Mon. Not. R. Astron. Soc. **363,** 563–580 (2005) doi:10.1111/j.1365-2966.2005.09460.x

The SCUBA Half-Degree Extragalactic Survey – I. Survey motivation, design and data processing

A. M. J. Mortier,[1⋆] S. Serjeant,[1] J. S. Dunlop,[2] S. E. Scott,[2] P. Ade,[3] D. Alexander,[17] O. Almaini,[4] I. Aretxaga,[5] C. Baugh,[6] A. J. Benson,[7] P. N. Best,[2] A. Blain,[8] J. Bock,[9] C. Borys,[8] A. Bressan,[10] C. Carilli,[11] E. L. Chapin,[5,14] S. Chapman,[8] D. L. Clements,[12] K. Coppin,[13] M. Crawford,[2] M. Devlin,[14] S. Dicker,[14] L. Dunne,[4] S. A. Eales,[3] A. C. Edge,[6] D. Farrah,[15] M. Fox,[12] C. Frenk,[6] E. Gaztañaga,[5,16] W. K. Gear,[3] E. Gonzales-Solares,[17] G. L. Granato,[10] T. R. Greve,[8] J. A. Grimes,[7] J. Gundersen,[18] M. Halpern,[13] P. Hargrave,[3] D. H. Hughes,[5] R. J. Ivison,[2,19] M. J. Jarvis,[7] T. Jenness,[20] R. Jimenez,[14] E. van Kampen,[2,21] A. King,[12] C. Lacey,[6] A. Lawrence,[2] K. Lepage,[13] R. G. Mann,[2] G. Marsden,[13] P. Mauskopf,[3] B. Netterfield,[22] S. Oliver,[23] L. Olmi,[24] M. J. Page,[25] J. A. Peacock,[2] C. P. Pearson,[26] W. J. Percival,[2] A. Pope,[13] R. S. Priddey,[27] S. Rawlings,[7] N. Roche,[2] M. Rowan-Robinson,[12] D. Scott,[13] K. Sekiguchi,[28] M. Seigar,[20,29] L. Silva,[30] C. Simpson,[31] I. Smail,[6] J. A. Stevens,[19] T. Takagi,[1] G. Tucker,[32] C. Vlahakis,[3] I. Waddington,[23] J. Wagg,[5] M. Watson,[33] C. Willott[34] and M. Vaccari[12]

[1] *Centre for Astrophysics and Planetary Science, School of Physical Sciences, University of Kent, Canterbury, Kent CT2 7NR*
[2] *Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ*
[3] *Cardiff School of Physics and Astronomy, Cardiff University, 5, The Parade, Cardiff CF24 3YB*
[4] *The School of Physics and Astronomy, University of Nottingham, University Park, Nottingham NG7 2RD*
[5] *Instituto Nacional de Astrofísica, Óptica y Electrónica (INAOE), Apartado Postal 51 y 216, 72000 Puebla, Puebla, Mexico*
[6] *Institute for Computational Cosmology, University of Durham, South Road, Durham DH1 3LE*
[7] *Department of Astrophysics, Denys Wilkinson Building, Keble Road, Oxford OX1 3RH*
[8] *Caltech, 1200 E. California Blvd, CA 91125-0001, USA*
[9] *Jet Propulsion Laboratory, 4800 Oak Grove Drive, Pasadena, CA 91109, USA*
[10] *Osservatorio Astronomico di Padova, Vicolo dell'Osservatorio, 5, I-35122 Padova, Italy*
[11] *National Radio Astronomy Observatory, PO Box O, Socorro, NM 87801, USA*
[12] *Astrophysics Group, Blackett Laboratory, Imperial College, Prince Consort Rd., London SW7 2BW*
[13] *Department of Physics & Astronomy, University of British Columbia, 6224 Agricultural Road, Vancouver, British Columbia V6T 1Z1, Canada*
[14] *Department of Physics & Astronomy, University of Pennsylvania, 209 South 33rd Street, Philadelphia, PA 19104-6396, USA*
[15] *Infrared Processing Analysis Center, California Institute of Technology, Jet Propulsion Laboratory, Pasadena, CA 91125, USA*
[16] *Institut d'Estudis Espacials de Catalunya, IEEC/CSIC, c/Gran Capita 2-4, 08034 Barcelona, Spain*
[17] *Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA*
[18] *Experimental Cosmology Lab, Department of Physics, University of Miami, 1320 Campo Sano Drive, Coral Gables, FL 33146, USA*
[19] *UK ATC, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ*
[20] *Joint Astronomy Centre, 660 N. A'ohōkū Place, University Park, Hilo, HI 96720, USA*
[21] *Institute for Astrophysics, University of Innsbruck, Technikerstr. 25, A-6020 Innsbruck, Austria*
[22] *Department of Astronomy and Astrophysics, University of Toronto, 60 St. George Street, Toronto, Ontario M5S 3H8, Canada*
[23] *Astronomy Centre, University of Sussex, Falmer, Brighton BN1 9QH*
[24] *Osservatorio Astrofisico di Arcetri, Largo E. Fermi 5, I-50125 Firenze, Italy*
[25] *Mullard Space Science Laboratory (MSSL), University College London, Holmbury St. Mary, Dorking, Surrey RH5 6NT*
[26] *Institute of Space and Astronautical Science (ISAS), Yoshinodai 3-1-1, Sagamihara, Kanagawa 229 8510, Japan*
[27] *Department of Physics, Astronomy & Mathematics, University of Hertfordshire, College Lane, Hatfield, Hertfordshire AL10 9AB*
[28] *Subaru Telescope, National Astronomical Observatory of Japan, 650 North A'ohoku Place, Hilo, HI 96720, USA*
[29] *Department of Physics & Astronomy, University of California Irvine, 4129 Frederick Reines Hall, Irvine, CA 92697-4575, USA*
[30] *Osservatorio Astronomico di Trieste, Via Tiepolo 11, I-34131 Trieste, Italy*
[31] *Department of Physics, University of Durham, South Road, Durham DH1 3LE*
[32] *Department of Physics, Brown University, 182 Hope Street, Box 1843, Providence, RI 02912, USA*
[33] *Department of Physics & Astronomy, University of Leicester, Leicester LE1 7RH*
[34] *Herzberg Institute of Astrophysics, National Research Council, 5071 West Saanich Road, Victoria, British Columbia V9E 2E7, Canada*

⋆E-mail: amjb2@star.kent.ac.uk

Accepted 2005 July 22. Received 2005 June 9; in original form 2004 September 6

ABSTRACT

The Submillimetre Common-User Bolometer Array (SCUBA) Half-Degree Extragalactic Survey (SHADES) is a major new blank-field extragalactic submillimetre (submm) survey currently underway at the James Clerk Maxwell Telescope (JCMT). Ultimately, SHADES aims to cover half a square degree at 450 and 850 μm to a 4σ depth of $\simeq 8$ mJy at 850 μm. Two fields are being observed, the Subaru/*XMM–Newton* Deep Field (SXDF) ($02^{\rm h}18^{\rm m} - 05^{\circ}$) and the Lockman Hole East ($10^{\rm h}52^{\rm m} + 57^{\circ}$). The survey has three main aims: (i) to investigate the population of high-redshift submm galaxies and the cosmic history of massive dust-enshrouded star formation activity; (ii) to investigate the clustering properties of submm-selected galaxies in order to determine whether these objects could be progenitors of present-day massive ellipticals; and (iii) to investigate the fraction of submm-selected sources that harbour active galactic nuclei. To achieve these aims requires that the submm data be combined with cospatial information spanning the radio-to-X-ray frequency range. Accordingly, SHADES has been designed to benefit from ultra-deep radio imaging obtained with the Very Large Array (VLA), deep mid-infrared observations from the *Spitzer* Space Telescope, submm mapping by the *Balloon-borne Large Aperture Submillimetre Telescope* (*BLAST*), deep near-infrared imaging with the United Kingdom Infrared Telescope, deep optical imaging with the Subaru Telescope and deep X-ray observations with the *XMM–Newton* observatory. It is expected that the resulting extensive multiwavelength data set will provide complete photometric redshift information accurate to $\delta z \lesssim 0.5$ as well as detailed spectral energy distributions for the vast majority of the submm-selected sources. In this paper, the first of a series on SHADES, we present an overview of the motivation for the survey, describe the SHADES survey strategy, provide a detailed description of the primary data-analysis pipeline and demonstrate the superiority of our adopted matched-filter source-extraction technique over, for example, Emerson-II style methods. We also report on the progress of the survey. As of 2004 February, 720 arcmin^2 had been mapped with SCUBA (about 40 per cent of the anticipated final total area) to a median 1σ depth of 2.2 mJy per beam at 850 μm (25 mJy per beam at 450 μm), and the source-extraction routines give a source density of 650 ± 50 sources deg^{-2} $> 3\sigma$ at 850 μm. Although uncorrected for Eddington bias, this source density is more than sufficient for providing enough sources to answer the science goals of SHADES, once half a square degree is observed. A refined reanalysis of the original 8-mJy survey Lockman hole data was carried out in order to evaluate the new data-reduction pipeline. Of the 17 most secure sources in the original sample, 12 have been reconfirmed, including 10 of the 11 for which radio identifications were previously secured.

Key words: galaxies: evolution – galaxies: formation – galaxies: starburst – cosmology: observations – infrared: galaxies – submillimetre.

1 INTRODUCTION

Theories of galaxy formation and evolution, embedded within hierarchical structure-formation models, can describe many of the observed features of galaxies (Cole et al. 2000; Granato et al. 2000; Hatton et al. 2003). While local galaxies can in some cases still provide constraints on the high-redshift populations (Panter, Heavens & Jimenez 2003; Heavens et al. 2004), the bulk of the constraints on models of galaxy evolution come either from the integral constraint from the far-infrared background (e.g. Dwek et al. 1998; Gispert, Lagache & Puget 2000, and references therein), or directly from high-redshift galaxy surveys (e.g. Steidel et al. 1999), the most ground-breaking of which were the Canada–France Redshift Survey (Lilly et al. 1995) and *Hubble Deep Field* North (HDF-N) (Williams et al. 1996).

Such optical surveys have led to a great deal of progress in understanding the assembly of stellar populations, and hierarchical galaxy-formation models are in increasingly good agreement with many (but not all) of these observations (e.g. Guiderdoni et al. 1998; Blain 1999; Kauffmann et al. 1999; van Kampen, Jimenez & Peacock 1999; Cole et al. 2000; Somerville, Primack & Faber 2001).

However, the discovery of a substantial population of faint submillimetre (submm) galaxies (Smail, Ivison & Blain 1997; Barger et al. 1998; Hughes et al. 1998) has posed serious problems for

the current generation of galaxy-formation models based on hierarchical structure growth. Models of the optical/ultraviolet (UV) spectral energy distributions (SEDs) of the galaxy population in the HDF-N (Thompson, Weymann & Storrie-Lombardi 2001) predict only submilliJansky/μJy-level 850-μm flux densities, but the submm point sources in the HDF-N field have 850-μm flux densities of several mJy (Hughes et al. 1998; Borys et al. 2003; Serjeant et al. 2003a). This shows that there is a population of star-forming galaxies that are heavily obscured by dust and have much higher infrared luminosities than would be inferred from the optical/UV observations alone. If these galaxies are at high redshifts (as current data imply), and if their emission is powered by star formation with a standard solar neighbourhood initial mass function (IMF), then their observed 850-μm flux densities of several mJy imply star formation rates $\sim$1000 $M_\odot$ yr^{-1}. Moreover, the far-infrared (FIR) luminosity density implied by the submm galaxy population suggests that these infrared-luminous galaxies contributed several tens of per cent of the volume-averaged star formation density at $z \simeq 2$ (e.g. Smail et al. 1997; Barger, Cowie & Sanders 1999).

One attractive interpretation of the submm galaxy population is that these violently star-forming galaxies are progenitors of present-day massive ellipticals (Hughes et al. 1998; Scott et al. 2002). There are four main pieces of evidence in support of this (e.g. Dunlop 2002). First, the star formation rates inferred from the submm flux densities are sufficient to construct the stellar population of even the most-massive elliptical galaxy in $\sim$1 Gyr; secondly, the K-band morphologies of submm-selected galaxies resemble those of radio galaxies, which locally are hosted in giant ellipticals (e.g. Lutz et al. 2001); thirdly, the comoving number density of bright submm sources in the redshift interval $z \simeq 2$–3 is comparable with the present-day number density of bright $>$2–3 $L^\star$ ellipticals (Scott et al. 2002); and fourthly, tentative detections of clustering suggest that the submm galaxies trace the collapse of rare, high density overdensities at high redshift (Almaini et al. 2003; Greve et al. 2004; Blain et al. 2004b). Furthermore, the high dynamical masses suggested by CO observations imply massive systems, and the gas masses implied by the CO luminosities suggest extensive star formation (Genzel et al. 2003); also, the dynamical gas and stellar masses estimated in the rest-frame optical/UV for these galaxies indicate that they are both massive and gas-rich and already contain significant stellar population (Smail et al. 2004; Swinbank et al. 2004).

A complication to this interpretation is that theory suggests a less-direct relation between local galaxies and their high-redshift antecedents. Cold dark matter (CDM) simulations inevitably predict that massive galaxies today are assembled hierarchically from a large number of smaller fragments that existed at high redshift. Conversely, the majority of the rare early collapsing objects at $z \simeq 2$–3 should be found inside massive galaxies at the present (Baugh et al. 1998). If this is true, then it suggests a more detailed set of questions that new, larger submm surveys should attempt to settle: (i) what fraction of present-day massive ellipticals have merged with at least one Submillimetre Common-User Bolometer Array (SCUBA) galaxy? (ii) what fraction of SCUBA galaxies will end-up in a present-day massive elliptical? and (iii) if the answer to the second question is close to 100 per cent, how close are the SCUBA galaxies to the end of the merger process? There are claims (Bell et al. 2004) that the total mass of stars in ellipticals has roughly doubled since redshift $z = 1$. If this increase applies to the most luminous ellipticals, this would mean that most present ellipticals were only in the earliest phases of assembly at $z = 2$–3. There is thus some uncertainty over whether a given SCUBA galaxy represents a late phase in the construction of an elliptical, or an early phase in the assembly of its nucleus. Uncertainties in the true star formation rate and the possible lifetime of the starburst contribute to this uncertainty, which can only be addressed statistically by looking at the population as a whole.

