EMU: Evolutionary Map of the Universe

Ray P. Norris[A,AM], *A. M. Hopkins*[B,AJ], *J. Afonso*[C], *S. Brown*[A], *J. J. Condon*[D], *L. Dunne*[E], *I. Feain*[A], *R. Hollow*[A], *M. Jarvis*[F,AL], *M. Johnston-Hollitt*[G], *E. Lenc*[A], *E. Middelberg*[H], *P. Padovani*[I], *I. Prandoni*[J], *L. Rudnick*[K], *N. Seymour*[L], *G. Umana*[M], *H. Andernach*[N], *D. M. Alexander*[U], *P. N. Appleton*[O], *D. Bacon*[P], *J. Banfield*[A], *W. Becker*[Q], *M. J. I. Brown*[R], *P. Ciliegi*[S], *C. Jackson*[A], *S. Eales*[T], *A. C. Edge*[U], *B. M. Gaensler*[V,AJ], *G. Giovannini*[J], *C. A. Hales*[A,V], *P. Hancock*[V,AJ], *M. T. Huynh*[W], *E. Ibar*[X], *R. J. Ivison*[X,Y], *R. Kennicutt*[Z], *Amy E. Kimball*[D], *A. M. Koekemoer*[AA], *B. S. Koribalski*[A], *Á. R. López-Sánchez*[B,AK], *M. Y. Mao*[A,B,AB], *T. Murphy*[V,AJ], *H. Messias*[C], *K. A. Pimbblet*[R], *A. Raccanelli*[P], *K. E. Randall*[A,V], *T. H. Reiprich*[AD], *I. G. Roseboom*[AE], *H. Röttgering*[AF], *D. J. Saikia*[AG], *R. G. Sharp*[AH], *O. B. Slee*[A], *Ian Smail*[U], *M. A. Thompson*[F], *J. S. Urquhart*[A], *J. V. Wall*[AI], and *G.-B. Zhao*[P]

[A] CSIRO Astronomy & Space Science, Epping, NSW 1710, Australia
[B] Australian Astronomical Observatory, Epping, NSW 1710, Australia
[C] Centro de Astronomia e Astrofísica da Universidade de Lisboa, Observatório Astronómico de Lisboa, Tapada da Ajuda, 1349-018 Lisboa, Portugal
[D] National Radio Astronomy Observatory, Charlottesville, VA 22903, USA
[E] School of Physics & Astronomy, University of Nottingham, Nottingham, NG7 2RD, UK
[F] Centre for Astrophysics Research, University of Hertfordshire, Hatfield, Hertfordshire, AL10 9AB, UK
[G] School of Chemical & Physical Sciences, Victoria University of Wellington, Wellington 6140, New Zealand
[H] Astronomisches Institut, Ruhr-Universität Bochum, 44801 Bochum, Germany
[I] European Southern Observatory, D-85748 Garching bei München, Germany
[J] INAF-IRA, 40129 Bologna, Italy
[K] Department of Astronomy, University of Minnesota, Minneapolis, MN 55455, USA
[L] University College London, Mullard Space Science Laboratory, Dorking, Surrey, RH5 6NT, UK
[M] INAF-Catania Astrophysical Observatory, 95123 Catania, Italy
[N] Depto. de Astronomía, Universidad de Guanajuato, Guanajuato, C.P. 36000, GTO, Mexico
[O] NASA Herschel Science Center, Caltech, Pasadena, CA 91125, USA
[P] Institute of Cosmology and Gravitation, University of Portsmouth, Portsmouth, PO1 3FX, UK
[Q] Max-Planck Institut für extraterr. Physik, 85741 Garching, Germany
[R] School of Physics, Monash University, Clayton, VIC 3800, Australia
[S] INAF – OABO, 40127 Bologna, Italy
[T] School of Physics and Astronomy, Cardiff University, Cardiff, CF24 3AA, UK
[U] Department of Physics, Durham University, Durham, DH1 3LE, UK
[V] Sydney Institute for Astronomy, School of Physics, The University of Sydney, NSW 2006, Australia
[W] International Centre for Radio Astronomy Research, University of WA, Crawley, WA 6009, Australia
[X] UK Astronomy Technology Centre, Royal Observatory, Edinburgh EH9 3HJ, UK
[Y] Institute for Astronomy, University of Edinburgh, Edinburgh, EH9 3HJ, UK
[Z] Institute of Astronomy, University of Cambridge, Cambridge, CB3 0HA, UK
[AA] Space Telescope Science Institute, Baltimore MD 21218, USA
[AB] School of Mathematics & Physics, University of Tasmania, Hobart, TAS 7001, Australia
[AD] Argelander Institute for Astronomy, Bonn University, 53121 Bonn, Germany
[AE] Dept of Physics and Astronomy, University of Sussex, Falmer, East Sussex, BN1 9RH, UK
[AF] Leiden Observatory, Leiden University, 2300 RA Leiden, The Netherlands
[AG] National Centre for Radio Astrophysics, TIFR, Pune 411 007, India
[AH] Mount Stromlo Observatory, Canberra, ACT 2611, Australia
[AI] Dept. of Physics and Astronomy, University of British Columbia, Vancouver, B.C. V6T 1Z1, Canada
[AJ] ARC Centre of Excellence for All-sky Astrophysics, University of Sydney, Sydney, NSW 2006, Australia
[AK] Department of Physics and Astronomy, Macquarie University, North Ryde, NSW 2109, Australia
[AL] Physics Department, University of the Western Cape, Cape Town 7535, South Africa
[AM] Corresponding author. Email: ray.norris@csiro.au

Abstract: EMU is a wide-field radio continuum survey planned for the new Australian Square Kilometre Array Pathfinder (ASKAP) telescope. The primary goal of EMU is to make a deep (rms $\sim 10\,\mu$Jy/beam) radio continuum survey of the entire Southern sky at 1.3 GHz, extending as far North as $+30°$ declination, with a resolution of 10 arcsec. EMU is expected to detect and catalogue about 70 million galaxies, including typical star-forming galaxies up to $z \sim 1$, powerful starbursts to even greater redshifts, and active galactic nuclei to the edge of the visible Universe. It will undoubtedly discover new classes of object. This paper defines the science goals and parameters of the survey, and describes the development of techniques necessary to maximise the science return from EMU.

1 Introduction

1.1 Background

Deep continuum surveys of the radio sky have a distinguished history both for discovering new classes of object and for providing radio counterparts to astronomical objects studied at other wavelengths. The earliest large surveys, such as the 3C catalogue (Edge et al. 1959) and the Molonglo Reference Catalogue (Large et al. 1981), gave us the first insight into the physics of radio galaxies and radio-loud quasars, but were insufficiently sensitive to detect any but the nearest radio-quiet or star-forming galaxies. Later radio surveys reached flux densities where normal star-forming galaxies were detected, but were still largely dominated by radio-loud active galactic nuclei (AGN). Only very long integrations in narrow deep fields made it possible to start probing star-forming galaxies beyond the local Universe. This paper describes a planned survey, EMU (Evolutionary Map of the Universe), which will reach a similar sensitivity ($\sim 10\,\mu$Jy/beam) as those deep surveys, but over the entire visible sky. At that sensitivity, EMU will be able to trace the evolution of galaxies over most of the lifetime of the Universe.

Fig. 1 shows the major 20-cm continuum radio surveys. The largest existing radio survey, shown in the top right, is the wide but shallow NRAO VLA Sky Survey (NVSS), whose release paper (Condon et al. 1998) is one of the most cited papers in astronomy. The most sensitive existing radio survey is the deep but narrow Lockman Hole observation (Owen & Morison 2008) in the lower left. All current surveys are bounded by a diagonal line that roughly marks the limit of available telescope time of current-generation radio telescopes. The region to the left of this line is currently unexplored, and this area of observational phase space presumably contains as many potential new discoveries as the region to the right.

The Square Kilometre Array (SKA) is a proposed major internationally funded radio telescope (Dewdney et al. 2009) whose construction is expected to be completed in 2022. It will be many times more sensitive than any existing radio telescope, and will answer fundamental

questions about the Universe (Carilli & Rawlings 2004). It is likely to consist of between 1000 and 1500 15-meter dishes in a central area of diameter 5 km, surrounded by an equal number of dishes in a region stretching up to thousands of kilometres.

The Australian SKA Pathfinder (ASKAP) is a new radio telescope being built both to test and develop aspects of potential SKA technology, and to develop
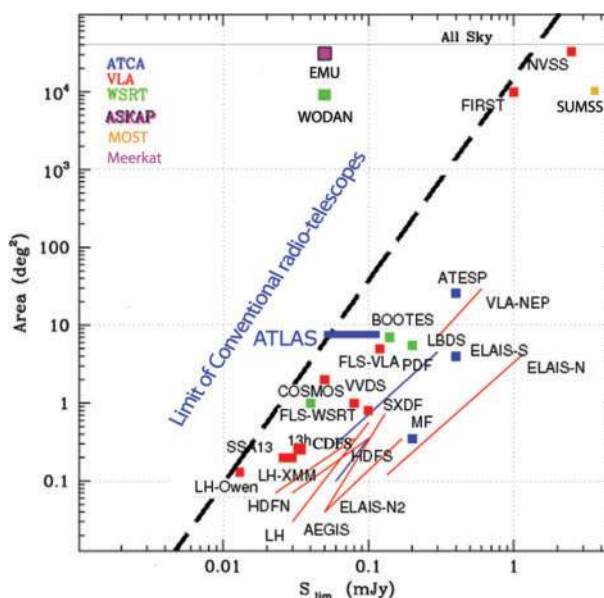

Figure 1 Comparison of EMU with existing deep 20 cm radio surveys. Horizontal axis is 5-σ sensitivity, and vertical axis shows the sky coverage. The diagonal dashed line shows the approximate envelope of existing surveys, which is largely determined by the availability of telescope time. The squares in the top left represent the EMU survey, discussed in this paper, and the complementary WODAN (Röttgering et al. 2010b) survey which has been proposed for the upgraded Westerbork telescope to cover the sky north of $+30°$. Surveys represented by diagonal lines are those which range from a wide shallow area to a smaller deep area. The horizontal line for ATLAS extends in sensitivity from the intermediate published data releases (Norris et al. 2006; Middelberg et al. 2008a; Hales et al. 2011) to the final data release (Banfield et al. 2011).

SKA science. ASKAP is being built on the Australian candidate SKA site in Western Australia, at the Murchison Radioastronomy Observatory, with a planned completion date of late 2012. In addition to developing SKA science and technology, ASKAP is a major telescope in its own right, likely to generate significant new astronomical discoveries.

1.2 ASKAP

ASKAP (Johnston et al. 2007, 2008; Deboer et al. 2009) will consist of 36 12-metre antennas spread over a region 6 km in diameter. Although the array of antennas is no larger than many existing radio telescopes, the feed array at the focus of each antenna is revolutionary, with a phased-array feed (PAF: Bunton & Hay 2010) of 96 dual-polarisation pixels, designed to work in a frequency band of 700–1800 MHz, with an instantaneous bandwidth of 300 MHz. This will replace the single-pixel feeds that are almost universal in current-generation synthesis radio telescopes. As a result, ASKAP will have a field of view up to 30 deg^2 enabling it to survey the sky up to thirty times faster than existing synthesis arrays, and allowing surveys of a scope that cannot be contemplated with current-generation telescopes. To ensure good calibration, the antennas are a novel 3-axis design, with the feed and reflector rotating to mimic the effect of an equatorial mount, ensuring a constant position angle of the PAF and sidelobes on the sky. The pointing accuracy of each antenna is significantly better than 30 arcsec.

The ASKAP array configuration (Gupta et al. 2008) balances the need for high sensitivity to extended structures (particularly for neutral hydrogen surveys) with the need for high resolution for continuum projects such as EMU. To achieve this, 30 antennas follow a roughly Gaussian distribution with a scale of ∼700 m, corresponding to a point spread function of ∼30 arcsec using natural weighting, with a further six antennas extending to a maximum baseline of 6 km, corresponding to a point spread function of ∼10 arcsec using uniform weighting. The positions of the antennas are optimised for *uv* coverage (i.e. coverage in the Fourier plane) between declination −50° and +10°, but give excellent *uv* coverage between declination −90° and +30°.

The PAF is still under development, but the performance of prototypes gives us confidence that the EMU survey is feasible as planned. The PAF will consist of 96 dual-polarisation receivers, each with a system temperature ∼50 K, which are combined in a beam-former to form up to 36 beams. Each of these beams has the same primary beam response as a single-pixel feed (∼1.2° full-width half-maximum at 1.4 GHz), distributed in a uniform grid across an envelope of 30 deg^2. The optimum weighting and number of beams is still being studied, but the current expectation is that 36 beams will be used for EMU, with the sensitivity over the 30 deg^2 field of view (FOV) expected to be uniform to ∼20%. This will be improved to <10% uniformity by dithering, with no significant loss of sensitivity, so that the images from the 36 beams can be jointly imaged and deconvolved as a single image covering the FOV.

Consequently, it is expected that the telescope will dwell on one position in the sky for 12 hours, reaching an rms sensitivity of ∼10 μJy/beam over a ∼30 deg^2 FOV. The strategy for achieving this is still under development, and is discussed in §3.6.

Although high spatial resolution is essential for EMU, the short spacings of ASKAP also deliver excellent sensitivity to low surface brightness emission, which is essential for a number of science drivers such as studies of radio emission from nearby clusters (§2.7). The ∼10 μJy/beam rms continuum sensitivity in 12 hours is approximately constant for beam sizes from 10 to 30 arcsec, then increases to ∼20 μJy/beam for a 90 arcsec beam and ∼40 μJy/beam for a 3 arcmin beam.

Science data processing (Cornwell et al. 2011) will take place in an automated pipeline processor in real time. To keep up with the large data rate (∼2.5 GB/s, or 100 PB/year), all science data processing steps, from the output of the correlator to science-qualified images, spectra, and catalogues, are performed in automated pipelines running on a highly distributed parallel processing computer. These steps include flagging bad data, calibration, imaging, source-finding, and archiving.

A typical ASKAP field will contain about 50 Jy of flux in compact or slightly resolved sources. ASKAP can observe the entire visible 20 cm continuum sky to an rms sensitivity of ∼1 mJy/beam in one day, so that initial observations will produce a global sky model (an accurate description of all sources stronger than ∼1 mJy) which significantly simplifies subsequent processing, as strong sources will be subtracted from the visibility data before processing. This sky model also means that antenna complex gains can be self-calibrated in one minute without any need to switch to calibrator sources. It is expected that individual receiver gains will be sufficiently stable that the dominant causes of antenna complex gain variation (i.e. ionosphere and troposphere) will be common to all pixels, so that a gain solution in one beam of an antenna can be transferred to other beams of that antenna.

In continuum mode, ASKAP will observe a 300 MHz band, split into 1 MHz channels, with full Stokes parameters measured in each channel. The data will be processed in a multi-frequency synthesis mode, in which data from each channel are correctly gridded in the *uv* plane. As well as producing images and source catalogues, the processing pipeline will also measure spectral index, spectral curvature, and all polarisation products across the band.

Completion of the Boolardy Engineering Test Array (BETA), which is a 6-antenna subset of ASKAP, is expected in late 2011. BETA will be equipped with 6 PAFs, and all the necessary beamformers, correlators, and processing hardware to produce images over the full 30 deg^2 field. The primary goal of BETA is to enable engineering tests and commissioning activities while the

remaining ASKAP hardware is being constructed. If engineering commissioning proceeds as expected, science observations on BETA will commence in 2012, on a small number of test fields on which good radio-astronomical and ancillary data already exist. These test BETA observations will be used to debug and fine-tune not only ASKAP, but also the processes for handling the data.

The full ASKAP array is expected to be commissioned in early 2013, and the science surveys are expected to start in late 2013. There is no proprietary period on ASKAP data, with all data being placed in the public domain after quality control, so data are expected to start flowing to the astronomical community by the end of 2013.

Expressions of interest for ASKAP survey projects were sought in November 2008, and full proposals were solicited in mid-2009 (Ball et al. 2009). Of the 38 initial expressions of interest, ten proposals were eventually selected, with two, EMU (Evolutionary Map of the Universe) and WALLABY (Wide-field ASKAP L-band Legacy All-sky Blink SurveY: Koribalski et al. 2011), being selected as highest priority. EMU is an all-sky continuum survey, while WALLABY is an all-sky survey for neutral hydrogen. ASKAP design is now being driven by the requirement to maximise the science return from these ten projects, with a particular focus on maximising the science from EMU and WALLABY.

It is planned that EMU, WALLABY, and some other projects will observe commensally, i.e., they will agree on an observing schedule, and will observe the sky in both continuum and HI modes at the same time, splitting the two data streams into two separate processing pipelines. More information on all the ASKAP projects, including links to their individual websites, can be found at `http://askap.org`.

1.3 EMU

The primary goal of EMU is to make a deep (10 μJy/beam rms) radio continuum survey of the entire Southern sky, extending as far North as +30°. EMU will cover roughly the same fraction (75%) of the sky as the benchmark NVSS survey (Condon et al. 1998), but will be 45 times more sensitive, and will have an angular resolution (10 arcsec) 4.5 times better. Because of the excellent short-spacing *uv* coverage of ASKAP, EMU will also have higher sensitivity to extended structures. The sky coverage of EMU is shown in Fig.2, and the EMU specifications are summarised in Table 1. Like most radio surveys, EMU will adopt a 5-σ cutoff, leading to a source detection threshold of 50 μJy/beam. EMU is expected to generate a catalogue of about 70 million galaxies, and all radio data from the EMU survey will be placed in the public domain as soon as the data quality has been assured.

Currently, only a total of about 5 deg^2 of the sky has been surveyed at 20 cm to the planned 10 μJy/beam rms of EMU, in fields such as the *Hubble*, *Chandra*, COSMOS and Phoenix deep fields (Huynh et al. 2005; Miller et al.
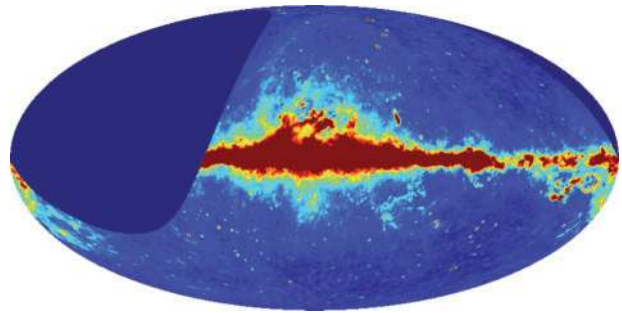

Figure 2 A representation of the EMU sky coverage in Galactic coordinates overlaid on 23 GHz WMAP data (Gold et al. 2011). The dark area in the top left is the part of the sky **not** covered by EMU.

Table 1. EMU Specifications

Instantaneous FOV	30 deg^2
Area of survey	Entire sky south of +30° dec.
Synthesised beamwidth	10 arcsec FWHM
Frequency range	1130–1430 MHz
RMS sensitivity	10 μJy/beam
Total integration time	~1.5 years[1]
Number of sources	~70 million

[1]The primary specification is the sensitivity, rather than the integration time. If for any reason ASKAP is less sensitive than expected, EMU will increase the integration time rather than lose sensitivity. Conversely, an increase in sensitivity of ASKAP may reduce the total integration time.

2008; Schinnerer et al. 2007; Hopkins et al. 2003; Biggs & Ivison 2006; Morrison et al. 2010), with a further 7 deg^2 expected in the immediate future as part of the ATLAS survey (Norris et al. 2006; Middelberg et al. 2008a; Hales et al. 2011; Banfield et al. 2011).

Surveys at this depth extend beyond the traditional domains of radio astronomy, where sources are predominantly radio-loud galaxies and quasars, into the regime of star-forming galaxies. At this depth, even the most common active galactic nuclei (AGN) are radio-quiet AGNs, which make up most of the X-ray extragalactic sources. As a result, the role of radio astronomy is changing. Whereas most traditional radio-astronomical surveys had most impact on the niche area of radio-loud AGNs, current radio-astronomical surveys are dominated by the same galaxies as are studied by optical and IR surveys, making radio-astronomical surveys such as EMU an increasingly important component of multi-wavelength studies of galactic evolution.

Because only a small area of sky has been surveyed to the depth of EMU, it is difficult to estimate precisely how many galaxies it will detect. Most surveys to this sensitivity cover only a small area of sky, so that source counts at this level are significantly affected by sample variance, completeness, and bias issues. Our estimate for the number per deg^2 above a flux density of 50 μJy/beam is based on an extrapolation from source counts at higher flux densities (2263 sources/deg^2; Jackson 2005), the compilation shown in Fig. 3 (2278 sources/deg^2), and
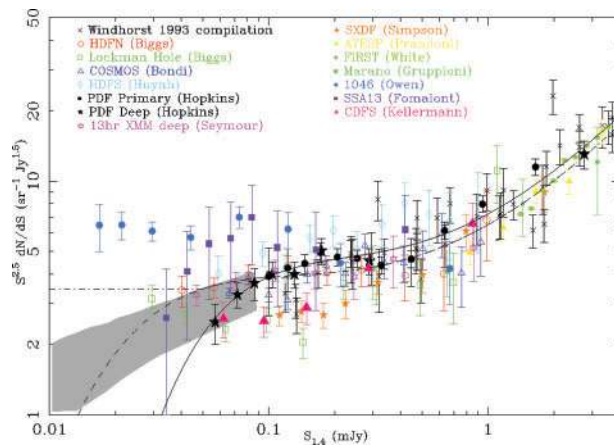
Figure 3 Distribution of differential radio source counts at 1.4 GHz, based on and updated from the distribution shown in Hopkins et al. (2003). The solid curve is the polynomial fit from Hopkins et al. (2003), the dashed curve is an updated polynomial fit and is the one used to estimate the EMU source numbers. The horizontal dot-dashed line represents a non-evolving population in a Euclidean universe. The shaded region shows the prediction based on fluctuations due to weak confusing sources (a 'P(D) analysis') from Condon (1974); Mitchell & Condon (1985).

Figure 4 Differential fraction of star-forming galaxies as a function of 1.4 GHz flux density, from a selection of recent deep surveys. Shaded boxes, and the two lines for Padovani et al., show the range of uncertainty in the survey results. Arrows indicate constraints from other surveys. These results show that the fraction of star-forming galaxies increases rapidly below 1 mJy and, at the 50 μJy survey limit of EMU, about 75% of sources will be star-forming galaxies.

the COSMOS survey (2261 sources/deg^2; Scoville et al. 2007; Schinnerer et al. 2007). These three figures are in good agreement and predict a total of $\sim$70 million sources in EMU, which is therefore the number adopted throughout this paper.

Estimating the fraction of these radio sources which are AGN is difficult. Below 1 mJy, star-forming galaxies start to become a major component of the 1.4 GHz source counts, dominating below $\sim$0.15 mJy (Seymour et al. 2008; Ibar et al. 2009), but, even at these levels, there is still a significant proportion of low-luminosity AGNs (Jarvis et al. 2004; Afonso et al. 2005, 2006; Norris et al. 2006; Simpson et al. 2006; Smolcic et al. 2008; Seymour et al. 2008; Mignano et al. 2008; Padovani et al. 2009).

