Legacy e-MERLIN Multi-band Imaging of a complete Nearby Galaxy Sample (LeMMINGS)

Coordinators:
R. J. Beswick (`Robert.Beswick@manchester.ac.uk`) & I. M$^{\rm c}$Hardy(`imh@astro.soton.ac.uk`)
Addresses:
R. J. Beswick, Jodrell Bank Centre for Astrophysics, The University of Manchester, M13 9PL, UK
I. M$^{\rm c}$Hardy, School of Physics & Astronomy, University of Southampton, SO17 1BJ, UK

Abstract: The two processes which dominate the appearance of our universe are star-formation and accretion. Star-formation (SF) is fundamental to the formation and evolution of galaxies whilst accretion provides a major power source in the universe, dominating the emission from distant quasars as well as from nearby X-ray binary systems. The feedback between these two processes is also crucial, e.g. in reconciling the observed galaxy luminosity function with the predictions from the standard hierarchical clustering models. Radio observations provide by far the best single diagnostic of these two processes, providing a direct view of SF even in dusty environments and allowing detection of AGN and measurement of their accretion rate at bolometric luminosities far below anything detectable at higher energies. A large, statistically complete, sample of galaxies such as we propose here, provides the perfect laboratory to study for the first time not only these two processes, but also to quantify how they interact in different types of galaxies. e-MERLIN is perfectly tuned to such studies as, even at low frequencies, its pc-scale resolution, coupled with spectral information, allows almost unambiguous discrimination of even faint AGN inside large SF regions. Specifically here we will carry out a complete census of SF and AGN activity as a function of galaxy mass, morphology and spectral type, black hole mass and luminosity. We will thus calibrate other SF indicators, e.g. IR and Hα luminosity, and constrain patterns in jet strength compared to merger histories. Additionally, we will determine whether ultra-luminous X-ray sources may come from intermediate mass black holes.

The broad philosophy of this legacy programme will be to provide the definitive parsec-scale, μJy sensitivity radio images of a large sample of well-known galaxies in the nearby universe. As such this project will both address numerous key science questions regarding SF and activity in galaxies and is specifically designed to be a lasting Legacy data-set for the wider community, with the sample selected to maximize multi-wavelength coverage and consequently the amount of future legacy science achievable.

Full list of team members:

Susanne Aalto *(Onsala Space Observatory, Sweden)*,
Paul Alexander *(Cambridge)*,
Willem Baan *(ASTRON, The Netherlands)*,
Elias Brinks *(Hertfordshire)*,
Stephane Corbel *(CEA Saclay, France)*,
Tom Dwelly *(Southampton)*,
Danielle Fenech *(UCL)*,
Jack Gallimore *(Bucknell, USA)*,
Melvin Hoare *(Leeds)*,
Rob Kennicutt *(Cambridge)*,
Johan Knapen *(IAC, Tenerife, Spain)*,
Elmar Körding *(Southampton/Saclay, France)*,
Tom Maccarone *(Southampton)*,
Sera Markoff *(Amsterdam, The Netherlands)*,
Carole Mundell *(LJMU)*,
Alison Peck *(ALMA, Chile)*,
Miguel Perez-Torres *(IAA, Spain)*,
D. J. Saikia *(NCRA, India)*,
Ralph Spencer *(JBCA, Manchester)*,
Ian Stewart *(JBCA, Manchester)*,
Philip Uttley *(Southampton)*,
Martin Ward *(Durham)*,
Antxon Alberdi *(IAA, Spain)*
Megan Argo *(Curtin, Australia)*
Rob Beswick *(JBCA, Manchester)*
John Conway *(Onsala Space Observatory, Sweden)*
Phil Diamond *(JBCA, Manchester)*
Janine van Eymeren *(JBCA, Manchester)*
Jay Gallagher *(Wisconsin, USA)*
Dave Green *(Cambridge)*
Sebastian Jester *(MPIA Heidelberg, Germany)*
Hans-Rainer Klöckner *(Oxford)*
Christian Knigge *(Southampton)*
Ian M$^{\rm c}$Hardy*(Southampton)*
Jon Marcaide *(Valencia, Spain)*
Ivan Marti-Vidal *(Valencia, Spain)*
Tom Muxlow *(JBCA, Manchester)*
Alan Pedlar *(JBCA, Manchester)*
Cristina Romero-Canizales *(IAA, Spain)*
Eva Schinnerer *(MPIA Heidelberg, Germany)*
Ian Stevens *(Birmingham)*
Michele Thornley *(Bucknell, USA)*
Fabian Walter *(MPIA Heidelberg, Germany)*
Jeremy Yates *(UCL)*

1 Science Justification

1.1 OVERVIEW

Much of the intense star-formation (SF) and many supermassive black holes (SMBH) are heavily obscured by dense regions of dusty gas. For example, even normal SF zones are obscured in the FUV and the most intense areas can be totally obscured to any direct light from stellar sources or associated HII regions. While such zones are the birthplace of key astrophysical objects ranging from AGN to compact massive star clusters, it is difficult to probe their internal structures, a problem exacerbated by their small spatial scales of often <10 pc. Furthermore the study of compact star clusters and HII regions demonstrates that dense modes of SF can be found throughout the disks of galaxies. Thus while these types of processes are key evolutionary drivers, they are accessible in suitable spatial detail only within nearby samples of galaxies.

Radio emission offers a crucial means to understand dense SF regions and their relationships to AGN. Radio observations provides a direct view of dense SF regions and radio luminosity is an excellent star-formation rate (SFR) indicator. For the majority of supermassive black holes (SMBHs) in the nuclei of galaxies, accretion rates ($\dot{m}$) are low and radiation is produced inefficiently ($L \propto \dot{m}^2$), resulting in undetectably low optical through X-ray luminosities. However such SMBHs produce jets with easily detectable radio emission, providing a measure of their accretion rate.

All giant galaxies are hybrid objects having, to a greater or lesser extent, emission from SF regions and from accretion onto a SMBH. The strong correlation between the mass of the SMBH and the galaxy bulge mass (Magorrian et al 1998) indicates a strong correlation, averaged over the age of the universe, between SFR and black hole growth. However whether the two processes directly correlate, or whether one lags the other, or whether the galaxy type (spiral/elliptical, LINER/AGN/absorption line) affects the interaction is unknown. Does AGN feedback turn off SF in all galaxies, or only in those with very luminous AGN? Does it matter if the feedback energy comes in the form of a fast, highly collimated, jet or a less collimated Seyfert-type outflow? Is high SF activity required to fuel AGN accretion and are galactic bars efficient in terms of bringing fuel to SMBH and nuclear SF regions?

Better measures of the astrophysics of AGN-SF interactions in the local universe are essential for understanding our observations of hybrid objects at high redshifts. In this survey we will measure SFRs and accretion rates in a complete sample of galaxies of wide morphological and spectral type. Hence we will determine how the local galactic environment affects SFR and AGN activity and, in particular, the crucial interaction between those two fundamental processes. For example, what mechanism produces powerful jets and how do they relate to their surroundings?

This survey will also address many important specific points. For example, what mechanism produces powerful jets? Spin? There are hints from Balmaverde and Capetti (2006) that low luminosity AGN (LLAGN) are more radio loud if their galaxy bulges have a flat 'core' type light profile. Such a profile could be produced in a major merger, which might also give rise to large black hole spin. Our survey will greatly improve on the earlier observations and should give a definite answer. Also, is there a real physical difference between type 1 (broad line) and type 2 (narrow line) AGN or are observed differences purely a function of orientation-dependent obscuration? There are suggestions (Stevens et al 2005; Strong et al 2004) of higher SF and molecular gas content around type 2 AGN, implying a larger obscuring torus. Our observations will definitively measure SF products in galaxies and will probe the neutral and molecular gas which fuels both SF and accretion.

Although observations of the distant universe can tell us about the average evolution of SF and accretion rate with cosmic time, they do not have the spatial resolution to unravel the interaction between SFR and accretion in individual galaxies. Observations of our own Galaxy tell us about SF in individual molecular clouds and about accretion onto solar-mass sized compact objects but reveal nothing about how SFR or accretion onto SMBHs varies in galaxies of different mass or morphology. Only by observations of a large sample of nearby galaxies can we sample a wide enough range of galactic properties, with high enough angular resolution to answer the questions raised above. Here, by way of its unique combination of high resolution and high sensitivity, e-MERLIN has a vital role to play.

