The Astrophysical Journal, 690:1250–1263, 2009 January 10 doi:10.1088/0004-637X/690/2/1250

PHOTOMETRIC REDSHIFT AND CLASSIFICATION FOR THE *XMM*–COSMOS SOURCES*

M. Salvato[1], G. Hasinger[2,3], O. Ilbert[3], G. Zamorani[4], M. Brusa[2], N. Z. Scoville[1], A. Rau[1], P. Capak[1,5], S. Arnouts[6], H. Aussel[7], M. Bolzonella[4], A. Buongiorno[2], N. Cappelluti[2], K. Caputi[8], F. Civano[9], R. Cook[1,10,11], M. Elvis[9], R. Gilli[4], K. Jahnke[12], J. S. Kartaltepe[3], C. D. Impey[13], F. Lamareille[2], E. Le Floch[3], S. Lilly[8], V. Mainieri[14], P. McCarthy[15], H. McCracken[16], M. Mignoli[4], B. Mobasher[17], T. Murayama[18], S. Sasaki[19], D. B. Sanders[3], D. Schiminovich[4], Y. Shioya[20], P. Shopbell[1], J. Silverman[8], V. Smolčić[1], J. Surace[5], Y. Taniguchi[20], D. Thompson[21], J. R. Trump[13], M. Urry[22], and M. Zamojski[5,23]

[1] California Institute of Technology, MC 105-24, 1200 East California Boulevard, Pasadena, CA 91125, USA
[2] Max Planck Institut für extraterrestrische Physik, Giessenbachstrasse 1, 85748 Garching, Germany
[3] Institute for Astronomy, University of Hawaii, 2680 Woodlawn Drive, Honolulu, HI 96822, USA
[4] Istituto Nazionale de Astrofisica (INAF)–Osservatorio Astronomico di Bologna, via Ranzani 1, I-40127 Bologna, Italy
[5] *Spitzer* Science Center, California Institute of Technology, Pasadena, CA 91125, USA
[6] CFHT corporation, 65-1238 Mamalahoa Hwy, Kamuela, HI 96743, USA
[7] CEA/DSM-CNRS, Universite' Paris Diderot, DAPNIA/SAp, Orme des Merisiers, 91191, Gif-sur-Yvette, France
[8] Department of Physics, Eidgenossiche Technische Hochschule (ETH), CH-8093 Zurich, Switzerland
[9] Harvard-Smithsonian Center for Astrophysics 60 Garden St., Cambridge, MA 02138, USA
[10] Department of Physics and Astronomy, University College London, Gower Street, London, WC1E 6BT, UK
[11] Physics Department, Brown University, Box 1843, Providence, RI 02912, USA
[12] Max Planck Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany
[13] Steward Observatory, University of Arizona, Tucson, AZ 85721, USA
[14] European Southern Observatory, Karl-Schwarzschild-str. 2, 85748 Garching, Germany
[15] Carnegie Observatories, 813 Santa Barbara Street, Pasadena, CA, 91101, USA
[16] Institut d'Astrophysique de Paris, UMR 7095, CNRS, Université Pierre et Marie Curie, 98 bis Boulevard Arago, F-75014 Paris, France
[17] Department of Physics and Astronomy, University of California, Riverside, CA 92521, USA
[18] Astronomical Institute, Graduate School of Science, Tohoku University, Aramaki, Aoba, Sendai 980-8578, Japan
[19] Physics Department, Graduate School of Science and Engineering, Ehime University, 2-5 Bunkyo-cho, Matsuyama 790-8577, Japan
[20] Research Center for Space and Cosmic Evolution, Ehime University, Bunkyo-cho 2-5, Matsuyama 790-8577, Japan
[21] LBT Observatory, University of Arizona, 933 North Cherry Avenue, Tucson, AZ 85721-0065, USA
[22] Department of Physics, Yale University, P.O. Box 208121, New Haven, CT 06520-8121, USA
[23] Department of Astronomy, Columbia University, MC2457, 550 W. 120 St. New York, NY 10027, USA
Received 2008 August 15; accepted 2008 September 9; published 2008 December 8

ABSTRACT

We present photometric redshifts and spectral energy distribution (SED) classifications for a sample of 1542 optically identified sources detected with *XMM* in the COSMOS field. Our template fitting classifies 46 sources as stars and 464 as nonactive galaxies, while the remaining 1032 require templates with an active galactic nucleus (AGN) contribution. High accuracy in the derived photometric redshifts was accomplished as the result of (1) photometry in up to 30 bands with high-significance detections, (2) a new set of SED templates, including 18 hybrids covering the far-UV to mid-infrared, which have been constructed by the combination of AGNs and nonactive galaxies templates, and (3) multiepoch observations that have been used to correct for variability (most important for type 1 AGNs). The reliability of the photometric redshifts is evaluated using the subsample of 442 sources with measured spectroscopic redshifts. We achieved an accuracy of $\sigma_{\Delta z/(1+z_{\rm spec})} = 0.014$ for $i^*_{AB} < 22.5$ ($\sigma_{\Delta z/(1+z_{\rm spec})} \sim 0.015$ for $i^*_{AB} < 24.5$). The high accuracies were accomplished for both type 2 (where the SED is often dominated by the host galaxy) and type 1 AGNs and QSOs out to $z = 4.5$. The number of outliers is a large improvement over previous photometric redshift estimates for X-ray-selected sources (4.0% and 4.8% outliers for $i^*_{AB} < 22.5$ and $i^*_{AB} < 24.5$, respectively). We show that the intermediate band photometry is vital to achieving accurate photometric redshifts for AGNs, whereas the broad SED coverage provided by mid-infrared (*Spitzer*/IRAC) bands is important to reduce the number of outliers for normal galaxies.

Key words: catalogs – galaxies: active – methods: miscellaneous

1. INTRODUCTION

It is now well established that active galactic nuclei (AGNs) play a significant role in cosmic evolution—tracing the growth of super massive black holes in galaxy nuclei and providing important energetic feedback to the host galaxies and the intergalactic medium (IGM). Characterization of the AGN properties and their redshifts is therefore of paramount importance to understanding cosmological evolution. The mass of the central super massive black hole (SMBH) directly correlates with the

* Based on observations with the NASA/ESA *Hubble Space Telescope*, obtained at the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc. (AURA) under NASA contract NAS 5-26555. Also based on observations made with the *Spitzer Space Telescope*, which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under NASA contract 1407. Also based on data collected at: the Subaru Telescope, which is operated by the National Astronomical Observatory of Japan; the *XMM-Newton*, an ESA science mission with instruments and contributions directly funded by ESA member states and NASA; the European Southern Observatory under Large Program 175.A-0839, Chile; Kitt Peak National Observatory, Cerro Tololo Inter-American Observatory and the National Optical Astronomy Observatory, which are operated by AURA under cooperative agreement with the National Science Foundation; and the Canada–France–Hawaii Telescope with MegaPrime/MegaCam operated as a joint project by the CFHT Corporation, CEA/DAPNIA, the NRC and CADC of Canada, the CNRS of France, TERAPIX and the University of Hawaii.

mass and the velocity dispersion of the bulge of its host galaxy (Gebhardt et al. 2000; Ferrarese & Merritt 2000) and these correlations apparently persist both for normal, less active galaxies and for AGNs (McLure & Dunlop 2002). This is strongly suggesting a coevolution of galaxies and their nuclear BHs. Unfortunately, given the low fraction of galaxies with high-luminosity AGNs and the fact that those with high luminosity may give a biased view of the overall importance of the AGN feedback, the codependent evolution of galaxies and BHs has been pursued very little at high redshift, where the majority of both galaxy and AGN evolution occurs. Here we address this problem—attempting to identify a much larger sample of AGNs over a range of luminosity and to provide classifications and redshifts for much fainter samples.

At present, no criterion exists to identify all AGNs in a field using a single photometric band. Optical, infrared, radio, narrow bands, and X-ray techniques each yield incomplete, partially overlapping, subsamples of the overall AGN population. A complete census can only be achieved by combining the results of the various selection techniques (G. Zamorani et al. 2009, in preparation). Nevertheless, the X-ray band appears to be very efficient and rather complete for most classes of AGN (Brandt & Hasinger 2005).

The COSMOS survey (Scoville et al. 2007b) provides full multiwavelength data sets in 30+ bands from X-ray to radio in 2 deg^2, and we focus here on *XMM*–COSMOS selected sources (Hasinger et al. 2007) using spectral energy distributions (SEDs) to disentangle between nonactive galaxies and AGNs and to provide accurate photometric redshifts. Photometric redshifts for normal (inactive) galaxies can reach an accuracy of $\sigma_{\Delta z/(1+z)} = 0.05$ (Grazian et al. 2006; Ilbert et al. 2006) or better (Gabasch et al. 2004; Wolf et al. 2004) and most recently, better than 0.02 (Ilbert et al. 2009). In contrast, for AGNs the most accurate results have been 0.06–0.1 (Zheng et al. 2004; Polletta et al. 2007; Rowan-Robinson et al. 2008) for objects at redshift $z < 1$.

Uncertainties in the SED fitting of AGNs come from two major sources. First, AGNs can vary with time and quasi-simultaneous photometry is required to obtain a snapshot SED. Second, the observed emission is often a superposition of an AGN and a host galaxy. Depending on the properties of the host and the relative power of the nucleus, the different spectral bands can be dominated by either of the two components. For example, NGC 3079 is a Seyfert 2 galaxy, identified on the basis of its optical nucleus and a radio jet, yet its *Spitzer*/IRS spectrum is often used as a template for nonactive, starburst galaxies in the infrared (Weedman et al. 2006a). On the other hand, Mrk231, a Seyfert 1 and an ultraluminous infrared galaxy (ULIRG), exhibits strong silicate absorption in the IRS spectrum and is therefore highly obscured with much AGN (and starburst) luminosity reemitted in the far-infrared. Clearly, to analyze samples which, as *XMM*–COSMOS, contain both galaxies and AGNs, one needs: (1) a fine sampling and full spectral coverage with photometry spanning from the UV to mid-infrared and (2) SED templates representing the full range of AGNs and QSOs, together with variable-mixing fractions of AGN and host galaxy emission.

Only recently, with *Spitzer* and *Galaxy Evolution Explorer* (*GALEX*) providing full SED templates of AGNs beyond the optical bands (e.g., Polletta et al. 2007), such analysis has become possible at moderately high redshift.

