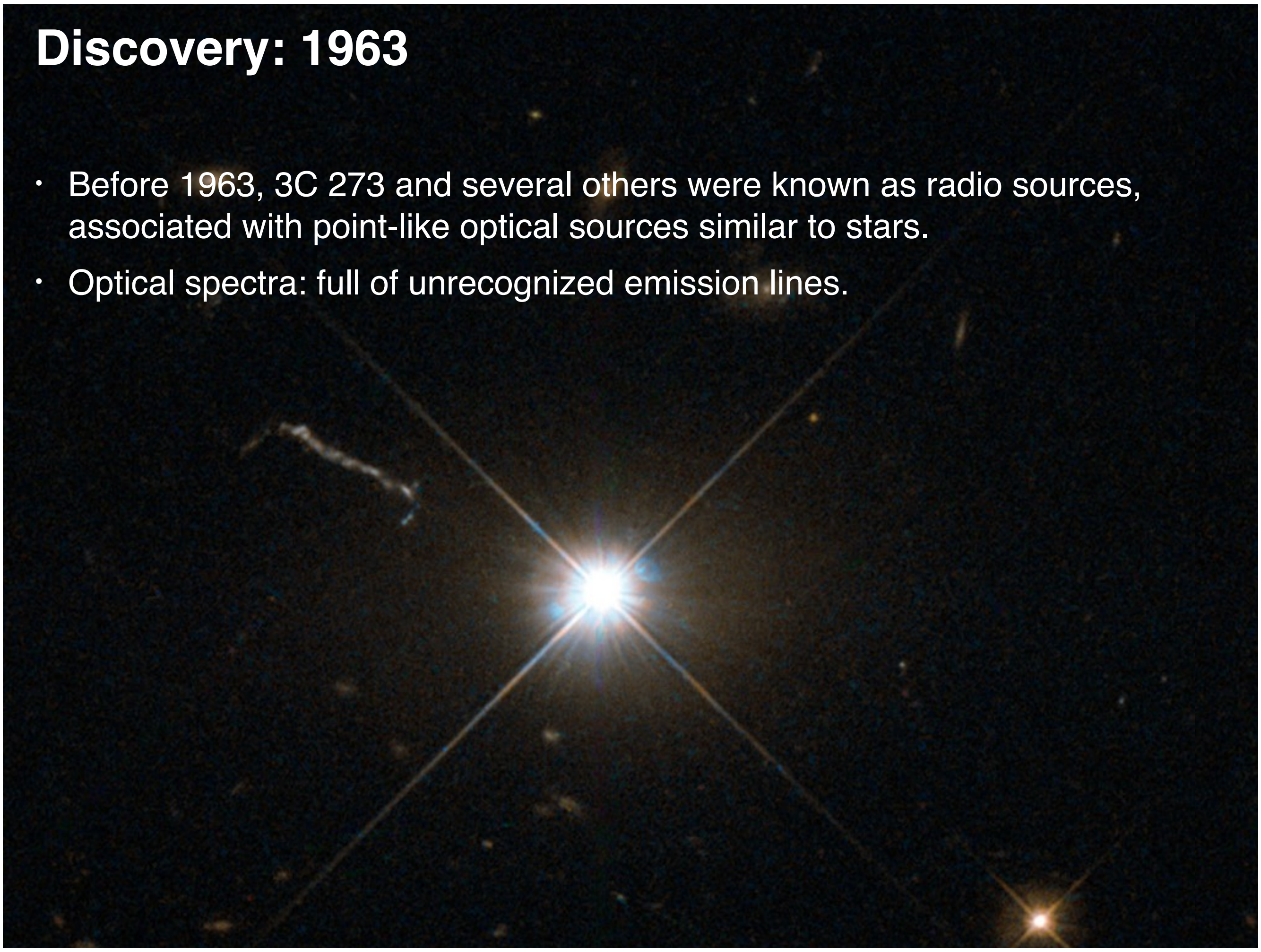
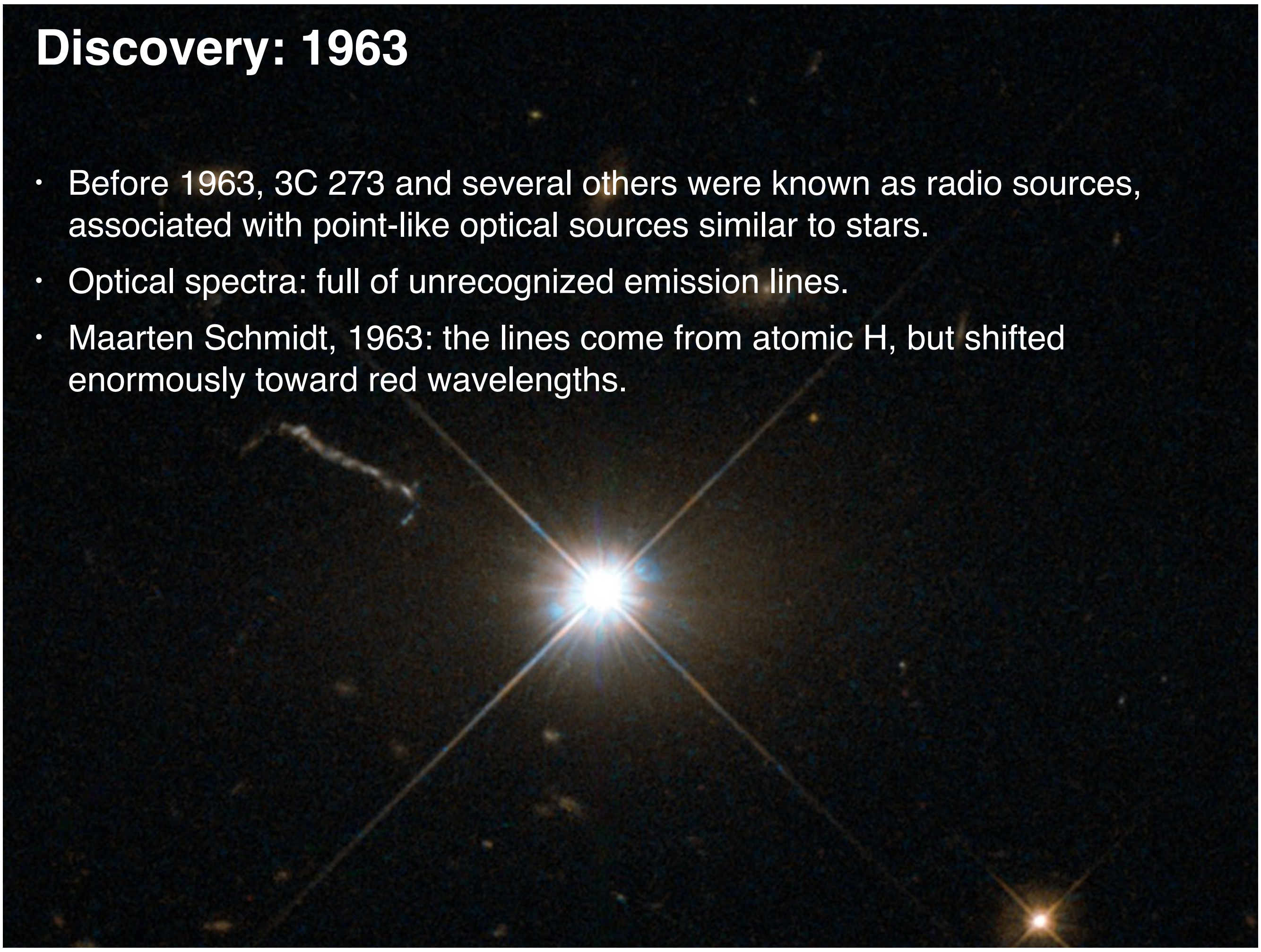
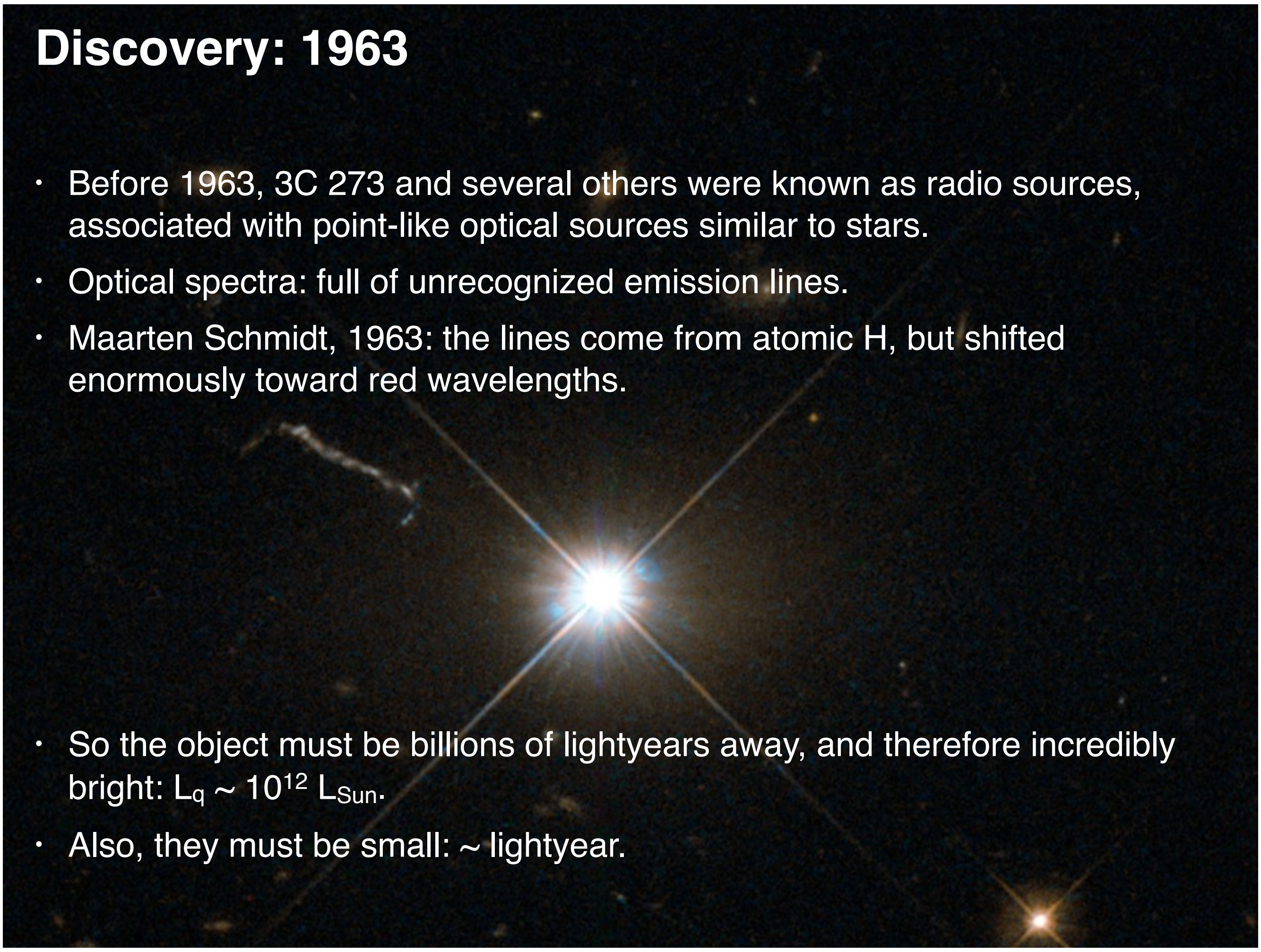

Quasars
Maria Süveges
Max Planck Institute for Astronomy, Heidelberg
Astro@Stats
8 September 2017, Padova


Discovery: 1963
• Before 1963, 3C 273 and several others were known as radio sources, associated with point-like optical sources similar to stars.
• Optical spectra: full of unrecognized emission lines.


