Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

Mon. Not. R. Astron. Soc. **427,** 3103–3133 (2012) doi:10.1111/j.1365-2966.2012.22089.x

Accreting supermassive black holes in the COSMOS field and the connection to their host galaxies

A. Bongiorno,[1,2⋆] A. Merloni,[1] M. Brusa,[1] B. Magnelli,[1] M. Salvato,[1] M. Mignoli,[3] G. Zamorani,[3] F. Fiore,[2] D. Rosario,[1] V. Mainieri,[4] H. Hao,[5] A. Comastri,[3] C. Vignali,[3,6] I. Balestra,[1] S. Bardelli,[3] S. Berta,[1] F. Civano,[7] P. Kampczyk,[8] E. Le Floc'h,[9] E. Lusso,[10] D. Lutz,[1] L. Pozzetti,[6] F. Pozzi,[3] L. Riguccini,[9] F. Shankar[11] and J. Silverman[12]

[1]*Max-Planck-Institut für extraterrestrische Physik (MPE), Giessenbachstrasse 1, Garching bei München, D-85748, Germany*
[2]*INAF-Osservatorio Astronomico di Roma, Via di Frascati 33, 00040, Monteporzio Catone, Rome, Italy*
[3]*INAF-Osservatorio Astronomico di Bologna, Via Ranzani 1, 40127, Bologna, Italy*
[4]*ESO, Karl-Schwarschild-Strasse 2, Garching bei München, D-85748, Germany*
[5]*SISSA, Via Bonomea 265, I-34136 Trieste, Italy*
[6]*Dipartimento di Astronomia, Università di Bologna, via Ranzani 1, I-40127 Bologna, Italy*
[7]*Harvard Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA*
[8]*Institute of Astronomy, ETH Zurich, CH-8093, Zurich, Switzerland*
[9]*Laboratoire AIM, CEA/DSM-CNRS-Université Paris Diderot, IRFU/Service d'Astrophysique, Bât. 709, CEA-Saclay, 91191 Gif-sur-Yvette Cedex, France*
[10]*Max-Planck-Institut für Astronomie, Königstuhl 17, Heidelberg, D-69117, Germany*
[11]*GEPI, Observatoire de Paris, CNRS, Univ. Paris Diderot, 5 Place Jules Janssen, F-92195 Meudon, France*
[12]*Institute for the Physics and Mathematics of the Universe (IPMU), University of Tokyo, Kashiwanoha 5-1-5, Kashiwa-shi, Chiba 277-8568, Japan*

Accepted 2012 September 9. Received 2012 September 5; in original form 2012 May 10

ABSTRACT

Using the wide multiband photometry available in the Cosmic Evolution Survey (COSMOS) field, we explore the host galaxy properties of a large sample of active galactic nuclei (AGNs; ∼1700 objects) with $L_{\rm bol}$ ranging from 10^{43} to 10^{47} erg s^{-1}, obtained by combining X-ray and optical spectroscopic selections. Based on a careful study of their spectral energy distributions, which have been parametrized using a two-component (AGN+galaxy) model fit, we have derived dust-corrected rest-frame magnitudes, colours and stellar masses of the obscured and unobscured AGN hosts up to high redshift ($z \lesssim 3$). Moreover, for the sample of obscured AGNs, we have also derived reliable star formation rates (SFRs). We find that AGN hosts span a large range of stellar masses and SFRs. No colour-bimodality is seen at any redshift in the AGN hosts, which are found to be mainly massive, red galaxies. Once we have accounted for the colour–mass degeneracy in well-defined mass-matched samples, we find a residual (marginal) enhancement of the incidence of AGNs in redder galaxies with lower specific SFRs. We argue that this result might emerge because of our ability to properly account for AGN light contamination and dust extinction, compared to surveys with a more limited multiwavelength coverage. However, because these colour shifts are relatively small, systematic effects could still be considered responsible for some of the observed trends. Interestingly, we find that the probability for a galaxy to host a black hole that is growing at any given 'specific accretion rate' (i.e. the ratio of X-ray luminosity to the host stellar mass) is almost independent of the host galaxy mass, while it decreases as a power law with $L_{\rm X}/M_*$. By analysing the normalization of such a probability distribution, we show how the incidence of AGNs increases with redshift as rapidly as $(1+z)^4$, which closely resembles the overall evolution of the specific SFR of the entire galaxy population. We provide analytical

⋆E-mail: angela.bongiorno@oa-roma.inaf.it

fitting formulae that describe the probability of a galaxy of any mass (above the completeness limit of the COSMOS) to host an AGN of any given specific accretion rate as a function of redshift. These can be useful tools for theoretical studies of the growing population of black holes within galaxy evolution models. Although AGN activity and star formation in galaxies do appear to have a common triggering mechanism, at least in a statistical sense, within the COSMOS sample, we do not find any conclusive evidence to suggest that AGNs have a powerful influence on the star-forming properties of their host galaxies.

Key words: catalogues – surveys – galaxies: active – galaxies: evolution – galaxies: fundamental parameters.

1 INTRODUCTION

Since their discovery more than 40 yr ago, active galactic nuclei (AGNs) that are powered by accretion on to supermassive black holes (SMBHs) have mostly been considered to be rare and exotic objects with peculiar (and extreme) astrophysical properties. In the past decade, studies of the local Universe have established the presence of SMBHs in the nuclei of virtually all galaxies with a bulge/spheroidal component (e.g. Kormendy & Richstone 1995; Kormendy & Bender 2011). This has dramatically changed our perception of this class of objects. Because the cosmological growth of SMBHs is mostly because of the accretion of matter during their active phases (Soltan 1982), we are led to conclude that most bulges have been through a phase of strong nuclear activity (Marconi et al. 2004; Merloni & Heinz 2008; Shankar 2009). Moreover, as the energy released in the process of accretion on to the SMBH can be higher than the total binding energy of a massive galaxy, an AGN might represent a key ingredient of the formation and evolution of all galaxies. The question of whether, and how, nuclear black holes influence their host galaxies, and vice versa, has become a major focus in studies of structure formation and evolution.

Indeed, many observational discoveries do support the notion that there is a close link between the growth of a SMBH and the assembly of the host galaxy, the most crucial of which are the following.

(i) The tight correlation between the mass of the central SMBH and the various properties of their host galaxies, such as luminosity (Kormendy & Richstone 1995; McLure & Dunlop 2001; Marconi & Hunt 2003), stellar mass (Magorrian et al. 1998) and velocity dispersion (Ferrarese & Merritt 2000; Gebhardt et al. 2000; Tremaine et al. 2002; Kormendy & Bender 2011).

(ii) As populations, SMBHs and galaxies seem to evolve with redshift in a similar way. In fact, the AGN evolution is luminosity-dependent with low-luminosity AGNs reaching a peak in their space density later in the history of the Universe (i.e. at lower redshift) than higher luminosity AGNs (Miyaji, Hasinger & Schmidt 2000; Fiore et al. 2003; Ueda et al. 2003; Cirasuolo, Magliocchetti & Celotti 2005; Hasinger, Miyaji & Schmidt 2005; La Franca et al. 2005; Bongiorno et al. 2007). This pattern is similar to the so-called 'cosmic downsizing' of star-forming galaxies (Cowie et al. 1996; Menci et al. 2008), also seen in spheroidal galaxies (Cimatti, Daddi & Renzini 2006; Thomas et al. 2010).

(iii) The shape of the integrated AGN activity over cosmic time is similar to that of the global SFR with a peak at $z \sim 2$ and a rapid decline at both lower (e.g. Dickinson et al. 2003; Merloni 2004; Hopkins, Richards & Hernquist 2007; Shankar, Weinberg & Miralda-Escudé 2009) and higher (e.g. Franceschini et al. 1999; Wilkins, Trentham & Hopkins 2008; Brusa et al. 2009a) redshifts.

(iv) Finally, luminous AGNs show a correlation between the AGN bolometric luminosity and the star formation (SF) of the host galaxy over more than five orders of magnitude in luminosity (Netzer 2009), albeit with a substantial scatter. However, possible deviations from such behaviour have been reported (Shao et al. 2010; Mullaney et al. 2012; Page et al. 2012; Santini et al. 2012; Trichas et al. 2012).

While the evidence of a tight link between the formation of galaxies and the growth of the SMBHs at their centres is plentiful and clear, the physical processes behind this interplay remain unclear. In particular, the fundamental question of whether AGN-driven feedback processes are ultimately responsible for determining the global properties of the galaxy population (e.g. luminosity functions, colour–magnitude distributions and the evolution thereof), or whether, on the contrary, they happen to be triggered and fuelled as a byproduct of SF activity and the morphological evolution of their hosts, still remains unanswered. This is of fundamental importance for understanding both galaxy assembly and the accretion density of the Universe. The work we present here is motivated by the desire to obtain robust, and statistically significant, constraints on the possible scenarios for the coevolution of galaxies and AGNs.

A number of models (e.g. Somerville, Primack & Faber 2001; Granato et al. 2004; Monaco & Fontanot 2005; Springel, Di Matteo & Hernquist 2005; Croton et al. 2006; Hopkins et al. 2006; Schawinski et al. 2006; Cen & Chisari 2011) have been developed in recent years to explain this coevolution and to describe the main mechanisms that fuel the central SMBH and build the galaxy's bulge. Some of these semi-analytical models and hydrodynamical simulations (e.g. Springel et al. 2005; Hopkins et al. 2006; Menci et al. 2008) invoke major mergers of gas-rich galaxies as the main fuelling mechanism. Alternative mechanisms have been discussed in the literature, including minor mergers (e.g. Johansson, Burkert & Naab 2009), disc instabilities (e.g. Genzel et al. 2008) and recycled gas from dying stars (e.g. Ciotti, Ostriker & Proga 2010). Different fuelling mechanisms are usually associated with different luminosity ranges. For example, major mergers are invoked to trigger bright quasars (e.g. Hopkins et al. 2008), while the role of major mergers in triggering more typical ($\sim L_*$) AGNs is called into question (e.g. Cisternas et al. 2011; Schawinski et al. 2011; Kocevski et al. 2012). In contrast, secular processes seem to be enough to fuel low-luminosity AGNs (i.e. Seyfert-like objects; Hopkins et al. 2008; Allevato et al. 2011). Because of the large mismatch in physical scales between AGNs and galaxies, all these models have to rely on specific assumptions regarding the mechanism responsible for the link between the nuclear activity, which releases most of its energy on the scale of few Schwarzschild radii R_S ($\sim 10^{-5}$ pc), and the stellar population on much larger scales ($\sim$ a few kpc, some $\sim 10^8$ times R_S). This mechanism is, in most cases, an energetic feedback from

the central engine that deposits the energy liberated by the accretion process within the host galaxy, or its dark matter halo (Silk & Rees 1998). However, the physical description of the feedback itself and its effect on the AGN host galaxy are very different among different models.

In particular, because of the complex interplay between gas inflow, outflow and nuclear obscuration during active AGN phases, a correct and complete identification of unobscured, obscured and highly obscured AGNs at all redshifts (and especially in the $z = 1$–3 interval, where most of the feedback is expected to occur) is therefore crucial for a comprehensive understanding of the still little explored phase of the common growth of SMBHs and their host galaxies.

Generally speaking, two different approaches can be followed to test the predictions of different theoretical models of AGN feedback and to constrain the physical mechanisms at play: a detailed study of the physical properties of relatively small but well-defined samples of selected sources or a statistical investigation of a large sample of AGNs, which includes different subclasses. Within the Cosmic Evolution Survey (COSMOS), the first approach has been followed in our previous papers (i.e. Merloni et al. 2010; Cisternas et al. 2011; Lusso et al. 2011; Mainieri et al. 2011), while the latter approach is adopted here.

In order to study growing SMBHs and their influence on the host galaxies over different periods of their cosmic evolution, it is necessary to compile a sample of AGNs (both unobscured and obscured) that is as complete as possible, spanning a wide range in luminosity and redshift. X-ray surveys have been proved to be the most efficient way to compile nearly unbiased samples of Compton-thin AGNs (Brandt & Hasinger 2005), but they might miss Compton-thick AGNs (e.g. Comastri & Brusa 2008; La Massa et al. 2009). However, even though it allows a straightforward selection of unobscured AGNs and is less biased against Compton-thick AGNs, the optical spectroscopic selection can fail to select those AGNs for which the nuclear optical/infrared (IR) light is diluted by their host galaxy. To overcome these limitations, we decided to consider and combine X-ray and optically selected samples of AGNs, compiling a highly homogeneous and representative sample of obscured and unobscured AGNs selected in the COSMOS field over a wide redshift baseline ($0 < z < 4$).

Other selection criteria, for example, from IR colour–colour diagrams (Stern et al. 2005; Donley et al. 2012) or from radio surveys (e.g. Smolčić et al. 2008), can provide additional AGNs and are needed in order to obtain a full picture of the AGN–galaxy coevolution (e.g. Hickox et al. 2009). However, because X-ray and optical selections make up more than 50 per cent of the full AGN population in COSMOS (Brusa et al., in preparation), for the purpose of this paper, we focus on only these two methods, quantifying, whenever relevant, the incompleteness because the other samples have not been included.

The major observational challenge in any comprehensive study of the AGN–galaxy coevolution is the accurate separation of the AGN and galaxy emission components, at all optical–IR wavelengths. This is a crucial step for a number of reasons. Depending on the intrinsic spectral energy distribution (SED) of the nuclear (AGN) and of the stellar light, and because of their respective level of extinction, inaccurate deblending might not only hamper any precise determination of the physical properties of the galaxy, but might also mask AGN signatures and bias our view of the SMBH growth.

To this end, we have used a two-component SED fitting procedure. The observed SED is fitted with a large grid of models based on a combination of the Richards et al. (2006, hereafter R06) AGN template and several host-galaxy models (i.e. synthetic spectra created from the stellar population synthesis models of Bruzual & Charlot 2003, hereafter BC03). Given the wide multiwavelength coverage available in the COSMOS field, this fitting technique allows us to decompose, typically with a high level of confidence, the entire SED into nuclear AGN and host galaxy components, and to derive robust measurements of the physical properties of the host galaxy (e.g. rest-frame colours, stellar mass, *K*-band luminosity and SFR). The strength of the SED fitting method is that, given sufficiently wide photometric coverage, it is applicable to all AGNs, obscured and unobscured, independent of their luminosity. In particular, once a comprehensive set of templates for the SED components is chosen, the method can be applied (almost) blindly to any detected object in a multiwavelength survey, irrespective of the specific selection criteria, reaching an accuracy that depends crucially on the number of bands and depth of the available photometric catalogues. In this respect, COSMOS is a uniquely suited field for our investigation.

The structure of the paper is as follows. In Section 2, we introduce our sample, before proceeding to a discussion of the two-component SED fitting method in Section 3. In Section 4, we discuss the physical properties of the AGN host galaxies (i.e. colours, masses and SFRs) compared to normal galaxies. In Sections 4.3 and 4.4, we focus on the incidence of the AGNs as a function of their host properties (i.e. mass and SFR, respectively). We present a final discussion and a summary of our main results in Sections 5 and 6, respectively. Additionally, in Appendix A, we discuss the effect on the derived properties of the host galaxy if the AGN component is not properly subtracted, while in Appendix B we show a detailed comparison between two SFR indicators: the ultraviolet (UV)–optical SED fitting and the method based on the far-infrared (FIR) band. In Appendix C, we provide two tables with the AGN catalogue and the parameters derived from the SED fitting (e.g. rest-frame magnitudes, masses and SFRs), together with a short explanation of the columns. These tables show only 16 lines as an example of the tables published online (see Supporting Information). Throughout this paper, we use the standard cosmology ($\Omega_{\rm m} = 0.3$, $\Omega_\Lambda = 0.7$, with $H_0 = 70$ km s^{-1} Mpc^{-1}). Magnitudes, if not differently specified, are *AB* magnitudes.

2 AGN SAMPLES FROM COSMOS

In order to construct a sample of AGNs that is as complete as possible, including both obscured and unobscured objects, first we selected *XMM*–COSMOS point-like sources (Hasinger et al. 2007; Cappelluti et al. 2009). To these, we added a smaller sample of purely optically selected AGNs from the zCOSMOS bright spectroscopic survey (Lilly et al. 2009), spanning a lower luminosity range for AGNs (Bongiorno et al. 2010; Mignoli et al., in preparation). In this section, we give the details of our sample selection.

2.1 X-ray selected sample

The *XMM*–COSMOS catalogue includes ∼1800 point-like X-ray sources. The X-ray catalogue has been presented by Cappelluti et al. (2009), while the optical identifications and multiwavelength properties have been discussed by Brusa et al. (2010).[1] The sources

[1] In this paper, we use an improved version of the *XMM*–COSMOS identification catalogue with respect to the one published by Brusa et al. (2010). The main difference is the availability of new photometric redshifts from Salvato et al. (2011). The updated version of the *XMM*–COSMOS identification catalogue can be downloaded from http://www.mpe.mpg.de/XMMCosmos/xmm53_release/.

of the *XMM*–COSMOS catalogue of Brusa et al. (2010) have been classified as normal galaxies and obscured and unobscured AGNs, according to their properties (i.e. X-ray luminosities, spectra and multiwavelength SEDs; Salvato et al. 2011). Here, we consider only those sources classified as AGNs, both unobscured and obscured, as we describe below.

Optical spectra are available for more than half of the sample from different instruments: Visible Multi-Object Spectrograph/Very Large Telescope (VIMOS/VLT; zCOSMOS project; Lilly et al. 2007, 2009), the Inamori Magellan Areal Camera and Spectrograph (IMACS; Trump et al. 2007, 2009), Multiple Mirror Telescope (MMT; Prescott et al. 2006) and Deep Imaging Multi-Object Spectrograph (DEIMOS)/Keck-II (PIs: Scoville, Capak, Salvato, Kartaltepe, Mobasher).

From the *XMM*–COSMOS subsample for which spectra are available, we have identified 430 sources that show broad (FWHM $>$ 2000 km s^{-1}) emission lines in their optical spectra, which we classify as type 1 (unobscured). Furthermore, 402 spectroscopically confirmed AGNs were instead classified as type 2 (obscured), by making use of both their spectral properties and X-ray luminosities. In particular, the class of type 2 AGNs includes (i) all sources that do not show broad emission lines, and for which the ratio between their high-ionization narrow lines indicates AGN activity (Baldwin, Phillips & Terlevich 1981), and (ii) when the high ionization lines are not detected in the observed wavelength range, all the objects without broad lines and with rest-frame (2–10 keV) X-ray luminosity greater than 2×10^{42} erg s^{-1}.

For about half of the *XMM*–COSMOS sources (723 objects), only photometric redshifts are available (the photo-z AGN sample). However, these are very accurate with $\sigma_{\Delta z/(1+z)} \sim 0.015$ and a fraction of outliers of just 5.8 per cent (Salvato et al. 2009, 2011). Thus, the *XMM*–COSMOS sources for which no spectra are available have been classified as unobscured type 1 AGNs in 172 cases, and as obscured type 2 AGNs in 551 cases. In practice, the source classification for photo-z AGNs was performed following Salvato et al. (2011) on the basis of the template[2] that best describes their SED, and applying the same threshold in X-ray luminosity ($L_{[2-10]\,\mathrm{keV}} > 2 \times 10^{42}$ erg s^{-1}).

Summarizing, the *XMM*–COSMOS sample analysed in this paper comprises 602 unobscured type 1 AGNs (430 with spectroscopic redshift and broad emission lines; 172 with photometric redshift) and 953 obscured type 2 AGNs (402 with spectroscopic redshift and 551 with photometric redshift), for a total of 1555 X-ray selected AGNs.

2.2 Optically selected Seyfert 2 galaxies: diagnostic diagrams

For type 1 AGNs, the X-ray and optical selection match very well; that is, in addition to the 430 spectroscopic broad-line AGNs detected by *XMM*, there are only 25 ($\sim$6 per cent) objects purely optically selected from the zCOSMOS bright survey, which are not detected in the X-ray band by *XMM* (the number goes down to just 14, $\sim$3 per cent, if we also consider *Chandra* detections). This is not the case for obscured type 2 AGNs, for which a more substantial fraction ($>$25 per cent) that show no X-ray counterparts can be selected from optical surveys. Moreover, while the redshift and luminosity distributions of X-ray and optically selected type 1 AGNs are very similar, for type 2 objects this is not the case, and different bands can allow us to retrieve different subclasses of sources. For this reason, we also include in this study obscured type 2 AGNs selected purely on optical properties. Because optically selected AGNs (Seyfert 2 galaxies) tend to have much lower luminosities (see Fig. 10), this allows us to enlarge the analysed luminosity range.

The sample of optical type 2 AGNs was selected in the following way. From the final 20 000 VIMOS/VLT zCOSMOS bright survey, and in analogy with the procedure adopted by Bongiorno et al. (2010) for the first 10 000 spectra (10k; Lilly et al. 2009), we have selected 392 sources using the standard diagnostic diagrams ([O III]/Hβ versus [N II]/Hα and [O III]/Hβ versus [S II]/Hα; Baldwin et al. 1981, the BPT diagram) to isolate AGNs in the redshift range $0.15 < z < 0.45$. We have used the diagnostic diagram [O III]/Hβ versus [O II]/Hβ (Rola, Terlevich & Terlevich 1997; Lamareille et al. 2004, the 'blue' diagram) to extend the selection to higher redshifts ($0.5 < z < 0.92$). These diagnostic diagrams allow the classification of emission-line objects as Seyfert 2 galaxies, candidate Seyfert 2 galaxies and low-ionization nuclear emission-line region (LINER) galaxies. Here, we include only the 105 objects classified as secure Seyfert 2 galaxies (the LINER sample is studied separately by Tommasin et al. 2012). Of these 105 AGNs, 18 have X-ray counterparts and are therefore already included in the X-ray selected sample described above. Hence, from the optical selection based on the spectral diagnostic diagrams, we have isolated 87 pure optically selected Seyfert 2 that do not show any X-ray emission at the depth of our *XMM* observations.

2.3 Optically selected obscured AGNs: the [Ne V emission-line sample

A relatively new method to select obscured AGNs in optical surveys uses the detection of the [Ne V] line, whose presence is a distinctive feature of AGN activity (e.g. Gilli et al. 2010). Using this criterion on the same 20 000 zCOSMOS bright spectra, Mignoli et al. (in preparation) have identified 95 obscured type 2 AGNs, 11 of which have already been identified through the emission-line diagnostic diagrams. Of the remaining 84 sources, 24 have X-ray counterparts and are hence already included in the X-ray selected sample described above. Thus, from the [Ne V]-based optical selection, we have isolated an additional 60 pure optically selected Seyfert 2 galaxies that do not show any X-ray detection.

In total, the optically selected obscured AGNs amount to 189 sources, of which 147 are purely optically selected with no *XMM* detection.[3] Because these two criteria are both based on spectral features, the number of spectra available for obscured AGNs is artificially higher than for the unobscured AGNs.

2.4 Total AGN sample and the parent galaxy sample

The final combined sample analysed in this paper consists of 1702 AGN: 602 type 1 AGNs (430 of these have optical spectra) and 1100 type 2 AGNs (549 of these have spectra). Fig. 1 shows a graphical representation of the sample and its basic substructure, highlighting

[2] This was chosen from a library of hybrid templates obtained by combining the observed galaxy templates from Polletta et al. (2007) with different levels of AGN contamination (Salvato et al. 2011).

[3] We note here that seven of the 147 purely optically selected AGNs are actually detected in the deeper *Chandra* observation covering the central part of the COSMOS field (Elvis et al. 2009; Civano et al. 2012). In order to keep selection effects under control and to exploit the full 2 deg^2 area of COSMOS, in this paper we limit the X-ray analysis to the *XMM*–COSMOS sample.

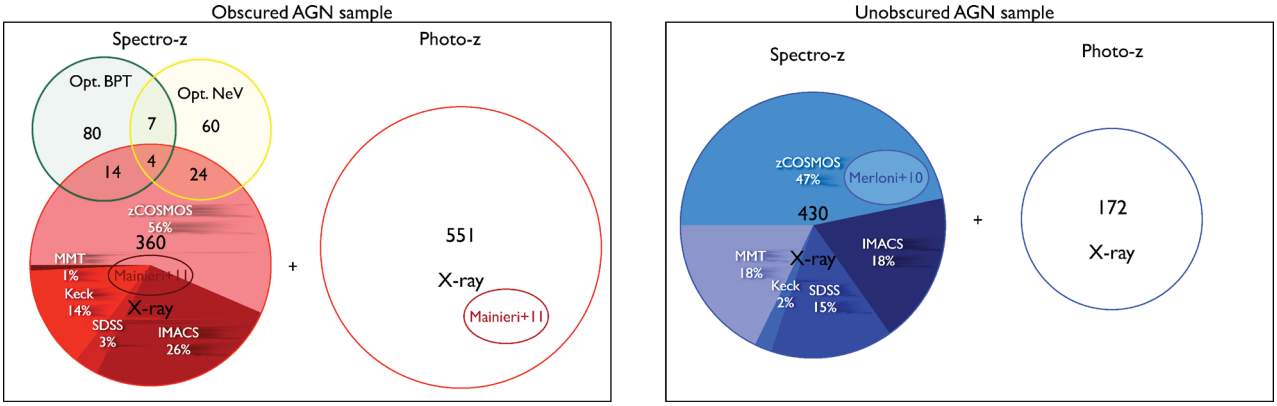

Figure 1. Schematic view of the analysed subsamples for obscured (left) and unobscured (right) AGNs. Different circles represent the different methods used to select them. Blue and red circles correspond to X-ray unobscured and obscured AGNs, respectively, and green and yellow circles denote optically selected AGNs using the diagnostic diagrams and the [Ne v] line, respectively. Open and filled symbols are for photometric and spectroscopic redshifts, respectively, whose origin is highlighted with the slices of different colour gradations. The combined analysed sample consists of 1702 AGNs: 602 type 1 AGNs (of which 430 have spectra) and 1100 type 2 AGNs (of which 549 have spectra).

