The Astrophysical Journal Supplement Series, 173:494–502, 2007 December

CLUSTERING PROPERTIES OF REST-FRAME UV-SELECTED GALAXIES. I. THE CORRELATION LENGTH DERIVED FROM *GALEX* DATA IN THE LOCAL UNIVERSE

Bruno Milliard,[1] Sébastien Heinis,[1,2] Jérémy Blaizot,[1,3] Stéphane Arnouts,[1] David Schiminovich,[4] Tamás Budavári,[2] José Donas,[1] Marie Treyer,[1] Michel Laget,[1] Maurice Viton,[1] Ted K. Wyder,[5] Alex S. Szalay,[2] Tom A. Barlow,[5] Karl Forster,[5] Peter G. Friedman,[5] D. Christopher Martin,[5] Patrick Morrissey,[5] Susan G. Neff,[6] Mark Seibert,[5] Todd Small,[5] Luciana Bianchi,[7] Timothy M. Heckman,[2] Young-Wook Lee,[8] Barry F. Madore,[9,10] R. Michael Rich,[11] Barry Y. Welsh,[12] Sukyoung K. Yi,[8] and C. K. Xu[5]

Received 2006 August 22; accepted 2007 October 16

ABSTRACT

We present the first measurements of the angular correlation function of galaxies selected in the far (1530 Å) and near (2310 Å) ultraviolet from the *GALEX* survey fields overlapping SDSS DR5 in low Galactic extinction regions. The area used covers 120 deg^2 (*GALEX* Medium Imaging Survey) down to magnitude AB = 22, yielding a total of 100,000 galaxies. The mean correlation length is $\sim 3.7 \pm 0.6$ Mpc, and no significant trend is seen for this value as a function of the limiting apparent magnitude or between the *GALEX* bands. This estimate is close to that found from samples of blue galaxies in the local universe selected in the visible and similar to that derived at $z \simeq 3$ for LBGs with similar rest frame selection criteria. This result supports models that predict antibiasing of star-forming galaxies at low redshift and brings an additional clue to the downsizing of star formation at $z < 1$.

Subject headings: stars: formation — ultraviolet: galaxies

1. INTRODUCTION

In the current paradigm of structure formation, the bulk of the most massive systems form in a cold dark matter–dominated universe by the merging of less massive units formed earlier. In parallel to this hierarchical evolution, recent observations point to "downsizing," namely, the fact that in galaxies having high baryonic masses the bulk of stars formed at high redshift ($z \gtrsim 1$), while in galaxies having low baryonic masses the bulk of stars formed at lower redshift (Cowie et al. 1996; Heavens et al. 2004; Bundy et al. 2006; Jimenez et al. 2005; see also De Lucia et al. 2006 and Neistein et al. 2006 for results from simulations). The star formation efficiency shows a strong decline at $0 < z < 1$, as measured by the evolution of the star formation rate density (Hopkins & Beacom 2006; Lilly et al. 1996; Schiminovich et al. 2005; Sullivan et al. 2000; Wilson et al. 2002). These epochs also see the bulk of the build-up of the bimodality in galaxy properties of the local universe, which is apparent in their color distribution (Baldry et al. 2004) morphologies (Kauffmann et al. 2004), spectral class (Madgwick et al. 2002), and spatial distribution (Budavari et al. 2003). Understanding the full picture is complex as this evolution is the result of the interplay of several physical processes (Faber et al. 2005) and the combination of the effects of initial galaxy formation conditions ("nature") with galaxy evolution events ("nurture") (Kauffmann et al. 2004). In this context, tracers that measure over cosmic time galaxy populations selected with homogeneous physical criteria are of primary interest. They help compare observations to simulation predictions over a large range of redshifts with reduced uncertainties and allow a study of the redshift evolution of galaxy properties derived from different surveys.

The ultraviolet (UV) range of the spectrum meets these conditions: UV luminosities provide a good measure of recent star formation within galaxies (Kennicutt 1998), modulo attenuation by dust, and have been widely used at high redshifts to study the properties of the Lyman break galaxies (LBGs; Giavalisco & Dickinson 2001; Shapley et al. 2003; Steidel et al. 1995). As large amounts of data are now becoming available at lower redshifts as part of the *GALEX* surveys (Martin et al. 2005), the rest-frame UV spectral domain is presently well sampled over the full $0 < z < 6$ redshift range. Furthermore, comparison of results from high- and low-z UV-selected samples is eased by the fact that the UV luminosity density fractions[13] probed at high and low z are similar (Heinis et al. 2007, hereafter Paper II), due to the strong luminosity evolution of the UV luminosity function (Arnouts et al. 2005). Noticeably, during the epochs probed by *GALEX* the properties of active star-forming galaxies show a very fast evolution.

The wealth of UV-selected data now available at low redshifts enables statistical studies in the context of the downsizing of star formation, and in particular searches for links between the star

[1] Laboratoire d'Astrophysique de Marseille, BP 8, Traverse du Siphon, 13376 Marseille Cedex 12, France.
[2] Department of Physics and Astronomy, The Johns Hopkins University, Homewood Campus, Baltimore, MD 21218.
[3] Max Planck Institut für Astrophysik, D-85748 Garching, Germany.
[4] Department of Astronomy, Columbia University, New York, NY 10027.
[5] California Institute of Technology, MC 405-47, 1200 East California Boulevard, Pasadena, CA 91125.
[6] Laboratory for Astronomy and Solar Physics, NASA Goddard Space Flight Center, Greenbelt, MD 20771.
[7] Center for Astrophysical Sciences, The Johns Hopkins University, 3400 North Charles Street, Baltimore, MD 21218.
[8] Center for Space Astrophysics, Yonsei University, Seoul 120-749, Korea.
[9] Observatories of the Carnegie Institution of Washington, 813 Santa Barbara Street, Pasadena, CA 91101.
[10] NASA/IPAC Extragalactic Database, California Institute of Technology, Mail Code 100-22, 770 South Wilson Avenue, Pasadena, CA 91125.
[11] Department of Physics and Astronomy, University of California, Los Angeles, CA 90095.
[12] Space Sciences Laboratory, University of California at Berkeley, 601 Campbell Hall, Berkeley, CA 94720.

[13] We define the UV luminosity density fraction of a given sample as the ratio of the UV luminosity density encompassed by the sample over the total UV luminosity density at the same redshift.

formation properties and galaxy environment in terms of galaxy or dark matter density. Here we focus on the evolution with redshift of the link of star formation with dark matter and particularly the evolution of the class of dark matter halos hosting actively star-forming galaxies since $z \sim 1$. This can be achieved by the study of the clustering of galaxies: at high redshift, LBGs studies show that UV-selected galaxies inhabit high galaxy density regions (Steidel et al. 1998) and are strongly biased with respect to the underlying dark matter, with more actively star-forming galaxies being more biased (Adelberger et al. 2005; Giavalisco & Dickinson 2001; Foucaud et al. 2003; see Giavalisco 2002 for a review on the properties of LBGs). We propose to extend such studies to low redshifts using similar selection criteria. This paper is the first in a series and presents the methods and first results of angular clustering measurements from *GALEX* data. Section 2 presents the data sets and the derivation of the redshift distributions. Section 3 presents two methods to derive the angular correlation function from a set of fields and a discussion about the behavior of these methods with respect to photometry inhomogeneity. In § 4 we present our results on the angular correlation functions and correlation lengths. To provide the crucial link to dark matter halos, we use the analytical Mo & White (2002) formalism that we present in § 4.3. We end with a short discussion in § 5.

