

## Status and perspectives of the CRAFTS extra-galactic HI survey

Kai Zhang<sup>1,2,3\*</sup>, JingWen Wu<sup>1,2,3\*</sup>, Di Li<sup>1,2,3\*</sup>, Marko Krčo<sup>1,3</sup>, Lister Staveley-Smith<sup>4,5</sup>,  
NingYu Tang<sup>1,3</sup>, Lei Qian<sup>1,3</sup>, MengTing Liu<sup>1,2,3</sup>, ChengJin Jin<sup>1,3</sup>, YouLing Yue<sup>1,3</sup>, Yan Zhu<sup>1,3</sup>,  
HongFei Liu<sup>1,3</sup>, DongJun Yu<sup>1,3</sup>, JingHai Sun<sup>1,3</sup>, GaoFeng Pan<sup>1,3</sup>, Hui Li<sup>1,3</sup>, HengQian Gan<sup>1,3</sup>,  
Rui Yao<sup>1,3</sup>, and FAST Collaboration

<sup>1</sup>National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100101, China;

<sup>2</sup>University of Chinese Academy of Sciences, Beijing 100049, China;

<sup>3</sup>CAS Key Laboratory of FAST, National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100101, China;

<sup>4</sup>International Centre for Radio Astronomy Research (ICRAR), The University of Western Australia, Perth 6907, Australia;

<sup>5</sup>ARC Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), The University of Western Australia, Perth 6907, Australia

Received January 28, 2019; accepted March 4, 2019; published online March 21, 2019

The Five-hundred-meter Aperture Spherical radio Telescope (FAST) is expected to complete its commissioning in 2019. FAST will soon begin the Commensal Radio Astronomy FasT Survey (CRAFTS), a novel and unprecedented commensal drift scan survey of the entire sky visible from FAST. The goal of CRAFTS is to cover more than 20000 deg<sup>2</sup> and reach redshift up to about 0.35. We provide empirical measurements of the beam size and sensitivity of FAST across the 1.05 to 1.45 GHz frequency range of the FAST L-band array of 19-beams (FLAN). Using a simulated HI-galaxy catalogue based on the HI Mass Function (HIMF), we estimate the number of galaxies that CRAFTS may detect. At redshifts below 0.35, over  $6 \times 10^5$  HI galaxies may be detected. Below the redshift of 0.07, the CRAFTS HIMF will be complete above a mass threshold of  $10^{9.5} M_{\odot}$ . FAST will be able to investigate the environmental and redshift dependence of the HIMF to an unprecedented depth, shedding light onto the missing baryon and missing satellite problems.

### HI regions and 21-cm lines, radio telescopes and instrumentation, statistical and correlative studies of properties

PACS number(s): 47.55.nb, 47.20.Ky, 47.11.Fg

**Citation:** K. Zhang, J. W. Wu, D. Li, M. Krčo, L. Staveley-Smith, N. Y. Tang, L. Qian, M. T. Liu, C. J. Jin, Y. L. Yue, Y. Zhu, H. F. Liu, D. J. Yu, J. H. Sun, G. F. Pan, H. Li, H. Q. Gan, R. Yao, and FAST Collaboration, Status and perspectives of the CRAFTS extra-galactic HI survey, *Sci. China-Phys. Mech. Astron.* **62**, 959506 (2019), <https://doi.org/10.1007/s11433-019-9383-y>

## 1 Introduction

HI, or neutral hydrogen is the most abundant species in the universe. Observations of the HI 21 cm line can reveal a variety of information on galaxies, including their redshift, HI mass, central velocity, dynamical mass, and so on [1]. Wide-field blind extragalactic HI sky surveys can be implemented

owing to the advent of multi-beam receivers which greatly improve the efficiency and allow the surveys to cover cosmologically significant volumes. From the HIPASS (HI Parkes All Sky Survey [2, 3]), to the recently completed ALFALFA (Arecibo Legacy Fast ALFA [4-7]) survey, blind surveys have played a key role in exploring the HI distribution and properties of nearby galaxies.

The HI mass function (HIMF) depicts the cosmic number density per bin of HI mass [8] and is found to be well fitted

\*Corresponding authors (Kai Zhang, email: [zk3kw2n@nao.cas.cn](mailto:zk3kw2n@nao.cas.cn); JingWen Wu, email: [jingwen@nao.cas.cn](mailto:jingwen@nao.cas.cn); Di Li, email: [dili@nao.cas.cn](mailto:dili@nao.cas.cn))

by the Schechter function [9]. Faint-end slopes of the HIMF are closely related to the “missing satellite problem” [10, 11]. At high redshift ( $z > 1.5$ ), the total mass density of HI, or  $\Omega_{\text{HI}}$ , can be deduced from the absorption lines of quasar optical spectra through damped-Ly $\alpha$  objects [12, 13], while at lower redshifts  $\Omega_{\text{HI}}$  can be evaluated by the HIMF [14–16]. The HIMF is very important in understanding the galactic HI evolution as a function of redshift, and serves as a test of theoretical cosmological simulations [17, 18]. With large sampled populations and spatial volumes blind HI surveys are ideal for measuring the HIMF and for studying the environmental dependence of the HIMF [14, 16, 19, 20]. FAST and ASKAP (The Australian Square Kilometre Array Pathfinder) will carry out the next generation of HI blind sky surveys [21, 22], to enable the study of the evolution of the HIMF and  $\Omega_{\text{HI}}$  up to higher redshifts.

As the largest filled-aperture single-dish telescope in the world, FAST [23, 24] is expected to finish its commissioning in 2019 [25, 26]. The Commensal Radio Astronomy FasT Survey (CRAFTS) [25] will simultaneously conduct an extragalactic HI, galactic HI imaging, pulsar search, and Fast Radio Burst (FRB) search surveys. The predictions [21] show FAST would be an excellent instrument for large-scale HI surveys. Based on test observations taken during FAST commissioning, we use updated parameters of FAST and the latest HIMF from ALFALFA survey to predict the capability of CRAFTS to detect HI galaxies.

We briefly introduce the CRAFTS survey plan in sect. 2, and summarize several key parameters of FAST in sect. 3. The sensitivity of CRAFTS with regards to the extragalactic HI survey is discussed in sect. 4. The number of HI galaxies CRAFTS may detect is discussed in sect. 5. We will discuss the impact of confusion on CRAFTS in sect. 6, and a summary is provided in sect. 7. The cosmological distances calculation and cosmological corrections in this paper refer to refs. [27, 28], and we assume  $H_0$  (Hubble constant) =  $73 \text{ km s}^{-1} \text{ Mpc}^{-1}$ ,  $\Omega_m$  (density parameter of matter) = 0.25 and  $\Omega_\Lambda$  (density parameter of dark energy) = 0.75.

