GALAXY FORMATION

Week 9

A UNIVERSE OF THINGS

- We have covered how the Universe evolves with time, how temperatures change and how density fluctuations grow and eventually collapse.
- The final step is to form the objects we actually see in the Universe; stars, planets, black holes, neutron stars, people,etc.
- All of these things form in galaxies. Thus galaxies are the main structural block in the Universe. While their are some things that aren't galaxies, the vast majority of astrophysical objects are born and live in galaxies.

Thus image, called the Hubble Ultra Deep Field, is the deepest picture ever taken of space. What you see are thousands and thousands of galaxies and a few stars. The Universe, when one looks deep enough, is a collection of galaxies.

The galaxies show a wide range of properties, the colors, sizes, shapes and total brightnesses vary over a wide range of values. A theory of galaxy formation should be able to explain all these properties.

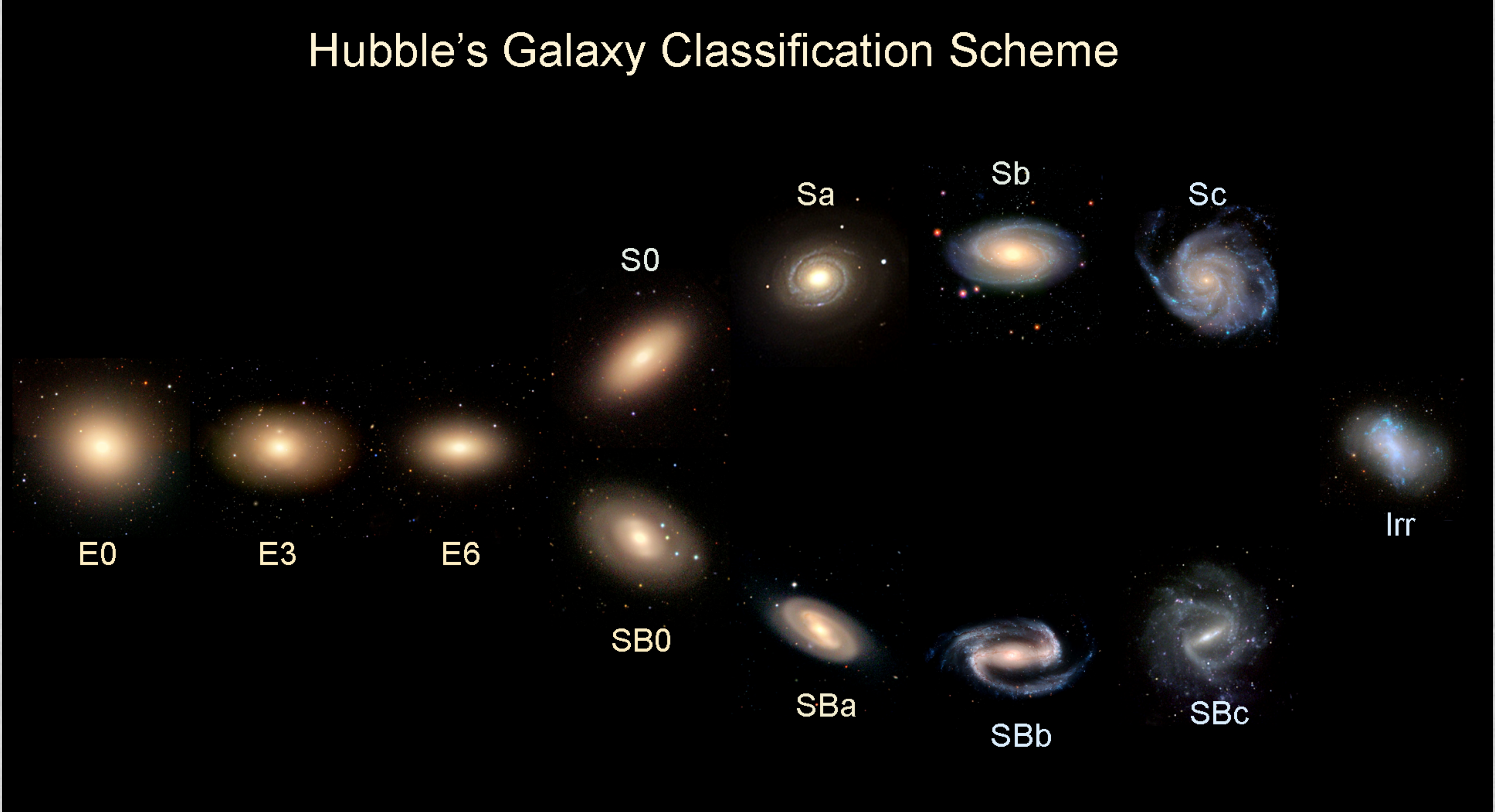
Hubble Ultra Deep Field
Hubble Space Telescope • Advanced Camera for Surveys

NASA, ESA, S. Beckwith (STScI) and the HUDF Team

STScI-PRC04-07a

GALAXY PROPERTIES

- Galaxies can be broadly classified into two type; spirals and ellipticals.
- Spiral galaxies or disk galaxies are disk like in shape, stars are mostly in rotational motion. These galaxies contain gas and dust and have ongoing star formation. The are blueish in color.
- Elliptical galaxies are as the name suggests ellipsoidal in shape, stellar motions are random providing pressure support. There is negligible star formation, gas and dust. The stars are old and the colors red.



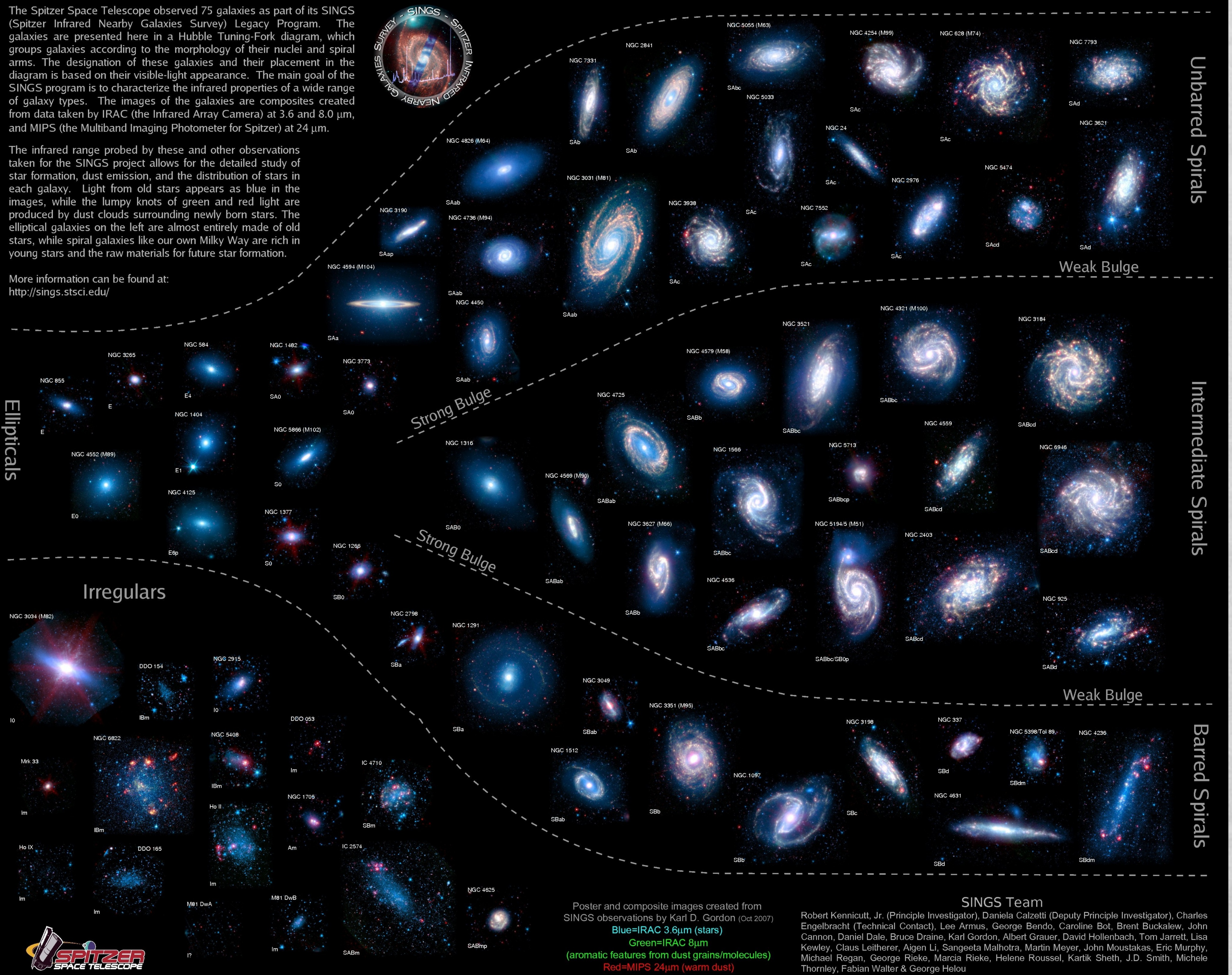
Hubble originally made these classification and put them together in a tuning fork. Today we don't think the distinction between barred and unbarred spirals matters. Also we don't think that as ellipticals become more elliptical that they become more like spirals. Today they are just treated as separate groups or divided by disk-to-bulge ratio.

The Spitzer Infrared Nearby Galaxies Survey (SINGS) Hubble Tuning-Fork

The Spitzer Space Telescope observed 75 galaxies as part of its SINGS (Spitzer Infrared Nearby Galaxies Survey) Legacy Program. The galaxies are presented here in a Hubble Tuning-Fork diagram, which groups galaxies according to the morphology of their nuclei and spiral arms. The designation of these galaxies and their placement in the diagram is based on their visible-light appearance. The main goal of the SINGS program is to characterize the infrared properties of a wide range of galaxy types. The images of the galaxies are composites created from data taken by IRAC (the Infrared Array Camera) at 3.6 and 8.0 µm, and MIPS (the Multiband Imaging Photometer for Spitzer) at 24 µm.

The infrared range probed by these and other observations taken for the SINGS project allows for the detailed study of star formation, dust emission, and the distribution of stars in each galaxy. Light from old stars appears as blue in the images, while the lumpy knots of green and red light are produced by dust clouds surrounding newly born stars. The elliptical galaxies on the left are almost entirely made of old stars, while spiral galaxies like our own Milky Way are rich in young stars and the raw materials for future star formation.

