Predictions for ASKAP neutral hydrogen surveys

Alan R. Duffy,1,2* Martin J. Meyer,1 Lister Staveley-Smith,1 Maksym Bernyk,3 Darren J. Croton,3 Bärbel S. Koribalski,4 Derek Gerstmann1 and Stefan Westerlund1

1International Centre for Radio Astronomy Research, The University of Western Australia, Crawley, WA 6009, Australia
2School of Physics, University of Melbourne, Parkville, VIC 3010, Australia
3Centre for Astrophysics & Supercomputing, Swinburne University of Technology, PO Box 218, Hawthorn, VIC 3122, Australia
4CSIRO Astronomy and Space Science, Australia Telescope National Facility, Box 76, Epping, NSW 1710, Australia

Accepted 2012 August 24. Received 2012 August 24; in original form 2012 May 23

ABSTRACT

The Australian Square Kilometre Array Pathfinder (ASKAP) will revolutionize our knowledge of gas-rich galaxies in the Universe. Here we present predictions for two proposed extragalactic ASKAP neutral hydrogen (H\textsc{i}) emission-line surveys, based on semi-analytic models applied to cosmological \textsc{N}-body simulations. The ASKAP H\textsc{i} All-Sky Survey, known as Widefield ASKAP L-band Legacy All-sky Blind surveY (WALLABY), is a shallow 3π survey (z = 0–0.26) which will probe the mass and dynamics of over 6 × 10^5 galaxies. A much deeper small-area H\textsc{i} survey, called Deep Investigation of Neutral Gas Origins (DINGO), aims to trace the evolution of H\textsc{i} from z = 0 to 0.43, a cosmological volume of 4 × 10^7 Mpc^3, detecting potentially 10^5 galaxies. The high-sensitivity 30 antenna ASKAP core (diameter ∼2 km) will provide an angular resolution of 30 arcsec (at z = 0). Our simulations show that the majority of galaxies detected in WALLABY (87.5 per cent) will be resolved. About 5000 galaxies will be well resolved, i.e. more than five beams (2.5 arcmin) across the major axis, enabling kinematic studies of their gaseous discs. This number would rise to 1.6 × 10^5 galaxies if all 36 ASKAP antennas could be used; the additional six antennas provide baselines up to 6 km, resulting in an angular resolution of 10 arcsec. For DINGO this increased resolution is highly desirable to minimize source confusion, reducing confusion rates from a maximum of 10 per cent of sources at the survey edge to 3 per cent. We estimate that the sources detected by WALLABY and DINGO will span four orders of magnitude in total halo mass (from 10^{11} to 10^{15} M_☉) and nearly seven orders of magnitude in stellar mass (from 10^5 to 10^{12} M_☉), allowing us to investigate the process of galaxy formation across the last four billion years.

Key words: galaxies: evolution – galaxies: luminosity function, mass function – radio lines: galaxies.

1 INTRODUCTION

Neutral hydrogen (H\textsc{i}) is an ubiquitous tracer of large-scale structure in the Universe. It allows us to study the physical and dynamical processes within galaxies, including the kinematic properties of structures such as bars, discs and warps. Each galaxy H\textsc{i} spectrum provides a large set of galaxy properties, e.g. the systemic velocity, the integrated flux density and the velocity width. These are used to derive the galaxy distance, its gas mass and its total dynamical mass, respectively. The gas mass is also a good indicator of ongoing star formation.

The evolution of H\textsc{i} is of fundamental importance to understanding the build-up of both the stellar and gas masses within galaxies and the method by which galaxies accrete their material. Due to the inherent signal weakness of the 21-cm hyperfine splitting transition, the detection of H\textsc{i} emission in distant galaxies requires high-resolution and high-sensitivity observations. It is a crucial window into galaxy formation over time; and while we wait for the next generation of large-scale H\textsc{i} surveys, we will explore their potential via \textsc{N}-body simulations.

In Table 1 we present a, non-exhaustive, list of large area H\textsc{i} surveys that have either been completed or are currently ongoing, as well as several future surveys of note.

Several large-scale H\textsc{i} surveys were obtained with the 64-m Parkes telescope, made possible by the innovative 21-cm multi-beam system which consists of 13 dual-polarization feed horns and a powerful correlator (Staveley-Smith et al. 1996). Most prominent among them is the H\textsc{i} Parkes All Sky Survey (HIPASS; Barnes et al. 2001). We also include the (unfinished) northern counterpart,
the H\textsc{i} Jodrell All Sky Survey (HIJASS; Lang et al. 2003) which utilized a four-beam receiver. We list two major surveys currently underway in Table 1. The first is the Arcroce Legacy Fast ALFA Survey (ALFALFA) which utilizes the Arcroce L-Band Feed Array (ALFA) seven-beam receiver on the 305-m Arcroce dish. The second is the Effelsberg-Bonn H\textsc{i} Survey (EBH\textsc{i}; Kerp et al. 2011), carried out with the recently installed seven-beam system on the 100-m dish. Finally we present two future surveys in the table: the first is the Chinese-built Five-hundred metre Aperture Spherical Telescope (FAST; Nan 2006) which utilizes a 19-beam receiver and Square Kilometre Array (SKA1).

There are three precursor instruments to the SKA: the Murchison Widefield Array (MWA; Lonsdale et al. 2009), the Meer-Karoo Array Telescope (MeerKAT; Booth et al. 2009) and the Australian SKA Pathfinder (ASKAP; Johnston et al. 2008; Deboer et al. 2009).

Here we will focus on planned H\textsc{i} surveys with ASKAP, which is currently under construction in the Murchison Radio Astronomy Observatory in Western Australia. ASKAP will consist of 36 antennas (12-m diameter), of these 30 antennas are located within a 2-km diameter circle. ASKAP’s large field of view – 30deg – provided by novel phased array feeds (Chippendale et al. 2010) makes ASKAP a 21-cm survey machine.

In the following sections we introduce two ASKAP H\textsc{i} surveys: the shallow H\textsc{i} All-Sky Survey [known as Widefield ASKAP H\textsc{i} Blind surveY (WALLABY); Koribalski & Staveley-Smith 2009] and the deep, but small-area H\textsc{i} survey [known as Deep Investigation of Neutral Gas Origins (DINGO); Meyer 2009].

1.1 WALLABY

The WALLABY\textsuperscript{2} is a large project led by Bärbel Koribalski and Lister Staveley-Smith. WALLABY proposes to observe ~75 per cent of the sky (−90° < δ < +30°) out to a redshift of z = 0.26. In the WALLABY proposal (Koribalski & Staveley-Smith 2009), it is estimated that 500,000 galaxies will be detected over the full survey area (assuming an angular resolution of 30 arcsec), of these ~1000 galaxies will be spatially well resolved (i.e. >5 arcmin in angular extent or >10 beams).

To achieve a much higher angular resolution of 10 arcsec the full ASKAP configuration (36 antennas with baselines up to 6 km) is needed. Given the high computational cost of spectral line imaging of large volumes at such high resolution, we are considering to obtain ‘postage stamps’. These are high-resolution mini cubes formed at the position and velocity of particularly interesting galaxies determined a priori. This is possible because the 21-cm data are collected by the full array and will serve many survey science projects. For simplicity in this work we assume that all galaxies will be imaged using the 6-km baselines as we are also interested in probing the issue of resolving out objects and hence studying what fraction of galaxies become non-detected.

The WALLABY goals, outlined in detail in the ASKAP Survey Science Proposal (Koribalski & Stavely-Smith 2009), are to examine the properties, environment and large-scale distribution of gas-rich galaxies. In summary, WALLABY will study galaxy formation and the missing satellite problem in the Local Group, evolution and star formation in galaxies, mergers and interactions in galaxies, the H\textsc{i} mass function and its variation with local environment, processes governing the evolution and distribution of cool gas at low redshift, and the nature of the cosmic web.

WALLABY will also be able to investigate cosmological parameters. For example, we will be able to measure the matter power spectrum in the local Universe. Furthermore, we should be able to constrain the equation of state of dark energy to better than 20 per cent (Duffy, Moss & Staveley-Smith 2012b). Other tests that have been proposed for such a large H\textsc{i} survey are studying the coherent bulk flows of galaxies on large scales (e.g. Burke & Taylor 2004; Abate et al. 2008), the measurement of Baryonic acoustic oscillations and the Hubble constant (Beutler et al. 2011), the measurement of the rms mass fluctuations, σ, and growth factor, f (Beutler et al. 2012), and the surface brightness dimming of objects with intrinsic brightness, the so-called Tolman test (Khdekar & Chakraborti 2011).

