



Machine Learning Approach for the Search of High-Confidence Blazar Candidates and their Multiwavelength Counterparts

Sabrina Einecke, Dominik Elsässer, Wolfgang Rhode

European Week of Astronomy and Space Science

Prague

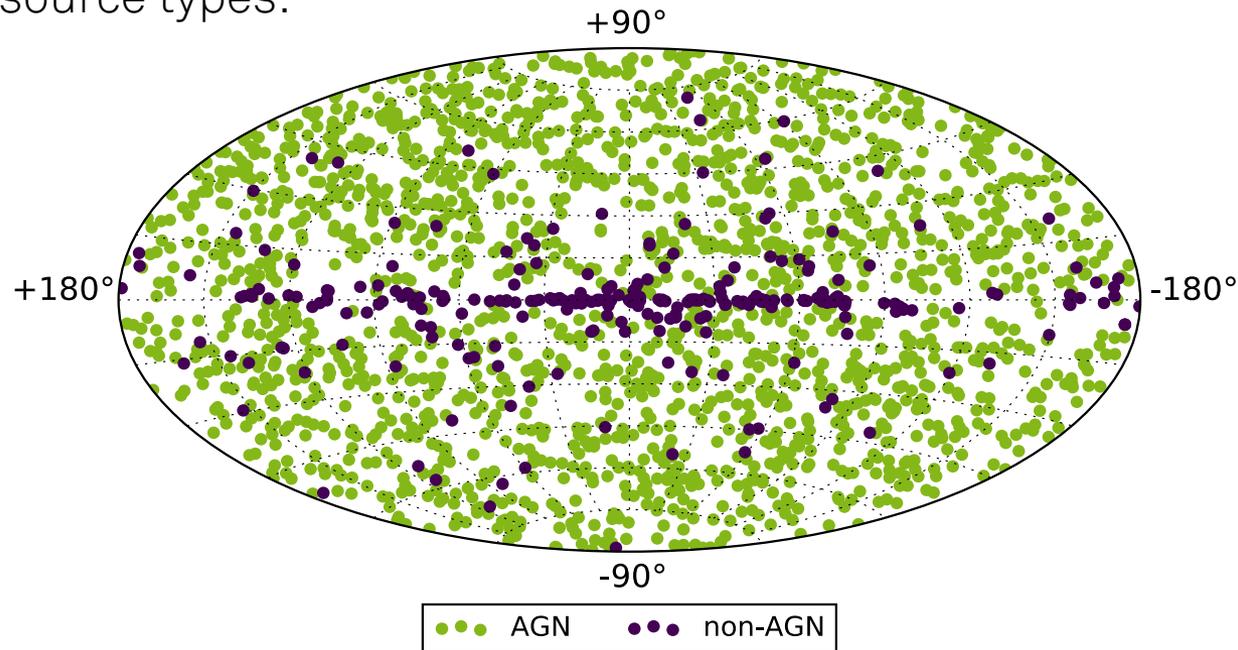
June 29th, 2017

Third *Fermi*-LAT Source Catalog (3FGL)

- Information about fluxes, variability, ...
- 3033 point sources

Sources **associated** with source types:

- 1745 AGNs
- 279 non-AGNs



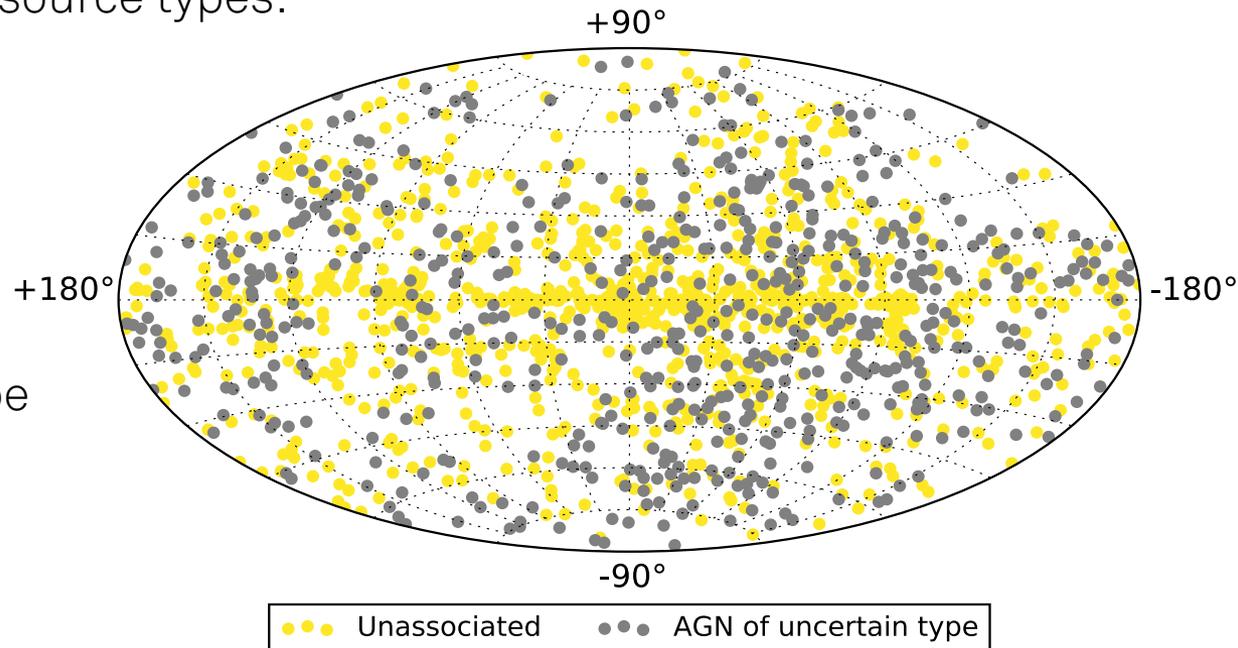
Third *Fermi*-LAT Source Catalog (3FGL)

- Information about fluxes, variability, ...
- 3033 point sources

Sources **associated** with source types:

