Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **523,** 233–250 (2023) https://doi.org/10.1093/mnras/stad1450
Advance Access publication 2023 May 16

A statistical analysis of the structure of the interstellar medium in the disc of the Milky Way

M. L. Bates★ and A. P. Whitworth
School of Physics and Astronomy, Cardiff University, Cardiff CF24 3AA, UK

Accepted 2023 May 5. Received 2023 May 5; in original form 2022 November 10

ABSTRACT

We construct Convolutional Neural Networks (CNNs) trained on exponentiated fractional Brownian motion (xfBm) images, and use these CNNs to analyse Hi-GAL images of surface density in the Galactic Plane. The CNNs estimate the Hurst parameter, $\mathcal{H}$ (a measure of the power spectrum), and the scaling exponent, $\mathcal{S}$ (a measure of the range of surface densities), for a square patch comprising $[\mathcal{N} \times \mathcal{N}] = [128 \times 128]$, $[64 \times 64]$, or $[32 \times 32]$ pixels. The resulting estimates of $\mathcal{H}$ are more accurate than those obtained using Δ-variance. We stress that statistical measures of structure are inevitably strongly dependent on the range of scales they actually capture, and difficult to interpret when applied to fields that conflate very different lines of sight. The CNNs developed here mitigate this issue by operating effectively on small fields (small $\mathcal{N}$), and we exploit this property to develop a procedure for constructing detailed maps of $\mathcal{H}$ and $\mathcal{S}$. This procedure is then applied to Hi-GAL maps generated with the PPMAP procedure. There appears to be a bimodality between sightlines with higher surface density ($\gtrsim \mathbf{32\,M_\odot\,pc^{-2}}$), which tend to have higher $\mathcal{H}$ ($\gtrsim 0.8$) and $\mathcal{S}$ ($\gtrsim 1$); and sightlines intercepting regions of lower surface density ($\lesssim \mathbf{32\,M_\odot\,pc^{-2}}$), which tend to have lower $\mathcal{H}$ ($\lesssim 0.8$) and $\mathcal{S}$ ($\lesssim 1$); unsurprisingly the former sightlines are concentrated towards the Galactic Midplane and the Inner Galaxy. The surface density PDF takes the form $\mathrm{d}P/\mathrm{d}\Sigma \propto \Sigma^{-3}$ for $\mathbf{\Sigma \gtrsim 32\,M_\odot\,pc^{-2}}$, and on most sightlines this power-law tail is dominated by dust cooler than ~ 20 K, which is the median dust temperature in the Galactic Plane.

Key words: methods: data analysis – methods: statistical – stars: formation – ISM: clouds.

1 INTRODUCTION

The processes that regulate the evolution of the Galactic interstellar medium (ISM) – magnetohydrodynamic turbulence, self-gravity, non-LTE chemistry, radiation transport, etc. – are non-linear, and therefore the overall structure of the ISM is chaotic. However, the structure of the ISM appears to be approximately statistically self-similar over a wide range of length scales. Consequently many attempts have been made to estimate a fractal dimension, $\mathcal{D}_\mathcal{E}$ (e.g. Bazell & Desert 1988; Falgarone, Phillips & Walker 1991; Elmegreen 1997; Stutzki et al. 1998; Elia et al. 2014). Here $\mathcal{D}_\mathcal{E}$ is the fractal dimension of an image in $\mathcal{E}$-dimensional space, and is essentially a measure of how efficiently structures fill that space (Mandelbrot & Cannon 1984). The fractal dimension is a non-integer number, with possible values ranging from $\mathcal{D}_\mathcal{E} = \mathcal{E} - 1$ to $\mathcal{D}_\mathcal{E} = \mathcal{E}$ (e.g. the perimeter area dimension), or from $\mathcal{D}_\mathcal{E} = \mathcal{E}$ to $\mathcal{D}_\mathcal{E} = \mathcal{E} + 1$ (e.g. the box-counting dimension). It has the advantage that (i) it is straightforward to estimate $\mathcal{D}_2$ from a 2D image; (ii) if the 2D image is the projection of a statistically isotropic 3D field, and $\mathcal{D}_2$ is well defined, then it is likely that the underlying 3D field is also fractal, and its dimension, $\mathcal{D}_3$, can be constrained. In this paper, we consider 2D images of surface density, derived from dust optical-depth estimates.

Various methods have been adopted to estimate $\mathcal{D}_2$, in particular perimeter-area analysis of contoured images to estimate $\mathcal{D}_{2:\mathrm{PA}}$ (Bazell & Desert 1988; Dickman, Margulis & Horvath 1990; Falgarone et al. 1991; Williams, Blitz & McKee 2000; Marchuk et al. 2021), and box-counting analysis to measure $\mathcal{D}_{2:\mathrm{BC}}$ (Mandelbrot & Cannon 1984; Sanchez, Alfaro & Perez 2005; Federrath, Klessen & Schmidt 2009; Elia et al. 2018). These two fractal dimensions should be related by $\mathcal{D}_{2:\mathrm{PA}} = \mathcal{D}_{2:\mathrm{BC}} - 1$ (Voss 1988; Vogelaar & Wakker 1994; Stutzki et al. 1998).

Many of the above analyses have sought to determine a single fractal dimension that obtains over a large dynamic range. In other words they have assumed that the observed structures subscribe to a mono-fractal hierarchy. However, the turbulent cascade operating in the interstellar medium is complicated. First, turbulent energy is injected by many different processes, on many different length and time-scales, and frequently anisotropically. Secondly, the turbulent cascade appears to deliver objects with increasingly strong self-gravity, and evolving proportions of solenoidal and compressive modes, as it progresses to smaller length scales. Consequently the fractal dimension may be scale-dependent, and may vary with position.

A more general approach that admits this possibility is to assume that there is an ensemble of interwoven structures with different fractal dimensions, and hence that the overall structure is multifractal (e.g. Chappell & Scalo 2001; Elia et al. 2018; Robitaille et al. 2020b). This has the advantage that it delivers a more detailed description, and therefore facilitates more detailed intercomparison between different observed images (for example, different patches on the sky, or the same patch observed at different wavelengths), and between observed images and synthetic images (for example, the results of numerical simulations). However, the basic products of

★ E-mail: matthew.bates@astro.cf.ac.uk

such analyses are rather abstract, and usually involve distributions of parameters. These distributions can sometimes be reduced to a few numbers, but the interpretation of those numbers in terms of physics is usually still quite difficult, as acknowledged in Chappell & Scalo (2001). We give an example in Appendix A.

In addition, these methods can in general only be applied to regular rectangular images comprising a large number of pixels. This means it is likely that the user is analysing a combination of regions that have very different intrinsic structures, and are being combined in unknown proportions within an arbitrary frame. For example, the analyses in Elia et al. (2018) treat most of a single Hi-GAL tile (see Molinari et al. (2010) and Section 4 for further details of the Hi-GAL survey). Therefore they include lines of sight through dense star-forming regions, lines of sight through very diffuse gas, and lines of sight through regions between these two extremes.

An alternative method for quantifying statistically the structure of a 2D image is to estimate its Δ-variance (Stutzki et al. 1998). Δ-variance analysis assumes that the underlying structure subscribes, at least approximately, to a fractional Brownian motion (fBm) distribution (Peitgen & Saupe 1988; Stutzki et al. 1998; Elia et al. 2014). It involves computing the variance, $\sigma^2_\Delta(L)$, of the image after it has been convolved with an isotropic filter, $\odot_L$, of length scale L. The Hurst parameter of the image, $\mathcal{H}$, is then given by

$$2\,\mathcal{H} = \frac{\mathrm{dln}\left(\sigma^2_\Delta\right)}{\mathrm{dln}(L)} \;=\; \beta \;-\; 2, \tag{1}$$

where β is the index of the power spectrum. Basically $\mathcal{H}$ measures how smooth the underlying structure is; large $\mathcal{H}$ means that the image is dominated by extended structures; small $\mathcal{H}$ means that the image is dominated by compact structures. The corresponding perimeter-area and box-counting fractal dimensions for a 2D image are $\mathcal{D}_{2:\mathrm{PA}} = 2 - \mathcal{H}$ and $\mathcal{D}_{2:\mathrm{BC}} = 3 - \mathcal{H}$.

fBm images have both negative and positive values. To create synthetic images that can be compared with real surface density images (which, apart from noise, are everywhere positive-valued) we follow Elmegreen, Kim & Staveley-Smith (2001) and generate an fBm image, then exponentiate this fBm image using a scaling parameter $\mathcal{S}$, thereby creating an exponentiated fractional Brownian motion (xfBm) image. These xfBm images have a log-normal distribution of surface density (cf. Robitaille et al. 2020b). The width of the log-normal distribution is regulated by $\mathcal{S}$ (see equation 2); large $\mathcal{S}$ gives a broad distribution of surface density, and small $\mathcal{S}$ a narrow one.

Indeed, we would argue that two parameters may be the minimum required to characterize the statistics of a 2D image: one (here $\mathcal{H}$) to describe the degree to which the contours of the image are convoluted; and one (here $\mathcal{S}$) to describe the range of values represented by the contours. (Other parameters might reflect the degree of 'nestedness', and any intrinsic anisotropy.)

In the sequel, synthetic xfBm images are used to train a Convolutional Neural Network (CNN), which is then used to analyse Hi-GAL images of the surface density in the Galactic Plane, in terms of the distributions of Hurst parameter, $\mathcal{H}$, scaling parameter, $\mathcal{S}$, and surface density Σ. The plan of the paper is as follows.

In Section 2 we describe how we generate xfBm images with different combinations of $\mathcal{H}$ and $\mathcal{S}$. In Section 3 we explain how we use these images to train efficient Convolutional Neural Networks (CNNs) that return reliable estimates of $\mathcal{H}$ and $\mathcal{S}$, and we compare the performance of these CNNs with the performance of Δ-variance. In Appendix B we explain why the Hurst parameter is a valid descriptor for an xfBm field. In Section 4 we describe the high-resolution Hi-GAL images of surface density to which we apply the CNNs from Section 3; the Point Process Mapping (PPMAP) technique used to obtain these images is outlined in Appendix C. In Section 5 we describe the procedure for generating detailed maps of $\mathcal{H}$ and $\mathcal{S}$, emphasizing the issues that derive from the finite range of angular scales captured. In Section 6 we illustrate the procedure by applying it to the three Hi-GAL tiles closest to the Galactic Centre, and analyse the results. In Section 7 we apply the procedure to the whole Galactic Plane, and discuss the resulting statistics. In Section 8 we summarize the main conclusions.

We adopt $\mathcal{H}$ and $\mathcal{S}$ to characterize surface density structures, because they are parameters with a straightforward visual meaning (as illustrated in Section 2, Fig. 1). We are not challenging the possibility that Hi-GAL images may be more accurately characterized as multifractal. We are simply exploring an alternative procedure that has merit by virtue of its ability (i) to work on very small patches of an image, and (ii) to deliver metrics that, although essentially monofractal, admit a relatively simple interpretation.

2 CONSTRUCTING EXPONENTIATED FRACTIONAL BROWNIAN MOTION IMAGES

The procedure used to generate square 2D xfBm images is described in detail in Bates, Whitworth & Lomax (2020). Here we summarize briefly the five stages involved. Square 2D xfBm images are characterized by four parameters: the Hurst parameter, $\mathcal{H}$, the scaling parameter, $\mathcal{S}$, the number of pixels along one side of the square, $\mathcal{N}$, and a random seed, $\mathcal{R}$. For this study we explore $0 \le \mathcal{H} \le 1$, $0 \le \mathcal{S} \le 3$, and $\mathcal{N} = 32,\ 64\,,$ or 128.

In Stage 1, we construct a periodic fBm image $f(\mathcal{H}{:}\boldsymbol{r})$ for the specified $\mathcal{H}$, using the spectral synthesis method introduced in Peitgen & Saupe (1988), by first generating a power spectrum, $\hat{f}(\mathcal{H}{:}\boldsymbol{r})$, and then performing an inverse discrete Fourier transform. Here $\boldsymbol{r}$ takes integer values in the range $1 \le \boldsymbol{r} \le 4\mathcal{N}$.

In Stage 2, the fBm image is exponentiated, using the specified scaling parameter, $\mathcal{S}$,

$$g(\mathcal{HS}{:}\boldsymbol{r}) = \exp\left\{\frac{\mathcal{S}\, f(\mathcal{H}{:}\boldsymbol{r})}{\left\langle f^2(\mathcal{H}{:}\boldsymbol{r})\right\rangle^{1/2}}\right\}. \tag{2}$$

This transforms the normally distributed fBm image into a lognormal image, thereby ensuring that the image is everywhere positive.

In Stage 3, we cut out the central $\mathcal{N}\times\mathcal{N}$ patch, so as to have a non-periodic image, $g(\mathcal{HS}{:}\boldsymbol{r})$. We do this because the real images that we wish to mimic are not periodic.[1]

In Stage 4, we add random noise to each pixel. The noise values are uniformly distributed between 0 per cent and 5 per cent of the standard deviation, σ_g, of the xfBm image, $g(\mathcal{HS}{:}\boldsymbol{r})$.

In Stage 5, extreme pixels are filtered out. Specifically, any pixel with $g > \mu_g + 2.5\sigma_g$ is replaced with $\mu_g + 2.5\sigma_g$, where μ_g is the mean of the xfBm image.

Fig. 1 shows nine xfBm images generated with different combinations of $\mathcal{H}$ and $\mathcal{S}$ but the same random seed. Consequently the large-scale structure has the same features in all images. All the images are rendered with the same surface density colour-table.

[1] In Bates et al. (2020) we first shifted the centre of mass of the periodic image to the geometric centre, and then cut out the central $\mathcal{N}\times\mathcal{N}$ patch. This was because the procedure developed there was designed to analyse individual molecular clouds and star clusters. Here, we are concerned with arbitrarily positioned patches, and so this centring step is not appropriate. We therefore refer to the images generated here as *un-centred*. This is the only fundamental difference from the images presented in Bates et al. (2020).

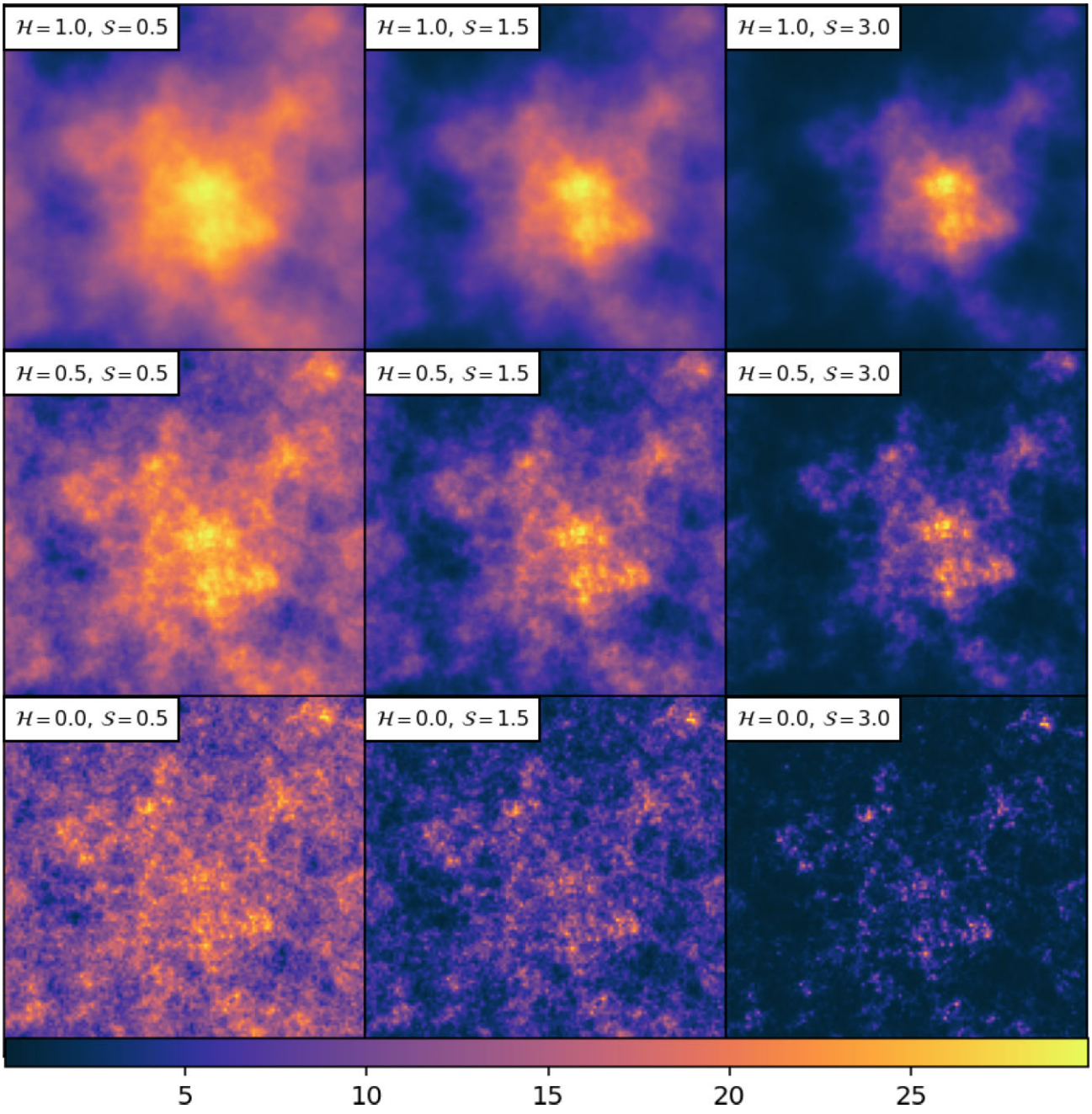

Figure 1. The surface densities of square xfBm images generated using different combinations of $\mathcal{H}$ and $\mathcal{S}$ but the same seed – so that the same structures are visible on all nine images, and the differences between the images are entirely due to differences in $\mathcal{H}$ and $\mathcal{S}$. In order to make this more obvious, we have centred the images on the highest density peak, but the images used for training the CNNs are uncentred. All images have 128×128 pixels. The images in each column have the same value of $\mathcal{S}$, but $\mathcal{H}$ decreases from top to bottom and therefore the amount of small-scale structure increases from top to bottom. The images in each row have the same value of $\mathcal{H}$, but $\mathcal{S}$ increases from left to right and therefore the range of surface densities increases from left to right. The colour represents the cube-root of the surface density, in order to capture the large dynamic range. The scale of the colour bar is arbitrary, but the same for all images.

