

天体物理研究概论 《二》

星系的统计性质

沈世银

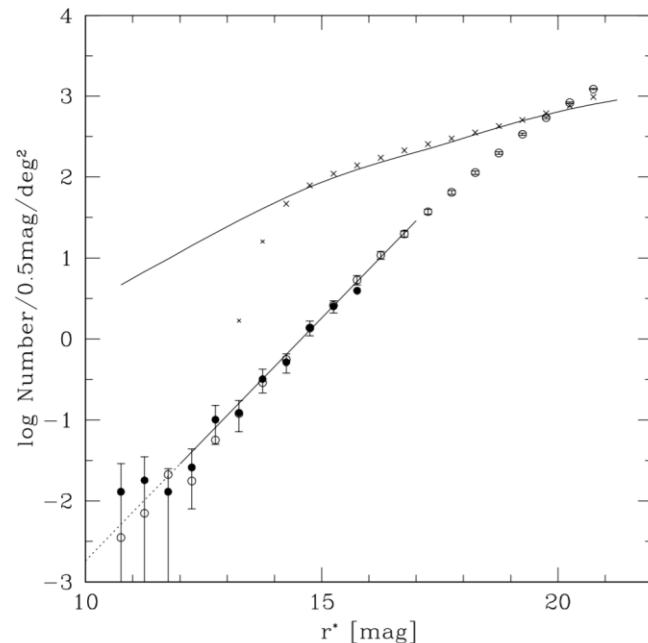
目录

- 星系性质的分布函数
 - 星等、光度、质量
 - 不同波段
 - 不同质量成份
 - 空间分布
- 条件分布函数
- 标度律 (scaling relation)

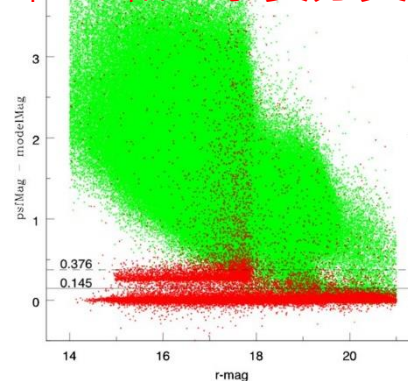
(一) 星系性质的分布函数

1、星系的数目分布 (number counts)

- $N(m)$: 单位立体角里面单位星等星系的数目
- 如果欧几里得空间, 星系的数密度是常数
 - $\text{Log } N(m) = 0.6 m + \text{const.}$
 - 奥博斯佯谬
- 星系计数
 - 星系光度函数及演化
 - 宇宙学
 - 银河系尘埃消光



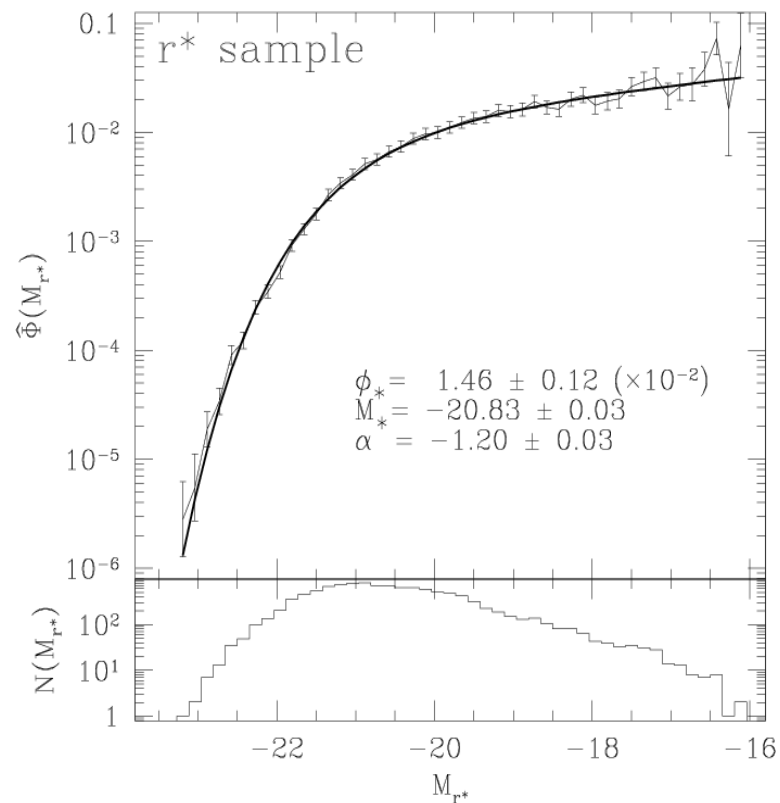
恒星和星系要分类



2、星系的光度函数 (luminosity function)

• 星系的光度函数

- 最基本的分布函数
- 单位体积单位光度 (星等) 范围内星系的数目
 - 暗星系多, 亮星系少
- 直接观测到的星系数目:
亮星系多, 暗星系少

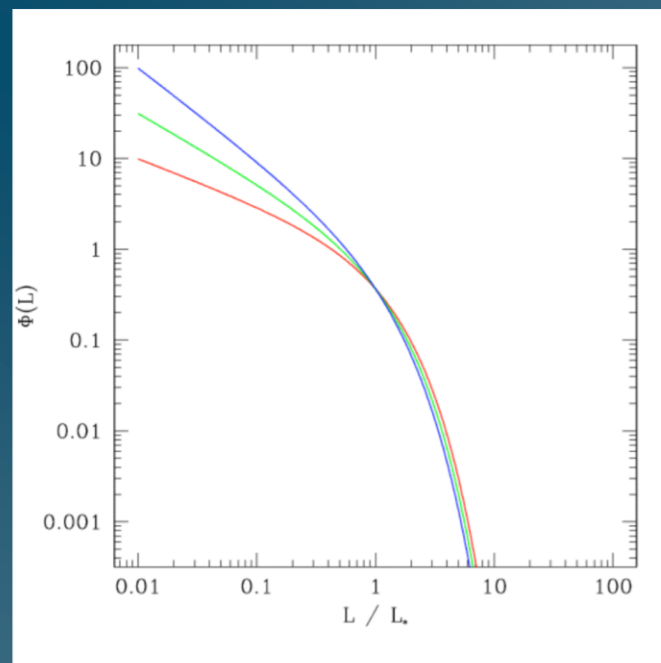


Schechter 函数

$$\Phi(L)L = \Phi_* \left(\frac{L}{L_*}\right)^\alpha \exp(-L/L_*) \frac{L}{L_*}$$

- Φ_* (光度密度)
- α (暗段斜率)
- L_* (特征光度)

$\alpha = -0.5$ (red), $\alpha = -0.75$ (green), $\alpha = -1$ (blue):



• 星系的数密度

- 如果 $\alpha < -1$, n 发散

$$n = \int_L^\infty \Phi(L) dL$$

• 星系的光度密度

- 由 L_* 星系主导

$$\int_L^\infty \Phi(L) L dL$$

$$\phi(M) = \frac{\ln 10}{2.5} \phi_* \left(10^{0.4(M^* - M)}\right)^{(\alpha+1)} \exp\left[-10^{0.4(M^* - M)}\right]$$

光度函数的计算

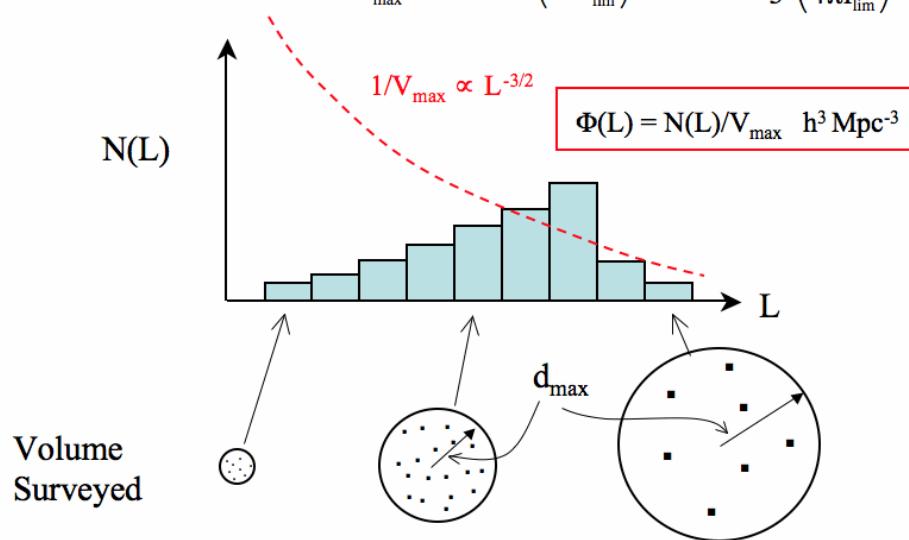
- 体积完备样本

- 星等完备样本

- SDSS主星系样本 $r < 17.77$
- 亮星系更容易被观测到
- V_{\max} :
 - 星系在星等完备样本中能观测到的最大体积范围
 - $\langle V/V_{\max} \rangle \sim 0.5$

$1/V_{\max}$ corrections for Malmquist bias

Flux limit f_{lim} $f_{\text{lim}} = \frac{L}{4\pi d_{\text{max}}^2}$ $d_{\text{max}} = \left(\frac{L}{4\pi f_{\text{lim}}}\right)^{1/2}$ $V_{\text{max}} = \frac{4\pi}{3} \left(\frac{L}{4\pi f_{\text{lim}}}\right)^{3/2}$



最大似然法

- [arXiv:astro-ph/9707246v1](https://arxiv.org/abs/astro-ph/9707246v1) 22 Jul 1997
- Sandage, Tammann & Yahil (1979): STY

$$p(M_i, z_i) = \frac{\phi(M_i)}{\int_{-\infty}^{M_{faint}(z_i)} \phi(M) dM}. \quad \mathcal{L} = \prod_{i=1}^{N_g} p(M_i, z_i).$$

- stepwise maximum likelihood method (SWML)
 - Efstathiou, Ellis & Peterson (1988, EEP)
 - 非参数化的光度函数形式

$$\phi(M) = \phi_k, \quad M_k - \Delta M/2 < M < M_k + \Delta M/2, \quad k = 1, \dots, N_p.$$

光度函数计算中的细节

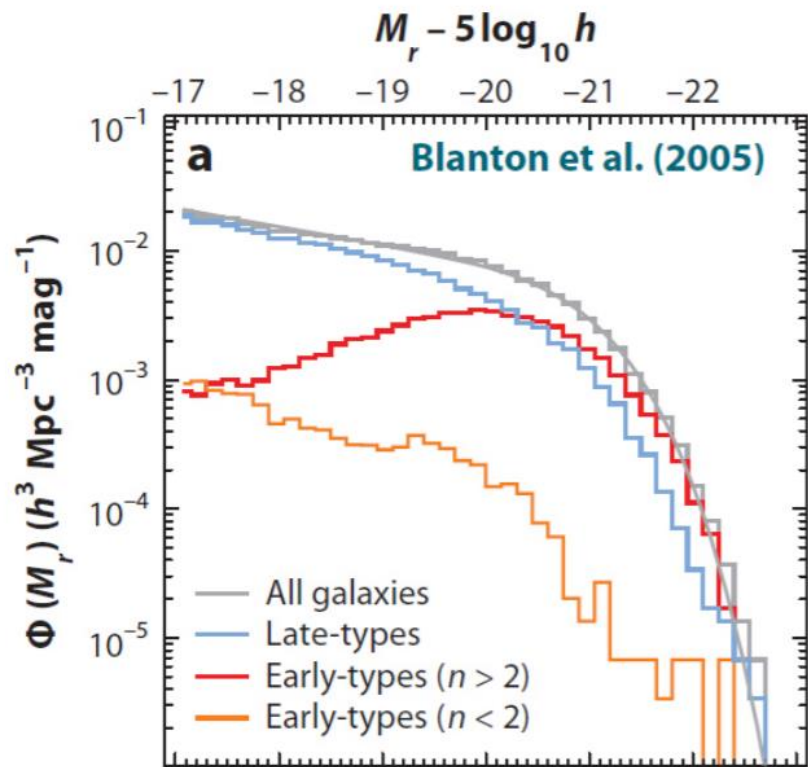
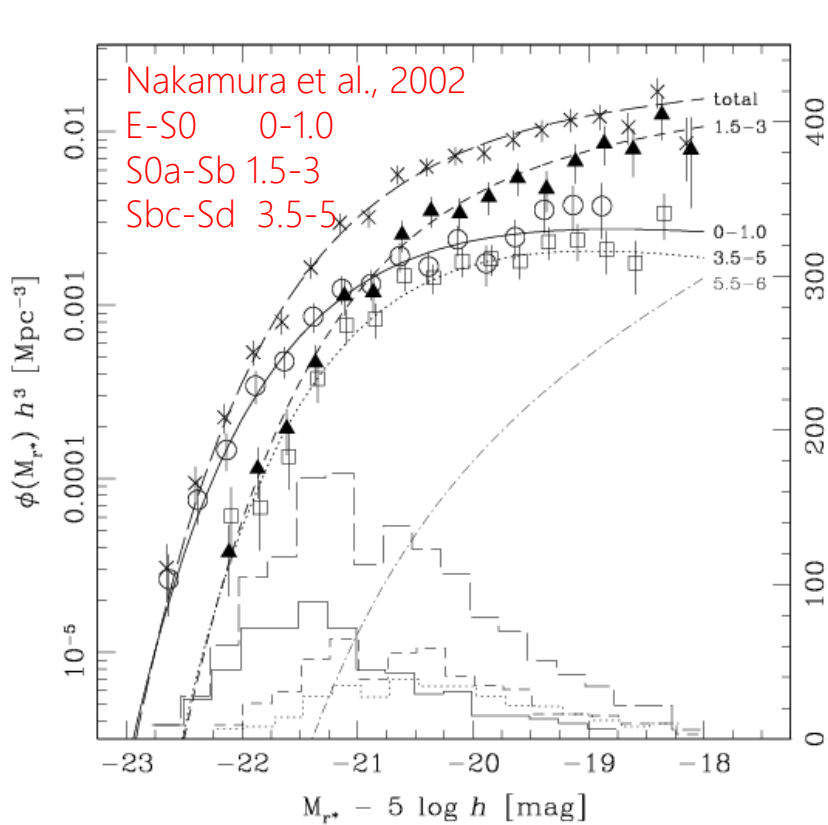
- K改正
- 演化效应改正
- 完备性改正
 - 低表面亮度星系

不同波段的光度函数

Table 15.1. Galaxy luminosity functions in different bands.

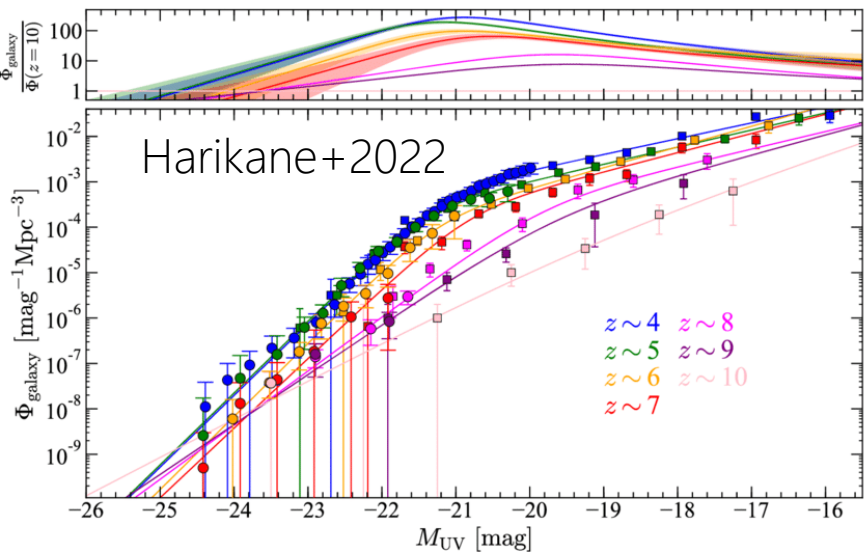
Band	α	$\mathcal{M}^* - 5 \log h$	$\phi^*/(h^3 \text{ Mpc}^{-3})$	Reference
<i>u</i>	-0.92 ± 0.07	-17.93 ± 0.03	$(3.05 \pm 0.33) \times 10^{-2}$	Blanton et al. (2003)
<i>g</i>	-0.89 ± 0.03	-19.39 ± 0.02	$(2.18 \pm 0.08) \times 10^{-2}$	Blanton et al. (2003)
<i>r</i>	-1.05 ± 0.01	-20.44 ± 0.01	$(1.49 \pm 0.04) \times 10^{-2}$	Blanton et al. (2003)
<i>i</i>	-1.00 ± 0.02	-20.82 ± 0.02	$(1.47 \pm 0.04) \times 10^{-2}$	Blanton et al. (2003)
<i>z</i>	-1.08 ± 0.02	-21.18 ± 0.02	$(1.35 \pm 0.04) \times 10^{-2}$	Blanton et al. (2003)
<i>J</i>	-1.10 ± 0.04	-22.85 ± 0.04	$(0.71 \pm 0.01) \times 10^{-2}$	Jones et al. (2006)
<i>H</i>	-1.11 ± 0.04	-23.54 ± 0.04	$(0.72 \pm 0.01) \times 10^{-2}$	Jones et al. (2006)
<i>K</i>	-1.16 ± 0.04	-23.83 ± 0.03	$(0.75 \pm 0.01) \times 10^{-2}$	Jones et al. (2006)
<i>b_J</i>	-1.21 ± 0.03	-19.66 ± 0.07	$(1.61 \pm 0.08) \times 10^{-2}$	Norberg et al. (2002b)
Near-UV	-1.12 ± 0.10	-17.77 ± 0.15	$(1.22 \pm 0.20) \times 10^{-2}$	Budavári et al. (2005)
Far-UV	-1.10 ± 0.12	-17.20 ± 0.14	$(1.30 \pm 0.20) \times 10^{-2}$	Budavári et al. (2005)

