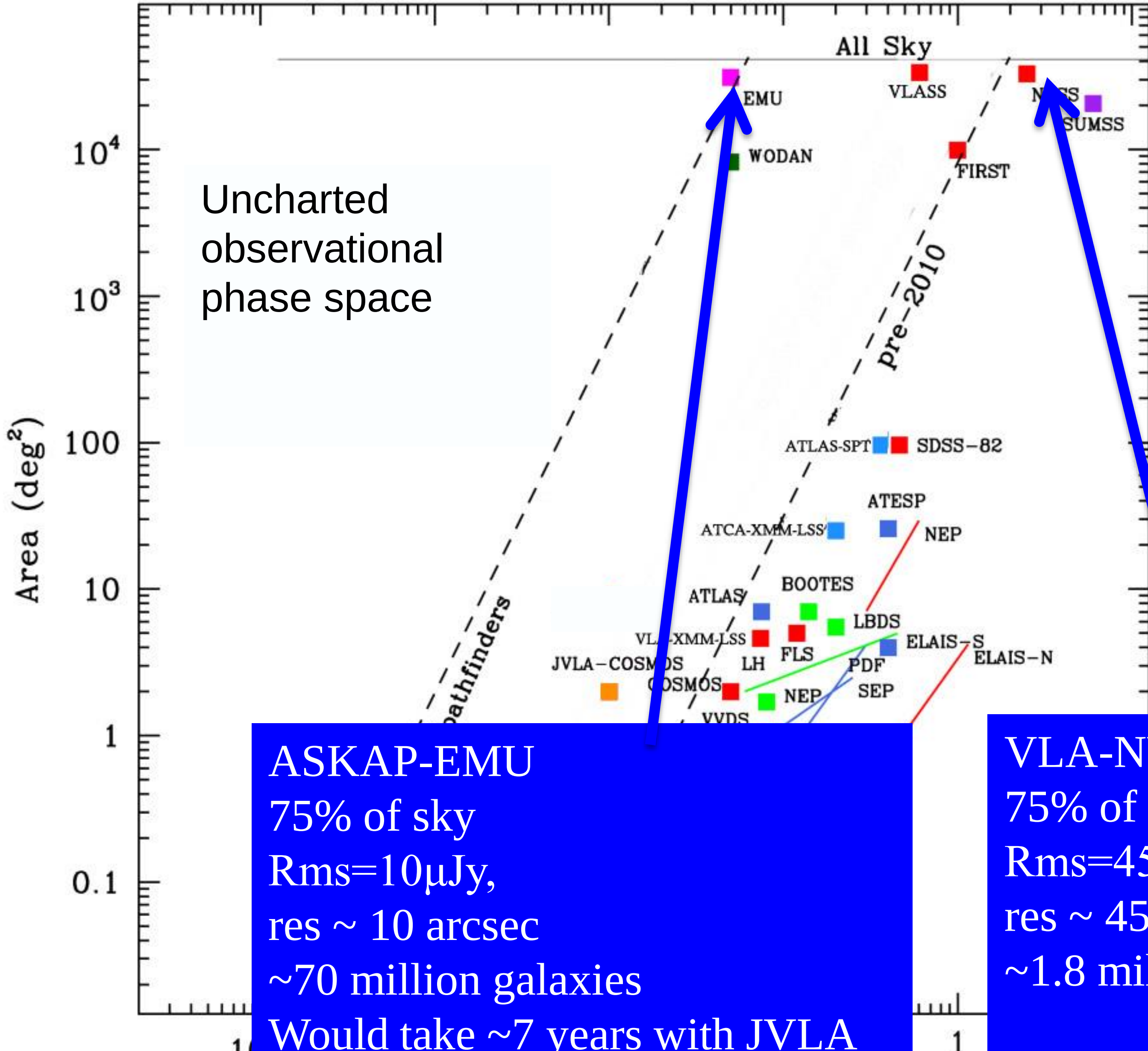
The EMU project consists of about
300 scientists in 21 countries
http://askap.pbworks.com/TeamMembers



ASKAP-EMU
75% of sky
Rms=10µJy,
res ~ 10 arcsec
~70 million galaxies
Would take ~7 years with JVLA

VLA-NVSS
75% of sky
Rms=450µJy,
res ~ 45 arcsec
~1.8 million galaxies

5σ Sensitivity (mJy)

CSIRO

EMU Overview

Deep radio image of 75% of the sky (to declination +30°)

Frequency range: 1100-1400 MHz

40 x deeper than NVSS (the largest existing radio survey)

- 10 μJy rms across the sky

5 x better resolution than NVSS (10 arcsec)

Better sensitivity to extended structures than NVSS

Will detect and image ~70 million galaxies at 20cm

c.f. 2.5 million detected over the entire history of radio-astronomy so far

All data to be processed in pipeline

Images, catalogues, cross-IDs, to be placed in public domain

Survey starts 2017(?)

How does EMU differ from earlier surveys?

1. Scale – increases the number of known radio sources by a factor of ~30

2. Will not be dominated by AGN – about half the galaxies will be normal SF galaxies

3. Ambition – includes:

- Cross-identification with optical/IR catalogues
- Ancillary data (redshifts etc)
- Key science projects as an integral part of the project

4. Uses “Large-n astronomy” techniques

5. Explicitly includes “discovering the unexpected”

EMU and its pathfinders

Comparison: NVSS
3π sr, rms=450 μJy
1.8 million galaxies

ATCA – ATLAS
(2006-2013)
6 antennas single-pixel

7 sq deg
Rms=15 μJy
6000 galaxies

ATCA - SCORPIO
(2014-2016)
6 antennas single-pixel
Galactic (b=0)

4 sq deg
Rms=30 μJy
Hundreds of
Galactic objects

ATCA – ATLAS - SPT
(2013-2016)
6 antennas single-pixel

100 sq deg
Rms=40 μJy
30,000 galaxies
300 clusters?

ASKAP – early science
(2016)
12 antennas MkII PAF

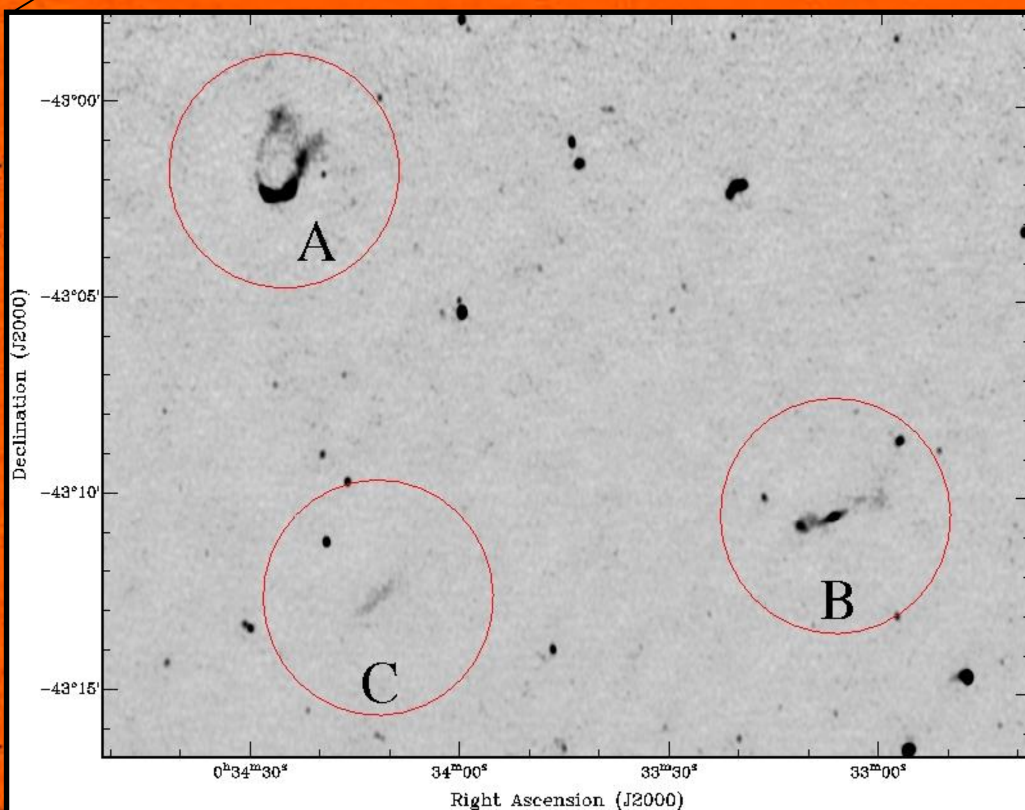
1000 sq deg
Rms=30 μJy
0.5 million galaxies

