

Rigorous luminosity function determination in the presence of a background: theory and application to two intermediate redshift clusters[★]

S. Andreon,^{1†} G. Punzi² and A. Grado³

¹INAF–Osservatorio Astronomico di Brera, Milano, Italy

²INFN and Scuola Normale Superiore, Pisa, Italy

³INAF–Osservatorio Astronomico di Capodimonte, Napoli, Italy

Accepted 2005 March 31. Received 2005 March 9; in original form 2005 January 14

ABSTRACT

In this paper we present a rigorous derivation of the luminosity function (LF) in the presence of a background. Our approach is free from the logical contradictions of assigning negative values to positively defined quantities and avoids the use of incorrect estimates for the 68 per cent confidence interval (error bar). It accounts for Poisson fluctuations ignored in previous approaches and does not require binning of the data. The method is extensible to more complex situations, does not require the existence of an environment-independent LF, and clarifies issues common to field LF derivations. We apply the method to two clusters of galaxies at intermediate redshift ($z \sim 0.3$) with among the deepest and widest K_s observations ever taken. Finally, we point out the shortcomings of flip-flopping magnitudes.

Key words: methods: statistical – galaxies: clusters: general – galaxies: evolution.

1 INTRODUCTION

The luminosity function (LF), i.e. the number of galaxies per unit luminosity and volume, is one of the fundamental quantities of observational cosmology: it is interesting in its own right, and it is a necessary ingredient (weight) in most cosmological measures dealing with galaxies. The history of the LF determination dates back to Zwicky (1957) at least. This debate with Hubble (1936); Zwicky (1951) around the shape of the LF is one of the pillars of the history of LF determination.

With the advent of large surveys, such as the 2dF (Folkes et al. 1999), the Sloan Digital Sky Survey (SDSS; York et al. 2000) and the Virtual Observatory, samples grow by orders of magnitude, and it is nowadays common to deal with more than 1000 galaxies when computing the LF. However, at the extremes of absolute magnitude ranges or in special environments or for certain galaxy types, the number of galaxies is often low. Methods used for the LF computation have also improved along the years (see citations in Section 3).

In Andreon (2004) we have shown how much the neglected observer prior influences the found result (error and confidence interval; an example along the same lines is presented in Blanton et al. 2003). This paper has a twofold aim: to improve the method in the LF determination and to apply it to the best data (useful for the LF determination) ever taken in the K band.

The paper is organized as follows. In Section 2, we present the data and the data reduction. In Section 3, we present a new statistical method. In Section 4 we derive the LF. A discussion and a summary are presented in Section 5.

We assume $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $\Omega_\Lambda = 0.7$ and $\Omega_M = 0.3$.

2 DATA AND DATA REDUCTION

2.1 AC 114 and AC 118

AC 114 and AC 118 are among the most observed clusters at intermediate redshift. Discovered by Couch & Newell (1984) and later by Abell, Corwin & Olowin (1989), they have been the focus of extensive studies: spectroscopic observations (e.g. Couch & Sharples 1987), near-infrared imaging (e.g. Barger et al. 1996), *Hubble Space Telescope* (*HST*) observations (e.g. Couch et al. 1998), mass determination through gravitational lensing experiments (Smail et al. 1997), galaxy evolution studies (Barger et al. 1996; Couch et al. 1998; Jones, Smail & Couch 2000; Couch et al. 2001), etc. AC 114 is a regular massive cluster, whereas AC 118 is a massive merging system. A detailed description of these two clusters may be found in the above-mentioned papers.

AC 114 observations were carried out at the 3.5-m New Technology Telescope (NTT) with SOFI (Moorwood, Cuby & Lidman 1998) for four nights during the autumn of 1998. SOFI is equipped with a 1024×1024 pixel Rockwell ‘Hawaii’ array. In its large field mode, the pixel size is 0.292 arcsec and the field of view $5 \times 5 \text{ arcmin}^2$. The field was observed in the near-infrared K_s passband ($\lambda_c = 2.2\mu$; $\Delta\lambda \sim 0.3\mu$) during four photometric nights with good seeing (FWHM < 0.8 arcsec). The total useful exposure time is

[★]Based on observations collected at the European Southern Observatory, Chile, and, in part, on observations with the NASA/ESA *Hubble Space Telescope*.