子样本的光度函数

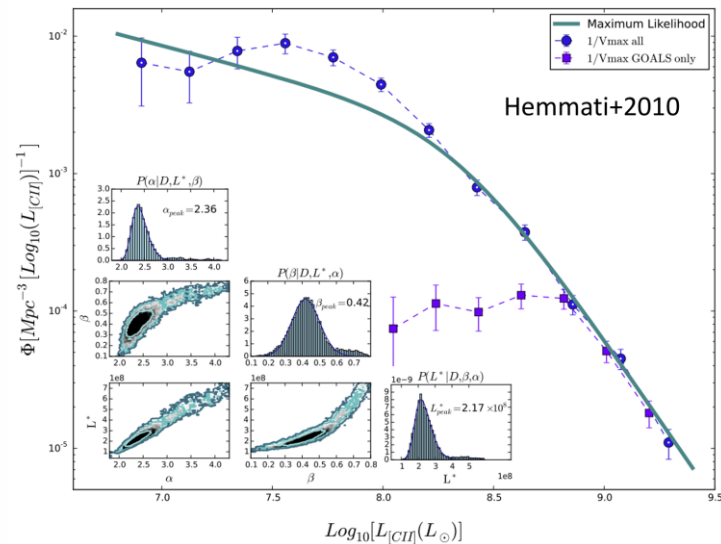


不同形态星系的光度函数

- 不同红移的光度函数

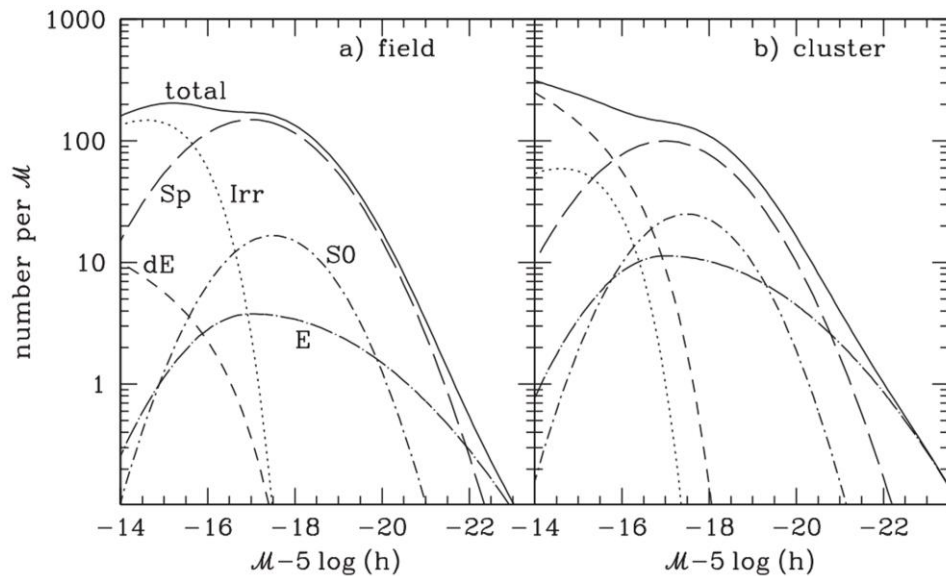
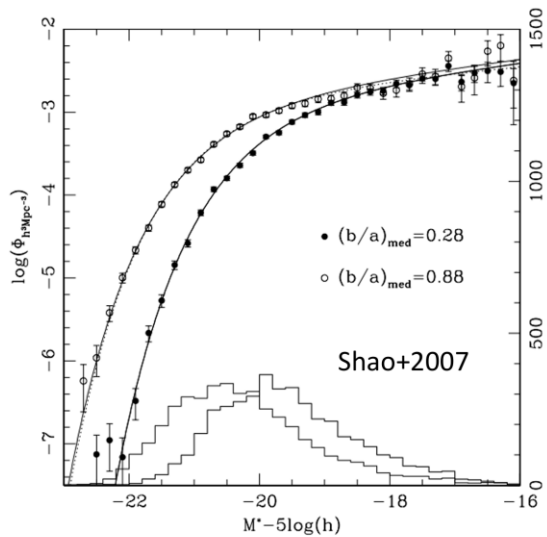


- 发射线流量的光度函数



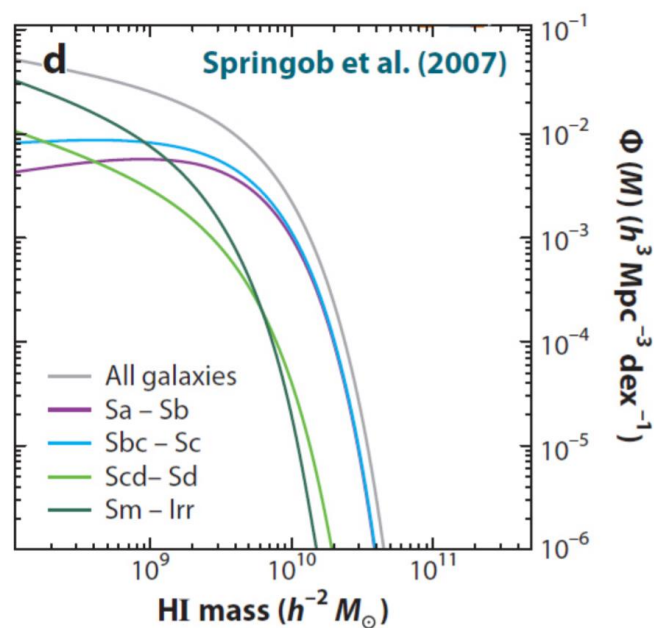
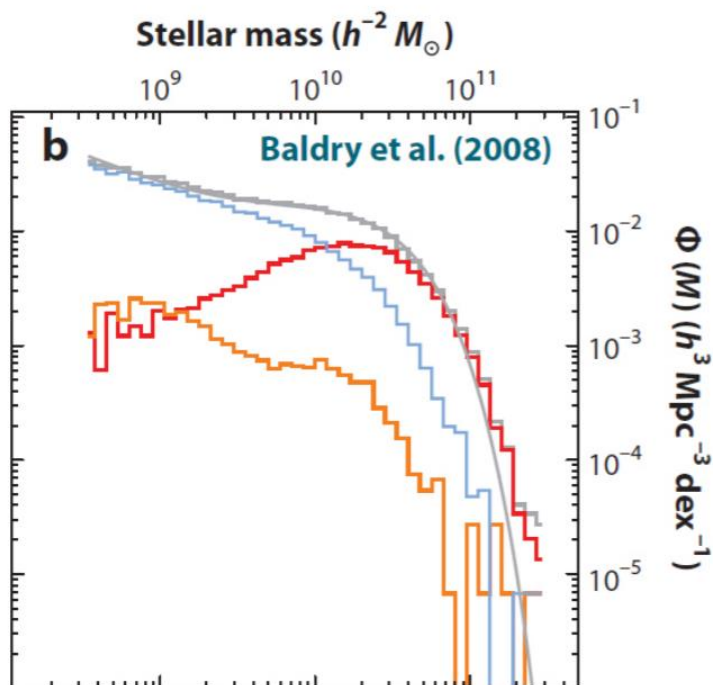
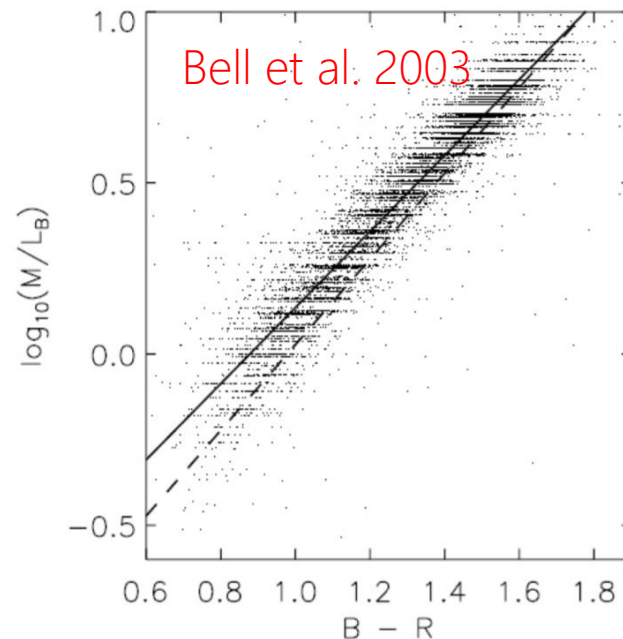
- 光度函数与环境

- 不同倾角旋涡星系的光度函数



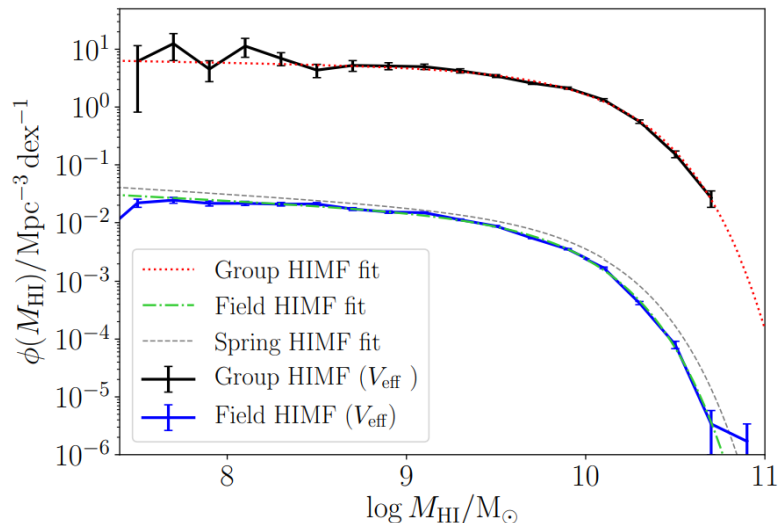
2.1 质量函数

- $\Phi(M)$
 - 恒星质量, 气体质量等
 - 质光比: 从颜色推算
 - 星族, IMF

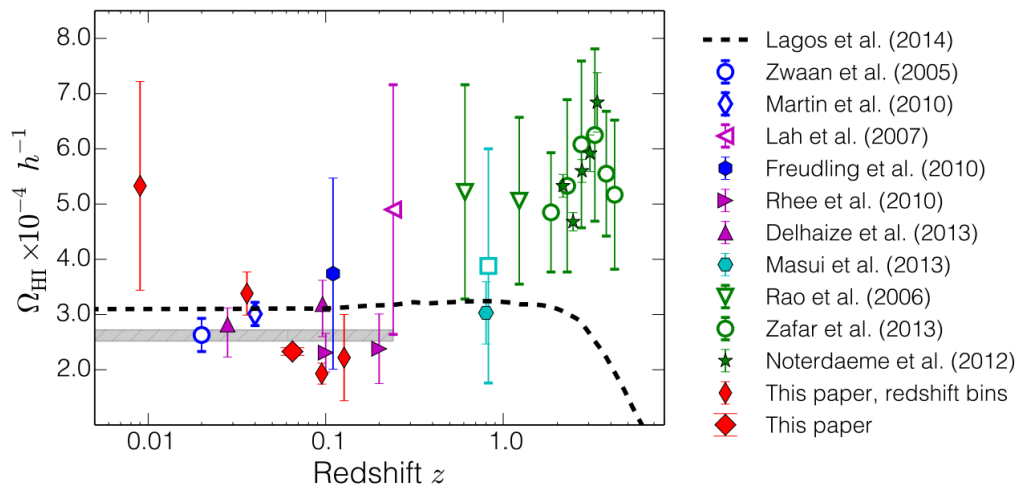


HI mass function

Jones et al. 2020

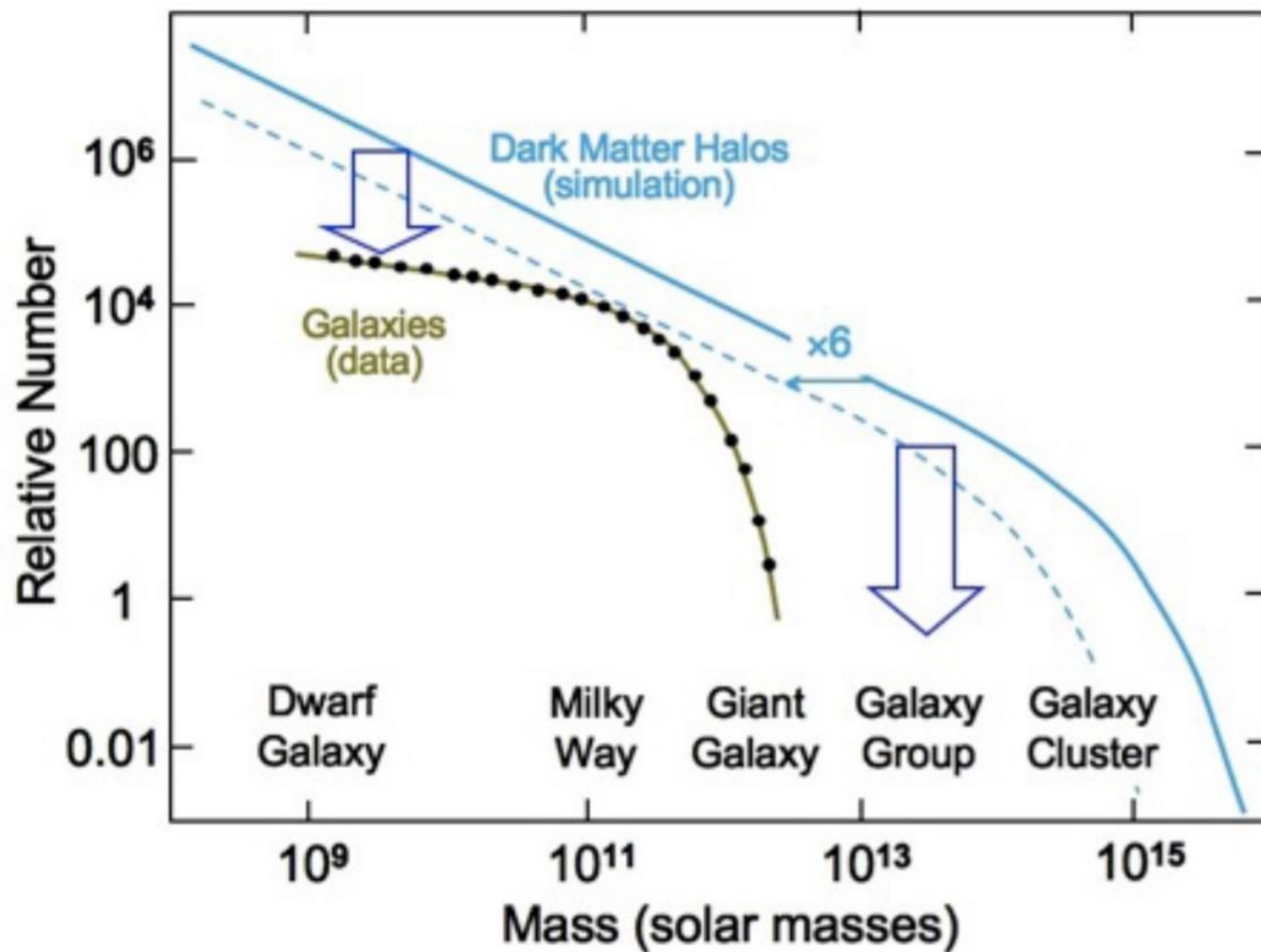


Hoppmann et al. 2015



结合HI和恒星质量函数可以更好的理解星系中的恒星形成过程

Halo and Galaxy Mass Distributions



作业1

- 单位平方度内，有多少 $r < 20.5$ mag 的星系？
- SDSS给出的本地星系的Schechter光度函数的参数为 $M_r^* = -20.83$, $\alpha = -1.0$, 请问单位体积范围内 $M_r = -16$ mag的星系数目是 $M_r = -22$ mag星系数目的多少倍？SDSS是星等极限样本 ($r < 17.77$) ,请问在SDSS的观测样本中, $M_r = -22$ mag的星系数目是 $M_r = -16$ mag星系数目的多少倍？
 - 根据以上光度函数，画出一个 $M_r = 20$ mag星系的红移分布概率（不考虑星系的K改正和演化效应）

3. 星系的空间分布

<http://arxiv.org/abs/1202.6633>

- 两点相关函数

- 三维 $dP_{12} = \bar{n}^2 \{1 + \xi_{12}(r)\} dV_1$

- 二维 $dP = N[1 + \omega(\theta)]d\Omega$

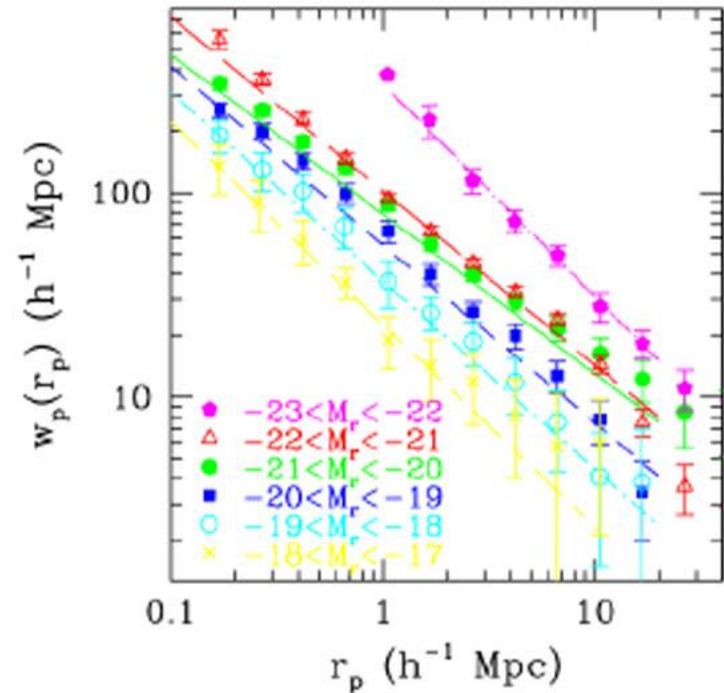
$$\hat{\xi}_{DP}(r) = \frac{N_R}{N_D} \frac{DD(r)}{DR(r)} - 1 \quad \text{Davis and Peebles (1983)}$$

- 幂率函数

$$\xi(r) = \left(\frac{r}{r_0}\right)^{-\gamma}$$

- r_0 : 相关长度
 - 参数和样本 (光度/质量) 相关

- 自相关, 交相关



红移畸变 (redshift distortion)

v_1 and v_2 are defined to be the redshift positions of a pair of galaxies

- $s=(v_1 - v_2)$ to be the redshift space separation
- $l = 1/2 (v_1 + v_2)$ to be the mean distance to the pair

