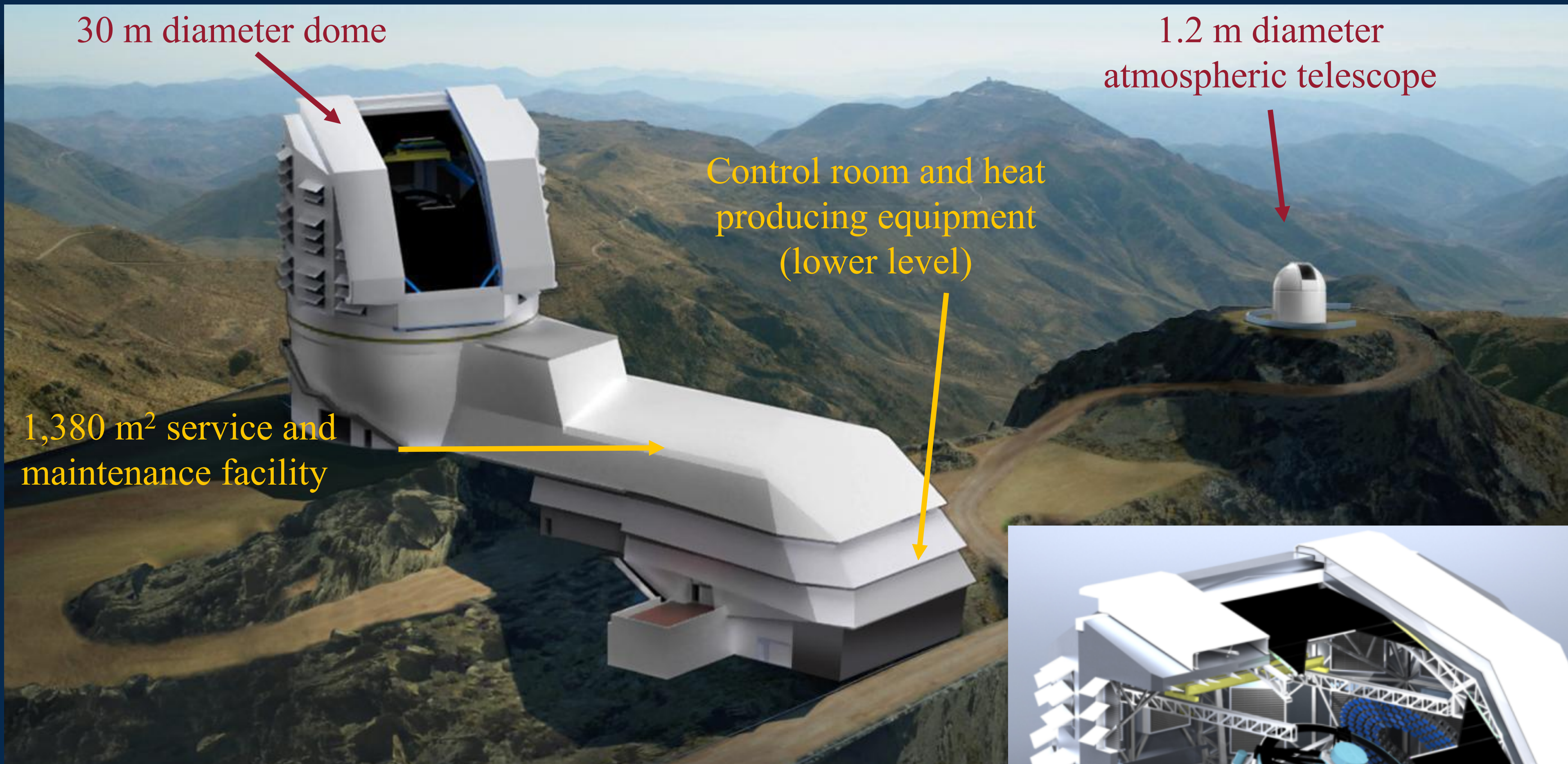
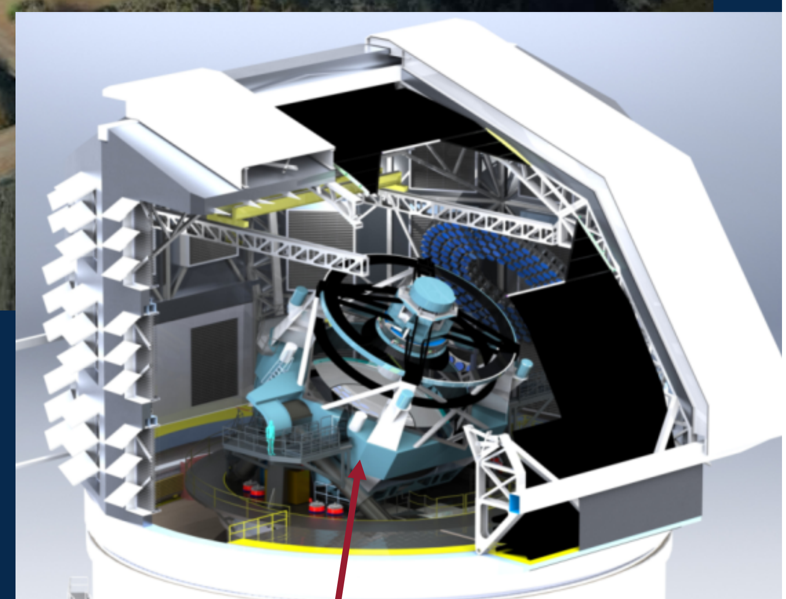
Quasar Discovery Space for LSST and Obscured Quasars at High Redshift

Gordon Richards
Drexel University

Credits:
LSST: Niel Brandt, Ohad Shemmer, Tina Peters, LSST AGN SC
Type-2: John Timlin, Joe Hennawi, Angelica Rivera

LSST: A Digital Color Movie of the Universe





350 ton telescope

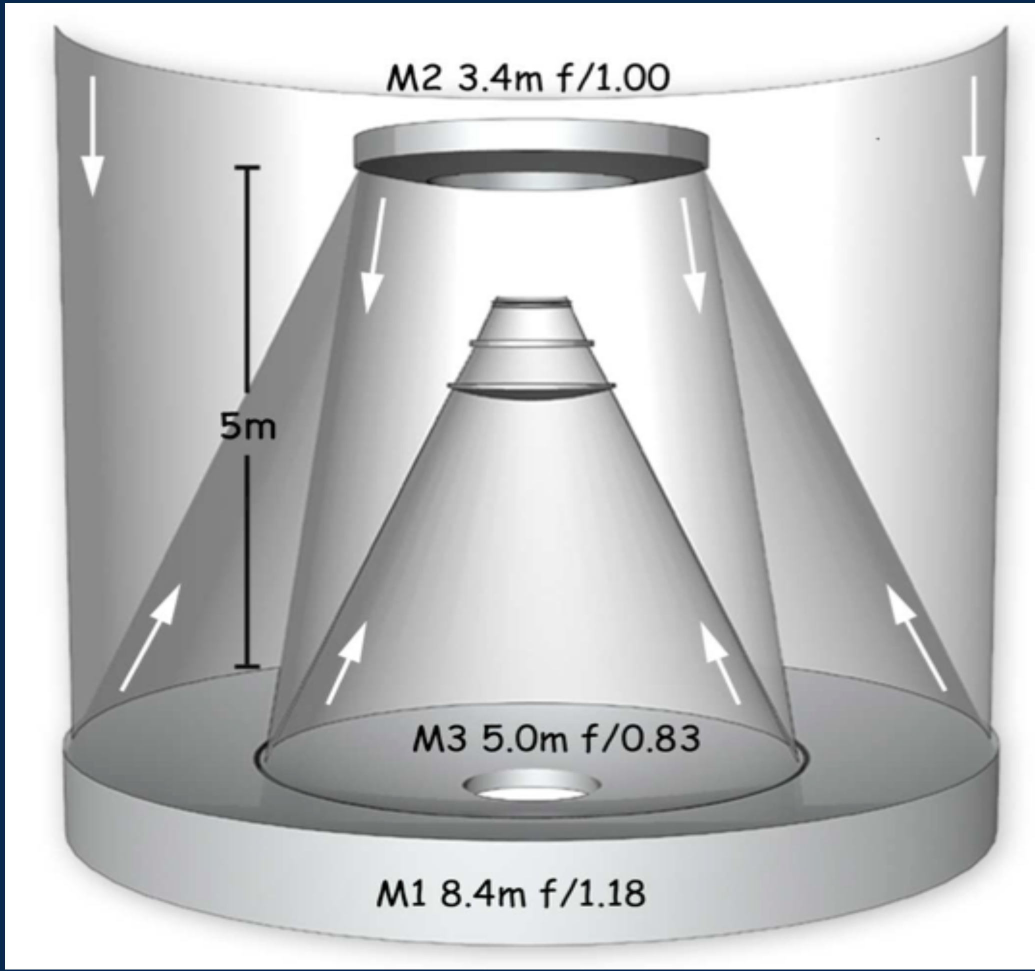
A catalog of 20 billion stars and 20 billion galaxies!

LSST Quickly Becoming Real

Time-lapse movie:
https://www.lsst.org/news/see-whats-happening-cerro-pachon

LSST: A Brief Summary

A public optical/NIR survey of ~ half the sky in the *ugrizy* bands to *r* ~ 27.5 based on ~ 820 visits over a 10-year period.



8.4 m, 6.7 m effective - 10 deg^2 - 3.2 Gpix camera

Wide

The observable southern sky. Each exposure covers 50 full Moons.

Fast

Whole observable sky scanned every 3-4 nights.

Deep

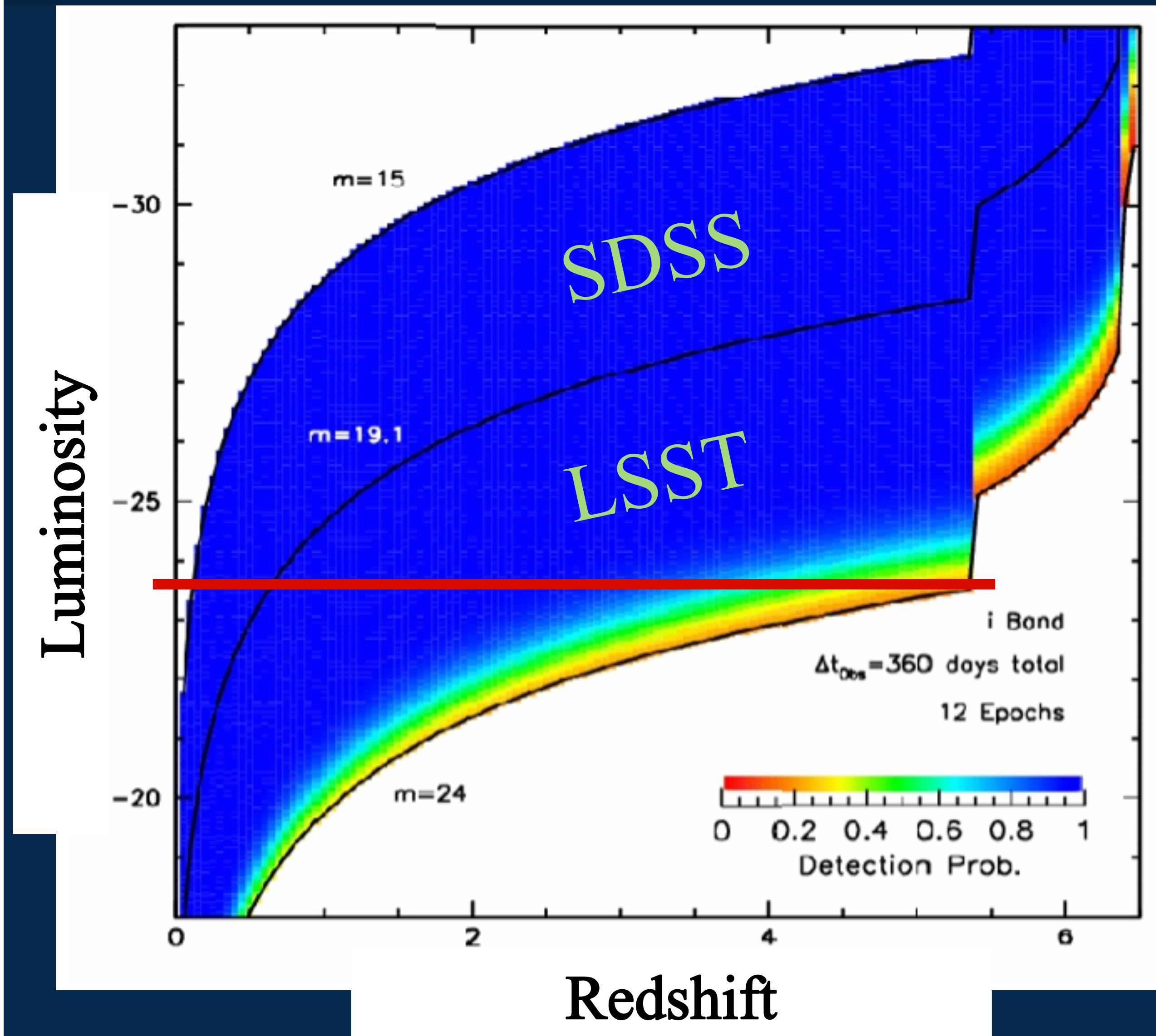
10-100 times deeper than other very wide-field surveys.

See arXiv:0805.2366 for more details.

LSST: AGN Expectations

- LSST will identify 10-30 million quasars and AGNs using data solely from the LSST project.
- Overwhelming statistics to investigate AGN evolution (e.g., as a function of environment: voids to superclusters).

