Evolution of the Far-Infrared–Radio Correlation and Infrared SEDs of Massive Galaxies over $z = 0 – 2$

N. Bourne,1⋆ L. Dunne,1 R. J. Ivison,2,3 S. J. Maddox,1 M. Dickinson4 and D. T. Frayer5

1 School of Physics & Astronomy, University of Nottingham, University Park, Nottingham NG7 2RD
2 UK Astronomy Technology Centre, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ
3 Institute for Astronomy, University of Edinburgh, Blackford Hill, Edinburgh EH9 3HJ
4 National Optical Astronomy Observatory, Tucson, AZ 85719, USA
5 National Radio Astronomy Observatory, PO Box 2, Green Bank, WV 24944, USA

ABSTRACT

We investigate the far-infrared–radio correlation (FRC) of stellar-mass-selected galaxies in the Extended Chandra Deep Field South using far-infrared imaging from Spitzer and radio imaging from the Very Large Array and Giant Metre-Wave Radio Telescope. We stack in redshift bins to probe galaxies below the noise and confusion limits. Radio fluxes are $K$–corrected using observed flux ratios, leading to tentative evidence for an evolution in spectral index. We compare spectral energy distribution (SED) templates of local galaxies for $K$–correcting FIR fluxes, and show that the data are best fit by a quiescent spiral template (M51) rather than a warm starburst (M82) or ULIRG (Arp220), implying a predominance of cold dust in massive galaxies at high redshift.

In contrast we measure total infrared luminosities that are consistent with high star-formation rates. We observe that the FRC index ($q$) does not evolve significantly over $z = 0 – 2$ when computed from $K$-corrected 24 or 160-$\mu$m photometry, but that using 70-$\mu$m fluxes leads to an apparent decline in $q$ beyond $z \sim 1$. This suggests some change in the SED at high redshift, either a steepening of the spectrum at rest-frame $\sim 25 – 35\mu$m or a deficiency at $\sim 70\mu$m leading to a drop in the total infrared/radio ratios. We compare our results to other work in the literature and find synergies with recent findings on the high-redshift FRC, high specific star-formation rates of massive galaxies and the cold dust temperatures in these galaxies.

Key words: galaxies: high-redshift – galaxies: evolution – galaxies: ISM – infrared: galaxies – radio continuum: galaxies.

1 INTRODUCTION

One of the most exciting research areas in observational astronomy at this time is the field of far-infrared astronomy. Our understanding of extragalactic sources in the far-infrared (FIR) and sub-millimetre (sub-mm) to millimetre regimes has improved exceptionally over the past decade, thanks to such instruments as ESA’s Infrared Space Observatory (ISO, launched 1995; [Kessler et al., 1996]) and NASA’s Spitzer Space Telescope (launched 2003; [Werner et al., 2004]), alongside ground-based instruments such as the Sub-millimetre Common-User Bolometer Array (SCUBA; [Holland et al., 1999]) and the Max Planck Millimetre Bolometer Array (MAMBO; [Kreysa et al., 1998]), both commissioned in the late 1990’s. One of the most important FIR instruments to date is the Multiband Imaging Photometer for Spitzer (MIPS; [Rieke et al., 2004]), which, with Spitzer’s 0.85m mirror provides imaging with diffraction limited resolutions of 6, 18 and 40 arcsec in three broad bands centred at 24, 70 and 160$\mu$m respectively.

Recent advances in the resolution of the cosmic infrared background (CIB) by stacking into images from Spitzer (see e.g. [Dole et al., 2006; Dye et al., 2007; Chary & Pope, 2010]) and BLAST (The Balloon-born Large Aperture Sub-mm Telescope; [Marsden et al., 2009]) have enabled an improved understanding of the history of star formation and galaxy formation and evolution, and this will be further improved by ongoing work (such as [Berta et al., 2010]) with ESA’s new Herschel Space Observatory ([Pilbratt et al., 2010]).

Most of the stellar mass in the local universe is concentrated in the most massive galaxies ($M_\star \gtrsim 10^{11}M_\odot$; [Kauffmann et al., 2003]) and observations show that these
have been in place since \( z \sim 1 \) (Dickinson et al. 2003; Bundy, Ellis & Conselice 2005; Pérez-González et al. 2008; Taylor et al. 2007; Collins et al. 2006). The significant increase in density of luminous (\( L_{8-1000\mu m} \gtrsim 10^3 L_\odot \)) and ultra-luminous (\( L_{8-1000\mu m} \gtrsim 10^8 L_\odot \)) infrared galaxies (LIRGS and ULIRGs) from the local universe to \( z \sim 2 - 3 \) (e.g., Daddi et al. 2005a; Caputi et al. 2008) is thought to reveal the formation stages of these latter-day giants, which apparently formed in a remarkably short time between \( 1 \lesssim z \lesssim 3 \), in an antithesis to the paradigm of hierarchical structure formation (e.g., de Lucia et al. 2006).

In the FIR–sub-mm regime the dominant source of the extragalactic background light (after the cosmic microwave background) is thermal continuum emission from interstellar dust, which is mainly composed of polycyclic aromatic hydrocarbons (PAHs), graphites and silicates, typically less than \( \sim 0.25\mu m \) in size (Draine & Li 2007; Draine et al. 2007). The FIR emission in star-forming galaxies is thought to arise both from cold dust in the large-scale ‘cirrus’ component of the interstellar medium (ISM), and from warmer dust in and around star-forming regions (e.g., de Jong et al. 1984; Helou 1986). If there is a sufficient level of dust-shrouded star formation then the FIR emission is dominated by this ‘warm’ dust component, which has characteristic temperatures of around 30–50K (Dunne & Eales 2001; Sajina et al. 2006; Dye et al. 2007; Pascale et al. 2009). The dust is heated to these temperatures by the ultraviolet (UV) radiation field from hot O and B type stars, which are present only in regions of ongoing or recent star formation (‘recent’ meaning within the lifetime of these short-lived stars, \( \lesssim 10^9 \) yr; Kennicutt 1998). For this reason, the total IR luminosity (\( L_{\text{TIR}} = L_{8-1000\mu m} \)) can be used as a tracer of star-formation rates (SFRs) in galaxies, often using one or a number of FIR fluxes (such as the Spitzer MIPS bands) or sub-mm fluxes to estimate \( L_{\text{TIR}} \) (Kennicutt 1998).

One issue with using the FIR as a SFR tracer is the contribution from cold dust in the ISM, which is heated by older stars in the disk of the galaxy, and is therefore unrelated to star formation (Calzetti et al. 2010). Contamination of samples by galaxies hosting active galactic nuclei (AGN) is another problem, as AGN also emit UV radiation which can heat dust in the torus. AGN-heated dust is generally hotter than dust heated in star-forming regions, so the thermal spectrum peaks at a shorter wavelength, and mid-infrared (MIR) fluxes (including Spitzer’s 24-\( \mu m \) band, as well as shorter wavelengths) would be boosted. Mid-infrared fluxes are also affected more uncertainly by the emission features of PAH molecules, which are ubiquitous in star-forming galaxies (e.g., Leger & Puget 1984; Roche et al. 1992; Lutz et al. 1998; Alistarmandola, Hudgins & Sandford 1998), and the 10-\( \mu m \) silicate absorption trough, so the use of 24-\( \mu m \) fluxes as SFR indicators at high redshifts is subject to some contention (see e.g., Dale et al. 2002; Calzetti et al. 2007; Daddi et al. 2007; Papovich et al. 2007; Young, Bendo & Lucero 2009; Rieke et al. 2009).

Another part of a galaxy’s SED that can be used as a SFR indicator is the radio luminosity. Non-thermal radio continuum emission from star-forming galaxies originates from type II supernova remnants (SNRs), the endpoints of the same massive short-lived stars that heat the dust via their UV radiation. This connection with dust heating is important because it leads to the well-known (but not fully understood) FIR–Radio correlation (FRC; van der Kruit 1973; Rickard & Harvey 1984; Hellen, Soifer & Rowan-Robinson 1985; Condon 1992, etc.). The FRC is linear, remarkably tight and holds for a wide range of galaxy types over at least five orders of magnitude in luminosity (Yun, Reddy & Condon 2001). It can be explained in terms of ongoing star formation producing hot massive (\( M > 8M_\odot \)) stars: while the FIR flux is emitted from dust heated by these stars, the radio emission arises from synchrotron radiation by cosmic ray (CR) electrons accelerated in the SNRs of the dying stars. The non-thermal radio emission is smeared out through the galaxy as the relativistic CR electrons travel through the galaxy over lifetimes of \( \tau \sim 10^8 \) years, during which they emit synchrotron radiation via interactions with the galactic magnetic field (as described by Condon 1992, and shown observationally by Murphy et al. 2008). A shallower thermal component is also present in the radio spectrum due to bremsstrahlung radiation from electrons in H ii regions, but this becomes dominant only at high frequencies (\( \gtrsim 30 \) GHz) and at 1.4 GHz only comprises \( \lesssim 10\% \) of the radio flux (Condon 1992).

It is difficult however to explain the linearity and tightness of the correlation between thermal FIR luminosity and non-thermal radio luminosity. ‘Minimum energy’ estimates of magnetic fields in galaxies (Burbidge 1955) imply a large variation between normal galaxies and extreme starbursts like Arp220. To explain a constant FIR/radio ratio between such disparate systems, complex physical solutions need to be provided, for example invoking strong fine-tuning to regulate electron escape and cooling timescales, or short cooling timescales with magnetic fields \( \sim 10 \) times stronger than the ‘minimum energy’ argument suggests (Thompson et al. 2008).

Voelk (1989) first suggested a ‘calorimeter’ model whereby both UV light from massive stars and CR electrons from SNRs are (a) proportional to the supernova rate, and (b) efficiently absorbed and reprocessed within the galaxy, so that the respective energy outputs in FIR re-radiation and radio synchrotron would both be tied to the supernova rate. This theory requires a correlation between the average energy density of the radiation field and the magnetic field energy density. Voelk (1989) argues this is plausible if the origin of the magnetic field is a turbulent dynamo effect, since the turbulence would be largely caused by the...
activity of massive stars, and hence correlated with the supernova rate. Alternative non-calorimetric models include those of Helou & Bica (1992), using a correlation between disk scale height and the escape scale length for CR electrons; and Niklas & Beck (1997), in which the FRC is driven by correlations with the overall gas density and equipartition of magnetic field and CR energy. 

Eckl (2002) argues for a ‘conspiracy’ to diminish both the FIR and radio emission originating from star formation in low luminosity galaxies when compared with luminous \( \sim L_\star \) galaxies – without this the relationship would not remain linear over the full luminosity range. Similarly, the calorimeter model of Lacki, Thompson & Quataert (2010) invokes conspiracies in low and high gas surface density regions to maintain the relationship. The physical origin of the FRC therefore is still an open question. A full review of the theories is beyond the scope of this study, but a more detailed discussion of the literature can be found in Vlahakis, Eales & Dunne (2007), and a more in-depth treatment is provided in the numerical work of Lacki et al. (2010).