Separation between two galaxies across (r_p) and along (π) the line of sight

$$\pi = \frac{s \cdot l}{|l|},$$
$$r_p = \sqrt{s \cdot s - \pi^2}.$$

• 小尺度

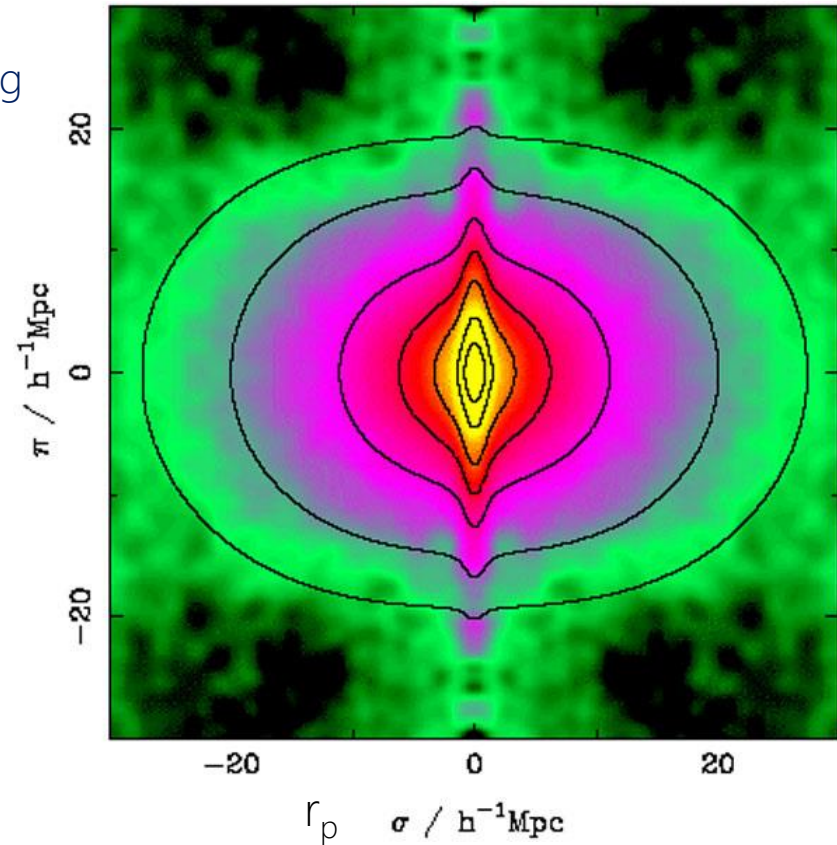
- Finger of God

星系团中星系的随机运动

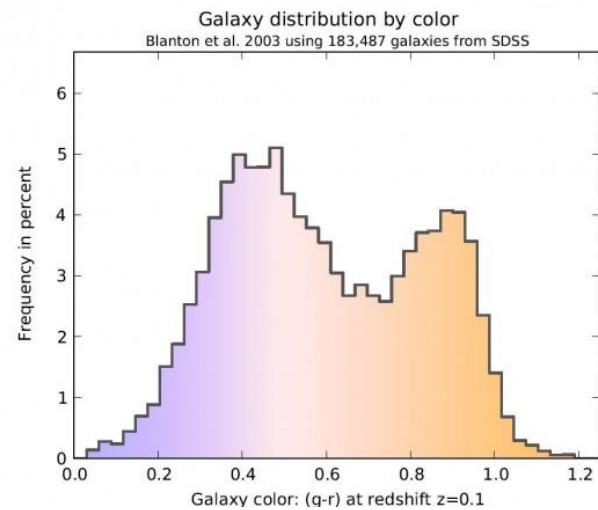
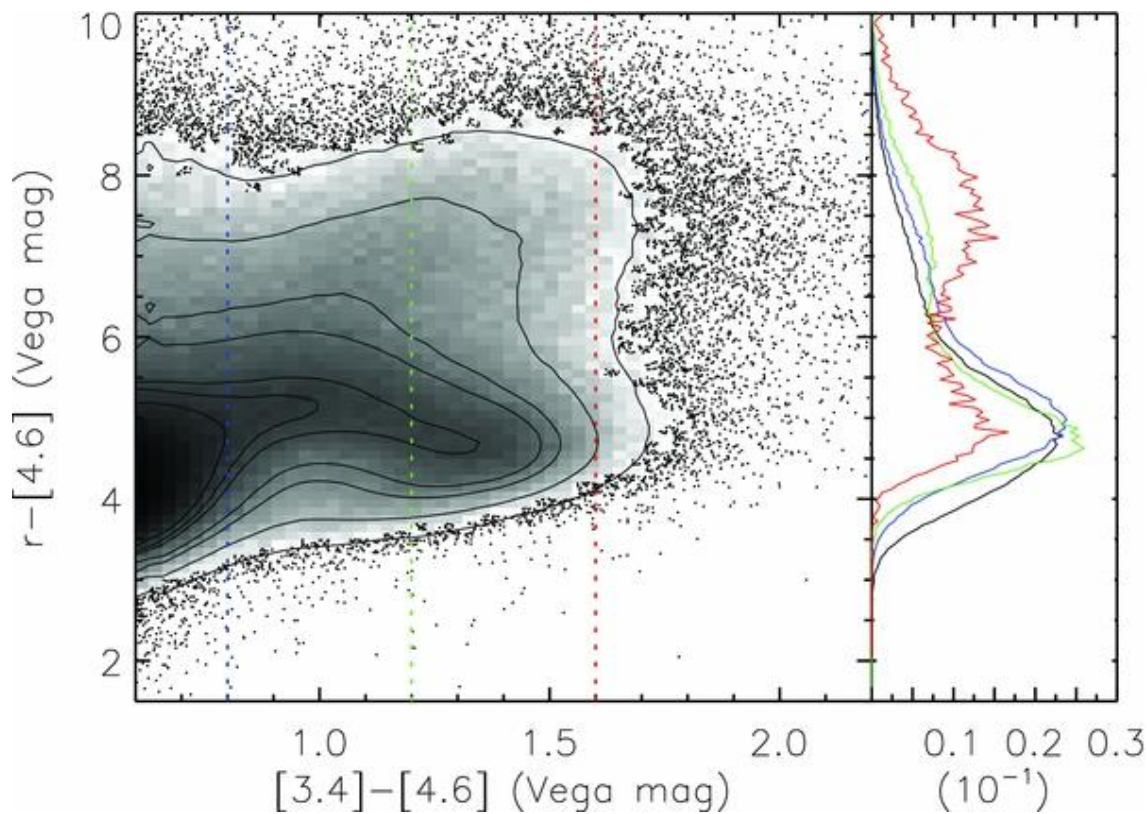
• 大尺度

- Kaiser effect

大尺度上星系的引力成团效应



4、星系的颜色分布



为什么是双峰分布?

(二) 星系的条件分布函数

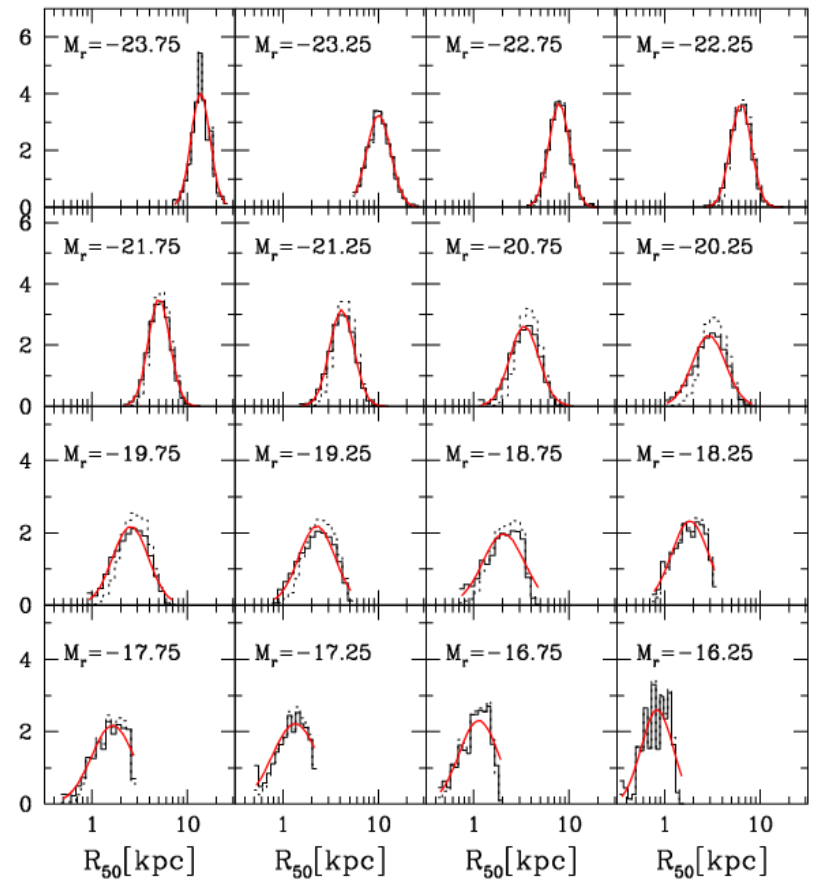
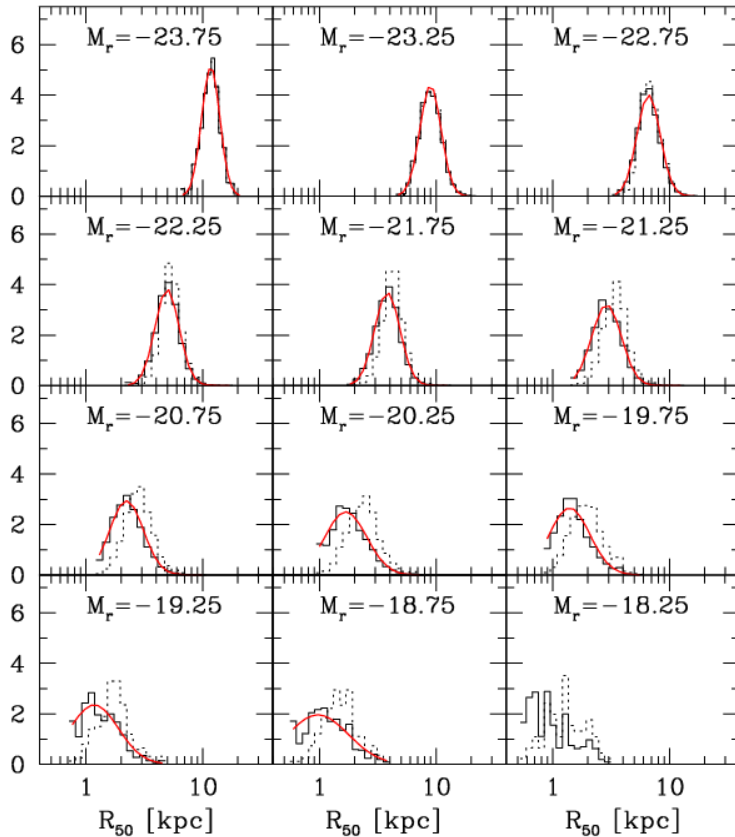
- 第一参量：质量
 - 可以是恒星质量（观测研究）
 - 也可以是暗晕质量（模型研究）

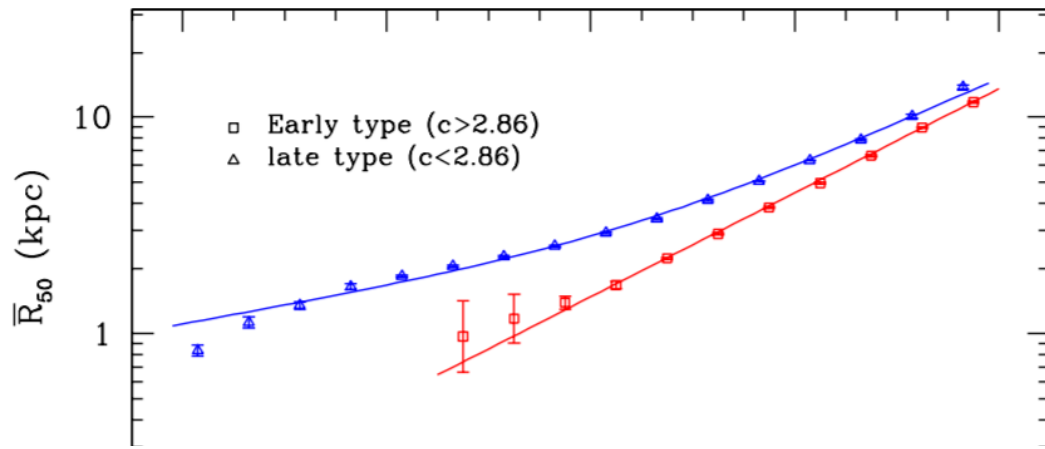
- 其它参量
 - $f(R|M)$
 - $f(\text{color}|M)$
 -

星系尺度的分布函数 (Shen et al. 2003)

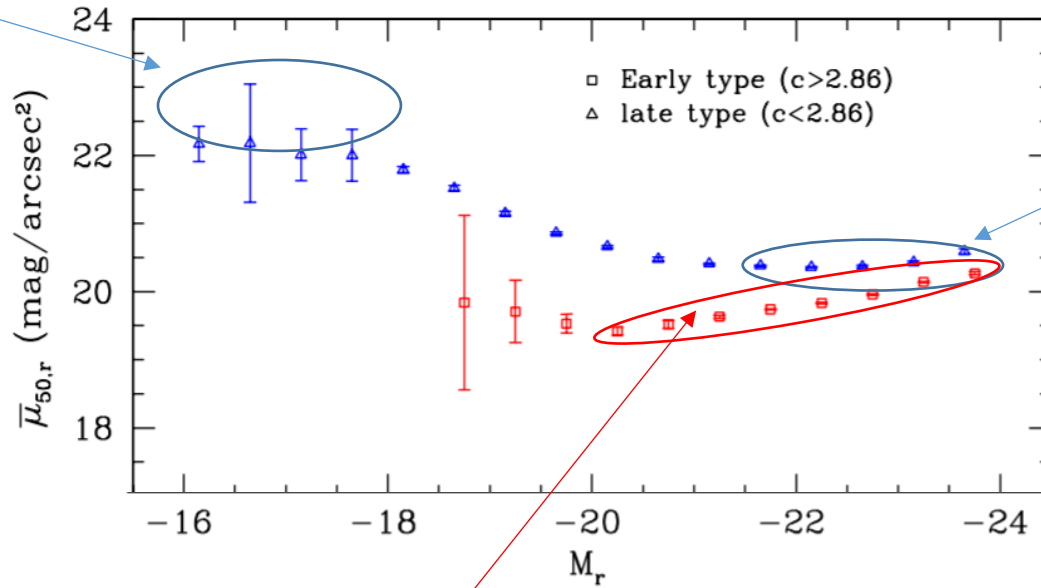
Early type $c > 2.86$

Late type $c < 2.86$





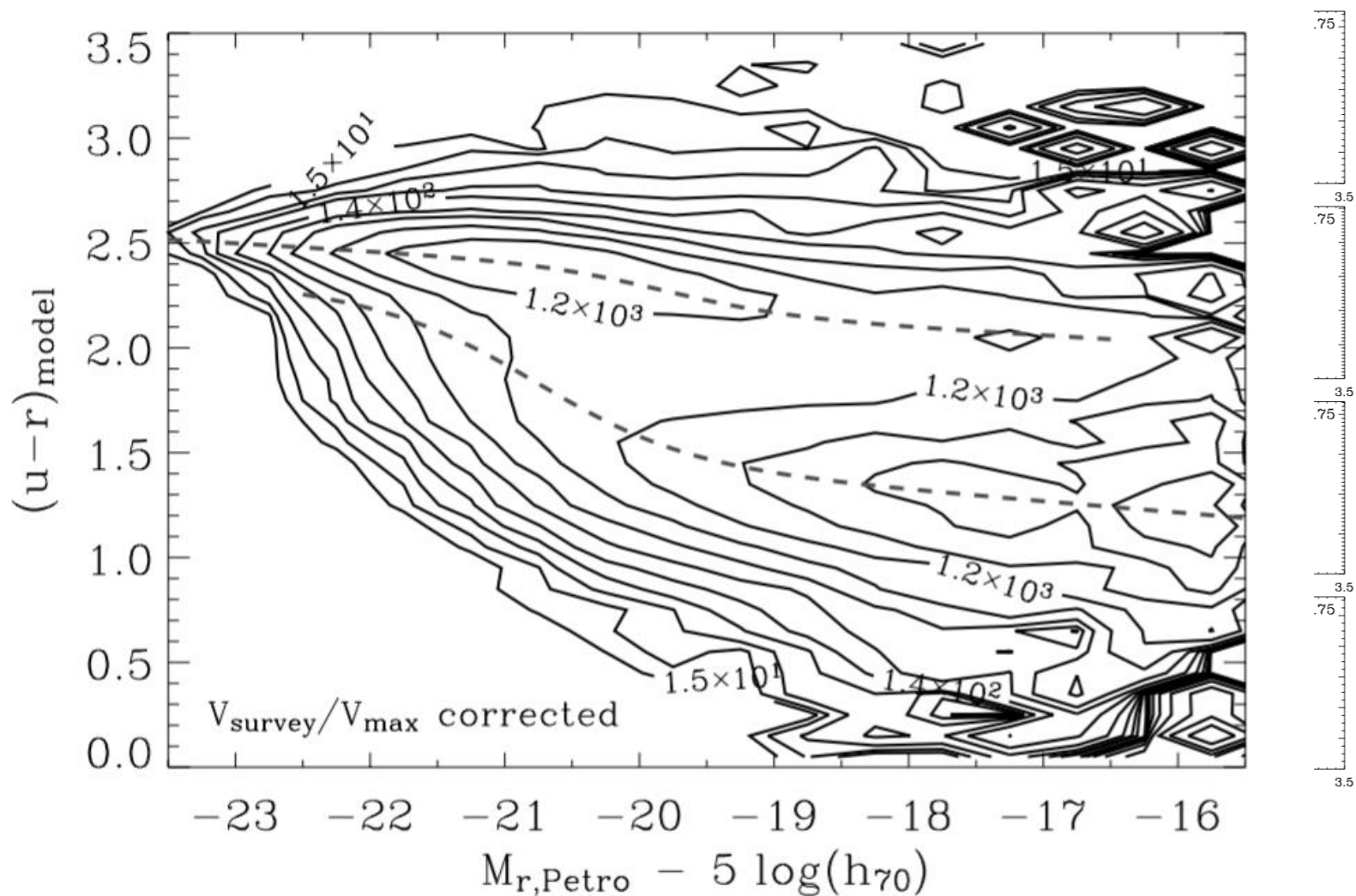
Low surface brightness galaxies

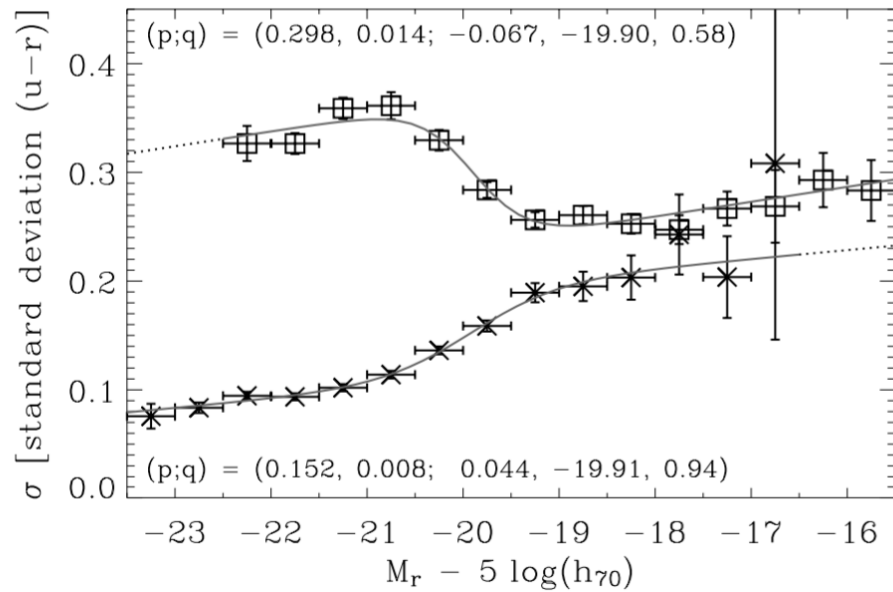
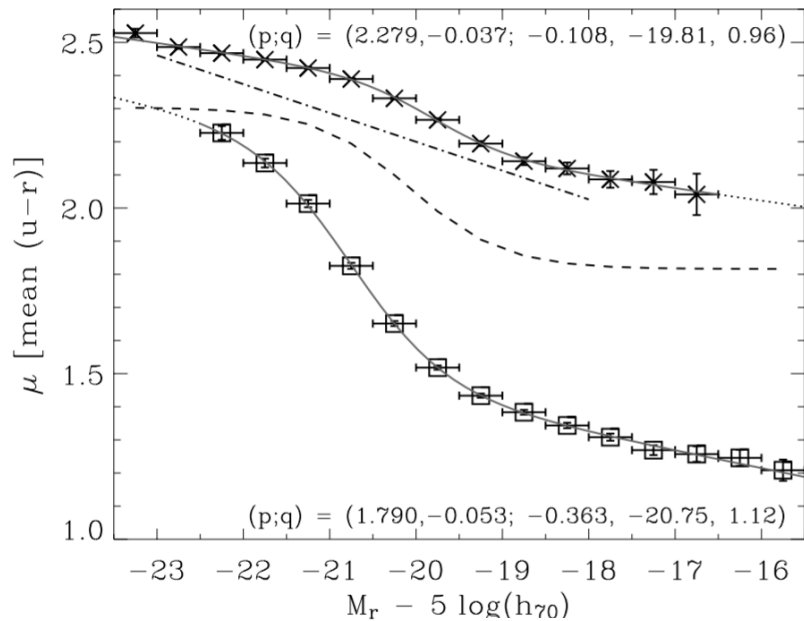


