Peering towards Cosmic Dawn
Proceedings IAU Symposium No. 333, 2018
V. Jelić & T. van der Hulst, eds.

doi:10.1017/S1743921318000625

Extragalactic optical and near-infrared foregrounds to 21-cm epoch of reionisation experiments

Matt J. Jarvis[1,2], Rebecca A.A. Bowler[1] and Peter W. Hatfield[1]

[1]Oxford Astrophysics, Denys Wilkinson Building, Keble Road, Oxford OX1 3RH, UK
email: matt.jarvis@physics.ox.ac.uk

[2]Astrophysics Group, Department of Physics, University of the Western Cape, Private Bag X17, Bellville 7535, South Africa

Abstract. Foreground contamination is one of the most important limiting factors in detecting the neutral hydrogen in the epoch of reionisation. These foregrounds can be roughly split into galactic and extragalactic foregrounds. In these proceedings we highlight information that can be gleaned from multi-wavelength extragalactic surveys in order to overcome this issue. We discuss how clustering information from the lower-redshift, foreground galaxies, can be used as additional information in accounting for the noise associated with the foregrounds. We then go on to highlight the expected contribution of future optical and near-infrared surveys for detecting the galaxies responsible for ionising the Universe. We suggest that these galaxies can also be used to reduce the systematics in the 21-cm epoch of reionisation signal through cross-correlations if enough common area is surveyed.

Keywords. galaxies: evolution, galaxies: high-redshift, cosmology: early universe, cosmology: large-scale structure of the universe

1. Introduction

Understanding the extragalactic foregrounds, in order to obtain the cleanest signal from 21-cm emission in the epoch of reionisation (EoR), is a key issue for current and future low-frequency radio telescopes. The foregrounds that are considered are usually just those based on the radio information, imaged either directly from the low-frequency radio telescopes themselves, or from other higher frequency radio surveys. These provide flux information and the spatial distribution of discrete radio sources, which help in subtracting such foreground sources from the EoR signal. However, much more information is present in these surveys, including the clustering information and the flux-distribution (or source counts), as a function of redshift.

Furthermore, very little information is currently used from surveys beyond the radio waveband. As we move towards the era where radio surveys are becoming ever more sensitive with relatively high resolution, then this is understandable, as one might imagine that all of the relevant information for the removing foreground is present in these data. However, one could also take the view that if the information is available, then why not use it. This may be a particularly profitable avenue to explore as we move towards deeper and deeper observations at low-frequencies to detect the EoR HI power spectrum, but do not have the necessary spatial information at radio wavelengths to fully understand the extragalactic foreground emission.

In these proceedings we highlight some possible ways to incorporate information from multi-wavelength (non-radio surveys) in order to help mitigate the foreground contamination in current and future EoR surveys. We also highlight the gains that can be made