An ongoing strand of research at the current time is the investigation of the FRC at high redshifts and low fluxes, in particular whether there is any evolution (e.g. Appleton et al. 2004; Frayer et al. 2006; Ibar et al. 2008; Garn et al. 2009; Seymour et al. 2009; Ivison et al. 2010; Sargent et al. 2010a,b). Measurements of any evolution (or lack thereof) would improve the accuracy of FIR-/radio-estimated SFRs at high redshift, and could shed light on the mechanism governing the FRC, as well as highlighting differences in the physical and chemical properties of star-forming galaxies at high and low redshift (Seymour et al. 2004).

In the current work we investigate the FRC over a range of redshifts, for a sample that is not limited by FIR- or radio flux. Using Spitzer FIR data and radio data from the Very Large Array (VLA) and Giant Metre-Wave Radio Telescope (GMRT), we quantify the FIR–Radio Correlation as a function of redshift in massive galaxies selected from a near-infrared (NIR) survey of the Extended Chandra Deep Field South (ECDFS). We use Equation 1 to define the ‘\( q_{IR} \)’ index, which quantifies the FRC as the logarithmic ratio between a monochromatic FIR flux (\( S_{\nu, IR}, \) e.g. at 24, 70 or 160\( \mu \)m), and 1.4-GHz radio flux (\( S_{\nu, 1.4 \text{ GHz}} \)).

\[
q_{IR} = \log_{10} \left( \frac{S_{\nu, IR}}{S_{\nu, 1.4 \text{ GHz}}} \right)
\] (1)

We also investigate the effects of using different FIR bands to quantify the FRC, and the effects of assumptions about the SEDs of the galaxies in the sample. We employ a ‘stacking’ methodology to recover sufficient signal-to-noise ratios on faint objects to obtain measurements of the average properties of the sample. The data are described in Section 2, while the binning and stacking methodologies are described in Section 3. The analysis of SEDs and application of K- corrections is covered in Section 4 and the results are analysed and discussed in Section 5. A concordance cosmology of \( \Omega_m = 0.27, \Omega_\Lambda = 0.73, H_0 = 71 \text{ km s}^{-1} \text{ Mpc}^{-1} \) is assumed throughout.

2 DATA

The ECDFS is a \( \sim 0.25 \text{ deg}^2 \) square centred at \( 3\text{h}32\text{m}30\text{s}, -27\text{d}48\text{'}20\text{''} \) (J2000). It is a much-studied region of sky, with a rich body of published data and studies of extragalactic sources at a broad range of wavelengths stretching from X-ray to radio regimes. We use radio imaging data at 1.4 GHz from the VLA, with a typical rms sensitivity across the map of 8\( \mu \)Jy beam\(^{-1}\) and beam dimensions of \( 2.8 \times 1.8 \text{ arcsec} \) (Miller et al. 2008). Imaging data were obtained at 610 MHz from the GMRT, reaching a sensitivity of \( 40\mu \)Jy beam\(^{-1}\) with a 6.5 \times 5.4 arcsec beam (Ivison et al. 2010a). For the FIR, Spitzer MIPS images at 24, 70 and 160\( \mu \)m were obtained from the FIDEL survey (DR3; Dickinson et al., in preparation).

To look at a range of galaxy types over a range of redshifts we must give careful thought to how the galaxies are selected. For example, selecting radio-bright galaxies will naturally favour active radio galaxies, while selection at 24\( \mu \)m is likely to favour galaxies with dusty starbursts and/or obscured AGN components. These biases will affect the measured value of \( q_{IR} \) (see e.g. Sargent et al. 2010a). There is however a good body of evidence that distant massive galaxies in a range of phases of star-formation and nuclear activity can be effectively selected in NIR filters at \( \sim 2\mu \)m. This part of the spectrum is minimally affected by dust absorption, AGN and other components, and hence is relatively insensitive to the ‘type’ of galaxy or the shape of its SED. Furthermore it is insensitive to the age of the stellar population (hence SFR), because the light is dominated by old main-sequence stars that make up the bulk of the stellar mass in all galaxies. Thus NIR luminosity is primarily dependent on stellar mass only (Glazebrook et al. 1997; Gardner 1995).

In the ECDFS there exists NIR data from the Spitzer Infrared Array Camera (IRAC; Fazio et al. 2004), of which the two shortest wavelength bands (3.6 and 4.5\( \mu \)m) can also be used as an effective tracer of stellar mass (e.g. Serjeant et al. 2008). A catalogue of IRAC sources matched with optical–NIR photometry in the Multimwavelength Survey by Yale–Chile (MUSYC; Gawiser et al. 2006) is collected in Spitzer’s IRAC and MUSYC Public Legacy of the ECDFS (SIMPLE; Damen et al., in preparation). The catalogue was extracted from IRAC 3.6 and 4.5-\( \mu \)m images, and sources for which the mean of the 3.6 and 4.5-\( \mu \)m AB magnitudes \( (3.6 + 4.5)/2 < 21.2 \) were selected, giving a catalogue of 3841 sources. The source extraction and selection is described by Damen et al. (2009). In SIMPLE the IRAC sources have been matched to multiwavelength counterparts in the MUSYC catalogue, which contains photometry in UBVRi’z’JHK bands. Stars have been identified and excluded from the catalogue using the colour criterion \( J – K < 0.04 \), and potential AGN were removed by excluding any matches with Chandra X-ray sources (Virani et al. 2006). Photometric redshifts were collated for all objects in the sample from COMBO-17 (Wolf et al. 2004), and by using the EAZY code (Brammer, van Dokkum & Coppi 2008) as described by Damen et al. (2009). Damen et al. compared the photometric redshifts to spectroscopic ones where available, and showed that the median \( (z_{\text{spec}} - z_{\text{phot}})/(1+z) \) is 0.033 (0.079 at \( z \geq 1 \)). As described in Section 5, we divide the sample into bins with sizes \( \Delta z/(1+z) \sim 0.2 – 0.4 \), so it is safe
to neglect these photometric uncertainties. The final catalogue used in this work contains 3529 sources with photometric redshifts up to \( z = 2 \), in the region of the ECDFS defined by the rectangle \( 52^\circ 51' 48'' < \text{RA} < 53^\circ 25' 14'', -28^\circ 03' 27'' < \text{Dec} < -27^\circ 33' 22'' \) (J2000).

3 STACKING METHODOLOGY

3.1 Sample Selection and Redshift Binning

Since sources are selected by their NIR flux across a range of SED types, many are likely to be faint or undetectable at the wavelengths of interest. In order to probe the evolution of fluxes as a function of redshift, we stack galaxies into seven bins in redshift and measure median fluxes in each bin. The great advantage of this technique is the gain in signal-to-noise ratio, as combining many sources reduces the random noise while maintaining the average level of the signal. This gain is at the expense of knowledge of the individual galaxies, but with careful application of criteria when binning the galaxies, and with a large enough sample, it can reveal properties of galaxies below the noise and confusion levels. The technique has been used to great effect many times in the literature; for example by [Serjeant et al. 2004; Dole et al. 2006; Ivison et al. 2007; Takagi et al. 2007; White et al. 2007; Papovich et al. 2007; Dunne et al. 2009].

We do not know the distribution of fluxes in the stacks, but since we select massive galaxies with unknown SEDs at a range of redshifts, we may expect to be prone to some outliers. For example, radio-bright AGN have unusually high radio fluxes and are outliers on the FRC. We cannot be certain that these have been successfully removed from the sample by cross-matching with the Chandra catalogue, as we know that there is limited overlap between X-ray and radio-selected AGN samples [Rovilos et al. 2007; Pierce et al. 2010; Griffith & Stern 2010, although see Section 5.1]. We therefore used the median statistic to represent the properties of the typical galaxies in each stack, because unlike the mean, the median is resistant to outliers [Gott et al. 2001; White et al. 2007; Carilli et al. 2007; Dunne et al. 2009].

To study redshift evolution, we divided the sample according to the photometric redshifts in the catalogue, and stacked each bin into the radio and FIR images. The sample was split into the redshift bins given in Table 1.

The most significant sampling bias that we expect to see is that of stellar mass. The stellar masses of galaxies in the catalogue have been estimated by [Damen et al. 2009], by SED-fitting with a Kroupa (2001) initial mass function (IMF). The effect of Malmquist bias is that the median mass in the sample is lower at lower redshifts (because low mass galaxies dominate the population), but is higher at higher redshifts where only the more massive galaxies are detectable. This is illustrated in Fig. 1 and we test the effect on our results in Section 5.

### Table 1. Redshift bins and statistics of catalogue

| Bin | Boundaries | Median (z) | Count N |
|-----|------------|------------|---------|
| ALL | 0.00 \( \leq z \) < 2.00 | 0.73 | 3172 |
| ZB0 | 0.00 \( \leq z \) < 0.40 | 0.21 | 528 |
| ZB1 | 0.40 \( \leq z \) < 0.61 | 0.53 | 529 |
| ZB2 | 0.61 \( \leq z \) < 0.73 | 0.67 | 529 |
| ZB3 | 0.73 \( \leq z \) < 0.96 | 0.87 | 528 |
| ZB4 | 0.96 \( \leq z \) < 1.20 | 1.06 | 529 |
| ZB5 | 1.20 \( \leq z \) < 1.42 | 1.29 | 264 |
| ZB6 | 1.42 \( \leq z \) < 2.00 | 1.61 | 265 |

![Figure 1. Scatter plot of stellar masses in the catalogue as a function of photometric redshift, with the divisions between the redshift bins marked as dashed lines. Large circles mark the median mass and redshift in each bin.](image-url)

### 3.2 Stacking into the VLA and GMRT Radio Images

Radio images have pixel units of Janskys (Jy) per beam, and adhere to the convention whereby each pixel value is equal to the flux density of a point source located at that position. The assumption that the pixel value at the position of each catalogue object gives the correct radio flux density for that source is generally good, though requires a small correction to give the total integrated flux of the source. This integrated-flux correction accounts for any sources being extended over more than one beam, and also for any astrometric offset between the catalogue coordinates and the radio source position. Since we are stacking, it is suitable to consider the overall effect on the median stacked source, so the integrated-flux correction is calculated from the stacked “postage-stamp” image of the full sample. This image is created by cutting out a 41 pixel square centred on each source, and stacking the images by taking the median value of each pixel. The integrated flux is calculated

---

2 As an indication of the need to stack, we roughly estimated the minimum SFRs detectable at high redshift, using the noise levels in Table 2 and calculating SFRs that would be derived from 5σ detections using the methodology described in Sections 4 and 5. These are, from the MIPS and radio fluxes respectively, 870 & 130M⊙yr\(^{-1}\) at \( z = 1 \), and 4900 & 870M⊙yr\(^{-1}\) at \( z = 2 \).

3 41 pixels corresponds to 49.2, 164, 328, 20.5 and 61.5 arcsec in the 24-µm, 70-µm, 160-µm, 1.4-GHz and 610-MHz bands respectively.
Figure 2. Postage-stamp images of the stacked targets in (from left to right) 24µm, 70µm, 160µm, 1.4 GHz, 610 MHz. The top row is the stacked full sample; other rows are the stacks of redshift bins ZB0–ZB6 (from top to bottom) as described in Section 3.1. Images are coloured by flux in Jy pixel$^{-1}$ units (MIPS) and Jy beam$^{-1}$ units (radio), and colour scales are included below each. Contours of [3, 5, ... 17, 19] times the pixel noise are overlaid. Each tile is a square 41 pixels on a side, which corresponds to 49.2, 164, 328, 20.5 and 61.5 arcsec for the five bands respectively.
using the AIPS package JMFIT and the correction is simply the ratio of this to the value of the peak (central) pixel in the image.

