



ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA

## ARCHIVIO ISTITUZIONALE DELLA RICERCA

### Alma Mater Studiorum Università di Bologna Archivio istituzionale della ricerca

THE CHANDRA COSMOS LEGACY SURVEY: OVERVIEW AND POINT SOURCE CATALOG

This is the final peer-reviewed author's accepted manuscript (postprint) of the following publication:

*Published Version:*

THE CHANDRA COSMOS LEGACY SURVEY: OVERVIEW AND POINT SOURCE CATALOG / Civano, F.; Marchesi, S.; Comastri, A.; Urry, M.C.; Elvis, M.; Cappelluti, N.; Puccetti, S.; Brusa, M.; Zamorani, G.; Hasinger, G.; Aldcroft, T.; Alexander, D.M.; Allevato, V.; Brunner, H.; Capak, P.; Finoguenov, A.; Fiore, F.; Fruscione, A.; Gilli, R.; Glotfelty, K.; Griffiths, R.E.; Hao, H.; Harrison, F.A.; Jahnke, K.; Kartaltepe, J.; Karim, A.; Lamassa, S.M.; Lanzuisi, G.; Miyaji, T.; Ranalli, P.; Salvato, M.; Sargent, M.; Scoville, N.J.; Schawinski, K.; Schawinski, E.; Silverman, J.; Smolcic, V.; Stern, D.; Toft, S.; Trakhenbrot, B.; Treister, E.; Vignali, C.. - In: THE ASTROPHYSICAL JOURNAL, ISSN 0004-637X, ELETTRONICO, 819, 1(2016), pp. A62.1-A62.18. [10.3847/0004-637X/819/1/62]

*Published:*

DOI: <http://doi.org/10.3847/0004-637X/819/1/62>

*Terms of use:*

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>).  
When citing, please refer to the published version.

(Article begins on next page)



This is the final peer-reviewed accepted manuscript of:

F. Civano et al, THE CHANDRA COSMOS LEGACY SURVEY: OVERVIEW AND POINT SOURCE CATALOG, The Astrophysical Journal, 2016, 819, 62.

The final published version is available online at: <https://doi.org/10.3847/0004-637X/819/1/62>

Rights / License:

The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

*This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>)*

***When citing, please refer to the published version.***

THE *Chandra* COSMOS LEGACY SURVEY: OVERVIEW AND POINT SOURCE CATALOG

F. CIVANO<sup>1,2</sup>, S. MARCHESI<sup>1,3,2</sup>, A. COMASTRI<sup>4</sup>, M.C. URRY<sup>1</sup>, M. ELVIS<sup>2</sup>, N. CAPPELLUTI<sup>4</sup>, S. PUCCETTI<sup>5</sup>, M. BRUSA<sup>3,4</sup>, G. ZAMORANI<sup>4</sup>, G. HASINGER<sup>6</sup>, T. ALDCROFT<sup>2</sup>, D. M. ALEXANDER<sup>7</sup>, V. ALLEVATO<sup>8</sup>, H. BRUNNER<sup>9</sup>, P. CAPAK<sup>10,11</sup>, A. FINOGENOV<sup>8</sup>, F. FIORE<sup>12</sup>, A. FRUSCIONE<sup>2</sup>, R. GILLI<sup>4</sup>, K. GLOTFELTY<sup>2</sup>, R. E. GRIFFITHS<sup>13</sup>, H. HAO<sup>2</sup>, F. A. HARRISON<sup>14</sup>, K. JAHNKE<sup>15</sup>, J. KARTALTEPE<sup>16,17</sup>, A. KARIM<sup>18</sup>, S. M. LAMASSA<sup>1</sup>, G. LANZUISI<sup>3,4</sup>, T. MIYAJI<sup>19,20</sup>, P. RANALLI<sup>21</sup>, M. SALVATO<sup>9</sup>, M. SARGENT<sup>22</sup>, N. J. SCOVILLE<sup>11</sup>, K. SCHAWINSKI<sup>23</sup>, E. SCHINNERER<sup>15,24</sup>, J. SILVERMAN<sup>25</sup>, V. SMOLCIC<sup>26</sup>, D. STERN<sup>27</sup>, S. TOFT<sup>28</sup>, B. TRAKHENBROT<sup>23</sup>, E. TREISTER<sup>29</sup>, C. VIGNALI<sup>3,4</sup>

Nov 30, 2015

ABSTRACT

The *COSMOS-Legacy* survey is a 4.6 Ms *Chandra* program that has imaged 2.2 deg<sup>2</sup> of the COSMOS field with an effective exposure of  $\simeq 160$  ks over the central 1.5 deg<sup>2</sup> and of  $\simeq 80$  ks in the remaining area. The survey is the combination of 56 new observations, obtained as an X-ray Visionary Project, with the previous C-COSMOS survey. We describe the reduction and analysis of the new observations and the properties of 2273 point sources detected above a spurious probability of  $2 \times 10^{-5}$ . We also present the updated properties of the C-COSMOS sources detected in the new data. The whole survey includes 4016 point sources (3814, 2920 and 2440 in the full, soft and hard band). The limiting depths are  $2.2 \times 10^{-16}$ ,  $1.5 \times 10^{-15}$  and  $8.9 \times 10^{-16}$  erg cm<sup>-2</sup> s<sup>-1</sup> in the 0.5-2, 2-10 and 0.5-10 keV bands, respectively. The observed fraction of obscured AGN with column density  $> 10^{22}$  cm<sup>-2</sup> from the hardness ratio (HR) is  $\sim 50_{-16}^{+17}\%$ . Given the large sample, we compute source number counts in the hard and soft bands, significantly reducing the uncertainties of 5-10%. For the first time, we compute number counts for obscured (HR>0.2) and unobscured (HR<0.2) sources and find significant differences between the two populations in the soft band. Due to the un-precedent large exposure, *COSMOS-Legacy* area is 3 times larger than surveys at similar depth and its depth is 3 times fainter than surveys covering similar area. The area-flux region occupied by *COSMOS-Legacy* is likely to remain unsurpassed for years to come.

*Subject headings:* catalogs – cosmology: observations – galaxies: evolution – quasars: general – surveys – X-rays: general

<sup>1</sup> Yale Center for Astronomy and Astrophysics, 260 Whitney Avenue, New Haven, CT 06520, USA

<sup>2</sup> Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA

<sup>3</sup> Dipartimento di Fisica e Astronomia, Università di Bologna, viale Berti Pichat 6/2, 40127 Bologna, Italy

<sup>4</sup> INAF-Osservatorio Astronomico di Bologna, via Ranzani 1, 40127 Bologna, Italy

<sup>5</sup> ASDC-ASI, Via del Politecnico, 00133 Roma, Italy

<sup>6</sup> Institute for Astronomy, 2680 Woodlawn Drive, University of Hawaii, Honolulu, HI 96822, USA

<sup>7</sup> Centre for Extragalactic Astronomy, Department of Physics, Durham University, South Road, Durham DH1 3LE, UK

<sup>8</sup> Department of Physics, University of Helsinki, Gustaf Hällströmin katu 2a, FI-00014 Helsinki, Finland

<sup>9</sup> Max-Planck-Institut für extraterrestrische Physik, Giessenbachstrasse 1, D-85748 Garching bei München, Germany

<sup>10</sup> Infrared Processing and Analysis Center (IPAC), 1200 East California Boulevard, Pasadena, California 91125, USA

<sup>11</sup> California Institute of Technology, 1200 East California Boulevard, Pasadena, California 91125, USA

<sup>12</sup> INAF - Osservatorio Astronomico di Roma, via Frascati 33, Monte Porzio Catone, I00040, Italy

<sup>13</sup> Physics & Astronomy Dept., Natural Sciences Division, University of Hawaii at Hilo, 200 W. Kawili St., Hilo, HI 96720, USA

<sup>14</sup> Cahill Center for Astronomy and Astrophysics, California Institute of Technology, 1216 E. California Blvd, Pasadena, CA, 91125 USA

<sup>15</sup> Max Planck Institute for Astronomy, Königstuhl 17, D-69117 Heidelberg, Germany

<sup>16</sup> National Optical Astronomy Observatory, 950 N. Cherry Ave, Tucson AZ 85719, USA

<sup>17</sup> School of Physics and Astronomy, Rochester Institute of Technology, 84 Lomb Memorial Dr., Rochester, NY 14623, USA

<sup>18</sup> Argelander-Institut für Astronomie, Universität Bonn, Auf

dem Hügel 71, D-53121 Bonn, Germany

<sup>19</sup> Instituto de Astronomía sede Ensenada, Universidad Nacional Autónoma de México, Km. 103, Carret. Tijuana-Ensenada, Ensenada, BC, Mexico

<sup>20</sup> University of California San Diego, Center for Astrophysics and Space Sciences, 9500 Gilman Drive, La Jolla, CA 92093-0424, USA

<sup>21</sup> AASARS, National Observatory of Athens, 15236 Penteli, Greece

<sup>22</sup> Astronomy Centre, Department of Physics and Astronomy, University of Sussex, Brighton, BN1 9QH, UK

<sup>23</sup> Institute for Astronomy, Department of Physics, ETH Zurich, Wolfgang-Pauli-Strasse 27, CH-8093 Zurich, Switzerland

<sup>24</sup> National Radio Astronomy Observatory, Pete V. Domenici Science Operations Center, 1003 Lopezville Road, Socorro, NM 87801, USA

<sup>25</sup> Kavli Institute for the Physics and Mathematics of the Universe (Kavli IPMU, WPI), Todai Institutes for Advanced Study, The University of Tokyo, Kashiwa 277-8583, Japan

<sup>26</sup> Department of Physics, University of Zagreb, Bijenička cesta 32, HR-10000 Zagreb, Croatia

<sup>27</sup> Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109, USA

<sup>28</sup> Dark Cosmology Centre, Niels Bohr Institute, University of Copenhagen, Juliane Mariesvej 30, DK-2100 Copenhagen, Denmark

<sup>29</sup> Universidad de Concepción, Departamento de Astronomía, Casilla 160-C, Concepción, Chile

## 1. INTRODUCTION

One of the most active but least known epochs in astrophysics is the period between re-ionization ( $z \geq 8$ , i.e. when the Universe was less than  $\sim 0.6$  Gyr old), where the growth of structures becomes highly non-linear and the first stars form, and  $z \sim 2$  ( $\sim 3.25$  Gyr old), where major virialization occurs and star formation (SF) and supermassive black hole (SMBH) accretion peak.

At these early times, the pre-cursors of the clusters and groups seen at  $z \lesssim 1$  have low density and are much larger, on both physical (Mpc) and observed (arcminutes) scales. Surveys for these large scale structures become rapidly more efficient as the dimension of the survey exceeds the structure's typical sizes ( $\sim 15'$ ). Large area surveys (several times  $15'$  wide) are essential for the detection of these structures, which cannot be seen in smaller area surveys, however deep.

The equatorial  $2 \text{ deg}^2$  COSMOS area (Scoville et al. 2007a) is the deepest, most complete survey accessible to both hemispheres (notably by both ALMA and the Karl G. Jansky VLA) and large enough to find high-redshift clusters. A significant investment of 640 *HST* orbits (Scoville et al. 2007b, Koekemoer et al. 2007), 620h of *Spitzer* (Capak et al. in prep.), 260h of *Herschel* (Lutz et al. 2011), 750 hrs of JVLA (Schinnerer et al. 2004, 2007, 2010 and Smolcic et al. 2014 and in prep.), over 300 nights of large ground-based telescopes VLT, Keck, Subaru, VISTA for both imaging and spectroscopy (Lilly et al. 2009, Hasinger et al. in prep., Taniguchi et al. 2007, McCracken et al. 2014) has been made in this field.

The first homogeneous coverage in the X-rays of the whole COSMOS field was obtained with the XMM-*Newton* satellite (1.5 Megaseconds; Hasinger et al. 2007, Cappelluti et al. 2009, Brusa et al. 2010). These observations have been crucial for characterizing the most luminous Active Galactic Nuclei (AGN) in COSMOS (e.g. Brusa et al. 2010, Mainieri et al. 2011, Allevato et al. 2011, Lusso et al. 2012, among others). The obscured AGN population of the COSMOS field can be studied by jointly using the XMM-COSMOS data with the  $\sim 3$  Megaseconds of *NuSTAR* (Harrison et al. 2012) time available, which led to the discovery of a Compton-thick AGN (with obscuration exceeding  $10^{24} \text{ cm}^{-2}$  equivalent hydrogen column density) in the field (Civano et al. 2015), that was not recognized as such by XMM-*Newton* alone. Instead, for the faint and the high- $z$  AGN population, that could be responsible for the re-ionization of the Universe (see Giallongo et al. 2015), *Chandra* (Weisskopf et al. 2002) is the preferred instrument. Indeed, the large (1.8 Ms) *Chandra* COSMOS survey (C-COSMOS; Elvis et al. 2009, E09; Puccetti et al. 2009, P09; Civano et al. 2012) has already contributed significantly to the study of the early epochs of the Universe, finding: three luminous AGN residing in protoclusters between  $z \sim 4.55$  and 5.3 (Capak et al. 2011); the largest sample of X-ray selected  $z > 3$  quasars in a contiguous field (81 sources, Civano et al. 2011); a precocious SMBH in a normal size galaxy at  $z > 3$  (Trakhtenbrot et al. 2015); and AGN correlation lengths of  $7h^{-1} \text{ Mpc}$  ( $\sim 10'$ ) at  $z \sim 1-2$ , Allevato et al. 2011). However, C-COSMOS only covered  $\frac{1}{4}$  of COSMOS at  $\sim 160$  ks depth, plus  $0.5 \text{ deg}^2$  at  $\sim 80$  ks depth (Fig. 1, green squares).

We present here the *Chandra COSMOS-Legacy* survey<sup>30</sup>, the combination of the old C-COSMOS survey with 2.8 Ms of new *Chandra* ACIS-I (Garmire et al. 2003) observations ( $56 \times 50$  ks pointings) approved during *Chandra* Cycle 14 as an X-ray Visionary Project (PI: F. Civano; program ID 901037). *COSMOS-Legacy* uniformly covers the  $\sim 1.7 \text{ deg}^2$  COSMOS/*HST* field at  $\sim 160$  ksec depth, expanding on the deep C-COSMOS area (dashed green square in Fig. 1) by a factor of  $\sim 3$  at  $\sim 3 \times 10^{-16} \text{ erg cm}^{-2} \text{ s}^{-1}$  ( $1.45$  vs  $0.44 \text{ deg}^2$ ), for a total area covered of  $\sim 2.2 \text{ deg}^2$ .

This paper is the first in a series and presents the main properties of the survey and the X-ray point source catalog to be followed by a paper on the multiwavelength identification of the X-ray sources by Marchesi et al. (in preparation). In Section 2, we present the observations and tiling strategy. In Section 3, we detail all the steps of the data processing, including astrometric corrections, exposure and background map production. The data analysis procedure is instead described in Section 4, with some references and comparison with the one adopted for C-COSMOS as explained in P09. The point source catalog and the source properties are presented in Section 4.1. Last, in Section 5 and 6, the survey sensitivity and the number counts in both soft and hard band, and also dividing the sources in obscured and unobscured, are presented.