LSST: (Single-epoch) Depth

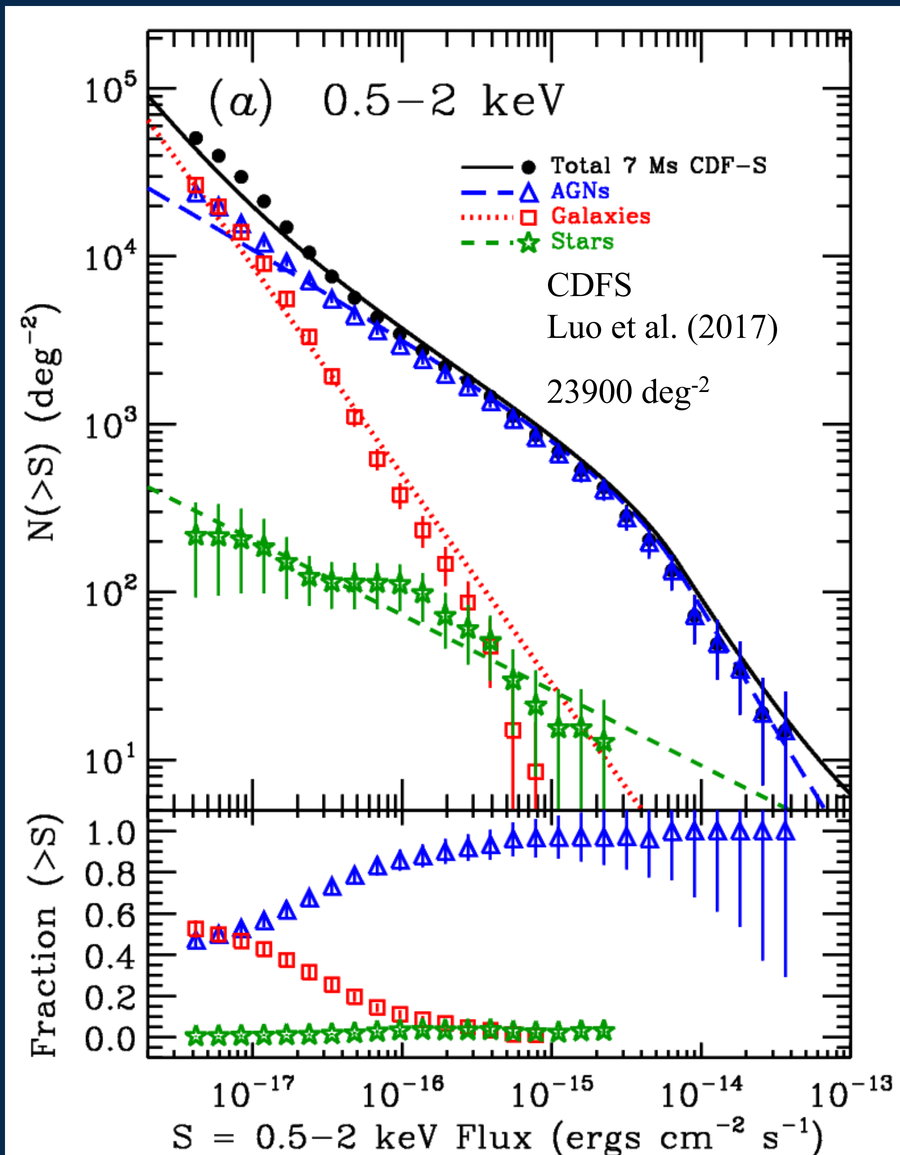


Single-epoch LSST ~100x deeper than SDSS

i~19 vs. i~24

M_i ~-24 vs. -28 from z=0 to ~5.5

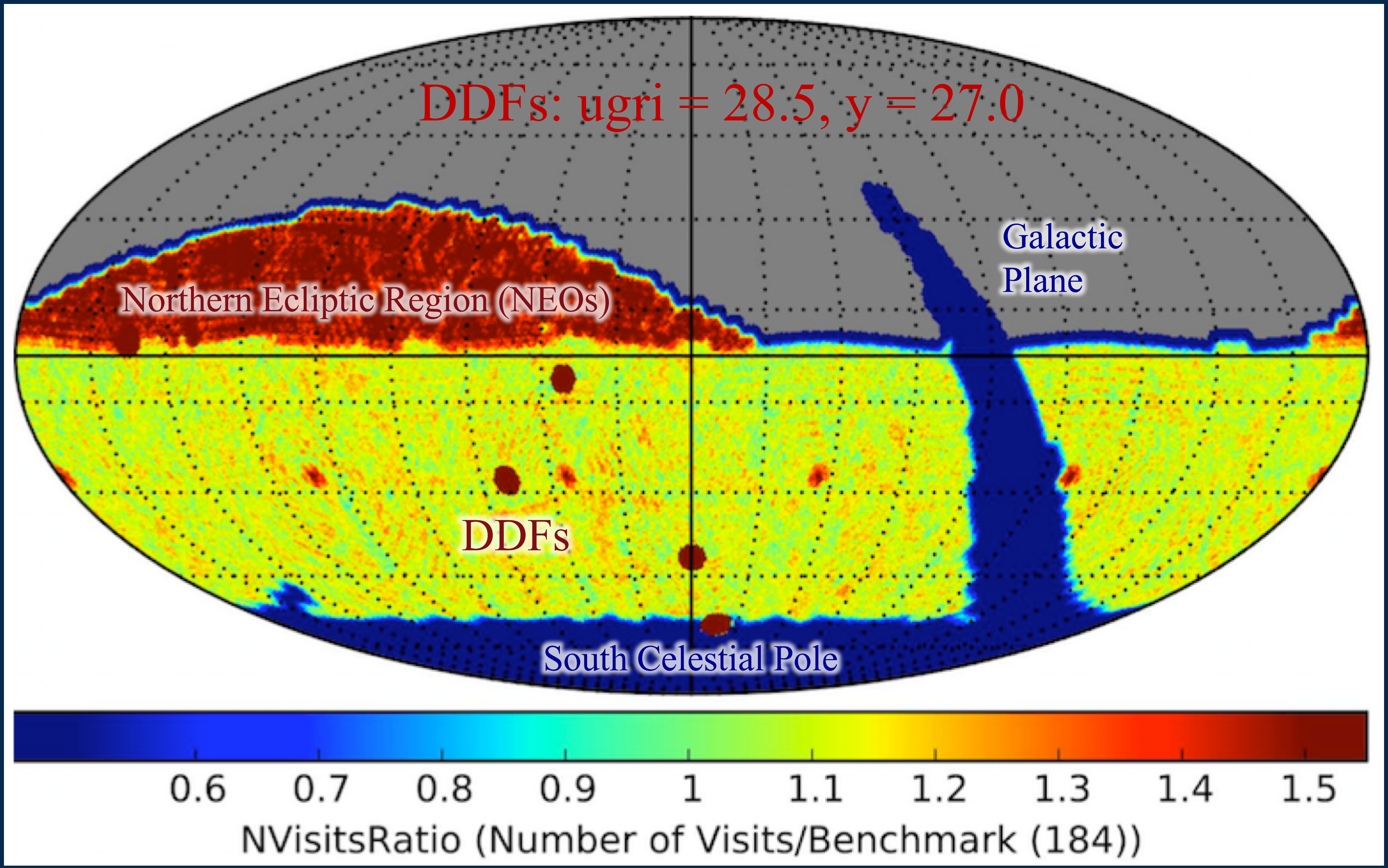
LSST: AGN Expectations



- ~300 million AGNs *detected* by LSST
- Obscuration and host-galaxy dilution will hinder AGN *selection*.
- Recognize ~100 million as such by combing with multi-wavelength observations.

LSST: Coverage

Operations Simulation of *r*-Band Visits



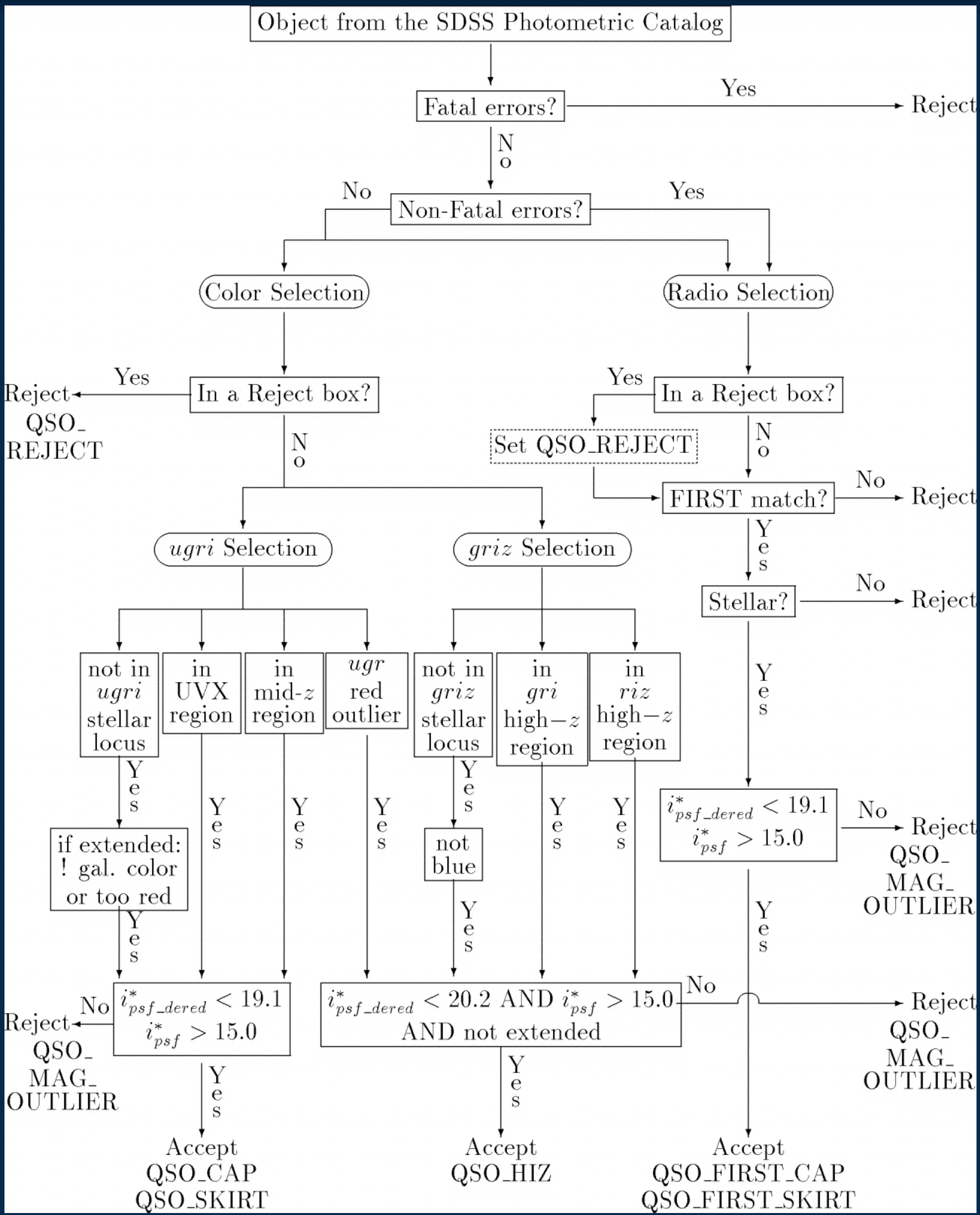
Finding AGNs: The Old Way

Complex, Glorified Color Cuts

Perfectly fine when coupled with spectroscopy to find specific needles in haystack.

Completely inadequate for LSST

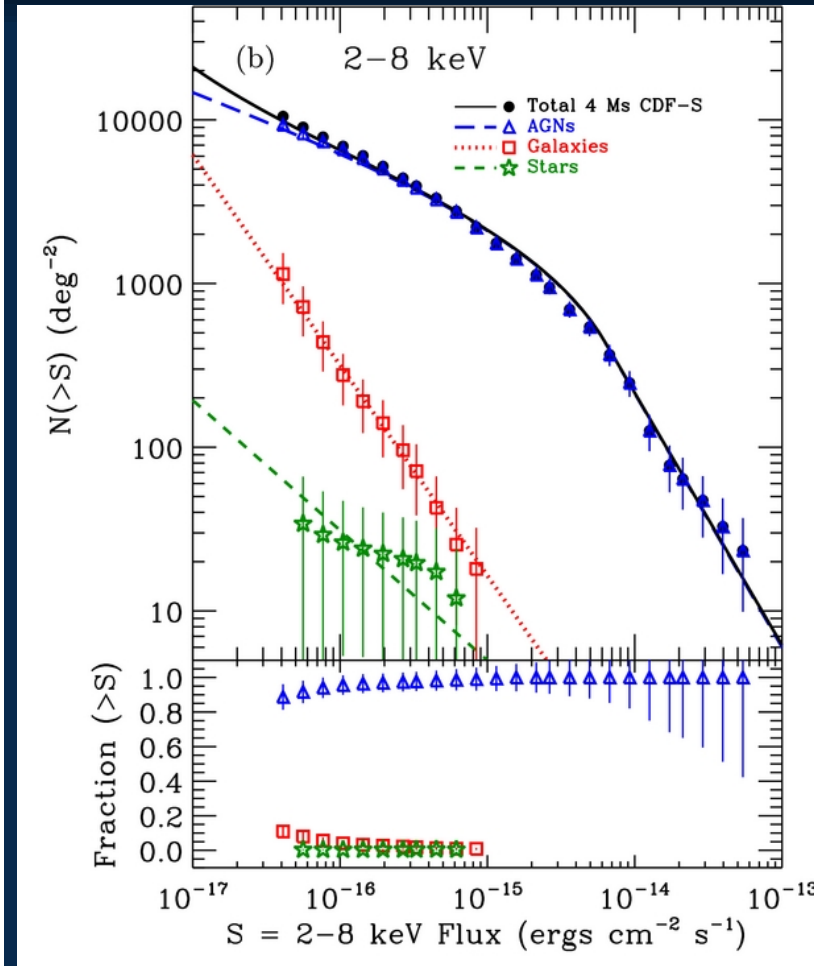
Richards et al. 2002



5 Classes of AGNs to Find

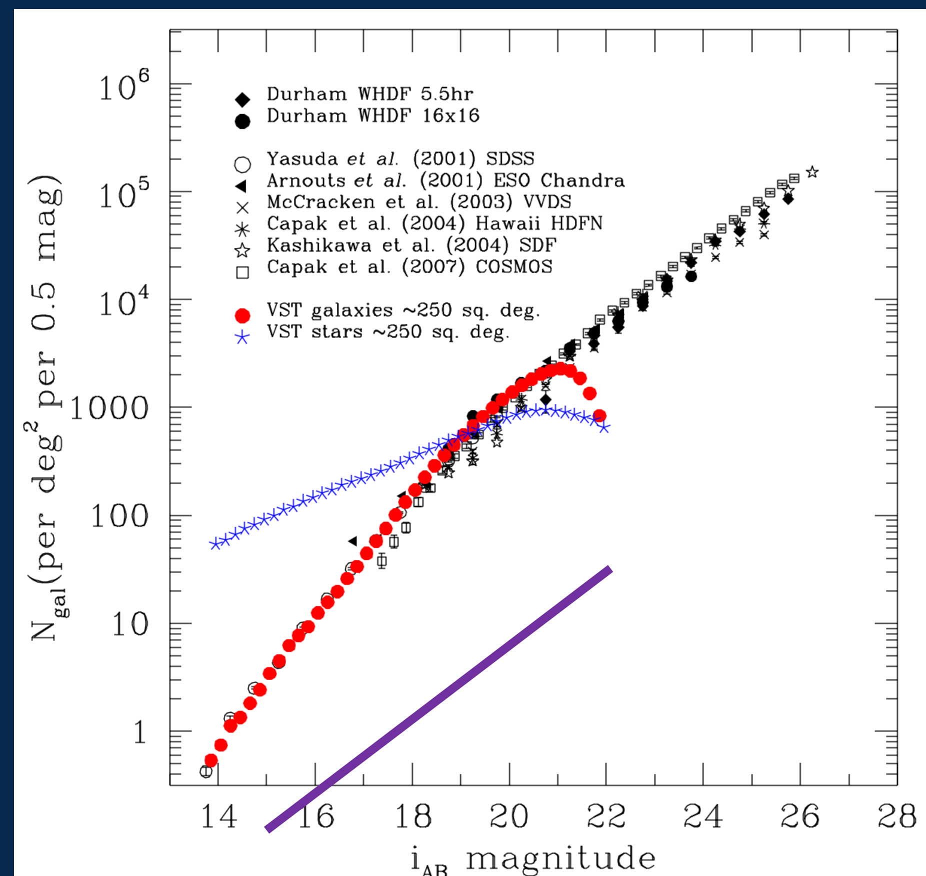
1. unobscured quasars
2. lower-luminosity AGNs
3. very high-z quasars
4. obscured quasars/AGNs
5. transient BH fueling events