The three images on the top row all have $\mathcal{H} = 1$, so the power is concentrated in the long wavelength modes and there is little small-scale structure. Conversely, the three images on the bottom row all have $\mathcal{H} = 0$, so the power is concentrated in the short wavelength modes and there is lots of well-defined small-scale structure. (The same trend is also seen in fig. 5 of Dib et al. (2020), except that they use β to characterize the power spectrum, instead of $\mathcal{H}$, see equation 1). Likewise, the three images in the left-hand column all have $\mathcal{S} = 0.5$, so there is a rather small range of surface density and little contrast in the image, whilst the three images in the right-hand column have $\mathcal{S} = 3.0$, a large range of surface density, and therefore high contrast.

These images are computationally cheap to generate. On a personal computer, ∼65 128×128-pixel images, or ∼130 64×64-pixel images, or ∼550 32×32-pixel images, can be generated in one second. However, given an xfBm image, it is more complicated

Table 1. The architectures of CNN:128, CNN:64, and CNN:32. CNN:$\mathcal{N}$ takes a normalized $\mathcal{N} \times \mathcal{N}$ pixel image and outputs estimates for $\mathcal{HS}$, by passing the image through $\log_2(\mathcal{N}) - 2$ convolutional layers and five fully connected layers. Each convolutional layer outputs the convolution of a 3×3 window with a trainable kernel, and is followed by a max pooling layer that outputs the maximum value of a 2×2 sliding window with a step of 2. For each operation we give the number of trainable parameters.

Layer	Operation	CNN:128 Output size	Trainable Params.	CNN:64 Output size	Trainable Params.	CNN:32 Output size	Trainable Params.
Input	input layer	$128 \times 128 \times 1$	0	$64 \times 64 \times 1$	0	$32 \times 32 \times 1$	0
Convolutional 1	3×3 kernel	$126 \times 126 \times 256$	2560	$62 \times 62 \times 256$	2560	$30 \times 30 \times 256$	2560
MaxPooling 1	2×2 max-pool	$63 \times 63 \times 256$	0	$31 \times 31 \times 256$	0	$15 \times 15 \times 256$	0
Convolutional 2	3×3 kernel	$61 \times 61 \times 256$	590 080	$29 \times 29 \times 256$	590 080	$13 \times 13 \times 256$	590 080
MaxPooling 2	2×2 max-pool	$30 \times 30 \times 256$	0	$14 \times 14 \times 256$	0	$6 \times 6 \times 256$	0
Convolutional 3	3×3 kernel	$28 \times 28 \times 256$	590 080	$12 \times 12 \times 256$	590 080	$4 \times 4 \times 256$	590 080
MaxPooling 3	2×2 max-pool	$14 \times 14 \times 256$	0	$6 \times 16 \times 256$	0	$2 \times 2 \times 256$	0
Convolutional 4	3×3 kernel	$12 \times 12 \times 256$	590 080	$4 \times 4 \times 256$	590 080	–	–
MaxPooling 4	2×2 max-pool	$6 \times 6 \times 256$	0	$2 \times 2 \times 256$	0	–	–
Convolutional 5	3×3 kernel	$4 \times 4 \times 256$	590 080	–	–	–	–
MaxPooling 5	2×2 max-pool	$2 \times 2 \times 256$	0	–	–	–	–
Flatten	to 1D layer	$1 \times 1 \times 1024$	0	$1 \times 1 \times 1024$	0	$1 \times 1 \times 1024$	0
Dense 1	fully connected	$1 \times 1 \times 512$	262 400	$1 \times 1 \times 256$	262 400	$1 \times 1 \times 256$	262 400
Dense 2	fully connected	$1 \times 1 \times 512$	65 792	$1 \times 1 \times 256$	65 792	$1 \times 1 \times 256$	65 792
Dense 3	fully connected	$1 \times 1 \times 512$	65 792	$1 \times 1 \times 256$	65 792	$1 \times 1 \times 256$	65 792
Dense 4	fully connected	$1 \times 1 \times 512$	65 792	$1 \times 1 \times 256$	65 792	$1 \times 1 \times 256$	65 792
Dense 5	fully connected	$1 \times 1 \times 512$	65 792	$1 \times 1 \times 256$	65 792	$1 \times 1 \times 256$	65 792
Output	a channel each for $\mathcal{H}$ and $\mathcal{S}$	$1 \times 1 \times 2$	514	$1 \times 1 \times 2$	514	$1 \times 1 \times 2$	514
		–	–	–	–	–	–
Totals			2 888 962		2 298 882		1 708 802

to estimate $\mathcal{H}$ and/or $\mathcal{S}$. The Δ-variance procedure introduced by Stutzki et al. (1998) and refined by Ossenkopf, Krips & Stutzki (2008), is able to estimate $\mathcal{H}$ quite accurately, provided the power spectrum has sufficient dynamic range, say $\mathcal{N} \gtrsim 128$, but it is not designed to estimate $\mathcal{S}$.

Strictly speaking, the Hurst parameter, $\mathcal{H}$ (or equivalently, the power-spectrum exponent, $\beta = 2[1 + \mathcal{H}]$) is a property of the underlying 2D fBm image, rather than the final xfBm image. However, the power-spectrum basically reflects the amount of structure on different spatial scales (rather than the absolute heights of those structures), and this is not greatly changed by exponentiation. This was shown by Bates et al. (2020) (see their section 3.3 and Fig. 4). For pure 128×128 fBm fields, Δ-variance returned the correct β to within 1 per cent. If the images were rendered non-periodic, the uncertainty increased to 9 per cent, but the mean was still very close to the initial value. And if the images were then exponentiated, the uncertainty increased to 18 per cent, but again the mean was still very close to the initial value. Stutzki et al. (1998) obtained a very similar result. They took an fBm field and squared it (an alternative way to render the field positive definite), thereby producing an 'fBm^2' field. Then, by applying Δ-variance to the fBm^2 field, they retrieved a β value very close to the one used to generate the original fBm field. We speculate further on the reasons for this in Appendix B.

3 CONVOLUTIONAL NEURAL NETWORKS FOR ANALYSING XFBM IMAGES

In this section we describe the procedure used to train CNN:128, CNN:64, and CNN:32 (where CNN:$\mathcal{N}$ is a CNN applicable to a square image with $\mathcal{N} \times \mathcal{N}$ square pixels). The architectures of CNN:128, CNN:64, and CNN:32 are presented in Table 1. The architecture of CNN:128 is similar to the 128×128 CNN developed in Bates et al. (2020), but the dimensions of the convolutional and max-pooling layers are different. The reader is referred to Bates et al. (2020) for an explanation of what the different layers do.

For each CNN, we execute 100 training epochs, and in each training epoch we use 20 000 newly generated and unique synthetic xfBm images (thus 2000 000 different images in total). The xfBm images have random values of $\mathcal{H}$ distributed uniformly on the interval $0 < \mathcal{H} < 1$, and random values of $\mathcal{S}$ distributed uniformly on the interval $0 < \mathcal{S} < 3$.

During a training epoch, these images are fed through the CNN in batches of 32, and the training loss-function of the batch,

$$\mathcal{L}_{\text{BATCH TRAINING}} = \left\langle \left[F\left(\boldsymbol{x}_{\text{INPUT}}\right) - \boldsymbol{y}_{\text{KNOWN}} \right]^2 \right\rangle_{32}, \tag{3}$$

is computed. In equation (3), the elements of $\boldsymbol{y}_{\text{KNOWN}}$ are the known values of $\mathcal{H}$ and $\mathcal{S}$ for the batch, and the elements of $F\left(\boldsymbol{x}_{\text{INPUT}}\right)$ are the corresponding estimates of $\mathcal{H}$ and $\mathcal{S}$. After the processing of each batch, the parameters of the CNN are tuned to minimize $\mathcal{L}_{\text{BATCH TRAINING}}$, using the method of gradient descent.

At the end of each training epoch, a further 200 new synthetic xfBm images are generated and the validation loss-function,

$$\mathcal{L}_{\text{EPOCH VALIDATION}} = \left\langle \left[F\left(\boldsymbol{x}_{\text{INPUT}}\right) - \boldsymbol{y}_{\text{KNOWN}} \right]^2 \right\rangle_{200}, \tag{4}$$

is computed and monitored. By virtue of using unique images at each epoch, $\mathcal{L}_{\text{EPOCH VALIDATION}}$ decreases more-or-less monotonically, albeit increasingly slowly.

After 100 epochs, the training and validation cycle is terminated (because $\mathcal{L}_{\text{EPOCH VALIDATION}}$ is no longer decreasing significantly) and the accuracy of the CNN is evaluated. For this purpose, 10 000 new $\mathcal{N} \times \mathcal{N}$ xfBm images are generated (again with random values of $\mathcal{H}$ distributed uniformly on the interval $0 < \mathcal{H} < 1$, and random values of $\mathcal{S}$ distributed uniformly on the interval $0 < \mathcal{S} < 3$). These images are then fed through the CNN and the resulting estimates of $\mathcal{H}$ and $\mathcal{S}$, recorded.

Fig. 2 shows the results of these evaluations. The results for CNN:128 are shown in the left-hand column, those for CNN:64 in the middle column, and those for CNN:32 in the right-hand column. The panels on the top row show estimated $\mathcal{H}$ plotted against true

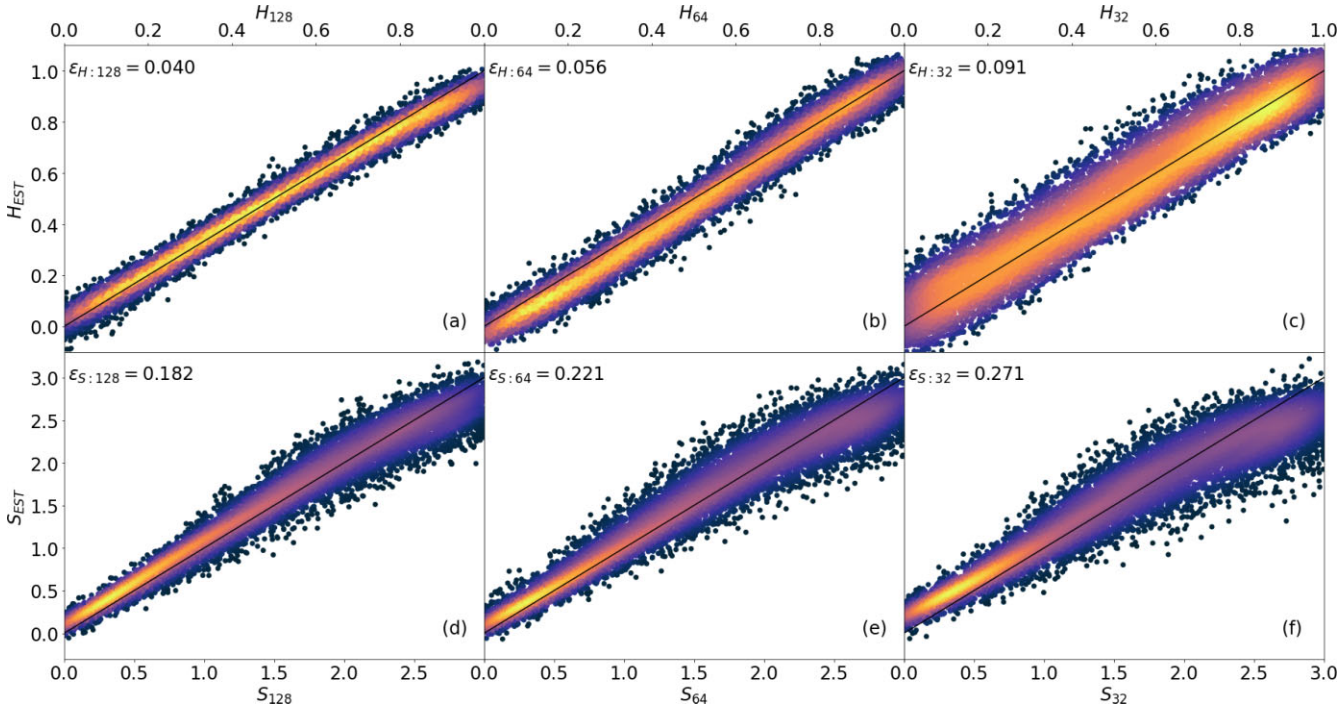

Figure 2. The fidelity of the different CNNs. The panels on the top row display estimated $\mathcal{H}$ against true $\mathcal{H}$ for 10 000 different random combinations of true $\mathcal{H}$ (distributed uniformly on the interval [0,1]) and true $\mathcal{S}$ (distributed uniformly on the interval [0,3]). The panels on the bottom row display the corresponding values of estimated $\mathcal{S}$ against true $\mathcal{S}$. The panels in the left-hand column show the results for CNN:128; those in the middle column, CNN:64; and those in the right-hand column, CNN:32. In each panel the black line represents exact correspondence, and the root-mean-square error is given in the top left-hand corner.

Table 2. The root-mean-square errors for CNN:128, CNN:64, and CNN:32. The numbers in square brackets on the $\epsilon_{\mathcal{H}:\mathcal{N}}$ row are the corresponding errors for Δ-variance, evaluated as in Bates et al. (2020).

$\mathcal{N}$	128	64	32
$\epsilon_{\mathcal{H}:\mathcal{N}}$	0.040 [0.097]	0.056 [0.116]	0.091 [0.245]
$\epsilon_{\mathcal{S}:\mathcal{N}}$	0.182	0.221	0.271

$\mathcal{H}$. The panels on the bottom row show estimated $\mathcal{S}$ plotted against true $\mathcal{S}$. In each plot the black line represents exact agreement. The root-mean-square errors are given in the corner of each plot, and are summarized in Table 2. Given the small dynamic range of angular scales on the the xfBm images, and their lack of periodicity, these are acceptable statistics. We did explore the possibility of applying a spline correction to the results, to improve the fit further, and in some cases the fits did look better. However, in terms of the root-mean-square error, the improvement was very small. This is because the systematic errors (which are due to the non-uniqueness of the CNN) are smaller than the random ones (which are due to the lack of periodicity and the limited dynamic range). In the sequel we use CNN:64 and CNN:32, for reasons explained in Section 5.

The uncertainties obtained here for CNN:128 are significantly smaller than those obtained for the 128 × 128 network developed in Bates et al. (2020). This is because here, at each epoch, we generate new xfBm images, both for training and validation, whereas in Bates et al. (2020) we used the same images at each epoch, and simply divided them randomly between training and validation.

In Table 2 we also give root-mean-square errors for $\mathcal{H}$ when evaluated on the same images using Δ-variance. We see that these errors are significantly larger than the corresponding errors obtained with the CNNs, and deteriorate faster with decreasing $\mathcal{N}$. This is because the dynamic range of length scales over which $\mathrm{d}\ln\left(\sigma_\Delta^2\right)/\mathrm{d}\ln(L)$ can be estimated faithfully is very sensitive to $\mathcal{N}$, and in effect vanishes for $\mathcal{N} \lesssim 16$.