In this paper we also consider a proposed Northern hemisphere H\textsc{i} survey that, when combined with WALLABY, will provide a true H\textsc{i} all-sky survey. A proposal has been submitted to the Astron Westerbork Synthesis Radio Telescope (WSRT) facility to carry out the Westerbork Northern Sky H\textsc{i} Survey (WNSHS\textsuperscript{3}). This project is led by Guyla Jozsa, and will target the northern sky using new phased array feeds that can instantaneously observe 8 deg\textsuperscript{2}. WNSHS as proposed will likely be slightly deeper survey than WALLABY with a resolution similar to the full 6-km baseline of ASKAP. The science goals for WNSHS are similar, and due to filling in the sky coverage

\begin{table}
\centering
\begin{tabular}{|c|c|c|c|}
\hline
Survey & Sky area (°) (resolution) & Velocity range (km s\textsuperscript{-1}) & Detections & Reference \\
\hline
HIPASS\textsuperscript{1} & 29 343 (15.5 arcmin) & −1280 < cz < 12 700 & ~5300 & Barnes et al. (2001) \\
HIJASS & 11000 (12 arcmin) & −1280 < cz < 12 700 & 222 & Lang et al. (2003) \\
ALFALFA\textsuperscript{2} & 7000 (4 arcmin) & −1600 < cz < 18 000 & ~15 000 & Haynes et al. (2011) \\
EBH\textsc{i} S & 21 400 (9 arcmin) & cz < 21000 & Underway & Kerp et al. (2011) \\
FAST\textsuperscript{3} & 4000 (3 arcmin) & z < 0.4 & 2 × 10\textsuperscript{6} & Duffy et al. (2008) \\
SKA\textsuperscript{4} & 20 000 (12 arcsec) & z < 1.5 & 10\textsuperscript{9} & Abdalla & Rawlings (2005) \\
\hline
\end{tabular}
\caption{We summarize here the key large area H\textsc{i} surveys that have been completed, or near to completion, as well as future projects. \textsuperscript{1}We have included several catalogues: the HIPASS Bright Galaxy Catalogue (Koribalski et al. 2004), the Southern hemisphere H\textsc{i} catalogue (Meyer et al. 2004) and its northern extension (δ < 25°; Wong et al. 2006). \textsuperscript{2}For ALFALFA detections we quote the values from the 40 per cent catalogue which at the time of writing was the most complete published results. \textsuperscript{3}The survey considered here consisted of 600 s integration times on source, lasting a year. \textsuperscript{4}The SKA as considered in Abdalla & Rawlings (2005) would contain 50 per cent of the collecting area within 5 km and we therefore use this to determine the typical resolution although in reality baselines may extend to three orders of magnitude larger.}
\end{table}
missed by ASKAP, entirely complimentary to WALLABY. As currently designed there are two survey modes for WNSHS, the first is a limited integration time of 4 h per observation, which would perfectly match the expected sensitivity for the WALLABY survey. The second is a deeper, 12 h pointing, which would decrease the flux limits by 1.7 and open up a much deeper exploration (more in line with the deep ASKAP survey, DINGO, discussed later). We consider both survey modes, with a 4 h WNSHS model + WALLABY, termed ‘ALL SKY’ and the deeper WNSHS case on its own. In this study we will show that the ALL SKY combination of WALLABY and the 4 (12) h pointing WNSHS could potentially detect of the order of 800k (>10^3) galaxies.

1.2 DINGO

The DINGO^4 survey is led by Martin Meyer and consists of deep and ultradeph phases which differ in area and depth. In the first phase, the survey proposes to target five non-contiguous fields, 150 deg^2 in total, out to z = 0.26. While the redshift range is the same as for WALLABY, the integration per field is 500 h (>60 times longer) providing 8 times better sensitivity. Where feasible the target fields will be selected to overlap with the Galaxy And Mass Assembly (GAMA; Driver et al. 2009) survey. The second phase is proposed to consist of two ultradeph fields, 60 deg^2 in total, over the redshift range z = 0.1-0.43. Very long integration times with 2500h per field will enable DINGO UDEEP to probe the evolution of H I over the last 4 billion years of cosmic time.

DINGO is designed to probe the H I universe out to the maximum redshift possible, enabling the evolution in key cosmological parameters (such as the cosmic H I density) and the H I mass function to be measured. As well as this it will sample a sufficient volume to measure the two-point correlation function and the halo occupation distribution function as a function of redshift. Overlaps with GAMA fields will provide matching stellar properties for the DINGO H I detections. Alternatively, for the H I non-detections in these fields, the optical redshifts from GAMA enable the possibility of H I spectral stacking to extend the effective limit of DINGO.

1.3 ASKAP H I predictions

Estimating the performance of the proposed ASKAP surveys, WALLABY and DINGO, is a challenging theoretical problem due to both the large cosmic volumes probed by these surveys (3.26 and 0.04 Gpc^3, respectively) and the low detection mass threshold (∼10^5 M⊙).

Previous work on simulating H I in cosmological simulations using fully hydrodynamical simulations (e.g. Popping et al. 2009; Altay et al. 2011; Duffy et al. 2012a) has been limited to smaller volumes (<100 h^(-1) Mpc)^3 and hence is unsuitable for making accurate predictions of the large variety of structure found with ASKAP. Instead we can use N-body simulations, which are computationally cheaper to run, and hence can simulate larger regions of the Universe. Creating galaxy properties in simulations that do not explicitly track the gas is called semi-analytic modelling. The low computational cost of running a galaxy formation model atop an existing N-body simulation is such that the various parameters in the model can be tuned to successfully recreate numerous observational constraints.

Several recent attempts have been made to split the cold gas from the semi-analytic model into atomic and molecular hydrogen components (e.g. Obreschkow & Rawlings 2009a; Power, Baugh & Lacey 2010; Lagos et al. 2011). Using observational the constraint, from Blitz & Rosolowsky (2006) and Leroy et al. (2008), that the molecular–atomic ratio H_2/H I depends on the local interstellar medium pressure, Obreschkow & Rawlings (2009a) found that the H I mass function does not strongly evolve strongly until z > 1. Using a similar methodology, Power et al. (2010) found that this conclusion holds for several semi-analytic models. Lagos et al. (2011) differ slightly in that they calculate the H_2/H I ratio in the semi-analytic model itself and form stars from the molecular component alone. The non-evolving prediction of the H I mass function holds as before, with some interesting new predictions at low H I masses <10^6 h^(-1) M⊙ which are below the typical thresholds for the galaxies found in ASKAP but which the SKA might detect in large numbers.

In this study we seek a simple recipe to guarantee that the H I masses in the simulation are similar to those observed locally. Hence, rather than adopting one of the previously mentioned techniques, we argue for a more transparent method. One possibility could be to randomly assign H I masses based on the locally observed H I mass function (e.g. Zwaan et al. 2005; Martin et al. 2010) to dark matter haloes in an existing N-body simulation. However, there are several advantages in using the galaxy properties from semi-analytic modelling. First, confusion estimates based on realistic clustering of galaxies become possible. In addition, we can probe the effects of the modest evolution in the H I abundance of the galaxies as their cold gas mass changes in the simulation. Finally, one can examine the galaxy stellar properties from the catalogue with the H I-selected sample.

To these ends we make use of a semi-analytic catalogue of galaxies to form a mock light cone for ASKAP as detailed in Section 2, with the method by which we create H I gas masses for each galaxy in the catalogue given in Section 2.1. We consider the sensitivity of a radio telescope in Section 2.3 and the impact that realistic spatial and spectral features for the simulated galaxies will have when observed by instruments such as ASKAP in Sections 2.3.2 and 2.3.4. In Sections 1.1 and 1.2, we detail the anticipated number counts for WALLABY and DINGO, respectively, along with the ability of ASKAP to probe large distributions of galaxy properties in such a flux-limited sample. The issue of confusion of H I sources is considered in Section 4.1 for the particular case of DINGO. We consider the problems of identifying optical counterparts to the undetected H I sources to enable the stacking of their spectra in Section 4.3. We compare the H I surveys in Section 5. Finally we conclude in Section 6 and emphasise the need for zoom-in, high-resolution images around galaxy detections, so-called ‘postage stamps’, to limit the effects of confusion on source counts as well as guarantee a large sample of resolved galaxies that can be modelled using their velocity field information.

2 METHODOLOGY

The galaxy catalogue is created using the semi-analytic model (SAM) of Croton et al. (2006) to produce galaxies based on the underlying dark-matter-only Millenium Simulation (Springel et al. 2005). This sample of galaxies accurately recreates the observed stellar mass function with a combination of supernovae feedback and, crucially to this model, feedback at the high-mass end from active galactic nuclei. The cosmology used in the Millennium Simulation is (Ω_m = 0.25, Ω_Λ = 0.75, Ω_b = 0.045, σ_8 = 0.9 and

© 2012 The Authors, MNRAS 426, 3385–3402
Monthly Notices of the Royal Astronomical Society © 2012 RAS
This differs from the current best-fitting cosmological model based on measurements of supernovae and the cosmic microwave background (Komatsu et al. 2011), most significantly in the relatively high $\sigma_8$ value adopted in the simulation. This will have the effect of creating haloes that, at a fixed total mass of $10^{11} (10^{-4}) h^{-1} M_\odot$, are ~30 per cent (20 per cent) too concentrated (Duffy et al. 2008). We ignore the effects of baryons on the dark matter halo, the baryonic back-reaction, which typically increase the concentration but, with sufficient feedback (e.g. by an accreting supermassive blackhole), gas can be expelled to actually reduce the concentration relative to the dark-matter-only simulation (e.g Duffy et al. 2010). At all masses the haloes will be ~5 per cent too spherical but with identical spin distributions (Macciò, Dutton & van den Bosch 2008) relative to the ‘true’ cosmology.

We then utilized the Theoretical Astrophysical Observatory (TAO),\(^5\) which is a cloud-based web application which produces mock catalogues from different cosmological simulations and galaxy models in the form of a light cone (detailed in Bernyk et al., in preparation) to create an all-sky galaxy catalogue extending to $z = 0.26$ for the WALLABY survey and to $z = 0.43$ for the narrower but deeper DINGO survey. Although the DINGO surveys are not contiguous fields (they are placed around the sky to lessen cosmic variance), we make the simulated light cone a single field of the same total area; this will not have any impact on the expected galaxy counts. This tool creates a light cone through the simulation volume, stitching a randomly rotated cube to the far edge to minimize the repeated structure in a cone. Objects that are broken up by this rotation of the cube are separately reattached to preserve the one- and two-halo terms. The closest snapshot in redshift from the full simulation catalogue is used where possible to accurately trace the evolution of galaxies. This resource provides stellar, dark matter and cold gas masses (amongst many other quantities) for all galaxies within the light cone.

We use the flux limit calculations from Duffy et al. (2012b) and figures for ASKAP given in Table 2 for several $\text{H}^\text{i}$ surveys. In Table 3 (and 4) we provide survey parameters and summarize the results of our simulations. The respective surveys are WALLABY (as described in Section 1.1), All-Sky (WALLABY plus a Northern hemisphere extension with the same resolution and sensitivity, essentially the 4-h WNSHS survey), WNSHS (the 12-h deep survey with 1.7$\alpha$ better sensitivity than WALLABY) and finally DINGO DEEP and UDEEP (as given in Section 1.2).

