

Astr 511: Galaxies as galaxies

Winter Quarter 2017, University of Washington

Mario Jurić & Željko Ivezić

Lecture 4:

Luminosity and mass functions

Outline

- Galaxies as seen by SDSS
- Luminosity function: basic concepts
- Methods for estimating LF from data
- Stellar mass function in the Milky Way

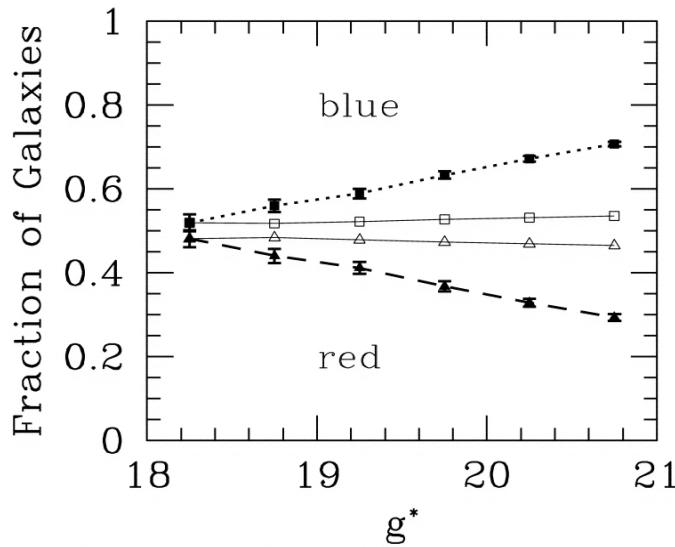
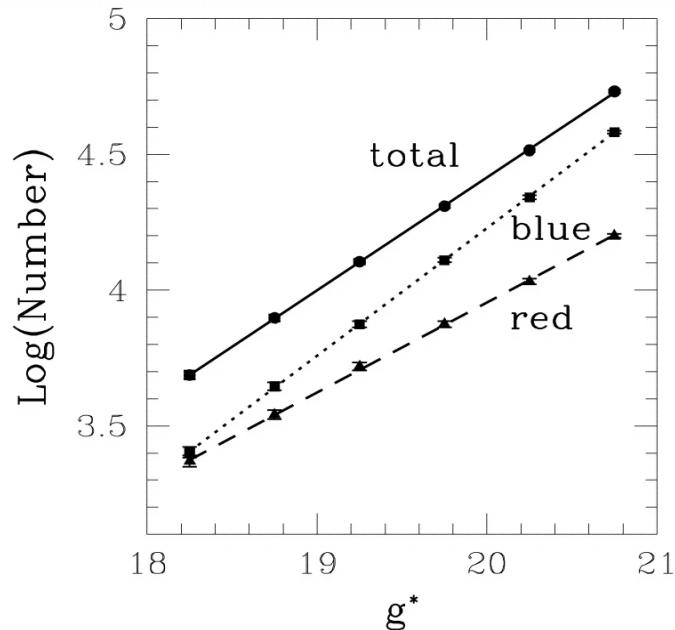
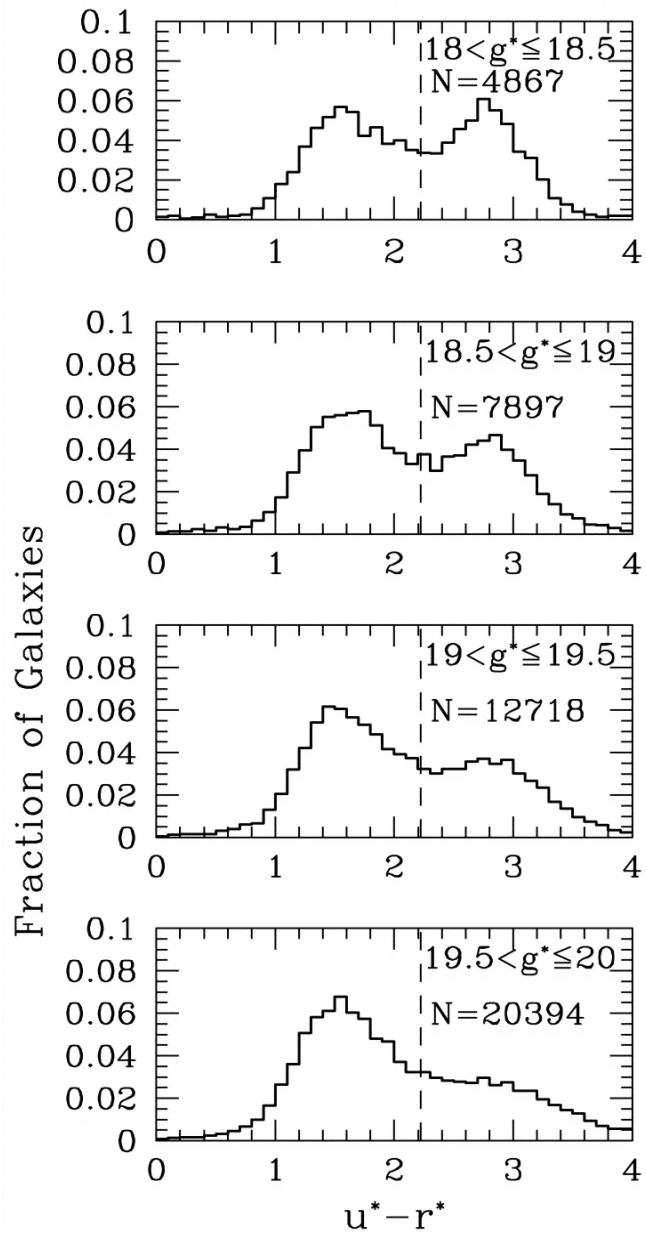
Galaxies in SDSS-I

- Imaging Survey
 - 10,000 deg² (1/4 of the full sky)
 - 5 bands (ugriz: UV-IR), 0.02 mag photometric accuracy
 - < 0.1 arcsec astrometric accuracy
 - 100,000,000 stars and 100,000,000 galaxies

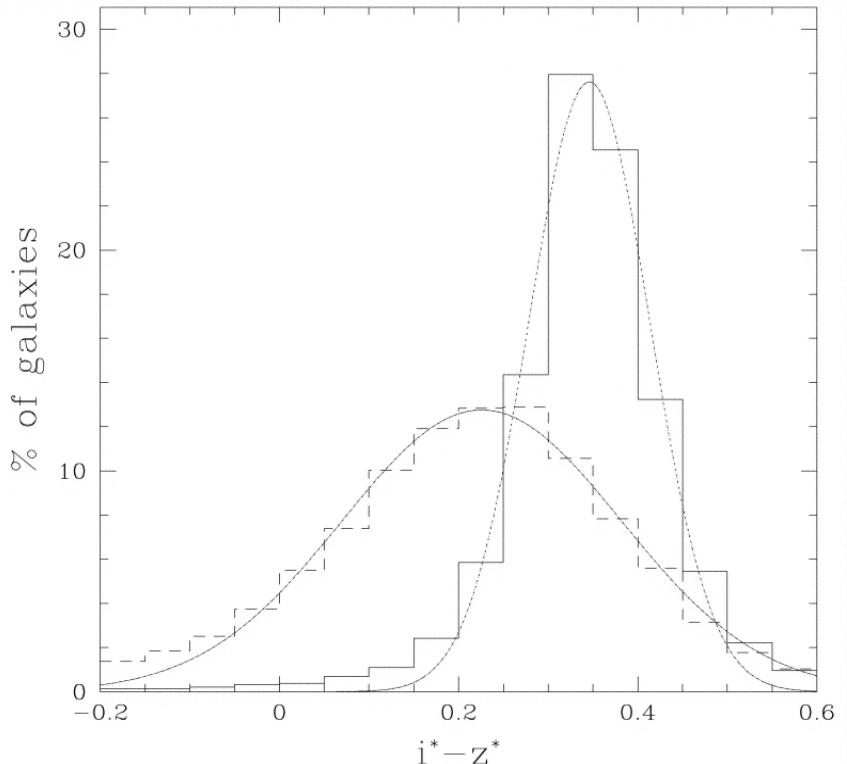
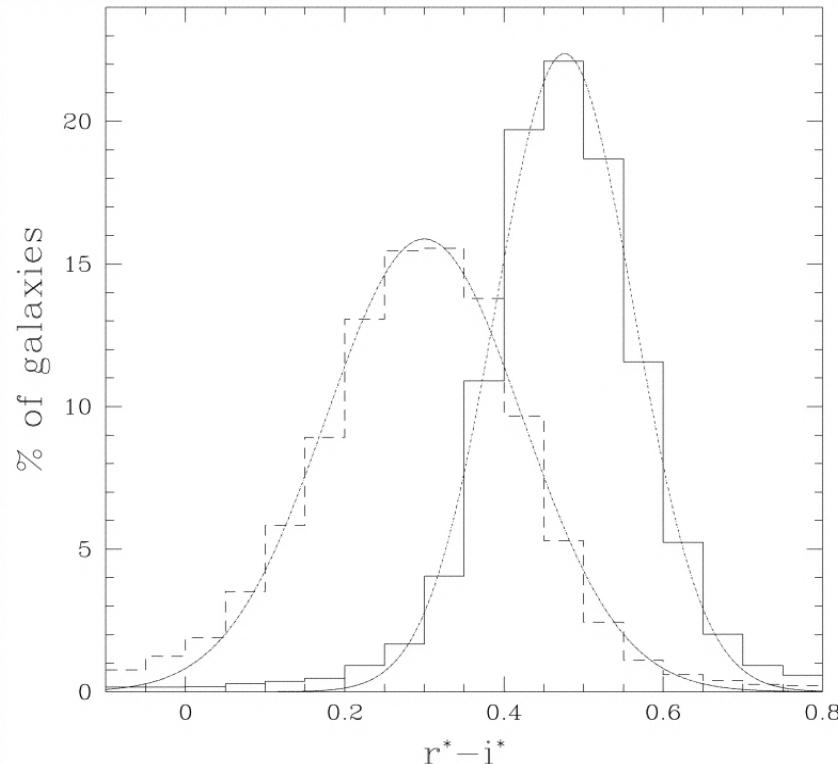
- Spectroscopic Survey
 - 1,000,000 galaxies
 - 100,000 quasars
 - 100,000 stars

SDSS Spectroscopic Galaxy Survey

- Two samples:
 1. the “main” galaxy sample ($r_{Pet} < 17.77$): ~ 1 million spectra
 2. luminous red galaxy sample (LRG, cut in color-magnitude space): $\sim 100,000$ spectra
- Distance estimate allows the determination of luminosity function (Blanton et al. 2001)
- Spectra are correlated with morphology (and colors)



Strateva et al. (2001, AJ 122, 1861): **bimodal $u - r$ color distribution** ($u - r$ is similar to $U - V$)



The broad band SEDs of galaxies are nearly one-dimensional family.

“Everything is correlated with everything” (Blanton et al. 2003)

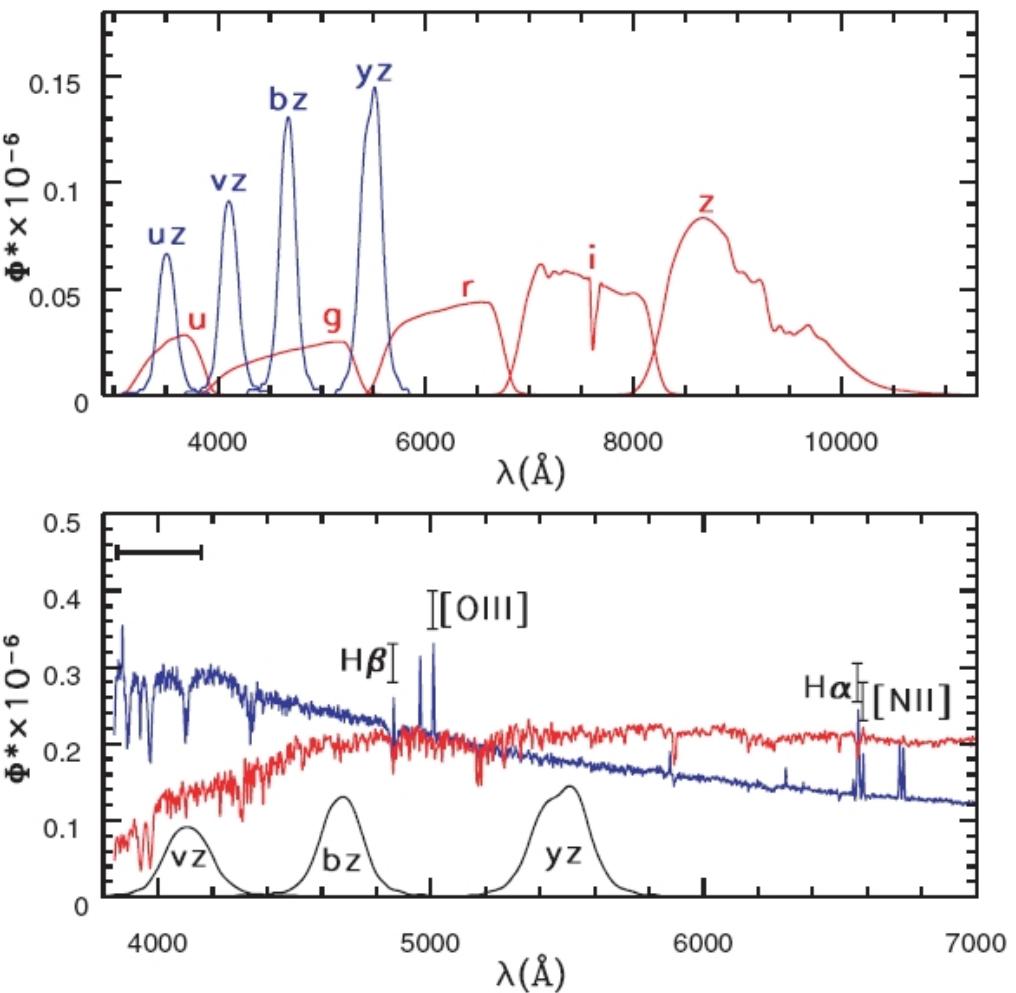


Figure 1. The top panel shows renormalized (see equation 3) filter transmission curves for the SDSS photometric system ($ugriz$), and for the Strömgren photometric system. The bottom panel emphasizes the 3800–7000 Å region. The two spectra are typical for blue and red galaxies, and the four labelled emission lines are used to separate star-forming from AGN galaxies. The horizontal bar in the top left-hand corner marks the wavelength region used in the analysis by Kauffmann et al. (2003).

- Star formation vs AGN
- It is not easy to distinguish galaxies with active star formation from those that harbour an AGN (active galactic nuclei; a.k.a. black hole with an accretion disk)
- We can use the emission line strength to separate them: H_α , H_β , $[NII]$, $[OIII]$
- Physical origin: AGN have power-law spectra, so they have more UV photons than even the hottest stars; as a result, for a given $[OIII]/H_\beta$ ratio, AGNs have larger $[NII]/H_\alpha$ ratio than star-forming galaxies