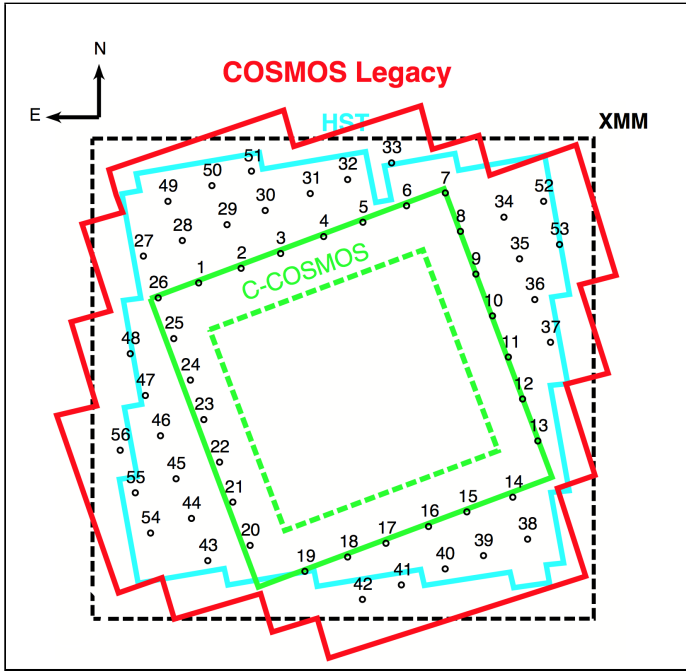
We assume a cosmology with  $H_0 = 71 \text{ km s}^{-1} \text{ Mpc}^{-1}$ ,  $\Omega_M = 0.3$  and  $\Omega_\Lambda = 0.7$ , and magnitudes are reported in the AB system if not otherwise stated. Throughout this paper, we make use of *J2000.0* coordinates. The data analysis is performed in three X-ray bandpasses 0.5-2 keV (soft band, S), 2-7 keV (hard band, H), and 0.5-7 keV (full band, F), while sensitivity and fluxes have been computed in the 0.5-2, 2-10 and 0.5-10 keV bands for an easy comparison with other works in the literature.

## 2. OBSERVATIONS

The half-a-field shift tiling strategy was designed in order to cover uniformly, in depth and point spread function size, the COSMOS *Hubble* area (cyan outline in Fig. 1; Scoville et al. 2007b), by combining the old C-COSMOS observations (green outline in Fig. 1) with the new *Chandra* ones (red outline in Fig. 1). To achieve this, 56 ACIS-I pointings (numbered black points in Fig. 1) were used, 11 of which were scheduled as two or more separate observations because of satellite constraints, for a total of 68 pointings. Moreover, the observing roll angle was constrained to be within  $70 \pm 20$  degree or  $250 \pm 20$ . We summarize the main properties of the new *Chandra* COSMOS Legacy observations in Table 1.

The observations took place in four blocks: November, 2012 to January, 2013; March to July, 2013; October, 2013 to January, 2014; and March, 2014. The mean net effective exposure time per field was 48.8 ks, after all the cleaning and reduction operations (see Section 4). The maximum exposure was 53 ks (observation 15227) while the minimum exposure was 45.2 ks (combined observations 15208 and 15998).

<sup>30</sup> Throughout the paper we use the term C-COSMOS to refer to the original survey of the inner field, and the name *Chandra COSMOS-Legacy* survey to refer to the full, combined survey, including the new data presented here.



**Figure 1.** *COSMOS-Legacy* tiling (red) compared to the area covered by HST (cyan), C-COSMOS (green solid: total area; green dashed: deeper area) and XMM-COSMOS (black). The ordering numbers of new observations are marked (See Table 1).

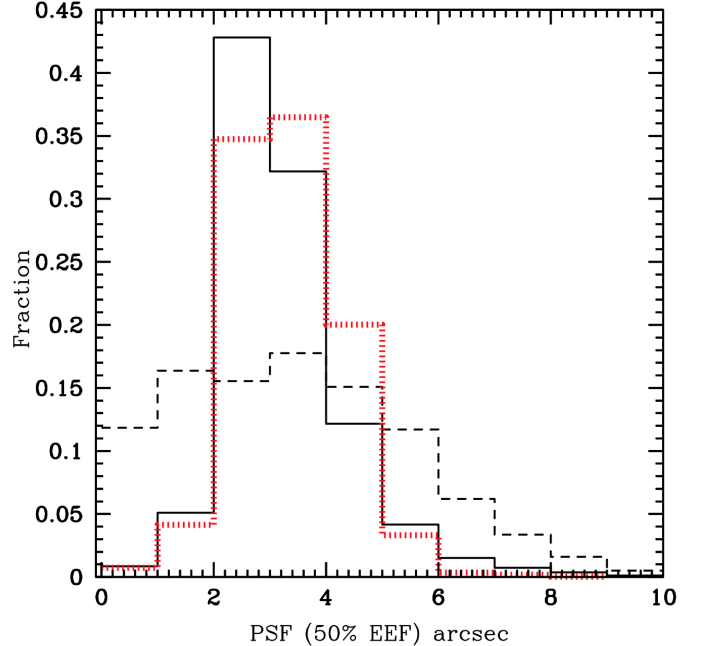
The sequence of the observations was designed to start from the N-E top corner tile of C-COSMOS moving towards W and proceeding clockwise around the central C-COSMOS area, in such a way that the outer frame of the C-COSMOS survey overlaps with the inner frame of the new *Chandra* observations. The tiling number and the total area covered is shown in Figure 1.

Using this tiling strategy we achieve an approximately uniform combined point spread function (PSF) across the survey. The mean combined PSF width (size at 50% of the encircled energy fraction, EEf, in the 0.5-7 keV band; see Section 5 for details on the PSF maps), weighted on the exposure, peaks at around 3'' (see Figure 2). As shown in Figure 2, 80% of the field has a PSF in the range 2''-4''. As a comparison, in a single-pointed survey (regardless of exposure time), the PSF size distribution is flat, and although  $\sim 30\%$  of the field has a PSF  $< 2''$  the PSF can reach a substantially larger size ( $> 4''$ ) in 40% of the field.

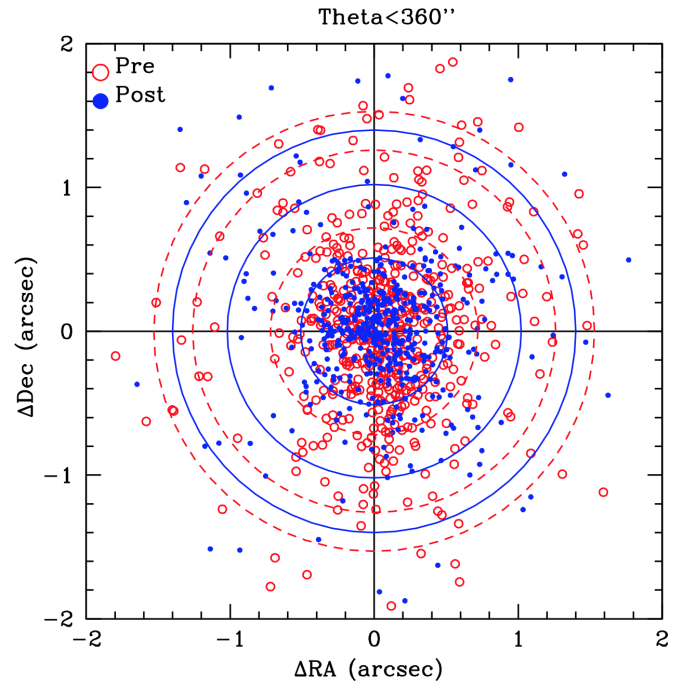
### 3. DATA PROCESSING

The data reduction was performed following the procedures described in E09 for C-COSMOS, using standard *Chandra* CIAO 4.5 tools (Fruscione et al. 2006) and CALDB 4.5.9. We also reprocessed the 49 C-COSMOS observations in order to use them in concert with the new observations for source detection in the area where the new observations overlap with the old ones and to compute the sensitivity of the whole survey (see the comparison between fluxes in Section 4.1.3).

We used the `chandra_repro` reprocessing script, which automates the CIAO recommended data processing steps and creates new level 2 event files, applying the *VFAINT* mode for ACIS background cleaning to all the observations. We then performed the following steps before



**Figure 2.** Normalized distribution of the combined point spread function (50% of EEf in 0.5-7 keV) size in arcsecond measured in *COSMOS-Legacy* (solid histogram) and in a single pointing survey (dashed line). In red, the distribution of the combined PSF (the mean value) for all the detected sources.



**Figure 3.** The X-ray to I-band separation ( $\Delta RA$ ,  $\Delta dec$ ) in arcsecond for X-ray sources within  $6'$  from the aim point detected in each single observations before (red open circles) and after (blue solid circles) the aspect correction. The circles encompass 68%, 90% and 95% of the sources before (red dashed) and after (blue solid) the correction.

starting data analysis: astrometric correction and reprocessing of all the observations to a standard frame of reference using the new aspect solution (Section 3.1); mo-

**Table 1**  
COSMOS Legacy Survey (CLS) Observation Summary

Field <sup>a</sup>	Obs. ID	RA	Dec.	Date	Exp. time second	Roll (deg)
CLS_1	15207	150.544451	2.499045	2012-11-25	14883	70.2
	15590	150.544402	2.499094	2012-11-23	14893	70.2
	15591	150.544454	2.499065	2012-11-25	19828	70.2
CLS_2	15208	150.415643	2.543225	2012-12-07	22985	70.2
	15598	150.415625	2.543213	2012-12-08	22193	70.2
CLS_3	15209	150.295749	2.588083	2012-12-03	23775	70.2
	15600	150.295747	2.588106	2012-12-05	21795	70.2
CLS_4	15604	150.164741	2.639752	2012-12-10	20988	70.2
	15210	150.164738	2.639709	2012-12-16	24365	70.2
CLS_5	15211	150.045569	2.682903	2012-12-13	23572	70.2
	15605	150.045586	2.682879	2012-12-15	21801	70.2
CLS_6	15212	149.913425	2.732850	2012-12-21	25249	70.2
	15606	149.913418	2.732845	2012-12-23	25219	70.2
CLS_7	15213	149.796052	2.772968	2013-01-01	49435	62.2
CLS_8	15214	149.751287	2.655331	2013-01-03	45983	61.75
CLS_9	15215	149.704144	2.525446	2013-01-07	49437	63.2
CLS_10	15216	149.654208	2.399733	2013-01-16	46459	56.7
CLS_11	15217	149.627509	2.272922	2013-03-23	46057	265.2
CLS_12	15218	149.584767	2.145874	2013-03-22	46475	265.2
CLS_13	15219	149.538688	2.017596	2013-03-30	49432	261.6
CLS_14	15220	149.614659	1.846399	2013-04-04	49924	60.1
CLS_15	15221	149.753949	1.801935	2013-04-10	49431	58.2
CLS_16	15222	149.870306	1.757718	2013-04-04	49407	60.0
CLS_17	15223	149.999623	1.706079	2013-04-17	50905	55.2
CLS_18	15224	150.115609	1.664373	2013-04-19	49426	55.2
CLS_19	15225	150.245495	1.621716	2013-04-05	49631	59.8
CLS_20	15226	150.411336	1.697830	2013-06-21	49428	250.2
CLS_21	15227	150.463753	1.829216	2013-05-02	53051	50.2
CLS_22	15228	150.504029	1.950647	2013-04-30	49432	50.2
CLS_23	15229	150.551660	2.080265	2013-05-10	49012	52.2
CLS_24	15230	150.592692	2.199969	2013-05-08	49429	52.2
CLS_25	15231	150.642972	2.325853	2013-05-13	48446	51.0
CLS_26	15232	150.690403	2.449284	2013-05-16	35085	50.6
	15649	150.690409	2.449315	2013-06-03	15251	50.65
CLS_27	15233	150.734924	2.575875	2013-05-21	46476	50.20
CLS_28	15234	150.616710	2.623373	2013-05-22	5439	50.20
	15653	150.594589	2.629150	2014-01-16	44895	58.21
CLS_29	15235	150.480833	2.671101	2013-06-01	49440	48.31
CLS_30	15236	150.364822	2.714260	2013-06-01	49435	48.23
CLS_31	15237	150.228563	2.765929	2013-06-08	25246	50.65
	15655	150.228550	2.765907	2013-06-10	24466	50.65
CLS_32	15238	150.114727	2.808579	2013-06-09	49429	50.65
CLS_33	15239	149.981160	2.858739	2013-06-11	49430	50.20
CLS_34	15240	149.617459	2.695912	2013-10-15	48450	77.09
CLS_35	15241	149.593992	2.566795	2014-03-28	48600	260.21
CLS_36	15242	149.547344	2.443553	2013-06-22	49432	50.20
CLS_37	15243	149.499113	2.312060	2013-07-05	47985	50.20
CLS_38	15244	149.547442	1.723759	2014-01-21	47461	53.21
CLS_39	15245	149.680897	1.673516	2014-01-23	49437	53.21
CLS_40	15246	149.796705	1.629086	2013-10-22	48850	75.21
CLS_41	15247	149.953115	1.578784	2014-03-18	49545	267.21
CLS_42	15248	150.047419	1.537353	2013-11-13	49438	71.61
CLS_43	15249	150.516513	1.655025	2013-11-29	45635	70.21
CLS_44	15250	150.566083	1.783479	2013-12-12	49315	70.21
CLS_45	15251	150.612991	1.904134	2013-12-03	29702	67.91
	16544	150.613008	1.904121	2013-12-04	19830	67.91
CLS_46	15252	150.660018	2.034094	2013-12-14	49434	70.21
CLS_47	15253	150.707972	2.162869	2014-01-28	49132	53.21
CLS_48	15254	150.753963	2.289683	2014-01-29	49139	53.21
CLS_49	15255	150.661405	2.741395	2014-03-24	49435	260.21
CLS_50	15256	150.504801	2.795740	2014-01-13	49943	59.21
CLS_51	15257	150.384246	2.838987	2014-01-04	49435	61.85
CLS_52	15258	149.497504	2.746858	2014-01-01	49432	62.27
CLS_53	15259	149.451159	2.620733	2014-01-27	49435	53.21
CLS_54	15260	150.690908	1.740589	2014-01-05	22793	60.21
	16562	150.690921	1.740576	2014-01-25	26736	60.21
CLS_55	15261	150.736977	1.863191	2014-01-18	46474	59.21
CLS_56	15262	150.782957	1.992193	2014-01-12	50236	59.21

saic and exposure map creation in three standard *Chandra* bands (Section 3.2): 0.5-7 keV, 0.5-2 keV and 2-7 keV; background map creation, using a two-components model to take into account both the cosmic background contribution and the instrumental one (Section 3.3).

### 3.1. Astrometric corrections

Even though *Chandra* data astrometry is accurate to  $0.6''$  (at 90% confidence, see Proposer User Guide<sup>31</sup> Chapter 5), in order to produce a sharp X-ray mosaic and to match the positions of X-ray sources with the optical catalog for which the positional accuracy is  $\sim 0.2''$  (Capak et al. 2007, Ilbert et al. 2009, Laigle et al. submitted), we performed source detection on each individual observation to register them to a common optical astrometric frame. This work has been done on the new observations and also on the C-COSMOS outer frame fields overlapping with the new data. We generated a list of detected sources using the CIAO wavelet source detection tool `WAVDETECT` on each single observation binned at  $1''$  and adopted a false-positive detection probability threshold corresponding to  $\sim 10$  spurious sources per field. Of the detected sources (on average 150 sources per field), we considered in each field those with significance  $> 3.5\sigma$  and within  $360''$  from the aim point. In *Chandra* data, the positional accuracy of significant sources is  $< 1''$  even at  $10'$  off-axis and it is energy independent (K. Glotfelty private communication). Therefore, choosing sources within  $6'$  of the aim-point provides a sample of sources with very good centroid estimate ( $< 0.3''$ ) for astrometric purposes. Using the CIAO tool `reproject_aspect`, these sources were then compared to the CFHT *MegaCam* catalog of i-band selected sources (McCracken et al. 2012) with optical magnitudes in the range 18-23 AB *mag*. At least 4 sources in each field, not on the same side of the aim-point, are needed to compute meaningful rotational and translation transformations. In our analysis, we used on average 12 sources (up to 22 sources per field), with 75% of the fields having more than 10 sources used to perform the reproject.