Keeping AGNs from being Elusive



Lehmer et al. 2012

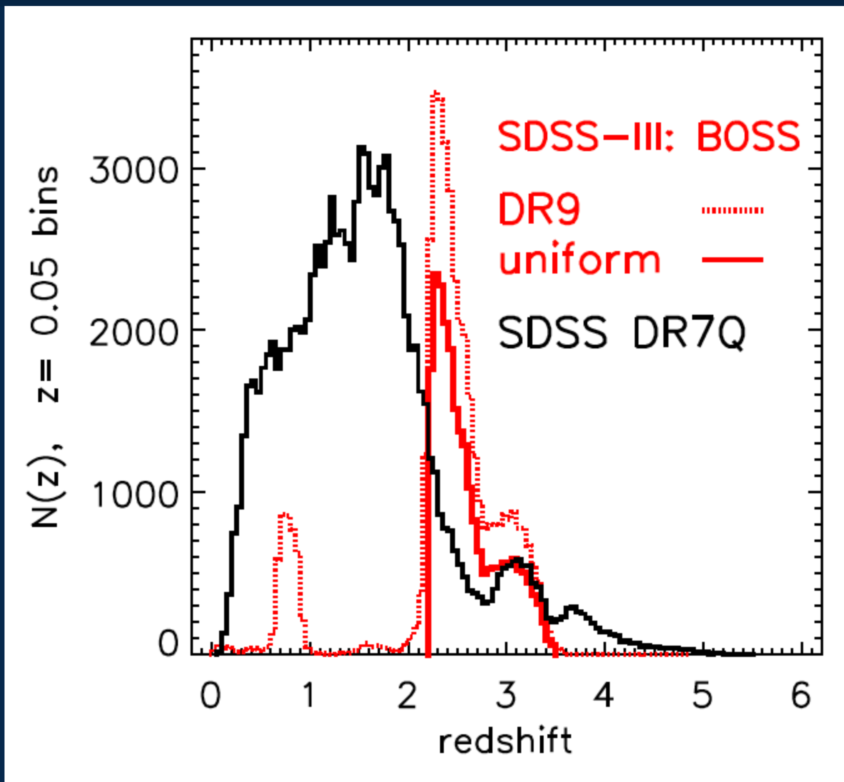
Bright Optical sources are not necessarily AGN



Bright X-ray sources are AGN

Shanks et al. 2015

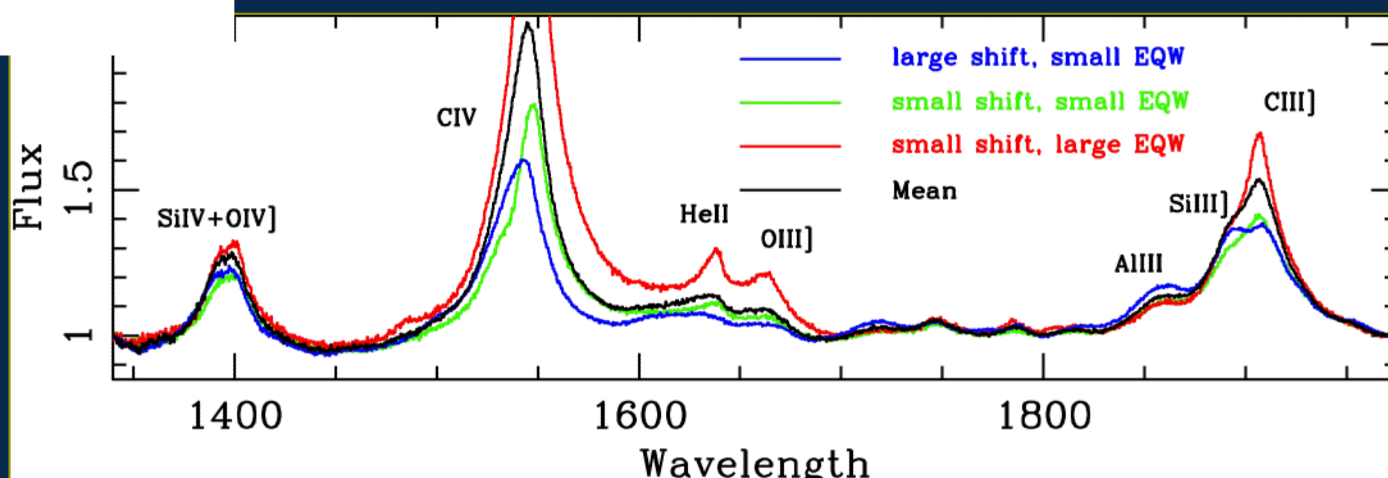
Keeping AGNs from being Elusive



Ross et al. 2013

Avoid redshift "holes"

Be sensitive to the full diversity of the population



Richards et al. 2011

5 Classes of AGNs to Find

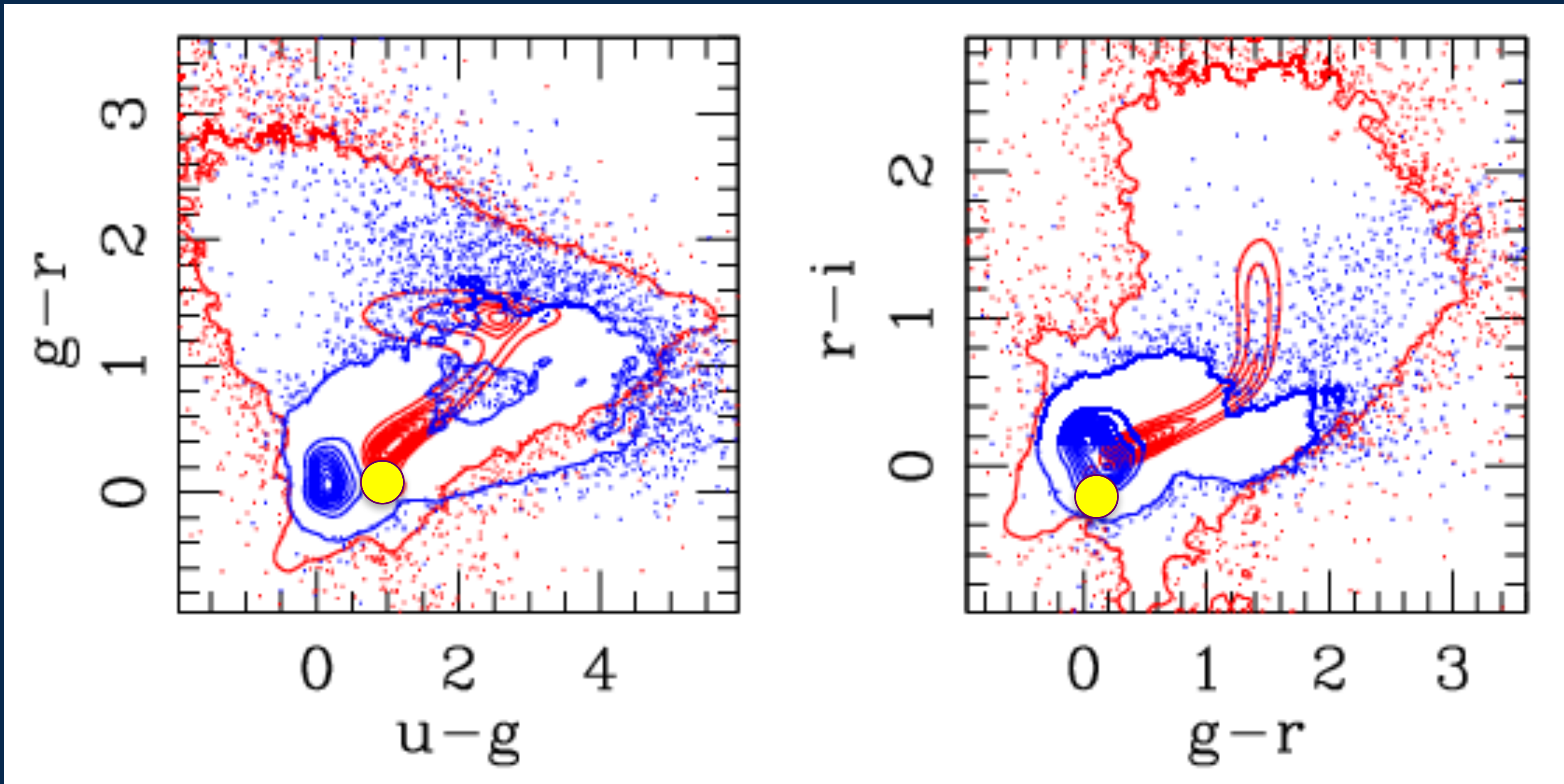
1. unobscured quasars
2. lower-luminosity AGNs
3. very high-z quasars
4. obscured quasars/AGNs
5. transient BH fueling events

LSST: AGN Selection

Multicolor selection in *ugrizy* from z = 0-7.5

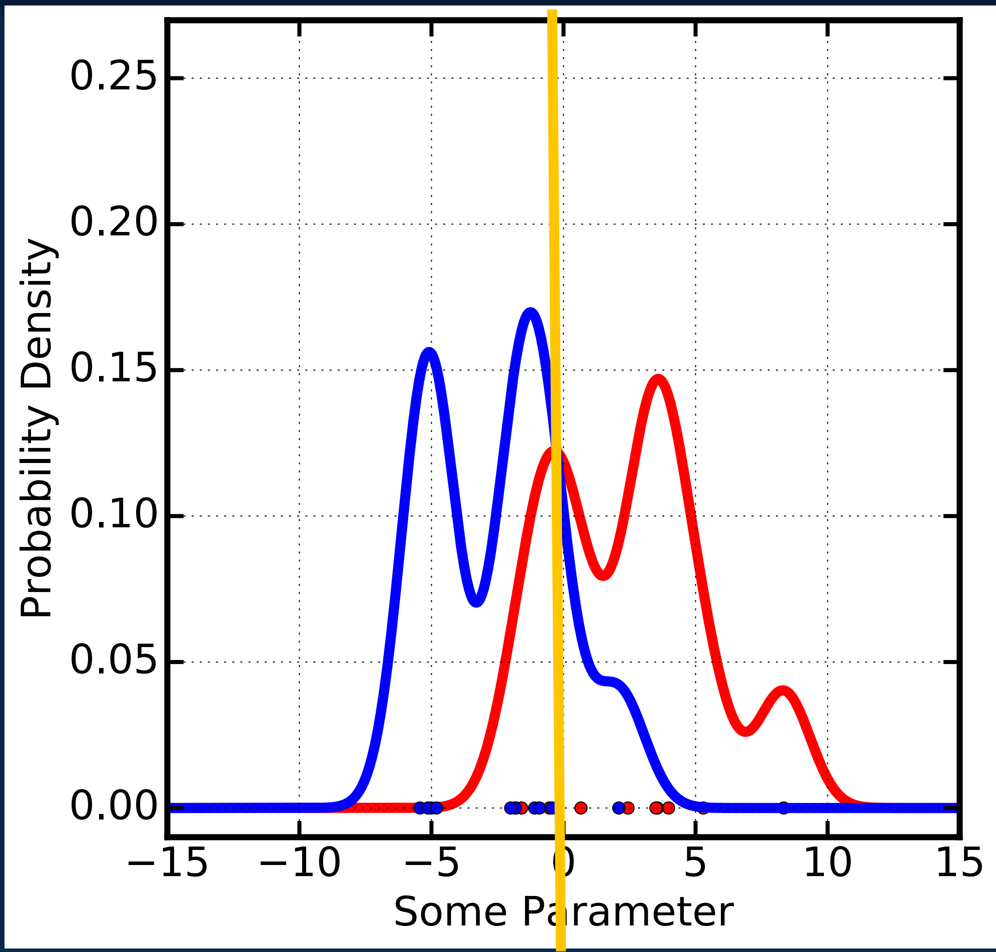
- Works best when $L_{AGN} > L_{Host}$

KDE: Richards et al. 2004, 2009ab, 2015



Given two training sets, Here quasars and stars (non-quasars), and an unknown object, which class is more likely?

