Chapter 2

Probing large scale turbulence through radio interferometric observation of H I specific intensity

In this chapter, we discuss the means of probing the large scale structures in the galaxies using radio interferometric observations. We discuss here various statistical estimators to probe the coherent structure and dynamics of the ISM. Finally, we discuss the Visibility Moment Estimator for H I column density and velocity power spectrum (VME) that is used in the next chapter.

2.1 H I as a probe for structure and dynamics

Different components of the ISM can be probed by various spectral lines which measure the structures and dynamics of the ISM. From the spectral lines of elements, not only the density distributions of the elements in the ISM but also their dynamics can be traced due to the doppler shift of the line. Several strong spectral lines originated from ISM are Ca^+, Na, CO etc (Goldreich and Kwan, 1974; Wilson et al., 1970; Zuckerman and Palmer, 1974). But these elements constitute very little in the ISM and present unevenly in the entire galactic disc. Around 70% of the ISM constitute atomic hydrogen (H I). H I is evenly distributed around the galaxy and can be used as a tool to probe the structure and dynamics of the ISM. Here we use the hyperfine structure transition line (ν_0 =1420.40575 MHz, λ =21cm) or 21cm spectral line to observe H I .

External spiral galaxies are located around small patches in the sky. We define cartesian coordinates with origin at the galactic centre, where the position vector is defined as $\vec{r} = (x, y, z)$. Here z is the line of sight of direction and $\vec{R} = (x, y)$ is a vector in the sky plane, perpendicular to the line of sight direction. The specific intensity of radiation of 21cm line emission coming from neutral hydrogen in the galactic disc is defined as follows (Draine, 2011),

$$I(\vec{\theta}, \nu) = I_0 \int \mathrm{d}z\, n_{\mathrm{HI}}(\vec{r})\phi(\nu), \qquad I_0 = \frac{3}{16} h\nu_0 A_{21} \tag{2.1}$$

where A_{21} is the Einstein coefficient of 21cm radiation. Here $\vec{\theta}$ is the angular seperation in the sky plane defined as $\vec{\theta} = \frac{\vec{R}}{D}$, where D is the distance to the galaxy, n_{HI} is the number density of atomic hydrogen and $\phi(\nu)$ is the line shape function. The latter is normalized to unity when integration over all frequency is considered. Owing to the internal dynamics of the galaxy, the hydrogen gas in different parts of the galaxy have different velocities and hence observed frequency of the emission depends on the Doppler shift of the line. Hence the specific intensity can be written as a function of the Doppler velocity v as,

$$I(\vec{\theta}, v) = I_0 \int \mathrm{d}z\, n_{\mathrm{HI}}(\vec{r})\phi(v), \qquad \phi(v) = \exp\left(-\frac{[v - v_z(\vec{r})]^2}{2\sigma^2}\right) \tag{2.2}$$

where $v_z(\vec{r})$ is the line of sight velocity and σ gives the velocity dispersion.

The information of density and velocity can be extracted by calculating the moments of the intensity distrubution $I(\vec{\theta}, v)$. The first three moments are defined as follows,

$$
\begin{aligned}
M_0(\vec{\theta}) &= \int \mathrm{d}v \, I(\vec{\theta}, v) = I_0 N_{\mathrm{HI}}(\vec{\theta}), && (2.3)\\
M_1(\vec{\theta}) &= \frac{\int \mathrm{d}v \, I(\vec{\theta}, v) v}{M_0(\vec{\theta})} && (2.4)\\
M_2(\vec{\theta}) &= \sqrt{\frac{\int \mathrm{d}v \, I(\vec{\theta}, v)(v - M_1(\vec{\theta}))^2}{M_0(\vec{\theta})}}. && (2.5)
\end{aligned}
$$

The $N_{\mathrm{HI}}(\vec{\theta})$ is the column density of atomic hydrogen. The zeroth moment $M_0(\vec{\theta})$ gives us the column density distribution of neutral atomic hydrogen. The first moment $M_1(\vec{\theta})$ gives the density weighted line of sight velocity distribution. The second moment $M_2(\vec{\theta})$ is the density weighted velocity dispersion fields. Hence by observing the H I 21cm emission we can measure the column density and line of sight velocity of atomic hydrogen present in the galaxy. This is the widely used definition of the H I moment maps. If the specific definition of moments are adapted during the calculation, we separately define them in adequate places.

We here introduce to the reader several parameters of a typical disc galaxy. The disc galaxies have a morphology of a disc, where the radius of H I disc extends typically upto $\sim$ 40 kpc with a width of a few 100pc. Galaxies are inclined at different angles in the plane of the sky where we are observing. Hence a galactic plane makes a certain angle with the line of sight direction z and we call this inclination angle i. The observed column density and line of sight velocity information from the observing 21cm emission is also affected by the inclination angle. A typical spiral galaxy is similar to a cylindrical disc in shape with some finite disc thickness. But due to the inclination effect, what we observe is the projection of the actual galactic disc along the line of sight direction. This makes the circular galactic plane looks more like an ellipse. The major axis of this observed ellipse is differently oriented to the North-South direction in the sky. We define the angle from the

North Celestial pole to the major axis in the East direction as position angle PA and this is unique for each galaxy. The disc thickness is called scale height which can be different for different galaxies. Measuring scale height is not straightforward due to limitations in observing the three-dimensional structure of the galactic disc. Only edge-on galaxies whose inclination angles are near 90° give access to the study of vertical structures in the disc including measurement of scale heights. In disc galaxies, H I gas is distributed to much extended to the outskirts of the stellar disc where stars are absent. H I scale length is defined as the radial extent of H I disc in the galaxies. All the disc galaxies are found to have large H I scale lengths with the respective stellar scale length.

2.2 Radio Interferometric observation

The 21cm emission line belongs to the radio regime of the electromagnetic spectrum. Hence it is obvious that we use radio telescopes in order to observe this radiation. Typically resolution of a single radio dish depends on its diameter and the observing wavelength λ. Single dish radio telescopes have limited resolutions to probe large scale structures in the sky. Later the introduction of radio interferometers paves the way for the revolution in radio wave observations. Radio interferometers are composed of numerous array elements called antennas or tiles which all together act as a single dish radio telescope, that can produce high resolution radio images. Very Large Array (VLA), Giant Metrewave Radio Telescope (GMRT), Westerbok Synthesis Radio Telescope (WSRT) telescopes are some of the important radio interferometers which have been used by the radio astronomy community worldwide since 1990s. Upcoming facilities like the Square Kilometer Array (SKA) telescope and its precursors like Australian SKA Pathfinder (ASKAP), MeerKAT, and Murchison Widefield Array (MWA) are next generation radio telescopes which give much higher resolution and sensitive radio data.

In radio interferometers, the collection of many antennas is arranged in a specific pattern on the surface of the earth. Each antenna records the electric field of the electromagnetic wave coming from a particular direction of the sky. The physical size of the individual antenna, the electronic receiver chain and the observing wavelengths limit the sensitivity of the antenna as well as the entire interferometer to a limited portion of the sky. This is known as the field of view (FOV) of the interferometer. Electric fields from each antenna pair are correlated and recorded. This quantity is known as the visibility function. The pair of antennas for which the visibility is recorded is called a baseline. The baseline vector $\vec{U}$ is the ratio of instantaneous projected separation of the antenna pair on a plane perpendicular to the direction of the incoming wave from the sky to the observing wavelengths. Clearly, the visibilities are functions of the baseline vector and observing frequency, i.e, $V(\vec{U},\nu)$. The visibility function is essentially the spatial coherence function of the electric field measured by the antenna pairs. The van Cittert-Zernike theorem states that the spatial coherence function or visibility function when measured confining to a plane and only in a small region of the sky, is the complex Fourier transform of the intensity or sky brightness distribution (Thompson et al., 2017). Hence we have,

$$V(\vec{U},\nu) = \mathcal{FT}\{I(\vec{\theta},\nu)\}. \tag{2.6}$$

Ideally, by taking the inverse Fourier transform of the visibility function we can reconstruct the intensity or sky brightness distribution. However, in real observations, visibilities can be measured only at the discrete values $\vec{U}_i$, which correspond to the physical baselines offered by the antenna pairs. We introduce a function $S(\vec{U})$ to capture this sampling of the baselines by the interferometer:

$$S(\vec{U}) = \sum_{i=1}^{N_b} \delta_D(\vec{U} - \vec{U}_i). \tag{2.7}$$

Here N_b is the total number of baselines. Note that in the rest of this particular section, we do not explicitly show the frequency dependence of visibility functions in equations for the sake of simplicity. Fourier relation between the intensity and measured visibilities are modified as follows (Sault and Conway, 1999),

$$V(\vec{U}) = \tilde{I}(\vec{U})S(\vec{U}) + \mathcal{N}(\vec{U}), \tag{2.8}$$

were $\tilde{I}(\vec{U})$ is the Fourier transform of the sky brightness distribution $I(\vec{\theta})$ and $\mathcal{N}(\vec{U})$ is the measurement noise. The discrete visibilities are the complex quantity directly measured by the radio interferometers. Note that, to completely describe a real measurement, we need to also scale the first term on the right hand side of the eqn (2.8) by the gain of the interferometer. To estimate the specific intensity at different points in the sky, reconstruction of the sky brightness distribution is necessary from the observed visibilities.

Inverse Fourier transform of the measured visibility is called the dirty image:

$$I_D(\vec{\theta}) = I(\vec{\theta}) \otimes B_D(\vec{\theta}). \tag{2.9}$$

The sampled visibilities are weighted in the baseline plane and grided for faster computation. Here $B_D(\vec{\theta})$ is the Inverse Fourier transform of the weighted sampling function $S(\vec{U})$ and essentially the Point Spread Function (PSF) of the interferometer. The weighting schemes are discussed shortly. The symbol $\otimes$ denotes convolution here. We have neglected the measurement noise for simplicity. The PSF of the interferometer is complicated and is often called the dirty beam. Reconstruction of the sky brightness distribution is essentially a deconvolution of the interferometer PSF from the dirty image. Since the sampling function can be quite discrete, irregular and incomplete, the interferometer PSF can be quite complicated thereby making the deconvolution procedure non-trivial. Different algorithms have been devised for this purpose including types of CLEAN (Clark, 1980;

Cotton, 1979; Högbom, 1974; Schwab, 1984), Maximum Entropy Image Reconstruction (MEM) (Narayan and Nityananda, 1986) and RESOLVE (Junklewitz et al., 2016). In this paper, we focus on the CLEAN algorithm, which is the most widely used algorithm in the radio astronomy community to date (Sault and Conway, 1999).