In this paper, we derive photometric redshifts for the *XMM*–COSMOS sources—mostly dominated by an AGN (see Section 5)—achieving an accuracy comparable to that obtained for nonactive galaxies.

In Section 2, we describe the sample and the photometric data set. The variability analysis is presented in Section 3. In Section 4, the procedure used for the AGN classification and photometric redshift determination is discussed. The results are given in Section 5 and the description on the released catalog is given in Section 6. The discussion and the conclusions end the paper in Sections 7 and 8. Throughout this work we use AB magnitudes and assume that $H_0 = 70$ km s^{-1}Mpc^{-1}, $\Omega_\Lambda = 0.7$, and $\Omega_M = 0.3$.

2. THE DATA SET

2.1. The XMM–COSMOS Sample

The COSMOS field has been observed with *XMM-Newton* for a total of ~ 1.55 Ms (Hasinger et al. 2007; Cappelluti et al. 2007, 2009) at the homogeneous depth of ~ 50 ks. The final catalog includes point-like sources detected above a given threshold with a maximum likelihood detection algorithm in at least one of the soft (0.5–2 keV), hard (2–10 keV) or ultrahard (5–10 keV) bands down to limiting fluxes of 5×10^{-16}, 3×10^{-15}, and 5×10^{-15} erg cm^{-2} s^{-1}, respectively (see Cappelluti et al. 2007, 2009 for more details). The adopted threshold (Lik > 10) corresponds to a probability of $\sim 4.5 \times 10^{-5}$ that a source is a background fluctuation. A detailed X-ray of optical association using COSMOS optical, near- and mid-infrared catalogs was performed using the maximum likelihood (ML) technique (Brusa et al. 2007, 2009). For the subfield of the *XMM*–COSMOS covered also by *Chandra*, the optical identifications have been augmented with the more accurate *Chandra* positions (F. Civano et al. 2009, in preparation; Elvis et al. 2009). As a result, the sample can be divided into the following subgroups.

1. 1528 sources (78.9%) have unique and secure optical counterparts within 5.″0 from the *XMM* coordinates. We will refer to these sources as the "secure sample."
2. 52 sources are brighter than $i^* = 16.5$ and are saturated in many (if not all) optical bands. 37 of these are associated with bright stars, and one is a galaxy with spectroscopic redshift at 0.122. These 38 sources are removed from the input catalog as all the photometry is saturated. We will refer to the remaining 14 sources as the "bright sample."
3. Ten sources are undetected in the optical catalog and the identification of the counterpart is done in the near- and mid-infrared using the *K* band (H. McCracken et al. 2009, in preparation) and IRAC catalogs.[24] For this "optically undetected sample," not more than five photometric points are available.
4. 233 sources have an ambiguous optical cross-identification. For these objects, the ML technique yielded two sources with similar probability. This sample is referred to as having "ambiguous counterparts."

The photometric redshift determination is tuned using only the "secure" and the "bright" samples (1528 + 14 = 1542 sources); all of the results and the discussion below refer only to them. The optically undetected sample will be discussed separately in Section 5.3, while for the sample with ambiguous counterparts, we will provide photometric redshift for the primary and the secondary counterparts in the catalog, but they are not included in the redshift distribution or in the discussion.

[24] http://irsa.ipac.caltech.edu/data/COSMOS/

2.2. Multiwavelength Data

For the 1542 *XMM* sources associated with an optical/infrared counterpart, the photometric redshift has been computed using broad and intermediate band photometry in the range $1350\,\text{Å} < \lambda < 8\,\mu\text{m}$ using *GALEX*, Subaru, Sloan Digital Sky Survey (SDSS), Canada–France–Hawaii Telescope (CFHT), and *Spitzer*/IRAC.

The optical/near-infrared photometry is extracted from the recently updated COSMOS photometric catalog (P. Capak et al. 2008, in preparation). Compared to the previous catalog (Capak et al. 2007), the new photometry implements: (1) a more accurate algorithm for source detection, (2) new, deeper (and with better seeing) u^*- and K_S-band imaging from CFHT, and (3) new J-band data (UKIRT) and 12 additional intermediate and two narrow bands (from Subaru; Taniguchi et al. 2007; N. Taniguchi et al. 2009, in preparation). The fluxes were measured in fixed apertures of 3.″0 diameter, on point-spread function (PSF) matched images (FWHM of 1.″5). Simulations (Capak et al. 2007) show that the flux in 3.″0 aperture corresponds to 0.759 of the total flux for a point-like source.

For each optical/infrared counterpart of a source in the *XMM*–COSMOS sample, the closest *GALEX* source from the deblended, PSF-fitted *GALEX*–COSMOS catalog of Zamojski et al. (2007) was considered for inclusion in the photometry data set. Only 625 sources do have a *GALEX* counterpart.

IRAC photometry was also extracted from the IRAC–COSMOS catalog. In this case, the possible IRAC counterpart of each *XMM* source was taken to be the closest IRAC source within 1.″0 of the optical counterpart. For approximately 5% of the *XMM*–COSMOS sources, blending of multiple IRAC sources makes it impossible to reliably extract a correct IRAC flux. Additional 5% fall in masked regions in the IRAC images, and therefore lack mid-infrared photometry.

The 3.″8 aperture fluxes given in the COSMOS–IRAC catalog (aperture 2 in the catalog) were then translated to total fluxes by division by factors of 0.765, 0.740, 0.625, 0.580 for 3.6 μm, 4.5 μm, 5.8 μm, and 8 μm, respectively (see the readme file associated to the catalog and Surace et al. 2005 for details[25]).

Lastly, we multiply the total flux from *GALEX* and IRAC by 0.759, in order to rescale the photometry to the optical and near-infrared one. All photometry has been transformed to the AB magnitude system. The available bands, depth, and observing epochs are summarized in Table 1.

3. VARIABILITY ANALYSIS

AGNs exhibit nonperiodic flux variations on timescales of minutes to decades (see Wold et al. 2007, for a review). Thus, photometry collected nonsimultaneously will be affected by variability and may be not representative of a snapshot SED. Wolf et al. (2004) pointed out that ignoring variability can significantly increase the errors in photometric redshift estimations of AGNs.

The COSMOS optical photometry has been acquired in five epochs distributed over $\sim$ 6 years (see Column 7 in Table 1). Within each epoch (apart from 2005), several individual filter observations were distributed over less than three months, covering the whole optical range. The photometry in each filter has generally a very high calibration accuracy. This allows to study and correct AGN time variability over timescales of years, while shorter variability cannot be easily addressed.

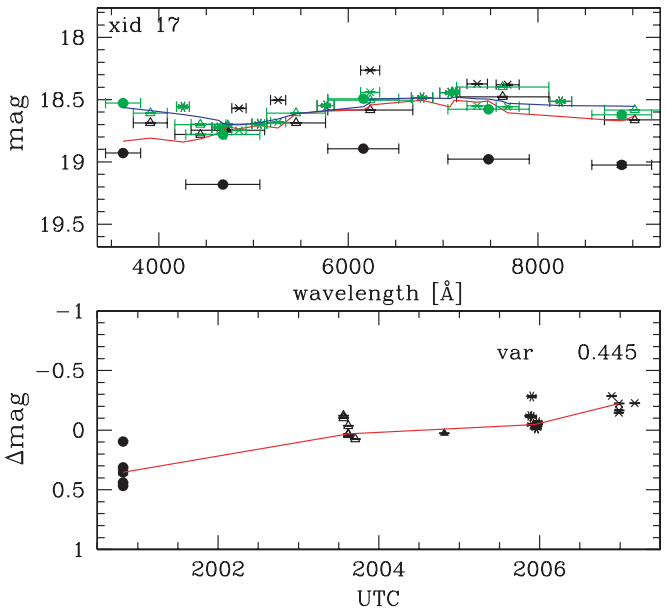

Figure 1. Time variability for the source with XID 17. Bottom: the source was observed in five epochs, each epoch marked with a different symbol. The five groups are, from left to right: SDSS, Subaru broad-band images (Subaru$_{BB}$), CFHT i band, the first set of intermediate bands Subaru images (Subaru$_{IB1}$), and the second epoch of intermediate bands Subaru images (Subaru$_{IB2}$). The optical photometry varied up to 0.6 mag over the six years of COSMOS data acquisition. The red line connects the median values for each group of observations. Top: the photometric points of the source before (black) and after (green) our correction. The red and blue lines represent the smoothed spectral shape of the object before and after the correction for variability (see text for more details).

Comparing the extracted fluxes at different epochs, a number of objects show clear time variability. As an example, in Figure 1 we show the variability of source XID 17 and the correction that we apply. In the upper panel, the raw optical photometry of this source is shown as black data points with different symbols for each epoch. There is obviously a large scatter between the photometry at different epochs. The time variability correction proceeded in the following way.

1. In order to flatten out the overall spectral shape to first order, the wavelength dependence of the optical magnitudes was simply smoothed by a running average filter. The result of this smoothing is shown as the red solid line in the upper panel of Figure 1.
2. This smoothed spectral shape was then subtracted from each individual data point in order to calculate the relative magnitude differences shown in the lower panel of Figure 1. Here the magnitude differences Δm are plotted against the observation date of the five separate epochs, n, of contemporaneous observations. For a constant source, all points in this panel should be within the measurement uncertainties around $\Delta m = 0$. In this particular example, the source XID 17 apparently brightened by more than 0.6 magnitude over the six years.
3. For the epochs $n = 1$, 2, 4, and 5, the median Δm_n of the magnitude differences of all optical bands observed in this epoch was calculated. For epoch $n = 3$, which consists of only one data point, Δm_3 was computed by a linear interpolation between $n = 2$ and $n = 4$.
4. The correction offset for a given epoch n is the difference, $\text{offset}_n = \Delta m_n - \Delta m_{n=4}$, where $n = 4$ is the epoch of the Subaru$_{IB1}$ and IRAC observations (December 2005–March

[25] Available at http://swire.ipac.caltech.edu/swire/astronomers/publications/SWIRE2_doc_083105.pdf

Table 1
Photometric Coverage and Depth of the COSMOS Data Set Used for the Photometric Redshift Estimation