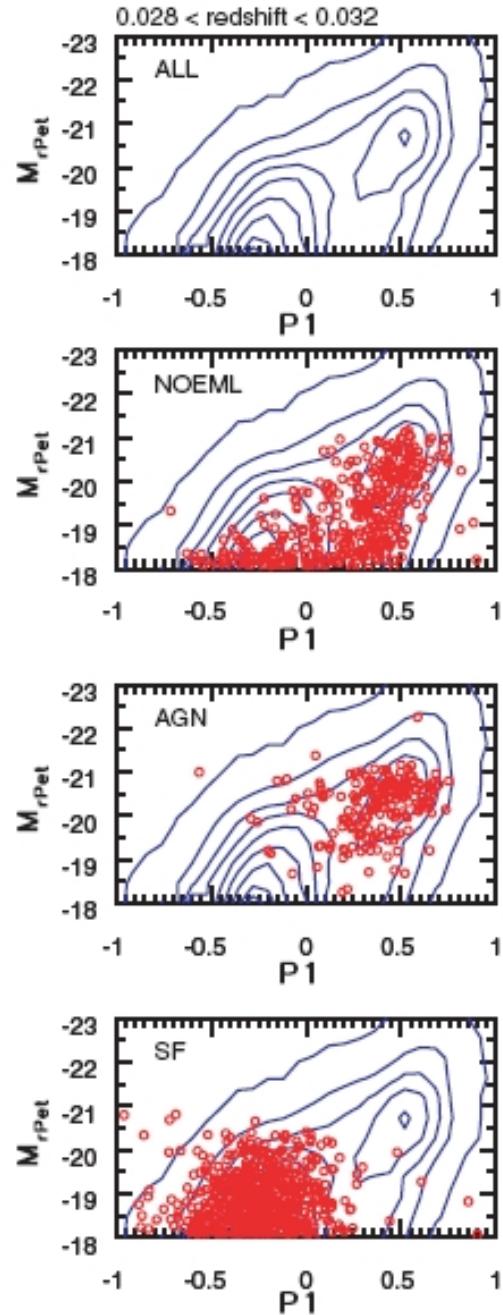
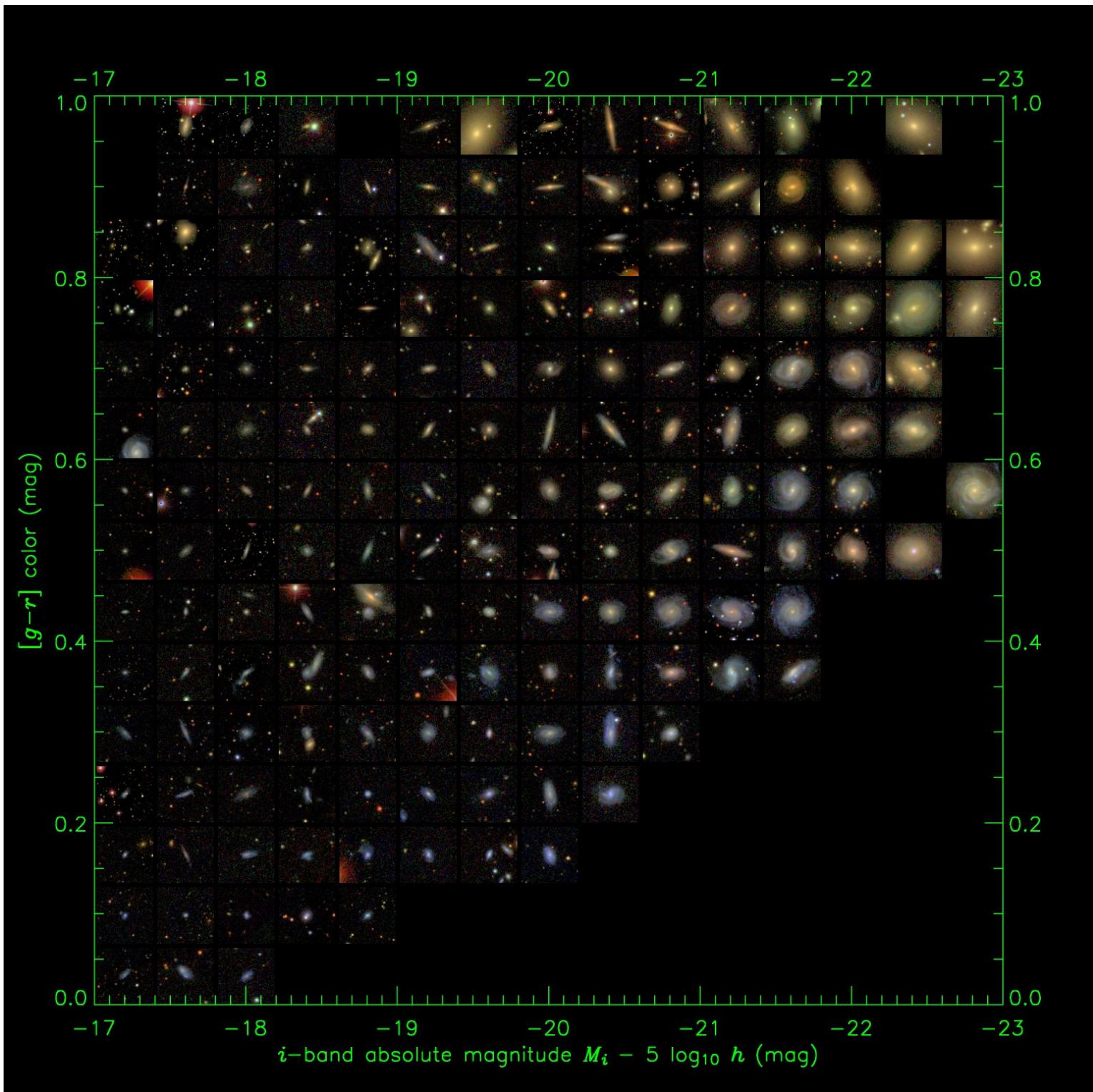
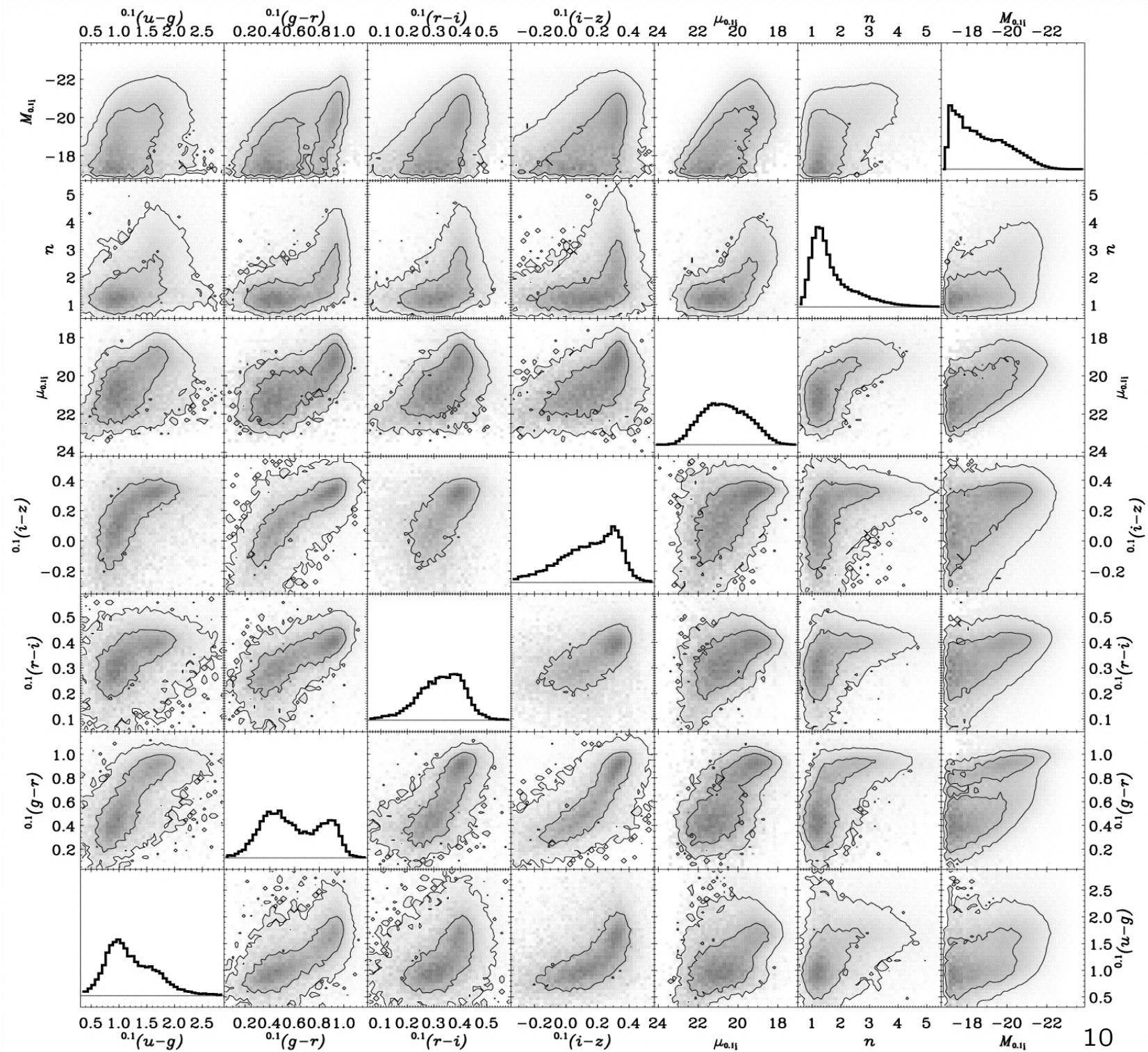


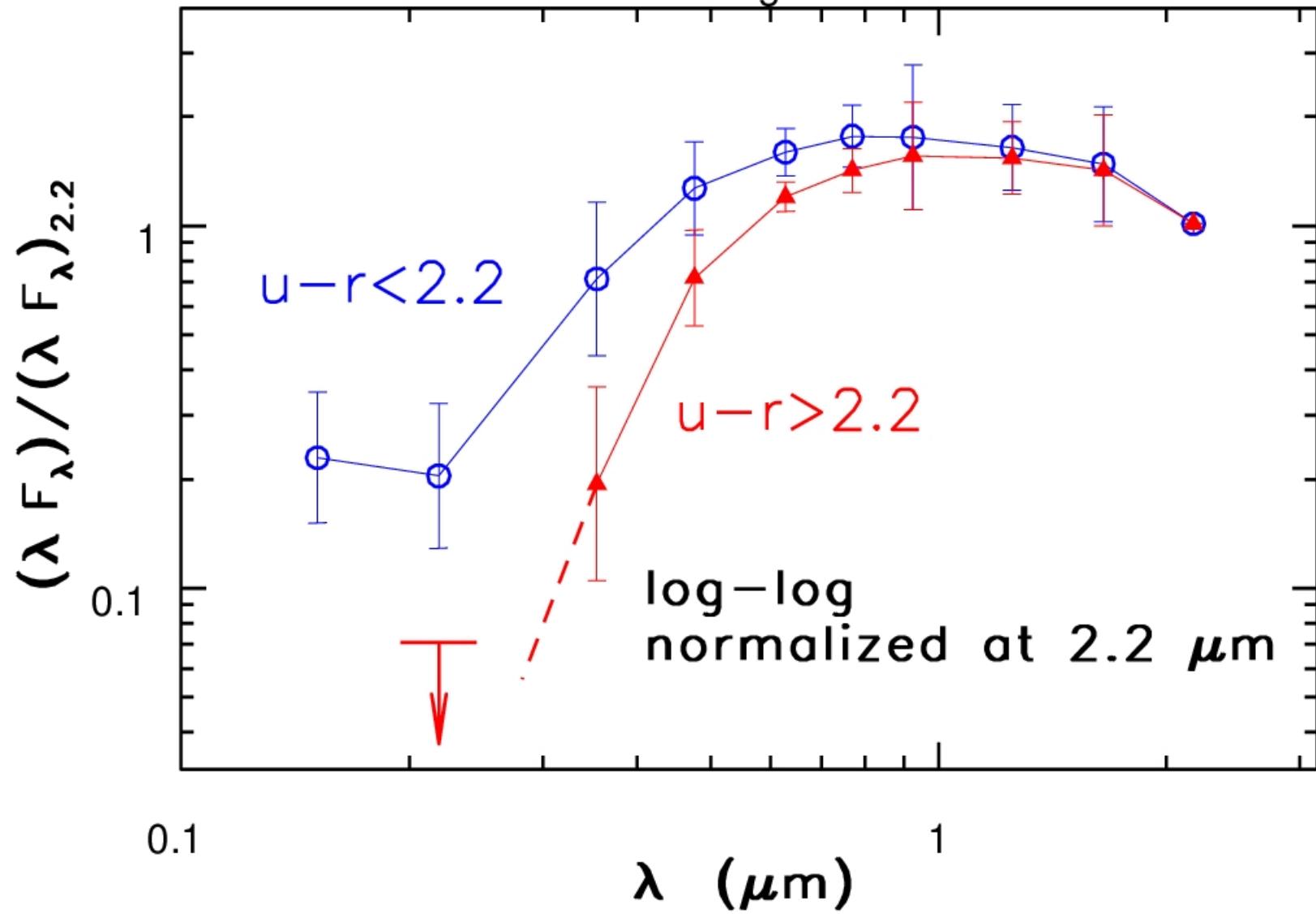
Figure 21. The colour–absolute magnitude diagram for SDSS main galaxies. The top panel shows the absolute r -band magnitude as a function of the $P1$ rest-frame Strömgren colour for galaxies with $0.028 < z < 0.032$. The remaining panels compare the distribution of all galaxies (contours) to the distributions of three subsamples (symbols) selected using emission lines: galaxies without emission lines in the upper middle panel, AGN galaxies in the lower middle panel, and star-forming galaxies in the bottom panel.

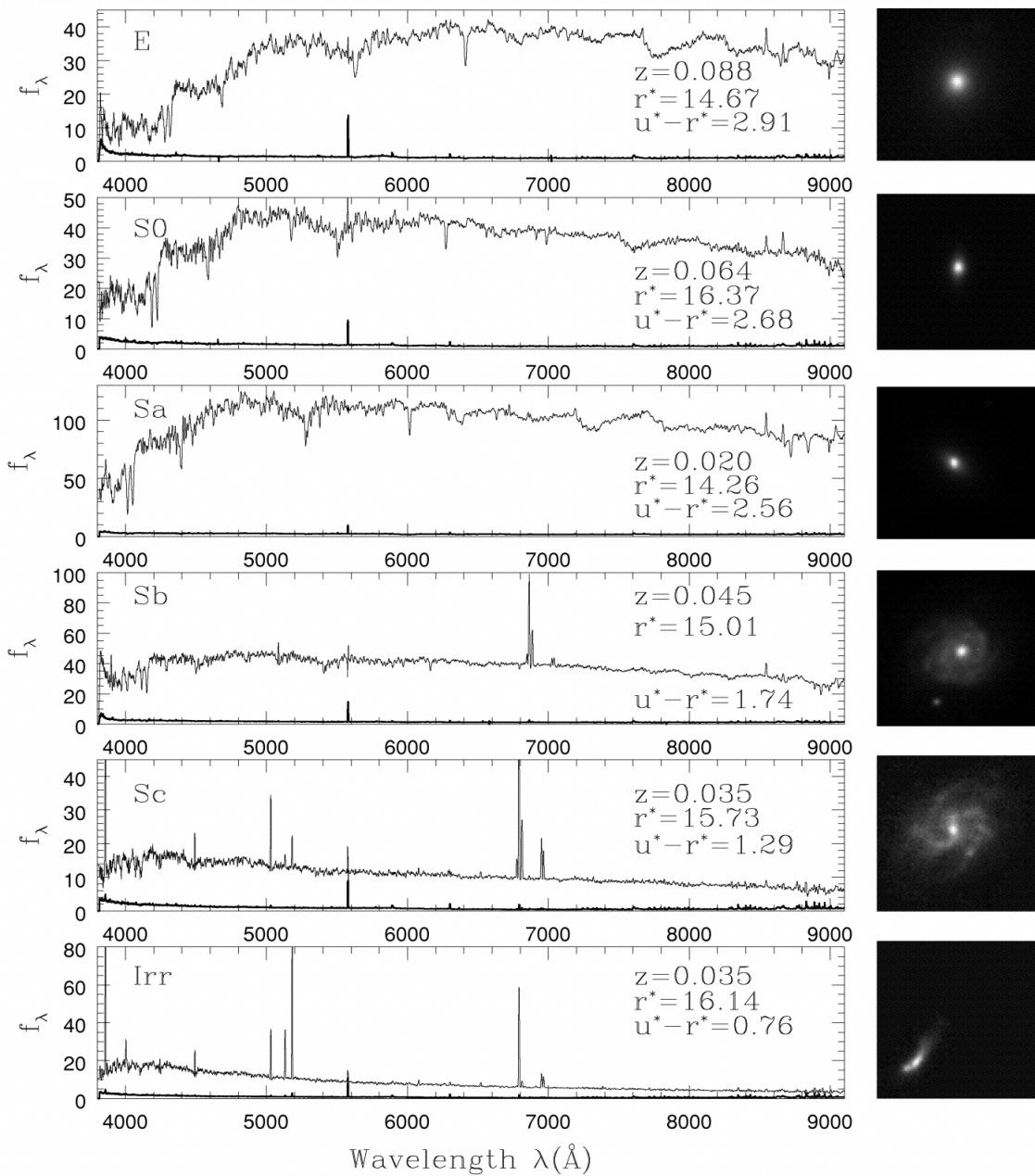
- **Correlations of Galaxy Parameters**
- Many physical parameters are correlated with each other; for example, the luminosity, concentration of the light profile, and spectral line strengths are correlated with colors
- In the color-color space, galaxies form a very thin locus: **the SEDs of galaxies are nearly one-dimensional family** (at the level of ~ 0.03 mag)
- Next page: SDSS sample from Blanton et al. (2003); the quantities are $u - g$, $g - r$, $r - i$ and $i - z$ colors, surface brightness, Sersic index, and absolute magnitude in the r band; the grayscale plots show galaxy distribution in 2D diagrams, together with distributions of each individual quantity (histograms).



