Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

Mon. Not. R. Astron. Soc. **417,** 1192–1209 (2011) doi:10.1111/j.1365-2966.2011.19336.x

A Monte Carlo approach to evolution of the far-infrared luminosity function with BLAST

Gaelen Marsden,[1⋆] Edward L. Chapin,[1] Mark Halpern,[1] Guillaume Patanchon,[2] Douglas Scott,[1] Matthew D. P. Truch,[3] Elisabetta Valiante,[1] Marco P. Viero[4,5] and Donald V. Wiebe[1]

[1]*Department of Physics & Astronomy, University of British Columbia, 6224 Agricultural Road, Vancouver, BC V6T 1Z1, Canada*
[2]*Université Paris Diderot, Laboratoire APC, 10, rue Alice Domon et Léonie Duquet 75205 Paris, France*
[3]*Department of Physics & Astronomy, University of Pennsylvania, 209 South 33rd Street, Philadelphia, PA 19104, USA*
[4]*California Institute of Technology, 1200 E. California Blvd., Pasadena, CA 91125, USA*
[5]*Department of Astronomy & Astrophysics, University of Toronto, 50 St. George Street, Toronto, ON M5S 3H4, Canada*

Accepted 2011 June 28. Received 2011 June 2; in original form 2010 October 5

ABSTRACT

We constrain the evolution of the rest-frame far-infrared (FIR) luminosity function out to high redshift, by combining several pieces of complementary information provided by the deep Balloon-borne Large-Aperture Submillimeter Telescope surveys at 250, 350 and 500 μm, as well as other FIR and millimetre data. Unlike most other phenomenological models, we characterize the uncertainties in our fitted parameters using Monte Carlo Markov Chains. We use a bivariate local luminosity function that depends only on FIR luminosity and 60-to-100 μm colour, along with a single library of galaxy spectral energy distributions indexed by colour, and apply simple luminosity and density evolution. We use the surface density of sources, Cosmic Infrared Background (CIB) measurements and redshift distributions of bright sources, for which identifications have been made, to constrain this model. The precise evolution of the FIR luminosity function across this crucial range has eluded studies at longer wavelengths (e.g. using SCUBA and MAMBO) and at shorter wavelengths (e.g. with *Spitzer*), and should provide a key piece of information required for the study of galaxy evolution. Our adoption of Monte Carlo methods enables us not only to find the best-fitting evolution model, but also to explore correlations between the fitted parameters. Our model-fitting approach allows us to focus on sources of tension coming from the combination of data sets. We specifically find that our choice of parametrization has difficulty fitting the combination of CIB measurements and redshift distribution of sources near 1 mm. Existing and future data sets will be able to dramatically improve the fits, as well as break strong degeneracies among the models. Two particular examples that we find to be crucial are: obtaining robust information on redshift distributions and placing tighter constraints on the range of spectral shapes for low-luminosity ($L_{\rm FIR} < 10^{10}\,{\rm L}_{\odot}$) sources.

Key words: stars: formation – galaxies: evolution – galaxies: high-redshift – submillimetre: galaxies.

1 INTRODUCTION

It is now known that a significant fraction of the total light produced by stars and active galactic nuclei (AGN) throughout cosmic history is absorbed by dust and re-radiated thermally at much longer wavelengths. This light was first observed at low angular resolution by the *COBE* satellite as the diffuse Cosmic Infrared Background (CIB; Puget et al. 1996; Fixsen et al. 1998). Over the last decade it has been largely resolved into point sources at wavelengths ∼100–1000 μm, demonstrating that it is predominantly produced by individual galaxies (Dole et al. 2006; Pope 2007; Serjeant et al. 2008; Marsden et al. 2009; Pascale et al. 2009). Surveys with the *IRAS*, *ISO* and *Spitzer* satellites show that most of the shorter-wavelength light is produced by galaxies at redshifts $z \lesssim 1$, while ground-based submillimetre (submm) surveys have found most of the longer-wavelength light to be produced by more distant objects. Recent surveys by the Balloon-borne Large Aperture

⋆E-mail: gmarsden@phas.ubc.ca

Submillimeter Telescope (BLAST) at 250, 350 and 500 μm, a precursor to the *Herschel*/SPIRE surveys that are now well underway, have shown that the transition from low to high redshifts as one observes at longer wavelengths occurs gradually across the 250–500 μm band.

The fact that there is a transition to higher-redshift sources observed at longer wavelengths is not a surprise. Many groups have predicted this general behaviour using simple parametrized models for the evolution of local far-infrared (FIR) and submm galaxy luminosity functions. The data typically fit at these wavelengths include the surface density of sources as a function of brightness (source counts) and redshift information, when available (e.g. Blain & Longair 1993; Guiderdoni et al. 1997; Blain et al. 1999; Chary & Elbaz 2001; Malkan & Stecker 2001; Rowan-Robinson 2001; Lagache, Dole & Puget 2003; Dole, Lagache & Puget 2003; Lagache et al. 2004; Le Borgne et al. 2009; Valiante et al. 2009). These *phenomenological* models may be thought of as the simplest fitting functions available, since they typically include only two main ingredients: spectral energy distribution (SED) templates, to relate observed flux densities in different bands given luminosities and redshifts; and some evolutionary form for the luminosity function, to produce greater numbers of objects at higher redshifts – typically luminosity or density evolution following a power law in $(1 + z)$.

As observational data at ∼10–1000 μm have improved, in terms of wavelength coverage, survey area and depth, many authors have added greater complexity to their models. For example, it is now common to divide up the local luminosity function into multiple galaxy populations, assigning different SEDs to each, and then evolving the populations independently (e.g. Rowan-Robinson 2001; Lagache et al. 2003), or assuming some relation between the IR luminosity and the AGN content (Valiante et al. 2009). With this added freedom, such models can simultaneously fit the longer-wavelength 850–1200 μm SCUBA/MAMBO/AzTEC number counts and approximate redshift distributions, as well as the shallower *IRAS*/*ISO*/*Spitzer* surveys in the FIR (60–200 μm), and more recently the deep 24 μm *Spitzer* surveys (e.g. Lagache et al. 2004; Valiante et al. 2009; Rowan-Robinson 2009). Despite these successes, we note that, prior to the first measurements, predictions for the number counts in the BLAST/SPIRE bands varied widely (Patanchon et al. 2009, hereafter P09). We believe there are two main reasons for these discrepancies. First, the number of parameters associated with the multiple discrete populations is large, potentially leading to significant uncertainties in any part of the spectrum that lacks observational constraints. Secondly, it is common knowledge that rest-frame dust temperature, and hence bolometric luminosity and total dust mass, is degenerate with source redshift (e.g. Blain, Barnard & Chapman 2003); a 'redder' object could either be a cooler local galaxy, or a warmer, more distant galaxy. For this reason, assumptions about the SED shapes for each galaxy population, and the potential for evolution in these shapes, can significantly affect the results.

An alternative to phenomenological modelling of the data is the *ab initio* approach, or solutions to the forward problem: simulate as much of the physics of galaxy formation and evolution as possible, and evolve the model forward in time from the big bang, tweaking model parameters to fit observables (e.g. Baugh et al. 2005). Currently, such models usually incorporate the numerical evolution of the dark matter distribution, and adopt a range of recipes to assign galaxies to over-densities in this distribution, both as a function of mass and as the dark matter halo merger histories. Observational constraints include the cosmic microwave and infrared backgrounds, the luminosity functions and spatial clustering of local galaxy populations, and, more recently, their observed surface density and redshift distributions.

For the purpose of understanding galaxy evolution theory, *ab initio* models presently offer the most complete toolbox, combining a wide variety of physics into a single coherent model. However, they are still quite limited in the precision of inference that they allow, since the simulations lack the necessary resolution, and therefore assumptions about smaller-scale physics must be made (such as the details of star formation within molecular clouds, growth of black holes, and the interaction between the galaxy and the inter-galactic medium). These assumptions are guided by intuition and simple tuning recipes based, for example, on information from higher-resolution hydrodynamic simulations (e.g. Narayanan et al. 2010). These limitations result in a large number of parameters (undoubtedly with many partial degeneracies which can be tuned to match the sub-grid physics), and it can be extremely difficult to estimate meaningful uncertainties.

In this paper we consider a phenomenological model with more modest goals. Unlike *ab initio* models, and more recent multi-population phenomenological models that seek to fit the widest range of data available (e.g. attempting to connect the submm to FIR and mid-IR galaxy populations, as in Chary & Elbaz 2001; Lagache et al. 2004; Valiante et al. 2009), we focus our efforts on data that constrain *only* the evolution of the rest-frame FIR peak ($\lambda \sim$ 60–200 μm), for galaxies at redshifts up to $z \sim 4.5$. This redshift range encompasses the bulk of the 850–1200 μm submm galaxy (SMG) population which peaks near $z \sim 2.5$ (Aretxaga et al. 2003; Chapman et al. 2003a, 2005; Chapin et al. 2009b), as well as the most distant spectroscopically confirmed examples (Capak et al. 2008; Schinnerer et al. 2008; Coppin et al. 2009; Daddi et al. 2009). We are therefore only attempting to fit data that directly probe dust-reprocessed radiation from the most active star-forming galaxies, from their formation epoch to the present day. We also explicitly set out to determine whether a *single* galaxy population with a simple evolutionary form can reproduce the observed data across the peak in the CIB. Most authors have concluded, through fitting 'by hand', that multiple populations with independent evolutionary forms are required by the data; however, an exhaustive non-linear search of parameter space has never been performed as in this work to determine: (i) whether a more complicated model is indeed necessary; and if so (ii) identify precisely where the tension is coming from to probe the types of new models, and/or new data that would be required to fit such models.

For local galaxies, the region of the spectrum we are considering is quite smooth. At wavelengths $\gtrsim$100 μm, SED shapes resemble modified blackbodies, $S_\nu \propto \nu^\beta B_\nu(T_{\rm obs})$, although at slightly shorter wavelengths they are typically brighter and flatter due to a combination of opacity effects, and sensitivity to warmer dust grains. We therefore do not need to pay special attention to tuning the MIR spectra of our models, e.g. stochastically heated small dust grains, including polycyclic aromatic hydrocarbons, or PAHs (Draine & Li 2001), as has been necessary to fit the deep *Spitzer* 24 μm data (Chary & Elbaz 2001; Lagache et al. 2004; Valiante et al. 2009; Rowan-Robinson 2009). We do, however, include a range of SED shapes in our analysis, with a distribution characterized by the single $C \equiv \log_{10}(S_{60}/S_{100})$ colour near the peak of the rest-frame FIR emission (which is a good indicator of the FIR peak wavelength) and its well-known correlation with FIR luminosity, $L \equiv L_{\rm FIR}$, the integrated luminosity from 42.5 to 122.5 μm (Soifer & Neugebauer 1991; Chapman et al. 2003b; Chapin, Hughes & Aretxaga 2009a, hereafter C09a).

We combine this simple SED shape parametrization with the local *IRAS* luminosity function as our local boundary condition. We then evolve this local bivariate distribution, $\Phi(L, C)$, using only luminosity and density evolution, to fit the submm–FIR data. While the idea of incorporating a correlation between luminosity and FIR colour is not new (e.g. Lagache, Dole & Puget 2003; Lewis, Chapman & Helou 2005; Valiante et al. 2009), our reduced number of parameters allows us to fully explore the parameter space using Monte Carlo Markov Chains (MCMC). To our knowledge, there are only two other published attempts to fully characterize the uncertainties in a phenomenological model – Kelly, Fan & Vestergaard (2008) and Le Borgne et al. (2009), only the latter of which was concerned with submm–FIR surveys. In addition, a concurrent study by Béthermin et al. (2011) uses methods similar to those presented here. Another feature of our analysis that sets it apart from earlier work is our focus on the potential evolution in the correlation between luminosity and colour, since it is degenerate with redshift and heavily influences conclusions about the dust-enshrouded star formation rate history.

Throughout this paper, we consider two basic models: one in which the local correlation holds at high redshift; and a second in which the correlation undergoes luminosity evolution (making galaxies of a given luminosity in the past appear cooler than at the present day). This is an area for which BLAST data, and newer SPIRE data, provide the strongest constraints on this crucial part of the spectrum. This approach has allowed us both to: choose the model that best-fits existing data; and clearly indicate what future data are required to break the remaining degeneracies in parameter space.

A basic assumption that we make, as with all models of this type, is that high-redshift luminosity functions smoothly evolve over time to produce the modern-day $z = 0$ luminosity functions. If there is a significant galaxy population that existed in the early Universe, but is completely absent in the local Universe (even as a faint tail), our model will not give plausible results. We also emphasize the fact that our model *will not* give useful predictions for data far from the rest-frame FIR peak, such as the 24 μm source counts. Further work would be needed in order to achieve this, and in all likelihood, require additional model parameters.

In Sections 2.1 and 2.2, we describe the local boundary conditions of our model – the local luminosity and colour–luminosity distributions, and our adopted SED templates. The parametrization of the redshift evolution is described in Section 2.3, and the connection of this model to observed quantities (such as number counts, background intensities and redshift distributions) is provided in Section 2.4. The data sets that we use to fit the model are given in Section 3 and our fitting procedure is summarized in Section 4. The results of the fits are presented in Section 5, and Section 6 discusses the implications of and future improvements to the model.

Throughout this paper a standard cosmology is adopted, with $\Omega_{\rm M} = 0.272$, $\Omega_\Lambda = 0.728$ and $H_0 = 70.4\,{\rm km\,s^{-1}\,Mpc^{-1}}$ (Komatsu et al. 2011).

2 EVOLUTION OF THE LUMINOSITY FUNCTION

2.1 Local luminosity and colour distributions

In most past phenomenological models, authors have used either the local 60 μm luminosity function (e.g. Saunders et al. 1990) or the local 850 μm SCUBA luminosity function (Dunne et al. 2000). The former is one of the most well-studied luminosity functions, based on the all-sky *IRAS* 12, 25, 60 and 100 μm survey. Since SCUBA was not sensitive enough to conduct a survey over a significant portion of the sky, pointed follow-up of an *IRAS* galaxy sample was employed for the latter. This technique is adequate, provided that no significant population of cool ($\lesssim$25 K) galaxies exist in the local Universe. BLAST traces the peak of dust emission at high redshift, as *IRAS* does locally; emission at 850 μm comes from the Rayleigh–Jeans part of the spectrum. For these reasons, we choose an *IRAS*-based luminosity function as the basis for our model.

