The Astrophysical Journal Supplement Series, 172:396–405, 2007 September

THE *XMM-NEWTON* WIDE-FIELD SURVEY IN THE COSMOS FIELD. V. ANGULAR CLUSTERING OF THE X-RAY POINT SOURCES[1]

Takamitsu Miyaji,[2] Giovanni Zamorani,[3] Nico Cappelluti,[4] Roberto Gilli,[3] Richard E. Griffiths,[2] Andrea Comastri,[3] Günther Hasinger,[4] Marcella Brusa,[4] Fabrizio Fiore,[5] Simonetta Puccetti,[5] Luigi Guzzo,[6] and Alexis Finoguenov[4]

Received 2006 April 24; accepted 2006 June 9

ABSTRACT

We present the first results of the measurements of angular auto-correlation functions (ACFs) of X-ray point sources detected in the *XMM-Newton* observations of the $\sim$2 deg^2 COSMOS field (COSMOS *XMM-Newton*). Significant positive signals have been detected in the 0.5–2 (SFT) band, in the angle range of 0.5′–24′, while the positive signals were at the $\sim$2 and $\sim$3 σ levels in the 2–4.5 (MED) and 4.5–10 (UHD) keV bands, respectively. Correctly taking integral constraints into account is a major limitation in interpreting our results. With power-law fits to the ACFs without the integral constraint term, we find correlation lengths of $\theta_c = 1.9'' \pm 0.3''$, $0.8''^{+0.5''}_{-0.4''}$, and $6'' \pm 2''$ for the SFT, MED, and UHD bands, respectively, for $\gamma = 1.8$. The inferred comoving correlation lengths, also taking into account the bias by the source merging due to *XMM-Newton* PSF, are $r_c \approx 9.8 \pm 0.7$, $5.8^{+1.4}_{-1.7}$, and $12 \pm 2\ h^{-1}$ Mpc at the effective redshifts of $\bar{z}_{\rm eff} \approx 1.1$, 0.9, and 0.6 for the SFT, MED, and UHD bands, respectively. If we include the integral constraint term in the fitting process, assuming that the power law extends to the scale length of the entire COSMOS *XMM-Newton* field, the correlation lengths become larger by $\sim$20%–90%. Comparing the inferred rms fluctuations of the spatial distribution of AGNs $\sigma_{8,\rm AGN}$ with those of the underlying mass, the bias parameters of the X-ray source clustering at these effective redshifts are in the range $b_{\rm AGN} = 1.5$–4.

Subject headings: galaxies: active — quasars: general — X-rays: galaxies

Online material: color figures

1. INTRODUCTION

Results from recent X-ray surveys have made very significant contributions to understanding formation and evolution of supermassive black holes (SMBHs) at galaxy centers. In particular, studies of X-ray luminosity function and its evolution have been providing the most reliable current estimates of the accretion history to SMBH. One of the most important findings in recent years is that luminous active galactic nuclei (AGNs) arise earlier in the history of the universe than lower luminosity ones (Ueda et al. 2003; Hasinger et al. 2005; Barger et al. 2005; La Franca et al. 2005). This suggests that more massive SMBHs have been formed earlier in the universe, and reside quiescently at the centers of giant elliptical galaxies in the later epochs, while more numerous, less massive SMBHs have been formed and accreted later in the history of the universe.

Clustering properties of AGNs and their evolution with redshift provide yet additional clues to understanding the accretion processes onto the SMBHs. These give clues to environments of AGN activities. In the framework of the cold dark matter (CDM) structure formation scenario, clustering properties or the bias of AGNs over a sufficiently large scale $b_{\rm AGN} = (\delta\rho/\rho)_{\rm AGN}/(\delta\rho/\rho)_{\rm mass}$ may be related to the typical mass of dark matter halos in which they reside (Mo & White 1996; Sheth et al. 2001). At the same time, the mechanisms of triggering the AGN activity, which might be closely related to galaxy interactions and/or merging (Menci et al. 2004; Di Matteo et al. 2005), yield a clustering of AGNs and can therefore be inferred from the clustering analysis.

Since strong X-ray emission is a typical feature of AGN activity, X-ray surveys provide most efficient means of constructing comprehensive complete samples of AGNs without contamination from the light in the stellar population of host galaxies. In particular, surveys in the harder ($E > 2$ keV) X-ray band, such as those available from *XMM-Newton*, are very efficient in finding not only unobscured AGNs, which are relatively easy to select also in the optical bands, but also obscured ones, which are difficult to select with optical selection criteria alone. This is important because most of the accretion ($\sim$80%; Comastri et al. 1995; Gilli et al. 2001; Ueda et al. 2003) occurs in AGNs obscured by gas (in X-ray bands) and dust (in the optical bands). While one approach in investigating the environment of AGNs is to measure AGN overdensities around known clusters of galaxies (Cappi et al. 2001; D'Elia et al. 2004; Cappelluti et al. 2005), a more common and direct measure can be obtained by calculating auto-correlation functions (ACFs) of well-defined samples of X-ray-selected AGNs.

While small-number statistics limits the accuracy of the clustering measurements of X-ray-selected AGNs, there are a number of reports on the detection of the correlation signals. Samples based on the *ROSAT* All-Sky survey have mainly constrained the clustering properties of type 1 local AGNs at $z \lesssim 0.3$. The correlation lengths resulting from the angular (Akylas et al. 2000) and 3D (Mullis et al. 2004) analyses of these samples are 6–7 h^{-1} Mpc.[7] Due to the wide redshift distribution, it is more

[1] Based on observations obtained with *XMM-Newton*, an ESA science mission with instruments and contributions directly funded by ESA Member States and NASA.

[2] Department of Physics, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213.

[3] INAF-Osservatorio Astronomico di Bologna, via Ranzani 1, I-40127 Bologna, Italy.

[4] Max-Planck Institut für extraterrestrische Physik, Postfach 1312, 85741 Garching, Germany.

[5] INAF-Osservatorio Astronomico di Roma, via Frascati 33, I-00100 Monteporzio, Italy.

[6] INAF-Osservatorio Astronomico di Brera, via Bianchi 46, I-23807 Merate (LC), Italy.

[7] Throughout this paper we adopt $(H_0,\ \Omega_M,\ \Omega_\Lambda) = (100\ h$ km s^{-1} Mpc^{-1}, 0.3, 0.7).

TABLE 1
X-Ray Sources and Sensitivity Limits

Band (keV)	Number	$CR_{lim,min}$–$CR_{lim,max}$ (counts s^{-1})	$S_{X,lim}$ Range (ergs s^{-1} cm^{-2})[a]	Area (deg^2)
SFT	1037	7.0×10^{-4}–2.2×10^{-2}	6.7×10^{-16}–2.1×10^{-14}	1.43
MED	545	7.0×10^{-4}–2.5×10^{-2}	4.6×10^{-15}–1.6×10^{-13}	1.56
UHD	151	9.0×10^{-4}–2.2×10^{-2}	8.7×10^{-15}–1.8×10^{-13}	1.25

[a] Flux range corresponding to the CR limits. The conversions have been made, following C07, to fluxes in 0.5–2.0, 2.0–10, and 5.0–10 keV assuming power-law spectra of photon indices $\Gamma = 2.0$, 1.7, and 1.7 for the SFT, MED, and UHD bands, respectively.

difficult to obtain clustering signal in deeper X-ray surveys before redshifts for a complete set of X-ray sources are obtained. Nevertheless, Basilakos et al. (2004, 2005) measured strong angular correlation signals in their *XMM-Newton*/2dF survey, which covers a total area of 2 deg^2 over two fields. They obtained a correlation length of $\sim 7.5\ h^{-1}$ Mpc in physical units for both soft and hard X-ray-selected sources, suggesting a clustering evolution that is fixed in the proper coordinate between $z \sim 0$ and $z \gtrsim 1$. At much fainter X-ray fluxes, Gilli et al. (2005) analyzed the projected-distance correlation function $w(r_p)$ for the X-ray sources with spectroscopic redshifts in the Chandra Deep Fields North (CDF-N) and South (CDF-S). They found significantly different clustering properties in these two fields, suggesting a cosmic variance effect due to redshift spike features (Gilli et al. 2003). Recently, Yang et al. (2006) made detailed analysis on their 0.4 deg^2 *Chandra* Large Area Synoptic X-Ray Survey (CLASXS) supplemented by CDF-N. With spectroscopic redshifts for a good portion of the sources, they explored the clustering properties in different redshift and luminosity bins as well as intrinsic absorption bins. They found the evolution of bias with redshift but they did not find significant dependence in the clustering properties of X-ray-selected AGNs based on either luminosity or intrinsic absorption.