†E-mail: andreon@brera.mi.astro.it

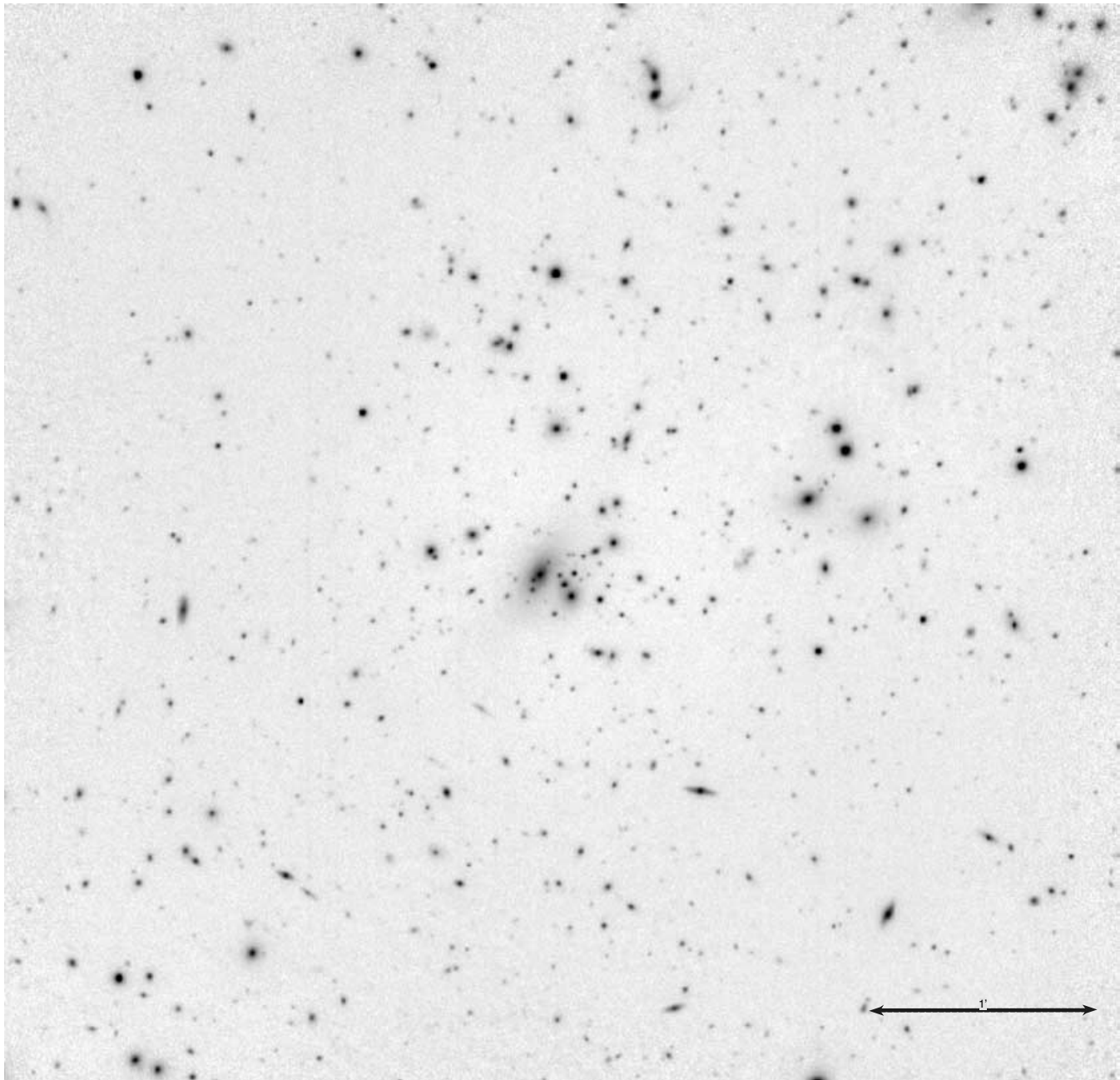


Figure 1. The K_s -band image of AC 114. The field of view is $\sim 5 \times 5$ arcmin². North is up and east is to the left.

18 840 s, resulting from the co-addition of many short jittered exposures. Photometric calibration has been obtained by observing a few standard stars, interspersed with AC 114 observations, taken from the list of Infrared NICMOS Standard Stars published in Persson et al. (1998). Fig. 1 shows the final K_s image of AC 114. This image has a seeing of 0.8 arcsec.

AC 118 observations were carried out with the same instrument, the night after AC 114 observations, and are fully described in Andreon (2001, hereafter Paper I).

All images have been reduced as in Paper I. Shortly, they are flat-fielded by flaton–flatoff. In order to test the accuracy of the flat-fielding, a standard star has been observed in eight chip locations. The root-mean-square (rms) variation of its magnitude is 0.008 mag. Because the rms deviation is small, our images do not require a

supplementary illumination correction. The background has been removed by using ECLIPSE (Devillard 1997), taking advantage of the telescope nodding during the observations. Images have been combined using the task IMCOMBINE under IRAF using integer pixel shifting.

2.2 Control field: CDF-S and HDF-S

As control field we use the Hubble Deep Field South 1 and 2 (HDF-S) images (Da Costa et al. 1998), already used for AC 118, and therein described, supplemented by Chandra Deep Field South (CDF-S) images (Rengelink et al. 1998; Vandame et al. 2001). We only remind that all these images have been taken with the same instrument, filter and telescope as the AC 114 and AC 118 images;

Table 1. The data.

	AC 114	AC 118	CDF-S	HDF-S
Exposure time (min)	314	265	80–180	180–300
Seeing (FWHM, arcsec)	0.73	0.75	~0.8	0.9–1.0
Completeness magnitude ($\phi = 4.4$ arcsec)	20.3	20.5	19.5–20.0	20.25
Area (arcmin ²)	23.7	23.7	242.0–45.0	47.0

these cluster images are interspersed to control field images, hence ensuring an almost perfect homogeneity of the data. The basic data reduction of control and science fields is based on the same software (ECLIPSE). Two major differences occur: science data have not been resampled, in order not to correlate the noise of adjacent pixels; and science data are combined with more attention to flux (see Paper I for details), allowing us to claim a better photometric calibration precision for cluster images (better than 0.01 mag) than other authors claim for the control field (around 0.05 mag).

The 16 SOFI pointings of the CDF-S guarantee a large area coverage (242 arcmin²) down to $K_s = 19.5$ and hence a good determination of the galaxy counts in the control field. Three of them, covering 45 arcmin², are exposed longer and reach $K_s = 20$, hence supplementing the 47 arcmin² of the HDF-S down to $K_s = 20.25$ used in Paper I. At $K_s < 18$ HDF-S shows a marginally high overdensity with respect to CDF-S. Therefore, we arbitrary remove the bright part of the HDF-S galaxy counts (which, in any case, carry a negligible weight, given the small observed area of the HDF-S).

Table 1 shows a summary of all observations. The area coverage of the CDF-S alone is larger than the latest published galaxy counts (Cristóbal-Hornillos et al. 2003), down to their completeness limit ($K \sim 19.5$).

2.3 Photometry and flip-flopping magnitudes

Objects have been detected by using SExtractor version 2.2 (Bertin & Arnouts 1996). For AC 114 we made use of the rms map for a clean detection, as we did for AC 118. Because of the varying exposure time across the field of each image, due to the dithering, we consider here only the central square areas listed in Table 1.

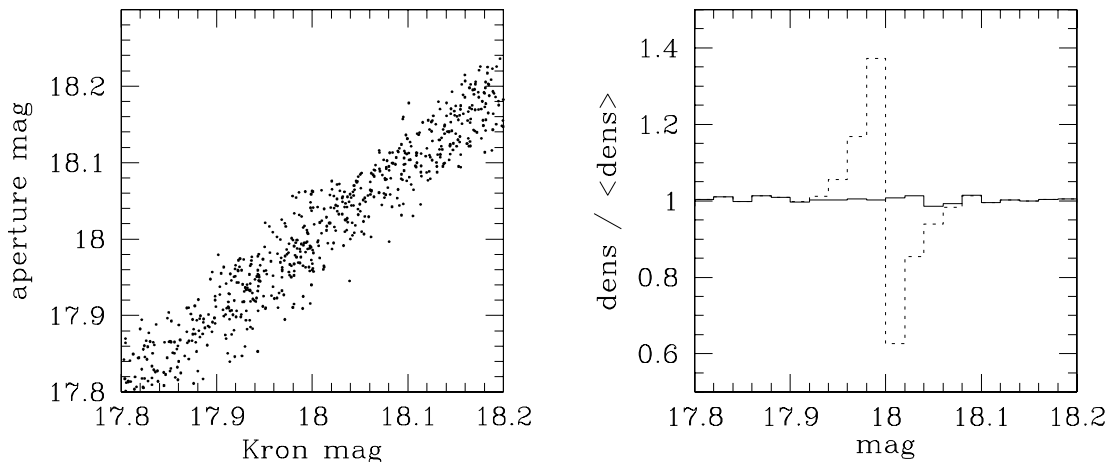


Figure 2. The magnitude bias in a Monte Carlo simulation with Gaussian noise ($\sigma = 0.03$ mag; shown in the left panel), adopting an orthogonal (solid histogram) or vertical cut (dotted histogram).

