Exploring XMM-ATLAS with the ARCHES tools

Francisco J. Carrera
(IFCA, CSIC-UC, Spain)

Ioannis Georgantopoulos (NOA, Greece)
Piero Ranalli (Lund Obs., Sweden)
Manolis Rovilos (Patent Office, UK)

Paris, France, 30-November- 2015

Exploring the ARCHES cross-correlation tool with XMM-ATLAS

Francisco J. Carrera
(IFCA, CSIC-UC, Spain)

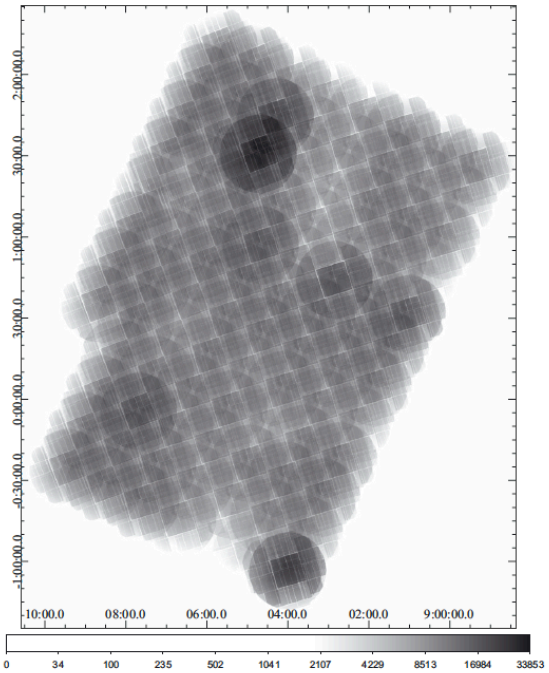
Ioannis Georgantopoulos (NOA, Greece)
Piero Ranalli (Lund Obs., Sweden)
Manolis Rovilos (Patent Office, UK)

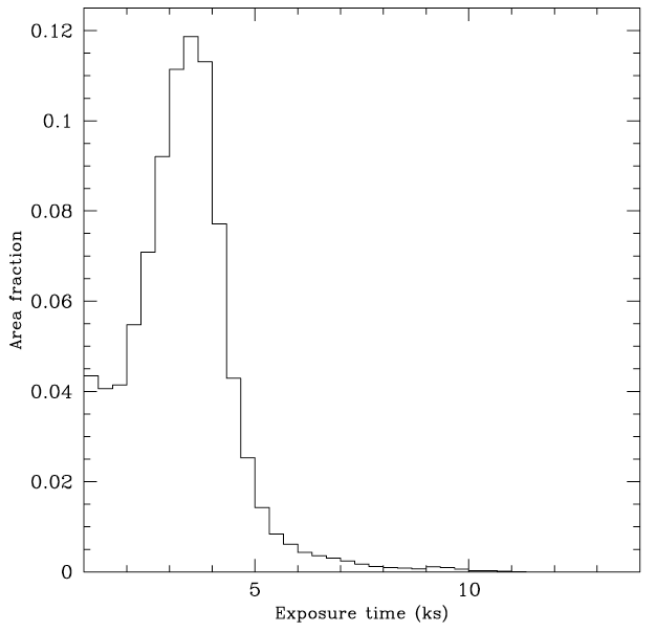
Paris, France, 30-November- 2015

Outline

- Introduction
  - XMM-ATLAS
  - Other catalogues: SDSS, VIKING, WISE
  - Astrometry
  - The ARCHES cross-correlation tool
- SDSS in full area:
  - Improvements
  - Checks
- Conclusions

