FROM PITFALLS TO PROMISE

Present and future of high-energy catalogue cross-correlations.

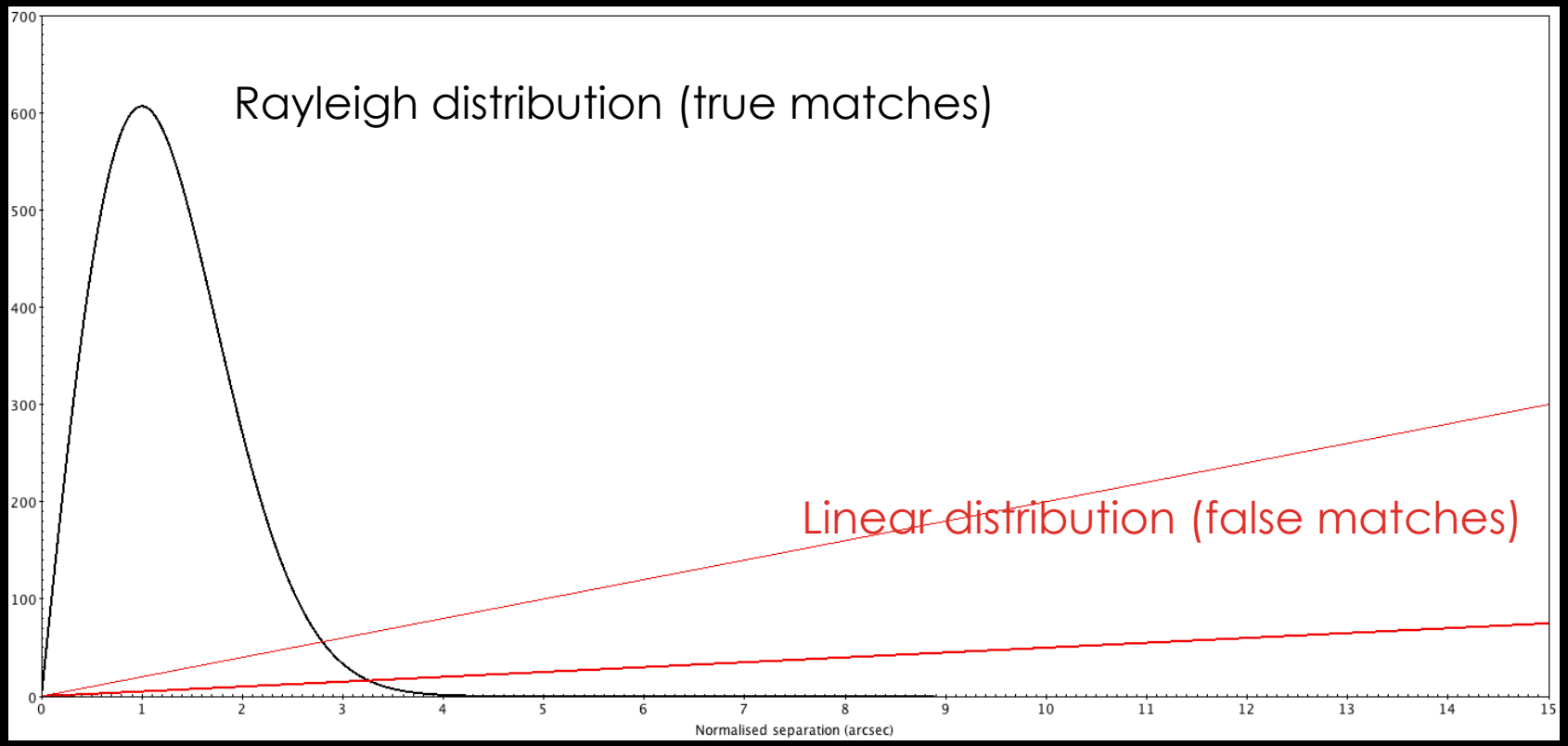
Beatriz Mingo (The Open University)

OUTLINE

- Cross-correlations
  - Do-s and don't-s
  - Why we need a statistical approach
  - Relevance to high energies
- Bias, bias, bias!
- Synergies
  - New and future surveys (beyond high energies)
    - Radio

THE 2-BODY PROBLEM

Theory:
When cross-matching two catalogues, keeping all matches.

