

# Turbulence in compact to giant H II regions

J. García-Vázquez<sup>1</sup>★, William J. Henney<sup>2</sup> and H. O. Castañeda<sup>1</sup>†

<sup>1</sup>Escuela Superior de Física y Matemáticas, Instituto Politécnico Nacional, U.P. Adolfo López Mateos, Zacatenco, Ciudad de México, México C.P. 07738

<sup>2</sup>Instituto de Radioastronomía y Astrofísica, Universidad Nacional Autónoma de México, Antigua Carretera a Pátzcuaro #8701,

Ex-Hda. San José de la Huerta, Morelia, Michoacán, México C.P. 58089

Accepted XXX. Received YYY; in original form ZZZ

## ABSTRACT

Radial velocity fluctuations on the plane of the sky are a powerful tool for studying the turbulent dynamics of emission line regions. We conduct a systematic statistical analysis of the H $\alpha$  velocity field for a diverse sample of 9 H II regions, spanning two orders of magnitude in size and luminosity, located in the Milky Way and other Local Group galaxies. By fitting a simple model to the second-order spatial structure function of velocity fluctuations, we extract three fundamental parameters: the velocity dispersion, the correlation length, and the power law slope. We determine credibility limits for these parameters in each region, accounting for observational limitations of noise, atmospheric seeing, and the finite map size. The plane-of-sky velocity dispersion is found to be a better diagnostic of turbulent motions than the line width, especially for lower luminosity regions where the turbulence is subsonic. The correlation length of velocity fluctuations is found to be always roughly 2% of the H II region diameter, implying that turbulence is driven on relatively small scales. No evidence is found for any steepening of the structure function in the transition from subsonic to supersonic turbulence, possibly due to the countervailing effect of projection smoothing. Ionized density fluctuations are too large to be explained by the action of the turbulence in any but the highest luminosity sources. A variety of behaviors are seen on scales larger than the correlation length, with only a minority of sources showing evidence for homogeneity on the largest scales.

**Key words:** HII regions – ISM: kinematics and dynamics – turbulence

## 1 INTRODUCTION

Photoionized regions around high-mass stars (H II regions) show highly vigorous dynamics as a result of the star formation process and the energy and momentum injected by the newly-formed stars. Rather than manifesting a simple pattern such as expansion, infall, or rotation, the motions are frequently disordered or “turbulent” and must be characterised by statistical techniques.

In relatively low luminosity regions, such as the nearby Orion Nebula, the disordered motions are approximately transonic, with typical velocities of order 5 to 10 km s<sup>-1</sup> (Castañeda 1988; García-Díaz et al. 2008). In larger, higher luminosity regions, such as 30 Doradus in the Large Magellanic Cloud, the velocities are significantly supersonic, of order 30 km s<sup>-1</sup> (Torres-Flores et al. 2013a; Castro et al. 2018). The same tendency of increasing velocity dispersion ( $\sigma$ ) and luminosity ( $L$ ) continues up to the scale of entire galaxies with an approximate relation  $L \propto \sigma^\alpha$  that spans more than 5 orders of magnitude in luminosity with a power-law index  $\alpha = 3$  to 7 (Terlevich & Melnick 1981; Rozas et al. 2006; Chávez et al. 2014; Moiseev et al. 2015). However, it is not known whether a single physical mechanism underlies this relationship at all scales. The relative importance of gravity and the various stellar feedback mechanisms (heating, direct radiation pressure, stellar winds, cluster winds) are unclear in

many cases and are frequently disputed (Krumholz & Burkhardt 2016; Melnick et al. 2021).

One of the simplest ways to measure the velocity dispersion of an H II region is to use the Doppler width of a strong emission line, such as the optical hydrogen recombination line H $\alpha$   $\lambda$ 6563 (e.g., Roy et al. 1986). We will use  $\sigma_{\text{los}}$  to denote the root-mean-square (RMS) line-of-sight velocity dispersion determined in this way. Unfortunately, there are many processes that contribute to this width in addition to the line-of-sight turbulent velocity fluctuations, such as thermal and fine-structure broadening, instrumental broadening, dust scattering, and large-scale expansion (see Rozas et al. 2006 and García-Díaz et al. 2008 for detailed discussion). All except the last of these can in principle be approximately corrected for, which is a reasonable strategy to apply in the case of high-luminosity regions where the turbulent width is expected to be larger than the correction terms. However, for lower luminosity regions the turbulent velocities are much smaller, which limits the accuracy of such a correction process (see section 3.4 of Arthur et al. 2016).

An alternative way to study turbulent motions is to measure the fluctuations on the plane of the sky of the velocity centroids of an emission line (Von Hoerner 1951). We will denote the RMS magnitude of these fluctuations by  $\sigma_{\text{pos}}$ . In addition to being unaffected by the various nuisance broadening mechanisms discussed in the previous paragraph, this also offers the opportunity to study the spatial scale of the fluctuations by measuring average velocity differences as a function of the angular separation between two points. Different

★ E-mail: jgarcia1600@alumno.ipn.mx

† Deceased.

mathematical tools can be used to study these spatial fluctuations, such as the autocorrelation function (Lagrois & Joncas 2011) and  $\Delta$ -variance (Ossenkopf et al. 2006). In this work, we will concentrate on the second order structure function,  $B(r)$  (see section 3.1), which has been used in many previous studies of velocity fluctuations in H II regions in our own Galaxy (Münch 1958; Castañeda 1988; Roy & Joncas 1985; O'Dell & Wen 1992; Medina et al. 2014) and in external galaxies (Feast 1961; Medina Tanco et al. 1997; Lagrois & Joncas 2009; Lagrois & Joncas 2011; Melnick et al. 2021).

In this paper, we employ archival data from a wide variety of integral field and multi-longslit spectrographic datasets to analyze spatially resolved H $\alpha$  velocity maps for a sample of nine H II regions (Figure 1), ranging in size from 0.5 pc to 200 pc and in H $\alpha$  luminosity from  $10^{37}$  erg s $^{-1}$  to  $10^{39}$  erg s $^{-1}$ . The observations and the physical characteristics of our sample are described in section 2. For studying the turbulence we apply a uniform methodology to the analysis of the structure function by fitting a simple functional form that consists of a power law of slope  $m$  at small separations, transitioning to a constant value at separations larger than an autocorrelation length  $r_0$ . This is described in section 3. We also take into account three observational limitations of the datasets: the finite angular resolution,  $s_0$  (set by pixel spacing or atmospheric seeing); the finite map size,  $L_{\text{box}}$ ; and a contribution of spatially uncorrelated noise,  $B_0$ , to the observed structure function. In Appendix A we use simulated velocity maps to derive appropriate functional forms for the effects of  $s_0$ ,  $L_{\text{box}}$  and  $B_0$ . In our fits, we marginalize over these nuisance parameters to obtain robust credibility limits for the parameters of astrophysical interest:  $\sigma_{\text{pos}}$ ,  $r_0$ , and  $m$ . The results of these fits are presented in section 4. In section 5 we discuss the relation of our results to previous studies and the correlations that we find between the structure function parameters and other properties of the H II regions in our sample. Finally, in section 6 we present our conclusions.

## 2 OBSERVATIONS OF THE H II REGION SAMPLE

Previous investigations of the centroid velocity structure function in H II regions have used a variety of methodologies, which makes it difficult to compare results between different regions. Arthur et al. (2016) summarise historical results for the Orion Nebula in their Table 5. The most significant differences are seen when different emission line tracers are used, but even when using the same line, there is some variance between different studies in the derived values for both the velocity dispersion  $\sigma_{\text{pos}}$  and power-law slope  $m$ . It is therefore worthwhile to employ a uniform approach across a variety of different regions. To that end, we have selected 9 H II regions, covering a broad range in size and luminosity, for which good quality mapping of the H $\alpha$  line exists in the literature or in data archives. Figure 1 shows optical images of each region in our sample and Table 1 lists their most important physical parameters. The derived centroid H $\alpha$  velocity maps are shown in Figure 2. Our sample includes three Milky Way regions (at approximate distance of 0.5 to 2 kpc), two regions in the Magellanic Clouds (50 to 60 kpc), and four regions in more distant galaxies of the Local Group (500 to 800 kpc). Further details of the observations are given in the following section.

### 2.1 Spectroscopic datasets

Wherever possible, we calculate the centroid velocity  $V_c$  directly as the normalized first velocity moment of the spectral line intensity profile for each plane-of-sky position  $x$  on the nebula:  $V_c(x) =$

$M_1(x)/M_0(x)$ . The  $k$ th unnormalized moment  $M_k$  of the continuum-subtracted spectral intensity profile  $I(v)$  is defined as

$$M_k = \int_{\Delta v} I(v) v^k dv \quad (1)$$

where the Doppler velocity  $v$  is calculated according to the VOPT convention (Eq. [32] of Greisen et al. 2006):

$$v \equiv c(\lambda - \lambda_0)/\lambda_0, \quad (2)$$

where  $\lambda$  is the observed wavelength,  $\lambda_0$  is the rest wavelength, and  $c$  is the speed of light. The spectral window  $\Delta v$  is chosen to be just large enough to include the entire H $\alpha$  emission line. So long as the signal-to-noise is high enough, this mean velocity can be calculated to a much higher precision than the nominal velocity resolution of the spectrograph. In a few cases, as noted below, we are working with already reduced data, which are provided in the form of Gaussian fits to the line profiles. In the cases where multiple Gaussian components have been fitted, we take the flux-weighted mean velocity of these components to be the centroid.

The observed RMS spectral line width  $\sigma'$  can likewise be found from the second velocity moment as

$$\sigma'(x)^2 = [M_2(x)/M_0(x)] - V_c(x)^2. \quad (3)$$

However, many different broadening mechanisms contribute to this width, including the finite spectral resolution  $\sigma_{\text{ins}}$ , fine-structure splitting  $\sigma_{\text{fs}}$ , and thermal Doppler broadening  $\sigma_t$ . In the approximation that all these broadening mechanisms are independent and described by Gaussian profiles, then the dispersion of line-of-sight velocities due to bulk motions  $\sigma_{\text{los}}$  can be found by subtracting these contributions in quadrature from the observed width:

$$\sigma_{\text{los}}^2 = \sigma'^2 - \sigma_{\text{ins}}^2 - \sigma_{\text{fs}}^2 - \sigma_t^2. \quad (4)$$

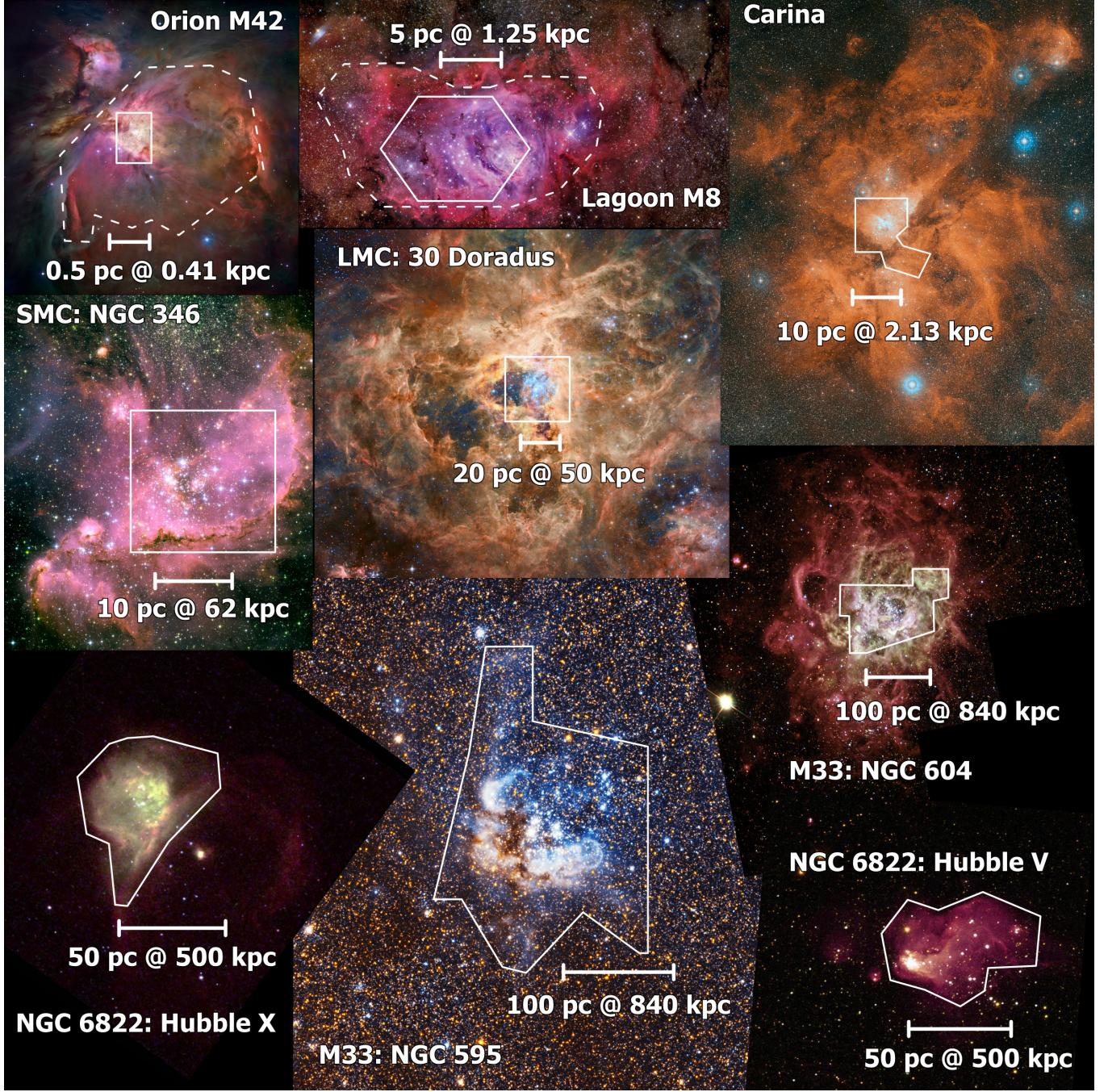
See sec 5.1 of García-Díaz et al. (2008) for further details.

#### 2.1.1 Longslit echelle spectroscopy

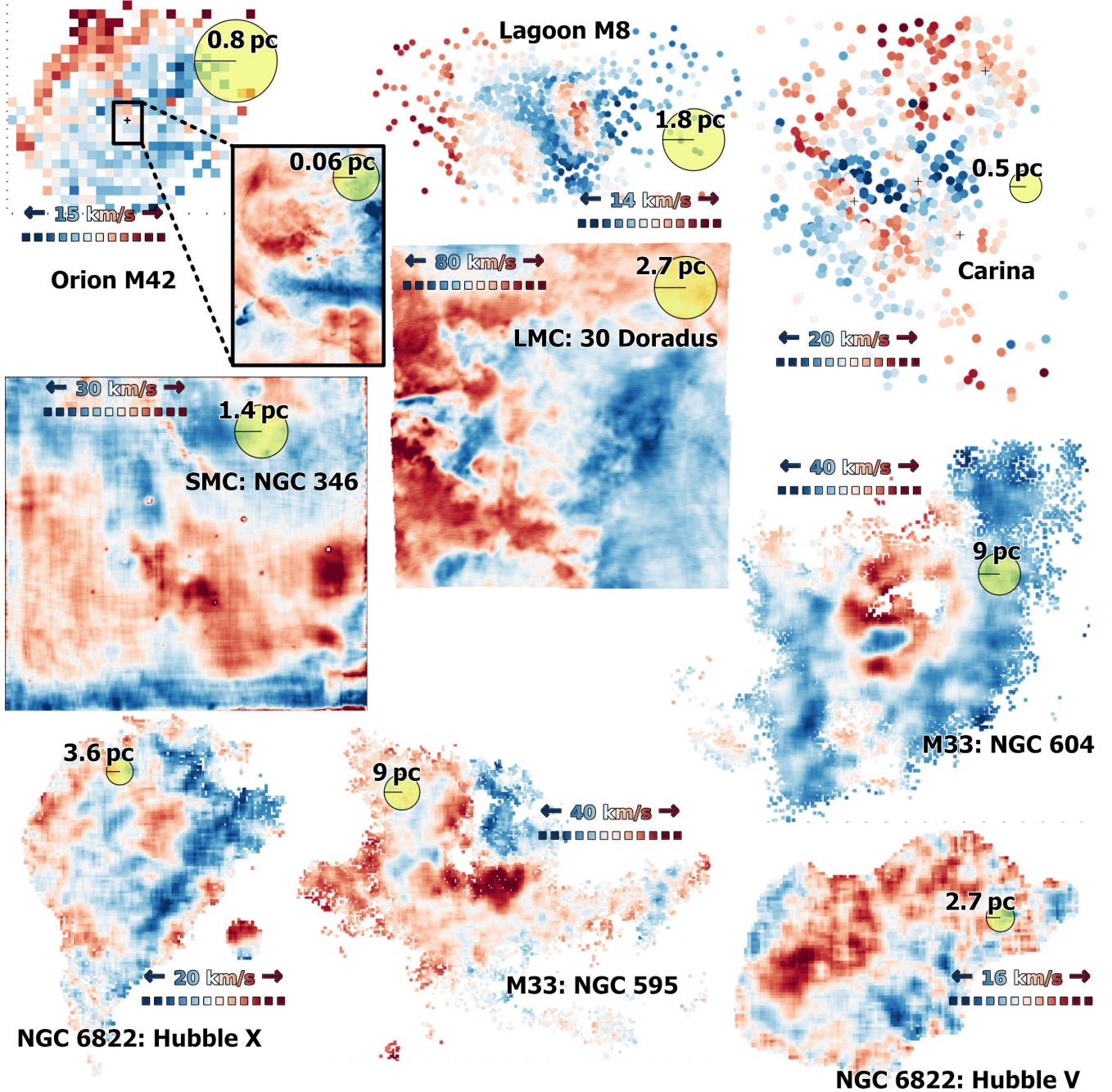
For the Orion Nebula, we use data obtained with the echelle spectrograph attached to the 4 m telescope at Kitt Peak National Observatory (KPNO) initially published in Doi et al. (2004). These observations cover a 3 arcmin  $\times$  5 arcmin region of the central (Huygens) part of the Orion Nebula. The data consist of 96 North–South orientated 300 arcsec slits, uniformly spaced at 2 arcsec intervals with a width of 0.8 arcsec with a velocity resolution of 8 km s $^{-1}$ . For the centroid velocities, we use the intensity-weighted mean velocities calculated by García-Díaz et al. (2008), who used additional observations with East–West oriented slits to refine the inter-slit velocity calibration. Unlike in the previous analysis of this dataset by Arthur et al. (2016), we do not mask out regions affected by known high-velocity outflows. This decision was made for consistency with the analysis of more distant regions, where such a masking out is not possible.

#### 2.1.2 Fabry-Pérot observations

We use the velocity field presented in Hanel (1987) to analyze the Extended Orion Nebula (EON henceforth; Güdel et al. 2008). The EON is an extended region in southwest direction from the Huygens region which presents a diminution in the brightness. The observational data was obtained using a Fabry-Pérot interferometer with a étalon separation of 0.5 mm on the 106 cm-Cassegrain telescope at Observatorium Hoher List. A number of fifteen interferograms have been taken with different pointing directions of the telescope's optical axis in the nebula with an exposure time between 10 and 40



**Figure 1.** The sample of nine H II regions employed in this study, arranged from top to bottom in order of distance. White shapes show the approximate extents of the emission line maps that we use for studying the ionized gas kinematics. Scale bars show the equivalent linear length scales at the distance to each region. For Orion, two fields are shown: the Extended Orion Nebula, sampled at a scale of 0.1 pc, and the central Huygens region, sampled at a scale of 0.004 pc. All images are combinations of optical narrow-band and wide-band filters. In most cases, red/orange/pink represents H $\alpha$  or other emission lines emitted by the ionized gas, while blue/white represents continuum starlight from young high-mass stars. In some cases, green (for Hubble X and NGC 604) or blue (for NGC 595) represents the [O III]  $\lambda 5007$  emission line. *Image credits as follows.* **Orion Nebula:** HST Treasury Program on the Orion Nebula Cluster (Robberto et al. 2013). **M8 Lagoon:** R. Jay Gabany. **Carina:** ESO/Digitized Sky Survey 2, Davide De Martin. **NGC 346:** NASA, ESA and A. Nota (ESA/STScI) (Nota et al. 2006). **30 Doradus:** Robert Gendler, Roberto Colombari, Hubble Legacy Archive, European Southern Observatory. **Hubble X:** C. R. O'Dell (Vanderbilt University), NASA and The Hubble Heritage Team (STScI/AURA). **Hubble V:** C. R. O'Dell (Vanderbilt University) and L. Bianchi (Johns Hopkins University) and Osservatorio Astronomico, Torinese, Italy), NASA, ESA, and The Hubble Heritage Team (STScI/AURA). **NGC 595:** NASA, ESA, and M. Durbin, J. Daleanton, and B. F. Williams (University of Washington). **NGC 604:** NASA and The Hubble Heritage Team (AURA/STScI).



**Figure 2.** Maps of the mean H $\alpha$  velocity for each of the regions in our sample. All velocities are relative to the systemic velocity of each region, with more negative velocities shown in blue and more positive velocities in red. The range of velocities is different for each region, as indicated on the individual maps. The correlation length of the velocity fluctuations in each region is shown as the radius of a yellow circle and labeled with its value in parsecs.

minutes. The exposures are overlapped and fall into one square of a grid with a width of 1 arcmin centered at  $\theta^1$ Ori C.

### 2.1.3 FLAMES multi-fiber spectroscopy

Archival data from Damiani et al. (2016) and Damiani et al. (2017) are used to study the Carina Nebula and Lagoon Nebula, respectively. These were obtained as a by-product of a study of young stars in their respective regions as part of the Gaia-ESO Spectroscopic Survey

(Gilmore et al. 2012; Randich et al. 2013) using VLT/FLAMES with the GIRAFFE and UVES spectrographs (Pasquini et al. 2002). The spectra are from multiple discrete fiber positions in each nebula (866 positions in Carina; 1089 in the Lagoon), visible as colored disks in Figure 2. The angular separation between fibers varies across the maps with average nearest-neighbor distance of  $(21 \pm 17)$  arcsec. The spectral resolution ranges from  $6 \text{ km s}^{-1}$  (UVES) to  $16 \text{ km s}^{-1}$  (GIRAFFE). We do not have access to the individual spectra, but instead use the Gaussian fits to the line profiles from (Damiani et al. 2016,

2017), obtained from data tables downloaded from the CDS Vizier service.<sup>1</sup> For the Lagoon, we use single-Gaussian fits, while for Carina we take the flux-weighted mean of the blue and red component of two-Gaussian fits.