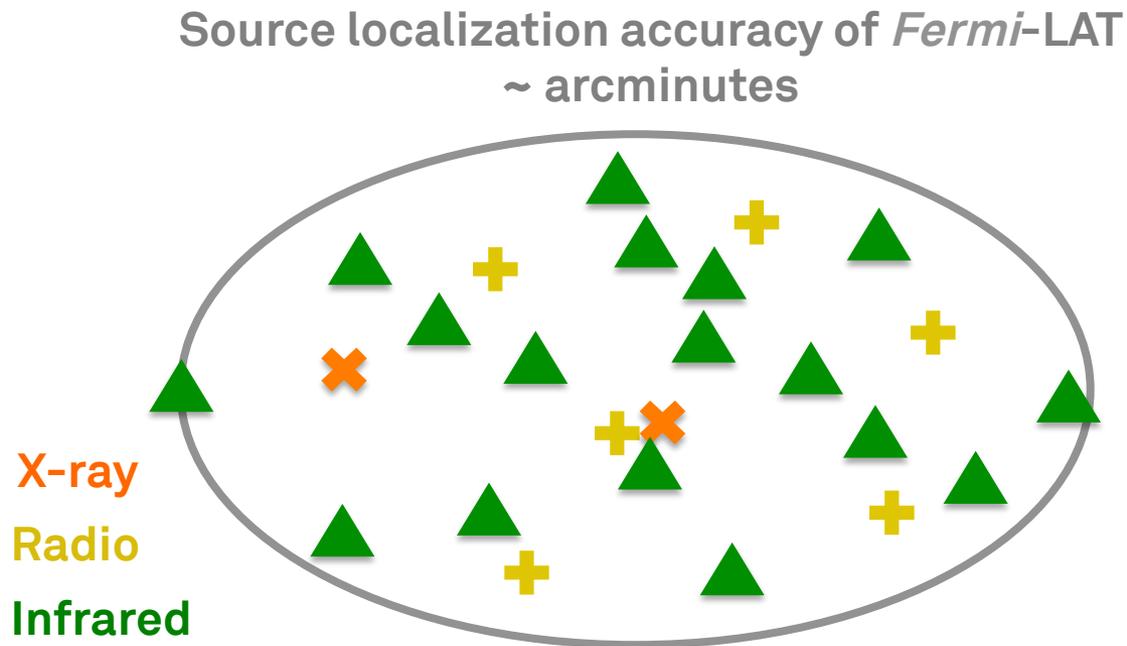
- 1745 AGNs
- 279 non-AGNs

- 1010 **unassociated**
- 573 associated to AGNs of **uncertain** type



Additional Task

Which are the corresponding counterparts in other wavelengths?



Motivation

Idea

- Extension to MWL properties
 - Additional source class-specific characteristics
 - Perfectly suited for machine learning
- Determination of most likely corresponding counterpart at same time

Aim

- Assign AGN classes to unassociated sources
(2 classes: AGN / non-AGN)
- Assign blazar classes to unassociated and uncertain sources
(3 classes: BLL / FSRQ / non-blazar)
- Link counterparts to unassociated sources



Improve knowledge of population of gamma-ray emitting objects

Motivation

Idea

- Extension to MWL properties
 - Additional source class-specific characteristics
 - Perfectly suited for machine learning
- Determination of most likely corresponding counterpart at same time