ASKAP – EMU
(2017-2018)
30-36 antennas MkII PAF

3π sr
Rms=10 μJy
70 million galaxies

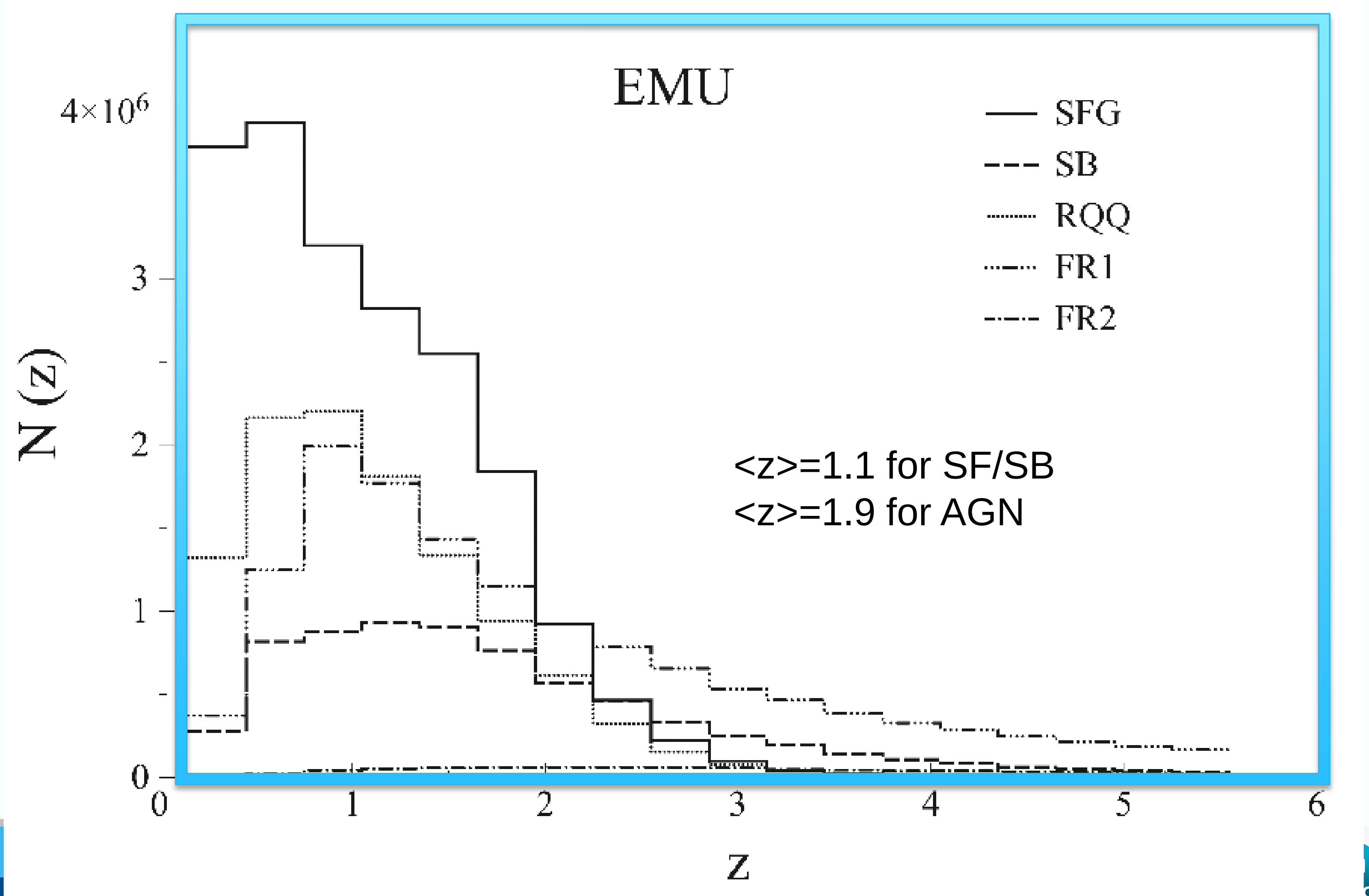
CSIRO

The EMU Pathfinder:
ATLAS=Australia Telescope Large Area Survey
7 sq deg to rms=15 μJy



Mao et al. 2010MNRAS.406.2578M

Redshift distribution of EMU sources



Based on SKADS (Wilman et al; 2006, 2008)

EMU Project Structure

The EMU Management Team

- **Ray Norris** (Project Leader)
- **Andrew Hopkins & Nick Seymour** (Project Scientists)
- **Anna Kapinska** (Project Manager)
- **Josh Marvil** (ASKAP Early Science Team Leader)
- **Ian Heywood** (Project Guru)
- **Kate Chow** (Editor, EMU newsletter)

The EMU's EGG (Emu Genius Group)

- The consultative group of ~30 active, involved, EMU members

The EMU team

- ~300 team members from 21 countries

Early Science Projects

Development Projects

Collaboration Projects

Key Science Projects

ASKAP Early Science

ASKAP Continuum Early Science

~800 hours of ASKAP-12 time, on a shared risk basis

Consensus that priority should be given to 3-band observations

- Adds value to – not superseded by - EMU

Expressions of interest were invited, and grouped into fields

Priority provisionally assigned taking into account:

- the number and quality of science projects using the field
  - Non-commensal projects given max allocation of one field
- the extent to which the observation will help debug and commission ASKAP
- the extent to which the projects use ASKAPs unique capability, and in particular the wide field of view
- the extent to which the projects helps drive an EMU key science project

Results in:

- Cosmology Field: 2000 sq deg, single band, 200 hours
- All other fields use 3 bands

Requested fields include:

SPT

CDFS

GAMA

ELAIS-S1

COSMOS

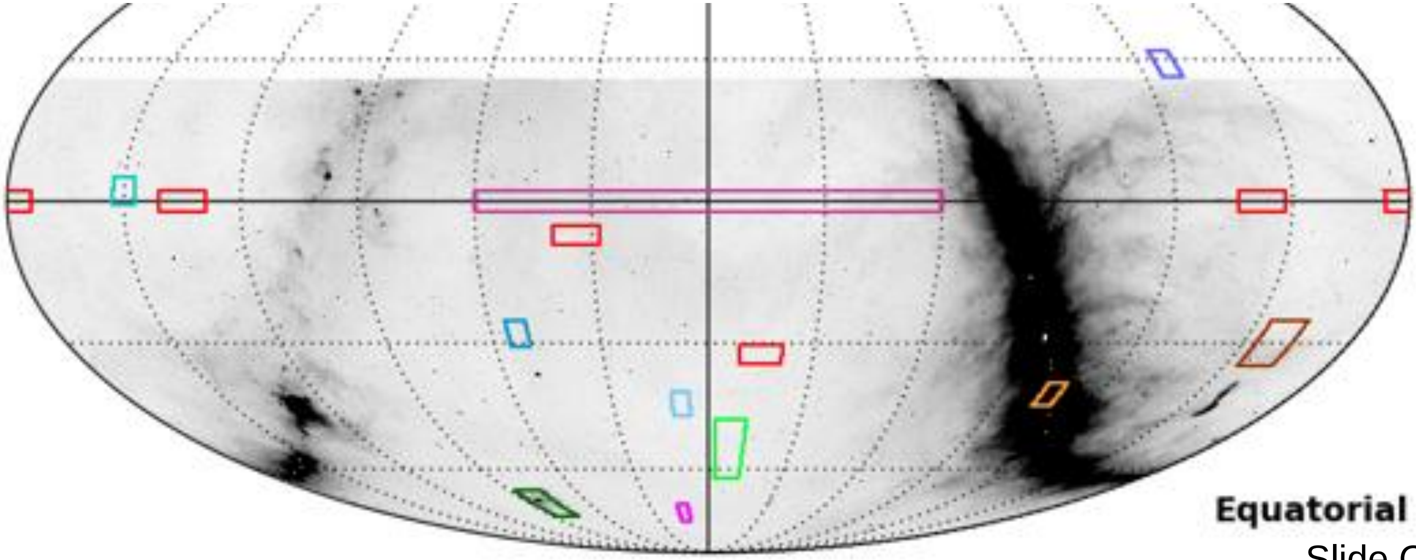
Stripe 82

Scorpio

Shapley

SPARCS-N

SMC & LMC



Slide Courtesy of Josh Marvil

Current ranked list of fields

Name	RA	Dec	# of pointings	# of hours	Area in Deg^2	Bands	Projects	Priority	Notes
Cosmology	02:30:00 (0<RA<5)	-52:30:00 (-65<dec<-25)	60	200	2000	800-1100	3,13,14,17,22	A	1
GAMA(1)	23:00:00	-32:30:00	2-3	60-90	60-90	all	4,5,6,7,17,27	A	2
GAMA(2)	12:00:00 09:00:00 14:32:00 02:16:00	00:00:00 00:00:00 00:00:00 -07:00:00	8-12	240-360	240-360	all	4,5,6,7,17,27	B	2
SPT	00:00:00 (23<RA<0)	-55:00:00 (-60<dec<-50)	4	120	120	all	13,14,22	A	1
Shapley(1)	13:27:57	-31.30.09	1	30	30	all	20	A	3
Shapley(2)	12:54:18 13:47:28	-29.01.16 -32.50.59	2	60	60	all	20	B	3
Low-z clusters(1)	22:49:57	-64.25.46	1	30	30	all	19	A	
Low-z clusters(2)	21:52:19	-19.34.42	1	30	30	all	19	B	
CDFS	03:31:00	-28.00.00	1	30	30	all	6,11,12,17	A	
ELAIS	00:34:00	-43.40.00	1	30	30	all	6,12,17	A	
COSMOS	10:00:00	02:10:00	1	30	30	all	11	A	
SPARCS-N	15:30:00	29:00:00	1	30	30	all	11	A	
SCORPIO	17:00:00	-41.30.00	1	30	30	all	10	A	
Galactic-2	15:54:21	-53.43.19	1	30	30	all	9	A	
SMC	00:53:00	-72:50:00	1	30	30	all	21	A	
LMC	05:23:00	-69:45:00	4	120	120	all	21	B	
early polarisation	TBD	TBD	1	30	30	all	POSSUM	A	

EMU Development Projects

Examples of EMU Development Projects

Developers earn co-authorship on key science papers

- Ensure the EMU database satisfies our storage and access needs (both CASDA and value–added, and interact

See Bernd Vollmer talk on Friday

- Develop, set up, and implement the data quality/validation process
- Ensure ASKAPSOFT imaging satisfies EMU needs
- See what special imaging is needed for the Galactic Plane
- Ensure ASKAPSOFT source extraction satisfies EMU needs