4 HIGH-RESOLUTION HI-GAL IMAGES OF SURFACE DENSITY ALONG THE GALACTIC PLANE

To analyse the structure of the interstellar medium in the disc of the Milky Way, we use images of surface density based on observations from the Herschel infrared Galactic Plane (Hi-GAL) survey (Molinari et al. 2010). The Hi-GAL survey divides the Galactic Plane into 163 [2.4° × 2.4°] tiles, each of which has been observed by *The Herschel Space Observatory* (Pilbratt et al. 2010) in the PACS bands centred at 70 and 160 μm (Poglitsch et al. 2010) and the SPIRE bands centred at 250, 350, and 500 μm (Griffin et al. 2010). Each individual tile is labelled with the integer Galactic longitude closest to its centre. Thus, for example, the tile towards the Galactic Centre is labelled ℓ 000, and the tile towards the Anti-Centre is labelled ℓ 180.

Images of surface density have been generated using the PPMAP analysis procedure (Marsh, Whitworth & Lomax 2015). In the first instance, the version of PPMAP developed by Marsh et al. (2017) delivers images of the contributions, $\Delta\tau_{300:q}$, to the total dust optical-depth at 300 μm, from dust at twelve different representative temperatures, T_q. These images are then added to obtain images of the total optical-depth of dust at 300 μm, τ_{300}.

The PPMAP analysis procedure is outlined more fully in Appendix C, and here we simply stress two points. (i) $\Delta\tau_{300:q}$ represents the contribution from a range of dust temperatures around T_q; for example, $T_5 = 15.6$ K represents dust with temperature in the range $14.3\,\mathrm{K} < T \leq 16.9\,\mathrm{K}$, and $T_6 = 18.4$ K represents dust with temperature in the range $16.9\,\mathrm{K} < T \leq 20.0\,\mathrm{K}$ (see Table C1). (ii)

The raw PPMAP images have $\phi_{\rm min} = 8.5$ arcsec pixels, corresponding to a linear scale of $0.041\,{\rm pc}\,[{\rm D\,kpc^{-1}}]$ for a structure at distance D along the line of sight.

We eschew the convention of converting the total dust optical depth into a column density of molecular hydrogen, since this appears to make little sense on at least two counts. First, on many lines of sight, even close to the Galactic Midplane, the fraction of hydrogen that is actually molecular is low and poorly constrained by observation. Under most circumstances, molecular hydrogen (as inferred by its presumed association with carbon monoxide) is only the dominant form of hydrogen in the interstellar medium if the volume and surface densities are sufficiently high. The exact threshold above which molecular hydrogen becomes the dominant form of hydrogen is quite complicated to formulate (e.g. Mac Low & Glover 2012), but reasonable working values are $\rho \gtrsim 2 \times 10^{-21}\,{\rm g\,cm^{-3}}$, $\boldsymbol{\Sigma \gtrsim 50\,g\,cm^{-2}}$ (equivalently $A_{\rm V} \gtrsim 2$). Consequently on most of the lines of sight in the Hi-GAL survey, a significant fraction of the hydrogen is atomic, as can be seen by comparing maps from the Leiden/Argentina/Bonn 21 cm Survey (Kalberla et al. 2005, and references therein) with the ^{12}CO(1-0) survey of the whole Galactic Plane by Dame, Hartmann & Thaddeus (2001) or the Galactic Ring Survey of Jackson et al. (2006).

Secondly, it is frequently implied, or even presumed, that the temperatures derived from far-infrared Spectral Energy Distributions are gas-kinetic temperatures. In reality they are dust-vibrational temperatures, and – unless the density is high, $n_{\rm H_2} \gtrsim 10^5\,{\rm cm^{-3}}$, the gas-kinetic temperature can be very different from (often much higher than) the co-spatial dust-vibrational temperature. In other words, it is important to be mindful that what is being observed is dust and not gas.

However, in order to make the PPMAP IMAGES more easily visualized, they are presented in terms of surface density (in ${\rm M_\odot\,pc^{-2}}$). The surface density is derived on the assumption that the net mass opacity of dust at 300 μm (i.e. cross-section per unit mass of everything) is

$$\kappa_{300} = [0.11\,{\rm cm^2\,g^{-1}}] \;=\; [2.1 \times 10^{-5}\,{\rm pc^2\,M_\odot^{-1}}]\,. \tag{5}$$

It follows that the presented total surface densities are given by

$$\Sigma = \frac{\tau_{300}}{\kappa_{300}} \;=\; [4.8 \times 10^4\,{\rm M_\odot\,pc^{-2}}]\,\tau_{300}\,. \tag{6}$$

Similarly the temperature differential surface densities are given by

$$\Delta\Sigma_q = [4.8 \times 10^4\,{\rm M_\odot\,pc^{-2}}]\,\Delta\tau_{300:q}\,. \tag{7}$$

Since this involves a single, well-defined conversion factor, the images can easily be recalibrated if a different value of κ_{300} is deemed appropriate.

None the less, even this is problematic, since κ_{300} is likely to vary with location in the Galaxy (and therefore along the line of sight), in particular (i) due to spatial variations in interstellar metallicity, and (ii) due to variations in the mix of dust types in different environments. Even if one neglects the likelihood that the mix of dust types changes with metallicity, it is still to be expected that the amount of dust increases, more-or-less monotonically, with metallicity. There is at least a fourfold variation in metallicity between the central regions of the Galaxy and the outer regions observed by *Herschel* (e.g. Wenger et al. 2019), and therefore presumably a comparable – or even greater (e.g. Galliano et al. 2021) – variation in the abundance of dust. It is also to be expected that the properties of dust will vary between different phases of the interstellar medium, due to coagulation and ice-mantle growth (e.g. Ormel et al. 2009, 2011). In short, it is important to keep in mind that the fundamental quantity derivable from long-wavelength dust-continuum emission – as detected by *Herschel* or any other far-infrared telescope – is the dust optical-depth (here τ_{300}). Any other quantity derived therefrom is dependent on uncertain conversion factors, like the one we adopt in equations (6) and 7.

Fig. 3 shows maps of temperature differential surface density for the Galactic Centre tiles ($\ell\,358$, $\ell\,000$, $\ell\,002$), and for the dust-temperature intervals for which there is significant emission ($T \leq 12.1$ K, 12.1 K $< T \leq 14.3$ K, etc.).

In the $\ell\,002$ tile there is little dust at $T \leq 14.3$ K or $T > 23.6$ K. Consequently the signal to noise in the corresponding maps (panels (a) and (f) in Fig. 3) is much lower than in the other temperature differential maps, and so the corresponding estimates of surface density are also extremely uncertain. This is the reason why there is a discontinuity on these maps, between the $\ell\,002$ tile and the adjacent $\ell\,000$ tile. We note that in the 14.3 K $< T \leq 16.9$ K, 16.9 K $< T \leq 20.0$ K, and 20.0 K $< T \leq 23.6$ K maps, where there is higher signal-to-noise, there is no discernible discontinuity. Hence there is also no discernible discontinuity on the map of total surface density, Fig. 4.

The dust in the $\ell\,002$ tile is also significantly cooler than that in the other two tiles, by of order $\Delta T \sim 3$ K. If we adopt the canonical relationship $T \propto U^{1/[4+\beta_{\rm D}]}$, where U is the ambient radiant energy density, and $\beta_{\rm D}$ is the far-infrared/sub-millimetre emissivity index, with $\beta_{\rm D} \sim 2$, this corresponds to the ambient radiation field being roughly two times stronger in the $\ell\,000$ and $\ell\,358$ tiles than in the $\ell\,002$ tile

Fig. 4 shows the total surface density for the three Hi-GAL tiles closest to the Galactic Centre ($\ell\,002$, $\ell\,000$, $\ell\,358$). These maps are obtained by summing the contributions from all the temperature intervals,

$$\Sigma = \sum_{q=1}^{q=Q} \{\Delta\Sigma_q\}\,, \tag{8}$$

where Q is the total number of discrete temperatures, T_q (i.e. by adding all the maps in Fig. 3).

In the sequel we first use the CNNs developed in Section 3 to analyse the maps of total surface density (e.g. Fig. 4), and in Section 7 we show that there appear to be two structural modes: one (the *undulating* mode) representing lines of sight with low surface density, and one (the *monolithic* mode) representing lines of sight with high surface density. We then use the temperature differential information (e.g. Fig. 3) to show that there is no clear evidence for a distinction between these two structural modes on the basis of their distributions of dust-vibrational temperature.

5 MAPS OF $\mathcal{H}$ AND $\mathcal{S}$

5.1 Notation for estimates of $\mathcal{H}$ and $\mathcal{S}$

This section is necessarily rather laboured, but it makes an important point that is critical and should be kept in mind. Any procedure for quantifying the structure of an image of the interstellar medium must be explicit about the range of angular scales that is meaningfully represented. There cannot be any a priori expectation that the resulting metrics (fractal dimension, multifractal spectrum, Hurst parameter, scaling parameter, power spectrum, etc.) will be the same if different ranges are considered, even if the procedure is notionally perfect.

This is illustrated by the maps of $\mathcal{H}$ and $\mathcal{S}$ generated and analysed here. These maps depend on two user-controlled parameters: (i) the angular size of the pixels, $\phi_{\rm min}$; and (ii) the size of the CNN used,

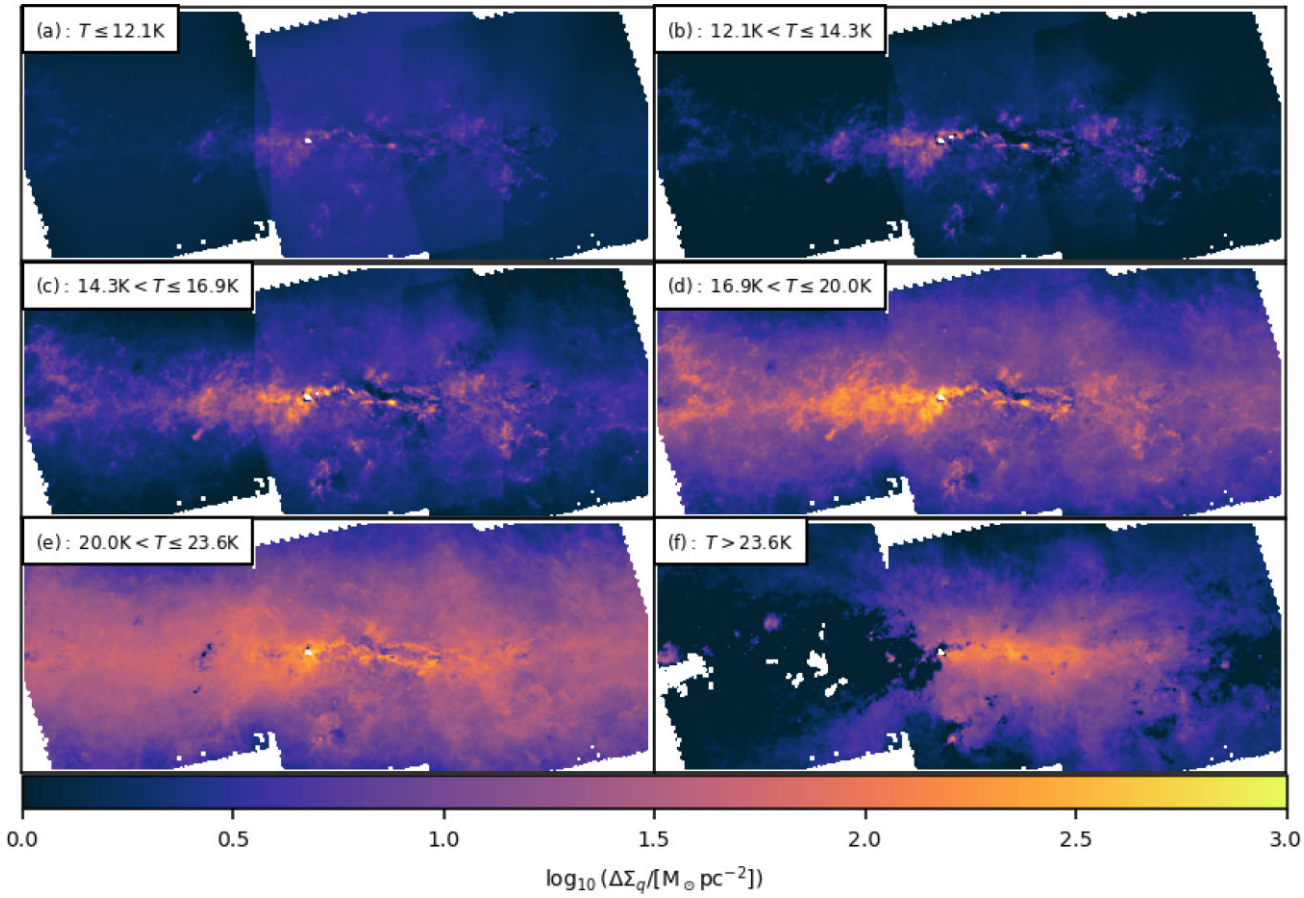

Figure 3. Images of the temperature-differential surface density, $\Delta\Sigma_{300:q}$, for the three Hi-GAL tiles towards the Galactic Centre (ℓ 002, ℓ 000, ℓ 358), and different ranges of dust temperature: (a) $T \leq 12.1$ K (represented by $T_1 = 8.00$ K, $T_2 = 9.45$ K, and $T_3 = 11.2$ K, see Table C1); (b) 12.1 K $< T \leq 14.3$ K (represented by $T_4 = 13.2$ K); (c) 14.3 K $< T \leq 16.9$ K (represented by $T_5 = 15.6$ K); (d) 16.9 K $< T \leq 20.0$ K (represented by $T_6 = 18.4$ K); (e) 20.0 K $< T \leq 23.6$ K (represented by $T_7 = 21.7$ K); and (f) $T > 23.6$ K (represented by $T_8 = 25.7$ K, $T_9 = 30.3$ K, $T_{10} = 35.8$ K, $T_{11} = 42.3$ K, and $T_{12} = 50.0$ K). Surface densities are derived from PPMAP estimates of the temperature-differential dust optical depth at 300 μm, using a standard conversion factor (see equation 7). The tiles have been combined using the REPROJECT package (Robitaille, Deil & Ginsburg 2020a). The co-ordinates, (b, ℓ), are omitted in order to increase the size of each panel, but can be inferred from Fig. 4, which shows exactly the same region.

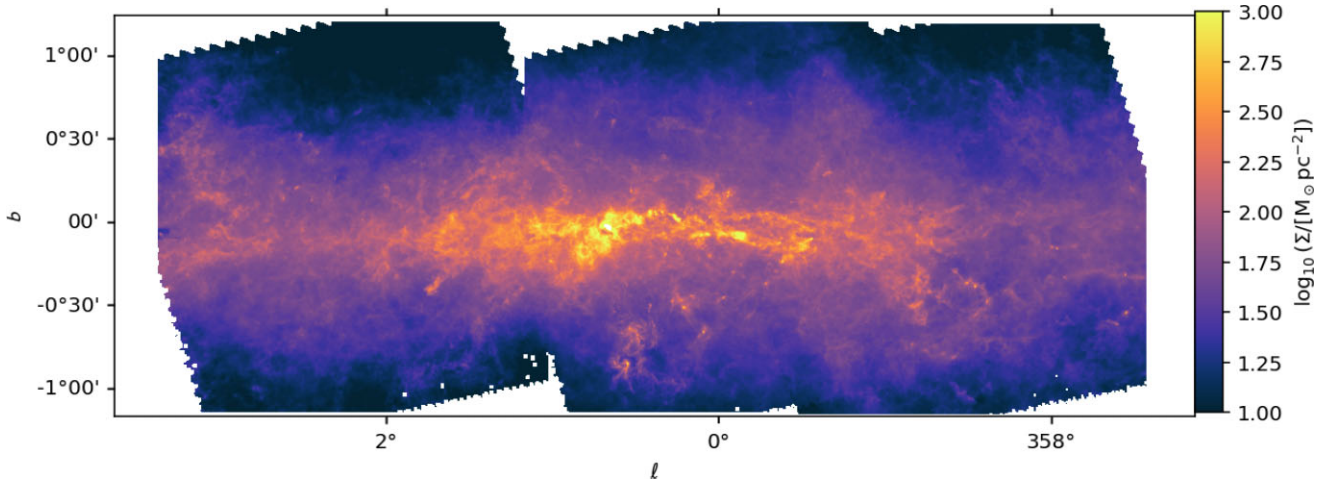

Figure 4. Images of the total surface density in the ℓ 002, ℓ 000, and ℓ 358 Hi-GAL tiles. These tiles cover an approximately 7° × 2° region in the direction of the Galactic Centre. Surface densities are derived from PPMAP estimates of the total dust optical depth at 300 μm, using a standard conversion factor (see equation 6). The three tiles have been combined using the REPROJECT package (Robitaille et al. 2020a).

i.e. the number of pixels, $\mathcal{N}$, along one side of the patch on which the CNN operates.