For most of the interferometers, the visibilities are not measured at regular intervals (like in a regular grid) but rather at apparently random locations in the baseline plane. Hence, values of the visibility function are estimated through interpolations at regular intervals in the baseline plane to initiate FFT. This process is called gridding. Further, different grid points in the baseline plane usually have a different number of measurements of the visibilities and they can be given different weights while constructing the image. This process is called weighting. Note that, some of these grid points may not have any measurements and a part of the baseline plane may be completely void of measured visibilities. This is termed the incomplete baseline coverage of the interferometer. The weighting schemes try to partially rectify the effect of the incomplete coverage. To evaluate the Fourier transform of the measured visibilities, the observed visibilities are weighted and put in regularly spaced grid points. Two extreme weighting schemes used are called uniform and natural weightings respectively. In the case of natural weighting, all the measurements are given the same weight and added together. This type of weighting emphasizes the part of the visibility plane where more measurements are present. Note that, when a weighting scheme is used, the effective dirty beam is to be considered as the Fourier transform of the sampling function multiplied by the weighting kernel. As for most of the interferometers, the baseline coverage is better at lower values of $\mid \vec{U} \mid$ and the PSF derived from the naturally weighted visibilities is broader with fewer side lobes. On the other hand, in a different weighting scheme called uniform weighting, the visibilities are weighted by the local density of measured visibilities before gridding. This type of weighting results in a narrower PSF but has a higher power in the side lobes. The robust weighting scheme designed by Briggs (1995) tries to combine the natural and uniform weightings and is used mostly based on the need of the particular imaging problem.

2.3 Brief overview of two point correlation statistics

The large-scale distribution of H I extracted by various one point statistics of sky brightness carries important information about its interplay with star formation in spiral and dwarf galaxies. The information about the existence of stochastic random fluctuations in the density and velocity and their origin can only be extracted by finding the two point correlation statistics of the intensity. Two-point statistics of any field quantifies the scale dependence of fluctuations in it (Elmegreen and Scalo, 2004). Here we discuss various two point statistics.

2.3.1 Autocorrelation function and Structure function

For a two dimensional scalar field $A(\vec{\theta})$, the structure function and the autocorrelation function evaluate the two-point statistics in the $\vec{\theta}$ plane. We use the 2D cartesian vector $\vec{\theta}$ to denote points in a tangential plane centered at a direction in the cellestial hemisphere. Autocorrelation function of homogenous and isotropic field $A(\vec{\theta})$ is given as,

$$\zeta_A(|\Delta\vec{\theta}|) = \left\langle A(\vec{\theta})A(\vec{\theta} + \Delta\vec{\theta})\right\rangle \tag{2.10}$$

Similarly the structure function of order 2 field A is defined as,

$$S_2(|\Delta\vec{\theta}|) = \left\langle \left[A(\vec{\theta}) - A(\vec{\theta} + \Delta\vec{\theta})\right]^2\right\rangle \tag{2.11}$$

where in both cases the angular bracket denotes an average over an ensemble.

2.3.2 Power spectrum

The power spectrum of field A is defined as

$$\left\langle \tilde{a}^*(\vec{U})\tilde{a}(\vec{U}')\right\rangle = \delta_{2D}(\vec{U}-\vec{U}')P(\vec{U}) \tag{2.12}$$

where $\vec{U}$ is Fourier conjugate of $\vec{\theta}$ and $\tilde{a}(\vec{U})$ is Fourier transform of $A(\vec{\theta})$. It can be shown that the power spectrum is the Fourier transform of the auto correlation function. For a Gaussian random field, all these different estimators contain the same information. For turbulence genrated random fluctuations, these estimators are expected to follow power laws (Federrath et al., 2009; Fleck, 1996) with the amplitude and the slope having information about the nature of the turbulence and its source.

Power spectrum is a commonly used quantifier of turbulence statistics, where the column density power spectrum and in some cases the line of sight velocity power spectrum can be measured. Kolmogorov (1941) formulated the incompressible fluid turbulence where the slope of the velocity power spectrum is predicted to be $-5/3$. Much of the information about ISM turbulence in the Galaxy is known from the measurements of power spectra in H I emission (Crovisier and Dickey, 1983; Green, 1993) and absorption (Chepurnov et al., 2015; Deshpande et al., 2000; Elmegreen et al., 2001; Roy et al., 2010). Begum et al. (2006) have estimated the power spectra of H I intensity fluctuations for the Large and Small Magellanic Clouds, DDO 210 and inferred the existence of large scale coherent structures in those dwarf galaxies suggesting long-range turbulence energy cascade. Dutta et al. (2009) and later Maier et al. (2017); Zhang et al. (2012) have found the existence of large scale turbulence cascades in a large number of dwarfs. Turbulence induces fragmentation in gas and helps to form molecular clouds. The turbulence energy cascade carries energy to smaller scales adding to the available pressure. Thus in general turbulence regulates the star formation process in the ISM (Mac Low and Klessen, 2004).

Dutta et al. (2008) estimated the H I power spectra of the external spiral galaxy NGC 628, where they find that the power spectra follow a power law with a slope of -1.6 over a length scale ranging from 800 pc to 8 kpc indicating a possible large scale energy input in the turbulence cascade. This was followed by measurements of H I intensity fluctuation power spectra of the nearly face-on galaxy NGC 1058 (Dutta et al., 2009) and the harassed galaxy NGC 4254 (Dutta et al., 2010) in the Virgo cluster. In a more comprehensive study, Dutta and Bharadwaj (2013); Dutta et al. (2013) observed collectively for 18 spiral galaxies in their sample the intensity fluctuation power spectra follow power law over a range of 300 pc to 16 kpc. They find that half of the galaxies in their sample follow a power law with a slope between -1.5 and -1.9.

Most of the studies listed above estimate the H I specific intensity fluctuation power spectra, which is directly related to the H I column density (Dutta and Bharadwaj, 2013). The range of length scales over which the density fluctuation power assumes a power law provides us with the range of turbulence cascade, the injection scale of energy and to some extent the nature of turbulence forcing (Federrath et al., 2009). However, to infer the momentum and energy input and the rate of energy transfer by the turbulence cascade the velocity statistics of the turbulence are essential. Spectral line observations, like the H I , let us estimate the line of sight velocity statistics. In this regard line of sight velocity dispersions are investigated in literature (Ianjamasimanana et al., 2015; Tamburro et al., 2009). However, the two-point statistics of the velocity fluctuations carries more information about the turbulent dynamics. Several estimators of the two-point statistics of the turbulent velocity fluctuations are devised (Esquivel and Lazarian, 2005; Lazarian and Pogosyan, 2000, 2006; Padoan et al., 2001) in literature and are successfully used for estimating the velocity power spectra of the Galaxy and the nearby dwarf galaxies (Nestingen-Palm et al., 2017; Zhang et al., 2012). These studies find evidence of large scale driving for the ISM turbulence in SMC and LMC (Chepurnov et al., 2015; Stanimirović and

Lazarian, 2001; Szotkowski et al., 2019) with injection scale as high as 2 kpc generating supersonic turbulence. Dutta (2015) discussed the limitations of most of these techniques to estimate the turbulence velocity power spectra for the external spiral galaxies.

2.4 Efficacy of image and baseline based estimators for radio interferometric observations[1]

Basic observables in interferometers are visibilities. As discussed earlier, various algorithms are used to reconstruct the sky brightness distribution. Here we do a comparative study of image and visibility based estimators for large scale distribution of column density and its power spectrum. We use simulated observations of the model sky to access the efficacy of different estimators of the above two quantities.

2.4.1 Simulating model visibility dataset

We proceed by simulating an observation with a known sky model based on the observed H I emission from the nearby external spiral galaxies (similar models are used in Dutta et al. (2009)). The specific intensity distribution from such a galaxy is given by

$$I(\vec{\theta}) = W(\vec{\theta}) \left[\bar{I} + \delta I(\vec{\theta}) \right], \tag{2.13}$$

where $W(\vec{\theta})$ quantifies the large-scale distribution of the H I column density in the sky and is normalized as $\int W(\vec{\theta}) \mathrm{d}\vec{\theta} = 1$. We call this the window function. The quantity $\bar{I}$ is the total intensity coming from the entire galaxy and $\delta I(\vec{\theta})$ corresponds to zero mean random fluctuations in the specific intensity. In the case of the ISM of spiral galaxies, such fluctuations arise as a result of compressible fluid turbulence therein. The ensemble average

---

[1] Work presented in this section is published in Dutta and Nandakumar (2019)

of the fluctuations $\delta I(\vec{\theta})$ is zero. Hence, at any given $\vec{\theta}_0$ in the galaxy, the ensemble average of the term $W(\vec{\theta})\delta I(\vec{\theta})$ vanishes. Here $W(\vec{\theta}_0)$ is the value at the window function at that $\vec{\theta}_0$. The window function falls to zero at large $\vec{\theta}$, which tapers the random fluctuations $\delta I(\vec{\theta})$ at large $\vec{\theta}$ eventually ensuring zero specific intensity arising from the sky beyond the galactic extent.

Modeling window function

The H I profile of a spiral galaxy is dominated by the radial variation in H I column density. However, azimuthal variations, like spiral arms, and rings are also seen. We use the shapelet decomposition of the specific intensity to model its large-scale structure which is given by $W(\vec{\theta})\bar{I}$. Shapelets are defined as a set of localized basis functions with different shapes (Refregier, 2003), we use Gaussian weighted Hermite polynomials in polar coordinates here. In terms of these shapelet basis $S_{nm}(\vec{\theta},\beta)$ and the shapelet coefficients f_{nm}, the specific intensity can be written as

$$I(\vec{\theta}) = \sum_{n=0}^{\infty}\sum_{m=-n}^{n} f_{nm}S_{nm}(\vec{\theta},\beta). \tag{2.14}$$

Here β is called the shapelet scale. Different orders n of shapelet coefficients represent different scales of the specific intensity with higher orders representing the smaller scale variations in it. We can represent the large scale component of the specific intensity by restricting the order n to smaller values. We define the window function as

$$W(\vec{\theta}) = \frac{\sum_{n=0}^{N}\sum_{m=-n}^{n} f_{nm}S_{nm}(\vec{\theta},\beta)}{\int \mathrm{d}\vec{\theta}\sum_{n=0}^{N}\sum_{m=-n}^{n} f_{nm}S_{nm}(\vec{\theta},\beta)\Big/\int \mathrm{d}\vec{\theta}}. \tag{2.15}$$

Here the $\vec{\theta}$ integrals are taken over the entire galactic disc. Consideration of choosing the parameters N and β will be discussed shortly.