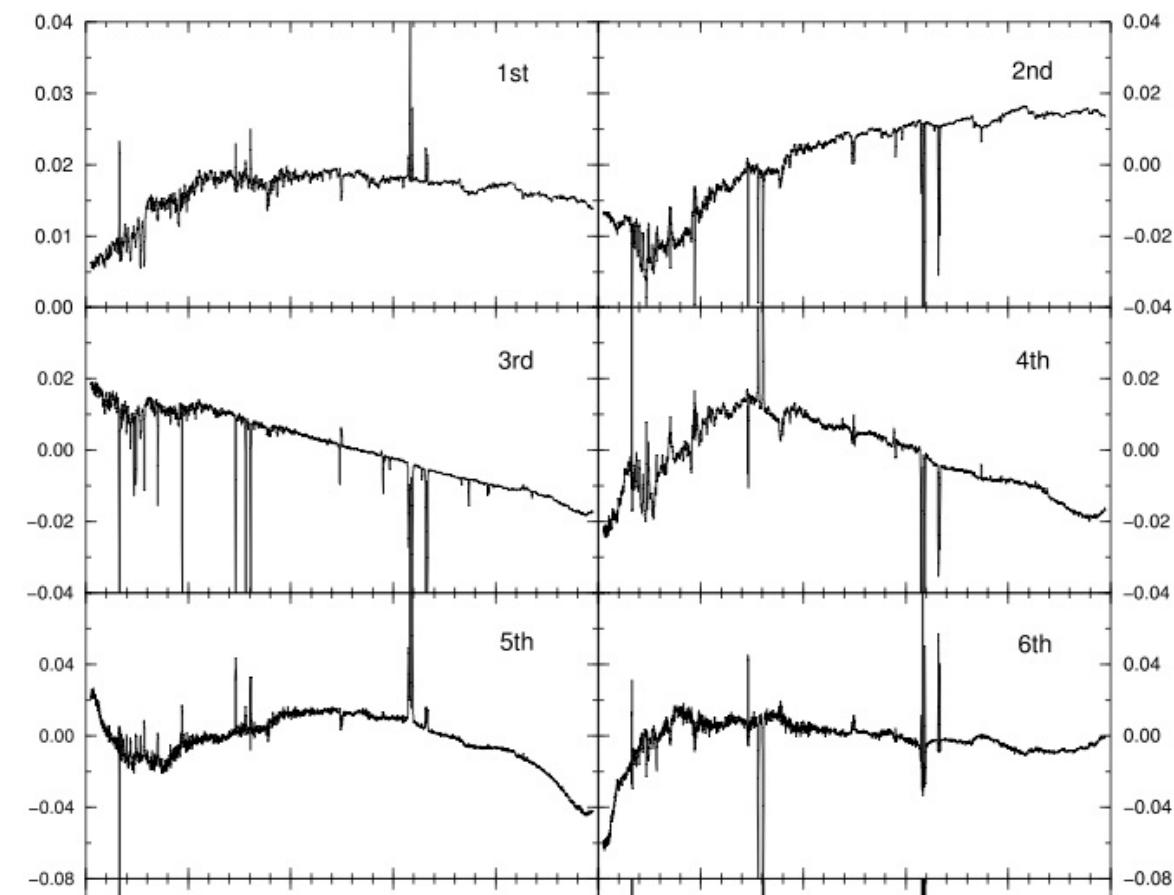


GALEX/SDSS/2MASS SEDs of galaxies

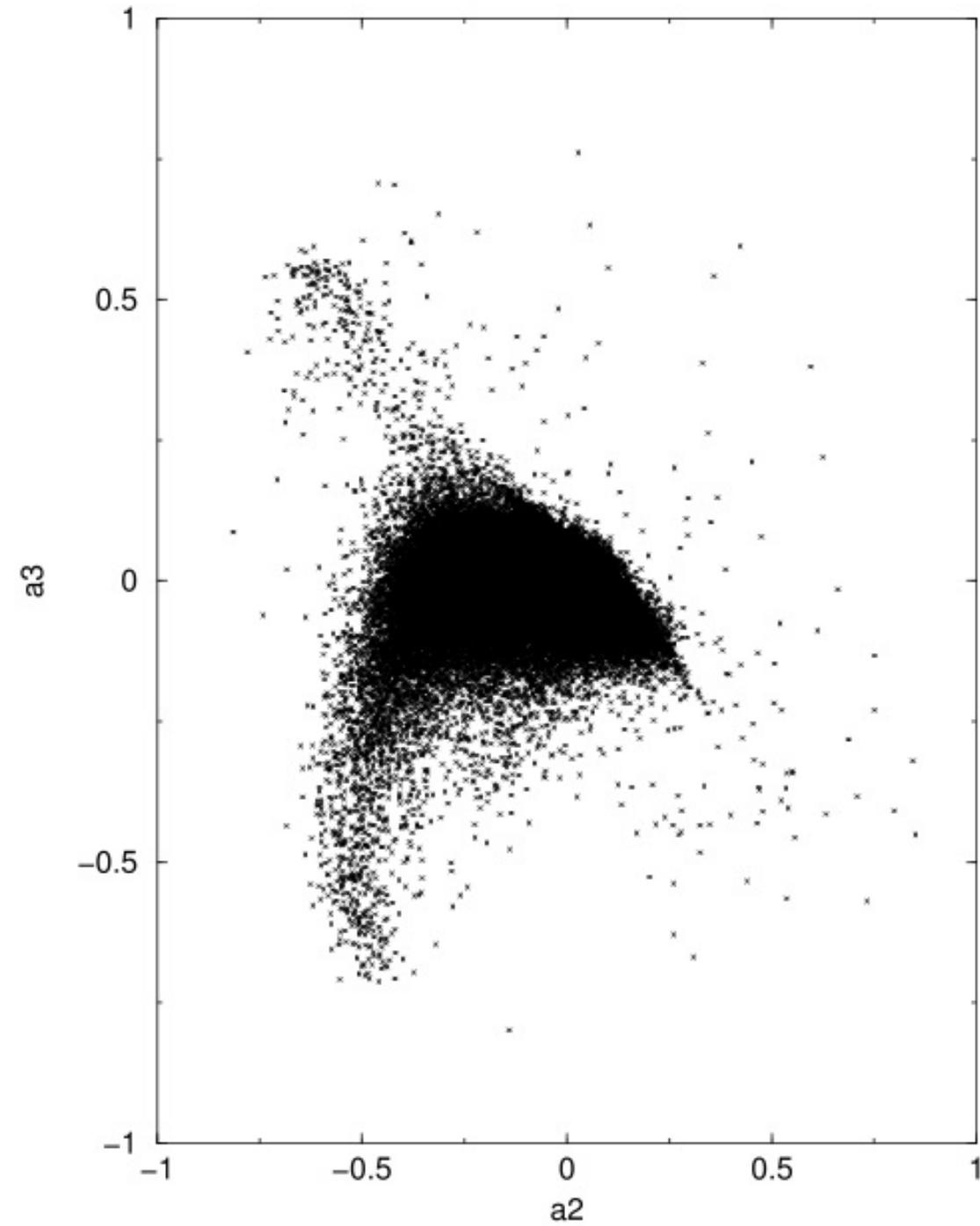




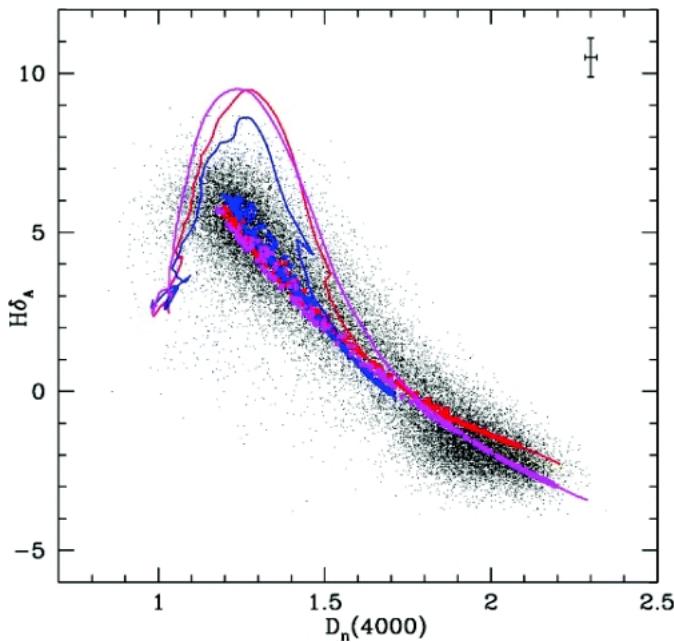
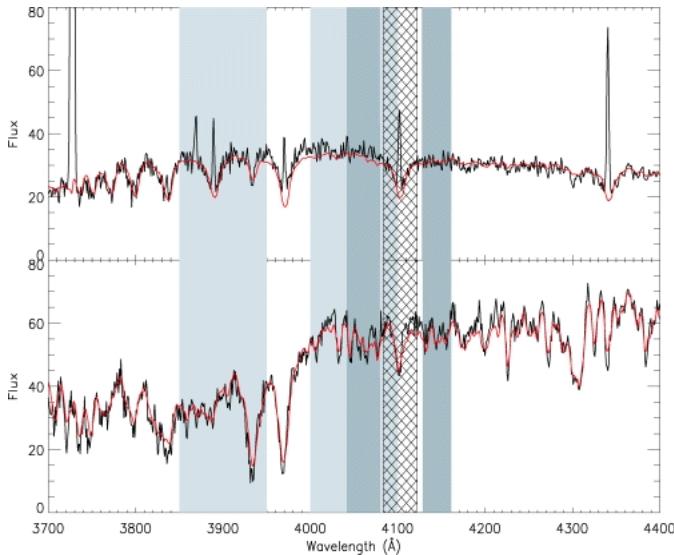
- Spectra are correlated with morphology
- Principal component analysis: spectra form a low-dimensional family: it is possible to describe most of variance using only 2 parameters (Yip et al. 2004)



- Spectra are correlated with morphology Principal component analysis: eigenspectra can be non-negative, which is not physical; there are other mathematical methods, such as non-negative matrix factorization (also, non-linear methods known as manifold learning, e.g. locally linear embedding and isometric mapping), see astroML for code.



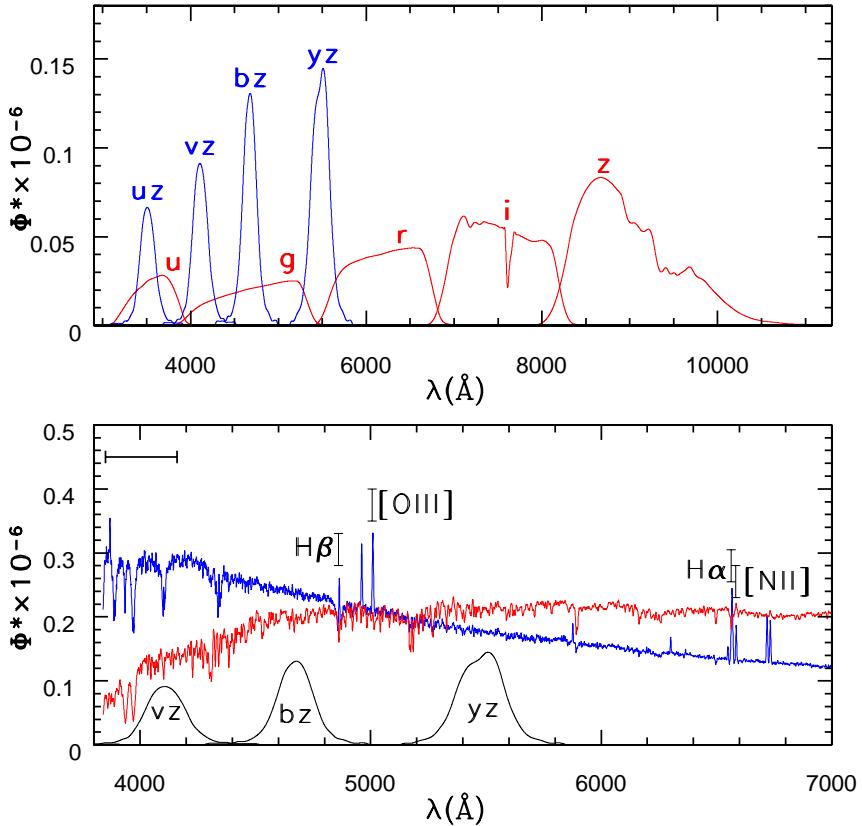
- Spectra are correlated with morphology
- Principal component analysis: spectra form a low-dimensional family: it is possible to describe most of variance using only 2 parameters (Yip et al. 2004)
- What physical quantities can we extract from SDSS spectra of galaxies?



Spectral analysis

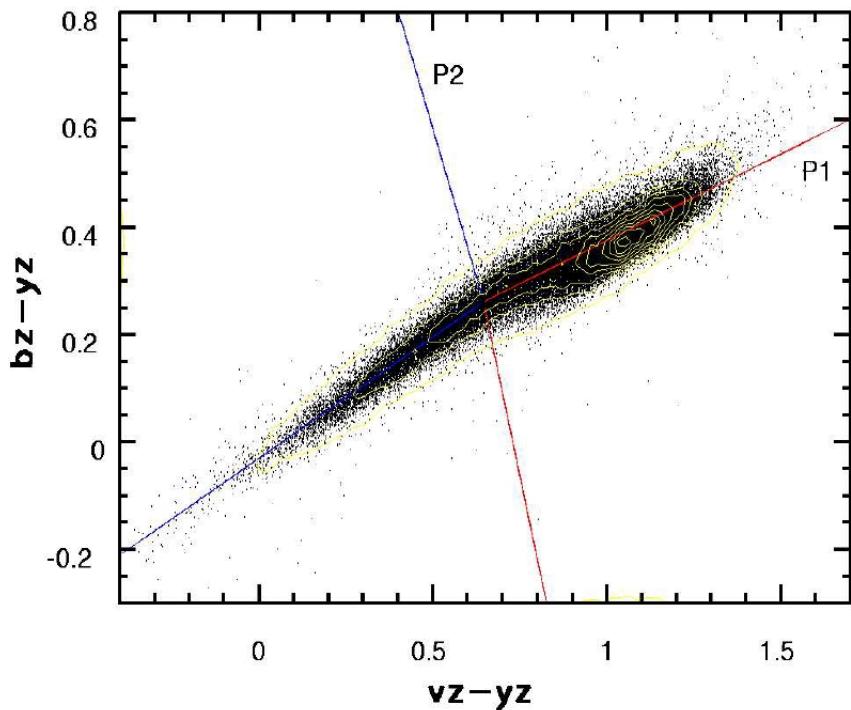
- Kauffmann et al. (2003, 2004): model-dependent estimates of stellar mass and dust content using H_δ , D_{4000} and broad-band colors
- From the position in the $H_\delta - D_{4000}$ diagram, get a model-dependent estimate of stellar mass-to-light ratio, and using measured luminosity get stellar mass. The measured luminosity is corrected for the dust extinction estimated from the discrepancy between the model-predicted and measured broad-band colors.

Narrow-band rest-frame colors of galaxies in SDSS

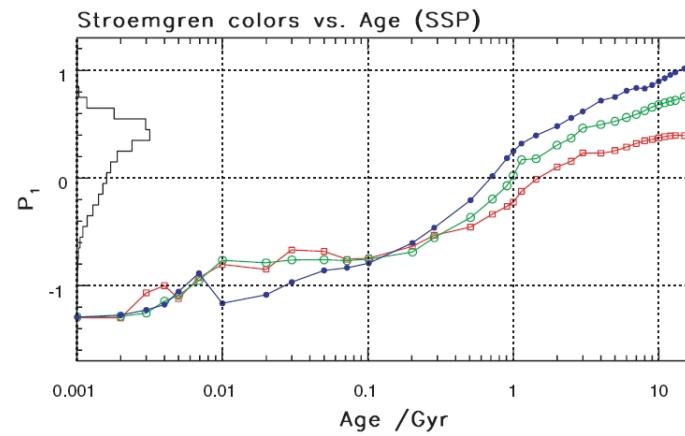
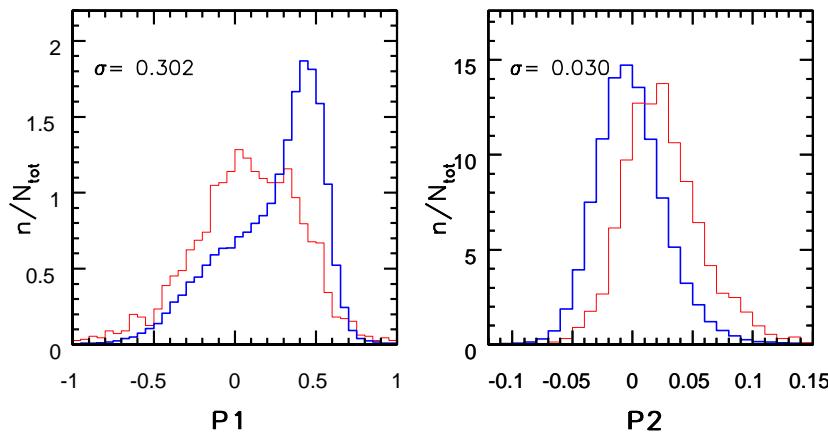


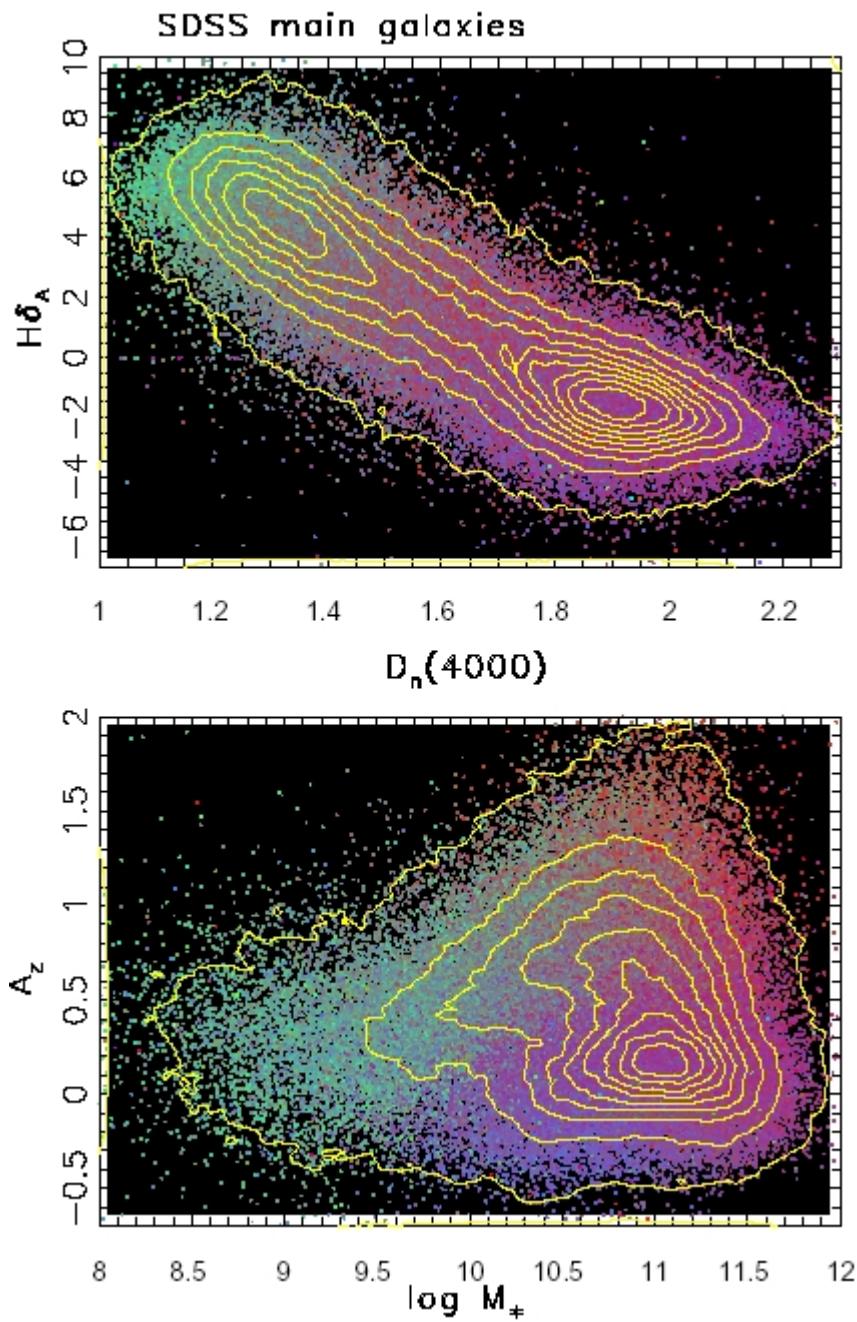
- Strömgren colors: designed for estimating effective temperature, metallicity and gravity for stars; narrow band (200 Å)
- SDSS spectra can be used to synthesize narrow-band rest-frame colors for galaxies
- Can we determine age and metallicity of galaxies?

Narrow-band rest-frame colors of galaxies in SDSS



- Galaxies form a narrow locus in rest-frame colors: SEDs are nearly a one-dimensional family, with the Hubble type controlling the position along the locus (Smolčić et al. 2006, MNRAS 371, 121)
- The locus width is only 0.03 mag; the offset in the “narrow” direction is correlated with the dust content
- In order to demonstrate that this width is only 0.03 mag, an accurate survey such as SDSS was needed!