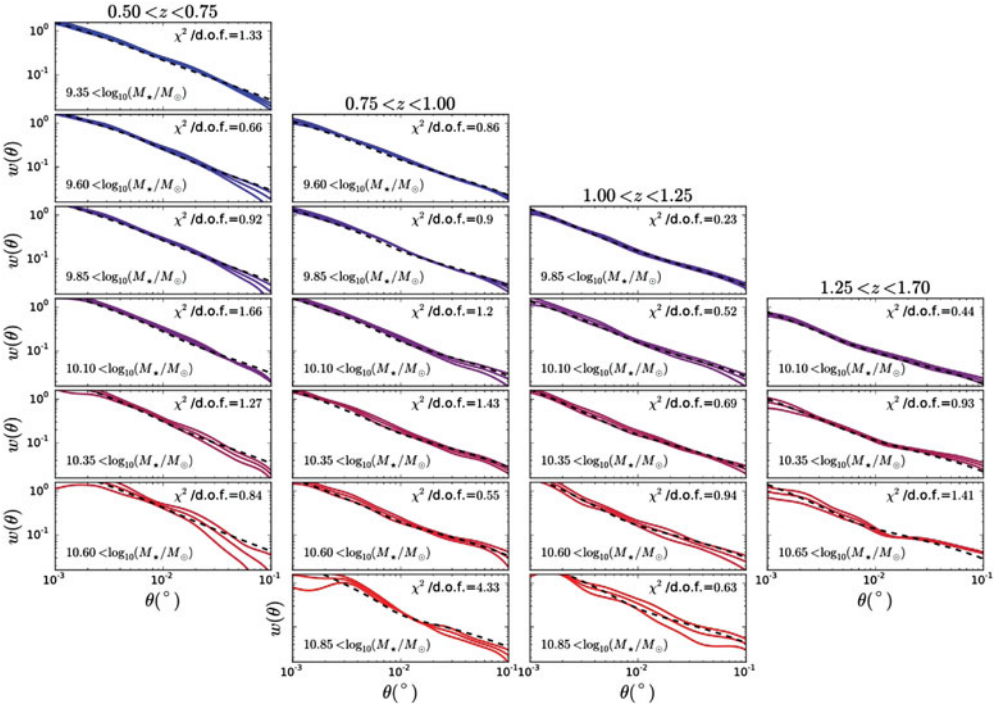

Figure 1. Two-point correlation function for galaxies selected from the VIDEO survey as a function of stellar mass and redshift (taken from Hatfield *et al.* 2016).

by cross-correlating the EoR signal with galaxies discovered within the EoR from ground- and space-based optical and near-infrared surveys.

2. Clustering of extragalactic foregrounds

The galaxy correlation function, or power spectrum, can be considered as being comprised of a 1-halo term, which probes the excess pairs of galaxies within a dark matter halo, and a 2-halo term which traces the excess pairs of galaxies between dark matter haloes. This transition between the intra-halo and inter-halo clustering has been widely modelled using the halo-occupation distribution (HoD) model (e.g. Cooray & Sheth 2002, Zheng *et al.* 2005). The shape of the galaxy correlation function is also imprinted on the extragalactic foregrounds of the EoR. This foreground contamination at radio wavelengths predominantly arises from synchrotron emission from active galactic nuclei and star-forming galaxies. These galaxies predominantly lie at $1 < z < 3$, the peak of both the star-formation rate density (e.g. Madau & Dickinson 2014), and the AGN-accretion-related luminosity density (e.g. Aird *et al.* 2015). Unfortunately, using the low-frequency radio data to directly detect and model the extragalactic radio source population to the required depths is hindered by confusion, thus higher resolution data at radio wavelengths is advantageous (see e.g. Hale *et al.* 2018), but much could also be learnt from optical and near-infrared (and other wavelength) data if it can be used as a prior on the expected radio emission.

Traditional radio surveys have in the past been sensitive to the AGN population, and there are several studies which show correlations between accretion rate and the ability to form radio jets (e.g. Rawlings & Saunders 1991, Willott *et al.* 1999, Hardcastle

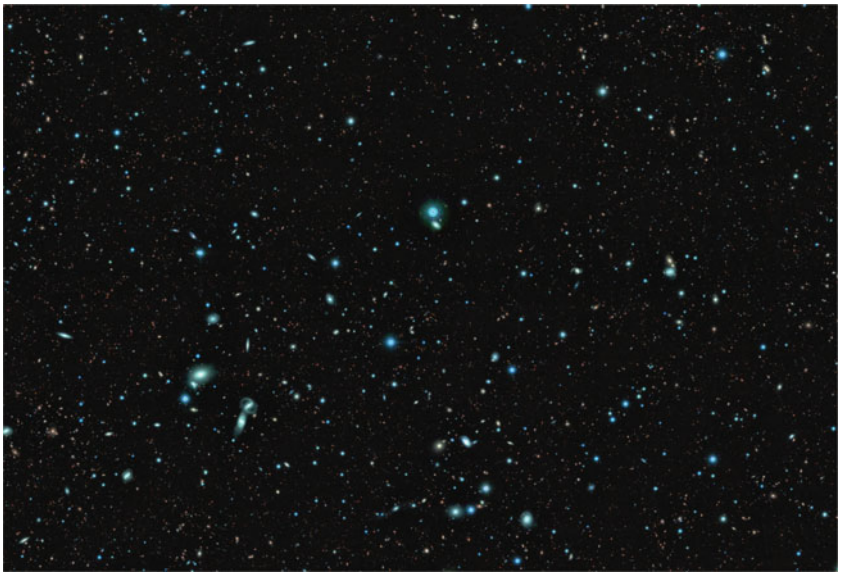

Figure 2. Near-infrared YJK_s-colour image of ~ 400 arcmin^2 of the VIDEO-XMM-LSS field.

