(Evolution) and identification of (high-redshifts) AGN: lessons for WFXT from COSMOS and CDFS

Marcella Brusa
MPE

contributions from many people
WFXT Goal: evolution of high-z sources

Current knowledge of high-z (> 3) AGN from X-ray (and optical) surveys (see talks by R. Gilli & F. Fiore)

Challenges:
1) Statistics --> large area surveys
2) Identifications --> lessons from XMM/Chandra surveys
3) Redshifts --> multiwavelength follow-up

Resources needed....
Number Statistics
X-rays from high-z Quasars

1990-1994: pioneering works with ROSAT
Wilkes+92, Elvis+94, Bechtold+94 (record QSO z=4)

2002-2005: Chandra/XMM contribution
Follow-up of optically SDSS QSOs
Brandt+02, Mathur+02, Vignali+03, 05 (record QSO z=6.4)
XMM-COSMOS z>3 QSOs (Brusa et al. 09)

The number of high-z AGN detected so far

<table>
<thead>
<tr>
<th>SDSS*</th>
<th>X-ray sel. $^$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$z &gt; 3$</td>
<td>8000</td>
</tr>
<tr>
<td>$z &gt; 4$</td>
<td>1500</td>
</tr>
<tr>
<td>$z &gt; 5$</td>
<td>150</td>
</tr>
<tr>
<td>$z &gt; 6$</td>
<td>10</td>
</tr>
</tbody>
</table>

X-rays needed to get the LF faint end (more representative of the whole high-z population)
Lg(Lx)>44 QSO:
same behaviour of optically selected bright
QSOs (logLx~45)

based on 40 QSOs from XMM-COSMOS
XMM and Chandra z>3 QSOs

Log(Lx)>44 QSO:
same behaviour of optically selected bright QSOs (logLx~45)
based on 40 QSOs from XMM-COSMOS Brusa, Comastri et al. 09, ApJ

High fluxes (>5x10^{-16} cgs):
data and predictions robust; to have same statistics of SDSS need to survey >200 deg2 at COSMOS depth

Low fluxes (<10^{-16} cgs):
data scarce, predictions uncertain [CDFS analysis predict a factor of ~2 more than extrapolations]
Expectations

From Brusa+09 [Gilli et al. 2007 model]

<table>
<thead>
<tr>
<th>z range</th>
<th>limiting flux</th>
<th>constant</th>
<th>decline</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt; 3</td>
<td>&gt; 4 \times 10^{-15}</td>
<td>230</td>
<td>75</td>
</tr>
<tr>
<td>&gt; 10^{-15}</td>
<td>80</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>&gt; 4 \times 10^{-15}</td>
<td>14</td>
<td>6.2</td>
<td></td>
</tr>
<tr>
<td>&gt; 10^{-14}</td>
<td>1.8</td>
<td>0.75</td>
<td></td>
</tr>
<tr>
<td>&gt; 4</td>
<td>&gt; 4 \times 10^{-15}</td>
<td>80</td>
<td>13</td>
</tr>
<tr>
<td>&gt; 10^{-15}</td>
<td>30</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>&gt; 4 \times 10^{-15}</td>
<td>3</td>
<td>0.5</td>
<td></td>
</tr>
<tr>
<td>&gt; 10^{-14}</td>
<td>0.6</td>
<td>0.07</td>
<td></td>
</tr>
</tbody>
</table>

From WFXT white paper

<table>
<thead>
<tr>
<th>WFXT</th>
<th>flux(lim)</th>
<th>deg2</th>
<th>z&gt;3</th>
<th>z&gt;6</th>
</tr>
</thead>
<tbody>
<tr>
<td>wide</td>
<td>4x10^{-15}</td>
<td>20.000</td>
<td>1.26x10^5</td>
<td>500</td>
</tr>
<tr>
<td>medium</td>
<td>5x10^{-16}</td>
<td>3.000</td>
<td>2.25x10^5</td>
<td>1000</td>
</tr>
<tr>
<td>deep</td>
<td>3x10^{-17}</td>
<td>100</td>
<td>3(6)x10^4</td>
<td>300 (&gt;300)</td>
</tr>
</tbody>
</table>
## Expectations