The angular size, $\phi_{\rm min}$, of the square pixels on the PPMAP image determines the angular size of the smallest structures that can be resolved, which is $\sim 2\phi_{\rm min}$. It is only 'user-controlled' in the sense that the user can chose to rebin the raw image, thereby degrading the resolution but increasing the signal-to-noise. Here we use both raw PPMAP images (with 8.5 arcsec pixels, see Section 4 and Appendix C) and also rebinned PPMAP images (with 17 arcsec pixels).

The size, $\mathcal{N}$, of the square CNNs used determines the angular size of the patches to be analysed, $\phi_{\rm max} = \mathcal{N}\phi_{\rm min}$, and hence the size of the largest structures that can be resolved, i.e. $\lesssim \phi_{\rm max}$.

In order to distinguish values of $\mathcal{H}$ or $\mathcal{S}$ produced with different combinations of $\phi_{\rm min}$ and $\mathcal{N}$, we denote them $\mathcal{H}_{\mathcal{N}\times\phi_{\rm min}}$ and $\mathcal{S}_{\mathcal{N}\times\phi_{\rm min}}$, where in the subscripts to $\mathcal{H}$ and $\mathcal{S}$ it is implicit that $\phi_{\rm min}$ is measured in arcseconds. For example, $\mathcal{H}_{32\times8.5}$ is an estimate of $\mathcal{H}$ based on an image with 8.5 arcsec pixels, using CNN:32. It therefore registers structures with angular scale between 8.5 arcsec and 32 × 8.5 arcsec = 272 arcsec, and is different from (say) $\mathcal{H}_{32\times17}$ which registers structures with angular scale between 17 arcsec and 32 × 17 arcsec = 544 arcsec. Only under the idealized circumstance that (i) the field being analysed is an exact xfBm field with dynamic range extending from $\ll$8.5 arcsec to $\gg$544 arcsec, and (ii) CNN:32 has zero error ($\epsilon_{\mathcal{H}:32} = 0$), are the two estimates necessarily the same, $\mathcal{H}_{32\times8.5} = \mathcal{H}_{32\times17}$.

In reality, observed fields are not exact xfBm fields. Consequently, even though observed fields are noisy, and as a result values of $\mathcal{H}$ and $\mathcal{S}$ estimated by the CNNs have errors (see Table 2), some of the difference between different estimates (for example, $\mathcal{H}_{\mathcal{N}\times\phi_{\rm min}}$ and $\mathcal{H}_{\mathcal{N}'\times\phi'_{\rm min}}$, with $\mathcal{N} \neq \mathcal{N}'$ and/or $\phi_{\rm min} \neq \phi'_{\rm min}$) may be attributable to real physical factors. For example, if the power spectrum flattens with decreasing wavenumber, $\mathcal{H}_{32\times17}$ should (all other things being equal) be smaller than $\mathcal{H}_{32\times8.5}$.

5.2 Constructing maps of $\mathcal{H}$ and $\mathcal{S}$

To produce maps of $\mathcal{H}_{\mathcal{N}\times\phi_{\rm min}}$ and $\mathcal{S}_{\mathcal{N}\times\phi_{\rm min}}$, we first specify the pixel size, $\phi_{\rm min}$, and, if the raw surface density image has smaller pixels than required, we rebin the image accordingly.

Next we create a square $\mathcal{N} \times \mathcal{N}$ grid with $\mathcal{N}^2$ square pixels each having dimension $\phi_{\rm min} \times \phi_{\rm min}$. This grid is placed on the surface density image, thereby isolating a small $\phi_{\rm max} \times \phi_{\rm max}$-patch, where

$$\phi_{\rm max} = \mathcal{N}\phi_{\rm min}, \tag{9}$$

and CNN:$\mathcal{N}$ is then used to estimate a single value of $\mathcal{H}$ and a single value of $\mathcal{S}$ for this patch.

The grid is then re-positioned and the process repeated. The grid is always positioned so that it only encloses whole pixels, and all possible positions of the grid are considered. For the relatively small values of $\mathcal{N}$ considered here ($\mathcal{N} = 32$ and 64), there are $\sim 2 \times 10^6$ different grid positions per tile when the raw PPMAP images are used (with $\phi_{\rm min} = 8.5$ arcsec), and $\sim 5 \times 10^5$ different grid positions when rebinned images with $\phi_{\rm min} = 17$ arcsec are used.

Once all possible positions of the grid have been treated, each pixel is allocated values of $\mathcal{H}_{\mathcal{N}\times\phi_{\rm min}}$ and $\mathcal{S}_{\mathcal{N}\times\phi_{\rm min}}$ which are the means of the values for all the grid positions that contain that pixel. Since the grids have an angular extent of $\sim\phi_{\rm max}$ (see equation 9), and therefore pull in information from pixels that far away, the resolution of the resulting maps of $\mathcal{H}_{\mathcal{N}\times\phi_{\rm min}}$ and $\mathcal{S}_{\mathcal{N}\times\phi_{\rm min}}$ is $\sim\phi_{\rm max}$.

There is evidently a compromise to be made here. For a given map, and hence a given $\phi_{\rm min}$, increasing $\mathcal{N}$ increases the range of scales that can be captured. If we are dealing with a perfect monofractal, an increased range of scales is useful, since it allows a more accurate evaluation of the Hurst parameter. This is because the notion of monofractality involves self-similarity (strictly self-affinity) across a large dynamic range, and so it is easier to measure when there is a large dynamic range. This is particularly true when the field analysed is non-periodic, because under this circumstance the larger scales do not accurately reflect the underlying self-similarity.

However, as shown by Elia et al. (2014, 2018), the Hi-GAL tiles are clearly not monofractal, and the statistics of substructure are likely to vary, both over an individual Hi-GAL tile (between lines of sight that intersect dense molecular gas, and lines of sight that intercept very little) and between different Hi-GAL tiles. Increasing $\mathcal{N}$ increases $\phi_{\rm max}$ (see equation 9), and thereby degrades the resolution, increasing the likelihood that lines of sight with very different statistics are analysed together. There is therefore an advantage to using a relatively small $\mathcal{N}$, provided that the resulting estimates of $\mathcal{H}$ and $\mathcal{S}$ are sufficiently accurate.

In Section 3 we have shown that CNN:64, and even CNN:32, return acceptably accurate values of $\mathcal{H}$ and $\mathcal{S}$ (see Fig. 2 and Table 2). The compromise we adopt is therefore to use CNN:64 and CNN:32 in the sequel, in order to obtain better resolution than would be obtained with CNN:128. Specifically we consider the three combinations of $\mathcal{N}$ and $\phi_{\rm min}$ whose properties are summarized in Table 3.

Table 3. The parameters regulating different estimates of $\mathcal{H}$ and $\mathcal{S}$. The first and second columns give the symbols used for these estimates. The third column gives the dynamic range of angular scales that is registered, $\mathcal{N}$, i.e. the number of pixels along one side of the grid. The fourth column gives the pixel size of the surface density image used, $\phi_{\rm min}$. The fifth column gives the resolution of the resulting maps, $\phi_{\rm max}$.

$\mathcal{H}_{\mathcal{N}\times\phi_{\rm min}}$	$\mathcal{S}_{\mathcal{N}\times\phi_{\rm min}}$	$\mathcal{N}$	$\phi_{\rm min}$	$\phi_{\rm max}$
$\mathcal{H}_{64\times8.5}$	$\mathcal{S}_{64\times8.5}$	64	8.5 arcsec	544 arcsec
$\mathcal{H}_{32\times17}$	$\mathcal{S}_{32\times17}$	32	17 arcsec	544 arcmin
$\mathcal{H}_{32\times8.5}$	$\mathcal{S}_{32\times8.5}$	32	8.5 arcsec	272 arcsec

6 THE TILES TOWARDS THE GALACTIC CENTRE

We illustrate these issues by analysing the three Hi-GAL tiles close to the direction of the Galactic Centre (ℓ 358, ℓ 000, and ℓ 002). One should be mindful that in these tiles, as in all the other Hi-GAL tiles, the observed intensities, and hence also the estimated surface densities, reflect integrals along the line of sight. Consequently on many lines of sight there may be confusion, due to contributions from different structures superposed along the line of sight. At the same time we note that, if the structures intercepted are truly monofractal, as our analysis assumes, there is no problem with confusion, since the statistical properties of a monofractal structure are independent of resolution, and therefore independent of the distance to that structure.

The two panels on the top row of Fig. 5 show $\mathcal{H}_{64\times8.5}$ and $\mathcal{S}_{64\times8.5}$, i.e. estimates of $\mathcal{H}$ and $\mathcal{S}$ obtained with $\phi_{\rm min} = 8.5$ arcsec (the raw PPMAP image) and CNN:64, for the three Galactic Centre tiles. The maps capture the statistics of substructure between $\sim$8.5 arcsec and $\sim\phi_{\rm max} = 64 \times 8.5$ arcsec = 544 arcsec. The angular resolution of the maps is $\sim\phi_{\rm max}$, and for a region at distance D along the line of sight, this angular resolution corresponds to a linear resolution of $L \sim 13.2\,{\rm pc}\,[D/5\,{\rm kpc}]$. In these tiles a significant fraction of the emission is likely to come from structures at the Galactic Centre at distance $D \sim 8$ kpc, and hence for these structures the linear resolution is $L \sim 21$ pc.

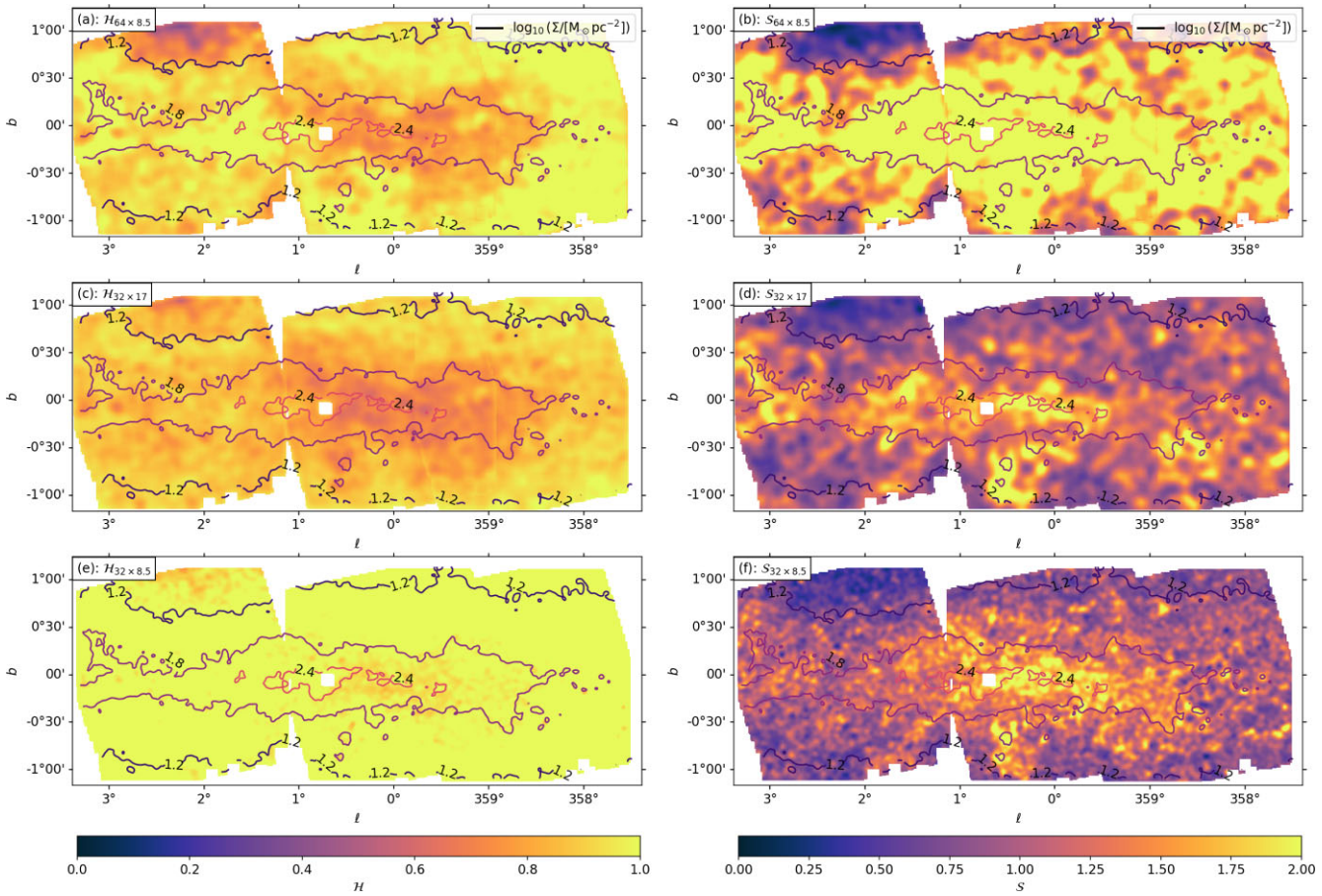

Figure 5. Maps of $\mathcal{H}$ and $\mathcal{S}$ in the three tiles close to the direction of the Galactic Centre (tiles ℓ 358, ℓ 000, ℓ 002). The left-hand panels show $\mathcal{H}$, and the right-hand panels show $\mathcal{S}$. The panels on the top row show $\mathcal{H}_{64\times8.5}$ and $\mathcal{S}_{64\times8.5}$, i.e. values generated by applying CNN:64 directly to the high-resolution PPMAP images; these maps have an angular resolution of 544 arcsec and sample a dynamic range of angular scales, from ∼8.5 to ∼544 arcsec (i.e. × 64). The panels on the middle row show $\mathcal{H}_{32\times17}$ and $\mathcal{S}_{32\times17}$, i.e. values generated by first rebinning the PPMAP image by a linear factor of two, and then applying CNN:32; these maps have an angular resolution of 544 arcsec (the same as those on the row above), but a smaller dynamic range of angular scales, from ∼17 to ∼544 arcsec (i.e. × 32). The panels on the bottom row show $\mathcal{H}_{32\times8.5}$ and $\mathcal{S}_{32\times8.5}$, i.e. values generated by applying CNN:32 directly to the high-resolution PPMAP image; these maps have an angular resolution of 272 arcsec (i.e. finer than that of the maps on the two rows above), and they register angular scales from ∼8.5 to 272 arcsec (again a range of × 32, but registering smaller angular scales than those on the row above, and disregarding larger angular scales). See the text for further details. The background contours represent surface density, as shown in Fig. 4, and are labelled with the value of $\log_{10}(\Sigma/[\mathrm{M}_\odot\ \mathrm{pc}^{-2}])$, i.e. 1.2, 1.8, and 2.4. The tiles have been combined using the REPROJECT package (Robitaille et al. 2020a).

We note that on the panels of Fig. 5, there are sometimes linear structures marking the join between adjacent tiles, for example between tiles ℓ 000 and ℓ 358 in Fig. 5(a), at positive latitudes. These features arise because the values of $\mathcal{H}$ and $\mathcal{S}$ for a given pixel are obtained by averaging over a patch of pixels on the same tile, as explained in Section 5.2. For two adjacent pixels on opposite sides of the join between two tiles, this averaging pulls in values of $\mathcal{H}$ and $\mathcal{S}$ from completely different patches (because they are on completely distinct tiles). Therefore such features are inevitable at joins where there is a gradient in the overall background on scales of order $\phi_{\max}$. In contrast, for adjacent pixels on the same tile, the associated patches are offset by at most $\sqrt{2}\phi_{\min}$.

The two panels on the middle row of Fig. 5 show $\mathcal{H}_{32\times17}$ and $\mathcal{S}_{32\times17}$, i.e. estimates of $\mathcal{H}$ and $\mathcal{S}$ obtained with $\phi_{\min} = 17$ arcsec (a rebinned version of the PPMAP image) and CNN:32. The angular resolution of these maps is $\sim\phi_{\max} = 32 \times 17$ arcsec $= 544$ arcsec, the same as the maps on the top row. However, because the pixels are larger ($\phi_{\min} = 17$ arcsec rather than $\phi_{\min} = 8.5$ arcsec), $\mathcal{H}_{32\times17}$ and $\mathcal{S}_{32\times17}$ only capture the statistics of substructures between ∼17 and ∼544 arcsec.