Seymour et al. (2008) have presented the most comprehensive attempt so far to divide radio sources into AGN and SF galaxies, and their result, together with other recent estimates, is shown in Fig. 4. From these we estimate that about 75% of EMU sources will be star-forming galaxies.

To estimate the redshift distribution of AGN and SF galaxies, we use the SKADS simulation (Wilman et al. 2008, 2010), shown in Fig. 5. About 50 million of the EMU sources are expected to be star-forming galaxies (see §2.1) at redshifts up to $z \sim 3$, with a mean redshift of $z \sim 1.08$. The remainder are AGNs with a mean $z \sim 1.88$, and extend up to $z \sim 6$. However, if any FRII (Fanaroff & Riley 1974) galaxies exist beyond that redshift (e.g. $L \sim 3.3 \times 10^{25}$ W Hz^{-1} at $z = 10$), EMU will detect them.

Confusion of radio sources, discussed more thoroughly in §3.6.1, is well-understood at this level, since previous surveys have already imaged small areas of sky to this depth and beyond.
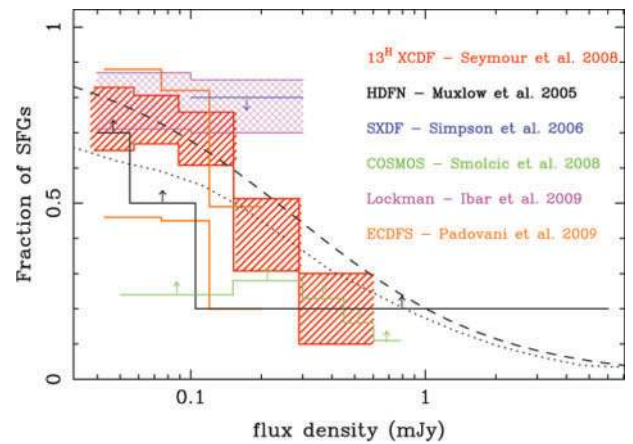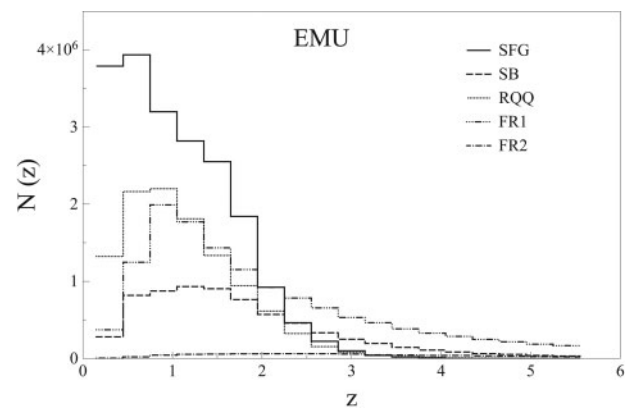

Figure 5 Expected redshift distribution of EMU sources, based on the SKADS simulations (Wilman et al. 2008, 2010). The five lines show the distributions for star-forming galaxies (SFG), starburst galaxies(SB), radio-quiet quasars (RQQ), and radio-loud galaxies of Fanaroff–Riley types I and II (FRI & FR2; Fanaroff & Riley 1974). Vertical scale shows the total number of sources expected to be detected by EMU.

EMU differs from many previous surveys in that a goal of the project is to cross-identify the detected radio sources with major surveys at other wavelengths, and produce public-domain VO-accessible catalogues as 'value-added' data products. This is facilitated by the growth in the number of large southern hemisphere telescopes and associated planned major surveys spanning all wavelengths, discussed below in §3.9.

1.4 Science

Broadly, the key science goals for EMU are:

- To trace the evolution of star-forming galaxies from $z = 2$ to the present day, using a wavelength unbiased by dust or molecular emission,

- To trace the evolution of massive black holes throughout the history of the Universe, and understand their relationship to star formation,
- To use the distribution of radio sources to explore the large-scale structure and cosmological parameters of the Universe, and to test fundamental physics,
- To determine how radio sources populate dark matter haloes, as a step towards understanding the underlying astrophysics of clusters and haloes,
- To create the most sensitive wide-field atlas of Galactic continuum emission yet made in the Southern Hemisphere, addressing areas such as star formation, supernovae, and Galactic structure,
- To explore an uncharted region of observational parameter space, with a high likelihood of finding new classes of object.

Table 1 gives an overview of EMU specifications. In addition to the well-defined scientific goals outlined above, and the obvious legacy value, the large EMU dataset will include extremely rare objects, which is only made possible by covering large areas.

A challenge for EMU will be the lack of spectroscopic redshifts, since no existing or planned redshift survey can cover more than a tiny fraction of EMU's 70 million sources. As discussed in §3.11, ~30% of EMU sources will have multi-wavelength optical/IR photometric data at the time of data release, increasing to ~70% in 2020. We expect these to provide accurate photometric redshifts for the majority of star-forming galaxies in EMU, and a minority of AGN (for which photometric redshifts tend to be unreliable). In addition, many of the EMU sources will have 'statistical redshifts', which are valuable for some statistical tests. For example, most polarised sources are AGNs (mean $z \sim 1.88$), while most unpolarised sources are star-forming galaxies (mean $z \sim 1.08$). More precise statistical redshifts can be derived where optical/IR photometry is available, as discussed in §3.11.

A further goal of EMU is to test and develop strategies for the SKA. Many aspects of ASKAP, such as the automated observing, calibration, and data reduction processes, and the phased-array feeds, are potential technologies for the SKA, and it will be important to test whether these approaches deliver the planned results.

1.5 Relationship to Other Surveys

The following radio surveys are particularly complementary to the scientific goals of EMU.

- The WODAN survey (Röttgering et al. 2010b) has been proposed for the Westerbork telescope which is currently being upgraded with a phased array feed (Oosterloo et al. 2009). WODAN will cover the northern 25% of the sky (i.e. North of declination +30°) that is inaccessible to ASKAP, to an rms sensitivity of 10 μJy/beam and a spatial resolution of 15 arcsec. Together, EMU and WODAN will provide full-sky 1.3 GHz imaging at ~10–15 arcsec resolution to an rms noise level of 10 μJy/beam, providing an unprecedented sensitive all-sky radio survey as a legacy for astronomers at all wavelengths. The WODAN survey will overlap with EMU by a few degrees of declination to provide a comparison and cross-validation, to ensure consistent calibration, and to check on completeness and potential sources of bias between the surveys.
- The LOFAR continuum survey (Röttgering et al. 2010a) will cover the northern half of the sky (i.e. North of declination 0°) with the new LOFAR telescope operating at low frequencies (15–200 MHz). LOFAR will be especially complementary to WODAN and EMU in surveying the sky at high sensitivity and resolution but at a much lower frequency.
- The MIGHTEE survey (van der Heyden & Jarvis 2010) on the Meerkat telescope (Jonas 2009) will probe to much fainter flux densities (0.1–1 μJy rms) over smaller areas ($\sim$35 deg^2) at higher angular resolution, providing the completeness as a function of flux density for the EMU and WODAN Surveys. The higher sensitivity and resolution will enable exploration of the AGN and star-forming galaxy populations to higher redshifts and lower luminosities.
- POSSUM (Gaensler et al. 2010) is an all-sky ASKAP survey of linear polarisation. It is expected that POSSUM will be commensal with EMU, and that the two surveys will overlap considerably in their analysis pipelines and source catalogues. POSSUM will provide a catalogue of polarised fluxes and Faraday rotation measures for approximately 3 million compact extragalactic sources. These data will be used to determine the large-scale magnetic field geometry of the Milky Way, to study the turbulent properties of the interstellar medium, and to constrain the evolution of intergalactic magnetic fields as a function of cosmic time. POSSUM will also be a valuable counterpart to EMU, in that it will provide polarisation properties or upper limits to polarisation for all sources detected by EMU.
- FLASH (Ball et al. 2009) is an ASKAP survey whose goal is to detect extragalactic neutral hydrogen absorption. To do so it will observe at frequencies outside the 1130–1430 MHz band of EMU, thus yielding valuable spectral index information for those sources common to both surveys.
- DINGO (Ball et al. 2009)is an ASKAP survey whose goal is to detect faint extragalactic neutral hydrogen emission, and to do so it will spend many days on one ASKAP pointing. As a byproduct, it will thus provide sensitive continuum images over smaller areas (several tens of deg^2), allowing EMU to explore fainter flux densities in an optimal tiered survey structure, and also to quantify the effects of confusion at this level. However, the continuum images from DINGO will be severely confusion-limited at flux densities below a few μJy/beam. It may be possible to transcend this limit by subtracting known sources from the image, such as those star-forming galaxies which are seen in infrared images and whose radio flux can be predicted

using the IR–radio correlation. However, this challenge is currently external to the core EMU project.

- VAST (Chatterjee et al. 2010) is an ASKAP survey that will observe partly commensally with EMU, with the goal of detecting transients and variable sources. EMU has no planned transient capability, since all information on variability of EMU sources will be available from VAST. This separation enables each of EMU and VAST to focus on its specific science goals, although significant coordination between the projects will clearly be essential.

- WALLABY (Koribalski et al. 2011) is an HI survey which will deliver high-sensitivity spectral line data over the same area of sky as EMU, and will observe commensally with EMU. Observations will give a velocity coverage of $-2{,}000$ to $+77{,}000\,\mathrm{km\,s}^{-1}$ ($z = 0 - 0.26$) and velocity resolution of $4\,\mathrm{km\,s}^{-1}$. The angular resolution for WALLABY will be 30 arcsec, a factor of three lower than EMU, as computing resources to make the large spectral line data cubes are restricted to baselines shorter than ~ 2 km. Nearly all the $\sim 5 \times 10^5$ sources detected by WALLABY will also be detected by EMU, and WALLABY will provide an HI redshift for each of these, adding significantly to the redshift information for low-redshift EMU sources.

This paper defines the EMU survey, setting out its science goals in §2, and identifying the challenges to achieve these goals. §3 describes how these challenges are being addressed in the EMU Design Study, and §4 describes the survey operational plan, primary data products, and the data release plans and policy.

2 EMU Science Goals

2.1 Star-Forming Galaxies and AGNs

The fraction of star-forming galaxies as a function of flux density is shown in Figure 4. Of the ~ 70 million sources detected by EMU to a 5σ limit of $50\,\mu$Jy, about 20 million galaxies are expected to be dominated by AGNs, and 50 million to be dominated by star formation (SF). However, there is considerable overlap between the two classes, with composite AGN/SF galaxies becoming more common at low flux densities (e.g. Chapman et al. 2003; Norris et al. 2006; Seymour et al. 2008).

It is unclear what fraction of putative SF galaxies have a significant AGN component. However, the excellent agreement in star-formation rate between radio and other star-forming indicators (e.g. Cram et al. 1998; Bell 2003) suggests that an AGN is not a major contributor to the radio emission in such galaxies.

Detected AGNs and star-forming galaxies span a significant fraction of the age of the Universe, almost reaching the era of re-ionisation for radio AGNs and the most extreme starbursts. Particularly at high redshift, both AGN and star formation processes are likely to be important in a large fraction of galaxies, but neither the fraction of the luminosity (bolometric and radio)

generated by each process, nor how they are influenced by feedback, is currently known. EMU will explore the evolution of these populations and its dependence on galaxy mass, environment, SF history, and interaction/merger history. It will quantify these effects in detail, by providing a deep homogeneously selected sample of both AGN and SF galaxies over the majority of cosmic history, unbiased by dust obscuration, and so provide a comprehensive overview of galaxy evolution.

The EMU analysis pipeline, which will encompass automated multi-wavelength cross-identification of sources between the EMU catalogue and other complementary surveys, will also include a variety of measures appropriate for distinguishing between, and quantifying the proportions of, AGN and star-forming activity. These will include:

- Radio morphology (e.g. Biggs & Ivison 2008; Biggs, Younger & Ivison 2010),
- Radio spectral index (e.g. Ibar et al. 2009, 2010),
- Radio–far-infrared ratio,
- Radio–near-infrared ratio,
- Radio polarisation, from the POSSUM project (Gaensler et al. 2010),
- Radio variability, from the VAST project (Chatterjee et al. 2010),
- optical and IR colours, to be used in SED analysis and comparison with templates.

These diagnostics and the algorithm which encodes them will be trained prior to the EMU data using the ATLAS data (Mao et al. 2011b).

2.2 Evolution of Star-Forming Galaxies

Tracing when and where stars formed across cosmic time is one of the key questions in galaxy evolution today. There is now evidence that star formation depends both on galaxy stellar mass (Feulner et al. 2005; Juneau et al. 2005; Papovich et al. 2006; Mobasher et al. 2009, see Figure. 7) and environment (Lewis et al. 2002; Gomez et al. 2003), and these dependencies also evolve with time (Elbaz et al. 2007). Massive galaxies appear to form their stars early and quickly, progressively becoming less active after $z \sim 2$, while lower mass galaxies become dominant at lower redshifts. As a result, the cosmic SFR density is dominated by progressively less massive galaxies in less dense environments at lower redshifts. This process is known as cosmic 'downsizing' (Cowie et al. 1996), although it has been challenged by results from Zheng et al. (2007) and Dunne et al. (2009).

EMU, with 50 million SF galaxies and a large collection of complementary data, will accurately determine the true dependence of star formation on mass, environment, and redshift, in all types of galaxy from normal star-forming galaxies up to the most UltraLuminous InfraRed Galaxies (ULIRGs). Radio observations have a considerable advantage over most other diagnostics of star
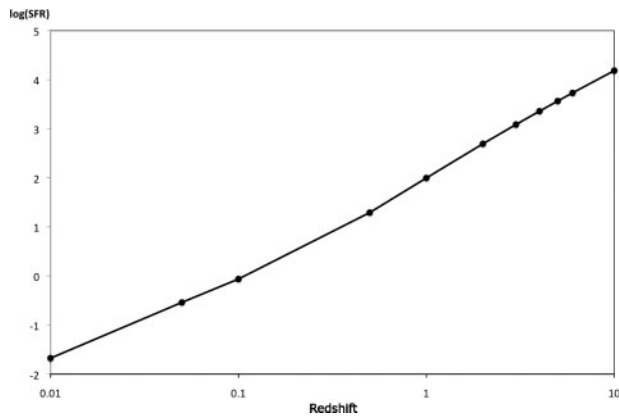
Figure 6 Star formation rate (in $M_\odot$/yr) of individual star-forming galaxies detectable by EMU at $5\,\sigma$ as a function of redshift. The calculation of SFR follows Bell (2003) modified following Seymour et al. (2008), and the k-correction assumes a spectral index of -0.7.

formation as radio luminosities are directly related to star formation rates (Condon 1992; Haarsma et al. 2000; Bell 2003) through a calibration reliable at least up to $z \sim 2$ (Garn et al. 2009), and no corrections for absorption by gas or dust are required. However, to measure star formation rates from radio luminosities requires that any AGNs are removed from radio samples, and this can largely be achieved using a variety of indicators, as discussed in §2.1. For tracing evolution, EMU sources will necessarily be limited to those for which redshifts can be measured or estimated, as discussed in §3.11. Fortunately, photometric redshifts prove to be very accurate for SF galaxies (see e.g. Rowan-Robinson et al. 2008) and a large fraction of star-forming galaxies detected by EMU will eventually have reliable photometric redshifts.

EMU's surface brightness sensitivity of 0.06 K is well below the median face-on surface brightness of a typical spiral galaxy of ~ 1 K at 1.4 GHz. So, unlike previous large radio surveys, EMU will have the brightness sensitivity needed to detect nearly all resolved normal spiral galaxies. Luminous starburst galaxies, with star formation rates (SFRs) around $100\,M_\odot\,yr^{-1}$, will be detectable to $z \sim 2$ (see Fig. 6), while ordinary disk galaxies like the Milky Way, with SFRs of only a few $M_\odot\,yr^{-1}$, will be visible to $z \sim 0.3$, and many of these will have spectroscopic redshifts measured by WALLABY. Star-formation rates of classes of galaxy at higher redshifts can be estimated using stacking (see §3.10), provided that AGNs have been eliminated from the stacked sample.

EMU will therefore trace the co-moving SFR density (e.g., Lilly et al. 1996; Madau et al. 1996; Hopkins 2004; Hopkins & Beacom 2006) up to high redshift with greater accuracy than previously possible, and thereby determine when most of the stars in the Universe formed. The wide area of this survey will overcome the problems of low number statistics and sample variance sometimes associated with deep but small-area radio surveys (Hopkins et al. 2003).
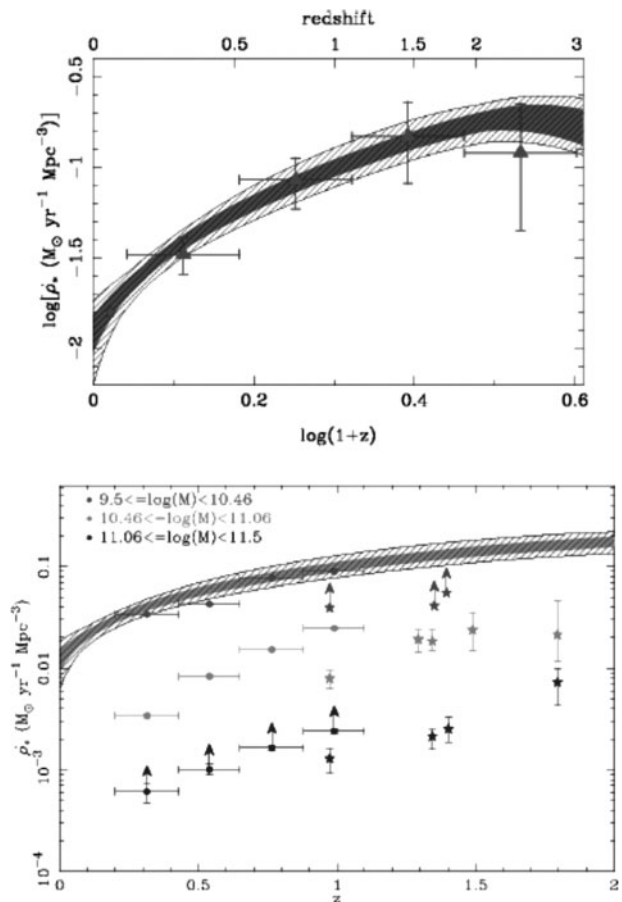

Figure 7 Evolution with redshift of the star formation rate density (SFRD) in galaxies, $\dot{\rho}_*$. The grey and shaded regions in both panels are the 1-σ and 3-σ confidence region fits to the compilation of SFRD data from Hopkins & Beacom (2006). Top: Measured SFRD taken from Mauch & Sadler (2007) and Seymour et al. (2008). Bottom: The SFRD shown as a function of galaxy mass, adapted from Fig. 7 of Mobasher et al. (2009).

2.3 Evolution of AGNs

AGNs play a major role in the framework of galaxy formation. During their short active lifetime, they release an enormous amount of energy in the form of ionising radiation or relativistic jets, which can have a significant effect on the host galaxy and its surroundings.

The peak of QSO activity took place at $z \sim 2$ (e.g. Schmidt 1968; Shaver et al. 1996; Croom et al. 2004, Hasinger et al. 2005), at epochs when SF activity was also extreme. Intriguingly, embedded AGNs have been found in 20–30% of $z \sim 2$ massive star-forming galaxies (Daddi et al. 2007). It has even been suggested that the AGN jets could be responsible for the formation of galaxies at high redshift (Elbaz et al. 2009; Klamer et al. 2004).

At lower redshifts, AGNs appear to downsize, in a similar way to that of SF galaxies, so that the peak of the AGN activity appears to shift significantly to lower redshifts for lower power AGNs. This suggests (Cowie et al. 1996) that a feedback mechanism couples AGNs to galaxy evolution (Hasinger et al. 2005; Bongiorno et al. 2007; Padovani et al. 2009), via supernovae, starbursts or

AGNs (Croton et al. 2006; Bower et al. 2006). AGN jets either push back and heat the infalling gas, reducing the cooling flows building up the galaxy, or shock-heat and collapse the gas clouds, inducing star formation. The AGN energetic feedback appears to be an important ingredient for reproducing the galaxy stellar mass function (Croton et al. 2006; Best et al. 2006), and the remarkable black hole versus bulge mass (or velocity dispersion) correlation (Gebhardt et al. 2000; Ferrarese & Merritt 2000; Springel et al. 2005).

Recent work work on low-z 3CR galaxies by Ogle et al. (2010) suggest that radio jets can lead to significant shock-heating of the host molecular disk, leading to a significant enhancement of warm molecular hydrogen emission. Despite their large molecular masses, these systems seem to be very inefficient at making stars, suggesting that shock-heating by radio jets can play a negative-feedback role in star formation in these systems.

Low-power radio-loud AGNs ($P < 10^{25}$ WHz^{-1}) appear to be quite different from the high-power radio-loud AGNs. Many do not show the luminous narrow lines expected in the framework of AGN unification (see e.g. Hine & Longair 1979; Laing et al. 1994; Jackson & Rawlings 1997). They also lack the expected infrared emission from a dusty torus (see e.g. Whysong & Antonucci 2004; Ogle et al. 2006) and do not show accretion-related X-ray emission (Hardcastle et al. 2006; Evans et al. 2006). The sum of this observational evidence suggests that low-power radio sources correspond to a distinct phase of AGN, accreting through a radiatively inefficient 'radio mode', rather than the radiatively efficient accretion 'quasar mode' typical of optical or X-ray selected AGNs (Croton et al. 2006; Bower et al. 2006). The physical reasons behind these two different accretion modes are still unclear (e.g. Hardcastle et al. 2007; Herbert et al. 2010).

The key questions related to AGN evolution that EMU will address include: (1) The relationship between AGN and SF activity; (2) The evolution of low and intermediate power AGNs, exploring the so-called 'AGN cosmic downsizing' scenario, (3) The relative contribution of different accretion regimes (radio versus quasar modes) in low and intermediate luminosity radio AGNs, their evolution with redshift, and the role played by the environment; (4) The relative contribution of radiative versus jet-driven (kinetic) feedback to the global AGN feedback in models of galaxy formation; and (5) the evolution of radio galaxies as a function of environment, spectral index, redshift, radio luminosity, and source size.

2.4 High-Redshift AGNs and the Epoch of Re-ionisation

There have been many attempts to find very high redshift radio galaxies (De Breuck et al. 2001; Jarvis et al. 2001a,b; Best et al. 2003; Cruz et al. 2006, 2007) but all suffer from the difficulty of finding these extremely rare galaxies. EMU will identify numerous radio galaxies at $z < 4$ and isolate the most distant objects by cross-matching with large near- and mid-infrared surveys, as described in §3.9. EMU will also provide spectral indices and (with POSSUM) polarisation information which can also help identify high-redshift AGNs.

High-redshift AGNs from EMU will constrain the existence and demographics of the most massive galaxies over the history of the Universe (e.g. Jarvis et al. 2001b,c; Wall et al. 2005). These sources provide constraints on the co-evolution of galaxy bulges and central supermassive black holes (Magorrian et al. 1998; McLure et al. 2006), trace (proto-)clusters at early times (e.g. Miley et al. 2006), and may be crucial in determining the impact that powerful radio activity has on the host galaxy (e.g. Croton et al. 2006; Bower et al. 2006) and its larger-scale environment (Rawlings & Jarvis 2004; Gopal-Krishna & Wiita 2001; Elbaz et al. 2009).