## 2 The CRAFTS survey plan

When inactive, FAST’s surface is a partial sphere composed of over 4400 panels connected to over 2250 actuators with an aperture of 500 m. While observing, the actuators behind the reflector deform the illuminated surface into a 300 m-aperture paraboloid, with an rms (root mean square) uncertainty of  $\sim 4 \text{ mm}$  which is sufficient for observations at frequencies up to 5 GHz [23]. In the case of FAST, the most efficient way to conduct a large commensal survey is to use drift scans which have the added benefit of minimizing gain fluctuations due

to variations in the surface and feed position. CRAFTS will observe galactic HI, extra-galactic HI, pulsars, and FRBs simultaneously while drifting. CRAFTS will conduct two full passes surveys similarly to ALFALFA [4], which should help alleviate the influence of Radio Frequency Interference (RFI) and make low SNR (Signal-to-Noise Ratio) detections more reliable. It will take CRAFTS approximately 220 full days of observation to complete each pass [25].

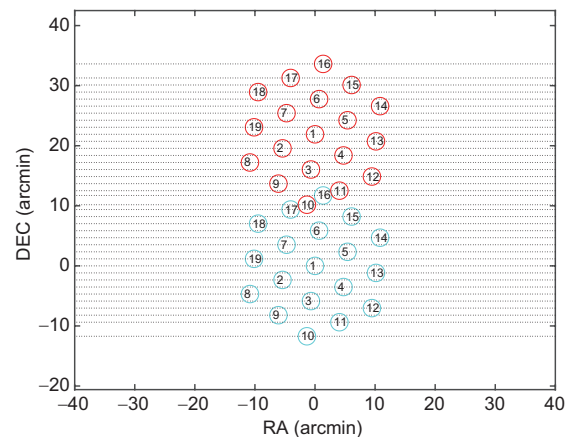
While drifting, the 19-beam receiver (the FAST L-band Array of 19-beam, or FLAN), will be rotated by  $23.4^\circ$  to achieve a super-Nyquist sampling. The spacing of each scan would be  $21.9'$  in declination to fill the gap of two outer beams so that the survey would be nearly uniformly covered as described in ref. [25]. Figure 1 shows a sketch of two adjacent drift scans of FLAN using this orientation.

## 3 Key FAST parameters

FAST is located at a latitude of  $25.6529^\circ\text{N}$ . The effective illuminated aperture size is about 300 m up to a zenith angle of  $26.4^\circ$ , and can be partially illuminated up to FAST’s maximum zenith angle of  $40^\circ$ . Thus the total sky visible from FAST is approximately  $20000 \text{ deg}^2$  within a declination range between  $-14^\circ$  and  $66^\circ$ .

The FLAN bandwidth (1.05 to 1.45 GHz) corresponds to a maximum redshift of 0.35 for the 21 cm HI line. The channel width of FLAN for the extragalactic HI survey is about 7.6 kHz, corresponding to a velocity resolution of 1.6 km/s for HI.

We use FAST commissioning data to estimate its beam-size, gain and system temperature. Much of the data are taken from ref. [26] and the parameters are all obtained from center beam of the FLAN.



**Figure 1** (Color online) An example sketch of two adjacent drift scans of FLAN using the orientation used by CRAFTS. Blue and red circles with a diameter of 2.9 arcmin represent the position and the size of beams in two drift scans. The dotted lines show the drifting tracks of individual beams.

Due to the wide bandwidth of FLAN and the limited feed cross-section size, the electric field of the aperture antenna tends to be uniformly distributed at low frequencies. While at high frequencies, the distribution function of electric field that illuminates the aperture resembles a Gaussian function. Therefore the FLAN beam size cannot be simply estimated by assuming it is inversely proportional to frequency; at low frequencies the beamsize is smaller than otherwise predicted. The FLAN FWHM beamsize of center beam can be accurately estimated by

$$\left(\frac{\theta}{\prime}\right) = \left[\frac{m1(ZA)}{\prime}\right] \left(\frac{\text{Freq}/1250}{\text{MHz}}\right)^{-1} + \left(\frac{m2}{\prime}\right) \left[\left(\frac{\text{Freq}}{\text{MHz}}\right) - 1250\right], \quad (1)$$

where  $\theta$  is beamsize in arcmin, Freq is frequency in MHz,  $m1$  is beamsize at the frequency of 1250 MHz in arcmin and  $m2$  is the correction factor in arcmin. From observations,  $m1$  is approximately 3.24 arcmin and  $m2$  is approximately  $5.475 \times 10^{-4}$  arcmin when zenith angle is below  $26.4^\circ$ . Figure 2(a) shows how beamsize varies with frequency when zenith angle is below  $26.4^\circ$ .

For zenith angles above  $26.4^\circ$  the aperture is only partially illuminated and thus the beamsize will increase. When zenith angle increases, the illuminated aperture size will decrease,

we assume that the illuminated aperture size decreases linearly when zenith angle increases and thus  $m1$  is proportional to zenith angle. For  $m2$  we assume it does not change. Figure 2(b) shows the relation between  $m1$  and zenith angle. Thus when zenith angle is above  $26.4^\circ$ ,  $m1$  can be approximated by

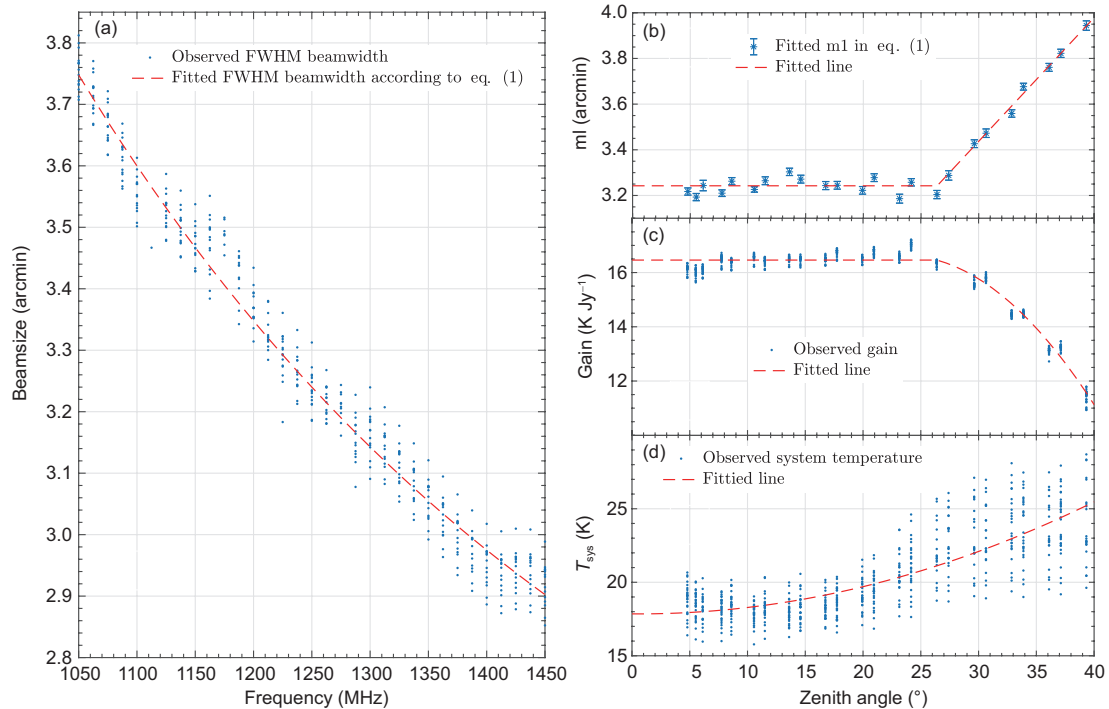
$$\left(\frac{m1}{\prime}\right) = 3.24 + 5.41 \times 10^{-2} \left[\left(\frac{ZA}{^\circ}\right) - 26.4\right], \quad (2)$$

where  $ZA$  is zenith angle in degree.