The critical importance of e-MERLIN observations Due to its high sensitivity and high angular resolution e-MERLIN is, and will remain for the foreseeable future, the best radio facility in the world for the study of SF and accretion and is perfectly matched to this programme. Our sample (described below) contains galaxies from $\sim$ 3 to 100Mpc. Even at the median distance of our more distant, shallower, sub-sample (20Mpc), a linear resolution of <4 pc will be achieved at C-band, easily resolving SNR and HII regions, and equivalent to sub-torus scales within the nuclei of nearby AGN. In our deeper subsample with a 5-σ detection limit of 14μJy bm^{-1} we will detect sources down to $\sim 10^{16}$ W Hz^{-1} at C-band, tracing far down into the radio luminosity function of both accretion dominated and SF related sources (see Fig 1). Moreover the field of view, will be $\gtrsim$3 arcmin for all of our observations, adequately covering all of the optical extent of the large majority of galaxies in the sample. The EVLA simply does not have the resolution, and VLBI does not have adequate sensitivity to extended structure, in the 1 to 7 GHz frequency range where synchrotron and free-free radio emission on physically important size scales within nearby galaxies is dominant.

In order to separately understand accretion and SF, we must be able to discriminate between signitures of the two. Radio morphology, particularly when, as here, coupled with spectral index information, is the single most important discriminator, and is especially strong in nearby galaxies where parsec-scale resolution is available. However no single diagnostic can provide unambiguous discrimination and other important tools such as radio/mid-IR ratio, near IR colours, X-ray luminosity and spectrum and optical spectral type are crucial. Taken together, these various diagnostics provide excellent discrimination (e.g. see Seymour et al 2008; Smolic et al 2008). Our sample has been designed to maximise the availability of these cross-wavelength discriminators.

Specific key science aims: Whilst the breadth and range of science that will come from these data will be extremely extensive, as is required from any legacy project, the key scientific aims of this project can be distilled into a few primary objectives:-

- To measure levels of SF activity in galaxies of all morphological, luminosity and spectral population mix. These measurements, and in particular the resulting systematic and unbiased census of individual SF products (RSNe, SNR, HII regions and alike), will be used as a direct extinction-free tracer of SF, and hence be used to help calibrate commonly used SFR indicators such as IR and Hα.

- A complete census of AGN activity and jet structures in galaxies of all optical types, including LINERs and absorption line galaxies as well as broad line AGN, which will be cross-correlated against levels of ongoing SF.

- A serendipitous parsec-scale imaging survey of the cold ISM using its HI absorption and maser emission. This survey will constrain the content and composition of cold gas present in the immediate nuclear region of galaxies as well as its kinematics. Both of which are basically unknown on these size scales and can be addressed with e-MERLIN well before ALMA or the SKA are fully operational.

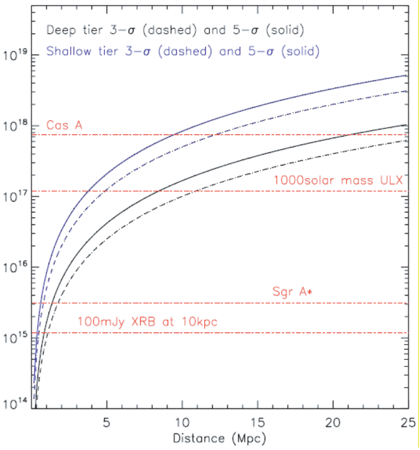

Figure 1: Luminosity sensitivity thresholds at 3 and 5-σ for both the deep (black) and statistical (blue) parts of this survey, along with specific luminosities of known sources and typical classes of sources.

This proposal: In order to address the key questions regarding SF and accretion in the local universe, we propose a comprehensive legacy programme surveying the radio emission from a large, unbiased and complete sample of 291 nearby galaxies. This project will consist two of closely related samples: a moderately deep snapshot survey of an unbiased sample of 280 galaxies selected from the Palomar Bright Galaxies Survey (Ho et al 1997), and an extremely deep survey of a nearby sub-set of 41 galaxies selected from the statistical sample to encompass the entire range of luminosities, scales and galaxy-types. The size of this sample follows the diversity of galaxy properties; it is large enough to provide statistically reality even though reality requires it cannot be statistically complete. This sub-sample has been specifically selected to have particularly superb ancillary data available across all wavebands, and in particular these targets have been selected to have *Spitzer* (SINGS) and *Herschel* (KINGFISH) coverage plus X-ray (*Chandra* and *XMM*), molecular

CO (IRAM HERA), HI (VLA THINGS), UV (GALEX NGS) data along with proposed EVLA continuum (KINGFISH follow-up) and extensive optical coverage.

We will observe all of the sample for at least 48 minutes in both L and C bands to provide a large, and statistically complete, sample covering a range of optical properties, to a low radio flux limit. However to study the very lowest surface brightness SF regions, and the very lowest luminosity AGN, we will observe a subset of 41 of our sample for 6 hours at both L and C band which will providing both superb *UV* coverage and high sensitivity.

These two sets of observations are highly complementary. The larger statistical sample will mainly probe the nuclear AGN and higher surface brightness starburst regions over all galaxy types, with large enough numbers that we can determine the way in which AGN and SF emission varies with galaxy type, and how the two emission processes interact. The deep sample will also sample a range of galaxy types in an unbiased, although not quite statistically complete, manner, allowing for a much more detailed study of, e.g. starburst and low luminosity AGN jet emission. Thus it will be possible to study the physics of the interaction between AGN emission and SF in much greater detail with the deep sample.

1.2 Core Science Themes

The core themes that this project will address are those of SF, and accretion and the fuelling of these two processes. In the three following subsections we outline *some* of the primary science drivers and products that will arise from this survey.

1.2.1 Star-formation and the compact radio source population of galaxies

Calibration of star-formation levels in Galaxies: The SF history, along with current levels of SF, within individual galaxies still remains a crucial physical parameter which observations are only now beginning to accurately characterise. Traditionally observations of optical line, IR and the global synchrotron emission from galaxies have often been used as proxies for SF. However, these tracers do have some fundamental flaws. In many galaxies optical emission lines are heavily obscured towards their centres, thus can require potentially large corrections. Whilst IR emission, which essentially traces the light from young stars reprocessed by dust and re-radiated at longer wavelengths, relies upon empirical interpretation of physically complex processes in order for it to be related to SF. Global radio synchrotron emission also provides an alternative, and extinction-free, indication of SF. However, the link between radio synchrotron emission and SF is also via complex physical mechanisms which are general calibrated using the radio-to-infrared correlation (Yun, Reddy & Condon, 2001; Bell 2003; Beswick *et al.* 2008). The long-standing critical issue which remains is how to calibrate either global radio or IR emission as a measure of SF, and how this calibration varies as a function of galaxy type and environment (Kennicutt 1998).

Highly sensitive, milliarcsecond radio images of nearby galaxies provide one method by which observations can directly address this issue in a way which is independent of complex physical emission mechanisms. Whereas low resolution radio observations of normal and SF galaxies trace the diffuse radio emission primarily resulting from charged particles that have escaped from old supernova remnants (SNR) (Condon 1992), with adequate sensitivity, high resolution observations can be used to systematically characterise the populations of individual SF products on a galaxy by galaxy basis, critically resolving away the diffuse emission. This population census can hence be used to directly infer the levels of SF.

At high angular resolution each individual external galaxy can be considered as a laboratory containing a large sample of discrete radio sources, which can be studied in a systematic way. When observed with μJy sensitivities at frequencies between 1 and 7 GHz this source population, with the exception of accretion dominated objects and in particular AGN (see section 1.2.2), will consist exclusively of sources related to various key phases of the stellar evolutionary sequence. This population will be a mixture of sources from the early stages of SF, such as compact HII regions, through superstar clusters (SSCs), and stellar end-points like X-ray binaries, planetary nebulae, supernovae (SNe) and their remnants (SNR).

All parts of this project will make significant contributions in this area. The larger statistical sample of galaxies will be observed at C-band to a 1-σ depth of 14 μJy bm^{-1}, which at the median distance of sources in this sample (20Mpc) will detect a significant fraction of SNR and HII regions, the majority of will be spatially resolved. However in the much deeper observations (3μJy bm^{-1} and 8μJy bm^{-1}, 1-σ at C and L-band respectively) all of these objects, such as those seen in well-studied nearby galaxies such as M82 (see Fig. 2), will be detected out to the maximum distance of this sub-sample (25 Mpc). In the closer sources the luminosity detection threshold will be 1.6×10^{17} W Hz at C-band, several times fainter than Cas A, implying that this survey will compile a complete census, at both frequencies, of all young radio SNRs (see Fig 1).