The two panels on the bottom row of Fig. 5 show $\mathcal{H}_{32\times8.5}$ and $\mathcal{S}_{32\times8.5}$, i.e. estimates of $\mathcal{H}$ and $\mathcal{S}$ obtained with $\phi_{\min} = 8.5$ arcsec (the raw PPMAP image) and CNN:32. The angular resolution of these maps is $\sim\phi_{\max} = 32 \times 8.5$ arcsec $= 272$ arcsec $\simeq 4.5$ arcmin. These maps therefore have better resolution than the maps on the two rows above, but they only capture the statistics of substructures between ∼8.5 and ∼272 arcsec. The corresponding linear resolution at distance D is $L \sim 6.6\,\mathrm{pc}\,[D/5\,\mathrm{kpc}]$, thus ~ 10.6 pc for structures at the Galactic Centre.

The solid contours on the panels of Fig. 5 represent lines of constant surface density. They are labelled with the value of $\log_{10}(\Sigma/[\mathrm{M}_\odot\ \mathrm{pc}^{-2}])$, i.e. 1.2, 1.8, and 2.4. On lines of sight outside the contour labelled '1.2' (equivalently, lines of sight with surface density $\Sigma \lesssim 16\,\mathrm{M}_\odot\ \mathrm{pc}^{-2}$, sub-millimetre optical depth, $\tau_{300} \lesssim 0.0003$, and visual extinction of $A_\mathrm{V} \lesssim 0.7$), PPMAP surface densities tend to become less reliable, i.e. the uncertainties returned by PPMAP are

sometimes comparable with the expectation values. However, lines of sight compromised in this way are only a very small fraction of the total.

PPMAP returns NaN ('Not-a-Number') for pixels that do not contain enough signal to support a meaningful evaluation of the surface density. These pixels are mainly towards the edges of the tiles, but there are a few in the interiors of the tiles. If the $\mathcal{N} \times \mathcal{N}-$grid overlaps a pixel with NaN, the CNN cannot operate. Consequently these pixels do not have estimates of $\mathcal{H}$ or $\mathcal{S}$, and surrounding pixels have estimates of $\mathcal{H}$ and $\mathcal{S}$ that are more uncertain. An example of a region that is void for this reason can be seen near $\ell = 0.75^\circ$ on all the panels of Fig. 5.

6.1 Comparing maps of $\mathcal{H}$ and $\mathcal{S}$ with the same angular resolution but different dynamic range of angular scales

If we treat $\mathcal{H}_{32\times17}$ and $\mathcal{S}_{32\times17}$ (middle row of Fig. 5) as our reference, then $\mathcal{H}_{64\times8.5}$ and $\mathcal{S}_{64\times8.5}$ (top row of Fig. 5) have the same resolution ($\phi_{\rm max} = 544$ arcsec) but capture a larger dynamic range of angular scales ($\phi_{\rm max}/\phi_{\rm min} = 64$, rather than 32, because they have smaller $\phi_{\rm min}$). Specifically, substructures between ∼8.5 and ∼17 arcsec come into the reckoning when estimating $\mathcal{H}_{64\times8.5}$ and $\mathcal{S}_{64\times8.5}$ but are ignored when estimating $\mathcal{H}_{32\times17}$ and $\mathcal{S}_{32\times17}$.

Comparing the maps of $\mathcal{H}_{64\times8.5}$ (Fig. 5a) and $\mathcal{H}_{32\times17}$ (Fig. 5c), there is both an overall morphological similarity and a systematic shift, in the direction $\mathcal{H}_{64\times8.5} > \mathcal{H}_{32\times17}$. This implies a paucity of small-scale substructure between ∼8.5 and ∼17 arcsec (which at the Galactic Centre corresponds to linear scales between ~ 0.3 pc and ~ 0.6 pc). Consequently the power spectrum steepens a little at large wavenumber, and this increases $\mathcal{H}$ slightly.

The maps of $\mathcal{S}_{64\times8.5}$ (Fig. 5b) and $\mathcal{S}_{32\times17}$ (Fig. 5d) are morphologically very similar, but $\mathcal{S}_{64\times8.5}$ is systematically larger than $\mathcal{S}_{32\times17}$, implying that the small-scale substructures between ∼8.5 and ∼17 arcsec have higher surface density than the more extended substructures on the map, as would be expected if the small-scale substructures tend to be nested within the larger ones.

6.2 Comparing maps of $\mathcal{H}$ and $\mathcal{S}$ with the same dynamic range of angular scales but different angular resolution

If we continue to treat $\mathcal{H}_{32\times17}$ and $\mathcal{S}_{32\times17}$ (middle row of Fig. 5) as our reference, then $\mathcal{H}_{32\times8.5}$ and $\mathcal{S}_{32\times8.5}$ (bottom row of Fig. 5) capture the same dynamic range of angular scales ($\phi_{\rm max}/\phi_{\rm min} = 32$) but with finer resolution ($\phi_{\rm max} = 272$ arcsec, rather than 544 arcsec). Substructures between 8.5 and 17 arcsec come into the reckoning when estimating $\mathcal{H}_{32\times8.5}$ and $\mathcal{S}_{32\times8.5}$ but are ignored when estimating $\mathcal{H}_{32\times17}$ and $\mathcal{S}_{32\times17}$. Conversely, substructures between 272 and 544 arcsec come into the reckoning when estimating $\mathcal{H}_{32\times17}$ and $\mathcal{S}_{32\times17}$ but are ignored when estimating $\mathcal{H}_{32\times8.5}$ and $\mathcal{S}_{32\times8.5}$.

Comparing the maps of $\mathcal{H}_{32\times8.5}$ (Fig. 5e) and $\mathcal{H}_{32\times17}$ (Fig. 5c), the two maps have comparable morphology, but $\mathcal{H}_{32\times8.5}$ (typically in the interval [0.8,1.0]) is systematically larger than $\mathcal{H}_{32\times17}$ (typically in the interval [0.6,0.8]). Again this is attributable to a steepening of the power spectrum at large wavenumber, and a flattening at small wavenumber.

If we compare the maps of $\mathcal{S}_{32\times8.5}$ (Fig. 5f) and $\mathcal{S}_{32\times17}$ (Fig. 5d), they are quite similar, apart from the fact that $\mathcal{S}_{32\times8.5}$ has visibly better resolution (∼272 arcsec $\simeq 0.075^\circ$, so 32 resolution elements across a tile) than $\mathcal{S}_{32\times17}$ (∼544 arcsec $\simeq 0.151^\circ$, so 16 resolution elements across a tile). In other words, if Fig. 5(f) were re-gridded to the resolution of Fig. 5(d), the two maps would look very similar. Since these maps reflect the scaling of surface density (i.e. $\bar{\Sigma}(L)$, the mean surface density on scales of order L) – and by implication the scaling of volume density (i.e. $\bar{\rho}(L)$, the mean volume density on scales of order L; cf. Larson 1981) – this suggests that these scaling laws are self-similar across the full range of scales from 8.5 to 544 arcsec.

7 ANALYSIS OF THE WHOLE GALACTIC PLANE

7.1 The distributions of $\mathcal{H}$ and $\mathcal{S}$ for all Hi-GAL tiles

The top row of Fig. 6 shows the distributions of (a) $\mathcal{H}_{32\times8.5}$, (b) $\mathcal{H}_{32\times17}$, and (c) $\mathcal{H}_{64\times8.5}$ for all the non-NAN pixels on all the Hi-GAL tiles. The middle row of Fig. 6 shows the corresponding distributions of (d) $\mathcal{S}_{32\times8.5}$, (e) $\mathcal{S}_{32\times17}$, and (f) $\mathcal{S}_{64\times8.5}$. The bottom row of Fig. 6 shows plots of (h) $\mathcal{H}_{32\times8.5}$ against $\mathcal{S}_{32\times8.5}$, (i) $\mathcal{H}_{32\times17}$ against $\mathcal{S}_{32\times17}$, and (j) $\mathcal{H}_{64\times8.5}$ against $\mathcal{S}_{64\times8.5}$.

These distributions reproduce the trends found for the tiles towards the Galactic Centre, in Section 6. Specifically, (i) the power spectrum is curved, i.e. flatter at small wavenumbers, and steeper at large wavenumbers; (ii) the larger the range of scales represented (i.e. the larger $\mathcal{N} = \phi_{\rm max}/\phi_{\rm min}$), the larger the value of $\mathcal{S}$. We must also be mindful that the different pixels represented in Fig. 6 correspond to a variety of lines of sight, with a relatively small fraction intercepting dense and highly structured star-forming molecular clouds, and a larger fraction intercepting more quiescent regions at higher Galactic latitudes.

Consider first the distributions of $\mathcal{H}_{32\times8.5}$ (Fig. 6a) and $\mathcal{S}_{32\times8.5}$ (Fig. 6d). Scales between 8.5 and 272 arcsec are represented, whilst scales between 272 and 544 arcsec are not, so the power spectrum is biased towards larger wavenumbers, and is relatively steep, i.e. large $\mathcal{H} \sim 1$. The dynamic range is relatively small (272 arcsec/8.5 arcsec = 32), so $\mathcal{S}$ is also relatively small ($\mathcal{S} \sim 0.9$).

Next consider the distributions of $\mathcal{H}_{32\times17}$ (Fig. 6b) and $\mathcal{S}_{32\times17}$ (Fig. 6e). Scales between 17 and 544 arcsec are represented, whilst scales between 8.5 and 17 arcsec are not, so the power spectrum is biased towards smaller wavenumbers, and is somewhat flatter, i.e. smaller $\mathcal{H}$. The dynamic range is again relatively small (544 arcsec/17 arcsec = 32), so $\mathcal{S}$ is again relatively small ($\mathcal{S} \sim 0.9$). There are indications of bimodality in the distributions, and Fig. 6(h) suggests one concentration of pixels with $(\mathcal{H}_{32\times17}, \mathcal{S}_{32\times17}) = (0.9, 1.0)$, and a second with $(\mathcal{H}_{32\times17}, \mathcal{S}_{32\times17}) = (0.2, 0.5)$.

Finally consider the distributions of $\mathcal{H}_{64\times8.5}$ (Fig. 6c) and $\mathcal{S}_{64\times8.5}$ (Fig. 6f). Scales between 8.5 and 544 arcsec are represented, so the power spectrum reflects both smaller and larger wavenumbers and the dynamic range is larger (544 arcsec/8.5 arcsec = 64). There are again indications of bimodality, and in general $\mathcal{H}_{64\times8.5}$ increases with increasing $\mathcal{S}_{64\times8.5}$, as shown on Fig. 6(i).

The bimodality reflects two distinct types of line of sight. One type is characterized by a mode which we label the 'monolithic mode', with relatively large angular scales dominant (large $\mathcal{H} \gtrsim 0.8$) and quite high scaling parameter (large $\mathcal{S} \gtrsim 1$). The second type is characterized by a mode which we label the 'undulating mode', with a broader range of angular scales (intermediate $\mathcal{H} \lesssim 0.8$) and lower scaling parameter ($\mathcal{S} \lesssim 1$). With CNN:32 operating on the un-rebinned PPMAP images ($\phi_{\rm min} = 8.5$ arcsec $\phi_{\rm max} = 272$ arcsec; panels 6a, 6d, and 6g), the distinction between the two types of line of sight is weak; the small dynamic range ($\mathcal{N} = 32$) makes it hard to capture the broad range of angular scales of the undulating mode, and the small $\phi_{\rm max}$ makes it hard to capture the large angular scales of the monolithic mode. With CNN:32 operating on the rebinned PPMAP images ($\phi_{\rm min} = 17$ arcsec, $\phi_{\rm max} = 544$ arcsec; panels 6b, 6e and 6h), the

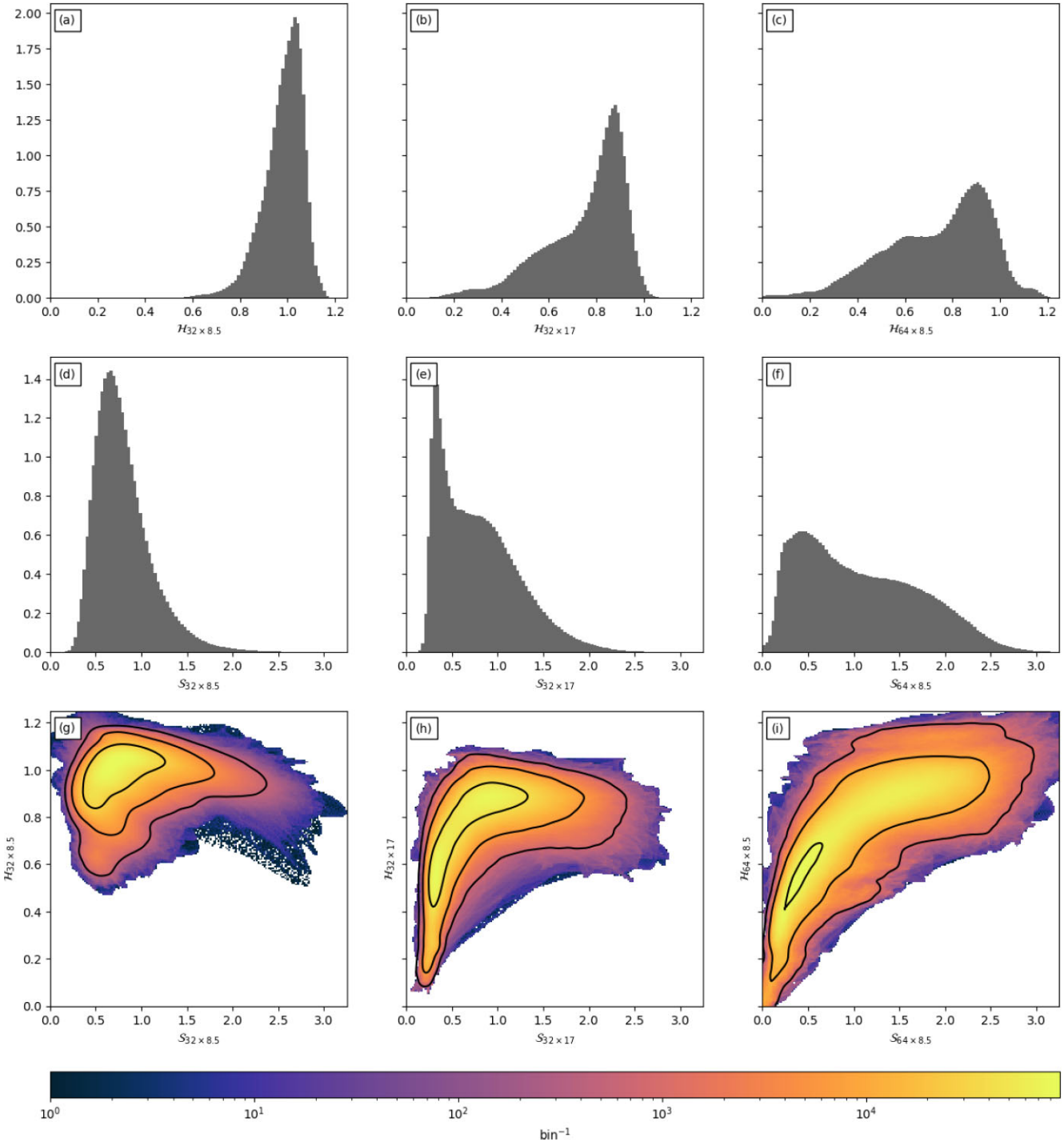

Figure 6. Top row, left to right: the distributions of (a) $\mathcal{H}_{64\times8.5}$, (b) $\mathcal{H}_{32\times17}$, and (c) $\mathcal{H}_{32\times8.5}$. Middle row, left to right: the distributions of (d) $\mathcal{S}_{64\times8.5}$, (e) $\mathcal{S}_{32\times17}$, and (f) $\mathcal{S}_{32\times8.5}$. Bottom row, left to right: the 2D distributions of (g) $\mathcal{H}_{64\times8.5}$ against $\mathcal{S}_{64\times8.5}$, (h) $\mathcal{H}_{32\times17}$ against $\mathcal{S}_{32\times17}$, and (i) $\mathcal{H}_{32\times8.5}$ against $\mathcal{S}_{32\times8.5}$. On each panel there is a value or point for each of the $\sim 2 \times 10^8$ pixels on the Hi-GAL maps. On panels (a) through (c), the bin size is $\Delta\mathcal{H} = 6.70 \times 10^{-3}$, and on panels (d) through (f) the bin size is $\Delta\mathcal{S} = 1.80 \times 10^{-2}$. On panels (a) through (f) the ordinate values have been divided by 10^7, and give the number of pixels in that bin. The contours on panels (g), (h), and (i) represent 10^7, 10^8, and 7×10^8 points per bin, respectively.

bimodality starts to emerge, because the monolithic modes are better captured by the larger $\phi_{\rm max}$. Finally, with CNN:64 operating on the un-rebinned PPMAP images ($\phi_{\rm min} = 8.5$ arcsec, $\phi_{\rm max} = 544$ arcsec; panels 6c, 6f, and 6i), the bimodality becomes even clearer, because the undulating modes are now better captured by the larger dynamic range ($\phi_{\rm max}/\phi_{\rm min} = 64$).

