Estimating statistics of sky brightness using radio interferometric observations

RAA 2019 Vol. **19** No. **4**, 60(18pp) doi: 10.1088/1674–4527/19/4/60

http://www.raa-journal.org *http://iopscience.iop.org/raa*

Research in Astronomy and Astrophysics

Estimating statistics of sky brightness using radio interferometric observations

Prasun Dutta and Meera Nandakumar

Department of Physics, IIT (BHU) Varanasi, 221005 India; *pdutta.phy@itbhu.ac.in, meeranandakr.rs.phy17@itbhu.ac.in*

Received 2018 June 5; accepted 2018 November 5

Abstract Radio interferometers are used to construct high resolution images of the sky at radio frequencies and are the key instruments for accessing the statistical properties of the evolution of neutral hydrogen over cosmic time. Here we use simulated observations of the model sky to assess the efficacy of different estimators of the large-scale structure and power spectrum of the sky brightness distribution. We find that while the large-scale distribution can be reasonably estimated using the reconstructed image from interferometric data, estimates of the power spectrum of the intensity fluctuations calculated from the image are generally biased. This bias is found to be more pronounced for diffuse emission. The visibility based power spectrum estimator, however, gives an unbiased estimate of the true power spectrum. This work demonstrates that for an observation with diffuse emission the reconstructed image can be used to estimate the large-scale distribution of the intensity, while to estimate the power spectrum, visibility based methods should be preferred. With the upcoming experiments aimed at measuring the evolution of the power spectrum of the neutral hydrogen distribution, this is a very important result.

Key words: instrumentation: interferometers — methods: data analysis — techniques: interferometric

1 INTRODUCTION

Radio interferometers are the key instruments used to map spatially resolved sky brightness fluctuations at radio frequencies (Taylor et al. 1999). Within radio frequencies, different emission mechanisms exist (Rybicki & Lightman 1985), including continuum emission by synchrotron radiation, 21-cm line emission from neutral hydrogen gas or H I to name a few. Various statistics of the sky brightness distribution are evaluated from the radio interferometric data to investigate different properties of Galactic and extragalactic sources including the magnetic field as well as relativistic electron distributions in galaxies (Sukumar & Allen 1989; Basu & Roy 2013), the morphology and dynamics of H I in nearby dwarf and spiral galaxies (Begum et al. 2008; Walter et al. 2008; Dutta et al. 2010a, 2009, 2013), the origin and evolution of radio jets and lobes in radio galaxies (Nandi & Saikia 2012), structures of supernova remnants (Roy et al. 2009), cosmological evolution of H I (Ghosh et al. 2011), etc.

A class of these investigations uses large-scale distribution of the source emission to infer local physical properties like the morphology of supernova remnants, radio galaxies and spiral galaxies, the radial distribution of H I and rotational velocities or the relation between star formation and gas distribution in spiral galaxies, from one-point statistics like the mean or median of the specific intensity at different positions within the source. On the other hand, two-point statistics, like the autocorrelation function, structure function or power spectrum of the sky brightness fluctuations carry important physical information like the properties of magnetohydrodynamic (MHD) turbulence in supernova remnants or hydrodynamic turbulence in the H I in galaxies and evolution of the distribution of H I during the cosmic dawn, epoch of reionization and post-reionization era. In most of these cases, a power spectrum of the specific intensity is evaluated from the observed interferometric data (Dutta & Bharadwaj 2013; Zhang et al. 2012).

Radio interferometers inherently measure a quantity called visibility, a complex transform of the sky brightness distribution. Roughly speaking, for many of the scientific cases described above, the visibilities can be approximated as the Fourier transform of the sky brightness distribution measured at certain spatial frequencies. Estimation of the one-point statistics of the sky brightness distribution re-

quires a faithful reconstruction of it (the reconstruction is usually referred to as the image) from the visibilities. As the visibilities are often not measured at all spatial frequencies, reconstruction of the image is not straightforward and involves complicated algorithms to deconvolve the effect of this incomplete measurement.

One approach to estimating the two-point statistics is to use the reconstructed image and calculate either the structure function or the autocorrelation function in the image plane, or estimate the power spectrum of the image in its Fourier conjugate plane. We shall call the estimators of two-point statistics, that rely on the reconstructed image, the *image based estimators* (e.g. Zhang et al. 2012; Walker et al. 2014; Grisdale et al. 2017). Using image based estimators, it is possible to evaluate the two-point statistics of a part of the astrophysical source under consideration. This is essential in some particular cases, viz., in the correlation of star formation with turbulence in the interstellar medium (ISM), a variation of the MHD turbulence in the arm and inter arm regions of spiral galaxies (Basu & Roy 2013), etc. However, any problem in image reconstruction is likely to show up as artifacts in the image based estimation of the two-point statistics and needs to be investigated thoroughly before using for scientific inference. On the other hand, a different class of estimators of two-point statistics is used in the literature where visibilities are directly implemented to estimate the power spectrum of the sky brightness distribution (Dutta et al. 2009; Choudhuri et al. 2016). As this does not need image reconstruction, these estimators are more direct and are not prone to artifacts of image reconstruction. We shall call these estimators the *visibility based estimators*. As the visibility based estimators implement the visibilities directly, they cannot be used to estimate the two-point statistics of a part of the image. However, in a few particular cases, these limitations can be overcome by applying suitable techniques (see e.g., Dutta et al. 2010a).

Image reconstruction from the visibilities is a long-standing problem (Thompson et al. 2017), and we shall discuss it in detail in the next section. The objective here is to have the best guess of the sky brightness distribution from limited observations that an interferometer provides. The question is how accurately the one and two-point statistics can be computed from these reconstructions. As the reconstruction process requires various input parameters from astronomers, the outcome is not unique (see e.g., Cornwell et al. 1999; Condon et al. 1998; Becker et al. 1995; White et al. 2007; Taylor et al. 1999). In this paper, we investigate the efficacy of different estimators for the first and second order statistics of radio interferometric data using simulated observations. Engaging with all the different classes of radio interferometric observations as well as all the different imaging techniques is not possible within the scope of a single paper. Here we have focused on a particular problem of estimating the first and second order statistics of the H I emission of spiral galaxies. Our results are directly applicable to, but not limited to, cases involving the same class of problems.

The rest of the paper is structured as follows: Section 2 describes H I observations and reconstruction of the image; simulation of the interferometric observation is discussed in Section 3; Section 4 focuses on the visibility and image based statistical estimators; we present the result and analysis of the simulated data in Section 5; and we conclude in Section 6.

2 BRIEF OVERVIEW OF RADIO INTERFEROMETRIC OBSERVATION

Here we give a brief overview of interferometric observations; interested readers may refer to Taylor et al. (1999) for further details. Radio interferometers are a collection of many array elements, called antennas, arranged in a specific pattern on the surface of the Earth[1]. 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 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 $\boldsymbol{U}$ is the ratio of instantaneously projected separations 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, i.e., $V(\boldsymbol{U})$. However, they can be measured only at the discrete values $\boldsymbol{U_i}$, which correspond to the physical baselines offered by all the antenna pairs. We introduce a function $S(\boldsymbol{U})$ to capture this sampling of the baselines by the interferometer

$$S(\boldsymbol{U}) = \sum_{i=1}^{N_b} \delta_D(\boldsymbol{U} - \boldsymbol{U_i}) \ . \tag{1}$$

[1] In principle, there can be "zero spacing" interferometers (Mahesh et al. 2014), interferometers with a very small number of baselines and the antenna can be kept in space. However, most of the interferometers used today consist of many antennas on the surface of the Earth.