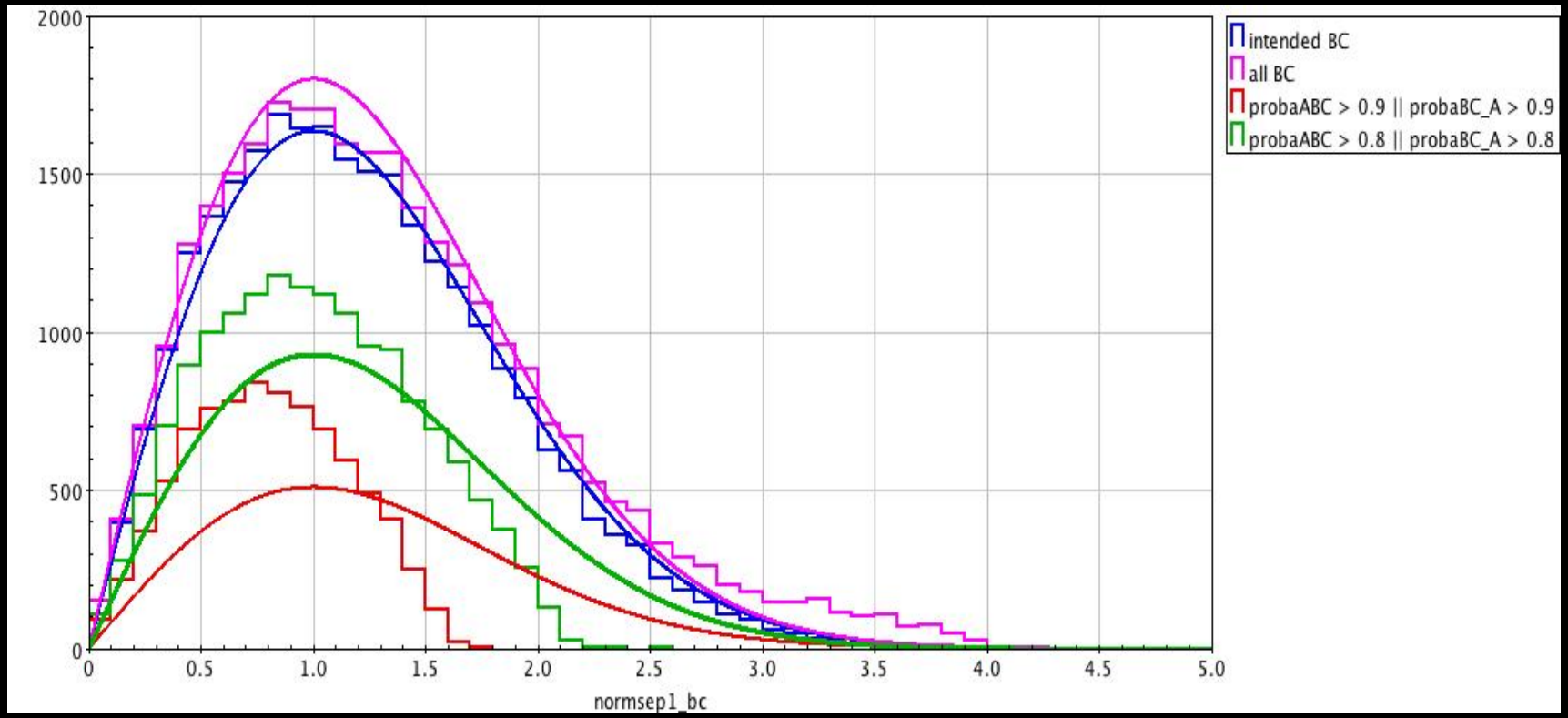


The Open University

Credit: Simon Rosen

THE 2-BODY PROBLEM

Practice:
Simulated catalogues, keeping only first match (true or false)
When cutting at given probability, the distribution changes -> lost matches