A higher integrated/peak ratio might be expected in the lowest-redshift bins if a large number of resolved sources were included in the stack. We therefore tested whether the correction varied significantly between different redshift bins, by constraining the beam centroid position and orientation and measuring integrated-/peak-flux ratios in each postage stamp. Without these constraints the fitted fluxes would be more prone to flux boosting by noise peaks, as the signal-to-noise in these postage stamps is low. In both 1.4 GHz and 610 MHz however, the result for each of the bins was consistent (within 1σ) with the result for the stack of all sources, indicating that the stacked sources were unresolved in all bins. Hence for all bins we used the integrated-flux corrections from the full stack (which have the smallest errors), i.e. 1.55 ± 0.08 for 1.4 GHz and 1.05 ± 0.12 for 610 MHz.

A further correction would need to made to point-source fluxes measured in the 1.4 GHz image to account for bandwidth smearing (BWS), an instrumental effect caused by the finite bandwidth of the receiver resulting in sources appearing more extended with increasing angular distance from the centre of the pointing. Since integrated flux is conserved in BWS, the effect is corrected in our data by the integrated-flux correction (this explains the large correction at 1.4 GHz which would otherwise appear to be inconsistent with unresolved sources).

### 3.3 Stacking into the Spitzer FIDEL Images

Measuring fluxes in the MIPS images requires a different technique, due to the large point-spread function (PSF) which results from the diffraction-limited resolution of MIPS. The centre of the PSF can be described by a roughly Gaussian profile, with full widths at half maximum (FWHM) of 6, 18 and 40 arcsec in the 24, 70 and 160-µm bands respectively (Rieke et al. 2004). The outer portion of the 24-µm PSF is less predictable and it is known to vary between different pointings and different source colours. For this reason, and to allow for potentially resolved sources, we chose to measure 24-µm fluxes by aperture photometry, and adopted an aperture of radius 13 arcsec (the radius for aperture photometry at 24µm) to measure total fluxes in Jy. Due to poorer resolution in the 70 and 160-µm maps, it is sufficiently accurate to measure point-source fluxes by applying a correction to the central pixel value: the factors used were 43.04 and 46.86 at 70 and 160-µm respectively. This converts the fluxes to units of Jy beam”⁻¹, and accounts for large-scale emission in the wings of the PSF, as well as a colour correction. No further correction is required to measure total (integrated) fluxes as we can confidently assume that none of the sources is larger than the beam in these two bands.

In stacking the FIDEL images it was necessary to exclude objects close to the edges of the map where the noise was higher, to ensure that noise in the stacks reduced as 1/√N, and to prevent gradients being introduced into the postage stamps. This was achieved by placing lower limits on integration time. Limits were chosen based on stacks of random positions, resulting in the exclusion of 3.5, 8.6 and 9.9 percent of the 3529 catalogue sources in the 24, 70 and 160-µm bands respectively. Because these cuts are based on integration time alone, there is no correlation with the nature of the sources themselves, so no systematic effect on the measured properties of the galaxies will be introduced.

Postage-stamp images of the stacked bins in the MIPS and radio maps are shown in Fig. 2, including noise contours as described in the following section.

#### 3.4 Analysis of Random Errors in Stacked Flux Measurements

Random errors in any flux measurement arise from noise in the image. Simplistically these errors might be expected to arise purely from the variance of pixel values in the map, σ², and the error on N stacked measurements is then given by σ/√N. This assumption is valid for the radio images, so it is sufficient to use the rms values at the corresponding positions on the rms map. In the MIPS data maps however, pixel covariance provides a non-negligible contribution to the error, so the rms maps are not sufficient. In order to measure the total random error on a measured flux we chose random positions in the sample region of the map and selected those that fell on empty regions of sky. This was tested by taking an aperture of radius 13 arcsec (the radius for aperture photometry at 24µm) around each position and measuring the standard deviation of pixel values in that aperture. If the aperture contained any pixels that deviated from the aperture mean by more than 2.5σ, then the position was discarded. The positions were also required to be separated and not overlapping. Thus positions were chosen to represent regions of empty sky with no sources. The number of positions used was chosen to be 500, to roughly match the sample size of bins used for stacking sources. For the 24µm case, where aperture photometry was used, the fluxes at the 500 positions were measured in exactly the same aperture as was used for source photometry, and the standard deviation of these sky fluxes was taken to represent the random error on an individual aperture measurement. Repeating the analysis with different random catalogues of varying sizes produced consistent results, and the distribution of flux values in the random catalogues was verified to be Gaussian with high certainty.

The analysis was repeated for the other two MIPS maps and also both radio maps, using the central (brightest) pixel for flux measurement since this was the method used for source photometry in those bands. The resulting error or noise values are given in Table 2. The radio error values were close to the average value in the rms map, confirming that the radio error is equal to the rms value. These values were multiplied by the aperture correction or integrated-flux correction in each case to represent the error on a flux measurement in Janskys. We checked that the noise in a
Figure 3. Results of measuring noise in stacks of empty sky positions. The noise was calculated by stacking $N$ random positions, repeating 500 times, and taking the standard deviation of the 500 stacked fluxes (see text for details). Results of these stacks are shown as symbols, and are in the same units as in Table 2. Lines show the least-squares fit to the results of each band with a power-law index of $-0.5$, representing the expected reduction of noise as $1/\sqrt{N}$ for each band. The agreement between results and expectations is good.

In using the median to represent the fluxes of $N$ sources in some bin, we must also consider the width of the distribution of fluxes in that bin: if this is larger than the estimated measurement error $1/\sqrt{N}$ then the latter is a poor indicator of uncertainty on the quoted median. For this reason we estimated 1σ uncertainties on the median following the method of Gott et al. (2001) and compared them to the estimated measurement error in each stack (the value in Table 2 divided by $\sqrt{N}$). At 160µm, 610 MHz and 1.4 GHz we found the two to be about equal (see Table 5), confirming that the flux errors we have estimated cover the distribution of fluxes in the bins. At 24µm the uncertainty on the median was around three times the size of the estimated errors, and at 70µm around twice the size, indicating that in these bands the flux distribution in each bin was somewhat broader than the estimated errors allowed for. It is possible that the method described in preceding paragraphs systematically underestimates the noise in these images, as a result of the constraints used to identify empty ‘sky’ apertures. Those constraints were designed to distinguish true readnoise on the detector from confusion noise in the sky, but the 24-µm image in particular is highly confused, meaning that the constraints could lead to correlation in the empty apertures stacked, and increase the chance of underestimating the noise. For the analysis of stacked results in this paper we therefore quote the uncertainties on the median following the Gott et al. (2001) method.

Figure 4. Autocorrelation function $W(\theta)$ of the SIMPLE catalogue. The dashed line is a power-law fit with index -1.15. Error bars are twice the Poisson errors, as described in the text.

3.5 Background Subtraction and Clustering Analysis

A similar methodology to the random error analysis was used to measure the background value to be subtracted. The method described above chooses empty apertures containing just sky, which is simplistically what needs to be subtracted before performing aperture photometry, but in the case of MIPS the combination of high source density and low resolution require that source confusion is also accounted for. When stacking, the random boosting of fluxes on individual sources will average out to a constant correction that can be included in the background. Hence when measuring the background for subtraction, a catalogue of random positions were chosen and stacked, without any criteria on the existence or otherwise of sources close to these random positions. On average the random catalogue should coincide with sources with the same probability as the source catalogue does. In this context we use the term ‘sources’ in the general sense, meaning any source of flux in the images that we stack into, i.e. any object that could boost a measured flux at a given position. At this stage we are making the assumption that source clustering does not play a part. Stacks of 3500 random positions (to match the sample size of the source catalogue) were made and repeated 1000 times in each of the three MIPS and two radio maps. The mean of the 1000 stacked fluxes was taken to be the background value, and the standard error was taken to be the uncertainty; results are given in Table 2.

Any clustering of the sources in the catalogue would lead to an increased probability of confusion for a catalogue source compared with a random position, hence with increased clustering the background subtraction becomes increasingly less effective. In order to estimate the size of this effect we would ideally need to understand the correlation function of sources in each image on scales smaller than the beam size. Since this is not possible, we made the assumption that the correlation of sources in the images that we stack into is approximately the same as that in the source catalogue. This may not fully account for confusion if the sources in the image are more clustered than the IRAC (cat-
linear regression given by fused sources to a measured flux, hence a correction factor. We used the empirical xFLS PRF available on the SSC website:

\[ \text{error on a single flux measurement; for } 24\mu m \text{ this is the noise on a corrected aperture flux, for } 70 \text{ and } 160\mu m \text{ it is the noise on a corrected point source flux, and for the radio it is the noise in a beam and does not include the integrated-flux correction. Background levels are in the same units, these are the values subtracted from the median source fluxes. Errors on background fluxes are standard errors from 1000 measurements as described in Section 6.}

Table 2. Information on the FIDEL and radio images. Measured total noise values represent the 1σ error on a single flux measurement; for 24µm this is the noise on a corrected aperture flux, for 70 and 160µm it is the noise on a corrected point source flux, and for the radio it is the noise in a beam and does not include the integrated-flux correction. Background levels are in the same units, these are the values subtracted from the median source fluxes. Errors on background fluxes are standard errors from 1000 measurements as described in Section 6.

| Band    | Pixel scale | PSF/Beam FWHM | Noise level | Background level |
|---------|-------------|---------------|-------------|------------------|
| 24µm    | 1.2         | 5.9           | 62          | −37.00 ± 0.04    |
| 70µm    | 4.0         | 18            | 1,200       | +2.2 ± 0.8       |
| 160µm   | 8.0         | 40            | 20,000      | +2,000 ± 10      |
| 1.4 GHz | 0.5         | 2.8 × 1.5     | 8.83        | −0.014 ± 0.005   |
| 610 MHz | 1.5         | 7.7           | 71.9        | −0.01 ± 0.03     |

Table 2

- The cross-correlation function \( W(\theta) \) is given by Equation 3: a convolution of the point-response function (PRF) with the 13-arcsec radius aperture, to give a curve of growth which represents the contribution of a background source to the aperture as a function of angular separation \( \theta \). This function is then substituted for the Gaussian beam profile in Equation 4.

- The robustness of the results was tested by checking against the method of Maseddi et al. (2006), which gave indistinguishable results.

Thus we calculated the average across all the data centres, of the excess probability of confusion with any of the reference centres. The correction to stacked flux was then calculated in the same way as described above, using \( W_{D,E} \) in Equation 4. It should be noted that using this estimate of the fractional contribution involves the implicit assumption that the average flux of background sources is equal to the average stacked flux. Since we can only correct for confusion with catalogue sources by this method (i.e. to avoid double-counting) this is a reasonable assumption.