Detailed models of the hierarchical assembly of galaxies based on standard prescriptions for gas cooling, star formation and feedback, in general, do not predict large numbers of SCUBA galaxies with star formation rates $\sim$1000 $M_\odot$ yr^{-1}. However, Baugh et al. (2005) have shown that the observed number counts and redshift distributions of the submm sources can be reproduced in CDM models if the star formation in these objects is occurring with a top-heavy IMF (e.g. Larson 1998), implying that the same submm flux can be produced with star formation rates $\sim$5 times lower than for a Salpeter IMF. An alternative model that can explain the submm galaxies within the framework of CDM is that of Granato et al. (2004), who propose modifications to the treatments of gas cooling and feedback, as opposed to modifications to the IMF. Finally, the treatment of virialization and the survival of subhaloes in the model of van Kampen et al. (1999) produces star formation histories, which allow for much higher star formation rates at early times (especially in bursts), and therefore predicts sufficient number of submm galaxies at high redshifts (van Kampen 2003). The SCUBA Half-Degree Extragalactic Survey (SHADES) will provide the means to distinguish between these and other alternatives (van Kampen et al. 2005).

Of course, the radical possibility that the hierarchical orthodoxy may be flawed in some way is worth keeping open, in order that the standard theory can be properly verified. Furthermore, even the interpretation of all SCUBA sources as dusty starbursts could still be subject to future revision. Although the majority of the bright submm sources have secure extragalactic identifications, and in many cases redshifts, it remains possible (albeit increasingly unlikely) that a fraction of submm objects could be Galactic in origin (Sciama 2000; Lawrence 2001). Even accepting that the sources are extragalactic, the starburst model is not unchallenged: Efstathiou & Rowan-Robinson (2003) and Kaviani, Haehnelt & Kauffmann (2003) have successfully modelled the emission of some submm sources as extended cirrus in galaxies heated by the interstellar radiation field. This is possible partly because the far-infrared colour temperatures are not well constrained with existing data. This interpretation is supported by observational evidence from Farrah et al. (2004), using submm detections of a galaxy at $z \simeq 0.5$ that hosts a Type Ia supernova. The rest-frame optical luminosities and colours of submm galaxies require such cirrus galaxies to be more heavily extinguished than their local counterparts – Efstathiou & Rowan-Robinson find a best-fitting extinction coefficient of $A_V \simeq 1$–3 for the high-redshift galaxies as opposed to $A_V \simeq 0.4$–0.9 for the local galaxies in their sample. Such an interpretation may be testable using high-resolution observations with the Plateau de Bure Interfermometer, and will be conclusively answered using the Atacama Large Millimeter Array (ALMA) with a spatial resolution of $\lesssim$0.1 arcsec.

At present, the determination of the redshift distribution and clustering of the submm galaxy population offers the best available method to constrain the properties of the overdensities hosting bright submm galaxies, and hence to differentiate between alternative models of galaxy evolution (van Kampen et al. 2005). To this end, the SHADES (http://www.roe.ac.uk/ifa/shades/) consortium is mapping 0.5 deg^2 with the SCUBA at the James Clerk Maxwell Telescope (JCMT) on Mauna Kea, Hawaii. SHADES expects to produce a complete sample of $\sim$300 bright submm sources at 850 μm.

The survey is being carried out in fields with abundant supporting multiwavelength data (see Section 3.1 for full details).

In this paper, the first in a series of papers on SHADES, we present an overview of the motivation for the survey, discuss the adopted observing strategy and describe the operation of, and first results from the primary SHADES data-reduction pipeline. However, we stress that three additional and independent reductions of the SHADES data are currently underway within the consortium, and that we aim to take advantage of these multiple reductions to maximize the robustness of the final 850-μm source list. These alternative reductions and the outcome of cross-referencing the resulting maps and source lists will be presented elsewhere. The current paper is structured as follows. Section 2 describes the aims of the SHADES survey in more detail and Sections 3 and 4 present the data acquisition and analysis methods being used. Section 5 outlines the source-extraction methods under development for this large survey; in Section 6, we present the progress of the survey so far and Section 7 concludes the paper. Throughout, we assume a cosmology with $\Omega_{\rm M} = 0.3$, $\Omega_\Lambda = 0.7$ and a Hubble constant of $H_0 = 72\,{\rm km\,s^{-1}\,Mpc^{-1}}$.

2 MOTIVATION FOR THE SURVEY

2.1 Background: the problem of cross-identifications

The faint optical/near-infrared identifications of submm galaxies (e.g. Lilly et al. 1999; Barger, Cowie & Richards 2000; Clements et al. 2004; Ivison et al. 2002; Smail et al. 2002; Serjeant et al. 2003a,b; Webb et al. 2003; Borys et al. 2004; Dunlop et al. 2004; Wang, Cowie & Barger 2004) and the broad (∼10–15 arcsec) beams of the largest current submm/millimetre-wave telescopes together present difficulties for the unambiguous identification of submm galaxies (Hughes et al. 1998).

However, extensive long-term efforts towards identification made in the radio (e.g. Ivison et al. 2004), supported in some cases by millimetre-wave interferometry (e.g. Downes et al. 1999; Gear et al. 2000; Lutz et al. 2001), have produced radio identifications for ∼50–70 per cent of the brighter submm sources. These radio detections have been successfully exploited to derive accurate (i.e. subarcsec) positions for submm galaxies, thus facilitating further spectroscopic study. As a result, spectroscopic redshifts for approximately 90 radio-detected submm sources have been published to date (Ivison et al. 1998, 2000; Chapman et al. 2003, 2005).

The spectroscopic follow-up is of course biased against those sources at redshifts where no spectral features fall within the spectroscopic range, most notably $1.2 < z < 1.7$, and may also exclude the highest-redshift objects (due to the less favourable K-correction in the radio waveband compared to the submm for redshifts $z \gtrsim 3$). Together these effects result in moderate incompleteness in the final redshift surveys (Chapman et al. 2005) even if all identifications are robust.

Recently, rapid detections of SCUBA galaxies have been made with the *Spitzer* Space Telescope in integrations of only ∼10 min (Charmandaris et al. 2004; Egami et al. 2004; Frayer et al. 2004; Ivison et al. 2004; Serjeant et al. 2004). This small sample of identifications in conjunction with the abundant *Spitzer* coverage of our fields shows the potential for identification and follow-up of the submm sources.

Despite the radio and the *Spitzer* data, some sources may remain unidentified. Lack of a robust radio or mid-infrared indentification could have five origins: (i) the submm source could be spurious; (ii) the source could be severely flux boosted (Eddington 1913, see Section 6.2); (iii) the radio/far-infrared (IR) emission could be significantly more extended than the Very Large Array (VLA)-synthesized beam; (iv) the characteristic dust temperatures could be low; or (v) the source could be at very high redshift.

Nevertheless, armed with lower-resolution observations with the VLA along with the low-resolution submm imaging to be provided by the *Balloon-borne Large Aperture Submillimetre Telescope* (*BLAST*) (Hughes et al. 2002, see Section 3.1) and mid-infrared imaging from the *Spitzer* Wide-Area Infrared Extragalactic Legacy Survey (SWIRE) (Lonsdale et al. 2003, 2004, see Section 3.1), we anticipate being able to distinguish between these five alternatives for most of the apparently unidentified sources.

The large positional uncertainty of the SCUBA sources may also lead to unreliable identifications. A measure of this is the cross-identification limit, which we define to be one random source per 10 search areas. At this surface density, the likelihood of a spurious identification is given by $p \lesssim 1 - \exp^{-N(>S)\pi r^2} \simeq 0.1$, where r is 7 arcsec (the half width at half-maximum of the JCMT beam at 850 μm) and $N(>S)$ is the cumulative source counts of other objects in the identification catalogue with fluxes greater than S. For the *Spitzer* Infrared Array Camera (IRAC) 3.6- and 4.5-μm bands, the resulting cross-identification limits are 58 and 60 μJy, respectively, well above the SWIRE sensitivity and confusion limits (shown in Table 1). For the IRAC 5.8- and 8.0-μm and *Spitzer* Multiband Imaging Photometer for *Spitzer* (MIPS) 24-μm bands, the cross-identification limits are 48, 44 and 120 μJy, similar to the SWIRE sensitivities but still well above the confusion limit (Fazio et al. 2004; Marleau et al. 2004). The 70- and 160-μm cross-identification limits of 0.93 and 9.1 mJy are well below the confusion and SWIRE sensitivity levels (Dole et al. 2004). This highlights the need for careful analyses to identify the SHADES galaxies at wavelengths shorter than 24 μm. Our abundant multiwavelength coverage, especially optical and near-infrared, can be used to find objects by looking at outliers in colour space (e.g. Smail et al. 2002; Webb et al. 2003; Pope et al. 2005).

To illustrate the potential power of the multifrequency/multifacility approach adopted for SHADES, we have investigated the properties of four template galaxies, whose SEDs have been normalized to the SHADES survey depth of 8 mJy at 850 μm, as the assumed redshift is varied. Table 1 shows the flux densities expected from our four template galaxies. Note the similarity of the SWIRE limits to the Arp 220 model SED flux densities and also to the depth required for reliable identifications. Fig. 1 shows the four example model SEDs as the assumed redshift is varied from $z = 0$ to $z = 4$. The key points illustrated by these models can be summarized as follows.

(i) The Arp 220 SED with the SHADES normalization shows that BLAST detections of Arp 220-like SEDs should be possible to $z \simeq 2$–3 and that detections at 3.6 and 4.5 μm in conjunction with non-detections at 5.8, 8 and 24 μm can be used to differentiate Arp 220-like SEDs from other forms. The relative number counts of Huang et al. (2004) show that higher wavelength drop-outs do occur; Le Floc'h et al. (2004) using preliminary *Spitzer* MIPS (24 μm) results suggest that the SEDs of high-redshift ($z \geqslant 1$) sources in the Lockman Hole East and the Extended Groth Strip are well fitted by an Arp 220-like model.

(ii) M82 is another prototypical star-forming galaxy and galaxies with M82-like SEDs should be visible at BLAST depths out to $z \simeq 3$ and *Spitzer* SWIRE depths out to $z \simeq 4$ using the IRAC (3.6–8 μm) wavebands.

Table 1. Comparison of the *Spitzer* survey flux-density limits with galaxy populations normalized to the SHADES depth of 8 mJy at 850 μm. Also listed are the effective cross-identification limits (as described in the text).

Wavelength	3.6 μm	4.5 μm	5.8 μm	8.0 μm	24 μm	70 μm	160 μm
SWIRE 5σ in 500 s[a]	3.7 μJy	5.3 μJy	48 μJy	37.7 μJy	106 μJy	6 mJy[a]	50 mJy[a]
$S_{\rm conf}$[b]	1.5 μJy[d]	1.5 μJy[d]	8.3 μJy[d]	5.4 μJy[d]	88 μJy[e]	5.3 mJy[f]	48 mJy[f]
$N_{\rm src}$[c]	1.7[d]	1.4[d]	0.84[d]	0.81[d]	0.13[e]	0.009[f]	0.003[f]
ARP 220 $z = 2^{\star}$	22 μJy	23 μJy	26 μJy	17 μJy	0.21 mJy	3.7 mJy	62 mJy
ARP 220 $z = 4^{\star}$	3.5 μJy	5.2 μJy	6.1 μJy	6.9 μJy	6.0 μJy	0.17 mJy	4.3 mJy
NGC 1068 $z = 2^{\star}$	9.6 μJy	22 μJy	55 μJy	0.16 mJy	3.3 mJy	16 mJy	32 mJy
NGC 1068 $z = 3^{\star}$	2.6 μJy	3.9 μJy	10 μJy	32 μJy	0.86 mJy	5.4 mJy	12 mJy
M82 $z = 2^{\star}$	74 μJy	0.11 mJy	0.15 mJy	0.11 mJy	1.4 mJy	9.6 mJy	58 mJy
M82 $z = 4^{\star}$	4.2 μJy	10 μJy	19 μJy	36 μJy	38 μJy	1.1 mJy	5.3 mJy
HR10 $z = 2^{\star}$	71 μJy	0.12 mJy	0.15 mJy	0.11 mJy	0.36 mJy	6.6 mJy	44 mJy
HR10 $z = 4^{\star}$	2.4 μJy	6.6 μJy	19 μJy	39 μJy	20 μJy	0.28 mJy	5.0 mJy

[a]SWIRE limits from Lonsdale et al. (2004). The SWIRE team is currently reassessing the observing strategy due to the degraded MIPS sensitivities. The values shown here for 70- and 160-μm sensitivities are pre-flight estimates only. [b]Confusion limit from observed source counts, of one source per 40 SCUBA beams (7-arcsec radius circle). [c]Number of observed sources per SCUBA 850-μm beam (7-arcsec radius circle) greater than the SWIRE limit. [d]Number counts from Fazio et al. (2004). [e]Number counts from Marleau et al. (2004). [f]Number counts from Dole et al. (2004).