Galaxy are extended objects, and hence their luminosity depends on the way their borders are defined. We improve our magnitude definition with respect to Paper I; here, we adopt Kron magnitudes (see Kron 1980 for the exact definition, and Bertin & Arnouts 1996 for a software implementation) for bright [$(K_{\text{Kron}} + K_{\text{aper}})/2 < 18$ mag] objects and aperture (in a 4.4-arcsec aperture) otherwise. In Paper I, the cut was performed along one of the axes (K_{Kron}), and not orthogonally to the $K_{\text{Kron}}-K_{\text{aper}}$ relationship, spuriously producing a density variation in the unbinned distribution of galaxy counts (not actually used in that paper, but used here) due to the spread around the $K_{\text{Kron}}-K_{\text{aper}}$ relationship. A Monte Carlo simulation of what occurs is shown in Fig. 2. On the left, we show a linear relationship between the Kron and aperture magnitude, with a Gaussian small scatter ($\sigma = 0.03$ mag) and no bias. On the right, we histogram the galaxy counts with a cut orthogonal to the Kron–aperture magnitude relationship (solid histogram), and at a fixed Kron magnitude (dotted histogram). The latter histogram presents a huge (40 per cent) variation near the ‘bridge’ magnitude, absent when an orthogonal cut is performed.

Why not use Kron magnitudes at all magnitudes then, as in many of the works in the literature? The reason is found in the SExtractor manual. The Kron magnitude is measured in two different ways depending on the measured object radius: it is a true Kron magnitude for objects larger than a radius threshold and an aperture magnitude for fainter objects. Such a measure, and potentially all flip-flopping magnitudes (such as ‘auto’ or ‘best’ magnitude) distorts the luminosity distribution (i.e. the galaxy counts) near the ‘bridge’ point.

The magnitude completeness is defined as the magnitude where objects start to be lost because their measured central brightness is lower than the detection threshold (see Garilli, Maccagni & Andreon 1999 and Paper I for details). For AC 114, the (5σ) limiting magnitude is $K_s \sim 20.3$ mag in a 4.4-arcsec aperture. For simplicity and for excess of caution, we consider here only $K_s < 20.0$ mag objects.

Objects are classified according to their compactness, by using the SExtractor stellar classifier. Almost the whole area of AC 114 studied here has been observed by the *HST* mosaic (Couch et al. 1998). Galaxies are resolved (i.e. not point-like) objects at the *HST* resolution. The comparison of our ground-based classification and *HST* images of the same objects confirms the goodness of our ground-based star/galaxy classification because a few galaxies, out of hundreds, are misclassified.

2.4 Comparison to literature photometry

AC 114 has been observed in the K' band by Barger et al. (1996) and by Stanford et al. (2002). Stanford et al. (2002) measure aperture magnitudes (in a 5-arcsec aperture). Our Kron magnitudes agree well with them, with no photometric offset and a typical scatter of 0.2 mag.

Barger et al. (1996) measure pseudo-total magnitudes on images taken with an instrument having a large pixel size (0.79 arcsec). Our magnitudes are brighter than theirs by 0.18 mag, the offset being potentially due to their quite large pixel size and worse seeing (between 1.1 and 1.7 arcsec).

All objects listed in Barger et al. (1996) or Stanford et al. (2002) are present in our catalogue, as expected, because our images are much deeper. Instead, several objects, brighter than the completeness magnitude of Barger et al. (1996) or Stanford et al. (2002), are missing in their catalogues.

3 LUMINOSITY FUNCTION: STATISTICAL METHODS

3.1 Background

We are here faced with the classical problem of determining two extended (integral >1) density probability functions, one carrying the signal (the cluster LF) and the other being due to a background [background galaxy counts (BKGs)] from the observations of many individual events (the galaxies luminosities), without knowledge of which event is the signal and which is background.

Traditionally, the cluster LF is computed as the difference between galaxy counts in the cluster and control field directions (Zwicky 1957; Oemler 1974), i.e. after binning the events (galaxy magnitudes) in magnitude bins. In performing such a computation:

- (i) galaxy counts are binned in magnitude bins (of arbitrary width);
- (ii) galaxy counts in the control field direction are subtracted from counts in the cluster direction in order to obtain the cluster contribution alone;
- (iii) in order to estimate the error on the cluster LF, approximate Poisson errors (i.e. \sqrt{n}) and, in some cases, over-Poisson errors due to large-scale structure are added in quadrature, under the (approximate) hypothesis that they are Gaussian distributed.

Binning has several advantages:

- (i) it allows us to ‘see’ how data are distributed (or better, to ‘see’ the data distribution convolved by the binning function);
- (ii) it allows a quick analysis of the data;
- (iii) it allows us to calculate the goodness of fit in a simple way, using χ^2 ; and
- (iv) it provides a correct result at large signal-to-noise ratio.

The bin width is arbitrary, but recently Takeuchi (2000) has suggested a legitimate rule when bins are all of the same width: the Akaike information criterion can be used for optimally choosing the number of bins. However, when galaxy counts change by three order of magnitudes, as is usual in computing LFs, such an approach is optimal, on average, but far from optimal in the (faint) bins populated by thousand of galaxies or in the almost empty (bright) bins.

Indeed, it would be preferable to avoid any binning of the data for the following reasons.

- (i) No matter which amplitude bin is chosen, it tends to be too wide in crowded regions and too narrow in low-populated regions.

Adaptive binning (i.e. of variable width) depending on the local population is a possible solution, which, however, shares the problems listed below, and requires a more elaborate fitting algorithm (an appropriate convolution of the fitting function).

(ii) The negative LF (as well as background galaxy counts) makes no sense (because both functions are positively defined), and hence any determination allowing the LF to be negative has a dubious meaning. Binning, coupled with background subtraction, may produce such occurrences: it may happen that, because of statistical fluctuations, the counts in the control field direction are larger than those in the cluster direction, leaving a negative number of galaxies, for a positively defined quantity. Are a negative number of galaxies ever seen? Although negative values are often consistent with zero, they cannot be simply ignored or set to zero, otherwise a significant bias would occur. For example, the integral over the LF, the cluster richness and the luminosity density are systematically overestimated.

(iii) Binning frequently produces error bars on LF crossing the $LF = 0$ line, considering the possibility of a negative number of galaxies (which the authors are still not ready to accept).

(iv) Binning in high dimensions (here we have, for example, six to nine dimensions; see equation 3) makes the data sparse, no matter how large the sample is, especially when the galaxy density changes by three orders of magnitude from bright to faint magnitudes. As mentioned, low-populated regions are a problem for several reasons.

(v) Binning implicitly assumes that no change is occurring inside the bin, and it occurs only at the bin boundaries (the idea of continuity is lost in binning). For example, LFs are often measured in redshift bins (assuming they do not evolve inside the redshift bin), and then compared among them to look for a redshift evolution, which, according to the logic of the people making such a comparison, occurs at the bin boundary only, and with ‘quantum jumps’ (see Andreon 2004 for details).

(vi) Binning makes a rigorous statistical analysis a nightmare: errors are not Gaussian distributed (when the number of objects inside a bin is small, and in a multidimensional space there are always bins low populated), linear least-squares fits (such as χ^2) badly fail and give biased results when the number of objects inside the bin is small (Wheaton et al. 1995). The latter work suggests fitting ‘one count at a time’, i.e. not to bin at all. Furthermore, having observed n_0 , the 68 per cent confidence interval is not given by $[n_0 - \sqrt{n_0}, n_0 + \sqrt{n_0}]$ when n_0 is small (see, for example, Gehrels 1986, or statistical textbooks).