The gain is about 16.46 K/Jy for the center beam when zenith angle is below  $26.4^\circ$  and is largely independent of frequency over the bandwidth covered by FLAN. Gain decreases for zenith angles greater than  $26.4^\circ$  due to the surface being only partially illuminated as shown in Figure 2(c). When zenith angle is above  $26.4^\circ$ , the gain may be estimated by a second degree polynomial such as:

$$\left(\frac{G}{\text{K/Jy}}\right) = 16.46 - 0.02 \left[\left(\frac{ZA}{^\circ}\right) - 26.4\right]^2 - 0.12 \left[\left(\frac{ZA}{^\circ}\right) - 26.4\right], \quad (3)$$

where  $G$  is the gain in K/Jy and  $ZA$  is zenith angle in degree. At a zenith angle of  $40^\circ$ , the gain will decrease to about 11.0 K/Jy, with a loss of approximately 33%.



**Figure 2** (Color online) (a) Beamsize versus frequency when zenith angle is below  $26.4^\circ$ , the red dashed line represents the fitted result from eq. (1). (b) The blue asterisks represent the fitted result of  $m1$  by eq. (1) at different zenith angles, which tells the beamsize at the frequency of 1250 MHz, while the red dashed line represents the fitted line from eq. (2). (c) The gain versus zenith angle, the red dashed line represents the fitted result from eq. (3). (d) System temperature versus zenith angle, the red dashed line represents the fitted result from eq. (4). The parameters here are all obtained from center beam of the 19-beam receiver (FLAN).

Figure 2(d) shows the relation between system temperature and zenith angle. We use a second-order polynomial function to estimate the relation between system temperature and zenith angle with

$$\left(\frac{T_{\text{sys}}}{\text{K}}\right) = 4.9 \times 10^{-3} \left(\frac{\text{ZA}}{\circ}\right)^2 + 5.5 \times 10^{-3} \left(\frac{\text{ZA}}{\circ}\right) + 17.85, \quad (4)$$

where  $T_{\text{sys}}$  is system temperature in K, and ZA is zenith angle in degree. At zenith angles above  $26.4^\circ$  the feed is tilted back slightly so as to avoid spill over. However, radiation from the surrounding mountain peaks enters the nearer sidelobes and increases the system temperature. A ground screen made from wire mesh was proposed to reduce  $T_{\text{sys}}$  for FAST when the zenith angle is large, which is designated as the ‘‘backward illumination’’ mode [29].

Figure 3 shows an example spectrum of AGC11820 taken by FAST on 2018/10/22.

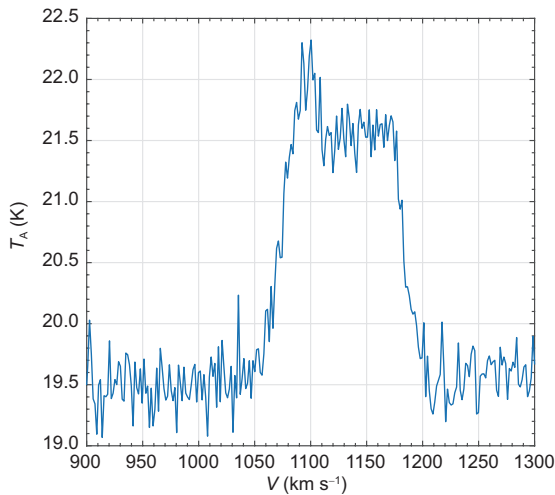
## 4 Sensitivity calculation

### 4.1 Effective integration time

The effective integration time of a drift scan survey reflects the equivalent rms noise by combining data from multiple observations, which may be estimated by

$$t_{\text{obs}}/t_{\text{eff}} = \Omega_{\text{obs}}/\Omega_{\text{b}}, \quad (5)$$

where  $t_{\text{eff}}$  is the effective integration time for each point in the sky within the observed region,  $t_{\text{obs}}$  is the total observation time,  $\Omega_{\text{obs}}$  is the total solid angle of observation, and  $\Omega_{\text{b}}$  is the solid angle of the beam. We assume all 19 beams of FLAN have same beam pattern, which can be estimated as an

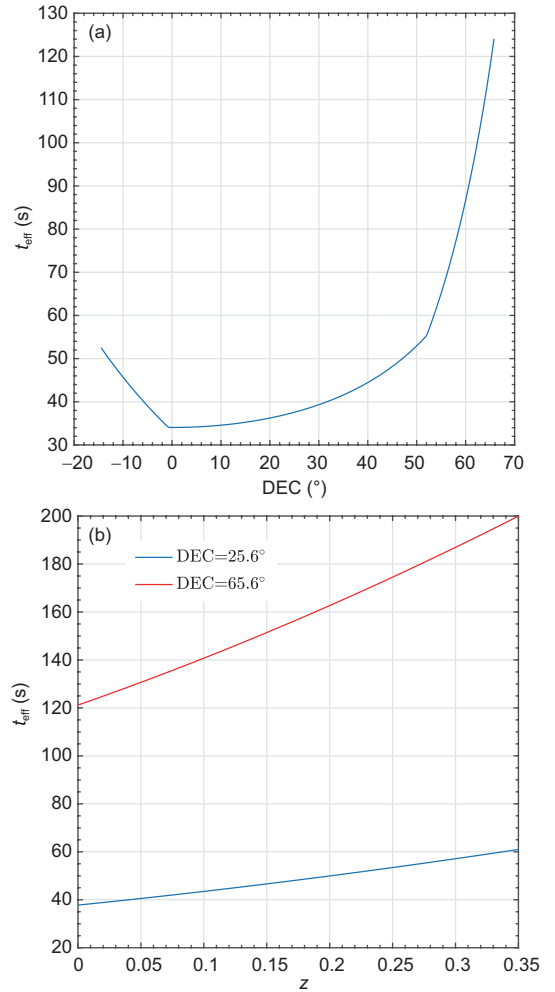


**Figure 3** (Color online) A spectrum of AGC11820 taken by FAST on 2018/10/22. The horizontal axis represents radial velocity and the vertical axis represents antenna temperature.