- We calculated the corrections for three redshift ranges, shown in Table 3 by grouping the bins as follows: 0.0 \( \leq \) \( z \) \( < \) 0.6; 0.6 \( \leq \) \( z \) \( < \) 1.0; 1.0 \( \leq \) \( z \) \( < \) 2.0. Errors in the table were calculated using standard formulae for the propagation of errors, with the error bars on \( W_{D,E}(\theta) \) as shown in Fig 4. The results in the table are not surprising; 160µm has the lowest resolution therefore the greatest confusion, 24µm suffers more than 70µm because aperture photometry is used, and the radio images have sufficiently high resolution to largely avoid confusion.

\[ W_{D,D}(\theta) = \frac{DD - 2 DR - RR}{RR} \]  
Results are shown in Fig. 1 which includes a fit by linear regression given by \( W(\theta) = 0.000269 \theta^{-1.15} \), where \( \theta \) is in degrees. By dividing the region into four equal quadrants and comparing the scatter between results in each, we found that the standard error was a factor 2.0 larger than the simplistic Poisson errors. We therefore quote error bars on all correlation functions of twice the Poisson error.

The strong clustering implies a significant correction to the measured fluxes from stacking catalogue sources. The correction accounts for the flux contribution from any background sources separated by some angular distance \( \theta \) from the target. This contribution, as a fraction of the average source flux, is given by Equation 3, a convolution of the correlation function \( W(\theta) \) with the beam profile for the corresponding band (assumed to be Gaussian, \( \exp\left(-\theta^2/2\sigma^2\right) \), with \( \sigma = \text{FWHM}/(2\sqrt{2\ln 2}) \), scaled by the number density of background sources \( n \).

\[ F = n \int_0^\infty W(\theta)e^{-\theta^2/2\sigma^2}2\pi\theta d\theta \]  
This equation gives the average contribution of confused sources to a measured flux, hence a correction factor of \( 1/(1+F) \) must be applied to stacked fluxes.

For the 24-µm case a slightly different convolution must be used because aperture photometry is used. The contribution of a background source to a flux measurement then depends not only on where it falls on the beam profile, but on how much of its beam falls within the aperture. We computed the convolution of the 24-µm point-response function (PRF) with the 13-arcsec radius aperture, to give a curve of growth which represents the contribution of a background source to the aperture as a function of angular separation \( \theta \). This function is then substituted for the Gaussian beam profile in Equation 4.

This method corrects a stacked flux using the average probability of confusion from another source at separation \( \theta \), scaled by the amount of flux expected from a distance \( \theta \) from the centre of the beam. When correcting stacks of individual redshift bins, we must assume the same level of clustering in each bin if we are to use the autocorrelation of the full catalogue. To account for the probability of confusion of a target from a particular redshift range, while accounting for the contribution from background sources at all redshifts, we must consider the cross-correlation of the sources in the particular range (the ‘data’ centres, \( D \)) with the full catalogue (the ‘reference’ centres, \( E \); see Fig. 5). A modification of the Landy & Szalay (1993) method was used to calculate the cross-correlation function \( W_{D,E}(\theta) \), given by Equation 4.

\[ W_{D,E}(\theta) = \frac{DE - 2 DR - RR}{RR} \]  

We used the empirical xFLS PRF available on the SSC website: http://ssc.spitzer.caltech.edu/mips/psf.html.
Figure 5. Comparing cross-correlation functions of three redshift ranges (D) with the reference catalogue (E) being the full SIMPLE catalogue: (a) \( W_{D,E}(\theta) \) where \( D \) is the subset in \( 0.0 \leq z < 0.6 \); (b) \( W_{D,E}(\theta) \) where \( D \) is the subset in \( 0.6 \leq z < 1.0 \); (c) \( W_{D,E}(\theta) \) where \( D \) is the subset in \( 1.0 \leq z < 2.0 \). On each of (a)-(c) the black line shows the fit to the autocorrelation function of the full catalogue for comparison. Error bars are twice the Poisson errors, as described in the text. The power-law fits to the four functions are shown in (d), where the slope has been fixed by the fit to the full catalogue.

Table 3. Correction factors (C) to stacked source fluxes to remove contribution from correlated background sources, using autocorrelation of full catalogue and cross-correlations of three redshift ranges with the full catalogue. Errors were calculated as described in the text. Corrected flux \( S_{\text{stack,corr}} = S_{\text{stack}} \times C \)

| Band     | 0.0 ≤ z < 0.6 | 0.6 ≤ z < 1.0 | 1.0 ≤ z < 2.0 | All       |
|----------|---------------|---------------|---------------|-----------|
| 24µm     | 0.86 ± 0.05   | 0.74 ± 0.04   | 0.80 ± 0.04   | 0.79 ± 0.03 |
| 70µm     | 0.90 ± 0.04   | 0.80 ± 0.03   | 0.86 ± 0.04   | 0.84 ± 0.02 |
| 160µm    | 0.81 ± 0.05   | 0.66 ± 0.04   | 0.74 ± 0.05   | 0.72 ± 0.03 |
| 1.4 GHz   | –             | –             | –             | 1.000 ± 0.001 |
| 610 MHz   | –             | –             | –             | 1.000 ± 0.002 |

4 GALAXY SEDS AND K–CORRECTIONS

4.1 Radio K–Correction

Observed fluxes were converted to rest-frame (emitted) monochromatic luminosities using Equation (5) which contains a bolometric K-correction \( K(z) \), accounting for the shift of the spectrum in relation to the receiver, and a further bandwidth correction \([1+z]^{-1}\), accounting for the stretching of the spectrum in relation to the bandwidth of the receiver (\( d_L \) is the luminosity distance to the source, while \( z \) is its redshift).

\[
L_{\nu,\text{em}} = 4\pi d_L^2 S_{\nu,\text{obs}} K(z) [1+z]^{-1}
\] (5)

The radio spectrum can be assumed to follow a simple
power law \( S_\nu \propto \nu^\alpha \) resulting from the sum of the non-thermal synchrotron and thermal bremsstrahlung components; the power law index is typically \( \alpha \approx -0.8 \) for star-forming galaxies (Condon 1992), although steeper indices might be expected in AGN-dominated sources (Ibar et al. 2004). The \( K \)–correction to a monochromatic flux with a power law spectrum is given by Equation 6, which is independent of the filter transmission function.

\[
K(z) = [1 + z]^{-\alpha}
\]

The radio spectral index for each bin was evaluated using the stacked fluxes in the two radio bands, \( S_{610\, \text{MHz}} \) and \( S_{1.4\, \text{GHz}} \) in Equation 7 (which follows from \( S_\nu \propto \nu^\alpha \)), and is plotted in Fig. 4.

\[
\alpha = \frac{\log \left( \frac{S_{610\, \text{MHz}}}{S_{1.4\, \text{GHz}}} \right)}{\log (610/1400)}
\]  

These spectral indices were used to \( K \)–correct each measured radio flux using Equation 6, taking the observed median index for each bin to calculate \( K \)–corrections for all sources in that bin.

In Fig. 4 we note an apparent evolution to steeper radio slopes at increasing redshift in our sample. A linear least-squares fit to the \( \alpha(z) \) values gives a slope of \(-0.39 \pm 0.15\); the slope is non-zero at the 2.6\( \sigma \) level. This apparent trend is an unexpected result, and it is noteworthy that it was not observed in the stacked 24-\( \mu \)m sample of Ivison et al. (2010a), who used the same radio data and stacking technique; although their spectral indices do cover a similar range. The possible implications are discussed in Section 5.1 of this paper, but we note that using a single spectral index of \(-0.74\) for \( K \)–corrections leads to a slight rise in the \( q \) indices in the three highest redshift bins (a change of \( \delta q = +0.17 \) for the last bin at \( \langle z \rangle = 1.6 \)).

### 4.2 Infrared \( K \)–Correction

In the mid-/far-infrared part of the spectrum sampled by the MIPS bands, the assumption of a simple power law is not valid and \( K \)–corrections must be calculated by evaluating Equation 8, which defines the \( K \)–correction as the ratio of intrinsic luminosity to observed for a general filter transmission profile \( T_\nu (\nu) \).

\[
K(z) = \frac{\int_0^\infty T_\nu (\nu)L_\nu (\nu)\,d\nu}{\int_0^\infty T_\nu (\nu)L_\nu (\nu[1+z])\,d\nu[1+z]} \tag{8}
\]

This requires knowledge of both the filter transmission function \( T_\nu (\nu) \) and the SED \( L_\nu (\nu) \). A well-studied local galaxy can be used as a template for high-redshift galaxies; commonly used templates in FIR studies include Arp220 and M82, which are IR-luminous and therefore considered to be

---

8 The transmission functions for the MIPS filters are available on the Spitzer Science Center website at [http://ssc.spitzer.caltech.edu/mips/spectral_response.html](http://ssc.spitzer.caltech.edu/mips/spectral_response.html)
Evolution of the FIR–Radio Correlation and IR SEDs

5 RESULTS AND DISCUSSION

5.1 Evolution of Radio Properties of the Sample

Our results indicate a significant increase in radio luminosity with redshift (see Table 4), which is to be expected if the radio emission is related to star formation, due to the increase in star-formation activity in the most massive galaxies from the local universe back to \( z \sim 2 \). The apparent evolution in radio spectral index over the redshift range (Fig. 6) is more surprising and, notwithstanding the large error bars, hints at a fundamental change in the sample demographic, with different sources dominating the radio luminosity at \( z < 1 \) and \( z > 1 \) respectively. The most likely potential contaminant is radio flux from AGN, which would have a different spectral index than that from star formation, and would also have the effect of boosting radio luminosity. Radio-loud AGN source counts are known to evolve strongly at \( z > 1 \) (e.g. [Wall et al. 2003]).

The effect that AGN contamination would have on the median spectral indices is not entirely straightforward. While flat (\( \alpha \gtrsim -0.5 \)) spectra are associated with radio-quiet quasars or low-luminosity AGN ([Bondi et al. 2007]; [Huynh, Jackson & Norris 2009]), steep spectra (\( \alpha < -1 \)) have frequently been used to select powerful radio galaxies at high redshift (generally \( z \gtrsim 2 \); e.g. [de Breuck et al. 2004]; [Pedani 2002]; [Cohen et al. 2004]). This is because AGN radio spectra are flat at low frequencies but steepen at high frequency, hence steep slopes are observed when the spectrum is highly redshifted (the frequency of the turnover varies, depending on properties such as magnetic field strength and electron density; [Huynh et al. 2007]). The evolving spectral indices seen in Fig. 6 could therefore be a sign of AGN dominating the radio signal at higher redshifts.

Matches with the Chandra X-ray catalogue ([Virani et al. 2004]) have been removed from the sample, reducing the likelihood of contamination from unobscured AGN. But overlap between X-ray and radio AGN is known to be small (e.g. [Pierce et al. 2010]). In order to identify any AGN that are obscured or undetected in X-rays, we looked to the \( \text{MIR} \) fluxes from the IRAC catalogue. IRAC colours have been shown to provide some limited diagnostics for selecting obscured AGN based on the rest-frame MIR slope (e.g. [Lacy et al. 2004]; [Stern et al. 2004]; [Alonso-Herrero et al. 2006]; [Donley et al. 2007, 2008]).