We therefore opt for an unbinned fit of ‘one galaxy at a time’, following the path put forward by Sandage et al. (1979, hereafter STY), where it was assumed that no background, no evolution and no environment dependence were present. In Lin et al. (1999), a monolithic (i.e. independent of luminosity) extension has been introduced under the assumption of no background at all (i.e. the redshift of each galaxy is known). Andreon (2004) removed the monolithic evolution, allowing galaxies of different luminosity to evolve by different amounts, still in the absence of background.

In the present paper, we allow the presence of background galaxies, unrelated to the cluster, i.e. we present how the LF can be computed when the individual membership of galaxies is unknown. However, we assume (as STY) a LF universality (i.e. an LF independent of environment). The reason is mainly technical, not theoretical: the formalism introduced below is easily extensible to such a case (for example, following the parametrization with environment or redshift outlined in Andreon 2004), but coding it is quite complex.

In order to account for observations of different quality (depth, area, etc.) a determination using several data sets (each having bounds in magnitude or area) is allowed, as in Efstathiou et al. (1988, hereafter EEP). For example, the HDF-S observations used are actually left-censored (i.e. we have no data to the left of) at $K_s = 19.0$ and right-censored (i.e. we have no data at the right of) at $K_s = 20.25$.

The method naturally converges to results obtained when data are binned, when binning can be done, i.e. when the number of objects for the bin is large and the Gaussian approximation occurs.

3.2 Adopted approach

3.2.1 Using extended likelihood and properly accounting for background

Our approach is based on a single likelihood function, which accounts simultaneously for all available data, cluster and control fields. The use of extended likelihood keeps the normalization usually lost in other methods. We do not require that the observed background in the cluster line of sight is ‘average’ (or typical), but only that it is drawn from the same parent distribution from which the background in the control field is drawn.

Given j data sets (say, cluster 1, cluster 2, ... field 1, field 2, ...) each composed of N_j galaxies, we maximize the extended likelihood \mathcal{L} given by

$$\ln \mathcal{L} = \sum_{\text{data sets } j} \left(\sum_{\text{galaxies } i} \ln p_i - s_j \right). \quad (1)$$

Here, p_i is the (extended, because integral is not 1) probability of the i th galaxy of the j th data set to have m_i , i.e. $p_i = p(m_i)$; s is the integral of function ϕ over the range $[\text{mag}_{\text{left},j}, \text{mag}_{\text{lim},j}]$, i.e. the expected number of galaxies, given the model. In

$$s = \int_{\text{mag}_{\text{left},j}}^{\text{mag}_{\text{lim},j}} \phi(m) dm, \quad (2)$$

$\text{mag}_{\text{lim},j}$ is the limiting magnitude of the j th data set, and $\text{mag}_{\text{left},j}$ is the limiting magnitude at the bright end (in the case of left-bounded magnitude values) of the j th data set. For example, if in the sample $K < 10$ galaxies are filtered out (because saturated, or because such galaxies would cause trouble for the instrument by, say, occupying a large fraction of the field of view), it will be $\text{mag}_{\text{left}} = 10$ for that sample.

ϕ is the sum of a power law (accounts for the background contribution) and a Schechter (1976) function:

$$p_i = \phi(m) = \delta_c \Omega_j \phi^* 10^{0.4(\alpha_c+1)(m^*-m)} e^{-10^{0.4(m^*-m)}} + \Omega_j 10^{a+b*(m-20)+c*(m-20)^2}. \quad (3)$$

Here, $\delta_c = 1$ for cluster data sets, $\delta_c = 0$ for the other data sets, a, b, c describe the shape of the galaxy counts in the reference field direction, and Ω_j is the studied solid angle. The number ‘20’ is there for numerical convenience. If galaxy counts have a more complex magnitude distribution, more coefficients (or any other parametrization) can be used to describe the shape distributions. Analogously, if the cluster LF is more complex than a Schechter function, say a sum over the LFs of the individual morphological types (e.g. Andreon 1998), the Schechter function in equation (3) can be replaced with the reader’s favourite function without affecting the overall approach.

The above approach neglects the effect of large-scale structure, and it is justified when the variance due to large-scale structure is

much lower than the Poissonian variance. For $K_s > 12$ mag, and for a solid angle as small as one single SOFI field of view (about 20–25 arcmin²), the variance due to large-scale structure, computed according to Huang et al. (1997), is less than 1 per cent of the Poissonian variance and can be safely neglected.

The cluster LF is given by the Schechter (or favourite) function with parameters that maximize the likelihood. Confidence contours may be computed using the likelihood ratio theorem. The 68 and 95 per cent confidence contours for two interesting parameters are computed from $2\Delta \ln \mathcal{L} = 2.3$ and 6.17, respectively (Avni 1976; Wilks 1938, 1963; Cash 1979; Press et al. 1993). The 68 per cent confidence interval for a single parameter is computed using $2\Delta \ln \mathcal{L} = 1$ (Avni 1976; Wilks 1938, 1963; Cash 1979; Press et al. 1996). We note the approximate nature of these and that some regularity conditions are required (see Protassov et al. 2002 for astronomical related references). The large (>1000) number of galaxies and the absence of borders near the best-fitting parameters guarantee that the hypothesis on which the likelihood ratio theorem is based is satisfied for the data used in the present paper. Regularity conditions are not always satisfied when dealing with the Butcher–Oemler effect (Andreon, in preparation).

For the goodness of fit, we adopt the Persson χ^2 test, accurately described in section 14.3 of Press et al. (1993) for Poissonian distributed quantities. The Persson χ^2 is, in the long run, χ^2 -distributed with the number of degrees of freedom, ν , equal to the number of the bins minus the number of parameters of the fit function. The test is applied on galaxy counts, not on the difference of galaxy counts in the cluster and control field directions. The goodness-of-fit estimation requires us to bin the data.

3.2.2 How to find a global minimum?

The maximum likelihood can be found using simulated annealing methods (e.g. Press et al. 1996), because the desired global maximum is often hidden among many, poorer, local maxima in high-dimensional spaces. For larger dimension problems, it is computationally more efficient to use the Markov Chain Monte Carlo method (e.g. Dunkley et al. 2005).

Best-fitting parameters are determined all together at once; we avoid the procedure used by other authors of fitting the control field counts, and, once we have found the best-fitting parameters for the control field, we switch to fitting the cluster counts by keeping the background parameters fixed. The above procedure does not guarantee finding the global minimum.

Such a global fitting also accounts for a difference in the observed value of background counts in cluster and reference field directions.

3.3 Where we improve with respect to previous approaches

In this section we summarize the differences between our approach and previous approaches.

3.3.1 Sandage, Tammann & Yahil

The approach of STY and other maximum likelihood approaches do not deal with the presence of a background, and hence cannot be used when the individual membership of galaxies is unknown.

