Bayesian photometric redshifts of blended sources

Daniel M. Jones* and Alan F. Heavens

Astrophysics Group and Imperial Centre for Inference and Cosmology, Imperial College London, Kensington, London SW7 2AZ, UK

Accepted 2018 November 29. Received 2018 November 23; in original form 2018 August 8

ABSTRACT
Photometric redshifts are necessary for enabling large-scale multicolour galaxy surveys to interpret their data and constrain cosmological parameters. While the increased depth of future surveys such as the Large Synoptic Survey Telescope (LSST) will produce higher precision constraints, it will also increase the fraction of sources that are blended. In this paper, we present a Bayesian photometric redshift (BPZ) method for blended sources with an arbitrary number of intrinsic components. This method generalizes existing template-based BPZ methods, and produces joint posterior distributions for the component redshifts that allow uncertainties to be propagated in a principled way. Using Bayesian model comparison, we infer the probability that a source is blended and the number of components that it contains. We extend our formalism to the case where sources are blended in some bands and resolved in others. Applying this to the combination of LSST- and Euclid-like surveys, we find that the addition of resolved photometry results in a significant improvement in the reduction of outliers over the fully blended case. We make available blendz, a Python implementation of our method.

Key words: methods: statistical – galaxies: distances and redshifts – cosmology: observations.

1 INTRODUCTION
Current photometric surveys such as CFHTLens (Heymans et al. 2012), KiDS (de Jong et al. 2013), and DES (Dark Energy Survey Collaboration 2016) image galaxies over large volumes of the Universe to probe the growth of structure and the distribution of matter on large scales. Through techniques such as galaxy clustering and cosmic shear, these surveys are able to constrain cosmological parameters and conduct tests of the standard ΛCDM cosmological model (e.g. Kilbinger et al. 2013; Dark Energy Survey Collaboration 2018).

These observational tests require the redshift distribution of the sample to make model predictions for comparison. Additional information can also be obtained by considering the redshift dependence using tomography, where galaxies are placed into one of several redshift bins (e.g. Hu 1999; Petri, May & Haiman 2016; Joudaki et al. 2018). However, the large number of sources required to constrain cosmological parameters to high precision makes obtaining spectroscopic redshifts for the entire sample unfeasible. As a result, photometric redshifts are a vital part of the cosmological analysis pipeline of galaxy surveys.

Photometric redshifts seek to infer the redshift of galaxies from noisy observations of their flux in several broad-band filters. They provide an alternative to spectroscopic redshifts that requires less telescope time, at the expense of a reduction in precision. As a result, photometric redshifts can be applied to galaxies too faint and samples too large for spectroscopic observations. There are two general classifications for photometric redshift methods: Bayesian methods, introduced by Benítez (2000), marginalize over all templates to produce a posterior redshift distribution. This correctly accounts for the uncertainty in the galaxy template that is ignored by maximum likelihood methods. Bayesian methods also include prior distributions that can reduce catastrophic outliers.

Template-based methods use a set of galaxy spectra that are assumed to be representative of every galaxy they are applied to. These templates are redshifted and integrated over the response function of each filter to produce predictions of observed fluxes. These predictions are then used to infer the redshift from the observed fluxes. Maximum likelihood methods (e.g. Bolzonella, Miralles & Pelló 2000; Ilbert et al. 2006) find the best-fitting template by minimizing χ² to estimate the redshift. Bayesian methods, introduced by Benítez (2000), marginalize over all templates to produce a posterior redshift distribution. This correctly accounts for the uncertainty in the galaxy template that is ignored by maximum likelihood methods. Bayesian methods also include prior distributions that can reduce catastrophic outliers.

Empirical methods estimate redshifts by fitting for the mapping between flux and redshift from a set of training data, rather than specifying it a priori through a template set. This mapping is typically found using machine learning methods such as neural networks (e.g. Collister & Lahav 2004; Sadeh, Abdalla & Lahav 2016), Gaussian processes (e.g. Way & Srivastava 2006; Almosallam, Jarvis & Roberts 2016), and random forests (e.g. Carliles et al. 2010; Carrasco Kind & Brunner 2013). These methods are exam-
samples of supervised learning; they require large data sets of fluxes and associated spectroscopic redshifts that are representative of the sample they are applied to.

If representative data are available, empirical methods are typically more accurate than template-based methods (Hildebrandt et al. 2010). However, redshift estimates of galaxies not represented by the training data are much less reliable (Beck et al. 2017). The common case where spectroscopic training data are shallower than the photometry can lead to biases where the redshifts of high-redshift galaxies are underestimated (Rivera et al. 2018).

In practice, template-based and empirical methods are not so distinct. The priors of Bayesian methods typically include a set of parameters that are fitted using a set of training data (e.g. Benitez 2000; Schmidt & Thorman 2013, also see Section 2.6). In addition, recent applications of photometric redshifts have used hybrid methods that combine a template-based approach with machine learning methods (e.g. Speagle & Eisenstein 2017; Duncan et al. 2018; Leistedt et al. 2018).

In addition to these methods, clustering redshift methods (e.g. Newman 2008; Ménard et al. 2013; Schmidt et al. 2013) cross-correlate the angular positions of a photometric sample with a spectroscopic sample to estimate the redshift distribution. Clustering redshift methods do not model fluxes as a function of redshift, instead only using the spatial information of photometric data. As such, clustering redshifts are complementary to other photometric redshift methods; Cawthon et al. (2017) use clustering redshifts to calibrate biases from other photometric redshift methods, for example.

Ensuring that photometric redshifts are accurate and precise is necessary for obtaining unbiased constraints on cosmological parameters. Huterer et al. (2006) found that future tomographic surveys would require the mean of each redshift bin to be known to a precision of 0.003, though this requirement can be reduced by self-calibration (e.g. Huterer et al. 2006; Sun, Zhan & Tao 2015; Samuroff et al. 2017) and combining weak lensing data with other cosmological probes such as baryonic acoustic oscillations (LSST Science Collaboration 2009). Photometric redshifts are also important in the calibration of other systematics. Multiplicative biases in the measurement of shear can be detected and corrected for, provided that photometric redshifts of galaxies in the sample are unbiased (Hoekstra, Viola & Herbonnet 2017). Weak lensing shape measurement biases can themselves also be redshift dependent; without unbiased redshift estimates to make corrections, these can lead to biases of a few percent in the cosmological parameters $\sigma_8$ and $w_0$ (Sembolini et al. 2009).

Another key part of precision cosmology is an accurate understanding of uncertainties in parameter constraints. To enable this, uncertainties arising from each step of the analysis should be accounted for and propagated onwards. In cosmological analyses, this is typically accomplished using a Bayesian framework (e.g. Hildebrandt et al. 2017; Troxel et al. 2017), allowing these uncertainties to be combined and marginalized over for the final constraints. It is therefore essential that photometric redshift methods provide not only point estimates of redshifts, but also a measure of their uncertainties.

The uncertainty associated with a redshift estimate can be represented by a single number, i.e. a point estimate with an error bar. However, doing this necessitates making an assumption about how the error is distributed. Uncertainties in photometric redshifts can be highly non-Gaussian, and so are poorly described by a single number such as the variance. Photometric redshift methods that instead characterize their results using a probability distribution function (PDF) can capture all of this information.

Photometric redshifts can also suffer from degeneracies that result in high-redshift galaxies having similar colours to those at low redshifts (e.g. Graham et al. 2018). As a result, several well-separated redshifts are plausible, and an accurate representation of the uncertainty should reflect this. While this can be easily described with a multimodal PDF, a single number can be misleading. Error bars that cover the full range of parameter space between the low- and high-redshift estimates do not show that redshifts between these are disfavoured, inflating uncertainties.

Several photometric redshift methods are able to produce PDFs as their result. Bayesian template-based methods (e.g. Benitez 2000) produce a posterior distribution, a PDF of the model parameters conditioned on the data and any model assumptions. In addition to the galaxy redshift, these model parameters can include other quantities of interest such as the galaxy template (Feldmann et al. 2006). A joint posterior over all of these parameters contains information about the uncertainty of each, including any correlation between them. Machine learning methods can also produce PDFs by utilizing ensemble techniques, where the predictions of several models are combined to produce a distribution. Examples of this technique include the combination of decision trees in a random forest (Carrasco Kind & Brunner 2013) and committees of neural networks constructed with different network architectures and initialized randomly (Firth, Lahav & Somerville 2003).

Future galaxy surveys such as the Large Synoptic Survey Telescope (LSST; Ivezić et al. 2008) will obtain extremely high-precision constraints on cosmological parameters. By utilizing deeper photometry, these surveys will probe greater volumes than previously, resulting in an increased number density of galaxies imaged. While this increased depth drives the high precision these surveys will achieve, it also increases the fraction of objects that overlap with others along the line of sight, known as blending (Chang et al. 2013).

Most existing deblending methods do not utilize the colour information from photometry, instead using the spatial information contained in an image from a single band. The commonly used SExtractor (Bertin & Arnouts 1996) searches for adjacent pixels on a flux-thresholded map that separate into disjoint regions as the threshold is increased. Doing this for many thresholds allows each pixel to be assigned to a single object, contributing the entirety of its flux to that object. The SDSS deblender (Lupton 2005) lifts this restriction, splitting the flux proportionally between objects based on object templates. These templates are constructed by finding peaks in the image and assuming symmetry around them, comparing pairs of pixels and setting them to be equal. Profile fitting methods (e.g. Peng et al. 2002; Robotham et al. 2017) forward model the image using physical profile models, deblending by directly fitting for the galaxy properties. In far-infrared astronomy, blending is common due to the reduced angular resolution of these instruments compared to optical telescopes. As a result, galaxies that are well resolved in optical observations may become blended in the far-infrared. Deblending methods designed for this case such as Hurley et al. (2017) can use the unblended observations to place strong priors on the number and position of sources. We refer to this mix of blended and unblended observations as partial blending throughout this paper.

The ability for most deblending methods to successfully identify blended galaxies depends on their angular separation, Galaxies
with too small an angular separation are instead identified as single sources. Dawson & Schneider (2014) estimate that 45–55 per cent of sources in LSST will be blended, with 15–20 per cent of all sources being misidentified as a single source, referred to as ambiguously blended objects.

Blending of sources can have an impact that is significant for constraining cosmological parameters. Dawson et al. (2016) estimate that ambiguously blended objects in LSST will result in an increase in shear noise of 14 per cent for the deepest photometry ($i < 27$) and 7 per cent for the gold standard sample ($i < 25.3$). Since these ambiguous blends are difficult to separate due to their small angular separation, deblending methods that incorporate colour information could be beneficial. Recent deblending methods such as MuSCADeT (Joseph, Courbin & Starck 2016) and SCARLET (Melchior et al. 2018) incorporate this colour information by using wavelet transforms, enforcing that the representation of components is sparse in this space.

