A&A 433, 855–873 (2005)
DOI: 10.1051/0004-6361:20041340
© ESO 2005

Astronomy & Astrophysics

X-ray spectra of *XMM-Newton* serendipitous medium flux sources

S. Mateos[1,5], X. Barcons[1], F. J. Carrera[1], M. T. Ceballos[1], A. Caccianiga[2], G. Lamer[3], T. Maccacaro[2], M. J. Page[4], A. Schwope[3], and M. G. Watson[5]

[1] Instituto de Física de Cantabria (CSIC-UC), 39005 Santander, Spain
[2] Osservatorio Astronomico di Brera, via Brera 28, 20121 Milano, Italy
[3] Astrophysikalisches Institut Potsdam, An der Sternwarte 16, 14482 Potsdam, Germany
[4] Mullard Space Science Laboratory, UCL, Holmbury St Mary, Dorking, Surrey RH5 6NT, UK
[5] Department of Physics and Astronomy, University of Leicester, LE1 7RH, UK
e-mail: sm279@star.le.ac.uk

Received 21 May 2004 / Accepted 14 December 2004

Abstract. We report on the results of a detailed analysis of the X-ray spectral properties of a large sample of sources detected serendipitously with the *XMM-Newton* observatory in 25 selected fields, for which optical identification is in progress. The survey covers a total solid angle of ~3.5 deg^2 and contains 1137 sources with $\sim 10^{-15} < S_{0.5-10} < 10^{-12}$ erg cm^{-2} s^{-1} with good enough spectral quality as to perform a detailed X-ray spectral analysis of each individual object. We find evidence for hardening of the average X-ray spectra of the sources towards fainter fluxes, and we interpret this as indicating a higher degree of photoelectric absorption amongst the fainter population. Absorption is detected at 95% confidence in 20% of the sources, but it could certainly be present in many other sources below our detection capabilities. For Broad Line AGNs (BLAGNs), we detect absorption in ~10% of the sources with column densities in the range $10^{21}-10^{22}$ cm^{-2}. The fraction of absorbed Narrow Emission Line galaxies (NELGs, most with intrinsic X-ray luminosities $>10^{43}$ erg s^{-1}, and therefore classified as type 2 AGNs) is significantly higher (40%), with a hint of moderately higher columns. After correcting for absorption, we do not find evidence for a redshift evolution of the underlying power law index of BLAGNs, which stays roughly constant at $\Gamma \sim 1.9$, with intrinsic dispersion of 0.4. A small fraction (~7%) of BLAGNs and NELGs require the presence of a soft excess, that we model as a black body with temperature ranging from 0.1 to 0.3 keV. Comparing our results on absorption to popular X-ray background synthesis models, we find absorption in only ~40% of the sources expected. This is due to a deficiency of heavily absorbed sources (with $N_{\rm H} \sim 10^{22}-10^{24}$ cm^{-2}) in our sample in comparison with the models. We therefore conclude that the synthesis models require some revision in their specific parameters.

Key words. X-rays: general – surveys – galaxies: active

1. Introduction

Extensive studies with different satellites have been performed during the last decades to unveil the origin of the cosmic X-ray background radiation (XRB). In the 0.5–2 keV band, the *ROSAT* deep surveys in the Lockman Hole (Hasinger et al. 1998) resolved ~70–80% of the XRB into discrete sources down to a flux limit of $\sim 10^{-15}$ erg cm^{-2} s^{-1}. The deepest *Chandra* observations have now resolved ~90% of the soft XRB reaching a flux limit of $\sim 2 \times 10^{-16}$ erg cm^{-2} s^{-1} (Moretti et al. 2003).

In the hard X-ray sky, important contributions to the population studies were performed by the *ASCA* (Georgantopoulos et al. 1997; Cagnoni et al. 1998; Boyle et al. 1998; Ueda et al. 1998, 1999; Della Ceca et al. 1999) and *BeppoSAX* (Giommi et al. 2001) satellites. They resolved ~30% of the 2–10 keV hard XRB (Ueda et al. 1998; Fiore et al. 2001) down to fluxes of 5×10^{-14} erg cm^{-2} s^{-1}. *Chandra* and *XMM-Newton* have reached fluxes 100 times fainter, resolving ~50–70% of the hard XRB (Mushotzky et al. 2000; Barger et al. 2001; Fiore et al. 2001; Giacconi et al. 2001; Hasinger et al. 2001; Tozzi et al. 2001; Baldi et al. 2002; Worsley et al. 2004). The fraction of resolved X-ray background has recently been increased to ~90% in the 2–10 keV and ~94% in the 0.5–2 keV bands by Cowie et al. (2002) and Moretti et al. (2003), respectively.

Optical follow-up of the faint X-ray sources has revealed that the bulk of the XRB is made up of type 1 and type 2 AGNs. The results of optical identification programmes suggest that the fractional contribution of type 2 AGNs is becoming increasingly important towards fainter fluxes (e.g. Fiore et al. 2003).

Now that the discrete origin of the XRB has been confirmed all over the X-ray energy band (0.1–10 keV), efforts are being concentrated on the study of the source populations that make up the XRB. The nature and cosmic evolution of these sources is still controversial (Barger et al. 2002; Mainieri et al. 2002; Fiore et al. 2003; Ueda et al. 2003).

At the faintest X-ray fluxes an important contribution to the XRB emission is expected to come from type 2 AGN (Setti & Woltjer 1989; Gilli et al. 2001; Comastri et al. 1995).

Article published by EDP Sciences and available at http://www.aanda.org or http://dx.doi.org/10.1051/0004-6361:20041340

This has been confirmed observationally (Mainieri et al. 2002). The distribution of intrinsic absorption among the sources has only been measured observationally for the brightest sources at low redshifts, but is still unknown at high redshifts. Recent *Chandra* and *XMM-Newton* studies have revealed that the widely accepted correlation between the optical obscuration and X-ray absorption (type 1 AGN being unabsorbed and type 2 AGN being absorbed) is not always true. Several cases of type 1 AGN suffering from significant X-ray absorption have been found (Mittaz et al. 1999; Fiore et al. 2001; Page et al. 2001; Schartel et al. 2001; Tozzi et al. 2001; Mainieri et al. 2002; Brusa et al. 2003; Page et al. 2003; Carrera et al. 2004; Perola et al. 2004), as well as Seyfert 2 galaxies unabsorbed in X-rays (Pappa et al. 2001; Panessa et al. 2002; Barcons et al. 2003).

The motivation for this work is to exploit the large collecting area and wide field of view (FOV) of *XMM-Newton* to perform a detailed spectral analysis of X-ray selected sources down to medium-faint fluxes. This study will allow us to put strong observational constraints on the average X-ray spectral properties of the source population that dominates the X-ray emission at the flux level of our survey, typically $\sim 10^{-14}$ erg cm^{-2} s^{-1} in the 0.5–10 keV energy band. Therefore we have conducted a large (total solid angle of $\sim$3.5 deg^2) and medium to deep (0.5–10 keV flux in the range $\sim 10^{-15}$–10^{-12} erg cm^{-2} s^{-1}) survey of sources detected with the *XMM-Newton* observatory in the 0.2–12 keV energy band. We take advantage of an ongoing identification effort of sources in the fields under study, which enable us to obtain spectral properties of representative samples of identified sources. This work complements an earlier study of the X-ray spectral properties of a large sample of sources at similar fluxes, but where the count rates detected in the XMM-Newton standard energy bands (0.2–0.5, 0.5–2, 2–4.5, 4.5–7.5, 7.5–12 keV) were used to analyse the broad band X-ray spectral shape of the sources (Mateos et al. 2003).

The paper is organised as follows: Sect. 2 describes the X-ray data and the current status of the identifications; Sect. 3 describes the extraction of the X-ray spectral products; Sect. 4 describes the results of the spectral fitting; Sect. 5 discusses the average spectral properties of our sources; the X-ray spectral properties of sources identified as ALGs are described in Sect. 6; Sect. 7 compares our observational results with specific XRB synthesis models; the results of our analysis are summarised in Sect. 8. Throughout this paper we have adopted the currently favoured concordance cosmology parameters H_0 = 70 km s^{-1} Mpc^{-1}, $\Omega_{\rm m} = 0.3$ and $\Omega_\Lambda = 0.7$. The luminosity values that we use in the paper are the intrinsic luminosities of the objects corrected for absorption. All one-parameter uncertainties are computed with a delta chi-squared of 2.706, equivalent to 90% confidence region for a single parameter.

2. Observations

In this study we use 25 observations selected from the public *XMM-Newton* data archive. They cover a total solid angle of $\sim$3.5 deg^2. To obtain a clean extragalactic sample of objects and to maximise the survey efficiency, the X-ray observations were selected according to the following criteria: sky positions at high galactic latitudes ($|b| > 20°$); observations for which we had data from the European Photon Imaging Camera (EPIC)-pn detector in FULL-FRAME MODE (i.e. the EPIC-pn full FOV covered). We avoided fields containing bright point or extended targets where the wings of the target can reduce substantially the area of the survey covered. The current analysis has been performed within the framework of the tasks of the *XMM-Newton* Survey Science Centre[1].

The X-ray fields are listed in Table 1 together with their observational details.

2.1. X-ray data reduction

All the observations were processed through the *XMM-Newton* pipeline processing system that uses a suite of tasks created specifically for the analysis of *XMM-Newton* data, namely the Science Analysis System (Gabriel et al. 2004). The current version at the time of the analysis being v5.3.3. The pipeline process provides calibrated event files, images and exposure maps in five standard X-ray energy bands (0.2–0.5, 0.5–2, 2–4.5, 4.5–7.5 and 7.5–12 keV) covering the energy range where the data is best calibrated. The X-ray images are created with a pixel size of $4'' \times 4''$ and cover a FOV of $\sim 30'$ diameter. The exposure maps give information on the spatial efficiency of each camera including the energy dependent mirror vignetting. The pipeline products also provide source lists for each observation and for each EPIC camera (pn, MOS1 and MOS2).

The source detection is run independently for each of the EPIC detectors, simultaneously on the five energy bands defined above. The area of the images where the sources are searched for is defined by the SAS task `emask`, that creates a map where CCD gaps and bad pixels/columns are excluded. The source detection process starts with the task `eboxdetect` (local mode) that performs a sliding box cell detection of sources with a minimum likelihood of detection of 8. This provides an input list of source positions masked out by the SAS task `esplinemap` to build background maps for each detector and energy band. `Eboxdetect` (map mode) is run again, but this time the background is taken from the maps created by `esplinemap`. This improves significantly the detection sensitivity. Finally the task `emldetect` performs maximum likelihood Point Spread Function (PSF)[2] fits to the distribution of counts of the sources detected by `eboxdetect`, obtaining for each object the total (0.2–12 keV) count rate and the likelihood of detections (total and on each energy band). For a description of the application of Maximum Likelihood analysis to detection of sources see Cruddace et al. (1987). The sources with a total detection likelihood higher than 10 are included in the final list.

2.2. X-ray source list

The geometrical shadowing of half of the sky X-ray emission received by the MOS1 and MOS2 detectors (deviated to

[1] See `http://xmmssc-www.star.le.ac.uk`

[2] *XMM-Newton* mirror modules have a PSF with *FWHM* of $\sim 6''$.

Table 1. Observational details of the *XMM-Newton* fields.

Field	RA	Dec	b^a	N_H^{Gal}	Rev./Obs. id	Filter[b]			GTI[c](ksec)		
	(J2000)	(J2000)	(deg)	(10^{20} cm^{-2})		M1	M2	pn	M1	M2	pn
A2690	00:00:30.3	–25:07:30.0	–78.90	1.84	0088/0125310101	M	M	M	19	19	21
CL0016 + 16	00:18:33.0	+16:26:18.0	–45.54	4.07	0194/0111000101	M	M	M	31	31	29
					0194/0111000201	M	M	M	5	5	3
G133 – 69 pos 2	01:04:00.0	–06:42:00.0	–69.35	5.19	0104/0112650501	Th	M	Th	23	23	18
G133 – 69 pos 1	01:04:24.0	–06:24:00.0	–68.68	5.20	0188/0112650401	Th	M	Th	23	23	20
SDS – 1b	02:18:00.0	–05:00:00.0	–59.75	2.47	0118/0112370101	Th	Th	Th	47	48	40
					0119/0112371001	Th	Th	Th	51	51	43
SDS – 3	02:18:48.0	–04:39:13.0	–59.35	2.54	0121/0112370401	Th	Th	Th	21	21	15
					0121/0112371501	Th	Th	Th	7	7	4
SDS – 2	02:19:36.0	–05:00:00.0	–58.91	2.54	0120/0112370301	Th	Th	Th	50	50	40
Mkn 3	06:15:36.3	+71:02:15.0	22.72	8.82	0158/0111220201	M	M	M	54	54	44
MS 0737	07:44:04.5	+74:33:49.5	29.57	3.51	0063/0123100101	Th	Th	Th	39	39	20
					0063/0123100201	Th	Th	Th	19	19	20
CL0939 + 472	09:43:00.0	+46:59:30.0	48.88	1.24	0167/0106460101	Th	Th	Th	49	49	43
S5 0836 + 716	08:41:24.0	+70:53:41.0	34.43	2.98	0246/0112620101	M	M	M	4	4	25
B2 1028 + 31	10:30:59.1	+31:02:56.0	59.79	1.94	0182/0102040301	Th	Tck	Th	26	26	23
B2 1128 + 31	11:31:09.4	+31:14:07.0	72.03	2.00	0175/0102040201	Th	Tck	Th	19	23	13
Mkn 205	12:21:44.0	+75:18:37.0	41.67	3.02	0075/0124110101	M	M	M	50	50	37
MS1229.2 + 6430	12:31:32.0	+64:14:21.0	53.05	1.98	0082/0124900101	Th	Th	Th	34	34	29
HD 117555	13:30:47.0	+24:13:59.0	80.67	1.16	0199/0100240101	M	M	M	29	29	25
					0205/0100240201	M	M	M	36	36	33
UZ Lib	15:32:23.0	–08:32:05.0	36.57	8.97	0210/0100240801	M	M	M	23	23	23
PKS 2126 – 158	21:29:12.2	–15:38:41.0	–42.39	5.00	0255/0103060101	M	M	M	22	22	16
PKS 2135 – 147	21:37:45.2	–14:32:55.4	–43.85	4.70	0254/0092850201	M	M	M	15	16	28
MS2137 – 23	21:40:15.2	–23:39:41.0	–47.50	3.50	0254/0008830101	Th	Th	Th	13	14	10
PB5062	22:05:09.8	–01:55:18.0	–43.28	6.17	0267/0012440301	Th	Th	Th	31	31	28
LBQS 2212 – 1759	22:15:31.6	–17:44:05.7	–52.92	2.39	0355/0106660401	Th	Th	–	33	33	–
					0355/0106660501	Th	Th	Th	8	5	8
					0356/0106660601	Th	Th	Th	103	103	91
PHL 5200	22:28:30.4	–05:18:55.0	–49.97	5.26	0269/0100440101	Tck	Tck	Tck	45	45	43
IRAS 22491 – 18	22:51:49.4	–17:52:23.2	–61.42	2.71	0267/0081340901	M	M	M	22	22	20
EQ Peg	23:31:50.0	+19:56:17.0	–39.14	4.25	0107/0112880301	Tck	Tck	Tck	14	14	12

[a] Galactic latitude.
[b] Blocking filters: Th: Thin at 40 nm A1; M: Medium at 80 nm A1; Tck: Thick at 200 nm A1.
[c] Good time intervals after removal of flares.