Henceforth, our discussion will be limited to the observation of H I 21-cm radiation from nearby spiral galaxies. Typically, these galaxies are about $\sim 2-20$ Mpc away and hence their extent in the sky is limited to $\sim 40'$ or lower. Interferometers with a relatively narrow FOV of about one degree are sufficient to observe these galaxies. Existing radio interferometers like the VLA[2], GMRT[3], etc., have the required FOV. Here, we use the vector $\boldsymbol{\theta}$ to denote the angular separation between any point in the FOV with respect to the center of the FOV. Correspondingly, the baseline vector can also be taken as two dimensional and can be considered as the Fourier conjugate of $\boldsymbol{\theta}$. In such approximation, the visibilities become the Fourier transform of the sky brightness distribution (Sault & Conway 1999) sampled at the baselines where the interferometric measurements are done, i.e.,

$$V(\boldsymbol{U}) = \tilde{I}(\boldsymbol{U})S(\boldsymbol{U}) + \mathcal{N}(\boldsymbol{U}) \ , \tag{2}$$

where $\tilde{I}(\boldsymbol{U})$ is the Fourier transform of the sky brightness distribution $I(\boldsymbol{\theta})$ and $\mathcal{N}(\boldsymbol{U})$ is the measurement noise. This is the quantity directly measured by radio interferometers. Note that, to completely describe a real measurement, we need to also scale the first term on the right hand side of Equation (2) by the gain of the interferometer. However, here we assume that a proper calibration procedure is followed to take care of the effect of the gain. Note that we have neglected the effect of the antenna primary beam here. This is justified if the angular extent of the galaxy is much smaller than the full width at half maximum (FWHM) of the primary beam of the antenna. On the other hand, if the angular extent of the galaxy is larger, the primary beam can be included in the window function as described in Equation (4) in the next section.

The visibilities, as given by Equation (2), are used directly in the visibility based estimators of the two-point statistics of the sky brightness distribution. However, to estimate the one-point statistics at different points in the sky or to estimate the two-point statistics using an image based estimator, reconstruction of the sky brightness distribution is necessary from the observed visibilities.

The inverse Fourier transform of the measured visibility is called the dirty image

$$I_D(\boldsymbol{\theta}) = I(\boldsymbol{\theta}) \otimes B_D(\boldsymbol{\theta}) \ . \tag{3}$$

Here $B_D(\boldsymbol{\theta})$ is the Inverse Fourier transform of the weighted sampling function $S(\boldsymbol{U})$ 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 often called the dirty beam as it has secondary maxima around the center, known as the side lobes. 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, and often irregular and incomplete, the interferometer PSF can be quite complicated thereby making the deconvolution procedure nontrivial. Different algorithms have been devised for this purpose, including types of CLEAN (Högbom 1974; Cotton 1979; Clark 1980; Schwab 1984), Maximum Entropy Image Reconstruction (MEM) (Narayan & 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 & Conway 1999).

Ever since the first version of CLEAN was outlined in Högbom (1974), it has been widely used and also widely evolved. In CLEAN, the sky image is assumed to be a collection of point sources. The algorithm relies on estimating the brightness and position of all the point sources in the sky from the image using an iterative procedure. This is achieved in different ways in different variations of CLEAN. Here we have utilized the Cotton-Schwab variant of CLEAN (Clark 1980). 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 implemented, the effective dirty beam is to be considered as the Fourier transform of the sampling function multiplied by the weighting kernel. As with most interferometers, the baseline coverage is better at lower values of $|\boldsymbol{U}|$, and the PSF derived from the naturally weighted visibilities is broader but 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. An iterative weighting scheme called 'adaptive weighting' is discussed in Yatawatta (2014) to assist high bandwidth and

[2] NRAO-VLA: Very Large Array, New Mexico

[3] Giant Metrewave Radio Telescope, NCRA-TIFR

high dynamic range imaging by minimizing the PSF variation across frequencies while maximizing the sensitivity. The detailed process of gridding, different weightings and their effects are discussed in Thompson et al. (2017). We have used the default gridding convolution function in the task IMAGR in AIPS[4] which implements a prolate spheroid wave function. Note that the weighting schemes are used to reduce the artifacts in the image that may arise from the incomplete baseline coverage. However, they may introduce additional effects, which one needs to assess and quantify before applying the reconstructed image for scientific inferences.

The efficacy of the CLEAN reconstruction algorithm, and its limitations and known artifacts are discussed in detail in Thompson et al. (2017); we highlight a few relevant points here. It has been shown in Schwarz (1978) and Schwarz (1979) that in the absence of noise and in the case when a lower number of point sources is required to construct the sky brightness distribution than the independent measurement of visibilities, CLEAN reduces to a least squares image estimation procedure. However, in a real scenario, particularly with extended sources and measurement noise, CLEAN does not produce the unique reconstruction of the sky. The reconstructed image depends on many user selectable parameters, like the loop gain, size of the dirty beam patch used in the minor cycle, the weighting and tapering schemes, etc. For example, it is recommended to use a smaller loop gain and a larger patch of the dirty beam in the minor cycles in the case of diffuse emissions (Taylor et al. 1999).

The uncorrelated measurement noise in the visibilities gives rise to correlated noise in any reconstructed image and cannot be avoided. This is also the source of correlated noise present in image based estimates of the power spectrum in the case of incomplete baseline coverage (see Junklewitz et al. 2015).

3 SIMULATING A MODEL VISIBILITY DATASET

In this article, we are interested in investigating the efficacy of different statistical estimators used to interpret radio interferometric data. We proceed by simulating an observation with a known sky model based on the observed H I emission from nearby external spiral galaxies (similar models are used in Dutta et al. 2009). We write the specific intensity distribution from such a galaxy as a function of the angle $\boldsymbol{\theta}$ (for simplicity the x and y components of the vector $\boldsymbol{\theta}$ can be considered along the local directions of Right Accession and Declination) from the center of the galaxy as

$$I(\boldsymbol{\theta}) = W(\boldsymbol{\theta}) \left[\bar{I} + \delta I(\boldsymbol{\theta})\right] \ , \tag{4}$$

where $W(\boldsymbol{\theta})$ quantifies the large-scale distribution of the H I column density in the sky and is normalized as $\int W(\boldsymbol{\theta}) d\boldsymbol{\theta} = 1$. We call this the window function.[5] The quantity $\bar{I}$ is the total intensity coming from the entire galaxy and $\delta I(\boldsymbol{\theta})$ 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.

3.1 Modeling the 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. 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 the shapelet basis $S_{nm}(\boldsymbol{\theta}, \beta)$ and the shapelet coefficients f_{nm}, the specific intensity can be written as

$$I(\boldsymbol{\theta}) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} f_{nm} S_{nm}(\boldsymbol{\theta}, \beta) \ . \tag{5}$$

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 define the window function as

$$W(\boldsymbol{\theta}) = \frac{\sum_{n=0}^{N} \sum_{m=-n}^{n} f_{nm} S_{nm}(\boldsymbol{\theta}, \beta)}{\int d\boldsymbol{\theta} \sum_{n=0}^{N} \sum_{m=-n}^{n} f_{nm} S_{nm}(\boldsymbol{\theta}\ , \beta)}, \tag{6}$$

where we truncate the summation in Equation (5) to a certain value of N and hence capture only the large-scale structure of the sky brightness distribution. Notice also that the window function so defined is normalized such that $\int W(\boldsymbol{\theta}) d\boldsymbol{\theta} = 1$. Exact considerations for choosing the parameters N and β will be discussed shortly.

[4] NRAO AIPS: Astrophysical Image Processing System

[5] Since we are mostly interested in estimating the power spectrum of $\delta I(\boldsymbol{\theta})$, in general, the window function can be thought of as a multiplication of the galaxy window function with the primary beam of the interferometer. However, in most cases the angular extent of the galaxy is smaller than the primary beam and the latter can be ignored.