For Level 6 data

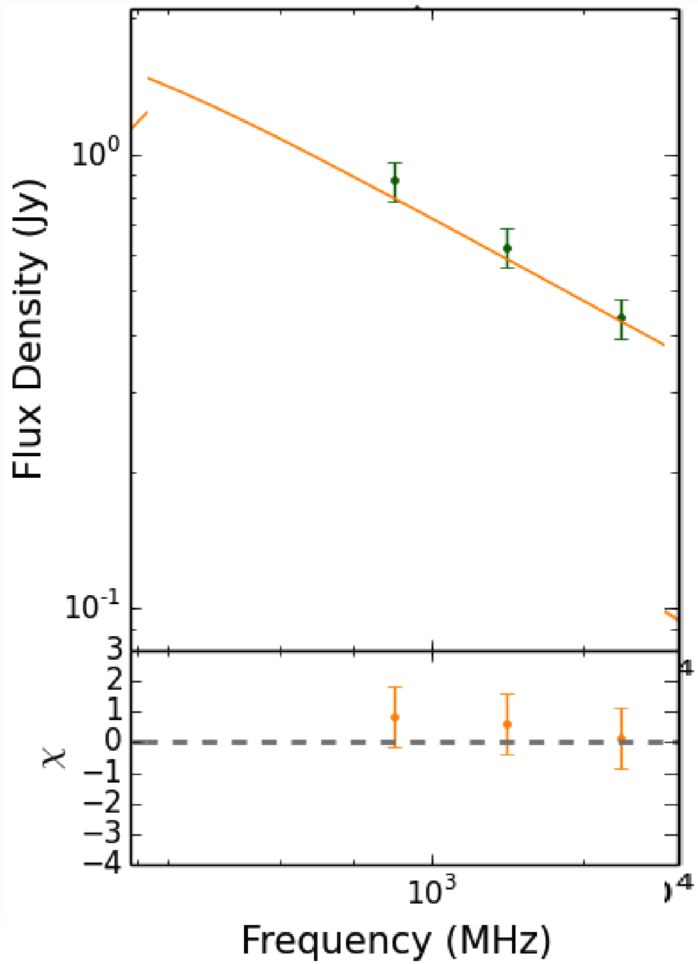
- Develop algorithms for extract

See Chiara Ferrari talk on Friday

For Level 7 data

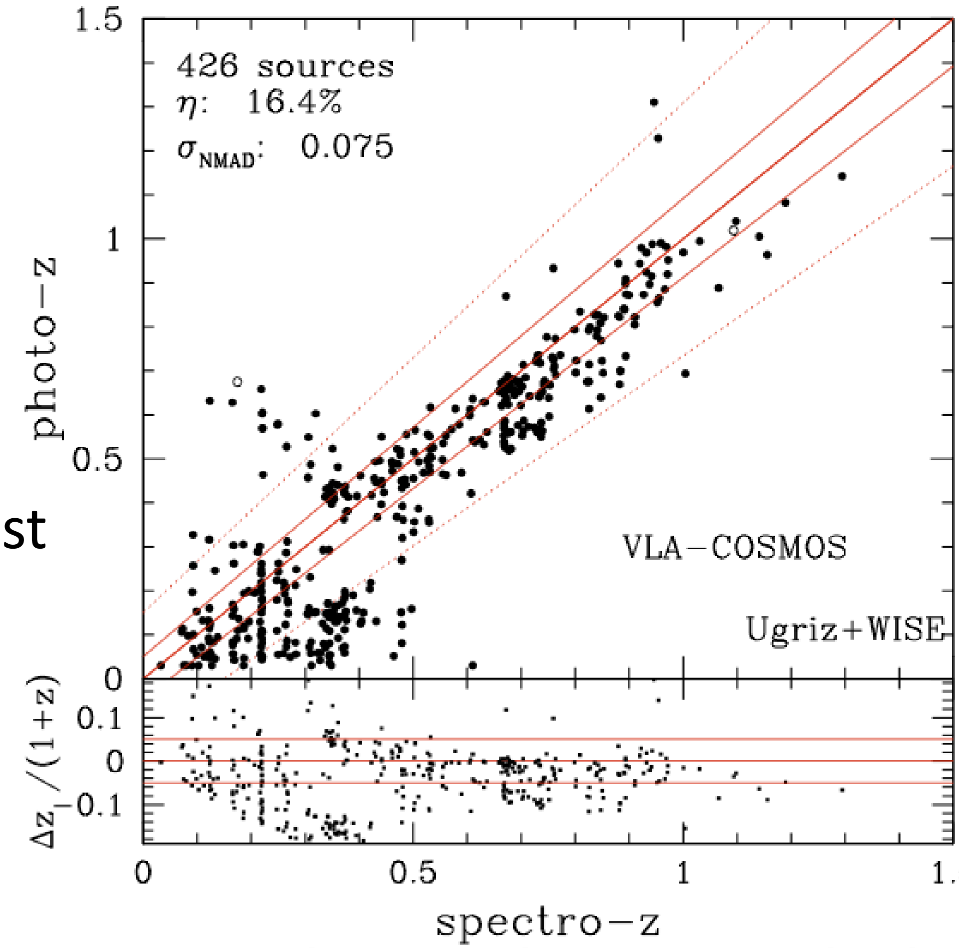
- Develop the self-ID and cross-ID algorithms
- Develop an "optimum photo-z algorithm" for all EMU and an optimum photo-z strategy for those smaller areas of EMU covered by other surveys such as DES
- Develop techniques for Statistical redshifts & Spatial Cross-correlation redshifts
- Explore other EMU applications for Machine Learning

"There's nothing as useless as a radio source"
Really?



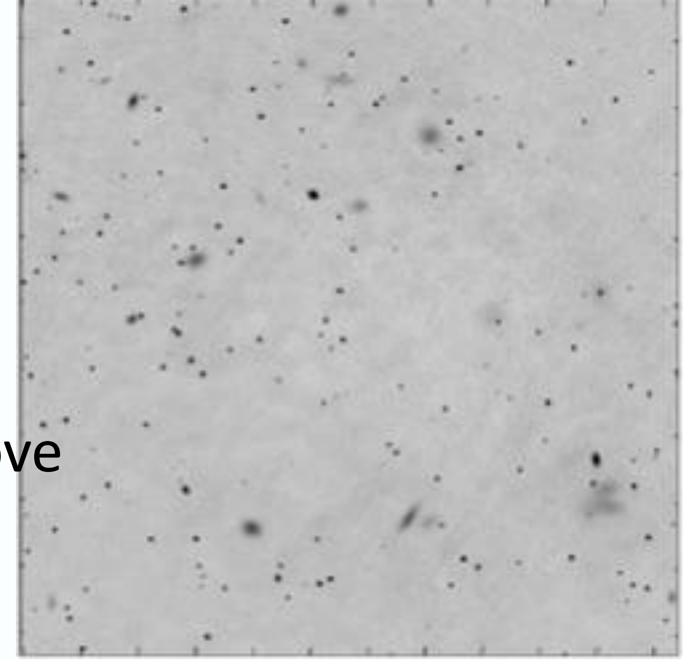
EMU (Statistical) Redshifts

- We will measure spectroscopic redshifts for only ~2% of EMU sources (WALLABY, TAIPAN, etc)
- Even photometric redshifts are hard to do well (SkyMapper)
  - But many of our science goals don't need accurate z's – they just need a redshift bin
- Several machine-learning algorithms are being tried (e.g. kNN, right)



From Salvato, Zinn, ++ in preparation)

Compact Source Extraction



Initial data challenge study showed existing source extraction algorithms not up to the job.

Need to build on this, identify problems, improve existing algorithms or develop a better one.

Work in progress by Matt, Josh, Andrew.

Publications of the Astronomical Society of Australia (PASA)
© Astronomical Society of Australia 2015; published by Cambridge University Press.
doi: 10.1017/pas.2015.xxx.

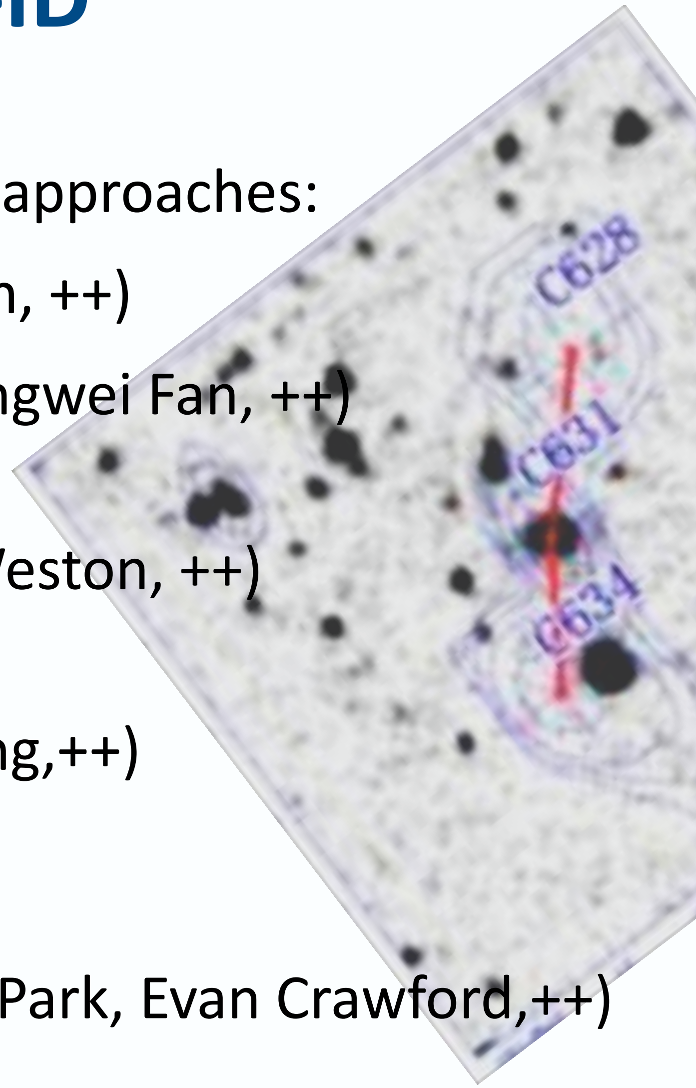
The ASKAP/EMU Source Finding Data Challenge