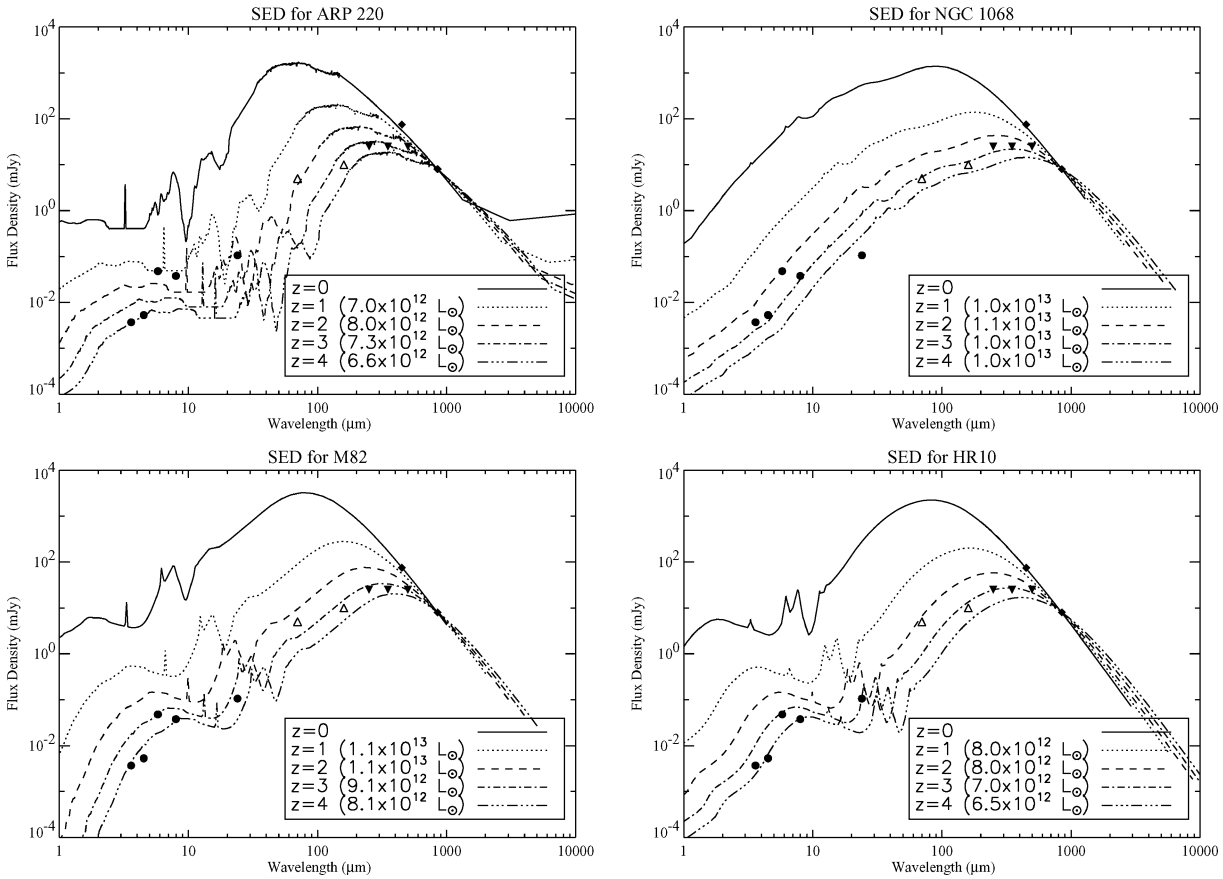

Figure 1. Model spectral energy distribution (SED) plots of template galaxies with survey sensitivities overplotted. Filled circles: SWIRE sensitivities 500 s, 5σ (Lonsdale et al. 2004). Open triangles: the pre-flight *Spitzer* sensitivities for 70 and 160 μm are shown whilst the SWIRE team reassesses its observing strategy due to the degraded MIPS sensitivities, although currently these are expected to be two to three times worse. Filled triangles: the pre-flight BLAST confusion limit estimates [derived from Rowan-Robinson (2001)] of 25 mJy; the sensitivity estimates (Hughes et al. 2002) are in the range 15–25 mJy. Filled diamonds: SHADES survey sensitivities of 8 mJy at 850 μm and 75 mJy at 450 μm. Models have been normalized to the SHADES survey depth of 8 mJy at 850 μm. Top left-hand panel: Arp 220 – a highly obscured local starburst galaxy [note: it is also a ULIRG Elbaz et al. (2002)]. Top right-hand panel: NGC1068 – a typical Seyfert galaxy (Efstathiou, Hough & Young 1995). Bottom left-hand panel: M82 – an irregular dusty star-forming galaxy (Efstathiou et al. 2000). Bottom right-hand panel: HR10 – a high-redshift and extremely-red galaxy (Takagi, Hanami & Arimoto 2004). Bolometric (3–1000 μm rest frame) luminosities for each model, as normalized, are shown in the legends.

(iii) The SED of HR10, a high-redshift and extremely red galaxy, is well matched out to redshifts of $z \simeq 2$–3 with BLAST and $z \simeq 3$–4 at shorter wavelengths. HR10 would be defined as an ultraluminous infrared galaxy (ULIRG) by its infrared luminosity.

(iv) The SED of NGC 1068, dominated by an AGN-heated dusty torus, is discussed in more depth in Section 2.2.3

These models illustrate that the mid-infrared and far-infrared flux-density limits for the coverage of the SHADES fields will detect submm galaxies to redshifts $z \lesssim 4$, comparable to that from the deepest radio integrations. We can confirm these expectations using the properties of existing SCUBA galaxies, which have been detected at 24 μm (Egami et al. 2004; Frayer et al. 2004). These have been found to have 24-μm flux densities ranging from 80 to 2800 μJy, although the median flux is 230 μJy and half of the detections lie in the range 160–370 μJy. Such measurements compare favourably with the anticipated SWIRE 5σ survey limit of 105 μJy at 24 μm (Lonsdale et al. 2004). The deeper *Spitzer* Guaranteed Time Observations (GTO) data may also detect many of the submm galaxies at other wavelengths.

2.2 SHADES science goals

2.2.1 Measurement of the cosmic history of massive dust-enshrouded star formation activity

A key constraint in the phenomenology of galaxy formation is the cosmic history of dust-enshrouded star formation, observed via the evolution of far-infrared luminosity density. The flat submm selection function is well known to give submm galaxy surveys an informative constraint on the far-infrared luminosity density throughout most of the Hubble Volume (e.g. Franceschini et al. 1991; Blain & Longair 1996). Spectroscopic redshifts or photometric redshift estimates are clearly essential to constrain the evolving far-infrared luminosity density, but the identifications have proved challenging (see above). Moreover, even when secure identifications are available, the optical or near-infrared spectroscopic follow-up observations are time-consuming, typically requiring >2 h of integration on each target with 8–10 m class telescopes, and are not guaranteed of success (e.g. Chapman et al. 2003; Simpson et al 2004).

However, even in the absence of an optical or near-infrared identification, photometric redshifts can still be derived on the basis of long-wavelength (i.e. radio to far-infrared) photometry. The simplest of these methods uses a single colour ratio between observations at 850 μm and 1.4 GHz (Carilli & Yun 1999; Dunne, Clements & Eales 2000; Rengarajan & Takeuchi 2001) to discriminate between low- and high-z star-forming galaxies. This method applied to a sample of 30 sources (Ivison et al. 2002) produces estimated redshifts accurate to $\delta z \simeq 0.4$–1.5.

Despite the dust temperature–redshift degeneracy present in far-infrared (FIR) sources (Blain 1999; Blain, Barnard & Chapman 2003), extension of this method to include multiple colour information in the radio to submm bands allows the derivation of photometric redshifts with accuracies of $\delta z \simeq 0.5$ or better (Hughes et al. 2002; Yun & Carilli 2002; Aretxaga et al. 2003; Wiklind 2003; Aretxaga, Hughes & Dunlop 2005), even taking into account the full range of dust properties present in the local galaxy analogues. The most recent comparison of photometric and spectroscopic redshifts for blank-field submm sources yields a dispersion of $\delta z \sim 0.3$ when three or more long-wavelength detections are available (Aretxaga et al. 2005). Uncertainties due to the temperature–redshift degeneracies could be further reduced by performing spectroscopic calibration upon a representative subset of the sample (Blain et al. 2004a).

For those sources for which near-infrared or mid-infrared identifications are secured, a complementary (and in principle independent) method of estimating redshifts is to use the *Spitzer* Space Telescope and corresponding ground-based near-infrared observations to identify the position of the redshifted 1.6-μm peak of the near-infrared stellar continuum using the IRAC 3–8 μm bands (Simpson & Eisenhardt 1999; Sawicki 2002). This method gives an indication of which objects lie at redshift $z \gtrsim 1.5$ and obtains redshift estimates accurate to $\delta z \sim 0.5$ at $z \lesssim 1.5$ independent of the BLAST and VLA measurements. Where optical photometry is available, the photometric redshift is more accurate still (Pope et al. 2005).

Given the need to break the sample into a few redshift bins, while maintaining statistically useful numbers in each bin, we require of the order of 100 sources in total in order for the $N(z)$ histograms to be able to differentiate between available models (van Kampen et al. 2005). The combined multiwavelength data are expected to yield a redshift resolution of better than $\delta z \simeq 0.5$ for 300 galaxies, using spectroscopically calibrated photometric redshifts (Aretxaga et al. 2005).

2.2.2 Determination of whether SCUBA galaxies are progenitors of present-day massive elliptical galaxies

The high inferred star formation rates of submm galaxies, while consistent with expectations for protoellipticals, nevertheless do not provide unambiguous evidence that they are progenitors of massive elliptical galaxies. This is both because the duration of the starburst is unknown and because the derived star formation rate is sensitive to the assumed temperature and IMF of the mass distribution (Larson 1998; Baugh et al. 2005). In contrast, the clustering of bright SCUBA sources on scales of up to ∼10 Mpc offers a potentially very powerful constraint on the nature of the submm population (Percival et al. 2003), given even a relatively broad ($\delta z \simeq 0.5$) constraint on the estimated redshifts of individual sources (van Kampen et al. 2005). As discussed in van Kampen et al. (2005), to sample the submm population over such scales requires a degree-scale survey (1° is 29 Mpc at $z = 2$ in our adopted cosmology). Therefore, the desire to probe scales approaching 10 Mpc provides one of the primary motivations for the eventual areal coverage goal of SHADES (i.e. half a square degree).

Previous blank-sky surveys carried out with SCUBA have identified a total of ∼100 sources in disparate areas on the sky. These survey fields include the HDF (Hughes et al. 1998; Borys et al. 2003; Serjeant et al. 2003a; Borys et al. 2004; Wang et al. 2004), the Hawaii deep-field regions (Barger et al. 1998, 1999), the Lockman Hole East and the Elais N2 region in the 8-mJy survey (Fox et al. 2002; Scott et al. 2002), the *Spitzer* northern continuous viewing zone (Sawicki & Webb 2005) and the Canada–France Redshift Survey (CFRS) fields by the Canada–United Kingdom Deep Submillimeter Survey (CUDSS) (Eales et al. 1999). However, the resulting composite existing 'sample' of submm sources spans a wide range in intrinsic luminosity and is distributed between many small fields imaged to varying depths. As a result, it has proved impossible to derive unambiguous constraints from the apparent clustering strength of these sources (Scott et al. 2002; Borys et al. 2003). By providing a complete and homogeneous sample of the most luminous submm sources in two wide-area fields, SHADES aims

to provide the first robust constraints of the clustering properties of the submm galaxy population.

2.2.3 Determination of the fraction of SCUBA sources that harbour obscured active galactic nuclei

The *Spitzer* 3.6–160 μm data from the SWIRE and GTO surveys (Lonsdale et al. 2003, 2004; Egami et al. 2004; Huang et al. 2004) sample the SEDs in the rest-frame near-infrared and mid-infrared. The latter is sensitive to the presence of active galactic nuclei (AGN) dust tori, making the *Spitzer* data key to determining the AGN bolometric fraction in these high-infrared luminosity galaxies (e.g. Almaini, Lawrence & Boyle 1999; Efstathiou, Rowan-Robinson & Siebenmorgen 2000; Farrah et al. 2002). Most torus models show warm colour temperatures, which therefore would not contribute significantly to the submm flux.

X-ray visible AGN and submm sources are only rarely coincident in shallow X-ray observations (Bautz et al. 2000; Fabian et al. 2000; Almaini et al. 2003; Waskett et al. 2003) and yet seem to trace similar structures on arcminute scales in the Elais N2 field (Almaini et al. 2003). Possibly, these two populations represent different and relatively short-lived phases in the formation of massive objects at high redshift. The majority of SCUBA sources may contain a massive and active black hole that is too heavily obscured to be detected with the current X-ray surveys. Such Compton-thick objects may be associated with the formation of super-massive black holes. Some X-ray objects have been found to have a submm source associated with them (e.g. Barger et al. 2001) and Alexander et al. (2003, 2005) suggest that even those submm galaxies without an X-ray counterpart could contain low-luminosity AGN. However, as Alexander et al. point out, the submm emission in these galaxies would be dominated by star formation and not the AGN, so any estimate of star formation rates would not be affected.

The NGC 1068 SED in Fig. 1, normalized to the SHADES sensitivity, demonstrates that hyperluminous AGN dust tori in the SHADES survey, with the AGN dominating even the submm flux, should be detectable in the *Spitzer* 24–160 μm and BLAST 250–500 μm bands at redshifts of up to $z = 2$. In this extreme limiting case, the SED is dominated by the dust torus with little contribution from circumnuclear star formation. Ivison et al. (2004) and Egami et al. (2004) have already shown that *Spitzer* 24-μm photometry is efficient at demonstrating the presence of AGN in cases where the AGN makes a much smaller bolometric contribution, compared to that of the star formation; this AGN detection is particularly effective when combined with photometry in the IRAC bands (Ivison et al. 2004).

3 DATA ACQUISITION

3.1 The survey fields

The half square degree to be covered by SHADES is split between two survey fields – the Lockman Hole East (field centre approximately $10^{\rm h}52^{\rm m}28^{\rm s}$ $+57^\circ22'20''$) and the Subaru/*XMM–Newton* Deep Field (SXDF) ($02^{\rm h}18^{\rm m}00^{\rm s}$ $-$ $05^\circ00'00''$). The fields were chosen for their low Galactic cirrus (100-μm surface brightness $\sim$1 MJy sr^{-1}) and benefit from abundant cospatial multiwavelength data. The declination of the source fields and their spread in right ascension (RA) make these fields observable for the majority of the year using the JCMT as well as being accessible to BLAST, *Spitzer*, the VLA, UKIRT, Subaru, Keck and Gemini Telescopes. The decision to include an equatorial field was partly driven by the desire to provide a submm source catalogue accessible to ALMA.

Observations in the Lockman Hole East are being extended around the 151 arcmin^2 observed as part of the SCUBA 8-mJy survey (Scott et al. 2002). The direction and shape of this extension are driven primarily by the existence of ultra-deep VLA imaging at 1.4 GHz.

The specific choice of the SXDF as the equatorial field was motivated by the existence of deep *XMM–Newton* imaging, associated deep VLA observations at 1.4 GHz and existing deep multicolour optical imaging obtained with the Subaru Telescope.

The near-infrared (J, H, K) imaging for both fields will be provided by the new UK Infrared Telescope Wide Field Camera (UKIRT WFCAM) instrument as part of the United Kingdom Infrared Deep Sky Survey (UKIDSS). Specifically, the Ultra Deep Survey (UDS) component of UKIDSS will cover 0.77 deg^2 to $K \simeq 23$ in the SXDF, while the Deep Extragalactic Survey (DXS) in UKIDSS will map the SHADES Lockman field to $K \simeq 21$.

Further multiwavelength coverage comes from BLAST (500, 350 and 250 μm) surveys in our fields and two *Spitzer* surveys – the SWIRE Legacy Survey (Lonsdale et al. 2003, 2004) and the GTO data (Egami et al. 2004; Huang et al. 2004). The BLAST survey is expected to be confusion limited at all wavelengths, corresponding to a 5σ flux density at the confusion limit of approximately 25 mJy in all three bands (derived from Rowan-Robinson 2001). The SWIRE survey flux density limits are shown in the first row of Table 1. Various pointed follow-ups are also underway including using the Submillimetre High Angular Resolution Camera II (SHARC-II) at the Caltech Submillimeter Observatory (CSO).