With the corrected aspect solution, we reprocessed the level 1 data using `chandra_repro` and performed the `WAVDETECT` detection again to compute the new separation between X-ray and optical positions. The resulting standard deviation on the shift computed from the detected sources within  $6'$  is  $0.36''$  and  $0.51''$  on the RA and Dec, respectively. After matching all the X-ray fields to the same astrometric optical frame, 95% of the X-ray sources used for the astrometry correction have a distance to their optical counterpart smaller than  $1.4''$ , 10% lower than the value before the correction ( $1.53''$ ). The improvement in the position increases to 20% when considering 90% of the sources ( $1.26''$  to  $1.02''$ ) and 30% when considering a smaller sample of 68% of the sources (from  $0.72''$  to  $0.51''$ ; see Figure 3). This is consistent with, and slightly better than, what was found for C-COSMOS (see E09, Figure 6).

### 3.2. Exposure maps and data mosaic creation

We created exposure maps in three bands using the standard CIAO procedure. The spectral model used

for the map creation is a single power-law with slope  $\Gamma=1.4$ <sup>32</sup> and Galactic absorption  $N_H=2.6\times 10^{20}$   $\text{cm}^{-2}$  (Kalberla et al. 2005). Instrument maps, generated with `MKINSTMAP` for each CCD in each observation, were used as input files for the `MKEXPMAP` tool, which computes an exposure map for each CCD separately. These exposure maps were combined in a single exposure map for each observation using `DMREGRID` with a binning of 2 pixels.

Figure 4 shows a composite image of the effective exposure time (in seconds) in the full band for both the new observations (left) and the whole *COSMOS-Legacy* (right). As can be seen, the central  $1.5 \text{ deg}^2$ , covering almost entirely the HST area, have a uniform depth of  $\simeq 160$  ks.

The data mosaic image was created in three bands using the HEASoft `addimages` tool, which adds together a set of images using sky coordinates. In Figure 5, the three color image, created by combining the exposure corrected images in three non-overlapping bands (0.5-2.0 keV, 2.0-4.5 keV, and 4.5-7.0 keV as red green and blue, respectively) is shown. The combined image was then Gaussian smoothed with a 3 pixel radius. A filter was then applied to isolate sources from the background level, as well as to increase the contrast and color vibrancy of those sources. This process was repeated 3 times.

### 3.3. Background maps creation

The *Chandra* background consists of two different components: the cosmic X-ray background and a quiescent instrumental background due to interactions between the ACIS-I CCD detectors and high-energy particles. We followed the procedure described in Cappelluti et al. (2013) to create background maps, which we used for the selection of reliable sources in our detection procedure and for the computation of the sensitivity curves.

The background maps were computed for each observation separately in the full, soft and hard bands. We ran `WAVDETECT` with a threshold parameter `sighthresh=10-5`, corresponding to  $\sim 100$  spurious sources per field (see Section 3.1), large enough to select also sources with significant signal only in stacked emission. We then removed these sources from the science images by excising a region corresponding to the source size (using a  $3\sigma$  value) as computed by the detection tool. We then uniformly distributed the remaining counts, rescaled by the ratio between the whole area of the observation and the area without the removed sources. These files were then used as initial background.

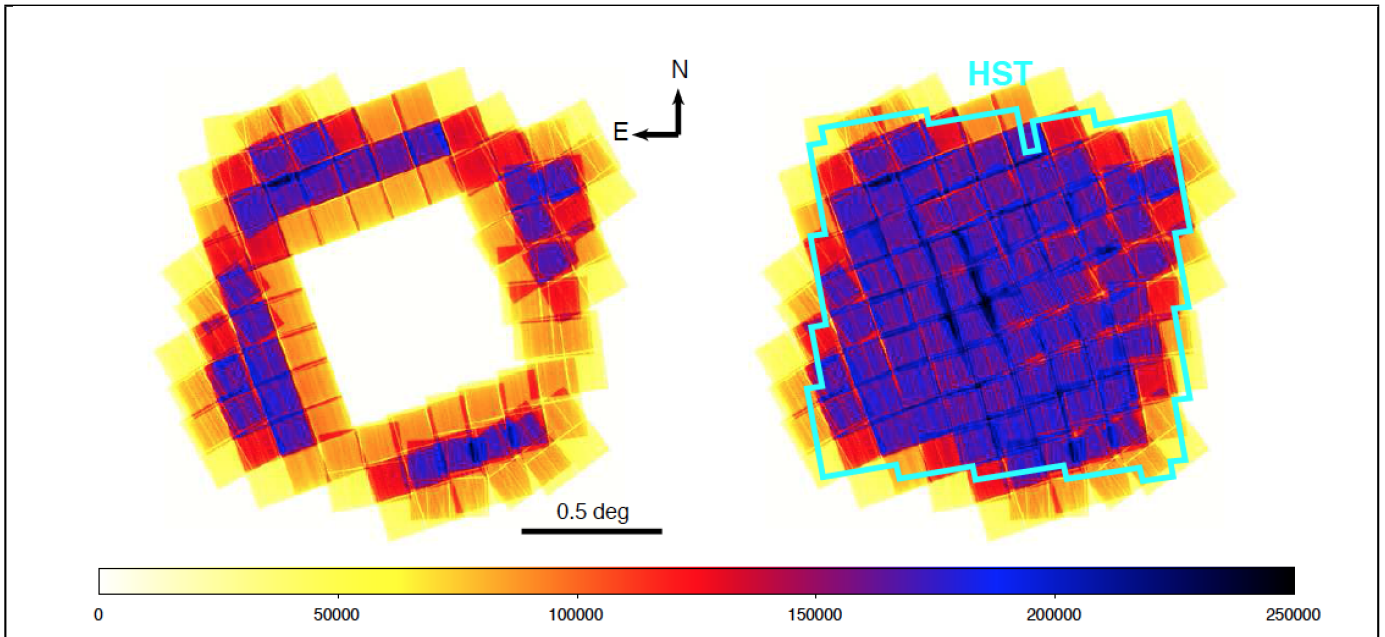
We then downloaded “stowed background” data from the *Chandra* archive<sup>33</sup>. Stowed background files are particle-only background files and are obtained when the ACIS detector is out of the focal plane. These files were then rescaled using the procedure described in Hickox & Markevitch (2006): we measured the ratio between the number of counts in our initial background ( $C_{data}$ ) and in the stowed image ( $C_{stow}$ ) in the energy range 9.5-12 keV. In this band, the effective area of *Chandra* is  $\simeq 0$

<sup>32</sup> The choice of such a spectral slope is not only because of consistency with E09 and P09, but it is also the slope of the cosmic X-ray background (e.g., Hickox & Markevitch 2006) and therefore well represents a mixed distribution of obscured and unobscured sources at the fluxes covered by *COSMOS-Legacy*.

<sup>31</sup> [http://cxc.cfa.harvard.edu/proposer/POG/html/chap5.html#tth\\_flg33](http://cxc.cfa.harvard.edu/proposer/POG/html/chap5.html#tth_flg33)

<sup>33</sup> <http://cxc.harvard.edu/ciao/threads/acisbackground/>





**Figure 4.** The mosaic of exposure maps for the new observations (left) and for the whole *COSMOS-Legacy* survey (right) in the full band. The color bar gives the achieved effective exposure in units of seconds. We reached a uniform coverage of  $\sim 160$  ks over the full HST area (cyan polygon).

and consequently all the counts have a non-astrophysical origin.

The stowed background, rescaled to our data by  $C_{data}/C_{stow}$ , was then subtracted from the initial background to obtain a first version of the cosmic X-ray background. The counts of this map were then renormalized using the exposure maps to create an exposure-corrected cosmic background.

Finally, we performed a Monte Carlo simulation using the exposure-corrected cosmic X-ray background and the stowed background as input files. We simulated 1000 images for each of the two backgrounds using the IDL routine *poidev* to obtain a Poissonian realization of each map, and then we obtained our final homogeneous background map adding together the two mean simulated images. In order to use these maps for sensitivity computations and in our detection algorithm, a Gaussian smoothing (with a scale of 20 pixels) was applied to this final background map using the FTOOL *fgauss*.

The distribution of the computed background (in counts/arcsec<sup>2</sup>) in the three bands is reported in Figure 6. The overall background count distribution is consistent with the one found in C-COSMOS (see Figure 4 of P09). In the full band the main peak is at around 0.13 counts/arcsec<sup>2</sup> and this corresponds to the deepest part of the exposure. In C-COSMOS, the deep and shallow areas were roughly the same size and therefore the background distribution had two clear peaks of approximately the same height, while in *COSMOS-Legacy*, the area with higher exposure is 3 times larger than the shallow area. This is represented in the background distribution as well. The number of background counts is consistent with the expectation for *Chandra* given the distribution of our exposure times.

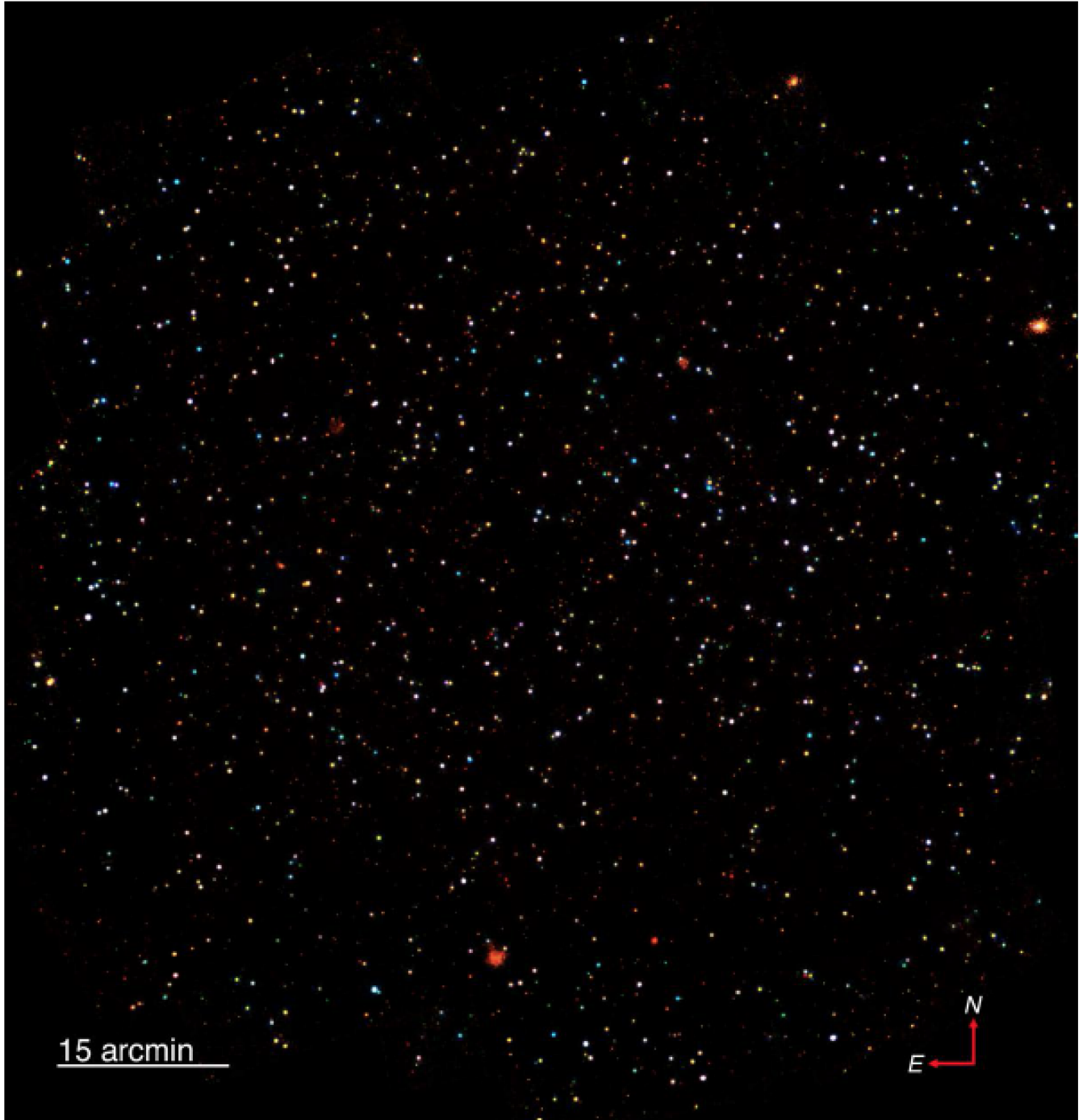
#### 4. DATA ANALYSIS: SOURCE DETECTION AND PHOTOMETRY

The analysis presented in the following focuses only on point sources. A parallel effort on the detection of extended sources will be presented by Finoguenov et al. (in preparation). To avoid contamination by extended sources (Finoguenov et al. 2007, Kettula et al. 2013) and visually inspected all the brightest ( $L_X > 10^{41}$  erg s<sup>-1</sup> in 0.5-2 keV) ones to check if a point source is detected inside them by *Chandra*.

Puccetti et al. (2009) extensively discussed and compared different source detection techniques concluding that the best procedure for C-COSMOS was a combination of *PWdetect* (Damiani et al. 1997) and the *Chandra Emldetect* (*CMLdetect*) Maximum Likelihood algorithm. As shown by P09 using extensive simulations, one of the strongest features of *PWdetect* is its ability to locate X-ray sources with extreme accuracy ( $0.02'' \pm 0.15''$ , P09 Table 1), while *CMLdetect* is the best tool to perform source photometry and derived source significance. The *COSMOS-Legacy* survey shares the same tiling layout, exposure time per field and roll angle range of C-COSMOS, hence, we can follow the P09 procedure and use the same significance threshold for source detection.

The original version of *CMLdetect*, called *emldetect* (Crudace et al. 1988, Hasinger et al. 1993), is part of the *XMM-Newton* SAS package and is based on a code originally developed for *ROSAT* data. *CMLdetect* has been adapted to run on *Chandra* data by replacing the *XMM-Newton* PSF library with the *Chandra* one (see Krumpel et al. 2015 for another application of *CMLdetect*). Moreover, this new tool can also work with different PSFs simultaneously.