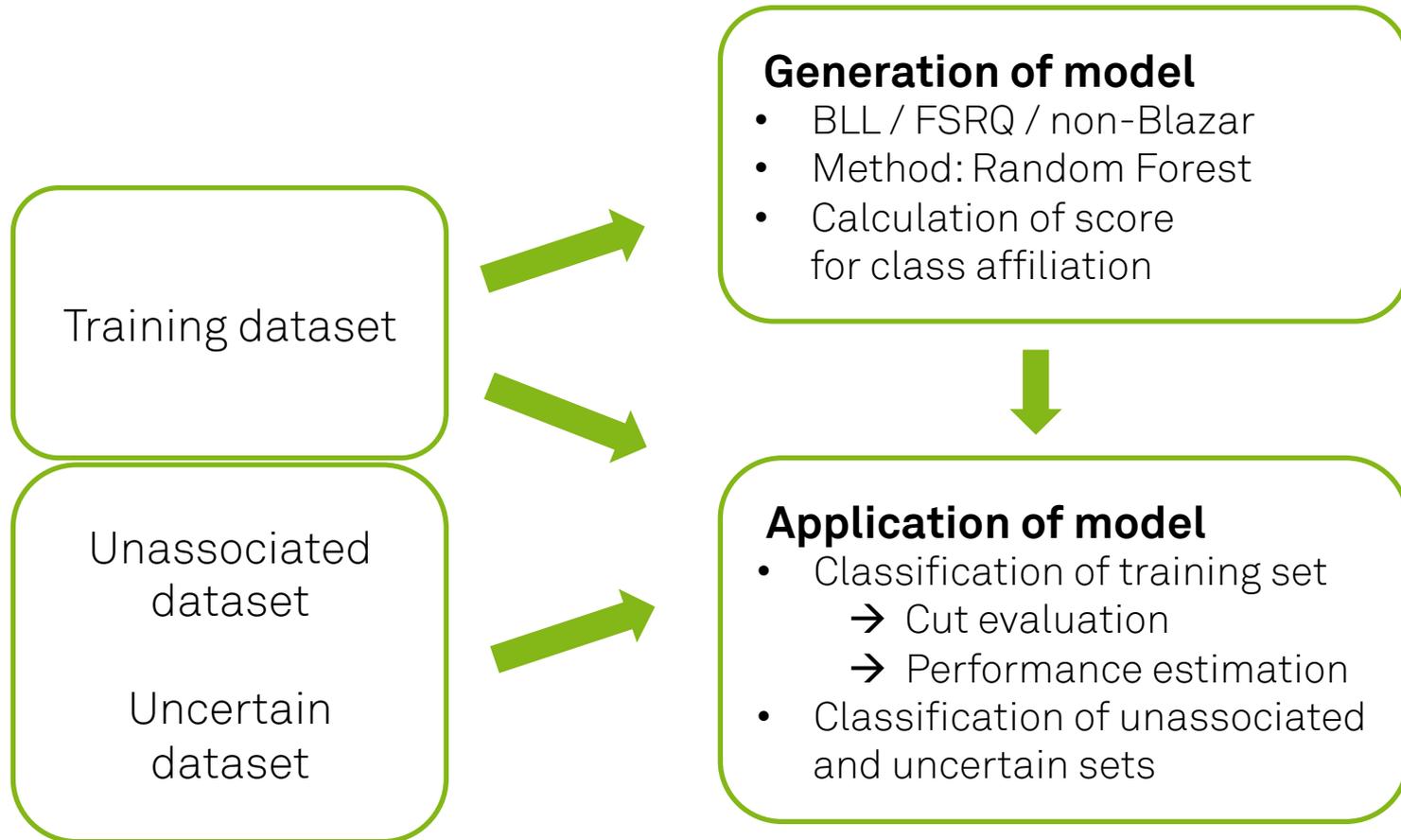
Aim

- Assign AGN classes to unassociated sources
(2 classes: AGN / non-AGN)
- **Assign blazar classes to unassociated and uncertain sources
(3 classes: BLL / FSRQ / non-blazar)**
- **Link counterparts to unassociated sources**



Improve knowledge of population of gamma-ray emitting objects

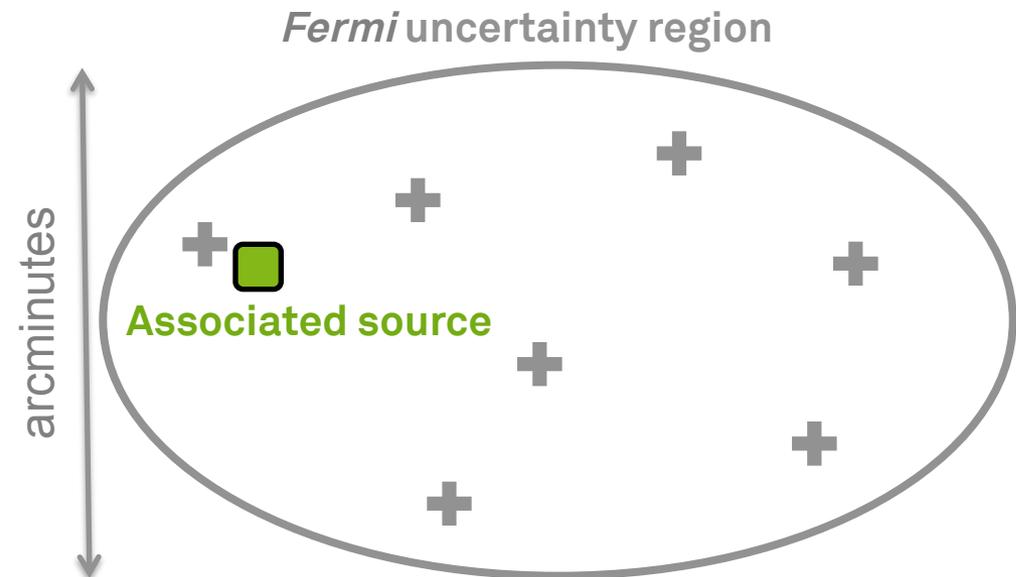
Classification procedure



Creation of Training Dataset

Example: Combination of *Fermi* with *WISE* sources

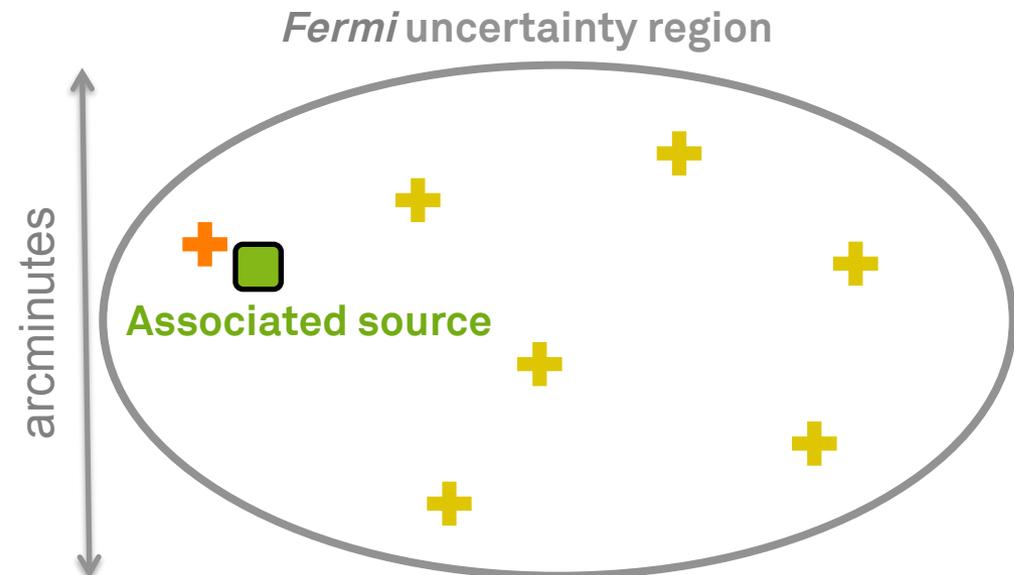
- Search for all *WISE* sources within *Fermi* uncertainty region



Creation of Training Dataset

Example: Combination of *Fermi* with *WISE* sources

- Search for all WISE sources within *Fermi* uncertainty region
- Choose WISE source next to associated source (< 6 arcsec)
→ **Source type class** (BLL, FSRQ)
- Remaining WISE sources → **Background class**



Creation of Training Dataset

Example: Combination of *Fermi* with *WISE* sources

- Search for all WISE sources within *Fermi* uncertainty region
- Choose WISE source next to associated source (< 6 arcsec)
→ **Source type class** (BLL, FSRQ)
- Remaining WISE sources → **Background class**

<i>Fermi</i> name	WISE name	<i>Fermi</i> features	WISE features	Combined features	Class
Fermi #1	WISE #1	BLL
Fermi #1	WISE #2	Non-Blazar
Fermi #1	WISE #3	Non-Blazar
Fermi #1	WISE #4	Non-Blazar
Fermi #2	WISE #10	FSRQ
Fermi #2	WISE #11	Non-Blazar