Filter	Telescope	Effective λ (Å)	FWHM (Å)	zp_{corr} (mag_{AB})	Depth (mag_{AB})	Time of Observations (UTC)
u^*	CFHT	3911.0	538	0.05	26.50	Jan 2004–Apr 2007
B_J	Subaru	4439.6	807	−0.24	27.00	Jan 2004
V_J	Subaru	5448.9	935	−0.09	26.60	Feb 2004
g^+	Subaru	4728.3	1163	0.02	27.00	Jan 2005
r^+	Subaru	6231.8	1349	0.00	26.80	Jan 2004
i^+	Subaru	7629.1	1489	0.02	26.20	Jan 2004
z^+	Subaru	9021.6	955	−0.04	25.20	Jan 2004
J	UKIRT	12444.1	1558	0.12	23.70	Mar 2006
K_S	CHFT	21480.2	3250	−0.05	23.70	Mar 2007
i^*	CHFT	7629	1460	−0.01	24.00	Jan 2004
u	SDSS	3623	557	0.00	22.00	Jul 2001
g	SDSS	4677	1171	0.00	22.20	Jul 2001
r	SDSS	6159	1123	0.00	22.00	Jul 2001
i	SDSS	7480	1291	0.00	21.30	Jul 2001
z	SDSS	8884	943	0.00	20.50	Jul 2001
IA427	Subaru	4256.3	207	0.04	25.82	Jan 2006
IA464	Subaru	4633.3	218	0.01	25.65	Feb 2006
IA484	Subaru	4845.9	229	0.00	25.60	Jan 2007
IA505	Subaru	5060.7	231	0.00	25.55	Feb 2006
IA527	Subaru	5258.9	242	0.03	25.62	Jan 2007
IA574	Subaru	5762.1	272	0.08	25.61	Jan 2007
IA624	Subaru	6230.0	301	0.00	25.60	Dec 2006
IA679	Subaru	6778.8	336	−0.18	25.60	Feb 2006
IA709	Subaru	7070.7	316	−0.02	25.65	Feb 2007
IA738	Subaru	7358.7	324	0.02	25.60	Jan 2007
IA767	Subaru	7681.2	364	0.04	25.60	Mar 2007
IA827	Subaru	8240.9	344	−0.02	25.39	Jan 2006
IRAC1	*Spitzer*	35262.5	7412	0.00	23.90	Jan 2006
IRAC2	*Spitzer*	44606.7	10113	0.00	23.30	Jan 2006
IRAC3	*Spitzer*	56764.4	13499	0.01	21.30	Jan 2006
IRAC4	*Spitzer*	77030.1	28397	−0.17	21.00	Jan 2006
FUV	*GALEX*	1551.3	231	0.31	25.69	Feb 2004
NUV	*GALEX*	2306.5	789	−0.02	25.99	Feb 2004

2006). Each point in the plot represents the deviation of the photometry in a given filter observed in epoch n, from m_n.

5. In order to derive the closest approximation to a snapshot SED, we corrected for variability subtracting the value of $offset_n$ from each of the multiepoch photometry data points, resulting in the green data points in the upper panel of Figure 1.
6. Similar to the red curve, the thin blue solid line in the upper panel of Figure 1 shows the smoothed spectral shape of the variability corrected magnitudes.

In the case of source XID 17, the variability correction clearly resulted in a significant reduction of the scatter in the photometry, and thus an improvement in the photometric redshift determination. As discussed above, this correction scheme only works for slow time variability and there are examples of sources, where a significant time variability could not be reduced by this method. However, given the temporal sampling of the diverse data sets, it is the best that could be done and this correction was found to significantly reduce the number of outliers for the spectroscopic subsample (see Section 7).

As anticipated from the known variability timescales of AGNs, the *XMM*–COSMOS sources exhibited a variety of behaviors—some sources evolved gradually, while others displayed more extreme outbursts. We quantify the observed variability with the parameter,

$$\mathrm{VAR} = \sqrt{\sum_{n=1,2,3,5} (\Delta m_n - \Delta m_4)^2}. \quad (1)$$

In Figure 2, the distribution of VAR for all the sources (black solid line) is shown.

4. FITTING TECHNIQUE

To derive the intrinsic SEDs and redshifts of the observed sources, we used the publicly available LePhare code (Arnouts, Ilbert).[26] For this fitting, a library of expected intrinsic rest-frame SEDs of the source population is supplied and a χ^2 minimization is used to solve for the most likely redshift, SED type and intrinsic extinction. A user-supplied prior can also be used. In the following, we describe the transformation of the template SEDs to the observed photometric bands, the selection of template SEDs, the adopted extinction law, and the control of source variability which is important for AGNs.

The accuracy of the photometric redshift fitting was evaluated using a subsample of 396 *XMM* sources with $i^* \leqslant 22.5$ and secure spectroscopic redshift. 90 sources had optical spectroscopy from SDSS (York et al. 2000; Adelman-McCarthy et al. 2008),

[26] http://www.oamp.fr/people/arnouts/LE_PHARE.html

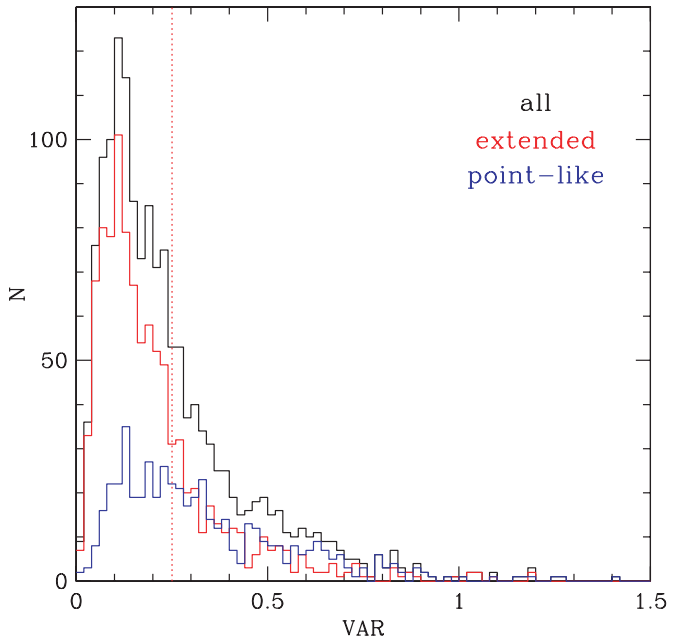

Figure 2. Histogram of the parameter VAR for all the sources (black solid line) extended (red) and point-like sources (blue). The vertical dotted line at VAR = 0.25 defines our adopted separation between host- (low VAR) and AGN- (high VAR) dominated sources.

while 306 were observed within the COSMOS collaboration (235 observed with IMACS and MMT; Trump et al. 2007, 2009; Prescott et al. 2006; 71 observed by *z*COSMOS with VIMOS; Lilly et al. 2007). The spectroscopic sample was used to train/evaluate the fitting procedure and the accuracy quantified as $\sigma_{\Delta z/(1+z_{\rm spec})} = 1.48 \times {\rm median}|z_{\rm phot} - z_{\rm spec}|/(1 + z_{\rm spec})$ (Hoaglin et al. 1983), where $z_{\rm phot}$ and $z_{\rm spec}$ are the photometric and spectroscopic redshifts, respectively. As the reliability of photometric redshifts trained on local/bright spectroscopic samples may not hold for low-luminosity or high-redshift sources for which the SED templates may no longer be representative, we have also performed an a posteriori test using spectroscopy which became available only recently for 46 optically faint, *K*-band, and *Spitzer*/MIPS 24 μm selected sources (S. Lilly et al. 2008, in preparation; J. Kartaltepe et al. 2009, in preparation). These additional spectroscopic redshifts indicated a high success rate, with correctly recovered redshift also for these faint sources (see Section 5.5).

4.1. Transformation of Template SEDs to Observed Photometric System

To reduce the spectra template to the same photometric system of the observed sources, all SEDs were convolved with the filter transmission curves, which also include the instrument efficiency. The same zero points derived for the new COSMOS photometric redshift catalog of non-X-ray-detected galaxies (Ilbert et al. 2009) were used for the *XMM* sources. This zero-point calibration is extensively explained in Ilbert et al. (2006) and reported in Column 5 of Table 1. The convolved template SEDs were then redshifted in $\Delta z = 0.02$ wide bins from $z = 0.02$ to $z = 8$ and parabolically interpolated for higher redshift accuracy. The opacity of the intergalactic medium was taken into account as described by Madau (1995).

4.2. Galaxy and AGN Templates

Ideally, a template library should represent all possible SEDs. However, the larger the number of templates, the higher the risk of degeneracy, which can then lead to an increasing number of catastrophic failures. To identify a suitable library, we run LePhare on the spectroscopic sample of 396 sources with $i^*_{AB} < 22.5$. Here, we fixed the redshifts to the spectroscopic values and let the code select the best-fitting SED template. Initially, the entire set of 28 templates from SWIRE[27] was used, including elliptical and spiral galaxies, starbursts, ULIRG, low-luminosity AGNs and QSOs (QSO1, QSO2, TQSO1, BQSO1, etc.; see Polletta et al. 2007 for details on these templates). We also added two composite low- and high-luminosity QSO templates from SDSS and one template of a blue, star-forming galaxy from the library used by Ilbert et al. (2009), for a total of 31 templates.

First, we trimmed this list by discarding all templates that were not selected for any of the 396 sources in the spectroscopic sample. Next, we selected for each galaxy type the template which was often most used. Accordingly, the S0 and the I22491 templates were singled out to represent passive galaxies and starburst/ULIRGs, respectively. This process reduced the original number of SEDs from 31 to 12.

For a large number of objects, significant discrepancies between the photometry and templates occurred longward of

Table 2
Final Template Library Used in This Work

ID	Name[a]	Type
1	BC03	Blue starforming[b]
2	S0	S0[c]
3	Sb	Spiral b[c]
4	Spi4	Spiral c[c]
5	M82	Starburst[c]
6	I22491	Starburst/ULIRG[c]
7	Sey18	Seyfert 1[c].8
8	Sey2	Seyfert 2[c]
9	S0-90_QSO2-10	Hybrid[d]
10	S0-80_QSO2-20	”
11	S0-70_QSO2-30	”
12	S0-60_QSO2-40	”
13	S0-50_QSO2-50	”
14	So-40_QSO2-60	”
15	S0-30_QSO2-70	”
16	S0-20_QSO2-80	”
17	S0-10_QSO2-90	”
18	Mrk231	Seyfert 1, BALQSO[c]
19	I22491-90_TQSO1-10	Hybrid[d]
20	I22491-80_TQSO1-20	”
21	I22491-70_TQSO1-30	”
22	I22491-60_TQSO1-40	”
23	I22491-50_TQSO1-50	”
24	I22491-40_TQSO1-60	”
25	pl_I22491-30_TQSO1-70	”
26	pl_I22491-20_TQSO1-80	”
27	pl_I22491-10_TQSO1-90	”
28	pl_QSOH	QSO high lum.[e]
29	pl_QSO	QSO low lum.[f]
30	pl_TQSO1	QSO high IR lum.[c]

Notes.
[a] Names are as in the original libraries.
[b] Ilbert et al. (2009).
[c] Polletta et al. (2007).
[d] This work.
[e] composite #33 at http://www.sdss.org/dr5/algorithms/spectemplates.
[f] composite #30 at http://www.sdss.org/dr5/algorithms/spectemplates.