Freeman law 1970

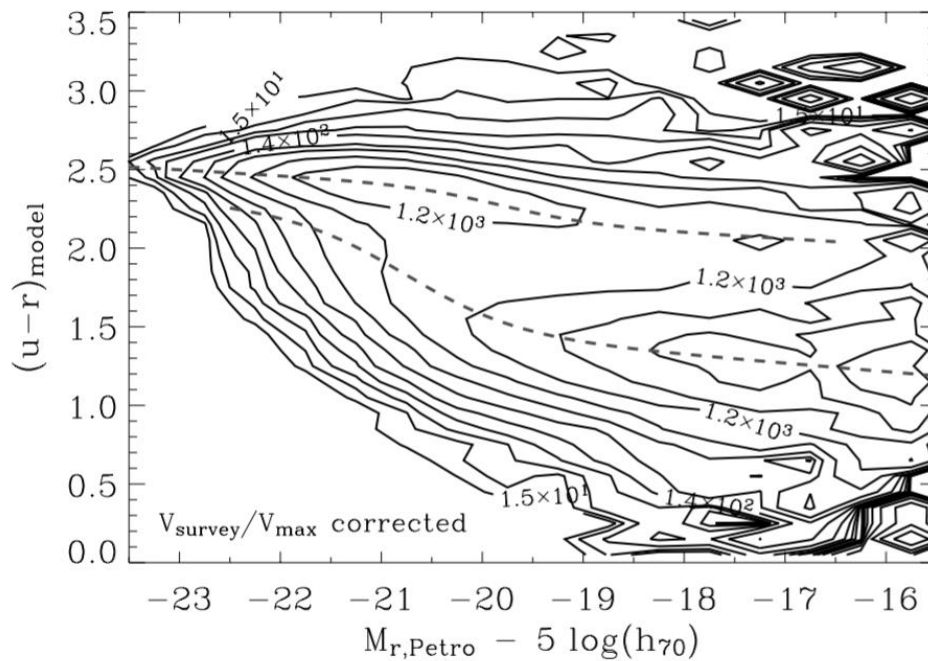
Kormendy relation 1977

星等颜色分布(Baldry et al. 2004)



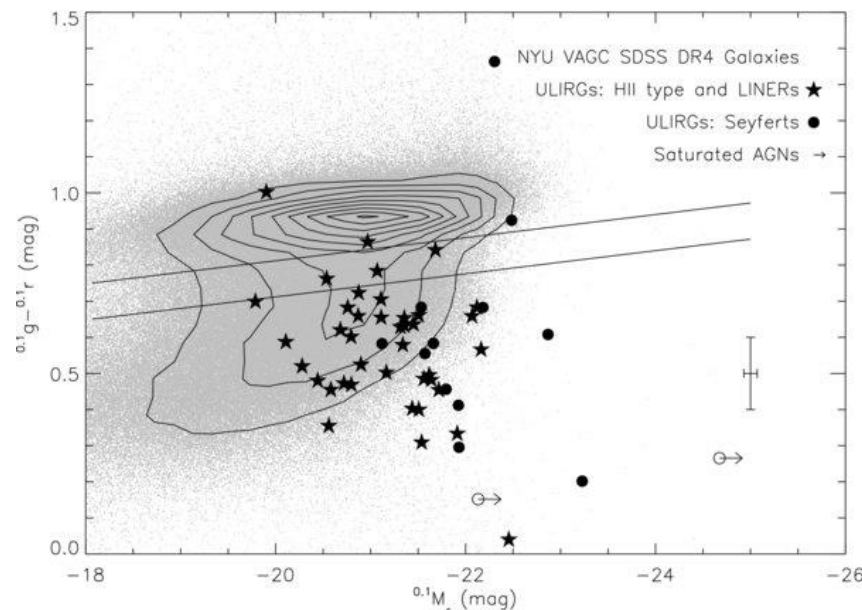
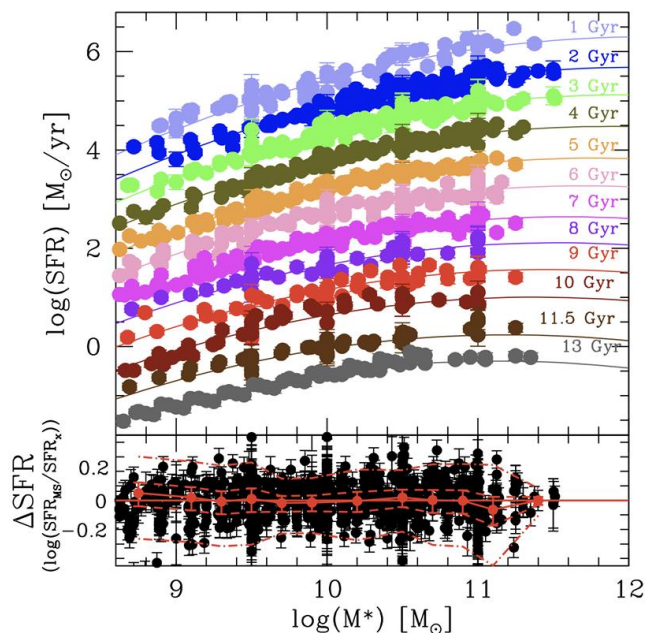


Baldry et al. 2004



颜色-星等关系

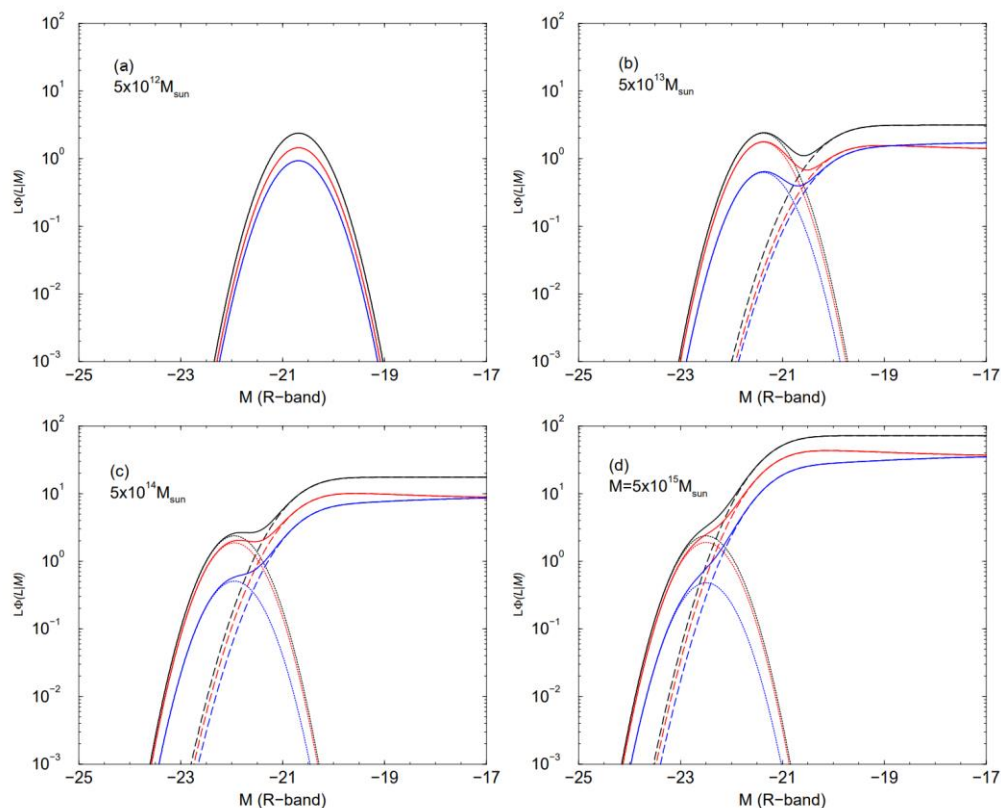
- sSFR-质量 (光度) 关系
- SFR-质量 (光度) 关系
 - Main sequence



基于暗晕质量的Conditional luminosity function model (Yang et al., 2003, 2005)

- 暗晕 → 星系群
 - 主暗晕, 子暗晕
 - 中央星系, 卫星星系
 - 中央星系不一定是**最亮星系**
- Halo Occupation Distribution (HOD)
- 条件分布函数和两点相关函数结合
 - 从统计上完全Mock星系的分布

Cooray 2005



(三) 标度率 (scaling relation)

- 定理

- $F=ma$

- 量纲不一致的相关关系

- 弥散

- 观测误差
 - 内禀弥散

3.1 动力学标度关系

- 星系中各种组分在质量上的关系
 - Tully-Fisher关系
 - Faber-Jackson关系
 - Fundamental plane
 - M-sigma 关系
- 注意：大城市里面酒吧多，书店也多
 - 相关性 不等于 因果性

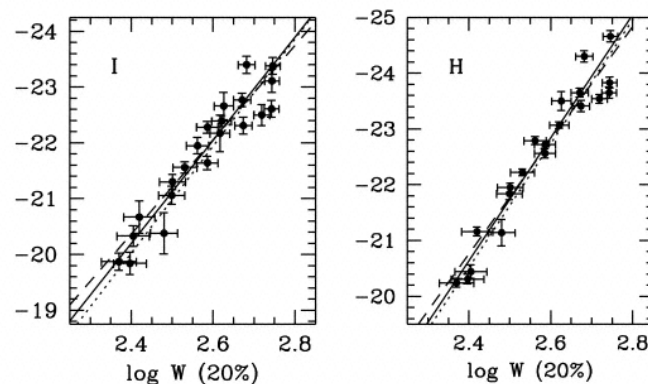
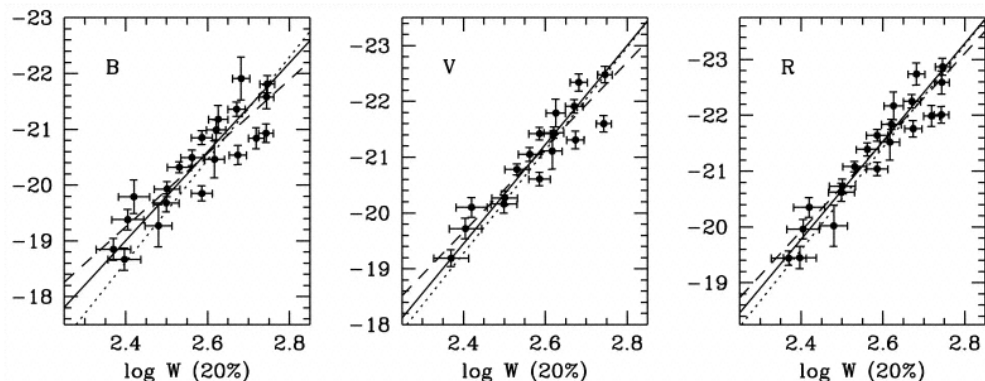
旋涡星系的 Tully-Fisher关系

- $M = a \log W + b$
 - $L = a * V^\alpha$ ($\alpha \approx 3-4$)
 - 不同波段斜率不同
 - 弥散: ~ 0.2 个星等
 - M 是重子质量, 弥散更小

- 测定星系的距离
 - 速度的测量不依赖于距离

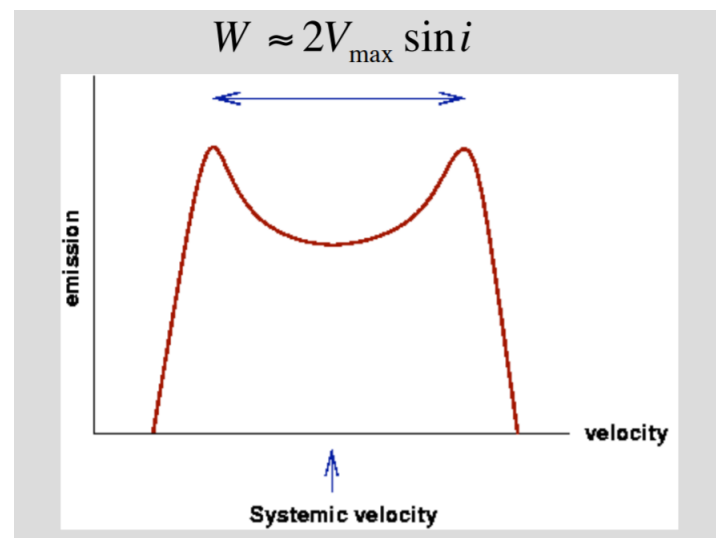
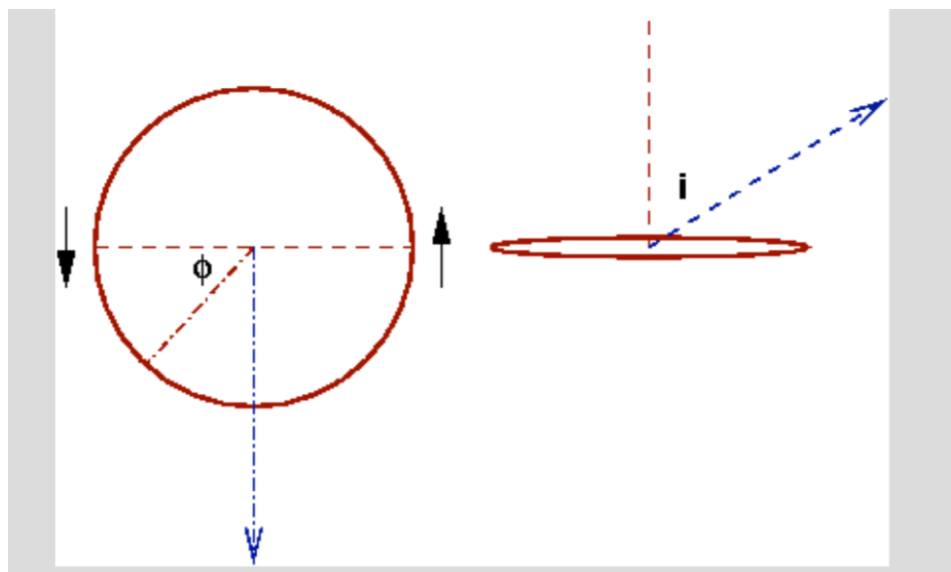
$$L_H \approx 3 \times 10^{10} \left(\frac{V_{\max}}{196 \text{ km/s}} \right)^{3.8} L_{H, \text{solar}}$$

*BVRIH*_{0.5} Tully-Fisher relations for spiral galaxies with Cepheid distances, using 20% line width. Solid lines represent the bivariate fits, while the dotted and dashed lines represent inverse and direct fits, respectively.