XMM-ATLAS



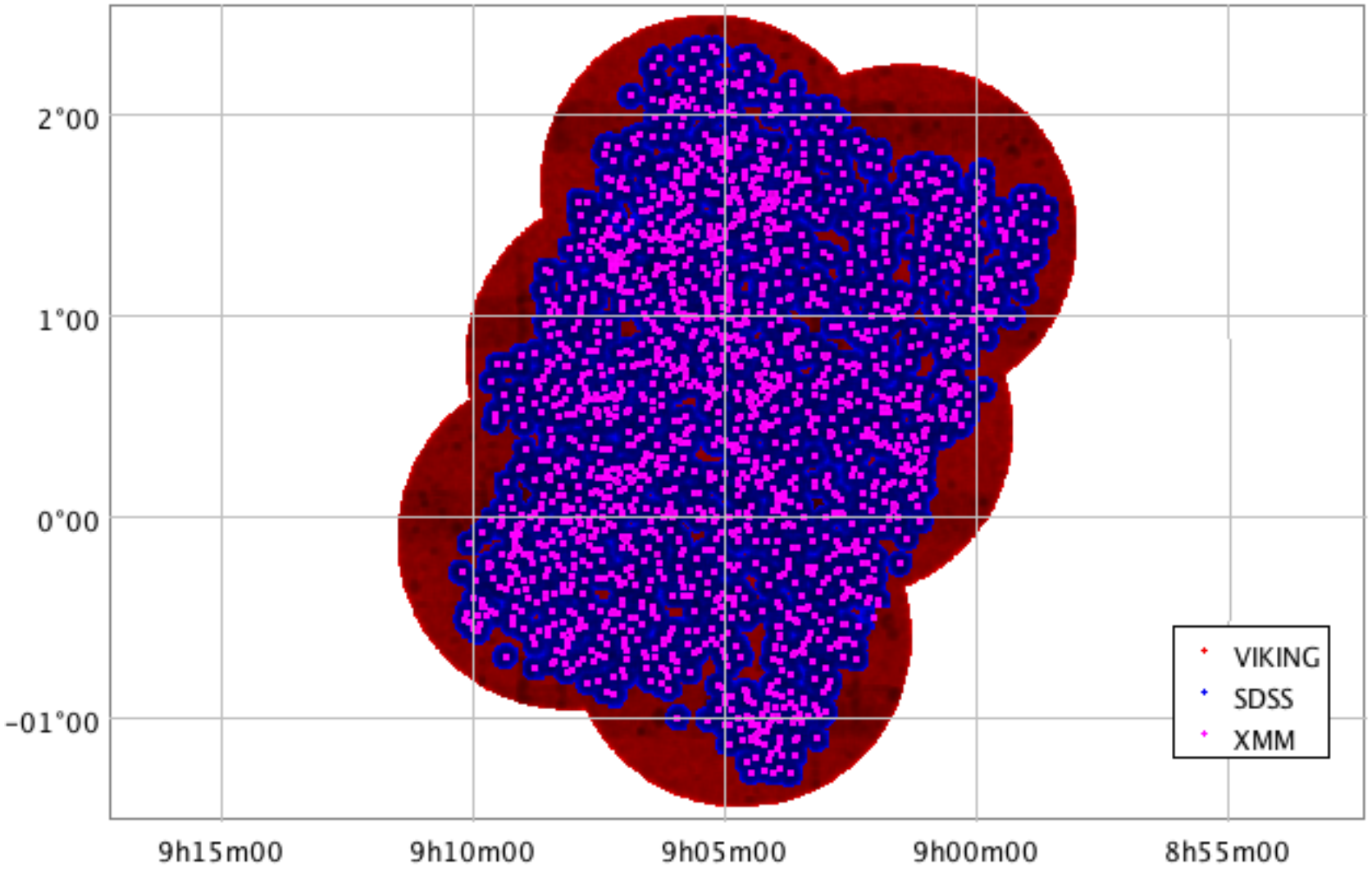


- Centered in Herschel-ATLAS SDP (Rigby+11)
  - (09:04:30,+00:34:00)
- Ranalli+15: source catalogue
  - 7.1deg^2 , 336ks total: mode ~3.5ks
  - Wavelet+`emldetect` source search: 1816 sources
    - in three bands: 0.5-2, 2-8, 0.5-2keV

