THE CLUSTERING OF *XMM-NEWTON* HARD X-RAY SOURCES

S. BASILAKOS,[1] A. GEORGAKAKIS,[1] M. PLIONIS,[1,2] AND I. GEORGANTOPOULOS[1]

ABSTRACT

This Letter presents the clustering properties of hard (2–8 keV) X-ray–selected sources detected in a wide-field (≈ 2 deg^2), shallow [$f_X(2$–8 keV$) \approx 10^{-14}$ ergs cm^{-2} s^{-1}], and contiguous *XMM-Newton* survey. We perform an angular correlation function analysis using a total of 171 sources to the above flux limit. We detect an ~ 4 σ correlation signal out to 300″ with $w(\theta < 300″) \simeq 0.13 \pm 0.03$. Modeling the two-point correlation function as a power law of the form $w(\theta) = (\theta_0/\theta)^{\gamma-1}$, we find $\theta_0 = 48.9^{+15.8}_{-24.5}$ arcsec and $\gamma = 2.2 \pm 0.30$. Fixing the correlation function slope to $\gamma = 1.8$, we obtain $\theta_0 = 22.2^{+9.4}_{-8.6}$ arcsec. Using Limber's integral equation and a variety of possible luminosity functions of the hard X-ray population, we find a relatively large correlation length, ranging from $r_0 \sim 9$ to 19 h^{-1} Mpc (for $\gamma = 1.8$ and the *concordance* cosmological model), with this range reflecting also different evolutionary models for the source luminosities and clustering characteristics. The relatively large correlation length is comparable to that of extremely red objects and luminous radio sources.

Subject headings: cosmology: observations — galaxies: active — large-scale structure of universe — quasars: general — surveys — X-rays: diffuse background

1. INTRODUCTION

It is well known that the study of the distribution of matter on large scales, using different extragalactic objects, provides important constraints on models of cosmic structure formation. Since active galactic nuclei (AGNs) can be detected up to very high redshifts, they provide information on the underlying mass distribution as well as on the evolution of large-scale structure (see Hartwick & Schade 1990; Basilakos 2001 and references therein).

The traditional indicator of clustering, the angular two-point correlation function, is a fundamental and simple statistical test for the study of any extragalactic mass tracer and is relatively straightforward to measure from observational data. The overall knowledge of the AGN clustering using X-ray data comes mostly from the soft (≤ 3 keV) X-ray band (Boyle & Mo 1993; Vikhlinin & Forman 1995; Carrera et al. 1998; Akylas et al. 2000; Mullis 2002), which is, however, biased against absorbed AGNs. Hard X-ray surveys (≥ 2 keV) play a key role in our understanding of how the whole AGN population, including obscured (type II) AGNs, trace the underlying mass distribution. Furthermore, understanding the spatial distribution of type II AGNs is important since they are among the main contributors of the cosmic X-ray background (Mushotzky et al. 2000; Hasinger et al. 2001; Giacconi et al. 2002).

Recently, Yang et al. (2003), performing a counts-in-cells analysis of a deep ($f_{2-8\,\mathrm{keV}} \sim 3 \times 10^{-15}$ ergs s^{-1} cm^{-2}) *Chandra* survey in the Lockman Hole northwest region, found that the hard-band sources are highly clustered with $\sim 60\%$ of them being distributed in overdense regions. The *XMM-Newton*, with ~ 5 times more effective area, especially at hard energies, and ~ 3 times larger field of view, is an ideal instrument for clustering studies of X-ray sources.

In this Letter, we estimate for the first time the angular correlation function of the *XMM-Newton* hard X-ray sample. Using Limber's equation and different models of the luminosity func-

tion for these sources, we derive the expected spatial correlation function, which we compare with that of a variety of extragalactic populations. Hereafter, all H_0-dependent quantities will be given in units of $h \equiv H_0/100$ km s^{-1} Mpc^{-1}.