Table 1. Analysed subsamples of obscured and unobscured AGNs. Bold numbers indicate the total number of sources while the rest highlight the sources in common between two or more classification methods. The combined analysed sample consists of 1702 AGNs: 602 unobscured AGNs (of which 430 have spectra) and 1100 obscured AGNs (of which 549 have spectra).

Obscured AGN sample				
	Spec. z			Phot. z
	X-ray	Opt. BPT	Opt. Ne v	X-ray
X-ray	**360**	18	28	**551**
Opt. BPT	18	**105**	11	
Opt. Ne v	28	11	**95**	
Unobscured AGN sample				
	Spec. z			Phot. z
X-ray	**430**			**172**

the classification, the method used and the overlaps between the various subsamples (see also Table 1). The colour code adopted in Fig. 1 is used throughout the paper: blue and red colours denote X-ray unobscured and obscured AGNs, respectively; green denotes those Seyfert 2 galaxies selected optically using the diagnostic diagrams; yellow denotes obscured AGNs selected via the [Ne v] line emission. Open and filled symbols are used for photometric and spectroscopic redshifts, respectively. The full sample can be found in the online material (see Supporting Information). An example of the published table is given in Table C1.

Because our aim in this paper is to study the properties of the AGN's host galaxy and to compare them to those of normal galaxies, throughout the paper we use the galaxy sample selected by the Infrared Array Camera (IRAC) in the COSMOS area as a comparison 'parent' galaxy sample. For the whole galaxy sample, photometric redshifts, masses and SFRs are available (Ilbert et al. 2010). Moreover, in order to disentangle multiple dependences and thus to isolate the real differences in the properties of galaxies hosting an AGN from those that do not, we have also created two matched parent samples: one matched using mass and the other matched using the i-band apparent magnitude. In both cases, we matched each AGN with three non-AGNs chosen from the whole IRAC sample. For the mass-matched sample, we have considered the (z, M) plane. For each AGN with a given redshift and host galaxy stellar mass (see Section 3.5.1), we extracted from the parent sample the three non-AGN objects that lie closest in the (z, M) plane to the AGN, allowing a maximum difference of a factor of 10 in mass. The same procedure was followed to build the sample matched using the i-band, this time working on the (z, i) plane, allowing a maximum difference of a factor of 4 in luminosity.

2.5 Multiwavelength observations

The COSMOS is a multiwavelength observational project over $1.4 \times 1.4\,\mathrm{deg}^2$ of equatorial field centred at (RA, Dec.)$_{J2000}$ = (150.1083, 2.210) (Scoville et al. 2007). The COSMOS field is characterized by an unprecedented deep multiwavelength coverage, and in order to compile the SED for the AGN sample described above, we have exploited the large amount of multiwavelength data available in the field. For most of the sources, we have constrained the SED using 14 different bands that encompass optical to mid-infrared (MIR) wavelengths. More specifically, for all our sources, we have used six Subaru bands (B, V, g, r, i, z; from Taniguchi et al. 2007) and the Canada–France–Hawaii Telescope (CFHT) U, J and K bands (Capak et al. 2007; McCracken et al. 2010). All but a few (∼97 per cent) of our sources are detected in all four *Spitzer*/IRAC bands (3.6, 4.5, 5.8 and 8.0 μm; Sanders et al. 2007; Ilbert et al. 2010) and ∼80 per cent of them are also detected in the 24-μm Multiband Imaging Photometer for *Spitzer* (MIPS) band (Le Floc'h et al. 2009). Thus, for most of the sources, we can constrain the SED in a very large wavelength interval, ranging from ∼3800 Å (U_{CFHT}) to 24 μm (MIPS). Moreover, for a few per cent of the sample, we can also rely on longer wavelength detections. In particular, for the X-ray selected sample, 109 sources have been detected at 70 μm (MIPS; Sanders et al. 2007), 216 are visible at 100 μm and 201 at 160 μm (77 of these are detected in both wavelengths) by the *Herschel* Space Observatory as part of the Photodetector Array Camera and Spectrometer (PACS) Evolutionary Probe (PEP;[4] Lutz et al. 2011) guaranteed time key programme.

[4] http://www.mpe.mpg.de/ir/Research/PEP/index.php

3 SPECTRAL ENERGY DISTRIBUTIONS OF COSMOS AGNS

In this section, we describe the main tool in our analysis (i.e. the study of the overall SED of the entire AGN sample). The approach we follow is to consider, with as few a priori assumptions as possible, each and every observed SED in the sample as the result of the superposition of the nuclear AGN emission and of the galactic stellar light entering the fixed apertures of 3 arcsec diameter used to derive all optical/IR photometry (Salvato et al. 2009).

The typical SED of a pure quasi-stellar object (QSO), shown in the right panel of Fig. 2, is characterized by two bumps in the UV and IR regimes (Sanders et al. 1989; Elvis et al. 1994; R06), which create a dip at around 1 μm. The UV bump is interpreted as thermal emission from the accretion disc (Czerny & Elvis 1987). The IR bump is thought to be a result of the absorption of intrinsic AGN UV, X-ray and optical radiation by dusty clouds in the AGN torus on ∼pc scales, which subsequently reradiate this energy at IR frequencies (Barvainis 1987). Despite the apparent large scatter in the observed SED of large QSO/AGN samples (R06), there appears to be little evidence for substantial systematic changes in the shape of such an intrinsic SED (apart from those introduced by dust extinction) across wide luminosity ranges. This has recently been demonstrated by the analysis of the broad Hα line of selected local AGNs from the Seventh Data Release (DR7) of the Sloan Digital Sky Survey (SDSS; Stern & Laor 2012).

However, the optical SED of a galaxy (for example, see the left panel of Fig. 2) is usually modelled as the integrated light of the stellar populations of the galaxy, which are generated by different star formation histories (SFHs). Galaxy SEDs peak typically at around 1 μm for a very wide range of SFR histories.

In most AGNs (with the exception of the most luminous), both components contribute significantly and the global SED is the result of the combination of the SEDs of the central QSO and the host galaxy (Hao et al. 2009; Merloni et al. 2010; Lusso et al. 2011). How much these two components contribute to the global SED in detail depends on a number of factors, the most important being their relative luminosity and the level of obscuration affecting each of them.

3.1 Modelling the emission components of the AGN and host galaxy using SED fitting

Fixing the redshift of the source to the spectroscopic or photometric redshift, we fit the observed fluxes $f_{\rm OBS}$ with a two-component model, using a combination of AGN and host-galaxy templates (see, for example, similar approaches by Pozzi et al. 2010; Lusso et al. 2011); see also Appendix A for a comparison with the standard SED fitting method without the AGN component. This fitting technique, initially presented by Bongiorno et al. (2007), has already been successfully used in the analysis of smaller subsamples of AGNs in the COSMOS field (i.e. Merloni et al. 2010; Mainieri et al. 2011). A more complex analysis, including multiple dust emission components, has been carried out by Lusso et al. (2012). As mentioned previously, the SEDs of the AGN and the galaxy are shaped in a way that allows a relatively easy decoupling, with the galaxy's emission peak falling in the wavelength range of the AGN dip at ∼1 μm (rest frame). To be specific, we assume that

$$f_{\rm obs} = c_1 f_{\rm AGN} + c_2 f_{\rm GAL}. \quad (1)$$

Using a library of galaxy and AGN templates, we find the best normalization parameters c_1 and c_2 that reproduce the observed flux of each object, by minimizing χ^2. In type 1 unobscured AGNs, we expect the AGN component to dominate in the optical and IR bands with some degree of contribution from the host galaxy, while the optical continuum of type 2 obscured AGNs is dominated by the host-galaxy emission with the AGN component rising mainly at an IR wavelength. Note that we are not implying that both components

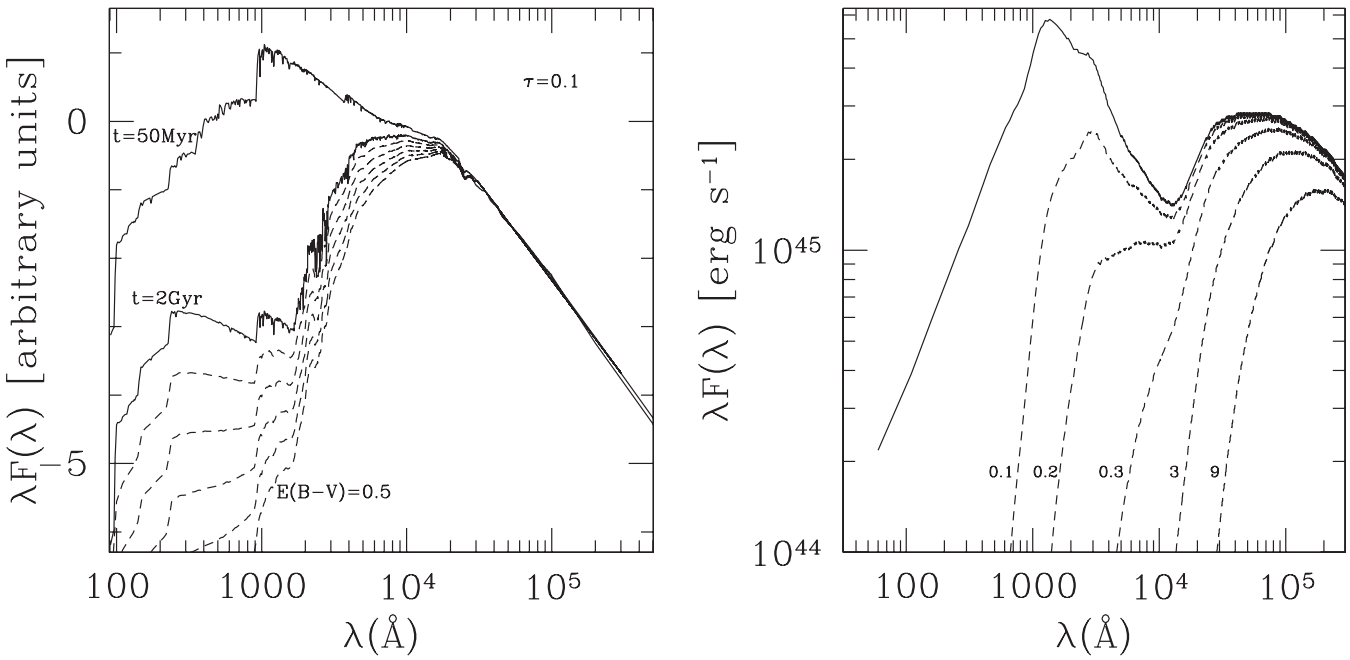

Figure 2. Left: examples of galaxy templates derived from the models of BC03. The two solid curves correspond to a model with $\tau = 0.1$ and with ages of 50 Myr and 2 Gyr, respectively. The dashed curves correspond to the latter SED dust-reddened using equation (3) with $E_{\rm g}(B-V) = [0.1, 0.2, 0.3, 0.4, 0.5]$. Right: AGN template by R06 (solid line). The dashed curves correspond to the same SED dust-reddened using equation (4) with different $E_{\rm a}(B-V)$ values [0.1, 0.2, 0.3, 3, 9].

are always detectable. Solutions in which the SED is a pure QSO (in the case of a type 1 AGN) or a galaxy with a negligible contribution from the central AGN (in the case of a type 2 AGN) are both possible.

3.2 Library templates

The data are fitted with a large grid of AGN and galaxy templates. For the AGN component, we adopt the R06 mean QSO SED, as derived from the study of 259 IR-selected quasars with both SDSS and *Spitzer* photometry (see the right panel of Fig. 2). The R06 SED is an extension and is consistent with the original Elvis et al. (1994) QSO SED. For the galaxy component, we have generated a library of synthetic spectra using the models of stellar population synthesis of BC03. We have assumed a universal initial mass function (IMF) from Chabrier (2003) and we have built 10 exponentially declining SFHs SFR $\propto e^{-t_{age}/\tau}$ with e-folding times, τ, ranging from 0.1 to 30 Gyr and a model with constant SF. For each of the SFHs, the SED was generated for a grid of 13 ages ranging from 50 Myr to 9 Gyr, subject only to the constraint that the age should be smaller than the age of the Universe at the redshift of the source. Some examples of the used galaxy templates are given in the left panel of Fig. 2.

3.2.1 Extinction

Both the AGN and the galaxy templates can be affected by dust extinction. The observed flux can be written as

$$f_{\rm obs}(\lambda) = f_{\rm int}(\lambda) \times 10^{-0.4A_\lambda}, \quad (2)$$

where $f_{\rm obs}$ and $f_{\rm int}$ are the observed and intrinsic fluxes, respectively, and A_λ is the extinction at a given wavelength λ described by $A_\lambda = k(\lambda)E(B-V)$, with the colour excess $E(B-V)$ and the reddening curve $k(\lambda)$.

Galaxies

Recent studies on high-redshift galaxies and SF obscured by dust have shown the importance of the reddening in the high-redshift universe. Therefore, dust extinction, produced inside the galaxies themselves, is an important effect to be taken into account.

As the reddening curve for our galaxy templates, we have chosen the Calzetti law (Calzetti et al. 2000), which is the most used attenuation curve in high-redshift studies. The Calzetti law is an empirical relation derived using a small sample of low-redshift starburst galaxies. It has the following form

$$k_{\rm g}(\lambda) = \begin{cases} 2.659\left(-2.156 + \frac{1.509}{\lambda} - \frac{0.198}{\lambda^2} + \frac{0.011}{\lambda^3}\right) + R_V \\ 2.659\left(-1.857 + \frac{1.040}{\lambda}\right) + R_V \end{cases}, \quad (3)$$

with $R_V = 4.05$. The two equations are valid for $0.12 \leq \lambda \leq 0.63$ μm and $0.63 \leq \lambda \leq 2.20$ μm, respectively. Below and above the valid wavelength range, the slope is computed by extrapolating to lower and higher wavelengths the value of $k_{\rm g}(\lambda)$ obtained by interpolating between 0.11–0.12 and 2.19–2.20 μm, respectively.

In the fitting procedure, we have considered for the galaxy component only $E_{\rm g}(B-V)$ values in the range $0 \leq E_{\rm g}(B-V) \leq 0.5$. Following Fontana et al. (2006) and Pozzetti et al. (2007), we impose the prior $E_{\rm g}(B-V) < 0.15$ if $t_{\rm age}/\tau > 4$ (i.e. we exclude the models implying large dust extinctions in the absence of a significant SFR, $t_{\rm age}/\tau > 4$).

In the left panel of Fig. 2, we show two examples of galaxy templates with $\tau = 0.1$, and $t_{\rm age} = 50$ Myr and 2 Gyr (solid lines). For the latter template, we also show the corresponding dust-reddened templates obtained by applying the above equations with different $E_{\rm g}(B-V)$ values (dashed lines).

Active galactic nuclei

The extinction of the nuclear AGN light has been modelled with a Small Magellanic Cloud-like dust-reddening law from Prevot et al. (1984) for which

$$k_{\rm a}(\lambda) = 1.39\lambda_{\mu\rm m}^{-1.2}. \quad (4)$$

We have considered $E_{\rm a}(B-V)$ values in the ranges $0 \leq E_{\rm a}(B-V) \leq 1$ and $0.3 \leq E_{\rm a}(B-V) \leq 9$ for unobscured and obscured type 1 AGNs, respectively, in steps of $\Delta E_{\rm a}(B-V) = 0.1$. In the right panel of Fig. 2, we show the AGN template (solid line) and the corresponding dust-reddened templates obtained for different $E_{\rm a}(B-V)$ values. As seen in this figure, the upper limit of $E_{\rm a}(B-V) = 9$ corresponds to a very extinct AGN SED, and an even higher upper limit would not change the SED shape much.

Taking into account the described extinction laws, we can rewrite equation (1) as

$$f_{\rm obs} = c_1 f_{\rm AGN}^{\rm int}(\lambda) \times 10^{-0.556\ \lambda_{\mu m}^{-1.2} E_{\rm a}(B-V)} + c_2 f_{\rm GAL}^{\rm int}(\lambda) \times 10^{-0.4 k_{\rm g}(\lambda) E_{\rm g}(B-V)}, \quad (5)$$

with $k_{\rm g}(\lambda)$ described by equation (3).

3.3 Results of the spectral energy distribution fitting

Given the wide multiwavelength coverage, the fitting technique described above allows us to decompose the entire SED into a nuclear AGN and a host-galaxy component and to derive robust measurements of the properties of both the AGN and the host galaxy, for almost all (98.6 per cent) of the objects in the original sample.[5] Some examples of the SED fitting for unobscured (upper panels) and obscured (bottom panels) AGNs are shown in Fig. 3, while the left panel of Fig. 4 shows the median SED chosen as the best fit in the SED fitting for unobscured (blue) and obscured (red) AGNs, together with the 10th, 25th, 75th and 90th percentiles of the samples. Finally, the right panel of the Fig. 4 highlights the median AGN contribution to the global SED at different wavelengths for unobscured (blue continuum line) and obscured (red continuum line) AGNs. The shaded areas mark the 25th and 75th percentiles, while the long dashed lines mark the 10th and 90th percentiles of the samples. From this figure, we see that, for obscured AGNs, three-quarters of the objects have an AGN contribution in the optical bands smaller than 20 per cent. Thus, any residual AGN contamination (or AGN oversubtraction) should not affect the determination of the rest-frame optical colours of the host by more than about 0.1 mag. However, at rest-frame wavelengths of about 1 μm, three-quarters of type 1 AGNs have AGN fractions smaller than about 50 per cent. This implies that, even for unobscured objects, our method allows a robust determination of the total stellar masses of the host (see also Merloni et al. 2010).

Based on the best-fitting solution, we have derived the AGN luminosity and the rest-frame magnitudes, colours, stellar mass content

[5] In fact, because of an insufficient photometric coverage, five sources have been excluded from the analysis. For 18 sources, no good solution could be found because of problems in the photometry or in the SED; the reduced χ^2 in these cases is large (>20). For these sources, no physical parameters could be derived.

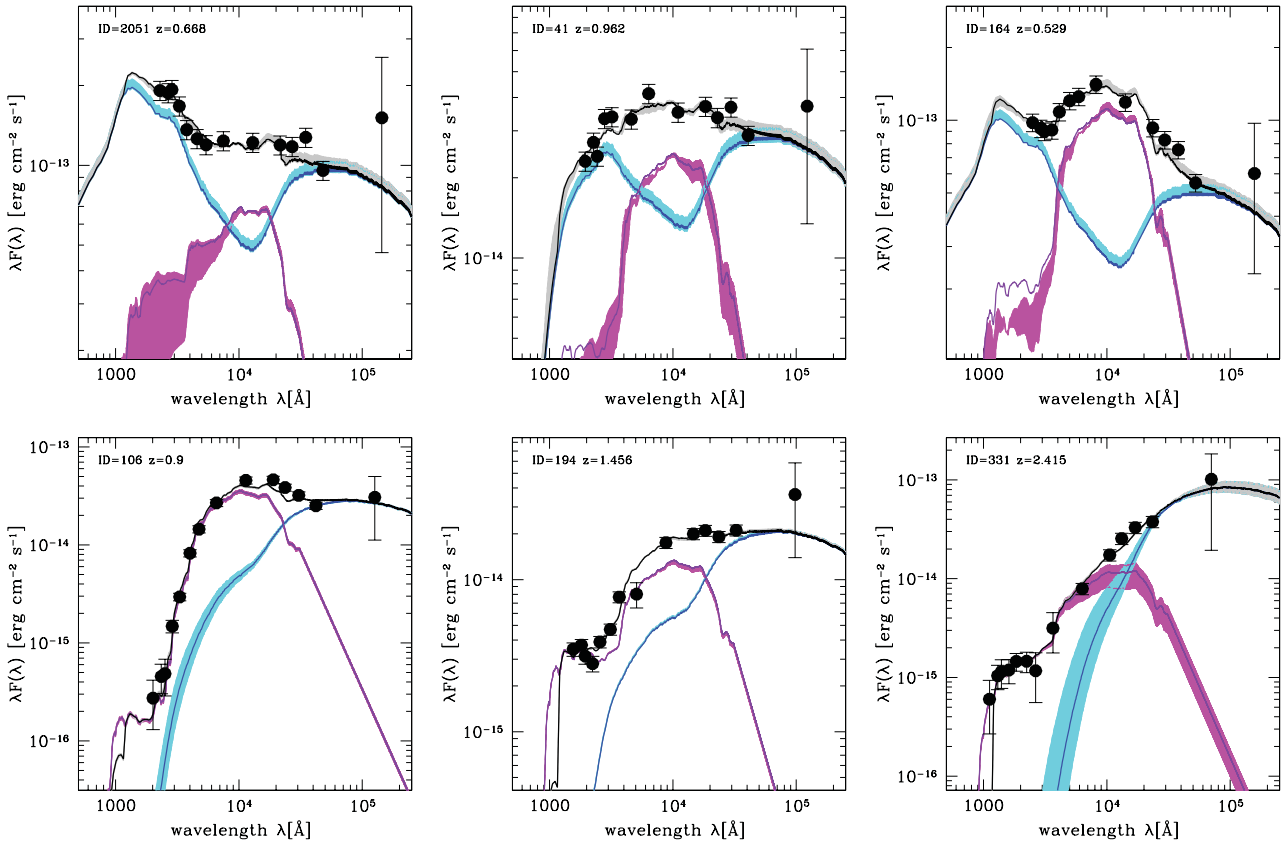

Figure 3. Examples of SED decompositions for unobscured (upper panels) and obscured (bottom panels) AGNs. Black circles are rest-frame fluxes corresponding to the observed bands used to constrain the SED. Purple and blue lines correspond, respectively, to the galaxy and the AGN template found as a best-fitting solution through χ^2 minimization, while the black line shows their sum. Pink and cyan shaded areas show the range of the SED template library within 1σ of the best-fitting template; light grey denotes their sum.

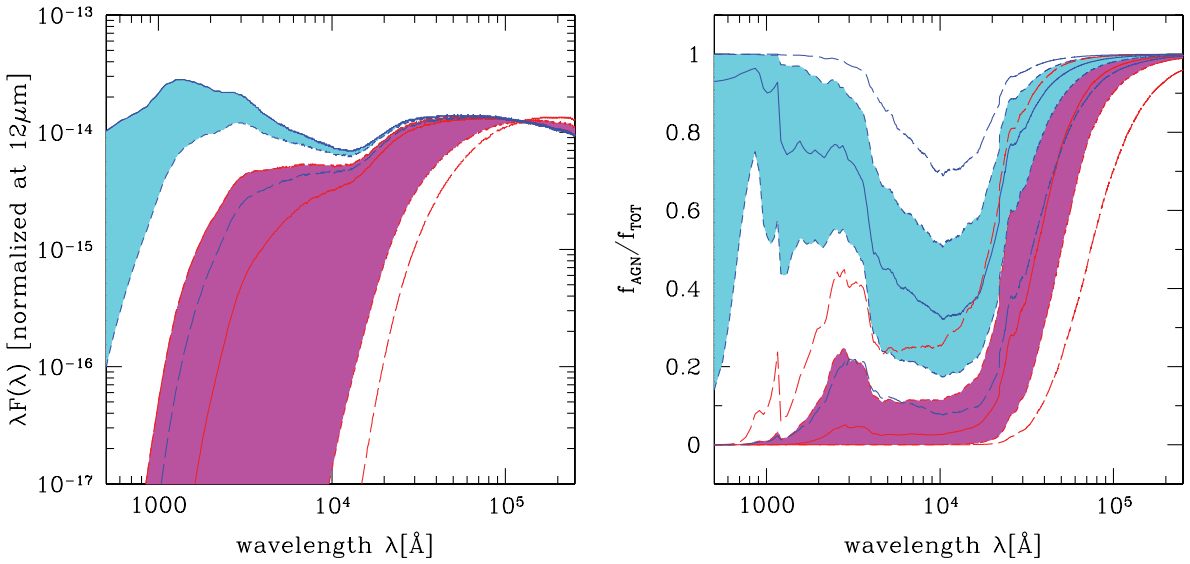

Figure 4. Left: median AGN SED chosen as the best fit for our sample of unobscured (blue continuum line) and obscured (red continuum line) AGNs, all normalized at rest-frame 12 μm. Shaded areas correspond to the range within the 25th and 75th percentiles, while long dashed lines show that within the 10th and 90th percentiles. Right: median AGN contribution to the total SED at different wavelengths for unobscured (blue continuum line) and obscured (red continuum line) AGNs. Shaded areas correspond to the range within the 25th and 75th percentiles, while long dashed lines show that within the 10th and 90th percentiles.

and SFRs of the galaxies. The 1σ errors on the best-fitting parameters are computed by considering the range of values corresponding to the solutions for which $\Delta\chi^2 = \chi^2(\mathrm{sol}) - \chi^2(\mathrm{best}) \leq 1.0$, corresponding to 1σ in the case of one parameter of interest (Avni 1976).

In some cases, we are not able to derive a measurement for the galaxy or the AGN components, but we can still define a meaningful upper limit on the unconstrained component. In unobscured type 1 AGNs, if the rest-frame K-band luminosity of the host galaxy is smaller than 10 per cent of the AGN luminosity in the same rest-frame band, we assign an upper limit to the SED component of the host galaxy. This is the case for 34 type 1 AGNs for which the value of the upper limit of the galaxy K-band luminosity is taken as the highest possible value of the parameter within the uncertainty. In these cases, we also assign an upper limit to the object's total stellar mass, assuming the median mass-to-light ratio of all other objects in the sample. No rest-frame magnitudes in the U and B bands are derived in these cases.