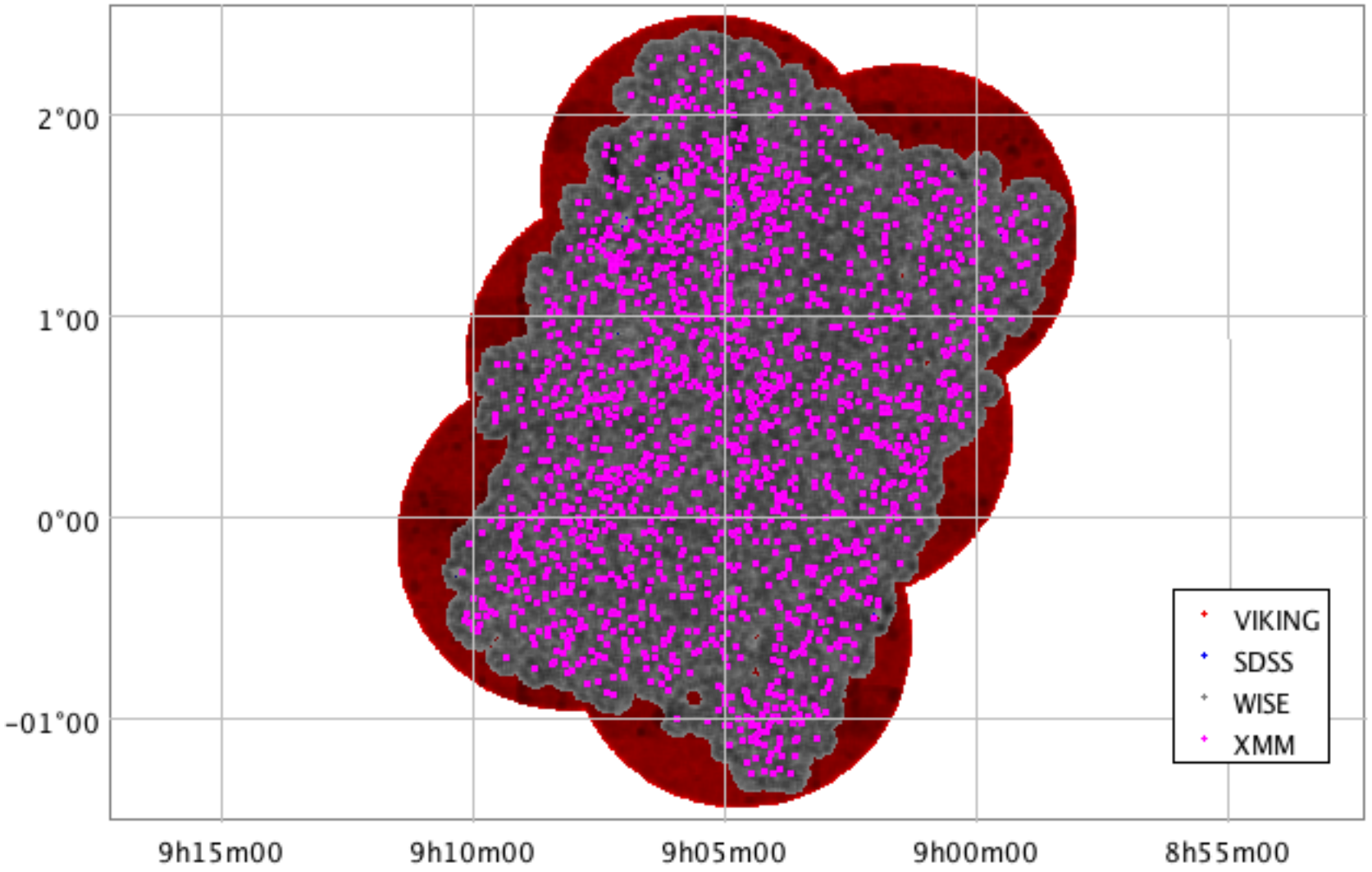
Other catalogues

- SDSS DR9 (Adelman-McCarthy+09): ugriz
  - 161131 sources in overlapping area (111404 `clean`=1)
- VIKING DR1 (Edge & Sutherland 2014): zYJHKs
  - band-merged source catalogue
  - 736187 sources in ~overlapping area (421850 `pNoise`<0.5 && nBands≥2)
- WISE (Cutri+2012): 3.4, 4.6, 12 & 22 μm
  - 68147 sources in overlapping area (all det. 5σ in ≥1 band)

Other catalogues



Other catalogues



Astrometry

- SDSS DR9: adding 0.1″ in quadrature to the RA,Dec pos. errors
- For XMM-ATLAS using SDSS DR12 QSO as reference
  - Filtering QSOs on good z quality
  - Filtering XMM-ATLAS pointlike `ext058`<0.001
  - 71 pairs within 5″: <dRA>=-0.02″ <dDec>=-0.2″ sigma~1.5″
    - Ranalli+15 rectified w.r.t. SDSS DR7 QSOs, residual difference?
  - Adding 1.5″ in quadrature to `radec_err` from `emldetect`
- For VIKING using SDSS DR9 point sources as reference
  - Filtering on VIKING point sources
  - 51014 pairs within 1″: <dRA>~<dDec>~0.035″ sigma~0.1″
  - Setting `errpos=0.1″` (no positional error in catalogue)
- WISE used as in catalogue

The ARCHES cross-correlation tool

- `archesxmatch`: Full N-dimensional symmetric cross-correlation of catalogues (see F.-X. Pineau's talk)
- It uses:
  - Source positions
  - Source positional errors ($1\,\sigma$)
  - Catalogue area
  - Sky densities of pairs, triplets...
- It provides:
  - List of tuples within some user-chosen $3\,\sigma$ distance
  - Probabilities of all combinations of catalogues for that tuple
  - List of five highest probabilities (when defined): `maxProbaVal1...maxProbaVal5`
  - Marginalised 2D probabilities for tuples in ≥3 catalogues:
    - e.g.: `margProba_AB` for A,B,C: $P_{marg}(AB)=P(ABC)+P(AB\_C)$

ARCHES xcorr. tool: X and 3 other cats

- For simplicity, let's consider N=4 (XMM-ATLAS+3 cats.)
  - A=XMM, B=SDSS, C=VIKING, D=WISE
- "Left-join": each XMM-ATLAS source will be considered afresh with each new catalogue: X-ray-centric
- Output tuples could have up to four dimensions:
  - No source in any other catalogue within limit `nPos=1`
  - A source in 1 other catalogue, `nPos=2`: AB,A_B; AC,A_C; AD,A_D
  - One source in each of 2 other catalogues, `nPos=3`:
    - ABC,AB_C,AC_B,A_BC≡(A)(BC),A_B_C≡(A)(B)(C)
    - ABD,AB_D,AD_B,A_BD,A_B_D
    - ACD,AC_D,AD_C,A_CD,A_C_D
    - ~~BCD,BC_D,BD_C,B_CD,B_C_D~~ (remember, X-ray-centric)
  - One source in each of 3 other catalogues , `nPos=4`:
    - ABCD
    - ABC_D,ABD_C,ACD_B
    - AB_CD,AB_C_D,AC_BD,AC_B_D,AD_BC,AD_B_C
    - A_BCD,A_BC_D,A_BD_C,A_B_CD,A_B_C_D

ARCHES xcorr. tool: X and 3 other cats

- X-ray-centric: Considering only relationship with XMM source
  - Using for classification (for the time being) only maximum probability max[P(*)]=P_{max}≡`maxProbaVal1`

		nPos			
		1	2	3	4
Only X	X___	All	A_B, A_C, A_D	A_BC,A_B_C, A_CD,A_C_D, A_BD,A_B_D	A_BCD,A_BC_D,A_B_CD,A_B_C_D, A_BD_C
X+1cat	XS__	-	AB	AB_C,AB_D	AB_CD,AB_C_D
	X_V_	-	AC	AC_B,AC_D	AC_BD,AC_B_D
	X__W	-	AD	AD_B,AD_C	AD_BC,AD_B_C
X+2cat	XSV_	-	-	ABC	ABC_D
	X_VW	-	-	ACD	ACD_B
	XS_W	-	-	ABD	ABD_C
X+3cat	XSVW	-	-	-	ABCD

ARCHES xcorr. tool: X and 3 other cats

		nPos							
		1	2			3			4
Only X	X___	-	A_B	A_C	A_D	A_BC, A_B_C	A_CD, A_C_D	A_BD, A_B_D	A_BCD, A_BC_D, A_B_CD, A_B_C_D, A_BD_C
X+1cat	XS__	-	AB	-	-	AB_C	-	AB_D	AB_CD, AB_C_D
	X_V_	-	-	AC	-	AC_B	AC_D	-	AC_BD, AC_B_D
	X__W	-	-	-	AD	-	AD_C	AD_B	AD_BC, AD_B_C
X+2cat	XSV_	-	-	-	-	ABC	-	-	ABC_D
	X_VW	-	-	-	-	-	ACD	-	ACD_B
	XS_W	-	-	-	-	-	-	ABD	ABD_C
X+3cat	XSVW	-	-	-	-	-	-	-	ABCD