[27] http://www.iasf-milano.inaf.it/ polletta/templates/swire_templates.html

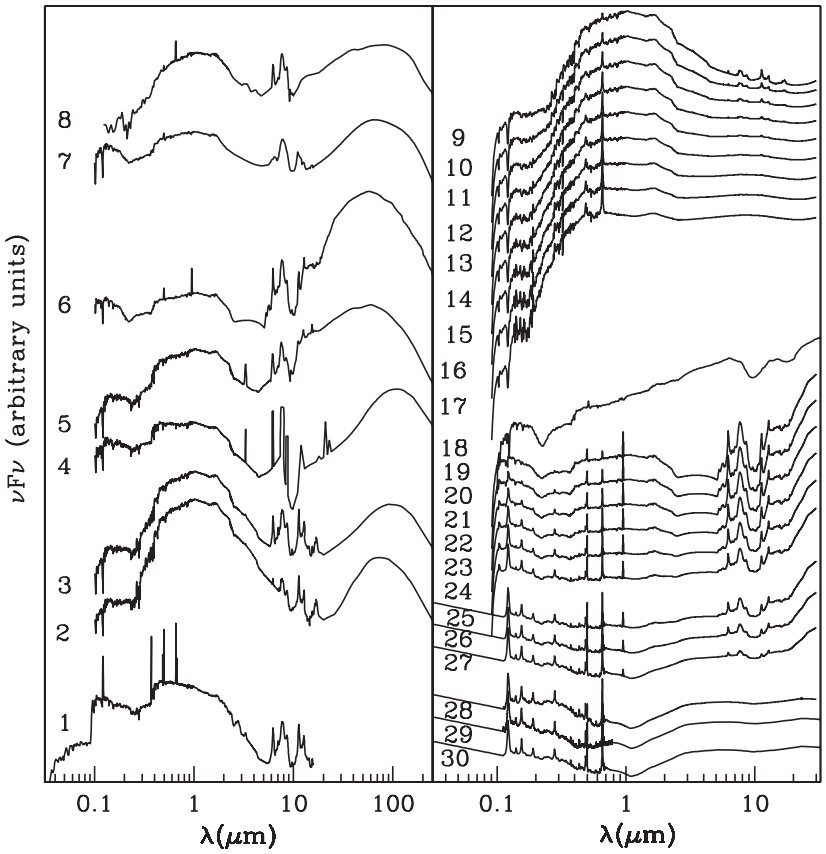

Figure 3. Visualization of the final choice of templates used in this paper. Numbers correspond to the IDs in Table 2.

1 μm. This suggests that the SEDs are not properly representing the mid-infrared emission component. The SED templates are from spectra of nearby galaxies where the nuclear flux is well sampled, but the host galaxy is not fully included in slits or fibers. In contrast, for all *XMM* sources at intermediate/high redshift, their host galaxy sizes are comparable to seeing the achievable with ground-based optical telescopes and our photometric data should include a larger fraction of the host galaxy emission (Szokoly et al. 2004). Thus, for the *XMM* sources analyzed here, the SED templates should very likely have a larger contribution from the host galaxy relative to the nuclear emission component.

To better account for the host galaxy component, for each pair of AGN (type 1, type 2, QSO1, QSO2) and host galaxy type (elliptical, spiral, starburst), hybrid templates were constructed. Here, a pair of SEDs was normalized by the integrated fluxes in different passbands. For each normalization, nine hybrids with varying ratios (90:10, 80:20,. . ., 20:80, 10:90) between both components were built. After recalculating the photometric redshifts using these hybrid SEDs, two sets of templates provided the smallest $\sigma_{\Delta z/(1+z_{\rm spec})}$ and the number of outliers, η (defined as percentage of objects for which $|z_{\rm phot} - z_{\rm spec}| > 0.15(1 + z_{\rm spec})$. These sets of hybrids are as follows.

1. Starburst/ULIRG I22491 plus TQSO1 templates, both normalized between 5000 Å and 5200 Å (models 19–27 in Table 2 and in Figure 3).
2. S0 plus QSO2 templates, both normalized between 0.9 μm and 1 μm (models 9–17 in Table 2 and in Figure 3).

Lastly, a number of sources showed discrepancies between the observed and the predicted flux in the UV. This is likely the result of intrinsic absorption in the empirical templates. Thus, we extended the QSO-dominated templates 25–30 (templates with prefix "pl") into the UV using a power law component with spectral index $\alpha = -0.56$ (Scott et al. 2004). Our final library includes 30 SEDs.

4.3. Intrinsic Extinction

To allow for the intrinsic reddening, two extinction laws were tested: the SMC extinction derived by Prevot et al. (1984) and the starburst galaxy extinction law obtained by Calzetti et al. (2000). We explored $E(B - V)$ values from 0 to 0.5 in steps of 0.05. The two laws provide a similar photometric redshift accuracy ($\sigma_{\Delta z/(1+z_{\rm spec})} = 0.014$), but with the SMC extinction law, the number of outliers is reduced by 1.1%. In principle, highly obscured AGNs might have $A_V > 1$. However, none of the fits required $E(B - V) > 0.4$ and 66% of the sources were fit with $E(B - V) \leqslant 0.05$. Less than 20 sources required $E(B - V) > 0.35$. This result is perhaps not totally surprising, since many of our templates are empirical and thus already include some extinction.

4.4. Luminosity Prior

Applying a luminosity prior during the SED fitting can be important to avoid unlikely combinations of source type, luminosity, and redshift which can lead to wrong solutions for the photometric redshift. Quasars are usually defined as point-like sources which have an absolute magnitude of $M_B \leqslant -23$. For lower luminosity AGNs, the host galaxy contributes to the observed flux and the object may be classified as extended in good seeing data or, even more, in *Hubble Space Telescope* (*HST*) data. In the spectroscopic sample, we did not find any point-like source with an absolute magnitude M_B fainter than -20, and thus we adopted this value as the lower limit on the

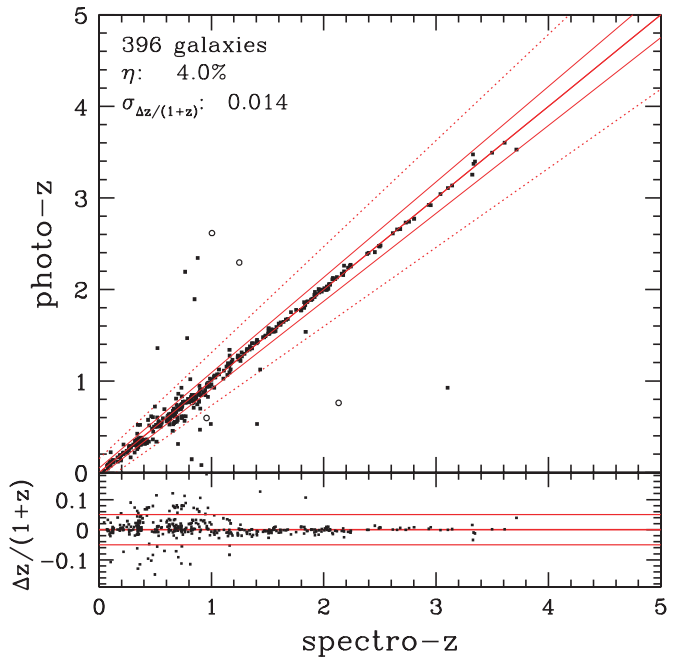

Figure 4. Comparison of photometric and spectroscopic redshifts for the 396 sources with $i^*_{AB} < 22.5$. Empty circles represent sources for which the second peak in the redshift probability distribution agrees with the spectroscopic redshift. The lines correspond to (1) $z_{\rm phot} = z_{\rm spec}$ (thick solid), (2) $z_{\rm phot} = z_{\rm spec} \pm 0.05(1 + z_{\rm spec})$ (solid) and (3) $z_{\rm phot} = z_{\rm spec} \pm 0.15(1 + z_{\rm spec})$ (dotted).

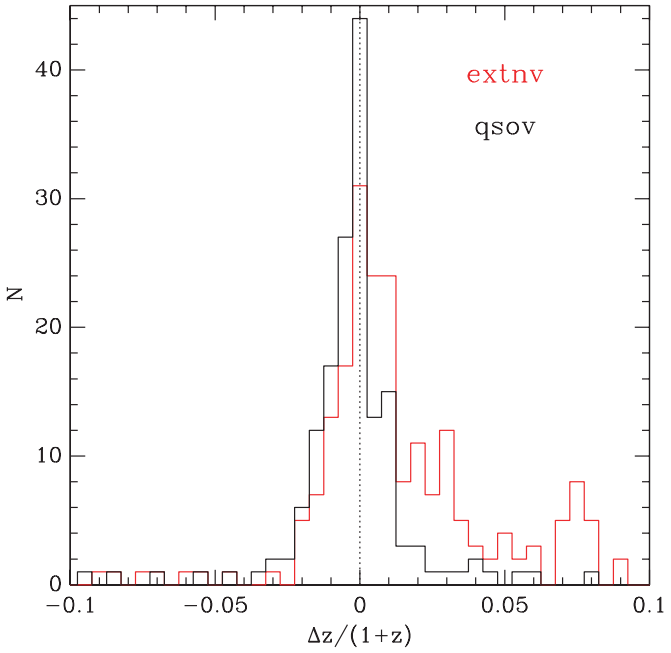

Figure 5. Histogram of $\Delta z/(1 + z)$ for the "extnv" (red) and "qsov" (black) samples (see text for details).

luminosity prior. While it is known that low-luminosity AGNs may have M_B fainter than our limit, to use a lower threshold of $M_B = -19$ (as used by Polletta et al. 2007 and Rowan-Robinson et al. 2008) would increase the degeneracy between the galaxy and AGN templates, and thus the outliers from 6.3% to 8.0%.