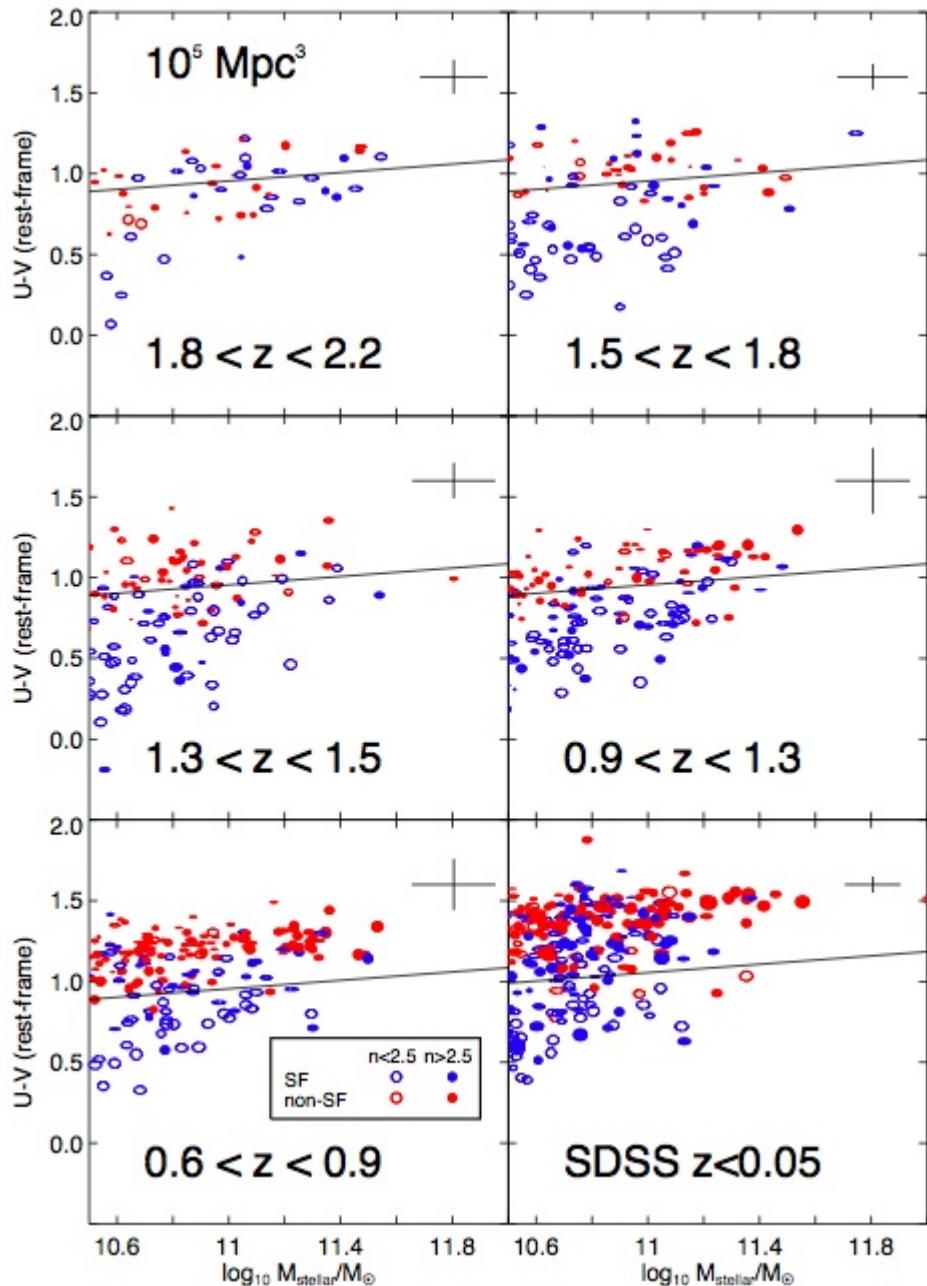




Narrow-band rest-frame colors of galaxies in SDSS

- Left: color-coded by P1 from the previous page
- Many observables are correlated!
- E.g. the rest-frame colors with the position in the $H_\delta - D_{4000}$ diagram
- This implies: galaxy mass and colors are well correlated – the bimodal distribution in colors reflects a characteristic galaxy mass (stars only): $\sim 3 \times 10^{10} M_\odot$.
- This characteristic galaxy mass probably marks the transition between different physical processes that control galaxy formation and evolution

Galaxy Evolution



- SDSS observations are essentially a snapshot at $z = 0$ (now!).
- Bell et al. (2012, ApJ 753, 167): “We use HST/WFC3 imaging from the CANDELS Multicycle Treasury Survey, in conjunction with the Sloan Digital Sky Survey, to explore the evolution of galactic structure for galaxies with stellar masses $> 3e10 M_{\odot}$ from $z = 2.2$ to the present epoch, a time span of 10 Gyr. We explore the relationship between rest-frame optical color, stellar mass, star formation activity and galaxy structure.”
- Left: (fig. 4): The evolution of $U - V$ rest-frame color as a function of stellar mass, in redshift bins (the symbol size scales with observed galaxy size, and are color-coded by star-formation: quiescent galaxies are red). ¹⁹

- **Morphology-density relation:** red elliptical galaxies are found in regions more populated by galaxies than blue spiral galaxies (Dressler 1980, ApJ 236, 351; Postman & Geller 1984, ApJ 281, 95)
- SDSS-based visualization (Cowan & Ivezić 2008, ApJ 674, L13):

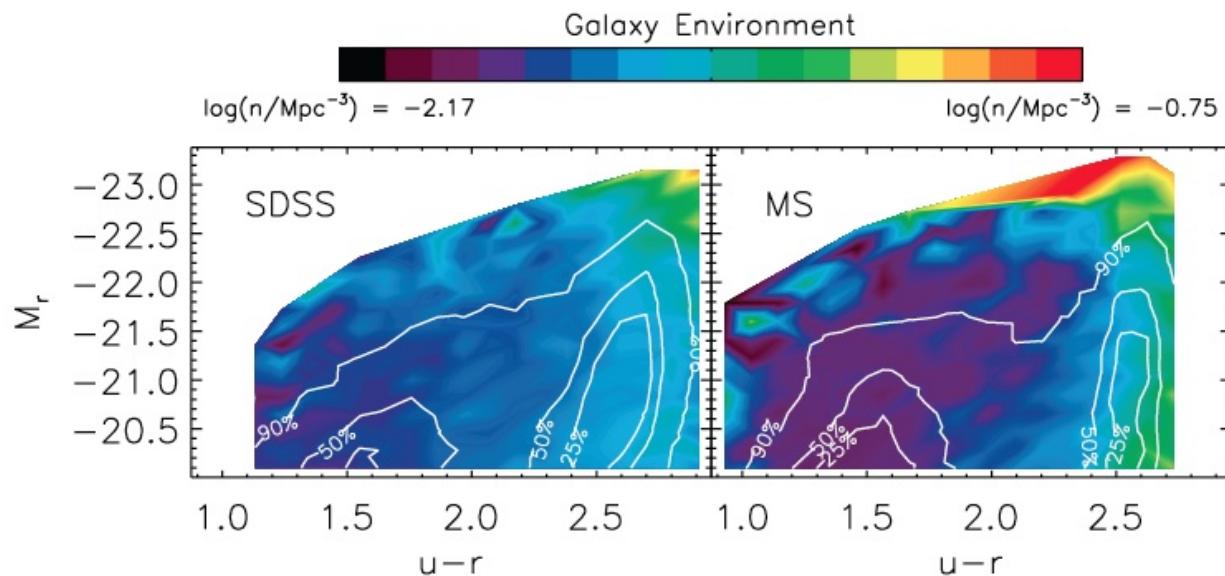
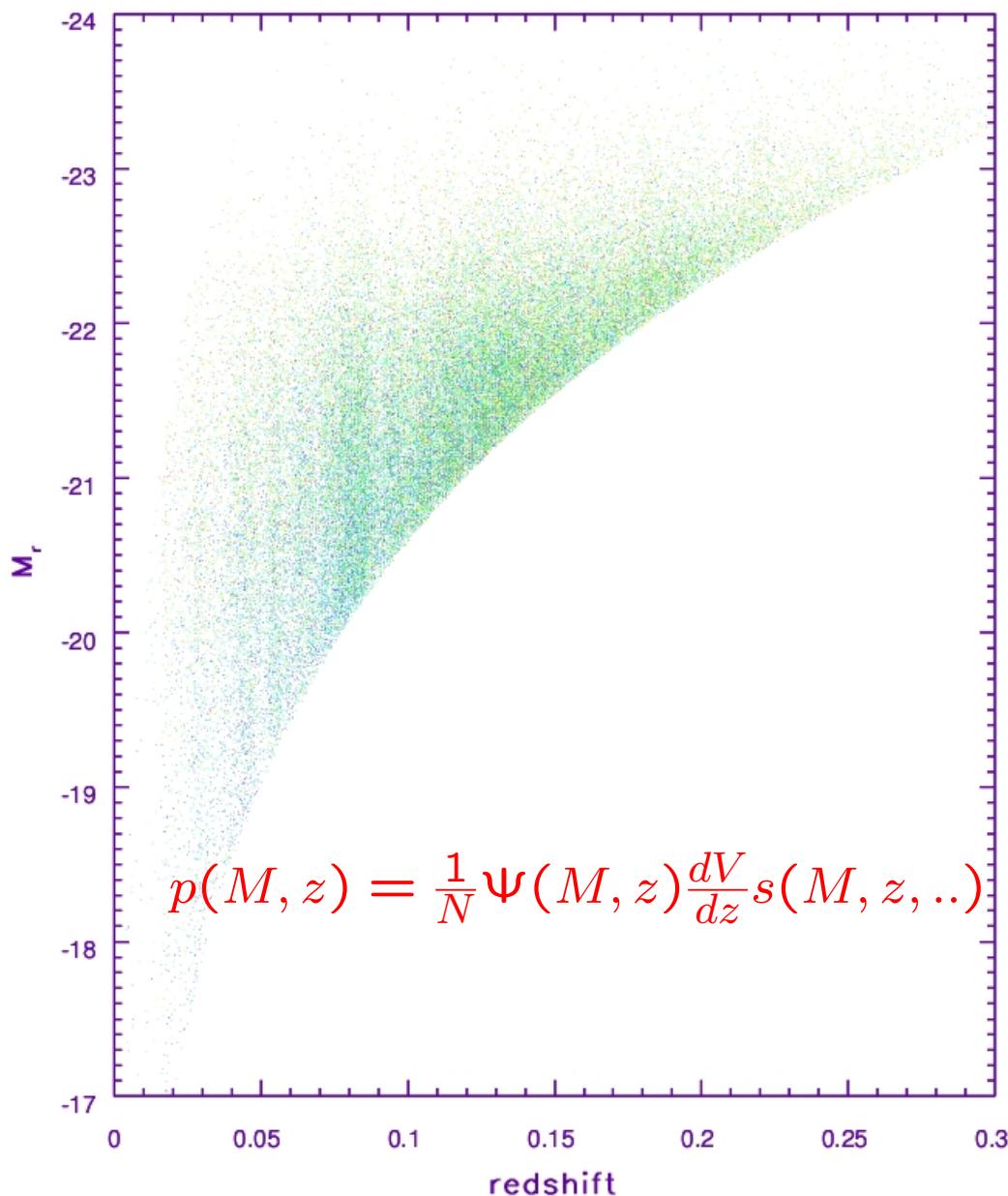


FIG. 5.—Color-magnitude diagram for SDSS and MS galaxies in the left and right panels, respectively. The labeled white lines show which regions on the plot are most populated (these are complete volume-limited samples), and the color-coded background shows the median local environment around galaxies with a given color and magnitude (dark corresponds to low densities; bright corresponds to high densities).

Luminosity Function

- Luminosity Function is the distribution in the luminosity–position plane; how many galaxies per unit interval in luminosity and unit volume (or redshift): $\Psi(M, z)$
- Imagine a tiny area with the widths ΔMr and Δz centered at some Mr and z in the plot to the left: count the number of galaxies, divide by $\Delta Mr \Delta z$, correct for the fraction of sky covered by your survey, and for the selection probability (a function of Mr , z , and possibly many other parameters): this gives you $\Psi(M, z)$.



Luminosity Function

- Luminosity Function is the distribution in the luminosity–position plane; how many galaxies per unit interval in luminosity and unit volume: $\Psi(M, z)$
- Often, this is a separable function: $\Psi(M, z) = \Phi(M) n(z)$, where $\Phi(M)$ is the absolute magnitude (i.e. luminosity) distribution, and $n(z)$ is the number volume density.
- Luminosity is a product of flux and distance squared (ignore cosmological effects for simplicity): $L = 4\pi D^2 F$
- The samples are usually *flux-limited* (meaning: all sources brighter than some flux limit are detected) – the minimum detectable luminosity depends on distance: $L > 4\pi D^2 F_{min}$, or for absolute magnitude $M < M_{max}(D)$ (c.f. the first plot)

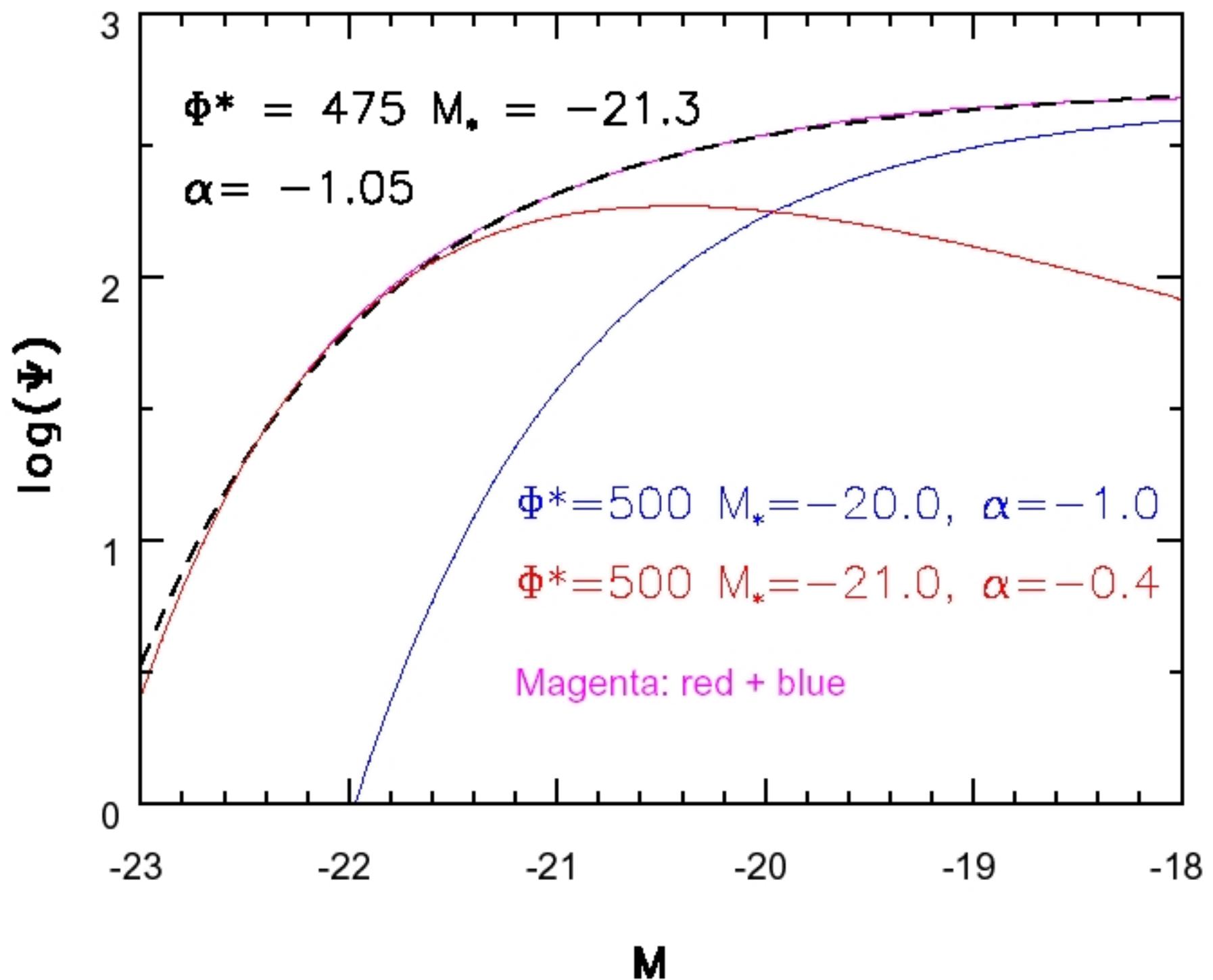
Schechter Function

Galaxy luminosity distribution resembles a power-law, with an exponential cutoff. This distribution is usually modeled by Schechter function:

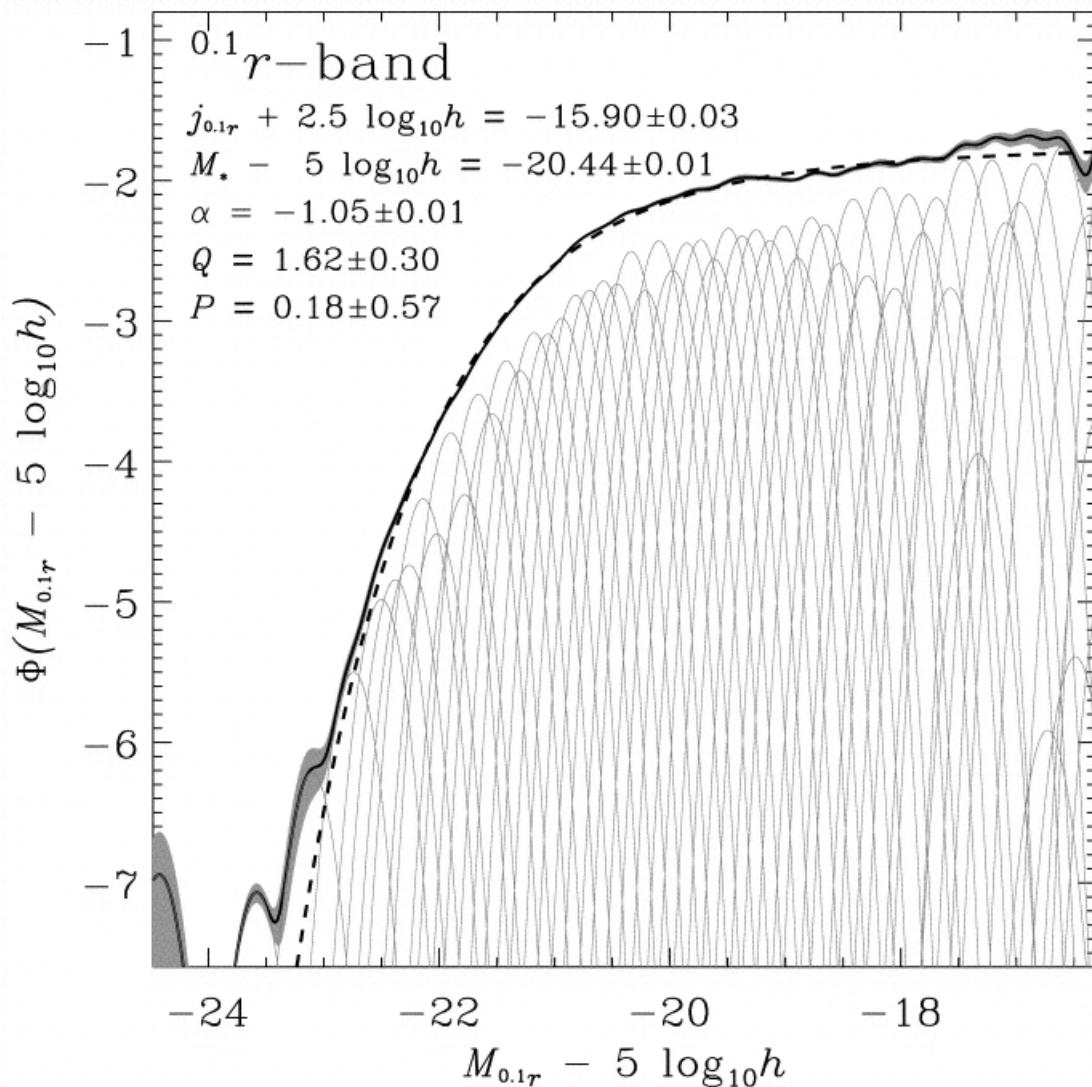
$$\Phi(L) = \Phi^* \left(\frac{L}{L_*} \right)^\alpha e^{-L/L_*} \quad (1)$$

Or using absolute magnitudes:

$$\Phi(Mr) = 0.4\Phi^* e^{-0.4(\alpha+1)(Mr-M^*)} e^{-e^{-0.4(Mr-M^*)}} \quad (2)$$



The LF in the SDSS r band

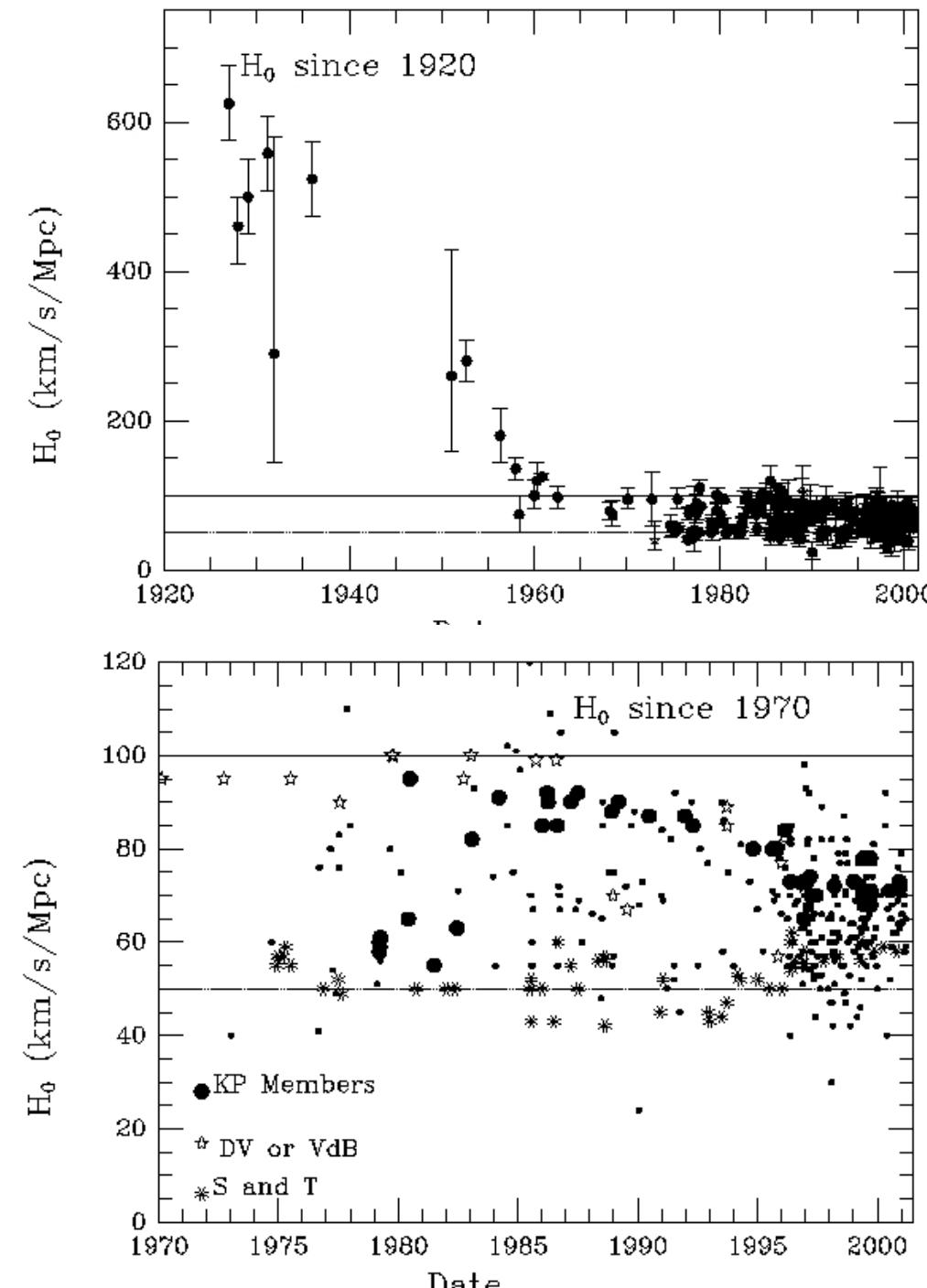


Note: this LF cannot be expressed as $\Phi(M, z) = f(M) g(z)$
– not separable!

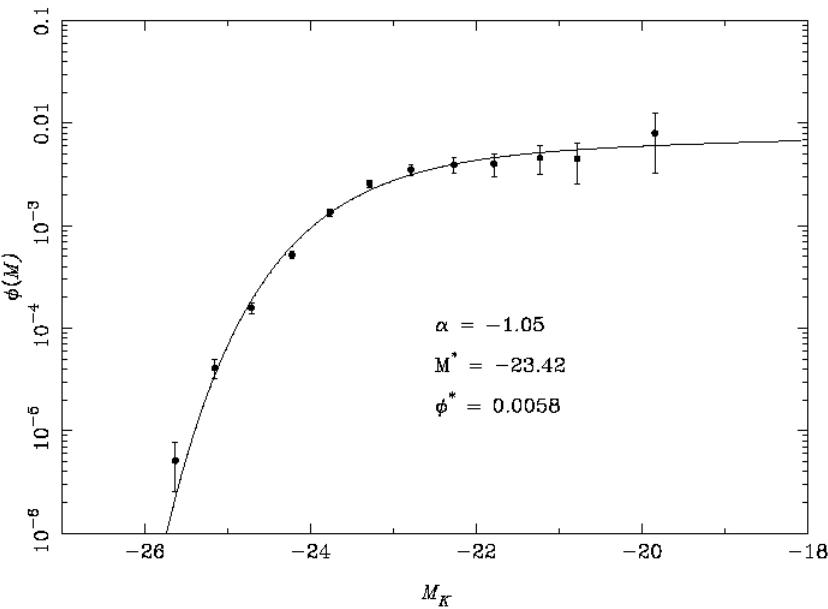
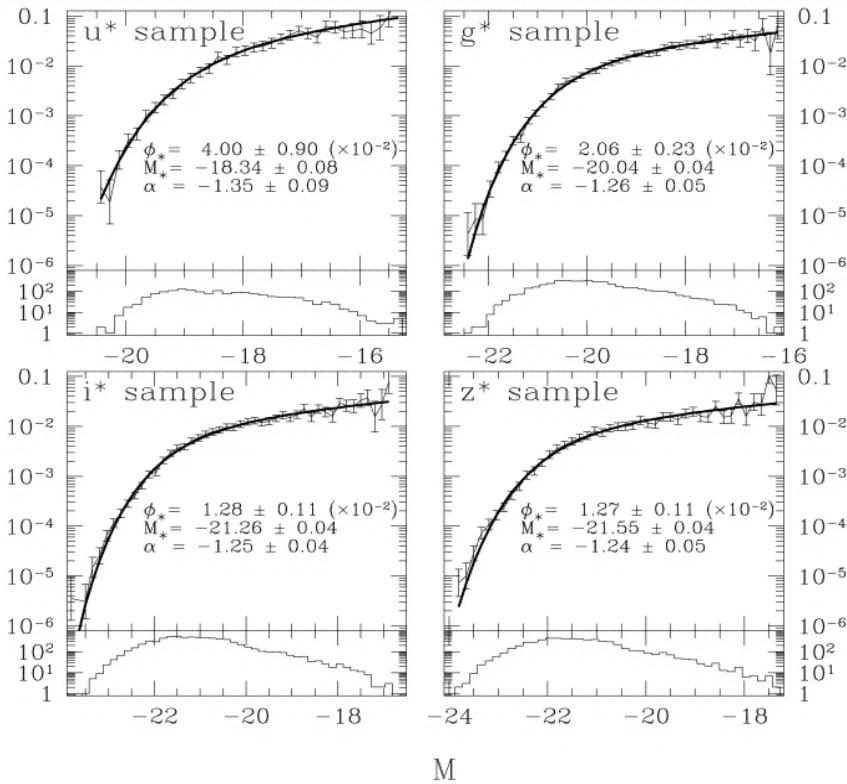
- The thick solid line is the SDSS r band luminosity function, and the gray band is its uncertainty.
- The dashed line is a Schechter-like fit that also includes the effects of changing luminosity and the number density with time (i.e. distance, or redshift). $Q > 0$ indicates that galaxies were more luminous in the past, and $P > 0$ that galaxies were more numerous in the past. For detailed discussion, see Blanton et al. 2003 (Astronomical Journal, 592, 819-838)

H_0 as “a function of time”

- the first three points: Lemaitre (1927), Robertson (1928), Hubble (1929), all based on Hubble's data
- the early low value (290 km/s/Mpc): Jan Oort
- the first major revision: discovery of Population II stars by Baade
- the very recent convergence to values near 65 ± 10 km/sec/Mpc
- the best Cepheid-based value for the local H_0 determination is 71 ± 7 km/s/Mpc, the WMAP5 value based on cosmic microwave background measurements: 72 ± 3 km/s/Mpc.
- WMAP9: 69.3 ± 0.8 km/s/Mpc, and the Planck Mission: $H_0 = 67.8 \pm 0.8$ km/s/Mpc ($h = 0.678$)
- Thusly, $-5\log_{10}(h) = 0.84$ mag!

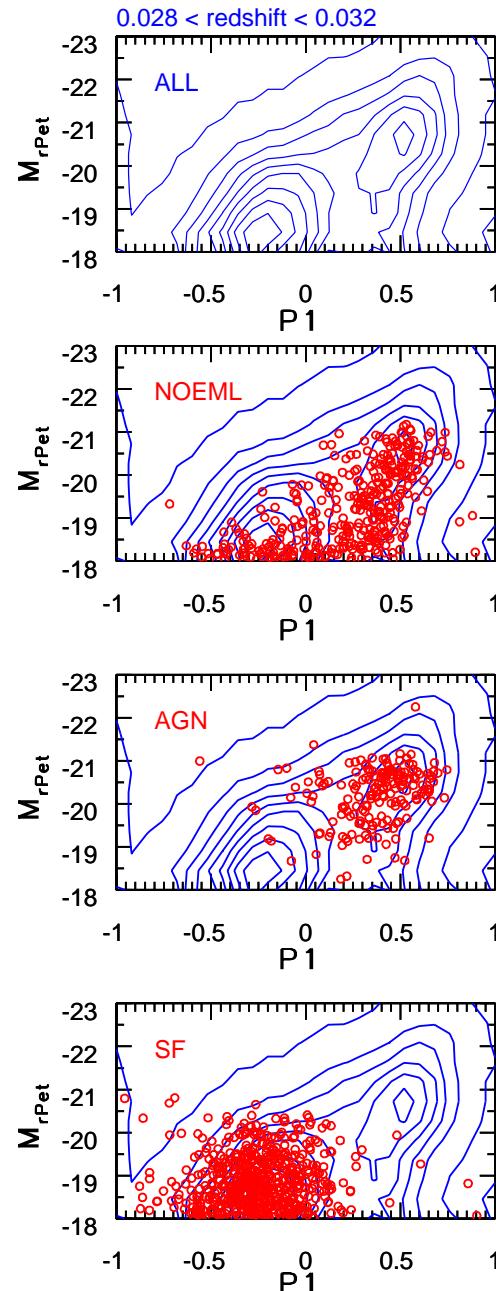


Copyright SAO 2001



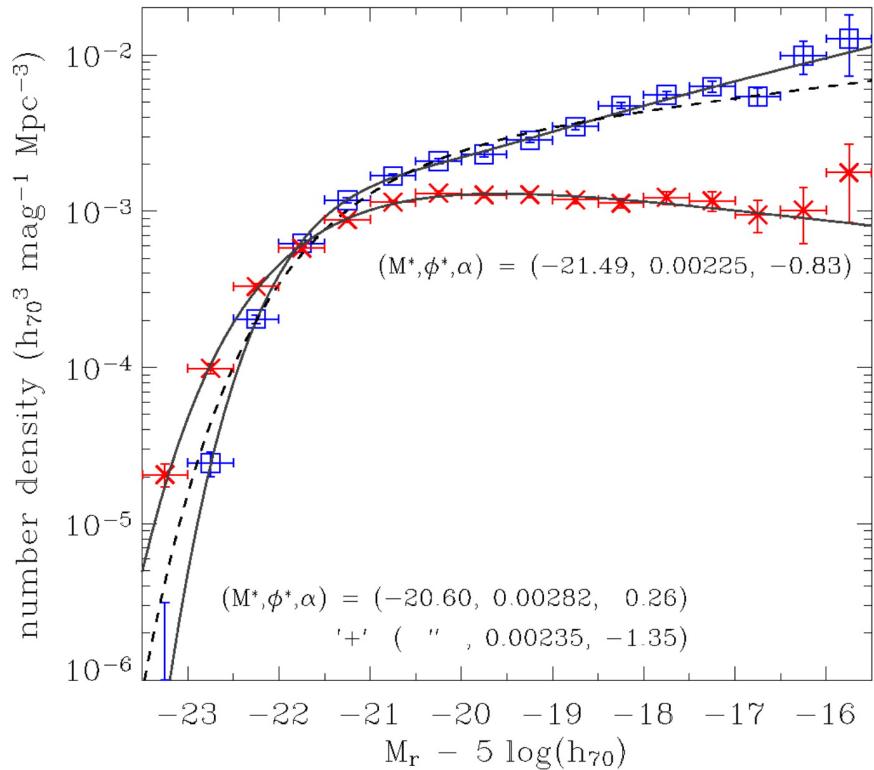
The dependence of LF on wavelength

- Top: SDSS ugi* bands
- Bottom: 2MASS K band
- The Schechter function is still a good fit, but best-fit parameters vary.
- Since the SEDs of galaxies are nearly one-dimensional families, once the LF for a sample selected by color or morphology is known, the LFs at other wavelengths can be simply obtained by shifting the M axis by the appropriate color difference.
- This doesn't work for the LFs in the top four panels because they are computed for the whole sample.