2.1 Estimating $\text{H}^\text{i}$ in galaxies

The Croton et al. (2006) model produces a list of cold gas masses for the galaxies that we can attempt to break into neutral atomic, molecular hydrogen and ionized hydrogen components located near to the galaxy. At the high-mass end we would envisage significant fractions of neutral hydrogen to be in molecular form (Leroy et al. 2008; Saintonge et al. 2011), with the low-mass end predominantly atomic hydrogen with a component of ionized hydrogen due to the ionizing cosmic ultraviolet background (e.g. Quinn, Katz & Efstathiou 1996). The cold gas–$\text{H}^\text{i}$ mass fraction has been previously studied in semi-analytic works as both a constant; for example, Power et al. (2010) found 0.54 after considering He abundances and typical ionization fractions of the gas, as well as a function of the halo properties themselves (e.g. Obreschkow & Rawlings 2009b). As we are interested in creating a realistic mock catalogue rather than understanding the $\text{H}^\text{i}$ properties of galaxies we apply a reduction ratio $R = M_{\text{H}^\text{i}}/M_{\text{cold}}$ to convert cold gas to $\text{H}^\text{i}$ so as to recover the observed $\text{H}^\text{i}$ mass function.\(^6\) We find that a functional form suggested by Yang, Mo & van den Bosch (2003) for creating a conditional luminosity function by assigning stellar mass to dark matter haloes is particularly suitable for this conversion:

$$R = \left( \frac{M_{\text{H}^\text{i}}}{M_{\text{cold}}} \right) \left[ \left( \frac{M_{\text{cold}}}{M_*} \right)^{-\alpha} + \left( \frac{M_{\text{cold}}}{M_*} \right)^{\beta} \right]^{-1},$$

(1)

where the fraction of $\text{H}^\text{i}$ in the cold gas is $(M_{\text{H}^\text{i}}/M_{\text{cold}})_0 = 0.41$ at a characteristic cold gas mass $M_* = 8.8 \times 10^{10} M_\odot$ with faint-end slope $\alpha = 0.52$ and bright-end slope $\beta = 0.56$ for a sample of galaxies in the redshift range $0.02 < z < 0.07$ centred on the mean redshift of the WALLABY survey $z \approx 0.05$ (as calculated in Duffy et al. 2012a). We refer to the survey predictions using this conversion scheme as Fixed $R$. Interestingly, the normalization of the fraction for the low-redshift sample is similar to the value suggested by Power et al. (2010).

As Power et al. (2010) demonstrated the cosmic density of cold gas within semi-analytic models typically increased by 0.2 dex over the redshift range $z = 0 - 0.43$ and hence we can expect that our predictions with Fixed $R$ will demonstrate mild evolution in the resultant $\text{H}^\text{i}$ mass function. However, we can also refit the mass

\[^{5}\text{http://tao.it.swin.edu.au/mock-galaxy-factory}\]

\[^{6}\text{In this work we take the ALFALFA mass function (Martin et al. 2010) to model the galaxies on but note that this typically results in more $\text{H}^\text{i}$-rich galaxies than if we used HIPASS (Zwaan et al. 2005). The latter mass function results in 16 per cent fewer galaxies, e.g. in the DINGO DEEP survey.}\]
Table 3. We summarize here the proposed survey specific values for the two aspects of DINGO (DEEP and Ultra-DEEP), WALLABY and WNSHS as described in the text. For Ultra-DEEP we consider two evolution models: a varying H\textsc{i}-cold gas fraction such that $\Omega_{H_{\text{i}}}$ remains constant and a fixed H\textsc{i}-cold gas fraction that results in a modern increase in $\Omega_{H_{\text{i}}}$ (the latter given in bold face). For reference we include the 1\,$\sigma$ noise levels in each survey for two bandwidth measures of 3.86 km s\textsuperscript{-1} and 100 kHz. Note that with the final decision on the baseline weighting still to be decided, the high-resolution mode is conservatively set to the same flux limits as the low-resolution case (therefore the number of resolved systems is a lower limit). We list the median and quartile limits of the H\textsc{i} diameter of the galaxies in each survey, along with the value for those galaxies detected (in the low-resolution survey only). We consider what fraction of these systems are resolved, i.e. have an angular diameter greater than the angular resolution of the telescope (the low-resolution value is 30 arcsec). These are then separated into the expected number of galaxies that will be resolved by more than 1, 3, 5, 10 or 20 beams (in parenthesis are the percentage of detected galaxies in each case). We then consider a postage-stamp reimaging around each of these detections at higher resolution using 10 arcsec resolution available with the full 6 km baseline of ASKAP. For the case of WNSHS the highest resolution imaging is 13 arcsec. This extra resolution makes a significant difference to the final number counts of well-resolved systems and argues for a high-resolution reimaging for strong galaxy detections. The survey number counts are presented in Table 4.

| Parameter | WALLABY | ALL SKY | WNSHS | DINGO DEEP | DINGO UDEEP |
|-----------|---------|---------|-------|------------|-------------|
| Sky obs (deg\textsuperscript{2}) | 30940 | 41252 | 10313 | 150 | 60 |
| Survey (h) | 9600 | 14800 | 16900 | 2500 | 5000 |
| Redshift range | 0–0.26 | 0–0.26 | 0–0.26 | 0–0.26 | 0.1–0.43 |

1\,$\sigma$ noise (mJy km s\textsuperscript{-1})

- $\Delta v = 3.86$ km s\textsuperscript{-1}
  - $D_{H_{\text{i}}}$ (kpc) 14±3
- $\Delta v = 100$ kHz
  - $D_{H_{\text{i}}}$ (kpc) 14±3

Detected $D_{H_{\text{i}}}$ (kpc) 42±10

| Low resolution | WALLABY | ALL SKY | WNSHS | DINGO DEEP | DINGO UDEEP |
|----------------|---------|---------|-------|------------|-------------|
| > 1 beam [$N_{\text{gal}}$ (per cent)] | 542706 (87.5) | 726346 (87.5) | 327454 (62.4) | 4926 (9.0) | 579 (1.4), 855 (1.4) |
| > 3 beams [$N_{\text{gal}}$ (per cent)] | 30859 (5.0) | 41397 (5.0) | 11557 (2.2) | 152 (0.3) | 0 (0), 0 (0) |
| > 5 beams [$N_{\text{gal}}$ (per cent)] | 6205 (1.0) | 8394 (1.0) | 2416 (0.5) | 38 (0.07) | 0 (0), 0 (0) |
| > 10 beams [$N_{\text{gal}}$ (per cent)] | 694 (0.1) | 946 (0.1) | 271 (0.07) | 2 (0.004) | 0 (0), 0 (0) |
| > 20 beams [$N_{\text{gal}}$ (per cent)] | 74 (0.01) | 109 (0.01) | 37 (0.01) | 0 (0) | 0 (0), 0 (0) |

| High resolution | WALLABY | ALL SKY | WNSHS | DINGO DEEP | DINGO UDEEP |
|-----------------|---------|---------|-------|------------|-------------|
| > 1 beam [$N_{\text{gal}}$ (per cent)] | 247038 (100) | 323156 (100) | 356447 (100) | 4926 (9.0) | 431904 (76.5) |
| > 3 beams [$N_{\text{gal}}$ (per cent)] | 239653 (97.0) | 310109 (96.0) | 153328 (43.0) | 4926 (10.4) | 759 (1.5), 855 (1.6) |
| > 5 beams [$N_{\text{gal}}$ (per cent)] | 146280 (59.2) | 183811 (56.9) | 31759 (8.9) | 955 (2.0) | 47 (0.10), 37 (0.07) |
| > 10 beams [$N_{\text{gal}}$ (per cent)] | 22149 (9.0) | 26310 (8.1) | 3761 (1.1) | 115 (0.2) | 0 (0), 0 (0) |
| > 20 beams [$N_{\text{gal}}$ (per cent)] | 2476 (1.0) | 2868 (0.9) | 432 (0.1) | 10 (0.02) | 0 (0), 0 (0) |

Table 4. We summarize here the proposed survey specific values for the two aspects of DINGO (DEEP and Ultra-DEEP) and WALLABY (Koribalski & Staveley-Smith 2009) as well as a northern extension plus WALLABY (‘ALL SKY’) and the WNSHS survey, similar to Table 3. We present two numbers for the predicted galaxy counts, and their mean redshift, reflecting the effects of including the reduction of signal-to-noise ratio by spatially resolved galaxies, as demonstrated in Fig. 3. The brackets ignore this effect and therefore have a larger galaxy number count. There are second numbers in boldface for DINGO UDEEP which is our evolutionary model. The H\textsc{i}–cold gas fraction that results in a modern increase in $\Omega_{H_{\text{i}}}$ and hence more H\textsc{i}–cold gas mass conversion for $z = 0$ but applied to the high-redshift cold gas mass function, which increases with redshift in agreement with Power et al. (2010). There is more cold gas at high redshift, and hence more H\textsc{i}–cold gas mass function is allowed to evolve by retaining the same H\textsc{i}–cold gas mass conversion for $z = 0$ but applied to the high-redshift cold gas mass function, which increases with redshift in agreement with Power et al. (2010). There is more cold gas at high redshift, and hence more H\textsc{i} which results in the ultra-deep survey outperforming the wider angle, lower redshift DINGO survey in terms of number of galaxies detected.
Table 4 – continued