Type 1 AGNs and QSOs can be identified first by their point-like appearance in the image or by the amount of variability as defined in Section 3. For the sources detected in the ACS/FW814 filter and brighter than $\mathrm{FW814}_{AB} = 24$, the peak surface brightness is measured and used to separate extended and point-like sources (Leauthaud et al. 2007). For the 500 sources missing this parameter, the star/galaxy separation computed on the best-seeing Subaru *R*-band image has been used to separate extended and point-like sources. The subsample of sources with an extended optical counterpart includes 989 sources, while the subsample of point-like sources includes 553 members.

We also applied the prior in luminosity to sources with high variability (Section 3). For extended sources, the variability parameter VAR can additionally be used to disentangle between host- and AGN-dominated sources. As expected, the majority of the extended sources (Figure 2, red distribution) exhibit only low levels of variability (VAR < 0.25). However, a tail of higher VAR is present for the low redshift type 1 AGN, where apparently both the host galaxy and the variability of the nucleus are detected. In order to identify the best threshold, we run the photometric redshift code using a wide range of VAR parameters above which the luminosity prior was applied also to extended objects. The prior is always used for point-like sources. The smallest number of outliers was achieved using a separation at VAR = 0.25. Accordingly, we found 252 extended sources in the *XMM*–COSMOS sample with VAR > 0.25.

In the following, we will refer to the point-like and the variable sources (805) as the "qsov" sample. To this sample we applied the luminosity prior. All the remaining extended sources (737) with VAR $\leqslant$ 0.25 form the "extnv" sample. No prior is applied to the latter sample.

5. RESULTS

5.1. Photometric Redshift Accuracy

Following the procedure described in the previous sections, we obtained photometric redshifts for all 1542 optical counterparts of the *XMM*–COSMOS sample described in Section 2. For the 396 bright spectroscopically confirmed galaxies, we obtained a final accuracy of $\sigma_{\Delta z/(1+z_{\rm spec})} = 0.014$ and $\eta = 4.0\%$ outliers (see Figure 4). Considering the "qsov" and "extnv" samples separately, the accuracies are $\sigma_{\Delta z/(1+z_{\rm spec})} = 0.012$ and $\sigma_{\Delta z/(1+z_{\rm spec})} = 0.019$, while the outliers are $\eta = 6.3$ % and $\eta = 2.3\%$, respectively (see the first row of Table 4).

Figure 5 shows the distribution of $(z_{\rm spec} - z_{\rm phot})/(1 + z_{\rm spec})$ for the "extnv" and "qsov" samples. The histogram for the "extnv" sample is broader, not perfectly centered at 0, and with a small asymmetric tail toward positive Δz (i.e., $z_{\rm phot} < z_{\rm spec}$). We found that the scatter could be reduced for the "extnv" sample by adding additional normal galaxy SED templates, but this also increases the degeneracy and the number of outliers.

We were also able to characterize the nature of the small number of outliers and understand the reasons why the redshifts were not correctly assigned. First, six out of the 16 outliers (37.5%) are blended with a nearby object, closer than 1.″5, while the spectroscopic follow-up was centered on the right counterpart, identified on our ACS images. In the entire *XMM*–COSMOS catalog, we flagged 34 cases as potentially problematic, due to the presence of a close source. Additionally, two of the remaining outliers miss photometry in the IRAC band (i.e., the source is either partially or completely in a masked area) and the remaining optical bands were not sufficient to find the correct redshift solutions. Lastly, we found that for strongly varying sources (VAR > 0.25), the available u^* fluxes were likely the problem—the adopted u^* magnitudes are an average between four runs of observations taken yearly between 2004 and 2007.

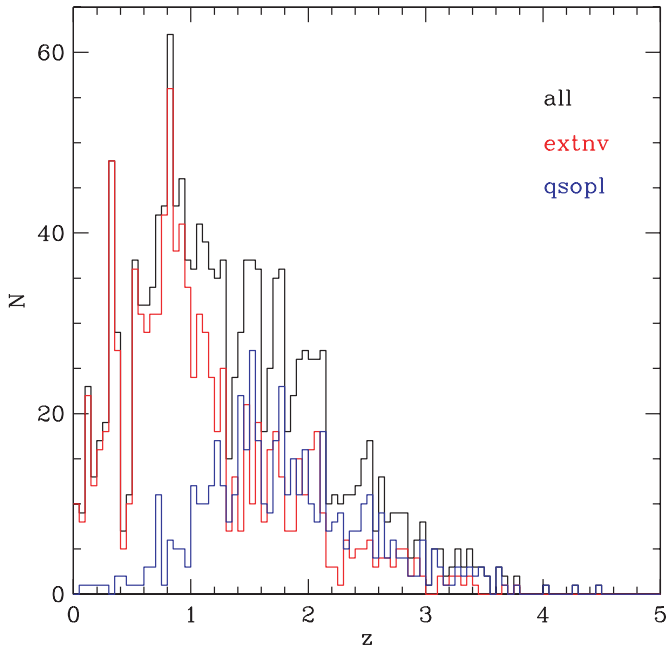

Figure 6. Photometric redshift distribution for the entire sample (black line), "extnv" (red) and "qsov" (blue) samples.

Adding 0.2 mag in quadrature to the measurement uncertainty associated with the u^*-band photometry for the outliers, three of the remaining eight outliers would now have the correct redshift due to the reduced weighting of the u^* band; an additional one would have the second comparable but lower peak of the redshift probability distribution at the right redshift. We decided against systematically reducing the weighting of the u^* band for the entire sample, as this would degrade the quality of the final result.

5.2. Redshift and Templates Distribution

Figure 6 shows the photometric redshift distribution for the complete *XMM* sample. As expected, the large majority of the sources at $z < 1$ are extended and not strongly time variable, while most of the point-like and varying sources are at $z > 1$. The redshift distribution shows significant peaks at $z \sim 0.35$, ~ 0.8, corresponding to spectroscopically confirmed galaxy overdensities in the COSMOS field (Scoville et al. 2007a; Finoguenov et al. 2007; Lilly et al. 2007; Gilli et al. 2009). This suggests that the clustering of the normal galaxies and the AGNs follow each other.

The distribution of the selected SED templates as a function of redshift is shown in Figure 7. Templates of normal galaxies (templates 1–6) or type 2 low-luminosity AGNs (e.g., Seyfert 1.8, template 7) are most frequent at $z < 2$. At higher redshift, mainly QSO templates (templates 28–30) have been selected. For some of the lower redshift galaxies, the detected X-ray emission may, in fact, arise from star-forming regions rather than nuclear activity, while at higher redshifts, the X-ray emission is predominantly from AGNs.

In Section 5.4, we show that the best-fit SED templates are generally consistent with other properties of the host galaxies. We note that Figure 7 clearly reveals that templates 14–17, with an increasing contribution of the QSO2 SED, were rarely selected. However, running the photometric redshift fitting without these templates resulted in a significant increase in outliers, indicating that for a few sources these templates are essential. In summary, we classify 464 sources as best fit by a nonactive galaxy (templates 1–6), while 1032 are best fit with a template which includes an AGN component (templates 7–30).

5.3. Optically Undetected Sample

As discussed in Section 2, *XMM*–COSMOS sources are not detected in our optical broad or intermediate filters. However, unique counterparts for these sources have been identified in the *K* band and in the IRAC catalogs. These are potentially very

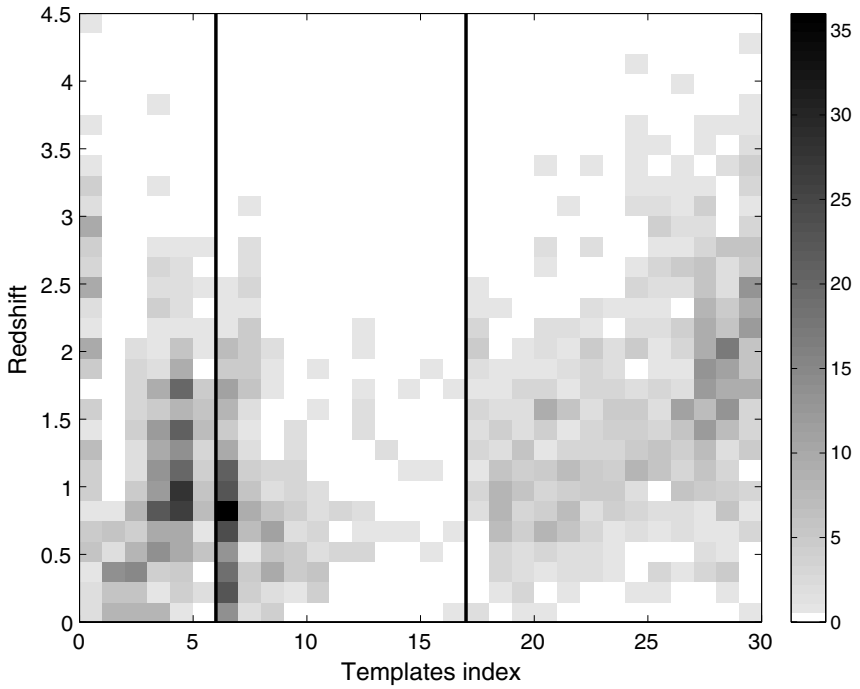

Figure 7. Density map of SED template distribution as a function of photometric redshift.