The BLAST 500-μm data are expected to be deeper than the SCUBA 450-μm data because the SCUBA observations are deliberately restricted to be conducted in only grade 2–3 weather ($\tau_{225\,\rm GHz} \simeq 0.05$–$0.10$; this restriction was adopted to allow other, smaller programmes on the JCMT to best exploit the rare, grade-1 conditions). Because the 450-μm SCUBA data will be of only moderate quality, we expect few sources to be detected in the 450-μm maps. Therefore, in this paper, we focus primarily on the 850-μm data.

3.2 SCUBA technical information

SCUBA is composed of two arrays of bolometers that view the same region of sky simultaneously, a long wave array of 37 bolometers used at 850 μm and a short wave array of 91 detectors used at 450 μm. The pixels are arranged in a hexagonal pattern with the feedhorns close-packed. The bolometer performance improves with decreasing temperature, so SCUBA has a helium cooling system to improve detector sensitivity. Thermal noise from the sky and local surroundings dominates at millimetre wavelengths.

The SHADES survey is conducted in jiggle-mapping mode. The SCUBA bolometers instantaneously undersample the sky but are dithered in a 64-point pattern to ensure that, overall, the sky is sampled at the Nyquist frequency or better at both wavelengths. Further details on SCUBA can be found in Holland et al. (1999).

Both the terrestrial atmosphere and thermal emission from the telescope contribute to a strong background ($\sim$1 Jy arcsec^{-2}) and the atmospheric part of this emission varies rapidly. By rapidly chopping the secondary mirror and nodding the entire telescope, these effects are reduced, leaving a residual atmospheric noise, which is common mode to and therefore uniform across the whole array of bolometers.

Jiggle-maps are coadded to improve the signal-to-noise ratio (S/N) allowing measurements of signals that are tens to hundreds of thousands of times fainter than the background.

3.3 The observing strategy

The survey makes use of close-packed hexagonal geometry to place jiggle-maps in an interleaved positioning scheme. This provides as uniform a noise level across the survey field as practicable, as well as ensuring that each sky position is covered by multiple bolometers. Fig. 2 shows one interleaved 'tripos' pattern, illustrating the central triangular region covered uniformly by the three hatched jiggle-maps.

The tripos system of three overlapping maps is observed for each of the six different chop-throw and chop position angle (PA) combinations. This is a significant departure from the previous survey strategy most similar to SHADES that adopted for the 8-mJy survey (Fox et al. 2002; Scott et al. 2002). We specifically chose the six chop throws motivated by the Emerson-II chop-throw methodology (Emerson 1995), that is, chop throws of 30, 44 and 68 arcsec at position angles (PAs) of zero and 90° in right ascension/declination coordinates. This choice of chop-throw and PA ensures that no Fourier mode larger than the beam and smaller than the largest chop-throw is lost entirely. The source extraction from these multiple chop-throw maps is discussed below (Section 5). By observing each of the chop-throw/PA combination maps at different airmasses, the noise levels in each tripos should ensure an even coverage. Thus, one chop-throw/chop PA combination will always be observed within a particular airmass range; while it would be possible to balance the airmasses evenly over the chops in order to avoid correlating the chop strategy with airmass, our adopted strategy makes observing decisions easy enough to maintain a low error rate. In addition, the jiggle-maps at each tripos position are observed in a priority order, which keeps the survey area approximately circular at any time, in order to minimize the perimeter area, which will have lower S/N.

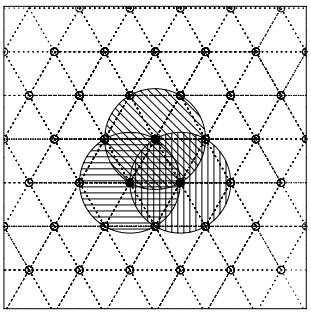

Figure 2. The 'tripos' positioning scheme. Individual jiggle-maps (2.6-arcmin diameter) are represented as large hatched circles; only three of these are plotted for clarity, but the centres of the others are marked with small open circles. This pattern of overlapping jiggle maps, in conjunction with observing different chop-throw/PA combinations at different airmasses, provides an approximately uniform coverage over the map. These three hatched jiggle maps, with centres marked by filled circles, comprise one 'tripos'. Only these three positions contribute to the data within the triple-hatched triangle formed by the filled circles. Any other triangular region is covered by its own combination of three jiggle-maps. The jiggle-map positions shown by the open circles are observed sequentially in a spiral manner from the centre of the map outwards; the resulting roughly circular areal coverage minimizes the perimeter, which has lower coverage and so lower S/N than the rest of the survey.

In the SXDF, the observing strategy has been to extend the region from the centre of the field in a spiral manner. The area of Lockman Hole East already covered by the 8-mJy survey has not been repeated, since the noise level of this area is already at the required depth. Therefore, as discussed above, the new SCUBA data have been extended around the existing 8-mJy Lockman Hole East data.

4 THE DATA-REDUCTION PROCESS

We discuss here the Interactive Data Language (IDL) reduction pipeline developed for SHADES from the original 8-mJy survey reduction process (Scott et al. 2002; Serjeant et al. 2003a). This process flat-fields the data, combines the data from the individual chops and nods and corrects for the effect of the atmosphere using an extinction correction routine. The pipeline also corrects for noisy bolometers and fast transient spikes in the data, for example, cosmic rays (see below). In order to extract spatial information, the jiggle-map data can be regridded on to rectangular coordinates. The improvements and variations on the 8-mJy survey pipeline made for the SHADES data are listed below; for further details of this reduction method see e.g. Serjeant et al. (2003a).

4.1 Extinction corrections

Sky absorption caused by water vapour in the atmosphere was previously removed from the data using smoothed CSO sky-dip measurements extrapolated to 850 and 450 μm [using the conversion relations given by Archibald et al. (2002)] or by SCUBA sky-dip measurements when insufficient CSO data were available. CSO tau sky monitors take sky-dip readings at 225 GHz every 10 min. SCUBA sky-dips have the benefit of being at the observation wavelength, but are taken at a different time to the observations. The new pipeline preferentially uses the water vapour radiometer data (WVM; Wiedner et al. 2001), available since 2000 July, which are taken at the same time and azimuth as the observation at a frequency of 183 GHz every 6 s. These WVM data have been converted into 225-GHz values, allowing for the elevation corrections, by Weferling (private communication). The WVM data are smoothed over 18 s – a time-scale similar to that of the nod – and applied to the data. We have undertaken a comparison of the WVM, SCUBA sky-dip and CSO sky-dip data, and find that they agree to within 10 per cent when converted to 225-GHz opacity. This corresponds to derived flux-density uncertainties of a few per cent. The sky extinction correction from the WVM has the advantage of being able to reliably fit variations, which occur over the course of a single map as well as over the course of the night. It is therefore used for all nights where the necessary data are available.

4.2 Bolometer and sky-noise analysis

In what follows, we refer to a single measurement (i.e. the average of eight chop cycles subtracted between the two nod positions) from a single bolometer as a 'readout'. Each readout represents approximately 2 s of integration time.

The noise in our data varies with time and between bolometers. As the noise was found to vary significantly on time-scales significantly longer that ∼100 readouts, we decided to measure the noise in 128-readout groups for each bolometer. This is achieved by fitting a Gaussian to the histogram of readouts of a 128-readout group. Outliers in these readout distributions at the $>3\sigma$ level are flagged as glitches. This filters out cosmic rays. Instrument systematics should be removed by the nodding technique. We find, nevertheless, that

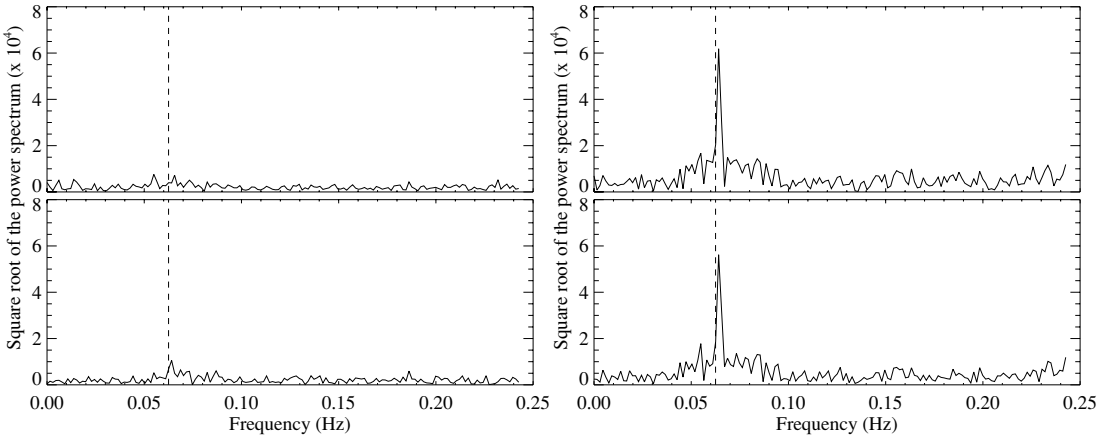

Figure 3. The spike in the power spectrum of the bolometer readouts. The left-hand column of plots shows the power spectrum of a typical (in this case the central) bolometer, before (top) and after (bottom) sky subtraction. The right-hand column of plots shows the same for a bolometer affected by the power spike, showing clearly the spike at a period of roughly 16 samples (at the dotted line) and the fact that the power spike itself has not changed.

there is still residual sky emission in the data, caused by the sky-level varying on time-scales less than that of the nod. This level is difficult to characterize, but appears to be common to all bolometers in the array, taking into consideration the noise level of individual bolometers. The average in the form of the mode of the data is found by fitting a Gaussian to the data from all bolometers for each of the long/short wavelength bolometer groups for each time-step. The DC-offset over the long/short array is then removed for each time-step. Sky levels are evaluated independently at each wavelength so that two flux readings can be taken at different wavelengths from the data. Higher-order terms of the sky level are not removed in this method as they have been found not to be significant.

A significant difference from the 8-mJy survey is the presence of a spike in the power spectrum of some of the bolometers at a period of roughly 16 samples, for dates from the end of 2002 onwards as seen in Fig. 3 (Borys et al. 2004; Webb et al. 2004). Tests have shown that this effect is only significant for our data between 2002 December and 2003 June. This effect presents difficulties for the sky-noise removal programme because the frequency is the same as that of the jiggle pattern itself. It has the effect of skewing the distribution of the sky noise, so that it is not correct to remove the same DC sky level for both the bolometers with the noise spike and those without. Where a significant number (greater than 10) of bolometers are affected in each map, the bolometers with and without the noise spike are treated as individual data sets for the sky-subtraction analysis. The process of bolometer-noise measurement and sky subtraction is then iterated.

In the method described in this paper, we have not attempted to remove the noise spike, but have simply chosen to isolate the affected bolometers in our sky-subtraction. Therefore, we do not knowingly allow the presence of the spike to distort the sky-subtraction process. As a result our assigned noise levels may not be optimal, although sources lost to the higher noise regions should be accounted for statistically using completeness simulations. An alternative noise-spike treatment, which seeks to remove the noise spike from the data time-stream, has been applied to the SHADES data, and the results are almost identical.

Implicitly these algorithms assume the S/N of point sources to be negligible in the bolometer readouts, which can readily be seen to be the case given the thousands of readouts that contribute to the 3–10σ sources in submm surveys. However, this is not true for calibrators, so for these the noise measurement and deglitching stages are replaced by assigning a noise value that consists of an arbitrary-magnitude noise equivalent flux-density (NEFD) value, which scales as the square root of the total integration time on each position on the sky (Archibald et al. 2002).

4.3 Flux calibration

Calibration maps of Mars, Uranus and secondary objects (the compact non-variable sources 16293−2422, CRL2688, CRL618, HL Tau and the variable sources IRC+10216 and OH231.8) have been used to calculate the flux conversion factors (FCF) for each night (Jenness et al. 2000). The calibration maps are taken with the same chop throws as the observations, and give information on the gain of the telescope as it changes throughout the night in response to dish temperature and environmental effects. The FCF is particularly variable at times around sunrise/sunset. Therefore, calibration observations at all three chop-throws and positions are taken at the start and end of each observing block, around sunrise/sunset. Extra, single chop-throw observations are also made when many hours of observations have been taken without calibrating.

A source-extraction routine identical to that used in the final source-extraction analysis is used to find the integrated voltage reading of the calibration source, and this is compared to standard flux values, including adjustments for the known variability of IRC+10216 and OH231.8 (Jenness et al. 2002) to calculate the FCF. This gives an FCF representing Jy V^{-1} for the total flux of a point source (see Section 5.1). This calibration factor corresponds to the typical notation of Jy per beam. Calibration factors can have an error of up to 10 per cent because of variation with time and measurement of the sources' integrated flux. The typical variation in FCF across any SHADES shift is approximately 5 per cent, though this is larger when data are taken around sunrise or sunset, or when observing is extended into daylight hours. The FCF data are time-interpolated over the night, as opposed to using the mean value for each half of the night, as was done for the 8-mJy survey.

Some FCFs do not follow the trend of the data for the rest of the night. The most common reason for this is because the calibrator is extended. Mars behaved as a non-point source object due to its recent proximity to the Earth and IRC+10216 has a CO/dust envelope of at least 1 arcmin in extent. These observations were not used to calculate the night's FCFs. Approximately 10 per cent of our calibration maps are affected in this way. When abnormal FCFs

occur, the calibration changes throughout the night are tested using the pointing observations of the secondary calibrators named above.

Where there are no usable calibration values or no agreement between the pointing observations and calibration observations, standard calibration values are used. Standard calibration values were evaluated using our own sample of calibrator and pointing maps, reduced in IDL. These standard FCFs are approximately 20 per cent lower than the JCMT standards using a SCUBA User Reduction Facility (SURF) reduction, due to the differences in the method used to rebin the data. This systematic difference demonstrates the importance of treating the calibrators in the same way as the sources. One test of these FCFs is to compare the flux densities evaluated for sources with, for example, the flux densities calculated using the 8-mJy survey-reduction method; the flux densities were found not to be systematically affected. Tests were also carried out on the calibration observations to investigate the effect of a noise-weighted source-extraction as opposed to replacing the NEFD with a fixed arbitrary value. The differences were much less than the calibration factor errors for the respective wavelengths ($\sim$1 per cent at 850 μm and $\sim$10 per cent at 450 μm).