RGS gratings) makes the EPIC-pn camera a factor of two more sensitive than the MOS cameras overall. Therefore, to optimise the sensitivity of the survey, we used the pipeline EPIC-pn source lists. In the cases where we had more than one observation for a field we used the source list of the deepest exposure.

We defined circular regions (with radius between 16 to 160 arcsec), to exclude from the images the areas contaminated by the emission of the targets. For the objects detected near CCD gaps the uncertainties in their coordinates and flux may be larger than the statistical uncertainties as given by the pipeline products. In order to get rid of these sources with somewhat uncertain source parameters, we have excluded zones near CCD gaps, widening them in size up to the radius that contains 80% of the PSF on each detector point. We then used these masks to remove from the source list all the objects detected close to CCD gaps. For each field and observation we have examined the images searching for spurious sources, like detections in hot pixels or bright segments not removed by the pipeline process. We have checked carefully detections close to the tails of bright and/or extended objects, that in most cases turned out to be spurious. The visual screening process removed ~10% of the objects from the original list. The total number of sources detected within these specifications was 2145.

2.3. Source detection sensitivity

The variety of data that we have used in our study, make it difficult to calculate the source detection sensitivity as a function of the X-ray flux, because in principle it depends on the exposure times of the fields and their Galactic absorbing columns, N_H (Zamorani et al. 1988).

The *XMM-Newton* EPIC-pn detector is sensitive to X-ray photons at energies ranging from ~0.2 to ~12 keV.

The sensitivity is maximum between 0.5 and 4.5 keV, where it is almost independent of energy, i.e., objects with different spectral shapes are detected with similar efficiency. Outside this energy interval, the sensitivity is a strong function of energy.

Because the band of detection of photons is wide, we do not expect the use of fields with different $N_{\rm H}^{\rm Gal}$ (ranging from 1 to 9×10^{20} cm^{-2}) to introduce any significant bias (e.g. detections in the fields with the higher values of Galaxy absorption biased towards hard[3] spectrum sources) in the broad band spectral properties of our objects at the flux limits of our survey.

However, we can see in Table 1, that the exposure times vary significantly from field to field. At 0.5–10 keV fluxes below $\sim 10^{-14}$ erg cm^{-1} s^{-1}, where only a few fields contribute to the source list, we expect our survey to be biased against objects with spectra peaking outside the 0.5–4.5 keV energy range, for example highly obscured AGNs.

To build our source lists, the *XMM-Newton* source detection algorithm was run simultaneously in the five standard *XMM-Newton* energy bands. Therefore we have analysed sources detected in at least one of these energy bands. This might be important, because the sensitivity varies significantly from band to band, and hence, different source populations may contribute at different flux levels.

In order to calculate the source detection efficiency function, we have carried out simulations. A detailed description of how we obtained this selection function is given in Appendix A. There is indeed a mild dependence of the selection function on the spectral slope Γ at any given flux, and we will correct for this effect.

Figure 1 shows the distribution of our sources in bins of 0.5–10 keV flux. The fluxes were obtained from the best fit model of each object (see Sect. 5.4). The selection of fields with a range of exposure times results in a gradual reduction of the sky coverage as we go to fainter fluxes, and is responsible for broadening the flux distribution below $\sim 10^{-14}$ erg cm^{-2} s^{-1}. Our sample is likely to be representative of the dominant X-ray source population down to a flux of $\sim 10^{-14}$ erg cm^{-2} s^{-1}. For comparison we have also plotted in Fig. 1 the distribution of flux for the sources with optical spectroscopic identifications.

The goal of the present paper is not to produce complete source lists to build up source counts or luminosity functions (which will be the subject of a forthcoming paper), but rather to study the X-ray spectral properties of representative samples within the medium flux range, where it is known that ~50% of the 0.5–2 keV (and ~40% of the 2–8 keV, see Bauer et al. 2004) accretion power in the Universe is produced. For this purpose, none of the above possible biases in the source list has any significant impact on our results, after appropriate corrections are applied.

2.4. Optical spectroscopic identifications

The selected fields are being followed-up in the optical band as part of the *XMM-Newton* Survey Science Centre XID identification programme (Watson et al. 2001). The observing time was awarded by a project named AXIS[4] ("An *XMM-Newton* International Survey", Barcons et al. 2002). The optical observations consist of high quality multicolour imaging and optical spectra of the brightest sources (the optical observations and spectroscopic identifications will be described in detail in a forthcoming paper). This is an ongoing process, whose first results were presented in Watson et al. (2001) and Barcons et al. (2002). The full list of identified sources will be made available via the *XMM-Newton* Science Archive when completed down to a 0.5–4.5 keV flux of 2×10^{-14} erg cm^{-2} s^{-1}. Two identifications in the A2690 field were from Piconcelli et al. (2002). For the time being, ~11% (232) of the sources in our sample have been identified via optical spectroscopy. We have optical identifications within all fields except one (MS1229), the fraction of identified sources per field ranging from ~1.4% to ~39%.

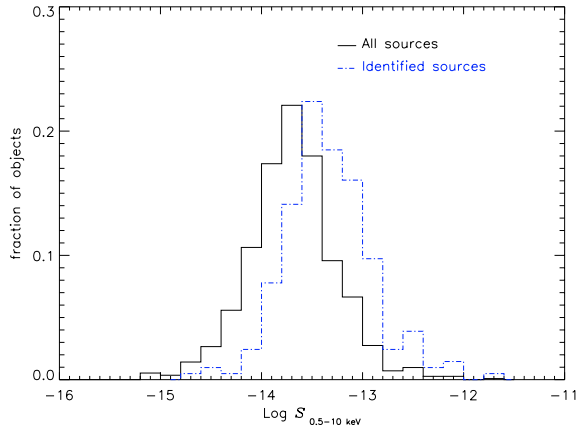

Fig. 1. $S_{0.5-10\ \rm keV}$ distribution for the whole sample (solid line) and the sub-sample of identified sources (dashed-dot line).

The objects were classified according to their optical emission line properties. Sources with spectra showing strong broad emission lines ($FWHM > 1000$ km s^{-1}) were classified as Broad Line AGNs (BLAGNs). If only narrow emission lines were seen ($FWHM < 1000$ km s^{-1}), the sources were classified (purely in optical spectroscopic terms) as Narrow Emission Line Galaxies (NELGs). Objects with galaxy-like optical spectra and no apparent emission lines were classified as absorption line galaxies (ALGs), some of which are associated to clusters. Table 2 lists the number of identified sources, along with those which have good quality X-ray spectra (see below).

Of all the identified sources, we have only analysed the spectra of BLAGNs, NELGs and ALGs.

Figure 2 shows the redshift distributions of these sources. BLAGNs are detected up to redshifts of ~3, while at the flux limit of our optical identifications we do not detect NELGs and ALGs at redshifts above ~0.5. Similar distributions, at similar fluxes, have been found previously (e.g. Green et al. 2004 for the ChaMP survey). In spite of the incomplete identifications, we do expect the identified sources to constitute a "representative" sample of the underlying population of sources at X-ray fluxes above $\sim 2 \times 10^{-14}$ erg cm^{-2} s^{-1}, where the identified fraction of objects is much higher (50%).

[3] We use the term "soft spectra" when the measured spectral photon index, Γ, is found to be ~2, the typical value for unabsorbed AGNs. We consider a spectra to be hard if $\Gamma \leq 1.5$.

[4] http://venus.ifca.unican.es/~xray/AXIS/

Table 2. Breakdown of the sources with optical spectroscopic identifications. (BLAGN: broad line AGNs; NELG: narrow emitting line galaxies; ALG: absorption line galaxies).

Object class	N^a	N good^b
BLAGN	149	141
NELG	32	29
ALG	10	7
Stars	40	32
BL Lac	1	1
Total	232	210

[a] Number of sources identified.
[b] Number of sources identified that fulfil the quality thresholds applied to the spectra (see Sect. 3).

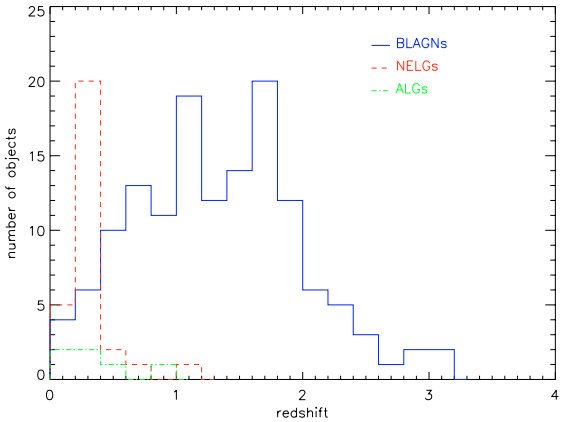

Fig. 2. BLAGNs (solid line), NELGs (dashed line) and ALGs (dashed-dot line) redshift distributions.

3. Extraction of X-ray spectral products

We have cleaned the calibrated event files with the same filtering expressions that were used by the pipeline to create the EPIC X-ray images. This filtering removes time intervals contaminated by soft proton flares, cosmic ray tracks and spurious noise events not created by X-rays. For each object in the catalogue the SAS task `region` was used to define the source and background extraction regions.

The source region shape was defined as a circle centred on the source coordinates and with a typical radius of ~14–30″, depending on the source position within the detector. The background region was defined as an annulus centred at the source position, with inner radius equal to the source extraction radius and outer radius three times the inner radius. If there was overlap with nearby sources, the sizes of source and/or background regions were reduced to eliminate the overlap. If nearby sources were inside the background region their extraction regions were excluded from the background region.

Source and background spectra were then extracted with the SAS task `evselect`. Single and double events were included for the pn detector and single-quadruple events for the MOS detectors in the energy band 0.2–12 keV. Further, only events with the highest spectral quality (`FLAG=0`) were included for the EPIC-pn data.

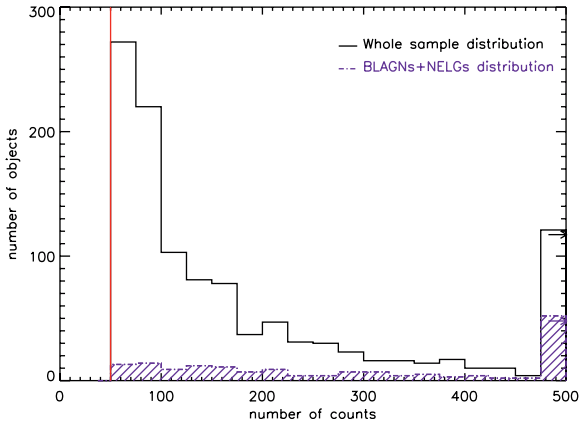

Fig. 3. Distributions of background subtracted counts (MOS+pn) for the objects in our sample (solid histogram) and for the objects classified as BLAGNs or NELGs (shaded dot-dashed histogram). The vertical line indicates one of the quality thresholds applied to the spectra. For clarity, the last bin in the plots includes the spectra with more than 500 counts.

In order to perform a proper spectral analysis, we created for each source the redistribution matrix file (RMF) and the ancillary response file (ARF), using the SAS tasks `rmfgen` and `arfgen` respectively. To maximise the signal to noise ratio, we have combined MOS1 and MOS2 source and background spectra and the corresponding response matrices. Merged source and background spectra were obtained by adding the individual spectra. Backscale values (size of the regions used to extract the spectra) and calibration matrices for the combined spectra were obtained weighting the input data with the exposure times.

For the fields with more than one observation, we added all the MOS and pn data for each epoch but MOS and pn data were not merged because of their different responses[5].

In order to permit the χ^2 minimisation technique in the fitting process, the raw spectra were binned such that each bin contained ≥10 counts (source plus background). We then imposed two quality thresholds on the grouped spectra selected for analysis: number of bins (MOS+pn) ≥ 5, and number of background subtracted counts (MOS+pn) ≥ 50.

This filtering process resulted in a final sample of 1137 sources having X-ray spectra of sufficient quality to analyse them individually (see Table 2 for the distribution of identified sources).

Figure 3 shows the quality of the selected spectra as a function of the 0.2–12 keV number of counts for the whole sample of sources (solid-line histogram) and for the objects identified as BLAGNs and NELGs (shaded histogram). The vertical line indicates the quality filter that we applied to select the spectra appropriate for the analysis. In Fig. 4 we present, as an example, the X-ray spectra of some of the objects that we have analysed.