2. THE SAMPLE

The hard X-ray–selected sample used in the present study is compiled from the *XMM-Newton* Two-Degree Field (2dF) survey. This is a shallow (2–10 ks per pointing) survey carried out by the *XMM-Newton* near the north Galactic pole [NGP; R.A.(J2000.0) = $13^\mathrm{h}41^\mathrm{m}$, decl.(J2000.0) = $00°00′$] and the south Galactic pole [SGP; R.A.(J2000.0) = $00^\mathrm{h}57^\mathrm{m}$, decl.(J2000.0) = $-28°00′$] regions. A total of 18 *XMM-Newton* pointings were observed equally split between the NGP and the SGP areas. However, a number of pointings were discarded because of elevated particle background at the time of the observation. This results in a total of 13 usable *XMM-Newton* pointings covering an area of ≈ 2 deg^2. A full description of the data reduction, source detection, and flux estimation is presented by Georgakakis et al. (2003, 2004). For the two-dimensional correlation analysis presented in this Letter, we use the hard (2–8 keV) band catalog of the *XMM-Newton* 2dF survey. This comprises a total of 171 sources above the 5 σ detection threshold to the limiting flux of $f_X(2$–8 keV$) \approx 10^{-14}$ ergs s^{-1} cm^{-2}. Note that our hard X-ray sources consist of a mixture of QSOs and relatively nearby ($z < 0.8$) galaxies with red colors $g - r > 0.5$, which are most probably associated with obscured low-luminosity AGNs (Georgantopoulos et al. 2004).

3. CORRELATION FUNCTION ANALYSIS

3.1. *The Angular Correlation*

The clustering properties of the hard X-ray–selected sources are estimated using the two-point angular correlation function $w(\theta)$ defined by $dP = n^2[1 + w(\theta)]d\Omega_1 \, d\Omega_1$, where dP is the joint probability of finding two sources in the solid angle elements $d\Omega_1$ and $d\Omega_2$ separated by angle θ and n is the mean surface density of sources. For a random distribution of sources, $w(\theta) = 0$. Therefore, the angular correlation function provides a measure of galaxy density excess over that expected for a

[1] Institute of Astronomy and Astrophysics, National Observatory of Athens, I. Metaxa and B. Pavlou, P. Penteli, Athens 152 36, Greece; svasil@astro.noa.gr, age@astro.noa.gr, ig@astro.noa.gr.
[2] Instituto Nacional de Astrofisica, Optica y Electronica, Apartado Postal 51 y 216, Puebla, Puebla 72000, Mexico; mplionis@inaoep.mx.
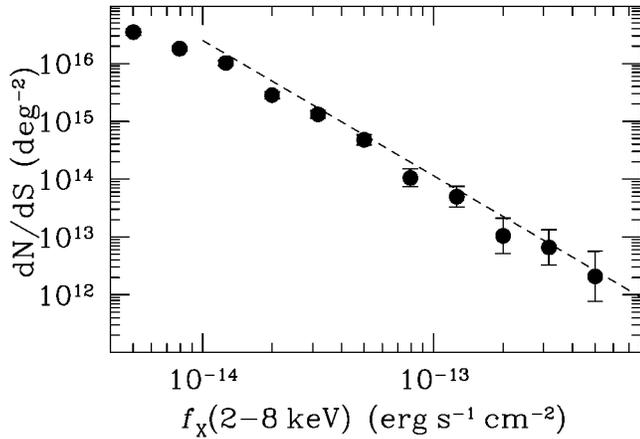
Fig. 1.—Hard-band (2–8 keV) differential number counts from the present survey in comparison with those of Baldi et al. (2002; *dashed line*).
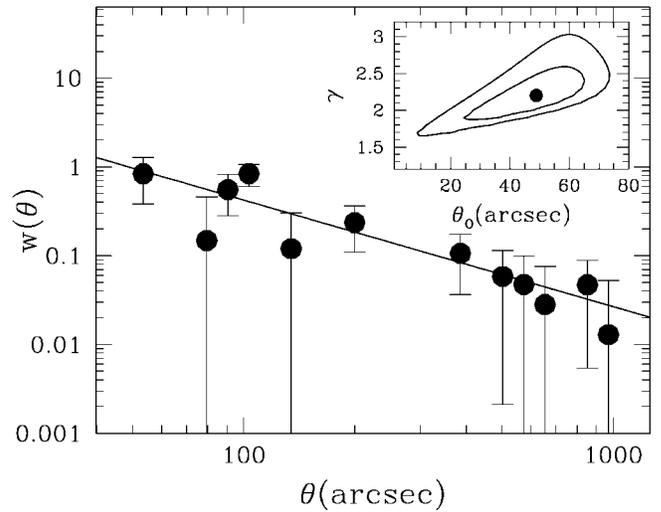

Fig. 2.—Two-point angular correlation function of the hard (2–8 keV) X-ray sources. The line represents the best-fit power law $w(\theta) = (\theta_0/\theta)^{\gamma-1}$ with $\theta_0 = 48\farcs9$ and $\gamma = 2.2$. *Inset*: Iso-$\Delta\chi^2$ contours in the γ-θ_0 parameter space.

random distribution. A variety of estimators of $w(\theta)$ have been used over the years (see Infante 1994).