The dependence of LF on galaxy type

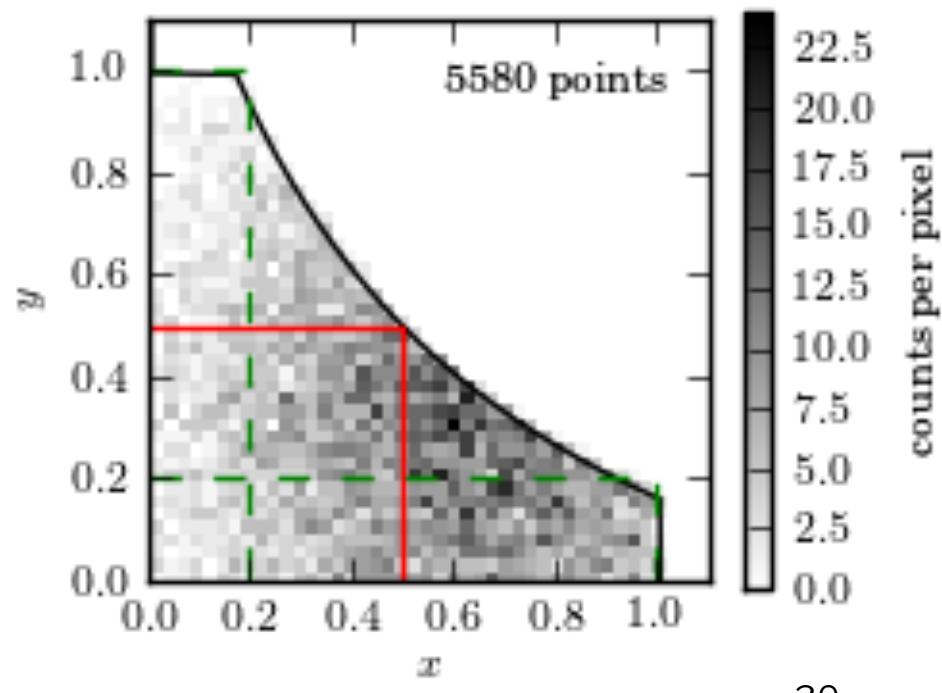
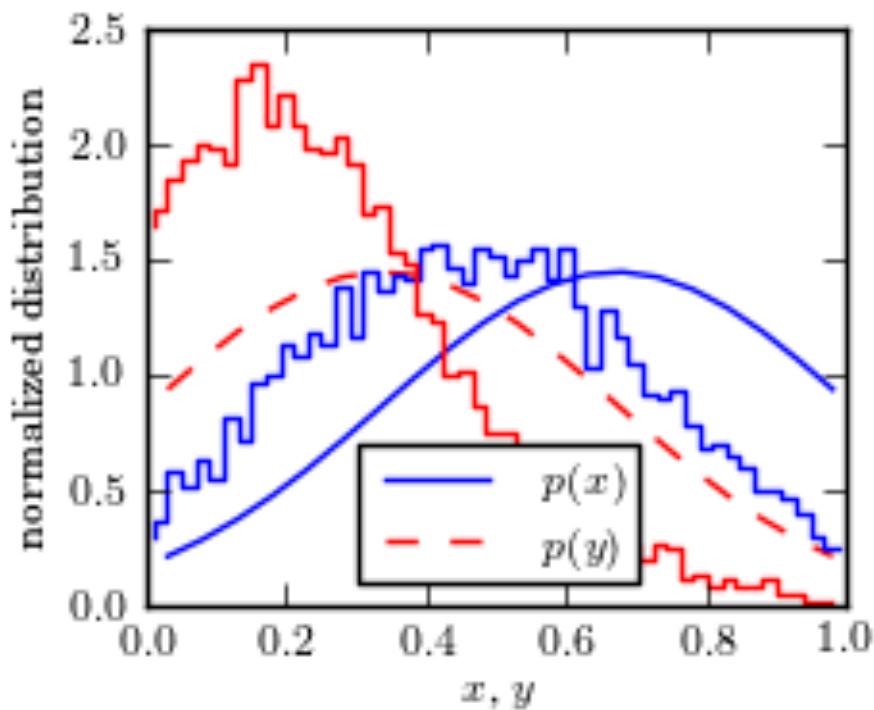
- The top panel shows the distribution of SDSS galaxies in the absolute magnitude – color plane (in a narrow redshift range)
- In the bottom three panels, the same distribution is compared to the distributions for subsamples selected by their emission line properties
- Note that the most luminous galaxies ($M_r < -20$) are predominantly red ($P_1 > 0.2$), while faint galaxies ($M_r > -19$) are blue ($P_1 < 0.2$)



The dependence of LF on galaxy type

- The comparison of LFs for blue and red galaxies (from Baldry et al. 2004, ApJ, 600, 681-694)
- The red distribution has a more luminous characteristic magnitude and a shallower faint-end slope, compared to the blue distribution
- The transition between the two types corresponds to stellar mass of $\sim 3 \times 10^{10} M_{\odot}$
- The differences between the two LFs are consistent with the red distribution being formed from major galaxy mergers.

- Toy example: the x-y “measurements” are truncated (similar to a flux-limited sample in astronomy).
- Histograms of measured quantities are biased for a truncated sample.
- Unbiased distributions can be obtained using various methods (V/V_{max} method, C^- method, maximum likelihood method)

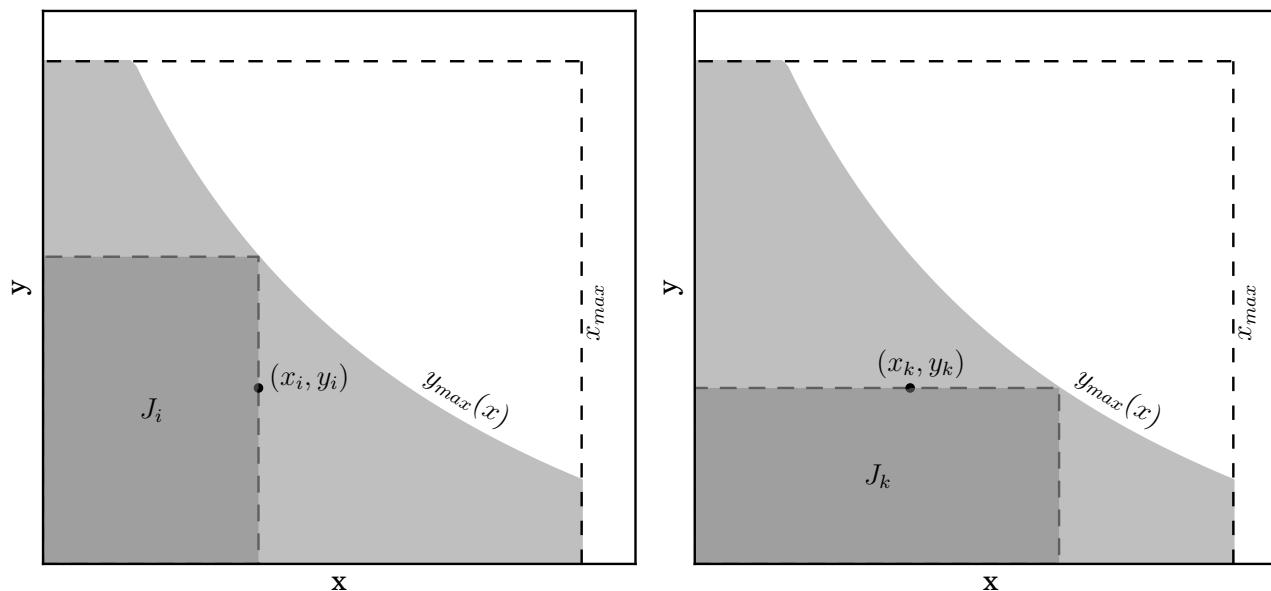


The C^- method for estimating LF

- Lynden-Bell (1971, MNRAS 155, 95); a non-parametric method that works for separable LFs, $\Psi(L, z) = \Phi(L)n(z)$
- practically all non-parametric methods can be reduced to the C^- method (Petrosian 1992)
- parametric methods are usually based on maximizing likelihood (e.g. Marshall 1985)
- the simplest and most famous method, the V_{max} method (Schmidt 1968), requires binning in two axes simultaneously, while with the C^- method data is binned only one axis at a time (e.g. Fan et al. 2001)
- How do we know that separable LF is a good guess for our data?

C⁻ method

- Given a set of measured pairs (x_i, y_i) , with $i = 1 \dots N$, and known relation $y_{max}(x)$, estimate the two-dimensional distribution, $n(x, y)$, from which the sample was drawn. Assume that measurement errors for both x and y are negligible compared to their observed ranges, that x is measured within a range defined by x_{min} and x_{max} , and that the selection function is 1 for $0 \leq y \leq y_{max}(x)$ and $x_{min} \leq x \leq x_{max}$, and 0 otherwise.

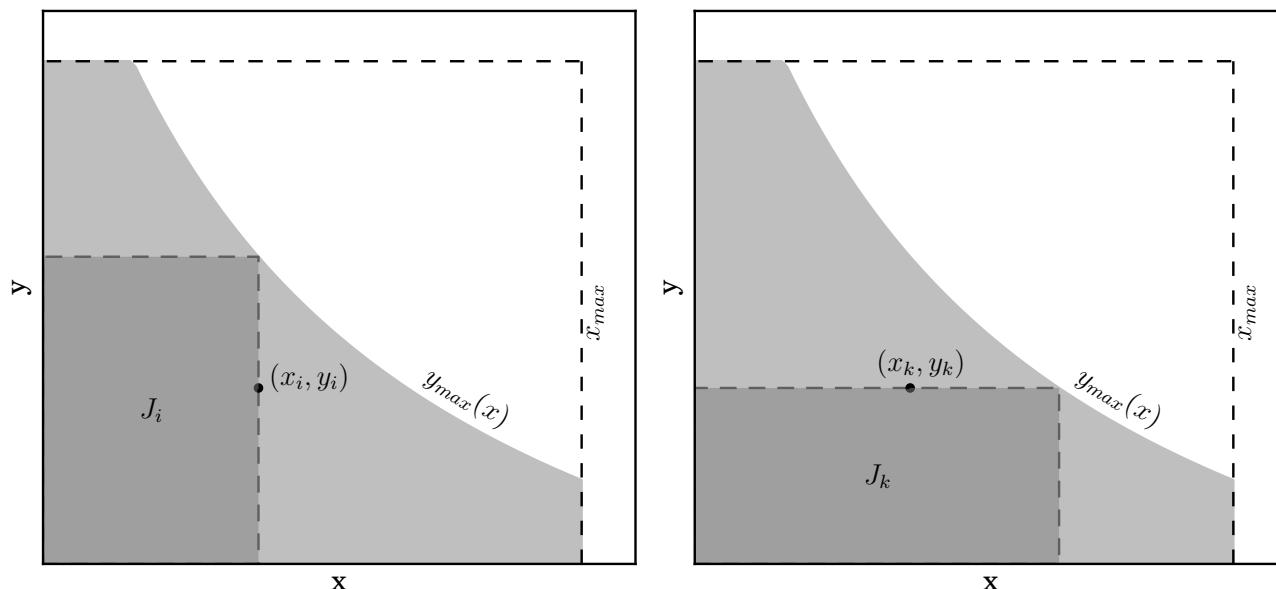


C⁻ method

- C^- method is applicable when the distributions along the two coordinates x and y are uncorrelated, that is, when we can assume that the bivariate distribution $n(x, y)$ is separable

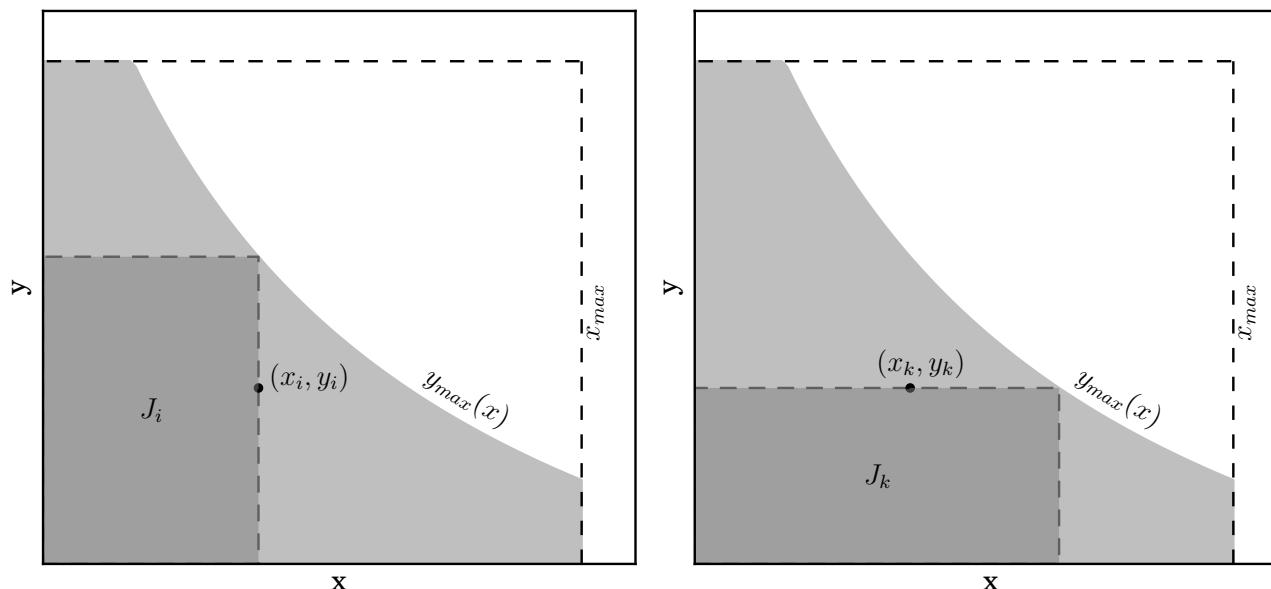
$$n(x, y) = \Psi(x) \rho(y). \quad (3)$$

Therefore, before using the C^- method we need to demonstrate that this assumption is valid.



C⁻ method

- Define a *comparable* or *associated* set for each object i such that $J_i = \{j : x_j < x_i, y_j < y_{max}(x_i)\}$; this is the largest x -limited and y -limited data subset for object i , with N_i elements (see the left panel).
- Sort the set J_i by y_j ; this gives us the rank R_j for each object (ranging from 1 to N_i)



C⁻ method

- Define a *comparable* or *associated* set for each object i such that $J_i = \{j : x_j < x_i, y_j < y_{max}(x_i)\}$; this is the largest x -limited and y -limited data subset for object i , with N_i elements.
- Sort the set J_i by y_j ; this gives us the rank R_j for each object (ranging from 1 to N_i)
- Define the rank R_i for object i in *its* associated set: this is essentially the number of objects with $y < y_i$ in set J_i .
- If x and y are truly independent, R_i must be distributed *uniformly* between 0 and N_i .

C⁻ method

- If x and y are truly independent, R_i must be distributed *uniformly* between 0 and N_i .
- In this case, it is trivial to determine the expectation value and variance for R_i : $E(R_i) = E_i = N_i/2$ and $V(R_i) = V_i = N_i^2/12$. We can define the statistic

$$\tau = \frac{\sum_i (R_i - E_i)}{\sqrt{\sum_i V_i}} \quad (4)$$

If $\tau < 1$, then x and y are uncorrelated at $\sim 1\sigma$ level!

C⁻ method

Assuming that $\tau < 1$, it is straightforward to show using relatively simple probability integral analysis (e.g., see Appendix in Fan et al. 2001), as well as the original Lynden-Bell's paper, how to determine cumulative distribution functions. The cumulative distributions are defined as

$$\Phi(x) = \int_{-\infty}^x \Psi(x') dx', \quad (5)$$

and

$$\Sigma(y) = \int_{-\infty}^y \rho(y') dy'. \quad (6)$$

Then,

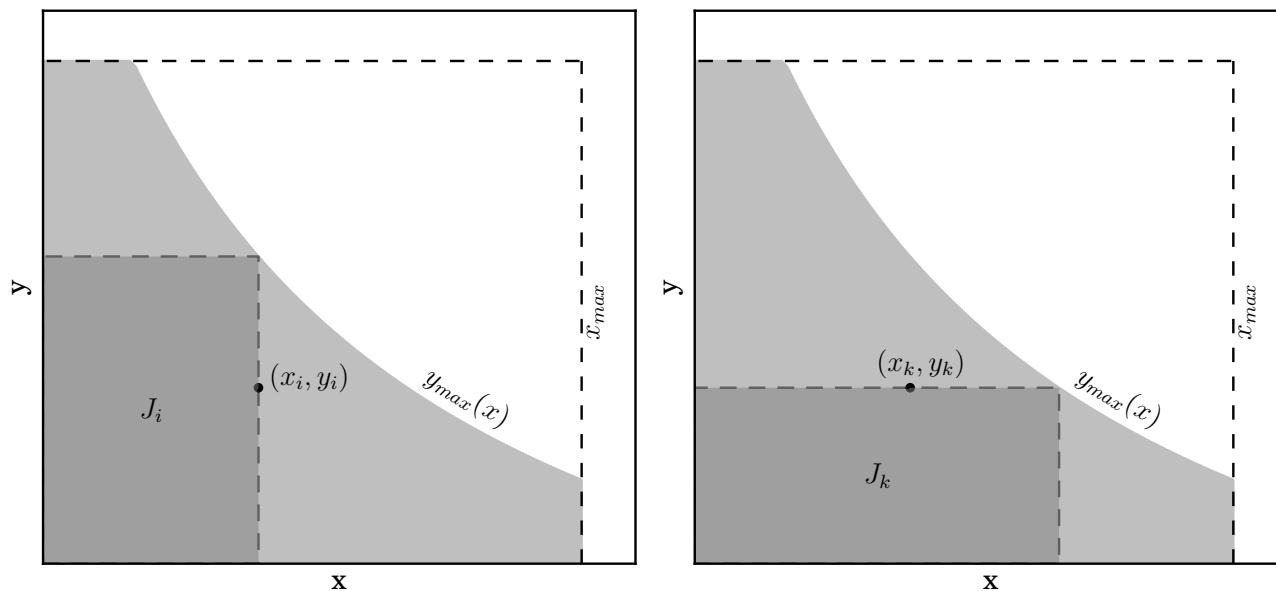
$$\Phi(x_i) = \Phi(x_1) \prod_{k=2}^i (1 + 1/N_k) \quad (7)$$

where it is assumed that x_i are sorted ($x_1 \leq x_k \leq x_N$).