Modeling $\bar{I}$ and $\delta I(\vec{\theta})$

The power spectrum of the specific intensity fluctuations $\delta I(\vec{\theta})$ is defined as

$$\left\langle \delta \tilde{I}^*(\vec{U})\delta \tilde{I}(\vec{U}') \right\rangle = \delta_D(\vec{U} - \vec{U}')P(U), \tag{2.16}$$

where $U = \mid \vec{U} \mid$. The angular brackets denote ensemble averages. Observations in our Galaxy (Crovisier and Dickey, 1983; Green, 1993) and external dwarf and spiral galaxies (Begum et al., 2006; Dutta and Bharadwaj, 2013; Dutta et al., 2009, 2013; Elmegreen et al., 2001) suggest that the H I specific intensity fluctuations can be modelled as a Gaussian random distribution having a power law power spectrum, i.e, $P(U) \propto U^{\alpha}$. We define the ratio of the total intensity $\bar{I}$ to the standard deviation of these fluctuations $\sigma_{\delta I}$ as $\mathcal{R}$, i.e,

$$\mathcal{R} = \frac{\bar{I}}{\sigma_{\delta I}}. \tag{2.17}$$

We use the parameters α and $\mathcal{R}$ to simulate zero-mean Gaussian random numbers with a given power law power spectrum to represent $\delta I(\vec{\theta})$.

Modeling the specific intensity distribution $I(\vec{\theta})$

We model the window function based on the large-scale structure of the face-on spiral galaxy NGC 628. We decompose the column density or moment zero map (natural weighted) of NGC 628 taken from THINGS (Walter et al., 2008) survey data product [2] in terms of its shapelet coefficients and use the first few shapelets to model the window function. We choose the largest shapelet order N and the shapelet scale β as follows. Considering a given value of β, we construct the zeroth-order shapelet ($N = 0$, Gaussian function) from the moment zero map of NGC 628 and estimate the mean square difference between the moment zero map and this basic shapelet. The lowest mean square difference

---

[2] THINGS: The H I Nearby Galaxy Survey data product: http://www.mpia.de/THINGS/Data.html

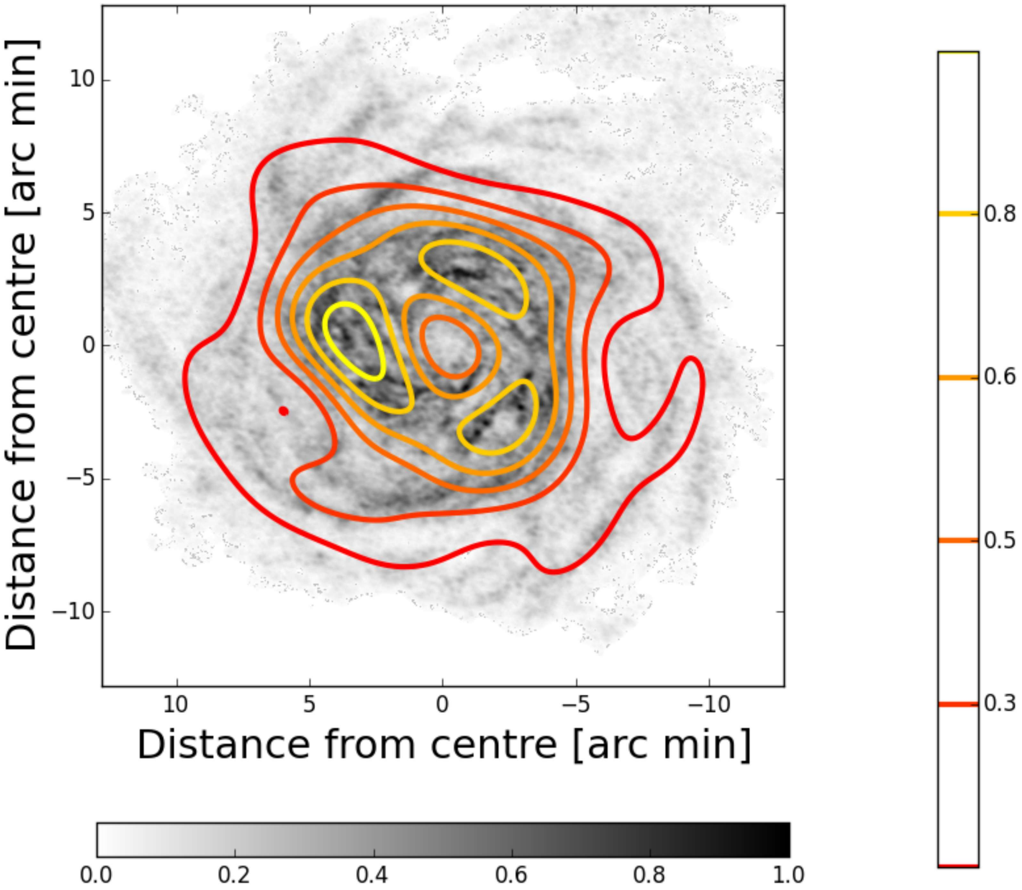

Fig. 2.1 Greyscale image showing the natural weighted moment zero map of the galaxy NGC 628. Contours represent the model window function based on this map (see eqn (2.15)). Note that the range of values in the pixels of these maps differ. We have scaled the pixel values to keep the maximum pixel value as unity for both the maps.

corresponds to $\beta = 240''$ for the galaxy NGC 628. Dutta et al. (2013) found that the intensity fluctuations in the galaxy NGC 628 are dominated by the window function at angular scales $< 240''$. We found that for $N <= 12$ the shapelet coefficients do not have significant structures at angular scales $< 240''$. Hence we use $\beta = 240''$ and $N = 12$ to construct the model window function. The Greyscale image in the figure 2.1 represents the moment zero map (natural weighted) of NGC 628 from the THINGS archive. We show the model window function based on this moment zero map as red contours in the same figure. The range of values in the pixels of the moment zero map and the window function differ. We have scaled the pixel values to keep the maximum pixel value as unity for both the maps to make the comparison easier. Figure 2.2 shows the power spectrum of a model sky brightness distribution $I(\vec{\theta})$ (eqn (2.13)) in blue dashed line. The values of the model parameters are $\mathcal{R} = 5,\ \alpha = -1.5$. The green dot-dashed line shows a power law of slope -1.5. Power spectra of only the first term $W(\vec{\theta})\bar{I}$, is shown with a black solid line. Clearly, for baselines lesser than $1\ k\lambda$, the power spectrum of the model $I(\vec{\theta})$ is dominated by the window function, while for baselines greater than $1\ k\lambda$, the power spectrum of the model image $I(\vec{\theta})$ follows a power law with $\alpha = -1.5$.

Dutta et al. (2013) has estimated the power spectra of 18 spiral galaxies from the THINGS sample using a visibility based estimator. They found that power spectra follow power laws with α ranging between -0.3 to -2.2. Moreover, 9 of the 18 galaxies in their sample have α in between -1.5 to -1.8. We choose three values of α for our model sky image : $[-0.5, -1.5, -2.0]$. Dutta and Bharadwaj (2013) found that $\mathcal{R}$ varies between 5 to 10 for the six galaxies they analyzed. We consider two values of $\mathcal{R}$ here: $[5, 10]$.

Using the above parameters we generate six model sky specific intensity distributions in a square grid of 1024^2 with each grid element representing a $1.5'' \times 1.5''$ patch in the sky.

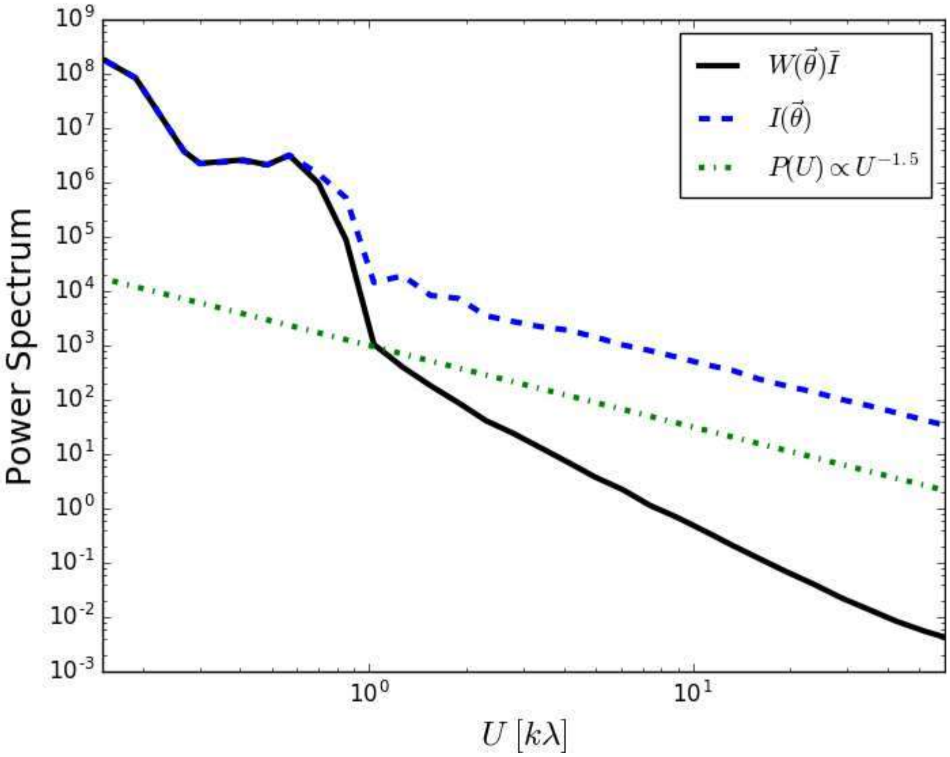

Fig. 2.2 Power spectrum for the first term $W(\vec{\theta})\bar{I}$ in eqn. (2.13) (solid black line) is compared with the power spectrum of $I(\vec{\theta})$ for the model image (blue dashed line) with $\mathcal{R} = 5$ and $\alpha = -1.5$. Observe that for $U > 1\ k\lambda$ there is a significant deviation between the two. The green dot-dash line corresponds to a power law with index -1.5. Clearly, the power spectra of $I(\vec{\theta})$ follow a power law for $U > 1\ k\lambda$ and has the slope of -1.5

.