Deblending methods that produce a set of component-separated maps are useful for later applying existing analysis techniques designed for individual components. However, splitting the analysis in this way can lose information, such as the correlation between deblending parameters and the parameters in a subsequent analysis. An analysis method that jointly constrains parameters directly from blended data provides a self-consistent, principled way to characterize and propagate this information.

In this paper, we present a method that generalizes the Benítez (2000) Bayesian photometric redshift (BPZ) method to the case of blended observations. This is a template-based method where the task of determining the component redshifts is cast as a Bayesian parameter inference problem. The product of such an inference is a joint posterior distribution of the redshift and magnitude of each component in the blended source. This distribution characterizes the complete statistical uncertainty in the result in a way that can be propagated through the rest of the cosmological analysis. Determining the number of components in an observed source, i.e. whether or not it is blended, is treated as a model comparison problem. In this way, our method allows the identification of blended sources from aperture photometry alone.

Throughout, we use source to refer to the (possibly) blended system that is observed, and component to refer to the underlying physical objects that make up this source. For parameters defined for each component in a source, we index over component using Greek letters and indicate the collection of these using sets, i.e. \{θ\} ≡ {θα, θβ, . . . , θN}. Vector quantities defined for each filter band are in bold \(q\), and observed quantities are denoted with a hat \(\hat{q}\). Where necessary, quantities defined for a specific number of components are distinguished by a subscript number in brackets, i.e. \(q^{(1)}\) is the definition of \(q\) for a single component. A summary of our notation is provided in Table 1.

This paper is organized as follows. In Section 2, we describe our formalism for estimating redshifts as a parameter inference problem, describing its application to partially blended systems in Section 3. In Section 4, we discuss our inference methods, detailing how we use model comparison to identify blended objects in Section 4.1. In Section 5, we test our method on simulated observations. Section 6 describes a test of our method on the Galaxy And Mass Assembly survey (GAMA; Baldry et al. 2017) blended sources catalogue (Holwerda et al. 2015), for which spectroscopic redshifts are available. We conclude in Section 7.

2 BLENDING PHOTO-Z FORMALISM

2.1 Flux model

In the same way as other template-based photometric redshift methods, we assume that each observed component is well represented by one of a set of \(T\) templates. Each template \(t\) is defined by its rest-frame spectral flux density \(F_t(\lambda_{\text{em}})\) as a function of the emitted wavelength \(\lambda_{\text{em}}\). This template is redshifted and observed through a broad-band filter \(b\), the response of which is denoted \(W_{b}(\lambda_{\text{obs}})\) as a function of observed wavelength \(\lambda_{\text{obs}}\).

The flux of template \(t\), at redshift \(z\) and observed in band \(b\), is then given by

\[
T_{t,b}(z) = \frac{1}{c g^{\text{AB}} C_{b}} \int_{0}^{\infty} F_t \left( \frac{\lambda}{1+z} \right) W_{b}(\lambda) d\lambda.
\]

(1)

where \(g^{\text{AB}} = 3631\) Jy is the zero-point of the AB-magnitude system and the normalization 

\[
C_{b} \equiv \int_{0}^{\infty} \frac{W_{b}(\lambda)}{\lambda} d\lambda.
\]

By including \(g^{\text{AB}}\), our fluxes are dimensionless throughout, and the conversion between magnitudes and fluxes defined in the way is given by \(F = 10^{0.4m}\).

This template is then scaled by a normalization \(a\) so that the flux of an object modelled with template \(t\), at a redshift \(z\) and observed in band \(b\), is given by

\[
F_{t,b}^{(a)}(z,a) = a T_{t,b}(z).
\]

(2)

We model the flux of blended sources as a linear combination of individual component fluxes. For a blend of \(N\) components, the flux observed in band \(b\) is given by

\[
F_{t,b}^{(N)}(z,\{a\}) = \sum_{a=1}^{N} a_{a} T_{t,b}(z_{a}).
\]

(3)

where \(a_{a}\) is the normalization for component \(a\). For the reasons specified in Section 2.3, we sample \(m_{0,a}\), the apparent magnitude of each component in the reference band \(b_{0}\) rather than this normalization directly. The normalization \(a_{a}\) is then defined such that the model flux in the reference band is equal to \(m_{0,a}\). Thus, the model flux is given by

\[
F_{t,b}^{(N)}(z,\{m_{0}\}) = \sum_{a=1}^{N} 10^{-0.4m_{0,a}} T_{t,b}(z_{a}).
\]

(4)

2.2 Fully blended posterior

For a fixed number of components, photometric redshift determination is a parameter inference problem; we wish to infer the joint posterior distribution of the redshifts and apparent magnitudes of each component given a data vector \(\vec{D}\) of broad-band fluxes. This data vector is split into two parts \(\vec{D} = (\hat{F}, \tilde{F})\), where \(\hat{F}\) is the flux of the reference band and \(\tilde{F}\) is the vector of the remaining \(B - 1\) fluxes. This is done since the normalization of each component is defined in the reference band, and it is the flux of this band on which the priors are conditioned.

Following BPZ (Benítez 2000), we set the flux of non-detections to zero. Likewise, bands that are not observed are given a flux of zero, with the corresponding error set to an extremely large value. As discussed in Section 2.4, we assume that sources are selected to zero. Likewise, bands that are not observed are given a flux of zero, with the corresponding error set to an extremely large value. As discussed in Section 2.4, we assume that sources are selected
index can take a value $1 \leq t \leq T$ and components may share the same template, so there are $T^B$ of these sets to marginalize over, giving

$$P([z], [m_0]|\hat{F}, \hat{F}_0, \chi, N) = \sum_{i=1}^{T^B} P([z], [t], [m_0]|\hat{F}, \hat{F}_0, \chi, N).$$

(5)

We have emphasized that our posterior is defined for a fixed number of components by conditioning on $N$. In the general case where this number is unknown a priori, it can be inferred from the data; this is discussed in Section 4.1. We have also made the dependence on cosmological parameters, which are required for converting between distance and redshift, explicit in the above expression. These parameters are denoted by $\chi = \{\Omega_m, \Omega_\Lambda, H_0\}$ for brevity. Applying Bayes rule, the posterior becomes

$$P([z], [m_0]|\hat{F}, \hat{F}_0, \chi, N) \propto \sum_{i=1}^{T^B} P([z], [t], [m_0], N)P([z], [t], [m_0]|\chi, N).$$

(6)

Since only the prior is dependent on cosmological parameters, we have removed the conditioning on $\chi$. We then factorize the likelihood so that it is split in the same way as the data vector, giving

$$P([z], [m_0]|\hat{F}, \hat{F}_0, \chi, N) \propto \sum_{i=1}^{T^B} P([z], [t], [m_0]|\hat{F}, \hat{F}_0, \chi, N)$$

$$\times P(\hat{F}_0|[m_0], N)P([z], [t], [m_0]|\chi, N).$$

(7)

Since the magnitude of each component in the reference band is a sampled parameter in the posterior, our model for the reference-band flux is simply the sum of these after converting from magnitudes to fluxes. As a result, the conditioning on $\{z\}$ and $\{t\}$, in the reference band likelihood is unnecessary and so has been removed. We assume that the error on the observed reference-band flux is normally distributed with variance $\sigma_b^2$. Thus, the reference band likelihood is given by

$$P(\hat{F}_0|[m_0], N) = \frac{1}{\sqrt{2\pi\sigma_b^2}} \exp \left( -\frac{(\hat{F}_0 - \sum_{i=1}^{N} 10^{-0.4m_0,b})^2}{2\sigma_b^2} \right),$$

(8)

where $m_{0,a}$ is the sampled reference band magnitude for component $a$. Similarly, we use an uncorrelated multivariate Gaussian likelihood for $\hat{F}$,

$$P(\hat{F}|[z], [t], [m_0], N) = \prod_{b=1}^{B} \frac{1}{\sqrt{2\pi\sigma_b^2}} \exp \left( -\frac{(\hat{F}_b - F_{(b)}^{(N)}([z], [m_0]))^2}{2\sigma_b^2} \right),$$

(9)

where $F_{(b)}^{(N)}([z], [m_0])$ is the model flux specified in equation (4) and $\sigma_b^2$ is the variance on the observed flux in band $b$.

2.3 Separating the joint prior

We now develop the prior so that it can be written in terms of individual components. We start by separating the joint prior into a product over priors on redshift, template, and magnitude. Removing unnecessary conditioning, the joint prior becomes

$$P([z], [t], [m_0]|\chi, N) = P([z]|\chi, N)P([t]|\chi, N)$$

$$\times P([m_0]|N)P([m_0]|N).$$

(10)

This splitting up of the joint prior is similar to the approach of Benítez (2000). There are two important differences, however. First, we include a prior on the apparent magnitude of each component. This differs from the approach of Benítez (2000), who considers the magnitude on which the redshift and template priors are conditioned to be exactly the observed reference band magnitude. The uncertainty in the scaling of the template is then represented by marginalizing over a normalization factor with an assumed flat prior. However, while this normalization is not defined as such, it is
acting to set the apparent magnitude of the source in the reference band. This magnitude is a quantity about which prior information is known.

The prior information on the apparent magnitude of components is particularly important in the blended case, as we need to consider more than just the overall magnitude of the source. The individual magnitudes of each component are necessary for scaling the model fluxes when predicting the model flux $F_{[m_0]}(z, \{m_0\})$. In addition, motivated by existing galaxy observations and following Benítez (2000), our redshift and template priors for each component are independent, i.e. galaxy types are not correlated. This allows us to make the assumption that the redshift of each component cannot be assumed to be independent, i.e. galaxy types are not correlated. This allows us to split the template prior as

$$P([t]|m_0, N) = \prod_{\alpha=1}^{N} P(t_\alpha|m_{0,\alpha}).$$

We also make the assumption that the redshift of each component depends only on its own type, not the types of other components. The redshifts of each component cannot be assumed to be independent however, as galaxies are distributed in a correlated way. The additional probability of finding $N$ galaxies within a separation $r$ over a random Poisson process is described by galaxy correlation functions of up to order $N$ (Peebles 2001). We denote the combination of correlation functions describing this extra correlation as $\xi_x^{(N)}(z)$, i.e. the excess probability for two galaxies is given by

$$1 + \xi_x^{(2)}(z_\alpha, z_\beta) \equiv 1 + \xi(r_{\alpha\beta}),$$

where the separation $r_{\alpha\beta} \equiv |\vec{r}_\alpha - \vec{r}_\beta|$ is the comoving distance between components $\alpha$ and $\beta$. In the two-component case, only the two-point correlation function $\xi(r)$ is necessary. However, for three galaxies, higher order correlation functions are needed, i.e.

$$1 + \xi_x^{(3)}(z_\alpha, z_\beta, z_\gamma) \equiv 1 + \xi(r_{\alpha\beta}) + \xi(r_{\beta\gamma}) + \xi(r_{\alpha\gamma}) + \xi(r_{\alpha\beta\gamma}),$$

where $\xi(r_{\alpha\beta}, r_{\beta\gamma}, r_{\alpha\gamma})$ is the connected three-point galaxy correlation function.