### From Brusa+09 [Gilli et al. 2007 model]

<table>
<thead>
<tr>
<th>$z$ range</th>
<th>limiting flux</th>
<th>constant</th>
<th>decline</th>
</tr>
</thead>
<tbody>
<tr>
<td>$z &gt; 3$</td>
<td>$&gt; 4 \times 10^{-15}$</td>
<td>230 deg$^{-2}$</td>
<td>75 deg$^{-2}$</td>
</tr>
<tr>
<td>$z &gt; 4$</td>
<td>$&gt; 4 \times 10^{-15}$</td>
<td>80 deg$^{-2}$</td>
<td>30 deg$^{-2}$</td>
</tr>
</tbody>
</table>

### From WFXT white paper

<table>
<thead>
<tr>
<th>$z$ range</th>
<th>limiting flux</th>
<th>deg$^2$</th>
<th>$z &gt; 3$</th>
<th>$z &gt; 6$</th>
</tr>
</thead>
<tbody>
<tr>
<td>wide</td>
<td>$4 \times 10^{-15}$</td>
<td>20,000</td>
<td>$1.26 \times 10^5$</td>
<td>500</td>
</tr>
<tr>
<td>all sky</td>
<td>$10^{-14}$</td>
<td>30,000</td>
<td>$2.25 \times 10^4$</td>
<td>30</td>
</tr>
<tr>
<td>medium</td>
<td>$4 \times 10^{-15}$</td>
<td>400</td>
<td>$2.5 \times 10^3$</td>
<td>4</td>
</tr>
</tbody>
</table>
**Expectations**

Table 3. Expected numbers of \(z>3\) QSOs

<table>
<thead>
<tr>
<th>(z) range</th>
<th>limiting flux</th>
<th>constant(^a)</th>
<th>decline(^b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(z&gt;3)</td>
<td>(&gt;4 \times 10^{-15})</td>
<td>230 deg(^{-2})</td>
<td>75 deg(^{-2})</td>
</tr>
<tr>
<td>(&gt;10^{-15})</td>
<td>80</td>
<td></td>
<td>30</td>
</tr>
<tr>
<td>(&gt;4 \times 10^{-16})</td>
<td>14</td>
<td></td>
<td>6.2</td>
</tr>
<tr>
<td>(&gt;10^{-14})</td>
<td>1.8</td>
<td></td>
<td>0.75</td>
</tr>
<tr>
<td>(z&gt;4)</td>
<td>(&gt;4 \times 10^{-15})</td>
<td>80</td>
<td>13</td>
</tr>
</tbody>
</table>

From Brusa+09 [Gilli et al. 2007 model]

From WFXT white paper

Message (1):

* High-\(z\) QSOs in WFXT --> statistics will be even few orders of magnitude **larger** than SDSS

A LOT OF HIGH-Z AGN!

<table>
<thead>
<tr>
<th>eROSITA</th>
<th>flux(lim)</th>
<th>deg2</th>
<th>(z&gt;3)</th>
<th>(z&gt;6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>all sky</td>
<td>(10^{-14})</td>
<td>30,000</td>
<td>2.25x10(^4)</td>
<td>30</td>
</tr>
<tr>
<td>medium</td>
<td>4x10(^{-15})</td>
<td>400</td>
<td>2.5x10(^3)</td>
<td>4</td>
</tr>
</tbody>
</table>
Identification issues
(whole X-ray population)
Counterparts Identifications

(some references: Sutherland & Saunders 1988, Ciliegi et al. 2003, Brusa et al. 2005)
Counterparts Identifications

(some references: Sutherland & Saunders 1988, Ciliegi et al. 2003, Brusa et al. 2005)

TOOLS:

1) a statistical, powerful, method, the “Likelihood Ratio Technique” (Sutherland & Sanders 1992) widely used in several Chandra/XMM surveys in recent years