4. Spectral analysis

The data were fitted with the XSPEC (version 11.2) software package. The spectral analysis has been carried out fitting

[5] See *XMM-Newton* handbook at `http://xmm.vilspa.esa.es`

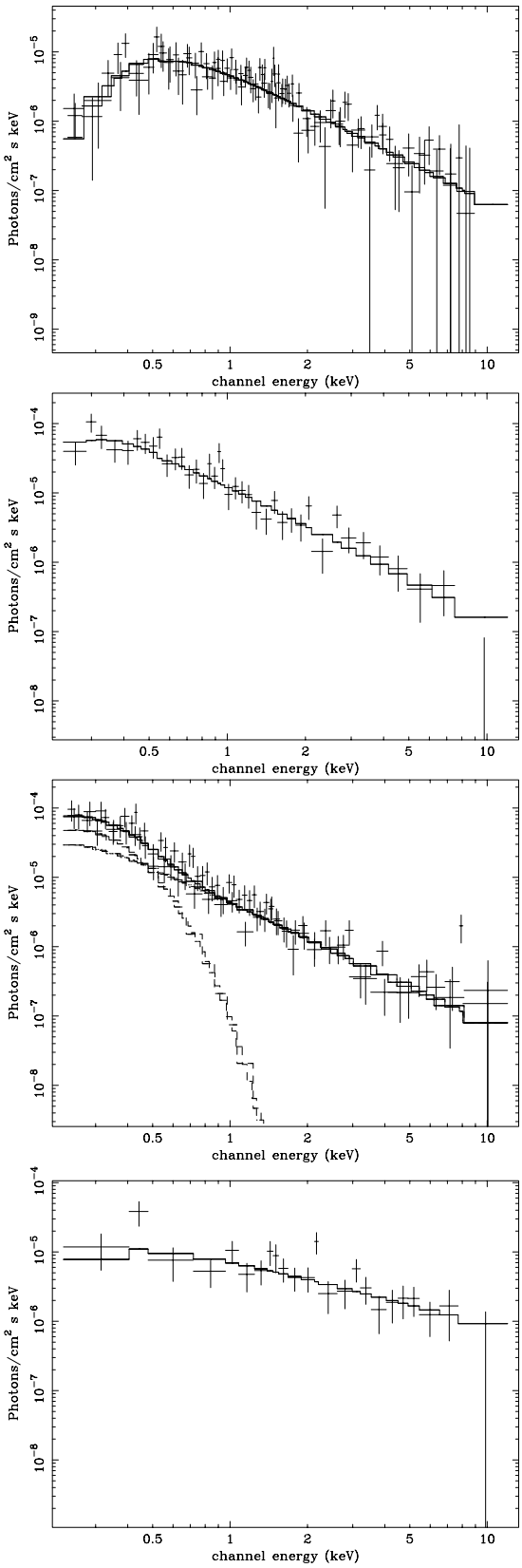

Fig. 4. From top to bottom unfolded spectra of **a) XMMU J221454.9-173949**, absorbed BLAGN; **b) XMMU J084132.3+704757**, unabsorbed NELG ($L_{2-10\,\mathrm{keV}} \sim 6 \times 10^{42}$ erg s^{-1}); **c) XMMU J33106.2+241325**, BLAGN with soft excess component (see Sect. 5.3 for details); **d) XMMU J021908.3-044731**, flat spectrum ($\Gamma = 0.94^{+0.23}_{-0.25}$) object, still not identified.

simultaneously MOS and pn spectra, forcing the models to have the same parameters in both instruments, including the normalisation (we do not expect significant calibration mismatches between instruments in the energy interval considered and for the quality level of our data).

We quantified the results of the spectral fits in terms of the null hypothesis probability $P(\chi^2)$, i.e., the probability that the observed spectrum is derived from the parent model under scrutiny. We consider a model not good enough if $P(\chi^2)$ is less than 5%.

To test the significance of additional input parameters the F-test has been applied with an adopted significance threshold of 95%. As is shown in Protassov et al. (2002), if the conditions to use the F-test statistic are not satisfied, the false positive rate might differ from the value expected for the selected confidence level. One of the conditions that must be satisfied, is that the null values of the additional parameters, should not be in the boundary of the set of possible values of the parameter. We are using the F-test to calculate the significance of detection of absorption, and one of the possible values of the absorbing column is zero. Hence, because one of the conditions for using the F-test is not satisfied, we do not know whether the fraction of spurious detections in our data differs from the 5% expected. To calculate this number, we carried out simulations of unabsorbed spectra, and then we calculated the number of cases where we detected absorption with a confidence level above 95%. We have detected absorption in ~2% of the simulated spectra. As this fraction is not significantly different from the expected value, we decided to be conservative, and hence, we have continued with 5% as the number of spurious detections, or false positive rates.

4.1. Single power law (model A)

We have started the analysis fitting a single power law to the background-subtracted spectra, hereafter model A. This model has two free parameters, the normalisation and the continuum slope Γ. To account for the effect of the galactic neutral hydrogen absorption along the line of sight, a fixed photoelectric absorption component was included. The Galaxy H column density values for each field, $N_{\mathrm{H}}^{\mathrm{Gal}}$, were extracted from the HI map of Dickey & Lockman (1990). For the selected acceptance level, 139 (12%) objects had statistically unacceptable fits with this model, the expected number for the selected confidence level being 57. Within the sub-samples of identified sources, model A could not be accepted for 10 (7%) BLAGNs and 9 (31%) NELGs.

4.2. Single power law with excess absorption (model B)

Synthesis models of the XRB, based on AGN unification schemes (Setti & Woltjer 1989; Madau et al. 1994; Comastri et al. 1995; Pompilio et al. 2000; Gilli et al. 2001) predict a large population of highly absorbed AGNs. Hence we expect a significant fraction of our objects to exhibit absorption in excess of the Galactic value. An indication of absorption is to

measure a flat spectral slope with model A. To study excess absorption, we fitted a model with two local absorbers, one with the column density fixed at the Galactic value and the other one as a free parameter $N_{\rm H}^{\rm obs}$ (model B). Note that this last term is essentially indistinguishable from an absorber with column density $\sim N_{\rm H}^{\rm obs}\,(1+z)^{2.7}$ (Barger et al. 2002; Longair 1992) at the redshift of the source, z, since absorption edges would not be detected in our highly grouped, moderate signal-to-noise spectra. This model has three free parameters, the normalisation, the continuum spectral slope Γ, and the equivalent absorbing column at $z = 0$, $N_{\rm H}^{\rm obs}$. The top panel in Fig. 5 compares the $P(\chi^2)$ distributions obtained with model A with the results from model B for the whole sample of sources. We see that the number of unacceptable fits has been reduced by a factor of ~1.5, implying the presence of intrinsically absorbed objects.

4.3. Single power law with intrinsic absorption (model C)

As an alternative, more physical parameterisation of the absorbing column of the BLAGN, NELG and ALG, we instead fitted photoelectric absorption at the redshift of the source (model C). The model has three free parameters, the normalisation, the continuum spectral slope, Γ, and the rest frame absorption, $N_{\rm H}^{\rm intr}$. Figure 5 compares the results of the fits obtained with model A and model C, for BLAGNs and NELGs. For a detailed description of the spectral properties of the 7 ALGs analysed, see Sect. 6. In the case of the BLAGNs, we did not find a substantial improvement in the overall quality of the fits, suggesting that the majority of the BLAGNs do not require intrinsic absorption in excess of the Galactic value. Therefore, for the statistically unacceptable fits, more complicated models might be needed to reproduce the spectra of these sources. On the contrary, the observed significant improvement in the quality of the NELG spectral fits obtained with model C, suggests that a large fraction of NELGs are intrinsically absorbed.

4.4. Absorbed power law and soft excess (model D)

The soft excess has been shown to be a common feature of the X-ray spectra of Seyfert 1 galaxies (Turner & Pounds 1989). The physical origin of this emission is still unclear, since its determination depends on a knowledge of the shape of the power law and the quantity of absorption. It has been interpreted as primary emission from the accretion disc, gravitational energy released by viscosity in the disc, or as secondary radiation from the reprocessing of hard X-rays in the surface layers of the disc.

We have searched for soft excess among the BLAGNs and NELGs. We first have fitted their spectra with a power law and a low energy black body component at the redshift of the source, absorbed by the Galaxy. The soft excess emission, parameterised as a black body, adds to the model two parameters, the black body temperature and normalisation.

We have compared the χ^2 of the fit with the values obtained with model A to search for the cases where the significance of a χ^2 improvement was above 95%. From the 170 sources analysed (141 BLAGNs and 29 NELGs) we found evidence

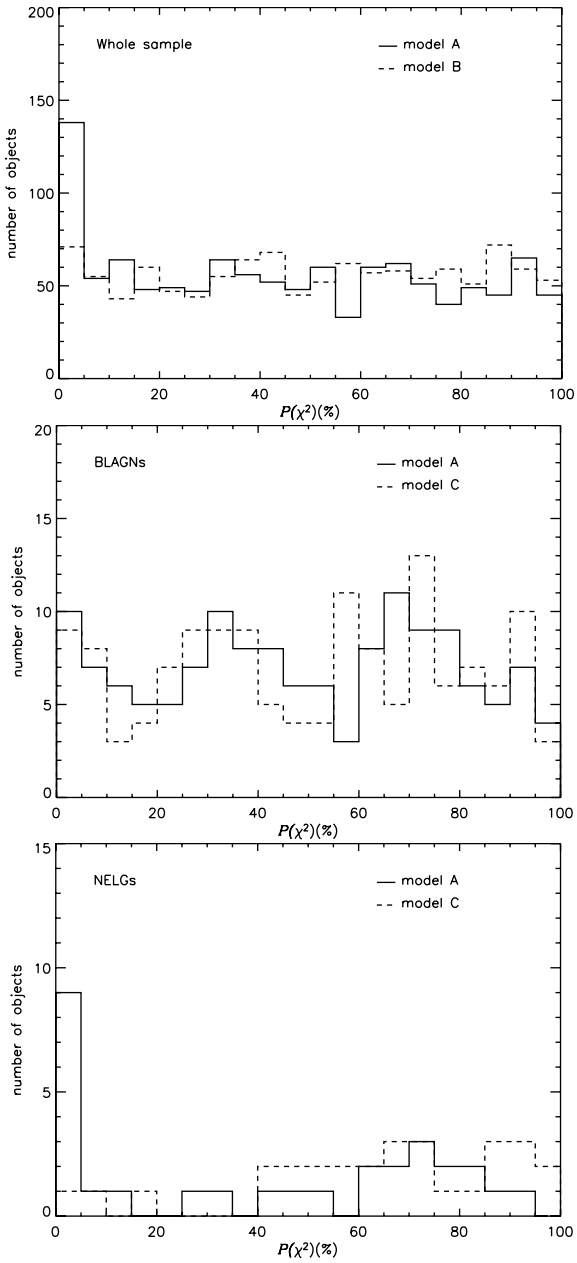

Fig. 5. Comparison of the null hypothesis probability distributions, $P(\chi^2)$, obtained from model A (simple power law, solid lines) and models B and C (observer frame absorbed power law and rest frame absorbed power law models, dashed lines). *From top to bottom*: whole sample of sources, BLAGNs and NELGs.

of soft excess emission in 12 sources (~7%), 10 BLAGNs and 2 NELGs.

As explained before, if soft excess is present in the spectra of the objects, it must be modelled properly to prevent the continuum spectral slope Γ and/or the intrinsic absorption $N_{\rm H}^{\rm intr}$ being incorrectly determined. Therefore we have analysed the X-ray spectra of these 12 sources in more detail. All the spectra have a large enough number of bins, 35 < bins < 317, to use a model with five free parameters: a power law and a low energy black body component, both absorbed by the Galaxy ($N_{\rm H}^{\rm Gal}$; fixed) and by absorption ($N_{\rm H}^{\rm intr}$; free) at the redshift of the

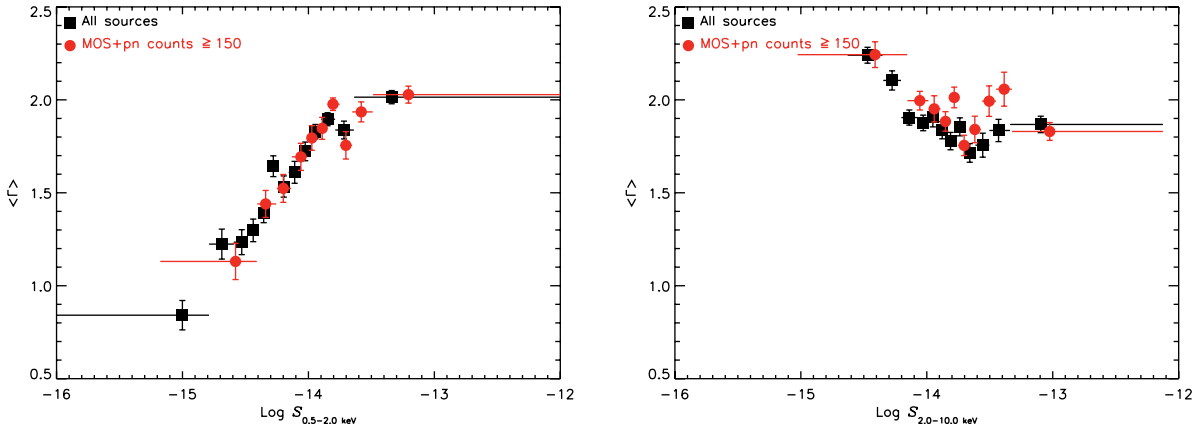

Fig. 6. Weighted photon index versus $S_{0.5-2\ \mathrm{keV}}$ and $S_{2-10\ \mathrm{keV}}$ from model A (simple power law). The results obtained for the objects with X-ray spectra with more than 150 counts (background subtracted) are included for comparison. Error bars correspond to 90% confidence.

sources (hereafter, model D). Results from this specific exercise are presented in Sect. 5.3.

5. Results of spectral fitting

5.1. The continuum shape

Although we know that a large fraction of sources in our list are absorbed, we have started the study with the results that we obtained with model A to allow comparison with spectral analyses of data with low signal to noise. This model, as was explained before, does not take into account the effects of absorption or soft excesses.

To calculate the objects average photon index, $\langle\Gamma\rangle$, we have used the standard formula for the weighted mean,

$$\langle\Gamma\rangle = \sum P_i \times \Gamma_i$$

where the weight, P_i, of each individual best fit value, Γ_i, is a function of the error obtained from the fit, σ_i,

$$P_i = \frac{1/\sigma_i^2}{\sum\left(1/\sigma_i^2\right)}.$$

To calculate the uncertainty in $\langle\Gamma\rangle$ we have used the standard deviation (Bevington 1992)

$$\sigma^2 = \frac{1}{(\mathrm{N}-1)} \sum P_i \times (\Gamma_i - \langle\Gamma\rangle)^2$$

that includes the measurement errors, σ_i, and the dispersion of each Γ_i from the estimated value $\langle\Gamma\rangle$.