It is well known that in the STY approach the normalization is lost (i.e. ϕ^*). This situation is not typical of maximum likelihood methods in general, and, in fact, the normalization is kept in our approach, which also gives rigorous 68 per cent confidence intervals; this is a good reason to adopt it.

3.3.2 *Efstathiou, Ellis & Peterson*

The EEP method does not deal with the presence of a background, and hence cannot be used when the individual membership of galaxies is unknown. Furthermore, EEP need to bin the data. Section 3.1 explains why we dislike binning the data.

3.3.3 *Wrong Poisson errors for small n*

As mentioned, LF determinations derived by binning the data in magnitude bins and by computing the cluster contribution as a straightforward cluster – field difference have error bars that are difficult to compute, because for small n , the 68 per cent confidence interval is not given by $[n - \sqrt{n}, n + \sqrt{n}]$ (e.g. Gehrels 1986), and the 68 per cent confidence interval on the difference is not given by the quadrature sum of the 68 per cent confidence intervals of the two addenda.

Because we do not bin the data, we avoid dealing with these incorrect expressions.

3.3.4 *Binning but forgetting to marginalize the model over the bin*

Several LF methods bin the data in mag. Obviously, the change does not occur at the border bin. One should therefore marginalize (integrate) the model LF over the quantity binned. Such a rule is used in several papers for the ‘mag’ quantity (e.g. Paolillo et al. 2001), but not systematically by all authors. Said simply, some authors sincerely believe that inside the bin there is only one ‘mag’ and they compute errors as if this belief is true. However, when describing the LF these authors do not write that the LF is a sum of delta function, each one centred at the centre of the bin, but a smooth function, in logical contradiction with having assumed a sum of delta functions.

3.3.5 *Forgetting s*

The s term in the likelihood is required, as long as Poisson fluctuations are allowed. If absent, or replaced by the observed number of galaxies, Poisson fluctuations at each m are allowed, but Poisson fluctuations of the total number of objects are not.

In particular, neglecting s in the presence of small signals (i.e. the only occasion when statistics is actually required) is dangerous, in the sense that even meaningless results can be found (for example, a negative number of cluster galaxies), and usually leads to underestimating the uncertainty on the parameters. Overlooking s is quite standard in the astronomical community, in the LF computation, in recent detections of cluster of galaxies jointly using (*ROSAT*) X-ray photons and (SDSS) galaxy catalogues, etc.

Popesso et al. (2004) adopt a maximum likelihood method but they replace s with the number of observed galaxies (see their equation 4 and related comments). Their algorithm fails to find a reasonable best-fitting parameter in several cases (look for M^* = 0.00 values in their table 2); the error on the best-fitting parameter is found in some cases to be less than 0.005 (for example, for RXC J0747.0+4131), a precision never previously achieved not even with a 10 times larger sample. M^* of a $z = 0.78$ cluster (RXC J1140.3+6609) can be computed with good accuracy using about 50-s exposures at a 2.5-m telescope, when its brightest galaxies are marginally detected, if any. Replacing s with the number of observed galaxies may produce failures in finding reasonable values for the best-fitting parameters and may give strongly underestimated uncertainties. The s term, prescribed by the extended likelihood approach, does not allow similar situations to occur.

The latter work disagrees with us in computing the LF from incomplete samples without accounting for incompleteness.

3.3.6 *Dissenting views*

The measure of the LF by using the statistical subtraction of background has been criticized by Toft, Soucail & Hjorth (2003), who suggest an alternative way to compute the LF ‘without having to make uncertain statistical corrections to account for foreground and background contamination’. A similar statement is repeated in Toft et al. (2004), and in Blanton et al. (2005), because ‘background subtraction [is] an uncertain procedure’.

First of all, it is unclear to us why the background subtraction is uncertain. It is known with a degree of accuracy that it depends on the available data, as other experimental quantities.

Secondly, Toft et al. (2003) replace it with a photometric redshift selection plus a correction for galaxies lost in the selection. Such a correction is uncertain, because it requires us to know the distribution of spectral types at the observed redshifts, and the spectral templates at these redshifts. Both are unknown, at the difference of background counts that are known, because computed in a control field.

The solution of Blanton et al. (2005), instead, is to adopt a method (EEP) which assumes that the LF is independent of environment (equation 2.3 of EEP) for a sample in which the dependence of the LF on environment is flagrant (fig. 15 in Blanton et al. 2005).

Therefore, we cannot agree with their criticisms of the background subtraction methods, and with their proposed solution.

The background subtraction method has been criticized by Valotto, Moore & Lambas (2001), claiming that the presence of a background overdensity in the cluster line of sight favours the cluster detectability and biases the slope of the LF. The above occurs often in their simulations, because ‘many of the clusters found in two dimensions have no significant three-dimensional counterparts’, as they claim. In nature, instead, most of (and perhaps all) the clusters whose LFs are computed have three-dimensional counterparts (i.e. when spectroscopy is performed, the cluster is confirmed), which simply means that their simulations are not an accurate reproduction of nature. Therefore, their criticism does not apply to actual data used for the LF measure, but eventually applies to cases where the cluster detection is doubtful. Furthermore, the LF of a large sample of clusters in Paolillo et al. (2001), selected in two dimensions by Abell (1958) and background subtracted in the way criticized by Valotto et al. (2001), is equal to the LF of another large sample of clusters (Garilli et al. 1999), which is X-ray selected and, according to Valotto & Giovanelli (2004), does not suffer bias. Therefore, the effect of the bias (if it exists) is negligible for the data sets actually used. Finally, in the few cases when a cluster LF is determined by performing a spectroscopic survey deep enough to probe the LF slope, the derived LF is equal within the errors to that derived by using a statistical background subtraction (e.g. for the Coma cluster; Mobasher et al. 2003 versus Andreon & Cuillandre 2002).

4 APPLICATION OF THE METHOD, THE COMPOSITE LUMINOSITY FUNCTION AND DISCUSSION

We apply the method to the data presented in Section 2. Table 2 lists the best-fitting parameters and errors for Schechter parameters.

Figs 3 (for AC 114) and 4 (for AC 118) show the galaxy counts in the control field direction (lower points in the lower panel) and in the cluster line of sight (upper points in the lower panel), and

Table 2. Best-fitting parameters, errors and goodness of fit. a , b and c describe the shape of galaxy counts (equation 3), whereas α , M^* and ϕ^* describe the shape of the cluster LF (equation 3). Units: when inserted in equation 3, a , b and c provide galaxy counts in units of deg^{-2} . The latter are also the units of ϕ^* . M^* is given in mag units. In the first three lines, there are more decimals than precision allows, to avoid truncation errors. The last three lines quote values including all (0.5 mag wide) bins or, in parentheses, excluding bins with less than 10 galaxies.