V的测量：谱线展宽

$$V_r(R, i) = V_{sys} + V(R) \sin i \cos \phi$$



The "Physics" of Tully-Fisher

gravity: $V^2 = \frac{GM}{R} \Rightarrow M \sim RV^2$

mass-to-light ratio: $M = L \left(\frac{M}{L} \right)$

surface brightness: $\Sigma_{SB} = \frac{L}{\text{area}} \sim \frac{L}{R^2} \Rightarrow L \sim R^2 \Sigma$

so
 $m \sim M$

$$RV^2 \sim L \left(\frac{M}{L} \right)$$

$$\sqrt{\frac{L}{\Sigma}} V^2 \sim L \left(\frac{M}{L} \right)$$

$$L \sim \frac{V^4}{\Sigma \left(\frac{M}{L} \right)^2}$$

$$M_{\text{halo}} = \rho_{200} \cdot \frac{4}{3} \pi R_{200}^3$$

$$V_c^2 = \frac{GM_{\text{halo}}}{R_{200}}$$

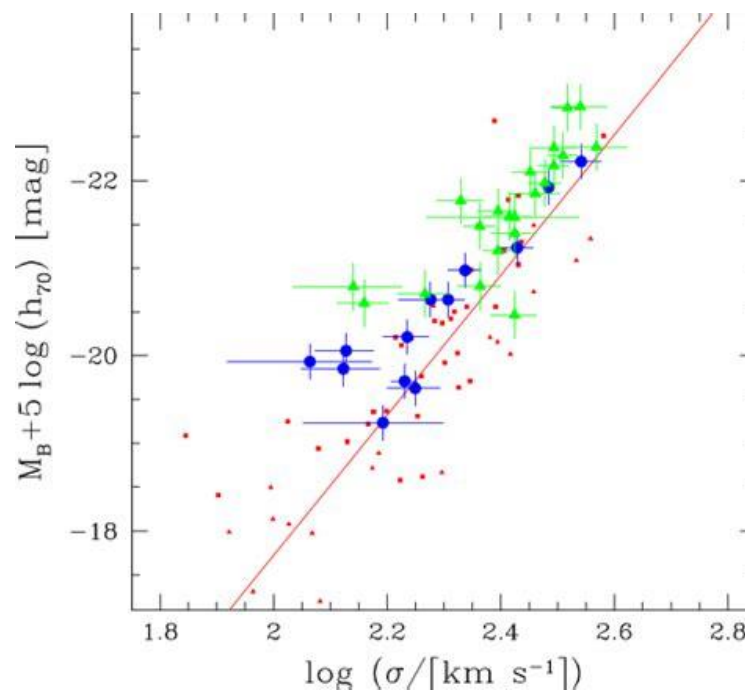
$$V_c \propto R_{200}$$

$$M_{\text{halo}} \propto V_c^3$$

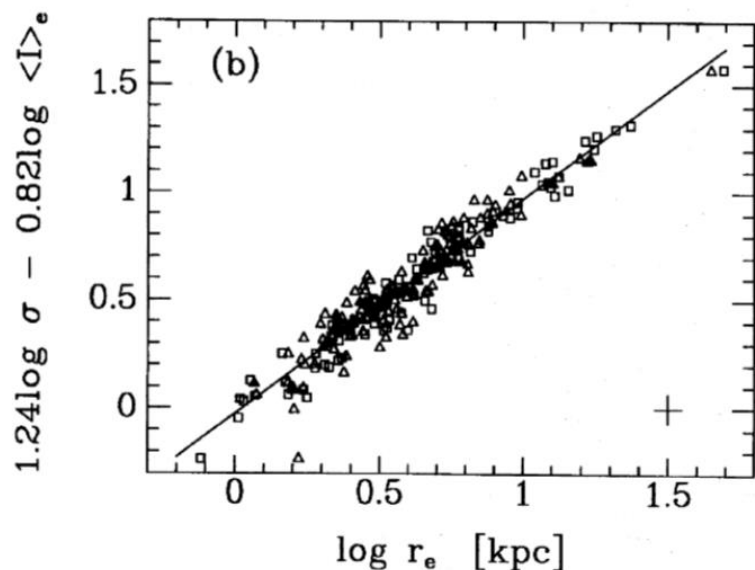
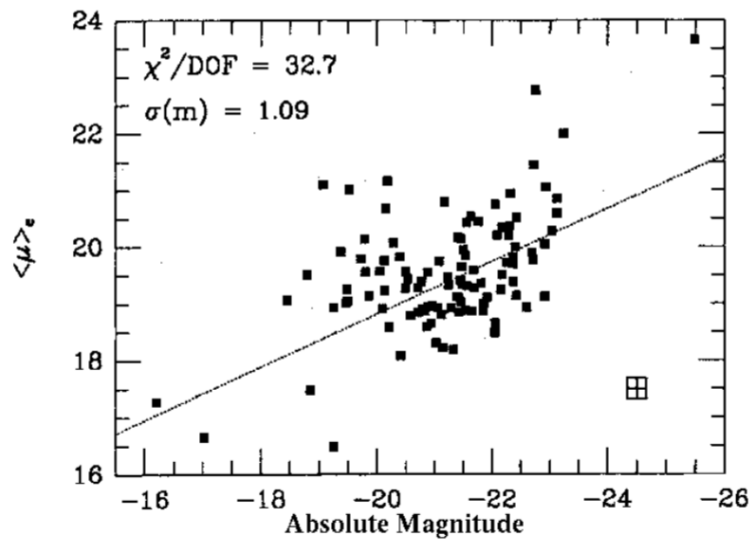
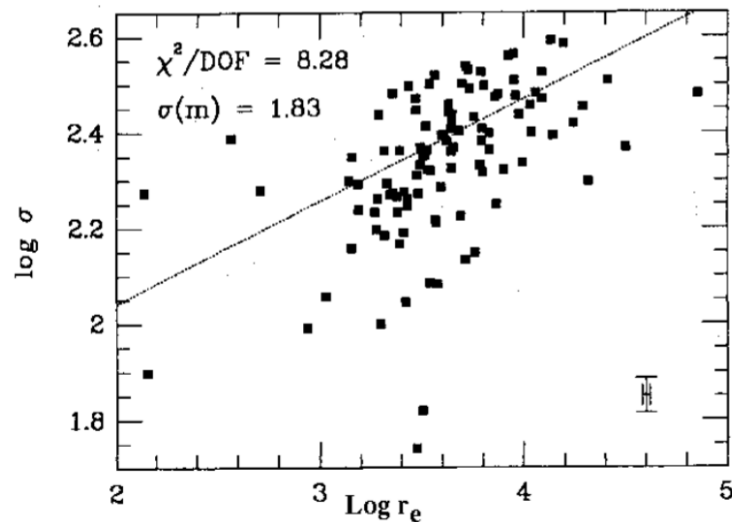
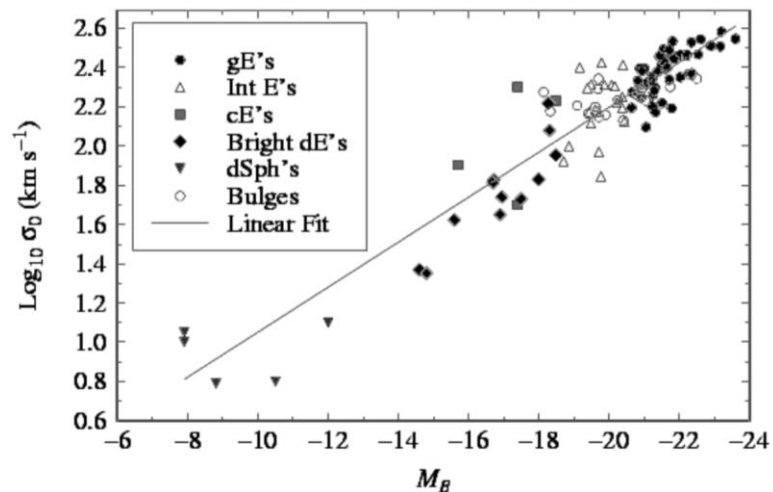
$$M^* \text{ --- } M_{\text{halo}}$$

椭圆星系的Faber-Jackson关系

- $L \propto \sigma^4$
 - σ : 恒星的速度弥散度
 - 关系的弥散 $\sim 0.4\text{mag}$



基本面 (fundamental plane)



维里定理 $2K + U = 0$

$$U = -\frac{3}{5} \frac{GM^2}{R} \quad (\text{半径为 } R, \text{ 质量为 } M) \text{ 等密度}$$

$$K = \frac{3}{2} M \omega^2$$

$$\omega^2 \propto G \frac{M}{R}$$

设 M/L 常数 $M \propto L$?

表面重力 $u = \frac{L}{\pi R^2}$

M : 常数?

$$L \propto \frac{\omega^4}{G^2 M}$$

基本面

$$\lg R_e = a \lg \sigma_0 + b \lg I_e + c$$

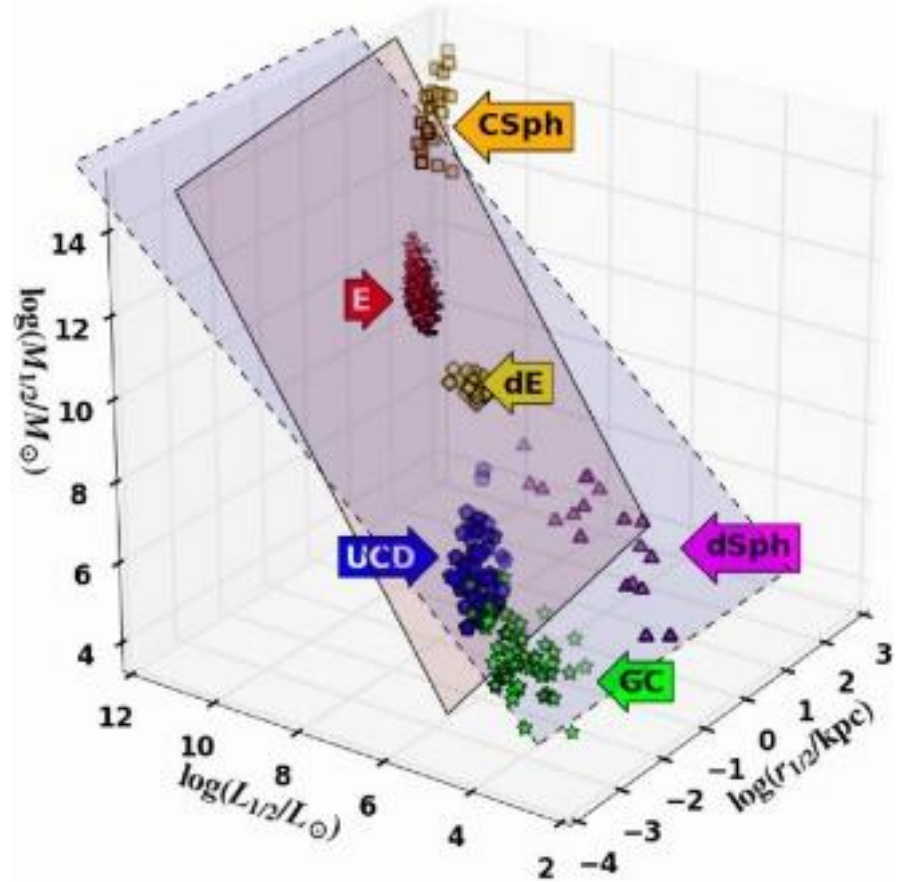
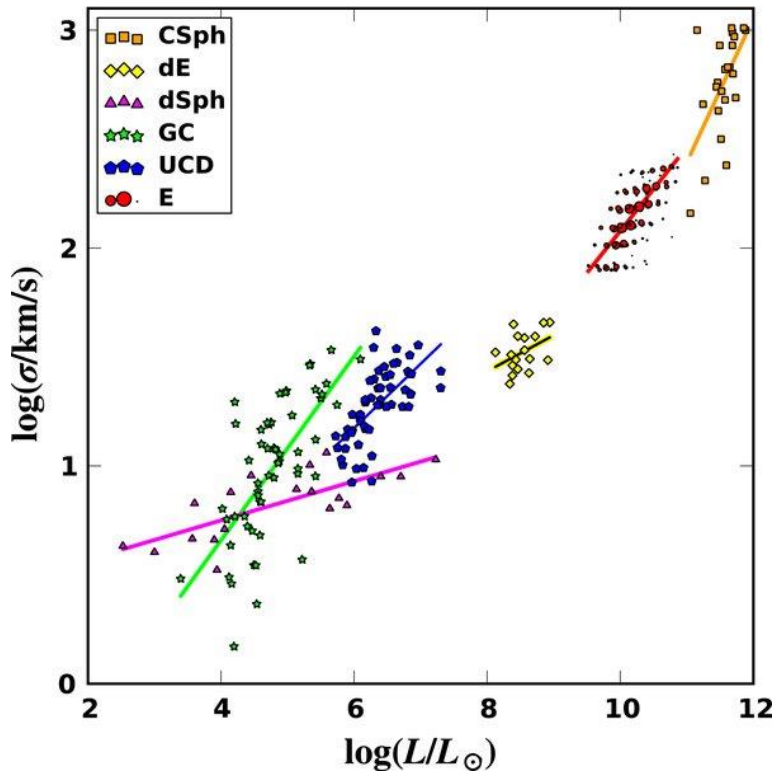
- 可以测定距离
 - l_e , σ_0 的测量不依赖于距离
- 偏离维里面 ($a=2, b=-1$)
 - 星系的质光比是常数吗?

$$\sigma_r^2 = \frac{1}{3} \gamma \frac{GM_{\text{dyn}}}{\frac{4}{3}R_h} = \frac{1}{2} G \gamma \frac{\frac{1}{2}L}{R_h^2} \frac{M_{\text{dyn}}}{L} R_h \propto IY R_h \quad . \quad (3)$$

其中, Y 为系统的质光比, $Y = M_{\text{dyn}}/L$, L 为系统的光度; I 为半质量半径处对应的平均面亮度, $I \propto \frac{1}{2}L/R_h^2$ 。质光比可以分解成重子物质比例的倒数 (M_{dyn}/M_* , M_* 为系统的恒星质量) 与恒星质光比 (M_*/L) 的乘积: $M_{\text{dyn}}/L = M_{\text{dyn}}/M_* \times M_*/L$ 。若质光比 Y 为常数, 则式 (3) 可以表示为:

$$\sigma^2 \propto I R_h \quad . \quad (4)$$

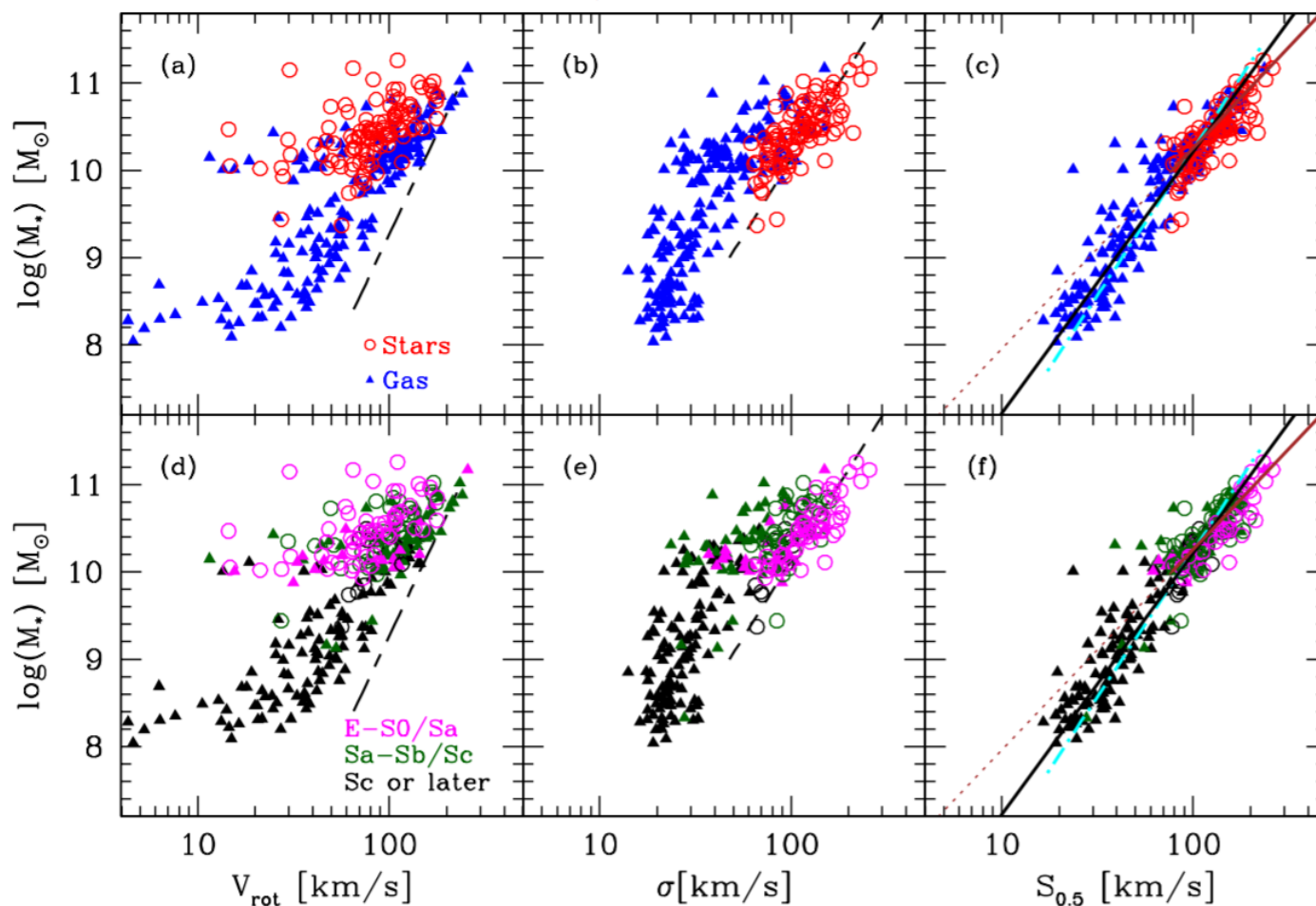
其它球状动力学系统？



Dynamical scaling relation

arXiv1410.3931

$$S_{0.5} = \sqrt{0.5V_{rot}^2 + \sigma^2}$$



stellar mass - binding energy relation

(Shi et al. arXiv:2109.00114)

$$E_{\text{binding,e}} \approx \frac{GM_{\text{dyn,e}}^2}{R_e} \equiv \frac{R_e V_e^4}{G}$$
$$= 4.62 \times 10^{48} \text{ erg} \left(\frac{V_e R_e^{0.25}}{\text{km s}^{-1} \text{ kpc}^{0.25}} \right)^4.$$

- V_e is the velocity of a circular orbit at R_e due to the total gravity from dark matter, stars and gas: $V_e = GM(R < R_e)/R_e$.
- The light-weighted velocity dispersion within the effective radius can be transformed to V_e by multiplying with a factor of 5/2 (Equation 19 of Cappellari et al. 2006).