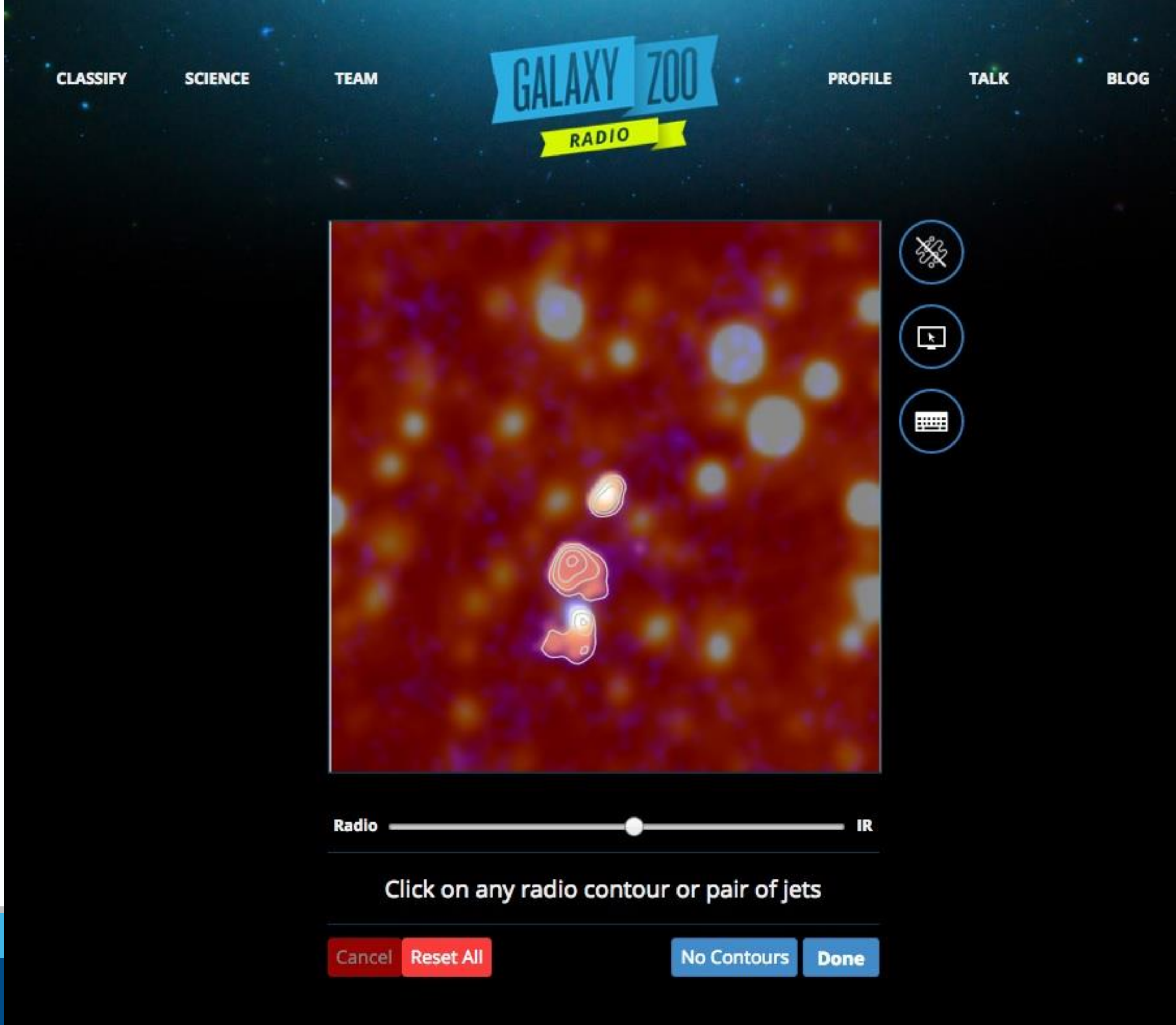
A. M. Hopkins[1,*], M. T. Whiting[2], N. Seymour[3], K. E. Chow[2], R. P. Norris[2], L. Bonavera[4], R. Breton[5], D. Carbone[6], C. Ferrari[7], T. M. O. Franzen[3], H. Garsden[8], J. Gonzalez-Nuevo[4], C. A. Hales[9], P. J. Hancock[3,10,11], G. Heald[12,13], D. Herranz[4], M. Huynh[14], R. J. Jurek[2], M. Lopez-Caniego[15,4], M. Massardi[16], N. Mohan[17], S. Molinari[18], E. Orrù[12], R. Paladino[19,16], M. Pestalozzi[18], R. Pizzo[12], D. Rafferty[20], H. J. A. Röttgering[20], L. Rudnick[21], E. Schisano[18], A. Shulevski[12,13], J. Swinbank[22,6], R. Taylor[23,24], A. J. van der Horst[25,6]

EMU classification and cross-ID development projects

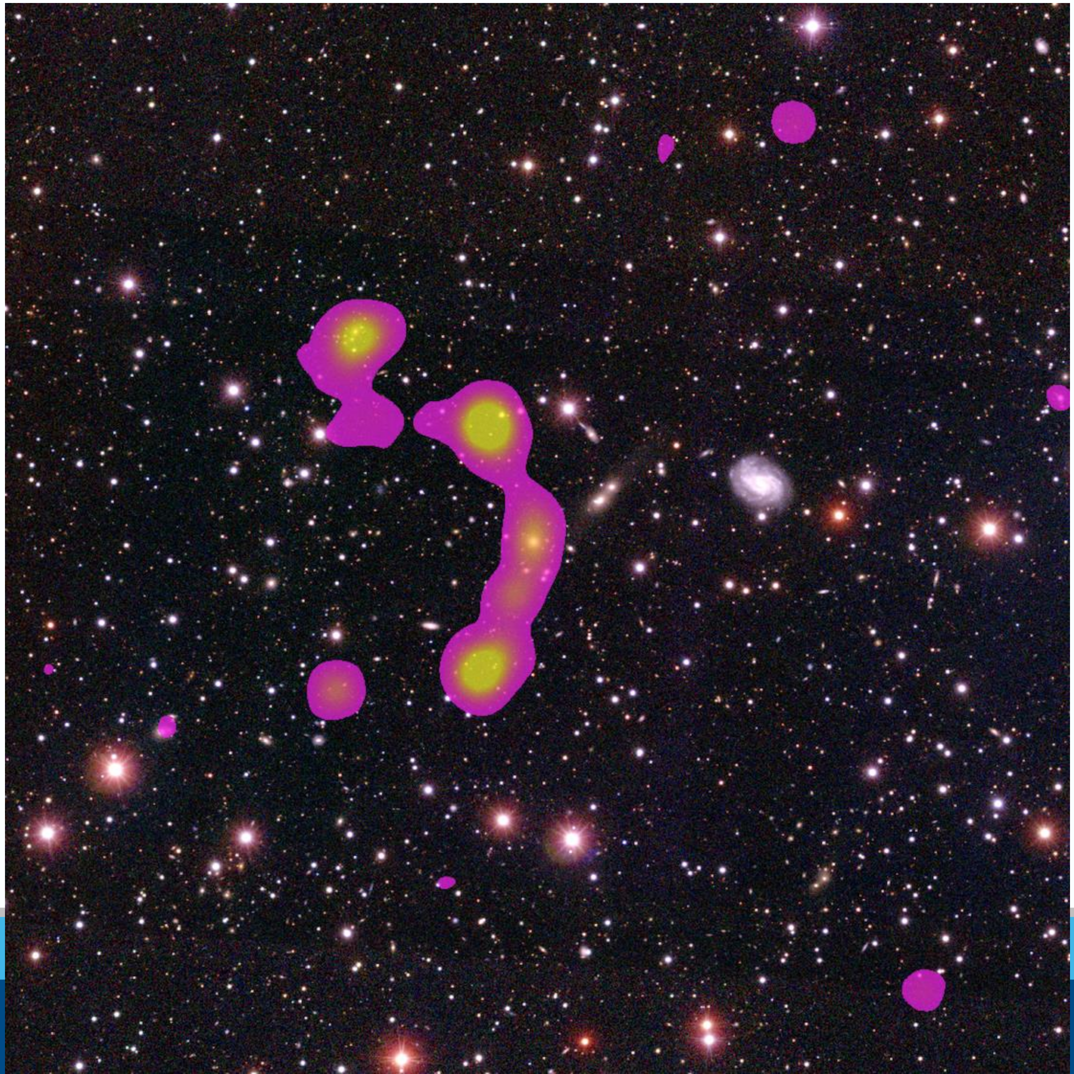
Currently using ATLAS as a training set for 5 approaches:

1. “Expert by eye” classification (Jesse Swan, ++)
2. Bayesian approach (Tamas Budavari, Dongwei Fan, ++)
   - See Fan+2015 2015MNRAS.451.1299F
3. Likelihood Ratio (Nick Seymour, Stuart Weston, ++)
   - See Weston+2016 in prep.
4. Radio Galaxy Zoo (Julie Banfield, Ivy Wong,++)
   - See Banfield+2015MNRAS.453.2326B
   - Over 1 million cross-IDs
5. Machine Learning (Ray Norris, Laurence Park, Evan Crawford,++)
   - Part of the WSU WTF Machine Learning hub




CLASSIFY
SCIENCE
TEAM
GALAXY ZOO
RADIO
PROFILE
TALK
BLOG
Radio
IR
Click on any radio contour or pair of jets
Cancel
Reset All
No Contours
Done

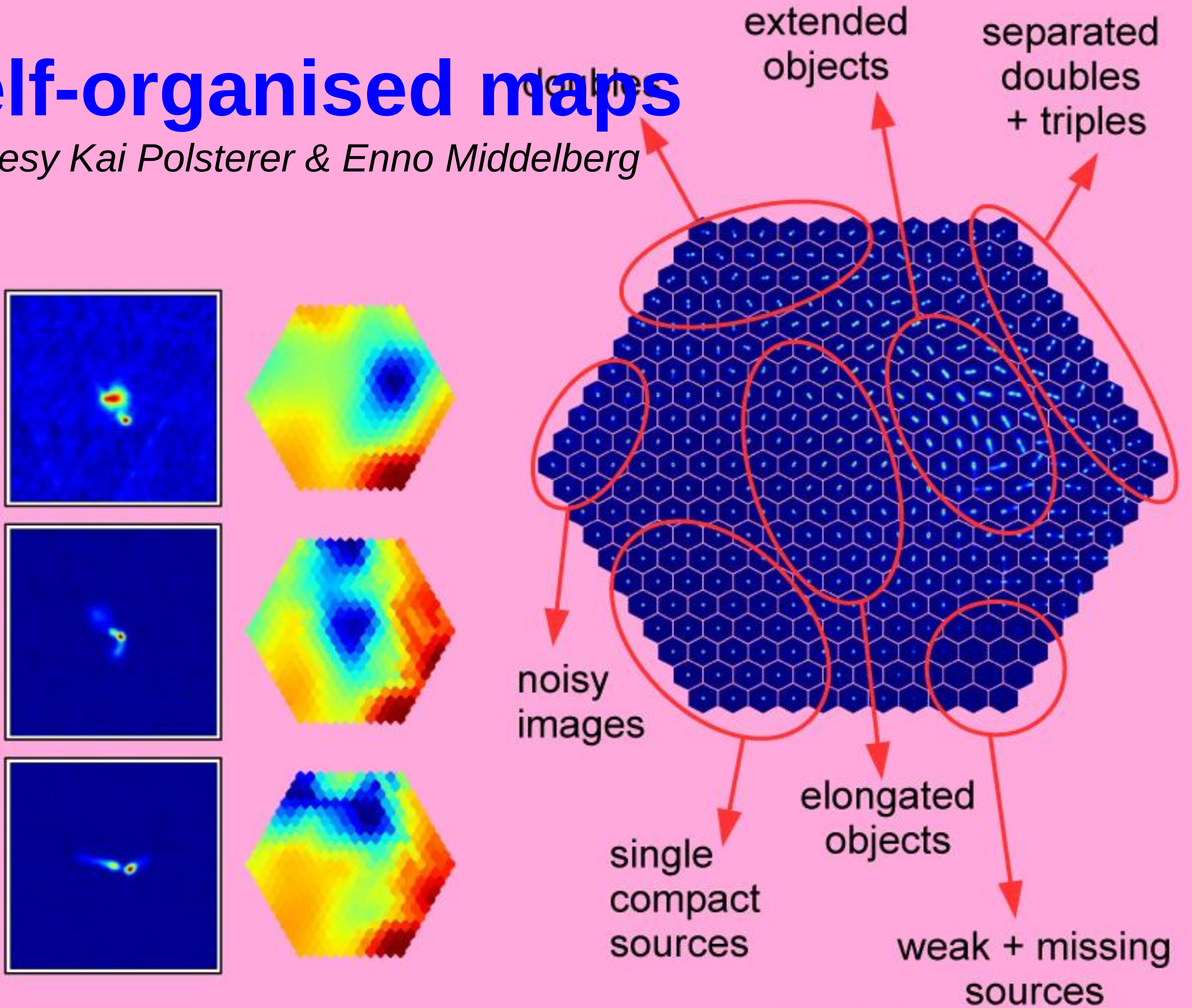
New result from Radio Galaxy Zoo: citizen scientists discover Giant WAT + cluster