More information can be found at:
http://sings.stsci.edu/



Poster and composite images created from
SINGS observations by Karl D. Gordon (Oct 2007)
Blue=IRAC 3.6µm (stars)
Green=IRAC 8µm
(aromatic features from dust grains/molecules)
Red=MIPS 24µm (warm dust)

SINGS Team

Robert Kennicutt, Jr. (Principle Investigator), Daniela Calzetti (Deputy Principle Investigator), Charles Engelbracht (Technical Contact), Lee Armus, George Bendo, Caroline Bot, Brent Buckalew, John Cannon, Daniel Dale, Bruce Draine, Karl Gordon, Albert Grauer, David Hollenbach, Tom Jarrett, Lisa Kewley, Claus Leitherer, Aigen Li, Sangeeta Malhotra, Martin Meyer, John Moustakas, Eric Murphy, Michael Regan, George Rieke, Marcia Rieke, Helene Roussel, Kartik Sheth, J.D. Smith, Michele Thornley, Fabian Walter & George Helou

BULGE/DISK CONTINUUM

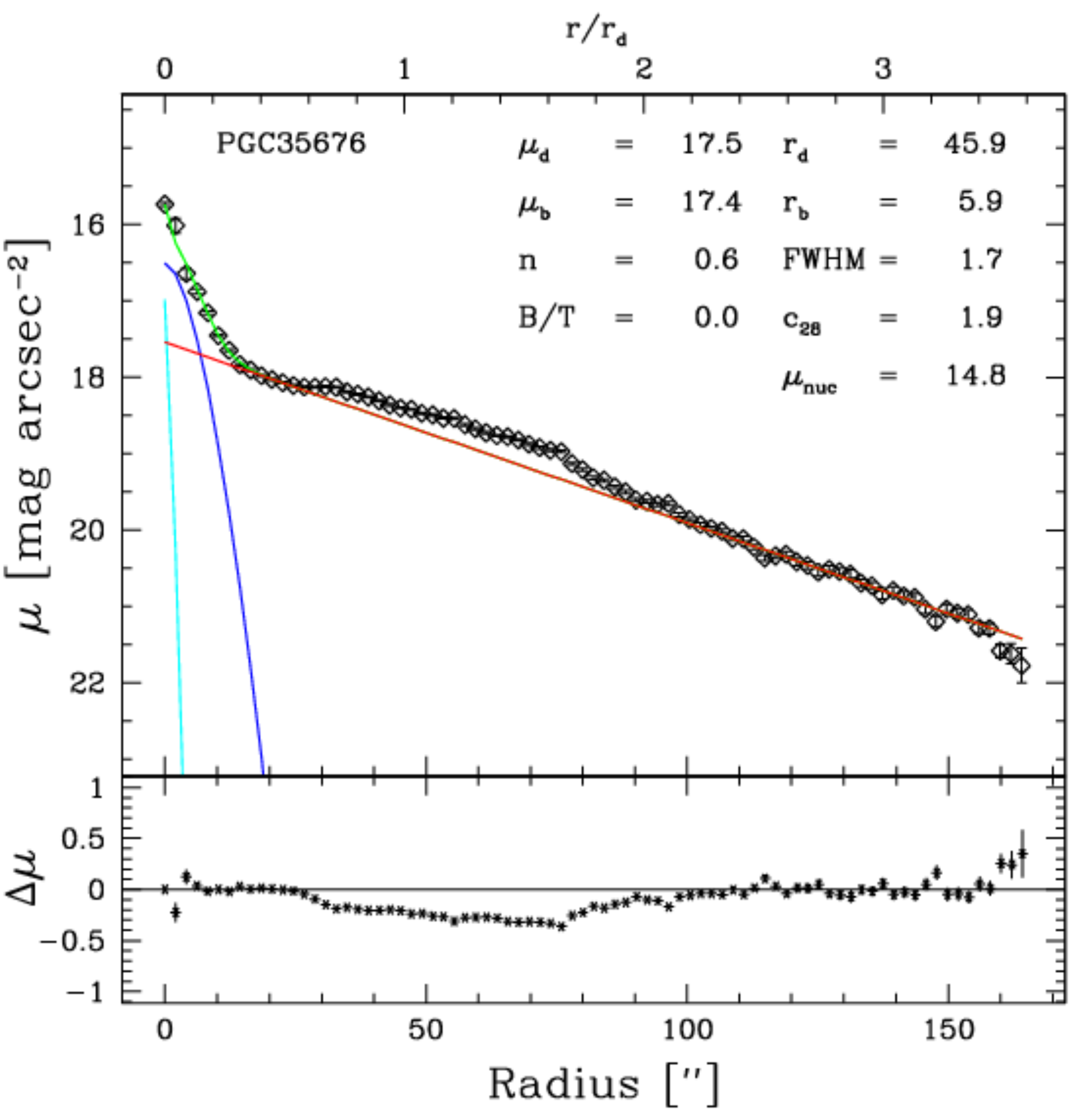
- Now days it is thought that this two type classification is too simple and that instead there is a continuum where galaxies are part disk and part bulge.
- There are pure disk and pure bulge galaxies that correspond to the spirals and ellipticals, but there are also galaxies in between.
- Often people still try to cut a sample into two types and the two type simplification is still widely used.

LIGHT PROFILES

- The light profiles of galaxies can be measured as a function of elliptical isophotes and primarily two types are found.
- Disks are well fit by an exponential light profile with a scale radius R_s.
- Ellipticals and large bulges are fit by the de Vaucouleurs profile or $\exp((R/R_e)^{1/4})$. Where R_e is called the effective radius, but it is the half light radius.
- The fraction of disk and bulge is usually determined by fitting both profiles and judging their relative contribution to the total light. A generalization called the Sersic profile where 1/4 is replaced by 1/n can be used to fit intermediate cases.

The light profile of a galaxy fit with an exponential profile and a bulge profile.

Note the clear change in the light profile with radius. Bulge profiles fall off very fast with radius, disks are much more extended.

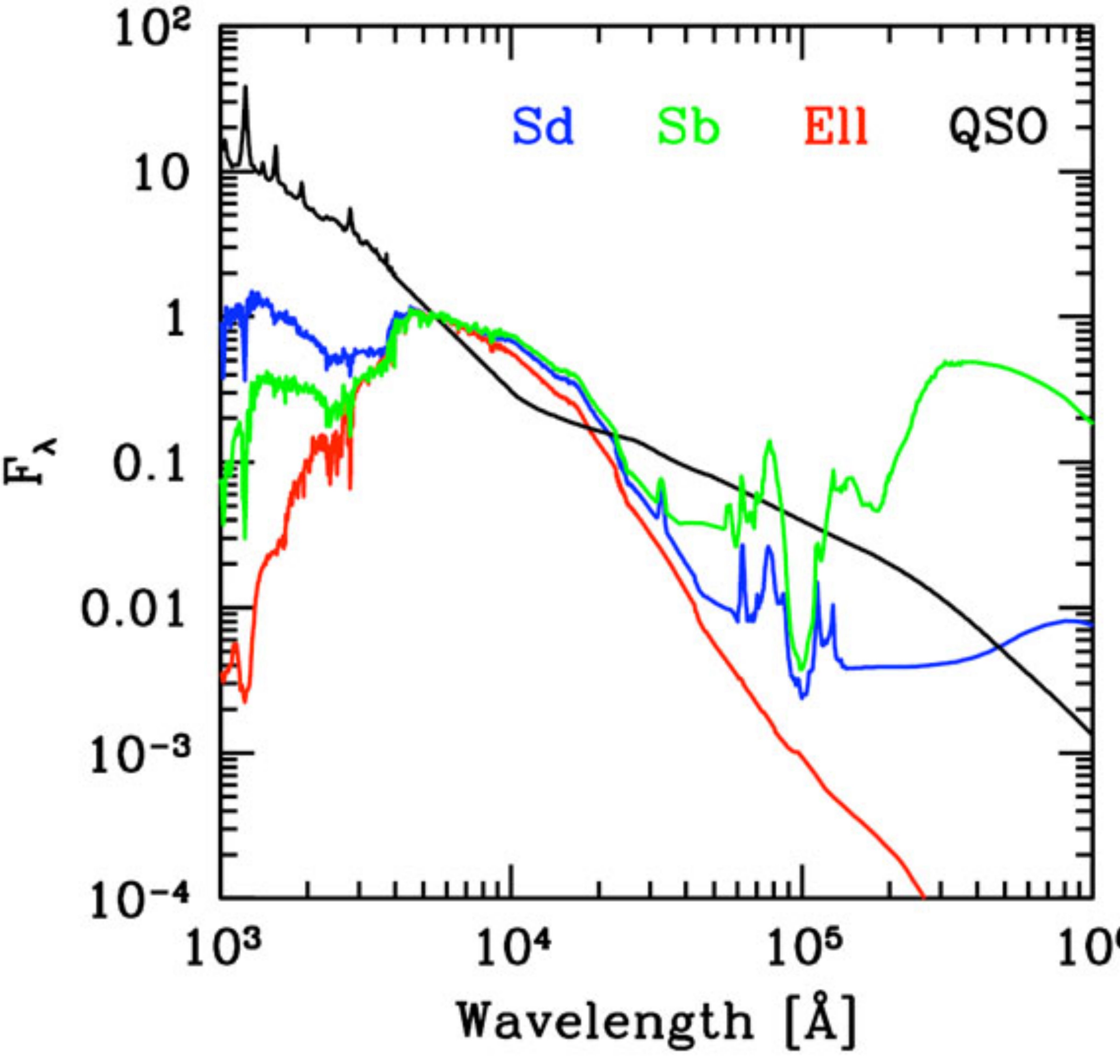


BASIC GALAXY PROPERTIES

- Every galaxy can be described by a few basic properties. The fundamental properties are the most interesting as far as theory goes, but most of those can't be directly observed.
- The spectral energy distribution (SED) of a galaxy is the amount of light at every wavelength. With a model this can be converted into a galaxy's stellar mass and star formation history. The mass mostly controls its overall luminosity while the history determines its color.
- A galaxy's "current" star formation rate can be measured with a number of indicators.

The SEDs of galaxy show a wide range of behavior which is why galaxies have different colors.

Disk type galaxies are bluer because they have ongoing star formation which emits at uv and blue wavelengths.



BASIC GALAXY PROPERTIES

- A galaxy's stellar mass, star formation rate and star formation history are basic properties.
- The size of its bulge and disk and the relative amount of each are also fundamental.
- The amount of cold gas and its metallicity, the mass of the super massive black hole in a galaxy core and the rotational velocity and velocity dispersion of a galaxy all should be fundamental too.

$$\{m_*, \dot{m}_*, r_b, r_d, B/T, v_c, \sigma, m_g, Z, m_{bh}\}$$

OTHER PROPERTIES

- Galaxies have many other properties that we believe are less indicative of their basic properties.
- These include the degree of spiral structure or bar in a galaxy, the inclination of a disk and ellipticity of a bulge, the amount of dust in a galaxy, warps in the disk and other small perturbations.
- Thus we will consider the basic properties as describing galaxies and these others as noise.

OBSERVATIONAL STUDIES

- Observationally the statistical study of galaxies mostly consists of describing the distribution of galaxy properties and looking for relationships between them.
- Common measurements are; the galaxy stellar mass function (gsmf), the galaxy size function, the galaxy velocity function and the gas mass function.
- The galaxy correlation function is also an important fundamental observation that can tell us about the clustering of galaxies in terms of other properties.