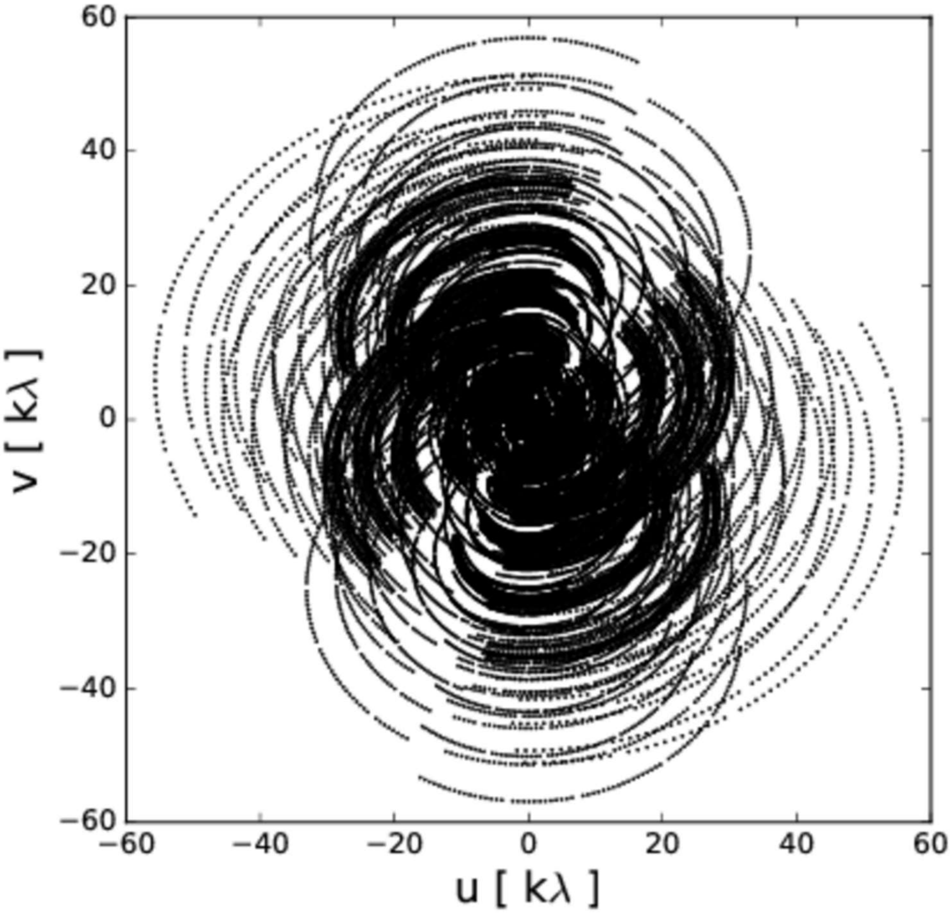

Fig. 2.3 Figure showing the sampling function for the simulated observation presented here. Black points are the places in the baseline plane where the visibilities are measured.

2.4.2 Simulated visibility data

To simulate radio interferometric observations and generate random group visibilities from the above sky model, we need to choose a particular array configuration of the interferometer. We model our telescope based on the GMRT array configuration. We scale the antenna coordinates to half their original values. This decreases the largest baseline available for the array to $60\ k\lambda$ (instead of $\sim\ 120\ k\lambda$ for the original GMRT array configuration) at 21 cm and hence also reduces the effective resolution of the array. Note

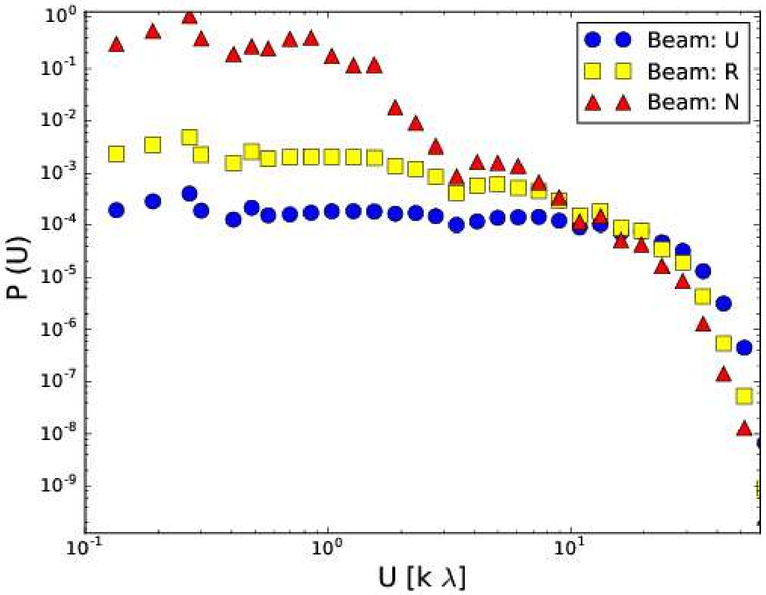

Fig. 2.4 Figure showing the power spectrum of the dirty beam for the uniform, robust and natural weighted beams for the sampling function given in Figure 2.3.

that, this compromise is made to increase the computational speed and does not limit our analysis of assessing the efficacy of different estimators. We choose the declination of the source to be $+54°$, which produces a fairly good uv-coverage. For each of the six model-specific intensity distributions, we perform eight hours equivalent of simulated observation. Figure 2.3 shows the sampling function corresponding to this simulated observation. We also show the power spectrum of the dirty beam corresponding to the three different weighting schemes in figure 2.4. The power spectra of the three beams are significantly different. We use discrete Fourier transform to calculate the complex visibility values at the sampled baseline positions from the model sky images. In principle, the measurement noise at each baseline as well as its effect in the one and two point statistics can be arbitrarily decreased by increasing the total observation time. Hence, we do not include any measurement noise in our model observations for simplicity.

2.4.3 One Point statistics: large-scale distribution of specific intensity

The large-scale distribution of H I carries important information about its interplay with star formation in spiral and dwarf galaxies. Most spiral galaxies show visible depression of H I near the central part owing to the large star formation rate (Wang et al., 2014). Bagetakos et al. (2011) have observed the presence of H I holes at small knots of star formation in the discs of the spiral galaxies. Dwarf galaxies like GR 8 have a clear spatial correlation between star formation rate and H I column density (Begum and Chengalur, 2003). These studies require evaluating the locally averaged H I intensity distribution from the observed visibilities. An estimate of the window function can be achieved by performing local average of the specific intensity as $\langle I(\vec{\theta})\rangle = \bar{I}W(\vec{\theta})$.

We reconstruct the specific intensity distribution corresponding to the observed visibilities using CLEAN as discussed in the previous section. In this work, we explore various user-defined parameters with CLEAN to assess their effect on the reconstructed image. These will be discussed in detail in the next section. For each reconstructed image, we evaluate the window function using its shapelet coefficients. Following the same arguments as discussed in section 3.3, we use the shapelet scale to be $240''$ and the first 12 shapelet coefficients to represent the window function. To distinguish the window function estimated from the reconstructed image from the model window function, we mention the former by $W_C(\vec{\theta})$ for the rest of the analysis. Further, for the model ($W(\vec{\theta})$) as well as the reconstructed windows ($W_C(\vec{\theta})$), we estimate the azimuthally averaged window function defined as

$$W_A(\theta) = \frac{1}{2\pi}\int\limits_0^{2\pi} W(\theta, \phi)\mathrm{d}\phi, \tag{2.18}$$

in different bins of θ. Here (θ, ϕ) are the polar components of the vector $\vec{\theta}$. We use the standard deviation of the values of the reconstructed window $W_C(\vec{\theta})$ in each azimuthal bin to represent the statistical fluctuations associated with the estimated value of $W_A(\theta)$ in the corresponding bin. This gives us a robust way of comparing the estimates of the window function with that of the model.

2.4.4 Power Spectum estimators

Power spectrum estimators from the interferometric data can be categorized into two classes the *image based estimator* and the *visibility based estimator*.

Image based power spectrum : Here we use the reconstructed images from the deconvolution of measured visibilities by radio interferometers for calculating the power spectrum. To distinguish the reconstructed image from the sky brightness distribution $I(\vec{\theta})$, we shall denote the former by $I_C(\vec{\theta})$. Since the image is already evaluated at regular grids in $\vec{\theta}$, a two dimensional FFT can be used to estimate the Fourier transform of $I_C(\vec{\theta})$. As the interferometers are mostly not sensitive at baselines lower than a certain value, they effectively do not measure the first term in the eqn (2.13) and we may write

$$\tilde{I}_C(\vec{U}) = \tilde{W}(\vec{U}) \otimes \delta\tilde{I}(\vec{U}) + \mathcal{B}_I(\vec{U}), \tag{2.19}$$

where $\tilde{W}(\vec{U})$ represents the Fourier transform of the window function and $\otimes$ denotes the convolution. The quantity $\mathcal{B}_I(\vec{U})$ jointly represents any artifacts introduced in the image reconstruction procedure and effective noise in $\tilde{I}_C(\vec{\theta})$ resulting from the measurement noise. We correlate $\tilde{I}_C(\vec{U})$ at each baseline, which gives

$$P_C(\vec{U}) = \langle \tilde{I}_C(\vec{U})\tilde{I}_C^*(\vec{U})\rangle = \mid \tilde{W}(\vec{U}) \mid^2 \otimes P(\vec{U}) + P_{\mathcal{B}}(\vec{U}), \tag{2.20}$$

where $P_{\mathcal{B}}(\vec{U})$ is related to both $\mathcal{B}_I(\vec{U})$ and $\tilde{I}(\vec{U})$. The angular brackets above denotes the ensemble average of many realizations of the sky. In practice, we assume statistical isotropy and choose azimuthal bins to perform this average. Hence the image based azimuthally averaged power spectrum estimator is given as

$$P_I(U) = \frac{1}{2\pi} \int\limits_0^{2\pi} \tilde{I}_C(U,\phi)\tilde{I}_C^*(U,\phi)\mathrm{d}\phi, \tag{2.21}$$

where (U, ϕ) are the polar components of the vector $\vec{U}$. As discussed before (see figure 2.2), the window function represents large-scale variation of the specific intensity and hence at baselines $U >> 1/\theta_0$, where θ_0 is the extent of the galaxy, the window function can be treated as a delta function $\delta_D(\vec{U})$. Hence in the absence of $P_{\mathcal{B}}$, at baselines $U >> 1/\theta_0$, the quantity $P_C(\vec{U}) \sim P(\vec{U})$ and $P_I(U)$ gives an estimate of the power spectrum of the sky brightness fluctuations. The quantity $P_{\mathcal{B}}(\vec{U})$ is a manifestation of the incomplete baseline coverage and the different techniques incorporated in the image reconstruction process. Arguably it depends on the user chosen parameters in the CLEAN (or other algorithms) and hence needs to be evaluated and subtracted from the above equation to estimate the power spectrum in an unbiased way. Unfortunately, a separate estimation of $P_{\mathcal{B}}$ is almost always impossible. Interestingly, such a bias is grossly ignored in literature where image based estimators are used (Walker et al., 2014; Zhang et al., 2012). Further, incomplete baseline coverage also makes the measurement noise correlated in $\tilde{I}_C(\vec{\theta})$ and introduces a non zero $P_{\mathcal{B}}(\vec{U})$ (see Dutta (2011) for detail). However, as we have not considered the measurement noise in our simulation, we refrain from investigating this effect here. Estimating the errors in the reconstructed image is not straightforward and only a Monte-Carlo based technique can be effectively used (Sault and Conway, 1999). Following that, the errors in the image based estimates of the power spectrum are also non-trivial. We use the variation of the power spectrum values in different $\vec{U}$ inside an annular bin to represent the error in the image based power spectrum estimator. Additionally, at smaller baselines independent estimates of the power spectrum reduce and the sample variance dominates. The sample variance is given by $P_I/\sqrt{N_g}$, where N_g is the number of independent estimates of P_I in a given annular bin. We also add this in quadrature to represent the error in the image based estimator of the power spectrum.