KDE Methodology

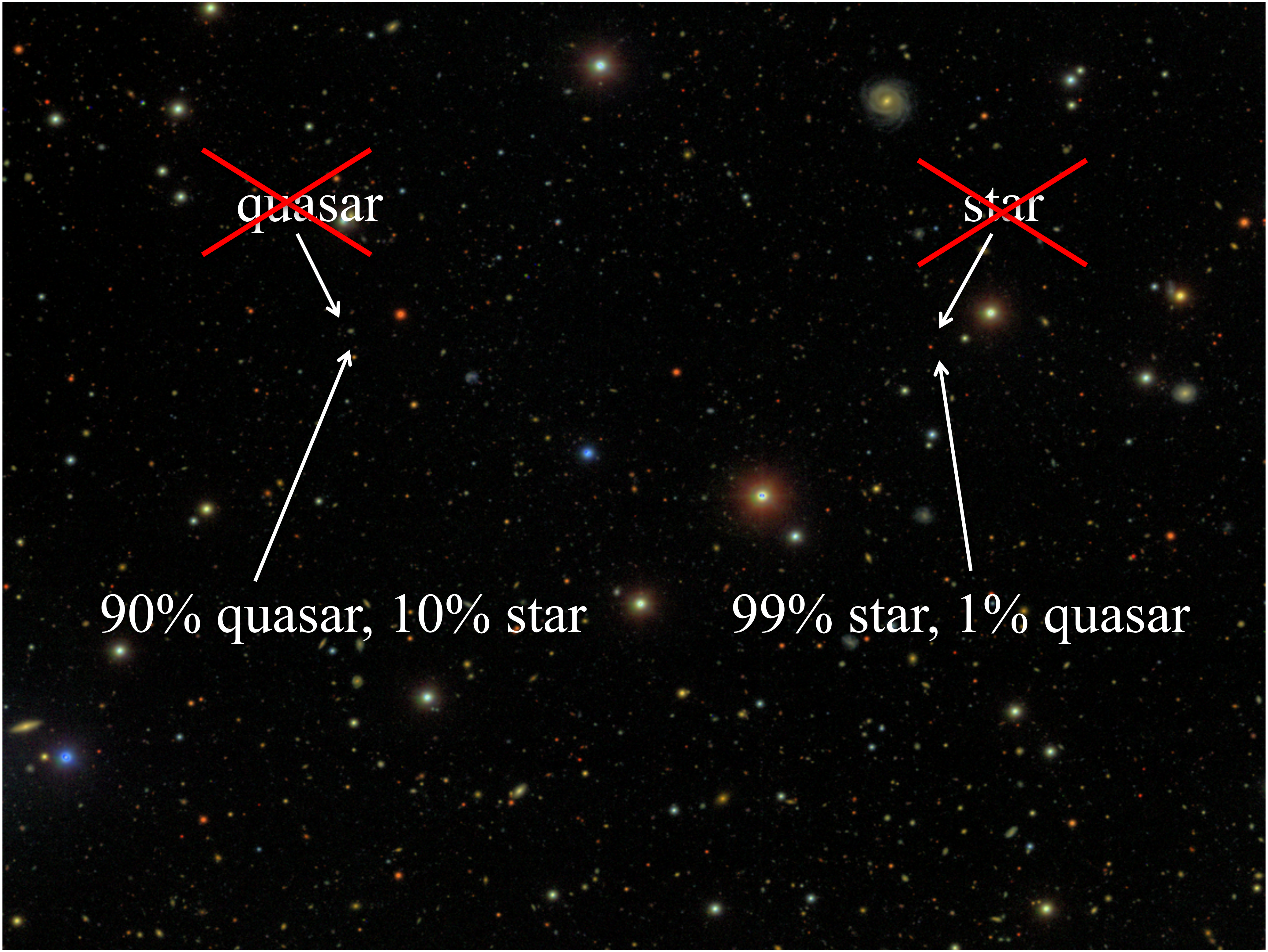


- Compare new object to N-dimensional PDFs of training sets

- Can include multi-wavelength data

Star if P(Star|x)>0.5,
QSO if P(Star|x)<0.5

$$P(Star \mid x) = \frac{P(x \mid Star)P(Star)}{P(x \mid Star)P(Star) + P(x \mid QSO)P(QSO)}$$


~~quasar~~
~~star~~
90% quasar, 10% star
99% star, 1% quasar

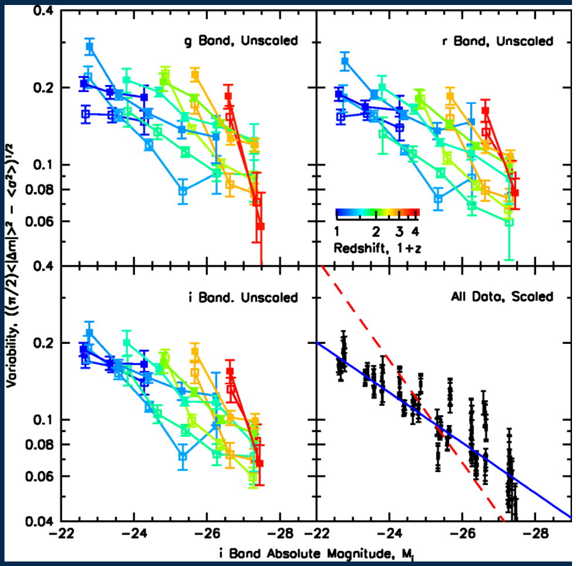
LSST: AGN Selection

Multicolor selection in *ugrizy* from z = 0-7.5

- Works best when $L_{\mathrm{AGN}} > L_{\mathrm{Host}}$

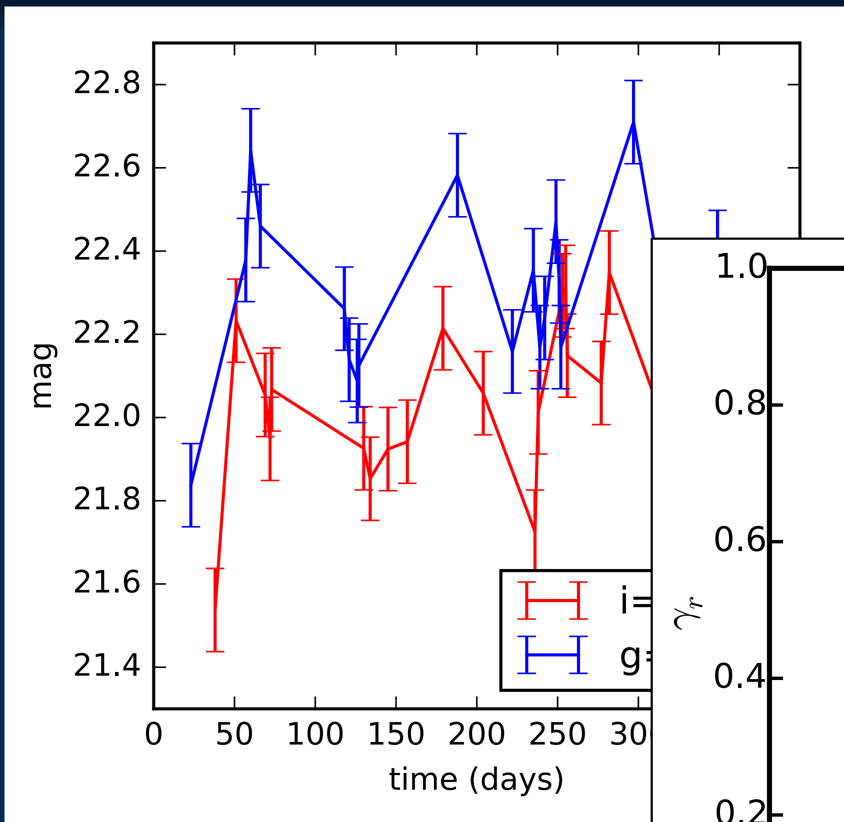
Variability selection

- 55-185 samplings per band over 10 yr
- Highly effective complement to color selection
- Need to assess effectiveness when $L_{\mathrm{AGN}} \sim L_{\mathrm{host}}$

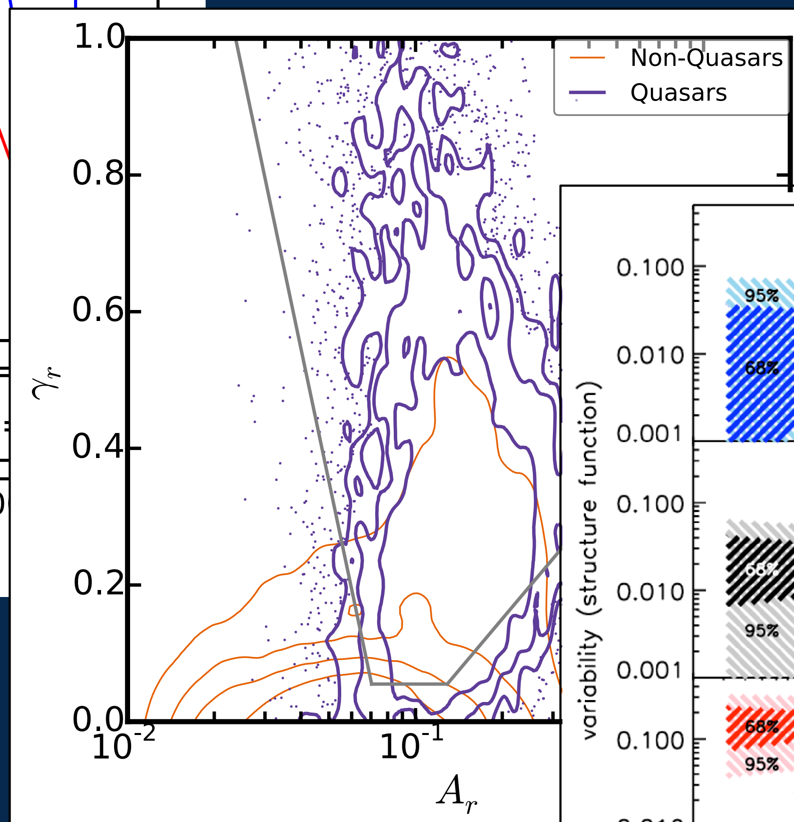


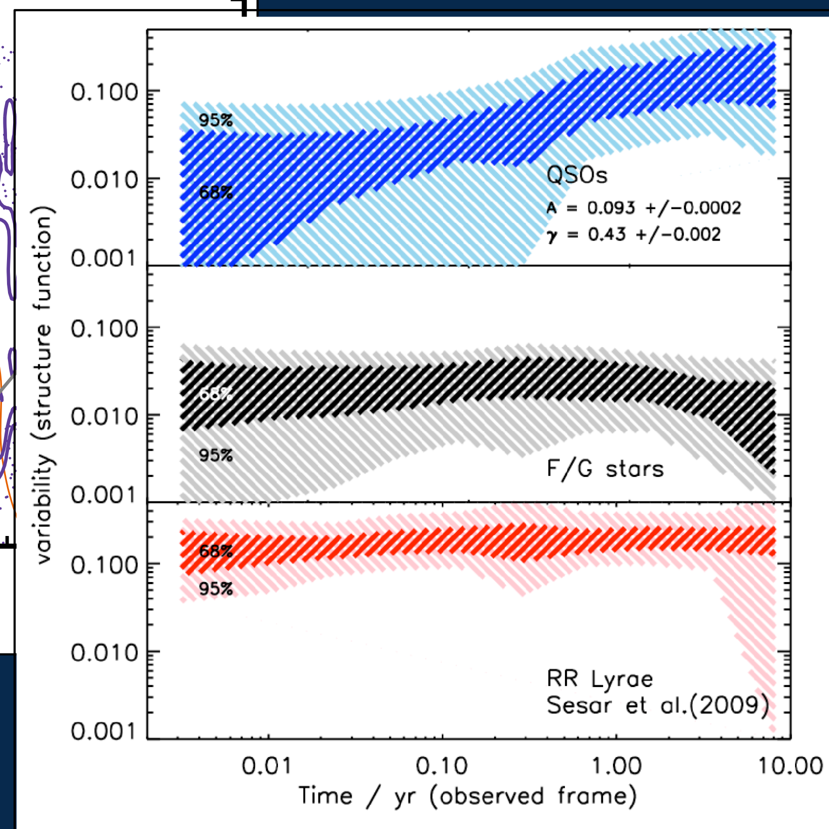
Vanden Berk et al. 2004

Selection by Variability



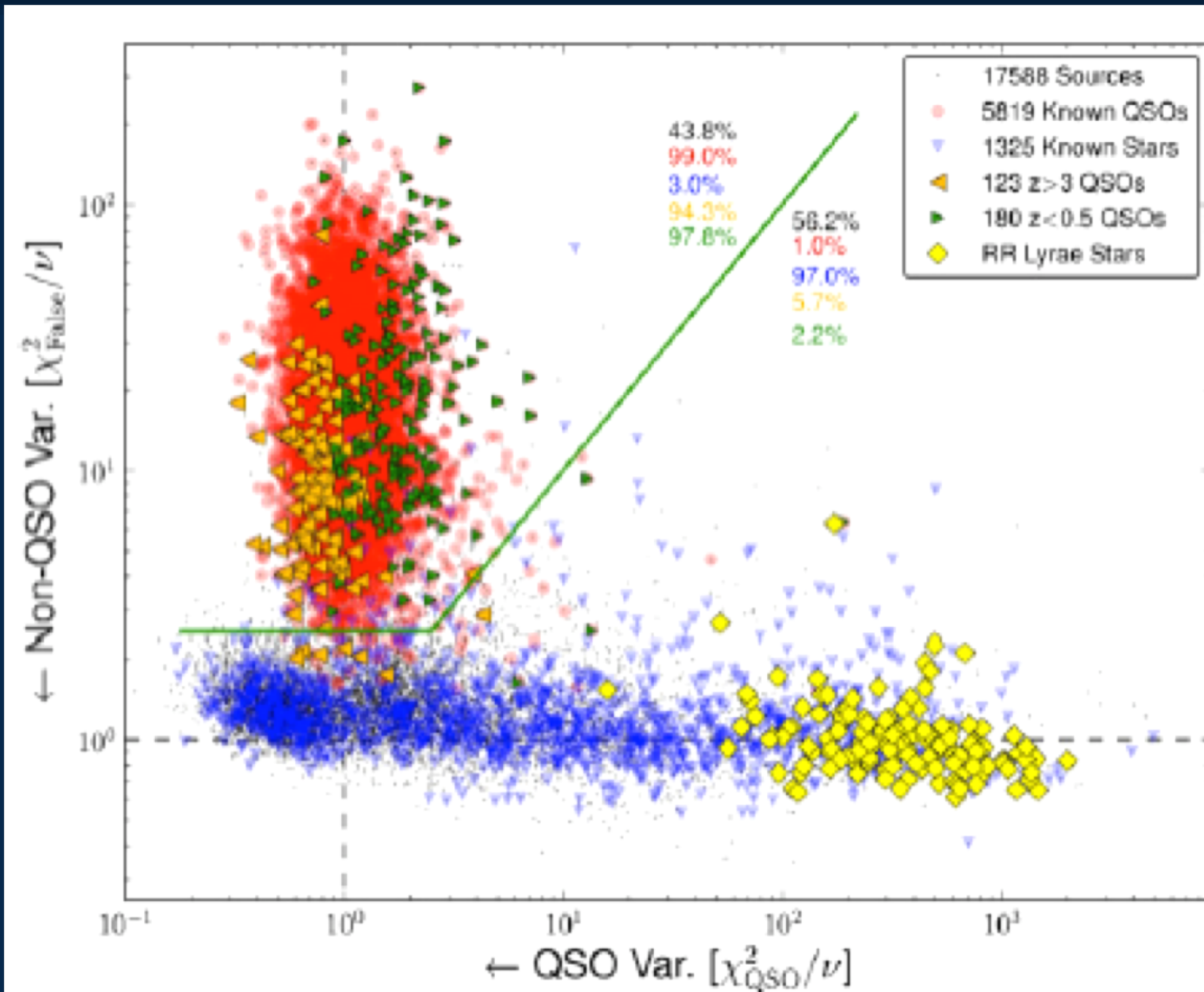
Peters et al. 2015





Schmidt et al. 2010

Selection by Variability



Butler & Bloom 2011

Pros: new source of information, cuts in likelihood space

Cons: need to parameterize the light curves, still cuts, and still variability only, LLAGNs