Following these definitions, we find our sources to have a continuum emission, from 0.2 to 12 keV, with an average slope of $\langle\Gamma\rangle = 1.86 \pm 0.02$ (the arithmetic mean being 1.70 ± 0.02). To study the dependence of the objects average spectrum on the X-ray flux, we have calculated the values of $\langle\Gamma\rangle$ in bins of 0.5–2 and 2–10 keV flux (each bin containing the same number of objects). Figure 6 shows the results that we have obtained. We see that the average spectrum of the sources becomes harder with decreasing 0.5–2 keV flux, with $\langle\Gamma\rangle$ going from 2.01 ± 0.03 at $\sim 4.4 \times 10^{-14}$ erg cm^{-2} s^{-1} to 0.95 ± 0.12 at $\sim 10^{-15}$ erg cm^{-2} s^{-1}. This trend[6] (harder sources at fainter 0.5–2 keV fluxes) has been found in previous X-ray spectral analyses (e.g., Vikhlinin et al. 1999; Mittaz et al. 1999; Tozzi et al. 2001; Mainieri et al. 2002). Its origin is still not clear, although it could be explained as a rapid increase in the intrinsic absorption of the objects with decreasing flux, as it is predicted by some AGN synthesis models (see Gilli et al. 2001).

It is not possible to see the same effect using the 2–10 keV fluxes, where Γ seems to become softer at fainter fluxes (see Fig. 6). It is important to note that this softening of Γ is not a real property of our objects. It comes from using different energy bands simultaneously in the source detection. As we go to fainter 2–10 keV fluxes, it is more difficult to detect sources with flat spectral slopes (hard objects) because their spectra peak at energies above $\sim$2 keV where we know the sensitivity of the X-ray detectors decreases rapidly. Hence, the number of objects only detected in the soft band becomes more important at fainter 2–10 keV fluxes, specially below $\sim 2 \times 10^{-14}$ erg cm^{-2} s^{-1}. Note that the last bin at the faintest fluxes (see Fig. 6) is located at a value of $\langle\Gamma\rangle$ of more than 2.5. The objects that contribute to this bin have spectra with very few counts and therefore the values of Γ are poorly constrained.

We have studied in more detail whether the observed hardening of Γ with decreasing soft flux could be affected by any bias introduced during our analysis. We first have checked that the same result is obtained when only sources with more than 150 counts (circles in Fig. 6) are used. Therefore, the sources with poor spectral quality are not introducing any bias in our result.

We expect the majority of the observations that we have used, not to be deep enough as to detect sources with a 0.5–2 keV flux of $\sim 10^{-15}$ erg cm^{-2} s^{-1} in the soft band. Hence, it could be possible that our $\langle\Gamma\rangle$ becomes harder with decreasing flux partly due to the contribution to each bin from bright hard sources not detected in the soft band. This can be checked by repeating the plot of Fig. 6 with those sources selected in the soft or hard bands (detection likelihoods above 10 on each

[6] The same dependence, i.e., hardening of $\langle\Gamma\rangle$ at fainter 0.5–2 keV fluxes, is observed if $\langle\Gamma\rangle$ is calculated with the arithmetic mean, therefore the result is not an artifact of the weighted mean.

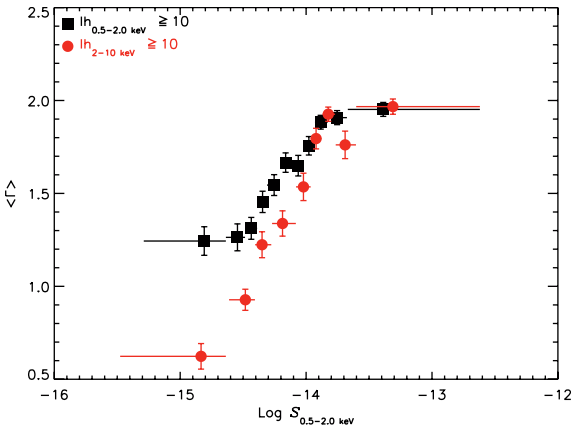

Fig. 7. Weighted photon index versus $S_{0.5-2\,\mathrm{keV}}$ from model A (simple power law) for sources with detection likelihood ≥10 in the 0.5–2 keV band (squares) and for sources with detection likelihood ≥10 in the 2–10 keV band (circles). Error bars correspond to 90% confidence.

band). The results, plotted in Fig. 7, show a very similar dependence of Γ with flux. Only the ⟨Γ⟩ of the bin at the faintest soft fluxes in Fig. 6, could be affected or even dominated by the contribution of bright hard sources not detected in the soft band. It is interesting to note that the hardening of Γ is more important for objects detected in the 2–10 keV band (see Fig. 7). If absorption is producing the hardening of Γ, this result suggests that we are seeing the same population of objects in the soft and hard bands, the only difference being that the objects detected in the 2–10 keV band are more absorbed on average.

Finally we have used the source detection efficiency function, W(Γ, S) (see Appendix A), to correct our results for all the biases of our study. The de-biased average values of Γ at each bin of flux were obtained weighting each individual value, Γ_i with the function:

$$P_i = \frac{1\,/\left(\sigma_i^2 \times W\left(\Gamma_i, S_i\right)\right)}{\sum\, 1\,/\left(\sigma_i^2 \times W\left(\Gamma_i, S_i\right)\right)}$$

where the sum is calculated with all the sources falling in the specified bin in the Γ, S plane (for a detailed explanation see Appendix A).

The results are shown in Fig. 8, the left plot for the whole sample of sources, and the right plot for the sources detected in the soft band. For comparison we have calculated the ⟨Γ⟩ values weighted with the fitting errors in the same flux bins (stars in the plots). We still see the hardening of ⟨Γ⟩ with the 0.5–2 keV flux, so this dependence is indeed a real property of the population of the sources analysed.

5.2. Excess absorption

The fit to the X-ray spectra with a single power law (model A) yields a large number of sources with statistically unacceptable fits. Models B and C reduce significantly the number of unacceptable fits as it is shown in Fig. 5, suggesting that a substantial fraction of our sources could be X-ray absorbed, in agreement with the predictions of the standard synthesis models of the XRB (Comastri et al. 1995; Gilli et al. 2001).

Table 3. Fractions of F-test absorbed objects and significances.

Type	$N_{\rm tot}$	$N_{\rm abs}$	f_-^{+a}	F^b
All	1137	245	$0.17\,^{+0.02}_{-0.02}$	0.1335
BLAGNs	141	16	$0.06\,^{+0.05}_{-0.04}$	0.0089
NELGs	29	13	$0.40\,^{+0.15}_{-0.14}$	0.1740
ALG	7	3	$0.40\,^{+0.20}_{-0.20}$	0.0400

[a] The f values are the mode of the function $P(f)$ (see Sect. 5.2 for details). The errors are 90%.
[b] F is where $P(f \geq F) = 0.9973$ (3σ).

To search for absorbed objects we used the F-test, measuring the improvement in the χ^2 of the fits after adding absorption to the model. A 95% confidence cutoff is adopted for absorption to be detected. In Table 3, we show the total number of objects of different types, the number of absorbed objects of that type, and the fraction of absorbed objects of each type with its uncertainty and significance.

For a sample of N objects, of which n are absorbed, the Bayesian posterior probability of the fraction of absorbed objects is $P(f) \propto (f + f_{\rm spu})^n (1 - f - f_{\rm spu})^{N-n}$ (from the binomial distribution), where $f_{\rm spu}$ is the fraction of spurious detections (for our sources 5%). We used the mode of $P(f)$ to estimate the fraction (corrected from spurious detections) of absorbed objects within each class of sources. The errors in the fractions were calculated integrating $P(f)$ from the mode value in the two directions until we obtained 45% of the probability (90% errors). For every "class" of sources, we measured the confidence of detection of absorption integrating the $P(f)$ distributions from $f = 1$ down to the value where the integral is equal to 0.9973 (or 3σ). These values, listed in the last column of Table 3, indicate the fraction of absorbed objects that we have detected within each class of sources with a significance of more than 3σ. For example, we know with a confidence of more than 3σ that at least ~1% of BLAGNs and ~17% of NELGs are absorbed.

It appears that the fraction of absorbed objects varies substantially among different classes of object, with the NELGs having the highest fraction. We have used the method described in Stevens et al. (2005) to study whether our samples of BLAGNs and NELGs have different fractions of absorbed objects. Assuming that BLAGNs and NELGs come from the same parent population, the fraction of absorbed objects is

$$P(f) \propto \left(f + f_{\rm spu}\right)^{(n+m)} \left(1 - f - f_{\rm spu}\right)^{(N+M-n-m)}$$

where M and N are the number of sources in the two samples and m and n the number of detections on each sample.

The probability of detection of i objects in the first sample and j in the second one is given by

$$P(i, j; N, M) = \binom{M}{j}\binom{N}{i} \int_0^1 \mathrm{d}f\, P(f) \left(f + f_{\rm spu}\right)^{(i+j)} \times \left(1 - f - f_{\rm spu}\right)^{(N+M-i-j)} .$$

The fractions of absorbed objects in BLAGNs and NELGs will be different if the probability of detecting $\leq m$ in the first sample and $\geq n$ in the second one is low, i.e. if

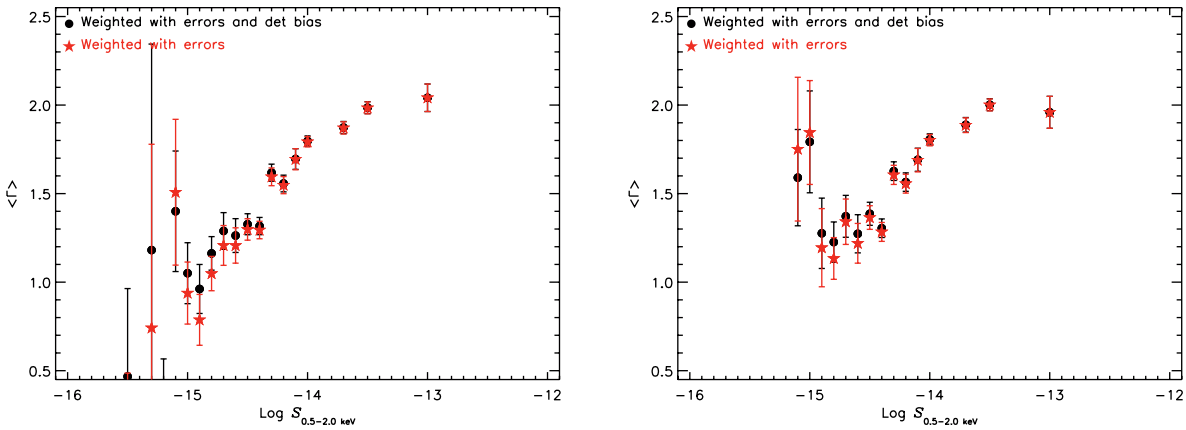

Fig. 8. Comparison of the weighted photon index versus $S_{0.5-2\,\mathrm{keV}}$ from model A (simple power law) with the values obtained after correcting for the source list biases. Left for the whole sample of objects and right for the sub-sample of sources with detection likelihood ≥10 in the 0.5–2 keV band. Error bars correspond to 90% confidence.

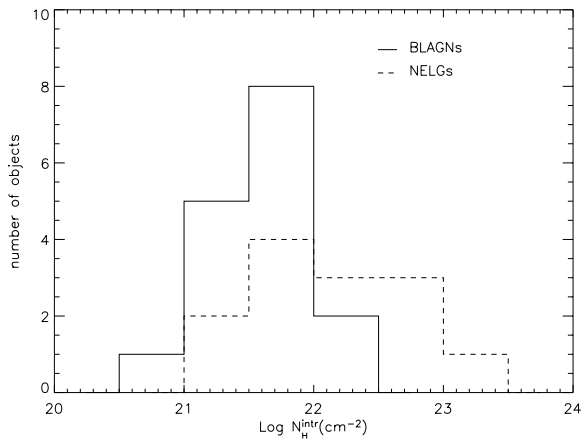

Fig. 9. Distributions of intrinsic absorption for BLAGNs (solid histogram) and NELGs (dashed histogram).

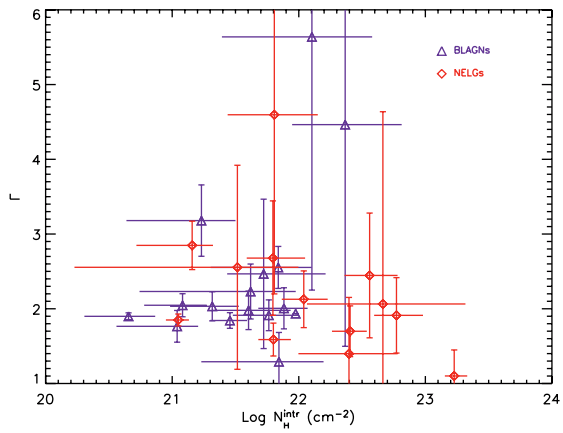

Fig. 10. Photon index versus intrinsic absorption for absorbed (F-test significance ≥ 95%) BLAGNs (triangles) and NELGs (diamonds).

$P(\geq n, \leq m; N, M) = \sum_{j=0}^{M} \sum_{i=n}^{N} P(i, j; N, M)$ is low. We find that $P(m \leq 16, n \geq 13; 141, 29) = 0.000013$. Therefore we see that the fraction of absorbed objects in BLAGNs and NELGs is different with a significance of more than 3σ. We cannot distinguish the fraction of absorbed objects among NELGs and ALGs (the probability of them being different is not significant).

Figure 9 shows the intrinsic $N_{\mathrm{H}}^{\mathrm{intr}}$ distributions measured for the absorbed BLAGNs and NELGs. BLAGNs show column densities $4.5 \times 10^{20} < N_{\mathrm{H}}^{\mathrm{intr}} < 2.3 \times 10^{22}$ cm^{-2} while NELGs seem to cover a much wider range of absorption, with $1.1 \times 10^{21} < N_{\mathrm{H}}^{\mathrm{intr}} < 1.7 \times 10^{23}$ cm^{-2}. However, the small number of objects does not allow us to find a significantly different (in terms of the Kolmogorov-Smirnov test) distribution of absorbing columns between BLAGNs and NELGs.

