Mon. Not. R. Astron. Soc. **427,** 816–827 (2012) doi:10.1111/j.1365-2966.2012.22047.x

Automated measurement of redshifts from mid-infrared low-resolution spectroscopy

Antonio Hernán-Caballero★

Instituto de Física de Cantabria, CSIC-UC, Avenida de los Castros s/n, 39005 Santander, Spain

Accepted 2012 September 4. Received 2012 September 4; in original form 2012 June 26

ABSTRACT

Obtaining accurate redshifts from mid-infrared (MIR) low-resolution ($R \sim 100$) spectroscopy is challenging because the wavelength resolution is too low to detect narrow lines in most cases. Yet, the number of degrees of freedom and diversity of spectral features are too high for regular spectral energy distribution (SED) fitting techniques to be convenient. Here we present a new SED-fitting-based routine for redshift determination that is optimized for MIR low-resolution spectroscopy. Its flexible template scaling increases the sensitivity to slope changes and small-scale features in the spectrum, while a new selection algorithm called maximum combined pseudo-likelihood (MCPL) provides increased accuracy and a lower number of outliers compared to the standard maximum-likelihood (ML) approach. Unlike ML approach, the MCPL approach searches for local (instead of absolute) maxima of a 'pseudo-likelihood' (PL) function, and combines results obtained for all the templates in the library to weed out spurious redshift solutions. The capabilities of the MCPL approach are demonstrated by comparing its redshift estimates to those of the regular ML approach and to the optical spectroscopic redshifts of a sample of 491 *Spitzer*/Infrared Spectrograph spectra from extragalactic sources at $0 < z < 3.7$. The MCPL approach achieves a redshift accuracy $\Delta(z)/(1+z) < 0.005$ for 78 per cent of the galaxies in the sample compared to 68 per cent for the ML approach. The rate of outliers [$\Delta(z)/(1+z) > 0.02$] is 14 per cent for the MCPL approach and 22 per cent for the ML approach. χ^2 values for ML solutions are found to correlate with the signal-to-noise ratio of the spectra, but not with redshift accuracy. By contrast, the peak value of the normalized combined PL (γ) is found to provide a good indication on the reliability of the MCPL solution for individual sources. The accuracy and reliability of the redshifts depend strongly on the MIR SED. Sources with significant polycyclic aromatic hydrocarbon emission obtain much better results compared to sources dominated by active galactic nucleus continua. Nevertheless, for a given γ the frequency of accurate solutions and outliers is largely independent of their SED type. This reliability indicator for MCPL solutions allows to select subsamples with highly reliable redshifts. In particular, a $\gamma > 0.15$ threshold retains 79 per cent of the sources with $\Delta(z)/(1+z) < 0.005$ while reducing the outlier rate to 3.8 per cent.

Key words: methods: data analysis – catalogues – galaxies: distances and redshifts – infrared: galaxies.

1 INTRODUCTION

Finding a galaxy's redshift typically requires the identification of narrow emission or absorption lines in a medium- or high-resolution spectrum (spectroscopic redshift). Alternatively, broad features of the spectral energy distribution (SED) are revealed by multiwavelength photometry, and can be used to obtain photometric redshifts. Spectroscopic redshifts are accurate but very time-consuming, while photometric redshifts offer limited accuracy [typically in the $\Delta z/(1+z) \sim 0.01$–0.1 range] and suffer from colour degeneracies that may lead to catastrophic errors (Fernández-Soto, Lanzetta & Yahil 1999; Benítez 2000).

Halfway between the two is low-resolution spectroscopy (LRS), which at $R \sim 50$–100 provides a compromise between sensitivity and spectral resolution. LRS has become common in space-based infrared (IR) missions, particularly *Spitzer* with the Infrared Spectrograph (IRS; Houck et al. 2004), and later *AKARI* with its

★E-mail: ahernan@ifca.unican.es

Near-infrared Camera/spectrometer (IRC) (Murakami et al. 2007; Onaka et al. 2007).

Future IR missions will also provide LRS capabilities, including the Mid-InfraRed Instrument (MIRI) onboard *JWST* (Wright et al. 2008) and two instruments for *SPICA*: the Mid-InfRAred Camera with/without LEns (Wada & Kataza 2010) and the SpicA FAR-infrared Instrument (Goicoechea et al. 2011).

Since the spectrum is spread over a smaller number of pixels, the mid-IR (MIR) 5–35 μm LRS offers higher continuum sensitivity with a spectral resolution still capable of resolving many features used for diagnostics, like the polycyclic aromatic hydrocarbon (PAH) bands and absorption bands from silicates, water ice and carbon monoxide, among others.

The LRS has also the potential to yield redshifts with accuracies intermediate between those of medium-resolution spectroscopy and photometric redshifts, since the theoretical redshift resolution is proportional to the wavelength resolution: $\Delta z/(1+z) \sim \Delta\lambda/\lambda$. Nevertheless, narrow spectral lines are marginally unresolved at LRS resolutions, and this outweighs for them the sensitivity advantage over higher resolution spectroscopy, since the lines get washed out by the continuum emission. Because of this dilution, narrow lines are clearly detected only in high signal-to-noise ratio (S/N) spectra or in sources with large equivalent width (EW). Therefore, fine structure lines observed at the LRS are not suitable for spectroscopic redshift determination in the general case.

In MIR spectra, the PAH and silicate bands are routinely used to estimate the redshifts of optically faint sources (e.g. Houck et al. 2007; Yan et al. 2007; Farrah et al. 2009; Weedman & Howick 2009). However, since they often show complex morphologies, it requires a visual inspection to properly identify them, particularly in sources with high obscuration or where a strong active galactic nucleus (AGN) continuum reduces the contrast of the features. In addition, spectra with low S/N make it difficult to identify individual features even for the trained eye.

When spectroscopic redshifts are not workable, the backup strategy is photometric redshifts. The multiple photometric redshift techniques developed can be classified in two main groups: those based on 'learning' with a large training set (e.g. Connolly et al. 1995; Brunet et al. 1997; Wang, Bahcall & Turner 1998; Collister & Lahav 2004; Wadadekar 2005; Carliles et al. 2010) and those based on 'SED fitting' with a set of spectral templates (e.g. Baum 1962; Koo 1985; Gwyn & Hartwick 1996; Lanzetta, Yahil & Fernandez-Soto 1996; Sawicki, Lin & Yee 1997; Fernández-Soto et al. 1999; Benítez 2000; Bolzonella, Miralles & Pelló 2000; Le Borgne & Rocca-Volmerange 2002; Babbedge et al. 2004; Feldmann et al. 2006).

SED fitting works well with broad-band photometry in the optical and near-IR (NIR) because the SED of galaxies in this range shows little diversity. In normal galaxies, the SED is dominated by starlight, and it can be successfully modelled by the combination of a few stellar populations obscured by a screen of dust (e.g. Bruzual & Charlot 1993, 2003; Silva et al. 1998) or compared to a small set of empirical (e.g. Assef et al. 2008) or semi-empirical (e.g. Coleman, Wu & Weedman 1980; Ilbert et al. 2006) templates.

Even if the galaxy hosts a low-luminosity AGN or an obscured AGN of any luminosity, it has little impact on the broad-band SED of the galaxy. Only quasars produce continuum emission strong enough to dominate the optical–NIR SED of the galaxy, and they become a hassle for photometric redshift routines (Hatziminaoglou, Mathez & Pelló 2000; Richards et al. 2001). This is because their power-law SED does not provide high-contrast features, and the broad emission lines require good coverage with narrow- or intermediate-band filters to be identified (Benítez et al. 2009; Abramo et al. 2012; Matute et al. 2012).

In the MIR, SED fitting is far more problematic because the output of galaxies arises from several independent processes, including the Rayleigh–Jeans tail of stellar emission, thermal emission from hot and warm dust heated by the (active) nucleus, fluorescence of PAH molecules, radiative transitions of ionized and neutral atoms, rotational and vibrational transitions of H_2 and other molecules, and non-thermal radiation from radio sources (AGNs, supernovae, masers).

Furthermore, there is high dispersion in the correlation between emission from the stellar component and the interstellar medium, and population synthesis codes do not yet reproduce accurately spectral features at wavelengths $\lambda \gtrsim 5$ μm, in particular PAH emission (Brodwin et al. 2006). In practice, adding photometric points at wavelengths $\lambda \gtrsim 5$ μm to an optical–NIR SED does not improve the accuracy of the redshift solution. Nevertheless, Rowan-Robinson et al. (2008) successfully apply a two-step method to fit photometry longwards and shortwards of 5 μm with two separate sets of templates, and Negrello et al. (2009) obtain $\Delta(z)/(1+z) < 0.1$ for 90 per cent of sources in the range $0.5 < z < 1.5$ using a combination of *ISO*, *Spitzer* and *AKARI* photometry in the 3.6–24 μm range.

Template fitting can produce very accurate redshift estimates with the MIR LRS if one important issue is addressed. Because of the diversity of MIR SEDs and the large number of data points in the spectrum (compared to photometric SEDs), it cannot be expected that every source in a survey will find an accurate model of its MIR SED in the template library. As a consequence, the standard approach of SED fitting photometric redshifts (i.e. shifting and scaling of the template, and a χ^2 minimization to find the best fit) needs to be modified. This is because it favours the templates that best reproduce the overall shape of the continuum even if the smaller scale features (the ones capable of producing an accurate redshift) are poorly fitted or misplaced.

A good match at a certain redshift between small-scale spectral features of the spectrum and a template will be signalled by a sudden drop in the value of the χ^2 statistics relative to values for similar redshifts. This may not be the absolute minimum in χ^2 if the shape of the continuum is somewhat different for the spectrum and template, and there may be more than one such dip if one or more features produce partial matches by chance.