Throughout the paper a ΛCDM cosmology is assumed with matter density $\Omega_m = 0.3$, vacuum energy density $\Omega_\Lambda = 0.7$, and a Hubble parameter $h = 0.7$, where $H_0 = 70$ km s^{-1} Mpc^{-1}. All correlation length values taken from the literature have been converted accordingly using equation (4) in Magliocchetti et al. (2000).

2. PRIMARY FIELDS SELECTION

We use Medium Imaging Survey (MIS) fields from the *GALEX* Release 2 (GR2), which allows us to probe the clustering of faint sources in FUV and NUV at three limiting apparent AB magnitudes 22.0, 21.5, and 21.0. The magnitudes we refer to are corrected for Galactic extinction using Schlegel maps (Schlegel et al. 1998) and the Cardelli et al. (1989) extinction law, unless specified otherwise. The average color excess in the fields, derived from the Schlegel et al. maps, ranges from 0.08 to 0.12. The extinction coefficients $A_{\rm FUV}/E(B-V)$ and $A_{\rm NUV}/E(B-V)$ are, respectively, 8.29 and 8.61.

We start with 348 MIS *GALEX* fields overlapping SDSS DR5, of which only a subset will be kept because of Galactic extinction (see § 3.2). We only include sources within a 0.5° radius from the field center, since artifacts concentrate near the periphery of the field of view (see Morrissey et al. 2005) and photometric accuracy decreases beyond this limit. We used only objects within the *GALEX* primary resolution. We used SDSS masks to mask out holes, bright stars and trails; we used also *GALEX* masks, as well as additional ones to mask out resolved galaxies or artifacts not predicted by the *GALEX* pipeline. Only objects with an SDSS match within 4″ are kept, and the closest match is used. Galaxies are assumed to be SDSS galaxies (*type* = 3 following the morphological classification of Lupton et al. 2001 and Stoughton et al. 2002). To check the effects of possible residual QSOs in our sample, we removed from our sample AGN-dominated objects as objects classified QSOs by a template fitting procedure.[14] Impacts on the results are found negligible; hence, we do not remove QSOs classified objects from the sample in the following.

[14] Le Phare: available and documented at http://www.oamp.fr/arnouts/LE_PHARE.html.

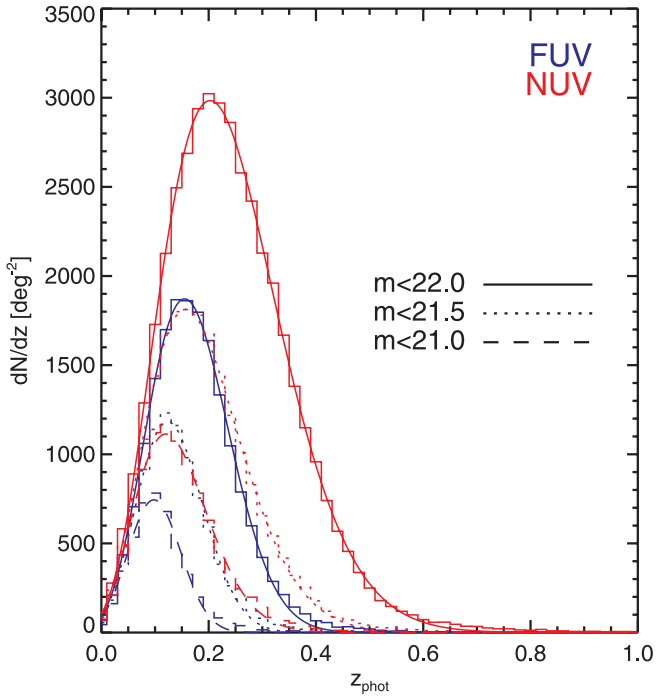

FIG. 1.—Derived redshift distributions for the two *GALEX* bands using three magnitude cuts (*histograms*). The solid curves show the best fitting $N(z)$ convolved by a Gaussian with $\sigma = 0.03$ (see text).

2.1. *Redshift Distributions*

To get the redshift distributions of the samples for the different magnitude cuts, we use the polynomial fit method described in Connolly et al. (1995, 1997). This method requires first to be trained with a spectroscopic sample. We train on 6 bands (NUV from *GALEX*, as all objects do not have FUV photometry, and the 5 SDSS bands) with 17,843 objects from the sample described in § 2 having SDSS spectroscopic redshifts. We then apply the coefficients derived from the training set to the whole sample.

We performed a simple correction for the broadening due to photometric redshifts errors by assuming that the photometric redshifts errors follow a normal distribution independent of the object magnitudes and redshifts, with the standard deviation $\sigma = 0.03$ measured using all the available spectroscopic redshifts from the SDSS. We check that the standard deviation does not vary with apparent magnitude using our photometric redshift estimation on *GALEX* fields with SDSS overlap and the independent and deeper spectroscopy from Papovich et al. (2006).

Following Efstathiou et al. (1991) the parent distribution of the true redshifts is described by the following parametric shape:

$$\frac{dN}{dz} = A_z \left(\frac{z}{z_c}\right)^2 \exp\left[-\left(\frac{z}{z_c}\right)^n\right]. \quad (1)$$

We fit this shape convolved by a Gaussian with $\sigma = 0.03$ to the observed photometric redshift distribution. Figure 1 shows the Gaussian-convolved best estimate $N(z)$, along with the measured distribution for the 174 least extinct fields (see § 3.2), and Table 1 lists the parameters of the true distributions.

3. CHOICE OF ACF ESTIMATION METHOD

3.1. *Methods*

Given that the *GALEX* data are extracted from relatively large numbers of similar exposures, the question arises as to the most

TABLE 1
SAMPLE DESCRIPTION, POWER-LAW BEST-FIT PARAMETERS AND COMOVING CORRELATION LENGTHS

	FUV			NUV		
PARAMETER	22. mag	21.5 mag	21. mag	22. mag	21.5 mag	21. mag
N_{gal}[a]	44651	22655	11418	99368	48274	22948
$\bar{z}$[b]	0.18	0.15	0.12	0.25	0.21	0.16
A_z[c]	5485.96	3616.28	2417.72	8255.5	5084.1	3291.14
z_c[c]	0.16	0.13	0.11	0.18	0.13	0.1
n[c]	2.18	2.29	2.76	1.67	1.61	1.66
$A_w \times 10^3$ (deg$^\delta$)	$8.0^{+1.7}_{-1.5}$	$9.7^{+3.7}_{2.7}$	$10.4^{+7.2}_{-4.4}$	$3.2^{+0.5}_{-0.5}$	$4.6^{+1.1}_{-0.8}$	$6.3^{+0.9}_{-0.9}$
δ	0.80 ± 0.05	0.80 ± 0.08	0.75 ± 0.12	0.89 ± 0.04	0.88 ± 0.05	0.84 ± 0.09
r_0 (Mpc)	$4.2^{+0.5}_{-0.4}$	$3.7^{+0.7}_{-0.6}$	$2.8^{+0.9}_{-0.7}$	$4.0^{+0.3}_{-0.3}$	$3.7^{+0.4}_{-0.3}$	$3.3^{+0.7}_{-0.5}$

NOTES.—The amplitude and slope of best-fit power laws to the angular correlation function, and hence the comoving correlation length account for the integral constraint correction (see text). No attempt to remove residual QSOs from photometric redshifts is performed here.
[a] Number of galaxies in the samples.
[b] Mean photometric redshift.
[c] Parameters of the true best-fit redshift distribution (see text).

appropriate method to retrieve the available information. A first, straightforward approach to measure the ACF from a group of nonoverlapping fields is to treat all of them as disjoint subfields of one large, discontinuous field and to apply the Landy & Szalay (1993, hereafter LS93) estimator on it:

$$w_{\rm CF} = \frac{\rm DD - 2DR + RR}{\rm RR}, \qquad (2)$$

where DD, DR, and RR are, respectively, the number of data-data, data-random, and random-random pairs from all fields (including cross pairs from different subfields), normalized by the suitable pair numbers. In the case of this composite field method (CF), the number of random points is fixed for the global field, and not for each individual *GALEX* field. This is the ideal method, which in principle allows one to extract all the available information. In particular, this method reduces the integral constraint bias and the noise, especially at large angular separations.