	AC 114	AC 118	AC 114 + AC 118
a	4.37095547	4.37931252	4.37309837
b	0.303063065	0.312422931	0.305745333
c	-0.0223323945	-0.0203797854	-0.0216595493
α	-1.30 ± 0.07	-1.03 ± 0.02	-1.15 ± 0.05
M^*	15.04 ± 0.32	15.72 ± 0.21	15.43 ± 0.14
ϕ^*	$6.3 \pm 1.8 \cdot 10^3$	$1.47 \pm 0.28 \cdot 10^4$	$1.04 \& 1.03 \cdot 10^4$
χ^2	39.0 (21.1)	34.2 (12.3)	49.0 (24.7)
ν	32 (15)	33 (15)	50 (24)
$P(\geq \chi^2)$	0.20 (0.15)	0.40 (0.65)	0.53 (0.40)

a joint fit to cluster and control field counts. For display purposes only, we show points and error bars computed with usual recipes, although we make no use of them in our analysis (parameter or errors determination). The fit is performed on the unbinned distributions, whereas we bin them for display purposes only. According to astronomical standard practice, the error bars in the lower panels have a width given by \sqrt{n}/Ω_j . The points in the upper panels of Figs 3

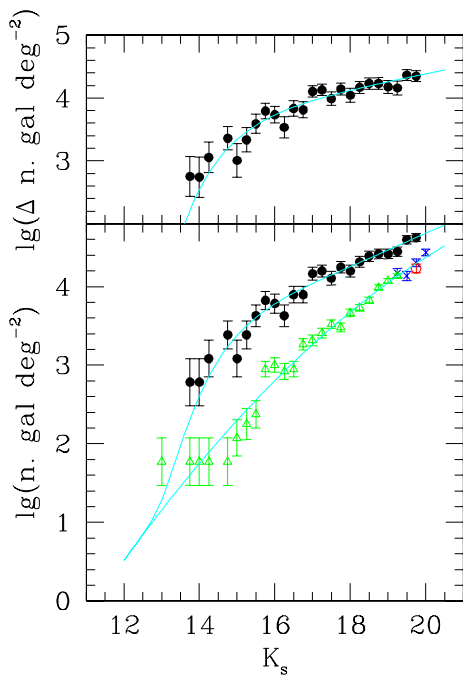


Figure 3. Upper panel: the K_s -band LF of AC 114. Lower panel: galaxy counts in the AC 114 line of sight (solid upper points, in black) and in the control field (lower points, colour-coded and type-coded: green triangles = CDF-S; red empty circles = deep part of CDF-S; blue crosses = HDF-S). Note that because of crowding some points are to be seen on the plot. Incomplete bins (i.e. those that cover a magnitude interval not completely explored by the observations) are not plotted. The cyan lines are best joint fit to the control field and cluster line-of-sight directions on unbinned data. Data are binned in the figure for display purposes only, and are computed as described in the text. Note that bins are 0.25 mag wide, half the usual bin width.

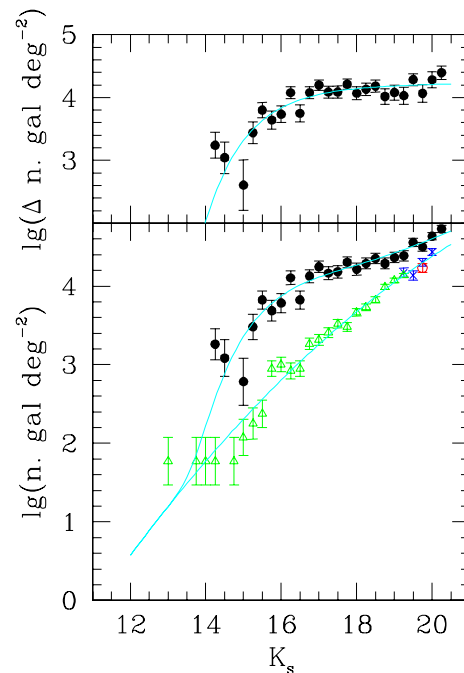


Figure 4. As Fig. 3, but for AC 118.

and 4 mark the algebraic difference between the galaxy counts in the cluster direction and the best-fitting background counts. When the difference is negative (at $K < 14$ mag, plus a few points at fainter mag) the result cannot be plotted, because the scale requires a positive argument for the logarithm. Error bars in the upper panels of Figs 3 and 4 mark the square root of the variances of the minuend and subtrahend.

The Schechter curve is instead the rigorous derivation of the cluster LF, drawn with the best-fitting parameters found on the (cluster+field, field) data sets. It is not a best fit to the cluster – field difference, as detailed in Section 3.2.2, and it is positively defined at every magnitude, at the variance of the above-mentioned ‘data’ points. Nevertheless, the curves nicely describe the (approximately computed) cluster contribution, especially at large signal-to-noise, because here the two approaches converge by definition. At $K_s = 14.75$ mag for AC 118, the model predicts a number of galaxies similar to the data points of adjacent bins, but the above-mentioned difference takes an unphysical value, the unpleasant situation discussed in Section 3.1.

The fits are good, in the sense that the probability of finding a worse χ^2 is large (Table 2).

These LF determinations are among the deepest at the studied (large) area ever measured (see fig. 10 in Paper I), to the best of our knowledge. We hope that our LF method makes them also the most rigorously determined.

A question naturally arises. Are our improvements formally correct but of null importance? After all, the best-fitting model passes through the cluster contribution, even if approximately computed. So, why should we bother with all these apparently useless complications?

Our method does not produce puzzling results, and it is the appropriate choice when puzzling results are found, for example when the following are observed:

- (i) negative star formation rates, which, according to the authors ‘lack physical sense’ (Rojas et al. 2004);

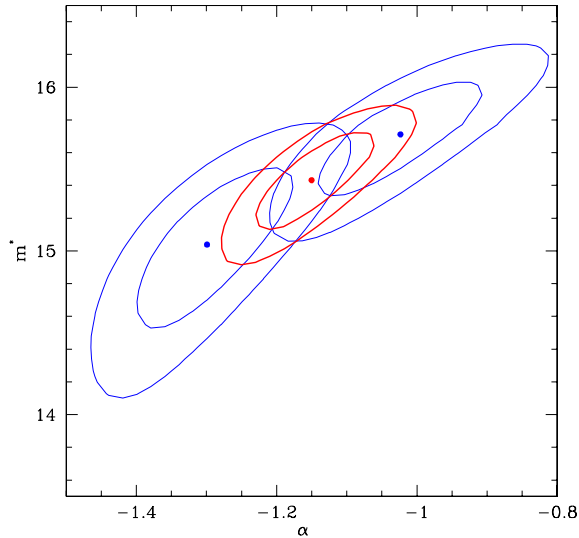


Figure 5. The 68 and 95 per cent confidence contours for α and m^* . The contours at the left, right and centre concern AC 114, AC 118 and the combined AC 114 + AC 118 sample, respectively.

- (ii) negative flux densities (for some SCUBA sources; Smith et al. 2001, their table 1);
- (iii) clusters with negative blue fraction (Butcher & Oemler 1984, their fig. 3);
- (iv) clusters with negative masses (at least in some magnitude/radial bin; see, for example, Hansen et al. 2005, their fig. 5, top-left panel).