In Fig. 10 we plot \( S_{\lambda_{\text{obs}}} / S_{\lambda_{\text{obs}}} \) against \( S_{\lambda_{\text{obs}}} / S_{\lambda_{\text{obs}}} \) for all the objects in the sample. In this plot, the majority of objects have \( S_{\lambda_{\text{obs}}} / S_{\lambda_{\text{obs}}} < 1 \), well separated from the region occupied by AGN at intermediate to high redshifts.
used where they are $\geq$ the median 1

The representative error bar in the top right shows limit (indicated by arrows). IRAC fluxes are from the SIMPLE

duced by confusion at 24 $\mu$m, but positions in the horizontal axis are reliable as signal-to-noise is good in both 4.5$\mu$m and 8.0$\mu$m and confusion noise is much lower. Tracks of Arp220 and Mrk231 are overlaid in grey, showing the locus of each on the diagram as a function of redshift between 0 and 2. This plot can be used as an AGN/starburst diagnostic since AGN have been shown to lie to the right, with $S_{8.0\mu m}/S_{4.5\mu m} > 2$ at intermediate to high redshifts (as the prototype Mrk231 does) while starbursts (such as Arp220) lie above and to the left (e.g. Ivison et al. 2004, Pope et al. 2008, Coppin et al. 2010). With the exception of bin ZB0 ($0 \leq z < 0.4$), our data lie well to the left of the plot, indicating flat spectral slopes at $< 8.0\mu$m (compared with Arp220) and negligible AGN contamination. The scatter in bin ZB0 is attributed to strong PAH emission at 7.7$\mu$m which makes the diagnostic unreliable at low redshift (although it is noted that strong PAH emission is generally associated with star formation and not AGN).

The width of the redshift bins and the use of photometric

detected at 24$\mu$m, we use the 5$\sigma$ upper limit (indicated by arrows). IRAC fluxes are from the SIMPLE
catalogue. The representative error bar in the top right shows the median 1$\sigma$ errors. Some scatter in the vertical axis is introduced by confusion at 24$\mu$m, but positions in the horizontal axis are reliable as signal-to-noise is good in both 4.5$\mu$m and 8.0$\mu$m and confusion noise is much lower. Tracks of Arp220 and Mrk231 are overlaid in grey, showing the locus of each on the diagram as a function of redshift between 0 and 2. This plot can be used as an AGN/starburst diagnostic since AGN have been shown to lie to the right, with $S_{8.0\mu m}/S_{4.5\mu m} > 2$ at intermediate to high redshifts (as the prototype Mrk231 does) while starbursts (such as Arp220) lie above and to the left (e.g. Ivison et al. 2004, Pope et al. 2008, Coppin et al. 2010). With the exception of bin ZB0 ($0 \leq z < 0.4$), our data lie well to the left of the plot, indicating flat spectral slopes at $< 8.0\mu$m (compared with Arp220) and negligible AGN contamination. The scatter in bin ZB0 is attributed to strong PAH emission at 7.7$\mu$m which makes the diagnostic unreliable at low redshift (although it is noted that strong PAH emission is generally associated with star formation and not AGN).

5.2 The Observed and $K$-corrected FIR–Radio Correlation as a Function of Redshift

The $q$ index was calculated for each stack and for each FIR band $i$, substituting the monochromatic FIR flux $S_{i}$ (Jy) for $S_{10}$ in Equation 1. Details of the stacked fluxes and $q$ ratios can be found in Table 5 at the end of this paper. Fig. 10 displays the calculated $q$ values as a function of redshift, both before and after $K$-corrections were applied. Fluxes were $K$-corrected individually, rather than after stacking, using the photometric redshift of each source, and for the radio, the stacked spectral index measured in the corresponding bin, and for the FIR, the corrections shown in Fig. 5.

The left panel of Fig. 10 reveals a very slight downward trend of $q$ in each of the three MIPS bands, with a more bumpy evolution in $q_{24}$. The anomalies in the observed $q_{24}$ evolution can be explained in terms of the MIR SED, which in star-forming galaxies often contains broad PAH emission features at 3.3, 6.2, 7.7, 8.6, 11.3 and 12.7$\mu$m (e.g. Roche et al. 1991, Genzel et al. 1998, Armus et al. 2002). These redshifted features can account for the boosting of $q_{24}$ in the bins centred at redshifts 0.9 and 1.6. Similarly the dip in $q_{24}$ at redshift 1.3 can be attributed to the broad 10-$\mu$m silicate trough redshifted into the 24-$\mu$m passband.

The width of the redshift bins and the use of photometric...
Evolution of the FIR–Radio Correlation and IR SEDs

Figure 10. Far-infrared–radio relation ($q$) as a function of redshift for the three MIPS bands (from top to bottom, 24µm, 70µm, 160µm). Ratios of stacked observed fluxes are plotted on the left. On the right we show ratios of stacked $K$–corrected fluxes: radio fluxes are $K$–corrected using the measured spectral index for each bin, and infrared fluxes are $K$–corrected using the four templates described in Section 4.2. Vertical error bars represent estimated 1σ uncertainties on the stacked medians, following Gott et al. (2001); they indicate the spread of the data, and are of similar size or larger than the noise, as described in Section 3.4. Horizontal bars in the left-hand panels mark the full width of each bin. Horizontal bars are omitted from the right-hand panels for clarity.

Redshifts accounts for the breadth of redshifts over which these features appear to have an effect.

The apparent evolution of $q$ with redshift that is expected from various SED templates has been plotted over observed data by several authors including Ibar et al. (2008), Seymour et al. (2009) and Sargent et al. (2010a). In this work we choose instead to plot the $K$–corrected $q$–values in an attempt to emphasise the ‘excess’ evolution that may reveal intrinsic changes in the rest-frame flux ratios of galaxies within the sample at different redshifts.
the least evolution in all three q indices. K–corrections using the Arp220 template give rise to an increasing q21, due to the steeper MIR slope (\( \lambda \lesssim 24\mu m \)). In q160 the Arp220, Mrk231 and M82 templates all exacerbate the downward trend towards high redshift, while the M51 template removes it, as a result of the cooler dust temperature (longer wavelength of the peak) in M51. However, none of the templates removes the trend in q70, and this could be attributed to a real evolution in rest-frame flux ratios, or a steeper spectral slope at \( \lambda \lesssim 70\mu m \) in the galaxies sampled, in comparison to the templates chosen.

5.3 Infrared Spectral Energy Distributions

Interpretation of our results is evidently subject to the assumptions made about the ‘average’ or typical SED of the sample. It is possible to better constrain the FIR K–correction by analysing the evolution of MIPS flux ratios (colours) as a function of redshift. These colours are sensitive to the position of the peak of the thermal dust emission, hence the temperature of the emitting dust, as well as the slope of the SED on the short-wavelength side of the peak.

Fig. 11 shows the evolution in observed MIPS colours with redshift, plotted over the expected tracks for each of the SED templates, and reveals that the SED most consistent with observed colours at all redshifts is M51.

The important factor distinguishing the M51 template from the others used is the position of the peak of the SED at a longer wavelength. M51 is a quiescent star-forming galaxy with an IR SED dominated by cold dust, and evidence from Fig. 11 therefore points to a cold dust temperature for the galaxies in our sample, at least in the first three redshift bins. The 70–160-\( \mu m \) colour is directly sensitive to the position of the peak at low redshifts, but it is clear from the middle panel of Fig. 11 that over the last four bins the M51 and Arp220 templates are barely distinguishable in this colour space, so we cannot draw conclusions on the dust temperatures at \( z \gtrsim 0.8 \). This is because at these redshifts both bands are shortward of the peak of even the hottest IR SED, and probe the slope on the Wien side. The 70–160-\( \mu m \) colours of the high-redshift bins are consistent with the steeper slopes of M51 and Arp220, and not with the shallower slopes of M82 and Mrk231 (similarly the 24–70-\( \mu m \) and 24–160-\( \mu m \) colours rule out Arp220, due to its strong PAH emission). These steeper slopes are potentially an indication of a stronger contribution from ‘cold’ dust (in the ambient ISM) relative to ‘hot’ dust (in Hii regions associated with star formation) or a dearth of emission from very small grains (VSGs); alternatively they could even result from extremely optically thick systems where the SED is steepened by MIR dust attenuation. In this case, however, we would expect to see a stronger 10-\( \mu m \) silicate absorption feature such as that evident in the Arp220 K–correction at \( z \sim 1.5 \). Our stacked colours are not consistent with such a strong absorption which reduces the likelihood that optically thick MIR emission is responsible for the steeper rest-frame MIR slope at high redshift.

Cold dust temperatures are nevertheless consistent with the conclusions of studies such as Chapman et al. (2003) and Pope et al. (2004, 2008) for high-redshift sub-mm galaxies (SMGs), and Symeonidis et al. (2009), Sevmore et al. (2010) and Giovannoli et al. (2010) for 70-\( \mu m \)-selected galaxies at \( z \lesssim 1 \). There are also parallels with a recent detailed study of two massive K–selected galaxies at \( z \sim 2 \) by Muzzin et al. (2010), who fitted SEDs to data from Spitzer, BLAST and LABOCA (Siringo et al. 2010) instruments. Their best fits were star-formation-dominated SEDs with \( L_{TIR} \sim 10^{13}L_{\odot} \), but with cold dust temperatures, in contrast to ULIRGs in the local universe. Similar cool SEDs have also been determined for SMGs out to \( z \sim 1 \) from BLAST and Herschel studies (Dye et al. 2009, Amblard et al. 2010).

For our sample of massive galaxies, we expect to probe the epoch of stellar mass buildup at \( z > 1 \). Indeed, Table 4 shows that both radio– and IR-derived SFRs in our bins do reach high values beyond this redshift. It seems a reasonable assumption that the IR and radio luminosities are dominated by star-forming activity, since we do not expect a significant contamination from AGN-dominated galaxies in the sample (see Section 5.1). We observe therefore that despite the tendency towards higher luminosities (and SFRs) in the sample at increasing redshifts, there is no evidence for a change in the SED towards the templates of local high-SFR galaxies such as Arp220 or M82.

In Table 4 we show indicative TIR luminosities, derived for each bin using the rest-frame FIR luminosities in the

![Figure 11](image-url)
MIPS bands, scaled up to $L_{8-1000\mu m}$ assuming the M51 template. This was done by using the luminosities in the three MIPS bands simultaneously to find the best-fitting normalization of the M51 template. There will be some systematic uncertainties in the calibration, and for an idea of the size of these we consider another method to estimate $L_{\text{TIR}}$ from MIPS luminosities. Dale & Helou (2002) offer one such formula for $L_{8-1000\mu m}$, calibrated for a large sample of normal star-forming galaxies with a range of morphologies, colours and FIR luminosities (see also Dale et al. [2001]), we therefore consider it appropriate for M51–like SEDs. The uncertainty on this calibration was shown to be $\sim 25\%$ by Draine & Li (2007), and we find that using Dale & Helou’s method yields values well within $25\%$ (typically 6%, but as much as 16% for the highest-redshift bin) of those found using the M51 template over the same range. Hence assuming systematic errors of $25\%$ on $L_{\text{TIR}}$ is reasonable.

Notwithstanding these uncertainties, the results imply that the typical galaxies sampled have quiescent IR SEDs at low $z$, but rapidly evolve towards higher IR luminosities at increasing $z$. By $z \sim 2$ they appear to reach ULIRG luminosities, as star-formation activity becomes significantly more prevalent in massive galaxies at these redshifts (e.g. Daddi et al. [2005]). The rise in luminosity with redshift which we observe (from $10^{10}$ to $10^{12}L_\odot$) may be partially attributed to increasing median stellar mass with redshift. This cannot be the full story though, since assuming a linear relationship between stellar mass and $L_{\text{TIR}}$ implies an increase by a factor of 19, whereas $L_{\text{TIR}}$ increases by a factor of $\sim 180$, and $L_{1.4\,\text{GHz}}$ by $\sim 220$ over the redshift range. Indeed we know that $L_{\text{TIR}}$ is linked not to stellar mass itself, but to SFR, which is well-known to rise with increasing redshift (Lilly et al. [1996]; Madau et al. [1996]; Pérez-González et al. [2005]; Damen et al. [2007]; Magnelli et al. [2008]).