In addition, it is likely that several templates have at least some features in common with the spectrum. Each of them will produce a drop in χ^2 at the actual redshift of the source, while spurious alignments can occur at different redshifts for each template. Thus, filtering the redshift values at which different templates obtain local minima of χ^2, and then combining them in a way that favours strong dips as well as frequently repeated redshifts, allows us to obtain a redshift solution that is much more robust than finding the absolute minimum of χ^2 for any template.

In this work, a routine for redshift estimation from the MIR LRS based on these principles is presented, and its capabilities demonstrated using a large sample of extragalactic sources with both optical and *Spitzer*/IRS spectroscopy.

The outline of this paper is as follows. Section 2 presents the algorithm for redshift estimation and explains the features that depart from regular χ^2 minimization routines. Section 3 describes the sample selection and Section 4 the template library. Section 5 evaluates the accuracy and reliability of the redshift estimates obtained and their dependency on the MIR type of the source. Section 6 briefly summarizes the main conclusions of this paper.

2 THE METHOD

The redshift determination algorithm described here is implemented by zCOLORS (Redshift COde for LOw-Resolution Spectroscopy), developed by the author. It obtains the redshift of a source as well as an estimate of its reliability by comparing its MIR spectrum (hereafter, the spectrum) with a set of spectral templates (hereafter, the templates).

Let $F(\lambda)$ be the flux density of the spectrum as a function of the observed wavelength λ, and $S_k(\lambda')$ the flux density of the kth template as a function of the rest-frame wavelength λ'.

Both the spectrum $F(\lambda)$ and templates $S_k(\lambda')$ are resampled to a common grid of wavelength values $\lambda_i = \lambda_0 e^{\beta i}$. The parameter β determines the spectral resolution $R = 1/\beta$ of the resampled spectrum and templates. The uniform spacing in $\log\lambda$ is convenient for computational efficiency reasons, since such sampling ensures that the set of redshift values $z_j = e^{\beta j} - 1$, also evenly spaced in $\log(1 + z)$, verify $\lambda_{i+j} = \lambda_i(1 + z_j)$ for every $\{i, j\}$. This allows us to obtain the observed frame templates for redshift z_j, sampled at the same wavelengths as $F(\lambda)$, by just shifting one position the indices of the templates for z_{j-1}, with no need for a new resampling or interpolation.

For every redshift z_j and template k, the routine performs a least-squares fit in which the template flux is scaled to fit the spectrum. Since the overall continuum slope provides little information on the redshift of the source, the flexibility of the fit is increased by allowing for a wavelength-dependent scaling factor, that is, the spectrum $F(\lambda)$ is fitted to a function of the form

$$f_k(\lambda, z_j) = [a_k(z_j) + b_k(z_j)\log\lambda] S_k\left(\frac{\lambda}{1 + z_j}\right), \quad (1)$$

where a and b are free parameters.

This flexible scaling helps the templates obtain better fits even if their continuum slope is somewhat different from that of the spectrum. As a consequence, the fit becomes more sensitive to small-scale features in the spectrum.

In a standard template fitting, the goodness of fit is quantified by the reduced χ^2 statistics:

$$\chi^2_k(z_j) = \frac{1}{N_{jk} - 2}\sum_i^{N_{jk}}\left(\frac{F(\lambda_i) - f_k(\lambda_i, z_j)}{\sigma_i}\right)^2, \quad (2)$$

where N_{jk} is the number of λ values in which the (resampled) spectrum and the redshifted template k overlap at redshift z_j, and σ_i is the 1σ uncertainty in $F(\lambda_i)$.

The likelihood of a given redshift and template pair (z, T) is then $\mathscr{L}_T(z) \propto e^{-\chi^2_T(z)}$, and assuming all templates and redshifts have the same probability a priori, the maximum likelihood (ML) solution is simply the (z, T) pair that maximizes $\mathscr{L}_T(z)$.

The ML solution also assumes that the template set is exhaustive (includes all possible types), but this condition is difficult to meet with samples of MIR spectra because of the high number of physical processes involved.

When none of the templates is an accurate model of the spectrum, the χ^2 minimization favours those templates that best reproduce the overall shape of the continuum even if small-scale features are poorly fitted, simply because the latter only affect a small fraction of the λ_i. Such behaviour is unwelcome, since the narrow spectral features are crucial to obtain an accurate redshift estimate, while the continuum curvature and slope changes only provide rough redshift indications.

To overcome this limitation, we propose a new algorithm for finding the most probable redshift value, called 'maximum combined pseudo-likelihood' (MCPL).

The main features of the MCPL algorithm are: (i) it searches for local – instead of absolute – maxima in $\mathscr{L}_T(z)$; and (ii) it combines information on the local maxima found by all templates to produce a pseudo-likelihood distribution as a function of redshift.

The rationale behind this approach is that the broad-band SED of the source determines the general shape of $\mathscr{L}_T(z)$, while narrow spectral features cause high-frequency variations in $\mathscr{L}_T(z)$ as they correlate (or not) with features in the template. A good correlation of several features at a certain redshift produces a sharp peak in $\mathscr{L}_T(z)$ that signals a candidate redshift solution. The spurious alignment of a few features or noise spikes in the spectrum and the template can also produce a peak in $\mathscr{L}_T(z)$ at a wrong redshift. This peak can even be higher than the $\mathscr{L}_T(z)$ value at the actual redshift if the template is a poor model for the spectrum. However, such chance alignments tend to appear at different redshifts for each template, while the legitimate peak always occurs at the same (actual) redshift. Because of this, combining information on the position and strength of local maxima produced by all the templates has the potential to yield a more robust redshift estimate compared to considering only the best-fitting template.

This idea is implemented by a 'filter' function that zeroes all values of $\mathscr{L}_T(z)$ except those corresponding to local maxima:

$$\mathscr{L}^*_T(z) = \begin{cases} \mathscr{L}_T(z) & \text{if local maximum} \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

The combined, filtered likelihood distribution is then the sum over all the templates:

$$\mathscr{L}^*(z) = \sum_i^{N_T} \mathscr{L}^*_i(z). \quad (4)$$

The filtering implies that each template promotes only those redshift values at which it finds a (partial) correspondence of features with the spectrum.

The information provided by the continuum is not lost, though, since the height of the local peaks in $\mathscr{L}_T(z)$ still indicates the goodness of fit between spectrum and template at those particular redshifts.

Since the template set cannot reproduce all the possible combinations of spectral features (i.e. it is not complete), even the best-fitting template at the correct redshift is not in general an accurate model for the intrinsic spectrum of the source. In other words, the differences between spectrum and template cannot be accounted for by flux uncertainties alone. Therefore, the reduced χ^2 increases with increasing S/N, and produces $\chi^2 \gg 1$ even for fits that a visual inspector would consider very good.

The exponential dependency of $\mathscr{L}$ on χ^2 implies that the peak with lowest χ^2 usually dominates the resulting combined $\mathscr{L}^*(z)$, making the contribution from other peaks negligible and turning the MCPL algorithm into a simple ML algorithm.

To overcome this undesired effect, the likelihood function $\mathscr{L}$ can be substituted with a 'pseudo-likelihood' q that is proportional to the inverse of the reduced chi-squared statistics:

$$q_T(z) = \frac{a(z, T, \theta)}{\chi^2_T(z)} \quad (5)$$

where the *prior* $a(z, T, \theta)$ captures all the information about the source to be weighted in the selection of the best redshift estimate, such as the observed flux density in a given band, or the a priori probability of any (z, T) combination.

Using $q_T(z) \propto 1/\chi^2_T(z)$ instead of $\mathscr{L}_T(z) \propto \mathrm{e}^{-\chi^2_T(z)}$ still favours the lowest value of χ^2, but lets other local peaks have some influence on the final solution.

The 'combined pseudo-likelihood' function $Q(z)$ is then defined as

$$Q(z) = \sum_i^{N_T} q_i^*(z)\,, \tag{6}$$

where $q_i^*(z)$ is the filtered $q_T(z)$ for template i in which all values other than local maxima have been zeroed.

Variations among templates in the profile of a resolved spectral feature like the 7.7-μm PAH complex or the ∼10 μm silicate feature cause that different templates produce peaks at slightly different redshifts. This results in tight clusters of nearby peaks in the combined $Q(z)$. If the redshift difference between the peaks in a cluster is comparable to the theoretical redshift resolution the spectrum is capable of, it can be assumed that all these peaks represent the same redshift solution, although with some dispersion.

To compensate for this, $Q(z)$ is convolved with a Gaussian kernel (whose full width at half-maximum is twice the redshift resolution) to produce a smoothed $Q_s(z)$. The final solution is then the redshift $z_{\rm best}$ that maximizes $Q_s(z)$.

In regular SED fitting ML photometric redshifts, error bars for the redshift estimate can be computed using the $\Delta\chi^2$ method (e.g. Anvi 1976; Bolzonella et al. 2000). This method assumes the probability distribution for the minimum of $\chi^2(z)$ ($\chi^2_{\rm min}$) is the χ^2 distribution for n degrees of freedom (Press et al. 1992). Nevertheless, even for broad-band photometric redshifts the χ^2 distribution is not a realistic description of the actual redshift uncertainties, because the model is not linear in the fitting parameters (namely the redshift) and there are degeneracies between redshifts and galaxy SED types (Oyaizu et al. 2008).

Like $\chi^2_{\rm min}$, the peak value of $Q_s(z)$ depends mainly on the S/N of the spectrum, and does not provide a direct estimation of the reliability of the redshift solution. Still, our results with MIR spectra indicate that the value of $Q_s(z_{\rm best})$, if normalized to the integral over the entire range of redshifts,

$$\gamma = Q_s(z_{\rm best})/I, \qquad I = \int_{zmin}^{zmax} Q_s(z')\,{\rm d}z', \tag{7}$$

provides valuable information regarding the strength of the redshift solution. A value of γ close to 1 indicates that the peak at $z_{\rm best}$ clearly dominates the $Q_s(z)$ distribution, and thus the redshift estimate should be reliable. Conversely, a very low γ indicates that there are many secondary peaks with similar strength, and the redshift estimate is unreliable.