Infrared-Faint Radio Sources (IFRS) are probably a particular class of radio source that are characterised by a very high radio–infrared ratio ($S_{20\,cm}/S_{3.6\,\mu m} > 500$) and a low infrared flux density. First identified by Norris et al. (2006), there is strong circumstantial evidence that they are high-redshift radio galaxies, based on their SED (Garn & Alexander 2008; Huynh et al. 2010), their steep spectral index (Middelberg et al. 2011), their VLBI cores (Norris et al. 2007; Middelberg et al. 2008b), and their extreme faintness in the infrared ($S_{3.6\,\mu m} < 1.2\,\mu Jy$; Norris et al. 2011a). Observations suggest ∼5 IFRSs occur per deg^2 at a flux limit of ∼100 μJy (Norris et al. 2006; Middelberg et al. 2008a), implying that EMU will detect at least 1.5×10^5, but the lack of corresponding deep infrared data will prevent their identification from other unremarkable non-detections. However, many will be located in smaller deep infrared fields such as those observed with HERMES (Oliver et al. 2010), which should yield several thousand IFRSs, enabling us to compile solid statistical data on their distribution, spectral index, and polarisation. Thousands more will be selected as candidate high-redshift AGNs through their steep spectrum and polarisation (see §3.11), and it expected that they will turn out to be an important class of high-redshift AGN.

A major objective of current extragalactic astronomy is to understand the Epoch of Reionisation (EoR), when ultraviolet photons from the first stars and quasars ionised the primordial neutral hydrogen. This process can in principle be studied by measuring the neutral hydrogen fraction in the early Universe using the 21 cm forest in front of a bright distant source (Carilli et al. 2002), analogous to the Ly-α forest seen against bright quasars at lower redshifts (e.g. Peroux et al. 2005. Such observations require a population of $z < 6$ radio-loud background sources, but the highest redshift radio galaxy currently known lies at only $z = 5.2$ (van Breugel et al. 1999). EMU provides an excellent opportunity to identify such sources, producing a large sample of distant radio sources for investigating the formation and evolution of the most massive galaxies at the highest redshifts.

2.5 CSS and GPS sources

Compact Steep Spectrum (CSS) sources are typically defined to be radio sources with a spectral index < -0.5, and with a size less than about 20 kpc, and hence of sub-galactic dimensions. Gigahertz Peaked Spectrum (GPS) sources are radio sources whose spectrum reaches a maximum in the frequency range 1–10 GHz. Typically they are significantly smaller (hundreds of pc) than CSS sources, and confined to the circumnuclear regions of the host galaxy.

Approximately 10% and 30% of bright centimetre wavelength sources are GPS and CSS sources, respectively. If similar statistics hold for AGNs at low flux densities, EMU will discover $\sim$2 million GPS sources and $\sim$5 million CSS sources. Most studies of CSS and GPS objects have been confined to sources selected from strong source surveys, and current samples only probe down to $\sim$10 mJy at 1.4 GHz (Tschager et al. 2003; Snellen et al. 1998, 1999; Randall et al. 2011). EMU will offer a complete, faint sample of these intriguing objects, providing a probe into the evolution of young radio AGNs, and showing how these objects fit into models of galaxy evolution.

There is a consensus (O'Dea 1998; Fanti 2009a,b; Morganti et al. 2009; Snellen et al. 2009) that GPS and CSS sources represent the start of the evolutionary path for large-scale radio sources. It is generally accepted that most GPS sources evolve into CSS sources, which gradually transform into the largest radio sources known, Fanaroff–Riley Type I and II galaxies (Fanaroff & Riley 1974), depending on their initial luminosity. These sources offer an ideal resource to investigate early galaxy evolution and formation, as well as AGN feedback, as they are young AGNs but also have star formation occurring due to interactions and mergers (Labiano et al. 2008; Morganti 2008). Measurements of component advance speeds for a few compact sources yield ages of about 10^3 yr, while spectral studies indicate ages less than about 10^5 yr (O'Dea 1998; Owsianik et al. 1999; Murgia et al. 1999).

These sources are thought to be fuelled by the infall of gas to the supermassive black holes, triggered by interactions with companions and mergers. CSS and GPS sources show evidence, from both their structural and polarisation asymmetries, that the jets are interacting with an asymmetric external environment (Saikia et al. 2003; Cotton et al. 2003). CSS and GPS sources show evidence of H_I absorption more often than other radio-loud AGNs, with the GPS objects showing the highest incidence of absorption (Gupta et al. 2006, and references therein). The EMU sample of CSS and GPS sources over a large redshift range will enable the variation of the H_I properties with cosmic epoch to be determined.

Several CSS and GPS objects show evidence of diffuse extended emission which may be a remnant of an earlier cycle of activity. An important goal of AGN physics is to understand the episodic nature of nuclear or jet activity
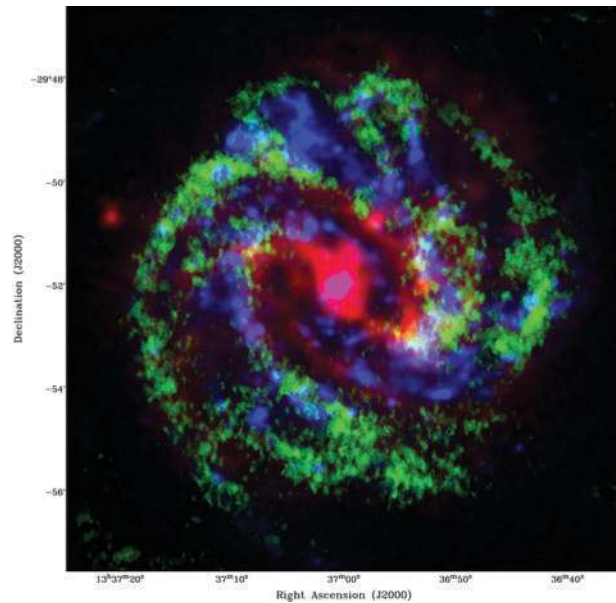

Figure 8 This multi-wavelength composite image of the inner part of the spiral galaxy M 83 highlights the synergy between the two key ASKAP large survey science projects: WALLABY and EMU. The ATCA 20-cm radio continuum emission is shown in red, the ATCA HI distribution in green, and the GALEX FUV emission in blue (Koribalski et al. 2008; Gil de Paz et al. 2007).

(Saikia et al. 2009). EMU, with its high surface brightness sensitivity, will be an ideal survey to probe the existence of such diffuse emission for a large number of sources, and constrain the time scales of episodic activity for these compact objects.

2.6 Low Redshift Galaxies, and Synergies with WALLABY

WALLABY is expected to detect HI emission from $\sim 5 \times 10^5$ galaxies to a depth of $z = 0.05$, with massive galaxies detected out to $z = 0.25$. The majority of these galaxies will be spiral, with typical HI masses of a few times $10^9 M_\odot$. Their large gas reservoir fuels star formation, implying that all spiral galaxies detected by WALLABY will have 20-cm radio continuum emission detectable by EMU. As a result, WALLABY is expected to contribute $\sim 5 \times 10^5$ redshifts to EMU.

In return, EMU will be able to measure star-formation rates of the galaxies studied by WALLABY, enabling detailed studies of the factors that influence star-formation rates in the local Universe. A goal of WALLABY is to measure local and global star-formation rates for all gas-rich spirals and compare their SF and HI distributions (see Fig. 8).

Although most WALLABY data will have a 30-arcsec resolution, high-resolution (10 arcsec) HI 'postage stamps' will be obtained of particularly interesting nearby galaxies, allowing a more detailed analysis and comparison with data at other wavelengths.

2.7 Galaxy Clusters

Several different types of diffuse radio emission are associated with clusters of galaxies (Kempner et al. 2004),
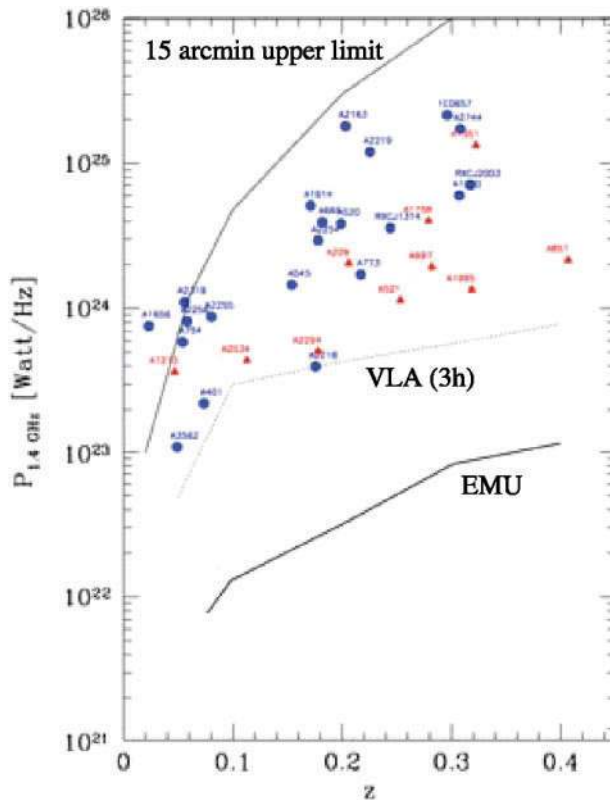
Figure 9 1.4 GHz radio power of detected cluster haloes as a function of redshift showing the detection limits of previous cluster observations, adapted from Giovannini et al.(2009), and the 5σ calculated detection limit of EMU, assuming a halo with a diameter of 1 Mpc, using the calculated ASKAP sensitivity on different scale sizes discussed in §1.2. The upper line shows the limit corresponding to a scale size of 15 arcmin, which is approximately the largest object that can be imaged with the VLA unless single-dish data are added to the interferometry data. At high redshifts, sensitivity may be limited by confusion, although we expect to overcome this by subtracting off compact sources.

including haloes around the centres of clusters, relics (representing shocks from cluster–cluster collisions) at the periphery, and tailed radio galaxies which are an important barometer of the intra-cluster medium. These three classes of radio source are important as tracers of clusters, and are also diagnostics of the physics of clusters, particularly when combined with X-ray data. However, the number of detected cluster radio sources is limited by current telescope sensitivities (see Fig. 9). EMU will not only give us large samples of such sources, but will push beyond the present limits to detect diffuse sources with a range of powers over a larger redshift range, greatly improving our understanding of these sources.

Most clusters have been found either through X-ray surveys (Rosati et al. 1998; Romer et al. 2001; Pierre et al. 2003) or by searching for over-densities in optical colour–position space (Gladders & Yee 2005; Wilson et al. 2008; Kodama et al. 2007). As a result, tens of thousands of clusters are currently known, but only a few at $z > 1$

(Wilson et al. 2008; Kodama et al. 2007), with the highest redshift at $z = 2.07$ (Gobat et al. 2011).

2.7.1 Haloes

Radio haloes are found in clusters and groups of galaxies, indicating synchrotron emission powered by diffuse and faint (0.1-1 µG) magnetic fields and relativistic particles. So far about 35 radio haloes are known with $z < 0.4$ and only 2 at $z \sim 0.5$, generally discovered by making deep radio surveys of hot and bright X-ray clusters (Venturi et al. 2008; Giovannini et al. 2009). A strong correlation is present between the total halo radio power and the cluster X-ray luminosity (Giovannini et al. 2009). Since the cluster mass and X-ray luminosity are correlated it follows that halo radio power correlates with the cluster mass (Feretti 2000; Govoni et al. 2001). Brunetti et al. (2009), Cassano et al. (2010), Schuecker et al. (2001) have suggested that radio haloes in the centres of clusters are distributed bimodally, with haloes generally found only in those clusters which have recently undergone a merger, resulting in a disturbed appearance at X-ray wavelengths.

The ATLBS survey (Subrahmanyan et al. 2010), which has surveyed 8.4 deg^2 to an rms sensitivity of 80 µJy/beam on a scale size of 50 arcsec at 1.4 GHz, has detected tens of diffuse sources, of which about 20 have been tentatively identified as cluster and group haloes (Saripalli et al. 2011). EMU will have even better sensitivity to low-surface brightness structures than ATLBS, so if the ATLBS numbers are confirmed, then EMU will discover $\sim 6 \times 10^4$ cluster and group haloes, which significantly increases the number of known clusters. An important science goal will then be to compare the X-ray properties (luminosity, temperature and, for the brighter clusters, morphologies) of these radio-selected clusters to those of the X-ray selected population from the eROSITA all-sky X-ray survey (Predehl et al. 2010).

2.7.2 Relics

On the periphery of clusters, elongated radio 'relics' are found, which probably represent the signatures of shock structures generated in cluster mergers (Röttgering et al. 1997; Brown et al. 2011a; van Weeren et al. 2010). They provide important diagnostics for the dynamics of accretion and mergers by which clusters form (Barrena et al. 2009). Large populations of these structures will appear in EMU, and are likely to lead to new cluster identifications, especially beyond $z \sim 0.5$. For example, only 44 radio relics are currently known (Hoeft et al. 2011), and few have been discovered in current radio surveys because of the relatively poor sensitivity of most surveys to low–surface brightness structures. One probable relic has been discovered in the seven square degrees of ATLAS (Middelberg et al. 2008a; Mao et al. 2010). Since EMU will have greater sensitivity to such low–surface brightness structures than ATLAS, this suggests that EMU should detect >4000 relics, although this

number is clearly very uncertain. As a means of finding clusters, it is less effective than other radio and X-ray techniques, but will be invaluable for studying shock structures accompanying cluster mergers. Furthermore, relics show evidence of ordered large-scale magnetic fields in the periphery of galaxy clusters, in regions with a very low density of galaxies and thermal gas.

2.7.3 Tailed Radio Galaxies

Head tail, wide-angle tail (WAT), and narrow-angle tail galaxies (collectively named 'tailed radio galaxies') are believed to represent radio-loud AGNs in which the jets are distorted by the intra-cluster medium (Mao et al. 2009, 2010). They can also contribute to diffuse emission (Rudnick & Lemmerman 2009), especially after the jets from the nucleus of the host galaxy have turned off; EMU's high resolution and sensitivity, especially when combined with polarisation information from the commensal POSSUM survey (Gaensler et al. 2010), will allow us to distinguish these from the (largely unpolarised) cluster-wide halo emission. Even more importantly, such tailed galaxies can be detected out to high redshifts (Wing & Blanton 2011; Mao et al. 2010), providing a powerful diagnostic for finding clusters. From the WATs discovered in the ATLAS fields, Mao et al. (2011c) and Dehghan et al. (2011) have estimated that EMU will detect at least 26000 WATs, and possibly as many as 2×10^5, depending on their luminosity function and density evolution.

2.7.4 AGN Feedback in Galaxy Clusters and Mini-Halo Sources

The relatively cool and dense gas at the centres of many galaxy clusters and groups emits copious X-ray radiation by thermal bremsstrahlung and line emission. In the absence of external sources of heating, this high emission should lead to very rapid cooling ($t_{cool} < 1\,\mathrm{Gyr}$) and very high rates of mass deposition onto the central cluster galaxy (up to $\sim 1000\,\mathrm{M_\odot/yr}$), in turn causing very high star-formation rates and strong X-ray line emission (e.g. Fabian & Nulsen 1977; Cowie & Binney 1977; Peterson & Fabian 2006; McNamara & Nulsen 2007). The lack of such obvious observational signatures (e.g. Peterson et al. 2001, 2003) implies that some central source of heating must be present. The most plausible source of heating is feedback from the central AGN. EMU will enable a statistically significant study of the correlation between the radio emission from AGNs at the cluster centre with the thermal and non-thermal cluster properties, and explore how it evolves with redshift.

Moreover a few cooling-core clusters exhibit signs of diffuse synchrotron emission that extends far from the dominant radio galaxy at the cluster center, forming what is referred to as a mini-halo. These diffuse radio sources are extended on a moderate scale (typically $\sim 500\,\mathrm{kpc}$) and, like large-scale haloes, have a steep spectrum and a very low surface brightness. Because of a combination of small angular size and the strong radio emission of the

central radio galaxy, the detection of a mini-halo requires data of a much higher dynamic range and resolution than those in available surveys. The study of radio emission from the center of cooling-core clusters is of great importance, not only in understanding the feedback mechanism involved in the energy transfer between the AGN and the ambient medium (e.g. McNamara & Nulson 2007), but also in the formation process of the non-thermal mini-haloes. The energy released by the central AGN may also play a role in the formation of these extended structures (e.g. Fujita et al. 2007).

2.7.5 The Impact of EMU on Cluster Research

The key impacts of EMU are likely to be:

- An increase in the number of known clusters beyond the few tens of thousands currently known. EMU will detect at least 3×10^4 new clusters (Mao et al. 2011c; Norris et al. 2011c; Dehghan et al. 2011), which will roughly double the number of known clusters. Depending on the redshift distribution and luminosity function of cluster radio sources, EMU may detect as many as $\sim 2 \times 10^5$ cluster sources (Mao et al. 2011c; Norris et al. 2011c). eROSITA is also expected to detect $\sim 10^5$ clusters at X-ray wavelengths (Predehl et al. 2010; Pillepich et al. 2011), and comparison of these two complementary surveys will be transformational.

- The detection of clusters at high redshifts. In principle, radio sources can be used to detect clusters even beyond $z = 1$, where current constraints on large-scale structure are weaker, and traditional detection techniques like X-ray surveys and use of the Red Cluster Sequence (Gladders & Yee 2005) become less effective. A few clusters have already been detected up to high redshift using radio detections (Blanton et al. 2003; Wing et al. 2011), but extrapolation beyond $z = 1$ is uncertain because the luminosity function and evolution of cluster radio sources are unknown. Furthermore, inverse Compton cooling of electrons by the cosmic microwave background is expected to quench their synchrotron radio emission at $z \gg 1$, although confidence in this expectation is reduced by the failure to detect a similar effect in the far IR–radio correlation at high redshift (Mao et al. 2011a).

- The exploration of radio properties of clusters detected at other wavelengths. It will be important to compare the properties of radio-selected clusters from the unbiased EMU survey to those of the X-ray selected population from surveys such as the eROSITA all-sky X-ray survey (Predehl et al. 2010), and the Sunyaev–Zeldovich surveys made with the South Pole Telescope (Williamson et al. 2011), Atacama Cosmology Telescope (Marriage et al. 2010) and Planck (Planck Collaboration 2011) surveys. EMU will enable not only statistical approaches (e.g. halo occupation distribution modelling: Peacock & Smith 2000) but also directly observed overdensities using tracers such as WATs (Mao et al. 2010).

- The detection of new forms of cluster radio emission. For example, high-resolution numerical simulations (Battaglia et al. 2009) predict an additional network of weak shocks throughout the cluster volume which can only be seen with sufficient resolution and sensitivity, but which will be detectable with the short spacings of ASKAP.

Additional cluster-related questions to be addressed by EMU include:

- Do the radio properties of AGNs depend more on the properties of their host galaxy or those of their local environment, such as gas temperature and cooling time (e.g. Mittal et al. 2009)?
- Are galaxies of a given mass more likely to host a radio source if they are a central or satellite galaxy within a halo?
- Is star formation truncated in haloes above a given mass by AGN feedback or virial shock heating?
- How does the observed correlation and indicated feedback cycle between cluster-central radio AGNs and the cooling intracluster medium evolve with redshift (e.g. Santos et al. 2010)?

2.8 Cosmic Filaments, and the Warm–Hot Intergalactic Medium (WHIM)

Approximately half of the Universe's baryons are currently 'missing' (Cen & Ostriker 1999) and are likely to be contained in the elusive Warm–Hot Intergalactic Medium (WHIM), where temperatures of 10^5–10^7 K make them extremely difficult to detect (Fraser-McKelvie et al. 2011). By subtracting the contribution of compact radio sources, studies of low–surface brightness emission with EMU can illuminate the otherwise invisible baryons associated with large-scale structure. Perhaps the most important diffuse radio sources are those illuminating shock structures in the WHIM, which occur during infall onto and along large-scale structure filaments (Miniati et al. 2001; Dolag et al. 2008).

Because of the sensitivity and short spacings of ASKAP, EMU will be able to detect faint radio emission from cosmological filaments, increasing our knowledge of the physical properties of large-scale structures. The shortest spacing of ASKAP, 20 m (resulting in sampling of the uv plane down to 8 m in a joint deconvolution), ensures that EMU will be sensitive to structures as large as 1°, while even larger scales can be imaged by adding in single-dish data.

In a few cases, the synchrotron emission from filaments may be detected directly. One putative example is the bridge of radio emission extending over 1 Mpc between the central halo of the Coma cluster and the peripheral relic source $1253 + 275$ (Kim et al. 1989, Kronberg et al. 2007). Another is the radio emission from ZwCl $2341.1 + 0000$ which is a linear structure some 6 Mpc in length (Bagchi et al. 2002; Giovannini et al. 2010).

Even though EMU will be able to detect only the brightest shock structures in the WHIM directly, statistical characterisation of an ensemble of fainter shock emission is possible (Keshet et al. 2004; Brown et al. 2010). Faint synchrotron emission due to WHIM shocks is highly correlated with large-scale structure (LSS) on ~1–4 Mpc scales (Pfrommer et al. 2007; Ryu et al. 2008; Skillman et al. 2008, 2010), and can thus be detected below the EMU noise limits through LSS cross-correlation (Brown et al. 2010) using the statistical redshifts described in §3.11, or in some cases high-quality photometric or spectroscopic redshifts. The sky coverage of EMU coupled with its high sensitivity to low–surface brightness emission makes it ideal for detecting the synchrotron cosmic web at cosmologically important redshifts $0.1 < z < 1.0$ (Brown et al. 2011b), corresponding to ~0.1–0.5 Mpc on arcmin scales.

2.9 Cosmology and Fundamental Physics

2.9.1 Background

The existence of dark energy is one of the most profound problems in cosmology. The evidence for its presence is indirect: it is implied by the relation between Type Ia supernovae and the Hubble constant (Riess et al. 1998; Perlmutter et al. 1999), by the combined inference of flat geometry ($\Omega_{tot} = 1$) from cosmic microwave background (CMB) measurements and low mass density $\Omega_{mass} \sim 0.3$) from large-scale structure measurements (Cole et al. 2005; Eisenstein et al. 2005), and from galaxy cluster measurements (Vikhlinin et al. 2009; Mantz et al. 2010; Allen et al. 2008).

A further puzzle is the nature of gravity. While general relativity (GR) is consistent with all current observational measurements (e.g. Uzan 2003), it fails to connect gravity with the other fundamental forces, models for which lead to hypotheses such as 'brane-worlds' (Brax & van de Bruck 2003) in which gravity is modified by leakage into dimensions other than the brane dimensions. Such models predict deviations from GR which have not yet been fully tested.

EMU will be able to perform independent tests of models of dark energy and deviations from GR using a combination of (i) the auto-correlation of radio source positions, (ii) the cross-correlation of radio sources with the CMB (the late-time integrated Sachs–Wolfe (ISW) effect), and (iii) cross-correlation of radio sources with foreground objects (cosmic magnification).

We assume throughout most of this section that redshifts of EMU sources are **not** available, and at the end consider the impact if statistical redshifts were to become available. Throughout this section, we assume a 5σ detection limit, which is expected to be reliable for the EMU survey, while Raccanelli et al. (2011) use a conservative 10σ limit, which gives rise to slightly different constraints.