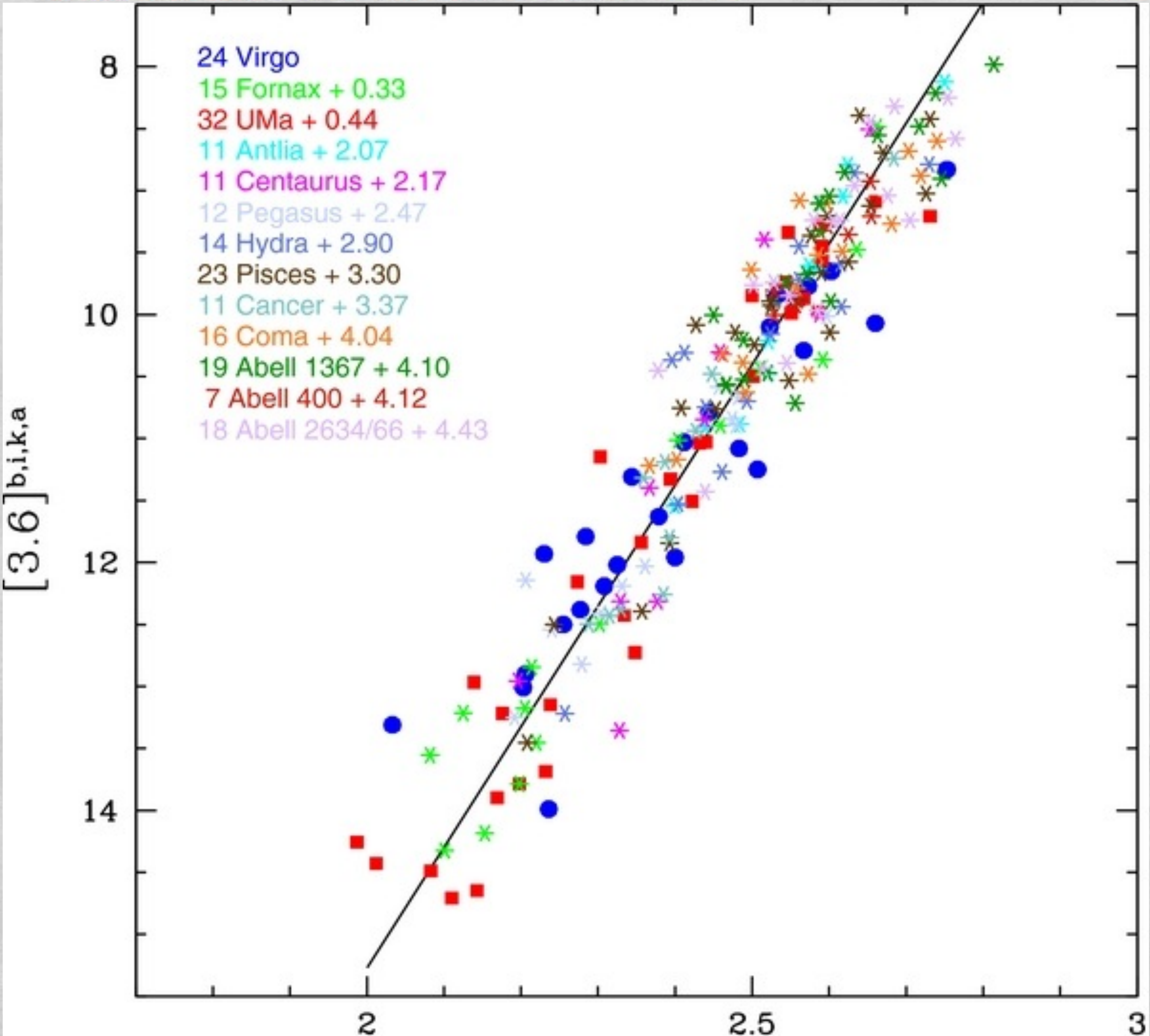
RELATIONS BETWEEN PROPERTIES

- Relations between properties are even more useful for a theoretical understanding of galaxies because they tell us multiple properties are controlled by the same physical mechanism.

- We have already introduced the Tully-Fisher and Faber-Jackson relations between galaxy mass and velocity as well as the fundamental plane for ellipticals.

- In addition there is the Kormandy relation between elliptical galaxy size and mass, the baryonic Tully-Fisher between relation stellar plus gas mass and velocity, the mass metallicity relation and the m-σ relation between bulge velocity and black hole mass.

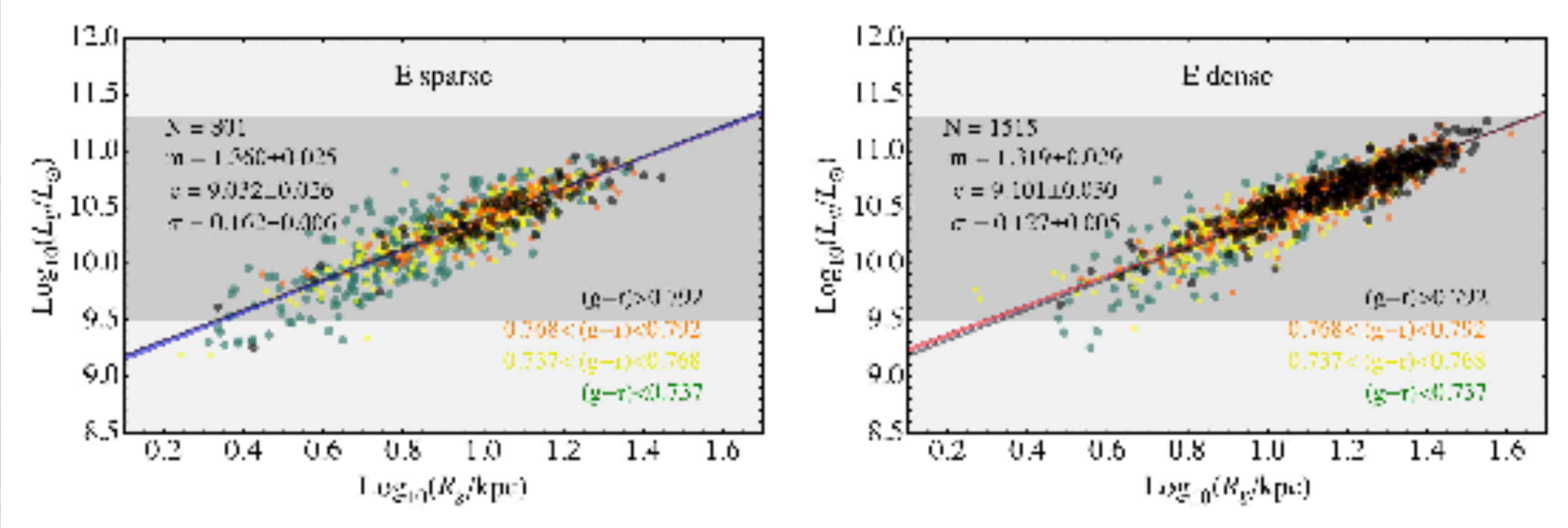
TULLY-FISHER



A relationship between galaxy rotational velocity (only good for disks) and luminosity (probably mass).

This makes sense since the more massive the system the stronger the gravity the faster stars would have to orbit.

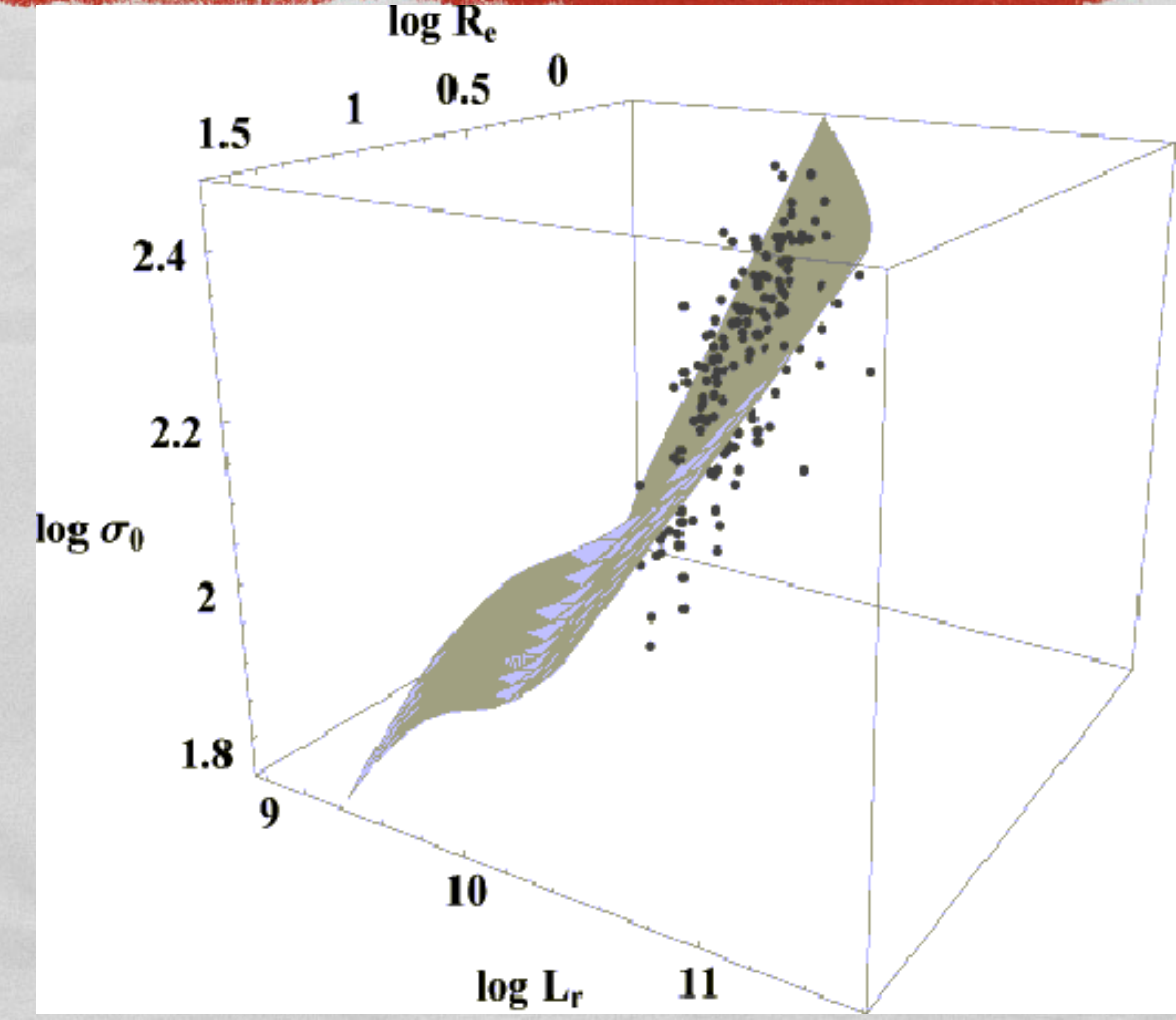
LUMINOSITY - SIZE



The size of elliptical galaxies increases with increasing luminosity (mass). Not entirely unexpected. This is often called the Kormandy relation after its discoverer. The situation with spirals is somewhat more complicated. There is an overall luminosity size relation, but the scatter at a given luminosity is much greater. This suggests size is more tied to mass for spheroids than for disks.