Another related parameter, useful to estimate the degree of degeneracy, is the ratio R between the $Q_s(z)$ values for the highest and second-highest peaks. A ratio close to 1:1 indicates peaks of comparable strength and reveals a significant probability of catastrophic redshift error.

3 SAMPLE SELECTION

The spectra used here were selected from the *Spitzer*/IRS ATLAS project (ATLAS-IRS; Hernán-Caballero & Hatziminaoglou 2011), which compiles MIR spectra and ancillary data from 739 extragalactic sources at $0 < z < 3.7$.

The parent sample is composed of all the ATLAS-IRS sources with a known spectroscopic redshift from optical or NIR spectroscopy ($z_{\rm spec}$). To ensure enough spectral coverage, 20 sources observed in only one of the four IRS modules were discarded. 11 additional sources were selected as templates (see Section 4) and removed from the sample to avoid circularity issues.

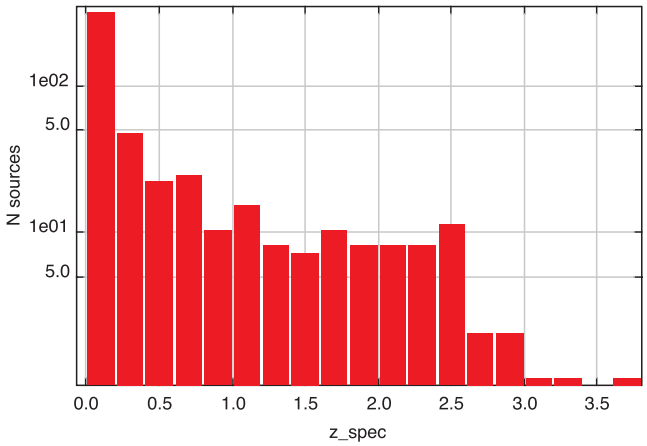

Figure 1. Redshift distribution for the 491 sources in the main sample.

The information content in each spectrum was computed using the net significance ($\mathscr{N}$), defined as the maximum cumulative S/N of the spectrum (Pirzkal et al. 2004). The three sources with lowest net significance values ($\mathscr{N} < 100$, corresponding to the median S/N per pixel $\lesssim$0.8) were also removed from the sample.

The remaining 491 sources constitute the main sample of this work. Their redshift distribution is shown in Fig. 1.

About half of the sample is at redshift $z_{\rm spec} < 0.15$, while only ∼20 per cent is at $z_{\rm spec} > 1$. This distribution is similar to that of the entire set of extragalactic sources with *Spitzer*/IRS spectroscopy and known spectroscopic redshifts (Lebouteiller et al. 2011).

321 sources (65 per cent of the sample) are classified as AGNs in the optical, including 124 optical quasars, 46 obscured quasars, 12 type 1 Seyferts, 73 intermediate-type Seyferts, 60 type 2 Seyferts and 56 low-ionization nuclear emission-line regions (LINERs).

In the MIR, 285 sources (58 per cent of the sample) are classified as AGN dominated, while 181 are starburst dominated and 10 are classified as composites with roughly equal contributions from the AGNs and starbursts to the IR output of the galaxy (see Hernán-Caballero & Hatziminaoglou 2011, for further details on the optical and MIR classifications).

4 TEMPLATES

To obtain reliable redshift estimates, it is essential that the templates cover as much as possible the rest-frame spectrum of the source at any redshift, since insufficient overlap between the spectrum and template increases the probability of obtaining a good fit at a wrong redshift due to chance alignments of spectral features.

For spectra observed in the four IRS channels (5–35 μm), a redshift search range $0 < z < 4$ implies that the templates should ideally span the entire 1–35 μm rest-frame range. In practice, shorter wavelength coverage suffices, as long as the template and spectrum share enough overlap in the entire redshift search range.

A large number of spectral templates were generated based on IRS spectra from ATLAS-IRS and from the Cornell Atlas of Spitzer/IRS Spectra (CASSIS; Lebouteiller et al. 2011) using the following methods:

(i) For a sample of local luminous and ultraluminous IR galaxies, their *AKARI* 2.5–5 μm spectrum (Imanishi et al. 2010) is used to extend their IRS spectrum down to rest frame ∼2 μm.

(ii) Composite spectra of several samples of $z > 1$ quasars (selected by their rest-frame 3.6–5.6 μm continuum slope and/or strength of the silicate feature) are used to extend the IRS spectrum of individual, lower redshift quasars with good S/N.

(iii) Individual and composite spectra of $z > 0.5$ radio galaxies are used to extend individual spectra of low-redshift radio galaxies.

(iv) Another template is produced from the composite spectra of radio galaxies with strong emission in the [S IV] 10.51 μm line.

(v) An early-type non-active galaxy template is obtained by extending the IRS spectrum of NGC 5011 with the elliptical galaxy template from Coleman et al. (1980).

In addition, archival *ISO*/SWS 2–45 μm spectra of NGC 1068, M82 and the Circinus galaxy are also included as templates.

A selection procedure was devised to identify the best performing templates. First, the filtered pseudo-likelihood function, $q_{i,j}(z)$, is calculated for every pair $\{i, j\}$ of template and spectrum. Then, subsets of templates are given a score based on the number of accurate $[\Delta z/(1+z) < 0.01]$ redshift solutions obtained with that particular subset. A penalization factor depending on the number of templates is also included to discourage large template sets.

An iterative process finds the subset that maximizes the score by randomly adding or removing one or several templates from the subset, keeping only those changes that increase the score until no further increases are possible. This process is run several times to ensure it always converges to the same template set.

Note that since the MCPL redshift solution for any source depends on the whole template set (and not just the template obtaining the best fit) the optimization would discard a template that produces good fits for a few unusual sources if it degrades the solution for many others. Nevertheless, it is unlikely that a chance alignment of some features in a template that does not match the overall SED of a source can produce a peak in $q_{i,j}(z)$ strong enough to overshadow the combined peaks produced by the remaining templates at the actual redshift.

The final set, containing 22 templates, is shown in Fig. 2 and listed in Table 1. The templates derived from spectra of sources in the sample produce, as expected, very good fits for these particular sources. To avoid misrepresenting the actual accuracy of the method, these sources are removed from the sample.

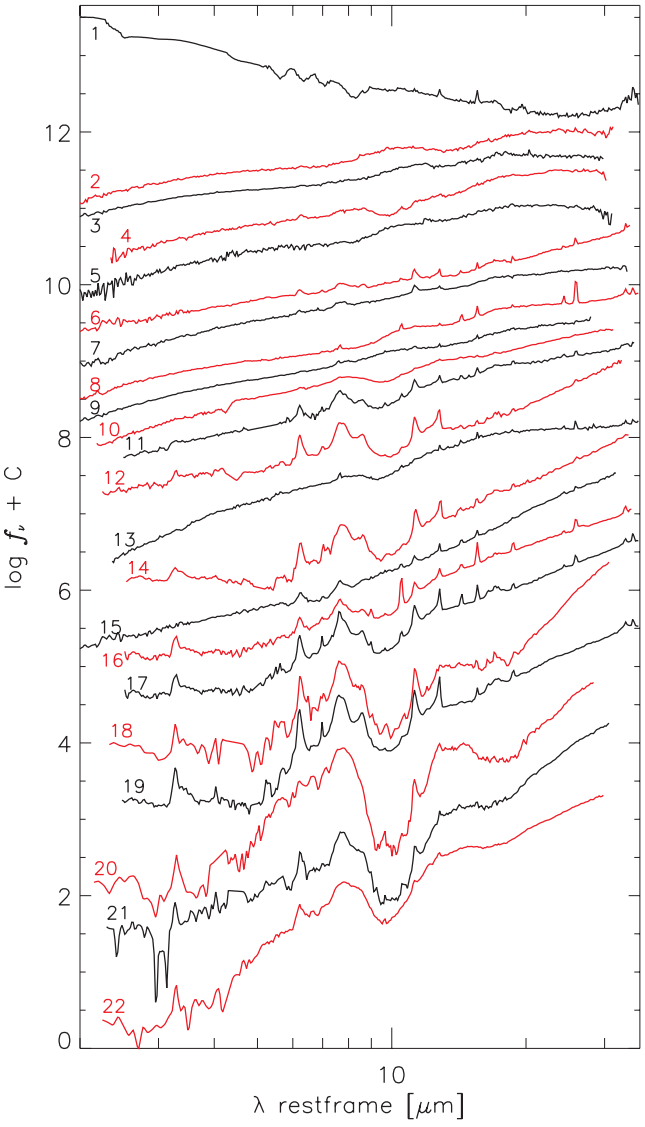

Figure 2. The 22 spectral templates used by the SED-fitting routine, ordered by increasingly red continuum. The numbers correspond to the row indices in Table 1. The alternative red and black colours are used for clarity only.

5 RESULTS

zCOLORS was run on the sample of 491 spectra selected from the ATLAS-IRS with the template set described in the previous section.

The spectra and templates are resampled to a constant spectral resolution $R = 500$, which provides a redshift resolution $\frac{\Delta z}{1+z} = 0.002$. In test runs, higher resolution values increase the computational cost with no significant gain in accuracy of the redshift solutions. The search range for redshifts is $-0.05 \leq z \leq 4$, with the extension to small negative values being important to properly identify the peak of $q_T(z)$ in nearby galaxies ($z \sim 0$).