By firstly detecting and then identifying the physical nature of these objects using a combination of radio morphologies and spectral indices, alongside extensive multi-wavelength ancillary data, this programme will provide the first detailed extinction-free census of SF products within nearby galaxies. The majority of core-collapse SNe evolve to form long-lived radio SNR, hence this statistically well-constrained census combined with information regarding the sizes and hence canonical ages of SNR (derived directly from this project), can be used to directly infer levels of SF in individual galaxies (e.g Pedlar 2002; Fenech et al 2008).

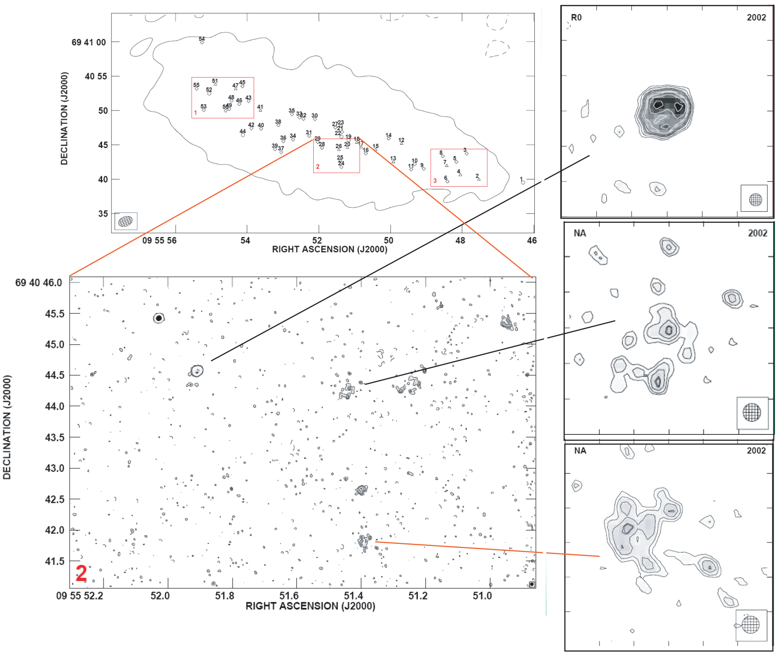

Figure 2: Current state-of-the-art, extremely deep 5GHz MERLIN observations of M82. With a rms sensitivity of 17 μJy bm^{-1} these observations required 8-days of integration and detected in total 55 SNR and HII regions (Fenech *et al.* 2008). These data have a sensitivity equivalent to our 'shallowest' $\sim$1 hr observations, whilst our deep 6 hr observations will reach > 6 times deeper with vastly increased image fidelity.

Importantly this survey will also identify populations of sources which trace earlier stages in stellar evolution, such as HII regions and SSCs, which can be used to place useful constraints on the levels of SF at various phases in the evolution of individual galaxies. When compared with other wavelength tracers, which probe different ranges of SF age and different spatial regions, these radio diagnostics will provide significant new insights.

Whilst this goal will be achievable on a galaxy-by-galaxy basis, the power and importance of this survey only arises from its large size and the available complementary multi-wavelength data-sets. By combination of these direct radio tracers of SF products, with other multi-wavelength SF proxies, significant constraints will be placed on their calibration and interpretation, which will have important implications across a wide range of observational astrophysics. Critically, the individual galaxies within this programme have been chosen to span a complete range of types and levels of both historical and ongoing SF, thus allowing this census of SF products to be applied over the wide range of luminosity and environment parameter space inhabited by galaxies.

Tracing the IR-radio correlation down the luminosity function: Radio and IR observations provide some of the best methods to study SF at all redshifts. Across many orders of magnitude in luminosity the emission from galaxies in these two wavebands have been shown to be very tightly correlated (Yun et al, 2001; Condon 1992). However, it has recently been suggested that for the lowest luminosity galaxies there may be some deviation from the tight well-known radio-IR correlation (Bell 2003; Boyle et al. 2007, Beswick et al 2008). Equally importantly, it is as yet unclear where this correlation breaks in terms of physical size-scale. In the few nearby spiral galaxies which have been studied so far, the IR-radio correlation has been shown to still be valid on kiloparsec scales (Murphy et al 2006). This programme of observations will provide a radio complement to both existing *Spitzer* and granted *Herschel* observations, which will both extend further the luminosity range of the studied systems, by including compact dwarf galaxies, and allow the relationship of the radio and IR emission process to be investigated on pc-scales.

Studies of individual components: At the depth and linear resolution that will be achieved by these observations, the studies of both galaxies and individual radio sources within them will be revolutionised. This will result in a wealth of legacy science which can be exploited, by this consortium and the wider astronomical community, especially when combined with existing multi-wavelength data. Whilst it is impossible to completely gauge discoveries and science that will arise from any survey which makes such a large leap in sensitivity and angular resolution, it is certain that these areas will include:

- Determination of the effect of environment on SF and AGN activity
- To study the demographics of compact nuclear SF with AGN type, i.e. to determine whether type 2 AGN show stronger SF from a larger dusty torus.
- Determine radio luminosity functions for both SF galaxies and AGN as a function of optical luminosity to provide a template on which the evolution of more distant galaxy samples may be based.
- Determination of the relationship between AGN and SF activity, particularly AGN fuelling and AGN feedback into the SF process.
- Measurements of the sizes of many hundreds of individual SNR in galaxies, providing 'epoch-zero' measurements which with future e-MERLIN observations can be used to determine SNR expansion velocities.
- Physics of compact HII regions as a function of galaxy environment.
- Detection of transient radio sources, such as new supernovae.
- Radio detection and imaging of planetary nebulae in nearby galaxies.

1.2.2 A complete census of accretion power in nearby galaxies

Accretion on to SMBH is one of the most significant energy sources in the universe, with the potential to clear star-forming gas from galactic bulges and even to regulate the growth of entire galaxies in galaxy clusters (e.g. di Matteo et al 2005). The mechanism for this feedback is mechanical, through jets and outflows powered by accretion. However, despite the importance of SMBH activity in regulating galaxy formation, comparatively little is known about SMBH activity towards low radiative luminosities. This is a significant gap in our understanding of feedback and the role of SMBH in galaxy growth and evolution, since it is now known that mechanical jet power can be energetically more significant than supernova feedback even at low AGN luminosities (Nagar et al. 2005; Koerding et al. 2008). The major difficulty in studying low luminosity AGN is precisely their low *radiative* output with respect to their surrounding host galaxy, especially since many LLAGN are embedded in nuclear SF regions (Ho et al. 1997). The problem is compounded because at low-luminosities AGN become radiatively inefficient, and lose the strong optical and X-ray signatures commonly used to identify AGN activity. However, just because an AGN is radiatively inefficient does not make it mechanically inefficient, in fact at low luminosities the total power is almost certainly mechanically dominated via the jet. And yet the only samples of low-luminosity SMBH activity studied to date have been optically selected, identified primarily from the Palomar spectroscopic survey of nearby galaxy nuclei (Ho et al. 1997). Even follow-up radio surveys, which should be sensitive to the jet power hidden in other wavebands, have been selected only from the subset of the Palomar sample which are *optically identified* as AGN (Filho et al. 2006; Nagar et al. 2004; Ho & Ulvestad 2001).

There is a pressing need to carry out a complete census of SMBH accretion power in the local universe, using deep e-MERLIN observations of a flux-limited sample of galaxies. Our statistical survey provides the ideal sample, since although the existing Palomar optical spectroscopy is available to aid our interpretation we make no *prior* selection on whether a galaxy is optically classified as active or not. Sensitive, high spatial-resolution radio observations are the best way to detect SMBH accretion in the local universe, primarily because of sensitivity. If we apply the 'fundamental plane' of accreting black holes relating X-ray luminosity, radio luminosity and BH mass (e.g. Falcke, Körding & Markoff 2004; Körding et al 2006; Merloni et al. 2003), then the 50 μJy 5-σ detection limit which we will reach in this survey corresponds to an X-ray flux of 5×10^{-15} erg s cm^{-2}, for a 10^7 M$_\odot$ BH at 20 Mpc (the median distance of the shallow sample), or a flux of 2×10^{38} erg s^{-1}, which is impossible to discern against the background of nuclear X-ray binaries in a galaxy nucleus. Based on the observed X-ray-[OIII] correlation in AGN (Heckman et al. 2005), the corresponding [OIII] line flux is only 7×10^{-17} erg s cm^{-2}, a factor 20 lower than the detection limit for emission lines in the Palomar spectroscopic survey.