### 5.4 Evolution in Specific Star Formation Rates

The SFRs given in Table 4 were calculated using the formulae of Bell [2003], which assume a Salpeter [1955] IMF:

$$
SFR_{\text{TIR}} = \begin{cases} 
1.57 \times 10^{-10} L_{\text{TIR}} \left( 1 + \sqrt{L_{\text{TIR}}/10^{10}} \right), \\
L_{\text{TIR}} > 10^{11} \\
1.17 \times 10^{-10} L_{\text{TIR}} \left( 1 + \sqrt{L_{\text{TIR}}/10^{9}} \right), \\
L_{\text{TIR}} \leq 10^{11} 
\end{cases}
$$

(9)

where $L_{\text{TIR}}$ is in units of $L_\odot$, $L_{\text{TIR}}$ is given in units of $L_\odot$, $L_{1.4\,\text{GHz}}$ in W Hz$^{-1}$ and $L_{\odot} = 6.4 \times 10^{22}$ W Hz$^{-1}$. These conversions were applied to the stacked $L_{\text{TIR}}$ and $L_{1.4\,\text{GHz}}$ to obtain SFRs and to stacked $L_{\text{TIR}}/M_*$ and $L_{1.4\,\text{GHz}}/M_*$ to obtain SSFRs (where $M_*$ are the individual stellar masses). All SFRs were converted to a Kroupa [2001] IMF by subtracting 0.2 dex, following Damen et al. [2009], in order to ensure consistency with the stellar masses used. Radio- and TIR-derived SFRs appear to be roughly in agreement; the TIR values are generally higher although mostly they are within the broad error bars given by the calibration of $L_{\text{TIR}}$. Agreement naturally depends upon the value of $q_{\text{TIR}}$ as a function of redshift being equal to the local value (e.g. the median in Bell’s [2003] sample was 2.64). This will be discussed in the next section. We note that using a constant radio $K$-correction based on the overall median spectral index of $-0.74$ reduces radio SFRs in the last three bins but does not improve the agreement overall.

Fig. 12 (black solid points) shows the median specific SFRs (SSFRs: calculated source-by-source as the radio SFR divided by the stellar mass) as a function of redshift. It is immediately apparent that most SSFRs increase strongly with redshift, indicating a rise in star-formation efficiency within the sample at increasing look-back times, a result seen many times in the literature (e.g. Cowie et al. [1996]; Madau, Pozzetti & Dickinson [1998]; Brinchmann & Ellis [2000]; Bauer et al. [2002]; Feulner et al. [2003]; Pérez-González et al. [2008]; Dunne et al. [2003]; Damen et al. [2009]; Oliver et al. [2010b]). The black open

### Table 4. Stacked galaxy properties derived from measured fluxes: Rest-frame 1.4-GHz luminosities; SFRs derived from $L_{1.4\,\text{GHz}}$ using the Bell (2003) calibration; Total IR luminosities ($L_{\text{TIR}} = L_{1.4\,\text{GHz}} + L_{\text{FIR}}$) estimated directly from the $K$-corrected MIPS fluxes using the M51 template; SFRs derived from $L_{\text{TIR}}$ using the Bell (2003) calibration; Corresponding $q_{\text{TIR}}$ values calculated as described in the text.

| $z$ Range | $(z)$ | $L_{1.4\,\text{GHz}}$, W Hz$^{-1}$ | SFR$_{1.4\,\text{GHz}}$, $M_\odot$/yr$^{-1}$ | $L_{\text{TIR}}$, $L_\odot$ | SFR$_{\text{TIR}}$, $M_\odot$/yr$^{-1}$ | $q_{\text{TIR}}$ |
|----------|------|-----------------|-----------------|-----------------|-----------------|-----------------|
| 0.00–0.40 | 0.21 | $(1.03 \pm 0.13) \times 10^{24}$ | 0.57 $\pm$ 0.08 | $(5.8 \pm 1.4) \times 10^{9}$ | 0.61 $\pm$ 0.17 | 2.76 $\pm$ 0.12 |
| 0.40–0.61 | 0.53 | $(1.21 \pm 0.06) \times 10^{22}$ | 4.23 $\pm$ 0.21 | $(5.7 \pm 1.4) \times 10^{10}$ | 4.8 $\pm$ 1.3 | 2.68 $\pm$ 0.11 |
| 0.61–0.73 | 0.67 | $(2.15 \pm 0.11) \times 10^{22}$ | 7.50 $\pm$ 0.37 | $(9.9 \pm 2.5) \times 10^{10}$ | 8.1 $\pm$ 2.1 | 2.67 $\pm$ 0.11 |
| 0.73–0.96 | 0.87 | $(3.15 \pm 0.19) \times 10^{22}$ | 10.98 $\pm$ 0.69 | $(1.72 \pm 0.43) \times 10^{11}$ | 18.4 $\pm$ 4.7 | 2.75 $\pm$ 0.11 |
| 0.96–1.20 | 1.06 | $(7.48 \pm 0.72) \times 10^{22}$ | 26.1 $\pm$ 2.5 | $(3.27 \pm 0.82) \times 10^{11}$ | 34.1 $\pm$ 8.8 | 2.65 $\pm$ 0.12 |
| 1.20–1.42 | 1.29 | $(1.52 \pm 0.14) \times 10^{23}$ | 53.0 $\pm$ 4.9 | $(6.2 \pm 1.6) \times 10^{11}$ | 64 $\pm$ 16 | 2.62 $\pm$ 0.12 |
| 1.42–2.00 | 1.61 | $(3.63 \pm 0.35) \times 10^{23}$ | 127 $\pm$ 12 | $(1.07 \pm 0.27) \times 10^{12}$ | 169 $\pm$ 28 | 2.48 $\pm$ 0.12 |
This lack of evolution is in agreement with previous studies of 24-µm-selected (e.g. Appleton et al. 2004) and radio-selected samples (e.g. Ibar et al. 2008). Somewhat surprisingly however, some evolution is still apparent in q70 after M51 K-correction, at around 3σ significance.

Some of the anomalies in the K-corrected q24 and q70 graphs could be due to the MIR spectrum and/or the radio K-correction. The z ≈ 0.9 bin for example coincides with the redshifted PAH feature at 12.7µm, and the boost in K-corrected 24-µm flux at this redshift might be a sign of strong PAH emission in the sources. The radio K-correction could also play a part, since in this bin the measured spectral index is relatively flat. This explanation appears likely since a similar bump is apparent in q70 at the same redshift. Furthermore, there are particularly low values of q24 and q70 in the z ≈ 1.6 bin, which coincides with the steepest measured spectral index. Repeating the stacking analysis using a constant spectral index of −0.74 ± 0.07 for radio K-corrections was found to have a small effect on both of these bins, changing each q index by −0.06 dex at z ≈ 0.9 and +0.17 dex at z ≈ 1.6. Similarly the values in the intermediate bins at z ≈ 1.1 and 1.3 were raised by 0.07 and 0.06 respectively (changes in the low-redshift bins were negligible); however this still leaves a decline of 2σ significance in q70 when the M51 template is used. Clearly it is the FIR SED which dominates the evolution of monochromatic q indices, and not the radio spectrum.

One factor that could account for this decline in q70 is a steepening of the continuum slope shortward of 70µm, relative to the M51 template. The SED in the MIR region (10µm ≤ λ ≤ 70µm) is thought to be dominated by emission from very small grains (VSGs, with radii ≤10nm) with fluctuating temperatures resulting from a mixture of thermal and single-photon heating (Desert et al. 1990). A steepening of the the slope shortward of 70µm might be due to an increase in the FIR (≈ 100µm) luminosity (dominated by big grains) relative to the VSG contribution at shorter wavelengths, although this is not clear from the MIPS flux ratios (Fig. 11).

In this context it is interesting to compare with the results of Seymour et al. (2008), who measured 70-µm fluxes for a sample of faint radio sources and reported a decrease in observed q70 with redshift (both for detected sources and stacks), which is not fully accounted for by the K-correction of any single model SED. Seymour et al. concluded that their stacked data show a discrepancy at 0.5 ≤ z ≤ 1.5 between increasing total L_{TIR} values (estimated from radio luminosities) and decreasing q70, implying a change in the ULIRG SED at high redshift. Whatever the cause, it seems plausible that these two samples are similarly affected.

Some of the first results from Herschel provide further tantalising evidence for some change in star-formation activity at high redshifts: Rodighiero et al. (2010) stacked into 100 and 160-µm imaging from the PACS (Poglitsch et al., 2010) Evolutionary Probe (PEP; Berta et al., 2010) at the positions of IRAC (4.5-µm) sources that were optically classified as star-forming and undetected in the 160-µm image, divided into bins of stellar mass and redshift. They found that SSFRs (derived from IR+UV luminosities) followed a power-law trend with mass, with an index of −0.25 ± 0.11 at z < 1, in agreement with SSFRs from radio stacking (Dunne et al. 2008; Pannella et al. 2009), but that the index steepened to −0.50 ± 0.16 at 1 < z < 2, deviating from
Evolution of the FIR–Radio Correlation and IR SEDs

the radio results. A change in the IR SED or $q_{\text{TIR}}$ would be expected to produce such a deviation between SSFRs derived respectively from IR and radio (as is suggested by our data in Table 4).

In Fig. 13 we plot the $q$ indices calculated from $L_{\text{TIR}}$ (listed in Table 5) using Equation 11 (Helou et al. 1983):

$$q_{\text{FIR}} = \log \left( \frac{L_{\text{FIR}}}{3.75 \times 10^{12}} \right) - \log \left( \frac{L_{\text{1.4 GHz}}}{\text{W Hz}^{-1}} \right)$$

Here we substitute $L_{\text{TIR}}$ for $L_{\text{FIR}} = L_{40-120\mu m}$, (as in Bell 2003 and Ivison et al. 2010a, for example), and this difference should be noted when comparing to other work. As an indication, the ratio of $L_{\text{TIR}}/L_{\text{FIR}}$ in the M51 template is 2.1 (which implies $q_{\text{TIR}} - q_{\text{FIR}} = 0.32$), but this ratio is likely to be variable since much of the longer wavelength emission can include contributions from dust heated by older stellar populations (as discussed for example by Bell 2003).

The results for $q_{\text{TIR}}$ are shown in Fig. 13 alongside the median result of Bell (2003) of $q_{\text{TIR}} = 2.64 \pm 0.02$ for a FIR+UV-selected sample of star-forming galaxies at $z \approx 0.0$. We see that our results are generally a little higher than this value at $z < 1$, and it is only due to an apparent evolution in our results that they are more in agreement at high redshift. The slight discrepancy is just within the errors allowed by our TIR normalization, and is likely to result from a systematic difference in the assumptions made about the SEDs and the associated calibration of TIR.