| S/N | Distribution | WALLABY | ALL SKY | WNSHS | DINGO DEEP | DINGO UDEEP |
|-----|--------------|---------|---------|--------|------------|-------------|
|     |              | Ω_H^1   | Ω_H^1   |        | Ω_H^1, Fixed | Ω_H^1, Fixed |
| 10  | n_{gal}      | 196 119 (272341) | 262 328 (364399) | 176 959 (216 630) | 219 80 (22 933) | 18 407 (18934), 20 654 (21 183) |
|     | z ≥ 0.036    | 0.036 (0.043) | 0.036 (0.043) | 0.049 (0.055) | 0.103 (0.105) | 0.184 (0.186), 0.186 (0.187) |
| 15  | n_{gal}      | 94 767 (151174) | 127 064 (202187) | 89 993 (120 789) | 12 402 (13 235) | 9027 (9443), 10142 (10 658) |
|     | z ≥ 0.028    | 0.028 (0.035) | 0.029 (0.035) | 0.039 (0.045) | 0.087 (0.090) | 0.167 (0.170), 0.167 (0.170) |
| 20  | n_{gal}      | 54 688 (99285) | 73 349 (132934) | 54 558 (79 653) | 8230 (8977) | 4997 (5422), 5868 (6282) |
|     | z ≥ 0.023    | 0.023 (0.031) | 0.023 (0.031) | 0.032 (0.040) | 0.077 (0.080) | 0.157 (0.160), 0.156 (0.159) |
| 25  | n_{gal}      | 35 031 (71613) | 47 015 (95879) | 36 585 (57 474) | 5941 (6620) | 3090 (3416), 3611 (3981) |
|     | z ≥ 0.020    | 0.020 (0.028) | 0.020 (0.028) | 0.028 (0.035) | 0.069 (0.073) | 0.150 (0.152), 0.148 (0.152) |
| 30  | n_{gal}      | 23 808 (54753) | 32 070 (73334) | 25 874 (44 165) | 4576 (5188) | 2049 (2334), 2427 (2713) |
|     | z ≥ 0.018    | 0.018 (0.025) | 0.018 (0.025) | 0.025 (0.032) | 0.063 (0.068) | 0.144 (0.147), 0.143 (0.146) |

function in several redshift slices such that we can recover a non-evolving Ω_H^1; we have calculated these, fairly arbitrary, conversions for the higher redshift sample in DINGO UDEEP. We call this conversion the Fixed Ω_H^1 model. We include the different predictions for the Fixed RΩ_H^1 cases in Table 3. However, in all plots and discussions in the text we will consider only the latter case as it is both the most conservative in terms of galaxies detected and in better agreement with the data, which typically shows little evolution in cosmic H_I density to z = 3 (e.g. Rao, Turnshek & Nestor 2006; Lah et al. 2007; Noterdaeme et al. 2009; Prochaska & Wolfe 2009; Meiring et al. 2011). As shown in Fig. 1 this method creates an H_I mass function in excellent agreement with the data.

This simple, yet effective, method produces a complex H_I–halo mass distribution that we examine in Fig. 2. In the top panel we consider the histogram distribution of the H_I–halo mass ratio, finding that the majority of systems have an H_I mass that is approximately 1 per cent of the total halo mass. This distribution sharply drops off to higher ratios with a strongly asymmetric tail to lower ratios which extends to systems incredibly H_I poor. If we consider which systems are, relatively, the most devoid of H_I we find (from the...
Figure 3. We consider the number of galaxies distributed as a function of signal-to-noise ratio for the DINGO DEEP survey. The H\textsc{i} masses of the galaxies are converted from the cold gas mass from the Crotan et al. (2006) semi-analytic model as given by equation (1) and converted into an H\textsc{i} signal-to-noise ratio by applying equations (4) and (11). The grey solid histogram is the initial source population for all galaxies with cold gas masses above 10\(^8\)M\(_\odot\) (found by Power et al. 2010 to be the resolution limit of the Millennium Simulation). We then calculate the observed galaxy counts for a DINGO survey in black dotted, with a special case in red dash for galaxies that are spatially resolved which suffer a loss of signal as described in Section 2.3.3. Galaxies that are larger or face-on will be more resolved, although the impact on numbers for a deep survey like DINGO is marginal. Several objects lose nearly an order of magnitude in signal-to-noise ratio as a result of this effect which may impact kinematic studies of these objects which demand more robust data than a simple detection.

2.2 Resolution limits of the simulation

For such low H\textsc{i} detection limits as proposed in the DINGO and WALLABY surveys there may be a potential issue with the resolution limits of the semi-analytic catalogue itself. This is because we have conservatively set the minimum H\textsc{i} mass of the catalogue to be the same as the cold gas mass limit, i.e. 10\(^8\)M\(_\odot\). We can try and estimate the number of sources we may mass by using an observed H\textsc{i} mass function and extrapolating below this mass limit. To that end we use the empirical technique presented in Duffy et al. (2012b), which include realistic rotation widths and H\textsc{i} line profiles.

Using empirical relations from Duffy et al. (2012b) we find that the DINGO DEEP survey would detect 64 859 galaxies of these 2871 have an angular extent greater than 30 arcsec. When we impose a limiting H\textsc{i} mass of 10\(^8\)M\(_\odot\) we find that 60 314 galaxies are detected, of these 2785 are larger in angular extent than 30 arcsec. This means that by constraining the semi-analytic catalogue to a limiting H\textsc{i} mass of 10\(^8\)M\(_\odot\), we underestimate the DINGO DEEP survey by approximately 4500 galaxies, or nearly 8 per cent. Of these missed objects the majority are unresolved, with only 1 per cent having an angular extent greater than 30 arcsec.

For the low-redshift WALLABY survey the full catalogue with no mass cut is estimated to be 672 660 galaxies and 471 129 of these are resolved by the 2-km baseline mode of ASKAP. When a limiting H\textsc{i} mass of 10\(^8\)M\(_\odot\) is assumed, the galaxy count drop to 631 409. Of these objects, 462 495 (522 981) are resolved by an angular beam of 30 arcsec. In other words, we are potentially underestimating the WALLABY catalogue by 40k galaxies, or 7.5 per cent similar to the case of DINGO DEEP, but nearly a quarter of these galaxies would be marginally resolved.

The predictions for the high-redshift DINGO UDEEP catalogue are essentially unchanged by setting the limiting H\textsc{i} mass of 10\(^8\)M\(_\odot\). Therefore, we caution the reader that the catalogue created in this work will likely underpredict the lower redshift DINGO DEEP and WALLABY surveys by 8 per cent, with the vast majority of these systems being unresolved point sources for the case of DINGO but as many as a quarter of this missing population marginally resolved in the WALLABY survey. With these caveats in mind we now consider the manner in which we can use the semi-analytic catalogue to create a realistic H\textsc{i} survey.

2.3 Creating H\textsc{i} surveys

The monochromatic luminosity, \(L_\nu\), from a cloud of H\textsc{i} (of mass \(M_{\text{HI}}\)) with a 21-cm line emission line profile \(\phi(v)\) (a narrow function of unit area, where the frequency \(\nu\) is close to the rest frequency, \(\nu_{12} = 1.42040575\) GHz) is

\[
L_\nu = \frac{3}{4} h \nu v_1 A_{12} \frac{M_{\text{HI}}}{m_\text{H}} \phi(v),
\]

where \(A_{12}\) is the Einstein coefficient for the H\textsc{i} spin-flip transition. We can relate this to the total integrated flux of the source in terms of the bolometric luminosity \(S_{\text{TOT}} = L_{\text{TOT}}/[4\pi d(z)^2]\) (equation 3.89 from Peacock 1999) at a luminosity distance of \(d_l = D(1 + z)\), where \(D\) is the comoving distance and \(L_{\text{TOT}} = \int L_\nu \text{d}v\). The total flux is the flux density, \(S_\nu\), integrated over frequency:

\[
S_{\text{TOT}} = \int S_\nu \text{d}v = \frac{3}{16\pi} \frac{h \nu v_1 A_{12}}{d_l^2} \frac{M_{\text{HI}}}{m_\text{H}} \int \phi(v) \text{d}v.
\]

This equation simplifies to

\[
S_{\text{TOT}} \approx \frac{M_{\text{HI}}}{\text{M}_\odot} 1 \text{ Jy Hz}^{-1} 49.8 \left(\frac{d_l}{\text{Mpc}}\right)^{-2}.
\]

In the case of a boxcar line profile, \(S_{\text{TOT}} = S_\nu \Delta v\), where \(\Delta v\) is in the observed frame of the galaxy. Note that a given rest-frame frequency or velocity width is measured by the observer to be \((1 + z)\) narrower (i.e. a galaxy will span fewer observed channels at higher redshifts for the same intrinsic rotation width). This becomes important when calculating the signal-to-noise ratio (S/N) of the observation in the following section.

We emphasize that this formula implicitly assumes an optically thin approximation for the H\textsc{i}, i.e. there is negligible self-absorption of the H\textsc{i} flux, in the galaxies themselves. This is a valid assumption for massive late-type systems with mean correction factors of 19 per cent for Sb, 16 per cent for Sbc and Sc, and 4 per cent for Sa and Sab galaxies found in Zwaan et al. (1997) after averaging over angle of inclinations from correction factors found in Haynes & Giovanelli (1984). In Haynes & Giovanelli (1984) smaller dwarf systems had a larger correction factor; however, the majority of the WALLABY and DINGO detections lie above these mass limits. Therefore, to a good approximation we can ignore the effects of inclination on the detectability of the galaxy in terms of total flux received from the galaxy. There are observational consequences of the inclination of the galaxy in terms of the observed velocity width and the resolving out of extended sources which we consider in Section 2.3.3 that are fully consistent with this optically thin assumption.
2.3.1 Detection limits

As detailed in Duffy et al. (2008, and references therein), the expected thermal noise for a dual-polarization single-beam, single-dish telescope is given by

$$\sigma_{\text{noise}} = \sqrt{\frac{k T_{\text{sys}}}{A}} \frac{1}{\sqrt{\Delta t / \Delta v}}$$

(5)

for an observing time of $\Delta t$ within a bandwidth $\Delta v$, which is assumed to be the ASKAP frequency resolution of 18.31 kHz (a velocity width $\Delta V = 3.86 \text{ km s}^{-1}$ at $z = 0$), where $k = 1380 \text{ Jy m}^2 \text{ K}^{-1}$ is the Boltzmann constant and $T_{\text{sys}}$ is the system temperature.