2) combined information from different wavebands (optical / K-band / IR)
Counterparts Identifications

(some references: Sutherland & Saunders 1988, Ciliegi et al. 2003, Brusa et al. 2005)

TOOLS:

1) a statistical, powerful, method, the “Likelihood Ratio Technique” (Sutherland & Sanders 1992) widely used in several Chandra/XMM surveys in recent years

2) combined information from different wavebands (optical / K-band / IR)

\[ LR = f(r) \times q(m) / n(m) \]

- \( f(r) \) = distance term (distance X-cp + positional errors)
- \( n(m) \) = background distribution
- \( q(m) \) = overdensities of the counterparts
Counterparts Identifications

(some references: Sutherland & Saunders 1988, Ciliegi et al. 2003, Brusa et al. 2005)

TOOLS:

1) a statistical, powerful, method, the “Likelihood Ratio Technique” (Sutherland & Sanders 1992) widely used in several Chandra/XMM surveys in recent years

2) combined information from different wavebands (optical / K-band / IR)

\[
LR = \frac{f(r) \times q(m)}{n(m)}
\]

- \(f(r)\) = distance term (distance X-cp + positional errors)
- \(n(m)\) = background distribution
- \(q(m)\) = overdensities of the counterparts

The LR is computed for each source in each band (I,K,3.6micron...)

The procedure gives, for each band, the most likely counterpart; in case of \(\geq 2\) equally likely counterparts (in the same and/or from different bands) all the cp are considered (“ambiguous”)

Important for XMM sources (at almost all fluxes) and Chandra sources mostly at \(F<10^{-15}\)
Counterparts Identifications
Counterparts Identifications

XMM-COSMOS (note: XMM PSF worse than WFXT....)

BREAKDOWN:

85% unique associations; 15% ambiguous associations at F>1e-15

95% unique + 5% ambiguous associations at fluxes of the WFXT wide survey
statistical properties of primary and secondary within ambiguous sources are indistinguishable - in most cases the two sources have same optical / K-band magnitudes
Counterparts Identifications

XMM-COSMOS (note: XMM PSF worse than WFXT....)

BREAKDOWN:

85% unique associations; 15% ambiguous associations at F>1e-15

95% unique + 5% ambiguous associations at fluxes of the WFXT wide survey
statistical properties of primary and secondary within ambiguous sources are indistinguishable - in most cases the two sources have same optical / K-band magnitudes

RELIABILITY* of the method [“a posteriori” test on XMM-COSMOS id using Chandra]

98.7% [only 9/712 unique sources resulted associated to the wrong optical cp]

99.6% [only 1/245 unique sources at fluxes of the WFXT wide survey]

(*see discussion in XMM-COSMOS ID paper; Brusa et al. to be subm)
Counterparts Identifications

XMM-COSMOS (note: XMM PSF worse than WFXT....)

BREAKDOWN:

85% unique associations; 15% ambiguous associations at F>1e-15

95% unique + 5% ambiguous associations at fluxes of the WFXT wide survey statistical properties of primary and secondary within ambiguous sources are indistinguishable - in most cases the two sources have same optical / K-band magnitudes

RELIABILITY*

98.7%

99.6%

HST/ACS  

Chandra  

WFXT (simulation)

(*see discussion in XMM-COSMOS ID paper; Brusa et al. to be subm)
Counterparts Identifications
Counterparts Identifications

CDFS (note: Chandra PSF better than WFXT....)

BREAKDOWN (Luo et al. 2Ms, AJ, submitted):

96% identifications, 4% not identified at fluxes comparable to WFXT deep
Counterparts Identifications

CDFS (note: Chandra PSF better than WFXT....)

BREAKDOWN (Luo et al. 2Ms, AJ, submitted):

96% identifications, 4% not identified at fluxes comparable to WFXT deep

[ among identifications: 90% “unique” associations; 10% ambiguous associations] statistical properties of primary and secondary within ambiguous sources are different - in most cases the two sources have very different SED....
Counterparts Identifications

CDFS (note: Chandra PSF better than WFXT....)