Machine learning

- **Multiclass** problem: BLL vs. FSRQ vs. non-Blazar
- Classification based on several features:
spectral index, spectral curvature, variability, hardness ratio, flux, ...
- **Feature Selection:**
 - Removal of correlated features
(Spearman's rank corr.)
 - Recursive feature elimination
(Mean AUC as score)
- Cross **validation:**
Calculation of performance values,
e.g. AUC, purity, efficiency,
with One-Vs-Rest strategy

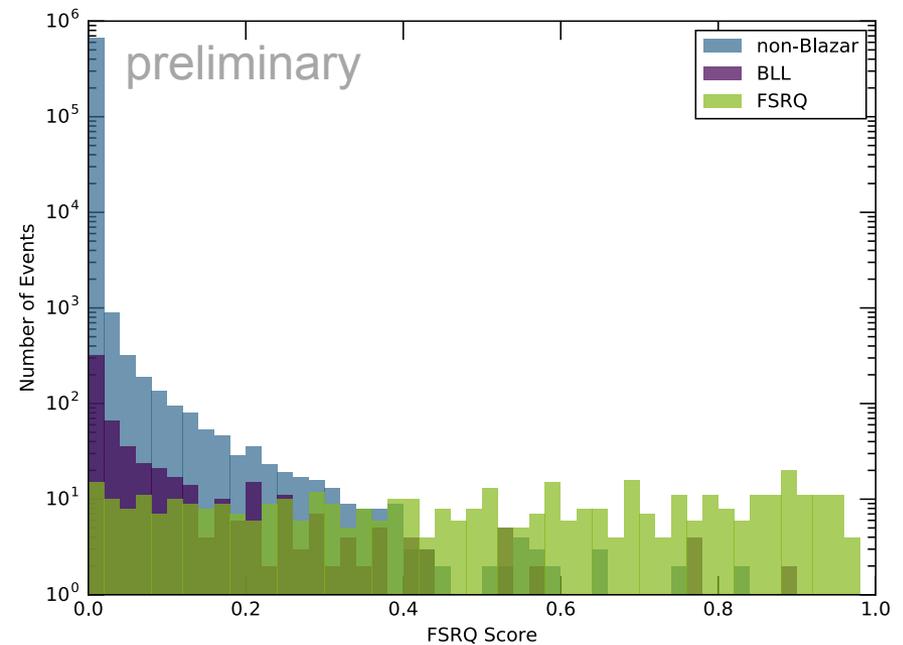
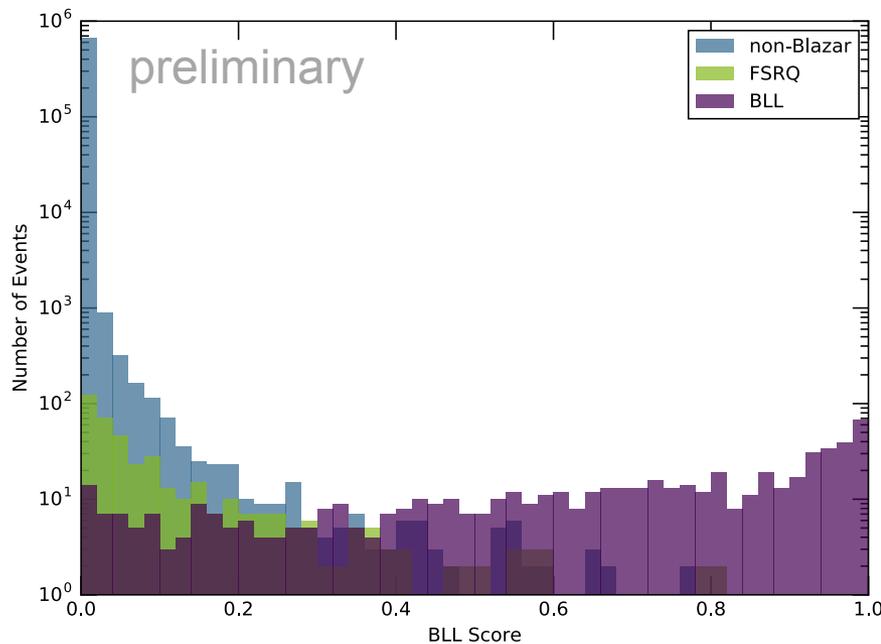


Classification of Training Set

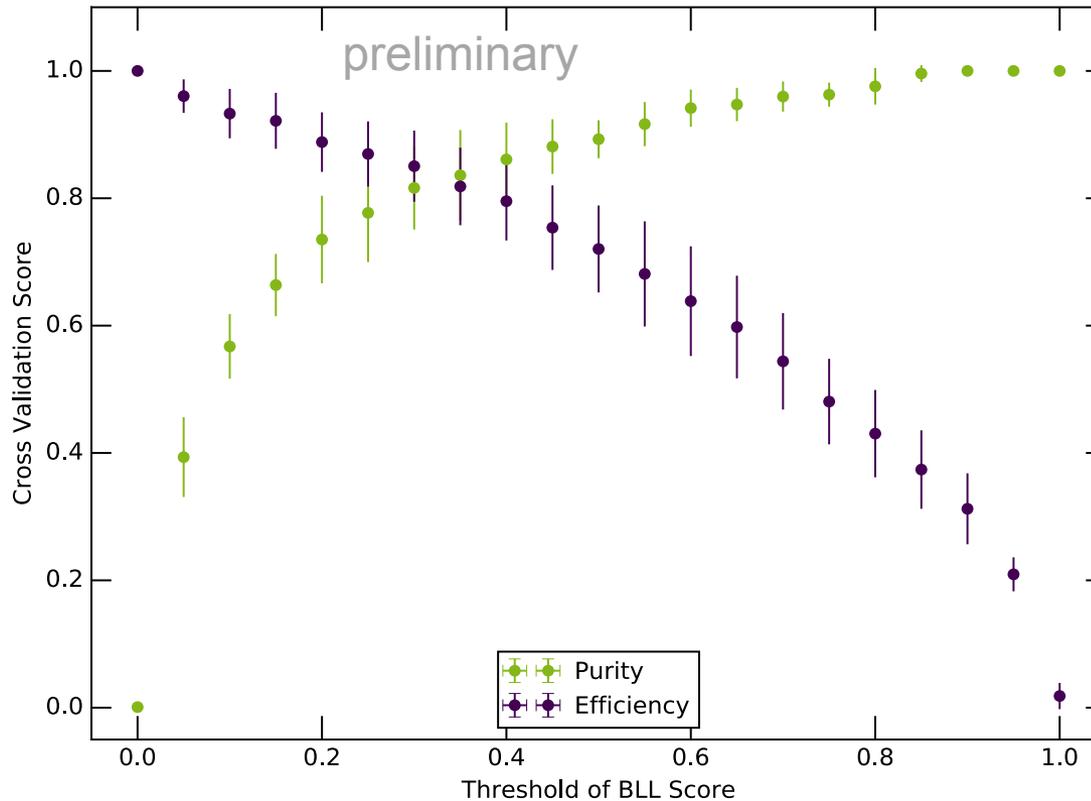
Combination of *Fermi* with *WISE* sources