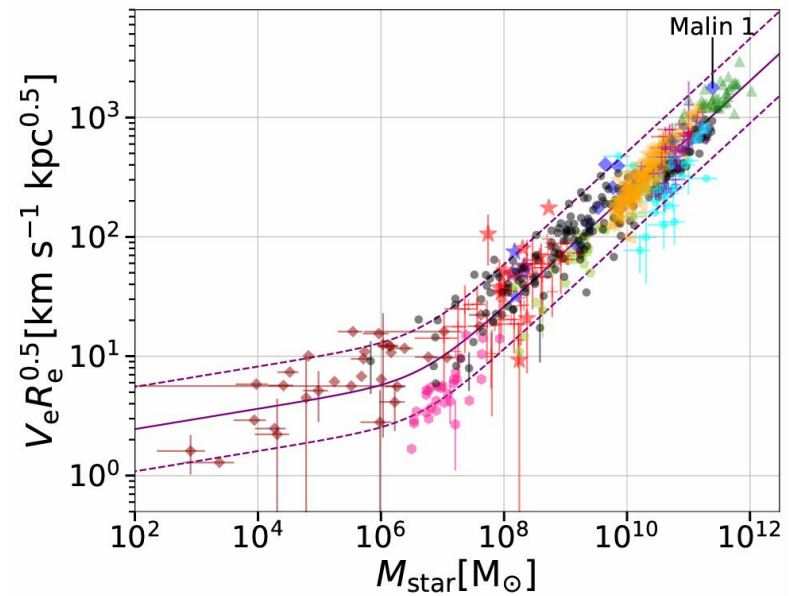
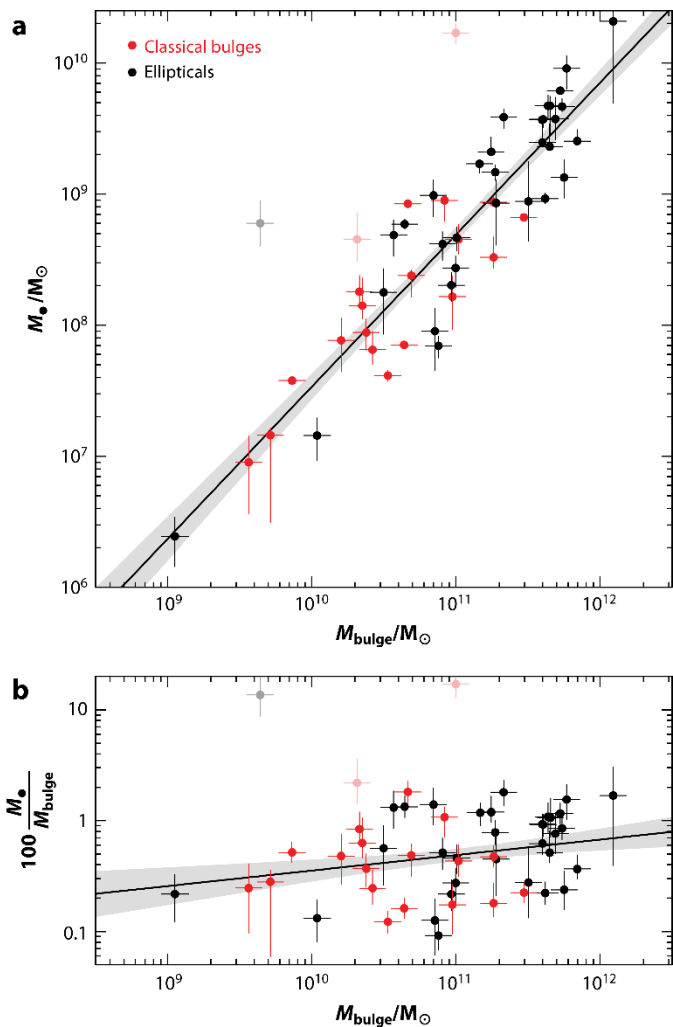
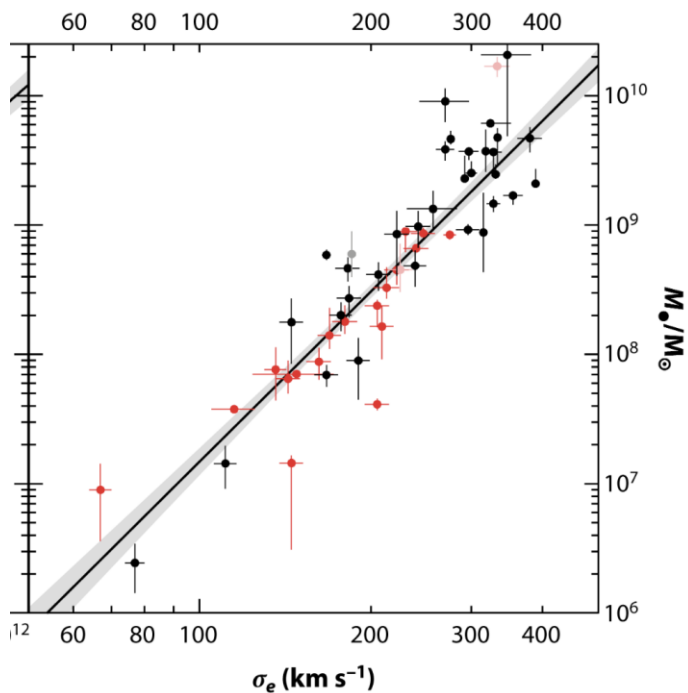


Figure 4. $V_e R_e^{0.5}$ that represents the square root of the dynamical mass within R_e , as a function of the galaxy stellar mass. For clarity, error bars that are smaller than 0.15

黑洞的M-sigma关系

- $M_{\text{BH}} \sim \sigma^{(4-5)}$
- $M_{\text{BH}} \sim (0.001-0.01) M_{\text{bulge}}$
 - 伪核球有更小的黑洞质量

$$100 \left(\frac{M_{\bullet}}{M_{\text{bulge}}} \right) = \left(0.49^{+0.06}_{-0.05} \right) \left(\frac{M_{\text{bulge}}}{10^{11} M_{\odot}} \right)^{0.15 \pm 0.07}; \quad \text{intrinsic scatter} = 0.28 \text{ dex.}$$

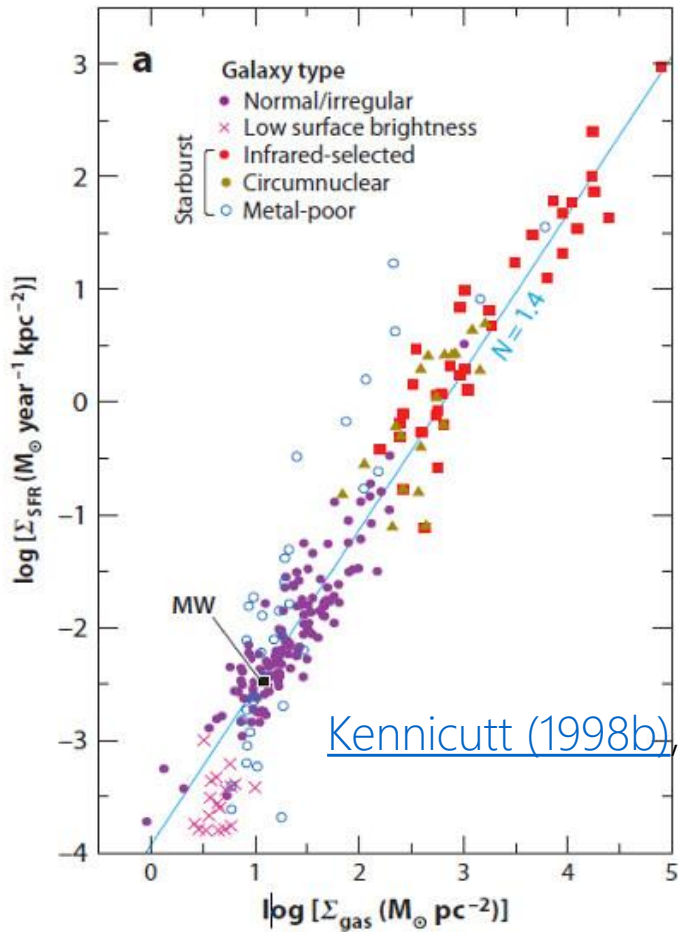


作业2

- 某旋涡星系和椭圆星系，其观测到的I波段的视星等都是18等；半光度半径内的表面亮度都是 20mag/arcsec^2 ；其中旋涡星系（图像椭率为0.5）观测得到的HI的速度展宽是 200km/s ，椭圆星系观测到的中心速度弥散度是 200km/s ，请问这两个星系的距离分别是多少？
 - 旋涡星系的 Tully-Fisher关系以及椭圆星系的 Fundamental Plane的内秉弥散在0.2个星等左右，请问以上距离计算的误差是多少？

3.2 恒星形成相关

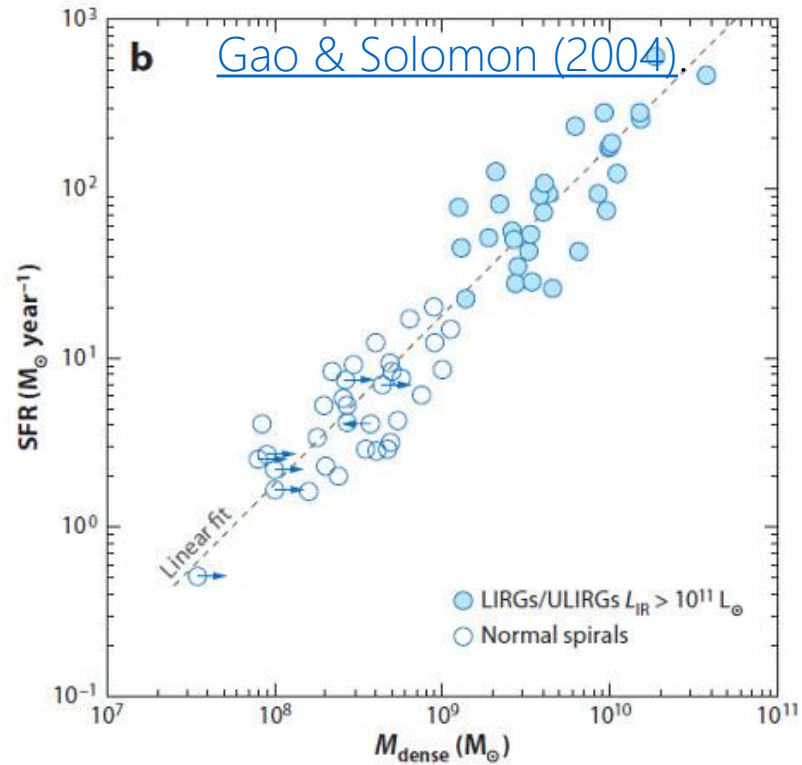
Kennicutt–Schmidt law

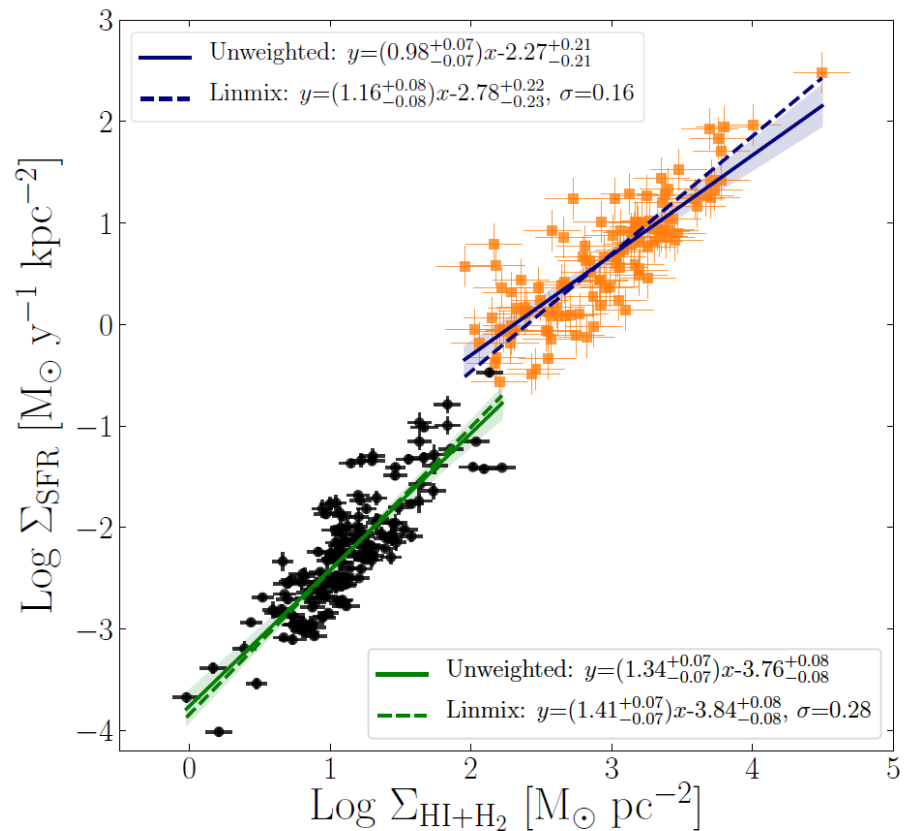
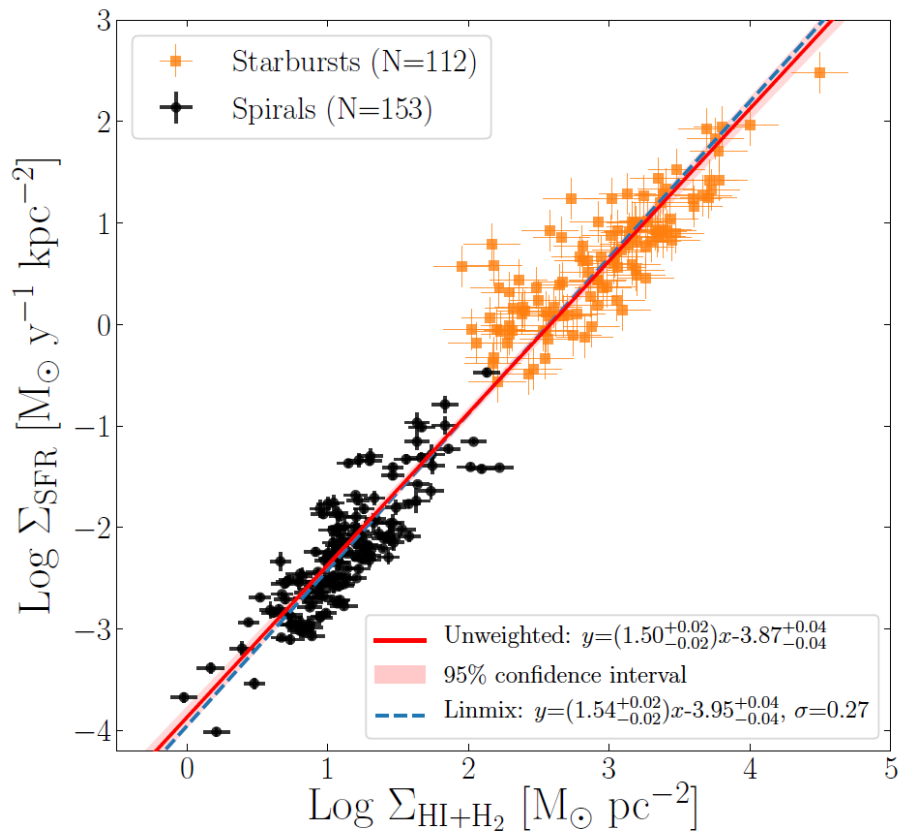


