 showed the largest deviation from the average flux of the source (points at $y\sim0.7$ and $y\sim0.5$ in Fig. 1). For this reason, we **associate** an uncertainty of 30% on the integrated flux of the sources AGAL320.881−00.397, AGAL326.661+00.519 and AGAL327.119+00.509, and of 50% for sources AGAL329.066−00.307 and AGAL342.484+00.182, observed immediately after these two scans on G327.3−0.6.

[2] http://www.iram.fr/IRAMFR/GILDAS

[3] www3.mpifr-bonn.mpg.de/div/submmtech/heterodyne/champplus/champ_efficiencies.16-09-14.html

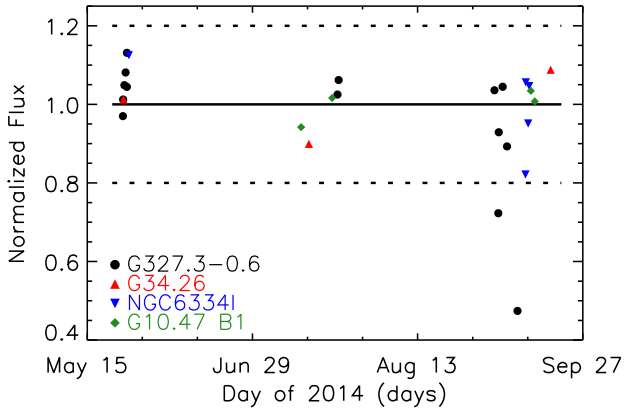


Fig. 1: Ratio between the daily integrated flux and the average flux of the CO (6–5) transition for the calibration sources observed during the campaign. A solid horizontal line is placed at 1.0 and the dashed lines indicate a deviation of 20% from unity.

3.2. FLASH$^+$ *observations*

The FLASH$^+$ (Klein et al. 2014) heterodyne receiver on the APEX telescope was used to observe the **central positions of the CHAMP$^+$ maps in CO (4–3) on 2011 June 15 and 24, August 11 and 12. Table 1 summarises the observational setup. The observations were performed in position switching mode with an offset position of 600″ in right ascension for sky-subtraction.** Pointing and focus were checked on planets at the beginning of each observing session. The pointing was also regularly checked during the observations on Saturn and on hot cores (G10.62, G34.26, G327.3−0.6, NGC6334I and SGRB2(N)). The average PWV varied from 1.10 to 1.55 mm per day with a median value of 1.29 mm. The system temperatures of the observations **ranged from 650 to 2150 K**.

The single-pointing observations were processed using GILDAS/CLASS software. The data were binned to a final spectral resolution of 2.0 km s^{-1} and **a fitted line was subtracted to stablish a straight baseline**. The antenna temperatures were converted to $T_{\rm mb}$ by assuming beam and forward efficiencies of **0.60** and 0.95, respectively. The resulting CO (4–3) spectra have an average rms **noise** of **0.36 K**. The uncertainty on the integrated flux of FLASH$^+$ data was estimated to be ∼20% based on the continuum flux of the sources observed during the pointing scans.

3.3. *Spatial convolution of the mid-J CO data*

The CO (6–5) and CO (7–6) data **were** convolved to a common angular resolution of 13.″4, matching the beam size of the single-pointing CO (4–3) observations. The resulting spectra are shown in Appendix B. **The rms of the convolved spectra range between 0.12–1.60 K, 0.09–1.03 K and 0.33–2.80 K for the CO (4–3), CO (6–5) and CO (7–6) transitions, respectively; with mean and standard deviation values consistent with the values reported in Table 1.** Since our sources are not homogeneously distributed in distance **(see Sect. 2)**, spectra convolved to the same angular resolution of 13.″4 sample linear scales between 0.06 and 0.84 pc. **In order to study the effect of any bias introduced by sampling different linear scales within the clumps, the CO (6–5) and CO (7–6) data were also convolved** to the same linear scale, $\ell_{\rm lin}$, of ∼0.24 pc, which corresponds to

Table 1: Summary of the observations.

Trans.	E_u (K)	Freq. (GHz)	Instr.	η_{mb}	Beam size ($''$)	ΔV (km s^{-1})	T_{sys} (K)	rms (K)	Map size	Number of observed sources
CO (4–3)	55	461.04	FLASH$^+$	**0.60**	13.4	0.953	1398 ± 761	**0.36±0.24**	–	98
CO (6–5)	116	691.47	CHAMP$^+$	0.41	**9.6**	0.318	1300 ± 250	**0.20±0.08**	80″×80″	99
CO (7–6)	155	806.65	CHAMP$^+$	0.34	**8.2**	0.273	5000 ± 1500	**0.87±0.26**	80″×80″	99

Notes. The columns are as follows: (1) observed transition; (2) **upper-level** energy of the transition; (3) rest frequency; (4) instrument; (5) main-beam efficiency (η_{mb}); (6) beam size at the rest frequency; (7) spectral resolution; (8) **mean systemic temperature of the observations**; (9) **mean rms of the data using the single-pointing CO (4–3) spectra and the spectra extracted at central position of the CO (6–5) and CO (7–6) maps; (10) size of the maps; (11) number of observed sources per transition (AGAL301.136+00.226 was not observed in CO (4–3))**; .

an angular size θ **(in radians) of**:

$$\theta = \tan^{-1}\left(\frac{\ell_{lin}}{d}\right) \quad (1)$$

that depends on the distance of the source. The choice of ℓ_{lin} **is driven by the nearest source, AGAL353.066+00.452, for which the part of the map with a relatively uniform rms**[4] **corresponds to a linear scale of ∼0.24 pc.** Since we are limited by the beam size of the CO (6–5) observations (∼10″), the same projected length can be obtained only for sources located **at distances** up to ∼5.0 kpc. This **limit** defines a subsample of **72 clumps (10** 70w**, 20** IRw**, 26** IRb **and 16** H II **regions)**.

The rest of the paper focuses on the properties of the full TOP100 **sample based on the spectra convolved to 13.″4. The properties of the distance-limited subsample differ from those of the 13.″4 data only for the line profile (see Sect. 4.2). A detailed comparison between the CO line luminosity and the properties of the clumps for the distance-limited sample is presented in Appendix C.2.**

3.4. *Self-absorption and multiple velocity components*

The CO spectra of several clumps show a double-peak profile close to the ambient velocity (e.g., AGAL12.804−00.199, AGAL14.632−00.577, and AGAL333.134−00.431, see Fig. B.1). These complex profiles could arise from different velocity components in the beam or could be due to self-absorption given the likely high opacity of CO transitions close to the systemic velocity. To distinguish between these two scenarios, the 13.″4 CO spectra obtained in Sect. 3.3 were compared to the C^{17}O (3–2) data from Giannetti et al. (2014) observed with a similar angular resolution (19″). In the absence of C^{17}O observations (AGAL305.192−00.006, AGAL305.209+00.206 and AGAL353.066+00.452), the C^{18}O (2–1) profiles were used. Since the isotopologue line emission is usually optically thin (cf. Giannetti et al. 2014), it provides an accurate determination of the systemic velocity of the sources and, therefore, can be used to distinguish between the presence of multiple components or self-absorption in the optically thick ^{12}CO lines. Thus, when C^{17}O or C^{18}O show a single peak corresponding in velocity to a dip in CO, we consider the CO spectra to be affected by self-absorption. Otherwise, if also the isotopologue data show a double-peak profile, the emission is likely due to two different velocity components within the beam. From the comparison with the CO isotopologues, we found 83 clumps with self-absorption features in the CO (4–3) line, 79 in the CO (6–5), 70 in the CO (7–6) transition. These numbers indicate that higher-J CO transitions tend to be less affected by self-absorption features when compared to the lower-J CO lines. Finally, only 15 objects do not display self-absorption features in any transitions. The CO spectra affected by self-absorption features are flagged with an asterisk symbol in Table A.6.

To assess the impact of self-absorption on our analysis, in particular on the properties derived from the integrated flux of the CO lines, we compared the observed integrated intensity of each CO transition with the corresponding values obtained from the Gaussian fit presented in Sect. 3.5. First, we tested the impact of self-absorption on the analysis presented in Sect. 4.3 by comparing the observed integrated intensity of the CO lines with the corresponding value from the Gaussian fit. This comparison indicated that self-absorption changes the offsets and the scatter of the data but not the slopes of the relations between the CO emission and the clump properties. Then, we investigated the ratio between the observed and the Gaussian integrated intensity values as a function of the evolutionary classes of the TOP100 **sample. We found that 95% of the sources exhibit ratios between 0.7 and 1.0 for all three lines. We also note a marginal decrease on the ratios from the earliest** 70w **class (∼1.0) to** H II **regions (∼0.8), indicating that self-absorption does not significantly affect the results presented in the following sections. We further investigated the effects of self-absorption by studying the sub-sample of ∼15 sources not affected by self-absorption (that is, the sources that are not flagged with an asterisk symbol in Table A.6) and verified that the results presented in the following sections for the full sample are consistent with those of this subsample, although spanning a much broader range of clump masses and luminosities.**

Five sources (see Appendix C.1) show a second spectral feature in the ^{12}CO transitions and in the isotopologue data of Giannetti et al. (2014) shifted in velocity from the rest velocity of the source. We compared the spatial distribution of the integrated intensity CO (6–5) emission with the corresponding ATLASGAL 870 μm images (see Fig. C.1) for these five clumps. We found that in all sources the morphology of the integrated emission of one of the two components (labelled as C2 in Tables A.2 to A.4) has a different spatial distribution than the dust emission at 870 μm and, thus, is likely not associated with the TOP100 **clumps. These components are excluded from any further analysis in this paper.**

3.5. *Gaussian decomposition of the CO profiles*

The convolved CO spectra were fitted using multiple Gaussian components. The fits were performed interactively using the

[4] **The area outside of the central 60″×60″ region of the map is covered by only one pixel of the instrument, resulting in a larger rms.**

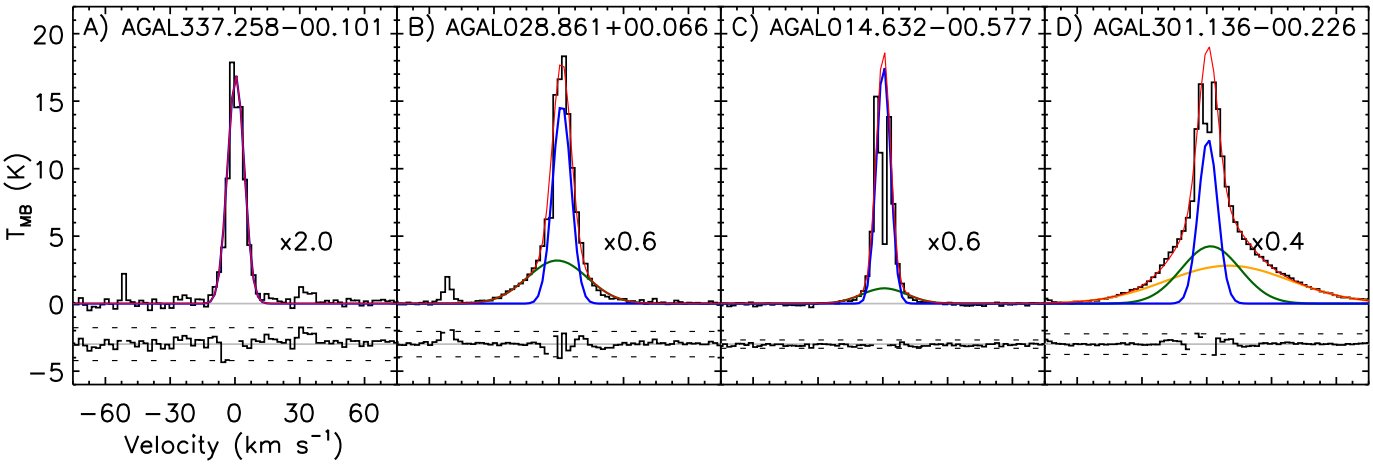


Fig. 2: Gaussian decomposition for the CO (6–5) line using up to 3 components. Panel A: single Gaussian fit; Panel B: two-components fit; Panel C: two-components fit (channels affected by self absorption are masked); Panel D: three-components fit (channels affected by self absorption are masked). The spectra, fits and residuals were multiplied by the factor shown in each panel. The Gaussian fits are shown in blue, green and yellow, ordered by their line width; the sum of all components is shown in red. The grey lines indicate the baseline and the dashed horizontal lines placed in the residuals correspond to the 3-σ level. Self-absorption features larger than 5-σ were masked out from the residuals.

Table 2: Results of the Gaussian fit of the CO spectra convolved to **a common angular resolution of 13.″4**.

Transition	One comp.	Two Comp.	Three comp.
CO (4–3)	28 (11,11,3,3)	67 (3,20,28,16)	3 (0,0,2,1)
CO (6–5)	12 (6,4,1,1)	58 (8,20,20,10)	29 (0,7,12,10)
CO (7–6)	35 (14,14,6,1)	53 (0,16,24,13)	11 (0,1,3,7)

Notes. The columns are as follows: (1) CO transition, (2)–(4) number of sources fitted using (2) a single Gaussian component, (3) two and (4) three components. In each column, the values in parenthesis indicate the corresponding number of 70w, IRw, IRb and H II regions.

Table 3: **Statistics of the FWHM of the Gaussian components fitted on the CO line profiles convolved to a common angular resolution of 13.″4.**

			FWHM (km s^{-1})			
Transition	Class.	N	Range	Mean	Median	σ
CO (4–3)	Narrow	32	3.23-7.47	6.14	6.66	1.23
	Broad	134	7.5-91.2	21.6	15.3	16.9
CO (6–5)	Narrow	46	3.05-7.48	5.94	6.31	1.18
	Broad	161	7.5-97.1	24.3	17.8	18.5
CO (7–6)	Narrow	32	2.00-7.38	5.86	6.28	1.24
	Broad	140	7.5-94.4	28.1	17.1	27.2

Notes. The table presents the statistics on the FWHM of the Narrow and Broad Gaussian velocity component, classified according to Sect. 3.5. The columns are as follows: (1) referred CO transition; (2) classification of the Gaussian component; (3) number of fitted components per class (N); (4) the minimum and maximum value per class; (5) the mean, (6) the median, and (7) the standard deviation of the distribution.

MINIMIZE task in CLASS/GILDAS. A maximum number of three Gaussian components per spectrum was adopted. Each spectrum was initially fitted with one Gaussian component: if the residuals had substructures larger than 3-σ, a second or even a third component was added. In case of self-absorption (see Sect. 3.4), the **affected channels** were masked before performing the fit. Any residual as narrow as the final velocity resolution of the data (2.0 km s^{-1}) was ignored. In particular for CO (4–3), absorption features shifted in velocity from the main line are detected in several sources. These features are likely due to emission in the reference position**, and were also masked before fitting the data. Examples of the line profile decomposition are given in Fig. 2.**

Each component was classified as Narrow (N) or Broad (B), adopting the scheme from San José-García et al. (2013). **According to their definition, the narrow component has a full-width at half maximum (FWHM) narrower than 7.5 km s^{-1}, otherwise it is classified as a broad component. Results of the Gaussian fit are presented in Tables A.2–A.4.** In several cases, two broad components are needed to fit the spectrum. For the CO (6–5) data, **29** of the profiles required 3 components and, thus, two or three components have received the same classification. In these cases, they were named as, e.g., B1, B2; ordered by their width. The C2 features mark secondary velocity components not associated with TOP100 clumps (see Sect. 3.4).

As a consequence of high opacity and self-absorption, the Gaussian decomposition of the line profile can be somewhat dubious. In some cases, and in particular for the CO (4–3) transition, the fit is unreliable (e.g., AGAL305.192−00.006 and AGAL333.134−00.431 in Fig. B.1). The sources associated with unreliable Gaussian decomposed CO profiles (21, 13 and 13 for CO (4–3), CO (6–5) and CO (7–6), respectively) are flagged with an asterisk symbol in Tables A.2 to A.4 and their data are not included in the analysis presented in Sect. 4, as well as the integrated properties of their line profiles (e.g., their integrated intensities and corresponding line luminosities, see Sect. 4.3).

The general overview of the fits are given in Table 2 and the statistics of FWHM of the Narrow and Broad Gaussian components are listed in Table 3. The spectrum of each source with its corresponding decomposition into Gaussian components is presented in Fig. B.1.

4. Observational results

The whole sample was detected above a 3-σ threshold in the single-pointing CO (4–3) data (source AGAL301.136−00.226 was not observed with FLASH$^+$) and in the 13.″4 CO (6–5) **and CO (7–6) spectra, with three** 70w **sources (AGAL030.893+00.139, AGAL351.571+00.762 and AGAL353.417−00.079) only marginally detected above the 3-σ limit in CO (7–6).**

In the rest of this section we characterise the CO emission towards the TOP100 sample through the maps of CO (6–5) (Sect. 4.1) and the analysis of the CO line profiles **for the spectra convolved to 13.″4** (Sect. 4.2). In Sect. 4.3, we compute the CO line luminosities and compare them with the clump properties. **Finally, in Sect. 4.4 we compute the excitation temperature of the gas.**

4.1. The extent of the CO emission

In Fig. 3 we present examples of the integrated intensity maps of the CO (6–5) emission as a function of the evolutionary class of the TOP100 **clumps. The CO (6–5) maps of the full** TOP100 **sample are presented in Appendix B (see Fig. B.1).**

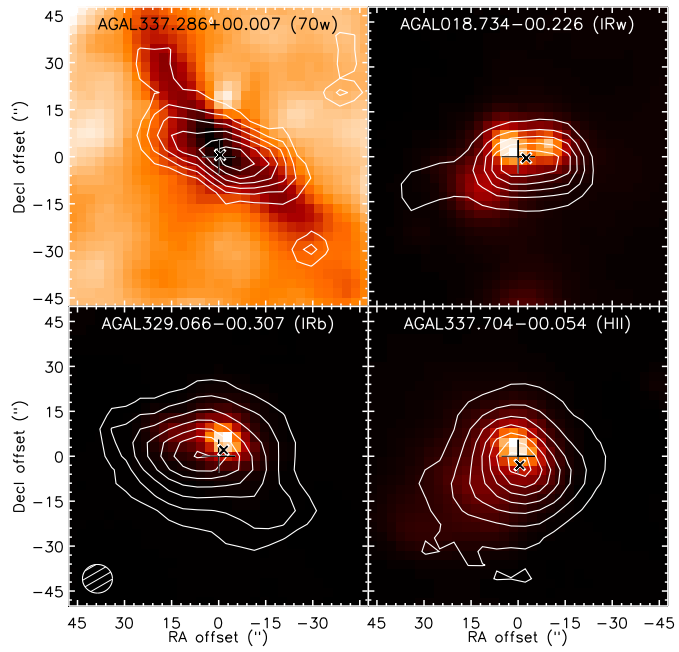


Fig. 3: **Distribution of the CO (6–5) emission of four representative clumps of each evolutionary classes of the** TOP100 **clumps. The CO contours are presented on top of the *Herschel*/PACS maps at 70 μm. The CO contours correspond to the emission integrated over the full-width at zero power (FWZP, with the velocity range shown in the bottom right side of the image) of the CO (6–5) line, and the contour levels are shown from 20% to 90% of the peak emission of each map, in steps of 10%. The (0,0) position of the map is shown as a + symbol, the position of the CSC source from Contreras et al. (2013) is shown as a $\times$ symbol.**

We estimated the linear size of the CO emission, Δs, defined as the average between the maximum and minimum elongation of the 50% peak intensity contour level of the CO (6–5) integrated intensity (see Table A.5). The linear sizes of the CO emission ranges between 0.1 and 2.4 pc, with a median value of 0.5 pc. In order to investigate if Δs varies with evolution, we performed a non-parametric two-sided Kolmogorov-Smirnov (KS) test between pairs of classes (i.e. 70w **vs.** IRb**;** IRw **vs.** H II**). The subsamples were considered statistically different if their KS rank factor is close to 1 and associated with a low probability value, p, ($p \leq 0.05$ for a significance $\geq 2\,\sigma$). Our analysis indicates that there is no significant change in the extension of CO with evolution (KS ≤ 0.37, $p \geq 0.05$ for all comparisons).**

The CO extent was further compared with the bolometric luminosity, $L_{\rm bol}$, and the mass of the clumps, $M_{\rm clump}$, reported by König et al. (2017). The results are presented in Fig. 4. Δs shows a large scatter as a function of $L_{\rm bol}$ while it increases with $M_{\rm clump}$ ($\rho = 0.72$, $p < 0.001$ for the correlation with $M_{\rm clump}$ $\rho = 0.42$, $p < 0.001$ for $L_{\rm bol}$, where the ρ is the Spearman rank correlation factor and p its associated probability). This confirms that the extent of the CO emission is likely dependent of the amount of gas within the clumps, but not on their bolometric luminosity.

We also derived the extent of the 70 μm emission ($\Delta s_{70\,\mu{\rm m}}$) towards the 70 μm-bright clumps using the *Herschel*-PACS images by applying the same procedure used for CO (6–5). The values are also reported in Table A.5. Figure 5 compares the extent of the CO (6–5) emission with that of the 70 μm emission towards the 70 μm bright clumps. The extent of the emission of CO (6–5) and of the 70 μm continuum emission are correlated (Fig. 5, $\rho = 0.67$, $p < 0.001$), and in the majority of cases, the points are located above the equality line, suggesting that the gas probed by the CO (6–5) transition tends to be more extended than the dust emission probed by the PACS data towards the 70 μm-bright clumps.

4.2. Line profiles

In the majority of the cases, the CO profiles are well fit with two Gaussian components, one for the envelope, one for high-velocity emission (see Table 2). A third component is required in some cases, in particular for the CO (6–5) data, which have the highest signal-to-noise ratio. We note that the majority of sources fitted with a single Gaussian component are in the earliest stages of evolution (70w **and** IRw **clumps), suggesting that the CO emission is less complex in earlier stages of high-mass star formation. We also detect non-Gaussian high-velocity wings likely associated with outflows in most of the CO (6–5) profiles. A detailed discussion of the outflow content in the** TOP100 **sample and of their properties will be presented in a forthcoming paper (Navarete et al., in prep.).**

To minimise biases due to different sensitivities in the analysis of single spectra, we computed the average CO spectrum of each evolutionary class and normalised it by its peak intensity. The spectra were shifted to 0 km s^{-1} using the correspondent $V_{\rm lsr}$ given in Table A.1. Then, the averaging was performed by scaling the intensity of each spectrum to the median distance of the subsample ($d = 3.26$ kpc for the distance-limited sample, $d = 3.80$ kpc for the full sample). The resulting spectra of the 13.″4 dataset are shown in Fig. 6 while those of the distance-limited subsample are presented in Appendix C.2 (Fig. C.2). While the 13.″4 data show no significant difference between the average profiles of IRw **and** IRb **classes, in the distance-limited subsample the width (expressed through the full width at zero power, FWZP, to avoid any assumption on the profile) and the intensity of**

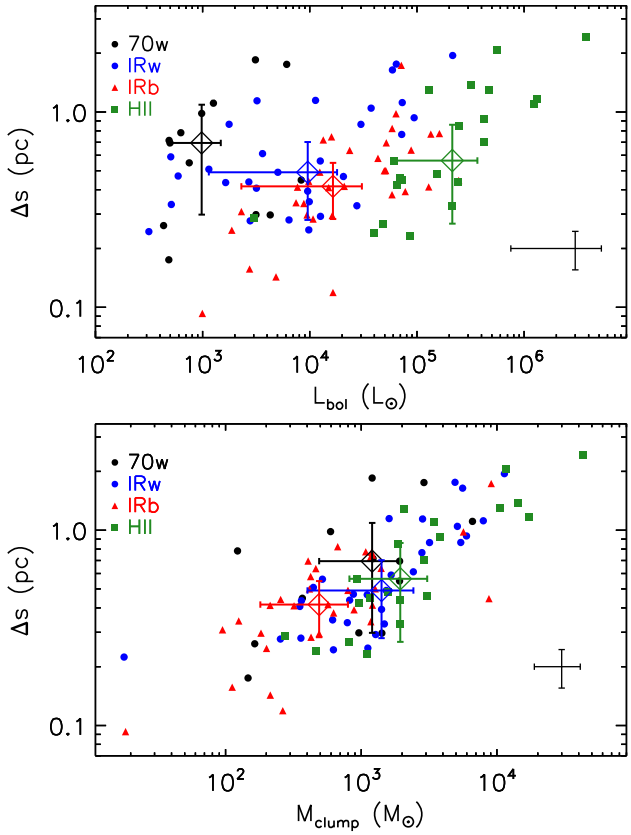


Fig. 4: **Size of the CO (6–5) emission towards the** TOP100 **sample versus the bolometric luminosity (top) and the mass (bottom) of the sources. The median values for each class are shown as open diamonds and their error bars correspond to the absolute deviation of the data from their median value. The typical uncertainty is shown by the error bars on the bottom right of each plot.**

Table 4: Statistics on the CO line profiles**, convolved to a common angular resolution of 13.″4**.

	FWZP (km s^{-1})			
Transition	Range	Mean	Median	Standard deviation
CO (4–3)	10-140	49	44	24
CO (6–5)	14-162	62	56	33
CO (7–6)	4-142	39	30	27

Notes. The table presents the FWZP of the CO lines. The mid-J CO lines were convolved to a common angular resolution of 13.″4. The columns are as follows: (1) referred CO line; (2) the minimum and maximum values, (3) the mean, (4) the median, and (5) the standard deviation of the distribution.

the CO lines progressively increase with the evolution of the sources (from 70w **clumps towards** H II **regions) especially when the normalised profiles are considered. The difference between the two datasets is due to sources at large distances (d > 12 kpc; AGAL018.606−00.074, AGAL018.734−00.226 and AGAL342.484+00.182) for which the observations sample a much larger volume of gas. The increase of line width with evolution is confirmed by the analysis of the individual FWZP values of the three CO lines, presented in Table A.6 (see Table 4 for the statistics on the full** TOP100 **sample).**

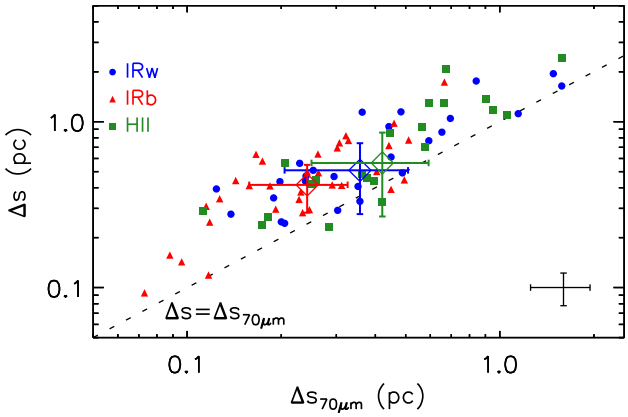


Fig. 5: **Size of the CO (6–5) emission versus the size of the 70 μm emission from the *Herschel*-PACS images towards the 70 μm-bright clumps. The dashed line indicates $y = x$. The median values for the 70 μm-bright classes are shown as open diamonds and their error bars correspond to the absolute deviation of the data from their median value. The typical uncertainty is shown by the error bars on the bottom right of each plot.**

Despite the possible biases in the analysis of the line profiles (e.g., different sensitivities, different excitation conditions, complexity of the profiles), our data indicate that the CO emission is brighter and broader in late evolutionary phases. This was already seen in G327.3−0.6 (AGAL327.293−00.579 in our sample) by Leurini et al. (2013), who mapped a larger area of the source and showed that the mid- and high-J CO (e.g., CO (10–9)) emission has a larger contribution from a photon dissociation region (PDR) associated with a H II **region rather than from younger sources (a hot core and an infrared dark cloud present in the area). Similarly, Cubick et al. (2008) suggested that the main contributors to the Galactic FIR line emission are PDRs surrounding high-mass young stars.**

4.3. The CO line luminosities

The intensity of the CO profiles ($S_{\rm int}$, in K km s^{-1}) was computed by integrating the CO emission over the velocity channels within the corresponding FWZP range. Then, the line luminosity ($L_{\rm CO}$, in K km s^{-1} pc^2) of each CO line was calculated using Eq. 2 from Wu et al. (2005), assuming a source of size equal to the beam size of the data (see Sect. 3.3). The derived $L_{\rm CO}$ values are reported in Table A.6. The errors in the $L_{\rm CO}$ values are estimated by error propagation on the integrated flux and considering an uncertainty of 20% in the distance. The median values of $L_{\rm CO}$, $L_{\rm bol}$, $M_{\rm clump}$ and L/M, the luminosity-to-mass ratio, per evolutionary class are summarised in Table 5. We also performed the same analysis on the data convolved to a common linear scale of 0.24 pc (assuming the corresponding angular source size of 0.24 pc to derive the line luminosity) and no significant differences in the slope of the trends were found. Therefore, the distance-limited sample will not be discussed any further in this section.

In Fig. 7 we show the cumulative distribution function (CDF) of the line luminosities for the three CO transitions: $L_{\rm CO}$ increases from 70w **sources towards** H II **regions. Each evolutionary class was tested against the others by comput-**

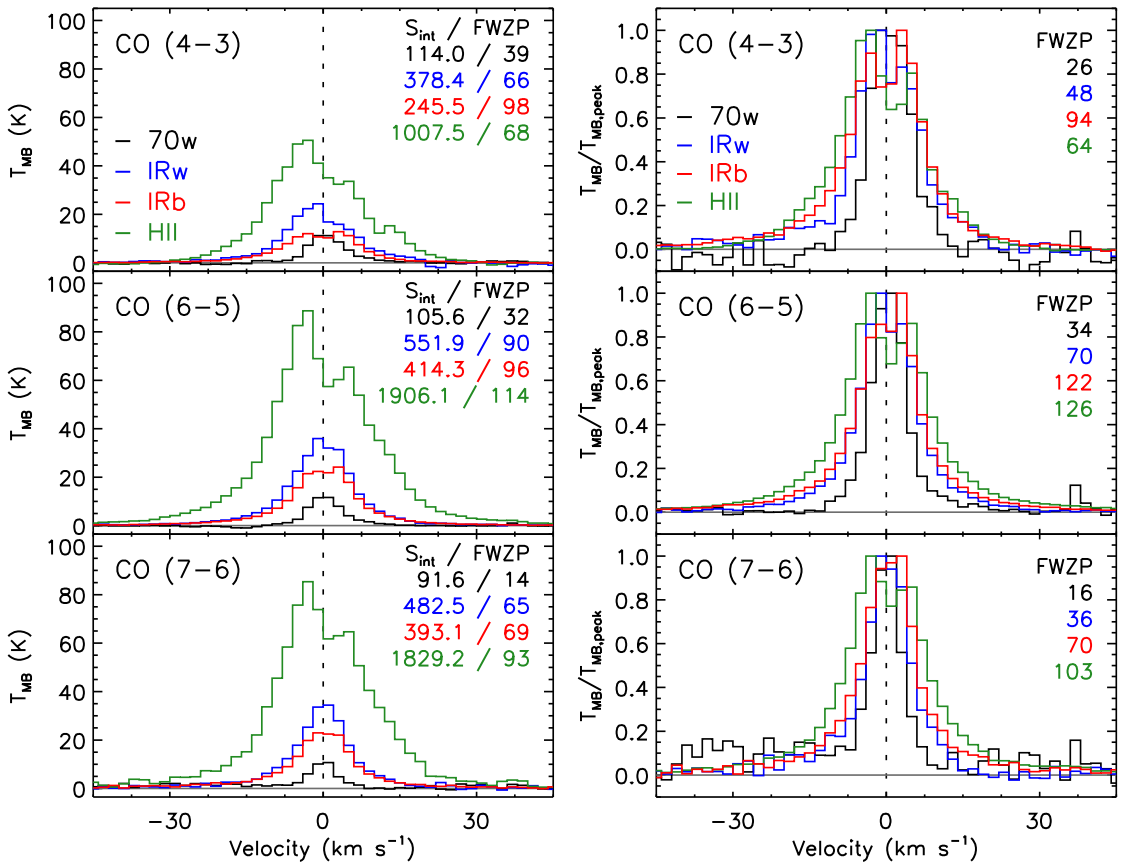


Fig. 6: **Left:** Average CO (4–3), CO (6–5) and CO (7–6) spectra convolved to 13.″4 beam of each evolutionary class scaled to the median distance of the whole sample ($d = 3.80$ kpc). **Right:** Same plot, but the average CO spectra are normalised by their peak intensity. The baseline level is indicated by the solid grey line and the black dashed line is placed at 0 km s^{-1}. The FWZP of the profiles are shown in the upper right side of the panels (in km s^{-1} units), together with the integrated intensities ($S_{\rm int}$, in K km s^{-1} units) of the CO profiles shown in the left panels.

Table 5: Median values per class of the clump and CO profile properties.

Property	70w	IRw	IRb	H II
$L_{\rm bol}$ (10^3 L$_\odot$)	1.26±0.83	9.6±8.4	16.5±1.4	21.4±1.5
$M_{\rm clump}$ (10^3 M$_\odot$)	1.22±0.70	1.4±1.1	0.49±0.31	1.9±1.1
L/M (L$_\odot$/M$_\odot$)	2.58±0.93	9.0±6.8	40±23	76±28
$L_{\rm CO\,(4-3)}$ (K km s^{-1} pc^2)	9.8±7.9	20±14	23±16	52±41
$L_{\rm CO\,(6-5)}$	5.1±4.0	15±11	19±12	51±40
$L_{\rm CO\,(7-6)}$	4.7±3.7	11.8±8.6	14.8±9.7	54±42
FWZP$_{\rm CO\,(4-3)}$ (km s^{-1})	24.0±8.0	40±12	54±14	54±8.0
FWZP$_{\rm CO\,(6-5)}$	26.0±6.0	42±14	72±22	94±26
FWZP$_{\rm CO\,(7-6)}$	12.0±2.0	24.0±6.0	40±14	66±18
$T_{\rm ex}$ (K)	22.4±3.8	33.8±6.7	45±15	85±12

Notes. The median and the absolute deviation of the data from their median value are shown for the clump properties (bolometric luminosity, mass and luminosity-to-mass ratio), for the line luminosity and full width at zero power of the low-J and mid-J CO profiles convolved to the same angular size of 13.″4**, and for the excitation temperature of the CO (6–5) emission**.

ing their two-sided KS coefficient (see Table 6). The most significant differences are found when comparing the earlier and later evolutionary classes (70w and IRb, 70w and H II, $\rho \geq 0.66$ for the CO (6–5) line), while no strong differences are found among the other classes (KS ≤ 0.5 and $p \geq 0.002$ for the CO (6–5) transition). These results indicate that, although we observe an increase on the CO line luminosity from 70w clumps towards H II regions, no clear separation is found in the intermediate classes (IRw and IRb).

We also plotted $L_{\rm CO}$ against the bolometric luminosity of the clumps (Fig. 7), their mass and their luminosity-to-mass ratio (Figs. 8 and 9 for the CO (6–5) line). The L/M ratio is believed to be a rough estimator of evolution in the star formation process for both low- (Saraceno et al. 1996) and high-mass regimes (e.g., Molinari et al. 2008), with small L/M values corresponding to embedded regions where (proto-)stellar activity is just starting, and high L/M values in sources with stronger radiative flux and that have accreted most of the

Table 6: Kolmogorov-Smirnov statistics of the CO line luminosity as a function of the evolutionary class of the clumps.

Classes	CO (4–3)	CO (6–5)	CO (7–6)
70w-IRw	0.32, $p = 0.24$	0.44, $p = 0.03$	0.46, $p = 0.02$
70w-IRb	0.34, $p = 0.16$	0.66, $p < 0.001$	0.66, $p < 0.001$
70w-H II	0.55, $p = 0.007$	0.74, $p < 0.001$	0.83, $p < 0.001$
IRw-IRb	0.21, $p = 0.46$	0.22, $p = 0.38$	0.24, $p = 0.26$
IRw-H II	0.36, $p = 0.06$	0.44, $p = 0.01$	0.45, $p = 0.007$
IRb-H II	0.42, $p = 0.02$	0.48, $p = 0.002$	0.51, $p = 0.002$

Notes. The rank KS and its corresponding probability (p) are shown for each comparison. A p-value of < 0.001 indicate a correlation at 0.001 significance level. p-values of 0.05, 0.002 and < 0.001 represent the ~ 2, 3 and $> 3\,\sigma$ confidence levels.

Table 7: Parameters of the fits of $L_{\rm CO}$ as a function of the clump properties.

Transition	Property	α	β	ϵ
CO (4–3)	$L_{\rm bol}$	$-0.81^{+0.21}_{-0.19}$	0.52±0.05	0.40
	$M_{\rm clump}$	$-1.24^{+0.20}_{-0.18}$	0.87±0.06	0.33
	L/M	$1.03^{+0.13}_{-0.14}$	0.21±0.10	0.60
CO (6–5)	$L_{\rm bol}$	$-1.36^{+0.17}_{-0.16}$	0.63±0.04	0.28
	$M_{\rm clump}$	-1.49±0.21	0.90±0.07	0.36
	L/M	$0.73^{+0.10}_{-0.12}$	0.41±0.10	0.57
CO (7–6)	$L_{\rm bol}$	$-1.67^{+0.17}_{-0.15}$	0.69±0.04	0.23
	$M_{\rm clump}$	-1.53±0.23	0.89±0.08	0.43
	L/M	0.56±0.11	0.54±0.10	0.56

Notes. The fits were performed by adjusting a model with three free parameters in the form of $\log(y) = \alpha + \beta \log(x) \pm \epsilon$, where α, β and ϵ correspond to the **intercept, the slope** and the intrinsic scatter, respectively.

mass (Molinari et al. 2016; Giannetti et al. 2017; Urquhart et al. 2018). In addition, the L/M ratio also reflects the properties of the most massive young stellar object embedded in the clump (Faúndez et al. 2004; Urquhart et al. 2013a). The fits were performed using a Bayesian approach, by adjusting a model with three free parameters (the intercept, α, the slope, β, and the intrinsic scatter, ϵ). In order to obtain a statistically reliable solution, we computed a total of 100 000 iterations per fit. The parameters of the fits are summarised in Table 7. The correlation between $L_{\rm CO}$ and the clump properties was checked by computing their Spearman rank correlation factor and its associated probability (ρ and p, respectively, see Table 8). Since $L_{\rm CO}$ with $L_{\rm bol}$ and $M_{\rm clump}$ have the same dependence on the distance of the source, a partial Spearman correlation test was computed and the partial coefficient, $\rho_{\rm p}$, was obtained (see Table 8).