2-CATALOGUE X-MATCH

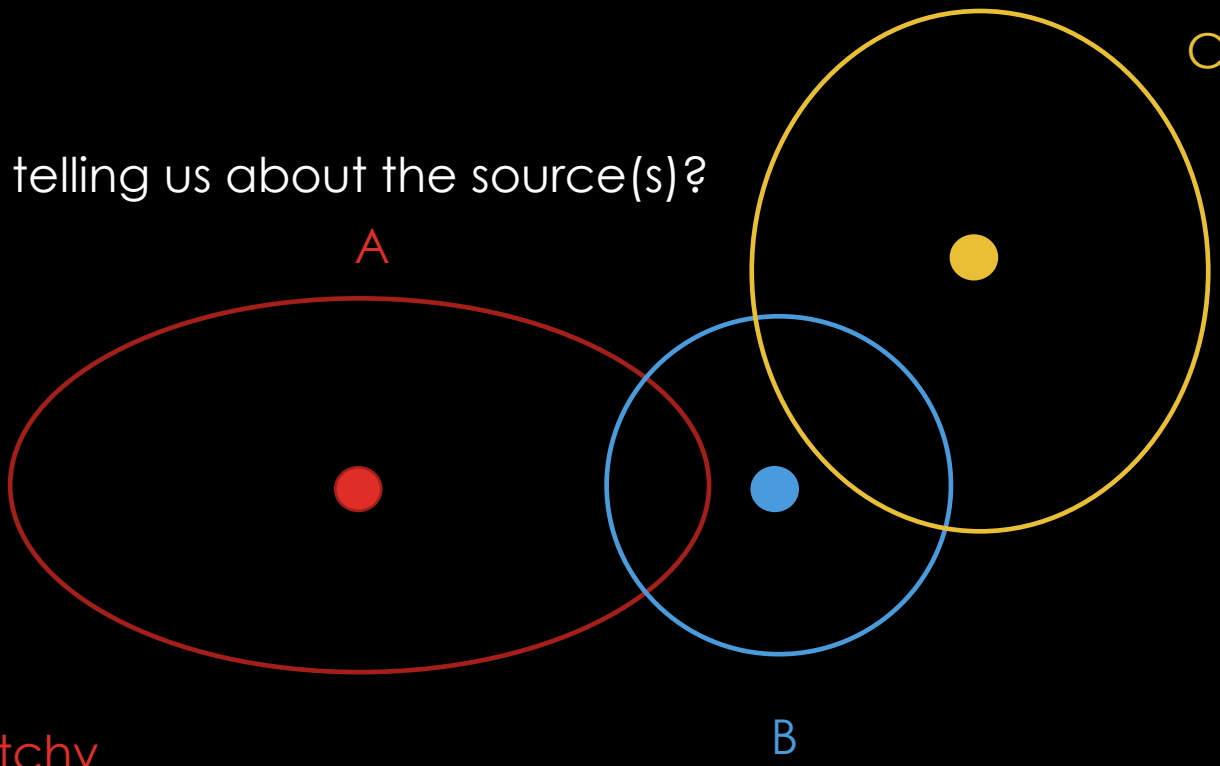
- Great preparatory work for multiple catalogue matches: gives you an idea of bias/systematics.
- Do:
  - Check sky densities of both catalogues
  - Match sky areas (for sky densities + border effects)
  - Check systematics: error variations, coverage fluctuations, astrometry, calibration… (read the documentation!)
  - Carefully assess number of spurious matches
- Don't:
  - Load both tables on TOPCAT, get the nearest match, and run away with it

The Open University

Don't do this by hand!!!!

N-CATALOGUE X-MATCH

A+B=OK
B+C=OK
A+C= ☹
What is this telling us about the source(s)?