Although best in the ideal case, the CF method requires precise homogeneity of the data and may not be robust in practice. We therefore introduce another estimator, which we define as the following pair-weighted average (PW) of the ACF measured in each field individually:

$$w_{\rm PW}(\theta) = \frac{\sum_i \widetilde{\rm RR}_i(\theta) w_i(\theta)}{\sum_i \widetilde{\rm RR}_i(\theta)}, \qquad (3)$$

where w_i is the ACF estimated from field i alone computed with the LS93 estimator and $\widetilde{\rm RR}_i$ is the number of random-random pairs in the random catalog constructed for this field (see the Appendix for a derivation of this formula (eq. [3] from $w_{\rm CF}$). The $\widetilde{\rm RR}_i$ term involves pair numbers and field geometry information. The PW method is by construction insensitive to field-to-field fluctuations—and thus best suited for the peculiar MIS geometry. A drawback of the PW method is the increase of the integral constraint (IC) bias because of the smaller angular extent of the field[15] as well as an increase in the noise. The integral constraint can be relatively well corrected for using its estimate given by LS93. To compute it, we assume that the real correlation function is a power law $A_w\theta^{-\delta}$ and we fit $A_w\theta^{-\delta} - I(A_w, \delta)$ to the data, where $I(A_w, \delta) = 1/\Omega^2 \int_\Omega A_w\theta^{-\delta} d\Omega_1 d\Omega_2$, integrated over a *GALEX* field. This method is similar to that used by Roche & Eales (1999), except that δ is left as a free parameter. In the following, "PW method" will refer to the IC-corrected technique. We have checked the accuracy of the above correction of the IC bias using a 100 deg^2 synthetic catalog derived from `GalICS` (Hatton et al. 2003; Blaizot et al. 2005). The ACF has been computed with the CF and the PW methods from 50 randomly positioned fields of radius 0.5°. The `GalICS`-magnitude cut was chosen to obtain approximately the same mean number of galaxies as found using the NUV < 22 cut. The results of the CF and PW methods have been found undistinguishable for the model catalogs (Fig. 2).

[15] In the case of independent fields, the IC in the PW method is typically higher than that of the CF method by a factor of the number of fields.

3.2. *Systematic Effects*

To test the sensitivity of the CF and PW methods to systematics, we used a statistical approach to decide whether the photometry of a given field is drawn from the same distribution than

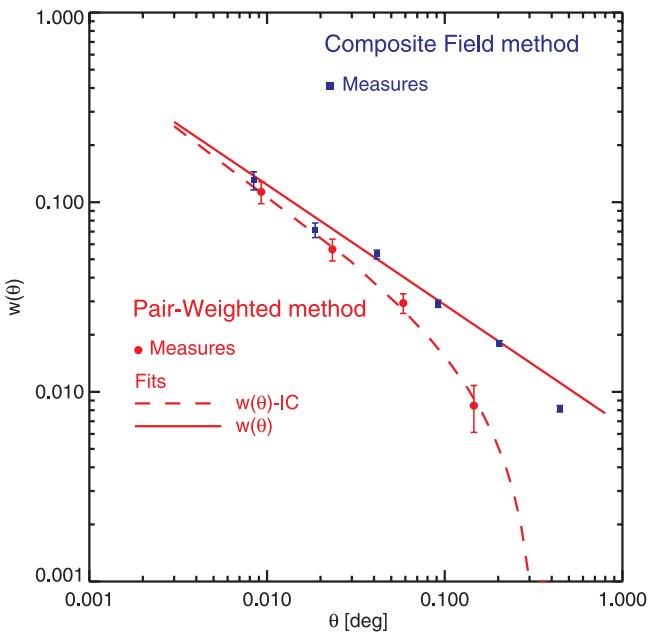

FIG. 2.—Validation of the method used to correct for the integral constraint bias. The ACF is computed from 50 randomly chosen fields in a synthetic catalog with the CF (*filled squares*) and the PW (*filled circles*) methods. The dashed (solid) line shows the best fit of the PW result uncorrected (corrected) for the integral constraint bias (see text for details).

