The Quest for the Missing Dust: I – Restoring Large Scale Emission in *Herschel* Maps of Local Group Galaxies

Christopher J. R. Clark [1], Julia C. Roman-Duval [1], Karl D. Gordon [1], Caroline Bot [2] and Matthew W. L. Smith [3]

[1] *Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, Maryland 21218-2463, United States of America*
[2] *Observatoire Astronomique de Strasbourg, Université de Strasbourg, UMR 7550, 11 rue de l'Université, F-67000 Strasbourg, France*
[3] *School of Physics and Astronomy, Cardiff University, Queen's Buildings, The Parade, Cardiff, Wales, CF24 3AA, United Kingdom*

(Received 2021-06-02; Revised 2021-07-16; Accepted 2021-07-20)

Submitted to (and accepted for publication in) ApJ

ABSTRACT

Because the galaxies of the Local Group have such large angular sizes, much of their diffuse, large-angular-scale emission is filtered out by the *Herschel* data reduction process. In this work, we restore this previously missed dust in *Herschel* observations of the Large Magellanic Cloud, Small Magellanic Cloud, M 31, and M 33. We do this by combining *Herschel* data (including new reductions for the Magellanic Clouds), in Fourier space, with lower-resolution data from all-sky surveys (*Planck*, IRAS, and COBE) that did not miss the extended emission. With these new maps, we find that a significant amount of emission was missing from uncorrected *Herschel* data of these galaxies; over 20% in some bands. Our new photometry also resolves the disagreement between fluxes reported from older HERITAGE Magellanic Cloud *Herschel* reductions, and fluxes reported from other telescopes. More emission is restored in shorter wavelength bands, especially in the galaxies' peripheries, making these regions 20–40% bluer than before. We also find that the *Herschel*-PACS instrument response conflicts with the all-sky data, over the 20–90′ angular scales to which they are both sensitive, by up to 31%. By binning our new data based on hydrogen column density, we are able to detect emission from dust at low ISM densities (at $\Sigma_{\rm H} < 1\,{\rm M}_\odot{\rm pc}^{-2}$ in some cases), and are able to detect emission at much lower densities (a factor of 2.2 lower on average, and more than a factor of 7 lower in several cases) than was possible with uncorrected data.

Keywords: Dwarf galaxies (416), Far infrared astronomy (529), Interstellar dust (836), Local Group (929), Submillimeter astronomy (1647), Astronomy data reduction (1861).

1. INTRODUCTION

The life cycle of interstellar dust in galaxies is highly dynamic. Dust grains are understood to be primarily manufactured through stellar death – by core-collapse supernovæ (Barlow et al. 2010; Matsuura et al. 2011; Gomez et al. 2012) and asymptotic giant branch stars (Höfner & Olofsson 2018). Grains are then processed in the interstellar environment. In denser regions of the InterStellar Medium (ISM), dust can coagulate (Stepnik et al. 2003), and gas-phase metals can accrete onto the grains (Köhler et al. 2015; Zhukovska et al. 2016; Jones et al. 2017), decreasing the gas-to-dust ratio (G/D). Conversely, the reduced shielding in low-density environs increases the rate of photo-destruction of dust by high-energy photons from massive young stars (Boulanger et al. 1998; Beirão et al. 2006) and sputtering by supernovæ shocks (Jones 2004; Bocchio et al. 2014; Slavin et al. 2015), returning metals to the gas phase.

Dust is found to have properties that differ between environments (Cardelli et al. 1989; Gordon et al. 2003). The elemental composition of dust, as inferred from observed depletions of elements from the gas phase, is known to vary markedly between regions (Jenkins 2009; Parvathi et al. 2012; Roman-Duval et al. 2021); and ice spectral features are found in certain areas, indicating

Corresponding author: Christopher J. R. Clark
cclark@stsci.edu

the formation of icy mantles (Boogert et al. 2015). The dependence of dust's emissivity with wavelength – typically expressed in terms of emissivity spectral index, β – is known to vary within galaxies (Smith et al. 2012; Kirkpatrick et al. 2014; Rigby et al. 2018), and traces variations in the physical properties of grains (Demyk et al. 2017a,b; Ysard et al. 2018). In some situations, β seems to exhibit a 'break' at submillimetre (submm) wavelengths, manifesting as submm excess emission. This excess is most commonly found in dwarf galaxies (Galliano et al. 2003; Bot et al. 2010; Rémy-Ruyer et al. 2013; Gordon et al. 2014), and in the periphery of larger late-type galaxies (Paradis et al. 2012; Hunt et al. 2015); the presence of submm excess is therefore associated with environments of lower density and of lower metallicity. Disentangling the interplay between the relative influence of density and metallicity on dust evolution therefore requires as much data as possible, that samples a wide range in both parameters.

The Local Group is the prime laboratory for understanding internal processes of galaxies, including as those governing dust and the ISM. When observing astrophysical processes in the Local Group, we can enjoy exceptionally high-fidelity data, but also benefit from being able to place our observations in the broader context of the entire galaxies within which those astrophysical processes are observed. Whilst we can of course study the Milky Way with resolution that can't be matched in external galaxies, such studies are complicated by the fact that distances to Milky Way features are often uncertain (especially for the ISM), large swathes of the Galaxy are obscured or confused from our observing position, and ultimately we are only observing one type of environment – a high-mass high-metallicity moderately-star-forming spiral galaxy.

1.1. *Challenges of FIR–Submm Observations in the Local Group*

Studies of dust emission in Local Group galaxies often suffer from the specific complexities of Far-InfraRed (FIR) and submm observing. Instruments in this regime generally operate by scanning the sky, with the resulting detector timelines compared and combined in the data reduction process to produce maps. However, the change in flux density recorded by the detector as it scans will not only be due to emission from the target source(s), but can also be caused by drift in instrumental temperature, varying atmospheric foreground emission (for ground-based telescopes), $1/f$ noise, and other effects. This will introduce considerable artefacts on larger angular scales in the resulting maps (see Roussel 2012, Piazzo et al. 2015, and references therein).

A standard way of minimising such artefacts is by observing a wide area of background devoid of emission from the target source. Source emission is then constrained relative to this background by comparing overlapping scans that cover both (Meixner et al. 2013). However, observing large amounts of sky is not possible for all observing programmes. Plus, background scanning will struggle to save observations from noise that manifests at the same angular scales as the emission from the target object, or where the 'empty' background also contains large-scale structure. It is practically impossible to recover emission on scales larger than the size of the map being scanned.

Another technique for removing large scale instrumental artefacts consists of high-pass filtering the detector timelines (Griffin et al. 2010; Roussel 2012). However, this runs the risk of removing genuine large-scale astronomical emission in the observations (Pascale et al. 2011; Valiante et al. 2016). Additionally, some degree of ringing is also likely to be introduced around areas of compact bright emission (Chapin et al. 2013; Kirk et al. 2018); this ringing will manifest on the same scale as the applied filter.

Ultimately, most reduction treatments for suppressing large scale artefacts and noise from FIR–submm scan data will result in astrophysical emission on large angular scales *also* being filtered out of the observations. This is particularly problematic for the Local Group, as it contains the most extended galaxies on the sky – observations of which therefore stand to suffer the most from any removal of FIR–submm emission on larger angular scales. Of particular concern is that the dust giving rise to this filtered-out flux will be the most diffuse dust in these galaxies, found in their outer regions and other low-density environments. This diffuse dust is likely to have distinct properties not found in denser areas, and is important for understanding the evolution of the ISM with environment, especially with regards to G/D (Roman-Duval et al. 2017). The loss of this flux will therefore systematically skew our understanding of the ISM.

There are, however, certain FIR–submm observations that are unaffected by the large scale noise and filtering issues. All-sky surveys with absolute photometric calibration, such as the COsmic Background Explorer (COBE; Boggess et al. 1992) and *Planck* (Planck Collaboration et al. 2011a), can can accurately capture emission on all angular scales, down to their resolving limit. However, these observatories have resolutions at least an order of magnitude worse than that achieved by larger telescopes such as the *Herschel* Space Observatory (Pilbratt et al. 2010); but, of course, larger telescopes are

the ones whose data most suffers from suppression of large-scale astrophysical emission[1].

This situation – high-resolution data that is missing large-scale flux, and low-resolution data that preserves it – is one that radio astronomers handle frequently. When performing multi-dish interferometry, it is common practice to use low-resolution single-dish data to restore the large-scale flux to which the high-resolution interferometric data is not sensitive, by combining the two datasets in Fourier space; this process is often referred to as 'feathering'. Despite being a long-established technique in radio interferometry (Bajaja & van Albada 1979), it has only rarely been applied to single-dish FIR–submm observations (eg, Csengeri et al. 2016, Abreu-Vicente et al. 2017, Smith et al. *subm.*).

1.2. *Paper Overview*

In this paper, we use the 'feathering' Fourier combination approach to produce corrected versions of the *Herschel* maps of the Local Group galaxies M 31, M 33, the Large Magellanic Cloud (LMC), and the Small Magellanic Cloud (SMC).

In Section 3, we detail all of the high- and low-resolution input data we employ, including our new *Herschel* reductions. To feathering together two observations requires us to first infer how the emission detected by the low-resolution telescopes *would appear* if observed in the bandpass used by the high-resolution telescopes; we describe this process in Section 4. In Section 5, we cover the specifics of our Fourier combination process. In Section 6, we describe how we subtract Milky Way foreground emission from the maps we produce; then in Section 7, we present some initial results obtained with our new maps, exploring the properties of the newly-restored dust emission in our target galaxies.

2. SAMPLE GALAXIES

For this work, we produced feathered 100–500 μm maps of the Local Group galaxies M 31, M 33, the LMC, and the SMC. We explored the possibility of extending our reprocessing to other extended nearby galaxies, such as M101 and M51. However, as discussed in Section 5, feathering together two observations requires there to be a range of angular scales over which they are both sensitive. A consequence of this is that the target source must be well-resolved in both sets of observations, otherwise the low-resolution data will only be providing information about emission on angular scales larger than the size of the target source. Even the most extended galaxies outside the Local Group, such as M101, are at best only marginally resolved in the available low-resolution data, meaning that feathering cannot be relied upon to restore extended emission missed by *Herschel*. On the other hand, the fact that these more distant galaxies have much smaller angular sizes means that they should not be susceptible to the removal of extended flux. In Sections 5 and we examine the scales at which flux does begin to be lost, and do indeed find it to be at large enough scales that galaxies outside the Local Group should not be effected.

Table 1. Basic properties of the Local Group galaxies considered in this work. Values taken from the NASA/IPAC Extragalactic Database; except for axial ratios and position angles, taken from the HyperLEDA database.

	M 31	M 33	LMC	SMC
α (J2000)	10.69°	23.46°	80.89°	13.16°
δ (J2000)	+41.27°	+30.66°	-69.76°	-72.80°
Distance (kpc)	790	840	50	62
Hubble Type	SAb	SAcd	SBm	Irr
R_{25} (arcmin)	89	32	323	151
R_{25} (kpc)	20.5	7.5	5	2.5
Pos. Angle (deg)	35	23	170	45
Axial Ratio	2.57	1.70	1.17	1.66

The four Local Group galaxies which are suitable for feathering represent a broad range of galaxy types and properties: M 31 gives us a high-mass disc galaxy passing through the green valley (Mutch et al. 2011); M 33 is a lower-mass gas-rich spiral featuring the highest star-formation efficiency in the Local Group (Gardan et al. 2007); the LMC is a 0.5 $Z_\odot$ galaxy on the spiral / dwarf-irregular border hosting the Local Group's most aggressive site of star formation (Schneider et al. 2018; Ruiz-Lara et al. 2020); and the SMC is a highly-disturbed 0.25 $Z_\odot$ dwarf galaxy displaying the unusual ISM properties common to low-mass low-metallicity systems (Jenkins & Wallerstein 2017; Murray et al. 2019a). There is clear value to producing corrected versions of the high-resolution FIR–submm observations of these key local laboratories. The basic properties of each of these galaxies is provided in Table 1, whilst FIR–submm colour images of each are shown in Figure 1.

3. INPUT DATA

The obvious low-resolution data to use for restoring large-scale emission to *Herschel* observations are the all-sky surveys produced by *Planck*, and the InfraRed

[1] The *Spitzer* (Werner et al. 2004) Multiband Imaging Photometer (MIPS; Rieke et al. 2004) instrument was also remarkably good at preserving large-scale emission, including for the Magellanic Clouds (Meixner et al. 2006), but instead suffered from non-linearity issues at 160 μm (Meixner et al. 2013).

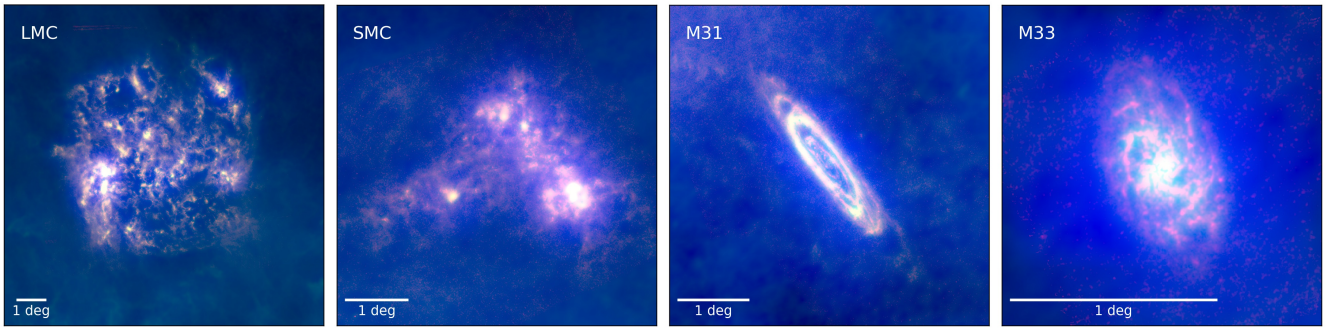

Figure 1. FIR–submm three-colour images of the galaxies we study in this work, also illustrating the range of resolution and sensitivity in the data we employ. Blue is mapped to COBE-DIRBE 100 μm data (FWHM = 0.7°), green is mapped to *Planck* 350 μm data (FWHM = 4.6′), and red is mapped to *Herschel*-SPIRE 500 μm data (FWHM = 35″).

Astronomical Satellite (IRAS; Neugebauer et al. 1984). The wavelength coverage of IRAS and *Planck* fully encompasses our bands of interest for *Herschel*, and their angular resolution is still high enough compared to *Herschel* that there is good overlap in the angular scales to which both instruments are sensitive. An additional complication arises from the fact that IRAS has *also* been found to have flux discrepancies on large angular scales, and lacks independent absolute calibration (see Section 3.4). We therefore take a two-stage Fourier combination approach. First, we feather together the IRAS data with COBE data. Then, in the second step, the *Planck* and rectified IRAS data is feathered with the *Herschel* data to produce our final maps.

In this section we describe all of the input data we employ. For the *Herschel* PACS and SPIRE data, this includes a description of how we created our new reductions for the observations of the Magellanic Clouds. For each instrument, we also describe corresponding Point Spread Function (PSF), as these are of particular importance to the feathering process.

3.1. *Herschel-PACS Data*

The Photodetector Array Camera and Spectrometer (PACS; Poglitsch et al. 2010) on board *Herschel* performed photometric imaging in 3 FIR bands, at 70, 100, and 160 μm (although the design of the filter wheel meant that only two bands could be observed simultaneously; 160 μm plus one of the others); the typical[2] FWHM resolution in these bands is 9″, 10″, and 13″ respectively.

For M 31, these used the PACS data from *Herschel* programme `GT1_okrause_4`, as re-reduced by M. W. L Smith (*Priv. Comm.*). The detector timeline reduction was performed using the *Herschel* Interactive Processing Environment (`HIPE`; Ott 2010) v12[3]. Map-making was then performed using the `Scanamorphos` pipeline (Roussel 2012, 2013) v24.0, a pipeline designed to be particularly effective at preserving larger-scale emission, and minimise negative bowl features.

For M 33, we retrieved the latest reductions from the *Herschel* Science Archive (HSA[4]). Specifically, we used the Standard Product Generation (SPG; the standardised automated reductions provided by the HSA) maps, for which the detector timeline reduction was performed by the HSA using `HIPE` v14, and for which the map-making was carried out using the `JScanam` pipeline (a modified version of `Scanamorphos`, now included with `HIPE`, designed to be robust when run in an automated way; Graciá-Carpio et al. 2017). At 70 and 160 μm, we used the PACS data from *Herschel* programme `OT2_mboquien_4`; at 100 μm, where data from that programme was not available, we instead used the shallower (although wider-area) data from programme `KPOT_ckrame01_1`.

3.1.1. *PACS Re-Reduction for the Magellanic Clouds*

For the Magellanic Clouds, we performed our own re-reductions. The Magellanic Clouds were observed by *Herschel* as part of the *Herschel* Inventory of The Agents of Galaxy Evolution (HERITAGE; Meixner et al. 2013) key programme (`KPOT1_mmeixner1_1`). The HERITAGE PACS data observed both the LMC and SMC at 100 and 160 μm. HERITAGE employed an unusual 'basket-weave' observing strategy, where the full width of each Magellanic Cloud was observed with long continuous scans legs in alternating directions; then 6 months later, a matching set of orthogonal cross-scans was observed. This approach was taken for several reasons: it

[2] The PACS beam size was dependent upon observing mode; the values given here are roughly average values, to be representative.

[3] `HIPE` v12 uses the same photometric calibration products as `HIPE` v15.0.1, the most recent version as of the time of writing.

[4] http://www.cosmos.esa.int/web/herschel/science-archive

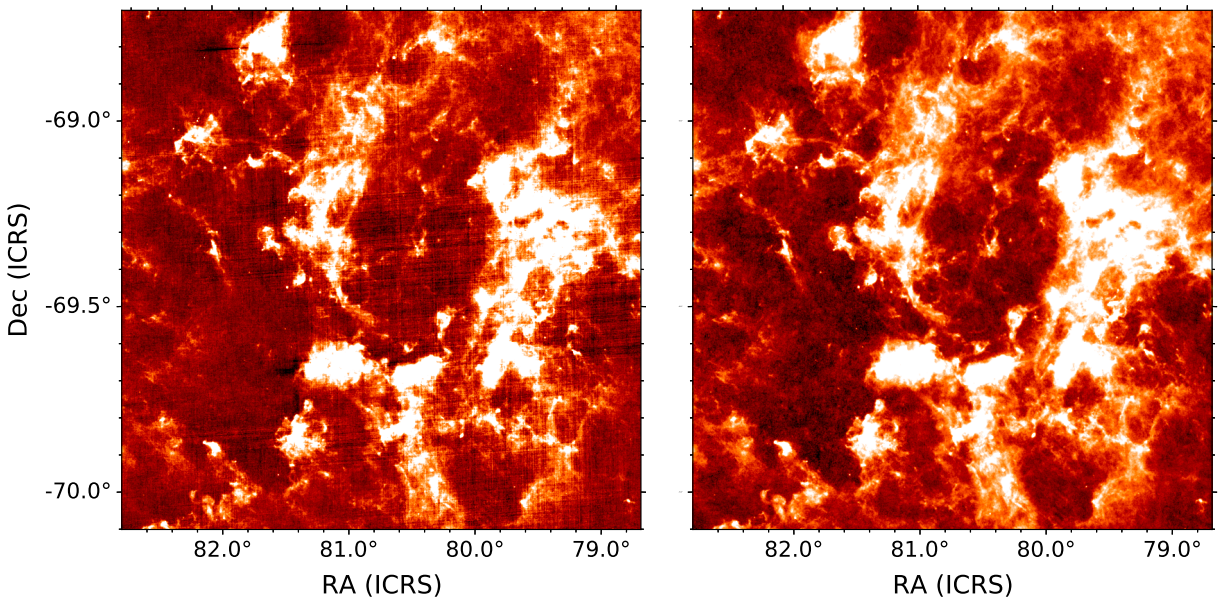

Figure 2. Comparison of the old HERITAGE reduction of the PACS data (*left*), to our new reduction (*right*), for a portion of the LMC at 100 μm. Note the significant artefacts in the old reductions, such as the cross-hatching, and the linear negative features around bright regions. Also note the improved sensitivity in the new reduction revealing faint features not visible in the old map. (NB: The new map shown here does *not* incorporate the restored large-scale emission, as covered in Section 5.)

ensured that all scans sampled the 'empty' sky beyond the target galaxies; it maximised the number of cross-scans that each scan overlapped, to allow instrumental variation to be disentangled more effectively; and it increased the fraction of the observing time that was spent integrating along scan legs, as opposed to time spent in turnaround or slewing.

However, this observing strategy had the unintended consequence of leaving the PACS data severely affected by $1/f$ noise, arising from the extremely long scan legs (8° for the LMC, and 6° the SMC), which are the longest of any *Herschel* observations. In the PACS detector timelines, instrumental baseline drift entirely dominates over astrophysical signal. Whilst the redundancy from the multiple orthogonal cross-scans means this drift can be well accounted for, doing so requires holding all of the scans in memory, which was not plausible at the time of the original HERITAGE data release. Instead, they adopted a number of other strategies to remove the baseline drift from the scans; see Section 3 of Meixner et al. (2013) for details. The baseline drift removal strategies employed by Meixner et al. (2013) led to significant artefacts around bright sources (illustrated in the left panel of Figure 2), and relied upon tying the surface brightness at the end of each scan leg to a linear interpolation from IRAS and COBE data, leading to conspicuous discontinuities.

We therefore needed to perform our own new reductions of the HERITAGE observations of the Magellanic Clouds. We obtained the detector timelines for each individual observation from the HSA, where they had already been reduced using SPG pipeline in `HIPE` v14, with the most recent calibration products. To create combined maps from these processed timelines, we used the `UNIMAP` pipeline (Piazzo et al. 2015), v7.1.0. `UNIMAP` is a generalised least-squares map-maker, specifically designed to handle data where the detector timelines suffer from $1/f$ noise. Whilst most other PACS map-makers (such as `JScanam`) require specific matched pairs of scan and cross-scan observations in order to function, `UNIMAP` can operate with arbitrary sets of overlapping scans – a necessary feature given the HERITAGE observing strategy. Recent releases of `HIPE` provide the option to call `UNIMAP` for the map-making stage of PACS data reduction, and this was how we used `UNIMAP` to produce maps for the LMC and SMC, with all processing options set to the default settings. This makes our LMC and SMC reductions substantially similar to those produced by the *Herschel* Science Centre as part of its PACS 'Highly Processed Data Products' data release (Calzoletti 2017), which also used `UNIMAP`. However, we use a slightly newer version of `HIPE` than Calzoletti (2017); also, their reductions for the SMC break the data into separate maps for the bar and bridge of the SMC, whereas we produce one contiguous map. A comparison of the old

HERITAGE reduction and our new reduction for a portion of the SMC is shown in Figure 2. As an example, our PACS 100 μm reduction for the LMC is shown in the right panel of Figure 3.

3.1.2. *Herschel-PACS PSF*

For the PACS PSF, we used the azimuthally-averaged PSFs produced by Aniano et al. (2011)[5]. For the LMC and SMC, the fact that the observations conducted during multiple epochs, spaced at six month intervals means that the effective PSF will be averaged over multiple orientations, meaning that the standard PACS PSF is unsuitable, as it has a great deal of azimuthal variation (Bocchio et al. 2016) that will not be reflected in our maps. For consistency, we therefore also use the Aniano et al. (2011) PSFs when working with the data for M 31 and M 33.

3.2. *Herschel-SPIRE Data*

The Spectral and Photometric Imaging REceiver (SPIRE; Griffin et al. 2010) on board *Herschel* provided simultaneous photometric imaging in 3 submm bands, centred at 250, 350, and 500 μm; these bands achieved resolutions of 18″, 25″, and 36″ FWHM, respectively.

For M 31, we used the SPIRE data from *Herschel* programme `GT1_jfritz_1`, as re-reduced by M. W. L Smith (*Priv. Comm.*). The detector timeline reduction was performed using the Bright Galaxy Adaptive Element (BRIGADE; Smith et al. 2012; Smith 2013) pipeline, run with `HIPE` v12[3]. In the case of M 31, this data has fewer thermal drift artefacts than the SPG reductions from the HSA.

For M 33, we took the latest SPG reductions from the HSA, which use the up-to-date `HIPE` v14 photometric calibrations. These maps use the data from *Herschel* programme `KPOT_ckrame01_1`.

In all instances, the data were reduced with the relative gains of the SPIRE bolometers optimised for detecting extended emission, using the values for the beam area values provided in `HIPE` v15; specifically, 469.7, 831.7, and 1793.5 arcsec^2 at 250, 350, and 500 μm respectively. Additionally, all maps were produced using the SPIRE de-striper, to mitigate large scale artefacts arising from instrumental drift.

3.2.1. *SPIRE Re-Reduction for the Magellanic Clouds*

For the LMC and SMC, we again used the data from HERITAGE (*Herschel* programme `KPOT1_mmeixner1_1`). We had to perform our own reductions, as the HSA cannot automatically generate reductions for observations with unusual scan strategies, like HERITAGE. The original HERITAGE reductions relied upon setting the surface brightness at the end of each scan leg to 0 – making the final maps insensitive to emission at their edges, such as might arise from cirrus, extended Magellanic Cloud dust, etc.

We obtained the detector timelines for all of the HERITAGE observations from the HSA, where they had already been reduced through the SPG pipeline using the `HIPE` v14 calibrations. We merged all of the scan legs for all of the observations for each galaxy, and then carried out map-making using the SPIRE destriper.