In the right panel of Fig. 7, we show the CO line luminosity versus the bolometric luminosity of the TOP100 **clumps. The plot indicates that $L_{\rm CO}$ increases with $L_{\rm bol}$ over the entire $L_{\rm bol}$ range covered by the** TOP100 **clumps ($\sim 10^2$-10^6 L$_\odot$). The Spearman rank test confirms that both quantities are well correlated for all CO lines ($\rho \geq 0.67$, with $p < 0.001$), even when excluding the mutual dependence on distance ($\rho_p \geq 0.68$). The results of the fits indicate a systematic increase in the slope of $L_{\rm CO}$ versus $L_{\rm bol}$ for higher-J transitions: 0.52±0.05, 0.63±0.04 and 0.69±0.04 for the CO (4–3), CO (6–5) and CO (7–6), respectively. For the CO (6–5) and CO (7–6) lines, however, the slopes are consistent in within 2-σ. Concerning the dependence of the CO luminosity on $M_{\rm clump}$ (see**

Table 8: Spearman rank correlation statistics for the CO line luminosity as a function of the clump properties towards the TOP100 sample.

Property	CO (4–3)	CO (6–5)	CO (7–6)
$L_{\rm bol}$	0.67, $p < 0.001$; $\rho_p = 0.68$	0.83, $p < 0.001$; $\rho_p = 0.87$	0.87, $p < 0.001$; $\rho_p = 0.89$
$M_{\rm clump}$	0.72, $p < 0.001$; $\rho_p = 0.32$	0.74, $p < 0.001$; $\rho_p = 0.47$	0.75, $p < 0.001$; $\rho_p = 0.52$
L/M	0.20, $p = 0.09$	0.39, $p < 0.001$	0.43, $p < 0.001$

Notes. The rank ρ and its corresponding probability (p) are shown for each comparison. A p-value of < 0.001 indicate a correlation at 0.001 significance level. p-values of 0.05, 0.002 and < 0.001 represent the ~ 2, 3 and $> 3\,\sigma$ confidence levels. For $L_{\rm bol}$ and $M_{\rm clump}$, the partial correlation coefficient, ρ_p, is also shown.

Fig. 8), the partial correlation tests indicates that the distance of the clumps plays a more substantial role in the correlation found between $L_{\rm CO}$ and $M_{\rm clump}$ ($0.32 \leq \rho_p \leq 0.52$) than in the correlations found for $L_{\rm CO}$ vs. $L_{\rm bol}$. Finally, we do not find any strong correlation between the CO line luminosity and L/M ($\rho \leq 0.43$ for all transitions) although the median $L_{\rm CO}$ values per class do increase with L/M (Fig. 9). These findings are discussed in more detail in Sect. 5.

4.4. The excitation temperature of the CO gas

The increase of $L_{\rm CO}$ with the bolometric luminosity of the source (see Fig. 7) suggests that the intensity of the CO transitions may depend on an average warmer temperature of the gas in the clumps due to an increase of the radiation field from the central source (see for example van Kempen et al. 2009a). To confirm this scenario, we computed the excitation temperature of the gas, $T_{\rm ex}$, and compared it with the properties of the clumps.

Ideally, the intensity ratio of different CO transitions well separated in energy (e.g., CO (4–3) and CO (7–6)) allows a determination of the excitation temperature of the gas. However, most of the CO profiles in the TOP100 **clumps are affected by self-absorption (see Sect. 3.4), causing a considerable underestimate of the flux especially in CO (4–3) and leading to unreliable ratios. Moreover, the CO (6–5) and CO (7–6) lines are too close in energy to allow a reliable estimate of the temperature. Alternatively, the excitation temperature can be estimated using the peak intensity of optically thick lines. From the equation of radiative transport, the observed main beam temperature ($T_{\rm mb}$) can be written in terms of $T_{\rm ex}$ as:**

$$T_{\rm mb} = \frac{h\nu}{k} \left[J_\nu(T_{\rm ex}) - J_\nu(T_{\rm bg}) \right] \left[1 - exp\,(-\tau_\nu) \right] \quad (2)$$

where $J_\nu(T) = \left[\exp(h\nu/kT) - 1\right]^{-1}$, $T_{\rm bg}$ is the background temperature and τ_ν is the opacity of the source at the frequency ν. In the following, we include only the cosmic background as background radiation. Assuming optically thick emission ($\tau_\nu \gg 1$), $T_{\rm ex}$ is given by:

$$T_{\rm ex} = \frac{h\nu/k}{\ln\left[1 + \frac{h\nu/k}{T_{\rm mb} + (h\nu/k) J_\nu(T_{\rm bg})}\right]} \quad (3)$$

We computed $T_{\rm ex}$ using the peak intensity of the CO (6–5) line from the Gaussian fit (Sect. 3.5) and also from its maxi-

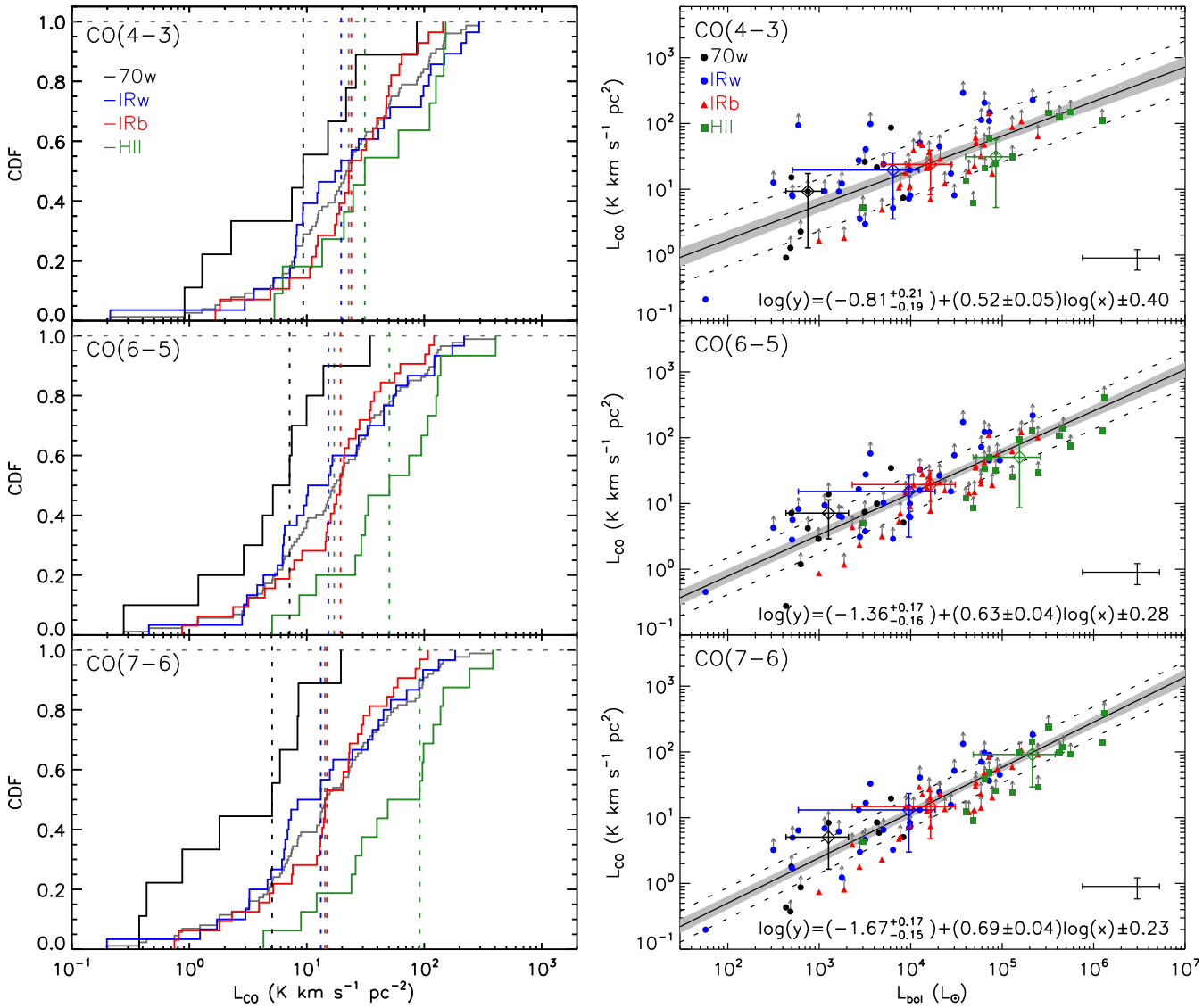


Fig. 7: **Left panels: Cumulative distribution function of the line luminosity of the CO (4–3) (upper panel), CO (6–5) (middle) and CO (7–6) emission (bottom) towards the** TOP100 **sample. The median values per class are shown as vertical dashed lines in their corresponding colours. Right** panels: Line luminosity of the same CO J-transitions versus the bolometric luminosity of the TOP100 sources. The median values for each class are shown as open diamonds and their error bars correspond to the absolute deviation of the data from their median value. Points having an upward arrow indicate a self-absorption feature in the spectrum convolved to 13″.4 and correspond to a lower limit. The typical error bars are shown at the bottom right side of the plots. The black solid line is the best fit, the light grey shaded area indicates the 68% uncertainty, and the dashed lines show the intrinsic scatter (ϵ) of the relation.

mum observed value. Since CO (6–5) may be affected by self-absorption, the maximum observed intensity likely results in a lower limit of the excitation temperature. The values derived using both methods are reported in Table A.7. $T_{\rm ex}$ derived from the peak intensity of the Gaussian fit ranges between 14 and 158 K, with a median value of 37 K. The analysis based on the observed intensity delivers similar results ($T_{\rm ex}$ values range between 14 and 147 K, with a median value of 34 K).

The temperature of the gas increases with the evolutionary stage of the clumps and is well correlated with $L_{\rm bol}$ ($\rho = 0.65$, $p < 0.001$, see Fig. 10a). No significant correlation is found with $M_{\rm clump}$ ($\rho = 0.07$, $p = 0.52$). On the other hand, the excitation temperature is strongly correlated with L/M ($\rho = 0.70$, $p < 0.001$), suggesting a progressive warm-up of the gas in more evolved clumps. We further compared the $T_{\rm ex}$ values obtained from CO with temperature estimates based on other tracers ($C^{17}O$ (3–2), methyl acetylene, CH_3CCH, ammonia, and the dust, Giannetti et al. 2014, 2017; König et al. 2017; Wienen et al. 2012). All temperatures are well correlated ($\rho \geq 0.44$, p,< 0.001), however, the warm-up of the gas is more evident in the other molecular species than in CO (cf. Giannetti et al. 2017).

5. Discussion

5.1. Opacity effects

In Sect. 3.4 we found that self-absorption features are present in most of the CO spectra analysed in this work. To ad-

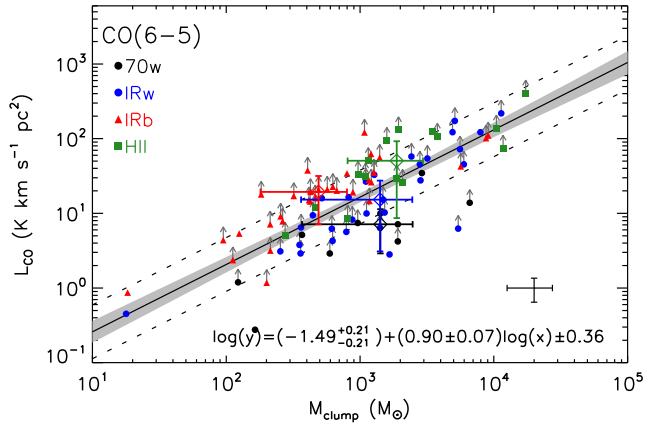


Fig. 8: Same as the right panel of Fig. 7, but displaying the line luminosity of the CO (6–5) emission as a function of the mass of the clumps.

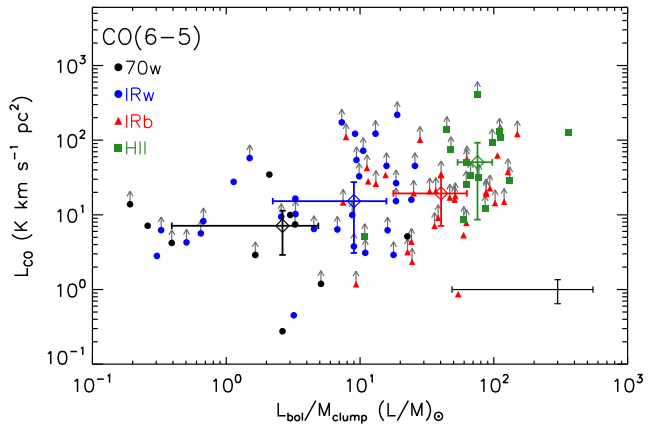


Fig. 9: Same as the right panel of Fig. 7, but displaying L_{CO} of the CO (6–5) line as a function of the L/M ratio of the TOP100 sources.

Table 9: Kolmogorov-Smirnov statistics of the excitation temperature of the CO (6–5) line as a function of the evolutionary class of the clumps.

Classes	Observed	Gaussian
70w-IRw	0.51, $p = 0.008$	0.54, $p = 0.01$
70w-IRb	0.87, $p < 0.001$	0.80, $p < 0.001$
70w-H II	1.00, $p < 0.001$	1.00, $p < 0.001$
IRw-IRb	0.51, $p < 0.001$	0.43, $p = 0.005$
IRw-H II	0.94, $p < 0.001$	0.93, $p < 0.001$
IRb-H II	0.60, $p < 0.001$	0.63, $p < 0.001$

Notes. The rank KS and its corresponding probability (p) are shown for each comparison. A p-value of < 0.001 indicate a correlation at 0.001 significance level. p-values of 0.05, 0.002 and < 0.001 represent the ~ 2, 3 and $> 3\,\sigma$ confidence levels.

dress this, we investigated the effects of self-absorption on our analysis and concluded that they are negligible since more than 80% of the CO integrated intensities are recovered in the majority of the sources (Sect. 3.5). We also verified that the steepness of the relations between L_{CO} and the clump properties is not affected by self-absorption.

In addition, the CO lines under examination are certainly optically thick, and their opacity is likely to decrease with J. Indeed, the comparison between L_{CO} for different CO transitions and the bolometric luminosity of the clumps (see Sect. 4.3) suggests a systematic increase in the slope of the relations as a function of J (see Table 7 for the derived power-law indices for L_{CO} versus L_{bol}). Such a steepening of the slopes with J is even more evident when including the relation found by San José-García et al. (2013) for CO (10–9) line luminosity in a complementary sample of the TOP100. For the CO (10–9) transition, they derived $\log(L_{CO}) = (-2.9\pm0.2) + (0.84\pm0.06)\log(L_{bol})$, which is steeper than the relations found towards lower-J transitions reported in this work. In Sect. 5.4 we further discuss SJG13 results by analysing their low- and high-mass YSO subsamples. Our findings suggest that there is a significant offset between the subsamples, leading to a much steeper relation between L_{CO} and L_{bol} when considering their whole sample. However, the individual subsamples follow similar power-law distributions, with power-law indices of (0.70±0.08) and (0.69±0.21) for the high- and low-mass YSOs, respectively.

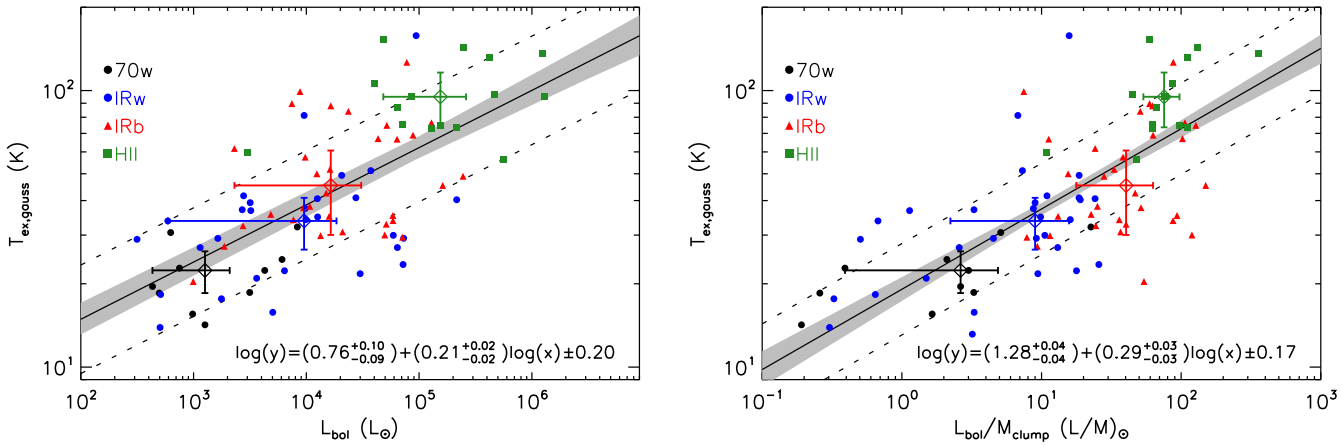
In Fig. 11, we present the distribution of the power-law indices of the L_{CO} versus L_{bol} relations, β_J, as a function of their corresponding upper-level J number, J_{up}. We include also the datapoint from the $J_{up} = 10$ line for the high-mass sources of SJG13 (see also discussion in Sect. 5.4. The best fit to the data, $\beta_J = (0.40\pm0.11)+(0.04\pm0.02)\,J_{up}$, confirms that the power-law index β_J gets steeper with J.

The fact that the opacity decreases with J could result in different behaviours of the line luminosities with L_{bol} for different transitions. This effect was recently discussed by Benz et al. (2016) who found that the value of the power-law exponents of the line luminosity of particular molecules and transitions depends mostly on the radius where the line gets optically thick. In the case of CO lines, the systematic increase on the steepness of the L_{CO} versus L_{bol} relation with J (see Table 7) suggests that higher J lines trace more compact gas closer to the source and, thus, a smaller volume of gas is responsible for their emission. Therefore, observations of distinct J transitions of the CO molecule, from CO (4–3) to CO (7–6) (and even higher J transitions, considering the CO (10–9) data from SJG13), suggest that the line emission arises and gets optically thick at different radii from the central sources, in agreement with Benz et al. (2016).

5.2. Evolution of CO properties with time

In Sect. 4.3, we showed that L_{CO} does not correlate with the evolutionary indicator L/M. This result is unexpected if we consider that in TOP100 the evolutionary classes are quite well separated in L/M (with median values of 2.6, 9.0, 40 and 76 for 70w, IRw, IRb and H II regions, respectively, König et al. 2017).

Previous work on SiO in sources with similar values of L/M as those of the TOP100 (e.g., Leurini et al. 2014 and Csengeri et al. 2016) found that the line luminosity of low-excitation SiO transitions does not increase with L/M, while the line luminosity of higher excitation SiO lines (i.e., $J_{up} > 3$) seems to increase with time. Those authors interpreted these findings in terms of a change of excitation conditions with time which is not reflected in low excitation transitions. This effect likely applies also to low- and mid-J CO lines with relatively low energies (≤ 155 K); higher-J CO transitions could be more sensitive to changes in excitation since they have



(a) (b)

Fig. 10: Excitation temperature of the CO (6–5) line versus the bolometric luminosity of the TOP100 clumps (Left) and the luminosity-to-mass ratio (Right). The excitation temperature was derived using the peak of the Gaussian fit of the CO profiles. The median values for each class are shown as open diamonds and their error bars correspond to the median absolute deviation of the data from their median value. The black solid line is the best fit, the light grey shaded area indicates the 68% uncertainty, and the dashed lines show the intrinsic scatter (ϵ) of the relation. The best fit to the data is indicated by the filled black line.

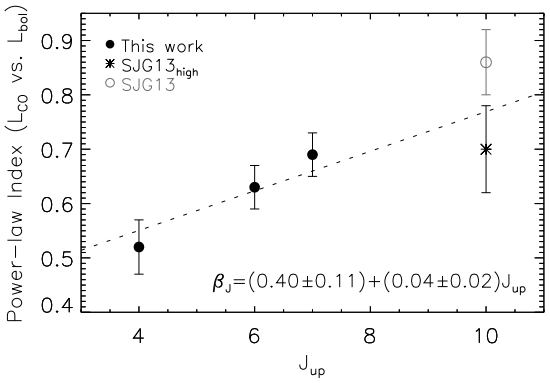


Fig. 11: Power-law indices of the L_{CO} versus L_{bol} relations for different J transitions as a function of the upper-level J number. The β indices from Table 7 (filled black circles) are plotted together with data from **SJG13** (open grey circle, excluded from the fit) and the exponent derived for their high-luminosity subsample ($*$ symbol, see Fig. 12). The best fit is indicated by the dashed black line.

upper level energies in excess of 300 K (e.g., CO (10–9) or higher J CO transitions). This hypothesis is strengthened by our finding that the excitation temperature of the gas increases with L/M (see Fig. 10b). This scenario can be tested with observations of high-J CO transitions which are now made possible by the SOFIA telescope in the range J_{up} = 11–16. Also *Herschel*-PACS archive data could be used despite their coarse spectral resolution.

5.3. Do embedded H II regions still actively power molecular outflows?

In Sect. 4.2 we showed that H II regions have broad CO lines (see Fig. 6) likely associated with high-velocity outflowing gas. The H II sources in our sample are compact/unresolved objects in the continuum emission at 5 GHz. **Their envelopes still largely consist of molecular gas and have not yet been significantly dispersed by the energetic feedback of the YSOs**. Our observations suggest that high-mass YSOs in this phase of evolution still power molecular outflows and are therefore accreting. This result is in agreement with the recent study of **Urquhart et al. (2013b) and** Cesaroni et al. (2015) who suggested that accretion might still be present during the early stages of evolution of H II regions based on the finding that **the Lyman continuum luminosity of several H II regions appears in excess of that expected for a zero-age main-sequence star with the same bolometric luminosity. Such excess could be due** to the so-called flashlight effect (e.g., Yorke & Bodenheimer 1999), where most of the photons escape along the axis of a bipolar outflow. Indeed Cesaroni et al. (2016) further investigated the origin of the Lyman excess looking for infall and outflow **signatures** in the same sources. They found evidence for both phenomena although with low-angular resolution data. Alternatively, the high-velocity emission seen in CO in the TOP100 in this work and in SiO in the ultra-compact H II regions of Cesaroni et al. (2016) could be associated with other younger unresolved sources in the clump and not directly associated with the most evolved object in the cluster. Clearly, high angular resolution observations **(e.g., with ALMA)** are needed to shed light on the origin of the high-velocity emission and confirm whether indeed the high-mass YSO ionizing the surrounding gas is still actively accreting.

5.4. CO line luminosities from low- to high-luminosity sources

In this section, we further study the correlation between L_{CO} and the bolometric luminosity of the clumps for different CO transitions to investigate the possible biases that can arise when comparing data with very different linear resolutions. This is important in particular when comparing galactic observations to the increasing number of extragalactic studies of mid- and high-J CO lines (e.g., Weiß et al. 2007; Decarli et al. 2016). We use results from SJG13 for the high energy CO (10–9) line (with a resolution of ∼ 20″) and from the CO (6–5) and CO (7–6) transitions observed with APEX by vK09. The sources presented by SJG13 cover a broad range of luminosities (from < 1 $L_\odot$ to ∼ 10^5 $L_\odot$) and are in different

evolutionary phases. On the other hand, the sample studied by vK09 consists of eight low-mass YSOs with bolometric luminosities $\lesssim 30\,L_\odot$.

To investigate the dependence of the line luminosity in different CO transitions on L_{bol} from low- to high-mass star-forming clumps, we first divided the sources from SJG13 into low- ($L_{bol} < 50\,L_\odot$) and high-luminosity ($L_{bol} > 50\,L_\odot$) objects. In this way and assuming the limit $L_{bol} = 50\,L_\odot$ adopted by SJG13 as a separation between low- and intermediate/high-mass YSOs, we defined a sub-sample of low-mass sources (the targets of vK09 for CO (6–5) and CO (7–6), and those of SJG13 with $L_{bol} < 50\,L_\odot$ for CO $J = 10 - 9$) and one of intermediate- to high-mass clumps (the TOP100 **for the mid-J CO lines and the sources from SJG13 with $L_{bol} > 50\,L_\odot$ for CO (10–9)). The limit $L_{bol} = 50\,L_\odot$ was chosen as separation between low- and intermediate/high-mass YSOs following SJG13.**

In the upper panels of Fig. 12 we compare our data with those of vK09 for the CO (6–5) and CO (7–6) transitions. We could not include the sources of **SJG13** in this analysis because observations in the CO (6–5) or CO (7–6) lines are not available. We calculated the CO line luminosity of their eight low-mass YSOs using the integrated intensities centred on the YSO on scales of ∼0.01 pc (see their Table 3). **In order to limit biases due to different beam sizes, we recomputed the CO luminosities from the central position of our map at the original resolution of the** CHAMP$^+$ **data (see Table 1), probing linear scales ranging from ∼0.04 to 0.6 pc.** We performed three fits on the data: we first considered only the original sources of **vK09 and the** TOP100 **separately, and then combined both samples**. The derived power-law indices of the CO (6–5) data are 0.59 ± 0.25 and 0.59 ± 0.04 for the low- and high-luminosity subsamples, respectively. Although the power-law indices derived for the two sub-samples are consistent within 1-σ, the fits are offset by **roughly one order of magnitude (from** -2.75 **to** -1.54 **dex)**, indicating that L_{CO} values are systematically larger towards high-luminosity sources. Indeed, the change on the offsets explains reasonably well the steeper power-law index found when combining both subsamples (0.74 ± 0.03). Similar results are found for CO (7–6), although the difference between the offsets are slight smaller (∼0.8 dex). **The bottom panel of Fig. 12 presents the CO (10–9) luminosity for the SJG13 sample with the best fit of their low- and high-luminosity sources separately. The derived power-law indices are 0.69 ± 0.21 and 0.70 ± 0.08 for the low- and high-luminosity subsamples, respectively. The fits are offset by roughly 0.3 dex, which also explains the steeper slope of 0.84 ± 0.06 found by SJG13 when fitting the two sub-samples simultaneously.**

We interpret the shift in CO line luminosities between low- and high-luminosity sources as a consequence of the varying linear resolution and sampled volume of gas of the data across the L_{bol} axis. In high-mass sources, mid-J CO lines trace extended gas (see the maps presented in Figs. B.1 for the TOP100**) probably due to the effect of clustered star formation. Since the data presented in Fig. 12 are taken with comparable angular resolutions, the volume of gas sampled by the data is increasing with L_{bol} because sources with high luminosities are on average more distant. For the CO (10–9) data, the two subsamples are likely differently affected by beam dilution. In close-by low-mass YSOs, the CO (10–9) line is dominated by emission from UV heated outflow cavities (van Kempen et al. 2010) and therefore is extended. In high-mass YSOs, the CO (10–9) line is probably emitted in the inner part of passively heated envelope (Karska et al. 2014) and therefore could suffer from beam dilution. This could explain the smaller offset in the CO (10–9) line luminosity between the low- and high-luminosity subsamples.**

SJG13 and San José-García et al. (2016) found a similar increase in the slope of the line luminosity of the CO $(10 - 9)$ **and of H_2O transitions versus L_{bol} when including extragalactic sources (see Fig. 14 of SJG13 and Fig. 9 of San José-García et al. 2016). These findings clearly outline the difficulties of comparing observations of such different scales and the problems to extend Galactic relations to extragalactic objects.**

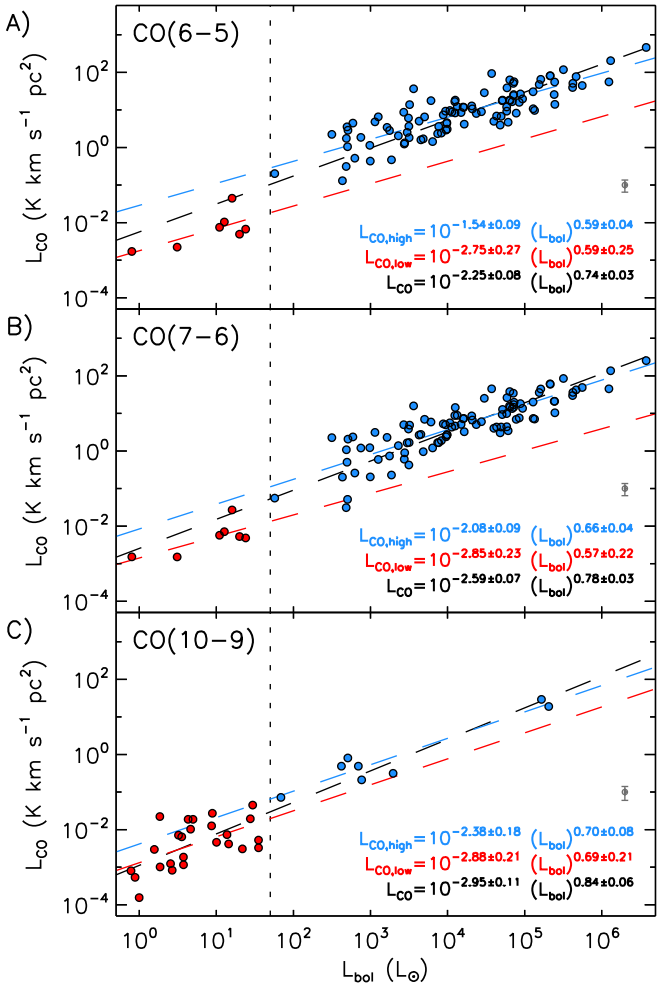


Fig. 12: CO line luminosity as a function of the bolometric luminosity for the CO (6–5) (upper panel), CO (7–6) (middle) and CO (10–9) (bottom) transitions. The fits were performed on the whole dataset (all points, shown with black contours), on the low- and high-luminosity subsamples (points filled in red and blue, respectively). The CO (6–5) and CO (7–6) data towards low-luminosity sources are from van Kempen et al. (2009a); the CO (10–9) data are from San José-García et al. (2013). The dashed vertical line at $L_{bol} = 50\,L_\odot$ marks the transition from low- to high-**mass YSOs**. The typical error bars are shown in the bottom right side of the plots.

6. Summary

A sample of 99 sources, selected from the ATLASGAL 870 μm survey and representative of the Galactic population **of star-**

forming clumps in different evolutionary stages (from 70 μm-weak clumps to H II regions), was characterised **in terms of** their CO (4–3), CO (6–5) and CO (7–6) emission.

We first investigated the effects of different linear resolutions on our data. By taking advantage of our relatively high angular resolution maps in the CO (6–5) and CO (7–6) lines, we could study the influence of different beam sizes on the observed line profiles and on the integrated emission. We first convolved the CO (6–5) **and CO (7–6)** data to a common linear size of ~0.24 pc using a distance limited sub-sample of clumps and then to a common angular resolution of 13.″4**, including the single-pointing CO (4–3) data**. We verified that the results **typically** do not depend on the spatial resolution of the data, at least in the range of distances sampled by our sources. The only difference between the two methods is found when comparing the average spectra for each evolutionary class: indeed, only when using spectra that sample the same **volume of** gas (**i.e.,** same linear resolution) it is possible to detect an increase in line width from 70w **clumps to** H II **regions, while the line widths of each evolutionary class are less distinct** in the spectra smoothed to the same angular size **due to sources at large distances (>12 kpc)**. This result is encouraging for studies of large samples of SF regions across the Galaxy based on single-pointing observations.

The analysis of the CO emission led to the following results:

1. All the sources were detected in the **CO (4–3), CO (6–5) and CO (7–6) transitions.**
2. **The spatial distribution of the CO (6–5) emission ranges between 0.1 and 2.4 pc. The sizes of mid-J CO emission display a moderate correlation with the sub-mm dust mass of the clumps, suggesting that the extension of the gas probed by the CO is linked to the available amount of the total gas in the region. In addition, the CO (6–5) extension is also correlated with the infrared emission probed by the *Herschel*-PACS 70 μm maps towards the 70 μm-bright clumps.**
3. The CO profiles can be decomposed using up to three velocity components. The majority of the spectra are well fitted by two components, one narrow (FWHM < 7.5 km s^{-1}) and one broad; 30% of the sources need a third **and broader** component for the CO (6–5) line profile.
4. **The FWZP of the CO lines increases with the evolution of the clumps (with median values of 26, 42, 72 and 94 for** 70w**,** IRw**,** IRb **clumps and** H II **regions, respectively, for the CO (6–5) transition).** H II regions are often associated with broad **velocity** components, with FWHM values up to ~100 km s^{-1}. This suggests that accretion, resulting in outflows, is still undergoing in the more evolved clumps of the TOP100.
5. The CO line luminosity increases with the **bolometric** luminosity and the mass of the sources, although it does not seem to increase with L/M.
6. The dependence of the CO luminosity as a function of the bolometric luminosity of the source seems to get steeper with J. This likely reflects the fact that higher J CO transitions are more sensitive to the temperature of the gas and likely **arise** from an inner part of the envelope.
7. **The excitation temperature of the clumps was evaluated based on the peak intensity of the Gaussian fit of the CO (6–5) spectra. We found that T_{ex} increases as a function of the bolometric luminosity and the luminosity-to-mass ratio of the clumps, as expected for a warming up of the gas from** 70w **clumps towards** H II **regions.**
8. The observed CO emission towards more luminous and distant objects likely originates from multiple sources within the linear scale probed by the size of beam (up to 0.84 pc), thus, are systematically larger than the emission from resolved and nearby less luminous objects, from which the CO emission is integrated over smaller linear scales (~0.01 pc). We found that the line luminosity of the CO lines shows similar slopes as a function of the bolometric luminosity for **low-mass and high-mass star-forming sources**. However, as a consequence, the distribution of the CO line luminosity versus the bolometric luminosity follows steeper power-laws when combining low- and high-luminosity sources.

Acknowledgements. FN thanks to *Fundação de Amparo à Pesquisa do Estado de São Paulo* (FAPESP) for support through processes 2014/20522-4 and 2013/11680-2. T.Cs. acknowledges support from the *Deutsche Forschungsgemeinschaft, DFG* via the SPP (priority programme) 1573 'Physics of the ISM'. We thank the useful comments and suggestions made by an anonymous referee that led to a much improved version of this work.

References

Benz, A. O., Bruderer, S., van Dishoeck, E. F., et al. 2016, A&A, 590, A105 5.1
Bergin, E. A. & Tafalla, M. 2007, ARA&A, 45, 339 1
Beuther, H., Walsh, A., Schilke, P., et al. 2002, A&A, 390, 289 1
Carey, S. J., Noriega-Crespo, A., Mizuno, D. R., et al. 2009, PASP, 121, 76 2
Cesaroni, R., Pestalozzi, M., Beltrán, M. T., et al. 2015, A&A, 579, A71 5.3
Cesaroni, R., Sánchez-Monge, Á., Beltrán, M. T., et al. 2016, A&A, 588, L5 5.3
Contreras, Y., Schuller, F., Urquhart, J. S., et al. 2013, A&A, 549, A45 2, 3, A, B.1, C.1
Csengeri, T., Leurini, S., Wyrowski, F., et al. 2016, A&A, 586, A149 2, 5.2
Csengeri, T., Urquhart, J. S., Schuller, F., et al. 2014, A&A, 565, A75 A, B.1
Cubick, M., Stutzki, J., Ossenkopf, V., Kramer, C., & Röllig, M. 2008, A&A, 488, 623 4.2
de Graauw, T., Helmich, F. P., Phillips, T. G., et al. 2010, A&A, 518, L6 1
Decarli, R., Walter, F., Aravena, M., et al. 2016, ApJ, 833, 70 5.4
Faúndez, S., Bronfman, L., Garay, G., et al. 2004, A&A, 426, 97 4.3
Giannetti, A., Leurini, S., Wyrowski, F., et al. 2017, A&A 2, 4.3, 4.4
Giannetti, A., Wyrowski, F., Brand, J., et al. 2014, A&A, 570, A65 2, 3.4, 4.4, A, B, B.1
Gómez-Ruiz, A. I., Gusdorf, A., Leurini, S., et al. 2012, A&A, 542, L9 1
Grellmann, R., Preibisch, T., Ratzka, T., et al. 2013, A&A, 550, A82 1
Güsten, R., Baryshev, A., Bell, A., et al. 2008, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 7020, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 10 3.1
Güsten, R., Nyman, L. Å., Schilke, P., et al. 2006, A&A, 454, L13 1
Hosokawa, T. & Omukai, K. 2009, ApJ, 691, 823 1
Karska, A., Herpin, F., Bruderer, S., et al. 2014, A&A, 562, A45 5.4
Kasemann, C., Güsten, R., Heyminck, S., et al. 2006, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 6275, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 0 3.1
Kim, W.-J., Wyrowski, F., Urquhart, J. S., Menten, K. M., & Csengeri, T. 2017, A&A, 602, A37 2
Klein, T., Ciechanowicz, M., Leinz, C., et al. 2014, IEEE Trans. on Terahertz Science and Technology, 4, 588 3.2
König, C., Urquhart, J. S., Csengeri, T., et al. 2017, A&A, 599, A139 2, 4.1, 4.4, 5.2, A, C.2, C.2
Krumholz, M. R., Klein, R. I., McKee, C. F., Offner, S. S. R., & Cunningham, A. J. 2009, Science, 323, 754 1
Leurini, S., Codella, C., López-Sepulcre, A., et al. 2014, A&A, 570, A49 5.2
Leurini, S., Wyrowski, F., Herpin, F., et al. 2013, A&A, 550, A10 4.2
Leurini, S., Wyrowski, F., Wiesemeyer, H., et al. 2015, A&A, 584, A70 1
Molinari, S., Pezzuto, S., Cesaroni, R., et al. 2008, A&A, 481, 345 4.3
Molinari, S., Schisano, E., Elia, D., et al. 2016, A&A, 591, A149 4.3
Molinari, S., Swinyard, B., Bally, J., et al. 2010, PASP, 122, 314 2
Mottram, J. C., van Dishoeck, E. F., Kristensen, L. E., et al. 2017, A&A, 600, A99 1
Rosen, A. L., Krumholz, M. R., McKee, C. F., & Klein, R. I. 2016, MNRAS, 463, 2553 1
San José-García, I., Mottram, J. C., Kristensen, L. E., et al. 2013, A&A, 553, A125 1, 3.5, 5.1, 12
San José-García, I., Mottram, J. C., van Dishoeck, E. F., et al. 2016, A&A, 585, A103 5.4
Saraceno, P., Andre, P., Ceccarelli, C., Griffin, M., & Molinari, S. 1996, A&A, 309, 827 4.3
Schuller, F., Menten, K. M., Contreras, Y., et al. 2009, A&A, 504, 415 1
Schulz, A., Henkel, C., Beckmann, U., et al. 1995, A&A, 295, 183 1

Tang, X. D., Henkel, C., Wyrowski, F., et al. 2017, ArXiv e-prints [arXiv:1711.10012] 2
Urquhart, J. S., König, C., Giannetti, A., et al. 2018, MNRAS, 473, 1059 2, 4.3
Urquhart, J. S., Moore, T. J. T., Schuller, F., et al. 2013a, MNRAS, 431, 1752 4.3
Urquhart, J. S., Thompson, M. A., Moore, T. J. T., et al. 2013b, MNRAS, 435, 400 5.3
van Kempen, T. A., Kristensen, L. E., Herczeg, G. J., et al. 2010, A&A, 518, L121 5.4
van Kempen, T. A., van Dishoeck, E. F., Güsten, R., et al. 2009a, A&A, 507, 1425 1, 4.4, 12
van Kempen, T. A., van Dishoeck, E. F., Hogerheijde, M. R., & Güsten, R. 2009b, A&A, 508, 259 1
Weiß, A., Downes, D., Neri, R., et al. 2007, A&A, 467, 955 5.4
Wienen, M., Wyrowski, F., Schuller, F., et al. 2012, A&A, 544, A146 4.4
Wu, J., Evans, II, N. J., Gao, Y., et al. 2005, ApJ, 635, L173 4.3
Yorke, H. W. & Bodenheimer, P. 1999, ApJ, 525, 330 5.3
Zhang, X., Lee, Y., Bolatto, A., & Stark, A. A. 2001, ApJ, 553, 274 1
Zinnecker, H. & Yorke, H. W. 2007, ARA&A, 45, 481 1

Appendix A: Full tables

Here, we present the full version of tables shown in the paper. Table A.1 presents the properties of the observed clumps. Tables A.2, A.3 and A.4 present the Gaussian components of each source for the CO (4–3), CO (6–5) and CO (7–6) transitions, respectively. **The extension of the CO (6–5) emission is listed in Table A.5.** Table A.6 displays the integrated properties of the CO lines studied in this work. **Finally, the excitation temperature derived from the CO (6–5) spectra is presented in Table A.7.**

Table A.1: Properties of the TOP100 sources.