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

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

$$\left\langle \delta\tilde{I}^{*}(\boldsymbol{U})\delta\tilde{I}(\boldsymbol{U}')\right\rangle = \delta_D(\boldsymbol{U}-\boldsymbol{U}')P(U)\ , \qquad (7)$$

where $U = |\boldsymbol{U}|$. The angular brackets denote ensemble averages. Observations in our Galaxy (Crovisier & Dickey 1983; Green 1993) and external dwarf and spiral galaxies (Elmegreen et al. 2001; Begum et al. 2006; Dutta et al. 2009, 2013; Dutta & Bharadwaj 2013) suggest that the H I specific intensity fluctuations can be modeled 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.,

$$\mathcal{R} = \frac{\bar{I}}{\sigma_{\delta I}}\ . \qquad (8)$$

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

3.3 Modeling the Specific Intensity Distribution $I(\boldsymbol{\theta})$

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 (naturally weighted) of NGC 628 taken from the THINGS (Walter et al. 2008) survey as a data product[6] 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 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 $N \leq 12$ 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 Figure 1 represents the moment zero map (naturally 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 differs. 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 displays the power spectrum of a model sky brightness distribution $I(\boldsymbol{\theta})$ (Eq. (4)) with a blue dashed line. The values of the model parameters are $\mathcal{R} = 5,\ \alpha = -1.5$. The green dot-dashed line indicates a power law with slope -1.5. The power spectrum of only the first term, $W(\boldsymbol{\theta})\bar{I}$, is signified with a black solid line. Clearly, for baselines lower than $1\ k\lambda$, the power spectrum of the model $I(\boldsymbol{\theta})$ is dominated by the window function, while for baselines greater than $1\ k\lambda$, the power spectrum of the model image $I(\boldsymbol{\theta})$ 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 and -2.2. Moreover, 9 of the 18 galaxies in their sample have α between -1.5 and -1.8. We choose three values of α for our model sky image: $[-0.5, -1.5, -2.0]$. Dutta & Bharadwaj (2013) found that $\mathcal{R}$ varies between 5 and 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.

3.4 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[7]. 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 that this compromise is made to increase the computational speed and does not limit our analysis of 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 the equivalent of eight hours of simulated observation.

[6] THINGS: The HI Nearby Galaxy Survey
data product: `http://www.mpia.de/THINGS/Data.html`

[7] GMRT original array configurations can be seen in `http://www.gmrt.ncra.tifr.res.in/gmrt_hpage/Users/doc/obs_manual`

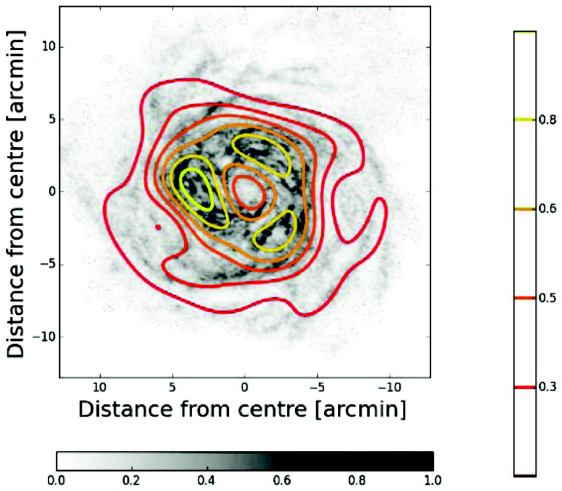


Fig. 1 Greyscale image showing the naturally weighted moment zero map of the galaxy NGC 628. Contours represent the model window function based on this map (see Eq. (6)). Note that the range of values in the pixels of these maps differs. We have scaled the pixel values to keep the maximum pixel value as unity for both the maps.

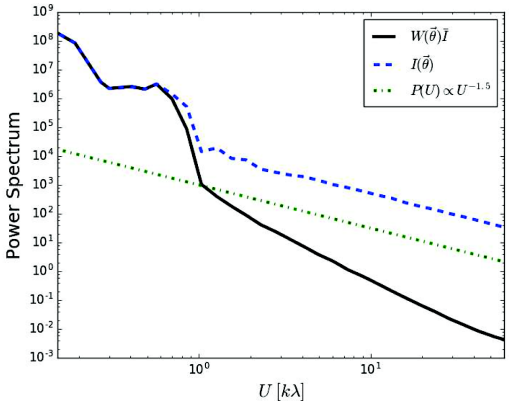


Fig. 2 Power spectrum for the first term $W(\boldsymbol{\theta})\bar{I}$ in Equation (4) (*solid black line*) is compared with the power spectrum of $I(\boldsymbol{\theta})$ for the model image (*blue dashed line*) with $\mathcal{R} = 5$ and $\alpha = -1.5$. Observe that for $U > 1\ k\lambda$ there is a significant deviation between the two. The *green dot-dashed line* corresponds to a power law with index -1.5. Clearly, the power spectra of $I(\boldsymbol{\theta})$ follow a power law for $U > 1\ k\lambda$ with a slope of -1.5.

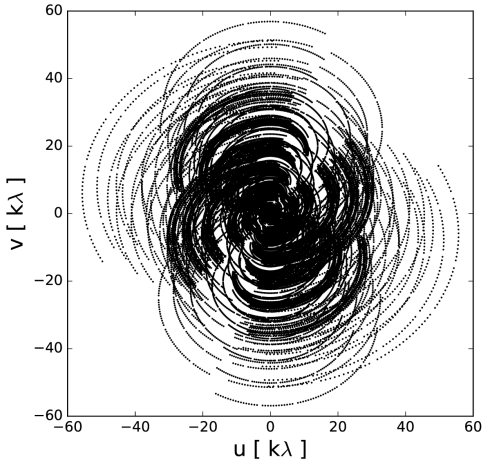


Fig. 3 Figure showing the sampling function for the simulated observation presented here. *Black points* are the places in the baseline plane where the visibilities are measured.

Figure 3 displays 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 4. Clearly, the power spectra of the three beams are significantly different. We use a 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 on 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.

4 DIFFERENT STATISTICAL ESTIMATORS

4.1 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 large star formation rates (Wang et al. 2014). Bagetakos et al. (2011) have observed the presence of H I holes at small knots of star formation in the disks of spiral galaxies. Dwarf galaxies like GR 8 have a clear spatial correlation between star formation rate and H I column density (Begum & 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 a local average of the specific intensity as $\langle I(\boldsymbol{\theta})\rangle = \bar{I}W(\boldsymbol{\theta})$.

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 take the shapelet scale to be $240''$ and the first 12 shapelet coefficients to represent the window function. To distinguish between the window function estimated from the reconstructed image and the model window function, we mention the earlier $W_C(\boldsymbol{\theta})$ for the rest of the analysis. Further, for the model ($W(\boldsymbol{\theta})$) as well as the reconstructed windows ($W_C(\boldsymbol{\theta})$), we estimate the azimuthally averaged window function defined as

$$W_A(\theta) = \frac{1}{2\pi}\int_0^{2\pi} W(\theta,\phi)d\phi \tag{9}$$

in different bins of θ. Here (θ,ϕ) are the polar components of the vector $\boldsymbol{\theta}$. We use the standard deviation of the values of the reconstructed window $W_C(\boldsymbol{\theta})$ 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 those of the model.

4.2 Two-Point Statistics: Power Spectrum

Two-point statistics of any field quantifies the scale dependence of fluctuations in it. There are several quantifiers for two-point statistics. For a two dimensional field $A(\boldsymbol{\theta})$, the structure function and the autocorrelation function evaluate the two-point statistics in the $\boldsymbol{\theta}$ plane while its power spectrum evaluates the two-point statistics in a plane Fourier conjugate to $\boldsymbol{\theta}$, like the baseline plane $\boldsymbol{U}$ (see Elmegreen & Scalo 2004 for more detailed analysis of these estimators). For a Gaussian random field, all these different estimators contain the same information. We restrict ourselves to measuring the power spectrum of the sky brightness fluctuations, i.e. the quantity $\delta I(\boldsymbol{\theta})$ here. As discussed before, power spectrum estimators from the interferometric data can be categorized into two classes, *image based estimators* and *visibility based estimators*. We give a brief description of these estimators here.