All templates are assumed to have the same a priori probability. The only *prior* introduced is a luminosity limit, aimed at preventing bright sources from obtaining high redshift estimates that would imply unrealistically high luminosities. The luminosity limit is conservatively put at $\nu L_\nu = 5 \times 10^{47}\,\mathrm{erg\,s^{-1}}$, which is just above the most luminous source in the sample. For every source, the redshift ($z_{\rm cut}$) corresponding to this luminosity is found, and the *prior* $a(z)$ is then defined by

$$a(z) = \begin{cases} 1 & \text{if } z \leq z_{\rm cut} \\ 0 & \text{if } z > z_{\rm cut} \end{cases} \quad (8)$$

The luminosity limit achieves an ∼30 per cent reduction in the number of catastrophic errors $[\Delta z/(1+z) > 0.1]$ relative to a flat *prior*. This suggests a more elaborate *prior* including luminosity limits that depend on the SED and observed flux of the source could probably help to further reduce degeneracies.

Table 2 contains the redshift solutions obtained by the ML and MCPL algorithms for all the sources in the sample.

5.1 Accuracy of redshift solutions

The accuracy of the $z_{\rm IRS}$ estimates is evaluated by comparing them to the redshifts from optical or NIR spectroscopy ($z_{\rm spec}$). The error in the $z_{\rm IRS}$ value is represented by $d = (z_{\rm IRS} - z_{\rm spec})/(1 + z_{\rm spec})$ and its modulus, $\delta = |d|$, defines the accuracy of the redshift solution.

Table 1. IR templates.

No.	Template ID	Redshift	Type	Source
1	NGC 5011	0.001 05	El-2	CWW_E + IRS[a]
2	PG 1116+215	0.1765	QSO (S1.0)	QSO_comp + IRS
3	PG 1048+342	0.167 13	Elliptical (S1.0)	QSO_comp + IRS[b]
4	2MASS J12324114+1112587	0.249	NLAGN	QSO_comp + IRS
5	2MASXi J2222211+195947	0.21	Seyfert (S1.0)	IRS + IRS[c]
6	PG 1440+356	0.079	QSO (S1n)	QSO_comp + IRS
7	[HB89] 1415+451	0.1136	RQQ (S1.0)	QSO_comp + IRS
8	3C 120	0.03301	BLRG	QSO_comp + IRS
9	PG 1700+518	0.292	QSO (S1.0)	QSO_comp + IRS
10	IRAS 23060+0505	0.173	ULIRG (S1h)	*AKARI* + IRS
11	NGC 7674	0.028 92	Seyfert (S1h)	*AKARI* + IRS
12	IRAS 08559+1053	0.148	ULIRG (S2)	*AKARI* + IRS
13	NGC 1068	0.003 79	LIRG (S2)	*ISO* SWS
14	NGC 7130	0.016 15	LIRG (S1.9)	*AKARI* + IRS
15	PG 0157+001	0.163 11	QSO (S1.5)	QSO_comp + IRS
16	RG strong [S IV]	0.000 00	RG composite	IRS composite[d]
17	ESO 339-G011	0.0192	LIRG (S2)	*AKARI* + IRS
18	IRAS 22206−2715	0.1314	ULIRG (H II)	*AKARI* + IRS
19	MCG-02-01-051	0.0271	LIRG (H II)	*AKARI* + IRS
20	IRAS 11582+3020	0-223	ULIRG (LINER)	*AKARI* + IRS
21	IRAS 12359−0725	0.138	ULIRG (H II/LINER)	*AKARI* + IRS
22	IRAS 12018+1941	0.1686	ULIRG (S2)	*AKARI* + IRS

[a]Optical–NIR elliptical galaxy template from Coleman et al. (1980).
[b]Glikman + blue AGN composite.
[c]Extended at shorter wavelengths using the IRS spectrum of the broad absorption line (BAL) quasar ELAISC15 J003059−442133 ($z = 2.101$).
[d]Composite spectrum of radio galaxies with strong emission in the [S IV] 10.51 μm line.

Table 2. Results for individual sources. This is a sample of the full table which is available as Supporting Information with the online version of the article.

Source ID (1)	RA (J2000.0) (2)	Dec. (J2000.0) (3)	z_{spec} (4)	z_{MCPL} (5)	z_{ML} (6)	γ (7)	R (8)	$\log \mathcal{N}$ (9)	$\log \chi^2_{ML}$ (10)	MIR class (11)
UGC 00006	00:03:09.60	+21:57:37.0	0.0220	0.0222	0.0202	0.62	9.83	5.30	2.79	MIR_SB
[HB89] 0003+199	00:06:19.50	+20:12:10.0	0.0260	0.0284	0.0284	0.20	2.19	3.75	0.20	MIR_AGN1
2MASX J00070361+1554240	00:07:03.60	+15:54:24.0	0.1140	0.1118	1.5595	0.15	1.44	4.48	1.64	MIR_AGNx
III Zw 002	00:10:30.80	+10:58:13.0	0.0898	0.0898	0.0876	0.17	2.45	4.79	1.99	MIR_AGN1
NGC 0017	00:11:06.50	−12:06:26.0	0.0196	0.0202	0.0202	0.52	19.96	4.11	1.73	MIR_SB
2MASX J00114330−0722073	00:11:43.30	−07:22:07.0	0.1180	0.1185	0.1185	0.82	17.71	4.99	2.02	MIR_SB
2MASX J00212652−0839261	00:21:26.50	−08:39:26.0	0.1280	0.1320	0.1320	0.46	9.23	4.97	2.48	MIR_AGN2
LBQS 0018−0220	00:21:27.30	−02:03:33.0	2.5960	0.8183	2.5958	0.11	1.80	3.60	0.31	MIR_AGN
2MASX J00215355−7910077	00:21:57.00	−79:10:14.0	0.0728	0.0304	0.0304	0.13	1.47	4.62	0.92	MIR_AGN
PG 0026+129	00:29:13.60	+13:16:03.0	0.1420	0.1388	0.2585	0.13	1.06	4.16	1.22	MIR_AGN1

Notes. Column (1): source identification in NED. Columns (2) and (3): J2000 RA and Dec. Column (4): spectroscopic redshift from NED. Columns (5) and (6): MCPL and ML redshift solutions, respectively. Column (7): reliability parameter for the MCPL solution. Column (8): ratio of γ values for the first and second MCPL redshift solutions. Column (9): logarithm of the net significance. Column (10): logarithm of the reduced χ^2 for the ML solution. Column (11): MIR classification of the source.

Fig. 3 shows the correlation between z_{IRS} and z_{spec} values for both, the ML and MCPL selection algorithms. 86 per cent of MCPL solutions and 78 per cent of ML solutions are enclosed within the parallel lines that represent accuracy $\delta < 0.02$, but typical accuracies are much higher.

The number of outliers ($\delta > 0.02$) is 69 for the MCPL algorithm and 106 for the ML algorithm. Excluding them, errors for MCPL (ML) solutions average 2.0×10^{-5} (-4.5×10^{-5}) with a standard deviation of 0.0033 (0.0046). This indicates there is no significant bias in the redshift estimates, and the typical errors are just a fraction of the spectral resolution of the IRS ($\delta\lambda/\lambda \sim 0.008$–$0.016$, depending on the wavelength). The distribution of redshift errors is approximately Gaussian, with full width at half-maximum 0.0047 for the MCPL approach and 0.0056 for the ML approach (see Fig. 4).

The cumulative distribution of δ for MCPL and ML solutions is presented in Fig. 5. Both distributions show similar trends, with a rapid growth in the number of sources up to $\delta \sim 0.005$ and much slower growth at higher values. The curve for MCPL solutions is consistently over ML solutions in the entire δ range, but the separation is larger at $\delta \sim 0.005$. The number of sources in the range $0.02 < \delta < 0.1$ is 52 for the ML approach versus 29 for the MCPL approach, indicating a higher prevalence of low-accuracy solutions in the ML approach. Catastrophic redshift errors ($\delta > 0.1$) are obtained in 40 and 54 sources for the MCPL and ML algorithms, respectively.

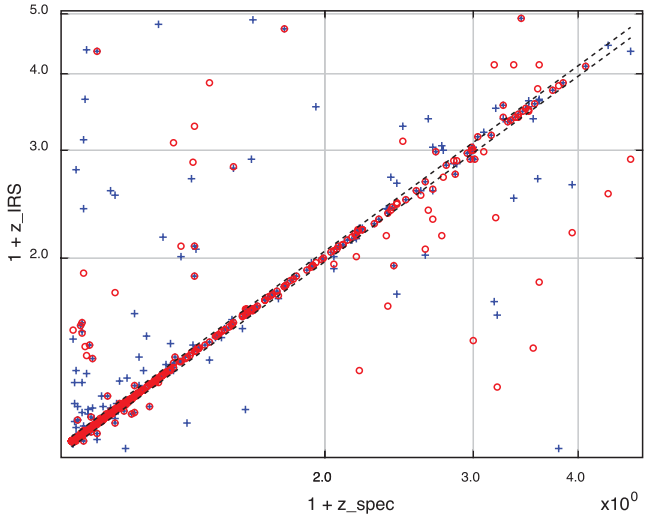

Figure 3. Comparison between the redshifts derived from template fitting ($z_{\rm IRS}$) and those from optical spectroscopy ($z_{\rm spec}$). The plus signs mark solutions from the ML algorithm, while the open circles mark those of the MCPL selection algorithm. The dotted lines enclose those sources with redshift accuracy $\delta < 0.02$.