Banfield+2016, MNRAS, in press

Self-organised maps

courtesy Kai Polsterer & Enno Middelberg



EMU Collaboration Projects

Examples of EMU Collaboration Projects

Meerkat-MIGHTEE	Ray Norris
eROSITA	Nicolas Clerc et al.
SkyMapper	Julie Banfield
WISE	Tom Jarrett
VHS	(TBD)
LSST	Amy Kimball
MWA	(TBD)
Taipan	Ray Norris
DES/OZ-DES	Nick Seymour
POSSUM	(TBD)
XXL	Vernesa Smolcic
Radio Galaxy Zoo	Julie Banfield
FP7-HELP	(TBD)
CTA	(TBD)

EMU Key Science Projects

EMU Key Science Projects

EMU Key Science Projects	Project Leaders
EMU Value-Added Catalogue	Nick Seymour
Characterising the Radio Sky	Ian Heyw
EMU Cosmology	Davi
Cosmic We	
Clusters of Galaxies	ie Johnston-Hollitt Chiara Ferrari
Cosmic star formation history	Andrew Hopkins
Radio-loud AGN	Anna Kapinska
Radio AGN in the EoR	Jose Afonso
Radio-quiet AGN	Isabella Prandoni
Local Univers	Josh Marvil, Michael Brown
The Gal ne	Roland Kothes
SCOR dio Stars	
ning Data for the Unexpected	Ray Norris

How does the Universe olve?

See Shea Brown/Tessa Vernstrom talks on Thursday

See Francesco Cavallero talk on Friday

EMU Cosmology
(leader: David Parkinson)

The four probes of EMU Cosmology ("large-n" cosmology)

1. "Auto-correlations of radio data

- -> spatial power spectrum

2. Cross-correlation between (z<0.5) optical foreground galaxies and (<z>~1.5) EMU galaxies

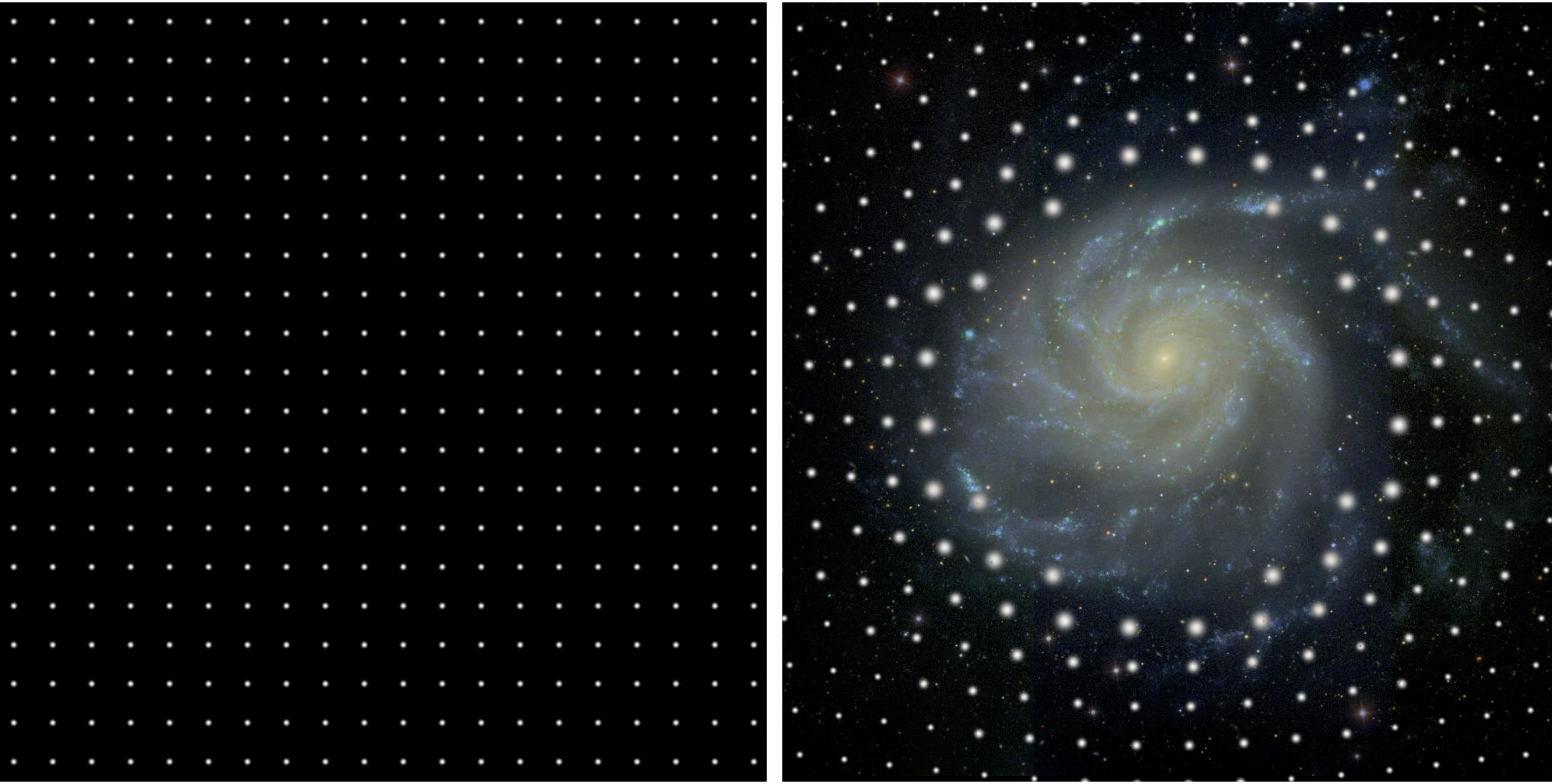
- -> cosmic magnification at low z
- Only needs 2 redshift bins

3. Cross-correlation between EMU galaxies and CMB ($\theta < 1^\circ$)

- -> cosmic magnification at high z
- Doesn't need redshifts
- Good match to EMU

4. Cross-correlation between EMU density and CMB ($\theta \sim 10^\circ$)
   - Using Integrated Sachs-Wolfe effect
   - Standard ΛCDM predicts no high-z ISW
   - Massive neutrinos do predict high-z ISW