LSST: AGN Selection

Multicolor selection in *ugrizy* from z = 0-7.5

- Works best when $L_{\mathrm{AGN}} > L_{\mathrm{Host}}$

Variability selection

- 55-185 samplings per band over 10 yr
- Highly effective complement to color selection
- Need to assess effectiveness when $L_{\mathrm{AGN}} \sim L_{\mathrm{host}}$

Astrometry

- Minimizes confusion with stars
- Proper Motion: will reach ~1mas/yr at r~24
- DCR: Improves photo-z

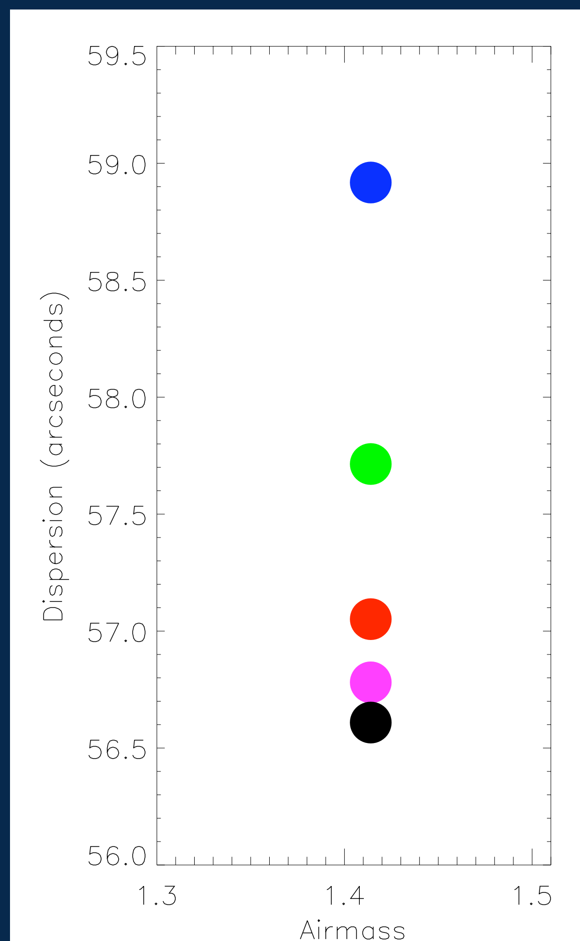
Astrometric Data

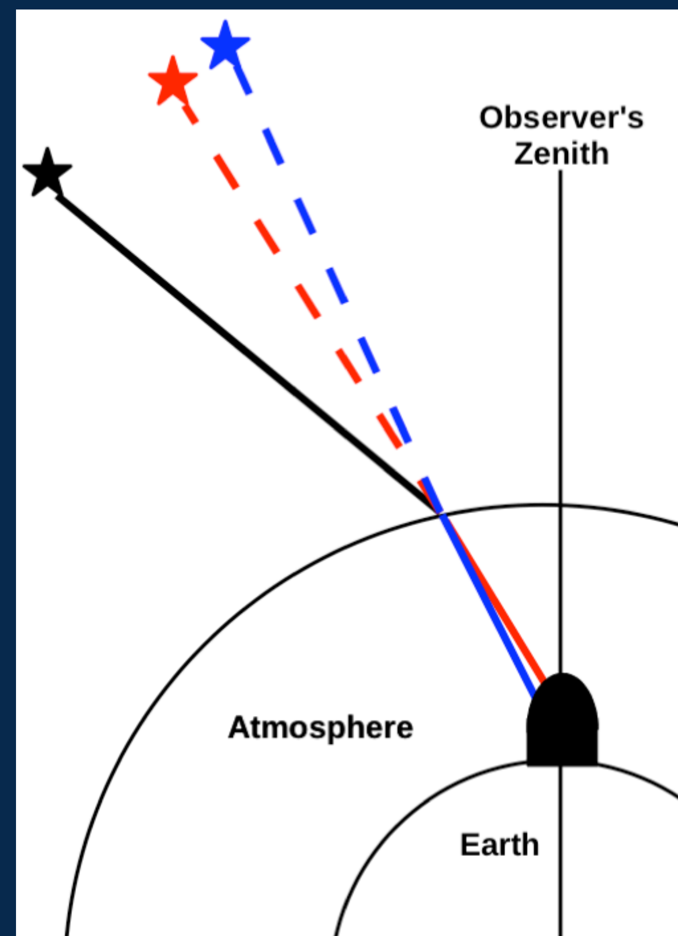
Is the object moving?
If not can we learn something from Differential Chromatic Refraction?

In short: use the atmosphere as a prism

Only works in u and g

Kaczmarczik et al. 2009
Peters et al. 2016

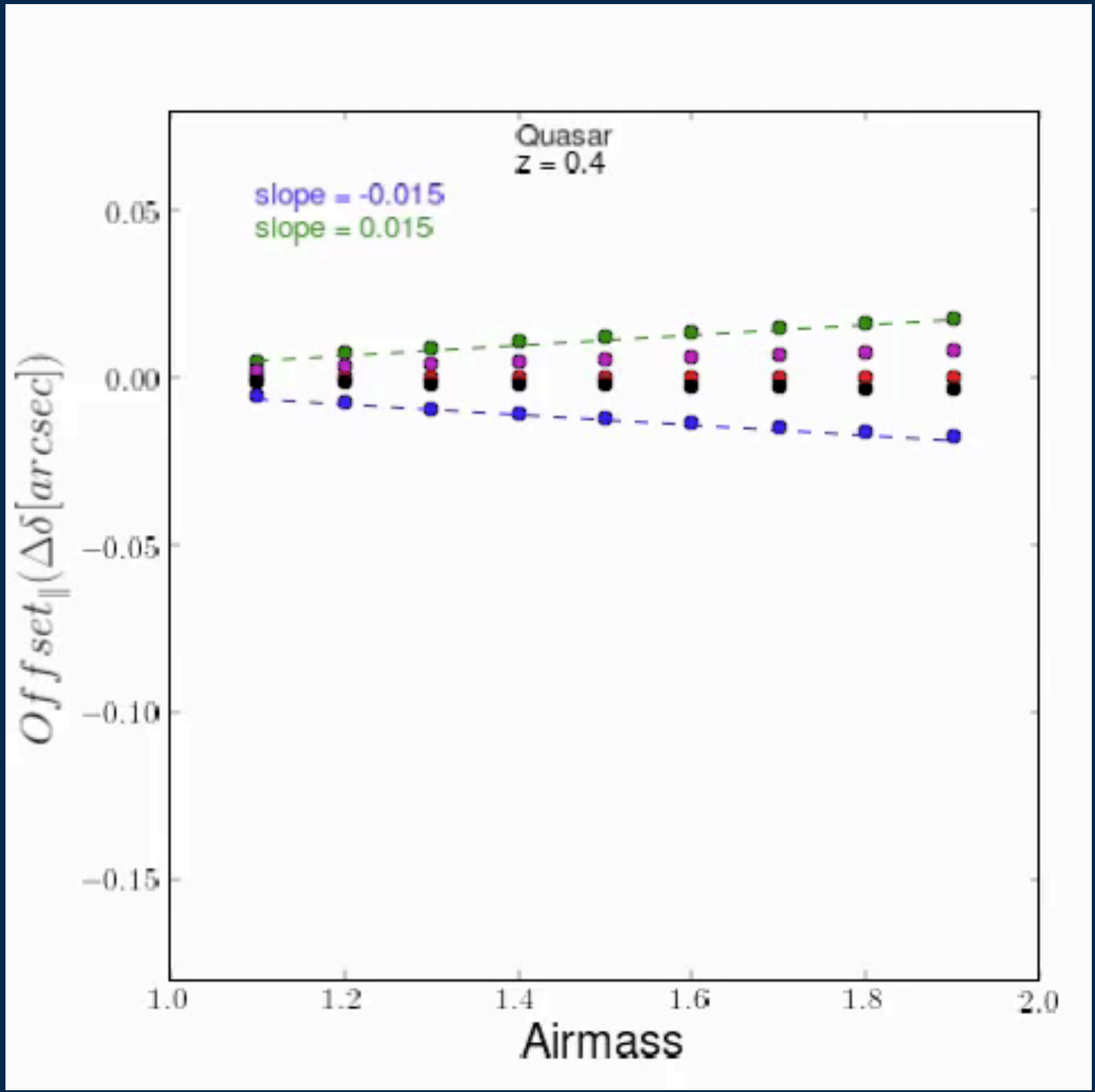




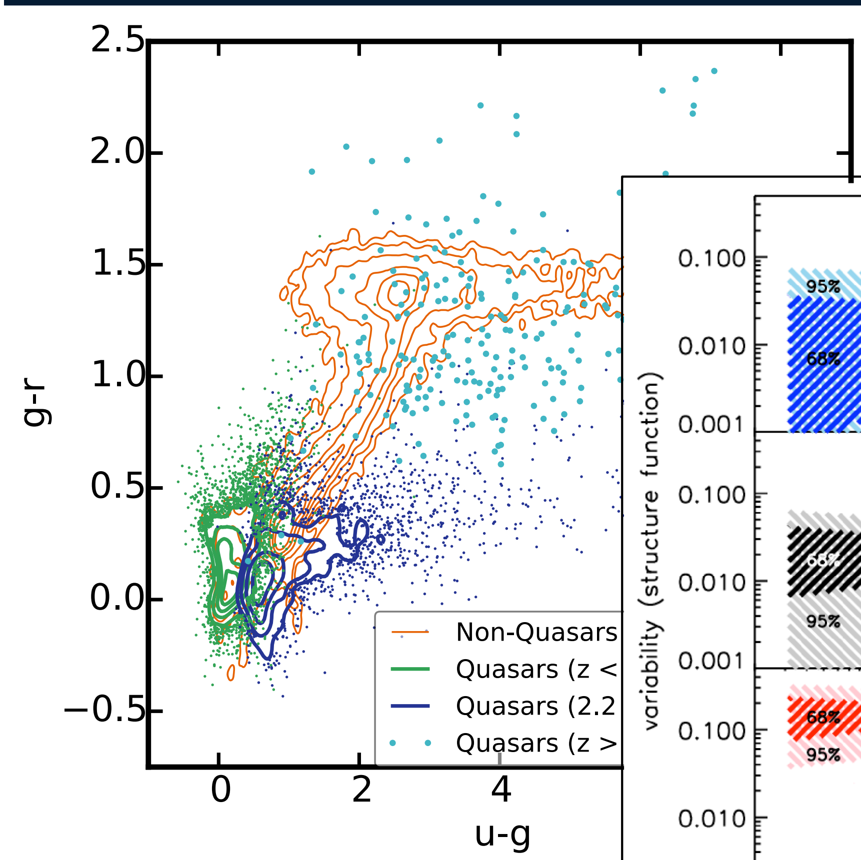
Differential Chromatic Refraction

Is a function of redshift and airmass.

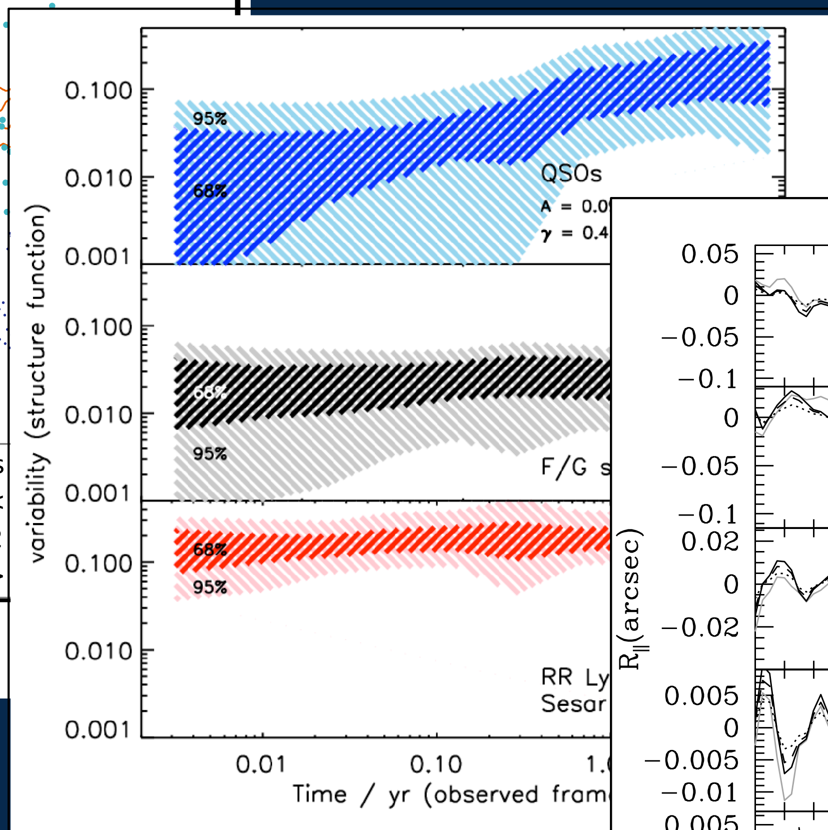
Can aid both selection and photo-z.