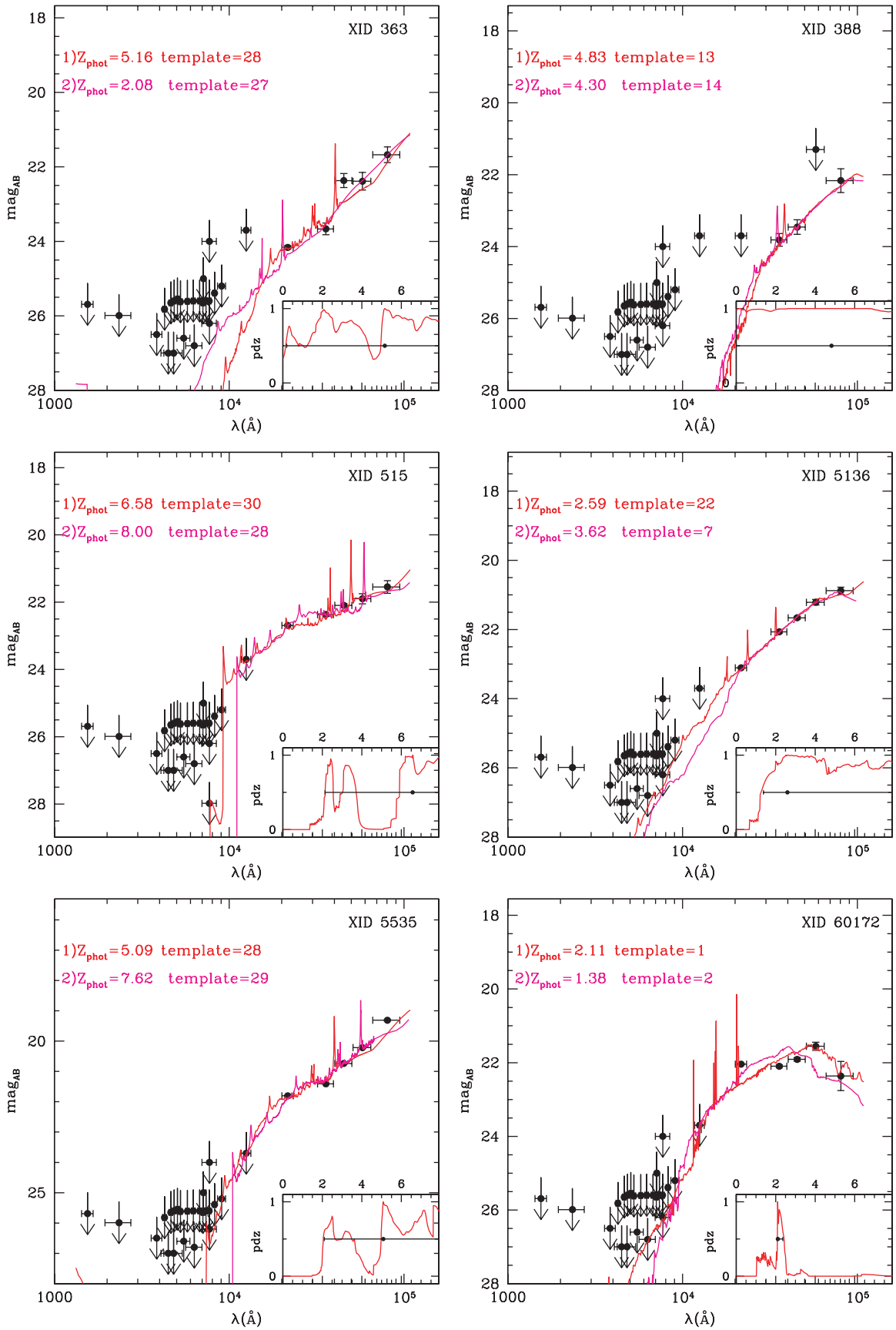

Figure 8. SED fitting and photometric redshift estimation for the isolated six *K*-band-detected sources for which only five (or less) photometric points are available. The redshift probability distribution is shown in the inserts. Although the 1σ errors (black horizontal bars in the insert) are so large that a solution at $z \sim 2$ cannot be excluded, the best formal fit redshift is greater than 4.

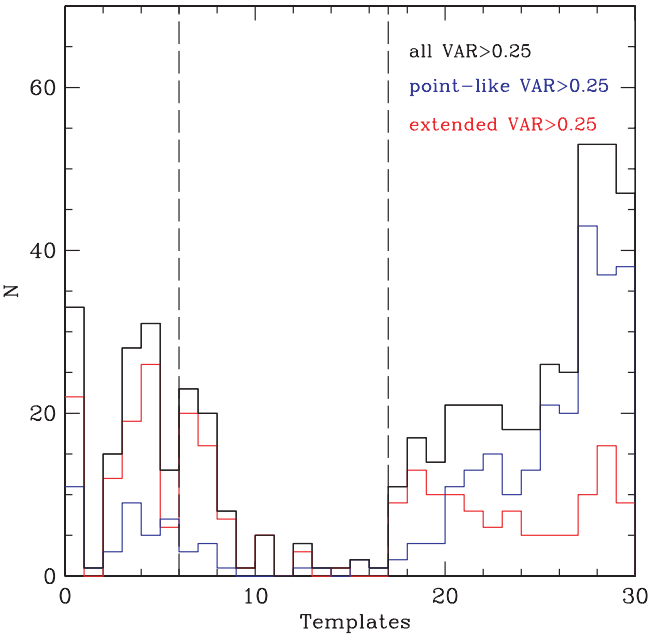

Figure 9. Distribution of the best-fit SED templates for all variable sources (black), for the point-like (blue) and the extended (red) subsamples. The vertical lines indicate the separation between templates of nonactive galaxies (templates 1–6), type 2 AGNs or with a QSO2 component (templates 7–17), and type 1 AGNs, QSOs or with a QSO component (templates 18–30). The bulk of the sources is best fit with a template of the third group.

high redshift sources ($z > 3$; Koekemoer et al. 2007; Brusa et al. 2009), as suggested for similar MIPS 24 selected objects (Houck et al. 2005; Weedman et al. 2006b; Yan et al. 2007; Sajina et al. 2007). However, doubts on this interpretation have recently been presented by Brand et al. (2008), who observed 16 X-ray-loud, optically faint sources with *Spitzer*/IRS and measured redshifts $z \sim 2$ for all targets.

Our redshift estimates for six of the 10 sources in the *XMM*–COSMOS optically faint sample is shown in Figure 8. We do not report the results for the remaining four sources, as strong contamination from a nearby object does not allow accurate photometry. In four of the six cases, the formal best-fit redshift is higher than 4, but the redshift probability distribution function (see insert in each figure) indicates that there are not enough constraints to reject a solution at lower redshift. In particular, $z \sim 2.0$–2.5 would be acceptable for all of them.

5.4. Reliability of the Best-Fit SED Templates

It is reasonable to ask if the selected SED templates are consistent with other properties of the sources as revealed in the COSMOS multiwavelength data set. For example, (1) are the highly varying sources mostly fit with a template, including a type 1 AGN contribution? and (2) were the sources likely to be "host-dominated" on the basis of their low X-ray luminosity better fit by templates with little overall AGN contribution?

Figure 9 clearly shows that most (∼62%) of the strongly varying sources are best fit with type 1 AGN or QSO-dominated hybrid SEDs, while ∼ 25% are best fit with one of the nonactive galaxies. Similar reassurance is found in the correlation of type 1, type 2, and normal galaxies, with X-ray properties like luminosity and hardness ratio. For example, Figure 10 shows the distribution of variable (right panel) and nonvariable (left panel) sources in the *XMM*–color plane. Type 1 AGNs are expected to be in a narrow region with HR1 ∼ –0.5, while type 2 AGNs and normal galaxies have higher HR1 values (Hasinger et al. 2001; Mainieri et al. 2002; Della Ceca et al. 2004; Hasinger et al. 2007; Cappelluti et al. 2009). Our classification is well consistent with the distribution on the *XMM* sources in the X-ray hardness ratio plan (see caption of Figure 10). Note that for some sources, the X-ray flux is obtained by stacking up to nine pointings taken over three years. Thus, the determination of the HR values can be effected by spectral variability.

The 2–10 keV X-ray luminosity is commonly used to distinguish AGNs ($L_X > 10^{42}$ erg s^{-1}) from starburst or normal galaxies ($L_X < 10^{42}$ erg s^{-1}). As we show in Figure 11, the large majority of low-luminosity sources are nonvariable and a host-dominated template was selected for most of them. A large fraction of high-luminosity objects is instead varying and the best fit is obtained with type 1 and QSO SEDs. A number of sources with ($L_X > 10^{42}$ erg s^{-1}) have VAR < 0.25 and are fit by a normal galaxy template. This is consistent with cases in which the optical nuclear activity is completely hidden by the host galaxy.

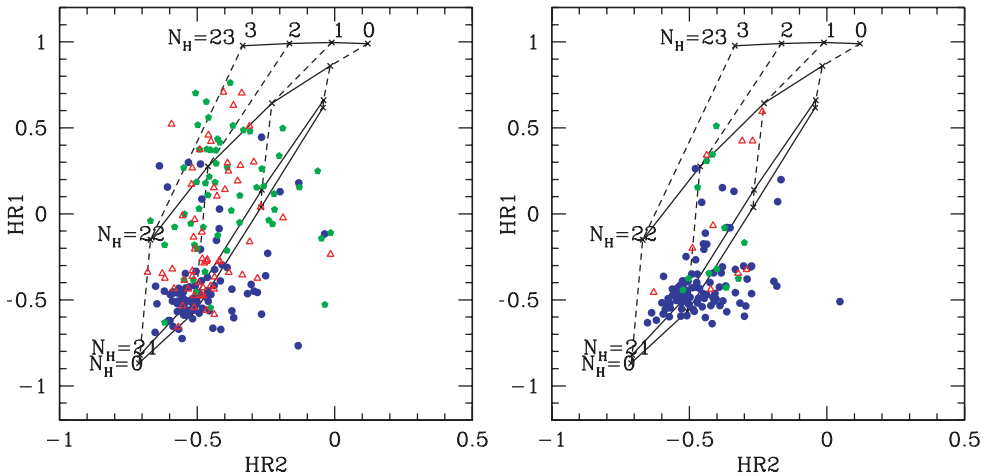

Figure 10. Left: X-ray hardness ratios of all *XMM*–COSMOS nonvarying sources (VAR < 0.25) and with ΔHR1 < 0.25 and ΔHR2 < 0.25. The hardness ratios are defined as HR1 = $(F_{(2-10\,\mathrm{keV})} - F_{(0.5-2\,\mathrm{keV})})/(F_{(2-10\,\mathrm{keV})} + F_{(0.5-2\,\mathrm{keV})})$ and HR2 = $(F_{(5-10\,\mathrm{keV})} - F_{(2-10\,\mathrm{keV})})/(F_{(5-10\,\mathrm{keV})} + F_{(2-10\,\mathrm{keV})})$. The best-fit templates are color-coded as follows: blue filled circles (Type 1 and QSO1 hybrids: templates 18–30), red open triangles (Type 2 and QSO2 hybrids: templates 7–17) and green open circles (normal galaxies: templates 1–6). The solid lines represent lines of constant N_H and the dotted lines are lines of equal photon index. Right: same as left, for sources with VAR > 0.25.