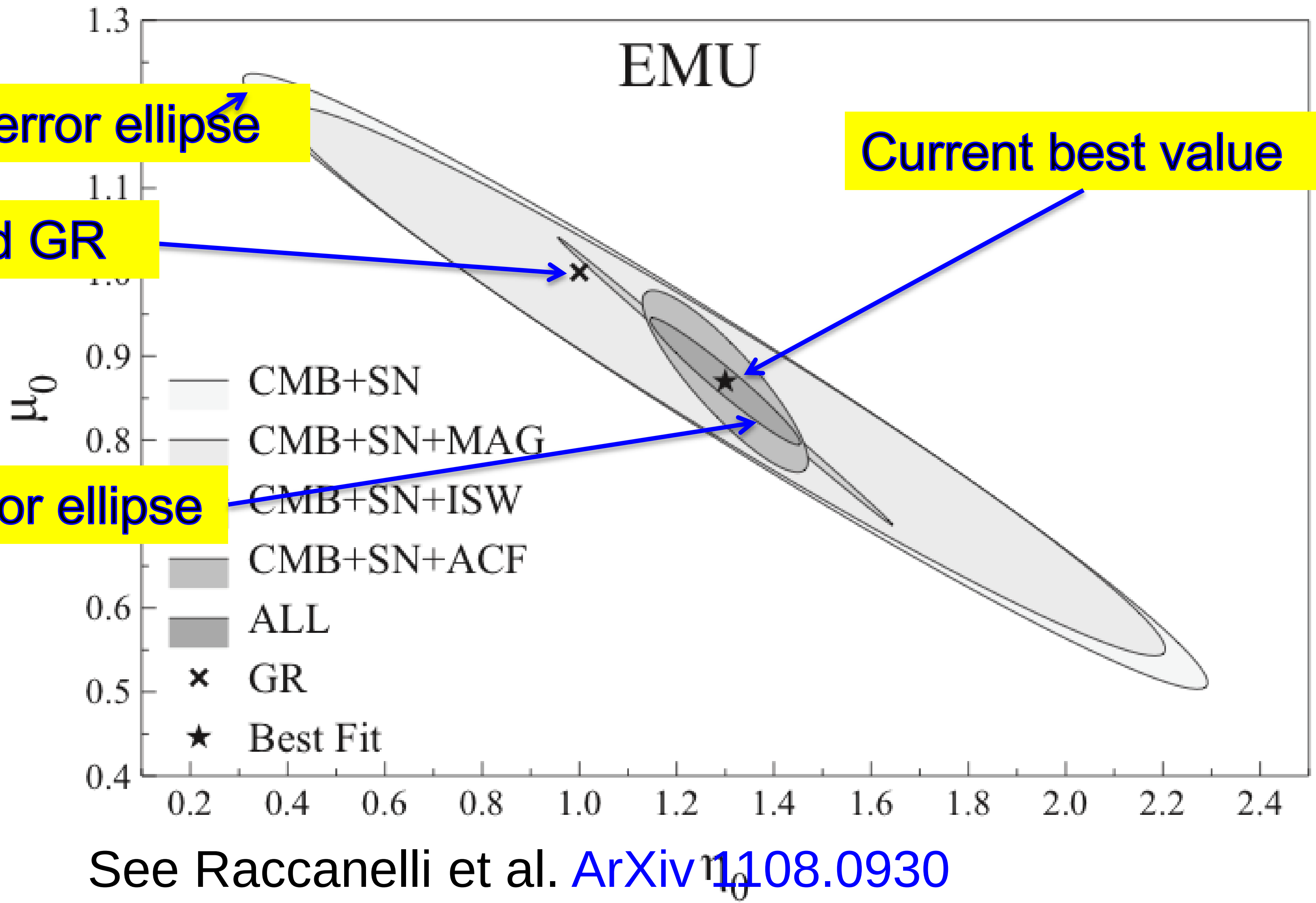
Cosmic Magnification



Don’t need to know individual z’s, just the z-distribution

BUT do need to ensure there is no overlap between samples

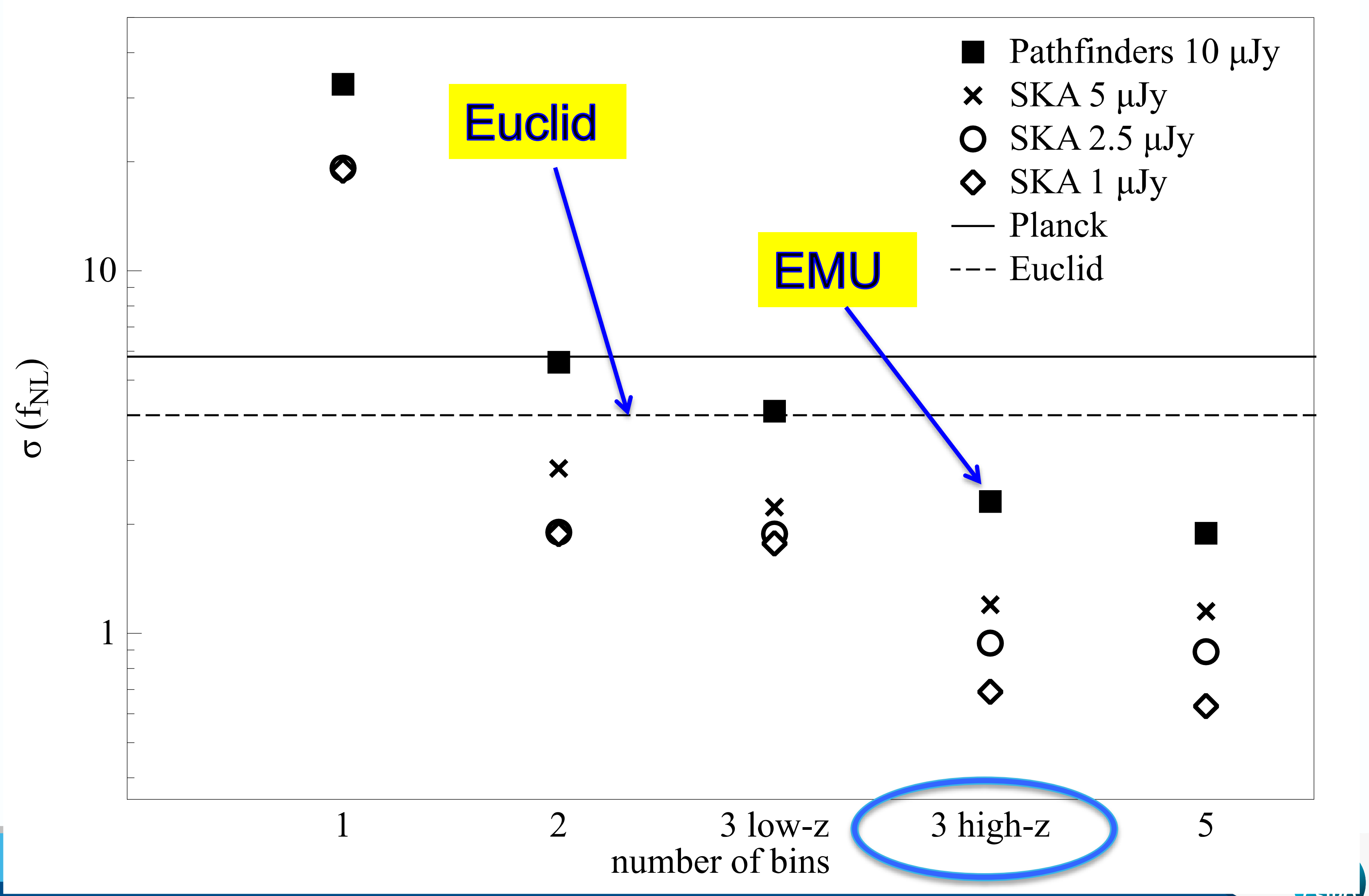
Modified Gravity from ISW (assuming no redshifts)



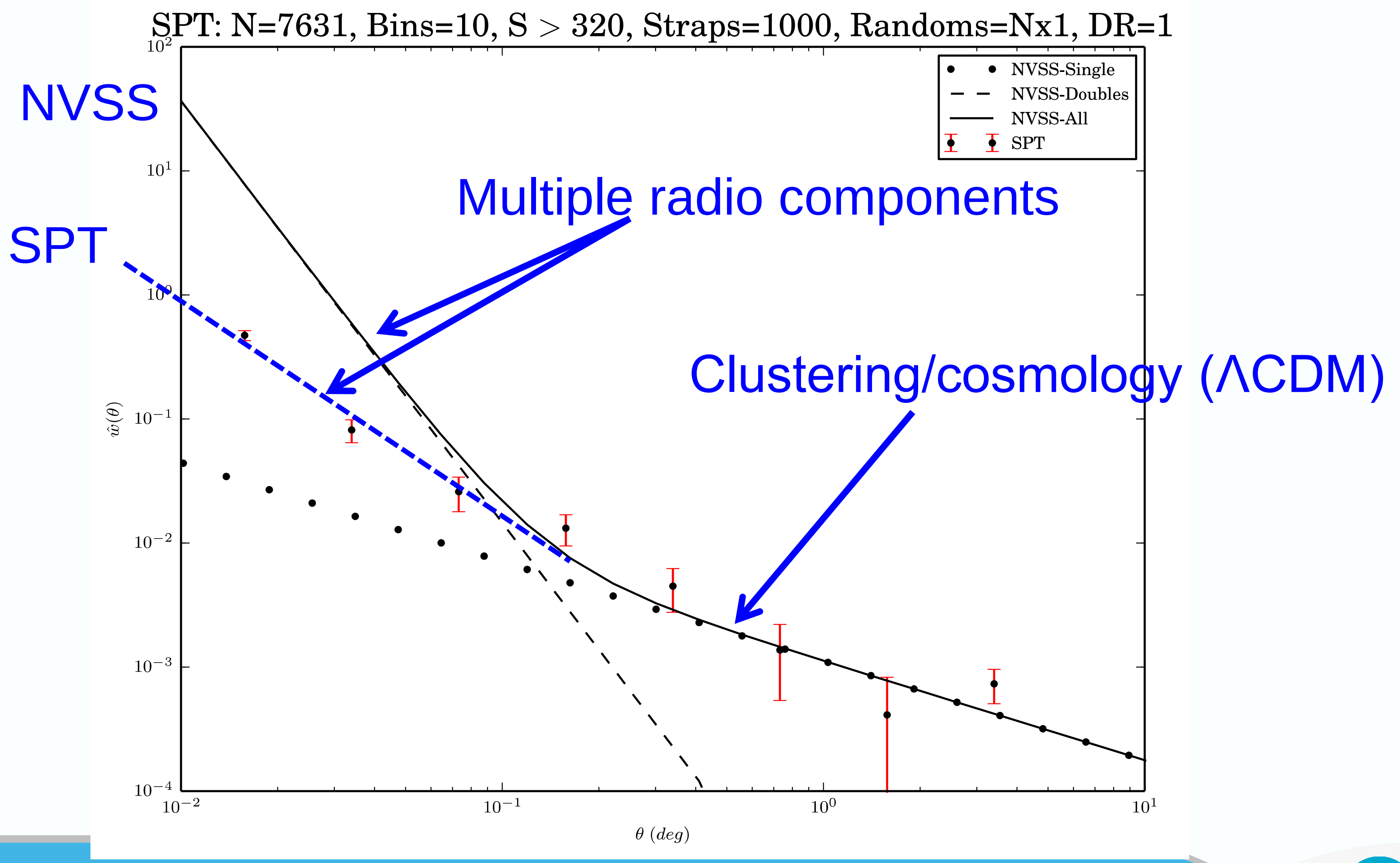
See Raccanelli et al. ArXiv 1108.0930

Non-gaussianity

Raccanelli et al., 2014, arXiv1406.0010



Progress report from PhD student Glen Rees
E.g Spatial autocorrelation function on ATLAS-SPT



Clusters KSP

(leaders: Chiara Ferrari & Melanie Johnston-Hollitt)

Abstract: radio observations are really good at tracing the properties of clusters. But our current small, biased, samples may be misleading us. EMU will detect > 100,000 clusters.

Kaustuv Basu (Uni. Bonn) suggests that up to 80% of clusters may contain halos

Radio halos in galaxy clusters are Mpc scale diffuse synchrotron sources. Currently we don't know exactly what fraction of clusters host radio halos, and what is their distribution with cluster masses and redshift.