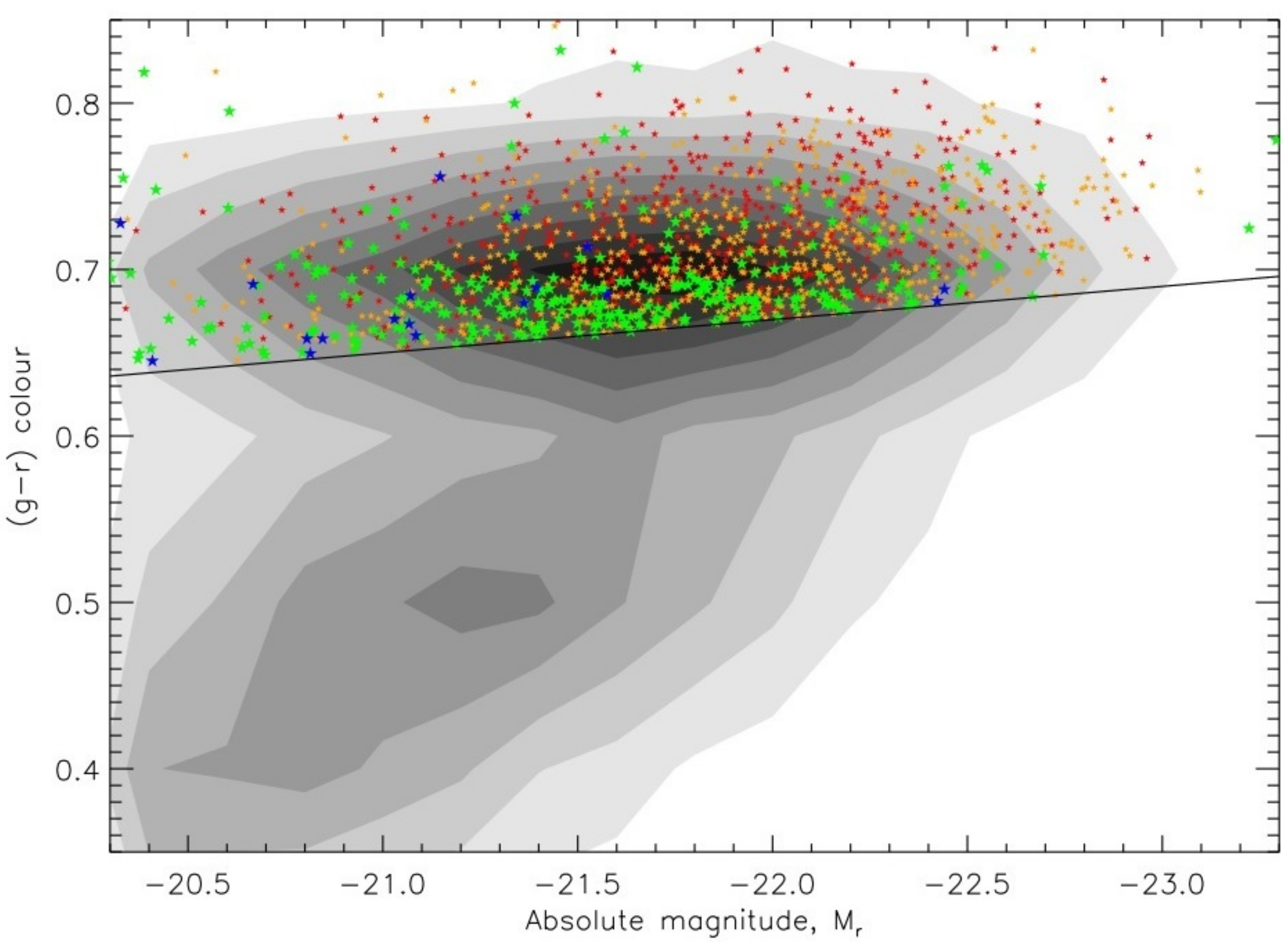
FUNDAMENTAL PLANE

A relationship for elliptical galaxies between their velocity dispersion, luminosity and size. Since we know velocity and mass are related and that mass and size are related, this basically says that a tighter relation can be found when all three are used.



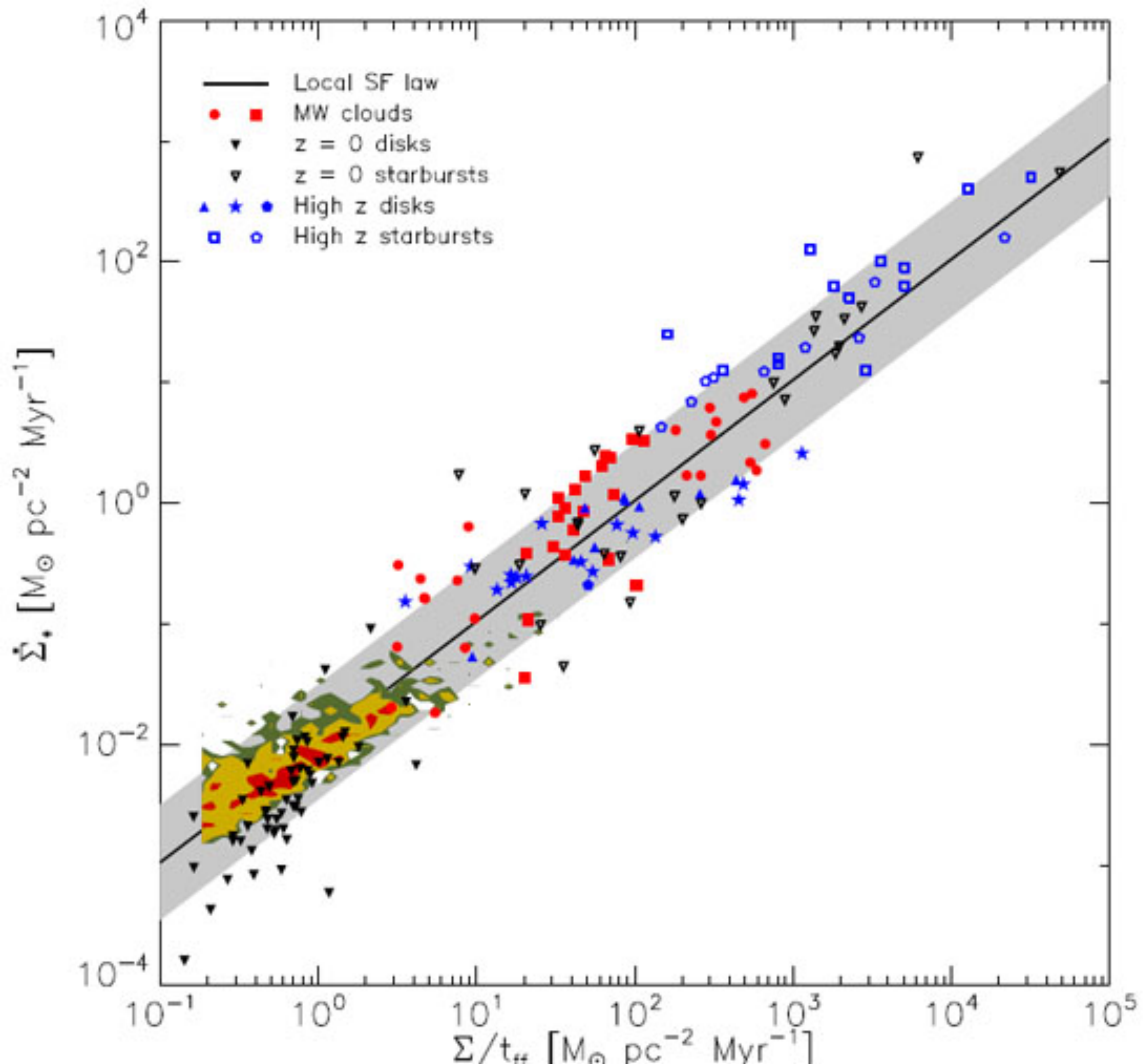
COLOR - MAGNITUDE

The color magnitude diagram shows that star formation history (color) correlates with mass. It also shows that the two type classification isn't that bad, since there are two loci in this distribution. Massive galaxies are red, elliptical and old.