The excess probability term $\xi_x^{(N)}(z)$ is defined in the posterior as a function of the component redshifts $\{z\}$, though the galaxy correlation function $\xi$ (and higher order correlations) is defined in terms of comoving separation $r$. We therefore need to convert between the redshifts of each component and the comoving distance separating them. The line-of-sight comoving distance as a function of redshift is given by (e.g. Hogg 1999)

$$r(z) = \frac{c}{H_0} \int_0^z \frac{dz'}{E(z')}.$$

where, neglecting radiation density and rewriting $\Omega \equiv \Omega_m + \Omega_\Lambda$,

$$E(z) = \sqrt{\Omega_m(1 + z)^3 + (1 - \Omega)(1 + z)^2 + \Omega_\Lambda}.$$

We assume a flat Planck $^1$ (Planck Collaboration XIII 2016) cosmology throughout; $\Omega_m = 0.3065$, $\Omega_\Lambda = 0.6935$, and $H_0 = 67.9$ km s$^{-1}$ Mpc$^{-1}$.

However, the comoving distance separating components will depend not only on their redshifts, but also on their angular separation on the sky. As a result, we derive an effective correlation function $\xi_{\text{eff}}$ that takes this angular dependence into account.

Consider the case of a two-component blend, as shown in Fig. 1. The two components are at comoving distances $r_\alpha$ and $r_\beta$ from the observer, with separation $\Delta r \equiv r_\beta - r_\alpha$. From the definition of the correlation function, we can write the ratio of the expected number of galaxies in a region with clustering $N^\alpha$ and that without $N^\theta$ as

$$1 + \xi_{\text{eff}} \equiv \frac{N^\alpha}{N^\theta}.$$

Given that these components are blended, there is some maximum angular separation $\theta$ between them; we assume this to be small. We therefore compare the expected number of galaxies in a disc of width $dr$ and radius $\rho_{\text{max}} = r_\beta \theta$. The expected number without clustering is given by

$$N^\theta_{\text{disc}} = n \pi \sigma^2 \theta^2.$$

To find the expected number with clustering, we integrate over the disc using the volume element of an annulus with radius $\rho$, i.e.

$$N^\theta_{\text{disc}} = \int_{\rho=0}^{r_{\beta}\theta} n[1 + \xi(r)] 2\pi \rho d\rho dr.$$

Thus, writing $r = \sqrt{\Delta r^2 + \rho^2}$, the ratio becomes

$$1 + \xi_{\text{eff}} \equiv \frac{2 \pi \sigma^2}{r_{\beta}^2 \theta^2} \int_{\rho=0}^{r_{\beta}\theta} [1 + \xi(\sqrt{\Delta r^2 + \rho^2})] \rho d\rho.$$

\footnote{We use the TT + low\text{\,}T + lensing + ext values from Table 4.}

Figure 1. Diagram showing the set-up of the $\xi_{\text{eff}}$ calculation. We assume that two galaxies, represented by grey circles, will be blended if their angular separation is within $\theta$. Given that these two galaxies are blended, the galaxy at a comoving distance $r_\beta$ will lie within the disc.

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure1.png}
\caption{Diagram showing the set-up of the $\xi_{\text{eff}}$ calculation. We assume that two galaxies, represented by grey circles, will be blended if their angular separation is within $\theta$. Given that these two galaxies are blended, the galaxy at a comoving distance $r_\beta$ will lie within the disc.}
\end{figure}
As described below, the effect of clustering is small. As a result, we adopt a simple power law for the two-point correlation function,

\[ \xi(r) \propto \left( \frac{r}{r_0} \right)^{-\gamma}. \]  

(20)

Inserting this into equation (19) and integrating, the effective correlation function is given by

\[ \xi_{\text{eff}}(r_{\alpha}, r_{\beta}) = \frac{r_0^2}{(1 - \gamma) r_0^2 \theta^2} \left[ \left( \frac{\Delta r^2 + r_0^2 \theta^2}{r_0^2} \right)^{1-\gamma} - \left( \frac{\Delta r^2}{r_0^2} \right)^{1-\gamma} \right]. \]  

(21)

The effect of the strength of clustering evolving with redshift can be included in this formalism by allowing the parameters \( r_0 \) and \( \gamma \) to vary with redshift (e.g. Soltan 2016). We test the effect of this on the redshift inference by using a toy model where \( \gamma = 1.92 \) is kept constant, while \( r_0 \) linearly varies between \( r_0 = 5 \text{ Mpc} \ h^{-1} \) at redshift \( z = 2 \) and \( r_0 = 6 \text{ Mpc} \ h^{-1} \) at redshift \( z = 0.5 \), with a linear extrapolation outside of this range. Since the value of \( \xi_{\text{eff}} \) is non-negligible only when \( z_\alpha \approx z_\beta \), this interpolation of \( r_0 \) is evaluated using \( z_\alpha \) only.

We then simulated two-component blends from a prior with \( \xi_{\text{eff}} \) included as described in Section 5. Results assuming \( \xi_{\text{eff}} = 0 \) showed negligible differences from those where the effect was included. At the population level, the RMS scatter defined in equation (45) changed by 0.205 per cent between results including and excluding the correlation function. There were also negligible changes to the results at the individual source level. A comparison of maximum a posteriori results in each case is shown in Fig. 2. The vast majority of sources show negligible differences, and visually inspecting the posteriors with larger changes shows these are highly multimodal, with modes of comparable heights. In these cases, small differences in the posteriors result in larger differences in point estimates as the maximum a posteriori value moves between modes. This is a limitation of point estimates, and can be mitigated by using the full information content of the posterior distributions, which do not vary strongly.

Due to the small effect, our results throughout include a simple non-evolving correlation function with \( r_0 = 5 \text{ Mpc} \ h^{-1} \) and \( \gamma = 1.77 \) (Peebles 2001). A plot of this is given in Fig. 3. Inserting the correlation function allows us to write the joint redshift prior separated by component as

\[ P([z]|\{t_i\}, \{m_0\}, x, N) = [1 + \xi_{\text{eff}}^{(N)}([z])] \prod_{\alpha=1}^{N} P(z_\alpha | t_\alpha, m_{0,\alpha}). \]  

(22)

We separate the joint magnitude prior by assuming that the only correlation between the component magnitudes is from the effect of a selection function \( S([m_0]) \) applied to the total magnitude, as discussed in Section 2.4. The magnitude prior can then be written as

\[ P([m_0]|N) = S([m_0]) \prod_{\alpha=1}^{N} P(m_{0,\alpha}). \]  

(23)

Finally, we impose a sorting condition. Without this, the components would be exchangeable, i.e. swapping the component labels \( \alpha, \beta, \ldots \) would have no effect on the prediction of the model. As a result, the marginalized posterior for the redshift of a single component would contain contributions from every component in the source, as demonstrated in Fig. 4. Imposing a sorting condition on either the magnitudes or the redshifts would have the same effect of breaking the exchangeability of the components. In our tests, sorting by redshift produced posteriors that recovered the true redshift more successfully. However, in high-redshift samples, there is an intrinsic colour degeneracy that can occasionally cause problems with a redshift sorting condition.

The Lyman break and Balmer break are absorption features occurring at 912 and 3650 Å, respectively. If photometry over a sufficiently wide wavelength range is not available, a Lyman break at high redshift can be confused with a Balmer break at low redshift (e.g. Graham et al. 2018). If the sample is deep enough that these high-redshift solutions are not unlikely a priori, this can cause bimodal posteriors and contribute to catastrophic outliers (Brimijoule et al. 2008).

Consider the case of a two-component blend where the redshift of one component is well constrained but the other has a bimodal
as surveys typically select sources by imposing cuts on the apparent
account for the selection effect of the survey observing it. Galaxy
is given by sorting is used. In summary, the posterior for the fully blended case
where the redshift object depends on which of the two degenerate peaks is
these two modes, it will appear in the 1D marginal distributions
posterior. If the well-constrained redshift happens to lie between
zα = 0.31, zβ = 1.19 as indicated by the orange lines.
As a result of the exchangeability, the 2D marginal redshift distribution is
symmetric about the dashed black line, and each 1D posterior contains a
distinct peak for each component.

Figure 4. By not imposing a sorting condition, the components in a source
are exchangeable. This is demonstrated here for a simple two-component
blend with redshifts zα = 0.31, zβ = 1.19 as indicated by the orange lines.
As a result of the exchangeability, the 2D marginal redshift distribution is
symmetric about the dashed black line, and each 1D posterior contains a
distinct peak for each component.

posterior. If the well-constrained redshift happens to lie between
these two modes, it will appear in the 1D marginal distributions of each component redshift, as whether it is the lower or higher redshift object depends on which of the two degenerate peaks is
being sampled. In this case, sorting by magnitudes would result in
a posterior more representative of the underlying system, where the
redshift of one component is well constrained while the other has
two well-separated modes. We did not find this to be a problem in
our tests however, and so apply redshift sorting throughout.

The sorting condition Λα is imposed by introducing Heaviside step functions Θ into the product over components, and is defined as

\[
\Lambda_\alpha = \begin{cases} 
\Theta(q_\alpha - q_a) & \text{for } \alpha = 1 \\
\text{otherwise},
\end{cases}
\]

where \(q\) is either \(z\) or \(m_0\) depending on whether redshift or magnitude sorting is used. In summary, the posterior for the fully blended case is given by

\[
P(\{z, [m_0] \} | F, F_0, \chi, N) \propto \sum_{i=1}^{N} P(F | \{z\}, \{t_i\}, \{m_0\}, N) \\
\times P(F_0 | \{m_0\}) \left[ 1 + \frac{\delta(N)(\{z\})}{2} \right] S(\{m_0\}) \\
\times \prod_{\alpha=1}^{N} \Lambda_\alpha P(z_\alpha | t_\alpha, m_{0,\alpha}) P(t_\alpha | m_{0,\alpha}) P(m_{0,\alpha}).
\]

(25)

2.4 Accounting for selection effects

When considering the total apparent magnitude of a source, we must
account for the selection effect of the survey observing it. Galaxy
surveys typically select sources by imposing cuts on the apparent
magnitude they observe \(m < m_{\text{lim}}\) since they cannot observe arbitrarily faint sources. As we are sampling intrinsic magnitudes rather than observed magnitudes, these selection effects do not impose a hard cut in our magnitude prior.