Running the destriper on the SPIRE observations of the LMC and SMC was a complex proposition. Firstly, the destriper can struggle to achieve good results when a map contains extremely bright regions – which the Magellanic Clouds certainly do (30 Doradus, LHA 120-N 55A, etc). The destriper provides the option to specify regions that may prove problematic, so they can be masked from the destriping algorithm; we therefore manually identified and masked 8 such regions in the LMC, and 6 in the SMC (listed in Appendix B). Secondly, the destriper is highly memory-intensive. The scan timelines are compared simultaneously, through an iterative-map-making process with which the destriper isolates instrumental drift; as such, a change in the value of any one pixel can potentially change the value of every other pixel. The memory requirements for carrying out this process for the HERITAGE observations of the Magellanic Clouds were exceptional. In particular, for the LMC, destriping all three SPIRE bands required 1 petabyte-hour of memory, with instantaneous memory usage peaking at almost 10 terabytes for the 250 μm data. This is in all likelihood the most computationally-intensive *Herschel* data reduction that will ever be carried out (whilst larger SPIRE maps do exist, they can be produced by reducing individual tiles of scans and cross-scans, and then mosaicking them together; the HERITAGE observing strategy precludes this).

3.2.2. *Herschel-SPIRE PSF*

In keeping with PACS, and for the same reasons given in Section 3.1.2, also we used the azimuthally-averaged PSFs produced by Aniano et al. (2011)[5] for SPIRE.

3.3. *Planck Data*

The *Planck* satellite (Planck Collaboration et al. 2011a) surveyed the entire sky in 9 submm–microwave, bands from 350 μm to 10 mm. Its mission to map the cosmic microwave background required its instruments to have highly accurate absolute calibration, with sensitivity to emission on angular scales spanning the whole

[5] http://www.astro.princeton.edu/~ganiano/Kernels.html

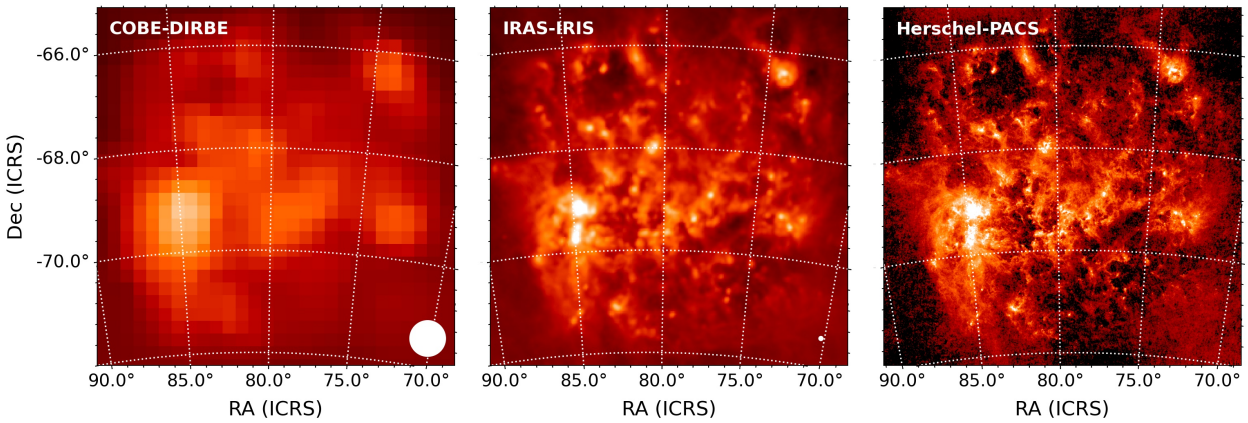

Figure 3. The 100 μm observations of the LMC, from the three instruments that observed at that wavelength; DIRBE (*left*), IRIS (*centre*), and PACS (*right*). These IRIS and PACS images have not had large-scale emission restored by feathering. The beam FWHM for each instrument is shown by the white circle in each panel; the PACS beam is not visible at this scale. All three images have been matched to approximately the same colour scale; the dark 'holes' in the PACS image, and the inconsistent surface brightness levels around the galaxy's outskirts, illustrate the difficulty PACS suffered in retrieving larger-scale diffuse emission.

range from its primary beam up to the all-sky dipole (Planck Collaboration et al. 2016a,b). This makes it perfectly suited to providing the absolutely-calibrated low-resolution data we need at wavelengths $\geq 350\,\mu$m.

We retrieved the 2018 *Planck* data release all-sky intensity maps from the *Planck* Legacy Archive[6]. Specifically we obtained the 350 and 550 μm maps, both from *Planck*'s high frequency instrument (Planck Collaboration et al. 2018). These maps have had the zodiacal light contribution removed, are already provided in our desired surface brightness units of MJy sr^{-1}, and are calibrated assuming a constant νS_ν reference spectrum. The telescope achieves FWHM resolution of 4.6′ and 4.8′ at 350 and 550 μm respectively.

For each of our galaxies, we produced a square cutout map centred on the target, with a width equal to 20 times its R_{25} (the isophotal semi-major axis at which the optical surface brightness falls beneath 25 mag arcsec^2; see Table 1). We made these cutouts by reprojecting from the HEALPix projection (Górski et al. 2005) of the all-sky maps to a standard east–north gnomic tan projection, using the `Python` package `reproject`.

3.3.1. *Planck PSF*

Different parts of the sky were scanned a varying number of times by *Planck*, and each with different scan orientations, so the *Planck* PSF varies as a function of sky position. The *Planck* Legacy Archive is able to provide the average *Planck* PSF for a given area of sky. We obtained the average PSF for each of our sample galaxies, with the average taken over the $10D_{25}$ square cutout area. In reality, the effective PSF will be slightly different over different parts of our sample galaxies, but these differences will be extremely small – the variation in the solid angle of the *Planck* beam across the sky has a standard deviation of <0.8% in the bands we use (Planck Collaboration et al. 2016c). Performing feathering with a positionally-varying PSF would add exceptional complication to our work for very little gain. The average PSFs we obtained for each galaxy well match the average all-sky PSFs, but differ in terms of the structure of the wings.

3.4. *IRAS-IRIS Data*

The InfraRed Astronomical Satellite (IRAS; Neugebauer et al. 1984) mapped 96% of the sky in 4 IR bands, from 12–100 μm. The IRAS observations at 100 μm provide the necessary shorter-wavelength counterpart to the *Planck* data, with a well-matched angular resolution of 5′. Together, IRAS and *Planck* cover the entire wavelength range of the *Herschel* data we are reprocessing.

There exist two sets of all-sky IRAS maps; the Iras Sky Survey Atlas (ISSA; Wheelock et al. 1994), and the Improved Reprocessing of the Iras Survey (IRIS; Miville-Deschênes et al. 2005). For our purposes, both sets of maps have pros and cons. The newer IRIS maps have been de-striped, de-glitched, zodiacal light subtracted, and had their absolute flux calibration pegged to that of COBE. However, Bot et al. (*in prep.*) have found that the IRIS maps nonetheless show considerable (up to ∼20%) deviation from the COBE data over large angular scales (> a few degrees) in certain regions,

[6] https://pla.esac.esa.int/maps

such as the Magellanic Clouds – precisely where we need to be able to depend on large-scale fidelity in the IRAS data. Whilst the ISSA maps do not suffer from this specific systematic, they do however lack the improved absolute calibration as the IRIS data, and have a highly non-linear response function that varies with both the angular scale and surface brightness of the emission being observed. Indeed, the ISSA explanatory supplement (Wheelock et al. 1994) suggests that users should assume photometric uncertainties of up to 30% and 60% at 60 μm and 100 μm respectively.

For this reason, we opt to use IRIS for our source of IRAS maps. But this requires us to first find a way to correct the extended surface brightness deviations in the IRIS data around the Magellanic Clouds. Fortunately, this problem can be corrected in exactly the same manner as the poor preservation of extended emission in the *Herschel* data.

First, we correct the IRIS data by feathering it with data from an even-lower-resolution telescope, one that has accurate absolute calibration and preservation of large-scale-emission. Then, we use this corrected IRIS data (in concert with the *Planck* data) to provide the low-resolution counterpart to correct the *Herschel* and *Spitzer* data in a second round of feathering, to produce our final high-resolution maps. The obvious choice for the even-lower-resolution data with which to correct the IRIS maps is COBE, which we describe below, in Section 3.5.

We retrieved the IRAS-IRIS data we used from the InfraRed Science Archive (IRSA[7]). For this work, we only required the IRAS 100 μm maps. For each our galaxies, we produced a square cutout map centred on the target, with a width equal to 20 times its R_{25}. We made these cutouts by mosaicking together all of the individual IRIS maps that covered the region of interest – by first projecting these individual maps to a shared pixel grid (using the `Python` package `reproject`), then taking the mean value in any pixel where multiple maps overlapped. This is the same process used in the mosaicking scripts released by the IRIS team[8]. As an example, the IRIS 100 μm map of LMC is shown in the centre panel of Figure 3.

3.4.1. *IRAS-IRIS PSF*

Unfortunately, the IRIS PSF is not especially well characterised. Because IRIS maps are constructed from scanning data, the instrumental PSF is worse in the cross-scan direction (Wheelock et al. 1994). To mitigate this, the IRIS reductions were made such that all data was smoothed to this worst, cross-scan direction, rendering the effective PSF circular. Miville-Deschênes et al. (2005) verified that the power spectrum of the resulting maps was compatible with the PSF being Gaussian, but performed no further characterisation.

We attempted to determine the true 100 μm IRIS PSF empirically, by stacking on the positions of bright point sources from the IRIS Point Source Catalog (Beichman et al. 1988). However, the noise in the stack was high enough that we were unable to trace the PSF past 20′, beyond which the noise dominated. We tried limiting the stack to only sources with higher Signal-to-Noise Ratios (SNR), or to sources out of the Galactic plane, but in these cases the reduction in the number of sources being stacked meant that the noise in the stack remained similarly poor. However, we were able to verify that the radial profile of the IRIS PSF is well-modelled by a Gaussian with a FWHM of 4.4′ out to the maximum radius we could trace. This value compares well to the 4.3′ FWHM stated in Miville-Deschênes et al. (2005); we therefore opt to follow that official value, and use a Gaussian with FWHM of 4.3′ as our 100 μm IRIS PSF.

3.5. *COBE-DIRBE Data*

The Diffuse InfraRed Background Experiment (DIRBE; Silverberg et al. 1993) instrument on COBE was a single-pixel absolute photometer that observed the entire sky in 10 bands from 1.25–240 μm, at an instrumentally-limited resolution of 0.7°. DIRBE required accurate absolute calibration for COBE's mission to map both the cosmic infrared background and cosmic microwave background. As a result, its accuracy has been well characterised (Fixsen et al. 1997), and DIRBE serves as the primary photometric reference standard for extended emission across most of its wavelength range (Miville-Deschênes et al. 2005; Gordon et al. 2006).

We retrieved the all-sky COBE-DIRBE maps at 100, 140, and 240 μm from the Legacy Archive for Microwave Background Data Analysis (LAMBDA[9]). Specifically, we used the Zodical-Light-Subtracted Mission Average Maps. These all-sky are presented in the pre-HEALPix quadrilateralised spherical cube projection (Torres et al. 1989). We used this all-sky data to create square cutouts, of width 30°, centred on each of our sample galaxies. To reproject the maps from the quadrilateralised spherical cube projection to the standard gnomonic TAN projection for our cutouts, we used the algorithm supplied by the COBE team[9], albeit adapted

[7] https://irsa.ipac.caltech.edu/data/IRIS/images/

[8] https://www.cita.utoronto.ca/~mamd/IRIS/IrisDownload.html

[9] https://lambda.gsfc.nasa.gov/product/cobe/dirbe_products.cfm

from the original `FORTRAN` code into `Python` (Oliver Lomax, *priv. comm.*). As an example, the DIRBE 100 μm map of LMC is shown in the left panel of Figure 3.

We note that there is an alternate set of DIRBE maps hosted at IRSA, stored in a HEALPix projection. However, the surface brightness in these maps, compared to those hosted at LAMBDA, do not agree. When regridded to the same projection to allow direct comparison, they can disagree by up to $\pm$15%, particularly surrounding bright sources. It seems that this is due to a slight difference in resolution between the two maps. In communication with personnel at IRSA, it appears that these maps were produced and delivered by the *Planck* team as part of their data releases (potentially just for quick-look purposes). As the maps hosted at LAMBDA are the original and official maps delivered by the COBE team, we opt to use those. People seeking DIRBE data online should be aware that the different archives provide differing versions of the data.

3.5.1. *COBE-DIRBE PSF*

The resolution of DIRBE is instrument-limited, not diffraction-limited, and is dictated by how the single pixel of DIRBE scanned a given sky position numerous times, along numerous direction. However, the DIRBE data products hosted on LAMBDA only include the instrument's *scanning* PSF – the spread function response from a single pass by the detector, in a particular scanning direction. We, however, need the effective PSF resulting from all the scans in all the directions. We therefore produced maps of the effective DIRBE PSFs for ourselves, by making an azimuthally-averaged version of each band's one-dimensional scanning PSF. The azimuthal average should well replicate the effect of the >100 scans (Hauser et al. 1998) DIRBE conducted at every point on the sky. Our resulting PSFs have the smoothed top-hat shape expected of the DIRBE beam (Kashlinsky et al. 1996; Barreiro et al. 2004), with the correct width of 0.7°.

4. SED-DRIVEN INTERPOLATION

In order to correctly feather together two observations, those observations must be observing the sky with the same wavelength sensitivity. For instance, we cannot simply feather SPIRE 500 μm and *Planck* 550 μm observations together; although their spectral responses have significant overlap (having bandwidths of $\sim \lambda/3$), they will nonetheless sample different parts of any source SED, meaning observed brightness will vary as a function of source spectrum. This is even true for bands with the same effective wavelengths, such as the DIRBE, IRAS, and PACS 100 μm bands – each has a different response function, so the relative brightnesses reported by each will differ by source spectrum.

Therefore, we need to model the underlying source spectrum. Specifically, we take the approach of modelling the SED in each pixel of a given set of low-resolution maps, and then use this to predict what emission *would* be seen by the high-resolution filters. These predicted maps have the resolution (and preserved large-scale emission) of the low-resolution maps, translated into the high-resolution bands. This method also allows us to create these predicted maps for high-resolution bands that have no direct equivalent amongst the low-resolution data; for instance, by modelling the 100, 350, and 550 μm IRAS and *Planck*, we can predict the low-resolution emission that would be observed in the SPIRE 250 μm filter. In effect, we are performing a 'physically motivated interpolation' to infer the emission in the high-resolution bands.

The FIR–submm emission in our observations will have an SED that is dominated by dust emission. This dust SED can be approximated by a Modified BlackBody (MBB), whereby the dust mass opacity varies with wavelength according to the emissivity law:

$$\kappa(\lambda) = \frac{\kappa(\lambda_{ref})}{\lambda_{ref}^{-\beta}} \lambda^{-\beta} \tag{1}$$

where κ is the dust mass opacity (ie, the grain absorption cross section per unit mass) at wavelength λ, B is the Planck function evaluated at wavelength λ and dust temperature T_d, and Σ_d is the dust column density.

The surface brightness, $S(\lambda)$, of dust emission at a given wavelength, λ, , can be expressed by:

$$S(\lambda) = \kappa(\lambda)\ B(\lambda, T_d)\ \Sigma_d \tag{2}$$

where $\kappa(\lambda_{ref})$ is the dust mass opacity at some reference wavelength (λ_{ref}), and β is the dust emissivity spectral index. Gordon et al. (2014) showed that the FIR–submm dust SED of the Magellanic clouds is particularly well fit by a Broken-Emissivity Modified BlackBody (BEMBB). In this model, the value of β changes at some break wavelength λ_{break}, allowing us to fit SEDs that exhibit 'submm excess' emission above what would be predicted from a MBB alone (Galliano et al. 2003; Bot et al. 2010), by changing to a shallower β at longer wavelengths. The BEMBB emissivity law therefore takes the form:

$$\kappa(\lambda) = \frac{\kappa(\lambda_{ref})}{\lambda_{ref}^{-\beta}} E(\lambda) \tag{3}$$

where

$$E(\lambda) = \begin{cases} \lambda^{-\beta_1} & \text{for } \lambda < \lambda_{break} \\ \lambda^{\beta_2-\beta 1}\,\lambda^{-\beta_1} & \text{for } \lambda \geq \lambda_{break} \end{cases} \quad (4)$$

for which β_1 is the value of β at wavelengths $< \lambda_{break}$, and β_2 is the value of β at wavelengths $\geq \lambda_{break}$.

Adopting a BEMBB model allows our SED fitting to be flexible. In the portions of the Magellanic Clouds and M 33 known to exhibit submm excess (Planck Collaboration et al. 2011b; Gordon et al. 2014; Relaño et al. 2018), it will provide a good fit where a MBB model would have had larger (and systematic) residuals, Meanwhile, in areas where the SED would already have been well fit by a MBB model (or when fitting data at wavelengths too short to be sensitive to break; see Section 4.2), a BEMBB model can simply have $\beta_1 = \beta_2$. Because all of our data is at wavelengths $> 100\,\mu$m, we do not need to be concerned about the power-law emission that results from stochastic heating small grains at $\lesssim 70\,\mu$m (Boulanger & Perault 1988; Desert et al. 1990). The BEMBB model should therefore provide reliable interpolation to predict the surface brightness expected in the filters of interest.

Note that although we are modelling the SED with physical parameters (such as the dust temperature, mass surface density, emissivity spectral index, etc), we are not, at this point, at all concerned with the actual physical implications of any of the models – we are only performing the SED fitting to achieve our physically motivated interpolation. We are therefore unconcerned by the fact that each pixel will contain a combination of emission from the target galaxy and the foreground Milky Way cirrus, or by the fact that some pixels might be more properly modelled by some other model (two temperature, etc) – as long as our SED modelling provides good fits for reliable interpolation, it has achieved its purpose. We only consider the various model parameters insomuch as they provide a diagnostic for the success of the fitting.

For $\kappa(\lambda_{ref})$, we use the Roman-Duval et al. (2017) value of $1.24\,\mathrm{m^2\,kg^{-1}}$ at $\lambda_{ref} = 160\,\mu$m. For our purposes, the specific value of $\kappa(\lambda_{ref})$ is essentially arbitrary; however, using the Roman-Duval et al. (2017) value allows ease of comparison to their IRAS and *Planck* `DustBFF` analysis of the Magellanic Clouds, which provides a useful additional check.

4.1. *SED fitting with DustBFF*

We perform our SED fitting using the Dust Brute Force Fitter (`DustBFF`; Gordon et al. 2014). `DustBFF` is a grid-based SED fitting code. This is the most efficient solution for fitting the SED of large numbers of pixels in a set of maps.

Table 2. SED model grid parameter ranges and step sizes, used for pixel-by-pixel SED fitting using `DustBFF`, with a BEMBB model. For the logarithmically-spaced parameters Σ_d and T_c, we also give the percentage difference between grid steps.

Parameter	Minimum	Maximum	Step
Σ_d ($\mathrm{M_\odot\,pc^{-2}}$)	$10^{-3.5}$	10^{1}	0.025 dex (9.6%)
T_d (K)	10	50	0.04 dex (5.9%)
β_1	0	3	0.1
[a] λ_{break} (μm)	150	350	25
[a] e_{500}	-0.5	2.0	0.075

[a] When fitting only DIRBE data, we set $e_{500} = 0$ and fix λ_{break}, effectively adopting a MBB model.

The BEMBB model implementation in `DustBFF` does not explicitly parameterise β_2, instead using the $500\,\mu$m excess, e_{500}, which describes the relative excess in flux at $500\,\mu$m, above what would be expected from a MBB model with no break (where negative values indicate a $500\,\mu$m deficit), given by:

$$e_{500} = \left(\frac{\lambda_{break}}{500\,\mu\mathrm{m}}\right)^{\beta_2-\beta_1} - 1 \quad (5)$$

With our BEMBB model, the free parameters are therefore Σ_d, T_d, β_1, λ_{break}, and e_{500}. The parameter grid we use is described in Table 2. When fitting DIRBE data alone (see Section 4.2), we set $e_{500} = 0$ and fix λ_{break}, effectively adopting a MBB model, as DIRBE data does not provide sufficient long wavelength coverage to constrain any SED break.

The grid spacing for Σ_d and T_c is base-10 logarithmic; this is to capture the wide dynamic range of possible dust mass surface densities, and the fact that the FIR–submm bands we are modelling are more sensitive to smaller changes in cooler dust temperatures than hotter ones. Our parameter grid effectively imposes a flat prior across the range of values modelled for each parameter (or flat in logarithmic space for Σ_d and T_c), with zero likelihoods outside these ranges. The full grid contains over 24 million models. `DustBFF` computes the probability of fitting the data for each model in the grid. The mathematical formalism under which `DustBFF` operates is provided in full in Section 4 of Gordon et al. (2014).

The `DustBFF` formalism incorporates a full covariance matrix, $\mathcal{C}$, given by:

$$\mathcal{C} = \mathcal{C}_{calib} + \mathcal{C}_{instr} \quad (6)$$

where $\mathcal{C}_{calib}$ is a matrix incorporating the covariances between bands due to calibration uncertainty, and $\mathcal{C}_{instr}$ is a matrix used to capture how the instrumental noise will effect the uncertainty for each model.

It is vital to take account of the correlated uncertainties, as they can lead to systematic errors in model results (Veneziani et al. 2013). For instance, if correlated error in the calibration of the long-wavelength points of a dust SED leads to those fluxes being recorded as greater than they are, this would be erroneously interpreted as favouring a flatter emissivity slope β; accounting for the correlated uncertainties via $\mathcal{C}_{calib}$ protects against this. We provide $\mathcal{U}_{uncorr}$ and $\mathcal{U}_{corr}$ for each instrument in their respective subsections below. This effect can be accounted for by explicitly parameterising the correlated uncertainties (Galliano et al. 2011; Clark et al. 2019), or by treated them as hyperparameters in a hierarchical Bayesian model (Kelly et al. 2012), or by including them as off-diagonal terms in the covariance matrix (Smith et al. 2012; Gordon et al. 2014), as we do here.

We also incorporate the effect of the instrumental noise in the observations, via the matrix $\mathcal{C}_{instr}$ (which was not included in the original Gordon et al. 2014 presentation of `DustBFF`). This allows our fitting to take account of the fact that models which predict a higher surface brightness will be less effected by uncertainty, simply because the noise will be lower relative to the emission. All of the observations we will be fitting using `DustBFF` have effectively flat noise over the map areas being modelled. We therefore calculate the diagonal values of $\mathcal{C}_{instr}$ by taking the median of each observation's uncertainty map; the off-diagonal elements are zero. We provide $\mathcal{C}_{instr}$ for each instrument in their respective subsections below, with a different $\mathcal{C}_{instr}$ calculated for each galaxy's observations.

In previous uses of `DustBFF`, $\mathcal{C}$ also incorporated the correlated uncertainty between bands arising from the uncertainty on the background subtraction (Gordon et al. 2014; Roman-Duval et al. 2017; Chastenet et al. 2017). However, as we are concerned with the entirety of the flux recorded in each pixel, we are not performing background subtraction, so background subtraction covariance is not present.

The full posterior probability distributions returned by `DustBFF` are extremely large, containing the probabilities of all 24 million models, calculated for every single pixel (of which there are are over one million, in the case of our *Planck* and IRIS maps of the LMC). This is an impractical amount of data to handle, especially given that the vast majority of the models will have essentially nil probability for any given pixel. Therefore, for each pixel, we draw 500 random models from the posterior (with replacement), where the probably of any given model being drawn is proportional to its posterior probability. These 500 realisations of the posterior are what we use to constrain parameter medians, uncertainties, covariances, etc, in each pixel, and a propagated throughout our analysis. For our fiducial model in each pixel, we use the Maximum A-Posteriori likelihood (MAP) model; our fiducial output maps are comprised of the MAP model values for each pixel.

4.2. *SED Fitting with DIRBE to Infer Large-Scale Emission in IRIS 100 μm Band*

As described in Section 3, we require DIRBE data to correct large-scale artefacts in the IRIS maps, and feathering together these two datasets first requires fitting the DIRBE SED, pixel-by-pixel, to infer how the emission observed by DIRBE would appear in the IRIS 100 μm band. To do this, we model the pixel-by-pixel SEDs of the DIRBE 100, 140, and 240 μm (ie, DIRBE bands 8, 9, and 10) observations of our target galaxies, using `DustBFF`.

All of the data are already at the same 0.7° instrumentally-limited DIRBE resolution, so no convolution was required prior to fitting.

We fit the SED of each pixel using `DustBFF`. We kept $e_{500} = 0$ and fixed λ_{break} at an arbitrary value, because DIRBE does not provide sufficiently long-wavelength coverage to constrain these parameters. Moreover, because we are only performing this fitting to predict the IRIS flux at 100 μm, constraining the model at the longest wavelengths is less vital (unlike for the IRIS–*Planck* SED fitting in Section 4.3). This resulted in a grid of 79 200 models.

The uncertainty on these bands' calibration is dominated by the 10% uncertainty on the absolute gain, which is shared between the bands; the remaining 1–3% uncertainty for each band represents the photometric repeatability error arising from fluctuations in the relative gains, and is uncorrelated between bands. Therefore, each element of the DIRBE $\mathcal{U}_{corr}$ matrix has a value of 0.1; whilst the diagonal elements of the $\mathcal{U}_{uncorr}$ matrix have values of 0.03, 0.01, and 0.02 (with matrix rows/columns representing the 100, 140, and 240 μm values respectively).

We calculated the target galaxies' covariance matrices for instrumental noise, $\mathcal{C}_{instr}$, by taking the average absolute of each band's uncertainty map as the diagonal elements. The diagonal elements for each galaxy and band are given in Table 3; the off-diagonals are all zero. As can be seen, each band has similar noise across all maps.