4.2.1 Image based power spectrum estimator

Image based estimators use the reconstructed image to estimate the power spectrum. To distinguish the reconstructed image from the sky brightness distribution, we shall denote the former by $I_C(\boldsymbol{\theta})$. Since the image is already evaluated at regular grids in $\boldsymbol{\theta}$, a two dimensional fast Fourier transform can be used to estimate the Fourier transform of $I_C(\boldsymbol{\theta})$. As the interferometers are mostly not sensitive at baselines lower than a certain value, they effectively do not measure the first term in Equation (4) and we may write

$$\tilde{I}_C(\boldsymbol{U}) = \tilde{W}(\boldsymbol{U}) \otimes \delta\tilde{I}(\boldsymbol{U}) + \mathcal{B}_I(\boldsymbol{U})\ , \tag{10}$$

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

at each baseline, which gives

$$P_C(\boldsymbol{U}) = \langle \tilde{I}_C(\boldsymbol{U})\tilde{I}^*_C(\boldsymbol{U})\rangle = \mid \tilde{W}(\boldsymbol{U}) \mid^2 \otimes P(\boldsymbol{U}) + P_{\mathcal{B}}(\boldsymbol{U}) \ , \tag{11}$$

where $P_{\mathcal{B}}(\boldsymbol{U})$ is related to both $\mathcal{B}_I(\boldsymbol{U})$ and $\tilde{I}(\boldsymbol{U})$. The angular brackets above denote 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

$$P_I(U) = \frac{1}{2\pi} \int_0^{2\pi} \tilde{I}_C(U,\phi)\tilde{I}^*_C(U,\phi) d\phi \ , \tag{12}$$

where (U,ϕ) are the polar components of the vector $\boldsymbol{U}$. As discussed before (see Fig. 2), the window function represents large-scale variation of the specific intensity and hence at baselines $U \gg 1/\theta_0$, where θ_0 is the extent of the galaxy, the window function can be treated as a delta function $\delta_D(\boldsymbol{U})$. Hence in the absence of $P_{\mathcal{B}}$, at baselines $U \gg 1/\theta_0$, the quantities $P_C(\boldsymbol{U}) \sim P(\boldsymbol{U})$ and $P_I(U)$ give an estimate of the power spectrum of the sky brightness fluctuations. The quantity $P_{\mathcal{B}}(\boldsymbol{U})$ 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 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 largely ignored in literature where image based estimators are used (Zhang et al. 2012; Walker et al. 2014). Further, incomplete baseline coverage also makes the measurement noise correlated in $\tilde{I}_C(\boldsymbol{\theta})$ and introduces a nonzero $P_{\mathcal{B}}(\boldsymbol{U})$ (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 & Conway 1999). Following that, 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 $\boldsymbol{U}$ 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.

4.2.2 Visibility based power spectrum estimator

The visibility based power spectrum estimators use the directly measured visibilities and do not require image reconstruction (Bharadwaj & Sethi 2001). Since the Fourier transform of the first term in Equation (4) is mostly not measured by an interferometer (see the discussion above), the measured visibilities can be written as

$$V(\boldsymbol{U}) = \left[W(\tilde{\boldsymbol{U}}) \otimes \delta\tilde{I}(\boldsymbol{U})\right] S(\boldsymbol{U}) + \mathcal{N}(\boldsymbol{U}) \ . \tag{13}$$

The visibility correlation gives

$$\begin{aligned}\langle V(\boldsymbol{U})V^*(\boldsymbol{U})\rangle = \mid W(\tilde{\boldsymbol{U}}) \mid^2 \otimes P(U) \mid S(\boldsymbol{U}) \mid^2 \\ + \mid \mathcal{N}(\boldsymbol{U}) \mid^2 \ .\end{aligned} \tag{14}$$

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

$$P_V(U) = \int_0^{2\pi} \langle V(\boldsymbol{U})V^*(\boldsymbol{U})\rangle d\phi \Big/ \int_0^{2\pi} \mid S(\boldsymbol{U}) \mid^2 d\phi. \tag{15}$$

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 $\mid \mathcal{N}(\boldsymbol{U}) \mid^2$ 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).

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 to 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 an abrupt

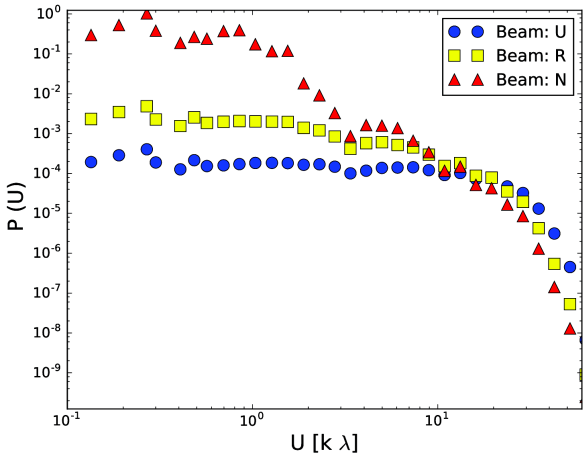


Fig. 4 Figure showing the power spectrum of the dirty beam for the uniform, robust and naturally weighted beams for the sampling function given in Fig. 3.

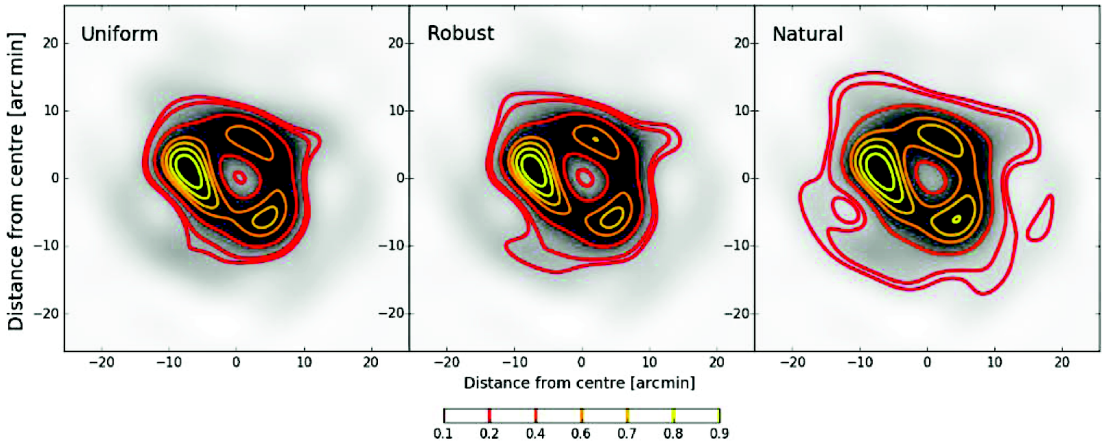


Fig. 5 Window functions of the reconstructed images using different weighting schemes are shown in contours against the greyscale plot of the window function of the model image corresponding to $\mathcal{R} = 5,\ \alpha = -1.5$. The scales of the contours are relative here.

fall in the baseline coverage at large baseline we use a Gaussian taper with the UVTAPER parameter set to 45 $k\lambda$ 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 $k\lambda$. Tapering down-weights the visibilities at the larger baselines and hence may have an effect on 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 natural weighting, uniform weighting and 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 naturally 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 $\sim 63°$, $\sim 68°$ and $\sim 70°$ respectively. These images are used for further analysis.