*PWdetect* was developed to properly treat *Chandra* data with PSF varying across the field and it is based on the wavelet transform (WT) of the X-ray image. A WT is the convolution of an image with a “generating



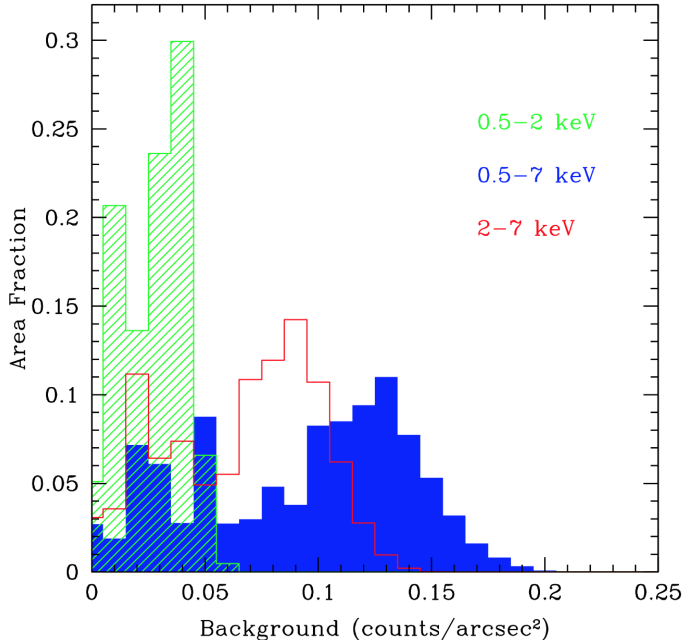
png

**Figure 5.** Three color image of the whole *COSMOS-Legacy* field (0.5-2.0 keV, 2.0-4.5 keV, and 4.5-7.0 keV as red green and blue, respectively).

wavelet” kernel which depends on position and length scale (a free parameter). For this survey, and for *Chandra* data in general, the length scale varies from  $0.5''$  to  $16''$  in steps of  $\sqrt{2}$ . These steps cover all possible *Chandra* PSFs (the largest are those at large off axis angle  $\theta_i$ ). Both radial and azimuthal PSF variations are accounted for by *PWDetect*, which first assumes a Gaussian PSF and then corrects by a PSF shape factor, calibrated with respect to source positions on the CCD.

*PWDetect* works on stacked observations only if co-

aligned (same aim point and roll-angle), as is the case for 11 of our fields which are observations split into multiple parts. Therefore, *PWDetect* was run on each of our new 56 fields setting the detection limit to  $3.8\sigma$  corresponding to a probability of a spurious detection to  $\simeq 10^{-4}$  with the aim of creating a large catalog of detections to be fed to *CMLDetect*. Also, given that the outer frame of C-COSMOS overlaps with the new survey, we run *PWDetect* on 20 old fields (fields 1-1 to 1-6, 1-6 to 6-6, 6-6 to 6-1 and last 6-1 to 1-1 as in Table 3 of E09).



**Figure 6.** Distributions of background counts per square arcsecond in the full (solid blue histogram), soft (shaded green histogram) and hard (empty red histogram) bands.

For overlapping regions between different pointings, we performed a positional cross correlation (using a  $2''$  radius) and if a source was detected in more than one field, we chose the position of the source at the smallest  $\theta_i$ , i.e. the one with the best PSF. We performed a visual inspection of all the sources having multiple matches within  $5''$ . About 90% of the pairs in the range  $2-5''$  were actually false detections, mainly caused by PSF tail detection of bright sources.

The positions obtained with `PWDetect` were then fed as input to `CMLDetect`, to obtain photometric information and significance for each source. We ran `CMLDetect` allowing the detection only of point-like sources. `PWDetect` can be used to obtain net counts, rates and fluxes, but we opted to use `CMLDetect` because it can work on a mosaic, while `PWDetect` cannot. Moreover, P09 has shown that `PWDetect` count rates are systematically less accurate than those of `CMLDetect` (the median ratio between the output detected and input simulated count rates ranges from 86 to 94% for `PWDetect` versus 97 to 105% for `CMLDetect`, independently from the energy). `CMLDetect` performs a simultaneous maximum likelihood PSF fitting for each input candidate source, previously obtained using `PWDetect`, to all images at each position and, working on a mosaic, can provide a refined position of the source and count rates. This procedure was run in three bands: full (0.5-7 keV), soft (0.5-2 keV) and hard (2-7 keV). With the goal of not missing close pairs, we run `CMLDetect` allowing to slightly change the input position provided by `PWDetect`.

The best-fit maximum likelihood parameter in `CMLDetect`, `DET_ML`, is related to the Poisson probability that a source candidate is a random fluctuation of the background ( $P_{random}$ ), as follows:

$$DET\_ML = -\ln(P_{random}). \quad (1)$$

As a consequence, sources with small values of `DET_ML` have high values of  $P_{random}$  and are then likely to be background fluctuations. We chose a threshold significance value of  $2 \times 10^{-5}$ , that corresponds to `DET_ML`=10.8, i.e., a source needs to have `DET_ML`>10.8 in at least one of the three bands to be included in the final catalog. This value is the same used in C-COSMOS and represents the best compromise between completeness and reliability as shown by P09 in Figure 11 and 12. 75% of the sources detected by `PWDetect` in a single field with `DET_ML`>10.8 and fed to `CMLDetect`, were found to be above the threshold in output.

To improve the final completeness of the catalog, we also search for less significant sources, up to about 100 times higher  $P$ , which corresponds to a threshold `DET_ML`=6. Similarly to what was done in C-COSMOS, sources with `DET_ML` in the range 6 to 10.8 and are only considered in this catalog if these have `DET_ML`>10.8 in another band. Sources with `DET_ML`<6 are considered undetected.

P09 performed extensive Monte Carlo simulations for C-COSMOS to both test the detection and photometry strategy as well as to determine the completeness and reliability of the source catalog at the chosen `DET_ML` threshold. Given that *COSMOS-Legacy* is the scaled up version of C-COSMOS (in area and exposure), the same analysis was followed, therefore we infer that the completeness and reliability of the catalog are the same. Therefore, the chosen `DET_ML` threshold implies a completeness of 87.5% and 68% for sources with at least 12 and 7 full band counts, of 98.2% and 83% for the soft band, 86% and 67% for the hard band. At this significance level and the same count limits, the reliability is  $\sim 99.7\%$  for the three bands.

#### 4.1. Point source catalog

##### 4.1.1. Source numbers

We positionally matched the three single-band, `CMLDetect` output catalogs (including all the sources to `DET_ML`=6) to one another using a cross-correlation radius of  $3''$ . We first matched the full band detected source catalog to the soft band one, then the full with the hard band catalog and finally the soft and hard band one. We performed a visual inspection of the whole sample, and we also made use of the catalog of optical/IR identifications (presented in a companion paper, Marchesi et al. in press), to solve ambiguous cases. After the visual inspection, we found that  $<1\%$  of the matches are actually fake associations and all related to sources at the outer edges of the survey with rather wide point spread function, therefore with large positional error. Overall, the mean (median) separation between detections of the same source in two different bands is  $0.43''$  ( $0.23''$ ) for full to soft and  $0.41''$  ( $0.23''$ ) for full to hard, with 90% of the matches within  $1''$ . For soft to hard matches, the mean (median) separation is instead  $0.73''$  ( $0.56''$ ), with  $\sim 80\%$  within  $1''$ . The source position is determined in the full band for all the sources detected in the full band; if a source is not detected the full band, the soft band position is used. The hard band position is used for sources

detected in the hard band only.

In Table 2, we report the total number of new sources for each combination of bands, while in Table 4 we report the number of sources detected in each band at the two adopted thresholds ( $\text{DET\_ML} > 10.8$  and  $6 < \text{DET\_ML} < 10.8$ ). The number of detections with  $\text{DET\_ML} > 10.8$  in at least one of three X-ray bands is 2273. The number of expected spurious sources with  $\text{DET\_ML} > 10.8$  is reported in each band for two count limits in Table 3.

In the area where the new data overlap with the outer C-COSMOS frame, the exposure time is now double the previous mean exposure time (142 ks versus 72 ks), and 385 new sources are detected in addition to the 694 sources already in E09. For the last 694 sources with doubled exposure time, 676 have been detected in the new data as well. The eighteen C-COSMOS sources not detected in the new data had  $\text{DET\_ML}$  values in E09 in the three bands close to the threshold ( $\text{DET\_ML} < 15$ ); moreover, 10 of them were detected only in 2 out of 3 bands in E09 and the remaining eight were detected only in 1 band.

In Table 2, we include the number of sources in each combination of bands for the C-COSMOS area including the new data and also in parentheses the number of sources as in E09. The same old and new numbers are included in Table 4. In this paper we provide also an updated catalog of the C-COSMOS sources with larger exposure in the total data. Among the 676 C-COSMOS sources with new data, only  $\sim 1.5\%$ ,  $\sim 2\%$  and  $\sim 3\%$  in the full, soft and hard band, respectively, have a  $\text{DET\_ML}$  value which is below the threshold in the combined data while it was above the 10.8  $\text{DET\_ML}$  threshold in the C-COSMOS catalog, confirming the reliability of the detection method, and the consistency between the analysis performed in E09 and P09 and the one performed here. The actual fraction of sources with  $\text{DET\_ML}$  lower in the combined dataset than in C-COSMOS is 14% in the full and 10% in the soft and hard bands. On average, sources with lower  $\text{DET\_ML}$  in the new dataset are in an area of the field where the ratio  $(\text{exp\_new} - \text{exp\_old})/\text{exp\_old}$  is 40% lower than the average ratio of the sources in the catalog, therefore the discrepancy could be explained with source variability.

The total number of sources summing the two datasets is reported in the last column of Table 2. Adding the new observations, we more than double the sample with respect to C-COSMOS, obtaining a catalog of 4016 sources, the largest sample of X-ray sources homogeneously detected and with uniform multiwavelength data (see Section 7 for a discussion and Marchesi et al. in press). In comparison, other contiguous surveys with similar area in the literature have about 20% fewer sources than *COSMOS-Legacy* (see 3362 sources in Stripe 82 by LaMassa et al. 2013a,b and 2015; 3293 in X-Bootes by Murray et al. 2005; 2976 in X-DEEP2 by Goulding et al. 2012).

In Figure 7, we show the signal-to-noise ratio ( $\text{SNR} = \text{count rate}/\text{count rate error}$ ) as a function of the  $\text{DET\_ML}$  for the new sources with  $\text{DET\_ML} > 10.8$ . In excellent agreement with the finding in C-COSMOS, the  $\text{SNR}$  increases smoothly with increasing  $\text{DET\_ML}$ , with a dispersion of a factor of 2 at both low and high  $\text{DET\_ML}$  values.

#### 4.1.2. Source positional errors

To compute the positional errors associated with the X-ray centroids given in the catalog ( $\sqrt{\sigma_{R.A.}^2 + \sigma_{dec}^2}$ ), we followed the prescription of P09 defining  $\text{err\_pos} = r_{PSF}/\sqrt{S}$ , where  $S$  is the number of net (i.e. background subtracted) source counts in the full band in a circular region of radius  $r_{PSF}$  containing 50% of the encircled energy in the observation where the source is at the smallest off-axis angle. The positional errors are generally in very good agreement with those resulting from *CMLDetect*. In Figure 8, the positional error distribution is presented for all the new sources (black solid), the old C-COSMOS sources (red dashed) and the updated C-COSMOS distribution (blue dotted). The sources plotted in the lowest bin are those with positional error values actually smaller than  $0.1''$ , which we set to  $0.1''$ , consistently with the work done in P09. These sources with small positional error are just very bright objects (with  $\sim 240$  mean full band counts; see next section).

The peak of the new sources distribution is  $\sim 0.6''$  and 85% of the sources have a positional error  $< 1''$ , while C-COSMOS source distributions peaks at around  $0.4''$ . This difference (the somewhat larger positional errors for the sources detected with the new data than for those detected in C-COSMOS) is due to the fact that, as shown in Fig. 9, the net counts distribution for the sources in the new data peaks at a lower value than for the C-COSMOS sources (therefore giving a smaller denominator in the formula of the positional error).

#### 4.1.3. Source counts and fluxes

The count rates in three bands reported here were obtained with *CMLDetect*. Vignetting and quantum efficiency were taken into account when measuring the effective exposure time. The count rate error at 68% confidence level was computed using the equation

$\text{err\_rate} = \frac{\sqrt{C_S \cdot 90\% + (1+a) \times B_{90\%}}}{0.9 \times T}$ , where  $C_S$  are the source net counts estimated by aperture photometry, using, for each observation where the source was detected, an extraction radius including 90% of the EEF;  $B$  are the background counts estimated in the same aperture on the background maps used in *CMLDetect* and corrected with a factor  $a=0.5$ , introduced to account for the uncertainties on the background estimation in a given position (see P09);  $T$  is the vignetting corrected exposure time.

In Figure 9, the net count distributions for the new sources in three bands are compared to those in E09 (C-COSMOS old) and also to the updated counts distribution of C-COSMOS (C-COSMOS new). The total is the sum of the new detections plus the updated C-COSMOS. The median (mean) value of net counts in the whole dataset in full, soft and hard bands is 30, 20 and 22 (80, 60, 43), respectively, compared to C-COSMOS where we had 33, 22 and 23 (88, 65, 46). The total number of net counts for the 676 C-COSMOS sources also detected in the new dataset is on average 60-80% larger than the number of counts in C-COSMOS only. As a consequence, the updated C-COSMOS count histograms in Figure 9 are all shifted to a higher numbers of counts. While in the full band the peak of the distribution is still around 30 counts, we more than double the number of sources with more than 70 full band counts, for which

it is possible to perform individual X-ray spectral analysis, from 390 (Lanzuisi et al. 2013) to  $\sim 950$  sources in *COSMOS-Legacy*.

The fluxes were obtained from the count rates using the relation  $F=R \times (CF \times 10^{-11})$ , where  $R$  is the count rate in each band and  $CF$  is the energy conversion factor computed using the CIAO tool *srcflux*, assuming a power-law spectrum with slope  $\Gamma=1.4$  and a Galactic column density  $N_H=2.6 \times 10^{20} \text{ cm}^{-2}$ . Due to the fact that the observations have been taken in two different *Chandra* cycles, i.e. Cycle 8 for C-COSMOS and Cycle 14 for the new data, we used as  $CF$  a weighted mean of the factors in the different cycles, depending on the exposure time for each source accumulated in each cycle, to take into account for its variation ( $\sim 15\%$  between the two cycles). The Cycle 14 (Cycle 8<sup>34</sup>)  $CF$  are 1.71 (1.57), 7.40 (6.34) and 3.06 (3.04) counts  $\text{erg}^{-1} \text{ cm}^2$  for 0.5-10, 0.5-2 and 2-10 bands, respectively. The conversion factors are sensitive to the assumed spectral shape: for  $\Gamma=2$ , there is a change of 40% in the full band  $CF$ , of  $\sim 5\%$  in the soft band and of  $\sim 20\%$  in the hard band.