#### 2.1.4 MUSE integral field spectroscopy

For the Magellanic Cloud regions 30 Doradus and NGC 346 we use data obtained with the MUSE spectrograph (Bacon et al. 2010, 2014) on the VLT. Each exposure consists of a  $300 \times 300$  pixel spectral image with a plate scale of  $0.2 \text{ arcsec pixel}^{-1}$  and a spectral resolution of  $110 \text{ km s}^{-1}$  at  $\text{H}\alpha$ . For 30 Doradus, four separate exposures are mosaicked to give a square field of size 2 arcmin (Castro et al. 2018). We use centroid velocities of single Gaussian fits to the  $\text{H}\alpha$  line.<sup>2</sup> For NGC 346, we use a single field that was observed in 2016 as part of ESO observing program 098.D-0211(A) (P.I. W.-R. Hamann). We obtained the pipeline-reduced data cubes from the ESO data archive<sup>3</sup> and extracted the  $\text{H}\alpha$  profile using tools in the MPDAF python library.<sup>4</sup>

#### 2.1.5 TAURUS-II Fabry-Perot interferometer

The archival data for the extragalactic giant regions (GEHRs), NGC 604, NGC 595, Hubble X and Hubble V was obtained with the Fabry Perot TAURUS-II instrument (Gordon et al. 2000) on the 4.2 m William Herschel Telescope (WHT) and was retrieved from the La Palma archive<sup>5</sup>. The IPCS-II detector was used with a etalon of  $125 \mu\text{m}$ . Each data cube is  $256 \times 256 \times 100$  with a spatial scale of  $0.26 \text{ arcsec pixel}^{-1}$  with a total integration time of 3600 s (36 s per frame). For more details on the observational setup see table 1 of Sabalisck et al. (1995). The reduction and analysis was done with the packages TAUCAL y TAUFRITS (Lewis & Unger 1992) and the emission spectrum was adjusted with a single Gaussian fit. The NGC 604 observations were previously reported and analyzed in Sabalisck et al. (1995), Medina Tanco et al. (1997) and Melnick et al. (2021).

## 2.2 Physical properties of the sample regions

Table 1 summarizes the properties of the regions in our sample with data taken from the literature. Further details on each individual source and their references are given in Appendix C.

## 3 PLANE-OF-SKY VELOCITY FLUCTUATIONS

Figure 2 shows maps of the  $\text{H}\alpha$  centroid velocity  $V_c$  for each H II region in our sample, calculated as described above in section 2.1 and after subtracting the mean systemic velocity,  $\langle V_c \rangle$  in each case. The overall amplitude of the plane-of-sky velocity fluctuations can be characterized by a dispersion,  $\sigma_{\text{pos}}$ , defined as:

$$\sigma_{\text{pos}}^2 = \left\langle [V_c(\mathbf{x}_i) - \langle V_c \rangle]^2 \right\rangle, \quad (5)$$

<sup>1</sup> <https://doi.org/10.26093/cds/vizier.35910074> and <https://doi.org/10.26093/cds/vizier.36040135>.

<sup>2</sup> Data kindly provided by Norberto Castro Rodríguez, priv. comm.

<sup>3</sup> [http://archive.eso.org/wdb/wdb/eso/eso\\_archive\\_main/query?prog\\_id=098.D-0211\(A\)](http://archive.eso.org/wdb/wdb/eso/eso_archive_main/query?prog_id=098.D-0211(A))

<sup>4</sup> <https://mpdaf.readthedocs.io/en/latest/>

<sup>5</sup> <http://casu.ast.cam.ac.uk/casuadc/ingarch>

**Table 1.** Summary of properties of our H II region sample. Diameters are taken from Table 2 of Kennicutt (1984). References for other properties are given in Appendix C. Mean non-thermal linewidth values ( $\langle \sigma_{\text{los}} \rangle$ ) shown here are corrected considering equation (4) which takes into account  $\sigma_{\text{fs}}$ .

H II Region	Distance, $d$ [kpc]	Diameter, $D_{\text{H II}}$ [pc]	$\log L(\text{H}\alpha)$ [erg s $^{-1}$ ]	$\langle \sigma_{\text{los}} \rangle$ [km s $^{-1}$ ]
Orion	$0.44 \pm 0.02$	$5 \pm 0.5$	37.18	$9.9 \pm 1.2^{\text{a}}$
Lagoon	$1.25 \pm 0.12$	$25 \pm 2.5$	37.47	$11.2 \pm 1.6$
Carina	$2.35 \pm 0.05$	$200 \pm 20$	39.01	$18.6 \pm 3.3^{\text{a}}$
30 Dor	$50.0 \pm 0.2$	$370 \pm 37$	39.46	$21.7 \pm 2.2$
NGC 346	$62 \pm 1$	$220 \pm 22$	38.77	$9.6 \pm 1.0$
Hubble X	$490 \pm 40$	$160 \pm 16$	38.21	$10.0 \pm 0.02$
Hubble V	$490 \pm 40$	$130 \pm 13$	38.3	$9.8 \pm 0.03$
NGC 595	$820 \pm 30$	$400 \pm 40$	38.95	$16.5 \pm 0.1$
NGC 604	$820 \pm 30$	$400 \pm 40$	39.42	$17.5 \pm 0.30$

<sup>a</sup>Value obtained from our observations.

where the average is performed over all observed points  $i$  in a given map. Note that  $\sigma_{\text{pos}}$  is also the RMS width of the probability density function (PDF) of  $V_c$ . This PDF for each region is shown in Figure 3 on a common velocity scale and arranged in order of decreasing  $\sigma_{\text{pos}}$ . It can be seen that the amplitude of the velocity fluctuations varies considerably between regions, with  $\sigma_{\text{pos}}$  in 30 Dor being more than 5 times higher than in nearby regions such as Orion.

All of the PDFs are significantly non-Gaussian and many show evidence of a multi-modal structure. Some, such as Orion core and Hubble V, show a single dominant velocity component, but most show several distinct peaks, as many as four in the case of Carina. There is no clear tendency for the number of identifiable velocity components to increase with increasing velocity dispersion. Instead, both the widths of the individual peaks and the separation between them seem to increase in tandem. For most regions the PDFs are approximately symmetrical, but two regions show significant skewness: Orion core has a smooth asymmetrical tail towards the blue, whereas NGC 346 has one towards the red.

## 3.1 The second-order structure function

In order to probe the dependence of the velocity fluctuations on spatial scale, the primary tool that we employ is the second order structure function of differences in velocity centroids,  $B(r)$ , which is a function of the scalar separation or lag,  $r$ , between two points on the plane of the sky:

$$B(r) = \left\langle [V_c(\mathbf{x}_j) - V_c(\mathbf{x}_i)]^2 \right\rangle_{|\mathbf{x}_j - \mathbf{x}_i| \approx r}. \quad (6)$$

The averaging is performed over all pairs of points  $(i, j)$  whose scalar separation  $|\mathbf{x}_j - \mathbf{x}_i|$  is close to  $r$ , irrespective of the orientation of the separation vector. In practice, we achieve this by binning the separations with a constant logarithmic width of 0.05 dex.

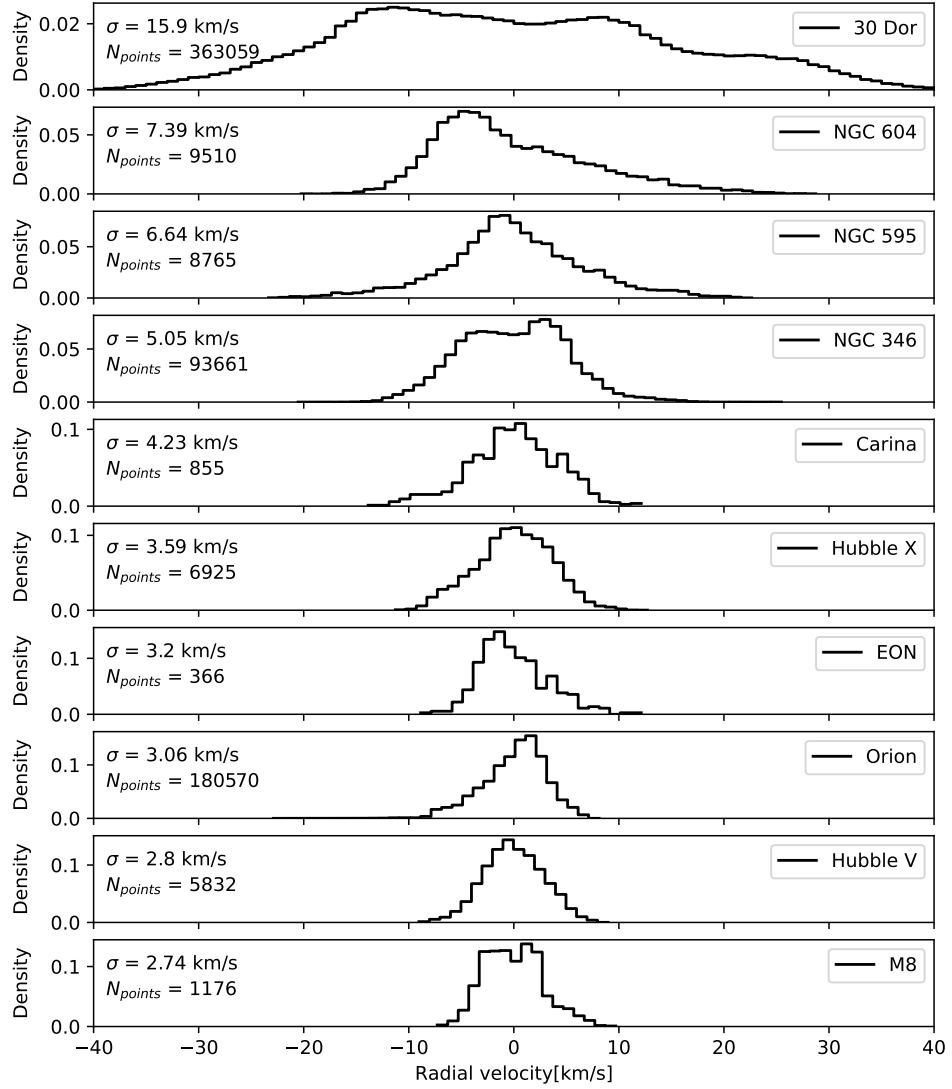
We will also employ the related quantity of the normalized spatial autocovariance or autocorrelation function:

$$C(r) = \frac{1}{\sigma_{\text{pos}}^2} \left\langle [V_c(\mathbf{x}_j) V_c(\mathbf{x}_i)] \right\rangle_{|\mathbf{x}_j - \mathbf{x}_i| \approx r}. \quad (7)$$

If the fluctuations are perfectly spatially homogeneous, then the two quantities are related (Scalo 1984) as:

$$B(r) = 2\sigma_{\text{pos}}^2 [1 - C(r)]. \quad (8)$$

In less ideal situations, then  $B(r)$  is to be preferred since it is relatively



**Figure 3.** Histograms of the plane-of-sky variation in the H $\alpha$  centroid velocities. Each panel shows the fraction of the observed spatial points (pixels or fiber positions) with a given velocity offset from the mean systemic velocity of each region. The RMS width  $\sigma_{\text{pos}}$  of each distribution is marked, as is the number of spatial points. The bin width is  $1 \text{ km s}^{-1}$  and the regions are arranged in order of decreasing  $\sigma_{\text{pos}}$ .

unaffected by non-stationary effects such as large-scale linear gradients (Scalo 1984). Nonetheless,  $C(r)$  is more amenable to heuristic reasoning in some cases, which we will take advantage of below.

### 3.2 A heuristic model for the structure function

A common property of homogeneous fluctuating velocity fields is that neighboring points tend to have similar velocities ( $C(r) \approx 1$  for small  $r$ ), whereas points that are far apart may have very different velocities ( $C(r) \ll 1$  for large  $r$ ). The value of the separation that corresponds to the transition between these two regimes is called the correlation length,  $r_0$ . In the simplest case, two points separated by  $r \gg r_0$  have totally uncorrelated velocities in the sense that knowledge of the velocity at the first point is of no help in predicting the velocity at the second point. At scales smaller than  $r_0$ , the fluctuations often show a power-law behavior as a function of  $r$ .

In order to capture these two behaviors, we therefore propose the following idealized 2-parameter model for the autocorrelation

function:

$$C_{\text{mod}}(r; r_0, m) = 2^{-(r/r_0)^m} \quad (9)$$

in which  $r_0$  is the correlation length (see above) and  $m$  is the power-law slope at small scales. This is constructed so that  $C_{\text{mod}}(r) = 1/2$  at  $r = r_0$ , while the exponential form ensures that  $C(r)$  rapidly approaches zero for larger separations. We assume the validity of equation (8) to determine the structure function from this model autocorrelation function:

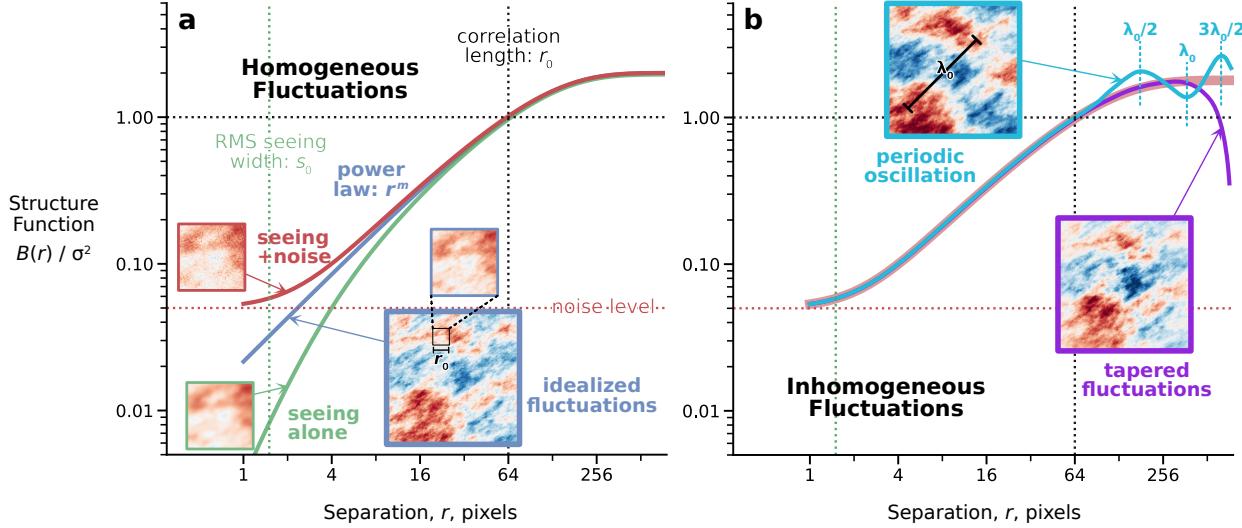
$$B_{\text{mod}}(r) = 2\sigma_{\text{pos}}^2 \left[ 1 - 2^{-(r/r_0)^m} \right] \quad (10)$$

This has the following properties:

1. Small scales:  $B_{\text{mod}}(r) \propto r^m$  for  $r \ll r_0$ ;
2. Correlation scale:  $B_{\text{mod}}(r_0) = \sigma_{\text{pos}}^2$ ;
3. Large scales:  $B_{\text{mod}}(r) \rightarrow 2\sigma_{\text{pos}}^2$  for  $r \gg r_0$ .

An example is shown by the blue line in Figure 4a.

In a previous paper, we used a different functional form (See



**Figure 4.** Example of the model structure function for homogeneous fluctuations, together with a realization of the corresponding velocity field on a  $512 \times 512$  pixel grid. (a) Idealized structure function (blue line) for the parameters  $m = 1$  and  $r_0 = 64$  pixels. Red and green lines show the effects of observational limitations at small scales due to seeing and noise. (b) Examples of inhomogeneous effects at large scales, which are not included in the model structure function. Note how hard it is to detect by eye the difference between these fields and the one shown in panel a, despite the important changes in the large-scale end of the structure function.

Fig. 13 of Arthur et al. 2016):  $C(r) = 1/[1 + (r/r_0)^m]$ , as proposed in Scalo (1984), which behaves identically to equation (9) in the first two regimes, but is more gradual in its approach to the large-scale asymptote of  $2\sigma_{\text{pos}}^2$ . However, we find that equation (9) provides a much improved fit to the observed structure functions of our sample regions (see following section).

Slightly different definitions of the correlation length are found in the literature. For instance the definition of Jaupart & Chabrier (2022) in two dimensions is

$$\ell_c^2 = \frac{1}{4} \int C(r) d^2r = \frac{\pi}{2} \int_0^\infty r C(r) dr \quad (11)$$

while the “integral length scale” (Pope 2000) is defined as

$$\mathcal{L} \equiv \int_0^\infty C(r) dr. \quad (12)$$

For our model structure function, these evaluate to

$$\mathcal{L} = \frac{\Gamma(1/m)}{m (\ln 2)^{1/m}} r_0 \quad (13)$$

and

$$\ell_c^2 = \frac{\pi \Gamma(2/m)}{2m (\ln 2)^{2/m}} r_0^2 \quad (14)$$

where  $\Gamma$  is the usual Gamma function. For example, if  $m = 1$  these become  $\mathcal{L} \approx 1.443r_0$  and  $\ell_c \approx 1.808r_0$ .

### 3.2.1 Effects of observational limitations at small spatial scales

Two observational effects can modify the observed structure function at the smallest scales. The first is the blurring of the observed image by atmospheric seeing, which tends to reduce  $B(r)$  for small  $r$ . In Appendix A2, we perform numerical experiments with synthetic velocity fields and find that a good approximation to the effects of seeing is to multiply the model structure function by a factor:

$$S(r) = \left[ \frac{1}{1 + 1.25s_0/r_0} \right] \left[ \frac{1}{1 + (2.6s_0/r)^{1.5}} \right], \quad (15)$$

where  $s_0$  is the RMS width of the seeing profile<sup>6</sup>, and  $r_0$  is the correlation length. An example is shown by the green line in Figure 4a. The effect of seeing is to make the structure function bend down away from the idealized power law, becoming increasingly steep at scales smaller than  $s_0$ , but having a noticeable effect at scales up to  $10 \times s_0$ .

The second effect is the presence of instrumental noise in the observational measurements, such as Poisson noise from photon-counting statistics. Although this affects all scales, it is most noticeable for small separations, where the intrinsic structure function is smallest. If the noise is spatially uncorrelated (equal and independent uncertainties in the velocity measurement of each pixel), then its contribution to the structure function is independent of separation, which means it can be represented as a constant term  $B_{\text{noise}}$ .

Combining both the effects of seeing and noise yields the corrected model structure function:

$$\tilde{B}_{\text{mod}}(r) = B_{\text{mod}}(r) S(r) + B_{\text{noise}} \quad (16)$$

An example of this equation is shown by the red line in Figure 4. Note that seeing and noise have opposite effects on the slope of the structure function at intermediate scales: seeing tends to steepen the slope, while noise tends to flatten it. This can lead to a degeneracy between these two parameters when fitting observations over a limited range of separations.

### 3.2.2 Effects of observational limitations at large spatial scales

In Appendix A1 we consider the effects on the structure function of the finite size  $L$  of the observational map. If the true correlation length  $r_0$  is more than a small fraction of the map size (approximately,  $r_0 > 0.1L$ ), then the apparent velocity variance measured directly

<sup>6</sup> Note that the full-width half maximum (FWHM) seeing width is  $2(2 \ln 2)^{1/2}s_0 \approx 2.35s_0$ .

**Table 2.** Bounds of allowed values for parameters in model fits

Parameter	Lower	Upper
$\sigma^2$	0.25 $\max[B_{\text{obs}}]$	2 $\max[B_{\text{obs}}]$
$r_0$	0.01 $L$	2 $L$
$m$	0.5	2.0
$s_0$	0.1 arcsec	1.5 arcsec
$B_{\text{noise}}$	0	3 $\min[B_{\text{obs}}]$

Note:  $\max[B_{\text{obs}}]$  and  $\min[B_{\text{obs}}]$  are over all bins in the observed structure function with  $r < L/2$ .

from the data is less than the true  $\sigma_{\text{pos}}^{-2}$  of the underlying fluctuations (see Figure A2). This is because the observed map is not large enough to sample the full range of velocity variations. At the same time, the apparent correlation length is smaller than the true value. However, this does not mean that the model structure function needs to be modified. In the appendix we show that the model fitting is capable of determining the true fluctuation variance and true correlation length (both to within 10%), so long as  $L > 3 r_0$ . This is a significant improvement over direct empirical measurement, which requires  $L > 10 r_0$  in order to be accurate (compare panels a and b of Figure A2).

### 3.2.3 Additional effects omitted from the model structure function

For separations that are comparable to the diameter  $D$  of the H II region, then the supposition of homogeneity may break down. For instance, if the amplitude of the velocity fluctuations is higher in the core of the region than in the outskirts, then the autocorrelation function will be U-shaped since all separations with  $r > D/2$  must be between two points in the outskirts, which will tend to be relatively well correlated simply because both velocities will be close to the mean. This yields a corresponding downturn in the structure function at the largest scales.

Another example is if there were a periodic velocity fluctuation, with spatial wavelength  $\lambda$ , then the autocorrelation function would become negative for  $r \approx \lambda/2, 3\lambda/2, \dots$ , which yields periodic peaks in the structure function. Both effects are illustrated in Figure 4b. Neither of them can be captured by our model structure function, which assumes that the autocorrelation is strictly positive and monotonically decreasing with  $r$ . We deal with this limitation by not attempting to model scales larger than  $L/2$ , where the influence of these effects will be greatest.

## 4 STRUCTURE FUNCTION FITS

In order to provide a uniform description of the velocity fluctuations, we fit the model structure function  $\tilde{B}_{\text{mod}}(r)$  of equation (16) to each H II region in our sample. Results are shown in Figures 5, 7, and 8, arranged in order of increasing distance. Best-fit values for the model parameters, together with 95% credibility range, are given in Table 3. After describing the methods used to measure the structure function and fit the model (section 4.1), we give details of the results in the example case of the inner Orion Nebula (4.2) and a broad overview of the results for other regions (4.3). Further details of the results for individual regions are given in Appendix C.