However, for 279 obscured type 2 AGNs that are not detected in the 24-μm MIPS band, we can provide only upper limits on the AGN component (and thus on the AGN luminosity). This is because, in heavily obscured sources, the AGN component rises almost exclusively at near-infrared (NIR) wavelengths. Without the 24-μm detection, it is not possible to determine reliably the AGN contribution to the optical/IR SED. For these sources, we have derived an upper limit on the 12-μm rest-frame AGN luminosity (see below) using the 24-μm detection limit and applying the appropriate k-correction. Here, we note that, in these cases, we can still use the observed X-ray (or emission-line) luminosity to infer the level of AGN strength in the source.

Summarizing, from the initial sample of 1702 sources, the SED fitting was successfully applied to 1679 sources. Of these objects, 34 type 1 AGNs have only an upper limit in the parameters of the galaxy component (mass and rest-frame M_K) while no U- and B-band rest-frame magnitudes were assigned. For 279 type 2 AGNs, only an upper limit in the parameters of the AGN component ($L_{12.3\mu\mathrm{m}}$ and thus L_{bol}) could be derived. All the derived parameters are available online (see Supporting Information). An example of the published table can be found in Table C2.

3.4 AGN–galaxy mixing diagram

Based on our SED fitting procedure described in Section 3.3, Fig. 5 shows the rest-frame NIR and optical slopes of the total best-fitting

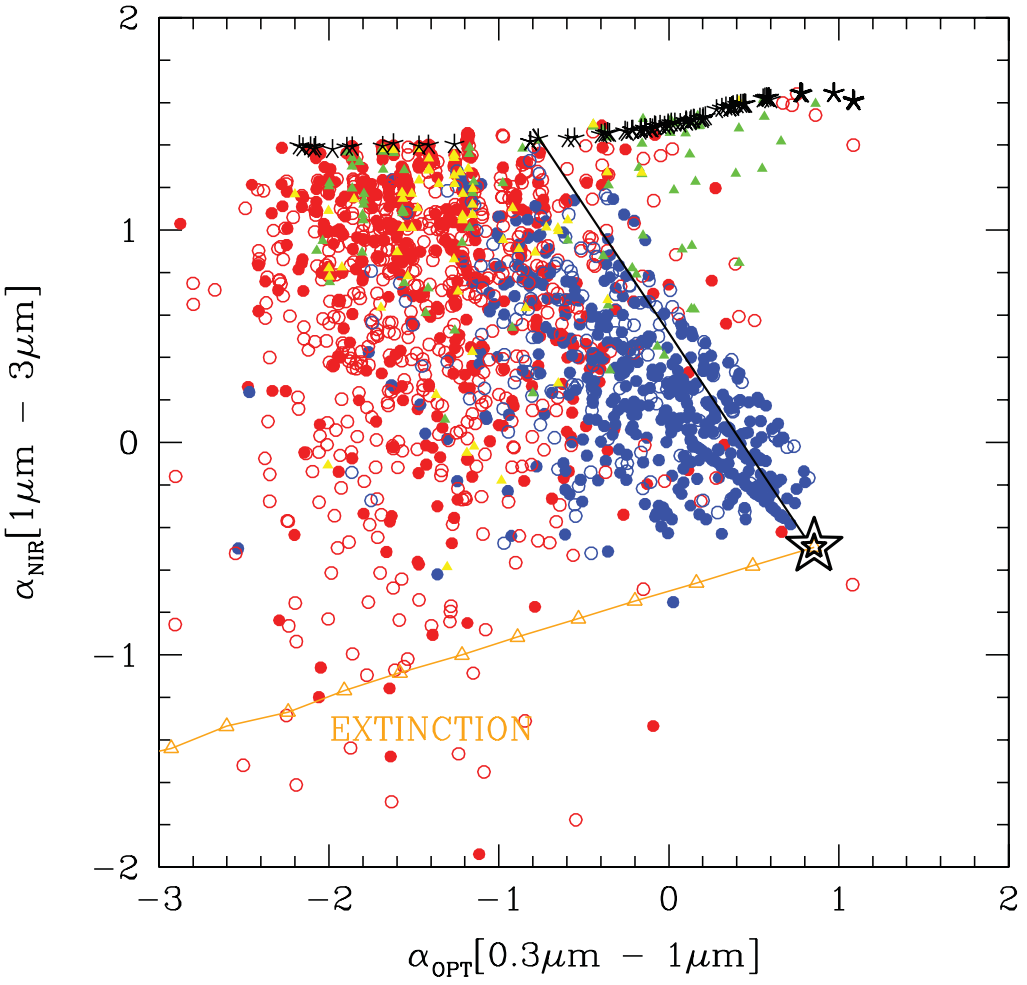

Figure 5. Mixing diagram from Hao et al. (2010), showing the slope plot of the entire sample of analysed AGNs. Circles correspond to *XMM*–COSMOS obscured (red) and unobscured (blue) AGNs; green and yellow triangles show the AGNs optically selected using the BPT diagram and the Ne v line, respectively. Open and filled circles correspond to photometric and spectroscopic redshifts, respectively. The big star shows the R06 mean SED, while the asterisks show different BC03 galaxy templates. The black line indicates the boundaries of the possible slopes obtained by mixing the R06 template with different fractions (0–100 per cent) of a given galaxy template. The orange line shows the reddening vector of R06 with the triangles corresponding to steps of 0.1 in $E(B-V)$.

SED of the sources in our sample (adapted from Hao et al. 2010). For each source, we plot the slopes on either side of the rest-frame 1 μm, which is approximately the inversion point in the typical QSO SED in a $\nu f(\nu)$ versus ν representation (i.e. where the exponent of a local power-law fit to the SED changes sign). In particular, we have computed the IR slope ($\alpha_{\rm NIR}$) between 1 and 3 μm, and the optical slope ($\alpha_{\rm OPT}$) between 0.3 and 1 μm (for more details, see Hao et al. 2010; Elvis et al. 2012). The locations of the sources in this colour–colour space are shown in Fig. 5, where circles represent *XMM*–COSMOS obscured (red) and unobscured (blue) AGNs, while triangles correspond to the optically selected sources without an X-ray counterpart (green for AGNs selected with the diagnostic diagrams and yellow for Ne V-selected AGNs). Open and filled circles correspond to photometric and spectroscopic redshifts. Note that the same symbols and colour code are used throughout the paper.

In Fig. 5, the big star indicates the SED of our pure type 1 AGN template, for which $\alpha_{\rm OPT} = 0.85$ and $\alpha_{\rm NIR} = -0.49$ (R06). Starting from this point, the orange line shows the R06 template when reddening is applied; the orange triangles correspond to steps in $E(B-V)$ of 0.1. The asterisks show the positions in this diagram of different galaxy templates derived from the stellar population synthesis models of BC03: from passive to extreme starbursts going from left to right. The black line indicates the R06 template when mixed with different fractions (0–100 per cent) of a given galaxy template. This mixing curve defines the boundaries of the possible slopes obtained by mixing the R06 SED with this particular galaxy template. This plot shows the three main components of the typical SED of the analysed sources (AGN, host galaxy and extinction) and their relative contributions to the total SED. It serves as a powerful illustration of the level of contamination of the optical/IR SEDs of AGNs, at least at the luminosity levels probed by the COSMOS. As seen in Fig. 5, for type 1 unobscured AGNs the galaxy contamination is less predominant than for type 2 obscured AGNs, where high values of extinction are associated with high contamination of the host galaxy, which can dominate in the 0.3–3 μm interval. Moreover, there are no objects along the AGN extinct line at $E(B-V) > 1$, showing that when the AGN absorption is high, the galaxy contribution becomes important.

Most of the sources lie on the left-hand side of the $\alpha_{\rm OPT}$–$\alpha_{\rm NIR}$ plot, and only a small percentage (and only obscured AGNs) lie in the upper-right corner. This is because, in nature, galaxies with high SFRs also show typically high levels of obscuration. This means that, while we do include non-extinct starburst templates in our library, these solutions are almost never chosen because they are unrealistic (unless $t_{\rm age}$ or $\tau/t_{\rm age}$ are very low, as is the case for the few points in that region of the plot).

Unobscured AGNs in this region of the plot have been interpreted by Hao et al. (2010) as hot dust-poor (HDP) quasars (i.e. normal quasars with relatively weak IR bumps). However, for the obscured AGNs, another possible explanation is that these sources are normal AGNs hosted by starburst galaxies, which steepen the UV slope. Looking at the SEDs of all these sources one by one, we have found that all but four of them are not detected at 24 μm and their SED in the UV–optical part is dominated by a starburst galaxy. Most of these sources are Seyfert 2 galaxies optically selected using the BPT diagrams. Because of the errors associated with the line measurements, some of them can be misclassified normal star-forming galaxies. Alternatively, they could be very faint AGNs hosted in starburst galaxies.

3.5 Measuring AGN host stellar masses and star formation rates

3.5.1 Stellar masses

The SED fitting procedure allows us to estimate the total stellar mass and the SFR of the AGN host galaxies. Knowing the IMF, for a given SFH, each combination of τ and $t_{\rm age}$ (see Section 3.1) is, in fact, uniquely associated with the value of the specific star formation rate (sSFR = SFR/M; see Section 4.4) and, given the normalization of the template, the total stellar mass.

Thus, for our AGN COSMOS sample, we have derived a robust estimate of the host galaxy stellar masses for 1650 objects (out of a total number of 1702 attempted SED fits). As explained in Section 3.1, for 34 type 1 AGNs, we could derive only an upper limit on the stellar mass, while for 18 objects the SED fit failed because of problems in the photometry or in the SED shape, so no mass estimation was possible.

The top-left panel of Fig. 6 shows the stellar mass distribution of the host galaxies for the entire AGN sample (open histogram), divided into type 1 (hatched blue) and type 2 (hatched red) objects. The bottom panel shows a detail of the M_* distribution in the different subsamples, colour-coded as usual. Masses range mainly from 10^{10} to $10^{11.5}$ $\rm M_\odot$ with a peak at $10^{10.9}$ $\rm M_\odot$. No significant difference is found between the type 1 and type 2 AGN host population, which on average show the same host galaxy mass distribution. Interestingly, the hosts of optically selected Seyfert 2 galaxies without X-ray counterparts show a tail in the distribution at lower masses down to $10^{8.5}$ $\rm M_\odot$.

The right panel of Fig. 6 shows the stellar masses of the AGN hosts as a function of redshift, superimposed on the underlying distribution of stellar masses of non-AGNs in the COSMOS area, drawn from Ilbert et al. (2010) and matched to the AGN population in redshift and i-band apparent magnitude (see Section 2.4).

3.5.2 Star formation rate

The measurement of the host galaxy SFR is subject to considerably larger uncertainties than that of stellar mass, in particular for type 1 AGNs. In general, the most reliable way to compute the SFR is to sum up the rate of unobscured SF (emitted in the UV) with the obscured SF re-emitted in the (mid- and far-) IR by dust ($\rm SFR_{UV+IR}$). Our optical/IR SED fitting procedure relies only on the UV emission measurements, which traces only the unobscured SF, which is however scaled up by the dust correction factor computed using the full SED to account for the latter contribution.

The issue of the reliability of UV-based indicators of SFRs in galaxies has long been debted. With the advent of the *Herschel* telescope, sensitive FIR measurements of cold dust heated by star-forming processes have become available for deep and medium-deep surveys of the extragalactic sky, providing the first direct handle of $\rm SFR_{UV+IR}$ in large samples of galaxies (e.g. Rodighiero et al. 2010). For complete samples of AGNs selected in both X-rays and optical spectroscopic surveys, however, the final answer to the question of what are the intrinsic star-forming properties of the AGN hosts requires SFR indicators that are sensitive to both star-forming and passive/quiescent galaxies. This is not possible even for *Herschel*, whose sensitivity to individual galaxies does not reach into the quiescent population at the redshifts ($z > 1$) of deep extragalactic surveys, such as COSMOS (e.g. Shao et al. 2010;

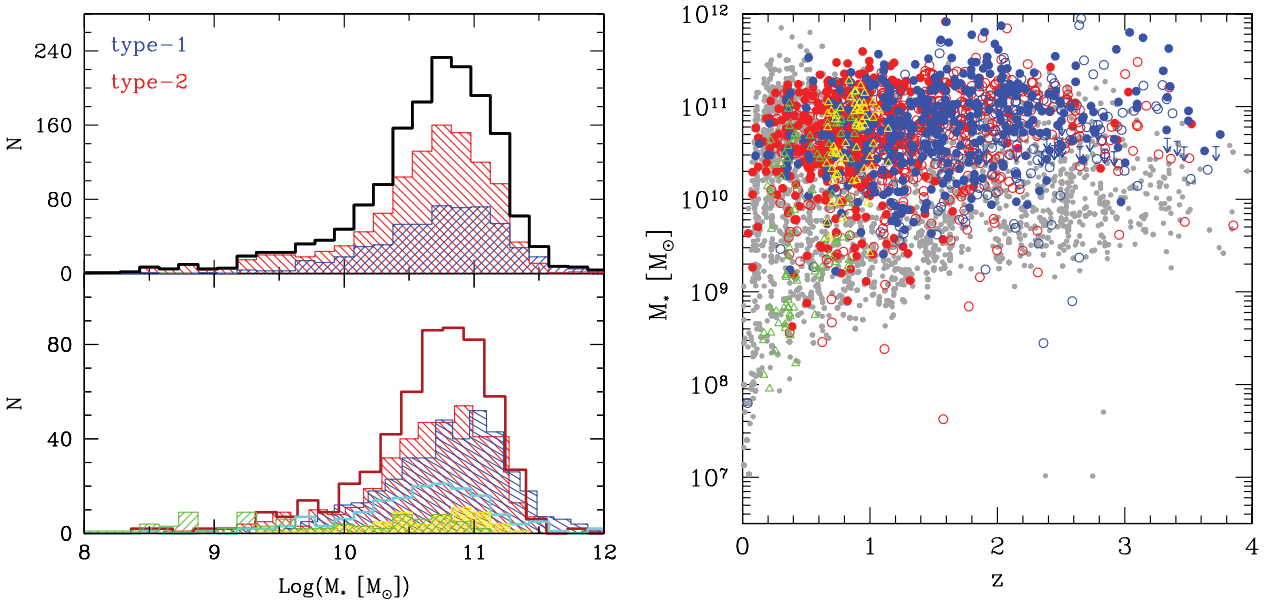

Figure 6. Left: in the lower panel, we show the stellar mass distribution of the AGN host galaxies for the different analysed subsamples: shaded blue and red for spectroscopically confirmed X-ray type 1 and type 2 AGNs, respectively; empty cyan and red for X-ray selected type 1 and type 2 AGNs, respectively, with photometric redshift; dashed green and yellow for AGNs optically selected using the BPT diagram and the Ne v line, respectively, without X-ray detection. The top panel shows the entire sample (thick open histogram), also divided into type 1 (shaded blue) and type 2 (shaded red) AGNs. Right: host galaxy stellar masses versus redshift for the entire AGN sample with the usual colour-coded subclass division. We also show in grey the distribution of masses for the parent galaxy population matched using the *i*-band apparent magnitude.

Mullaney et al. 2012). Indeed, only ∼10 per cent of the AGNs in our sample are detected above the 3σ limits of the PEP catalogues in the COSMOS field. In Appendix B, we compare the SFR estimated from the SED fitting with the one derived using other methods (i.e. FIR observation and optical emission lines). We find that above $\sim 20\,M_{\odot}\,yr^{-1}$, the agreement between the SFR derived from the SED and the one derived from the FIR bands is quite good, while at lower SFR ($<20\,M_{\odot}\,yr^{-1}$) the disagreement becomes evident, with SFR(FIR) being systematically higher than SFR(SED). We have interpreted this discrepancy as the combination of a possible contamination from SF in the NIR not properly taken into account in the SED fitting, together with a non-negligible AGN contamination to the FIR bands not taken into account in the FIR SFR estimates(e.g. Rosario et al. 2012). This would imply that, while our SFR estimates can in some cases be underestimated, the FIR SFR can be overestimated, especially at lower SFRs where the AGN contamination would represent a significant fraction of the IR emission.

In any case, the very nature of our objects and the SED decomposition technique, with which we analyse them, require some special care in the interpretation of the SFR measurements. The main problem with type 1 (unobscured) AGNs is that the accretion disc emission strongly contributes in the UV band (see the upper panels of Fig. 3), introducing a degeneracy in the SED fitting between the UV emission from SF and from the central AGN. In practical terms, this means that, in some cases, the SED fitting can give two significantly different solutions with a similar χ^2: an unobscured AGN, with a prominent big blue bump (BBB) dominating in the UV range, with a passive/moderately star-forming galaxy or, conversely, a moderately obscured AGN with a strongly star-forming galaxy, which substantially contributes to the observed UV emission. Because it is not possible to disentangle, a priori, this kind of degeneracy, we do not take into consideration the SFR measurements of type 1 AGNs in the further discussion.

In contrast, because in obscured AGNs the UV emission is suppressed by obscuration (see, as an example, the bottom panels of Fig. 3), the UV range is clean of AGN contamination and a reliable estimate of the SFR can be derived from the SED.

Fig. 7 shows the host galaxy SFR distribution in three redshift bins for 1090 obscured type 2 AGNs (thick solid line) and for the different subsamples with the usual colour code. AGN host galaxies span a very wide range of SFRs, with X-ray selected AGNs populating the entire range, while optically selected Seyfert 2 galaxies (which are mainly at low redshifts) are hosted predominantly in galaxies with low SFRs (but very low masses, as shown in Fig. 6). This can highlight a physical difference in the hosts of optically selected Seyfert 2 galaxies, or it can be, at least partly, driven by the fact that a high SFR implies strong emission-line dilution, which causes a certain fraction of AGNs to be missed by the emission-line ratio (BPT) selection.

3.6 Measuring nuclear $L_{12\,\mu m}$ and bolometric luminosities

Before proceeding to a comprehensive discussion of the observable (and physical) properties of the host galaxies of the AGNs selected in COSMOS, we present some basic intrinsic properties of the AGN population itself. This will also be a useful reminder that, in terms of the physical characteristics of growing black holes and of their immediate (nuclear) environments, every large survey selects families of objects occupying specific parts of the generic parameter space identified by the distance of the source, its accretion luminosity, its BH mass and the level of nuclear obscuration.

To begin with, we note that from the best-fitting (obscured or unobscured) AGN component in the SED, we can derive an estimate of the AGN luminosity at (rest frame) 12 μm, which can be used as a reliable estimator of the AGN bolometric luminosity. The intrinsic AGN 12-μm luminosity is calculated by integrating the AGN component obtained from the best-fitting SED over a narrow

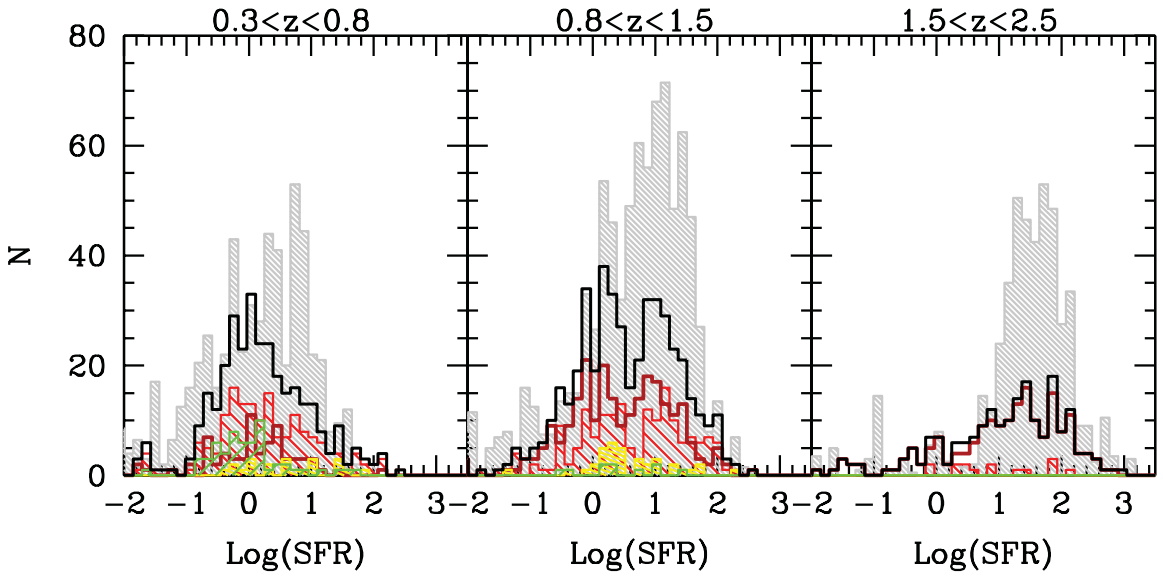

Figure 7. SFR distributions in three redshift bins of the AGN host galaxies for the entire type 2 AGN sample (thick open histogram) and divided into different subsamples: shadow and empty red for X-ray spectroscopically confirmed AGNs and with only photometric redshift type 2 AGNs, respectively; dashed green and yellow for AGNs optically selected using the BPT diagram and the Ne v line, respectively, without X-ray detection. The grey histogram shows as a comparison the distribution of the mass-matched parent sample. To make the comparison easier, the latter has been divided by a factor of 2.

(i.e. 1 μm) passband centred on 12 μm. By construction, and given the shape of the BC03 templates used to fit the galaxy component in every system, the MIR band is assumed to be completely dominated by the nuclear AGN light reprocessed in the obscuring molecular structure on a ∼1–10 pc scale. Such a drastic assumption is not likely to be correct for a number of individual objects, which might have a significant galaxy contribution (not included in the BC03 templates) at 12 μm. However, globally, this is not unreasonable for the studied AGN sample, as suggested in Fig. 8, where the AGN 12-μm luminosity is plotted against the intrinsic (de-absorbed) X-ray [2–10] keV luminosity[6] for the *XMM* selected sample of ∼1550 objects with X-ray detections and $L_{12\,\mu m}$ measurements.

Using a sample of local Seyfert galaxies (both Seyfert 1 and Seyfert 2) at high spatial resolution, Gandhi et al. (2009) found that the uncontaminated nuclear MIR continuum of an AGN closely correlates with the X-ray [2–10] keV AGN emitted powers over three orders of magnitude in luminosity. To be specific, the best-fitting correlation they obtained, considering only the 22 well-resolved sources, is shown as a thin green line in Fig. 8: $\log(L_{MIR}/10^{43}) = (0.19 \pm 0.05) + (1.11 \pm 0.07)\log(L_X/10^{43})$. Our data follow the Gandhi relation. As a comparison, the fit to our data, performed adopting the ordinary least-squares bisector estimate recommended by Isobe et al. (1990), which gives identical weighting to all sources, gives $\log(L_{MIR}/10^{43}) = (0.10 \pm 0.03) + (1.24 \pm 0.03)\log(L_X/10^{43})$ for the entire sample and $\log(L_{MIR}/10^{43}) = (0.29 \pm 0.03) + (1.13 \pm 0.03)\log(L_X/10^{43})$ if we consider only the spectroscopic sample. The fits are shown as black and grey solid thick lines, respectively, in Fig. 8, together with the fit for the type 1 (blue long-dashed line) and type 2 (red short-dashed line) AGN subsamples. The good agreement between the predicted and the measured 12.3-μm luminosities is also shown in Fig. 9, where we show the histogram of the ratio between the measured 12.3-μm luminosity and the one predicted by the Gandhi et al. (2009) relation for type 2 (top panel) and type 1 (bottom panel) AGNs. We note that, while the distribution of the ratio between the measured and the predicted MIR luminosity

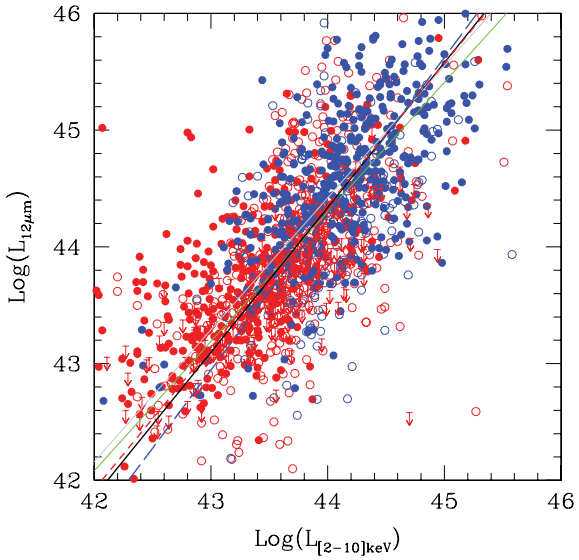

Figure 8. X-ray [2–10] keV versus MIR 12-μm luminosities for the sample of ∼1550 AGNs with X-ray detection and measured $L_{12\,\mu m}$. Red circles are spectroscopic (filled) and photometric (open) type 2 AGNs, while type 1 AGNs are represented with filled (spectroscopic) and open (photometric) blue circles. The thin solid green diagonal line is the [2–10] keV versus 12.3-μm luminosity relation found by Gandhi et al. (2009) for the sample of 22 well-resolved local Seyfert galaxies. The thick solid lines are the fitted correlation considering our entire sample (black) and only the spectroscopic sources (grey), while the long-dashed blue and short-dashed red lines correspond to the type 1 and type 2 subsamples.