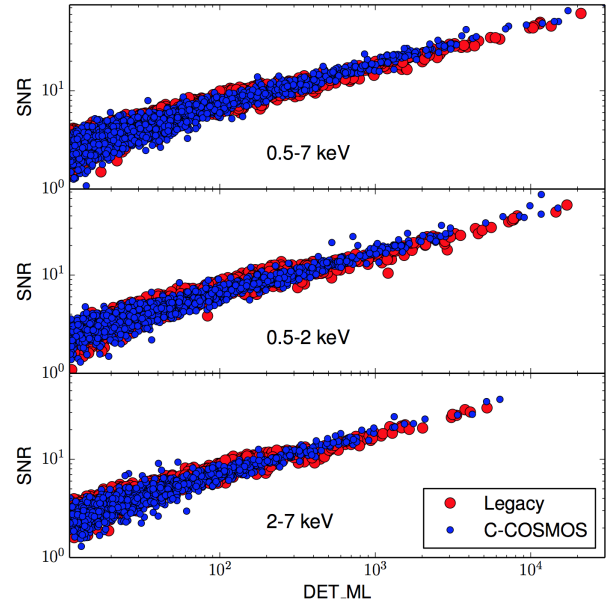
For the 676 C-COSMOS sources detected in the new data as well, we computed new total X-ray fluxes. In Figure 10 the normalized distribution of ratios between total and old fluxes are plotted for the three bands. From Gaussian fitting of the distributions, we find centroids at  $(F_{new}/F_{old}) = 1.06, 1.11, 0.99$  and standard deviations of  $\sim 0.50, \sim 0.55$  and  $\sim 0.40$  in full, soft and hard band, respectively, showing a good agreement between old and new fluxes. The distributions show wings to both negative and positive values. Malmquist bias is most likely responsible for the negative wing, while variability for the positive one.

The distributions of X-ray fluxes for the whole *Chandra* COSMOS Legacy survey in the full, soft and hard bands is shown in Figure 11, where it is also compared with C-COSMOS (the new version with just the updated fluxes, given the excellent agreement) and XMM-COSMOS. The new survey is about  $\sim 2.5$  times deeper than XMM-COSMOS in the 0.5-2 keV band and  $\sim 2$  times in the 2-10 keV band, and more than doubles the number of C-COSMOS sources in the same flux range. In the same Figure we compare our data with the 4 Ms CDFS (Xue et al. 2011) and the large area Stripe82 survey (LaMassa et al. 2013a,b) source flux distributions, respectively to the left and to the right of *COSMOS-Legacy* flux distribution. The combination of the three surveys (the deepest, the intermediate and among the widest, see also Section 7) allows to cover more than 4 orders of magnitude in flux.

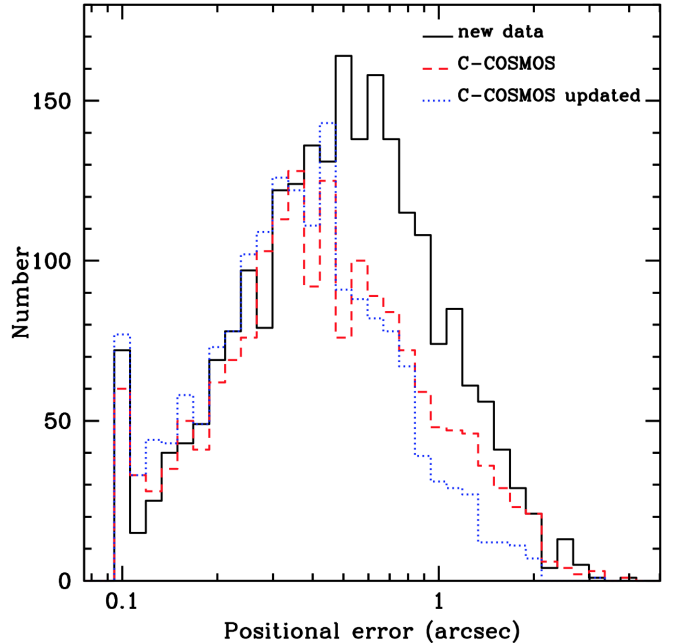
Upper limits (90% confidence level) on net counts, count rates and fluxes are given for all sources found in one band but not detected in another band. The upper limits were computed with the same procedure adopted for C-COSMOS and largely described in P09, to which we refer for a complete description.

#### 4.1.4. Hardness ratio analysis

<sup>34</sup> The  $CF$  used for C-COSMOS and reported in E09 and P09 (computed using the online tool PIMMS) slightly differ from the one used here cause the latter are now computed with the most updated response matrix. The difference is  $\sim 15\%$  and it reflects on the final fluxes for all C-COSMOS sources.

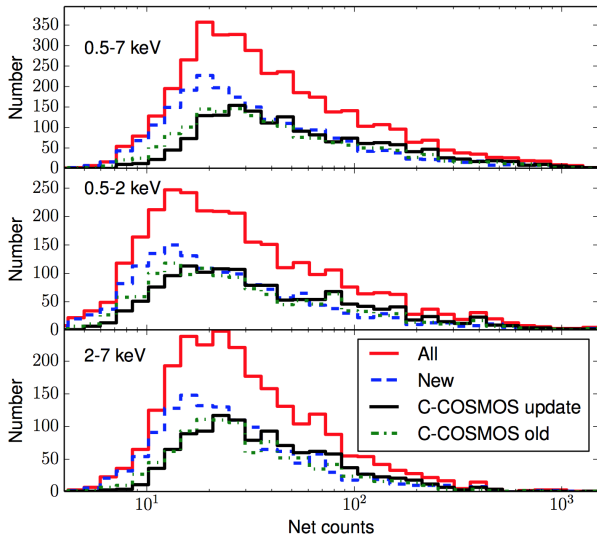


**Figure 7.** Signal-to-noise ratio as a function of DETML for sources detected in three bands. The new *Chandra* sources are plotted as red circles, the C-COSMOS sources as blue ones. We plot only sources with  $\text{DET.ML} > 10.8$ .

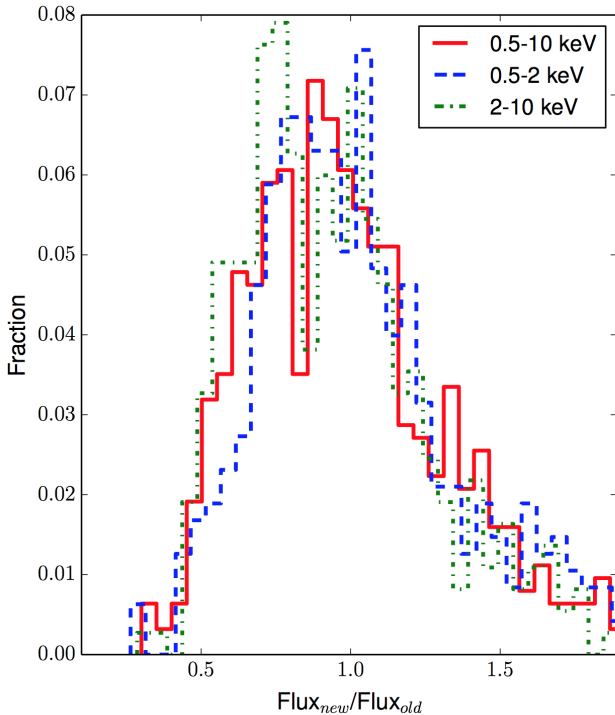


**Figure 8.** Positional error distribution for the new *COSMOS-Legacy* data (black solid line), the original C-COSMOS (red dashed line) and the updated C-COSMOS (blue dotted line).

In order to provide a rough estimate of the X-ray spectral shape of the sources, in particular of the intrinsic obscuration (see Marchesi et al. in press), for all the sources in the catalog, including the C-COSMOS sources, we computed the hardness ratio defined as  $HR = \frac{H-S}{H+S}$ , where  $H$  are the net counts in the hard band and  $S$  are

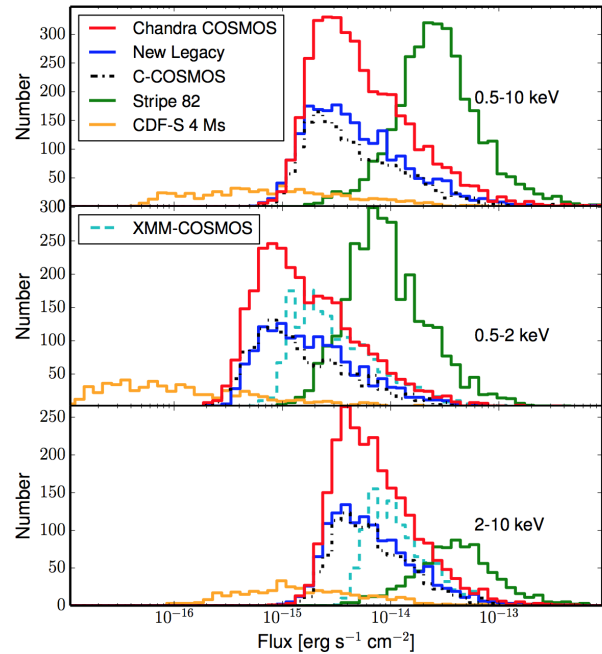


**Figure 9.** Source count distributions in three bands: 0.5-7 keV (top), 0.5-2 keV (center) and 2-7 keV (bottom) for *COSMOS-Legacy* (solid red), new data only (blue dashed) C-COSMOS old (green dot-dashed) and updated (black solid). Sources with upper limit have not been included.



**Figure 10.** Normalized distributions of ratios between new and old fluxes of the 676 sources detected in C-COSMOS and also in the new data at  $\text{DET\_ML} > 10.8$ . Sources with upper limit have not been included.

those obtained in the soft band. Given the low number of counts for most of the sources (see Figure 9), we used BEHR (Bayesian Estimation of Hardness Ratios, Park et al. 2006) which is particular effective in the low count



**Figure 11.** Flux distributions for sources detected in 0.5-10 keV (top), 0.5-2 keV (center) and 2-10 keV (bottom) bands for *COSMOS-Legacy* (solid red), new data only (blue solid), C-COSMOS updated (black dotted) and XMM-COSMOS (cyan dashed) sources. We also include the CDFS 4 Ms source flux distribution (orange) and the Stripe82 sources (green). Sources with upper limit have not been included.

regime, not needing a detection in both bands to work.

We extracted aperture photometry counts from each observation where the source was detected, using the PSF radius at encircled energy fraction (EEF)=0.9. We also extracted the background counts from the same observations, using an annulus with  $r_{min} = r_{PSF} + 8$  pixels and  $r_{max} = r_{PSF} + 40$  pixels, where  $r_{PSF}$  is the PSF radius at encircled energy fraction (EEF)=0.95 (in pixel). In the background extraction, we excluded the contamination by other nearby detected sources using an exclusion radius equal to  $r_{PSF}$ . Total counts, background counts and the ratio between the sum of background areas and the sum of source areas, both in soft and hard bands, were then fed as input parameters to BEHR.

For most sources (>3000) BEHR finds a detection on the HR, and for 989 sources an upper or lower limit (616 and 371 sources respectively). The typical error on the HR is  $\sim 0.2$ . In Figure 12, we plot the distribution of the HRs for the measured values (black solid line), for the lower limits (red) and the upper limits (blue). The mean (median) HR value is -0.09 (-0.17) for the measured values and it moves to lower values when including upper and lower limits (-0.11 and -0.19 for the mean and the median, respectively). A Gaussian fit returns a peak at -0.20 with a  $1\sigma$  dispersion of 0.32, however a single Gaussian is not clearly a best representation of the HR distribution. A double Gaussian fit returns a peak at -0.31 and one at 0.12 with a  $1\sigma$  dispersion of 0.18 and 0.38, respectively.

Hardness ratio is not a fully reliable measurement of

Bands	New	C-COSMOS	Legacy
F+S+H	1140	1047 (922)	2187
F+S	536	397 (474)	933
F+H	448	231 (257)	679
F	121	49 (73)	170
S	21	17 (32)	38
H	7	2 (3)	9
<b>Total</b>	<b>2273</b>	<b>1743 (1761)</b>	<b>4016</b>

**Table 2**

Number of sources with  $\text{DET\_ML} > 10.8$  in at least one band, for each combination of X-ray bands. The “New” column includes all the new detected sources. The columns labelled as “C-COSMOS” include the updated numbers, using the information from the new data and in parenthesis also the old numbers as in Elvis et al. (2009).

Bands	New		C-COSMOS		Legacy	
	> 7	> 12	> 7	> 12	> 7	> 12
F	5	5	6	6	12	11
S	4	3	4	3	9	7
H	3	3	4	3	8	7

**Table 3**

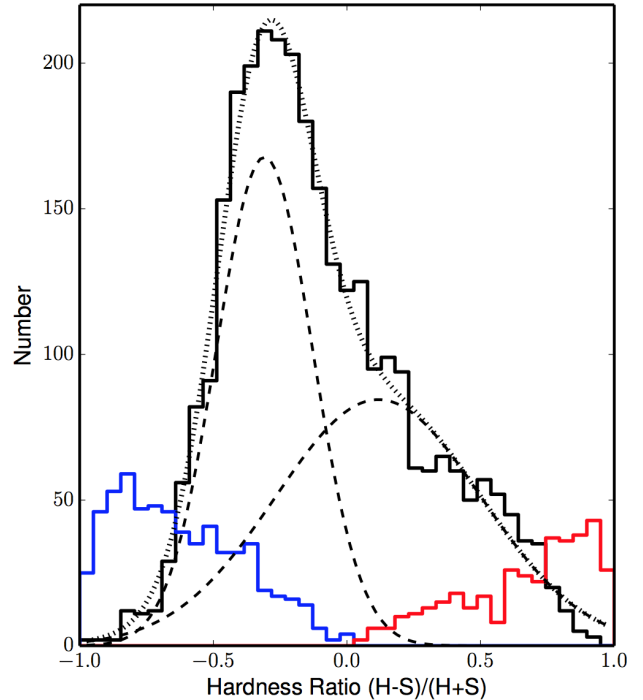
Number of spurious sources with  $\text{DET\_ML} > 10.8$  with at least 12 and 7 full band counts, corresponding to a reliability of 99.7% for the new data, the old C-COSMOS data (as in P09 Section5) and in the whole *COSMOS-Legacy*.

obscuration, because of the complexity of the spectral shape, the large error bars due to low counts statistic and the redshift dependency (see Marchesi et al. in press), however it is possible to roughly assume an HR value to divide the sources in obscured and unobscured. We use here  $\text{HR} = -0.2$ , which has been shown to be a fair value to separate sources with column densities above and below  $10^{22} \text{ cm}^{-2}$  (Lanzuisi et al. 2013, Civano et al. 2012) at all redshifts. A total of 1993 sources,  $50^{+17}_{-16}\%$  of the entire sample (errors have been computed using HR  $1\sigma$  errors) are therefore classified as obscured. Tentatively, the double Gaussian fit of the HR distribution could also be interpreted as to be due from two populations of sources, the obscured population peaking at positive HRs and the unobscured population peaking at negative HR. The broad dispersion of the Gaussian peaking at positive HR could be due to high redshift obscured sources whose HR would be negative even if obscured. A more detailed analysis on the obscured AGN fraction is presented in Marchesi et al. (in press).

#### 4.1.5. Source catalog

The catalog released with this paper contains all the measurements discussed above. In Table 5, we show the columns of the catalog of the new 2273 sources (named as “lid” in column 1) combined with the updated C-COSMOS catalog of 1743 sources (named as “cid” in column 1). The catalog will also be stored in the COSMOS web site at the *COSMOS-Legacy* project page<sup>35</sup>. Data products, including exposure and events mosaics, are available in the dedicated page<sup>36</sup> at the same website.

#### 4.2. Matching with XMM-COSMOS catalog



**Figure 12.** HR distributions for the whole sample (black), upper limits (blue) and lower limits (red). The dotted line is the sum of the double Gaussian fitting. The dashed lines are the two Gaussian resulting from the fitting.