### 4.1 Technical details of structure function measurement and fitting

For the observational values ( $B_{\text{obs}}(r)$ ; blue symbols), the structure function is the average value of equation (6) over all pairs of points

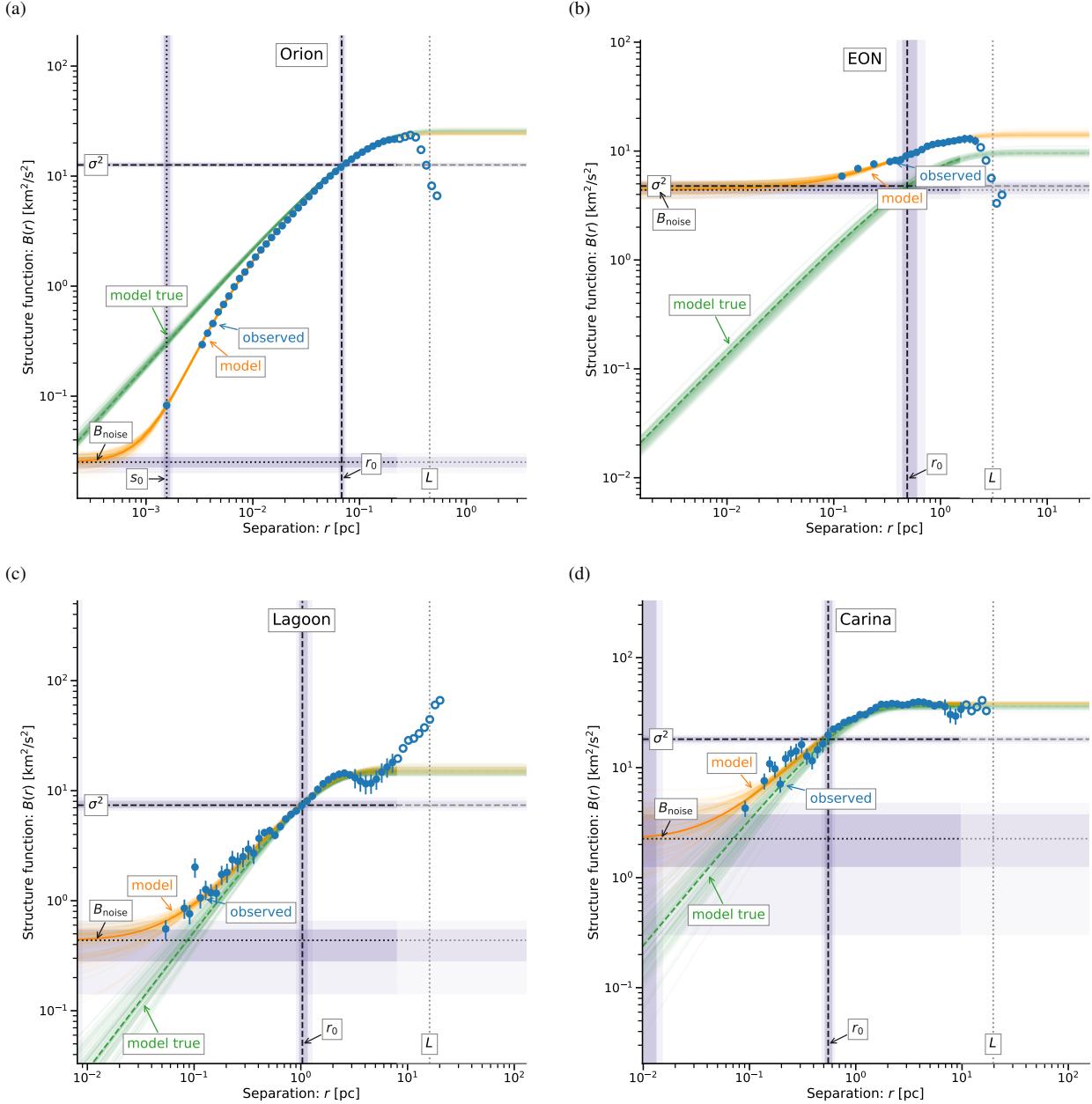
that contribute to each radial bin. We judge the dominant source of error in  $B_{\text{obs}}(r)$  to be systematic, rather than random, so we set the uncertainty of each bin (blue error bars) as a constant fraction (2 to 8%) of the observed value, as listed in column 3 of Table D1 (see supplementary material). In some sources (EON, and the regions observed using the FLAMES and the TAURUS instruments) the uncertainty is increased by a factor of order two for a small number of bins at the smallest and largest scales (columns 4 and 5 of Table D1). For most sources a fixed bin size of 0.05 dex is used, but for those observed with multi-fiber spectroscopy (Carina and M8, see section 2.1.3), where coverage is sparse at small separations, adjacent bins were merged to ensure at least 100 pairs contribute to each bin. Bins with separation less than  $L/2$  are used to constrain the model fitting (filled symbols), while larger separations (open symbols) are not used since they are more likely to be affected by a breakdown of the homogeneity assumption (see section 3.2.3).

Non-linear weighted least square fitting of the model is performed using the Levenberg-Marquardt algorithm (Moré 1978) as implemented in the `lmfit` Python library (Newville et al. 2014). The best-fit model  $\tilde{B}_{\text{mod}}(r)$  is shown (heavy orange solid line), together with the corresponding underlying ‘‘true model’’  $B_{\text{mod}}(r)$  (heavy dashed green line), which does not include the effects of seeing or noise. Note that the true model  $B_{\text{mod}}(r)$  depends on only 3 parameters:  $\sigma^2$ ,  $r_0$ , and  $m$ , whereas  $\tilde{B}_{\text{mod}}(r)$  additionally depends on the observational nuisance parameters:  $s_0$  and  $B_{\text{noise}}$ . The best-fit values of each of these parameters are shown by horizontal and vertical lines in the figure and are summarised in Table 3.

The posterior distributions of model parameters that are consistent with the observations for each region are estimated using Markov Chain Monte Carlo (MCMC) ensemble sampling (Goodman & Weare 2010) as implemented in the `emcee` Python library (Foreman-Mackey et al. 2013). A uniform prior distribution is assumed between upper and lower bounds for each parameter, as given in Table 2. To help ensure convergence of the MCMC algorithm, we used chain lengths that exceeded 50 times the estimated autocorrelation length for each parameter, which typically required of order 50 000 samples. Thin translucent lines in the figures show structure functions using the parameters of a random selection of 100 posterior samples from the MCMC chain, both for  $\tilde{B}_{\text{mod}}(r)$  (orange) and  $B_{\text{mod}}(r)$  (green). This gives an estimate in the uncertainty about the best-fit model. Credibility intervals for the parameters are estimated from percentiles of the posterior distribution and are indicated in the figures by shaded gray boxes around the best-fit parameter values (heavy shading for 68% interval; light shading for 95% interval). The 95% credibility interval for each parameter and for each source is also given in Table 3. Figure 6 gives an example for Orion of the pairwise correlations in the posterior distributions of the model parameters, plotted using the `corner` Python library (Foreman-Mackey 2017). Corresponding plots for the remaining sources and a table for the fitting parameters are given in Appendix D.

### 4.2 Example results of model fitting to the inner Orion Nebula

This dataset is among the highest quality of those in our sample, with more than  $10^5$  spatial points and a factor of more than 500 between the smallest and largest separations. As a result, the observationally derived structure function (Figure 5a) is very smooth and the model fit is very well constrained, as evidenced by the tight credibility limits on the model parameters. The derived correlation length  $r_0 = 0.068$  pc is 7 times smaller than the box size, which indicates that there should be a moderate finite map effect (section 3.2.2 and Appendix A1), but it is well within the range where the model fit can give reliable



**Figure 5.** Second-order structure functions for velocity centroid images of the H $\alpha$  emission line from Galactic H II regions. The observed structure function  $B_{\text{obs}}(r)$  is shown by blue symbols, with filled symbols indicating those points that are used to constrain the model fits. The full model is shown in orange, while the underlying ideal model (without the effects of seeing or noise) is shown in green. See text for details of the fitting process. (a) Inner Orion Nebula. (b) Extended Orion Nebula. (c) M8, the Lagoon Nebula. (d) Carina Nebula.

measurements (see Figure A2). The derived seeing parameter  $s_0 = 0.002$  pc is more than 30 times smaller than  $r_0$ , meaning that the seeing has a negligible effect at the correlation scale and above. Figure 6 shows the covariances between parameters in the model fit, which in some cases show significant correlations. For example,  $\sigma^2$  and  $r_0$  are positively correlated, whereas  $m$  and  $s_0$  are negatively correlated.

#### 4.3 Results of model fitting to other sources

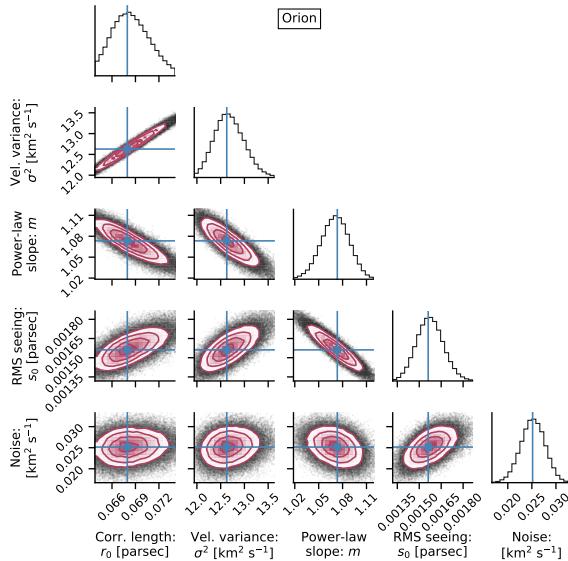
The structure functions for the remaining Galactic sources are shown in Figure 5b–d. The data quality for these sources is not so high as for

the inner Orion Nebula, with the result that the model parameters are not so well constrained. In particular, the relatively coarse spacing between nearest-neighbor spatial points (see sections 2.1.2 and 2.1.3) means that the seeing has no effect on the observed structure function, so that  $s_0$  is indeterminate in the model fits. In addition, the effects of noise are also much greater, as is apparent from visual inspection of the velocity fields (top row of Figure 2). The principal effect of this on the model fits is to increase the uncertainty in the power law slope  $m$ . The most extreme case is the Extended Orion Nebula (Fig. 5b), where noise makes a significant contribution to the structure function at all scales. The resultant degeneracy between parameters means that it is not possible to obtain a satisfactory fit if all parameters are allowed

**Table 3.** Best-fit model parameters and 95% credibility intervals for fits to observed structure functions

Region	$\sigma_{\text{pos}}^2$ [km $^2$ s $^{-2}$ ]	$B_{\text{noise}}$ [km $^2$ s $^{-2}$ ]	$s_0$ (RMS) [pc]	$r_0$ [pc]	$L$ [pc]	$m$ [-]	$L/r_0$ [-]	$s_0/r_0$ [-]	$s_0$ (FWHM) [arcsec]
30 Dor	297 $^{+40}_{-20}$	6.2 $^{+1.6}_{-1.6}$	0.12 $^{+0.06}_{-0.03}$	4.0 $^{+1.0}_{-0.5}$	31	0.85 $^{+0.08}_{-0.13}$	8	0.03	1.2 $^{+0.7}_{-0.4}$
NGC 604	84 $^{+22}_{-10}$	< 1.0	1.6 $^{+1.0}_{-0.6}$	12 $^{+6}_{-5}$	173	0.77 $^{+0.20}_{-0.22}$	14	0.13	1.0 $^{+0.5}_{-0.4}$
NGC 595	53 $^{+5}_{-2}$	< 3.3	< 1.0	11 $^{+1}_{-1}$	196	1.36 $^{+0.10}_{-0.15}$	18	< 0.04	< 0.5
NGC 346	33 $^{+3}_{-2}$	0.8 $^{+0.07}_{-0.07}$	0.08 $^{+0.02}_{-0.02}$	2.4 $^{+0.3}_{-0.2}$	20	0.95 $^{+0.05}_{-0.07}$	8	0.04	0.7 $^{+0.2}_{-0.1}$
Carina	18 $^{+2}_{-2}$	< 4.7	< 0.010	0.6 $^{+0.1}_{-0.1}$	20	1.16 $^{+0.28}_{-0.18}$	33	< 0.01	< 3.4
Hubble X	15 $^{+3}_{-1}$	< 0.8	< 0.55	4.0 $^{+0.5}_{-0.2}$	78	1.02 $^{+0.06}_{-0.23}$	20	< 0.07	< 0.5
Orion	13 $^{+1}_{-1}$	0.030 $^{+0.005}_{-0.005}$	0.0020 $^{+0.0002}_{-0.0002}$	0.068 $^{+0.006}_{-0.004}$	0.5	1.07 $^{+0.03}_{-0.04}$	7	0.02	1.8 $^{+0.2}_{-0.2}$
Hubble V	10 $^{+3}_{-1}$	< 0.65	< 0.67	3.6 $^{+0.5}_{-1.0}$	61	0.81 $^{+0.07}_{-0.28}$	17	< 0.14	< 0.65
Lagoon	7 $^{+1}_{-1}$	0.45 $^{+0.2}_{-0.3}$	< 0.010	1.0 $^{+0.2}_{-0.1}$	16	1.26 $^{+0.20}_{-0.20}$	16	< 0.01	< 3.5
EON	5 $^{+0.6}_{-0.4}$	4.4 $^{+0.9}_{-0.7}$	0 <sup>a</sup>	0.5 $^{+0.2}_{-0.1}$	3	1.0 <sup>a</sup>	6	0 <sup>a</sup>	0 <sup>a</sup>

<sup>a</sup>Assumed value.



**Figure 6.** Example corner plot of covariances between model parameters of fits to the H $\alpha$  structure function. The example shown is for Orion, while figures for the remaining regions are given in Appendix D. Plots on the diagonal show the 1-dimensional histogram of the posterior distribution of each parameter (labeled at bottom), as calculated by the MCMC method, assuming a uniform prior distribution within the limits given in Table 2. Off-diagonal plots show the 2-dimensional histogram of the joint posterior distribution each pair of parameters.

to vary, so we instead chose to fix the power slope at  $m = 1.0$ , which is close to the median value obtained for the other sources.

The structure functions for sources in the Magellanic Clouds are shown in Figure 7 and those in more distant galaxies in Figure 8. The Magellanic cloud sources have generally high data quality due to the large number of independent spatial points in the MUSE observations (section 2.1.4). The relatively poor spectral resolution means that the  $B_{\text{obs}}(r)$  cannot be measured down to such low values as in the inner Orion Nebula, but the larger amplitude of the fluctuations in

these sources means that the fitted model parameters are nevertheless tightly constrained.

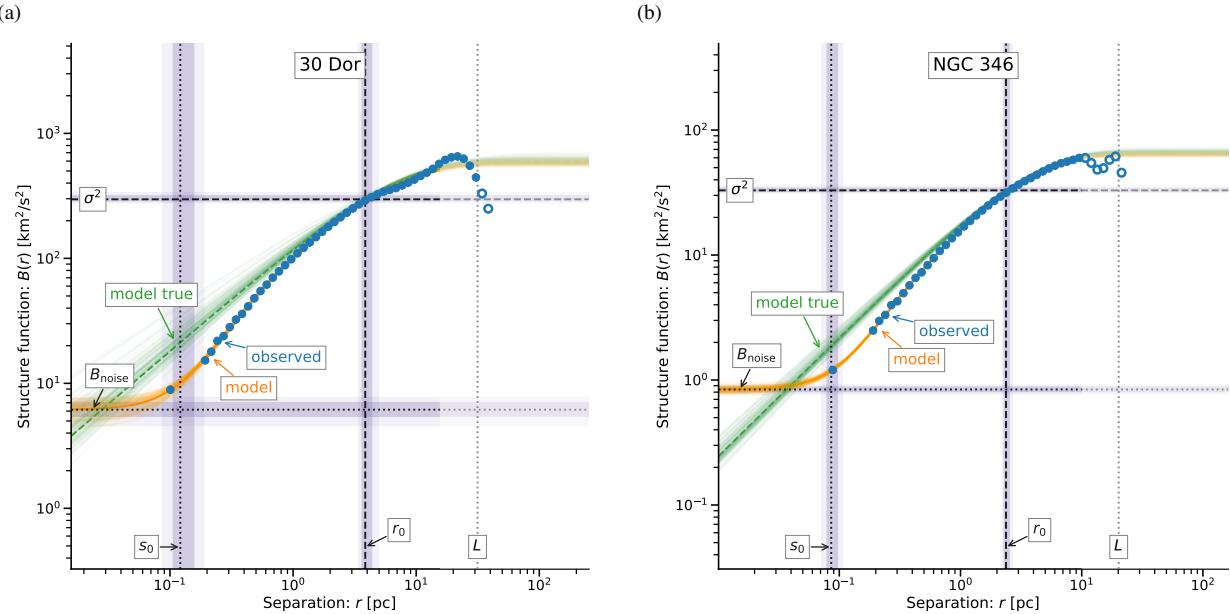
For the more distant sources, the principal observational limitation is the seeing. In three of the four sources (NGC 604, Hubble V and Hubble X), the derived values of  $s_0$  exceed 10% of the correlation length. This implies that the true power law slope  $m$  is less steep than would be naively inferred from the observations (compare the green and orange curves in Figure 8), but at the same time the degeneracies between parameters lead to large uncertainties in the determination  $m$ . On the other hand, the values of  $r_0$  and  $\sigma^2$  are still well-constrained.

#### 4.4 On the reasonableness of the derived seeing widths and noise

In the model fitting we allow the nuisance observational parameters  $s_0$  and  $B_{\text{noise}}$  to vary freely within a wide range (Table 2). This is necessary because of the heterogeneous nature of our source datasets and the fact that the details of the observational conditions are not available to us in all cases. However, it is worthwhile to perform a sanity check on the values that are implied by our model fits. To that end, the last column in Table 3 lists the FWHM seeing width for each fit in units of arcsec. Disregarding the 5 sources where only upper limits can be determined, the mean and standard deviation are  $(1.17 \pm 0.20)$  arcsec, which is perfectly consistent with expectations for seeing from ground-based observations.

However, there are a few outlier values that deserve closer attention. NGC 595, Hubble V, and Hubble X all have model-derived upper limits to the seeing FWHM of about 0.6 arcsec, which seems unrealistically small, especially given the higher value derived for NGC 604, which was observed with the same instrument. Unfortunately, we do not have access to the observing logs to check this, so we suggest extra caution in interpreting the fit results for these objects.

The inner Orion Nebula has the largest inferred seeing FWHM of  $(1.8 \pm 0.2)$  arcsec, which is nearly twice as high as the typical seeing measured during the observations (Doi et al. 2004). However, this can be at least partially explained by the fact that the velocity map is constructed by interpolating individual longslit observations onto a



**Figure 7.** Same as figure 5 except for Magellanic Cloud H II regions. (a) 30 Doradus in the LMC. (b) NGC 346 in the SMC.

regular grid (section 2.1.1). The seeing-limited resolution will only be achieved along the slits, whereas the resolution in the perpendicular direction is determined by the slit spacing of 2 arcsec. Given this, the model-derived value of  $s_0$  for this source is not unreasonable.

On the other hand, it is perhaps significant that both the inner Orion Nebula and 30 Doradus, the two highest quality datasets in our sample, should both have a fitting-derived seeing width that is slightly larger than expected. Since our intrinsic model  $B_{\text{mod}}(r)$  assumes a single power law for scales  $< r_0$ , the only way of accommodating a steepening of  $B_{\text{obs}}(r)$  at small separations is via the seeing term  $S(r)$ . But if the true  $B(r)$  really does steepen at small scales, then the seeing will be overestimated in the models. Unfortunately, the currently available data are insufficient to definitively decide this question.

For most of our sources  $B_{\text{noise}}/\sigma_{\text{pos}}^2 < 0.03$  and the noise has almost no effect on the structure function, apart from at the very smallest separations, resulting in a negligible influence on the other model parameters. The exceptions are the Galactic regions Lagoon, Carina, and especially the Extended Orion Nebula, for which the noise is sufficiently large that the slope of  $B_{\text{obs}}(r)$  is significantly shallower than the inferred slope  $m$  of the true model  $B_{\text{mod}}(r)$ . These sources also show the largest fluctuations  $B_{\text{obs}}(r)$  around the smooth fit at intermediate scales (see in particular Figure 5d), which is probably due to the relatively small number of spectra.

#### 4.5 Evidence for inhomogeneity at the largest scales

At scales larger than  $L/2$  we see a variety of different shapes for the structure functions of our sample regions. These points are excluded from our model fits but we give a qualitative description in this section. In some cases, such as Carina and Hubble X,  $B_{\text{obs}}(r)$  remains flat at a value of  $\approx 2\sigma_{\text{pos}}^2$ , which is consistent with uncorrelated homogeneous fluctuations at the largest scales.

For other regions, such as the Orion Nebula (both inner and outer) and 30 Doradus, there is a clear downturn in  $B_{\text{obs}}(r)$  at the largest separations. As discussed above in section 3.2.3, this is what would

be expected if the velocity fluctuations were inhomogeneous, with a larger amplitude in the center of the map and a smaller amplitude in the outskirts. A similar behavior, although not so marked, is shown by Hubble V and NGC 595. The Orion Nebula is a particularly interesting case since the  $\sigma_{\text{pos}}^2$  derived from the high-resolution dataset of the inner nebula on scales  $\approx 0.1$  pc is roughly twice as large as that derived from the lower resolution dataset of the Extended Orion Nebula on scales  $\approx 1$  pc (compare panels a and b of Figure 5). This is further evidence that the amplitude of velocity fluctuations increases towards the center of the nebula in this source.

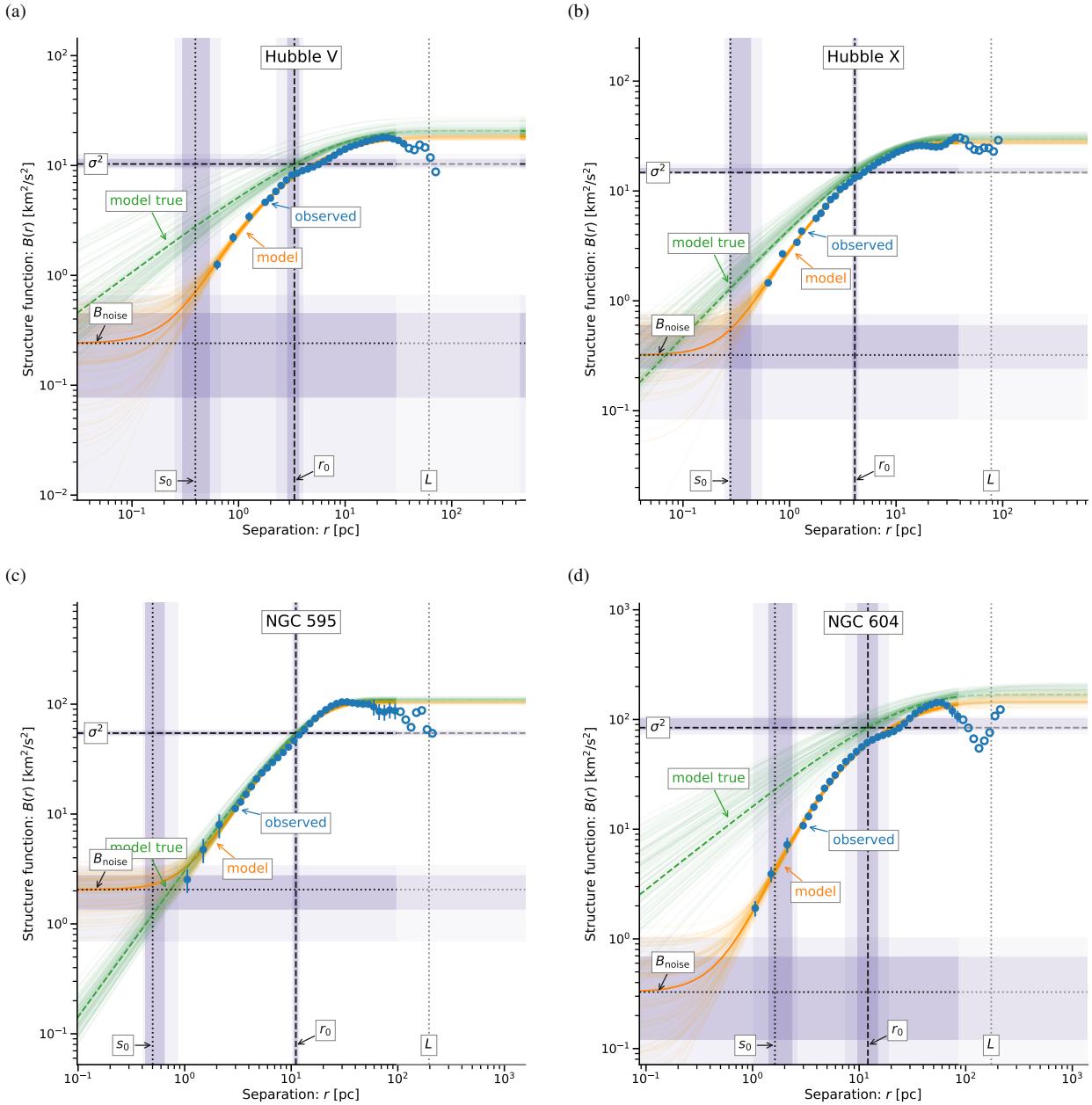
In contrast, other regions, such as the Lagoon and NGC 346, show  $B_{\text{obs}}(r)$  increasing at the largest separations, reaching values significantly larger than  $2\sigma_{\text{pos}}^2$ . This can be evidence for a large-scale linear velocity gradient across the region (see Figure 13 of Arthur et al. 2016) or a periodic fluctuation with  $\lambda > r_0$  (see section 3.2.3 above). This latter effect may also explain the behavior of NGC 604, which shows oscillatory behavior of  $B_{\text{obs}}(r)$ , similar to the light blue line in Figure 4b.