et al. 2006, Fernandes *et al.* 2011), and thus produce synchrotron emission. However, the accretion rate only maps onto observed optical luminosity for the unobscured quasar population, whereas if orientation-based unification is valid, then around 50 per cent of the population would be radio bright but optically faint, due to the optical emission only arising from the host galaxy stellar population. Moreover, it is now clear that the radio emission is not directly related to the accretion rate in the regime where the accretion is inefficient (e.g. Heckman & Best 2014). Fortunately, at the depth of low-frequency EoR data, the source counts are dominated by star-forming galaxies (e.g. Condon 1992, Wilman *et al.* 2008) and not AGN. As such it is possible to relate the optical emission to that expected at radio wavelengths.

It has been know for many years that there is a very tight correlation between the far-infrared and the radio emission (e.g. van der Kruit 1971, Condon *et al.* 1991, Yun *et al.* 2001, Jarvis *et al.* 2010). Under the assumption that the far-infrared emission is the result of the obscured star-formation, and the radio emission is due to synchrotron radiation resulting from supernova remnants, then both wavelengths trace the underlying star-formation rate within galaxies. More recently, many studies have utilised the extensive multi-wavelength data over large areas of the sky to determine the full spectral energy distribution of galaxies, from the UV through to the far-infrared (e.g. Smith et al. 2012). Such data provide a robust measurement of the total star-formation rate within galaxies, both unobscured from the UV, and obscured from the far-infrared.

Radio surveys covering this plethora of multi-wavelength data have subsequently been used to show that the radio emission is directly related to the total star formation rate in the galaxies (e.g. Davies et al. 2012). Given this, it is feasible to not only measure the expected contribution of star-forming galaxies to the total radio flux in EoR deep fields, but also determine the angular clustering, and use such information on the prior for the foreground contamination.

In Fig. 1 we show the measured angular correlation function for galaxies selected from the VISTA Deep Extragalactic Observations (VIDEO) Survey (Fig. 2; Jarvis *et al.* 2013). This shows that the shape and amplitude of the correlation function is dependent

on both the mass of the galaxy, as well as the redshift. However, the star-formation is related to the galaxy mass for the majority of galaxies that lie on the star-formation main sequence (e.g. Noeske *et al.* 2007, Whitaker *et al.* 2012, Johnston *et al.* 2015), as such some information about the spatial clustering of the extragalactic foregrounds can be obtained from data of this kind.

More informative however, would be a direct tracer of the star-forming and star-burst galaxies that produce the bulk of the radio emission in deep radio surveys. Such information can be obtained from deep rest-frame UV and/or deep far-infrared surveys. However, far-infrared surveys are generally only sensitive to the most extreme starbursting galaxies at high redshift, due to having low angular resolution and suffering from confusion. Although ALMA overcomes this, the ability to survey large areas is very limited. As such the information on the expected radio emission is limited or incomplete. However, full spectral energy distribution modelling from the optical through to the near infrared, to estimate star-formation rates allows information on the expected radio emission from such galaxies to gained.

Thus measuring the clustering of star-forming galaxies from optical and near-infrared surveys naturally informs on the expected clustering signal from the faint radio sources that lie in the foreground to the EoR 21-cm signal.

It also worth noting that there is likely to be significant cross-clustering between bright radio-loud AGN and the fainter star-forming populations, particularly at $z > 1$. This is because radio-loud AGN have been known for many years to trace some of the most overdense regions of the high-redshift universe. The overdensities or proto-clusters can also contain a large number of star-forming galaxies, i.e. there is additional information about the 1-halo term in the cross-correlation (e.g. Hatfield *et al.* 2017). Therefore, an additional step in accounting for the clustering of foreground galaxies, emitting at radio wavelengths could be made by considering such cross-correlations. An example of the cross-correlation signal between passive and star-forming galaxies can be seen in Fig. 3, where it is found that the low-mass star-forming galaxies at $z > 1$ tend to avoid the dark matter haloes that contain a high-mass passive galaxy at their centre. This kind of effect is present on the angular scales over which EoR experiments aim to detect the 21-cm power spectrum, and thus could imprint false signals on the EoR measurement unless properly accounted for.