Thompson (1999) showed that for a dual-polarization interferometer, the noise is reduced by a further $\sqrt{2}$ if the area $A$ is taken to be that of a single element, i.e.

$$\sigma_{\text{noise}} = \frac{k T_{\text{sys}}}{A} \frac{1}{\sqrt{\Delta t / \Delta v}}$$

(6)

so that for a full interferometer with $N(N - 1)/2$ baseline permutations

$$\sigma_{\text{noise}} = \sqrt{2} \frac{k T_{\text{sys}}}{A_{\text{eff}}} \frac{1}{\sqrt{N(N - 1)\Delta t / \Delta v}}$$

(7)

The effective area of an ASKAP dish is the geometric area of a 12-m diameter dish, $a$, reduced by the aperture efficiency, expected to be $\alpha_{\text{eff}} \approx 0.8$ (Johnston et al. 2008). Therefore, equation (7) can be rewritten as

$$\sigma_{\text{noise}} = \sqrt{2} \frac{k T_{\text{sys}}}{A_{\text{eff}}} \frac{1}{\sqrt{\Delta t / \Delta v}}$$

(8)

where $A_{\text{eff}} = \alpha_{\text{eff}} a \sqrt{N(N - 1)}$. We assume $N = 30$ dishes within the 2-km core since, due to processing limitations, the full ASKAP array of 36 dishes is unlikely to be initially available. Note the well-known similarity between equations (5) and (8) for large $N$.

The flux density limit for an observation, $S_{\text{lim}}$, depends on the required S/N:

$$S_{\text{lim}} = \frac{(S/N) \sigma_{\text{noise}}}{\sqrt{N_{\text{ch}} \Delta v}}$$

(9)

A galaxy is less easily detected if its flux is spread across a number of frequency channels, resulting in lower flux density. For a boxcar profile of rest-frame velocity width $W$, a line profile will be spread over $N_{\text{ch}} = (W/\Delta V)(1 + z)$ channels, where $\Delta V$ is the $z = 0$ ASKAP velocity resolution of $3.86 \text{ km s}^{-1}$. As the noise is assumed to be Gaussian the uncertainty in our measurement of the mean flux density is reduced by the square root of the number of independent channels, or samples. Hence, for detection

$$S_{\text{TOT}} \text{(Jy Hz)} > \frac{(S/N) \sigma_{\text{noise}}}{\sqrt{N_{\text{ch}} \Delta v}} N_{\text{ch}} \Delta v$$

(10)

We can thus observe a galaxy if it has integrated flux greater than the integrated noise with a S/N cut-off:

$$S_{\text{TOT}} \text{(Jy Hz)} > \frac{(S/N) \sigma_{\text{noise}} \Delta v}{\sqrt{N_{\text{ch}}}}$$

(11)

where we note again that the frequency bandwidth $\Delta v$ is fixed by the ASKAP correlator to be $18.31 \text{ kHz}$ and that the number of channels an observed object spans will decrease as $(1 + z)$ due to the redshifting of the H I line, aside from any additional evolutionary effects.

2.3.2 Disc diameter

As considered in Duffy et al. (2012b) a potential issue when using interferometers with high angular resolution is the resolving out of galaxies more extended than the beam. With at least 2 km baselines assumed for the initial survey phase of ASKAP, we will certainly have to consider extended faint sources. We now summarize the procedure used in Duffy et al. (2012b) to determine this issue. Although we have accurate H I galaxy masses in the SAM catalogue, we do not have a well-defined disc size and therefore utilize an empirically derived relation between this mass and the observed H I diameter, $D_{\text{HI}}$ (defined to be the region inside which the H I surface density is greater than $1 \text{ M}_\odot \text{ pc}^{-2}$). From Broeils & Rhee (1997) and Verheijen & Sancisi (2001) we have

$$D_{\text{HI}} / \text{kpc} = \left( \frac{M_{\text{HI}}}{M_{\text{zoom}}} \right)^{\gamma}$$

(12)

where they find an index of $\gamma = 0.55$ and a normalization mass of $M_{\text{zoom}} = 10^{8.5} \text{ M}_\odot$. We then convert this diameter to an on-sky angular scale using the angular diameter distance $d_L(z)$. Note that this relation is of critical importance to our results and we are confident that this relationship is well supported as more recent observations by Noordermeer et al. (2005) have found identical best-fitting results to Broeils & Rhee (1997) and Verheijen & Sancisi (2001) for an entirely new galaxy sample, observed with a different telescope. Although these works are limited to higher column densities than might be probed by the deeper H I survey DINGO, we believe that it is still appropriate to use equation (12) as the majority of the H I signal will be from the high column density material described by this relation.

The distribution of the typical diameters of the galaxies found in the various ASKAP H I surveys after selection effects are taken into account is shown in Table 3 as well as the fraction of the galaxies in each survey which are resolved by ASKAP.

To test the sensitivity of our results to this relation, we have refit the Noordermeer et al. (2005) results for the case of a uniform average surface density of $1 \text{ M}_\odot \text{ pc}^{-2}$, i.e. $\gamma = 0.5$, and find a normalizing mass of $M_{\text{zoom}} = 10^{8.3} \text{ M}_\odot$ with only 10 per cent greater reduced $\chi^2$ than the case where $\gamma$ is unconstrained. In using this more theoretically motivated diameter–mass relation our overall number counts, after considering resolving of extended sources, decrease by less than 1 per cent in the case of DINGO, typically the detected galaxies are of the order of 1 kpc larger than the case where $\gamma$ is free to vary. We are therefore confident that this critical assumption of a fixed H I diameter–mass relation is both supported by the data and that small variations around this best-fitting solution make per cent-level modifications to the numbers quoted in Table 3 and hence can be ignored. Evolution in the H I diameter–mass relation as modelled in Obreschkow et al. (2009) was found to be negligible.

2.3.3 Telescope resolution

The angular resolution of ASKAP will be initially limited to 30 arcsec at 21-cm wavelength using the central 2-km core. Typically, for radio telescopes, the beam area increases like $\lambda^2 \propto (1 + z)^2$. However, this is not the case for the ASKAP phased array feeds which have a roughly fixed covering area as a function of redshift.

The effect of resolution on the detectability of galaxies depends on the efficiency and nature of galaxy-finding algorithms. Development of several algorithms is underway for ASKAP and preliminary testing has been undertaken by Popping et al. (2012). For the present purposes, we will make the assumption that the ASKAP source finders will be able to combine neighbouring independent pixels of a spatially resolved galaxy to improve S/N in the same manner as for a spectrally resolved galaxy (see equation 10). We therefore assume

© 2012 The Authors, MNRAS 426, 3385–3402

Monthly Notices of the Royal Astronomical Society © 2012 RAS
that we have optimal spatial smoothing of our sources in the sky plane.

The number of pixels the galaxy spans is approximated by the ratio of the galaxy on-sky area, $A_{\text{gal}}$, and the beam area $A_{\text{beam}}$, each of which is calculated below. In practice, the beam is convolved with the galaxy, meaning that even when a galaxy is just unresolved by the telescope (i.e. covers 1 pixel) we would still expect to lose signal. This convolution is represented by an additional factor of unity added to the number of beams, reducing the S/N by $\sqrt{2}$ in this matched case. Therefore the total flux of the galaxy, i.e. the left-hand side of equation (11), is reduced by the geometric factor $\sqrt{1 + A_{\text{gal}}/A_{\text{beam}}}$ (an example of this is shown in Fig. 3).

To calculate the galaxy area we approximate the true 3D galaxy profile by an elliptical cylinder with a spectral profile that is independent of position. We further assume a random inclination to the observer, $\theta$, where the case of $\theta = 0$ corresponds to face-on and $\theta = \pi/2$ to edge-on. The projected area of the galaxy is $\pi (D_{\text{H}i}/2)^2 (B/A)$, where A and B are the major and minor axes, respectively, the ratio of which, B/A, is equal to $\cos (\theta)$, although in calculating this we limit the smallest axis ratio for spirals to 0.12 in accordance with Masters, Giovanelli & Haynes (2003).

For the natural Gaussian antenna distributions described in Staveley-Smith (2006) and modelled in Gupta, Johnston & Feain (2008), the full width at half-maximum beam extent for ASKAP is $\Omega_{\text{FWHM}} \approx 1.42/2000 \text{ m}$. The beam area, $A_{\text{beam}}$, is therefore given by $\pi \Omega_{\text{FWHM}}^2 / (4 \ln 2)$, which we compare to the area of the galaxy [which has been multiplied by the square of the cosmological angular diameter distance, i.e. we have included $(1+z)^4$ dimming].

For the current design of ASKAP, with a 2–km baseline, this results in 12.5 per cent of galaxies in WALLABY being resolved out. This effect is particularly an issue at the high-$z$ end of the survey where galaxies are faint (see Fig. 4). Overall, despite being more spatially resolved, face-on galaxies are easier to detect in the ASKAP surveys considered here. This is shown in Fig. 5, where a small increase of face-on systems are recovered from the DINGO survey relative to the input uniform cosme distribution.

We note that a natural antenna weighting would result in a 20 per cent increase in flux sensitivity for the high-resolution case of ASKAP, with the additional six dishes being used. However, it is still undecided which weighting scheme will be employed and hence we conservatively assume the same flux limit for both the high- and low-resolution ASKAP images.