## 5 DISCUSSION

### 5.1 Comparison with previous structure functions

Comparison with prior studies is complicated by the diversity of methodologies that have been employed. There is almost universal agreement on how to measure the velocity variance  $\sigma_{\text{pos}}^2$ , although even here there can be small differences according to whether the centroid velocities come from gaussian fitting or velocity moments, and whether or not smooth large-scale trends are first removed. Also, the field of view of the region would have an influence on  $\sigma_{\text{pos}}^2$ . Arthur et al. (2016) compared multiple studies of the inner Orion Nebula, finding agreement to within 20% in  $\sigma_{\text{pos}}^2$  determinations for emission lines that trace the bulk of the fully ionized gas.

There is less agreement in the literature on the best way define the correlation length,  $r_0$ . Our own definition corresponds to the lag where the autocorrelation function has fallen to a value of one-half,



**Figure 8.** Same as figure 5 except for H II regions in more distant Local Group Galaxies. (a) Hubble V in NGC 6822. (b) Hubble X in NGC 6822. (c) NGC 595 in M33. (d) NGC 604 in M33.

$C(r_0) = 1/2$  (see section 3.2), whereas other common choices are  $C(r) = e^{-r}$  (Miville-Deschénes et al. 1995) or the *total decorrelation lag*  $\tau_0$  (Lagrois & Joncas 2011), which corresponds to  $C(\tau_0) = 0$ . One disadvantage of this latter definition is that it can be sensitive to non-homogeneous behavior of the velocity field at large scales. The conversion between these different conventions will also depend on the value of the power-law slope, but we typically find that  $\tau_0/r_0 = 2$  to 3.

The power law slope,  $m$ , itself is probably the parameter that is most sensitive to methodological assumptions. The fundamental issue is that the observed structure function tends to show a negative curvature in log-log space at intermediate scales:  $d^2 \log B / d^2 \log r < 0$ . This is caused primarily by the steepening due to seeing at small scales together with the natural asymptotic flattening,  $B(r) \rightarrow 2\sigma_{\text{pos}}^2$ .

at large scales. As a result, the measured slope depends on the exact range of separations over which the power law is fitted. So long as the spatial dynamic range of the observations is high enough, then a reliable slope can nevertheless be obtained, as shown by Arthur et al. (2016) for the case of the Orion Nebula. However, this requires that  $r_0$  should be at least an order of magnitude higher than the larger of the seeing width or the minimum separation between spatial points. This condition is not satisfied for roughly half of our sample regions, and the same is true for many previous studies. In such cases, our model-based approach is a more reliable way of determining the slope.

Specific comparisons for individual sources are given in Appendix C. In summary, we agree with previous studies in cases where there is a good match in the exact area covered and in the angular

resolution. In other cases, large discrepancies are found, which are probably caused by a combination of methodological differences and real variations due to the inhomogeneity of the velocity fluctuations on the largest scales.

## 5.2 Scaling relations for turbulence parameters

We look for correlations and scaling relations between the physical parameters of the H II regions and the parameters of the fluctuating velocity fields that we derived in section 4. To accomplish this we implement a hierarchical Bayesian model for linear regression using the `linmix` Python library, which is based on the procedure of Kelly (2007). We fit relations of the form  $Y = aX + b$  where  $X = \log_{10} x$  and  $Y = \log_{10} y$  with  $x, y \in \{d, D_{\text{HII}}, L(\text{H}\alpha), \sigma_{\text{pos}}, r_0, m\}$  (taken from the physical properties in Table 1 and the derived turbulent parameters of Table 3). Thus the fitted relation between the original variables is of the power-law form:  $y = 10^b x^a$ . Note that, unlike with an Ordinary Least Squares method, the Kelly method uses an errors-in-variables model that allows for measurement uncertainty in both  $X$  and  $Y$ . Since the main focus of this paper is the turbulent velocity field, we restrict consideration to correlations that include one of the three structure function parameters,  $\sigma_{\text{pos}}$ ,  $r_0$ , or  $m$ , as the  $y$  variable. In addition, we fit a strictly linear relationship between the plane-of-sky and line-of-sight velocity dispersions:  $X = \sigma_{\text{pos}}$ ,  $Y = \sigma_{\text{los}}$ . The reason for not taking the logarithm in this instance is both to facilitate comparison with previous studies and to allow for a zero-point offset in the relation.

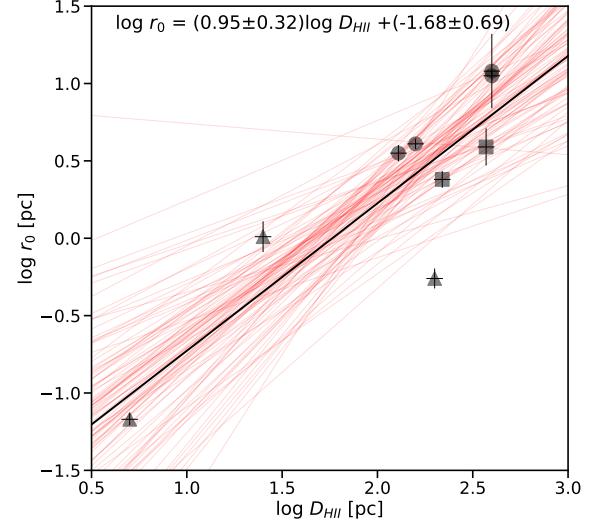
In the following analysis we omit the EON region since the model fits were relatively poorly constrained by the observations. For physical properties where uncertainties are not reported (size and luminosity), we assume an error of 10% of the value. The same criteria is used for the uncertainty values of the mean non-thermal linewidths for 30 Doradus and NGC 346. For the turbulent parameters we use the relative error from our model fits, considering the  $2\sigma$  interval.

Results for the most significant correlations, as measured by the Pearson correlation coefficient  $r$ , are given in Table 4 in descending order of  $r^2$ . We find three cases of highly significant correlations with significance level  $p < 0.01$ , which are illustrated in Figures 9–11:  $r_0$  versus  $D_{\text{HII}}$ ,  $\sigma_{\text{los}}$  versus  $\sigma_{\text{pos}}$ , and  $\sigma_{\text{pos}}$  versus  $L(\text{H}\alpha)$ . Each of these is discussed in turn in the following sections. The remaining correlations are at best marginally significant ( $p > 0.05$ ), including all correlations of the power law slope  $m$  with any other parameter, but this is itself interesting and is also given its own section below.

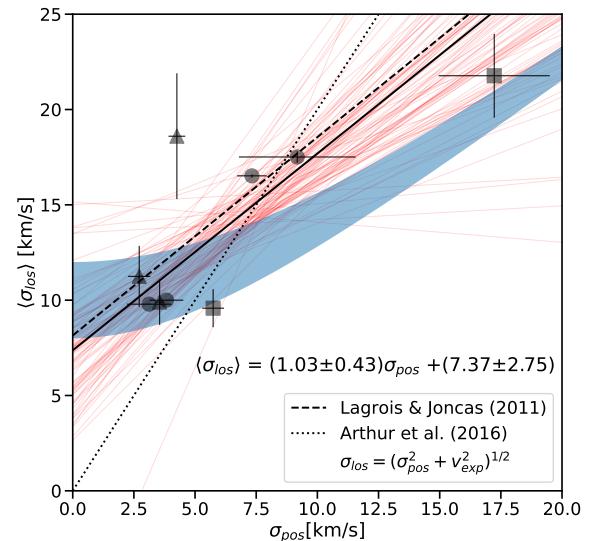
### 5.2.1 Diameter vs correlation length

Figure 9 shows the relation between the correlation length  $r_0$  of the velocity fluctuations and the diameter  $D_{\text{HII}}$  of the H II region. The relationship we find is  $\log r_0 = (0.95 \pm 0.32) \log D_{\text{HII}} - (1.68 \pm 0.69)$ , which is very close to linear, implying that the correlation length is a constant fraction of the diameter:  $r_0/D_{\text{HII}} = 0.02 \pm 0.01$ .

However,  $r_0$  is also well correlated with the observed map size,  $L$ , so it is important to check that the correlation lengths that we measure are true properties of the regions, and are not just artifacts of the finite map sizes. In Appendix A1 we show that  $L$  has a significant effect on the naive apparent  $r_0$  when  $r_0 > 0.1L$ , but that an accurate  $r_0$  can still be recovered from model fitting so long as  $r_0 < 0.20L$ . From our results (Table 3) we find that  $r_0 \approx (0.08 \pm 0.01)L$ , which is comfortably small enough that the measured correlation lengths should be reliable. A clear indication that the derived  $r_0$  is not real would be a structure function that never flattens off, but rather keeps



**Figure 9.** The relationship between correlation length of velocity fluctuations,  $\log_{10} r_0$ , and H II region diameter,  $\log_{10} D_{\text{HII}}$  derived from our results. Symbols indicate Galactic (triangles), Magellanic (squares) and extragalactic (circles) regions.



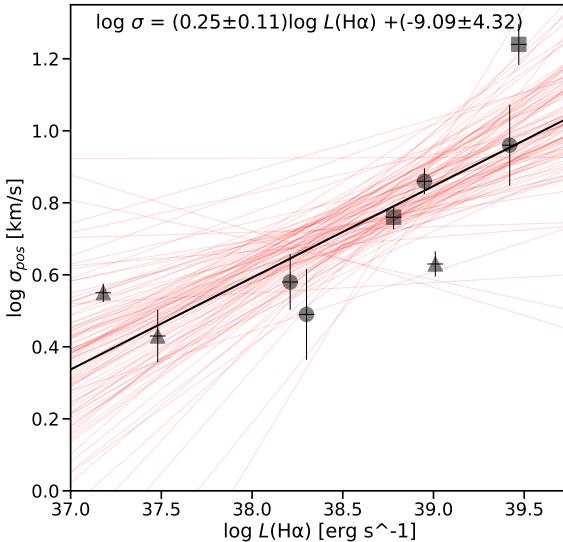
**Figure 10.** The relationship between line-of-sight non-thermal velocity dispersion,  $\sigma_{\text{los}}$ , and plane-of-sky dispersion of velocity centroids,  $\sigma_{\text{pos}}$  derived from our results. Symbols indicate Galactic (triangles), Magellanic (squares) and extragalactic (circles) regions. The black solid line is the median linear regression, with thin red lines showing 100 random samples from the posterior distribution to indicate the uncertainty in the regression. The dashed line shows the linear fit obtained by Lagrois et al. (2011) for a sample of small-to-intermediate size H II regions in the M33 galaxy.

rising at all scales. We do not see such a behavior in any of our regions. The most similar is the Lagoon Nebula (Figure 5c), where the structure function does flatten at intermediate scales, but then rises again for separations of order  $L$ . It may be that there is a second, outer correlation scale in this region, which we fail to measure because the observed map is too small.

The small value of  $r_0/D_{\text{HII}}$  that we obtain means that it would be feasible to study the spatial variation of the turbulence parameters within a given region. This is because parts of the nebula that are sep-

**Table 4.** Linear regressions values in the form  $Y = aX + b$  between our turbulent parameters obtained using the chi-square statistic and properties of each region (Table 1). The fifth column,  $r$ , is the Pearson correlation coefficient and the last column is the  $p$ -value. This results were obtained using the procedure in Kelly (2007).

$Y$	$X$	$a$	$b$	$r$	$p$
$\log r_0$	$\log D_{\text{HII}}$	$0.95 \pm 0.32$	$-1.68 \pm 0.69$	0.86	<b>0.003</b>
$\log \sigma_{\text{pos}}$	$\log L(\text{H}\alpha)$	$0.25 \pm 0.11$	$-9.09 \pm 4.32$	0.81	<b>0.008</b>
$\sigma_{\text{los}}$	$\sigma_{\text{pos}}$	$1.03 \pm 0.43$	$7.37 \pm 2.75$	0.78	<b>0.010</b>
$\log \sigma_{\text{pos}}$	$\log D_{\text{HII}}$	$0.26 \pm 0.18$	$0.19 \pm 0.39$	0.64	0.06
$\log \sigma_{\text{pos}}$	$\log r_0$	$0.18 \pm 0.19$	$0.67 \pm 0.14$	0.50	0.17
$\log m$	$\log d$	$-0.02 \pm 0.04$	$0.04 \pm 0.08$	-0.41	0.27
$\log m$	$\log \sigma_{\text{pos}}$	$-0.10 \pm 0.20$	$0.08 \pm 0.15$	-0.37	0.33
$\log m$	$\log r_0$	$-0.02 \pm 0.07$	$0.01 \pm 0.05$	-0.29	0.45



**Figure 11.** The relationship between plane-of-sky dispersion of velocity centroids,  $\log_{10} \sigma_{\text{pos}}$ , and  $\text{H II}$  region luminosity,  $\log_{10} L(\text{H}\alpha)$ , derived from our results. Symbols indicate Galactic (triangles), Magellanic (squares) and extragalactic (circles) regions.

arated by a distance  $\gg r_0$  are essentially independent. In principle, the variation of  $\sigma_{\text{pos}}$  could be measured in ‘‘macro-pixels’’ of size  $2r_0$ , but a minimum size of  $3r_0$  is necessary to reliably determine  $r_0$  itself (Appendix A1), and it is quite possible that  $r_0$  might vary across the face of the nebulae. The best that could be done with the current  $L \times L$  velocity maps would therefore be about  $3 \times 3$  pixels, but this could be increased to about  $15 \times 15$  pixels if the entire  $D_{\text{HII}} \times D_{\text{HII}}$  region were mapped.

The energy injection or driving scale of turbulence,  $\ell_E$ , is commonly identified with the ‘‘outer scale’’ of the fluctuations (Havercorn et al. 2004; Chepurnov & Lazarian 2010). The outer scale has been given diverse definitions (Klipp 2014) but most commonly is either the integral scale  $\mathcal{L}$ , or the scale at which the structure function becomes flat. These are both of the same order as the correlation length  $r_0$  as we define it in section 3.2, but are typically larger by a factor of 1.5 to 2 (for example, equation (14) for the integral scale). However, high resolution numerical simulations of driven isothermal turbulence (Federrath et al. 2021) show a somewhat larger gap of a factor of 4 between the correlation length and the driving scale. In either case, our results imply that the turbulence in  $\text{H II}$  regions is driven on relatively small scales, between 2% and 5% of the diameter of the  $\text{H II}$  region.

### 5.2.2 The constancy of the power law slope

The slope  $m$  of the power-law portion of the structure function shows no significant correlation with any other parameter (see Table 4), with a weighted mean value of  $m = 1.03$ . The dispersion in fitted  $m$  values is 1.6 times higher than expected on the basis of the confidence limits for the individual regions, implying that there may be a real intrinsic dispersion of  $\pm 0.17$  in the slopes. However, a  $\chi^2$  test shows that this is only marginally significant ( $p = 0.11$ ), so we cannot rule out that all the slopes are identical.

For incompressible Kolmogorov turbulence, the predicted slope of the three-dimensional structure function is  $m_{3D} = 2/3$ , but intermittency and compressibility (Schmidt et al. 2008) can increase this to  $m_{3D} \approx 0.8$  in the subsonic regime, and up to  $m_{3D} = 1$  in the supersonic regime. Comparison with our measurements is complicated by projection smoothing (Scalo 1984), which implies that  $m_{3D}$  is only observed directly for separations larger than the line-of-sight depth of the emitting regions. For smaller separations, one potentially observes a steeper slope:  $m = m_{3D} + 1 + \delta\kappa$ , where  $\delta\kappa$  represents the effects of emissivity fluctuations<sup>7</sup> along the line of sight (Brunt & Mac Low 2004), and varies from  $\delta\kappa = 0$  in the incompressible limit to  $\delta\kappa = -1$  in the case of strongly driven supersonic turbulence. In this latter case, the effects of projection smoothing and emissivity fluctuations cancel out, leading once more to  $m \approx m_{3D}$ .

The  $\text{H II}$  regions of our sample span the transonic regime, from subsonic velocity fluctuations in the Galactic sources to supersonic fluctuations in the more luminous extragalactic sources. It is therefore not surprising that we see no significant correlation between  $m$  and  $\sigma_{\text{pos}}^2$  ( $p = 0.33$ , Table 4). Although theoretically one might expect a positive correlation of  $m_{3D}$  with the amplitude of the velocity fluctuations due to the transition from subsonic to supersonic turbulence (Galtier & Banerjee 2011), this will be offset by a decrease in  $\delta\kappa$  if the amplitude of the density fluctuations (and hence emissivity fluctuations) increases along with the velocity fluctuations.

Evidence for just such an increase is shown in Figure 12, which shows how the fractional width of the probability distribution of large-scale  $\text{H}\alpha$  surface brightness fluctuations varies with the RMS turbulent Mach number,  $M$ , of the three-dimensional velocity fluctuations. The Mach number is estimated by assuming that the plane-

<sup>7</sup> Note that literature on turbulence in molecular clouds discuss fluctuations of density since they are mainly concerned with emission lines whose emissivity is approximately linearly proportional to density. In the case of photoionized regions, the emissivity of important lines such as  $\text{H}\alpha$  is proportional to  $\text{density-squared}$ , so in adopting the results of Brunt & Mac Low (2004) we substitute emission measure (or its proxy,  $\text{H}\alpha$  surface brightness) in place of column density.

of-sky velocity dispersion is a good estimate of the turbulent fluctuations in one dimension and that the turbulence is isotropic, yielding:  $\mathcal{M} = \sqrt{3}\sigma_{\text{pos}}/c_s$ , where an ionized sound speed of  $c_s = 11 \text{ km s}^{-1}$  is assumed for all sources. The RMS fractional width of density fluctuations in a turbulent medium,  $\sigma_\rho/\rho_0 \equiv \langle \delta\rho^2 \rangle^{1/2}/\rho_0$ , is predicted to depend linearly on the Mach number as  $\sigma_\rho/\rho_0 = b\mathcal{M}$ , where  $b$  depends on the nature of the turbulent driving (Federrath et al. 2010), varying between  $b = 1/3$  for solenoidal driving and  $b = 1$  for compressive driving. This is shown by gray shading<sup>8</sup> in Figure 12h, where it can be seen that many of the sources with subsonic velocity dispersions ( $\mathcal{M} < 1$ ) fall to the right of the theoretical prediction, implying that other factors in addition to turbulence are responsible for producing the fluctuations in density and surface brightness in these regions (see also section 4.5 of Arthur et al. 2016). For our highest luminosity source of 30 Dor on the other hand, the velocity dispersion is clearly supersonic ( $\mathcal{M} \approx 2.7$ ) and, although the surface brightness fluctuations are also larger than in the other sources, they are consistent with being caused by turbulence-induced density fluctuations.

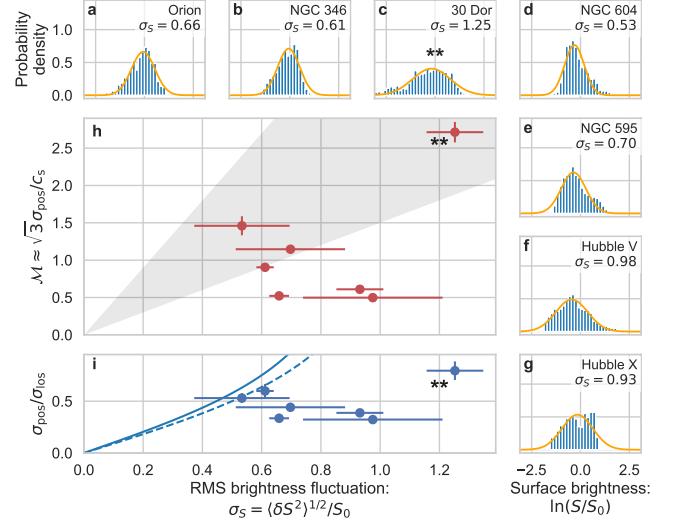
We find no clear evidence for a broken power-law in any of our structure functions, as has been claimed in previous studies. Although in some regions we do observe a steepening at the smallest scales (for example Figures 5a, 7ab), this can be accounted for in our model fits by the spatial smoothing caused by atmospheric seeing. This has a noticeable effect on the structure function at scales of up to 5 times the FWHM of the seeing (Appendix A2). There is a hint in the highest quality datasets that there may also be a real steepening of the structure function at small scales (see section 4.4), but higher spatial resolution observations are required in order to study this.

### 5.2.3 Plane-of-sky vs line-of-sight velocity dispersions

In Figure 10 we show the mean non-thermal line-of-sight velocity dispersion  $\langle \sigma_{\text{los}} \rangle$  (obtained from the spectral line profiles) versus the plane-of-sky dispersion of mean velocities  $\sigma_{\text{pos}}$  (obtained from our structure function fits). The relationship we obtain is  $\langle \sigma_{\text{los}} \rangle = (1.03 \pm 0.43)\sigma_{\text{pos}} + (7.4 \pm 2.8)$ , which can be compared with previous results using different types of samples. Lagrois et al. (2011) surveyed a sample of small-to-intermediate size H II regions in M33, finding  $\langle \sigma_{\text{los}} \rangle = (1.04 \pm 0.14)\sigma_{\text{pos}} + (8.15 \pm 0.93)$  for their sub-sample with higher resolution observations (21 sources). This is shown as the dashed line in Figure 10, which can be seen to be almost identical to our own result. Both fits are consistent with a slope of unity and an offset of  $8 \text{ km s}^{-1}$ . Arthur et al. (2016) calculate the relationship between  $\langle \sigma_{\text{los}} \rangle$  and  $\sigma_{\text{pos}}$  for different optical emission lines ([S II], [N II], H $\alpha$ , [O III]) in the Orion Nebula, finding  $\langle \sigma_{\text{los}} \rangle \approx 2\sigma_{\text{pos}}$ , which is shown by the dotted line in Figure 10. Although this line is a fair representation of the low- $\sigma$  points in our sample, it fails badly for highest dispersion source, 30 Doradus.