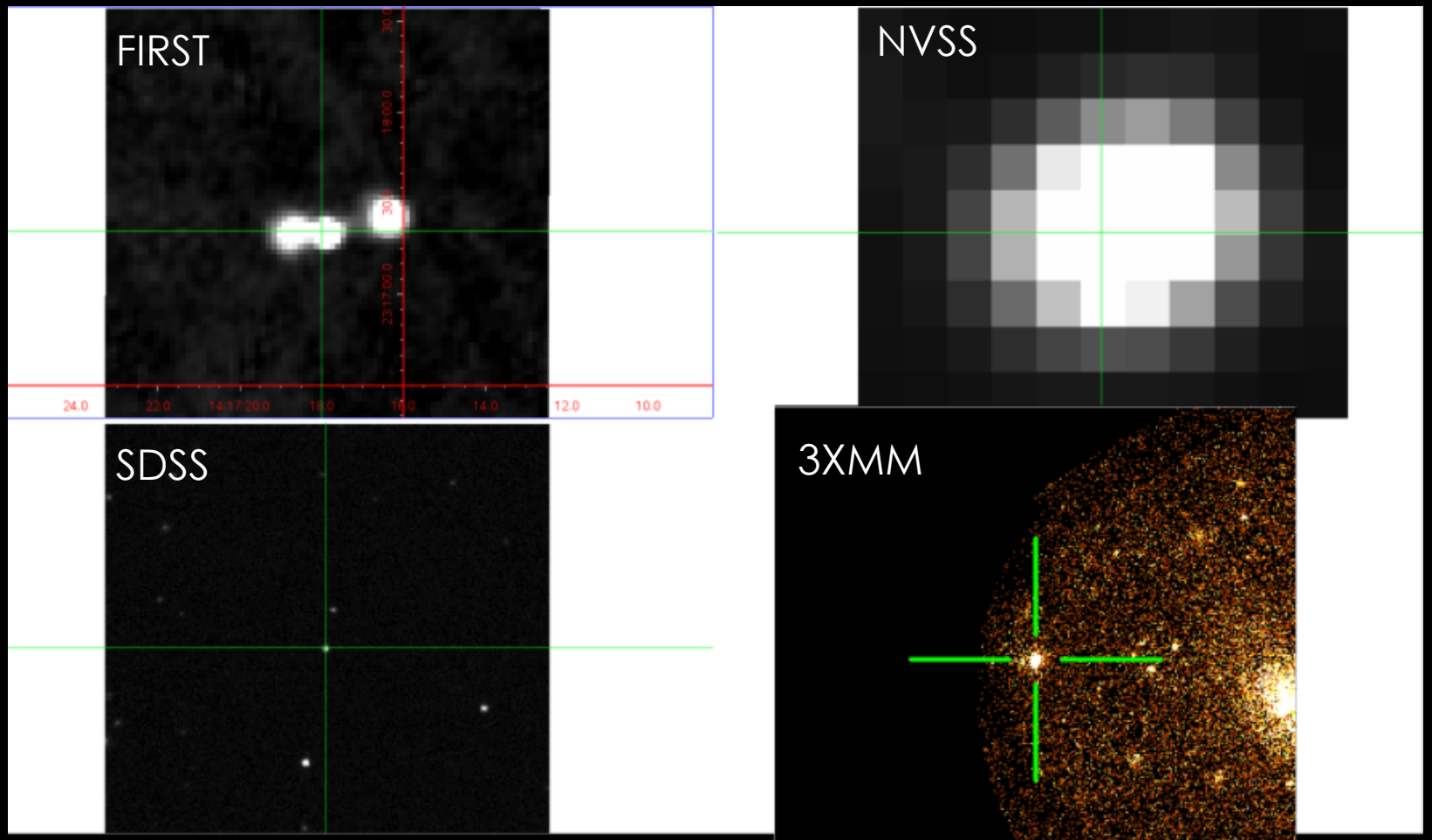


A is patchy
B uses very old/unreliable astrometry
C has RA/DEC dependent errors

The Open University

Mingo+ 2016

IT'S COMPLICATED!!!



(Your experience may vary…)

HIGH ENERGY CATALOGUES

- The good:
  - Relatively low sky density (less confusion)
  - Reliable astrometry
  - Underlying astrophysics quite well-understood (flux modelling is possible) -> all power laws, right? ☺
- The not so good:
  - Patchy (other than RASS, eventually eROSITA)
    - Difficult to establish upper limits, sky areas…
  - Relatively low resolution (other than Chandra)
    - More confusion
  - Still somewhat limited on extended sources

Current high-energy catalogues still hold a wealth of potential science!