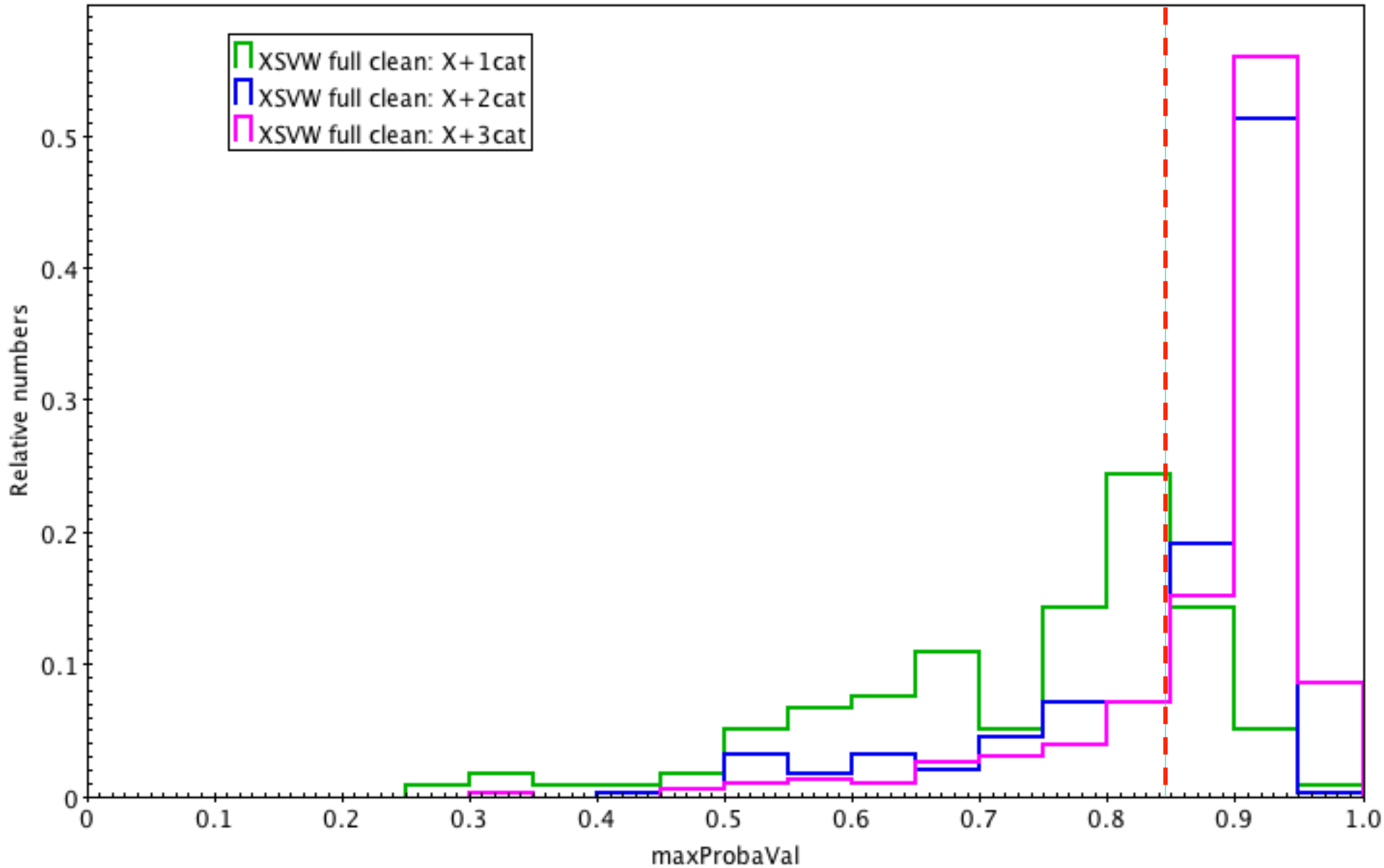
SDSS in full area: "clean" sample

- Full sample 1816 XMM sources in 7.1deg^2
- For each XMM sou. keeping only tuple with highest P_{max}
- Using "clean" sample:
  - SDSS: `clean=1`
  - VIKING: `pNoise`<0.9 & present in >1 band
- "Clean" xcorr better overall: more XMM with counterparts
  - Because probabilities also depend on density of X+1,2,3cat pairs

	Subsample	SDSS		SDSS clean	
Only X	X___	803	803	717	717
X+1cat	XS__	77	26	119	11
	X_V_		40		92
	X__W		11		16
X+2cat	XSV_	360	324	419	310
	X_VW		34		108
	XS_W		2		1
X+3cat	XSVW	576	576	516	561