At least three different explanations for the relation between  $\sigma_{\text{pos}}$

<sup>8</sup> In the figure, it is assumed that the relative fluctuations in surface brightness and density are equal, which is a result of the cancellation between two effects. First, that the volumetric H $\alpha$  emissivity  $E$  is proportional to density squared, so that  $\sigma_E/E_0 = 2\sigma_\rho/\rho_0$ . Second, that fluctuations in surface brightness  $S$  are related to those in emissivity by  $\sigma_S/S_0 = R^{1/2}\sigma_E/E_0$ , where  $R$  is the 2D-to-3D variance ratio (Brunt et al. 2010). For an emissivity power spectrum  $P(k) \sim k^{-3}$  and assuming a ratio of map size to correlation length of  $L/r_0 = 10$  (typical value from Table 3), we find  $R \approx 0.2$  and hence  $\sigma_S/S_0 \approx \sigma_\rho/\rho_0$ . We calculate  $\sigma_S/S_0$  by fitting a log-normal function to the probability distribution function of the surface brightness map after filtering out fluctuations on scales smaller than the velocity correlation length.



**Figure 12.** Panels a–g: Histograms (blue) of the probability density function (PDF) of the H $\alpha$  surface brightness of H II regions in our sample. Smooth lines (orange) show fitted log-normal distributions. In order to suppress fluctuations on scales smaller than the correlation length, surface brightness maps were rebinned by averaging into pixels of size  $0.5r_0$  before calculating the PDF. Panel h: RMS Mach number of three-dimensional velocity fluctuations as a function of the RMS fractional brightness fluctuation calculated from the PDFs. The gray shaded area shows the range of predicted values for the brightness fluctuations if they are entirely due to density fluctuations formed by driven turbulence with the observed Mach number. Panel i: Ratio of plane-of-sky to line-of-sight velocity dispersions as a function of the RMS fractional brightness fluctuations calculated from the PDFs. The solid and dashed lines show the expected relationship for the toy models of Appendix B.

and  $\langle \sigma_{\text{los}} \rangle$  have been proposed previously. The classic interpretation in terms of homogeneous incompressible turbulence (Von Hoerner 1951) is that  $\sigma_{\text{pos}}$  is reduced with respect to the “true” internal velocity dispersion due to averaging over multiple independent turbulent cells along the line of sight, which is another manifestation of projection smearing, as discussed above in section 5.2.2. Assuming that  $\langle \sigma_{\text{los}} \rangle$  measures the true dispersion, then one would expect to find  $\sigma_{\text{pos}}/\langle \sigma_{\text{los}} \rangle = (\mathcal{L}/H)^{1/2}$ , where  $H$  is the line-of-sight depth of the emitting gas and  $\mathcal{L}$  is the integral length scale of the fluctuations defined in section 3.2, which yields  $H/r_0 \approx 1.44(\langle \sigma_{\text{los}} \rangle/\sigma_{\text{pos}})^2$ . From the results of Figure 10 we would therefore find  $H/r_0 \approx 6$  for the lower dispersion sources, falling to  $H/r_0 \approx 2$  for the higher dispersion sources. However, such an analysis ignores the effect of density fluctuations, which would tend to reduce the effectiveness of projection smoothing compared with the incompressible case (see discussion of  $\delta\kappa$  above), so it is unclear if the results remain valid for the sources with supersonic velocity dispersion.

An alternative interpretation, as discussed in Lagrois et al. (2011), is that  $\langle \sigma_{\text{los}} \rangle$  includes a contribution due to ordered large-scale velocity gradients (for instance, radial expansion) in addition to the turbulence, whereas  $\sigma_{\text{pos}}$  includes only the turbulent motions. Assuming a constant value  $v_{\text{exp}}$  for the ordered contribution and that projection smoothing is negligible, one would predict  $\langle \sigma_{\text{los}} \rangle = (\sigma_{\text{pos}}^2 + v_{\text{exp}}^2)^{1/2}$ , which is shown by the blue shaded-area in Figure 10 with an upper limit of  $12 \text{ km s}^{-1}$  and lower limit of  $8 \text{ km s}^{-1}$  for  $v_{\text{exp}}$ . It can be seen that such a relationship reproduces the observationally derived values almost as well as the linear fit. On the other hand, we can rule out the possibility that  $v_{\text{exp}}$  might increase in proportion to the turbu-

lent velocity dispersion since that would predict that  $\langle \sigma_{\text{los}} \rangle \propto \sigma_{\text{pos}}$ , which is inconsistent with the observations as discussed above.

However, as pointed out by Arthur et al. (2016), the previous analysis is oversimplified since the presence of density fluctuations (or, more generally, emissivity fluctuations) means that the plane-of-sky dispersion of centroids  $\sigma_{\text{pos}}$  is not independent of the ordered large-scale velocity gradients along the line of sight. Indeed, the neglect of any projection smoothing for the turbulent contribution in the Lagrois et al. analysis is equivalent to assuming  $\delta\kappa = -1$ , which requires substantial density fluctuations. This leads to a third potential explanation of the  $\sigma_{\text{pos}}/\langle \sigma_{\text{los}} \rangle$  relationship. In the limit that the direct contribution of turbulent velocity fluctuations is negligible and that  $\langle \sigma_{\text{los}} \rangle$  is entirely due to ordered expansion along the line of sight, then a non-zero  $\sigma_{\text{pos}}$  can arise only from emissivity fluctuations, which cause different parts of the ordered velocity field to be picked out for different lines of sight. A simple model of this process is explored in Appendix B, where a relation between the surface brightness dispersion and the apparent velocity dispersions is derived. This is compared with the behavior of the ratio  $\sigma_{\text{pos}}/\langle \sigma_{\text{los}} \rangle$  for our observed sample in panel i of Figure 12, in which the solid and dashed lines show results for different assumptions about the effects of velocity gradients on  $\langle \sigma_{\text{los}} \rangle$ . Although the model does qualitatively reproduce the increase in  $\sigma_{\text{pos}}/\langle \sigma_{\text{los}} \rangle$  with increasing  $\sigma_S$ , the quantitative agreement is poor, with the model substantially overestimating the magnitude of plane-of-sky fluctuations as compared with the line-of-sight widths. This is evidence that the combination of a velocity gradient with emissivity fluctuations is insufficient on its own, and that we require an additional disordered velocity component in order to explain our observations.

In summary, our results are most consistent with a combination of expansion and disordered motions, in which the disordered component becomes increasingly dominant for the sources with higher velocity dispersion. We cannot rule out a role for either projection smoothing or emissivity fluctuations, but the contribution is probably minor in both cases.

#### 5.2.4 Luminosity vs centroid velocity dispersion

Figure 11 shows the observed relation between H $\alpha$  luminosity and plane-of-sky dispersion of velocity centroids for our sample, for which we obtain the power-law fit  $\log \sigma_{\text{pos}} = (0.25 \pm 0.11) \log L(\text{H}\alpha) - (9.1 \pm 4.3)$ . It is common to study the  $L-\sigma$  relationship, not only in individual H II regions, but also in H II galaxies, both in the local universe and at high redshift (Terlevich & Melnick 1981; Chávez et al. 2014). In such studies, the velocity dispersion is usually derived from the line width of the entire region ( $\sigma_{\text{los}}$ ) since the spatially resolved observations necessary for measuring  $\sigma_{\text{pos}}$  are not available.<sup>9</sup> A reliable empirical  $L-\sigma$  relationship allows the use of bright photoionized regions as standard candles in order to constrain cosmological parameters (Chávez et al. 2012; Wu et al. 2020; González-Morán et al. 2021). The luminosities of the regions used in these cosmological studies range from that of our own brightest source up to 100 times more luminous ( $L(\text{H}\alpha) = 10^{40}$  to  $10^{42} \text{ erg s}^{-1}$ ), although other studies cover intermediate luminosity

<sup>9</sup> The linewidth  $\sigma_{\text{los}}$  for the integrated emission of an H II region will be larger than the mean of the spatially resolved linewidths  $\langle \sigma_{\text{los}} \rangle$  since it also includes the plane-of-sky fluctuations  $\sigma_{\text{pos}}$ . Given the results of section 5.2.3, and assuming that the linear relationship of Figure 10 can be extrapolated to higher values, then the approximate relation is  $\sigma_{\text{los}} = \sqrt{2} \langle \sigma_{\text{los}} \rangle$  for  $\sigma_{\text{los}} \gg 10 \text{ km s}^{-1}$ .

regions that bridge the two ranges (Moiseev & Lozinskaya 2012; Yu et al. 2019).

Expressed in the customary form  $L \sim \sigma^\alpha$ , our result corresponds to  $\alpha = 4.0 \pm 1.6$ , which is smaller than the value of  $\alpha \approx 5$  that is typically found for more luminous regions (Moiseev et al. 2015; Wu et al. 2020) but not significantly so, given the uncertainty. Studies that rely on  $\langle \sigma_{\text{los}} \rangle$  generally find that  $\sigma$  becomes insensitive to luminosity below about  $L(\text{H}\alpha) = 10^{40} \text{ erg s}^{-1}$ , reaching a constant value of  $\sigma = 15 \text{ km s}^{-1}$  (Moiseev et al. 2015; Yu et al. 2019). On the other hand, by using  $\sigma_{\text{pos}}$  instead of  $\langle \sigma_{\text{los}} \rangle$ , we succeed in extending the same power-law relation down to  $L(\text{H}\alpha) = 2 \times 10^{37} \text{ erg s}^{-1}$  and  $\sigma = 3 \text{ km s}^{-1}$ . Given our analysis of the previous section, one possible explanation for this is that  $\sigma_{\text{pos}}$  is a better measure of the disordered turbulent motions within the H II regions than  $\langle \sigma_{\text{los}} \rangle$  since the latter is contaminated by large-scale ordered motions, which are always at least as large as the ionized sound speed.

## 6 CONCLUSIONS

We have used the second-order structure function of velocity centroids on the plane of the sky to carry out a systematic study of the turbulent kinematics in a sample of 9 H II regions spanning a wide range of luminosities and sizes. Our principal findings are as follows:

1. The velocity fluctuations range from subsonic (Mach number  $\mathcal{M} \approx 0.5$ ) in the case of low-luminosity Galactic regions, such as the Orion Nebula, up to supersonic ( $\mathcal{M} \approx 2.5$ ) in the case of giant extragalactic regions such as 30 Doradus in the Large Magellanic Cloud.
2. The correlation length of the velocity fluctuations is always a small fraction (roughly 2%) of the H II region diameter (section 5.2.1), implying that turbulent energy is injected on scales smaller than about 10% of the size of each region.
3. The power law slope of the structure function at scales smaller than the correlation length is  $m \approx 1.0 \pm 0.2$  for all regions, with no significant correlation with size, luminosity, or velocity amplitude (section 5.2.2). This can be explained as due to a fortuitous cancellation between, on the one hand, a steepening underlying velocity spectrum and, on the other hand, a reduced role for projection smoothing as the turbulent fluctuations pass from the subsonic to the supersonic regime.
4. The non-thermal component of the emission line widths ( $\langle \sigma_{\text{los}} \rangle$ ) is found to be proportional to the amplitude of the plane-of-sky velocity centroid fluctuations ( $\sigma_{\text{pos}}$ ) with a slope of unity, but an offset such that  $\langle \sigma_{\text{los}} \rangle \approx 10 \text{ km s}^{-1}$  when  $\sigma_{\text{pos}} = 0$  (section 5.2.3). This can be explained if  $\langle \sigma_{\text{los}} \rangle$  is affected by ordered velocity fields, such as radial expansion, which leave no imprint on  $\sigma_{\text{pos}}$ , implying that  $\sigma_{\text{pos}}$  is a better diagnostic of turbulent fluctuations than  $\langle \sigma_{\text{los}} \rangle$ , even though it is more challenging to measure.
5. The amplitudes of H $\alpha$  surface brightness fluctuations and turbulent velocity fluctuations are only weakly correlated (Figure 12h). For regions with supersonic turbulence these are consistent with the predictions for turbulence-induced density fluctuations, but for regions with subsonic turbulence the inferred density fluctuations are too large to be caused by turbulence alone.
6. On scales much larger than the correlation length, a variety of behaviors are seen (section 4.5). Some regions, such as Carina and Hubble X show uncorrelated homogeneous fluctuations at the largest scales, whereas others, such as Orion and 30 Doradus show evidence for radial gradients in the amplitude of the velocity fluctuations with more vigorous turbulence in the core of

the nebula than in the outskirts. In other regions, such as the Lagoon Nebula and NGC 346, there is evidence for quasi-periodic oscillations on scales similar to the size of the mapped region.

## ACKNOWLEDGEMENTS

Based in part on observations obtained at the Kitt Peak National Observatory, which is operated by the Association of Universities for Research in Astronomy, Inc., under cooperative agreement with the National Science Foundation. Based in part on observations made with the MUSE and FLAMES spectrographs on VLT telescopes at the La Silla Paranal Observatory, ESO, Chile under programme IDs 188.B-3002, 076.C-0888 and 098.D-0211. Based in part on observations made with the William Herschel Telescope operated by the Isaac Newton Group of Telescopes, located at the Spanish Roque de los Muchachos Observatory of the Instituto de Astrofísica de Canarias on the island of La Palma. We gratefully acknowledge financial support provided by Dirección General de Asuntos del Personal Académico, Universidad Nacional Autónoma de México, through grant Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica IN109823. We are grateful to Norberto Castro Rodríguez for providing maps of emission line velocity moments for 30 Dорадус derived from MUSE-VLT observations. JGV acknowledges and thanks CONACyT-Mexico for a PhD research scholarship. A sincere thanks to the anonymous referee who help us improve the quality of our work.

## DATA AVAILABILITY STATEMENT

All data and accompanying analysis programs used in this paper are available from the github repository <https://github.com/JavGVastro/PhD.Paper>.

## References

- Arias J. I., Barbá R. H., Maíz Apellániz J., Morrell N. I., Rubio M., 2006, *MNRAS*, 366, 739
- Arthur S. J., Medina S. N. X., Henney W. J., 2016, *MNRAS*, 463, 2864
- Bacon R., et al., 2010, in Proc. SPIE. p. 773508, doi:10.1117/12.856027
- Bacon R., et al., 2014, *The Messenger*, 157, 13
- Bensch F., Stutzki J., Ossenkopf V., 2001, *A&A*, 366, 636
- Bestenlehner J. M., et al., 2020, *MNRAS*, 499, 1918
- Bohuski T. J., 1973, *ApJ*, 183, 851
- Bosch G., Terlevich E., Terlevich R., 2002, *MNRAS*, 329, 481
- Brunt C. M., Mac Low M.-M., 2004, *ApJ*, 604, 196
- Brunt C. M., Federrath C., Price D. J., 2010, *MNRAS*, 403, 1507
- Castañeda H. O., 1988, *ApJS*, 67, 93
- Castro N., Crowther P. A., Evans C. J., Mackey J., Castro-Rodriguez N., Vink J. S., Melnick J., Selman F., 2018, *A&A*, 614, A147
- Chakraborty A., Anandarao B. G., 1999, *A&A*, 346, 947
- Chepurnov A., Lazarian A., 2010, *ApJ*, 710, 853
- Chávez R., Terlevich E., Terlevich R., Plonis M., Bresolin F., Basilakos S., Melnick J., 2012, *MNRAS*, 425, L56
- Chávez R., Terlevich R., Terlevich E., Bresolin F., Melnick J., Plonis M., Basilakos S., 2014, *MNRAS*, 442, 3565
- Cignoni M., et al., 2015, *ApJ*, 811, 76
- Damiani F., et al., 2016, *A&A*, 591, A74
- Damiani F., et al., 2017, *A&A*, 604, A135
- Danforth C. W., Sankrit R., Blair W. P., Howk J. C., Chu Y.-H., 2003, *ApJ*, 586, 1179
- De Marchi G., Panagia N., Sabbi E., 2011, *ApJ*, 740, 10
- Doh T., O'Dell C. R., Hartigan P., 2004, *AJ*, 127, 3456
- Dolphin A. E., Walker A. R., Hodge P. W., Mateo M., Olszewski E. W., Schommer R. A., Suntzeff N. B., 2001, *ApJ*, 562, 303
- Drisken L., Moffat A. F. J., Shara M. M., 1993, *AJ*, 105, 1400
- Eldridge J. J., Relaño M., 2011, *MNRAS*, 411, 235
- Esquivel A., Lazarian A., 2005, *ApJ*, 631, 320
- Esquivel A., Lazarian A., Horibe S., Cho J., Ossenkopf V., Stutzki J., 2007, *MNRAS*, 381, 1733
- Evans C. J., et al., 2011, *A&A*, 530, A108
- Fariña C., Bosch G. L., Barbá R. H., 2012, *AJ*, 143, 43
- Feast M. W., 1961, *Monthly Notices of the Royal Astronomical Society*, 122, 1
- Federrath C., Roman-Duval J., Klessen R. S., Schmidt W., Mac Low M. M., 2010, *A&A*, 512, A81
- Federrath C., Klessen R. S., Iapichino L., Beattie J. R., 2021, *Nature Astronomy*, 5, 365
- Foreman-Mackey D., 2017, *corner.py: Corner plots (ascl:1702.002)*
- Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, *PASP*, 125, 306
- Galtier S., Banerjee S., 2011, *Phys. Rev. Lett.*, 107, 134501
- García-Díaz M. T., Henney W. J., López J. A., Doi T., 2008, *Rev. Mex. Astron. Astrofis.*, 44, 181
- Gilmore G., et al., 2012, *The Messenger*, 147, 25
- González-Morán A. L., et al., 2021, *MNRAS*, 505, 1441
- Goodman J., Weare J., 2010, *Communications in Applied Mathematics and Computational Science*, 5, 65
- Gordon S., Koribalski B., Houghton S., Jones K., 2000, *MNRAS*, 315, 248
- Gouliermis D. A., Chu Y.-H., Henning T., Brandner W., Gruendl R. A., Hennekemper E., Hormuth F., 2008, *ApJ*, 688, 1050
- Greisen E. W., Calabretta M. R., Valdes F. G., Allen S. L., 2006, *A&A*, 446, 747
- Güdel M., Briggs K. R., Montmerle T., Audard M., Rebull L., Skinner S. L., 2008, *Science*, 319, 309
- Hanel A., 1987, *A&A*, 176, 347
- Haverkorn M., Gaensler B. M., McClure-Griffiths N. M., Dickey J. M., Green A. J., 2004, *ApJ*, 609, 776
- Heydari-Malayeri M., Selier R., 2010, *A&A*, 517, A39
- Hippelein H. H., 1986, *A&A*, 160, 374
- Hodge P., Lee M. G., Kennicutt Robert C. J., 1989, *PASP*, 101, 32
- Hunter D. A., Baum W. A., O'Neil Earl J. J., Lynds R., 1996, *ApJ*, 456, 174
- Hénault-Brunet V., et al., 2012, *A&A*, 545, L1
- Jaupart E., Chabrier G., 2022, *A&A*, 663, A113
- Kalari V. M., Horch E. P., Salinas R., Vink J. S., Andersen M., Bestenlehner J. M., Rubio M., 2022, arXiv e-prints, p. arXiv:2207.13078
- Kam Z. S., Carignan C., Chemin L., Amram P., Epinat B., 2015, *MNRAS*, 449, 4048
- Kelly B. C., 2007, *ApJ*, 665, 1489
- Kennicutt R. C. J., 1984, *ApJ*, 287, 116
- Kinson D. A., Oliveira J. M., van Loon J. T., 2021, *MNRAS*, 507, 5106
- Klipp C., 2014, *Boundary-Layer Meteorology*, 151, 57
- Koch E. W., Rosolowsky E. W., Boyden R. D., Burkhardt B., Ginsburg A., Loepky J. L., Offner S. S. R., 2019, *AJ*, 158, 1
- Krumholz M. R., Burkhardt B., 2016, *MNRAS*, 458, 1671
- Lagrois D., Joncas G., 2009, *ApJ*, 700, 1847
- Lagrois D., Joncas G., 2011, *MNRAS*, 413, 721
- Lagrois D., Joncas G., Drissen L., Arsenault R., 2011, *MNRAS*, 413, 705
- Levrier F., 2004, *A&A*, 421, 387
- Lewis J. R., Unger S. W., 1992, in Worrall D. M., Biemesderfer C., Barnes J., eds, *PASPVol. 25, Astronomical Data Analysis Software and Systems I*. p. 445
- Louise R., Monnet G., 1970, *A&A*, 8, 486
- Maggi P., et al., 2019, *A&A*, 631, A127
- Malumuth E. M., Waller W. H., Parker J. W., 1996, *AJ*, 111, 1128
- Martínez-Galarza J. R., Hunter D., Groves B., Brandl B., 2012, *ApJ*, 761, 3
- Medina-Tanco G. A., Sabalisck N., Jatenco-Pereira V., Opher R., 1997, *ApJ*, 487, 163
- Medina S.-N., Arthur S., Henney W., Mellema G., Gazol A., 2014, *MNRAS*, 445, 1797