axisymmetric Gaussian function, so  $\Omega_{\text{b}} = 19 \times 1.13 \times \theta^2$ , where  $\theta$  is the beamsize. If we assume the drifting speed at certain declination is constant and at the equator the drifting speed reaches its maximum, which is roughly  $0.25 \text{ arcmin/s}$ , thus the drifting speed of survey  $v_s/\text{arcmin s}^{-1} = 0.25 \times \cos(\text{DEC})$ , where DEC is the declination of the center beam of FLAN while drifting.  $\Omega_{\text{obs}}$  can be represented by  $\delta v_s t_{\text{obs}}$ , where  $\delta = 21.9'$  is the spacing in declination between two adjacent drift scans. Thus  $t_{\text{eff}}$  can be estimated by

$$\left(\frac{t_{\text{eff}}}{s}\right) = \frac{19 \times 1.13 \times \left(\frac{\theta}{\text{r}}\right)^2}{\left(\frac{0.25}{\text{r/s}}\right) \times \cos\left(\frac{\text{DEC}}{\circ}\right) \left(\frac{\delta}{\text{r}}\right)}. \quad (6)$$

If we overlook the peculiar motion of the earth and HI source, the frequency of HI line we observed is:  $\nu_{\text{obs}} = \nu_{\text{HI}}/(1+z)$ , where  $\nu_{\text{HI}} = 1420.4 \text{ MHz}$  is the rest-frame HI frequency, and  $z$  is the redshift of the source. Figure 4(a) shows the effective integration time after one pass as a function of



**Figure 4** (Color online) (a) shows  $t_{\text{eff}}$  versus DEC at redshift of 0 after one pass; (b) shows  $t_{\text{eff}}$  after one pass as a function of  $z$  at DEC of  $25.6^\circ$  and  $65.6^\circ$  in blue and red solid line respectively.

DEC at redshift of 0. The main factors that influence  $t_{\text{eff}}$  are the beamsize and  $1/\cos(\text{DEC})$ . When DEC ranges between  $-1.2^\circ$  and  $52^\circ$ ,  $\theta$  does not vary with DEC, so  $t_{\text{eff}}$  will change with  $1/\cos(\text{DEC})$ . When DEC is below  $-1.2^\circ$  or above  $52^\circ$ , the beamsize would increase linearly with ZA, so the growth rate of  $t_{\text{eff}}$  will become higher at higher zenith angle. Figure 4(b) shows how  $t_{\text{eff}}$  varies with  $z$  at two typical DEC's after one pass.

Figure 5 shows the effective integration time map of one-pass of CRAFTS in  $z$ -DEC plane based on eq. (6). The map shows that in most CRAFTS regions,  $t_{\text{eff}}$  is from 35 to 60 s, while at high zenith angle,  $t_{\text{eff}}$  will increase rapidly, especially at high declination regions, where the drifting speed is slower thus the duration time of source in beam will increase. CRAFTS plans to carry out a two-pass survey, so the ultimate effective integration time is twice the value in Figure 5.

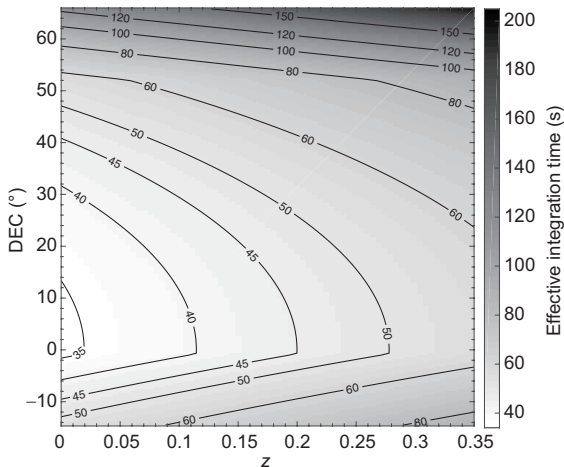
## 4.2 Sensitivity per channel

The rms noise per channel or  $S_{\text{rms}}$ , can be estimated by [6]

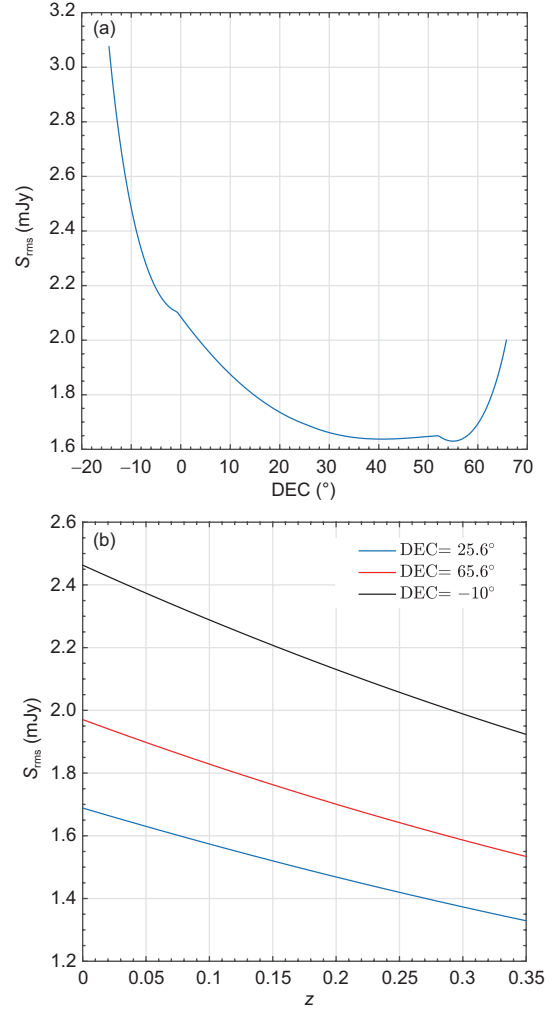
$$S_{\text{rms}} = \frac{T_{\text{sys}}}{G} \frac{1}{\sqrt{2f_t \Delta f_{\text{ch}} t_s}}, \quad (7)$$

where  $T_{\text{sys}}$  is system temperature in K,  $G$  is gain of the telescope in K/Jy,  $\Delta f_{\text{ch}}$  is single channel width in Hz and  $t_s$  is integration time in s, here we can use the effective integration time we calculated before. The factor  $f_t$  accounts for observation details like bandpass subtraction method, spectral smoothing and so on. For ALFALFA,  $f_t \approx 0.7$ , here we adopt this value for CRAFTS extragalactic HI survey.

Figure 6(a) shows  $S_{\text{rms}}$  as a function of DEC at redshift of 0. The main factor that influences  $S_{\text{rms}}$  at ZA below  $25.6^\circ$  is the effective integration time:  $S_{\text{rms}}$  will decrease along the increase of DEC. However, at high ZA, the increase of gain



**Figure 5** The effective integration time map of one-pass of CRAFTS in  $z$ -DEC plane. The pixel size is  $21.9'$  in declination and 0.005 in redshift.