ID	CSC Name	RA(J2000) (HH:MM:SS)	DEC(J2000) (DD:MM:SS)	Offset CSC (″,″)	GCSC Name	Offset GCSC (″,″)	$V_{\rm lsr}$ (km s^{-1})	d (kpc)	$\log\left(\frac{L_{\rm bol}}{L_\odot}\right)$	$\log\left(\frac{M_{\rm clump}}{M_\odot}\right)$	Class
1	AGAL008.684−00.367	18:06:23.27	−21:37:12.7	(+2.5,+6.8)	G008.6834−0.3675	(+1.9,+4.7)	38	4.8	4.44	3.17	IRw
2	AGAL008.706−00.414	18:06:36.81	−21:37:18.1	(−1.9,−1.7)	G008.7064−0.4136	(−1.3,+1.5)	38.9	4.8	2.70	3.22	IRw
3	AGAL010.444−00.017	18:08:44.94	−19:54:32.0	(−3.2,−4.7)	G010.4446−0.0178	(−0.8,−4.0)	73.1	8.6	4.05	3.21	IRw
4	AGAL010.472+00.027	18:08:38.24	−19:51:49.6	(−1.0,+0.6)	G010.4722+0.0277	(−3.1,+0.3)	67	8.6	5.67	4.02	H II
5	AGAL010.624−00.384	18:10:28.87	−19:55:47.4	(+0.0,−0.7)	G010.6237−0.3833	(−4.6,−0.6)	−2.0	5.0	5.63	3.58	H II
6	AGAL012.804−00.199	18:14:13.75	−17:55:31.2	(−5.4,−13.6)	G012.8057−0.1994	(−2.9,−9.2)	35.4	2.4	5.39	3.27	H II
7	AGAL013.178+00.059	18:14:00.77	−17:28:37.8	(+6.8,−2.1)	G013.1768+0.0599	(+2.9,−3.0)	49.2	2.4	3.92	2.57	70w
8	AGAL013.658−00.599	18:17:24.25	−17:22:11.9	(+1.2,+2.1)	G013.6570−0.5992	(+1.6,+0.7)	47.4	4.5	4.32	2.76	IRb
9	AGAL014.114−00.574	18:18:13.21	−16:57:17.4	(−0.5,−2.3)	G014.1145−0.5745	(+1.6,−2.0)	19.1	2.6	3.50	2.55	IRw
10	AGAL014.194−00.194	18:16:58.81	−16:42:15.6	(+0.1,−1.7)	G014.1944−0.1939	(−0.2,−0.6)	40.4	3.9	3.43	2.91	IRw
11	AGAL014.492−00.139	18:17:22.19	−16:25:00.3	(−0.4,+2.0)	G014.4918−0.1389	(+0.2,+0.6)	41.7	3.9	2.88	3.28	70w
12	AGAL014.632−00.577	18:19:14.82	−16:30:02.0	(+7.7,+0.2)	G014.6323−0.5763	(+6.4,+1.7)	16.8	1.8	3.44	2.40	IRw
13	AGAL015.029−00.669	18:20:22.64	−16:11:42.7	(−3.6,+5.7)	G015.0292−0.6706	(+4.2,+1.8)	19.4	2.0	5.13	3.08	IRb
14	AGAL018.606−00.074	18:25:08.35	−12:45:22.8	(−0.8,−0.5)	G018.6057−0.0747	(+0.8,−2.0)	43.8	4.3	2.77	2.94	IRw
15	AGAL018.734−00.226	18:25:56.21	−12:42:49.3	(−2.6,−0.4)	G018.7344−0.2261	(+1.1,−0.3)	39.5	12.5	4.86	3.90	IRw
16	AGAL018.888−00.474	18:27:07.58	−12:41:39.5	(+1.2,+1.5)	G018.8870−0.4741	(−0.3,−0.3)	65.1	4.7	3.51	3.45	IRw
17	AGAL019.882−00.534	18:29:14.71	−11:50:25.4	(−6.5,−1.6)	G019.8832−0.5347	(−1.1,−0.4)	42.9	3.7	4.09	2.90	IRb
18	AGAL022.376+00.447	18:30:24.22	−09:10:38.9	(−1.4,+3.2)	G022.3752+0.4472	(−1.6,+1.0)	51.5	4.0	2.50	2.80	IRw
19	AGAL023.206−00.377	18:34:55.09	−08:49:18.1	(−0.4,+1.2)	G023.2056−0.3772	(−3.7,+1.1)	77.4	4.6	4.10	3.11	IRw
20	AGAL024.629+00.172	18:35:35.71	−07:18:08.7	(−2.7,−7.6)	G024.6294+0.1731	(−2.9,−6.0)	115.8	7.7	3.70	3.18	IRw
21	AGAL028.564−00.236	18:44:17.89	−03:59:44.3	(+1.8,+5.1)	G028.5637−0.2358	(−0.2,+3.8)	87.5	5.5	3.25	3.73	IRw
22	AGAL028.861+00.066	18:43:46.20	−03:35:29.2	(−2.0,−3.8)	G028.8614+0.0664	(−0.0,−0.9)	105.4	7.4	5.21	3.03	IRb
23	AGAL030.818−00.056	18:47:46.60	−01:54:30.1	(+2.0,+4.4)	G030.8166−0.0561	(+1.2,+1.0)	99	4.9	4.80	3.75	IRb
24	AGAL030.848−00.081	18:47:55.43	−01:53:37.7	(−1.1,+7.0)	G030.8472−0.0817	(+3.7,+4.4)	97.9	4.9	3.49	3.08	70w
25	AGAL030.893+00.139	18:47:13.69	−01:45:07.6	(−6.1,+2.4)	G030.8930+0.1383	(−1.9,+2.6)	105.6	4.9	2.70	3.28	70w
26	AGAL031.412+00.307	18:47:34.40	−01:12:46.5	(−1.9,+3.5)	G031.4120+0.3076	(−4.8,+2.3)	98	4.9	4.84	3.49	H II
27	AGAL034.258+00.154	18:53:18.68	+01:14:58.5	(−3.9,+1.6)	G034.2572+0.1535	(−1.2,−0.3)	58	1.6	4.68	2.91	H II
28	AGAL034.401+00.226	18:53:18.78	+01:24:38.7	(+0.4,−1.7)	G034.4005+0.2262	(−2.6,−2.4)	57	1.6	3.48	2.44	H II
29	AGAL034.411+00.234	18:53:18.31	+01:25:24.6	(−2.8,−1.9)	G034.4112+0.2344	(−1.0,−0.0)	57	1.6	3.68	2.33	IRb
30	AGAL034.821+00.351	18:53:38.29	+01:50:28.2	(−3.2,−0.4)	G034.8206+0.3504	(−3.8,−2.1)	57.7	1.6	3.44	2.05	IRb
31	AGAL035.197−00.742	18:58:13.09	+02:40:38.9	(−2.6,−0.5)	G035.1976−0.7427	(−0.2,−0.7)	33.9	2.2	4.37	2.67	IRb
32	AGAL037.554+00.201	18:59:10.06	+04:12:18.5	(−0.9,−2.0)	G037.5537+0.2006	(−1.5,−3.8)	84	6.7	4.71	3.10	IRb
33	AGAL043.166+00.011	19:10:13.64	+09:06:16.7	(−3.5,−6.3)	G043.1668+0.0115	(−3.0,−2.3)	7	11.1	6.58	4.64	H II
34	AGAL049.489−00.389	19:23:43.30	+14:30:26.5	(+13.5,+3.2)	G049.4888−0.3882	(+12.9,+3.7)	57.9	5.4	5.75	4.07	H II
35	AGAL053.141+00.069	19:29:17.52	+17:56:22.3	(+3.4,−4.8)	G053.1415+0.0701	(+0.1,−1.2)	22	1.6	3.36	1.98	IRb
36	AGAL059.782+00.066	19:43:11.06	+23:44:05.4	(−2.4,−1.6)	G059.7830+0.0657	(−2.1,+0.6)	22	2.2	3.99	2.41	IRb
37	AGAL301.136−00.226	12:35:35.51	−63:02:30.5	(−10.5,−1.6)	G301.1365−0.2256	(−5.2,−1.4)	−37.0	4.4	5.33	3.29	H II
38	AGAL305.192−00.006	13:11:14.92	−62:47:26.6	(−14.3,+0.2)	G305.1935−0.0059	(−4.3,+0.1)	−33.5	3.8	4.10	2.71	IRw
39	AGAL305.209+00.206	13:11:13.72	−62:34:38.5	(+6.7,−3.6)	G305.2083+0.2063	(−2.3,−2.3)	−42.6	3.8	4.95	3.15	IRb
40	AGAL305.562+00.014	13:14:26.54	−62:44:27.3	(−1.7,+3.2)	G305.5628+0.0137	(−0.3,+1.5)	−39.8	3.8	4.71	2.61	IRb
41	AGAL305.794−00.096	13:16:34.48	−62:49:45.7	(−25.3,+5.1)	G305.7949−0.0965	(−18.6,+2.8)	−40.7	3.8	2.99	2.77	70w
42	AGAL309.384−00.134	13:47:22.79	−62:18:09.8	(+21.0,+0.1)	G309.3826−0.1332	(+7.4,+1.9)	−49.5	5.3	4.20	3.08	IRb
43	AGAL310.014+00.387	13:51:38.19	−61:39:17.3	(+3.0,+4.1)	G310.0135+0.3877	(−0.4,+4.0)	−41.6	3.6	4.70	2.62	IRb

Table A.1: continued.

ID	CSC Name	RA(J2000) (HH:MM:SS)	DEC(J2000) (DD:MM:SS)	Offset CSC (″,″)	GCSC Name	Offset GCSC (″,″)	$V_{\rm lsr}$ (km s^{-1})	d (kpc)	$\log\left(\frac{L_{\rm bol}}{L_\odot}\right)$	$\log\left(\frac{M_{\rm clump}}{M_\odot}\right)$	Class
44	AGAL313.576+00.324	14:20:08.44	−60:42:07.0	(−3.2,+2.0)	G313.5763+0.3243	(+1.1,+2.7)	−45.5	3.8	3.97	2.26	IRb
45	AGAL316.641−00.087	14:44:18.46	−59:55:17.8	(+4.9,+4.5)	G316.6403−0.0877	(+2.1,+3.0)	−19.1	1.2	3.00	1.26	IRb
46	AGAL317.867−00.151	14:53:16.76	−59:26:36.9	(−1.7,+4.1)	G317.8680−0.1514	(+6.0,+3.2)	−40.3	3.0	3.22	2.56	IRw
47	AGAL318.779−00.137	14:59:33.29	−59:00:36.6	(−0.8,+1.9)	G318.7790−0.1376	(+1.0,+1.0)	−38.5	2.8	3.81	2.56	IRw
48	AGAL320.881−00.397	15:14:33.61	−58:11:31.9	(−7.8,+1.5)	G320.8803−0.3970	(−9.1,+2.2)	−46.0	10.0	3.78	3.46	70w
49	AGAL326.661+00.519	15:45:03.20	−54:09:13.9	(−5.2,+2.7)	G326.6607+0.5190	(−4.7,+2.4)	−39.6	1.8	3.87	2.10	IRb
50	AGAL326.987−00.032	15:49:08.36	−54:23:04.7	(−0.4,−3.1)	G326.9871−0.0317	(−4.8,−1.3)	−58.2	4.0	3.06	2.65	IRw
51	AGAL327.119+00.509	15:47:33.56	−53:52:41.9	(−6.7,−2.0)	G327.1197+0.5099	(−8.4,+0.8)	−83.0	5.5	4.77	2.83	IRb
52	AGAL327.393+00.199	15:50:19.55	−53:57:04.6	(−10.6,−0.3)	G327.3928+0.1984	(−8.8,−2.2)	−88.7	5.9	4.13	3.07	IRb
53	AGAL329.029−00.206	16:00:31.50	−53:12:38.8	(+6.7,−14.3)	G329.0303−0.2022	(−2.7,−1.8)	−43.5	11.5	5.33	4.06	IRw
54	AGAL329.066−00.307	16:01:09.93	−53:16:05.4	(−2.6,+2.1)	G329.0656−0.3076	(−1.3,+1.0)	−42.2	11.6	4.85	3.96	IRb
55	AGAL330.879−00.367	16:10:20.62	−52:06:11.0	(−4.9,+3.9)	G330.8788−0.3681	(−5.2,+1.3)	−63.6	4.2	5.19	3.20	H II
56	AGAL330.954−00.182	16:09:53.25	−51:54:54.8	(−6.0,+0.1)	G330.9545−0.1828	(−5.0,+0.1)	−90.6	9.3	6.12	4.24	H II
57	AGAL331.709+00.582	16:10:06.19	−50:50:28.6	(+7.8,+0.9)	G331.7084+0.5834	(+0.7,+1.4)	−67.1	10.5	4.57	3.71	IRw
58	AGAL332.094−00.421	16:16:16.62	−51:18:26.0	(+0.2,+1.7)	G332.0946−0.4210	(+0.8,+2.3)	−56.5	3.6	4.77	2.80	IRb
59	AGAL332.826−00.549	16:20:10.69	−50:53:19.6	(+3.3,+4.1)	G332.8262−0.5493	(+5.2,+4.6)	−57.3	3.6	5.38	3.29	H II
60	AGAL333.134−00.431	16:21:02.20	−50:35:12.6	(+4.0,+2.3)	G333.1341−0.4314	(+4.7,−0.6)	−51.9	3.6	5.62	3.46	H II
61	AGAL333.284−00.387	16:21:30.64	−50:26:54.3	(+5.8,−4.7)	G333.2841−0.3868	(+4.7,−3.9)	−51.5	3.6	5.11	3.32	H II
62	AGAL333.314+00.106	16:19:28.79	−50:04:42.9	(−1.5,+1.2)	G333.3139+0.1057	(−3.4,−0.4)	−45.8	3.6	4.03	2.63	IRb
63	AGAL333.604−00.212	16:22:09.58	−50:06:01.1	(−0.6,+1.7)	G333.6036−0.2130	(−1.5,−1.4)	−47.1	3.6	6.09	3.54	H II
64	AGAL333.656+00.059	16:21:11.83	−49:52:16.7	(−4.5,+0.6)	G333.6563+0.0587	(+0.8,+0.4)	−84.8	5.3	3.63	3.15	70w
65	AGAL335.789+00.174	16:29:47.62	−48:15:51.4	(−5.0,−0.4)	G335.7896+0.1737	(+2.1,+0.0)	−49.5	3.7	4.31	3.04	IRw
66	AGAL336.958−00.224	16:36:17.29	−47:40:49.1	(−3.7,+3.2)	G336.9574−0.2247	(−0.1,+1.5)	−73.1	10.9	3.56	3.38	IRw
67	AGAL337.176−00.032	16:36:18.70	−47:23:24.5	(+2.5,+4.3)	G337.1751−0.0324	(+0.3,+2.3)	−70.0	11.0	4.77	3.75	IRw
68	AGAL337.258−00.101	16:36:56.58	−47:22:29.1	(−4.8,+1.2)	G337.2580−0.1012	(−0.9,+1.3)	−67.6	11.0	4.48	3.50	IRw
69	AGAL337.286+00.007	16:36:34.63	−47:16:50.9	(−0.7,+0.6)	G337.2860+0.0083	(−1.0,+2.9)	−105.8	9.4	3.10	3.82	70w
70	AGAL337.406−00.402	16:38:51.00	−47:27:58.8	(−0.2,+1.0)	G337.4052−0.4024	(+0.1,−0.6)	−40.1	3.3	4.93	3.04	H II
71	AGAL337.704−00.054	16:38:29.69	−47:00:38.2	(−0.8,−2.9)	G337.7045−0.0535	(+1.0,−0.3)	−46.5	12.3	5.50	4.15	H II
72	AGAL337.916−00.477	16:41:10.51	−47:08:06.7	(+2.5,+2.3)	G337.9154−0.4773	(+2.5,+1.5)	−39.6	3.2	5.11	3.08	IRb
73	AGAL338.066+00.044	16:39:28.79	−46:40:30.4	(−4.2,−4.2)	G338.0663+0.0445	(−5.2,−2.1)	−67.6	4.7	3.50	2.98	70w
74	AGAL338.786+00.476	16:40:22.30	−45:51:05.3	(+3.6,−0.4)	G338.7851+0.4767	(−3.3,+0.3)	−64.6	4.5	2.69	3.09	70w
75	AGAL338.926+00.554	16:40:34.50	−45:41:46.7	(−3.2,+4.9)	G338.9249+0.5539	(−2.7,+2.9)	−61.9	4.4	4.97	3.78	IRw
76	AGAL339.623−00.122	16:46:06.21	−45:36:49.5	(+6.0,+2.9)	G339.6225−0.1220	(+5.8,+3.8)	−34.2	3.0	4.18	2.50	IRb
77	AGAL340.374−00.391	16:50:02.85	−45:12:45.2	(−5.8,+3.1)	G340.3736−0.3904	(−6.7,+2.4)	−46.5	3.6	2.71	2.90	IRw
78	AGAL340.746−01.001	16:54:04.02	−45:18:46.7	(−7.1,+1.8)	G340.7456−1.0014	(−6.3,−0.4)	−29.3	2.8	3.89	2.33	IRb
79	AGAL340.784−00.097	16:50:15.36	−44:42:30.1	(−4.7,−2.1)	G340.7848−0.0968	(−5.5,+1.3)	−101.4	10.0	4.86	3.45	IRw
80	AGAL341.217−00.212	16:52:18.19	−44:26:53.1	(−3.2,−1.2)	G341.2179−0.2122	(−1.0,+0.7)	−43.2	3.7	4.21	2.69	IRb
81	AGAL342.484+00.182	16:55:02.31	−43:12:59.2	(+2.3,−2.0)	G342.4836+0.1831	(−0.4,−2.7)	−38.3	12.6	4.81	3.69	IRw
82	AGAL343.128−00.062	16:58:17.47	−42:52:09.3	(−3.1,+4.8)	G343.1271−0.0632	(−1.0,+2.0)	−30.4	3.0	4.86	3.06	H II
83	AGAL343.756−00.164	17:00:50.14	−42:26:14.7	(−0.8,+2.0)	G343.7559−0.1640	(−1.5,+2.3)	−27.9	2.9	4.00	2.79	IRw
84	AGAL344.227−00.569	17:04:07.71	−42:18:41.3	(+2.3,+0.8)	G344.2275−0.5688	(−0.4,+1.7)	−23.0	2.5	3.99	3.05	IRw
85	AGAL345.003−00.224	17:05:11.26	−41:29:06.6	(−3.6,−1.2)	G345.0029−0.2241	(−3.3,−1.2)	−27.1	3.0	4.81	2.99	H II
86	AGAL345.488+00.314	17:04:28.26	−40:46:26.1	(−0.6,+0.7)	G345.4871+0.3142	(−0.9,−0.6)	−17.3	2.2	4.79	2.97	H II

Table A.1: continued.

ID	CSC Name	RA(J2000) (HH:MM:SS)	DEC(J2000) (DD:MM:SS)	Offset CSC ($''$,$''$)	GCSC Name	Offset GCSC ($''$,$''$)	$V_{\rm lsr}$ (km s^{-1})	d (kpc)	$\log\left(\frac{L_{\rm bol}}{L_\odot}\right)$	$\log\left(\frac{M_{\rm clump}}{M_\odot}\right)$	Class
87	AGAL345.504+00.347	17:04:23.18	−40:44:23.3	(−2.3,−1.6)	G345.5045+0.3481	(−5.8,+0.4)	−17.1	2.3	4.64	2.63	IRb
88	AGAL345.718+00.817	17:03:06.25	−40:17:04.2	(+0.5,−2.1)	G345.7172+0.8176	(−1.9,−2.5)	−12.0	1.6	3.27	2.30	IRb
89	AGAL351.131+00.771	17:19:34.58	−35:56:47.7	(−0.6,+0.3)	G351.1314+0.7709	(+1.5,+1.8)	−5.6	1.8	2.80	2.09	70w
90	AGAL351.161+00.697	17:19:56.97	−35:57:52.2	(+7.0,+2.2)	G351.1598+0.6982	(+0.4,+0.4)	−6.5	1.8	3.94	3.07	IRb
91	AGAL351.244+00.669	17:20:19.14	−35:54:42.0	(−10.5,−0.3)	G351.2437+0.6687	(−8.1,−3.0)	−3.9	1.8	4.89	2.95	IRb
92	AGAL351.416+00.646	17:20:53.65	−35:47:00.8	(−8.3,−2.2)	G351.4161+0.6464	(−12.8,−0.1)	−6.0	1.3	4.60	2.67	H II
93	AGAL351.444+00.659	17:20:55.49	−35:45:07.8	(−12.7,−4.0)	G351.4441+0.6579	(−7.9,−6.9)	−5.0	1.3	3.98	3.15	IRw
94	AGAL351.571+00.762	17:20:51.05	−35:35:22.4	(−2.4,−2.9)	G351.5719+0.7631	(−0.7,+1.7)	−3.7	1.3	2.64	2.22	70w
95	AGAL351.581−00.352	17:25:25.30	−36:12:47.2	(−4.7,+2.1)	G351.5815−0.3528	(+0.3,+3.0)	−93.7	6.8	5.39	3.94	IRb
96	AGAL351.774−00.537	17:26:42.55	−36:09:21.5	(+2.7,+0.7)	G351.7747−0.5369	(+2.5,+3.2)	−2.7	1.0	4.22	2.42	IRb
97	AGAL353.066+00.452	17:26:13.58	−34:31:54.8	(−3.5,+1.3)	G353.0670+0.4519	(−0.4,+3.6)	−2.1	0.9	1.76	1.25	IRw
98	AGAL353.417−00.079	17:29:19.13	−34:32:14.6	(−1.3,+9.7)	G353.4173−0.0803	(+2.9,+6.8)	−54.2	6.1	3.65	3.25	70w
99	AGAL354.944−00.537	17:35:12.04	−33:30:28.0	(−3.3,+3.9)	G354.9437−0.5381	(−4.2,+1.0)	−5.5	1.9	2.68	2.17	70w

Notes. The columns are as follows: (1) ID of each source; (2) name from the ATLASGAL-CSC catalogue from Contreras et al. (2013); (3)–(4) Equatorial coordinates (J2000) of the central position of the CHAMP$^+$ maps; (5) offset between the CSC and the **central position** of the CHAMP$^+$ maps; (6) name from the ATLASGAL-GCSC catalogue from Csengeri et al. (2014); (7) offset between GCSC and the **central position** of the CHAMP$^+$ maps; (8) the local standard rest velocity ($V_{\rm lsr}$) from the C^{17}O (3–2) data from Giannetti et al. (2014); (9)–(11) distances, bolometric luminosities and clump masses from König et al. (2017); (12) classification of the clump: **70** μm **weak (**70w**), infrared weak (**IRw**),** infrared bright (IRb) or H II regions (H II) from König et al. (2017).

Table A.2: **Gaussian decomposition** of the CO (4–3) profiles towards the TOP100 sample.

ID	CSC Name	C_1	V_{peak} (km s^{-1})	FWHM (km s^{-1})	T_{peak} (K)	C_2	V_{peak} (km s^{-1})	FWHM (km s^{-1})	T_{peak} (K)	C_3	V_{peak} (km s^{-1})	FWHM (km s^{-1})	T_{peak} (K)
1	AGAL008.684−00.367	N	36.7	6.4	16.4	B	40.6	14.3	3.2	–	–	–	–
2	AGAL008.706−00.414	N	38.1	7.1	4.9	–	–	–	–	–	–	–	–
3	AGAL010.444−00.017*	B	73.8	9.3	3.1	C2	65.2	4.9	2.1	–	–	–	–
4	AGAL010.472+00.027*	N	64.6	6.2	22.8	B	68.0	12.7	11.1	–	–	–	–
5	AGAL010.624−00.384	B	−2.8	12.7	51.2	–	–	–	–	–	–	–	–
6	AGAL012.804−00.199*	B	36.6	14.9	43.6	–	–	–	–	–	–	–	–
7	AGAL013.178+00.059	B	48.3	11.8	9.2	–	–	–	–	–	–	–	–
8	AGAL013.658−00.599	N	48.6	7.1	8.7	B	46.4	49.5	1.6	–	–	–	–
9	AGAL014.114−00.574	N	19.6	5.1	15.1	B	21.1	16.7	1.0	–	–	–	–
10	AGAL014.194−00.194	B1	40.1	9.1	10.4	B2	39.7	23.4	3.4	–	–	–	–
11	AGAL014.492−00.139	N	40.1	7.5	7.9	B	27.2	70.9	0.7	–	–	–	–
12	AGAL014.632−00.577	B1	17.4	7.6	14.0	B2	22.1	13.8	1.2	–	–	–	–
13	AGAL015.029−00.669*	B	19.6	9.3	55.0	–	–	–	–	–	–	–	–
14	AGAL018.606−00.074	B	45.2	7.5	9.9	–	–	–	–	–	–	–	–
15	AGAL018.734−00.226	N	39.4	4.9	11.0	B	39.6	27.8	1.4	–	–	–	–
16	AGAL018.888−00.474	B	65.7	11.9	14.3	–	–	–	–	–	–	–	–
17	AGAL019.882−00.534	B1	44.3	11.0	18.2	B2	43.1	27.8	6.5	B3	86.6	29.2	0.3
18	AGAL022.376+00.447	N	53.3	3.4	20.3	B	56.9	16.9	3.8	–	–	–	–
19	AGAL023.206−00.377	B1	78.1	16.7	13.0	B2	87.5	64.1	1.4	–	–	–	–
20	AGAL024.629+00.172	N	114.0	3.2	3.3	B	116.7	13.9	2.5	–	–	–	–
21	AGAL028.564−00.236	B	87.2	8.4	5.9	–	–	–	–	–	–	–	–
22	AGAL028.861+00.066	N	107.4	6.9	13.8	B	104.0	25.5	5.0	–	–	–	–
23	AGAL030.818−00.056	B1	97.3	8.0	25.3	B2	101.7	19.2	5.6	–	–	–	–
24	AGAL030.848−00.081*	N1	102.9	7.0	3.6	N2	94.4	7.5	9.2	–	–	–	–
25	AGAL030.893+00.139	N	106.7	6.2	8.3	C2	94.3	2.0	4.2	–	–	–	–
26	AGAL031.412+00.307	N	98.2	7.3	24.7	B	96.5	20.5	3.0	–	–	–	–
27	AGAL034.258+00.154*	B1	57.7	8.8	25.9	B2	58.4	23.4	6.7	–	–	–	–
28	AGAL034.401+00.226	B1	57.5	11.6	19.1	B2	56.8	51.0	1.1	–	–	–	–
29	AGAL034.411+00.234	B1	57.9	10.7	14.3	B2	62.9	42.9	1.6	–	–	–	–
30	AGAL034.821+00.351*	B1	57.2	12.4	9.9	B2	61.5	56.2	1.5	–	–	–	–
31	AGAL035.197−00.742	B1	34.5	9.0	27.1	B2	31.2	15.7	11.9	–	–	–	–
32	AGAL037.554+00.201	N	85.7	7.1	11.3	B	82.7	16.9	7.2	–	–	–	–
33	AGAL043.166+00.011*	B1	−1.1	11.8	39.0	B2	13.2	17.3	24.0	B3	2.8	33.4	6.6
34	AGAL049.489−00.389	N	52.2	6.8	27.7	B	56.6	20.1	16.1	–	–	–	–
35	AGAL053.141+00.069*	B1	23.3	9.2	20.6	B2	23.4	42.1	1.4	–	–	–	–
36	AGAL059.782+00.066	B1	22.2	8.1	24.1	B2	21.0	26.0	3.0	–	–	–	–
37	AGAL301.136−00.226	–	–	–	–	–	–	–	–	–	–	–	–
38	AGAL305.192−00.006*	B	−30.5	7.8	6.3	–	–	–	–	–	–	–	–
39	AGAL305.209+00.206*	N	−44.6	6.1	15.7	B	−40.3	17.3	18.2	–	–	–	–
40	AGAL305.562+00.014	N	−38.8	6.7	25.1	B	−41.0	14.2	13.6	–	–	–	–
41	AGAL305.794−00.096*	B	−42.4	8.5	2.2	–	–	–	–	–	–	–	–
42	AGAL309.384−00.134	B	−50.6	11.8	13.2	–	–	–	–	–	–	–	–
43	AGAL310.014+00.387	B1	−43.8	12.3	9.0	B2	−53.6	29.1	3.0	–	–	–	–
44	AGAL313.576+00.324	B1	−47.9	9.3	10.0	B2	−40.4	17.0	4.5	–	–	–	–
45	AGAL316.641−00.087	N	−16.9	7.2	9.9	B	−22.2	26.9	2.7	–	–	–	–
46	AGAL317.867−00.151	B	−38.4	10.3	11.9	–	–	–	–	–	–	–	–
47	AGAL318.779−00.137	B1	−38.4	8.8	6.0	B2	−47.4	72.7	1.0	–	–	–	–
48	AGAL320.881−00.397	N	−45.5	6.4	11.7	–	–	–	–	–	–	–	–
49	AGAL326.661+00.519	N1	−39.7	3.8	31.6	N2	−38.0	7.2	13.7	–	–	–	–
50	AGAL326.987−00.032	B1	−57.9	9.1	7.0	B2	−57.6	14.4	2.5	–	–	–	–
51	AGAL327.119+00.509*	B1	−84.2	10.5	8.9	B2	−85.6	33.8	1.6	–	–	–	–
52	AGAL327.393+00.199	B1	−89.3	9.4	10.5	B2	−86.3	86.0	1.0	–	–	–	–
53	AGAL329.029−00.206	B1	−46.6	12.3	15.5	B2	−52.9	26.4	1.6	–	–	–	–
54	AGAL329.066−00.307	B1	−41.9	11.1	7.2	B2	−47.6	15.0	3.1	–	–	–	–
55	AGAL330.879−00.367*	B1	−64.7	17.5	26.7	B2	−78.6	32.2	7.0	–	–	–	–
56	AGAL330.954−00.182*	B1	−91.4	13.4	16.0	B2	−96.8	24.3	16.0	–	–	–	–
57	AGAL331.709+00.582	B1	−67.2	9.9	15.3	B2	−64.7	24.2	9.2	–	–	–	–
58	AGAL332.094−00.421	B1	−57.6	12.2	16.9	B2	−56.4	29.9	3.2	–	–	–	–
59	AGAL332.826−00.549*	N	−56.1	6.9	77.9	B	−59.0	19.3	22.0	–	–	–	–

Table A.2: continued.

ID	CSC Name	C_1	V_{peak} (km s^{-1})	FWHM (km s^{-1})	T_{peak} (K)	C_2	V_{peak} (km s^{-1})	FWHM (km s^{-1})	T_{peak} (K)	C_3	V_{peak} (km s^{-1})	FWHM (km s^{-1})	T_{peak} (K)
60	AGAL333.134−00.431*	B1	−47.4	12.3	18.4	B2	−58.8	20.6	26.8	–	–	–	–
61	AGAL333.284−00.387	B	−52.3	8.0	41.8	–	–	–	–	–	–	–	–
62	AGAL333.314+00.106	B1	−45.7	11.3	16.5	B2	−49.9	51.2	2.9	–	–	–	–
63	AGAL333.604−00.212	N	−46.8	7.2	25.2	B	−46.4	24.0	28.8	–	–	–	–
64	AGAL333.656+00.059	N	−83.9	6.5	10.7	–	–	–	–	–	–	–	–
65	AGAL335.789+00.174	B1	−50.0	10.2	13.6	B2	−49.7	24.2	9.4	–	–	–	–
66	AGAL336.958−00.224	B	−71.9	8.8	7.1	–	–	–	–	–	–	–	–
67	AGAL337.176−00.032	N	−70.4	4.5	5.9	B	−70.5	16.9	3.1	C2	−79.3	2.9	4.1
68	AGAL337.258−00.101	N	−69.1	6.2	6.1	–	–	–	–	–	–	–	–
69	AGAL337.286+00.007*	B	−102.6	11.0	2.7	C2	-73.3	3.9	3.8	–	–	–	–
70	AGAL337.406−00.402	B1	−41.3	7.6	33.7	B2	−35.6	30.6	4.5	–	–	–	–
71	AGAL337.704−00.054	N	−48.8	4.6	12.9	B	−52.8	20.2	2.9	–	–	–	–
72	AGAL337.916−00.477	B1	−41.0	12.7	21.7	B2	−47.4	38.4	12.2	B3	2.9	91.2	2.3
73	AGAL338.066+00.044	N	−68.9	7.3	4.2	B	−63.9	34.6	1.8	C2	−39.0	9.7	3.8
74	AGAL338.786+00.476	N	−62.4	6.8	7.2	–	–	–	–	–	–	–	–
75	AGAL338.926+00.554*	B1	−62.3	8.6	56.7	B2	−62.4	34.5	4.2	–	–	–	–
76	AGAL339.623−00.122	B1	−32.3	11.2	15.6	B2	−29.1	30.9	3.2	–	–	–	–
77	AGAL340.374−00.391	B1	−42.7	8.0	2.2	B2	−43.4	37.0	1.7	–	–	–	–
78	AGAL340.746−01.001	N	−29.4	7.4	15.5	B2	−23.8	13.7	3.5	–	–	–	–
79	AGAL340.784−00.097	B1	−102.4	9.7	10.0	B2	−102.6	35.5	1.0	–	–	–	–
80	AGAL341.217−00.212	N	−42.0	7.1	22.1	B	−48.5	32.5	3.8	–	–	–	–
81	AGAL342.484+00.182	B1	−42.4	8.2	7.5	B2	−47.5	15.3	4.0	–	–	–	–
82	AGAL343.128−00.062	B1	−29.2	15.7	25.1	B2	−29.3	35.4	10.5	–	–	–	–
83	AGAL343.756−00.164	B	−27.3	10.5	13.4	–	–	–	–	–	–	–	–
84	AGAL344.227−00.569	B1	−23.8	16.4	10.0	B2	−23.9	87.3	2.6	–	–	–	–
85	AGAL345.003−00.224	B1	−27.4	10.1	26.6	B2	−26.9	31.0	6.6	–	–	–	–
86	AGAL345.488+00.314*	B1	−18.7	9.6	28.8	B2	−15.1	11.3	7.0	–	–	–	–
87	AGAL345.504+00.347	N	−16.9	6.7	28.6	B	−16.6	20.7	15.1	–	–	–	–
88	AGAL345.718+00.817	B	−13.0	9.0	11.4	–	–	–	–	–	–	–	–
89	AGAL351.131+00.771	N	−5.6	5.0	15.3	–	–	–	–	–	–	–	–
90	AGAL351.161+00.697	B1	−4.7	10.9	28.5	B2	−11.4	19.9	13.3	–	–	–	–
91	AGAL351.244+00.669	B1	−3.3	9.3	47.8	B2	3.6	26.0	5.5	–	–	–	–
92	AGAL351.416+00.646	B1	−6.7	9.4	46.8	B2	−18.9	28.6	14.8	–	–	–	–
93	AGAL351.444+00.659	B1	−4.2	11.6	27.0	B2	−9.7	32.4	5.4	–	–	–	–
94	AGAL351.571+00.762	N	−3.6	6.0	8.1	–	–	–	–	–	–	–	–
95	AGAL351.581−00.352	B1	−95.6	13.5	9.0	B2	−108.6	54.1	1.6	–	–	–	–
96	AGAL351.774−00.537	B1	−1.7	17.4	28.4	B2	−10.7	53.3	7.2	–	–	–	–
97	AGAL353.066+00.452	N	−1.6	5.2	5.4	–	–	–	–	–	–	–	–
98	AGAL353.417−00.079*	B	−49.7	10.0	3.3	C2	−37.3	3.4	2.4	–	–	–	–
99	AGAL354.944−00.537	N	−5.9	5.4	5.9	–	–	–	–	–	–	–	–

Notes. The columns are as follows: (1)-(2) ID and CSC name of the source (given in Table A.1); **a $*$ symbol indicates the cases where the Gaussian fit is not reliable, as discussed in Sect. 3.5**; (3), (7), (11) the classification of each fitted Gaussian component (C_1, C_2 and C_3), into narrow (N), broad (B) or secondary (C2) components as discussed in the main text; (4)-(6), (8)-(10), (12)-(14) the central velocity (V_{peak}), full width at half maximum (FWHM) and peak temperature (T_{peak}) is presented for each component.

Table A.3: **Gaussian decomposition** of the CO (6–5) profiles towards the TOP100 sample.