2.9.2 Autocorrelations

The autocorrelation function (ACF) of source counts, also known as the two-point correlation of radio sources,

measures the degree of clustering of the radio sources (see e.g. Blake & Wall 2002; Raccanelli et al. 2011). Here we consider the ACF for EMU sources that we will measure in two dimensions on the sky, ignoring any potential availability of redshifts, and making no distinction between source types, which have different biases and number densities as a function of redshift.

Despite these limitations, the ACF is of value for two purposes. First, cosmological constraints from this probe can be combined with those from other probes to substantially improve the net cosmological constraint, as shown by the combined constraints in Figures 11 and 12. Second, the behaviour of the ACF on large scales constrains non-Gaussianity (i.e. the extent to which density fluctuations in the early universe were distributed with non-Gaussian noise) hence providing a window on early-Universe physics. Non-Gaussianity indicates a positive skewness of the matter density probability distribution, which would lead to an increased bias for large-scale haloes (see e.g. Dalal et al. 2008). EMU will be able to detect non-Gaussianity parameterized by the f_{NL} parameter at the level $f_{NL} \gtrsim 8$, (Raccanelli et al. 2011). Such a detection would be difficult to reconcile with a simple, single scalar field inflation model for the early Universe.

2.9.3 The Integrated Sachs–Wolfe Effect

The ISW effect can be understood as follows. Travelling from the last scattering surface to us, CMB photons pass through gravitational potential wells of large structures such as superclusters that lie along our line of sight. In an Einstein–de Sitter universe, the blueshift of a photon falling into a well is cancelled by the redshift as it climbs out. However, in a universe with dark energy or modified GR, the local gravitational potential varies with time, stretching the potential well while photons are traversing it, so that the blueshift on entry is not fully cancelled by the redshift on exit. Thus the net photon energy is increased, producing CMB temperature anisotropies which make the CMB appear slightly hotter in the direction of superclusters.

The effect is very weak, but can be detected by cross-correlating CMB maps with tracers of large-scale structure (Crittenden & Turok 1996) such as radio sources. The detection of the effect has been repeatedly confirmed (e.g. Giannantonio et al. 2008a, and references therein); for instance, galaxies in the Sloan Digital Sky Survey (SDSS York et al. 2000) have been cross-correlated with the CMB (Granett et al. 2008) to achieve a 4σ result. In the radio, cross-correlation by Raccanelli et al. (2008) of the NVSS radio galaxies with the CMB anisotropies also gave a tentative detection of the ISW effect. The NVSS detection is partly limited by shot noise, and so the far greater number of galaxies in EMU should achieve a robust measurement of the effect.

EMU will also be able to constrain the redshift evolution of the equation of state of dark energy, which would allow us to distinguish between different models of dark energy, such as a cosmological constant (Tegmark 2004),
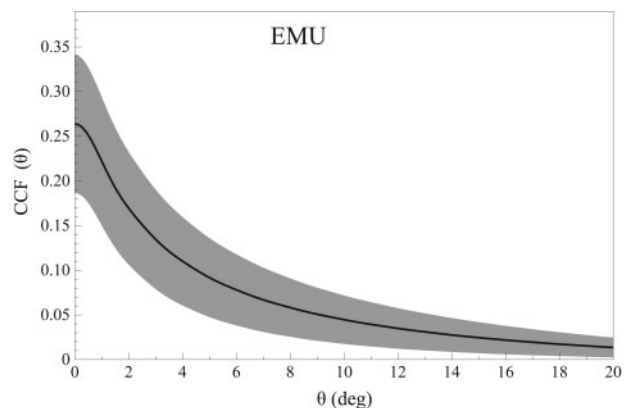

Figure 10 Predicted cross-correlation of the CMB (measured by WMAP etc.) with EMU sources, assuming a standard ΛCDM cosmology. The solid lines show the theoretical prediction, and the shaded region shows cosmic variance errors.

quintessence (Ziatev et al. 1999), early dark energy (Xia et al. 2009), or unified dark matter models (Bertacca et al. 2011).

The ISW effect is also sensitive to any modifications of gravity, allowing sensitive constraints on gravity, and can also be used to test models for the evolution of the clustering and bias of radio sources (Raccanelli et al. 2008), and to test models for the cosmological evolution of radio sources (Massardi et al. 2010). The predicted cross-correlation function for EMU sources is shown in Fig. 10. EMU will provide sensitive constraints on any modifications of gravity.

2.9.4 Cosmic Magnification

Large-scale structures along the line of sight to a distant source introduce distortions in the images of high-redshift sources as a result of gravitational lensing. The distortions by foreground (low-redshift) galaxies increase the apparent area occupied by background (high-redshift) galaxies, thus reducing the observed number density at a given flux density. On the other hand, the magnification increases the total flux density of unresolved high-redshift sources, thus increasing the observed number density at a given flux density. The size of the combined change in apparent number density due to these two opposing effects is sensitive to the assumed cosmological geometry and parameters of GR.

The combined effect, known as cosmic magnification, can be tested by cross-correlating number densities of low-redshift sources (e.g. selected from the optical surveys or from EMU star-forming galaxies) with number densities of high-redshift sources (selected from EMU). The effect was first unambiguously detected by Scranton et al. (2005) using SDSS foreground galaxies and quasars. The large database of EMU sources will develop this into a powerful technique for testing cosmological models.

2.9.5 Observational Tests with EMU

Deviations from GR or dark-energy physics will influence the auto-correlation, the ISW and cosmic

magnification (Zhao et al. 2010). By modelling the EMU source distribution and bias, Raccanelli et al. (2011) have shown that significant limits can be placed on cosmological parameters (such as w and w') that describe dark energy, and on parameters η and μ that describe modifications to GR (Pogosian et al. 2010). GR predicts that $\eta = \mu = 1$. Figs.11 and 12 show that, even without redshifts, EMU will place significant constraints on both dark energy and modified GR models.

Once shot noise is fully overcome, ISW measurements are limited only by cosmic variance. Substantial improvements over current measurements can therefore be achieved by using a bigger survey area. Combining the EMU survey with the WODAN survey will allow us to make a radio measurement of the ISW effect over the entire sky for the first time, perhaps leading to the highest-significance ISW measurement yet.

2.9.6 Tomographic Analysis

The discussion of cosmological tests above makes the conservative assumption that no redshifts are available for EMU sources. If statistical redshifts (§3.11) are available, the radio sources can be split into redshift bins for measuring the ISW effect or cosmic magnification, enabling a 'tomographic' analysis of auto- and cross-correlation data, leading to even more significant constraints than those discussed above. For example, strong polarised sources are known to have a high median redshift, even when they are undetected in optical surveys.

If there are enough sources in each bin to achieve a statistically significant ISW detection, EMU will enable a measurement of the redshift evolution of the ISW effect and so better constrain cosmological models. The results at low redshifts will give an independent confirmation of detections achieved at optical wavelengths. However, the radio sources extend to far higher redshifts, and so test the ISW at epochs that cannot be reached at other wavelengths: a detection of the ISW at high redshift would be a clear signature of a non-ΛCDM $+$ GR evolution (Raccanelli et al. 2011).

2.10 Galactic Science

The EMU survey will include the Galactic plane, thus creating a sensitive wide-field atlas of Galactic continuum emission, which can address several science goals including:

- A complete census of the early stages of massive star formation in the Southern Galactic plane,
- Understanding the complex structures of giant HII regions and the inter-relationship of dust, ionised gas and triggered star formation,
- Detection of the youngest and most compact supernova remnants to the edge of the Galactic disk, some of which may have exploded within the past century,
- Detection of supernova remnants, especially those detected by eROSITA but which are undetected by previous radio surveys,
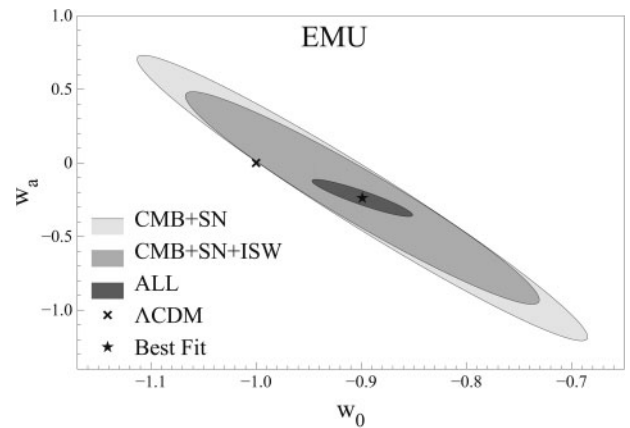

Figure 11 Predicted constraints from EMU on dark energy parameters, assuming no redshifts are available. The outer grey ellipse shows the current constraints from Type IA supernovae and CMB fluctuations. The intermediate ellipse shows the constraints from EMU using the ISW effect, providing only a modest improvement over existing measurements. The inner (dark) ellipse shows the constraints from EMU using all effects, including cosmic magnification, and shows a significant improvement over existing measurements. In particular, EMU will resolve the difference between the current best measurement (star) and the value (cross) predicted for a standard ΛCDM cosmology with non-evolving dark energy.

- Detection of planetary nebulae, which are the most abundant compact Galactic sources in the NVSS, and can be useful tools for measuring extinction and estimating the star-formation rate of stars too small to make SNe or HII regions (Condon & Kaplan 1998; Condon et al. 1999),
- Detection of radio stars and pulsars,
- Serendipitous discoveries, such as the radio flares from ultra-cool dwarfs found by Berger et al. (2001).

In providing a sensitive, high-resolution continuum image of the Galactic plane, EMU will complement the GASKAP HI Galactic HI survey (Dickey et al. 2010). Existing interferometric radio continuum surveys of the Galactic plane are either at high angular resolution but over a limited survey area, or cover a wide area at low angular resolution. For example, the MAGPIS (Helfand et al. 2006) and CORNISH surveys (Purcell & Hoare 2010) cover an area ~ 100 deg^2 at an angular resolution of 1–6 arcsec, while the International Galactic Plane Survey consists of a number of studies over several hundred deg^2 at a typical resolution of ~ 1 arcmin (McClure-Griffiths et al. 2005; Taylor et al. 2003; Stil et al. 2006; Haverkorn et al. 2006). EMU, with its full sky coverage, high sensitivity, and ~ 10 arcsec angular resolution, will bridge the gap between these two types of survey to reveal newly discovered populations of compact HII regions, planetary nebulae and young supernova remnants. When combined with the WODAN Northern Hemisphere survey (see §1.5), EMU will provide a complete census of centimetre-wave emission in the Galaxy.

At centimetre wavelengths the primary mechanisms for emission from Galactic objects are free–free emission from HII regions and planetary nebulae, synchrotron
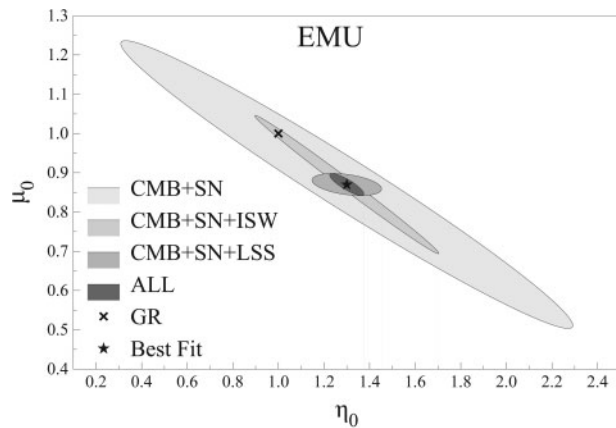
Figure 12 Forecast of constraints for modified gravity parameters, with symbols the same as for Fig. 11. The largest ellipse shows constraints from current measurements, and the smallest ellipse shows the constraints from EMU using all effects, but assuming no redshifts are available. EMU will resolve the difference between the current best measurement (star) and the value (cross) for standard general relativity, either confirming or ruling out some alternative models.

Figure 13 Spectral energy distribution of HII regions over the frequency ranges of the EMU & CORNISH surveys. The solid, dashed and dot-dashed lines represent the brightnesses of spherical isothermal and homogeneous hypercompact (HC), ultracompact (UC) and compact (C) HII regions at an assumed distance of 18 kpc. The red band shows the frequency range of the 5-GHz CORNISH survey and the blue band the range of the 1.3-GHz EMU survey. 5σ detection limits are indicated by horizontal lines.

radiation emitted by supernova remnants, and diffuse synchrotron emission emitted by relativistic cosmic-ray electrons accelerated by SNRs. The known radio populations of each of these types of object are limited by a combination of issues including the limited area covered by existing surveys, frequency-dependent selection bias (in the case of optically thick HII regions), or biases against large-scale structure introduced by limited *uv* coverage snapshot surveys.

Thermal emission from HII regions may be separated from non-thermal synchrotron emission by using the correlation between thermal free–free and infrared emission (e.g. Helfand et al. 2006; Thompson et al. 2006; Conti & Crowther 2004). The radio spectral index from EMU, and the polarisation data from POSSUM, will also be excellent discriminants between thermal and non-thermal emission. The resolution of EMU is particularly well-matched to the MIPSGAL 24 μm (Carey et al. 2009) and Hi-GAL 70 μm (Molinari et al. 2010) survey images, of which the latter is an effective tracer of the intensity of the exciting radiation field (Compiègne et al. 2010).

2.10.1. Star Formation

EMU's sensitivity allows access to all stages of the evolution of a compact HII region (hyper-compact, ultra-compact and compact), even though the 1.4-GHz continuum emission will be optically thick, and EMU's 10 arcsec resolution is insufficient to resolve the ultra-compact HII regions.

Ultracompact HII (UCHII) regions represent a young ($\sim10^5$ years old) phase in the development of an HII region (Churchwell 2002). Most UCHII regions have been discovered by snapshot VLA surveys at 5 GHz (e.g. Wood & Churchwell 1989; Kurtz et al. 1994; Walsh et al. 1998; Purcell & Hoare 2010; Urquhart et al. 2009).
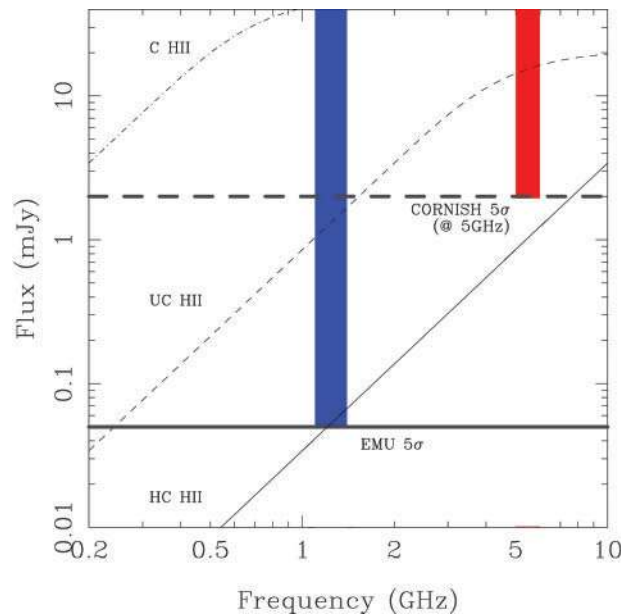
However, the snapshot surveys missed an entirely new class of optically thick HII regions known as hypercompact HII (HCHII) regions (kurtz 2005), and missed a diffuse component (kurtz et al. 1999, Longmore et al. 2009) in which many UCHII regions are embedded.

To build a complete census of such objects requires a radio survey that is sensitive to both faint objects and structure on scales of a few tens of arcsec. Fig. 13 shows the typical spectral energy distribution expected from three types of HII region (compact, ultracompact and hypercompact) located at a distance of 18 kpc. In the optically thick regime the spectral index of free–free emission is proportional to ν^2, falling to $\nu^{-0.1}$ in the optically thin regime (mezger et al. 1967). For compact HII regions this turnover occurs at ~1.5 GHz, whereas for the denser and more optically thick UCHII and HCHII regions the turnover frequency is shifted to higher values. This high-frequency turnover causes HCHII regions to be essentially undetectable with snapshot 5-GHz surveys (see Fig.13). Even though HCHII regions are extremely optically thick at 1.4 GHz, the sensitivity of EMU is sufficient to detect them at distances up to 18 kpc, except in regions limited by confusing strong sources. Consequently, EMU will be able to detect all types of HII region over most of the Galaxy, although higher-resolution follow-up observations with other arrays may be needed to separate closely associated ultracompact and hypercompact HII regions. Many of these observations will be made as part of higher-frequency surveys like CORNISH (Purcell et al. 2008) and MeerGAL (Thompson et al.

2011). Together with these surveys, EMU will be able to determine the turnover frequencies (and hence electron densities) of these regions.

2.10.2 Supernova Remnants

Only ~275 supernova remnants have so far been identified within the Milky Way (Green 2009), out of an estimated total population of between 500 and 1000 remnants (Helfand et al. 2006). Although the Green (2009) catalogue represents the result of more than ~50 years of intensive searches with the world largest radio telescopes, it is considered to be incomplete and strongly biased by selection effects, such as the lack of compact (and hence young) and faint remnants, both of which should be addressed by the high sensitivity and angular resolution of EMU. There also exist strong synergies between EMU and the high-energy observatories HESS, *Fermi* & *Chandra*, which may be used to confirm EMU non-thermal candidates as pulsar wind nebulae.

The increased surface-brightness sensitivity and higher angular resolution of EMU, combined with X-ray data from eROSITA (Predehl et al. 2010) should allow the identification of more than 200 faint and diffuse X-ray sources which are currently classified as supernova remnant candidates (Becker et al. 2009).

2.10.3 Stellar Radio Emission

In the last few years, the improvement of observational capabilities has led to the discovery of radio emission in a broad variety of stellar objects at diverse stages of their evolution, as shown in Fig. 14. In many cases, radio observations have revealed astrophysical phenomena not detectable by other methods (Güdel 2002).

Radio stars have so far been detected only by targeted observations directed at small samples of stars thought likely to be radio emitters. Existing observations suffer from limited sensitivity (e.g. no radio stars have been detected at the radio luminosity of the quiescent Sun, $L_{radio} \sim 10^4 \, \mathrm{WHz^{-1}}$), and suffer from a strong selection bias in that observations have been targeted to study a particular aspect of stellar radio emission. Consequently, it is difficult to forecast the impact of EMU on the field of stellar radio astronomy.

Much of our knowledge of radio stars comes from the study of active stars and binary systems (Slee et al. 1987; Drake et al. 1989; Umana et al. 1993, 1995; Seaquist et al. 1993; Seaquist & Ivison 1994; Umana et al. 1998; Budding et al. 1999). The strongest radio emissions appear to be associated with mass loss and magnetic phenomena and are often highly variable (Güdel 2002; Dulk 1985; White 2004). Non-thermal radio emission can also originate from shocks of colliding winds in massive binaries (Dougherty & Williams 2000).

Transient events are also observed as narrow-band, rapid, intense and highly polarised (up to 100%) radio bursts in stellar objects that have a strong (and often variable) magnetic field and a source of energetic
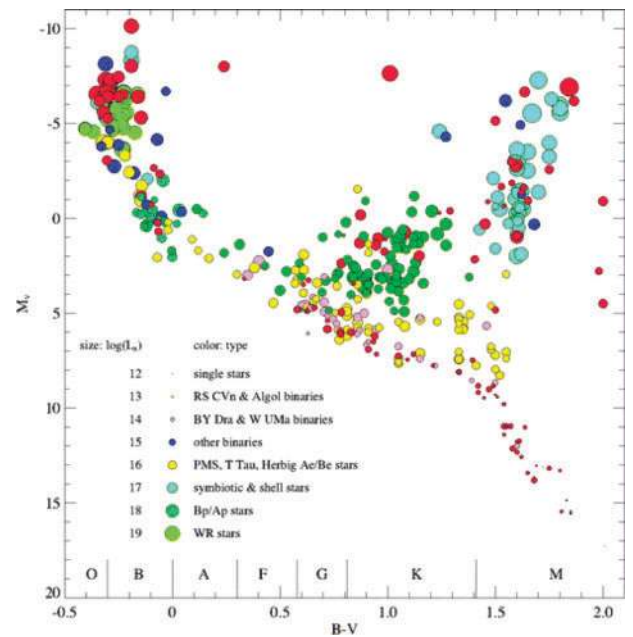

Figure 14 H–R diagram of all known radio stars, taken from Güdel et al. (2002). Radio stars can be found in almost every segment of the stellar H–R diagram.

particles, including RS CVns and flare stars (Slee et al. 2008), brown dwarfs (Hallinan et al. 2008), and chemically peculiar stars (Trigilio et al. 2000, 2008; Ravi et al. 2010). These radio flares have generally been interpreted as a result of coherent emission mechanisms, including electron cyclotron maser emission and plasma emission. This emission has been detected in only a few tens of stars, partly because of the limited sensitivity of available instruments, but mainly due to the absence of a deep all-sky radio survey.

These imperfect statistics suggest that EMU will provide an unbiased sample of several thousand stellar radio sources, enabling a detailed investigation into the physics of radio stars. This will include stellar physical parameters, magnetic activity, the fraction of active single stars and binary objects that show coherent emission, the time-scales of its variability and their relationship to stellar parameters. These are important diagnostics for studies of stellar magnetospheres, and can probe fundamental parameters such as magnetic field intensity and topology, and electron energy distribution.

However, it will be essential to distinguish radio stars from extragalactic sources using properties such as radio polarisation, SEDs, and the proper motions available from Hipparcos, TYCHO-2 and UCAC3 (Helfand et al. 1999; Kimball et al. 2009).

2.10.4 Pulsars

The number of pulsars at high Galactic latitude is poorly known, and Cameron et al. (2011) have found that none of the IFRS detected in the ATLAS survey (Norris et al. 2006; Middelberg et al. 2008a) are pulsars, implying that the surface density of pulsars with a continuum flux

level $> 150\,\mu$Jy at high Galactic latitude is <1 per 7deg^2. On the other hand, G. Hobbs (private communication) has estimated on the basis of other surveys that about one pulsar with a continuum flux level $> 50\,\mu$Jy should occur roughly every four deg^2, in which case EMU will detect about 8000 pulsars, which exceeds the total number of currently known pulsars. However, it will be difficult to distinguish these from other continuum sources. Diagnostics will include (i) lack of an optical/IR identification, (ii) steep spectrum, (iii) polarisation, especially circular polarisation, (iv) sidelobes caused by variability and scintillation during the observation. Candidates satisfying these criteria will be searched for pulsar emission using the ASKAP COAST (Ball et al. 2009) project or conventional single-dish pulsar searches.

2.11 Unexpected Outcomes

Experience has shown that whenever the sky is observed to a significantly greater sensitivity, or a significantly new volume of observational phase space is explored, new discoveries are made. Even ATLAS, which expanded the phase space of wide–deep radio surveys by only a factor of a few, identified a previously unrecognised class of object (IFRS: see §2.4). Because EMU will be much more sensitive than any previous large-scale radio survey, it is likely to discover new types of object, or new phenomena. Although it is impossible to predict their nature, we might reasonably expect new classes of galaxy, or perhaps even stumble across new Galactic populations. Furthermore, the large EMU dataset, covering large areas of sky, will be able to identify extremely rare objects.