Mixed-Attribute Selection

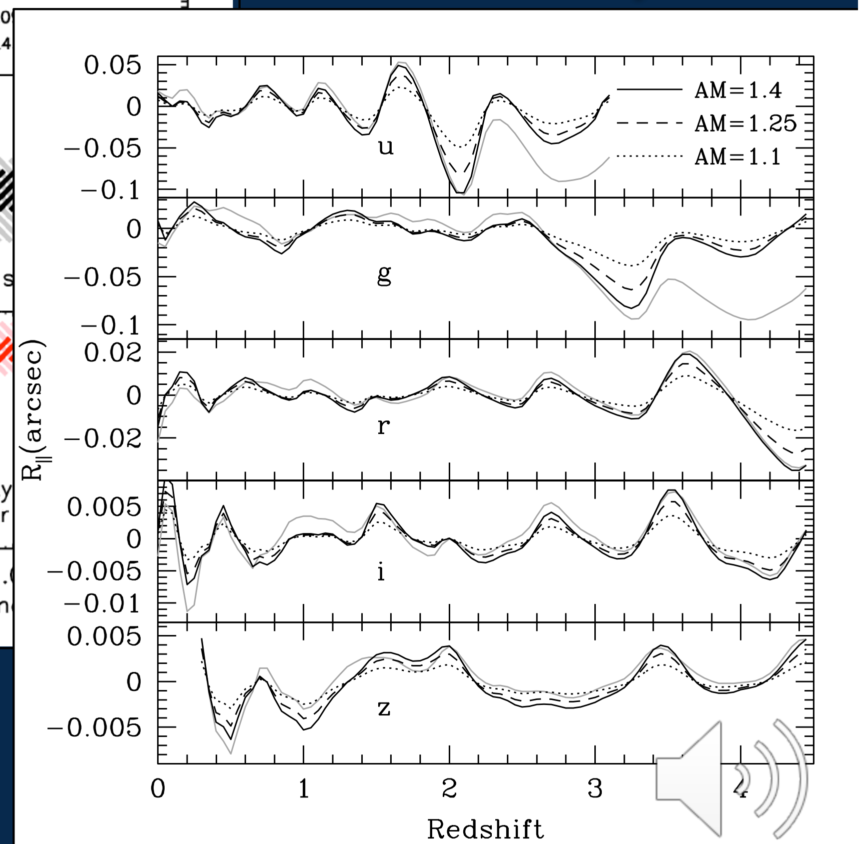


colors



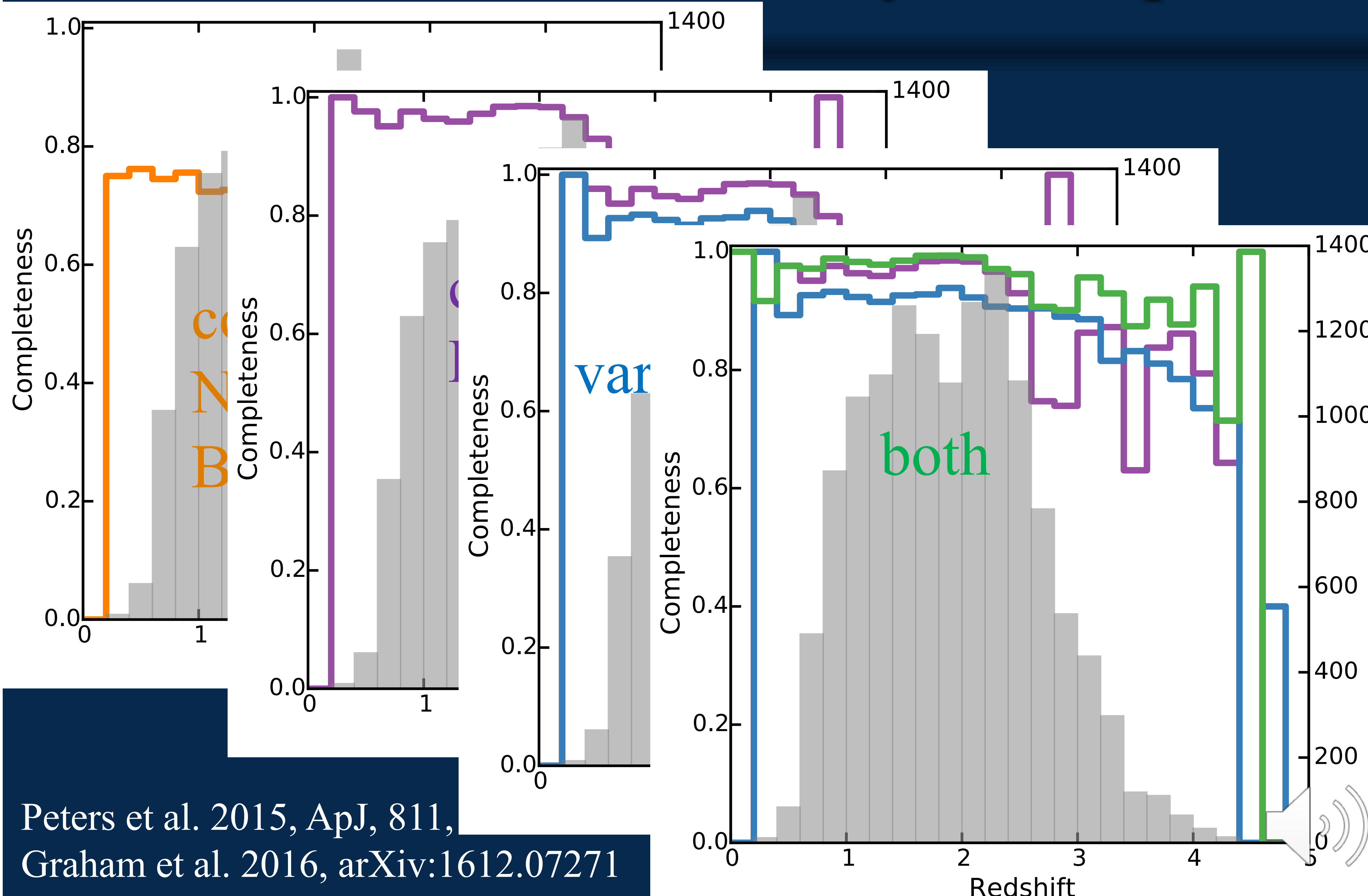
variability

astrometry



Also proper motion and multi-wavelength data.

Color AND Variability Examples



Peters et al. 2015, ApJ, 811,
Graham et al. 2016, arXiv:1612.07271

Kitchen-sink Selection

Simultaneously use all of these attributes to identify accretion SMBHs:

- color
- variability
- astrometry (differential chromatic refraction and proper motion)
- brightness
- Galactic latitude
- morphology
- probability of belonging to another class
- multi-wavelength matching

Machine Learning Algorithms

Need to start using Machine Learning algorithms (e.g., Scikit-Learn).
Links are to some work I have started.
Scikit-Learn itself is not up to the task (but is a good place to start).

```
In [1]: from astropy.table import Table
        import numpy as np

In [ ]: # Read in training file
        data = Table.read('GTR-ADM-QSO-ir-testhighz_findbw_lup_2016_starclean.fits')
        Xtrain = np.vstack([ data['ug'], data['gr'], data['ri'], data['iz'], data['zs1'], data['s1s2']]).T
        ytrain = np.array(data['labels'])

In [ ]: # "Whiten" the data
        from sklearn.preprocessing import StandardScaler
        scaler = StandardScaler()
        Xtrain_scaled = scaler.fit_transform(Xtrain)

In [ ]: # Instantiate the Random Forest classifier
        from sklearn.ensemble import RandomForestClassifier
        rfc = RandomForestClassifier(n_estimators=10, max_depth=15, min_samples_split=2, n_jobs=-1, random_state=42)
        rfc.fit(Xtrain_scaled, ytrain)
```

https://github.com/gtrichards/QuasarSelection
https://github.com/gtrichards/PHYS_T480 (based on Ivezic et al.)

Photo-z

- Without spectroscopy, accurate and precise photo-z estimates will be crucial

- Determine Photo-z Probability Distribution Function (PDF)

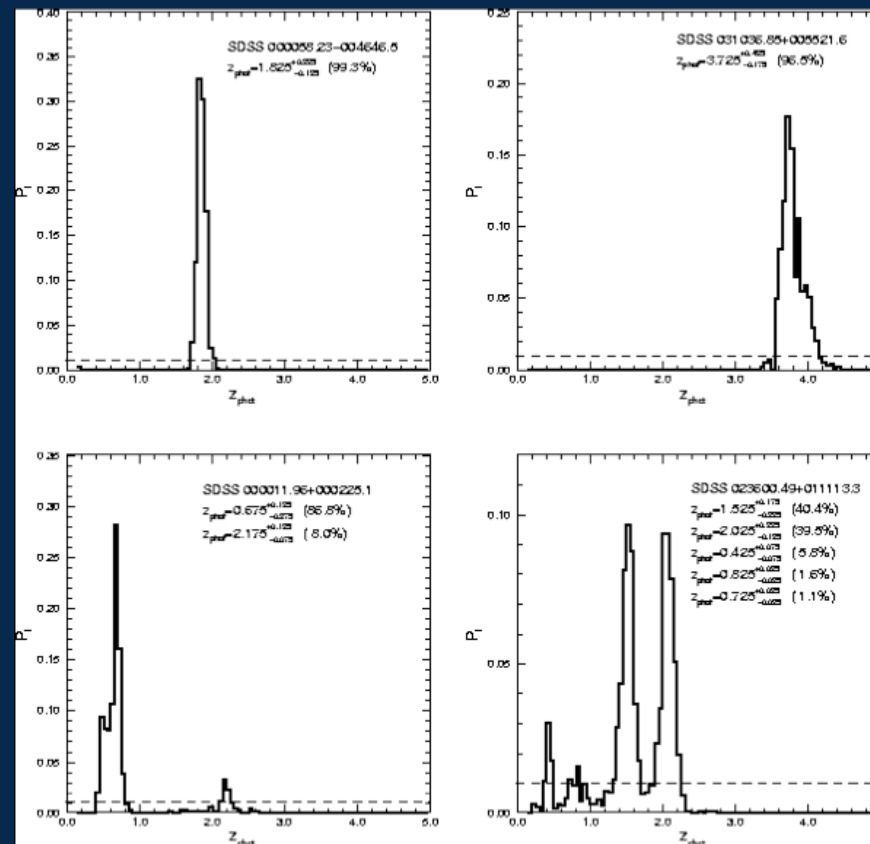


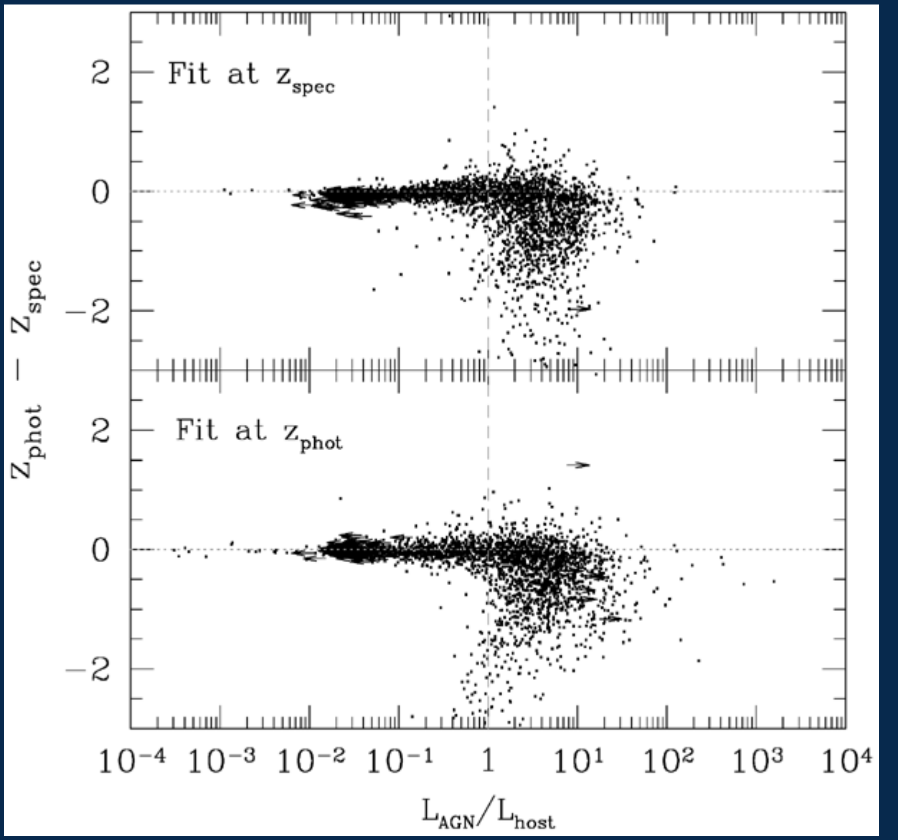
Photo-z

- Without spectroscopy, accurate and precise photo-z estimates will be crucial

- Determine Photo-z Probability Distribution Function (PDF)

- Need smooth transition between host- and nuclear-dominated sources (where template and empirical methods work best, respectively).