ALLWISE source catalog

- 3.4, 4.6, 12 and 22 μm
- 747 634 026 sources



Performance Estimation



Purity:
Ratio of correctly and
falsely classified
BLLs

Efficiency:
Ratio of correctly
classified BLLs and
all BLLs

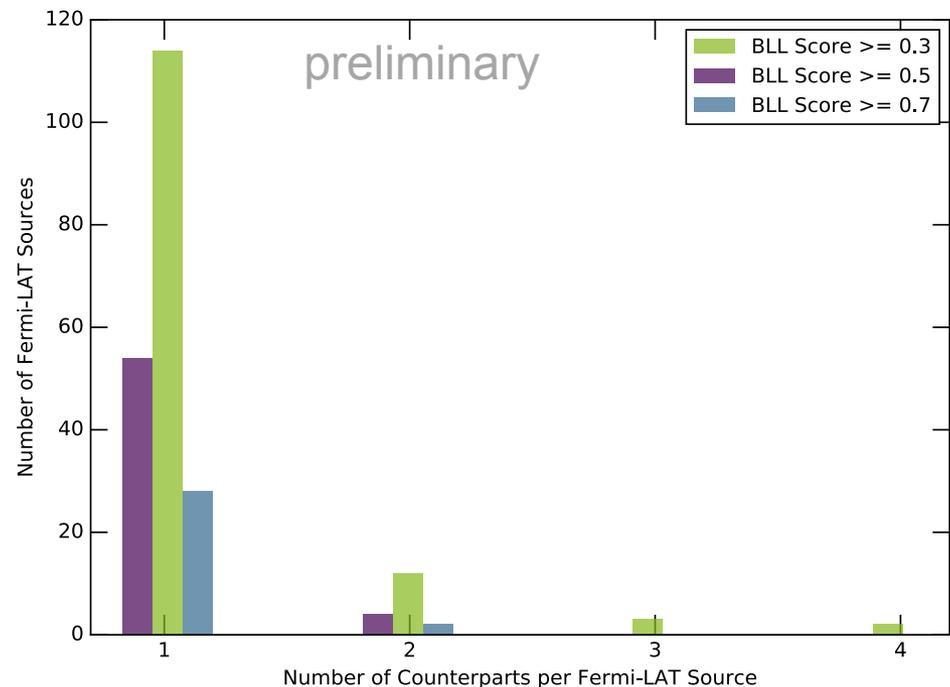
Choice of Score Cut

- Considering one additional waveband (Gamma+Infrared):

High score threshold for **large purity**

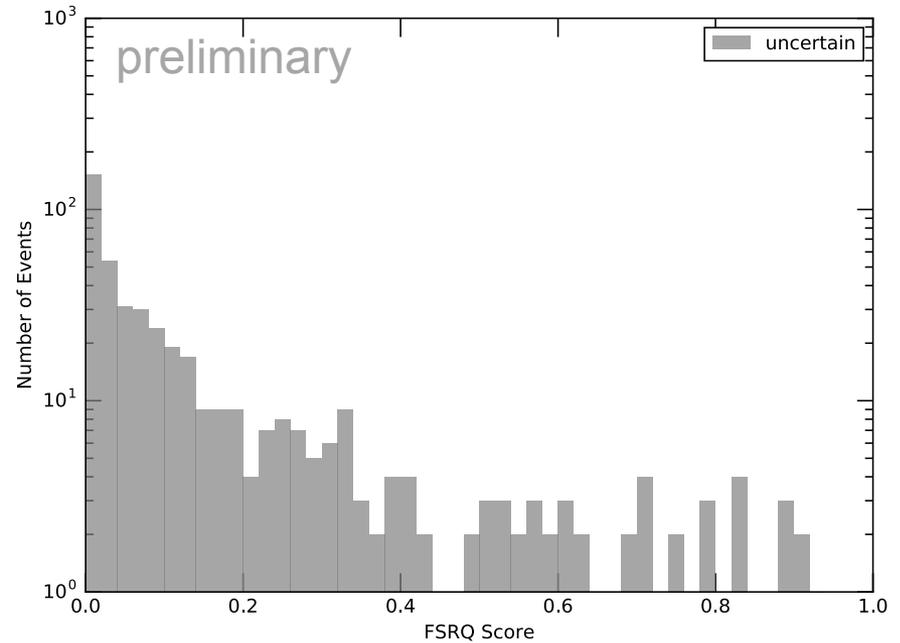
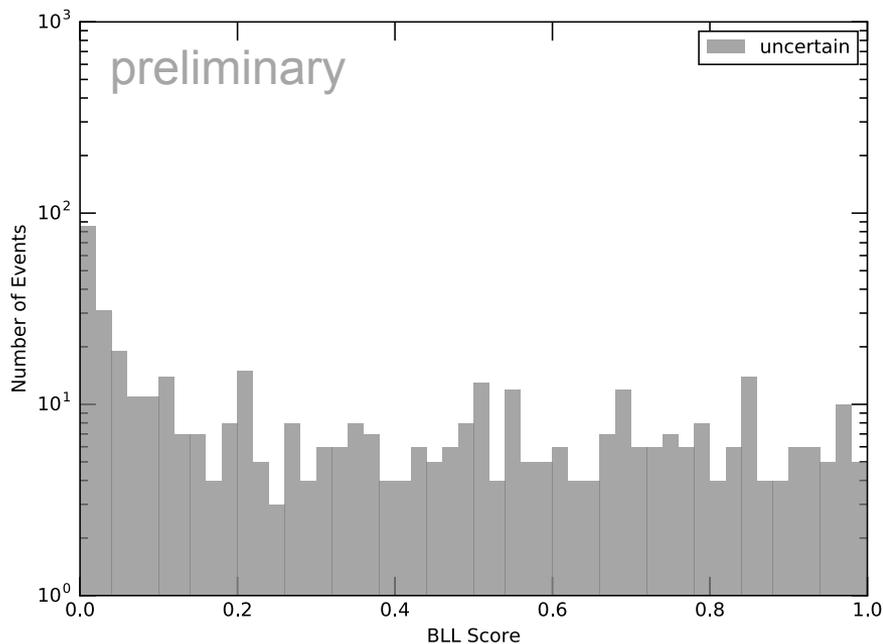
- Considering multiple additional wavebands (Gamma+Infrared, Gamma+X-ray, Gamma+Radio, ...):