One of the main aims of the COSMOS (Scoville et al. 2007) project is to trace the evolution of the large-scale structure of the universe with an unprecedented accuracy and redshift baseline. The *XMM-Newton* Survey, covering the entire COSMOS field (COSMOS *XMM-Newton*; Hasinger et al. 2007), is one of the most extensive *XMM-Newton* survey programs conducted so far. In the first-year *XMM-Newton* observations, about 1400 X-ray point sources have been detected and cataloged (Cappelluti et al. 2007, hereafter C07), which are dominated by AGNs at redshifts $0.7 < z < 2$ (Brusa et al. 2007; Trump et al. 2007).

In this paper, we report the first results of our investigations on the large-scale structure through an angular ACF analysis of the X-ray point sources detected in COSMOS *XMM-Newton*, as a preview of more detailed studies in the near future. Our future studies include the derivation of the direct three-dimensional (3D) correlation function using redshift information already available for a large portion of the X-ray sources and the analysis of the cross-correlation of the X-ray sources with galaxies in the multiwavelength COSMOS catalog.

The outline of the paper is as follows. In § 2, we explain the selection of our samples of X-ray sources to be used in the correlation function analysis, which are subsets of those described in C07. Details of the calculations, including the ACF estimator, the random sample, and power-law fits, are presented in § 3. The deprojection to the 3D correlation function is presented in § 4. The results are discussed in § 5. We summarize our conclusions in § 6

2. SAMPLE SELECTION

Our samples consist of the X-ray sources detected in the first-year *XMM-Newton* observations of the COSMOS field. The source detection, construction of the sensitivity maps, and source counts are described in C07. The X-ray source catalogs in three energy bands, corresponding to energy channels of 0.5–2 (SFT), 2–4.5 (MED), and 4.5–10 (UHD) keV are used.

For the angular ACF studies, we have applied further selection criteria to the C07 sources to minimize the effects of possible systematic errors in the sensitivity maps. The applied criteria for this kind of analysis should be stricter than those adopted for the derivation of the $\log N$-$\log S$ function, because localized systematic errors may cause spurious clustering of X-ray sources. In order to do this, we have scaled up the original sensitivity map to

$$CR_{lim,acf} = \max\left(aCR_{lim,C07} - b,\ CR_{lim,min}\right), \qquad (1)$$

where $CR_{lim,C07}$ is the limiting count rate (in counts s^{-1}) in the original C07 sensitivity map and $CR_{lim,acf}$ is the sensitivity map used for the ACF analysis. After a number of trials, the scaling coefficients (a, b) have been set to (1.33, 1×10^{-4}), (1.40, 0.0), and (1.44, 4×10^{-4}) for the SFT, MED, and UHD bands, respectively. We have excluded the area where $CR_{lim,acf}$ exceeds $CR_{lim,max}$ (low-exposure areas close to the field borders). Those X-ray sources with CRs below $CR_{lim,acf}$ at the source position have been excluded from the ACF analysis. After these screenings, the numbers of sources for the ACF analysis are 1037, 545, and 151 for the SFT, MED, and UHD bands, respectively. While the sensitivity in the soft band is the best among the three bands, some X-ray sources are hard enough that they are detected in only MED and/or UHD bands. These are mainly highly obscured AGNs. Out of the 545 (151) MED (UHD) band sources after the screening process, 59 (13) have not been detected in the SFT band, and only one UHD sources have escaped the detection in the MED band (before the screening process). The numbers of the X-ray sources, values of $CR_{lim,min}$ and $CR_{lim,max}$ and the total areas used for the ACF analysis are summarized in Table 1.

3. ANGULAR CORRELATION FUNCTION CALCULATION

3.1. *The ACF Calculation*

In calculating the binned ACF, we have used the standard estimator by Landy & Szalay (1993):

$$w_{est}(\theta_i) = (DD - 2DR + RR)/RR, \qquad (2)$$

where DD, DR, and RR are the normalized numbers of pairs in the ith angular bin for the data-data, data-random, and random-random samples, respectively. In addition, we use the symbols D and R to represent the data and random samples, respectively.

Expressing the actual numbers of pairs in these three combinations as $n_{\mathrm{pair},DD}(\theta_i)$, $n_{\mathrm{pair},DR}(\theta_i)$ and $n_{\mathrm{pair},RR}(\theta_i)$, the normalized pairs are expressed by

$$DD = n_{\mathrm{pair},DD}(\theta_i)/[N_D(N_D-1)],$$

$$DR = \frac{1}{2} n_{\mathrm{pair},DR}(\theta_i)/(N_D N_R),$$

$$RR = n_{\mathrm{pair},RR}(\theta_i)/[N_R(N_R-1)], \quad (3)$$

where N_D and N_R are the numbers of sources in the data and random samples, respectively. The number of objects in the random sample has been set to 20 times of that in the data sample. This makes the variance of the second and third terms of equation (2) negligible in the error budget of w_{est}.

Our *XMM-Newton* observations have varying sensitivity over the field. In order to create a random sample, which takes the inhomogeneity of the sensitivity over the field into account, we have taken the following steps.

1. Make a random sample composed of N_R objects, where N_R is an integer times N_D.
2. For each random object, assign a count rate from a source from the data sample. The assignments are made in sequence so that the CR distributions of the random and the data sample objects are exactly the same.
3. For each random object, assign a random position in the field. If the sensitivity-map value at this position ($\mathrm{CR}_{\mathrm{lim,acf}}$ from eq. [1], in units of counts s^{-1}) is larger than the assigned CR, find a different position. Repeat this until the position is sensitive enough to detect a point source with the assigned CR.

As a check on this procedure, we have also calculated ACFs using two other methods of generating random samples. The second method is to assign the count rates to the random sources drawn from a $\log N$-$\log S$ relation (e.g., Moretti et al. 2003), instead of copying the count rates of the data sample. Then the source is placed at a random position in the field. If the sensitivity limit at this position is higher than the assigned CR, this random source is rejected. Another more sophisticated and computationally demanding method is to generate random sources based on the $\log N$-$\log S$ relation, down to a flux level much lower than the sensitivity limit of our observations. These sources are then fed into a simulator, taking into account *XMM-Newton* instrumental effects, including position-dependent PSF, exposure maps, and particle background. The entire first-year XMM-COSMOS image has been simulated, and the same source detection procedure as that applied to the actual data has been applied on the simulated data. A random sample R is generated from 10 simulated COSMOS *XMM-Newton* fields. Using the above two methods for random sample generations did not alter the results significantly. In the following analysis, we show the results obtained by the first method.

3.2. *Error Estimation and Covariance Matrix*

We have estimated the errors using the variance of the calculated ACF by replacing D in equation (2) by random samples. A random sample with the same number of objects and the same set of count rates as D has been drawn independently from R in equation (2). We denote the random sample as a replacement of D during the error search by R' to distinguish from R. For each angular bin, a 1 σ error has been calculated as a standard deviation from resulting ACFs calculated from $N_{\mathrm{run}} = 80$ different R' samples, which is then multiplied by a scaling factor $[1 + w(\theta_i)]^{1/2}$ (hereafter referred to as "the scaled random errors"). This scaling factor corrects for the difference between the errors in the null-hypothesis case, obtained from R', and those in the presence of the correlation signal. This is in line with the observation that the error in each bin of the ACF calculated using the Landy & Szalay (1993) estimator is approximately a Poisson fluctuation of the number of data-data pairs in the bin; i.e., $\sigma \sim [1 + w(\theta)]/(n_{\mathrm{pair},DD})^{1/2}$. The standard deviation of the null-hypothesis ACF, obtained by replacing D with an R', is $\sigma_{\mathrm{ran}} \sim 1/(n_{\mathrm{pair},R'R'})^{1/2}$. The scaling factor can be obtained by using the relation, $n_{\mathrm{pair},DD} \approx [1 + w(\theta)] n_{\mathrm{pair},R'R'}$.

In correlation functions, the errors in different angular bins are not independent from one another, and correlations among the errors have to be taken into account when we make function fits. Thus, we have also estimated full covariance matrix in order to represent the correlations among errors by

$$M_{\mathrm{cov},ij} = \sum_k \left[w_{\mathrm{R},k}(\theta_i) - \langle w_R(\theta_i)\rangle\right] \times \left[w_{\mathrm{R},k}(\theta_j) - \langle w_R(\theta_j)\rangle\right]/N_{\mathrm{run}} \times [1 + w_{\mathrm{est}}(\theta_i)]^{1/2}\left[1 + w_{\mathrm{est}}(\theta_j)\right]^{1/2}, \quad (4)$$

where $w_{\mathrm{R},k}(\theta_i)$ is the ACF value for the kth random run (k runs through 1 to $N_{\mathrm{run}} = 80$), $\langle w_R(\theta_i)\rangle$ is their mean value, evaluated at the center of the i-th angular bin θ_i and $w_{\mathrm{est}}(\theta_i)$ is from equation (2). The square roots of the diagonal elements of equation (4) are the scaled random errors discussed above. The covariance matrix calculated in equation (4) is used later in § 3.4. Strictly speaking, equation (4) only takes into account the correlations of errors for the random cases, but not the correlation of errors due to clustering. One way of explaining this is that removing/adding one source (by a Poisson chance) affects multiple angular bins and this is represented by nondiagonal elements of equation (4). On the other hand, correlation of errors due to large-scale structure or the cosmic variance is not represented by this. If we observe another part of the sky, we sample different sets of large-scale structures, such as filaments and voids. Since the existence or nonexistence of one such structure affects multiple angular bins, there should be additional contribution to the nondiagonal elements of the correlations among errors in different angular bins. Equation (4), based on many random samples, thus includes the former type of the correlation of errors but not the latter. One way to include also the latter effect is to use the Jackknife resampling technique, as was done by, e.g., Zehavi et al. (2004). However, the Jackknife resampling requires to divide the sample into many statistically independent regions, which is not practically possible in our case.