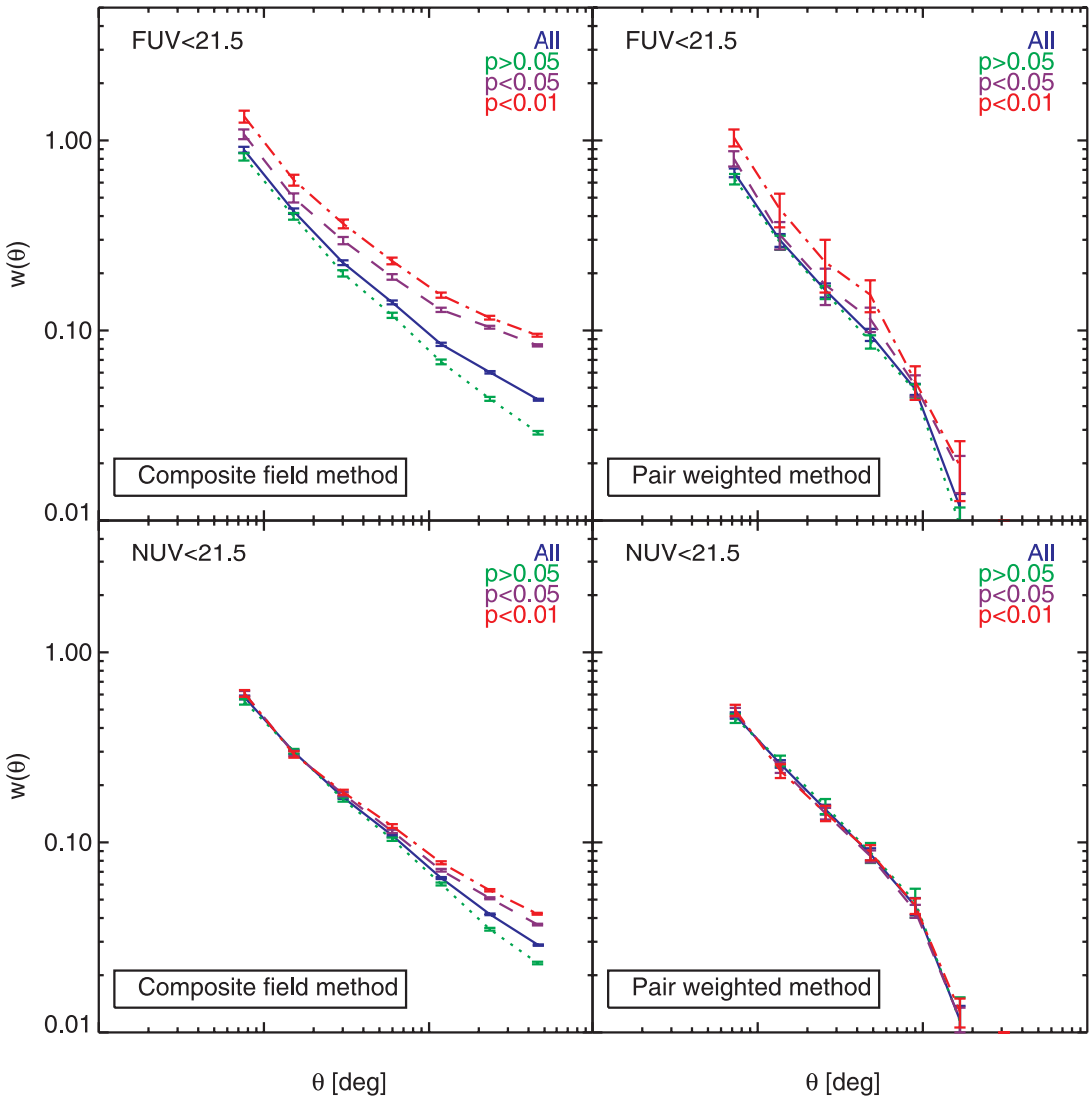

FIG. 3.— Angular correlation function for the CF (*left*) and PW (*right*) methods (see text). Upper panels show FUV < 21.5 selection and lower panels NUV < 21.5. The solid line shows the ACF for the whole sample; the other curves represent the ACF of the fields according to the probability that their magnitude distribution is drawn from the sample distribution than the whole sample (*dot-dashed line*, *dashed line*, $p < 0.01$; $p < 0.05$; *dotted line*, $p > 0.05$).

the photometry of the whole sample. To this aim, we used the Mann-Whitney test, which is independent of the size of the input samples; we use only objects with *type* $= 3$ during this process to avoid strong star contamination. For each field, we create a test sample built from the magnitudes of the objects in this field, and a control sample from the magnitudes of the objects that belong to all other fields. The Mann-Whitney test provides as output the probability that these two distributions are the same. We show in Figure 3 the ACFs of fields grouped according to their value of this probability, using the CF method (*left*) or the PW method (*right*), for FUV < 21.5 (*top*) or NUV < 21.5 (*bottom*). The results obtained from the CF and PW methods show significant differences. The CF method results show an excess of power especially at large scales; the amplitude of this excess increases as the probability that the field photometry is the same than the overall sample photometry decreases. Conversely, the results of the PW method are fairly insensitive to photometry inhomogeneities; there is an overall power excess at FUV < 21.5 for the fields with a probability lower than 0.01, but the ACF of the whole sample is very similar to the ACF of the best fields ($p > 0.05$).

There are several sources of systematic errors, which, although their individual effects are weak, may, combined with each other, yield the trends observed. Similar trends are observed when the fields are binned according to the mean Galactic extinction. However, the cross-correlation between galaxies and dust maps using both CF and PW methods is found at least 5 times lower than the autocorrelation at scales where the latter is positive ($\theta \lesssim 0.2^\circ$), and no obvious trend was found between the amplitudes of this cross-correlation and the Galactic extinction. On the other hand, the amplitude of the cross-correlation function between galaxies and background maps is higher in fields with higher mean Galactic extinction. Inhomogeneities may also arise from photometry drift with time, but sources drifted less than 0.1 mag (Morrissey et al. 2007) over the whole *GALEX* mission; a drift of this amplitude has a small effect on the CF method, as expected from tests on mock catalogs. Studying sources observed several times in overlapping regions shows that field-to-field fluctuations are less than 10% beyond what is expected from Poisson statistics. Note, however, that this result is based on a few sources per field located at the edges of the field, where photometry accuracy decreases. Star contamination can lower the amplitude of the PW method, as an

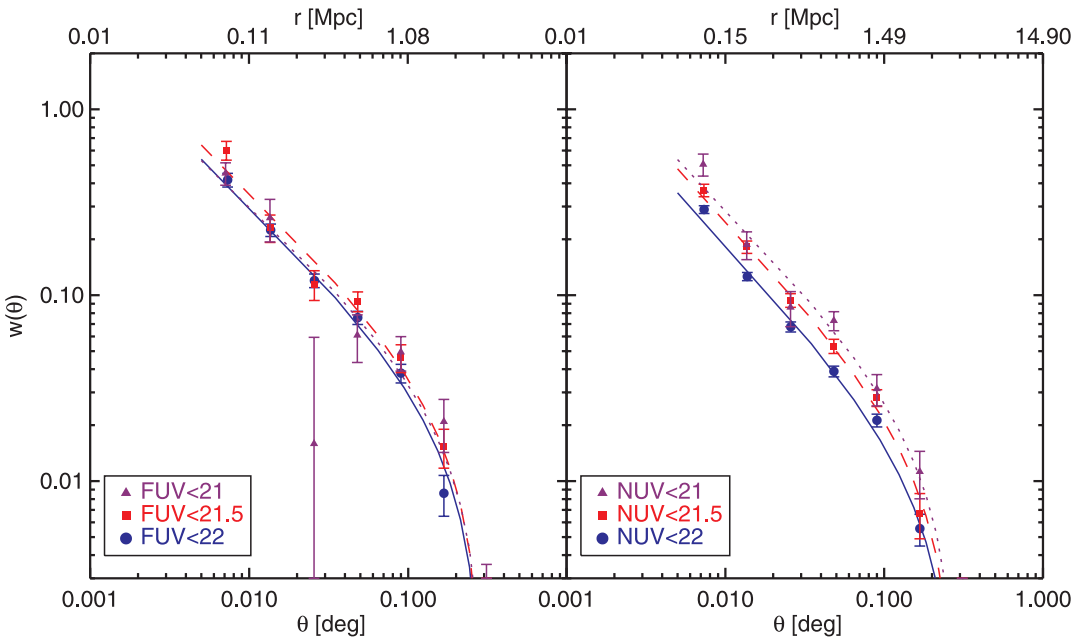

FIG. 4.—Angular correlation function measured in the 174 *GALEX* fields with the lowest Galactic extinction, for the FUV (*left*) and NUV (*right*) *GALEX* bands at three magnitude cuts: *circles*, $m_{\rm UV} < 22$; *squares*, $m_{\rm UV} < 21.5$ and *triangles*, $m_{\rm UV} < 21$. Dashed lines show the power-law best fit uncorrected for integral constraint. The upper axis shows the comoving distances corresponding to the angular scales at $z = 0.15$. No attempt to remove residual active nuclei by photometric redshift template fitting has been made beyond the SDSS classification.

addition of an uncorrelated population, while it may contribute to the effects observed with the CF method, given the variations of star counts with Galactic latitude. According to template-fitting–based classification, the fraction of stars in SDSS objects with *type* = 3 is 2% in NUV and 8% in FUV; we checked that this has small effect on the PW method.