STATISTICAL CROSS-MATCHING

- Needed to extend probability calculation to 3+ catalogues
- HE have often led the way
- Simplest: proximity-based probability (weighed by errors)
  - "Top-hat" function (inside all OK, outside none OK)
    - Good at low sky densities, do not impose model bias (though catalogue biases always present)
      - Mann+ 1997, Brusa+ 2005

STATISTICAL CROSS-MATCHING

- Forced photometry
  - Complex, requires raw data, but v. helpful with confusion
- Add astrophysical information (to avoid pairing e.g. stars with galaxies, and minimise saturation effects, bias towards brightest counterparts…)
  - Necessary at higher sky densities, but model-dependent
    - E.g. Sutherland & Saunders 1992, Budavari & Szalay 2008, Naylor + 2013, Wilson & Naylor 2017
  - Magnitude+colour as Bayesian priors:
    - Pineau+ 2015, 2017 → XMatch
    - Salvato+ 2018 → NWAY

See also talk by Dachen Ling – combination of astrophysical information + positional cross-match

CAVEATS

- Statistics != results: don't expect the tools to magically produce science. Results need interpretation
- Beware the bias: even with high-confidence results
  - With astrophysical model-based matching, think about selection effects
  - Catalogue biases are always there! X-match algorithm can't deblend a source for you or eliminate confusion (e.g. Wilson & Naylor 2017), turn an upper limit into a detection, or magically eliminate field borders. No need to re-write catalogue paper, but mention how those biases influence your results.

FUTURE X-MATCHES

- Gaia
  - Improve astrometry, better star/AGN separation (see e.g. Salvato+ 2018), incorporate proper motions
- LSST
  - For all your optical/NIR astrophysical needs
  - Will generate huge volume of data
    - New challenges in terms of variability, source counterparts, extended source matching…
- ELT + precursor surveys (instruments: SINFONI, HIRES, GMOS, FMOS, FLAMES, WEAVE, 4MOST, MOONS…)
  - More optical spectra, yay!

FUTURE X-MATCHES

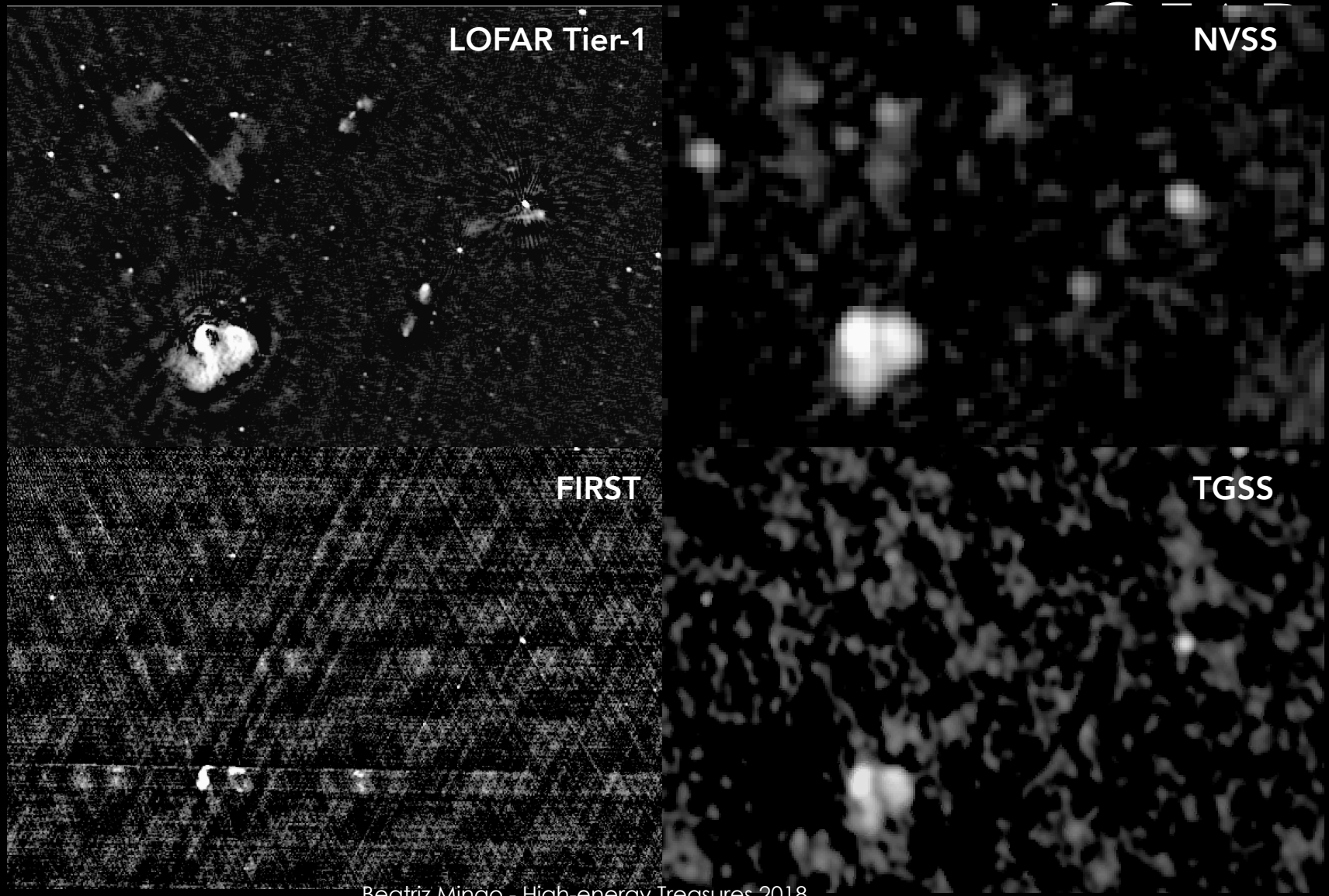
- Euclid?
  - Better cosmology, all-sky, good z estimations, but lower spatial resolution.
- JWST (Hubble 2.0)
  - Deep, exquisite positioning & resolution (no atmosphere!), but won't have the survey capability of LSST
- See also talk by P. Evans about transient identification!
  - The way we are doing x-matching is changing, with higher demand for real-time matches and follow-up