We matched the *COSMOS-Legacy* sources with those in XMM-COSMOS (Cappelluti et al. 2009). There are 1714 secure XMM-COSMOS sources with at least one counterpart in *COSMOS-Legacy*, 824 of which have at least one counterpart in the new data. There are 46 XMM-COSMOS sources outside the area covered by *COSMOS-Legacy* (see Fig. 1) and 126 with no *Chandra* counterparts. In summary, 93% of the XMM-COSMOS sources within the *COSMOS-Legacy* area have at least one *Chandra* counterpart.

The 126 sources with no *Chandra* counterparts can be divided in three groups: the 25 sources (20%) with *Chandra* exposure  $< 40$  ks; the 60 sources (48%; 13 of these sources have also *Chandra* exposure  $< 40$  ks) with XMM-COSMOS  $\text{DET\_ML} < 15$  in all of the three bands (0.5-2 keV, 2-8 keV, 4.5-8 keV); last, the 54 sources with XMM-COSMOS  $\text{DET\_ML} > 15$  in at least one band and *Chandra* exposure  $> 40$  ks. For the first group, the low exposure time could be the reason of the non detection, while for the second a non-detection in *Chandra* can be explained with a flux fluctuation within the flux uncertainty. We visually inspected the sources in the last group and we found that seven of them are located inside a bright cluster, and therefore have been not resolved into point sources by our analysis. For the remaining 47 sources the *Chandra* signal is weak or negligible, and therefore these sources could be candidate variable AGN. In particular, XMM-ID 30748 has  $\text{DET\_ML}$  20 times larger than the detection threshold in XMM-COSMOS: this source was detected only in the 0.5-2 keV band, with a flux  $F = 2.7 \times 10^{-15} \text{ erg s}^{-1} \text{ cm}^{-2}$  and a photometric red-

<sup>35</sup> <http://irsa.ipac.caltech.edu/data/COSMOS/tables/chandra/>

<sup>36</sup> <http://irsa.ipac.caltech.edu/data/COSMOS/images/chandra/>

Band	DET_ML $\geq$ 10.8			6 < DET_ML < 10.8		
	New	C-COSMOS	Legacy	New	C-COSMOS	Legacy
Full (F)	2146	1667 (1655)	3813	99	57 (71)	156
Soft (S)	1538	1382 (1340)	2920	159	79 (88)	238
Hard (H)	1325	1115 (1017)	2440	271	165 (165)	436

**Table 4**

Number of sources detected in each band at the two adopted thresholds. The columns labelled as C-COSMOS include the updated numbers using the information from the new data and, in parenthesis, also the old numbers as in Elvis et al. (2009).

No.	Field	Note
1	Name	<i>Chandra</i> source name
2	R.A.	<i>Chandra</i> Right Ascension (J2000, hms)
3	DEC	<i>Chandra</i> Declination (J2000, dms)
4	pos_err	Positional error [arcsec]
5	DET_ML_F	maximum likelihood detection value in 0.5-7 keV band
6	rate_F	0.5-7 keV count rate [counts s <sup>-1</sup> ]
7	rate_F_err	0.5-7 keV count rate error [counts s <sup>-1</sup> ]
8	flux_F	0.5-10 keV flux [erg cm <sup>-2</sup> s <sup>-1</sup> ]
9	flux_F_err	0.5-10 keV flux error [erg cm <sup>-2</sup> s <sup>-1</sup> ]
10	snr_F	0.5-7 keV S/N Ratio
11	exptime_F	0.5-7 keV exposure time [ks]
12	cts_ap_F	0.5-7 aperture photometry counts [counts]
13	cts_ap_F_err	0.5-7 aperture photometry counts error [counts]
14	DET_ML_S	maximum likelihood detection value in 0.5-2 keV band
15	rate_S	0.5-2 keV count rate [counts s <sup>-1</sup> ]
16	rate_S_err	0.5-2 keV count rate error [counts s <sup>-1</sup> ]
17	flux_S	0.5-2 keV flux [erg cm <sup>-2</sup> s <sup>-1</sup> ]
18	flux_S_err	0.5-2 keV flux error [erg cm <sup>-2</sup> s <sup>-1</sup> ]
19	snr_S	0.5-2 keV S/N Ratio
20	exptime_S	0.5-2 keV exposure time [ks]
21	cts_ap_S	0.5-2 aperture photometry counts [counts]
22	cts_ap_S_err	0.5-2 aperture photometry counts error [counts]
23	DET_ML_H	maximum likelihood detection value in 2-7 keV band
24	rate_H	2-7 keV count rate [counts s <sup>-1</sup> ]
25	rate_H_err	2-7 keV count rate error [counts s <sup>-1</sup> ]
26	flux_H	2-10 keV flux [erg cm <sup>-2</sup> s <sup>-1</sup> ]
27	flux_H_err	2-10 keV flux error [erg cm <sup>-2</sup> s <sup>-1</sup> ]
28	snr_H	2-7 keV S/N Ratio
29	exptime_H	2-7 keV exposure time [ks]
30	cts_ap_H	2-7 aperture photometry counts [counts]
31	cts_ap_H_err	2-7 aperture photometry counts error [counts]
32	hr	Hardness ratio
33	hr_lo_lim	Hardness ratio 90% lower limit
34	hr_up_lim	Hardness ratio 90% upper limit

**Table 5**

Data fields in the catalog.

shift  $z=2.71$ . Despite being interesting and worth further analysis on the variability, this is beyond the scope of this paper.

There are 58 XMM-COSMOS sources that have been resolved by the smaller *Chandra* PSF into two distinct sources using a maximum radius of 10'' for the match. Two XMM-COSMOS sources have been resolved into three *Chandra* sources using a maximum radius of 10''. As a comparison, 25 XMM-COSMOS sources (Brusa et al. 2010) were resolved into two separate C-COSMOS sources. More details on the optical counterparts of the XMM-COSMOS sources resolved in two *Chandra* ones are given in Marchesi et al. (in press).

There is a good agreement between XMM-COSMOS and *Chandra* fluxes. We rescaled the *Chandra COSMOS-Legacy* fluxes using the same slope used for XMM-COSMOS ( $\Gamma=2$  in soft band and  $\Gamma=1.7$  in hard band) and we found that the median value of the ratio  $flux_{XMM}/flux_{Chandra}$  is 1.13 in soft band and 1.22 in hard band.

## 5. SKY COVERAGE AND SURVEY SENSITIVITY

The sky coverage of a survey is the area covered as a function of the flux limit. We computed it in three bands (0.5-10, 0.5-2 and 2-10 keV) using the exposure and background maps (see Section 3.2 and 3.3) produced for the source detection, and assuming a power-law spectrum

with  $\Gamma=1.4$  and Galactic  $N_H=2.6\times 10^{20}$  cm<sup>-2</sup>. X-ray observations have a flux limit that changes over the field of view because the *Chandra* PSF changes in both size and shape as a function of the distance from the aim point and because the effective area is vignettted. In this survey, where the total coverage is obtained using multiple overlapping pointings, every source was observed in up to six different positions on the detector, resulting in a quite uniform average PSF (Figure 2).

The procedure we used to compute *COSMOS-Legacy* survey sky coverage is closely similar to that used by P09 for C-COSMOS, but makes use of a PSF map for each observation instead of an analytical form of the PSF as function of the off-axis angle. This is a more time consuming approach but one that returns a more detailed sensitivity map, which can be valuable in other studies (e.g., clustering analysis and correlation functions) or simply for source photometry (Section 4.1.4).

For each observation we made use of the CIAO tools `mkpsfmap` and `dmimgadapt` to create a background map convolved with the PSF map in such a way that at each position of the map, the count value corresponds to the number of counts in an aperture corresponding to 50% of the encircled energy fraction at that position.

For each position of the entire mosaic (applying a binning of 8 pixels for computing time purposes), we computed the minimum number of counts  $C_{min}$  needed to



exceed the background fluctuations, assuming the same probability for spurious sources (i.e., DET\_ML threshold) used in the C-COSMOS and *COSMOS-Legacy* catalogs for the Poisson statistics, i.e.  $2 \times 10^{-5}$ . We used the relation

$$P_{Poisson} = e^{-B} \sum_{k=C_{min}}^{\infty} \frac{B^k}{k!} = 2 \times 10^{-5}, \quad (2)$$

where  $B$  is the total background counts computed at each position of the grid, by summing the background counts in each observation covering that given position. Equation 2 is solved iteratively to find  $C_{min}$ ; then the count rate limit,  $R_{lim}$  is obtained using

$$R_{lim} = \frac{C_{min} - B}{f_{psf} \times T_{exp}}, \quad (3)$$

where  $T_{exp}$  is the total, vignetting corrected, exposure time at each position on the grid, while  $f_{psf}$  is the encircled count fraction of the PSF. In C-COSMOS, this value was tuned to reproduce the simulation results and then it was fixed to  $f_{psf}=0.5$ , however any number in the range 0.5-0.9 produced similar results with variations of the order of few percent in the resulting sensitivity.

Finally, we converted the count rate limit  $R_{lim}$  into the flux limit using the same conversion factors used for the sources in the catalog based on the position (see Section 4.1.3). We also computed the sensitivity for only the C-COSMOS area with the same method and obtained the same sensitivity as published in E09 and P09.

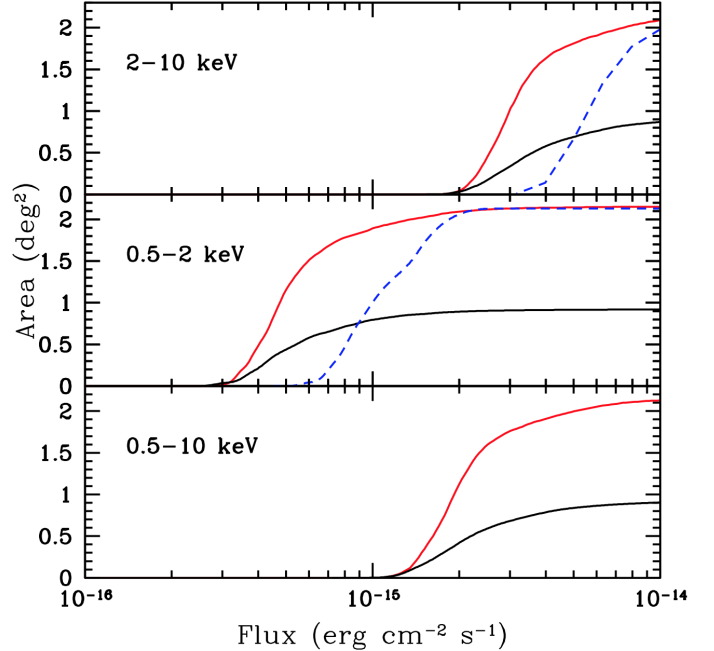
The sky coverage of the *Chandra* COSMOS Legacy survey in the three energy bands is shown in Figure 13. We compare our results with those of C-COSMOS (black solid lines) and XMM-COSMOS (blue dashed lines): the new survey covers a similar area to XMM-COSMOS and almost three times the area of C-COSMOS at faint fluxes (e.g.,  $\sim 5 \times 10^{-15}$  erg cm $^{-2}$  s $^{-1}$  in the soft band) and  $\sim 2$  times at bright fluxes (e.g.,  $> 10^{-15}$  erg cm $^{-2}$  s $^{-1}$  in the soft band).

We have verified that the limits at 20% (50%) completeness for the Legacy catalog are consistent with those computed and reported in Table 2 of P09 and assuming the changes in CF used here and explained in Section 4.1.3 of  $1.5$  ( $1.9$ )  $\times 10^{-15}$ ,  $3.9$  ( $4.9$ )  $\times 10^{-16}$  and  $2.5$  ( $3.1$ )  $\times 10^{-15}$  erg cm $^{-2}$  s $^{-1}$  in the F, S, and H bands. At this limit, *COSMOS-Legacy* increases by a factor of 3 the area covered with respect to C-COSMOS.

## 6. NUMBER COUNTS

The  $\log N$ - $\log S$  relation, i.e. the number of sources  $N(> S)$  per square degree detected at fluxes brighter than a given flux  $S$  (erg s $^{-1}$ cm $^{-2}$ ), provides a first estimate of source space density as a function of flux and therefore information on the cosmic population to compare with different models of population synthesis. Given that multiple  $\log N$ - $\log S$  curves have been published in the literature, it is also a standard check to validate the many calibration steps used to produce a catalog of X-ray point like sources.

We constructed the  $\log N$ - $\log S$  curve for *COSMOS-Legacy* in both the 0.5-2 keV and 2-10 keV bands. Following P09, we included only sources with DET\_ML $>10.8$  and we applied a cut in SNR ( $> 2$  and  $>$



**Figure 13.** Area-flux curve for *COSMOS-Legacy* (red solid line) in 2-10 keV (top), 0.5-2 keV (center) and 0.5-10 keV (bottom) bands. The coverage of C-COSMOS (black solid line) and XMM-COSMOS in the 0.5-2 keV and 2-10 keV bands (Cappelluti et al. 2009; dashed blue line) are shown for comparison.

2.5 in soft and hard) to limit the Eddington bias effect, which could have a significant (up to 30-50%) contribution at the lowest fluxes. This choice avoids sources with large statistical uncertainties on their fluxes and limits the errors due to the sky coverage uncertainties at the faint end. With the adopted thresholds in SNR, the agreement measured in P09 between simulations input and output  $\log N$ - $\log S$  is better than 5%. The procedure used by P09 is consistent with the one applied by Luo et al. (2008) on *Chandra* Deep Field South data. The number of sources not included because of the SNR cut is  $\sim 1\%$  in the soft and  $\sim 5\%$  in the hard band.

The adopted SNRs imply the following flux limits:  $2.7 \times 10^{-16}$  erg s $^{-1}$  cm $^{-2}$  in the 0.5-2 keV band and  $1.8 \times 10^{-15}$  erg s $^{-1}$  cm $^{-2}$  in the 2-10 keV band. These are the same flux limits of C-COSMOS, which is expected given that the new observations have the same maximum exposure. The final number of sources used here for the number counts with the above constraints are 2758 in the soft band (1309 from C-COSMOS and 1449 from the new sample) and 2243 in the hard band<sup>37</sup> (1056 from C-COSMOS and 1187 from the new sample).

We show the results obtained with these source selections in Figure 14 (top panels): the normalized Euclidean curves, i.e. with  $N(> S)$  multiplied by  $S^{1.5}$ , are presented in order to enhance the differences between different surveys. In the same figure we include the C-COSMOS (E09) and XMM-COSMOS points (Cappelluti et al. 2009). We also compare our  $\log N$ - $\log S$  relationships with those from previous X-ray surveys, span-

<sup>37</sup> We also applied a cut in exposure time at 40 ks in the hard band to limit sources (65 in total) at the edges of the field with high background level.

ning from wide (Stripe82 XMM: LaMassa et al. 2013b; 2XMMi: Mateos et al. 2008), to moderate (XDEEP2: Goulding et al. 2012), to small areas (4 Ms CDFS: Lehmer et al. 2012). As XDEEP2 and CDFS define their hard band in a slightly different energy range, we converted their energy to 2-10 keV to perform an adequate comparison.

*COSMOS-Legacy*  $\log N$ - $\log S$  covers 3 and 2.5 orders of magnitude in flux in the soft and hard band, respectively, with 2-8% errors at fluxes  $< 1 - 3 \times 10^{-14}$  erg cm $^{-2}$  s $^{-1}$ , respectively. The excellent statistics allows to considerably reduce the uncertainties (20-30%) in the number counts also at bright fluxes, which are now  $\simeq 40\%$  smaller than in C-COSMOS.