Historically, discoveries of new classes of objects occur when an open-minded researcher, intimate with the telescope and with their science, recognises something odd in their data (Ekers 2010; Bell-Burnell 2009). EMU is unlikely to be different: people will make unexpected discoveries while carefully using EMU data to test hypotheses. It is therefore arguable that this process needs no planning: any new class of object in the data will eventually be discovered anyway. On the other hand, given the large volume of data, it is possible that a class of objects will lie undiscovered for decades because it didn't happen to fall within the selection criteria of any astronomer.

It is therefore important for EMU to plan to identify new classes or phenomena, rather than hoping to stumble across them. EMU is therefore taking the novel approach of developing data-mining techniques to identify those objects that don't fall into one of the known categories of astronomical object. A data-mining project (named Widefield ouTlier Finder, or WTF) is being established that will attempt to assign each object in the EMU catalogue to a known class of astronomical object, using the available cross-identifications to compare colours, luminosities, and any other available data. It will then identify those that are outliers or which depart systematically from known examples. Most such outliers will simply represent

bad data, and are thus valuable in their own right for debugging ASKAP, while a few may be exciting new discoveries.

2.12 Legacy Value

The largest existing radio survey, shown in Fig. 1, is the NRAO VLA Sky Survey (NVSS), whose release paper (Condon et al. 1998) is one of the most-cited papers in astronomy. EMU will cover the same fraction ($\sim$75%) of the sky as NVSS, but will be 45 times more sensitive, 4.5 times higher in resolution, with higher sensitivity to extended structures. As a result, EMU will detect $\sim$38 times as many sources as NVSS. More importantly, the greater sensitivity means that EMU breaks into a different regime. While most NVSS sources are radio-loud AGNs, EMU will provide estimates of star formation rate and radio-mode accretion activity in the galaxies currently being studied by mainstream astronomers at all other wavelengths.

The legacy value of large radio surveys depends on the ability to obtain a radio image of any object being studied at other wavelengths. Thus the legacy value depends critically on the area of sky covered: a survey covering the entire sky has more than simply twice the scientific value of a survey covering half the sky. For example, the number of powerful (FRII-type) radio sources predicted at the epoch of reionisation is only $\sim$100 over the entire sky. The ability to identify radio counterparts (or upper limits) for any future observation (for example the highest redshift gamma-ray burst or supernova host galaxy) is essential. Serendipitous discoveries, too, are most likely to be maximised by surveying the greatest possible area. As a result, EMU extends as far North as $+30°$ declination, and, together with WODAN (see §1.5), will cover the entire sky. The combined EMU and WODAN catalogues are expected to become the primary radio source catalogue for all astronomers, and will not be superseded until after the SKA begins operation.

3 EMU Design Study

3.1 The Need for a Design Study

In the two years prior to commencement of the EMU survey, the EMU design study is addressing the following challenges:

- Developing an optimum observing strategy, while taking into account the needs of potentially commensal surveys,
- Working with the ASKAP computing group to ensure an optimum processing pipeline, and an optimum source-extraction algorithm, including extraction of extended sources,
- Addressing technical issues of dynamic range, confusion, etc.,
- Developing an optical/IR identification pipeline, including a citizen science project,
- Exploring the use of statistical redshifts (see §3.11),

- Developing the data access process, and identifying potential data issues,
- Developing 'value-added' projects to maximise the science return from EMU,
- Refining the EMU science goals, to ensure that the EMU survey is optimised to address them, and planning early science papers.

The EMU design study is carried out through a number of working groups who each take responsibility for a number of goals, milestones, and deliverables. The overall design study goal is to have a preliminary processing and analysis pipeline in place by November 2011, with a fully-functioning pipeline in place by November 2012.

3.2 The Role of ATLAS in the EMU Design Study

An important component of the EMU design study is the ATLAS project (Norris et al. 2006; Middelberg et al. 2008a; Hales et al. 2011; Banfield et al. 2011), which has a sensitivity, resolution and science goals similar to EMU, but covers a much smaller area of sky (7 deg^2 surrounding the CDFS and ELAIS-S1 fields). The ATLAS survey, together with earlier observations of the HDFS (Norris et al. 2005; Huynh et al. 2005), is being used as a test-bed for many of the techniques being developed for EMU, as well as guiding the development of the science goals and survey strategy.

For example, the prototype EMU source extraction and identification pipeline will be used for the final ATLAS data release in late 2011. During the course of the ATLAS project, dynamic ranges approaching $\sim 10^5$ have been reached, which are among the highest achieved in radio astronomy, and imaging artefacts have been encountered which have been rectified in some specific cases, and a solution for the general case is being tackled. Such imaging problems must be solved not only for EMU, but for other SKA Pathfinder projects.

3.3 Science Data Processing

The data from the correlator will be reduced in an automated processing pipeline (Cornwell et al. 2011), which includes the following steps for EMU:

- Flag for known radio frequency interference, working from a database of known RFI sources,
- Identify unknown radio frequency interference, saving candidate identification, and identify and flag further bad data,
- Solve for calibration parameters (i.e. frequency-dependent complex gains) using least squares fits of the predicted visibility (from previously obtained model of the sky) to the observed visibility,
- Apply calibration parameters, predicting forward from the last previous solution,
- Average visibility data to required temporal and spectral resolution,
- Construct an image by (i) gridding the data using convolutional resampling, (ii) Fourier transforming to the image plane, and (iii) deconvolving the point-spread function,
- Find sources in the resulting image or cube,
- Save science data products to the ASKAP science data archive.

The measured flux densities from most radio surveys have a typical accuracy of 10%, although the calibration accuracy should, in principle, enable flux densities to be measured to an accuracy of $\sim 1\%$. This latter figure has been adopted as the target for EMU, and the EMU Design Study will investigate the reason that most surveys fail to achieve this figure. A number of instrumental effects and completeness corrections have been identified (Hales et al. 2011), and may need to be applied to the data, including the following.

- The primary beam model must be accurate to $\sim 1\%$ so that the data can be corrected for the primary beam response to this accuracy. In the case of ASKAP, this will vary across the PAF, and so needs to be determined for each of the 36 beams.
- Position-dependent bandwidth smearing (chromatic aberration) over the mosaicked image due to the finite bandwidth of frequency channels (Condon et al. 1998; Ibar et al. 2009). Surprisingly, mosaicked images suffer more from bandwidth smearing than pointed observations, since any location in the image, even at the centre of a pointing, may include many contributions from adjacent pointings in which bandwidth smearing is significant. Thus, even for tight angular spacing between pointings in a mosaic, bandwidth smearing will always be non-zero at any location over the final mosaiced image. Hales et al. (2011) found that bandwidth smearing in central regions of the ATLAS mosaicked images, which used 8-MHz channels, typically caused a 10% decrement in peak flux due to this overlapping effect. Even with ASKAP's 1-MHz channels, this correction will be necessary if 1% accuracy is to be achieved.
- Position-dependent time-average smearing over the mosaicked image due to finite integration time and Earth rotation (Bridle & Schwab et al. 1999). This is thought not to be a significant issue for ASKAP, but is included here for completeness.
- Clean bias in measured fluxes, which may be mitigated by suitable weighting schemes (to control beam side-lobe properties) and cleaning depth.

In addition, there are observational effects which may bias the fluxes of individual sources and/or the statistical properties of the sample as a whole, as follows.

- The probability of measuring a faint source located on a noise peak is higher than the probability of measuring a strong source located in a noise trough, because faint sources are more numerous (Hogg & Truner 1998). This results in a bias on all measured source fluxes, which decreases with increasing signal-to-noise ratio. When considered in terms of source counts this is

known as Eddington bias (Eddington 1913; Simpson et al. 2006). Hales et al. (2011) have shown that Eddington bias can also be a sensitive probe of number counts below the flux limit of a survey.

- Resolution bias due to the lack of sensitivity to resolved sources, which can be manifested in two ways. First, the lack of short baselines can limit the maximum angular scale observable, although this is not likely to be a problem for EMU. Second, faint resolved sources may have integrated fluxes sufficient to be included in the final catalogue, but may be missed in the source extraction process because their peak fluxes fall below the signal-to-noise source detection threshold (e.g. Prandoni et al. 2001). This necessitates a resolution bias correction.

- Sensitivity will generally vary across an image, so that the area surveyed to any limit is a function of that limit. Since source counts must be normalised by the area-surveyed, a completeness correction must be applied to account for the position-dependent sensitivity to faint sources across the mosaicked image (e.g. Bondi et al. 2003), which must also take into account the position-dependent time-average and bandwidth-smearing effects (Hales et al. 2011). For the same reason, large variations in sensitivity across the field are likely to lead to uncertainties in source counts, and so it is essential to measure the sensitivity across the field to high accuracy, and keep counts this is known as Eddington it as uniform as possible.

3.4 Simulations and Imaging Pipeline

The goals of the simulations and imaging pipeline working group are:

- To develop a realistic extragalactic simulated sky,
- To develop a realistic Galactic simulated sky,
- To work with ASKAP engineers to choose the optimum PAF configuration and weighting scheme,
- To ensure that ASKAP imaging processes are robust to Galactic-plane observations,
- To optimise imaging algorithms and parameters,
- To ensure that EMU data will reach a dynamic range of 10^5,
- To ensure that EMU-measured flux densities are accurate to 1%.

3.4.1 Simulation Analysis

Simulations have been performed using the SKA design study (SKADS) simulated sky (Wilman et al. 2008) as initial input. This sky is augmented by the addition of extended, diffuse, or complex sources, which are not fully represented in the SKADS sky. This sky is then used to generate uv data similar to that which will be produced by ASKAP, using the antenna, beam and array characteristics. Thermal noise is added to the visibilities but calibration is assumed to be ideal. The data are then processed using the *ASKAPsoft* processing pipeline
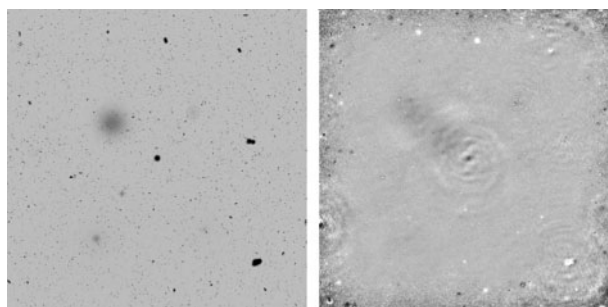

Figure 15 Part of the simulated sky as observed by the simulated ASKAP telescope and processed using *ASKAPsoft*. (Left) The input model sky. (Right) The difference between the observed sky image and the input sky showing artefacts caused by the observing process and deconvolution errors. The intensity scale of both images is increased to highlight these errors.
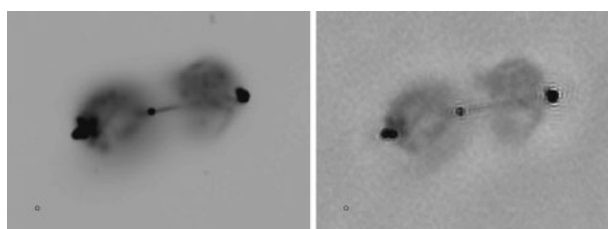

Figure 16 (Left) The simulated sky around Pictor-A as observed by the simulated ASKAP telescope and processed using *ASKAPsoft*. The greyscale ranges from -0.01 to 0.03 Jy/beam, with a peak of 0.9 Jy/beam. (Right) The difference between the observed sky image and the input sky showing artefacts. This image is enhanced by a factor of ten to show up artefacts, so the greyscale ranges from -0.001 to 0.003 Jy/beam.

(Cornwell et al. 2011) which is the prototype version of the ASKAP processing software. Current simulations use all 36 antennas, the full 300-MHz bandwidth, and 8 hours of on-source time with approximately uniform weighting. A typical observed simulated sky is shown in Fig. 15. Results so far suggest that the majority of fields will reach the required sensitivity and dynamic range, and the rest of this section considers those few remaining fields containing strong sources, such as that shown in Fig. 16, where this may not be the case.

The analysis of the simulations focusses on the difference image, which is the difference between the dirty image and the input model convolved with the point-spread function (PSF). In regions containing no strong sources, the rms noise in the difference image is close to the value predicted from receiver performance (i.e. 10 μJy for an EMU image). However, it is higher in regions containing bright or extended sources, because of the following effects.

- Insufficient sampling in the image domain, particularly for bright point sources that are not pixel-centred Briggs & Cornwell (1992). Simulations suggest that 4 pixels/beam will limit the dynamic range to ~2000–3000 in these regions, and so ~5–6 pixels across the beam are planned to achieve the required 10^5 dynamic range.

- Insufficient scales for the multi-scale CLEAN. The three scales used in earlier simulations were insufficient to deconvolve the more extended sources in the field, and so five scales (as used in the simulation results shown in Fig. 15) will be implemented in the final processing pipelines.
- A large CLEAN threshold. A CLEAN threshold of 1 mJy/beam was used in the simulation, which will leave residual lobes of ~1 μJy/beam around sources slightly weaker than 1 mJy/beam. A lower value of CLEAN threshold will be used for ASKAP.

As described in §1.2, ASKAP will generate a sky model of sources stronger than ~1 mJy, which will be subtracted from the observed *uv* data before imaging. As a result, many of the effects described above are not expected to be significant in ASKAP data, but simulations are continuing to test this hypothesis.

3.4.2 Future Simulations

A number of simulations are needed to ensure that the pipeline will deliver EMU science:

- Simulate low–surface brightness large-scale structures (e.g. Fig. 16) and test the effectiveness of the pipeline in retaining these structures,
- Ensure that preconditioning (see §3.4.3) can recover the source structures required for EMU science,
- Simulate faint (well below the rms) low–surface brightness large-scale structures to check that they can be recovered using correlation and stacking techniques,
- Test deconvolution effectiveness on sources with curved radio spectra.

3.4.3 Tapering

Radio synthesis imaging offers a trade-off between naturally weighted images (to optimise sensitivity) and uniformly weighted images (to optimise resolution). The ASKAP processing software will use preconditioning (Cornwell et al. 2011), rather than traditional weighting, and multi-scale clean (Cornwell 2008), which reduces the impact of these effects. Nevertheless, in effect both naturally weighted and uniformly weighted data may be required, implying two processing pipelines for continuum data. Extended structures may require an even more heavily tapered weighting. Future simulations will explore potential solutions, including the use of Briggs robust weighting that may provide an optimal compromise between the two. They will also explore the possibility of off-line processing to generate more heavily tapered images than those produced in the real-time pipeline, and to determine the optimal on-line weighting to facilitate this with minimum loss of information.

3.5 Source Extraction and Measurement

Several radio source extraction tools are widely used, such as the *Miriad* / AIPS Gaussian fitting routines IMSAD, SAD and VSAD, SExtractor (Bertin & Arnouts

1996), SFIND (Hopkins et al. 2002) and Duchamp (Whiting 2008). There are also numerous specific tools, such as HAPPY (a modified version of SAD used in the FIRST survey (White et al. 1997)), a machine-learning back-end to VSAD, used to construct the SUMSS catalogue (Mauch et al. 2003), BDSM (used for the LOFAR source-finding, N. Mohan, in prep), and the flood-fill algorithm being used in the ATLAS Data Release 2 (Murphy et al. 2007; Hales et al. 2011).

However, none of these is fully adequate for EMU, particularly (i) in extended sources, (ii) in the presence of artefacts from bright sources, and (iii) in images whose signal-to-noise ratio varies across the image. It appears that an optimum strategy has not yet been constructed, particularly for extended sources, and there is still no truly general-purpose and robust source-finding tool appropriate for the application to large-area radio surveys. Moreover, there has not been a published analysis that thoroughly and quantitatively compares the details of source-finding (from defining thresholds, identifying pixels to fit, fitting approaches, approaches to multiple-component and extended sources, etc.) as implemented in different existing tools.

Since the size of the EMU survey precludes manual intervention in the source extraction process, it is necessary to build efficient source extraction algorithms that are robust to varying signal-to-noise across the image, sensitive to source scales ranging from unresolved through complex and multiple-component sources, to extended low–surface brightness structures, and that require no manual intervention. The design study is therefore building on existing expertise to identify the optimum approach to source-finding, and produce a robust continuum source-finder for both point sources and extended emission.

The primary goal is to develop an algorithm (or suite of algorithms) that will be implemented both in the *ASKAP-soft* data-reduction pipeline and also in at least one of the standard radio-astronomical reduction packages, such as *casa*, as well as being published in a journal paper. The software tool or tools will also be made available for use by all members of the astronomical community for processing data from other instruments.

Specific goals are to develop optimum source identification and measurement strategies and algorithms, quantify the effects and limitations of confusion, and identify mitigation strategies for source-measurement of overlapping or multiple sources.

The problem is split into two parts, one dealing with point sources (including marginally extended sources still well-modelled by Gaussian fits), and the other with extended and complex source structures. These initial approaches will be extended to include more sophisticated investigations, contrasting wavelet or other decomposition approaches with machine-learning and genetic algorithms, in particular for the extended and complex source identification case, where information available at other wavelengths may also be introduced.
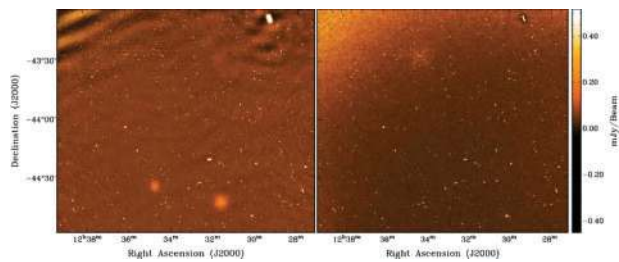
Figure 17 Left: a portion of the original SST3 continuum simulation, including artefacts in the top left-hand corner caused by a strong source just outside the field; Right:result of applying a spatial filter to remove large-scale structure. Note that the diffuse Gaussians in the South are also completely removed. Both images have the same scaling, with an rms of ∼15 µJy/beam.

An important part of this project is to build linkages with other survey groups through SPARCS (see §3.13) so that the various efforts underway worldwide can be coordinated, taking advantage of best practice where established, and combining resources where practical.

3.5.1 Compact Source Extraction and Measurement

The point-source investigation is comparing different background subtraction approaches (e.g., the polynomial mesh of SExtractor and the erosion/dilation algorithm of Rudnick (2002)), different approaches to thresholding (e.g. simple 'canonical' thresholds of $n\sigma$, and false discovery rate (FDR) methods (Hopkins et al. 2002)), and different approaches to Gaussian fitting.

This includes a comparison of existing source-finding and measurement tools, to identify their advantages and disadvantages, and to determine whether improvements can be incorporated into the existing ASKAP source-finder (Duchamp: Whiting 2008). For each tool, the completeness (fraction of artificial sources identified) and reliability (fraction of identified sources that are true model sources, as opposed to false detections) are measured as a function of flux density.

While all source-detection algorithms become incomplete and unreliable at low signal-to-noise, detailed comparisons can identify whether particular source-finders are more reliable at fainter levels. Furthermore, algorithms with the same completeness and reliability can produce different source catalogues. How these catalogues differ, and how they affect the science of the ASKAP projects, is currently being studied (Hancock et al. 2011).

The primary result so far has been to recognise the importance of accurately measuring the background noise level, which may vary across the field because of artefacts, diffuse emission, and varying sensitivity, and subtracting any varying low–surface brightness emission. Given the size of ASKAP images, this background variation often appears on scales larger than the filters coded into standard source-finders. Errors on the measured flux densities are clearly correlated with the background structure, which can be removed using a filter such as that shown in Fig. 17. However, it may be challenging to determine an appropriate scale size for the filter that can be used automatically and robustly in EMU observations.

The next step is to understand why sources are identified by one source-finder but not by others. This is often caused by the way tools group pixels into objects, such as objects assigned as extended by SExtractor, but measured as multiple separate sources by Gaussian fitting tools such as VSAD and Sfind. This analysis will then be extended to explore those objects that are poorly fitted by any of the source-finding tools, to produce a list of recommendations for implementation in the ASKAP source-detection pipeline to achieve optimal source extraction and measurement.

3.5.2 Extended and Complex Source Extraction and Measurement

The investigation of optimising detectability of extended structure involves exploring a variety of filters that maximise the detectability of extended structure, such as a spatial scale filter (Rudnick & Lemmerman 2009) to identify extended emission, as shown in Fig. 17.

Other approaches include the use of the Hough transform (Ballard 1981) to measure circular or elliptical structures (such as SN remnants, or radio relic sources), and using compressed sensing imaging techniques to radio interferometric data (Wiaux et al. 2009). The study focusses on metrics to compare techniques, and implementing suitable model sources in the simulations. Each technique seems to hold promise for different kinds of structure and measures, and the outputs are now being quantified and compared, to identify a suitable implementation for the ASKAP pipeline.

3.6 Survey Strategy

One challenge in planning a large survey like EMU is to develop the optimal survey strategy to achieve the primary science goals, while taking into account the capabilities and limitations of ASKAP and the need for the most efficient use of telescope time. Prior to commissioning, several key issues must be considered.

3.6.1 Source Density and Confusion

Condon (2008) has shown that a beamsize of 10 arcsec is near-optimum for ASKAP in terms of the trade-off between resolution and confusion, and that the rms sensitivity is increased by ∼10% by confusion noise. This is consistent with previous observations (Hopkins et al. 2003; Norris et al. 2005; Huynh et al. 2005) which have shown that the confusion level at the rms sensitivity (10 µJy/beam) and resolution (10 arcsec) of EMU is ∼50 synthesised beams per source, so that only a small fraction of EMU sources will be affected by confusion. Nevertheless, several science goals of EMU, in particular those focusing on diffuse emission and stacking experiments, rely on solid understanding and mitigation of confusion and noise variations.

Simulations conducted by the ASKAP computing team, supplemented by model diffuse sources, confirm that confusion is not the limiting factor at 10 arcsec, but
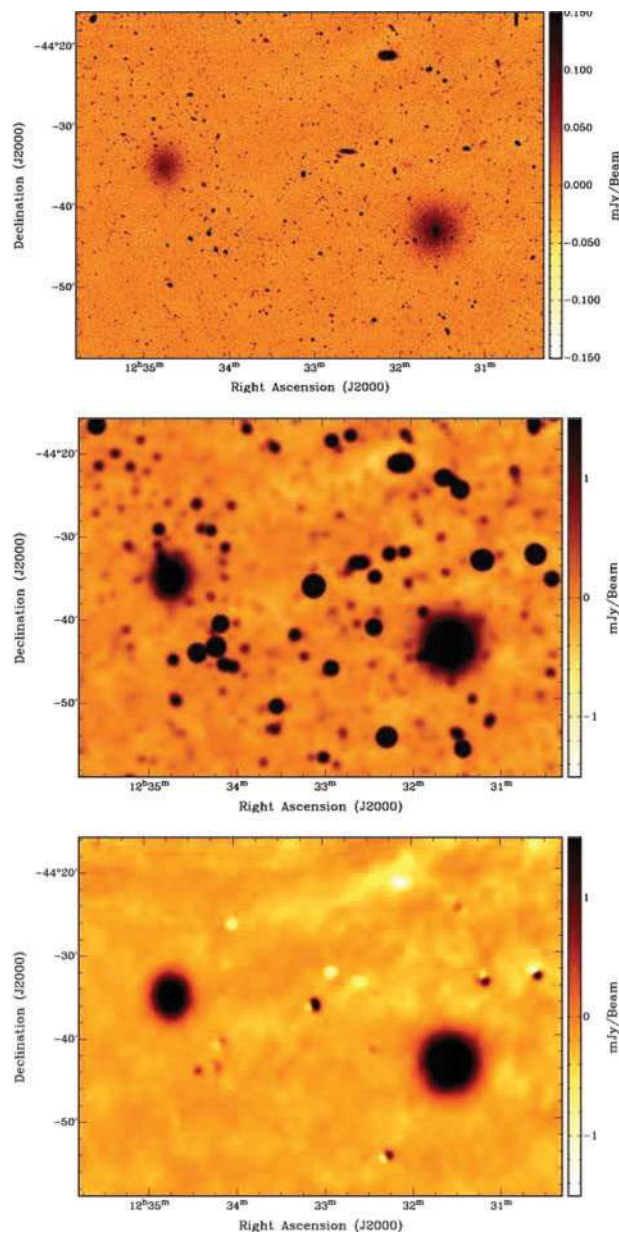
Figure 18 Top: Original 10-arcsec resolution ASKAP simulation with two diffuse radio haloes injected; Centre: Same as above convolved to 60 arcsec resolution. The image is confusion-limited at 40–100 μJy/beam; Bottom: Same as above but with the SKADS source model subtracted (with errors). This image reaches thermal noise on small scales, with a typical rms of 20–50 μJy/beam.

will be an issue at the 1 arcmin resolution ideal for the detection of large-scale diffuse emission. Compact source subtraction and filtering is being explored in order overcome this issue. Initial experiments (e.g., Fig. 18) have improved our understanding of these issues, and will inform the diagnostics needed in both the science and imaging simulations.