5.1 Window Function

For each of the reconstructed images, 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 5 show W_C corresponding 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 1 demonstrates that the naturally 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 $\boldsymbol{\theta}$; we may denote the values in the grid points as $W[i,j]$ and $W_C[i,j]$ respectively. We use the quantity

$$\chi = \frac{\sum\limits_{i,j} \left(W[i,j] - W_C[i,j]\right)^2}{N_G^2 \sqrt{\sigma_W^2 + \sigma_C^2}} \tag{16}$$

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[8]. The quantity χ gives the mean square deviation between the model and the estimated window functions. Clearly, 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 actually 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 value of χ lies between 0.004 and 0.01 in the six models discussed here. These numbers can be considered as references while interpreting the results.

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 have a bias, however, the natural weighting scheme produces the best estimate of the window function.

Figure 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. Color bands around each of the estimates correspond to 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 natural (triangle) weightings reproduce the azimuthally averaged window within errors, however, the points corresponding to the robust weighting (square) are systematically offset from the model window function.

Figure 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 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 PSF. Since the window function represents the large scale distribution, the natural weighting is expected to reproduce it better. Note that the actual values of χ do not carry a robust probabilistic interpretation here. Moreover, for the models with a relatively lower amplitude of 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.

5.2 Power Spectrum

We estimate power spectra from the model images using the image based estimator. Note that the model images do not have any artifacts 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 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 the visibility based power spectrum follows the reference spectrum quite well and overall bias is minimized. The large error bars in small and larger baselines are indicative of fewer independent measurements at those baselines. The visibility based power spectrum estimator assumes a 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 by $1/T(U)^2$. We plot

[8] The quantity χ used here has no probabilistic interpretation and should not be confused with a functional used in most maximum likelihood estimations.

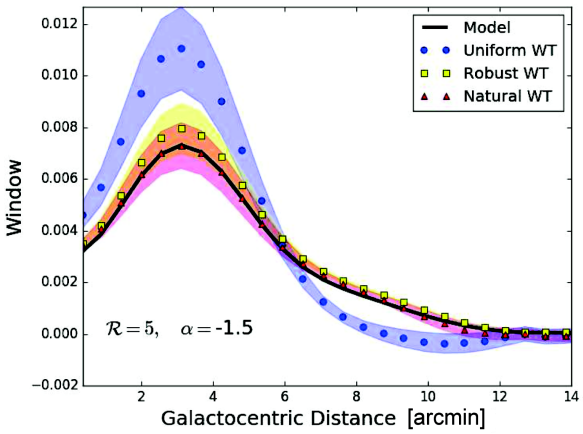


Fig. 6 This figure compares the azimuthally averaged window functions estimated using different weighting schemes with that of the input image. The color bands for each estimate represent the error.

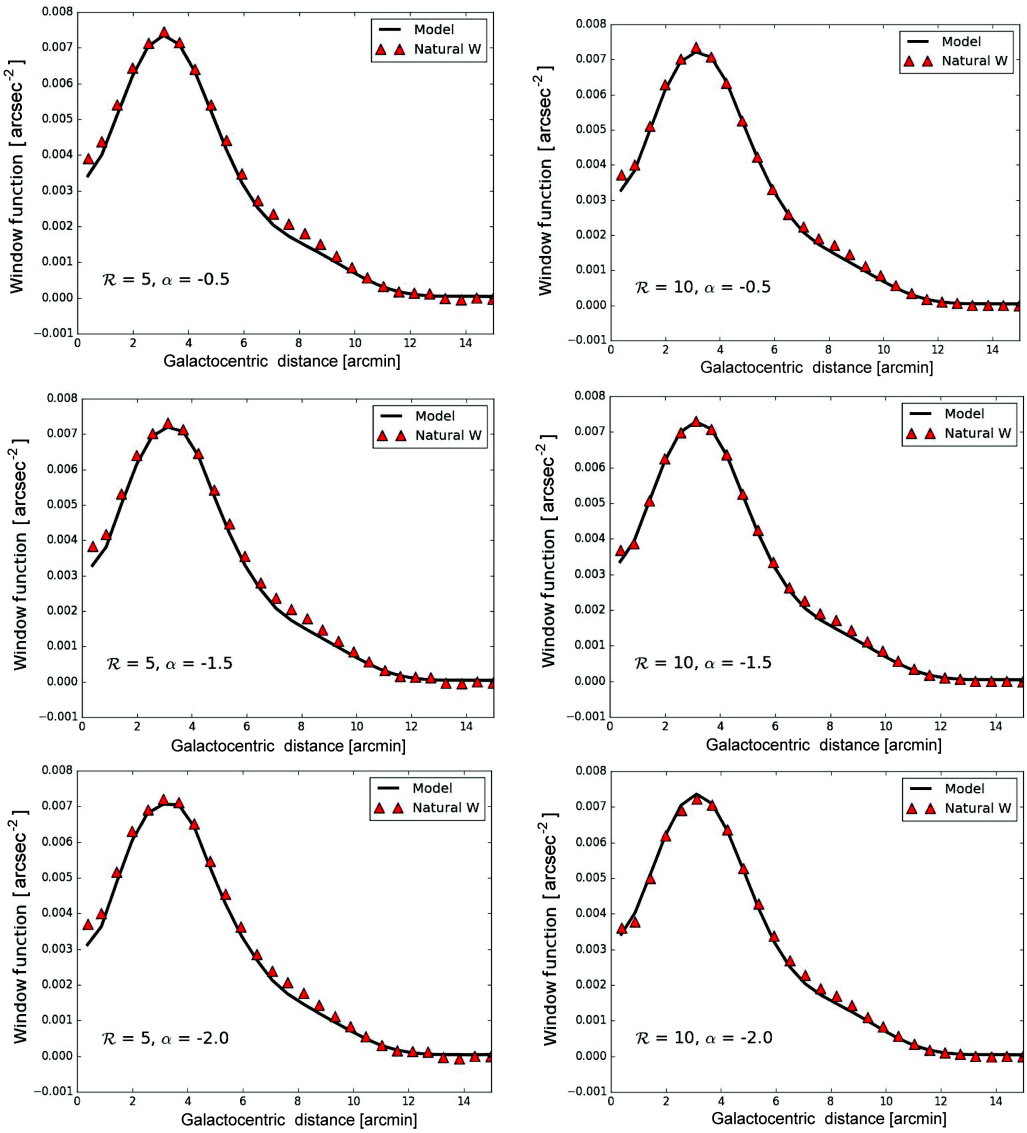


Fig. 7 Comparisons of the azimuthally averaged window functions estimated from the input model (*black solid line*) with the reconstructed naturally weighted image (*red triangles*) are shown for all six simulations.

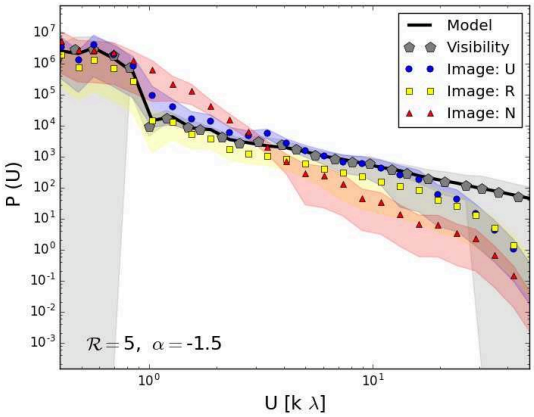


Fig. 8 This figure compares the power spectra estimated using the visibility based estimator and the image based estimator for all three different weighting schemes against the model power spectrum. The *shaded area* corresponds to the error associated with each estimate.