Low score threshold for **large efficiencies**



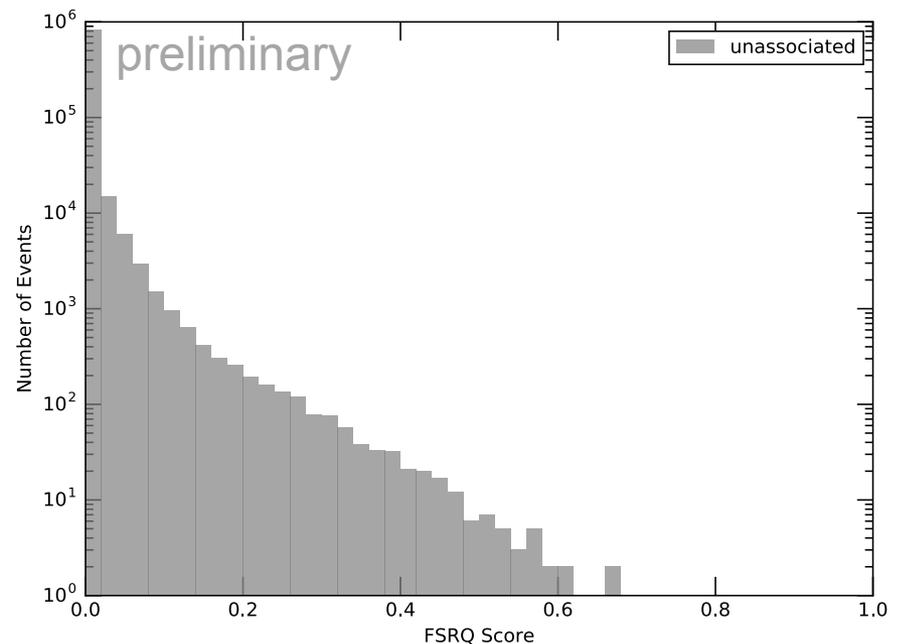
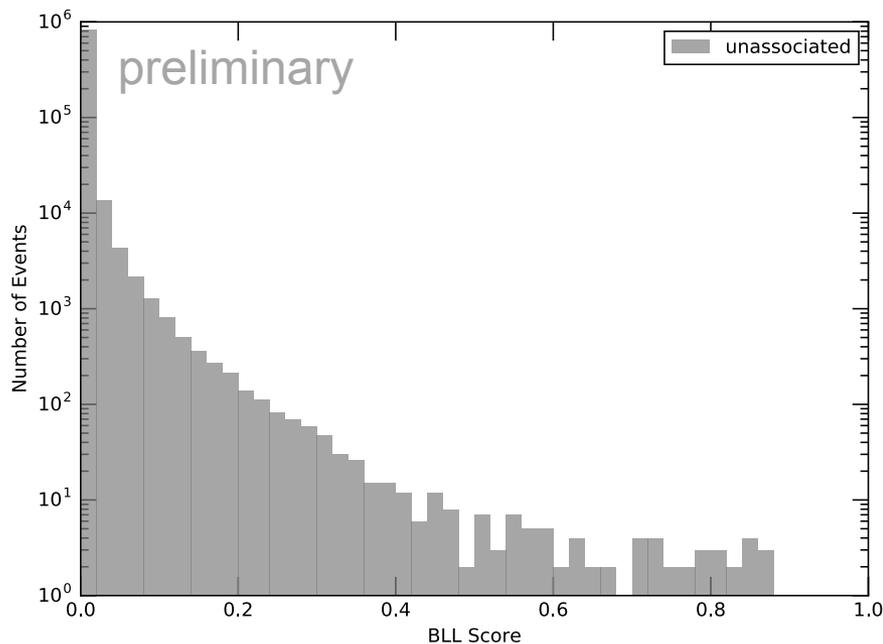
Classification of AGNs of Uncertain Type

- BLL score > 0.3
 - Precision $\sim 80\%$
 - 229 candidates
- FSRQ score > 0.5
 - Precision $\sim 80\%$
 - 43 candidates



Classification of Unassociated Sources

- BLL score > 0.3
 - Precision $\sim 80\%$
 - 243 candidates
- FSRQ score > 0.5
 - Precision $\sim 80\%$
 - 28 candidates



Conclusion and Prospects

- Development of new method to assign blazar classes and link multiwavelength counterparts
- BLL / FSRQ / non-blazar and AGN / non-AGN model creation and performance evaluation
- Individual models for particular catalogs
- Combination of multiple models from different wavelengths and instruments
- Application of models to unassociated sources and AGNs of uncertain type
- Prospects:
 - Dark matter searches
 - Populations studies