[6] We calculate this on the basis of the observed fluxes in the soft (0.5–2 keV) and hard (2–10 kev) X-ray energy bands, of the known redshift, and assuming, as the X-ray spectral model, an intrinsic power law of slope $\Gamma = 1.9 \pm 0.2$, absorbed by neutral gas.

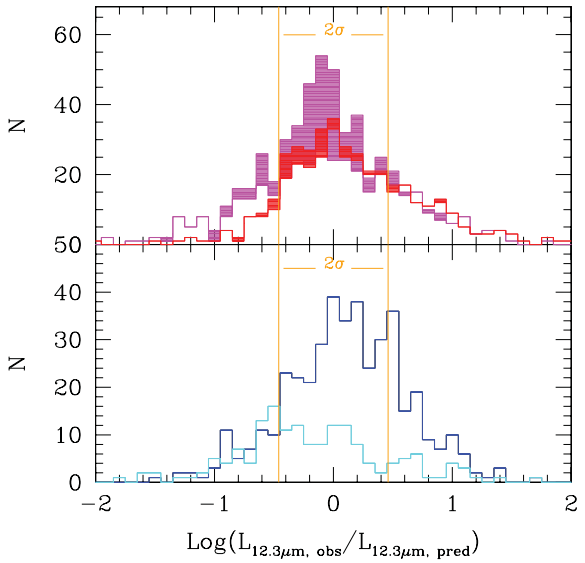

Figure 9. Histograms of the ratio between the measured 12.3-μm luminosity and that predicted by the Gandhi et al. (2009) relation on the basis of the observed 2–10 keV X-ray luminosity. In the top panel, we show type 2 AGNs, with the red and purple histograms marking objects in the spectroscopic and photometric redshift samples, respectively; filled histograms are for upper limits. In the bottom panel, we show the type 1 AGNs, with the dark and cyan histograms marking objects in the spectroscopic and photometric redshift samples, respectively. Yellow histograms in both panels show the outliers in the mixing diagram of Fig. 5 (see text for more details). The two solid orange lines delimit 2σ from the Gandhi relation (mean equal to zero and standard deviation $\sigma = 0.23$).

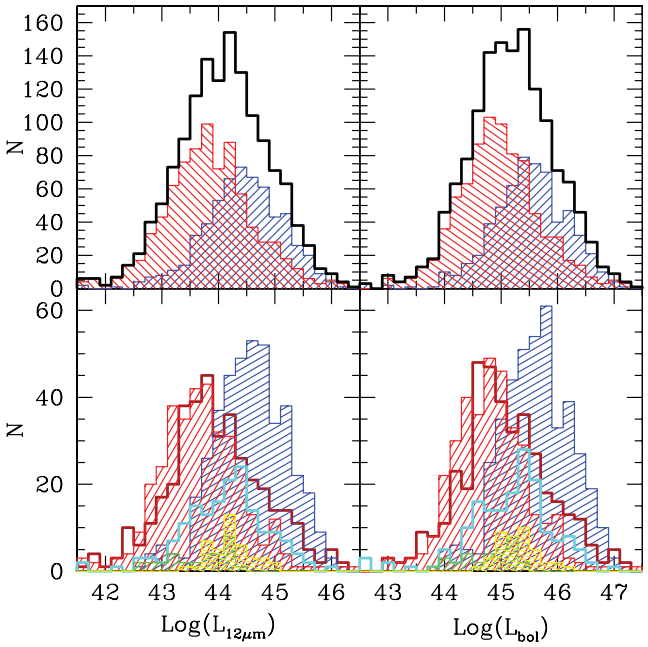

Figure 10. Upper panels: MIR and bolometric luminosity distribution of the entire sample (thick open histogram) and for the type 1 (shaded blue) and type 2 (shaded red) AGN subsamples separately. Bottom panels: MIR and bolometric luminosity distribution of the different subsamples included in our AGN sample: shaded blue and red for X-ray spectroscopically confirmed type 1 and type 2 AGNs, respectively; empty cyan and red for X-ray selected type 1 and type 2 AGNs, respectively, with photometric redshifts; dashed green and yellow for AGNs optically selected using the BPT diagram and the Ne v line, respectively, without X-ray detection.

is centred at zero,[7] the observed scatter is much larger than the one measured by Gandhi et al. (2009) (2σ from the Gandhi relation is delimited with orange lines; mean equal to zero and standard deviation $\sigma = 0.23$). If we consider only the sources with a spectroscopic redshift (black solid line), we do see a clear asymmetry in the distributions, as expected because of the unaccounted contribution of SF-related emission in the NIR (see also Lusso et al. 2011). Indeed, while ∼61 per cent of the objects are found within 2σ of the distribution, the tail extending to lower values includes 94 sources (12 per cent), while for a higher number of sources (211; 27 per cent of the sample) the observed MIR luminosities are significantly higher than those predicted. However, this is not the case for the photometric sample. A possible explanation of this effect might involve the effect of the uncertainties in the photometric classification in some of the AGNs for which we miss spectroscopic information. Indeed, we have tested that all of the small fraction of outliers in Fig. 5 (i.e. type 1 AGNs with $\alpha_{\rm OPT} < -1$ and type 2 AGNs with $\alpha_{\rm OPT} > -1$) show a broader distribution in the $L_{12.3\,\mu{\rm m,obs}}/L_{12.3\,\mu{\rm m,pred}}$ ratio, with a slight skew towards values less than unity.

Some of the observed differences can also be attributed to the fact that we are extending the Gandhi relation, which was derived for a sample of local Seyfert galaxies, to a sample spanning a much wider range of luminosity and redshift. However, it is also plausible that in a fraction of these sources, the SED-fitting procedure overestimates the nuclear contribution, which results in a higher $L_{12.3\,\mu{\rm m,obs}}$ compared to the predicted $L_{12.3\,\mu{\rm m,pred}}$. How much of this scatter is inherently a result of the physical conditions of the AGN and torus clouds, as compared to observational selection effects, remains an important unresolved issue, which is beyond the scope of this paper. However, the comparison points out that, on average, the 12.3-μm luminosity derived from the SED fitting is a reasonably good measure of the MIR AGN luminosity. Indeed, the assumption that the MIR emission is dominated by the AGN emission because of accretion on to the central black hole rather than SF from the host galaxy is plausible for most of the sources.

One clear advantage of adopting the rest-frame MIR luminosity, $L_{12\,\mu{\rm m}}$, as a universal AGN power estimator for our sample is that, contrary to the [2–10] keV luminosity, this quantity was measured homogeneously for all the sources, including the objects with no X-ray emission. Hence, it can be used to compare the AGN power in the different subsamples and to derive, in a homogeneous way, the bolometric luminosity for the entire sample.[8] This is shown in Fig. 10 where both the 12-μm and the derived bolometric luminosityusing the Hopkins et al. (2007) conversions are shown for the different subsamples (bottom panels) and for the full sample divided into obscured and unobscured AGNs (upper panels). Bolometric luminosities range from 10^{43} to 10^{47} erg s^{-1}.

Optically selected Seyfert 2 galaxies (green histogram) are among the faintest sources in our sample extending down to $L_{\rm bol} = 10^{43}$ erg s^{-1}. Because their 12-μm luminosities are of the same order as the X-ray selected obscured sources, the lack of X-ray detection could be because these sources lie in the tail of the Gandhi relation

[7] This contrasts with Lusso et al. (2011), who found a shift of 0.2 towards higher values when considering the ratio between the total (AGN+GAL) SED and the Gandhi predicted values.

[8] However, for a more detailed study of the bolometric luminosities of *XMM*–COSMOS AGNs, we refer the reader to Lusso et al. (2010, 2011).

(Fig. 9). Alternatively, some of these sources can be Compton-thick AGNs, as shown by Gilli et al. (2010) for the [Ne v]-based sample. Because of the high $N_{\rm H}$ absorption, they are not detected in the X-ray band.

4 PHYSICAL PROPERTIES OF COSMOS AGN HOST GALAXIES

4.1 Rest-frame colours

The unique strength of the COSMOS AGN sample at hand lies with the possibility, by means of our SED fitting technique, of reliably removing the AGN component from the overall SED of any AGN host. In this section, we discuss the rest-frame observables (absolute magnitudes and colours) of the AGN host-galaxy component, which are, by construction, little contaminated by the AGN colours. We also investigate the properties of the AGN host galaxies compared to normal galaxies.

Fig. 11 shows the rest-frame $U - B$ colours versus the B-band absolute magnitude (the colour–magnitude diagram, CMD) for our sample of AGN hosts, divided into three different redshift bins: $0.3 < z \leq 0.8$, $0.8 < z \leq 1.5$ and $1.5 < z \leq 2.5$. The grey contours show, for comparison, the whole galaxy sample drawn from Ilbert et al. (2010) in the same redshift intervals, while grey points correspond to the mass-matched parent sample described in Section 2.4. We have chosen to show both the measured values of the $U - B$ colours (upper row) and the extinction-corrected values, where the value of the extinction is that obtained for the galaxy component from the best-fitting SED decomposition (see Section 3.2.1). This is an important test, because galaxies with red colours can be either dusty star-forming galaxies or galaxies with old stellar populations. Any interpretation of the locations of AGN hosts in the CMD can

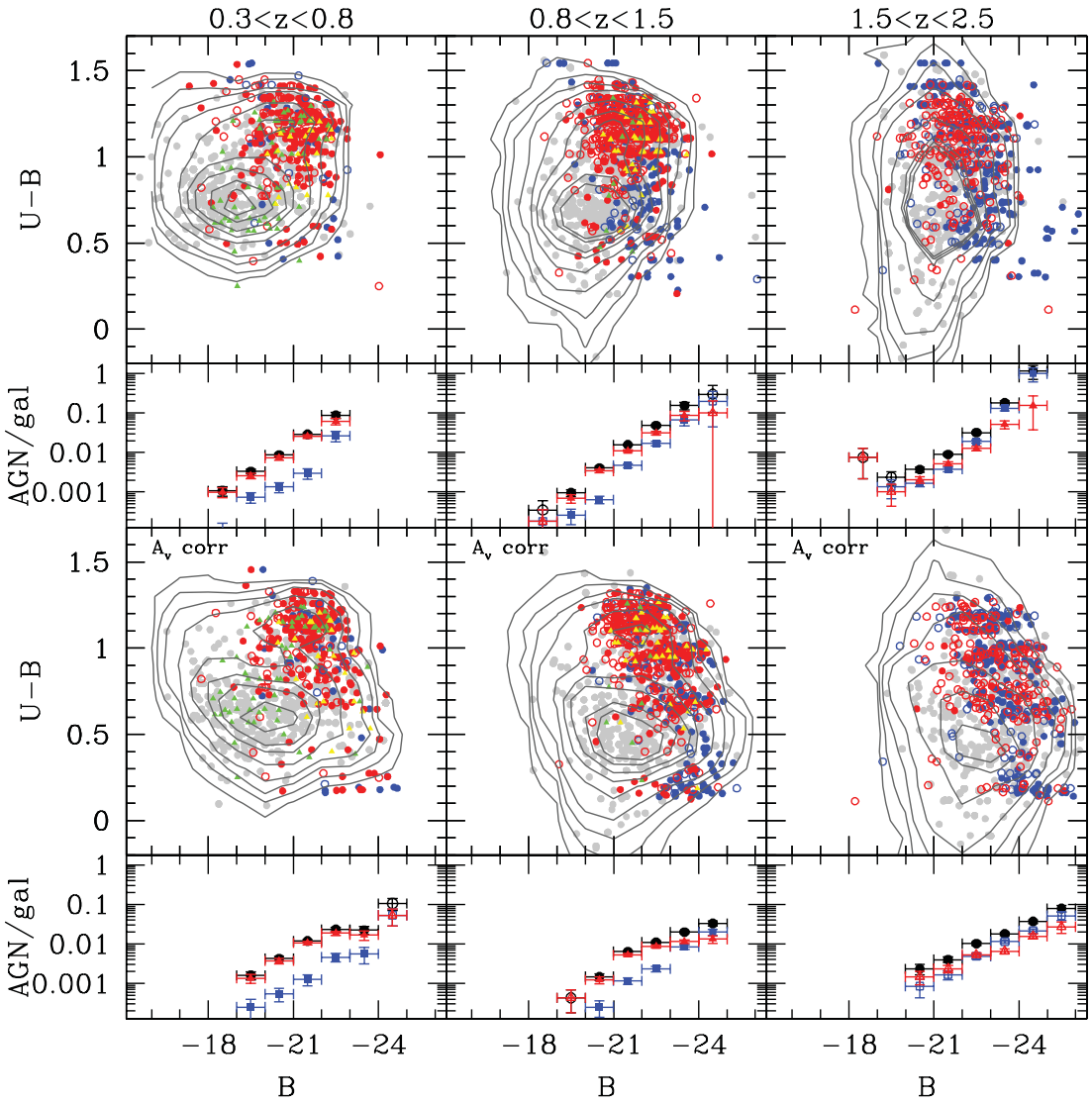

Figure 11. Upper panels: $U - B$ versus $M_{\rm B}$ CMD of the AGN host galaxies of unobscured (blue circles) and obscured (red circle) X-ray AGNs and optically selected (green and yellow triangles) AGNs in three redshift bins. Open and filled circles correspond to photometric and spectroscopic redshifts, respectively. Grey contours show the whole galaxy sample drawn from Ilbert et al. (2010), while grey points correspond to the mass-matched parent sample described in Section 2.4. Lower panels: same as above but now $M_{\rm B}$ and $U - B$ are dereddened using $E(B - V)$ derived from the SED fit. The bottom part of each panel shows the fraction of galaxies with an AGN as a function of B magnitude, considering all AGNs (black circles), only type 1 AGNs (blue squares) and only type 2 AGNs (red triangles). For these fractions, we consider only X-ray selected AGNs and the whole galaxy sample.

be strongly affected by the treatment of dust extinction in these systems (e.g. Cardamone et al. 2010). Our rich and comprehensive multiwavelength data set is very well suited to attempt a robust determination of the large-scale extinction properties of the galaxies in the various samples.

First of all, a note of caution is needed because, at any given host-galaxy magnitude, the colours of the type 1 hosts are more and more unreliable, the higher the nuclear AGN luminosity. In this respect, very blue colours of bright type 1 AGNs (see, in particular, the clusters of sources at $U - B \simeq 0.1$–0.3 in all extinction-corrected CMDs) should not be trusted too much, because they most likely indicate objects for which the SED decomposition, and thus the AGN blue light subtraction, might give uncertain and possibly unreliable results. We return to this point when discussing the inferred SFR properties of the AGN hosts.

We now discuss the CMD of the general population of AGN hosts. Underneath each of the diagrams in Fig. 11, we show the fraction of all AGNs (black), type 1 AGNs (blue) and type 2 AGNs (red) relative to the whole galaxy population in the corresponding redshift bin as a function of *B*-band magnitude. For consistency, in the computation of these fractions, we have restricted the analysis to the X-ray detected sample, the completeness of which can be properly assessed in a relatively simple way (see Section 4.3). However, we note that the only effect of the pure optical sources is to slightly flatten the slope at lower redshifts. In agreement with previous studies (Nandra et al. 2007; Brusa et al. 2009b), the AGNs are almost exclusively hosted in bright ($M_{\rm B} < -20$) galaxies and the fraction of galaxies hosting an AGN increases with higher luminosities.

No significant difference is found in the host's colours of unobscured and obscured AGNs, in contrast with what is expected from the most popular models of AGN fuelling and feedback (e.g. Hopkins et al. 2008). However, this agrees with the standard unified model (Antonucci 1993). As a matter of fact, the relative fractions of type 1 and type 2 AGNs change only slightly with *B*-band luminosity in each redshift interval, but the overall fractions of obscured and unobscured AGNs are markedly different in the three redshift bins. This is an effect introduced by the well-known anticorrelation between obscured AGN fraction and luminosity (e.g. Hasinger 2008) in combination with the almost constant (X-ray) flux limit of the *XMM*–COSMOS survey, which implies that high-redshift sources are more luminous than low-redshift sources. A comprehensive discussion of the incidence of obscured and unobscured AGNs in the COSMOS field as a function of luminosity, redshift and host-galaxy properties will be presented in a separate work (Merloni et al., in preparation).

We find no evidence of colour bimodality for AGN hosts, which populate mainly the red and intermediate parts of the bimodal colour–magnitude distribution observed for normal galaxies. Interestingly, at $0.3 < z < 0.8$, the low-luminosity, optically selected (from the BPT diagram) Seyfert 2 galaxies without X-ray counterparts (green triangles) seem to behave in a different way; they are mainly hosted in galaxies populating the blue and the intermediate regions. The effects of dust-extinction correction on the AGN distribution in the CMD are clearly visible in the lower panels of Fig. 11, where some objects populating the red sequence have moved to the intermediate $U - B$ region, and some that were in the intermediate region now lie in the blue region (see also Cardamone et al. 2010). However, in contrast to what was found by Cardamone et al. (2010), no clear bimodality is seen in any of the three panels, even after dust correction is applied. More importantly, it is clear from Fig. 11 that dust-extinction correction affects more severely the overall colour distribution of the parent galaxy sample. This can be seen in the fact that the AGN fraction in the *B*-band most luminous galaxies in the sample decreases once dust correction is included.

Indeed, at any redshift, AGN hosts are found mainly in red galaxies, albeit with a substantial tail towards intermediate and bluer colours (see also Fig. 12). In many previous studies, AGNs have been described as populating preferentially an intermediate region, the so-called 'green valley' (Silverman et al. 2009), with the possible interpretation that they live in quenched galaxies going through the transition phase between the active and passive SF phases. However, most of these works were focused on small subsamples of the AGN population. More importantly, the AGN component was not subtracted in the computation of the host colours, leading to a possible blue contamination from the AGN, which in fact would push red hosts into the green valley. Moreover, the net effect of the dust correction applied systematically, and in a uniform way, to both the parent and the AGN host-galaxy samples is to decrease the fraction of AGNs in the most star-forming objects.

In general, the often contrasting interpretations of many previous studies of the CMD of AGN hosts from deep multiwavelength surveys (Nandra et al. 2007; Silverman et al. 2009; Cardamone et al. 2010; Xue et al. 2010; Rosario et al. 2011) demonstrate that studying their position within optical CMDs is hardly a robust way to unveil the star-forming nature of AGN hosts. This is because the interpretation of optical colours in terms of SF properties of a galaxy can suffer from various, degenerate, biases, especially when dealing with distant AGN hosts. To counter such degeneracies, at least partly, in Section 4.4, we discuss in more detail how AGN hosts are distributed in the physical parameter space of stellar mass and SFR.

4.2 Host galaxy colour–mass diagrams

Fig. 12 shows the position of our AGN sources in the colour–mass diagram, where the colours ($U - V$) are corrected for extinction. Black contours show the whole galaxy sample drawn from Ilbert et al. (2010), while grey points correspond to the parent sample matched using the *i*-band apparent magnitude, as described in Section 2.4. Comparing with the whole galaxy population, we find that, at any redshift, AGNs are preferentially detected in massive galaxies, in agreement with previous studies (Brusa et al. 2009b; Silverman et al. 2009; Xue et al. 2010; Mullaney et al. 2012). The trend is still present if we use the sample matched using the *i*-band as a comparison sample. In Section 4.3, we investigate in greater detail the evolution of the AGN fraction as a function of the redshift, the X-ray luminosity and the stellar mass of the host, and of the ratio of these two latter quantities. However, we note, here, that except for few unobscured AGNs in the intermediate zt bin, showing extreme (and probably unreliable) blue colours, obscured and unobscured AGNs overlap very well in the colour–mass plane.

Because massive hosts tend to be redder, the trends with mass and with colours, seen in Section 4.1, should go in parallel. In many previous studies, it has been pointed out how, by removing the mass dependency (i.e. using a mass-matched control sample), the trend with colour disappears (Silverman et al. 2009; Xue et al. 2010; Mainieri et al. 2011; Rosario et al. 2011). However, this is not the case in our study. Fig. 13 shows the extinction-corrected $U - B$ colour histogram for AGN hosts in the usual three redshift bins, for three different ranges of stellar masses, all above the completeness threshold of the Ilbert et al. (2010) parent galaxy sample, divided into unobscured (shaded blue) and obscured (shaded red) AGNs and compared to the mass-matched galaxy sample. Galaxies hosting an

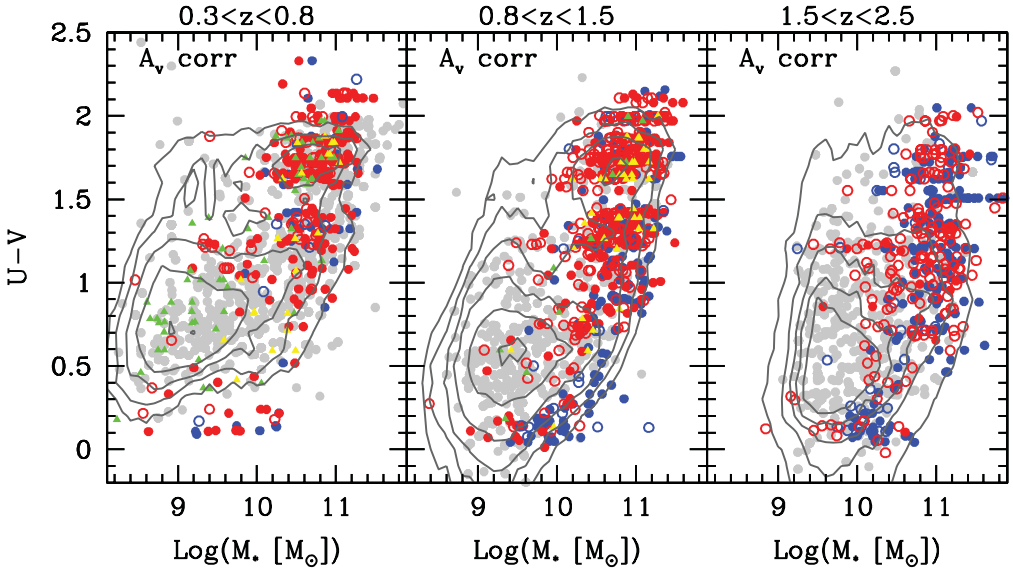

Figure 12. Extinction-corrected colour ($U - V$) versus mass diagram of the AGN host galaxies of unobscured (blue circles) and obscured (red circle) X-ray AGNs and optically selected (green and yellow triangles) AGNs in three redshift bins. Open and filled circles correspond to photometric and spectroscopic redshifts, respectively. Grey contours show the whole galaxy sample drawn from Ilbert et al. (2010), while grey points correspond to the parent sample matched using the *i*-band apparent magnitude, as described in Section 2.4.

AGN tend to be redder than normal galaxies of comparable stellar mass (the median colour shift ranges between ∼0.1 and ∼0.25).

Following our discussion in Section 4.1, we can speculate about the two main reasons for the discrepancy between our colour distributions of AGN hosts in mass-matched samples and previous studies. First of all, our SED decomposition technique allows a more reliable correction for AGN light pollution that would typically move the colour of an AGN host towards the blue, star-forming sequence. Secondly, because of the unique multiwavelength coverage of the COSMOS field, we are able to extinction-correct, in a systematic way, all galaxy colours, both in the parent sample and in the AGN host sample. We find that moving dusty star-forming galaxies back to the blue cloud with such a correction effectively decreases the incidence of AGNs among star-forming galaxies, while, conversely, the relative fraction of AGN hosts in truly passive galaxies is enhanced. However, we caution that, because these colour shifts are relatively small, systematic effects because of a non-optimal AGN–galaxy decomposition (AGN oversubtraction; see Section 3.3) or because of the specific choice of stellar population templates, could still be held responsible for some of this observed trend.

4.3 Incidence of AGNs as a function of the stellar masses of host galaxies

Let us now move to a quantitative assessment of the fraction of AGNs as a function of the stellar mass of the host galaxy. In order to do so, we need to take carefully into account the selection function of our survey. For this reason, in this section, we consider only the X-ray selected AGNs, taking advantage of the well-defined exposure map of the *XMM*–COSMOS survey (Cappelluti et al. 2009).

So far, one of the clearest pieces of evidence emerging from a variety of comparative studies of AGNs and host galaxies in large multiwavelength surveys is the fact that the likelihood of finding an AGN is a very strong function of the galaxy stellar mass, with AGN fractions (or, equivalently, duty cycles) increasing with the increasing stellar mass of the host galaxy. This is a particularly strong effect for radio-selected AGNs (Best et al. 2005; Smolčić 2009), but is has also been confirmed for X-ray selected AGNs (Brusa et al. 2009b; Xue et al. 2010; Mullaney et al. 2012).