Since we are interested in the range of SED shapes that produce both the rest-frame submm and FIR emission, we also use the distribution of $C \equiv \log_{10}(S_{60}/S_{100})$ colours as a function of luminosity. The observed correlation, with no corrections for observational biases, was measured by Soifer & Neugebauer (1991) and has been used in some phenomenological models (e.g. Lagache et al. 2003). An attempt was made to measure the full bivariate luminosity–colour distribution using the $1/V_{\rm max}$ technique (Schmidt 1968) by Chapman et al. (2003b). However, we instead use the updated version of this distribution from C09a, which incorporates additional corrections for the *IRAS* bandpasses, a bias against detecting cooler galaxies in the original 60 μm flux-limited sample, and redshift evolution. We make one minor alteration to the distribution; as noted in C09a (and other previous authors, e.g. Saunders et al. 1990; Lawrence et al. 1999), the faint-end of the luminosity function is biased high by the local over-density of galaxies. Since the joint density of galaxies as a function of luminosity and colour is formulated in C09a as $\Phi(L, C) = \Phi(L)\,p(C|L)$ (i.e. the product of a pure luminosity function, with the conditional probability of a galaxy having a colour C given a luminosity L), we simply replace $\Phi(L)$ with the measurement from Saunders et al. (1990), which is valid because they used an estimator that is insensitive to this over-density (and which was shown in C09a to be consistent at luminosities $L > L_*$). Here (and throughout), L is defined to be the integrated 42.5–122.5 μm FIR luminosity. Throughout, we follow the convention that $\Phi(L)$ is the number density of objects per unit luminosity and $\phi(L)$ is the number density per decade of luminosity.

2.2 SED library

As mentioned in the Introduction, the shapes of the submm–FIR SEDs of most galaxies in the local Universe are reasonably well parametrized by a simple two-parameter modified blackbody (in addition to the normalization), $S_\nu \propto \nu^\beta B(\nu, T)$, where $B(\nu, T)$ is the Planck function. While a range of values of β have been measured, they typically do not vary much from a canonical value of $\beta = 1.5$. Furthermore, β and $T_{\rm obs}$ are highly anti-correlated in the fits. Therefore, it is plausible that a single parameter can accurately describe most of the spread in the shapes of locally observed FIR SEDs – in our case we use the FIR colour C. In further support of this simple parametrization, Dunne et al. (2000) showed that it was possible to map the 60 μm luminosity function to the 850 μm luminosity function by adopting SEDs that follow the observed correlation between temperature and luminosity. Similarly, Serjeant & Harrison (2005) fit temperatures to *IRAS* 60 and 100 μm data with a fixed $\beta = 1.3$, and for each object estimated their 850 μm brightnesses. They found, using these predicted 850 μm flux densities, that they could also map the *IRAS* luminosity function to the 850 μm luminosity function.

In the spirit of these earlier analyses, we seek a set of SED templates that can transform the *IRAS* colour–luminosity distribution to local luminosity functions at other adjacent wavelengths, namely: *IRAS* 12 μm (Fang et al. 1998); ISOCAM 15 μm (Xu 2000); *IRAS* 25 μm (Shupe et al. 1998) and SCUBA 850 μm (Dunne & Eales

2001).[1] The shortest wavelength data that we will attempt to fit are the observed source counts at 70 μm. It is therefore important to match the local luminosity functions at wavelengths as short as 12 μm, since this wavelength is redshifted to 70 μm at $z \sim 5$ (we only expect a minor contribution to the 12–15 μm SEDs from more complicated emission mechanisms, and this will only impact the 70 μm measurements for $z \gtrsim 4.5$ sources). On the long-wavelength side, by achieving consistency between the *IRAS* and 850 μm luminosity functions, we can expect to reasonably interpolate the rest-frame luminosity functions in the BLAST bands that are bracketed by these wavelengths. Again, we emphasize that we are *not* attempting to fit 24 μm number counts, which would require accurate modelling of the data at shorter wavelengths ($\lesssim 8$ μm) where the SEDs are considerably more complicated.

We have examined and rejected three common SED models for individual galaxies. First, we attempted to use single-temperature modified blackbodies. We produced a library of SEDs with fixed values of β and a range of temperatures. We then predicted each of the monochromatic luminosity functions, $\Phi(L_\nu)$, from $\Phi(L, C)$,

$$\Phi(L_\nu) = \int \Phi(L(L_\nu, C), C) \frac{\partial L(L_\nu, C)}{\partial L_\nu} \mathrm{d}C, \quad (1)$$

where $L(L_\nu, C)$ is the FIR luminosity for an SED in our library of colour C, normalized to the luminosity density L_ν at a frequency ν. Similar to Dunne et al. (2000) and Serjeant & Harrison (2005), we were able to obtain good agreement between the *IRAS* and 850 μm luminosity functions using $\beta = 1.5$. However, unsurprisingly, this simplistic model is a poor fit to the shorter wavelength data, since the Wien tail of our single-temperature SEDs falls off considerably more rapidly than for real galaxies, which contain a mixture of dust at different temperatures and compositions. Even the apparent plausibility of our SEDs for the longer-wavelength data can be deceiving; an effective $\beta = 1.5$ single-temperature spectrum can also be produced by the superposition of a range of dust populations with a steeper $\beta = 2.0$ at a range of appropriately selected temperatures. This fact serves to remind us that no simple physical meaning should be attached to these model parameters; the modified blackbody is only a convenient fitting function.

Next, we tested two more realistic SED libraries (each spanning the submm–IR wavelengths of interest) that are commonly used in the submm and FIR literature: the templates from Chary & Elbaz (2001) that were fit to data spanning 0.44–850 μm, and which were used in their phenomenological model, based on evolution of the 15 μm luminosity function; and the templates of star-forming galaxies from Dale et al. (2001) that were fit to *IRAS* and *ISO* data. The Chary & Elbaz (2001) templates provided a reasonable extrapolation to the luminosity functions at 12–25 μm, but led to a significant over-prediction at 850 μm. The Dale et al. (2001) SEDs performed much better at 850 μm, but led to moderate over-predictions at 12–25 μm. Due to these shortcomings, we decided to produce our own SED templates that vary smoothly as a function of C.

The basis of our SED library is the model of Draine & Li (2007). Their parametrization includes: (i) a set of mid-IR templates as a function of the PAH abundance, $q_{\rm PAH}$; and (ii) longer-wavelength templates for cooler thermal emission that is composed of dust heated both by a single low-intensity radiation field, $U_{\rm min}$, and by a second component heated by a range of radiation intensities from $U_{\rm min}$ to $U_{\rm max}$, where a factor γ gives the fraction of the total dust emission produced by this second component. We experimented with these four parameters ($q_{\rm PAH}$, $U_{\rm min}$, $U_{\rm max}$ and γ) to produce a sequence of smoothly varying SEDs as a function of C that resulted in good extrapolations to the monochromatic luminosity functions on both sides of the FIR peak.

We found that the results did not depend particularly heavily on $q_{\rm PAH}$ and so we simply fix it to an intermediate value of 2.50 (from a possible range spanning 0.10–4.58). The value of $U_{\rm min}$ effectively sets the apparent temperature of the coolest dust, and larger values of $U_{\rm max}$ and γ increase the temperature and fraction of the hotter dust (i.e. together these parameters control most of the submm–FIR SED shape). We obtained good fits for our extrapolated luminosity functions by fixing $U_{\rm max} = 10^4$ (from a possible range spanning 10^3–10^6), and stepping through all 22 of the supplied templates, corresponding to values of $0.1 \leq U_{\rm min} \leq 25.0$. We simultaneously increased γ logarithmically from 10^{-9} to 0.4 over the 22 levels. Finally, for this set of SEDs, we found that the values of C only spanned −0.50 to 0.10, thereby missing some of the warmest values identified in the local Universe, and to a lesser extent, some of the cooler values (see fig. 4 in C09a). Since the Rayleigh–Jeans side of the Draine & Li (2007) models resembles modified blackbodies with $\beta = 1.5$, we simply extended the library to a larger value of $C = 0.5$ by taking the $C = 0.10$ SED and adding on modified blackbodies with $\beta = 1.5$, temperatures T ranging from 47 to 98 K, and normalized such that they pass through the $C = 0.10$ model at 100 μm. At the red end, we simply assign the $C = -0.50$ SED to galaxies with $C < -0.50$ (i.e. galaxies that appear to have temperatures $T \lesssim 27$ K using a $\beta = 1.5$ modified blackbody are truncated at $T = 27$ K). Fig. 1 shows the complete set of SEDs. The transition to the extrapolated $C > 0.10$ SEDs is obvious as an abrupt increase in the density of templates that peak at wavelengths $\lambda \lesssim 30$ μm.

While it does not affect the results of the model fits, we also add on a power-law radio component to the SEDs, based on the FIR-radio correlation of Ivison et al. (2010b). These authors measure

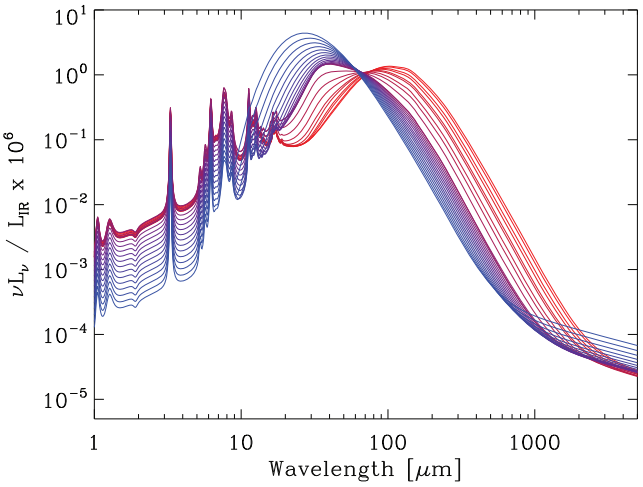

Figure 1. SEDs used in our model, generated from a range of templates presented in Draine & Li (2007). The 60-to-100 μm colour C of these SEDs ranges from −0.5 to 0.5, indicated by a colour gradient from red (cool, $C = -0.5$) to blue (warm, $C = 0.5$). The abrupt increase in the density of SEDs which peak at $\lambda \lesssim 30$ μm is caused by an extrapolation from the warmest Draine & Li (2007) model (with $C = 0.1$) to even larger values of C by adding modified blackbodies with $\beta = 1.5$ and temperatures ranging from 47 to 98 K. A radio component is added on, based on the FIR-radio correlations; this dominates at $\lambda \gtrsim 0.5$–2 mm.

[1] Since each study adopted slightly different values of H_0, we have corrected the luminosities and volumes for the value $H_0 = 70.4\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ used in this paper.

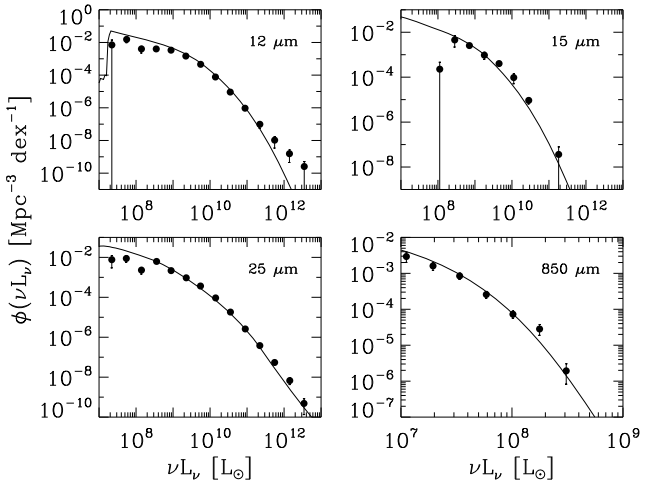

Figure 2. The local luminosity density functions derived from our adopted local luminosity function $\Phi(L)$, colour–luminosity function $p(C|L)$ and SED library, described in Sections 2.1 and 2.2. The 12, 15, 25 and 850 μm data sets are from Fang et al. (1998), Xu (2000), Shupe et al. (1998) and Dunne et al. (2000), respectively.

the quantity $q_{\rm IR}$, the ratio of the rest-frame 8–1000 μm flux to the 1400 MHz flux density,

$$q_{\rm IR} \equiv \log\left(\frac{S_{\rm IR}/3.75\times10^{12}}{\rm W\,m^{-2}}\right) - \log\left(\frac{S_{1400\,\rm MHz}}{\rm W\,m^{-2}\,Hz^{-1}}\right), \quad (2)$$

and also a power-law index $\alpha_{\rm rad}$. We add the radio continuum using $q_{\rm IR} = 2.41$ and $\alpha_{\rm rad} = -0.4$ at wavelengths longer than 100 μm.[2] This has a small effect on the number count predictions at millimetre wavelengths, presented in Section 5.

Fig. 2 shows the derived local monochromatic luminosity functions using the above SED library. The good correspondence indicates that we have found a plausible two-parameter model for the distribution of galaxy luminosities and colours spanning 12–850 μm. Later, in Section 6.1, we offer further evidence that this SED library is reasonable over the wavelength range of interest by comparing the colours of galaxies sampled from our best-fitting model with real surveys that span 24–850 μm.

2.3 Extension to high redshift

We use a simple extension of the local luminosity function $\Phi(L, C)$ to high redshift, incorporating parametric forms for density $\rho(z)$, luminosity $g(z)$ and colour–luminosity $h(z)$ evolution:

$$\Phi(L, C, z) = \rho(z) \times \Phi\left(\frac{L}{g(z)}\right) \times p\left(C\,\middle|\,\frac{L}{h(z)}\right). \quad (3)$$

Here, $\Phi(L, C, z)$ is the comoving luminosity evolution function, with units $L_\odot^{-1}$ Mpc^{-3}. Then the number of galaxies in a cell centred at (L, C, z) is

$$N = \Phi(L, C, z)\,\Delta L\,\Delta C\,\frac{{\rm d}V}{{\rm d}z}\,\Delta z, \quad (4)$$

where ΔL, ΔC and Δz are the dimensions of the cell.

We note that if $\Phi(L)$ were a power law, ρ and g would be completely degenerate (this is a well-known problem; e.g. Saunders et al. 1990). Our $\Phi(L)$ has both a break (at $L_* = 4\times10^9\,L_\odot$) and a low-luminosity cut-off (at $L_{\rm cut} = 10^8\,L_\odot$). These features, combined with our use of a colour–luminosity correlation, serve to break some of the degeneracy between the parametric functions, although strong correlations remain. We have not fully tested the effect of varying the low-luminosity cut-off (although see Section 6.5), but we note that the total luminosity, $\int L\Phi(L){\rm d}L$, integrated from $L_{\rm cut}$ to infinity, is ∼90 per cent of the full integral.