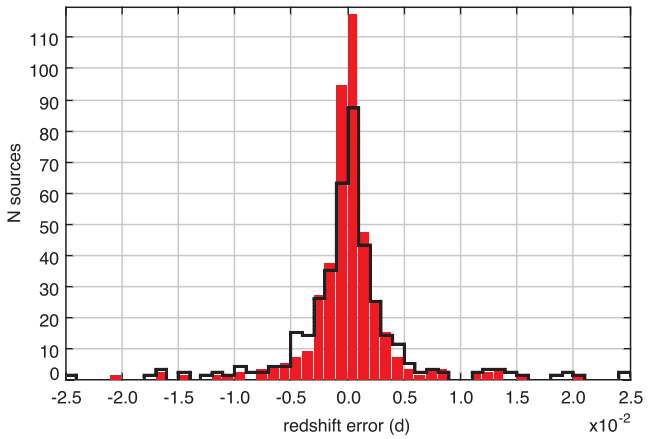

Figure 4. Distribution of redshift errors [$d = (z_{\rm IRS} - z_{\rm spec})/(1 + z_{\rm spec})$] obtained using the ML (solid line) and MCPL (solid bars) selection algorithms.

Comparison of ML and MCPL solutions for individual sources reveals that in two-thirds of the sample (318 sources) δ values from both algorithms are within 10 per cent of each other. In another 132 sources, the MCPL solution is clearly more accurate (sometimes by several orders of magnitude), while only in 37 cases is the ML solution significantly better.

The accuracy advantage of MCPL solutions over ML solutions is clearer at low redshift: at $z < 0.5$ MCPL solutions outperforms ML solutions in 96 cases versus 16 for ML solutions, while at $z > 1$ they are levelled, with each of them winning in 19 cases. This is probably a consequence of the decrease in the average S/N with redshift. A lower S/N reduces the contrast of the legitimate peak in $Q_{\rm s}(z)$ and makes it easier for spurious peaks to obtain comparable strength, increasing the risk of degeneracies.

Fig. 6 represents the ratio between δ values of the MCPL and ML solutions versus the 'degeneracy parameter' (R), defined as the ratio between the highest and second-highest peaks in $Q_{\rm s}(z)$. Most sources with ML solutions significantly more accurate than MCPL solutions have $\log(R) < 0.15$, indicating an extreme degree

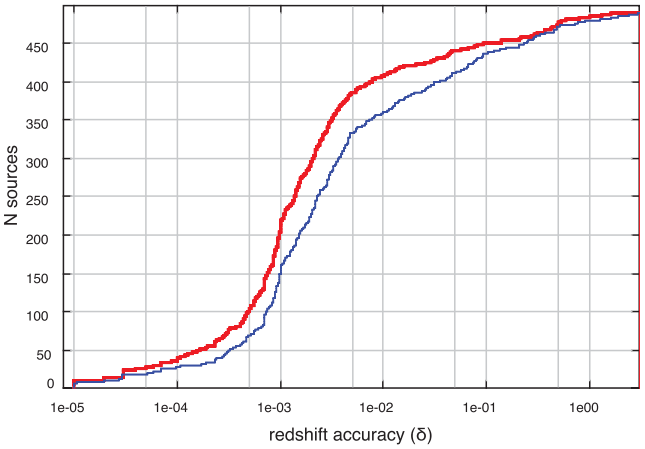

Figure 5. Cumulative distribution of the accuracy parameter $\delta = |z_{\rm IRS} - z_{\rm spec}|/(1+z_{\rm spec})$. The thin solid line represents results for the ML selection algorithm, while the thick solid line corresponds to those of the MCPL algorithm.

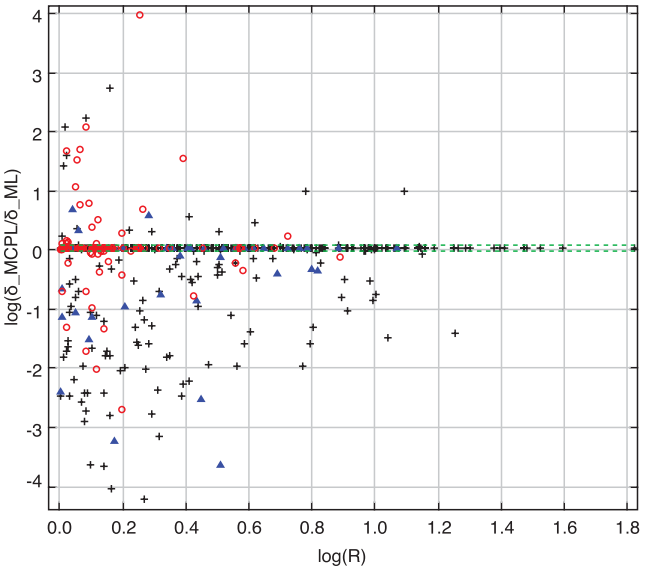

Figure 6. Logarithm of the ratio between accuracies of the MCPL and ML redshift solutions versus degeneracy parameter (R) of the MCPL solution for sources at $z_{\rm spec} < 0.5$ (plus signs), $0.5 < z_{\rm spec} < 1$ (solid triangles) and $z_{\rm spec} > 1$ (open circles). The negative values indicate a higher accuracy for the MCPL solution compared to the ML one. The dotted lines enclose the region corresponding to the MCPL and ML solutions within 10 per cent of each other, which contains 65 per cent of the sources in the sample.

of degeneracy in $Q_{\rm s}(z)$. In these sources, the two highest peaks in $Q_{\rm s}(z)$ have very similar strength, and it is thus no surprise that the MCPL algorithm chooses sometimes the wrong one. Nevertheless, even with a very degenerate $Q_{\rm s}(z)$ the MCPL algorithm offers higher reliability than the ML one: from 138 sources with $\log(R) < 0.15$, the MCPL solution is accurate ($\delta < 0.02$) while the ML solution is an outlier ($\delta > 0.02$) in 26 cases, versus only seven the other way around. In another 63 cases, both are accurate and in the remaining 40 both are outliers.

5.2 Reliability of individual solutions

Apart from obtaining a high rate of accurate solutions, it is important to know the reliability of individual solutions. In Section 2, we

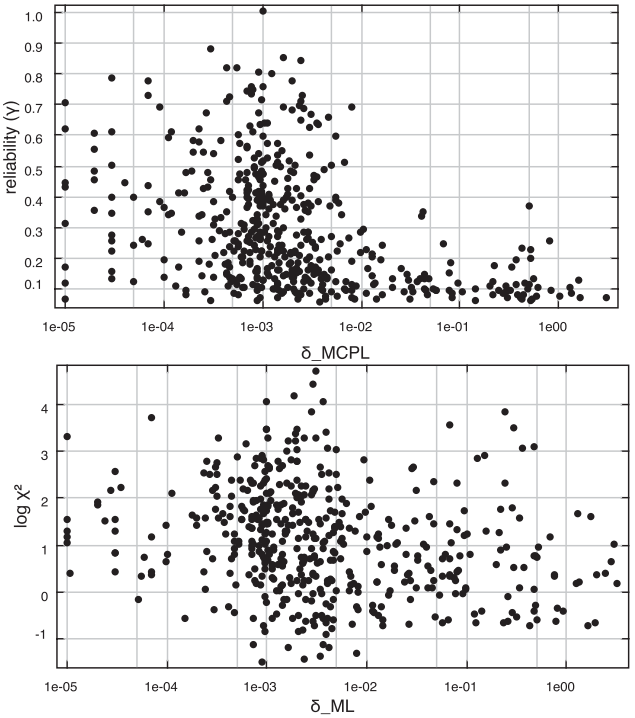

Figure 7. Top panel: relation between the reliability (γ) and accuracy (δ) parameters for MCPL solutions. Bottom panel: logarithm of the reduced χ^2 statistic versus δ for the ML solutions.

anticipated that the γ parameter can provide such information for the MCPL algorithm.

The upper panel in Fig. 7 represents γ versus δ for the MCPL solutions of the 491 sources in the sample. Accurate solutions obtain γ values spanning the whole ∼0.05–1 range, while outliers concentrate at $\gamma \lesssim 0.15$, with few cases above that value.

Table 3 demonstrates that the reliability of MCPL solutions increases monotonically with γ, both in terms of the dispersion and median of δ values and in terms of the frequency of outliers. This implies that by setting an appropriate threshold value for γ and selecting only the sources that surpass it, it is possible to obtain subsamples of sources with very reliable redshift estimates, albeit at a cost of completeness.

In contrast, χ^2 values for ML solutions (lower panel in Fig. 7) do not show an increase with δ. The higher χ^2 values occur preferably in sources with accurate redshifts, because these are usually the ones with higher S/N spectra. In other words, the χ^2 statistic correlates

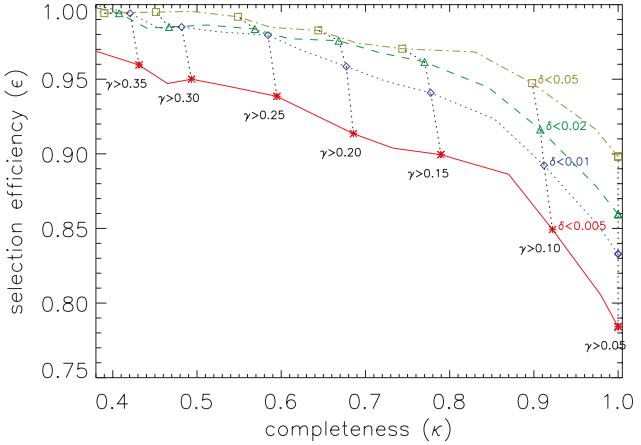

Figure 8. Selection efficiency versus completeness as a function of the threshold value used for the reliability parameter γ for redshift solutions with accuracy $\delta < 0.005$ (solid line), $\delta < 0.01$ (dotted line), $\delta < 0.02$ (dashed line) and $\delta < 0.05$ (dot–dashed line).

with the S/N of the spectrum, because at high S/N differences in the profile and strength of spectral features between spectrum and template are evident, while a very noisy spectrum blurs its features to the point that some of the templates can be considered an accurate model even at the wrong redshift. Therefore, the value of the absolute minimum in $\chi^2(z)$ cannot be used to identify the reliable ML solutions.