The slight decline in our $q_{\text{TIR}}$ values with redshift is described by an error-weighted least-squares fit given by $q_{\text{TIR}} \propto (1 + z)^{\gamma}$, with $\gamma = -0.11 \pm 0.07$. Note that stacking with the mass limit log$(M) > 10.5$ gives very similar results, fit by the index $\gamma = -0.18 \pm 0.10$. In comparison, the 24µm sample of Ivison et al. (2010a) showed evidence for evolution over redshifts from 0 to 3, with an error-weighted least-squares fit of the same form given by $\gamma = -0.15 \pm 0.03$. Most recently, Ivison et al. (2010a) showed that a sample of LIRGS detected by Spitzer and stacked into Herschel imaging at 100, 160, 250, 350 and 500µm appear to exhibit an evolution in $q_{\text{TIR}}$ over $z = 0 - 2$, with $\gamma = -0.04 \pm 0.03$ (or $-0.26 \pm 0.07$, discounting their 16 galaxies at $z < 0.5$ which were poorly matched in $L_{\text{TIR}}$ to the higher-redshift bins).

A slight decline of a similar scale ($\sim 0.35$ dex) in $q_{\text{TIR}}$ with redshift ($0 < z < 1.4$) was also observed by Sargent et al. (2010a), in the median IR/radio ratios of their sample jointly selected in the IR and radio. However this was at low (2σ) significance and the possibility of intrinsic evolution was rejected by the authors because the median at $z \sim 1.4$ was within the scatter of their low-$z$ value, and moreover because the average at $z > 2.5$ was very similar to the local value. Instead they considered that their sample was more contaminated by AGN at increasing redshifts, and that the hot dust in these AGN caused $q_{24}$ ratios to remain constant, while lower abundances of cold dust caused $q_{70}$ and $q_{160}$ to fall. It is interesting to note that we similarly observe constant $q_{24}$ and falling $q_{70}$ and $q_{160}$, but our observation of constant $q_{160}$ defies a similar explanation.

In a second paper, Sargent et al. (2010b) extended their earlier work using two volume-limited subsets of the joint sample: ULIRGs, and sources populating the bright end of the luminosity function defined by Magnelli et al. (2009). They showed that for both of these IR-bright populations, $q_{\text{TIR}}$ was constant out to redshift 1.4. Following a correction for increased scatter in their data beyond this redshift, they concluded that it remained constant out to redshift 2.\footnote{We note that our data are not affected by the bias in $q$ described by Sargent et al. (2010a,b) due to our selection in the IRAC bands.}

This result disagrees with that of Ivison et al. (2010a), which is flux-limited as opposed to volume-limited, showing the potential importance of selection effects.

The decline in our values of $q_{\text{TIR}}$ could still be caused by the same effect that introduces the decline in $q_{70}$, since the template would underestimate both if the true SEDs were steeper at $\lambda \lesssim 70\mu m$. Alternatively if $q_{\text{TIR}}$ really declines at high redshift then something must be causing galaxies to emit less in the IR relative to the radio at increasing redshifts. This could mean either a reduction in optical depth, causing more UV photons to escape, or an increase in the confinement and/or reprocessing efficiency of CR electrons leading to stronger radio emission. This latter possibility cannot be ignored in the light of our observation that radio spectral indices steepen at redshifts $z \gtrsim 1$, since Lacki et al. (2010) predict that steeper radio spectra are a sign of increasing electron calorimetry in normal galaxies.

In spite of these considerations, we remind the reader that the evolution in $q_{\text{TIR}}$ is at low significance (similar to that of Sargent et al. 2010a), and our data are consistent within 1.5σ with a non-evolution. There is also the potential for some bias introduced by the variation of spectral index with redshift: applying a constant spectral index of $-0.74$ to the radio $K$–corrections reduces the $q_{\text{TIR}}$ evolution to a level that is indistinguishable from being constant: $\gamma = 0.03 \pm 0.07$. The evolution in $q_{70}$ however is not fully removed by this change. Using the measured spectral indices in each redshift bin we fit $q_{70} \propto (1 + z)^{770}$ with $\gamma_{770} = -0.15 \pm 0.05$, while using the constant spectral index we find $\gamma_{770} = -0.10 \pm 0.05$. Nevertheless, it is emphasised that the measured spectral indices should give the most accurate $K$–correction, and Fig. 20 shows that the overall median of $-0.74$ is certainly not inappropriate to represent the flux ratios in all of the bins.

6 CONCLUSIONS

We have studied the FRC as a function of redshift for NIR-selected massive galaxies in the ECDFS, a sample which is unbiased by star-formation activity. We used a stacking analysis to evaluate the ratios of median FIR/radio fluxes of all galaxies in the sample, divided into redshift bins. This technique traces the typical objects in the population of massive galaxies from low redshift back to their formation epoch. A thorough analysis of clustering of the sample was used to correct for the differential effects of confusion in the three FIR bands. $K$–corrections were derived in the radio and FIR using ratios of observed fluxes, ensuring as much as possible a self-consistent analysis. A mass-limited sub-sample was also stacked to confirm the robustness of the results to Malmquist bias.

The results for $q_{24}$, $q_{70}$ and $q_{160}$ show a slight decline in the observed relations, not dissimilar to the results of previous studies, which can be largely accounted for by
the FIR K-correction using an M51 template. After K-correction $q_\text{TIR}$ is the only monochromatic index to show signs of evolution, suggesting that the 70-\(\mu\)m K-correction may be less effective as a result of a steep slope in the SED from \(\sim 25 - 35\mu\)m (corresponding to \(z \sim 1 - 2\)) compared with M51.

Observed MIPS colours at all redshifts are more consistent with the M51 template compared with hotter starburst galaxy templates, indicating that the typical IR SEDs of stellar-mass-selected galaxies at redshifts up to \(\sim 0.8\) (at least) appear to be dominated by cold dust. At higher redshifts it is not possible to constrain the dust temperature with MIPS colours, although it is still clear that M51 is the closest template. In contrast to this, both radio and total IR luminosities rise significantly with increasing redshift, as do derived SFRs. Specific SFRs similarly rise steeply, in agreement with results in the literature (Cowie et al. 1996; Madan et al. 1998; Brinchmann & Ellis 2000; Bauer et al. 2003; Feulner et al. 2005; Pérez-González et al. 2008; Dunne et al. 2004; Damen et al. 2009; Pannella et al. 2009; Oliver et al. 2010).

The stacked radio data reveal tentative evidence for an evolution in radio spectral index across the redshift range, an unexpected result that implies some change in the radio loss processes in our sample towards higher redshifts. The most likely explanation seems to be a shift towards greater inverse-Compton losses of the CR electrons at \(z > 1\), supporting the predictions of Lacki & Thompson (2010).

Overall our results show evidence that the FRC, measured from 24-\(\mu\)m fluxes or 160-\(\mu\)m fluxes closer to the FIR peak, remains roughly constant up to \(z \sim 2\), corresponding to 10 Gyr of cosmic time. This is similar to the conclusions of recent studies including Ibar et al. (2008), Garn et al. (2009), Younger et al. (2009), Ivison et al. (2010a) and Sargent et al. (2010a). The issue is clouded however by measurements at 70-\(\mu\)m, which appear to show a declining index with redshift, and when combined into a total IR luminosity, likewise show a slight decline (at low significance). This most likely implies a steeper spectral slope at wavelengths around 25 – 35\(\mu\)m (compared with the M51 template), leading to insufficient 70-\(\mu\)m K-corrections. But a true evolution in the ratios of 70-\(\mu\)m/radio luminosity and of TIR/radio luminosity is plausible, considering the apparent increase in electron-calorimetry behaviour at \(z > 1\), and considering the fact that rest-frame 24, 70 and 160-\(\mu\)m fluxes can arise from different components of the dust in a galaxy. It is also consistent with the results of Seymour et al. (2003) for $q_\text{TIR}$ and Ivison et al. (2010a/b) using BLAST/Herschel and Spitzer observations to measure $q_\text{TIR}$.

Constraining the FIR SED is one of the greatest problems in understanding the FRC and the dust emission in general from high redshift star-forming galaxies. Upcoming surveys with Herschel, such as the Herschel Multi-tiered Extragalactic Survey (HerMES; Oliver et al., 2010a) and the Herschel Astrophysical Terahertz Large Area Survey (H-ATLAS; Eales et al., 2010) are anticipated to revolutionise our understanding of these topics by providing deep and wide observations spanning the peak of FIR emission across the history of cosmic star formation.

Acknowledgements

The authors are grateful to Jacqueline Monkiewicz of the National Optical Astronomy Observatory, Tucson, and Miwa Block of the University of Arizona, for reducing the IRAC observations to measure $q_\text{TIR}$. The stacked radio data reveal tentative evidence for an evolution in radio spectral index across the redshift range, an unexpected result that implies some change in the radio loss processes in our sample towards higher redshifts. The most likely explanation seems to be a shift towards greater inverse-Compton losses of the CR electrons at \(z > 1\), supporting the predictions of Lacki & Thompson (2010).

Overall our results show evidence that the FRC, measured from 24-\(\mu\)m fluxes or 160-\(\mu\)m fluxes closer to the FIR peak, remains roughly constant up to \(z \sim 2\), corresponding to 10 Gyr of cosmic time. This is similar to the conclusions of recent studies including Ibar et al. (2008), Garn et al. (2009), Younger et al. (2009), Ivison et al. (2010a) and Sargent et al. (2010a). The issue is clouded however by measurements at 70-\(\mu\)m, which appear to show a declining index with redshift, and when combined into a total IR luminosity, likewise show a slight decline (at low significance). This most likely implies a steeper spectral slope at wavelengths around 25 – 35\(\mu\)m (compared with the M51 template), leading to insufficient 70-\(\mu\)m K-corrections. But a true evolution in the ratios of 70-\(\mu\)m/radio luminosity and of TIR/radio luminosity is plausible, considering the apparent increase in electron-calorimetry behaviour at \(z > 1\), and considering the fact that rest-frame 24, 70 and 160-\(\mu\)m fluxes can arise from different components of the dust in a galaxy. It is also consistent with the results of Seymour et al. (2003) for $q_\text{TIR}$ and Ivison et al. (2010a/b) using BLAST/Herschel and Spitzer observations to measure $q_\text{TIR}$.

Constraining the FIR SED is one of the greatest problems in understanding the FRC and the dust emission in general from high redshift star-forming galaxies. Upcoming surveys with Herschel, such as the Herschel Multi-tiered Extragalactic Survey (HerMES; Oliver et al., 2010a) and the Herschel Astrophysical Terahertz Large Area Survey (H-ATLAS; Eales et al., 2010) are anticipated to revolutionise our understanding of these topics by providing deep and wide observations spanning the peak of FIR emission across the history of cosmic star formation.