In the present study, we use the estimator (see Efstathiou et al. 1991)

$$w(\theta) = f(N_{DD}/N_{DR}) - 1, \quad (1)$$

where N_{DD} and N_{DR} are the number of data-data and data-random pairs, respectively, at separations θ and $\theta + d\theta$. In the above relation, f is the normalization factor $f = 2N_R/(N_D - 1)$, where N_D and N_R are the total number of data and random points, respectively. The uncertainty in $w(\theta)$ is estimated as $\sigma_w = \{[1 + w(\theta)]/N_{DR}\}^{1/2}$ (Peebles 1973). To account for the different source selection and edge effects, we have produced 100 Monte Carlo random realizations of the source distribution within the area of the survey by taking into account variations in sensitivity that might affect the correlation function estimate. Indeed, the flux threshold for detection depends on the off-axis angle from the center of each of the *XMM-Newton* pointings. Since the random catalog must have the same selection effects as the real catalog, sensitivity maps are used to discard random points in less sensitive areas (close to the edge of the pointings). This is accomplished, to the first approximation, by assigning a flux to the random points using the Baldi et al. (2002) 2–10 keV log N–log S (after transforming to the 2–8 keV band assuming $\Gamma = 1.7$). If the flux of a random point is less than 5 times the local rms noise (assuming Poisson statistics for the background), the point is excluded from the random data set. We note that the Baldi et al. (2002) log N–log S is in good agreement with the 2–8 keV number counts estimated in the present survey. This is demonstrated in Figure 1, where we plot our differential number counts and the best-fit relation of Baldi et al. (2002). Note that we have tested that our random simulations reproduce both the off-axis sensitivity of the detector as well as the individual field log N–log S.

We apply the correlation analysis, evaluating $w(\theta)$ in logarithmic intervals with $\delta \log \theta = 0.05$. The results are shown in Figure 2, where the line corresponds to the best-fit power-law model $w(\theta) = (\theta_0/\theta)^{\gamma-1}$ using the standard χ^2 minimization procedure in which each correlation point is weighted by its error. We find a statistically significant signal with $w(\theta < 300'') \simeq 0.13 \pm 0.03$ at the 4.3 σ and ~2.7 σ confidence levels using Poissonian or bootstrap errors, respectively. Note that the bootstrap errors probably overestimate the true uncertainty, es-

pecially in sparse samples (Fisher et al. 1994). Therefore, the true significance level is somewhere in between the above two values.

In the inset of Figure 2, we present the iso-$\Delta\chi^2$ contours (where $\Delta\chi^2 = \chi^2 - \chi^2_{\min}$) in the γ-θ_0 plane. The contours correspond to 1 σ ($\Delta\chi^2 = 2.30$) and 2 σ ($\Delta\chi^2 = 6.17$) uncertainties, respectively. The best-fit clustering parameters are

$$\theta_0 = 48.9^{+15.8}_{-24.5} \text{ arcsec}, \quad \gamma = 2.2 \pm 0.30, \quad (2)$$

where the errors correspond to 1 σ ($\Delta\chi^2 = 2.30$) uncertainties. Fixing the correlation function slope to its nominal value, $\gamma = 1.8$, we estimate $\theta_0 = 22.2^{+9.4}_{-8.6}$ arcsec.[3] Note that our results do not suffer from the *amplification bias*, which results from merging close source pairs when the point-spread function (PSF) size is larger than their typical separation (see Vikhlinin & Forman 1995). This is because the estimated θ_0-values are much larger than the *XMM-Newton* PSF size of 6″ FWHM.

Another systematic effect that could bias the angular two-point correlation function is that introduced by the so-called *integral constraint*. This results from the fact that the correlation function is estimated from a limited area, which in turn implies that over the area studied the relation $\iint w(\theta_{12}) d\Omega_1 d\Omega_2 = 0$ should be satisfied. We can attempt to estimate the resulting underestimation of the true correlation function by calculating the quantity $W = \int d\Omega_1 \int d\Omega_2 w(\theta_{12})/\int d\Omega_1 \int d\Omega_2$. Clearly, evaluating W necessitates a priori knowledge of the angular correlation function. A tentative value of W using a range of $w(\theta)$ given by varying within 1 σ our results (eq. [2]) is $W \simeq 0.02$. By adding W to our estimated (raw) $w(\theta)$ and fitting again the model correlation function, we find

$$\theta_0 \simeq 44'' \pm 20'', \quad \gamma \simeq 2 \pm 0.25, \quad (3)$$

consistent within the errors, with our uncorrected results (eq. [2]). If we fix $\gamma = 1.8$, we obtain $\theta_0 = 28'' \pm 9''$. Owing to the small effect of the integral constraint correction, we will use in the rest of the Letter our *raw* $w(\theta)$ estimates.