ID	CSC Name	C_1	V_{peak} (km s^{-1})	FWHM (km s^{-1})	T_{peak} (K)	C_2	V_{peak} (km s^{-1})	FWHM (km s^{-1})	T_{peak} (K)	C_3	V_{peak} (km s^{-1})	FWHM (km s^{-1})	T_{peak} (K)
1	AGAL008.684−00.367	N	38.0	7.2	24.9	B	42.4	17.8	3.3	–	–	–	–
2	AGAL008.706−00.414	N	38.3	6.3	3.4	–	–	–	–	–	–	–	–
3	AGAL010.444−00.017*	N	74.3	3.3	3.1	B	74.7	12.0	2.6	C2	65.4	2.0	1.7
4	AGAL010.472+00.027	N	66.0	7.1	72.7	B	69.0	14.9	13.3	–	–	–	–
5	AGAL010.624−00.384	B1	−2.4	11.4	108.1	B2	−2.5	24.1	7.8	–	–	–	–
6	AGAL012.804−00.199	B1	35.1	10.3	105.9	B2	39.8	18.5	28.4	–	–	–	–
7	AGAL013.178+00.059	B1	49.0	9.7	17.8	B2	59.6	26.0	0.8	–	–	–	–
8	AGAL013.658−00.599	N	48.2	5.5	12.2	B1	46.6	15.5	4.4	B2	42.1	62.3	1.4
9	AGAL014.114−00.574	N	19.8	6.3	22.2	B	19.6	20.0	3.0	–	–	–	–
10	AGAL014.194−00.194	N	39.8	5.3	13.6	B1	40.5	14.4	8.0	B2	35.8	29.1	1.8

Table A.3: continued.

ID	CSC Name	C_1	V_{peak} (km s^{-1})	FWHM (km s^{-1})	T_{peak} (K)	C_2	V_{peak} (km s^{-1})	FWHM (km s^{-1})	T_{peak} (K)	C_3	V_{peak} (km s^{-1})	FWHM (km s^{-1})	T_{peak} (K)
11	AGAL014.492−00.139	N	40.1	5.6	8.7	B	40.1	23.5	1.5	–	–	–	–
12	AGAL014.632−00.577	B1	18.2	7.9	26.4	B2	18.4	32.6	1.3	–	–	–	–
13	AGAL015.029−00.669*	N	20.9	6.9	75.3	B	18.6	10.9	60.3	–	–	–	–
14	AGAL018.606−00.074	N	45.8	4.4	15.0	B	45.6	10.4	4.8	–	–	–	–
15	AGAL018.734−00.226	N	40.1	6.4	12.9	B	40.1	26.6	3.2	–	–	–	–
16	AGAL018.888−00.474	B1	65.6	9.1	19.9	B2	69.4	20.7	3.1	–	–	–	–
17	AGAL019.882−00.534	N	45.3	6.5	18.5	B1	45.5	17.8	15.5	B3	64.1	70.9	1.8
18	AGAL022.376+00.447	N	52.5	3.1	11.6	B	53.8	10.4	4.1	–	–	–	–
19	AGAL023.206−00.377	B1	78.3	12.7	17.6	B2	77.8	37.3	3.5	–	–	–	–
20	AGAL024.629+00.172	B1	114.8	7.8	4.3	B2	120.5	40.1	0.5	–	–	–	–
21	AGAL028.564−00.236	N	86.1	5.6	4.6	B	87.6	16.1	1.6	–	–	–	–
22	AGAL028.861+00.066	B1	105.0	10.5	24.8	B2	102.0	32.0	6.4	–	–	–	–
23	AGAL030.818−00.056	N	97.5	6.6	35.9	B1	98.2	14.2	13.7	B2	100.3	52.5	2.4
24	AGAL030.848−00.081*	N	95.5	5.5	5.1	B	96.5	12.5	7.5	–	–	–	–
25	AGAL030.893+00.139	N	106.8	5.8	7.0	C2	97.0	5.2	4.5	–	–	–	–
26	AGAL031.412+00.307	N	97.5	7.3	42.8	B	100.2	29.3	4.2	–	–	–	–
27	AGAL034.258+00.154*	N	58.4	7.1	110.1	B1	57.9	15.1	23.2	B2	67.5	42.9	5.5
28	AGAL034.401+00.226	N	58.5	7.0	17.0	B1	58.0	12.1	26.1	B2	59.4	42.8	1.6
29	AGAL034.411+00.234	B1	57.9	10.1	19.6	B2	57.4	41.3	2.0	–	–	–	–
30	AGAL034.821+00.351	N	57.9	5.9	11.3	B	56.4	18.9	7.4	–	–	–	–
31	AGAL035.197−00.742	B1	34.3	9.5	46.2	B2	32.5	20.0	22.9	–	–	–	–
32	AGAL037.554+00.201	N	85.7	6.4	12.2	B	84.3	15.6	7.1	–	–	–	–
33	AGAL043.166+00.011*	B1	−0.5	12.9	56.0	B2	15.2	13.5	32.3	B3	2.1	42.3	13.1
34	AGAL049.489−00.389	N	50.8	6.4	34.1	B	58.6	29.3	10.9	–	–	–	–
35	AGAL053.141+00.069	N	22.3	7.2	37.2	B	21.0	17.0	10.8	–	–	–	–
36	AGAL059.782+00.066	B1	22.8	8.4	37.2	B2	22.0	24.6	5.3	–	–	–	–
37	AGAL301.136−00.226	B1	−38.3	9.9	34.4	B2	−37.3	31.1	14.0	B3	−28.8	65.4	10.4
38	AGAL305.192−00.006	N	−33.7	7.5	22.4	B	−34.4	19.9	4.6	–	–	–	–
39	AGAL305.209+00.206	B1	−42.4	10.9	43.7	B2	−40.1	33.2	10.1	–	–	–	–
40	AGAL305.562+00.014	N	−38.9	5.7	38.7	B	−40.0	14.9	21.2	–	–	–	–
41	AGAL305.794−00.096	B	−41.7	9.2	4.6	–	–	–	–	–	–	–	–
42	AGAL309.384−00.134	B1	−51.0	8.6	18.3	B2	−51.0	22.3	3.0	–	–	–	–
43	AGAL310.014+00.387	B1	−42.6	10.1	13.9	B2	−46.8	35.0	3.0	–	–	–	–
44	AGAL313.576+00.324	N	−46.7	7.4	19.4	B	−43.5	27.7	4.2	–	–	–	–
45	AGAL316.641−00.087	B1	−16.7	9.1	5.4	B2	−21.5	28.3	3.1	–	–	–	–
46	AGAL317.867−00.151	B	−39.9	11.2	15.7	–	–	–	–	–	–	–	–
47	AGAL318.779−00.137	N	−40.0	5.4	7.7	B	−38.7	20.7	2.2	–	–	–	–
48	AGAL320.881−00.397	N	−45.7	4.6	8.9	B	−44.6	10.5	2.9	–	–	–	–
49	AGAL326.661+00.519	N	−39.4	3.3	57.6	B	−38.4	7.6	18.1	–	–	–	–
50	AGAL326.987−00.032	N	−58.2	5.9	9.9	B1	−55.2	16.3	3.6	B2	−54.1	56.4	0.6
51	AGAL327.119+00.509	N	−83.9	6.3	18.1	B	−83.4	20.5	2.9	–	–	–	–
52	AGAL327.393+00.199	B1	−88.9	8.2	14.8	B2	−90.7	27.7	1.5	–	–	–	–
53	AGAL329.029−00.206	B1	−44.2	10.5	20.9	B2	−49.7	17.0	6.5	B3	−64.1	62.7	0.6
54	AGAL329.066−00.307	B1	−42.2	8.0	12.6	B2	−43.7	17.7	3.3	–	–	–	–
55	AGAL330.879−00.367	B1	−60.3	10.0	42.7	B2	−68.8	13.0	23.3	B3	−73.2	38.6	11.3
56	AGAL330.954−00.182	B1	−92.5	9.8	50.3	B2	−93.6	26.7	29.1	–	–	–	–
57	AGAL331.709+00.582	N	−66.1	5.7	8.6	B1	−66.3	13.4	14.9	B3	−63.9	36.3	2.9
58	AGAL332.094−00.421	B1	−57.6	12.4	12.5	B2	−54.5	23.7	9.2	–	–	–	–
59	AGAL332.826−00.549*	B1	−58.0	8.8	170.7	B2	−58.9	19.1	25.7	B3	−60.7	51.3	3.0
60	AGAL333.134−00.431*	B1	−50.0	9.6	34.7	B2	−55.9	22.3	47.5	–	–	–	–
61	AGAL333.284−00.387	B	−51.1	8.0	57.9	–	–	–	–	–	–	–	–
62	AGAL333.314+00.106	N	−45.1	7.1	11.1	B1	−47.1	15.2	12.2	B2	−45.9	64.9	1.2
63	AGAL333.604−00.212	B1	−44.1	11.9	85.0	B2	−48.1	25.3	38.4	–	–	–	–
64	AGAL333.656+00.059	N	−84.6	6.4	9.0	B	−84.5	23.4	0.8	–	–	–	–
65	AGAL335.789+00.174	N	−50.4	6.1	20.1	B1	−49.9	19.6	14.9	B2	−55.9	68.3	0.8
66	AGAL336.958−00.224	N	−72.4	4.8	6.7	B	−73.1	29.7	2.3	–	–	–	–
67	AGAL337.176−00.032	N	−69.8	3.4	12.2	B	−72.0	13.8	4.5	C2	−79.4	3.1	4.5
68	AGAL337.258−00.101	B	−67.8	9.6	9.2	–	–	–	–	–	–	–	–
69	AGAL337.286+00.007	N	−106.0	7.4	3.6	C2	−73.9	3.8	1.5	–	–	–	–
70	AGAL337.406−00.402	N	−40.7	6.8	69.8	B1	−40.3	23.5	8.7	B2	−28.2	83.2	3.1

Table A.3: continued.

ID	CSC Name	C_1	V_{peak} (km s^{-1})	FWHM (km s^{-1})	T_{peak} (K)	C_2	V_{peak} (km s^{-1})	FWHM (km s^{-1})	T_{peak} (K)	C_3	V_{peak} (km s^{-1})	FWHM (km s^{-1})	T_{peak} (K)
71	AGAL337.704−00.054*	N	−50.8	3.9	15.0	B1	−45.5	13.0	15.2	B2	−56.2	24.0	2.5
72	AGAL337.916−00.477	B1	−40.1	10.7	38.1	B1	−43.7	28.1	18.5	B2	−52.0	56.9	6.5
73	AGAL338.066+00.044	N	−69.3	4.9	4.5	B	−68.7	21.7	2.3	C2	−39.9	8.2	3.7
74	AGAL338.786+00.476	N	−63.8	7.1	6.0	–	–	–	–	–	–	–	–
75	AGAL338.926+00.554*	N	−62.5	6.4	131.4	B1	−60.4	16.3	11.3	B2	−62.5	49.1	2.5
76	AGAL339.623−00.122	N	−32.9	6.2	14.1	B1	−31.9	18.0	12.7	B2	−25.6	58.6	1.8
77	AGAL340.374−00.391	B1	−44.5	10.7	5.3	B	−38.7	50.8	1.3	–	–	–	–
78	AGAL340.746−01.001	N	−29.3	6.9	18.3	B1	−23.8	13.4	2.7	B2	−15.3	49.0	1.0
79	AGAL340.784−00.097	N	−101.4	6.9	8.1	B	−101.0	18.3	2.6	–	–	–	–
80	AGAL341.217−00.212	N	−43.5	4.1	26.6	B1	−41.5	12.3	7.5	B2	−44.9	37.9	3.6
81	AGAL342.484+00.182	N	−42.2	5.9	8.4	B	−45.1	17.2	6.4	–	–	–	–
82	AGAL343.128−00.062	B1	−29.2	16.5	46.2	B2	−26.2	39.8	14.1	–	–	–	–
83	AGAL343.756−00.164	B1	−27.2	8.4	19.1	B2	−25.5	23.2	1.2	–	–	–	–
84	AGAL344.227−00.569	B1	−22.3	10.1	17.8	B2	−27.8	29.3	6.1	–	–	–	–
85	AGAL345.003−00.224	B1	−26.9	7.5	51.2	B2	−27.3	20.6	17.8	B3	−13.4	97.1	2.6
86	AGAL345.488+00.314*	N	−18.7	7.1	87.0	B	−15.6	13.5	15.3	–	–	–	–
87	AGAL345.504+00.347	B1	−16.8	8.2	31.3	B2	−16.0	19.4	19.7	B3	−1.6	43.3	1.5
88	AGAL345.718+00.817	B	−12.2	7.8	14.4	–	–	–	–	–	–	–	–
89	AGAL351.131+00.771	N	−5.1	4.7	17.5	–	–	–	–	–	–	–	–
90	AGAL351.161+00.697	B1	−5.3	9.2	60.2	B2	−8.5	22.2	26.6	–	–	–	–
91	AGAL351.244+00.669	B1	−2.9	8.6	96.6	B2	0.6	25.4	13.4	B3	8.7	72.3	1.4
92	AGAL351.416+00.646	B1	−6.5	9.5	70.5	B2	−12.4	34.2	18.8	B3	−25.1	96.5	3.2
93	AGAL351.444+00.659	B1	−4.1	8.9	37.7	B2	−4.1	15.0	19.9	B3	−8.3	40.0	8.4
94	AGAL351.571+00.762	N	−4.1	3.5	8.2	–	–	–	–	–	–	–	–
95	AGAL351.581−00.352	B1	−95.4	10.7	23.9	B2	−97.2	29.2	10.2	B3	−131.9	70.0	0.6
96	AGAL351.774−00.537	B1	−1.3	10.6	36.0	B2	−0.8	26.3	22.9	B3	−10.2	61.3	14.9
97	AGAL353.066+00.452*	B	−1.0	7.6	16.5	–	–	–	–	–	–	–	–
98	AGAL353.417−00.079*	N	−55.6	2.7	1.9	B	−53.8	9.0	1.8	–	–	–	–
99	AGAL354.944−00.537*	N	−5.7	2.6	13.1	B	−5.2	8.8	2.0	–	–	–	–

Notes. The columns are as follows: (1)-(2) ID and CSC name of the source (given in Table A.1); **a $*$ symbol indicates the cases where the Gaussian fit is not reliable, as discussed in Sect. 3.5**; (3), (7), (11) the classification of each fitted Gaussian component (C_1, C_2 and C_3), into narrow (N), broad (B) or secondary (C2) components as discussed in the main text; (4)-(6), (8)-(10), (12)-(14) the central velocity (V_{peak}), full width at half maximum (FWHM) and peak temperature (T_{peak}) is presented for each component.

Table A.4: **Gaussian decomposition** of the CO (7–6) profiles towards the TOP100 sample.

ID	CSC Name	C_1	V_{peak} (km s^{-1})	FWHM (km s^{-1})	T_{peak} (K)	C_2	V_{peak} (km s^{-1})	FWHM (km s^{-1})	T_{peak} (K)	C_3	V_{peak} (km s^{-1})	FWHM (km s^{-1})	T_{peak} (K)
1	AGAL008.684−00.367	B1	38.3	9.4	15.1	B2	50.2	27.1	1.8	–	–	–	–
2	AGAL008.706−00.414	N	39.8	6.8	2.4	–	–	–	–	–	–	–	–
3	AGAL010.444−00.017	N	75.0	7.0	3.8	C2	66.6	3.7	1.8	–	–	–	–
4	AGAL010.472+00.027	N	66.1	7.2	63.4	B	70.6	20.1	8.2	–	–	–	–
5	AGAL010.624−00.384	B1	−2.4	9.7	88.4	B2	−1.6	17.4	30.2	–	–	–	–
6	AGAL012.804−00.199	B1	35.0	10.0	107.1	B2	39.8	17.4	28.4	–	–	–	–
7	AGAL013.178+00.059	B	49.2	11.0	16.0	–	–	–	–	–	–	–	–
8	AGAL013.658−00.599	N	48.3	6.5	10.0	B	43.6	69.9	2.9	–	–	–	–
9	AGAL014.114−00.574	N	19.4	4.2	20.4	B	19.9	14.3	6.3	–	–	–	–
10	AGAL014.194−00.194	B1	40.1	7.7	11.1	B2	41.3	29.2	3.6	–	–	–	–
11	AGAL014.492−00.139*	B	39.8	10.3	5.5	–	–	–	–	–	–	–	–
12	AGAL014.632−00.577	N	17.9	7.1	26.3	B	21.3	20.2	1.8	–	–	–	–
13	AGAL015.029−00.669*	B1	20.7	7.7	104.6	B2	18.8	13.7	25.0	–	–	–	–
14	AGAL018.606−00.074	B	45.2	7.5	10.2	–	–	–	–	–	–	–	–
15	AGAL018.734−00.226	N	40.4	6.7	10.1	B	44.9	46.0	2.6	–	–	–	–
16	AGAL018.888−00.474	N	65.8	5.7	13.2	B	66.0	18.0	6.2	–	–	–	–
17	AGAL019.882−00.534	N	45.1	6.6	13.1	B1	45.0	17.2	14.9	B2	65.7	96.9	2.3
18	AGAL022.376+00.447	N	53.3	6.3	7.6	–	–	–	–	–	–	–	–
19	AGAL023.206−00.377	B1	78.0	11.6	18.4	B2	78.2	50.0	3.6	–	–	–	–
20	AGAL024.629+00.172	B	115.8	7.8	3.7	–	–	–	–	–	–	–	–
21	AGAL028.564−00.236	N	86.4	4.0	3.3	–	–	–	–	–	–	–	–

Table A.4: continued.

ID	CSC Name	C_1	V_{peak} (km s^{-1})	FWHM (km s^{-1})	T_{peak} (K)	C_2	V_{peak} (km s^{-1})	FWHM (km s^{-1})	T_{peak} (K)	C_3	V_{peak} (km s^{-1})	FWHM (km s^{-1})	T_{peak} (K)
22	AGAL028.861+00.066	B1	104.6	12.2	21.5	B2	80.0	81.9	3.4	–	–	–	–
23	AGAL030.818−00.056	N	97.9	7.4	48.2	B	96.6	50.5	5.6	–	–	–	–
24	AGAL030.848−00.081*	B	96.2	10.9	9.5	–	–	–	–	–	–	–	–
25	AGAL030.893+00.139	N	106.3	2.0	7.2	C2	95.0	3.5	3.9	–	–	–	–
26	AGAL031.412+00.307*	B1	97.6	7.5	40.5	B2	97.4	85.6	2.7	–	–	–	–
27	AGAL034.258+00.154*	B1	57.7	8.9	71.7	B2	63.8	37.4	8.7	–	–	–	–
28	AGAL034.401+00.226	B1	58.0	10.3	29.6	B2	63.5	64.3	2.0	–	–	–	–
29	AGAL034.411+00.234	B1	58.2	10.4	16.3	B2	61.9	65.6	2.3	–	–	–	–
30	AGAL034.821+00.351	N	57.7	6.9	11.4	B	59.2	28.2	5.0	–	–	–	–
31	AGAL035.197−00.742	B1	34.6	8.9	38.8	B2	32.8	20.1	19.6	–	–	–	–
32	AGAL037.554+00.201	B1	85.5	7.7	12.8	B2	88.0	52.3	3.1	–	–	–	–
33	AGAL043.166+00.011*	B1	15.9	12.1	30.9	B2	0.2	14.1	49.8	B3	−1.9	70.9	6.5
34	AGAL049.489−00.389	N	51.1	5.8	36.3	B	59.2	27.1	17.3	–	–	–	–
35	AGAL053.141+00.069	N	22.4	6.3	30.5	B	21.1	16.0	11.8	–	–	–	–
36	AGAL059.782+00.066	B1	22.7	8.7	29.6	B2	21.6	28.0	4.0	–	–	–	–
37	AGAL301.136−00.226	B1	−38.4	9.5	37.8	B2	−36.8	32.7	14.8	B3	−22.9	70.3	7.8
38	AGAL305.192−00.006	B1	−33.9	7.9	15.2	B2	−34.1	24.5	3.0	–	–	–	–
39	AGAL305.209+00.206	B1	−42.5	10.7	45.0	B2	−37.5	43.1	6.8	–	–	–	–
40	AGAL305.562+00.014	N	−38.7	5.9	50.2	B	−42.7	17.8	11.3	–	–	–	–
41	AGAL305.794−00.096*	B	−42.5	11.1	5.7	–	–	–	–	–	–	–	–
42	AGAL309.384−00.134	B	−51.3	10.4	17.4	–	–	–	–	–	–	–	–
43	AGAL310.014+00.387	B1	−42.7	8.7	15.5	B2	−43.1	35.1	4.7	–	–	–	–
44	AGAL313.576+00.324	N	−46.8	6.4	19.4	B	−43.0	25.2	3.4	–	–	–	–
45	AGAL316.641−00.087	B	−17.8	17.1	5.9	–	–	–	–	–	–	–	–
46	AGAL317.867−00.151	B	−39.7	8.2	17.3	–	–	–	–	–	–	–	–
47	AGAL318.779−00.137	B	−38.3	7.9	10.3	–	–	–	–	–	–	–	–
48	AGAL320.881−00.397	N	−45.4	4.9	8.6	–	–	–	–	–	–	–	–
49	AGAL326.661+00.519	N	−39.4	4.0	69.0	–	–	–	–	–	–	–	–
50	AGAL326.987−00.032	B	−57.4	8.1	12.1	–	–	–	–	–	–	–	–
51	AGAL327.119+00.509	B	−84.0	9.9	11.4	–	–	–	–	–	–	–	–
52	AGAL327.393+00.199	B	−88.9	9.4	14.8	–	–	–	–	–	–	–	–
53	AGAL329.029−00.206	N	−43.8	7.1	15.8	B	−46.4	16.2	11.0	–	–	–	–
54	AGAL329.066−00.307	N	−42.0	6.7	12.4	B	−47.5	29.0	3.1	–	–	–	–
55	AGAL330.879−00.367	B1	−61.3	12.7	40.6	B2	−69.5	22.2	19.3	B3	−78.8	88.3	3.6
56	AGAL330.954−00.182	B1	−92.4	10.0	51.6	B2	−94.1	26.8	24.4	–	–	–	–
57	AGAL331.709+00.582	B1	−65.9	8.9	11.6	B2	−66.5	29.4	5.5	–	–	–	–
58	AGAL332.094−00.421	B1	−57.4	7.6	19.4	B2	−55.2	26.4	9.9	–	–	–	–
59	AGAL332.826−00.549*	B1	−57.8	12.0	95.4	B2	−59.6	33.9	7.0	B3	−64.9	118.0	0.9
60	AGAL333.134−00.431*	B1	−50.2	8.9	40.0	B2	−55.2	22.0	49.8	–	–	–	–
61	AGAL333.284−00.387	B	−51.2	7.6	55.9	–	–	–	–	–	–	–	–
62	AGAL333.314+00.106	B1	−45.9	10.7	16.9	B2	−58.7	90.9	2.2	–	–	–	–
63	AGAL333.604−00.212	B1	−44.1	11.9	94.3	B2	−49.0	24.1	40.8	–	–	–	–
64	AGAL333.656+00.059	B	−85.7	9.9	5.5	–	–	–	–	–	–	–	–
65	AGAL335.789+00.174	N	−49.7	5.2	20.1	B1	−50.6	19.5	13.3	B2	−72.3	116.4	1.2
66	AGAL336.958−00.224	N	−71.7	7.3	6.3	–	–	–	–	–	–	–	–
67	AGAL337.176−00.032	B1	−68.6	8.2	11.2	B2	−66.7	32.7	2.4	C2	−79.2	4.7	5.2
68	AGAL337.258−00.101	B	−68.1	10.1	8.6	–	–	–	–	–	–	–	–
69	AGAL337.286+00.007	B	−105.6	8.3	3.0	–	–	–	–	–	–	–	–
70	AGAL337.406−00.402	N	−40.9	5.4	75.5	B	−40.5	16.2	13.2	B	−24.9	61.1	2.0
71	AGAL337.704−00.054	B1	−47.2	11.1	24.7	B2	−40.4	68.9	3.0	–	–	–	–
72	AGAL337.916−00.477	B1	−40.3	10.6	40.8	B2	−46.1	33.1	18.4	B3	−46.9	113.4	3.0
73	AGAL338.066+00.044	B	−69.9	13.3	3.3	C2	−13.5	10.5	2.8	–	–	–	–
74	AGAL338.786+00.476*	B	−63.7	8.9	5.7	–	–	–	–	–	–	–	–
75	AGAL338.926+00.554*	B1	−62.3	8.6	43.2	B2	−60.9	26.9	6.5	–	–	–	–
76	AGAL339.623−00.122	N	−33.8	7.0	14.3	B	−32.2	21.6	9.7	–	–	–	–
77	AGAL340.374−00.391	B1	−44.3	8.2	6.6	B2	−43.5	61.0	1.6	–	–	–	–
78	AGAL340.746−01.001	B1	−28.9	7.5	15.6	B2	−18.6	30.2	3.4	–	–	–	–
79	AGAL340.784−00.097	B	−101.6	10.1	8.0	–	–	–	–	–	–	–	–
80	AGAL341.217−00.212	N	−43.1	5.6	28.2	B	−46.2	68.9	3.7	–	–	–	–
81	AGAL342.484+00.182	B1	−42.4	9.6	9.6	B2	−40.8	67.9	2.3	–	–	–	–

Table A.4: continued.

ID	CSC Name	C_1	V_{peak} (km s^{-1})	FWHM (km s^{-1})	T_{peak} (K)	C_2	V_{peak} (km s^{-1})	FWHM (km s^{-1})	T_{peak} (K)	C_3	V_{peak} (km s^{-1})	FWHM (km s^{-1})	T_{peak} (K)
82	AGAL343.128−00.062	B1	−29.1	18.4	47.8	B2	−23.3	52.3	8.2	–	–	–	–
83	AGAL343.756−00.164	B	−27.1	7.8	22.2	–	–	–	–	–	–	–	–
84	AGAL344.227−00.569	B1	−22.1	7.7	16.2	B2	−25.3	30.3	8.0	–	–	–	–
85	AGAL345.003−00.224	B1	−26.6	8.7	38.9	B2	−27.3	21.4	17.1	B3	−13.3	108.3	3.3
86	AGAL345.488+00.314*	N	−18.7	5.9	122.9	B	−17.5	11.9	24.7	–	–	–	–
87	AGAL345.504+00.347	B1	−16.7	9.4	37.4	B2	−14.1	24.9	9.2	–	–	–	–
88	AGAL345.718+00.817	N	−11.5	5.5	15.3	–	–	–	–	–	–	–	–
89	AGAL351.131+00.771	N	−5.4	4.7	13.7	–	–	–	–	–	–	–	–
90	AGAL351.161+00.697	B1	−6.4	11.0	79.2	B2	−8.9	28.0	11.5	–	–	–	–
91	AGAL351.244+00.669	B1	−2.9	8.4	108.6	B2	0.8	25.7	13.7	B3	19.9	120.2	2.0
92	AGAL351.416+00.646	B1	−7.0	8.5	59.0	B2	−11.7	31.7	17.0	B3	−19.5	76.3	5.4
93	AGAL351.444+00.659	B1	−4.3	11.0	45.6	B2	−6.9	28.0	12.2	–	–	–	–
94	AGAL351.571+00.762	N	−4.2	4.4	11.9	–	–	–	–	–	–	–	–
95	AGAL351.581−00.352	B1	−96.0	14.6	18.4	B2	−101.1	58.1	4.4	–	–	–	–
96	AGAL351.774−00.537	B1	−1.6	16.4	40.8	B2	−6.1	57.8	20.2	–	–	–	–
97	AGAL353.066+00.452	N	−1.2	6.6	10.6	–	–	–	–	–	–	–	–
98	AGAL353.417−00.079*	N	−55.4	5.1	6.1	–	–	–	–	–	–	–	–
99	AGAL354.944−00.537	N	−5.2	6.4	5.1	–	–	–	–	–	–	–	–

Notes. The columns are as follows: (1)-(2) ID and CSC name of the source (given in Table A.1); **a $*$ symbol indicates the cases where the Gaussian fit is not reliable, as discussed in Sect. 3.5**; (3), (7), (11) the classification of each fitted Gaussian component (C_1, C_2 and C_3), into narrow (N), broad (B) or secondary (C2) components as discussed in the main text; (4)-(6), (8)-(10), (12)-(14) the central velocity (V_{peak}), full width at half maximum (FWHM) and peak temperature (T_{peak}) is presented for each component.

Table A.5: Extension of the CO (6–5) and the *Herschel*-PACS 70 μm emission towards the TOP100 clumps.

ID	CSC Name	CO (6–5)					PACS 70 μm				
		$\Delta\theta_{max}$ ($''$)	$\Delta\theta_{min}$ ($''$)	$\Delta\theta_{avg}$ ($''$)	Δs_{avg} (pc)	σ_s (pc)	$\Delta\theta_{max}$ ($''$)	$\Delta\theta_{min}$ ($''$)	$\Delta\theta_{avg}$ ($''$)	Δs_{avg} (pc)	σ_s (pc)
1	AGAL008.684−00.367	16.8	11.7	14.2	0.331	0.085	16.6	14.2	15.4	0.357	0.077
2	AGAL008.706−00.414	30.5	20.3	25.4	0.589	0.130	–	–	–	–	–
3	AGAL010.444−00.017	38.1	17.2	27.6	1.146	0.256	11.8	11.4	11.6	0.483	0.108
4	AGAL010.472+00.027	41.2	21.1	31.2	1.292	0.280	16.6	15.4	16.0	0.663	0.142
5	AGAL010.624−00.384	47.3	29.7	38.5	0.924	0.193	24.5	22.5	23.5	0.564	0.116
6	AGAL012.804−00.199	88.5	57.9	73.2	0.852	0.173	49.7	26.8	38.2	0.446	0.091
7	AGAL013.178+00.059	47.3	29.7	38.5	0.448	0.094	–	–	–	–	–
8	AGAL013.658−00.599	25.9	12.5	19.2	0.416	0.101	14.2	12.6	13.4	0.291	0.064
9	AGAL014.114−00.574	45.8	19.6	32.7	0.407	0.089	46.6	9.9	28.2	0.352	0.077
10	AGAL014.194−00.194	27.5	18.8	23.1	0.438	0.098	14.2	11.1	12.6	0.239	0.053
11	AGAL014.492−00.139	36.6	21.9	29.2	0.549	0.119	–	–	–	–	–
12	AGAL014.632−00.577	36.6	25.8	31.2	0.277	0.059	16.6	14.6	15.6	0.138	0.030
13	AGAL015.029−00.669	91.5	68.9	80.2	0.770	0.155	40.7	27.6	34.2	0.328	0.067
14	AGAL018.606−00.074	24.4	20.3	22.3	0.470	0.105	12.6	10.3	11.5	0.240	0.054
15	AGAL018.734−00.226	24.4	12.5	18.5	1.117	0.272	23.7	14.2	19.0	1.146	0.243
16	AGAL018.888−00.474	62.5	36.8	49.7	1.141	0.235	17.0	14.6	15.8	0.363	0.078
17	AGAL019.882−00.534	35.1	20.3	27.7	0.492	0.107	16.2	13.4	14.8	0.263	0.057
18	AGAL022.376+00.447	13.7	11.7	12.7	0.244	0.065	11.4	9.9	10.6	0.205	0.047
19	AGAL023.206−00.377	13.7	12.5	13.1	0.292	0.076	14.2	13.0	13.6	0.303	0.066
20	AGAL024.629+00.172	13.7	12.5	13.1	0.491	0.128	13.8	12.2	13.0	0.488	0.107
21	AGAL028.564−00.236	38.1	27.4	32.8	0.865	0.183	–	–	–	–	–
22	AGAL028.861+00.066	27.5	15.7	21.6	0.775	0.178	14.6	13.8	14.2	0.511	0.111
23	AGAL030.818−00.056	51.9	30.5	41.2	0.979	0.204	20.1	18.6	19.4	0.459	0.096
24	AGAL030.848−00.081	91.5	64.2	77.8	1.849	0.374	–	–	–	–	–
25	AGAL030.893+00.139	38.1	20.3	29.2	0.694	0.151	–	–	–	–	–
26	AGAL031.412+00.307	19.8	18.8	19.3	0.458	0.105	17.4	14.2	15.8	0.375	0.080
27	AGAL034.258+00.154	41.2	29.7	35.5	0.268	0.056	26.8	21.3	24.0	0.182	0.038
28	AGAL034.401+00.226	51.9	24.3	38.1	0.288	0.061	15.8	13.8	14.8	0.112	0.024
29	AGAL034.411+00.234	19.8	18.0	18.9	0.143	0.033	13.0	12.2	12.6	0.096	0.021
30	AGAL034.821+00.351	27.5	14.1	20.8	0.157	0.037	11.8	11.4	11.6	0.088	0.020
31	AGAL035.197−00.742	88.5	31.3	59.9	0.636	0.132	15.8	15.4	15.6	0.166	0.035

Table A.5: continued.

ID	CSC Name	CO (6–5)					PACS 70 μm				
		$\Delta\theta_{\rm max}$ ($''$)	$\Delta\theta_{\rm min}$ ($''$)	$\Delta\theta_{\rm avg}$ ($''$)	$\Delta s_{\rm avg}$ (pc)	$\sigma_{\rm s}$ (pc)	$\Delta\theta_{\rm max}$ ($''$)	$\Delta\theta_{\rm min}$ ($''$)	$\Delta\theta_{\rm avg}$ ($''$)	$\Delta s_{\rm avg}$ (pc)	$\sigma_{\rm s}$ (pc)
32	AGAL037.554+00.201	18.3	12.5	15.4	0.501	0.125	13.0	11.8	12.4	0.404	0.090
33	AGAL043.166+00.011	61.0	29.0	45.0	2.423	0.505	33.6	25.3	29.4	1.584	0.324
34	AGAL049.489−00.389	114.5	43.1	78.8	2.066	0.440	26.6	24.8	25.7	0.674	0.147
35	AGAL053.141+00.069	61.0	18.8	39.9	0.309	0.067	15.8	13.8	14.8	0.115	0.025
36	AGAL059.782+00.066	62.5	21.9	42.2	0.442	0.094	15.0	12.2	13.6	0.143	0.031
37	AGAL301.136−00.226	18.3	12.5	15.4	0.329	0.082	20.1	19.3	19.7	0.421	0.088
38	AGAL305.192−00.006	35.1	25.8	30.4	0.561	0.120	13.0	11.8	12.4	0.229	0.051
39	AGAL305.209+00.206	45.8	23.5	34.7	0.638	0.136	15.0	13.4	14.2	0.262	0.057
40	AGAL305.562+00.014	48.8	26.6	37.7	0.695	0.146	17.0	15.8	16.4	0.302	0.064
41	AGAL305.794−00.096	77.8	29.0	53.4	0.983	0.205	–	–	–	–	–
42	AGAL309.384−00.134	38.1	19.6	28.8	0.746	0.163	13.4	10.3	11.9	0.307	0.069
43	AGAL310.014+00.387	30.5	26.6	28.5	0.500	0.107	14.2	13.4	13.8	0.242	0.053
44	AGAL313.576+00.324	18.3	14.1	16.2	0.297	0.072	11.4	9.5	10.4	0.192	0.044
45	AGAL316.641−00.087	19.8	12.5	16.1	0.093	0.023	13.4	11.8	12.6	0.073	0.016
46	AGAL317.867−00.151	41.2	19.6	30.4	0.435	0.095	14.6	13.0	13.8	0.198	0.043
47	AGAL318.779−00.137	25.9	15.7	20.8	0.280	0.065	–	–	–	–	–
48	AGAL320.881−00.397	38.1	34.4	36.2	1.753	0.366	–	–	–	–	–
49	AGAL326.661+00.519	44.2	33.7	39.0	0.343	0.071	15.0	13.8	14.4	0.127	0.028
50	AGAL326.987−00.032	33.6	19.6	26.6	0.509	0.112	15.0	11.4	13.2	0.253	0.056
51	AGAL327.119+00.509	36.6	25.0	30.8	0.823	0.176	12.6	11.4	12.0	0.322	0.072
52	AGAL327.393+00.199	29.0	21.1	25.0	0.719	0.158	17.8	13.0	15.4	0.442	0.095
53	AGAL329.029−00.206	50.3	19.6	35.0	1.946	0.423	39.5	13.8	26.6	1.484	0.313
54	AGAL329.066−00.307	35.1	26.6	30.8	1.732	0.369	12.6	11.1	11.9	0.665	0.149
55	AGAL330.879−00.367	25.9	21.9	23.9	0.482	0.106	18.6	17.4	18.0	0.362	0.076
56	AGAL330.954−00.182	29.0	22.7	25.9	1.168	0.254	21.3	20.9	21.1	0.954	0.198
57	AGAL331.709+00.582	21.4	19.6	20.5	1.046	0.237	13.8	13.4	13.6	0.695	0.152
58	AGAL332.094−00.421	22.9	20.3	21.6	0.377	0.085	13.4	13.0	13.2	0.231	0.051
59	AGAL332.826−00.549	29.0	21.1	25.0	0.437	0.096	24.5	20.9	22.7	0.396	0.082
60	AGAL333.134−00.431	45.8	34.4	40.1	0.700	0.145	43.0	22.9	33.0	0.575	0.118
61	AGAL333.284−00.387	86.9	61.0	73.9	1.291	0.261	35.5	32.4	34.0	0.592	0.120
62	AGAL333.314+00.106	16.8	15.7	16.2	0.283	0.068	14.2	12.6	13.4	0.234	0.051
63	AGAL333.604−00.212	73.2	53.2	63.2	1.103	0.224	69.5	51.7	60.6	1.058	0.213
64	AGAL333.656+00.059	12.2	11.0	11.6	0.297	0.082	–	–	–	–	–
65	AGAL335.789+00.174	29.0	23.5	26.2	0.467	0.101	17.4	15.8	16.6	0.295	0.063
66	AGAL336.958−00.224	12.2	11.0	11.6	0.612	0.168	8.7	8.3	8.5	0.449	0.110
67	AGAL337.176−00.032	41.2	20.3	30.8	1.641	0.357	43.4	15.8	29.6	1.579	0.329
68	AGAL337.258−00.101	18.3	14.1	16.2	0.864	0.210	12.6	11.8	12.2	0.653	0.145
69	AGAL337.286+00.007	33.6	14.9	24.2	1.109	0.256	–	–	–	–	–
70	AGAL337.406−00.402	15.3	14.1	14.7	0.232	0.058	18.2	17.8	18.0	0.284	0.060
71	AGAL337.704−00.054	24.4	21.9	23.1	1.376	0.304	16.6	13.8	15.2	0.903	0.195
72	AGAL337.916−00.477	30.5	22.7	26.6	0.413	0.090	20.5	19.7	20.1	0.312	0.065
73	AGAL338.066+00.044	13.7	12.5	13.1	0.298	0.078	–	–	–	–	–
74	AGAL338.786+00.476	33.6	32.1	32.8	0.715	0.151	–	–	–	–	–
75	AGAL338.926+00.554	61.0	26.6	43.8	0.934	0.196	21.7	19.7	20.7	0.442	0.092
76	AGAL339.623−00.122	35.1	21.1	28.1	0.410	0.089	13.4	11.8	12.6	0.184	0.041
77	AGAL340.374−00.391	22.9	15.7	19.3	0.336	0.078	–	–	–	–	–
78	AGAL340.746−01.001	36.6	25.0	30.8	0.412	0.088	12.2	11.4	11.8	0.158	0.035
79	AGAL340.784−00.097	16.8	14.9	15.8	0.767	0.186	12.6	11.8	12.2	0.593	0.132
80	AGAL341.217−00.212	18.3	14.9	16.6	0.295	0.071	14.2	13.4	13.8	0.246	0.054
81	AGAL342.484+00.182	35.1	22.7	28.9	1.758	0.379	14.2	13.4	13.8	0.841	0.183
82	AGAL343.128−00.062	42.7	18.0	30.4	0.447	0.099	18.2	17.0	17.6	0.259	0.055
83	AGAL343.756−00.164	29.0	20.3	24.6	0.347	0.076	13.8	13.0	13.4	0.189	0.041
84	AGAL344.227−00.569	27.5	13.3	20.4	0.249	0.059	17.8	15.0	16.4	0.200	0.043
85	AGAL345.003−00.224	39.7	18.0	28.9	0.422	0.093	17.8	16.6	17.2	0.251	0.053
86	AGAL345.488+00.314	61.0	43.8	52.4	0.564	0.115	20.1	18.2	19.2	0.206	0.043
87	AGAL345.504+00.347	68.6	37.6	53.1	0.579	0.119	16.6	15.4	16.0	0.174	0.037
88	AGAL345.718+00.817	38.1	27.4	32.8	0.248	0.052	16.2	15.0	15.6	0.118	0.025
89	AGAL351.131+00.771	91.5	86.1	88.8	0.783	0.158	–	–	–	–	–
90	AGAL351.161+00.697	42.7	34.4	38.6	0.340	0.071	28.4	23.3	25.8	0.228	0.047

Table A.5: continued.