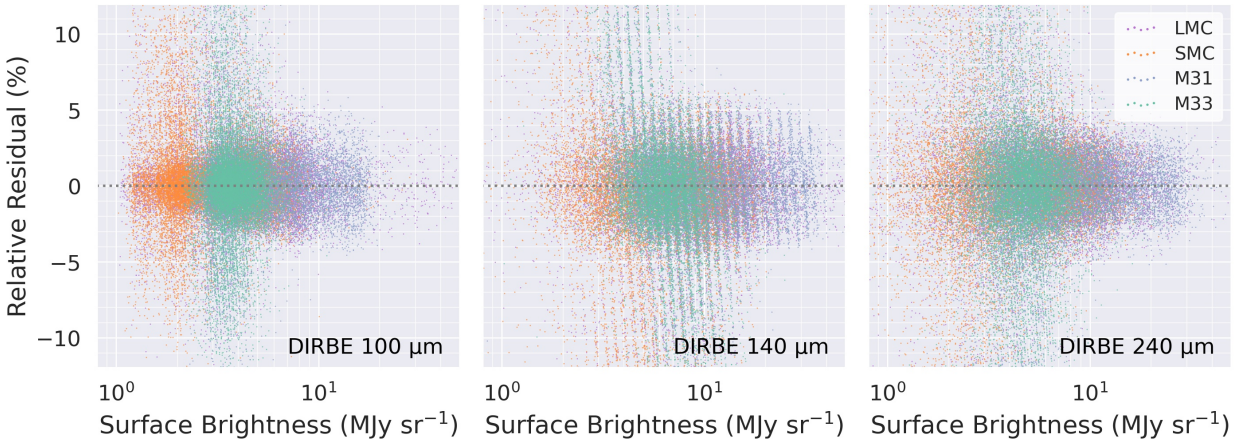

Figure 4. Relative residuals (ie, $S_{mod} - S_{obs}$) of `DustBFF` SED fits to DIRBE pixels for our target galaxies, for the MAP models. The plotted parameter space contains >98% of pixels. The striations, most apparent at 140 μm, are due to the discrete nature of the model grid.

Table 3. Diagonal elements of $\mathcal{C}_{instr}$, the instrumental noise covariance matrix for DIRBE SED fitting (off-diagonals are all zero). All values are in map units of MJy sr^{-1}.

Galaxy	100 μm	140 μm	240 μm
LMC	0.0529	1.90	1.08
SMC	0.0452	1.94	1.11
M 31	0.0589	2.10	1.18
M 33	0.0529	2.45	1.38

We validated the SED fits by checking the residual between the model and the observed surface brightness. In Figure 4, we plot the relative residuals, defined as $\Delta_{rel} = (S_{mod} - S_{obs})/S_{obs}$, for the MAP model of every pixel. For all galaxies and bands, the median Δ_{rel} was <0.7%, with >61% of pixels having $|\Delta_{rel}| < 5\%$. This indicates that `DustBFF` suffered no systematic problems fitting the data.

With the `DustBFF` SED fitting completed, we calculated how each model would appear if it were observed in the IRIS 100 μm band. This was done by convolving every model SED, in each pixel, through the IRAS 100 μm filter, colour-corrected to match the IRIS reference spectrum. The result was 500 images, representing the IRIS 100 μm fluxes predicted from the realisations of the posterior of the SED fit for every pixel, along with the fiducial image, containing the MAP model prediction for every pixel – all at DIRBE resolution. An illustration of this for a particular pixel can be seen in Figure 5.

Our predicted IRIS 100 μm band maps at the DIRBE resolution scale, with uncertainties constrained by our posterior sampling, are now suitable to be feathered with the the unfeathered IRIS maps. This process is described in Section 5.1.

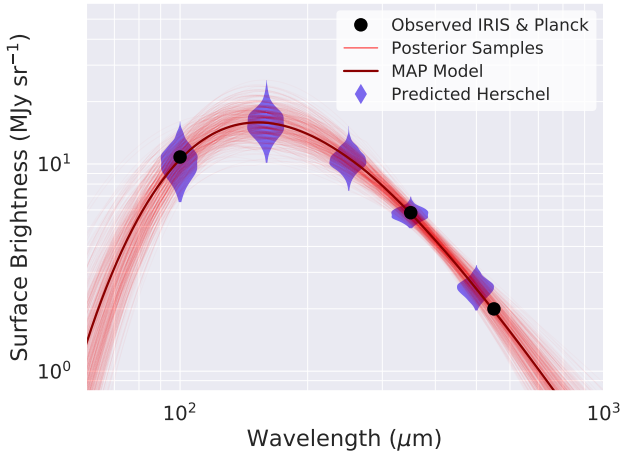

Figure 5. Example of how our `DustBFF` SED fitting is used to predict fluxes, for the case of using IRIS–*Planck* data to predict *Herschel* fluxes for the pixel located at $\alpha = 23.637°, \delta = 30.510°$, in the outskirts of M 33. The black circles show the observed surface brightness in the pixel, from our IRIS and *Planck* data. The thin red lines show 500 posterior samples from the `DustBFF` fit to the observations. The thick dark red line shows the MAP model. The blue violin plots show the distributions of predicted fluxes in each of the *Herschel* bands.

4.3. *SED Fitting with IRIS–Planck to Infer Large-Scale Emission in Herschel Bands*

We fit the SED of IRIS and *Planck* data in order to predict how the emission observed by these facilities would appear in the *Herschel* bands; these predicted maps are can then be used to feather together the high-

Table 4. Diagonal elements of $\mathcal{C}_{instr}$ instrumental noise covariance matrix for IRIS–*Planck* SED fitting (off-diagonals are all zero). All values are in map units of MJy sr^{-1}.

Galaxy	100 μm	350 μm	550 μm
LMC	2.108	0.0154	0.0166
SMC	1.276	0.0167	0.0180
M 31	1.906	0.0265	0.0287
M 33	1.615	0.0308	0.0329

and low-resolution data. Note that in this section we use *the feathered IRIS data, that has already been combined with the DIRBE data*, as described later in Section 4.2[10].

Each of the IRIS and *Planck* bands are at slightly different resolutions, so prior to fitting we convolved each to the 4.8′ poorest resolution of the *Planck* 550 μm band. We did this using conversion kernels creating with the `Python` package `photutils`, using the instrumental PSFs for each band.

We modelled the pixel-by-pixel SEDs of the IRIS 100 μm and *Planck* 350 and 500 μm, observations of our target galaxies, using `DustBFF`. We used the full grid of 24 235 200 models. The correlated uncertainty matrix used for the IRIS–*Planck* SED fitting was:

$$\mathcal{U}_{corr} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0.05 & 0.05 \\ 0 & 0.05 & 0.05 \end{bmatrix} \quad (7)$$

where the rows/columns represent the 100, 350, and 550 μm values respectively. The uncorrelated uncertainty matrix, $\mathcal{U}_{uncorr}$, had diagonal elements of 0.13, 0.014, and 0.011 (with off-diagonals of zero).

The 5% correlated error between the *Planck* bands come from uncertainty on the emission models of Uranus and Neptune, the calibrator sources for *Planck*'s high frequency instrument (Bertincourt et al. 2016). As the IRIS 100 μm band is calibrated independently from the *Planck* bands, it has no correlated uncertainty here. The 13% uncorrelated uncertainty on the IRIS 100 μm calibration is taken from Miville-Deschênes et al. (2005). This combines the 10% uncertainty on the absolute calibration of DIRBE, with which IRIS is calibrated; and an additional 3.7% from uncertainty in the colour corrections used by Miville-Deschênes et al. (2005) to relate IRIS to DIRBE. The 1.4% and 1.1% uncorrelated uncertainties on the *Planck* 350 and 500 μm bands are the statistical repeatability noise (Bertincourt et al. 2016).

The diagonal values of the instrumental noise covariance matrices, $\mathcal{C}_{instr}$, are given in Table 4, having been calculated from the average of the absolute values in the uncertainty map for each galaxy and band. For the 100 μm IRIS data, the uncertainty maps are part of the output from the DIRBE–IRIS feathering (detailed in Section 5.1).

As with the DIRBE SED fitting, we checked the SED fits by inspecting the relative residuals between the MAP model and observed surface brightnesses, for every pixel. This is plotted in Figure 6. There are definite systematic offsets apparent in the relative residuals. Fortunately, however, they remain very small in absolute terms. For IRIS 100 μm, the band with the worst offset, the median Δ_{rel} was only -0.91%, with a standard deviation of 0.99%, and >95% of pixels having $|\Delta_{rel}| < 2.1$%. For the *Planck* bands, median Δ_{rel} values are -0.40% and -0.29% at 350 and 500 μm respectively, with >95% having $|\Delta_{rel}| < 0.91$% in both cases. Given that the systematics are so small, being at the sub-percent level in all cases, with symmetrical distributions, we believe the SED fits to be satisfactory. It is odd that the offsets are negative in *all* bands, but we suspect this may arise from the logarithmic spacing of the dust mass surface density (and/or temperature) increments in the parameter grid.

Having completed the IRIS–*Planck* SED fitting[11], we used the resulting posterior to predict how the modelled emission would appear in the *Herschel* 100–500 μm bands, convolving each model through the *Herschel* filters, colour-corrected to match the *Herschel* reference spectrum. And example of this for the model SEDs of a particular pixel are shown in Figure 5. The output was 500 images, representing the *Herschel* fluxes in each band predicted from the realisations of the posterior of the SED fit for every IRIS–*Planck* pixel, along with that of the fiducial Maximum A-Posteriori (MAP) image, containing the MAP prediction for every pixel, all at DIRBE resolution.

We originally attempted this SED fitting process also using the *Planck* 850 μm band as well, to provide additional constraints for the models. However, we found that this resulted in considerably larger residuals, positive and negative, dominated by the structure of cosmic microwave background radiation, which is more prominent at 850 μm. For this reason, we opted to not use the *Planck* 850 μm data.

[10] We present the SED fitting of the feathered IRIS data here, before we describe the feathering of the IRIS data, in order to keep together our descriptions of all the flux-prediction SED-fitting.

[11] The IRIS–*Planck* SED fitting took about a week apiece for the Magellanic Clouds, and about a day each for M 31 and M 33, on a 3.2 GHz × 32 core computer.

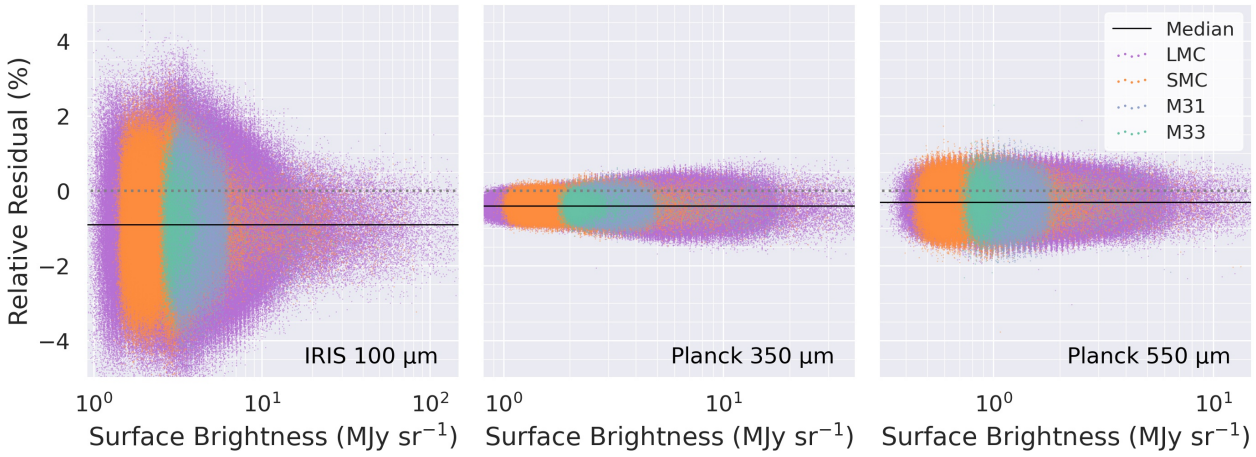

Figure 6. Relative residuals (ie, $S_{mod} - S_{obs}$) of `DustBFF` SED fits to IRIS–*Planck* pixels for our target galaxies, for the MAP models. Over 99% of pixels fall into the parameter space plotted. The median relative residual for each band is also plotted; for each band, the difference between the galaixes' medians was less than 1 part in 500 in all cases.

5. FEATHERING

The process of feathering is widely employed to allow low-resolution single-dish data to restore missing large-angular-scale emission to interferometric observations. Large-angular-scale emission is ofter missing from interferometric data due to the fact that no arrangement of dishes can sample the shortest spacings, therefore making interferometers insensitive to the lowest frequency parts of the *u-v* plane (ie, the lowest frequency parts of the Fourier domain). However, this shortcoming of interferometry can be overcome by replacing the poorly-sampled parts of the *u-v* plane with data from single-dish observations, which don't suffer from the short-spacing problem (Weiß et al. 2001; Stanimirovic 2002).

The feathering technique of combining high- and low-resolution data in Fourier space is especially well-suited to single-dish FIR–submm observations. In multi-dish interferometry, the *u-v* plane is sparsely sampled due to the fact that there are only a finite number of dishes (and therefore baselines). This complicates the process of successfully combining the data with single-dish observations to provide complete sampling of all relevant spatial frequencies. With both of our datasets being single-dish, however, the situation is more straightforward.

The low-resolution data provides full coverage of the *u-v* plane at spatial frequencies below the instrument's large PSF. The high-resolution data provides coverage of the *u-v* plane at spatial frequencies between the instrument's small PSF, up to some cutoff dictated by the observing and reduction strategy, above which emission is filtered (see Section 1.1). As long as the resolution of the low-resolution data is better than the largest scale to which the high-resolution data is sensitive – ie, as long as the spatial scales sampled by the two instruments have some overlap – then the observations can be combined via feathering.

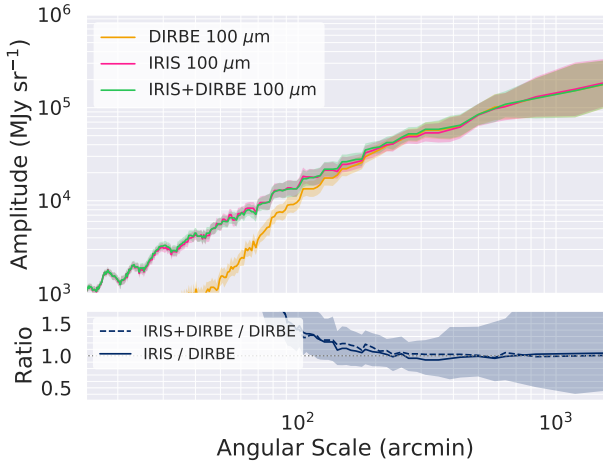

Figure 7. The power spectra of the DIRBE data, IRIS data, and IRIS+DIRBE feathered data, for the SMC. The shaded region around each line shows the uncertainty on the average binned spectrum amplitudes at each scale. The panel beneath the main plot shows the ratios of the IRIS amplitudes to the DIRBE amplitudes, and of the feathered IRIS+DIRBE amplitudes to the DIRBE amplitudes; the horizontal dotted line indicates a ratio of unity.

In practice, there are several steps to the feathering process:

1. Place the two observations on the same pixel grid.

2. Transform both observations to Fourier space.

3. Deconvolve the low-resolution data with the low-resolution beam.

4. Re-convolve the low-resolution data with the high resolution beam.

5. If possible, correct for any gain differences between the brightness scales of the observations.

6. Replace the low-frequency components of the high-resolution data with the low-frequency components of the low-resolution data, with some weighting function imposing a smooth transition between the two.

7. Transform the combined data back out of Fourier space.

For a more formal description of the mathematics behind the feathering process, see Vogel et al. (1984), Stanimirovic (2002), and Weiß et al. (2001).

As outlined in Section 1.2, producing our final data entails two stages of feathering. First, we combine DIRBE and IRIS data; secondly the resulting feathered IRIS maps, together with *Planck* data, are combined with *Herschel* data to produce our final maps. Because of differences between these various data sets, the specific implementations of the feathering process for each of the two stages vary somewhat; descriptions of both are provided in the following subsections. In Appendix C1, we perform in/out simulation tests to verify that the feathering methodology we use is effective.

The code we used to perform the feathering process is available online at: https://doi.org/10.5281/zenodo.4776266. That code includes a main function to conduct the feathering itself (`FourierCombine`), along with sub-functions to handle the cross-calibration and construction of a tapering function (`FourierCalibrate` and `FourierTaper`).

5.1. *Feathering Together IRIS and DIRBE*

The only band for which we feathered together DIRBE and IRIS data was at 100 μm. The DIRBE data we used are the maps we produced in Section 4.2, predicting the emission that DIRBE would have observed in the IRIS 100 μm band. For our fiducial feathered maps, we combine the IRIS cutout for each galaxy with the image produced from the MAP outputs for the pixel-by-pixel DIRBE SED-fitting (Section 4.2).

For each target galaxy, we reprojected this DIRBE data to the IRIS pixel grid. We also reprojected the DIRBE and IRIS PSF maps to have the same dimensions and pixel widths as the IRIS observation. As part of the reprojection, we also had to apodise the DIRBE maps, to remove pixel-edge artefacts from the re-gridded maps; we detail this in Appendix D.

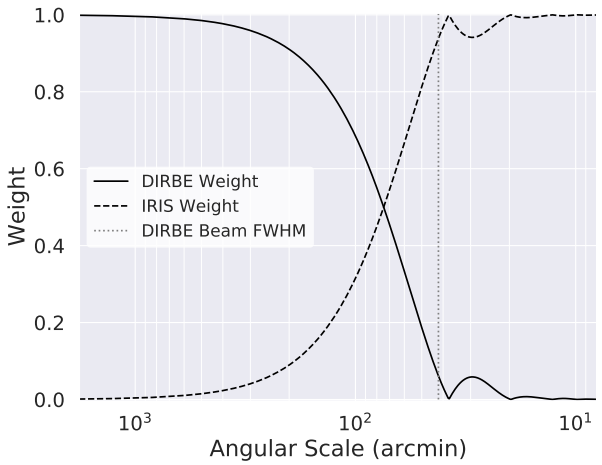

Figure 8. The weighting applied at different angular scales (ie, in Fourier space) when feathering together IRIS and DIRBE data. The 0.7° scale of the DIRBE beam FWHM is shown for reference. As the DIRBE beam is not Gaussian, the crossover point in Fourier space does not occur exactly at the angular frequency corresponding to 0.7° – rather, because the DIRBE beam is flatter than a Gaussian, with relatively more weight at larger radii, the crossover is also at larger radii. The beam non-Gaussianity is also what causes the slight ringing structure at smaller scales

With all of the data on the same pixel grid, we transform the observations and beams to Fourier space. Next, we deconvolve the DIRBE data with the DIRBE beam, then convolve it with the IRIS beam. This accounts for the brightness differences between the data sets that were caused *only* by the differences in their resolution; ie, where a given source would have had a lower peak brightness in the low-resolution data. The DIRBE data now has the same beam amplitude and effective resolution as the IRIS data[12], meaning the two can be directly compared and combined.

The spatial scales at which systematic response errors are observed in the IRIS data ($\sim 1°$; Bot et al., *in prep.*) are only somewhat larger than the spatial scales above which DIRBE is sensitive, as dictated by its 0.7° beam. The power spectra of both datasets, for the example of the SMC, are shown in Figure 7. The amplitude of the IRIS power spectrum is not considerably different from that of DIRBE at the largest scales, because the errors in

[12] The deconvolution of the low-resolution data introduces almost infinite noise to the low-resolution data at small spatial scales; however, because we will only be extracting large spatial scale information from this data when feathering, this causes no problems.

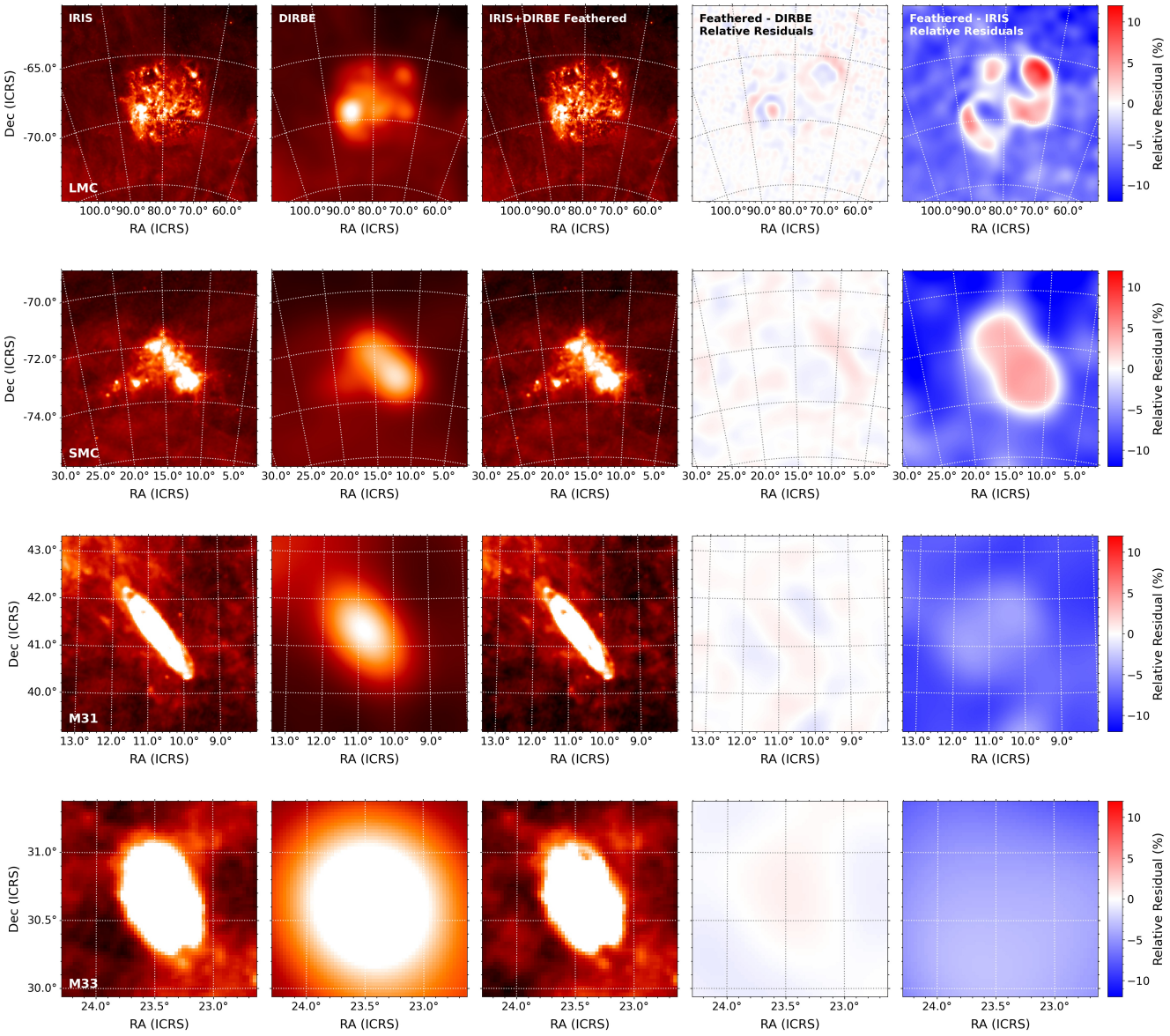

Figure 9. The results of feathering together IRIS and DIRBE data at 100 μm for each of our galaxies, along with diagnostic plots. In all cases, our fiducial data are shown. *1st Panel:* The un-corrected IRIS data. *2nd Panel:* The DIRBE data. *3rd Panel:* The combined IRIS+DIRBE data after feathering (all three surface brightness maps for each galaxy use the same colour scale). *4th Panel:* The relative residuals between the feathered IRIS+DIRBE data and the input DIRBE data, compared at the DIRBE resolution; positive residuals mean the feathered data has greater surface brightness than the DIRBE data, negative mean less. Ideally, this should be zero everywhere. *5th Panel:* The relative residuals between the IRIS+DIRBE data and the unfeathered IRIS data; positive residuals indicate where the IRISD surface brightness was increased by feathering, negative where it was decreased.

IRIS are localised only around certain parts of the LMC and SMC; the IRIS data is otherwise well-behaved. This can also be seen in the 5th column of Figure 9.

For feathering together the DIRBE and IRIS data in Fourier space, we chose to follow the weighting strategy used in the Common Astronomy Software Applications (CASA; McMullin et al. 2007) task `feather`[13]. In this approach, the Fourier transform of the low-resolution beam is used as the weighting function. Therefore, at the largest scales (ie, the zeroth Fourier mode), the DIRBE data is assigned a weight of 1. Going to smaller

[13] https://casa.nrao.edu/casadocs/casa-5.4.1/image-combination/feather

spatial scales, the weighting of the DIRBE data decreases, whilst the weighting of the IRIS data increases; the sum of both weights remains 1 everywhere. At scales smaller than the DIRBE beam FWHM, the weights cross over so that the IRIS data dominates, and becomes increasingly heavily weighted. This is illustrated in Figure 8.

The CASA algorithm minimises ringing by making the weighting transition as smooth as possible, tapering according to the low-resolution beam, using its FWHM as the natural crossover scale. A disadvantage of this approach is that the high-resolution data will nonetheless be assigned some (albeit minimal) weighting at the larger scales, where it has poor sensitivity (and vice-a-versa for the low-resolution data). However, If one were to instead attempt to 'squeeze' the transition region into a smaller range of scales, where the two datasets' sensitivity overlaps, there will be considerable ringing in the combined image if the transition window is too narrow. This is what would happen if we attempted to combine the DIRBE and IRIS data with such a transition window; the small-scale edge of such a window would be dictated by the 0.7° DIRBE beam, whilst the large-scale edge of the window would be dictated by the $\sim 1°$ scale of the IRIS response errors around the LMC and SMC.