3.3. *The Binned ACF Results*

The ACFs have been calculated for the three bands in logarithmically equally spaced bins with $\Delta \log \theta = 1/6$. The results are shown in Figure 1, where the top panels, composed of two layers of logarithmic plots with positive and negative parts ($\log |w(\theta)| > -2.8$ respectively), are attached together. The lower panels show fit residuals for the best-fit functions described in § 3.4. Changing the bin size did not change the clustering amplitude significantly.

The ACFs are presented with the scaled random errors. Positive signals have been detected down to $\theta \sim 0.5'$ in the 0.5–2 keV band and $\theta \sim 1'$ in the other bands. At the smallest scales, correlation signal is negative, probably due to confusion effects, where two sources separated by a distance comparable to or closer than

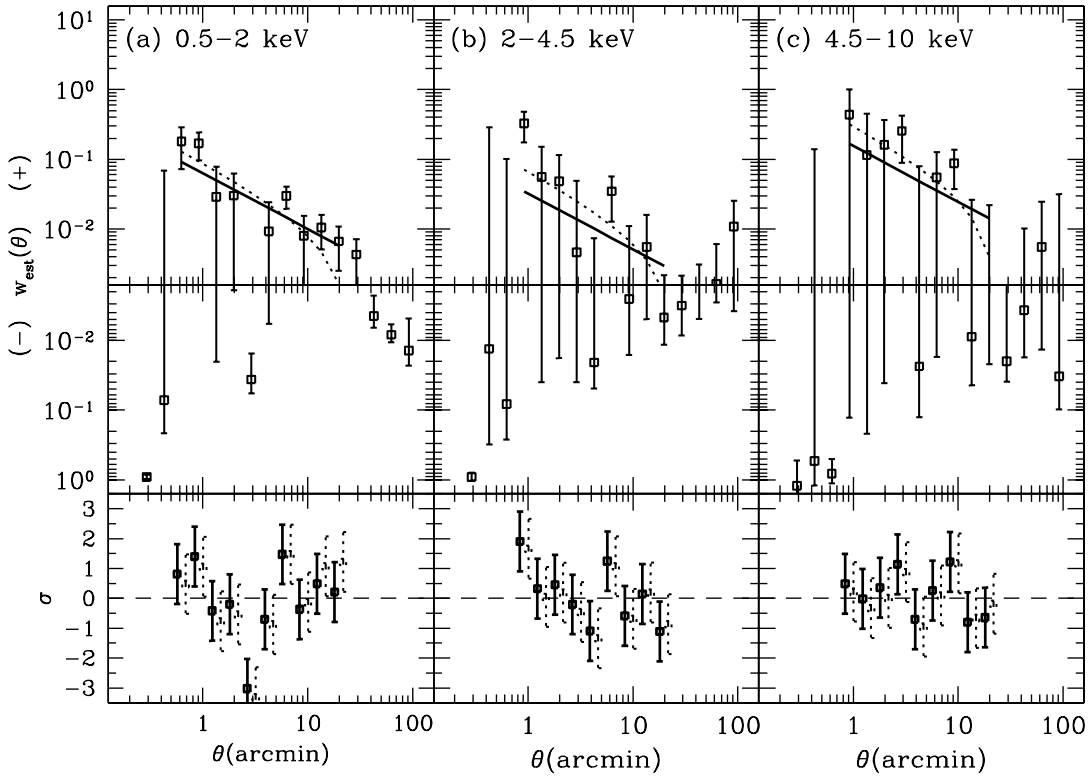

FIG. 1.—Binned estimated angular correlation functions $w_{\rm est}(\theta)$ plotted for the X-ray sources detected in the first-year *XMM-Newton* data in three standard energy bands as labeled. The vertical scale is logarithmic, where the positive and negative parts ($\log |w_{\rm est}(\theta)| > -2.8$) are attached together. The 1 σ errors are the diagonal components of eq. (4), i.e., the scaled random errors. The solid and dotted lines show the best-fit power-law models for $\gamma - 1 = 0.8$ without and with an integral constraint, respectively. The models are plotted in the range where the fits are made. Fit residuals in terms of σ has been also plotted in the lower panels for the two models (slightly offset for visibility) in the same line styles as the models. [*See the electronic edition of the Supplement for a color version of this figure*.]

the point spread function (PSF) cannot be detected separately in the source detection procedure and may well be classified as one extended source. In our current sample, the sources that have been classified as extended have been removed from the sample. The effect of this is discussed in detail in § 3.5. In the 0.5–2 and 2–4.5 keV bands, positive signals extend out to $\sim 20'$. Negative signals are seen at the largest angular scales, probably due to the integral constraint as discussed below.

3.4. *Power-Law Fits*

In order to make a simple characterization of our ACF results, we have fitted the ACF with a power-law model of the form

$$w_{\rm mdl}(\theta) = A\theta^{1-\gamma},$$

where γ is the slope index of the corresponding 3D correlation function. We use the normalization A as a fitting parameter rather than the angular correlation length $\theta_c = A^{1/(\gamma-1)}$, since this gives much better convergence of the fit.

The fits are made by minimizing χ^2_c. The subscript c denotes that the correlations between errors have been taken into account through the inverse of the covariance matrix:

$$\chi^2_c = \mathbf{\Delta}^T M^{-1}_{\rm cov} \mathbf{\Delta}, \tag{5}$$

where $\mathbf{\Delta}$ is a vector composed of $w_{\rm est}(\theta_i) - w_{\rm mdl}(\theta_i) + C$, $M_{\rm cov}$ is the covariance matrix calculated in equation (4) with $N_{\rm run} = 80$, and C is a constant to compensate for the integral constraint as discussed below.

Due to the finite area and the construction of $w_{\rm est}$ in equation (2) the estimated angular correlation function satisfies the integral constraint (e.g., Basilakos et al. 2005; Roche & Eales 1999):

$$\int\int w_{\rm est}\, d^2\Omega = 0. \tag{6}$$

This constraint usually results in $w_{\rm est}$ underestimating the true underlying angular correlation function by the constant C. Under an assumption that the true underlying $w(\theta)$ is a power law and is extended to the scale of the survey area, one can include C in the fitting process, where C can be uniquely determined by θ_c and γ by imposing the integral constraint (Roche & Eales 1999),

$$C = \sum_i A\theta_i^{1-\gamma} RR(\theta_i) / \sum_i RR(\theta_i), \tag{7}$$

where the sums are over angular bins and $RR(\theta_i)$ is the number of random-random pairs in the ith angular bin. The above assumption is not necessarily true. If residual systematic errors in the sensitivity maps are the main cause of the negative values at large angular separations, the determination of C shown above is not valid. However, this gives an approximate estimate of the degree of the underestimation by the integral constraint. This sets an limitation to the our angular ACF analysis, where the estimated C values are not negligible compared with the amplitude of the ACF signal. We have made fits with or without including the integral constraint.