A possible way forward in this area, and how to include such information in analyses of the 21-cm power spectrum could also come by determining the conditional luminosity function, or possibly more informative for EoR experiments, the conditional flux-density function. Where a formalism to provide a likelihood that a galaxy of a given flux or luminosity occupies a halo of a given mass. Such work could then directly feed into the expected clustering signal of all of the extragalactic foregrounds to the EoR.

3. Galaxies within the Epoch of Reionisation

The vast majority of current experiments to detect the 21-cm EoR signal rely on measuring the power spectrum or correlation function in relatively narrow and discrete spectral bands, or redshifts. As such, dedicated low-frequency radio telescopes with reasonable bandwidth and processing power can assess the impact of noise in a given spectral bin by cross-correlating with spectral bins that are relatively nearby, where one would not expect any cosmological signal, i.e. a signal at $z \sim 8$ should not correlate with the signal at $z \sim 8.2$. However, this may not mitigate broad band noise, or issues that affect different sub-bands in the same way. To overcome these issues, then cross-correlating observations from different telescopes should significantly improve matters, as one would

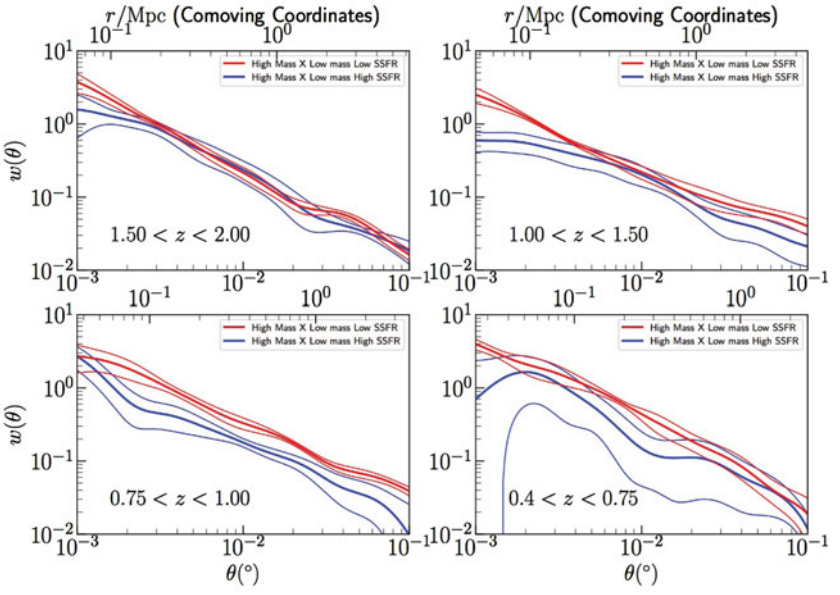

Figure 3. The cross-correlation function signal between low mass ($10.4 < \log_{10}(M_{min}/M_\odot) < 10.9$) and high mass ($10.9 < \log_{10}(M_{min}/M_\odot) < 11.4$) galaxies in VIDEO in four redshift bins. The low mass sample is selected to be passive ($\log_{10} sSFR < -11$) for the red curves, and star forming ($\log_{10} sSFR > -11$) for the blue curves. At the highest redshift, the two curves are very similar; the role of environment in determining star formation rate seems to be small. At $1 < z < 1.5$ the passive low-mass galaxies are more clustered around the massive galaxies, but only on small scales. In the $0.75 < z < 1$ and $0.4 < z < 0.75$ bins, the enhancement has reached the larger scales, and is substantial on all scales. (Adapted from Hatfield & Jarvis 2017).

not expect different telescopes to suffer exactly the same systematic telescope-related problems.

This could be achieved by cross-correlating signals from different 21-cm experiments, if the noise-related issues were just telescope dependent. Or one could cross-correlate signals from different observing runs, thus allowing possible removal of time-varying ionospheric noise. However, a completely independent tracer of the EoR are the galaxies that essentially supply the photons that contribute to reionising the Universe.