Consider a source with an intrinsic apparent magnitude exactly
equal to the survey magnitude limit. Assuming a normal distribution
for the observational error, the probability of observing this source is
1/2, since its observed apparent magnitude is equally likely to have
been scattered above and below the magnitude cut. However, since
objects in the sample have been detected by definition, we know the source must have been scattered brighter, effectively breaking the
symmetry of the error distribution. As a result, intrinsic apparent magnitudes around the magnitude limit are less probable and should
be downweighted.

To account for this, we follow the approach described in Leist-
edt, Mortlock & Peiris (2016) for including a selection effect. A
discrete variable \(D\) representing the fact that an object was detected
is introduced, and each term in the posterior is conditioned on it.
We assume that our selection effect is imposed on a single selection
band. Without loss of generality, we derive the effect by assuming
that the selection band is the reference band \(\beta\), and so only the reference band likelihood is affected. Conditioning on \(D\), the likelihood can be written using Bayes rule as

\[
P(F_0 | \{m_0\}, N, D) = \frac{P(D | F_0, \{m_0\}, N) P(F_0 | \{m_0\}, N)}{\int_{D} P(D | F_0, \{m_0\}, N) P(F_0 | \{m_0\}, N) \, d\hat{F}_0}.
\]

(26)

The numerator of equation (26) is equal to the likelihood defined
in equation (8) since the probability of detection for an object
that we know has been observed is \(P(D | F_0, \{m_0\}, N) = 1\) after
integrating over \(\hat{F}_0\), the denominator depends only on \(\{m_0\}\) and
represents the effect of the magnitude selection. We therefore choose
to write this term as part of the joint magnitude prior, defining the
selection effect

\[
S(\{m_0\}) = \int_{10^{-0.4m_{\text{lim}}}}^{\infty} P(D | F_0, \{m_0\}, N) \, d\hat{F}_0
\]

(27)

that appears in the posterior in equation (25). The selection is a hard
cut based on the observed flux, and so

\[
P(D | F_0, \{m_0\}, N, \cdot) = \begin{cases} 
1 & \text{for } \hat{F}_0 > 10^{-0.4m_{\text{lim}}} \\
0 & \text{otherwise}.
\end{cases}
\]

(28)

Thus, the integral becomes

\[
S(\{m_0\}) = \int_{10^{-0.4m_{\text{lim}}}}^{\infty} P(\hat{F}_0 | \{m_0\}) \, d\hat{F}_0.
\]

(29)

Since the reference band likelihood is assumed Gaussian, this
can be written in terms of the normal cumulative distribution function
as \(S(\{m_0\}) = 1 - \Phi(\hat{F}_0)\), where \(\Phi\) is defined for a Gaussian
distribution with mean \(\mu\) and standard deviation \(\sigma\) to be

\[
\Phi(x) = \frac{1}{2} \left[ 1 + \operatorname{erf} \left( \frac{x - \mu}{\sigma \sqrt{2}} \right) \right].
\]

(30)

Inserting this into equation (29), the effect of the magnitude selection can be written as

\[
S(\{m_0\}) = \frac{1}{2} - \frac{1}{2} \left[ 1 + \operatorname{erf} \left( \frac{10^{-0.4m_{\text{lim}}} - \sum_{\alpha=1}^{N} 10^{-0.4m_{0,\alpha}}}{\sigma_0 \sqrt{2}} \right) \right].
\]

(31)

By replacing the reference-band flux \(\sum_{\alpha=1}^{N} 10^{-0.4m_{0,\alpha}}\) with the model flux \(F_{[\alpha]}^{(N)}(\{z\}, \{m_0\})\), this selection can be performed on
discussed in Section 2.6. This was found to be $\phi = 0.705$, though the difference in results compared to fixing $\phi = 0.6$ was negligible.

Since the selection effect applies to the total source flux, individual components may be fainter than the survey magnitude limit, and so unobservable outside of a blended source. As a result, an analytic magnitude prior is required to describe the distribution of component magnitudes so that it can be used at faint magnitudes, where observations of individual components are unavailable.

For the reasons discussed in Section 4.2, we also apply a hard minimum and maximum cut to each component redshift $z_\alpha$ and component magnitude $m_{0,\alpha}$. This cut has little effect on the redshift priors, which already go towards zero at large redshift; the same is true of the magnitude prior at bright magnitudes. The faint end of the magnitude prior of the brightest component is also already forced towards zero by the selection function. This is because in a $N$-component blend, the flux of the brightest component must be at least $1/N$ that of the total source flux, by definition. The magnitudes of the other components are not constrained in this way however, and so this cut represents a sharp boundary in the prior.

In our tests, the results of the redshift estimation were not strongly dependent on the position of this faint-end magnitude cut $m_{\text{max}}$. However, the evidence calculation described in Section 4.1 is dependent on its position, as changing the position of the cut alters the prior volume integrated over in equation (42). As a result, the position of this cut must be decided; it defines the limit where a galaxy is considered to contribute to a blend, and is therefore problem-dependent.

In principle, one could consider a galaxy to be blended if another arbitrarily faint galaxy lies along the same line of sight. In practice however, observations have limited precision, and the flux of an extremely dim galaxy cannot be detected. In other words, a sufficiently dim galaxy should no longer be considered a blended component, but rather a contribution to the noise.

In practice, a simple method to set this cut is to fix it for the entire sample. However, the argument above suggests that this cut should be dependent on the noise of the observation, i.e. that $m_{\text{max}}$ should be set to the faintest magnitude that would have an observable effect. Fixing $m_{\text{max}}$ is effectively an assumption that the sample has sufficiently homogeneous noise properties that the change in this faintest magnitude is negligible. For a sample where this is not the case, the magnitude cut can be set as a $n\sigma_0$ flux deviation, i.e.

$$m_{\text{max}} = -2.5 \log_{10}(n\sigma_0),$$

where $\sigma_0$ is the error on the reference-band flux that varies for each source. In the tests in Section 6, we test both of these methods of setting $m_{\text{max}}$.

### 2.5 Specifying the priors

Like all Bayesian methods, the choice of priors should be problem dependent. For ease of comparison, we use the parametric forms given by Benítez (2000) with an additional magnitude prior. However, we stress that this choice is not a necessary one for our method and any joint $P(z, t, m_0)$ prior may be used.

The Benítez (2000) template and redshift priors are given by

$$P(t|m_0) = f_t e^{-k_t(m_0-m_{\text{min}})}$$

and

$$P(z|t, m_0) \propto z^{\alpha_t} \exp\left\{-\frac{z}{z_0 + k_{\text{red}}(m_0 - m_{\text{min}})}\right\},$$

respectively, where $m_{\text{min}}$ is the bright-end magnitude cut as described below. The parameters $\alpha_t$, $z_0$, $k_{\text{red}}$, $f_t$, and $k_t$ are set separately for early, late, and irregular template types. Their values are found using the procedure discussed in Section 2.6 and are listed in Table 2.

We use a magnitude prior given by

$$P(m_0) \propto 10^{m_{\text{min}}}.$$

The value $\phi = 0.6$ gives the expression for the expected galaxy number counts in a homogeneous, Euclidean universe (Yasuda et al. 2001), though we leave $\phi$ free to also be found using the procedure any band. The selection function would then also be dependent on the redshifts and templates, i.e. $S(z, t, |m_0|)$. This choice of selection band is included in the implementation described in Section 4.3.

A plot of this selection function for a galaxy from the GAMA blended sources catalogue, described in Section 6, is shown in Fig. 5.

### 2.6 Calibrating the priors using spectroscopic information

The joint prior is conditioned on a set of parameters $\theta$, i.e. $P(z, t, m_0|\theta)$, the posterior distribution of which we wish to infer. We can use spectroscopic information of a sample of galaxies from the population of interest to calibrate the above priors as suggested by Benítez (2000).

We assume here that this calibration is done with unblended galaxies, though this procedure can be extended to include blended galaxies too, provided that the number of components $N$ is known a priori. In that case, the reference band magnitudes of each component would need to be included as a parameter in this model, and either sampled along with $\theta$ or marginalized out of the posterior analytically.

---

### Table 2. The maximum a posteriori values of the prior parameters for the GAMA blended sources catalogue after calibrating using 26 782 unblended sources.

| Parameters | Early | Late | Irregular | Type-independent |
|------------|-------|------|-----------|------------------|
| $\alpha_t$ | 1.59  | 1.53 | 1.30      | –                |
| $z_0$      | 0.016 | 0.019| 0.066     | –                |
| $k_{\text{red}}$ | 0.048 | 0.048| 0.022     | –                |
| $k_t$      | 0.044 | 0.024| –         | –                |
| $f_t$      | 0.45  | 0.51 | –         | –                |
| $\phi$     | –     | –    | –         | 0.71             |
We consider a sample of \( G \) galaxies with photometry and spectroscopic redshifts \( \hat{z}_g \). These redshifts are assumed to be exact, i.e. we neglect the error on \( \hat{z}_g \). The set notation here now runs over each independently observed galaxy, not the blended components as before.

We start by writing this posterior as a marginalization over the photometric redshift model parameters for each galaxy and applying Bayes rule. Since the likelihood is independent of the prior parameters, we condition on \( \hat{z}_g \), the set notation here now runs over each independently observed galaxy, not the blended components as before.