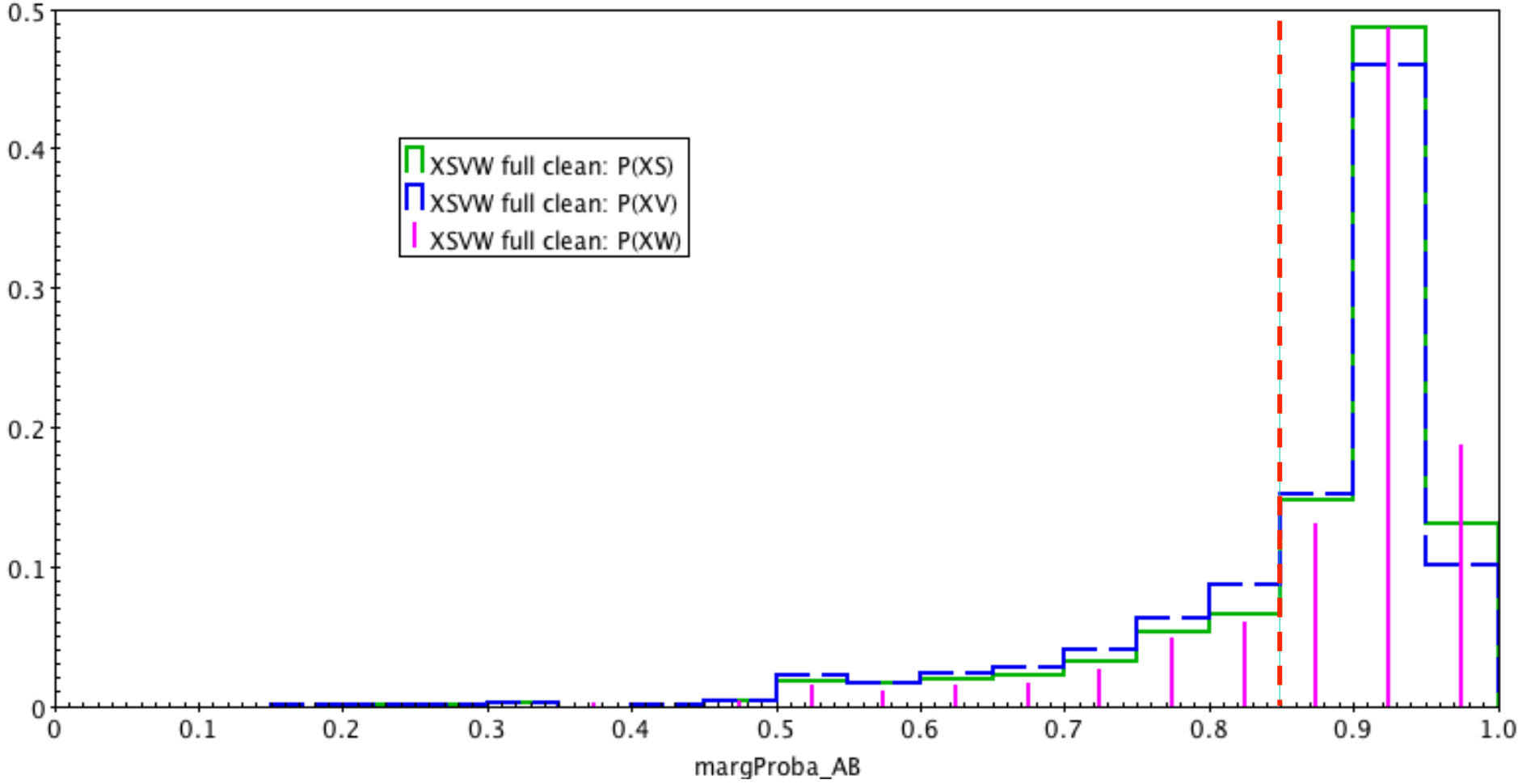
SDSS full clean: probability threshold

- Choosing threshold 85%: drop in P_{max} X+2,3cat



SDSS full clean: probability threshold

- Choosing threshold 85%: drop in marginalised probabilities too



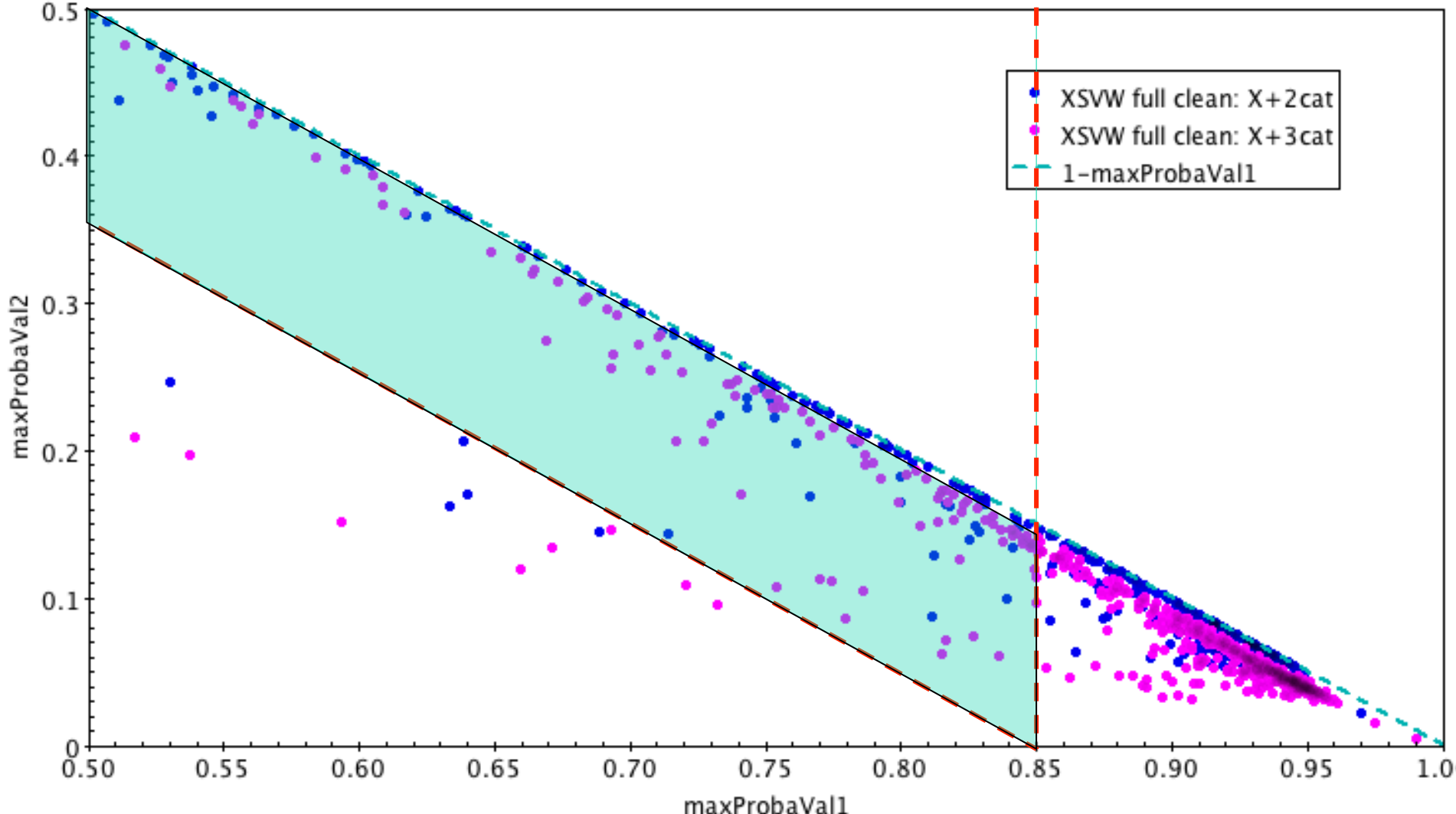
SDSS full clean: prob.≥85%

- ~2/3 of the X+2,3cat survive (by design)

	Subsample	SDSS clean prob.≥85%		SDSS clean	