BREAKDOWN (Luo et al. 2Ms, AJ, submitted):

96% identifications, 4% not identified at fluxes comparable to WFXT deep

[ among identifications: 90% “unique” associations; 10% ambiguous associations] statistical properties of primary and secondary within ambiguous sources are different - in most cases the two sources have very different SED....
Counterparts Identifications

CDFS (note: Chandra PSF better than WFXT....)

BREAKDOWN (Luo et al. 2Ms, AJ, submitted):

96% identifications, 4% not identified at fluxes comparable to WFXT deep

[ among identifications: 90% “unique” associations; 10% ambiguous associations] statistical properties of primary and secondary within ambiguous sources are different - in most cases the two sources have very different SED....
Counterparts Identifications

CDFS (note: Chandra PSF better than WFXT....)

BREAKDOWN (Luo et al. 2Ms, AJ, submitted):

96% identifications, 4% not identified at fluxes comparable to WFXT deep

[ among identifications:
90% "unique" associations; 10% ambiguous associations
statistical properties of primary and secondary within ambiguous sources are different - in most cases the two sources have very different SED....

HST/ACS K-band IRAC

Message (2):

* WFXT wide --> **identification "easy"** (straightforward) [also for eROSITA...]
  X-ray optical / X-ray - infrared correlations; low density of bkg sources

* WFXT deep --> secure identification **not trivial**
  high density of bkg sources; different sources emerges in different bands; 5” HEW really auspicable
At the limiting flux of the WFXT wide survey an optical coverage to $I \sim 24$ and infrared coverage to $K \sim 22$ would be enough...BUT this should be on the entire area...
Sensitivity of future large area surveys

~20000-30000 deg$^2$ – shallow sensitivity surveys

**PanSTARRS:**  
$I \sim 24.2$ (+grzy)

**LSST:**  
$I \sim 25.5$ (+ugrzy)

**EUCLID:**  
$K \sim 23.5$ (+zJH)

**LOFAR:**  
0.8 mJy at 120 MHz (= 0.1 mJy at 1.4 GHz)  
“radio” emitters (AGN and starburst)

(VISTA VHS, $K=20$, not enough..)
At the limiting flux of the WFXT deep survey an optical coverage to $I \sim 27$ and infrared coverage to $K \sim 24$ over 100 deg$^2$ is needed.
Sensitivity of future deep surveys

~20-100 deg² – deep sensitivity surveys
[need more coordination...]

**LSST:**
I~26.7 (+ugrzy) - over 500 deg²
cp for 90% of the sources

**EUCLID:**
K~25 (deep survey, on 50 deg² ...)

**PanSTARRS:**
I~28 (+grzy) - over 28 deg²

**VISTA VIDEO:**
K=23.5 (+zYJH) - over 15 deg²

**JWST**
Sensitivity of future deep surveys

\[ \sim 20-100 \text{ deg}^2 \] – deep sensitivity surveys
[need more coordination...]

**LSST:**

I\[\sim 26.7 \ (+ugrzy) \] - over 500 deg\(^2\)

cp for 90\% of the sources

**EUCLID:**

K\[\sim 25 \ (\text{deep survey, on } 50 \text{ deg}^2)\]

**PanSTARRS:**

I\[\sim 28 \ (+grzy) \] - over 28 deg\(^2\)

**VISTA VIDEO:**

K=23.5 (+zYJH) - over 15 deg\(^2\)

**JWST**

courtesy: R. Gilli

**Message (3):**

* currently planned and ongoing optical / IR all sky surveys can **do the game**

* IR **more important** than optical for faint and very high-z sources!
Redshift determination: how to pick up $z > 3$ (or $z > 6$) QSOs among million sources?
XMM-COSMOS $z>3$ QSOs

How to isolate them?

1) Get spectroscopy or photometric redshifts for all the (million) sources

2) Impose color pre-selection

3) Get redshifts from Iron line.... (only a fraction)

* historical note: the original XMM-COSMOS proposal claimed $\sim 160$ QSOs at $z>3$ in the survey....