We can then write the blended flux of \( g \) as

\[
P(\theta | \{ \hat{z}_g \}, \{ \hat{F}_g \}, \{ \hat{F}_0 \}) \propto \int d\hat{F}_g \int d\hat{F}_0 \delta(m_0)
\]

\[
\times \sum_{i=1}^{T} P(\{ \hat{z}_g \}, \{ \hat{F}_g \}, \{ \hat{F}_0 \}) | (\{ \hat{z}_g \}, \{ \hat{F}_g \}, \{ \hat{F}_0 \})\rangle P(\theta | \{ \hat{z}_g \}, \{ \hat{F}_g \}, \{ \hat{F}_0 \}).
\]

(36)

We apply product rule to separate the joint prior and remove unnecessary conditioning. We also assume that the galaxies in the sample are independent, and so all terms not shared across the population [i.e. \( P(\theta) \)] can be written as a product over galaxies. The posterior then becomes

\[
P(\theta | \{ \hat{z}_g \}, \{ \hat{F}_g \}, \{ \hat{F}_0 \}) \propto P(\theta) \prod_{g=1}^{G} \int d\hat{z}_g \int dm_0 \delta(m_0)
\]

\[
\times \sum_{t_g=1}^{T} P(\hat{F}_0 | m_0) P(\hat{F}_g | \hat{z}_g, t_g, m_0)
\]

\[
\times P(\hat{z}_g | m_0) P(\hat{z}_g, t_g, m_0 | \theta).
\]

(37)

By assuming that the spectroscopic redshifts are exact, the redshift likelihood can be written as a delta function, i.e. \( P(\hat{z}_g | m_0) = \delta(\hat{z}_g - z_g) \). We also assume that the error on the reference band magnitude is negligible, allowing us to write \( P(\hat{F}_0 | m_0) = \delta(m_0 - \hat{m}_0) \), where \( \hat{m}_0 = -2.5 \log_{10} (\hat{F}_0) \) is the reference-band flux of galaxy \( g \), converted to magnitudes. Replacing these likelihoods with delta functions, the marginalization can be done analytically using the sifting property of the delta function to give

\[
P(\theta | \{ \hat{z}_g \}, \{ \hat{F}_g \}, \{ \hat{F}_0 \}) \propto P(\theta) \prod_{g=1}^{G} \sum_{t_g=1}^{T} P(\hat{F}_g | \hat{z}_g, t_g, \hat{m}_0) P(\hat{z}_g, t_g, \hat{m}_0 | \theta).
\]

(38)

To find the prior parameters \( \theta \) that maximize this posterior, we use L-BFGS-B (Byrd et al. 1995), a local optimization algorithm that approximates the Hessian of the objective function and optimizes the parameters subject to simple box constraints; we use these constraints to ensure our parameters are positive. This method requires first-order derivatives, which we approximate through a finite-difference method.

The result of this procedure is an estimate of the maximum a posteriori values of the prior parameters. Throughout this paper, we use these values in the priors directly. In principle, these parameters could form part of a hierarchical model and be marginalized out as nuisance parameters. However, this would significantly increase the dimensionality of the parameter space to be sampled and, thus, the computation time required for each source. Table 2 lists the values of these prior parameters GAMA test described in Section 6. A plot of samples drawn from the resulting prior is shown in Fig. 6.

3 PARTIALLY BLENDED SOURCES

We can modify the formalism above for the case of sources for which every component does not contribute to every observation. We refer to these as partially blended sources. This can be the case when combining photometry from a wide range of wavelengths, e.g. optical and far-infrared observations. This partial blending may also occur for some sources observed in both ground-based and space-based surveys, as the latter does not suffer from atmospheric seeing and so can achieve a higher spatial resolution. An example of a pair of such surveys is LSST (Ivezic et al. 2008) and Euclid (Laureijs et al. 2011). Utilizing resolved photometry from Euclid could improve the precision of photometric redshifts of sources that are blended in the higher signal-to-noise observations of LSST. This possibility is explored using simulated observations in Section 5.2.

To generalize the method for this case, we introduce the measurement-component mapping \( \delta_{a,m} \), a \( N \times N_m \) matrix, where \( N_m \) is the number of measurements, a generalization of the number of bands in the fully blended case. This measurement-component mapping acts as an indicator variable, consisting only of zeros and ones indicating whether a particular component is present in a particular measurement.

An example of such a matrix is given below. Consider data containing \( N_m = 6 \) photometric measurements of \( N = 2 \) components. The first four measurements are of individually resolved components, while the final two measurements are blended. In a typical use case, we might expect the resolved measurements of each component to share filter bands, though the model does not require this. In this example, the measurement-component mapping is given by

\[
\delta = \begin{bmatrix}
1 & 0 & 1 & 0 & 1 & 1 \\
0 & 1 & 0 & 1 & 1 & 1
\end{bmatrix}.
\]

(39)

We can then write the blended flux of \( N \) components at a redshift \( z \) as

\[
F^{(N)}_{\chi, m, a}(z, \{ m_0 \}) = \sum_{a=1}^{N} \delta_{a,m} 10^{-0.4 m_{0,a}} T_{\alpha,a} T_{\gamma,a}(z_a).
\]

(40)

The only modification to the posterior of the fully blended case needed to accommodate the partial blending is to the sorting condition. As described in Section 2.3, the purpose of this condition is to prevent the exchangeability of components. However, this is not necessary in the partially blended case. Here, the components are intrinsically different as they appear individually in separate measurements and so are not exchangeable. As a result, we drop the sorting condition for the partially blended case, i.e. \( \Lambda_a = 1 \) over the entire parameter space. The posterior for the partially blended case is then given by

\[
P(\{ z \}, \{ m_0 \} | \hat{F}_g, \hat{F}_0, \chi, N, \delta) \propto \sum_{i=1}^{N} P(\hat{F}_i | \{ z \}, \{ t_i \}, \{ m_0 \}, N, \delta)
\]

\[
\times P(\hat{F}_0 | \{ m_0 \}) \left[ 1 + \xi(N) | \{ z \} \right] S(\{ m_0 \})
\]

\[
\times \prod_{a=1}^{N} P(z_a | t_a, m_{0,a}) P(t_a | m_{0,a}) P(m_{0,a}).
\]

(41)
2496  D. M. Jones and A. F. Heavens

Figure 6. Plot of the prior found for the test on the GAMA blended sources catalogue after calibrating using 26 782 unblended sources. The dashed line in the bottom panel shows a magnitude limit of \( r < 19.8 \).

4 INFERENCE USING NESTED SAMPLING

4.1 Determining the number of components with model comparison

The posteriors in equations (25) and (41) are defined for a specific number of components \( N \). In general however, this number of components is not known a priori. We therefore need a method to determine how many components are present in a source. Since our model is defined for a fixed number of components, we treat finding the number of components in a source as a model comparison problem.

Bayesian model comparison involves the calculation of the evidence \( Z \), an integral over the product of the prior and the likelihood (e.g. Trotta 2008). Given a data vector \( d \), a model \( m \), and a set of model parameters \( \{\theta\} \), the evidence is defined as

\[
Z \equiv P(d|m) = \int P(d|\{\theta\}, m)P(\{\theta\}|m)d\theta.
\]

(42)

This evidence term plays the role of the normalization of the posterior and so is typically ignored in parameter inference problems where this normalization is irrelevant. However, the evidence is the quantity of interest for model comparison problems. The ratio of the posterior probabilities of two models is proportional to the ratio of their evidences, a quantity known as the Bayes factor. By considering the number of components in a source as the model, we can write the relative probability of the source containing \( n \) components compared to \( m \) components as

\[
P_{n,m} = \frac{P(N = m | \hat{F}, \hat{F}_0)}{P(N = n | \hat{F}, \hat{F}_0)} = \frac{P(\hat{F}, \hat{F}_0 | N = n)}{P(\hat{F}, \hat{F}_0 | N = m)} \cdot \frac{P(N = n)}{P(N = m)}.
\]

(43)

Considering the cases of either isolated galaxies or blends of two components, the model prior ratio \( P(N = 2) / P(N = 1) \) represents the probability that a galaxy will be blended. Dawson & Schneider (2014) estimate the number of sources observed by LSST that will be blended by convolving Hubble Space Telescope images with a Gaussian point spread function (PSF) like that of LSST. They found this number to be 45–55 per cent of the total sources observed, with 15–20 per cent of observed sources classified as catastrophic blends that would be identified as single sources by fitting a
profile template to a galaxy image. Chang et al. (2013) estimate that the rejection of blended sources will reduce the number density of LSST sources by 16 per cent, though this estimate does not include the catastrophic blends of above. Studies such as these using existing high-resolution data or simulated observations can inform the blending prior ratio. Throughout this paper, we present results where this prior ratio is \( P(N = 2)/P(N = 1) = 1 \), i.e. we do not prefer either of the blended or single-component models a priori, though this information can be trivially included.

### 4.2 Nested sampling using MULTINEST

Calculating the evidence directly through numerical integration presents a difficult technical problem, particularly as the number of dimensions increases. To avoid this, we use Nested sampling (Skilling 2006), a Monte Carlo method for estimating the evidence while also sampling the posterior for parameter inference. Nested sampling reduces the problem of estimating the evidence to sampling a series of increasing likelihood thresholds, i.e. progressively smaller prior volumes nested within one another. Equation (42) can then be calculated using a one-dimensional quadrature integration method over this prior volume.

The computationally difficult part of the nested sampling algorithm is sampling a new point from within the potentially complicated boundary defined by the likelihood threshold. The MULTINEST sampler (Feroz, Hobson & Bridges 2009) does this efficiently by sampling from a collection of ellipses approximating this boundary rather than the prior itself. This collection of ellipses is formed by performing a clustering analysis on a fixed-sized set of the previous samples, known as the five points. A new sample is drawn from these ellipses, replacing the lowest likelihood point that is removed and stored as a posterior sample. Samples are rejected until the likelihood boundary is respected, though this occurs less frequently than when naively rejection sampling the prior.

The use of multiple ellipses when sampling has another distinct advantage in that it naturally enables efficient sampling of multimodal posteriors, since each mode is assigned a separate ellipse while low-probability regions between these modes are avoided. Multimodality is a feature that can cause difficulties for Markov Chain Monte Carlo (MCMC) samplers, as moving from one mode to another requires a move across the low-probability region separating them. As a result, these samplers can fail to explore the full posterior distribution, instead sampling only a single mode. We expect our problem to exhibit this multimodal behaviour due to the degeneracies described in Section 2.3, and so require a sampling method suited to this case.

The need for nested sampling methods to sample from the prior imposes some constraints on our choice of prior. MULTINEST natively samples from a unit side-length hypercube and these samples are transformed into samples of the prior using a prior transform function. However, due to the discrete marginalization over template, we cannot separate the posterior to define a prior transform function. As a result, we take the approach suggested by Feroz et al. (2009) of defining a uniform prior to sample from, and defining the “likelihood” for MULTINEST as our marginalized posterior.

This has two main effects. First, the sampling is likely to be less efficient, as the prior sampling step is not guided by the true prior, and so low-prior regions may be sampled frequently. Secondly, sampling from a uniform prior necessitates imposing a hard cut on the prior range of each parameter. Since the location of these cuts affects the value of the evidence \( Z \), they should not be imposed thoughtlessly. At high redshift and bright magnitudes, the priors tend to zero, meaning that the exact positions of these cuts have negligible effect on the evidence. However, this is not the case for the faint end of the magnitude priors; setting this cut is discussed in Section 2.5.