2.3.4 Spectral resolution

We calculate the velocity width, $W$, of the galaxy from the intrinsic line width, $W_e$, for a given angle of inclination $\theta$. The intrinsic line width of a galaxy, corrected for broadening, has been shown empirically to be related to the $H_1$ mass by (Briggs & Rao 1993; Lang et al. 2003)

$$W_e = \frac{420 \text{ km s}^{-1}}{W_{\theta}} = \left( \frac{M_{\text{HI}}}{10^{10} M_\odot} \right)^{0.3},$$

although we note that this relation shows a large dispersion, especially for dwarf galaxies. We have tested the consequences of this scatter by including an observed 0.2 dex lognormal dispersion in the velocities and find that 2 per cent more galaxies are detected (we henceforth ignore this negligible factor but note that for Tully–Fisher studies it will have to be modelled). The line width of a galaxy, $W_e$, which subtends an angle $\theta$ between its spin axis and the line of sight, can be computed using the Tully–Fouque rotation scheme (Tully & Fouque 1985):

$$(W_e \sin (\theta))^2 = W^2 + V_0^2 - 2 W V_0 \left[ 1 - e^{-\left( \frac{V_0}{100} \right)^2} \right] - 2V_{\theta}^2 e^{-\left( \frac{V_0}{100} \right)^2},$$

where $V_0 = 120 \text{ km s}^{-1}$ represents an intermediate transition between the small galaxies with Gaussian $H_1$ profiles in which the velocity contributions add quadratically and giant galaxies with a ‘boxy’ profile reproduced by the linear addition of the velocity terms. $V_\theta \approx 20 \text{ km s}^{-1}$ is the velocity width due to random motions in the disc (Rhee & van Albada 1996; Verheijen & Sancisi 2001).
In cases where $W_θ \gg V_c$, one can see that $W_θ = V_o + W_e \sin \theta$. For $\theta = 0$, one finds that $W_θ = V_o$, in other words the H_I dispersion in the disc, whereas for $\theta = \pi/2$ we recover $W_θ = V_o + W_e$ as expected.

The small additional broadening effect on the H_I profile due to the frequency resolution of the instrument (Bottinelli et al. 1990) is negligible for ASKAP and is ignored. Therefore for each galaxy we can uniquely assign an observed line width and angular extent on the sky for a random angle of inclination and its H_I mass (see Fig. 5).

3 WALLABY

With the large field of View of ASKAP it becomes feasible to scan the entire sky, in 8-h integrations, within several years, resulting in significant numbers of galaxies detected, $\sim$0.6 million as given in Table 4.

The effects of resolution on galaxy numbers, as shown in the high-resolution case in Table 4, cause a sharp drop in the detection rates. Nearly 0.3 million small, faint detections would now drop below the S/N threshold of the survey due to the resolving out of structure. As shown in Fig. 4, galaxies are lost throughout the entire redshift range of the survey, including marginally resolved systems at high redshift which now drop out of the survey entirely.

The creation of high-resolution images of low-resolution detections should clearly be limited to only the brightest sources; the exact threshold for which size and S/N detection should be rescaled is however entirely dependent on the performance of the source finder.

Furthermore source confusion, where multiple H_I sources overlap, is almost never an issue for WALLABY (we direct the reader to Section 4.1 where this technique is studied for the context of DINGO) in good agreement with our analytic estimates in Duffy et al. (2012b). Additionally we find that for WALLABY detections we can uniquely identify optical counterparts from a given optical photometric (spectroscopic) redshift catalogue to better than 97 per cent (99 per cent) as considered in Section 4.3 for the case of DINGO.

Although the proposed WALLABY survey extends to $z = 0.26$, the majority of galaxies detected at much lower redshifts ($z \sim 0.05$) and hence will have large angular extents on the sky, which can be expected to be well resolved by the 2-km baselines of ASKAP. The distribution of the detections as a function of proper physical diameter, for different redshift slices, is considered in the top panel of Fig. 6, with the on-sky angular extents of these systems in the middle panel and finally the number of ASKAP beams that these systems are resolved by in the bottom panel.

There are two interesting results from the top panel: the first is that the galaxies detected in a flux-limited survey are progressively larger in diameter (the solid vertical lines indicate the mean size of the sample) as this scales with H_I mass and hence flux. The second is that the WALLABY survey will detect orders of magnitude more galaxies within $z = 0.1$ than are detected in the redshift range beyond.

In the middle panel we consider the angular extent of all simulated galaxies in the WALLABY survey volume (in grey solid histogram) and the detected subset with and without resolution effects in black dotted and red dashed, respectively. This subset is typically several arcseconds in extent. The cumulative detection given by the red curve shows that at least 50 per cent of detections are larger than 30 arcsec. The blue vertical line indicates the resolution at 21-cm of the 2-km baseline of ASKAP; the intersection of this line with the cumulative curve indicates that the majority of ASKAP detections are resolved.

In the bottom panel of Fig. 6 we bin the galaxies by the number of ASKAP synthesized beams spanning the object. Although most (87.5 per cent) galaxies are only ‘marginally’ resolved with one or two beams, there are at least 5 per cent of all detections resolved by three beams (as given by the red curve) and 1 per cent of detections resolved with five beams. A non-negligible number of sources are extremely well resolved with over 30 beams across the galaxy.
The light cone distribution of these sources, in RA and redshift, is given in Fig. A1, with the cosmic web clearly visible across the volume surveyed, indicating that a wide range of environments, from clusters to voids, will be probed by the survey. This will in turn allow the measurement of the H I mass function as a function of environment, as well as various galaxy formation properties such as the stripping of star-forming gas within the hot haloes of clusters (e.g. Kilborn et al. 2009).

We now consider the type of galaxies that the WALLABY sample will contain in Fig. 7; this represents the true strength of using complex simulations to predict the galaxy catalogue as one can quantify the underlying galaxy population that a blind H I survey will probe. We consider several key galaxy properties, from top to bottom panels these are the total halo mass, H I mass, stellar mass and velocity width of the galaxies. From this figure it is clear that WALLABY can expect to detect systems ranging over four orders of magnitude in halo virial mass and seven orders of magnitude in stellar mass. The velocity width is approximately proportional to the square root of the mass enclosed and hence has a smaller dynamic range but we can expect to detect systems with velocities ranging from 20 to 1000 km s$^{-1}$. This broad selection of galaxies will be a unique resource to astronomers interested in galaxy formation as well offering powerful constraints on the nature of dark matter through structural analysis using the velocity widths of these systems.

4 DINGO

A great strength of ASKAP is the possibility of performing deep H I surveys to provide cosmologically representative samples of H I-rich galaxies beyond the local universe (>100 galaxies out to z = 0.43, as given in Table 4). The deeper H I surveys with ASKAP, termed DINGO, faces a unique challenge relative to the low-redshift WALLABY survey, namely the greater risk of confusion of H I sources at the increased distances which DINGO will probe relative to the case of WALLABY. We were able to ignore the issue of confusion for WALLABY (and have indeed checked that this assumption is valid) but must consider it for the deeper surveys before estimating the galaxy catalogues that DINGO will likely produce.

4.1 Source confusion

A potential issue in any galaxy survey is the ability to uniquely identify detections rather than artificially blending overlapping sources, in the sky plane and along the line of sight, as one object. In previous work (Duffy et al. 2012b), we argued with simple analytic calculations that the incidence of confusion was slight when one had a spectroscopic survey. With our galaxy catalogue we can make a much improved estimate by utilizing the actual on-sky angular diameter and velocity widths of the galaxies as well the realistic clustering of sources in space to determine whether they overlap as well as considering the improvements afforded by imaging with the 6-km baselines of ASKAP.

In Fig. 8 we consider the confusion rate, defined as the number of galaxies that overlap in angular projection (minimum on-sky angle is set by the telescope resolution, 30 arcsec or 10 arcsec) and along the line of sight divided by the total number of galaxies within a redshift slice $\Delta z = 0.02$. We note that we have an effective mass limit for this comparison of $M_{\text{lim}} = 10^{8.5} M_\odot$ at all times. If we considered galaxies below this H I mass limit, the confusion rate would naturally increase, yet little additional mass would be

Figure 7. In these plots we consider the underlying galaxy distribution, from top to bottom, as a function of total halo mass, H I mass, stellar mass and velocity width for a blind H I survey, WALLABY. As a result of the wide range in H I–halo mass ratios probed, Fig. 2, we expect the ASKAP surveys to probe stellar and halo masses over four orders of magnitude, as well as a wide range of halo velocity widths. We repeat this figure for the case of DINGO DEEP in Fig. 10.
Figure 8. We investigate the confusion rates of overlapping galaxies in both the DINGO DEEP and UDEEP surveys (left-hand and right-hand columns, respectively) with two possible dish configurations for ASKAP, a core design with only 2 km baselines and an extended 6 km case (giving 30 and 10 arcsec resolution shown in the top and bottom rows, respectively). To calculate the incidence rate of galaxies overlapping in redshift slices of width \( \Delta z = 0.02 \), we utilized the angular extent of the galaxies on the sky (if unresolved the angular diameter was fixed to be the telescope resolution) together with the velocity widths along the line of sight. We consider two cases of confusion; between all galaxies above a limiting H I mass \( M_{\text{HI}}^{\text{lim}} = 10^{8.5} \, M_\odot \), ‘All-All’ in black triangles, and the confusion rate of the detected galaxies in the survey with any galaxies, ‘Detect-All’ red diamonds. The former case is normalized by the number of galaxies in a redshift slice, and the latter by the number of detections in the same slice. Errors are Poissonian. The largest confusion rate is seen to be at the survey edge of the deepest survey, DINGO UDEEP (top right-hand panel), which reaches \( \sim 10 \) per cent if only 30 arcsec resolution is attainable. If a strategy of follow-up observations at 10 arcsec is adopted then this rate becomes negligible (\( < 3 \) per cent). Intriguingly, we note that the detected sources are more often confused than the average rate; this is discussed in the text.

Provided that the H I mass function has a faint-end slope sufficiently shallow that the total mass in systems below a given threshold mass converges, i.e. \( \alpha > -2 \), based on Zwaan et al. (2005) who find \( \alpha = -1.37 \pm 0.03 \pm 0.05 \) measured down to \( M_{\text{HI}} \approx 10^7 \, M_\odot \) which is sufficiently shallow for our conclusion to stand.