Most of these (and other) puzzling results originate from not fully accounting for the impact of a background and of its fluctuations in computing the quantity of interest. Either the analysis is rigorous, and we are sure that the result makes sense, or the correctness of the results cannot be guaranteed.

Fig. 5 shows 68 and 95 per cent confidence contours on m^* and α . With respect to confidence contours of AC 118 computed in Paper I, here confidence contours shrank by a factor of 2 because of the better determination of the background counts, and moved by one (old) sigma, because the new background counts no longer show a minor excess, with respect to a power law, at intermediate magnitudes.

Inspection of Fig. 5 shows that two LFs are approximately compatible at the 95 per cent confidence level. AC 114 is, if at all, steeper and brighter than AC 118, as comparison of Figs 2 and 3 also shows. Such a difference is expected, given the dependence of the best-fitting LF parameters on the environment, put forward in Paper I and in Andreon (2002), and the observed difference in the density distribution of the two clusters (compare Fig. 1 with fig. 1 of Paper I).

Although the inspection of the relative location of (α, M^*) values and contours is the standard astronomical way of comparing (α, M^*) values for different samples, a rigorous comparison of the two LFs, however, should follow another path. First of all, Fig. 5 shows that there are (α, M^*) values within both 95 per cent confidence contours, but it does not show whether these values are achieved for the same values of the parameters not plotted in the figure. For our LFs, the nuisance parameters are the background (a, b, c) parameters. For field LFs, the nuisance parameter ϕ^* is ‘hidden’ and the 95 per cent confidence contours of the two compared LFs may overlap, but for incompatible ϕ^* values. Secondly, a simple comparison of Fig. 5 may incorrectly leave the impression

of compatible LFs, when instead the two LFs are actually different. Consider, for example, the case of two LFs, very different but computed for a background known within a large uncertainty (leading to large confidence contours). The derived contours overlap each other, while a correct comparison of the data (see below) will show the two LFs to be different. Finally, and even in the absence of a background (or any nuisance parameter), the figure actually shows that there is a region of observed values of the α, M^* plane (the region where confidence contours cross) that can be drawn from two different true values of α, M^* at a given confidence, and not that a single pair of α, M^* may generate two observed α, M^* at that confidence; confidence contours give the probability of the data given the hypothesis and not vice versa. By the way, the contours for the two clusters are computed for different hypotheses and both cannot be true at the same time (the two pairs of best-fitting parameters are numerically different).

In order to establish if the two LFs differ, we can ask whether a fit of both clusters with a single set of α, m^* values is much worse than a fit with individual α, m^* values for each cluster. The likelihood ratio test (LRT) allows such a comparison, and without any need to bin the data. Our models to be compared are hierarchically nested¹ and regularity conditions needed to use the LRT hold in our case (but see Protassov et al. 2002 for a case where regularity conditions do not). The likelihood ratio is computed by taking the ratio of the maximum value of the likelihood function under the constraint of the null hypothesis (= one set of α, m^* values) to the maximum with that constraint relaxed (= two sets of α, m^* values). If the null hypothesis is true, then $2\Delta \ln \mathcal{L}$ (= twice the above ratio expressed in logarithm units) will be asymptotically χ^2 -distributed with the number of degrees of freedom equal to the difference in dimensionality of the two considered models.

Therefore, we modify equation (3) adding one more Schechter function, and we fit at once all data, once keeping one single set of m^*, α values for both clusters, and once leaving free m^*, α for each cluster independently. In each fit, both clusters share the same set of parameters describing the background, at variance of the fits discussed earlier in this section, when we did not constrain the parameters of the background to be the same in the AC 114 and AC 118 fits.

We measure $2\Delta \ln \mathcal{L} = 6.4$ for two (more) degrees of freedom. Therefore, under the null hypothesis (the two LFs having the same m^*, α parameters), there is about 5 per cent probability of observing a larger likelihood ratio, confirming the cursory inspection of AC 114 and AC 118 confidence contours previously mentioned. Such a probability is not small enough to reject the null hypothesis that the two LFs have the same m^*, α parameters at high confidence. We can therefore co-add the data of the two clusters and compute the composite LF, which is, actually, the likelihood under the null hypothesis just computed. The above path naturally solves the difficult procedure of averaging the LFs of the two clusters (or, more generally two data sets), rigorously accounting for the error on the relative normalization of the two LFs, often not even mentioned in astronomical papers.

Best-fitting values for the combined data set (= AC 114 + AC 118) are listed in Table 2, and m^*, α confidence contours are shown in Fig. 5. The fit is good, in the sense that the probability of finding a worse χ^2 is large (Table 2).

Although the use of a rigorous (and time-consuming) test leads to the same conclusion of a cursory inspection of AC 114 and

¹ For example, the allowed parameter values of one model must be a subset of those of the other model.

AC 118 confidence contours; the former guarantees a correctness that the latter does not, and therefore should be preferred.

5 SUMMARY AND COMMENTS

We have presented a rigorous method to measure the LF in the presence of background, extending previous methods to deal with a more complicate case, and including neglected terms. The approach does not suffer from logical inconsistencies (or limitations) present in previous approaches and puts on a sure foot claims of providing errors with the correct coverage (i.e. our errors are 68 per cent confidence intervals). We have applied the method to measure the LFs of two clusters of galaxies, using among the deepest and widest observations in the K_s band ever, and producing one of the best determinations of the LF in this band (and we hope one of the more rigorous ones, too). In passing, we show the bias of a flip-flopping definition of magnitude, and we note that several types of magnitudes are flip-flopping. Several of our comments are clearly not specific to cluster LFs and hold for the field LF too.

Distribution functions in the presence of background (such as the cluster LF in the absence of a redshift survey, but also the H α equivalent width distribution in the presence of a continuum) should:

- (i) be fitted without removing the background contribution, adding instead a background term to the model;
- (ii) be simultaneously fitted with the background distribution;
- (iii) use unbinned data;
- (iv) adopt the likelihood (not the conditional likelihood);
- (v) allow Poisson fluctuations of the whole sample (i.e. include the s term, as prescribed by the extended likelihood approach);
- (vi) avoid the use of \sqrt{n} in place of the 68 per cent confidence interval; and
- (vii) not compute the 68 per cent confidence interval by summing in quadrature the 68 per cent confidence intervals of the addenda.

Two sources of errors are negligible in our work, and therefore neglected. First, the error on the value of the input quantities, which in our case are magnitudes, but in other papers are magnitudes and densities. With the operated choices, magnitudes have negligible errors, and for this reason we have neglected their impact on the LF parameters. If this condition does not arise, it is necessary to convolve the fitting function by the error function, in the way described by Jeffreys (1938). Densities, instead, usually have large errors, as large as 100 per cent (for example, in some 2dF subsamples, from quantities quoted in Croton et al. 2005), simply because densities are computed by counting a small number of galaxies (e.g. as few as 1). Σ_5 , a measure of density derived from the distance of the fifth neighbour, used in some recent density estimates, has a ± 55 per cent error. The presence of large errors on the input quantity further complicates a rigorous determination of the LF and of its dependence on environment. Such a rigorous determination has not yet been published, to the best of our knowledge.