---

[3] The robustness of our results to the fitting procedure was tested using different bins (spanning from 10 to 20), and no significant difference was found.
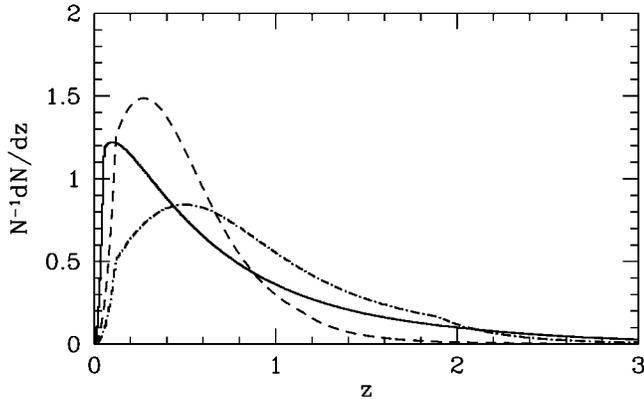
Fig. 3.—Predicted $N^{-1}\,dN/dz$ distribution for three different luminosity function models: (1) Boyle et al. (1998) with PLE (*continuous line*), (2) Ueda et al. (2003) with PLE (*dashed line*), and (3) Ueda et al. (2003) with LDDE (*dot-dashed line*).

Our results show that hard X-ray sources are strongly correlated, even more than the soft ones (see Vikhlinin & Forman 1995; Yang et al. 2003; S. Basilakos et al. 2004, in preparation). Our derived angular correlation length θ_0 is in rough agreement with, although somewhat smaller than (within 1 σ), the *Chandra* result of $\theta_0 = 40'' \pm 11''$ (Yang et al. 2003). The stronger angular clustering with respect to the soft sources either could be due to the higher flux limit of the hard *XMM-Newton* sample, resulting in the selection of relatively nearby sources, or could imply an association of our hard X-ray sources with high-density peaks of the underlying matter distribution. To test the latter suggestion, we have measured the cluster hard X-ray source cross-correlation function ($w_{c,\,\mathrm{hard}}$) using either the Goto et al. (2002) clusters, detected in the multicolor optical Sloan Digital Sky Survey (SDSS) data, or the X-ray clusters that we have detected on our fields (T. Gaga et al. 2004, in preparation). We find no significant cross-correlation signal [$w_{c,\,\mathrm{hard}}(\theta < 300'') \simeq -0.15 \pm 0.19$], a fact that weakens the suggestion of an association of the hard X-ray sources with high-density peaks. However, we must stress that the null result could be artificial, owing to small number statistics.

3.2. *The Spatial Correlation Length Using* $w(\theta)$

The angular correlation function $w(\theta)$ can be obtained from the spatial one, $\xi(r)$, through the Limber transformation (Peebles 1980). If the spatial correlation function is modeled as

$$\xi(r, z) = (r/r_0)^{-\gamma}(1 + z)^{-(3+\epsilon)}, \quad (4)$$

then for a flat universe the amplitude θ_0 in two dimensions is

related to the correlation length r_0 (see Efstathiou et al. 1991) in three dimensions through the equation

$$\theta_0^{\gamma-1} = H_\gamma r_0^\gamma \left(\frac{H_0}{c}\right)^\gamma \int_0^\infty \left(\frac{1}{N}\frac{dN}{dz}\right)^2 \frac{E(z)}{x^{\gamma-1}(z)}\,(1 + z)^{-3-\epsilon+\gamma}\,dz, \quad (5)$$

where $x(z)$ is the proper distance, $E(z) = [\Omega_m(1 + z)^3 + \Omega_\Lambda]^{1/2}$ is the element of comoving distance, and $H_\gamma = \Gamma(1/2)\Gamma\,[(\gamma - 1)/2]\,/\Gamma(\gamma/2)$. Note that if $\epsilon = \gamma - 3$ the clustering is constant in comoving coordinates (comoving clustering), while if $\epsilon = -3$ the clustering is constant in physical coordinates. We perform the above inversion in the framework of either the *concordance* ΛCDM cosmological model ($\Omega_m = 1 - \Omega_\Lambda = 0.3, H_0 = 70$ km s^{-1} Mpc^{-1}) or the Einstein–de Sitter model.