Visibility based power spectrum estimator: Since the Fourier transform of the first term in eqn. (2.13) is not measured mostly by an interferometer (see the discussion above), the measured visibilities can be written as

$$V(\vec{U}) = \left[\tilde{W}(\vec{U}) \otimes \delta\tilde{I}(\vec{U})\right] S(\vec{U}) + \mathcal{N}(\vec{U}). \tag{2.22}$$

The visibility correlation gives

$$\langle V(\vec{U})V^{*}(\vec{U})\rangle = \mid \tilde{W}(\vec{U}) \mid^{2} \otimes P(U) \mid S(\vec{U}) \mid^{2} + \mid \mathcal{N}(\vec{U}) \mid^{2}. \tag{2.23}$$

In absence of measurement noise, at baselines $U >> 1/\theta_0$, the visibility correlation gives $P(U) \mid S(\vec{U}) \mid^2$. The azimuthally averaged power spectrum then can be estimated as

$$P_V(U) = \int_0^{2\pi} \langle V(\vec{U})V^{*}(\vec{U})\rangle \mathrm{d}\phi \Bigg/ \int_0^{2\pi} \mid S(\vec{U}) \mid^2 \mathrm{d}\phi \tag{2.24}$$

In practice, it is estimated at discrete azimuthal bins. For most of the array configurations, the integral in the denominator of the above expression has a nonzero value. However, if the integral is zero in a particular bin the power spectrum is not evaluated at that bin. In realistic observations, the noise term $\mid \mathcal{N}(\vec{U}) \mid^2$ dominates and introduces a bias in power spectrum estimates. Dutta (2011) discusses the procedure to take care of this noise bias in detail. Since we do not have measurement noise in our simulation, we neglect this effect here. We estimate the errors in the visibility based power spectrum estimator following the calculations by Dutta (2011).

2.4.5 Analysis and Results

We use the task IMAGR in AIPS to reconstruct the sky brightness distribution for each of the six simulation sets. We discuss our analysis and results based on the simulation with parameters $\mathcal{R} = 5.0, \alpha = -1.5$ in detail and tabulate the results for all the models.

In choosing different parameters for CLEAN in the task IMAGR, we give particular emphasis on the fact that here we are interested in reconstructing the sky brightness distribution for diffuse emission. It is a common understanding that a smaller loop gain improves the reconstruction of extended sources (Thompson et al., 2017), however improvement for a gain < 0.01 is minimal. We choose a loop gain of 0.005. To tame the effect of abrupt fall in the baseline coverage at a large baseline we use a Gaussian taper with the UVTAPER parameter set to 45 kλ in IMAGR (for both u and v). This corresponds to a tapering function $T(U) = \exp\left[-U^2/2(30)^2\right]$, where U is measured in kλ. Tapering down-weights the visibilities at the larger baselines and hence may affect the power spectrum estimates. However, for a power law power spectrum the effect of tapering can be analytically reversed by multiplying the power spectrum by $1/T(U)^2$. We choose the pixel size for the image to be $1.5'' \times 1.5''$ in a grid of 1024^2. We use three different weighting schemes to weight and grid the visibility data, namely the natural weighting, the uniform weighting and the robust weighting. These are controlled mainly by the parameters ROBUST in AIPS. We have chosen ROBUST values of $(-5, 0, 5)$ to produce three different reconstructions of sky brightness from each simulated visibility dataset. For each IMAGR run, we manually stop the major cycles when the maximum and minimum pixels in the residual image are of similar value. The restoring beams for the uniform, robust and natural weighted images came out to be $4.1'' \times 3.9''$, $6.7'' \times 5.8''$ and $9.3'' \times 8.2''$ with the beam position angles $\sim 63^\circ, \sim 68^\circ$ and $\sim 70^\circ$ respectively. These images are used for further analysis.

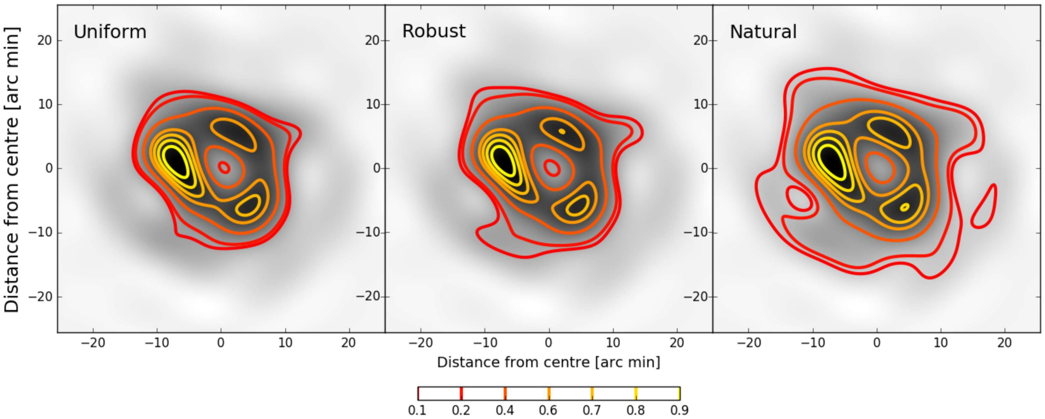

Fig. 2.5 Window functions of the reconstructed images using different weighting schemes are shown in contours against the grey scale plot of the window function of the model image corresponding to $\mathcal{R} = 5,\ \alpha = -1.5$. The scales of the contours are relative here.

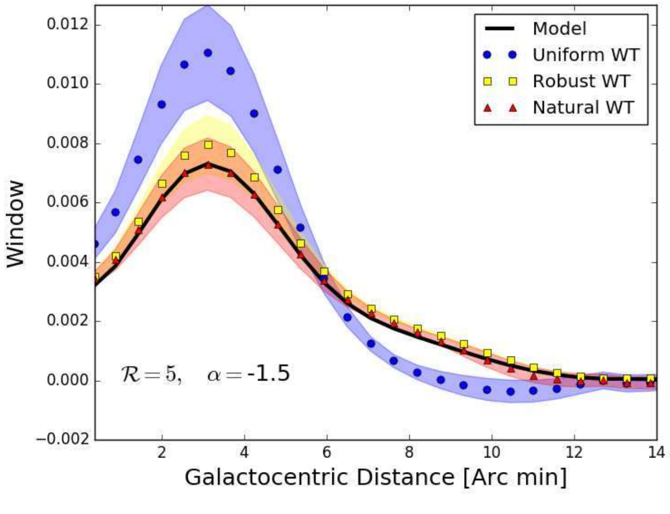

Fig. 2.6 This figure compares the azimuthally averaged window functions estimated using different weighting schemes with that of the input image. The colour bands for each estimate represents the error.

Window function

For each of the reconstructed image, we estimate the window function W_C and the azimuthally averaged profile W_A following the prescriptions given in the previous section. Contours in three panels of figure 2.5 show W_C corresponds to the uniform, robust and natural weighting schemes respectively. The Greyscale image in each panel corresponds to the model window function. A visual comparison of the contours of these three panels with those in figure 2.1 demonstrates that the natural weighted scheme best reproduces the model window. We shall quantify this statement shortly. Note that both the model window function W and the quantity W_C are estimated at Cartesian grids in $\vec{\theta}$, we may denote the values in the grid points as $W[i,j]$ and $W_C[i,j]$ respectively. We use the quantity

$$\chi = \frac{\sum\limits_{i,j} (W[i,j] - W_C[i,j])^2}{N_G^2 \sqrt{\sigma_W^2 + \sigma_C^2}} \tag{2.25}$$

to represent the deviation of W_C from W. Here σ_W and σ_C correspond to the standard deviation of the pixel values of $W[i,j]$ and $W_C[i,j]$ respectively and N_G gives the number of grid points along one axis[3]. The quantity χ gives the mean square deviation between the model and the estimated window functions. A lower value of χ corresponds to a better reconstruction of the window function. We have performed tests to check if the figure of merit, χ, captures the deviation between the estimates of the window function. In this test, we choose a fiducial model for the window function, values of $\mathcal{R}$ and α and generate different realizations of δI to simulate models of the sky brightness distribution. We then estimate the window function from each of these images and compare them with the fiducial window function using the figure of merit χ. If the window function estimated from any of these models exactly matches the fiducial model of the window, then the value of χ is zero. We found the χ value lies in between 0.004 to 0.01 within the six models

---

[3] The quantity χ used here has no probabilistic interpretation and should not be confused as a function used in the most maximum likelihood estimations

discussed here. These numbers can be considered as references while interpreting the result.

For the model with $\mathcal{R} = 5.0,\ \alpha = -1.5$ we found that the uniform weighting produces a window function with the largest value of $\chi = 2.63$, whereas the corresponding values of χ for robust weighting and natural weighting are 0.26 and 0.15 respectively. The window function estimated from the images produced with uniform and robust weighting has a bias, however, the natural weighting scheme produces the best estimate of the window function.

Figure 2.6 plots the azimuthally averaged window functions estimated from the model as well as the three reconstructed images. The black solid line corresponds to the azimuthally averaged window function of the model image. Colour bands around each of the estimates correspond to the variation of the window function within the respective azimuthal bins. The uniform weighting (circles) produces the largest deviation from the model window and also has the largest variation in each azimuthal bin as represented by the error band in the figure. Note that both the robust (square) and the natural weighting (triangle) reproduce the azimuthally averaged window within the errors, however, the points corresponding to the robust weighting (square) are systematically offset from the model window function.