Assef et al. 2010

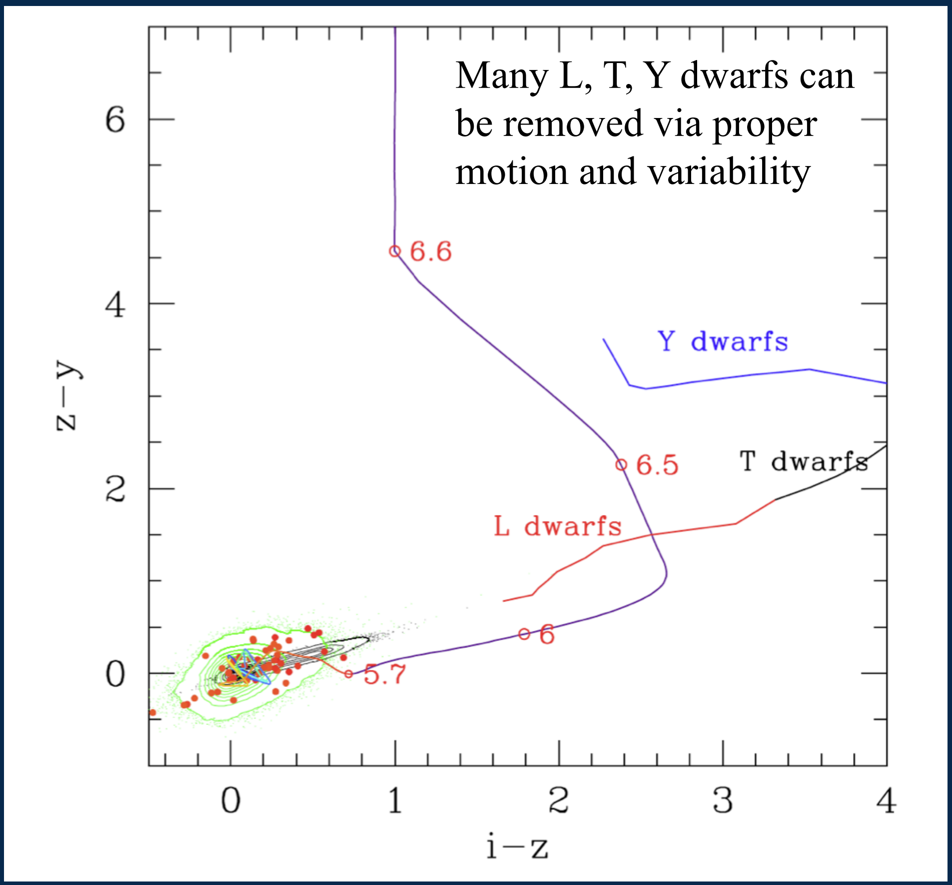


5 Classes of AGNs to Find

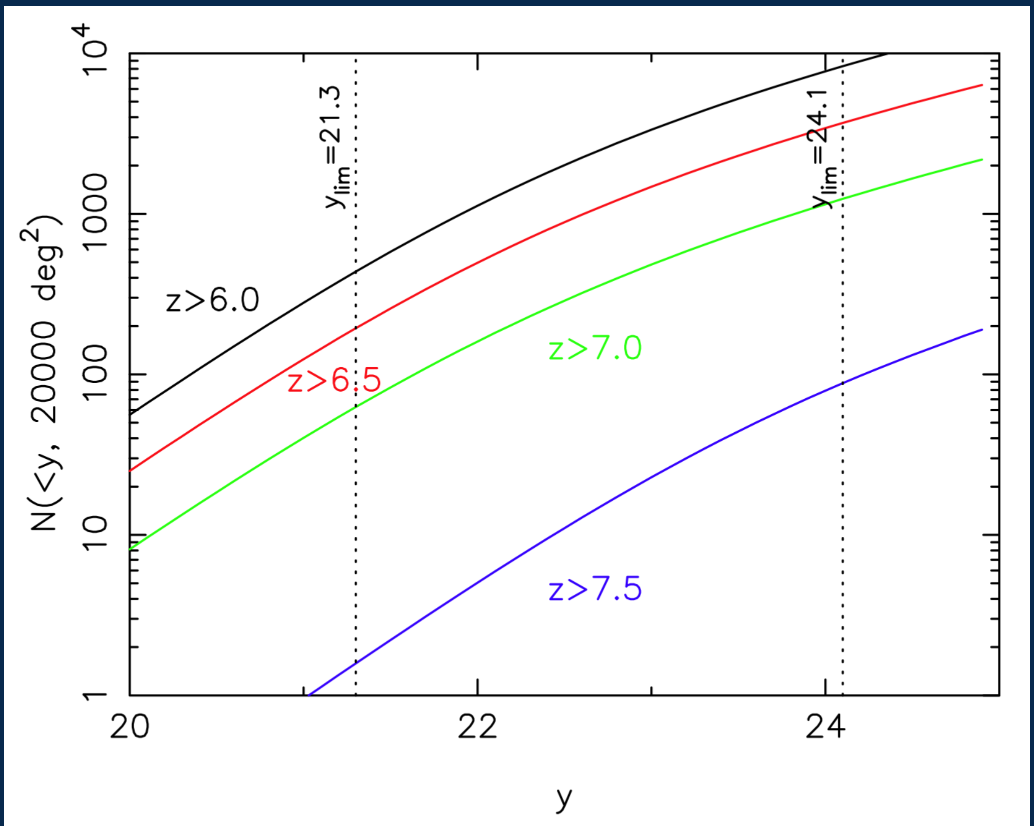
1. unobscured quasars
2. lower-luminosity AGNs
3. very high-z quasars
4. obscured quasars/AGNs
5. transient BH fueling events

High-Redshift AGN Selection

Colors of High-Redshift Quasars



Expected Numbers of $z > 6$ Quasars



- LSST alone will provide significant numbers of AGNs to $z \sim 7.5$ (to $L_{Opt} \sim 10^{44}$ erg/s)
- LSST+Euclid: ~1360 at z>7 and 24 at z>10
- LSST+WFIRST: ~1490 at z>7 and ~29 at z>10

5 Classes of AGNs to Find

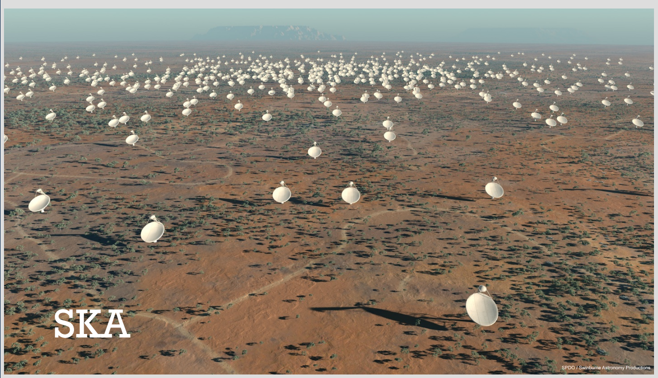
1. unobscured quasars
2. lower-luminosity AGNs
3. very high-z quasars
4. obscured quasars/AGNs
5. transient BH fueling events

Multi-wavelength Data

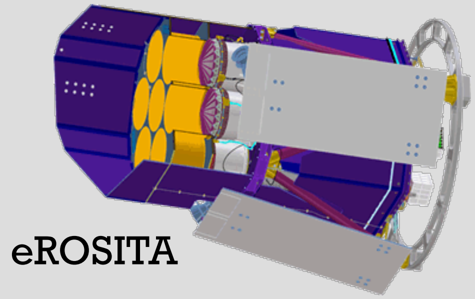
- The last way that LSST will identify BHs is by combining with multi-wavelength data.
- More generally data from “other facilities" (e.g., Euclid, GAIA are optical).
- Crucial for obscured quasars and low-luminosity AGNs.

Multi-wavelength AGN Selection

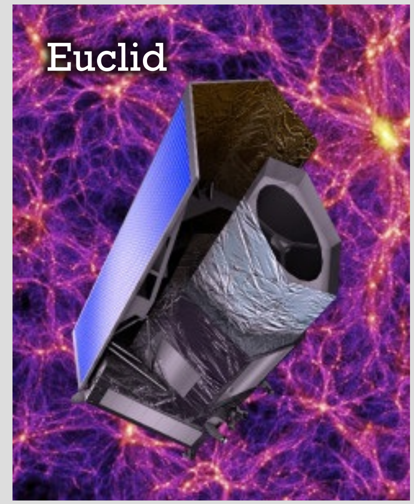
L_{R}, T_{b}, morphology



L_{X} and Γ_{X}



Infrared-optical colors



Need ambassadors

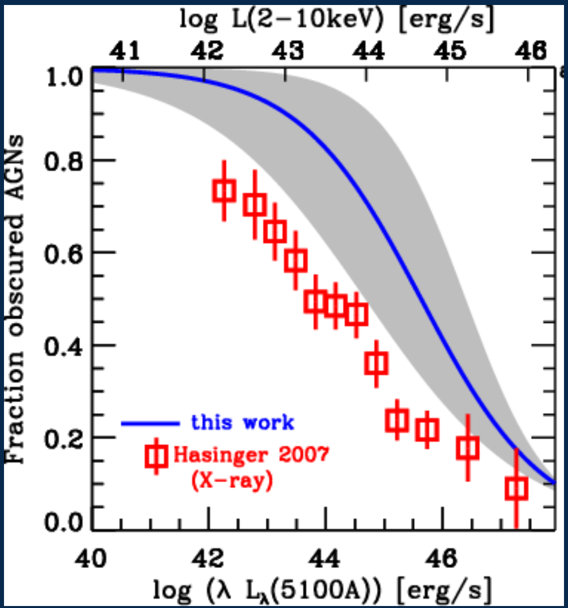
Issues: Resolution, Balkanization, Dropouts, Bandmerging

Type-2 Fraction as f(L,z)

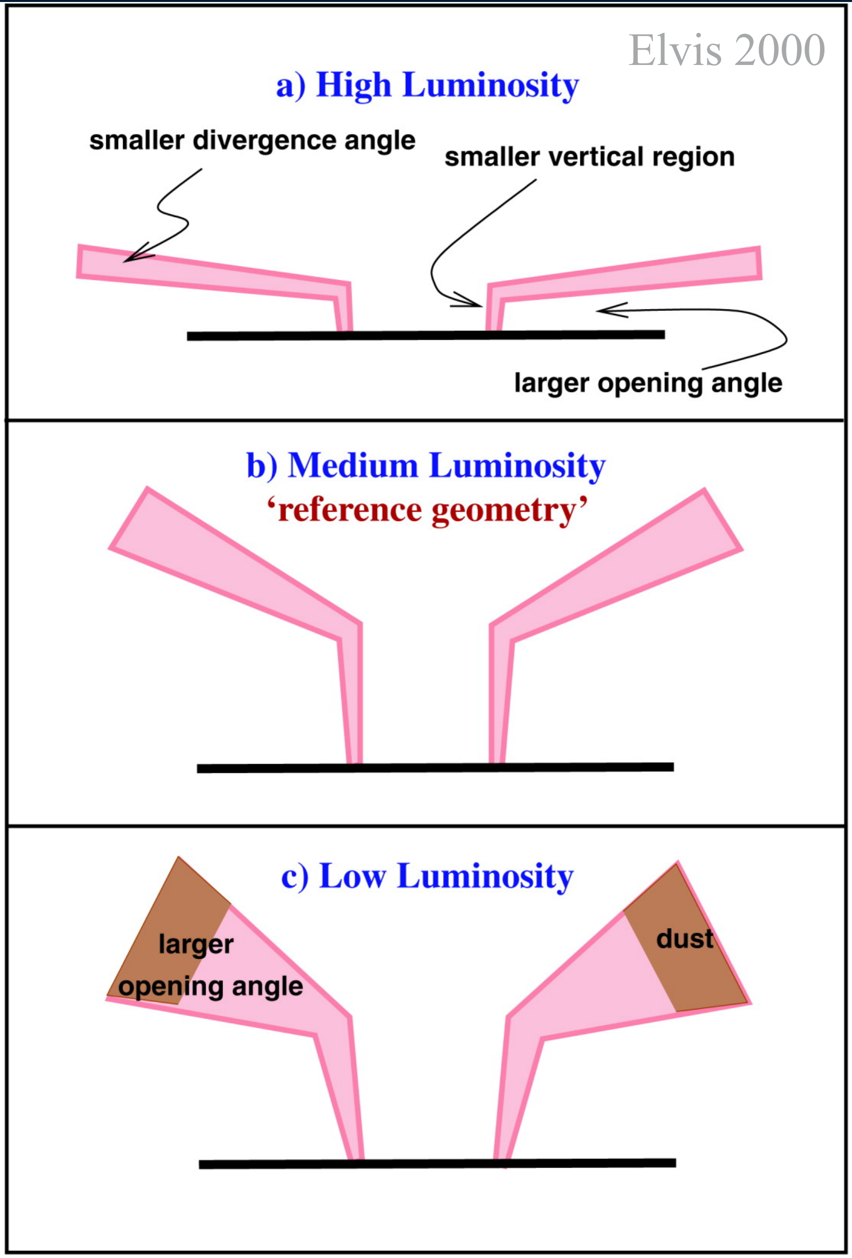
Complete census of SMBHs requires knowing the relative demographics of obscured and unobscured quasars.

Need to know it as a function of L and z.

Type-2:Type-1 generally 1-4:1



Maiolino et al. 2007



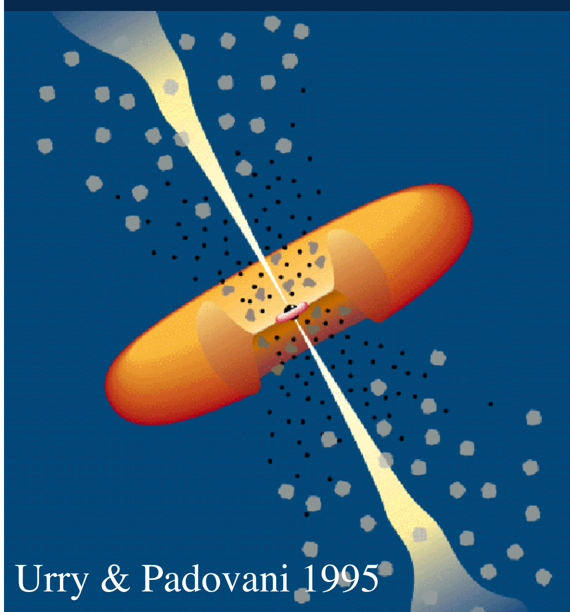
Where are the bright type-2 quasars at high-z???

- SDSS has ~43,000 z>1 quasars with i<19.1 (and >6000 with W4<8).

- If 50-75% of AGN are obscured, then where are the 43,000+ type-2s?