→ **Very Promising!**

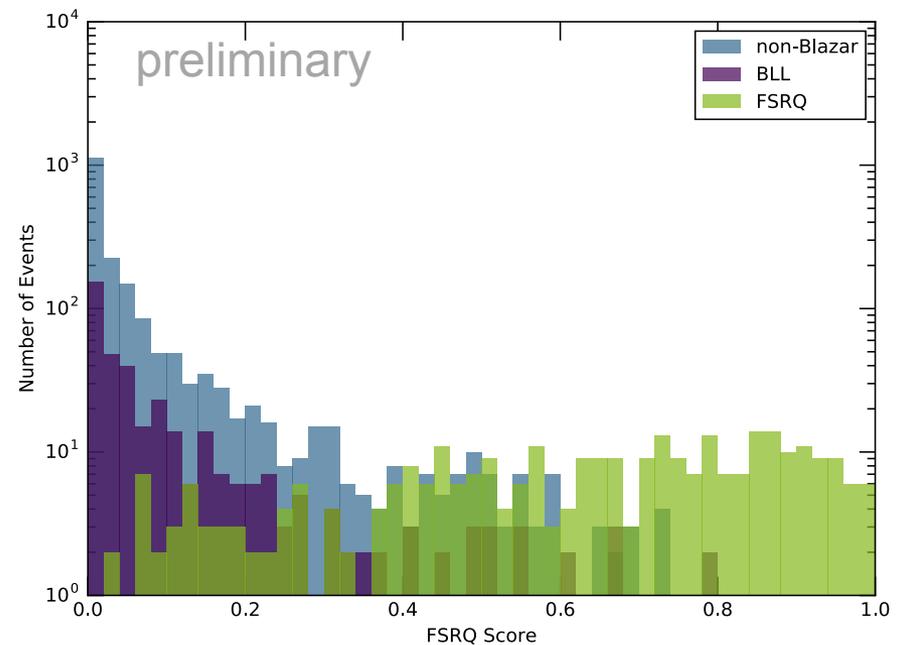
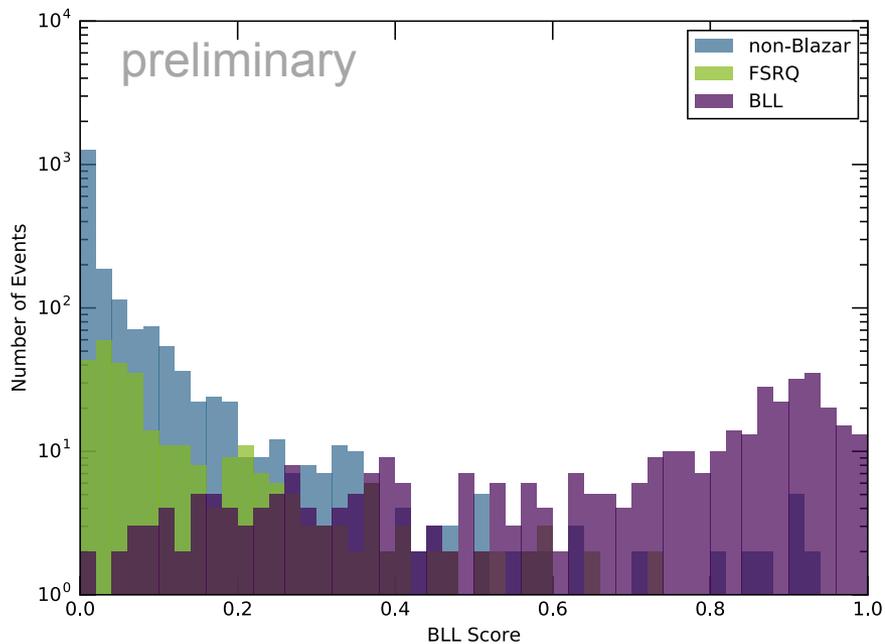
Backup

Classification of Training Set

Combination of *Fermi* with *Swift* X-ray sources

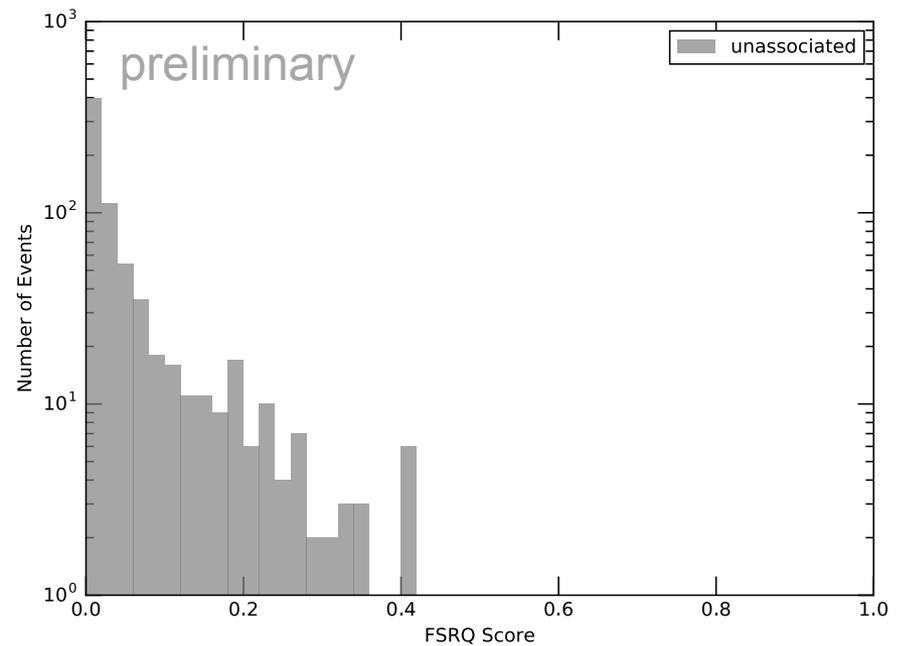
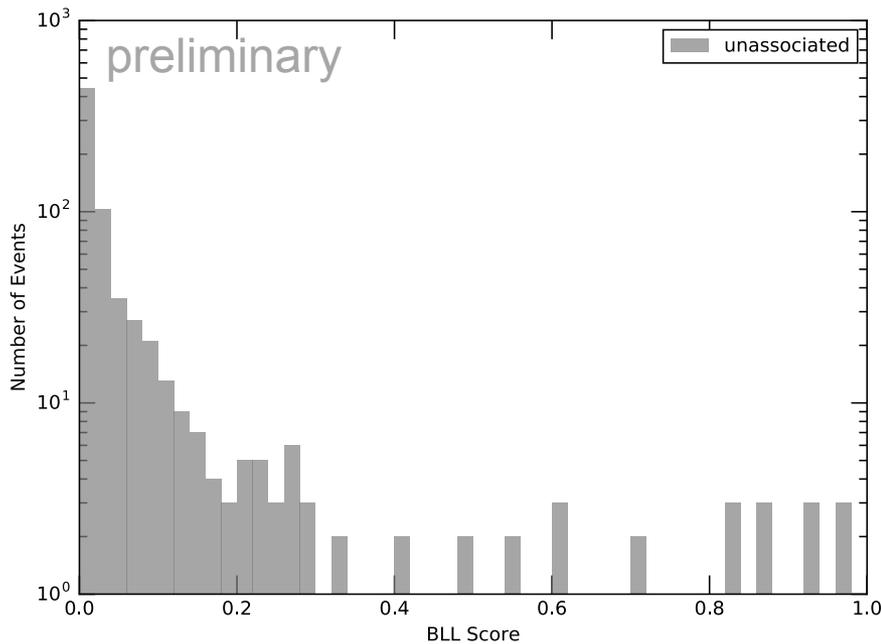
Swift point source catalog 1SXPS