Figure 2.7 shows the azimuthally averaged window function estimated for all the six models using natural weighting schemes (red triangles) against the same estimated from the model images (black solid line). Table 2.1 shows the values of χ for all the six model skies. A lower value of χ represents a better reproduction of the window function. Clearly, in all cases, the natural weighting gives the best reconstruction of the window function. It is observed that the value of χ in the case of uniform and robust weighting is considerably larger than the corresponding example of natural weighting. It is well known that the Natural weighting produces a wider point spread function. Since the window function

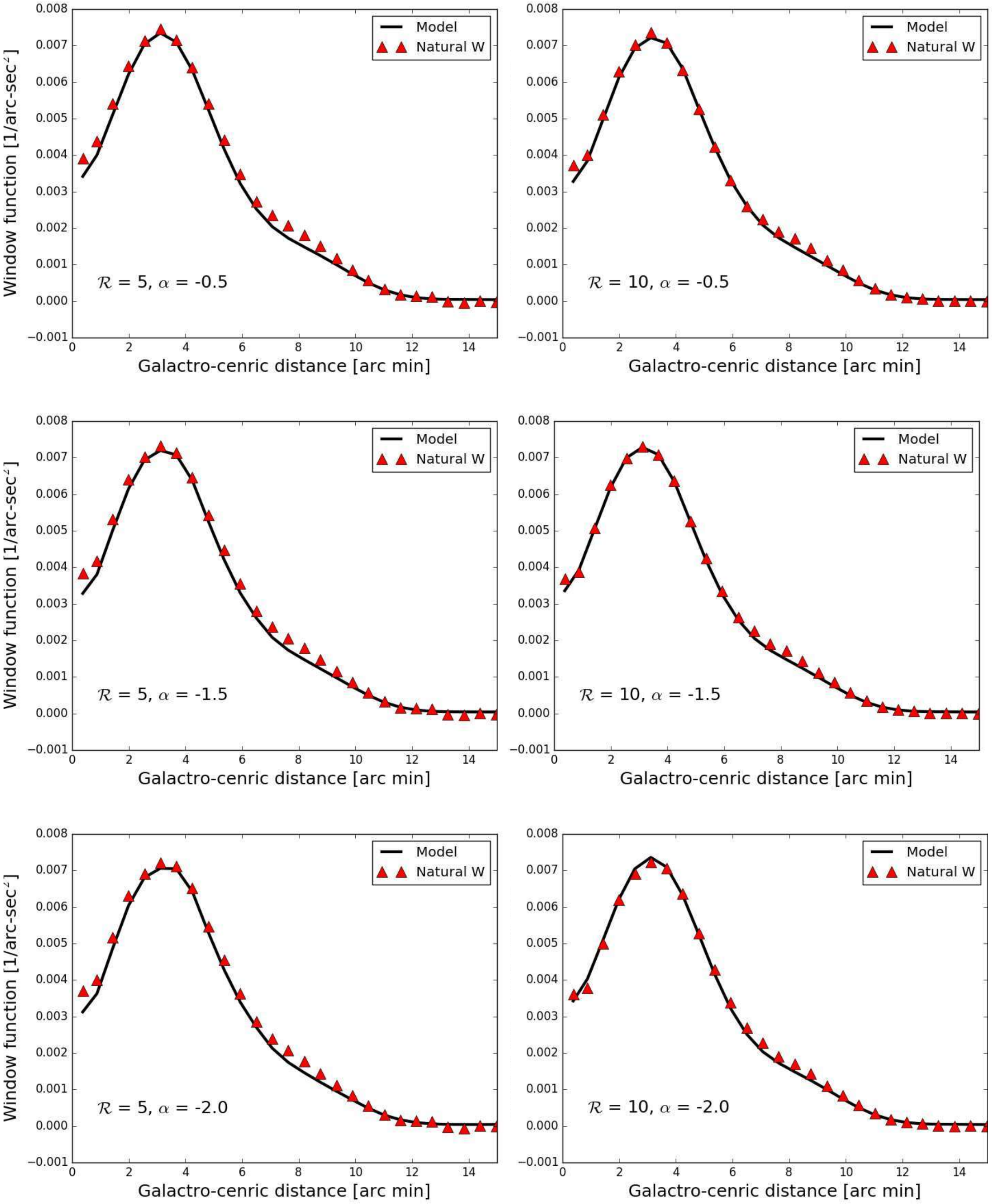

Fig. 2.7 Comparison of the azimuthally averaged window functions estimated from the input model (black solid line) and the reconstructed natural weighted image (red triangles) are shown for all six simulations.

		$\mathcal{R} = 5$			$\mathcal{R} = 10$		
α		U	R	N	U	R	N
-0.5	χ	2.77	0.42	0.02	1.36	0.1	0.004
	α_I	-1.4 ± 0.1	-1.9 ± 0.3	-2.1 ± 0.1	-1.3 ± 0.1	-1.4 ± 0.1	-1.7 ± 0.3
	α_V	-0.44 ± 0.07			-0.48 ± 0.08		
-1.5	χ	2.63	0.26	0.15	1.43	0.16	0.003
	α_I	-1.8 ± 0.1	-2.0 ± 0.1	-2.2 ± 0.1	-1.9 ± 0.1	-2.2 ± 0.2	-1.9 ± 0.3
	α_V	-1.5 ± 0.1			-1.4 ± 0.2		
-2.0	χ	3.76	0.22	0.02	1.16	0.08	0.004
	α_I	-2.1 ± 0.1	-2.3 ± 0.1	-2.3 ± 0.1	-2.1 ± 0.1	-2.4 ± 0.2	-2.1 ± 0.3
	α_V	-1.9 ± 0.1			-1.9 ± 0.2		

Table 2.1 Table summarising the result of comparison between different estimates of the window function and power spectrum. Efficacy to reproduce the model window function from an estimated image is quantified by χ (see eqn (2.25)). We tabulate the values of α_V and α_I for each model to access the merit of different estimators in the table. The headers 'U', 'R' and 'N' correspond to the uniform, robust and natural weighting schemes respectively.

represents the large scale distribution, the natural weighting is expected to reproduce it better. Note that the actual values of the χ do not carry a robust probabilistic interpretation here. Moreover, for the models with a relatively lower amplitude of the fluctuations in specific intensity ($\mathcal{R} = 10$), the χ values are systematically lower than the models with $\mathcal{R} = 5$. We conclude that with a careful choice of the imaging parameters for CLEAN it is possible to estimate the window function unbiasedly from the reconstructed image and it is best estimated when natural weighting is used.

Power Spectra

We estimate power spectra from the model images using the image based estimator. Note that, the model images do not have any artefacts that may arise from the reconstruction and hence this power spectrum can be considered as a reference. This is shown with a solid black line in figure 2.8. The power spectrum $P_V(U)$ is shown with grey pentagons in the same figure with the grey area indicating the error bars. It is quite clear that

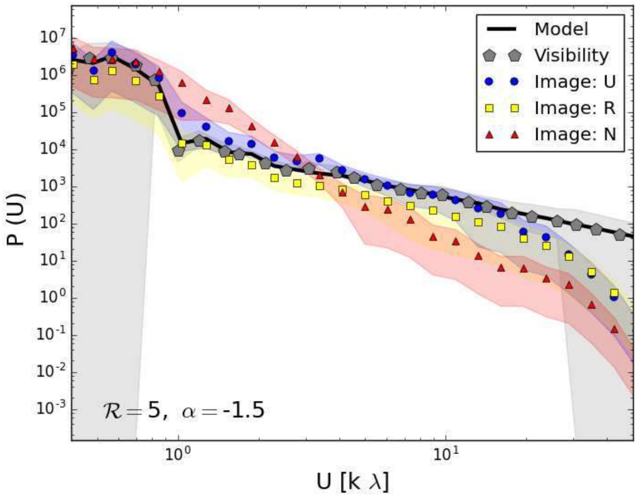

Fig. 2.8 This figure compares the power spectrum estimated using the visibility based estimator and the image based estimator for all three different weighting schemes against the model power spectrum. The shaded area corresponds to the error associated with each estimate.

the visibility based power spectrum follows the reference spectra quite well and overall bias is minimized. The large error bars in small and larger baselines are indicative of less independent measurements at those baselines. The visibility based power spectrum estimator assumes power law at baselines larger than $1\ k\lambda$.

We use the image based power spectrum estimator for all the three reconstructed images from the three weighting schemes. We correct each of these spectra for the effect of tapering by multiplying them with $1/T(U)^2$. We plot in the same figure with circular, square and triangular markers representing the uniform, robust and natural weightings respectively with the corresponding error bands. At longer baselines, the power falls drastically. This is an effect of the convolution of the CLEAN components with the restoring beam which produces correlation at the pixels at a scale smaller than the beam scale. The image based estimate with the natural weighting scheme is drastically different from the model. The power spectrum estimated using the reconstructed image with the