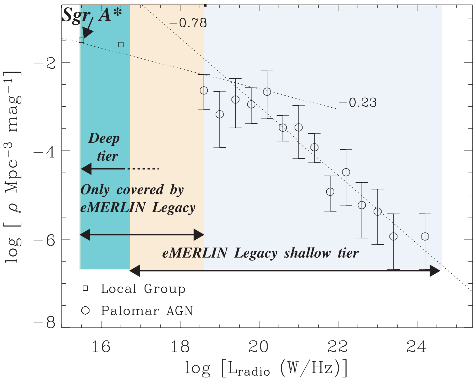

Figure 3: The radio luminosity function of the Palomar sample adopted from Nagar et al. 2005, and showing the extension to Sgr A* luminosities resulting from our survey. According to Nagar et al., there appears to be a break in the luminosity function around 10^{19} W Hz^{-1}, we will easily confirm this break, which will be important for determining where supernovae take over as being the most significant source of mechanical energy in galaxies.

Radio is also essential to unambiguously identify activity due to SMBH accretion. For example, even with more sensitive optical spectroscopy, it will be difficult to distinguish the weak optical emission from faint AGN against the contribution from nuclear SF. The equivalent [OIII] sensitivity for our snapshot survey corresponds to AGN line luminosities at 20 Mpc of order 10^{36} erg s^{-1}, making it extremely difficult using optical data to distinguish lowest luminosity AGN against any nuclear star-forming regions. In contrast, a clean separation of AGN and SF components can be made with the pc-scale resolution obtained by e-MERLIN, coupled with the leverage of L and C band fluxes to clearly distinguish the flat spectrum AGN cores.

Finally, using radio data circumvents the problem of where to draw the dividing line between a galaxy being ‘active’ or ‘normal’. This is because one can not only estimate the jet power from core radio luminosities but also measure the absolute mass accretion rate, which appears to be well correlated with radio luminosity for the lowest luminosity objects and where beaming effects do not play an important role, scaling as $L_R \propto \dot{M}^{1.4}$ (Körding et al. 2006). Thus we can classify all galaxies according to their SMBH accretion power, tying the lowest accretion ‘quiescent’ SMBH to the LINERs and classical Seyferts which make up almost half of the Palomar sample. The synergy between the shallow and deep observations is particularly important here. While the shallow tier will push the detection of LLAGN a factor 20 fainter than in the Palomar optical survey, the deep tier plays a key role in filling in the lowest luminosity end of activity from Sgr A* type activity upwards, due to its higher sensitivity and also lower median distance. For example, with the 14 μJy limit of the deep tier, accretion rates just a few times that in Sgr A* will be detectable out to $\sim$ 5 Mpc, encompassing $\sim$ 1/3 of the deep tier galaxies. Accretion rates of 10 times Sgr A* will be picked up by the snapshot survey at distances within 10 Mpc (about 30 galaxies). By combining the deep and shallow tiers we can measure the radio luminosity function of accretion from quiescent to fully active SMBH, and derive the associated accretion rate ‘luminosity function’ for SMBH accreting at rates similar to Sgr A* and above (see Fig. 1 and 3). Comparing this with the nuclear SFR which we can also derive from our survey, we can determine the accretion rate where jet power dominates the mechanical energy injected into galaxy bulges.

Besides making a complete survey of SMBH accretion in the local universe, one can use the results of our census for many exciting areas of research. Essentially, the results of our survey can be used for any work seeking to link SMBH activity with other galaxy properties, or putting SMBH activity into a wider context, e.g. through construction of luminosity functions and the implications for black hole growth and feedback. We list a few of these projects here:

The jet-host-galaxy connection: What causes the radio-loud/radio-quiet dichotomy? Recently, Balmaverde & Capetti (2006) have shown that LLAGN are systematically more radio loud if hosted by galaxy bulges with flat ‘core’ type light profiles rather than those with steeper ‘power-law’ type profiles. This result suggests an underlying mechanism for the radio-loud/radio-quiet dichotomy in AGN. One exciting possibility is that the core galaxies are the remnants of major mergers which have rapidly spinning SMBH as a result, leading to spin-powered jets. Alternatively, the difference could be due to the wider environment and type of accretion flow. We will widen the Balmaverde & Capetti result, limited to $>$ 1 mJy and only 3-5 arcsec resolution, to much lower luminosity ‘normal’ galaxies, whose environments and accretion flows should all be similar. If the same dichotomy persists, that would indicate that spin is indeed the origin of the most powerful jets; the results would clearly be very important for understanding AGN feedback.

AGN activity and nuclear star formation: Optical surveys find it impossible to cleanly distinguish *weak* AGN activity and nuclear SF, and X-ray observations are even worse. Our survey will allow us to measure weak AGN activity against the background of nuclear SF (including finally constraining whether LINERS nuclei are dominated by AGN or SF), and using the large sample statistics we can determine how SF affects AGN activity and vice versa, over a wide range of accretion rates. Does SF enhance SMBH activity, does nuclear SMBH activity suppress SF, perhaps through feedback, or are the accretion luminosity functions identical for galaxies with high and low nuclear SF?

Tracing intermediate mass Black holes and ULX sources: These data will give new means of searching for intermediate mass black holes (IMBHs) and testing claims that ultraluminous X-ray sources are IMBHs. Past work showed that several ultraluminous X-ray sources had radio flux upper limits which could be used, along with the fundamental plane relations for black hole activity, to constrain the masses of their black holes to being $\lesssim 1000\,\mathrm{M}_\odot$ (Körding, Colbert & Falcke 2005), using VLA data with rms noise of 14 μJy. This project will give black hole mass constraints about 10 times as small, allowing for meaningful tests of past claims from X-ray spectral fitting that the masses of black holes in ULXs are often $\gtrsim 1000\ M_\odot$. Additionally, the exquisite angular resolution of e-MERLIN, relative to that of the VLA, reduces the chance that a supernova remnant or an HII region will be mistaken for jet emission from the ULX itself (see e.g. Soria et al. 2006). An important piece of information for our understanding of ULXs also comes from the discovery of huge ionized bubble nebulae around a significant fraction of unobscured ULXs in nearby (∼2–5 Mpc) galaxies, that have allowed the (ionizing) X-ray luminosity to be measured independently (Pakull & Mirioni 2002; Kaaret et al 2004. For some of these sources, we are now also starting to observe large associated radio nebulae (Lang et al. 2007) that could be continuously energised by unseen jets produced by the ULX. As a calorimeter, the nebulae could provide a potential means to measure the total energy output of a ULX in all directions.

Searches for IMBHs which are not particularly bright X-ray sources can be made with this data set as well. The strongest cases for an intermediate mass black hole at the present time is in G1, the giant enigmatic star cluster in M 31 and the nucleus of M 33 (Wehner & Harris 2006). In M 31, G1 shows marginal dynamical evidence for containing an intermediate mass black hole (Gebhardt, Rich & Ho 2005), as well as X-ray (Trudolyubov & Priedhorsky 2004) and radio (Ulvestad, Greene & Ho 2007) flux densities consistent with what is expected for a 20,000 solar mass black hole. Searches for radio emission from star clusters and dwarf satellite galaxies around our target galaxies will provide constraints on what fraction of each class of object shows evidence for containing black holes of $\sim 1/1000$ of their total masses, in the same way that giant galaxies seem to do nearly universally.

1.2.3 Building blocks: the neutral and molecular ISM

In parallel to understanding the ongoing processes resulting from the birth and death of stars, and accretion, it is crucial that we understand the environment in which each of these sources is embedded. Radio, along with millimetre (CARMA, PdBI, SMA and ALMA) and infrared observations provide a unique means by which the physical relationship of the cold ISM (e.g. quiescent, entrained in winds etc) in nearby galaxies can be studied. This cold ISM in the form of neutral and molecular gas, and dust is the fuel for ongoing SF and accretion.

The built-in flexibility of the new e-MERLIN correlator means that it is possible to simultaneously observe several radio spectral lines alongside deep continuum observations, with no significant adverse effects on sensitivity (see Sec. 3). Using this facility, we will obtain high angular and velocity resolution observations of several important neutral and molecular species (HI, OH, H_2CO), via absorption and maser emission.