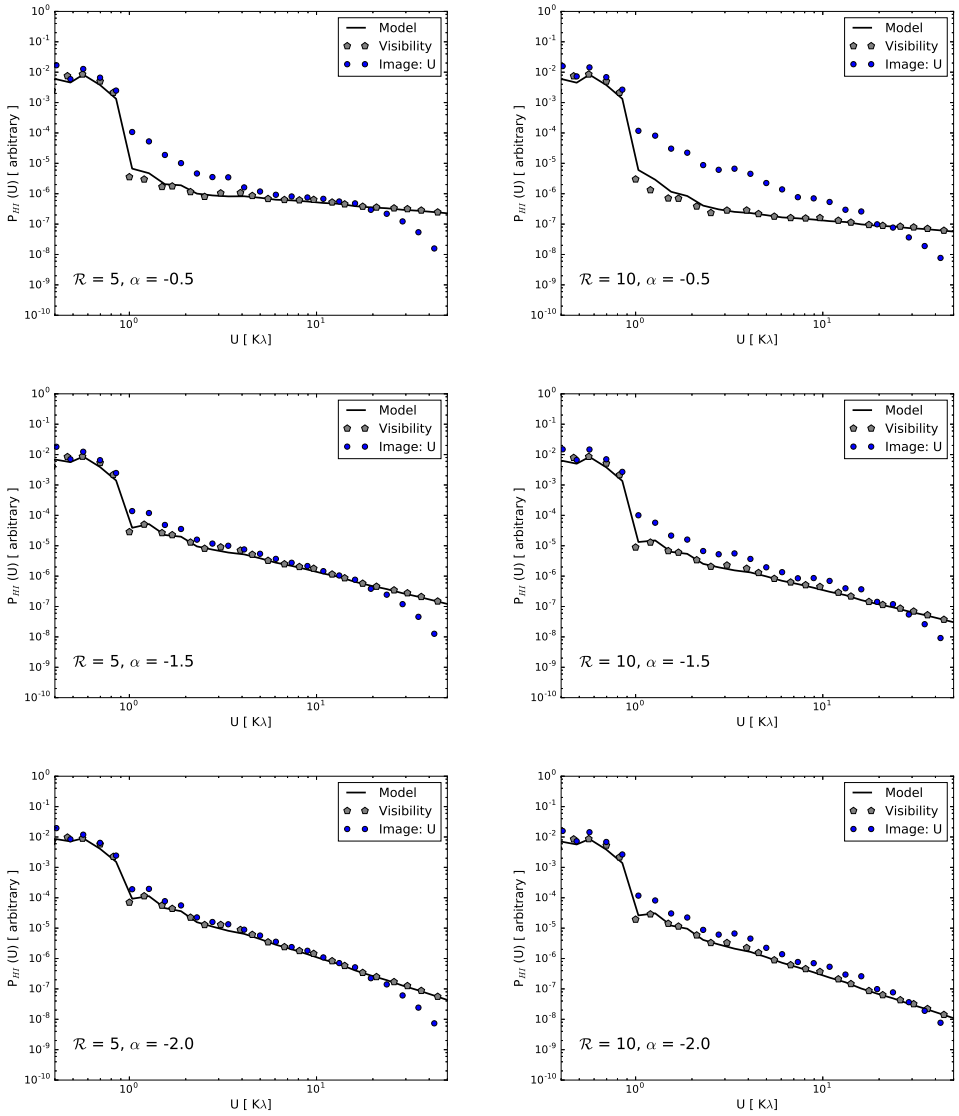


Fig. 9 Comparisons of the power spectra estimated from the input model (*black solid lines*), using the visibility based estimator (*grey pentagons*) and image based estimator from the uniform weighted image (*blue circles*) are shown for all six simulations. The amplitudes of the y-axes for different panels are arbitrary scales to keep the power spectra in the same range in all plots.

Table 1 Table summarizing the results of comparison between different estimates of the window function and the power spectrum. Efficacy at reproducing the model window function from an estimated image is quantified by χ (see Eq. (16)). We tabulate the values of α_V and α_I for each model to assess the merit of different estimators in the table. The headers 'U', 'R' and 'N' correspond to the uniform, robust and natural weighting schemes respectively.

		$\mathcal{R}=5$			$\mathcal{R}=10$		
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		

with circular, square and triangular markers in the same figure to represent 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 natural weighting scheme is drastically different from the model. The power spectrum estimated using the reconstructed image with the uniform weighting scheme almost follows the model power spectrum within 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 the image based estimates of the power spectra are, 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 varies as $-1.8 \pm 0.1, -2.0 \pm 0.1$ and -2.2 ± 0.1 for the uniform, robust and naturally 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 least bias among the image based estimators.

Figure 9 shows the power spectra 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 weightings. 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 that estimated using the visibility based estimator (α_V) in Figure 10. The 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 systematically smaller than the corresponding α_Vs with shallower power spectra having a systematically larger bias.

5.3 Application to Point Sources

Apart from the baseline coverage of the interferometer, efficacy of the image reconstruction may also depend on the structure of the sky brightness distribution itself. In the CLEAN algorithm, the sky is modeled 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 FOV. The

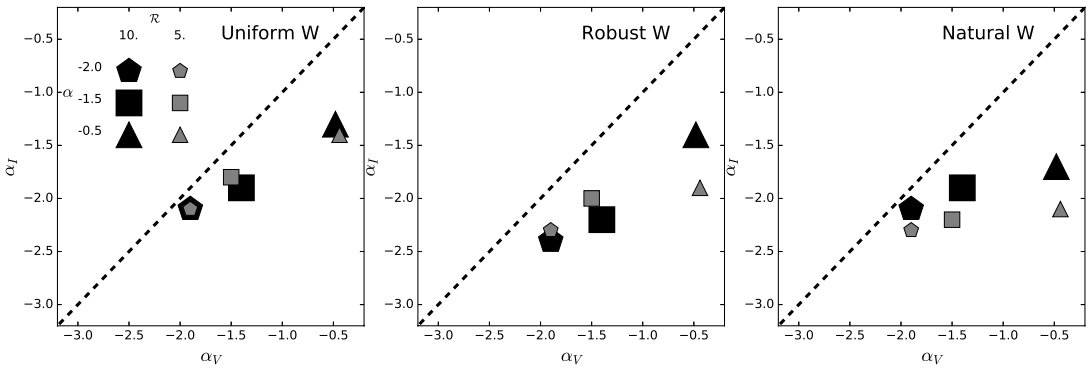


Fig. 10 The image and visibility based estimates of α are compared for different simulations and weighting schemes. In each panel, we plot the image based estimate α_I against the visibility based estimate α_V for all six simulations. The *dashed line* corresponds to an exact match. The meaning of different markers is given in the leftmost panel. Three panels in this figure correspond to three different weighting schemes, from left to right, giving uniform, robust and natural weightings respectively.