Discovery: 1963
Before 1963, 3C 273 and several others were known as radio sources, associated with point-like optical sources similar to stars.
Optical spectra: full of unrecognized emission lines.
Maarten Schmidt, 1963: the lines come from atomic H, but shifted enormously toward red wavelengths.

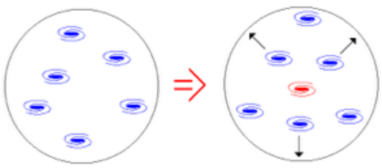

Discovery: 1963
Before 1963, 3C 273 and several others were known as radio sources, associated with point-like optical sources similar to stars.
Optical spectra: full of unrecognized emission lines.
Maarten Schmidt, 1963: the lines come from atomic H, but shifted enormously toward red wavelengths.
So the object must be billions of lightyears away, and therefore incredibly bright: $L_q \sim 10^{12} L_{Sun}$.
Also, they must be small: ~ lightyear.

Cosmology in 1963

Two main theories:

Steady-state

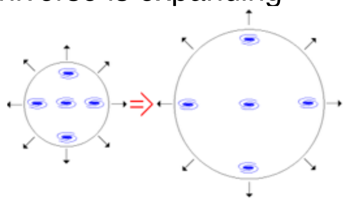
The flow is compensated by continuous local creation of



Satisfies the perfect cosmological principle:
the Universe is homogeneous and isotropic in space and time

Big Bang

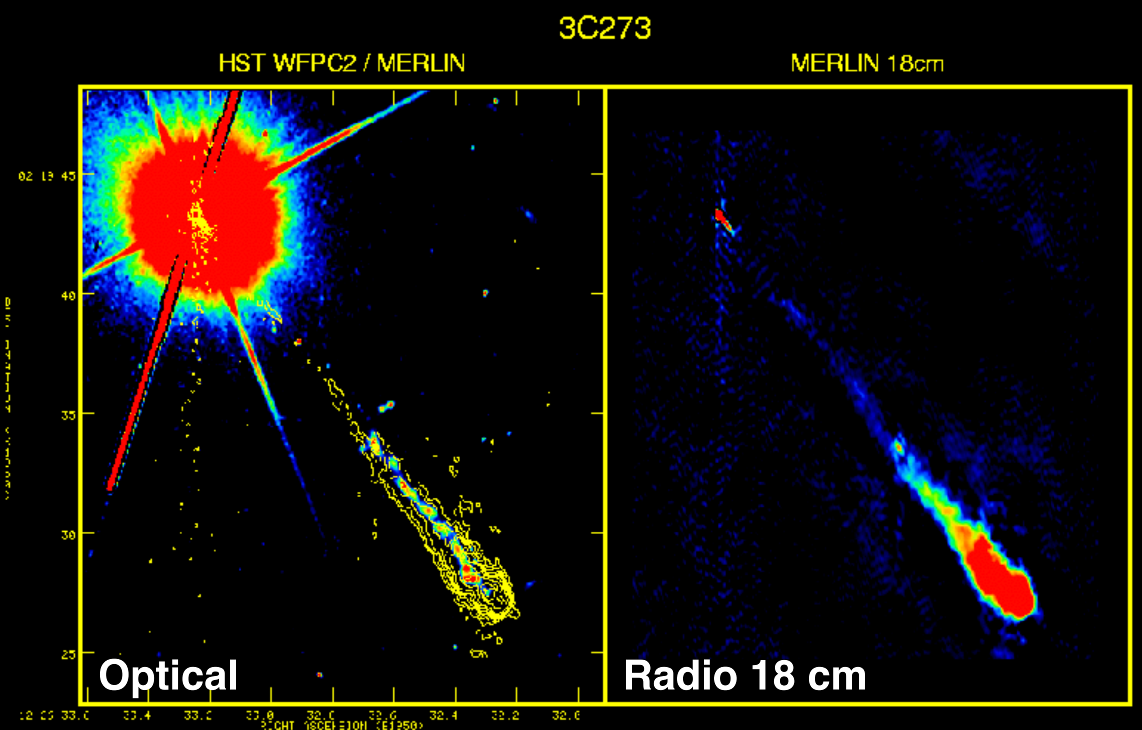
No creation of matter, the Universe is expanding



Satisfies the cosmological principle:
the Universe is homogeneous and isotropic in space

Quasars provide an argument against the perfect cosmological principle: they are rare now but were more frequent in the early Universe.

What do we know about the quasars? Often strong radio emitters

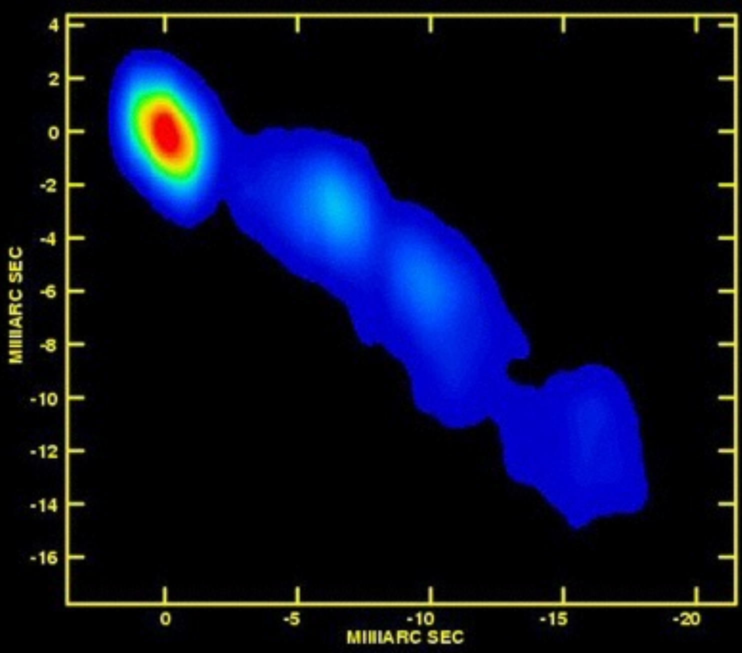


Radio image gallery of Jodrell Banks, A. Richards
http://www.jb.man.ac.uk/research/namgallery/

What do we know about the quasars? Strong X-ray emitters

Chandra X-ray Observatory

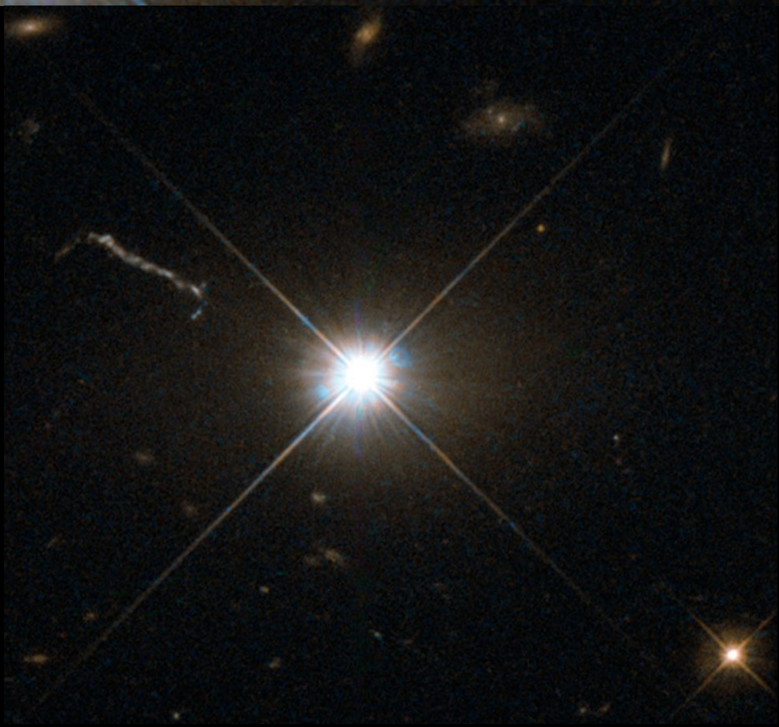
What do we know about the quasars? Interesting structures…



- Intermittent jet activity
- Knots, arcs, other structures
- Interaction with magnetic fields
- Interaction with inter-galactic matter
- Shock waves

Very Long Baseline Array
Image credit: NRAO

What do we know about the quasars? Interesting structures…



- Intermittent jet activity
- Knots, arcs, other structures
- Interaction with magnetic fields
- Interaction with inter-galactic matter
- Shock waves