4.4 Pointing corrections

Telescope pointing observations are taken at intervals during the evening to correct for the positioning of the telescope on a source with respect to its pointing model. Whilst these offsets are used during data acquisition, drifts between pointing observations cannot be corrected for at the telescope and so must be applied retroactively. Once applied, the SCUBA2MEM program (Jenness & Lightfoot 1998) takes them into account when exporting the positions of each bolometer in the time-series for further processing by the IDL pipeline. Due to the recent announcement of an error in the tracking model at the JCMT, which has affected data taken between 2000 August and 2003 April (Tilanus 2004), we are making pointing corrections to the maps currently reduced. The errors are in azimuth, but are elevation dependent. 35 per cent of the maps covered in this paper were taken during the affected period. However, the median absolute error of these when combining the offset between a pointing observation and its science observation is 0.53 arcsec, with only 15 per cent of the affected maps having an offset of $>$1.5 arcsec, which is of a similar level to the usual rms pointing error of 1.3 arcsec in each coordinate. The maximum error of any one map is 5 arcsec and less than 0.5 per cent of our total maps have a tracking error of $>$3 arcsec.

4.5 Making zero-footprint maps

The final images are produced using an optimal noise-weighted drizzling algorithm (Fruchter & Hook 2002) with a pixel size of 1 arcsec^2. This is the same method as that employed in the SCUBA 8-mJy survey (Scott et al. 2002) and HDF-N (Borys et al. 2003; Serjeant et al. 2003a) data reductions. Both output signal and noise maps were created, the signal in any single pixel being given by the noise-weighted average of the bolometer readouts for which this is the closest pixel. Unlike a standard shift-and-add technique, which takes the flux-density in each detector pixel and places it into the final map over an area equivalent to one detector pixel projected on the sky, drizzling takes the flux density and places it into a smaller area in the final map. Although this significantly reduces the S/N in each pixel, this approach helps preserve information on small angular scales, provided that there are enough observations to fill in the resulting gaps. The area in the coadded map receiving the flux from one detector pixel is termed the footprint. Our method is an extreme example of drizzling in which the footprint is selected to be as small as is practicable given the pointing errors involved (termed the 'zero-footprint'), effectively placing delta functions of flux into boxes 1 arcsec^2. This also allows us to represent each bolometer's view of the sky as an independent measurement, that is, there is no pixel-to-pixel cross-talk. The spatial variation in the sensitivity of the data is taken into account during source-extraction rather than at this map-making stage. For the SHADES data, separate zero-footprint maps are made for each chop-throw and PA.

A possible problem with this methodology is that it is lossy compression: by coadding the data, individual pixels are combined together and information about the distribution of flux readings is lost. However, it is possible during the running of the reduction pipeline to calculate statistics that help evaluate the self-consistency of the data contributing to each pixel and therefore the reliability of the sources that are extracted from the map.

5 THE IMAGE-PROCESSING PROCEDURE

5.1 Matched-filter source-extraction techniques

Convolution of the image with a point spread function (PSF) is the usual method for source-extraction and is the optimal point source filter in the case of uniform noise (von der Heide 1979; Eales et al. 1999, 2000). In the case of non-uniform noise, a method of minimization of χ^2 of the data with the PSF is used, which can be expressed as a convolution (for further details see Serjeant et al. 2003a). This method is optimal for point-source sensitivity, but is not optimal for spatial resolution and assumes that the sources will not be resolved or confused. These assumptions can hold for the 450-μm data, but are more difficult for the 850-μm data in which the beam size is larger and confusion noise is more of a problem.

Indeed, both by examining the data and through clustering arguments it is clear that some sources are partially blended and some are confused. For the science goals of SHADES, it is important that sources fairly close together on the sky can be properly separated, because otherwise much of the potential power of the first radial bin in the angular correlation function, $w(\theta)$, will be lost.

Starting from the simple case of unresolved point sources: generalizing the source-extraction methodology of Serjeant et al. (2003a), the best-fitting (minimum χ^2) point-source flux at any point on the sky is given by

$$F = \frac{\sum_c S_c W_c \otimes P_c}{\sum_c W_c \otimes P_c^2}, \tag{1}$$

where F is the best-fitting flux, c indicates the chop-throw and PA combination, S is the image signal, W is the reciprocal of image variance, P is the PSF and $\otimes$ indicates convolution. The S/N image is derived using

$$\frac{F}{\delta F} = \frac{\sum_c S_c W_c \otimes P_c}{\sqrt{\sum_c W_c \otimes P_c^2}}, \tag{2}$$

where δF is the error on F. In our case, P includes the positive beam and both negative sidelobes.

Simulations shown in Figs 5 and 6 indicate that this multiple chop source extraction yields similar S/N at the peaks as was inserted into the map at the source positions, while removing the problem of negative-chop holes coinciding with other sources.

Sources are currently identified as peaks in the S/N maps, using a connected pixel approach using the IDL LABEL_REGION routine

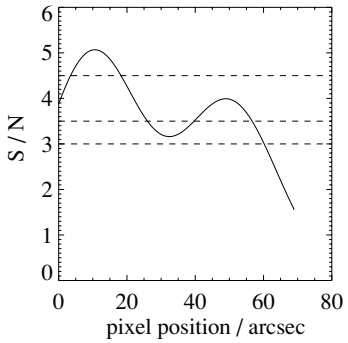

Figure 4. A connected pixels approach on point-source-filtered data with multiple threshold cuts can identify blended sources, which have two separate peaks. The 4.5σ cut would identify only the left-hand peak, whereas the 3σ cut would only identify the same. The 3.5σ cut would identify two peaks.

(http://www.rsinc.com) to find the peak. Sources were identified using multiple S/N threshold cuts between 3.0σ and the peak S/N of the map. These multiple cuts allow separation of blended sources, where there are separate peaks. (see Fig. 4).

It can be shown that the method of multiple threshold cuts can deblend Gaussian sources with equal fixed full width at half-maximum (FWHM) having minimum separation between peaks of $\sqrt{-2\,\sigma^2 \ln p}$ where FWHM $= 2\sqrt{2\ln 2}\sigma$ and $p =$ ratio of the individual sources' peak S/N where $0 < p \leqslant 1$. Sources are found with peaks $>3.0\sigma$ and we have a maximum S/N of $\sim$8. We would therefore consider peaks with separation less than $\sim$8.33 arcsec. This is comparable to the beamsize of the SCUBA beam and therefore this should only become a significant issue if sources are found with much higher S/N. For the purpose of our completeness and reliability simulations and for our future source count derivations from this analysis, we only use the method of multiple cuts described above.

Table 2. The best-fitting parameters for fitting two sources to the area around Lockman 850.1 and 850.8. The position and flux errors are formal 1σ errors computed from the covariance matrix.

Position	Flux (mJy)	S/N
10 52 01.31 + 57 24 44.4	5.2 ± 1.1	4.9
10 52 00.33 + 57 24 19.0	9.3 ± 1.2	8.0

Since source counts are not the only science goal, it may be useful to attempt deblending of source pairs, which are closer than the beamwidth. To that end a more sophisticated source deblending algorithm has been developed, building on that detailed in Scott et al. (2002). In the Scott et al. technique, the flux densities of all significant peaks are fitted simultaneously using a maximum-likelihood

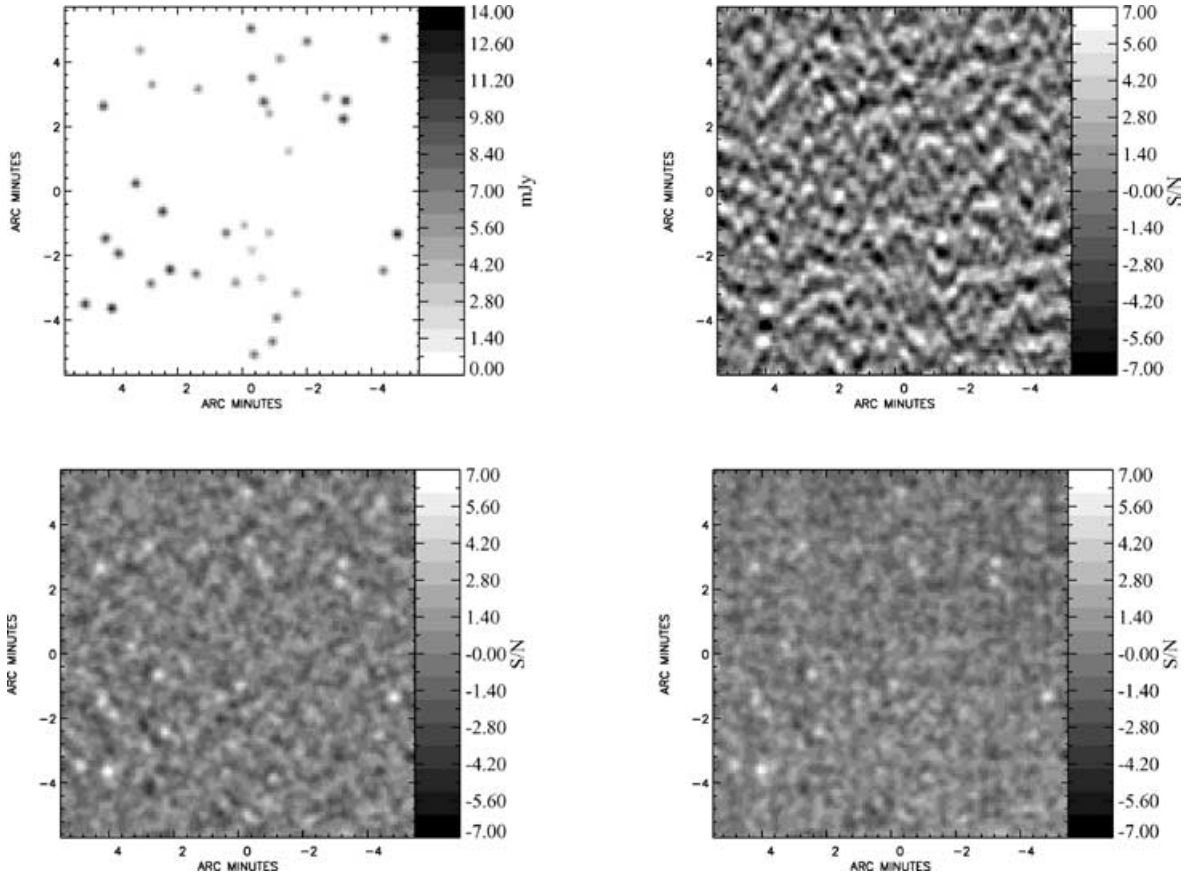

Figure 5. A simulation of source-extraction using the multi-chop technique. Top left: the input sources as they would appear with no noise or chopping. Top right: a source-extracted map using only one chop-throw/PA image. This is the method that was used for the 8-mJy survey, such that the source-extraction routine causes a double chop function to be seen. This causes blending of some of the sources with the chop functions of other sources. Bottom left: a source-extracted map using six chop throw and PA combinations, each with the individual map having $\sqrt{6}$ times the noise used in one chop image. Bottom right: the same simulated map was used in an Emerson-II deconvolution method to produce a map without significant chop residues.

technique at the positions of the peaks. We have extended this to fit to the spatial position as well as the flux, using the position and fluxes of sources from the source-extraction method as a starting point and fitting between one and six sources at each position. This becomes a computationally tractable problem when we consider only the pixels around the $>3\sigma$ peaks in the map that would contribute to the flux at that position.

Adding additional sources to the fits can improve the reduced χ^2, but these new sources may not necessarily represent the underlying distribution. Distinguishing between two close point sources and some other extended structure (such as a single extended source) implicitly involves constraining the quadropole moment of the image, and as Lucy (1992) has shown, the S/N requirements for constraining the higher-order multipoles are extremely stringent. Nevertheless, the positions of known sources may be refined using our multiwavelength data. Such refinement is optimal if one simultaneously fits all the relevant multiwavelength data, and this can be done within our methodology.

As an example of this methodology, we have attempted to deconvolve Lockman 850.1 and 850.8 using the 850 μm maps from the 8-mJy survey. The reported fluxes of these sources in the 8-mJy survey are 10.5 and 5.1 ± 1.3 mJy, respectively. We performed the multiple source fitting, and the best fit is given in Table 2, the covariance matrix of which shows a weak correlation of fluxes with positions and between the fluxes of the two sources.

5.2 Emerson-II deconvolution

An alternative method to construct images is to use the multiple chop strategy to recover the modes missing in any single chop, through the Emerson-II deconvolution algorithm (Emerson 1995; Jenness et al. 2000). This has been shown to be effective in reproducing sources in data with two chop throws and one PA, as was the observing strategy applied to the HDF (Hughes et al. 1998; Serjeant et al. 2003a). This is not the only possible methodology, but it has the advantage of having clear precedents in Galactic astronomy (e.g. Johnstone et al. 2000; Pierce-Price et al. 2000).

The difference image produced by the chop process during observation can be considered a convolution of the sky with the chop function. Recognizing that a convolution in real space is the same as multiplication in Fourier space, the Emerson-II algorithm reconstructs the image by effectively dividing by the chop function in Fourier space. Using this technique, modes in Fourier space are lost where the Fourier transform of the chop function is zero, but these can be filled in using modes from data taken at other chop angles. This method has the benefit of using the flux from the negative chop holes and *folding* it back on to the source position. The downside of this for the HDF data is that modes are lost altogether because the 30 and 45 arcsec are in a 2:3 ratio. In contrast, the chop strategy implemented for SHADES is well suited to this reduction method because the chop throws are incommensurate and the particular values of 30, 44 and 68 arcsec have been shown previously to work well with the Emerson-II technique. Fig. 5 shows an example of the source-extraction method of the primary pipeline, in which a number of sources, simulated using the 8-mJy survey source positions and fluxes, have been extracted using the multiple-chop source-extraction described above. Fig. 5 also shows the alternative method of reconstructing the image using the Emerson-II algorithm [the analogous algorithm for SCUBA scan maps is discussed in Johnstone et al. (2000)]. These two algorithms give superficially extremely similar reconstructions; this is partly because the multiple chop throws and PAs distribute the negative sidelobe fluxes over a large number of positions. Although the Emerson-II method may be useful for producing cosmetically clean images, we have not investigated the presence of possible artefacts made by the map-making process, and we prefer to use the direct method outlined in Section 5.1 to find sources.