To reduce the number of free parameters, we set $h(z) = [g(z)]^\alpha$. The parameter α controls the amount of colour–luminosity evolution in the model: with $\alpha = 1$, the right-hand side of equation (3) can be written as $\rho(z) \times \Phi(L/g(z), C)$; with $\alpha = 0$, the colour–luminosity relation does not evolve with redshift. The parametrization of $\rho(z)$ and $g(z)$ is discussed in Section 4.1. We explore models with α fixed to 1.0 and 0.0, as well as with α allowed to vary as a free parameter. In all cases, ρ and g are free parametrized functions. The consequences of varying α are significant and are discussed in Section 6.4.

2.4 Observables

Our evolving luminosity function can be integrated across the appropriate variables to provide observables from the model. We first change variables from intrinsic luminosity L to observed flux density S_ν:

$$f(S_\nu, C, z) = \Phi\left(L(S_\nu, C, z), C, z\right)\left(\frac{\partial L}{\partial S_\nu}\right)\left(\frac{{\rm d}V}{{\rm d}z}\right), \quad (5)$$

with

$$L(S_\nu, C, z) = \frac{4\pi\,D_{\rm L}^2(z)\,S_\nu}{(1+z)\,T(C, (1+z)\nu)}, \quad (6)$$

where $D_{\rm L}$ is the luminosity distance and $T(C,\ \nu)$ converts FIR luminosity L to luminosity density L_ν at rest-frame frequency ν for the SED template with colour C. The ${\rm d}V/{\rm d}z$ term converts the counts from number per unit volume to number per unit redshift.

We use four types of data in this analysis:

(i) differential number counts, calculated by integrating across colour and redshift.

$$\frac{{\rm d}N(S_\nu)}{{\rm d}S_\nu} = \int_0^\infty\int_{-\infty}^\infty f(S_\nu, C, z)\,{\rm d}C\,{\rm d}z; \quad (7)$$

(ii) background intensity (CIB), obtained by further integration over S_ν,

$$I_\nu = \int_0^\infty S_\nu\left[\frac{{\rm d}N(S_\nu)}{{\rm d}S_\nu}\right]{\rm d}S_\nu; \quad (8)$$

(iii) background intensity as a function of redshift of emitting sources,

$$\frac{{\rm d}I_\nu}{{\rm d}z} = \int_0^\infty\int_{-\infty}^\infty S_\nu\,f(S_\nu, C, z)\,{\rm d}S_\nu\,{\rm d}C; \quad (9)$$

(iv) number of sources brighter than $S_{\rm lim}$ as a function of redshift,

$$\left.\frac{{\rm d}N}{{\rm d}z}\right|_{S_\nu > S_{\rm lim}} = \int_{S_{\rm lim}}^\infty\int_{-\infty}^\infty f(S_\nu, C, z)\,{\rm d}S_\nu\,{\rm d}C. \quad (10)$$

3 DATA

We now describe each of the data sets which we use to constrain the model.

[2] The quantity $q_{\rm IR}$ has been updated to 2.40 using *Herschel*/SPIRE observations (Ivison et al. 2010a), but as the change is significantly smaller than the measurement errors, we have not updated our value.

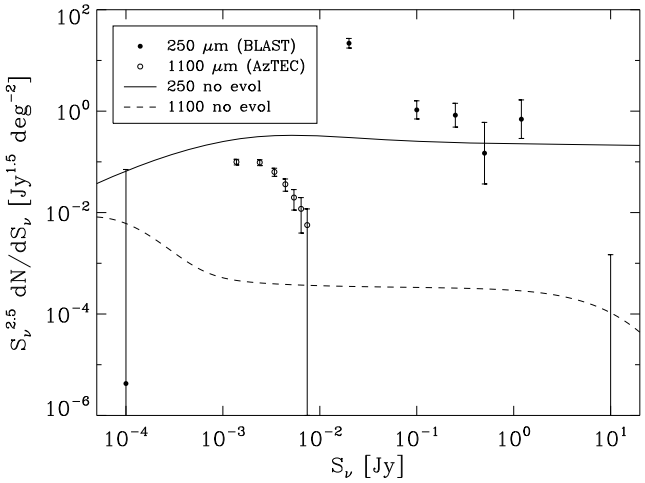

Figure 3. No-evolution models compared to data at 250 and 1100 μm. The error bars at 10^{-4} and 10^{1} are part of the BLAST 250 μm data set.

3.1 Number counts

From the choice of possible number counts, we use the following.

(i) *Spitzer* MIPS counts at 70 and 160 μm (Béthermin et al. 2010).

(ii) BLAST counts from table 3 of P09 (not constrained by the CIB) with covariance matrices (upper quadrants of tables 4–6 in P09). Since the BLAST counts are given as a series of nodes connected by power laws and not counts-in-bins, we treat these data differently than counts from other instruments. This is discussed further in Section 4.2. We have not accounted for the ∼10 per cent calibration errors, which are strongly correlated (Truch et al. 2009), or for errors due to cosmic variance, which may be significant at the bright end.

(iii) AzTEC 1.1 mm counts in the SHADES fields presented in Austermann et al. (2010). A covariance matrix is given; however, the correlations are very high, and the paper warns about over-interpreting measured correlations, so we use only the diagonal elements of the covariance matrix, ignoring correlations. This may over-weight the AzTEC counts, although a simple test (fitting the amplitude of a Schechter function with fixed shape) shows that the errors are reasonable.

In Fig. 3, we show the no-evolution counts (equation 3 with the evolutionary parameters ρ, g and h all set to 1.0). These are derived from the two-parameter local luminosity function described in Section 2.1 at 250 and 1100 μm compared to the measured counts by BLAST and AzTEC, respectively. We see that while the 250 μm counts are Euclidean at the bright end, they show strong evolution in the $S_\nu = 10$–100 mJy range. The 1100 μm counts show strong evolution over their full range.

3.2 Background intensity

We use observations of the CIB reported by Fixsen et al. (1998). We choose to fit the model at the same wavelengths at which the counts are used: 160, 250, 350, 500, 850 and 1100 μm. We assume 30 per cent band-independent errors (discussed further in Section 5.3).

3.3 Redshift distribution

We use dN/dz for SCUBA galaxies as measured by Chapman et al. (2005). They present a histogram of 73 sources (their fig. 4), but over-plot a model to show 'the likely effects of the sample selection.' We assign each of the nine histogram bins a Poisson error, then scale the histogram bin values and errors to fit their model. We fit a three-parameter Gaussian to the histogram and use the fitted amplitude to normalize the histogram. These scaled bin values and errors are used to constrain the shape of the model redshift distribution. We assume a limiting flux density $S_{\rm lim}$ of 5 mJy.

3.4 Data sets not used

A number of other relevant data sets exist that, for various reasons, we do not use to constrain the model. In most cases, the predictions of the best-fitting model are compared to the unused data sets in Section 5.

3.4.1 Counts

We have chosen to omit the SCUBA 850 μm number counts (e.g. Coppin et al. 2006) due to the fact that there is considerable tension between these measurements and those performed more recently with the AzTEC camera in a number of different fields at 1.1 mm. It has been noted since the very first surveys were undertaken with AzTEC that in order to scale its observed galaxy counts to those observed with SCUBA, each galaxy, roughly speaking, would have to be a factor of ∼3 brighter at 850 μm (e.g. Perera et al. 2008), whereas on an object-by-object basis, the individual galaxies appear to have a flux density ratio closer to 2, which is about what would be expected for a typical SED at redshift ∼2 (e.g. C09a; Perera et al. 2008). By itself, this comparison hints at a bias in counting, even though individual objects appear to be well-behaved. We note that as counting methods have improved, the downward correction for biases in the SCUBA counts has increased; the Coppin et al. (2006) measurement falls below essentially all of the previous estimates at $S_{850} > 2$ mJy, which were made using simpler techniques. Similarly, the methodology followed by the AzTEC team has also evolved over time and more biases have been discovered and corrected tending to lower the counts (e.g. Austermann et al. 2010; Downes et al. 2011). We also note that the counts at 870 μm measured with LABOCA fall considerably below those measured with SCUBA (Weiß et al. 2009).

We have found that it is impossible to fit a model that is consistent with the AzTEC 1.1 mm and SCUBA 850 μm counts simultaneously within the quoted uncertainties. We note that most recent phenomenological models have also only attempted to fit SCUBA counts, rather than including other counts near 1 mm (e.g. Lagache et al. 2003, 2004; Valiante et al. 2009; Le Borgne et al. 2009; Béthermin et al. 2011). We have explicitly checked that the model of Valiante et al. (2009), which fits the SCUBA 850 μm counts, significantly exceeds the AzTEC 1.1 mm counts, as we find here.

Taking all of these facts into consideration, there is strong evidence that the counts near 1 mm are biased by amounts that are not accurately characterized by published uncertainties. We feel that the best results will be obtained using the more recent, and more sophisticated, AzTEC measurements. This, however, is clearly an open subject that needs to be fully addressed in future assessments.

Number counts at 1.4 and 2.0 mm measured by the South Pole Telescope (SPT) were recently published by Vieira et al. (2010). They measured the bright end of the number counts, however, for which contributions by lensed galaxies are expected. Since our model does not include lensing, we do not use these counts to constrain our model.

3.4.2 Intensity in redshift slices

Pascale et al. (2009, table 2) presents the CIB at the BLAST wavelengths due to 24 μm sources in six redshift bins from $z \sim 0.4$ to 2. We attempted to use this data set to constrain the evolution model by comparing the values to the integral of the model over each redshift bin, but had difficulty reconciling this data set with the others; we found that, particularly at $z < 0.5$, the model preferred significantly lower values than the BLAST measurements. Since the analysis uses photometric redshifts, which may be unreliable, we have not included these data in our fits. We also note that the numbers of galaxies in their sample at these lower redshifts are extremely small; sampling variance due to clustering is certainly a large term that must be added to the quoted Poisson uncertainties.

3.4.3 Redshift distribution

A number of measurements of redshifts distributions have been made, including at: 170 μm by *ISO* (Patris et al. 2003; Dennefeld et al. 2005; Taylor et al. 2005); 250, 350 and 500 μm by BLAST (Dye et al. 2009; Dunlop et al. 2010; Chapin et al. 2011); and 1.1 mm by AzTEC (Chapin et al. 2009b). We assume limiting flux densities $S_{\rm lim}$ of 200, 40, 20 and 3.8 mJy for the 170, 250, 500 and 1100 μm distributions, respectively (we have not considered the 350 μm data set). We do not fit these dN/dz distributions, as the selection biases in producing the catalogues, due to the need for an optical counterpart to identify each source redshift, are difficult to quantify. This is also the case for the SCUBA distribution that we *do* use (although an attempt to correct the bias has been applied); however, we wanted to include at least one data set to constrain the redshift distribution, since the model is degenerate without it, and the SCUBA sample is the largest available.

4 MODEL FITTING

Given the data sets listed in Section 3, we map out the likelihood space of the luminosity evolution model using MCMC. This allows us to quote most likely parameter values along with errors and correlations between parameters.

4.1 Parametrization

We parametrize the evolution functions $\rho(z)$ and $g(z)$ as connected power laws at a series of nodes at a specific set of z_i. We have chosen six free parameters in each of ρ and z at a set of z_i spaced roughly linearly in $\log(1+z)$, spanning $z = 0$–5; the values used are z_i =0.1, 0.5, 1.0, 2.0, 3.5 and 5.0. The functions ρ and g are both fixed to 1.0 at $z = 0$ and to 10^{-12} at $z = 7$, which serves as a high-redshift cut-off. The evolution parameters $\log(\rho_i)$ and $\log(g_i)$ are then linearly interpolated in $\log(1+z)$.

4.2 Likelihood calculations

We calculate the likelihood of a model with a given set of evolution parameters based on χ^2,

$$\mathcal{L}(\{\xi_j\}) = \exp\left(-\frac{1}{2}\sum_i w_i \chi_{\rm i}^2\right), \quad (11)$$

where $\{\xi_j\}$ represents the $N_{\rm p}$ free parameters and the sum is over the $N_{\rm d}$ data sets. This formulation inherently assumes that the errors on the data sets are Gaussian distributed; this is not always a valid assumption, but it allows us to proceed in a straightforward manner.

In general, counts are treated as counts-in-bins. At each bin centre, we compare the model (equation 7) to the data. The full covariance matrix is used if available. However, as previously discussed, the BLAST $P(D)$ counts are treated differently: both the model and the data (connected power laws) are integrated between each pair of nodes, and the integrated values are compared. Errors (including correlations, which can be significant) on the data integrals are measured using Monte Carlo simulations, sampling from the chains produced by P09. We use Gaussians centred on the median values and with widths equal to half of the 68 per cent confidence regions. This is a reasonable description for all values except for the integrals between the two faintest bins at each wavelength, which have large positive tails.

For the SCUBA dN/dz, we compare the model to the normalized Gaussian discussed in Section 3.3 at the six redshift points z_i defined above. Errors at these points are obtained by propagating errors, including correlations, measured by the MCMC.

We are, in principle, free to set the relative weights w_i of each data set, but for now the weights are ignored ($w_i = 1$) and the relative importance of each data set depends on the number of measurements (and their errors) in each set.

4.3 Monte Carlos

We use CosmoMC[3] as our likelihood sampler (Lewis & Bridle 2002), using Metropolis–Hastings sampling. We run each model using four chains and run to an 'R-1' convergence of 0.003 to ensure accurate confidence limits. This typically requires $\sim 10^5$ samples per chain.

5 RESULTS

We fit 12 evolutionary parameters (ρ_i and g_i at the six z_i) to the data sets listed in Section 3. We find an important dependence on the parameter α, which governs the extent of evolution in the colour–luminosity relationship (Section 2.3); we have tested both $\alpha = 1$ (colour–luminosity evolution) and $\alpha = 0$ (no colour–luminosity evolution) and find that the $\alpha = 1$ model is a better fit to the data. We therefore concentrate on the $\alpha = 1$ model, but also show results for the $\alpha = 0$ model for comparison. The results of the MCMC analysis are presented in Figs 4–10. Throughout, the results for the $\alpha = 1$ model are in blue and for $\alpha = 0$ in red, with the best-fitting models in thick solid and thick dashed lines, respectively. Realizations of the model, including luminosity functions, counts and sample sources lists, are available at http://cmbr.phas.ubc.ca/model/.