Because of the limited signal-to-noise ratio, we were not able to constrain A and γ simultaneously. Thus, we have calculated the best-fit values and 1 σ confidence errors for the amplitude

TABLE 2
Results of the Power-Law Fits

Fit ID	Band (keV)	A[a] (arcmin$^{1/(\gamma-1)}$)	$\gamma-1$	θ_c[a] (arcsec)	C	$\theta_{\rm min}$ (arcmin)	$\theta_{\rm max}$ (arcmin)
S1	SFT	0.063 ± 0.008	0.8	1.9 ± 0.3	0	0.5	24
S2	SFT	0.093 ± 0.012	0.8	3.1 ± 0.5	8×10^{-3}	0.5	24
S3	SFT	0.034 ± 0.004	0.5	0.07 ± 0.02	0	0.5	24
S4	SFT	0.078 ± 0.010	0.5	0.37 ± 0.09	1.5×10^{-2}	0.5	24
S5	SFT	0.059 ± 0.009	0.8	1.7 ± 0.3	0	1.6	24
S6	SFT	0.089 ± 0.015	0.8	2.9 ± 0.6	7×10^{-3}	1.6	24
M1	MED	0.032 ± 0.015	0.8	$0.8^{+0.5}_{-0.4}$	0	0.7	24
M2	MED	0.071 ± 0.027	0.8	2.2 ± 1.0	5×10^{-3}	0.7	24
M3	MED	0.013 ± 0.008	0.5	$0.010^{+0.016}_{-0.009}$	0	0.7	24
M4	MED	0.048 ± 0.020	0.5	$0.14^{+0.14}_{-0.09}$	9×10^{-3}	0.7	24
M5	MED	0.021 ± 0.016	0.8	$0.47^{+0.49}_{-0.40}$	0	1.6	24
M6	MED	0.044 ± 0.029	0.8	$1.2^{+1.1}_{-0.9}$	3×10^{-3}	1.6	24
U1	UHD	0.15 ± 0.05	0.8	5.6 ± 2.3	0	0.5	24
U2	UHD	0.32 ± 0.09	0.8	14 ± 5	3×10^{-2}	0.7	24
U3	UHD	0.075 ± 0.024	0.5	$0.34^{+0.25}_{-0.18}$	0	0.7	24
U4	UHD	0.23 ± 0.07	0.5	$3.2^{+2.2}_{-1.6}$	5×10^{-2}	0.7	24
U5	UHD	0.080 ± 0.032	0.8	$2.5^{+1.3}_{-1.2}$	0	2.4	35
U6	UHD	0.17 ± 0.06	0.8	$6.5^{+3.0}_{-2.7}$	3×10^{-2}	2.4	35

[a] One σ errors are shown. The effects of PSF merging (§ 3.5) have not been taken into account.

for two fixed values of $\gamma - 1 = 0.8$ and 0.5. The former value is for the canonical value for local galaxies (e.g., Peebles 1980; Zehavi et al. 2004), the latter is approximately the slope found for X-ray sources in the Chandra Deep Fields (Gilli et al. 2005). The angular fit results are summarized in Table 2. In this table, fits with different bands and parameters are identified with a Fit ID. The angle range for the fits are $\theta_{\rm min} < \theta < \theta_{\rm max}$ and the boundaries are also shown in Table 2. For fit IDs S1–S4, M1–M4, and U1–U4, we have set $\theta_{\rm min} = 0.5'$ (SFT) or $0.7'$ (MED, UHD), which is the minimum at which ACF is still positive, and below which the ACF goes negative due to the *XMM-Newton* PSF. Likewise, we set $\theta_{\rm max} = 24'$, which is about the maximum scale where the ACF is still positive. The best-fit models for $\gamma - 1 = 0.8$ are overplotted in Figure 1 in the bin ranges included in the fits.

As another choice, we have set $\theta_{\rm min}$ and $\theta_{\rm max}$ in such a way that the range approximately corresponds to the projected comoving distance range of 1–16 h^{-1} Mpc (Fit IDs S5, S6, M5, M6, U5, and U6) at the effective median redshift of the sample ($\bar{z}_{\rm eff}$ discussed below in § 4). The rationale for the maximum scale is that, in our subsequent analysis, the correlation functions are converted to the root mean square (rms) density fluctuation with an 8 h^{-1} Mpc radius sphere (therefore the relevant maximum separation is 16 h^{-1}) in discussing bias parameters. The rationale for the minimum scale is to minimize the effects of nonlinearity in discussing the bias parameters and typical halo masses.

3.5. *Effects of Source Merging Due to PSF*

The "amplification bias," due to which the estimated ACF from sources detected in a smoothed image (e.g., by a finite PSF) is amplified with respect to the true underlying ACF, has been first noted and discussed by Vikhlinin & Forman (1995) in the context of the clustering of X-ray sources. This is caused by merging of multiple sources that are separated by distances comparable to or closer than the PSF. The effect of this bias depends on the true underlying angular correlation function and the number density of the sources. In principle, full simulations involving PSF smoothing and the source detection process are required to estimate the amount of this bias. Basilakos et al. (2005) took a simplified approach in estimating this effect on their ACF from their *XMM-Newton*/2dF survey. In order to estimate the size of the effect, they used particles sampled from a cosmological simulation. They simulated the *XMM-Newton* sources by merging all the particle pairs closer than $6''$. They then compared the angular ACFs from the particles themselves and the simulated *XMM-Newton* sources. As a result, they estimated that the measured angular correlation length is overestimated by 3%–4% due to the amplification bias.

In our case with COSMOS *XMM-Newton*, we have explicitly excluded sources that are classified as extended by the source detection procedure (C07). This causes most of the source pairs closer than $\sim 20''$ to disappear from the sample, since these pairs are classified as single extended sources. Because the exclusion of these sources can suppress the estimated angular correlation function, we use the term "PSF merging bias" rather than the amplification bias. Pairs of sources that are closer than $\sim 4''$ are, however, detected as a single point source. We have applied a similar approach to Basilakos et al. (2005) in estimating the effects of the PSF merging bias. We have sampled particles from the COSMOS-Mock catalog extracted from the Millenium simulation (Kitzbichler & White 2007) over ~ 2 deg^2 of the sky. Redshift, cosmological intrinsic redshift, and magnitudes in various photometric bands are provided for each mock galaxy in the catalog. We use the mock catalog to estimate the effect of the PSF merging bias in the "angular correlation function." Thus, the selected objects from the mock catalog for our simulation do not have to physically represent to the actual X-ray-selected AGNs. For our present purpose, we have chosen the mock galaxies in a redshift interval (roughly in the range $0.4 \lesssim z \lesssim 0.8$) and a magnitude range in such a way that the amplitude of the resulting angular ACF and the source number densities roughly match those of the X-ray samples. We have then created a simulated COSMOS *XMM-Newton* catalog as follows: (1) source pairs with separations smaller than $4''$ are merged into single sources and (2) pairs that are between $4''$ and $20''$ from each other are eliminated. We repeated this experiment 19 times and compared the mean angular ACFs from the original particles and that from simulated COSMOS *XMM-Newton* by making power-law fits to the mean ACFs. As a result we found that the ACF amplitudes

measured using the sources in our source detection procedure on the COSMOS *XMM-Newton* data are underestimated by 15% and 8% for the SFT and MED bands, respectively. Corrections for this effect have not been applied for the values in Table 2, but are considered in further discussions. The effects is negligible in the UHD band, due to the relatively low number density of the sources detected in this band, which made the average distance among neighboring sources much larger than the PSF.

4. IMPLICATION FOR THREE-DIMENSIONAL CORRELATION FUNCTION AND BIAS

4.1. *Deprojection to Real Space Correlation Function*

The 2D ACF is a projection of the real-space 3D ACF of the sources $\xi(r)$ along the line of sight. In the following discussions and thereafter, r is in comoving coordinates. The relation between the 2D (angular) ACF and the 3D ACF is expressed by the Limber equation (e.g., Peebles 1980). Under the usual assumption that the scale length of the clustering is much smaller than the distance to the object, this reduces to

$$w(\theta)N^2 = \int \left(\frac{dN}{dz}\right)^2 \int \xi\left(\sqrt{[d_A(z)\theta]^2 + l^2}(1+z)\right)\left(\frac{dl}{dz}\right)^{-1} dl\, dz, \quad (8)$$

where $d_A(z)$ is the angular distance, N is the total number of sources, and dN/dz is the redshift distribution (per z) of the sources. The redshift evolution of the 3D correlation function is customarily expressed by

$$\xi(r,\ z) = (r/r_{c,0})^{-\gamma}(1+z)^{-3-\epsilon+\gamma}, \quad (9)$$

where $\epsilon = -3$ and $\epsilon = \gamma - 3$ correspond to the case where the correlation length is constant in physical and comoving coordinates, respectively. In these notations, the zero-redshift 3D correlation length $r_{c,0}$ can be related to the angular correlation length θ_c by

$$r_{c,0}^{\gamma} = \left(N^2/S\right)\theta_c^{\gamma-1},$$

$$S = H_\gamma \int \left(\frac{dN}{dz}\right)^2 \left[\frac{c\, d\tau(z)}{dz}\right]^{-1} d_A^{1-\gamma}(1+z)^{-3-\epsilon}\, dz,$$

$$H_\gamma = \frac{\Gamma[(\gamma-1)/2]\Gamma(1/2)}{\Gamma(\gamma/2)}, \quad (10)$$

where $\tau(z)$ is the look-back time. Note that all dependence on cosmological parameters are included in $d_A(z)$ and $\tau(z)$. We also define the comoving correlation length

$$r_c(\bar{z}_{\rm eff}) = r_{c,0}(1+\bar{z}_{\rm eff})^{(-3-\epsilon+\gamma)/\gamma} \quad (11)$$

at the effective redshift $\bar{z}_{\rm eff}$, which is the median redshift of the contribution to the angular correlation (the integrand of the second term of eq. [10]).