Another approach is to use an iterative reconstruction scheme, motivated by cosmic microwave background map-making methods (e.g. Wright, Hinshaw & Bennett 1996) and successfully applied to SCUBA scan maps by Johnstone et al. (2000). The triple-beam pattern used in jiggle-map mode is more problematic for this approach, since each datum is the difference between one map value and the average of two others. Investigation of the iterative reconstruction of SHADES maps has been of only limited success (Lepage 2004).

5.3 Comparison of these two methods

To assess the relative merits of these two methods, we compare the completeness and reliability of sources found using simulations (see Fig. 6). The simulations have two components. First, a single Gaussian chop profile source is added at a random position in the real zero-footprint chop maps. We then attempt to recover the source (within 7 arcsec of the input position and within a factor of 2 of flux and S/N > 3.5) and information about the position and flux of the output source is retained. This gives information about the completeness and flux boosting. In the second method, a map is entirely simulated using the source counts of Scott, Dunlop & Serjeant (2005), the level of clustering seen in the 8-mJy survey (Scott et al. 2002), and Gaussian random noise is added using the real noise map as input. Sources are then extracted as described above and the results used for calculating the reliability of the extracted sources, which we here define as the number of 'real' sources found as a fraction of the total number of sources found by the source-identification process.

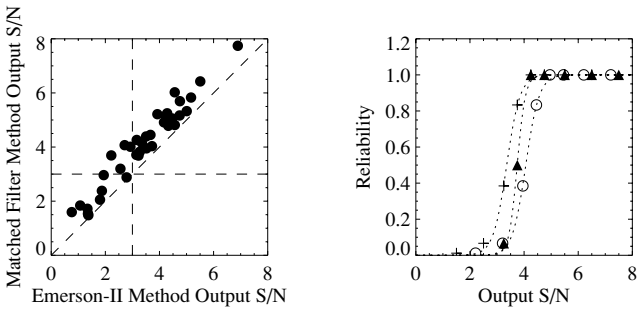

Figure 6. Left: comparison of the output S/N of the input sources for the matched-filter and Emerson-II deconvolution methods, for the simulations shown in Fig. 5. The dashed lines show S/N = 3 for each of the methods, and the line for which the two methods would have equal S/N. Right: the reliability of sources extracted from the matched-filter method (filled triangles) and Emerson-II deconvolved maps (crosses). This is based only on the sources in the top-left-hand panel, which is far fewer than the more extensive simulations in Fig. 8; however, since these are the same sources being extracted in both matched-filter and Emerson-II methods, the *relative* reliabilities can still be usefully compared. For a fixed S/N, the reliability appears to be higher for the raw Emerson-II deconvolution, but this is easily shown to be an artefact of the poorer S/N of Emerson-II for a fixed input flux (left). After adjusting for the poorer S/N in the Emerson-II method (open circles), the reliability is shown to be less for the Emerson-II deconvolved maps than the matched-filter method. Thus, for example, a 3.5σ source in an Emerson-II map is intrinsically brighter than a 3.5σ source in a matched-filter map, and it is unremarkable that such brighter sources can be extracted more reliably; however, an 8-mJy source extracted from an Emerson-II map has a lower reliability than an 8-mJy source from a matched-filter extraction.

The reliability of the Emerson-II method appears higher for a given S/N (crosses in Fig. 6) than the matched-filter method (filled triangles), but this is because the S/N for an individual source is lower in the Emerson-II method. After correcting for this S/N difference (open circles), it is clear that the matched filter has a higher reliability than Emerson-II for a given source (i.e. rather than a given S/N).

6 RESULTS AND ANALYSIS

6.1 Progress

A total of 1843 individual jiggle-maps, observed over 139 nights up to 2004 February, have been coadded into zero-footprint maps covering 720 arcmin^2. The median noise level in the 850-μm data is 2.2 mJy and that of the 450-μm data is 25 mJy. The data were taken within a range of weather conditions (JCMT grades 2–3) such that the mean $\tau_{850} = 0.27 \pm 0.07$ and $\tau_{450} = 1.45 \pm 0.44$. Ninety-two per cent of the data have τ_{850} in the range 0.20–0.48 and τ_{450} in the range 0.94–2.78.

The areas covered by the survey to 2004 February are shown in Fig. 7. Note the uniformity of the noise levels within the central regions of the maps as designed, with the higher-noise regions visible around the perimeter where further observations are due to be taken. Although the SHADES data are taken only within the fixed weather conditions outlined above, the data taken at the start for the SXDF were taken with consistently poorer than average weather (grade 3), hence the higher noise level at the centre of the SXDF map. However, the depth is consistent across that area and varies only by $\simeq 0.7$ mJy from the rest of the map. Also note the deep strip in the Lockman Hole East, taken during the SCUBA 8-mJy survey.

The source-extraction method can also yield sources near the edge of the map, which we choose to reject due to insufficient coverage/sampling and/or high noise values. These can be identified and removed from the source lists by creating a mask of the integration time convolved with the beam (which in our case rejects those regions near the edge of the map with little coverage) and rejection of those sources with large flux-density errors (typically greater than 10 mJy at 850 μm and greater than 100 mJy at 450 μm).

Using these criteria, simulations were carried out as described above in Section 5.3 for the full SHADES interim maps. These gave completeness, reliability and flux boosting effects as shown in Fig. 8. The expected completeness and reliability at 3σ, 3.5σ and 4σ are shown in Table 3.

Table 4 gives the number of sources at different significance cuts for each of the SHADES fields using the source-extraction method outlined above. In order to estimate the source-detection density,

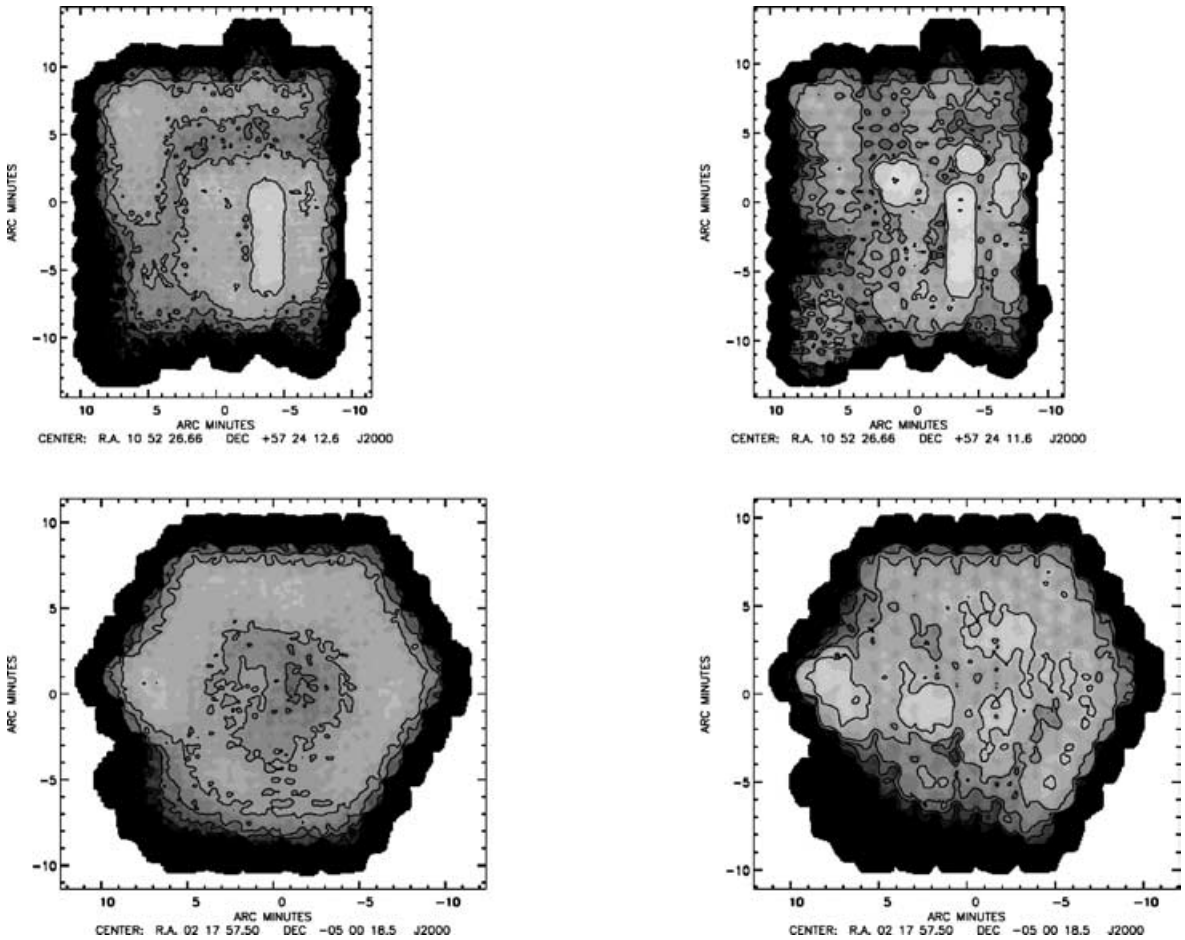

Figure 7. Noise maps of the (top row) Lockman Hole East and (bottom row) SXDF for data taken up to 2004 February. For the 850-μm data (left column), the lowest contour level is 0.71 mJy with higher contours spaced by 0.71 mJy to a maximum level shown in the map of 5 mJy. For the 450-μm data (right column), the lowest contour level is 7.1 mJy with higher contours spaced by 7.1 mJy to a maximum level in the map of 50 mJy. Note the high levels of noise around the edges of the map, which will be reduced when more observations are taken.

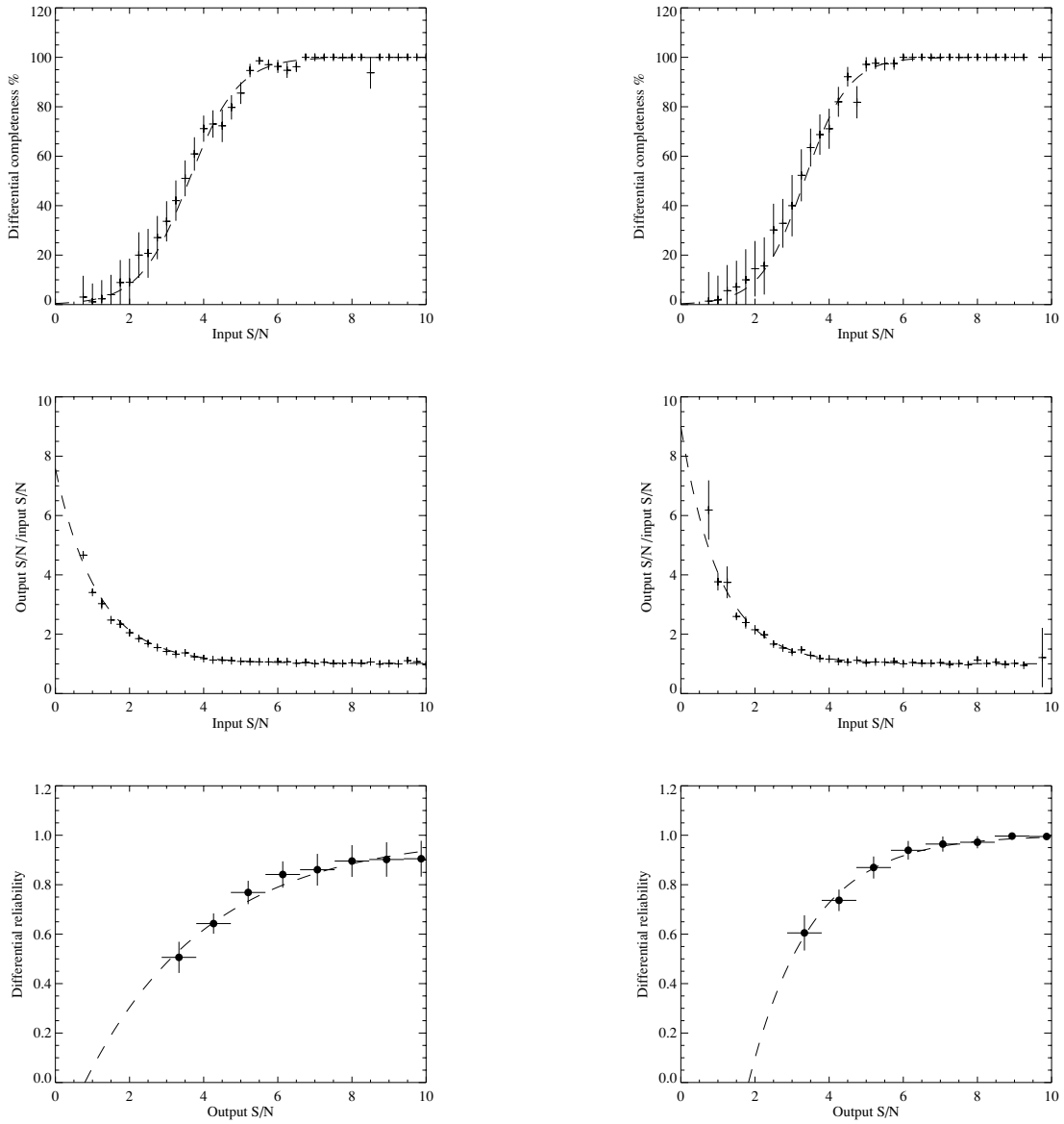

Figure 8. Completeness (top), flux boosting (middle) and reliability (bottom) for the Lockman Hole East (left) and SXDF (right).

Table 3. Comparison values of reliability (R), completeness (C), flux-boosting factor (B) and positional errors (D) at different threshold cuts for our two fields.

	R	C	B	D (arcsec)
LH				
3.0σ	48 per cent	29 per cent	1.47	2.90
3.5σ	56 per cent	46 per cent	1.31	2.85
4.0σ	62 per cent	65 per cent	1.21	2.79
SXDF				
3.0σ	50 per cent	36 per cent	1.45	2.89
3.5σ	63 per cent	57 per cent	1.28	2.73
4.0σ	73 per cent	76 per cent	1.18	2.57

we have conservatively considered only those sources that are also found in two independent reductions, our so-called consensus list (which will be discussed in future papers). This in effect means that, for sources of modest significance in the range 3–3.5σ, we only

Table 4. Number of sources at 850 μm greater than the S/N cut used in the source-extraction procedure. Note the steep negative slope. Numbers in brackets are the estimated number of statistically spurious sources assuming Gaussian noise.