We therefore opt instead to take advantage of the ringing minimisation provided by the CASA algorithm. We test this beam tapering method, by feathering simulated high- and low-resolution data, in Appendix C.

Comparison of the DIRBE and IRIS power spectra (Figure 7)) suggests there are no significant deviations between the brightness scales. IRIS was constructed to be cross-calibration to the DIRBE brightness scales, so this is not surprising (indeed, it seems that the IRIS response errors around the LMC and SMC are localised failures of this cross-calibration). Because the largest scales of the feathered maps are dictated entirely by the DIRBE data, the zero-level of the output maps will match that of DIRBE, by construction.

We carried out the feathering process 500 times, once for each sample of the DIRBE SED-fitting posterior interpolations produced in Section 4.2. We also feathered the fiducial DIRBE data with the unfeathered IRIS data to produce our fiducial IRIS+DIRBE map. For each of these 500 iterations, we permutated the surface brightness values in each IRIS pixel according to the IRIS uncertainty maps, assuming a random Gaussian noise distribution. The 500 resulting feathered maps therefore provide a bootstrapped Monte Carlo estimate of the uncertainty on the feathered output in each pixel.

5.1.1. *Validation of Feathering IRIS with DIRBE*

The results of the IRIS+DIRBE feathering are shown in Figure 9. The first three columns compare the IRIS, DIRBE, and feathered IRIS+DIRBE data for each galaxy. The last two columns show our two main validation tests for the feathering process.

The 4$^{\text{th}}$ column of Figure 9 shows the relative residuals between the DIRBE map and the feathered map; for this comparison, we convolved the feathered map to the DIRBE resolution. Ideally this residual should be 0% everywhere, as the feathered maps should conform to the DIRBE data at scales $\geq$ the DIRBE beam. There are slight ringing artefacts visible around the LMC and SMC, but at worst they are $\pm 3\%$ (for the LMC). This is much lower than either instruments' calibration uncertainty.

The very small residuals between the feathered and IRIS data also compare very favourably to the 5$^{\text{th}}$ column of Figure 9. This column shows the relative residuals between the feathered data and the unfeathered IRIS data; this illustrates the magnitude of the correction applied to the IRIS data by the feathering process. As before, this comparison is plotted at the resolution of DIRBE. The feathered data has significantly more surface brightness over the LMC and SMC, versus in the unfeathered IRIS data, correcting the response errors. The change in surface brightness is up to 10% for the LMC, and up to 5% for the SMC. M 31 and M 33 are not bright and/or extended enough to have suffered from the IRIS response issue, so do not display these changes in the feathered maps. The data for all four galaxies also shows a reduction in surface brightness over the background surrounding all galaxies, of $\sim 8\%$ in all cases; this is due to an apparent difference in the zero-level of the DIRBE and IRIS data.

The peak of 10% extra flux in bright regions, super-imposed over an $\sim 8\%$ reduction in background level, corresponds to an up-to 18% correction to the background-subtracted surface-brightness for the Magellanic Clouds. This correction is substantial, and clearly necessary to obtain good science analysis with the data; it is also in line with the $\approx 20\%$ IRIS response errors measured in other work (Bot et al., *in prep.*).

5.2. *Feathering Together Herschel and IRIS–Planck*

Here, we feather together our *Herschel* reductions with our inferred maps of large-scale emission in *Herschel* bands (produced using IRIS and *Planck* data in Section 4.3). We carried out this process for the 100, 160, 250, 350, and 500 μm data.

We first generated power spectra for of the IRIS–*Planck* and *Herschel* data for each of our galaxies and

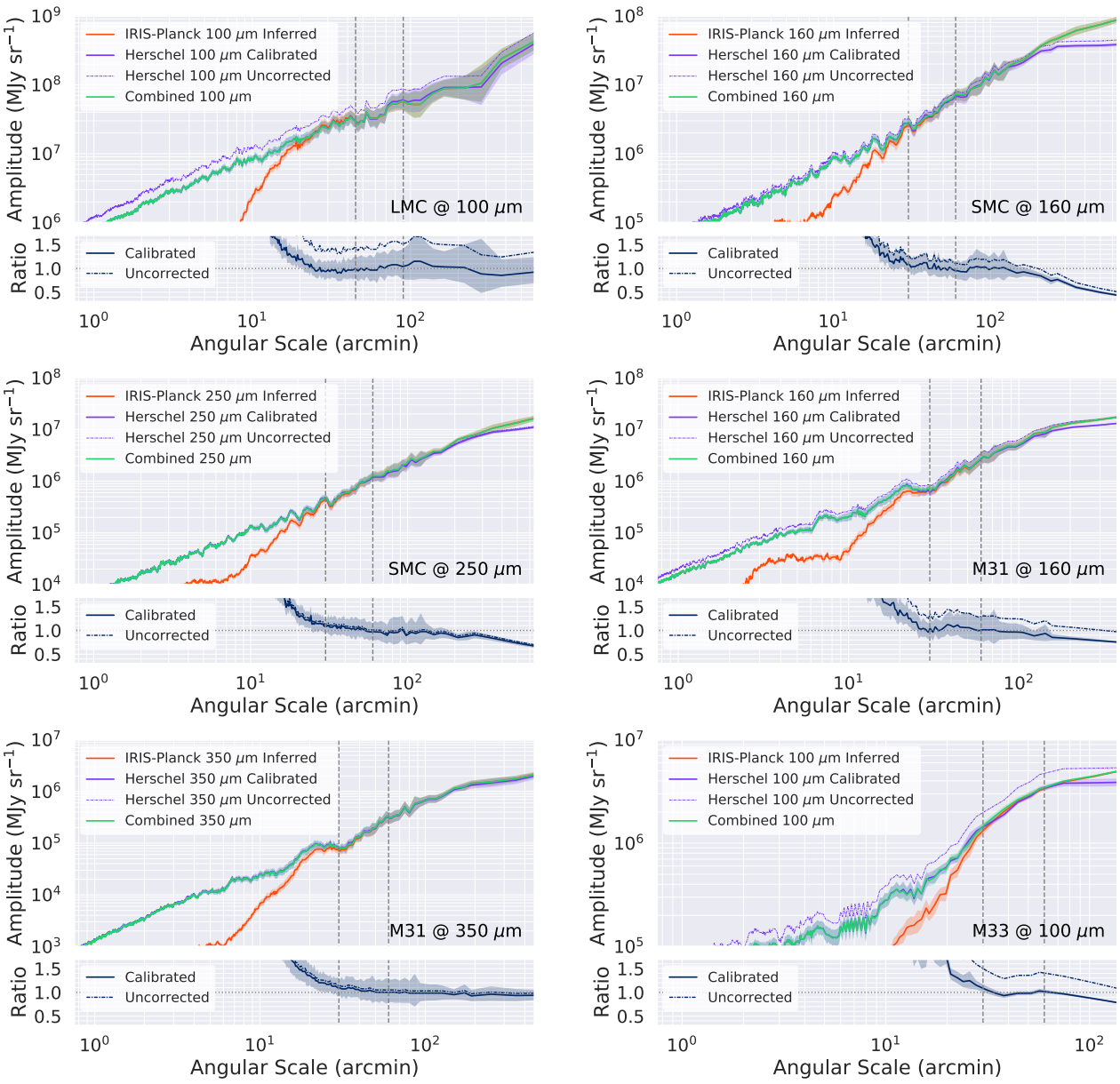

Figure 10. Example power spectra. The shaded region around each line shows the uncertainty on the average spectrum amplitude at each scale. Vertical dashed lines demark the 30′–60′ tapering and calibration window (45′–90′ for the LMC). In most bands the power in the *Herschel* data falls significantly beneath that in the IRIS–*Planck* data are large angular scales. Both the cross-calibrated and uncorrected *Herschel* power spectra are shown, to illustrate the offset between the two. The panel beneath each plot shows the ratio of the high-resolution *Herschel* amplitudes to the low-resolution IRIS–*Planck* amplitudes; the horizontal dotted line indicates a ratio of unity.

bands, some examples of which are shown in Figure 10); these plots also show the ratio between the amplitudes of the two power spectra. At the largest angular scales, the *Herschel* power spectrum tends to fall beneath the IRIS-*Planck* power spectrum, by as much as 50% (eg, see SMC 500˙μm panel of Figure 10).

We note that in some instances no power on large scales appears to be missing, such as at 500 μm for M 31 or M 33; this is potentially due to the smaller angular size of these galaxies, care taken during the reduction process, and the reduced filtering suffered needed at 500 μm thanks to its larger beam. However, it should be remembered that even if the power spectra for *Herschel* shows no deficit in power compared to IRIS–*Planck*, that does not preclude the possibility of the emission being distributed differently. For instance, as described in Section 5.2.1, the LMC, SMC, and M 31 data all show gradients in the foreground cirrus emission in the IRIS–

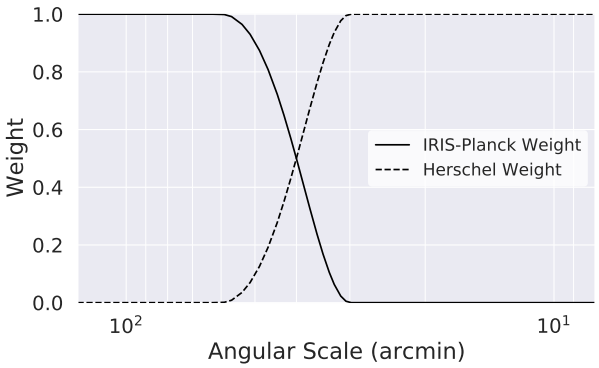

Figure 11. The weighting applied at different angular scales (ie, in Fourier space) when feathering together *Herschel* and IRIS–*Planck* data. Below 30′, the *Herschel* data has a weight of 1, and the IRIS–*Planck* data has a weight of 0; above 60′, these are reversed. Within the 30′–60′ tapering window, the weights smoothly swap, following a half cosine bell function.

Planck data that is considerably attenuated in the *Herschel* data; the presence (or not) of these gradients significantly effect the emission measured in the outskirts of our target galaxies after background subtraction (see Section 6). We therefore apply the full feathering process to all bands, even when no difference appears in the power spectra, to deal with these sorts of problems, and in order to treat all of our data consistently throughout.

Inspecting the power spectra in Figure 10 indicates that there is a reasonably wide range of spatial scales over which both sets of data are sensitive to emission. We therefore used a feathering technique that could take advantage of this. Within this window of overlapping spatial scales, a smooth tapering function was used to mediate the transition in weights, from using the low- to high-resolution data. At spatial scales larger than the taper window, only the low-resolution data was used; and at scales smaller, only the high-resolution data. The overlap in spatial sensitivity between both sets of data also made it possible to cross-calibrate them, to ensure that the low- and high-resolution data were both on the same brightness scale before being combined.

To find an appropriate tapering window, we inspected the power spectra, to identify a specific spatial scale interval for which the ratio between the high- and low-resolution amplitudes remained relatively constant. This constant ratio indicates that the two sets of data are maintaining a constant spatial sensitivity, except for a difference in flux calibration (if the ratio is 1, then the two datasets already have the same flux calibration). It was necessary to balance the benefits of a larger tapering window (less ringing in the feathered outputs; more values with which to perform brightness scale cross-calibration), versus the benefits of a narrower tapering window (less risk of extending sampling to spatial scales where either of the datasets starts to lose sensitivity).

Even within the tapering window, the ratio between the high- and low-resolution amplitudes does not remain perfectly constant. To check if there were systematic variations in the ratio within each tapering window, we fit straight lines to the amplitudes inside them; ideally the gradient of such a line should be zero. We performed a Monte-Carlo bootstrap resampling to evaluate the gradient uncertainty. For 17 of the 20 power spectra (ie, for each band for each galaxy), the gradient of the ratio within the tapering window was compatible with being 0, to within the uncertainty. For the remaining 3, the deviations were 1.03σ, 1.20σ, and 2.15σ. Given that 1 out of 20 *should* deviate by $>2\sigma$ on average for Gaussian statistics, we are satisfied that this indicates the amplitude ratios within the tapering windows are well-behaved.

After iterating through a range of options and inspecting the outputs, particularly the residual plots (see Section 5.2.1 and Figure 13), we found that a tapering window of 30′–60′ worked best; this range is marked in Figure 10. Specifically, within this window of overlapping sensitivity, we used a half cosine bell function (ie, half of a Hann window function; Harris 1978) to smoothly mediate the transition in weights applied to the datasets. This transition is illustrated in Figure 11.

Based upon a wide suite of feathering simulations, Kurono et al. (2009) find that the high- and low-resolution observations should have spatial scale overlap spanning at least a factor of 1.7, in order to achieve outputs with minimal errors (above this overlap factor, they find minimal, asymptotic improvements). Given that our 30′–60′ tapering window spans a factor of 2, we comfortably satisfy this criterion.

When feathering the PACS data for the LMC, we found that shifting the window function to larger angular scales was necessary, due to an apparent non-linearity in the response function of our IRIS–*Planck* data around the extremely high-surface-brightness star-forming complex of 30 Doradus. This response non-linearity caused the IRIS–*Planck* data to over-estimate the surface brightness at the centre of 30 Dor, and under-estimate it in the surrounding area (similar response issues around 30 Dor were discussed in Meixner et al. 2013). These errors appear at scales of up to ∼40′. As we describe in Appendix E, it appears that the PACS data is correctly recovering the emission around 30 Dor at these scales. By putting our tapering window at a larger scale, we ensure the emission from 30 Dor is correctly reproduced in the final feathered maps. We experimented with a range of tapering windows, and

Table 5. Cross-calibration correction factors applied to *Herschel* data for each band and galaxy, after comparison with power spectra of the corresponding IRIS–*Planck* data (within the tapering window), in order to place both on the same brightness scale prior to feathering together.

	LMC	SMC	M 31	M 33
100 μm	0.688 ± 0.006	0.705 ± 0.007	0.718 ± 0.014	0.723 ± 0.015
160 μm	0.797 ± 0.004	0.860 ± 0.006	0.777 ± 0.008	0.717 ± 0.010
250 μm	0.895 ± 0.002	0.962 ± 0.009	0.868 ± 0.011	0.850 ± 0.014
350 μm	0.943 ± 0.002	0.980 ± 0.006	0.955 ± 0.009	0.903 ± 0.024
500 μm	0.967 ± 0.003	0.995 ± 0.006	0.996 ± 0.011	0.928 ± 0.016

found that 45′–90′ stops emission being lost around 30 Dor, whilst still allowing the IRIS–*Planck* data to correct emission at larger scales. This maintains a tapering window spanning a factor of 2 in angular scale, so will satisfy the Kurono et al. (2009) factor >1.7 overlap criterion just as well as for the other galaxies.

To cross-calibrate the two datasets to the same brightness scale (ie, ensure the gain for both agrees), we divided the low-resolution amplitudes by the high-resolution amplitudes within the overlap window, then took the median of these ratios. We multiplied the high-resolution amplitudes (over all scales) by this ratio to fix them to the brightness scale of the low-resolution data (which has absolute calibration ultimately pegged by DIRBE and *Planck*). We estimated uncertainties on the cross-calibration factors by performing 100 bootstrap resamplings (with replacement) of all the amplitudes in the overlap window, and re-computing the correction factor each time; the standard deviation of these values was taken as the uncertainty.

The cross-calibration factors for each band and galaxy are listed in Table 5. At 500 μm, relatively little correction is necessary, with all factors being > 0.928. However, going to shorter wavelengths, especially into the PACS bands, larger corrections are necessary; the average correction at 100 μm is 0.706. The fact that PACS measures emission as being 20–30% brighter than IRIS (and *Planck*) is all the more striking given that we have shown that PACS can miss a considerable amount of emission at large scales. This means that PACS must be overestimating the brightness it *does* detect by even *more* than 20–30%. When comparing PACS and IRIS–*Planck* at a common resolution *without* applying the cross-calibration corrections, we found that regions of compact sources were indeed much brighter in PACS, over 40% so in some cases.

PACS is flux calibrated by reference to a set of 5 standard stars (Balog et al. 2014); the fact this calibration is based off point sources may explain why the calibration appears to struggle at the highly extended scales we are working with. Unlike SPIRE, PACS makes no corrections to the relative gains of the bolometers to account for the different illumination by the beam when making maps of extended emission. This may contribute to the need for a large gain correction. We also note that the two largest correction factors are for the LMC and SMC 100 μm maps; it is therefore possible that the issue is partly due to the `UNIMAP` pipeline, perhaps in how it handles the unusual scan strategy for the LMC and SMC observations.

It is conceivable that this calibration discrepancy is not present at scales smaller than the 6′ low-resolution beams, beneath which we cannot compare the *Herschel* data to the IRIS–*Planck* data. It is therefore possible that point source and compact sources *are* accurately represented in the unfeathered data. However, this would require a very abrupt shift from well-calibrated to poorly-calibrated as angular scale increased. If this is the case, then our cross-calibrated maps would make point sources too faint, whilst accurately representing extended emission (compared to the original maps making diffuse emission too bright, but accurately representing point sources).

Previous authors have compared *Herschel* data to all-sky survey data, such as from IRAS and *Planck*, to check for differences in gain. For instance, Molinari et al. (2016) compared *Herschel* and IRIS data for observations of the Galactic plane (which therefore contain highly extended emission), and didn't find evidence for gain discrepancies $> 10\%$. However, they made this comparison pixel-by-pixel, not in Fourier space, which could give undue weight to compact features. In contrast, Abreu-Vicente et al. (2017) *do* make such a comparison in Fourier space, also for Galactic plane fields – including data from Molinari et al. 2016 – and find average cross-calibration factors of 0.84, 0.88, 0.91, and 0.97 necessary at 160, 250, 350, and 500 μm respectively (measured between 7′–100′). These factors all lie within the range of values we find for each of these bands.

Regardless, the feathering process requires us to cross-calibrate the two datasets being combined to the same brightness scale, otherwise we find that the differences

lead to very pronounced artefacts – and over the scales we *are* able to probe, the *Herschel* response (especially for PACS) is too bright compared to IRIS–*Planck* (and therefore DIRBE also). Given that the science in this paper, and the subsequent papers in this series (Clark et al., *in prep.*), is concerned primarily with the diffuse ISM, our priority is that this is correctly calibrated – which is what our cross-calibration ensures.

We also note that the cross-calibration factors determined are quite robust against the specific choice of tapering window. As can be seen from inspecting Figure 10 (especially the plots for PACS bands), the offsets between the uncorrected *Herschel* amplitudes and the IRIS-*Planck* amplitudes stay remarkably constant, over all scales large enough for IRIS-*Planck* to be sensitive.

Before feathering, we reprojected the low-resolution data to pixel grid of the *Herschel* data in each band, apodising it to remove pixel-edge artefacts (as per Appendix D). For each band, both sets of data, and their corresponding beams, were Fourier transformed. The low-resolution data was deconvolved with its beam, then re-convolved with the band's corresponding *Herschel* beam. Having done this, we combined the two sets of data in Fourier space, using our tapering function to govern the transition within the tapering window.

As with IRIS+DIRBE, we performed the feathering 500 times for each galaxy and band; once for each sample of the SED-fitting posterior in each pixel. For each Monte Carlo iteration, we drew surface brightness values for each *Herschel* pixel according to the uncertainty map in each band, assuming Gaussian uncertainty. Recall that the uncertainties on the IRIS–*Planck* SED-fitting factored in the uncertainties on the IRIS+DIRBE feathering, which in turn factored in the uncertainties on the DIRBE SED fitting (and the IRIS photometric uncertainties). So the 500 bootstrap samples produced for each feathered map encompass the propagated uncertainty from all previous stages of our process.

The abrupt edge of the *Herschel* maps result in considerable edge effects at the borders of the feathered maps. We therefore constructed a mask to describe the portion of each feathered map far enough from the edge to not be affected. Specifically, we found that excluding a border of three times the low-resolution beam (ie, 14.5′) provided sufficient buffer. We exclude this region from our various diagnostic plots assessing the feathered data.

5.2.1. *Validation of Feathering Herschel with IRIS–Planck*

The results of feathering together the *Herschel* and IRIS–*Planck* data are shown in Figure 13, for the example case of the 250 μm data for each; corresponding figures for 100, 160, 350, and 500 μm are shown in Appendinx F. A zoomed-in view of the SMC wing, contrasting the unfeathered and feathered maps, is shown in Figure 12.

The 4$^{\rm th}$ column of these figures show the relative residuals between the IRIS–*Planck* map and the feathered map (with the feathered map convolved to the low-resolution beam for comparison); Ideally this residual should be 0% everywhere. The 5$^{\rm th}$ columns show the relative residuals between the feathered data and the unfeathered *Herschel* data, illustrating the degree to which the unfeathered was corrected by feathering; note that the zero level of these plots are somewhat arbitrary[14]. If the correction in the 5$^{\rm th}$ column is larger than the residual in the 4$^{\rm th}$ column, then the feathering process has done 'more good than harm'.

In Figures F4 and F5, in Appendix F, we can see that the average residual in the 4$^{\rm th}$ column is zero for every galaxy, as we would wish. There is some ringing apparent surrounding bright features, such a star-forming complex in the SMC wing, and the rings of M 31; these residuals peak in the ±5–8% range (the background regions where there is no ringing have residuals ≤1%). In contrast, the flux corrections in the 5$^{\rm th}$ column show that surface brightness was increased by 10–13% over most of the outskirts of the target galaxies. The distribution of emission restored by feathering is exactly as we would expect – regions of bright compact emission show little-to-no correction, whereas in the diffuse outer regions there has been significant addition. We can also see that feathering has corrected the flat backgrounds that *Herschel* reduction tends to impose, restoring the gradients due to foreground cirrus; this is especially clear to the north-west and south-east of the SMC.

These diagnostic plots show some noteworthy differences in the feathering results in different bands. The ringing in the residuals in the 4$^{\rm th}$ column are less pronounced in the other SPIRE bands, at 350 and 500 μm, (Figure F5), falling to ±3.5% at 500 μm. The flux corrections in the 5$^{\rm th}$ column also become smaller towards longer wavelengths. By 500 μm, little-to-no-correction is being made to M 31 and M 33, with only slight changes to the foreground gradient apparent. This is borne out by the power spectra, which show no significant power

[14] The zero levels for the plots of relative residuals, between the IRIS–*Planck* data and the unfeathered *Herschel* data, are relative and somewhat arbitrary. This is because the *Herschel* data has no absolute calibration. We have attempted to place the unfeathered *Herschel* maps on a zero level that matches that of the feathered data, by comparing the peaks of their pixel value distributions. But because the surface brightness distributions of the two sets of data are, different due to their different angular sensitivities, this is only approximate.

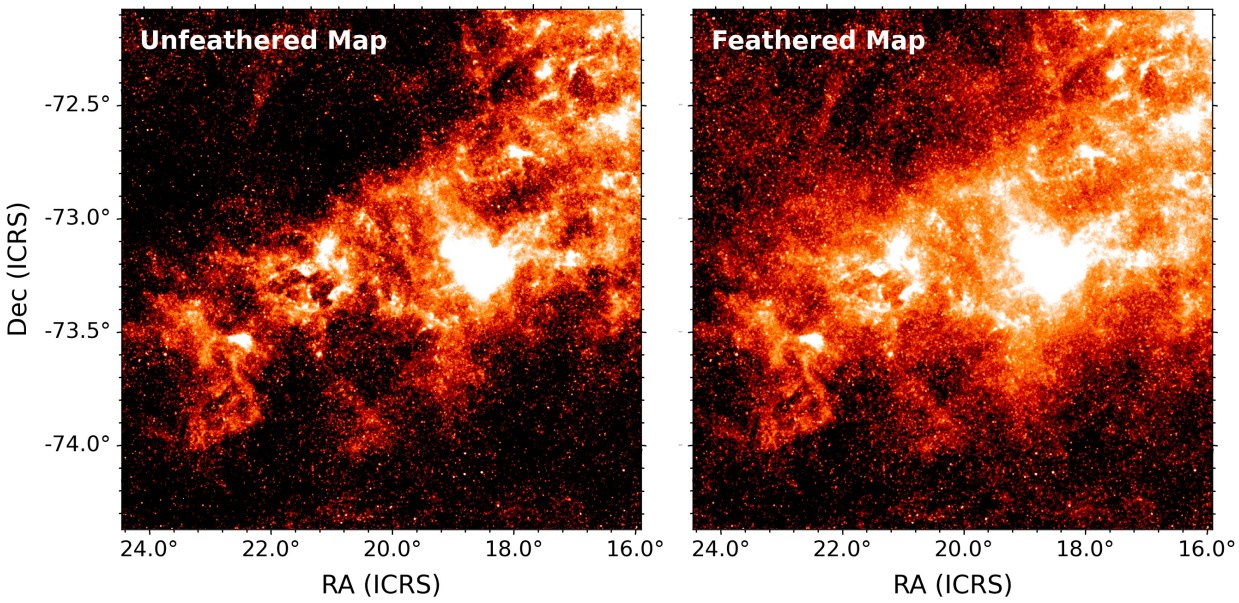

Figure 12. The southern portion of the SMC, centered on the SMC wing, as seen at 250 μm in both the unfeathered map (*left*) and our feathered map (*right*). Note that these images are on the same colour scale, and have both had foreground Milky Way emission subtracted, as per Section 6.

missing in the *Herschel* data, compared to the IRIS–*Planck* data at 500 μm for either galaxy (eg, see lower-left panel of Figure 10).