Since the MCPL algorithm offers higher redshift accuracy with a lower number of outliers, and also provides an indication of the reliability of the redshift solution, it can be considered superior to the ML algorithm for this purpose. In the remaining sections, only results from the MCPL algorithm will be discussed.

5.3 Selection efficiency and completeness

Let $N_g(D)$ be the number of sources with accuracy $\delta < D$, $N_s(G)$ the number of sources with $\gamma > G$, and $N_{sg}(D, G)$ the number of sources with $\delta < D$ and $\gamma > G$. The selection efficiency (ϵ) and the completeness (κ) are then defined by

$$\epsilon(G, D) = \frac{N_{sg}(D, G)}{N_s(G)}, \qquad \kappa(G, D) = \frac{N_{sg}(D, G)}{N_g(D)}. \tag{9}$$

As usual, there is a trade-off between completeness and selection efficiency, with either of them increasing only at the expense of the other. The relationship between ϵ and κ for a grid of values of D and G is shown in Fig. 8.

Table 3. Reliability of redshift measurements.

γ range	N sources	$\sigma(d)^a$	Median δ^b	$f(\delta < 0.02)^c$	$f(\delta > 0.02)^d$	$f(\delta > 0.1)^e$
0.05–0.1	73	0.0279	0.0040	0.53	0.47	0.30
0.1–0.15	80	0.0204	0.0020	0.73	0.28	0.12
0.15–0.2	49	0.0131	0.0019	0.88	0.12	0.08
0.2–0.3	89	0.0084	0.0011	0.96	0.04	0.03
0.3–0.5	123	0.0056	0.0009	0.98	0.02	0.01
0.5–1.0	77	0.0020	0.0009	1.00	0.00	0.00

[a]Standard deviation of redshift errors (excludes catastrophic errors).
[b]Median redshift accuracy (excludes catastrophic errors).
[c]Frequency of accurate redshift solutions.
[d]Frequency of outliers.
[e]Frequency of catastrophic errors.

Table 4. Dependency on MIR type.

Subset	MIR_SB	MIR_AGN	MIR_AGN1	MIR_AGN2	Total
$\delta < 0.02$	179	223	80	114	422
$\delta \geq 0.02$	2	62	33	10	69
$\gamma > 0.15$	175	149	54	86	338
$\gamma > 0.15$ and $\delta < 0.02$	173	139	47	85	325
$\gamma > 0.15$ and $\delta \geq 0.02$	2	10	7	1	13
Total	181	285	113	124	491

For D values in the range 0.005–0.05, the completeness has a very small (but consistent) dependency on D that reflects the reduction in the average γ for higher values of δ. The dependency on G is much stronger due to the large number of sources with low γ values.

Efficiencies for $D = 0.01$, 0.02 and 0.05 converge rapidly to $\epsilon = 1$ with increasing γ, because there are almost no sources with high γ values and $\delta > 0.01$. Nevertheless, for $D = 0.005$ there is slower growth, since some sources with very high γ values have accuracies $0.005 < \delta < 0.01$.

5.4 Dependency on the MIR SED

Differences in the MIR SED of starburst-dominated versus AGN-dominated sources cause important variations in the average accuracy of the redshift solutions depending on the MIR SED type.

Normal and starburst galaxies usually have very prominent PAH bands that are easily identified even in low-S/N spectra, and almost always obtain very accurate redshifts. On the other hand, sources dominated by AGN emission usually show a flat continuum with no high-contrast features, and the redshift determination relies on broad and shallow silicate features (in emission or absorption) and/or unresolved emission lines, making it much harder to distinguish the peak in $Q_s(z)$ corresponding to the actual redshift of the source.

In Hernán-Caballero & Hatziminaoglou (2011), the ratio $r_{\rm PDR}$ of the total PAH luminosity to the integrated rest-frame 5–15 μm luminosity is used to classify the ATLAS-IRS spectra into starburst dominated (MIR_SB, $r_{\rm PDR} > 0.15$) and AGN dominated (MIR_AGN, $r_{\rm PDR} < 0.15$). This is roughly equivalent to imposing a threshold EW ~ 0.2 μm in the EW of the 6.2- or 11.3-μm PAH bands, which corresponds to roughly equal contributions from the starburst and AGN to the IR luminosity of the galaxy (Hernán-Caballero et al. 2009). MIR_AGN sources are further separated into silicate emission (MIR_AGN1) and silicate absorption (MIR_AGN2) sources. Table 4 summarizes the accuracy and reliability statistics for these populations.

Sources dominated by star formation in their MIR spectra (MIR_SB) almost always obtain accurate redshifts. Among the 182 MIR_SB galaxies in the sample, MCPL solutions include only two outliers ($\delta > 0.02$), namely Murphy19 and NGC 4579. The optical redshift of the submillimetre galaxy Murphy19 (SDSS J123716.59+621643.9) is $z_{\rm spec} = 0.557$ (Wirth et al. 2004), but Murphy et al. (2009) give $z = 1.82$ based on the IRS spectrum. Although our solutions are consistent with the latter ($z_{\rm MCPL} = 1.806$; $z_{\rm ML} = 1.795$), the second-highest peak in $Q_s(z)$ is at $z = 0.5495$, indicating the optical redshift is confirmed with $\gtrsim$98 per cent confidence (see Section 5.5). NGC 4579 (M58) is a local spiral galaxy. The IRS spectrum contains emission from the LINER nucleus and its surroundings, and shows very intense H_2 lines combined with an unusual PAH spectrum (bright 11.3-μm PAH emission but very weak 6.2- and 7.7-μm bands). The lack of templates with significant H_2 emission is probably the cause of the wrong redshift solution for this source. In spite of that, the second-highest peak in $Q_s(z)$ coincides with the optical redshift.

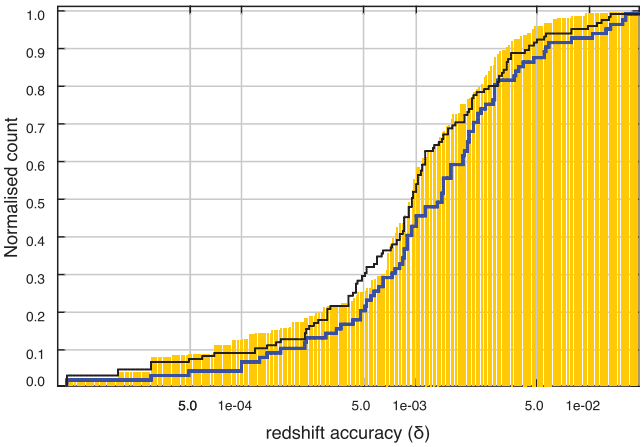

Figure 9. Normalized, cumulative distribution of δ values for MIR_SB (shaded area), MIR_AGN1 (thick solid line) and MIR_AGN2 (thin solid line) sources excluding outliers with $\delta > 0.02$.

Sources classified as MIR_AGN have redshifts that are much less reliable compared to those for MIR_SB. The overall outlier rate is 22 per cent, but there are strong variations in reliability among MIR_AGN subclasses: the fraction of outliers is 28 per cent for MIR_AGN1 versus 8 per cent for MIR_AGN2, and up to 40 per cent for MIR_AGN with no clear silicate emission or absorption. Nevertheless, if the outliers are removed, the accuracies for the remaining sources show very similar distributions in the MIR_SB, MIR_AGN1 and MIR_AGN2 subsamples (Fig. 9).

Further insight into the importance of the PAH bands for the accuracy (or lack thereof) of the redshift solution can be obtained from Fig. 10, which shows the strength of the PAH features, represented by $r_{\rm PDR}$, versus δ.

The subsample with $r_{\rm PDR} > 0.06$ includes by definition all the MIR_SB sources ($r_{\rm PDR} > 0.15$), the MIR composites ($r_{\rm PDR} \sim 0.15$), as well as some MIR_AGN sources with significant PAH emission. It comprises half of the total sample (242 sources) and has a 2 per cent rate of outliers and a median accuracy $\tilde{\delta} = 9.6 \times 10^{-4}$.

This demonstrates that detectable PAH emission is sufficient to obtain very accurate and reliable redshift estimates, even in high-redshift sources with poor-S/N spectra. However, PAHs are not the only feature capable of providing an accurate estimate, since there are many accurate redshifts down to $r_{\rm PDR} \sim 0$.

Fig. 11 shows the 9.7 μm apparent optical depth (a measurement of the strength of the silicate feature, see Hernán-Caballero & Hatziminaoglou 2011 for a discussion) versus δ for the sources with weak or undetected PAH bands ($r_{\rm PDR} < 0.06$). In this subsample, sources with silicate absorption ($\tau_{9.7} > 0$) are much more likely to obtain accurate redshifts than those with silicate emission

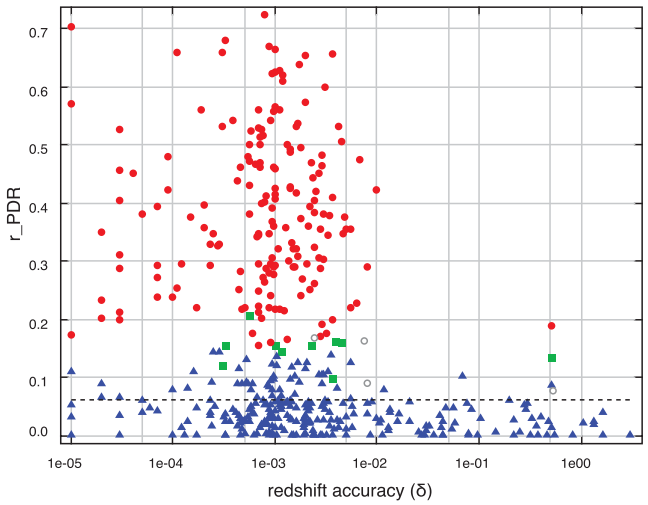

Figure 10. Ratio of the total PAH luminosity to integrated 5–15 μm luminosity ($r_{\rm PDR}$) versus redshift accuracy of the MCPL solution for sources classified as MIR_SB (solid circles), MIR_AGN (triangles), composite sources (squares), and sources with no MIR classification (open circles). The dashed line marks $r_{\rm PDR} = 0.06$.