This portion of the programme will result in a simultaneous observed imaging survey of all key hydroxyl transitions (1612 MHz, 1665 MHz, 1667 MHz and 1720 MHz), HI absorption and the formaldehyde K-doublet (4829 MHz). These data products will be used to trace gas distribution and dynamics, via absorption against radio continuum components and emission from maser regions and the physical relationship of the cold gas with regions of ongoing SF and nuclear activity (e.g. Mundell et al 1995; Wills et al 1998; Beswick et al 2001, 2003; Argo et al 2007). These HI absorption observations of the cold neutral medium (CNM) when combined with VLA HI observations from THINGS will also provide direct constraints on the fraction of neutral gas in different phases of the ISM (CNM vs WNM). By simultaneously imaging multiple molecular lines, probing different conditions, we will also be able to constrain the physical state of the ISM, and by comparison to our deep radio continuum imaging this will allow the ISM properties to be directly related to the nuclear and SF activity within galaxies.

1.3 Sample selection and observational request

The ideal sample of galaxies To achieve the aims of this survey we require radio observations of high sensitivity and resolution of a large, unbiased and complete, sample of galaxies covering a range of morphological and spectral type, luminosity and bulge structure. For optimum linear resolution these galaxies must be nearby. The best sample currently available in the northern hemisphere is the Palomar BG sample, selected with a simple magnitude limit of $M_B < 12.5$. There are 280 such galaxies above $\delta = 20^\circ$ (to optimise our snapshot *uv* coverage), with a median distance of 20 Mpc, which form the basis of our sample.

Complementing this large snapshot survey, a sub-sample of 41 nearby galaxies, selected from the complete Palomar sample, will be observed five times deeper at both L and C-band, with superb image fidelity. This sub-sample has been specifically selected to be as representative as possible of the wide variety of galaxies in the local universe, and has the most extensive available range of multi-wavelength ancillary data. This sub-sample includes all of the northern ($\delta > 0^\circ$) targets (so as to achieve the maximum coverage with ALMA) within the *Herschel* KINGFISH survey, the majority of which have high quality *Spitzer* and optical coverage, via SINGS, in addition to both molecular and neutral gas observations (IRAM and the VLA); *a more extensive summary of this ancillary data is presented in section 2 and Annex A*. This sub-sample will also be part of extensive EVLA continuum follow-up observations to KINGFISH which will provide both a low-spacing complement to these high resolution e-MERLIN observations as well as comparable resolution high frequency data. Crucially, all of these projects are led by members of this consortium, thus affording both access and expertise to facilitate extensive multi-wavelength science exploitation of these e-MERLIN data. These e-MERLIN observations are designed to provide the high sensitivity and high angular resolution data which is currently not available to these major surveys.

In summary, to achieve both the highest possible sensitivity and to preserve the statistical nature of this survey this programme will be formed from a large snapshot survey of 250 nearby galaxies which will be complemented by a much deeper full-imaging and spectral-line survey of a representative nearby sub-sample of 41 of these very nearby galaxies. These deep observations will be used to supplement and complete the full coverage of the larger statistical sample (in total 280 galaxies), without the need to repeat observations. Together these observations will provide a unique radio legacy atlas of local galaxies.

1.4 Summary

This comprehensive proposal will address some of the fundamental astrophysical questions regarding levels of SF and accretion processes while also revealing their interactions on poorly explored spatial scales in a large, statistical sample of galaxies. In particular, this legacy programme will make a complete imaging census of both AGN activity and jet structures and radio sources relating to key SF phases across all galaxy types and luminosities in the local universe.

Our legacy programme will provide the definitive parsec-scale, μJy sensitivity radio images of a representative sample of galaxies in the nearby universe. This project will address numerous key science questions regarding SF and activity in galaxies and is specifically designed to be a lasting Legacy data-set for the wider community, with the sample selected to maximize multi-wavelength coverage and consequently the amount of future legacy science achievable. The results will also bridge multi-wavelength information from X-rays to FIR, thereby providing a foundation for legacy science studies of processes in galaxies ranging from SF and SMBH growth to the acceleration of cosmic rays.

• Argo *et al.* 2007, MN, 380, 596 • Balmaverde & Capetti 2006, A&A, 477,35 • Bell, 2002, ApJ, 577, 150 • Beswick *et al.* 2001, MN, 325, 151 • Beswick *et al.* 2003, MN, 346, 434 • Beswick *et al.* 2008, MN, 385, 1143 • Boyle *et al.* 2007, MN, 376, 1182 • Condon 1992, ARA&A, 30, 575 • di Matteo *et al.* , 2005, Nat, 433,604 • Falcke *et al.* 2004, A&A, 414, 895 • Fenech *et al.* 2008, MN, in press • Filho *et al.* 2006, A&A, 451, 71 • Gebhardt *et al.* 2005, ApJ, 634, 1093 • Heckman *et al.* 2005, ApJ, 634, 161 • Ho *et al.* 1997, ApJ, 487, 568 • Ho & Ulvestad 2001, ApJS, 133, 77 • Kaaret *et al.* , 2004, MN, 351, 83 Kennicutt, 1998, ARA&A, 36, 189 • Körding *et al.* 2005, A&A, 436, 427 • Körding *et al.* 2006, MN, 372, 1366 • Körding *et al.* 2008, MN, 383,277 • Laing *et al.* 2007, ApJ, 666, 79 • Magorrian *et al.* 1998, AJ, 115, 2285 • Merloni *et al.* 2003, MN, 345, 1057 • Murphy *et al.* 2006, ApJ, 638, 157 • Nagar *et al.* 2005, A&A, 435, 521 • Pakull & Mirioni 2002, astro-ph/0202488 • Pedlar 2001, IAUS, 205, 366 • Seymour *et al.* 2008, MN, 386, 1695 • Smolic *et al.* 2008, ApJS, 177, 14 • Soria *et al.* 2006, MNRAS, 368, 1527 • Strong *et al.* 2004, MN, 352, 1151 • Trudolyubov & Priedhorsky 2004, ApJ, 616, 821 • Ulvestad *et al.* 2007, ApJ, 661, 151 • Wehner & Harris 2006, ApJ, 644, L17 • Wills *et al.* 1998, MN, 298, 347 • Yun, Reddy & Condon 2001, ApJ, 554, 803 •

2 Ancillary data sets and links with other legacy programmes

Links to other projects and ancillary data-sets: This legacy project has strong links with many major existing and planned legacy projects, investigating the properties of galaxies in the local universe. These links are driven by their complementary nature and the key science questions which they address. In particular this project has direct associations in terms of scientific goals, sample selection and personnel, with several large programmes (see also Annex A) in the infrared (KINGFISH, Herschel key programme and SINGS, Spitzer legacy), millimetre and radio spectral line (THINGS, VLA HI survey and an IRAM HERA CO survey) and in the radio continuum (via LOFAR surveys KSP and planned EVLA follow-ups to KINGFISH) wavebands. *This e-MERLIN legacy programme will provide the crucial, sensitive and high-resolution radio data, currently missing from this large multi-wavelength observing campaign.*

Supporting data-sets already in hand: Of the 280 galaxies within our large statistical sample most have extensive existing multi-wavelength coverage; 50% of the sample *Chandra*, and about 57% with *Spitzer* photometry (50% with spectroscopy) data. This sample is also very well covered by existing ground based optical observations.

The deep sub-sample selection of this Legacy survey has been specifically designed to match existing and ongoing legacy programmes on other instruments. Consequently, a uniquely extensive range of complementary multi-wavelength data-sets are available, either publicly or via this collaboration; including *Herschel* (KINGFISH, PI: Kennicutt), *Spitzer* (SINGS, PI: Kennicutt), complete ground based imaging of the sample in UBVRIJHK and Hα, and Paα and H-band using *HST* (ancillary data associated with SINGS), VLA HI emission data (THINGS, PIs: Walter, de Blok & Brinks), CO data via SONGS, *GALEX* Nearby Galaxies Survey, JCMT Local universe Survey, VHIKINGS private VLA HI survey (Mundell), CO emission IRAM-30m HERA (PI: Walter), WSRT and the UKIRT NGS. In addition, this sample has been chosen to have a large overlap with existing archival data (e.g. *HST, Chandra & XMM*).