Hubble Space Telescope
https://www.spacetelescope.org/images/potw1346a/

What do we know about the quasars? Variable objects

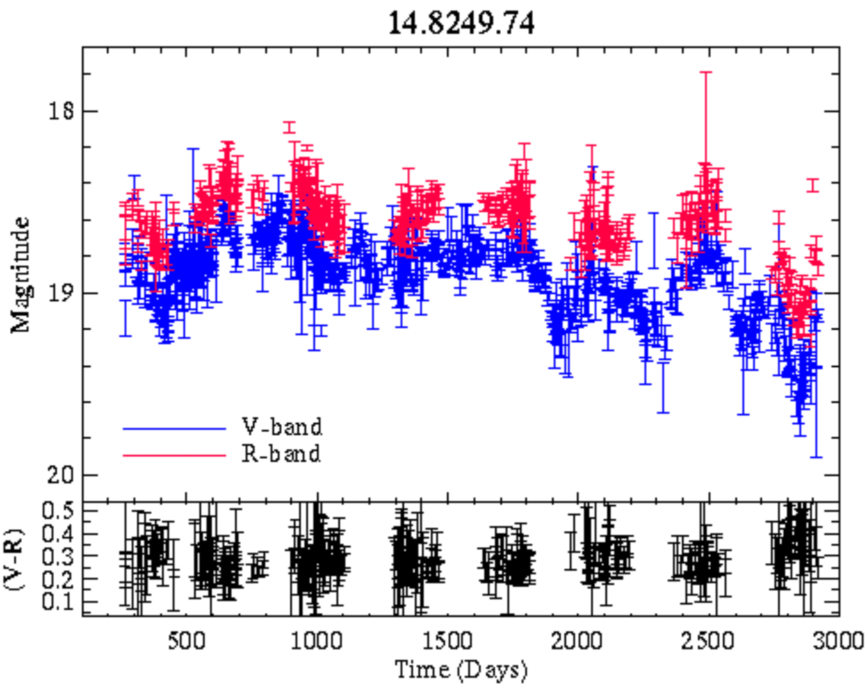
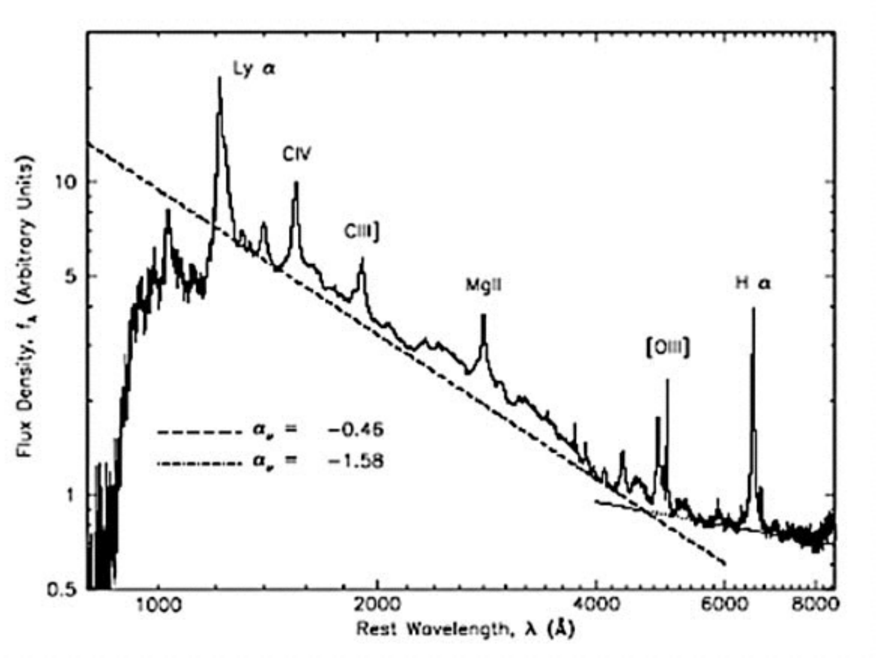


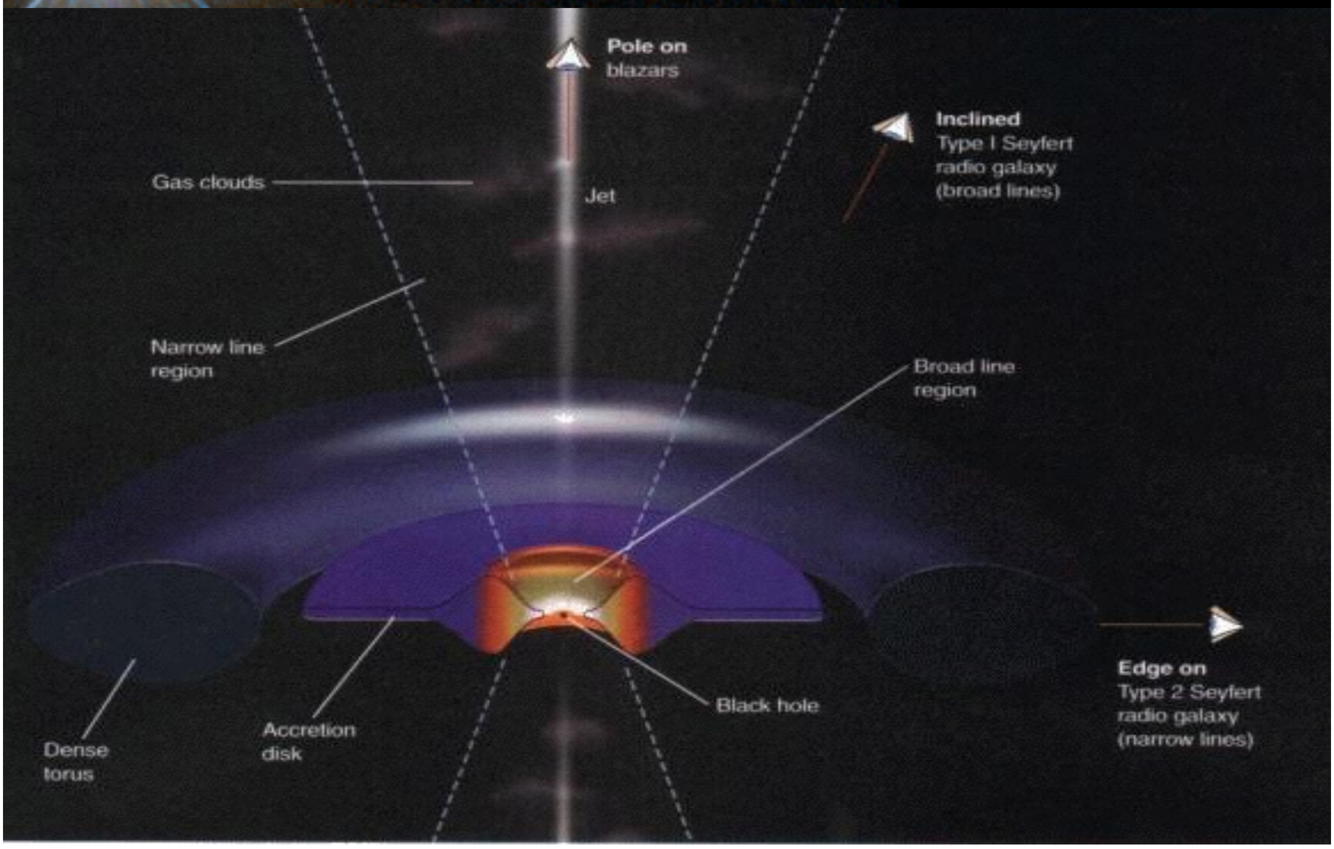
Image: http://www.astro.yale.edu/mgeha/MACHO/14.8249.74.html