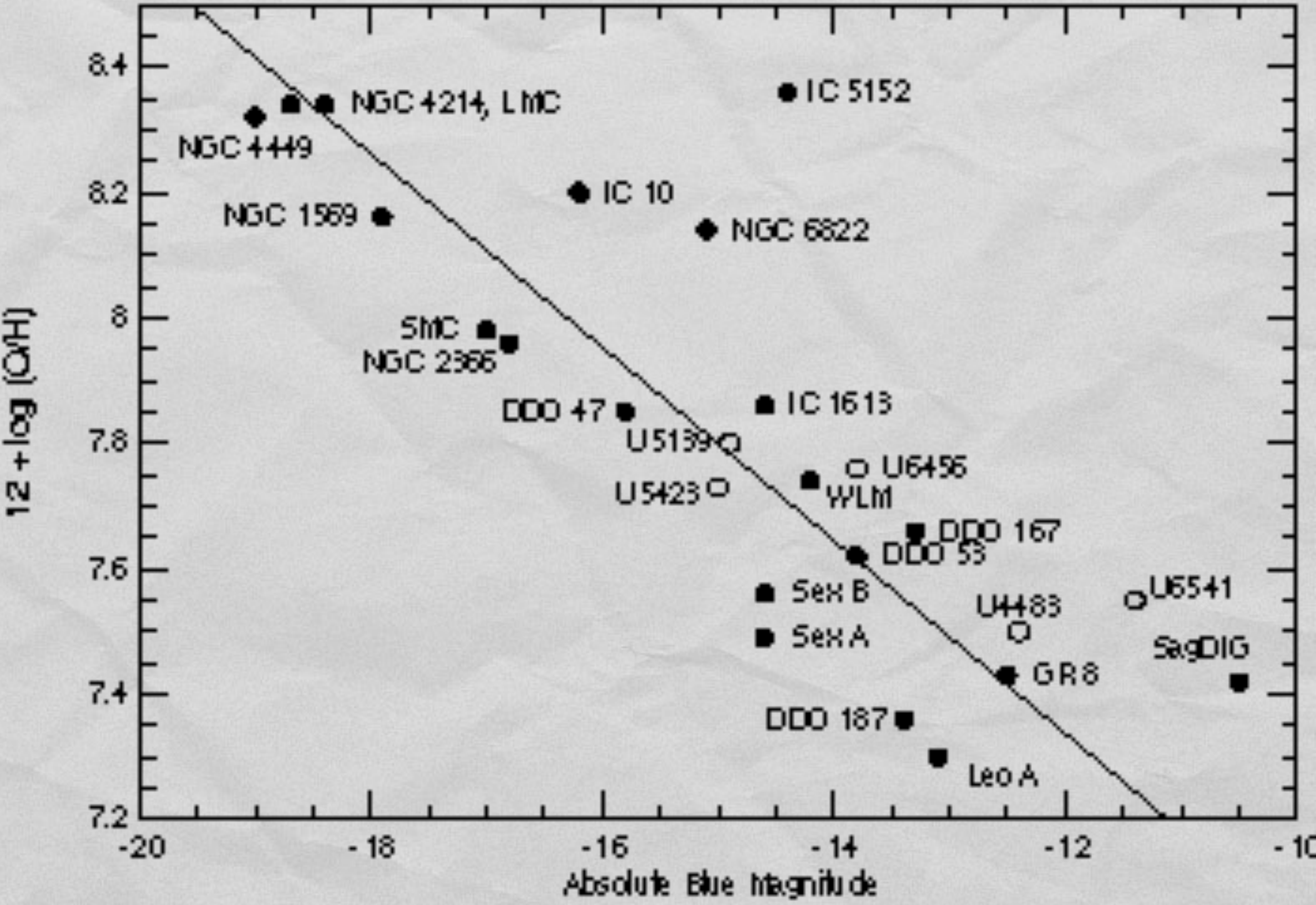
KENNICUT - RELATION

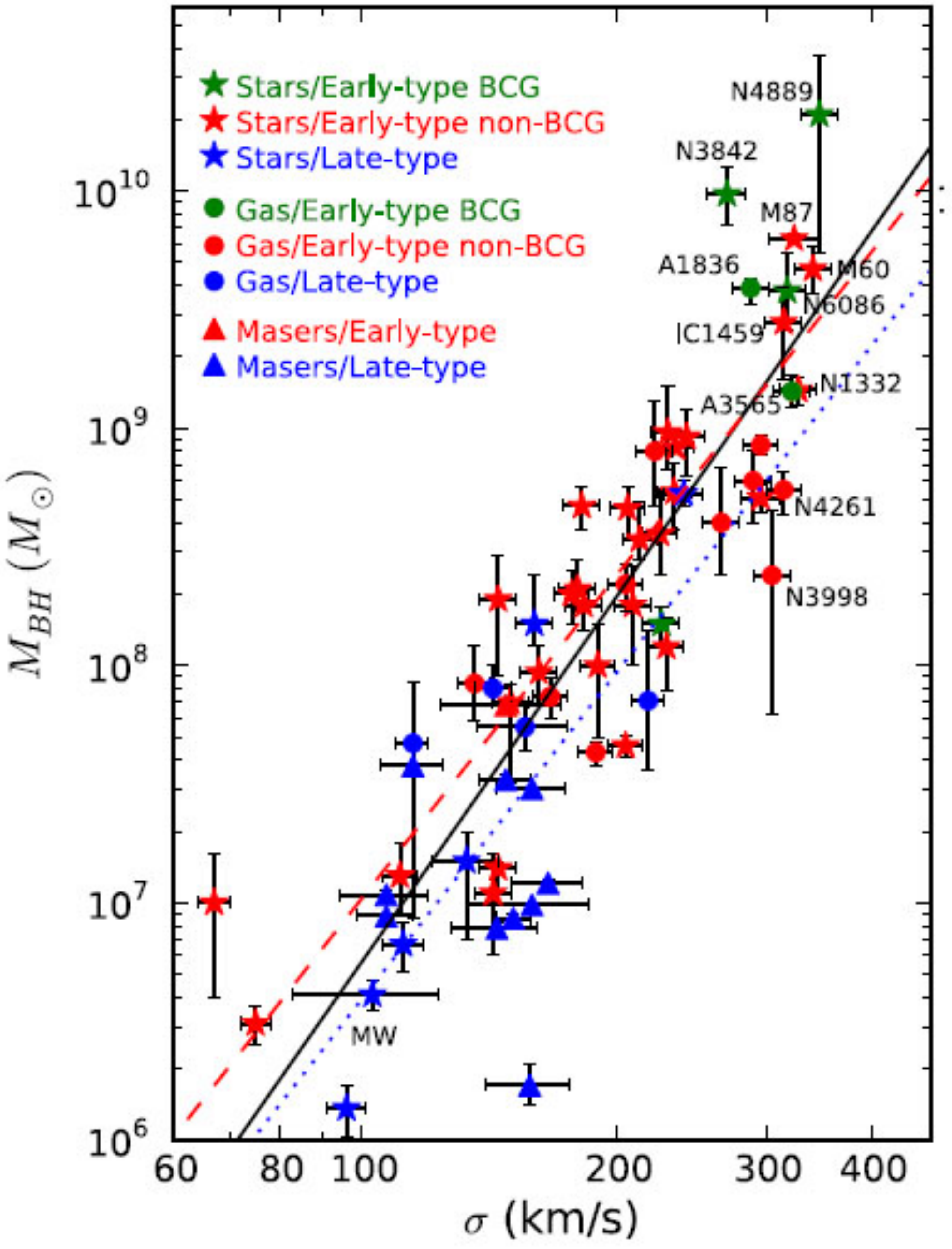
The surface density of star formation shows a tight relationship with the surface density of gas. Not entirely unexpected, since stars form from gas, but why surface density and why the power law?



MASS - METALICITY

Less massive galaxies have lower metallicities than high mass ones. They also have less stars, so this may not seem surprising, but they have a lower metallicity per star, not just overall.



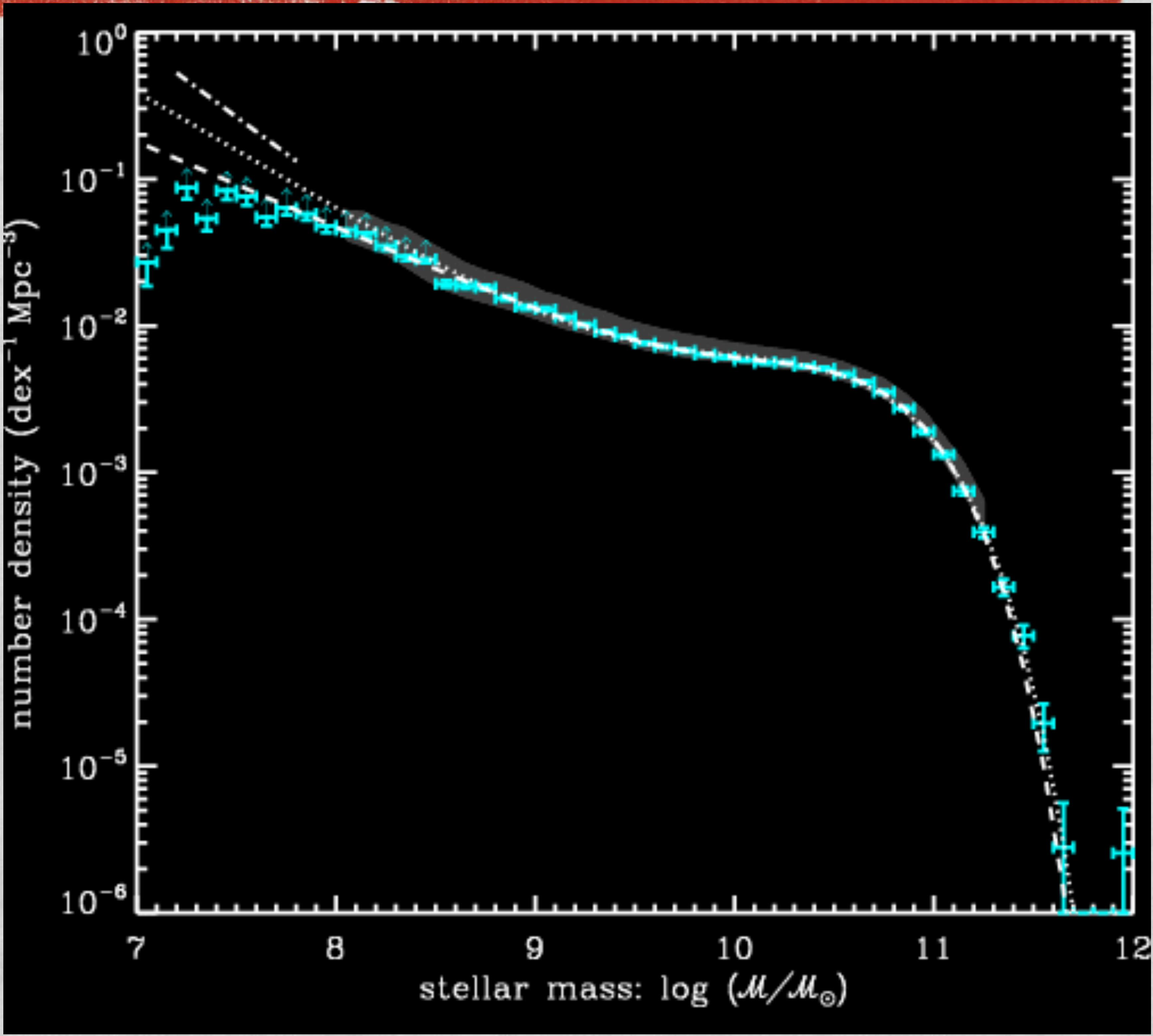


One of the most recently discovered relations is called the m-sigma relation, between the velocity dispersion of a galaxy's bulge and the mass of the super massive black hole in the galaxy.

This one is a little odd, that bigger black holes are in bigger galaxies makes sense, but why would the correlation be best with the bulge velocity dispersion instead of the galaxy's rotation velocity or stellar mass

GALAXY STELLAR MASS FUNCTION

- The galaxy stellar mass function is the number of galaxies in a small mass range per unit volume.
- It shows an exponential cutoff at higher masses, how can we understand this?



THE BASIC MODEL OF GALAXY FORMATION

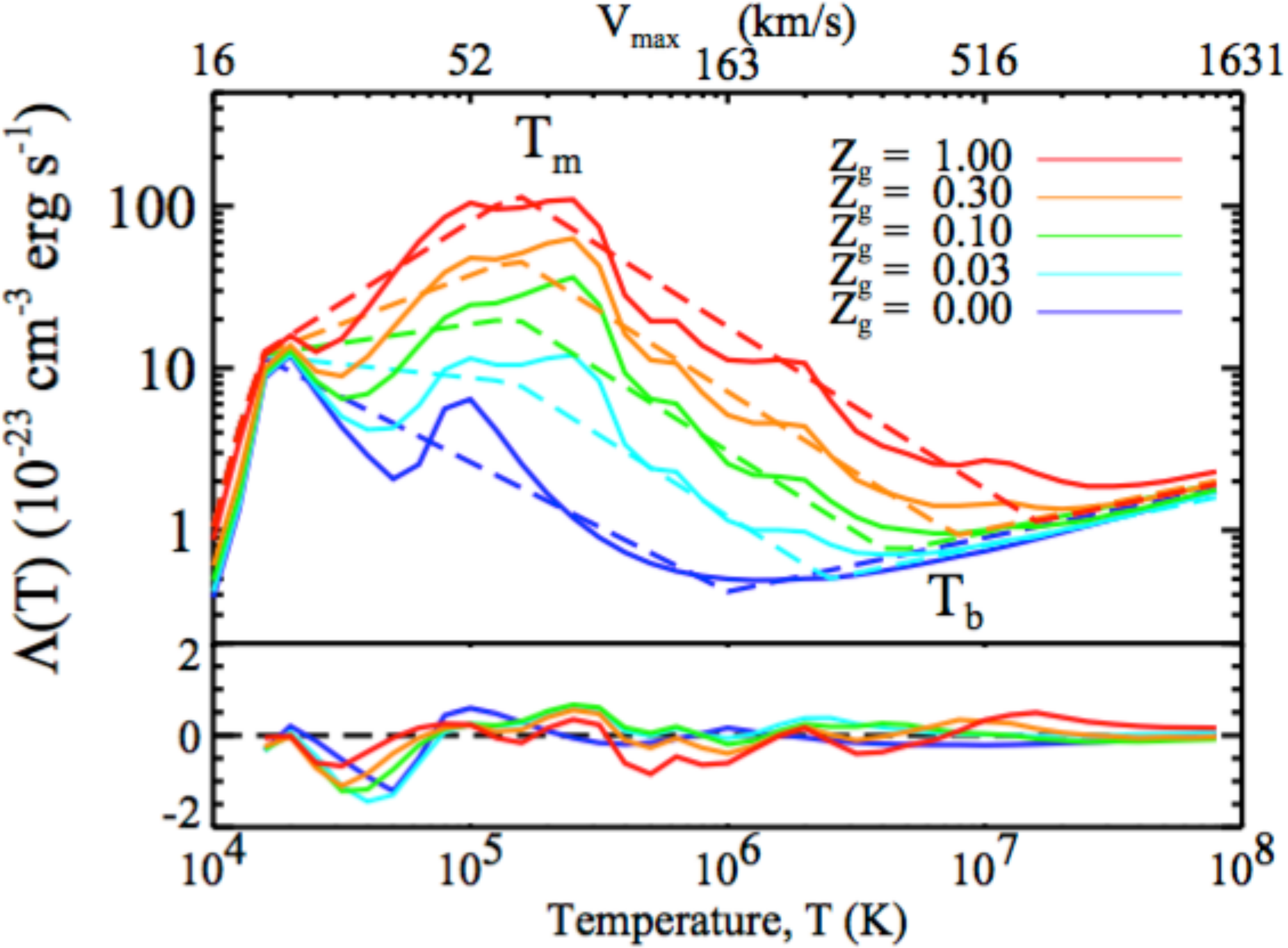
COOLING TIME

- How can we understand the exponential cutoff in galaxy masses. This question was addressed by Ostriker & Rees (1977), Silk (1977) and Binney (1977).

- They all made similar arguments that to form stars the cooling time of gas must be less than the dynamical time.