An essential ingredient of the deprojection process is the redshift distribution of the sources. At this stage, we do not yet have individual redshifts of a comprehensive set of the *XMM-Newton* complete sample. Thus, we use expected distributions from the X-ray luminosity functions and AGN population synthesis models. We use the model by Ueda et al. (2003; luminosity-dependent density evolution [LDDE]) for all bands. In calculating the redshift distribution, we have used the sensitivity map in units of CR and the actual *XMM-Newton* response function in each band. We also use the Hasinger et al. (2005) type 1 AGN soft X-ray luminosity function (SXLF) for the 0.5–2 keV for comparison. The redshift distributions of the X-ray sources predicted by these models are plotted in Figure 2. Both Ueda et al. (2003) and Hasinger et al. (2005) used samples with a very high identification completeness with redshifts measurements (>90%), at least down to the flux limits sampled by the COSMOS *XMM-Newton* survey. Thus, the effect on the expected redshift distribution due to the identification incompleteness is negligible. In calculating the 3D correlation functions, we use the fits with and without integral constraints. The angular correlation amplitude A has been multiplied by a correction factor due to PSF merging as discussed in § 3.5. We also use the fits with $\gamma = 1.8$. The calculated $r_{c,0}$ and the comoving correlation length at the effective median redshift $r_c(\bar{z}_{\rm eff})$ are listed in Table 3 for selected results. The errors on $r_{c,0}$ and $r_c(\bar{z}_{\rm eff})$ have been calculated for fixed γ and ϵ.

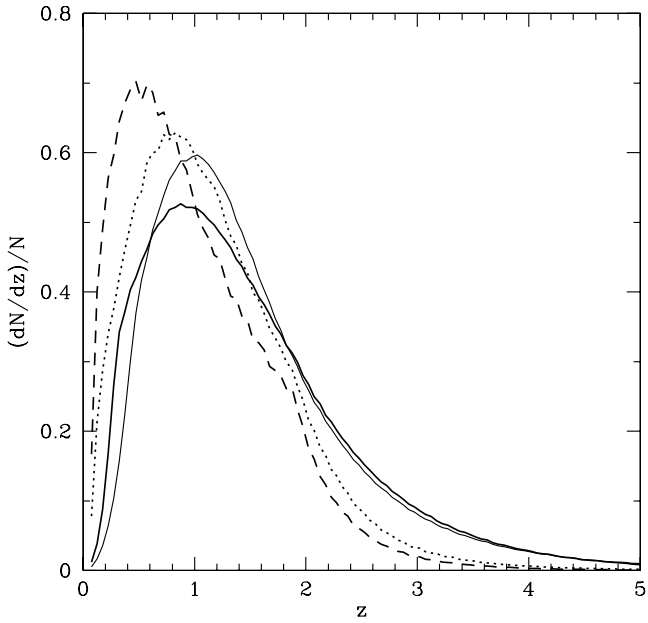

FIG. 2.—Model redshift distributions of the COSMOS-XMM sources in the 0.5–2 keV (*solid lines*), 2–4.5 keV (*dotted line*), and 4.5–10 keV (*dashed line*) ranges. The thicker and thinner solid lines correspond to the 0.5–2 keV band redshift distributions based on the Ueda et al. (2003) model and the Hasinger et al. (2005) soft X-ray luminosity function, respectively. [*See the electronic edition of the Supplement for a color version of this figure.*]

4.2. *Bias and Comparison with Other Works*

In order to estimate the bias parameter of the X-ray sources with respect to the underlying mass distribution, we calculate the rms fluctuation of the distribution of the X-ray sources in the sphere with a comoving radius of $r_{\rm max} = 8\ h^{-1}$ for the power-law model (e.g., § 59 of Peebles 1980),

$$\sigma_{8,\rm AGN}^2 = \int\int \xi(|\boldsymbol{r}_1 - \boldsymbol{r}_2|) dV_1 dV_2 / V^2 = (r_{\rm max}/r_c)^{-\gamma} J_2 \quad (12)$$

$$J_2 = 72/[(3-\gamma)(4-\gamma)(6-\gamma)2^{\gamma}]. \quad (13)$$

As discussed above, we have used the results from Fit IDs S5, S6, M5, M6, U5, and U6, where the fits are made to the angle range corresponding to $\approx$1–16 h^{-1} Mpc at $\bar{z}_{\rm eff}$. The corresponding quantity of the underlying mass distribution at $z = 0$, and σ_8 is one of the commonly used parameters in cosmology

TABLE 3
THREE-DIMENSIONAL CORRELATION LENGTHS

Fit ID	γ	ϵ	$\bar{z}_{\rm eff}$	$r_{c,0}$[a] (h^{-1} Mpc)	$r_c(\bar{z}_{\rm eff})$[a] (h^{-1} Mpc)	Model[b]
S1	1.8	−1.2	1.07	9.8 ± 0.7	9.8 ± 0.7	U03
S1	1.8	−1.2	1.11	9.4 ± 0.7	9.4 ± 0.7	H05
S1	1.8	−3.0	1.42	4.3 ± 0.3	10.4 ± 0.7	U03
S2	1.8	−1.2	1.07	12.1 ± 0.9	12.1 ± 0.9	U03
S5	1.8	−1.2	1.07	9.4 ± 0.8	9.4 ± 0.8	U03
S6	1.8	−1.2	1.07	11.8 ± 1.1	11.8 ± 1.1	U03
M1	1.8	−1.2	0.87	$5.8^{+1.4}_{-1.7}$	$5.8^{+1.4}_{-1.7}$	U03
M1	1.8	−3.0	1.13	$2.9^{+0.7}_{-0.8}$	$6.1^{+1.5}_{-1.8}$	U03
M2	1.8	−1.2	0.87	$9.0^{+1.8}_{-2.1}$	$9.0^{+1.8}_{-2.1}$	U03
M5	1.8	−1.2	0.87	$4.6^{+1.7}_{-2.5}$	$4.6^{+1.7}_{-2.5}$	U03
M6	1.8	−1.2	0.87	$6.9^{+2.2}_{-3.1}$	$6.9^{+2.2}_{-3.1}$	U03
U1	1.8	−1.2	0.60	$11.9^{+2.1}_{-2.4}$	$11.9^{+2.1}_{-2.4}$	U03
U1	1.8	−3.0	0.88	$6.6^{+1.1}_{-1.2}$	$12.5^{+2.2}_{-2.5}$	U03
U2	1.8	−1.2	0.60	19 ± 3	19 ± 3	U03
U5	1.8	−1.2	0.60	$8.4^{+1.7}_{-2.0}$	$8.4^{+1.7}_{-2.0}$	U03
U6	1.8	−1.2	0.60	$12.7^{+2.3}_{-2.7}$	$12.7^{+2.3}_{-2.7}$	U03

[a] Errors are 1 σ.
[b] Model designations: (U03), Ueda et al. 2003; (H05), Hasinger et al. 2005.

(Spergel et al. 2003). In order to compare our results with other similar works on a common ground, we calculated $\sigma_{8,\rm AGN}$ from power-law representations from literature and plotted them versus the effective redshift of each sample.[8]

For this comparison, we have used the best-fit correlation lengths and slopes (r_c, γ) from literature to estimate $\sigma_{8,\rm AGN}$ values and their 1 σ errors. Since each reference has a different method of presenting results, we take the following strategy in calculating $\sigma_{8,\rm AGN}$ and its 1 σ error.

1. If the referenced article gives confidence contours in the (r_c, γ) space, we calculate $\sigma_{8,\rm AGN}$ values for the nominal case as well as at each point in the $L = L_{\rm min} + 1$ contour, where L (with the best-fit value $L_{\rm min}$) is either χ^2 or a statistical estimator that varies as χ^2, e.g., Cash C estimator for Yang et al. (2006). The error range on $\sigma_{8,\rm AGN}$ is determined by the minimum and maximum values calculated from the points along the contour.
2. If no confidence contour in the (r_c, γ) space is given and there is a fit result with a fixed γ, we use this fit to calculate $\sigma_{8,\rm AGN}$. The 1 σ error in r_c is propagated from that of r_c.
3. If the article gives only best-fit (r_c, γ) values, and 1 σ errors on both parameters, the error of $\sigma_{8,\rm AGN}$ has been propagated from those of r_c and γ, neglecting possible correlation of errors between these two parameters. In this case, we may well have over- or underestimated the errors on $\sigma_{8,\rm AGN}$.

If the (r_c, γ) values are given in multiple evolution models, we use the one for which the correlation length is fixed in the comoving coordinates (i.e., $\epsilon = -3 - \gamma$). At about the effective median redshift of the sample, however, the correlation lengths calculated assuming different values of ϵ do not differ significantly. In the case of our present work, we see this by comparing the $r_c(\bar{z}_{\rm eff})$ values for $\epsilon = -1.2$ and -3 cases in Table 3. In addition, the value of $\sigma_{8,\rm AGN}$ is insensitive to the assumed value of γ. The change of $\sigma_{8,\rm AGN}$ is less than 0.1 between the assumed γ of 1.8 and 1.5 for our results in all the three bands. The results from literature we use for this comparison and the details of the conversion to $\sigma_{8,\rm AGN}$ are described below, roughly in order of redshift.