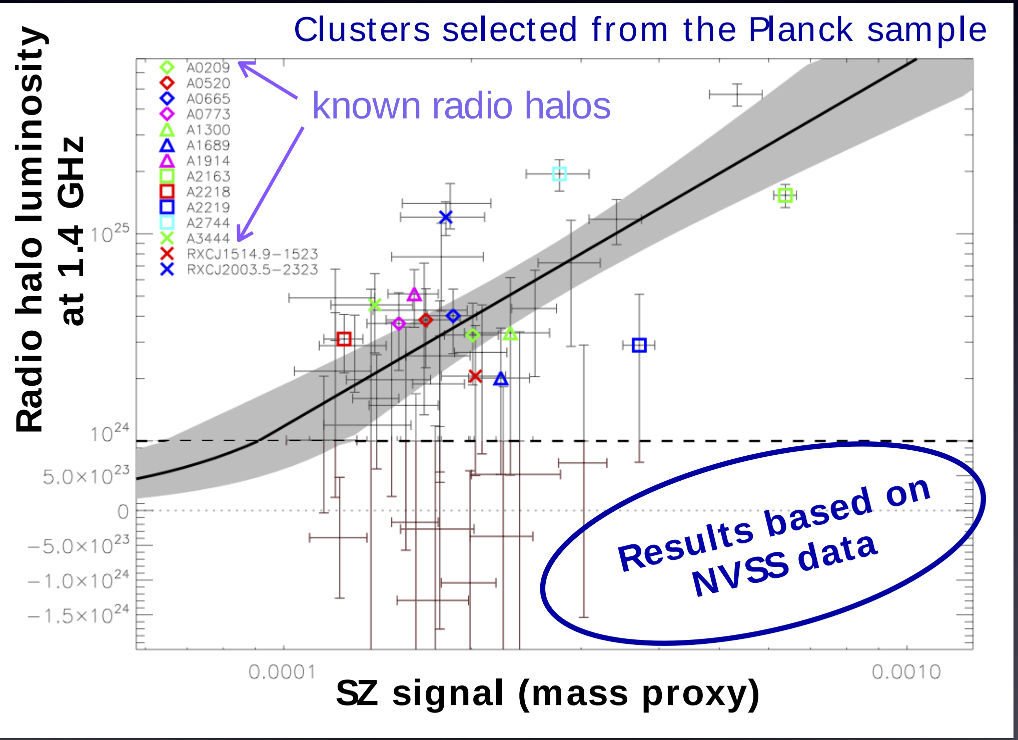
Proposed EMU Early Science Project 22

Need complete radio data for large and statistically representative galaxy cluster samples. Currently the radio analysis part had been possible only with the NVSS data (**Sommer & Basu 2014**).

NVSS based results are *noisy*, but hints at a promising and interesting new discovery:

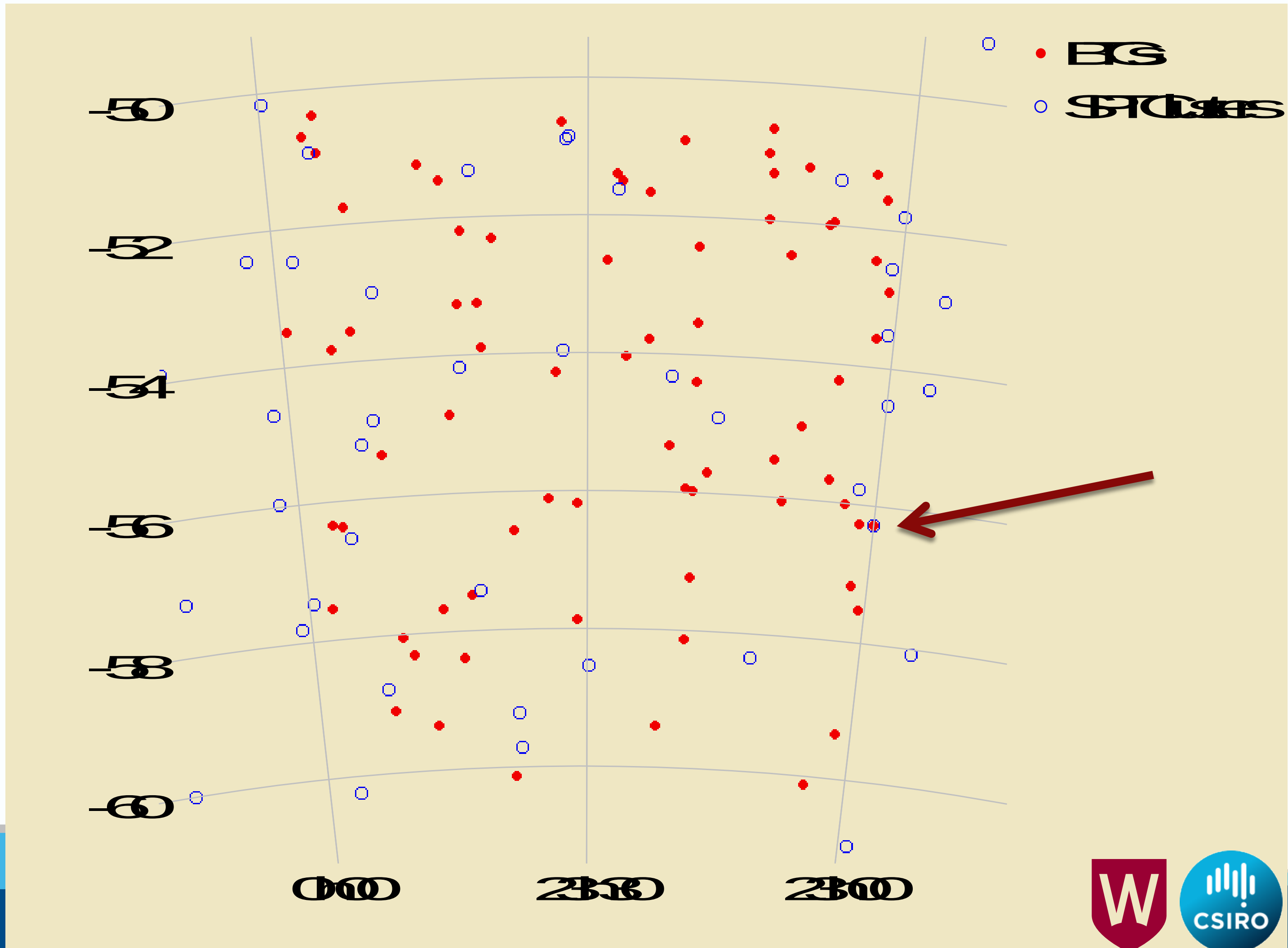
up to ~80% of massive clusters might be hosting radio halos!

Data from the ASKAP/EMU survey, with 40 times more sensitivity than NVSS and much better spatial dynamic range, **will settle this question comprehensively**. We will get radio halo statistics of large cluster samples, like that from the SPT survey, out to z~1.

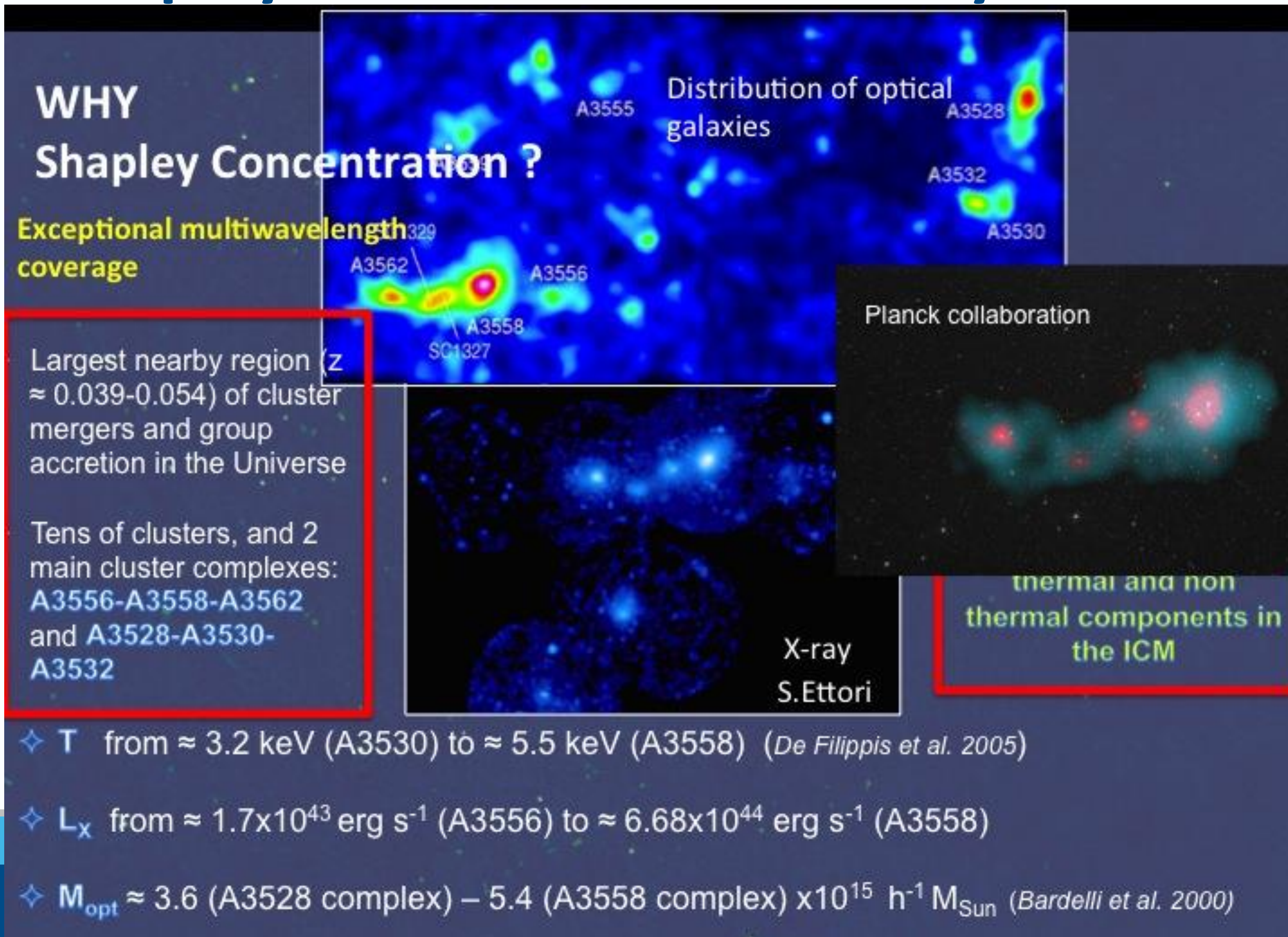


Ref: Basu, K. 2012, MNRAS, 421; Sommer, M. W. & Basu, K. 2014, MNRAS, 437

Andrew O'Brien (Western Sydney Uni) has shown that head-tail galaxies trace a different population of clusters from S-Z in the SPT field.