Secondly, we studied the whole galaxy population, and not a minority population. In the latter case, errors in the galaxy classification, even if coming infrequently, pollute the minority population with objects coming from the main sample. Let us consider an example: the LF of a population representing a tiny fraction, say 0.0003, as the fraction of local E+A galaxies. If classification errors concern a fraction of 0.0003 of the whole galaxy population (a very small fraction, indeed), any E+A sample studied is 50 per cent contaminated by objects unrelated to the class aimed to study, and one should not be surprised to ‘discover’ that the selected sample has an LF similar to that of the whole sample, being 50 per cent contaminated.

Such contamination should be accounted for in the LF computation, but it is usually not. Our approach of not subtracting the background from the data, but of adding a background term to the model, accounts for the uncertainty due to the mentioned contamination population.

REFERENCES

- Abell G. O., 1958, *ApJS*, 3, 211
 Abell G. O., Corwin H. G., Olowin R. P., 1989, *ApJS*, 70, 1
 Andreon S., 1998, *A&A*, 336, 98
 Andreon S., 2001, *ApJ*, 547, 62 (Paper I)
 Andreon S., 2002, *A&A*, 382, 821
 Andreon S., 2004, *A&A*, 416, 865
 Andreon S., Cuillandre J.-C., 2002, *ApJ*, 569, 144
 Avni Y., 1976, *ApJ*, 210, 642
 Barger A. J., Aragon-Salamanca A., Ellis R. S., Couch W. J., Smail I., Sharples R. M., 1996, *MNRAS*, 279, 1
 Bertin E., Arnouts S., 1996, *A&AS*, 117, 393
 Blanton M. R. et al., 2003, *ApJ*, 592, 819
 Blanton M. R., Lupton R. H., Schlegel D. J., Strauss M. A., Brinkmann J., Fukugita M., Loveday J., 2005, *ApJ*, submitted (astro-ph/0410164)
 Butcher H., Oemler A., 1984, *ApJ*, 285, 426
 Cash W., 1979, *ApJ*, 228, 939
 Couch W. J., Newell E. B., 1984, *ApJS*, 56, 143
 Couch W. J., Sharples R. M., 1987, *MNRAS*, 229, 423
 Couch W. J., Barger A. J., Smail I., Ellis R. S., Sharples R. M., 1998, *ApJ*, 497, 188
 Couch W. J., Balogh M. L., Bower R. G., Smail I., Glazebrook K., Taylor M., 2001, *ApJ*, 549, 820
 Cristóbal-Hornillos D., Balcells M., Prieto M., Guzmán R., Gallego J., Cardiel N., Serrano Á., Pelló, R., 2003, *ApJ*, 595, 71
 Croton D. J. et al., 2005, *MNRAS*, 356, 1155
 Da Costa L. et al., 1998, preprint (astro-ph/9812105)
 Devillard N., 1997, *The Messenger* 87, 19
 Dunkley J., Bucher M., Ferreira P. G., Moodley K., Skordis C., 2005, *MNRAS*, 356, 925
 Efstathiou G., Ellis R. S., Peterson B. A., 1988, *MNRAS*, 232, 431 (EEP)
 Folkes S. et al., 1999, *MNRAS*, 308, 459
 Garilli B., Maccagni D., Andreon S., 1999, *A&A*, 342, 408
 Gehrels N., 1986, *ApJ*, 303, 336
 Jeffreys H., 1938, *MNRAS*, 98, 190
 Hansen S. M., McKay T. A., Wechsler R. H., Annis J., Sheldon E. S., Kimball A., 2005, *ApJ*, submitted (astro-ph/0410467)
 Huang J.-S., Cowie L. L., Gardner J. P., Hu E. M., Songaila A., Wainscoat R. J., 1997, *ApJ*, 476, 12
 Hubble E., 1936, *ApJ*, 84, 158
 Jones L., Smail I., Couch W. J., 2000, *ApJ*, 528, 118
 Kron R. G., 1980, *ApJS*, 43, 305
 Lin H., Yee H. K. C., Carlberg R. G., Morris S. L., Sawicki M., Patton D. R., Wirth G., Shepherd C. W., 1999, *ApJ*, 518, 533
 Mobasher B. et al., 2003, *ApJ*, 587, 605
 Moorwood A., Cuby J.-G., Lidman C., 1998, *The Messenger*, 91, 9
 Oemler A. Jr., 1974, *ApJ*, 194, 1
 Paolillo M., Andreon S., Longo G., Puddu E., Gal R. R., Scaramella R., Djorgovski S. G., de Carvalho R., 2001, *A&A*, 367, 59
 Persson S. E., Murphy D. C., Krzemiński W., Roth M., Rieke M. J., 1998, *AJ*, 116, 2475
 Popesso P., Böhringer H., Brinkmann J., Voges W., York D. G., 2004, *A&A*, 423, 449
 Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P., 1993, *Numerical Recipes*. Cambridge Univ. Press, Cambridge
 Protassov R., van Dyk D. A., Connors A., Kashyap V. L., Siemiginowska A., 2002, *ApJ*, 571, 545
 Rengelink R. et al., 1998, preprint (astro-ph/9812190)
 Rojas R. R., Vogeley M. S., Hoyle F., Brinkmann J., 2004, *ApJ*, submitted (astro-ph/0409074)

- Sandage A., Tammann G. A., Yahil A., 1979, ApJ, 232, 352 (STY)
Schechter P., 1976, ApJ, 203, 297
Smail I., Ellis R. S., Dressler A., Couch W. J., Oemler A. J., Sharples R. M., Butcher H., 1997, ApJ, 479, 70
Smith I. A., Tilanus R. P. J., Wijers R. A. M. J., Tanvir N., Vreeswijk P., Rol E., Kouveliotou C., 2001, A&A, 380, 81
Stanford S. A., Eisenhardt P. R., Dickinson M., Holden B. P., De Propris R., 2002, ApJS, 142, 153
Takeuchi T. T., 2000, Ap&SS, 271, 213
Toft S., Soucail G., Hjorth J., 2003, MNRAS, 344, 337
Valotto C., Giovanelli R., 2004, AJ, 128, 115
Valotto C. A., Moore, B., Lambas D. G., 2001, ApJ, 546, 157
Vandame B. et al., 2001, preprint (astro-ph/0102300)
Wheaton W. A., Dunklee A. L., Jacobsen A. S., Ling J. C., Mahoney W. A., Radocinski R. G., 1995, ApJ, 438, 322
Wilks S., 1938, Ann. Math. Stat. 9, 60
Wilks S., 1963, Mathematical Statistics. Princeton Univ. Press, Princeton, NJ
York D. G. et al., 2000, AJ, 120, 1579
Zwicky F., 1951, PASP, 63, 61
Zwicky F., 1957, Morphological Astronomy. Springer-Verlag, Berlin

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