The feathering diagnostic plots for the PACS data at 100 and 160 μm (Figure F4) show even more dramatic corrections to the surface brightness, as can be clearly seen in the 5$^{\text{th}}$ columns. The restored surface brightness represents a >30% increase over large areas around all of the galaxies[15], for both PACS bands, once again in line with expectations from the power spectra in these bands (see Figure 10). Unfortunately, the noise in the residuals between the feathered data and IRIS–*Planck* data in the 4$^{\text{th}}$ columns can be considerable. The ringing around M 31 reaches 17% level for 160 μm, and 20% for 100 μm. Nonetheless, this is still much smaller than the correction applied by feathering, so the new maps are still more correct than the unfeathered maps, even with this residual noise. The only exception to this is the region of sky to the south-east of M 33, where there is a large ≈40% positive residual in the 4$^{\text{th}}$ column, whilst the correction in the 5$^{\text{th}}$ column is only reaches ≈30%. In general, large residuals in the 4$^{\text{th}}$ column are limited to the regions of empty background; within the galaxies (as outlined by the contours), the mean absolute residuals of the 4$^{\text{th}}$ column are ≈4.5% at 100 μm, and ≈4.0% at 160 μm. This compares favourably to the 7% photometric calibration uncertainty of the PACS instrument.

As noted in Section 5.1.1 and Appendix E, our IRIS–*Planck* maps do not correctly recover the emission around 30 Dor in the LMC. So the large residuals around 30 Dor in the 4$^{\text{th}}$ column of Figure F4 for PACS is expected (and indeed desirable), indicating where there is disagreement between the feathered data and the IRIS–*Planck* data, due to the PACS data correcting IRIS–*Planck* surface brightness artefacts at smaller scales.

Each of our feathered maps has an associated uncertainty map, which captures the uncertainties propagated through the entire process. In most bands, the increase in the uncertainty maps is negligible (especially in the PACS bands, where the instrumental uncertainty levels were already quite high). However, in the SPIRE bands the levels in the uncertainty maps can increase by a factor of a few, mainly because the instrumental noise levels in the original maps were very low to start with. Even with this increased uncertainty, the cirrus emission in faint sky regions away from our target galaxies is still consistently detected with SNR > 10 in the SPIRE bands.

6. FOREGROUND SUBTRACTION

The feathered maps produced in Section 5 include not only the restored large-angular-scale emission from our target galaxies, but also highly extended emission from the foreground of Milky Way cirrus emission. We must

[15] Note that this is separate from the cross-calibration corrections, as listed in Table 5, applied prior to feathering.

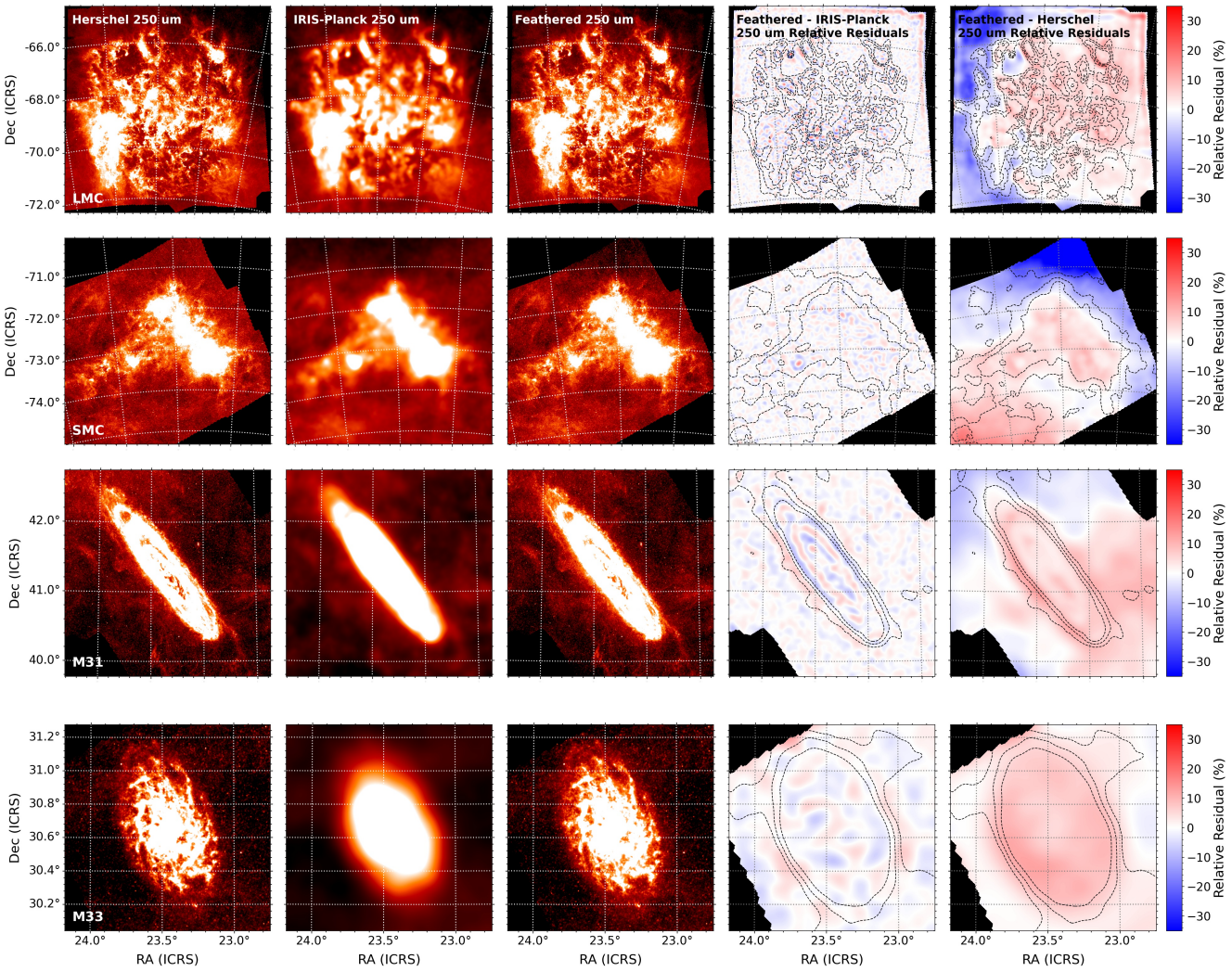

Figure 13. The results of feathering together *Herschel* with IRIS–*Planck* data at 250 μm for each of our galaxies, along with diagnostic plots. *1st Panel:* The un-corrected *Herschel* data. *2nd Panel:* The IRIS–*Planck* data. *3rd Panel:* The combined *Herschel* + IRIS–*Planck* data after feathering (all three surface brightness maps for each galaxy use the same colour scale). *4th Panel:* The relative residuals between the feathered data and the input IRIS–*Planck* data, compared at the IRIS–*Planck* resolution; positive residuals mean the feathered data has greater surface brightness than the DIRBE data, negative mean less. Ideally, this should be zero everywhere. *5th Panel:* The relative residuals between the IRIS–*Planck* data and the unfeathered *Herschel* data (to which the cross-calibration factors from Table 5 have been applied); positive residuals indicate where the IRIS–*Planck* surface brightness was increased by feathering, negative where it was decreased; note that the zero-level for this is somewhat arbitrary[14]. The contours in the 4$^{\rm th}$ and 5$^{\rm th}$ panels show SNR of 2, 5, and 10, from the IRIS–*Planck* map, using simple pixel noise in calculated in background regions.

subtract this foreground emission from the map before we can investigate the dust in our targets.

The Milky Way cirrus is highly structured, with much of this structure at angular sizes similar to our target galaxies. As a result, simple foreground subtraction with an annulus or similar will not do a good job of removing the foreground cirrus structure directly in front our targets. We therefore follow (Roman-Duval et al. 2017), and many previous works, in using all-sky maps of Galactic HI to trace the cirrus structures around, and in front of, our targets, and then subtracting them from our maps by comparison to their dust emission.

To do this, we used the all-sky HI data of the HI4PI survey (HI4PI Collaboration et al. 2016; Kalberla et al. 2020), which provides 21 cm coverage of the entire Milky Way velocity range, at 0.7° resolution. We obtained the HI4PI cubes that provide coverage towards our target galaxies. We inspected the velocity profiles of each, and found that the Milky Way HI emission was contained within the $-75 < v < 50\,{\rm km\,s^{-1}}$ velocity range. The exception to this is M 31, where the side rotating towards

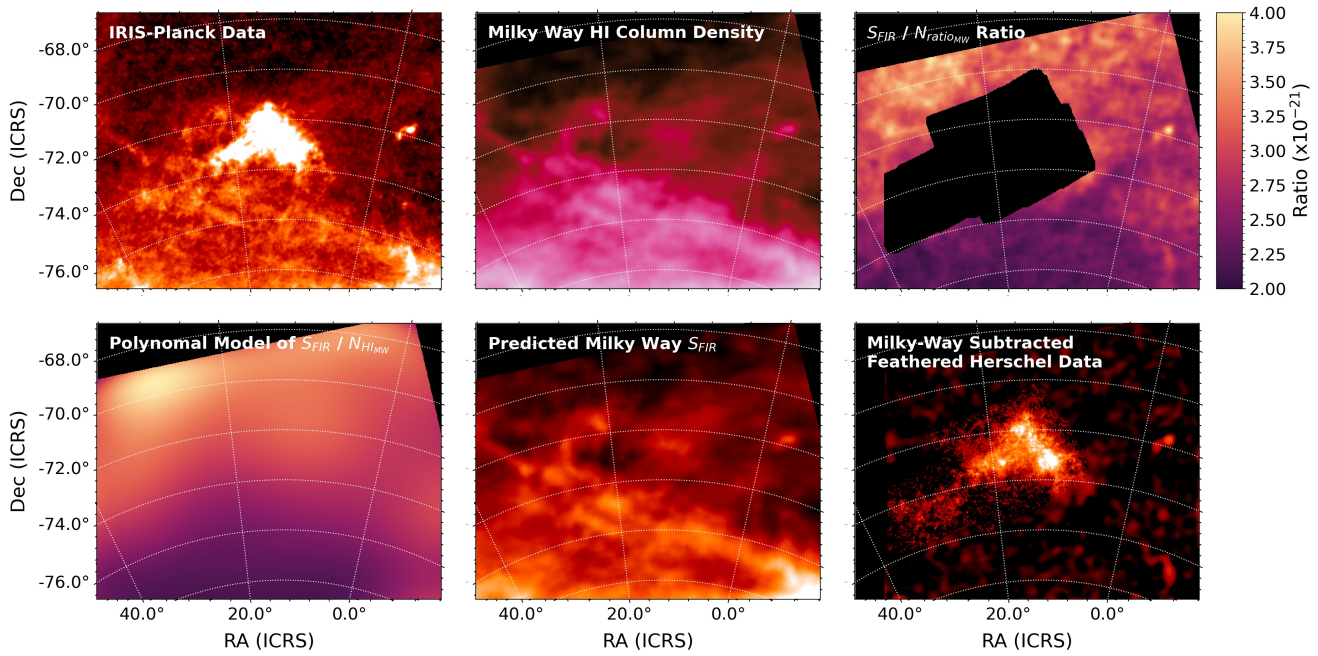

Figure 14. An example of our Milky Way foreground subtraction, for the case of the SMC at 500 μm. *Upper left:* Map of 500 μm surface brightness, S_{FIR} from our IRIS–*Planck* data. *Upper centre:* Column density of Galactic HI $N_{\mathrm{HI}_{MW}}$, from HI4PI 21 cm data in the Milky Way velocity range. *Upper right:* The ratio of FIR surface brightness to HI column density, $S_{FIR}/N_{\mathrm{HI}_{MW}}$. Footprint of *Herschel* observations is masked so that only the Milky Way values are traced. *Lower left:* Polynomial model of $S_{FIR}/N_{\mathrm{HI}_{MW}}$, allowing us to estimate the value of the ratio in the masked SMC region. *Lower centre:* Predicted foreground FIR surface brightness, via multiplying polynomial $S_{FIR}/N_{\mathrm{HI}_{MW}}$ model with the Galactic HI column density map. *Lower right:* Feathered *Herschel* data with predicted Galactic foreground FIR emission subtracted. The region with high-resolution *Herschel* coverage is visible in the centre, whilst regions with only low-resolution data available can be seen surrounding it.

us has ∼0 net velocity, making its HI emission confused with that of the Milky Way; we therefore handle foreground subtraction for M 31 differently, as described in Section 6.1. An example of the Milky Way HI emission, for the case of the SMC, is shown in the upper-left panel of Figure 14.

For the other galaxies, we compared the foreground HI column density to our predicted maps of the large-scale emission in each *Herschel* band, as inferred from our IRIS–*Planck* SED fitting in Section 4.3. These maps cover a much larger area than the footprint of the *Herschel* observations, spanning an area with a width of 3 times the *D25*. In each *Herschel* band, we first smoothed the map of predicted dust emission (with its resolution of 4.8′) to the 0.7° resolution of the HI data. We then produced a map of the ratio of the band's surface brightness to the HI column density, $S_{FIR}/N_{\mathrm{HI}_{MW}}$, for the region outside of the *Herschel* observation footprint; an example of this for the case of $S_{500\mu m}/N_{\mathrm{HI}_{MW}}$ around the SMC, is shown in the upper-right panel of Figure 14.

The maps of $S_{FIR}/N_{\mathrm{HI}_{MW}}$ show considerable variation, up to 75% (in line with Bianchi et al. 2017) – no doubt driven by changes in the Galactic dust-to-gas ratio, the dust temperature, etc. We could potentially use these ratio maps to simply calculate a constant ratio of FIR to HI column density for each band and field, and thereby impute the cirrus FIR surface brightness in front of our targets using the Galactic HI column for that area; however the large variation and structure in $S_{FIR}/N_{\mathrm{HI}_{MW}}$ means this would suffer large errors. Instead, we fit a 7$^{\mathrm{th}}$-order 2-dimensional polynomial to the $S_{FIR}/N_{\mathrm{HI}_{MW}}$ ratio map, to impute the variation of the ratio over the region observed by *Herschel*, which we mask out from the polynomial fitting; this is illustrated in the lower-left panel of Figure 14. We found that a 7$^{\mathrm{th}}$-order polynomial has sufficient freedom to capture the bulk of the structure in the $S_{FIR}/N_{\mathrm{HI}_{MW}}$ map without producing spurious features. We then use this polynomial model of the $S_{FIR}/N_{\mathrm{HI}_{MW}}$ ratio for a given band, along with the map of the Milky Way HI column over the field, to calculate the surface brightness of the foreground cirrus in that band, and subtract it from the feathered *Herschel* map; an example of this is shown in the lower-central and lower-right panels of Figure 14.

To assess the uncertainty in a subtracted map, we found the median value of the subtracted map outside

the *Herschel* footprint (ie, over the area where the polynomial was fitted). If the subtraction performed perfectly, this residual map should have a value of zero everywhere. For M 31, because of the less precise foreground subtraction we had to employ, the residuals are larger, with root-mean-square (RMS) residuals varying between bands from 8.3–9.3% (being larger in the PACS bands). For the other galaxies, the RMS residuals across all five bands average 6.8% for the LMC, 4.1% for the SMC, and 1.9% for M 33 – comparable for the photometric calibration uncertainties. The RMS for each map is always larger than the average residual for that map.

Because of the 0.7° resolution of HI4PI, our method is unable to subtract foreground structures smaller than 0.7°. The *total* amount of flux associated with any such smaller structures will be subtracted, however the result will be a compact structure in the final map, surrounded by a negative bowl. An example of this can be seen in the western edge of the final panel of Figure 14. Fortunately, the cirrus is sufficiently close to use that the vast majority of the foreground features in our maps are large enough to be subtracted well. Inspection of the HI data shows that there are no bright compact features in front of, or in the immediate proximity of, our target galaxies. Moreover, the worst negative bowl present in any of the foreground-subtracted maps is the one associated with the bright compact feature several degrees east of the SMC, visible in Figure 14, which at its worst point displays an 18% residual; the next worst is 13%, and all the rest are $< 10\%$. The RMS error on the HI4PI column densities is 6% outside of the Galactic plane, increasing to 10–15% in places (HI4PI Collaboration et al. 2016). As such, with only a few small exceptions, the magnitiude of the negative bowls is comparable to, or less than, the inherent noise in the HI maps.

This foreground subtraction will have also removed the emission of the Cosmic Infrared Background (CIB) from our maps. Technically the CIB will have manifested as a constant offset, in addition to the S_{FIR} emission per $N_{\mathrm{HI}_{MW}}$. In practice, our $S_{FIR}/N_{\mathrm{HI}_{MW}}$ will have captured this contribution as part of the S_{FIR} term. The fact that the residuals in the background of the subtracted maps are so close to zero indicates that the CIB was effectively removed.

6.1. *Foreground Subtraction for M 31*

Because the velocity of the HI emission from the Milky Way and M 31 overlap, we could not use the method described above for foreground subtraction. Fortunately, M 31 sits behind a region of Galactic cirrus that exhibits a fairly smooth gradient in surface brightness, so the more detailed subtraction method is not vital for success. Instead, we first masked the *Herschel* observation footprint from the predicted map of emission in each *Herschel* band derived from the IRIS-*Planck* SED-fitting, the same as above. However we then performed the polynomial modeling directly on the FIR maps, providing an estimate of the foreground emission in front of M 31 in each band, which we then subtracted from the feathered *Herschel* data. We calculated the uncertainty on this subtraction in the same manner as above.

6.2. *Foreground Subtraction for the Unfeathered Maps*

In order to compare our new, feathered *Herschel* maps to the unfeathered original maps, we also needed to foreground subtract the original maps. In the the case of the LMC, and of the PACS data for M 33, there is very little sky surrounding the target galaxies. Without sufficient area of empty sky, it was not possible to perform the same sort of subtraction as we did in Section 6, using Galactic HI data.

Nor can we use the IRIS–*Planck* FIR data, as the large-scale emission that fills the sky in that data is simply not present in the unfeathered *Herschel* data. For instance, the sky to the far south east of the SMC in the *Herschel*-PACS data has the same surface brightness as the sky to the far north west, due to the flattening imposed by the *Herschel* reduction – but in the IRIS–*Planck* data, the surface brightness of these areas of sky differs by more than a factor of 2, due to variation diffuse Galactic emission. Subtracting this from the unfeathered *Herschel* maps would lead to major negative features in large portions of the map. However, the unfeathered *Herschel* maps do preserve *some* large scale foreground emission (especially for M 31 and the SMC), so we cannot simply treat the background as being flat.

The tried-and-true method of placing an annulus around the target source to measure the sky level is also not ideal here. We are interested in diffuse emission at galaxy outskirts, and any emission in or beyond a sky annulus would therefore be subtracted. And because our targets are large (and sometimes irregular) galaxies, in relatively tight maps, any sky annulus we place would run into this problem. Instead, we mask out the central regions of each unfeathered *Herschel* map, leaving only a border along the inside edge of the *Herschel* footprint[16], with width equal to 5% the R_{25} of the target galaxy. We then fit a 1st-order 2-dimensional polynomial (ie, a tilted plane) to this border data, to model the foreground emission, which we then subtract. This

[16] Specifically within the same footprint region defined at the end of Section 5.2, to exclude bright artefacts in the turnaround regions, etc.

Table 6. The integrated fluxes we measure for our sample galaxies in each band with our feathered maps, compared to those measured from the unfeathered maps. Note that the new fluxes incorporate the cross-calibration corrections from Table 5, whereas the unfeathered fluxes do not; this correction imposes to a 14–31% reduction in flux in the PACS bands, and a 0–16% reduction in the SPIRE bands.

Band	LMC (kJy)		SMC (kJy)		M 31 (kJy)		M 33 (kJy)	
(μm)	Feathered	Unfeathered	Feathered	Unfeathered	Feathered	Unfeathered	Feathered	Unfeathered
100	188 ± 19	196	16.7 ± 1.4	15.7	3.37 ± 40	3.31	1.13 ± 0.08	1.62
160	232 ± 22	196	20.1 ± 1.6	18.5	7.29 ± 0.83	7.00	1.77 ± 0.13	1.80
250	140 ± 12	142	12.9 ± 0.8	10.7	6.18 ± 0.65	5.46	1.29 ± 0.08	1.31
350	71.2 ± 6	71	7.11 ± 0.48	6.23	3.43 ± 0.35	3.00	0.71 ± 0.04	0.71
500	28.5 ± 3	29	3.24 ± 0.23	3.01	1.40 ± 0.14	1.28	0.30 ± 0.02	0.31

maximizes our ability to recover diffuse emission in the outskirts of the target galaxies.

6.3. *The Change in Total Flux*

Our new maps change the photometry measured for the galaxies in our sample. The global fluxes we record are given in Table 6, which also lists the fluxes measured from the unfeathered maps, for comparison.

For the LMC and SMC, these fluxes were measured by taking the total flux within the *Herschel* footprint of the foreground-subtracted feathered maps. We opted not to place a tighter aperture around the sources in order to not exclude any faint emission at large radii. Our foreground selection should be equally good across the entire *Herschel* footprint, so no systematic bias should be suffered from measuring the fluxes this way. Nonetheless, this will drive up the 'aperture noise' associated with the resulting fluxes.

For M 31 and M 33, the fluxes were measured within elliptical apertures, centred on the coordinates given in Table 1. For M 31, the aperture had $a = 126'$, $b = 46'$, $\theta = 38°$; for M 33, the aperture had $a = 46'$, $b = 28'$, and $\theta = 20°$ (where a, b, and θ are the semi-major axis, semi-minor axis, and position angle, respectively). These apertures were designed to comfortably contain all the potential extended emission, especially as visible in the SPIRE maps[17].

For the uncertainty on the recorded fluxes, we use the fractional uncertainty on the foreground subtraction, the cross-calibration correction, added in quadrature to the instrumental calibration uncertainty (7% for PACS bands, 5.5% for SPIRE).

[17] For the feathered maps, the photometric apertures extended beyond the feathered footprint in some bands (due to small PACS observing footprint, and the masked feathering edge effect areas). For these pixels, we use the low-resolution IRIS–*Planck* predictions of emission in the *Herschel* band.

For the SPIRE bands, the feathered fluxes are in close agreement with the unfeathered fluxes for the LMC and M 33. For the SMC, the feathered fluxes are 9–21% brighter; and for M 31, the feathered SPIRE fluxes are 9–14% brighter (differences decreasing towards longer wavelengths, as would be expected).

In the PACS bands, there is relatively little change in the measured fluxes. Only the 160 μm flux for the LMC, and the 100 μm flux for M 33, differ by larger than the uncertainty. In these cases, the feathered fluxes are actually less than the unfeathered measurements, by 18% and 43% respectively. We suspect this change is due to the feathered maps allowing better foreground subtraction; a filament of foreground cirrus passes east-west over M 33, and a larger cloud of cirrus extends over the galaxy's southwest quadrant. For the other galaxies, where the change in PACS flux *is* within the uncertainty, the feathered fluxes are brighter in 5 out of 6 of them, for an average change of +7%.

Note, however, that all of these comparisons are occurring after the application of the cross-calibration correction factors during the feathering process, which have therefore *not* been applied to the unfeathered maps. These correction factors range impose14–31% reduction in flux in the PACS bands, and a 0–16% reduction in the SPIRE bands (Table 5). Therefore, had these corrections not been applied, the feathered data would have been much brighter, in almost all instances.

In short, whilst the surface brightness measured at any given point in our target galaxies will often have changed dramatically between the unfeathered and feathered maps (ie, Figures 13, F4, and F5, the *total* measured flux tends to be quite similar – thanks to the reduction in flux due to the cross-calibration correction being comparable to the increase in flux from restored emission.

6.4. *Total Flux Compared to the Literature*

For the LMC, our PACS fluxes agree within the uncertainties with those previously reported from IRAS (Rice

et al. 1988), COBE (Israel et al. 2010), and *Spitzer* (Dale et al. 2009), but differ from those reported by Meixner et al. (2013) with the HERITAGE data. At 100 μm, we find a flux 18% larger than Meixner et al. (2013), and at 160 μm we find a flux 16% less than Meixner et al. (2013). For SPIRE, however, our LMC fluxes agree closely with those of Meixner et al. (2013) (which agrees well with COBE at 240–250 μm; Israel et al. 2010).

For the SMC, our PACS fluxes once again agree with IRAS (Rice et al. 1988), COBE (Israel et al. 2010), and *Spitzer* (Leroy et al. 2007; Dale et al. 2009), and once again differ from those reported by Meixner et al. (2013) with the HERITAGE data; ours being 14% brighter at 100 μm, and 49% brighter at 160 μm. Our SPIRE fluxes for the the SMC are also larger than those reported by Meixner et al. (2013); by 29%, 21%, and 12%, at 250, 350, and 500 μm respectively. Whereas our 250 μm flux agrees well with the published 240 μm COBE flux (Israel et al. 2010). We note that the Meixner et al. (2013) SMC SPIRE fluxes are, however, a close match to our unfeathered fluxes.

For the Magellanic Clouds, it appears that matching the calibration of our feathered data to well-calibrated all-sky surveys has fixed the discrepancy between photometry previously measured from *Herschel*, and photometry measured from other facilities (including all-sky surveys).

For M 31, our fluxes agree with those reported by Viaene et al. (2014) to within 10% at 350 μm, and to with 5% in all other bands; this is within our uncertainties in all cases. The agreement in the PACS bands indicates that the additional flux recovered in our maps closely matches the relative decrease in flux imposed by our cross-calibration, as suggested in Section 6.3.

For M 33, our SPIRE fluxes agree to within 6.3% with those reported by (Hermelo et al. 2016); this is within the mutual uncertainties in each case. For the PACS bands, however, our fluxes are 12% and 18% fainter at 100 and 160 μm, respectively. This indicates that relatively less flux was restored by our feathering, as comapred to the cross-calibration decrease; this is not surprising, given that M 33 is by far the most compact of our sources, and therefore will have suffered the least emission lost by filtering.