All these tests suggest that there is some source of field-to-field variations in our data, likely due to a combination of zero-point calibrations, background fluctuations (correlated with Galactic extinction), etc. The PW method is fairly insensitive to any systematics, as expected, and we are thus confident that it is a robust estimator. It is this method we chose to use in the rest of the paper.

Conservatively, for the remainder of the paper, we restrict the analysis to the 174 fields with the lowest Galactic extinction ($\langle E(B-V)\leq 0.04\rangle$). The number of galaxies at the different limiting magnitude cuts are given in Table 1.

The characteristics of the UV dust attenuation in galaxies are not known to be correlated with the large-scale structure or the galaxy density (even in the extreme cases of clusters; see, e.g., Boselli & Gavazzi 2006); the effect of internal dust has thus been taken as an uncorrelated noise source on the UV fluxes and its effect on the ACF neglected. This allows direct comparison with clustering studies of high-redshift rest-frame UV-selected galaxies.

4. ANGULAR CORRELATION FUNCTION AND CORRELATION LENGTH

4.1. *Measurements*

We use the PW method described in § 3 to measure $\omega(\theta)$ from the 174 fields with the lowest Galactic extinction, using logarithmic-width bins of $\Delta\log\theta = 0.27$, and $\theta_{\rm min} = 0.005^\circ$ and $\theta_{\rm max} = 0.4^\circ$, which probes scales in the range 0.05–4 comoving Mpc at the median redshift ($z = 0.15$) of the samples considered here. The results are plotted in Figure 4; the 1 σ error bars represent internal scatter derived from jackknife resampling of the 174 *GALEX* fields used for the ACF. In order to check for any instrumental contribution to the ACF such as residual nonuniformities of the sensitivity across the field of view, the PW method has been applied to stars, selected as objects whose SDSS counterparts with *type* = 6. For stars we find no significant deviation from a null correlation function.

We fit the results using the method described in § 3. Our best fits for the different samples are given in Table 1, where θ is expressed in degrees. The error bars on A_w and δ are the projected $(\chi^2_{\rm min}+1)$ contour.

To derive the comoving correlation length, r_0, we used the Limber equation (Peebles 1980) with the true deconvolved redshift distributions (see in § 2.1). The results are given in Table 1. The uncertainties on r_0 have been assumed to be the extreme excursions of r_0 in the projection in the (r_0, δ) and (r_0, A_w) planes of the χ^2 contour at the 68% probability in the (A_w, δ) plane (see Fig. 5).

4.2. *Comparison with Previous Studies*

Given the error bars, the slopes δ found for the different magnitude cuts in the two bands are compatible with a constant value $\delta \simeq 0.81 \pm 0.07$. This is steeper than reported in several studies based on blue galaxies at low z (Budavari et al. 2003; Zehavi et al. 2002; Madgwick et al. 2003), and rest-frame UV-selected galaxies at higher redshifts (Adelberger et al. 2005; Porciani & Giavalisco 2002), all of them consistent with a value of $\delta \simeq 0.6$. However, our measurement is in agreement with Giavalisco et al. (1998), Giavalisco & Dickinson (2001), and Foucaud et al. (2003). Moreover, Coil et al. (2004) noticed a steepening of the slope not only for the reddest but also for the bluest galaxies of their samples. We discuss in Paper II the dependence of δ on UV luminosity.

With an average comoving correlation length 3.7 ± 0.6 Mpc at $z \sim 0.2$, the present *GALEX* data sets confirm the low clustering of the rest-frame UV-selected galaxies in the local universe observed by Heinis et al. (2004). The new mean value is 25% lower, although both measurements agree within error bars. Assuming the average values of r_0 and δ quoted above, the corresponding

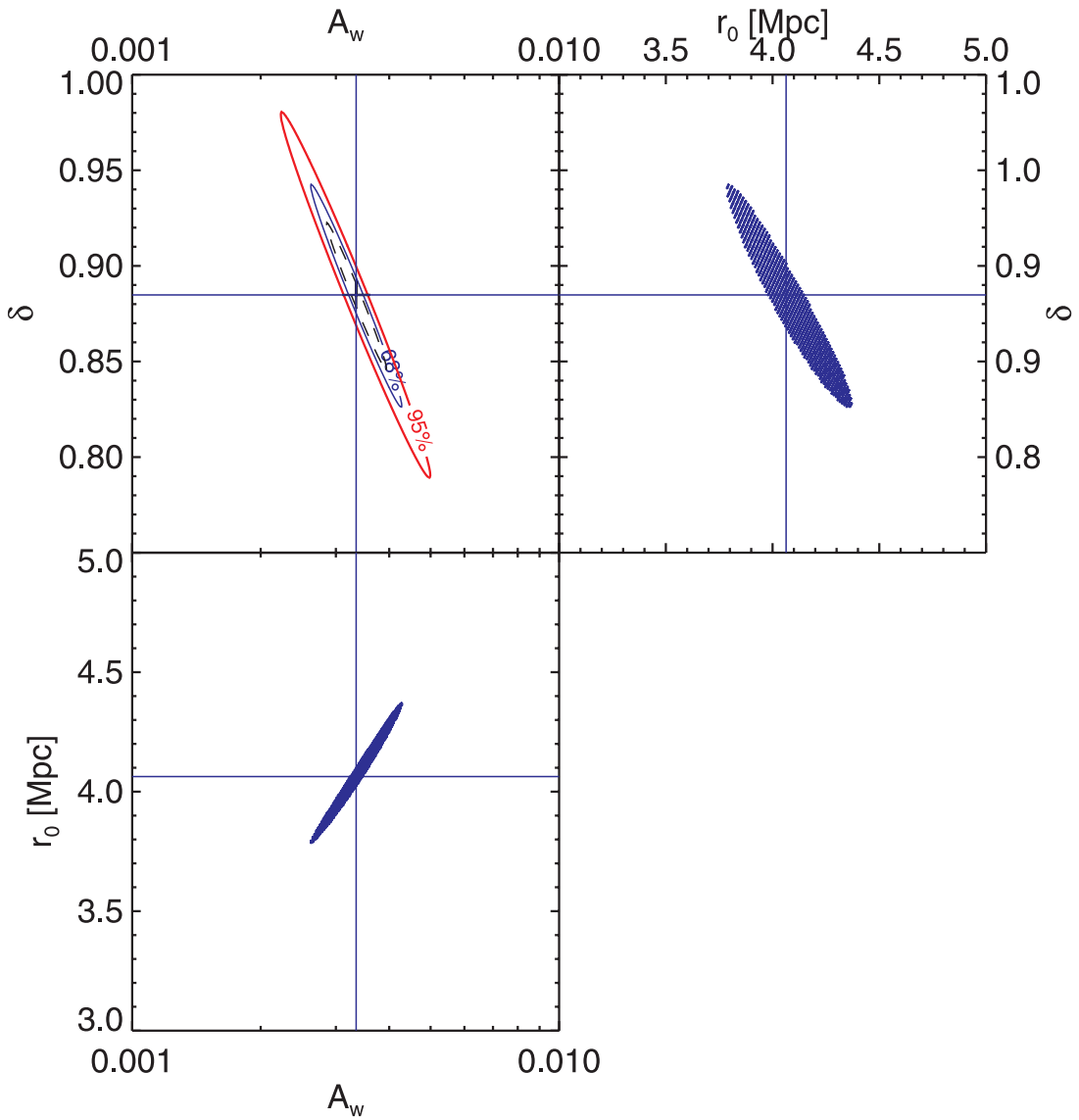

FIG. 5.—χ^2 contours and derivations of errors bars on A_w, δ and r_0 for the NUV < 22 result. *Top left*: contours of constant χ^2 in the (A_w, δ) plane. The inner (outer) solid line corresponds to the 68.3% (95.4%) confidence level. The dashed line shows the $\chi^2_{\rm min} + 1$ contour; its projections on the axes give the error bars on A_w and δ. *Top right*: comoving r_0 as a function of δ given A_w using the values of A_w and δ included in the 68.3% confidence level. *Bottom left*: comoving r_0 as a function of A_w given δ using the same values. The errors bars on r_0 are the extrema of this distribution. The solid line intersections show the location of the best fit in each panel.

bias defined at 8 Mpc by $b_8 = \sigma_{8,g}/\sigma_{8,m}$ (e.g., Magliocchetti et al. 2000) is 0.61 ± 0.09 at $z = 0.2$, a significant antibias.