### 4.3 blendz package

We have written a Python package blendz to perform the redshift inference of blended sources described in Sections 2 and 3, and the identification of the number of components using model comparison described in Section 4.1. The package supports analysis of blends with an arbitrary number of components using either the included or user-supplied template sets. The output of such an analysis is a set of samples from the joint posterior for each number of components considered, and an estimate of the Bayes factor for model comparison. The model comparison can then easily include a model prior through multiplication of the Bayes factor.

The package is also written in an object-oriented way, allowing the user to easily redefine the priors. While the supplied prior is used in this work with galaxies of either early, late, or irregular types, it is written to be calibrated and used with any number of possible types. For blended sources more than two components, the excess probability term \( \xi_f^{(N)} \) is defined recursively to use the correct combination of two-point terms and assumes higher order correlations are negligible.

Documentation and instructions for installation can be found at [http://blendz.readthedocs.io](http://blendz.readthedocs.io). The package can also be immediately installed from the official Python Package Index² by using the `pip install blendz` command. Finally, the source is available in a git repository hosted at [https://github.com/danmichaeljones/blendz](https://github.com/danmichaeljones/blendz).

## 5 RESULTS FROM MOCK OBSERVATIONS

### 5.1 Fully blended sources

As an initial test of the method, we used a Monte Carlo simulation to create a set of mock photometric observations to test our method against. These mock observations simulate an optical survey using the six LSST optical filters \( u, g, r, i, z, Y \) (LSST Science Collaboration 2009), with an \( r \)-band magnitude selection of \( m_{\text{lim}} = 24 \). We also applied hard cuts to the component magnitudes of \( m_{\text{min}} = 19 \) and \( m_{\text{lim}} = 26 \). We then generated 1000 sources, each of which is a blend of two components in all bands. This was done by sampling a prior describing this distribution of objects using the MCMC sampler emcee (Foreman-Mackey et al. 2013) to generate the true parameters \( \{z\}, \{i\}, \{m_0\} \) for each simulated source. A plot of this prior distribution, plotted using `corner.py` (Foreman-Mackey 2016), is shown in Fig. 7.

The effect of the selection function and the faint-end magnitude cut can be seen clearly in the two-peaked shape of the marginal distributions of \( m_{\alpha,0} \) and \( m_{\alpha,1} \). The brighter-magnitude peak is a result of the selection function. In the single-component case, this would cause the prior to tend to zero at faint magnitudes. In the two-component case however, the magnitude priors of each component extend beyond \( m_{\text{lim}} \) as the selection effect is applied to the combined

²[https://pypi.org/]
Weedman (1980), two starburst templates from Kinney et al. (1996), the template set of Coe et al. (2006), containing one early-type, two marginal distributions as a hard diagonal cut. The redshift sorting condition can be seen as a hard diagonal cut in the marginal distributions of \(z_\alpha\) and \(z_\beta\) by redshift. The redshift sorting condition can be seen as a hard diagonal cut in the marginal distributions of \(z_\alpha\) and \(z_\beta\). This component blend must, by definition, contribute at least half of the total flux. As a result, the selection effect prevents the magnitude of this component from being too faint. Since we impose the sorting condition on redshifts, the brightest component in a source is not exclusively the lower redshift one, this action of the sorting condition causes the brighter peak in the marginal distributions of both \(m_\alpha\) and \(m_\beta\). If we instead impose the sorting condition on the magnitudes, these distributions become unimodal. Fig. 7 also shows the effect of the redshift sorting condition in the \((z_\alpha, z_\beta)\) marginal distribution as a hard diagonal cut.

The model fluxes for these sampled parameters were then generated using the template responses defined in Section 2.1. We use the template set of Coe et al. (2006), containing one early-type, two late-type, and one irregular-type templates from Coleman, Wu & Weedman (1980), two starburst templates from Kinney et al. (1996), and two starburst templates from Coe et al. (2006). This same template set is then used during the inference. This allows a test of the method without the effect of unrepresentative templates, a source of error that is not unique to the case of blended sources.

Finally, we add an observational error to each observation. The flux error in band \(b\) is randomly drawn from an uncorrelated, zero-centred normal distribution \(\sigma_b \sim N(\sigma_0, \Sigma)\). The noise is set for all sources to be the final 1\(\sigma\) depth expected from LSST (Ivezic et al. 2008). We use these noisy observations to draw samples from the one- and two-component posteriors to test both the redshift determination and model comparison performance, setting the prior to the true distribution the photometry was sampled from.

Fig. 8 shows two examples of the 4D posterior that is the output from our method for each sample. For plotting purposes, the number of live points used for sampling is larger than that used for the inference and model comparison results throughout this paper. However, the change in the results is negligible. The left-hand panel shows an example of a well-constrained source with a unimodal posterior. This posterior shows correlations between the component parameters; this is expected, since the total flux of each band that is well constrained by the observations is split between the components. Reducing the model flux in a band of one component will result in a compensation in the other component, correlating their parameters.

The right-hand panel of Fig. 8 shows a particularly prominent example of the curved degeneracies that can arise in the blended posteriors. This is due to the total magnitude of the source being well constrained by noisy observations, while component magnitudes are not themselves observable. This leads to a degeneracy that is curved due to the non-linearity of adding magnitudes. A result of this curved degeneracy is bimodality in the marginalized posterior of \(z_\beta\). However, there is still significant probability density around the true redshift, highlighting the importance of not compressing the information content of a full posterior distribution into only a small set of numbers.

Fig. 9 shows a comparison of the photometric estimation of the component redshifts against their true simulated values. Point estimates of the redshift \(z_{\text{MAP}}\) are obtained by taking the maximum a posteriori (MAP) value of each component redshift posterior, marginalizing over the other three parameters. The method recovers the true redshift of each component from simulated photometry well. The performance of photometric redshift methods is often summarized by the RMS scatter \(\sigma_{\text{RMS}}\). We first define the normalized error for galaxy \(g\)

\[
\delta z_g = \frac{\tilde{z}_g - z_{\text{MAP},g}}{1 + \tilde{z}_g},
\]

where each galaxy \(g\) is a single component of a blended source. Writing the total number of galaxies in our test catalogue as \(N_g\), we then define the RMS scatter as

\[
\sigma_{\text{RMS}} = \sqrt{\frac{1}{N_g} \sum_g (\delta z_g)^2}.
\]

Computing this quantity for our mock blended observations, we find an RMS scatter of \(\sigma_{\text{RMS}} = 0.163\). This compares to a scatter of \(\sigma_{\text{RMS}} = 0.0267\) for photometric redshifts of mock observations of single sources, demonstrating the added difficulty of blending.

This scatter can be improved by excluding sources with photometric redshifts that, using the uncertainty information of the posterior distribution, are identified as untrustworthy. This is done by comparing a summary statistic against a threshold that controls the stringency of the test; we use the standard deviation of redshift marginal-posterior samples \(\sigma_a\) separately for each component, though a variety of summary statistics are available. Keeping only sources with \(\sigma_a < 0.2\sigma\alpha\), the RMS scatter is reduced to \(\sigma_{\text{RMS}} = 0.064\), with 37 per cent of sources removed. The effect of this is shown in the bottom row of Fig. 9.

The percentage of outliers can also be quantified. Outliers are defined as sources where either component has an error \(|z_{\text{MAP}} - \tilde{z}_g| \geq 0.15(1 + \tilde{z}_g)\). For the full set of mock observations, this percentage was found to be 18.6 per cent. By keeping only sources with \(\sigma_a < 0.2\sigma\alpha\) as described above, the percentage of outliers falls to only 6.0 per cent.

The results of the detection of blends are also shown in the centre panels of Fig. 9. By using equation (43), we calculate \(P_{2,1}\), the relative probability that a source is a two-component blend compared to a single source. The interpretation of this probability is problem-dependent; a probability of \(\ln(P_{2,1}) > 0\) indicates a preference towards the source being blended, while a threshold to \(\ln(P_{2,1}) > 5\)
Figure 8. The 4D posterior distribution output from our method for two example sources. The true parameter values are shown in orange. The left-hand panel shows a well-constrained source with some correlations between components, though the true redshift is well recovered. The right-hand panel shows an example of a bimodal posterior that can arise in photometric redshift problems.

Figure 9. Scatter plot comparing the maximum a posteriori point estimates from the photometric redshift estimation with the true redshifts for the mock observations. The left-hand panels distinguish the components, with $z_\alpha$ plotted with closed blue markers and $z_\beta$ plotted with open green markers. The centre panels show the blend identification, with sources identified as blends plotted with closed purple markers, and those misidentified as single sources plotted with open red markers. The right-hand panels show a 2D histogram of the combined sample. The panels in the top row show the results for the full mock catalogue, while the bottom row only includes sources where the standard deviation of samples from each redshift marginal posterior is sufficiently small, $\sigma_\alpha \leq 0.2\sigma$. The dashed lines in each panel show an error of $0.15(1 + z)$.
indicates strong evidence (Kass & Raftery 1995). Likewise, probabilities of \( \ln(P_{2,1}) < 0 \) and \( \ln(P_{2,1}) < -5 \) indicate a preference and strong evidence for the single source case, respectively. As the blended and single thresholds are pushed more positive and negative, respectively, there are sources with values of \( \ln(P_{2,1}) \) that fall between these thresholds. In these cases, the source is assigned neither label.

As described in Section 4.1, we assume the relative prior probability of a blend to be \( P(N = 2)/P(N = 1) = 1 \), i.e. we give no preference to either model. Under this assumption, the method identifies 92.7 per cent of sources as blends and 7.3 per cent as single sources. Increasing the threshold to strong evidence, we find that 89.9 per cent of sources are identified as blends and 0.2 per cent as single sources; the remaining 9.9 per cent fall between these thresholds. The distribution of the relative probability of blending and the effect on blend identification of changing the threshold are shown in Fig. 10.

These results show that the method can both recover the redshifts from broad-band observations of blended objects, and detect the blending of a large fraction of these objects from their photometry alone. In addition, the output from these tests is not just point estimates of redshifts, but the full four-dimensional posterior distributions that capture the correlations between components that can be lost by working with component-separated maps. These are the results of simulated observations, however; real data has the complications that capture the correlations between components that can be lost by working with component-separated maps. These are the results of simulated observations, however; real data has the complication that the flux model is no longer exact, i.e. the templates are not perfectly representative of all galaxies observed. As such, we test the method on real data in Section 6.