In the soft band, there is an excellent agreement between our survey and previous works below  $S \sim 10^{-14}$  erg cm $^{-2}$  s $^{-1}$ . At brighter fluxes instead, the uncertainties are larger due to the low number of detections (65 sources in *COSMOS-Legacy*). A larger spread is observed when comparing results from different surveys, due to the fact that bright sources can be properly sampled only with extremely large areas ( $> 5$ -10 deg $^2$ ). In the hard band instead, *COSMOS-Legacy* number counts agree with other surveys at faint fluxes, while at the bright end (i.e.  $S > 2 \times 10^{-14}$  erg cm $^{-2}$  s $^{-1}$ ), the *COSMOS-Legacy* counts are in the upper envelope of the spread.

We also compare our results with predictions of two different phenomenological models, Gilli et al. (2007) and Treister et al. (2009), assuming column densities in the interval  $N_H = 10^{20-26}$  cm $^{-2}$  and redshift  $z = 0-6$ . In Fig. 14 (bottom panels), we show the ratio of *COSMOS-Legacy* number counts to both models in the soft and hard bands (left and right). At the faint end of the soft band, i.e., up to fluxes  $\sim 10^{-14}$  erg cm $^{-2}$  s $^{-1}$ , our results are in agreement with the Gilli et al. (2007, solid points) model prediction within 1-5%, while the Treister et al. (2009, open points) model (open points) slightly under-predicts the counts by 5-10% in the same flux range. At bright fluxes, where the sample is limited by the statistics, the differences between models and data becomes larger, even exceeding 10%. In the hard band, both models reproduce well the observed data within 5% below  $> 2 \times 10^{-14}$  erg cm $^{-2}$  s $^{-1}$  and the difference becomes more pronounced at bright fluxes ( $> 10\%$  at fluxes  $> 5 \times 10^{-14}$  erg cm $^{-2}$  s $^{-1}$ ).

The Gilli et al. (2007) and Treister et al. (2009) models are based on different assumptions on the fraction of obscured sources and on the assumed luminosity and redshift dependences. Therefore, their differences are more marked when considering obscured and unobscured sources separately. We used the hardness ratio, as defined in Section 4.1.4, to divide the sample using  $HR > 0.2$  for obscured sources and  $HR < 0.2$  for unobscured sources. In the soft (hard) band there are 1057 (1332) obscured sources and 1701 (911) unobscured ones. In Figure 15, we present the number counts in the soft and hard bands (left and right) for both obscured (red) and unobscured (blue) sources. A clear difference is observed in the number counts of obscured and unobscured in the soft band, where we observe a ratio of up to  $\sim 10$  at bright fluxes, while it almost disappears in the hard band, where the ratio is very small at all fluxes. This im-

plies that the difference must be dictated by obscuration effects.

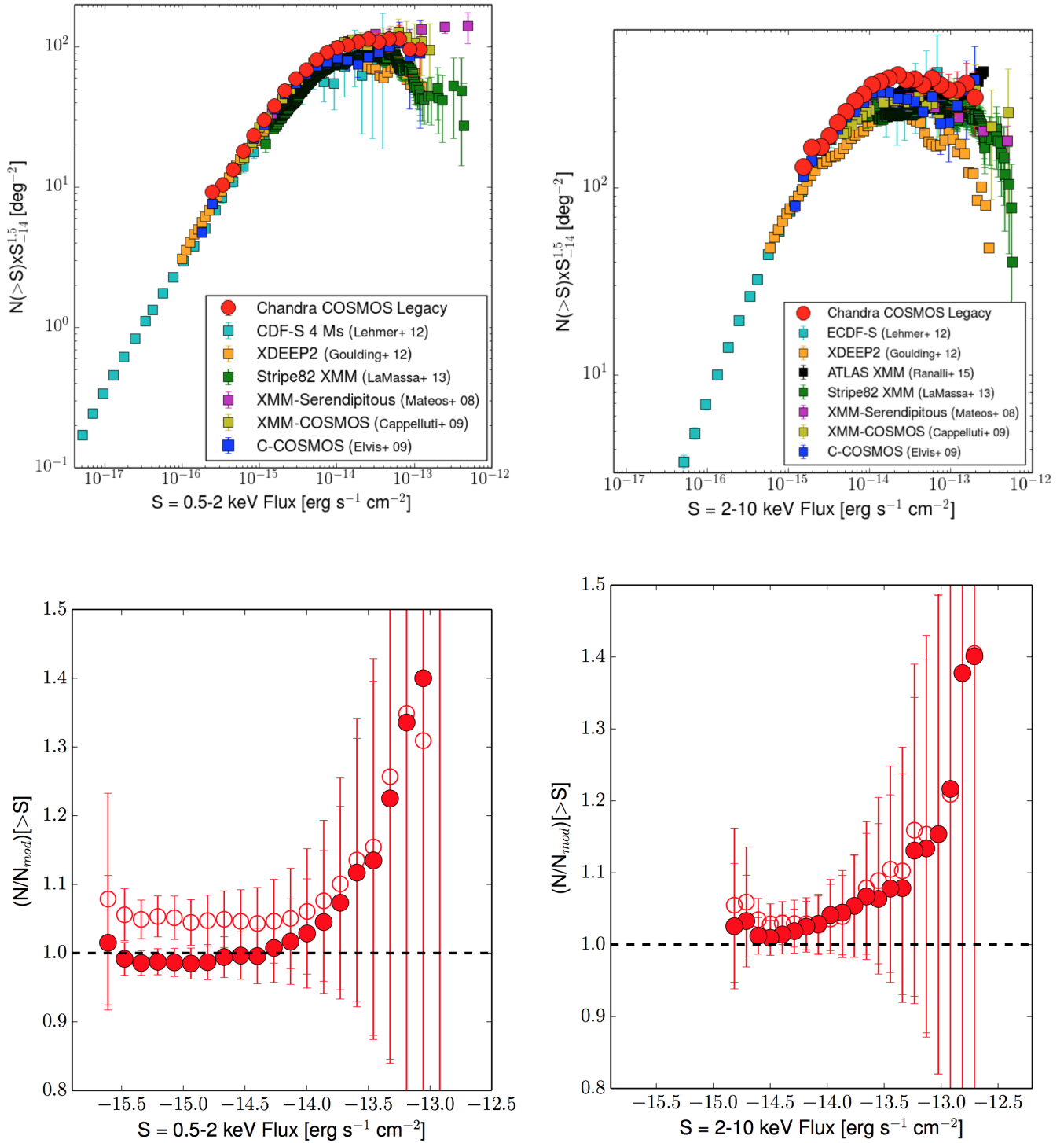
The models from Gilli et al. (2007, solid line) and Treister et al. (2009, dashed line) assuming column densities above and below  $10^{22}$  cm $^{-2}$  (red and blue respectively) are plotted in the same Figure. In the soft band, both predictions of the number of unobscured sources are in agreement within 5% with our data, up to fluxes of  $\sim 3 \times 10^{-14}$  erg cm $^{-2}$  s $^{-1}$ , while the difference becomes larger for obscured sources ( $> 10$ -20%), with both models over predicting the number of sources at all fluxes. In this last case, the Treister et al. (2009) model predictions are generally worse than those of the Gilli et al. (2007) model, by 5-10%. In the hard band instead, model predictions are in general excellent agreement with our data (differences  $< 5\%$  up to fluxes of  $5 \times 10^{-14}$  erg cm $^{-2}$  s $^{-1}$ ), for both samples above and below  $HR = 0.2$ .

Overall, these discrepancies between data and models are totally expected given that as for example a different spectral model could change source fluxes and sky coverage, and that the spectral parameters in the Gilli et al. and Treister et al. models are different from those used in this work. Therefore, despite all the underlying assumptions, the differences between observed number counts and phenomenological models are remarkably small (2-5%; see also LaMassa et al. 2013a for a discussion on discrepancies between data and population synthesis models).

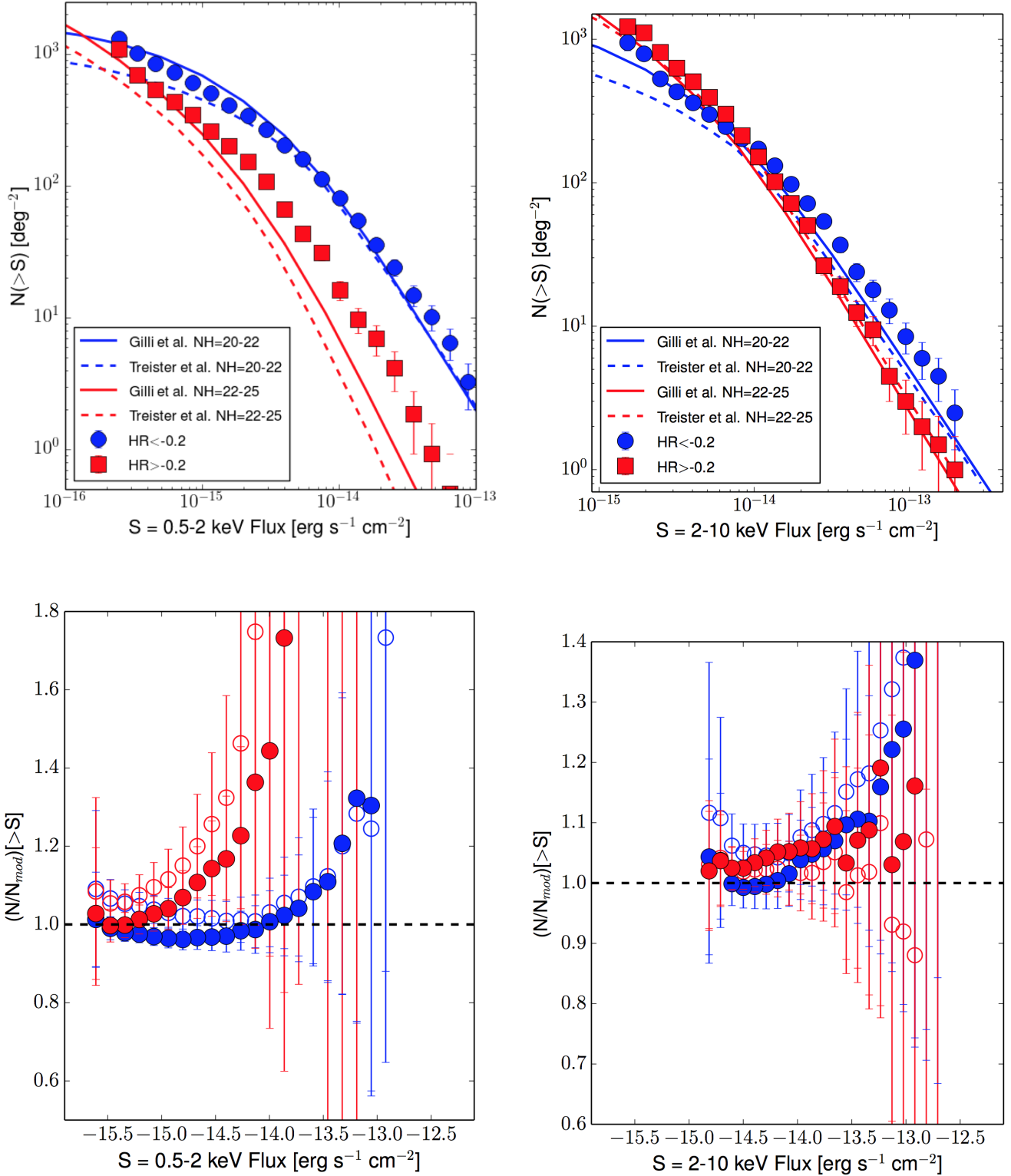
## 7. SUMMARY AND CONCLUSIONS

In this paper we presented *COSMOS-Legacy*, a 2.2 deg $^2$  *Chandra* survey of the COSMOS field. We employed a total of 4.6 Ms of exposure time, including 1.8 Ms already published by E09 plus 2.8 Ms obtained as an X-ray Visionary Project during *Chandra* Cycle 14. The new data comprise 56 overlapping observations which, added to the 36 C-COSMOS pointings, yield a relatively uniform coverage of  $\sim 160$  ksec over the whole *Hubble*-covered area. By construction, the survey flux limit is the same of C-COSMOS, computed in three bands using the same approach of P09.

We followed the same procedure used and tested by P09 combining standard CIAO tools for the data reduction, and PWDetect and CMLDetect for the data analysis, including the source detection and photometry. We also performed aperture photometry for consistency with the E09 and P09 analysis. The analysis was performed on the new *Chandra* data and also on the outer C-COSMOS frame, overlapping with the new observations. Given that the survey properties (exposure, roll angle and background counts) are consistent with C-COSMOS ones, we used the same probability threshold for the source detection corresponding to  $DET\_ML = 10.8$ . At this limit, we detected 2273 sources that were not previously detected in C-COSMOS, by combining detections in the full, soft and hard bands. 385 of these sources were detected in the area overlapping with C-COSMOS: in the same area we have also found 676 of the 694 old detections, while 18 sources were not detected again. The total number of sources in *COSMOS-Legacy* is 4016. The source properties, including counts, count rates, fluxes in three bands (full, soft and hard) and hardness ratio computed using a Bayesian approach are reported in an online table published with this paper.



**Figure 14.** Euclidean normalized  $\log N$ - $\log S$  curves in 0.5-2 keV (top-left) and 2-10 keV (top-right) bands. The *COSMOS-Legacy* curve for all sources with  $\text{DET\_ML} > 10.8$  and  $\text{SNR} > \text{SNR}_{\text{lim}}$  is plotted in red circles. Results from previous works are plotted (see label in the plot). The ratio of *COSMOS-Legacy* number counts to Gilli et al. (2007, red solid) and Treister et al. (2009; red empty) models are plotted in the soft and hard bands (bottom left and bottom right). The source number counts are multiplied by  $(S/10^{14})^{1.5}$  to highlight the deviations from the Euclidean behavior.



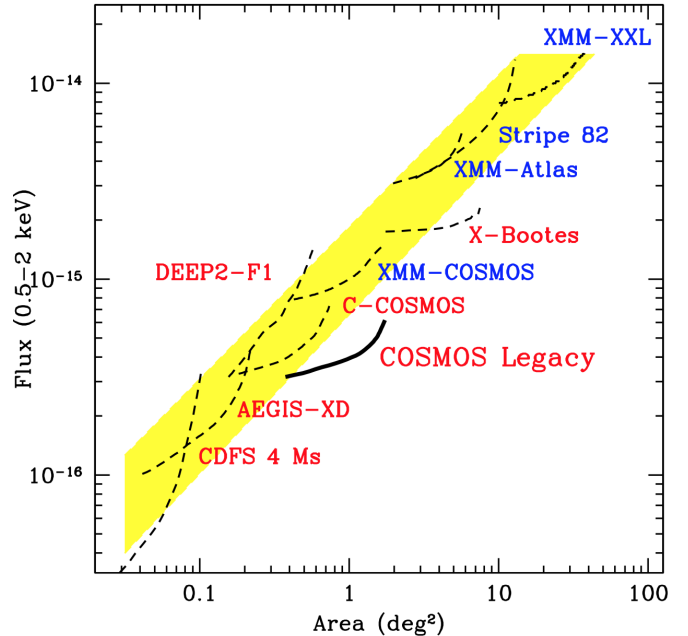
**Figure 15.** Number counts in soft (top-left) and hard (top-right) bands for sources with  $\text{HR} > -0.2$  (red squares) and  $< -0.2$  (blue circles) plotted with the Gilli et al. (solid) and Treister et al. (dashed) models with two different column density ranges  $> 10^{22} \text{ cm}^{-2}$  in red and  $< 10^{22} \text{ cm}^{-2}$  in blue. The ratio of *COSMOS-Legacy* number counts to Gilli et al. (2007; solid) and Treister et al. (2009; empty) models are plotted in the soft and hard bands (bottom left and bottom right).