7. INITIAL RESULTS

Our new *Herschel* maps allow us to probe the FIR–submm emission from our target galaxies with a combination of accuracy and resolution that has not previously been possible. Here, we examine the surface brightness properties of the galaxies in the new maps, in comparison with the old maps, to see what differences there are – and especially to inspect those properties out to low surface brightnesses that the old data could not probe.

Even with our feathered maps, the diffuse dust emission in the galaxies' peripheries is still faint, and typically has low SNR for individual pixels. Therefore, to get beneath the noise level, and study dust at densities otherwise too low to detect, we binned together the emission from pixels that have similar hydrogen column density.

7.1. *Hydrogen Surface Density Maps*

We constructed $\Sigma_{\rm H}$ maps for each of our galaxies, using 21 cm maps to trace the HI, and CO maps to trace the H_2. For the LMC, we used the HI data of Kim et al. (2003) and CO data of Wong et al. (2011). For the SMC, we used the HI data of Stanimirovic et al. (1999), and the CO data of Mizuno et al. (2001). For M 31 we used the HI data of Braun et al. (2009), and the CO data of Nieten et al. (2006). For M 33, we used the HI data of Koch et al. (2018), CO map of Gratier et al. (2010); Druard et al. (2014). We calculated molecular gas surface density, $\Sigma_{\rm H2}$ using the standard relation:

$$\Sigma_{\rm H2} = \alpha_{\rm CO} I_{\rm CO(1-0)} \tag{8}$$

where $I_{\rm CO(1-0)}$ is the velocity-integrated main-beam brightness temperature of the CO(1-0) line (in $\rm K\,km\,s^{-1}$), $\alpha_{\rm CO}$ is the CO-to-H_2 conversion factor (in $\rm K^{-1}\,km^{-1}\,s\,M_\odot\,pc^{-2}$).

All of the CO observations we used were of the CO(1-0) line, except for that of M 33, which was of CO(2-1). For that data, we inferred $I_{\rm CO(1-0)}$ by applying a line ratio, $r_{2:1} = I_{\rm CO(2-1)}/I_{\rm CO(1-0)}$. In spiral galaxies, $r_{2:1}$ varies radially (Casoli et al. 1991; Leroy et al. 2009). We therefore use the procedure laid out in Section 3.4 of Clark et al. (2019), where $r_{2:1}$ is determined as a function of galactocentric radius as a fraction of the R_{25}, using measurements made by Leroy et al. (2009) – varying from 1.0 at the galaxy centre, to 0.55 at the R_{25}. We thereby computed $I_{\rm CO(1-0)}$ for each pixel in the M 33 $I_{\rm CO(2-1)}$ map according to its position.

To compute $\Sigma_{\rm H2}$ for M 31 and M 33, we used the standard Milky Way value of $\alpha_{\rm CO} = 3.2\,\rm K^{-1}\,km^{-1}\,s\,pc^{-2}$, which tends to be applicable to high-metallicity spiral

galaxies in general[18] (see Saintonge et al. 2011, Bolatto et al. 2013, and references therein). For the LMC and SMC, we used α_{CO} values of 6.4 and 21 $K^{-1}\,km^{-1}\,s\,pc^{-2}$ respectively (Bolatto et al. 2013). From Σ_{H2} we then calculated H_2 surface density

All of the HI and H_2 maps we use have resolutions in the region of 1′, being quite similar both to each other, and to our *Herschel* maps – varying from 0.2′ for the M 33 H_2 data, to 2.6′ for the SMC H_2 data. All of the HI maps incorporate both interferometric and single-dish observations, ensuring that we won't miss any diffuse atomic structure due to lack of short-spacing data.

To make our maps of Σ_H, we convolved the HI and H_2 maps for each galaxy to whichever resolution was worst out of the two maps, re-gridded them to the same projection, then added them together to produce the combined map of hydrogen surface density.

7.2. *Tracing Dust Surface Brightness Down to the Lowest Hydrogen Surface Densities*

Dust and gas are typically well mixed in the ISM, across a wide range of surface densities (Hildebrand 1983; Eales et al. 2012; Williams et al. 2018). Given that ISM properties are strongly driven by density (see Section 1 and references therein), dust found at different locations, but sharing a given Σ_H, should have similar properties.

We convolved our Σ_H and foreground-subtracted *Herschel* maps to the limiting resolution for each galaxy. For the LMC and SMC, the limiting data were the Σ_H maps, with resolutions of 1′ for the LMC and 2.6′ for the SMC. For M 31 and M 33, the limiting data was *Herschel* at 500 μm, with its 36″ resolution. We then defined bins of Σ_H, each 0.025 dex wide. For each bin, all of the Σ_H pixels within that density range were identified, and all of the corresponding *Herschel* pixels had their mean surface brightnesses calculated, within each band.

For comparison, we also performed this process for our unfeathered maps – plus, for the LMC and SMC, the original HERITAGE maps (Meixner et al. 2013), to allow us to evaluate the difference from using the newer reduction calibrations, pipelines, and feathering. For the purposes of an 'apples-to-apples' comparison, we applied

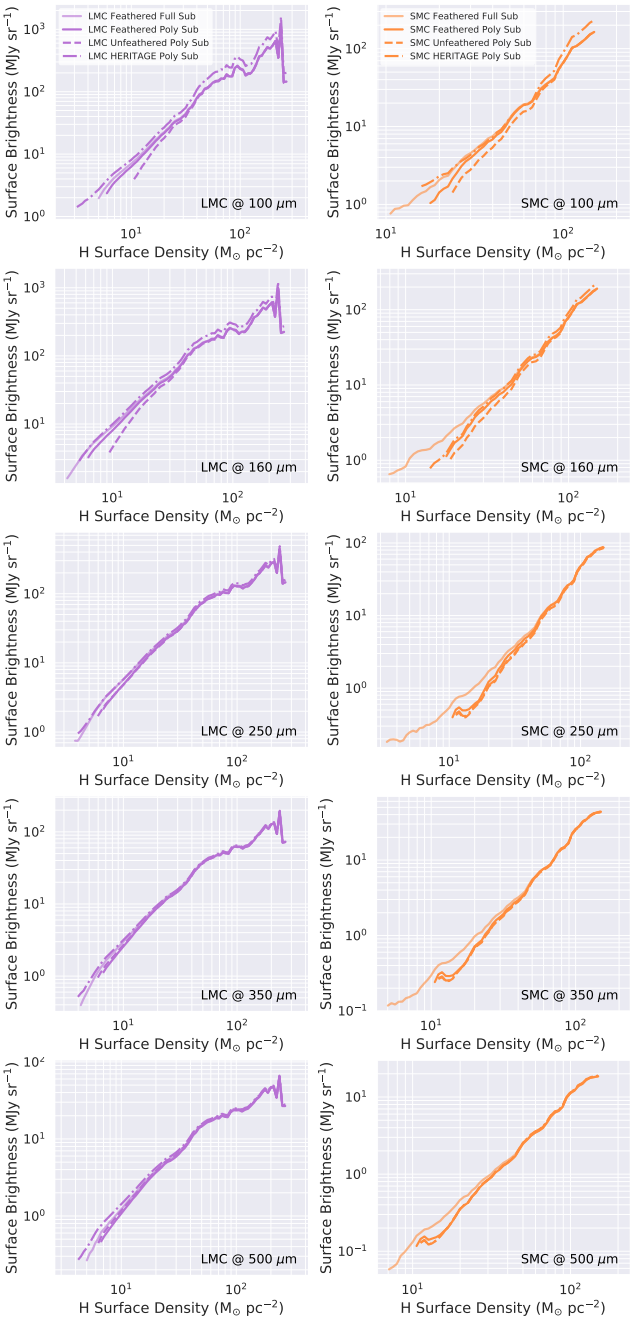

Figure 15. Plots of surface brightness (averaged in bins of H density) against H surface density, for the LMC and SMC. In each plot, we show this relation for of full-foreground-subtracted feathered data, and for the polynomial-foreground-subtracted feathered, unfeathered, and HERITAGE data (to allow these three to be compared fairly, with the same foreground subtraction). We plot each relation down until the bin in which < 67% of the binned pixels have positive values; below this point, the line can 'bounce' around due to the noise.

[18] For M 33, some authors have α_{CO} higher than the Milky Way value (Bigiel et al. 2010; Druard et al. 2014). However, Gratier et al. (2010) find a value within 10% of that of the Milky Way. Plus Rosolowsky et al. (2003) (who do not measure an an absolute α_{CO}, only its' apparent relative variation) find that α_{CO} in the highest-metallicity regions of M 33, at $> 1.4\,Z_\odot$, is not systematically different from that in regions at $< 0.5\,Z_\odot$; given that α_{CO} at sigh high metallicities should be comparable to that of the Milky Way, it suggests this is not significantly different elsewhere in M 33.

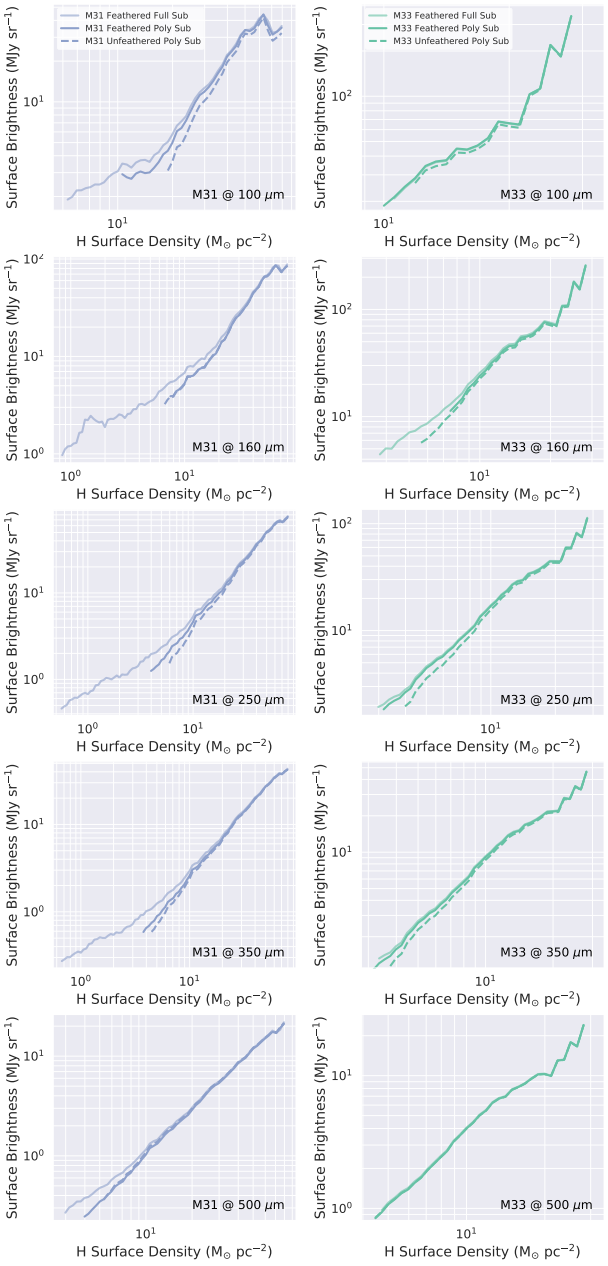

Figure 16. Same as for Figure 15, but for M 31 and M 33 (so therefore without HERITAGE data plotted).

the polynomial-based background subtraction used for our unfeathered maps in Section 6.2 to the feathered and HERITAGE maps also. By comparing the results for the feathered maps when using the polynomal versus full foreground subtractions, we are able to determine how much additional low-surface-brightness emission is retrievable thanks to the improved foreground subtraction that is only possible with the feathered maps.

In Figures 15 (LMC and SMC) and 16 (M 31, and M 33), we compare the relationship between HI surface density and surface brightness for the polynomial-subtracted feathered, unfeathered, and HERITAGE maps, along with the feathered maps that had undergone our full HI-based foreground subtraction. Note that it is hard to draw firm conclusions from differences between the HERITAGE reductions and our reductions, as HERITAGE used a much older version of HIPE (v7 for HERITAGE, compared to v12+ for our reductions), and there were some differences in calibration (eg, non-linearity corrections were added for PACS, bolometer relative gains were altered for SPIRE, and beam sizes were updated for both). However, it is still valuable to see what H column each dataset is able to trace down to.

We only perform the binning for this comparison for pixels that are covered in all of the datasets (*Herschel*, HI, and CO) for each given galaxy. As such, our ability to retrieve binned *Herschel* emission here is a lower limit (ie, we would be able to probe even deeper if able to bin over the entire map area).

For the SMC (right of Figure 15) and M31 (left of Figure 16), we see particularly impressive gains with the new data. For the SMC, averaged across all five bands, the full-foreground-subtracted feathered data can detect dust emission to a H surface density a factor of 2 lower than the HERITAGE data, and a factor of 2.3 lower than with the unfeathered data. Similarly, for M 31, the feathered data allows us to detect dust emission to a mean factor of 5 lower in density (reaching a factor of 10 lower at 250 μm).

For the LMC (left of Figure 15), we see that there are rather minimal difference between the surface brightnesses, down to the lowest H density for which we are able to trace dust emission. In each band, both the polynomial-foreground-subtracted HERITAGE data or full-foreground-subtracted feathered data were each able to trace down to similar H densities. This may be due to the fact that the LMC has a relatively abrupt 'edge' to its ISM disc, and the fact that the *Herschel* map is relatively tight compared to this disc, with little sky around the edges. There are also smaller gains for M 33 in most bands, only probing to H densities 25% lower, on average, than the unfeathered data (although the PACS bands specifically do go to 50% lower densities). As with the LMC, the M 33 maps are comparitively small, with M 33 lying close to the edges.

We also note that the full-foreground-subtracted feathered data also traces dust emission out much lower H densities than the polynomial-foreground-subtracted feathered data. This well demonstrates the value of the fact that feathering the *Herschel* maps with all-sky survey maps allows the combined maps to properly trace the large-scale foreground cirrus – therefore mak-

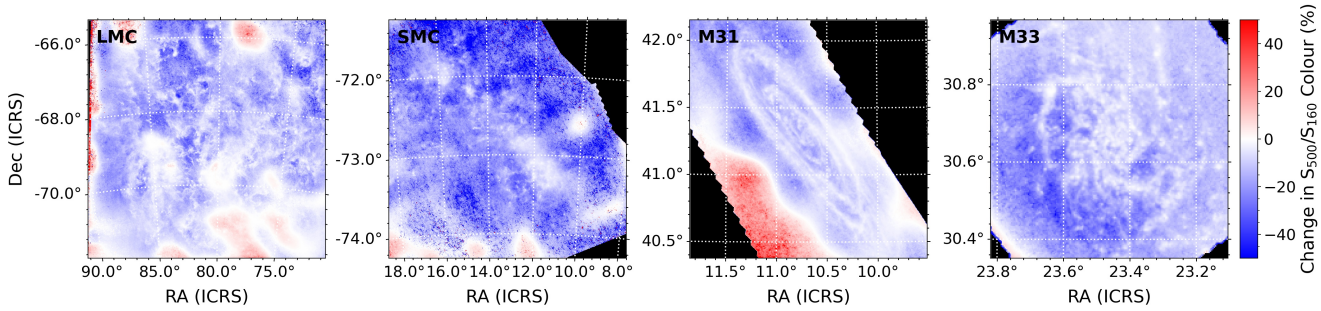

Figure 17. Illustration of the change in the S_{500}/S_{160} colour for each of our sample galaxies. Due to the greater amount of emission restored in the shorter-wavelength data, the bulk of the area is now bluer, especially in areas of more diffuse emission.

ing it possible to accurately subtract it by comparison to Galactic HI data.

For all galaxies and bands, our new data allows us to trace dust emission down to ISM surface densities of $< 10\,M_\odot pc^{-2}$. It is interesting that the surface brightness versus H density relations do not drop off with decreasing H density; instead, the gradients remain the same, or even flatten out. This is suggestive of the dust abundance in these low-density environments – which will be the central focus of the second paper in this series (Clark et al., *in prep.*).

7.3. *Much Bluer Galaxy Peripheries*

The magnitude and distribution of the emission restored by the feathering process varies from band to band for each galaxy. In general, our feathering process restored more diffuse emission in the shorter-wavelength data than at longer wavelengths (where there was much less filtering). This restoration was most prominent at the galaxies' outskirts. As a result, the emission at the edges of these galaxies is much bluer than it was in the unfeathered data.

This is demonstrated in Figure 17, which shows how much the S_{500}/S_{160} colour (ie, the 500 μm to 160 μm surface brightness ratio) has changed, between the unfeathered maps and the feathered maps. Bright compact regions show no change in colour, as they suffered little filtering in the unfeathered data. However, diffuse regions, especially at the galaxies' peripheries, are much bluer in the feathered data, with S_{500}/S_{160} often falling by 20-30% over large areas. Note that this increase in blueness is *despite* the fact that the cross-calibration in Section 5.2 reduced brightnesses in the PACS bands by 25–35% (see Table 5).

Dust temperatures derived from SED fitting FIR–submm data are strongly driven by colour (Bendo et al. 2012). As such, the increased blueness of the new data can be expected to affect dust temperatures – and also affect dust masses, and β (due to the temperature–β degeneracy; Kelly et al. 2012; Galliano 2018). Exploring the dust SED properties will be a central focus of the second paper in this series (Clark et al., *in prep.*).

8. CONCLUSION

We have produced new *Herschel* maps for the Local Group galaxies M 31, M 33, the LMC, and the SMC. These are some of the most heavily-studied galaxies in the sky, and represent key local laboratories on which we base our understanding of many systems across cosmic time. Because of these galaxies' large angular sizes, the standard *Herschel* reductions are vulnerable to severe loss of diffuse emission on large angular scales, filtered out as part of the reduction pipeline. To remedy this, we combined the latest *Herschel* reductions, in Fourier space, with *Planck*, IRAS, and DIRBE data, to restore any lost emission. From producing and analyzing these new 'feathered' maps, our key findings are:

- Large amounts of diffuse dust emission was indeed missing from standard reductions, especially around the outskirts of the sample galaxies. Our new maps restore this missing dust. This restoration was particularly pronounced in the shorter-wavelength bands.

- By restoring this diffuse emission, we find significantly different global fluxes measured for our galaxies, as compared to those measured from the unfeathered maps, by over 15% in many cases.

- However, from cross-calibrating the power spectra of *Herschel* data with that of absolutely-calibrated all-sky survey data, we find that *Herschel*-PACS maps seem to over-estimate the brightness of large-scale emission by 20–30% (in line with similar findings by Abreu-Vicente et al. 2017). Our new data corrects for this.

- Previous *Herschel* photometry for the Magellanic Clouds, from HERITAGE (Meixner et al. 2013),

conflicted with published photometry from lower-resolution facilities such as COBE, IRAS, and *Spitzer*, with differences of up to 50%. Photometry from our new maps resolve this disagreement, with our new *Herschel* fluxes generally agreeing well with values from the other facilities

- When binning together pixels in the feathered maps according to hydrogen surface density, we find that we can detect dust emission down to ISM densities of $\Sigma_{\rm H} < 1\,{\rm M}_{\odot}{\rm pc}^{-2}$ in some cases, and to at least $\Sigma_{\rm H} < 10\,{\rm M}_{\odot}{\rm pc}^{-2}$ for all galaxies and bands. The quality of our feathered maps allows us to detect such emission to lower densities than with unfeathered data (down to densities a factor of 2.2 lower, on average), or with previous HERITAGE data for the Magellanic Clouds (a factor of (down to densities 50% lower, on average), with particularly large improvements for the SMC and M 31. We find no indication that the dust emission drops off more sharply at these lower ISM densities.

- Because the restoration of large-scale emission was greater in the shorter-wavelength bans, the galaxies' far-infrared colours are now much bluer over large areas, especially in their peripheries. Specifically, we find the S_{500}/S_{160} colour becoming $>20\%$ bluer almost everywhere there is not bright compact emission present.

With this new data in hand, the immediate future work will be to study the variation of dust abundance relative to gas, and how this relates to grain-growth, deplations, etc. This will be the focus of paper II in this series (Clark et al., *in prep.*). Additional investigations that will be enabled by this data include a study of how FIR dust emissivity compares to UV–optical extinction, as revealed by *Hubble* Space Telescope imaging programs such as Scylla (Murray et al. 2019b) and the Panchromatic *Hubble* Andromeda Treasury (Dalcanton et al. 2012), to trace variations in dust properties and composition; and constraining the dust mass opacity coefficient, κ_d, by comparison to gas and metallicity data, at spatial resolutions high enough to overcome significant temperature mixing (Galliano et al. 2011; Priestley & Whitworth 2020).

The code used to carry out the feathering process presented here is available online at: https://doi.org/10.5281/zenodo.4776266. The maps themselves will be released alongside paper II of this series.

ACKNOWLEDGMENTS

The authors thank Oliver Lomax for sharing his adaptation of the old COBE-DIRBE reprojection algorithm from `FORTRAN` into `Python`, and thank Yumi Choi and Thomas Williams for valuable discussions. The authors also thank the anonymous referee for their thoughtful and helpful comments.

CJRC and JR-D acknowledge financial support from the National Aeronautics and Space Administration (NASA) Astrophysics Data Analysis Program (ADAP) grant 80NSSC18K0944.

This work used the Extreme Science and Engineering Discovery Environment (XSEDE; Towns et al. 2014), which is supported by National Science Foundation grant number ACI-1548562. Specifically, this work used XSEDE Bridges-Large at the Pittsburgh Supercomputing Center, through allocation AST190060.

This research made use of `Astropy`[19], a community-developed core `Python` package for Astronomy (Astropy Collaboration et al. 2013, 2018). This research made use of `reproject`[20], an `Astropy`-affiliated `Python` package for image reprojection. This research made use of `photutils`[21], an `Astropy`-affiliated `Python` package for detection and photometry of astronomical sources (Bradley et al. 2020). This research made use of `NumPy`[22] (van der Walt et al. 2011; Harris et al. 2020), `SciPy`[23] (Jones et al. 2001), and `Matplotlib`[24] (Hunter 2007). This research made use of the `pandas`[25] data structures package for `Python` (McKinney 2010). This research made use of `turbustat`[26], a `Python` package for turbulence statistics (Koch et al. 2019). This research made use of `iPython`, an enhanced interactive `Python` (Pérez & Granger 2007).

This research made use of sequential colour-vision-deficiency-friendly colourmaps from `cmocean`[27] (Thryng et al. 2016) and `CMasher`[28] (van der Velden 2020)

This research made use of `UNIMAP` (Piazzo et al. 2012, 2015; Piazzo et al. 2015; Piazzo et al. 2016a,b; Piazzo 2017), a development of the `ROMAMAP` pipeline (Traficante et al. 2011).

This research made use of `TOPCAT`[29] (Taylor 2005), an interactive graphical viewer and editor for tabular data.

This research made use of `Montage`[30], which is funded by the National Science Foundation under Grant Number ACI-1440620, and was previously funded by the NASA's

[19] https://www.astropy.org/
[20] https://reproject.readthedocs.io
[21] https://photutils.readthedocs.io
[22] https://numpy.org/
[23] https://scipy.org/
[24] https://matplotlib.org/
[25] https://pandas.pydata.org/
[26] https://turbustat.readthedocs.io/en/latest/
[27] https://matplotlib.org/cmocean/
[28] https://cmasher.readthedocs.io
[29] http://www.star.bris.ac.uk/~mbt/topcat/
[30] https://montage.ipac.caltech.edu/

Earth Science Technology Office, Computation Technologies Project, under Cooperative Agreement Number NCC5-626 between NASA and the California Institute of Technology.

This research made use of the SIMBAD database[31]; Wenger et al. 2000) and the VizieR catalogue access tool[32] (Ochsenbein et al. 2000), both operated at CDS, Strasbourg, France. This research has made use of the NASA/IPAC Extragalactic Database[33] (NED), operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with NASA. This research made use of the HyperLEDA database[34] (Makarov et al. 2014).

This research makes use of data from *Planck*, a project of the European Space Agency, which received support from: ESA; CNES and CNRS/INSU- IN2P3-INP (France); ASI, CNR, and INAF (Italy); NASA and DoE (USA); STFC and UKSA (UK); CSIC, MINECO, JA, and RES (Spain); Tekes, AoF, and CSC (Finland); DLR and MPG (Germany); CSA (Canada); DTU Space (Denmark); SER/SSO (Switzerland); RCN (Norway); SFI (Ireland); FCT/MCTES (Portugal); ERC and PRACE (EU).

Herschel is an ESA space observatory with science instruments provided by European-led Principal Investigator consortia and with important participation from NASA. The *Herschel* spacecraft was designed, built, tested, and launched under a contract to ESA managed by the *Herschel*/*Planck* Project team by an industrial consortium under the overall responsibility of the prime contractor Thales Alenia Space (Cannes), and including Astrium (Friedrichshafen) responsible for the payload module and for system testing at spacecraft level, Thales Alenia Space (Turin) responsible for the service module, and Astrium (Toulouse) responsible for the telescope, with in excess of a hundred subcontractors.