The implications of the choice for α, along with fits with α as a free parameter, are discussed further in Section 6.4. Although it is not shown here, the free α model leads to an intermediate value of $\alpha = 0.62 \pm 0.04$, with correlation coefficients (to the other parameters) as large as 0.6; the correlations are largest for the low-redshift parameters and are nearly zero for the high-redshift parameters ($z_i > 1$).

5.1 Parameters

The best-fitting evolution functions $\rho(z)$ and $g(z)$, along with 68 and 95 per cent confidence intervals, are shown in Fig. 4. Both models

[3] http://cosmologist.info/cosmomc/

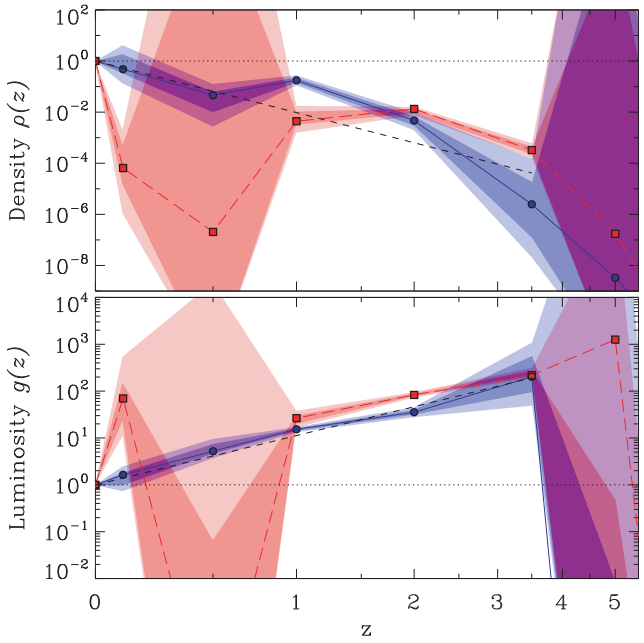

Figure 4. The parameter values, which are fit to the data sets shown in Figs 6–10. Above, density evolution, $\rho(z)$, and below, luminosity evolution $g(z)$. The *C*–*L* evolution ($\alpha = 1$) and no *C*–*L* evolution ($\alpha = 0$) models are shown, in blue and red, respectively. The values of ρ and g at $z = 0$ and $z = 7$ are fixed to 1.0 and 10^{-12}, respectively. The values in between, at z = 0.1, 0.5, 1.0, 2.0, 3.5 and 5.0, are free parameters in the model. The blue (red) symbols connected by solid (long-dashed) lines indicate the best-fitting parameter values for the $\alpha = 1$ ($\alpha = 0$) models. The coloured bands represent the 68 and 95 per cent confidence regions. We note that ρ_i and g_i at each z_i are highly anti-correlated (see Fig. 5). Representative power laws, $\rho(z) = (1 + z)^{\gamma}$ and $g(z) = (1 + z)^{\delta}$ up to a cut-off redshift $z = 3.5$, with $\gamma = -6.7$ and $\delta = 3.5$, have been over-plotted as black short-dashed lines.

show clear negative density evolution and positive luminosity evolution with increasing redshift. Both models remain well constrained up to $z = 3.5$, with the exception of the $z = 0.5$ node in the $\alpha = 0$ model; this feature is discussed further in Section 5.2.1. The data do not constrain the models above $z = 3.5$.

To guide the eye, we show (as a thin-dashed line) for both density and luminosity functions a representative single power law, $\rho(z) = (1 + z)^{\gamma}$ and $g(z) = (1 + z)^{\delta}$ up to a cut-off redshift $z = 3.5$. These values have not been fit to the data, but instead fit to the best-fitting $\alpha = 1$ values. The power indices of the lines shown are $\gamma = -6.7$ and $\delta = 3.5$. A generic feature of our models is therefore a combined trend of *negative density evolution* and *positive luminosity evolution* as one observes FIR-selected galaxies further into the past.

Density evolution may be considered an indication of the overall galaxy merger rate as a function of time. If galaxy formation were to follow a simple bottom-up scenario, such as in the case of dark matter halo merger histories, the smallest bodies are created first, and over time merge together to form a smaller number of galaxies (i.e. positive density evolution). Luminosity evolution does not change the total number of galaxies in the Universe, but rather their brightness distribution. Under the previous scenario, one might naively expect a large number of less luminous galaxies to merge together in the past, and produce a smaller number of more luminous sources in the present (i.e. negative luminosity evolution combined with positive density evolution). However, precisely the opposite behaviour is observed in deep extra-galactic surveys at different wavelengths. This apparent 'cosmic downsizing' (as first noted by Cowie et al. 1996) has been a topic of great interest to theorists, and our modelling results continue to support the trends observed at other wavelengths: regardless of the parametrization, there were fewer, but more luminous FIR-bright galaxies in the past. For the most luminous of these galaxies to have formed in the early Universe, most of their stars must have been created in a relatively short burst (less than 1 Gyr for galaxies at $z \gtrsim 3.0$, given the available time since the first generation of stars during re-ionization). The FIR-luminous galaxies presumably faded slowly over time as the more massive, luminous stars completed their life-cycle, giving way to lower-mass, cooler and less luminous stars.

The parameter distributions and correlations are shown in Fig. 5. The distributions for the $\alpha = 1$ model are shown in the lower-left, and for the $\alpha = 0$ model in the upper-right. The diagonal elements show the distributions of the free parameters, $\rho_1, \ldots, \rho_6, g_1, \ldots, g_6$. Above and below the diagonals are the joint distributions of all pairs of parameters. We note a few features: (i) the $z_1 = 0.5$ parameters for the $\alpha = 0$ model are unconstrained, as was noted above; (ii) the $z_6 = 5.0$ parameters are unconstrained by the data for both models, also noted above; (ii) for both ρ and g, neighbouring parameters (e.g. ρ_i and ρ_{i+1}) are anti-correlated; and (iv) ρ_i and g_i (at the same z_i) are strongly anti-correlated.

5.2 Comparison to data

Comparisons of the model to the constraining data sets are shown in Figs 6–10.

5.2.1 Number counts

Euclidean-normalized differential number counts at a range of wavelengths are shown in Figs 6 and 7 (for the $\alpha = 1$ and $\alpha = 0$ models, respectively). The counts derived from the best-fitting models are shown as solid lines, with the coloured regions showing the 68 and 95 per cent confidence regions. Also shown are the contributions to the counts from sources binned by redshift, along with the 68 per cent confidence regions. The counts at 850, 1400 and 2000 μm are shown for illustrative purposes only, and have not been used in the fits.

Comparing the counts derived from the $\alpha = 1$ and 0 models, we see that, with the exception of 160 μm, which the $\alpha = 0$ model is not able to reproduce, the two models give very similar counts through the flux density regions covered by data. We see, however, that the curves are formed from completely different redshift distributions; the $\alpha = 1$ model shows a prominent $z = 0.1$–0.5 component at all wavelengths, while this component is completely sub-dominant in the $\alpha = 0$ model. This fact provides an explanation for why the $z = 0.5$ density and luminosity evolution parameters in the $\alpha = 0$ model are completely unconstrained – since sources at this redshift do not contribute appreciably to the observed counts at any wavelength, there is only an upper limit to the strength of sources in this redshift range.

We remind the reader that the 250, 350 and 500 μm BLAST counts are fit using the integral of dN/dS_{ν} across a set of bins. An example fit (250 μm for the $\alpha = 1$ model) is presented in Fig. 8. This shows very clearly that the two models are able to produce essentially the same counts. We also point out that the bin-to-bin correlations are in some cases quite strong (up to $\rho = 0.95$), and that the χ^2 calculated from the full covariance matrix can be quite different from what would be inferred using the diagonal elements

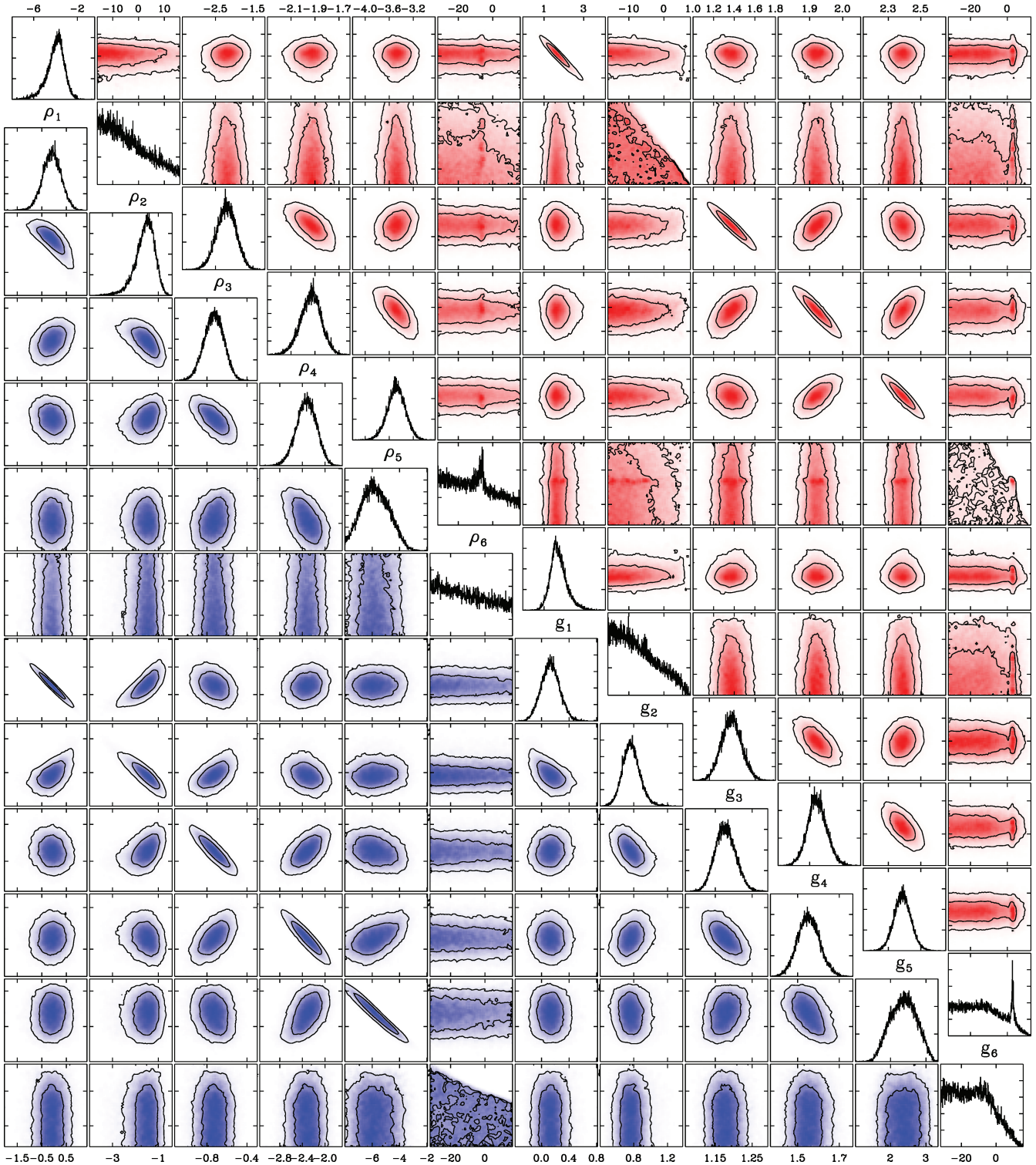

Figure 5. Likelihood space of $\log(\rho_i)$ and $\log(g_i)$ for the two models, as sampled using MCMC. The *C–L* evolution ($\alpha = 1$) and no *C–L* evolution ($\alpha = 0$) models are shown, in blue (lower-left) and red (upper-right), respectively. For each set, the diagonal shows the likelihood of each of the free parameters (ρ_i followed by g_i, from top-left to bottom-right, for $z_i = 0.1$, 0.5, 1.0, 2.0, 3.5 and 5.0). Below/above the diagonal are the correlations between all pairs of parameters. Contours are the 68 and 95 per cent confidence regions. We note that ρ and g at each z_i are highly anti-correlated, and there are moderate anti-correlations between adjacent redshift nodes. For both data sets, the $z = 5$ nodes are not constrained, and the $z = 0.5$ nodes are unconstrained for the $\alpha = 0$ model, as can be seen in Fig. 4. A spike is apparent in the $\alpha = 0$, $z = 5$ nodes; this feature appears to be consistent across and throughout the chains, but as it is uncorrelated with the other more well-constrained, lower-redshift parameters, we believe it has no effect on model predictions.

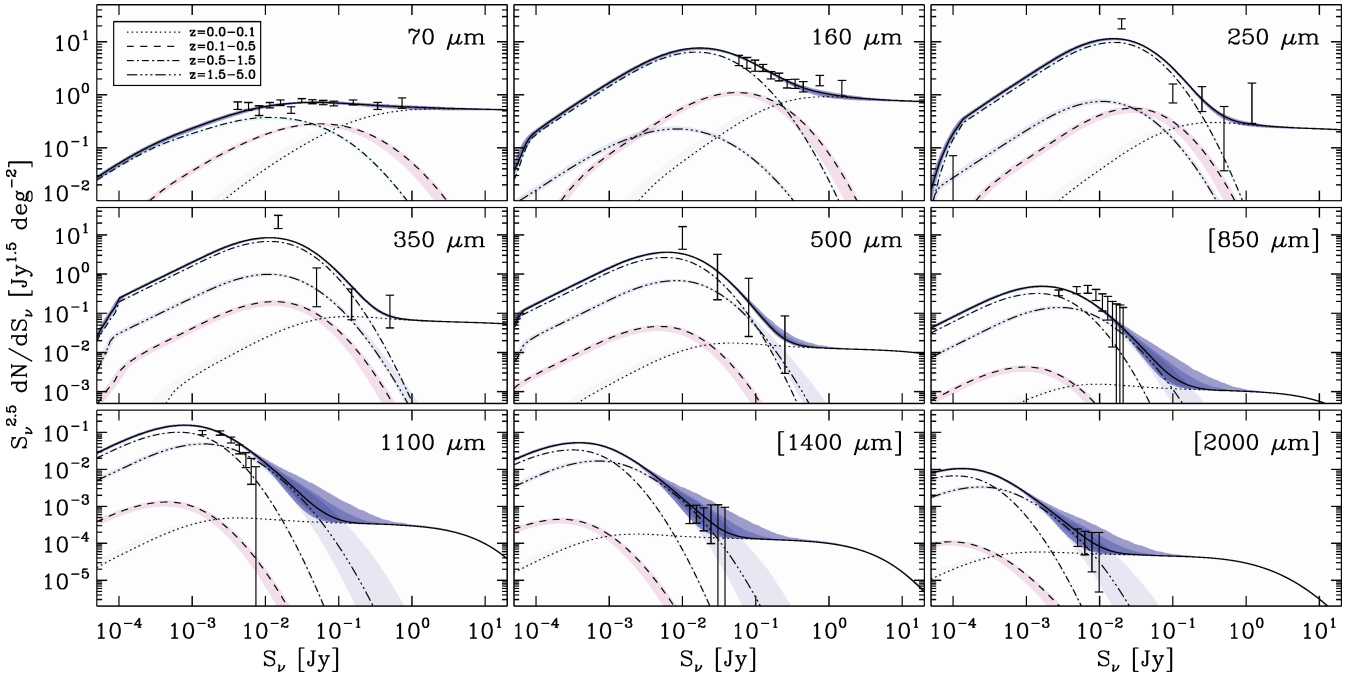

Figure 6. Euclidean-normalized differential number counts. The solid lines are the counts derived at each wavelength from the best-fitting $\alpha = 1$ model. The coloured bands represent the 68 and 95 per cent confidence regions. The contribution of sources at $z = 0.0$–0.1, 0.1–0.5, 0.5–1.5 and 1.5–5.0 are shown as dotted, dashed, dot–dashed and triple-dot–dashed lines, respectively. The coloured bands indicate the 68 per cent confidence regions. We note that the $z = 0.1$–0.5 component is sub-dominant at all flux densities and wavelengths, except for a narrow region in the shortest bands. The BLAST data sets (250–350 μm) include poorly constrained points at low and high flux density that are not included in the range shown here (see Fig. 8). Counts at 850 μm (SCUBA) and 1.4 and 2.0 mm (SPT) are shown for comparison only, and are not used in the fits.