Grazian et al. (2004) calculated the correlation function of 392 optically selected QSOs from the Asiago-ESO/RASS QSO Survey (AERQS) with $\bar{z}_{\rm eff} = 0.062$. They found the nominal values of $r_c = 8.6 \pm 2.0\ h^{-1}$ with $\gamma = 1.56$. And in the low-redshift end, Akylas et al. (2000) calculated the correlation function of X-ray-selected AGNs from the *ROSAT* All-Sky Survey (RASS) with a median redshift of $\bar{z} = 0.15$. Their correlation length for $\gamma = 1.8$ of $r_c = 6.5 \pm 1.0\ h^{-1}$ Mpc for the Einstein–de Sitter universe is increased by 5% to convert it to our adopted cosmology. Mullis et al. (2004) found $r_c = 7.4^{+1.8}_{-1.9}\ h^{-1}$ Mpc for $\gamma = 1.8$ in their *ROSAT* North Ecliptic Pole Survey (NEPS) AGNs with median redshift for the contribution to the clustering signal of $\bar{z}_{\rm eff} = 0.22$. Basilakos et al. (2004, 2005) in their 2 deg^2 *XMM-Newton* survey, with shallower flux limits than COSMOS *XMM-Newton*, found $r_c = 16.4 \pm 1.3\ h^{-1}$ at $\bar{z} = 1.2$ and $r_c = 19 \pm 3\ h^{-1}$ at $\bar{z} = 0.75$ for the 0.5–2 and 2–8 keV, respectively. A recent work by Puccetti et al. (2006) on the central $\sim$0.6 deg^2 region of the ELAIS-S1 field, covered by four mosaicked *XMM-Newton* exposures with $\sim$50–60 ks each, also measured angular ACFs of X-ray point sources. For fixed $\gamma = 1.8$, they found $r_c = 12.8 \pm 4.2\ h^{-1}$ at $\bar{z} = 1.0$ and $r_c = 17.9 \pm 4.8\ h^{-1}$ at $\bar{z} = 0.85$ for the 0.5–2 and 2–10 keV bands, respectively.

The correlation functions on the deepest X-ray surveys on the Chandra Deep Fields–South (CDF-S; $\bar{z} = 0.84$) and North (CDF-N; $\bar{z} = 0.96$) by Gilli et al. (2005) gave, for fixed $\gamma = 1.4$, $r_c = 10.4 \pm 0.8\ h^{-1}$ and $5.1^{+0.4}_{-0.5}\ h^{-1}$ Mpc, respectively. We use the results from their AGN samples. For all of the above samples, the errors on the $\sigma_{8,\rm AGN}$ have been calculated using method *b*.

An extensive redshift-space correlation function was made by Yang et al. (2006) who made use of the data from a combination of the CLASXS and CDF-N surveys, with a significant portion of the X-ray sources having measured spectroscopic redshifts. We have used their (s_0, γ) confidence contours, where s_0 is the redshift space comoving correlation length, in the four redshift bins with median redshifts of $\bar{z} = 0.45$, 0.92, 1.26, and 2.07 to estimate $\sigma_{8,\rm AGN}$ using method *a*. In this conversion, we have corrected for the redshift distortion by dividing the redshift space $\sigma_{8,\rm AGN}$ value by $(1.3)^{1/2}$ (Marinoni et al. 2005; Yang et al. 2006). Extensive clustering studies of QSOs from the 2dF QSO redshift survey (2QZ) have been made using both the projected-distance correlation function approach (Porciani et al. 2004) and the redshift space 3D correlation function approach (Croom et al. 2005). We have converted the nominal r_0-γ values and confidence contours in three redshift bins at $\bar{z}_{\rm eff} = 1.06$, 1.51, and 1.89 by Porciani et al. (2004) to $\sigma_{8,\rm AGN}$ using method *a*. In converting the Croom et al. (2005) (s_0, γ) results in 10 redshift bins ranging from $z = 0.5$ to 2.5, we have used method *c* and the redshift distortion correction has been made in the same way as we have done to the Yang et al. (2006) results.

Figure 3*a* shows the $\sigma_{8,\rm AGN}$ values as a function of the look-back time $\tau(z)$ for our default cosmology from our analysis results both without and with integral constraints. We also overplot $\sigma_{8,\rm AGN}$ values calculated from the results found in literature as detailed above. In order to compare them with those of the underlying mass distribution, we have also plotted the $\sigma_8 D(z)$ values from the linear theory (e.g., Carroll et al. 1992; Hamilton 2001), normalized to 0.75 at $z = 0$ (Spergel et al. 2003).[9] This curve has been

[8] Some authors give the median redshift of the number distribution of the X-ray sources, while we and some others give median redshift of the contribution to the clustering signal. We denote the former by $\bar{z}$ and the latter by $\bar{z}_{\rm eff}$. We do not make a distinction between these in Fig. 3.

[9] We use the latest value of σ_8 as of writing this paper obtained from http://lambda.gsfc.nasa.gov/product/map/current/parameters.cfm for the ΛCDM model derived from all data sets.

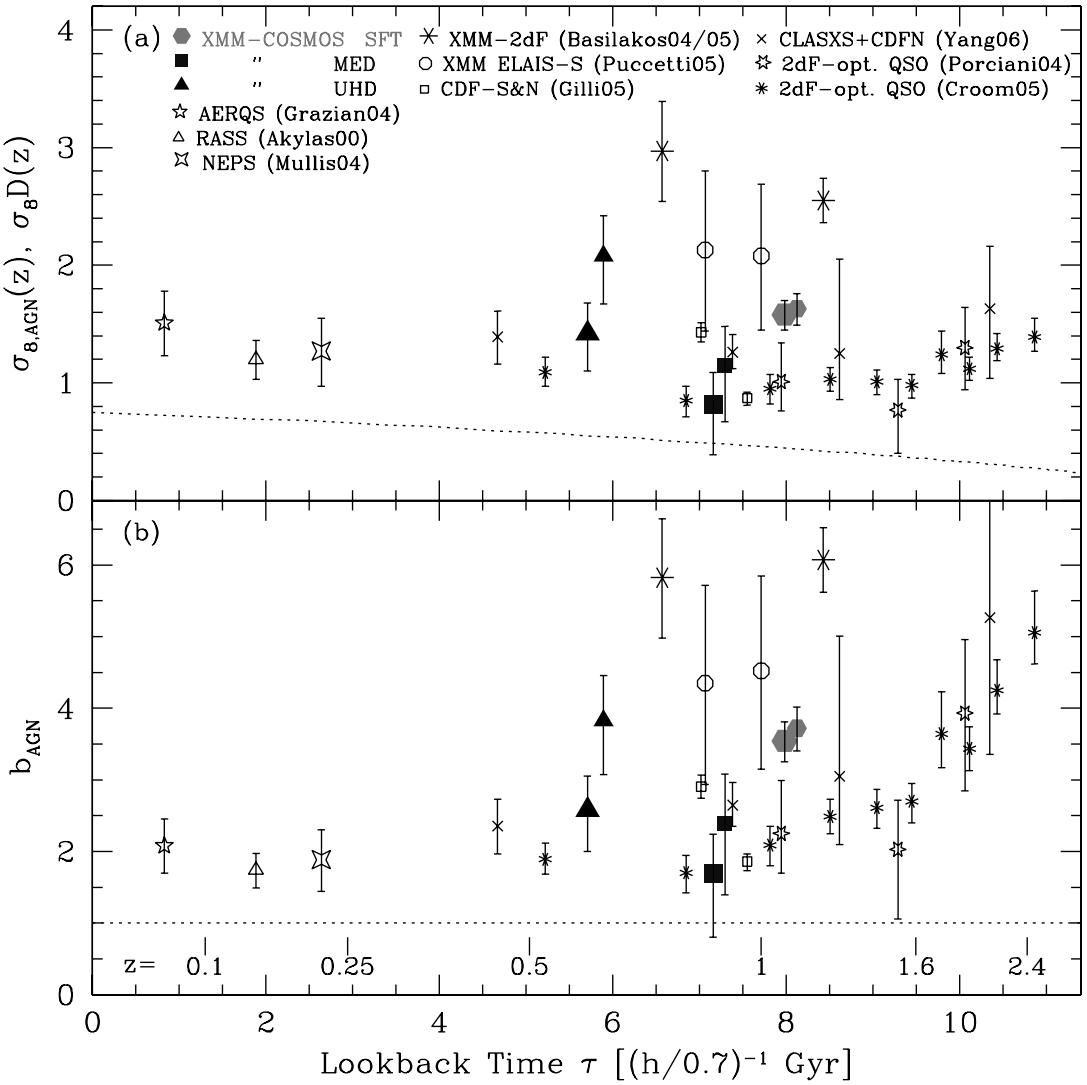

FIG. 3.—(*a*) The $\sigma_{8,\,\mathrm{AGN}}$ of the X-ray sources/AGNs inferred by the power-law fits to the correlation functions from this work and literature, plotted against the look-back time corresponding to the effective median redshift of the samples. All error bars indicate 1 σ errors. The results from this work for the fits without integral constraints are plotted with large solid symbols as labeled. Those with integral constraints are also plotted with smaller symbols and at positions slightly shifted rightward for visibility. The $\sigma_{8,\,\mathrm{AGN}}$ values calculated from various results in literature (see text) are also shown as labeled (abbreviated as the first author followed by the last two digits of the publication year). The dotted line shows $\sigma_8 D(z)$ for the mass in the linear theory normalized to 0.9 at $z = 0$. (*b*) The bias parameter $b_{\mathrm{AGN}} = \sigma_{8,\,\mathrm{AGN}}(z)/[\sigma_8 D(z)]$ is plotted as a function of the effective redshift. The meaning of the symbols are the same as panel (*a*). We also show redshift ticks at the bottom of the figure. [*See the electronic edition of the Supplement for a color version of this figure*.]