To be confident that the detected intrinsic absorptions among identified sources were not affected by the quality of their spectra, we have plotted Γ versus $N_{\mathrm{H}}^{\mathrm{intr}}$ to search for any correlation between the two spectral parameters. Such a correlation would be expected if the detections of absorption were spurious and linked to very large values of Γ. The results are shown in Fig. 10. The plot shows no apparent correlation between Γ and $N_{\mathrm{H}}^{\mathrm{intr}}$.

A Kolmogorov-Smirnov comparison between the redshift distributions of unabsorbed and absorbed BLAGNs does not indicate a significant difference between the two. The same result is obtained when comparing the 2–10 keV luminosity distributions of absorbed and unabsorbed BLAGNs and NELGs (we use the 2–10 keV energy band because it is less affected by absorption). We conclude that X-ray absorption does not occur preferentially at any particular redshift (in BLAGNs) or X-ray luminosity.

Finally, we have tested whether the lack of optical broad emission lines in unabsorbed NELGs can be explained in terms of AGN/host galaxy contrast (Lumsden et al. 2001; Moran et al. 2002; Page et al. 2003; Severgnini et al. 2003). To this end, we have compared the 2–10 keV luminosity distributions of two sub-samples of unabsorbed BLAGNs and NELGs detected with redshifts ranging from 0.17 to 0.7, where the redshift distributions of BLAGNs and NELGs overlap (see Fig. 2). The Kolmogorov-Smirnov significance of the two distributions being different is ~99%. Furthermore, the average luminosities of the selected BLAGNs and NELGs were found to be

Table 4. Spectral parameters of sources with a soft excess component detected.

Name [a]	Class	redshift	Γ^{+}_{-}	$N_{\rm H}^{\rm intr}$ [b] (10^{20} cm^{-2})	kT^{+}_{-} (keV)	$S_{0.5-2}$ (%)	$L_{0.5-2}$ (%)	model C [c]	model D [d]
XMMU J001831.7 + 162924	BLAGN	0.551	$2.35^{+0.17}_{-0.18}$	$4.15^{+11.25}_{-2.37}$	$0.186^{+0.259}_{-0.126}$	15.95	23.39	Not abs.	abs.
XMMU J021808.2 − 045844	BLAGN	0.712	$2.08^{+0.04}_{-0.05}$	$<10^{20}$	$0.142^{+0.164}_{-0.111}$	6.55	19.49	Not abs.	Not abs.
XMMU J021830.5 − 045622	BLAGN	1.430	$2.28^{+0.10}_{-0.18}$	$<10^{20}$	$0.129^{+0.223}_{-0.082}$	1.89	30.08	Not abs.	Not abs.
XMMU J061728.6 + 710600	BLAGN	0.219	$1.68^{+0.56}_{-0.37}$	$<10^{20}$	$0.214^{+0.253}_{-0.111}$	39.54	44.57	Not abs.	Not abs.
XMMU J061730.9 + 705955	NELG	0.200	$2.02^{+1.13}_{-1.00}$	$<10^{20}$	$0.207^{+0.272}_{-0.142}$	44.14	49.31	abs.	Not abs.
XMMU J074311.9 + 742935	BLAGN	0.332	$2.05^{+0.03}_{-0.05}$	$<10^{20}$	$0.119^{+0.130}_{-0.106}$	12.22	22.12	Not abs.	Not abs.
XMMU J12226.1 + 752616	NELG	0.238	$1.82^{+0.14}_{-0.10}$	$11.17^{+2.43}_{-2.30}$	$0.146^{+0.208}_{-0.112}$	12.30	22.72	abs.	abs.
XMMU J13316.2 + 241326	BLAGN	0.174	$1.83^{+0.45}_{-0.26}$	$<10^{20}$	$0.085^{+0.116}_{-0.064}$	14.71	25.20	Not abs.	Not abs.
XMMU J212904.5 − 154448	BLAGN	0.431	$1.29^{+0.27}_{-0.36}$	$14.21^{+32.59}_{-5.81}$	$0.205^{+0.251}_{-0.142}$	54.55	72.74	Not abs.	abs.
XMMU J221515.1 − 173224	BLAGN	1.165	$2.06^{+0.37}_{-0.35}$	$<10^{20}$	$0.212^{+0.233}_{-0.192}$	24.75	49.69	Not abs.	Not abs.
XMMU J221523.6 − 174321	BLAGN	0.956	$2.15^{+0.17}_{-0.23}$	$<10^{20}$	$0.112^{+0.234}_{-0.021}$	1.86	19.16	Not abs.	Not abs.
XMMU J222823.6 − 051306	BLAGN	0.758	$1.52^{+0.14}_{-0.15}$	$15.61^{+12.39}_{-7.20}$	$0.262^{+0.379}_{-0.171}$	35.78	54.55	Not abs.	abs.

[a] *XMM-Newton* source name.
[b] Intrinsic absorption measured with model D.
[c] Detection of intrinsic absorption with model C (F-test comparison of the χ^2 from model A and model C).
[d] Detection of intrinsic absorption with model D (F-test comparison of the χ^2 from model A and model D).

43.7 ± 0.16 and 43.2 ± 0.11 (in log units). Therefore, we have seen that our unabsorbed NELGs are on average significantly less luminous than unabsorbed BLAGNs at similar redshifts, and hence, our data are consistent with the explanation of non detection of broad optical lines in unabsorbed NELGs as an AGN/ host-galaxy contrast effect.

5.3. Soft excess

We have previously shown that ∼7% of the sources identified as BLAGNs or NELGs exhibit significant soft excess emission. To measure the broad band spectral parameters of these sources, and in particular the underlying power law index Γ we have used model D, as described in detail in Sect. 4.

The results of the fits are listed in Table 4, where for each source we have included its IAU-style *XMM-Newton* source name, the optical classification, redshift, and the spectral parameters measured with model D. We have also included the soft excess contributions to the flux and luminosity in the soft band and the F-test results obtained from the comparison of χ^2 from model A to χ^2 from model C and model D. We now discuss some particular sources in detail:

XMMU J001831.7+162924, XMMU J222823.6-051306, XMMU J212904.5-154448: we obtained an improvement in the spectral fits with model C, but the F-test significance was in all cases below 95%, therefore the sources were classified as unabsorbed. A proper modelling of the soft excess emission with model D allowed us to find significant absorption above the Galactic value in all the sources.

XMMU J221515.1-173224: we obtained an unusually steep spectral slope with model A, $\Gamma = 2.67^{+0.05}_{-0.06}$. The quality of the fit did not improve significantly with model C. With model D we obtained a significant improvement in the fit (>99.99%). This model did not require absorption in excess of the Galactic value. The measured temperature of the soft excess emission was $kT = 0.05^{+0.03}_{-0.01}$, below the typical values found for the rest of the sources. The fitted continuum was still significantly steep, with $\Gamma = 2.60^{+0.04}_{-0.11}$.We tried to find a more typical value of Γ fitting the continuum slope at high energies (>2 keV) with model A. We obtained a value of Γ significantly lower, $2.06^{+0.37}_{-0.35}$ and in agreement with the average spectral slopes observed in unabsorbed AGN. With Γ fixed to the above value the fit did not improve with model C, but with model D we still obtained an F-test significance of >99.99%. The model did not require absorption in excess of the Galactic value, therefore we believe the improvement is due entirely to the soft excess component. We adopted this model as the best fit model for the source although the χ^2 of the fit was slightly worse that the value obtained with Γ fully variable.

XMMU J061730.9+705955: we found a significant column density with model C, hence the source was classified as absorbed. However we obtained a very steep spectral slope, $\Gamma = 2.87^{+0.35}_{-0.31}$. With model D we obtained a better fit of the data and more typical values for the spectral parameters, although model D did not require intrinsic absorption. We have chosen model D in preference to model C for this source.

It is interesting to note that for some of the sources that we have analysed, we found very flat spectral slopes with model A, but they all required significant quantities of soft excess emission. This is the case for **XMMU J12226.1+752616** and **XMMU J222823.6-051306**, with measured model A spectral slopes of $\Gamma = 1.52^{+0.04}_{-0.04}$ and $\Gamma = 1.78^{+0.11}_{-0.11}$ respectively.

5.4. Results from best fit model

To test whether the spectral hardening observed using model A is due to an increase in intrinsic absorption at the faintest fluxes, we compared Γ (from the best fit model in terms of the F-test results) to the soft X-ray flux for each source. The best fit model can be model A or model B for the unidentified sources, and

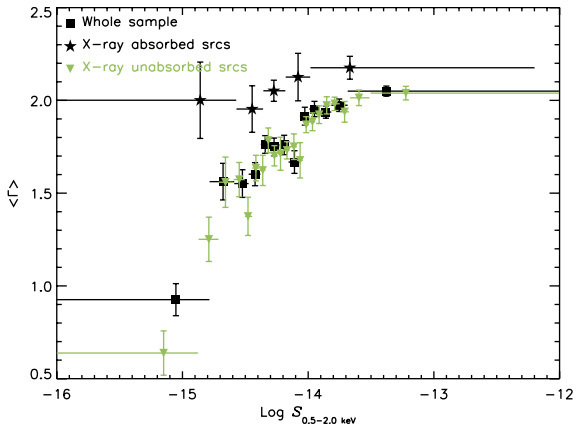

Fig. 11. $\langle\Gamma\rangle$ versus $S_{0.5-2\,\mathrm{keV}}$ obtained from the best fit model of each object. Squares denote the whole sample; stars absorbed sources (F-test ≥ 95.0%) and triangles unabsorbed sources (F-test < 95.0%).

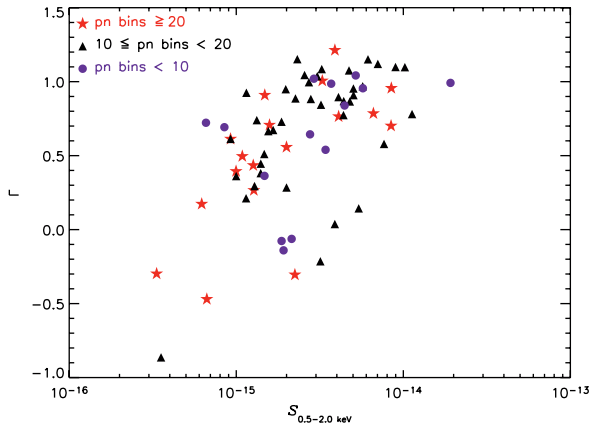

Fig. 12. Γ versus $S_{0.5-2\,\mathrm{keV}}$ for objects with flat spectrum (Γ positive error bar <1.5). We have grouped the objects in terms of the number of bins in their pn spectrum.

model A or model C for the identified sources (or model D for BLAGNs and NELGs when soft excess emission was detected). Using the best fit model for each object we obtained a value of $\langle\Gamma\rangle = 1.96 \pm 0.01$ (the unweighted value was 2.08 ± 0.04). The dependence of $\langle\Gamma\rangle$ with the soft flux is plotted in Fig. 11, where we see that:

1. The average continuum of absorbed sources is significantly softer than the continuum of unabsorbed sources, especially at the faintest fluxes, with $\langle\Gamma\rangle = 2.15 \pm 0.04$ (the unweighted value being 3.12 ± 0.12[7]) for the absorbed sources and $\langle\Gamma\rangle = 1.94 \pm 0.01$ (the unweighted value being 1.79 ± 0.02) for the unabsorbed sources.
2. There is no hardening of the average spectral slope for the absorbed sources. The hardening of Γ seen in these objects when fitted with the SPL model is likely to be an effect of absorption only.
3. For the sources where absorption was not detected in terms of the F-test, we still observe a significant hardening of the average spectral slope with the X-ray flux.

Since the sources where absorption is detected do not exhibit any significant variation of the underlying spectral index with flux, does this hold true for the remaining sources? We first wanted to check whether any effect could lead to an apparent flattening of the spectral shape towards fainter fluxes when absorption is included in the fitting model. To this goal, we have generated a number of fake spectra with an underlying power law $\Gamma = 2$ at various fluxes spanning those of our sample, and different amounts of absorption at $z = 0$.

We have then first fitted these spectra with models A and B, and then we have compared the χ^2 of the fits with the F-test. We found a range of low values of $N_{\mathrm{H}}^{\mathrm{obs}}$ where absorption was not significantly detected when model A was replaced by model B (F-test < 95%). In these cases, the best fit value of Γ (i.e., the value obtained with model A) was indeed lower than the input one, and it can reach values in the range $\Gamma \sim 1$–2.

[7] The unweighted mean is very high and different from the value that we obtained for the weighted mean, because there are a number of absorbed objects for which $\Gamma > 2.5$ but with large errors bars.

A second result of this exercise is, if we hold the absorption in the simulated spectra but vary the flux, at fainter fluxes the typical values of Γ fitted are marginally harder than at higher fluxes (and then fitted values of N_{H} tend to be lower than the input ones). We believe this effect to be due to a degradation of the spectral resolution of the X-ray spectrum, produced by the necessary grouping in very wide bands to achieve the desired statistics (10 counts/bin). This leads, in particular, to a smoothing of the model count rate function (which is the effective area times the input spectrum, convolved with the redistribution matrix) in the whole spectral range, but in particular in the vicinity of its peak at soft energies where most of the counts are detected. This produces an undesired slight hardening of the spectrum.

However, this effect is unlikely to produce the hardening of the unabsorbed sources that we see. The reason is that in this case we should see the lowest quality spectra (i.e., those with the smaller number of bins) being more prone to this effect, but this is not the case. Figure 12 shows the sources that we detect in the real sample with flat spectra (best fit Γ error bar upper limit below 1.5), grouped in terms of the number of bins in their EPIC-pn spectra. We clearly see that the hardening is not correlated with the quality of the spectra, and therefore the above effect does not dominate.