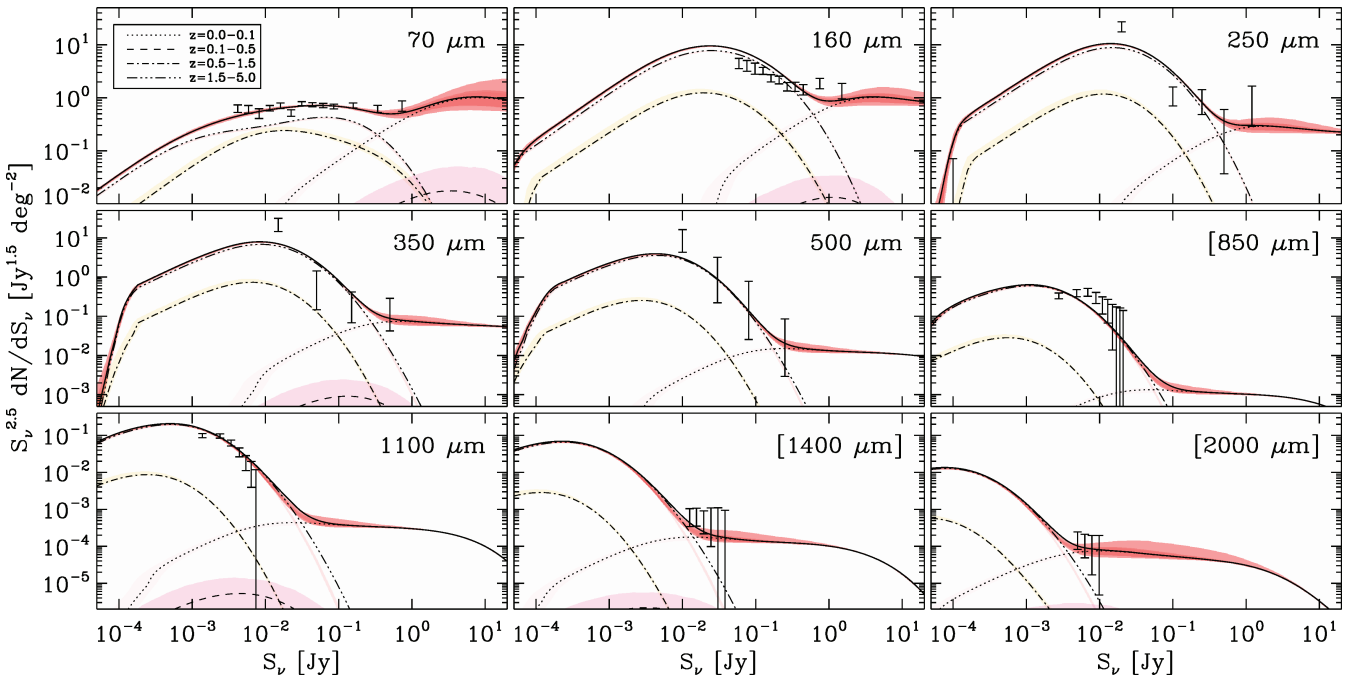

Figure 7. Same as Fig. 6, but for $\alpha = 0$. Note that, with the exception of 160 μm, which is not well-fit, the shape of the $\alpha = 0$ counts is very similar to the $\alpha = 1$ counts (Fig. 6), but the contribution from the various redshift ranges is completely different. In particular, the $z = 0.1$–0.5 contribution never dominates at any band or flux density.

alone. This is particularly true at 500 μm, where $\chi^2 = 45$; if we ignore the correlations, we instead find $\chi^2 = 7$.

Although best fits have been found, it is worth pointing out that none of the overall fits is formally 'good'; in particular, the model is quite low at the $S_\nu = 10$–20 mJy BLAST counts and is high compared to 1100 μm counts at the bright end. We believe that, due to the choice of parametrization in the P09 $P(D)$ counts, these points are biased high. This comes about since a connected power law is

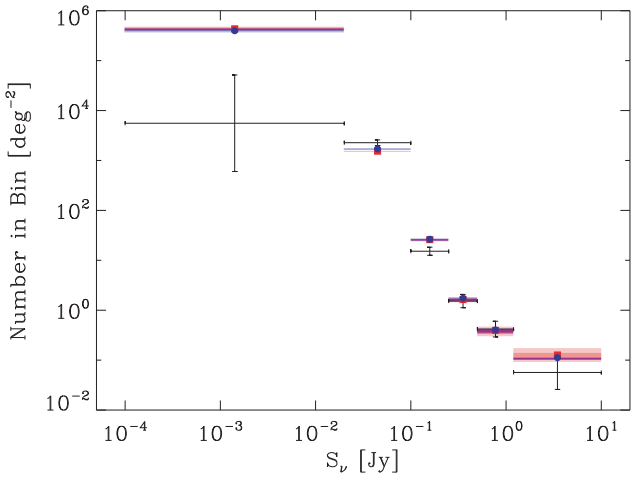

Figure 8. The BLAST 250 μm $P(D)$ counts integrated between nodes compared to models (colour-coded as in the previous figures). The horizontal error bars indicate the size of the bins (the bin edges correspond to the flux densities of the nodes listed in P09 and plotted in Figs 6 and 7), while the vertical bars show the errors determined from Monte Carlo sampling of the P09 covariance matrix.

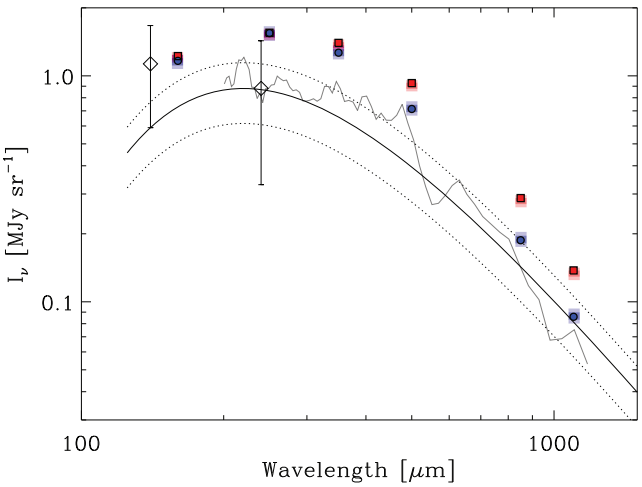

Figure 9. The CIB as measured by FIRAS (Fixsen et al. 1998, solid line) with representative 30 per cent errors (dotted lines). The circular and square points represent the CIB derived from the best-fitting models, $\alpha = 1$ and $\alpha = 0$, respectively. The coloured rectangles represent the 68 and 95 per cent confidence regions. Also shown, but not used in the fits, are the FIRAS (solid grey line) and DIRBE/WHAM measurements (diamonds) by Lagache et al. (2000). Our models tend to over-predict the background values; we discuss the reason for this in Section 6.

not able to reproduce the curvature displayed by the model. Recent results from *Herschel*/SPIRE using $P(D)$ to measure counts (Glenn et al. 2010), which go to fainter flux densities than the BLAST counts, are low compared to BLAST, supporting this hypothesis (which we discuss further at the end of Section 5.3).

Another assumption that may be biasing the fits is that the three BLAST measurements are correlated with one another, both by instrumental noise and by the fact the same part of the sky has been observed. For both of these reasons, it would be desirable to fit the model directly to the maps via multi-band $P(D)$. This will be the focus of a later paper.

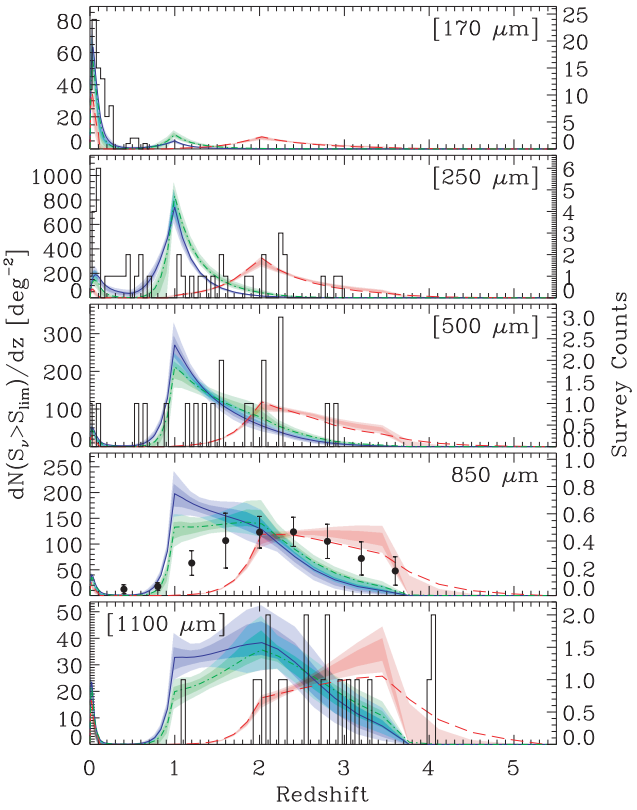

Figure 10. The redshift distribution of number counts above a flux density limit. The thick solid/long-dashed/dot–dashed curves show the predictions of the best-fitting $\alpha = 1/\alpha = 0$/free α models. The coloured bands represent the 68 and 95 per cent confidence regions. At 170 μm (*ISO*), 250 and 500 μm (BLAST) and 1.1 mm (AzTEC), redshift histograms are shown. At 850 μm (SCUBA), we show the scaled histogram of Chapman et al. (2005) with Poisson error bars, as described in Section 3.4.3. Redshift distributions at 170 (Patris et al. 2003; Dennefeld et al. 2005; Taylor et al. 2005), 250, 500 μm (Chapin et al. 2011) and 1.1 mm (Chapin et al. 2009b) are shown for comparison only, and are not used in the fits.

5.2.2 *Integrated brightness (CIB)*

The CIB as measured by FIRAS is used as a constraint on the integral of intensity over redshift at a range of wavelengths. This is shown in Fig. 9 for both models. The best-fitting models are shown as filled symbols, with the 68 and 95 per cent confidence regions shown as coloured rectangles. We see that both models are high compared to FIRAS at nearly all wavelengths. It may be interesting to note that the models appear to agree with the Lagache et al. (2000) curve slightly better than the Fixsen et al. (1998) curve.

As with the correlations between bands in the BLAST counts, we have also ignored correlations between bands in the FIRAS measurement. Proper treatment of these correlations would likely reduce the overall χ^2 of the models.

5.2.3 *Redshift distribution*

In Fig. 10, we show a variety of dN/dz measurements: 170 μm from *ISO*; 250 and 500 μm from BLAST; 850 μm from SCUBA; and 1.1 μm from AzTEC. As discussed in Section 3.4, the selection function for these redshift distributions is very poorly quantified, so we do not make full use of these counts as constraints on the model. However, to provide at least some direct redshift constraint, we use the approximated SCUBA distribution, and at the other wavelengths

simply show the counts predicted by the model compared to the data.

5.3 Goodness-of-fit

Table 1 lists the best-fit ting χ^2, along with the contribution from each data set, for the $\alpha = 1$, $\alpha = 0$ and the free α models (Columns 3, 9 and 10, respectively). The number of data points in each data set is listed in Column 2. The total χ^2 and number of degrees of freedom (DOF) for each fit are listed along the bottom two rows. Since there are 62 data points and 12 (13) free parameters, there are 50 (49) DOF for the fixed (free) α models. The reduced χ^2, $\chi^2_{\rm r} = \chi^2/N_{\rm dof}$, is unfortunately high, 3.8, 6.8 and 3.5 for the $\alpha = 1$, $\alpha = 0$ and free α models, respectively. Clearly, the $\alpha = 1$ model is a much better fit to the data than the $\alpha = 0$ model (but see Section 6.5). Adding an extra parameter for the free α model also appears to be justified; even so, we focus on the $\alpha = 1$ model, to keep things simple.

We explore the effects of re-fitting the model omitting various data sets in order to probe the 'strain' on the model due to any particular data set. We have run the $\alpha = 1$ model on all data sets: (i) excluding 250 and 500 μm counts; (ii) excluding 1100 μm counts; (iii) using 850 μm counts *instead of* 1100 μm counts; (iv) excluding the 850 μm dN/dz distribution and (v) excluding the CIB. The results of these tests are discussed here.

(i) The first test (Column 4) was meant to explore the effects of correlations between the BLAST data sets; however, it is not clear how much of the improvement in $\chi^2_{\rm r}$ is due to correlations and how much is due to the lessening of tension between BLAST and 1100 μm. The value of $\chi^2_{\rm r}$ is slightly lower, at 3.1, and we see that the 1100 μm counts and CIB are fit much better, although the dN/dz agreement is much worse.