The past decade has seen a huge increase in the number of galaxies discovered at $z > 6$. Deep imaging surveys combining optical and near-infrared imaging from the Hubble Space Telescope have paved the way, enabling tight constraints on the shape of the faint end of the galaxy luminosity function (e.g. McLure *et al.* 2013, Bouwens et al. 2015, Finkelstein *et al.* 2015). Measuring the bright end of the $z > 6$ galaxy luminosity function requires larger areas of sky than is feasible with HST. The surveys conducted with the sensitive, wide-field near-infrared cameras on ground-based 4- and 8-m class telescopes, such as UKIRT-WFCAM and VISTA, have pushed this field forward (e.g. Bowler et al. 2015).

Given that the current and planned EoR experiments aim to survey areas of 10s to 1000s of square degrees on the sky, cross-correlating 21-cm experiments with HST-sized surveys will not provide the necessary information. It is the wider shallower surveys that can play a significant role in this area.

There are two principal methods of finding galaxies at high redshift, the first relies on colour selection in broad-band filters, selecting galaxies which have a sharp break in

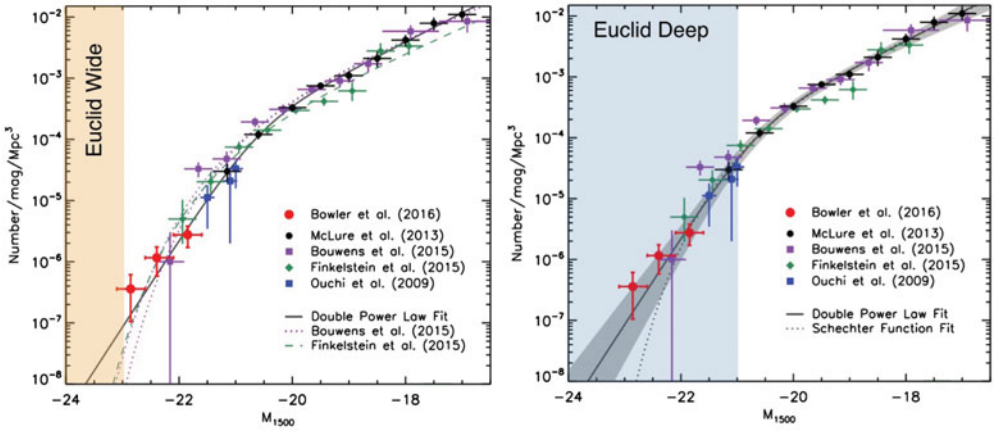

Figure 4. Current constraints on the $z \sim 7$ galaxy luminosity function from the studies listed in the legend. The left (right)-hand panel shows the magnitudes that will be accessible with the wide (deep) surveys with *Euclid*. (Figure adapted from Bowler *et al.* 2015).

the spectral energy distribution shortward of the Lyman break (e.g. Guhathakurta *et al.* 1990). The second relies on identifying plausible line-emitting galaxies in narrow-band filters (e.g. Ouchi et al. 2010). At high redshift the line that is most reliable is Lyα, due to its large equivalent width and its accessibility to optical and near-infrared telescopes at high redshift. Both offer powerful means to identify the highest-redshift galaxies. However, they both have strengths and weaknesses. Selecting Lyman-break galaxies at $z > 6$ requires significantly deeper data shortward of the Lyα line to remove low-redshift contaminants. Narrow-band selection of line emitters relies on broad-band photometry to enable identification of a spectral break shortward of the line and to estimate the equivalent width. However, contamination by galaxies at lower redshifts, where longer rest-frame wavelength lines fall in the narrow band filter, hinder forming completely reliable samples. Furthermore, due to the narrowness of the filters, only relatively small redshift ranges can be explored in a given filter. A third method of performing blind spectroscopy with slitless spectroscopy may provide additional information, but the required depth coupled with confusion issues may restrict its usefulness.

Current and future ground and space-based observatories have begun to constrain the galaxy luminosity function brightward of the knee of the galaxy luminosity function Bowler *et al.*(2015).