$$\Sigma_{SFR} = A \Sigma_{gas}^N$$

> Kpc尺度

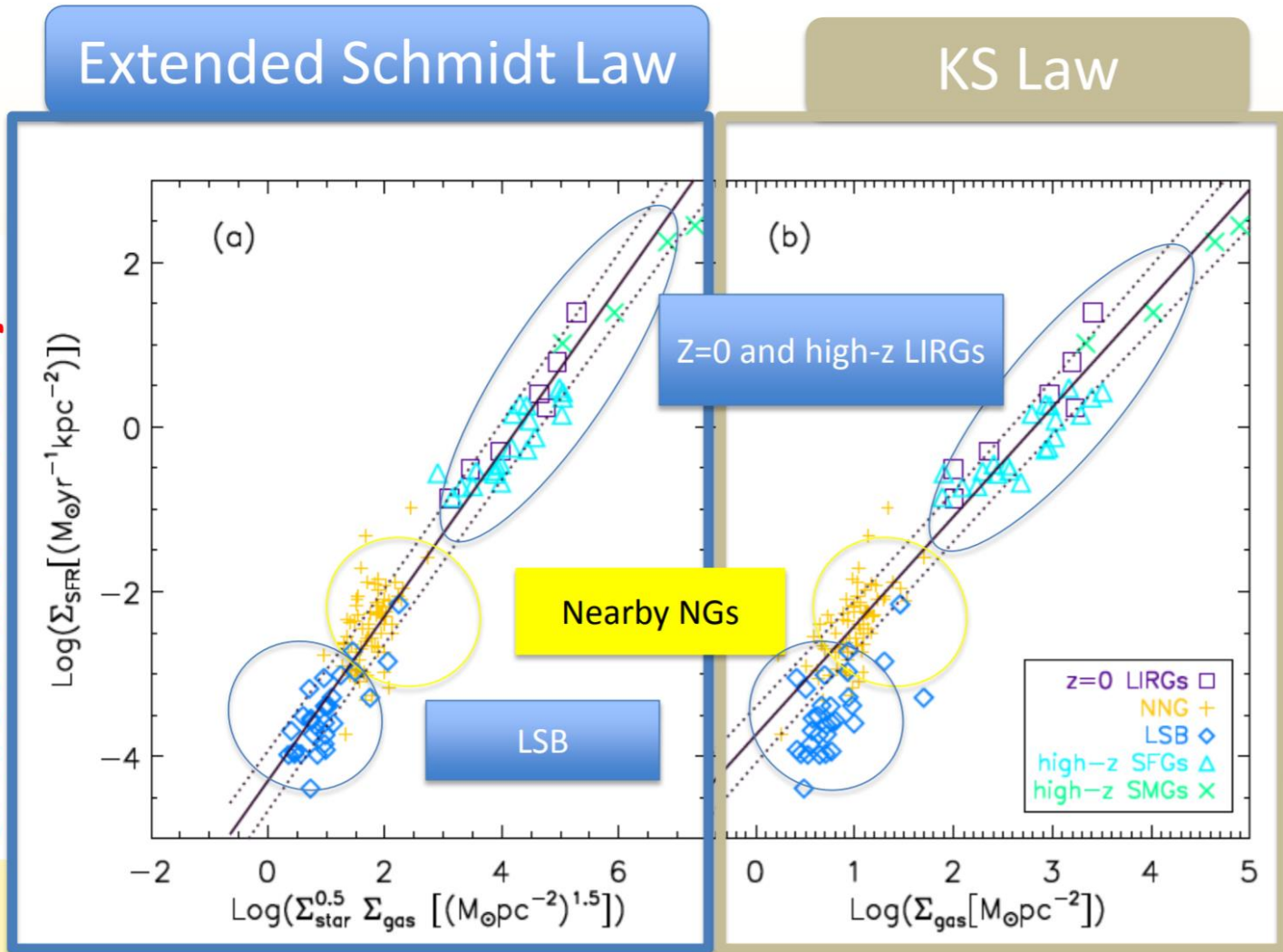
N=1.4



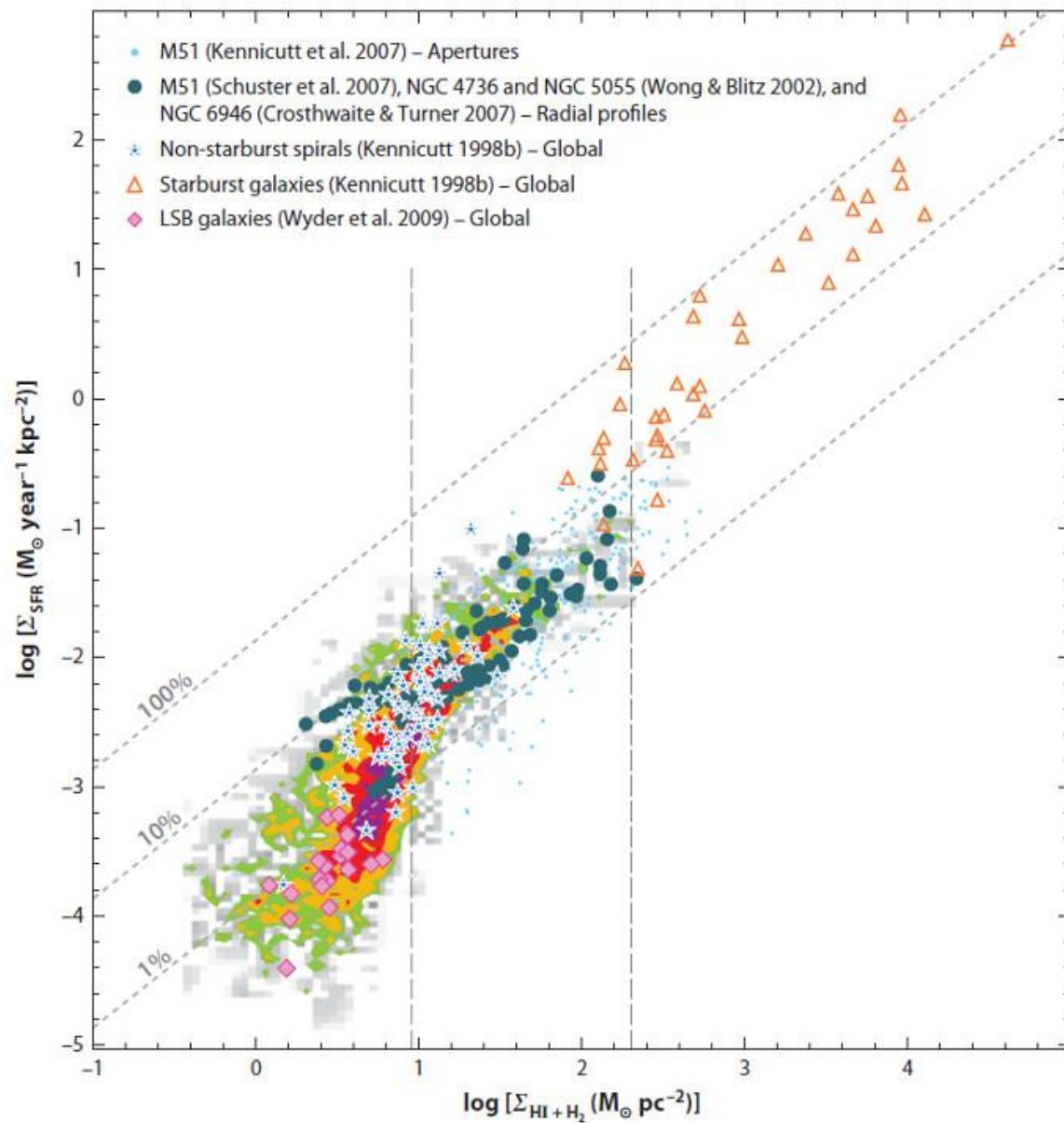


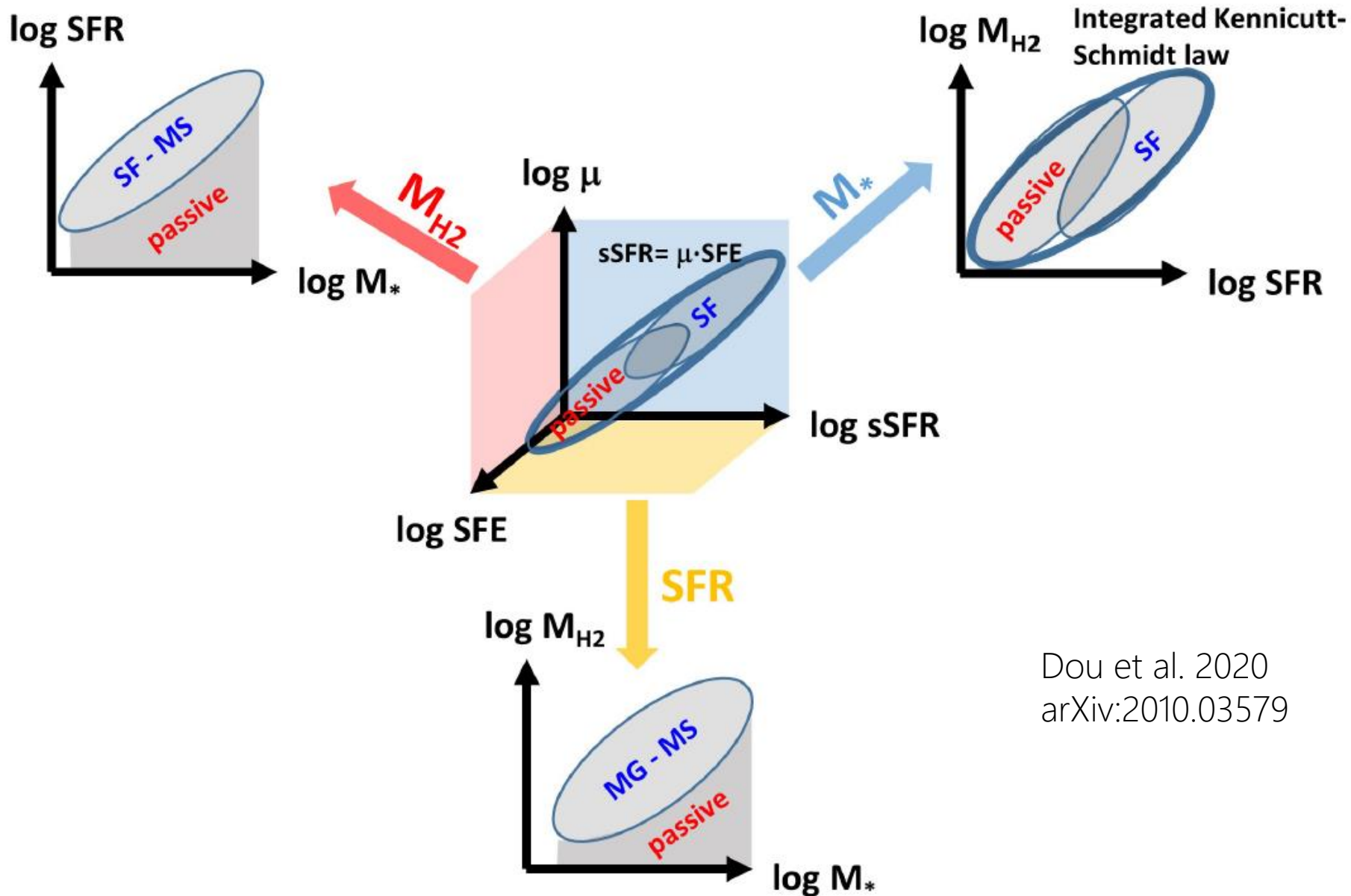
Extended Schmidt law (Shi et al. 2011)

SFR Surface Density



Sub-kpc 尺度



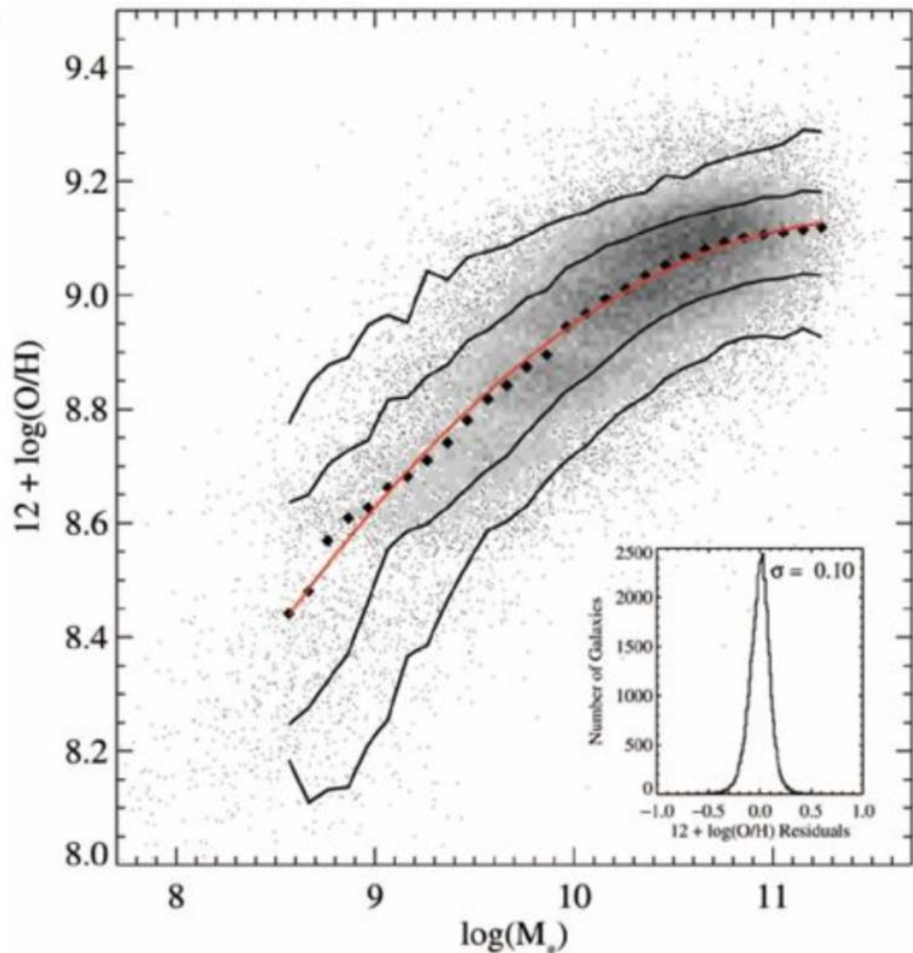


Dou et al. 2020
arXiv:2010.03579

3.3 金属丰度相关

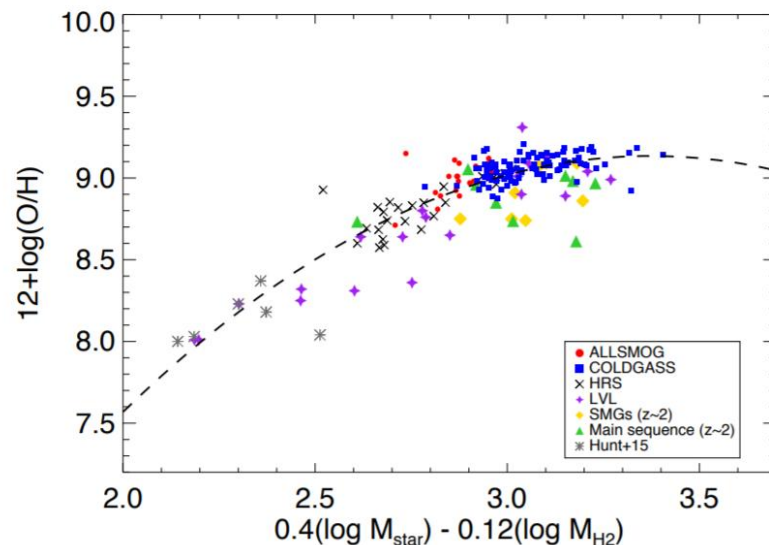
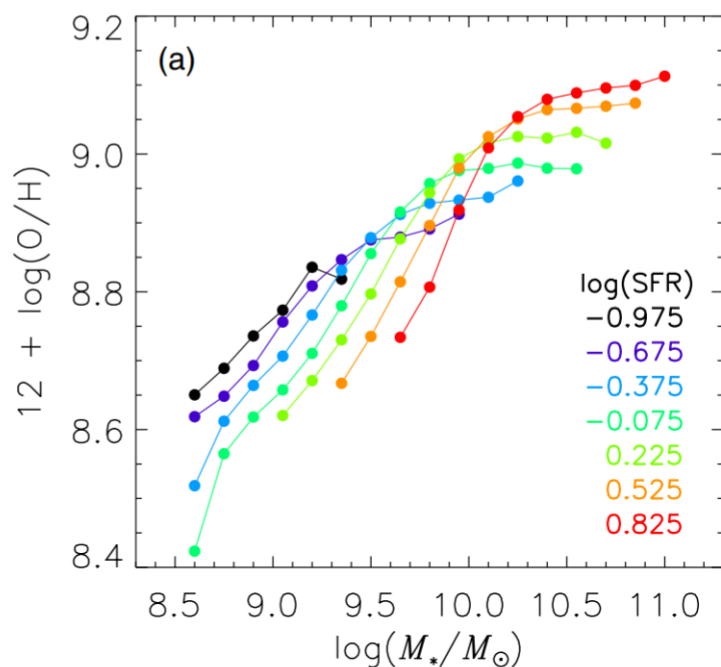
质量---金属丰度关系

- 大质量星系
 - 势井更深
- 其它模型：
 - 大质量星系的恒星形成效率更高

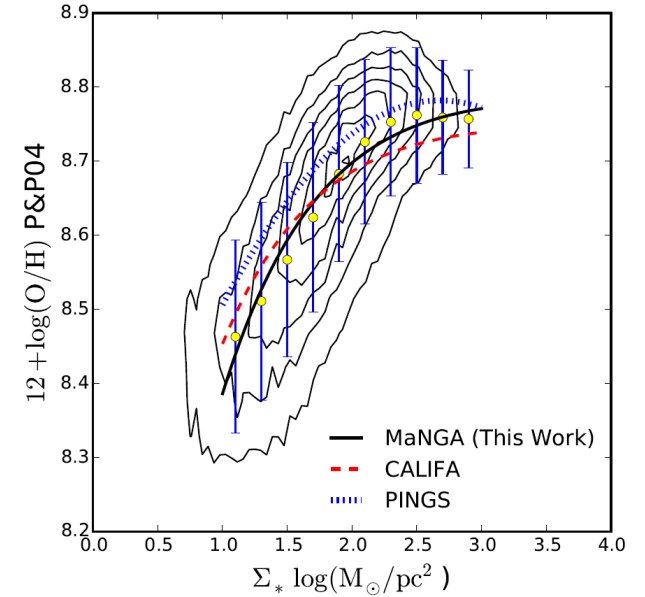
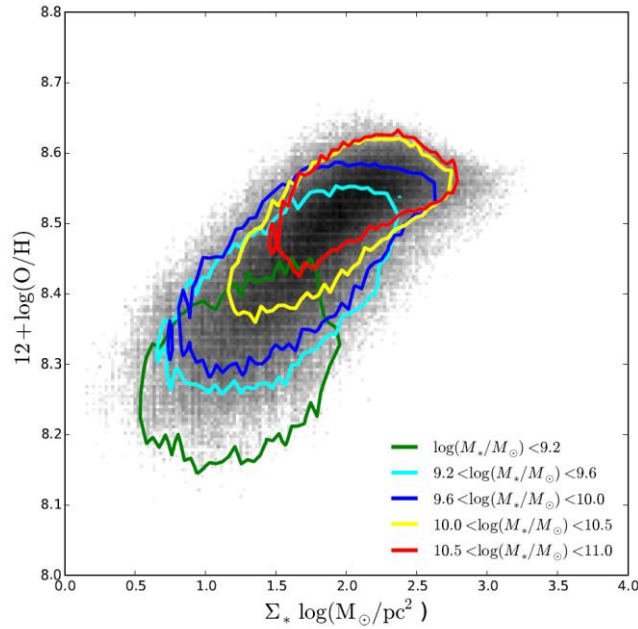


Fundamental metallicity relation

- M_* , gas-phase metallicity, SFR (Mannucci et al.2010)
 - 小质量端：气体内落稀释金属丰度
 - 大质量端：恒星形成星系富含气体

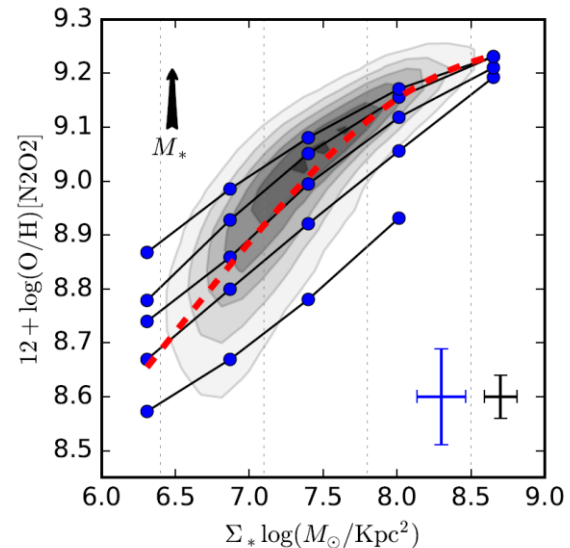
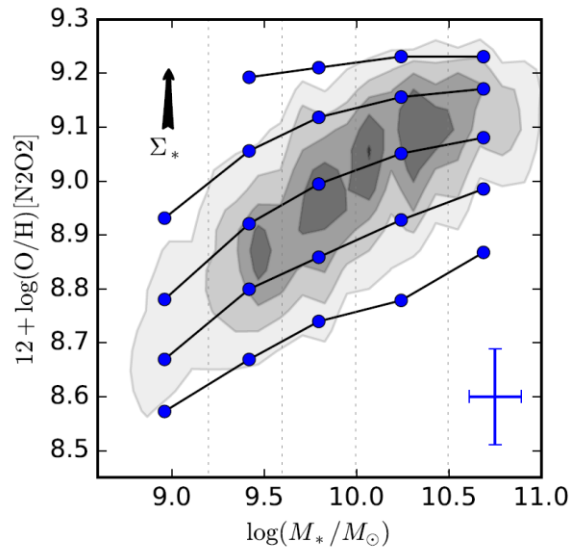


resolved MZ relation: **Surface mass density** Barrera-Ballestero et al. 2016



Gao et al. 2018 $(M_* - \Sigma_* - Z)$ relation

5



3.4 线性标度率的拟合

$$Y = ax + b$$

Famous linear scaling relations in astronomy

- period -luminosity relation of Cepheids
- $M_{\text{BH}} - \sigma$ relation
- Tully-Fisher ($L - V_{\text{max}}$) relation
- Fundamental plane of ellipticals
- $L - T$, $L - \sigma$ relation of groups and clusters
- All are statistical scaling relations, none of them are first principle like $F=ma$

Nature of the scaling relations

- Observables: (x_i, y_i) with error $(\Delta_{x,i}, \Delta_{y,i})$
- First, we should find some correlations, e.g. rank analysis
- To the first order, all the correlations are linear
 - $Y = a \cdot X + b + \sigma$
 - σ is the intrinsic scatter, may not be a constant
- Observables maybe biased
 - e.g. some low-luminosity galaxies are not observed
- Some observables may only be upper limits
 - e.g. we only get the upper-limit of L_x of some cluster

Ordinary Linear regression OLS($y|x$)

- y_i with measurement error σ_i

$$\chi^2(a, b) = \sum_{i=1}^N \left(\frac{y_i - a - bx_i}{\sigma_i} \right)^2$$

Code: *fit* in numeric recipes

Error on both x and y

$$\chi^2(a, b) = \sum_{i=1}^N \frac{(y_i - a - bx_i)^2}{\sigma_{y_i}^2 + b^2 \sigma_{x_i}^2}$$

Code: *fitexy* in numeric recipes

b ~ *biased to infinity*

Chi2 statistics

- 不考虑绝对值的 χ^2 : 带权重的最小二乘拟合
- $\chi^2/N \sim 1$: 完美模型+拟合
- $\chi^2/N \gg 1$: 不完美的模型 or 拟合
- $\chi^2/N \ll 1$: 数据太差无法有效区分模型

different linear regression methods

- Linear regression in astronomy I
(1990, ApJ,364,104)

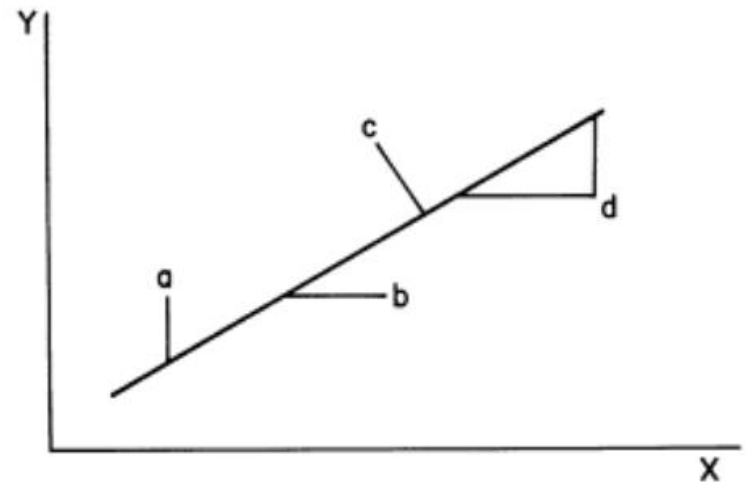


FIG. 1.—Illustration of the different methods for minimizing the distance of the data from a fitted line: (a) OLS($Y|X$), where the distance is measured vertically; (b) OLS($X|Y$), where the distance is taken horizontally; (c) OR, where the distance is measured vertically to the line; and (d) RMA, where the distances are measured both perpendicularly and horizontally. No illustration of the OLS bisector is drawn in this figure.