Obscured Quasars

- Irony is that the largest (robust) type-2 AGN samples are from the optical.
  - Zakamska et al. 2003: 291
  - Reyes et al. 2008: 887
  - Alexandroff et al. 2013: 145 (at z>2)
  - Yuan et al. 2016: 2758 (at z<1)



Urry & Padovani 1995

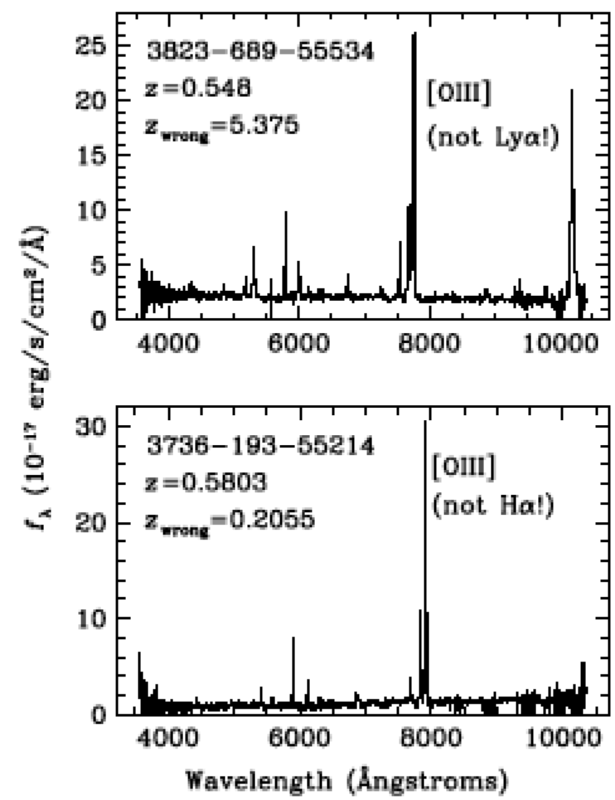
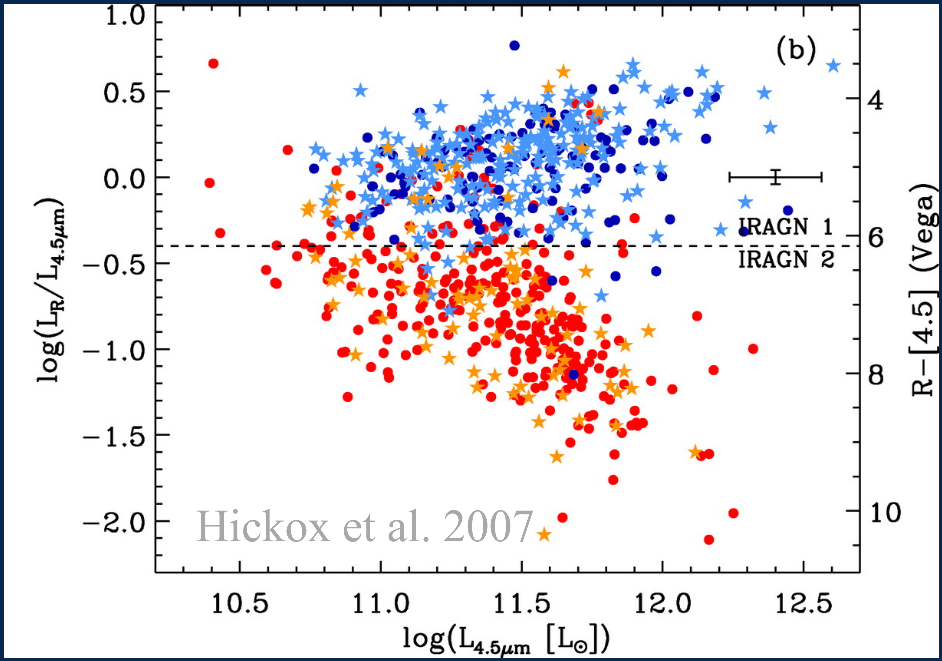


Figure 5. Example type 2 quasars identified assuming that the BOSS pipeline mistook [O III] for another strong emission line: for Lyα in the top panel (true redshift $z_{true} = 0.548$) and for Hα in the bottom panel (true redshift $z_{true} = 0.5803$). Each quasar is indicated with its plate, fiber, and Modified Julian Date (MJD (see § 2.4). These spectra have been smoothed with a five-pixel boxcar.

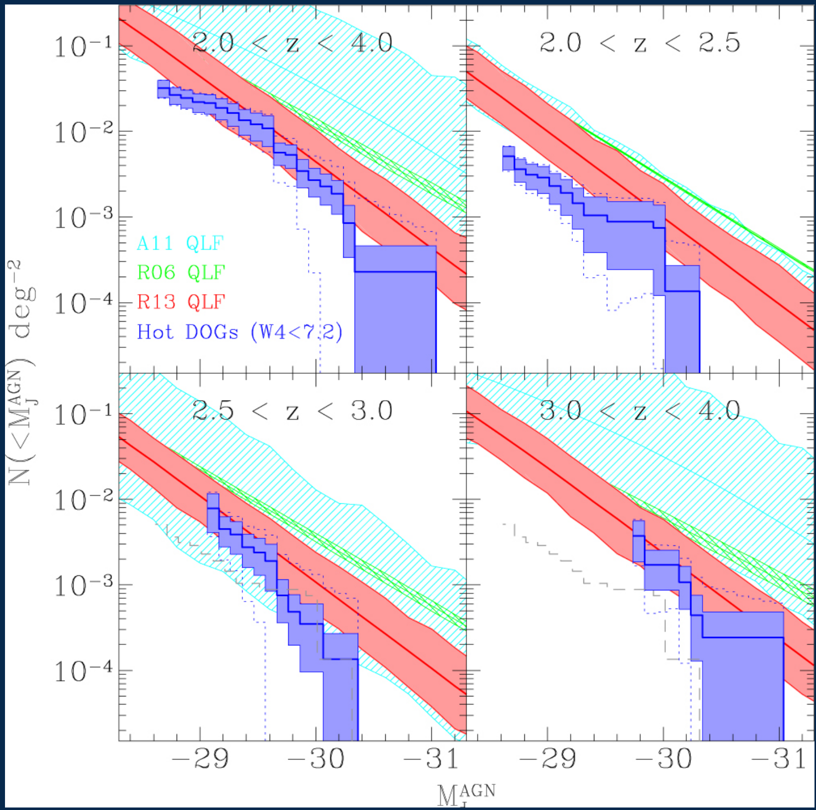
Yuan et al. 2016

High-z Type-2 AGNs

Large numbers that have been identified **photometrically** at other wavelengths.



Assef et al. 2015



What does the literature say?

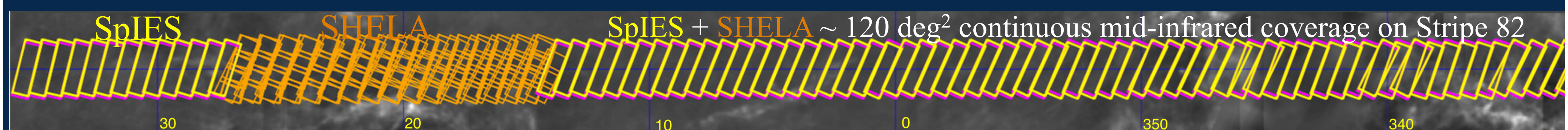
Literature (Over)-Simplification

Ratio of type-2 to type-1 at high-z (relatively bright IR sources)

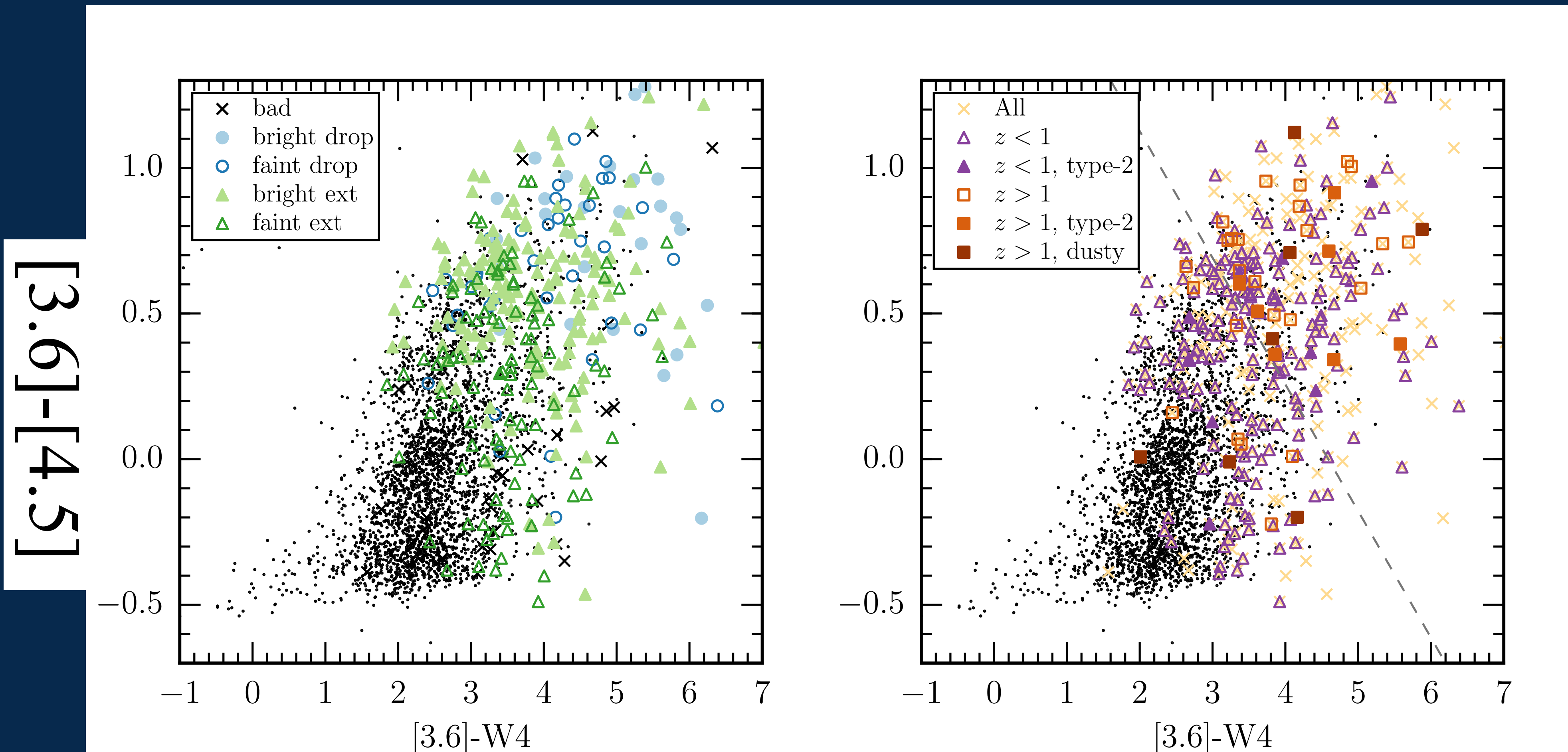
- Photometric selection: ~2:1
  (e.g., Martinez-Sansigre et al.2005, Polletta et al. 2008, Assef et al. 2015)
- Spectroscopic confirmation: ~1:1-2
  (e.g., Lacy et al. 2013)
- Many high-z type-2 candidates are either not high-z or type-2

Our Selection

- Very bright WISE source (W4<8)
- W4 as isotropic as possible for full-sky
- SDSS dropout (or very red r-W4)
- Matched to Spitzer data (for positional accuracy)
- 164 deg^2, mostly SpIES data on SDSS Stripe 82



Our Bright, High-z Sample



Left: coded by selection; Right: coded by spectra
Black points are Lockman Hole for reference

373 Candidates in 164 deg^2

We found:

At z>1

10 type-2 quasars

7 type-1 reddened quasars

4 normal type-1's

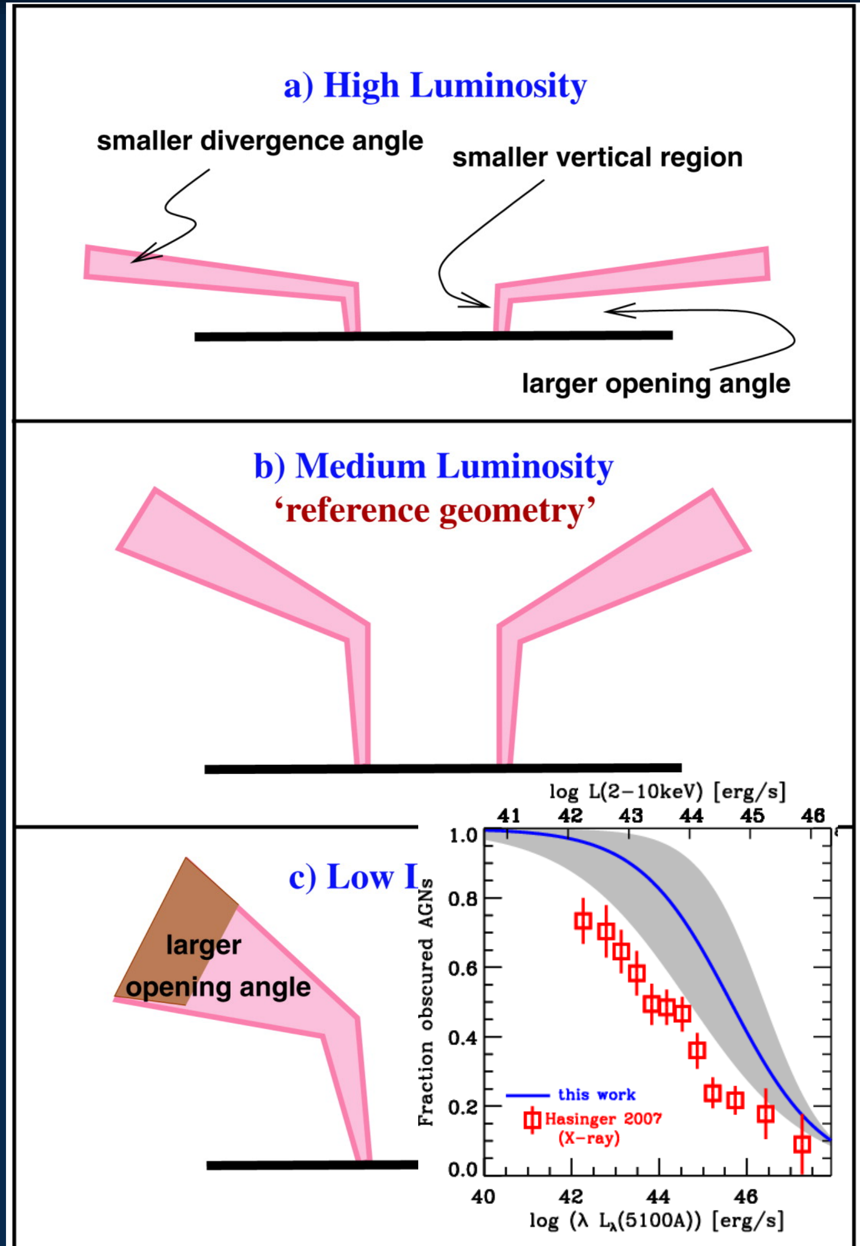
17-114 confirmed/likely obscured at z>1: 0.1-0.7 deg^{-2}

Similarly-selected type-1s: 0.75 deg^{-2}

Type-2 Fraction as f(L,z)

Conclusion:

- For z>1 and W4<8, type-1 density at least as high as type-2. Maybe much higher.
- Either missing "easy" type-2s or important physics (e.g., lack of NLR at high-L).
- Need bigger samples for LSST training sets.



Conclusions

- LSST will *detect* some 300 million AGNs
- Will need to take full advantage of all available information (esp. multi-wavelength) to *identify* ~100 million as AGNs
- High-z quasars (esp. obscured) are a particular challenge (but with great potential)