Planned follow-up observations associated with this Legacy project: Complementing this e-MERLIN programme, associated observations of portions of this sample are being planned with the EVLA (Schinnerer), European-LOFAR (Beswick and Conway) and the GMRT alongside various optical programmes, to be initiated as a direct consequence of this project, including using WIYN (Gallagher) and the GranTeCan and WHT on La Palma (Knapen), using the Fabry-Perot instrument GHaFaS on the latter. In addition, mm-wave continuum and spectral line observations, in support of this programme, are being planned, including CO2-1 and 1mm continuum surveys using existing sub-mm/mm arrays SMA (Peck), CARMA and PdBI along with single dish mm-wave lines and continuum using APEX (Aalto), JCMT, OSO-20m and other appropriate instruments (Baan) which will add greatly to our understanding of the ISM in the nuclear regions of these sources. These follow-up observations will be coordinated with this proposal, providing both new ancillary data, and supplementing existing data to gain complete coverage of our targets.

Synergies with other e-MERLIN Legacy proposals: This e-MERLIN legacy project bridges our understanding of SF and accretion on small scales within our galaxy and how these processes shape the evolution and appearance of entire galaxies. Consequently this project occupies a vital position; placing detailed studies of both accretion and SF processes in our own Galaxy within the context of the local galaxies, and providing local universe analogies to aid the physical interpretation of observations of high redshift galaxies.

Important scientific and technical synergies between various other e-MERLIN legacy programmes already exist, with several key co-investigators playing leading roles in other relevant projects. These include **EMERGE deep fields survey** (Muxlow [PI], Beswick, M^{c}Hardy, Diamond, Dwelly, Perez-Torres), **AGATE & e-QUATE** (Beswick), **LIRGI** (Conway [PI], Alberdi, Baan, Beswick, Diamond, Klöckner, Perez-Torres, Romero-Canizales), **COBRaS: Massive and young stars in the Galaxy** (Yates, Fenech, Diamond, Stevens), **Feedback processes in Massive star formation** (Hoare [PI], Diamond) and **High-energy astrophysics: ToO observations** (Beswick, Spencer, Körding). These links will be further fostered to insure both technical and scientific collaboration with the rest of the legacy programme at all stages, and that technical developments are not duplicated.

3 Technical justification

The technical strategy of this proposal, including required sensitivities and sample size, is driven by the scientific requirements to both observe a statistically large and well selected sample of galaxies and to provide imaging with a low enough luminosity sensitivity to fully trace and characterise the faint radio source populations of objects within individual galaxies. In order to fulfill these science aims in the most efficient manner, this programme will utilise both snapshot imaging techniques to observe a large sample with moderately high sensitivity and good image fidelity, and complete imaging runs of a sub-sample of sources, which will be 5 times deeper and with exquisite image fidelity. All observations will be made in L and C-band, across matching frequency bands, and in full-polarisation mode; thus providing both spectral index and polarisation information for every target. We will utilise all 16 available sub-bands of width 32 and 128 MHz correlated into 128 and 256 channels per polarisation product, incorporating 7 and 3-bit digitisation, at L and C-band respectively, apart from those dedicated for spectral line use. This mode will allow a field of view of up to ∼37′and ∼9′at L and C-band to be imaged with only a few percent radial smearing.

In total 291 galaxies will be observed in both L and C-band, with 250 sources observed only in snapshot mode and 41 observed with full 6 hr long imaging runs. A summary of this request is listed below, including calibration overheads.

	Number of targets	Integration per source	Estimated rms (μJy bm^{-1})	Luminosity (at median dist)	Total time inc. overheads
Snapshot (L-band)	250	48 m	38	1.8×10^{18} W Hz^{-1}	325 hr
Snapshot (C-band)	250	48 m	15	7.2×10^{17} W Hz^{-1}	325 hr
Deep (L-band)	41	∼4.8 hr	8	7.5×10^{16} W Hz^{-1}	246 hr
Deep (C-band)	41	∼4.8 hr	3	2.8×10^{16} W Hz^{-1}	246 hr
Grand total	291				1142 hr

Table 1: Summary of observing request. The inclusion of the Lovell Telescope is requested for the deep part of this survey only.

Snapshot imaging: In order to produce the high image fidelity required to reliably reconstruct the complex source structure expected from each galaxy, snapshot imaging run they will be segmented into at least 6 scheduling blocks each observed at different hour angles in order to, as completely as possible, fill the *uv* plane (see Fig 4).

At both C and L-band each source will be observed across the largest available e-MERLIN bandwidths. In the L-band (1.3-1.7 GHz) the precise position of individual side-bands and number of bits recorded will be optimised at the time of the observations to account for the presence of any interference. At C-band these snapshot images will utilise the full 2.048 GHz e-MERLIN bandwidth at the bottom end (4.5 – 6.5 GHz) of the available spectrum. We opt to observe in this specific frequency range since it maximises the sensitivity of certain elements of the array, notably the Defford telescope, and to match the observed frequency band of our deep observations.

We do **not** request the inclusion of the Lovell telescope within the snapshot portion of this legacy project. Whilst the inclusion of the Lovell telescope does increase the overall array sensitivity by a factor of 2-3, its relatively slow drive speeds would result in a significant increase in calibration overheads.

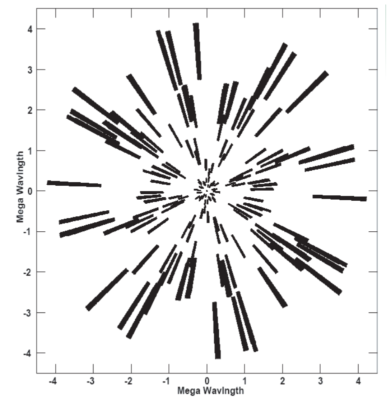

Figure 4: Simulated Fourier plane coverage of 4.5-6.5 GHz e-MERLIN snapshot image of total length 48 minutes split into 6 evenly spaced hour angle cuts for a source at declination +70 degrees.

Efficient scheduling: The optimum way to schedule this part of the survey, including phase referencing and calibration, is observe each source at both L and C-band before repeating this cycle for the next source and so on (see Table 2). This scheduling scenario would provide both quasi-simultaneous multi-frequency observations plus will minimise the telescope drive times. However, whilst the e-MERLIN system has this capability, each band change requires the rapid switching of receiver systems involving significant mechanical operations at each telescope. Thus in order to minimise potential telescope failure, we adopt a more conservative strategy of observing in a single frequency band per observing track, with observations made on adjacent days. In total this will result in a slightly increased time request (which is outlined in Table 1) but a significantly reduced risk of telescope failure.

L-band			C-band			C-band			...
Cal1	Target1	Cal1	Cal1	Target1	Cal1	Cal2	Target2	Cal2	..
4m	8m	2m	2m	8m	2m	4m	8m	2m	...

Table 2: Optimal multi-frequency snapshot scheduling blocks, including nominal on-source integration times. In the full schedule each source would be observed 6 times at evenly spread hour angles resulting in a total of 48 minutes of observations. Individual flux and bandpass calibrators would be scheduled daily and shared between all sources. Note that this scheduling scenario assumes rapid changes in observing bands.

Deep imaging: Each of the 41 targets within the deep sub-sample will be observed to a sensitivity 5 times deeper than the snapshot images. This depth will be achieved by observing each target for a total of 6 hr including phase reference calibration using the full eMERLIN array. To maximise the sensitivity achieved in this portion of the survey we request the inclusion of the Lovell telescope in both bands. In total this will result in approximately 5 hrs of on-source integration in each frequency band resulting in a rms sensitivity of $\sim$3 and 8 μJy bm^{-1} at C and L-band respectively.

For each individual source, observations will be split into three 2-hr long blocks, at both L and C-band, which will be interleaved within an individual observing track. This strategy will build up the full on-source integration time, whilst maximising the hour angle coverage and hence imaging capabilities. This will also provide near-simultaneous radio spectral index information about all sources and allow the sample to be built-up from complete observations of individual targets, facilitating rapid science exploitation of the survey prior to it being fully completed.

The field of view of these observations will be limited by the primary beam response on the Lovell Telescope baselines. The maximum field of view of the array including the Lovell telescope is $\sim$12 and $\sim$3 arcmin (HPBW) at L and C-band respectively. This FoV will allow the complete imaging of the central regions of interest of all of our target galaxies. However this FoV can be significantly increased by discarding data on the Lovell telescope baselines allowing the full FoV of the 25-m dishes to be imaged.