The redshift distribution dN/dz and the predicted total number, N, of the X-ray sources that enter in equation (5) can be found using the hard-band luminosity functions of Ueda et al. (2003) and of Boyle et al. (1998). We also use different models for the evolution of the hard X-ray sources: a pure luminosity evolution (PLE) or the more realistic luminosity-dependent density evolution (LDDE; Ueda et al. 2003). In Figure 3, we show the expected redshift distributions of the hard X-ray sources for three different luminosity functions and evolution models. Both the Boyle et al. (1998) and Ueda et al. (2003) luminosity functions with PLE give relatively similar dN/dz distributions. However, the LDDE model gives a dN/dz distribution shifted to much larger redshifts with a median redshift of $\bar{z} \simeq 0.75$ (see also Table 1).

For the comoving clustering model ($\epsilon = \gamma - 3$) and using the LDDE evolution model, we estimate the hard X-ray source correlation length to be $r_0 = 19 \pm 3\ h^{-1}$ Mpc and $r_0 = 13.5 \pm 3\ h^{-1}$ Mpc for $\gamma = 1.8$ and $\gamma = 2.2$, respectively. If $\epsilon = -3$, the corresponding values are $r_0 = 11.5 \pm 2\ h^{-1}$ Mpc and $r_0 = 6 \pm 1.5\ h^{-1}$ Mpc, respectively. In Table 1, we present the values of the correlation length, r_0, resulting from Limber's inversion for different luminosity function and evolution models.

These estimated clustering lengths (for $\gamma = 1.8$) are a factor of $\gtrsim 2$ larger than the corresponding values of the Lyman break galaxies (Adelberger 2000), the 2dF (Hawkins et al. 2003) and SDSS (Budavari et al. 2003) galaxy distributions, as well as the 2QZ QSOs (Croom et al. 2001). However, the most luminous, and thus nearer, 2QZ subsample ($18.25 < b_j < 19.80$) has a larger correlation length ($\sim 8.5 \pm 1.7\ h^{-1}$ Mpc) than the overall sample (Croom et al. 2002), in marginal agreement with our $\epsilon = -3$ clustering evolution results.

The large spatial clustering length of our hard X-ray sources can be compared with that of extremely red objects (EROs) and luminous radio sources (Roche et al. 2003; Overzier et al. 2003; Röttgering et al. 2003), which are found to be in the

TABLE 1
CORRELATION LENGTHS FOR DIFFERENT MODELS

Luminosity Function	Evolution Model	$(\Omega_m, \Omega_\Lambda)$	$r_0\ (1.8, -1.2)$ (h^{-1} Mpc)	$r_0\ (1.8, -3)$ (h^{-1} Mpc)	$r_0\ (2.2, -0.8)$ (h^{-1} Mpc)	$r_0\ (2.2, -3)$ (h^{-1} Mpc)	$\bar{z}$
Boyle	No evolution	(1, 0)	11.5 ± 2.00	9.0 ± 1.5	7.3 ± 1.2	6.0 ± 1.5	0.45
Ueda	No evolution	(1, 0)	9.5 ± 1.5	7.5 ± 1.0	6.5 ± 1.5	5.0 ± 1	0.40
Boyle	PLE	(1, 0)	13.0 ± 3.0	10.0 ± 2.0	8.0 ± 2.0	6.8 ± 1.5	0.50
Ueda	**PLE**	**(0.3, 0.7)**	13.0 ± 2.0	9.3 ± 1.6	9.0 ± 2.0	6.7 ± 1.5	**0.45**
Ueda	**LDDE**	**(0.3–0.7)**	19.0 ± 3.0	11.5 ± 2.0	13.5 ± 3	8.5 ± 2	**0.75**
Ueda	LDDE	(1, 0)	15.0 ± 2.5	7.8 ± 1.5	11.0 ± 2.5	6.0 ± 1.5	0.80

NOTES.—The table shows the hard X-ray source correlation length for different pairs of (γ, ϵ) and for the different luminosity functions and evolution models. The last column indicates the predicted median redshift, from the specific luminosity function used. Boldface delineates the preferred cosmological model and the most updated luminosity function.

range $r_0 \simeq 12$–15 h^{-1} Mpc. The possible association of EROs with high-z massive elliptical galaxies and of luminous radio sources with protoclusters (for a review, see Röttgering et al. 2003 and references therein) suggests that our hard X-ray sources could trace the high peaks of the underlying mass distribution (see also Yang et al. 2003).