We computed the source number counts in both the soft and hard bands and we find good agreement between our results and other surveys in the literature as listed above. The large number of sources in *COSMOS-Legacy* (20% or more than the sources in other contiguous surveys) allows to constrain the number of counts at medium fluxes ( $\sim 10^{-15}$  erg cm $^{-2}$  s $^{-1}$ ) with 10% errors and to reduce the uncertainties on the normalization at bright fluxes where discrepancies between different surveys still exist. The combination of *COSMOS-Legacy* with other surveys at fainter and brighter fluxes allows to cover more than 4 orders of magnitude in flux.

Using the hardness ratio we measure a fraction of obscured sources of  $50_{-16}^{+17}\%$ , defined as sources with  $HR > -0.2$ , corresponding to column density  $> 10^{22}$  cm $^{-2}$  at all redshifts, despite the uncertainties on the classification due to complex spectral modeling not taken into account in this work (see Wilkes et al. 2009, 2013). For the first time, we computed the number counts for obscured and unobscured sources separately using the hardness ratio as an indication for obscuration ( $HR = -0.2$  corresponding to the separation between  $>$  and  $< 10^{22}$  cm $^{-2}$ ). The large number of sources in each sample (about a thousand or more) allows to compute the number counts for the two populations and reveals a larger difference (in both normalization and shape) in the soft band, while a very small if not absent difference in the hard band is observed (the normalization is consistent, while we can observe a small difference in shape). Given the large range of luminosities and redshifts probed by *COSMOS-Legacy* this can be interpreted as a difference in orientation rather than an intrinsic difference due to an evolutionary state between obscured and unobscured sources.

In Figure 16, the area – flux parameter space of the most recent *Chandra* and XMM-*Newton* surveys (CDFS 4Ms, Xue et al. 2011; AEGIS-XD, Nandra et al. 2015; XDEEP2-F1, Goulding et al. 2012; C-COSMOS, E09; XMM-COSMOS, Cappelluti et al. 2009; X-Bootes, Murray et al. 2005; XMM-Atlas, Ranalli et al. 2015; Stripe 82, LaMassa et al. 2013a,b and LaMassa et al. 2015; XMM-XXL, PI: Pierre, see also Pierre et al. 2004) is presented. Most surveys lie on a locus (yellow shaded area) determined by our current X-ray telescope capabilities. *COSMOS-Legacy* is exploring a new region off this locus, an additional factor 2-3 deeper at the areas it covers, by using a total exposure time which is unusually large (4.6 Ms total) for that given area flux combination, and preparing for surveys with future facilities. The X-Bootes survey also explores a region off the survey locus, but at brighter fluxes and over a larger area.

In future decades, with facilities like *eROSITA* (Merloni et al. 2012), *Athena* (Nandra et al. 2013) and *X-ray Surveyor* (Vikhlinin et al. 2012), it will be possible to explore a new region of area-flux parameter space, moving away from the current survey locus towards the bottom right corner of Figure 16. For example, *Athena+* will perform a multi-tiered survey and given the combination of large effective area and field of view will enable X-ray surveys to be carried out two orders of magnitude faster than XMM-*Newton* and *Chandra* (see Figure 2 of Aird et al. 2013). With a *Chandra*-like resolution over  $10'$ , *X-ray Surveyor* will be able to cover the same *COSMOS-Legacy* area at the same flux in only 55 ksec (A. Vikhlinin private



**Figure 16.** Area–flux curves for *Chandra* (red) and XMM-*Newton* (blue) contiguous X-ray surveys. Each survey has been plotted using each sensitivity curve starting from the flux corresponding to the area that is 80% of the maximum area for that survey to the flux corresponding to the 20% of the total area. The plotted surveys are: CDFS 4Ms (Xue et al. 2011), XDEEP2-F1 (Goulding et al. 2011), AEGIS-XD (Nandra et al. 2015), C-COSMOS (E09), XMM-COSMOS (Cappelluti et al. 2009), X-Bootes (Murray et al. 2005), XMM-Atlas (Ranalli et al. 2015), Stripe 82 (LaMassa et al. 2013a,b, 2015), XMM-XXL (PI: Pierre; see also Pierre et al. 2004). The survey locus described in the last section is drawn in yellow.

communication), 80 times faster than *Chandra*.

Thanks to the large area covered at considerable depth, *COSMOS-Legacy* can now address those questions for which a large number of detected X-ray sources at a medium depth with uniform multiwavelength coverage and almost complete redshift information is needed. The excellent positional accuracy allows to obtain multiwavelength identifications and photometric redshifts for 96% of the sources (Marchesi et al. in press). We are currently working on papers on the X-ray luminosity function with a focus on the high redshift universe (Marchesi et al. in prep.), the X-ray spectral analysis and X-ray variability of the bright sample with a focus on the hunt for obscured sources (Lanzuisi et al. in prep.), the multiwavelength spectral energy distribution fitting with host galaxy properties (mass and star formation rates) for both optically classified as obscured and unobscured sources (Suh et al. in prep.), clustering measurement and dark matter halo mass (Allevato et al. in prep.) and, finally, a catalog of X-ray extended sources (Finoguenov et al. in prep.).

The wide area and the availability of extensive multiwavelength data in the COSMOS field enable us to probe the average X-ray emission of objects not individually detected by *Chandra*, therefore beyond the flux limit, through a stacking analysis. The combined *Chandra COSMOS-Legacy* dataset is now fully implemented

in the web-based *Chandra* stacking tool CSTACK<sup>38</sup>. This enables us to investigate the X-ray properties of differently selected samples, such as optical selected galaxies (e.g. Mezcuca et al. 2015, in press, finding indications of weak AGN activity in low mass non-elliptical galaxies), highly obscured AGN selected using both infrared or radio criteria, and early AGN populations at  $z > 5$ .

## 8. ACKNOWLEDGMENTS

The authors thank the referee for the useful comments. F.C. and M.C.U. gratefully thank Debra Fine for her support to women in science, the *Chandra* EPO team, in particular J. De Pasquale, for creating the true color X-ray image. This research has made use of data obtained from the *Chandra* Data Archive and software provided by the *Chandra* X-ray Center (CXC) in the CIAO application package.

This work was supported in part by NASA Chandra grant number GO7-8136A (F.C., S.M., V.A., M.E., H.S.); PRIN-INAF 2014 "Windy Black Holes combing galaxy evolution" (A.C., M.B., G.L. and C.V.); the FP7 Career Integration Grant "eEASy": "Supermassive blackholes through cosmic time: from current surveys to eROSITA-Euclid Synergies" (CIG 321913; M.B. and G.L.); UNAM-DGAPA Grant PAPIIT IN104113 and CONACyT Grant Científica Básica #179662 (T.M.); Collaborative Research Council 956, sub-project A1, funded by the Deutsche Forschungsgemeinschaft (A.K.); NASA contract NAS8-03060 (T.A., A.F., K. G); the Greek General Secretariat of Research and Technology in the framework of the programme Support of Postdoctoral Researchers (P.R.); NASA award NNX15AE61G (R.G.); the Science and Technology Facilities Council through grant code ST/I001573/1 (D.M.A.); the Swiss National Science Foundation Grant PP00P2.138979/1 (K.S.); the Center of Excellence in Astrophysics and Associated Technologies (PFB 06), by the FONDECYT regular grant 1120061 and by the CONICYT Anillo project ACT1101 (E.T.); the European Union's Seventh Framework programme under grant agreements 337595 (ERC Starting Grant, "CoSMass") and 333654 (CIG, AGN feedback; V.S.) S.T. is part of The Dark Cosmology Centre, funded by the Danish National Research Foundation. B.T. is a Zwicky Fellow.

## REFERENCES

- Aird, J., Comastri, A., Brusa, M., et al. 2013, arXiv:1306.2325  
 Allevato, V., Finoguenov, A., Cappelluti, N., et al. 2011, ApJ, 736, 99  
 Brusa, M., Civano, F., Comastri, A., et al. 2010, ApJ, 716, 348  
 Capak, P., Aussel, H., Ajiki, M., et al. 2007, ApJS, 172, 99  
 Capak, P. L., Riechers, D., Scoville, N. Z., et al. 2011, Nature, 470, 233  
 Cappelluti, N., Brusa, M., Hasinger, G., et al. 2009, A&A, 497, 635  
 Cappelluti, N., Kashlinsky, A., Arendt, R. G., et al. 2013, ApJ, 769, 68  
 Civano, F., Brusa, M., Comastri, A., et al. 2011, ApJ, 741, 91  
 Civano, F., Elvis, M., Brusa, M., et al. 2012, ApJS, 201, 30  
 Cruddace, R. G., Hasinger, G. R., & Schmitt, J. H. 1988, European Southern Observatory Conference and Workshop Proceedings, 28, 177  
 Damiani, F., Maggio, A., Micela, G., & Sciortino, S. 1997, ApJ, 483, 350  
 Civano, F., Hickox, R. C., Puccetti, S., et al. 2015, ApJ, 808, 185  
 Elvis, M., Civano, F., Vignali, C., et al. 2009, ApJS, 184, 158  
 Finoguenov, A., Guzzo, L., Hasinger, G., et al. 2007, ApJS, 172, 182  
 Fruscione, A., McDowell, J. C., Allen, G. E., et al. 2006, Proc. SPIE, 6270, 62701V  
 Garmire, G. P., Bautz, M. W., Ford, P. G., Nousek, J. A., & Ricker, G. R., Jr. 2003, Proc. SPIE, 4851, 28  
 Giallongo, E., Grazian, A., Fiore, F., et al. 2015, A&A, 578, A83  
 Gilli, R., Comastri, A., & Hasinger, G. 2007, A&A, 463, 79  
 Goulding, A. D., Forman, W. R., Hickox, R. C., et al. 2012, ApJS, 202, 6  
 Hasinger, G., Burg, R., Giacconi, R., et al. 1993, A&A, 275, 1  
 Hasinger, G., Cappelluti, N., Brunner, H., et al. 2007, ApJS, 172, 29  
 Hickox, R. C., & Markevitch, M. 2006, ApJ, 645, 95  
 Ilbert, O., Capak, P., Salvato, M., et al. 2009, ApJ, 690, 1236  
 Kalberla, P. M. W., Burton, W. B., Hartmann, D., et al. 2005, A&A, 440, 775  
 Kettula, K., Finoguenov, A., Massey, R., et al. 2013, ApJ, 778, 74  
 Koekemoer, A. M., Aussel, H., Calzetti, D., et al. 2007, ApJS, 172, 196  
 Krumpel, M., Miyaji, T., Brunner, H., et al. 2015, MNRAS, 446, 911  
 LaMassa, S. M., Urry, C. M., Glikman, E., et al. 2013, MNRAS, 432, 1351  
 LaMassa, S. M., Urry, C. M., Cappelluti, N., et al. 2013, MNRAS, 436, 3581  
 LaMassa, S. M., Urry, C. M., Cappelluti, N., et al. 2015, arXiv:1510.00852  
 Lanzuisi, G., Civano, F., Elvis, M., et al. 2013, MNRAS, 431, 978  
 Lehmer, B. D., Xue, Y. Q., Brandt, W. N., et al. 2012, ApJ, 752, 46  
 Lilly, S. J., Le Brun, V., Maier, C., et al. 2009, ApJS, 184, 218  
 Luo, B., Bauer, F. E., Brandt, W. N., et al. 2008, ApJS, 179, 19  
 Lusso, E., Comastri, A., Simmons, B. D., et al. 2012, MNRAS, 425, 623  
 Lutz, D., Poglitsch, A., Altieri, B., et al. 2011, A&A, 532, A90  
 Mainieri, V., Bongiorno, A., Merloni, A., et al. 2011, A&A, 535, A80  
 Marchesi, S., Civano, F., Elvis, M., et al. 2015, arXiv:1512.01105  
 Mateos, S., Warwick, R. S., Carrera, F. J., et al. 2008, A&A, 492, 51  
 McCracken, H. J., Milvang-Jensen, B., Dunlop, J., et al. 2012, A&A, 544, A156  
 Mezcuca, M., Civano, F., Fabbiano, G., Miyaji, T., & Marchesi, S. 2015, arXiv:1511.05844  
 Murray, S. S., Kenter, A., Forman, W. R., et al. 2005, ApJS, 161, 1  
 Nandra, K., Laird, E. S., Aird, J. A., et al. 2015, arXiv:1503.09078  
 Nandra, K., Barret, D., Barcons, X., et al. 2013, arXiv:1306.2307  
 Park, T., Kashyap, V. L., Siemiginowska, A., et al. 2006, ApJ, 652, 610  
 Pierre, M., Valtchanov, I., Altieri, B., et al. 2004, Journal of Cosmology and Astroparticle Physics, 9, 011  
 Puccetti, S., Vignali, C., Cappelluti, N., et al. 2009, ApJS, 185, 586  
 Ranalli, P., Georgantopoulos, I., Corral, A., et al. 2015, A&A, 577, A121  
 Schinnerer, E., Carilli, C. L., Scoville, N. Z., et al. 2004, AJ, 128, 1974  
 Schinnerer, E., Sargent, M. T., Bondi, M., et al. 2010, ApJS, 188, 384  
 Schinnerer, E., Smolčić, V., Carilli, C. L., et al. 2007, ApJS, 172, 46  
 Scoville, N., Abraham, R. G., Aussel, H., et al. 2007, ApJS, 172, 38  
 Scoville, N., Aussel, H., Brusa, M., et al. 2007, ApJS, 172, 1  
 Smolčić, V., Ciliegi, P., Jelić, V., et al. 2014, MNRAS, 443, 2590  
 Taniguchi, Y., Scoville, N., Murayama, T., et al. 2007, ApJS, 172, 9  
 Trakhtenbrot, B., Urry, C. M., Civano, F., et al. 2015, Science, 349, 168  
 Treister, E., Urry, C. M., & Virani, S. 2009, ApJ, 696, 110  
 Vikhlinin, A., Reid, P., Tananbaum, H., et al. 2012, Proc. SPIE, 8443, 844316  
 Weisskopf, M. C., Brinkman, B., Canizares, C., et al. 2002, PASP, 114, 1  
 Wilkes, B. J., Kilgard, R., Kim, D.-W., et al. 2009, ApJS, 185, 433  
 Wilkes, B. J., Kuraszkiewicz, J., Haas, M., et al. 2013, ApJ, 773, 15  
 Xue, Y. Q., Luo, B., Brandt, W. N., et al. 2011, ApJS, 195, 10

user=guest, password=guest and see the explanatory manual. As

<sup>38</sup> See <http://lambic.astrosen.unam.mx/cstack>. Login as

of writing this paper, stacking analyses utilizing the C-COSMOS dataset is publicly available. Analyses involving the whole *Chan-*

*dra COSMOS-Legacy* dataset is still proprietary and will be public in due course.