Only X	X___	1049	1049	717	717
X+1cat	XS__	24	5	119	11
	X_V_		15		92
	X__W		4		16
X+2cat	XSV_	296	216	419	310
	X_VW		80		108
	XS_W		0		1
X+3cat	XSVW	447	447	516	561

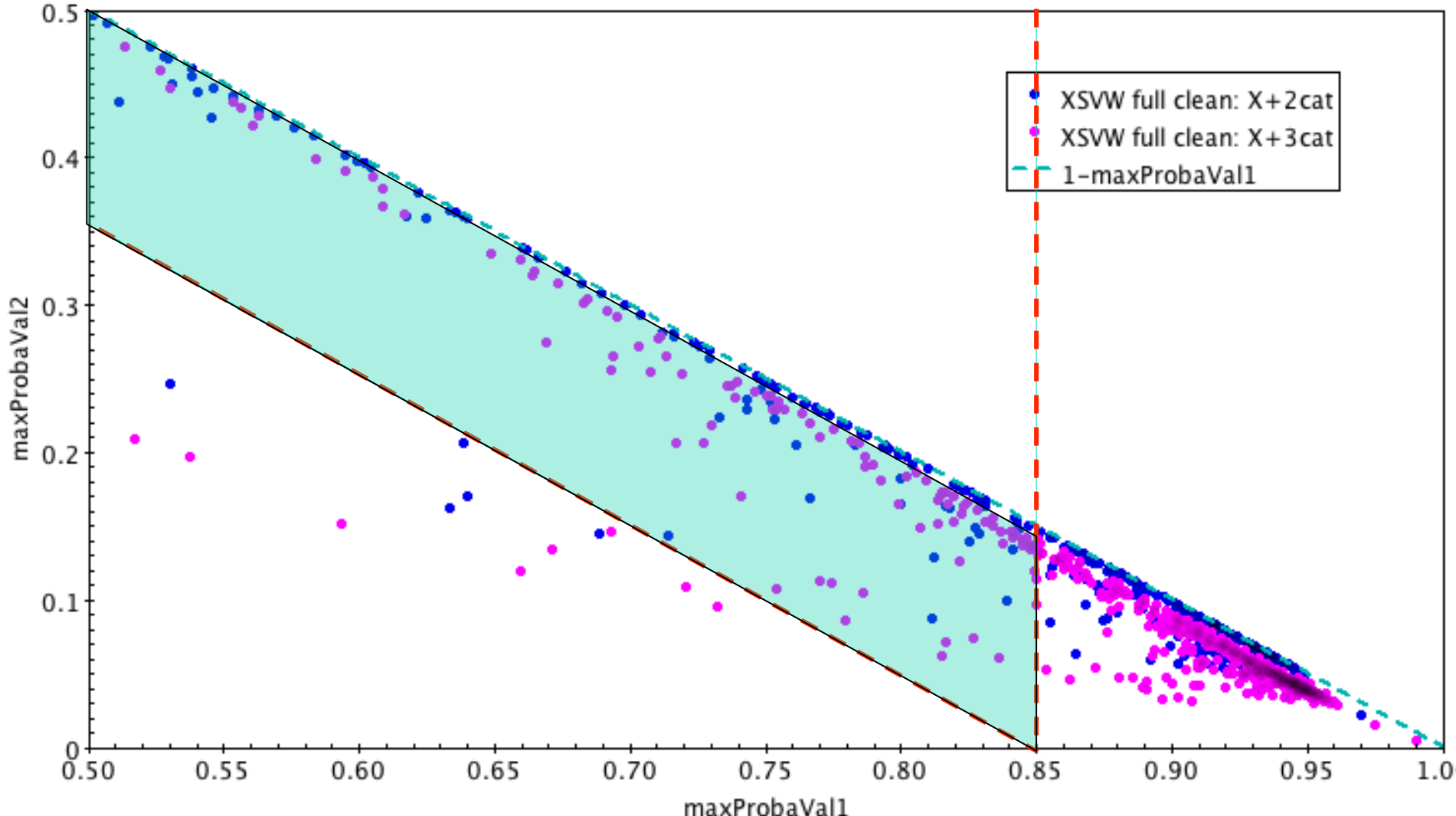
SDSS full clean: best tuple, 1st & 2nd

- For the tuple with the highest P_{max}, comparing P_{max} and second hightest prob. P_{2nd}=`maxProbaVal2`
- What do we do about those with $P_{max}+P_{2nd} \geq 0.85$?