S/N	N(>S/N) Lockman	N(>S/N) SXDF
1.0	1773 (1530)	1372 (1128)
2.0	648 (220)	524 (162)
3.0	98 (13)	106 (10)
3.5	46 (2.25)	40 (1.66)
4.0	24 (0.3)	16 (0.2)
5.0	6 (0)	4 (0)
6.0	2 (0)	2 (0)
6.5	1 (0)	1 (0)
Median 1σ noise at 850 μm	2.28 mJy	2.14 mJy

consider those that are reproduced in all four reductions undertaken so far.

In the Lockman Hole East map, 69 sources have been identified with S/N $\geqslant$ 3 of which 47 have S/N $\geqslant$ 3.5. In the SXDF, 61 sources have been identified with S/N $\geqslant$ 3, of which 53 have S/N $\geqslant$ 3.5. Thus, with 40 per cent of the data taken, SHADES has produced a sample of 100 sources at 850 μm with S/N $\geqslant$ 3.5.

The implied surface density of sources with S/N $\geqslant 3\sigma$ is therefore 653 ± 57 sources deg^{-2}. The error on the estimated source density is calculated using the Poisson error on the data (i.e. the number of sources observed), corrected for the area considered. The source density is likely to be an underestimate since parts of the maps have only been covered by a single chop at this time and consequently have a higher noise level. However, the uncorrected source density is consistent with a surface density of point sources more than sufficient for the science goals described in this paper.

Fig. 9 shows the S/N histogram of the new maps. A Gaussian can be fitted to the S/N data, though at high S/N levels an excess above this fit shows there are real sources in the map. This is seen to be the case with the SXDF observations; in the Lockman Hole, the effect of the combination of the deep strip (at one chop throw only) and the noisier edges is evident from the higher excess of pixels at higher S/N compared to the SXDF. The Lockman Hole deep strip has one chop throw only, so we also see an excess at large negative S/N due to the sidelobes of the sources in this deep strip. This negative excess is also present in the SXDF maps, but at a much lower level because our chop strategy deliberately reduces the chop holes (see Fig. 5). If there were no sources in the map, the plot should follow the Gaussian shape. Instead, it is possible to see the statistical detections at higher S/N as an excess of pixels with that S/N, especially at 850 μm. Also, assuming Gaussian random noise, we would expect to find 0.62 per cent of the survey beams to contain spurious detections at the 2.5σ level. This would mean that if there were no real sources in the map, we would expect to see a total of about 60 spurious sources in the Lockman Hole maps so far, and 44 spurious sources in the SXDF assuming a beam size of 14 arcsec at 850 μm. Instead we detect approximately 270 source candidates at $\geqslant 2.5\sigma$ in the Lockman Hole map and 250 source candidates in the SXDF map. We have extended this analysis to other σ cuts in Table 4.

6.2 Comparison with the 8-mJy survey

A reanalysis of the 8-mJy survey Lockman Hole East data was carried out in order to test the new pipeline. Of the 21 published sources at greater than 3.5σ (Scott et al. 2002), 4 (LE850.9, 10, 15 and 20) were rejected by Ivison et al. (2002) because they have $\sigma_{850} > 3$ mJy. This leaves 17. Of these, 12 were found using the new analysis method. Of the 12 8-mJy survey sources with S/N greater than 3.5 that have been reproduced, two lack a radio ID (LE850.4 and 11). The possible implications of this were outlined in Section 2.1. For those sources reproduced, Table 5 shows the S/N of the source in the new reduction. For those sources not reproduced, the S/N at the position of the original 8-mJy survey 'detection' is shown.

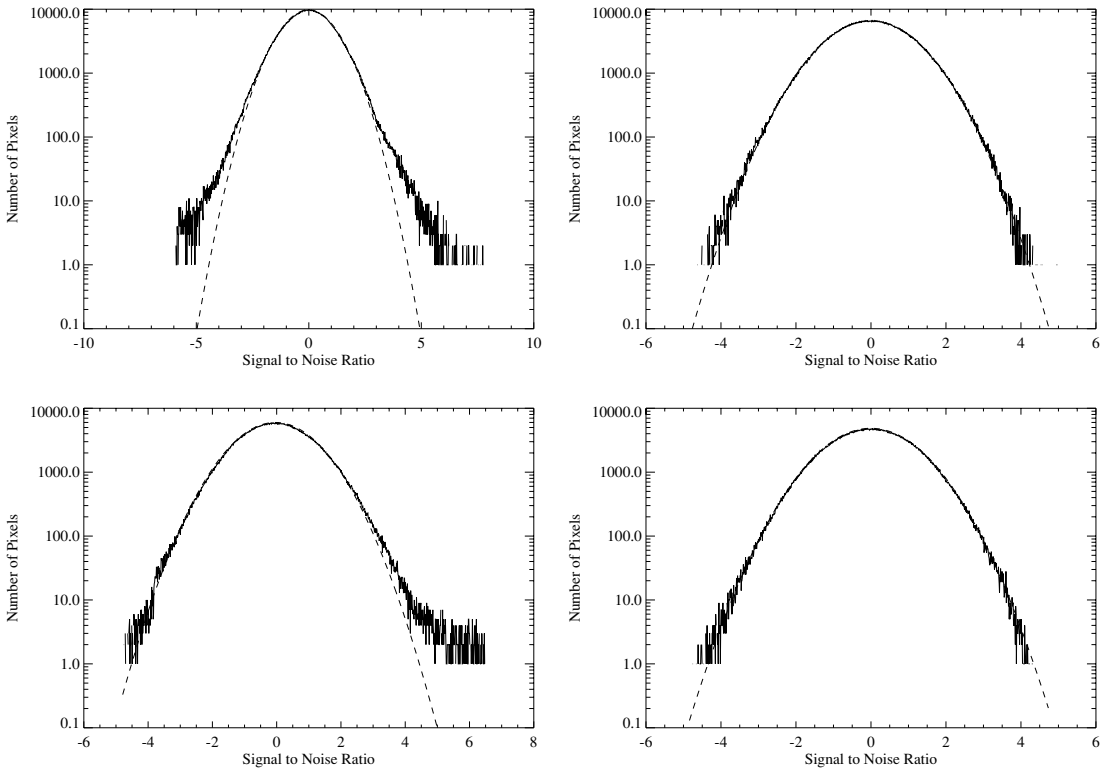

Figure 9. S/N histograms of the point-source-filtered SHADES maps, with a Gaussian fit to the data plotted as a dashed line. The noise levels are derived from the bolometer readouts using the methodology described in Section 4.2, which are used in the map-making (Section 4.5) and source-extraction (Section 5.1) procedures. Top left-hand panel: Lockman Hole 850-μm map – mean −0.01, variance 1.07. Top right-hand panel: Lockman Hole 450-μm map – mean −0.01, variance 1.02. Bottom:Left SXDF 850-μm map – mean −0.04, variance 1.15. Bottom left-hand panel: SXDF 450-μm map – mean −0.00, variance 1.11. Note the very close fit to the data at 450 μm and the excess S/N at 850 μm.

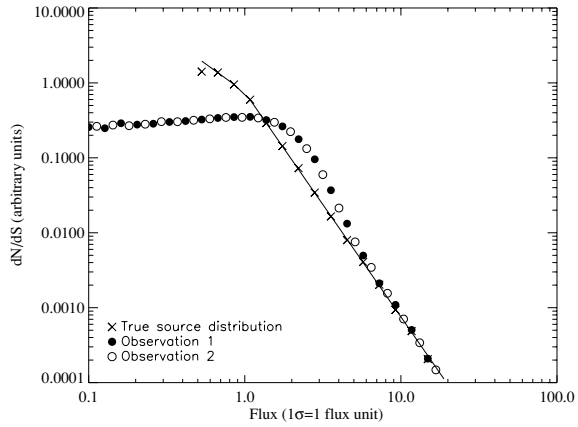

Figure 10. A simulation of Eddington bias for a sample of 10 000 sources with 1σ equal to one flux unit. Although flux boosting does occur for a fraction of the simulated sources, the actual sources detected between 3σ and 3.5σ vary between individual observations because of the random effect of the noise. Of the 10 000 sources simulated here, ∼2600 were found with S/N between 3σ and 3.5σ in the first and/or second observations, but only 535 were found with S/N in the range 3σ–3.5σ in the first observation and $>3\sigma$ in the second. Therefore, when applying a significance threshold, the lower-significance source lists derived from each observation are markedly different, despite the fact that both observations yield clearly consistent Eddington-biased source counts.

Of the sources previously detected between 3 and 3.5σ, only two have been confirmed as peaks in the S/N distribution, and one source has positive flux at the position with S/N of greater than 3. Extra chop data in the area have confirmed one source at a significance level greater than 3.5σ (LE850.29).

Only one of the nine missing sources originally found at a significance level greater than 3.5σ has a robust radio identification. This is source LE850.6, which has been resolved into two sources, but neither of the new peaks lies within 7 arcsec of the original source position. Note that further analysis of the Ivison et al. (2002) radio identifications in the Lockman Hole was carried out by Greve et al. (2004) and their results are used here. Source LE850.13 has been lost using this data-reduction method and does not return when all six chop data are added to the region.

One possibility is that the missing Scott et al. (2002) sources were spurious, perhaps due to weak bolometer noise spikes that escaped clipping in the earlier analysis. However, this cannot be the whole explanation since there is a weak positive signal in the positions of the missing sources [as quoted by (Scott et al. 2002); see Table 5] in our refined reanalysis.

One other possibility that could lead to *apparently* spurious sources, in the sense that the sources seem to be non-repeatable, is Eddington bias (Eddington 1913). This bias, sometimes confused with Malmquist bias (Teerikorpi 1997), describes a feature of observing a population of objects above a given flux limit with a negative-sloping source count. Random errors in the flux measurements of the objects can systematically alter the source counts measured above the flux limit, such that more sources have their flux densities boosted above the flux limit, than those that fall below the limit. This makes there appear to be more sources close to the flux limit than in the true population, the effect being more pronounced for lower S/N thresholds. A simulation of this effect on 10 000 sources can be seen in Fig. 10. When observations are repeated, different sources are boosted above the flux limit while others again drop below the flux limit. The number of sources observed in each of the first and second observations with a S/N in the range 3–3.5σ is ∼2600, but the number of sources with S/N in the range 3–3.5σ in the first observation and $>3\sigma$ in second is only 535. In this way, it is possible to have up to 80 per cent of the 3–3.5σ sources not appearing in both samples. This 80 per cent is made up of those sources that are flux boosted in the first observation and not in the second, or vice versa, and it will happen even if all the 3σ–3.5σ lists are 100 per cent reliable. By reanalysing the noise in the 8-mJy survey maps, we have reweighted the observations, and effectively resampled the noise, although the effect will probably be smaller than that shown in Fig. 10 because the raw data are the same and it is the dominant sky noise that has been remeasured. This might explain why the lower significance 3σ–3.5σ source lists from the 8-mJy survey data

Table 5. The S/N of the sources using the new data-reduction software, applied to the 8-mJy survey data only. The S/N quoted is of the source that corresponds to the old source (less than 7 arcsec from the original 8-mJy survey source). Radio identifications are from Greve et al. (2004) and selected such that the probability that an apparent radio-source identification could be the result of chance is $p < 0.05$. Of the nine sources in the original S/N $\geqslant 3.5$ list that have now been lost, four have already been rejected by Ivison et al. (2002) as likely reduction artefacts or seriously flux-boosted sources extracted from the high-noise regions of the original 8-mJy maps (LE850.9, 10, 15 and 20, marked by an asterisk). Source LE850.6 has been resolved into two blended sources, but does not strictly have a peak within 7 arcsec of the original source. The new sources found from our reduction are named with the prefix LE850.M, and do not appear in the current consensus lists. Not listed here are the eight sources found by our algorithm that *do* appear in the consensus reduction.

8-mJy source	8-mJy S/N	New S/N	Radio ID?
LE850.1	8.1	8.0	Yes
LE850.2	5.2	5.1	Yes
LE850.3	5.1	3.9	Yes
LE850.4	5.0	5.1	No
LE850.5	4.6	2.0	No
LE850.6	4.5	2.8	Yes
LE850.7	4.5	4.0	Yes
LE850.8	4.4	3.8	Yes
LE850.9★	4.2	0.8	No
LE850.10★	4.2	1.5	No
LE850.11[a]	4.1	4.3	No
LE850.12	4.0	4.3	Yes
LE850.13	3.7	1.4	No
LE850.14	3.6	5.2	Yes
LE850.15★	3.6	1.8	No
LE850.16	3.6	4.3	Yes
LE850.17	3.5	1.3	No
LE850.18	3.5	4.2	Yes
LE850.19	3.5	1.7	No
LE850.20★	3.5	0.2	No
LE850.21	3.5	4.0	Yes
LE850.27	3.4	3.3	N/A
LE850.29	3.3	4.7	N/A
LE850.32	3.2	3.2	N/A
LE850.M1	–	3.88	Yes
LE850.M2	–	3.77	?
LE850.M3	–	3.73	N/A
LE850.M4	–	3.51	N/A

[a]Source rejected in Ivison et al. (2002) for its high noise value but considered robust here.

reduction and the SHADES data reduction contain only three sources in common.

15 new $\geqslant 3.5\sigma$ sources have been identified using the new source-extraction method in the 8-mJy survey area, of which only seven are not also in the consensus lists. Three of these 15 have noise greater than 3 mJy. One of the new sources in the consensus list was first reported in full detail in Serjeant et al. (2004) with two possible *Spitzer* detections, which agree strongly with the radio identifications presented there. Although not detected at 450 and 1200 μm, weak positive signals of 1.6σ and 1.8σ have been found at this source position in both maps. The four sources *not* in the consensus list and with noise less than 3 mJy are listed in Table 5.

A future paper will perform a detailed comparison of the pipelines over all the SHADES area, that is, not just restricted to the 8-mJy Lockman Hole Field.

7 CONCLUSIONS AND SURVEY PROGRESS

To 2004 February, the SHADES areas covered a region of 720 arcmin^2 (approximately 40 per cent of the total expected area) using 1843 individual jiggle maps that have been observed over 139 nights during the period 1998 March–2004 February.

Source extraction from these maps indicates at least 653 ± 57 sources deg^{-2} having S/N >3 at 850 μm in these survey fields (uncorrected for Eddington bias), consistent with a surface density of point sources more than sufficient for the science goals described in this paper.