In Fig. 4 we show the current constraints on the galaxy luminosity function at $z \sim 7$, highlighting the relative dearth of observational constraints at the very bright end, but see Ono *et al.*(2017). This is due to the need for large areas of sky in order to find these rare objects, but also imaging data that is still deep enough to reliably detect these galaxies, without being highly contaminated by galaxies at lower redshifts. This will improve over the next few years as the combination of deep optical and near-infrared covering approximately an order of magnitude larger area than the combined UltraVISTA and UKIDSS-UDS surveys becomes available from a combination of the dark energy survey, HyperSuprimeCam and VIDEO data. Coupled with the unique power of HyperSuprimeCam for surveying large-swathes of sky with near-infrared narrow-band filters to detect line-emitters in the EoR (Ouchi *et al.* 2017), we will thus be edging closer to enabling cross-correlations between 21-cm experiments and optical/near-infrared galaxy surveys.

However, even with these, the total area surveyed will still be relatively limited compared to the current and future 21-cm experiments. Therefore, the large area of sky that

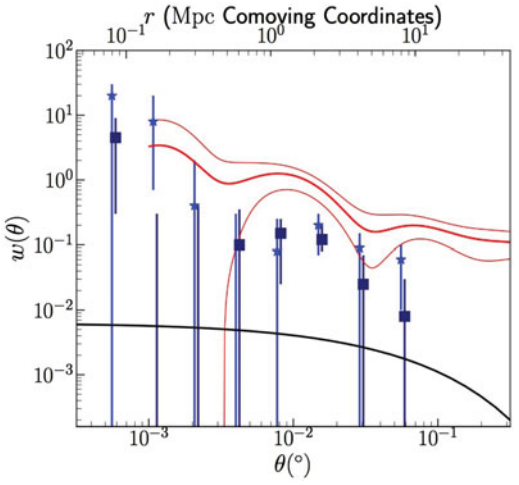

Figure 5. Measurements of the $z \sim 6$ angular correlation function. The red curve shows the results of Hatfield *et al.*(2017) for the brightest Lyman-break galaxies from ground based surveys (1σ uncertainties represented by the thinner red curves), and the lower luminosity galaxies from the study of Harikane *et al.*(2016). The dark matter angular correlation function is shown as the black curve. (Figure adapted from Hatfield *et al.* 2017).

will be surveyed with the *Euclid* satellite will be an extremely valuable, and a powerful complement to the new generation of 21-cm EoR experiments. In Fig. 4 we highlight the region of the $z \sim 7$ galaxy luminosity function that *Euclid* will probe with a wide-area (15,000 deg^2) and deep (40 deg^2) survey of Lyman-break galaxies, both potentially detecting several thousand galaxies at $z \sim 7$. Such numbers would allow cross-correlations between 21-cm power spectra and the galaxies at this epoch, not only helping mitigate the issues of only using the auto-correlation, but also providing a direct link between the bright-end of the source population responsible for reionisation, and the neutral and ionised hydrogen regions detected with low-frequency radio facilties. Additional information could then be gleaned from considering the line emitting galaxies observed with a combination of the *Euclid* slitless spectroscopy survey and ground-based narrow-band imaging. Indeed, the first measurements of the clustering of galaxies at $z > 6$ have recently been made (see Fig. 5; Harikane *et al.* 2016, Hatfield *et al.* 2017), showing that the brightest Lyman-break galaxies trace the most massive dark matter haloes at this epoch, and potentially where the signal from the EoR will be the strongest at 21 cm.

4. Summary

We have discussed how current and future ground- and space-based optical and near-infrared surveys can play an important role in the detection and understanding of the epoch of reionisation, when combined with the new generation of 21-cm experiments. Given the technical challenges of detecting the EoR at 21 cm it appears prudent to consider the vast swathe of information on the extragalactic foregrounds that is available from non-radio wavelengths. In particular, understanding the clustering of the star-forming population which make up the faint end of the radio source counts could be important. They comprise the faint foreground population that may provide additional prior information, which helps assure one that any features in the 21-cm power spectrum are indeed from the EoR, and not due to incomplete subtraction of a clustered population at lower redshift.

We then go on to highlight how current and future optical and near-infrared surveys can play a key role in confirming the 21-cm signal via cross-correlations with the galaxies within the EoR. The depth and breadth of survey required for this is ambitious, but the envisaged surveys with *Euclid* and *WFIRST*, in addition to ground-based narrow-band surveys for Ly-α emitters could and should play a role in allowing us to obtain a complete understanding of the reionisation history of the Universe.