What do we know about the quasars? Spectra

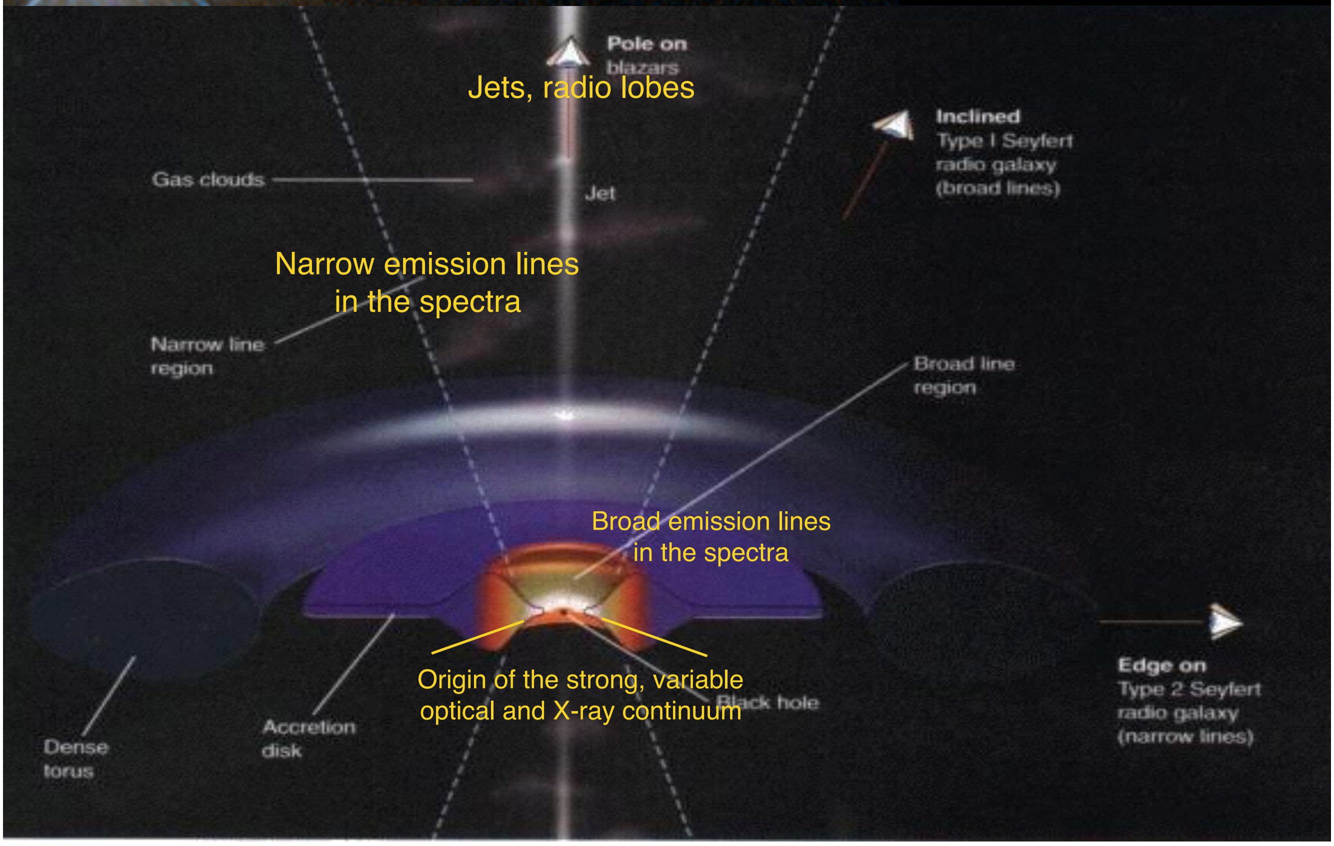


D. Vanden Berk & al., AJ, 122, 549-564, 2001

Structure



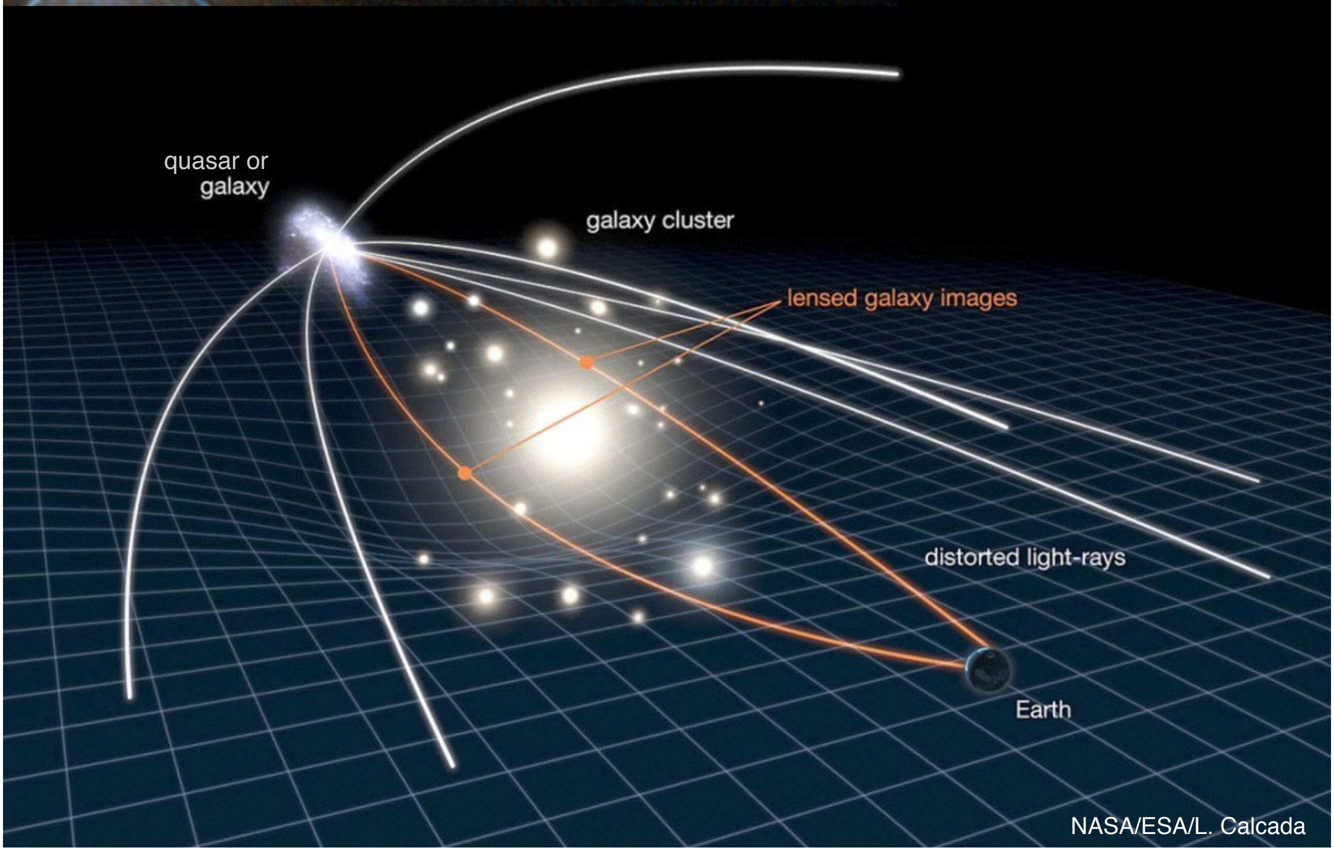
Structure



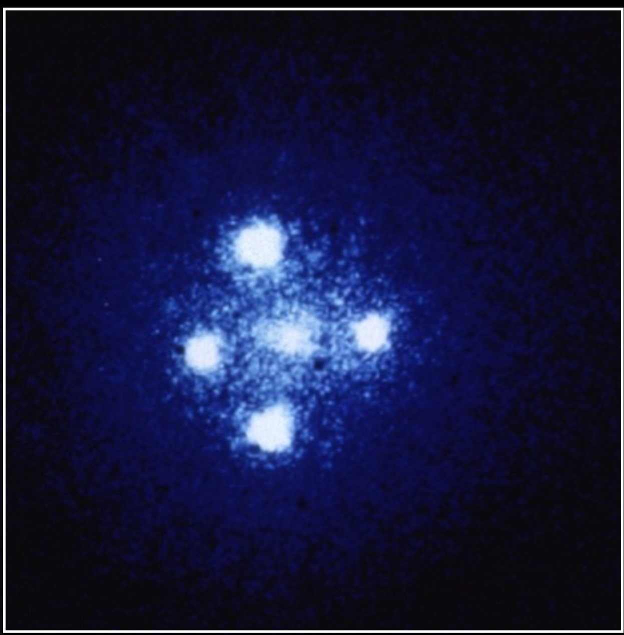
What do we know about the quasars? Likely a stage in galaxy evolution

NASA, A. Martel (JHU), H. Ford (JHU), M. Clampin (STScI), G. Hartig (STScI), G. Illingworth (UCO/Lick Observatory), the ACS Science Team and ESA; http://hubblesite.org/gallery/album/entire_collection/pr2003003a/

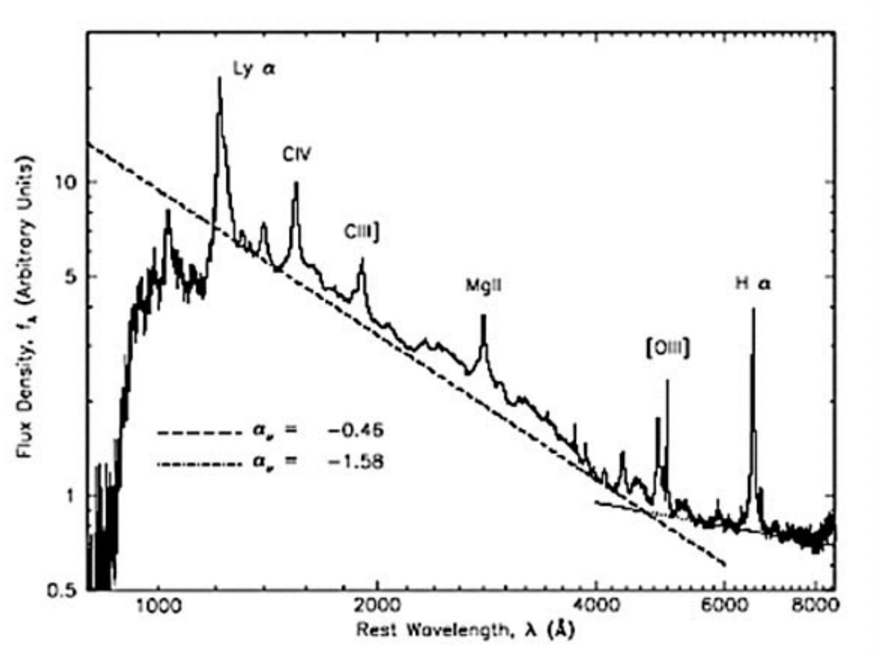
What do we know about the quasars? Gravitational lensing



What do we know about the quasars? Gravitational lensing

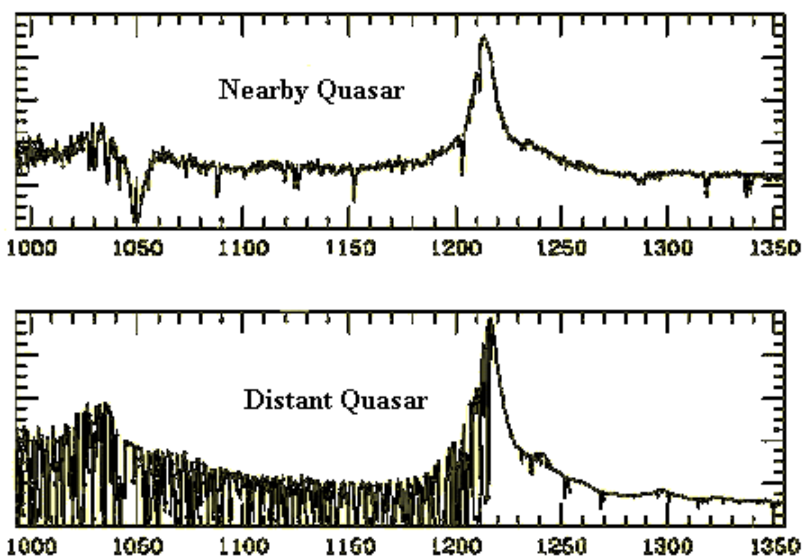


What do we know about the quasars?
The Lyman-α forest



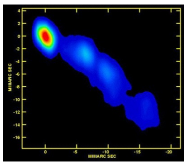
D. Vanden Berk & al., AJ, 122, 549-564, 2001

What do we know about the quasars?
The Lyman-α forest

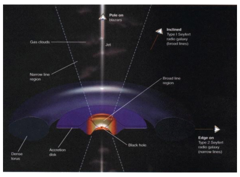


B. Keel, http://pages.astronomy.ua.edu/keel/agn/forest.html
Original data: HST Faint Object Spectrograph/Keck I HIRES

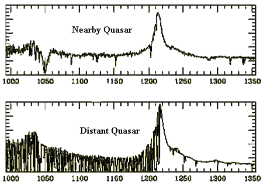
So they are interesting because of….