- Melnick J., Moles M., Terlevich R., Garcia-Pelayo J.-M., 1987, MNRAS, 226, 849
- Melnick J., Tenorio-Tagle G., Telles E., 2021, A&A, 649, A175
- Miville-Deschénes M., Joncas G., Durand D., 1995, ApJ, 454, 316
- Miville-Deschénes M.-A., Levrier F., Falgarone E., 2003, ApJ, 593, 831
- Moiseev A. V., Lozinskaya T. A., 2012, MNRAS, 423, 1831
- Moiseev A. V., Tikhonov A. V., Klypin A., 2015, MNRAS, 449, 3568
- Moré J. J., 1978, in Watson G. A., ed., Numerical Analysis. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 105–116
- Münch G., 1958, Reviews of Modern Physics, 30, 1035
- Newville M., Stensitzki T., Allen D. B., Ingargiola A., 2014, LMFIT: Non-Linear Least-Square Minimization and Curve-Fitting for Python, doi:10.5281/zenodo.11813, <https://doi.org/10.5281/zenodo.11813>
- Nota A., et al., 2006, ApJ, 640, L29
- O'Dell C. R., 2001, ARA&A, 39, 99
- O'Dell C. R., Henney W. J., 2008, AJ, 136, 1566
- O'Dell C. R., Wen Z., 1992, ApJ, 387, 229
- O'Dell C. R., Townsley L. K., Castaneda H. O., 1987, ApJ, 317, 676
- O'Dell C. R., Wen Z., Hu X., 1993, ApJ, 410, 696
- O'Dell C. R., Hodge P. W., Kennicutt Robert C. J., 1999, PASP, 111, 1382
- Ossenkopf V., Esquivel A., Lazarian A., Stutzki J., 2006, A&A, 452, 223
- Pasquini L., et al., 2002, The Messenger, 110, 1
- Pietrzyński G., et al., 2013, Nature, 495, 76
- Pope S. B., 2000, Turbulent Flows. Cambridge University Press, doi:10.1017/CBO9780511840531
- Prisinzano L., Damiani F., Micela G., Sciotino S., 2005, A&A, 430, 941
- Randich S., Gilmore G., Gaia-ESO Consortium 2013, The Messenger, 154, 47
- Robberto M., et al., 2013, ApJS, 207, 10
- Roy J.-R., Joncas G., 1985, ApJ, 288, 142
- Roy J. R., Arsenault R., Joncas G., 1986, ApJ, 300, 624
- Rozas M., Richer M. G., López J. A., Relaño M., Beckman J. E., 2006, A&A, 455, 539
- Sabalisk N. S. P., Tenorio-Tagle G., Castaneda H. O., Muñoz-Tunon C., 1995, ApJ, 444, 200
- Scalo J. M., 1984, ApJ, 277, 556
- Schmidt W., Federrath C., Klessen R., 2008, Phys. Rev. Lett., 101, 194505
- Sibbons L. F., Ryan S. G., Cioni M. R. L., Irwin M., Napiwotzki R., 2012, A&A, 540, A135
- Simón-Díaz S., Herrero A., Esteban C., Najarro F., 2006, A&A, 448, 351
- Simon M., Cassar L., Felli M., Fischer J., Massi M., Sanders D., 1984, ApJ, 278, 170
- Smith N., 2006, ApJ, 644, 1151
- Smith N., Brooks K. J., 2008, The Carina Nebula: A Laboratory for Feedback and Triggered Star Formation. p. 138
- Smith M. G., Weedman D. W., 1972, ApJ, 172, 307
- Stutzki J., Bensch F., Heithausen A., Ossenkopf V., Zieliński M., 1998, A&A, 336, 697
- Terlevich R., Melnick J., 1981, MNRAS, 195, 839
- Tiwari M., Menten K. M., Wyrowski F., Pérez-Beaupuits J. P., Wiesemeyer H., Güsten R., Klein B., Henkel C., 2018, A&A, 615, A158
- Tiwari M., Menten K. M., Wyrowski F., Giannetti A., Lee M. Y., Kim W. J., Pérez-Beaupuits J. P., 2020, A&A, 644, A25
- Tomita A., Ohta K., Saito M., 1993, Publications of the Astronomical Society of Japan, 45, 693
- Torres-Flores S., Barbá R., Maíz Apellániz J., Rubio M., Bosch G., Hénault-Brunet V., Evans C. J., 2013a, A&A, 555, A60
- Torres-Flores S., Barbá R., Maíz Apellániz J., Rubio M., Bosch G., Hénault-Brunet V., Evans C. J., 2013b, A&A, 555, A60
- Tothill N. F. H., Gagné M., Stecklum B., Kenworthy M. A., 2008, The Lagoon Nebula and its Vicinity. p. 533
- Von Hoerner S., 1951, Zeitschrift für Astrophysik, 30, 17
- Walborn N. R., Barbá R. H., Sewiło M. M., 2013, AJ, 145, 98
- Woodward C. E., et al., 1986, AJ, 91, 870
- Wu Y., Cao S., Zhang J., Liu T., Liu Y., Geng S., Lian Y., 2020, ApJ, 888, 113
- Ye T., Turtle A. J., Kennicutt R. C. J., 1991, MNRAS, 249, 722

Yu X., et al., 2019, MNRAS, 486, 4463

## APPENDIX A: DEGRADATION OF THE STRUCTURE FUNCTION DUE TO OBSERVATIONAL LIMITATIONS

To characterize the impact of the finite map size and the atmospheric seeing on the structure function we perform a series of experiments using artificial turbulent velocity maps. We quantify how these observational limitations affect the determination of the structure function parameters, such as the correlation length  $r_0$  and the velocity field variance  $\sigma_{\text{pos}}^2$ . The synthetic velocity maps are created using a modified version of the `make_3dfield` command from the `turbustat` Python library (Koch et al. 2019), which creates a three-dimensional fractional Brownian Motion field (Miville-Deschénes et al. 2003) based on a power-law energy spectrum in wavenumber  $k$ ,  $E(k) \propto k^{-n_{3D}}$ , and with random phases. Our modification multiplies the energy spectrum by a low-wavenumber exponential taper:  $e^{-1/2\pi r_0 k}$ . As a result, the generated fields become uncorrelated at scales larger than  $r_0$  and the resulting structure function closely approximates our model form (section 3.2).

In order to create a 2D plane-of-sky velocity map, we first create 3D cubes of fluctuating velocity and emissivity with `make_3dfield`, then transform to a PPV (position-position velocity) cube using the `make_ppv` command from `turbustat`, assuming thermal broadening with FWHM of  $20 \text{ km s}^{-1}$ , as appropriate for a hydrogen line at  $T \approx 10^4 \text{ K}$ . Finally, the first velocity moment map is obtained by integrating along the velocity axis.

For the emissivity field, we assume a log-normal distribution with RMS fractional width  $\sigma_E/E_0 = 1.0$ , which is typical of our sources (see footnote 8 and Figure 12 above). We further assume that there is no correlation between the emissivity and velocity fluctuations and that their spatial correlation length and power law indices are equal.<sup>10</sup>

The 3-dimensional 2nd-order structure function slope is related to the power law index as  $m_{3D} = n_{3D} - 3$ , while the equivalent relation in two dimensions is  $m_{2D} = n_{2D} - 2$  for separations smaller than the line-of-sight depth of the emitting region. So long as the emissivity fluctuations are weak and uncorrelated with the velocity fluctuations, then  $n_{2D} = n_{3D}$  (Miville-Deschénes et al. 2003; Levrier 2004), but if these conditions are not satisfied, then the structure function of the velocity centroids does not purely reflect the velocity power spectrum (Brunt & Mac Low 2004; Esquivel & Lazarian 2005; Ossenkopf et al. 2006; Esquivel et al. 2007) and one has  $m_{2D} = n_{3D} - 2 + \delta\kappa$  where  $\delta\kappa$  is a correction factor that accounts for the contribution to the structure function from column density fluctuations and cross terms (see section 5.2.2).

Brunt & Mac Low (2004) found that  $\delta\kappa$  is a function of the Mach number for hydrodynamic turbulent box simulations. In principle, we could calculate  $\delta\kappa$  for the fractional Brownian motion models that we use here by using the analytic machinery of Esquivel et al. (2007), but we prefer instead to use a more empirical approach. We vary  $n_{3D}$  until the resultant structure function slope matches a particular value in the range  $m_{2D} = 0.8 \rightarrow 1.2$  that encompasses our observational results.<sup>11</sup> At each step, we estimate  $m_{2D}$  by fitting our

<sup>10</sup> We have investigated departures from these assumptions and found that they make very little difference to the results. In particular, we have varied  $\sigma_E/E_0$  between 0.0 (constant emissivity) and 2.0 (very strong fluctuations, as seen in 30 Doradus), and found that the fitting parameters for the effects of seeing (see below) change by less than 20%.

<sup>11</sup> Note that  $m_{2D}$  is written simply as  $m$  in the main text.

idealized model structure function (equation (10), without the seeing and noise terms) to the structure function of the simulated velocity map. In the case of a uniform emissivity field ( $\sigma_E/E_0 = 0$ ), we find  $\delta\kappa \approx 0$ , which is maximal projection smoothing, exactly as expected for a case that mimics incompressible turbulence (see Appendix of Miville-Deschénes et al. 2003). As the amplitude of the emissivity fluctuations increase, we find that  $\delta\kappa$  becomes increasingly negative, with  $\delta\kappa \approx -0.3$  for  $\sigma_E/E_0 = 1$  (the case that we illustrate here) and  $\delta\kappa \approx -0.55$  for  $\sigma_E/E_0 = 2$ . This is qualitatively consistent with the results in Figure 12a of Brunt & Mac Low (2004), although a quantitative comparison is hard to make, since they are not holding  $m$  constant in the way we do here.

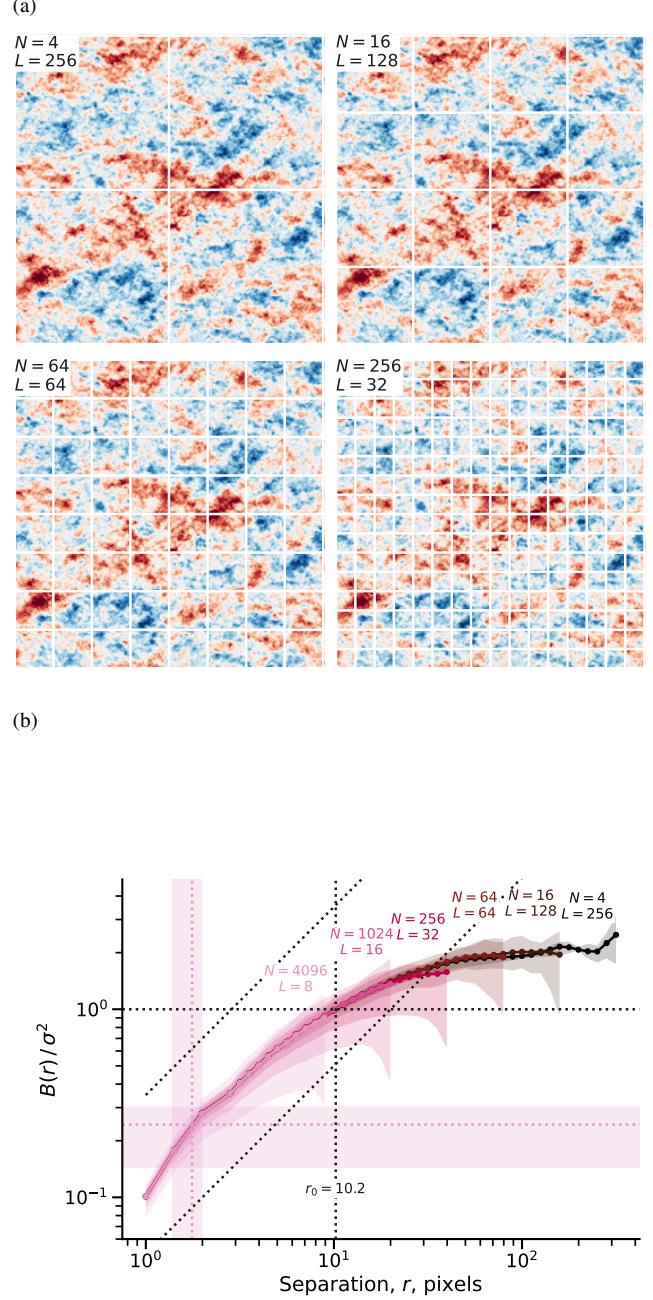
The maps generated by `turbustat` are spatially periodic, but we make them non-periodic by always dividing the full map into at least four tiles, which are each analyzed separately.

## A1 Finite box effects

Figure A1a shows how we test for finite map effects by taking a simulated 1st moment map of size  $512 \times 512$  pixels and iteratively dividing the full map into  $2 \times 2$  tiles. On the  $j$ th iteration we have  $N = 4^j$  tiles, each of size  $L \times L$  pixels with  $L = 2^{9-j}$ . The figure shows the result for  $j = 1, 2, 3$ , and 4, in which the full map was generated with a correlation length  $r_0 = 32$  pixels. We then calculate the individual structure function for each tile and find the mean and standard deviation for each level of subdivision, showing the results in Figure A1b. It can be seen that the mean structure functions for each level of subdivision (shown by colored dots and solid lines) overlap very closely with one another, apart from deviations due to edge effects when  $r > L/2$ . On the other hand, there is considerable dispersion between individual tiles (shown by light colored shading) for the cases where  $L < r_0$ .

We next empirically measure the velocity variance and apparent correlation length for each tile. The velocity variance  $\sigma^2$  is found by directly averaging over the pixels in the tile using equation (5), while the apparent correlation length  $r_0$  is found as the separation  $r$  for which  $B(r) = \sigma^2$ , using the tile's own empirically measured structure function and variance. In Figure A2a we show how these quantities deviate from the “true” values (as measured on the full  $512 \times 512$  map) as a function of the tile size  $L$ . It can be seen that the true values are successfully recovered so long as the map size is more than ten times the correlation length, but the apparent values of  $r_0$  and  $\sigma^2$  are smaller than the true values when  $L < 10r_0$ , increasingly so as  $L$  becomes smaller. This can be intuitively understood as due to the fact that a tile that is small compared with the correlation length will fail to sample the full range of velocity fluctuations, leading to an underestimate of  $\sigma^2$ . Then the fact that  $B(r)$  increases with  $r$  in the power law regime means there will be a concomitant underestimate of  $r_0$ .

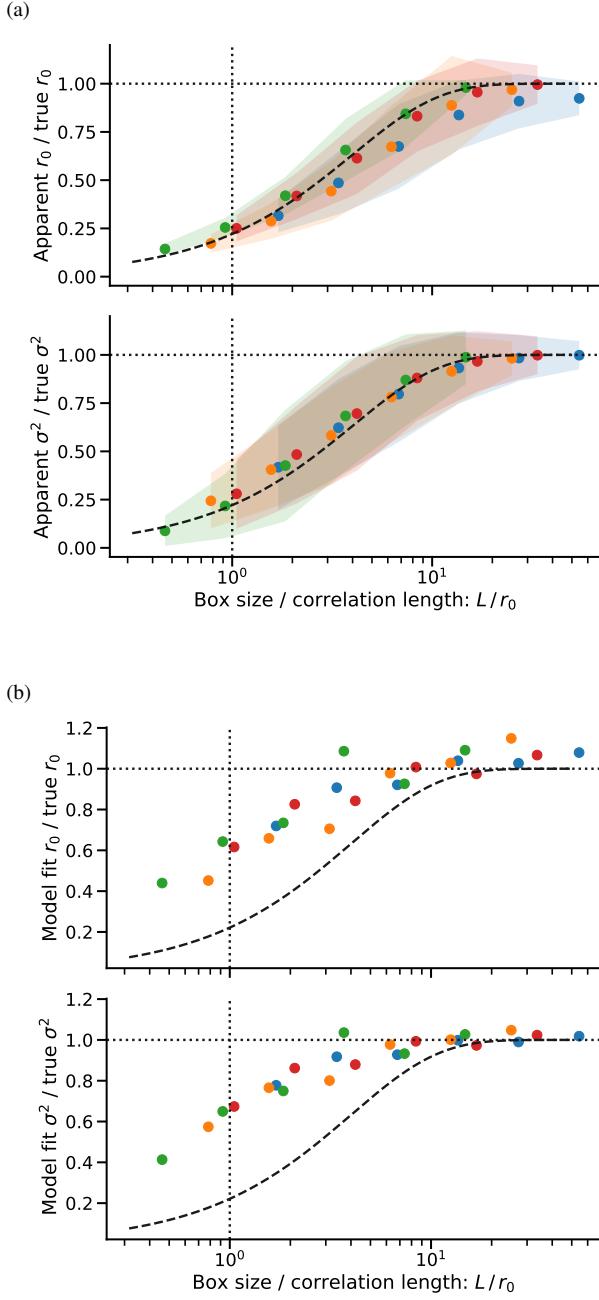
However, it is possible to do a considerably better job of recovering the true  $\sigma^2$  and  $r_0$  by fitting our heuristic model structure function of equation (10). This is shown in Figure A2b, where it can be seen that the true parameters are reliably recovered so long as  $L > 5r_0$ , which is a roughly 0.5 dex improvement over the purely empirical method. In order for the model fitting to work, the tile size must be large enough that the structure function begins to show some flattening at the largest scales, which is why the results are still inaccurate for the smallest values of  $L$  that sample only the rising power-law portion of  $B(r)$ .



**Figure A1.** Effects of finite box size on the structure function. (a) Construction of simulated turbulent velocity fields of different sizes by repeated division of an initial field of size  $512 \times 512$  pixels. The  $j$ th level of division yields  $N = 4^j$  fields, each of linear size  $L \times L$  where  $L = 2^{9-j}$ . The color scale is the same for all maps and it considers  $-3$  to  $+3$  times the standard deviation. (b) Resultant structure functions. Colored lines show the average  $B(r)$  over all sub-images, while the shaded areas show the standard deviation of  $B(r)$ .

## A2 Effects of seeing

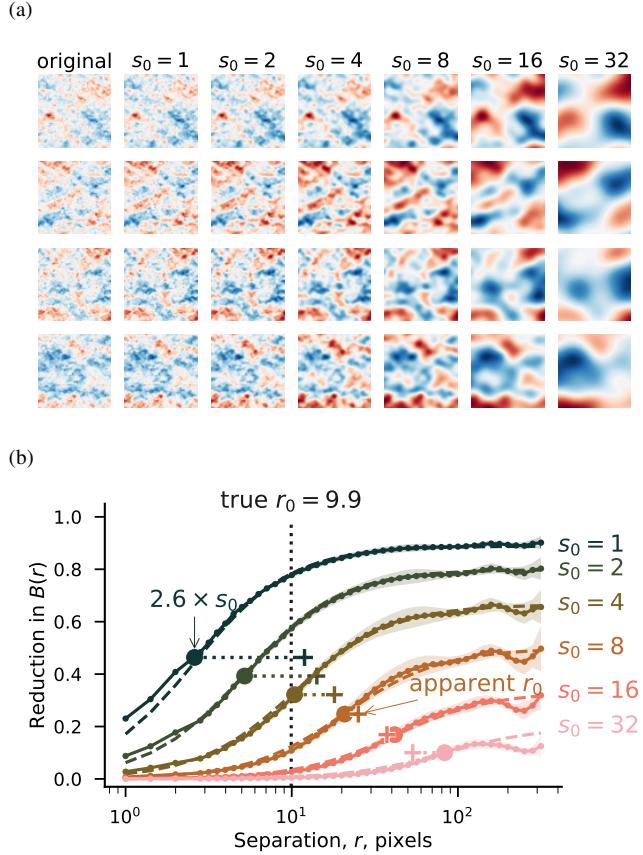
Ground-based optical observations are affected by atmospheric seeing, which limits the angular resolution of any observations. This changes the observed structure function by suppressing fluctuations between the most closely separated points. Bensch et al. (2001) stud-



**Figure A2.** Reduction of apparent velocity variance and correlation length due to finite box size. (a) Empirical measurements of  $\sigma^2$  and  $r_0$ . Colored dots show mean results from different simulated velocity fields, with different power law slopes. Colored shading shows the inter-quartile range over the different sub-images. The dashed line is an approximate analytic fit to these results ( $1 - e^{-L/4.0r_0}$ ). (b) Determination of  $\sigma^2$  and  $r_0$  from fitting the model of equation (9). Colored dots are mean results from different simulated velocity fields, while the dashed line is the same as in (a).

ied the analogous problem of beam smearing in the context of the  $\Delta$ -variance method (Stutzki et al. 1998), deriving a leading order approximation for the effect, but an exact analytic treatment is difficult. We therefore prefer to empirically study the effect of seeing using the synthetic velocity maps described above.

Figure A3a illustrates how four different synthetic 1st-moment maps of size  $256 \times 256$  pixels and correlation length  $r_0 = 32$  pixels



**Figure A3.** Effects of seeing on the structure function. (a) Each row shows a different simulated random velocity field on a  $256^2$  grid (see text for details). The left column shows the original field, while the remaining columns show the effects of smoothing by a gaussian kernel with RMS width  $s_0$  from 1 to 32 pixels. The same color scale is used for all columns, ranging from  $-3$  to  $+3$  times the standard deviation of the unsmoothed map. (b) Relative change in the second-order structure function due to the gaussian smoothing. Solid lines with symbols show the average over the 4 maps of  $B(r, s_0)/B(r, s_0 = 0)$  as a function of separation  $r$ , with a separation of  $2.6s_0$  indicated by a large filled circle on each curve. Dashed lines shows the empirical fit discussed in the text. The dotted line shows the correlation length of the original fields, while colored plus symbols show the apparent correlation length of the smoothed fields.

are affected by progressively broader gaussian seeing, with RMS width  $s_0 = 1$  to 32 pixels. Note that the seeing is applied to each velocity channel of the PPV emissivity cube before integrating to find the 2D velocity map. We calculate the structure function  $B(r)$  for each map and show in Figure A3b the average ratio between  $B(r)$  and the value in the absence of seeing, with different colors representing different values of  $s_0$ . Results are shown for a power-law slope of  $m = 1.0$  but we find very similar behaviour for other values of  $m$  within the range seen in our observations. For small values of  $s_0$ , the principle effect is seen at the smallest separations but the structure function is noticeably depressed for separations as large as  $r = 10s_0$ . As  $s_0$  increases to become comparable with  $r_0$ , even the largest separations are significantly affected. At the same time the apparent correlation length begins to increase, as indicated by the plus symbols in the figure.

We find that the following simple analytic function reproduces

these results to within better than 10%:

$$S(r; s_0, r_0) = \left[ \left( 1 + a_1 \frac{s_0}{r_0} \right) \left( 1 + \left( a_2 \frac{s_0}{r} \right)^{a_3} \right) \right]^{-1}, \quad (\text{A1})$$

which is shown by dashed lines in Figure A3b. The exact values of the constants  $a_1$ ,  $a_2$ ,  $a_3$  are found to depend slightly on the power law slope  $m$  and the magnitude of the emissivity fluctuations  $\sigma_E/E_0$ . We adopt the compromise values  $a_1 = 1.25$ ,  $a_2 = 2.6$ , and  $a_3 = 1.5$ , which fit the numerical results to within 20% for  $0.8 < m < 1.2$  and  $0.5 < \sigma_E/E_0 < 2.0$ . We therefore use this equation to represent the seeing in our structure function fits (section 4), with  $s_0$  as an additional free parameter for the models.

## APPENDIX B: APPARENT VELOCITY FLUCTUATIONS FROM EMISSIVITY FLUCTUATIONS

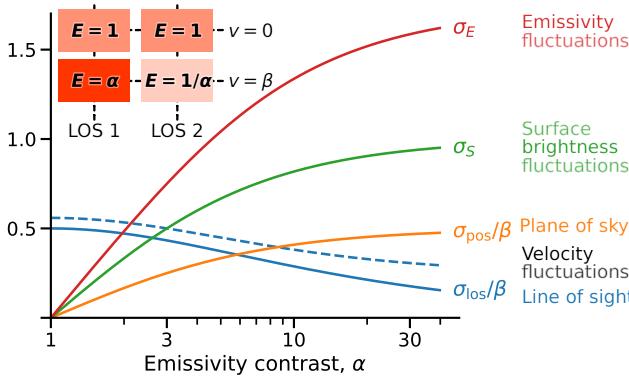
In this appendix we describe a simple toy model to illustrate how fluctuations of emissivity, combined with an ordered velocity gradient along the line of sight, can give rise to apparent fluctuations on the plane of the sky of the mean velocity. The model is illustrated in the inset of Figure B1. It consists of two overlapping emission layers, A and B, with velocities along the line of sight  $v_A = 0$  and  $v_B = \beta$ , respectively, which are observed along two lines of sight (LOS), 1 and 2. The emissivity of layer A is  $E_{A1} = E_{A2} = 1$  for both LOS, while the emissivity of layer B is  $E_{B1} = \alpha$  for LOS 1 and  $E_{B2} = \alpha^{-1}$  for LOS 2, where  $\alpha$  is a positive constant. Both layers are assumed to have unit depth along line of sight, such that the surface brightness for the two LOS are given by  $S_1 = 1 + \alpha$  and  $S_2 = 1 + \alpha^{-1}$ .