It is well known that blue galaxies are less strongly correlated than red ones, and not surprisingly the small correlation length found in this study is comparable to that measured for blue galaxies in the local universe: Coil et al. (2004) report a comoving r_0 of 2.54 ± 0.37 Mpc for the class of blue galaxies defined by ($0.2 < R - I < 0.4$) in their visible-selected sample, which spans the redshift range 0.3–0.6, with the lowest correlation length among all of their galaxy subsamples. The *GALEX* rest-frame UV-selected galaxies are nevertheless even less correlated than the galaxy class T4 (bluest of 4 classes) from Budavari et al. (2003) for which they derive a r_0 of 6.44 ± 0.27 Mpc. Hawkins et al. (2001) computed the redshift-space correlation function from far-infrared–selected galaxies in the local universe ($z \sim 0.03$). Converted to real-space, their estimate of the correlation length ($r_0 = 5 \pm 0.33$ Mpc) of the hotter galaxies, i.e., the most star-forming, is higher than ours from UV-selected galaxies.

Low-redshift rest-frame UV-selected galaxies possess correlation lengths slightly lower than those derived from high-z rest-frame UV-selected samples (see Fig. 6). Note that the comparison with results from higher z samples is not straightforward because of the UV luminosity segregation (Giavalisco & Dickinson 2001; Adelberger et al. 2005; Zehavi et al. 2005; Norberg et al. 2002): brighter objects are more clustered than fainter ones. The *GALEX* samples are the faintest of the rest-frame UV-selected samples considered here: the mean absolute magnitudes of the FUV and NUV samples are $M_{\rm FUV} = -18.3$ and $M_{\rm NUV} = -18.8$, while the LBG samples of Adelberger et al. (2005), Arnouts et al. (2002), Foucaud et al. (2003), and Giavalisco & Dickinson (2001) are all brighter than $M_{\rm UV} = -20$. We study in details the luminosity dependence of clustering within the *GALEX* samples in Paper II.

4.3. *Comparison with Dark Matter Halo Clustering Predictions*

In this section we use the formalism described by Mo & White (2002) to compute the correlation length of dark matter halos

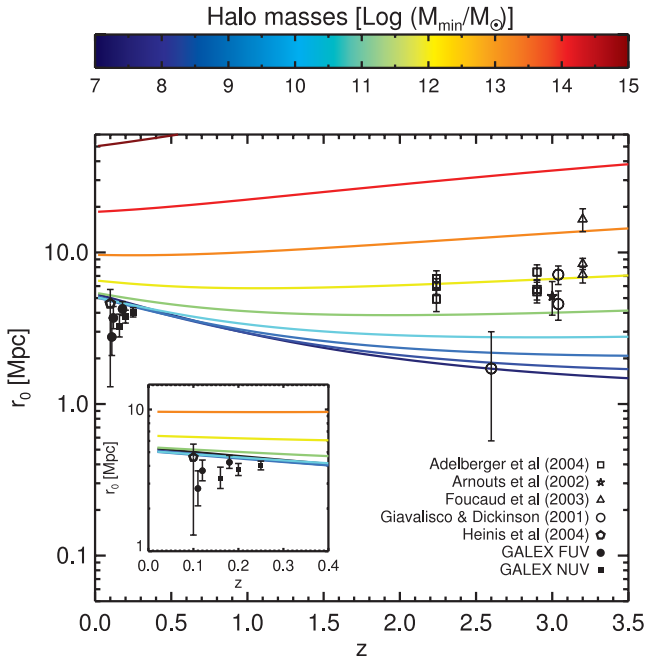

FIG. 6.—Comparison of the evolution with redshift of the correlation lengths obtained from rest-frame UV-selected samples with the correlation lengths of dark matter halos more massive than $M_{\rm min}$ (color-coded). The inset shows the low-z points.

(DMHs) above a given mass as a function of redshift. We assume that the spatial correlation function of DMHs with masses greater than $M_{\rm min}$ at a redshift z is well approximated by a power law:

$$\xi(r,\ M_{\rm min},\ z) = \left[\frac{r}{r_{0_h}(M_{\rm min},\ z)}\right]^{-\gamma_h} \qquad (4)$$

where $r_{0_h}(M_{\rm min},\ z)$ is the correlation length of such halos. The Mo & White (2002) formalism provides analytical equations for the abundance and the bias factor of the halos, $n(M,\ z)$ and $b(M,\ z)$, given their mass and redshift. The effective bias of the halos more massive than a minimum mass $M_{\rm min}$ at a redshift z is then given by

$$b_{\rm eff}(M_{\rm min},\ z) = \frac{\int_{M_{\rm min}}^{\infty} b(M,\ z) n(M,\ z)\, dM}{\int_{M_{\rm min}}^{\infty} n(M,\ z)\, dM}. \qquad (5)$$

The rms density fluctuations of the halos is linked to the rms density fluctuations of the underlying mass at 8 h^{-1} Mpc by $\sigma_{8,h}(M_{\rm min},\ z) = b_{\rm eff}(M_{\rm min},\ z)\sigma(z)_{8,m}$, where the subscripts h and m denote, respectively, halos and underlying mass, and $\sigma(z)_{8,m} = \sigma(0)_{8,m} D(z)$ (see Mo & White 2002) with $\sigma(0)_{8,m} = 0.9$. The correlation length of the DMHs with masses $M > M_{\rm min}$ at z is then obtained using (e.g., Magliocchetti et al. 2000)

$$\sigma_{8,h}(M_{\rm min},\ z) = \sqrt{C_{\gamma_h}\left[\frac{r_{0_h}(M_{\rm min},\ z)}{8}\right]^{\gamma_h}} \qquad (6)$$

where $C_\gamma = 72/[(3-\gamma)(4-\gamma)(6-\gamma)2^\gamma]$. We assumed that the slope of the spatial correlation function of the halos is $\gamma_h = 1.8$, after having checked that the results are rather insensitive to the adopted value if $1.5 < \gamma_h < 2.5$. Figure 6 shows the redshift evolution of the correlation length of DMH with masses $10^7\ M_\odot < M_{\rm min} < 10^{15}\ M_\odot$.