We conclude that the hardening towards fainter fluxes is real. This still leaves two options for its origin: intrinsic flattening of the power law or undetected low level absorption. Although the current data do not allow us to distinguish between them, there is nothing inconsistent with the simplest hypothesis which is that these sources are absorbed at a level which is undetectable in the current data.This is also supported by the results of Fig. 6 (right plot) where we do not see hardening of Γ with the 2–10 keV flux down to $\sim 2 \times 10^{-14}$ erg cm^{-2} s^{-1}. If absorption is producing the hardening effect that we see in our objects, then we do not expect to see the same effect using 2–10 keV fluxes, because they are less affected by absorption. However, deeper observations are needed to assess this point for sources at fluxes similar to those of the sources in our sample. The results of a detailed X-ray spectral analysis of the sources that were detected with *XMM-Newton* in

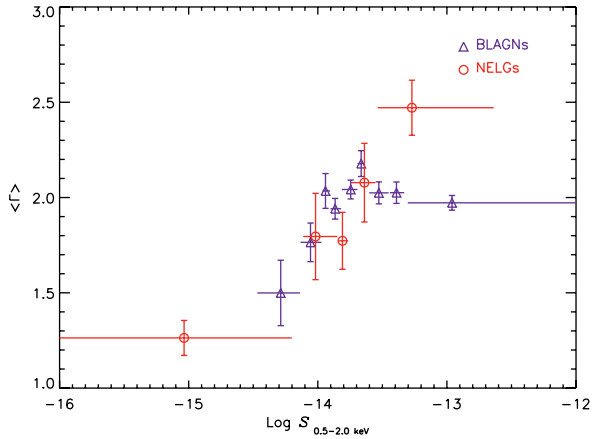

Fig. 13. $\langle\Gamma\rangle$ from best fit model versus $S_{0.5-2\,\mathrm{keV}}$ for BLAGNs (triangles) and NELGs (circles). The hardening of the spectral slope is evident even among the sub-samples of AGNs. Error bars correspond to 90% confidence level.

a deep observation in the Lockman Hole field will be published in Mateos et al. 2005 (in preparation).

The average spectral indices for the best fit model in the sub-samples of BLAGNs and NELGs are 1.99 ± 0.02 and 2.35 ± 0.09, both in agreement with the canonical value measured for nearby unabsorbed BLAGNs (Nandra & Pounds 1994). Both sub-samples of sources exhibit the same dependence with the soft X-ray flux as shown in Fig. 13. We still detect a clear correlation between the spectral slope and the soft X-ray flux.

5.5. Photon index intrinsic dispersion

In all the above analysis we have observed a clear dispersion in the measured spectral slope of our objects, even within the sub-samples of BLAGNs and NELGs, with the largest scatter in Γ being observed for the NELGs. To test whether the observed scattering is intrinsic to the sources we have followed the procedure described in Nandra & Pounds (1994) and Maccacaro et al. (1988), where the distribution of source spectral indices is assumed to be well reproduced by a Gaussian distribution of mean $\langle\Gamma\rangle$ and dispersion $\sigma_{\langle\Gamma\rangle}$. The best estimates of $\langle\Gamma\rangle$ and $\sigma_{\langle\Gamma\rangle}$ are obtained with the maximum likelihood (ML) technique. We have calculated the distribution of the spectral slopes that we obtained from the best fit model for the whole sample and the sub-samples of BLAGNs and NELGs. The results are shown in Table 5. In all the classes we measured a significant dispersion in spectral slopes. The large error in Γ obtained for NELGs did not allow us to differentiate the average spectral index of NELGs from that of BLAGNs. We see that the mean values of Γ computed with the ML and the weighted mean are very different, in particular for the NELGs. The reason is that with the ML method objects with untypical Γ (i.e. outliers) and not very large errors will increase the dispersion of the fitted gaussian distribution rather than significantly affecting the mean.

Our results do not show any trend in the underlying power law of the NELGs to be harder than that of the BLAGNs even at the faintest fluxes. Previous results based on *ROSAT* data (see Almaini et al. 1996; Romero-Colmenero et al. 1996) do find a flatter average spectral slope for NELGs (∼1.5) than for BLAGN (∼2). Our data shows that the spectral differences between these two classes of sources are due mostly (if not totally) to differing amounts of absorption rather than to differences in the underlying power law spectra. Therefore many objects classified as NELGs may contain active nuclei.

Figure 14 shows the $\langle\Gamma\rangle$ versus $\sigma_{\langle\Gamma\rangle}$ contour diagrams for the sub-samples of BLAGNs (solid lines) and NELGs (dashed lines), clearly showing the existence of an intrinsic dispersion in Γ with a confidence of more than 3σ.

Table 5. Comparison of the weighted mean spectral photon index obtained for different types of sources with the values from the Maximum Likelihood analysis. The best fit spectral slopes for each object were used in the calculations.

	Maximum likelihood		Weighted mean
Sample	$\langle\Gamma\rangle$	σ	$\langle\Gamma\rangle$
Whole sample	$1.86^{+0.02}_{-0.03}$	$0.36^{+0.01}_{-0.04}$	1.96 ± 0.01
BLAGN	$1.98^{+0.04}_{-0.04}$	$0.21^{+0.05}_{-0.04}$	1.99 ± 0.02
NELG	$1.85^{+0.16}_{-0.17}$	$0.36^{+0.14}_{-0.10}$	2.35 ± 0.09

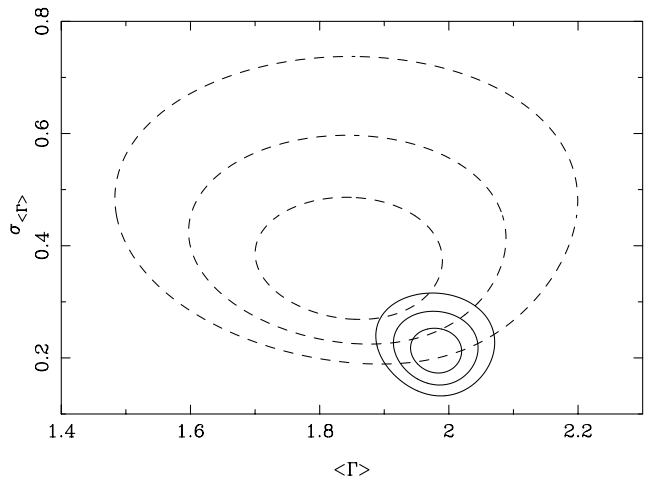

Fig. 14. Contour diagrams for the average spectral slope and intrinsic dispersion of our objects obtained from the Maximum Likelihood analysis (see Sect. 5.5). The spectral slopes obtained from the best fit model were used in the calculations. Solid line for BLAGNs and dashed line for NELGs. The contours are defined as $\Delta\chi^2$ = 2.3, 6.17 and 11.8 corresponding to standard 1, 2 and 3σ confidence regions for two parameters.

5.6. Spectral cosmic evolution

No correlation was found between Γ and the rest-frame 0.5–10 keV luminosities, as shown in Fig. 15. For the sub-sample of BLAGNs we studied the measured Γ that we obtained from each best fit model as a function of the redshift, but we did not find any clear tendency between the two parameters as can be seen in Fig. 16.

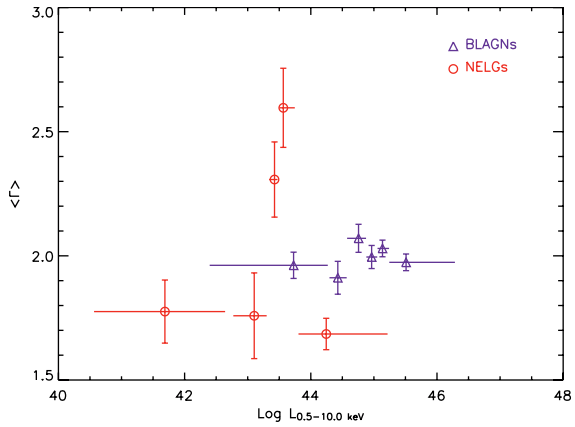

Fig. 15. Weighted $\langle\Gamma\rangle$ versus $L_{0.5-10\,\mathrm{keV}}$ for BLAGNs (triangles) and NELGs (circles). Error bars correspond to 90% confidence level.

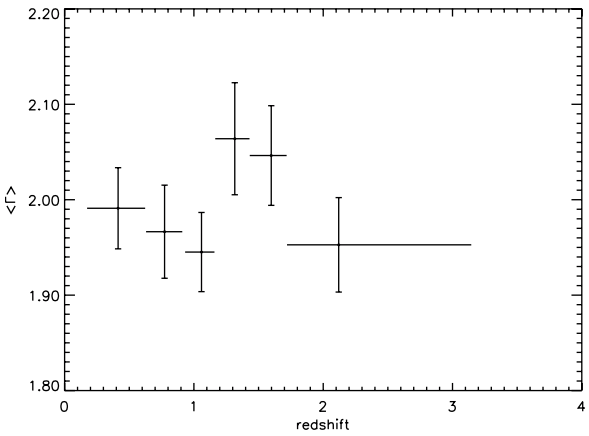

Fig. 16. $\langle\Gamma\rangle$ versus redshift for BLAGNs. We have used the values of Γ from the best fit model for each object (i.e. with absorption and soft excess if required). Error bars correspond to 90% confidence level.

We have studied the redshift distribution of the measured rest frame absorbing column densities for the absorbed sub-samples of BLAGNs and NELGs. The results, shown in Fig. 17, show no evidence of evolution of $N_{\mathrm{H}}^{\mathrm{intr}}$ with redshift. However, the ranges of the 2D parameter space that we fully sample with our source list are unclear.

We expect the minimum detected absorption to increase with z. This can be easily explained in terms of absorption being redshifted towards lower energies and further decreased by a factor $(1+z)^{2.7}$ as we move upwards in z.

To study if there exists a maximum detected column (only the brightest sources being actually detectable when heavy absorption is present), and if this limit depends on redshift, we have carried out simulations as follows. A 3D parameter space was defined with the axes being redshift, luminosity and intrinsic (rest-frame) absorption. For each parameter we have defined a grid of values covering the observational limits established by our sources. First, for each point in the $N_{\mathrm{H}}^{\mathrm{intr}}$-$z$ plane we simulate a spectrum for each BLAGN or NELG source in our sample, using its spectral response matrices and exposure time and assuming $\Gamma = 2$. As explained in Sect. 3 we have only analysed those spectra with more than 50 counts (background subtracted). We use this limit to calculate the required

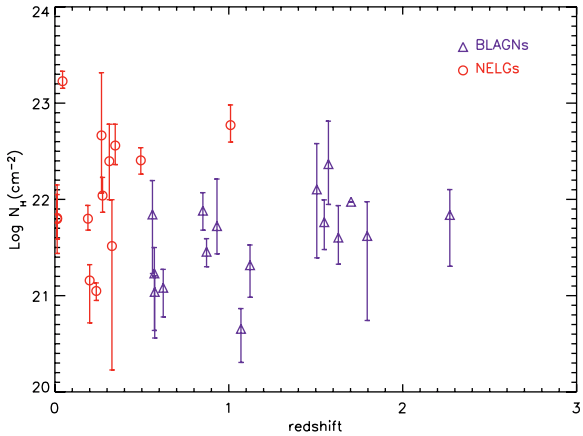

Fig. 17. Intrinsic absorption as a function of redshift for absorbed (F-test significance ≥95%) BLAGNs (triangles) and NELGs (circles).

minimum luminosity for each source to fulfil this constraint. Luminosities lower than this one would produce a spectrum outside our quality bounds.

Second, for a number of luminosities above this limit (still, for each value of intrinsic absorbing column and z) we simulate a spectrum for every source. These spectra were analysed following the same procedure that we used for the real spectra giving a full coverage of the above 3D parameter space.

We then examined the fitted values of $N_{\mathrm{H}}^{\mathrm{intr}}$ as a function of redshift z for a range of luminosities. The simulations confirm that the minimum detection of absorption depends strongly on z for the reason explained above. We also found two limits to the maximum absorption detected as a function z. The first one is the source detection itself: heavy absorption reduces significantly the flux of the source and makes detection very difficult, especially at high redshifts. The second limit is the detection of absorption: even if the source is detected, if its luminosity is not high enough we will measure an absorption below the maximum value established by the detection limit.

This strong dependence on luminosity and redshift makes it very difficult to derive the real absorbing column distribution in terms of the fitted values. The best way we found to tackle this problem is to compare our results with the predictions of specific models for the distribution of absorbing columns, folding the input values through the same effects as the real data. This is discussed in Sect. 7 in the framework of unified AGN models for the XRB.

6. ALGs X-ray spectra

The objects with galaxy-like optical spectra not showing emission line signatures were classified as absorption line galaxies (ALG).

The inferred high X-ray luminosities obtained from the analysis of their X-ray spectra strongly suggests that an AGN is the source of the X-ray emission in at least 5 out of the 7 objects. The peculiar properties of these objects have been explained by a very high column density that would be obscuring the AGN broad and narrow line regions. This hypothesis has found observational support, because many of the ALGs

studied have been found to be X-ray obscured (Mainieri et al. 2002; Comastri et al. 2002). However there are several examples of ALGs for which the obscuration model does not apply. In these cases other possible explanations have been suggested, e.g. that a BL Lac could be present in the core of some ALGs (Brusa et al. 2003). Alternatively some objects might host an AGN with a non-standard accretion disk (Yuan & Narayan 2004). A final possibility is that some ALGs may have normal (unabsorbed) AGNs that could be hidden if the host galaxy is bright enough (Severgnini et al. 2003; Page et al. 2003).

We now describe the X-ray spectral properties of the ALGs found in our list:

XMMU J122017.7+752216 is an optically normal galaxy that looks extended in the X-ray image. Its soft X-ray emission is best reproduced with a thermal model with temperature $kT = 0.64^{+0.04}_{-0.04}$ keV and metallicity $Z = Z_{\odot}$. The fit does not improve leaving Z as a free parameter. However, the high energy X-ray spectrum is best modelled with an unabsorbed power law of spectral slope $\Gamma = 1.87^{+0.12}_{-0.11}$. We cannot confirm whether the source of the hard X-ray emission is an AGN due to the low value of the hard X-ray luminosity inferred from this model ($L_{2-10} = 3.1 \times 10^{39}$ erg s^{-1}).

The nature of **XMMU J021822.1-050612** was studied in detail by Severgnini et al. (2003). They found the X-ray spectrum to be best modelled with an absorbed power law.

The X-ray spectra of **XMMU J010320.2-064157** and **XMMU J084211.6+710145** were best reproduced with an obscured (F-test values 99.7% and 95.0%) AGN with $\Gamma = 2.39^{+1.36}_{-0.96}$, $N_{\rm H} = 4.9^{+5.28}_{-2.53} \times 10^{22}$ cm^{-2} and $L_{2-10} = 2 \times 10^{44}$ erg s^{-1} for the first object, and $\Gamma = 1.82^{+0.16}_{-0.15}$, $N_{\rm H} = 9.3^{+8.5}_{-6.8} \times 10^{20}$ cm^{-2} and $L_{2-10} = 3.88 \times 10^{43}$ erg s^{-1} for the second object.