It then follows that the fractional standard deviation of the emissivity and surface brightness are given by:

$$\sigma_E = \frac{[3(\alpha + \alpha^{-1})^2 - 4(1 + \alpha + \alpha^{-1})]}{2 + \alpha + \alpha^{-1}}^{1/2} \quad (\text{B1})$$

$$\sigma_S = \frac{|\alpha - 1|}{1 + \alpha}, \quad (\text{B2})$$

while the RMS line width and standard deviation of the mean veloc-



**Figure B1.** Toy model of emissivity variations in an ordered velocity field, showing co-variation of fractional dispersion of emissivity,  $\sigma_E$ , and surface brightness,  $\sigma_S$ , together with plane-of-sky dispersion in mean velocity,  $\sigma_{\text{pos}}$ , and the mean RMS line width,  $\langle \sigma_{\text{los}} \rangle$ , both in units of the inter-layer velocity difference  $\beta$ .

ties on the plane of the sky are given by

$$\sigma_{\text{pos}} = \frac{|\alpha - 1|\beta}{2(1 + \alpha)} \quad (\text{B3})$$

$$\langle \sigma_{\text{los}} \rangle = \frac{\alpha^{1/2}\beta}{1 + \alpha}. \quad (\text{B4})$$

These are all shown in Figure B1 as a function of the emissivity contrast  $\alpha$ . It can be seen that  $\sigma_E$ ,  $\sigma_S$ , and  $\sigma_{\text{pos}}/\beta$  all increase with  $\alpha$ , reaching asymptotic values of  $\sqrt{3}$ , 1, and 0.5, respectively as  $\alpha \rightarrow \infty$ .<sup>12</sup> The line width  $\langle \sigma_{\text{los}} \rangle/\beta$  behaves in the opposite sense, falling from a maximum value of 0.5 when  $\alpha = 0$  towards zero as  $\alpha \rightarrow \infty$ , a result of the emission becoming increasingly dominated by one or the other of the two layers for high emissivity contrast. The dashed line shows the effect of including a small velocity gradient within each layer that contributes an additional width of  $\beta/4$ , which is assumed to add in quadrature with  $\langle \sigma_{\text{los}} \rangle$ .

The predictions of this model are compared with observations in panel i of Figure 12.

## APPENDIX C: NOTES ON INDIVIDUAL H II REGIONS

For each region or pair of regions in our sample, we give a more detailed description of the physical properties (supplementing section 2.2) together with additional details of the structure function (supplementing section 4) and comparison with previous results from the literature (supplementing section 5.1).

### C1 Orion Nebula

The Orion Nebula (M42) is the prototype of a galactic H II region located at a distance of 440 pc ( $1'' = 0.002$  pc; O'Dell & Henney 2008). The O7 V star  $\theta^1$  Ori C is the most luminous and hottest star with an ionizing flux  $Q_0$  of  $1.3 \times 10^{49}$  photons s<sup>-1</sup> (Simón-Díaz et al. 2006), dominating the ionization of the nebula, which has a luminosity in H $\alpha$  of  $1.5 \times 10^{37}$  erg s<sup>-1</sup>. With 3 other B-type stars they form the Trapezium cluster. The bright region that surrounds these massive ionizing stars is called the Huygens region. The position of the Orion Nebula on the near side of the Orion Molecular Cloud (OMC-1) makes it an ongoing site of star formation with a large population of young stars, some of which are sources of stellar jets and Herbig-Haro objects (O'Dell et al. 1993). Its physical properties and kinematics, along with its stellar population is documented in O'Dell (2001).

#### C1.1 Observed structure function and model fit

**Orion core** Figure 5a shows the structure function of the Orion core. A detailed description of the model fit is given in section 4.2 and will not be repeated here. At scales larger than the correlation length, a maximum  $B(r)$  value of  $23 \text{ km}^2 \text{ s}^{-2}$  is seen at  $r = 0.30$  pc, beyond which the structure function declines to  $7 \text{ km}^2 \text{ s}^{-2}$  at  $r = 0.5$  pc. This “hook” behavior can be explained by a radial gradient in the amplitude of the velocity fluctuations (see section 3.2.3).

An alternative approach to determine the velocity variance and correlation length is to empirically measure  $\sigma_{\text{pos}}^2$  from the observations and then to use linear interpolation on the measured structure function to find the point where  $B(r_0) = \sigma_{\text{pos}}^2$ . This yields

<sup>12</sup> To reach higher relative brightness dispersion  $\sigma_S > 1$  would require including more than two lines of sight in the model.

$\sigma_{\text{pos}}^2 = 9 \text{ km}^2 \text{s}^{-2}$  and  $r_0 = 0.05 \text{ pc}$ , as compared to the model-derived values of  $\sigma_{\text{pos}}^2 = 13 \text{ km}^2 \text{s}^{-2}$  and  $r_0 = 0.07 \text{ pc}$ . The main reason for the discrepancy is the finite box effect discussed in section A1 above, where it is shown that for homogeneous fluctuations it is necessary to sample a map of size  $L > 10r_0$  in order to empirically determine  $\sigma_{\text{pos}}^2$ . In the case of the Orion core observations,  $L = 7r_0$ , so the fluctuations are not fully sampled, leading to the empirical  $\sigma_{\text{pos}}^2$  being an underestimate. However, given the evidence for a radial gradient in the amplitude of the velocity fluctuations, it could be argued that the finite box effect is not so important as it would be for homogeneous fluctuations and it is instead the model-derived value of  $\sigma_{\text{pos}}^2$  that is an overestimate.

**EON** Figure 5b shows the structure function of the Extended Orion Nebula, which is ten times larger than the core in linear size, but observed at a much lower angular resolution (see Figures 1 and 2). The data for this region are the lowest quality of all the regions in our sample, and there are too many degeneracies between the parameters to achieve a credible fit using the full model. We therefore fix the power slope at  $m = 1$ , which implies that noise must make a significant contribution at all scales in order to reproduce the observed flat structure function. In turn, this implies that the true velocity variance is  $\sigma_{\text{pos}}^2 = 5 \text{ km}^2 \text{s}^{-2}$ , which is lower than the  $7 \text{ km}^2 \text{s}^{-2}$  that would be naively derived from the observations (compare green and orange lines in Figure 5b). Comparing the extended nebula with the core for the size scales where they overlap (0.1 to 0.5 pc), one sees that  $B(r)$  is roughly four times smaller in the EON than in the core for the same value of  $r$ . This is further evidence for a decline in the fluctuation amplitude with radius, as mentioned above. Also consistent with this is a hook-like downturn in  $B(r)$  for  $r > 2 \text{ pc}$  for the EON region. The fact that the derived  $r_0$  is larger for the EON region than the core (0.5 pc versus 0.07 pc) means that the correlation length is also inhomogeneous, tending to increase from the center to the outskirts of the nebula.

### C1.2 Comparison with previous studies

Arthur et al. (2016) calculated the structure function of different emission lines in the inner Orion Nebula with varying degrees of ionization. They found that higher ionization lines show a larger velocity dispersion and a steeper structure function slope. They also make a comparison with previous work (see Table 5 of their paper), finding a broad agreement in derived structure function parameters and their dependence on ionization energy. The power-law index obtained by Arthur et al. (2016) is  $m \sim 1.17 \pm 0.08$  for H $\alpha$ , with a correlation length of  $\sim 0.05 \text{ pc}$  and velocity variance  $\sigma_{\text{pos}}^2 = 9.4 \text{ km}^2 \text{s}^{-2}$ . This compares with our own values of  $m = 1.07 \pm 0.03$ ,  $r_0 = (0.068 \pm 0.004) \text{ pc}$  and  $\sigma_{\text{pos}}^2 = (13 \pm 1) \text{ km}^2 \text{s}^{-2}$ . Since they are based on the same observational data, the differences of order 30% between the Arthur et al. (2016) results and our own are entirely due to the different methodologies employed. In the case of  $\sigma_{\text{pos}}^2$ , their value is lower than ours since they masked out the locations of known high-velocity Herbig-Haro flows, whereas we do not.

## C2 Lagoon Nebula

The Lagoon Nebula (M8, NGC 6523) is located at a distance of 1250 pc ( $1'' = 0.006 \text{ pc}$ ; Prisinzano et al. 2005). It contains the young stellar cluster NGC 6530 and the region is illuminated by

different massive stars of spectral types O and B, the hottest being 9 Sgr, type O4V (Damiani et al. 2017). The optically brightest part of the nebula is denoted the “Hourglass nebula”, which is illuminated by the O7 star Herschel 36 (Woodward et al. 1986). These stars gives an ionizing flux  $Q_0 = 2.6 \times 10^{49} \text{ photons s}^{-1}$  (2 times higher than in the Orion Nebula) and the luminosity in H $\alpha$  is  $2.95 \times 10^{37} \text{ erg s}^{-1}$  (Kennicutt 1984). The mean non-thermal linewidth in H $\alpha$  is  $(11.7 \pm 1.6) \text{ km s}^{-1}$  (Bohuski 1973). Molecular gas and ongoing star formation are present in the region around Herschel 36 (Arias et al. 2006; Tiwari et al. 2018) and the lower-luminosity region M8 East (Simon et al. 1984; Tiwari et al. 2020). The properties of the nebula are reviewed by Tothill et al. (2008).

### C2.1 Observed structure function and model fit

Figure 5c shows the structure function of the Lagoon nebula. The model fit gives  $\sigma_{\text{pos}}^2 = (7 \pm 1) \text{ km}^2 \text{s}^{-2}$ ,  $r_0 = (1.0 \pm 0.2) \text{ pc}$ , and  $m = 1.26 \pm 0.2$ . In this case, there is good agreement between the model value and empirical value of  $\sigma_{\text{pos}}^2$ , indicating that finite box effects are unimportant. On the other hand, the model-derived  $m$  is steeper than would be naively derived from the observations, which is a symptom of the effects of noise. At separations smaller than the correlation length, the structure function shows irregular oscillations about the mean ascending trend, which are a symptom of the relatively small number of spatial points in this source.

At scales much larger than the correlation length the structure function begins to increase again, reaching a value of  $66 \text{ km}^2 \text{s}^{-2}$  at 20 pc, which is greater than  $9\sigma_{\text{pos}}^2$ . As discussed in section 4.5, this may be a sign of periodicity or a large-scale gradient in the velocity field.

### C2.2 Comparison with previous studies

The first investigation of the Lagoon nebula structure function is from Louise & Monnet (1970), who use Fabry-Perot observation of the H $\alpha$  line to probe scales between 0.7 pc and 7 pc. They find a rising power-law behavior with slope of  $m = 0.64$ , which is slightly smaller than the value we derive. However, given the very coarse binning in separation of their structure function, and the fact that they do not account for noise or the flattening of the structure function at large scales, their results are broadly consistent with ours, with the value of  $\sigma_{\text{pos}}^2 = 16 \text{ km}^2 \text{s}^{-2}$  at  $r = 3.5 \text{ pc}$  being almost identical to our own value at that scale.

More recent measurements of the structure function of the Lagoon Nebula have used the [O III]  $\lambda 5007$  line and have probed smaller scales. O’Dell et al. (1987) used longslit observations at a small number of position angles to study the structure function over an area  $2 \times 2 \text{ arcmin}^2$ , centered on the nebula’s principal ionizing star, 9 Sgr. However, it is hard to compare their results to ours, as they carry out a Gaussian decomposition of the line profile into two components, A and B, and analyse the structure function of the two components separately. The total variance of component A (the stronger and narrower of the two) is  $2 \text{ km}^2 \text{s}^{-2}$  at scales  $< 1 \text{ pc}$ , which is significantly smaller than our value of  $7 \text{ km}^2 \text{s}^{-2}$ , but this is not surprising since they are probing scales smaller than the correlation length and therefore suffer from a severe finite-box effect. Component B on the other hand shows a much larger variance of  $13 \text{ km}^2 \text{s}^{-2}$  and an essentially flat structure function.

Another [O III]  $\lambda 5007$  study (Chakraborty & Anandaraao 1999) calculated the structure function in the Hourglass region of the nebula: an area of  $1 \times 1 \text{ arcmin}^2$  centered at Herschel 36. As compared to the

data used in this paper, which spans between 0.1 and 10 pc), their results cover smaller scales, between 0.012 and 0.3 pc. Chakraborty & Anandarao (1999) report a rising power-law behavior with slope of  $m = 0.46$  over the entire range. No flattening of the structure function at the largest scales was seen, but that is to be expected if the correlation length is larger than the size of the region mapped. The amplitude of the fluctuations is much higher ( $\sigma_{\text{pos}}^2 = 130 \text{ km}^2 \text{ s}^{-2}$ ) than we find for the wider nebula ( $\sigma_{\text{pos}}^2 = 7 \text{ km}^2 \text{ s}^{-2}$ ), suggesting an even more extreme version of the inhomogeneity that we find in the Orion Nebula. This is supported by the fact that the Hourglass region also shows much larger line widths than the rest of the nebula (Chakraborty & Anandarao 1999), which is probably a symptom of the ongoing active star formation in that region.

### C3 Carina Nebula

The Carina nebula is a large star-forming complex located at a distance of 2350 pc ( $1'' = 0.01$  pc; Smith 2006). It contains young clusters like Trumpler 14 and 16, and Collinder 228 from the Car OB1 association, one of the largest OB associations in the Galaxy. This OB association has more than 65 stars earlier than B0, and several Wolf-Rayet (WR) stars and the most luminous star is the blue variable  $\eta$  Car. The ionizing flux  $Q_0$  is  $8.9 \times 10^{50}$  photons  $\text{s}^{-1}$  (Smith & Brooks 2008) (60 times higher compared to the Orion Nebula) and the luminosity in H $\alpha$  of  $10^{39}$  erg  $\text{s}^{-1}$ . As a high-mass star formation region, it presents several phenomena related to young stars, such as evaporating protoplanetary discs, erosion of large dust pillars, and the triggering of a second generation of embedded stars and Herbig-Haro objects (see Smith & Brooks 2008 and reference therein). This region is at an intermediate scale between nearby regions like Orion that are dominated by a single star, and giant star formation regions such as 30 Doradus.

#### C3.1 Observed structure function and model fit

Figure 5d shows the structure function of the Carina nebula. The model fit gives  $\sigma_{\text{pos}}^2 = (18 \pm 2) \text{ km}^2 \text{ s}^{-2}$ ,  $r_0 = (0.6 \pm 0.1)$  pc, and  $m = 1.2 \pm 0.3$ . Again there is good agreement between the model value and empirically measured value of  $\sigma_{\text{pos}}^2$ , indicating that finite box effects are unimportant. In addition, both noise and seeing are found to be unimportant. Similarly to the Lagoon nebula, the structure function of Carina also shows a non-uniform and irregular behavior below the correlation length, which can again be ascribed to the sparse spatial sampling. As a result of this, the power-law slope determination is not very precise.

On scales larger than the correlation length, the structure function of the Carina nebula is remarkably flat for  $r = 2$  to 6 pc, in good agreement with the prediction of homogeneous fluctuations (Figure 4a). But at the largest scales accessible to our observations, there is evidence for low-amplitude periodic fluctuations with a wavelength of  $\lambda \approx 8$  pc, which corresponds to the typical separation between the individual star clusters that make up the complex (for instance, between Tr 14 and Tr 16).

#### C3.2 Comparison with previous studies

We have found no previous studies of Carina's structure function in the literature.

### C4 30 Doradus

The giant H II region 30 Doradus (Tarantula Nebula) is located at a distance of 49.9 kpc ( $1'' = 0.24$  pc; Pietrzyński et al. 2013). It is a luminous star-forming region in the LMC (Large Magellanic Cloud) and is the most luminous complex in the Local Group (Kennicutt 1984). The star cluster that energizes the region is NGC 2070, with a central core (known as R136) with ionizing luminosity  $Q_0 = 2.5 \times 10^{51}$  photons  $\text{s}^{-1}$  (Bestenlehner et al. 2020), which produces an H $\alpha$  luminosity of  $2.9 \times 10^{39}$  erg  $\text{s}^{-1}$  (200 times the Orion Nebula) for the central ionized region of size 40 pc. The mean non-thermal linewidth in H $\alpha$  is  $22 \text{ km s}^{-1}$  (Torres-Flores et al. 2013b). Star formation is still ongoing in NGC 2070 (Walborn et al. 2013), although the star-formation rate peaked about 4 Myr ago (Cignoni et al. 2015) and it currently hosts some of the most massive stars known (Kalari et al. 2022) and several Wolf-Rayet stars (Evans et al. 2011). This nebula is commonly used as a comparison for extragalactic star-forming regions.

#### C4.1 Observed structure function and model fit

The 30 Doradus structure function is shown in Figure 7a. The model-derived parameters are  $\sigma_{\text{pos}}^2 = (297 \pm 40) \text{ km}^2 \text{ s}^{-2}$ ,  $r_0 = (4 \pm 1)$  pc, and  $m = 0.85 \pm 0.1$ . The empirically-derived variance is about 15% smaller, indicative of a moderate finite-box effect. On either measure, the region is a clear outlier in our sample, with a value of  $\sigma_{\text{pos}}^2$  that is more than three times larger than any other of our regions, as can be graphically appreciated from the velocity histograms in Figure 3.

At scales larger than the correlation length, the structure function has a maximum at  $r = 20$  pc and then falls at the largest scales. This behavior is similar to that seen in the Orion Nebula and might likewise be evidence for a decline of the fluctuation amplitude towards the outskirts. On the other hand, the sharp upward turn towards the maximum in  $B(r)$  is more reminiscent of the expectations for a periodic fluctuation (Figure 4b), in which case the wavelength would be of order the size of the map. Evidence for such a pattern can be seen directly in the velocity map (Figure 2) as a red-blue asymmetry between the south-east and north-west, which is also seen in the kinematics of the O stars (Hénault-Brunet et al. 2012) and has been interpreted as evidence for rotation of the cluster.

#### C4.2 Comparison with previous studies

Feast (1961) investigated the structure function in 30 Doradus by using spectra at  $N = 37$  different points, spread over a region of diameter 150 pc and with minimum separation of 4 pc. Velocities were measured by combining results from multiple spectral lines in the blue spectral region, the strongest being of [O II], [O III], and H $\beta$ . They found an essentially flat  $B(r)$  for  $r > 10$  pc with  $\sigma_{\text{pos}}^2 = 130 \text{ km s}^{-1}$ , which is roughly half the value that we determine for the central region of the nebula (of diameter 30 pc). This implies that the amplitude of the fluctuations is falling with radius, similar to what we inferred above for the Orion Nebula. For the smallest separations, of order 4 pc, they found that  $B(r)$  declined significantly, although an exact measurement was hampered by uncertainties in the correction for the effects of noise. This behavior is consistent with our own value of  $r_0 = 4$  pc, suggesting that the correlation length is not greatly different in the wider area covered by their study. This is a clear difference from the Orion Nebula, where we find a significantly larger correlation length in the extended nebula than in the core.

Melnick et al. (1987) used Fabry-Perot observations in H $\alpha$  and [O III] from Smith & Weedman (1972) to calculate the structure

function over a region of diameter 50 pc, which is smaller than in Feast (1961) but with a slightly larger number of spatial samples ( $N = 44$ ). They found an approximately flat  $B(r)$  for  $r > 20$  pc and a shallow rising power law, index  $\approx 0.16$ , for  $r = 5$  to 20 pc, with  $\sigma_{\text{pos}}^2 \approx 280 \text{ km s}^{-1}$ . This is similar to our own results over the same range of separations, including the local peak in  $B(r)$  at  $r = 20$  pc, which is clearly visible in Figure A2 of Melnick et al. (1987). The reason that Melnick et al. measure such a shallow slope for their smallest separations is that they are only sampling the range from  $1.25r_0 \rightarrow 5r_0$ , in which the structure function is tending toward its asymptotic value of  $2\sigma_{\text{pos}}^2$ . In order to measure the power-law portion of the structure function, it would be necessary to sample separations as small as 1 pc or 4 arcsec.

Melnick et al. (2021) used H $\alpha$  spectra from VLT FLAMES/GIRAFFE (see section 2.1.3) for  $N = 1668$  distinct fiber positions covering a region of diameter 300 pc (Torres-Flores et al. 2013a). They claim to find a completely flat structure function on scales from 3 to 200 pc, which seriously conflicts with our own observations at the lower end of this range. In particular, their two smallest separation bins are centered on  $r = 2.5$  and 7.5 pc, for which our own results yield  $B(r)/\langle\sigma_{\text{los}}\rangle^2 \approx 0.7$  and 1.3, respectively (our Figure 7a), whereas they find  $B(r)/\langle\sigma_{\text{los}}\rangle^2 \approx 2.0$  at all separations (see their Figure 15). We are at a loss to explain this discrepancy, which is inconsistent not only with our own results, but also with those of Feast (1961) and Melnick et al. (1987), who both found at least some evidence for a downturn in  $B(r)$  for  $r < 10$  pc. The result also seems to contradict the evidence of their own Figures 7 and 13, which show spatial coherence in the velocity field. If the structure function were truly flat, then blue shifted and red shifted profiles would be randomly mixed spatially, which is not observed.

## C5 NGC 346

NGC 346 is the most active star-formation region in the SMC (Small Magellanic Cloud) located at a distance of 62 kpc ( $1'' = 0.30$  pc ; Dolphin et al. 2001). It has more than 30 O stars that ionize N66, the largest H II region (De Marchi et al. 2011). The cluster has an ionizing output  $Q_0$  of  $5.1 \times 10^{50}$  photons s $^{-1}$  (40 times higher than the Orion Nebula) and the H II region has an H $\alpha$  luminosity of  $5.9 \times 10^{38}$  erg s $^{-1}$  (Heydari-Malayeri & Selier 2010; Kennicutt 1984). A supernova remnant is present in the region (J59.4–7210, Ye et al. 1991) but it does not overlap with the MUSE field that we observe here (Maggi et al. 2019). High-resolution H $\alpha$  spectroscopy of the H II region gives a non-thermal velocity width of  $\langle\sigma_{\text{los}}\rangle = 10.1 \text{ km s}^{-1}$  (Danforth et al. 2003). Gouliermis et al. (2008) propose an expanding H II region or bubble blown by the winds of the massive progenitor as a mechanism that shapes the recent star formation in this region, in addition to the photoionization from the OB stars. This mechanism is similar to shell-like H II regions, with a central cluster in a cavity, and with ongoing star formation triggered around their periphery.