4. CONCLUSIONS

In this Letter, we explore the clustering properties of hard (2–8 keV) X-ray–selected sources using a wide area (≈ 2 deg^2), shallow [$f_X(2$–8 keV$) \approx 10^{-14}$ ergs cm^{-2} s^{-1}] *XMM-Newton* survey. Using an angular correlation function analysis, we measure a clustering signal at the ~ 4 σ confidence level. Modeling the angular correlation function by a power law, $w(\theta) = (\theta_0/\theta)^{\gamma-1}$, we estimate $\theta_0 = 48.9^{+15.8}_{-24.5}$ arcsec and $\gamma = 2.2 \pm 0.30$. Fixing the correlation function slope to $\gamma = 1.8$, we estimate $\theta_0 = 22.2^{+9.4}_{-8.6}$ arcsec. When we use a variety of luminosity functions and evolutionary models, Limber's inversion provides correlation lengths that are in the range $r_0 \sim 10$–19 h^{-1} Mpc, typically larger than those of galaxies and optically selected QSOs but similar to those of strongly clustered populations, like EROs and luminous radio sources.

REFERENCES

Adelberger, K. 2000, in ASP Conf. Ser. 200, Clustering at High Redshift, ed. A. Mazure, O. Le Fèvre, & V. Le Brun (San Francisco: ASP), 13
Akylas, A., Georgantopoulos, I., & Plionis, M. 2000, MNRAS, 318, 1036
Baldi, A., Molendi, S., Comastri, A., Fiore, F., Matt, G., & Vignali, C. 2002, ApJ, 564, 190
Basilakos, S. 2001, MNRAS, 326, 203
Boyle, B. J., & Mo, H. J. 1993, MNRAS, 260, 925
Boyle, B. J., et al. 1998, MNRAS, 296, 1
Budavari, T., et al. 2003, ApJ, 595, 59
Carrera, F. J., Barcons, X., Fabian, A. C., Hasinger, G., Mason, K. O., McMahon, R. G., Mittaz, J. P. D., & Page, M. J. 1998, MNRAS, 299, 229
Croom, S. M., Boyle, B. J., Loaring, N. S., Miller, L., Outram, P. J., Shanks, T., & Smith, R. J. 2002, MNRAS, 335, 459
Croom, S. M., Shanks, T., & Boyle, B. J. 2001, MNRAS, 325, 483
Efstathiou, G., Bernstein, G., Katz, N., Tyson, J. A., & Guhathakurta, P. 1991, ApJ, 380, L47
Fisher, K. B., Davis, M., Strauss, M. A., Yahil, A., & Huchra, J. 1994, MNRAS, 266, 50
Georgakakis, A., Georgantopoulos, I., Stewart, G. C., Shanks, T., & Boyle, B. J. 2003, MNRAS, 344, 161
Georgakakis, A., et al. 2004, MNRAS, 349, 135
Georgantopoulos, I., et al. 2004, MNRAS, in press (astro-ph/0404048)

Giacconi, R., et al. 2002, ApJS, 139, 369
Goto, T., et al. 2002, AJ, 123, 1807
Hartwick, F. D. A., & Schade, D. 1990, ARA&A, 28, 437
Hasinger, G., et al. 2001, A&A, 365, L45
Hawkins, E., et al. 2003, MNRAS, 346, 78
Infante, L. 1994, A&A, 282, 353
Mullis, C. R. 2002, PASP, 114, 688
Mushotzky, R. F., Cowie, L. L., Barger, A. J., & Arnaud, K. A. 2000, Nature, 404, 459
Overzier, R. A., Röttgering, H., Rengelink, R. B., & Wilman, R. J. 2003, A&A, 405, 53
Peebles, P. J. E. 1973, ApJ, 185, 413
———. 1980, The Large-Scale Structure of the Universe (Princeton: Princeton Univ. Press)
Roche, N. D., Dunlop, J., & Almaini, O. 2003, MNRAS, 346, 803
Röttgering, H., Daddi, E., Overzier, R. A., & Wilman, R. J. 2003, NewA Rev., 47, 309
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., Sanders, D. B., & Steffen, A. T. 2003, ApJ, 585, L85