4.3.1 *AGN fraction as a function of X-ray luminosity*

First of all, we compute the probability of finding an AGN within a logarithmic bin of a given X-ray (2–10 keV, absorption-corrected) luminosity, $L_{\rm X}$, in different bins of stellar mass, by taking the ratio of the number of AGNs to the measured stellar mass of the host in that range, divided by the number of galaxies (of all morphological, colour and SF activity types) having stellar mass in the same range in the parent sample of Ilbert et al. (2010):

$$p_{\rm AGN}(L_{\rm X}|M_*, z) = \frac{{\rm d}N_{\rm AGN}(z)}{{\rm d}\log M_* {\rm d}\log L_{\rm X}} \frac{{\rm d}\log M_*}{{\rm d}N_{\rm gal}(z)}. \quad (6)$$

In practice, the sensitivity limits of the *XMM–Newton* observations vary spatially across the field (see fig. 5 of Cappelluti et al. 2009). Hence, to derive the fraction of galaxies hosting an AGN as a function of the stellar mass, we have followed the technique discussed in section 3.1 of Lehmer et al. (2007) (see also Mainieri et al. 2011). The $p_{\rm AGN}(L_{\rm X}|M_*, z)$ for each stellar mass and luminosity bin is determined by

$$p_{\rm AGN}(L_{\rm X}|M_*, z) = \frac{{\rm d}N_{\rm AGN}(L_{\rm X}|M_*, z)}{{\rm d}\bar{N}_{\rm gal}(M_*, z)\,{\rm d}\log L_{\rm X}}, \quad (7)$$

where ${\rm d}N_{\rm AGN}(L_{\rm X}|M_*, z)$ is the observed number of AGNs in each bin of X-ray luminosity, stellar mass and redshift. We have defined

$${\rm d}\bar{N}_{\rm gal}(M_*, z) = \frac{1}{{\rm d}N_{\rm AGN}(L_{\rm X}|M_*, z)} \sum_{i=1}^{{\rm d}N_{\rm AGN}} N_i(L_{{\rm X},i}|M_*, z). \quad (8)$$

Here, $N_i(L_{\rm X}|M_*, z)$ is the number of galaxies that could have been detected as X-ray sources at their position (given the known exposure map of the *XMM*–COSMOS survey) if they had the luminosity $L_{{\rm X},i}$ of the ith AGN in the given bin.

Fig. 14 shows the probability of galaxies hosting an AGN, $p_{\rm AGN}$, for different stellar mass bins in three redshift intervals. For lower

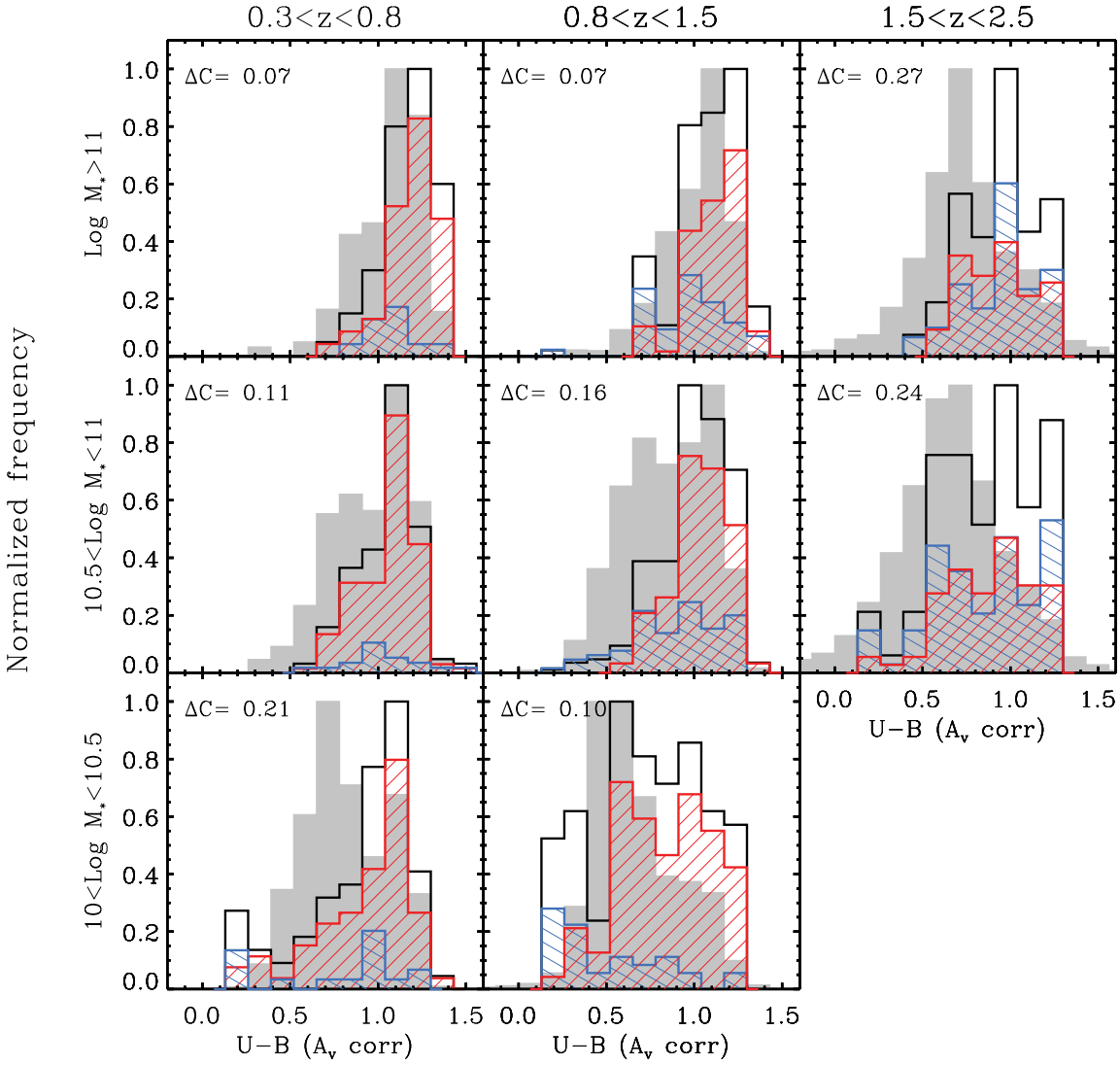

Figure 13. Rest-frame extinction-corrected $U - B$ colour (normalized) distribution of AGN host galaxies in three redshift bins (from left to right) and mass bins (from bottom to top). The black solid line corresponds to the total sample, which is further divided into unobscured (blue shaded histogram) and obscured (red shaded histogram) AGNs, while the filled grey histogram corresponds to the galaxy sample. For each redshift interval, we have considered only the masses above which the parent sample is complete in stellar mass for both passive and star-forming galaxies (Ilbert et al. 2010). In each panel, the median shift in colour $\Delta C \equiv \langle (U-B)_{\rm AGN} \rangle - \langle (U-B)_{\rm gal} \rangle$ is also reported.

mass points, completeness has been treated conservatively. That is, we have assumed that the parent galaxy sample is complete only above the mass completeness limit for quiescent galaxies as function of redshift, as reported by Ilbert et al. (2010); this translates into mass-limit completeness of about $M_* = 10^{9.7}$, $10^{10.1}$ and $10^{10.7}$ for the three redshift bins, respectively. In these cases, our AGN fractions are considered as upper limits.

The good statistics of our large, homogeneous and complete sample allows for an accurate study of AGN incidence, which accounts for both redshift and intrinsic AGN luminosity dependence (see Aird et al. 2012). In each panel, AGNs with different intrinsic X-ray luminosity are plotted with different colours. While the global trend with masses remains qualitatively the same in each luminosity bin, the normalization strongly decreases with luminosity (i.e. for a galaxy of a given mass, the probability to host an AGN is higher as the AGN X-ray luminosity decreases, and the same happens at any redshift). This is an obvious consequence of the shape of the AGN luminosity function: the more luminous AGNs are rare, while less luminous AGNs are much more common.

Very interestingly, as already shown by Aird et al. (2012), $p_{\rm AGN}(L_{\rm X}|M_*, z)$ as a function of galaxy stellar mass does show very similar profiles at all sampled luminosities. The AGNs, of any intrinsic luminosity, are more common, the more massive the host galaxy is. This is by no means a trivial fact; we could have easily expected that more luminous AGNs were to be found in more massive galaxies, and less luminous AGNs in lower-mass galaxies. In fact, this is the prediction of any model in which black hole growth in AGN phases occurs over relatively narrow Eddington ratio ranges. The very different picture we observe in Fig. 14 (as well as in Fig. 15) is a strong indication that AGNs of all masses should be characterized by a very broad distribution of Eddington ratios (e.g. Merloni & Heinz 2008). We quantify this statement by studying the AGN fractions as a function of the specific accretion rate in Section 4.3.2.

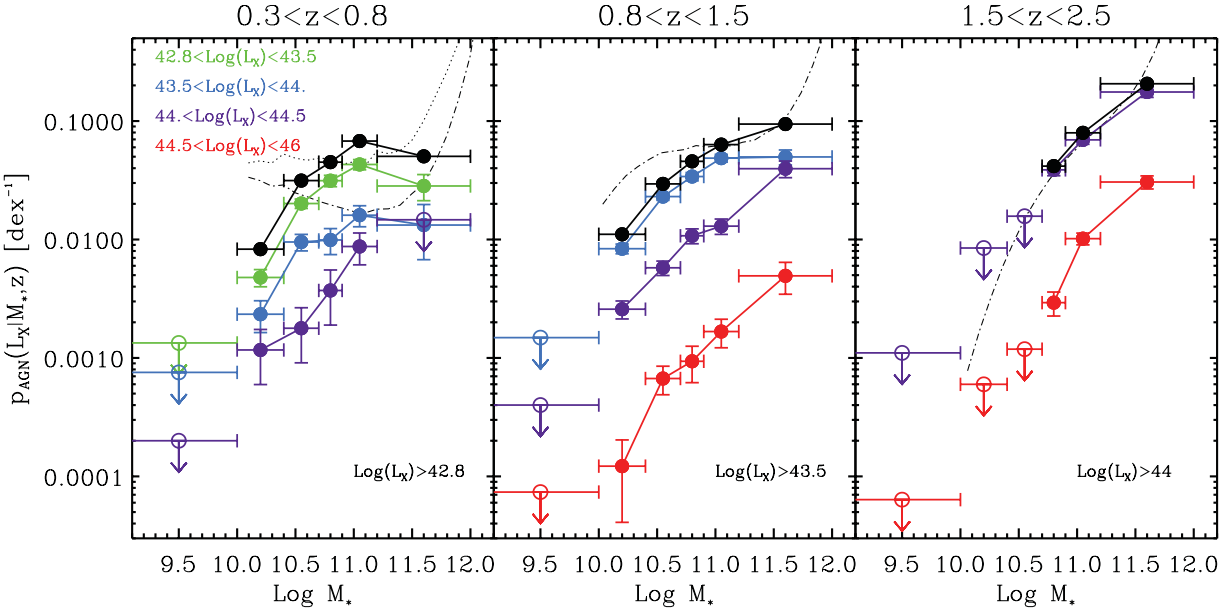

Figure 14. Probability for a galaxy to host an AGN of a given luminosity as a function of stellar mass in three redshift bins. Different colours show different X-ray luminosity bins: green for $\log(L_{[2-10\,\mathrm{keV}]}) = [42.8–43.5]$; blue for $\log(L_{[2-10\,\mathrm{keV}]}) = [43.5–44]$; purple for $\log(L_{[2-10\,\mathrm{keV}]}) = [44–44.5]$ and red for $\log(L_{[2-10\,\mathrm{keV}]}) = [44.5–46]$. The black points refer to the total fraction computed by summing all objects above the given luminosity thresholds. Upper limits correspond to the mass bins where the host galaxy parent sample is not complete, as determined by the completeness limits for quiescent galaxies in COSMOS, as reported by Ilbert et al. (2010). The thin dot-dashed line is the prediction for the total AGN fraction, computed in the same redshift bins, and above the same luminosity thresholds, from the 'AGN occupation' model of Fiore et al. (2012), while the dotted line in the first redshift bin is a modified version of it (see text for more details).

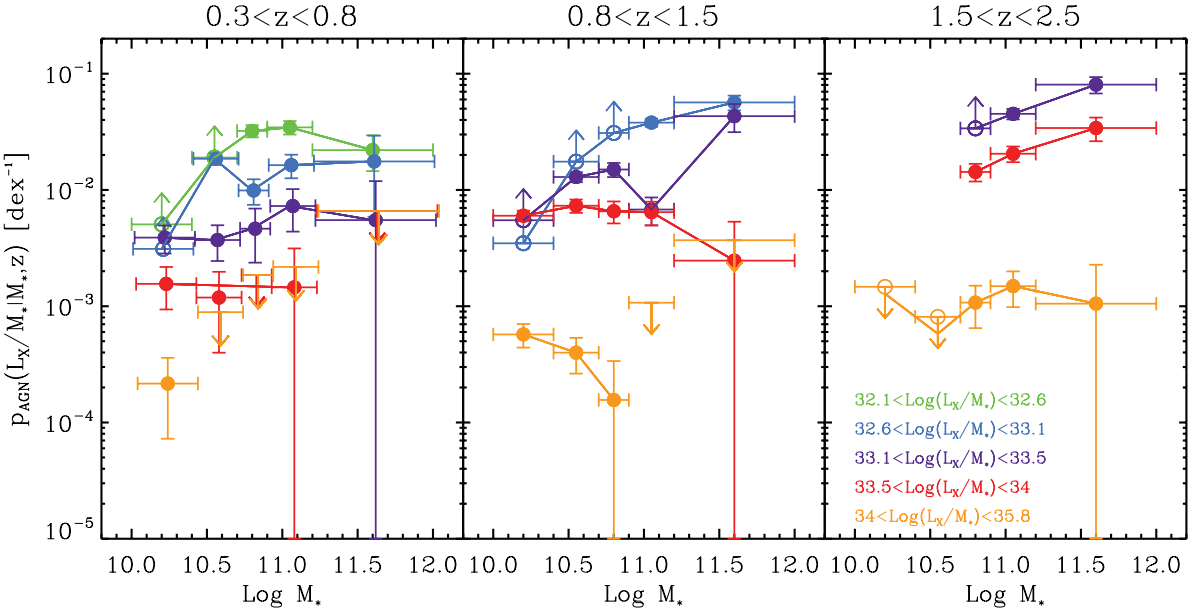

Figure 15. Probability for a galaxy to host an AGN of a specific accretion rate (per decade L_X/M_* ratio) as a function of stellar mass in three redshift bins. Different colours show different L_X/M_* intervals, as explained in the third panel. Empty symbols with arrows show upper and lower limits resulting from the combined effect of completeness in both L_X and M_*. Upper limits without empty circles mark the bins where no AGN was found; these were placed using the Gehrels (1986) formulae for Poisson statistics.

Fig. 14 also shows (black points) the total AGN fraction above a given luminosity threshold in each redshift interval; this is computed by summing up $p_{\mathrm{AGN}}(L_X|M_*, z)$ in different luminosity bins. As a comparison, the thin dot-dashed line is the prediction for the total AGN fraction, computed in the same redshift bins and above the same luminosity thresholds, from the 'AGN occupation' model of Fiore et al. (2012). We refer the reader to that work for more details. Here, we note that Fiore et al. (2012) have computed the AGN fraction by deriving an active SMBH mass function from the observed AGN luminosity function, assuming a given shape for the Eddington ratio distribution, and then comparing it to the observed galaxy mass function. In the two highest redshift bins,

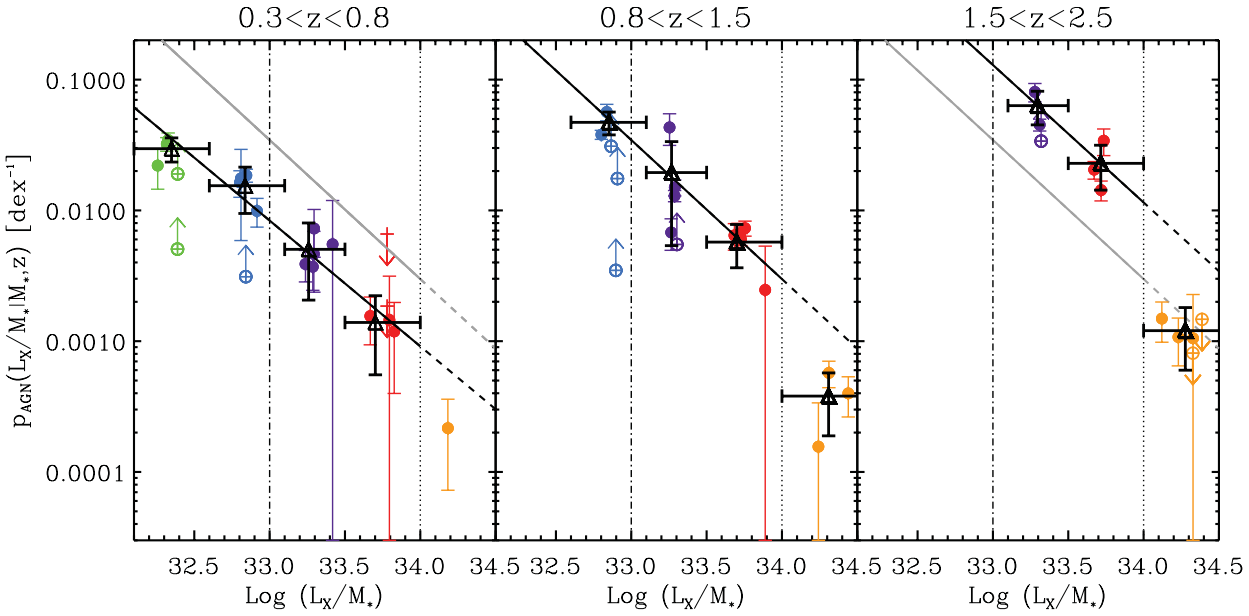

Figure 16. Probability for a galaxy to host an AGN of a given specific accretion rate (per decade L_X/M_* ratio) is plotted as a function of L_X/M_* for three redshift intervals. Coloured points have the same coding as Fig. 15 and plot the individual measurements, while black thick triangles are the mean values. The thick black line in each panel is a power-law fit to all the mean points obeying log (L_X/M_*) < 34. For ease of comparison, in each panel we also show the best fit to the central redshift interval. The vertical dotted (dot-dashed) lines mark the approximate location of objects at the Eddington limit (or 10 per cent of it) calculated assuming a constant bolometric correction $k_{bol} = 25$ and a constant host stellar to black hole mass ratio of $\mathcal{A} = 500$.

such a model reproduces our data remarkably well. However, severe discrepancies appear in our lowest redshift range. As we show in Section 4.3.2, it appears that the true Eddington ratio distribution of our AGN is better described by a simple power law, and the fact that the shape assumed in Fiore et al. (2012) is a lognormal can explain the discrepancy. In fact, while in the highest redshift bins, where our COSMOS AGN sample probes just a limited range of Eddington ratios, the observed distribution is hardly distinguishable from a lognormal distribution with appropriate parameters, in the lowest redshift range, the lognormal shape assumed by Fiore et al. (2012) underpredicts the number of low-luminosity AGNs hosted in massive galaxies ($10^{10.6} < M_* < 10^{11.6}$). Assuming a lognormal distribution with a peak at lower Eddington ratio would improve the prediction in the first redshift bin, as shown by the dotted line in Fig. 14, which corresponds to a distribution with $\lambda_{peak} = 0.03$.

4.3.2 *AGN fraction as a function of specific accretion rate*

By combining the available information (host galaxy stellar mass and AGN luminosity), we can study the probability that a galaxy will host an AGN as a function of mass in bins of specific accretion rate (i.e. the ratio L_X/M_*), which gives an approximate estimate of the rate at which the SMBH grows. To the extent that a proportionality between the black hole mass and the (total) host galaxy stellar mass can be assumed (and this is far from being clear; see Merloni et al. 2010), this ratio could give a rough measure of the black hole Eddington ratio, defined as the ratio between the AGN X-ray luminosity and the Eddington luminosity. To be specific, the Eddington ratio can be expressed as

$$\lambda_{Edd} = \frac{\mathcal{A} \cdot k_{bol}}{1.3 \times 10^{38}} \times \frac{L_X}{M_*}, \quad (9)$$

where k_{bol} is the 2–10 keV bolometric correction, and the factor $\mathcal{A}$ is a constant if the SMBH mass can be related to the host galaxy mass through scaling relations (with $\mathcal{A} \approx$ 500–1000; Magorrian et al. 1998; Häring & Rix 2004). Thus, for a mean bolometric correction[9] of $k_{bol} = 25$ and a constant host stellar to black hole mass ratio of $\mathcal{A} = 500$, a ratio of $L_X/M_* = 10^{34}$ would approximately correspond to the Eddington limit.

Fig. 15 shows how the probability $p_{AGN}(L_X/M_*|M_*, z)$, for a galaxy to host an AGN of a given specific accretion rate, scales with the stellar mass of the host. Here, of course, the issue of completeness is far more complicated. Depending on the L_X/M_* ratio interval considered, both the flux limit of the X-ray survey and the stellar mass completeness limit of the parent galaxy sample should be taken into account. Fig. 15 shows that, at any given specific accretion rate (a proxy, with all the caveats discussed above, of the Eddington ratio), the trend with mass almost disappears; the probability for a galaxy to host an AGN of a given L_X/M_* is almost independent of its mass but scales only with the ratio itself.

In Fig. 16, we plot $p_{AGN}(L_X/M_*|M_*, z)$ as a function of the specific accretion rate itself. As mentioned above, there is little dispersion among the points corresponding to different mass bins for the same L_X/M_* interval. We can assume that the AGN fraction at any given value of the L_X/M_* ratio is independent of host-galaxy mass, and we measure such a constant universal mean probability for every specific accretion rate and redshift bin. The thick black triangles in Fig. 16 show the mean values of $p_{AGN}(L_X/M_*|M_*, z)$. Indeed, all the data seem to be consistent with the idea that, at every redshift, there is a universal distribution function that describes the probability of a galaxy of any mass (above our completeness limit of $\sim 10^{10}$ M$_\odot$) to host an AGN of a given specific accretion rate L_X/M_*, independent of host stellar mass (Aird et al. 2012).

As we have anticipated, to the first order, the data are consistent with a power-law-like distribution and a clear break at values of

[9] Recently, Lusso et al. (2010, 2011) have shown that the 2–10 keV bolometric correction is likely to be an increasing function of the Eddington ratio itself; this is a further complication that we neglect here.

Table 2. The best-fitting slope and normalization for the universal probability distribution function $p_{\rm AGN}(L_{\rm X}/M_*, z)$ as a function of redshift for the entire X-ray selected sample, as shown in Fig. 16.

	All X-ray	
Redshift range	Log K	γ
$0.3 < z < 0.8$	29.50	0.96
$0.8 < z < 1.5$	33.41	1.06
$1.5 < z < 2.5$	33.46	1.04

$L_{\rm X}/M_*$ consistent with the Eddington limit. In Table 2, we report the values of the best-fitting power-law distribution for each of the three redshift bins, in the form:[10]

$$p_{\rm AGN}(L_{\rm X}/M_*|M_*, z) = K \left(\frac{L_{\rm X}}{M_*}\right)^{-\gamma}. \quad (10)$$

Compared to the values found by Aird et al. (2012), we find a steeper slope for the power-law index ($\approx$1 rather than $\approx$0.7). We speculate that this might have to do with the fact that, differently from Aird et al. (2012), we do not include AGNs from the deepest *Chandra* fields, which probe the fainter, shallower end of the X-ray luminosity function. A hint of such a flattening can be seen in the decline of $p_{\rm AGN}$ for the most massive galaxies at low X-ray luminosities in Fig. 14. These are objects with very low specific accretion rates, which are not complete in the COSMOS sample (they are shown as lower limits in Figs 15 and 16), and therefore we cannot properly assess the robustness of any change in slope at low $L_{\rm X}/M_*$. The possibility that $p_{\rm AGN}(L_{\rm X}/M_*|M_*, z)$ might break to a shallower slope at low values of the specific accretion rate could be tested, indirectly, by comparing the observed X-ray luminosity function of the AGNs with a convolution of the observed galaxy mass function and the probability distribution $p_{\rm AGN}(L_{\rm X}/M_*|M_*, z)$. We defer this investigation to future work.

Nevertheless, we see very clearly in Fig. 16 that the $p_{\rm AGN}(L_{\rm X}/M_*|M_*, z)$ normalization increases with redshift over the entire redshift range probed by our data. If we compute the value of $p_{\rm AGN}(L_{\rm X}/M_*|M_*, z)$ at $\log(L_{\rm X}/M_*) = 33.2$ (i.e. at a specific accretion rate approximately corresponding to $\approx$10 per cent of the Eddington rate, and where our data best constrain the distribution at all redshifts), we find that it evolves with redshift as fast as the sSFR of the overall population of galaxies (both star-forming and quiescent; see Karim et al. 2011) $\propto (1+z)^4$.

This very important fact, already pointed out, over a more limited redshift range, by Aird et al. (2012), implies that the rapidity with which a black hole grows is not influenced by the host galaxy mass. However, it is a strongly evolving function of redshift, with a redshift dependence that follows the overall evolution of the sSFR of the galaxy population. This also means that the trend with the host mass, visible in Fig. 14 and widely confirmed by many authors before us, is a trend induced by the fact that in more massive galaxies we are able to detect AGNs down to a lower intrinsic Eddington ratio, while in less massive galaxies only rapidly accreting black holes are bright enough to be observed, and their total number is therefore much smaller.

[10] We note that the fact that the observed distribution of specific accretion rates (or, for that matter, Eddington ratios) is typically found to be lognormal (e.g. Kollmeier et al. 2006; Netzer 2009; Trakhtenbrot et al. 2011; Lusso et al. 2012), rather than a power law, is because of the combined effects of a flux-limited survey, which depletes the number of low specific accretion rates at progressively lower masses, and because very massive galaxies, where slowly accreting objects can be detected, are exponentially rarer as the stellar mass increases. A thorough discussion of these selection effects on the determination of the accretion rate distribution function of AGNs can be found in Schulze & Wisotzki (2010) and Aird et al. (2012).