RADIO IS HIGH-ENERGY TOO

- LOFAR LoTSS: ~400 sq. deg. (HETDEX) → all N sky
  - ~325k sources
  - 71% with reliable host IDs, 94% z (limited by optical!)
  - 1st data release imminent! (Shimwell+, Williams+, Duncan+, in prep.), paper splash planned on A&A
- Relevant science:
  - Mapping nearby star formation
  - X-ray cluster/group finding (Croston+ 2017)
  - SNR studies
  - All the AGN science your hearts desire…
  - Galactic/extragalactic B field + particle acceleration studies
- Also Radio Galaxy Zoo, VLBI, eventually SKA…

The Open University

LoTSS: 6" resolution, 100x deeper than FIRST, maps compact +extended emission!

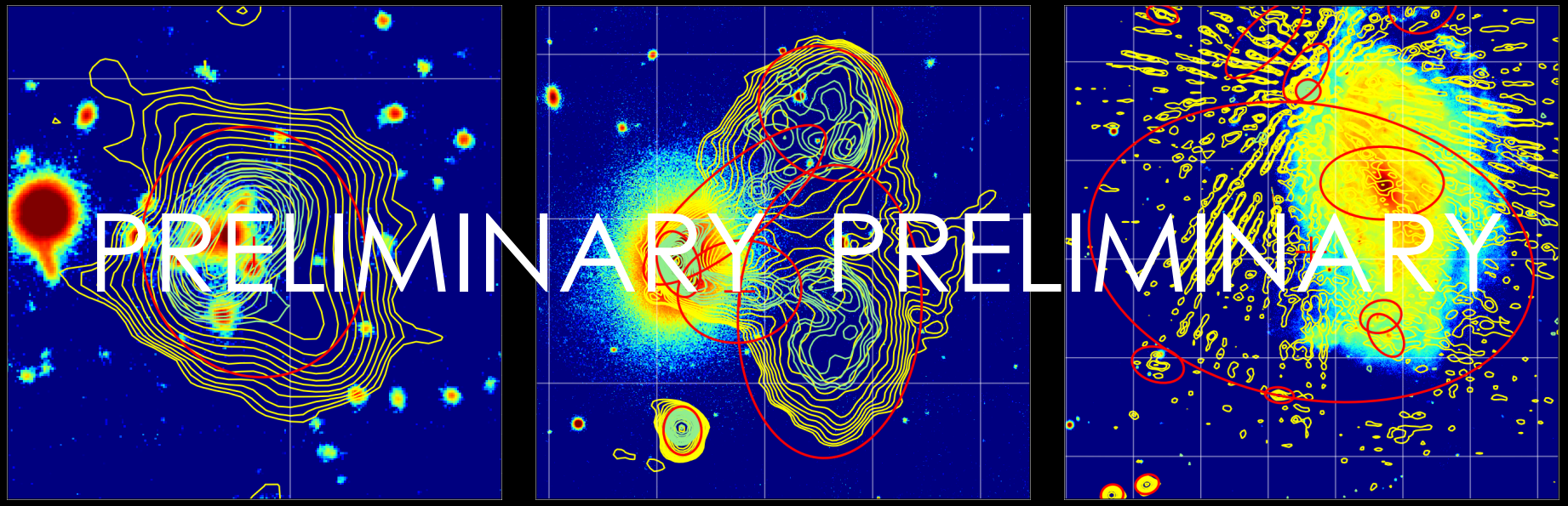


LOFAR GALAXY ZOO

Biggest challenge: find the host galaxies!

Easier for compact sources (maximum likelihood).

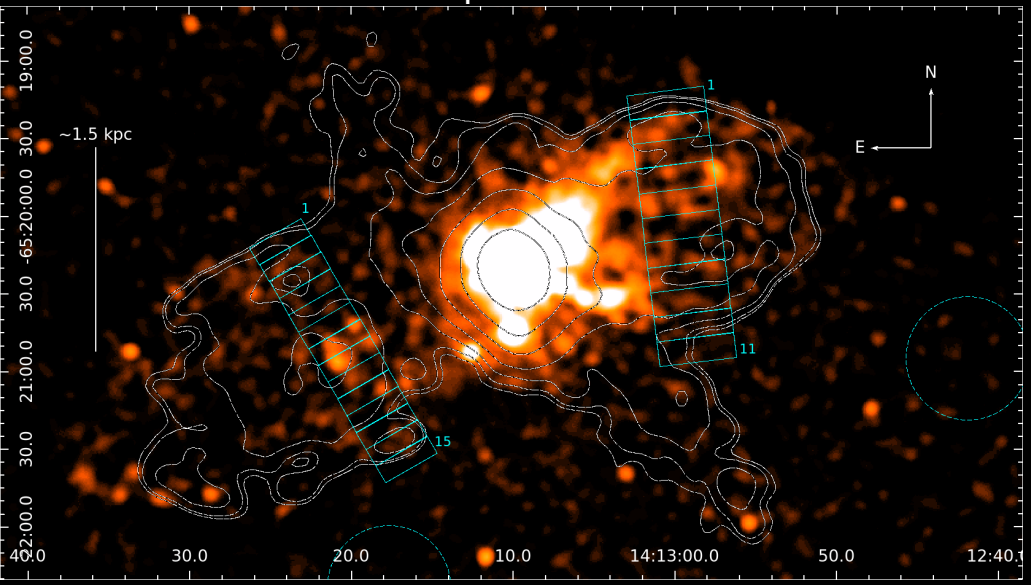
For extended (>12") sources: LOFAR Galaxy Zoo (LGZ)



LOFAR+FIRST contours, Pan-STARRS images (also WISE)