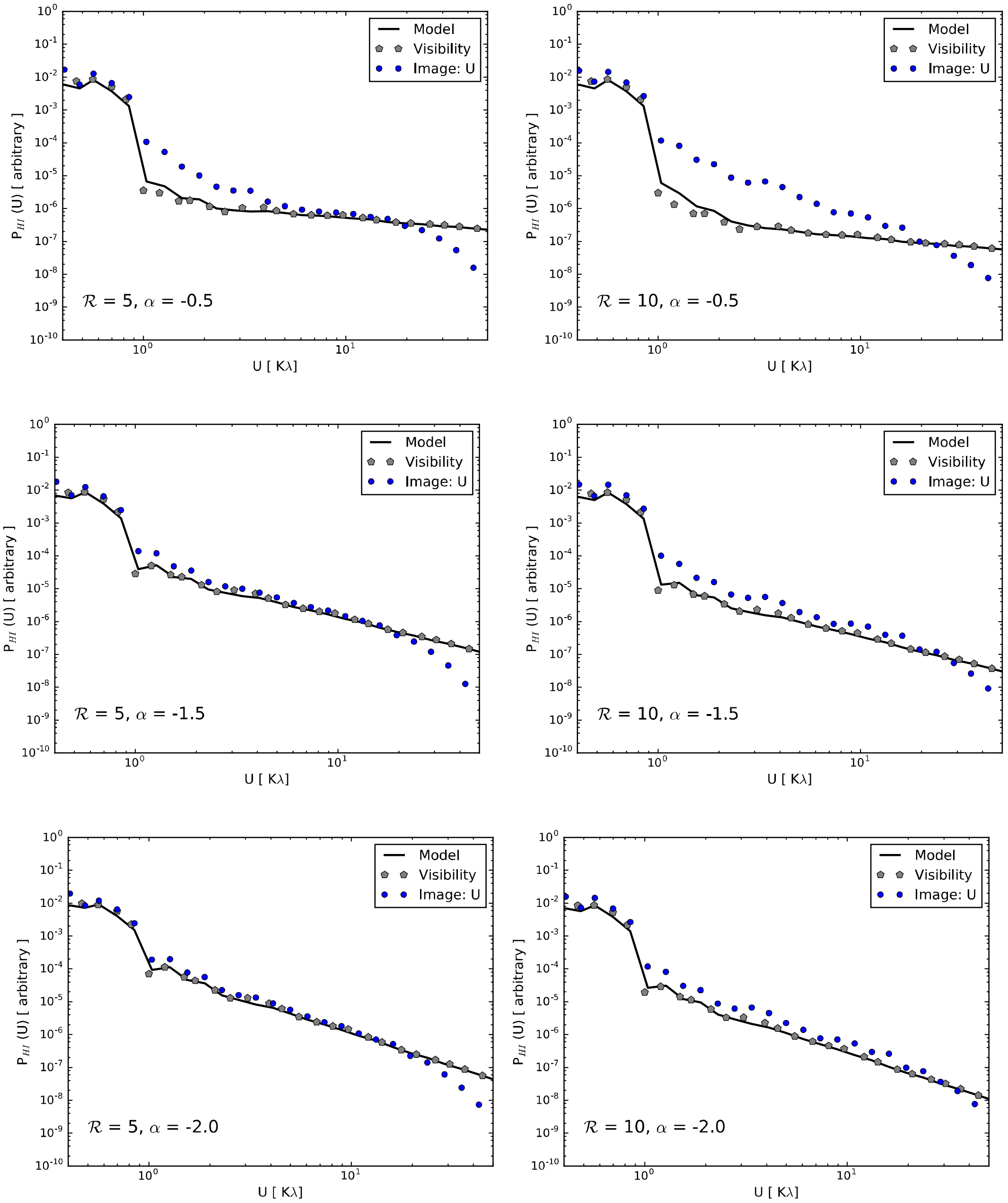

Fig. 2.9 Comparison of the power spectrum estimated from the input model (black solid line), using the visibility based estimator (grey pentagons) and image based estimators from the uniform weighted image (blue circles) are shown for all six simulations. The amplitudes of the y-axes for different panels are scaled arbitrarily to keep the power spectra in the same range in all plots.

uniform weighting scheme almost follow the model power spectra within the error bars in the baseline range of $1-20\ k\lambda$ and that with the robust weighting scheme is slightly different from the model. As discussed before, we expect the power spectrum to be a power law. To assess how good are the image based estimates of the power spectra, we fit (chi-square method) a power law function to these spectra between the baseline range $1-20\ k\lambda$ and find the best fit value of the power law slope with error bars. We find the power law index estimated from the image based power spectra vary as -1.8 ± 0.1, -2.0 ± 0.1 and -2.2 ± 0.1 for the uniform, robust and natural weighted images respectively. The best fit power law index for the visibility based estimate of the power spectra in the same baseline range is -1.5 ± 0.1. These numbers suggest that more or less all three image based estimators deviate from the power spectrum of the model sky. However, we must note that the uniform weighting scheme preserves the power spectrum of the model with the least bias among the image based estimators.

Figure 2.9 shows the power spectrum estimated from the uniform weighting schemes for all six of our models (blue circles) and from the visibility based estimators (grey pentagons) against the model power spectrum (black solid line). Clearly, the visibility based power spectrum reproduces the model power spectrum almost exactly, whereas the image based estimate of the power spectrum is biased. The bias is visually more prominent for larger values of α. Table 1 gives the values of the estimated α using visibility and image based estimators for all the different sky models and different weighting. To compare the result from all the six simulations we plot the different estimates of power law slope (α_I) from the image based estimator against those estimated using the visibility based estimator (α_V) in figure 2.10. Three panels in this image correspond to three different weighting schemes. Representations of different markers are given in the left panel. Clearly, for all models, the uniform weighting scheme performs the best. We also notice that the α_Is are

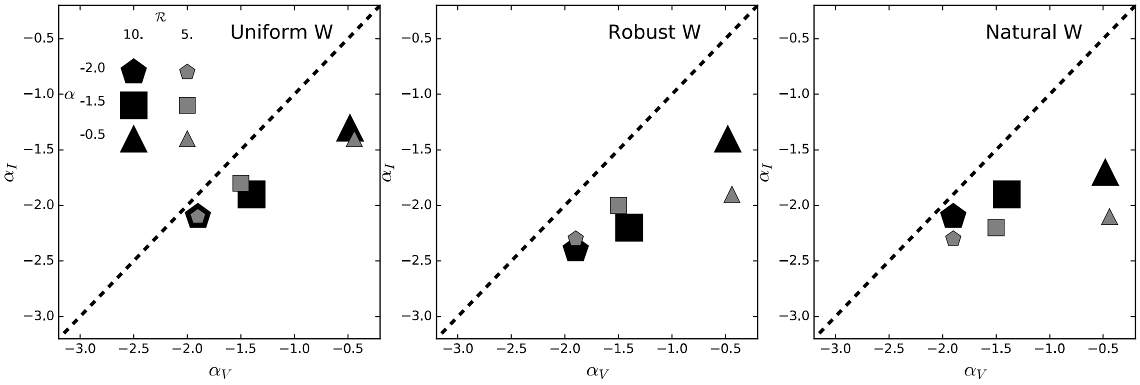

Fig. 2.10 The image and visibility based estimates of α are compared for different simulations and weighting schemes. In each panel, we plot the image based estimate α_I against the visibility based estimate α_V for all six simulations. The dashed line corresponds to an exact match. The meaning of different markers is given in the left most panel. Three panels in this figure correspond to three different weighting schemes, with a left to right giving uniform, robust and natural weightings respectively.

systematically smaller than the corresponding α_Vs with shallower power spectra having a systematically larger bias.

Apart from the baseline coverage of the interferometer, the efficacy of image reconstruction may also depend on the structure of the sky brightness distribution itself. In the CLEAN algorithm, the sky is modelled as a collection of point sources. If the observed sky is a set of isolated unresolved sources, then the visibility function is smooth across baselines. In such cases, CLEAN is supposed to give an unbiased estimate of the sky. On the other hand, for diffuse emission, the visibility function is expected to be patchy. Observations with inadequate baseline coverage will lack the full information to model the sky. To test, how much of the sky needs to be filled by sources to see the effect of the baseline coverage, we model the sky with a collection of point sources uniformly distributed in the field of view. The amplitude of these sources are varied randomly within a decade of flux density, the absolute flux scale is of no importance here. If we keep the number of point sources small, then they are expected to be isolated and the CLEAN must work well. On the other hand, if we increase the number of point sources, it would start

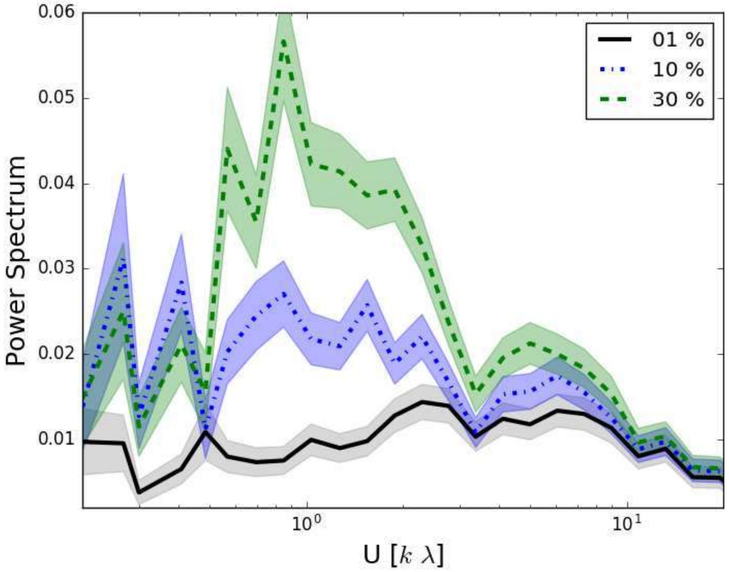

Fig. 2.11 The Figure shows the power spectra calculated from the residual images after the point source subtraction for the point sources influencing 1%, 10% and 30% of the field of view.

to simulate a diffuse emission and CLEAN may fail to reproduce an unbiased estimate of the sky. Considering a Gaussian PSF of the interferometer, we assume that each point source influences the nearby pixels within a circle of diameter equal to 2.5 times the full-width at half maxima of the PSF. How much of the field of view is covered by the point sources in this way gives a measure of how diffuse the emission is as seen by the telescope. We generate three model sky intensity distributions with 1%, 10% and 30% of the sky influenced by the point sources (using robust weighting with a ROBUST=0). Using these model images we simulate the visibilities and reconstruct images keeping the same baseline coverage as our previous simulations. We subtract the CLEAN component model of the point sources from the reconstructed image to get the residual maps. Power spectra of the residual maps for the three cases are shown in figure 2.11. The shaded region in cases show the corresponding errors. A flat power spectrum is expected if the residual image does not have any correlated noise. We see a significant systematic increase in the amplitude of the residual power spectrum with the number of sources. This demonstrates the limitation of image reconstruction from interferometric data with incomplete uv coverage in reproducing the structures of the diffuse sky.

In this work, we simulate H I observation of external spiral galaxies to test the efficacy of different estimators that are used to measure the statistical properties of the sky brightness distribution from the radio interferometers. In particular, we have investigated how well we can reconstruct the large-scale structure of the brightness given by the window function and the scale dependence of the structures given by the power spectrum of the intensity fluctuations. On the other hand, one can either use the visibilities directly to estimate the power spectrum, or first estimate the brightness distribution from the visibilities and then use those to estimate the power spectrum. Reconstruction of the brightness distribution is based on several algorithms. We find that with the Cotton-Schwab version of the CLEAN algorithm using natural weighting scheme, the window function is reasonably reproduced. The visibility based estimator of the power spectrum reproduces the model power spectrum unbiasedly. Amongst the image based estimators, the reconstructed image with uniform weighting scheme performs best, however, a general scale dependent bias is observed for all the image based estimators. As the radio astronomy community plans for larger interferometers, the visibility data volume is expected to grow large. One way of reducing the problem with large data volume is to perform online reconstruction of the images with the instantaneously available visibilities. Our investigation highlights the problems that may arise in a proper reconstruction of the sky statistics and emphasizes the need for recording the visibilities directly.