- A plasma of temperature T and number density n has a total energy of E~nkT. The rate at which it looses energy by cooling is $dE/dt = n_e^2 \Lambda(T,Z)$, where $\Lambda(T,Z)$ is a cooling function that depends on temperature and metallicity of the gas. So the time it takes for the gas to cool is

$$t_c = \frac{\frac{3}{2}\mu_e m_p k_B T}{n_e \Lambda(T,Z)}$$


V_{max} (km/s)
16
52
163
516
1631
T_m
T_b
Z_g = 1.00
Z_g = 0.30
Z_g = 0.10
Z_g = 0.03
Z_g = 0.00
$\Lambda(T)$ (10^{-23} cm^{-3} erg s^{-1})
100
10
1
2
0
-2
10^4
10^5
10^6
10^7
10^8
Temperature, T (K)

COOLING TIME

$$t_c = \frac{\frac{3}{2}\mu_e m_p k_B T}{n_e \Lambda(T,Z)}$$

- At high temperatures the cooling function goes to a power law and $\Lambda(T,Z) \sim T^{1/3}$. Thus if the gas densities are roughly the same $t_c \sim T^{2/3}$ and eventually there is so much energy in the gas that it can not cool in a dynamical time or even a Hubble time.
- In this way we can understand why clusters are mostly hot gas and why there is a sharp fall off in the galaxy mass function.

THE HALO MODEL

- White & Rees (1978) extended this model by considering gas cooling in dark matter halos.
- The circular velocity of a dark matter halo can be turned into a temperature. So the halo mass function can be made into a gas temperature function.
- Using the cooling time argument this then gives a stellar mass (assuming all cold gas becomes stars) for every halo. This relationship is not linear.
- In this simple model the galaxy properties are solely determined by the current halo properties.

ANGULAR MOMENTUM

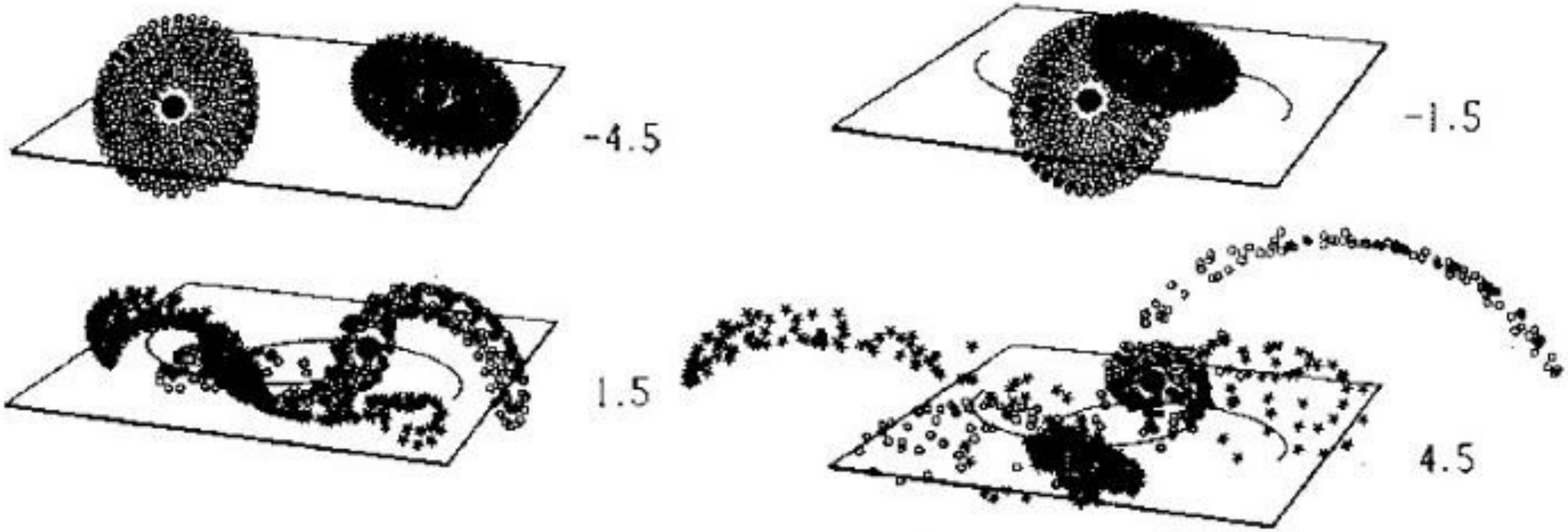
- The next important step in the classical model is that galaxy disk sizes are determined by angular momentum conservation.
- This idea (originally proposed by Mestel (1963)) was brought into the modern framework by Fall & Efstathiou (1980) who suggested that the angular momentum of a galaxy should be the baryon fraction of a dark matter halo times the angular momentum of a that halo.
- This gives disk sizes that are in pretty good agreement with observations.

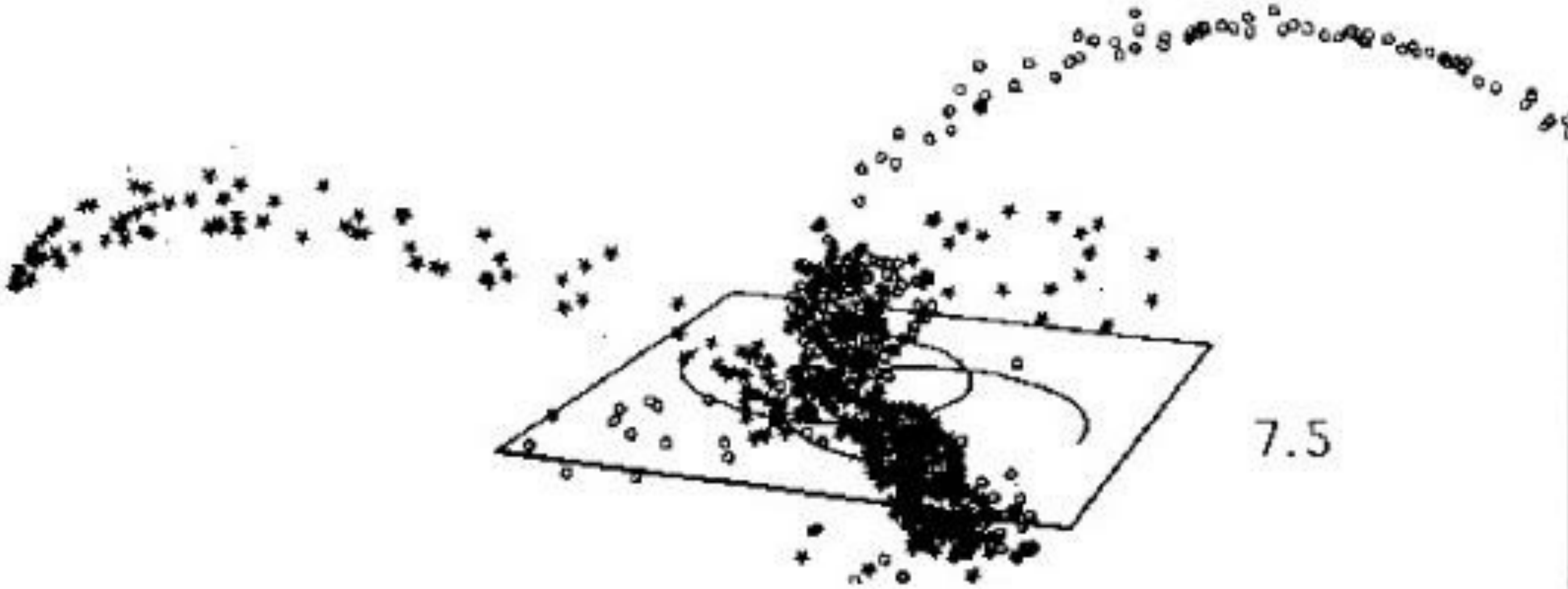
MERGERS

- The final piece of the basic picture is that when galaxies merge, disks are turned into spheroids.
- This was demonstrated in simulations by Toomre & Toomre (1972).
- It also makes intuitive sense, disks have low entropy, spheroids have higher entropy.

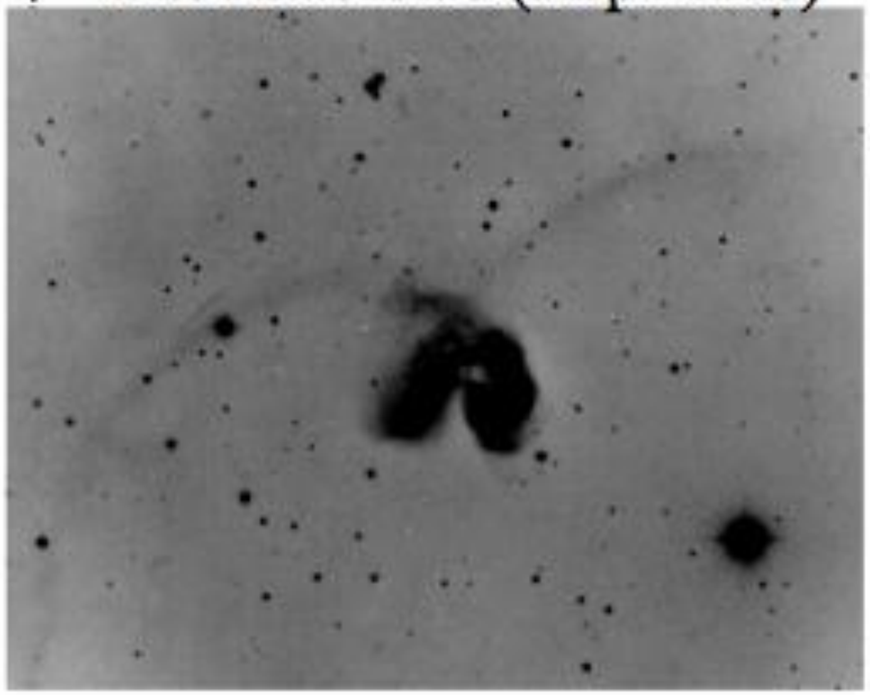
Numerical Simulations of Encounters

A. Toomre 1974





The Antennae (Arp Atlas)