References

Allamandola L. J., Hudgins D. M., Sandford S. A., 1999, ApJ, 511, L115
Alonso-Herrero A., et al., 2006, ApJ, 640, 167
Amblard A., et al., 2010, A&A, 518, L9+
Appleton P. N., et al., 2004, ApJS, 154, 147
Armus L., et al., 2007, ApJ, 656, 148
Bauer A. E., Drory N., Hill G. J., Feulner G., 2005, ApJ, 621, L89
Beck R., 2000, in Astronomy, physics and chemistry of \(H_3^+\) Vol. 358 of Royal Society of London Philosophical Transactions Series A, Magnetic fields in normal galaxies. pp 777–796
Bell E. F., 2003, ApJ, 586, 794
Berta S., et al., 2010, A&A, 518, L30+
Bondi M., et al., 2007, A&A, 463, 519
Brammer G. B., van Dokkum P. G., Coppi P., 2008, ApJ, 686, 1503
Pedani M., 2003, New Astronomy, 8, 805
Pierce C. M., et al., 2010, MNRAS, 405, 718
Pilbratt G., et al., 2010, A&A, 518, L1+
Poglitsch A., et al., 2010, A&A, 518, L2+
Pope A., et al., 2006, MNRAS, 370, 1185
Pope A., et al., 2008, ApJ, 675, 1171
Pérez-González P. G., et al., 2008, ApJ, 675, 234
Rickard L. J., Harvey P. M., 1984, AJ, 89, 1520
Rieke G. H., Alonso-Herrero A., Weiner B. J., Pérez-González P. G., Blaylock M., Donley J. L., Marcillac D., 2009, ApJ, 692, 556
Rieke G. H., et al., 2004, ApJS, 154, 25
Roche P. F., Aitken D. K., Smith C. H., Ward M. J., 1991, MNRAS, 248, 606
Rodighiero G., et al., 2010, A&A, 518, L25+
Rovilos E., Georgakakis A., Georganopoulos I., Alonso J., Koekemoer A. M., Mobasher B., Goudis C., 2007, A&A, 466, 119
Sajina A., Scott D., Dennefeld M., Dole H., Lacy M., Lagache G., 2006, MNRAS, 369, 939
Salpeter E. E., 1955, ApJ, 121, 161
Sargent M. T., et al., 2010a, ApJS, 186, 341
Sargent M. T., et al., 2010b, ApJ, 714, L190
Serjeant S., et al., 2004, ApJS, 154, 118
Serjeant S., et al., 2008, MNRAS, 386, 1907
Seymour N., Huynh M., Dwelly T., Symeonidis M., Hopkins A., McHardy I. M., Page M. J., Rieke G., 2009, MNRAS, 398, 1573
Seymour N., Symeonidis M., Page M. J., Huynh M., Dwelly T., McHardy I. M., Rieke G., 2010, MNRAS, 402, 2666
Silva L., Granato G. L., Bressan A., Danese L., 1998, ApJ, 509, 103
Siringo G., et al., 2009, A&A, 497, 945
Stern D., et al., 2005, ApJ, 631, 163
Strong A. W., Moskalenko I. V., Reimer O., 2000, ApJ, 537, 763
Symeonidis M., Page M. J., Seymour N., Dwelly T., Coppin K., McHardy I., Rieke G. H., Huynh M., 2009, MNRAS, 397, 1728
Takagi T., et al., 2007, MNRAS, 381, 1154
Taylor E. N., et al., 2009, ApJ, 694, 1171
Thompson T. A., Quataert E., Waxman E., Murray N., Martin C. L., 2006, ApJ, 645, 186
Vallee J. P., 1995, A&A, 296, 819
van der Kruit P. C., 1973, A&A, 29, 263
Vega O., Clemens M. S., Bressan A., Granato G. L., Silva L., Panuzzo P., 2008, A&A, 484, 631
Virani S. N., Treister E., Urry C. M., Gawiser E., 2006, AJ, 131, 2373
Vlahakis C., Eales S., Dunne L., 2007, MNRAS, 379, 1042
Voelk H. J., 1989, A&A, 218, 67
Wall J. V., Jackson C. A., Shaver P. A., Hook I. M., Kellermann K. I., 2005, A&A, 434, 133
Werner M. W., et al., 2004, ApJS, 154, 1
White R. L., Helfand D. J., Becker R. H., Glikman E., de Vries W., 2007, ApJ, 654, 99
Wolf C., et al., 2004, A&A, 421, 913
Young L. M., Bendo G. J., Lucero D. M., 2009, AJ, 137, 3053
Younger J. D., et al., 2009, MNRAS, 394, 1685
Yun M. S., Reddy N. A., Condon J. J., 2001, ApJ, 554, 803
### Table 5. Summary of stacking results.

| Bin (1) | ⟨z⟩ (2) | N_{stack} (3) | Band (4) | S_{obs}, µJy (5) | ±σ_{N}, µJy (6) | ±σ_{S}, µJy (7) | S/N (8) | q_{corr} (9) | ±σ_{S}, µJy (10) | S_k, µJy (11) | ±σ_{S}, µJy (12) | q_k (13) | ±σ_{S} (14) |
|---------|---------|---------------|----------|-----------------|----------------|----------------|--------|-------------|----------------|--------------|----------------|----------|------------|
| ALL 0.73 | 3172   | 24µm  | 145.0 | 1.3  | 4.8  | 111.5 | 0.97 | 0.02 | 254.6 | 7.7 | 1.47 | 0.03 |
| 70µm | 1637.2 | 21.9 | 34.3 | 47.3 | 1.85 | 0.05 | 2117.6 | 87.1 | 2.20 | 0.03 |
| 160µm | 8058.3 | 378.1 | 357.4 | 21.3 | 2.63 | 0.03 | 5031.5 | 336.2 | 2.51 | 0.04 |
| 1.4 GHz | 13.2 | 0.2 | 0.4 | 54.6 | 11.9 | 0.3 |
| 610 MHz | 24.4 | 1.3 | 1.4 | 18.1 | 21.3 | 1.1 |
| ZB0 0.21 | 528   | 24µm  | 186.4 | 3.1  | 9.1  | 59.5 | 1.09 | 0.05 | 223.2 | 14.2 | 1.44 | 0.07 |
| 70µm | 1674.3 | 54.3 | 72.7 | 30.8 | 2.07 | 0.05 | 2068.3 | 159.9 | 2.34 | 0.08 |
| 160µm | 9894.5 | 931.5 | 1023.8 | 10.6 | 2.73 | 0.07 | 5812.8 | 739.5 | 2.71 | 0.10 |
| 1.4 GHz | 12.8 | 0.6 | 1.2 | 21.5 | 11.4 | 1.2 |
| 610 MHz | 22.2 | 3.3 | 2.8 | 6.8 | 20.6 | 2.8 |
| ZB1 0.53 | 528   | 24µm  | 142.3 | 3.1  | 9.1  | 45.5 | 0.94 | 0.05 | 235.1 | 15.1 | 1.31 | 0.04 |
| 70µm | 1318.9 | 53.7 | 122.4 | 24.5 | 1.93 | 0.05 | 2305.5 | 242.4 | 2.29 | 0.04 |
| 160µm | 9235.8 | 946.8 | 1056.6 | 9.8 | 2.67 | 0.07 | 4981.4 | 674.9 | 2.63 | 0.05 |
| 1.4 GHz | 14.0 | 0.6 | 1.0 | 23.5 | 12.9 | 1.2 |
| 610 MHz | 22.8 | 3.3 | 2.5 | 6.9 | 19.5 | 2.1 |
| ZB2 0.67 | 529   | 24µm  | 144.3 | 3.1  | 10.2 | 46.1 | 0.97 | 0.05 | 251.6 | 20.1 | 1.35 | 0.04 |
| 70µm | 1120.2 | 53.6 | 125.7 | 20.9 | 1.88 | 0.06 | 2212.5 | 253.3 | 2.29 | 0.05 |
| 160µm | 7592.0 | 919.8 | 694.2 | 8.2 | 2.60 | 0.07 | 4152.7 | 674.9 | 2.71 | 0.08 |
| 1.4 GHz | 13.2 | 0.6 | 0.7 | 22.2 | 10.9 | 0.6 |
| 610 MHz | 20.5 | 3.3 | 2.8 | 6.2 | 16.4 | 2.2 |
| ZB3 0.87 | 529   | 24µm  | 150.5 | 3.1  | 11.0 | 48.1 | 1.03 | 0.05 | 259.2 | 21.2 | 1.47 | 0.04 |
| 70µm | 996.9 | 53.4 | 107.8 | 18.7 | 1.87 | 0.05 | 2212.5 | 246.8 | 2.37 | 0.05 |
| 160µm | 6505.4 | 932.5 | 851.5 | 7.0 | 2.58 | 0.08 | 3801.9 | 723.8 | 2.63 | 0.08 |
| 1.4 GHz | 12.1 | 0.6 | 0.6 | 20.3 | 9.2 | 0.6 |
| 610 MHz | 18.9 | 3.3 | 4.0 | 5.8 | 14.2 | 2.9 |
| ZB4 1.06 | 529   | 24µm  | 126.1 | 3.1  | 9.0  | 40.3 | 0.93 | 0.06 | 262.6 | 23.9 | 1.32 | 0.05 |
| 70µm | 834.9 | 53.4 | 90.2 | 15.6 | 1.78 | 0.06 | 2108.4 | 218.9 | 2.23 | 0.06 |
| 160µm | 7843.6 | 914.1 | 711.8 | 8.6 | 2.64 | 0.07 | 5288.1 | 625.9 | 2.60 | 0.07 |
| 1.4 GHz | 12.5 | 0.6 | 1.1 | 21.1 | 12.9 | 1.1 |
| 610 MHz | 25.7 | 3.3 | 2.8 | 7.8 | 25.1 | 2.6 |
| ZB5 1.29 | 265   | 24µm  | 110.5 | 4.4  | 9.8  | 25.0 | 0.77 | 0.06 | 286.3 | 27.7 | 1.29 | 0.05 |
| 70µm | 741.0 | 75.5 | 124.8 | 9.8 | 1.62 | 0.08 | 1976.6 | 354.0 | 2.12 | 0.08 |
| 160µm | 8484.4 | 1287.5 | 1102.0 | 6.6 | 2.57 | 0.07 | 6713.7 | 1010.8 | 2.64 | 0.08 |
| 1.4 GHz | 16.0 | 0.8 | 1.1 | 19.1 | 14.7 | 1.1 |
| 610 MHz | 33.0 | 4.6 | 3.7 | 7.1 | 30.7 | 3.4 |
| ZB6 1.61 | 264   | 24µm  | 168.2 | 4.4  | 17.7 | 38.0 | 0.95 | 0.06 | 325.3 | 24.7 | 1.20 | 0.05 |
| 70µm | 539.0 | 75.9 | 101.2 | 7.1 | 1.48 | 0.09 | 1489.7 | 278.4 | 1.81 | 0.10 |
| 160µm | 7404.0 | 1295.4 | 1205.1 | 5.7 | 2.50 | 0.09 | 7458.2 | 1078.3 | 2.56 | 0.10 |
| 1.4 GHz | 16.2 | 0.8 | 1.2 | 19.3 | 20.3 | 1.8 |
| 610 MHz | 40.8 | 4.7 | 3.5 | 8.8 | 48.5 | 4.2 |

(1) Redshift bin; (2) Median redshift; (3) number of objects in stack; (4) Band; (5) Median observed flux before any corrections; (6) Measured 1σ noise (reduced by √N); (7) Statistical 1σ uncertainty on median flux (following Gott et al. 2001); (8) Signal-to-noise ratio (S_{obs}/σ_{N}); (9) q index for IR band after clustering correction; (10) Error on q (using statistical uncertainty from column 7); (11) Median flux following clustering– and K–corrections (using M51 template for MIPS and the measured α(z) for radio); (12) Statistical uncertainty on corrected flux (as in column 7); (13) q index after K–corrections; (14) Statistical uncertainty on K–corrected q.