- Energy range: 0.3 – 10 keV
- 151 524 point sources



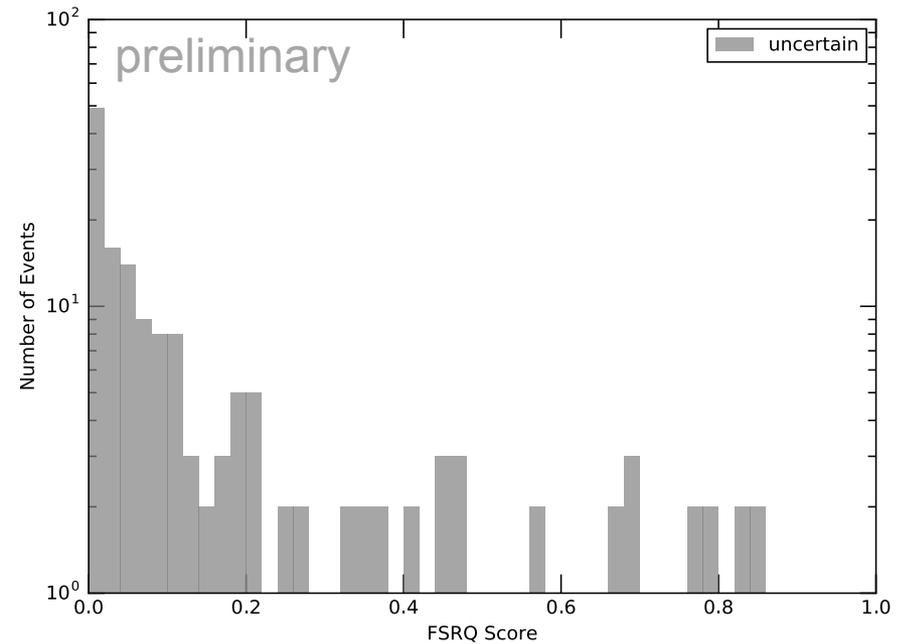
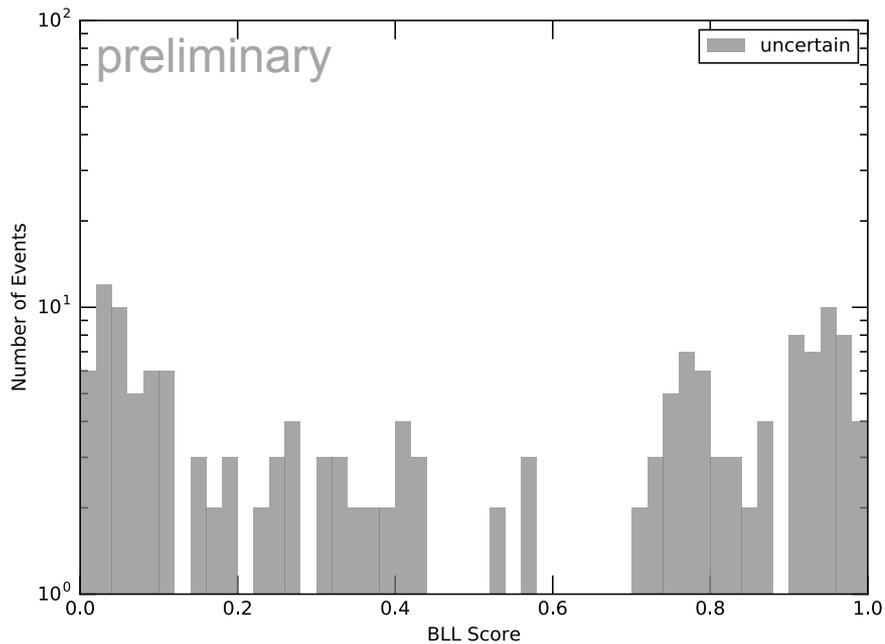
Classification of Unassociated Sources

- BLL score > 0.4
 - Precision $\sim 80\%$
 - 36 candidates
- FSRQ score > 0.5
 - Precision $\sim 80\%$
 - 2 candidates



Classification of AGNs of Uncertain Type

- BLL score > 0.4
- Precision $\sim 80\%$
- 90 candidates
- FSRQ score > 0.5
- Precision $\sim 80\%$
- 23 candidates



Combination of Results

Uncertain:

- **BLL:** 71 high-confidence candidates
from 90 (1SXPS) and 229 (ALLWISE) low-confidence candidates
- **FSRQ:** 17 high-confidence candidates
from 23 (1SXPS) and 43 (ALLWISE) low-confidence candidates

Unassociated:

- **BLL:** 32 high-confidence candidates
from 36 (1SXPS) and 243 (ALLWISE) low-confidence candidates
→ 17 candidates with distance < 20 arcsec
- **FSRQ:** 0 high-confidence candidates
from 2 (1SXPS) and 28 (ALLWISE) low-confidence candidates
→ 0 candidates

→ **Precision of 90% by combination of multiple catalogs!**