7.2 $\mathcal{H}$ and $\mathcal{S}$ as a function of surface density, Σ

The panels on the top row of Fig. 7 show the distributions of $\mathcal{H}_{32\times8.5}$, $\mathcal{H}_{32\times17}$, and $\mathcal{H}_{64\times8.5}$ against surface density, Σ. The panels on the bottom row show the corresponding distributions of $\mathcal{S}_{32\times8.5}$, $\mathcal{S}_{32\times17}$, and $\mathcal{S}_{64\times8.5}$ against Σ. In general $\mathcal{H}$ and $\mathcal{S}$ both increase with

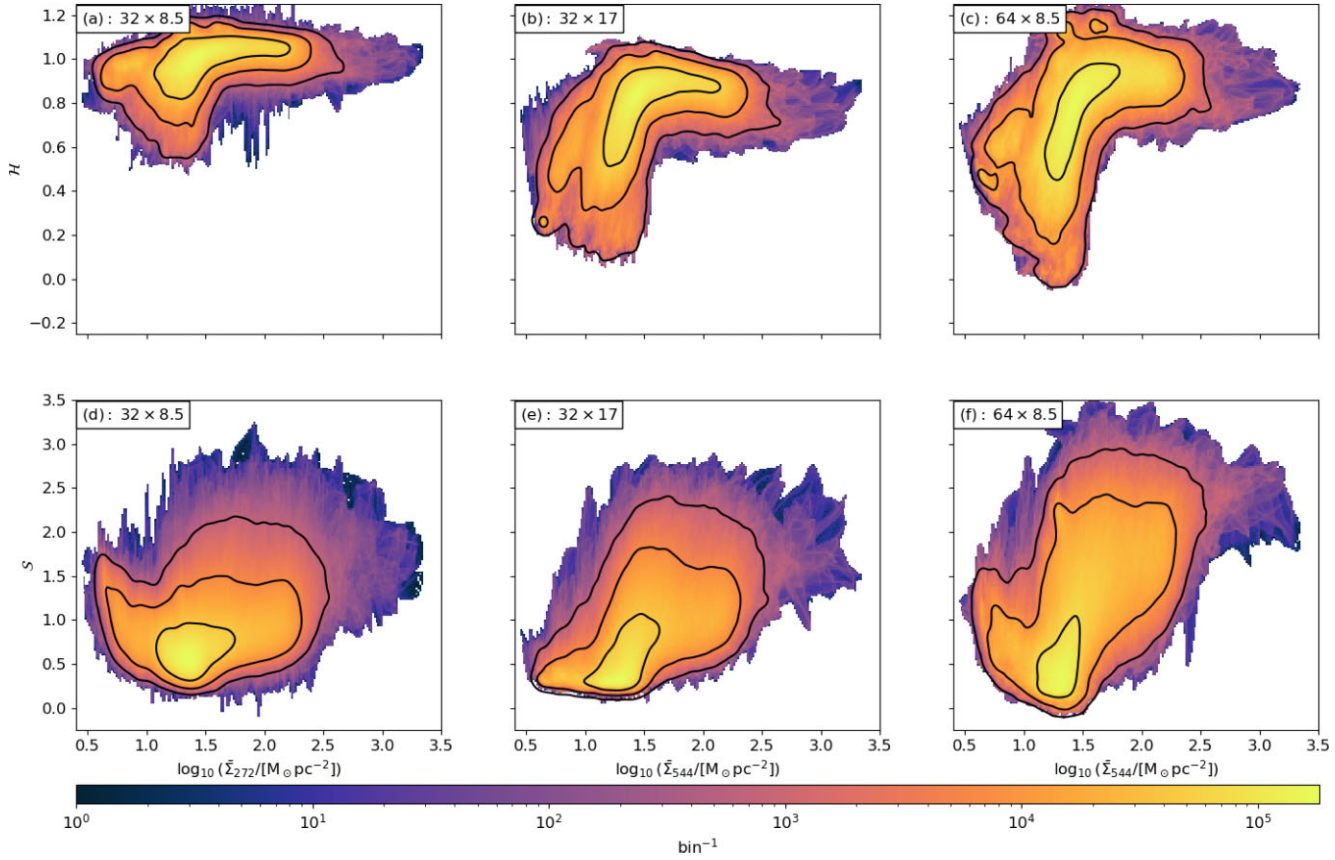

Figure 7. Top row, left to right: the distributions of (a) $\mathcal{H}_{64\times8.5}$, (b) $\mathcal{H}_{32\times17}$, and (c) $\mathcal{H}_{32\times8.5}$, against $\log_{10}(\Sigma/[\mathrm{M}_\odot\,\mathrm{pc}^{-2}])$, where Σ is the surface density, computed as described in Section 4. Bottom row, left to right: the distributions of (d) $\mathcal{S}_{64\times8.5}$, (e) $\mathcal{S}_{32\times17}$, and (f) $\mathcal{S}_{32\times8.5}$, against $\log_{10}(\Sigma/[\mathrm{M}_\odot\,\mathrm{pc}^{-2}])$. On each of these plots, there is a point for each of the $\sim 2 \times 10^8$ pixels on the Hi-GAL maps. The contour levels are the same as for Fig. 6 (i.e. 10^7, 10^8, and 7×10^8 bin^{-1}). The bin sizes are: $\Delta\mathcal{H} = 6.70 \times 10^{-3}$, $\Delta\mathcal{S} = 1.80 \times 10^{-2}$, and $\Delta \log_{10}(\Sigma) = 1.44 \times 10^{-2}$.

increasing surface density, especially for the maps that best capture the bimodality (i.e. 64 × 8.5 best, 32 × 8.5 worst). In other words, the substructure around pixels with large surface density tends to have a steeper power spectrum (power more concentrated towards large angular scales) and a larger range of surface density. Thus the monolithic mode tends to correlate with high surface density, $\Sigma \gtrsim 32\,\mathrm{M}_\odot\,\mathrm{pc}^{-2}$, and the undulating mode tends to correlate with low surface density, $\Sigma \lesssim 32\,\mathrm{M}_\odot\,\mathrm{pc}^{-2}$.

This in turn suggests that the monolithic mode describes molecular clouds where self-gravity is important (and therefore there may be imminent or on-going star formation) and a power-law distribution of surface density at high values (e.g. Kainulainen et al. 2009; Schneider et al. 2012; Girichidis et al. 2014; Schneider et al. 2015, 2022). In contrast, the undulating mode describes regions where turbulence maintains an approximately log-normal distribution of surface density, with a flatter power spectrum (smaller $\mathcal{H}$) and a narrower range (smaller $\mathcal{S}$).

7.3 $\mathcal{H}$ and $\mathcal{S}$ as a function of Galactic longitude, ℓ

The panels on the top row of Fig. 8 show the distributions of $\mathcal{H}_{32\times8.5}$, $\mathcal{H}_{32\times17}$, and $\mathcal{H}_{64\times8.5}$ against Galactic longitude, ℓ. The panels on the middle row show the distributions of $\mathcal{S}_{32\times8.5}$, $\mathcal{S}_{32\times17}$, and $\mathcal{S}_{64\times8.5}$ against ℓ. The panel on the bottom row shows the distribution of total surface density against ℓ. Each black dot gives the mean value of $\mathcal{H}$ (upper plots), $\mathcal{S}$ (middle plots), or $\log_{10}(\Sigma/[\mathrm{M}_\odot\,\mathrm{pc}^{-2}])$ (bottom plot) for the individual tile at that longitude. The vertical colour-coded line through a black dot gives the distribution of the corresponding quantity for the individual pixels in that individual tile. There are 163 tiles, hence 163 black dots on each plot. Each tile contains over 10^6 pixels, and therefore the colour-coded vertical distributions for individual tiles are well defined.

Since the larger values of $\mathcal{H}$ and $\mathcal{S}$ are associated with the monolithic mode, and hence dense molecular clouds, the distributions of $\mathcal{H}$, and to a lesser extent the distributions of $\mathcal{S}$, involve significantly lower values in directions towards the outer Galaxy ($90° \lesssim \ell \lesssim 270°$). As with all systematic trends, these signatures are weakest for the values obtained with $(\mathcal{N}, \phi_{\min}) = (32, 8.5)$, and strongest for those obtained with $(\mathcal{N}, \phi_{\min}) = (64, 8.5)$. As explained in Section 7.1, this is because $(\mathcal{N}, \phi_{\min}) = (64, 8.5)$ is most able to capture the monolithic mode, and $(\mathcal{N}, \phi_{\min}) = (32, 8.5)$ least able.

The red dashed lines in Fig. 8(g) mark the approximate longitudes of the tangent lines to spiral arms, according to Hou & Han (2014, 2015) (see the Caption to Fig. 8 for the identities of these lines). There are weak features at all these positions, but there are many other equally strong features. In general, these surface density peaks represent the accumulation of many contributing molecular clouds along the line of sight. An individual cloud at distance D, with cross-sectional area A on the sky, and therefore linear extent $L \simeq A^{1/2}$, only occupies a fraction

$$f \simeq 0.002 \left[\frac{L}{10\,\mathrm{pc}}\right]^2 \left[\frac{D}{5\,\mathrm{kpc}}\right]^{-2}$$

of the tile in which it is located. Therefore individual clouds are unlikely to be clearly identifiable on this plot.

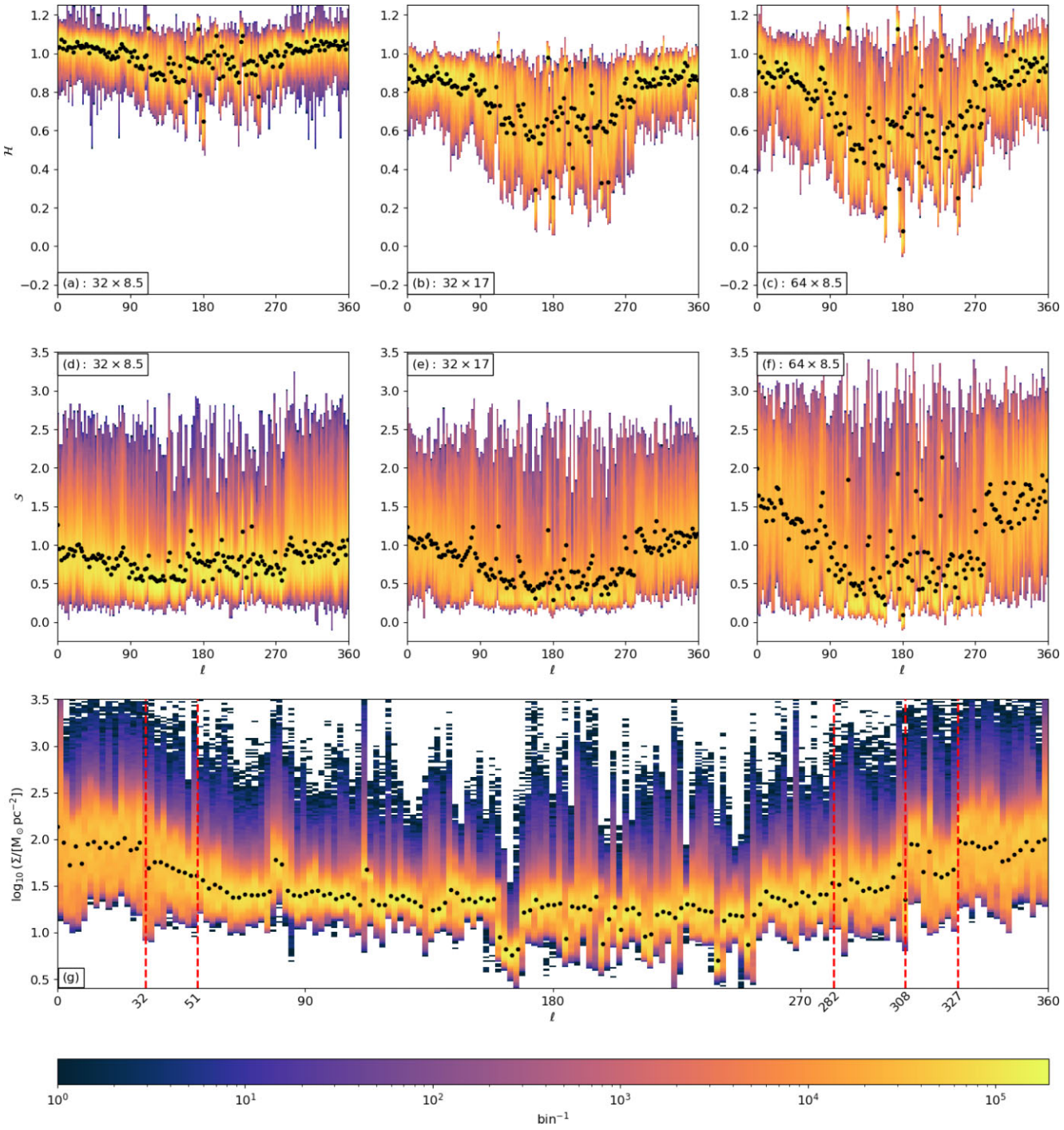

Figure 8. Top row, left to right: the distributions of (a) $\mathcal{H}_{64\times8.5}$, (b) $\mathcal{H}_{32\times17}$, and (c) $\mathcal{H}_{32\times8.5}$, in each tile, against the longitude, ℓ (in degrees). Middle row, left to right: the distributions of (d) $\mathcal{S}_{64\times8.5}$, (e) $\mathcal{S}_{32\times17}$, and (f) $\mathcal{S}_{32\times8.5}$, in each tile, against ℓ. Bottom row, panel (g), the distribution of total surface density in each tile, against ℓ. On each panel, the black dots represent the mean value in the individual tile at that longitude, and the colour-coded vertical lines through a black dot represents the distribution of surface densities for the individual pixels in that tile; there are $\sim10^6$ pixels in each tile. The red dashed lines represent the approximate longitudes of the tangent lines to spiral arms, according to Hou & Han (2014, 2015): $\ell \sim 32°$, Scutum Arm; $\ell \sim 51°$, Sagittarius Arm; $\ell \sim 282°$, Carina Arm; $\ell \sim 308°$, Centaurus Arm; $\ell \sim 327°$, Norma Arm.