[Linear Regression - statsmodels 0.15.0 \(+520\)](#)

Error on both x and y and with a constant intrinsic scatter σ

$$\ln L = -\frac{1}{2} \sum_i \ln (\sigma^2 + \sigma_{y,i}^2 + a^2 \sigma_{x,i}^2) - \sum_i \frac{[\hat{y}_i - (a\hat{x}_i + b)]^2}{2(\sigma^2 + \sigma_{y,i}^2 + a^2 \sigma_{x,i}^2)} + \text{constant.}$$

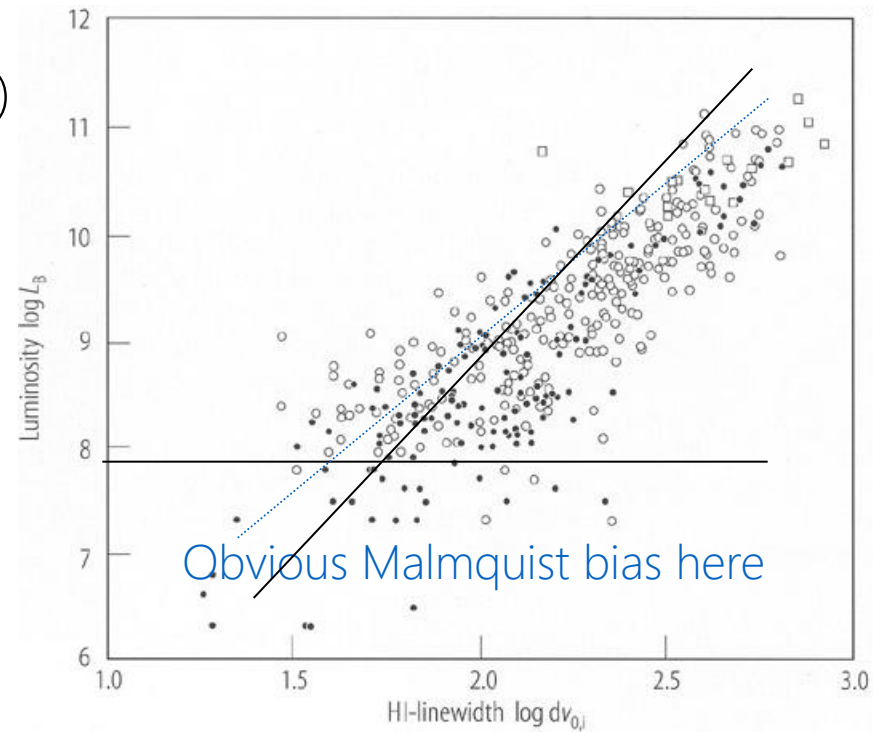
BCES : bivariate, correlate errors and scatter

(Akritas & Bershadsky, ApJ 470, 706 1996)

- Regression with correlated measurement errors and intrinsic scatter
 - allows for measurement errors on both variables
 - allows the measurement errors for the two variables to be dependent
 - allows the magnitudes of the measurement errors to depend on the measurements
- Intrinsic scatter: constant
- [bces · PyPI](#)

Eddington(Malmquist) bias

- Distance dependent observable
 - Eddington (1915), Malmquist(1920)
 - In magnitude limit sample, more faint source scattered in than bright source scattered out
- If we fit $M = a \log W + b$, a will be biased to smaller values
 - Fit $\log W = a' M + b'$ is better
 - At given M , no obvious in W

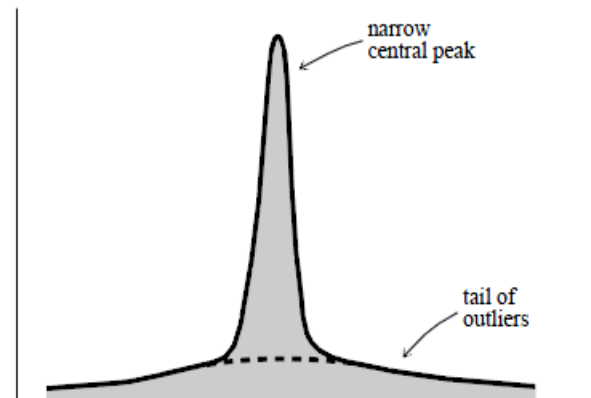


Robust estimation

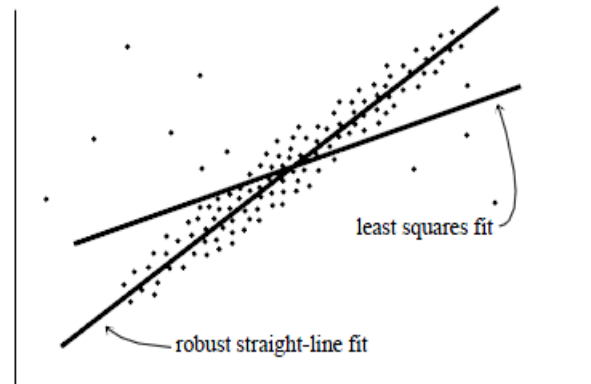
- Data with outlier

minimize over \mathbf{a}
$$\sum_{i=1}^N \rho \left(\frac{y_i - y(x_i; \mathbf{a})}{\sigma_i} \right)$$

$$\sum_{i=1}^N |y_i - a - bx_i|$$



(a)

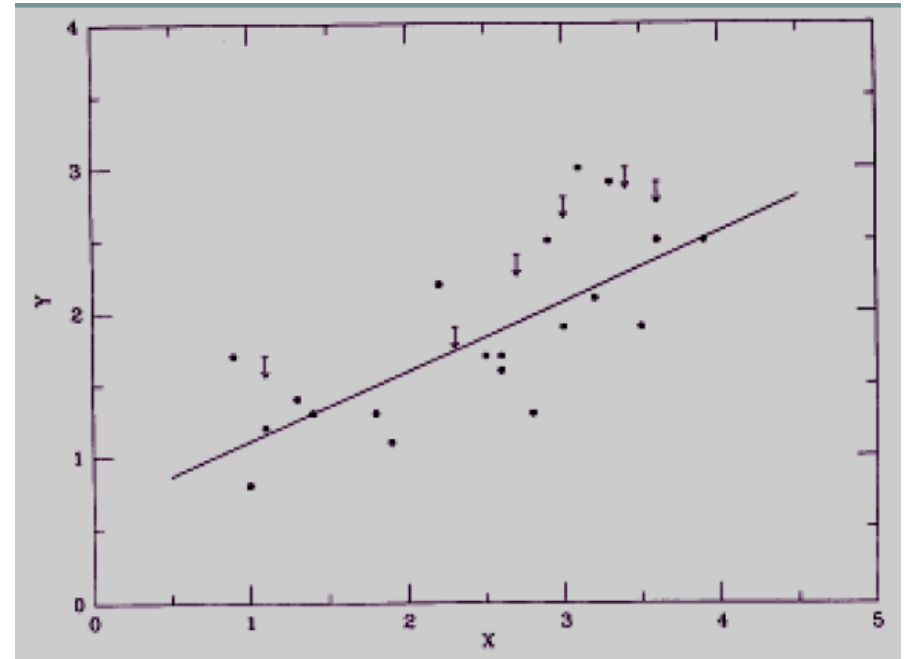
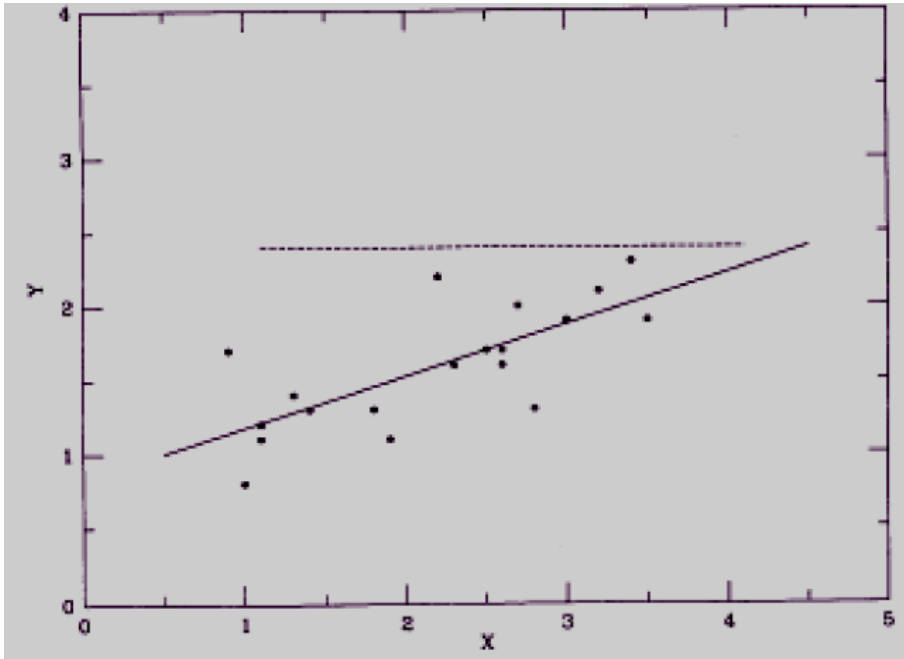


(b)

See *Numeric recipes* C15.7

Figure 15.7.1. Examples where robust statistical methods are desirable: (a) A one-dimensional distribution with a tail of outliers; statistical fluctuations in these outliers can prevent accurate determination of the position of the central peak. (b) A distribution in two dimensions fitted to a straight line; non-robust techniques such as least-squares fitting can have undesired sensitivity to outlying points.

Truncation, censored data

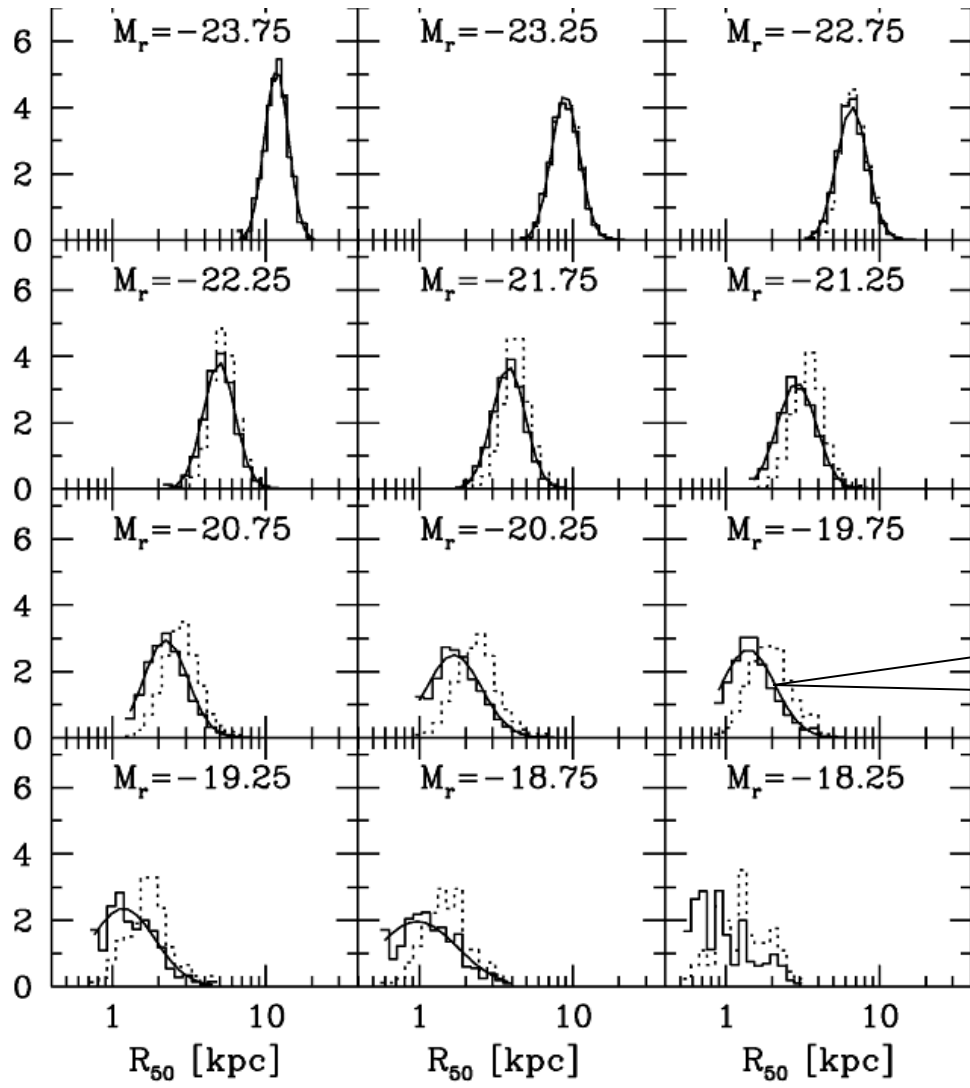


- Linear regression in astronomy (1992ApJ...397...55)
- Feigelson & Nelson (ApJ 1986)
- [GitHub - rsnemmen/asurv: ASURV: Astronomy SURVival analysis](https://github.com/rsnemmen/asurv)

A more straight-forward way

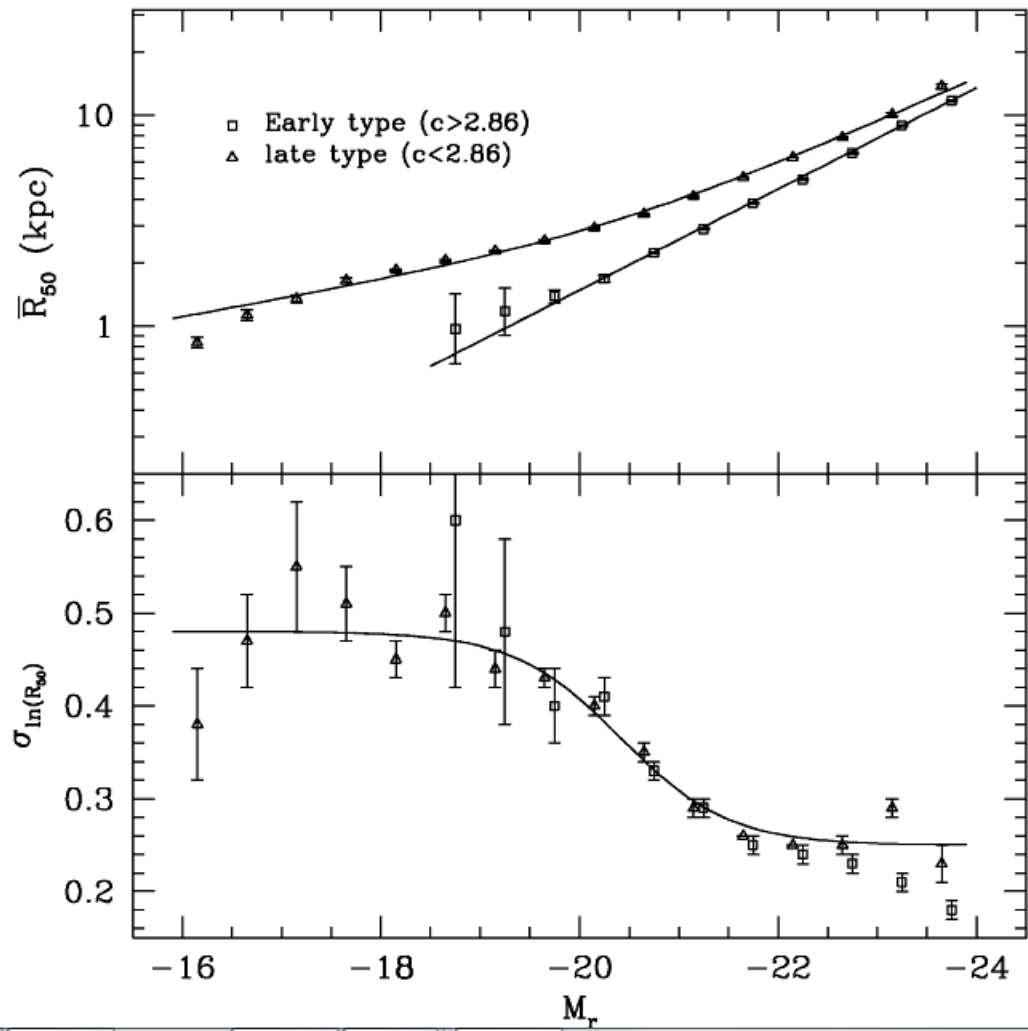
- First, at given bin of x , what is the distribution of y after correction for selection bias?
 - Is y Gaussian distributed? What is the scatter compared with its measurement error?
- Then what is the PDF($y|x$) changes as function of x
 - Is this relation linear or non-linear?
- What is $P(x)$? (luminosity/mass function)
- $P(x,y)=P(x)*P(y|x)$

L – R relation of galaxies (Shen et al. 2003)



We find, after correction for selection effect, at given M_r , $\text{Log } R$ is intrinsically Gaussian distributed.

Data is biased here



We plot $P(R|M)$ as function of M .

Intrinsic scatter is not a constant

补充阅读

- [Astrostat Club - cluster](#)
- [arXiv:2411.08747](#)
 - X的分布函数? X, Y的观测误差和X, Y之间存在关系 (泊松噪声)

Table 1. Comparison of Different Regression Methods

Method	$P(\mathbf{x})$	$P(\mathbf{x}_{err} \mathbf{x})$	$P(y \mathbf{x}, \boldsymbol{\theta})$	Optimization Objective
ML based method (This work)	NF	NF	Any	Likelihood/Posterior
KS-test based method (This work)	NF	NF	Any	p -value of 2D KS test
OLS/WLS	Linear	Likelihood
ODR/wODR	Uniform	...	Linear	Likelihood
mODR	Linear	Likelihood
LINMIX/ROXY	GMM	...	Linear/Any	Posterior
Leopy	Input (heuristic)	...	Any	Likelihood/Posterior
LtsFit	Linear	Likelihood

NOTE— $P(y_{err}|y)$ is modelled by the same method as $P(\mathbf{x}_{err}|\mathbf{x})$.

结题作业

- 本课程中收获的问题列表
- 某个星系（或问题）的综合论述

[Lecture Notes for Download – Frank van den Bosch](#)