shown to accurately represent the distribution of dark matter particles in the ΛCDM Hubble Volume Simulation (Marinoni et al. 2005). Figure 3*b* shows the inferred bias parameters $b_{\mathrm{AGN}} = \sigma_{8,\,\mathrm{AGN}}(z)/[\sigma_8 D(z)]$. The values of $\sigma_{8,\,\mathrm{AGN}}$ and b_{AGN} from this work are shown in Table 4.

There are other recently published angular auto-correlation measurements of X-ray point sources based on large-scale *XMM-Newton* surveys (e.g., Gandhi 2006; Carrera 2007). Comparisons with these will be made in a future publication.

5. DISCUSSION AND PROSPECTS

In this work, we used all the point sources above the scaled sensitivity threshold without further classification of the sources. We analyzed our results assuming that all the X-ray sources are AGNs. This, in practice, is a good approximation. Our preliminary identifications of the sources indicate that out of the 1037, 545, and 151 sources selected for the ACF analysis for the SFT,

TABLE 4
ESTIMATED $\sigma_{8,\,\mathrm{AGN}}$ AND BIAS

Fit ID	$\bar{z}_{\mathrm{eff}}$	$\sigma_{8,\,\mathrm{AGN}}$	b_{AGN}	$\log M_{\mathrm{halo}}$[a] ($h^{-1}\,M_\odot$)
Without Integral Constraint				
S5	1.07	$1.58^{+0.12}_{-0.13}$	3.5 ± 0.3	13.6 ± 0.1
M5	0.87	$0.82^{+0.27}_{-0.42}$	$1.7^{+0.6}_{-0.9}$	<13.3
U5	0.60	$1.42^{+0.26}_{-0.32}$	$2.6^{+0.5}_{-0.6}$	13.5 ± 0.2
With Integral Constraint				
S6	1.07	$1.63^{+0.13}_{-0.14}$	3.7 ± 0.3	13.6 ± 0.1
M6	0.87	$1.19^{+0.34}_{-0.50}$	$2.5^{+0.7}_{-1.0}$	$13.3^{+0.3}_{-0.7}$
U6	0.60	$2.08^{+0.34}_{-0.41}$	$3.8^{+0.6}_{-0.8}$	13.9 ± 0.2

[a] The error on M_{halo} reflects the statistical error on b_{AGN} only.

MED, and UHD bands, 20, 5, and 1, respectively, are apparent Galactic stars. Removing these sources from the analysis changed the results very little. Also our results are not likely to be heavily affected by the contamination of clusters/groups, since these sources are extended by $\gtrsim 20''$ (e.g., Finoguenov et al. 2007) and are likely to be classified as extended by the source detection procedure, and hence removed from our sample.

As seen in Figure 3, with an exception of the MED band, our analysis without integral constraints gives somewhat larger $\sigma_{8,\mathrm{AGN}}$ values than those obtained from results using 2dF optically selected QSOs by Porciani et al. (2004) and Croom et al. (2005) but in general agreement with the values from *Chandra* CLASXS+CDF-N by Yang et al. (2006) CDF-S by Gilli et al. (2005) and *XMM-Newton* results from Puccetti et al. (2006). Most likely due to the cosmic variance over a small field of view (FOV), the Gilli et al. (2005) result on CDF-N gave a significantly smaller correlation amplitude than their own CDF-S values, as well as our results. The angular ACFs from a shallower *XMM-Newton* survey by Basilakos et al. (2004, 2005) gave significantly larger $\sigma_{8,\mathrm{AGN}}$ values than other works in both 0.5–2 and 2–8 keV bands. The reason for their distinctively large value is unclear.

One of the interesting questions in investigating clustering properties of X-ray-selected AGNs is to investigate whether there is any difference in the environments of obscured and unobscured AGNs. Applying the population synthesis model of Ueda et al. (2003) to our sensitivity maps, only ~17% of the sources detected in the SFT band at $z \sim 1$ are obscured AGNs with $N_{\mathrm{H}} > 10^{22}$ cm^{-2}. The fraction increases to ~40% in the MED and UHD bands. A comparison of bias parameters between SFT band and MED band, which have similar $\bar{z}_{\mathrm{eff}}$ values, seems to show a lager bias parameter for the SFT sample. However, with the combination of statistical uncertainties and uncertainties in modeling the integral constraint in the MED band, we can only conclude that the bias of the obscured AGNs is not stronger than that of unobscured AGNs. In other works, Gilli et al. (2005), Yang et al. (2006), and Basilakos et al. (2004, 2005) did not find any statistical difference between the clustering properties of these two. Further studies involving the second-year COSMOS *XMM-Newton* data, which in effect doubles the *XMM-Newton* exposure over the COSMOS field, and redshift information of individual objects will probe into this problem further. Also with the accepted C-COSMOS program totaling 1.8 Ms of *Chandra* exposure, we will be able to probe the correlation functions to a much smaller scale, enabling us to investigate the immediate neighbor environments of these AGNs.

Our measured bias parameters based on the rms fluctuations in the 8 h^{-1} Mpc radius sphere are in the range $b_{\mathrm{AGN}} = 1.5$–4. The clustering properties of dark matter halos (DMH) depend on their mass (Mo & White 1996; Sheth et al. 2001), and we can estimate the typical mass of the DMHs in which the population of AGNs represented by our sample reside, under the assumption that the typical mass halo is the main cause of the AGN biasing. Following the approach of Yang et al. (2006) and Croom et al. (2005) who utilized the model by Sheth et al. (2001) we roughly estimate that the typical mass of DMH is $\sim 10^{13}$–10^{14} $M_{\odot}$ for our SFT and UHD samples (see Table 4). These are an order of magnitude larger than those estimated by Porciani et al. (2004) and Croom et al. (2005), probably reflecting the large bias parameters from our results.

One of the largest uncertainties in our analysis lies in the treatment of the integral constraint, because its effect is not negligible in our case compared with the ACF amplitudes in the range of our interest. Figure 3 shows that our results based on the fits with integral constraint, under an assumption that the fitted power-law behavior of the underlying $w(\theta)$ extends to the scale of the entire FOV, give a somewhat larger correlation amplitudes. This assumption may not be true. In addition, the apparent negative $w_{\mathrm{est}}(\theta)$ values at $\theta > 30'$ in Figure 1 may well be caused by remaining systematic errors. Thus, the interpretation of the angular correlation functions, where the signals are diluted by the projection along the redshift space, has a major limitation in correctly taking the integral constraint into account.

The situation will improve when redshift information on individual X-ray sources becomes available for a major and comprehensive set of the X-ray sources. This will enable us to calculate 3D correlation functions or projected-distance correlation functions in a number of redshift bins. With the line-of-sight dilution effect suppressed, we will be able to obtain a larger amplitude in the correlation signal, making the analysis much less subject to the uncertainties in the integral constraint. With optical follow-up programs underway on the COSMOS field through Magellan and z-COSMOS projects (Trump et al. 2007; Lilly et al. 2006), we are obtaining spectroscopic redshifts from a major fraction of the X-ray sources. At the time of writing this paper, we have been able to define a sample of 378 COSMOS *XMM-Newton* detected AGNs with measured spectroscopic redshifts (~30% of the X-ray point sources), with a median redshift of $z \sim 1$. Our preliminary analysis of these sources based on the projected-distance correlation function $w(r_p)$ gives a comoving correlation length of $r_c \approx 8$ h^{-1} Mpc and $\gamma = 1.6$, which is fully consistent with our results without the integral constraints. The results of a full analysis utilizing the redshift information will be presented in a future paper (Gilli et al. 2007, in preparation).