40* over $\sim 1650$ sources!!

2% of the XMM population
“complete” redshifts sample

XMM-COSMOS (almost 100%)

1640 XMM sources at $10^{-15}$ cgs

~840 “secure” spectroscopic redshifts (>50%)

~800 “good” photometric redshifts (Salvato et al. 2009)

Feasible only for small samples and/or when many optical/infrared filters are available. SDSS-like survey needed

key resources:

LSST (optical photometry);
EUCLID (IR photometry & spectra);
SDSSIII-BOSS (spectra)

Are depth/#of bands enough to get photz?

(Brusa et al. 2009, to be submitted)
Pre-selection based on colors (feasible from PanSTARRS and LSST multi-band photometry)

\[ z > 3 \] color-color selection v-i vs. b-v (proposed, e.g. in Casey et al. 2008, Siana et al. 2007)

U-dropout techniques (see Fabrizio talk)

8 objects would not have been selected
Pre-selection based on colors
(feasible from PanSTARRS and LSST multi-band photometry)

- z>3 color-color selection \(v-I\) vs. \(b-v\) (proposed, e.g. in Casey et al. 2008, Siana et al. 2007)
- U-dropout techniques (see Fabrizio talk)
- 8 objects would not have been selected
- High contaminant fraction (~50%)
Pre-selection based on colors
(feasible from PanSTARRS and LSST multi-band photometry)

From Willott+09

\[ z > 3 \] color-color selection \( v-I \) vs. \( b-v \) (proposed, e.g. in Casey et al. 2008, Siana et al. 2007)

U-dropout techniques (see Fabrizio talk)

8 objects would not have been selected

High contaminant fraction (~50%)

\[ z > 6 \] (optical): contaminants overwhelmingly more abundant (e.g. brown dwarfs 15x QSOS)

J-band used to discriminate between brown dwarfs and \( z = 6 \) QSOs, still spectroscopy success rate is ~20% (Jiang+09):

\[ z = 5.5 \]

\[ z = 6.7 \]
Pre-selection based on colors
(feasible from PanSTARRS and LSST multi-band photometry)

**Message (4):**
* multiwavelength information **mandatory**
* X-rays very efficient when coupled with other (less efficient) optical/IR criteria

Example: brown dwarfs are not X-ray emitters any match between an I-dropout and an X-ray source would immediately mark the object as a z>6 AGN!

z>3 color-color selection v-I vs. b-v (proposed, e.g. in Casey et al. 2008, Siana et al. 2007)  
U-dropout techniques (see Fabrizio talk)  
8 objects would not have been selected  
High contaminant fraction (~50%)

z>6 (optical): contaminants overwhelmingly more abundant (e.g. brown dwarfs 15x QSOS)

J-band used to discriminate between brown dwarfs and z=6 QSOs, still spectroscopy success rate is ~20% (Jiang+09):
Summary and WFXT Synergies
Summary and WFXT Synergies

Message (1):
* High-z QSOs in WFXT --> statistics will be even few orders of magnitude better than SDSS

Message (2):
* WFXT wide --> identification “easy” (straightforward) [also for eROSITA...]
  * WFXT deep --> secure identification not trivial
  5” HEW really auspicious

Message (3):
* currently planned and ongoing optical / IR all sky surveys can do the game (PanSTARRS, LSST, EUCLID, JWST...); deep coverage need coordination
  * IR more important than optical for faint and very high-z sources!

Message (4):
* multiwavelength information mandatory
* X-rays very efficient when coupled with other (less efficient) optical/IR criteria
Figure 2.1: The distribution of the $r$ band visits on the sky for one simulated realization of the baseline main survey. The sky is shown in Aitoff projection in equatorial coordinates and the number of visits for a 10-year survey is color-coded according to the inset. The two regions with smaller number of visits than the main survey (“mini-surveys”) are the Galactic plane (arc on the left) and the so-called “northern Ecliptic region” (upper right). The region around the South Celestial Pole will also receive substantial coverage (not shown here).