4.4 Incidence of AGNs as a function of the star-forming properties of their hosts

Studies in the local Universe (Brinchmann et al. 2004; Salim et al. 2005) have pointed out the existence of two main galaxy populations: star-forming galaxies, which are currently forming stars at a rate that has been found to be proportional to their mass, and quiescent galaxies, which are massive galaxies in which the SF has been quenched.[11] The sSFR can be defined as the ratio, sSFR = SFR/M_*, of (total) SFR to stellar mass of a galaxy. The inverse of the sSFR defines a time-scale for the formation of the stellar population of a galaxy, $t_{\rm SF} = {\rm sSFR}^{-1}$, which gives the time it would take a galaxy to double its stellar mass if forming stars at the observed rate.

In the sSFR–mass plane, star-forming galaxies can be divided into a star-forming sequence (the main sequence; Noeske et al. 2007), characterized by an almost constant (or slightly decreasing) sSFR as a function of stellar mass, and a high-mass region with progressively decreasing sSFR, the dying sequence, occupied by quiescent galaxies. In evolutionary terms, such a distribution is observed, almost unchanged, at any redshift, with the sequence of star-forming galaxies maintaining the same slope but increasing normalization as one moves towards higher redshift (Daddi et al. 2007; Elbaz et al. 2007; Noeske et al. 2007; Pannella et al. 2009). There are also more extreme objects, such as starburst galaxies and ultraluminous infrared galaxies (ULIRGs), which exhibit highly elevated SFRs (Sanders & Mirabel 1996) in galaxies within the same range of stellar mass of normal galaxies. Local ULIRGs are believed to be associated with rare major merger events (Sanders et al. 1988; Sanders & Mirabel 1996) and consequently distinct SF processes, but recent studies with the *Herschel* telescope have demonstrated that they represent a minor contribution to the total SFR density in the crucial redshift range $1.4 < z < 2.5$ (Rodighiero et al. 2011). Recently, Wuyts et al. (2011) have shown that a detailed morphological analysis of galaxies, as a function of their position in the sSFR–stellar mass plane, can provide a unique insight into the physical processes that shape and grow galaxies, and how the various modes of SF can be associated with them.

In light of the above results, it is obvious that a systematic study of the distribution of AGN host galaxies in the sSFR–stellar mass plane can reveal key pieces of evidence regarding the long sought-after association between SMBH growth and galaxy formation. Fig. 17 shows the position of our obscured (type 2) COSMOS AGN host galaxies in the sSFR–mass plane compared to the overall non-AGN galaxy population (grey dots). We have purposely avoided including type 1 objects in this diagram, because optical/UV-based SFR measures are too strongly affected by the unobscured accretion disc light in this class of AGN.

The purple, black and blue solid lines show the positions of the main sequence at $z \sim 0$ (Brinchmann et al. 2004), $z \sim 1$ (Elbaz et al. 2007) and $z \sim 2$ (Daddi et al. 2007), while the cyan dotted lines correspond to the age of the Universe $t_{\rm H}(z)$ at the minimum and maximum redshift considered in each panel. Galaxies of known

[11] Regardless of its physical cause, quenching is the process that causes the cessation of SF in some star-forming galaxies, which leads to the emergence of the so-called red sequence of passive galaxies (Strateva et al. 2001).

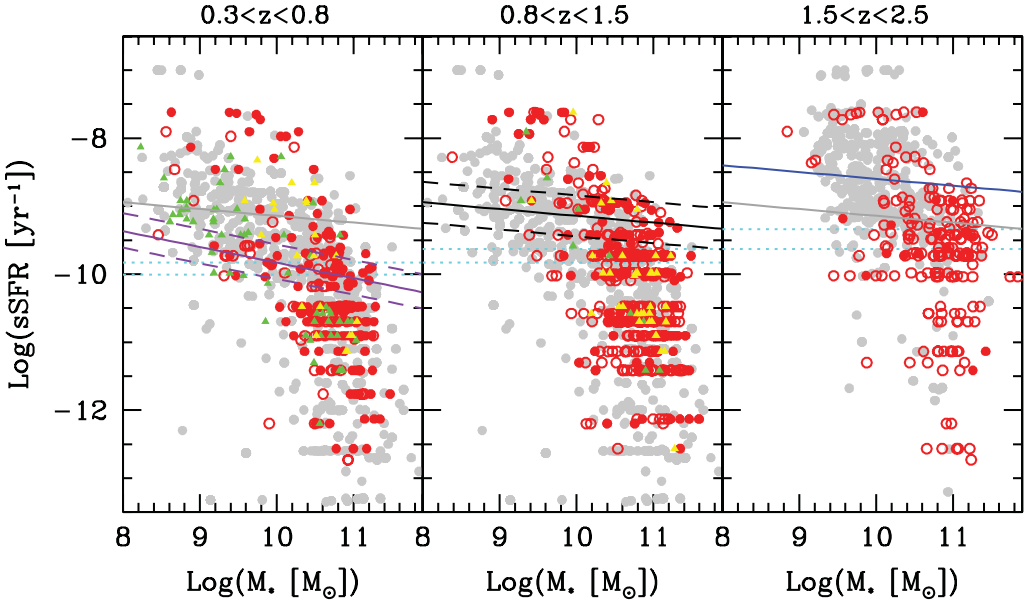

Figure 17. The sSFR rate versus stellar masses in three redshift bins for the host galaxies of type 2 AGNs compared to normal galaxies from the sample matched using the *i*-band apparent magnitude (grey points). The solid lines are the correlations found for blue star-forming galaxies: the purple line in the lowest redshift bin corresponds to $z \sim 0$ (Brinchmann et al. 2004), the black line in the intermediate redshift bin corresponds to $z \sim 1$ (Elbaz et al. 2007) and the blue line in the highest redshift bin corresponds to $z \sim 2$ (Daddi et al. 2007). In each panel, the $z \sim 1$ curve is reported in grey for reference. The cyan dotted horizontal lines correspond to the age of the Universe $t_H(z)$ at the minimum and maximum redshifts considered in each panel, and they indicate roughly the division between galaxies that are actively forming stars and those that are quiescent.

redshift and sSFR can, in fact, be divided into quiescent and star-forming galaxies, on the basis of their stellar mass doubling time-scale, t_{SF}: if $t_{SF} < t_H(z)$, where $t_H(z)$ is the age of the Universe at the redshift of the source, then the galaxy is defined as star-forming; if $t_{SF} > t_H(z)$, thenthe galaxy is defined as quiescent. The parent galaxy population from the S-COSMOS survey (Ilbert et al. 2010), matched using the *i*-band apparent magnitude, does show a broad distinction between a star-forming sequence and a dying sequence at the highest masses, which clearly evolves upwards (in sSFR) with increasing redshift.

Obscured AGN hosts, to first order, appear to populate the high-mass end of the distribution quite uniformly. We estimate that the fraction of AGNs hosted in starburst galaxies (defined as in Rodighiero et al. 2011 to be above four times the main sequence) is below 10 per cent, while the fraction of AGNs hosted in main-sequence galaxies range from 27 to 37 per cent in the different redshift bins. The rest of the AGNs (58–66 per cent) are found in quiescent galaxies. Considering only the most luminous sources, as done by Mainieri et al. (2011), we have found that at $z \sim 1$ about 62 per cent of type 2 QSOs are actively forming stars (i.e. hosted in starburst and main-sequence galaxies). As a comparison, using the FIR *Herschel* data in the GOODS field and excluding the galaxies in which a strong contribution from the AGN to FIR wavelengths was evident, Mullaney et al. (2012) have estimated for moderate luminous AGNs the following percentages: <10 per cent of AGNs are hosted in strongly starbursting galaxies, 79 per cent in main-sequence galaxies and 15 per cent in quiescent galaxies. The discrepancy between the two findings could be attributed to the differences (especially at low SFR) found between the SFR derived from SED fitting and the FIR SFRs, as shown in Appendix B. To test this in detail, we fit below and above 20 $M_\odot$ yr^{-1} the SFR_{SED}–SFR_{IR} relation found (see the left panel of Fig. B1), and we assign the corresponding SFR_{IR} to each of our SFR values. Now, we find that the fraction of AGNs hosted in strongly starbursting galaxies is 9 per cent, while 67 per cent are hosted in main-sequence galaxies and 24 per cent in quiescent galaxies, in much better agreement with the *Herschel*-based results.

Fig. 18 shows histograms of the sSFR distributions of AGN hosts and normal galaxies in bins of stellar mass and redshift. Once again, the AGNs do appear to follow the evolution of the sSFR with stellar mass and redshift, with just a marginal increase in the AGN fraction towards lower values of sSFR in almost all bins, indicating a slightly higher incidence of (obscured) AGNs in non-star-forming galaxies.

According to the models that assume that major mergers are the main AGN fuelling mechanism (e.g. Hopkins et al. 2008), we could have expected an enhancement of the SFR activity in obscured AGNs. However, this is not visible. From our analysis, there is no clear evidence that the COSMOS obscured AGN hosts are preferentially found among the extreme, most rapidly star-forming galaxies.

To analyse this in further detail, we divide the host galaxies into two classes, depending on the ratio between t_{SF} and t_H. According to this division of the total sample of obscured AGNs discussed here, the fractions of quiescent hosts (always computed for stellar masses above the completeness limit of the COSMOS IRAC survey, as discussed by Ilbert et al. 2010) are 75, 65 and 61 per cent in the redshift bins 0.3–0.8, 0.8–1.5 and 1.5–2.5, respectively. In total, about two-thirds of the obscured AGNs in COSMOS that we find in massive galaxies are hosted by objects with a stellar mass doubling time longer than the age of the Universe. These fractions are higher than those of the galaxy sample, where the fractions of quiescent hosts are 51, 41 and 32 per cent in the redshift bins 0.3–0.8, 0.8–1.5 and 1.5–2.5, respectively. Such a predominance of quiescent (or, better, not strongly star-forming hosts) is consistent with the predominantly red colours exhibited by the AGN hosts that we have discussed in Section 4.1.

In Section 4.3.2, we have computed the probability for a galaxy (above the stellar mass completeness threshold of the COSMOS

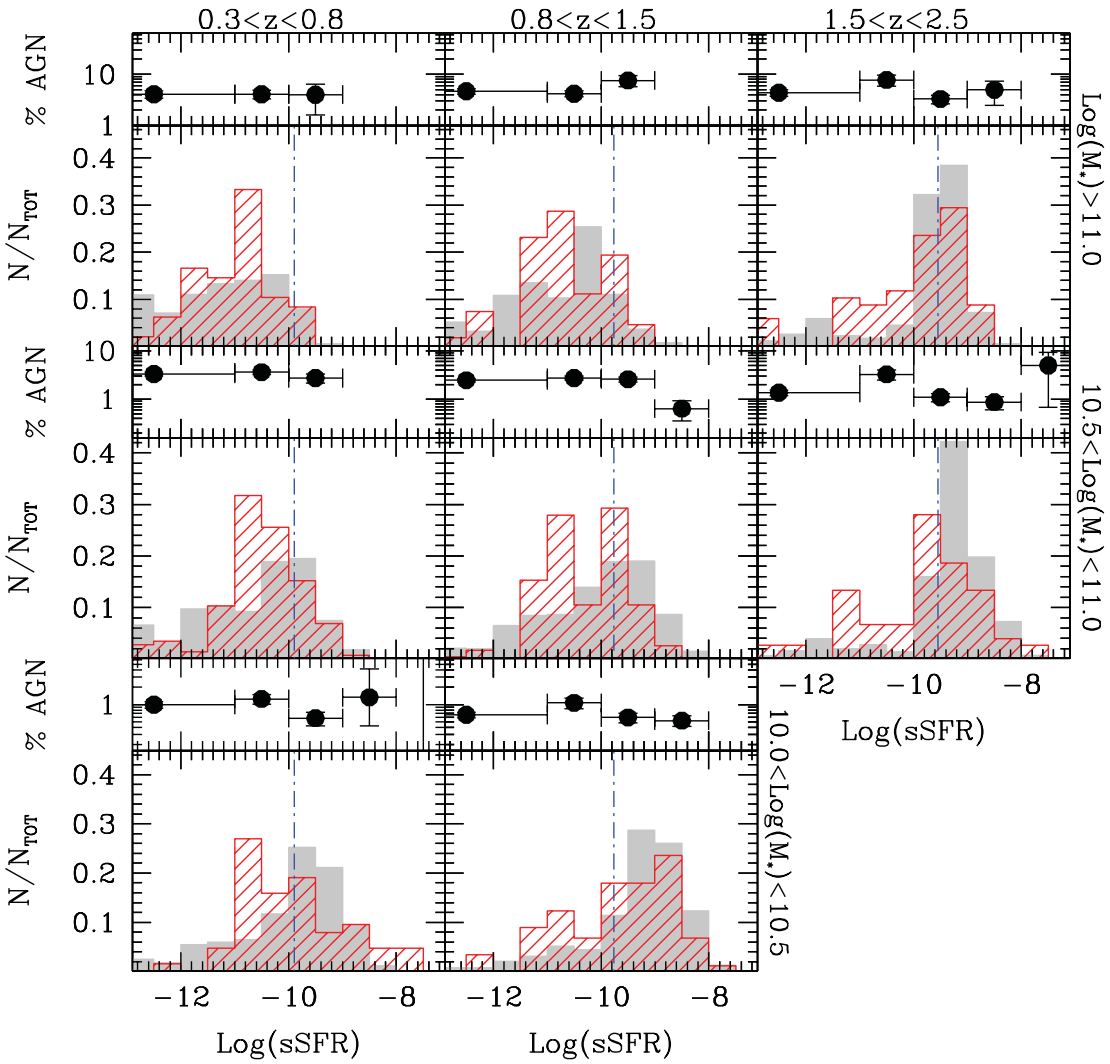

Figure 18. The sSFR distributions divided into three redshift bins and three mass ranges for type 2 AGN host galaxies (red histogram) and normal galaxies (grey histogram). The dot-dashed vertical blue lines are used to distinguish quiescent galaxies from star-forming galaxies, based on the ratio between the mass doubling time t_{SF} and the age of the Universe in any redshift bin. The upper panel of each histogram shows the observed AGN fractions.

IRAC survey) to host an AGN of a given specific accretion rate, irrespective of the SFR properties of the galaxies. For the subsample of AGNs that are obscured, for which a reliable and robust determination of the sSFR is possible, we have also computed separately the probability for star-forming and quiescent galaxies to host a (type 2) AGN, as a function of specific accretion rate L_X/M_*. The results are shown in Fig. 19, where we also report, for the sake of comparison, the overall p_{AGN} of Fig. 16. Taken at face value, the plot shows that there is little significant evidence for a difference in the probabilities for quiescent and star-forming galaxies to host type 2 AGNs, because the data points agree within the errors; here, these errors are representative of the dispersion of $p_{AGN}(L_X/M_*|M_*, z)$ among the different stellar mass bins. The slight increase in the probability for a passive galaxy to host an AGN is hinted at in the highest redshift bin. However, in the lowest redshift bin, star-forming galaxies appear to have a higher probability of hosting an AGN with low specific accretion rate; the number statistics is far too low to make any firmer statement. Even the size of our COSMOS sample is not large enough, once we have subdivided the sources into spectral class, X-ray luminosity, stellar mass and SF properties of the hosts.

5 DISCUSSION

Consistent with earlier studies, X-ray selected AGNs are found to reside mainly in massive galaxies, while optically selected Seyfert 2 galaxies, which are also low-luminosity galaxies, seem to prefer galaxies with smaller masses. The total stellar masses of the AGN host galaxies in our sample range from 10^{10} to $10^{11.5}$ $M_\odot$, with a peak at $10^{10.9}$ $M_\odot$. No difference is found between X-ray obscured and unobscured AGNs, which on average show the same mass distribution.

In accordance with their high stellar masses, AGN hosts are more likely to be found in galaxies with red colours. Many authors have pointed out that the trend with colour was just a mass-driven effect, and that removing the mass dependency would make it disappear (Silverman et al. 2009; Xue et al. 2010; Mainieri et al. 2011; Rosario et al. 2011). Indeed, several previous studies (e.g. Silverman et al. 2008) have claimed that AGNs are found mainly in a region of intermediate colours, the so-called green valley. This was interpreted as evidence of a close association between the migration of galaxies from the blue cloud to the red sequence and the feedback from AGNs. However, in most of these works, the AGN

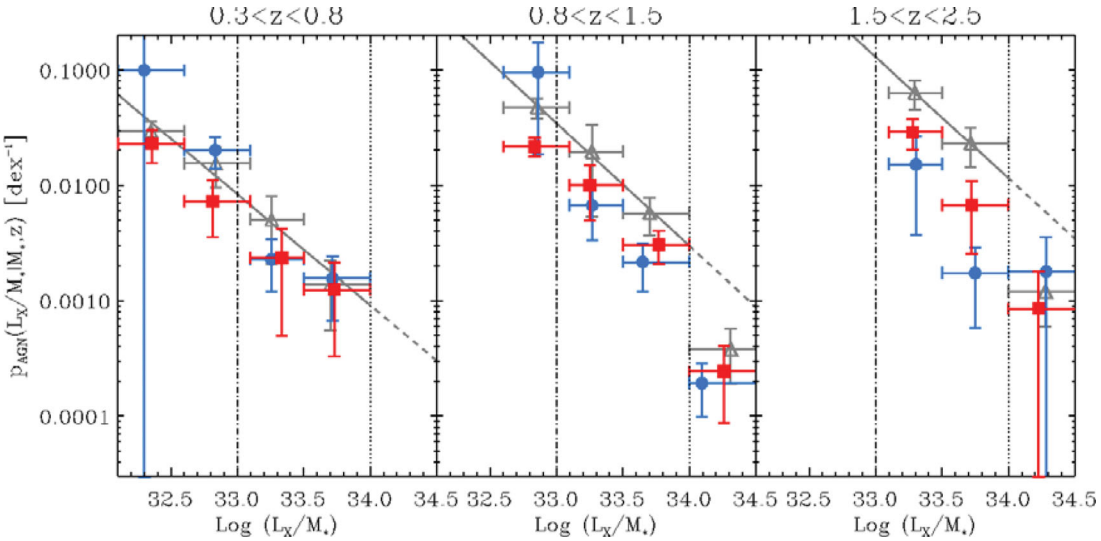

Figure 19. The average probability for a galaxy to host an AGN of a given specific accretion rate (per decade L_X/M_* ratio) is plotted as a function of L_X/M_* for three redshift intervals. In each panel, blue filled circles denote star-forming galaxies (i.e. those with $t_{SF} < t_H$) and red squares are for quiescent galaxies (i.e. those with $t_{SF} > t_H$). Solid lines of corresponding colours show the power-law fits to all points obeying $\log(L_X/M_*) < 34$. The grey symbols and lines are the probability distribution functions for the overall galaxy population (for all AGN types), as from Fig. 16. The vertical dotted (dot-dashed) lines mark the approximate location of objects at the Eddington limit (or 10 per cent of it) calculated assuming a constant bolometric correction $k_{bol} = 25$ and a constant host stellar to black hole mass ratio of $\mathcal{A} = 500$.

component was not subtracted in the computation of the host colours, and the dust-extinction correction could not be properly assessed. This led to a possible blue contamination from the AGN and a higher incidence of AGNs in the blue cloud and green valley. These two effects could explain our findings. In fact, in disagreement with the above studies, we find that compared to normal galaxies of the same mass, galaxies hosting an AGN are, on average, (slightly) redder and form stars less rapidly. It is only by properly taking into account both the dust-extinction properties and the contamination from the nuclear light of the entire galaxy population that we can assess the true incidence of AGNs as a function of SF in their hosts. However, we cannot exclude the fact that there is a difference in colour between low-luminosity and intermediate- to high-luminosity AGNs, and that these differences are, at least partially, responsible for the discrepancy between our results and previous studies based on deep fields (e.g. Silverman et al. 2008; Aird et al. 2012).

In agreement with the standard unified model (Antonucci 1993) and contrary to what is (perhaps naïvely) expected from the most popular models of AGN fuelling and feedback (Hopkins et al. 2008), and with a number of caveats regarding the robustness of colour determination in type 1 AGNs, we do not find any significant difference between the host colours of obscured and unobscured AGNs. We note that Pierce et al. (2010) and Rosario et al. (2011) have recently found possible evidence of a weak anticorrelation between the colour of the AGN hosts and the hydrogen column density of X-ray obscuring gas N_H, in the sense that redder AGNs seem to have, on average, higher values of N_H. Because our classification into obscured and unobscured AGNs is a complex one, and is mainly based on optical spectroscopy, it is not straightforward to compare our results with those of Pierce et al. (2010) and Rosario et al. (2011). However, a common wisdom is emerging that (at least mildly) obscured AGNs are not to be found among the most active star-forming galaxies.

When looking in detail at the relationships between AGN luminosity and host stellar masses, in agreement with Aird et al. (2012), we find that the probability of hosting an AGN of a given specific accretion rate L_X/M_* (a proxy for the SMBH Eddington ratio) is practically independent of the galaxy stellar mass. However, it does depend strongly on redshift and on the specific accretion rate itself (i.e. there is a strongly decreasing probability of hosting an AGN with increasing L_X/M_* ratio).

This also means that the higher AGN fraction is observed in massive red galaxies simply because it is easier to detect the more prevalent, low Eddington ratio AGNs in more massive galaxies, which are more likely to host the more massive SMBH.

The specific accretion rate distributions that we can indirectly infer from the observed probability for a galaxy to host an AGN are better described by a power-law function, the slope and normalization of which appear to increase with redshift. We note that, in studying the present-day growth rates of black holes in SDSS galaxies using the [O III] luminosity as a tracer of AGN accretion activity and stellar masses as a tracer of the BH masses, Kauffmann & Heckman (2009) have pointed out the existence of two trends in the Eddington ratio distribution as a function of the stellar population: a power law and a lognormal. They have interpreted these observed trends as the indication of two distinct accretion modes for black holes: the lognormal distribution identifies a mode in which the supply of gas is ample (i.e. feast), while the power-law distribution corresponds to a mode where accretion is limited by the available supply of gas (i.e. famine). Another interpretation has been proposed by Hopkins et al. (2008), who have identified the origin of the bimodality in the evolution of triggering rates (mergers) in combination with a luminosity-dependent AGN lifetime. Our result shows that at higher redshift the picture might be different, because we do not find a clear sign of a lognormal component in the specific accretion rate distribution, even when dividing the population of type 2 AGN hosts into actively star-forming and quiescent. This difference could be indicative of an interesting evolutionary effect to be further investigated, although we cannot exclude a possible selection effect, because the work from Kauffmann & Heckman (2009) is based on a pure optically

selected sample for which the AGN luminosity is computed based on the highly uncertain narrow-line [O III]-to-bolometric luminosity conversion. None the less, the apparent universality of the observed probability distribution $p_{\rm AGN}(L_{\rm X}/M_*|z)$ suggests that, at least in the redshift range populated by our COSMOS sample, the same physical processes might be responsible for triggering and fuelling AGN activity in galaxies of different masses.

We also find that the overall probability of hosting an AGN of any given $L_{\rm X}/M_*$ (as determined by the distribution normalization) increases rapidly with redshift as $(1+z)^4$. Such an evolution can be a result of either an increase in the number of sources accreting (vertical shift) or a shift of the characteristic Eddington ratio distribution towards higher values (horizontal shift). The second possibility would also explain the AGN downsizing (the comoving number density of high-luminosity objects peaks at higher redshifts than low-luminosity AGNs; Hasinger et al. 2005; La Franca et al. 2005; Bongiorno et al. 2007) as a consequence of the different Eddington ratio distributions at different redshifts.

The fact that the probability for a galaxy to host an AGN of a given specific accretion rate evolves with redshift in the same way as the overall evolution of the sSFR of the galaxy population (Karim et al. 2011) strongly suggests that AGN activity and SF are globally correlated, at least in a statistical sense. From a physical point of view, the drop with time in the global SFR density reflects a drop in the availability of cold gas in galaxies to fuel SF, and potentially AGNs. In that respect, AGNs seem to follow the evolution of the galaxy population as a whole, a possibility reinforced by our detailed study of the position of AGN hosts in the sSFR–stellar mass diagram (Section 4.4).

6 SUMMARY AND CONCLUSION

In this paper, we have studied a sample of ∼1700 obscured and unobscured AGNs, selected in the COSMOS field using X-ray and optical (spectroscopic) criteria. Of these, 602 are unobscured type 1 AGNs (430 with spectra) and 1100 obscured type 2 AGNs (549 with spectra) with $L_{\rm bol}$ ranging from 10^{43} to 10^{47} erg s^{-1}. Because of the wide COSMOS multiwavelength coverage, we have been able to study in detail the source SEDs and to fit the observed fluxes with a two-component model based on a combination of AGN and host-galaxy templates, accounting for dust extinction in both components independently. This method has allowed us to decompose the entire SED into a nuclear AGN and a host-galaxy component for almost all the sources, and to derive robust measurements of both the AGN and the host-galaxy properties (i.e. rest-frame colours, masses and SFRs). A test considering one-component (galaxy) versus two-component (galaxy+AGN) SED fitting techniques, discussed in Appendix A, has revealed that while the consideration of only the galaxy component can be used to roughly estimate the blue colours of obscured AGN hosts, caution has to be used when deriving red magnitudes (e.g. M_K) and physical parameters (e.g. stellar mass and SFR). On the contrary, for unobscured AGN hosts, no study should be performed without properly taking into account the AGN component. Even with this caution, we find that for unobscured AGNs, while total stellar masses of their host can be determined quite reliably, SFRs are affected too much by nuclear light contamination to be deemed robust.

Our main findings can be summarized as follows.

(i) The COSMOS AGN host galaxy masses range from 10^{10} to $10^{11.5}$ M$_\odot$ with a peak at $\approx 10^{10.9}$ M$_\odot$. No significant difference is found in the stellar mass distributions of X-ray obscured and unobscured AGN hosts. Only optically selected Seyfert 2 galaxies without X-ray counterparts populate the tail in the distribution at lower masses down to $10^{8.5}$ M$_\odot$.