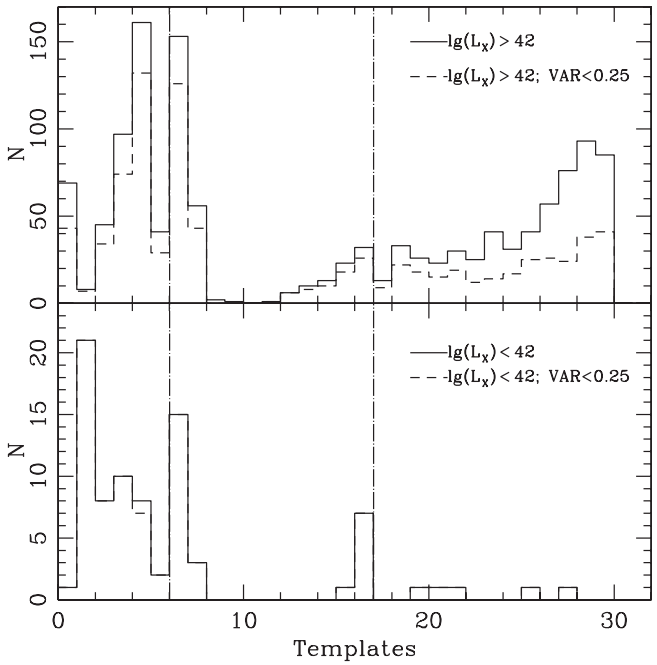

Figure 11. SED template distribution for all sources split on the basis of their X-ray luminosity (bottom: $L_X < 10^{42}$ erg s^{-1}, top: $L_X > 10^{42}$ erg s^{-1}) and variability (solid line: all, dashed: VAR < 0.25). Practically all sources with $L_X < 10^{42}$ erg s^{-1} are nonvariable (bottom panel).

5.5. Reliability of the Photometric Redshift for Faint Sources

Recently, the spectroscopic redshift for 46 faint sources $22.5 < i^*_{AB} < 24.5$ (average value $i^*_{AB} \sim 23.3$) became available. They are part of a Keck II spectroscopic follow-up of 24 μm selected sources (J. Kartaltepe et al. 2009, in preparation) and part of *z*COSMOS-*faint* (S. Lilly et al. 2008, in preparation). Figure 12 (left panel) shows the comparison between the photometric and spectroscopic redshifts for this sample. As these sources are faint, the photometric uncertainties are larger and we should expect lower photometric redshift accuracy. In fact, the accuracy becomes worse (although still comparable to the typical accuracy of photometric redshift for nonactive galaxies) by a factor of 1.5, from 0.014 to 0.023, while the number of outliers increases from 4.0% (for the bright sample) to 10.9%. However, two of the five outliers (open circles) do have a secondary maximum in the redshift probability distribution at the spectroscopic value.

As the reliability of the photometric redshift based on the comparison with a spectroscopic sample is limited to the redshift and magnitude ranges of the sample, we used the 1σ uncertainty (as derived from the redshift probability distribution function) to test the accuracy of the photometric redshift for all the *XMM*–COSMOS sources. Figure 12 (right panel) shows the 1σ uncertainties in the photometric redshift as a function of magnitude for sources with $0.2 < z < 1.25$ and $1.25 < z < 5$. Unlike normal galaxies (see, for example, Figure 9 in Ilbert et al. 2009), the uncertainties for the two redshift ranges behave in the same way and up to $i^+_{AB} \sim 24$ remain almost constant, of the order of 0.02. Only beyond $i^+_{AB} > 24$, the uncertainties increase and reach an amplitude of 0.2 at $i^+_{AB} \sim 26$. This is expected as AGNs show prominent emission lines, all over the spectrum. In particular, besides the well-known optical lines (H_{betza}, [O III], H_{alpha}, etc.), two strong emission lines in the blue part of the spectrum (C IV[1549 Å] and Mg II[2798 Å]) are available at a high redshift. The COSMOS survey, dense with intermediate band photometry, allows us to identify these lines and thus provides a reliable photometric redshift estimate over the full redshift range. The reliability of our photometric redshift is diminished only by the larger photometric errors at faint magnitudes.

5.6. Reliability of Star/Galaxy Separation

In addition to galaxy and AGN templates, stellar templates have been used to identify stars in the *XMM*–COSMOS sample. We constructed a spectral library of low-mass stars (Chabrier et al. 2000) and subdwarf O and B stars, white dwarfs, and binary systems (Bixler et al. 1991). The reliability of the star/galaxy separation was tested on the subsample with spectroscopic identification. We defined as star each point-like source with

$$\chi^2_{\rm star} \leqslant \chi^2_{\rm qso}, \quad (2)$$

and the remaining as extragalactic. In this way we classify 46 sources as stars, 28 of which are already spectroscopically confirmed. In summary, 439 (out of 396 + 46 = 442) galaxies and 24 (out of 28) stars were classified correctly. The completeness in identifying galaxies is 439/442 = 0.99 and 24/28 = 0.86 for the stars. Two of the three misclassified galaxies were missing IRAC photometry, and thus the possibility to disentangle between star and galaxies was limited.

6. RELEASE OF THE PHOTOMETRIC REDSHIFT CATALOG

An electronic version of the complete photometric redshift and classification catalog of the *XMM*–COSMOS sample is available. For each source, the coordinates, the photometric redshift, and the classification in star/galaxy is provided. A more comprehensive catalog is available on request.[28]

7. DISCUSSION

In this section, we analyze the impact of each contribution (new template library, correction for variability, photometric coverage) on our final result. We show that while the number of outliers can be decreased with the appropriate SED template library and the correction for variability, a good photometric coverage is essential in order to improve the accuracy of the photometric redshifts.

7.1. Relevance of SED Templates Library

An X-ray-selected sample of sources consists either of low-redshift star-forming galaxies or AGNs and QSOs, and the choice of templates must represent these populations. To verify this, we first compared our photometric redshifts with those obtained for the same sources by Ilbert et al. (2009) using the same photometric data set, but only normal galaxy SED templates. The result is shown in the first row of Table 3. For the "etxnv" sample, where the sources are galaxy-dominated, the library of Ilbert et al. (2009) provides an excellent accuracy, although the number of outliers is large. For the "qsov" sample dominated by AGNs and QSOs, the lack of appropriate templates leads to an expected failure.[29]

[28] ms@astro.caltech.edu

[29] In Ilbert et al. (2009), the new publicly available COSMOS photometric redshift catalog is presented. This catalog supersedes all previous ones (Mobasher et al. 2007; Ilbert et al. 2006) and is composed of 600,000 sources with $i^+_{AB} < 26$ and not detected in X-rays. In order to decrease the risk of degeneracies for this sample of normal galaxies, no AGN/QSO templates were used to compute the photometric redshifts.

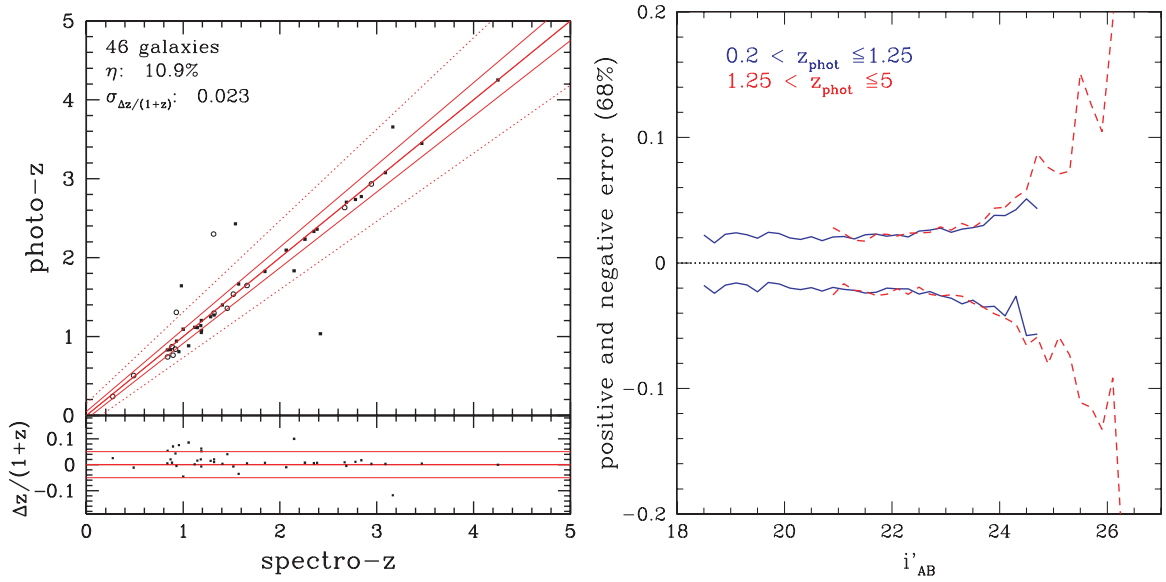

Figure 12. Left: same as Figure 4 for the 46 sources with $i^{+}_{AB} > 22.5$ with spectroscopic redshifts, which has been used for an a posteriori test. Right: 1σ error for the $z_{\rm phot}$ estimate as a function of the apparent magnitude in the redshift range $0.2 < z_{\rm phot} \leqslant 1.25$ (blue) and $1.25 < z_{\rm phot} \leqslant 5$ (red).

Table 3
Relevance of Variability Correction and Template Library

	extnv		qsov	
	$\sigma_{\Delta z/(1+z)}$	Outliers (%)	$\sigma_{\Delta z/(1+z)}$	Outliers (%)
Ilbert09[a]	0.014	14.0	0.419	71.4
Polletta07[b] + var. corr.	0.020	12.8	0.014	15.0
Salvato09[c], no var. corr.	0.017	4.1	0.013	11.6
Salvato09[c] + var.corr.	0.019	2.3	0.012	6.3

Notes.
[a] Library used in Ilbert et al. (2009).
[b] Library used in Polletta et al. (2007).
[c] Library used in this work.

The addition of AGN and QSO templates in the library does not automatically imply an improvement in the photometric redshifts. We verified this by computing the photometric redshift using the library of Polletta et al. (2007) which includes normal galaxies, AGNs, and QSOs. The result of the test is shown in the second row of Table 3. In this case, the accuracy is very good and comparable with the one obtained with our library. However, the number of outliers remain high. We then conclude that our library, constructed as described in Section 4.2, can be considered as representative of a typical population of X-ray-emitting galaxies.