(ii) Removing the 1100 μm counts (Column 5) greatly increases the goodness-of-fit, reducing $\chi^2_{\rm r}$ to 2.2. We see that the tension between the BLAST and 1100 μm data sets is greatly relieved, that the background is reduced at the shorter wavelengths and that dN/dz is allowed to fit nearly perfectly. We believe this is a strong clue for developing improved models, as we discuss in the next section.

(iii) Fitting the 850 μm counts instead of the 1100 μm counts (Column 6) greatly improves the fit compared to the full data set, to a reduced $\chi^2_{\rm r}$ to 2.2. This is because, compared to 1100 μm counts, the 850 μm counts are higher at the bright end and lower at the faint end, which allows better fits to the 500 and 850 μm counts. The 850 μm dN/dz distribution is allowed to fit well. We note, however, that the model significantly overpredicts the 1.4 and 2.0 mm counts (not shown here).

(iv) Removing the 850 μm dN/dz constraint (Column 7) also removes tension, in this case between BLAST and 1100 μm counts, although not to such a high degree as for (ii); here, $\chi^2_{\rm r} = 2.8$.

(v) Without the CIB constraints (Column 8), we see that the integrated background is entirely unbounded. This is because the CIB is the only available constraint on the faint end of the counts, which dominates the CIB if the faint-end counts are sufficiently steep. The quality of fit to all other data set is improved, with $\chi^2_{\rm r} = 2.6$. However, with no overall constraint at faint flux densities, the number of high-redshift sources is greatly increased. We note that this is not reflected in the 850 μm dN/dz constraint, since that data set includes only galaxies brighter than $S_{850} > 5$ mJy.

We have also run a fit of the $\alpha = 1$ model to test the hypothesis that the low-flux density BLAST nodes are biased high by the $P(D)$ parametrization, as discussed in Section 5.2.1. We have adjusted the 20/15/10 mJy nodes of the 250/350/500 μm counts to match the Glenn et al. (2010) *Herschel*/SPIRE $P(D)$ counts, and have also adjusted the values of the other nodes based on the BLAST counts

Table 1. Breakdown of χ^2 by data set for different models. The total χ^2 and number of degrees of freedom, $N_{\rm dof}$, are listed in the bottom two rows. In Columns 4–8, the $\alpha = 1$ model is re-fit excluding (or replacing) individual data sets. The χ^2 for excluded data sets are listed in parentheses.

Band (μm) data	No. Pts	$\alpha = 1$ All	No 250 and 500 counts	No 1100 counts	850 counts	No dN/dz	No CIB	$\alpha = 0$ all data	α free all data
				Counts					
70	13	12.2	14.4	13.6	13.5	14.3	10.1	15.6	14.6
160	11	6.5	13.9	7.7	8.4	8.9	5.0	71.9	8.3
250	6	10.0	(28.0)	13.8	10.7	13.1	11.2	19.5	14.0
350	5	16.7	20.5	17.3	17.8	15.9	7.7	47.5	18.5
500	5	45.3	(72.7)	13.8	25.6	31.5	28.5	81.8	48.1
850	10	(13.1)	(22.1)	(49.1)	13.7	(14.6)	(18.8)	(20.0)	(14.8)
1100	7	31.6	7.8	(916.9)	(359.0)	12.0	26.1	19.8	24.7
				Background					
160	1	4.0	2.5	1.7	2.5	1.8	(3.2×10^3)	5.1	1.9
250	1	7.2	4.0	4.0	4.7	4.3	(3.9×10^5)	7.3	4.6
350	1	9.6	3.9	7.2	6.9	6.7	(2.7×10^6)	14.2	7.7
500	1	7.2	2.0	7.8	5.8	4.6	(1.7×10^7)	20.4	6.4
850	1	1.1	0.0	2.9	1.1	0.2	(2.4×10^7)	11.4	1.1
1100	1	0.0	0.4	0.8	0.1	0.1	(3.6×10^6)	5.3	0.1
				dN/dz					
850	9	40.8	52.8	3.3	7.2	(238.5)	24.1	20.1	20.8
Total:		192.3	122.2	93.9	117.9	113.4	112.7	339.8	170.6
$N_{\rm dof}$:		50	39	43	53	41	44	50	49
$\chi^2_{\rm r}$:		3.8	3.1	2.2	2.2	2.8	2.6	6.8	3.5

covariance matrices. Additionally, we have doubled the errors on the lowest and highest flux density nodes, to compensate for the non-Gaussian error distributions (faint nodes) and cosmic variance (bright nodes). We find that, with this adjusted data set, χ^2 for the model is 80.3, with $\chi^2_r = 1.6$. This model is a much better fit to the 350, 500 and 1100 μm counts, the 850 μm redshift distribution, and CIB through the BLAST bands (although the model is still $\sim 1.5\sigma$ high). This test indicates that much of the tension in the model is due to the $P(D)$ parametrization, although this alone is not the sole cause of the over-predicted CIB.

To further test how much of this improvement is due to the 'adjustment' and how much is due to the doubled error bars, we also ran a test with the original BLAST counts, but with the errors on the lowest and highest flux density nodes doubled. For this test, we find $\chi^2 = 118$. This shows that a large fraction of the tension between data sets could be due to the apparently high data points at low flux density in the BLAST $P(D)$ counts, although underestimates of the errors at low and high flux density (possibly due to non-Gaussian error distributions and cosmic variance, respectively) may also be a significant factor.

6 DISCUSSION

We now look at some inferences that can be drawn from the model, discuss the implications of the colour–luminosity evolution degeneracy and data sets needed to resolve it, and consider other possible improvements, including the use of new data sets and modifications to the techniques.

6.1 Colour–colour distributions

In Section 2.2 it was shown that our SED library is consistent with the real spread in galaxy SEDs in the local Universe by using them to map our *IRAS*-based $\Phi(L, C)$ to the independent monochromatic luminosity functions at 12, 15, 25 and 850 μm. However, it is possible that our chosen SED shapes could conspire to produce these consistent *integral* quantities, while failing for individual galaxies. Now that we have an evolving model of the luminosity function in hand, we can go back to our best-fitting distribution, apply observational selection functions and compare the distribution of colours for *individual* galaxies from our model to those detected in real surveys (i.e. to verify the correlations between bands). For this test, we have relied on the two best examples that we could find in the literature of surveys with colour information spanning 24–850 μm: an $S_{70} > 5$ mJy *Spitzer* survey with cross-matched 24 and 160 μm observations taken as part of the Cosmological Evolution Survey (COSMOS, Kartaltepe et al. 2010); and the Submillimetre Local Universe Galaxy Survey (SLUGS), in which SCUBA was used to follow up $S_{60} > 5.24$ Jy *IRAS* galaxies at 450 and 850 μm (Dunne et al. 2000; Dunne & Eales 2001). Almost all of the SLUGS galaxies have luminosities $L > 10^{10}\,L_\odot$, with about half above $L > 10^{11}\,L_\odot$, and it is a truly local sample, with all of the objects lying at $z < 0.1$. The COSMOS sample is deeper and of higher redshift, though with a similar range of luminosities; about 50 per cent of the sources lie at $z > 0.5$, and virtually all of the objects have luminosities $L > 10^{10}\,L_\odot$, while ~70 per cent have luminosities $L > 10^{11}\,L_\odot$.

The comparison of our model to these data sets is shown in Fig. 11. The *Spitzer* COSMOS data show excellent correspondence with our model in terms of the 160-to-70 μm colour. Our theoretical distribution also broadly reproduces the main trend in the observed 24-to-70 μm correlation, although the scatter is slightly less than that observed in real galaxies. As stated in the Introduction, we have *not* attempted to fully reproduce the observed properties of 24 μm sources, since the SEDs of galaxies at those shorter wavelengths depend on other factors (which would then require more variables). We also note that since this is a higher-redshift sample, the 24 μm band will sample the PAH features for a small subset. The correspondence with the SLUGS galaxies is excellent, showing that the spread and correlation between the Rayleigh–Jeans side of our SED templates and the FIR peak is accurate.

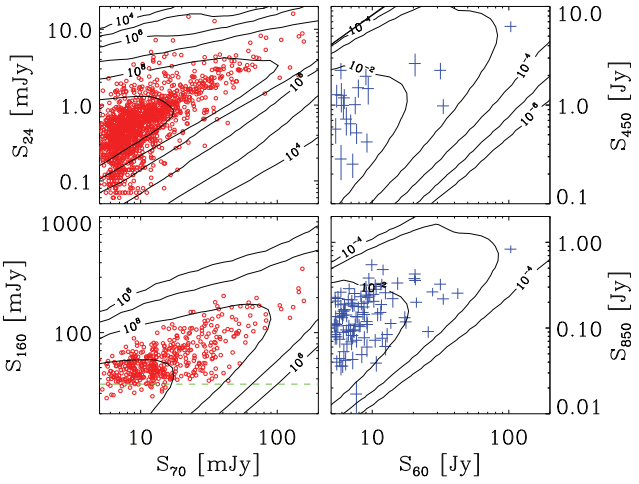

Figure 11. Comparison between the colours of galaxies drawn from our best-fitting evolving luminosity function (contours) with data from real surveys (symbols). The left-hand panels are galaxies selected at 70 μm above 5 mJy (red symbols, Kartaltepe et al. 2010). The top-left panel shows 24 μm versus 70 μm and the bottom-left panel shows 160 μm versus 70 μm. There is an approximate flux-limit of 36 mJy in the 160 μm data, which is indicated by a green dashed line. The right-hand panels compare our model with SCUBA 450 and 850 μm follow-up of 60 μm sources brighter than 5.24 Jy (blue symbols whose width and height indicate measurement errors; Dunne et al. 2000; Dunne & Eales 2001). The top-right panel shows 450 μm versus 60 μm, and the bottom right panel shows 850 μm versus 60 μm. In all panels, the contours indicate the density of sources predicted from the model, with units of $(\log \mathrm{Jy})^{-2}\,\mathrm{deg}^{-2}$. Unlike Fig. 2, which shows that our SED library and local colour–luminosity distribution can produce the correct *total numbers* of galaxies in several bands, this comparison validates our model for *individual objects*.

6.2 Star formation rate history

Star formation rate density (SFRD) can be derived from the model by calculating the luminosity density $L_{\mathrm{TIR}}(z)$:

$$L_{\mathrm{TIR}}(z) = \int_{-\infty}^{\infty} \mathrm{d}L \int_{-\infty}^{\infty} \mathrm{d}C \left[L \frac{\mathrm{d}L_{\mathrm{TIR}}}{\mathrm{d}L}(C) \right] \Phi(L, C, z), \qquad (12)$$

where L_{TIR} is the luminosity integrated over the total IR range, 8–1000 μm, and $\mathrm{d}L_{\mathrm{TIR}}/\mathrm{d}L$, which depends only on C, converts L to L_{TIR}. This can then be converted to SFRD using the simple relation given by Kennicutt (1998),

$$\mathrm{SFR}\,[\mathrm{M}_\odot\,\mathrm{yr}^{-1}] = 1.728 \times 10^{-10} L_{\mathrm{TIR}}\,[\mathrm{L}_\odot]. \qquad (13)$$

The resulting SFRD for the two models are shown in Fig. 12. The shapes are very different, with the $\alpha = 1$ model peaking at $z = 1$ and the $\alpha = 0$ peaking at $z = 2$, but with no SFRD whatsoever at $z < 1$. We believe that this is due to the data requiring a certain amount of cool-type galaxies; this can come either from intrinsically cool SEDs or from redshifting moderate-temperature SEDs so that they appear cooler. The $\alpha = 1$ model provides a higher fraction of cool galaxies with increasing redshift, while the $\alpha = 0$ model is required

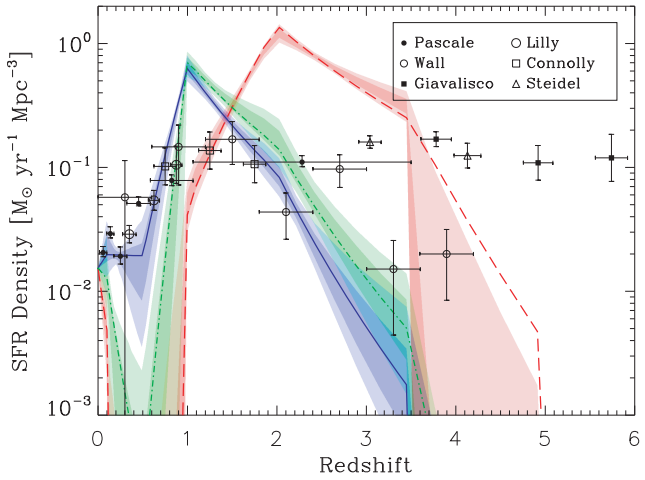

Figure 12. SFRD as a function of redshift. The $\alpha = 1$ (blue), $\alpha = 0$ (red) and free-α (green) models are shown. A sample of other estimated data points is shown; data are from Pascale et al. (2009), Wall et al. (2008), Giavalisco et al. (2004), Lilly et al. (1996), Connolly et al. (1997) and Steidel et al. (1999) (the latter four taken from Michałowski, Hjorth & Watson 2010, which has an extensive list of values in its appendix). Note that the jagged appearance in the model predictions is partly a result of the anti-correlations between adjacent redshift nodes in the fits (as seen in Fig. 5).

to compensate by strongly increasing the number of galaxies at high redshift; it is then necessary to *decrease* the number of galaxies at low-to-moderate redshift. The free α model is also shown (in green, dash–dotted line). Since the best-fitting α value is $\sim$0.6, it is not surprising that the SFRD curve falls between the $\alpha = 1$ and 0 models.

We show a small sampling of data along with the models. Below $z \sim 2$, the data are generally consistent with each other and are broadly in line with the $\alpha = 1$ model; this model in particular agrees well with the Wall, Pope & Scott (2008) compilation. Above $z \sim 2$, the data are inconsistent and it is hard to draw any firm conclusions, although we note that the Steidel et al. (1999) and Giavalisco et al. (2004) points are based on extinction-corrected UV measurements and do not necessarily bear any relation to SFRD estimates in the FIR. The $\alpha = 0$ model appears to be inconsistent with the data. We also note that the model with the 'adjusted' BLAST counts, as described at the end of Section 5.3, shows an SFRD curve that is slightly higher at $z = 0.5$, slightly lower at $z = 1$, and runs a bit flatter between $z = 2$ and 3.5 compared to the displayed $\alpha = 1$ model, fitting the Wall et al. (2008) compilation slightly better.