**Figure 6** (Color online) (a) shows  $S_{\text{rms}}$  as a function of DEC at redshift of 0 after one pass; (b) shows  $S_{\text{rms}}$  after one pass versus  $z$  at DEC of  $25.6^\circ$ ,  $65.6^\circ$ ,  $-10^\circ$  in blue, red and black solid line, respectively.

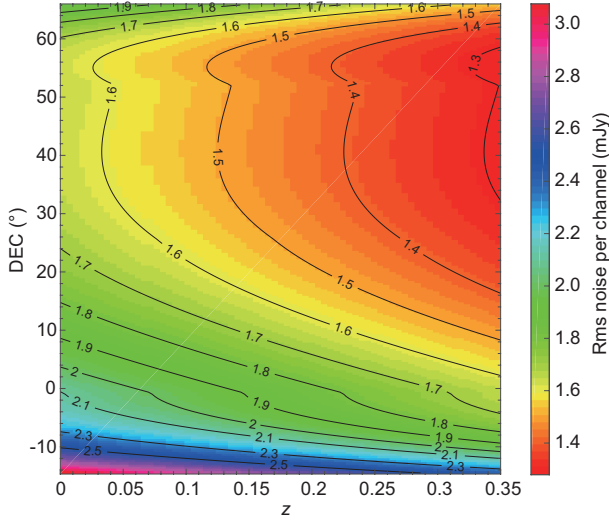
and  $T_{\text{sys}}$  will cause  $S_{\text{rms}}$  increase at a relatively higher rate. Figure 6(b) shows  $S_{\text{rms}}$  varies as a function of  $z$  at three typical DEC's.

$S_{\text{rms}}$  of one-pass of CRAFTS in  $z$ -DEC plane is shown in Figure 7. Figure 7 implies that  $S_{\text{rms}}$  will not fluctuate rapidly with DEC when DEC is above  $0^\circ$ . At higher redshift, the difference of  $S_{\text{rms}}$  between two adjacent drift scans will become larger. At DEC below  $10^\circ$ ,  $S_{\text{rms}}$  will increase rapidly when DEC decreases, which is caused by relatively low  $t_{\text{eff}}$ , high  $T_{\text{sys}}$  and low gain. For two-pass of CRAFTS,  $S_{\text{rms}}$  can be obtained by dividing the value in Figure 7 by square root of 2.

## 4.3 Flux limit for HI detection

The velocity line width of HI signal at source rest frame or  $\Delta V_{\text{rest}}$ , can be estimated by [28]





**Figure 7** (Color online) The rms noise per channel map for CRAFTS drifting one-pass in  $z$ -DEC plane, with same pixel size as Figure 5.

$$\Delta V_{\text{rest}} \simeq \frac{c(1+z)}{v_{\text{HI}}} \Delta v_{\text{obs}}, \quad (8)$$

where  $\Delta v_{\text{obs}}$  is the frequency line width of signal we observe,  $c$  is the speed of light and  $z$  is the redshift of the source. Here we assume the peculiar motion speed and line width of the source are much less than  $c$ . The relation between line width in source's rest frame,  $\Delta V_{\text{rest}}$ , and in the observed frame,  $\Delta V_{\text{obs}}$ , can be donated by  $\Delta V_{\text{rest}} = \Delta V_{\text{obs}} / (1+z)$ . Thus the corresponding observed velocity width in a single channel with channel width of  $\Delta f_{\text{ch}}$  while observing an HI source at redshift of  $z$ , can be estimated by

$$\Delta V_{\text{ch}} \simeq \frac{c(1+z)^2}{v_{\text{HI}}} \Delta f_{\text{ch}}. \quad (9)$$

The SNR of a signal can be estimated by [7]

$$\text{SNR} = \left( \frac{S^{V_{\text{obs}}}}{\text{Jy km s}^{-1}} \right) \left( \frac{\Delta V_{\text{obs}}}{\text{km s}^{-1}} \right)^{-1} f_{\text{smo}}^{1/2} \left( \frac{S_{\text{rms}}}{\text{Jy}} \right)^{-1}, \quad (10)$$

where  $S^{V_{\text{obs}}}$  is the integrated flux density of HI line in observed frame, we can make an assumption that HI line profile is a top-hat function, so  $S^{V_{\text{obs}}} = S_{\text{peak}} \Delta V_{\text{obs}}$ , where  $S_{\text{peak}}$  is peak flux of HI line and  $\Delta V_{\text{obs}}$  is the line width of HI line in observed frame. Usually  $\Delta V_{\text{obs}}$  can be described by  $W_{50}$ , which is measured at 50% level of each of two peaks.  $f_{\text{smo}}$  is the number of independent channels that signals can be smoothed over. For very broad HI signal profiles, the flux at two horns is much larger than that at center part of profiles, so direct smoothing will give diminishing returns [30]. We adopt the form of  $f_{\text{smo}}$  in ref. [30], which used the transition of completeness function of ALFALFA survey as the cutoff of  $f_{\text{smo}}$  and smoothed the signal to full line width to calculate SNR.  $f_{\text{smo}}$  is adopted as  $\Delta V_{\text{obs}} / \Delta V_{\text{ch}}$  while  $\Delta V_{\text{obs}} <$

$10^{2.5} \text{ km s}^{-1}$  and  $10^{2.5} / \Delta V_{\text{ch}}$  while  $\Delta V_{\text{obs}} \geq 10^{2.5} \text{ km s}^{-1}$ , where  $\Delta V_{\text{ch}}$  is the corresponding observed HI line velocity width in a single channel.

The flux limit of the telescope while observing HI galaxies at a given SNR can be denoted by

$$S_{\text{lim}} = \text{SNR} \cdot \frac{S_{\text{rms}}}{f_{\text{smo}}^{1/2}}. \quad (11)$$

If the peak flux of an HI galaxy signal is above  $S_{\text{lim}}$ , we can say the telescope could detect that galaxy.

## 5 Estimating the HI galaxy count

The HIMF tells us the relation between HI mass and number density of HI galaxy in our universe, which can be well fitted by Schechter function [9]. The HIMF could be expressed in the form of

$$\phi(M_{\text{HI}}) = \frac{dN_{\text{gal}}}{dV d\log_{10}(M_{\text{HI}})}, \quad (12)$$

$$= \ln(10) \phi_* \left( \frac{M_{\text{HI}}}{M_*} \right)^{\alpha+1} e^{-\left( \frac{M_{\text{HI}}}{M_*} \right)}, \quad (13)$$

where  $dN_{\text{gal}}$  is the average number of galaxies in comoving volume element  $dV$ ,  $\phi_*$  is the normalization constant,  $M_*$  is the ‘‘knee mass’’ and  $(\alpha + 1)$  is the low-mass slope, which is usually referred as the ‘‘faint end’’ of HIMF.