3.6.2 Choice of Observing Frequency

EMU will observe a single contiguous band at 1130–1430 MHz, which is a compromise between (i) the reduced PAF beam filling factor above 1400 MHz,

(ii) the poorer spatial resolution at lower frequencies, and (iii) the desire to observe commensally with the WALLABY project, for which the frequency range is determined by redshifted hydrogen. On the other hand, other factors (design constraints, choice of commensal observing programs) may necessitate revising this choice, and some continuum data may also be available as a byproduct of other ASKAP projects (e.g. POSSUM, DINGO, FLASH) . Therefore the EMU design study simulations are including observing frequency as a free parameter.

3.6.3 Observing Strategy

To achieve the expected sensitivity in EMU it will be necessary either to integrate at each pointing for ∼12 h and then move over to the next pointing, or take many short integrations at each pointing. The first option allows the required depth to be reached without requiring any historical data, and is robust to variability (on timescales >12 h), and changes in the dirty beam caused by instrumental failures. It will also allow an early identification of any instrumentation or systematic issues that may restrict our dynamic range. The second option is difficult to implement as it requires the processing pipeline to have access to historical data, but will detect variability, and will provide graceful degradation to array faults. For example, a hardware failure on a particular day may result in a reduction of sensitivity over several fields, rather than causing a total loss of one field. Option 1 is maximally commensal with WALLABY while option 2 is maximally commensal with VAST. It is likely that a variant of option 1 will be adopted, augmented by some dithering or interlacing to achieve the necessary sensitivity uniformity (see §3.6.6).

3.6.4 Tiling Strategy

Because of computational overheads, it will be inefficient for the ASKAP pipeline processor to combine data from overlapping pointings obtained on different dates. The loss of observing efficiency caused by overlapping pointings is minimised by the square feed array that has been adopted for ASKAP. Initial tiling experimentation has begun with tiling stripes of constant declination, motivating a coordinated EMU/WALLABY move towards a unified formalism.

The trade-offs between tiling strategies and PAF weighting schemes are not yet fully understood. Weighting schemes may be optimised for (i) maximum field of view, (ii) maximum sensitivity, (iii) uniform sensitivity, or (iv) best polarisation response, and that decision may influence the tiling strategy. A further factor is the need to combine (or 'stitch') images and catalogues across adjacent fields, which may also have some impact on tiling strategy.

3.6.5 Dynamic Range

Dynamic range may be the biggest factor in limiting the uniformity and completeness of EMU, and it drives

many of the ASKAP design requirements, such as the 3-axis telescopes, the strict pointing requirements, and the good gain stability of the antenna front-end subsystems. It also drives much of the current EMU design study, such as the simulations (see §3.4).

Each 30 deg^2 field will contain about two sources stronger than 1 Jy. The target rms noise is 10 µJy/beam, and so in principle a dynamic range of $\sim 10^5$ across the 30 deg^2 FOV is required. Although the effect of a strong source on adjacent beams is currently unclear, the 1 Jy sources will not be present in all the sub-beams within the 30 deg^2. Therefore, even if dynamic range limits some sub-beams, 10 µJy/beam rms should still be reached in the majority of sub-beams. Simulated ASKAP data is being used to test these effects and others, such as the effect of strong diffuse emission from the Galactic plane. Other issues include the optimum weighting of the PAF feeds, the required uniformity across the beam, the optimum observing strategy, and choice of observing parameters such as weighting. So far, simulations (see § 3.4) suggest that the dynamic range will not be limited by the data processing, but an unambiguous answer must await the characterisation of the PAFs in the BETA observations.

3.6.6 Sensitivity Uniformity

Sensitivity can in principle vary significantly across the field of view of a single pointing, although this will be mitigated by a planned dithering pattern that should reduce non-uniformity to below 10%. Nevertheless, it is important to estimate what the required uniformity is for EMU science goals before a final observing/dithering strategy is reached. We consider two cases.

- Galaxies, AGNs, & star formation: Completeness is a critical component, and science goals permit some sensitivity variations provided the target 10 µJy/beam rms is reached at the completeness cutoff. However, large sensitivity variations result in loss of observing time: a 20% gain variation in one spot would require the completeness threshold to be set 20% higher across the field, and since sensitivity $\propto$ (observing time)^2, this will result in $\sim$40% of the data being discarded.
- Diffuse emission: The scale of any sensitivity variations must be much smaller or much larger than the diffuse emission of interest, otherwise the variations act as confusion noise. Simulations have shown that variations larger than $\sim$20–25 arcmin can be effectively removed through spatial scale filtering. Noise variations on scales between 2 and 20 arcmin are the most problematic, though detailed knowledge and modelling of the variations will enable the data to be partially corrected for such variations before diffuse source extraction.

3.7 Commissioning and Science Verification

The EMU team will play an active role in the ASKAP science verification tests, which will include observations of a well-studied field with existing 20-cm data with the same sensitivity and resolution as EMU. In addition, that field will be observed regularly throughout the commissioning and operational life of ASKAP. Not only will such a process enable checks against images and results obtained with other instruments, but it will also provide a regular check on stability, and give warning of subtle problems that may arise during the lifetime of ASKAP.

These observations will also result in a set of radio data which can be co-added to provide an extremely deep field down to the confusion limit, and so the field should be chosen to maximise the availability of multi-wavelength data. Probably the best-studied field in the sky is that surrounding ATLAS-CDFS, which offers:

- 30 deg^2 which has been mapped with the VLA at 20 cm as part of NVSS to an rms of 0.45 mJy/beam,
- 4 deg^2 which has been mapped with the ATCA at 20 cm as part of ATLAS, to an rms of 10 µJy/beam, reproducing the rms, resolution, and dynamic range that are closely matched to the design study goals of ASKAP, including polarisation, spectral index, wide-field VLBI, and variability studies,
- Probably the most intensely studied field in the sky at optical IR and X-ray wavelengths, resulting in important science from the repeated (and hence deep) ASKAP observations,
- Areas of sky which have been mapped to an rms of 10 µJy/beam situated a few degrees away from sources with flux densities >1 Jy, which will be invaluable for checking dynamic range.

The first observations will take place on BETA (see § 1.2). Although the instantaneous sensitivity and spatial resolution of BETA are modest, it will be unmatched for surveying large diffuse structures until the advent of the full ASKAP array. It is not yet clear whether BETA will be fully committed to engineering tests, or whether some science might be achievable with BETA. For the latter eventuality, science targets have been chosen that can yield early science publications from ASKAP, including observations of clusters and superclusters including large diffuse galaxies, relics, and low–surface brightness shock structures.

Once the full ASKAP array is operational, the EMU survey will start with a number of stringent data quality checks before being released into the public domain, and this quality control process, described in §4.1, is the responsibility of the EMU team. In particular, it is planned to observe one or more fields that have been deeply surveyed by other telescopes, such as the ATLAS and COSMOS fields, to check for consistency of results and calibration. It is planned also to survey a large Northern field at a declination about +25° which will also have been observed with the EVLA, and which will subsequently be observed with Northern surveys such as WODAN to ensure consistency between the surveys.

Table 2. EMU Primary Data Products

UV DATA

uv data will be available in all four Stokes parameters, and in all 300 frequency channels.

IMAGES (in IQUV)

An intensity image,
A coverage and sensitivity map
An uncleaned residual image[a],
Images of spectral index and spectral curvature,
Point-spread functions across the image[b],
Detailed metadata for all the above

CATALOGUE OF COMPONENTS[c]

A unique IAU-registered component name,
Ra, Dec (J2000) of the peak,
A peak and integrated flux density,
The spectral index of the peak and integrated flux,
The parameters of a fitted Gaussian,
For failed Gaussian fits, an estimate of the extent of the component,
A postage-stamp image,
A history of changes to the catalogue entry.

---

[a]The residual image after cleaning, in which cleaned components, and sources in the ASKAP sky model, have not been added back in. This image will be valuable for stacking and for detecting diffuse emission.
[b]The point-spread function of the uncleaned emission, which will vary across the FOV. It is important to know this point-spread function at any point to calculate the statistics of stacked images (see §3.10).
[c]All measured quantities will be accompanied by an estimate of their standard error.

3.8 Data Format and Access

3.8.1 EMU Data Requirements

An EMU data requirements document has been continuously updated since the earliest days of EMU. The resulting list of primary data products (termed 'level 5 products' in ASKAP documentation), is summarised in Table 2. However, new requirements continue to arise. For example, a recently recognised requirement is to optimise EMU data delivery for stacking. To correctly interpret stacking results requires detailed knowledge of the point-spread function at each point of the image, resulting in further data requirements such as (i) storing uncleaned residual images without cleaned components being added back in, (ii) storing (or being able to reconstruct) an accurate representation of the dirty beam (which varies significantly across the field) at each point in the image.

In addition to the science-driven data requirements, discussed here, a number of statistical tests will be made on the data to check the data quality, as described in § 4.1.

3.8.2 Value-Added Data

EMU has an ambitious plan to provide source associations and identifications in near real-time, and place them in the public domain. These value-added data constitute the 'secondary data products', or, in ASKAP terminology, level 7 data. Their generation includes grouping radio components into sources (e.g. classical radio double and triple sources) and cross-identifying with optical and infrared surveys. EMU is also collaborating with the Galaxy Zoo group to extend this process into the 'citizen science' arena (see §3.9). To achieve this, collaborations have been initiated with the appropriate surveys, and discussions are taking place on how best to produce these associations and cross-identifications, using a probabilistic approach. In addition, NED (NASA Extragalactic Database) has agreed to accept a near–real time automated query for this process, and have also agreed to host the resulting source catalogue and associations.

3.8.3 Virtual Observatory Server

Management of primary ASKAP data will be the responsibility of the ASKAP Science Data Archive Facility (ASDAF). However, existing Virtual Observatory (VO) standards have not yet been fully implemented on a large radio survey, and challenges will probably be encountered when we do so. To identify these, an internal prototype VO server and associated MySQL database is being developed for ATLAS data and is now in regular use by ATLAS team members. It is expected that further VO development may include modifications to VO protocols, or a requirement for additional VO Tools. The EMU design study includes plans to place ATLAS cross-identifications and other data (e.g. spectroscopy) on the ATLAS VO server, as a step to identifying the EMU VO requirements.

3.9 Source Characterisation and Cross-Identification

The EMU component catalogue will be cross-matched in a near–real time automated pipeline with available large-area optical/IR surveys shown in Table 3. For each survey, the quoted sensitivity has been used to estimate the fraction of radio sources in EMU with $S_{20\,cm} > 50\,\mu Jy$ that will be detected by that survey, based on the measured flux densities of sources in COSMOS (Scoville et al. 2007; Schinnerer et al. 2007).

Experience with surveys such as ATLAS and COSMOS indicates that ∼5% of EMU radio sources will consist of multiple components such as radio double or triple sources with a core-lobe morphology. A key goal of the EMU design study is to develop an automated cross-identification pipeline, significantly more sophisticated than a simple nearest-neighbour algorithm, to group the 70 million radio components detected by EMU into 'sources', and cross-identify them with optical and infrared sources found by other surveys. The algorithm will be run (in an automated pipeline) on the primary components catalogue described in Table 2 to produce a secondary data product, the 'source' catalogue, in which components are grouped together to form sources, and are associated with optical/IR identifications. In cases where several cross-identifications are possible, each will be given in the source catalogue along with an estimated probability.

This design study builds on the experience developed through a number of international and multi-wavelength surveys in which EMU team members are named

Table 3. Key multi-wavelength surveys with which EMU data will be cross-identified (restricted to surveys larger than 1000 deg^2). All magnitudes in AB. The EMU detected column is the fraction of EMU sources that are detectable by the multi-wavelength survey to its 5σ limit. The survey matched column is the fraction of EMU sources in the previous column which are in the area of sky covered by the multiwavelength survey. The sensitivity shown for the WISE survey is for the 3.5 μm band.

Survey name	Area (deg^2)	Wavelength bands	Limiting mag.or flux[a]	EMU detected (%)	Survey matched (%)	Data release date
WISE[1]	40000	3.4, 4.6, 12, 22 μm	80 μJy	23	100	2012
Pan-Starrs[2]	30000	g, r, i, z, y	r < 24.0	54	50	2020
WALLABY[3,b]	30000	20 cm(HI)	1.6 mJy[c]	1	100	2013
LSST[4]	20000	u, g, r, i, z, y	r < 27.5	96	67	2020
Skymapper[5]	20000	u, v, g, r, i, z	r < 22.6	31	66	2015
VHS[6]	20000	Y, J, H, K	K < 20.5	49	66	2012
SDSS[7]	12000	u, g, r, i, z	r < 22.2	28	22	DR8 2011
DES[8]	5000	g, r, i, z, y	r < 25	71	17	2017
VST-ATLAS[9]	4500	u, g, r, i, z	r < 22.3	30	15	2012?
Viking[10]	1500	Y, J, H, K	K < 21.5	68	5	2012
Pan-Starrs Deep[2]	1200	0.5–0.8, g, r, i, z, y	g < 27.0	57	4	2020

[a]Denotes 5σ point source detection. However, in many cases, *a priori* positional information will enable 3σ data to be used, resulting in a higher detection rate.

[b]Being an HI survey, WALLABY will measure redshifts for all detected galaxies out to $z = 0.26$.

[c]per 4 km s^{-1} channel achieved in 8 hours integration.

[1]Wright et al. (2010)

[2]Kaiser et al. (2010)

[3]Koribalski et al. (2011)

[4]Ivezic et al. (2008)

[5]Keller et al. (2007)

[6]http://www.ast.cam.ac.uk/research/instrumentation.surveys.and.projects/vista

[7]Abazajian et al. (2009)

[8]The Dark Energy Survey Collaboration (2005)

[9]Shanks (2005)

[10]http://www.eso.org/sci/observing/policies/PublicSurveys/sciencePublicSurveys.html

investigators, and some of whom are developing similar tools for *Herschel*, *Spitzer* and SCUBA/SCUBA-2 cross-identifications (e.g. Roseboom et al. 2009, 2010; Smith et al. 2010; Chapin et al. 2010).

Given the expected source density, and redshift distribution, of both radio sources and potential matching catalogues (see Table 3), a likelihood-ratio matching technique (Sutherland & Saunders 1992; Smith et al. 2010) or similar (i.e. Budavari & Szalay 2009; Roseboom et al. 2009) should be adequate in the majority of cases, especially if source SED/colour information is incorporated (Roseboom et al. 2009; Chapin et al. 2010). The $\sim$5% of radio sources (see e.g. Norris et al. 2006) consisting of multiple components represent a considerable challenge to automated cross-identification algorithms, and our proposed algorithm will need to account for these. Some potential solutions (i.e. using geometric information in matching, simultaneous radio self-matching and ancillary data-matching) are currently being developed, and will be tested as part of the design study.

We will use Monte Carlo simulations to derive probable misidentification rates for various classes of source, and will publish the algorithm and the results of the simulations in a journal paper, as well as provide code to the astronomical community. During the early part of the design study, ATLAS data will be used to investigate possible algorithms. The later part of the design study will include setting up collaborations, ensuring access to appropriate surveys, constructing software for the pipeline processor, and running simulations to measure the misidentification rates.

Applying the limits shown in Table 3 to the sources detected in ATLAS and COSMOS, it is expected that, of the 70 million sources detected and catalogued by EMU, $\sim$40% will initially have identified counterparts in wide-field optical and IR surveys (e.g. SkyMapper, VHS, WISE, LSST) and a further few percent will be identified by smaller but deeper surveys such as HERMES (Seymour et al. 2011). This fraction will rise to $\sim$70% in 2020 with the release of the LSST data.

Of those 30 million sources with potential identifications initially, we estimate that $\sim$25 million will yield reliable cross-identifications using a simple likelihood ratio algorithm as discussed above, and a further 3 million will be handled satisfactorily by a more sophisticated cross-identification algorithm which will take account of the spectral energy distribution at other wavelengths. For example, three radio components in a line could either be three low-redshift star-forming galaxies, or a high-redshift FRII galaxy. Three IR galaxies with a star-forming SED coincident with the three radio components would indicate the first solution, whereas a bright elliptical component at the centre, and no identification with the outer components, would indicate the latter solution.

The remaining 2 million are likely to be extended, confused, or complex sources for which automated algorithms are unlikely to be able to make a confident identification. Ironically, these are the sources most likely

to yield the most interesting science, but a human brain will be required for an optimum cross-identification, and there are simply not enough human brains within the EMU team (or within the global community of professional astronomers) to look at each galaxy.

In cases where the automated pipeline fails, EMU will use a 'citizen science' approach to enlist the help of thousands of enthusiastic amateurs, in collaboration with the Galaxy Zoo group (Lintott et al. 2008) . Galaxy Zoo have used this technique very success ully for classification of SDSS galaxies, and have agreed to collaborate with EMU on creating a 'Radio Zoo'. An initial prototype Radio Zoo project will be established in mid-2011 using ATLAS data. If resources permit, machine-learning cross-identification will also be explored, following the approach of Djorgovski et al. (2010) who have extended the 'citizen science' approach by using it to train neural nets and other algorithms.

3.10 Stacking and Data-Intensive Research

EMU has an additional requirement for facilities equipped for data exploration and mining, and data-intensive research, such as the technique of stacking.

The process of stacking involves identifying a class of objects which are generally below the detection limit of a survey, but then combining (typically by taking a censored mean or median) the data at the position where such objects are expected in the survey. The noise tends to cancel, while any low level of flux density in the sources adds, resulting in a detection threshold very much lower than that of the unstacked survey. Stacking at radio wavelengths has been used very successfully (e.g. White 2007; Ivison et al. 2007; Ivison et al. 2010; Dunne et al. 2009; Messias et al. 2010; Bourne et al. 2011) on high-resolution data for the purpose of studying faint populations which are below the detection threshold of the radio image, and has proven to be a powerful tool for studying star formation rates, AGN activity, and measuring the fraction of the extragalactic background contributed by various source populations. For example, the radio–infrared correlation can be probed down to levels well below the threshold of any radio survey by averaging the radio flux density at the position of all galaxies with an infrared flux density within a particular bin (Boyle et al. 2007).

The unprecedented area–depth product of EMU makes it remarkably suitable for the stacking process. For example, stacking a sample of a million optically selected galaxies in the EMU data will result in a noise level of $\sim$10 nJy. Because of the wide area of EMU, even rare classes of source can be stacked, and it should also be possible to create stacked images of extended sources, such as clusters. However, the extent to which such deep stacking will be successful in EMU will depend on the extent to which imaging artefacts are cancelled by stacking, and the effectiveness of stacking EMU data will be simulated, and stacking science projects conducted using ATLAS, NVSS, and FIRST,

to see whether particular observing strategies or reduction algorithms are necessary.

In particular, it has been established that stacking can be biased by a number of effects such as failure to account for the point-spread function, particularly in the presence of confusion and in highly clustered fields (Bourne et al. 2011; Greve et al. 2010; Penner et al. 2011; Chary et al. 2008; Béthermin et al. 2010) and so stacking experiments and simulations will be conducted during the design study, (i) to determine what requirements these place upon the EMU data (e.g. storing the point-spread function at all positions across the primary beam) and (ii) to establish stacking algorithms and software which can be used robustly on EMU data.

Other examples of data-intensive research include:

- Identification of sources which do not fit into known categories of radio source, and so are likely to be artefacts or exciting new classes of source, as discussed in §2.11.
- Extraction of low–surface brightness emission cross-correlated with millions of galaxies selected from other surveys. For example, this can be used to detect the synchrotron emission from cosmic filaments or sheets.
- Cross-correlation of low-redshift galaxies with high-redshift galaxies, or the CMB, to test cosmology and fundamental physics as discussed in §2.9.

The computational facilities required for such data-intensive research are currently unfunded, but it is hoped that the computational facilities associated with the ASDAF will become available.

3.11 Measuring Redshifts

To interpret data from EMU, redshifts are invaluable. Most of the science goals discussed in §2 assume that redshifts will not be available for most EMU sources, although much of the EMU science can be greatly enhanced if redshifts *are* available. However, no existing or planned spectroscopic redshift survey can cover more than a tiny fraction of EMU's 70 million sources.

For nearby galaxies, H I redshifts will be available from WALLABY, which will provide $\sim$5$\times$10^5 redshifts, and smaller numbers will be provided by other redshift surveys such as SDSS (York et al. 2000) and GAMA (Driver et al. 2009). The remaining $\sim$99% of EMU galaxies will not have spectroscopic redshifts.

Photometric redshifts, in which SEDs of various templates are fitted to the measured multi-band photometry of target galaxies, are widely used as a surrogate for spectroscopic redshifts, with considerable success. As discussed below, $\sim$30% of EMU sources will have multi-wavelength optical/IR photometric data at the time of data release, increasing to $\sim$70% in 2020. The SED of star-forming galaxies has a characteristic shape which provides good fits to SED modelling, and so it is expected that the majority of EMU star-forming galaxies (whose host galaxies will mainly be low-redshift and

hence optically bright) will have reliable photometric redshifts.

Similarly, low-power AGN (such as Seyfert galaxies) and AGNs hosted by elliptical galaxies at $z < 1$ are also likely to yield reliable photometric redshifts.

High-redshift radio-loud AGNS present more of a challenge. Photometric redshifts are notoriously bad at fitting the power-law SEDs of AGNs, and hence strong radio-selected sources. For example, Rowan-Robinson et al. (2008) found a 99% success rate in photometric redshifts of galaxies in the SWIRE survey, while Norris et al. (2006) found only a 50% success rate in Rowan-Robinson's fits for radio sources within that sample, and a success rate indistinguishable from chance for radio sources at $z > 1$. On the other hand, Salvato et al. (2009) have achieved reliable photometric redshifts of AGNs in the COSMOS field by using high-quality photometry in a large number of bands, and taking account of effects such as source variability.