(ii) The host galaxies of obscured AGNs span a very wide range of SFRs, with X-ray selected AGNs populating the entire range, while optically selected Seyfert 2 galaxies are hosted mainly in galaxies with low SFRs (and very low masses). As extensively discussed in Appendix B, we find that above ∼20 M$_\odot$ yr^{-1} the SFRs derived from the SED and SFR derived from the *Herschel* FIR observations are in good agreement. At low SFR (<20 M$_\odot$ yr^{-1}), however, the disagreement becomes evident, with our SED fitting method indicating substantially lower SFR than the FIR SFR.

We interpret this discrepancy as the possible combination of two effects: (i) contamination from SF-related emission in the NIR not properly taken into account in the SED fitting; (ii) AGN contamination in the FIR bands not taken into account in the FIR SFR estimates. We have demonstrated that, at least for the *XMM*–COSMOS sample, which is dominated by high-luminosity AGNs, the driven effect is the AGN contamination on the 60-μm luminosities (see Appendix B for a detailed description). In contrast, this is low or even negligible for moderate and low-luminosity AGN samples.

(iii) In contrast to what was found by Cardamone et al. (2010), no clear bimodality is seen in the CMD of the COSMOS AGN hosts. At any redshift, the hosts of the AGN in our sample are mainly massive, redder galaxies. In many previous studies on AGN samples of much smaller size, AGNs were found to have a higher incidence in star-forming galaxies of blue colours (Silverman et al. 2009; Xue et al. 2010), once mass-matched samples of active and non-active galaxies are considered. Even when accounting properly for the mass–colour degeneracy, and when looking at the mass-matched sample, we do not confirm these findings with our much improved statistics. Even at fixed stellar masses, well above the completeness limit of the COSMOS sample, we find a slight trend for AGN hosts to be redder (by ∼0.1–0.2 mag) than the non-active galaxies. We argue that the main reason for such a discrepancy is the proper subtraction of the AGN light from our SED decomposition technique, complemented by a uniform and systematic treatment of dust extinction in samples of both parent galaxies and AGN host galaxies. However, we caution that degeneracies in CMDs can easily arise because of alternative choices of spectral templates and residual issues with the AGN–host decomposition technique.

(iv) For a galaxy of any given mass above the stellar mass completeness limit of the COSMOS survey, the probability of hosting an AGN of a given X-ray luminosity increases with stellar mass, and this holds true for any value of the X-ray luminosity to which we are sensitive. In fact, against the naïve expectation that more luminous AGNs should be found in more massive galaxies and less luminous AGNs in lower mass galaxies, we find that AGNs of any intrinsic luminosity are more common, the more massive the galaxy is.

(v) We have used the observable, the specific accretion rate (i.e. the ratio between X-ray luminosity and host galaxy stellar mass), as a proxy for the Eddington ratio of the growing black hole. By doing so, we find that the probability for a galaxy to host an AGN of any given Eddington ratio (or specific accretion rate) is roughly independent of the host galaxy stellar mass, but strongly decreases with increasing $L_{\rm X}/M_*$. We have also found clear evidence of a break in the specific accretion rate distribution at values consistent with the Eddington limit. These results imply that the higher incidence of AGNs observed in massive galaxies by essentially all multiwavelength surveys is just a consequence of the fact that objects with

low specific accretion rates are more common than those with high specific accretion rates. At the same time, low L_X/M_* objects in low-mass galaxies drop out of flux-limited AGN samples. This effect, closely associated with the broad Eddington ratio distribution of AGNs, was anticipated in Merloni & Heinz (2008) and has been confirmed observationally by Aird et al. (2012).

(vi) We have found that the probability for a galaxy to host an AGN of any given specific accretion rate strongly increases with redshift, approximately as $(1 + z)^4$ for a SMBH accreting at about 10 per cent of the Eddington limit, where most of the SMBH mass assembly is supposed to take place. Such a very strong evolutionary trend follows closely the overall evolution of the sSFR of the galaxy population (Karim et al. 2011). This is yet another strong indication that AGN activity and SF are globally correlated.

(vii) In order to study in detail the SF properties of the COSMOS AGN hosts, we have looked at the type 2 (obscured) AGNs only, in order to avoid any bias/contamination from the inaccurate subtraction of the dominant blue nuclear light in type 1 objects. Type 2 AGN hosts have, on average, the same, or slightly lower SFRs than non-active galaxies of the same mass and at the same redshift. We find that the fraction of AGNs hosted in strongly starbursting galaxies is very low (<10 per cent), while it is higher for main-sequence galaxies (27–37 per cent) and quiescent galaxies (58–66 per cent). Compared to *Herschel*-based results (e.g. Mullaney et al. 2012), we find a much higher fraction of quiescent hosts. We have tested whether the discrepancy between the two results can be explained by the differences in the SFR (especially below $20\,M_\odot\,yr^{-1}$) derived from SED fitting and FIR, which have been attributed to the combination of the two opposite effects explained above.

The probabilities for star-forming or quiescent galaxies to host a type 2 AGN also decline as a power law as a function of specific accretion rate. There are tantalizing hints of differences in the two populations and in their redshift dependence. However, more detailed investigations of the incidence of AGNs in galaxies as a function of redshift, stellar mass, SFR and nuclear luminosity are hampered by the limited statistics of the sample. Wide-area surveys probing volumes at $z > 1$, comparable to that explored by SDSS in the local Universe, will be needed to properly address the fundamental questions on the physical nature of the AGN–galaxy coevolution.

The most striking evidence that emerges from our study is the fact that AGN hosts do follow quite closely the overall evolution of the stellar build-up of the parent galaxy population and that they populate both star-forming and more passive galaxies almost equally.

We argue that, on the basis of the evidence presented in this paper, it is hard to find a strong indication that AGNs play a direct role in shaping the global properties of their host galaxies, or their evolution. The same, probably stochastic, process of AGN activation and triggering seems to be in place in galaxies of all masses, with a slight predominance in red galaxies with sSFRs below the nominal threshold that distinguishes actively star-forming galaxies from quiescent galaxies. Because of the robustness of this picture across the wide redshift range probed, we believe that whatever physical process is responsible for triggering and fuelling AGN activity, it must be the same between $z \sim 2.5$ and $z \sim 0.3$, but it must decrease in frequency or shift towards lower accretion rates. Although AGN activity and SF appear to have a common triggering mechanism, we do not find any conclusive evidence signalling that there is a powerful AGN influence on the SFRs of star-forming galaxies. The extent to which this lack of evidence can be meaningfully used to constrain and/or to rule out theoretical models of AGN feedback remains to be investigated.

ACKNOWLEDGMENTS

We thank the referee, Dr James Mullaney, for constructive comments that have helped to improve the manuscript and for the interesting discussion. We also thank R. Hyckox and K. Schawinski for useful discussions. AB is supported by the INAF Fellowship Programme. The *HST* COSMOS Treasury programme was supported through National Aeronautics and Space Administration (NASA) grant HST-GO-09822. This work is mainly based on observations obtained with *XMM–Newton*, an ESA Science Mission with instruments and contributions directly funded by ESA Member States and the USA (NASA), and with the European Southern Observatory under the Large Programme 175.A-0839, Chile. In Germany, the *XMM–Newton* project is supported by the Bundesministerium für Wirtschaft und Technologie/Deutsches Zentrum für Luft und Raumfahrt (BMWI/DLR, FKZ 50 OX 0001), the Max-Planck Society and the Heidenhain-Stiftung. In Italy, the *XMM–*COSMOS project is supported by ASI-INAF grants I/009/10/0 and ASI/COFIS/WP3110, I/026/07/0. Part of this work was supported by INAF PRIN 2010 'From the dawn of galaxy formation to the peak of the mass assembly'. We gratefully acknowledge the contribution of the entire COSMOS collaboration. More information on COSMOS is available at http://www.astro.caltech.edu/cosmos.

REFERENCES

Aird J. et al., 2012, ApJ, 746, 90
Allevato V. et al., 2011, ApJ, 736, 99
Antonucci R., 1993, ARAA, 31, 473
Avni Y., 1976, ApJ, 210, 642
Baldwin J. A., Phillips M. M., Terlevich R., 1981, PASP, 93, 5
Barvainis R., 1987, ApJ, 320, 537
Berta S. et al., 2011, A&A, 532, A49
Best P. N., Kauffmann G., Heckman T. M., Brinchmann J., Charlot S., Ivezić Ž., White S. D. M., 2005, MNRAS, 362, 25
Bongiorno A. et al., 2007, A&A, 472, 443
Bongiorno A. et al., 2010, A&A, 510, A56
Brandt W. N., Hasinger G., 2005, ARAA, 43, 827
Brinchmann J., Charlot S., White S. D. M., Tremonti C., Kauffmann G., Heckman T., Brinkmann J., 2004, MNRAS, 351, 1151
Brusa M. et al., 2009a, ApJ, 693, 8
Brusa M. et al., 2009b, A&A, 507, 1277
Brusa M. et al., 2010, ApJ, 716, 348
Bruzual G., Charlot S., 2003, MNRAS, 344, 1000 (BC03)
Calzetti D., Armus L., Bohlin R. C., Kinney A. L., Koornneef J., Storchi-Bergmann T., 2000, ApJ, 533, 682
Capak P. et al., 2007, ApJS, 172, 99
Cappelluti N. et al., 2009, A&A, 497, 635
Cardamone C. N., Urry C. M., Schawinski K., Treister E., Brammer G., Gawiser E., 2010, ApJ, 721, L38
Cen R., Chisari N. E., 2011, ApJ, 731, 11
Chabrier G., 2003, PASP, 115, 763
Chary R., Elbaz D., 2001, ApJ, 556, 562
Cimatti A., Daddi E., Renzini A., 2006, A&A, 453, L29
Ciotti L., Ostriker J. P., Proga D., 2010, ApJ, 717, 708
Cirasuolo M., Magliocchetti M., Celotti A., 2005, MNRAS, 357, 1267
Cisternas M. et al., 2011, ApJ, 741, L11
Civano F. et al., 2012, ApJS, 201, 30
Comastri A., Brusa M., 2008, Astronomische Nachrichten, 329, 122
Cowie L. L., Songaila A., Hu E. M., Cohen J. G., 1996, AJ, 112, 839
Croton D. J. et al., 2006, MNRAS, 365, 11
Czerny B., Elvis M., 1987, ApJ, 321, 305
Daddi E. et al., 2007, ApJ, 670, 156

Dickinson M., Papovich C., Ferguson H. C., Budavári T., 2003, ApJ, 587, 25
Donley J. L. et al., 2012, ApJ, 748, 142
Elbaz D. et al., 2007, A&A, 468, 33
Elbaz D. et al., 2011, A&A, 533, A119
Elvis M. et al., 1994, ApJS, 95, 1
Elvis M. et al., 2009, ApJS, 184, 158
Elvis et al., 2012, ApJ, 759, 6
Ferrarese L., Merritt D., 2000, ApJ, 539, L9
Fiore F. et al., 2003, A&A, 409, 79
Fiore F. et al., 2012, A&A, 537, A16
Fontana A. et al., 2006, A&A, 459, 745
Franceschini A., Hasinger G., Miyaji T., Malquori D., 1999, MNRAS, 310, L5
Gandhi P., Horst H., Smette A., Hönig S., Comastri A., Gilli R., Vignali C., Duschl W., 2009, A&A, 502, 457
Gebhardt K. et al., 2000, ApJ, 539, L13
Gehrels N., 1986, ApJ, 303, 336
Genzel R. et al., 2008, ApJ, 687, 59
Gilli R., Vignali C., Mignoli M., Iwasawa K., Comastri A., Zamorani G., 2010, A&A, 519, A92
Granato G. L., De Zotti G., Silva L., Bressan A., Danese L., 2004, ApJ, 600, 580
Grazian A. et al., 2007, A&A, 465, 393
Hao H. et al., 2009, Bull. Am. Astron. Soc., 41, 423
Hao H. et al., 2010, ApJ, 724, L59
Häring N., Rix H., 2004, ApJ, 604, L89
Hasinger G., 2008, A&A, 490, 905
Hasinger G., Miyaji T., Schmidt M., 2005, A&A, 441, 417
Hasinger G. et al., 2007, ApJS, 172, 29
Hickox R. C. et al., 2009, ApJ, 696, 891
Hopkins P. F., Hernquist L., Cox T. J., Di Matteo T., Robertson B., Springel V., 2006, ApJS, 163, 1
Hopkins P. F., Richards G. T., Hernquist L., 2007, ApJ, 654, 731
Hopkins P. F., Hernquist L., Cox T. J., Kereš D., 2008, ApJS, 175, 356
Ilbert O. et al., 2009, ApJ, 690, 1236
Ilbert O. et al., 2010, ApJ, 709, 644
Isobe T., Feigelson E. D., Akritas M. G., Babu G. J., 1990, ApJ, 364, 104
Johansson P. H., Burkert A., Naab T., 2009, ApJ, 707, L184
Karim A. et al., 2011, ApJ, 730, 61
Kauffmann G., Heckman T. M., 2009, MNRAS, 397, 135
Kennicutt R. C., Jr, 1998, ARAA, 36, 189
Kocevski D. D. et al., 2012, ApJ, 744, 148
Kollmeier J. A. et al., 2006, ApJ, 648, 128
Kormendy J., Bender R., 2011, Nat, 469, 377
Kormendy J., Richstone D., 1995, ARAA, 33, 581
La Franca F. et al., 2005, ApJ, 635, 864
Lamareille F., Mouhcine M., Contini T., Lewis I., Maddox S., 2004, MNRAS, 350, 396
La Massa S. M., Heckman T. M., Ptak A., Hornschemeier A., Martins L., Sonnentrucker P., Tremonti C., 2009, ApJ, 705, 568
Le Floc'h E. et al., 2009, ApJ, 703, 222
Lehmer B. D. et al., 2007, ApJ, 657, 681
Lilly S. J. et al., 2007, ApJS, 172, 70
Lilly S. J. et al., 2009, ApJS, 184, 218
Lusso E. et al., 2010, A&A, 512, A34
Lusso E. et al., 2011, A&A, 534, A110
Lusso E. et al., 2012, MNRAS, 425, 623
Lutz D. et al., 2011, A&A, 532, A90
McCracken H. J. et al., 2010, ApJ, 708, 202
McLure R. J., Dunlop J. S., 2001, MNRAS, 327, 199
Magnelli B., Elbaz D., Chary R. R., Dickinson M., Le Borgne D., Frayer D. T., Willmer C. N. A., 2009, A&A, 496, 57
Magnelli B., Elbaz D., Chary R. R., Dickinson M., Le Borgne D., Frayer D. T., Willmer C. N. A., 2011, A&A, 528, A35
Magorrian J. et al., 1998, AJ, 115, 2285
Mainieri V. et al., 2011, A&A, 535, A80
Marconi A., Hunt L. K., 2003, ApJ, 589, L21
Marconi A., Risaliti G., Gilli R., Hunt L. K., Maiolino R., Salvati M., 2004, MNRAS, 351, 169
Menci N., Fiore F., Puccetti S., Cavaliere A., 2008, ApJ, 686, 219
Merloni A., 2004, MNRAS, 353, 1035
Merloni A., Heinz S., 2008, MNRAS, 388, 1011
Merloni A. et al., 2010, ApJ, 708, 137
Miyaji T., Hasinger G., Schmidt M., 2000, A&A, 353, 25
Monaco P., Fontanot F., 2005, MNRAS, 359, 283
Mullaney J. R., Alexander D. M., Goulding A. D., Hickox R. C., 2011, MNRAS, 414, 1082
Mullaney J. R. et al., 2012, MNRAS, 419, 95
Nandra K. et al., 2007, ApJ, 660, L11
Netzer H., 2009, MNRAS, 399, 1907
Noeske K. G. et al., 2007, ApJ, 660, L43
Page M. J. et al., 2012, Nat, 485, 213
Pannella M. et al., 2009, ApJ, 698, L116
Pierce C. M. et al., 2010, MNRAS, 408, 139
Polletta M. et al., 2007, ApJ, 663, 81
Pozzetti L. et al., 2007, A&A, 474, 443
Pozzi F. et al., 2010, A&A, 517, A11
Prescott M. K. M., Impey C. D., Cool R. J., Scoville N. Z., 2006, ApJ, 644, 100
Prevot M. L., Lequeux J., Prevot L., Maurice E., Rocca-Volmerange B., 1984, A&A, 132, 389
Richards G. T. et al., 2006, ApJS, 166, 470 (R06)
Rodighiero G. et al., 2010, A&A, 518, L25
Rodighiero G. et al., 2011, ApJ, 739, L40
Rola C. S., Terlevich E., Terlevich R. J., 1997, MNRAS, 289, 419
Rosario D. J. et al., 2011, preprint (arXiv:1110.3816)
Rosario D. J. et al., 2012, A&A, 545, A45
Salim S. et al., 2005, ApJ, 619, L39
Salvato M. et al., 2009, ApJ, 690, 1250
Salvato M. et al., 2011, ApJ, 742, 61
Sanders D. B., Mirabel I. F., 1996, ARAA, 34, 749
Sanders D. B., Soifer B. T., Elias J. H., Neugebauer G., Matthews K., 1988, ApJ, 328, L35
Sanders D. B., Phinney E. S., Neugebauer G., Soifer B. T., Matthews K., 1989, ApJ, 347, 29
Sanders D. B. et al., 2007, ApJS, 172, 86
Santini P. et al., 2009, A&A, 504, 751
Santini P. et al., 2012, A&A, 540, A109
Schawinski K. et al., 2006, Nat, 442, 888
Schawinski K., Treister E., Urry C. M., Cardamone C. N., Simmons B., Yi S. K., 2011, ApJ, 727, L31
Schulze A., Wisotzki L., 2010, A&A, 516, A87
Scoville N. et al., 2007, ApJS, 172, 1
Shankar F., 2009, NewAR, 53, 57
Shankar F., Weinberg D. H., Miralda Escudé J., 2009, ApJ, 690, 20
Shao L. et al., 2010, A&A, 518, L26
Silk J., Rees M. J., 1998, A&A, 331, L1
Silverman J. D. et al., 2008, ApJ, 675, 1025
Silverman J. D. et al., 2009, ApJ, 696, 396
Smolčić V., 2009, ApJ, 699, L43
Smolčić V. et al., 2008, ApJS, 177, 14
Soltan A., 1982, MNRAS, 200, 115
Somerville R. S., Primack J. R., Faber S. M., 2001, MNRAS, 320, 504
Springel V., Di Matteo T., Hernquist L., 2005, ApJ, 620, L79
Stern J., Laor A., 2012, MNRAS, 423, 600
Stern D. et al., 2005, ApJ, 631, 163
Strateva I. et al., 2001, AJ, 122, 1861
Taniguchi Y. et al., 2007, ApJS, 172, 9
Thomas D., Maraston C., Schawinski K., Sarzi M., Silk J., 2010, MNRAS, 404, 1775
Tommasin S. et al., 2012, ApJ, 753, 155
Trakhtenbrot B., Netzer H., Lira P., Shemmer O., 2011, ApJ, 730, 7
Tremaine S. et al., 2002, ApJ, 574, 740
Trichas M. et al., 2012, ApJS, 200, 17
Trump J. R. et al., 2007, ApJS, 172, 383

Trump J. R. et al., 2009, ApJ, 696, 1195
Ueda Y., Akiyama M., Ohta K., Miyaji T., 2003, ApJ, 598, 886
Wilkins S. M., Trentham N., Hopkins A. M., 2008, MNRAS, 385, 687
Wuyts S. et al., 2011, ApJ, 742, 96
Xue Y. Q. et al., 2010, ApJ, 720, 368

APPENDIX A: TWO-COMPONENT VERSUS ONE-COMPONENT SED FITTING

In this appendix, we discuss the differences between the measured galaxy parameters that are obtained when using the SED fitting with two components (AGN and galaxy) compared to what would be obtained when ignoring the AGN component, as in other studies (e.g. Silverman et al. 2009; Rosario et al. 2011). The colour code for the symbols in the figures is the same as used throughout the paper.

A1 Rest-frame magnitudes and galaxy colours

Rest-frame magnitudes in any given band are computed by integrating the rest-frame galaxy template found as the best-fitting solution in the corresponding filter. Thus, they are related to the type of galaxy chosen as the best fit (which can introduce a difference of up to 4 mag) and, more importantly, to the normalization of the template, which is directly related to the total flux emitted by the source. For obscured AGNs, the total emission in the U and B bands ($\lambda_{\rm eff} \sim$ 3800–4500 Å) is hardly influenced by the nuclear emission, which is, at these wavelengths, partially or entirely absorbed (see the lower panels of Fig. 3). For this reason, the shape of the chosen galaxy at these wavelengths and the normalization of the template should not strongly change when considering (or not) the AGN component. In order to test the effect of using a one- or two-component fitting, we rerun our code using only galaxy templates. As shown in Fig. A1, the U and B rest-frame host galaxy magnitudes of obscured AGNs with [e.g. $U(\mathrm{SED_{AGN+GAL}})$] and without [e.g. $U(\mathrm{SED_{GAL}})$] the AGN component lie in a one-to-one relation with a roughly symmetric scatter of about 0.4 mag (red, yellow and green symbols). In contrast, the correlation is much worse for unobscured AGNs (blue circles) whose U- and B-band rest-frame magnitudes with and without an AGN can be up to more than 2 mag different. Moreover, the spread is not symmetric, and not using the AGN component always leads to a much brighter galaxy. This is because unobscured AGN emission is strong in the U and B bands (see the upper panels of Fig. 3); that is, not taking into account the AGN component corresponds to additional attributed flux to the galaxy, hence resulting in an overestimation of the galaxy luminosity. The conclusion is different if we look at the comparison in the K band ($\lambda_{\rm eff} \sim 2.2 \times 10^4$ Å). Unlike the U and B bands, at these wavelengths the contribution of AGN emission can be relevant, even in obscured sources (see the bottom-right panel of Fig. 3). For this reason, the K-band magnitude of both obscured and unobscured AGN hosts will be overestimated (i.e. up to 2 mag brighter) if we do not take into account the AGN component.

The objects whose host luminosities are less affected by the presence of an AGN component are the optically selected type 2 AGNs without X-ray emission (triangles). For these objects, the obscuration must be very high in order to suppress the AGN emission up to K-band wavelengths. As suggested in Section 3.6, these sources are likely to be Compton-thick AGNs, and hence a high level of obscuration is expected.

A2 Host-galaxy stellar mass and SFR

Apart from the rest-frame luminosity in various bands, the other two important parameters that we derive from the SED fitting are the stellar masses and the SFRs. Besides the dependence on the template normalization, these physical parameters are also linked to the principal parameters used in the SED fitting, that is, the e-folding time τ, the age since the beginning of SF $t_{\rm age}$ and, to some degree, the extinction $E(B-V)$. All SFHs considered in this paper

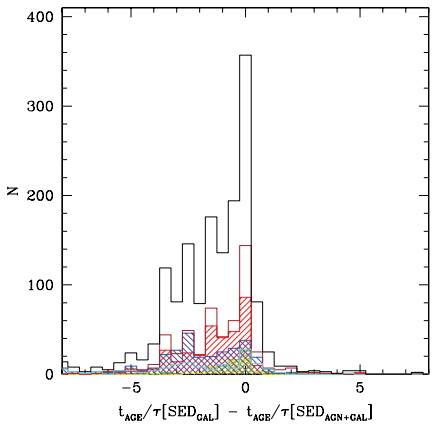

Figure A2. Comparison between the SED fitting parameters $t_{\rm age}/\tau$ obtained using the two-component SED fitting (GAL+AGN), presented in this paper, versus those obtained using a standard SED fitting considering only the galaxy component.

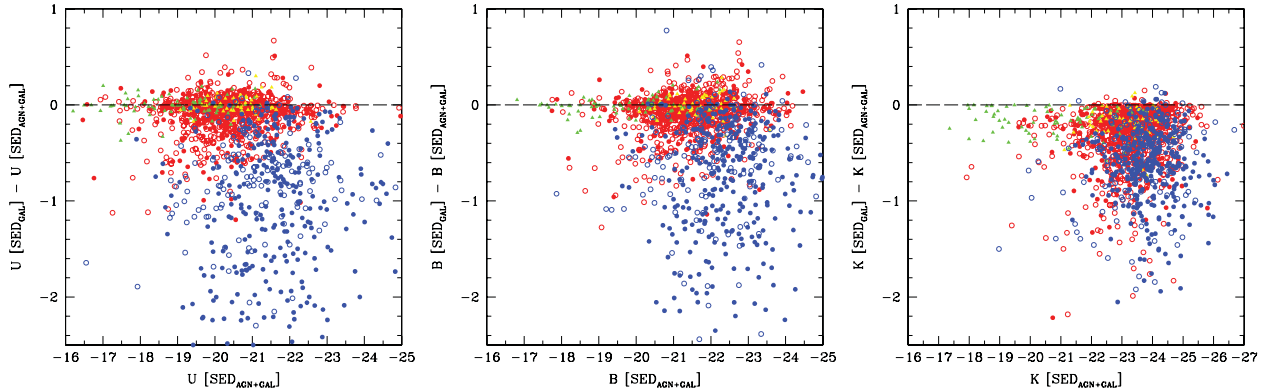

Figure A1. Comparison between the rest-frame U-, B- and K-band magnitudes of our AGN host sample, obtained using the two-component SED fitting presented in this paper, versus those obtained using a standard SED fitting considering only the galaxy component.