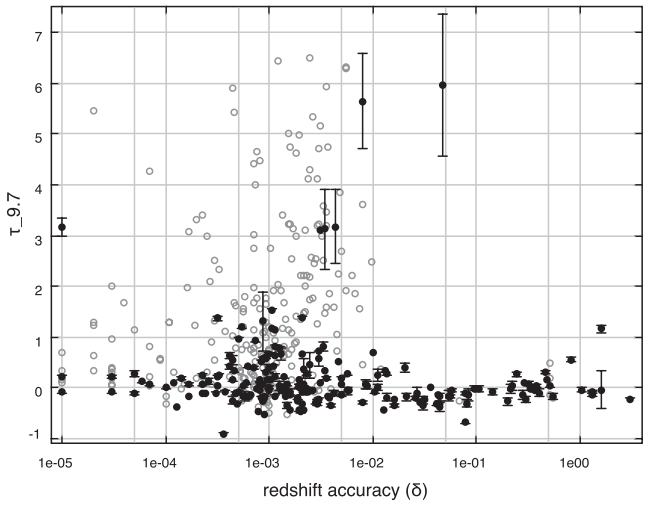

Figure 11. Optical depth at 9.7 μm ($\tau_{9.7}$) versus redshift accuracy. The solid symbols represent sources classified as MIR_AGN with very weak or absent PAH bands ($r_{\rm PDR} < 0.06$), while the open symbols represent the rest of the sample. The negative (positive) $\tau_{9.7}$ values indicate silicate emission (absorption).

(90 per cent versus 70 per cent, respectively, with $\delta < 0.02$), in spite of both populations having similar distributions of $r_{\rm PDR}$. This suggests that the silicate feature plays an important role in the redshift determination of sources with weak or no PAH emission. The diversity of shapes and lower contrast that the silicate feature presents when it appears in emission might be at least in part responsible for the decreased efficiency in these sources.

Albeit the rate of outliers is much higher in MIR_AGN sources compared to MIR_SB sources, the reliability of redshift estimates within a given γ interval is largely independent of the MIR classification. Fig. 12 shows the frequencies of highly accurate solutions ($\delta < 0.005$) and outliers ($\delta > 0.02$) as a function of γ for the MIR_AGN and MIR_SB populations separately. They are found to agree within their 90 per cent confidence limits.

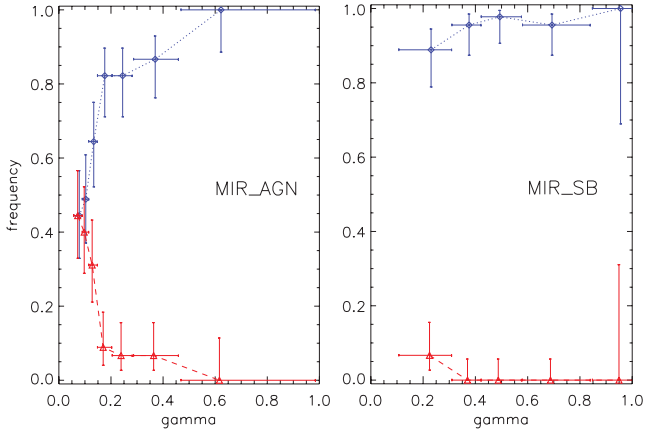

Figure 12. Frequency of highly accurate solutions ($\delta < 0.005$, blue diamonds) and outliers ($\delta > 0.02$, red triangles) as a function of γ for the sources classified as MIR_AGN (left-hand panel) and MIR_SB (right-hand panel). Each point represents a bin of γ containing 45 sources (except for the rightmost which contains the remainder). The horizontal bars represent the γ range covered by each bin, while the vertical error bars represent the 90 per cent confidence intervals calculated using the Wilson score formula for binomial distributions.

These confidence intervals can be used to put a lower limit on the probability of the redshift solution for a given source having accuracy better than some predefined value, or an upper limit on the probability of being an outlier. A more detailed model of such probabilities based on a much larger sample of MIR spectra from the CASSIS data base is under development, and will be presented elsewhere (Hernán-Caballero et al., in preparation).

5.5 Redshift degeneracies

There are 69 outliers ($\delta > 0.02$) in the sample, 40 of them with catastrophic errors ($\delta > 0.1$) in their redshift estimates.

Typical values of γ for the outliers are low, with 60 per cent of them below 0.15 (compared to just 20 per cent in the whole sample). They show multiple secondary peaks in the $Q_{\rm s}(z)$ distribution, with highest to second-highest ratios (R) in the range $1 < R < 3$, and in 90 per cent of cases verifying $R < 2$ (compared to only 36 per cent in the sources with $\delta < 0.02$). This suggests that most of these sources have degeneracy issues, and somehow spurious solutions obtain a $Q_{\rm s}(z)$ value higher than the peak for the actual redshift of the source.

Another possibility that deserves consideration is a wrong or inaccurate optical redshift. One way to rule out an error in the optical redshift is to search for a secondary peak in $Q_{\rm s}(z)$ that matches the $z_{\rm spec}$ value. If a significant peak is found very close to it, the optical redshift can be confirmed with high probability. On the other hand, the lack of a nearby secondary peak does not imply that $z_{\rm spec}$ is wrong, since spectra with very exotic MIR SEDs, strong artefacts or very low S/N could produce a very weak peak at the actual redshift that passes unnoticed.

To find out whether there are secondary solutions backing up the $z_{\rm spec}$ value for the 69 outliers, a routine finds all the peaks in the $Q_{\rm s}(z)$ distribution that verify $\gamma > 0.01$, and sorts them by their γ value. 48 outliers show at least one such secondary peak within $\delta < 0.02$ of the $z_{\rm spec}$ value. They are all listed in Table 5.

Since the redshift search range is very wide compared to the typical δ of these solutions, the probability of them occurring that close to $z_{\rm spec}$ fortuitously is low. If spurious peaks in $Q_{\rm s}(z)$ are

Table 5. Sources with accurate secondary redshift solutions. This is a sample of the full table which is available as Supporting Information with the online version of the article.

Source name (1)	$z_{\rm spec}$ (2)	$z_{\rm MCPL}$ (3)	$z_{\rm alt}$ (4)	$d_{\rm alt}$ (5)	$\gamma_{\rm alt}$ (6)	# (7)	P_r (8)
4C +29.31	0.3980	1.8457	0.3993	0.0009	0.070	2	0.0022
B3 1635+416	1.1790	1.0055	1.1423	−0.0169	0.085	3	0.0616
ELAISC15 J003014−430332	1.6540	1.3770	1.6480	−0.0022	0.067	2	0.0056
FLX J142644.33+333051.7	3.3550	1.5391	3.3568	0.0004	0.074	3	0.0015
GALEX 2533910445613399575	1.7070	1.9677	1.7069	−0.0000	0.087	4	0.0001
GOODS J123600.15+621047.5	2.0020	0.4593	2.0156	0.0045	0.037	7	0.0388
LBQS 0018−0220	2.5960	0.8183	2.5958	−0.0001	0.062	3	0.0002
MM J163655+4059	2.5920	3.1031	2.6030	0.0031	0.084	2	0.0076
SDSS J160250.95+545057.8	1.1970	0.3073	1.1943	−0.0012	0.053	3	0.0046
SDSS J161526.63+543005.9	1.3670	1.1681	1.4009	0.0143	0.054	3	0.0525

Notes. Column (1): source identifier in NED. Column (2): optical spectroscopic redshift. Column (3): MCPL redshift solution. Column (4): redshift of the secondary MCPL solution that matches $z_{\rm spec}$. Column (5): redshift error for the secondary solution. Column (6): reliability parameter for the secondary solution. Column (7): order number of the secondary solution after sorting by decreasing γ values. Column (8): probability of the secondary solution being spurious.

assumed to be randomly distributed in the redshift search range, the probability for a spurious solution obtaining accuracy δ or better is

$$P = \frac{2\delta}{\ln(1+z_{\rm max})}, \tag{10}$$

where $z_{\rm max}$ is the upper limit of the redshift search range for that particular source.

For a $Q_{\rm s}(z)$ distribution with several spurious peaks, the probability that at least one of the n highest peaks is within δ of $z_{\rm spec}$ just by chance is then

$$P_r(n) = 1 - (1-P)^n\,. \tag{11}$$

Probabilities P_r for the secondary solutions of these 48 outliers are listed in the last column of Table 5. 34 of them have $P_r < 5$ per cent, thus confirming the $z_{\rm spec}$ value with confidence $\gtrsim$95 per cent, while the other 14 have $0.05 < P_r < 0.25$, and some of them could be just random alignments.

5.6 Nature of outliers

Individualized inspection of outliers reveals a large number of radio galaxies and radio-loud quasars among them. The remainder are high-redshift sources (mostly quasars) with poor-S/N spectra.