C⁻ method

Analogously, if M_k is the number of objects in a set defined by $J_k = \{j : y_j < y_k, y_{max}(x_j) > y_k\}$ (see the right panel of figure below), then

$$\Sigma(y_j) = \Sigma(y_1) \prod_{k=2}^j (1 + 1/M_k). \quad (8)$$

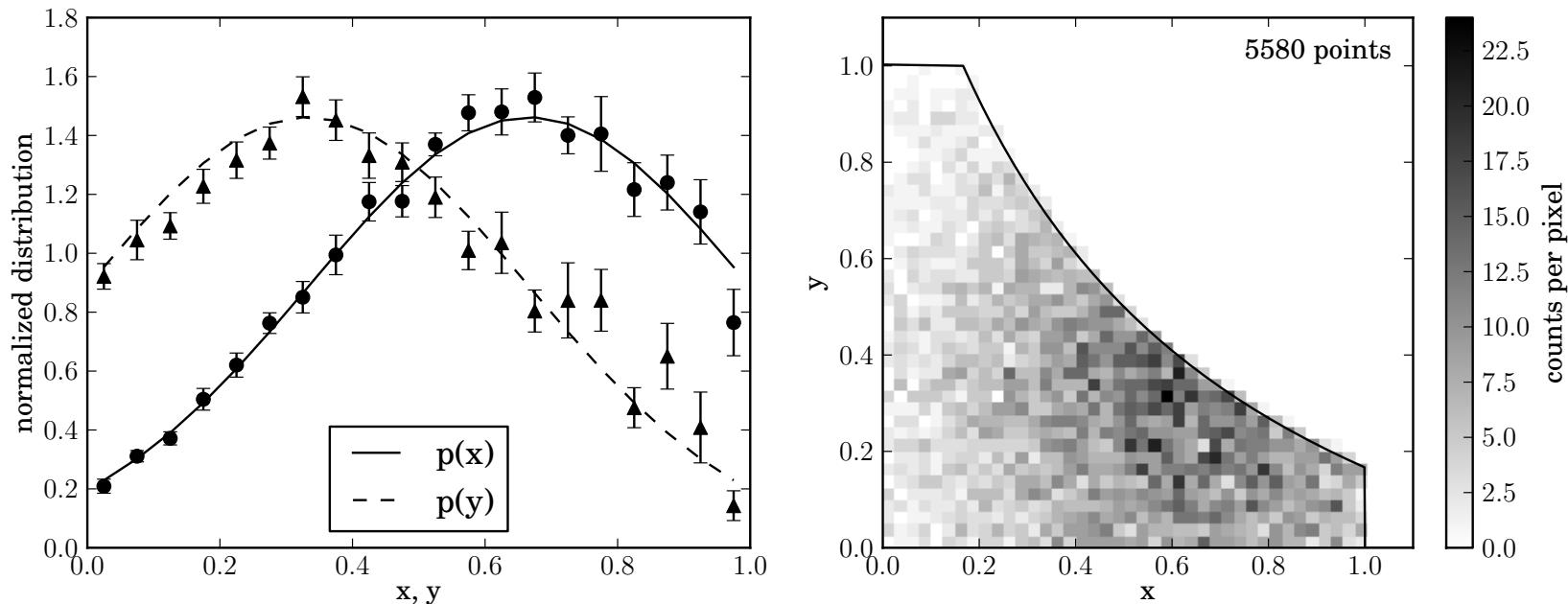


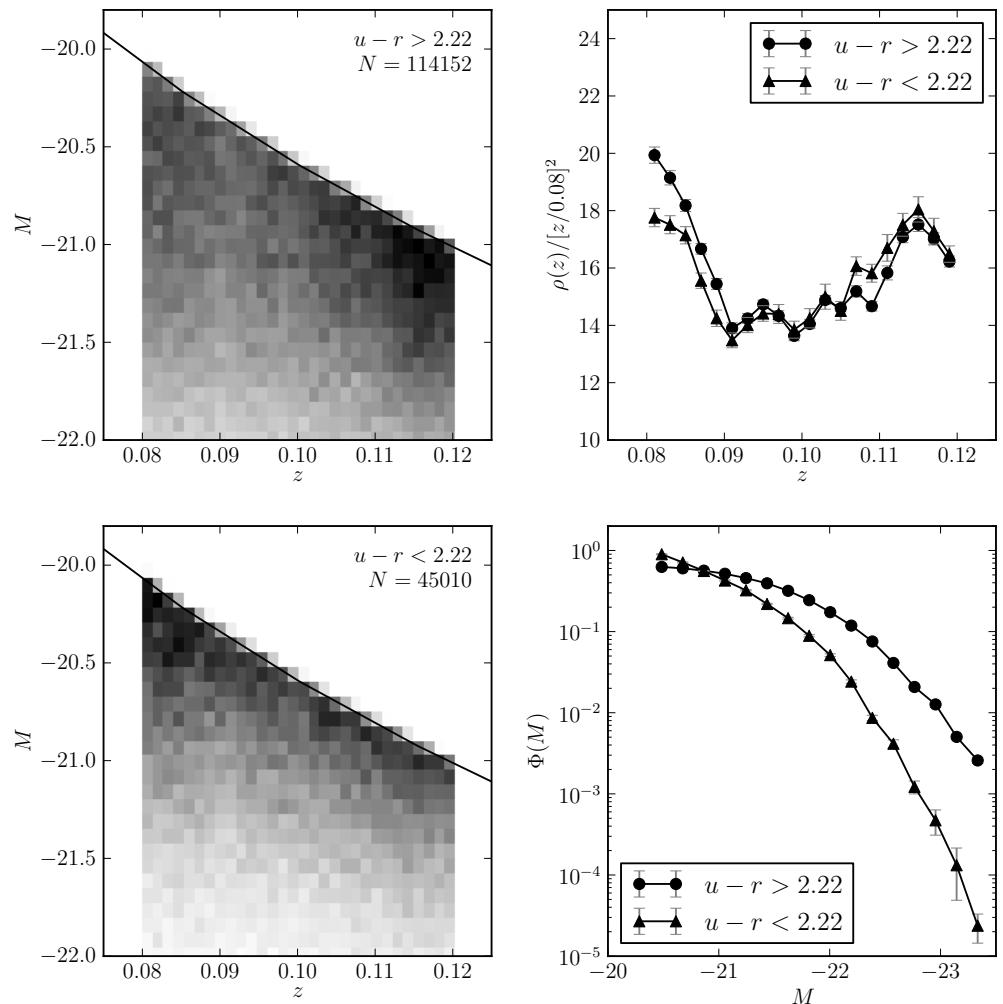
astroML implementation of the C^- method

- Note that both $\Phi(x_j)$ and $\Sigma(y_j)$ are defined on non-uniform grids with N values, corresponding to the N measured values.
- Essentially, the C^- method assumes a piece-wise constant model for $\Phi(x)$ and $\Sigma(y)$ between data points (equivalently, differential distributions are modeled as Dirac's δ functions at the position of each data point).
- As shown by Petrosian (1992), $\Phi(x)$ and $\Sigma(y)$ represent an optimal data summary.
- The differential distributions $\Psi(x)$ and $\rho(y)$ can be obtained by differentiating cumulative distributions in the relevant axis; an approximate normalization can be obtained by requiring that the total predicted number of objects is equal to their observed number.

- **astroML Book Figure 4.9:** The right panel shows a realization of truncated separable two-dimensional Gaussian distribution (with the truncation given by the solid line). The lines in the left panel show the true one-dimensional distributions of x and y , and the points are computed from the truncated data set using the C^- method (with error bars from 20 bootstrap resamples).

[http://astroml.github.com/
book_figures/chapter4/fig_lyndenbell_toy.html](http://astroml.github.com/book_figures/chapter4/fig_lyndenbell_toy.html)





- **astroML Book Figure 4.10:** The luminosity function for two $u-r$ color-selected subsamples of SDSS galaxies from the spectroscopic sample, with redshift in the range $0.08 < z < 0.12$ and flux limited to $r < 17.7$.
- The left panels show the distribution of sources as a function of redshift and absolute magnitude. The distribution $p(z, M) = \rho(z)\Phi(m)$ is obtained using Lynden-Bell's method, with errors determined by 20 bootstrap resamples, and shown in the right panels.

[http://astroml.github.com/
book_figures/chapter4/fig_lyndenbell_gals.html](http://astroml.github.com/book_figures/chapter4/fig_lyndenbell_gals.html)

Test of L-z Correlation. II

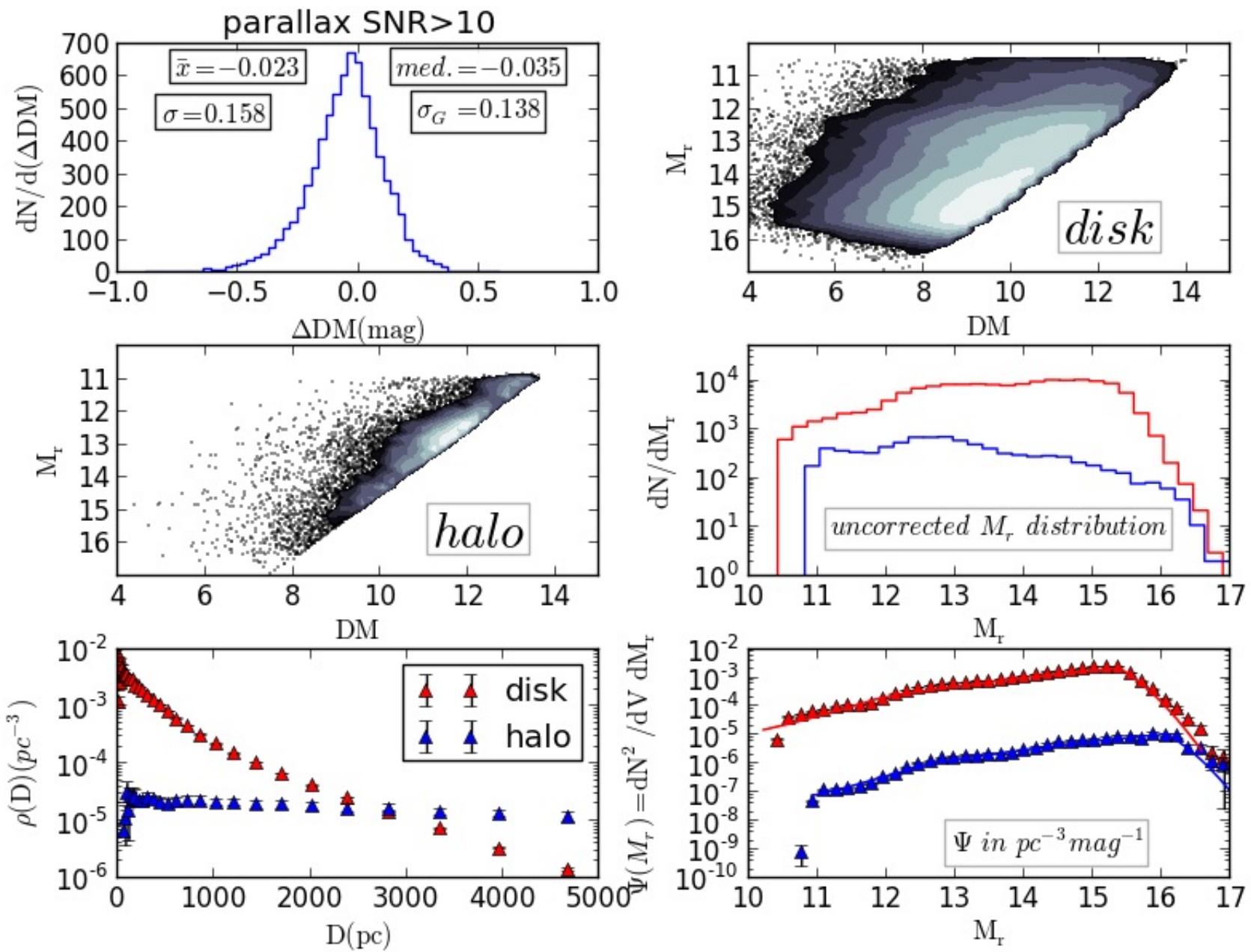
- In reality, the selection function is typically complex: $s(L, z|SED, \dots)$ (no sharp faint limit!)
- First define a generalized comparable set (Fan et al. 2001; AJ 121, 54) $J_i = \{j : L_j > L_i\}$; this is a luminosity limited data subset for object i
- Then generalize N_i to

$$T_i = \sum_{j=1}^{N_i} \frac{s(L_i, z_j|SED_j)}{s(L_j, z_j|SED_j)}, \quad (9)$$

and redefine the rank accordingly

$$R_i = \sum_{j=1}^{N_i} \frac{s(L_i, z_j|SED_j)}{s(L_j, z_j|SED_j)}, \quad (10)$$

for $z_j < z_i$. It follows that $E(R_i) = T_i/2$ and $V(R_i) = T_i^2/12$.



LF normalization

- The C⁻ method does not know (or need) details about our sample; in particular, it cannot give us the overall LF normalization!
- We will use HW#2 problem to discuss normalization in more detail; we can talk about three levels of normalization in this case:
 1. **The sample normalization:** if we didn't have the selection effects, how many objects would our sample contain?
 2. **Normalization to the full sky:** we need to know the sky coverage for our sample (and have arguments why we can extrapolate to the whole sky).
 3. **Extrapolation** from the volume probed by the sample **to** some other position; here, we want to know LF at $Z = 0$.

LF normalization: the sample normalization

If we didn't have the selection effects, how many objects would our sample contain?

- To recap, the cumulative luminosity (absolute magnitude) function is $\Phi_c(M_j)$ and the cumulative distance distribution is $n_c(D_j)$ where $j = 1 \dots N$.
- Both $\Phi_c(M_j)$ and $n_c(D_j)$ are direct outputs from the C⁻ method; let us renormalize them as $\Phi_c(M_N) = 1$ and $n_c(D_N) = 1$, where it is assumed that M_j and D_j arrays are sorted so that M_N and D_N are their maxima (btw, C⁻ would return $\Phi_c(M_N) = N$ and $n_c(D_N) = N$).
- The number of points, n , brighter than some arbitrary M^* and closer than D^* is then

$$n(M < M^* \text{ and } D < D^*) = C \Phi_c(M^*) n_c(D^*). \quad (11)$$

where we (still) don't know C (n.b. C is dimensionless).

LF normalization: the sample normalization

- Now, if make sure that M^* and D^* are **within our selection volume** (the implied apparent mag must be above our cutoff) and thus **unaffected by selection effects**, then we get C from

$$N^o(M < M^o \text{ and } D < D^o) = C \Phi_c(M^o) n_c(D^o), \quad (12)$$

which is almost the same expression as on the previous page, except that here $n(M < M^* \text{ and } D < D^*)$ is replaced by $N^o(M < M^o \text{ and } D < D^o)$: **the actual number of objects in our sample that satisfy this condition.**

- This is not mathematically optimal solution for C because N^o is a random variable, but with modern large samples this is nit-picking; the optimal procedure would integrate over the full sample, but nevertheless would still need to adopt an interpolation procedure for $\Phi_c(M)$ and $n_c(D)$...
- Given the real sample size, N , that is affected by selection effects, the “corrected” sample size is C !

LF normalization: the sample normalization

- The number of points per unit two-dimensional area, $dA = dM dD$, is then

$$\frac{d^2N}{dM dD} = C \left(\frac{d\Phi_c(M)}{dM} \right) \left(\frac{dn_c(D)}{dD} \right), \quad (13)$$

where we now know C and can easily take (numerical) derivatives $d\Phi_c(M)/dM$ and $dn_c(D)/dD$ (where $\Phi_c(M)$ and $n_c(D)$ came from C^- and are normalized to 1).

- The quantities in parenthesis are differential distribution functions.
- When normalizing to the full sky (step #2), we need to know the fraction of sky, f_{sky} , covered by our sample; if justified, we need to multiply C by $1/f_{sky}$: $C_{sky} = C/f_{sky}$.

LF normalization: extrapolation

How do we go from $d^2N/(dM dD)$ to volume density?

$$\Phi(M, D) \equiv \frac{d^2N}{dM dV} = \frac{d^2N}{dM dD} \frac{dD}{dV}, \quad (14)$$

where $dV = 4\pi D^2 dD$. We have two cases of interest:

- **Case 1:** We seek the volume density vs. D , $\rho(D)$, and we don't care about M distribution:

$$\rho(D) = \int \Phi(M, D) dM = \frac{C_{sky}}{4\pi D^2} \left(\frac{dn_c(D)}{dD} \right), \quad (15)$$

where we used the fact that $\int_{-\infty}^{\infty} \left(\frac{d\Phi_c(M)}{dM} \right) dM = \Phi_c(M_N) = 1$.

- Unit for $\rho(D)$ is the number of objects per (distance unit)³ (remember that $n_c(D)$ was dimensionless and normalized to unity at $D = D_N$; the unit comes from taking derivative with respect to D , $dn_c(D)/dD$).

LF normalization: extrapolation

- Given $\rho(D)$, we can fit some function to it and **extrapolate** to get $\rho(D = D_0)$ (and thus the ratio $\rho(D)/\rho(D = D_0)$ for any D_0 , including $\rho(D_0)/\rho(D_N)$).
- Case 2:** We want to know the M distribution at some $D = D_0$, call it $\psi(M|D = D_0)$ (e.g. $D_0 = 0$ corresponding to solar neighborhood, as discussed in this HW). First, at $D = D_N$ (recall $n_c(D_N) = 1$)

$$\psi(M|D = D_N) = \int \Phi(M, D)dD = C_{sky} \left(\frac{d\Phi_c(M)}{dM} \right). \quad (16)$$

- Then, extrapolating to D_0 (unit for ψ is the number of objects per mag; this is what we compare to the “true” LF)

$$\psi(M|D = D_0) = \psi(M|D = D_N) \frac{\rho(D_0)}{\rho(D_N)}. \quad (17)$$

Evolving Luminosity Function

- The C[–] method is simple and optimal, but it is valid only for uncorrelated variables (separable luminosity function). What do we do when the τ test suggests correlated variables?
- Recent work by Kelly, Fan and Vestergaard (2008, ApJ 682, 874) describes a powerful and completely general Bayesian approach (see their Appendix A for a nice introduction to Bayesian methodology). While too complex for homework, this is a fantastic method – if you ever come again across the problem of estimating a general multi-dimensional distribution that is sampled with non-negligible and possibly complex selection function, remember it!