Extensive multiwavelength coverage and the availability of a galaxy catalog also enables us to investigate the cross-correlation function between X-ray-selected AGNs and galaxies. By cross-correlating the X-ray-selected AGNs with 3 orders of magnitude more galaxies, we will be able to investigate the environments of the AGN activity in various redshifts with much better statistics.

6. CONCLUSIONS

We have presented the first results on the angular correlation functions (ACFs) of the X-ray-selected AGNs from the COSMOS *XMM-Newton* survey and reached the following conclusions.

1. A significant positive angular clustering signals has been detected in the 0.5–2 (SFT) bands in the angle range of 0.5′–20′, while in the 2–4.5 (MED) and 4.5–10 keV (UHD) bands, the positive signals are 2 and 3 σ, respectively. The robustness of the estimated correlation functions has been verified using different methods of generating random samples.

2. Power-law fits to the angular correlation function have been made, taking into account the correlation of errors. Correctly taking the integral constraint into account is a major limitation on interpreting the angular correlation function. For fits with fixed $\gamma - 1 = 0.8$ and without (with) the integral constraint term, we found correlation lengths of $\theta_c = 1\farcs9 \pm 0\farcs3$, $0\farcs8^{+0\farcs5}_{-0\farcs4}$, and $6'' \pm 2''$ ($3\farcs1 \pm 0\farcs5$, $2\farcs2 \pm 1\farcs0$, and $14'' \pm 5''$) for the SFT, MED, and UHD bands, respectively.

3. Due to *XMM-Newton* PSF, most of the source pairs closer than ~20″ are classified as single extended sources, and therefore excluded from the sample. This causes a bias in angular correlation function measurements. We have estimated this effect (the PSF merging bias) by simulations and found that the

estimated ACF underestimates the amplitude of the true underlying ACF by ∼15% and ∼8% for the SFT and MED bands, respectively

4. Using Limber's equation and the expected redshift distributions of the sources, we have found comoving correlation lengths of $r_c \approx 9.8 \pm 0.7$, $5.8^{+1.4}_{-1.7}$, and $12 \pm 2\ h^{-1}$ Mpc for $\gamma = 1.8$ at the effective redshifts of $\bar{z}_{\rm eff} \approx 1.1$, 0.9, and 0.55 for the SFT, MED, and UHD bands, respectively, for the fits without integral constraints, while 20%–90% larger correlation lengths have been obtained for the fits with integral constraints.

5. Using the fits in the angles corresponding to a projected distance range of 1–16 h^{-1} Mpc at the effective median redshift of the sample, we have calculated the rms fluctuations of the X-ray source distributions. Comparing them with that of the mass distribution from the linear theory, we find that the bias parameters of the X-ray sources are in the range $b_{\rm AGN} = 1.5$–4 at $0.5 < z < 1.2$.

6. If the bias mainly reflects the typical mass of dark matter halos in which these X-ray AGNs reside, their typical masses are 10^{13}–$10^{14}\ M_{\odot}$.

7. Further investigations utilizing redshifts of individual X-ray sources and/or involving cross-correlation function with galaxies taking advantage of the wealth of multiwavelength data are being conducted. The approved *Chandra* observations (C-COSMOS) on this field will enable us to probe into the clustering in much smaller scales and therefore into immediately neighboring environments of AGNs.

This work is based on observations obtained with XMM-Newton, an ESA science mission, with instruments and contributions directly funded by ESA member states and the US (NASA). Partial support of this work is provided by NASA LTSA Grant NAG5-10875 (T. M.) and the US *XMM-Newton* Guest Observer Support NNG04GG40G/NNG06GD77G (T. M., R. E. G.). In Germany, the *XMM-Newton* project is supported by the Bundesministerium für Bildung und Forschung/Deutsches Zentrum für Luft- und Raumfahrt, the Max-Planck Society, and the Heidenhain-Stiftung. Part of this work was supported by the Deutsches Zentrum für Luft- und Raumfahrt, DLR project numbers 50 OR 0207 and 50 OR 0405. In Italy, the COSMOS *XMM-Newton* project is supported by INAF and MIUR under grants PRIN/270/2003 and Cofin-03-02-23. The COSMOS Science meeting in 2005 May was supported in part by the NSF through grant OISE-0456439. We gratefully acknowledge the contributions of the entire COSMOS collaboration consisting of more than 70 scientists. In particular, we thank Manfred Kitbichler for allowing us to use his COSMOS-Mock catalogs and Chris Carilli for reviewing the manuscript. More information on the COSMOS survey is available at http://www.astro.caltech.edu/cosmos. It is a pleasure the acknowledge the excellent services provided by the NASA IPAC/IRSA staff (Anastasia Laity, Anastasia Alexov, Bruce Berriman, and John Good) in providing online archive and server capabilities for the COSMOS data sets.

Facilities: XMM

REFERENCES

Akylas, A., Georgantopoulos, I., & Plionis, M. 2000, MNRAS, 318, 1036
Barger, A. J., Cowie, L. L., Mushotzky, R. F., Yang, Y., Wang, W.-H., Steffen, A. T., & Capak, P. 2005, AJ, 129, 578
Basilakos, S., Georgakakis, A., Plionis, M., & Georgantopoulos, I. 2004, ApJ, 607, L79
Basilakos, S., Plionis, M., Georgakakis, A., & Georgantopoulos, I. 2005, MNRAS, 356, 183
Brusa, M., et al. 2007, ApJS, 172, 353
Cappelluti, N., Cappi, M., Dadina, M., Malaguti, G., Branchesi, M., D'Elia, V., & Palumbo, G. G. C. 2005, A&A, 430, 39
Cappelluti, N., et al. 2007, ApJS, 172, 341 (C07)
Cappi, M., et al. 2001, ApJ, 548, 624
Carrera, F. J., et al. 2007, A&A, in press (astro-ph/0703451)
Carroll, S. M., Press, W. H., & Turner, E. L. 1992, ARA&A, 30, 499
Comastri, A., Setti, G., Zamorani, G., & Hasinger, G. 1995, A&A, 296, 1
Croom, S. M., et al. 2005, MNRAS, 356, 415
D'Elia, V., Fiore, F., Elvis, M., Cappi, M., Mathur, S., Mazzotta, P., Falco, E., & Cocchia, F. 2004, A&A, 422, 11
Di Matteo, T., Springel, V., & Hernquist, L. 2005, Nature, 433, 604
Finoguenov, A., et al. 2007, ApJS, 172, 182
Gandhi, P., et al. 2006, A&A, 457, 393
Gilli, R., Salvati, M., & Hasinger, G. 2001, A&A, 366, 407
Gilli, R., et al. 2003, ApJ, 592, 721
———. 2005, A&A, 430, 811
Grazian, A., Negrello, M., Moscardini, L., Cristiani, S., Haehnelt, M. G., Matarrese, S., Omizzolo, A., & Vanzella, E. 2004, AJ, 127, 592
Hamilton, A. J. S. 2001, MNRAS, 322, 419
Hasinger, G., Miyaji, T., & Schmidt, M. 2005, A&A, 441, 417
Hasinger, G., et al. 2007, ApJS, 172, 29
Kitzbichler, M. G., & White, S. D. M. 2007, MNRAS, 376, 2
La Franca, F., et al. 2005, ApJ, 635, 864
Landy, S. D., & Szalay, A. S. 1993, ApJ, 412, 64
Lilly, S. J., et al. 2007, ApJS, 172, 70
Marinoni, C., et al. 2005, A&A, 442, 801
Menci, N., Fiore, F., Perola, G. C., & Cavaliere, A. 2004, ApJ, 606, 58
Mo, H. J., & White, S. D. M. 1996, MNRAS, 282, 347
Moretti, A., Campana, S., Lazzati, D., & Tagliaferri, G. 2003, ApJ, 588, 696
Mullis, C. R., Henry, J. P., Gioia, I. M., Böhringer, H., Briel, U. G., Voges, W., & Huchra, J. P. 2004, ApJ, 617, 192
Peebles, P. J. E. 1980, The Large Scale Structure of the Universe (Princeton: Princeton Univ. Press)
Porciani, C., Magliocchetti, M., & Norberg, P. 2004, MNRAS, 355, 1010
Puccetti, S., et al. 2006, A&A, 457, 501
Roche, N., & Eales, S. A. 1999, MNRAS, 307, 703
Scoville, N., et al. 2007, ApJS, 172, 1
Sheth, R. K., Mo, H. J., & Tormen, G. 2001, MNRAS, 323, 1
Spergel, D. N., et al. 2003, ApJS, 148, 175
Trump, R. J., et al. 2007, ApJS, 172, 383
Ueda, Y., Akiyama, M., Ohta, K., & Miyaji, T. 2003, ApJ, 598, 886
Vikhlinin, A., & Forman, W. 1995, ApJ, 455, L109
Yang, Y., Mushotzky, R. F., Barger, A. J., & Cowie, L. L. 2006, ApJ, 645, 68
Zehavi, I., et al. 2004, ApJ, 608, 16