REFERENCES

Abreu-Vicente, J., Stutz, A., Henning, T., et al. 2017, A&A, 604, A65, doi: 10.1051/0004-6361/201628891

Aniano, G., Draine, B. T., Gordon, K. D., & Sandstrom, K. 2011, PASP, 123, 1218, doi: 10.1086/662219

Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33, doi: 10.1051/0004-6361/201322068

Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, AJ, 156, 123, doi: 10.3847/1538-3881/aabc4f

Bajaja, E., & van Albada, G. D. 1979, A&A, 75, 251

Balog, Z., Müller, T., Nielbock, M., et al. 2014, Experimental Astronomy, 37, 129, doi: 10.1007/s10686-013-9352-3

Barlow, M. J., Krause, O., Swinyard, B. M., et al. 2010, A&A, 518, L138, doi: 10.1051/0004-6361/201014585

Barreiro, R. B., Hobson, M. P., Banday, A. J., et al. 2004, MNRAS, 351, 515, doi: 10.1111/j.1365-2966.2004.07802.x

Beichman, C. A., Neugebauer, G., Habing, H. J., Clegg, P. E., & Chester, T. J. 1988, in Infrared astronomical satellite (IRAS) catalogs and atlases. Volume 1: Explanatory supplemen, Vol. 1

Beirão, P., Brandl, B. R., Devost, D., et al. 2006, ApJ, 643, L1, doi: 10.1086/505027

Bendo, G. J., Boselli, A., Dariush, A., et al. 2012, MNRAS, 419, 1833, doi: 10.1111/j.1365-2966.2011.19735.x

Bertincourt, B., Lagache, G., Martin, P. G., et al. 2016, A&A, 588, A107, doi: 10.1051/0004-6361/201527313

Bianchi, S., Giovanardi, C., Smith, M. W. L., et al. 2017, A&A, 597, A130, doi: 10.1051/0004-6361/201629013

Bigiel, F., Bolatto, A. D., Leroy, A. K., et al. 2010, ApJ, 725, 1159, doi: 10.1088/0004-637X/725/1/1159

Bocchio, M., Bianchi, S., & Abergel, A. 2016, A&A, 591, A117, doi: 10.1051/0004-6361/201628665

Bocchio, M., Jones, A. P., & Slavin, J. D. 2014, A&A, 570, A32, doi: 10.1051/0004-6361/201424368

Boggess, N. W., Mather, J. C., Weiss, R., et al. 1992, ApJ, 397, 420, doi: 10.1086/171797

Bolatto, A. D., Wolfire, M., & Leroy, A. K. 2013, Annual Review of Astronomy and Astrophysics, 51, 207, doi: 10.1146/annurev-astro-082812-140944

Boogert, A. C. A., Gerakines, P. A., & Whittet, D. C. B. 2015, ARA&A, 53, 541, doi: 10.1146/annurev-astro-082214-122348

Bot, C., Ysard, N., Paradis, D., et al. 2010, A&A, 523, A20, doi: 10.1051/0004-6361/201014986

Boulanger, F., & Perault, M. 1988, ApJ, 330, 964, doi: 10.1086/166526

Boulanger, F., Abergel, A., Bernard, J. P., et al. 1998, in Astronomical Society of the Pacific Conference Series, Vol. 132, Star Formation with the Infrared Space Observatory, ed. J. Yun & L. Liseau, 15

Bradley, L., Sipőcz, B., Robitaille, T., et al. 2020, astropy/photutils: 1.0.1, 1.0.1, Zenodo, doi: 10.5281/zenodo.596036

Braun, R., Thilker, D. A., Walterbos, R. A. M., & Corbelli, E. 2009, ApJ, 695, 937, doi: 10.1088/0004-637X/695/2/937

Calzoletti, L. 2017, PACS Photometer Highly Processed Data Products with Unimap: Release Notes, Tech. Rep. HERSCHEL-HSC-DOC-2197, ESA

Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, ApJ, 345, 245, doi: 10.1086/167900

Casoli, F., Dupraz, C., Combes, F., & Kazes, I. 1991, A&A, 251, 1

Chapin, E. L., Berry, D. S., Gibb, A. G., et al. 2013, MNRAS, 430, 2545, doi: 10.1093/mnras/stt052

Chastenet, J., Bot, C., Gordon, K. D., et al. 2017, A&A, 601, A55, doi: 10.1051/0004-6361/201629133

Clark, C. J. R., De Vis, P., Baes, M., et al. 2019, MNRAS, 489, 5256, doi: 10.1093/mnras/stz2257

Csengeri, T., Weiss, A., Wyrowski, F., et al. 2016, A&A, 585, A104, doi: 10.1051/0004-6361/201526639

Dalcanton, J. J., Williams, B. F., Lang, D., et al. 2012, ApJS, 200, 18, doi: 10.1088/0067-0049/200/2/18

Dale, D. A., Cohen, S. A., Johnson, L. C., et al. 2009, ApJ, 703, 517, doi: 10.1088/0004-637X/703/1/517

Demyk, K., Meny, C., Lu, X. H., et al. 2017a, A&A, 600, A123, doi: 10.1051/0004-6361/201629711

Demyk, K., Meny, C., Leroux, H., et al. 2017b, A&A, 606, A50, doi: 10.1051/0004-6361/201730944

Desert, F.-X., Boulanger, F., & Puget, J. L. 1990, A&A, 237, 215

Druard, C., Braine, J., Schuster, K. F., et al. 2014, A&A, 567, A118, doi: 10.1051/0004-6361/201423682

Eales, S., Smith, M. W. L., Auld, R., et al. 2012, ApJ, 761, 168, doi: 10.1088/0004-637X/761/2/168

Eisenstein, D. J., Weinberg, D. H., Agol, E., et al. 2011, AJ, 142, 72, doi: 10.1088/0004-6256/142/3/72

[31] https://simbad.u-strasbg.fr/simbad/

[32] https://vizier.u-strasbg.fr/viz-bin/VizieR

[33] https://ned.ipac.caltech.edu/

[34] http://leda.univ-lyon1.fr/

Fixsen, D. J., Weiland, J. L., Brodd, S., et al. 1997, ApJ, 490, 482, doi: 10.1086/304906
Galliano, F. 2018, MNRAS, 476, 1445, doi: 10.1093/mnras/sty189
Galliano, F., Madden, S. C., Jones, A. P., et al. 2003, A&A, 407, 159, doi: 10.1051/0004-6361:20030814
Galliano, F., Hony, S., Bernard, J. P., et al. 2011, A&A, 536, A88, doi: 10.1051/0004-6361/201117952
Gardan, E., Braine, J., Schuster, K. F., Brouillet, N., & Sievers, A. 2007, A&A, 473, 91, doi: 10.1051/0004-6361:20077711
Gomez, H. L., Krause, O., Barlow, M. J., et al. 2012, ApJ, 760, 96, doi: 10.1088/0004-637X/760/1/96
Gordon, K. D., Clayton, G. C., Misselt, K. A., Landolt, A. U., & Wolff, M. J. 2003, ApJ, 594, 279, doi: 10.1086/376774
Gordon, K. D., Bailin, J., Engelbracht, C. W., et al. 2006, ApJL, 638, L87, doi: 10.1086/501046
Gordon, K. D., Roman-Duval, J., Bot, C., et al. 2014, ApJ, 797, 85, doi: 10.1088/0004-637X/797/2/85
Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, ApJ, 622, 759, doi: 10.1086/427976
Graciá-Carpio, J., Wetzstein, M., Roussel, H., & PACS Instrument Control Centre Team. 2017, in Astronomical Society of the Pacific Conference Series, Vol. 512, Astronomical Data Analysis Software and Systems XXV, ed. N. P. F. Lorente, K. Shortridge, & R. Wayth, 379
Gratier, P., Braine, J., Rodriguez-Fernandez, N. J., et al. 2010, A&A, 522, A3, doi: 10.1051/0004-6361/201014441
Griffin, M. J., Abergel, A., Abreu, A., et al. 2010, A&A, 518, L3, doi: 10.1051/0004-6361/201014519
Harris, C. R., Jarrod Millman, K., van der Walt, S. J., et al. 2020, Nature, 585, 537, doi: https://doi.org/10.1038/s41586-020-2649-2
Harris, F. J. 1978, in Proc. IEEE, 51–83
Hauser, M. G., Arendt, R. G., Berriman, G. B., et al. 1998, DIRBE Explanatory Supplement, Tech. rep., NASA Goddard Spaceflight Center
Hermelo, I., Relaño, M., Lisenfeld, U., et al. 2016, A&A, 590, A56, doi: 10.1051/0004-6361/201525816
HI4PI Collaboration, Ben Bekhti, N., Flöer, L., et al. 2016, A&A, 594, A116, doi: 10.1051/0004-6361/201629178
Hildebrand, R. H. 1983, QJRAS, 24, 267
Höfner, S., & Olofsson, H. 2018, A&A Rv, 26, 1, doi: 10.1007/s00159-017-0106-5
Hunt, L. K., Draine, B. T., Bianchi, S., et al. 2015, A&A, 576, A33, doi: 10.1051/0004-6361/201424734
Hunter, J. D. 2007, Computing in Science & Engineering, 9, 90, doi: 10.1109/MCSE.2007.55
Israel, F. P., Wall, W. F., Raban, D., et al. 2010, A&A, 519, A67, doi: 10.1051/0004-6361/201014073
Jenkins, E. B. 2009, ApJ, 700, 1299, doi: 10.1088/0004-637X/700/2/1299
Jenkins, E. B., & Wallerstein, G. 2017, ApJ, 838, 85, doi: 10.3847/1538-4357/aa64d4
Jones, A. P. 2004, in Astronomical Society of the Pacific Conference Series, Vol. 309, Astrophysics of Dust, ed. A. N. Witt, G. C. Clayton, & B. T. Draine, 347
Jones, A. P., Köhler, M., Ysard, N., Bocchio, M., & Verstraete, L. 2017, A&A, 602, A46, doi: 10.1051/0004-6361/201630225
Jones, E., Oliphant, T., Peterson, P., et al. 2001, SciPy: Open source scientific tools for Python. http://www.scipy.org/
Kalberla, P. M. W., Kerp, J., & Haud, U. 2020, A&A, 639, A26, doi: 10.1051/0004-6361/202037602
Kashlinsky, A., Mather, J. C., & Odenwald, S. 1996, ApJL, 473, L9, doi: 10.1086/310379
Kelly, B. C., Shetty, R., Stutz, A. M., et al. 2012, ApJ, 752, 55, doi: 10.1088/0004-637X/752/1/55
Kim, S., Staveley-Smith, L., Dopita, M. A., et al. 2003, ApJS, 148, 473, doi: 10.1086/376980
Kirk, H., Hatchell, J., Johnstone, D., et al. 2018, ApJS, 238, 8, doi: 10.3847/1538-4365/aada7f
Kirkpatrick, A., Calzetti, D., Kennicutt, R., et al. 2014, ApJ, 789, 130, doi: 10.1088/0004-637X/789/2/130
Koch, E. W., Rosolowsky, E. W., Boyden, R. D., et al. 2019, AJ, 158, 1, doi: 10.3847/1538-3881/ab1cc0
Koch, E. W., Rosolowsky, E. W., Lockman, F. J., et al. 2018, MNRAS, 479, 2505, doi: 10.1093/mnras/sty1674
Köhler, M., Ysard, N., & Jones, A. P. 2015, A&A, 579, A15, doi: 10.1051/0004-6361/201525646
Kurono, Y., Morita, K.-I., & Kamazaki, T. 2009, PASJ, 61, 873, doi: 10.1093/pasj/61.4.873
Leroy, A., Bolatto, A., Stanimirovic, S., et al. 2007, ApJ, 658, 1027, doi: 10.1086/511150
Leroy, A. K., Walter, F., Bigiel, F., et al. 2009, AJ, 137, 4670, doi: 10.1088/0004-6256/137/6/4670
Makarov, D., Prugniel, P., Terekhova, N., Courtois, H., & Vauglin, I. 2014, A&A, 570, A13, doi: 10.1051/0004-6361/201423496
Matsuura, M., Dwek, E., Meixner, M., et al. 2011, Science, 333, 1258, doi: 10.1126/science.1205983
McKinney, W. 2010, in Proceedings of the 9th Python in Science Conference, ed. S. van der Walt & J. Millman, 51 – 56
McMullin, J. P., Waters, B., Schiebel, D., Young, W., & Golap, K. 2007, in Astronomical Society of the Pacific Conference Series, Vol. 376, Astronomical Data Analysis Software and Systems XVI, ed. R. A. Shaw, F. Hill, & D. J. Bell, 127
Meixner, M., Gordon, K. D., Indebetouw, R., et al. 2006, AJ, 132, 2268, doi: 10.1086/508185
Meixner, M., Panuzzo, P., Roman-Duval, J., et al. 2013, AJ, 146, 62, doi: 10.1088/0004-6256/146/3/62
Miville-Deschênes, M.-A., Lagache, G., & Lagache, G. 2005, ApJS, 157, 302, doi: 10.1086/427938
Mizuno, N., Rubio, M., Mizuno, A., et al. 2001, PASJ, 53, L45, doi: 10.1093/pasj/53.6.L45
Molinari, S., Schisano, E., Elia, D., et al. 2016, A&A, 591, A149, doi: 10.1051/0004-6361/201526380
Murray, C. E., Peek, J. E. G., Di Teodoro, E. M., et al. 2019a, ApJ, 887, 267, doi: 10.3847/1538-4357/ab510f
Murray, C. E., Bot, C., Boyer, M. L., et al. 2019b, Scylla: a pure-parallel, multi-headed attack on dust evolution and star formation in ULLYSES galaxies, HST Proposal
Mutch, S. J., Croton, D. J., & Poole, G. B. 2011, ApJ, 736, 84, doi: 10.1088/0004-637X/736/2/84
Neugebauer, G., Habing, H. J., van Duinen, R., et al. 1984, ApJL, 278, L1, doi: 10.1086/184209
Nieten, C., Neininger, N., Guélin, M., et al. 2006, A&A, 453, 459, doi: 10.1051/0004-6361:20035672
Ochsenbein, F., Bauer, P., & Marcout, J. 2000, A&AS, 143, 23, doi: 10.1051/aas:2000169
Ott, S. 2010, in Astronomical Society of the Pacific Conference Series, Vol. 434, Astronomical Data Analysis Software and Systems XIX, ed. Y. Mizumoto, K.-I. Morita, & M. Ohishi, 139. https://arxiv.org/abs/1011.1209
Paradis, D., Dobashi, K., Shimoikura, T., et al. 2012, A&A, 543, A103, doi: 10.1051/0004-6361/201118740
Parvathi, V. S., Sofia, U. J., Murthy, J., & Babu, B. R. S. 2012, ApJ, 760, 36, doi: 10.1088/0004-637X/760/1/36
Pascale, E., Auld, R., Dariush, A., et al. 2011, MNRAS, 415, 911, doi: 10.1111/j.1365-2966.2011.18756.x
Pérez, F., & Granger, B. E. 2007, Computing in Science and Engineering, 9, 21, doi: 10.1109/MCSE.2007.53
Piazzo, L. 2017, IEEE Transactions on Image Processing, 26, 5232, doi: 10.1109/TIP.2017.2736421
Piazzo, L., Calzoletti, L., Faustini, F., et al. 2015, MNRAS, 447, 1471, doi: 10.1093/mnras/stu2453

Piazzo, L., Ikhenaode, D., Natoli, P., et al. 2012, IEEE Transactions on Image Processing, 21, 3687, doi: 10.1109/TIP.2012.2197009
Piazzo, L., Panuzzo, P., & Pestalozzi, M. 2015, Signal Processing, 108, 430, doi: https://doi.org/10.1016/j.sigpro.2014.09.039
Piazzo, L., Raguso, M. C., Calzoletti, L., Seu, R., & Altieri, B. 2016a, IEEE Transactions on Image Processing, 25, 4458, doi: 10.1109/TIP.2016.2592700
Piazzo, L., Raguso, M. C., Calzoletti, L., Seu, R., & Altieri, B. 2016b, in 2016 24th European Signal Processing Conference (EUSIPCO), 1553–1557
Pilbratt, G. L., Riedinger, J. R., Passvogel, T., et al. 2010, A&A, 518, L1, doi: 10.1051/0004-6361/201014759
Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2011a, A&A, 536, A1, doi: 10.1051/0004-6361/201116464
—. 2011b, A&A, 536, A17, doi: 10.1051/0004-6361/201116473
Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016a, A&A, 594, A8, doi: 10.1051/0004-6361/201525820
—. 2016b, A&A, 594, A5, doi: 10.1051/0004-6361/201526632
—. 2016c, A&A, 594, A7, doi: 10.1051/0004-6361/201525844
Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2018, arXiv e-prints, arXiv:1807.06207. https://arxiv.org/abs/1807.06207
Poglitsch, A., Waelkens, C., Geis, N., et al. 2010, A&A, 518, L2, doi: 10.1051/0004-6361/201014535
Priestley, F. D., & Whitworth, A. P. 2020, MNRAS, 494, L48, doi: 10.1093/mnrasl/slaa034
Relaño, M., De Looze, I., Kennicutt, R. C., et al. 2018, A&A, 613, A43, doi: 10.1051/0004-6361/201732347
Rémy-Ruyer, A., Madden, S. C., Galliano, F., et al. 2013, A&A, 557, A95, doi: 10.1051/0004-6361/201321602
Rice, W., Lonsdale, C. J., Soifer, B. T., et al. 1988, ApJS, 68, 91, doi: 10.1086/191283
Rieke, G. H., Young, E. T., Engelbracht, C. W., et al. 2004, ApJS, 154, 25, doi: 10.1086/422717
Rigby, A. J., Peretto, N., Adam, R., et al. 2018, A&A, 615, A18, doi: 10.1051/0004-6361/201732258
Roman-Duval, J., Bot, C., Chastenet, J., & Gordon, K. 2017, ApJ, 841, 72, doi: 10.3847/1538-4357/aa7067
Roman-Duval, J., Jenkins, E. B., Tchernyshyov, K., et al. 2021, ApJ, 910, 95, doi: 10.3847/1538-4357/abdeb6
Rosolowsky, E., Engargiola, G., Plambeck, R., & Blitz, L. 2003, ApJ, 599, 258, doi: 10.1086/379166
Roussel, H. 2012, Scanamorphos: Maps from scan observations made with bolometer arrays, Astrophysics Source Code Library. http://ascl.net/1209.012
—. 2013, PASP, 125, 1126, doi: 10.1086/673310
Ruiz-Lara, T., Gallart, C., Monelli, M., et al. 2020, A&A, 639, L3, doi: 10.1051/0004-6361/202038392
Saintonge, A., Kauffmann, G., Kramer, C., et al. 2011, MNRAS, 415, 32, doi: 10.1111/j.1365-2966.2011.18677.x
Schneider, F. R. N., Sana, H., Evans, C. J., et al. 2018, Science, 359, 69, doi: 10.1126/science.aan0106
Silverberg, R. F., Hauser, M. G., Boggess, N. W., et al. 1993, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 2019, Infrared Spaceborne Remote Sensing, ed. M. S. Scholl, 180–189, doi: 10.1117/12.157825
Slavin, J. D., Dwek, E., & Jones, A. P. 2015, ApJ, 803, 7, doi: 10.1088/0004-637X/803/1/7
Smith, M. W. L. 2013, PhD thesis, Cardiff University
Smith, M. W. L., Gomez, H. L., Eales, S. A., et al. 2012, ApJ, 748, 123, doi: 10.1088/0004-637X/748/2/123
Stanimirovic, S. 2002, in Astronomical Society of the Pacific Conference Series, Vol. 278, Single-Dish Radio Astronomy: Techniques and Applications, ed. S. Stanimirovic, D. Altschuler, P. Goldsmith, & C. Salter, 375–396. https://arxiv.org/abs/astro-ph/0205329
Stanimirovic, S., Staveley-Smith, L., Dickey, J. M., Sault, R. J., & Snowden, S. L. 1999, MNRAS, 302, 417, doi: 10.1046/j.1365-8711.1999.02013.x
Stepnik, B., Abergel, A., Bernard, J. P., et al. 2003, A&A, 398, 551, doi: 10.1051/0004-6361:20021309
Taylor, M. B. 2005, in Astronomical Society of the Pacific Conference Series, Vol. 347, Astronomical Data Analysis Software and Systems XIV, ed. P. Shopbell, M. Britton, & R. Ebert, 29
Thryng, K. M., Greene, C. A., Hetalnd, R. D., Zimmerle, H. M., & DiMarco, S. F. 2016, Oceanography. https://doi.org/10.5670/oceanog.2016.66
Torres, S., Aymon, J., Backus, C., et al. 1989, in Data Analysis in Astronomy, ed. V. di Gesu, L. Scarsi, P. Crane, J. H. Friedman, S. Levialdi, & M. C. Maccarone, 319
Towns, J., Cockerill, T., Dahan, M., et al. 2014, Computing in Science & Engineering, 16, 62, doi: 10.1109/MCSE.2014.80
Traficante, A., Calzoletti, L., Veneziani, M., et al. 2011, MNRAS, 416, 2932, doi: 10.1111/j.1365-2966.2011.19244.x
Valiante, E., Smith, M. W. L., Eales, S., et al. 2016, MNRAS, 462, 3146, doi: 10.1093/mnras/stw1806
van der Velden, E. 2020, Journal of Open Source Software, 5, 2004, doi: 10.21105/joss.02004
van der Walt, S., Colbert, S. C., & Varoquaux, G. 2011, Computing in Science & Engineering, 13, 22, doi: 10.1109/MCSE.2011.37
Veneziani, M., Piacentini, F., Noriega-Crespo, A., et al. 2013, ApJ, 772, 56, doi: 10.1088/0004-637X/772/1/56
Viaene, S., Fritz, J., Baes, M., et al. 2014, A&A, 567, A71, doi: 10.1051/0004-6361/201423534
Vogel, S. N., Wright, M. C. H., Plambeck, R. L., & Welch, W. J. 1984, ApJ, 283, 655, doi: 10.1086/162351
Weiß, A., Neininger, N., Hüttemeister, S., & Klein, U. 2001, A&A, 365, 571, doi: 10.1051/0004-6361:20000145
Wenger, M., Ochsenbein, F., Egret, D., et al. 2000, A&AS, 143, 9, doi: 10.1051/aas:2000332
Werner, M. W., Roellig, T. L., Low, F. J., et al. 2004, ApJS, 154, 1, doi: 10.1086/422992
Wheelock, S. L., Gautier, T. N., Chillemi, J., et al. 1994, IRAS sky survey atlas: Explanatory supplement, NASA STI/Recon Technical Report N
Williams, T. G., Gear, W. K., & Smith, M. W. L. 2018, MNRAS, 479, 297, doi: 10.1093/mnras/sty1476
Wong, T., Hughes, A., Ott, J., et al. 2011, ApJS, 197, 16, doi: 10.1088/0067-0049/197/2/16
York, D. G., Adelman, J., Anderson, Jr., J. E., et al. 2000, AJ, 120, 1579, doi: 10.1086/301513
Ysard, N., Jones, A. P., Demyk, K., Boutéraon, T., & Koehler, M. 2018, A&A, 617, A124, doi: 10.1051/0004-6361/201833386
Zhukovska, S., Dobbs, C., Jenkins, E. B., & Klessen, R. S. 2016, ApJ, 831, 147, doi: 10.3847/0004-637X/831/2/147

APPENDIX

Table A1. *Herschel* observation IDs we use for data reduction for each of our sample galaxies. Data from all bands observed by a given observation were used, unless otherwise noted.

Observation IDs	
LMC	
1342195668	1342195669
1342195683	1342195684
1342195707	1342195708
1342195712	1342195713
1342195728	1342202086
1342202087	1342202202
1342202203	1342202216
1342202217	1342202224
1342202225	1342202243
1342202244	
SMC	
1342192680	1342192681
1342192697	1342192698
1342192699	1342198565
1342198566	1342198590
1342198591	1342198863
1342205049	1342205050
1342205054	1342205055
1342205092	
M 31	
1342211294	1342211309
1342211319	1342213207
M 33	
1342189079[a]	1342189079
1342247408[a]	1342247409

[a] Used for PACS 100 μm only.

A. HERSCHEL OBSERVATIONS USED IN REDUCTIONS

In Sections 3.1 and 3.2, we discuss the *Herschel* PACS and SPIRE data used in this work. Here we present the specifics *Herschel* observation IDs that we used to reduce our data; they are listed in Table A1.

B. REGIONS MASKED DURING SPIRE DESTRIPING

In Section 3.2.1, we describe how, for our *Herschel*-SPIRE reduction process, we mask a number of bright regions in the LMC and SMC when running the destriper. This prevents artefacts being added, instead of removed, around these regions, and increases the speed with with the destriping algorithm can converge on a solution. These regions were identified manually, based on both their brightness, and their prominence relative to their surroundings. The positions and radii of the regions (all of which are circles) are listed in Table B2

Table B2. Positions and radii of circular regions masked when running the destriper during reduction of *Herschel*-SPIRE data for the LMC and SMC.

α (J2000)	δ (J2000)	Radius (arcmin)
LMC		
85.425	-69.551	53.8
82.776	-71.136	13.2
82.882	-68.509	15.5
83.592	-67.623	21.4
80.581	-67.942	14.0
81.422	-66.173	12.9
74.219	-66.379	21.4
72.905	-69.266	18.7
SMC		
11.849	-73.198	37.5
13.412	-72.715	38.6
15.468	-71.963	33.5
17.104	-72.803	33.6
19.715	-73.213	25.8
21.929	-73.454	25.8

C. FEATHERING IN-OUT TESTS

In order to test our feathering process, we performed a test whereby we generated three sets of test data: an objective map with *Herschel*-like resolution but with all large-scale emission preserved; a 'faux' *Herschel* map, with *Herschel*-like resolution and with large-scale emission removed; and a 'faux' *Planck* map, with no emission removed but with a low resolution. By feathering together the 'faux' *Herschel* and *Planck* maps, we should be able to re-create the objective map. How closely the feathered map matches the objective map allows us to evaluate how well the feathering has performed.