6.3 Fraction of CIB resolved by BLAST

If the CIB is in fact as high as indicated by the models ($\sim 2\sigma$ high compared to FIRAS at the BLAST wavelengths for the $\alpha = 1$ model; Fig. 9), the fraction of the CIB resolved by stacking the BLAST maps on 24 μm *Spitzer* sources would in fact be lower than is quoted by Marsden et al. (2009). Assuming the CIB values predicted by the $\alpha = 1$ model, $I_\nu = 1.5$, 1.3 and 0.7 MJy sr^{-1}, the stacked intensities measured by Marsden et al. (2009) correspond to 45, 45 and 50 per cent of the CIB, respectively. We note, however, that the CIB predicted by the model is likely driven high by the apparently high BLAST nodes, discussed in Section 5.2.1. Comparing to the model with the modified BLAST points, we find 55, 55 and 65 per cent, still significantly lower than the percentages quoted in Marsden et al. (2009), due to the fact that the model CIB is still $\sim 1.5\sigma$ high through the BLAST bands.

6.4 Colour–luminosity evolution

We have shown that, based on counts alone, the $\alpha = 1$ and $\alpha = 0$ models are nearly indistinguishable. The redshift distributions, however, are very different, with the $\alpha = 0$ model a better fit to the scaled Chapman et al. (2005) 850 μm redshift distribution. The SFRD figure, in particular, shows a large discrepancy between the models, with the data significantly preferring the $\alpha = 1$ model. To truly rule out one model or the other, however, direct measurements of the redshift distribution are needed.

In Fig. 13, we show the integrated 350 μm counts of sources brighter than 20 mJy in the colour–redshift plane for both models, with one-dimensional distributions shown in the side panels. The low-redshift distributions are of course exactly the same, but, as previously seen, the models are distinct at higher redshifts. We see that the $\alpha = 1$ model peaks at $C \sim -0.4$, $z \sim 1$, while the $\alpha = 0$ model peaks at $C \sim 0.1$, $z \sim 2$. A SPIRE redshift survey down to a flux limit of 20 mJy will allow us to unambiguously disentangle the two models.

When attempting to interpret the results of submm flux-limited samples, we note the strong bias towards detecting cooler, less luminous sources. Following the prescription of Chapin et al. (2011), we calculate the probability density in the colour–luminosity plane at two fixed redshifts, $z = 0.1$ and 1.0, for samples selected at 350 μm above a flux density of 20 mJy, for both the $\alpha = 1$ and $\alpha = 0$ models. The results are shown in Fig. 14. For reference, the locus of the 350 μm flux limit is shown as short-dashed lines at each redshift, and the thin solid line shows the local colour–luminosity correlation. The sample is biased downwards and to the left; this is due to the fact that the total number of sources

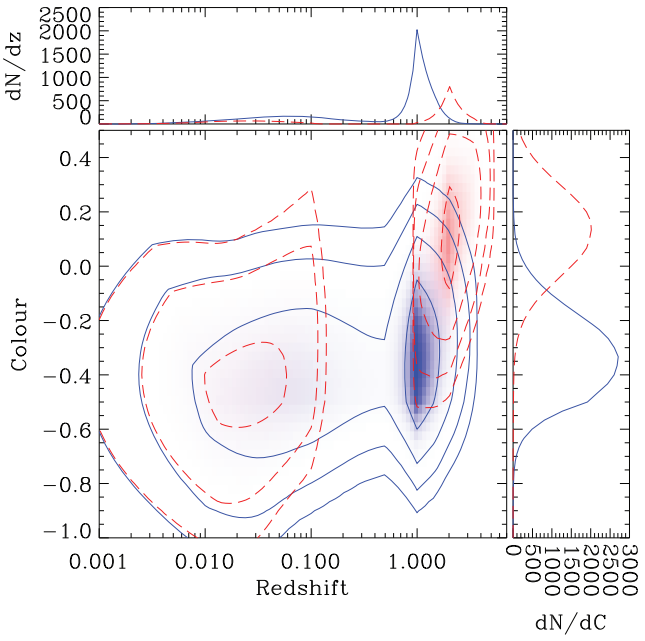

Figure 13. The density of 350 μm sources brighter than 20 mJy in the colour–redshift plane for both $\alpha = 1$ (blue, solid contours) and $\alpha = 0$ (red, long-dashed contours) models. The background intensity maps are linear and the contours are 1, 10, 100 and 1000 deg^{-2} per unit colour per unit redshift. The distributions marginalized over redshift and colour are shown at the right and top, respectively.

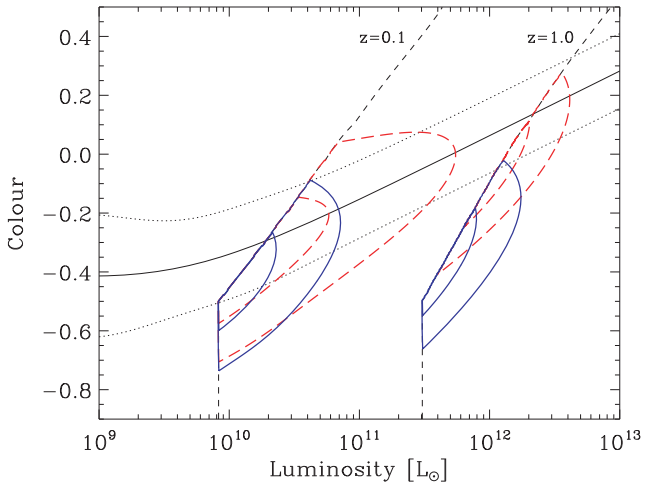

Figure 14. The density of 350 μm sources brighter than 20 mJy at a particular redshift in the colour–luminosity plane for both $\alpha = 1$ (blue, solid contours) and $\alpha = 0$ (red, long-dashed contours) models. We show the population at two redshifts, $z = 0.1$ and 1.0. The contours correspond to 0.1 and 0.5 of the maximum density. The thin short-dashed lines show the flux limits at the two redshifts. These lines are independent of luminosity below $C = -0.5$ due to our choice of SED library. The solid black line shows the local colour–luminosity relationship, $p(C|L)$ (see Section 2.1), with 1σ limits plotted as dotted lines. This figure demonstrates the strong bias towards cooler and lower-luminosity sources when compared to the rest-frame distribution, caused by our selection function.

increases toward lower luminosities, and that the selection function is *inclined* in the colour–luminosity plane; cooler sources may be detected at fainter luminosities, where there are considerably more objects. Chapin et al. (2011) used this argument to suggest that their submm-selected sample could be consistent with non-evolution of the colour–luminosity correlation (despite finding a tail of cool sources below the local colour–luminosity correlation). Similarly, included among the Science Demonstration Phase results from *Herschel* are some tentative indications that submm-selected sources are cooler in the past (e.g. Elbaz et al. 2010; Hwang et al. 2010); it will be necessary to model this selection effect to determine whether these observed trends are indeed real.

Of course, another possibility is that the colour–luminosity correlation evolves in a way completely different from the simple form assumed in this paper. For example, Symeonidis, Page & Seymour (2010) argue that the correlation simply broadens at high redshift, based on a sample of galaxies spanning the submm–FIR, rather than there being a trend towards lower temperatures in the past. While the observed correlation at high redshift may in fact be broader than the local correlation, we note that the combination of a tight colour–luminosity correlation with a submm selection effect can also broaden the apparent range of colours, as shown in Fig. 13, and certainly contributes, at least in part, to the effect seen by Symeonidis et al. (2010).

Based on the forward-modelling approach taken in this paper, which intrinsically accounts for these selection biases, we find that an evolution towards cooler temperatures in the past is the most likely scenario (i.e. the tail of cooler sources in the high-redshift submm-selected population is *even cooler* than what one might expect given the selection biases). This conclusion is in rough agreement with the model of Valiante et al. (2009), who also included luminosity evolution in the colour–luminosity correlation to produce cooler SEDs in the past, although they also included an extra population of cold local sources that are presumably missed in *IRAS* surveys. In contrast, Lewis et al. (2005) and Le Borgne et al. (2009) used fixed colour–luminosity correlations as a function of redshift, and also appear to obtain reasonable results. We note, however, that our result is based strongly on the imposed colour distribution at the faint end of the luminosity function, which we believe is possibly the cause of the over-predicted CIB. A modified colour–luminosity relationship could have a large effect on the relative merits of the $\alpha = 1$ and $\alpha = 0$ models. This is discussed further in the next section.

As already discussed, the nature of the evolution of the colour–luminosity correlation has a direct impact on our inferences about the total SFRD as a function of time. As shown in Fig. 12, a decrease in the typical temperature of galaxies at high redshift leads to a later peak formation epoch when compared with a non-evolving colour–luminosity scenario. Working from primarily SCUBA-selected samples, there is some evidence that FIR-luminous galaxies at high-redshift are indeed cooler, but also physically more extended, based on radio morphologies (e.g. Chapman et al. 2004), near–mid-IR colours (e.g. Hainline et al. 2009) and mid-IR spectra (e.g. Menéndez-Delmestre et al. 2009). These observations appear to be consistent with a local-Universe measurement that showed an anti-correlation between physical size as a function of luminosity and temperature (Chanial et al. 2007).

6.5 Reconciling the CIB with the SMG $z > 1$ redshift distribution

As has been seen in previous sections, the single-population model that we have attempted to fit fails, primarily, in reconciling the redshift distribution of ∼1-mm-selected SMGs (and therefore the $z > 1$ SFRD; Figs 10 and 12), with the intensity of the CIB (Fig. 9), despite doing a decent job of fitting the counts in the bands that we have considered (Figs 6 and 7). Generally speaking, we have found that models with lower (and therefore more consistent) values of the CIB push the ∼1 mm selected galaxies to lower (and less consistent) redshifts.

We believe most of the discrepancy is due to our lack of knowledge of the low-luminosity ($L < 10^{10}\,L_\odot$) galaxy SEDs. These faint galaxies are essentially undetected individually above the confusion limit in any of the existing blank-field ground-based ∼1 mm surveys (Blain et al. 2002), nor in any BLAST 250–500 μm (Dye et al. 2009; Chapin et al. 2011), or *Spitzer* 70 and 160 μm surveys (see e.g. the description of the COSMOS survey in Section 6.1). Even *IRAS* struggled to conduct unbiased surveys of such faint objects, since it was limited in sensitivity and also the nearby galaxies in the local over-density were limited by sampling variance (see Section 2.1). However, the integrated light from these faint but numerous galaxies *can* make a large contribution to the CIB.

Referring to fig. 4 of C09a, it is clear that the clean correlation between C and L broadens significantly between $10^{10}\,L_\odot$ and $10^{9}\,L_\odot$, and then it completely breaks down at even lower luminosities. Much of this effect is probably due to the fact that the 60/100 μm flux ratio is simply a poor indicator of the longer-wavelength SED of galaxies with extremely cool dust (i.e. these bands sample relatively warmer and potentially un-related dust). It is probably for this reason that other authors have found it necessary to modify by hand the distribution of low-luminosity/cool galaxies in their evolutionary models.

In addition, the faint end of the FIR luminosity function contains a significantly more heterogeneous collection of galaxies than the bright end. Most luminous infrared galaxies ($L > 10^{11}\,L_\odot$)

and virtually all ultra-luminous infrared galaxies ($L > 10^{12}\,L_{\odot}$) are merger-driven starbursts (Sanders & Mirabel 1996). In contrast, the faint end must include everything else, including mildly star-forming spirals like the Milky Way (relatively low luminosity despite the presence of dust), passively evolving spheroids (extremely faint due to the near complete lack of dust), but also smaller star-forming galaxies; blue compact dwarfs, for example, have dust and contain a number of active star-forming regions which may produce apparently large dust temperatures despite their lower luminosities.

To test the hypothesis that our lack of knowledge at these low luminosities plays a crucial role, we performed a simple test using our best-fitting $\alpha = 1$ model. First, we assigned all galaxies below $10^{10}\,L_{\odot}$ the *warmest* SED (largest value of C). This has the effect of *decreasing* the flux densities at $\lambda > 100\,\mu$m, and *increasing* the flux densities at $\lambda < 60\,\mu$m when extrapolating from the FIR luminosity for those galaxies. We then re-calculated the number counts, CIB spectrum and redshift distribution of 850 μm selected galaxies. While the modification to the SEDs of these faint galaxies had little impact on either the counts or redshift distribution, where data are available (again, since galaxies with such luminosities lie below current confusion limits), there was a *huge* change to the CIB. Its predicted spectrum became much warmer (peaking closer to 100 μm than 200 μm), and dropped significantly below the data at submm wavelengths. Conversely, assigning these low-luminosity galaxies the coolest SED has the opposite effect, pushing the predicted CIB peak to longer wavelengths, and exceeding the CIB by even more than with the regular model.

We then repeated this procedure using a lower luminosity threshold of $10^{9}\,L_{\odot}$. While the sense of the changes to the predicted quantities were the same, the impact is obviously smaller, and using the warmest SEDs for these faint galaxies, we obtain a reasonable fit to the CIB (slightly low at 200 μm, and then going through the data at longer submm wavelengths).

Based on this experiment, it is clear that how one treats galaxies at the faint-end of the FIR luminosity function is crucial, even although they are not well constrained observationally. While one could modify a phenomenological model *ad hoc* to evolve the faint-end significantly less than the bright-end as in Lagache et al. (2003) and Valiante et al. (2009), thus reducing the number of cool-galaxies at higher redshifts, the alternative of increasing the temperatures of fainter galaxies at *all* redshifts would have a similar effect. In the future one could use wide-area surveys such as H-ATLAS (Eales et al. 2010) to measure the SEDs of fainter nearby galaxies, and perhaps also with SCUBA-2 at 450 μm (since it will be capable of resolving most of the CIB directly into individual sources; Holland et al. (2006)). We also note that since our present model tends to pull down the SMG redshift distribution in order to improve the χ^2 of the CIB, our result that the $\alpha = 1$ model is better than the higher-redshift $\alpha = 0$ model is far from secure. Soon it should be able to perform tests such as Figs 13–14 to determine, at least for the more luminous objects, whether there really is evolution in the luminosity–temperature correlation using SPIRE data, particularly from HerMES.

6.6 Future improvements

Although in many ways our model is an improvement over previously published studies, throughout the paper we have discussed several shortcomings of the model and data sets used. Here we list some analysis techniques and future data sets that will improve the quality of the model.