Among the 21 outliers with no significant ($\gamma > 0.01$) secondary solutions within $\delta < 0.02$ of $z_{\rm spec}$, there are four local radio galaxies (3C 83.1, 3C 465, 3C 371 and 3C 390.3) and two intermediate-redshift radio-loud quasars (PG 2251+113 and 3C 295). The spectra of these six sources have high S/N, but are very different from each other. 3C 390.3 and PG 2251+113 show continuum emission peaking at ∼20 μm, a wide silicate emission feature and strong emission in the lines [Ne II] 12.81 μm, [Ne III] 15.55 μm and [O IV] 25.91 μm. 3C 83.1 and 3C 465 have a very weak MIR continuum dominated at $\lambda < 10$ μm by the Rayleigh–Jeans tail of stellar emission, and also show clear neon lines. 3C 371 is a flat spectrum radio source dominated by synchrotron emission in the MIR with no significant features. Finally, 3C 295 has a steep continuum lacking significant features and seems to have stitching issues in the LL2 module.

The redshift misidentification in all but the last two sources seems not to arise from a lack of spectral features capable of providing an accurate redshift estimate, but rather from an inadequate representation of these features in the set of templates used.

Another object, SWIRE J104354.82+585902.4, has conflicting optical redshift estimates: Trouille et al. (2008) give $z = 0.35$, while Weedman et al. (2006) give $z = 1.14\pm0.2$, much closer to the value 1.079 found here.

The remaining 14 sources are high-redshift ($z_{\rm spec} > 1$) quasars and IR galaxies with very low S/N spectra.

6 CONCLUSIONS

In this work, we apply a new SED-fitting algorithm to the problem of measuring redshifts in MIR low-resolution spectra. The algorithm is based on the same SED-fitting technique as applied to broad-band photometric redshifts, but with some important modifications that largely increase its efficiency with MIR spectra, namely a wavelength-dependent scaling factor for the template, which adds flexibility to the fit, and a novel algorithm for filtering and combining prospective redshift solutions, dubbed 'MCPL'.

The efficiency of the MCPL algorithm is compared to the regular ML algorithm using a sample of 491 *Spitzer*/IRS spectra for sources with accurate optical or NIR spectroscopic redshifts. The spectral templates used are obtained from *Spitzer*/IRS, *AKARI*/IRC and *ISO*/SWS spectroscopy of low-redshift galaxies, as well as composite templates of *Spitzer*/IRS spectra of higher redshift sources.

The MCPL algorithm offers superior performance compared to the ML algorithm both in terms of the number of highly accurate [$\Delta(z)/(1+z) < 0.005$] redshift solutions (78 per cent versus 68 per cent of the sample) and in terms of the number of outliers [$\Delta(z)/(1+z) > 0.02$; MCPL: 14 per cent, ML: 22 per cent]. Excluding outliers, the dispersion in the redshift errors is also lower for the MCPL algorithm: $\sigma = 0.0033$ (versus 0.0045 for the ML algorithm).

The reduced χ^2 statistic that determines the goodness of fit, often used to evaluate the reliability of the redshift solution, is found to correlate strongly with the S/N of the spectrum. High-S/N spectra obtain higher χ^2 values, indicating the differences between spectrum and template are more evident in them compared to low-S/N spectra. Nevertheless, the accuracy of the ML redshift solution does not correlate with χ^2, and thus cannot be directly used to estimate the confidence level of the redshift solution. On the other hand, the

normalized combined pseudo-likelihood (γ) offers a good indication of the reliability of the MCPL solution for individual spectra, with the median accuracy and rate of outliers both monotonically decreasing with increasing γ.

The fraction of accurate redshift solutions is much higher among sources classified as starbursts by their MIR emission compared to those classified as AGNs (2 per cent versus 21 per cent rate of outliers), thanks to the high contrast of the PAH emission bands, which are easily identified even in very low S/N spectra. The rate of outliers is also larger in AGNs with the 10 μm silicate feature in emission compared to those in absorption. Nevertheless, for any given γ range the accuracy of MCPL redshifts is largely independent of the MIR SED type.

Finally, we find that most outliers are radio galaxies, radio-loud quasars or high-redshift sources (mostly quasars) with very poor S/N. About two-thirds of the outliers show secondary MCPL solutions at the optical redshift. This indicates that degeneracy issues favoured spurious solutions in the selection process. This could be mitigated with templates that reproduce with greater fidelity the properties of these sources, in particular, radio galaxies.

ACKNOWLEDGMENTS

This work is based on observations made with the *Spitzer Space Telescope*, which is operated by the Jet Propulsion Laboratory, Caltech, under NASA contract 1407. The CASSIS is a product of the Infrared Science Center at Cornell University, supported by NASA and JPL. We wish to thank A. Alonso-Herrero, E. Hatziminaoglou, and the anonymous referee for useful comments that helped to improve this paper. AH-C is funded by the Universidad de Cantabria Augusto González Linares programme.

REFERENCES

Abramo L. R. et al., 2012, MNRAS, 423, 3251
Assef R. J. et al., 2008, ApJ, 676, 286
Avni Y., 1976, ApJ, 210, 642
Babbedge T. et al., 2004, MNRAS, 353, 654
Baum W. A., 1962, in McVittie G. C., ed., IAU Symp. Vol. 15, Problems of Extra-Galactic Research. Cambridge Univ. Press, Cambridge, p. 390
Benítez N., 2000, ApJ, 536, 571
Benítez N. et al., 2009, ApJ, 692, 5
Bolzonella M., Miralles J. M., Pelló R., 2000, A&A, 363, 476
Brodwin M. et al., 2006, ApJ, 651, 791
Brunner R. J., Connolly A. J., Szalay A. S., Bershady M. A., 1997, ApJ, 482, 21
Bruzual G., Charlot S., 1993, ApJ, 405, 538
Bruzual G., Charlot S., 2003, MNRAS, 344, 1000
Carliles S., Budavári T., Heinis S., Priebe C., Szalay A. S., 2010, ApJ, 712, 511
Coleman G. D., Wu C. C., Weedman D. W., 1980, ApJS, 43, 393
Collister A., Lahav O., 2004, PASP, 116, 345
Connolly A. J., Csabai I., Szalay A. S., Koo D. C., Kron R. G., Munn J. A., 1995, AJ, 110, 2655
Farrah D., Weedman D., Lonsdale C. J., Polletta M., Rowan-Robinson M., Houck J., Smith H. E., 2009, ApJ, 696, 2044
Feldmann R. et al., 2006, MNRAS, 372, 565
Fernández-Soto A., Lanzetta K. M., Yahil A., 1999, ApJ, 513, 34
Goicoechea J. R., Roelfsema P. R., Jellema W., Swinyard B. M., 2011, in Cernicharo J., Bachiller R., eds, Proc. IAU Symp. Vol. 280, The Molecular Universe. Cambridge Univ. Press, Cambridge
Gwyn S. D. J., Hartwick F. D. A., 1996, ApJ, 468, 77
Hatziminaoglou E., Mathez G., Pelló R., 2000, A&A, 359, 9
Hernán-Caballero A., Hatziminaoglou E., 2011, MNRAS, 414, 500
Hernán-Caballero A. et al., 2009, MNRAS, 395, 1695
Houck J. R. et al., 2004, in Mather J. C., ed., Proc. SPIE Vol. 5487, Optical, Infrared, and Millimeter Space Telescopes. SPIE, Bellingham, p. 62
Houck J. R., Weedman D. W., Le Floc'h E., Hao L., 2007, ApJ, 671, 323
Ilbert O. et al., 2006, A&A, 457, 841
Imanishi M., Nakagawa T., Shirahata M., Ohyama Y., Onaka T., 2010, ApJ, 721, 1233
Koo D. C., 1985, AJ, 90, 418
Lanzetta K. M., Yahil A., Fernandez-Soto A., 1996, Nat, 381, 759
Le Borgne D., Rocca-Volmerange B., 2002, A&A, 386, 446
Lebouteiller V., Barry D. J., Spoon H. W. W., Bernard-Salas J., Sloan G. C., Houck J. R., Weedman D., 2011, ApJS, 196, 8
Matute I. et al., 2012, A&A, 542, 20
Murakami H. et al., 2007, PASJ, 59, 369
Murphy E. J. et al., 2009, ApJ, 698, 1380
Negrello M. et al., 2009, MNRAS, 394, 375
Onaka T. et al., 2007, PASJ, 59, 401
Oyaizu H., Lima M., Cunha C. E., Lin H., Frieman J., 2008, ApJ, 689, 709
Pirzkal N. et al., 2004, ApJS, 154, 501
Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P., 1992, Numerical Recipes in C: The Art of Scientific Computing, 2nd edn. Cambridge Univ. Press, Cambridge
Richards G. T. et al., 2001, AJ, 122, 1151
Rowan-Robinson M. et al., 2008, MNRAS, 386, 697
Sawicki M. J., Lin H., Yee H. K. C., 1997, AJ, 113, 1
Silva L., Granato G. L., Bressan A., Danese L., 1998, ApJ, 509, 103
Trouille L., Barger A. J., Cowie L. L., Yang Y., Mushotzky R. F., 2008, ApJS, 179, 1
Wada T., Kataza H., 2010, in Oschmann J. M., Jr, Clampin M. C., MacEwen H. A., eds, Proc. SPIE Vol. 7731, Space Telescopes and Instrumentation 2010: Optical, Infrared, and Millimeter Wave. Am. Inst. Phys., New York, p. 77310U
Wadadekar Y., 2005, PASP, 117, 79
Wang Y., Bahcall N., Turner E. L., 1998, AJ, 116, 2081
Weedman D. W. et al., 2006, ApJ, 653, 101
Weedman D. W., Howick J. R., 2009, ApJ, 693, 370
Wirth G. D. et al., 2004, AJ, 127, 3121
Wright G. S. et al., 2008, in Oschmann J. M., Jr, de Graauw M. W. M., MacEwen H. A., eds, Proc. SPIE Vol. 7010, Space Telescopes and Instrumentation 2008: Optical, Infrared, and Millimeter. SPIE, Bellingham, p. 70100T
Yan L. et al., 2007, ApJ, 658, 778

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Table 2. Results for individual sources.
Table 5. Sources with accurate secondary redshift solutions.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

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