XMMU J214041.3-234719 has an unabsorbed X-ray spectrum with $\Gamma = 2.08^{+0.09}_{-0.08}$ and $L_{2-10} = 6.2 \times 10^{40}$ erg s^{-1}. Hence, from its X-ray spectrum alone, we would expect this source to be a BLAGN optically, rather that an ALG, according to the simplest AGN unified model.

A simple power law model can reproduce the spectra of **XMMU J233113.6+20056** (see Fig. 18) and **XMMU J122312.6+75247**, but the best fit spectral slopes were found to be well below the typical value found for unabsorbed AGNs ($\Gamma = 0.94^{+0.26}_{-0.27}$ and $\Gamma = 1.36^{+0.30}_{-0.28}$). In terms of the F-test none of the spectra required absorption in excess of the Galactic value, although with the absorbed model we obtained more typical spectral slopes. The low quality of the spectra has not allowed us to study in more detail whether the spectra of these sources is intrinsically flat or absorbed.

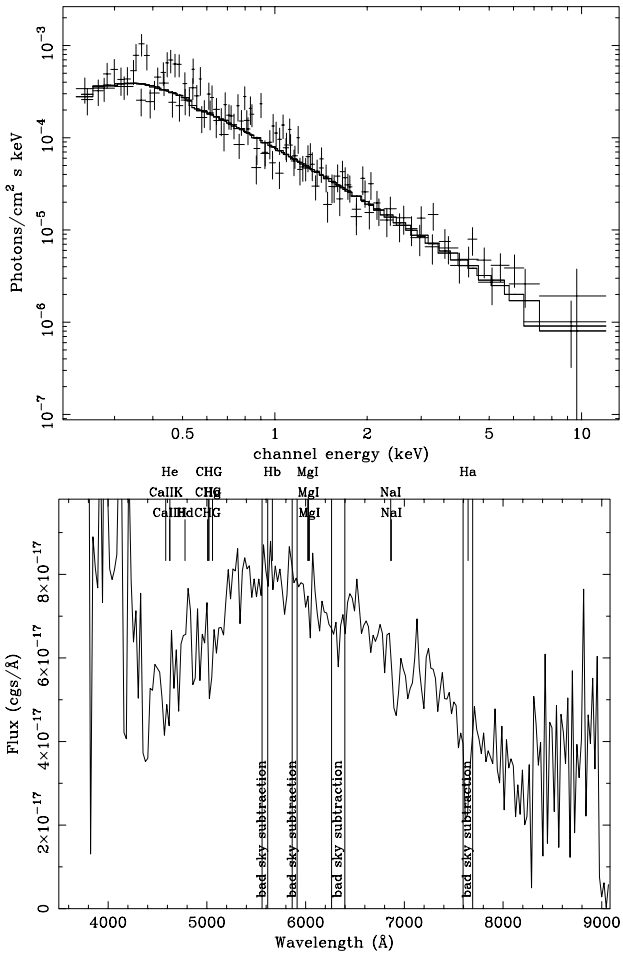

Fig. 18. *Top panel*: MOS+pn spectrum of the absorption line galaxy **XMMU J233113.6+20056** at $z = 0.378$. The X-ray spectrum was best fitted with a flat ($\Gamma \sim 0.944$) unabsorbed power law. *Bottom panel*: optical spectrum of the source, where no obvious broad lines were detected.

7. Testing population synthesis models

The standard population synthesis models of the XRB, based on unification schemes for AGN, can reproduce successfully the XRB with an appropriate mixture of unobscured and obscured AGN, as well as a number of observational constraints. The input parameters of these models are the type 1 and type 2 measured X-ray spectra and cosmological evolution, which is assumed to be identical. The X-ray luminosity function of absorbed objects is still unknown, but it is assumed to be the same as for unabsorbed objects. The ratio of absorbed to unabsorbed AGN comes out to be rather large in these models. A key descriptor of these models is the distribution of absorbing column densities among the sources. The $N_{\rm H}^{\rm intr}$ distribution is now observationally known for the local Seyfert galaxies (e.g., Risaliti et al. 1999), and in general these models assume it to be independent of redshift and luminosity.

The analysis that we have conducted has allowed us to measure the distribution of absorption in one of the largest samples of BLAGNs and NELGs studied up to now, and its cosmic evolution over a wide range of redshifts (see Figs. 9 and 17). In Sect. 5.6 we reported that the measured absorption is slightly lower than the real one, depending on the luminosity and redshift of the source. Hence, in order to compare the predictions of XRB models to our data, we assume an input distribution for $N_{\rm H}^{\rm intr}$ (independent of luminosity and redshift), and then fold it through the same detection effects as in our data.

We have used the $N_{\rm H}^{\rm intr}$ distributions adopted by Comastri et al. (1995), Gilli et al. (2001) and Pompilio et al. (2000), as input parameters to their analytical models of the XRB. We have simulated spectra for each one of the BLAGNs and NELGs

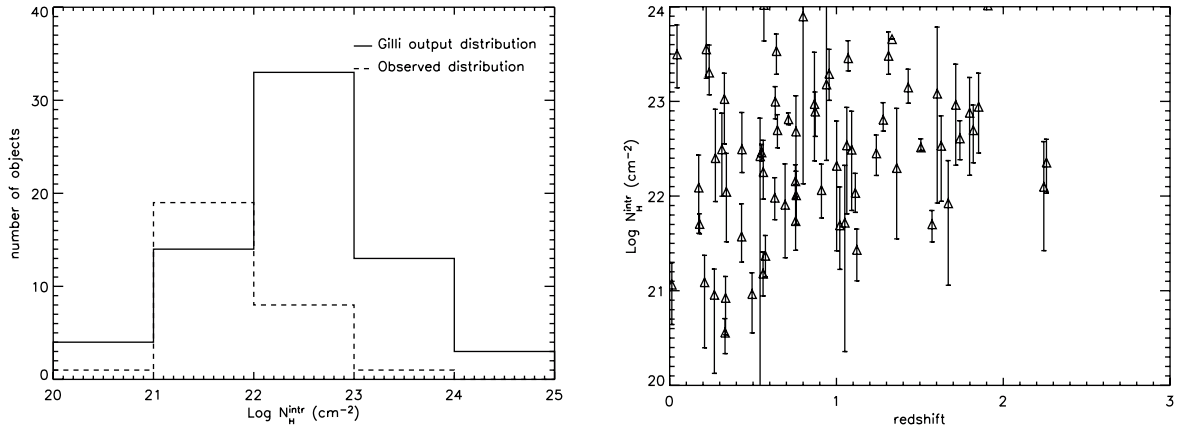

Fig. 19. *Left*: histograms of column density distributions in absorbed objects obtained from a simulated input distribution of $N_{\rm H}^{\rm intr}$ according to the model of Gilli et al. (2001). The dashed histogram represents the observed distribution for our sources. *Right*: $N_{\rm H}^{\rm intr}$ as a function of sources' redshifts from simulations.

included in our list, by fixing their redshifts, exposure times, position and luminosities. The slope of the continuum was assumed to be $\Gamma = 2$ for all the sources. For each source we extracted $N_{\rm H}^{\rm intr}$ values from the specific model distribution, until the simulated spectrum fulfilled the quality filters that we used for the real spectra (binned spectra with at least 5 bins and source net counts above 50). Indeed, heavily obscured sources do not pass, just as happens for real sources. We fitted the spectra with model A and model C and computed F-test significances for absorption, using the same calibration matrices.

Figure 19 compares the values of $N_{\rm H}^{\rm intr}$ measured in the real data and in the simulations for one of the XRB models (Gilli et al. 2001). Obviously, the adopted distribution is not able to reproduce the observational results. Very similar graphs are obtained for the models of Comastri et al. (1995) and Pompilio et al. (2000). For all three models there is a clear excess of sources with intrinsic absorptions $N_{\rm H}^{\rm intr} \geq 10^{22}$ cm^{-2}, which are not present in the real data. For an F-test confidence level of ≥95% we detect, in our real sample, absorption in 29 sources out of 179. The number of absorbed sources detected in the simulated data sets were 79, 67 and 71 for the Comastri et al. (1995), Gilli et al. (2001) and Pompilio et al. (2000) distributions, in all cases a factor of more than 2.5 over the observed value. This result is also evident from Fig. 19 (right panel), where the predicted distribution of absorbing columns with the Gilli et al. 2001 model is plotted against redshift. We find excess of absorbed objects all over the redshift interval up to redshifts of ∼2. Discrepancies between the predictions from synthesis models and what it is obtained from observations have been observed previously (i.e. Piconcelli et al. 2002; Caccianiga et al. 2004; Georgantopoulos et al. 2004; Perola et al. 2004).

In the recent "modified unified scheme" of Ueda et al. (2003) the AGN absorbing column density distribution depends on luminosity (not redshift dependence). The fraction of absorbed AGNs ($N_{\rm H} > 10^{22}$ cm^{-2}) decreases significantly with luminosity. This model will reduce the observed discrepancies fundamentally at high luminosities, and therefore at high redshifts. However, this model predicts a significant fraction of objects with $N_{\rm H} > 10^{22}$ cm^{-2} at the typical luminosities of our sample that we do not detect. The Ueda et al. (2003) model, although more in line with our results, still overpredicts the number of absorbed objects ($N_{\rm H} > 10^{22}$ cm^{-2}) especially at low redshifts.

One possible caveat is that some of these absorbed sources that we do not detect in the real spectra might be hiding within the unidentified sources. Our simulations indicate that ∼40% should have detectable absorption. However, even for the full sample of 1137 objects we only detect absorption in 10–20% of them. Assuming that the redshift distribution of the identified sources does not strongly deviate from the true one (and there is evidence that this is the case from other surveys), then 40% of absorbed sources predicted by the simulations should also be found in the whole sample, while only ∼20% are seen. This confirms that XRB synthesis models predict twice as many absorbed sources at our flux levels than we actually observe.

8. Conclusions

We have conducted a detailed analysis of the broad band spectral properties of a large sample of sources detected serendipitously with the *XMM-Newton* observatory. Various spectral models have been tested to reproduce the 0.2–12 keV emission of our sources. We summarise our main conclusions as follows:

1. Fitting the spectra of all the sources with a single power law we obtain a weighted mean of $\langle\Gamma\rangle = 1.86 \pm 0.02$, in agreement with previous surveys. However, for a significant number of sources, the quality of their fits was very poor, with 7–12% of the sources having statistically unacceptable fits (null hypothesis probability <5%).
2. We observed the continuum average spectral slope to become harder with decreasing 0.5–2 keV flux. This effect does not appear to be correlated with the quality of the spectral data. After correcting for all the biases that could be affecting our results, we still see the same dependence, therefore the hardening of Γ with decreasing soft flux is real.
3. Many objects in our sample exhibit absorption in excess of the Galactic value. In terms of the F-test (95% confidence level), excess absorption is found in 17% of

sources. Significant fractions of X-ray absorbed objects were found among NELGs (~40%), BLAGNs (~6%) and ALGs (~40%). This result clearly suggests that the relationship between optical identification and X-ray spectral signatures is still unclear and needs some revision.

4. The distributions of absorbing columns in BLAGNs and NELGs do not appear significantly different within the limitations of our sample, except for the overall factor that NELGs are 4 times more frequently absorbed than BLAGNs.
5. The typical absorption columns derived for BLAGNs ($N_{\rm H} < 10^{22}$ cm^{-2}) indicate that perhaps the host galaxy, or some sort of weak distant absorber might provide the absorbing gas, without blocking too much the line of sight to the Broad-Line region. This point deserves further exploration.
6. Using the best fit model for each object, i.e., including absorption and soft excess when required by the data, we still observe a substantial hardening of the average spectral index towards fainter soft fluxes. The same effect was also observed for the sub-samples of BLAGNs and NELGs. With the current data quality, we cannot reject that these sources have the same underlying power law ($\Gamma \sim 2$) with moderate absorption, nor can we reject that there is a population with genuinely flatter spectral index. However when hard fluxes are used to study this effect there is no evidence of a change of average spectral slope. This supports the hypothesis that absorption is producing the hardening of the average spectra of our sources, because hard band fluxes are less affected by absorption.
7. Using a maximum likelihood fit we have detected an intrinsic dispersion of the spectral slope, even within the subsamples of BLAGNs and NELGs, with the largest scattering observed for the NELGs.
8. We did not find a tendency of absorbed BLAGNs to occur preferentially at any particular redshifts. The same result is obtained when comparing the 2–10 keV luminosity distributions of absorbed and unabsorbed BLAGNs and NELGs. Therefore, no evidence is found of unabsorbed BLAGNs and NELGs having different X-ray luminosities to the absorbed objects.
9. We have seen that our unabsorbed NELGs are on average less luminous than the unabsorbed BLAGNs and hence our data are consistent with the explanation of the lack of broad lines in the optical spectra of unabsorbed NELGs as an AGN/ host-galaxy contrast effect.
10. A soft excess component was detected with a high significance in ~7% of BLAGNs and NELGs.
11. For XRB synthesis models based on the AGN unified scheme, simulations were conducted to test the distribution of absorbing columns from these models on the population of BLAGNs and NELGs. In general, the number of sources where absorption was detected in the simulated spectra (in about ~40% of them) was a factor of ~2.5 above the measured values for all models. Regardless of the incompleteness of the identifications, the XRB synthesis model prediction of detectable absorption in 40% of the sources is at odds with the overall observed rate of absorbed sources in the real sample.
12. The distribution of absorbing column densities predicted by the synthesis models and measured through the simulations is significantly skewed towards high column densities $N_{\rm H}^{\rm intr} \sim 10^{22}$–$10^{24}$ cm^{-2}. These highly absorbed sources, predicted by the XRB models at our flux levels, are not detected in our sample of objects.

Acknowledgements. S.M. acknowledges support from a Universidad de Cantabria fellowship. S.M., X.B., F.J.C. and M.T.C. acknowledge financial support from the Spanish Ministerio de Educación y Ciencia, under projects AYA2000-1690 and ESP2003-00812. This project was supported in part by the German BMBF under DLR grant 50 OX 0201. A.C. and T.M. acknowledge financial support from Italian Ministry of University and Scientific and Technological Research (MURST) through grant Cofin. We also thank the anonymous referee for his/her suggestions that helped to improve the manuscript considerably.