While the relatively sparse photometry available for most EMU sources (listed in Table 3) cannot generate such accurate photometric redshifts for high-redshift sources, it can provide an estimate of the probability of a given source being at a particular redshift, for about 30% of EMU galaxies. Such estimates are valuable for statistical purposes such as cosmological tests (see §2.9), and we term such redshifts 'statistical redshifts'.

Even a non-detection can carry useful information, and radio data themselves can add significantly to the choice of SED template, and hence to a probabilistic estimate of redshift. For example, an AGN template may be indicated by the radio morphology, the detection of polarisation, a radio spectral index that is significantly different from -0.7, or a low far infrared–radio ratio. Alternatively, high-redshift radio galaxies can be identified from their strong radio emission coupled with a K-band non-detection (Willott et al. 2003). Even the radio data alone can provide information about the probable redshift range. For example, a steep radio spectral index increases the probability of a high redshift (Breuck et al. 2002), while the angular size of some classes of galaxy can be correlated with redshift (e.g. Wardle & Miley 1974).

Radio polarisation data can also provide statistical redshift information (Norris et al. 2011c). The POSSUM project (Gaensler et al. 2010) will run commensally with EMU to generate a catalogue of polarised fluxes and Faraday rotation measures (or upper limits) for all sources detected by EMU. Sources that are strongly polarised are nearly always AGNs (Hales et al. 2011), and so have a mean $z \sim 1.88$, while unpolarised sources are mainly star-forming galaxies with a mean $z \sim 1.08$ (see Fig. 5). Consequently, cosmological tests may be made by treating unpolarised sources as a low-redshift screen in front of background high-redshift polarised sources.

A significant fraction of statistical redshifts are likely to be incorrect. A correction for this can be applied in a statistical study of a sufficient number of objects, provided that reliability and completeness are carefully calibrated in a small well-studied area with accurate spectroscopic redshifts. The ATLAS survey (Norris et al. 2006; Middelberg et al. 2008a) is expected to yield 16000 radio sources at its completion in 2011 (Banfield et al. 2011). Spectroscopy on these sources is already underway (Mao et al. 2011b) and other large spectroscopy proposals have also been submitted to measure redshifts for the ATLAS radio sources. This spectroscopy should enable the training of algorithms to yield, and correct, statisticalredshifts for ~30% of EMU sources.

A further initiative is a proposed project to make use of the wide-field spectroscopic capabilities of the UK Schmidt Telescope at Siding Spring. The aim of this project, named TAIPAN (Transforming Astronomical Imaging-surveys through Polychromatic Analysis of Nebulae), is to extend the depth of the 6dFGS (Jones et al. 2009)to achieve a complete spectroscopic survey over the full southern hemisphere, potentially approaching the depths of the SDSS spectroscopic survey. This would provide invaluable spectroscopic data to complement all wide-field ASKAP surveys, and especially EMU.

3.12 Imaging the Galactic Plane

Although EMU focuses primarily on extra-Galactic science, Galactic science is also an important secondary goal, as discussed in §2.10 above. Imaging the Galactic plane, or even imaging extra-Galactic sky near the Galactic plane, introduces additional issues which must be tackled in the design study, such as:

- What are the limitations in the depth of EMU due to the bright extended sources in the Galactic plane?
- How can the sensitivity to larger-scale structures such as HII regions be maximised?
- How close to the Galactic plane can the full sensitivity and dynamic range be achieved for the main extragalactic EMU survey?

The design study attempts to answer these questions by conducting simulations of the standard SKADS simulatedsky to which Galactic continuum images have been added.

3.13 SPARCS: SKA PAthfinders Radio Continuum Survey working group

Several next-generation radio telescopes and upgrades are already being built around the world, including LOFAR (The Netherlands), ASKAP (Australia), Meerkat (South Africa), EVLA (USA), eMERLIN (UK), the Allen Telescope Array (USA), and Apertif (The Netherlands). Most of these projects have multiple science goals, but they have one goal in common, which is to survey the radio continuum emission from galaxies, in order to understand the formation and evolution of galaxies over cosmic time, and the cosmological parameters and large-scale structures whichdrive it. In pursuit of this goal, the different teams are developing techniques such as multi-scale deconvolution, source extraction and classification, and

multi-wavelength cross-identification. Furthermore, these projects share specific science goals, which in most cases require further definition before a well-planned survey can be executed. Finally, each of these new instruments has different strengths, and coordination of surveys between them can help maximise the scientific results from each of them.

The SKA PAthfinder Radio Continuum Survey (SPARCS) working group has therefore been established, under the auspices of the SKA project, with the following goals:

- To coordinate developments of techniques, to avoid duplication of effort and ensure that each project has access to best practice,
- To facilitate cross-project discussions of the specific science goals, to ensure cross-fertilisation of ideas and optimum survey strategies,
- To coordinate the surveys in their choice of area, depth, location on the sky, and other survey parameters, to maximise the scientific return from the surveys.

SPARCS consists of a core group of the leaders of the radio continuum surveys such as EMU, and is also open to the estimated 200 other astronomers who are engaged in the continuum survey projects on SKA Pathfinders. The first meeting of SPARCS took place in February 2011 at the Lorentz Center in Leiden, The Netherlands, and was attended by about 60 astronomers. The Proceedings will be published as a multi-author paper (Norris et al. 2011b) in a refereed journal. SPARCS will continue to hold such meetings annually, as well as facilitating other interactions between members of the various survey groups.

4 EMU Observing and Operations

It is expected that the observations will take place commensally with other ASKAP surveys such as WALLABY and POSSUM, implying that observing strategies and schedules will need to be agreed on by the different teams. The actual observing is expected to be conducted autonomously, with observing schedules delivered to the telescope in advance, and telescope performance monitored throughout the observations. Little or no interaction is expected to take place between the astronomers and the telescope until the data products are delivered from the pipeline for quality control.

Observations are expected to take place on one field at a time, rather than revisiting each field many times, to minimise the effect of source variability and to minimise the large processing overheads associated with combining several epochs of data. Typically, ASKAP will survey one field of 30 deg^2 for 12 hours before moving to the next, although this may be complicated by the need for dithering to achieve a uniform sensitivity across the field.

4.1 Data Quality Control

Data produced by the *ASKAPsoft* processing pipeline will be subject to a number of data quality checks before being released into the public domain, and this quality control process is the responsibility of the EMU team. In real time, the statistics of each field will be compared with those obtained from other radio surveys, previous ASKAP pointings, and the expected performance of ASKAP, including:

- rms noise level,
- source counts of detected sources (i.e. do they fit published logN/logS distributions?),
- spectral indices,
- the circularly polarised source components (expected to be close to zero),
- rotation measures of linearly polarised source components,
- measured positions of cross-identified components, compared with their positions from multi-wavelength surveys.

In the early days of the survey, these automated tests will be followed by a visual inspection of each image for artefacts or other problems. However, it is expected that as experience and confidence grows with time, the visual inspections will be replaced by automated tests.

A further level of testing will be provided on a longer timescale by the cross-identification program, which will, for example, show up any systematic positional errors. Additional quality control data will come from the WTF program (section 2.11) which will not only identify astronomically interesting objects which do not fall into known classes, but will also identify errors in the EMU data which may masquerade as a new class of objects.

The final quality control step is carried out by external users. As all data is released into the public domain in near-real time, feedback will be encouraged from external users of EMU data which will enable the identification of rare or subtle problems in the data.

4.2 Data Release

The primary data products consist of the *uv* data, image data, and a catalogue of radio components, as listed in Table 2. These data products are delivered to the ASKAP Science Data Archive Facility (ASDAF) by the real-time processing pipeline, and are then subjected to the quality control processes described above. After passing quality control, the data will then be released into the public domain, with no proprietary period. Once routine observing is established, the time taken for the quality control step is expected to be no more than about a day, although in the early days of the survey the data quality control step is likely to take longer.

Because the time taken to process the primary data is comparable to the observing time, routine reprocessing of data is not an option. It is therefore important that the primary data products are as final as possible, and it is likely that there will be only one data release for each field.

The secondary data products consist of a catalogue of sources, each of which may consist of one or more

components in the primary catalogue, together with cross-identifications with optical/IR data. This will also be produced by an automated pipeline which will run asynchronously from the primary data processing pipeline. The secondary data products will also be placed in the public domain as soon as possible after processing, although the need to protect proprietary optical/IR data with privileged access is also recognised.

It is likely that as additional survey data become available over time, and as expertise increases, the cross-identification process will also be refined. Several releases of the secondary catalogue are therefore expected, each with a release number embedded in the source identifiers.

Both the primary and secondary data will be available from the ASDAF using Virtual Observatory (VO) tools such as Simbad and TopCat, and through simple cone-search and cut-out servers. It is expected that they will also be available from the NASA Extragalactic Database (NED) and other astronomical databases.

4.3 Outreach and Communications

EMU provides an excellent opportunity to inform and engage the public about new developments in radio astronomy and ASKAP. An EMU website aimed at the public is currently under development. Initially it will be a fairly traditional 'static' site but there is great potential for it to grow and incorporate new technologies as data become available, once ASKAP is operational. Linkage with the Radio Zoo site (see §3.9) should expose it to a large and diverse audience.

The advent of a new 'Australian curriculum' for both Science (K–10) and Physics (Years 11–12) (`http:www.australiancurriculum.edu.au`) over the next few years is an opportunity to develop new, targeted formal education resources, including utilising EMU data for the 'science understanding' and 'science inquiry skills' strands of the curriculum but examples of scientists involved in EMU and ASKAP for the 'science as a human endeavour' strand for students across Australia.

Communication within the EMU team is primarily via a closed wiki, augmented by an electronic newsletter producedthree times per year.

5 Conclusion

The most obvious benefit from EMU is the legacy survey. Its greater depth, greater resolution, and program of cross-identifications will impact all areas of astronomy. EMU's sensitivity means that the majority of objects in EMU will be star-forming galaxies rather than radio-loud AGNs. Thus EMU plays a key role in the continuing emergence of extra-Galactic radio astronomy from the niche that it once occupied, focussing largely on radio-loud AGNs, to one that impacts on all areas of astronomy. As a legacy survey, EMU will provide a resource to all astronomers, leveraging astronomical observations at all wavelengths

both through the VO server and also through data centres like NED and CDS. It is likely to find rare, interesting targets (such as the ~100 EoR radio galaxies) that can then be followed up by other observatories such as ALMA, ELT, JWST, and Meerkat, which are better suited to deep pointed observations.

In addition to the legacy program, this paper has described a number of science areas in which EMU is likely to have a major impact. For example, it will distinguish between alternative models of cosmology and general relativity, trace the evolution of star formation and AGNs, and significantly increase the number of known clusters of galaxies.

It is even possible that the greatest impact of EMU will be on science that hasn't yet been imagined. EMU will occupy a new region of observational phase space, and we plan to exploit this by mining the data for unexpected types of astronomical object.

EMU is novel in that cross-identifications with all available data at optical, infrared, and radio wavelengths will be produced by the EMU cross-identification pipeline, and will be placed in the public domain, as a service to the community, as soon as the data quality has been checked. They will be made available using the VO server (and thus immediately via services such as Aladin and TopCat), through the EMU web page, and also through ASDAF services.

EMU is an open collaboration and we welcome interactions and collaborations with other surveys and other instruments, to increase our common scientific productivity. We also encourage other individual scientists to join the EMU collaboration, to help generate even better science, by contacting Ray Norris. Further information on EMU can be found on `http:askap.org/emu`

Wajarri Yamatji people as the traditional owners of the Observatory site.

References

Abazajian, K. N. et al., 2009, ApJ Suppl., 182, 543

Afonso, J., Georgakakis, A., Almeida, C., Hopkins, A. M., Cram, L. E., Mobasher, B. & Sullivan, M., 2005, ApJ, 624, 135

Afonso, J., Mobasher, B., Koekemoer, A., Norris, R. P. & Cram, L., 2006, AJ, 131, 1216

Allen, S. W., Rapetti, D. A., Schmidt, R. W., Ebeling, H., Morris, R. G. & Fabian, A. C., 2008, MNRAS, 383, 879

Bagchi, J., Enßlin, T. A., Miniati, F., Stalin, C. S., Singh, M., Raychaudhury, S. & Humeshkar, N. B., 2002, New Astr., 7, 249

Ball, L. et al., 2009, ASKAP memo 26, http://www.atnf.csiro.au/SKA/ASKAP_SSP_Sep09_final.pdf

Ballard, D. H., 1981, Pattern Recognition, 13.2, 111

Banfield, J. K., et al., 2011, in preparation

Barrena, R., Girardi, M., Boschin, W. & Dasí, M., 2009, Astr. Ap., 503, 357

Battaglia, N., Pfrommer, C., Sievers, J. L., Bond, J. R. & Enßlin, T. A., 2009, MNRAS, 393, 1073

Becker, R. H., White, R. L. & Helfand, D. J., 1995, ApJ, 450, 559

Becker, W., 2009, Astrophysics and Space Science Library, 357, 91

Bell, E. F., 2003, ApJ, 586, 794

Bell-Burnell, J. 2009, in 'Accelerating The Rate of Astronomical Discovery', edited by Ray P. Norris & Clive L. N. Ruggles, 2009, PoS (SpS5)

Berger, E. et al., 2001, Natur, 410, 338

Bertacca, D., Raccanelli, A., Piattella, O. F, Pietrobon, D., Bartolo, N., Matarrese, S. & Giannantonio, T., 2011, J. Cosm. Astro-particle Phys., 3, 39

Bertin, E. & Arnouts, S., 1996, A&AS, 117, 393

Best, P. N., Arts, J. N., Röttgering, H. J. A., Rengelink, R., Brookes, M. H. & Wall, J., 2003, MNRAS, 346, 627

Best, P. N., Kaiser, C. R., Heckman, T. M. & Kauffmann, G., 2006, MNRAS, 368, L67

Béthermin, M., Dole, H., Beelen, A. & Aussel, H., 2010, Astr. Ap., 512, A78

Biggs, A. D. & Ivison, R. J., 2006, MNRAS, 371, 963

Biggs, A. D. & Ivison, R. J., 2008, MNRAS, 385, 893

Biggs, A. D., Younger, J. D. & Ivison, R. J., 2010, MNRAS, 408, 342

Blake, C. & Wall, J., 2002, MNRAS, 329, L37

Blanton, E. L. et al., 2003, AJ, 125, 1635

Bondi, M. et al., 2003, A&A, 403, 857

Bongiorno, A. et al., 2007, A&A, 472, 443

Boughn, S. & Crittenden, R., 2004, Natur, 427, 45

Bourne, N., Dunne, L., Ivison, R. J., Maddox, S. J., Dickinson, M. & Frayer, D. T., 2011, MNRAS, 410, 1155

Bower, R. G., Benson, A. J., Malbon, R., Helly, J. C., Frenk, C. S., Baugh, C. M., Cole, S. & Lacey, C. G., 2006, MNRAS, 370, 645

Boyle, B. J. et al., 2007, MNRAS, 376, 1182

Brax, P. & van de Bruck, C., 2003, Class.Quant.Grav., 20, R201 (also arXiv:hep-th/0303095)

Bridle, A. H. & Schwab, F. R., 1999, Synthesis Imaging in Radio Astronomy II, 180, 371

Briggs, D., & Cornwell, T. J., 1992, EVLA Memo 114

Brown, S. & Rudnick, L., 2009, AJ, 137, 3158

Brown, M. J. I. et al., 2008, ApJ, 682, 937

Brown, S., Farnsworth, D. & Rudnick, L., 2010, MNRAS, 402, 2

Brown, S., Duesterhoeft, J. & Rudnick, L., 2011, ApJ (Letters), 727, L25

Brown, S. et al., 2011, in preparation.

Brunetti, G., Cassano, R., Dolag, K. & Setti, G., 2009, Astr. Ap., 507, 661

Budding, E., Jones, K. L., Slee, O. B. & Watson, L., 1999, MNRAS, 305, 966

Bunton, J. D., & Hay, S. G., 2010, in International Conference on Electromagnetics in Advanced Applications (ICEAA), http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5651120

Cameron, A. et al., 2011, MNRAS, in press

Carey, S. J. et al., 2009, PASP, 121, 76

Carilli, C., & Rawlings, S., 2004, New Astronomy Reviews, Vol. 48, Elsevier.

Carilli, C. L., Gnedin, N. Y. & Owen, F., 2002, ApJ, 577, 22

Cassano, R., Ettori, S., Giacintucci, S., Brunetti, G., Markevitch, M., Venturi, T. & Gitti, M., 2010, ApJ (Letters), 721, L82

Cen, R. & Ostriker, J. P., 1999, ApJ, 514, 1

Chapman, S. C., Blain, A. W., Ivison, R. J. & Smail, I. R., 2003, Natur, 422, 695

Chapin, E. L. et al., 2011, MNRAS, 411, 505

Chary, R.-R., Cooray, A. & Sullivan, I., 2008, ApJ, 681, 53

Chatterjee, S. & Murphy, T., VAST Collaboration 2010, BAAS, 42, #470.12

Churchwell, E., 2002, ArAA, 40, 27

Cole, S. et al., 2005, MNRAS, 362, 505

Compiègne, M., Flagey, N., Noriega-Crespo, A., Martin, P. G., Bernard, J.-P., Paladini, R. & Molinari, S., 2010, ApJ (Letters), 724, L44

Condon, J. J., 1974, ApJ, 188, 279

Condon, J. J. 1988, in 'Galactic and Extragalactic Radio Astronomy', 2nd edn, eds. G.L. Verschuur and K.I. Kellermann. Also in http://nedwww.ipac.caltech.edu/level5/Sept04/Condon/Condon_contents.html

Condon, J. J., (1992), ARA&A, 30, 575

Condon, J. J., (2008), ASKAP memo 015, www.atnf.csiro.au/SKA/newdocs/condon_memo.pdf

Condon, J. J., Cotton, W. D., Greisen, E. W., Yin, Q. F., Perley, R. A., Taylor, G. B. & Broderick, J. J., 1998, AJ, 115, 1693

Condon, J. J. & Kaplan, D. L., 1998, ApJ Suppl., 117, 361

Condon, J. J., Kaplan, D. L. & Terzian, Y., 1999, ApJ Suppl., 123, 219

Conti, P. S. & Crowther, P. A., 2004, MNRAS, 355, 899

Cornwell, T. J., 2008, ISTSP, 2, 793

Cornwell, T. J., 2010, pers. comm.

Cornwell, T. J. et al., 2011, ASKAP Memo ASKAP-SW-0020

Cotton, W. D. et al., 2003, PASA, 20, 12

Cowie, L. L. et al., 1996, AJ, 112, 839

Cram, L., Hopkins, A., Mobasher, B. & Rowan-Robinson, M., 1998, ApJ, 507, 155

Crittenden, R. & Turok, N., 1996, Phys. Rev.Lett., 76, 575

Croom, S. M. et al., 2004, MNRAS, 349, 1397

Croton, D. J. et al., 2005, MNRAS, 356, 1155

Croton, D. J. et al., 2006, MNRAS, 365, 11

Cowie, L. L. & Binney, J., 1977, ApJ, 215, 723

Cruz, M. J. et al., 2006, MNRAS, 373, 1531

Cruz, M. J. et al., 2007, MNRAS, 375, 1349

Daddi, E. et al., 2007, ApJ, 670, 173

Dalal, N., Doré, O., Huterer, D. & Shirokov, A., 2008, Phys. Rev D., 77, 123514

The Dark Energy Survey Collaboration 2005, arXiv:astro-ph/0510346

Deboer, D. R. et al., 2009, IEEE Proceedings, 97, 1507

De Breuck, C. et al., 2001, AJ, 121, 1241

De Breuck, C. et al., 2002, AJ, 123, 637

Dehghan, S. et al., 2011, JApA, in press

Dewdney, P. et al., 2009, IEEE Proceedings, 97, 1482

Dickey, J. M., Gibson, S. J., Gomez, J. F., Imai, H., Jones, P. A., McClure-Griffiths, N. M., Stanimirovic, S. & van Loon, J. T., 2010, arXiv:1008.4640

Djorgovski, S. G., Mahabal, A., Donalek, C., Graham, M., Moghaddam, B., Drake, A. & Williams, R., 2010, BAAS, 42, #231.05

Dolag, K. et al., 2008, Sp. Sci. Rev., 134, 311

Dougherty, S. M. & Williams, P. M., 2000, MNRAS, 319, 143

Drake, S. A., Simon, T. & Linsky, J. L., 1989, ApJ Suppl., 71, 90

Driver, S. P. et al., 2009, A&G, 50, 12

Dulk, G. A., 1985, ARAA, 23, 169

Dunne, L. et al., 2009, MNRAS, 394, 3

Eddington, A. S., 1913, MNRAS, 73, 359

Edge, D. O., Shakeshaft, J. R., McAdam, W. B., Baldwin, J. E. & Archer, S., 1959, Mem. RAS, 68, 37

Eisenstein, D. J. et al., 2005, ApJ, 633, 560

Ekers, R. D., 2010 in 'Accelerating The Rate of Astronomical Discovery', edited by Ray P. Norris & Clive L. N. Ruggles, 2010, PoS (SpS5)

Elbaz, D. et al., 2007, Astr. Ap., 468, 33

Elbaz, D., Jahnke, K., Pantin, E., Le Borgne, D. & Letawe, G., 2009, Astr. Ap., 507, 1359

Evans, D. A., Lee, J. C., Kamenetska, M., Gallagher, S. C., Kraft, R. P., Hardcastle, M. J. & Weaver, K. A., 2006, ApJ, 653, 1121

Fabian, A. C. & Nulsen, P. E. J., 1977, MNRAS, 180, 479

Fanaroff, B. L. & Riley, J. M., 1974, MNRAS, 167, 31

Fanti, C., 2009a, Astron. Nachr., 330, 120

Fanti, R., 2009b, Astron. Nachr., 330, 303

Feretti, L., 2000, Invited review at IAU 199, arXiv:astro-ph/0006379

Ferrarese, L. & Merritt, D., 2000, ApJ, 539, L9

Feulner, G. et al., 2005, ApJ, 633, L9

Fraser-McKelvie, A., Pimbblet, K. A. & Lazendic, J. S. MNRAS, in press (arXiv:1104.0711)

Fujita, Y., Kohri, K., Yamazaki, R. & Kino, M., 2007, ApJ (Letters), 663, L61

Gaensler, B. M., Landecker, T. L. & Taylor, A. R., POSSUM Collaboration 2010, BAAS, 42, 515

Garn, T. & Alexander, P., 2008, MNRAS, 391, 1000

Garn, T., et al., 2009, MNRAS, (arXiv:0905.1218)

Gebhardt, K. et al., 2000, ApJ, 539, L13

Giannantonio, T., Scranton, R., Crittenden, R. G., Nichol, R. C., Boughn, S. P., Myers, A. D. & Richards, G. T., 2008a, 2008PhRvD, 77, l3520G (arXiv:0801.4380v2)

Gil de Paz, A. et al., 2007, ApJ Suppl., 173, 185

Giovannini, G. & Feretti, L., 2004, J. Korean Ast. Soc., 37, 323

Giovannini, G., Tordi, M. & Feretti, L., 1999, New Astr., 4, 141

Giovannini, G. et al., 2009, Astr. Ap., 507, 1257

Giovannini, G. et al., 2010, Astr. Ap., 511, L5

Giovannini, G. et al., 2011, JApA, in press

Gladders, M. D. & Yee, H. K. C., 2005, ApJ Suppl., 157, 1

Gobat, R. et al., 2011, Astr. Ap., 526, A133

Gold, B. et al., 2011, ApJ Suppl., 192, 15

Gomez, P. L. et al., 2003, ApJ, 584, 210

Gopal-Krishna, & Wiita, P. J., 2001, ApJ, 560, L115

Govoni, F., Feretti, L., Giovannini, G., Böhringer, H., Reiprich, T. H. & Murgia, M., 2001, Astr. Ap., 376, 803

Granett, B. R., Neyrinck, M. C. & Szapudi, I., 2008, ApJ (Letters), 683, L99

Green, D. A., 2009, Bulletin of the Astronomical Society of India, 37, 45

Greve, T. R. et al., 2010, ApJ, 719, 483

Güdel, M., ARA&A, 40, 217

Gupta, N., Salter, C. J., Saikia, D. J., Ghosh, T. & Jeyakumar, S., 2006, MNRAS, 373, 972

Gupta, N., et al., 2008, 'The Initial Array Configuration for ASKAP', ASKAP Internal Report.