The HI mass of a galaxy,  $M_{\text{HI}}$ , at redshift of  $z$  can be expressed by [1, 28]

$$\frac{M_{\text{HI}}}{M_{\odot}} = \frac{2.35 \times 10^5}{(1+z)^2} \left[ \frac{d_L(z)}{\text{Mpc}} \right]^2 \left( \frac{S^{V_{\text{obs}}}}{\text{Jy km s}^{-1}} \right), \quad (14)$$

where  $d_L(z)$  is the luminosity distance to the galaxy and  $S^{V_{\text{obs}}}$  is the velocity integral flux in observed frame, which can be approximated by  $S_{\text{peak}} W_{50}$  if we assume HI signal is a top-hat function. We can substitute  $S_{\text{peak}}$  by  $S_{\text{lim}}$  in eq. (11) to calculate the minimum HI mass or  $M_{\text{HI,lim}}$  FAST will detect. If we assume the sensitivity of ALFALFA is 2.4 mJy [16] at resolution of  $10 \text{ km s}^{-1}$ , for detecting an HI galaxy with a line width of  $200 \text{ km s}^{-1}$  at redshift of 0.05, in ALFALFA survey,  $M_{\text{HI,lim}}$  will be about  $10^{10} M_{\odot}$ ; while for two-pass CRAFTS,  $M_{\text{HI,lim}}$  will be about  $10^{9.3} M_{\odot}$  and  $10^{9.5} M_{\odot}$  at DEC of  $25.6^\circ$  and  $-14.0^\circ$ , respectively.

The method we use here to make mock catalogue is similar to ref. [30], we integrate the HIMF ranges from 6.2 to 14 times of  $\log_{10}(M_{\text{HI}}/M_{\odot})$  to obtain the average HI galaxy number density and use the normalized HIMF as the probability distribution function (PDF) of  $M_{\text{HI}}$ . We make a mock catalogue by using the PDF we obtain, the number of galaxy in mock catalogue is obtained by multiplying average number density with the volume of sky region we plan to observe. CRAFTS can detect galaxies with  $M_{\text{HI}}$  above  $M_{\text{HI,lim}}$

in mock catalogue. Following [21], we assume the observed line width of all galaxies is  $200 \text{ km s}^{-1}$ , which is a close approximation of average velocity width of HI galaxies.

Using this method, we derive that the total number of HI galaxies CRAFTS will detect in one-pass and two-pass is approximately  $4.0 \times 10^5$  and  $6.5 \times 10^5$ , respectively. The median redshift for CRAFTS after two passes is close to 0.07, and the HI mass limit for detecting a galaxy with line width of  $200 \text{ km s}^{-1}$  at that redshift is approximately  $10^{9.5} M_{\odot}$ . CRAFTS may not be able to detect HI galaxies efficiently at redshift above 0.2. We may need more passes or some strategies to increase SNR to reach higher redshift for HI galaxy detection. A detailed description and discussion of our simulations and the methodology in deriving the total number of HI galaxies in CRAFTS survey will be presented in a separated paper (Zhang et al. in prep.).

For the determination of “faint end” of HIMF, in ref. [31], the author discussed the potential of FAST in detecting low mass HI galaxies in local group, from their simulation results, two-pass CRAFTS would detect 32 galaxies with HI mass above  $10^5 M_{\odot}$  by assuming the effective integration time is 60 s, future detection on low mass galaxies will be a complementary test for current simulation models.

## 6 Confusion

Confusion is caused by source blending in the beam. Beam-size and HI mass limit will increase at high redshift, which will grow the impact of confusion to HI surveys. Confusion will cause biases on flux, line width and redshift measurement of HI galaxies, which will thus influence the calculation of survey products like correlation functions(CF), HIMF and HI line width functions(WF) [30]. According to previous study [21, 32], confusion might be the main factor that limits survey capability of FAST at high redshifts due to its relatively large beam size.

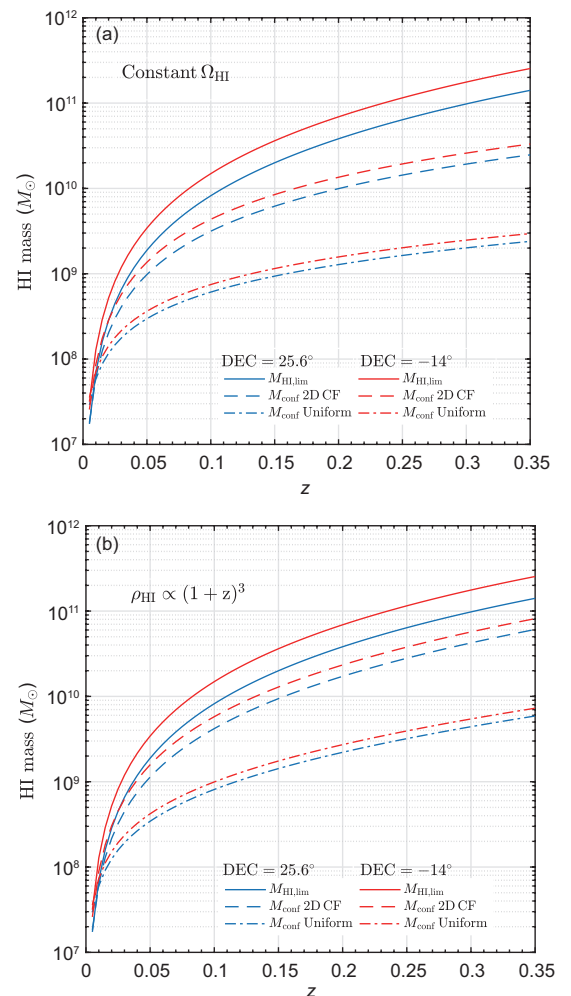
Here we use the model in ref. [32] to analyze the impact of confusion on CRAFTS. That work created a mock stack in the sky with velocity range of  $600 \text{ km s}^{-1}$ , which is the broadest velocity width of HI galaxy. They defined a cylinder in redshift space centered on the target being stacked, the radius of the cylinder is the length of beam size projected on the sky. Then they calculated the total HI mass in the volume of cylinder as an estimation of the confused HI mass that can not be separated from the spectrum of the target, which is referred to “confused mass” or  $M_{\text{conf}}$ . We compare this confused mass to the HI mass limit CRAFTS could detect. If  $M_{\text{conf}}$  is close to  $M_{\text{HI,lim}}$ , it will be difficult for CRAFTS to tell whether there is confusion source in the detected signal or not.

$M_{\text{conf}}$  is closely related with the beamsize and the

beamsize of FAST would vary with ZA and frequency. Figure 8 shows the comparison of  $M_{\text{conf}}$  and  $M_{\text{HI,lim}}$  of two-pass CRAFTS at two typical DEC. It seems that confusion will not affect CRAFTS in detecting sources too much, because  $M_{\text{conf}}$  is smaller than  $M_{\text{HI,lim}}$  at two typical DEC. Perhaps it is due to the limit of integration time and that the beamsize of FAST at high redshift will be smaller than usual.