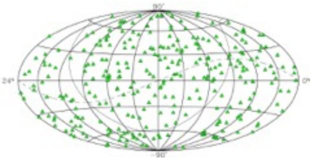
…their physics (physics under extreme conditions, accretion, tests of general relativity,…)



…their role in the history and evolution of the Universe (interaction between active galactic nuclei and galaxy, a stage in the early evolution of galaxies, …)



…cosmology (Lyman alpha forest, gravitational lensing)



...practical astronomy: a universal celestial “reference frame” and an “absolute” coordinate system

Detect them in survey data

Data:

- integrated photon flux in some wavelength bands (‘photometry’)
- spectra over some wavelength range
- time series of both above
- position and motion in the sky
- parallax
- morphology

Derived quantities:

- colours
- time series parameters
- line intensities in the spectrum
- distance

Detect them in Gaia data

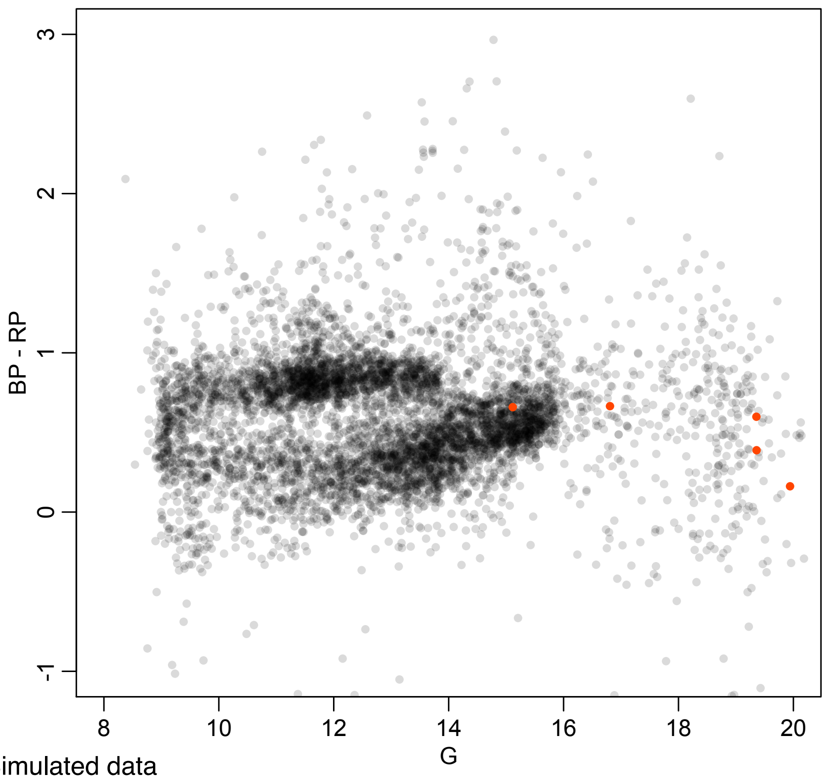
Data:

- integrated photon flux in some wavelength bands (‘photometry’)
- spectra over some wavelength range
- time series of both above
- position and motion in the sky
- parallax
- morphology

Derived quantities:

- colours
- time series parameters
- line intensities in the spectrum
- distance

Fraction of quasars

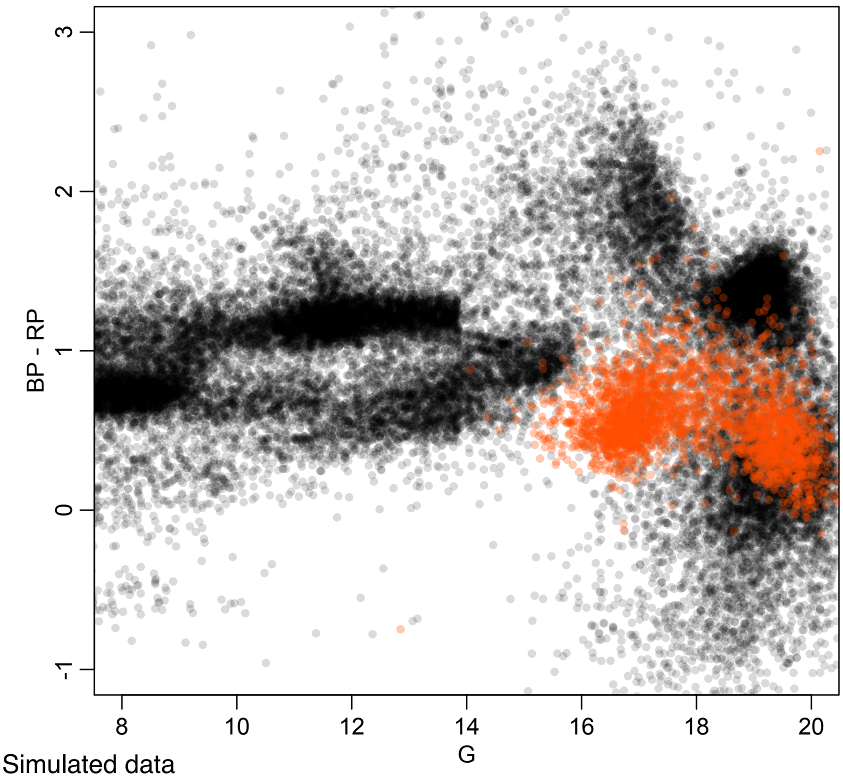


Simulated data

Expected proportions:
5 objects out of 10000

Plot: near-real fraction:
5 quasars, 6300 other

Fraction of quasars



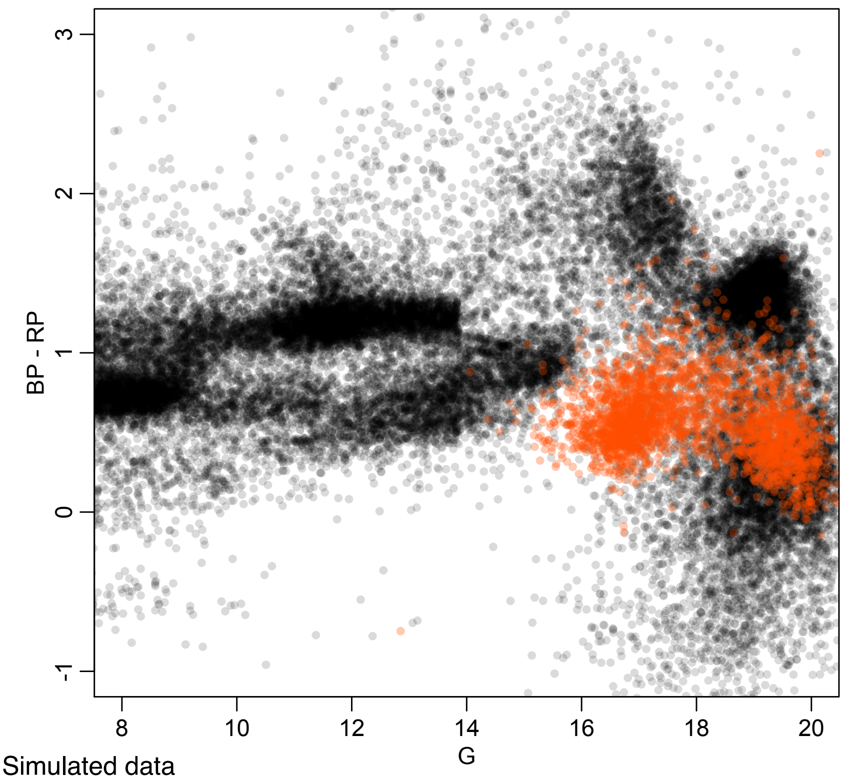
Simulated data

Expected proportions:
5 objects out of 10000

Plot: false proportions:

2600 quasars,
18600 others

What to train on? Real data?



Simulated data

Expected proportions:
5 objects out of 10000

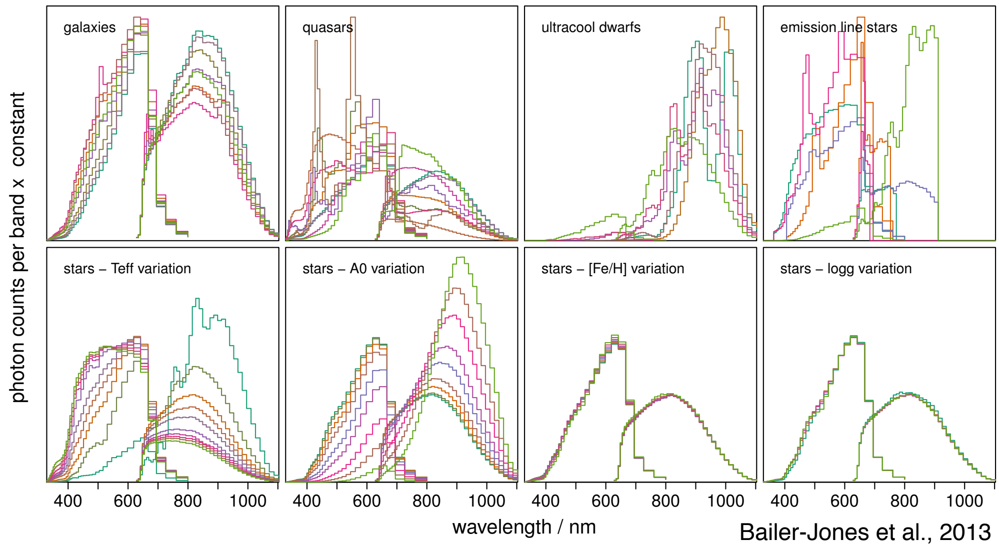
Plot: false proportions:

2600 quasars,
18600 others

In addition:
biases in the training set

- due to position in the Galaxy and intrinsic luminosity
- due to selection biases (scientific interests, observability, funding,…)

What to train on? Simulated spectra?

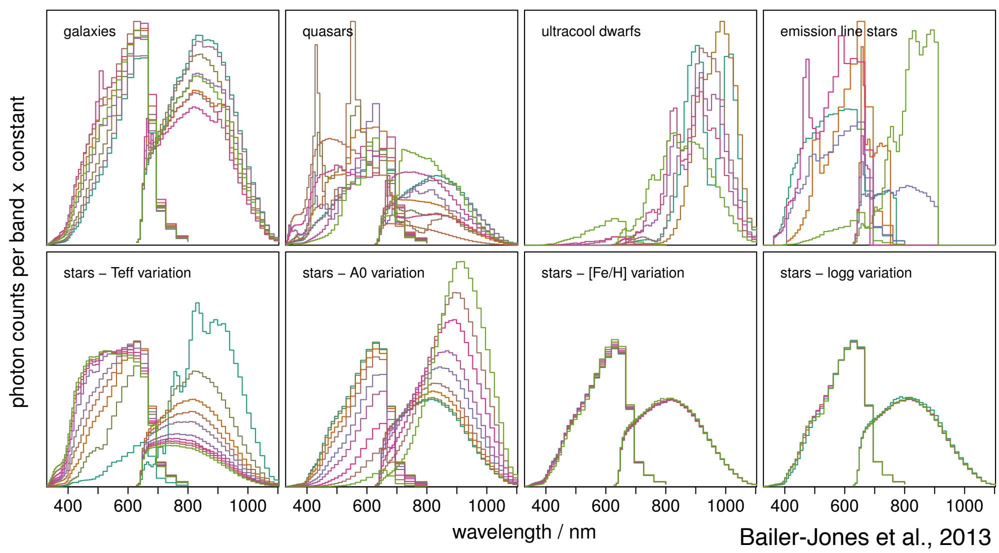


Bailer-Jones et al., 2013

Simulated training set:

- we can compute it on a grid
- we can add nominal noise

What to train on? Simulated spectra?