Spectral line: In tandem with these deep continuum observations we will use the flexibility of the e-MERLIN correlator to simultaneously observe several key radio spectral lines. In particular we will target the redshifted 21 cm line of HI, the four OH lines at L-band (1612, 1665, 1667 and 1720 MHz) and the formaldehyde K-doublet at 4829 MHz, all of which are situated within our observing band. To achieve this at L-band four of the 16 available sub-bands will be configured with a reduced band-width and centred upon the redshifted frequency of each line. One single sub-band can be used to simultaneously cover both 1665 and 1667 MHz OH maser lines. Each of these sub-bands will be observed with a total bandwidth of 16 MHz correlated into 1024 channels, resulting in a velocity resolution of $\sim$ 3 km s^{-1}. This will result in a 1-σ line sensitivity of $\sim$6 mJy bm^{-1} chan^{-1}, before smoothing. In total this will slightly reduce total available continuum bandwidth at L-band hence resulting in a few % loss in continuum sensitivity. At C-band one sub-band will be used to observe formaldehyde using a single 32 MHz sub-band correlated into 1024 channels resulting in a velocity resolution of $\sim$4 km s^{-1} and line rms sensitivity of $\sim$ 3 mJy bm^{-1} chan^{-1}, and reducing the overall continuum sensitivity at C-band by less than 1%.

Validation of observing strategy and phasing of observations: In order to fully optimise our observing schedules and set-ups, including testing of our snapshot imaging capabilities, we request that in the initial phase of operations that several of our highest priority deep imaging observations be made. This will allow us to test our initial pipeline-processing and immediately confirm the robustness of the observing plan, and in particular provide observational data to test our snapshot imaging strategy. In addition these data will form part of our project's early release observations, and will provide high impact early science for e-MERLIN.

4 Pipeline processing, archiving and data release

Data structure: All of the observations within this programme will be made in a pseudo-spectral line mode, providing full wide-field imaging capabilities. Additionally the observations from the deep sub-sample will also utilise the flexibility of the correlator to configure individual sub-bands for high velocity resolution spectral line coverage. All observations will be phase referenced, with regular bandpass and flux calibration. In total the data volumes will be typically $\sim$10-20GB hr^{-1}.

The data will initially be converted to a UV-FITS format compatible with AIPS or an EVLA/ALMA binary data format for processing in CASA. It is envisaged that, at least in the initial phases AIPS, along with bespoke algorithms and parseltongue pipeline scripts (developed by this team) will be used to process the data.

Processing and pipelining: Due to the complicated mixture of spectral-line and continuum data which will be gathered by this experiment we will initially separate these spectral-line components by sub-band. This will allow the spectral-line data to be independently analysed in a manageable way. In parallel the continuum sub-bands, including snapshot imaging, will be reduced using pipeline processes, including automated flagging routines (utilising telescope pointing data, T_{sys} measurements and auto-correlation spectra), initial amplitude calibration from T_{sys} data, data editing, and amplitude, phase and passband calibration applied individually by sub-band and as function of both frequency and time.

It is envisioned that these *calibrated* data will initially be imaged, sub-band by sub-band to further assess data quality, and to identify and subtract any confusing sources prior to full continuum imaging and any self-calibration. Each sub-band will be polarisation calibrated using bright calibration sources. The final imaging of each target will be undertaken using multiple facets, in order to image the full field, and multi-frequency deconvolution algorithms to fully account for spectral behaviour of sources.

This legacy team will work closely with the e-MERLIN team and other legacy projects to share resources in developing these techniques. In particular, we will use the large level of existing experience within the team to focus upon the development and implementation of multi-frequency deconvolution algorithms (Stewart, Fenech et al) and pipeline analysis techniques to deal with mixed continuum and spectral line data, in additional to the optimisation of e-MERLIN snapshot imaging surveys. Development in each of these areas will be required for many future e-MERLIN projects, both Legacy and standard proposals, and will be a benefical product of this programme. This team is already developing these multi-frequency imaging algorithms and is testing them on existing 10-band MERLIN MFS data between 4.5 and 6.5 GHz in preparation for e-MERLIN.

Planned data reduction centres: The large volumes of data that will be produced by this project necessarily mean that multiple data reduction centres will be required. As such we will spread the distribution of analysis of data amongst the team. In particular, this effort (including pipeline development) will be primarily focused at Manchester (Beswick, Muxlow, van Eymeren, Spencer, Diamond, Stewart, Pedlar), Southampton (M^{c}Hardy, Dwelly, Körding) and UCL (Yates, Fenech), providing both computational and man-power resources plus natural complementary synergies with other potential Legacy programmes (e.g. EMERGE and COBRaS). Radio experts from throughout the collaboration will aid with this data reduction, both at their home institutes and via extended visits to either JBCA or other data centres.

Final data products, archiving and data access: The final data products of this experiment will be both deep fully calibrated images and uv data-sets for each target source at each band observed. These data will include, full sensitivity wide-band images along with pixel-by-pixel spectral information for each field across the entire imaging band, and individual spectral-line cubes. Each of these data products will promptly be made public and will be VO-compliant.

Delivery schedule: In order to facilitate rapid dissemination of the results of this project we plan to implement a rapid and phased release of enhanced data products. These enhanced data products, will take the form of final images, cubes and spectral index information. We plan to phase the release of these data products, in line with when individual observations are made, making individual blocks of observations available to the full consortium within a few months of each observation and to the full community within a maximum of 12 months of the end of each semester's observations.

5 Management & Resource plan

Management structure: The overall management of this project will be split between the two project coordinators, Beswick and M^{c}Hardy. Beswick will take the primary role in overseeing the planning and scheduling of all observations, along with initial data pipeline processing, and facilitate overall science output from the deep sub-sample. M^{c}Hardy will take the key responsibility to insure the success and scientific robustness of the statistical sample and will assist with the coordination of additional required follow-up observations which will be led by other members of the team.

Core delivery teams and Data working groups: During this initial planning, algorithm and pipeline development and data acquisition phase, individual *data working groups* (detailed below) will be formed to undertake these key tasks and minimise duplication effort. Members of these teams, already working closely with the e-MERLIN team (which include Beswick and Muxlow), and will continue throughout the e-MERLIN commissioning phase and the duration of this project, to fully develop analysis strategies. It is intended that these teams will also work closely with other e-MERLIN Legacy programmes, providing a cross-fertilization of skills and resources, and to minimise any duplication of effort.

Core delivery teams	
Obs planning & scheduling	Beswick, M^{c}Hardy, Muxlow, van Eymeren, Körding
Imaging algorithm and pipeline development	Stewart, Fenech, Beswick, Muxlow, Klöckner
Continuum imaging	Beswick, Muxlow, Stewart, Conway, Green, Gallimore, Fenech, Körding, PDRA (soton), Marti-Vidal, van Eymeren, Yates
Spectral line imaging	Argo, Baan, Beswick, Brinks, Conway (+PHD), van Eymeren, Gallimore, Klöckner, Peck, Yates

Science working groups: The science exploitation phase of this project will be organised around a series of primary Science Working Groups (SWGs), under which specific research projects will be undertaken by smaller groups and individuals with specific interests and skills. These SWGs will be set up in the 6-12 months after the formal allocation of time to the project. This timescale should be sufficient to allow leading members to bid for additional support from STFC and other international funding bodies. Membership of the SWGs will be open to all and the intention is that at the minimum they will act as venues to minimise unnecessary duplication of effort and to foster collaboration.

Project Team: The scientific promise of this project has brought together a large and diverse team of experts. This team brings significant radio interferometric experience (both technical and observational) which will contribute to the success of this project and e-MERLIN as whole. Equally many of the collaboration will contribute dedicated observations with other instruments (see Sec. 2) and significant levels of PDRA and PhD support (see below and Sec. 4).

Additional resources: The consortium will, in addition to significant individual contributions by team members, bring significant PDRA and PhD student support to the project. In the first instance this will include, at least a dedicated PDRA in Manchester (van Eymeren), Southampton and Hertfordshire (pending continuation of STFC allocation). It is planned that this project will also benefit from and provide resources for several PhD students, who will play significant leadership roles in the science exploitation of this legacy project. Initially these will include PhD projects based at Manchester, Onsala, Saclay and Birmingham.