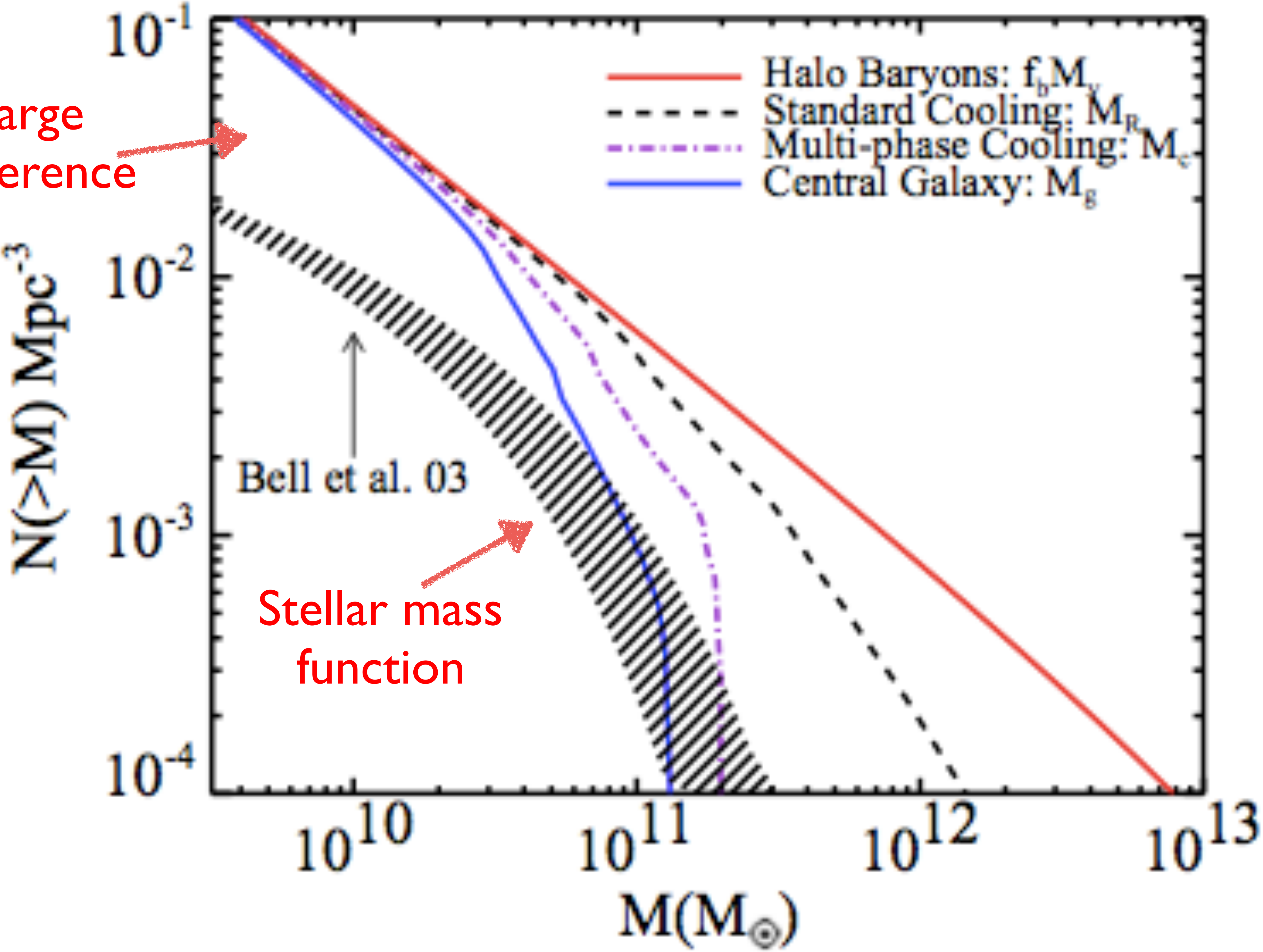


THE BASIC MODEL

- Gas cooling in dark matter halos sets the stellar mass of galaxies.
- Angular momentum conservation sets the sizes of disk galaxies.
- Mergers transform disk galaxies into spheroids.

THE PROBLEM WITH THE BASIC PICTURE

- The basic picture leaves many details to be worked out. However, one issue that immediately comes up is that just based on cooling the predicted stellar masses are way off.
- The solution to this is generally thought to be some form of additional energy that heats or removes gas from dark matter halos. This is generically called feedback because the energy most likely is a by product of the galaxy formation process.



large difference

Stellar mass function

SUPERNOVA FEEDBACK

- The first feedback solution proposed was by Dekel & Silk (1986). They argued that supernova would drive winds (Larson 1974) that would expel gas in low mass galaxies.

- This also could explain why low mass galaxies have low metallicities as the metals from the supernova would also be ejected.

- Trying to implement this in numerical simulations it looks like this may work in small galaxies, but that winds are not able to remove material in medium and large galaxies (Mac Low & Ferrara 1999).

OTHER FEEDBACKS

- Many other sources of feedback have been proposed:
  - The UV background (Efstathiou 1992) likely explains why low mass halos produce no stars at all.
  - Active galactic nuclei (Silk & Rees 1998) may keep massive galaxies from forming new stars. This feedback may also explain the m-sigma relation.
  - Other options, magnetic fields, cosmic rays, radiation pressure, have also been considered.
- Understanding the important sources of feedback and how they work is the main goal in galaxy formation.

TECHNIQUES

- There are three basic techniques used to model galaxy formation:
  - Analytic - One can give simple prescriptions that relate halo properties to galaxy properties. This has the advantage of simplicity, but also assumes history is unimportant.
  - Semi-analytic - This method uses a Monte-Carlo approach or N-body simulations to describe a halos merger history and then uses simple prescriptions for the gas physics.
  - Hydrodynamic Simulation - The cooling and motion of gas can be solved numerically, but star formation and feedback are always implemented subgrid, because the range of scales are just to large. The most believable method, but also the most difficult to explore new ideas with since simulations take ~3 or 4 months to run or longer.

ANALYTIC

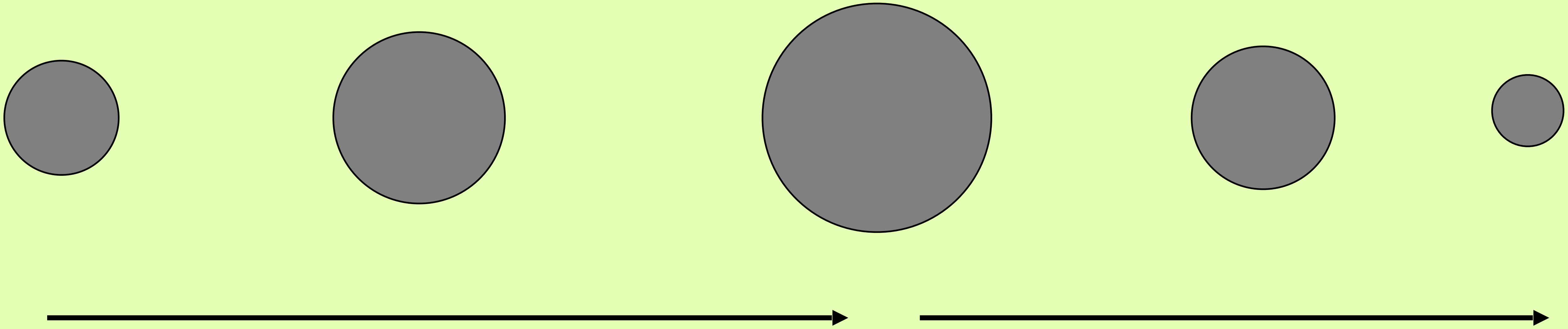
- Spherical Collapse Model
- Press-Schechter / Extended Press Schechter
- Disk Sizes

The Spherical Collapse model

We have examined the growth of perturbations, but what happens when the perturbations are no longer small. That is δ approaches 1.

We can solve this for the simple case of a uniform overdense region. This can be called a spherical top hat which basically just means a uniform over density over a spherical region. While not really physical this will give us a sense of what happens.

What will happen is that the perturbation will grow denser slowing its expansion. It will still expand, but eventually gravity will overcome the Hubble expansion and then the region will collapse.



Hubble expansion wins

Collapse of density region

The Spherical Collapse model

The equation for the evolution of a spherical over density is identical to that of a matter only universe as we have seen earlier.

$$\ddot{R} = -\frac{GM}{R^2}$$

The solution can be give by:

$$R(\theta) = A[1 - \cos(\theta)]$$

$$t(\theta) = B[\theta - \sin(\theta)]$$

(the “cycloid” solution)

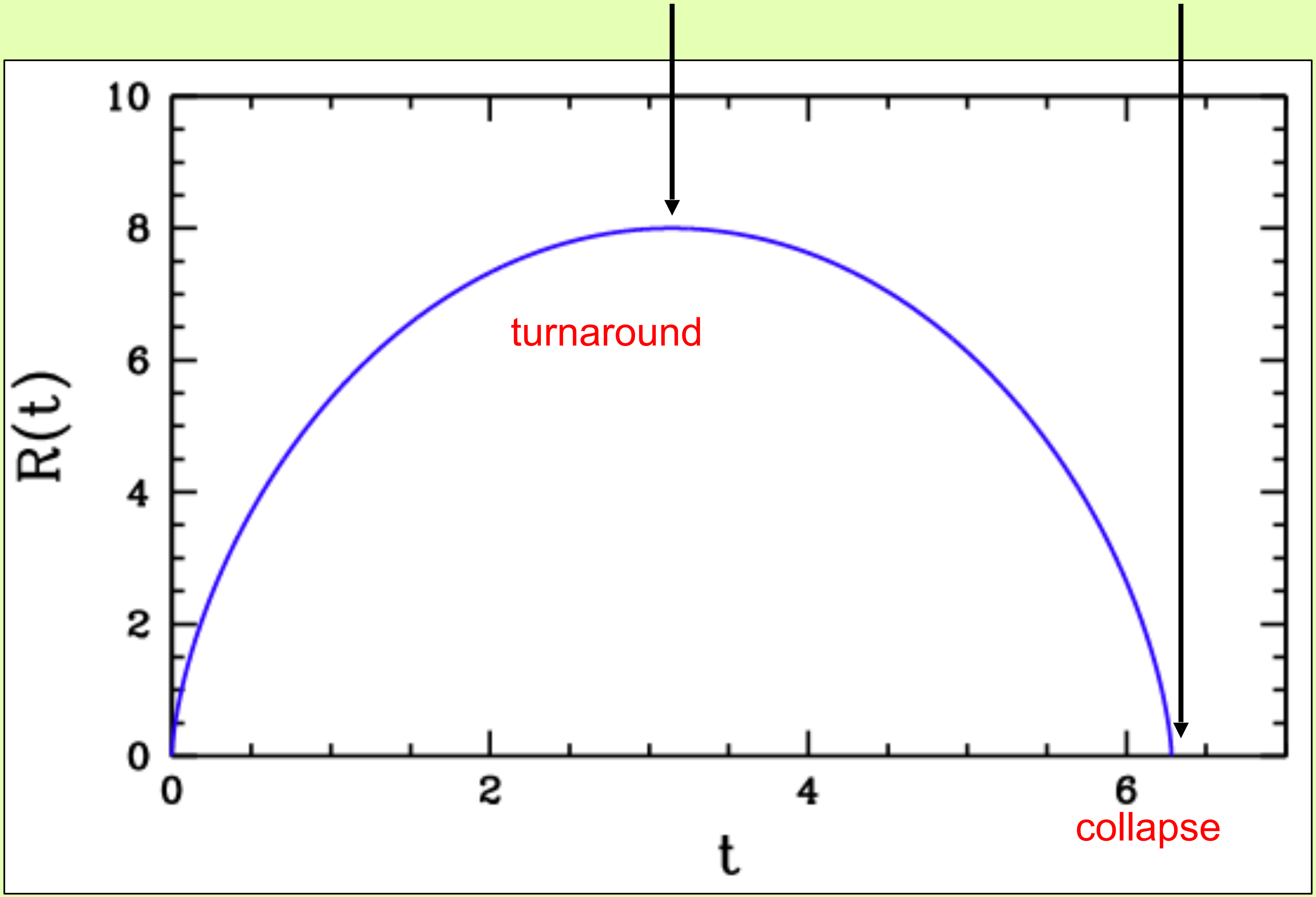
Where, $A^3 = GMB^3$

If describing the whole universe θ is the conformal time, but for a spherical region it is just a parameter.