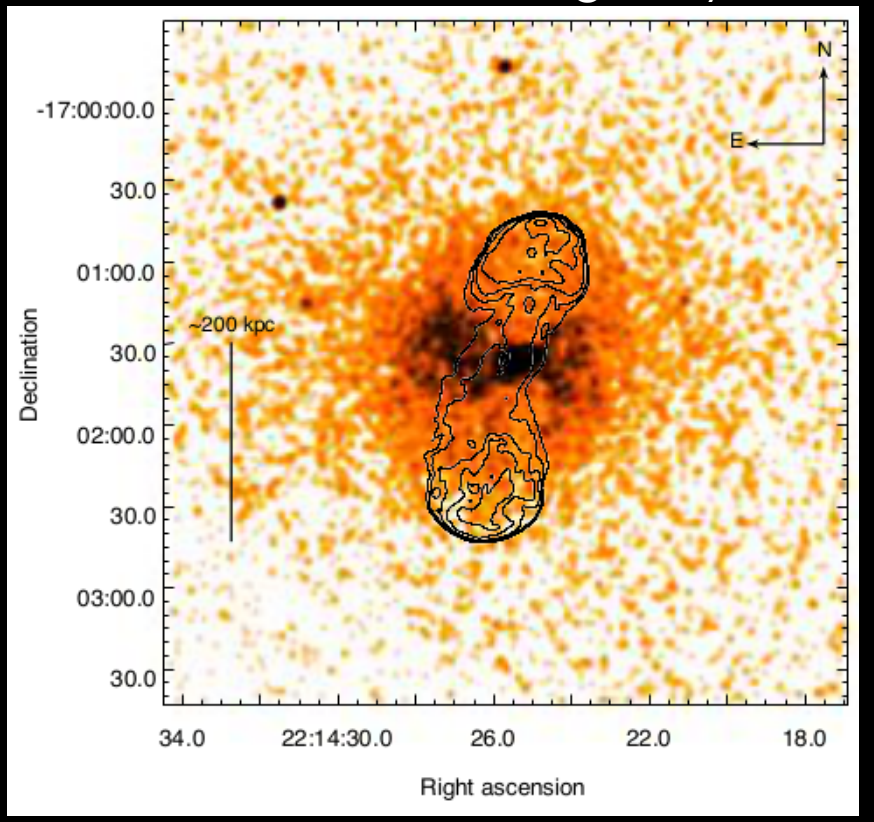
RADIO: UNDERSTAND POPULATIONS

Circinus: a radio-quiet AGN



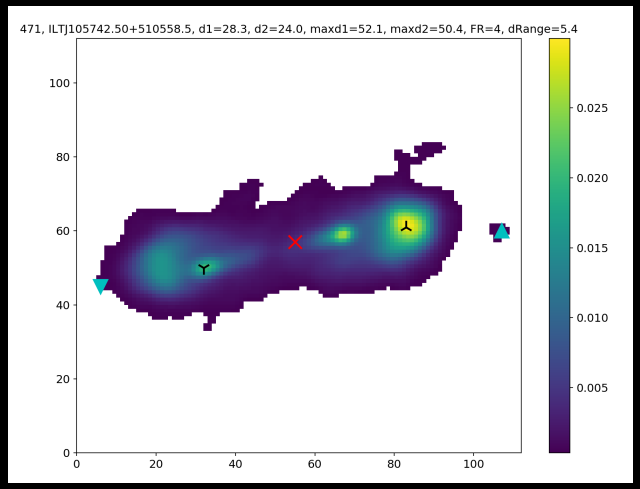
Mingo+ 2012
Hayashida+ 2013 (gamma-ray detection!)

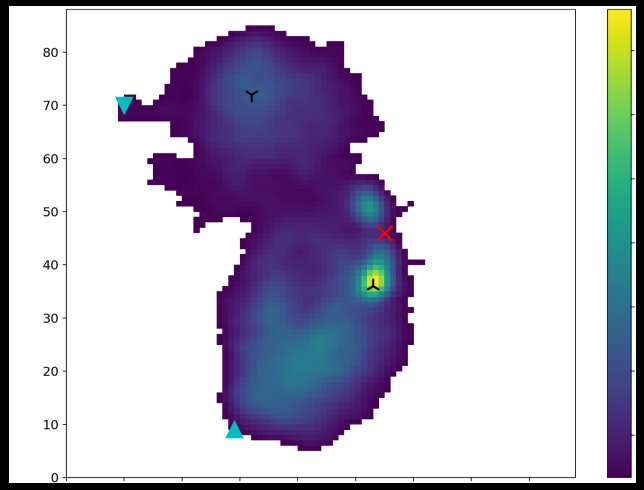
3C 444: a "non-active" galaxy



Croston+ 2011

The Open University





RESTARTERS

Our best candidates to directly constrain AGN cycles!
We need high energy (and optical, MIR) data to determine accretion modes!

CONCLUSIONS

- Multi-catalogue cross-match needs to be tackled carefully & systematically – check your biases!
  - See talks in this session!
- There is a wealth of potential science in current high-energy catalogues – other wavelengths are/will be producing great samples to x-match again!
- Please look at the radio, both to look for new sources and to learn how to deal with big data + extended sources!

Email: bmingo@extragalactic.info
Twitter: @OgNimaeb