This paper has outlined the SHADES survey goals, data-taking and data-reduction strategies, as part of which a new SCUBA reduction pipeline has been developed using IDL, based upon the reduction pipeline used for the 8-mJy survey. A test of this data-reduction and source-extraction method was made by comparing the new maps with the sources extracted from the original 8-mJy survey Lockman Hole East maps that form a subset of the SHADES maps. Of the 17 more secure ($\geqslant 3.5\sigma$) sources in the 8-mJy survey, 12 have been reconfirmed, of which 10 have radio identifications. 13 new candidate sources with S/N $\geqslant$ 3.5 and ($\sigma_{850} < 3.5$ mJy) have been identified using the new data-reduction method, nine of which appear in the consensus between other reductions, which will be discussed in future papers.

A full presentation and analysis of the submm source counts derived from the interim SHADES maps discussed here will be presented in a separate paper. This paper will also compare the results of the data-reduction and source-extraction described here, with the outcome of three further independent reductions/extractions carried out at IfA Edinburgh, INAOE Mexico, and UBC Vancouver.

The SHADES consortium maintains a public web page at http://www.roe.ac.uk/ifa/shades/.

ACKNOWLEDGMENTS

Major parts of this work were supported by PPARC and by the Canadian NSERC. This work is also supported by Nuffield Foundation Grant number NAL/00529/G. AJB, ACE, OA, PNB and IRS acknowledge the Royal Society for generous funding. DHH, IA and JW are supported by CONACYT grants 39953-F and 39548-F. CPP acknowledges a European Union Fellowship to Japan. EvK acknowledges funding by the Austrian Science Foundation FWF under grant P15868. The JCMT is operated by the Joint Astronomy Centre on behalf of the UK Particle Physics and Astronomy Research Council, the Canadian National Research Council and the Netherland Organization for Scientific Research.

REFERENCES

Alexander D. et al., 2003, AJ, 125, 383
Alexander D. M., Smail I., Bauer F. E., Chapman S. C., Blain A. W., Brandt W. N., Ivison R. J., 2005, Nat, 434, 738
Almaini O., Lawrence A., Boyle B. J., 1999, MNRAS, 305, L59
Almaini O. et al., 2003, MNRAS, 338, 303
Archibald E. N. et al., 2002, MNRAS, 336, 1
Aretxaga I., Hughes D. H., Chapin E. L., Gaztañaga E., Dunlop J. S., Ivison R. J., 2003, MNRAS, 342, 759
Aretxaga I., Hughes D. H., Dunlop J. S., 2005, MNRAS, 358, 1240
Barger A. J., Cowie L. L., Sanders D. B., Fulton E., Taniguchi Y., Sato Y., Kawara K., Okuda H., 1998, Nat, 394, 248
Barger A. J., Cowie L. L., Sanders D. B., 1999, ApJ, 518, L5
Barger A. J., Cowie L. L., Richards E. A., 2000, AJ, 119, 2092
Barger A. J., Cowie L. L., Mushotzky R. F., Richards E. A., 2001, AJ, 121, 662
Baugh C. M., Cole S., Frenk C. S., Lacey C. G., 1998, ApJ, 498, 504
Baugh C. M., Lacey C. G., Frenk C. S., Granato G. L., Silva L., Bressan A., Benson A. J., Cole S., 2005, MNRAS, 356, 1191
Bautz M. W., Malm M. R., Baganoff F. K., Ricker G. R., Canizares C. R., Brandt W. N., Hornschemeier A. E., Garmire G. P., 2000, ApJ, 543, L119
Bell E. F. et al., 2004, ApJ, 608, 752
Blain A. W., 1999, MNRAS, 309, 955
Blain A. W., Longair M. S., 1996, MNRAS, 279, 847
Blain A. W., Jameson A., Smail I., Longair M. S., Kneib J.-P., Ivison R. J., 1999, MNRAS, 309, 715
Blain A. W., Barnard V. E., Chapman S. C., 2003, MNRAS, 338, 733
Blain A. W., Chapman S. C., Smail I., Ivison R. J., 2004a, ApJ, 611, 52
Blain A. W., Chapman S., Smail I., Ivison R. J., 2004b, ApJ, 611, 725
Borys C., Chapman S., Halpern M., Scott D., 2003, MNRAS, 344, 385
Borys C., Scott D., Chapman S. C., Halpern M., Nandra K., Pope A., 2004, MNRAS, 355, 485
Carilli C. L., Yun M. S., 1999, ApJ, 513, L13
Charmandaris V. et al., 2004, ApJS, 154, 142
Chapman S., Blain A. W., Ivison, R. J., Smail I., 2003, Nat, 422, 695
Chapman S., Blain A. W., Smail I., Ivison R. J., 2005, ApJ, 622, 772
Clements D. et al., 2004, MNRAS, 351, 447
Cole S., Lacey C. G., Baugh C. M., Frenk C. S., 2000, MNRAS, 319, 168
Dole H. et al., 2004, ApJS, 154, 87
Downes D. et al., 1999, A&A, 347, 809
Dunlop J. S., 2002, in Metcalfe N., Shanks T., eds, ASP Conf. Ser. Vol. 283, A New Era in Cosmology. Astron. Soc. Pac., San Francisco, p. 381
Dunlop J. et al., 2004, MNRAS, 350, 769
Dunne L., Clements D. L., Eales S. A., 2000, MNRAS, 319, 813
Dwek E. et al., 1998, ApJ, 508, 106
Eales S., Lilly S., Gear W., Dunne L., Bond J. R., Hammer F., Le Févre O., Crampton D., 1999, ApJ, 515, 518
Eales S., Lilly S., Webb T., Dunne L., Gear W., Clements D., Yun M., 2000, AJ, 120, 2244
Eddington A. S., 1913, MNRAS, 73, 359
Efstathiou A., Rowan-Robinson M., 2003, MNRAS, 343, 322
Efstathiou A., Hough J. H., Young S., 1995, MNRAS, 277, 1134
Efstathiou A., Rowan-Robinson M., Siebenmorgen R., 2000, MNRAS, 313, 734
Egami E. et al., 2004, ApJS, 154, 130
Elbaz D., Cesarsky C. J., Chanial P., Aussel H., Franceschini A., Fadda D., Chary R. R., 2002, A&A, 384, 848
Emerson D., 1995, in Emerson D. T., Payne J. M., eds, ASP Conf. Ser. Vol. 75, Multi-Feed Systems for Radio Telescopes. Astron. Soc. Pac., San Francisco, p. 309
Fabian A. C. et al., 2000, MNRAS, 315, L8
Farrah D., Serjeant S., Efstathiou A., Rowan-Robinson M., Verma A., 2002, MNRAS, 335, 1163
Farrah D., Fox M., Rowan-Robinson M., Clements D., Afonso J., 2004, ApJ, 603, 489
Fazio et al., 2004, ApJS, 154, 39

Fox M. J. et al., 2002, MNRAS, 331, 839
Franceschini A., Toffolatti L., Mazzei P., Danese L., de Zotti G., 1991, A&AS, 89, 285
Frayer D. T. et al., 2004, ApJS, 154, 137
Fruchter A. S., Hook R. N., 2002, PASP, 114, 144
Gear W. K., Lilly S. J., Stevens J. A., Clements D. L., Webb T. M., Eales S. A., Dunne L., 2000, MNRAS, 316, L51
Genzel R., Baker A. J., Tacconi L. J., Lutz D., Cox P., Guilloteau S., Omont A., 2003, ApJ, 584, 633
Gispert R., Lagache G., Puget J.-L., 2000, A&A, 360, 1
Granato G. L., Lacey C. G., Silva L., Bressan A., Baugh C. M., Cole S., Frenk C. S., 2000, ApJ, 542, 710
Granato G. L., De Zotti G., Silva L., Bressan A., Danese L., ApJ, 600, 580
Greve T., Ivison R. J., Bertoldi F., Stevens J. A., Dunlop J. S., Lutz D., Carilli C. L., 2004, MNRAS, 354, 779
Guiderdoni B., Hivon E., Bouchet F. R., Maffei B., 1998, MNRAS, 295, 877
Hatton S., Devriendt J. E. G., Ninin S., Bouchet F. R., Guiderdoni B., Vibert D., 2003, MNRAS, 343, 75
Heavens A., Panter B., Jimenez R., Dunlop J., 2004, Nat, 428, 625
Holland W. et al., 1999, MNRAS, 303, 659
Huang J.-S. et al., 2004, ApJS, 154, 44
Hughes D. H. et al., 1998, Nat, 394, 241
Hughes D. H. et al., 2002, MNRAS, 335, 871
Ivison R. J., Smail I., Le Borgne J.-F., Blain A. W., Kneib J.-P., Bezecourt J., Kerr T. H., Davies J. K., 1998, MNRAS, 298, 593
Ivison R. J., Smail I., Barger A. J., Kneib J.-P., Blain A. W., Owen F. N., Kerr T. H., Cowie L. L., 2000, MNRAS, 315, 209
Ivison R. J. et al., 2002, MNRAS, 337, 1
Ivison R. J. et al., 2004, ApJS, 154, 124
Jenness T., Lightfoot J. F., 1998, in Albrecht R., Hook R. N., Bushouse H. A., eds, ASP Conf. Ser. Vol. 145, Astronomical Data Analysis Software and Systems VII. Astron. Soc. Pac., San Francisco, p. 216
Jenness T., Holland W. S., Chapin E., Lightfoot J. F., Duncan W. D., 2000, in Monset N., Veillet C., Crabtree D. ASP Conf. Ser. Vol. 216, Astronomical Data Analysis Software and Systems IX. Astron. Soc. Pac., San Francisco, p. 559
Jenness T., Stevens J. A., Archibald E. N., Economou F., Jessop N. E., Robson E. I., 2002, MNRAS, 336, 14
Johnstone D., Wilson C. D., Moriarty-Schieven G., Giannakopoulou-Creighton J., Gregersen E., 2000, ApJS, 131, 505
Kauffmann G., Colberg J. M., Diaferio A., White S. D. M., 1999, MNRAS, 303, 188
Kaviani A., Haehnelt M. G., Kauffmann G., 2003, MNRAS, 340, 739
Larson R. B., 1998, MNRAS, 301, 569
Lawrence A., 2001, MNRAS, 323, 147
Le Floc'h E. et al., 2004, ApJS, 154, 170
Lepage K., 2004, Masters thesis, Univ. of British Columbia
Lilly S. J., Le Fèvre O., Crampton D., Hammer F., Tresse L., 1995, ApJ, 455, 50
Lilly S. J., Eales S. A., Gear W. K. P., Hammer F., Le Fèvre O., Crampton D., Bond J. R., Dunne L., 1999, ApJ, 518, 641
Lonsdale C. et al., 2003, PASP, 115, 897
Lonsdale C. et al., ApJS, 154, 54
Lucy L., 1992, AJ, 104, 1260
Lutz D. et al., 2001, A&A, 378, 70
Marleau F. R. et al., 2004, ApJS, 154, 66
Panter B., Heavens A. F., Jimenez R., 2003, MNRAS, 343, 1145
Pearson C. P. et al., 2004, MNRAS, 347, 1113
Percival W. J., Scott D., Peacock J. A., Dunlop J. S., 2003, MNRAS, 338, L31
Pierce-Price D. et al., 2000, ApJ, 545, L121
Pope A., Borys C., Scott D., Conselice C., Dickinson M., Mobasher B., 2005, MNRAS, 358, 149
Rengarajan T. N., Takeuchi T. T., 2001, PASJ, 53, 433
Rowan-Robinson M., 2001, ApJ, 549, 745
Sawicki M., 2002, AJ, 124, 3050
Sawicki M., Webb T. M. A., 2005, ApJL, 618, 67
Sciama D. W., 2000, MNRAS, 312, 33
Scott S. E. et al., 2002, MNRAS, 331, 817
Scott S. E., Dunlop J. S., Serjeant S., 2005, MNRAS, submitted
Serjeant S. et al., 2003a, MNRAS, 344, 887
Serjeant S., Farrah D., Geach J., Takagi T., Verma A., Kaviani A., Fox M., 2003b, MNRAS, 346, L51
Serjeant S. et al., 2004, ApJS, 154, 118
Simpson C., Eisenhardt P., 1999, PASP, 111, 691
Simpson C., Dunlop J. S., Eales S. A., Ivison R. J., Scott S. E., Lilly S. J., Webb T. M. A., 2004, MNRAS, 353, 179
Smail I., Ivison R. J., Blain A. W., 1997, ApJ, 490, L5
Smail I., Ivison R. J., Blain A. W., Kneib J.-P., 2002, MNRAS, 331, 495
Smail I., Chapman S. C., Blain A. W., Ivison R. J., 2004, ApJ, 616, 71
Somerville R. S., Primack J. R., Faber S. M., 2001, MN 320, 504
Steidel C. C., Adelberger K. L., Giavalisco M., Dickinson M., Pettini M., 1999, ApJ, 519, 1
Swinbank A. M., Smail I., Chapman S. C., Blain A. W., Ivison R. J., Keel W. C., 2004, ApJ, 617, 64
Takagi T., Hanami H., Arimoto N., 2004, MNRAS, 355, 424
Teerikorpi P., 1997, ARA&A, 35, 101
Thompson R. I., Weymann R. J., Storrie-Lombardi L. J., 2001, ApJ, 546, 6945B
Tilanus R. P. J., 2004, JCMT Technical Report TR/001/106, http://www.jach.hawaii.edu/JACdocs/JCMT/tr/001/106
van Kampen E., 2003, in Plionis M., ed., Proceeding of the Multi-Wavelength Cosmology Conference. Mykonos, Greece, June 2003, (Kluwer) astro-ph/0310297
van Kampen E., Jimenez R., Peacock J. A., 1999, MNRAS, 310, 43
van Kampen E. et al., 2005, MNRAS, 359, 469
von der Heide, K., 1979, Optics Communications, 31, 279
Wang W.-H., Cowie L. L., Barger A. J., 2004, ApJ, 613, 655
Waskett T. J. et al., 2003, MNRAS, 341, 1217
Webb T. M. A., Lilly S. J., Clements D. L., Eales S., Yun M., Brodwin M., Dunne L., Gear W. K., 2003, ApJ, 597, 680
Webb T. M. A., Brodwin M., Eales S., Lilly S. J., 2004, ApJ, 605, 645
Wiedner M. C., Hills R., Carlstrom J. E., Lay O. P., 2001, ApJ, 553, 1036
Wiklind T., 2003, ApJ, 588, 736
Williams R. E. et al., 1996, AJ, 112, 1335
Wright E. L., Hinshaw G., Bennett C. L., 1996, ApJ, 458, L53
Yun M. S., Carilli C. L., ApJ, 2002, 568, 88

This paper has been typeset from a TEX/LATEX file prepared by the author.