In the framework of Halo Occupation Distribution (HOD) models (e.g., Berlind & Weinberg 2002; Cooray & Sheth 2002), recent studies have pointed out that the galaxy correlation function is likely to be the sum of two components. The first component dominates at small scales, describing the correlation of galaxies that are in the same halo, and the second component accounts for galaxies in different halos, dominating at large scales. The present sample, however, does not provide sufficient constraint to fit HOD models to our results—we plan to perform this in future papers. Assuming that the correlation function of DMHs is a power law at all scales and that each halo hosts at most one galaxy, a direct comparison of the *GALEX* results with the correlation lengths of dark matter halos (Fig. 6) shows that the UV-selected galaxies in our samples have the same correlation lengths as halos with masses lower than $M_{\rm min} = 10^{11}\ M_\odot$. At $z > 2$, clustering measurements from LBGs samples show that halos with comparable clustering strengths have $M_{\rm min} \gtrsim 10^{12}\ M_\odot$, as already mentioned by Adelberger et al. (2005) and Giavalisco & Dickinson (2001). These results suggest that the characteristic mass of halos hosting active star formation has decreased from $z = 3$. Note that taking luminosity evolution into account does not weaken this result, since at low redshifts UV-selected samples actually probe the same UV LD fraction as their high-redshift counterparts (see Paper II). This mass evolution, as well as the bias evolution, can be interpreted as additional evidence for the "downsizing" scenario of star formation, although it applies here to the mass of the underlying halo rather than to the baryonic mass. Note that these is a correlation between halo and galaxy mass (Shankar et al. 2006), although its scatter is expected to be stronger for star-forming galaxies (see, e.g., Yoshikawa et al. 2001). Theoretical studies also predict that the SFR increases with halo mass (e.g., at $z = 3$, Bouché et al. 2005); however, in presence of AGN feedback (e.g., Di Matteo et al. 2005; Croton et al. 2006), or when taking into account gravitational heating (Khochfar & Ostriker 2007), this trend reverses at lower redshift. The conclusions of this work are developed in Paper II.

5. CONCLUSIONS

We presented here the first clustering measurements from the *GALEX* data. These data provide a unique basis to statistical studies of star formation in galaxies at low redshift from their UV continuum. The same tracer can now be used in an homogeneous way over a large redshift range ($0 < z < 4$) to investigate the processes driving star formation evolution. We discussed the impact of photometric inhomogeneities on the clustering measurements and used a method insensitive to them. We measured the clustering by the angular correlation function and fitted our results with a power-law parameterization: $w(\theta) = A_w\theta^{-\delta}$. We derive steep slopes, $\delta \simeq 0.81 \pm 0.07$. Assuming photometric redshift estimation, we compute the correlation length, r_0. The results confirm the low clustering of UV-selected galaxies at low redshift ($r_0 = 3.7 \pm 0.6$ Mpc). Comparison with analytical modeling shows that active star forming at $z < 0.4$ present the same correlation lengths than DMHs with $M_{\rm min} < 10^{11}\ M_\odot$. This result is in agreement with the "downsizing" scenario.

It is with great pleasure that we thank Jean-Michel Deharveng for support and discussions. *GALEX* (*Galaxy Evolution Explorer*) is a NASA Small Explorer, launched in 2003 April. We gratefully acknowledge NASA's support for construction, operation, and science analysis for the *GALEX* mission, developed in cooperation with the Centre National d'Etudes Spatiales of France and the Korean Ministry of Science and Technology.

APPENDIX

PAIR-WEIGHTED AVERAGE ESTIMATOR FOR THE ANGULAR CORRELATION FUNCTION

In this section we discuss the weighted estimator presented in § 3. The expression can be derived directly from the definition of the LS93 estimator with some assumptions. The LS93 estimator is

$$w_{\rm LS}(\theta) = \frac{{\rm DD}(\theta) - 2{\rm DR}(\theta) + {\rm RR}(\theta)}{{\rm RR}(\theta)}, \tag{A1}$$

where DD, DR, and RR are normalized by the suitable pairs number

$${\rm DD} = \frac{2\widetilde{\rm DD}}{n_g(n_g-1)}, \tag{A2}$$

$${\rm DR} = \frac{\widetilde{\rm DR}}{n_g n_r}, \tag{A3}$$

$${\rm RR} = \frac{2\widetilde{\rm RR}}{n_r(n_r-1)}, \tag{A4}$$

where n_g is the number of galaxies in the sample and n_r is the number of random objects in the random sample. In the following we do not recall the θ dependence of the different quantities.

Let us consider the case of the CF method (see § 3) applied on N fields positioned on the sky in such a way that no cross pair between objects from different fields has to be accounted for in the computation of $w(\theta)$. The total number of pairs over all the fields in each angular bin can then be expressed using the number of pairs in each field:

$$\widetilde{\rm DD} = \sum_{i=1}^{N} \widetilde{{\rm DD}_i} = \sum_{i=1}^{N} \frac{n_{g_i}(n_{g_i}-1)}{2}\,{\rm DD}_i, \tag{A5}$$

where $\widetilde{{\rm DD}_i}$ is the number of data-data pairs and n_{g_i} the number of galaxies in the ith field. The same equations hold for $\widetilde{\rm DR}$ and $\widetilde{\rm RR}$.

When computing the ACF of one field individually, we consider 100 random samples with the same number of random points that galaxies in this field.[16] The quantity ${\rm RR}_i$ is then the average of the 100 computations. So $n_{g_i} = n_{r_i} = n_i$ and $n_g = n_r = n$. Then the LS93 estimator can be written

$$w = \left[\frac{2\widetilde{\rm RR}}{n(n-1)}\right]^{-1}\left[\frac{2}{n(n-1)}(\widetilde{\rm DD}+\widetilde{\rm RR}) - \frac{2}{n^2}\widetilde{\rm DR}\right]. \tag{A6}$$

Let us consider the term in brackets; with our assumptions it yields

$$\frac{2}{n(n-1)}\sum_i \frac{n_i(n_i-1)}{2}({\rm DD}_i + {\rm RR}_i) - \frac{2}{n^2}\sum_i n_i^2 {\rm DR}_i, \tag{A7}$$

then introduce the term ${\rm RR}_i/{\rm RR}_i$ in both sums:

$$\frac{2}{n(n-1)}\sum_i \frac{n_i(n_i-1)}{2}({\rm DD}_i + {\rm RR}_i)\frac{{\rm RR}_i}{{\rm RR}_i} - \frac{2}{n^2}\sum_i n_i^2 {\rm DR}_i \frac{{\rm RR}_i}{{\rm RR}_i}. \tag{A8}$$

The ACF of the ith field can be written as

$$\begin{aligned} w_i &= w_{1_i} + w_{2_i} + w_{3_i}, \\ w_{1_i} &= \frac{{\rm DD}_i}{{\rm RR}_i}, \\ w_{2_i} &= -2\frac{{\rm DR}_i}{{\rm RR}_i}, \\ w_{3_i} &= \frac{{\rm RR}_i}{{\rm RR}_i}. \end{aligned}$$

[16] In the case of the CF method, the total number of random points would also be fixed to n_g, but the number of random points in each field is allowed to be different of n_{g_i}.