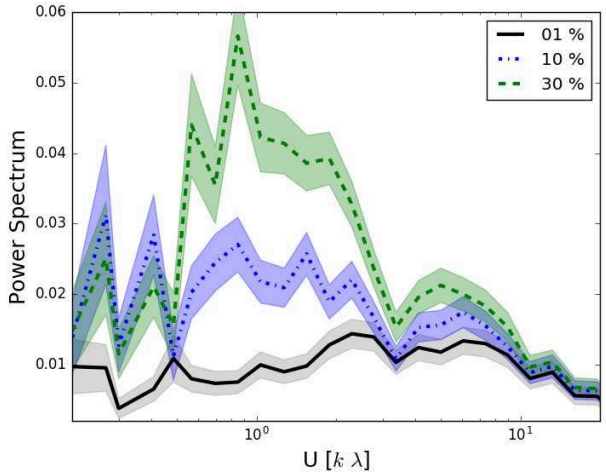
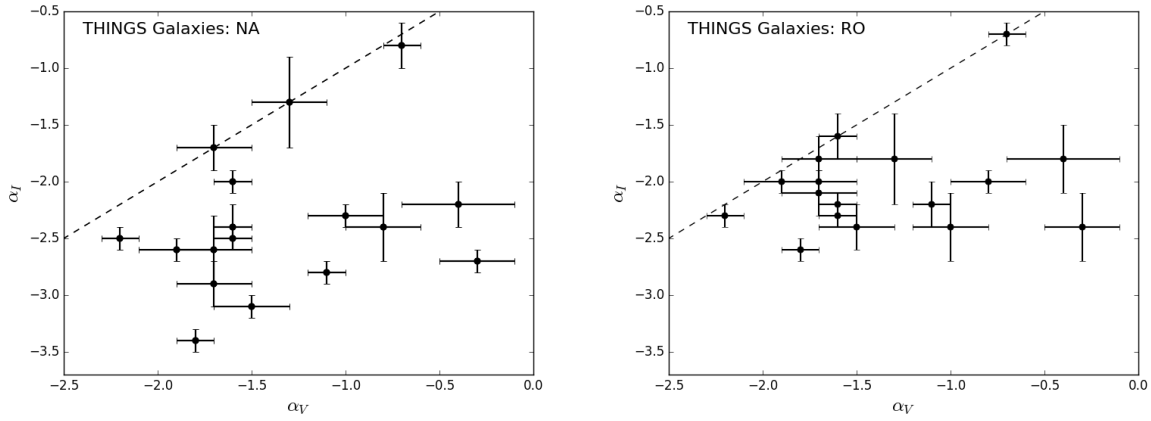


Fig. 11 The figure shows the power spectra calculated from the residual images after point source subtraction for the point sources influencing 1%, 10% and 30% of the FOV.

amplitudes 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 CLEAN must work well. On the other hand, if we increase the number of point sources, it would start 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 FWHM of the PSF. How much of the FOV 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 ROBUST=0). Using these model images, we simulate the visibilities and reconstruct images keeping the same baseline coverage as in 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 11. The shaded region of the cases indicates 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.



(a) (b)

Fig. 12 Comparison of the power law slopes calculated from the visibility and the image based power spectra for 18 galaxies in the THINGS sample using naturally weighted (NA in left) and robustly weighted (RO in right) images from the THINGS data archive. *Dashed lines* correspond to an exact match between the slopes.

5.4 Application to H I Power Spectrum of Galaxies

Walter et al. (2008) performed a survey of H I in a sample of 34 external galaxies using B, C and D configurations of the VLA. To reconstruct the images, they used the multi-scale version of the CLEAN algorithm with the natural and robust (with the ROBUST parameter set to 0.5) weighting schemes. The effective resolution of their reconstructed image with the robust weighting scheme is $\sim 6''$. Dutta et al. (2013) have estimated slope of the power spectrum (α_V) of the H I intensity fluctuation of 18 spiral galaxies from a THINGS sample using the same visibility based estimator as we used here. As our simulation suggests that the visibility based estimator estimates the true power spectrum of the specific intensity fluctuations, we implement the measured α_V as a proxy for the value of α for these galaxies. We applied the publicly available naturally and robustly weighted THINGS moment 0 maps of these 18 spiral galaxies to estimate the power spectrum using the image based estimator we have discussed here. These power spectra were fit well by power laws in a similar range of length scales as in the work of Dutta et al. (2013) and we estimate the corresponding power law slopes (α_I). We plot the values of α_V and α_I along with their error bars in the left and right panels of Figure 12 for the naturally and robustly weighted maps. The dashed line corresponds to $\alpha_V = \alpha_I$. As is clear for the majority of the galaxies, the data points lie away from the equality line for both cases of naturally and robustly weighted maps. The image based estimator systematically produces steeper spectra. As expected from our simulation result, biases in the robustly weighted maps are lower but still significant. Moreover, Figure 12 compares quite well with Figure 10, where a similar plot is made for the results of our simulation. As the THINGS archive does not provide any reconstructed moment 0 map estimated using the uniform weighting scheme, we do not show them here. Nevertheless, it is clear from our analysis that in general bias exists in the image based estimates of the power spectrum. Several authors, including but not limited to Zhang et al. (2012); Walker et al. (2014); Grisdale et al. (2017), have used the image based estimators to make inferences about the power spectrum of the sky brightness distribution and observe discrepancy with the visibility based estimators. We believe our investigation answers the reason for these discrepancies.

6 DISCUSSION AND CONCLUSIONS

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 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. In order to estimate the window function, it is essential to reconstruct the sky brightness distribution from the observed visibilities. On the other hand, one can either use the visibilities directly to estimate the power spectrum, or first estimate the brightness distribution from the visibil-

ities 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 the natural weighting scheme, the window function is reasonably reproduced. The visibility based estimator of the power spectrum reproduces the model power spectrum unbiasedly. Among 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.

It is clear that for an ideal interferometer with visibilities measured over the entire baseline-plane, that is with complete uv coverage, the sky brightness distribution can be estimated without any bias. In such a case the image based power spectrum estimators are expected to produce an unbiased result. Apparent reasons for the bias in the image based estimates of the power spectrum from a realistic interferometer could be the incompleteness and non-uniformity of the uv coverage of the interferometer. Different weighting schemes try to address these issues, however, as our result suggests, for the baseline coverage of the simulated visibilities used here, the weighting methods fail to reproduce the power spectrum of the model sky. Interestingly, the natural weighting gives the best estimate of the window function, while the power spectrum estimated from the image with uniform weighting provides the best approximation to the power spectrum. This can be understood in the following way. The natural weighting gives the same weight to all the measured visibilities in a grid. Since the baseline coverage of the interferometer is more complete at the shorter baselines and falls as the baseline increases, effective natural weighting produces a larger synthesized beam. Hence, it is expected that natural weighting would produce a better approximation to the large-scale distribution, while it models the small scale distribution poorly. The effect then is also a redistribution of the flux across different angular scales. Thereby, the power spectrum at large scales, that is at small baselines, is enhanced, whereas the power spectrum at large baselines is reduced. Effectively, the bias parameter $P_{\mathcal{B}}(\boldsymbol{U})$, as discussed in Equation (11), assumes positive values at smaller baselines and is negative at larger baselines. Uniform weighting, on the other hand, produces a smaller restoring beam and hence reproduces relatively more power at the longer baselines compared to the natural weighting. This is observed in the power spectrum estimates. We note here that we have not addressed the relationship between the incomplete baseline coverage and the observed bias in the image based estimators. This is not in the scope of the paper and requires further and detailed investigation.

In the case of large bandwidth observations with smooth specific intensity across frequencies, a method termed multifrequency synthesis can be effectively applied which drastically improves the effective baseline coverage and hence reduces the CLEAN artifacts. In this work, we model our sky brightness distribution based on H I emission from external spiral galaxies. In such a case, the specific intensity has structures as a function of frequency and hence the method of multifrequency synthesis cannot be used. In continuum observations with large bandwidth, however, multifrequency synthesis (Sault & Conway 1999; Bajkova 2008; Junklewitz et al. 2015) is possible and, hence, the image based power spectrum may be a useful tool. We have also restricted our analysis to relatively small FOV such that the effect of the w-terms is negligible. For a larger FOV, the w-term is known to restrict the baseline range over which the visibility based power spectrum can be used (Dutta et al. 2010b). Wakker & Schwarz (1988); Cornwell (2008) have introduced the method of multi-scale CLEAN for image reconstruction in the case of extended objects. Multiscale CLEAN tries to first find the large-scale structures from the image and then proceeds with smaller scales. While this has been successful and has been used in many cases (e.g. Greisen et al. 2009), the subjective choice of the scales at which CLEAN proceeds remains an open question. An advanced version of multi-scale CLEAN has been reported in Zhang et al. (2016), where adaptive methods are used in choosing the scales and loop gains. We note that this choice of scales may introduce artificial scale dependence in the power spectrum. Hence, we have not considered these algorithms in this work for power spectrum analysis.