7.4 Temperature differential surface densities

In the first instance, PPMAP delivers maps of temperature-differential surface density; these are then co-added to produce maps of total surface density. It is this fundamental aspect of the PPMAP procedure that underpins the high resolution of the surface density maps analysed in Sections 6 and 7, and allows us to compute a meaningful median dust-vibrational temperature (see below). The individual temperature-differential maps (Fig. 3) cannot usefully be analysed with the CNNs developed here, because it is unclear how the resulting Hurst parameters should be interpreted. Under most circumstances

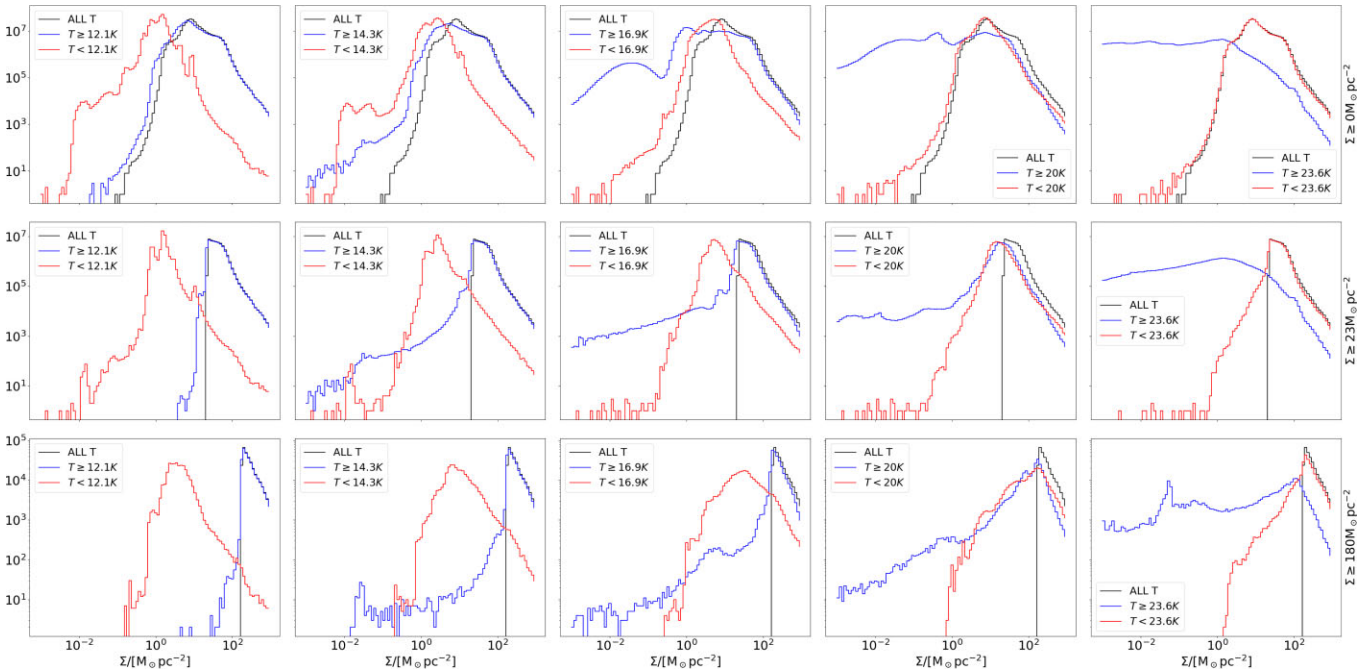

Figure 9. Distributions of surface density, $\Sigma/[\mathrm{M}_\odot\,\mathrm{pc}^{-2}]$, for dust at all temperatures (black curves), then just for the dust below a given separatrix-temperature, T_s (red curves), and just for the dust above T_s (blue curves). From left to right, the five columns have separatrix-temperatures $T_\mathrm{s} = 12.1\,\mathrm{K}$, 14.3 K, 16.9 K, 20.0 K, and 23.6 K. On the top row the plots show the results for all $\sim 2 \times 10^8$ pixels. On the middle row the plots show the results when we restrict consideration to the $\sim 5 \times 10^7$ pixels with $\Sigma > 23\,\mathrm{M}_\odot\,\mathrm{pc}^{-2}$; this threshold is equivalent to $A_\mathrm{V} \sim 1$. On the bottom row the plots show the results when we restrict consideration to the $\sim 3 \times 10^5$ pixels with $\Sigma > 180\,\mathrm{M}_\odot\,\mathrm{pc}^{-2}$; this threshold is equivalent to $A_\mathrm{V} \sim 8$, and corresponds to the approximate threshold above which low-mass star formation in Ophiuchus, Aquila/Serpens, and Taurus appears to be concentrated (cf. Johnstone, Di Francesco & Kirk 2004; Könyves et al. 2015; Marsh et al. 2016; Ladjelate et al. 2020; Howard et al. 2021). Both axes are scaled logarithmically.

it must be presumed that the different dust-vibrational temperatures represent dust in different layers of a structure (cloud, filament, or shell) that is either predominantly heated by externally incident radiation, resulting in a positive radial temperature gradient; or predominantly heated by internally generated radiation (e.g. from stars), resulting in a negative radial temperature gradient.

We can however explore whether the two structural modes (monolithic and undulating) that we have identified in Section 7.1 are associated with different temperature regimes, given that the monolithic mode appears to be correlated with higher surface densities than the undulating mode. The top row of Fig. 9 shows the distribution of surface density in different dust-vibrational temperature intervals. In each panel the interval is defined by specifying a separatrix temperature, T_s. The black histogram gives the distribution of surface density for all dust. The red histogram then gives the distribution of surface density only for dust with vibrational temperature $T < T_\mathrm{s}$, whilst the blue histogram gives the distribution only for dust with $T > T_\mathrm{s}$. Reading from left to right, there are five values of $T_\mathrm{s} = 12.1\,\mathrm{K}$, 14.3 K, 16.9 K, 20.0 K, and 23.6 K.

The middle row of Fig. 9 shows the corresponding distributions, when we limit consideration to pixels with surface density above $23\,\mathrm{M}_\odot\,\mathrm{pc}^{-2}$. In the solar vicinity, this corresponds to visual extinction $A_\mathrm{V} \gtrsim 1$ mag.

The bottom row of Fig. 9 shows the corresponding distributions, when we limit consideration to pixels with surface density above $180\,\mathrm{M}_\odot\,\mathrm{pc}^{-2}$. In the solar vicinity, this corresponds to visual extinction $A_\mathrm{V} \gtrsim 8$ mag, and hence to the surface density above which low-mass star formation in Ophiuchus, Aquila/Serpens, and Taurus is concentrated (e.g. Johnstone et al. 2004; Könyves et al. 2015; Marsh et al. 2016; Ladjelate et al. 2020; Howard et al. 2021).

Several features of Fig. 9 are noteworthy. First, the distribution of surface density for all dust is, de facto, independent of T_s, and hence the black distribution is the same for each plot on the same row of Fig. 9. On the middle and bottom plots, the vertical black line marks the surface density cut-off ($23\,\mathrm{M}_\odot\,\mathrm{pc}^{-2}$ or $180\,\mathrm{M}_\odot\,\mathrm{pc}^{-2}$).

Secondly, if we focus on a single pixel, both the the surface density corresponding to dust below the separatrix temperature, and the surface density corresponding to dust above the separatrix temperature, are necessarily smaller than the surface density of all dust. Consequently (i) they do not normally contribute to the same histogram-bin as the surface density of all dust (but rather to a lower bin), and (ii) one or both may actually contribute to a bin below the cut-off. This means that the blue and red distributions can, *and inevitably do,* at some surface densities lie above the black histogram, especially, but not exclusively, at low surface densities.

Thirdly, the median dust temperature is $\sim 20\,\mathrm{K}$, and this result is more-or-less independent of whether we apply a surface density cut. In other words, irrespective of whether we consider all lines of sight or only lines of sight with surface density above the thresholds ($23\,\mathrm{M}_\odot\,\mathrm{pc}^{-2}$ or $180\,\mathrm{M}_\odot\,\mathrm{pc}^{-2}$), half the dust is hotter than $\sim 20\,\mathrm{K}$ and half is colder. At first sight, this result seems rather surprising, given that in simple radiation transport models of externally heated clouds, employing symmetric geometries, the coldest dust is located on lines of sight with large surface densities – and therefore predominantly the monolithic mode.

However, we should be mindful of the fact that in nature lines of sight that intercept dense molecular gas (i) usually also intercept large columns of more diffuse atomic gas, in which the dust is not significantly shielded from the ambient radiation field and is therefore relatively warm; and (ii) often take in dust that is heated by

nearby star-formation activity, for example dust in photo-dissociation regions (PDRs).

Fourthly, except where the surface densities are very low, both the cooler dust *and* the warmer dust approximate rather well to power-law probability distributions at large surface densities,

$$\frac{\mathrm{d}P}{\mathrm{d}\Sigma} \propto \frac{1}{\Sigma^3}. \tag{10}$$

This corresponds to a slope of -2 on a plot of $\log_{10}\{\mathrm{d}P/\mathrm{d}\log_{10}(\Sigma/[\mathrm{M}_\odot\,\mathrm{pc}^{-2}])\}$ against $\log_{10}(\Sigma/[\mathrm{M}_\odot\,\mathrm{pc}^{-2}])$ (cf. Girichidis et al. 2014; Schneider et al. 2015). The low surface densities, which appear to subscribe to a log-normal distribution, are only visible in the plots on the top row. Once a surface density threshold is introduced, they are not adequately sampled.

8 SUMMARY

We have developed and implemented a family of CNNs that can be used to analyse xfBm substructure in a pixelated map, in terms of the Hurst parameter, $\mathcal{H}$, and a scaling parameter, $\mathcal{S}$. For 2D images, the Hurst parameter, $\mathcal{H}$, is related to the power-law exponent, $\beta \simeq 2(1+\mathcal{H})$, the perimeter-area fractal dimension, $\mathcal{D}_{2:\mathrm{PA}} \simeq 2-\mathcal{H}$, and the box-counting fractal dimension $\mathcal{D}_{2:\mathrm{BC}} \simeq 3-\mathcal{H}$. The scaling parameter, $\mathcal{S}$, measures the range of intensities on the map. The CNNs are very efficient, in terms of the computational resource required to apply them. In addition, they evaluate $\mathcal{H}$ more accurately and faster than Δ-variance, and can operate on smaller fields than Δ-variance (fewer pixels). The CNNs also return estimates of $\mathcal{S}$, which Δ-variance does not.

The CNNs developed here can operate on patches of $\mathcal{N}\times\mathcal{N} = 128\times128$, 64×64, or 32×32 pixels. If the pixels are square, with angular size $\phi_{\mathrm{min}}\times\phi_{\mathrm{min}}$, the patch captures structures with size between $\sim\phi_{\mathrm{min}}$ and $\sim\phi_{\mathrm{max}} = \mathcal{N}\phi_{\mathrm{min}}$; thus $\mathcal{N}$ regulates the dynamic range of the substructures that can be captured by the CNN, and the accuracy with which the parameters $\mathcal{H}$ and $\mathcal{S}$ are estimated; larger $\mathcal{N}$ gives greater accuracy. We stress that when estimating the statistics of substructure on maps that are not robust xfBm fields, it is essential to be mindful of the range of structures that can be captured, i.e. the range $[\phi_{\mathrm{min}},\phi_{\mathrm{max}}]$, and that statistics estimated using different ranges may be different for robust physical reasons.

Smaller $\mathcal{N}$ allows the user to isolate a smaller patch on the map being analysed, and hence to estimate more distinctly localized values of $\mathcal{H}$ and $\mathcal{S}$. This has considerable advantages when the map being analysed contains regions characterized by very different substructure, and this is certainly the case for maps of the surface density of the interstellar medium, such as the Hi-GAL maps derived from thermal dust emission that are considered here. We implement a procedure for obtaining values of $\mathcal{H}$ and $\mathcal{S}$ for an individual pixel by considering, and averaging over, all possible positions of the CNN patch which overlap that pixel.

To explore the trade-off between a larger dynamic range (larger $\mathcal{N}$) and finer resolution (smaller $\phi_{\mathrm{max}} = \mathcal{N}\phi_{\mathrm{min}}$), we first apply the 32×32 CNN to the raw PPMAP maps, which have 8.5 arcsec pixels (these results are labelled '32×8.5'); then we apply the 32×32 CNN to rebinned PPMAP maps, which have 17 arcsec pixels (these results are labelled '32×17'); and finally we apply the 64×64 CNN to the raw PPMAP maps (these results are labelled '64×8.5').

Regions of higher surface density ($\gtrsim 32\,\mathrm{M}_\odot\,\mathrm{pc}^{-2}$) are concentrated towards the inner Galaxy, and tend to have larger $\mathcal{H}\,(\gtrsim 0.8)$ (implying a steep power spectrum, with power concentrated on large angular scales) and larger $\mathcal{S}\,(\gtrsim 1)$ (a broad range of surface densities); we term this the monolithic mode. Regions of lower surface density are more ubiquitous, and tend to have lower $\mathcal{H}\,(\lesssim 0.8)$ (a shallower power spectrum, with power on a broader range of angular scales) and lower $\mathcal{S}\,(\lesssim 1)$ (a smaller range of surface density); we term this the undulating mode. These structural differences between different regions are most apparent when the analysis is performed with 64×8.5, and least apparent when the analysis is performed with 32×8.5.

We also analyse the distribution of surface density as a function of dust temperature, distinguishing the surface density associated with dust at temperatures less than, and more than, a user-defined separatrix temperature T_{s}. This is done for $T_{\mathrm{s}} = 12.1\,\mathrm{K}$, $14.3\,\mathrm{K}$, $16.9\,\mathrm{K}$, $20.0\,\mathrm{K}$, and $23.6\,\mathrm{K}$. The median dust temperature is $\sim 20\,\mathrm{K}$. For all T_{s}, both the cooler dust ($< T_{\mathrm{s}}$), and the warmer dust ($> T_{\mathrm{s}}$), approximate well to a power-law surface density distribution, $\mathrm{d}P/\mathrm{d}\Sigma \propto \Sigma^{-3}$ (corresponding to a slope of -2 on a log–log plot) at large surface density. The distribution of dust-vibrational temperature appears to be very similar for the two different structural modes (monolithic and undulating).

ACKNOWLEDGEMENTS

We thank the referee for their careful, thorough, and thought-provoking report, which helped us to improve significantly the original version of the paper. We also thank Sarah Ragan for helpful discussions. MLB gratefully acknowledges the receipt of a PhD studentship from the UK Science and Technology Facilities Council (STFC) through the Centre for Doctoral Training (CDT) in Data Intensive Science (ST/P006779/1). APW gratefully acknowledges the support of an STFC Consolidated Grant (ST/K00926/1). This work was performed using the computational facilities of the Advanced Research Computing at Cardiff (ARCCA) Division, Cardiff University.

DATA AVAILABILITY

The code used in this work and in generating the images can be found at https://github.com/MBates92/HGCNN. The data used and generated in this this work can be made available on request.

REFERENCES

Bates M. L., Whitworth A. P., Lomax O. D., 2020, MNRAS, 493, 161
Bazell D., Desert F. X., 1988, ApJ, 333, 353
Chappell D., Scalo J., 2001, ApJ, 551, 712
Dame T. M., Hartmann D., Thaddeus P., 2001, ApJ, 547, 792
Dib S. et al., 2020, A&A, 642, A177
Dickman R. L., Margulis M., Horvath M. A., 1990, ApJ, 365, 586
Elia D. et al., 2014, ApJ, 788, 3
Elia D. et al., 2018, MNRAS, 481, 509
Elmegreen B. G., 1997, ApJ, 477, 196
Elmegreen B. G., Kim S., Staveley-Smith L., 2001, ApJ, 548, 749
Falgarone E., Phillips T. G., Walker C. K., 1991, ApJ, 378, 186
Federrath C., Klessen R. S., Schmidt W., 2009, ApJ, 692, 364
Galliano F. et al., 2021, A&A, 649, A18
Girichidis P., Konstandin L., Whitworth A. P., Klessen R. S., 2014, ApJ, 781, 91
Griffin M. J. et al., 2010, A&A, 518, L3
Hou L. G., Han J. L., 2014, A&A, 569, A125
Hou L. G., Han J. L., 2015, MNRAS, 454, 626
Howard A. D. P., Whitworth A. P., Griffin M. J., Marsh K. A., Smith M. W. L., 2021, MNRAS, 504, 6157
Jackson J. M. et al., 2006, ApJS, 163, 145
Jáquez-Domínguez J. M. et al., 2023, preprint (arXiv:2304.04864)
Johnstone D., Di Francesco J., Kirk H., 2004, ApJ, 611, L45

Juvela M., 2023, preprint (arXiv:2304.05102)
Kainulainen J., Beuther H., Henning T., Plume R., 2009, A&A, 508, L35
Kalberla P. M. W., Burton W. B., Hartmann D., Arnal E. M., Bajaja E., Morras R., Pöppel W. G. L., 2005, A&A, 440, 775
Könyves V. et al., 2015, A&A, 584, A91
Ladjelate B. et al., 2020, A&A, 638, A74
Larson R. B., 1981, MNRAS, 194, 809
Mac Low M.-M., Glover S. C. O., 2012, ApJ, 746, 135
Mandelbrot B. B., Cannon J. W., 1984, Am. Math. Monthly, 91, 594
Marchuk A. A., Smirnov A. A., Mosenkov A. V., Il'in V. B., Gontcharov G. A., Savchenko S. S., Román J., 2021, MNRAS, 508, 5825
Marsh K. A., Whitworth A. P., Lomax O., 2015, MNRAS, 454, 4282
Marsh K. A. et al., 2016, MNRAS, 459, 342
Marsh K. A. et al., 2017, MNRAS, 471, 2730
Molinari S. et al., 2010, Publ. Astron. Soc. Pac., 122, 314
Ormel C. W., Paszun D., Dominik C., Tielens A. G. G. M., 2009, A&A, 502, 845
Ormel C. W., Min M., Tielens A. G. G. M., Dominik C., Paszun D., 2011, A&A, 532, A43
Ossenkopf V., Krips M., Stutzki J., 2008, A&A, 485, 719
1988, Peitgen H.-O., Saupe D.eds, The Science of Fractal Images. Springer, New York, Berlin
Pilbratt G. L. et al., 2010, A&A, 518, L1
Poglitsch A. et al., 2010, A&A, A518, L2
Robitaille T., Deil C., Ginsburg A., 2020a, reproject: Python-based astronomical image reprojection, Astrophysics Source Code Library, record ascl:2011.023
Robitaille J. F., Abdeldayem A., Joncour I., Moraux E., Motte F., Lesaffre P., Khalil A., 2020b, A&A, 641, A138
Sadavoy S. I. et al., 2012, A&A, 540, A10
Sanchez N., Alfaro E. J., Perez E., 2005, AJ, 625, 849
Schneider N. et al., 2012, A&A, 540, L11
Schneider N. et al., 2015, MNRAS, 453, L41
Schneider N. et al., 2022, A&A, 666, A165
Stutzki J., Bensch F., Heithausen A., Ossenkopf V., Zielinsky M., 1998, A&A, 336, 697
Vogelaar M. G. R., Wakker B. P., 1994, A&A, 291, 557
Voss R. F., 1988, in Peitgen H.-O., Saupe D., eds, The Science of Fractal Images. Springer, New York Berlin, p. 65
Wenger T. V., Balser D. S., Anderson L. D., Bania T. M., 2019, ApJ, 887, 114
Williams J. P., Blitz L., McKee C. F., 2000, in Mannings V., Boss A. P., Russell S. S.eds, Protostars and Planets IV. University of Arizona Press, Tucson, p. 97

APPENDIX A: MULTIFRACTAL ANALYSIS

As an illustration of the difficulty of interpreting multifractal spectra, we have revisited the plot of D_{+20} (fractal dimension of order +20) against D_{-10} (fractal dimension of order −10), which was explored by Elia et al. (2018) (their fig. 19), as one way of encapsulating the products of multifractal analysis. The order, θ, of the generalized fractal dimension, D_θ, can take any real value between $\theta = -\infty$ and $\theta = +\infty$. With large positive θ, D_θ reflects the scaling of the higher density (i.e. more contrasted) structures within the map. Conversely, with large negative θ, D_θ reflects the scaling of lower density structures. For details of how D_θ is computed, the reader is referred to Elia et al. (2018), where they show that a pure fBm field with a specific power-law exponent $\beta = 2.4$ (equivalently $\mathcal{H} = 0.2$), and a specific random seed, here labelled 'C' (hence a given overall structure, like the images in our Fig. 1) corresponds to a specific $(D_{-10}, D_{+20}) = (2.25, 1.96)$. However, this would seem to be of limited use, since a viable statistical parameter must be independent of the random seed used to generate a particular realisation.