2.5 Visibility moment estimators for column density and line of sight velocity

The visibility moment estimator (VME) for column density and turbulence velocity power spectrum is introduced in Dutta (2016) and a variant in Nandakumar and Dutta (2020). The angular extent of the galaxy is rather small and we consider a flat sky approximation here. The number density of neutral hydrogen $n_{\rm HI}$ in eq (2.2) can be written as

$$n_{\rm HI}(\vec{\theta}, z) = n_0\, G(\vec{\theta}, z)\, \left[1 + \delta n(\vec{\theta}, z)\right], \tag{2.26}$$

where n_0 is the average H I number density over the entire galaxy. Here we consider the centre of the galaxy and the tangent plane coinciding. In general, the galactic disc can have an arbitrary orientation with respect to the tangent plane we have considered here. The galaxy H I profile function $G(\vec{\theta}, z)$ times n_0 gives the large scale variation of the H I number density in the coordinate system discussed here, while $\int {\rm d}\vec{\theta} \int {\rm d}z G(\vec{\theta}, z) = 1$. Note that for any galaxy not oriented exactly face on, a circular disc will be seen in projection in the tangent plane. The term δn gives the random variation in the H I number density. For a nearly face-on thin disc (disc thickness l), the specific intensity of H I emission can be written as

$$I(\vec{\theta}, \nu) = I_0\, G(\vec{\theta})\, \left[1 + \delta n(\vec{\theta})\right] \phi(\vec{\theta}, \nu). \tag{2.27}$$

Here $I_0 = \frac{3h\nu_0 A_{21} n_0 l}{16\pi}$ gives the total H I flux from the galaxy, the quantities $G(\vec{\theta}) = G(\vec{\theta}, z = 0), \delta n(\vec{\theta}) = \delta n(\vec{\theta}, z = 0)$ and $\phi(\vec{\theta}, \nu) = \phi(\vec{\theta}, z = 0, \nu)$. The line shape function $\phi(\vec{\theta}, \nu)$ can be modelled as

$$\phi(\vec{\theta}, \nu) = \frac{1}{\sqrt{2\pi}\sigma_\nu} \exp\left[-\frac{1}{2}\left(\frac{\nu - \nu_0(1 - \frac{v_L(\vec{\theta})}{c})}{\sigma_\nu}\right)^2\right], \tag{2.28}$$

where σ_ν gives the width of the spectral line. The line of sight velocity $v_L(\vec{\theta})$ can be written in terms of the contribution from the tangential rotation velocity $v^\Omega(\vec{\theta})$ and the contribution from the random turbulent velocity $\delta v^T(\vec{\theta})$ as $v_L(\vec{\theta}) = v^\Omega(\vec{\theta}) + \delta v^T(\vec{\theta})$. For an exactly face on disc $v^\Omega(\vec{\theta}) = 0$. We define the $j^{\rm th}$ moment of the specific intensity and the local average of it as

$$M_j(\vec{\theta}) = \int {\rm d}\nu \, \nu^j \, I(\vec{\theta},\nu), \tag{2.29}$$

$$\langle M_j(\vec{\theta})\rangle = \int {\rm d}\vec{\theta}' M_j(\vec{\theta}')L(\vec{\theta}-\vec{\theta}'), \tag{2.30}$$

were $L(\vec{\theta})$ is the local averaging kernel. The convolution of $M_j(\vec{\theta})$ with $L(\vec{\theta})$ brings out the large scale features of $M_j(\vec{\theta})$. The quantities $\langle M_j(\vec{\theta})\rangle$ gives the $\vec{\theta}$ dependence of the large scale variation in the column density, i.e $G(\vec{\theta})$ for $j = 0$ and the line of sight component of the tangential rotational velocity $v^\Omega(\vec{\theta})$ with $j = 1$. Note that the definition of $M_j(\vec{\theta})$ used here does not correspond to the usual definition of the image moments.

2.5.1 Visibility moments

The angular extent of the external galaxies we consider here are rather small, we can approximate the visibilities as Fourier transform of the sky brightness distribution, i.e,

$$V(\vec{U},\nu) = \int {\rm d}\vec{\theta} e^{-i2\pi\vec{\theta}.\vec{U}} \, A(\vec{\theta})I(\vec{\theta},\nu), \tag{2.31}$$

where $A(\vec{\theta})$ is the beam pattern of each antenna or the primary beam of the telescope. Here we assume that during the observation, the antennas are tracking the galaxy, i.e, the origin of our coordinate system (θ_x, θ_y, z) is always at the centre of the field of view of observation. In general, the antenna beam pattern is a function of the observing frequencies. However, for the observation of H I emission from nearby external galaxies, the fractional bandwidth is small enough and the variation of $A(\vec{\theta})$ with observing frequencies can be ignored. The $j^{\rm th}$ moment of the visibility can be written as

$$V_j(\vec{U}) = \int {\rm d}\nu\, \nu^j\, V(\vec{U},\nu). \tag{2.32}$$

Using the eqn (2.29) we define the followings

$$W_j(\vec{U}) = \int {\rm d}\vec{\theta} e^{-i2\pi\vec{\theta}.\vec{U}}\, A(\vec{\theta})\langle M_j(\vec{\theta})\rangle. \tag{2.33}$$

The quantities $W_j(\vec{U})$ carry the information about the large scale variation of the specific intensity ($j = 0$) and the line of sight component of the tangential rotational velocity ($j = 1$) of the galaxy.

2.5.2 Power spectrum of column density and velocity

The power spectrum of the column density $P_{\rm HI}$ and the turbulent velocity P_v fluctuations are defined as

$$\langle \delta\tilde{n}(\vec{U}')^*\delta\tilde{n}(\vec{U})\rangle = \delta(\vec{U}-\vec{U}')P_{\rm HI}(\vec{U}') \tag{2.34}$$

$$\langle \delta\tilde{v}^T(\vec{U}')^*\delta\tilde{v}^T(\vec{U})\rangle = \delta(\vec{U}-\vec{U}')P_v(\vec{U}'). \tag{2.35}$$

Here $\delta\tilde{n}$ and $\delta\tilde{v}^T$ are the two dimensional Fourier transform (as used in eqn (2.31)) of the density and velocity fluctuations $\delta n(\vec{\theta}), \delta v^T(\vec{\theta})$ respectively. The symbol (*) denotes complex conjugate and $\delta(\vec{U})$ is the two dimensional Dirac delta function.

The function $V_0(\vec{U})$ has the information of the large scale variation of the intensity across the sky, as well as the fluctuations in the column density, whereas the function $W_0(\vec{U})$ depends only on the large scale variation of the intensity across the sky. We define $X_0(\vec{U}) = V_0(\vec{U}) - W_0(\vec{U})$, which gives the convolution of the column density fluctuation with $W_0(\vec{U})$, i.e,

$$X_0(\vec{U}) = W_0(\vec{U}) \otimes \delta\tilde{n}(\vec{U}). \tag{2.36}$$

The symbol (⊗) denotes convolution here. We use Q_0 to define the following autocorrelation function of X_0:

$$\begin{aligned} Q_0(\vec{U},\vec{U}') &= \langle X_0(\vec{U}')^* X_0(\vec{U})\rangle \\ &= \int \mathrm{d}\vec{U}_1 W_0(\vec{U}-\vec{U}_1) W_0(\vec{U}'-\vec{U}_1) P_{\mathrm{HI}}(\vec{U}_1). \end{aligned} \tag{2.37}$$

In the discussion of visibility based power spectrum estimator in Dutta et al. (2009) and in several earlier work (Bharadwaj and Sethi, 2001), it is demonstrated that such autocorrelation function decreases rapidly with increasing $\theta_0 \mid \vec{U}-\vec{U}' \mid$ for a galaxy with angular extent of θ_0 in the sky. At baselines $\mid U \mid > 1/\theta_0$, we may write from eqn (2.37)

$$P_{\mathrm{HI}}(U) = \lim_{|U|\theta_0\to\infty, \theta_0|\vec{U}-\vec{U}'|\to 0} \frac{Q_0(\vec{U},\vec{U}')}{\int \mathrm{d}\vec{U}' \mid W_0(\vec{U}') \mid^2}. \tag{2.38}$$

Hence, this works as an estimator for the H I column density fluctuations. Here we have also assumed that the fluctuations in the column density are statistically homogeneous and isotropic and hence are the only function of $U = \mid \vec{U} \mid$.

To estimate the power spectrum of the turbulent velocity fluctuations, we proceed as follows. We define $X_1(\vec{U}) = \frac{c}{\nu_0}\left[V_1(\vec{U}) - W_1(\vec{U})\right]$, where c is the speed of light. $X_1(\vec{U})$ can be written in terms of the column density and turbulent velocity fluctuations as

$$X_1(\vec{U}) = \frac{c}{\nu_0}\, W_1(\vec{U}) \otimes \delta\tilde{n}(\vec{U}) + W_0(\vec{U}) \otimes \delta\tilde{v}^T(\vec{U}). \tag{2.39}$$

The autocorrelation of X_1 gives

$$\begin{aligned} Q_1(\vec{U},\vec{U}') &= \langle X_1(\vec{U}')^* X_1(\vec{U})\rangle \\ &= \int \mathrm{d}\vec{U}_1 W_0(\vec{U}-\vec{U}_1)W_0(\vec{U}'-\vec{U}_1)P_v(\vec{U}_1) \\ &+ \left(\frac{c}{\nu_0}\right)^2 \int \mathrm{d}\vec{U}_1 W_1(\vec{U}-\vec{U}_1)W_1(\vec{U}'-\vec{U}_1)P_{\rm HI}(\vec{U}_1) \end{aligned} \tag{2.40}$$

Here we have assumed that the standard deviation in the relative fluctuations in the column density δn is smaller than unity. We also assume that the standard deviation of the line of sight component of the random turbulent velocity δv^T is smaller than the line of sight component of the tangential rotation velocity v^{Ω}. We neglect the contribution in $Q_1(\vec{U},\vec{U}')$ that arises from multiplication of δv^T and v^{Ω}. The power spectrum of H I velocity fluctuations can be written as

$$P_v(U) = \lim_{|U|\theta_0\to\infty,\theta_0|\vec{U}-\vec{U}'|\to 0} \frac{\left[Q_1(\vec{U},\vec{U}') - H(\vec{U},\vec{U}')\right]}{\int \mathrm{d}\vec{U}' \mid W_0(\vec{U}') \mid^2}, \tag{2.41}$$

where $H(\vec{U},\vec{U}')$ is the second term in the right hand side of eqn (2.40). Hence, this works as an estimator for the H I velocity fluctuations.

2.6 Summary

Following are the summary of this chapter.

- H I 21cm radiation is an effective probe to measure the structure and dynamics in the interstellar medium.
- Radio interferometers can be used for H I observations and the basis of their imaging techniques is discussed.

- To probe the turbulent stochastic fluctuations in the ISM and their origin, two point statistical estimators like power spectrum can be used.
- A controlled test on the efficacy of two types of power spectrum estimators; imaged based estimators and visibility based estimators are presented where it was shown that the visibility based estimator is trustworthy.
- Motivating from the conclusions of the above study, a variant of visibility moment estimator (VME) for column density and turbulence velocity power spectrum was introduced. The existence of turbulence cascade can be detected by implementing these estimators in real observational samples of galaxies.