Bailer-Jones et al., 2013

Simulated training set:

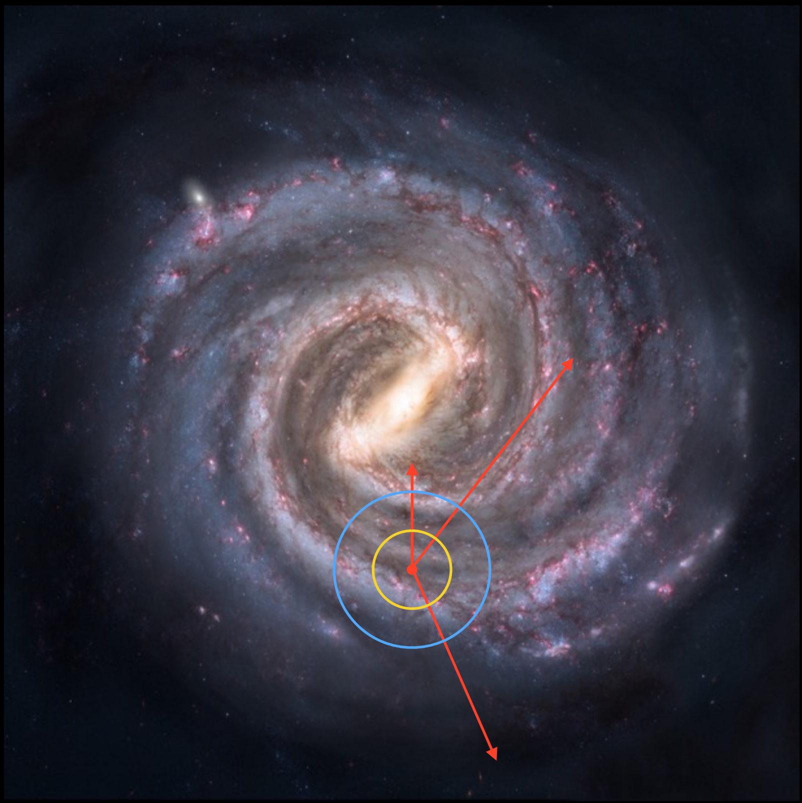
- we can compute it on a grid (but it will not follow the real distribution)
- we can add nominal noise (but it will not reproduce real artefacts)

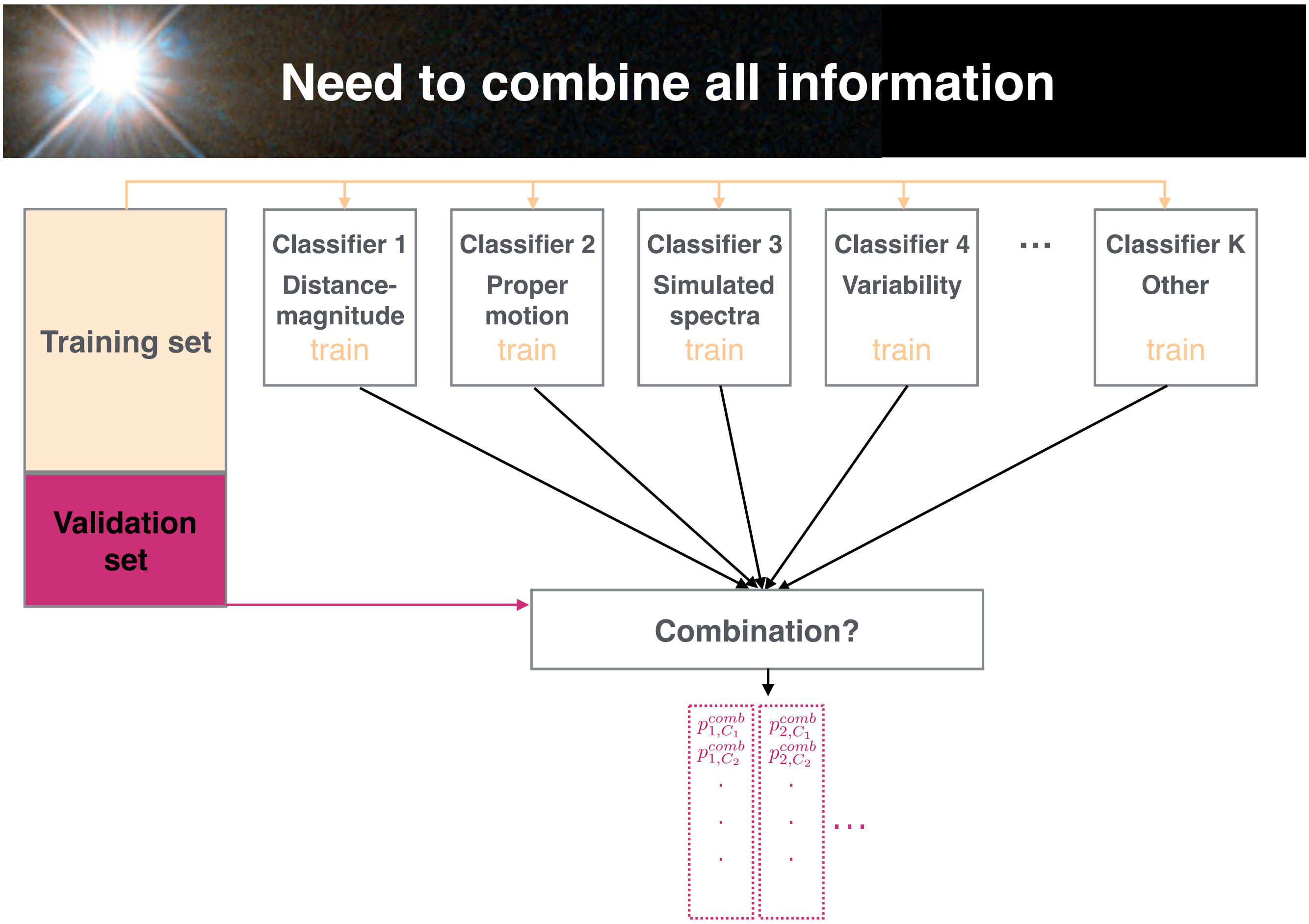
What to train on?
Use Galactic distributions?

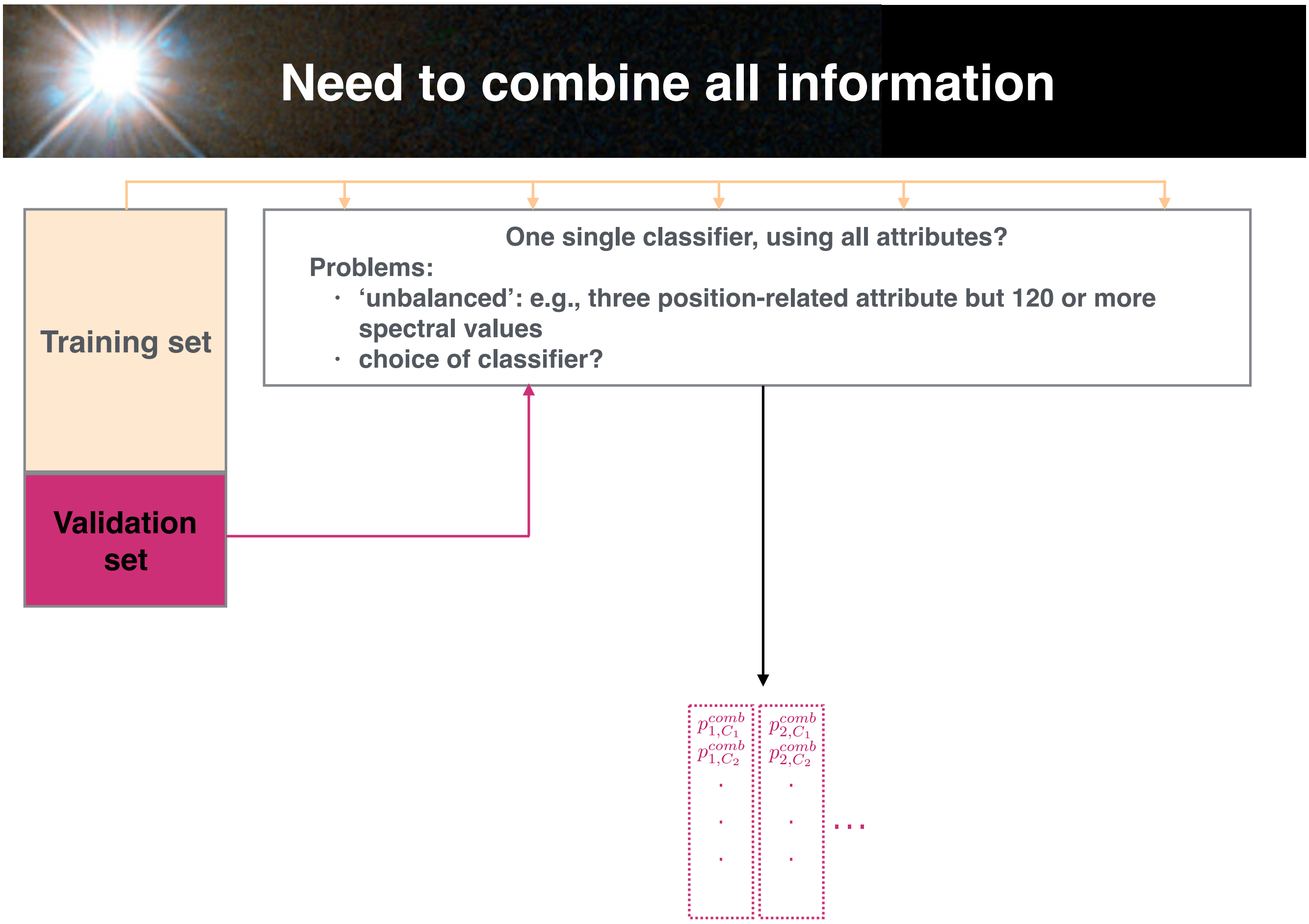
What to train on?
Use Galactic distributions?

What to train on?
Use Galactic distributions?