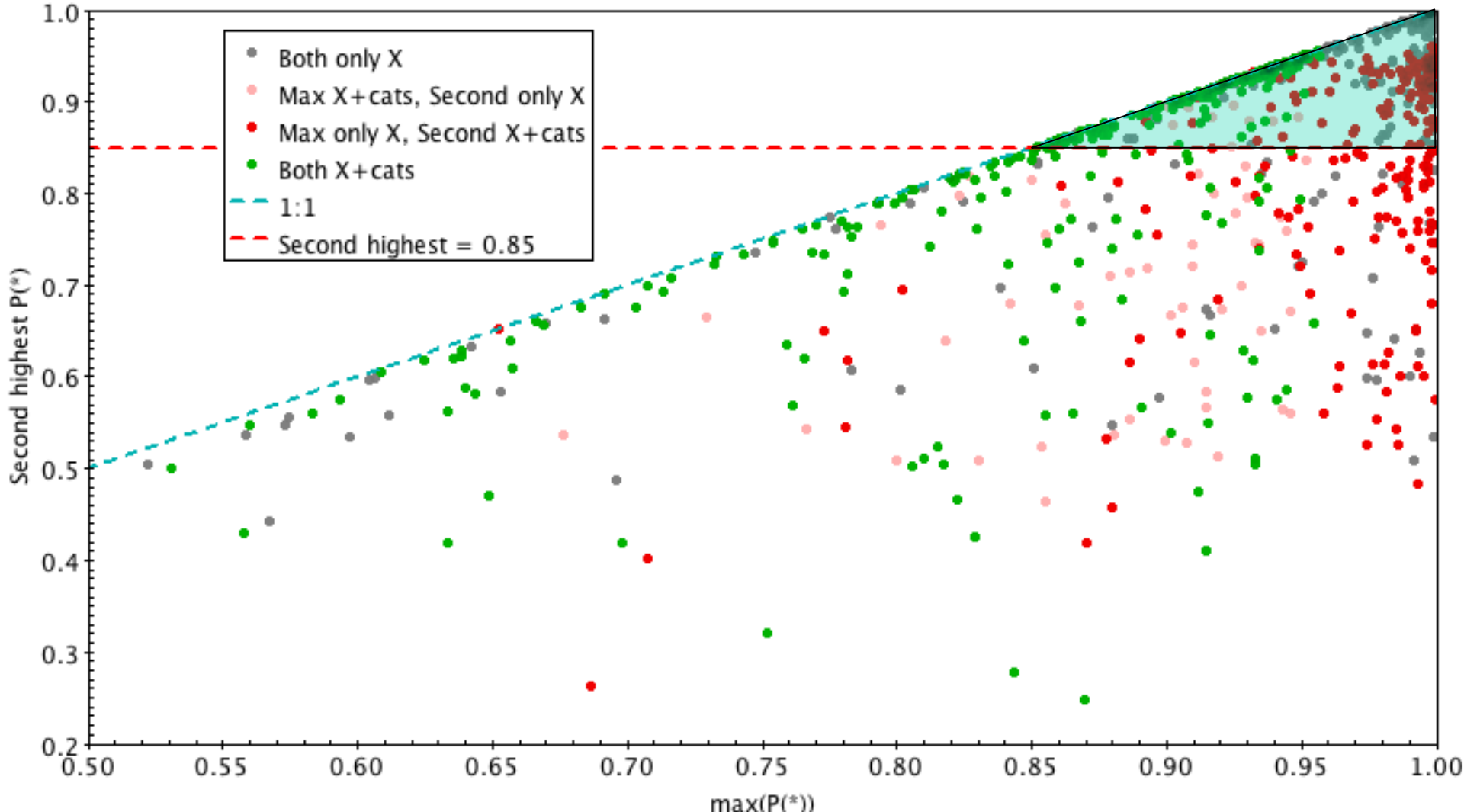
SDSS full clean: best tuple, 1st & 2nd

- What do we do about those with $P_{max}+P_{2nd} \geq 0.85$?
  - If compatible, could boost total probability above threshold: 0 cases
  - If not compatible, could change P_{max} combination: always ~ABC, A_BC