7.2. Relevance of Correction for Variability

In the previous section, we have shown the importance of the library to increasing the accuracy and reducing the number of outliers. The comparison has been performed on the photometric data set corrected for variability; thus, it is natural to ask how much the improved performances shown by our library are depending on it. In the third row of Table 3, our library has been used on the photometric data set without any correction for variability. While the accuracy of the photometric redshift did not change, the number of outliers is increased by a factor of 2 in both "extnv" and "qsov" subsamples, demonstrating the importance of the correction for variability.

7.3. Relevance of Spectral Coverage

To quantify the benefits of spectral coverage, we split the available photometric data into five subsets (*GALEX*, optical broad band, optical intermediate band, near-infrared, and IRAC). Next, we derived the photometric redshifts for the "extnv" and "qsov" samples using different combinations of these subsets and compared the results in terms of accuracy and number of outliers. We find that for all the combinations the accuracy is $\sigma_{\Delta z/(1+z_{\rm spec})} < 0.07$, which is good enough for most AGN studies. A summary is given in Table 4.

By far, the best result for the "extnv" sample was achieved using the entire spectral coverage. This is not surprising and is also discussed for normal galaxies in COSMOS (Ilbert et al. 2009). The 30-band photometry of COSMOS essentially resembles a low-resolution spectrum and a representative SED template library can be fitted with high reliability.

Table 4 also shows the importance of the intermediate optical filters that allow the precise localization of spectral features (e.g., emission and absorption lines, Balmer break). The elimination of these filters produces a significantly worse accuracy (by a factor of 3 for both the "qsov" and the "extnv" samples). At the same time, the number of outliers increases by a factor of ~ 3 for both samples.

In contrast, the removal of the broad optical bands photometry when the intermediate bands are available, does not change significantly the accuracy of the photometric redshift, neither for the "qsov" nor for the "extnv" samples. However, in both cases the number of outliers increases dramatically. The same is happening when the IRAC photometry is not considered for the "extnv" sample. As already pointed out by Rowan-Robinson et al. (2008), the presence of IRAC photometry is crucial in suppressing the number of outliers.

8. CONCLUSIONS

In this paper, we presented the photometric redshifts and classifications for 1542 *XMM*–COSMOS sources with optical counterparts. Using SED matching to up to 30 photometric bands, we identified 46 Galactic foreground stars, 464 sources best fit with a nonactive galaxies template, and 1032 best

Table 4
Relevance of Spectral Coverage and Photometric Resolution for the Variability Corrected Bright Spectroscopic Sample

Sample	*GALEX*	Optical	Optical	NIR	IRAC	Reliability	
		Broadband	IB and NB			Outliers(%)	$\sigma_{\Delta z/(1+z)}$
qsov						6.3	0.012
	√	√	√	√	√		
extnv						2.3	0.019
qsov						19.4	0.063
	√	√	...	√	√		
extnv						6.8	0.056
qsov						6.9	0.011
	√	√	√	√	...		
extnv						5.9	0.019
qsov						14.3	0.012
	√	...	√	√	√		
extnv						6.8	0.021

fit by a template which has an AGN contribution. A high reliability in recognizing stars from galaxies and AGNs was demonstrated using a spectroscopically identified subset of 470 sources. Overall, we reach a photometric redshift accuracy of $\sigma_{\Delta z/(1+z_{\rm spec})} = 0.014$ $\left(i^+_{AB} < 22.5\right)$ with 4.0% outliers. However, the most important result is perhaps the accuracy for the subsample of AGN-dominated sources ($\sigma_{\Delta z/(1+z_{\rm spec})} = 0.011$) which marks a significant improvement in photometric redshift estimation for these kind of sources. The number of outliers, however, is still nonnegligible (6.3% when considering the "qsov" sample alone).

The excellent photometric redshifts and classifications have been made possible by a number of factors. Here, the most important was likely to be the large spectral coverage from UV to 8 μm, densely sampled with broad, intermediate, and narrow filters. The 25 optical bands essentially mimic low-resolution spectra and are powerful in constraining the positions of spectral features. The near-mid-infrared coverage provides the spectral range required to better distinguish stars from galaxies and helps to constrain SED templates, especially for passive galaxies.

Also important was the recognition of and correction for variability. This is especially relevant for AGN-dominated sources, for which photometry collected over a large time span may dramatically vary from a snapshot SED. We applied a correction for variability by rescaling the optical photometry to a common epoch close to that of IRAC observations.

The correction decreased the number of outliers by a factor of 2. The significant improvement obtained in our sample with this correction demonstrates the importance of quasi-simultaneous observations for projects aiming for good photometric redshift estimates of AGNs. Future surveys will have to consider this point very carefully. Deep imaging in a common filter in each observing epoch would provide the possibility of a valuable first-order correction.

Finally, we used a revised set of SED templates to fully represent the entire range of different contributions of AGNs and their host galaxies. Empirical SED templates are often obtained from objects in the nearby universe where the nuclear and extended fluxes are easily resolved. Their application to high-redshift objects, where the two contributions are mixed together, may not always be appropriate. Our hybrid templates already showed a significant improvement, although further advancements are required.

The catalog of *XMM*–COSMOS-selected sources with photometric redshifts and classifications, together with the X-ray properties (Brusa et al. 2009) offers an unprecedented possibility to study the luminosity function and redshift distribution of AGNs. In combination with the immense multiband COSMOS data set and the new photometric redshifts for optically selected sources (Ilbert et al. 2009), studies of the environment of AGNs will also allow us to probe AGN-fuelling models, clustering, and evolution.

We gratefully acknowledge the contributions of the entire COSMOS collaboration consisting of more than 100 scientists. More information on the COSMOS survey is available at http://www.astro.caltech.edu/cosmos. We acknowledge the anonymous referee for helpful comments that improved the paper. We also acknowledge the hospitality of the Institute for Astronomy, Hawaii, and the support of the Aspen Center for Physics where the work started and the manuscript was completed. G.H. acknowledges support by the German Deutsche Forschungsgemeinschaft, DFG Leibniz Prize (FKZ HA 1850/28-1). This work was supported in part by NASA Grant GO7-8136A.

REFERENCES

Adelman-McCarthy, J. K., et al. 2008, ApJS, 175, 297
Bixler, J. V., Bowyer, S., & Laget, M. 1991, A&A, 250, 370
Brandt, W. N., & Hasinger, G. 2005, ARA&A, 43, 827
Brand, K., et al. 2008, ApJ, 680, 119
Brusa, M., et al. 2007, ApJS, 172, 353
Brusa, M., et al. 2009, ApJ, in press (arXiv:0809.2513)
Calzetti, D., et al. 2000, ApJ, 533, 682
Capak, P., et al. 2007, ApJS, 172, 99
Cappelluti, N., et al. 2007, ApJS, 172, 341
Cappelluti, N., et al. 2009, A&A, submitted
Chabrier, G., Baraffe, I., Allard, F., & Hauschildt, P. 2000, ApJ, 542, 464
Della Ceca, R., et al. 2004, A&A, 428, 383
Elvis, M., et al. 2009, ApJ, submitted
Ferrarese, L., & Merritt, D. 2000, ApJ, 539, L9
Finoguenov, A., et al. 2007, ApJS, 172, 182
Gabasch, A., et al. 2004, A&A, 421, 41
Gebhardt, K., et al. 2000, ApJ, 539, L13
Gilli, R., et al. 2009, A&A, in press (arXiv:0810.4769)
Grazian, A., et al. 2006, A&A, 449, 951
Hasinger, G., et al. 2001, A&A, 365, L45
Hasinger, G., et al. 2007, ApJS, 172, 29
Hoaglin, D. C., Mosteller, F., & Tukey, J. W. 1983, in Understanding Robust and Exploratory Data Analysis (Wiley Series in Probability and Mathematical Statistics), ed. D. C. Hoaglin, F. Mosteller, & J. W. Tukey (New York: Wiley)
Houck, J. R., et al. 2005, ApJ, 622, 105
Ilbert, O., et al. 2006, A&A, 457, 841
Ilbert, O., et al. 2009, ApJ, in press (arXiv:0809.2101)

Koekemoer, A., et al. 2007, ApJS, 172, 196
Leauthaud, A., et al. 2007, ApJS, 172, 219
Lilly, S. J., et al. 2007, ApJS, 172, 70
Madau, P. 1995, ApJ, 441, 18
Mainieri, V., Bergeron, J., Hasinger, G., Lehmann, I., Rosati, P., Schmidt, M., Szokoly, G., & Della Ceca, R. 2002, A&A, 393, 425
McLure, R. J., & Dunlop, J. S. 2002, MNRAS, 331, 795
Mobasher, B., et al. 2007, ApJS, 172, 117
Polletta, M., et al. 2007, ApJ, 663, 81
Prescott, M. K. M., Impey, C. D., Cool, R. J., & Scoville, N. Z. 2006, ApJ, 644, 100
Prevot, M., et al. 1984, A&A, 132, 389
Rowan-Robinson, M., et al. 2008, MNRAS, 386, 697
Sajina, A., Yan, L., Armus, L., Choi, P., Fadda, D., Helou, G., & Spoon, H. 2007, ApJ, 664, 713
Scott, J. E., Kriss, G. A., Brotherton, M., Green, R. F., Hutchings, J., Shull, J. M., & Zheng, W. 2004, ApJ, 615, 135
Scoville, N. Z., et al. 2007a, ApJS, 172, 150
Scoville, N. Z., et al. 2007b, ApJS, 172, 1
Surace, J. A., et al. 2005, SWIRE Data Release 2 (Pasadena, CA: Cal Tech) http://swire.ipac.caltech.edu/swire/astronomers/publications/SWIRE2_doc_083105.pdf
Szokoly, G. P., et al. 2004, ApJS, 155, 271
Taniguchi, Y., et al. 2007, ApJS, 172, 9
Trump, J. R., et al. 2007, ApJS, 172, 383
Trump, J. R., et al. 2009, ApJ, in press (arXiv:0811.3977)
Weedman, D., et al. 2006a, ApJ, 653, 101
Weedman, D. W., et al. 2006b, ApJ, 651, 101
Wold, M., Brotherton, M. S., & Shang, Z. 2007, MNRAS, 375, 989
Wolf, C., et al. 2004, A&A, 421, 913
Yan, L., et al. 2007, ApJ, 658, 778
York, D. G., et al. 2000, AJ, 120, 1579
Zamojski, M. A., et al. 2007, ApJS, 172, 468
Zheng, W., et al. 2004, ApJS, 155, 73