ID	CSC Name	CO (6–5)					PACS 70 μm				
		$\Delta\theta_{max}$ ($''$)	$\Delta\theta_{min}$ ($''$)	$\Delta\theta_{avg}$ ($''$)	Δs_{avg} (pc)	σ_s (pc)	$\Delta\theta_{max}$ ($''$)	$\Delta\theta_{min}$ ($''$)	$\Delta\theta_{avg}$ ($''$)	Δs_{avg} (pc)	σ_s (pc)
91	AGAL351.244+00.669	53.4	35.2	44.3	0.391	0.081	53.3	48.6	50.9	0.449	0.090
92	AGAL351.416+00.646	48.8	25.0	36.9	0.240	0.051	27.6	26.1	26.9	0.174	0.036
93	AGAL351.444+00.659	73.2	47.7	60.4	0.393	0.080	21.3	17.0	19.1	0.124	0.026
94	AGAL351.571+00.762	53.4	27.4	40.4	0.262	0.055	–	–	–	–	–
95	AGAL351.581−00.352	13.7	13.3	13.5	0.446	0.114	15.4	14.6	15.0	0.495	0.107
96	AGAL351.774−00.537	33.6	15.7	24.6	0.119	0.027	25.3	22.9	24.1	0.117	0.024
97	AGAL353.066+00.452	83.9	23.5	53.7	0.224	0.047	–	–	–	–	–
98	AGAL353.417−00.079	–	–	–	–	–	–	–	–	–	–
99	AGAL354.944−00.537	22.9	14.9	18.9	0.175	0.041	–	–	–	–	–

Notes. The extension of the CO emission is measured from the 50% peak contour on the maps presented in Appendix B. The columns are as follows: (1) ID of the source; (2) the CSC name of the ATLASGAL clump; (3) the maximum elongation of the CO emission (in arcseconds); (4) the minimum elongation of the CO emission (in arcseconds); (5) the average size of the CO emission (in arcseconds); (6) the linear size of the CO emission (in parsecs), considering the average between the data presented in columns (3)-(4) and taking into account the distance of the source from Table A.1; (7) error of the linear extent of the CO emission, considering an uncertainty of 1.5$''$on the angular sizes presented in columns (3)-(4); (8)-(12) same as columns (3)-(7) but for the extension of the *Herschel*-PACS 70 μm emission towards the 70 μm-bright clumps.

Table A.6: **Integrated** properties of the CO emission profiles **convolved to 13$''$.4**.

ID	CSC Name	CO (4–3)					CO (6–5)					CO (7–6)				
		rms	FWZP	S_{int}	L_{CO}		rms	FWZP	S_{int}	L_{CO}		rms	FWZP	S_{int}	L_{CO}	
1	AGAL008.684−00.367	0.16	32	79.3	17.38	*	0.23	38	69.9	15.26	*	1.28	24	71.8	15.69	*
2	AGAL008.706−00.414	0.12	18	37.5	8.21		0.18	20	12.9	2.82		0.82	10	7.9	1.72	
3	AGAL010.444−00.017	0.22	28	43.7	30.65		0.16	26	24.4	17.05		0.37	24	16.9	11.85	
4	AGAL010.472+00.027	0.28	42	228.7	160.37	*	0.43	44	197.1	137.81	*	0.91	32	170.6	119.26	*
5	AGAL010.624−00.384	0.19	44	535.0	125.76	*	0.15	60	458.1	107.34	*	0.87	38	421.0	98.65	*
6	AGAL012.804−00.199	0.15	52	452.5	25.01	*	0.15	60	529.0	29.14	*	0.43	46	532.4	29.32	*
7	AGAL013.178+00.059	0.13	66	134.9	7.45		0.22	40	93.4	5.15		0.90	24	92.0	5.07	
8	AGAL013.658−00.599	0.15	94	151.8	29.11	*	0.16	98	109.5	20.92	*	0.92	80	120.9	23.10	*
9	AGAL014.114−00.574	0.14	40	46.7	2.96	*	0.12	38	59.8	3.78	*	0.40	24	73.3	4.63	*
10	AGAL014.194−00.194	0.17	48	189.3	27.62		0.14	50	113.5	16.51		0.52	40	90.4	13.15	
11	AGAL014.492−00.139	0.19	44	64.9	9.32	*	0.18	28	29.4	4.21	*	1.09	10	17.3	2.47	*
12	AGAL014.632−00.577	0.17	44	110.2	3.54	*	0.19	40	96.7	3.10	*	0.40	26	94.0	3.01	*
13	AGAL015.029−00.669	0.17	36	564.1	21.22	*	0.21	38	580.8	21.77		1.14	26	588.9	22.08	
14	AGAL018.606−00.074	0.16	22	65.1	93.93	*	0.10	28	45.9	8.24	*	0.41	18	35.7	6.39	
15	AGAL018.734−00.226	0.19	48	99.2	148.27		0.17	50	81.8	121.87		0.86	28	60.6	90.32	
16	AGAL018.888−00.474	0.17	38	188.8	40.69		0.20	52	128.7	27.64		1.05	22	78.0	16.76	
17	AGAL019.882−00.534	0.19	88	389.1	50.00		0.50	94	267.3	34.24		0.64	84	229.7	29.42	
18	AGAL022.376+00.447	0.14	62	84.2	12.66	*	0.28	22	28.5	4.27	*	1.24	14	21.6	3.23	*
19	AGAL023.206−00.377	0.20	98	254.0	51.33	*	0.11	82	162.9	32.81	*	0.37	88	202.8	40.87	*
20	AGAL024.629+00.172	0.29	26	42.0	23.99	*	0.21	24	18.0	10.25	*	0.91	10	11.5	6.55	*
21	AGAL028.564−00.236	0.25	18	43.0	12.26	*	0.22	28	22.0	6.24	*	1.15	4	4.3	1.23	*
22	AGAL028.861+00.066	0.37	42	206.5	108.78	*	0.44	64	230.6	121.10	*	1.27	50	205.3	107.79	
23	AGAL030.818−00.056	0.27	48	207.0	47.69	*	0.22	108	185.2	42.53	*	0.90	66	209.8	48.17	*
24	AGAL030.848−00.081	0.14	24	99.2	22.85	*	0.21	32	61.0	14.00	*	0.99	22	45.6	10.48	*
25	AGAL030.893+00.139	0.20	22	65.9	15.17		0.21	28	31.1	7.14		1.45	6	7.9	1.80	
26	AGAL031.412+00.307	0.32	44	118.5	27.29	*	0.24	66	154.6	35.50	*	0.37	54	166.6	38.25	*
27	AGAL034.258+00.154	0.47	58	267.0	6.23	*	0.52	106	371.1	8.64	*	0.47	80	383.5	8.92	*
28	AGAL034.401+00.226	0.14	62	227.2	5.30	*	0.13	94	217.4	5.06	*	0.41	50	183.6	4.27	*
29	AGAL034.411+00.234	0.18	92	210.1	4.90	*	0.28	86	136.4	3.18	*	1.24	24	99.1	2.31	
30	AGAL034.821+00.351	0.18	104	156.9	3.66	*	0.19	50	101.5	2.36	*	1.34	22	77.5	1.80	
31	AGAL035.197−00.742	0.26	42	269.5	12.40	*	0.09	76	348.8	16.00	*	0.34	46	295.5	13.55	*
32	AGAL037.554+00.201	0.15	44	142.1	61.38	*	0.26	38	81.6	35.13	*	1.56	26	70.6	30.39	*
33	AGAL043.166+00.011	0.52	66	1128.2	1335.97	*	0.25	132	904.5	1067.60		1.59	84	771.2	910.26	
34	AGAL049.489−00.389	0.63	58	542.2	152.25	*	0.39	82	264.6	74.06	*	0.91	70	327.2	91.57	*
35	AGAL053.141+00.069	0.15	76	211.9	5.20	*	0.14	72	179.2	4.39	*	0.35	38	159.8	3.91	*
36	AGAL059.782+00.066	0.24	56	247.0	11.05	*	0.12	66	204.4	9.12	*	0.43	40	170.3	7.60	*
37	AGAL301.136−00.226	–	–	–	–		0.88	130	705.3	130.58	*	1.62	126	780.8	144.54	
38	AGAL305.192−00.006	1.57	12	50.7	7.0		0.66	30	115.4	15.94		1.06	26	95.4	13.18	

Table A.6: continued.

		CO (4–3)					CO (6–5)					CO (7–6)				
ID	CSC Name	rms	FWZP	L_{CO}	σ_L		rms	FWZP	L_{CO}	σ_L		rms	FWZP	L_{CO}	σ_L	
39	AGAL305.209+00.206	1.60	36	433.4	60.0	*	0.22	96	403.8	55.75	*	0.83	70	396.5	54.75	*
40	AGAL305.562+00.014	0.40	32	376.9	52.2	*	0.23	38	271.7	37.51	*	1.65	26	248.5	34.31	*
41	AGAL305.794−00.096	0.34	22	13.8	1.9	*	0.22	24	21.0	2.90	*	1.62	16	34.1	4.71	*
42	AGAL309.384−00.134	0.53	26	133.6	36.6	*	0.21	44	95.9	26.14	*	1.58	26	99.8	27.20	*
43	AGAL310.014+00.387	0.45	48	182.2	22.8	*	0.21	74	120.3	14.99	*	1.27	30	104.9	13.07	
44	AGAL313.576+00.324	0.35	42	174.5	23.9	*	0.21	60	131.0	17.90	*	1.18	28	104.3	14.25	*
45	AGAL316.641−00.087	0.35	54	122.7	1.7	*	0.22	40	64.1	0.87		1.42	30	55.2	0.75	
46	AGAL317.867−00.151	0.35	22	110.7	9.2	*	0.23	42	77.8	6.48	*	1.49	20	73.9	6.15	*
47	AGAL318.779−00.137	0.33	36	70.1	5.2	*	0.21	30	39.3	2.91	*	1.18	16	43.9	3.25	
48	AGAL320.881−00.397	0.43	28	90.6	86.4		0.16	26	36.4	34.59		0.91	10	20.5	19.53	
49	AGAL326.661+00.519	0.36	100	334.5	10.6		0.17	24	169.2	5.36		0.95	18	151.7	4.80	
50	AGAL326.987−00.032	0.28	28	62.0	9.2758	*	0.18	54	63.2	9.42	*	0.79	20	45.9	6.85	*
51	AGAL327.119+00.509	0.24	42	90.4	26.34	*	0.24	38	69.3	20.13	*	1.00	22	49.1	14.26	*
52	AGAL327.393+00.199	0.32	54	140.7	47.32		0.14	54	84.0	28.16		0.77	22	67.5	22.61	
53	AGAL329.029−00.206	0.25	42	179.2	226.91	*	0.23	94	173.0	218.34	*	0.55	38	145.3	183.49	
54	AGAL329.066−00.307	0.24	42	111.9	143.95	*	0.17	46	86.2	110.58	*	0.77	24	65.7	84.22	
55	AGAL330.879−00.367	0.38	76	716.6	118.97	*	0.33	110	570.3	94.37	*	0.66	92	583.6	96.58	*
56	AGAL330.954−00.182	0.57	54	519.4	432.79	*	0.22	102	488.0	405.33	*	0.43	70	463.0	384.55	*
57	AGAL331.709+00.582	0.44	48	275.5	293.09	*	0.16	76	163.3	173.12	*	0.85	40	125.3	132.84	*
58	AGAL332.094−00.421	0.38	56	257.4	32.01	*	0.18	68	183.5	22.74	*	0.76	44	185.0	22.93	*
59	AGAL332.826−00.549	0.46	48	416.1	51.74	*	0.26	94	400.4	49.62	*	0.55	76	439.0	54.40	*
60	AGAL333.134−00.431	0.47	52	711.5	88.46	*	0.15	68	686.7	85.11	*	0.79	50	738.4	91.51	*
61	AGAL333.284−00.387	0.29	28	250.9	31.20	*	0.19	76	208.6	25.85	*	0.76	30	193.8	24.02	*
62	AGAL333.314+00.106	0.37	82	318.1	39.55	*	0.18	88	156.5	19.40	*	0.83	32	112.5	13.94	*
63	AGAL333.604−00.212	0.37	52	892.8	111.00		0.18	68	1018.6	126.23		0.79	50	1109.5	137.50	
64	AGAL333.656+00.059	0.29	26	80.5	21.60		0.15	28	37.2	9.96		0.87	30	31.7	8.49	
65	AGAL335.789+00.174	0.28	54	350.0	45.23	*	0.15	94	206.5	26.60	*	0.82	46	189.9	24.46	*
66	AGAL336.958−00.224	0.35	40	86.2	98.45	*	0.23	60	50.6	57.63	*	0.86	22	29.0	32.95	
67	AGAL337.176−00.032	0.37	34	97.8	113.55	*	0.18	30	62.4	72.18	*	0.70	32	61.3	70.92	*
68	AGAL337.258−00.101	0.51	20	39.7	8.08	*	0.20	30	46.9	54.26	*	0.93	26	44.8	51.84	*
69	AGAL337.286+00.007	0.38	32	40.8	34.84	*	0.24	24	16.3	13.85	*	0.97	12	9.8	8.35	*
70	AGAL337.406−00.402	0.58	56	243.4	24.82	*	0.58	116	309.7	31.47	*	1.31	66	256.3	26.05	*
71	AGAL337.704−00.054	0.43	32	101.6	146.50	*	0.27	60	167.5	240.73	*	0.97	36	168.7	242.41	*
72	AGAL337.916−00.477	0.97	140	897.6	88.18	*	0.42	110	638.2	62.49		0.70	88	607.4	59.48	
73	AGAL338.066+00.044	0.44	58	124.0	26.18	*	0.22	40	35.4	7.44	*	1.15	12	15.3	3.23	*
74	AGAL338.786+00.476	0.33	18	50.6	9.78	*	0.16	20	19.0	3.66	*	0.84	14	17.8	3.43	*
75	AGAL338.926+00.554	0.49	58	245.0	45.50	*	0.20	86	244.5	45.26	*	1.00	50	242.8	44.95	*
76	AGAL339.623−00.122	0.45	62	242.4	21.07	*	0.25	78	197.3	17.10	*	0.89	42	147.9	12.81	*
77	AGAL340.374−00.391	0.35	48	63.0	7.84	*	0.20	56	45.5	5.64	*	0.94	22	40.2	4.98	*
78	AGAL340.746−01.001	0.29	32	164.1	11.99	*	0.16	68	97.7	7.12	*	1.22	22	71.2	5.18	*
79	AGAL340.784−00.097	0.35	40	114.6	109.73	*	0.17	42	47.4	45.25	*	0.74	24	38.1	36.35	*
80	AGAL341.217−00.212	0.36	68	176.0	22.75	*	0.17	86	160.8	20.71	*	0.91	58	159.2	20.50	*
81	AGAL342.484+00.182	0.43	34	136.9	206.90	*	0.14	54	81.1	122.10	*	0.70	30	64.8	97.53	*
82	AGAL343.128−00.062	0.37	82	680.6	60.34	*	0.24	104	572.0	50.55	*	1.20	76	552.5	48.82	*
83	AGAL343.756−00.164	0.51	34	99.6	8.04	*	0.19	42	77.3	6.22	*	1.10	20	91.4	7.35	*
84	AGAL344.227−00.569	0.52	104	321.7	19.60	*	0.27	96	164.0	9.96	*	0.91	40	137.9	8.37	*
85	AGAL345.003−00.224	0.55	58	236.5	20.70	*	0.94	108	381.1	33.24	*	1.27	96	454.0	39.59	*
86	AGAL345.488+00.314	0.34	28	278.9	13.19	*	0.24	56	232.6	10.96	*	1.13	34	232.2	10.94	*
87	AGAL345.504+00.347	0.42	52	388.7	18.88	*	0.31	72	297.8	14.42	*	1.03	40	251.7	12.19	*
88	AGAL345.718+00.817	0.37	24	78.0	1.82	*	0.20	24	50.6	1.18	*	0.96	12	34.9	0.81	*
89	AGAL351.131+00.771	0.31	12	71.4	2.27	*	0.17	16	37.6	1.19	*	0.91	12	27.5	0.87	*
90	AGAL351.161+00.697	0.38	70	567.6	18.04	*	0.17	84	461.7	14.63	*	0.62	64	468.4	14.84	*
91	AGAL351.244+00.669	0.34	50	544.6	17.30	*	0.22	146	604.2	19.14	*	0.81	66	641.9	20.33	*
92	AGAL351.416+00.646	1.02	72	785.2	13.53	*	0.99	130	699.7	12.01	*	0.93	142	706.1	12.12	*
93	AGAL351.444+00.659	0.55	56	418.2	7.20	*	0.20	94	371.2	6.37	*	1.25	66	408.6	7.02	*
94	AGAL351.571+00.762	0.41	18	52.9	0.91	*	0.30	14	16.1	0.28	*	2.85	12	25.2	0.43	*
95	AGAL351.581−00.352	0.55	60	144.4	64.24	*	0.35	84	228.9	101.52	*	1.07	54	206.4	91.53	*
96	AGAL351.774−00.537	0.85	100	740.7	7.11		0.88	162	816.2	7.81		1.37	96	778.0	7.44	
97	AGAL353.066+00.452	0.58	10	29.9	0.21	*	0.19	26	63.9	0.45	*	1.39	10	28.0	0.20	*
98	AGAL353.417−00.079	0.63	24	42.1	14.83	*	0.22	18	11.6	4.08	*	1.84	10	16.8	5.91	*

Table A.6: continued.

		CO (4–3)				CO (6–5)				CO (7–6)			
ID	CSC Name	rms	FWZP	L_{CO}	σ_L	rms	FWZP	L_{CO}	σ_L	rms	FWZP	L_{CO}	σ_L
99	AGAL354.944−00.537	0.62	12	36.8	1.29	0.23	22	19.0	0.66	1.42	8	10.7	0.37

Notes. The columns are as follows: (1)-(2) ID and CSC name of the source; (3) rms of the CO (4–3) data (in K), (4) the full width at zero power (FWZP, in km s^{-1}); (5) integrated intensity of the line ($S_{\rm int}$, in K km s^{-1}); (6) the line luminosity (L_{CO}, in K km s^{-1} pc^2) of the CO (4–3) spectra; (7) an asterisk mark indicates if the spectrum is contaminated by a self-absorption feature; the same properties computed for the CO (6–5) and CO (7–6) data are presented in columns (8)-(12) and (13)-(17), respectively.

Table A.7: Excitation temperature derived from the CO (6–5) spectra convolved to 13.″4.

ID	CSC Name	T_{ex}	σ_T	$T_{ex,gauss}$	σ_{T+}	σ_{T-}	
1	AGAL008.684−00.367	28.0	0.3	40.9	22.8	25.2	
2	AGAL008.706−00.414	14.2	0.3	13.9	0.3	0.3	
3	AGAL010.444−00.017	17.5	0.2	39.9	21.2	13.9	*
4	AGAL010.472+00.027	52.3	0.4	97.0	0.6	0.6	
5	AGAL010.624−00.384	84.9	0.2	131.5	27.9	28.0	
6	AGAL012.804−00.199	97.1	0.2	143.0	0.0	0.0	
7	AGAL013.178+00.059	31.0	0.2	32.1	7.4	7.7	
8	AGAL013.658−00.599	29.7	0.2	30.8	18.6	16.2	
9	AGAL014.114−00.574	24.4	0.1	39.3	18.3	19.8	
10	AGAL014.194−00.194	34.0	0.2	37.1	17.0	18.6	
11	AGAL014.492−00.139	20.5	0.2	22.8	6.1	6.8	
12	AGAL014.632−00.577	38.1	0.2	41.6	10.9	11.3	
13	AGAL015.029−00.669	146.9	0.2	67.2	41.2	25.5	*
14	AGAL018.606−00.074	26.1	0.1	33.8	6.1	6.3	
15	AGAL018.734−00.226	29.4	0.2	29.2	2.7	2.8	
16	AGAL018.888−00.474	36.8	0.2	36.8	6.6	6.8	
17	AGAL019.882−00.534	49.7	0.5	50.0	2.8	2.8	
18	AGAL022.376+00.447	20.4	0.4	29.0	13.1	15.2	
19	AGAL023.206−00.377	32.4	0.1	34.9	0.5	0.5	
20	AGAL024.629+00.172	15.8	0.3	15.8	7.8	11.3	
21	AGAL028.564−00.236	15.3	0.3	17.7	4.8	5.8	
22	AGAL028.861+00.066	46.1	0.5	45.4	10.7	11.0	
23	AGAL030.818−00.056	44.1	0.2	66.7	43.0	45.2	
24	AGAL030.848−00.081	25.8	0.2	30.4	14.9	10.0	*
25	AGAL030.893+00.139	18.8	0.3	18.5	6.8	9.1	
26	AGAL031.412+00.307	42.9	0.2	61.5	10.4	10.5	
27	AGAL034.258+00.154	87.4	0.5	54.1	31.3	19.8	*
28	AGAL034.401+00.226	52.5	0.1	59.5	47.8	23.8	
29	AGAL034.411+00.234	34.0	0.3	35.7	1.0	1.0	
30	AGAL034.821+00.351	29.3	0.2	32.4	2.2	2.2	
31	AGAL035.197−00.742	65.3	0.1	84.1	0.1	0.1	
32	AGAL037.554+00.201	28.3	0.3	32.8	3.4	3.5	
33	AGAL043.166+00.011	84.9	0.3	135.3	98.8	57.1	*
34	AGAL049.489−00.389	58.3	0.4	90.6	60.0	36.1	*
35	AGAL053.141+00.069	49.2	0.1	61.7	7.0	7.0	
36	AGAL059.782+00.066	51.7	0.1	57.4	0.4	0.4	
37	AGAL301.136−00.226	59.7	0.9	73.6	0.1	0.1	
38	AGAL305.192−00.006	33.5	0.7	40.6	23.9	26.9	
39	AGAL305.209+00.206	69.7	0.2	68.9	2.2	2.2	
40	AGAL305.562+00.014	74.5	0.2	74.8	2.9	3.0	
41	AGAL305.794−00.096	15.9	0.3	15.5	1.0	1.0	
42	AGAL309.384−00.134	32.1	0.2	35.1	5.8	5.9	
43	AGAL310.014+00.387	30.4	0.2	30.0	2.4	2.4	
44	AGAL313.576+00.324	35.5	0.2	37.7	3.7	3.8	
45	AGAL316.641−00.087	21.0	0.3	20.4	3.0	3.3	
46	AGAL317.867−00.151	28.2	0.3	29.2	0.4	0.4	
47	AGAL318.779−00.137	20.6	0.3	22.3	4.9	5.3	
48	AGAL320.881−00.397	24.0	0.2	24.5	4.5	4.8	
49	AGAL326.661+00.519	89.7	0.2	89.7	13.6	13.7	
50	AGAL326.987−00.032	23.8	0.2	27.0	14.4	19.1	

Table A.7: continued.

ID	CSC Name	$T_{\rm ex}$	σ_T	$T_{\rm ex,gauss}$	σ_{T+}	σ_{T-}	
51	AGAL327.119+00.509	28.2	0.3	33.8	6.1	6.3	
52	AGAL327.393+00.199	28.8	0.2	29.9	4.9	5.1	
53	AGAL329.029−00.206	36.0	0.3	40.3	7.8	7.9	
54	AGAL329.066−00.307	28.8	0.2	29.5	3.9	4.0	
55	AGAL330.879−00.367	64.9	0.3	75.0	0.5	0.5	
56	AGAL330.954−00.182	74.8	0.2	94.9	17.0	17.1	
57	AGAL331.709+00.582	37.3	0.2	51.3	29.3	18.6	∗
58	AGAL332.094−00.421	46.5	0.2	35.3	1.1	1.1	
59	AGAL332.826−00.549	85.2	0.3	76.0	48.1	29.5	∗
60	AGAL333.134−00.431	91.3	0.2	85.3	55.6	33.7	∗
61	AGAL333.284−00.387	67.4	0.2	72.8	0.7	0.7	
62	AGAL333.314+00.106	35.8	0.2	38.1	7.8	8.0	
63	AGAL333.604−00.212	140.5	0.2	136.3	1.2	1.2	
64	AGAL333.656+00.059	21.9	0.2	22.4	7.6	8.9	
65	AGAL335.789+00.174	40.3	0.2	49.3	27.8	29.8	
66	AGAL336.958−00.224	21.7	0.3	20.9	3.5	3.7	
67	AGAL337.176−00.032	30.0	0.2	29.9	5.2	5.4	
68	AGAL337.258−00.101	22.4	0.2	21.8	0.3	0.3	
69	AGAL337.286+00.007	15.2	0.3	14.2	0.9	1.0	
70	AGAL337.406−00.402	61.6	0.6	95.4	31.0	31.3	
71	AGAL337.704−00.054	41.7	0.3	80.5	51.7	31.5	∗
72	AGAL337.916−00.477	68.9	0.4	76.4	0.3	0.3	
73	AGAL338.066+00.044	17.4	0.3	18.6	2.1	2.3	
74	AGAL338.786+00.476	17.1	0.2	17.4	0.3	0.3	
75	AGAL338.926+00.554	51.4	0.2	62.3	37.5	23.4	∗
76	AGAL339.623−00.122	42.5	0.3	42.6	12.2	12.6	
77	AGAL340.374−00.391	17.1	0.3	18.3	4.7	5.6	
78	AGAL340.746−01.001	32.9	0.2	34.1	17.6	20.1	
79	AGAL340.784−00.097	21.4	0.2	23.5	2.5	2.6	
80	AGAL341.217−00.212	42.8	0.2	51.9	11.2	11.4	
81	AGAL342.484+00.182	27.4	0.2	27.1	1.1	1.1	
82	AGAL343.128−00.062	68.0	0.2	75.3	0.6	0.6	
83	AGAL343.756−00.164	31.5	0.2	34.1	14.1	15.3	
84	AGAL344.227−00.569	34.1	0.3	37.4	2.2	2.3	
85	AGAL345.003−00.224	63.4	1.0	86.9	30.5	30.8	
86	AGAL345.488+00.314	59.6	0.3	56.9	33.4	21.0	∗
87	AGAL345.504+00.347	63.4	0.3	67.0	145.2	141.7	
88	AGAL345.718+00.817	26.3	0.2	27.3	3.1	3.2	
89	AGAL351.131+00.771	26.5	0.2	30.6	0.2	0.2	
90	AGAL351.161+00.697	89.9	0.2	99.1	0.1	0.1	
91	AGAL351.244+00.669	112.8	0.2	126.5	16.3	16.3	
92	AGAL351.416+00.646	84.5	1.0	105.7	19.2	19.3	
93	AGAL351.444+00.659	62.2	0.2	81.2	45.1	46.1	
94	AGAL351.571+00.762	19.7	0.4	19.6	0.5	0.5	
95	AGAL351.581−00.352	40.4	0.4	49.0	4.3	4.3	
96	AGAL351.774−00.537	74.1	0.9	88.0	9.1	9.1	
97	AGAL353.066+00.452	28.7	0.2	13.2	4.9	3.6	∗
98	AGAL353.417−00.079	15.5	0.3	32.9	16.5	11.0	∗
99	AGAL354.944−00.537	16.6	0.3	20.6	8.9	6.2	∗

Notes. The columns are as follows: (1)-(2) ID and CSC name of the source (given in Table A.1); (3)-(4) excitation temperature (in K) and its uncertainty derived from the peak intensity of the CO (6–5) spectra; (5)-(7) excitation temperature (in K) and its upper and lower uncertainty derived from the peak intensity of the Gaussian fit of the CO (6–5) spectra; (8) an asterisks indicates the cases where the Gaussian fit is dubious and, therefore, the excitation temperature was obtained from the relation ($\log(T_{\rm ex}) = (0.75 \pm 0.10) + (0.21 \pm 0.02) \log(L_{\rm bol})$ (see Sect. 4.4 for further details).

Appendix B: CO spectra and CO (6–5) maps

In Fig. B.1 **we show the integrated CO (6–5) maps towards the** TOP100 **sample together with the CO (4–3) spectra from** FLASH$^+$ **observations, the convolved** CHAMP$^+$ **mid-J CO spectra using a fixed beam size of 13.″4 and the isotopologue C^{17}O (3–2) or C^{18}O (2–1) spectra from Giannetti et al. (2014).**

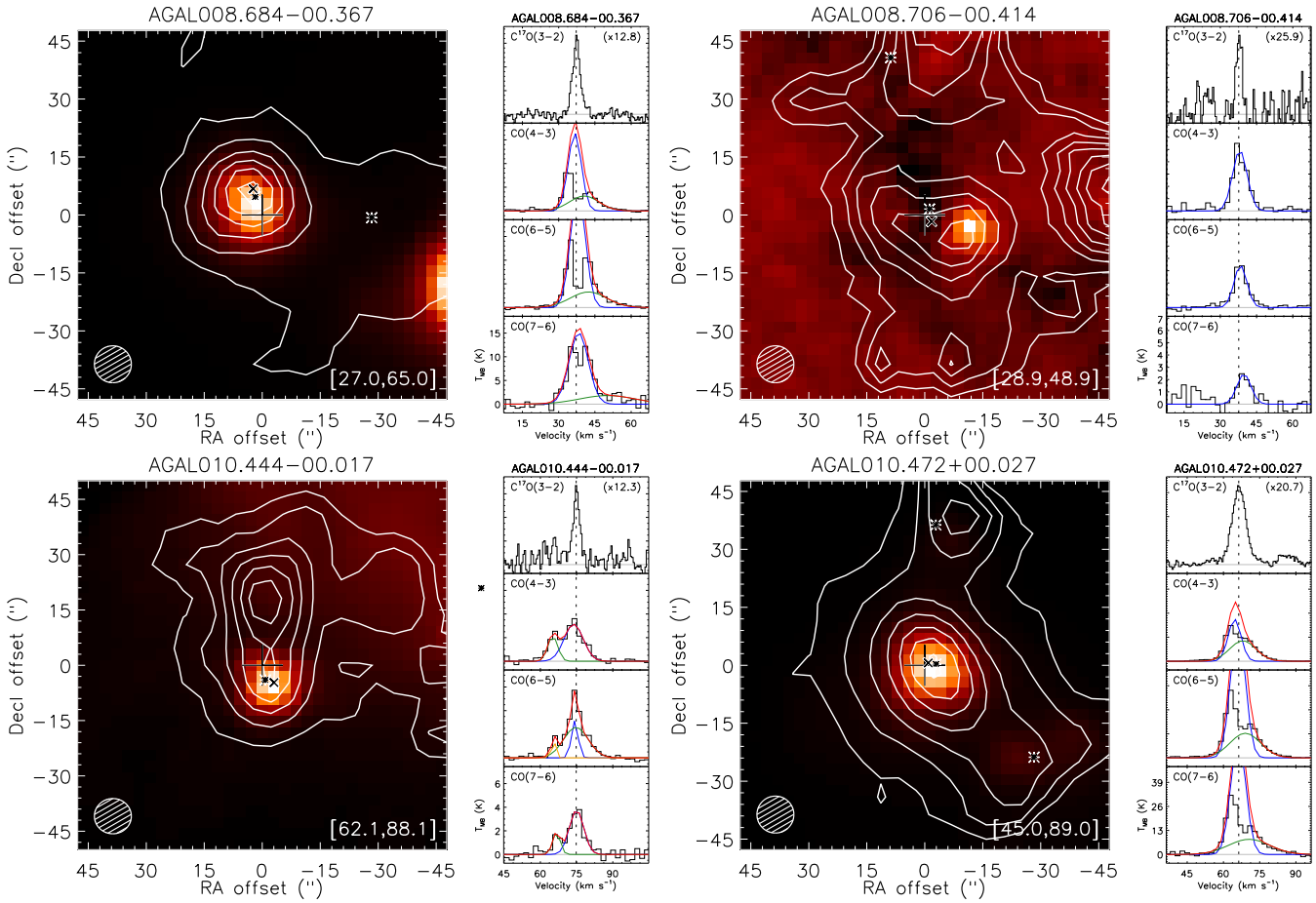


Fig. B.1: Left: False-colour *Herschel*/PACS images at 70 μm overlaid by the CO (6–5) emission contours towards the TOP100 source. The CO contours correspond to the emission integrated over the full-width at zero power (FWZP) of the CO (6–5) profile, the velocity range shown in the bottom right side of the image, and the contour levels are shown from 20% to 90% of the peak emission of each map, in steps of 10%. The (0,0) position of the map is shown as a + symbol, the position of the CSC source from Contreras et al. (2013) is shown as a × symbol and the dust continuum emission peaks from Csengeri et al. (2014) are shown as asterisks. **Right, from top to bottom panel: C^{17}O (2–1) or C^{18}O (1–0) from Giannetti et al. (2014), CO (4–3), CO (6–5) and CO (7–6) profiles** towards the TOP100 sample, convolved into a fixed beam size of 13.″4 (shown in black) and their fitted Gaussian components. The narrower Gaussian component fitted to the data is shown in blue, the second and broader component is shown in green and the third component is shown in yellow. The sum of all Gaussian components is shown in red, except for the cases where a single component was fitted. The vertical dashed black line is placed at the rest velocity ($V_{\rm lsr}$) of each source. The horizontal filled grey line displays the baseline of the data.

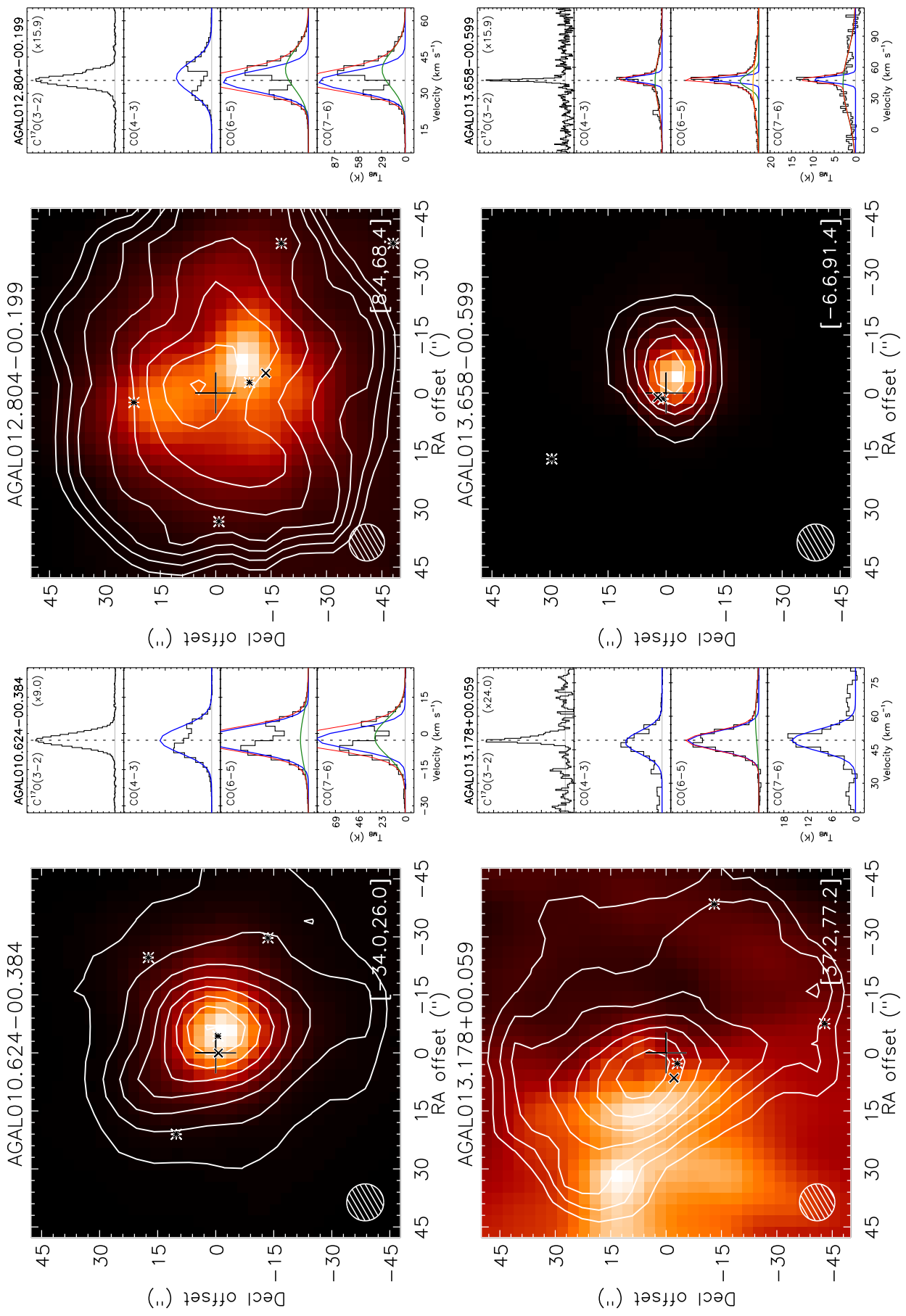


Fig. B.1: Continued.