This is at first a surprising result as often H I sources are often seen as the most weakly clustered population (e.g. Meyer et al. 2007; Martin et al. 2012) and hence the ‘Detect’ sample, which will be typically be more massive than ‘All’, should be just as confused as ‘All’. The 0.3 dex increase is in fact a combination of three possibly effects.

The first is that more massive objects do lie within denser environments, and while H I is a weakly correlated tracer there are tentative results that overdensities are still detectable over ‘field’ galaxies, as shown for the case of the Fornax cluster in Waugh et al. (2002). The second factor is that the semi-analytic model of Croton et al. (2006) did not model proximity effects for infalling objects within overdense regions through mechanisms such as gas stripping, harassment or starvation. Therefore we should expect a higher clustering signal than otherwise expected making our conclusions about confusion limits for the ASKAP surveys a conservative case.

The final factor is numerical in nature and is because we can only track objects to \( M_{\text{HI}}^{\text{lim}} = 10^{8.5} \, M_\odot \). This means that although massive galaxies in our catalogue have resolved satellites, objects at this limit will not have their own satellites. We can test the severity of this issue by choosing a mass cut of \( M_{\text{HI}}^{\text{lim}} = 10^9 \, M_\odot \) and rerunning the analysis, in this case the difference is lessened (as might be
expected since proportionally more galaxies above this limit in the volume are detected) but only by 0.1 dex leaving the additional discrepancy a likely combination of the first two explanations.

Within the entire DINGO DEEP sample of detections 766 (350) galaxies, or a rate of 1.4 per cent (0.7 per cent), are confused with galaxies of mass greater than $M_{H_1} = 10^{8.5} M_\odot$ with the ASKAP 30 arcsec (10 arcsec) beam configuration. Of the detections that are confused with other detections, i.e. self-confused, the rate is of course far lower with only 62 (34) galaxies, or a rate of 0.11 per cent (0.07 per cent), with the ASKAP 30 arcsec (10 arcsec) beam configuration. With a deeper galaxy sample one might expect the DINGO UDEEP survey to suffer a greater incidence of confusion within a fixed telescope beam resolution, and indeed 3971 (1015) galaxies, a rate of 4.9 per cent (1.3 per cent), suffer confusion for the 30 arcsec (10 arcsec) ASKAP beam. The self-confusion incidence in UDEEP is also slightly higher than in DEEP, with 406 (118) galaxies overlapping with other detected galaxies, a rate of 0.5 per cent (0.15 per cent).

### 4.2 H\textsc{i} detections

With the previous result that the overall number counts in DINGO will be only slightly affected by the issue of source confusion, we can consider the distribution of the detections as a function of redshift in Fig. 9 for the DEEP and UDEEP aspects of DINGO. We see immediately that the surveys complement each other, with the increased observing time over smaller areas enabling the UDEEP part of DINGO to extend the H\textsc{i} detections to $z = 0.43$. Note that there are several orders of magnitude more galaxies undetected in the survey volume with $M_{H_1} = 10^{8.5} M_\odot$. These can be stacked to create a measurable signal from individually undetected sources (after cross-correlating with optical surveys, the issue of identifying these galaxies is considered in Section 4.3).

The properties of the detected galaxies in DINGO are similar to that probed by WALLABY, with the caveat that the simulations only contain galaxies with $M_{H_1} \geq 10^{8.5} M_\odot$ and hence the likely low-mass systems in the latter are missing. In Fig. 10 we consider the total halo mass, H\textsc{i} mass, stellar mass and intrinsic rotational width of the sample of galaxies detected by DINGO DEEP. The survey detects four orders of magnitude in total mass, seven orders of magnitude in stellar mass and nearly two orders of magnitude in velocity width.

A significant science case for large area H\textsc{i} surveys is to measure the H\textsc{i} mass function; as is shown in Fig. 11, the ASKAP telescope will create a complete mass function above $M_{H_1} \approx 10^{10} M_\odot$ with the DINGO survey. This will enable the measurement of evolution in the high-mass end of the mass function across the entire redshift range, and to progressively lower masses for lower redshift ranges.

We now consider the physical extent of the sources detected by DINGO in Fig. 12 which are typically of smaller physical extent (top panel) at a given redshift than the corresponding WALLABY sources as a result of the greater integration time per field which allows fainter (and hence physically smaller) galaxies to be detected. As shown in the middle panel of Fig. 12 the majority of galaxies are unresolved (comparing the intersection of the cumulative red curve with the ASKAP 30 arcsec beam denoted by the vertical blue line). If we create a histogram of the number of synthesized beams the galaxies are resolved by (bottom panel), we see that less than a few per cent of galaxies are just resolved by more than one beam (DEEP and UDEEP resolve ≤9 and 1 per cent of galaxies, respectively).

We consider the case of observing these systems with the 10 arcsec synthesized beam when using 6-km baselines, as shown in Table 3, and find that now 95 per cent of objects are marginally resolved. Although not a key science case for DINGO there could be as many as 100 galaxies resolved by more than 10 beams which will be a valuable data set to compliment the WALLABY sample.

We illustrate the strength of DINGO as a deep survey instrument in Fig. 11 in which the large-scale structure is clearly visible out to the edge of the survey at $z \sim 0.43$ allowing the H\textsc{i} properties of galaxies to be studied as a function of both environment and redshift. These light cones are presented as one contiguous field, while in reality the DINGO fields are distributed across the sky, thereby reducing the effects of cosmic variance.

### 4.3 Matching optical catalogues

We can use our catalogue to estimate typical misidentification rates when uniquely assigning optical counterparts to our H\textsc{i} detections. This measurement differs slightly from the confusion rate discussed before. For the case where there are, say, two (three) possible counterparts one can always blindly choose one for the optical counterpart, statistically speaking, in this case we will be correct 50 per cent (33 per cent) of the time. If one considered priors such as the size of
In these plots we consider the underlying galaxy distribution, from top to bottom, as a function of total halo mass, H\(_I\) mass, stellar mass and velocity width for a deep H\(_I\) survey, DINGO DEEP. The range of properties probed is similar to WALLABY, given in Fig. 7.

These systems then we can be substantially more reliable than simple blind guessing. Therefore the misidentification rate is slightly lower than the confusion rate due to the possibility of randomly assigning objects from the list of possible counterparts.

In Duffy et al. (2012b) we considered two optical catalogues: one with photometric redshifts (a typical redshift error for such galaxies is \(\Delta z = 0.05\) as argued by Hildebrandt, Wolf & Benítez 2008) and a spectroscopic redshift sample. We estimated that in the latter case, the uncertainty as to which galaxy was the counterpart was when two galaxy rotation widths overlapped, hence we conservatively assumed \(\Delta z = 0.002\) (twice the width of an \(M^*\) galaxy which is the likely system detected by ASKAP at high redshift\(^8\)).

4.3.1 DINGO DEEP

In Fig. 13 we consider the case for the DINGO DEEP (UDEEP) survey in the left-hand (right-hand) columns. For both telescopes we consider a resolution of 30 arcsec in the top panel and the high-resolution imaging at 10 arcsec in the bottom panel. For each case we consider an optical catalogue with only photometric redshifts (\(\Delta z = 0.05\)) and one with spectroscopic redshifts (\(\Delta z = 0.002\)), results given by the red square and black diamond points, respectively. We find that the H\(_I\) galaxies surveyed with the 30 arcsec ASKAP beam at \(z = 0.05\) can be uniquely identified with a single optical counterpart, with photometric errors, more than 96 per cent of the time (top left-hand panel). However, the misidentification rate strongly rises to \(\sim 18\) per cent at the survey edge, \(z = 0.26\). If one has access to a spectroscopic survey then even at the edge of the survey optical counterparts can be correctly identified 95 per cent of the time.

---

\(^8\) Although individual velocity widths of the galaxies are available, and indeed were used in the confusion study of Section 4.1, we adopt the limiting spectroscopic width in accordance with previous estimates from Staveley-Smith (2008) and Duffy et al. (2012b).
at low redshift and it is therefore unsurprising that the misidentification rate with an optical catalogue is also much higher than in the case of the lower redshift DINGO DEEP.

We find that if one only has access to the 30 arcsec DINGO UDEEP source list and attempt to uniquely identify counterparts from a photometric catalogue (red squares top right-hand panel of Fig. 13), then at the beginning (end) of the survey the misidentification rate is 7 per cent (31.5 per cent). If using a spectroscopic optical catalogue (black diamonds top right-hand panel of Fig. 13), then the rates are 1–12 per cent across the survey.

As before if one images the field at 10 arcsec (bottom right-hand panel of Fig. 13) then even with a photometric catalogue the misidentification rate is never higher than 6 per cent (reached at $z = 0.43$) and when combined with a spectroscopic optical survey this would drop to 0 of the order of a percentage.

In conclusion, the rates of confusion for H I galaxies are negligible with ASKAP provided that one has access to the highest possible resolution of 10 arcsec synthesized beams. Key science goals that rely on the identification of the optical counterparts to H I sources can make use of photometric catalogues, suffering from less than 9 per cent misidentification rates (even out to $z = 0.43$ for the case of DINGO UDEEP) if one has the high-resolution imaging.

If ASKAP synthesized beams of 30 arcsec are used then DINGO DEEP can still identify an optical counterpart at $z = 0.05 (0.26)$ using a photo-z catalogue more than 96 (82.5) per cent of the time. For DINGO UDEEP the photo-z catalogue will suffer a misidentification rate of 5 (30) per cent at the start (end) of the survey. If a spectroscopic catalogue is available then even with 30 arcsec resolution the UDEEP sources will be uniquely identified at $z = 0.43$ 88 per cent of the time.

In conclusion, we find that the quality of the optical catalogue depends strongly on the ASKAP resolution. For DINGO DEEP and UDEEP science goals that can accommodate misidentification rates of the order of 5 per cent then a photometric redshift catalogue is sufficient with 10 arcsec resolution imaging from ASKAP. However, for certain science goals such as H I spectral stacking, spectroscopic optical catalogues are essential, irrespective of ASKAP resolution.