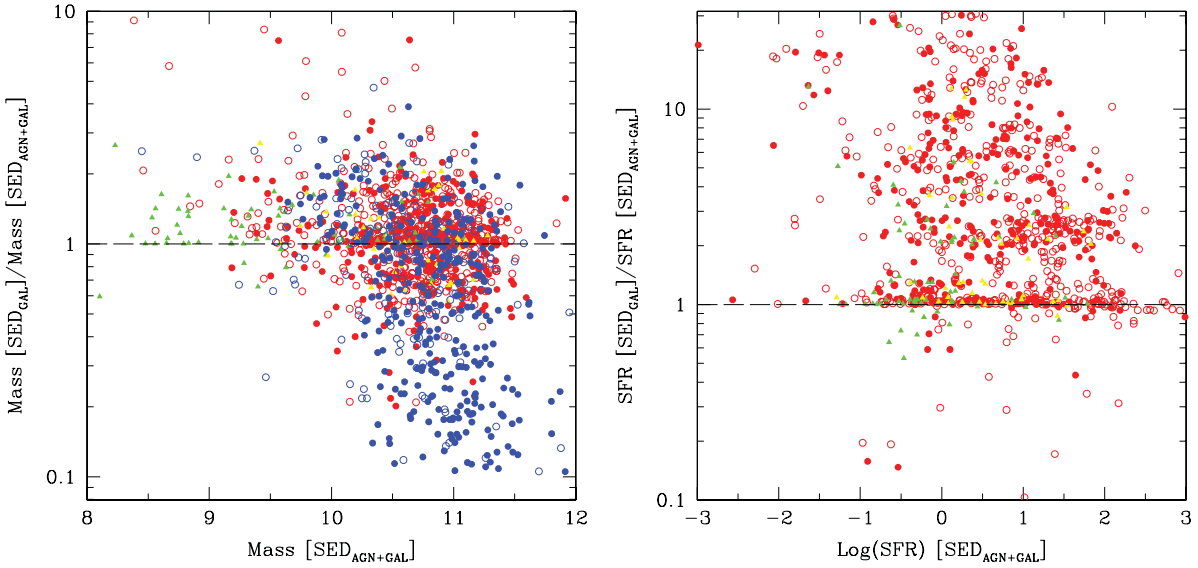

Figure A3. Comparison between the stellar masses and SFRs of our AGN hosts obtained using the two-component SED fitting presented in this paper versus those obtained using a standard SED fitting considering only the galaxy component.

have an exponentially decreasing shape with SFR $\propto e^{-t_{age}/\tau}$ (except one, which is constant). For such an exponentially decreasing SFH, a large value of t_{age}/τ implies a quasi-passive evolution with a negligible ongoing SF and it also implies that the majority of the stars in the galaxy are old (Grazian et al. 2007).

Fig. A2 shows the difference in the t_{age}/τ ratio when considering or not considering the AGN component in the SED fitting. In almost all cases, not considering the AGN component results in lower values of t_{age}/τ, which directly translates into galaxies with higher SFRs, as can be seen in the right panel of Fig. A3. For obscured type 2 AGNs in the COSMOS sample, this panel shows the ratio between the SFR obtained without the AGN component and once the AGN component is taken into account. For about 40 per cent of the objects, the two values agree within a factor of 2, while for the rest of the objects the SFR estimated without the AGN component is up to more than 20 times higher than the one obtained including the AGN.

Finally, the measured stellar masses show a non-negligible spread in both directions, for both obscured and unobscured AGNs. As shown in the left panel of Fig. A3, the masses computed not including the AGN component range from six times higher (two times if we consider only spectroscopic redshifts) to 10 times smaller (i.e. in unobscured AGNs).

Summarizing, for obscured AGNs, a standard galaxy SED fitting technique that does not include a nuclear component results in small differences in rest-frame U and B absolute magnitudes ($\sim$0.4 mag brighter) and larger differences for the K-band magnitudes ($\sim$0.8 mag brighter). The mass and SFR estimates are instead much more affected (up to a factor of 6 in mass, and SFR up to 20 times higher). This is not the case for optically selected type 2 AGNs with no X-ray counterparts for which, because of the extremely high obscuration, the differences in magnitudes, mass and SFR are relatively small when considering (or not) the AGN component. For unobscured AGNs, however, all derived quantities (absolute magnitudes, stellar mass and SFR) are strongly offset from the values obtained when taking into account the AGN component.

In conclusion, while the SED fitting technique that considers only the galaxy component can be used to roughly estimate the blue colours of obscured AGN hosts, caution has to be used when deriving red magnitudes (e.g. M_K) and physical parameters (e.g. stellar mass and SFRs). In contrast, for unobscured AGN hosts, no study can be performed without properly taking into account the AGN component.

APPENDIX B: COMPARISON BETWEEN DIFFERENT SFR INDICATORS

Evidences of SF in galaxies can be seen in both the UV and IR bands. The UV emission is a direct measure of the starlight produced by a mixture of young (O, B and A type) stars. Using the UV emission to measure the SFR has the limitation that part of the emitted radiation is absorbed, which is not taken into account. However, when using only the IR emission to constrain the SFR, the assumption is that a dominant fraction of the total bolometric stellar luminosity of the star-forming population is absorbed and reradiated as thermal IR dust emission. Thus, the most robust way to compute the SFR is to sum both the unobscured UV SF with the obscured UV SF re-emitted in the IR by dust (SFR$_{UV+IR}$).

Because of the advent of the *Herschel* Space Observatory, deep FIR observations of the COSMOS field are now available. Here, we use the deep COSMOS PACS 100- and 160-μm observations of the PEP[12] (Lutz et al. 2011) guaranteed time key programme, reaching a 3σ limit of $\sim$4 and $\sim$10 mJy at 100 and 160 μm, respectively (see Berta et al. 2011). Using these observations, we infer the obscured SFR of our host galaxies and test the accuracy of SFRs derived from SED fitting.

Only a small fraction of our sources ($\sim$10 per cent) are detected above the 3σ limits of the PEP catalogues. For these sources, we computed the SFR using their IR luminosities and assuming the L_{IR} to SFR conversion of Kennicutt (1998); that is, $SFR = 10^{-10}L_{IR}$ for a Chabrier IMF. The IR luminosity L_{IR} is derived from the monochromatic luminosity at the longest PACS wavelength available, converted using the IR SED templates of Chary & Elbaz (2001), because it has been proven to provide reliable conversions over a broad range of redshift (Elbaz et al. 2011). In the left panel of

[12] http://www.mpe.mpg.de/ir/Research/PEP/index.php

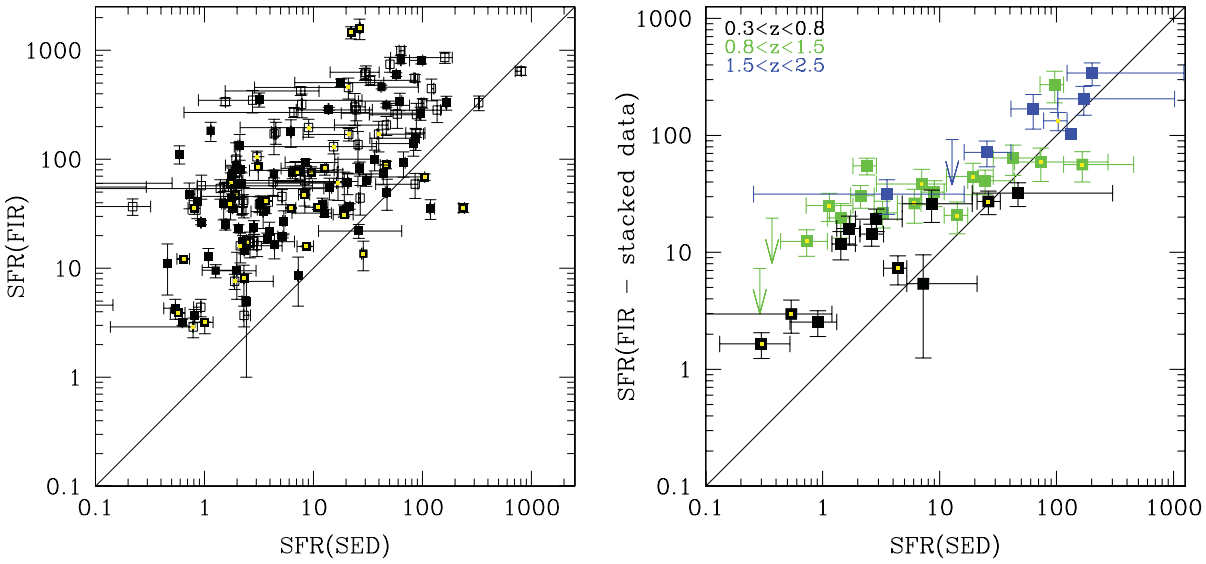

Figure B1. Left: comparison between the SFR derived from the SED fitting and the SFR derived from the FIR band for the sources detected in at least one FIR band. Open and filled squares correspond to objects with photometric and spectroscopic redshifts, respectively. Right: comparison between the SFR derived from the SED fitting and the SFR derived from the FIR using stacked data (binning in SFR_{SED} and redshift). Different colours correspond to different redshift ranges: $0.3 < z < 0.8$, black; $0.8 < z < 1.5$, green; $1.5 < z < 2.5$, blue.

Fig. B1, we show the comparison between the SFRs computed using the IR data and those computed with the two-component SED fitting. As can be seen, the majority of the sources lie above the one-to-one relation, with the SFR computed from the SED being lower than the one derived from the IR. This is driven by a selection effect that is a result of the fact that, given a broad distribution of SFR_{IR}/SFR_{SED} values, the *Herschel*-detected sources will be the most powerful ones corresponding, by definition, to the upper envelope of the SFR_{IR}/SFR_{SED} distribution.

However, to test the accuracy of the SFR derived from the SED fitting, we also have to rely on a statistical approach using a stacking analysis. Images of galaxies in a given redshift and SFR bin are stacked, allowing us to detect their typical FIR flux densities well below the detection limits of the current PACS observations (the noise of the stacked images scales roughly with the square root of the number of the stacked sources). We start by separating our galaxies into the usual three redshift bins (i.e. $0.3 < z < 0.8$, $0.8 < z < 1.5$ and $1.5 < z < 2.5$). Within each redshift bin, we then separate our galaxies per SFR bin. The sizes of the SFR bins are optimized in order to obtain the highest number of bins with clear PACS detections (i.e. $>3\sigma$). The first SFR bin is initialized with a size of 0.1 dex and an upper range equal to the highest SFR observed in the redshift bin. Then, we progressively decrease the value of the lower range of the bin (using steps of 0.1 dex), until we obtain a 3σ detection in at least one PACS passband. The lower range of the bin is then taken as the upper range of the next SFR bin, and the same procedure is repeated in order to find the size of the new bin. The process is stopped once the lower range of the current bin reaches the lowest SFR observed in the redshift bin. The photometry of the stacked images is measured using the PSF fitting analysis presented by Magnelli et al. (2009, 2011) and using the relevant PSFs and aperture corrections of the COSMOS field (Berta et al. 2011). Errors are measured using a standard bootstrap analysis, which takes into account both the photometric uncertainties of the stacked images and the intrinsic dispersion of the underlying galaxy population.

We derive the obscured SFR (i.e. SFR_{IR}) of each redshift and SFR bin using their stacked PACS flux densities and the Chary & Elbaz (2001) SED library. Where there is only one PACS detection (at 100 or 160 μm), SFR_{IR} is obtained simply using the scaled Chary & Elbaz (2001) SED library. Where there are two PACS detections (at 100 and 160 μm), we fit the PACS flux densities with the Chary & Elbaz (2001) SED library, leaving the normalization of each template as a free parameter. In both cases, SFR_{IR} is given by integrating the best Chary & Elbaz (2001) SED template from 8 to 1000 μm (L_{IR}) and using the standard SFR–L_{IR} relation of Kennicutt (1998) for a Chabrier (2003) IMF [$SFR\ (M_\odot\ yr^{-1}) = 1 \times 10^{-10} L_{IR}\ (L_\odot)$].

In the right panel of Fig. B1, for type 2 AGNs we compare the SFRs computed using the stacked IR data and that computed with the two-component SED fitting. Above $\sim 20\,M_\odot\ yr^{-1}$, the comparison is good, with the SFR derived from the SED fitting being slightly underestimated, while at low SFR ($<20\,M_\odot\ yr^{-1}$) the disagreement becomes evident. However, this trend seems to be in contrast with previous studies that were based on samples of normal galaxies. In fact, Santini et al. (2009) and Wuyts et al. (2011) have studied a sample of galaxies at high redshift (up to $z \sim 3$) and they have found an overall good correspondence between SFR_{UV+IR} and SFR_{SED} up to intermediate SFRs, while at the largest SFR ($>50\,M_\odot\ yr^{-1}$) the discrepancy was significant and hence worrisome. However, Wuyts et al. (2011) have shown that the deviation from a one-to-one relation between SFR_{UV+IR} and SFR_{SED} is most severe for galaxies that have a large SFR_{IR}/SFR_{UV} ratio. This is not the case for most of our sources (especially at high X-ray luminosities) for which the IR emission is dominated by the emission from hot dust heated by the AGN rather than young stars (see Section 3.6 and Fig. 8). However, for the sources for which this is not the case, we still assume that the IR band ($\sim$12 μm at $z \sim 1$) is dominated by the AGN emission. This assumption can lead us to underestimate the SFR, attributing the IR emission resulting from obscured young stars to the AGN component. However, the IR estimation of the SFR assumes zero contamination from the AGN component in the IR and attributes

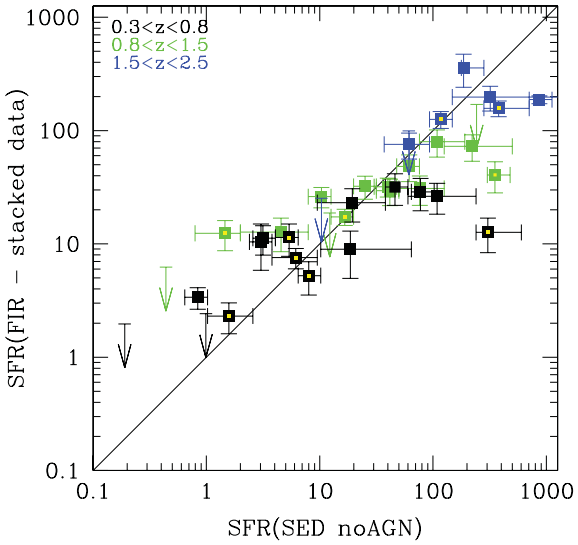

Figure B2. Comparison between the SFR derived from the SED fitting without an AGN component and the SFR derived from the FIR using stacked data (binning in SFR_{SED} and redshift). Different colours correspond to different redshift ranges: $0.3 < z < 0.8$, black; $0.8 < z < 1.5$, green; $1.5 < z < 2.5$, blue.

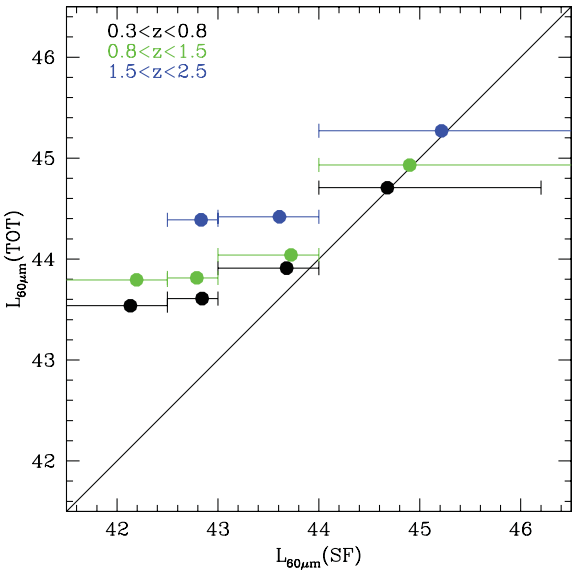

Figure B3. Mean total (AGN+SF) $L_{60\mu m}$ versus mean $L_{60\mu m}$ resulting from the SF contribution only, obtained using a mock catalogue of galaxies with different levels of SF and AGN contaminations (Rosario et al. 2012), statistically weighed using the observed L_{bol} distribution of our AGN sample in the usual three redshift bins. Error bars show the range in $L_{60\mu m}$(SF) used in the computation of the mean.

the whole IR flux to the emission of the young stars. Again, this assumption, which is reasonable in most cases, can have the effect of overestimating the SFR, and the effect would be higher at low SFR where the contamination from the AGN is a significant fraction of the IR emission. As explained below, we have performed two types of test to explore the above hypotheses. First of all, we compared the SFR obtained from the SED fitting without including the AGN component with the IR SFR (see Fig. B2). In this case, we find that there is a much better agreement between the two SFR indicators at low SFR. However, at high SFR, we no longer have an agreement (i.e the SFR from the SED fitting is higher than that derived from the FIR band; see Fig. B2). The result of this comparison indicates that the discrepancy seen in the right panel of Fig. B1 can be partly a result of the fact that the SED fitting method tends to overestimate the AGN component, thus underpredicting the SFR.

However, to test the AGN contamination of the FIR bands, we built a mock catalogue of galaxies with different levels of SF and AGN contaminations.[13] From this catalogue and using the computed weights, we calculated the average $L_{60\,\mu m}$ resulting from the AGN+SF as a function of $L_{60\,\mu m}$ resulting from the SF only. We find that, if we include the lowest luminosity tail of the AGN distribution (i.e. AGN with L_{bol} down to 10^{41} erg s^{-1}), there is hardly any difference between L_{SF} and L_{TOT} because the weighed mock is dominated by low-luminosity AGNs. In contrast, if we restrict the comparison to $L_{bol} > 10^{44}$ erg s^{-1}, we find that the 60-μm luminosity is contaminated by the AGN contribution, especially at low $L_{60\,\mu m}$. Because $L_{60\,\mu m}$ is commonly used as a proxy for L_{IR} and, therefore, SFR, this plot can be directly compared with the right panel of Fig. B1. As can clearly be seen, the two plots show very good agreement in the overall shape.

To perform a direct comparison with the observed trend in the right panel of Fig. B1, we have used the same mock catalogue as before, which this time is weighed using the observed L_{bol} distribution of our sample in the usual three redshift bins. The result is shown in Fig. B3, where we plot the mean total (AGN+SF) $L_{60\,\mu m}$ versus the mean $L_{60\,\mu m}$ resulting from the SF contribution only. As is clearly seen in the plot, the shape of the points in this figure and in Fig. B1 (right panel) are in perfect agreement. This demonstrates that the trend observed when comparing SFR–IR and SFR–SED is mostly driven by the effect of the AGN contamination on the 60-μm luminosities. This effect is strong for the *XMM*–COSMOS sample, which is dominated by high-luminosity AGNs, but it can be much smaller, or even negligible, when considering lower-luminosity samples, as demonstrated in the above test.

We conclude that the discrepancy seen in the right panel of Fig. B1 is the result of a combination of two effects: (i) the tendency for the SED fit to overestimate the AGN component, and thus to underpredict the SFR; (ii) the FIR overestimation of the SFR, especially at high AGN luminosities and low SFR, where the AGN contamination in the IR band is not negligible, as demonstrated by the tests performed using the mock catalogue.

APPENDIX C: AGN HOST GALAXY PARAMETERS DERIVED FROM THE TWO-COMPONENT SED FITTING

Here, as examples, we report the first 16 lines of the tables published online (see Supporting Information).

[13] The intrinsic AGN luminosity at 60 μm was estimated using the mean SED (normalized to the bolometric luminosity of the AGN) from Mullaney et al. (2011) and assuming a scatter around this value of 0.3 dex, which corresponds to the expected scatter in the SED shapes from Rosario et al. (2012). Because the simulation assumes that the AGNs have a uniform distribution in AGN bolometric luminosity, we used the bolometric luminosity function of Hopkins et al. (2007) to derive statistical weights for the AGN sample, depending on their L_{bol}. We applied them to obtain the real AGN distribution.

Table C1. Source identification. This is a sample of the full table, which is available as Supporting Information with the online version of the article. The columns are: (1) *XMM* ID from Cappelluti et al. (2009); (2) zCOSMOS ID (Lilly et al. 2009) when the object is in the BPT selected sample (Bongiorno et al. 2010); (3) zCOSMOS ID (Lilly et al. 2009) when the object is in the Ne v-selected sample (Mignoli et al., in preparation); (4) photometric ID from Ilbert et al. (2009); (5) RA; (6) Dec.; (7) *i*-band apparent magnitude; (8) spectroscopic redshift; (9) photometric redshift when the spectroscopy is not available (Salvato et al. 2009); (10) classification derived from the spectra, where 1 = unobscured AGNs, 2 = obscured AGNs classified from the X-ray (Brusa et al. 2010), 22.2 = obscured AGNs selected using the BPT diagram (Bongiorno et al. 2010) and 22.4 = obscured AGNs selected through the [Ne v] line (Mignoli et al., in preparation); (11) classification derived from the SED (Salvato et al. 2011) when the spectra are not available (1 = unobscured AGNs; 2 = obscured AGNs).

XMM ID	zC-ID(BPT)	zC-ID(NeV)	PHOT-ID	α_{J2000}	δ_{J2000}	I_{AB}	z_{spec}	z_{phot}	$Class_{spec}$	$Class_{phot}$
1	–	–	786683	150.10521	1.981182	19.12	0.373	–	1.0	−99.0
6	–	–	767213	150.17979	2.110378	18.34	0.360	–	1.0	−99.0
19	–	–	1024197	149.99362	2.258538	20.12	0.659	–	2.0	−99.0
26	–	–	1627881	150.39963	2.687948	18.25	0.216	–	2.0	−99.0
21	–	–	1425464	150.23079	2.578183	20.12	1.403	–	1.0	−99.0
63	826475	–	1067126	149.78189	2.139104	18.97	0.355	–	2.0	−99.0
181	–	–	751114	150.26744	2.055690	24.93	–	1.719	−99.0	2.0
182	–	–	1215214	150.32085	2.332960	25.48	–	2.030	−99.0	2.0
184	–	–	1453559	150.22397	2.550771	22.26	–	1.111	−99.0	2.0
185	–	–	808830	149.99526	2.006661	21.90	–	2.189	−99.0	1.0
–	803488	–	272411	150.65016	1.657098	20.43	0.587	–	22.2	−99.0
–	–	803886	286168	150.52974	1.725584	20.93	0.896	–	22.4	−99.0
–	–	804237	291946	150.42996	1.684480	22.30	1.000	–	22.4	−99.0
–	804862	–	303288	150.27359	1.778924	22.44	0.370	–	22.2	−99.0
–	–	804431	308347	150.38304	1.745181	22.06	0.702	–	22.4	−99.0
–	805059	–	311801	150.23288	1.724173	22.36	0.338	–	22.2	−99.0
...	...	...	...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...	...	...	...

Table C2. Parameters derived from the SED fitting. This is a sample of the full table, which is available as Supporting Information with the online version of the article. The columns are: (1), (2) and (3) as above; (4) extinction $E(B-V)$ of the AGN component; (5) rest-frame *U*-band magnitude of the host galaxy not corrected for extinction; (6) rest-frame *B*-band magnitude of the host galaxy not corrected for extinction; (7) host galaxy stellar mass; (8) flag indicating whether the previous value is a measurement (=0), an upper limit (=1) or if there is no measurement (=2); (9) host galaxy SFR; (10) flag indicating whether the previous value is a measurement (=0), an upper limit (=1) or if there is no measurement (=2 or = 3); (11) extinction $E(B-V)$ of the galaxy component.

XMM	zC(BPT)	zC(NeV)	E^{AGN}_{B-V}	$M^{gal}_{U_{jkc}}$	$M^{gal}_{B_{jkc}}$	$\log(M^{gal}_{*}\ (M_{\odot})$	$f_{M^{gal}_{*}}$	$\log\,(SFR^{gal}\ (M_{\odot}\ yr^{-1})$	$f_{SFR^{gal}}$	E^{gal}_{B-V}
1	–	–	0.0	−20.235	−20.641	9.238	0	−99.0	3	0.3
6	–	–	0.1	−21.655	−22.406	10.975	0	−99.0	3	0.0
19	–	–	0.3	−21.352	−22.254	11.158	0	1.459	0	0.2
26	–	–	5.6	−20.108	−20.983	10.546	0	0.364	0	0.0
21	–	–	0.1	−99.000	−99.000	10.537	1	−99.0	3	0.0
63	826475	–	8.7	−20.647	−21.582	10.994	0	0.861	0	0.1
181	–	–	0.6	−20.417	−21.728	11.051	0	1.628	0	0.5
182	–	–	0.6	−20.746	−22.006	10.984	0	−0.152	0	0.1
184	–	–	9.0	−21.045	−21.952	10.866	0	1.005	0	0.1
185	–	–	0.0	−23.113	−24.295	11.296	0	−99.0	3	0.0
–	803488	–	9.0	−21.056	−22.238	10.483	0	−0.818	0	0.0
–	–	803886	9.0	−21.468	−22.497	10.979	0	0.997	0	0.1
–	–	804237	3.3	−20.700	−21.656	10.454	0	1.536	0	0.4
–	804862	–	2.1	−18.115	−18.743	8.7610	0	−0.296	0	0.0
–	–	804431	0.3	−19.425	−20.458	10.325	0	−0.15	0	0.0
–	805059	–	0.3	−17.459	−18.166	8.6090	0	−0.622	0	0.0
...	...	...	...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...	...	...	...

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Table C1. Source identification.
Table C2. Parameters derived from the SED fitting.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

This paper has been typeset from a TEX/LATEX file prepared by the author.