To create these data, we started with Sloan Digital Sky Survey (SDSS; York et al. 2000; Eisenstein et al. 2011) r-band observations of M 101. Obviously this is not FIR data, but it does contains both highly compact (individual stars, HII-regions) and highly extended

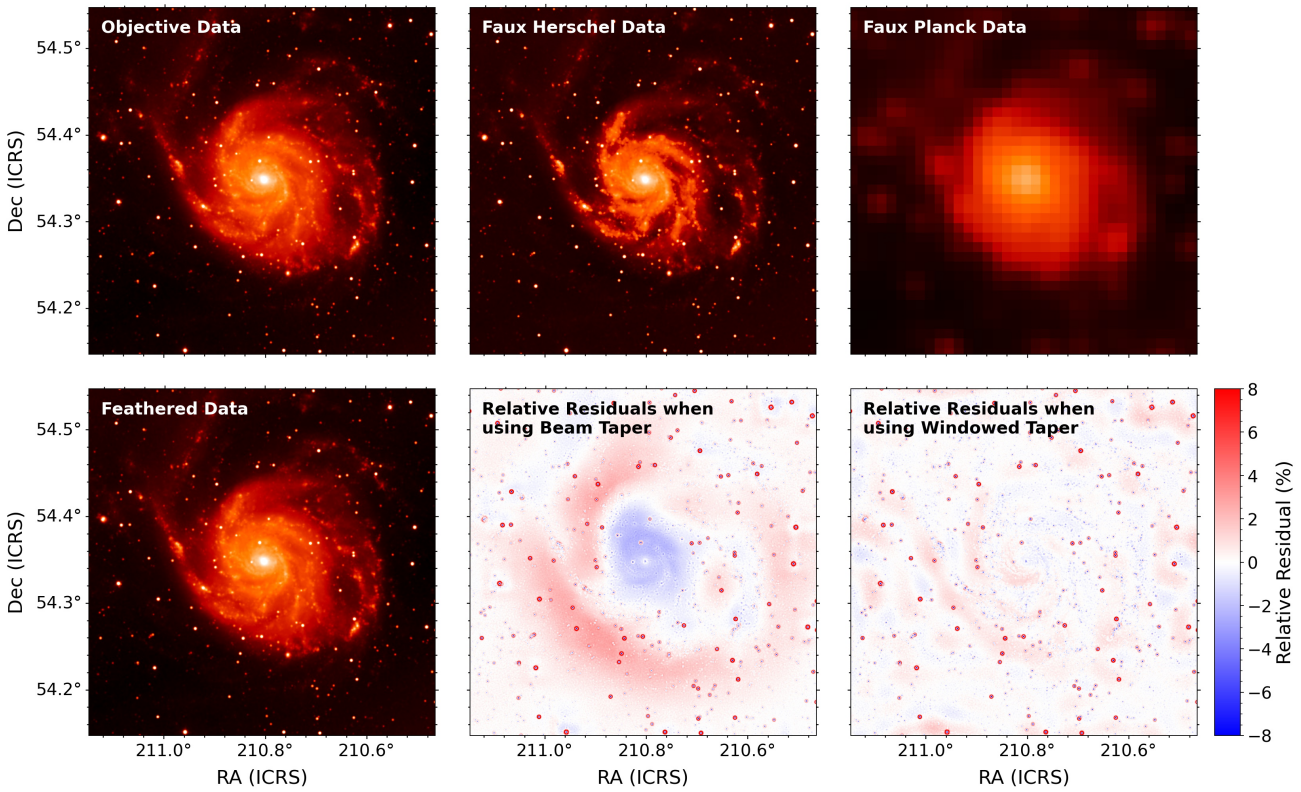

Figure C1. Illustration of our in-out test of our feathering process, as tested on SDSS r-band data of M 101. *Upper left:* The objective data, at 6″ resolution, with no emission filtered out. *Upper centre:* The 'faux' *Herschel* data, at 6″ resolution, with emission on scales $>200''$ filtered out. *Upper right:* The 'faux' *Planck* data, at 100″ resolution. *Lower left:* Feathered output, combining the two 'faux' observations, when using a windowed taper (the result when using a beam taper is visually extremely similar, so we show the one example). *Lower middle:* Relative residuals between the feathered map and the objective map, when using a beam taper. *Lower right:* Relative residuals between the feathered map and the objective map, when using a windowed taper.

(disc, stellar streams) emission. Plus, the SDSS resolution of ≈1.3″ is good enough, compared the *Herschel*- and *Planck*-like resolutions to which we will be smoothing it, that it can be treated as an effectively-infinite resolution 'ground truth'.

We convolved the SDSS r-band data to a resolution of 6″ to create our objective map. To create our 'faux' *Planck* map, we convolved the data to a resolution of 100″. Whilst 100″ is considerably better resolution than the ≈300″ resolution of *Planck*, M 101 is so compact that smoothing to that resolution would render it too close to being a point source to be useful for feathering. Plus, the factor of 16.7 difference in resolution between 6″ and 100″ well-matches the factor 16.3 difference in resolution between the *Herschel* 250 μm resolution of 18″, and the *Planck* 850 μm resolution of 293″ (to which we smoothed all of our IRIS–*Planck* data in Section 5.1).

To create out 'faux' *Herschel* map, we filtered large scale emission from the objective map; we did this by masking compact sources from the objective map, smoothing that map to a resolution of 200″, multiplying that map's surface brightnesses by 0.5, then subtracting this map from a copy of the original objective map. This produces a map with the same 6″ resolution as the objective map, but with emission on scales larger than 200″ filtered out (but with little negative bowling, much like real *Herschel* data). These three maps are shown in the top row of Figure C1.

We then feathered together the 'faux' *Herschel* and *Planck* maps. We did this two different ways. The first time we used the low-resolution beam to taper the transition from the high- to-low resolution data in Fourier space (as we did for feathering DIRBE with IRIS in Section 5.1). The second time we tapered within a window of angular scales to which both data were sensitive (as we did for feathering IRIS–*Planck* with *Herschel* in Section 5.1). Specifically, we used a tapering window of 120″ to 180″; this is above the 100″ resolution of the 'faux' *Planck* data, beneath and 200″ filtering scale of the 'faux' *Herschel* data. Plus the upper and lower

bounds of this window differ only by a factor of 1.5, compared to the factor of 2 difference for the window we actually use in Section 5.2, therefore the results in this test should be worse than those obtained in reality (specifically, a narrower window should cause more ringing), and so provide a conservative assessment.

To evaluate how accurately the feathering performed, we found the relative residual between the two output feathered maps, and the objective map. An example feathered output, and the two residual maps, are shown in the lower row of Figure C1.

When using the beam taper, there are some large-scale residuals; the surface-brightness in the centre of M 101 was underestimated, and the surface brightness in the outskirts of overestimated. However these effects are reasonably small, ranging from -2.5% to +3.0%. When using a tapering window, the residuals are much smaller in angular scale, tracing M 101's spiral arms, and also very small in magnitude, ranging from -0.9% to +1.1%.

This in-out test indicates that, for both methods, the total error introduced as a result of feathering is small – smaller than the instrumental calibration uncertainty for all instruments. Whilst the beam tapering seems to be somewhat less accurate, we are not able to window taper in the case of our DIRBE–IRIS feathering, as described in Section 5.1. As noted above, this in-out test has been designed to be conservative, and will likely be introducing slightly larger errors than our Local Group galaxy feathering will suffer in practice.

In the case of both methods, compact ringing residuals appear around bright point sources (in this case, stars), with a negative residual in the centre, surrounded by a ring of positive residual. However, total flux is conserved correctly overall for each source, and the entire artefact is always smaller than the instrumental PSF. Moreover, for our actual feathering with *Herschel* data, there are very few bright point sources of this kind, so we are not concerned about this kind of ringing affecting our final data products.

D. APODISATION

Unfeathered reprojection of low-resolution data to the pixel grid of high-resolution data can be problematic. An example of this, in the case of DIRBE and IRIS data, is shown in Figure D2. The DIRBE pixels are so large compared to the IRIS pixels ($150'$ versus $1.5'$) that the pixel edges remain extremely prominent in the reprojection. If this data is used for feathering, these pixel edge artefacts would still be partially visible in the final feathered data. A similar problem is encountered when reprojecting IRIS and *Planck* data to a *Herschel* pixel grid.

We prevent the problem by apodising the reprojected data – smoothing it with a kernel smaller than the low-resolution pixel size. This has the effect of suppressing the sharp pixel edges, whilst having only a very minor impact on the effective resolution of the low-resolution data. Specifically, we smooth the reprojected data with a Gaussian kernel with a standard deviation of $p_{low}/(p_{high}2\sqrt{2})$ pixels, where p_{low} and p_{high} are the pixel widths of the low- and high-resolutions maps respectively; this corresponds to 0.35 times the low-resolution pixel width. To make sure the feathering process can properly incorporate this change in the low-resolution effective PSF, we also convolved the low-resolution PSF with the apodisation kernel. There are other reprojection methods that result in a similarly 'smooth' output map (spline interpolation, etc) without requiring this after-the-fact apodisation; however, this method allows us to explicitly account for the resulting change in the PSF; this is important given that successful feathering depends upon the PSF being well-constrained.

E. IRAS-IRIS RESPONSE ARTEFACTS AROUND 30 DORADUS

Upon first producing feathered maps for the LMC in the 100 and 160 μm *Herschel*-PACS bands, we noticed large negative bowls around the 30 Doradus star-forming complex, especially at 100 μm. There was also considerable ringing around 30 Dor (and to a lesser extent around the star-forming region LHA 120-N 55A) in the maps of the residuals between the feathered data and the IRIS–*Planck* data. This ringing manifested as a large flux deficit around the perimeter of 30 Dor (hence the negative bowls), with some flux enhancement in the centre.

By comparing the differences in surface brightness between several of our datasets over this region, we realised that these artefacts were actually originating in the IRIS–*Planck* data. (Here, we again use the term 'IRIS–*Planck* data' to refer to the data that infers emission in the *Herschel* bands from SED-fitting to the IRIS and *Planck* maps.)

In the left panel of Figure E3, we plot the IRIS–*Planck* 100 μm surface brightness against the *Herschel*-PACS 100 μm / IRIS–*Planck* 100 μm surface brightness ratio; the datasets were convolved to the same resolution and projected to the same pixel grid before comparison, and limited to high-SNR pixels ($> 10\,\mathrm{MJy\,sr^{-1}}$ in IRIS–*Planck* 100 μm). For the most part the two maps agree quite closely, as would be expected, with the ratio remaining roughly constant (NB, there is no true zero level for the *Herschel*-PACS data, so the surface-

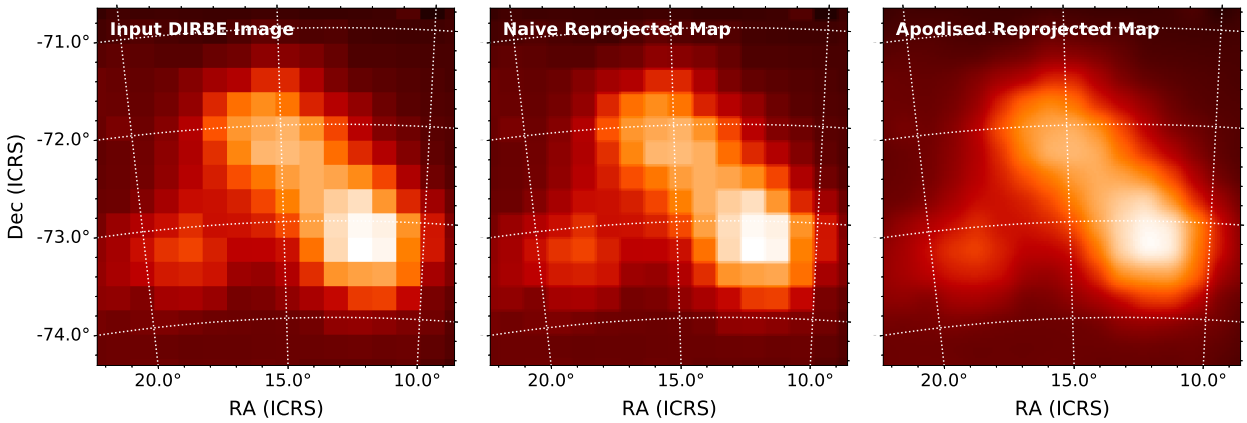

Figure D2. *Left:* DIRBE 100 μm data for the SMC. *Centre:* DIRBE data reprojected to the IRIS pixel grid; the DIRBE pixels have diameters 10 times greater than the IRIS pixels, so each 'big' pixel in this image is in fact made up of 10×10 IRIS-sized pixels, with the transitions between each 'big" pixel still being quite sharp. *Right:* The reprojected DIRBE data, apodised to remove the pixel edge artefacts.

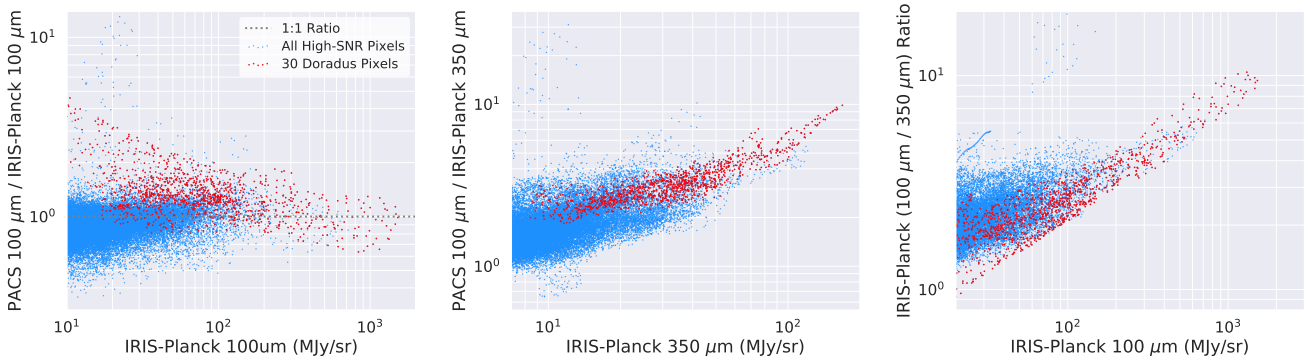

Figure E3. Plots used to diagnose the origin of artefacts around the 30 Doradus star forming region in the LMC. In each case, all data was convolved to the same resolution and projected to the same pixel grid before comparison. *Left:* IRIS–*Planck* 100 μm surface brightness plotted against the PACS 100 μm / IRIS–*Planck* 100 μm ratio. *Centre:* IRIS–*Planck* 350 μm surface brightness plotted against the PACS 100 μm / IRIS–*Planck* 350 μm ratio. *Right:* IRIS–*Planck* 100 μm surface brightness plotted against the IRIS–*Planck* 100 μm / IRIS–*Planck* 350 μm ratio. Pixels located with 0.5° of the centre of 30 Dor are highlighted in each plot.

brightness ratio values on the y-axis are unavoidably somewhat arbitrary). However for pixels within 0.5° of 30 Dor, highlighted in red, there is strong disagreement, with *Herschel*-PACS being much brighter than the IRIS–*Planck* for fainter pixels, and somewhat fainter for brighter pixels.

In the central panel of Figure E3, we plot IRIS–*Planck* 350 μm surface brightness against the *Herschel*-PACS 100 μm / IRIS–*Planck* 350 μm surface brightness ratio. The IRIS–*Planck* 350 μm data will naturally be dominated by the contribution of the *Planck* data to the IRIS–*Planck* SED fitting, and therefore should be well insulated against any IRIS-specific effects. This plot shows that the *Herschel*-PACS 100 μm data around 30 Dor is actually much better correlated with the IRIS–*Planck* 350 μm data than it was with the IRIS–*Planck* 100 μm data. We see that brighter pixels tend to have higher 100 μm/350 μm ratios (as expected due to colour effects), but that this trend is much the same for 30 Dor as for the rest of the LMC. This strongly indicates that the problem is with the IRIS–*Planck* 100 μm data, not the *Herschel*-PACS 100 μm data.

The right panel of figure Figure E3 plots the IRIS–*Planck* 100 μm surface brightness against the IRIS-*Planck* 100 μm/350 μm ratio. Once again, pixels in 30 Dor show markedly different behaviour than pixels in the rest of the LMC. Fainter pixels around 30 Dor can actually be *fainter* in IRIS–*Planck* 100 μm than in 350 μm; the inverse is true for the brighter pixels. This reflects what was seen in the first panel of the figure.

In short, the IRIS–*Planck* 100 μm data appears to be aberrant from both the *Herschel*-PACS 100 μm data *and* the IRIS–*Planck* 350 μm data.

The feathering processes was correcting the artefacts in the IRIS–*Planck* 100 μm data at smaller scales, where the erroneous emission in the IRIS–*Planck* map was replaced by the artefact-free emission in the *Herschel*-PACS map. But at larger scales, where the feathered map did not incorporate the *Herschel* emission, the artefacts persisted – hence the negative bowls. Our working assumption is that some aspect of our DIRBE+IRIS feathering gave rise to these artefacts – possibly because we had to use beam-mediated tapering, as opposed to a tapering window (see Section 5.1 and Appendix C). That is why the negatve bowls were strong in the 100 μm feathered map, and somewhat present in the 160 μm feathered map (where the impact of the 100 μm DIRBE+IRIS data on the inferred 160 μm surface brightness was lessened)

By extending the tapering window out to larger angular scales, we were able to increase the range of angular scales over which the artefacts were replaced by the artefact-free *Herschel*-PACS data. Fortunately, for the LMC, there is not an immediate, precipitous loss of power in the *Herschel*-PACS data as soon as one progresses to angular scales sampled by the IRIS–*Planck* (see Figure 10, and Section 5.1.1); rather, it seems that the *Herschel*-PACS data is mainly missing data on only the very largest scales. Therefore, we were able to employ a slightly larger tapering window to fix the artefacts, without excluding emission missed by *Herschel*-PACS. We experimented, to find a window that would remove the negative bowls from the final data whilst not causing residuals on the larger scales, and found that 45′–90′ appears to achieve this.

F. FEATHERING OUTPUTS

Here we provide illustrations of the outputs of our feathering *Herschel* data with IRIS–*Planck* data, along the associated diagnostic plots, at 100, 160, 350, and 500 μm (Figures F4 and F5). The 250 μm outputs are shown in Figure 13.

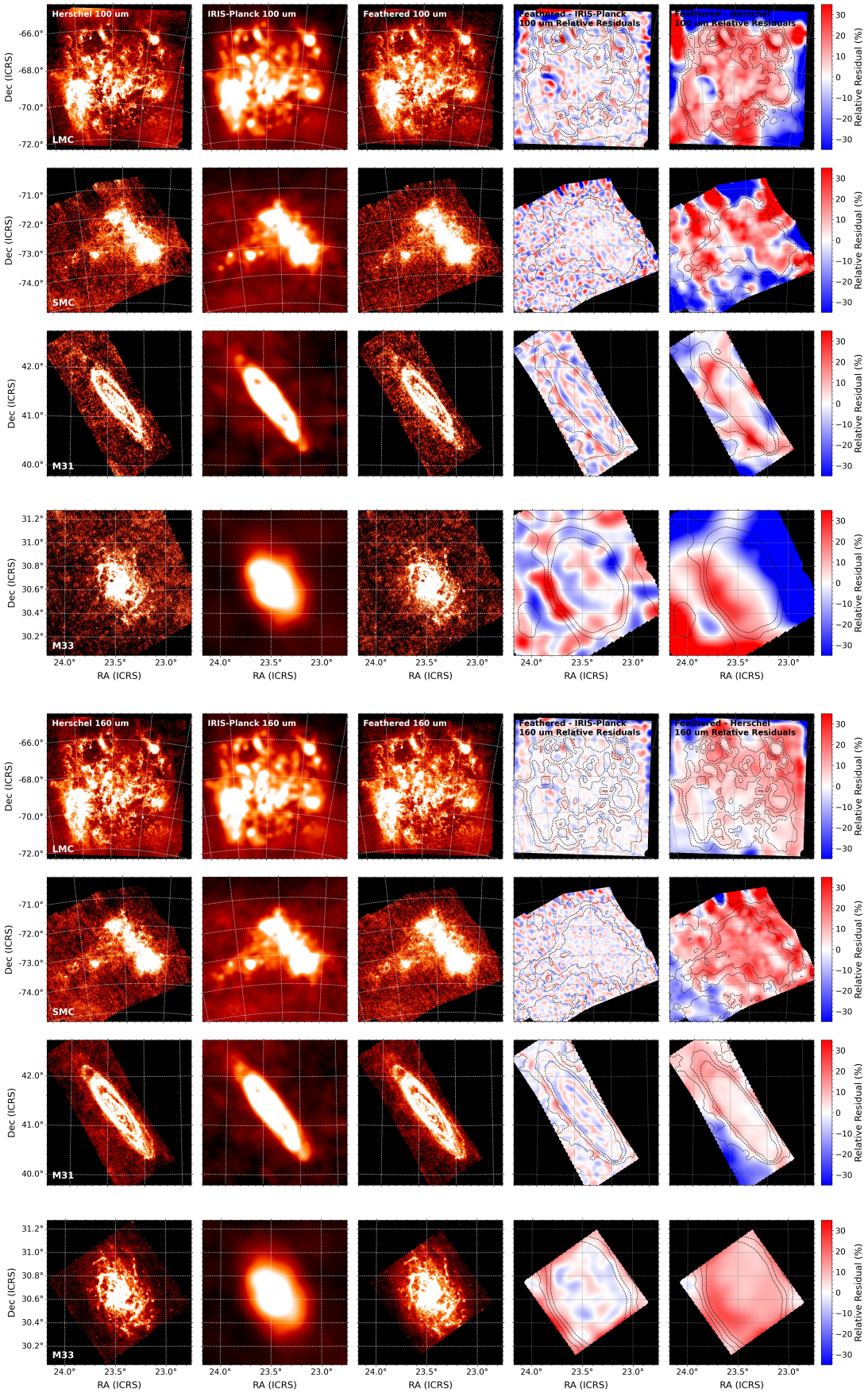

Figure F4. The results of feathering together *Herschel* with IRIS–*Planck* data at 100 μm (top four rows) and 160 μm (bottom four rows) for each of our galaxies, along with diagnostic plots. Description of panels same as per Figure 13.

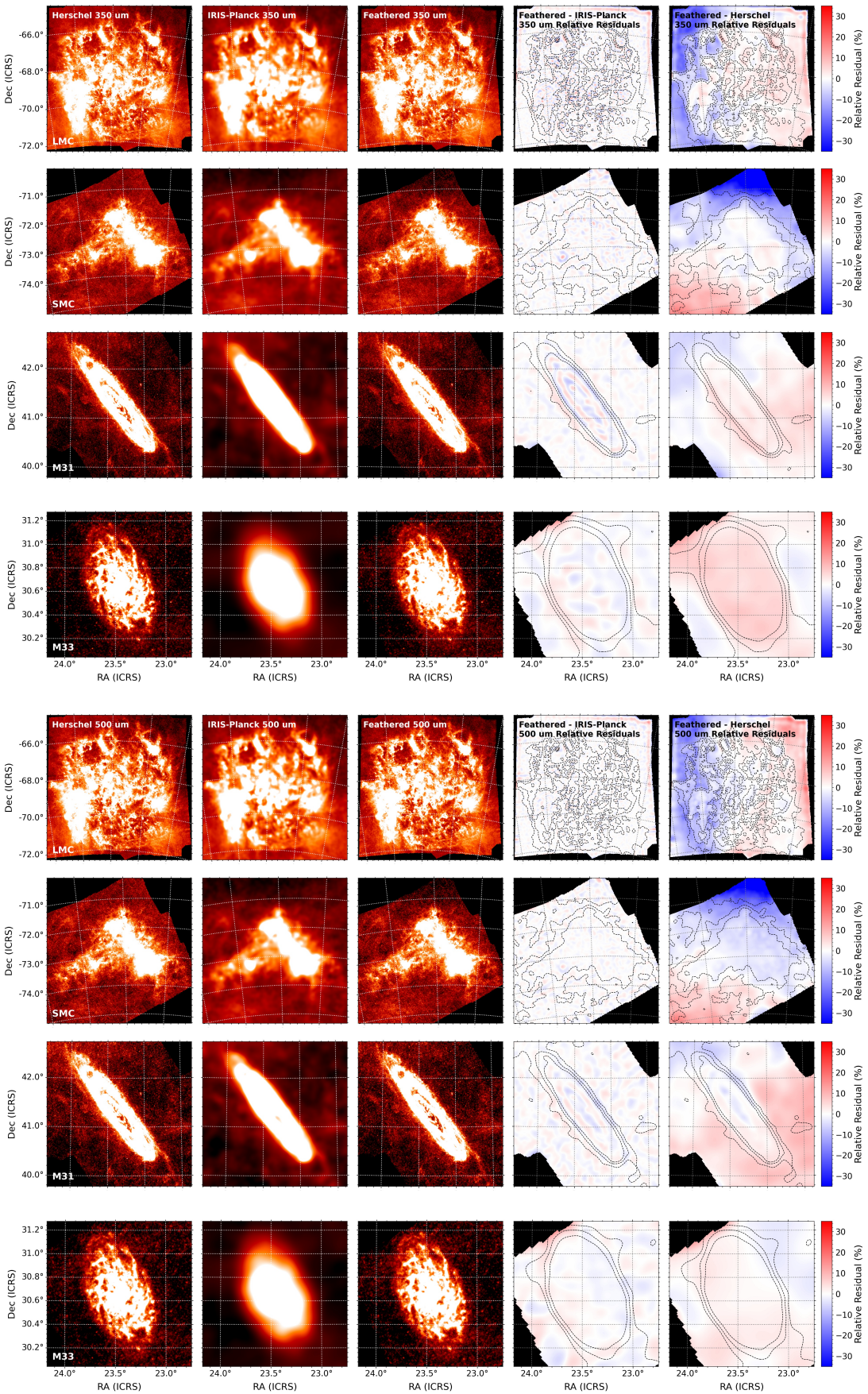

Figure F5. The results of feathering together *Herschel* with IRIS–*Planck* data at 350 μm (top four rows) and 500 μm (bottom four rows) for each of our galaxies, along with diagnostic plots. Description of panels same as per Figure 13.