Tiziana Venturi++ plans to study nearby Shapley clusters in ASKAP Early Science



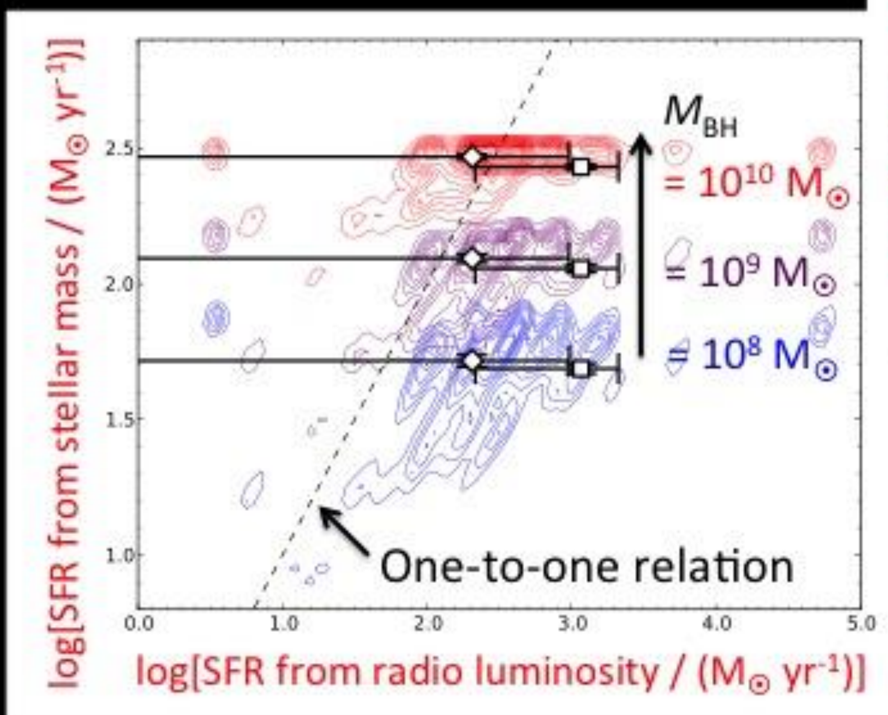
EMU Radio-quiet AGN KSP

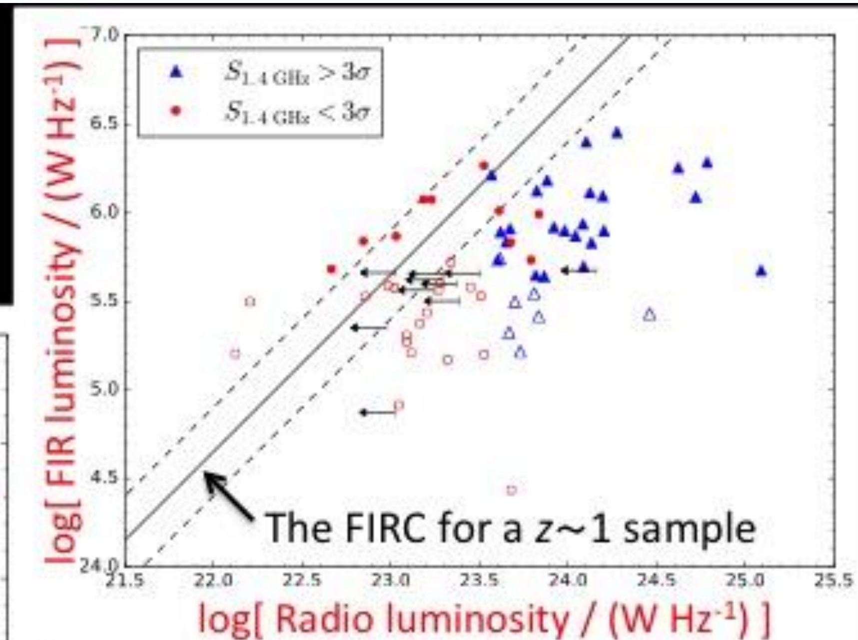
(Leader: Isabella Prandoni
- also see her talk on Monday)

Evidence that accretion dominates the radio emission in RQQs

Sarah White (ICRAR, Curtin University), Matt Jarvis (Oxford), Boris Häußler (ESO), Natasha Maddox (ASTRON),Eleni Kalfountzou (ESA),Martin Hardcastle (Herts)

A statistical comparison of star-formation rate estimates (White et al. 2015, arXiv:1410.3892)





Using the far-infrared to radio correlation (FIRC) to quantify the star-formation component (White et al., in prep.)

Sarah White, June 2016

mining radio survey data for the unexpected

WTF?

WTF = Widefield ouTlier Finder

The discovery of pulsars

Jocelyn Bell:

- explored a new area of observational phase space
- knew the instrument well enough to distinguish interference from signal
- observant enough to recognise a sidereal signature
- open minded – prepared for discovery
- within a supportive environment
- persistent

See Bell-Burnell (2009) PoS(sps5)014 for a personal perspective

Could Jocelyn Bell Discover the Unexpected in ASKAP data?

- Data volumes are huge – cannot sift by eye
- Instrument is complex – no single individual will be familiar with all possible artifacts
- ASKAP will be superb at answering well-defined questions (the "known unknowns")
- Humans won't be able to find the "unknown unknowns"
- Can we mine data for the unexpected, by rejecting the expected?

If not, ASKAP will not reach its full potential
i.e. it will not deliver value for money

Discoveries with HST

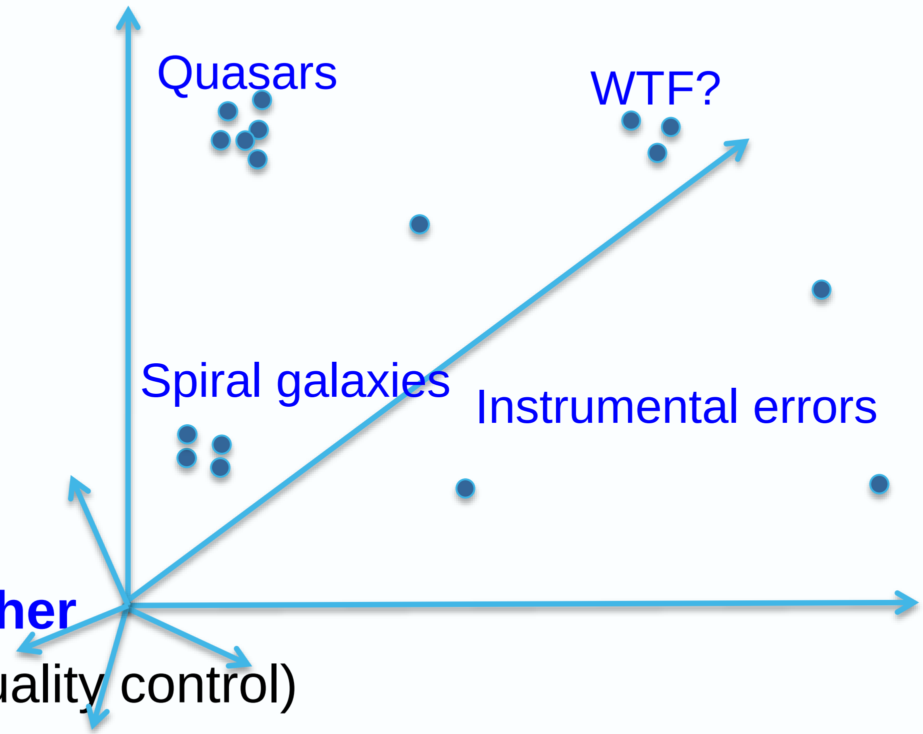
Project	Key project	Planned?	Nat. Geo. top ten?	Highly cited?	Nobel prize?
Use Cepheids to improve value of H0	✓	✓	✓	✓	
study intergalactic medium with uv spectroscopy	✓	✓			
Medium-deep survey	✓	✓			
Image quasar host galaxies		✓	✓		
Measure SMBH masses		✓	✓		
Exoplanet atmospheres		✓	✓		
Planetary Nebulae		✓	✓		
Discover Dark Energy			✓	✓	✓
Comet Shoemaker-Levy			✓		
Deep fields (HDF, HDFS, UDF, FF, etc)			✓	✓	
Proplyds in Orion			✓		
GRB Hosts			✓		

from Norris et al. 2013: arXiv1210.7521

Mining large data sets for the unexpected

WTF will work by searching the n-dimensional (large n) phase space of observables, using techniques (both supervised and unsupervised) such as

- Neural nets
- self-organised maps
- Decision tree approach
- Zoo approach
- Cluster analysis
- k-nearest-neighbours
- Bayesian approaches



Identified objects/regions will be either

- processing artifacts (important for quality control)
- statistical outliers of known classes of object (interesting!)
- New classes of object (WTF)

WTF Phase 1

- Received a grant from Amazon Web Services to develop WTF on the AWS cloud platform
- Goals:
  - Implement WTF, initially as an open challenge (c.f. Kaggle)
  - Evaluate AWS platform as a collaborative research environment
- Approach
  - Set up challenges consisting of data (images or tables) with embedded "EMU eggs"
  - Data include both simulations and real data
  - Invite ML and other algorithm groups to discover the EMU eggs
  - Develop visualisation tools to understand the process and data

First step: learn to use machine learning techniques to discover the expected before we discover the unexpected!

E.g.

- Radio-infrared cross-identification
- Morphological classification of radio sources (e.g. bent-tail)
- Source extraction (both compact and diffuse)
- Photometric redshifts
- Statistical redshifts

Putting example data and scripts on wiki for people to play

```
python src/wave_logistic.py --train data/train_list.txt --test data/test_list.txt
```

•Developing an EMU “hub” at Western Sydney University, focussing on machine learning and astroinformatics

•Developing machine learning projects as part of WTF

See our newsletter on http://tinyurl.com/emunews

What's in a name?

'WTF is that?' How we're trawling the Universe for the unknown

Posted: October 9, 2015 | **Author:** Nicholas Kachel | **Filed under:** Astronomy and Space, News | **Tags:** ASKAP | 1 Comment

The Australian Square Kilometre Array Pathfinder. Credit: Alex Cherney

Here's a challenge: how would you go about finding something if you didn't know what it was you were looking for?

No, this isn't an ancient riddle or one of those horrible corporate team building exercises. It's actually a very real problem being being faced by astronomers using our newest telescope, the Australian SKA Pathfinder (ASKAP).

What's in a name?