Appendix A: Sources' detection efficiency function

The use of different energy bands with different sensitivities to detect our sources, the selection of fields covering a wide range of exposure times (from ~10 to ~100 ks), and the quality filters applied to the spectra (especially the selection of spectra with MOS+pn counts ≥50) makes the calculation of the efficiency of source detection as a function of flux difficult. We know it will be a function of the X-ray flux, but it is important to know whether it also depends on the spectral shape of the sources.

We have conducted simulations to calculate the efficiency of detection, W(Γ, S): for a given number of sources with spectral parameters Γ and S, W(Γ, S) gives us the fraction of simulated objects that were detected, i.e., their spectra fulfilled the quality filters that we applied to our data, see Sect. 3). and with best fit spectral parameters being the same as the input (simulated) ones. This function can be used to correct observed source properties, for example the dependence of $\langle\Gamma\rangle$ with the X-ray flux, from all the effects listed above. We defined a grid of points in Γ-S covering the range of values measured for our objects (Γ from 0.5 to 3; S from 10^{-16} to 10^{-12} erg cm^{-2} s^{-1}).

We have simulated the same number of spectra on each grid point. For these simulations we have used the list of detected sources before applying the quality filters, i.e., 2145 sources have been simulated for each point in the Γ-S grid. The input parameters of each simulated spectrum were the exposure time, response matrices and Galactic absorption of each real source, and the values of Γ and S at each grid point. We used model A (see Sect. 4.1) to simulate all the spectra, therefore absorption is not included in the simulations. Once all the spectra were simulated, we applied the same quality filters that we used for our real data. Our source detection efficiency, at each grid point, was defined as

$$W(\Gamma, S) = \frac{N_{\rm det}(\Gamma, S) + N_{\rm det}(\Gamma', S')}{N_{\rm sim}}$$

where $N_{\rm sim}$ is the number of sources that were simulated at the grid point (Γ, S) (this number is the same for all the grid points).

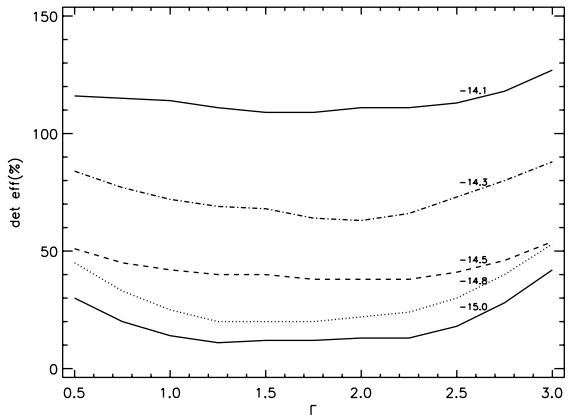

Fig. A.1. Dependence of the detection efficiency of sources as a function of Γ at different flux levels, obtained from simulations. Some values of the efficiency function can be above 100% because the fitted parameters may differ from the input ones (see text for details).

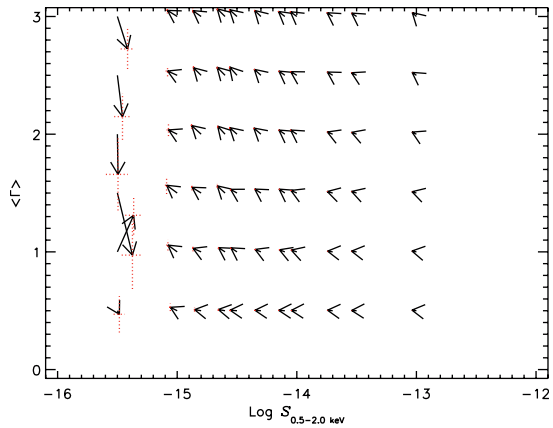

Fig. A.2. Representation of the average movements in the Γ-*S* plane obtained from simulations. The beginning point of each arrow indicates the input values of Γ and *S* in the simulations; the end points are the observed (fitted) average values of Γ and *S* for the detected sources.

It is important to notice that, there are two possible contributions to the final number of sources that will appear at a given grid point (Γ, *S*):

1. The number of simulated sources at the grid point (Γ, *S*) that, after fitting, remain at the same grid point, $N_{\rm det}(\Gamma, S)$.
2. The number of sources simulated at different grid points (Γ′, *S*′) that were detected, and with best fit spectral parameters (Γ, *S*), $N_{\rm det}$ (Γ′, *S*′).

The values of the detection efficiency as a function of Γ are plotted in Fig. A.1 for different fluxes. We observe that, at fluxes below $\sim 10^{-14}$ erg cm^{-2} s^{-1}, the efficiency of detection decreases significantly from Γ = 0.5 to Γ = 2, and then, starts to increase up to Γ = 3. The rise of efficiency at high values of Γ could be due to the sharp increase in the effective areas of the EPIC detectors at low energies combined with the increase in the relative contribution to the total counts at these energies for the steeper spectral slopes. It is interesting to notice that, although we have shown that the detection efficiency increases for Γ values above 2.5, we have not detected these super-soft sources in our sample, hence, they must be rare. The reason why the efficiency function is higher for objects with hard (Γ ∼ 0.5) spectral slopes (notice that the effective area of the X-ray detectors drops rapidly at energies below ∼1 keV) is that a source with a flat spectral slope has to be much brighter than one with a soft (Γ ∼ 2) spectral slope in order to have the same 0.5–2 keV flux. Therefore we will receive more counts in the 0.2–12 keV band from flatter objects, which means that it will better detect them and more easily pass our spectral quality filters.

Our simulations showed that there are movements of objects in the Γ-*S* plane, i.e., the fitted parameters may differ from the input ones. It is necessary to study how important this effect is and whether there are systematic shifts (for example a tendency to detect the sources with lower values of gamma at faint fluxes). Hence, we have calculated for each grid point the average values of Γ and *S*, in the same way as we did for the real sample, from all the sources simulated at that point that were detected (no matter at which grid point they were detected). The results, shown in Fig. A.2, confirm the existence of shifts in the Γ-*S* plane at the faintest fluxes: on average, the measured Γ values tend to be softer, and hence the observed 0.5–2 keV fluxes are lower. However, for fluxes $\geq 10^{-15}$ erg cm^{-2} s^{-1} these movements are not significant and tend to soften than harden spectra.

References

Akiyama, M., Ohta, K., Yamada, T., et al. 2000, AJ, 532, 700
Akiyama, M., Ueda, Y., Ohta, K., et al. 2003, ApJS, 148, 275
Almaini, O., Shanks, T., Boyle, B. J., et al. 1996, MNRAS, 282, 295
Baldi, A., Molendi, S., Comastri, A., et al. 2002, ApJ, 564, 190
Barcons, X., Carrera, F. J., Watson, M. G, et al. 2002, A&A, 382, 522
Barcons, X., Carrera, F. J., & Ceballos, M. T. 2003, MNRAS, 346, 897
Barger, A. J., Cowie, L. L., Mushotzky, R. F., et al. 2001, AJ, 121, 662
Barger, A. J., Cowie, L. L., Brandt, W. N., et al. 2002, AJ, 124, 1839
Bauer, F.E., Alexander, D. M., Brandt, W. N., et al. 2004, AJ, 128, 2048
Bevington, P. R., & Robinson, D. K. 1992, Data reduction and error analysis for the Physical sciences (McGraw Hill)
Boyle, B., Georgantopoulos, I., Blair, A. J., et al. 1998, MNRAS, 296, 1
Brusa, M., Comastri, A., Mignoli, M., et al. 2003, A&A, 409, 65
Caccianiga, A., Severgnini, P., Braito, V., et al. 2004, A&A, 416, 901
Cagnoni, I., Della Ceca, R., & Maccacaro, T. 1998, ApJ, 493, 54
Carrera, F. J., Page, M. J., & Mittaz, J. P. D. 2004, A&A, 420, 163
Ciliegi, P., Elvis, M., Wilkes, B. J., Boyle, B. J., & McMahon, R. G. 1997, MNRAS, 284, 401
Comastri, A., Setti, G., Zamorani, G., et al. 1995, A&A, 296, 1
Comastri, A., Fiore, F., Vignali, C., et al. 2001, MNRAS, 327, 781
Comastri, A., Mignoli, M., Ciliegi, P., et al. 2002, ApJ, 571, 771
Cowie, L. L., Garmire, G. P., Bautz, M. W., et al. 2002, ApJ, 566, L5-L8
Cruddace, R. G., Hasinger, G., & Schmitt, J. H. 1987, in Astronomy from Large Databases, ed. F. Murtaugh, & A. Heck, 177
della Ceca, R., Castelli, G., Braito, V., et al. 1999, AJ, 524, 674
Dickey, J. M., & Lockman, F. J. 1990, ARA&A, 28, 215
Feng, Y., & Narayan, R. 2004, ApJ, 612, 724
Fiore, F., Giommi, P., Vignali, C., et al. 2001, MNRAS, 327, 771
Fiore, F., Brusa, M., Cocchia, F., et al. 2003, A&A, 409, 79

Gabriel, C., Denby, M., Fyfe, D. J., et al. 2004, in Astronomical Data Analysis Software and Systems XIII, ed. F. Ochsenbein, M. Allen, & D. Egret, ASP Conf. Ser., 314, 759
Georgakakis, A., Georgantopoulos, I., Stewart, G. C., et al. 2003, MNRAS, 344, 161
Georgantopoulos, I., Stewart, G. C., Blair, A. J., et al. 1997, MNRAS, 291, 203
Georgantopoulos, I., Georgakakis, A., Akylas, A., et al. 2004, MNRAS, 352, 91
George, I. M., Turner, T. J., Yaqoob, T., et al. 2000, AJ, 531, 52
Giacconi, R., Rosati, P., Tozzi, P., et al. 2001, ApJ, 551, 624
Gilli, R., Salvati, M., & Hasinger, G. 2001, A&A, 366, 407
Giommi, P., Perri, M., & Fiore, F. 2000, A&A, 362, 799
Green, P. J., Silverman, J. D., Cameron, R. A., et al. 2004, ApJS, 150, 43G
Hasinger, G., Burg, R., Giacconi, R., et al. 1998, A&A, 329, 482
Hasinger, G., Altieri, B., Arnaud, M., et al. 2001, A&A, 365, L45
Jansen, F. A., Lumb, D., Altieri, B., et al. 2001, A&A, 365, L1
Longair, M. S. 1992, High Energy Astrophysics (Cambridge Univ. Press), Vol. 1
Lumsden, S. L., & Alexander, D. M. 2001, MNRAS, 328, L32
Maccacaro, T., Gioia, L. M., Wolter, A., et al. 1988, ApJ, 326, 680
Madau, P., Ghisellini, G., & Fabian, A. C. 1994, MNRAS, 270, L17
Mainieri, V., Bergeron, J., Hasinger, G., et al. 2002, A&A, 393, 425
Mateos, S., Barcons, X., Carrera, F. J., et al. 2003, AN, 324, 48
Mateos, S., et al. 2005, in preparation
Mittaz, J. P. D., Carrera, F. J., Romero-Colmenero, E., et al. 1999, MNRAS, 308, 233
Moran, E. C., Filippenko, A. V., & Chornock, R. 2002, ApJ, 579, L71
Moretti, A., Campana, S., Lazzati, D., & Tagliaferri, G. 2003, ApJ, 588, 696
Mushotzky, R. F., Cowie, L. L., Barger, A. J., et al. 2000, Nature, 404, 459
Nandra, K., & Pounds, K. A. 1994, MNRAS, 268, 405
Page, M. J., Mittaz, J. P. D., & Carrera, F. J. 2000, MNRAS, 318, 1073
Page, M. J., Mittaz, J. P. D., & Carrera, F. J. 2001, MNRAS, 325, 575
Page, M. J., McHardy, I. M., Gunn, K. F., et al. 2003, AN, 324, 101
Panessa, F., & Bassani, L. 2002, A&A, 394, 435
Pappa, A., Georgantopoulos, I., Stewart, G. C., et al. 2001, MNRAS, 326, 995
Perola., G. C., Puccetti, S., Fiore, F., et al. 2004, A&A, 421, 491
Piconcelli, E., Cappi, M., Bassani, L., et al. 2002, A&A, 394, 835
Piconcelli, E., Cappi, M., Bassani, L., et al. 2003, A&A, 412, 689
Pompilio, F., La Franca, F., & Matt, G. 2000, A&A, 353, 440
Protassov, R., van Dyk, D. A., Connors, A., Kashyap, V. L., & Siemiginowska, A. 2002, AJ, 571, 545
Risaliti, G., Maiolino, R., & Salvati, M. 1999, AJ, 522, 157
Romero-Colmenero, E., Branduardi-Raymont, G., Carrera, F. J., et al. 1996, MNRAS, 282, 94
Schartel, N., Schmidt, M., Fink, H. H., et al. 2001, A&A, 320, 696
Setti, G., & Woltjer, L. 1989, A&A, 224, L21
Severgnini, P., Caccianiga, A., Braito, V., et al. 2003, A&A, 406, 483
Silverman, J. D., Green, P. J., Aldcroft, T. D., et al. 2003, in AGN physics with the SDSS, ed. G. T. Richards, & P. T. Hall, in press [arXiv:astro-ph/0310907]
Stevens, J. A., Page, M. J., Ivison, R. J., et al. 2005, MNRAS, submitted
Tozzi, P., Rosati, P., Nonino, M., et al. 2001, AJ, 562, 42
Turner, T. J., & Pounds, K. A. 1989, MNRAS, 240, 833
Ueda, Y., Takahashi, T., Inoue, H., et al. 1998, Nature, 391, 866
Ueda, Y., Takahashi, T., Ishisaki, Y., et al. 1999, ApJ, 524, L11
Ueda, Y., Akiyama, M., Ohta, K., & Miyaji, T. 2003, ApJ, 598, 886
Vikhlinin, A., Forman, W., Jones, C., et al. 1995, ApJ, 451, 564
Watson, M. G., Auguères, J.-L., Ballet, J., et al. 2001, A&A, 365, L51
Worsley, M. A., Fabian, A. C., Barcons, X., et al. 2004, MNRAS, 352, L28
Zamorani, G., Gioia, I. M., Maccacaro, T., & Wolter, A. 1988, A&A, 196, 39