References

Aird, J., Coil, A. L., Georgakakis, A., *et al.* 2015, *MNRAS*, 451, 1892
Bowler, R. A. A., Dunlop, J. S., McLure, R. J., *et al.* 2015, *MNRAS*, 452, 1817
Bouwens, R. J., Illingworth, G. D., Oesch, P. A., *et al.* 2015, *ApJ*, 803, 34
Condon, J. J., Anderson, M. L., & Helou, G. 1991, *ApJ*, 376, 95
Condon, J. J. 1992, ARA&A, 30, 575
Cooray, A. & Sheth, R. 2002, *Physics Reports*, 372, 1
Davies, L. J. M., Huynh, M. T., Hopkins, A. M., *et al.* 2017, *MNRAS*, 466, 2312
Fernandes, C. A. C., Jarvis, M. J., Rawlings, S., *et al.* 2011, *MNRAS*, 411, 1909
Finkelstein, S. L., Ryan, R. E., Jr., Papovich, C., *et al.* 2015, *ApJ*, 810, 71
Guhathakurta, P., Tyson, J. A., & Majewski, S. R. 1990, ApJL, 357, L9
Hale, C. L., Jarvis, M. J., Delvecchio, I., *et al.* 2018, *MNRAS*, 474, 4133
Hardcastle, M. J., Evans, D. A., & Croston, J. H. 2006, *MNRAS*, 370, 1893
Harikane, Y., Ouchi, M., Ono, Y., *et al.* 2016, *ApJ*, 821, 123
Hatfield, P. W., Lindsay, S. N., Jarvis, M. J., *et al.* 2016, *MNRAS*, 459, 2618
Hatfield, P. W., Bowler, R. A. A., Jarvis, M. J., & Hale, C. L. 2017, arXiv:1702.03309
Hatfield, P. W. & Jarvis, M. J. 2017, *MNRAS*, 472, 3570
Heckman, T. M. & Best, P. N. 2014, ARA&A, 52, 589
Jarvis, M. J., Smith, D. J. B., Bonfield, D. G., *et al.* 2010, *MNRAS*, 409, 92
Jarvis, M. J., Bonfield, D. G., Bruce, V. A., *et al.* 2013, *MNRAS*, 428, 1281
Johnston, R., Vaccari, M., Jarvis, M., *et al.* 2015, *MNRAS*, 453, 2540
Madau, P. & Dickinson, M. 2014, ARA&A, 52, 415
McLure, R. J., Dunlop, J. S., Bowler, R. A. A., *et al.* 2013, *MNRAS*, 432, 2696
Noeske, K. G., Weiner, B. J., Faber, S. M., *et al.* 2007, ApJL, 660, L43
Ono, Y., Ouchi, M., Harikane, Y., *et al.* 2017, *PASJ*, (arXiv:1704.06004)
Ouchi, M., Mobasher, B., Shimasaku, K., *et al.* 2009, *ApJ*, 706, 1136
Ouchi, M., Shimasaku, K., Furusawa, H., *et al.* 2010, *ApJ*, 723, 869
Ouchi, M., Harikane, Y., Shibuya, T., *et al.* 2017, arXiv:1704.07455
Rawlings, S. & Saunders, R. 1991, *Nature*, 349, 138
Smith, D. J. B., Dunne, L., da Cunha, E., *et al.* 2012, *MNRAS*, 427, 703
van der Kruit, P. C. 1971, *A&A*, 15, 110
Willott, C. J., Rawlings, S., Blundell, K. M., & Lacy, M. 1999, *MNRAS*, 309, 1017
Wilman, R. J., Miller, L., Jarvis, M. J., *et al.* 2008, *MNRAS*, 388, 1335
Whitaker, K. E., van Dokkum, P. G., Brammer, G., & Franx, M. 2012, ApJL, 754, L29
Yun, M. S., Reddy, N. A., & Condon, J. J. 2001, *ApJ*, 554, 803
Zheng, Z., Berlind, A. A., Weinberg, D. H., *et al.* 2005, *ApJ*, 633, 791