## 7 Discussion and conclusion

In this paper, we quantify the science prospect of extragalactic HI detections based on a planned large-scale survey,



**Figure 8** (Color online) The comparison of  $M_{\text{HI,lim}}$  and  $M_{\text{conf}}$  of CRAFTS after two passes at DEC of  $25.6^\circ$  (in blue lines) and  $-14^\circ$  (in red lines). Solid lines represent  $M_{\text{HI,lim}}$  of CRAFTS, dashed lines represent  $M_{\text{conf}}$  including 2-dimensional correlation function measured from  $\alpha$ .40 catalogue, which tells the excess probability (above random) of two galaxies being separated by a given distance in a rather small scale [33]. The solid-dashed lines represent the  $M_{\text{conf}}$  if we assume HI galaxies are uniformly distributed. (a) and (b) represent two different results by assuming  $\Omega_{\text{HI}}$  is constant and  $\rho_{\text{HI}} \propto (1+z)^3$ , respectively.

namely, the CRAFTS. We use FAST commissioning data to estimate parameters including the beamsize, the gain and the system temperature. We make a mock catalogue based on the HIMF derived from ALFALFA by assuming the HIMF does not evolve with redshift, and HI galaxies are distributed uniformly in the nearby universe. We also study the potential impact of confusion to CRAFTS survey using model in ref. [32]. We summarized the expected results from CRAFTS as the following.

(1) FAST plans to complete a blind drift scan search of HI galaxies as part of CRAFTS, which covers DEC between  $-14^\circ$  and  $66^\circ$ , nearly  $20000 \text{ deg}^2$  of sky in two passes. The bandwidth of the receiver (FLAN) covers 1.05 to 1.45 GHz, reaching a redshift of 0.35. The channel width is 7.6 kHz, which corresponds to a velocity resolution of approximately  $1.6 \text{ km s}^{-1}$  at HI rest frequency.

(2) The main reflector of FAST can be fully illuminated when zenith angle is below  $26.4^\circ$ , with a beamsize of 2.95 arcmin at the HI rest frequency. At zenith angles above  $26.4^\circ$ , the main reflector is partially illuminated and beamsize will increase linearly with zenith angle.

(3) When zenith angle is below  $26.4^\circ$ , the gain of FAST is largely constant over the bandwidth of FLAN, which is close to  $16.5 \text{ K/Jy}$ . When zenith angle is above  $26.4^\circ$ , there will be a gain loss, at a zenith angle of  $40^\circ$ , the gain will decrease to approximately  $11.0 \text{ K/Jy}$ , with a loss of approximately 33%.

(4) System temperature of FAST will increase with zenith angle ranging from approximately 18 to 26 K.

(5) The rms noise per 7.6 kHz channel of two-pass CRAFTS in most sky regions will be between 1.1 and 1.5 mJy.

(6) From our predictions, CRAFTS may detect over  $6 \times 10^5$  HI galaxies at a median redshift of 0.07 due to its wide coverage and outstanding sensitivity performance. With such huge number of samples, we can make unprecedented progress in studying HI survey products like the HI mass function (HIMF), the Correlation Function (CF) and the HI Velocity Width Function (WF) and deepening our understanding of HI distribution and its property in the universe.

(7) CRAFTS will not be confusion limited for detecting HI galaxy because of its limited integration time and unusual smaller beam size at high redshift.

*This work was supported by the National Key R&D Program of China (Grant Nos. 2017YFA0402600, and 2016YFA0400702), the National Natural Science Foundation of China (Grant Nos. 11690024, 11725313, 11590783, and 11803051), the International Partnership Program of Chinese Academy of Sciences (Grant No. 114A11KYSB20160008), the CAS "Light of West China" Program, and the Young Researcher Grant of National Astronomical Observatories, Chinese Academy of Sciences. The author would like to thank helpful discussion with Michael G. Jones.*