6 Scope of project and its Legacy status

Nearby galaxies will always be of fundamental importance in almost all aspects of astrophysics and cosmology, e.g. galaxy formation and dynamics, black hole growth, simply because, being nearby, we can observe them with much better spatial resolution than any other galaxies. Thus aspects of galactic structure such as the light profiles of the bulges, the presence of counter-rotating gas and star streams, can be easily studied whereas it would be impossible to study these processes in more distant galaxies. Nearby galaxies form a vital link between the very detailed study of our own Galaxy and statistical, but not very deep, studies of large samples of very distant galaxies. Nearby galaxies are therefore always going to be the targets of major observing programmes with every new observatory that comes along (e.g., coming soon, Herschel, ALMA, JWST).

It is, of course, important that one chooses sensible samples of nearby galaxies. The galaxies that we select here have already been chosen, by many other authors, as targets of major observing programmes. The Palomar sample is the best defined nearby galaxy sample in the sky, being selected purely on a magnitude cut-off limit and has been widely observed in all wavebands. The deeper subset has been even more widely observed, as discussed earlier. The excellent selection criteria, coupled with the plethora of previous observations in other bands, make our sample an absolutely classic legacy sample.

Legacy fields have two main defining characteristics by which you can recognise them: other observers choose to make their own, different, observations in the same fields and archival researchers use them for different purposes to that for which the original legacy team made them. Guessing the aims of future archival researchers is largely a catch-22 situation as, if we knew what these good ideas were, we ought to have mentioned them. However it is not our intention to do more than the most cursory study of polarisation, e.g. in AGN jets, and other researchers may wish to make such a study as we will record full polarisation data. It may be possible to find features within each spectral band which are useful for studying galactic dynamics, although this possibility has not yet been investigated.

It is, however, very clear that future observers will probably want to make Herschel observations to fully study the temperature distribution of the cooler gas in our sample galaxies. Also, observers will undoubtedly want to use the EVLA, to combine with our observations, to improve sensitivity to low surface brightness structures. Future observers with JWST and ALMA will want to compare their high spatial resolution images with our matching resolution radio images of extended emission to improve upon our understanding of the radio/IR relationship in galaxies of differing SF rate or relate molecular gas structures to radio continuum. It is not hard to think of things to do with our sample with future observatories.

Annex A: The deep '6 hr' source sample

Name	Alt. ID	RA2000	DEC2000	SINGS	KINGFISH /GTO Herschel(*)	THINGS	IRAM (* pending)	Type	Nuc Type	D25 (arcmin)	Dist) (Mpc)	SFR $M_\odot\,yr^{-1}$
NGC0628	M74	01:36:41.77	+15:46:59	yes	yes	yes		SAc		10.5 x 9.5	11.4	4.0
NGC0925	UGC01913	02:27:16.88	+33:34:45.0	yes	yes	yes	yes	SABd	H	10.5 x 5.9	10.1	2.4
IC342		03:46:48.5	+68:05:46	no	yes			SAB	Sy2/H	21.4 x 20.9	2.4	
NGC1569		04:30:49.0	+64:50:53	no	no	yes		IBm		3.6 x 1.8		
NGC2146		06:18:37.7	+78:21:25	no	yes		yes*	SBab pec	SB	6.0 x 3.4	14.5	
NGC2403	UGC03918	07:36:51.40	+65:36:09.2	yes	yes*	yes		SABcd	H	21.9 x 12.3	3.5	1.3
NGC2841	UGC04966	09:22:02.63	+50:58:35.5	yes	yes	yes	yes	SAb	L/Sy	8.1 x 3.5	9.8	0.2
NGC2903		09:32:10.1	+21:30:04	no	no	yes	yes	SBbc	H	12.6 x 6.0	6.5	
NGC2976	UGC05221	09:47:15.46	+67:54:59.0	yes	yes	yes	yes	SAc	H	5.9 x 2.7	3.5	0.2
NGC3031	M81	09:55:33.17	+69:03:55.1	yes	yes*	yes		SAab	L	26.9 x14.1	3.5	1.1
NGC3034	M82	09:55:52.22	+69:40:46.9	yes	yes*			IO	SB	11.2 x 4.3	3.5	6.0
NGC3077		10:03:20.6	+68:44:04	no	yes	yes	yes*	I0pec	H	5.4 x 4.5	3.5	
NGC3190	UGC05559	10:18:05.64	+21:49:55.0	yes	yes		yes*	SAap	L	4.4 x 1.5	17.4	<0.1
NGC3184	UGC05557	10:18:16.98	+41:25:27.8	yes	yes	yes	yes	SABcd	H	7.4 x 6.9	8.6	1.2
NGC3198	UGC05572	10:19:54.90	+45:32:58.8	yes	yes	yes	yes	SBc		8.5 x 3.3	9.8	0.85
IC2574	UGC05666	10:28:21.2	+68:24:43	yes	yes	yes	yes	SABm		13.2 x 5.4	3.5	0.10
NGC3351	M95	10:43:57.73	+11:42:13.0	yes	yes	yes	yes	SBb	SB	7.4 x 5.0	9.3	1.2
NGC3627	M66	11:20:15.03	+12:59:29.6	yes	yes	yes		SABb	Sy2	9.1 x 4.2	8.9	6.9
NGC3938	UGC06856	11:52:49.45	+44:07:14.6	yes	yes		yes*	SAc		5.4 x 4.9	12.2	1.2
NGC4125	UGC07118	12:08:06.02	+65:10:26.9	yes	no			E6p		5.8 x 3.2	21.4	...
NGC4214		12:15:38.9	+36:19:40	no	no	yes	yes	Irr		8.5 x 6.6	4.3	
NGC4236	UGC07306	12:16:42.12	+69:27:45.3	yes	yes			SBdm		21.9 x 7.2	3.5	0.3
NGC4254	M99	12:18:49.63	+14:24:59.4	yes	yes		yes*	SAc		5.4 x 4.7	20.0	11.0
NGC4321	M100	12:22:54.90	+15:49:20.6	yes	yes		yes*	SABbc	L	7.4 x 6.3	20.0	5.5
NGC4449		12:28:11.2	+44:05:36	no	no	yes		Irr	SB	6.2 x 4.4	3.5	
NGC4450	UGC07594	12:28:29.63	+17:05:05.8	yes	no			SAab	L	5.2 x 3.9	20.0	0.5
NGC4536	UGC07732	12:34:27.13	+02:11:16.4	yes	yes		yes*	SABbc	H	7.6 x 3.2	25.0	3.7
NGC4552	M89	12:35:39.81	+12:33:22.8	yes	no			E	L	5.1 x 4.7	4.5	...
NGC4559	UGC07766	12:35:57.69	+27:57:35.1	yes	yes			SABcd	H	10.7 x 4.4	11.6	...
NGC4569	M90	12:36:49.80	+13:09:46.3	yes	yes		yes*	SABab	L/Sy	9.5 x 4.4	20.0	1.9
NGC4579	M58	12:37:43.60	+11:49:05.1	yes	yes		yes*	SABb	L/Sy	5.9 x 4.7	20.0	2.0
NGC4631	UGC07865	12:42:08.01	+32:32:29.4	yes	yes		yes*	SBd		15.5 x 2.7	9.0	3.3

Table 3 – continued from previous page

NGC4725	UGC07989	12:50:26.61	+25:30:02.7	yes	yes		yes*	SABab	Sy2	10.7 x 7.6	17.1	...
NGC4736	M94	12:50:53.06	+41:07:13.6	yes	yes	yes	yes	SAab	L	11.2 x 9.1	5.3	2.1
NGC4826	M64	12:56:43.76	+21:40:51.9	yes	yes	yes		SAab	Sy2	10.0 x 5.4	5.6	0.3
NGC5194	M51a	13:29:52.71	+47:11:42.6	yes	yes*	yes		SABbc	H/Sy2	11.2 x 6.9	8.2	5.4
NGC5457	M101	14:03:12.5	+54:20:55	no	yes	yes	yes*	SAB		28.8 x 26.9	4.9	
NGC5474	UGC09013	14:05:01.61	+53:39:44.0	yes	yes		yes*	SAcd	H	4.8 x 4.3	6.9	0.2
NGC5866	UGC09723	15:06:29.56	+55:45:47.9	yes	yes			S0		4.7 x 1.9	12.5	<0.1
NGC6946	UGC11597	20:34:52.34	+60:09:14.2	yes	yes	yes	yes	SABcd	H	11.5 x 9.8	5.5	2.2
NGC7331	NGC7331	22:37:04.10	+34:24:56.3	yes	yes	yes	yes	SAb	L	10.5 x 3.7	15.7	4.2