In this work, we have not included any thermal noise in our simulation. We justified it by considering the fact that for any experiment the thermal noise can be decreased arbitrarily by increasing the integration time. Each visibility measurement has noise associated with it as shown in Equation (13). Estimating the power spectrum by directly squaring the visibilities gives rise to a noise bias in the visibility correlation. Since the measurement noise arises due to electronics associated with different antennas, the noise at different baselines is not correlated. Hence, the noise bias in the power spectrum is expected to be independent of baseline. The visibility based estimator of the power spectrum we use here, however, does not correlate the visibilities in the same baseline, but at nearby baselines. This

effectively reduces the noise bias from the power spectrum estimates. The effect of measurement in the image based power spectrum estimator depends on the baseline coverage. Analytical calculation shows that uncorrelated thermal noise in the visibility measurements with incomplete baseline coverage appears as correlated noise in the image plane and adds to an additional bias in the image based estimates of the power spectrum on top of what is discussed here. The effect of noise on visibility and image based estimators of the power spectrum are discussed in detail in Dutta (2011); interested readers may have a look.

As the radio astronomy community plans for larger interferometers, the volume of visibility data is expected to grow progressively larger. One way of reducing the problem of 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 or estimating the power spectrum and recording. A visibility based power spectrum as discussed in this work has limited use when the power spectrum of a selected part of the telescope's FOV is required to be estimated. Choudhuri et al. (2016) has developed a visibility based Tapered Gridded Estimator (TGE) that uses a tapered window to reduce the response of the sky outside it. We have worked on modifications of the TGE limiting the tapering function to the required part of the FOV of interest and selectively estimating the power spectrum using the visibility based method; the results will be reported in a separate paper.

Acknowledgements PD would like to acknowledge numerous useful discussions with Wasim Raja and Aritra Basu. PD would also like to acknowledge Apurba Bera, Barnali Das, Saheb Ghosh and Surajit Mandal for providing gracious hospitality to finish writing the paper. PD acknowledges support of the DST-INSPIRE faculty fellowship from DST, India for this work. MN acknowledges the DST-INSPIRE fellowship for funding this work.

References

Bagetakos, I., Brinks, E., Walter, F., et al. 2011, AJ, 141, 23
Bajkova, A. T. 2008, Astronomy Reports, 52, 951
Basu, A., & Roy, S. 2013, MNRAS, 433, 1675
Becker, R. H., White, R. L., & Helfand, D. J. 1995, APJ, 450, 559
Begum, A., & Chengalur, J. N. 2003, AAP, 409, 879
Begum, A., Chengalur, J. N., & Bhardwaj, S. 2006, MNRAS, 372, L33
Begum, A., Chengalur, J. N., Karachentsev, I. D., Sharina, M. E., & Kaisin, S. S. 2008, MNRAS, 386, 1667
Bharadwaj, S., & Sethi, S. K. 2001, Journal of Astrophysics and Astronomy, 22, 293
Briggs, D. S. 1995, in Bulletin of the American Astronomical Society, 27, American Astronomical Society Meeting Abstracts, 1444
Choudhuri, S., Bharadwaj, S., Chatterjee, S., et al. 2016, MNRAS, 463, 4093
Clark, B. G. 1980, AAP, 89, 377
Condon, J. J., Cotton, W. D., Greisen, E. W., et al. 1998, AJ, 115, 1693
Cornwell, T., Braun, R., & Briggs, D. S. 1999, in Astronomical Society of the Pacific Conference Series, 180, Synthesis Imaging in Radio Astronomy II, eds. G. B. Taylor, C. L. Carilli, & R. A. Perley, 151
Cornwell, T. J. 2008, IEEE Journal of Selected Topics in Signal Processing, 2, 793
Cotton, W. D. 1979, AJ, 84, 1122
Crovisier, J., & Dickey, J. M. 1983, AAP, 122, 282
Dutta, P. 2011, arXiv:1102.4419
Dutta, P., Begum, A., Bharadwaj, S., & Chengalur, J. N. 2009, MNRAS, 398, 887
Dutta, P., Begum, A., Bharadwaj, S., & Chengalur, J. N. 2010a, MNRAS, 405, L102
Dutta, P., Guha Sarkar, T., & Khastgir, S. P. 2010b, MNRAS, 406, L30
Dutta, P., Begum, A., Bharadwaj, S., & Chengalur, J. N. 2013, NewAstro, 19, 89
Dutta, P., & Bharadwaj, S. 2013, MNRAS, 436, L49
Elmegreen, B. G., Kim, S., & Staveley-Smith, L. 2001, APJ, 548, 749
Elmegreen, B. G., & Scalo, J. 2004, ARAA, 42, 211
Ghosh, A., Bharadwaj, S., Ali, S. S., & Chengalur, J. N. 2011, MNRAS, 411, 2426
Green, D. A. 1993, MNRAS, 262, 327
Greisen, E. W., Spekkens, K., & van Moorsel, G. A. 2009, AJ, 137, 4718
Grisdale, K., Agertz, O., Romeo, A. B., Renaud, F., & Read, J. I. 2017, MNRAS, 466, 1093
Högbom, J. A. 1974, AAPS, 15, 417
Junklewitz, H., Bell, M. R., & Enßlin, T. 2015, AAP, 581, A59
Junklewitz, H., Bell, M. R., Selig, M., & Enßlin, T. A. 2016, AAP, 586, A76
Mahesh, N., Subrahmanyan, R., Udaya Shankar, N., & Raghunathan, A. 2014, arXiv:1406.2585
Nandi, S., & Saikia, D. J. 2012, Bulletin of the Astronomical Society of India, 40, 121
Narayan, R., & Nityananda, R. 1986, ARAA, 24, 127
Refregier, A. 2003, MNRAS, 338, 35
Roy, N., Bharadwaj, S., Dutta, P., & Chengalur, J. N. 2009, MNRAS, 393, L26
Rybicki, G. B., & Lightman, A. P. 1985, Radiative Processes in

Astrophysics (John Wiley & Sons)
Sault, R. J., & Conway, J. E. 1999, in Astronomical Society of the Pacific Conference Series, 180, Synthesis Imaging in Radio Astronomy II, eds. G. B. Taylor, C. L. Carilli, & R. A. Perley, 419
Schwab, F. R. 1984, in Indirect Imaging. Measurement and Processing for Indirect Imaging, ed. J. A. Roberts, 333
Schwarz, U. J. 1978, AAP, 65, 345
Schwarz, U. J. 1979, in Astrophysics and Space Science Library, 76, IAU Colloq. 49: Image Formation from Coherence Functions in Astronomy, ed. C. van Schooneveld, 261
Sukumar, S., & Allen, R. J. 1989, Nature, 340, 537
Taylor, G. B., Carilli, C. L., & Perley, R. A., eds. 1999, Astronomical Society of the Pacific Conference Series, 180, Synthesis Imaging in Radio Astronomy II
Thompson, A. R., Moran, J. M., & Swenson, Jr., G. W. 2017, Interferometry and Synthesis in Radio Astronomy (3rd edn.; Springer)
Wakker, B. P., & Schwarz, U. J. 1988, AAP, 200, 312
Walker, A. P., Gibson, B. K., Pilkington, K., et al. 2014, MNRAS, 441, 525
Walter, F., Brinks, E., de Blok, W. J. G., et al. 2008, AJ, 136, 2563
Wang, J., Fu, J., Aumer, M., et al. 2014, MNRAS, 441, 2159
White, R. L., Helfand, D. J., Becker, R. H., Glikman, E., & de Vries, W. 2007, APJ, 654, 99
Yatawatta, S. 2014, MNRAS, 444, 790
Zhang, H.-X., Hunter, D. A., & Elmegreen, B. G. 2012, APJ, 754, 29
Zhang, L., Zhang, M., & Liu, X. 2016, APSS, 361, 153