SDSS full clean: 1st & 2nd best tuples

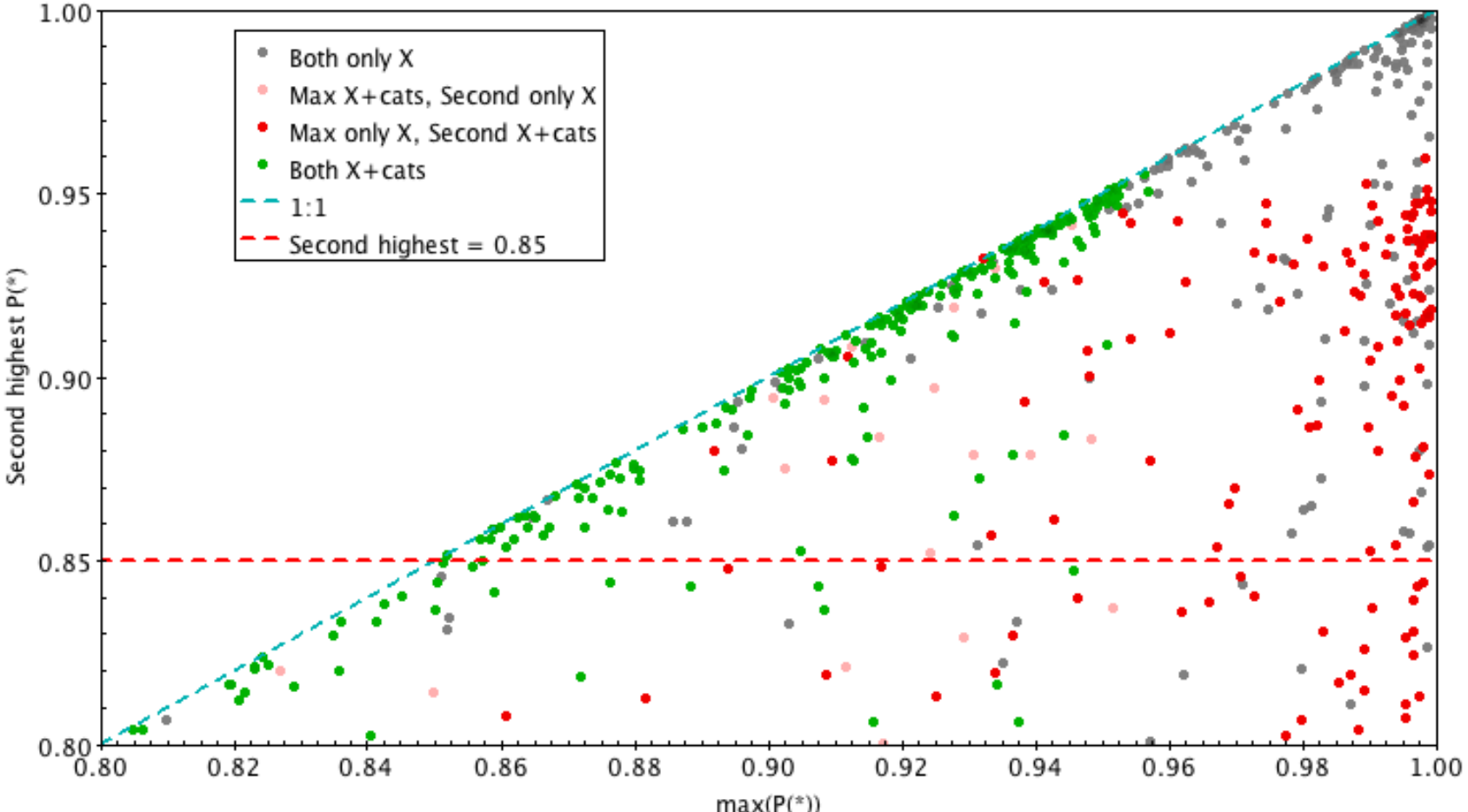
- For each XMM source, comparing tuples with 2 highest P_{max}: 801 unique XMM-ATLAS sources (1015 ≤1 tuple)
- What do we do about those with $P_{max}(2^{nd})$≥0.85?



SDSS full clean: 1st & 2nd best

- Counting:
  - Remember $P_{max} \geq P_{max}(2^{nd})$
  - Worry about $P_{max}(2^{nd}) \geq 0.85$?

1st	2nd	N	$P_{max}(2^{nd}) \geq 0.85$
Only X	Only X	230	-
X+cat	Only X	60	13
Only X	X+cat	192	97
X+cat	X+cat	319	192
Total		**801**	**302**



SDSS in full area: questions/issues

- Areas not quite matched for VIKING
- At the moment, for each tuple, using just max[P(*)], should we use marginalised probabilities instead? e.g.
  - For probs. from 3 catalogues: `margProba_A*`
  - For probs. from 4 catalogues: P(XSV_)=P(ABC_D)+P(ABCD)
  - In how many these marginalised probs. would change combination?
    - For XMM-ATLAS/SDSS/VIKING/WISE: 0
- At the moment, for each XMM source, using just the tuple with the highest max[P(*)], should we worry about the other tuples?
  - How many of those share SDSS/VIKING/WISE sources?
    - Each source from each catalogue can only belong to one tuple!
    - For XMM-ATLAS/SDSS/VIKING/WISE: most change VIKING, a few change also SDSS
  - In how many the marginalised probs. above would change order?
    - For XMM-ATLAS/SDSS/VIKING/WISE: 1 from X_VW to X__W, diff. VIKING

Conclusions

- Tool works well
- The input needs to be worked on:
  - Matched astrometry
  - Matched sky coverage
- The output needs to worked on:
  - Understand the probabilities and their meaning
  - Choose the one(s) that best suit what is needed
- Future for XMM-ATLAS cross-correlations:
  - Match XMM astrometry?
  - Match sky coverages (XMM-SDSS-VIKING-WISE, add KiDS)
  - Repeat cross-correlations
  - Marginalise 3,4 catalogue probabilities?
  - Get SEDs (ARCHES tool too)
  - Get photo-z (+ U. Napoli: machine learning)

4 catalogues: marginalised probs.

- X-ray-centric: Considering only relationship with XMM source

		Probabilities for `nPos`			
		1	2	3	4
Only X	X___	All	P(A_B) P(A_C) P(A_D)	P(A_BC)+P(A_B_C) P(A_CD)+P(A_C_D) P(A_BD)+P(A_B_D)	P(A_BCD)+P(A_BC_D) +P(A_B_CD)+P(A_B_C_D)+ P(A_BD_C)
X+1cat	XS__	-	`margProba_AB`		
	X_V_	-	`margProba_AC`		
	X__W	-	`margProba_AD`		
X+2cat	XSV_	-	-	P(ABC)	P(ABC_D)+P(ABCD)
	X_VW	-	-	P(ACD)	P(ACD_B)+P(ABCD)
	XS_W	-	-	P(ABD)	P(ABD_C)+P(ABCD)
X+3cat	XSVW	-	-	-	P(ABCD)