Haarsma, D. B., Partridge, R. B., Windhorst, R. A. & Richards, E. A., 2000, ApJ, 544, 641

Hales, C. A., et al., 2011, in prep.

Hallinan, G., Antonova, A., Doyle, J. G., Bourke, S., Lane, C. & Golden, A., 2008, ApJ, 684, 644

Hancock, P., Murphy, T. & Hopkins, A., 2011, in preparation.

Hardcastle, M. J., Evans, D. A. & Croston, J. H., 2006, MNRAS, 370, 1893

Hardcastle, M. J. et al., 2007, MNRAS, 376, 1849

Hasinger, G. et al., 2005, Astr. Ap., 441, 417

Haverkorn, M., Gaensler, B. M., McClure-Griffiths, N. M., Dickey, J. M. & Green, A. J., 2006, ApJ Suppl., 167, 230

Helfand, D. J., Velusamy, T., Becker, R. H. & Lockman, F. J., 1989, ApJ, 341, 151

Helfand, D. J., Schnee, S., Becker, R. H., White, R. L. & McMahon, R. G., AJ, 117, 1568

Helfand, D. J., Becker, R. H., White, R. L., Fallon, A. & Tuttle, S., 2006, AJ, 131, 2525

Herbert et al., 2010, MNRAS, 406, 1841

Hine, R. G. & Longair, M. S., 1979, MNRAS, 188, 111

Hoeft, M. et al., 2011, JApA, in press

Hogg, D. W. & Turner, E. L., 1998, PASP, 110, 727

Hopkins, A. M. & Beacom, J. F., 2006, ApJ, 651, 142

Hopkins, A. M. et al., 2002, AJ, 123, 1086

Hopkins, A. M. et al., 2003, AJ, 125, 465

Hopkins, A. M., 2004, ApJ, 615, 209

Huynh, M. et al., 2005, AJ, 130, 1373

Huynh, M. T., Norris, R. P. & Middelberg, M., 2010, ApJ, 710, 698

Ibar, E., Ivison, R. J., Biggs, A. D., Lal, D. V., Best, P. N. & Green, D. A., 2009, MNRAS, 397, 281

Ibar, E., Ivison, R. J., Best, P. N., Coppin, K., Pope, A., Smail, I. & Dunlop, J. S., 2010, MNRAS, 401, L53

Ivezic, Z., Tyson, J. A., Allsman, R., Andrew, J., Angel, R., & for the LSST Collaboration 2008, arXiv:0805.2366

Ivison, R. J. et al., 2007, ApJ, 660, L77

Ivison, R. J. et al., 2010, MNRAS, 402, 245

Jackson, N. & Rawlings, S., 1997, MNRAS, 286, 241

Jackson, C., 2005, PASA, 22, 36

Jarvis, M. J. et al., 2001a, MNRAS, 326, 1563

Jarvis, M. J. et al., 2001b, MNRAS, 326, 1585

Jarvis, M. J. et al., 2001c, MNRAS, 327, 907

Jarvis, M. J. & Rawlings, S., 2004, NewAR, 48, 1173

Jarvis, M. J., Teimourian, H., Simpson, C., Smith, D. J. B., Rawlings, S. & Bonfield, D., 2009, MNRAS, 398, L83

Johnston, S. et al., 2007, PASA, 24, 174

Johnston, S. et al., 2008, ExA, 22, 151

Johnston-Hollitt, M. et al., 2009, MNRAS, (submitted)

Jonas, J. L, 2009, IEEE Proceedings, 97, 1522

Jones, D. H. et al., 2009, MNRAS, 399, 683

Juneau, S. et al., 2005, ApJ, 619, L135

Kaiser, N. et al., 2010, Proc. SPIE 7733, 0E

Keller, S. C. et al., 2007, PASA, 24, 1

Kempner, J. C., Blanton, E. L., Clarke, T. E., Enßlin, T. A., Johnston-Hollitt, M. & Rudnick, L., 2004, The Riddle of Cooling Flows in Galaxies and Clusters of galaxies, 335, www.astro.virginia.edu/coolflow

Keshet, U., Waxman, E. & Loeb, A., 2004, ApJ, 617, 281

Kim, K.-T., Kronberg, P. P., Giovannini, G. & Venturi, T., 1989, Natur, 341, 720

Kimball, A. E., Knapp, G. R., Ivezić, Ž., West, A. A., Bochanski, J. J., Plotkin, R. M. & Gordon, M. S., 2009, ApJ, 701, 535

Klamer, I. J., Ekers, R. D., Sadler, E. M. & Hunstead, R. W., 2004, ApJ, 612, L97

Kodama, T. et al., 2007, MNRAS, 377, 1717

Koribalski, B. S., et al., 2008, in 'Galaxies in the Local Volume', Sydney, July 2007, Springer, 41

Koribalski, B. S. et al., 2011, PASA, in preparation.

Kronberg, P. P., Kothes, R., Salter, C. J. & Perillat, P., 2007, ApJ, 659, 267

Kurtz, S., Churchwell, E. & Wood, D. O. S., 1994, ApJ Suppl., 91, 659

Kurtz, S. E., Watson, A. M., Hofner, P. & Otte, B., 1999, ApJ, 514, 232

Kurtz, S., 2005, in 'Massive Star Birth: A Crossroads of Astrophysics', IAU Symp. 227, 111, Cambridge: Cambridge University Press

Labiano, A., O'Dea, C. P., Barthel, P. D., de Vries, W. H. & Baum, S. A., 2008, Astr. Ap., 477, 491

Laing, R. A., Jenkins, C. R., Wall, J. V. & Unger, S. W., 1994, The Physics of Active Galaxies, 54, 201

Large, M. I., Mills, B. Y., Little, A. G., Crawford, D. F. & Sutton, J. M., 1981, MNRAS, 194, 693

Lewis, I. et al., 2002, MNRAS, 334, 673

Lilly, S. J. et al., 1996, ApJ, 460, L1

Lintott, C. J. et al., 2008, MNRAS, 389, 1179

Liu, F. et al., 2008, Proc. SPIE, Vol. 7017, 70170M

Longmore, S. N., Burton, M. G., Keto, E., Kurtz, S. & Walsh, A. J., 2009, MNRAS, 399, 861

Madau, P. et al., (1996), MNRAS, 283, 1388

Magorrian, J. et al., (1998), AJ, 115, 2285

Mantz, A., Allen, S. W., Rapetti, D. & Ebeling, H., 2010, MNRAS, 406, 1759

Mao, M. et al., 2009, MNRAS, 392, 1070

Mao, M. et al., 2010, MNRAS, 406, 2578

Mao, M. Y., Huynh, M. T., Norris, R. P., Dickinson, M., Frayer, D., Helou, G. & Monkiewicz, J. A., 2011, ApJ, 731, 79

Mao, M., et al., 2011b, in preparation.

Mao, M., et al., 2011c, JApA, in press

Markevitch, M., 2011, arXiv:1010.3660

Marriage, T. A., et al., 2010, arXiv:1010.1065

Massardi, M. et al., 2010, MNRAS, 404, 532

Mauch, T. et al., 2003, MNRAS, 342, 1117

Mauch, T. & Sadler, E. M., 2007, MNRAS, 375, 931

McClure-Griffiths, N. M., Dickey, J. M., Gaensler, B. M., Green, A. J., Haverkorn, M. & Strasser, S., 2005, ApJ Suppl., 158, 178

McLure, R. J., Jarvis, M. J., Targett, T. A., Dunlop, J. S. & Best, P. N., 2006, MNRAS, 368, 1395

McNamara, B. R. & Nulsen, P. E. J., 2007, Ann. Rev. Astr. Ap., 45, 117

Messias, H., Afonso, J., Hopkins, A., Mobasher, B., Dominici, T. & Alexander, D. M., 2010, ApJ, 719, 790

Mezger, P. G. & Henderson, A. P., 1967, ApJ, 147, 471

Middelberg, E. et al., 2008a, AJ, 135, 1276

Middelberg, E., Norris, R. P., Tingay, S., Mao, M. Y., Phillips, C. J. & Hotan, A. W., 2008b, Astr. Ap., 491, 435

Middelberg, E., Norris, R. P., Seymour, N., Huynh, M., Johnston-Hollitt, M., Mao, M. Y., Hales, C. A. & Hotan,, 2010, Astr. Ap., in press

Mignano, A. et al., 2008, Astr. Ap., 477, 459

Miley, G. K. et al., 2006, ApJ, 650, L29

Miller, N. A., Fomalont, E. B., Kellermann, K. I., Mainieri, V., Norman, C., Padovani, P., Rosati, P. & Tozzi, P., 2008, ApJ Suppl., 179, 114

Miniati, F. et al., 2001, ApJ, 562, 233

Mitchell, K. J. & Condon, J. J., 1985, AJ, 90, 1957

Mittal, R., Hudson, D. S., Reiprich, T. H. & Clarke, T., 2009, Astr. Ap., 501, 835

Mobasher, B. et al., 2009, ApJ, 690, 1074

Molinari, S. et al., 2010, PASP, 122, 314

Morganti, R., 2008, in 'Extragalactic Jets: Theory and Observation from Radio to Gamma Ray', edited by Rector, T. A. & De Young, D. S., 210

Morganti, R., Emonts, B., Holt, J., Tadhunter, C., Oosterloo, T. & Struve, C., 2009, Astron. Nachr., 330, 789

Morrison, G. E. et al., 2010, ApJ Suppl., 188, 178

Murgia, M., Fanti, C., Fanti, R., Gregorini, L., Klein, U., Mack, K.-H. & Vigotti, M., 1999, Astr. Ap., 345, 769

Murphy, T., Mauch, T., Green, A., Hunstead, R. W., Piestrzynska, B., Kels, A. P. & Sztajer, P., 2007, MNRAS, 382, 382

Norris, R. P. et al., 2005, AJ, 130, 1358

Norris, R. P. et al., 2006, AJ, 132, 2409

Norris, R. P., Tingay, S., Phillips, C., Middelberg, E., Deller, A. & Appleton, P. N., 2007, MNRAS, 378, 1434

Norris, R. P., et al., 2009 in 'Proceedings of Panoramic Radio Astronomy'. Ed George Heald. PoS (PRA2009)033

Norris, R. P., 2010, in 'Accelerating The Rate of Astronomical Discovery', edited by Ray P. Norris & Clive L. N. Ruggles, 2010, PoS (SpS5)

Norris, R. P., et al., 2011a, ApJ, in press. arXiv1105.0960N

Norris, R. P., et al., 2011b, PASA, in preparation.

Norris, R. P. et al., 2011c, JApA, in press

O'Dea, C., 1998, PASP, 110, 493

Ogle, P., Whysong, D. & Antonucci, R., 2006, ApJ, 647, 161

Ogle, P., Boulanger, F., Guillard, P., Evans, D. A., Antonucci, R., Appleton, P. N. & Nesvadba, N., Leipski2010, ApJ, 724, 1193

Oliver, S. J. et al., 2010, Astr. Ap., 518, L21

Oosterloo, T., Verheijen, M., van Cappellen, W., Bakker, L., Heald, G. &, Ivashina, M., 2009, arXiv:0912.0093

Overzier, R. A., Röttgering, H. J. A., Rengelink, R. B. & Wilman, R. J., 2003, Astr. Ap., 405, 53

Owen, F. N. & Morison, G. E., 2008, AJ, 136, 1889

Owsianik, I., Conway, J. E. & Polatidis, A. G., 1999, NewAR, 43, 669

Padovani, P. et al., 2009, ApJ, 694, 235

Papovich, C. et al., 2006, ApJ, 640, 92

Peacock, J. A. & Smith, R. E., 2000, MNRAS, 318, 1144

Pen, U.-L., Staveley-Smith, L., Peterson, J. B. & Chang, T.-C., 2009, MNRAS, 394, L6

Penner, K. et al., 2011, MNRAS, 410, 2749

Perlmutter, S. et al., (1999), Astrophys. J., 517, 565

Peroux, C. et al., 2005, MNRAS, 363, 479

Peterson, J. R. et al., 2001, Astr. Ap., 365, L104

Peterson, J. R., Kahn, S. M., Paerels, F. B. S., Kaastra, J. S., Tamura, T., Bleeker, J. A. M., Ferrigno, C. & Jernigan, J. G., 2003, ApJ, 590, 207

Peterson, J. R. & Fabian, A. C., 2006, Phys. Rep., 427, 1

Pfrommer, C. et al., 2007, MNRAS, 378, 385

Pierre, M. et al., 2004, J. Cosm. & Astroparticle Phys., 9, 11

Pietrobon, D. et al., 2006, Phys. Rev. D, 74, 043524

Pillepich, A., Porciani, C. & Reiprich. T. H., 2011, in preparation.

Planck Collaboration 2011, arXiv:1101.2024

Pogosian, L. et al., 2010, Phys. Rev. D, 81, 104023

Prandoni, I., Gregorini, L., Parma, P., de Ruiter, H. R., Vettolani, G., Wieringa, M. H. & Ekers, R. D., 2001, Astr. Ap., 365, 392

Predehl, P. et al., 2010, Proc. SPIE, 7732,

Purcell, C. R., Hoare, M. G. & Diamond, P., 2008, Massive Star Formation: Observations Confront Theory, 387, 389

Purcell, C. R. & Hoare, M. G., 2010, Highlights of Astronomy, 15, 781

Raccanelli, A., Bonaldi, A., Negrello, M., Matarrese, S., Tormen G. & de Zotti, G., 2008, MNRAS, 386, 2161

Raccanelli, A., et al., 2011, in preparation

Randall, K. E. et al., 2011, MNRAS, , in press.

Ravi, V. et al., 2010, MNRAS, 408, L99

Rawlings, S. & Jarvis, M. J., 2004, MNRAS, 355, L9

Reich, W., Fürst, E. & Arnal, E. M., 1992, Astr. Ap., 256, 214

Riess, A. G. et al., 1998, AJ, 116, 1009

Romer, A. K. et al., 2001, ApJ, 547, 594

Rosati, P. et al., 1998, ApJ (Letters), 492, L21

Roseboom, I. G., Oliver, S., Parkinson, D. & Vaccari, M., 2009, MNRAS, 400, 1062

Roseboom, I. G. et al., 2010, MNRAS, 409, 48

Röttgering, H. J. A., Wieringa, M. H., Hunstead, R. W. & Ekers, R. D., 1997, MNRAS, 290, 577

Röttgering, H. et al., 2010a, PoS (ISKAF2010)050

Röttgering, H. et al., 2010b, www.astron.nl/radio-observatory/apertif-eoi-abstracts-and-con-tact-information

Rowan-Robinson, M. et al., (2008), MNRAS, 386, 697

Rudnick, L., 2002, PASP, 114, 427

Rudnick, L. & Lemmerman, J. A., 2009, ApJ, 697, 1341

Ryu, D. et al., 2008, Science, 320, 909

Saikia, D. J., Jeyakumar, S., Mantovani, F., Salter, C. J., Spencer, R. E., Thomasson, P. & Wiita, P. J., 2003, PASA, 20, 50

Saikia, D. J. & Jamrozy, M., 2009, Bulletin of the Astronomical Society of India,, 37, 63

Salvato, M. et al., 2009, ApJ, 690, 1250

Santos, J. S., Tozzi, P., Rosati, P. & Böhringer, H., 2010, Astr. Ap., 521, A64

Saripalli, L., et al., 2011, in preparation.

Schmidt, M., 1968, ApJ, 151, 393

Schinnerer, E. et al., 2007, ApJS, 172, 46

Schuecker, P., Böhringer, H., Reiprich, T. H. & Feretti, L., 2001, Astr. Ap., 378, 408

Scoville, N., 2007, ApJS, 172, 1

Scranton, R. et al., 2005, ApJ, 633, 589

Seaquist, E. R., Krogulec, M. & Taylor, A. R., 1993, ApJ, 410, 260

Seaquist, E. R. & Ivison, R. J., 1994, MNRAS, 269, 512

Seymour, N. et al., 2008, MNRAS, 386, 1695

Seymour, N. et al., 2011, MNRAS, 280,

Shanks, T., 2005, VST-ATLAS proposal, www.astro.dur.ac.uk/Cosmology/vstatlas/

Shaver, P. A., Wall, J. V., Kellermann, K. I., Jackson, C. A. & Hawkins, M. R. S., 1996, Natur, 384, 439

Simpson, C. et al., 2006, MNRAS, 372, 741

Skillman, S. W. et al., 2008, ApJ, 689, 1063

Skillman, S. W., Hallman, E. J., O'Shea, B. W., Burns, J. O., Smith, B. D. &, Turk, M. J., 2010, arXiv:1006.3559

Slee, O. B., Wilson, W. & Ramsay, G., 2008, PASA, 25, 94

Slee, O. B., Nelson, G. J., Stewart, R. T., Wright, A. E., Innis, J. L., Ryan, S. G. & Vaughan, A. E., 1987, MNRAS, 229, 659

Smith, D. J. B., Dunne, Maddox, 2010, arxiv:1007.5260

Smolcic, V. et al., 2008, ApJS, 177, 14

Snellen, I. A. G., Schilizzi, R. T., de Bruyn, A. G., Miley, G. K., Rengelink, R. B., Röttgering, H. J. A. & Bremer, M. N., 1998, Astr. Ap. Suppl., 131, 435

Snellen, I. A. G., Schilizzi, R. T., Miley, G. K., Bremer, M. N., Röttgering, H. J. A. & van Langevelde, H. J., 1999, NewAR, 43, 675

Snellen, I. A. G. et al., 2009, Astron. Nachr., 330, 297

Springel, V. et al., 2005, New Astr., 435, 629

Stil, J. M. et al., 2006, AJ, 132, 1158

Subrahmanyan, R., Ekers, R. D, Saripalli, L. & Sadler, E. M, 2010, MNRAS, 402, 2792

Sutherland, W. & Saunders, W., 1992, MNRAS, 259, 413

Sutherland, W., 2009, in Science with the VLT in the ELT Era, Astrophysics and Space Science Proceedings, 171

Taylor, A. R. et al., 2003, AJ, 125, 3145

Tegmark, M. et al., (2004), Phys. Rev. D, 69, 103501

Thompson, M. A., Hatchell, J., Walsh, A. J., MacDonald, G. H. & Millar, T. J., 2006, Astr. Ap., 453, 1003

Thompson, M. A., et al., in preparation

Trigilio, C., Leto, P., Umana, G., Buemi, C. S. & Leone, F., 2008, MNRAS, 384, 1437

Trigilio, C., Leto, P., Leone, F., Umana, G. & Buemi, C., 2000, Astr. Ap., 362, 281

Tschager, W., Schilizzi, R. T., Röttgering, H. J. A., Snellen, I. A. G., Miley, G. K. & Perley, R., 2003, PASA, 20, 75

Umana, G., Trigilio, C. & Catalano, S., 1998, Astr. Ap., 329, 1010

Umana, G., Trigilio, C., Tumino, M. & Catalano, S., Rodonó, M. 1995, Astr. Ap., 298, 143

Umana, G., Trigilio, C., Hjellming, R. M. & Catalano, S., Rodonó, M. 1993, Astr. Ap., 267, 126

Urquhart, J. S. et al., 2009, Astr. Ap., 501, 539

Uzan, J. P., 2003, Rev. Mod. Phys., 75, 403

vanBreugel, W. et al., 1999, ApJ, 518, L61

Van der Heyden, K., & Jarvis, M. J., 2010, MIGHTEE proposal to Meerkat

vanWeeren, R. J., Röttgering, H. J. A., Brüggen, M. & Hoeft, M., 2010, Science, 310, 347

Venturi, T. et al., 2008, Astr. Ap., 484, 327

Vikhlinin, A. et al., 2009, ApJ, 692, 1060

Wall, J. V. et al., 2005, Astr. Ap., 434, 133

Walsh, A. J., Burton, M. G., Hyland, A. R. & Robinson, G., 1998, MNRAS, 301, 640

Wardle, J. F. C. & Miley, G. K., 1974, Astr. Ap., 30, 305

White, R. L. et al., 1997, ApJ, 475, 479

White, R. L., Becker, R. H. & Helfand, D. J., 2005, AJ, 130, 586

White, S. M., 2004, NewAR, 48, 1319

White, R. L., Helfand, D. J. & Becker, R. H. et al., 2007, ApJ, 654, 99

Whiting, M. T., 2008; 'Galaxies in the Local Volume', ed. B.S. Koribalski & H. Jerjen, Astrophysics and Space Science Reviews, Springer, 343

Whysong, D. & Antonucci, R., 2004, ApJ, 602, 116

Wiaux, Y., Jacques, L., Puy, G., Scaife, A. M. M. & Vandergheynst, P., 2009, MNRAS, 395, 1733

Wilman, R. J. et al., 2008, MNRAS, 388, 1335

Wilman, R. J., Jarvis, M. J., Mauch, T., Rawlings, S. & Hickey, S., 2010, MNRAS, 405, 447

Williamson, R., et al., 2011, arXiv:1101.1290

Willott, C. J. et al., 2003, MNRAS, 339, 173

Wilson, G. et al., 2008, Infrared Diagnostics of Galaxy Evolution, 381, 210

Wing, J. D. & Blanton, E. L., 2011, AJ, 141, 88

Wood, D. O. S. & Churchwell, E., 1989, ApJS, 69, 831

Wright, E. L. et al., 2010, AJ, 140, 1868

Wyithe, J. S. B. & Loeb, A., 2009, MNRAS, 397, 1926

Xia, J.-Q., Viel, M., Baccigalupi, C. & Matarrese, S., 2009, J. Cosm. Astroparticle Phys., 09, 003

York, D. G. et al., 2000, AJ, 120, 1579

Zhao, G.-B., Giannantonio, T., Pogosian, L., Silvestri, A., Bacon, D. J., Koyama, K., Nichol, R. C. & Song, Y.-S., 2010, Phys. Rev. D, 81, 103510

Zheng, X. Z. et al., 2007, ApJ, 661, L41

Zlatev, I., Wang, L. & Steinhardt, P. J., 1999, PhRvL, 82, 896