### C5.1 Observed structure function and model fit

Figure 7b shows the structure function of NGC 346. The model-derived parameters are  $\sigma_{\text{pos}}^2 = (33 \pm 3) \text{ km}^2 \text{ s}^{-2}$ ,  $r_0 = (2.4 \pm 0.3)$  pc, and  $m = 0.95 \pm 0.07$ . These agree very well with the empirically derived parameters. At large scales, an oscillatory behavior in  $B(r)$  is observed for  $r > 10$  pc, consistent with a periodic pattern with wavelength  $\lambda \approx 15$  pc and amplitude  $6 \text{ km}^2 \text{ s}^{-2}$ . Inspection of the velocity field (Figure 2) shows that this pattern is predominantly oriented along the north-south direction and corresponds to the spacing between the filaments of molecular gas.

### C5.2 Comparison with previous studies

We have found no previous studies of NGC 346's structure function in the literature.

## C6 Hubble X and Hubble V in NGC 6822

The irregular dwarf galaxy, NGC 6822, is located at a distance of 500 kpc ( $1'' = 2.42$  pc; Sibbons et al. 2012) and hosts the H II regions Hubble X and Hubble V which are the brightest regions in the galaxy (Hodge et al. 1989).

Hubble V is located near the “Hodge OB 8” OB association with a number of  $\sim 36$  young stellar objects (Kinson et al. 2021) and its core is powered by an O3 V star (O'Dell et al. 1999). It has an ionizing output  $Q_0$  of  $1.7 \times 10^{50}$  photons s $^{-1}$  (13 times higher compared to the Orion Nebula) and has a H $\alpha$  luminosity of  $2 \times 10^{38}$  erg s $^{-1}$  (Bosch et al. 2002). The mean non-thermal linewidth in H $\alpha$   $\langle\sigma_{\text{los}}\rangle = (10.30 \pm 0.03) \text{ km s}^{-1}$  (Hippelein 1986).

Hubble X is located near the Hodge “OB 13” OB association with a number of  $\sim 29$  young stellar objects (Kinson et al. 2021) and its core is powered by an O4 V star (O'Dell et al. 1999). It has an ionizing output  $Q_0$  of  $1.4 \times 10^{50}$  photons s $^{-1}$  (10 times higher compared to the Orion Nebula) and has a H $\alpha$  luminosity of  $1.6 \times 10^{38}$  erg s $^{-1}$  (Bosch et al. 2002). The mean non-thermal linewidth in H $\alpha$  is  $\langle\sigma_{\text{los}}\rangle = (10.50 \pm 0.02) \text{ km s}^{-1}$  (Hippelein 1986). Tomita et al. (1993) studies the kinematics of both regions.

### C6.1 Observed structure function and model fit

In Figure 8a we show the structure function of Hubble V. The model-derived parameters are  $\sigma_{\text{pos}}^2 = (10 \pm 3) \text{ km}^2 \text{ s}^{-2}$ ,  $r_0 = (3.6 \pm 1.0)$  pc, and  $m = 0.8 \pm 0.3$ . The velocity variance is slightly higher than the empirically derived value of  $\sigma_{\text{pos}}^2 = 8 \text{ km}^2 \text{ s}^{-2}$  because the model-derived seeing is only 6 times smaller than the correlation length.

At the largest scales, the structure function turns downward, as seen in Orion. However, inspection of the velocity map suggests that the explanation may be different in this source: the velocity varies more strongly across the short axis of the nebula than along its long axis. There is also marginal evidence for weak periodic fluctuations with wavelength  $\lambda \approx 50$  pc.

The Hubble X structure function is shown in Figure 8b. The model-derived parameters are  $\sigma_{\text{pos}}^2 = (15 \pm 3) \text{ km}^2 \text{ s}^{-2}$ ,  $r_0 = (4.0 \pm 0.5)$  pc, and  $m = 1.0 \pm 0.2$ , which again is slightly higher than the empirically derived value of  $\sigma_{\text{pos}}^2 = 13 \text{ km}^2 \text{ s}^{-2}$  because the latter is reduced by the seeing.

At the largest scales the structure function is approximately flat, consistent with homogeneous fluctuations.

### C6.2 Comparison with previous studies

We have found no previous studies of the structure functions of Hubble V or X in the literature.

## C7 NGC 595 and NGC 604 in M33

M33 is the third largest galaxy in the Local Group, located at a distance of 820 kpc ( $1'' = 4.07$  pc; Kam et al. 2015) and is of morphological type SA(s)cd. The two brightest H II regions in the galaxy are NGC 604 and NGC 595.

NGC 595 is the second most luminous giant H II region in the M33 galaxy, excited by a stellar population comprising 250 OB stars, 13

hot supergiants, and 11 Wolf-Rayet stars (Drissen et al. 1993; Malumuth et al. 1996). The ionizing output  $Q_0$  is  $7.6 \times 10^{50}$  photons s $^{-1}$  (60 times higher compared to the Orion Nebula) and the H II region has a luminosity in H $\alpha$  of  $8.9 \times 10^{38}$  erg s $^{-1}$  (Bosch et al. 2002). The mean non-thermal linewidth in H $\alpha$  is  $\langle \sigma_{\text{los}} \rangle = (16.8 \pm 0.1)$  km s $^{-1}$  (Lagrois & Joncas 2009). Lagrois & Joncas (2011) present a kinematic study of the regions using optical emission lines and radio 21 cm line observations and find an H $\alpha$  shell morphology that is related to the stellar winds of the several massive stars located in its interior. They proposed champagne flows at the periphery of the molecular cloud leading to accelerated ionized material.

NGC 604 is the most luminous H II region in M33 and the second most luminous in the Local Group after 30 Dor. The dominant stellar cluster is located in the center of the region and contains 200 OB stars and 4 Wolf-Rayet stars (Hunter et al. 1996; Eldridge & Relaño 2011). The optically visible stars come from a population with an age of 4 Myr, although infrared observations detect the presence of ongoing massive star formation in the region (Fariña et al. 2012; Martínez-Galarza et al. 2012). The ionizing output is  $Q_0 = 2.2 \times 10^{51}$  photons s $^{-1}$  (175 times higher than the Orion Nebula) and the region has an H $\alpha$  luminosity of  $2.6 \times 10^{39}$  erg s $^{-1}$  (Bosch et al. 2002). The mean non-thermal linewidth in H $\alpha$  is  $\langle \sigma_{\text{los}} \rangle = (17.8 \pm 0.3)$  km s $^{-1}$  (Hippelein 1986). The morphology of the region is strikingly similar to that of 30 Dor (see Figure 1) and is visually dominated by a predominant big loop, many expanding shells and different size filaments.

### C7.1 Observed structure function and model fit

In Figure 8c we show the structure function of NGC 595. The model-derived parameters are  $\sigma_{\text{pos}}^2 = (53 \pm 5)$  km $^2$  s $^{-2}$ ,  $r_0 = (11.0 \pm 1.0)$  pc, and  $m = 1.36 \pm 0.15$ . The velocity variance is slightly higher than the observed value of  $44$  km $^2$  s $^{-2}$ , indicative of a moderate finite box effect. At the largest scales, the structure function trends slowly downwards, which may be due to a radial decline in the amplitude of the turbulent velocity field, but the effect is less marked than in other regions, such as Orion.

In Figure 8d we show the structure function of NGC 604. The model-derived parameters are  $\sigma_{\text{pos}}^2 = (84 \pm 20)$  km $^2$  s $^{-2}$ ,  $r_0 = (12 \pm 6)$  pc, and  $m = 0.8 \pm 0.2$ . Again the model-derived velocity variance is higher than the observed value of  $55$  km $^2$  s $^{-2}$ , which this time is due primarily to the fact that the derived correlation length is only 6 times higher than the RMS seeing width, so some of the small-scale fluctuation power is missing from the observations. Of all our sources, NGC 604 shows the largest difference between the inferred “true” structure function and the observed one, so the derived parameters should be treated with caution. At the largest scales, the structure function shows a pronounced minimum at  $r \approx 150$  pc, which may be due to a periodic oscillation with  $\lambda \approx 300$  pc, which is slightly larger than the size of the map. On the other hand, if we fit the model for scales  $r \leq 20$  pc, there is a decrease in the difference between the inferred “true” structure function and the observed one. With our current observations the fitting using less than half of the observational points cannot be justified; but with better quality observations and using our current model we would expect to derive more accurate parameters.

### C7.2 Comparison with previous studies

Lagrois & Joncas (2011) calculated the structure function for the H $\alpha$  emission line for NGC 595 over a region of diameter 300 kpc with

a resolution of 6.6 pc pixel $^{-1}$  (Lagrois & Joncas 2009), presenting results for separations  $r \geq 16.5$  pc. They found a total velocity variance of  $\sigma_{\text{pos}}^2 = (35 \pm 2)$  km $^2$  s $^{-2}$ , which is smaller than our derived value of  $\sigma_{\text{pos}}^2 = 53$  km $^2$  s $^{-2}$ . Our map covers a smaller region, so part of the reason for this discrepancy may be a radial decline in the velocity variance, as mentioned above. Another reason may be that the angular resolution of the Lagrois & Joncas (2009) map is of the same order as the correlation length, so some of the small-scale power is missing.

Lagrois & Joncas (2011) employ a novel technique based on filtering the bidimensional velocity field to remove large-scale non-turbulent motions before calculating the structure function. They find a total decorrelation length of  $\tau_0 = 43$  pc and a very steep power-law index of  $m = 1.55 \pm 0.10$ . As discussed in section 5.1,  $\tau_0$  is predicted to be 2–3 times larger than  $r_0$ , given the way that we define the correlation length, but even taking this into account, our derived value is still smaller than theirs by at least 50%. Although this may be due to the difference in the areas covered, it is more likely due to the lower resolution of the Lagrois & Joncas data.

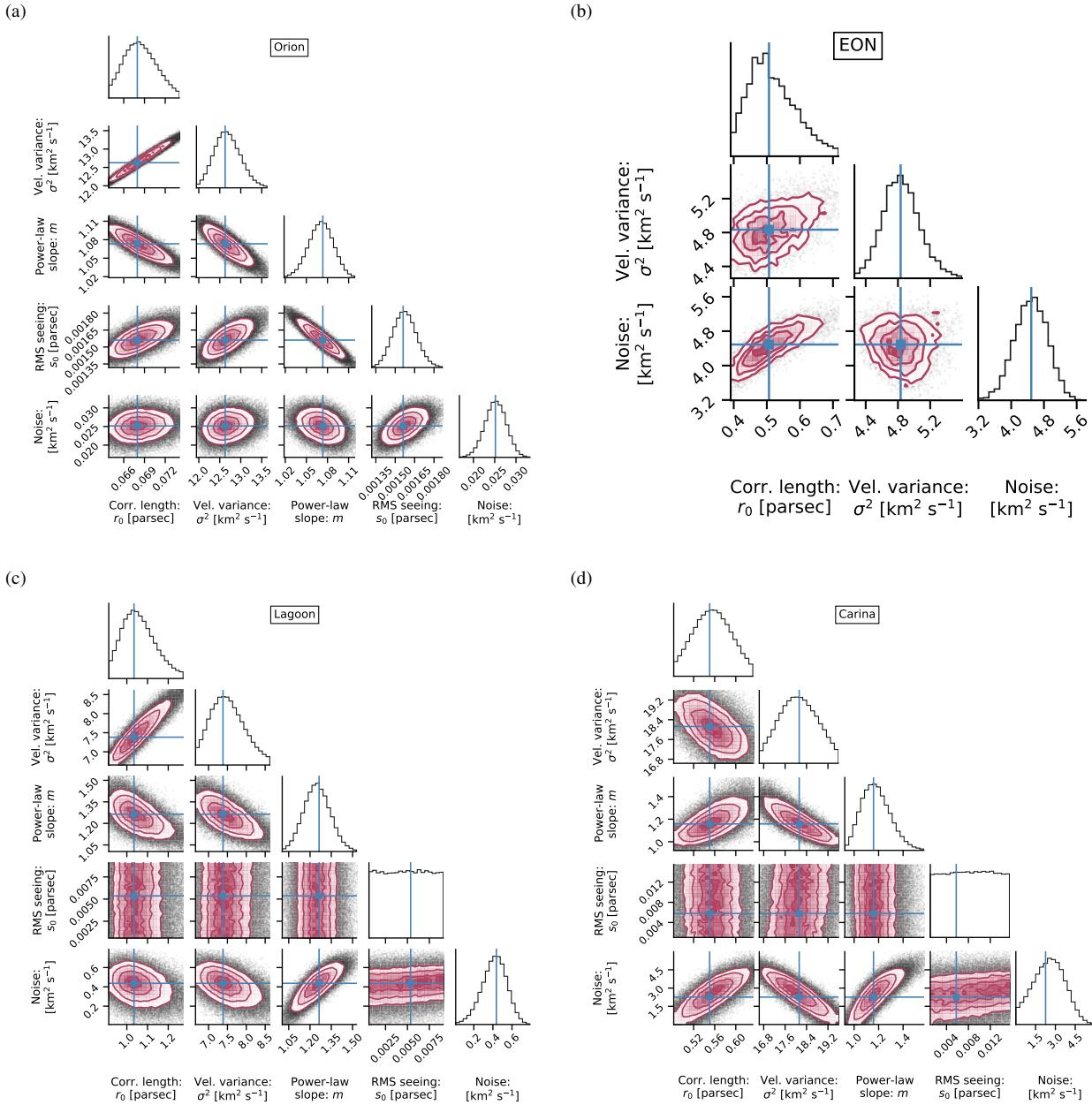
The power-law index is somewhat steeper than our own value, but in part this is due to differences in methodology: the width of the low-pass filter that Lagrois & Joncas use to remove large-scale motions is 8 pixels or 53 pc, which is comparable to the decorrelation length. As a result, subtracting the filtered data causes a clear steepening of the structure function, as can be seen in their Figure 6d. The unfiltered data can be seen to show a shallower slope, more consistent with our own results. Which result is more reliable is difficult to say, but could potentially be checked by experimenting on fake data in a similar fashion to our Appendix A.

All previous studies of the structure function in NGC 604 are based on the same TAURUS-II observational dataset that we use here (Sabalisk et al. 1995) and the results are unsurprisingly similar to our own, although there are some differences in methodology and interpretation. For instance, Medina Tanco et al. (1997) identify an apparent “knee” in the structure function at  $r \approx 10$  pc, which they interpret as evidence for a broken power law. However, in our own results the same feature is modeled as a combination of two effects: the natural flattening of the structure function for  $r > r_0$ , plus a steepening at small scales due to the smoothing effects of the seeing. Melnick et al. (2021) also analysed the same structure function and concluded that separations  $< 2$  arcsec should be discounted due to the seeing. The results of our experiments with synthetic observations (Appendix A2) suggest that the situation is even worse than Melnick et al. surmised and that scales up to 5 arcsec will be affected by the seeing. On the other hand, if the effects of seeing are incorporated into the model structure function, as we have done here, then the data at these separations do not need to be discarded entirely.

## APPENDIX D: ADDITIONAL COVARIANCE CORNER PLOTS FOR MODEL FITS

Supplementary online-only material.

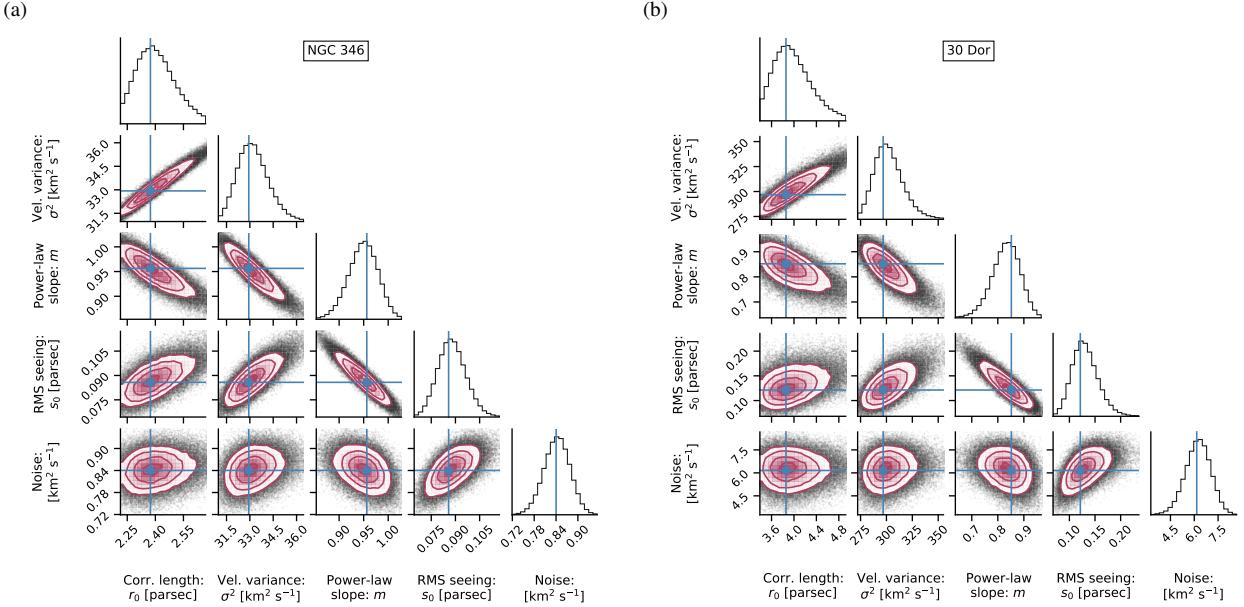
This paper has been typeset from a  $\text{\TeX}/\text{\LaTeX}$  file prepared by the author.



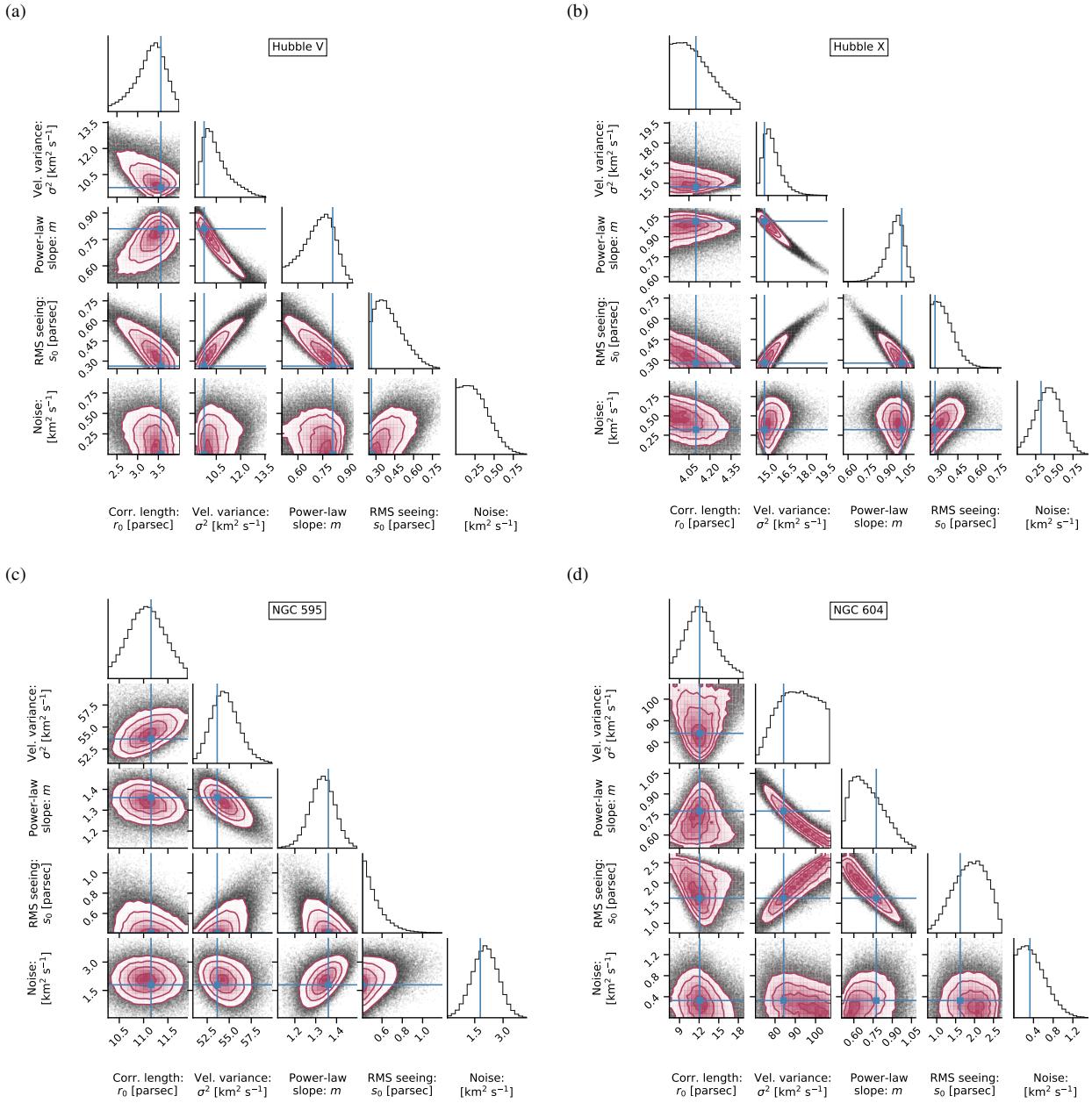
**Figure D1.** Corner plots of covariances between fitted model parameters of H $\alpha$  structure function for Galactic H II regions.

**Table D1.** Fitting parameters.

H II Region	No. points merged	Relatively uncertainty	Weight / No. points Small scales	Weight / No. points Large scales	Max. separation	Reduced chi-square	Autocorrelation time	Acceptance fraction
Orion	3	0.02	—	—	$0.5L$	0.82	60	0.55
EON	—	0.03	2.0/3	3.0/3	$0.7L$	0.62	40	0.64
Lagoon	2	0.08	2.5/16	2.5/8	$0.5L$	0.94	72	0.50
Carina	2	0.055	3.0/14	3.0/4	$0.5L$	0.94	72	0.50
30 Dor	5	0.07	—	—	$0.9L$	0.98	84	0.54
NGC 346	3	0.02	—	—	$0.5L$	0.83	64	0.54
Hubble V	—	0.05	2.0/3	—	$0.6L$	0.86	340	0.44
Hubble X	—	0.032	2.0/4	2.0/6	$0.5L$	0.73	79	0.49
NGC 595	—	0.055	4.5/3	3.0/5	$0.5L$	0.83	67	0.51
NGC 604	—	0.08	2.0/3	1.5/2	$0.5L$	0.92	112	0.45



**Figure D2.** Same as figure D1 except for Magellanic Cloud H II regions.



**Figure D3.** Same as figure D1 except for H II regions in more distant Local Group Galaxies.