- 1 M. S. Roberts, *Galaxies and the Universe*, edited by M. Sandage, and J. Kristan (University of Chicago Press, Chicago, 1975), p. 309.
- 2 D. G. Barnes, L. Staveley-Smith, W. J. G. de Blok, T. Oosterloo, I. M. Stewart, A. E. Wright, G. D. Banks, R. Bhathal, P. J. Boyce, M. R. Calabretta, M. J. Disney, M. J. Drinkwater, R. D. Ekers, K. C. Freeman, B. K. Gibson, A. J. Green, R. F. Haynes, P. Te Lintel Hekkert, P. A. Henning, H. Jerjen, S. Juraszek, M. J. Kesteven, V. A. Kilborn, P. M. Knezek, B. Koribalski, R. C. Kraan-Korteweg, D. F. Malin, M. Marquarding, R. F. Minchin, J. R. Mould, R. M. Price, M. E. Putman, S. D. Ryder, E. M. Sadler, A. Schroder, F. Stootman, R. L. Webster, W. E. Wilson, and T. Ye, *Mon. Not. R. Astron. Soc.* **322**, 486 (2001).
- 3 M. J. Meyer, M. A. Zwaan, R. L. Webster, L. Staveley-Smith, E. Ryan-Weber, M. J. Drinkwater, D. G. Barnes, M. Howlett, V. A. Kilborn, J. Stevens, M. Waugh, M. J. Pierce, R. Bhathal, W. J. G. de Blok, M. J. Disney, R. D. Ekers, K. C. Freeman, D. A. Garcia, B. K. Gibson, J. Harnett, P. A. Henning, H. Jerjen, M. J. Kesteven, P. M. Knezek, B. S. Koribalski, S. Mader, M. Marquarding, R. F. Minchin, J. O'Brien, T. Oosterloo, R. M. Price, M. E. Putman, S. D. Ryder, E. M. Sadler, I. M. Stewart, F. Stootman, and A. E. Wright, *Mon. Not. R. Astron. Soc.* **350**, 1195 (2004).
- 4 R. Giovanelli, M. P. Haynes, B. R. Kent, P. Perillat, A. Saintonge, N. Brosch, B. Catinella, G. L. Hoffman, S. Stierwalt, K. Spekkens, M. S. Lerner, K. L. Masters, E. Momjian, J. L. Rosenberg, C. M. Springob, A. Boselli, V. Charmandaris, J. K. Darling, J. Davies, D. G. Lambas, G. Gavazzi, C. Giovanardi, E. Hardy, L. K. Hunt, A. Iovino, I. D. Karachentsev, V. E. Karachentseva, R. A. Koopmann, C. Marinoni, R. Minchin, E. Muller, M. Putman, C. Pantoja, J. J. Salzer, M. Scodreggio, E. Skillman, J. M. Solanes, C. Valotto, W. van Driel, and L. van Zee, *Astron. J.* **130**, 2598 (2005).
- 5 M. P. Haynes, R. Giovanelli, A. M. Martin, K. M. Hess, A. Saintonge, E. A. K. Adams, G. Hallenbeck, G. L. Hoffman, S. Huang, B. R. Kent, R. A. Koopmann, E. Papastergis, S. Stierwalt, T. J. Balonek, D. W. Craig, S. J. U. Higdon, D. A. Kornreich, J. R. Miller, A. A. O'Donoghue, R. P. Olowin, J. L. Rosenberg, K. Spekkens, P. Troischt, and E. M. Wilcots, *Astron. J.* **142**, 170 (2011), arXiv: 1109.0027.
- 6 R. Giovanelli, and M. P. Haynes, *Astron. Astrophys. Rev.* **24**, 1 (2016), arXiv: 1510.04660.
- 7 M. P. Haynes, R. Giovanelli, B. R. Kent, E. A. K. Adams, T. J. Balonek, D. W. Craig, D. Fertig, R. Finn, C. Giovanardi, G. Hallenbeck, K. M. Hess, G. L. Hoffman, S. Huang, M. G. Jones, R. A. Koopmann, D. A. Kornreich, L. Leisman, J. Miller, C. Moorman, J. O'Connor, A. O'Donoghue, E. Papastergis, P. Troischt, D. Stark, and L. Xiao, *Astrophys. J.* **861**, 49 (2018), arXiv: 1805.11499.
- 8 F. H. Briggs, *Astron. J.* **100**, 999 (1990).
- 9 P. Schechter, *Astrophys. J.* **203**, 297 (1976).
- 10 A. Klypin, A. V. Kravtsov, O. Valenzuela, and F. Prada, *Astrophys. J.* **522**, 82 (1999).
- 11 A. M. Martin, E. Papastergis, R. Giovanelli, M. P. Haynes, C. M. Springob, and S. Stierwalt, *Astrophys. J.* **723**, 1359 (2010), arXiv: 1008.5107.
- 12 L. J. Storrie-Lombardi, and A. M. Wolfe, *Astrophys. J.* **543**, 552 (2000).
- 13 C. Péroux, R. G. McMahon, L. J. Storrie-Lombardi, and M. J. Irwin, *Mon. Not. R. Astron. Soc.* **346**, 1103 (2003).
- 14 M. A. Zwaan, M. J. Meyer, L. Staveley-Smith, and R. L. Webster, *Mon. Not. R. Astron. Soc.-Lett.* **359**, L30 (2005).
- 15 A. M. Martin, E. Papastergis, R. Giovanelli, M. P. Haynes, C. M. Springob, and S. Stierwalt, *Astrophys. J.* **723**, 1359 (2010), arXiv: 1008.5107.
- 16 M. G. Jones, M. P. Haynes, R. Giovanelli, and C. Moorman, *Mon. Not. R. Astron. Soc.* **477**, 2 (2018), arXiv: 1802.00053.
- 17 D. Obreschkow, D. Croton, G. De Lucia, S. Khochfar, and S. Rawlings, *Astrophys. J.* **698**, 1467 (2009), arXiv: 0904.2221.
- 18 A. R. Duffy, S. T. Kay, R. A. Battye, C. M. Booth, C. Dalla Vecchia, and J. Schaye, *Mon. Not. R. Astron. Soc.* **420**, 2799 (2012), arXiv: 1107.3720.



- 19 C. M. Moorman, M. S. Vogeley, F. Hoyle, D. C. Pan, M. P. Haynes, and R. Giovanelli, *Mon. Not. R. Astron. Soc.* **444**, 3559 (2014), arXiv: [1408.3392](#).
- 20 M. G. Jones, E. Papastergis, M. P. Haynes, and R. Giovanelli, *Mon. Not. R. Astron. Soc.* **457**, 4393 (2016), arXiv: [1510.07050](#).
- 21 A. R. Duffy, R. A. Battye, R. D. Davies, A. Moss, and P. N. Wilkinson, *Mon. Not. R. Astron. Soc.* **383**, 150 (2007), arXiv: [0707.2316](#).
- 22 A. R. Duffy, M. J. Meyer, L. Staveley-Smith, M. Bernyk, D. J. Croton, B. S. Koribalski, D. Gerstmann, and S. Westerlund, *Mon. Not. R. Astron. Soc.* **426**, 3385 (2012), arXiv: [1208.5592](#).
- 23 R. Nan, *Sci. China Ser. G-Phys. Mech. Astron.* **49**, 129 (2006).
- 24 R. Nan, D. Li, C. Jin, Q. Wang, L. Zhu, W. Zhu, H. Zhang, Y. Yue, and L. Qian, *Int. J. Mod. Phys. D* **20**, 989 (2011), arXiv: [1105.3794](#).
- 25 D. Li, P. Wang, L. Qian, M. Krco, A. Dunning, P. Jiang, Y. Yue, C. Jin, Y. Zhu, Z. Pan, and R. Nan, *IEEE Microwave* **19**, 112 (2018), arXiv: [1802.03709](#).
- 26 P. Jiang, Y. L. Yue, H. Q. Gan, R. Yao, H. Li, G. F. Pan, J. H. Sun, D. J. Yu, H. F. Liu, N. Y. Tang, L. Qian, J. G. Lu, J. Yan, B. Peng, S. X. Zhang, Q. M. Wang, Q. Li, D. Li, and FAST Collaboration, *Sci. China-Phys. Mech. Astron.* **62**, 959502 (2019), arXiv: [1903.06324](#).
- 27 D. W. Hogg, arXiv: [astro-ph/9905116](#).
- 28 M. Meyer, A. Robotham, D. Obreschkow, T. Westmeier, A. R. Duffy, and L. Staveley-Smith, *Publ. Astron. Soc. Aust.* **34**, e052 (2017), arXiv: [1705.04210](#).
- 29 C. J. Jin, K. Zhu, J. Fan, H. F. Liu, Y. Zhu, H. Q. Gan, J. L. Yu, Z. S. Gao, Y. Cao, and Y. Wu, in *The optics of the Five-hundred-meter Aperture Spherical radio Telescope: Proceedings of the International Symposium on Antennas and Propagation (ISAP)*, Nanjing, 2013.
- 30 M. G. Jones, E. Papastergis, M. P. Haynes, and R. Giovanelli, *Mon. Not. R. Astron. Soc.* **449**, 1856 (2015), arXiv: [1502.07359](#).
- 31 J. Li, Y. G. Wang, M. Z. Kong, J. Wang, X. Chen, and R. Guo, *Res. Astron. Astrophys.* **18**, 003 (2018), arXiv: [1711.09315](#).
- 32 M. G. Jones, M. P. Haynes, R. Giovanelli, and E. Papastergis, *Mon. Not. R. Astron. Soc.* **455**, 1574 (2016), arXiv: [1510.03854](#).
- 33 E. Papastergis, R. Giovanelli, M. P. Haynes, A. Rodríguez-Puebla, and M. G. Jones, *Astrophys. J.* **776**, 43 (2013), arXiv: [1308.2661](#).