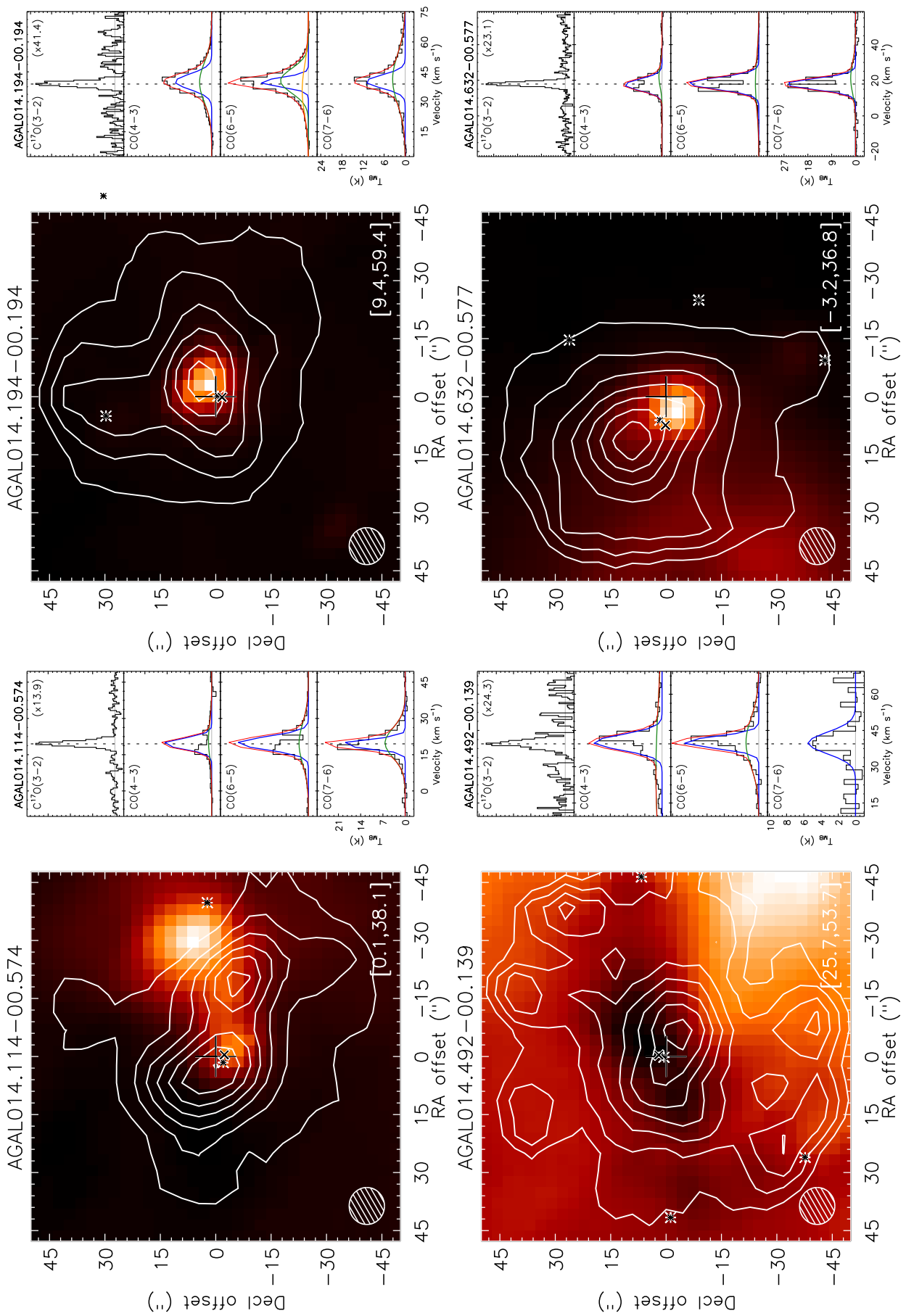


Fig. B.1: Continued.

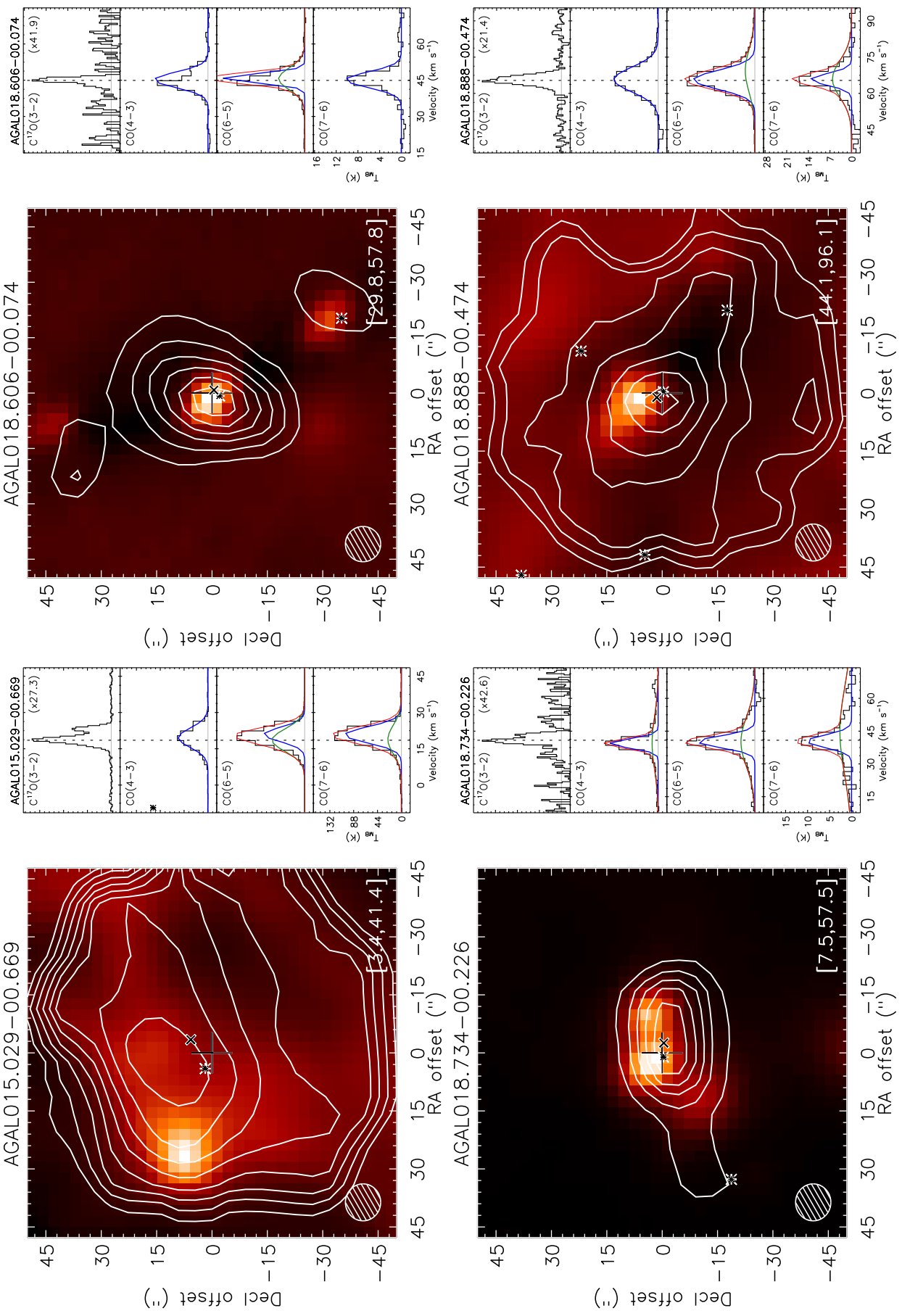


Fig. B.1: Continued.

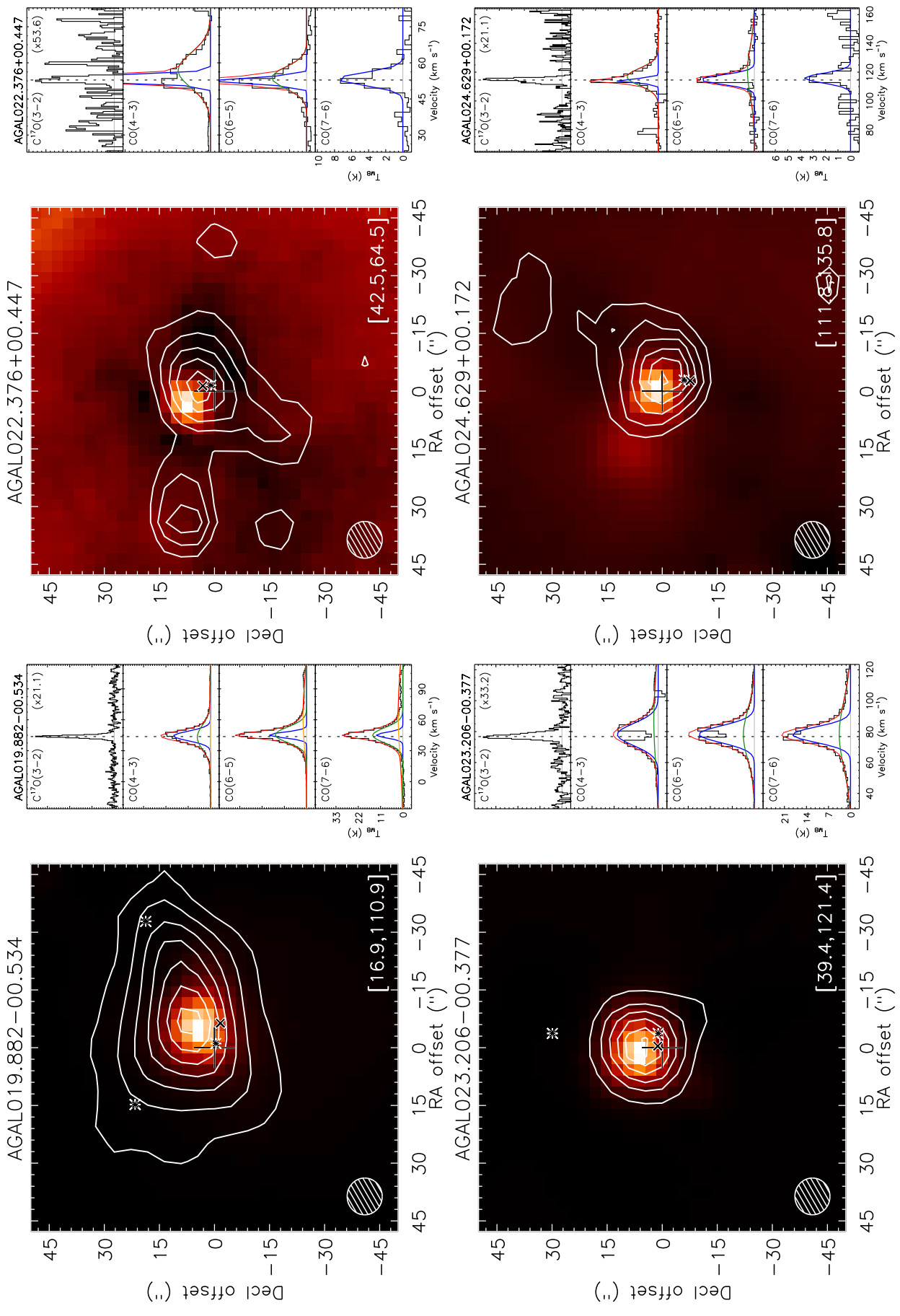


Fig. B.1: Continued.

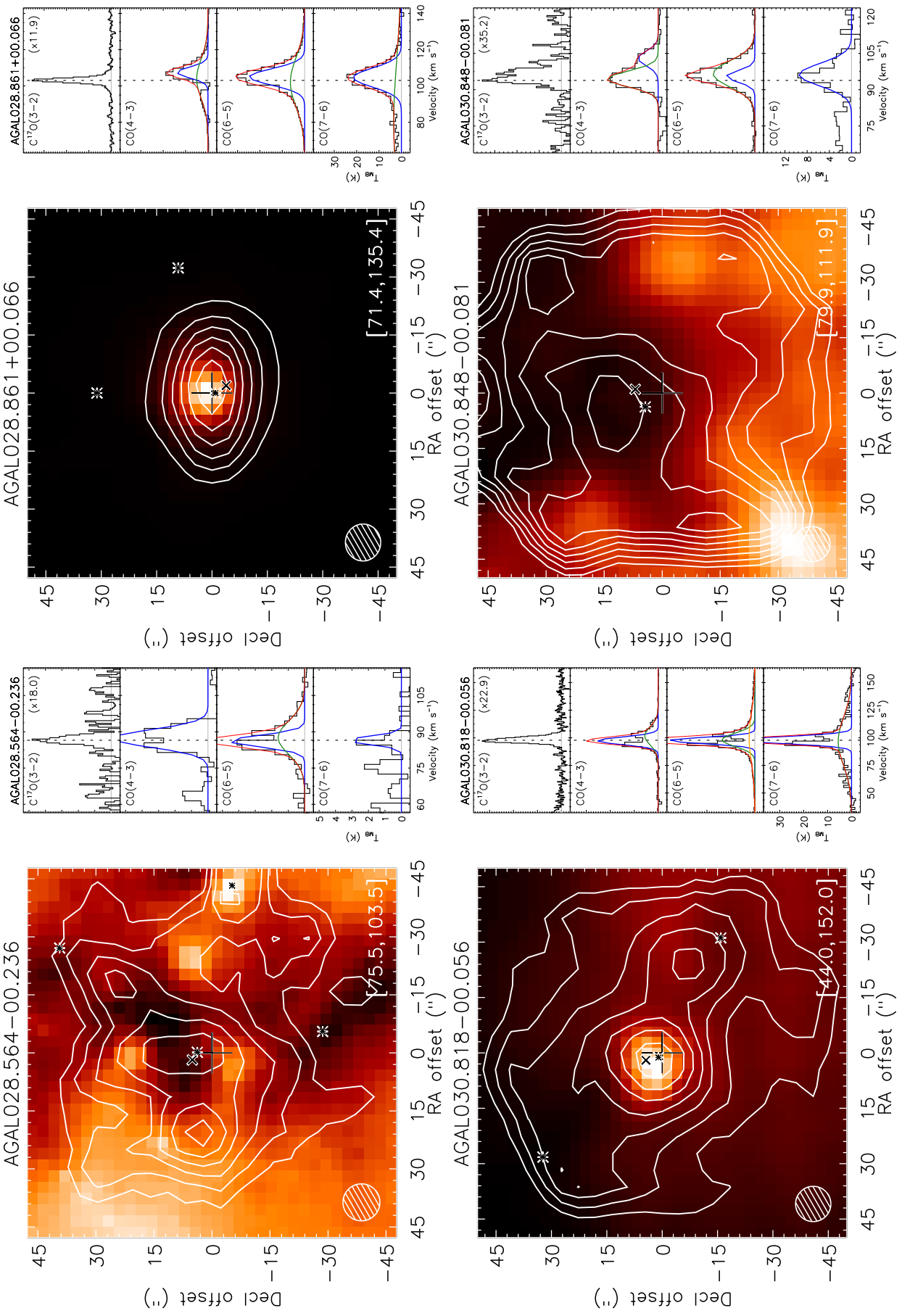


Fig. B.1: Continued.

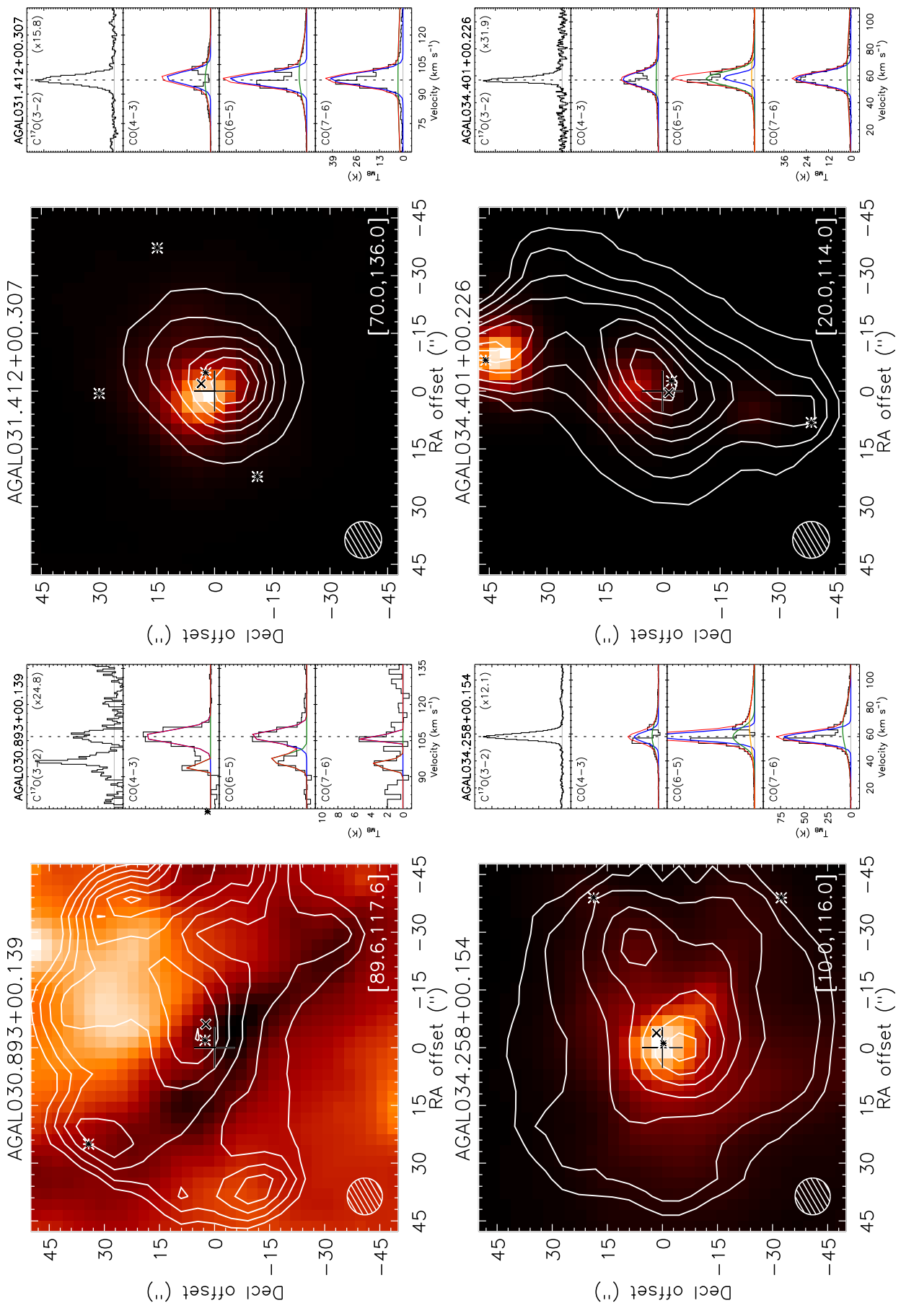


Fig. B.1: Continued.

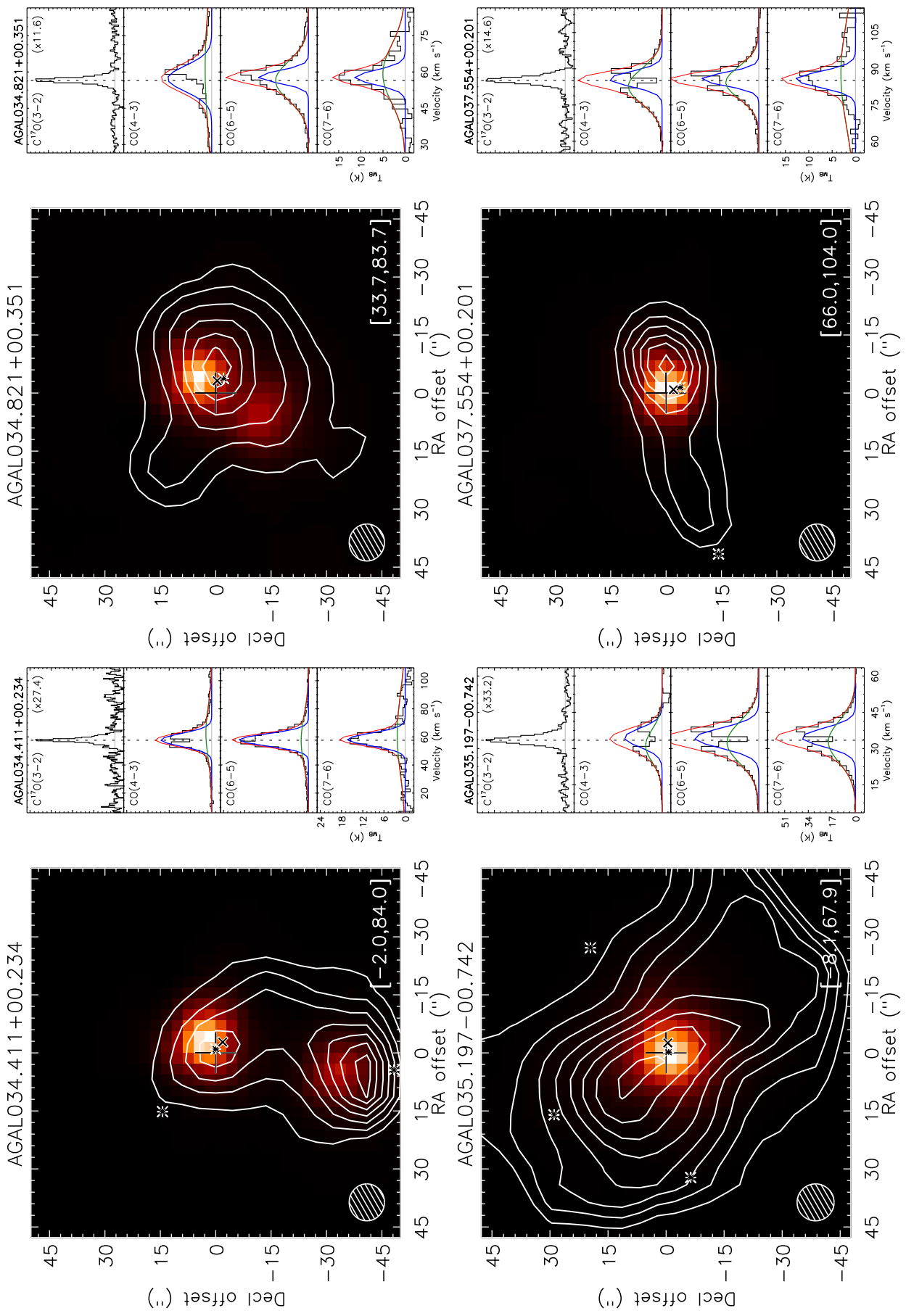


Fig. B.1: Continued.

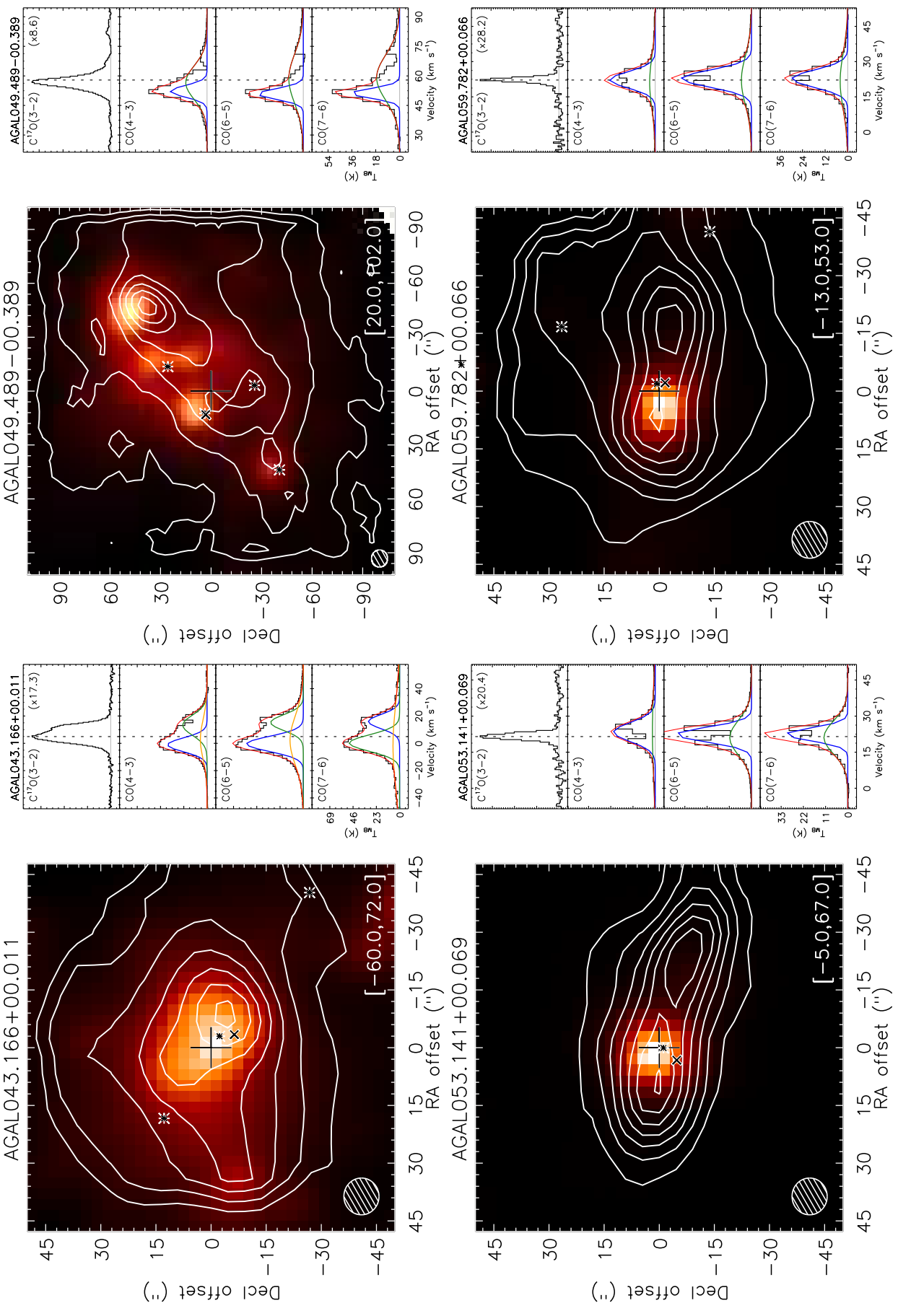


Fig. B.1: Continued.

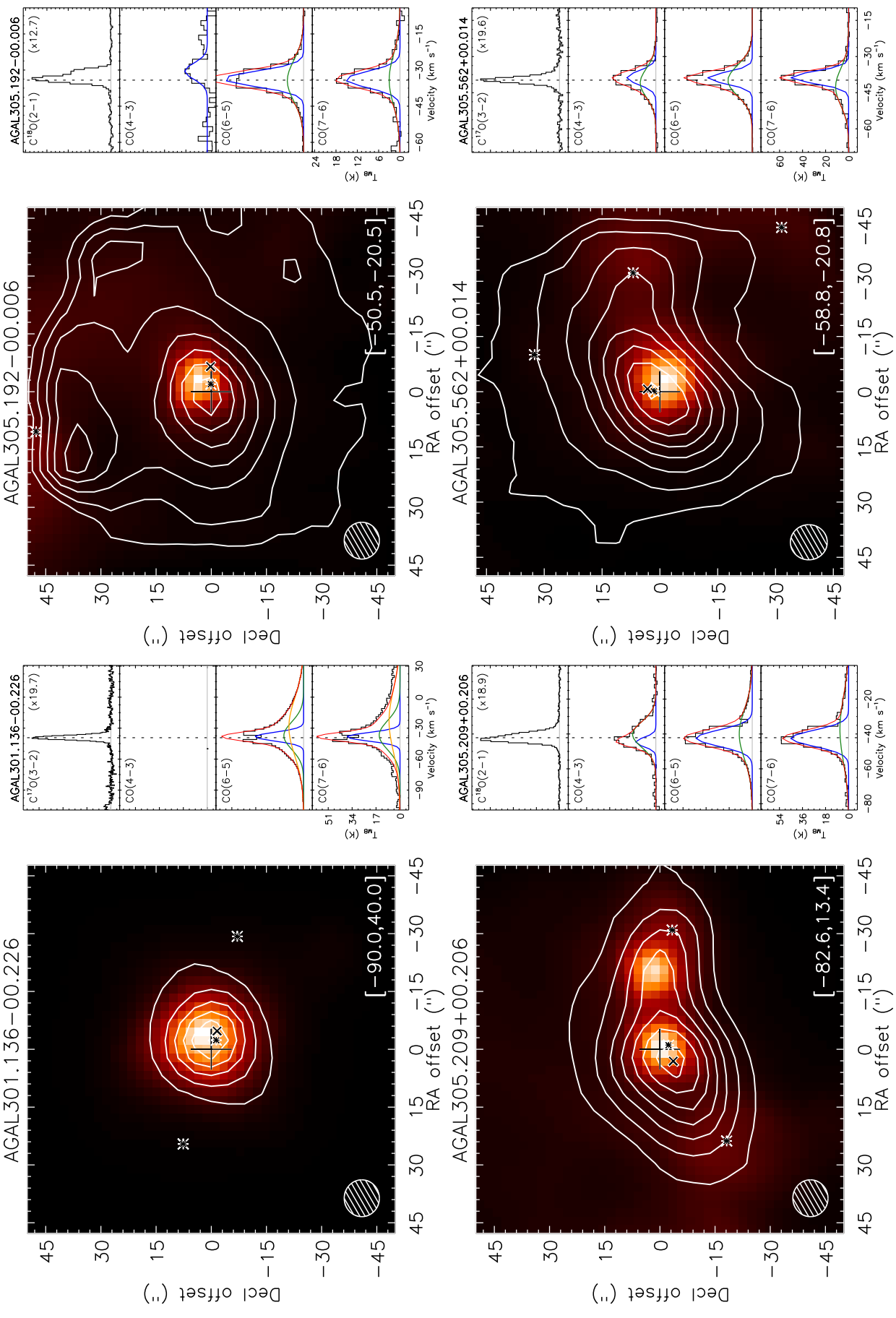


Fig. B.1: Continued.

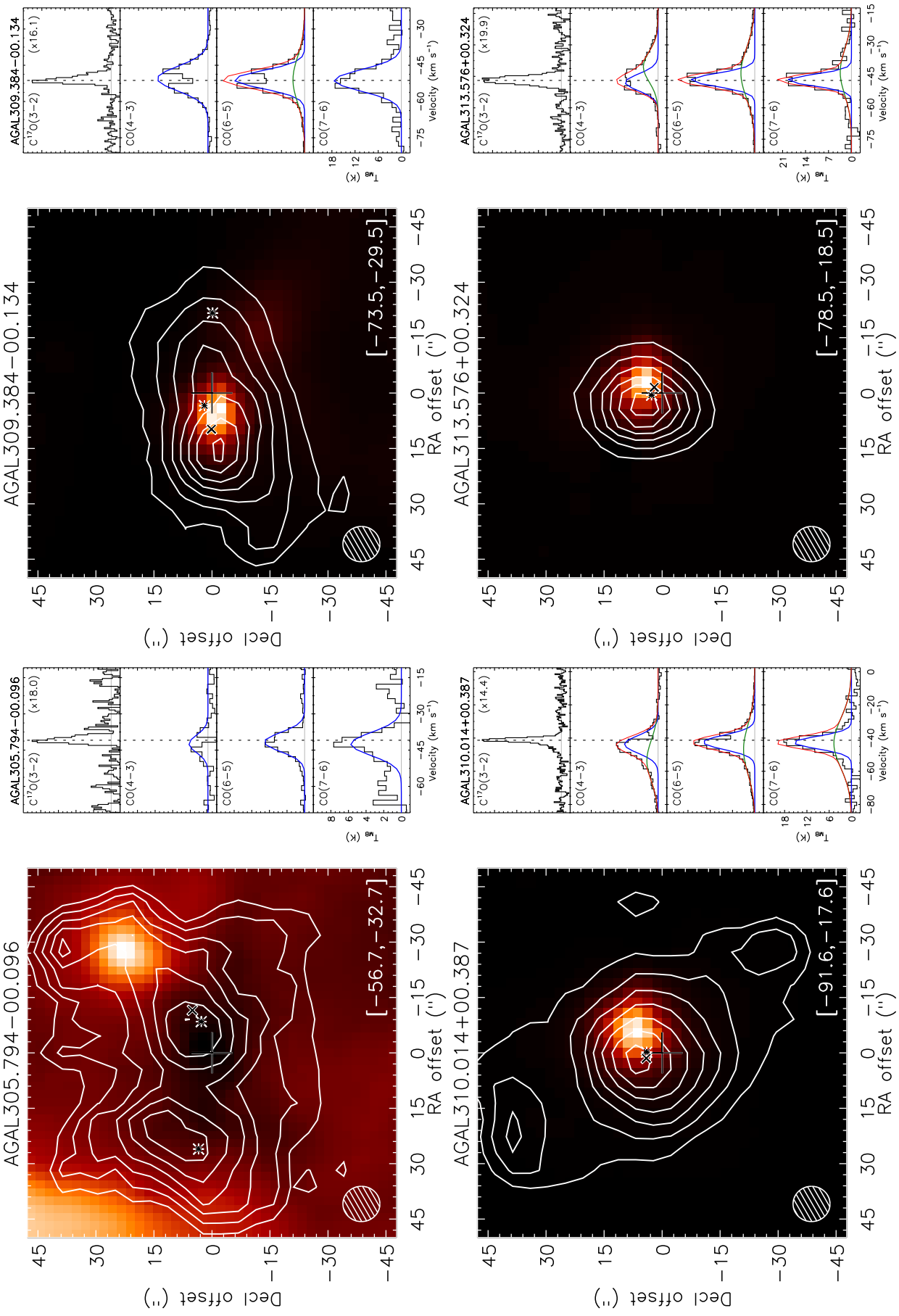


Fig. B.1: Continued.

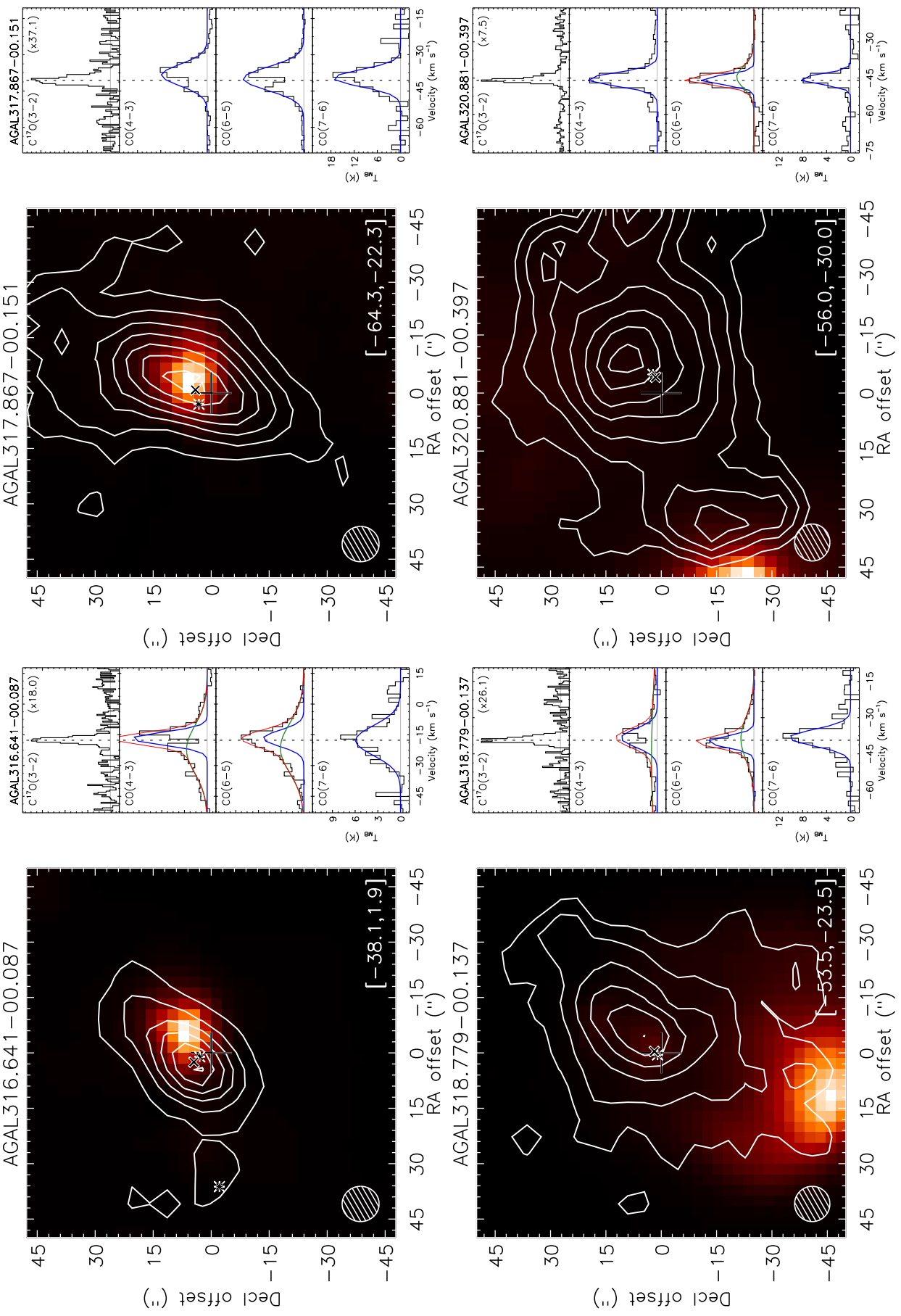


Fig. B.1: Continued.