5.2 Partially blended sources

To test the effect of adding resolved data, we created a set of mock photometric observations of two-component partially blended systems. These observations simulate the same six-band optical survey as described in Section 5.1, combined with a four-band optical and infrared space-based survey using the Euclid filters \( v, Y, J, H \) (Racca et al. 2016). This latter survey is assumed to have made resolved measurements of each component, while the former is fully blended as before. Thus, our partially blended data vector contains 14 fluxes for each source.

For a comparison with the partially blended results, we repeat the inference several times. First, we compare against the fully blended LSST-like case described above. Next, we compare to an inference using the resolved Euclid bands only, testing the effect of removing the difficulty of blending but using lower signal-to-noise data. Finally, we test against the case of using both the LSST- and Euclid-like data, but assuming that sources are blended in all bands. This allows us to separate the improvement as a result of adding resolved data from that of simply having more bands available.

For the fully blended bands, we reuse the simulated fluxes described in Section 5.1. For the resolved bands, we generate observed fluxes using the same randomly sampled source parameters. The observed fluxes are then generated using the flux model described in Section 2.1 with added observational errors drawn randomly from an uncorrelated, zero-centred normal distribution. The noise in these resolved bands is set to the final 1σ depths expected from Euclid observations (Laureijs et al. 2011).

We use the same prior as described in Section 5.1 for both the simulation and inference steps. The reference band over which the prior is defined is set to be the \( r \) band of the blended observations. This band is not present during the inference step using only the resolved data. As a result, we use the flux model from Section 2.1 to convert between \( r \) - and \( Y \)-band magnitudes before evaluating the prior.

Fig. 11 shows a comparison of the photometric redshift point estimates with the true simulated values for the four sets of inferences. As before, these point estimates are the MAP values of the marginal redshift posteriors.

The left-hand panel shows the fully blended case, the same results as Section 5.1. Again, the redshifts of many components are well recovered, though there is a significant fraction of outliers. The RMS scatter in this case was found to be \( \sigma_{\text{RMS}} = 0.163 \). The percentage of outliers, sources where either component has an error \( |z_{\text{MAP}} - z_s| \geq 0.15(1 + z_s) \), was found to be 18.6 per cent.

The centre-left panel of Fig. 11 shows the results for the resolved observations. Though finding the photometric redshift of resolved sources is an easier inference problem, this is counteracted by the significant reduction in the signal-to-noise of this data. As a result, we find an RMS scatter of \( \sigma_{\text{RMS}} = 0.212 \) with 55.0 per cent of sources marked as outliers.

The centre-right panel of Fig. 11 shows the results for combination of LSST- and Euclid-like data in the fully blended case. We find that the addition of the four Euclid bands significantly improves the precision of the redshift inference, which has an RMS scatter of \( \sigma_{\text{RMS}} = 0.073 \). The fraction of outliers has also improved significantly, with only 6.6 per cent of sources marked as outliers.

Finally, the right-hand panel of Fig. 11 shows the results for the partially blended case, combining the high-precision blended observations with the resolved data. Here, we find that the RMS
scatter has reduced to $\sigma_{\text{RMS}} = 0.065$, a factor of 2.5 improvement over the blended LSST-like data alone, and a factor of 1.2 over the combined LSST- and Euclid-like blended data. The percentage of outliers has also been reduced. Here, only 3.4 per cent of sources are found to be outliers, a factor of 5 improvement over the fully blended case, and a factor of 1.9 improvement over the combined LSST- and Euclid-like blended data.

While the most significant improvement was obtained through the increase in the number of bands, these results show that the quality of photometric redshifts of blended sources can be improved through the inclusion of resolved data. This is particularly apparent in the reduction of outliers. One advantage conferred by the addition of resolved data is a constraint on the relative magnitudes of blended components. In the fully blended case, the reference-band magnitude of each of the components must be inferred from the combined magnitude of the blended source only. This can lead to the degenerate distributions shown in Fig. 8. Adding resolved photometry can help to break this degeneracy by providing information about each component individually. The precision of the photometric redshift inferences is therefore improved.

An example of this phenomenon is shown in Fig. 12. The left-hand panel shows a corner plot of the posterior distribution for a fully blended source. The marginal distributions for each component redshift are highly multimodal, with well-separated redshifts occurring at distinct magnitudes. Though there is significant posterior density around the true redshifts, the MAP point estimate of $z_\beta$ would show a significant error, as the incorrect mode has a higher posterior. The right-hand panel of Fig. 12 shows the same source analysed in the partially blended case after the addition of the resolved photometry. Here, the width of the posterior has been significantly reduced by the removal of the incorrect mode. The posterior now shows that the redshift of the source has been well constrained, and the redshift point estimates would no longer have a large error.

### 6 GAMA BLENDED SOURCES CATALOGUE

The GAMA survey (Baldry et al. 2017) is a spectroscopic galaxy survey that observed 286 deg$^2$ of sky over several regions to a magnitude limit of between $r < 19$ and $r < 19.8$. In doing so, it obtained precise redshifts of > 150,000 sources. The observed regions were chosen to overlap with existing imaging surveys such as Sloan Digital Sky Survey (SDSS) (Stoughton et al. 2002) and VISTA Kilo-degree Infrared Galaxy (VIKING) Survey (Edge et al. 2013). As a result, the spectroscopic data are accompanied by a set of aperture-matched photometry covering nine filter bands $u, g, r, i, z, Y, J, H, K$ from optical to infrared wavelengths (Hill et al. 2011).

The GAMA blended sources catalogue (Holwerda et al. 2015) contains 280 sources from the GAMA survey that have been spectroscopically identified as blended objects. These were selected using an automated template-based spectrum fitting method (Baldry et al. 2014) that cross-correlates galaxy templates with the observed spectra to determine the galaxy redshift. Sources where two different redshifts showed strong cross-correlations were visually inspected, resulting in a selection of blended galaxies. The motivation of Holwerda et al. (2015) was the identification of strong lens candidates. However, a catalogue of spectroscopically identified blended galaxies with accompanying nine-band photometry gives us a useful test case for the blended photometric redshift estimation method on non-simulated photometry with secure redshifts available for both components.

We first calibrate the prior using the procedure described in Section 2.6. To do this, we used 26,782 unblended, well-observed galaxies. These were selected by enforcing every band to be free from SExtractor (Bertin & Arnouts 1996) error flags and excluding all galaxies in the blended source catalogue. The resulting prior from the calibration procedure is shown in Fig. 6. As discussed in Section 2.5, we test two methods of setting the faint-end magnitude cut $m_{\text{max}}$, first as a 5$\sigma$ flux deviation using equation (35), and secondly, fixing $m_{\text{max}} = 20.8$. Throughout, we refer to these as the sigma-$m_{\text{max}}$ case and fixed-$m_{\text{max}}$ case, respectively. We then proceed with the inference using the same template set described in Section 5.

The resulting redshift point estimates are shown in Fig. 13. While noisier than the simulated case, the method still recovers reasonable estimates; using equation (45), we find an RMS scatter of $\sigma_{\text{RMS}} = 0.156$ in both the sigma-$m_{\text{max}}$ and fixed-$m_{\text{max}}$ cases. This compares to the scatter for a set of unblended GAMA sources of $\sigma_{\text{RMS}} = 0.116$. Obtaining this value even without the added complication of blending suggests a mismatch between the sources and the template set.

We also compute the inferred blend probability $P_{2:1}$ for these galaxies. The distribution of these probabilities is shown in Fig. 14. As described in Section 5, $\ln(P_{2:1}) > 0$ and $\ln(P_{2:1}) > 5$ show a preference and strong evidence for a blended source, respectively.

3The templates used to fit the spectroscopic redshifts as described in Holwerda et al. (2015) do not cover the full wavelength range of the photometry. As a result, we do not use them for the photometric redshift inference.
while \( \ln(P_{2,1}) < 0 \) and \( \ln(P_{2,1}) < -5 \) show the same for the single source case. The distribution of the blend probability and the effect of the evidence threshold on blend identification are shown in Fig. 14.

In our tests of the \( \sigma_{\max} \) case, 71.6 per cent of sources showed a preference for being blended, with 28.4 per cent preferring a single source. At the strong evidence threshold, 60.4 per cent of sources are correctly identified as blends, with 18.2 per cent misidentified as single sources. The threshold to exclude misidentified sources completely in the \( \sigma_{\max} \) case is \( |\ln(P_{2,1})| < 12.5 \), with 50.7 per cent of sources identified as blends at this level.

The identification of blends was very similar in the \( m_{\max} \) case. We found that 71.1 per cent of sources showed a preference for being blended, and 28.9 per cent preferred a single source. At the strong evidence threshold, 60.4 per cent of sources are correctly identified as blends, with 18.2 per cent misidentified as single sources. The threshold to exclude misidentified sources completely in the \( m_{\max} \) case is \( |\ln(P_{2,1})| < 13.9 \), slightly higher than the \( \sigma_{\max} \) case. At this level, 48.0 per cent of sources are still correctly identified as blends.

These results show that photometric redshift estimates can be obtained for blended sources, and that the method can identify many blended sources from just their broad-band photometry. By adjusting the threshold of the probability \( P_{2,1} \), blended sources can be selected in a way that trades off completeness and purity.

Several techniques for improving the scatter of photometric redshift estimates have been proposed, such as rest-frame template error functions (Brammer, van Dokkum & Coppi 2008), iterative methods to modify templates to be more representative (Feldmann et al. 2006), using clustering-based redshift estimation to calibrate systematic biases using galaxies (Gatti et al. 2018) and intensity mapping observations (Alonso et al. 2017), and constructing priors in terms of physical galaxy properties (Tanaka 2015). While an investigation of these methods is beyond the scope of this paper, they could also be applied while using this method. This could help to reduce the scatter of the blended photometric redshift estimates to a level necessary for future surveys, while retaining the full information of the posterior for accurate error propagation.

### 7 CONCLUSIONS

Blended sources will become far more common in future galaxy surveys than are found currently due to increases in the depth of photometry and as a result, the number density of galaxies. We present a BPZ method that generalizes the existing BPZ (Benítez 2000) method to the case of blended observations. We derive a posterior for the redshift and magnitude of each component from the broad-band photometry alone. The joint posterior distribution of the redshifts of all components in a blend provides a complete accounting of the correlations in the final result, information that can be lost when separating components and estimating redshifts for each component separately. This uncertainty information is essential for obtaining accurate uncertainties on cosmological parameters that rely on the photometric redshift estimates. A Python implementation of the method, blendz, is available to download.