Stellar Mass Function

- Analogously to luminosity function, the mass function is the distribution of mass of stars, galaxies, etc.
- The term *Stellar mass function* can refer to the distribution of galaxies with respect to their stellar mass (mass of all their stars), or to the distribution of mass of stars in the Milky Way!
- The distribution of mass of stars in the Milky Way is often parametrized by a power law, $dN/dM \propto M^{-\alpha}$, with $\alpha = 2.35$ (called Salpeter function in this context; FYI: power law is called the Pareto distribution in statistics...)
- Kroupa, Tout & Gilmore (1993, MNRAS 262, 545) proposed a three-part power law

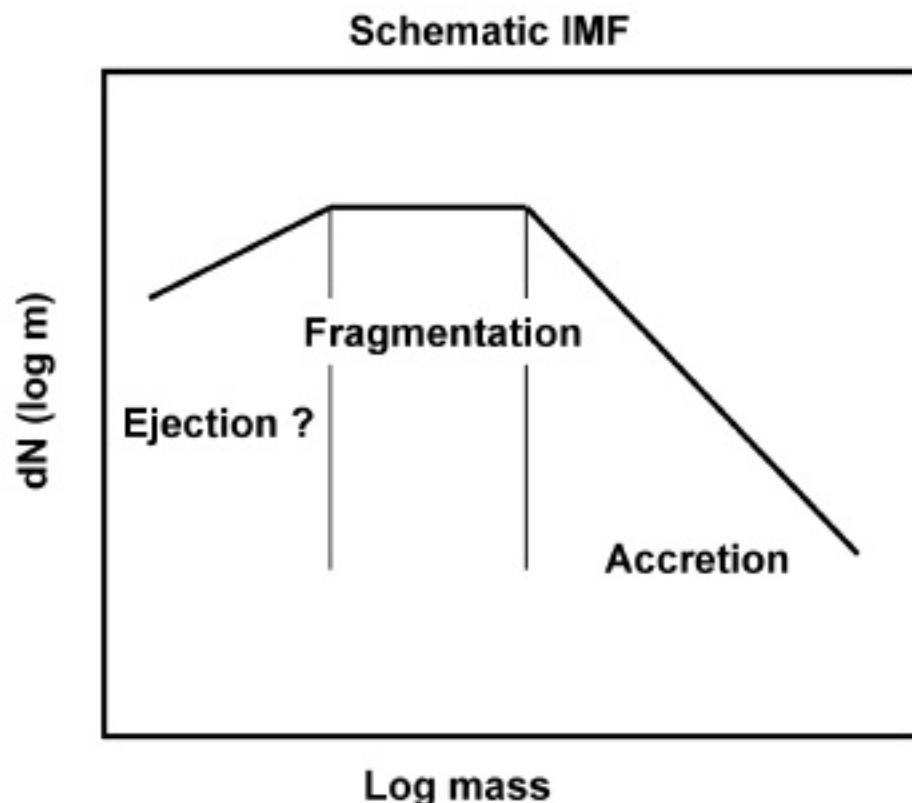


Fig. 11.— A schematic IMF showing the regions that are expected to be due to the individual processes. The peak of the IMF and the characteristic stellar mass are believed to be due to gravitational fragmentation, while lower mass stars are best understood as being due to fragmentation plus ejection or truncated accretion while higher-mass stars are understood as being due to accretion.

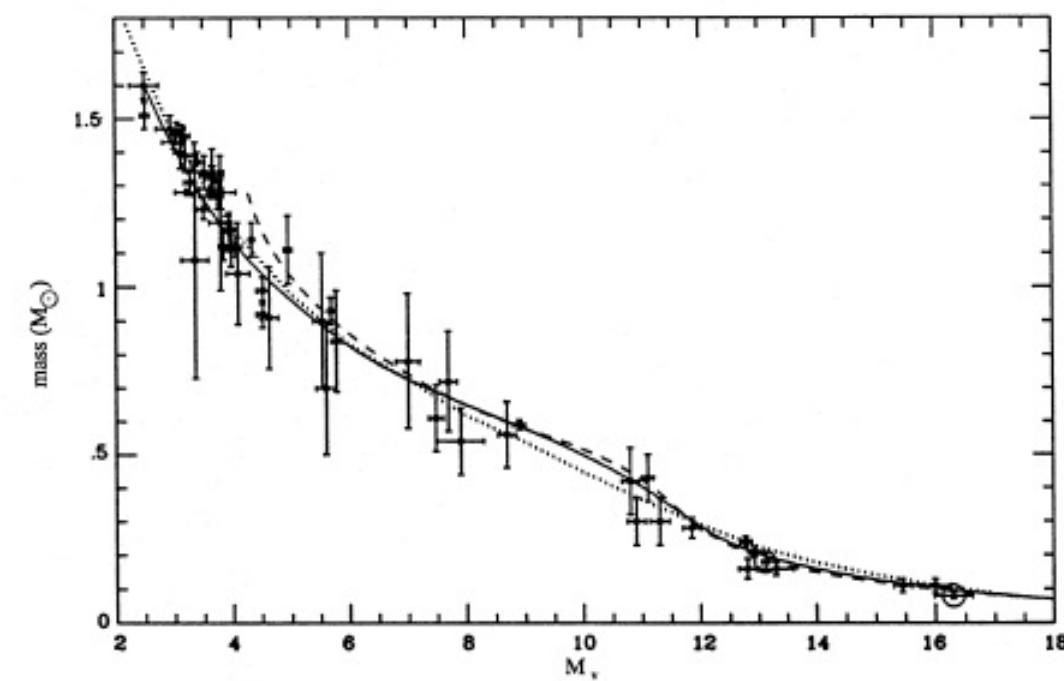
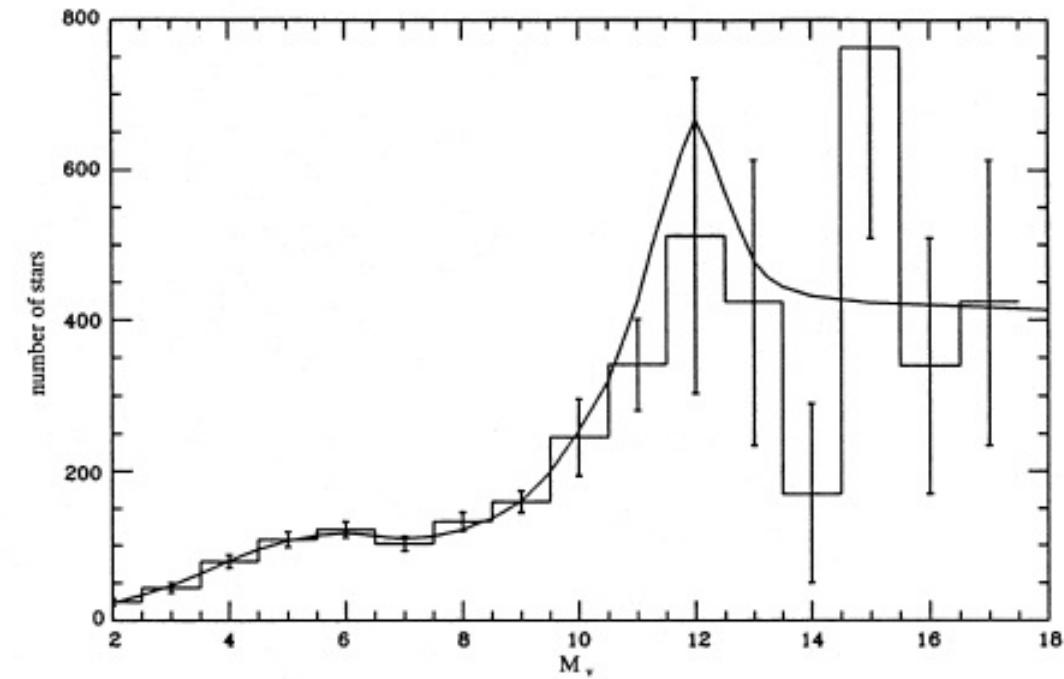
Initial Mass Function

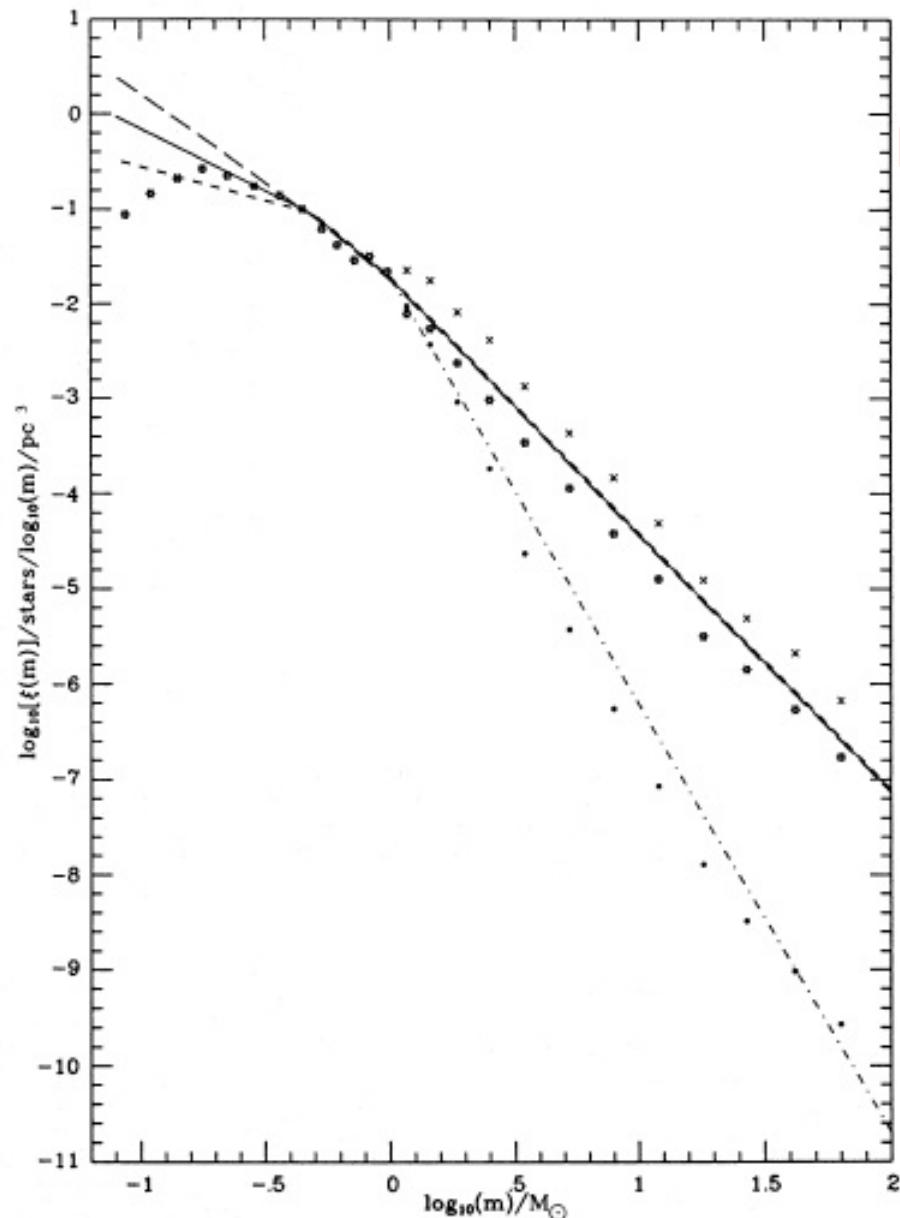
- An approximate understanding of the origin of different slopes.
- A hard problem to solve! (e.g. turbulence, magnetic fields...)

From W. Chen

Mass Function of Disk stars

- Determination of a three-part power law mass function by Kroupa, Tout & Gilmore (1993)
- Top: the measured number of stars per M_V bin
- Bottom: the mass-luminosity relation adopted in deriving the mass function





Mass Function of Disk stars

- Determination of a three-part power law mass function by Kroupa, Tout & Gilmore (1993)
- Present-day mass function (PDMF): dot-dashed line
- Initial mass function (IMF): solid line
- Note that the PDMF and IMF are equal below about 1 solar mass.

Figure 22. The stellar initial mass function (IMF) and present-day mass function (PDMF). The solid line represents the IMF given by equation (13), and the long- and short-dashed lines are for the cases $\alpha_1 = 1.85$ and 0.70 , respectively. The PDMF ($\alpha_3 = 4.5$, Section 2) is indicated by the dot-dashed line. At masses below about $1 M_\odot$ the PDMF equals the IMF. As a comparison, we show the PDMF derived by Scalo (1986) by solid dots. He corrects for stellar evolution; for a Galactic disc age of $T_0 = 9$ Gyr the IMF is indicated by stars, and for $T_0 = 12$ Gyr by crosses.

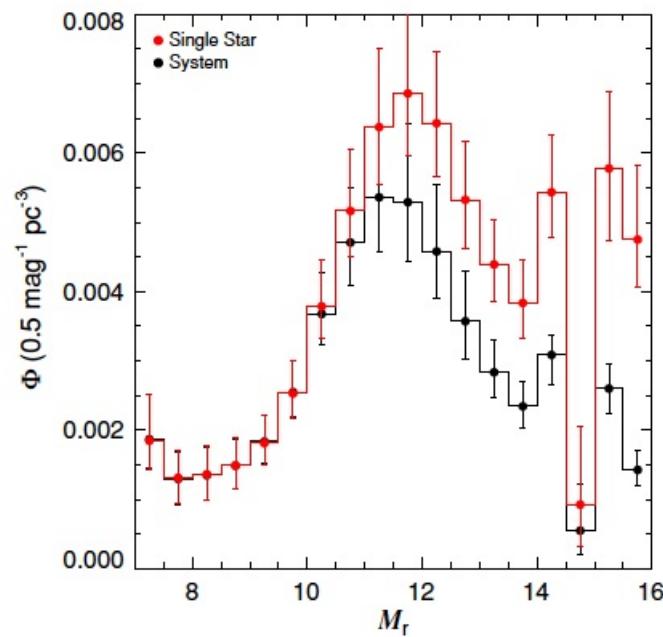


Figure 21. Single-star (red filled circles) and system (black filled circles) LFs. Note that the major differences between our system and single-star LFs occur at low luminosities, since low-mass stars can be companions to stars of any higher mass, including masses above those sampled here.

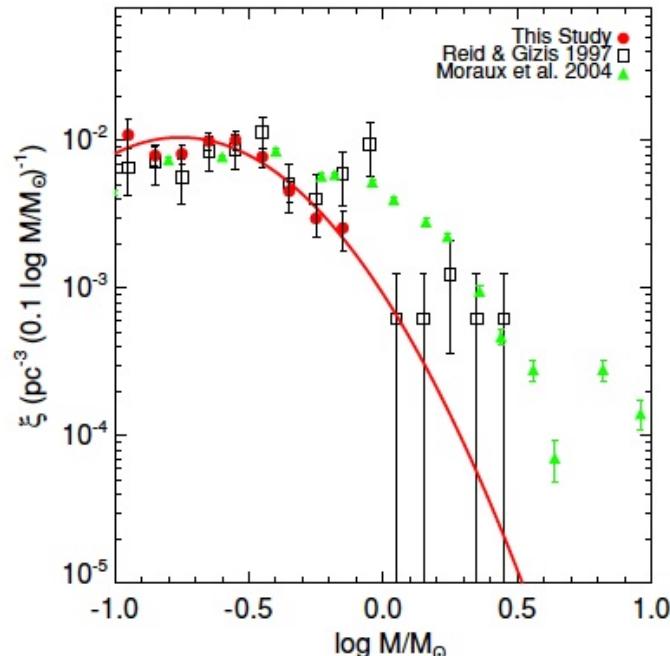
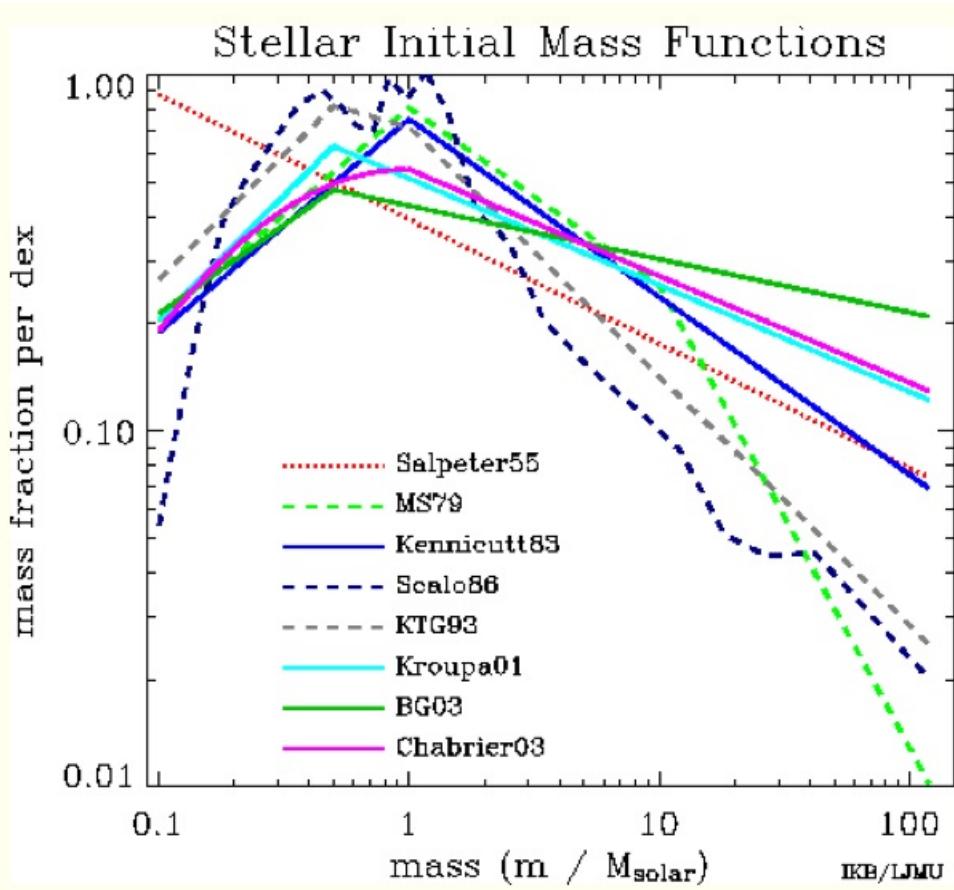


Figure 27. Shown are the single-star MF and best lognormal fit from this study (red filled circles and solid line), the Reid & Gizis (1997, open squares), MF (open squares), and the Pleiades MF Moraux et al. (2004, green triangles). The best fit extrapolated from our study systematically under-predicts the density at masses outside the bounds of our data.

Luminosity and Mass Function of Low-mass stars

- Bochanski et al. (2010, AJ 139, 2679)
- Based on SDSS data for 15 million low-mass stars!
- The mass range: 0.1–0.8 solar masses (corresponding to $7 < M_r < 16$)
- The turn-over and a local maximum well detected!
- Data well described by a log-normal distribution (over the probed mass range).



Initial Mass Function

- The stellar initial mass function (IMF) is used for computing stellar masses and colors of galaxies in cosmology.
- There is substantial variation between different estimates (left).
- Kroupa (2001) claimed a variable IMF (MNRAS 322, 231).

From Ivan Baldry:

<http://www.astro.ljmu.ac.uk/~ikb/research/imf-use-in-cosmology.html>