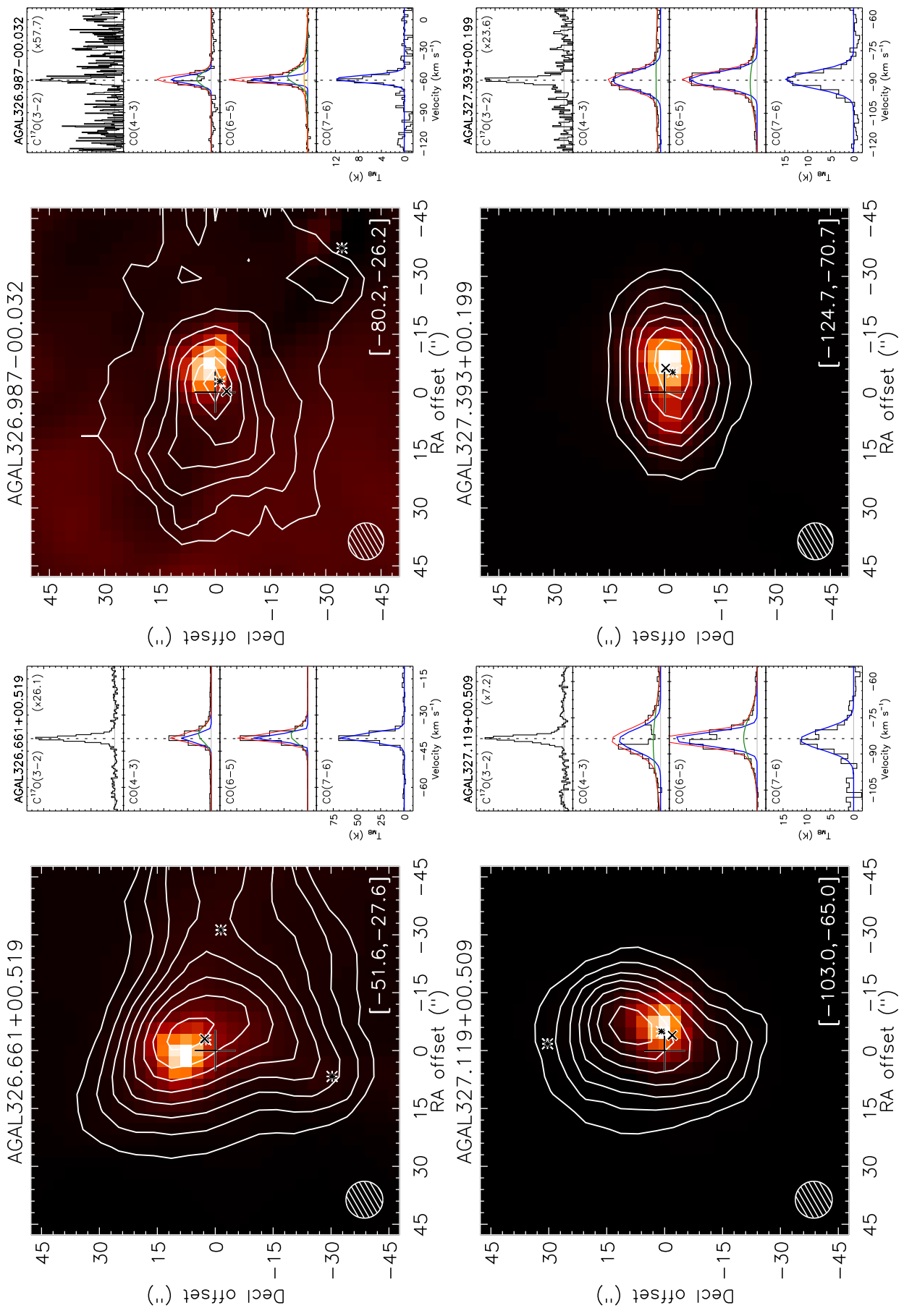


Fig. B.1: Continued.

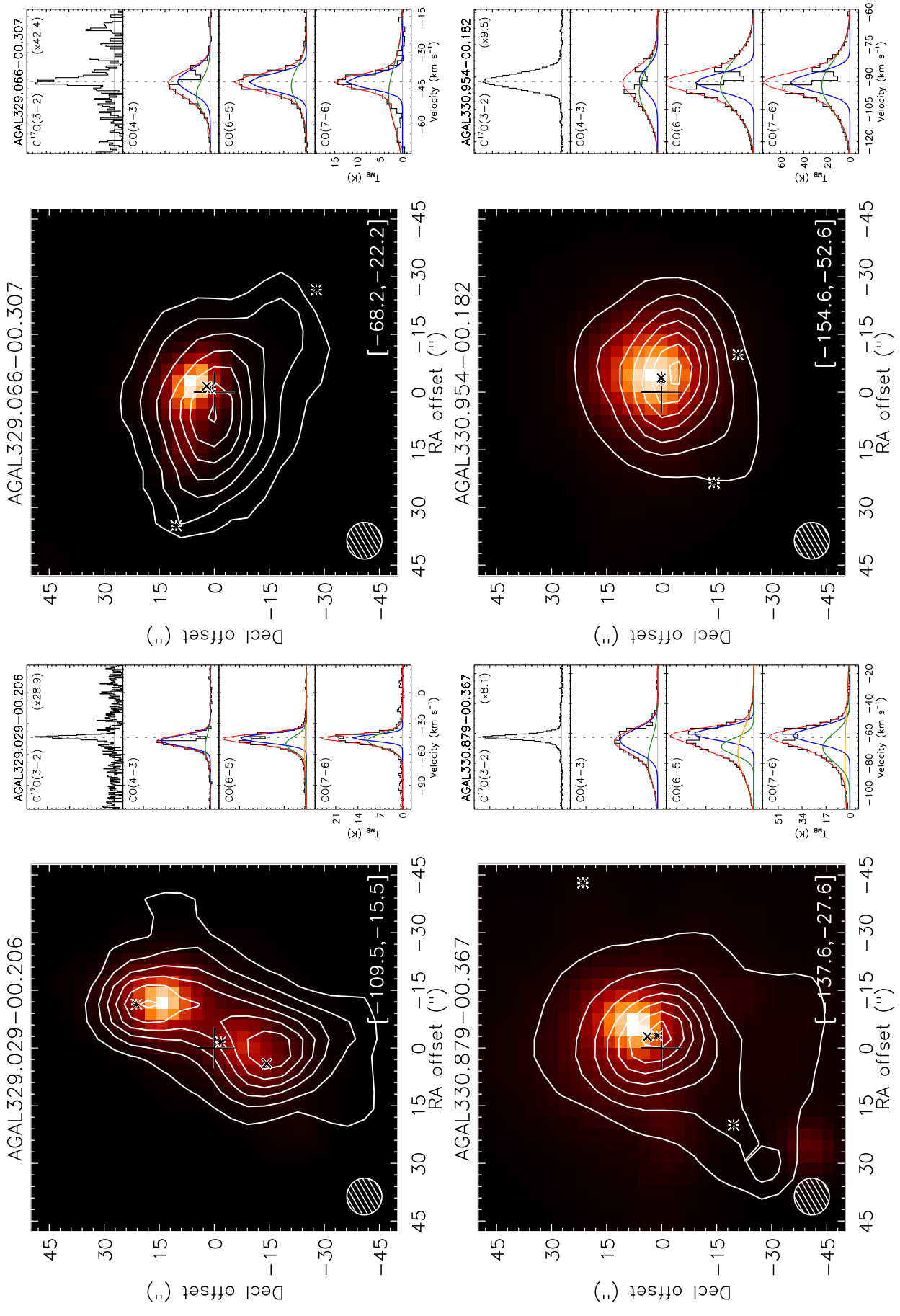


Fig. B.1: Continued.

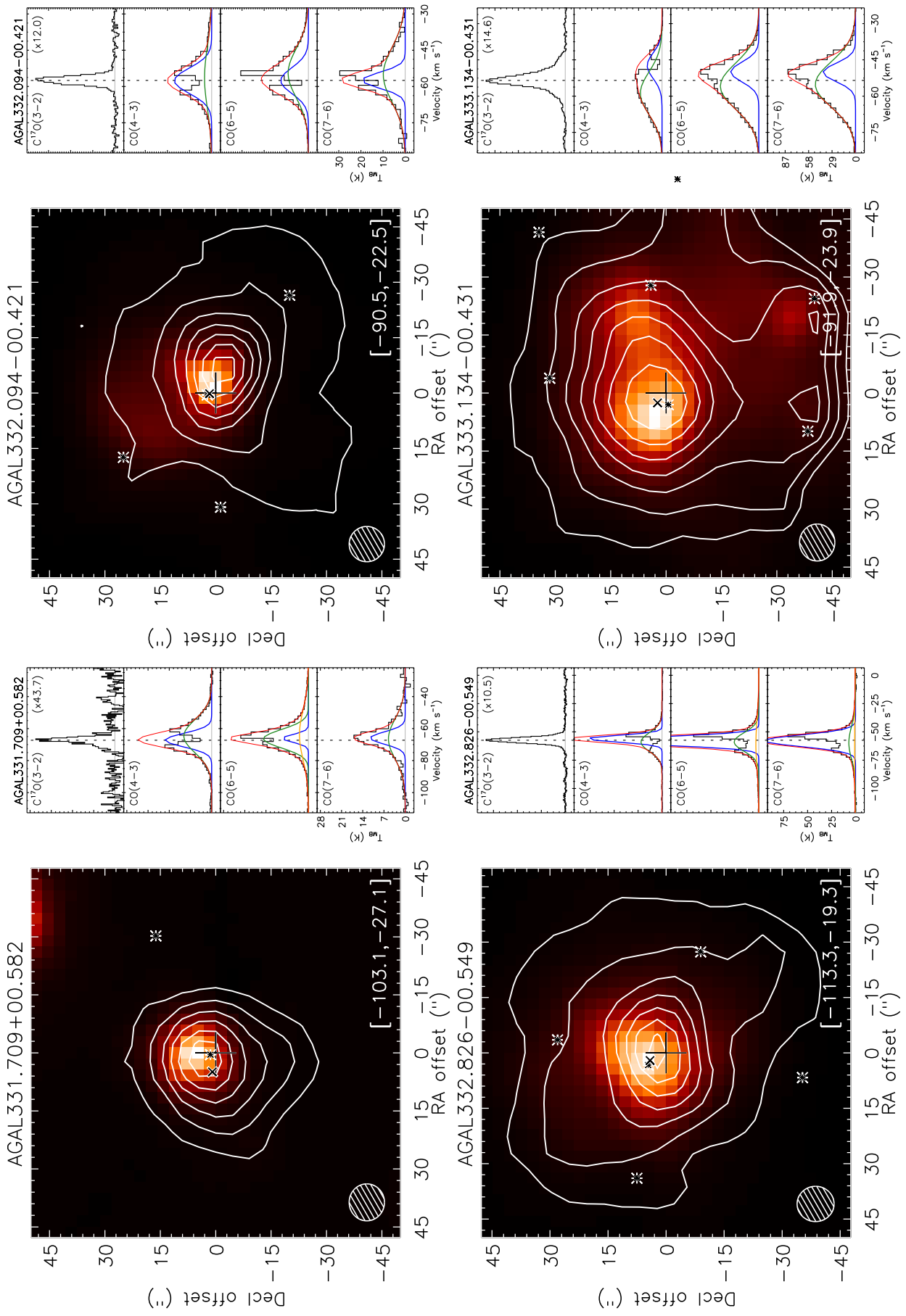


Fig. B.1: Continued.

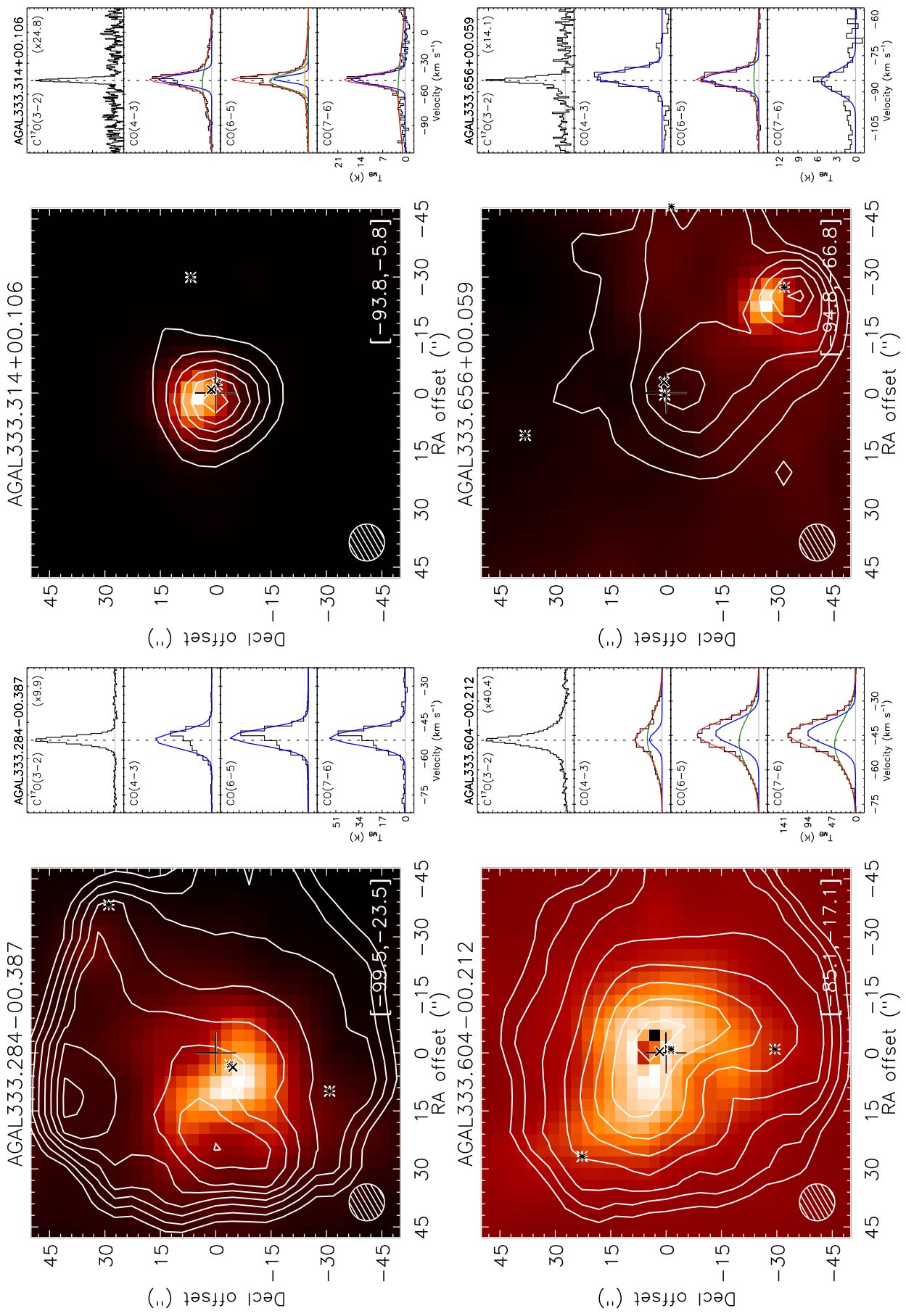


Fig. B.1: Continued.

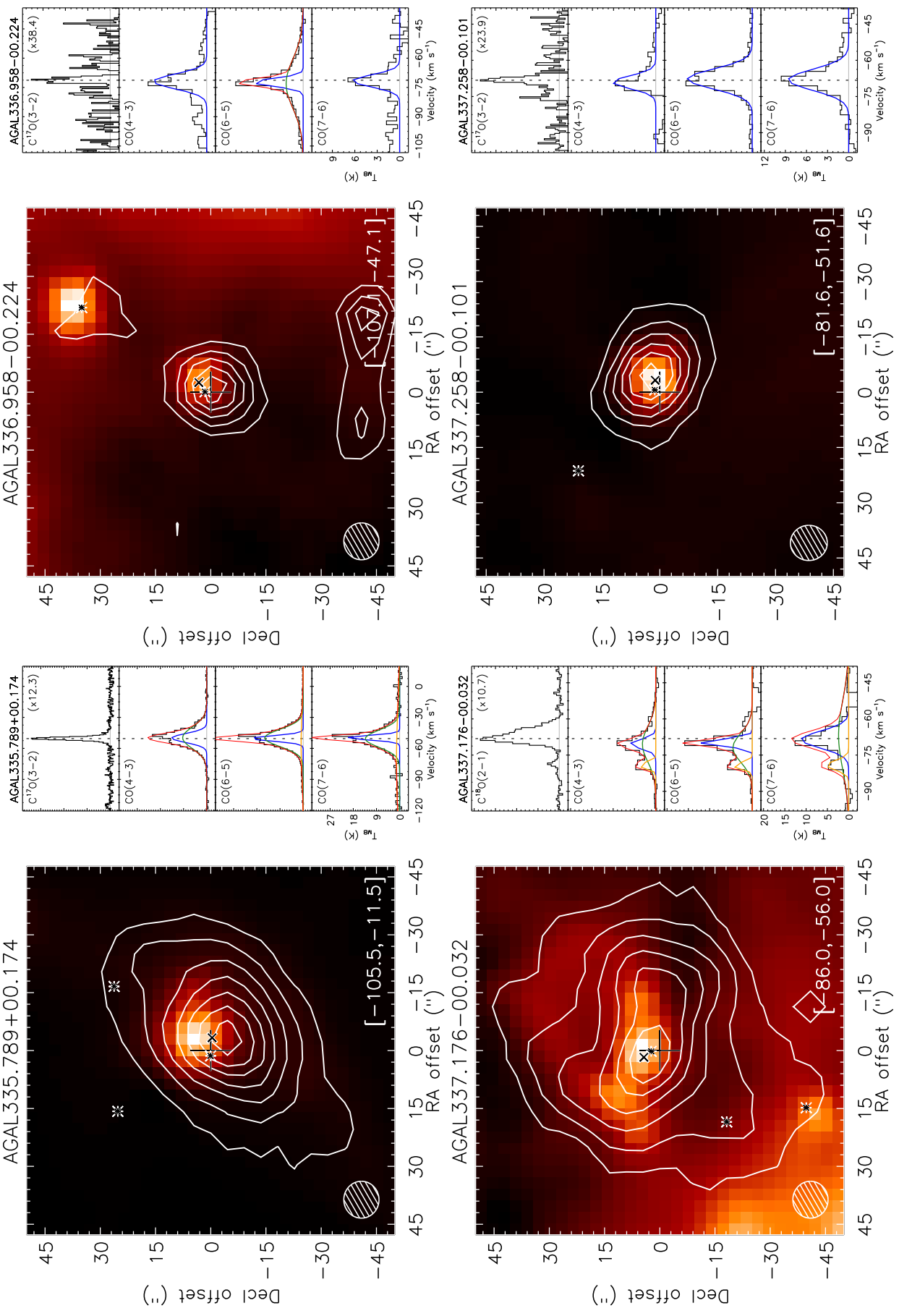


Fig. B.1: Continued.

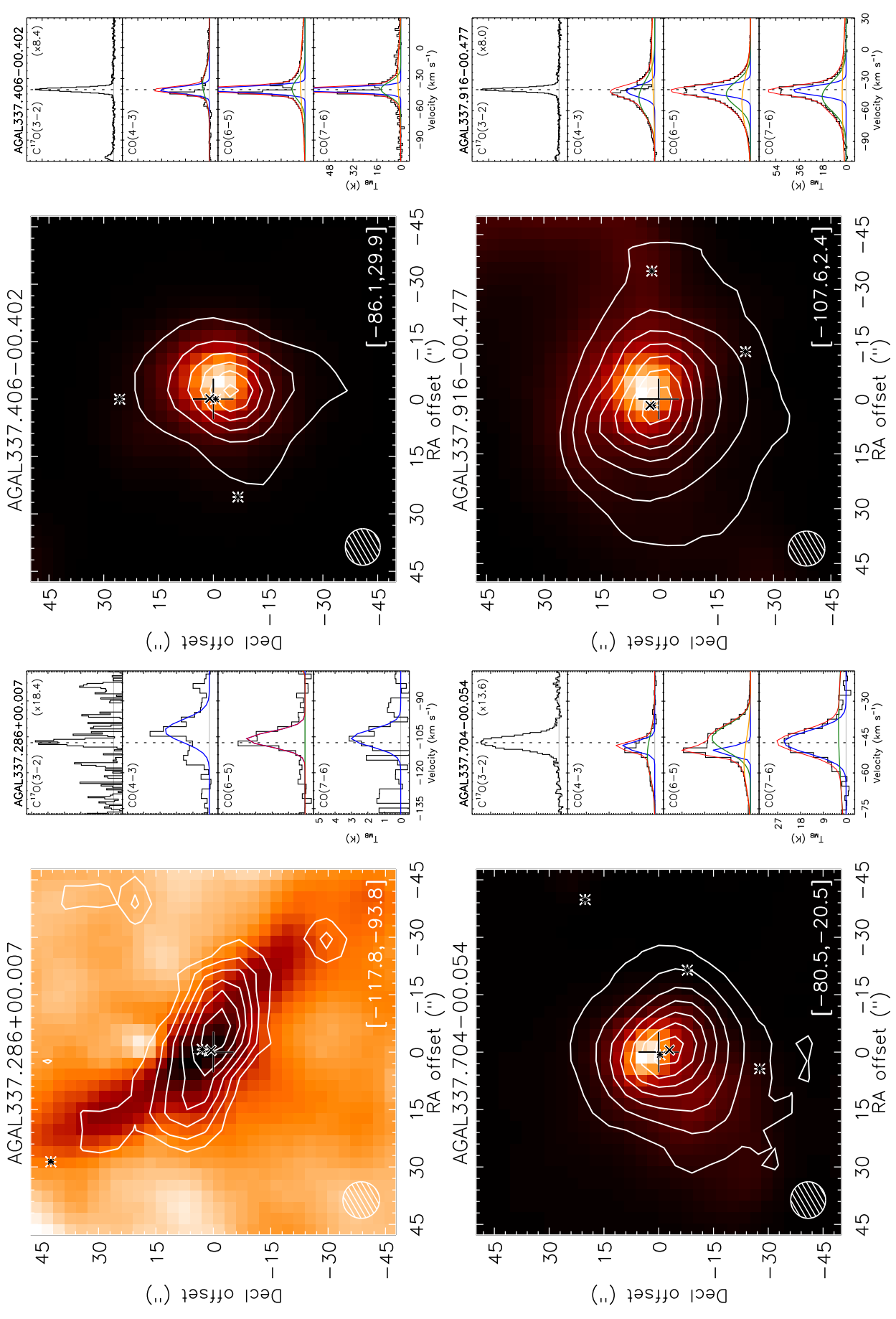


Fig. B.1: Continued.

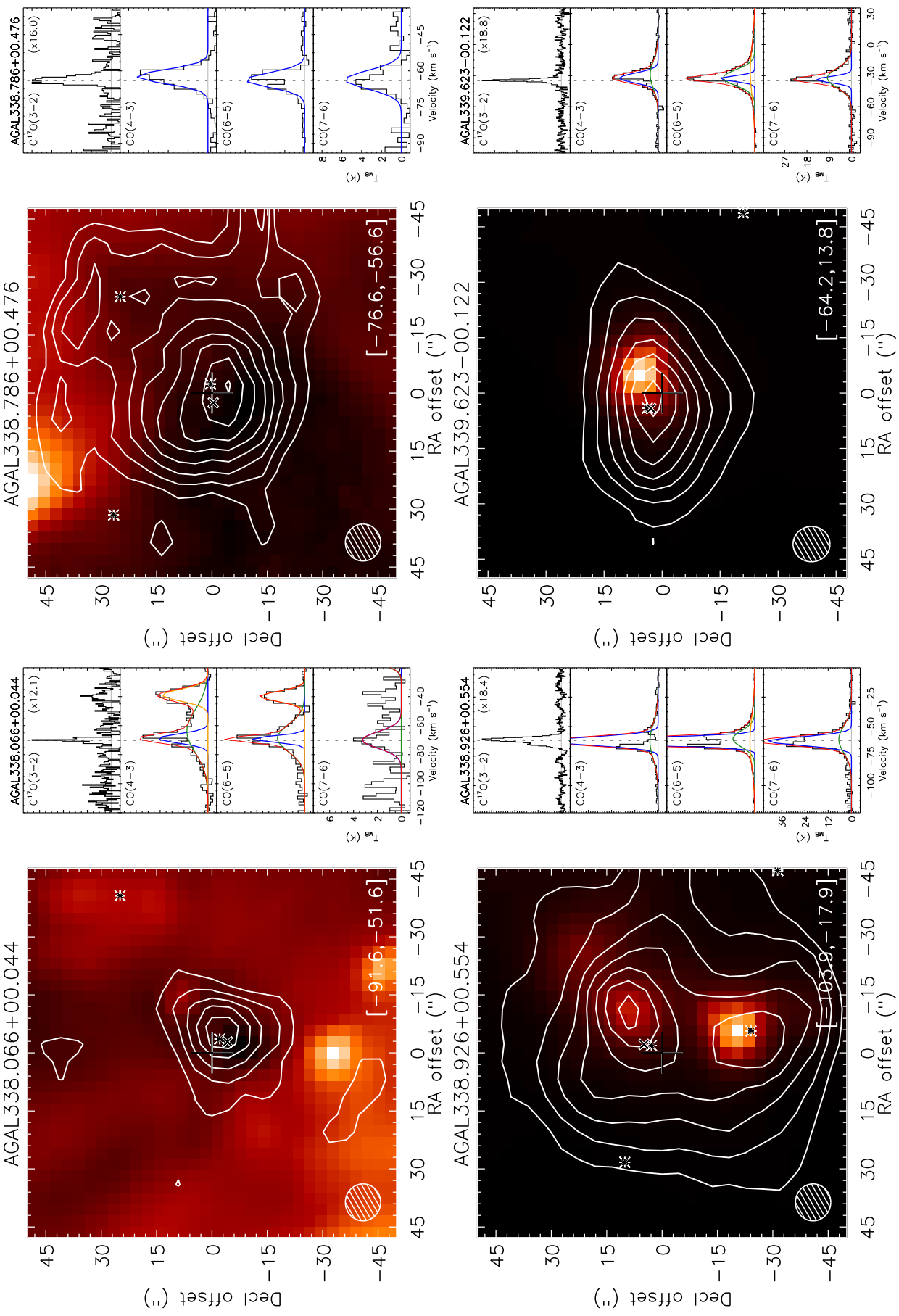


Fig. B.1: Continued.

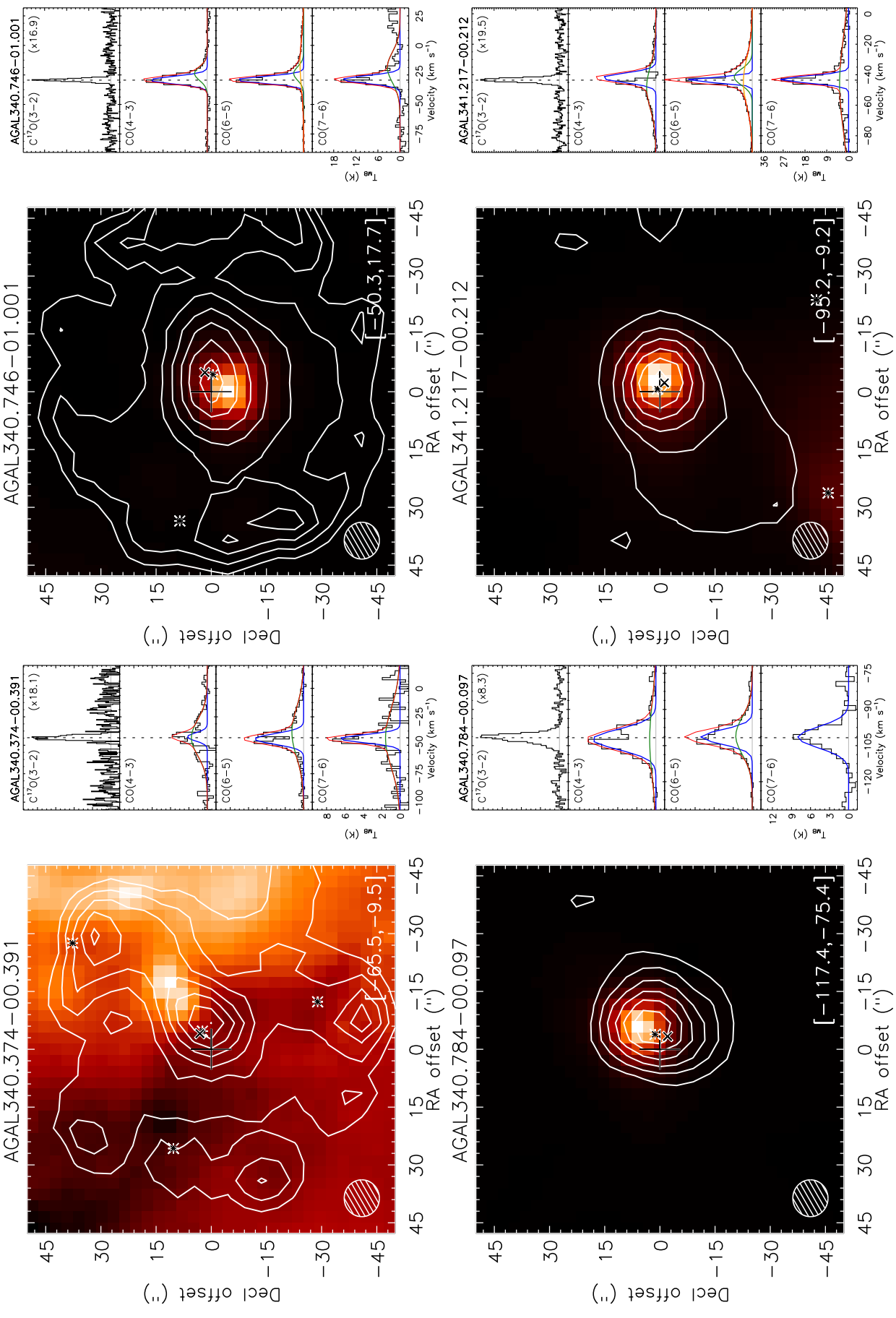


Fig. B.1: Continued.

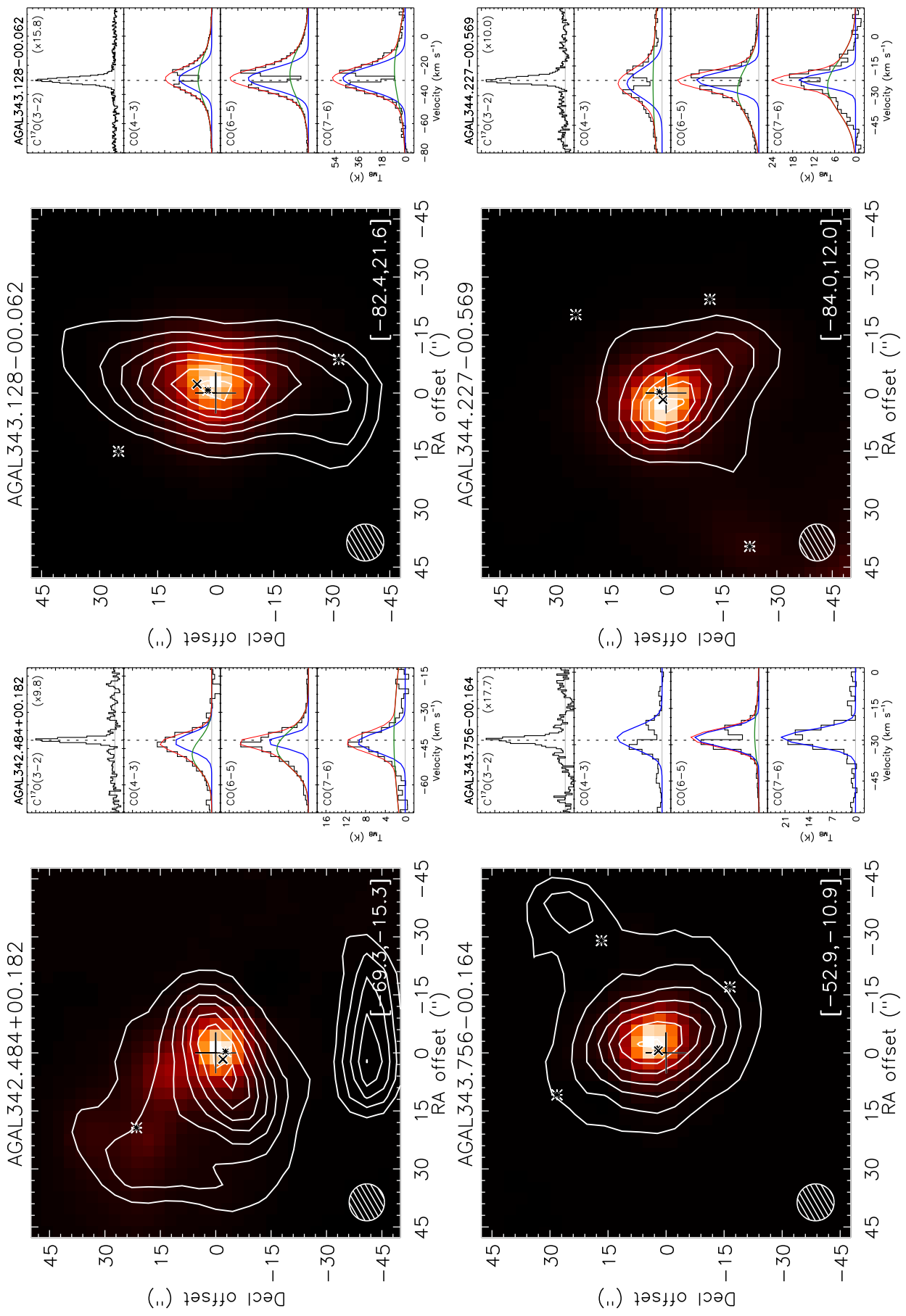


Fig. B.1: Continued.

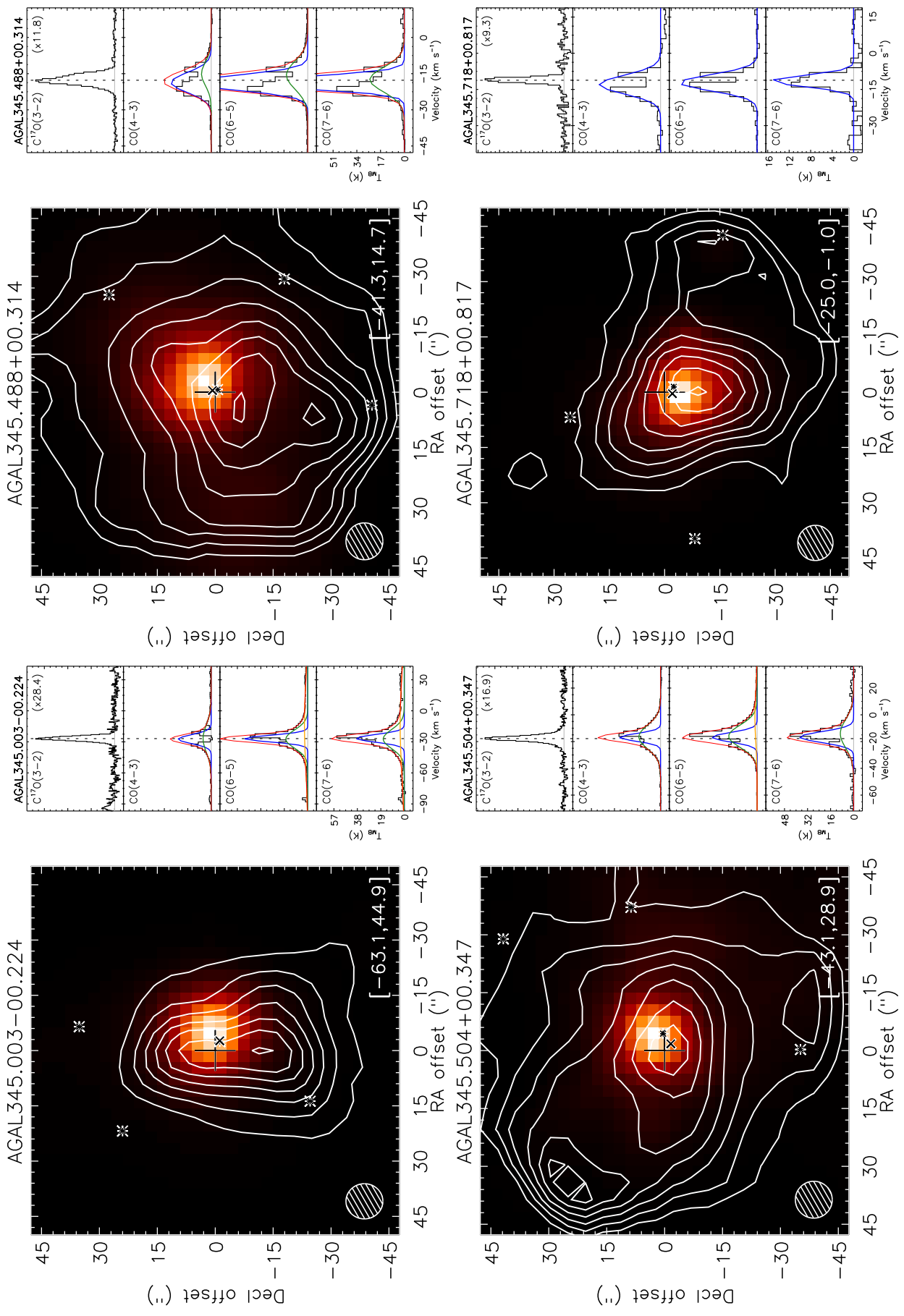


Fig. B.1: Continued.