(i) As discussed in Section 5.2.1, the BLAST $P(D)$ counts are inherently correlated. Instead of using the counts as an intermediate data set, we can use $P(D)$ *within* the model-fitting framework to fit the maps directly. This is of course computationally intensive, but certainly worth pursuing. It also has the advantage over the P09-style $P(D)$ in that the shape of the model is much more reflective of the shape of the true underlying counts, rather than imposing a connected power law on to the counts. Furthermore, as we noted in Section 3.4.1, the discrepancy between SCUBA 850 μm and AzTEC 1.1 mm counts may be due to the potentially biased counting of sources. $P(D)$ analyses do not suffer most of those problems.

(ii) New redshift distributions from SPIRE will go a long way in constraining the model, in particular breaking the colour–luminosity evolution degeneracy, and enabling more redshift nodes to be used.

(iii) Direct measurements of the local luminosity functions at FIR and submm wavelengths will give us a much better starting point for the evolving luminosity function, and will allow us to derive a better SED library. Wide-area PACS and SPIRE surveys will be able to provide these. We also hope that wide-area surveys will be capable of detecting and measuring the SEDs of galaxies with FIR luminosities in the range 10^{10}–$10^{11}\,L_{\odot}$ to solve the discrepancy between the redshift distribution of SMGs and the spectrum of the CIB.

(iv) Extend the model to the mid-IR by including a more sophisticated SED library. This may require additional SED parameters to include, for example, the AGN contribution (e.g. Valiante et al. 2009).

(v) Include lensing by adopting a similar approach to Paciga, Scott & Chapin (2009). Such a treatment is now necessary to explain the counts at mm wavelengths covering wide areas, and more recently, *Herschel*/SPIRE surveys.

(vi) More versatile modelling of the evolving SED distribution, in particular changing the width and shape of $P(C|L)$ as a function of L.

Finally, we note that a major part of the work in this paper went into developing likelihood expressions for the various data sets. To improve on our methodology for current and future surveys, it will be important to fully characterize uncertainties and correlations between data sets. For example, the differential number counts, integrated background and redshift distributions of sources in the same field will have a correlated cosmic variance term.

7 CONCLUSIONS

We have presented a sophisticated technique using MCMC to fit a simple evolving luminosity function to a range of FIR and submm data. We are able to measure errors on the evolutionary parameters and show the correlations between these parameters. The results of the model are available at http://cmbr.phas.ubc.ca/model/. We also show how the various data sets are in tension with one another and demonstrate the importance of redshift distributions.

An advantage of our approach over some other models in the literature is that we need only consider the evolution of one galaxy population with a single-parameter family of SEDs based on the correlation between the 60-to-100 μm rest-frame colour and FIR luminosity. While we find that, across the 70–1100 μm wavelength range, the counts can be fit using models with and without evolution towards cooler galaxy dust temperatures at higher-redshifts, there is significant tension between the spectrum of the CIB and the redshift distribution of SMGs. We believe that most of this discrepancy is caused by the presently unknown distribution of submm SEDs for

galaxies with luminosities $<10^{10}\,L_{\odot}$. Emerging data from *Herschel* will immediately help in two areas: (i) wide-area surveys, such as H-ATLAS, should enable us to measure the SEDs for at least a small sample of such objects; (ii) deeper surveys, such as HerMES, can potentially be used to search for evolution in the colour–luminosity correlation at the brighter end of the luminosity function, as indicated in Figs 13–14. Such data should obviate the need for *ad hoc* modifications to the low-luminosity region of the local luminosity function by, for example, using multiple uncorrelated galaxy populations. The reality is that there is a continuum of galaxy types, and it is our hope that we can gain more realistic insight into how they form and evolve using the simplest phenomenological models that are consistent with the data.

ACKNOWLEDGMENTS

We acknowledge the support of NASA through grant numbers NAG5-12785, NAG5-13301, and NNGO-6GI11G, the NSF Office of Polar Programs, the Canadian Space Agency, the Natural Sciences and Engineering Research Council (NSERC) of Canada, and the UK Science and Technology Facilities Council (STFC). We thank the authors of CosmoMC for providing an easy-to-use MCMC generic sampler. Thanks also to Matthieu Béthermin for useful discussions.

REFERENCES

Aretxaga I., Hughes D. H., Chapin E. L., Gaztañaga E., Dunlop J. S., Ivison R. J., 2003, MNRAS, 342, 759
Austermann J. E., Dunlop J. S., Perera T. A., Scott K. S., Wilson G. W., Aretxaga I., Hughes D. H., Almaini O., 2010, MNRAS, 401, 160
Baugh C. M., Lacey C. G., Frenk C. S., Granato G. L., Silva L., Bressan A., Benson A. J., Cole S., 2005, MNRAS, 356, 1191
Béthermin M., Dole H., Beelen A., Aussel H., 2010, A&A, 512, A78
Béthermin M., Dole H., Lagache G., Le Borgne D., Penin A., 2011, A&A, 529, A4
Blain A. W., Longair M. S., 1993, MNRAS, 264, 509
Blain A. W., Smail I., Ivison R. J., Kneib J.-P., 1999, MNRAS, 302, 632
Blain A. W., Smail I., Ivison R. J., Kneib J.-P., Frayer D. T., 2002, Phys. Rev., 369, 111
Blain A. W., Barnard V. E., Chapman S. C., 2003, MNRAS, 338, 733
Capak P., Carilli C. L., Lee N., Aldcroft T., Aussel H., Schinnerer E., Wilson G. W., Yun M. S., 2008, ApJ, 681, L53
Chanial P., Flores H., Guiderdoni B., Elbaz D., Hammer F., Vigroux L., 2007, A&A, 462, 81
Chapin E. L., Hughes D. H., Aretxaga I., 2009a, MNRAS, 393, 653 (C09a)
Chapin E. L. et al., 2009b, MNRAS, 398, 1793
Chapin E. L., Chapman S. C., Coppin K. E., Devlin M. J., Dunlop J. S., Greve T. R., Halpern M., Hasselfied M. F., 2011, MNRAS, 411, 505
Chapman S. C., Blain A. W., Ivison R. J., Smail I. R., 2003a, Nat, 422, 695
Chapman S. C., Helou G., Lewis G. F., Dale D. A., 2003b, ApJ, 588, 186
Chapman S. C., Smail I., Windhorst R., Muxlow T., Ivison R. J., 2004, ApJ, 611, 732
Chapman S. C., Blain A. W., Smail I., Ivison R. J., 2005, ApJ, 622, 772
Chary R., Elbaz D., 2001, ApJ, 556, 562
Connolly A. J., Szalay A. S., Dickinson M., Subbarao M. U., Brunner R. J., 1997, ApJ, 486, L11
Coppin K. et al., 2006, MNRAS, 372, 1621
Coppin K. E. K., Smail I., Alexander D. M., Weiss A., Walter F., Swinbank A. M., Greve T. R., Kovacs A., 2009, MNRAS, 395, 1905
Cowie L. L., Songaila A., Hu E. M., Cohen J. G., 1996, AJ, 112, 839
Daddi E. et al., 2009, ApJ, 694, 1517
Dale D. A., Helou G., Contursi A., Silbermann N. A., Kolhatkar S., 2001, ApJ, 549, 215
Dennefeld M., Lagache G., Mei S., Ciliegi P., Dole H., Mann R. G., Taylor E. L., Vaccari M., 2005, A&A, 440, 5
Dole H., Lagache G., Puget J.-L., 2003, ApJ, 585, 617
Dole H. et al., 2006, A&A, 451, 417
Downes T. P., Welch D. E., Scott K., Austermann J., Wilson G. W., 2011, preprint (arXiv e-prints)
Draine B. T., Li A., 2001, ApJ, 551, 807
Draine B. T., Li A., 2007, ApJ, 657, 810
Dunlop J. S., Ade P. A. R., Bock J. J., Chapin E. L., Cirasuolo M., Coppin K. E. K., Devlin M. J., Griffin M., 2010, MNRAS, 408, 2022
Dunne L., Eales S. A., 2001, MNRAS, 327, 697
Dunne L., Eales S., Edmunds M., Ivison R., Alexander P., Clements D. L., 2000, MNRAS, 315, 115
Dye S., Ade P. A. R., Bock J. J., Chapin E. L., Devlin M. J., Dunlop J. S., Eales S. A., Griffin M., 2009, ApJ, 703, 285
Eales S. et al., 2010, PASP, 122, 499
Elbaz D., Hwang H. S., Magnelli B., Daddi E., Aussel H., Altieri B., Amblard A., Andreani P., 2010, A&A, 518, L29
Fang F., Shupe D. L., Xu C., Hacking P. B., 1998, ApJ, 500, 693
Fixsen D. J., Dwek E., Mather J. C., Bennett C. L., Shafer R. A., 1998, ApJ, 508, 123
Giavalisco M. et al., 2004, ApJ, 600, L103
Glenn J., Conley A., Bethermin M., Altieri B., Amblard A., Arumugam V., Aussel H., Babbedge T., 2010, MNRAS, 409, 109
Guiderdoni B., Bouchet F. R., Puget J., Lagache G., Hivon E., 1997, Nat, 390, 257
Hainline L. J., Blain A. W., Smail I., Frayer D. T., Chapman S. C., Ivison R. J., Alexander D. M., 2009, ApJ, 699, 1610
Holland W. et al., 2006, in Zmuidzinas J., Holland W. S., Withington S., Duncan W. D., eds. Millimeter and Submillimeter Detectors and Instrumentation for Astronomy III. Proc. SPIE, Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conf. Vol. 6275. SPIE, Bellingham, p. 62751E
Hwang H. S., Elbaz D., Magdis G. E., Daddi E., Symeonidis M., Altieri B., Amblard A., 2010, MNRAS, 409, 75
Ivison R. J., Magnelli B., Ibar E., Andreani P., Elbaz D., Altieri B., Amblard A., Arumugam V., 2010a, A&A, 518, L31
Ivison R. J., Alexander D. M., Biggs A. D., Brandt W. N., Chapin E. L., Coppin K. E. K., Devlin M. J., Dickinson M., 2010b, MNRAS, 402, 245
Kartaltepe J. S. et al., 2010, ApJ, 709, 572
Kelly B. C., Fan X., Vestergaard M., 2008, ApJ, 682, 874
Kennicutt R. C., Jr, 1998, ARA&A, 36, 189
Komatsu E., Smith K. M., Dunkley J., Bennett C. L., Gold B., Hinshaw G., Jarosik N., Larson D., 2011, ApJS, 192, 11
Lagache G., Haffner L. M., Reynolds R. J., Tufte S. L., 2000, A&A, 354, 247
Lagache G., Dole H., Puget J.-L., 2003, MNRAS, 338, 555
Lagache G. et al., 2004, ApJS, 154, 112
Lawrence A. et al., 1999, MNRAS, 308, 897
Le Borgne D., Elbaz D., Ocvirk P., Pichon C., 2009, A&A, 504, 727
Lewis A., Bridle S., 2002, Phys. Rev. D, 66, 103511
Lewis G. F., Chapman S. C., Helou G., 2005, ApJ, 621, 32
Lilly S. J., Le Fevre O., Hammer F., Crampton D., 1996, ApJ, 460, L1
Malkan M. A., Stecker F. W., 2001, ApJ, 555, 641
Marsden G., Ade P. A. R., Bock J. J., Chapin E. L., Devlin M. J., Dicker S. R., Griffin M., Gundersen J. O., 2009, ApJ, 707, 1729
Menéndez-Delmestre K. et al., 2009, ApJ, 699, 667
Michałowski M., Hjorth J., Watson D., 2010, A&A, 514, A67
Narayanan D., Hayward C. C., Cox T. J., Hernquist L., Jonsson P., Younger J. D., Groves B., 2010, MNRAS, 401, 1613
Paciga G., Scott D., Chapin E. L., 2009, MNRAS, 395, 1153
Pascale E., Ade P. A. R., Bock J. J., Chapin E. L., Devlin M. J., Dye S., Eales S. A., Griffin M., 2009, ApJ, 707, 1740
Patanchon G., Ade P. A. R., Bock J. J., Chapin E. L., Devlin M. J., Dicker S. R., Griffin M., Gundersen J. O., 2009, ApJ, 707, 1750 (P09)
Patris J., Dennefeld M., Lagache G., Dole H., 2003, A&A, 412, 349
Perera T. A. et al., 2008, MNRAS, 391, 1227
Pope E. A., 2007, PhD thesis, Univ. British Columbia

Puget J.-L., Abergel A., Bernard J.-P., Boulanger F., Burton W. B., Desert F.-X., Hartmann D., 1996, A&A, 308, L5
Rowan-Robinson M., 2001, ApJ, 549, 745
Rowan-Robinson M., 2009, MNRAS, 394, 117
Sanders D. B., Mirabel I. F., 1996, ARA&A, 34, 749
Saunders W., Rowan-Robinson M., Lawrence A., Efstathiou G., Kaiser N., Ellis R. S., Frenk C. S., 1990, MNRAS, 242, 318
Schinnerer E. et al., 2008, ApJ, 689, L5
Schmidt M., 1968, ApJ, 151, 393
Serjeant S., Harrison D., 2005, MNRAS, 356, 192
Serjeant S., Dye S., Mortier A., Peacock J., Egami E., Cirasuolo M., Rieke G., Borys C., 2008, MNRAS, 386, 1907
Shupe D. L., Fang F., Hacking P. B., Huchra J. P., 1998, ApJ, 501, 597
Soifer B. T., Neugebauer G., 1991, AJ, 101, 354
Steidel C. C., Adelberger K. L., Giavalisco M., Dickinson M., Pettini M., 1999, ApJ, 519, 1
Symeonidis M., Page M. J., Seymour N., 2010, preprint (arXiv e-prints)
Taylor E. L., Mann R. G., Efstathiou A. N., Babbedge T. S. R., Rowan-Robinson M., Lagache G., Lawrence A., Mei S., 2005, MNRAS, 361, 1352
Truch M. D. P., Ade P. A. R., Bock J. J., Chapin E. L., Devlin M. J., Dicker S. R., Griffin M., Gundersen J. O., 2009, ApJ, 707, 1723
Valiante E., Lutz D., Sturm E., Genzel R., Chapin E. L., 2009, ApJ, 701, 1814
Vieira J. D., Crawford T. M., Switzer E. R., Ade P. A. R., Aird K. A., Ashby M. L. N., Benson B. A., Bleem L. E., 2010, ApJ, 719, 763
Wall J. V., Pope A., Scott D., 2008, MNRAS, 383, 435
Weiß A. et al., 2009, ApJ, 707, 1201
Xu C., 2000, ApJ, 541, 134

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