By inferring the redshifts of components directly from their blended photometry, the method presented here is directly appli-
Figure 13. Scatter plot comparing the maximum a posteriori point estimates from the photometric redshift estimation with the spectroscopic redshifts for sources from the GAMA blended sources catalogue. The left-hand panels distinguish the components, with $z_\alpha$ plotted with closed blue markers and $z_\beta$ plotted with open green markers. The centre panels show the blend identification, with sources identified as blends plotted with closed purple markers, and those misidentified as single sources plotted with open red markers. The right-hand panels show a 2D histogram of the combined sample. The panels in the top row show the results for the sigma-$m_{\text{max}}$ case, while those in the bottom row show the fixed-$m_{\text{max}}$ case. The dashed lines in each panel show an error of 0.15(1 + $z$).

Figure 14. Plots showing the differences in the model comparison results between the two methods tested of setting the faint-end magnitude cut $m_{\text{max}}$, labelled the sigma-$m_{\text{max}}$ and fixed-$m_{\text{max}}$ cases. The left-hand panel shows the distribution of the relative blend-to-single probability for the mock catalogue, with the inset showing the same distribution, zoomed around lower relative probabilities and binned more finely. The solid line shows the sigma-$m_{\text{max}}$ case, and the dashed line shows the fixed-$m_{\text{max}}$ case. The two right-hand panels show the percentage of sources assigned as either blended, single sources, or not assigned to either as the threshold for deciding between each label is changed.

The method presented here could also be combined with existing deblending methods that utilize the spatial information of images directly. These methods are complementary; image-based deblending methods are effective provided that components are sufficiently well separated. If this is not the case, there is too little spatial information to be able to separate components, and colour information is necessary. Combining these methods could allow future surveys to identify a greater proportion of blended sources, reduc-
ing their effects on cosmological constraints. Debiasing methods that also incorporate colour information would need to be combined with this method more carefully however, as the colour information would be used twice and thus the blending probabilities would not be independent. This method could instead be extended to incorporating imaging data by constructing a forward model of the galaxy in each band and constraining both morphology and redshift simultaneously.

ACKNOWLEDGEMENTS

We thank Andrew Jaffe, Daniel Mortlock, and Josh Greenslade for helpful discussions, and the referee Michael Troxel for helpful comments. DMJ acknowledges funding from STFC through training grant ST/N504336/1. GAMA is a joint European–Australasian project based around a spectroscopic campaign using the Anglo-Australian Telescope. The GAMA input catalogue is based on data taken from the SDSS and the UKIRT Infrared Deep Sky Survey. Complementary imaging of the GAMA regions is being obtained by a number of independent survey programmes including GALEX MIS, VST KiDS, VISTA VIKING, WISE, Herschel-ATLAS, GMRT, and ASKAP providing UV to radio coverage. The GAMA website is http://www.gama-survey.org/. Based on observations made with ESO Telescopes at the La Silla Paranal Observatory under programme ID 179.A-2004.

REFERENCES

Almosallam I. A., Jarvis M. J., Roberts S. J., 2016, MNRAS, 462, 726
Alonso D., Ferreirra P. G., Jarvis M. J., Moodley K., 2017, Phys. Rev. D, 96, 043515
Baldry I. K. et al., 2014, MNRAS, 441, 2440
Baldry I. K. et al., 2017, MNRAS, 474, 3875
Baldry I. K. et al., 2017, MNRAS, 474, 3875
Beck R., Lin C.-A., Ishida E. E. O., Gieseke F., de Souza R. S., Costa-Duarte Carliles S., Budavári T., Heinis S., Priebe C., Szalay A. S., 2010, ApJ, 712, 511
Byrd R. H., Lu P., Nocedal J., 1995, SIAM J. Sci. Comput., 16, 1190
Cawthon R. et al., 2017, MNRAS, 481, 2427
Cawthon R. et al., 2017, MNRAS, 481, 2427
Chawton R. et al., 2017, MNRAS, 481, 2427
Chang C. et al., 2013, MNRAS, 434, 2121
Coe D., Benítez N., Sánchez S. F., Jee M., Bouwens R., Ford H., 2006, AJ, 132, 926
Coleman G. D., Wu C.-C., Weedman D. W., 1980, ApJS, 43, 393
Collister A. A., Lahav O., 2004, PASP, 116, 345
Dark Energy Survey Collaboration, 2016, MNRAS, 460, 1270
Dark Energy Survey Collaboration, 2018, Phys. Rev. D, 98, 043526
Dawson W., Schneider M., 2014, Technical Report, Complementarity of LSST and WFIRST: Regarding Object Blending. Lawrence Livermore National Laboratory (LLNL), Livermore, CA
Dawson W. A., Schneider M. D., Tyson J. A., Jee M. J., 2016, ApJ, 816, 11 de Jong J. T. A. et al., 2013, The Messenger, 154, 44
Duncan K. J., Jarvis M. J., Brown M. J. I., Röttgering H. J. A., 2018, MNRAS, 477, 5177
Edge A., Sutherland W., Kuijken K., Driver S., McMahon R., Eales S., Emerson J. P., 2013, The Messenger, 154, 32
Feldmann R. et al., 2006, MNRAS, 372, 565
Feroz F., Hobson M. P., Bridges M., 2009, MNRAS, 398, 1601
Firth A. E., Lahav O., Somerville R. S., 2003, MNRAS, 339, 1195
Foreman-Mackey D., 2016, J. Open Source Softw., 24
Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306
Gatti M. et al., 2018, MNRAS, 477, 1664
Graham M. L., Connolly A. J., Ivezic Z., Schmidt S. J., Jones R. L., Juríč M., Daniel S. F., Yoachim P., 2018, AJ, 155, 1
Heymans C. et al., 2012, MNRAS, 427, 146
Hildebrandt H. et al., 2010, A&A, 523, A31
Hildebrandt H. et al., 2017, MNRAS, 465, 1454
Hill D. T. et al., 2011, MNRAS, 412, 765
Hoekstra H., Viola M., Herbonnet R., 2017, MNRAS, 468, 3295
Hogg D. W., 1999, preprint (arXiv:9905116
Holwerda B. W. et al., 2015, MNRAS, 449, 4277
Hurley P. D. et al., 2017, MNRAS, 464, 885
Huterer D., Takada M., Bernstein G., Jain B., 2006, ApJ, 647, L21
Ilbert O. et al., 2006, A&A, 511, 384
Ivezic Z. et al., 2008, preprint (arXiv:0805.2366)
Joseph R., Courbin F., Stacek J.-L., 2016, A&A, 589, A2
Joudaki S. et al., 2018, MNRAS, 474, 4894
Kass R. E., Raftrey A. E., 1995, J. Am. Stat. Assoc., 90, 773
Kalbinger M. et al., 2013, MNRAS, 430, 2200
Kinney A. L., Calzetti D., Bohlin R. C., McQuade K., Storchi-Bergmann T., Schmitt H. R., 1996, ApJ, 467, 38
Laureijs R. et al., 2011, preprint (arXiv:1110.3193
Leistedt B., Hogg D. W., Wechsler R. H., DeRose J., 2018, preprint (arXiv:1807.01391)
Leistedt B., Mortlock D. J., Peiris H. V., 2016, MNRAS, 460, 4258
LSST Science Collaboration, 2009, preprint (arXiv:0912.0201)
Lupton R. H., 2005, Technical Report, SDSS Image Processing I: The Deblander. Available at: wwwastro.princeton.edu/ rhl/photomisc/deb lender.pdf
Melchior P., MoolkampF., Jeree M., Armstrong R., Sun A.-L., Bosch J., Lupton R., 2018, Astron. Comput., 24, 129
Menard B., Scranon R., Schmidt M., Morrison C., Jeong D., Budavari T., Rahman M., 2013, preprint (arXiv:1303.4722
Newman J. A., 2008, ApJ, 684, 88
Peebles P. J. E., 2001, in Martínez V. J., Trimble V., Pons-Bordería M. J., eds, ASP Conf. Ser. Vol. 252, Historical Development of Modern Cosmology. Astron. Soc. Pac., San Francisco, p. 201
Peng C. Y., Ho L. C., Impey C. D., Rix H.-W., 2002, AJ, 132, 926
Petri A., May M., Haiman Z., 2016, Phys. Rev. D, 94, 063534
Planck Collaboration XIII, 2016, A&A, 594, A13
Racca G. D. et al., 2016, in Howard A. M., GiovannG. F., Makenzie L., Natalie B., Nicholas S., Edward C. T., eds, Proc. SPIE Conf. Ser. Vol. 9404, Space Telescopes and Instrumentation 2016: Optical, Infrared, and Millimeter Wave. SPIE, Bellingham, p. 99040O
Rhodes J. et al., 2017, ApJS, 233, 21
Rivera J. D., Moraes B., Abdalla F. B., Abdalla M. C. B., 2018, MNRAS, 477, 4330
Robothom A. S. G., Taranu D. S., Tobar R., Moffett A., Driver S. P., 2017, MNRAS, 466, 1513
Sadeh I., Abdalla F. B., Lahav O., 2016, PASP, 128, 104502
Samuroff S., Troxel M. A., Bridle S. L., Zuntz J., MacCrann N., Krause E., Eifler T., Kirk D., 2017, MNRAS, 465, L20
Schmidt S. J., Menard B., Scranon R., Morrison C., McBride K. C., 2013, MNRAS, 431, 3307
Schmidt S. J., Thornon P., 2013, MNRAS, 431, 2766
Semboloni E., Tereno I., van Waerbeke L., Heymans C., 2009, MNRAS, 397, 608
Skilling J., 2006, Bayesian Anal., 1, 833
Sołtan A. M., 2016, in van de Weygaert R., Shandarin S., Saar E., Einasto J., eds, Proc. IAU Symp. 308, The Zeldovich Universe: Genesis and Growth of the Cosmic Web. Kluwer, Dordrecht, p. 291
Speagle J. S., Eisenstein D. J., 2017, MNRAS, 469, 1186
Stoughton C. et al., 2002, AJ, 123, 485
Sun L., Zhan H., Tao C., 2015, preprint (arXiv:1512.00600)
Tanaka M., 2015, ApJ, 801, 20
Trotta R., 2008, Contemp. Phys., 49, 71
Troxel M. A. et al., 2017, Phys. Rev. D, 98, 043528
Way M. J., Srivastava A. N., 2006, ApJ, 647, 102

Yasuda N. et al., 2001, AJ, 122, 1104

This paper has been typeset from a \TeX/\LaTeX file prepared by the author.