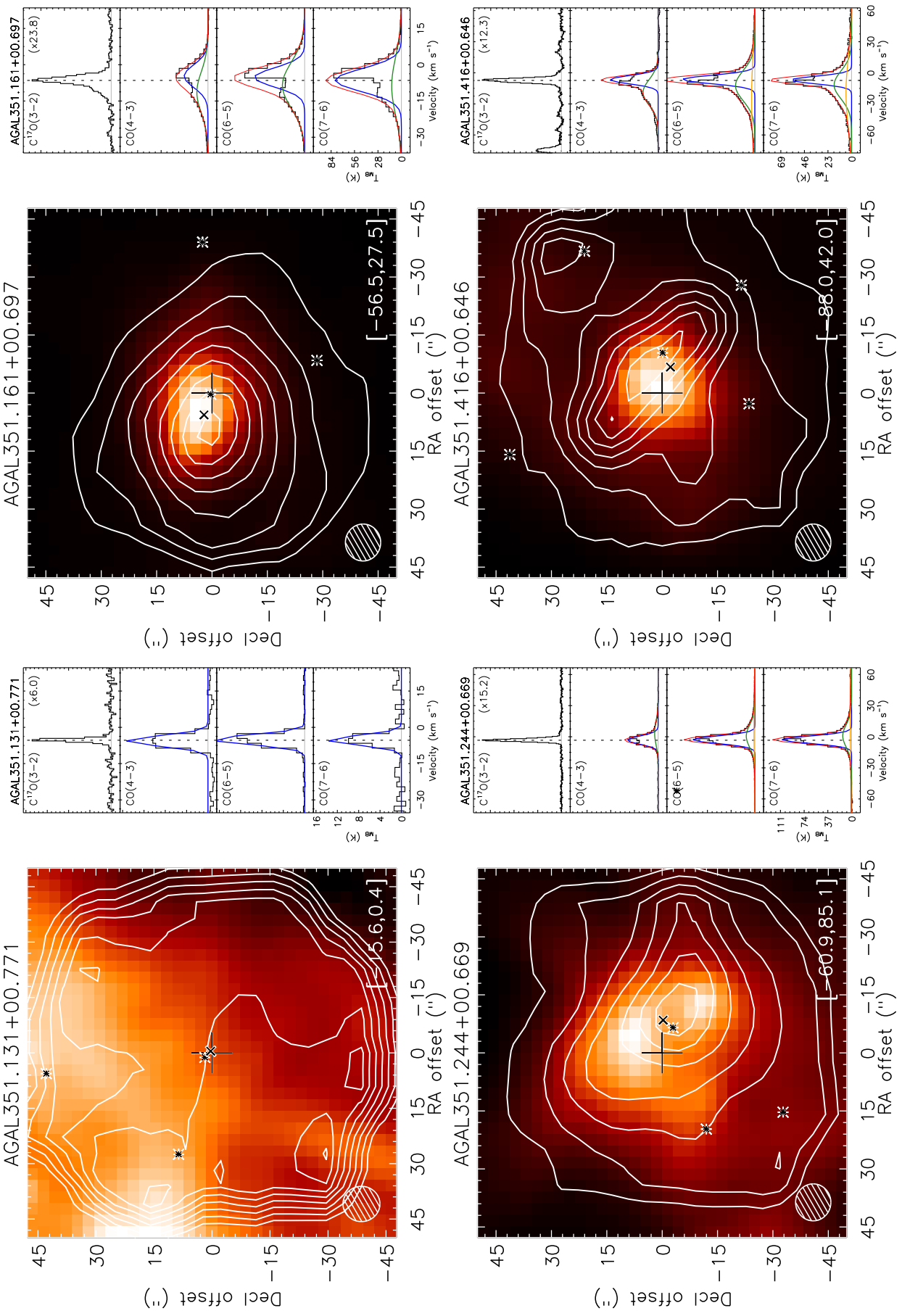


Fig. B.1: Continued.

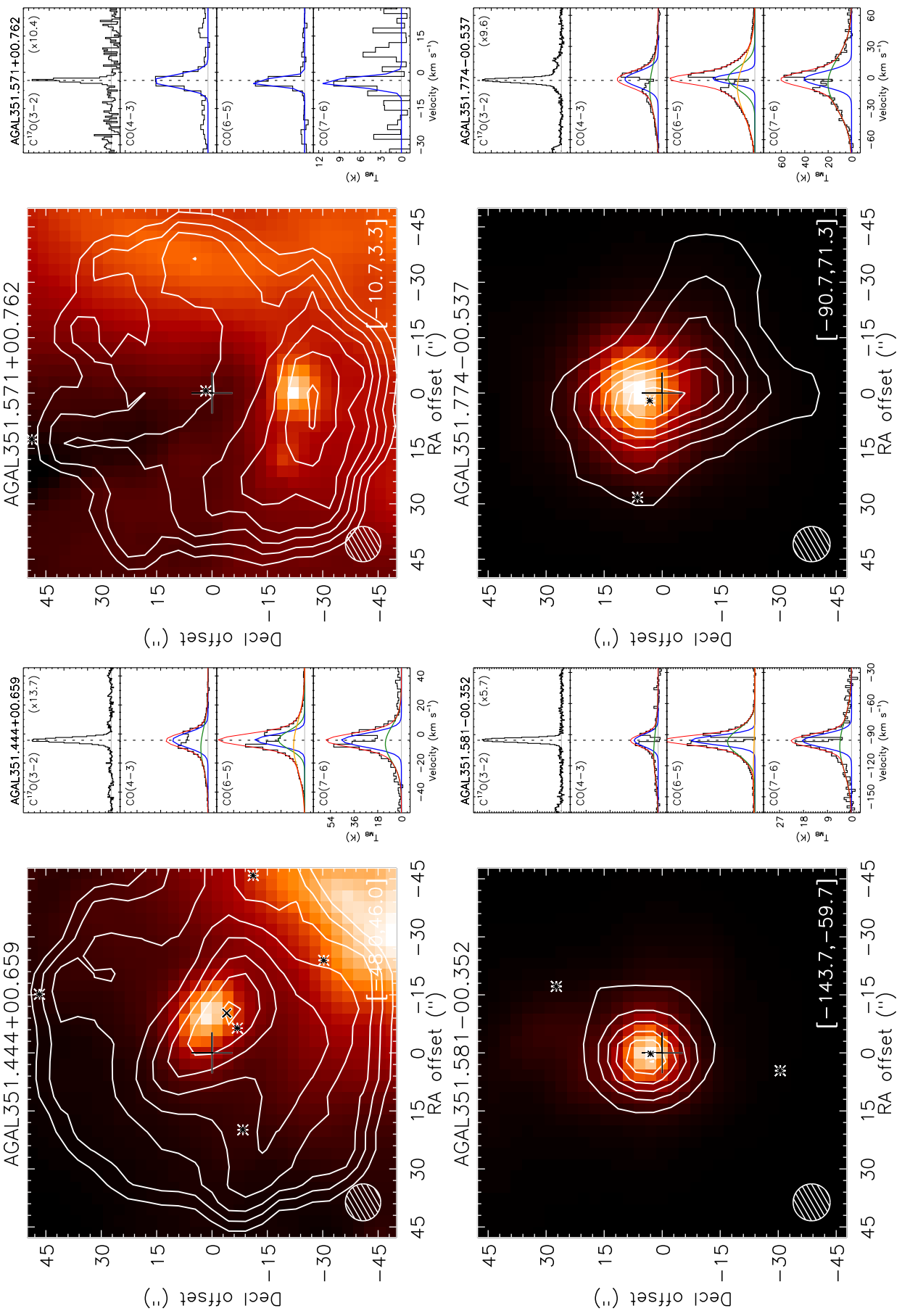


Fig. B.1: Continued.

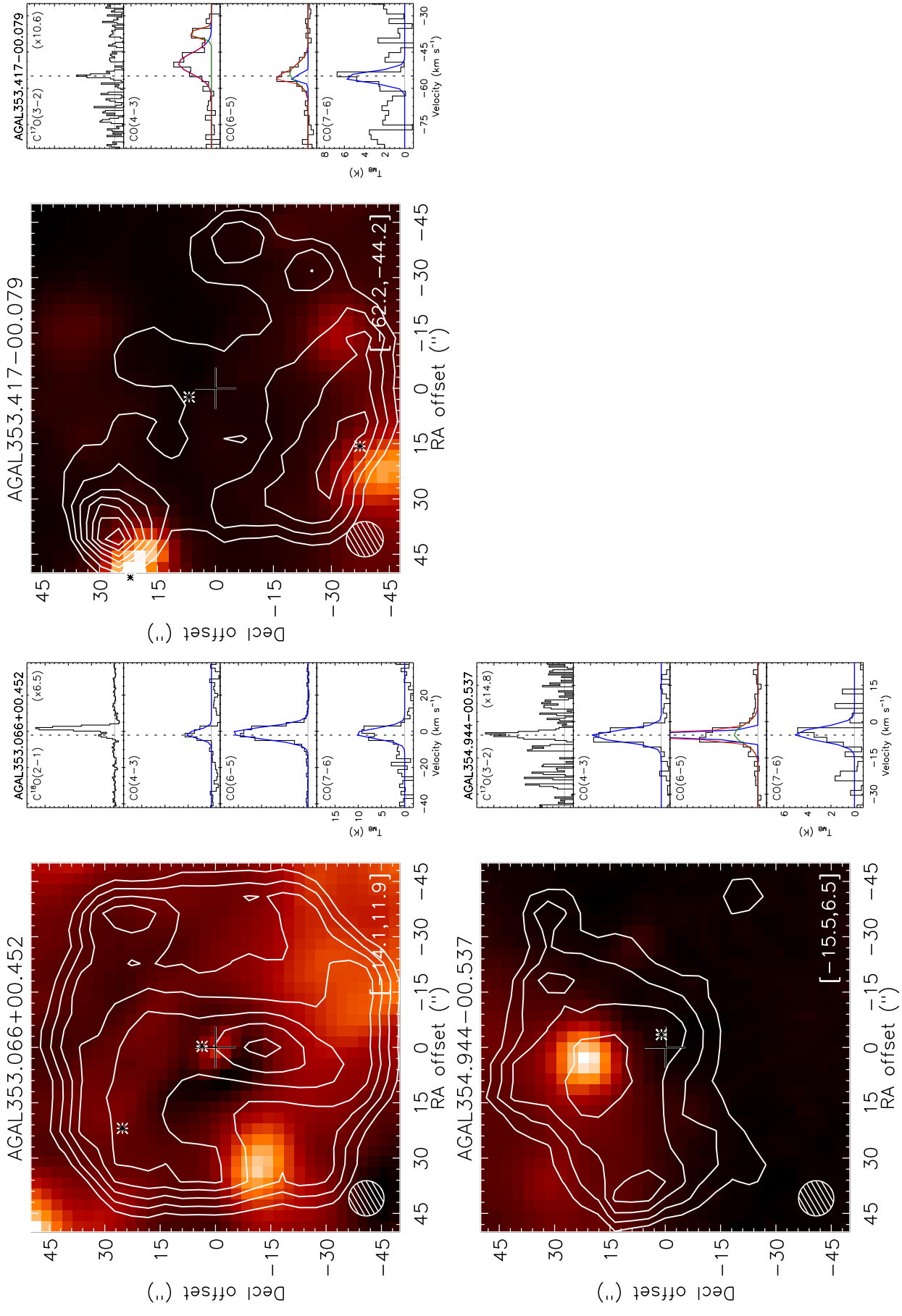


Fig. B.1: Continued.

Appendix C: Additional material

Appendix C.1: ***Integrated CO intensity maps of the secondary Gaussian components***

Figure C.1 presents the 870 μm LABOCA maps towards five TOP100 **clumps displaying secondary CO peaks in their spectra. The integrated CO (6–5) distribution of the two velocity components clearly shows that the two components trace different structures in the observed field.**

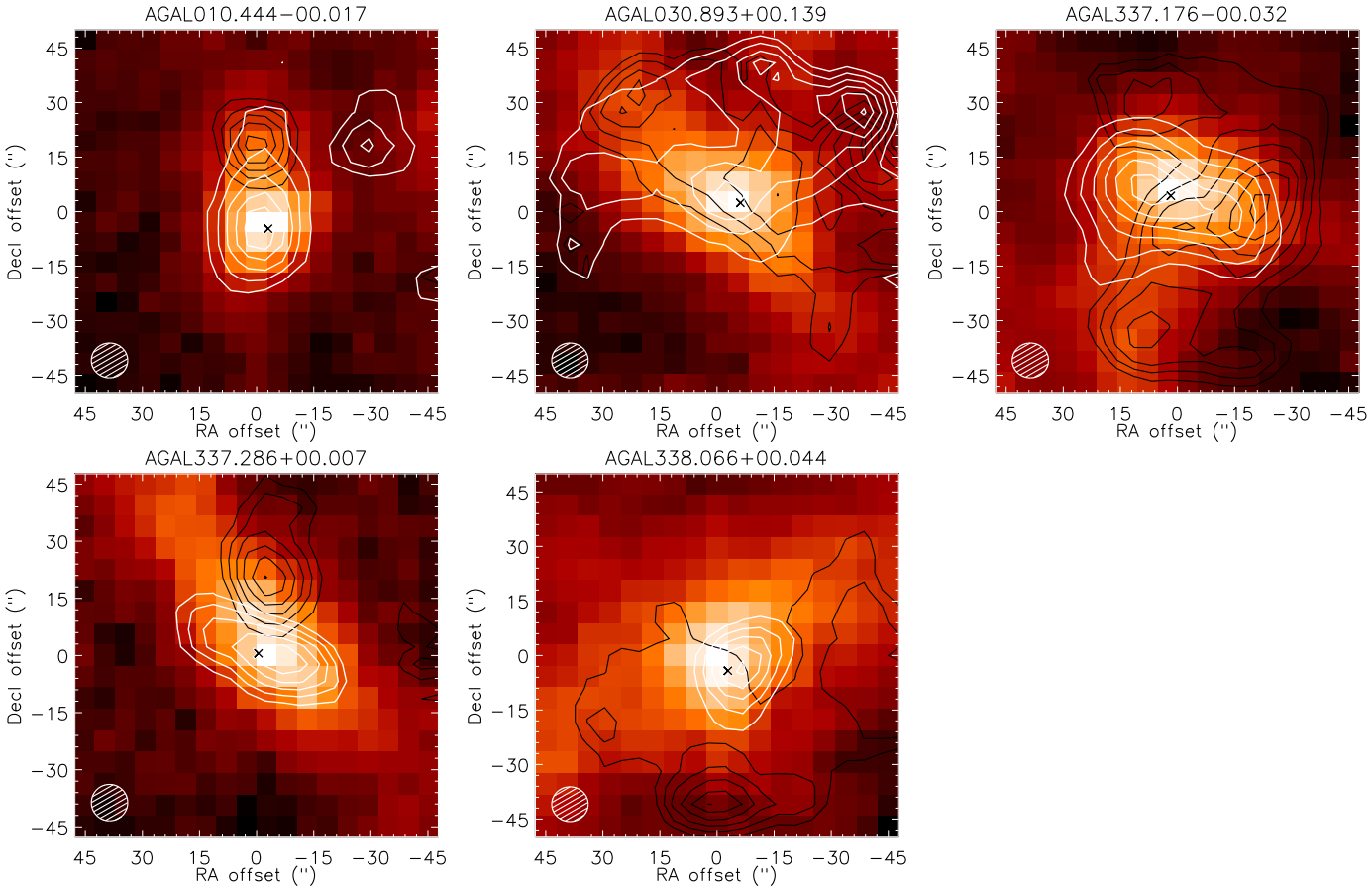


Fig. C.1: **APEX-LABOCA images at 870 μm overlaid by the CO (6–5) emission contours (C1 component in white, C2 component in black; in both cases, the contours are from 30% to 90% of the peak emission of the corresponding component in steps of 10%) towards the** TOP100 **clumps. The position of the CSC source from Contreras et al. (2013) is shown as a $\times$ symbol.**

Appendix C.2: Analysis of the mid-J CO emission in the distance-limited subsample

For completeness, in this section we present the analysis of the CO emission for the distance-limited sample (defined in Sect. 3.3).

Figure C.2 shows the average CO spectra per class integrated over a linear scale ($\sim$ 0.24 pc) for the distance-limited subsample. When compared to Fig. 6, which shows the same kind of spectra but for the full sample (using the spectra convolved to 13.″4, see Sect. 3.3), we found that the IRw **and** IRb **classes are much better separated in the distance-limited sample than in the full dataset smoothed to 13.″4. In fact, the** IRw **and** IRb **gets less distinguishable when including the outlier** IRw **clumps located at $d \geq 12$ kpc (AGAL018.606−00.074, AGAL018.734−00.226, AGAL342.484+00.182).**

Figure C.3 presents the CDF for CO (6–5) and CO (7–6) line luminosity. $L_{\rm CO}$ ranges from 2.7 to 284.0 K km s^{-1} pc^2 for the CO (6–5) transition, and 1.2 to 276.5 K km s^{-1} pc^2 for the CO (7–6) line. The KS tests indicate that the evolutionary classes are relatively more distinguishable based on the distance-limited subsample than on the full sample (excluding the comparison between IRw **and** IRb**, all tests indicated ranks KS $\geq$ 0.51, with $p \leq 0.001$ for both J transitions, see Table C.1).**

We also looked at the CO (6–5) and CO (7–6) line luminosities as function of the bolometric luminosity of the sources (see right panel of Fig. C.3), their mass and their luminosity-to-mass ratio (See Fig. C.4). Table C.2 lists the Spearman correlation factor, ρ, and its associated probability, p, for the CO line luminosity versus the bolometric luminosity, the clump mass and the luminosity-to-mass ratio of the clumps. For $L_{\rm bol}$ and $M_{\rm clump}$, the partial Spearman rank, excluding the dependency on the distance, is also provided. The ρ values are similar to those reported on Table 8 for the correlation between $L_{\rm CO}$ and $L_{\rm bol}$, based on the 13.″4 dataset, indicating no significant improvement on the correlation between these quantities.

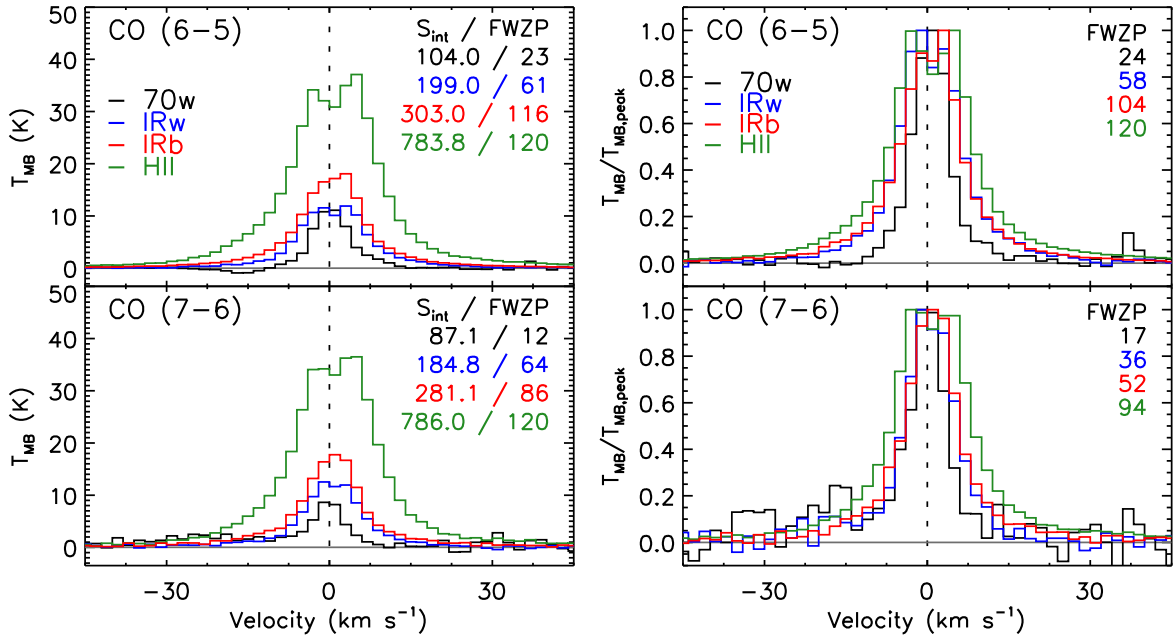


Fig. C.2: Left: Average CO (6–5) and CO (7–6) spectra of each ATLASGAL class scaled to the median distance of the distance-limited subsample ($d = 3.26$ kpc). Right: Same plot, but the average CO spectra were normalised by their peak intensity. The baseline level is indicated by the solid grey line. The black dashed line marks a velocity of 0 km s^{-1}. **The FWZP of the profiles are shown in the upper right side of the panels (in** km s^{-1} **units), together with the integrated intensity (**$S_{\rm int}$**, in** K km s^{-1} **units) of the CO profiles shown in the left panels.**

Table C.1: Kolmogorov-Smirnov statistics of the mid-J CO line luminosity for the distance-limited subsample as a function of the evolutionary class of the clumps.

Transition	70w-IRw	70w-IRb	70w-H II	IRw-IRb	IRw-H II	IRb-H II
CO (6–5)	0.60, $p < 0.001$	0.80, $p < 0.001$	0.95, $p < 0.001$	0.31, $p = 0.08$	0.71, $p <= 0.001$	0.51, $p = 0.001$
CO (7–6)	0.73, $p < 0.001$	0.80, $p < 0.001$	1.00, $p < 0.001$	0.34, $p = 0.04$	0.68, $p <= 0.001$	0.52, $p = 0.001$

Notes. The rank KS and its corresponding probability (p) are shown for each comparison. A p-value of < 0.001 indicate a correlation at 0.001 significance level. p-values of 0.05, 0.002 and < 0.001 represent the ~ 2, 3 and $> 3\,\sigma$ confidence levels.

The correlation with $M_{\rm clump}$, however, is weaker ($\rho \leq 0.44$, $p < 0.001$) than the one found towards the 13.″4 dataset ($\rho \geq 0.72$, $p < 0.001$, see Table 8). The weaker correlation between $L_{\rm CO}$ and $M_{\rm clump}$ on the distance-limited subsample might arise from the fact that the CO line luminosity is integrated over only a fraction of the beam used for estimating the mass of the clumps. Indeed König et al. (2017) used a minimum aperture size of 55.″1 for their study, while the minimum beam size adopted for the convolution of the CO was about 10″ (see Sect. 3.3).

We also found that $L_{\rm CO}$ is relatively better correlated with L/M for the distance-limited dataset ($\rho \geq 0.67$, $p < 0.001$ for all lines, see Table C.2) rather than the 13.″4 spectra ($\rho \leq 0.43$, $p \leq 0.09$ for the mid-J CO lines, see Table 8).

We compared the best fits obtained for the mid-J CO line luminosity convolved to the same linear scale with those derived using the 13.″4 data. Table C.3 reports the coefficients of the individual fits. We find that $L_{\rm CO}$ vs. $L_{\rm bol}$ follows a power-law distribution with indices of 0.53±0.03 and 0.59±0.03 for the CO (6–5) and CO (7–6) lines, respectively. Such power-law distributions are relatively less steeper than those derived towards the 13.″4 dataset, with indices of 0.63±0.04 and 0.69±0.04 for the same J transitions (see Table 7). The offset of the fits indicates the brightness of the CO emission is roughly 0.5 dex larger than the $L_{\rm CO}$ values derived using the spectra convolved to a 13.″4 beam. At least for the closest sources, such an increment in $L_{\rm CO}$ is expected due to the larger size of the beam corresponding to 0.24 pc. For example, at $d = 1.85$ kpc, the linear scale of 0.24 pc corresponds to a beam of 26.″8, which is sampling an area 4 times larger than the 13.″4 dataset.

Finally, we further investigated the effects of beam dilution on our results by integrating the CO emission over the full maps for those clumps where the aperture used by König et al. (2017) to derive the bolometric luminosity of the clumps, $\Delta\theta_{\rm ap}$, was larger or equal to the CO emission extension ($\Delta\theta \leq \Delta\theta_{\rm ap}$). Such criterion was satisfied for 92 of the 99 clumps. The best fit of the data indicates that $L_{\rm CO}$ increases with $L_{\rm bol}$ with a power-law index of 0.59 ± 0.03 and $0.68\,^{+0.05}_{-0.06}$ for the CO (6–5) and CO (7–6) transitions, consistently with the results based on the 13.″4 dataset (see Fig. 7). Similar results are also found when comparing $L_{\rm CO}$ versus $M_{\rm clump}$ and L/M. The overall results suggests that the analysis of the mid-J CO emission is robust in terms of beam dilution effects.

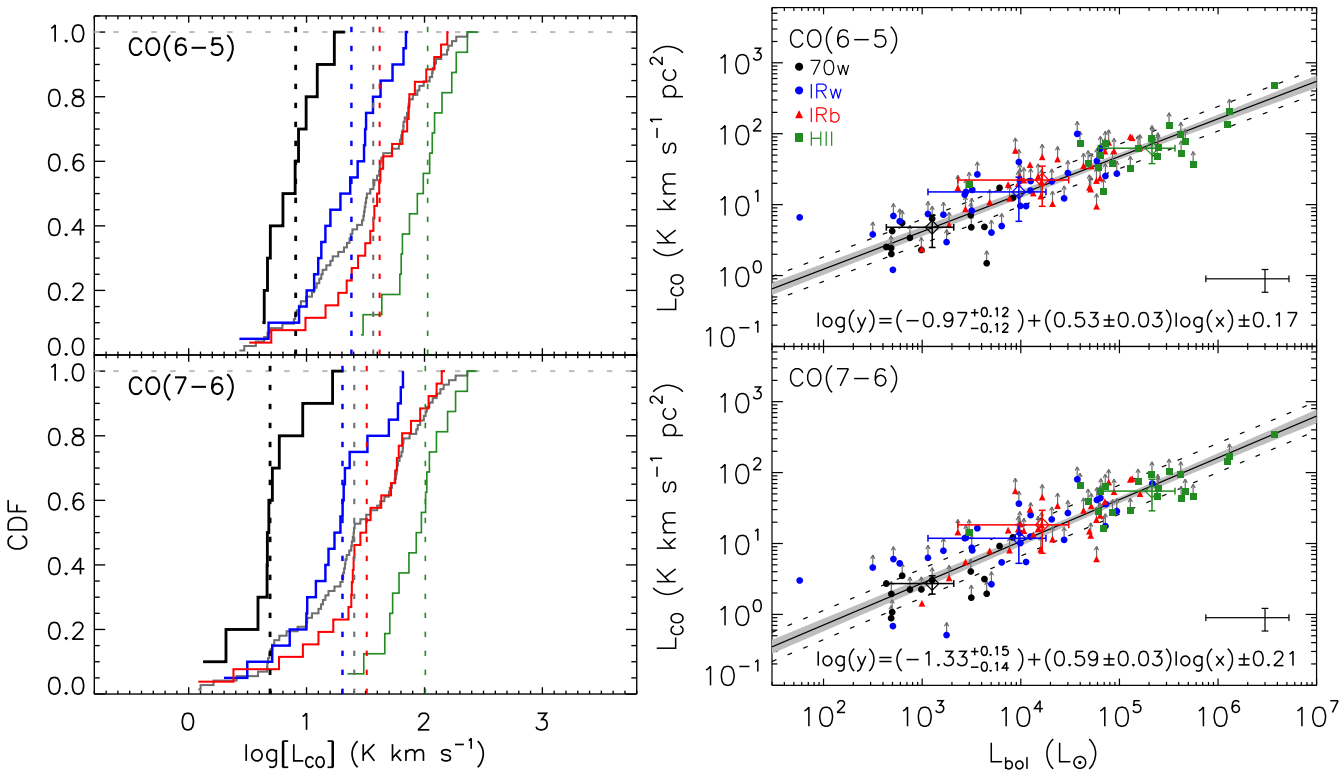


Fig. C.3: **Left panels: Cumulative distribution function (CDF) of the CO line luminosity derived using the spectra convolved to the same linear scale of 0.24 pc. The median values per class are shown as vertical dashed lines in their corresponding colours. Right: Line luminosity of the same CO lines versus the bolometric luminosity of the** TOP100 **clumps in distance-limited subsample. The median values for each class are shown as open diamonds and their error bars correspond to the absolute deviation of the data from their median value. Points having an upward arrow indicate a self-absorption feature in the spectrum and correspond to a lower limit. The typical error bars are shown in the bottom right side of the plots. The black solid line is the best fit, the light grey shaded area indicates the 68% uncertainty, and the dashed lines show the intrinsic scatter (ϵ) of the relation.**

Table C.2: Spearman rank correlation statistics for the CO line luminosity as a function of the clump properties towards the distance-limited subsample.

Property	CO (6–5)	CO (7–6)
$L_{\rm bol}$	0.86, $p<0.001$; $\rho_p=0.91$	0.87, $p<0.001$; $\rho_p=0.88$
$M_{\rm clump}$	0.44, $p<0.001$; $\rho_p=0.85$	0.41, $p<0.001$; $\rho_p=0.81$
L/M	0.67, $p<0.001$	0.79, $p<0.001$

Notes. The rank ρ and its corresponding probability (p) are shown for each comparison. A p-value of <0.001 indicate a correlation at 0.001 significance level. p-values of 0.05, 0.002 and <0.001 represent the ~ 2, 3 and $>3\,\sigma$ confidence levels. For $L_{\rm bol}$ and $M_{\rm clump}$, the partial correlation coefficient, ρ_p, is also shown.

Table C.3: Parameters of the fits of $L_{\rm CO}$, extracted within a common linear scale, as a function of the clump properties.

Transition	Property	α	β	ϵ
CO (6–5)	$L_{\rm bol}$	$-0.97^{+0.12}_{-0.12}$	0.53±0.03	0.17
	$M_{\rm clump}$	$-0.18^{+0.24}_{-0.23}$	0.48±0.08	0.43
	L/M	$0.68^{+0.05}_{-0.06}$	0.54±0.05	0.32
CO (7–6)	$L_{\rm bol}$	$-1.33^{+0.15}_{-0.14}$	0.59±0.03	0.21
	$M_{\rm clump}$	$-0.26^{+0.26}_{-0.29}$	0.47±0.08	0.51
	L/M	$0.47^{+0.06}_{-0.06}$	0.64±0.05	0.32

Notes. The fits were performed by adjusting a model with three free parameters in the form of $\log(y)=\alpha+\beta\log(x)\pm\epsilon$, where α, β and ϵ correspond to the intercept, the slope and the intrinsic scatter, respectively.

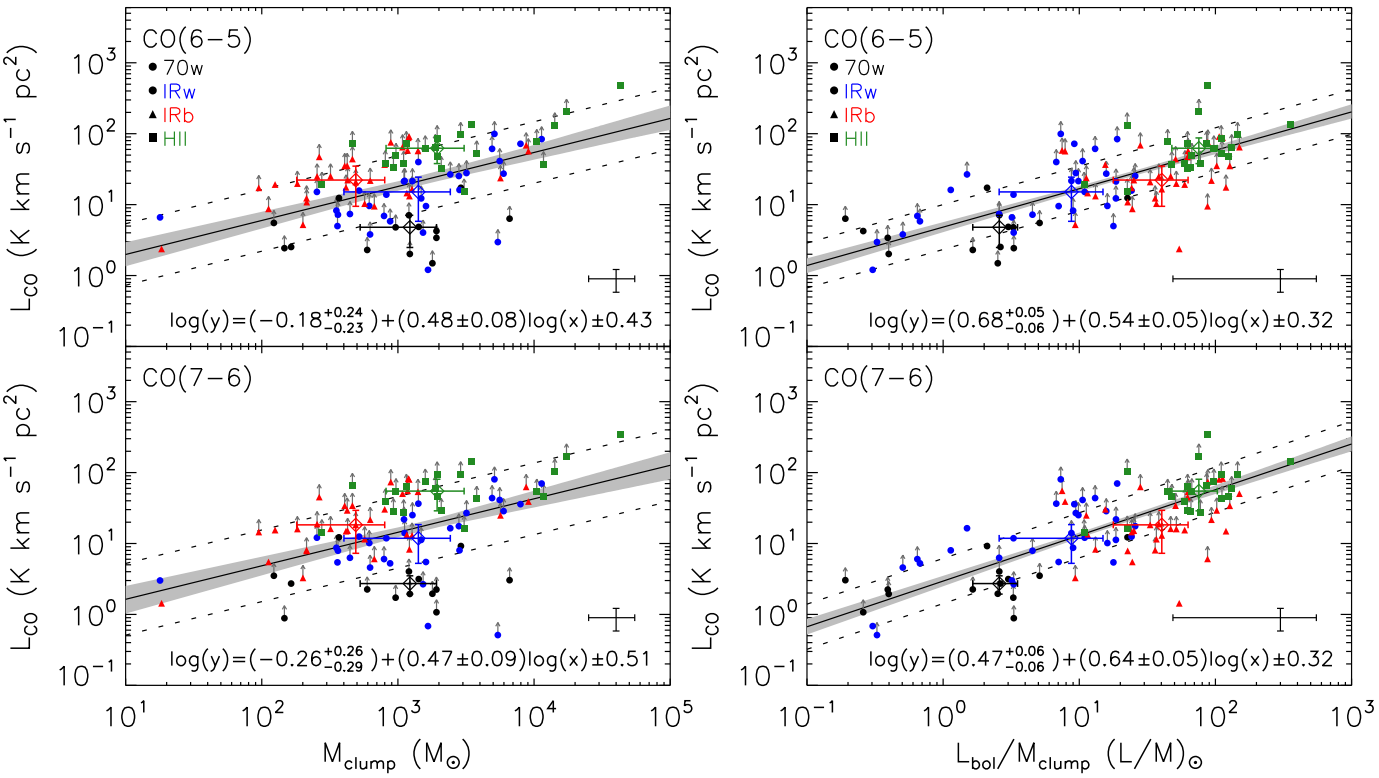


Fig. C.4: **The CO line luminosity derived using the spectra convolved to the same linear scale of 0.24 pc are shown versus their masses (left panels) and their luminosity-to-mass ratios (right). For a complete description of the plots, see Fig. C.3.**

Table C.4: Spearman statistics for the CO line luminosity as a function of the clump properties using the spectra integrated on the whole CHAMP$^+$ maps.

Transition	$L_{\rm bol}$	$M_{\rm clump}$	L/M
CO (6–5)	0.84, $p<0.001$	0.72, $p<0.001$	0.43, $p<0.001$
	$\rho_p=0.96$	$\rho_p=0.93$	
CO (7–6)	0.82, $p<0.001$	0.59, $p<0.001$	0.49, $p<0.001$
	$\rho_p=0.97$	$\rho_p=0.94$	

Notes. The rank ρ and it corresponding probability (p) are shown for each comparison. A p-value of <0.001 indicate a correlation at 0.001 significance level. p-values of 0.05, 0.002 and <0.001 represent the ~2, 3 and >3σ confidence levels.

C.2 Near-infrared survey of molecular H_2 jets

Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **450,** 4364–4398 (2015) doi:10.1093/mnras/stv914

A survey of extended H_2 emission from massive YSOs

F. Navarete,[1⋆] A. Damineli,[1] C. L. Barbosa[2] and R. D. Blum[3]

[1]*IAG-USP, Rua do Matão, 1226, 05508-900 São Paulo, SP, Brazil*
[2]*MCTI/Laboratório Nacional de Astrofísica, Rua Estados Unidos 154, CEP 37504-364 Itajubá, MG, Brazil*
[3]*NOAO, 950 N Cherry Ave, Tucson, AZ 85719, USA*

Accepted 2015 April 22. Received 2015 April 22; in original form 2014 October 28

ABSTRACT

We present the results from a survey, designed to investigate the accretion process of massive young stellar objects (MYSOs) through near-infrared narrow-band imaging using the H_2 ν=1–0 S(1) transition filter. A sample of 353 MYSO candidates was selected from the Red MSX Source survey using photometric criteria at longer wavelengths (infrared and submillimetre) and chosen with positions throughout the Galactic plane. Our survey was carried out at the Southern Astrophysical Research Telescope Telescope in Chile and Canada–France–Hawaii Telescope in Hawaii covering both hemispheres. The data reveal that extended H_2 emission is a good tracer of outflow activity, which is a signpost of accretion process on young massive stars. Almost half of the sample exhibit extended H_2 emission and 74 sources (21 per cent) have polar morphology, suggesting collimated outflows. The polar-like structures are more likely to appear on radio-quiet sources, indicating these structures occur during the pre-UCH II phase. We also found an important fraction of sources associated with fluorescent H_2 diffuse emission that could be due to a more evolved phase. The images also indicate only ∼23 per cent (80) of the sample is associated with extant (young) stellar clusters. These results support the scenario in which massive stars are formed by accretion discs, since the merging of low-mass stars would not produce outflow structures.

Key words: stars: early-type – stars: formation – stars: massive – stars: pre-main-sequence – ISM: jets and outflows – infrared: stars.

1 INTRODUCTION

The scenario of formation of low-mass stars ($M < 8$ M$_\odot$), through disc accretion controlled by magnetic field is already well understood (Shu, Adams & Lizano 1987). For high masses ($M > 8$ M$_\odot$), however, there are theoretical problems for applying the same scenario. From the observational point of view, there are too few bona fide candidates to test the models. This is due to the fact that massive young stellar objects (MYSOs) are short lived and occur inside large molecular clouds, subject to huge reddening. The relevant phases of massive stars formation are not seen directly at wavelengths shorter than the mid-infrared (MIR) and also to wavelengths longer than radio-mm. Shocked molecular outflows are 'smoking guns' of the accretion process and can be traced outside the central heavily obscured zone. The aim of this work is to use the H_2 ν=1–0 S(1) emission to constrain the possible models.

Even after decades of progress both theoretical and observational fronts, formation of massive stars remains an open question. The short time-scale for MYSO formation indicates that high rates of mass accretion are required. Two scenarios currently dominate the discussions, (a) in the first, high-mass stars are formed by accretion through a disc (Jijina & Adams 1996; Krumholz, McKee & Klein 2005); (b) in the second, high-mass stars are formed via coalescence of low-mass stars (Bonnell, Bate & Zinnecker 1998; Bonnell et al. 2001; Bally & Zinnecker 2005). In the last decade, a number of observations have been presented favouring the first scenario, among them Bik & Thi (2004) and Blum et al. (2004).

Low- and intermediate-mass stars are formed by the gravitational collapse of the parental giant molecular cloud (GMC), followed by the accretion process (Palla 1996). During the accretion phase, material is ejected as well via collimated bipolar jets. However, when a YSO reaches 8 M$_\odot$, the radiative flux becomes so intense (using $\phi = L/4\pi d^2$, the ratio between the radiative fluxes of an O5 and a B3 star – masses of ∼40 and ∼8 M$_\odot$, respectively – is ≈250) that it may interrupt the accretion flow. A process that constrains the outcoming radiation field to narrower angles may leave some room for the accretion process to continue in some directions. This seems to be the case for the outflows driven by young stars from a very broad mass range, as previous reported by several authors (Bachiller 1996; Bontemps et al. 1996; Shepherd & Churchwell 1996; Beuther et al. 2002; Wu et al. 2004). Outflows associated with high-mass objects are expected to be more energetic than the outflows observed in lower mass YSOs (Beuther et al. 2005; Zhang

⋆ E-mail: navarete@usp.br