Their fig. 19 is much more interesting because it shows the distribution of (D_{-10}, D_{+20}) for a large sample of trial fields with a range of different characteristics, and is able to distinguish some features quite clearly. (i) The HiGAL fields have $2.0 \lesssim D_{-10} \lesssim 2.3$, and $0.7 \lesssim D_{+20} \lesssim 1.8$, suggesting that the higher density structures have a broader range of scalings than the lower density ones. (ii) In contrast, their fBm fields have $2.0 \lesssim D_{-10} \lesssim 2.9$ and $D_{+20} \sim 1.95$ suggesting a broad range of scalings for the lower density structures, and an almost universal, marginally fractal scaling for the higher density ones, as should be expected. (c) Finally the fields from numerical simulations do not subscribe to either of these patterns. D_{+20} and D_{-10} tend to be anticorrelated, i.e. strong scaling of higher density structures tends to combine with weak scaling of lower density structures, and vice versa. It is therefore interesting to explore where xfBm fields lie on this plot.

Fig. A1 shows the results obtained for xfBm fields generated using the procedure described in Section 2, but with $[\mathcal{N} \times \mathcal{N}] = [10^2 \times 10^2]$ and $[10^3 \times 10^3]$. The fields analysed have been generated with three discrete values of $\mathcal{S}$, which are represented with different markers in Fig. A1: $\mathcal{S} = 0.1$, represented with filled circular dots; $\mathcal{S} = 1$, represented with stars; and $\mathcal{S} = 3$, represented with crosses. For each value of $\mathcal{S}$, we generate 1000 realizations with random

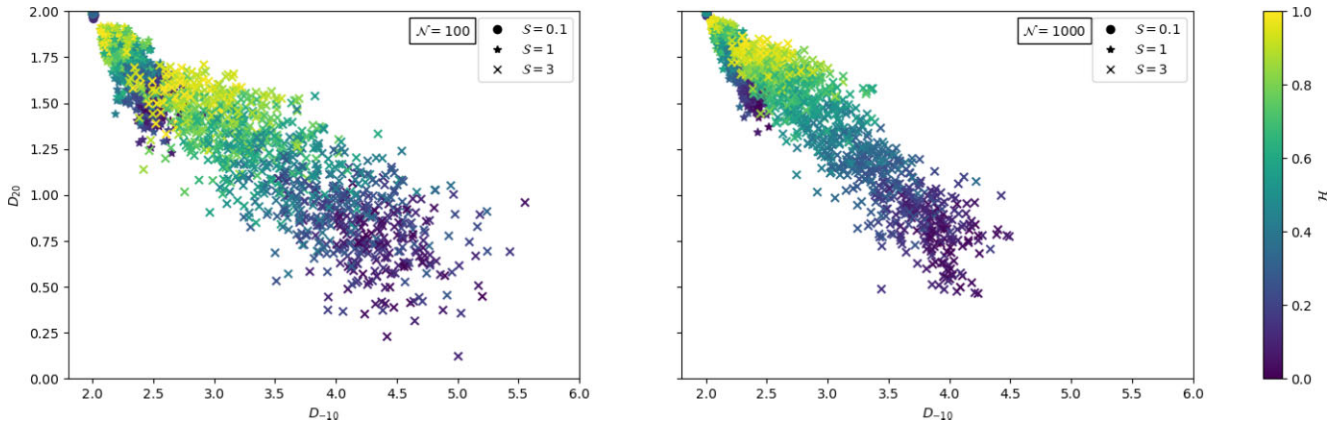

Figure A1. Plots of the fractal dimension of the order of +20 (D_{+20}) against the fractal dimension of the order of −10 (D_{-10}). Fields with $\mathcal{S} = 0.1$ are represented with solid circular dots; fields with $\mathcal{S} = 1$ are represented with stars; and fields with $\mathcal{S} = 3$ are represented with crosses. The colours of the symbols encode the value of $\mathcal{H}$, according to the colour-bar on the right. The left-hand plot is obtained with fields having $[\mathcal{N} \times \mathcal{N}] = [10^2 \times 10^2]$ pixels, and the right-hand plot with fields having $[10^3 \times 10^3]$ pixels.

Table C1. The discrete dust temperatures used in the Marsh et al. (2017) PPMAP analysis. The first row gives q, the dust temperature ID. The second row gives the discrete dust temperature, T_q. The third and fourth rows give, respectively, the lower limit, $0.920\,T_q$, and the upper limit, $1.087\,T_q$, for the range of temperatures represented by T_q. The lowest discrete temperature, $T_1 = 8.00$ K, must actually represent dust below its listed lower limit, '7.36 K', and dust at these very low temperatures is therefore poorly represented; however, the results suggest that the amount of dust at these very low temperatures is extremely small. Similarly, the highest discrete temperature, $T_{12} = 50.0$ K, must actually represent dust above its listed upper limit '54.3 K', and dust at these very high temperatures is therefore poorly represented; however, the results suggest that the amount of dust at these very high temperatures is again extremely small.

q	1	2	3	4	5	6	7	8	9	10	11	12
T_q/K	8.00	9.45	11.2	13.2	15.6	18.4	21.7	25.7	30.3	35.8	42.3	50.0
$> 0.920\,T_q$/K	'>7.36'	>8.69	>10.3	>12.1	>14.3	>16.9	>20.0	>23.6	>27.9	>33.0	>38.9	>46.0
$\leq 1.087\,T_q$/K	≤8.69	≤10.3	≤12.1	≤14.3	≤16.9	≤20.0	≤23.6	≤27.9	≤33.0	≤38.9	≤46.0	'≤54.3'

values of $\mathcal{H}$ distributed uniformly on the interval [0,1], and the markers are colour-coded according to the value of $\mathcal{H}$, using the colour-bar on the right-hand side of the figure. There is a clear trend on the (D_{-10}, D_{+20}) plane, and unsurprisingly it is somewhat better defined with the [$10^3 \times 10^3$] fields (right-hand plot) than with the [$10^2 \times 10^2$] fields (left-hand plot). However, given the scatter, the trend with increasing $\mathcal{S}$, from low D_{-10} and high D_{+20} to high D_{-10} and low D_{+20}, is almost parallel to the trend with decreasing $\mathcal{H}$. Consequently this plot has limited use *as a diagnostic of xfBm fields*. An xfBm image with given (D_{-10}, D_{+20}) could have a range of $\mathcal{H}$ and $\mathcal{S}$ values, satisfying the approximate relation

$$\mathcal{S}[1.2 - \mathcal{H}]1.5\left\{[D_{-10}-2]^2+[D_{+20}-2]^2\right\}, \quad 0 \leq \mathcal{H} \leq 1. \tag{A1}$$

In contrast, the CNN-based procedure developed here raises this degeneracy.

APPENDIX B: THE FIDELITY OF THE HURST PARAMETER FOR NON-FBM FIELDS

In Bates et al. (2020) and this paper, we have applied Δ-variance and CNNs to determine the Hurst parameters of fields that are not true fractional Brownian motion (fBm) fields – albeit that they are derived from true fBm fields (by exponentiation plus additional procedures to render them non-periodic and noisy; see Section 2). Therefore they always have a well defined underlying Hurst parameter. Similarly Stutzki et al. (1998) successfully apply Δ-variance to fields that have been obtained by squaring fBm fields. The question then arises as to why these techniques (Δ-variance and CNNs) are still able to produce reasonably accurate estimates of the underlying Hurst parameter.

We suggest here that this is because both techniques ultimately, if indirectly, measure the contour structure of the fields to which they are applied – as indeed do the area-perimeter and box-counting procedures for estimating fractal dimension. Neither exponentiation, nor squaring, alters the structure of the contours, and therefore the Hurst parameter is not changed very much, if at all. In the case of exponentiated (xfBm) fields, the height of the contours is of course changed by exponentiation, but this is reflected in the scaling parameter, $\mathcal{S}$.

Indeed, we suggest that statistical characterization of a 2D field necessarily requires two parameters, one reflecting the contour structure (here $\mathcal{H}$) and one reflecting the relative heights of the contours (here $\mathcal{S}$). Thus, in principle, a map of the Netherlands might have the same $\mathcal{H}$ as a map of Switzerland, but a very different $\mathcal{S}$. (In reality the $\mathcal{H}$ values for the two countries are probably very different because the geographical processes shaping their surfaces have been very different.)

APPENDIX C: THE PPMAP ANALYSIS PROCEDURE

The PPMAP analysis procedure used by Marsh et al. (2017) to analyse the Hi-GAL tiles assumes that the radiation detected in the *Herschel* bands is primarily thermal emission from dust; that the mass opacity coefficient of the emitting dust can be approximated by

$$\kappa_\lambda = \kappa_{300}\ [\lambda/300\,\mu\mathrm{m}]^{-2} \tag{C1}$$

(Sadavoy et al. 2012); and that the emission is optically thin. The intensity on a given line of sight is then given by

$$I_\lambda \simeq \int_{T=0}^{T=\infty} B_\lambda(T)\,\frac{\mathrm{d}\tau_\lambda}{\mathrm{d}T}\,\mathrm{d}T. \tag{C2}$$

In equation (C2), $B_\lambda(T)$ is the Planck Function; T is the dust-vibrational temperature;

$$\tau_\lambda(T) = \kappa_\lambda \int_{s=0}^{s=\infty} \rho(T, s)\,\mathrm{d}s \tag{C3}$$

is the optical-depth along the line of sight due to dust at vibrational temperatures less than T; and $\rho(T, s)$ is the density of dust with vibrational temperature less than T at distance s along the line of sight. Equation (C3) assumes that κ_{300}, and hence κ_λ, do not vary along the line of sight; otherwise κ_λ cannot be taken outside the integral.

The standard procedure for analysing multiwavelength images of dust emission (*Modified Blackbody Fitting*) assumes that all the dust along a given line of sight can be accurately represented by a single mean vibrational temperature, $\bar{T}$. Equations (C2) and (C3) then reduce to

$$I_\lambda \simeq B_\lambda\left(\bar{T}\right)\ \tau_{300}\ [\lambda/300\,\mu\mathrm{m}]^{-2}, \tag{C4}$$

where τ_{300} is the optical-depth at 300 μm, due to all the dust along the line of sight. To estimate $\bar{T}$ and τ_{300}, the observational data from the different bands is first rebinned to the coarsest resolution (for *Herschel* this is 36 arcsec at 500 μm) and then the spectral energy distribution is fit with equation (C4). This procedure tends to overestimate τ_{300}, because the overestimated contribution from warmer than average dust out-weighs the underestimated contribution from cooler than average dust (e.g. Marsh et al. 2015; Jáquez-Domínguez et al. 2023; Juvela 2023). It also ignores a lot of high-resolution information from the shorter wavelength *Herschel* bands.

The PPMAP analysis procedure is predicated on the assumption that on most lines of sight there is a significant range of dust temperatures, and that this range can be represented by Q discrete dust temperatures, between $T_{\rm MIN}$ and $T_{\rm MAX}$, *viz.*

$$T_1 = T_{\rm MIN}; \tag{C5}$$

$$T_q = F^2 T_{q-1}\,, \qquad 2 \le q \le Q; \tag{C6}$$

where

$$F = \left[T_{\rm MAX}/T_{\rm MIN}\right]^{1/2[Q-1]}\,; \tag{C7}$$

and hence $T_Q = T_{\rm MAX}$. Each discrete dust temperature T_q therefore represents a finite range of dust temperature, $T_q/F < T \le FT_q$.

The integral in equation (C2) is then approximated by a sum over the discrete dust temperatures,

$$I_\lambda \simeq \sum_{q=0}^{q=Q} \left\{B_\lambda\left(T_q\right)\,\Delta\tau_{300:q}\right\}\,[\lambda/300\,\mu{\rm m}]^{-2}\,. \tag{C8}$$

In equation (C8), $\Delta\tau_{300:q}$ is the contribution to the optical depth at 300 μm, from the dust at temperature T_q, or – strictly speaking – the dust in the temperature-range $T_q/F < T \le FT_q$, i.e.

$$\Delta\tau_{300:q} = \int_{T=T_q/F}^{T=FT_q} \frac{{\rm d}\tau_{300}}{{\rm d}T}\,{\rm d}T. \tag{C9}$$

Marsh et al. (2017) adopt $Q = 12$, $T_{\rm MIN} = 8.00$ K, and $T_{\rm MAX} = 50.0$ K. Hence $F = [50/8]^{1/22} = 1.087$. The resulting discrete dust temperatures, T_q, and the finite ranges that they represent are given in Table C1.

The PPMAP procedure starts by giving the optical-depth contributions, $\Delta\tau_{300:q}$, uncorrelated random values from a Gaussian distribution, and computing the resulting synthetic integrated intensities in the observed *Herschel* bands, Hb,

$$I_{\rm Hb} = \int_{\lambda=0}^{\lambda=\infty} I_\lambda\,\mathcal{F}_{\rm Hb}(\lambda)\,{\rm d}\lambda. \tag{C10}$$

where $\mathcal{F}_{\rm Hb}(\lambda)$ is the spectral response function in the Hb band. These synthetic intensities are then compared with the true observed intensities, but with the uncertainties artificially inflated, so that, despite the mismatch between the synthetic intensities and the true intensities being large, the fit is tolerable – in the sense that the adjustments to the $\Delta\tau_{300:q}$ values that are required to improve the fit are in the linear regime. This process is then repeated recursively with the uncertainties slowly but steadily reduced until they correspond to the true noise. PPMAP thereby returns expectation values and uncertainties for the optical-depth contributions, $\Delta\tau_{300:q}$.

By using the data at its native resolution, PPMAP delivers images with the finest resolution; for *Herschel* this is $\phi_{\rm min} = 8.5$ arcsec at 70 μm.[2] Thus each [2.4° × 2.4°] tile furnishes 24 images, each comprising [1440 × 1440] pixels, and representing the expectation values for $\Delta\tau_{300:q}$, and the corresponding uncertainties, $\sigma_{300:q}$, for the 12 different temperatures, T_q.[3] The effect of interpolating the long-wavelength emission to pixels that are smaller than the long-wavelength resolution is reflected in the uncertainties. By invoking a range of temperatures, the contributions from warmer than average dust and cooler than average dust are estimated more accurately. The estimated total optical-depth at the reference wavelength is

$$\tau_{300} = \sum_{q=1}^{q=Q} \left\{\Delta\tau_{300:q}\right\}. \tag{C11}$$

[2]The PACS beam at 70 μm is distorted by the parallel PACS/SPIRE fast-scan mode, and has dimensions ∼6 arcsec × 12 arcsec. Since the Hi-GAL maps combine Herschel observations from two orthogonal scan directions Marsh et al. (2017) adopt a circular beam with FWHM equal to the geometric mean of these dimensions (although strictly speaking the combined beam is not exactly circular).

[3]It is tempting to assume that the dust at different temperatures is concentrated at different locations along the line of sight, and sometimes this may be a reasonable approximation. However, it should be kept in mind that in reality some of the dust at different temperatures could simply be dust of different types mixed together in the same region, or even small transiently heated grains caught at different stages as they cool following a temperature spike. Under this circumstance, κ_{300} should be a sum of the mass-opacity coefficients of the different types of dust, weighted in proportion to their masses.

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