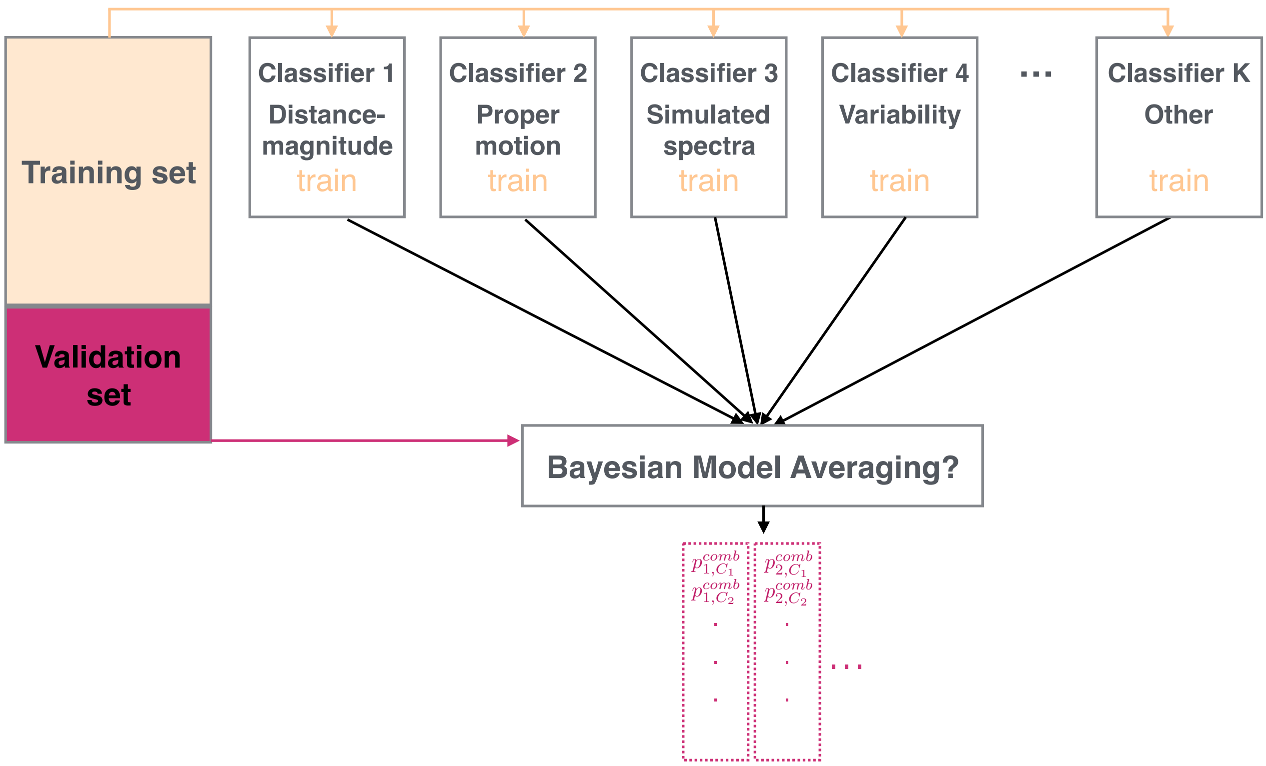
What to train on?
Use Galactic distributions?




Need to combine all information
Training set
Validation set
Classifier 1
Distance-magnitude
train
Classifier 2
Proper motion
train
Classifier 3
Simulated spectra
train
Classifier 4
Variability
train
...
Classifier K
Other
train
Combination?
p^{comb}_{1,C_1}
p^{comb}_{1,C_2}
p^{comb}_{2,C_1}
p^{comb}_{2,C_2}
...


Need to combine all information
Training set
Validation set
One single classifier, using all attributes?
Problems:
· 'unbalanced': e.g., three position-related attribute but 120 or more spectral values
· choice of classifier?
p^{comb}_{1,C_1}
p^{comb}_{1,C_2}
p^{comb}_{2,C_1}
p^{comb}_{2,C_2}
...

Need to combine all information



Bayesian Model Averaging

ξ parameter of interest (class in our case)

M_k classifier k

D data

posterior:

$$P(\xi \mid D) = \sum_{k=1}^{K} P(\xi \mid M_k, D) P(M_k \mid D)$$

probability of model k given D:

$$P(M_k \mid D) = \frac{P(D \mid M_k) P(M_k)}{\sum_{i=1}^{K} P(D \mid M_i) P(M_i)}$$

likelihood of D under model k:

$$P(D \mid M_k) = \int p(D|\theta_k, M_k) p(\theta_k | M_k) d\theta_k$$

Bayesian Model Averaging

ξ parameter of interest (class in our case)

M_k classifier k

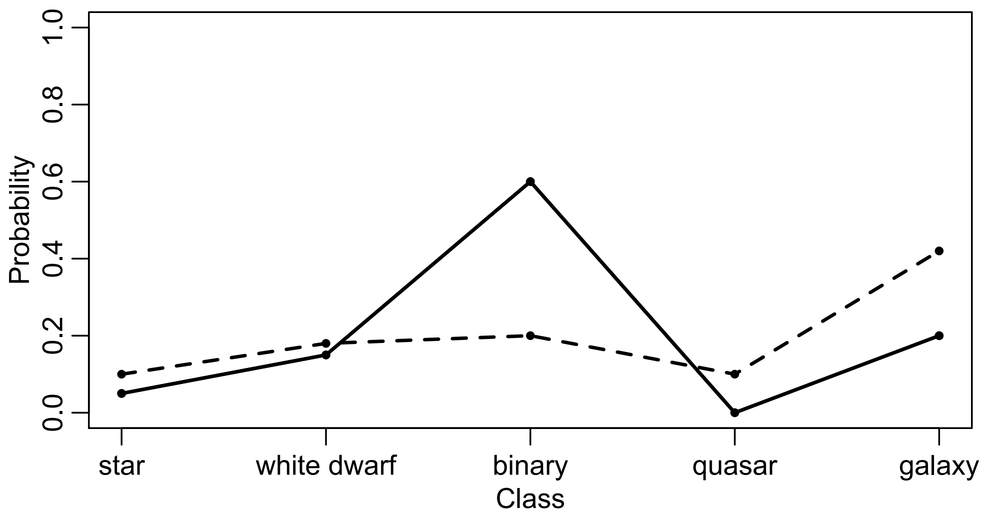
D data

posterior: $$P(\xi \mid D) = \sum_{k=1}^{K} P(\xi \mid M_k, D) P(M_k \mid D)$$

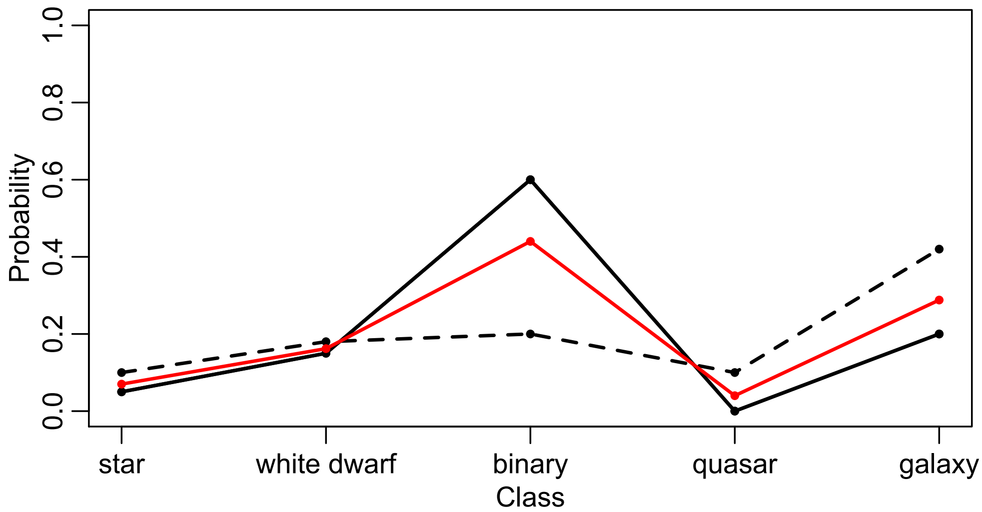
probability of model k given D: $$P(M_k \mid D) = \frac{P(D \mid M_k) P(M_k)}{\sum_{i=1}^{K} P(D \mid M_i) P(M_i)}$$

likelihood of D under model k: $$P(D \mid M_k) = \int p(D|\theta_k, M_k) p(\theta_k|M_k) d\theta_k$$

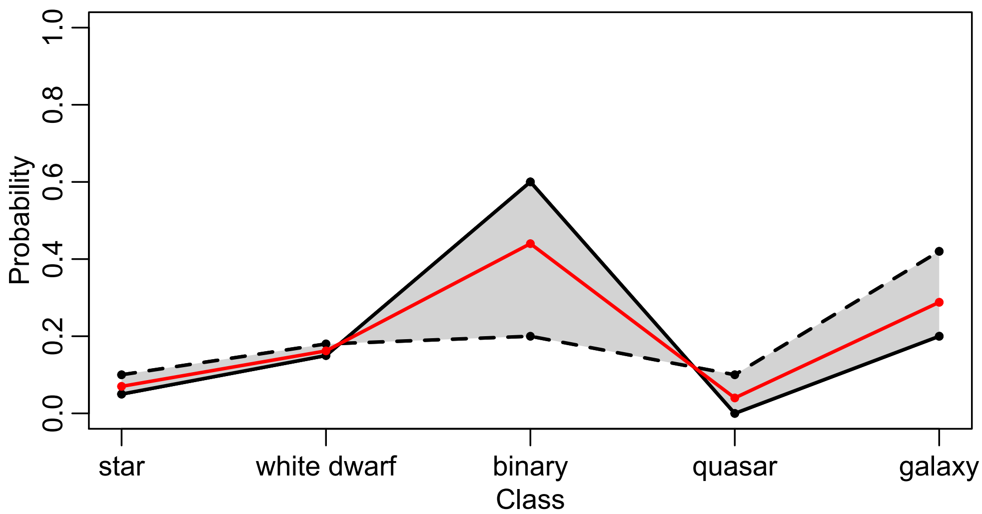
Bayesian Model Averaging



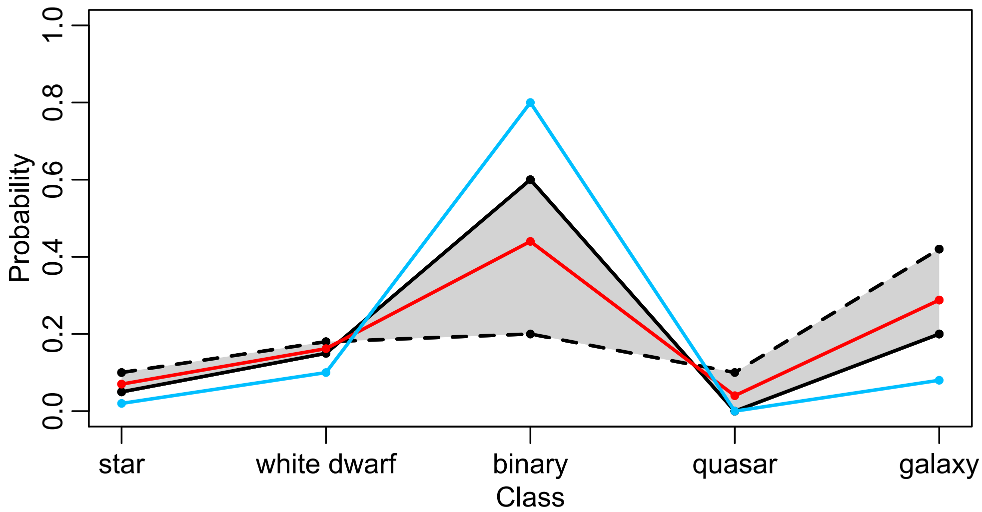
Bayesian Model Averaging



Bayesian Model Averaging



Bayesian Model Averaging



Hierarchical combination

Idea:

The methods (classifiers) work as mappings of the full information contained in the data into the space $\mathcal{P}$ of probability distributions over the classes.