5 COMPARISON OF SURVEYS

We combine the overall galaxy count distribution as a function of redshift for all three surveys (WALLABY, DINGO DEEP and UDEEP) in Fig. 14 which clearly demonstrates the ‘wedding cake’ design of successfully deeper, overlapping, surveys which produce significant galaxy samples of the order of $10^5$ systems per redshift bin across the entire redshift range $0 < z < 0.43$ of ASKAP.

The WALLABY survey will be able to find $\sim 6 \times 10^5$ galaxies at S/N = 5 which will be nearly two orders of magnitude more galaxies than the current H I survey all-sky surveys have been able to find. We have modelled the number counts as a function of S/N in Table 4, with a smooth drop in galaxy detections for more demanding thresholds. DINGO may, through the DEEP and UDEEP tiered approach, detect $\sim 10^5$ galaxies over the redshift range 0–0.43, covering 5 billion years of cosmic evolution.

We also considered a Northern hemisphere H I survey using the Westerbork array called WNSHS which was mid-way in flux sensitivity between WALLABY and DINGO DEEP. This facility can potentially detect $\sim 2 (5) \times 10^5$ galaxies dependent on the depth to which it surveys, making it an incredibly powerful data set in its own right. However, if one were to combine both WALLABY and WNSHS the potential science increases with a combined catalogue of over a million H I sources with integrated S/N of at least 5 with
Figure 13. We have calculated the misidentification rate of galaxies when assigning an optical counterpart to H\textsc{i} detections for the 30 and 10 arcsec resolutions of ASKAP (top and bottom panels) in the DINGO DEEP and UDEEP surveys (left-hand and right-hand columns, respectively). We considered two typical redshift errors that an optical catalogue would suffer; the first in red squares is a photometric redshift uncertainty of the position of the optical counterpart position (as argued by Hildebrandt et al. 2008 this is typically $\Delta z = 0.05$) and in black diamonds we consider a spectroscopic survey with typical uncertainties assumed to be twice the velocity width of an $M^*$ galaxy ($\Delta z = 0.002$) which will be detectable throughout the redshift range of the survey (Duffy et al. 2012b). The rates are normalized by the number of galaxies within the redshift slice, with Poissonian errors for each point. Within the greater volume of the ASKAP beam and optical photo-$z$ errors there are more potential optical counterparts. At the edge of the DINGO UDEEP survey, more than 30 per cent of galaxies will be misassigned. If there are follow-up observations at 10 arcsec the misidentification rate drops to $\sim 3$ per cent ($\leq 10$ per cent) for DINGO DEEP (UDEEP).

Figure 14. The expected galaxy number density in redshift bins of width $\Delta = 0.02$ for the DINGO DEEP, UDEEP and WALLABY surveys. For DINGO UDEEP we take the conservative case that there is no evolution in the H\textsc{i} mass function across the redshift probes, and for WALLABY we reduce the number of galaxies found by an order of magnitude for ease of comparison with the deeper surveys. Note the ‘wedding cake’ design of these overlapping surveys which ensure that significant numbers of galaxies are probed at each redshift.

spectroscopic redshifts and a full $4\pi$ sky coverage. This will enable bulk flow studies and velocity field probes using a contiguous, blind data set which will dramatically improve the power of this promising probe of cosmology and structure formation.

6 CONCLUSION

ASKAP is a uniquely powerful telescope, capable of cataloging $6 \times 10^5$ galaxies in the local Universe and probing the distribution of H\textsc{i} out to $z = 0.43$ with $5-6 \times 10^4$ galaxies (dependent on evolutionary possibilities), creating a more complete view of H\textsc{i} than ever before. The systems we expect to find will probe four orders of magnitude in total halo mass and seven orders of magnitude in stellar mass. This will greatly aid our understanding of galaxy formation as well as the nature of dark matter haloes. Furthermore, a subsample of the nearby detections are well resolved by ASKAP, with $\sim 6000$ ($\sim 700$) galaxies resolved by $>5$ ($>10$) beams, providing a valuable resource for the dynamical modelling of galaxies in 3D. However, if the full 6-km baseline for ASKAP is available and the galaxies are surveyed with 10 arcsec resolution then significantly more systems are well resolved, $\sim 1.5 \times 10^5$ ($2 \times 10^4$) galaxies resolved by $>5$ ($>10$) beams. This resolved catalogue can then probe the small-scale clustering of the dark matter and the influence that the baryons have at such scales (e.g. Duffy et al. 2010). As well as a crucial aid to studying the velocity structure of nearby objects the postage-stamp zoom-in’s around galaxy detections are crucial to lowering the confusion rate of the distant galaxy population to sub-per cent levels for all spectroscopic surveys with ASKAP. The postage stamps will also be crucial in making possible unambiguous assignments of optical counterparts to the H\textsc{i} detections, a key science case for multiwavelength surveys such as GAMA (Driver, Hill & Kelvin 2011).
As a community resource we make available catalogues for both surveys with the sky coordinates (RA and Dec.), redshift and observed redshift (i.e. including peculiar motion) as well as the stellar, halo, cold gas and H\textsc{i} mass and finally the velocity width of the system. This catalogue will be invaluable in guiding survey preparations for Tully–Fisher studies, velocity field probes as well as correlation function investigations to best utilize the formidable resource that ASKAP will represent for a wide range of astronomical fields.

ACKNOWLEDGMENTS

We would like to thank Martin Zwaan for his excellent comments in refereeing the paper. This work makes extensive use of the COYOTE\textsuperscript{\small{©}}\textsuperscript{\tiny{C}} libraries and ARD is grateful to David Fanning in creating such a fantastic (free) resource. Additionally ARD would like to thank Chris Power, Peder Norberg and Simon Driver for helpful discussions. DJC acknowledges receipt of a QEI Fellowship. ARD gratefully acknowledges the use of computer facilities purchased through a UWA Research Development Award. The Millennium Simulation used in this paper was carried out as part of the programme of the Virgo Consortium on the Regatta supercomputer of the Computing Centre of the Max-Planck-Society in Garching, and the Cosmology Machine supercomputer at the Institute for Computational Cosmology, Durham. The mock galaxy catalogues used in this work are available on request and were generated using the Theoretical Astrophysical Observatory (see http://tao.it.swin.edu.au/mock-galaxy-factory). High-resolution images and movies are available at http://ict.icrar.org/store/Movies/Duffy12c/.

REFERENCES

Abate A., Bridle S., Theodorou L. F. A., Warren M. S., Hendry M., 2008, MNRAS, 389, 1739
Abdalla F. B., Rawlings S., 2005, MNRAS, 360, 27
Altay G., Theuns T., Schaye J., Crighton N. H. M., Dalla Vecchia C., 2011, ApJ, 737, L37
Barnes D. G. et al., 2001, MNRAS, 322, 486
Beutler F. et al., 2012, MNRAS, 423, 3430
Beutler F. et al., 2011, MNRAS, 416, 3017
Broeils A. H., Rhee M.-H., 1997, A&A, 324, 877
Bottinelli L., Gouguenheim L., Fouque P., Paturel G., 1990, A&AS, 82, 391
Briggs F. H., Rao S., 1993, ApJ, 417, 494
Broeils A. H., Rhee M.-H., 1997, A&A, 324, 877
Burkey D., Taylor A. N., 2004, MNRAS, 347, 255
Chippendale A. P., O’Sullivan J., Reynolds J., Gough R., Hayman D., Haynes M. P., 2003, ATNF SKA Memo Series, 16
Chippendale A. P., O’Sullivan J., Reynolds J., Gough R., Hayman D., Haynes M. P., 2010, in IEEE Int. Symp., phased Array Systems and Technology (ARRAY), p. 648
Croton D. J. et al., 2006, MNRAS, 365, 11
Deboer D. R. et al., 2009, IEEE Proc., 97, 1507
Driver S. P. et al., 2009, Astron. Geophys., 50, 5.12
Driver S. P., Hill D. T., Kelvin L. S., 2011, MNRAS, 413, 971
Duffy A. R., Battye R. A., Davies R. D., Moss A., Wilkinson P. N., 2008, MNRAS, 383, 150
Duffy A. R., Schaye J., Kay S. T., Dalla Vecchia C., Battye R. A., Booth C. M., 2010, MNRAS, 405, 2161
Duffy A. R., Kay S. T., Battye R. A., Booth C. M., Dalla Vecchia C., Schaye J., 2012a, MNRAS, 420, 2799
Duffy A. R., Moss A., Staveley-Smith L., 2012b, PASA, 29, 202
Gupta N., Johnston S., Feain I., 2008, ATNF SKA Memo Series, 16

APPENDIX A: H\textsc{i} SURVEY LIGHT CONES
Figure A1. The light cone pie-plot for the shallow all-sky survey WALLABY and the two deeper surveys from DINGO. We have taken the entire declination range of the survey and projected it on to a redshift–RA 2D plot. Brightness is based on the number density of the sources in each map pixel. The cosmic web is clearly visible in this image, a key science driver of WALLABY is the measurement of the H\textsc{i} mass function in different environments. In this volume are over 0.6 million galaxy detections, which can be compared with HIPASS and ALFALFA which had approximately two orders of magnitude fewer detections. The DINGO survey will probe evolution in the high-mass end of the H\textsc{i} mass function over 4 billion years of cosmic time. Additionally, DINGO will overlap with existing GAMA fields to enable a wealth of multiwavelength data to be used when analysing the H\textsc{i} detections; as well as enabling H\textsc{i} spectral stacking at given optical spectroscopic redshifts to extend the H\textsc{i} detections. We reiterate that the DINGO fields are not actually contiguous as pictured here, but in fact are spaced across the sky. For the mock light cone we made the simplifying assumption that the fields were contiguous, this had no impact on the final number of galaxies predicted. High-resolution versions of this image and fly-through movies are available at http://ict.icrar.org/store/Movies/Duffy12c/.

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