Hence, equation (A8) becomes

$$\frac{2}{n(n-1)}\sum_i \frac{n_i(n_i-1)}{2}\mathrm{RR}_i(w_{1_i}+w_{3_i})+\frac{1}{n^2}\sum_i n_i^2\mathrm{RR}_i w_{2_i}. \tag{A9}$$

At this stage we also assume that $n_i \ll 1$ so that $n_i(n_i-1)\simeq n_i^2$, and hence $n(n-1)\simeq n^2$; equation (A9) yields

$$\frac{2}{n^2}\sum_i \widetilde{\mathrm{RR}}_i w_i. \tag{A10}$$

Coming back to equation (A6), we finally get

$$w_{\mathrm{PW}}(\theta)=\frac{\sum_i \widetilde{\mathrm{RR}}_i(\theta)w_i}{\sum_i \widetilde{\mathrm{RR}}_i(\theta)}. \tag{A11}$$

REFERENCES

Adelberger, K. L., Steidel, C. C., Pettini, M., Shapley, A. E., Reddy, N. A., & Erb, D. K. 2005, ApJ, 619, 697
Arnouts, S., et al. 2002, MNRAS, 329, 355
———. 2005, ApJ, 619, L43
Baldry, I. K., Glazebrook, K., Brinkmann, J., Ivezić, Ž., Lupton, R. H., Nichol, R. C., & Szalay, A. S. 2004, ApJ, 600, 681
Berlind, A. A., & Weinberg, D. H. 2002, ApJ, 575, 587
Blaizot, J., Wadadekar, Y., Guiderdoni, B., Colombi, S. T., Bertin, E., Bouchet, F. R., Devriendt, J. E. G., & Hatton, S. 2005, MNRAS, 360, 159
Boselli, A., & Gavazzi, G. 2006, PASP, 118, 517
Bouché, N., Gardner, J. P., Katz, N., Weinberg, D. H., Davé, R., & Lowenthal, J. D. 2005, ApJ, 628, 89
Budavari, T., et al. 2003, ApJ, 595, 59
Bundy, K., et al. 2006, ApJ, 651, 120
Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, ApJ, 345, 245
Coil, A. L., Newman, J. A., Kaiser, N., Davis, M., Ma, C., Kocevski, D. D., & Koo, D. C. 2004, ApJ, 617, 765
Connolly, A. J., Csabai, I., Szalay, A. S., Koo, D. C., Kron, R. G., & Munn, J. A. 1995, AJ, 110, 2655
Connolly, A. J., Szalay, A. S., Dickinson, M., Subbarao, M. U., & Brunner, R. J. 1997, ApJ, 486, L11
Cooray, A., & Sheth, R. 2002, Phys. Rep., 372, 1
Cowie, L. L., Songaila, A., Hu, E. M., & Cohen, J. G. 1996, AJ, 112, 839
Croton, D. J., et al. 2006, MNRAS, 365, 11
De Lucia, G., Springel, V., White, S. D. M., Croton, D., & Kauffmann, G. 2006, MNRAS, 366, 499
Di Matteo, T., Springel, V., & Hernquist, L. 2005, Nature, 433, 604
Efstathiou, G., Bernstein, G., Tyson, J. A., Katz, N., & Guhathakurta, P. 1991, ApJ, 380, L47
Faber, S. M., et al. 2005, preprint (astro-ph/0506044)
Foucaud, S., McCracken, H. J., Le Fevre, O., Arnouts, S., Brodwin, M., Lilly, S. J., Crampton, D., & Mellier, Y. 2003, A&A, 409, 835
Giavalisco, M. 2002, ARA&A, 40, 579
Giavalisco, M., & Dickinson, M. 2001, ApJ, 550, 177
Giavalisco, M., Steidel, C. C., Adelberger, K. L., Dickinson, M. E., Pettini, M., & Kellogg, M. 1998, ApJ, 503, 543
Hatton, S., Devriendt, J. E. G., Ninin, S., Bouchet, F. R., Guiderdoni, B., & Vibert, D. 2003, MNRAS, 343, 75
Hawkins, E., Maddox, S., Branchini, E., & Saunders, W. 2001, MNRAS, 325, 589
Heavens, A., Panter, B., Jimenez, R., & Dunlop, J. 2004, Nature, 428, 625
Heinis, S., Treyer, M., Arnouts, S., Milliard, B., Donas, J., Gal, R., Martin, D. C., & Viton, M. 2004, A&A, 424, L9
Heinis, S., et al. 2007, ApJS, 173, 503 (Paper II)
Hopkins, A. M., & Beacom, J. F. 2006, ApJ, 651, 142
Jimenez, R., Panter, B., Heavens, A. F., & Verde, L. 2005, MNRAS, 356, 495
Kauffmann, G., White, S. D. M., Heckman, T. M., Ménard, B., Brinchmann, J., Charlot, S., Tremonti, C., & Brinkmann, J. 2004, MNRAS, 353, 713
Kennicutt, R. C., Jr. 1998, ARA&A, 36, 189
Khochfar, S., & Ostriker, J. P. 2007, preprint (astro-ph/0704.2418)
Landy, S. D., & Szalay, A. S. 1993, ApJ, 412, 64 (LS93)
Lilly, S. J., Le Fevre, O., Hammer, F., & Crampton, D. 1996, ApJ, 460, L1
Lupton, R. H., Gunn, J. E., Ivezic, Z., Knapp, G. R., Kent, S., & Yasuda, N. 2001, in ASP Conf. Ser. 238, Astronomical Data Analysis Software and Systems X, ed. F. R. Harnden, Jr., F. A. Primini, & H. E. Payne (San Francisco: ASP), 269
Madgwick, D. S., et al. 2002, MNRAS, 333, 133
———. 2003, MNRAS, 344, 847
Magliocchetti, M., Bagla, J. S., Maddox, S. J., & Lahav, O. 2000, MNRAS, 314, 546
Martin, D. C., et al. 2005, ApJ, 619, L1
Mo, H. J., & White, S. D. M. 2002, MNRAS, 336, 112
Morrissey, P., et al. 2005, ApJ, 619, L7
———. 2007, ApJS, 173, 682
Neistein, E., van den Bosch, F. C., & Dekel, A. 2006, MNRAS, 372, 933
Norberg, P., et al. 2002, MNRAS, 332, 827
Papovich, C., et al. 2006, AJ, 132, 231
Peebles, P. J. E. 1980, The Large-Scale Structure of the Universe (Princeton: Princeton Univ. Press)
Porciani, C., & Giavalisco, M. 2002, ApJ, 565, 24
Roche, N., & Eales, S. A. 1999, MNRAS, 307, 703
Schiminovich, D., et al. 2005, ApJ, 619, L47
Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, ApJ, 500, 525
Shankar, F., Lapi, A., Salucci, P., De Zotti, G., & Danese, L. 2006, ApJ, 643, 14
Shapley, A. E., Steidel, C. C., Pettini, M., & Adelberger, K. L. 2003, ApJ, 588, 65
Steidel, C. C., Adelberger, K. L., Dickinson, M., Giavalisco, M., Pettini, M., & Kellogg, M. 1998, ApJ, 492, 428
Steidel, C. C., Pettini, M., & Hamilton, D. 1995, AJ, 110, 2519
Stoughton, C., et al. 2002, AJ, 123, 485
Sullivan, M., Treyer, M. A., Ellis, R. S., Bridges, T. J., Milliard, B., & Donas, J. 2000, MNRAS, 312, 442
Wilson, G., Cowie, L. L., Barger, A. J., & Burke, D. J. 2002, AJ, 124, 1258
Yoshikawa, K., Taruya, A., Jing, Y. P., & Suto, Y. 2001, ApJ, 558, 520
Zehavi, I., et al. 2002, ApJ, 571, 172
———. 2005, ApJ, 630, 1