Let classifier i be represented by the mapping $f_i : \mathbb{R}^D \longrightarrow \mathcal{P}$,

$$\boldsymbol{p}_i = f_i(x_1, \ldots, x_D; \theta_i)$$

(possibly with some classifier-specific tuning parameter vector θ_i).

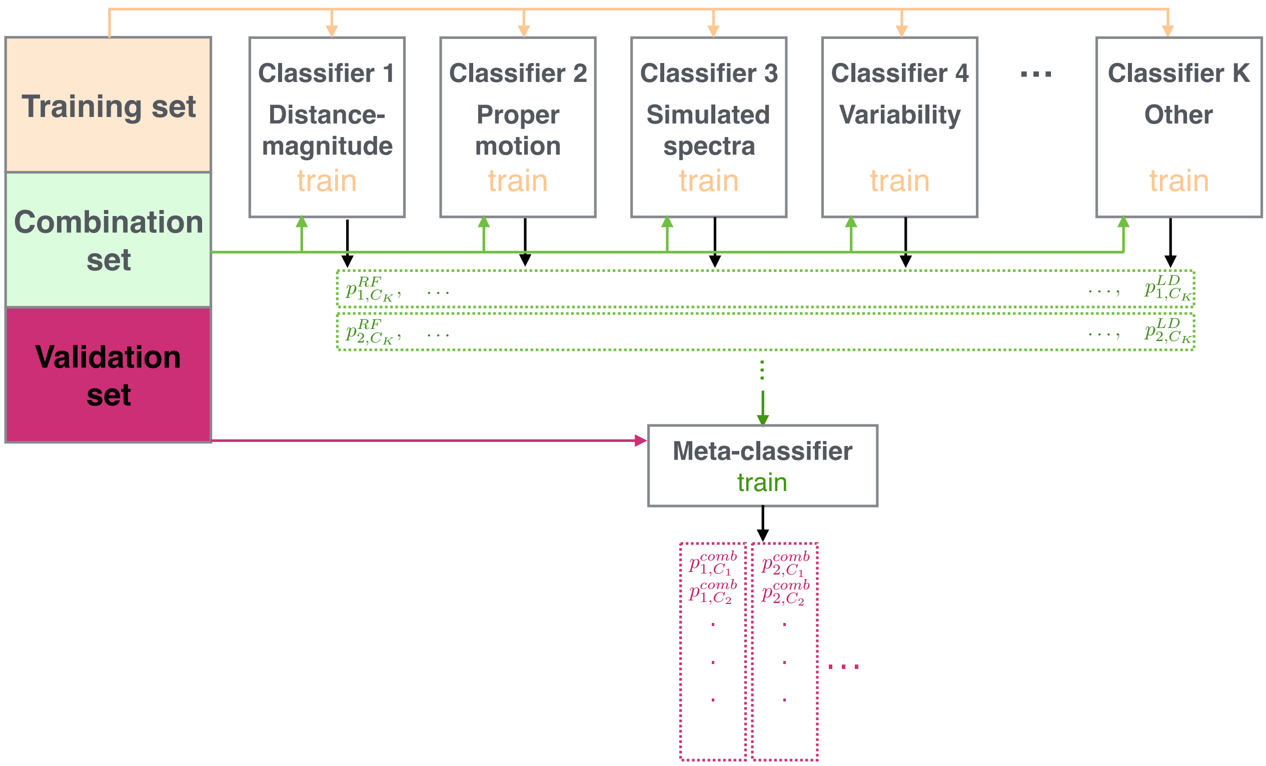
The $\boldsymbol{p}_i$ are in general not independent.

Then a second-level classifier can be defined as

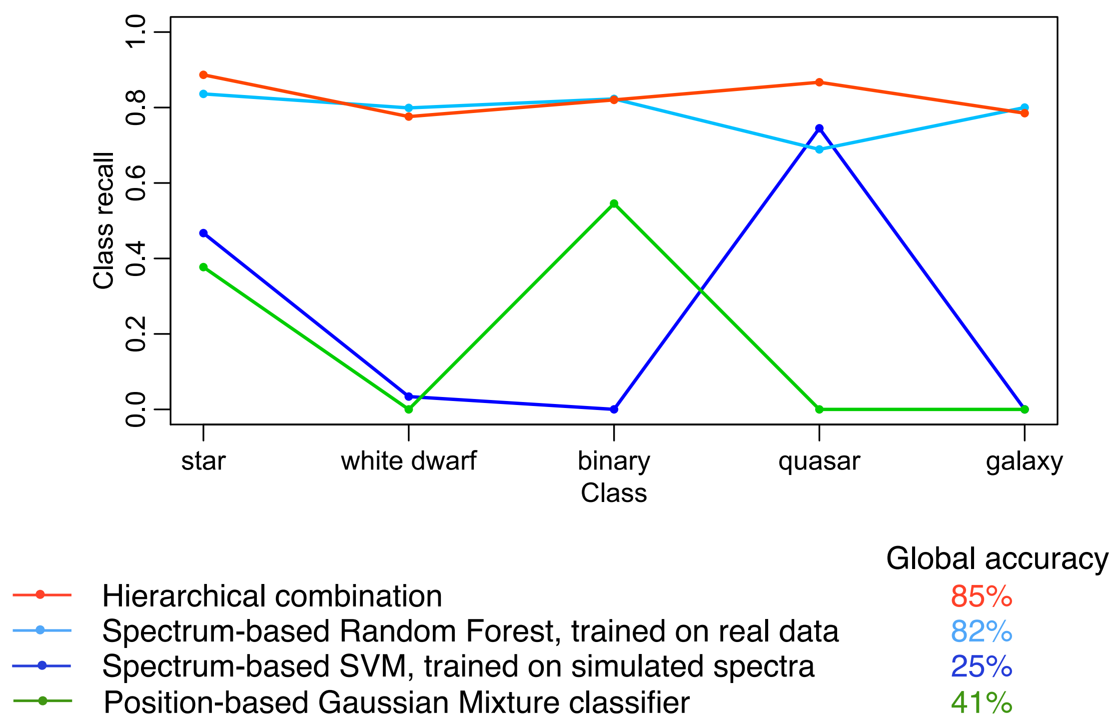
$$\begin{aligned} g : \mathcal{P}^K &\longrightarrow \mathcal{P} \\ \{\boldsymbol{p}_1, \ldots, \boldsymbol{p}_K\} &\mapsto \mathcal{P} \end{aligned}$$

Similar to stacking generalization (Wolpert 1992), which uses the point estimates, not the probability distributions.

Hierarchical combination



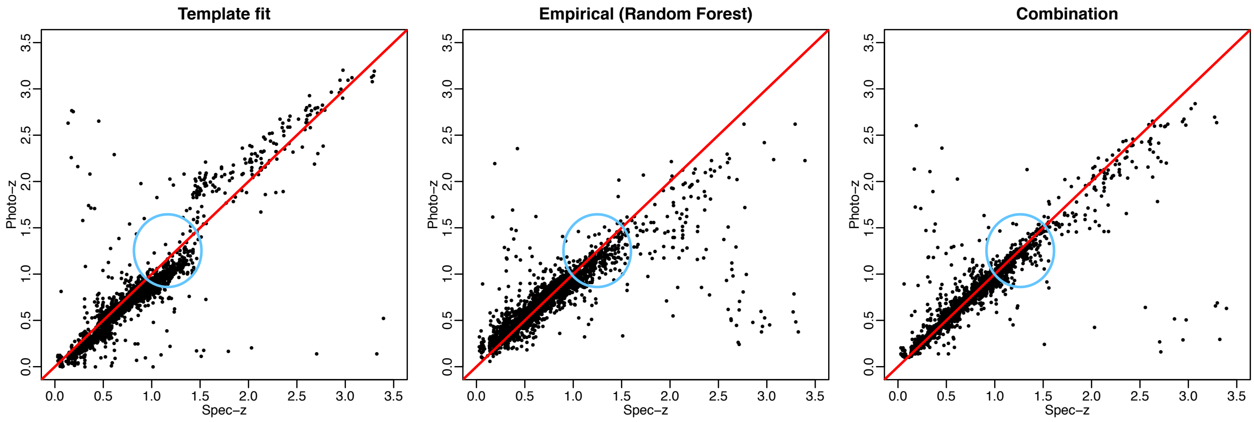
Hierarchical combination: results



Another application: photometric redshifts

Photometric redshift estimation: based on a few measurements of the brightness of a galaxy, estimate its redshift (that is, its distance).

Two basic kinds of methods: template fitting (based on theoretically prescribed spectra) and empirical (using observed real galaxies).



Summary

Whatever interesting objects we wish to pick out from survey data:

- Often they are rare.
- Often many sources of information: spectra, photometry, location & motion, morphology, time series behaviour observed in different wavelengths, etc.
- Often, applicable methods are of varying quality: many high-variance or biased, a few good…
- …and that, varying over the covariate space / object types / …

Combination seems not just a good idea, but necessary.

Hierarchical combination:

- **improves** on single-method analysis;
- is capable of **bias correction** (in case the training set contains relevant information);
- is **general** (applicable for data analysis where there are many optional methods, each with its different excellences and failures);