The Spherical Collapse model

$\theta = \pi$

$\theta = 2\pi$



turnaround

collapse

The Spherical Collapse model

Expand and only keep low order terms:

$$R = A\left[1-\cos(\theta)\right]$$

$$t = B\left[\theta-\sin(\theta)\right]$$

$$\cos(\theta) \approx 1-\frac{\theta^2}{2}+\frac{\theta^4}{24}-\ldots$$

$$\sin(\theta) \approx \theta-\frac{\theta^3}{6}+\ldots$$

$$R \approx A\left[\frac{\theta^2}{2}-\frac{\theta^4}{24}\right] = \frac{A}{2}\theta^2\left[1-\frac{\theta^2}{12}\right]$$

$$t \approx B\left[\frac{\theta^3}{6}\right] \qquad \rightarrow\ \theta \approx \left(\frac{6t}{B}\right)^{1/3}$$

$$\boxed{\rightarrow\ \mathrm{R} \approx \frac{A}{2}\left(\frac{6t}{B}\right)^{2/3}\left[1-\frac{1}{12}\left(\frac{6t}{B}\right)^{2/3}\right]}$$

The Spherical Collapse model

$$R(t) \approx \frac{A}{2}\left(\frac{6t}{B}\right)^{2/3}\left[1 - \frac{1}{12}\left(\frac{6t}{B}\right)^{2/3}\right]$$

Compare this to our previous linear theory result:

$$R(t) \approx a(t)\left[1 - \frac{1}{3}\delta(t)\right]$$

where:

$$a(t) = \left(\frac{3}{2}H_0 t\right)^{2/3}$$

and:

$$\delta(t) \propto t^{2/3}$$

The cycloid solution at small t agrees with linear theory.

The Spherical Collapse model

Turnaround

The sphere breaks away from general expansion and reaches a maximum radius at $\theta=\pi$. At this point, linear theory predicts that the density contrast is $\delta_{lin}=1.06$.

Collapse

The sphere collapses to a singularity at $\theta=2\pi$. This occurs when $\delta_{lin}=1.69$.

Virialization

Complete collapse never occurs in practice because the kinetic energy of collapse is converted into random motions. When the sphere has collapsed to half its maximum size, its kinetic energy is $K=-0.5U$, where U is the potential energy. This is the condition for equilibrium according to the virial theorem. This occurs at $\theta=3\pi/2$ when the density contrast is $\delta_{lin}=1.58$.

The Spherical Collapse model

If virialization occurs at 3π/2:

$$1+\delta_{\mathrm{vir}} \equiv \Delta_{\mathrm{vir}} = \frac{\rho}{\bar{\rho}} \approx 147$$

If virialization occurs at 2π:

$$1+\delta_{\mathrm{vir}} \equiv \Delta_{\mathrm{vir}} = \frac{\rho}{\bar{\rho}} \approx 178$$

More generally:

$$\Delta_{\mathrm{vir}} \approx 178\,\Omega_m^{-0.7}$$

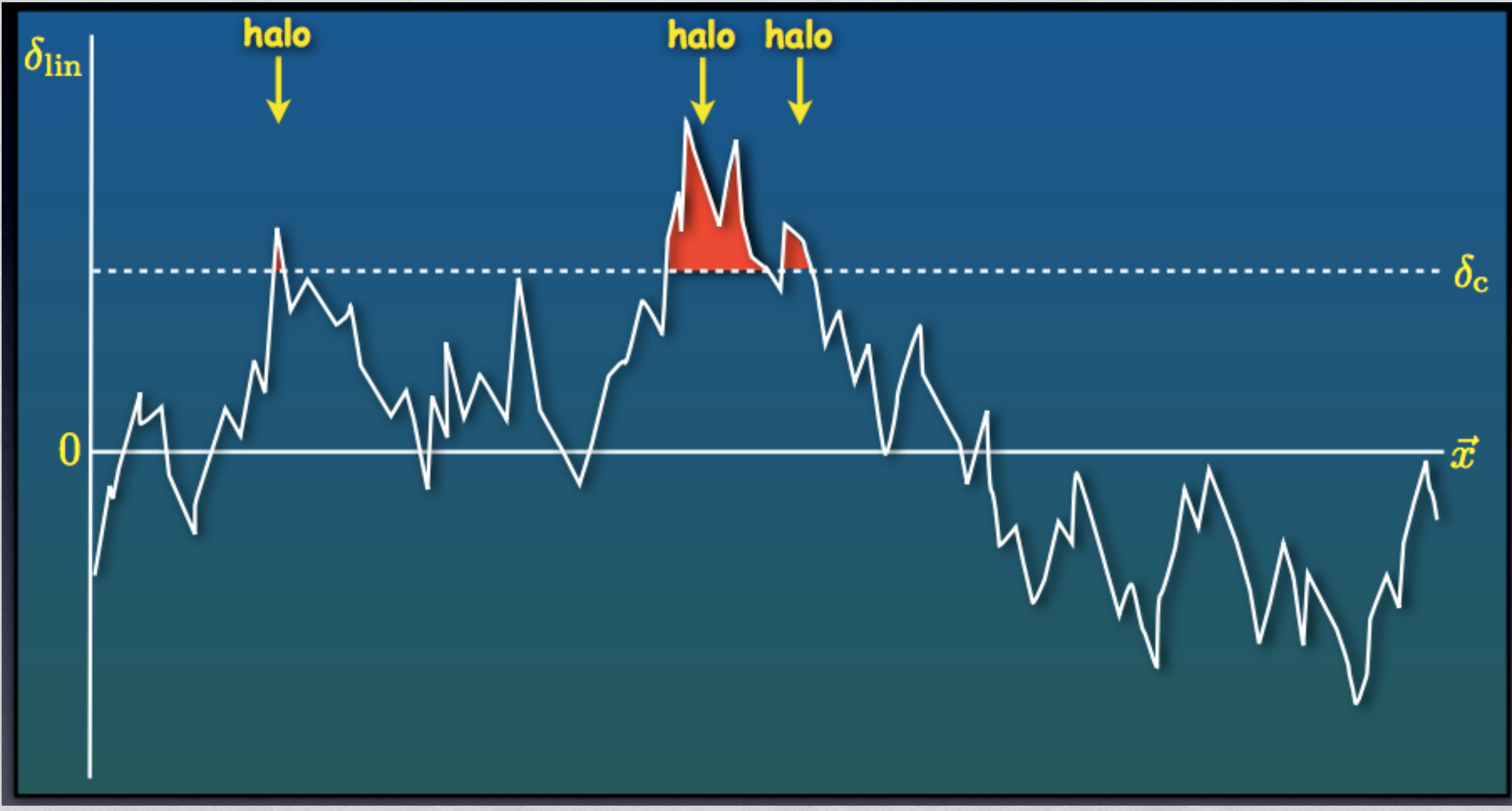
Even more generally (for flat matter + dark energy models):

$$\Delta_{\mathrm{vir}} \approx \left[18\pi^2 + 82\left(\Omega_m - 1\right) - 39\left(\Omega_m - 1\right)^2\right]\Omega_m^{-1}$$

Bryan & Norman (1998)

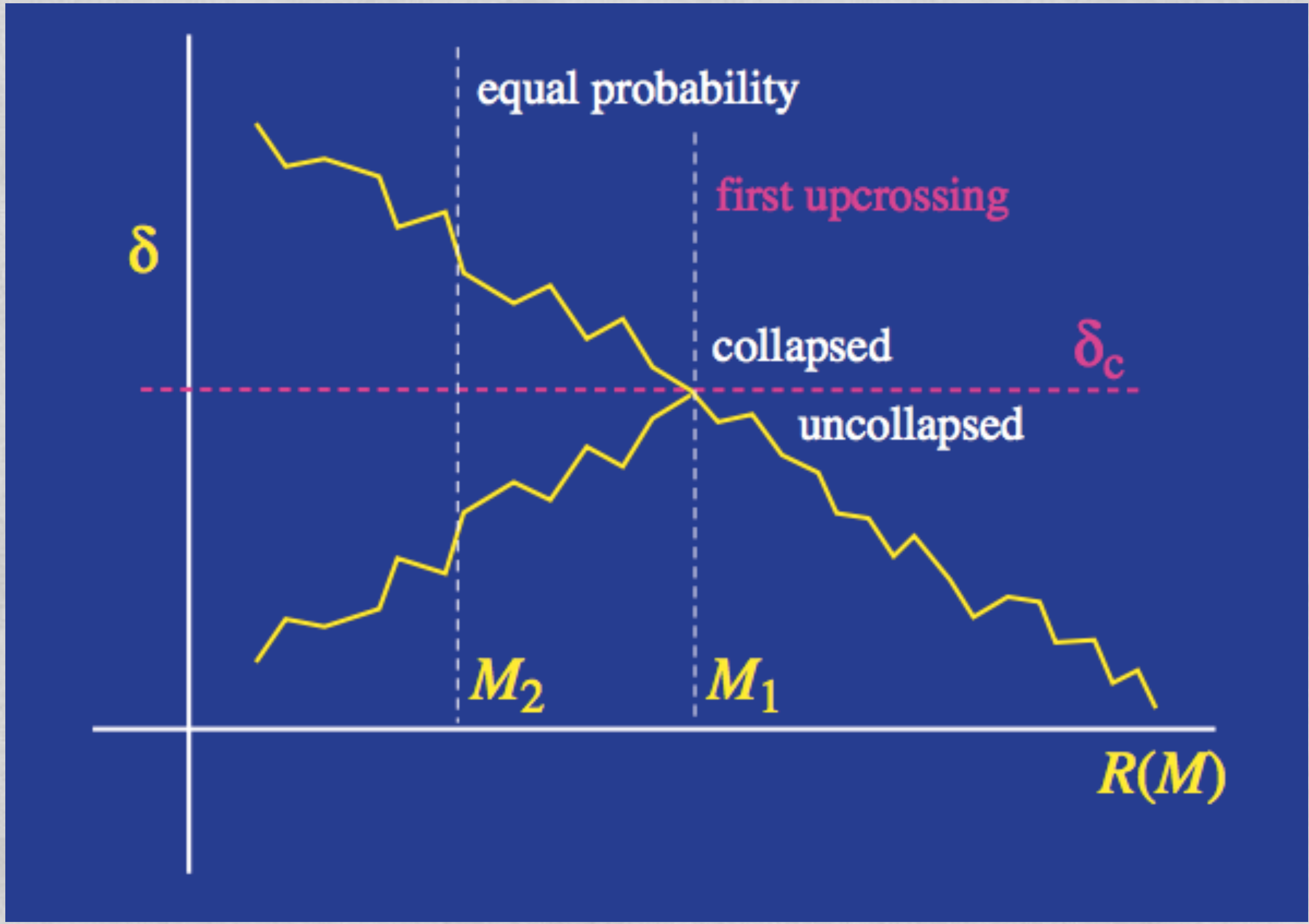
EXTENDED PRESS SCHECHTER

- As we've seen before, the spherical collapse model can then be used to estimate the halo mass function by exploring the variance and asking how often you get a mass fluctuation greater than δ_{crit}, where δ_{crit} comes from the spherical collapse model. This is called Press-Schechter.
- This idea was extended and improved later on with the realization that within every mass (or spatial) scale lies all smaller scales. So while a region will have a value δ_1 on one scale it has values δ_2,δ_3,δ_4 on different scales. However, these values will be correlated since the mass depends on the average overdenstiy which combines all smaller scales.


halo
halo
halo
$\delta_{\rm lin}$
$\delta_{\rm c}$
0
$\vec{x}$

EXTENDED PRESS SCHECHTER

- With the right filter the gaussian nature of the fluctuations gives that the value of δ will be a random walk.



Some objects of mass M_1 will have just collapsed. Some had a region in them that was already collapsed of mass M_2. In this way one can build a merger tree. Note that it is purely statistical. If one wants to start with some choice of perturbations then evolving them into the future requires simulations.

DISK SIZES

- Finally we can build an analytic model of disk sizes based on the assumption of specific angular momentum conservation (Fall & Efstathiou 1980).
- Angular momentum must be conserved. Specific angular momentum conservation means that each mass element individually conserves its angular momentum, there is no angular momentum exchange.
- If this is the case then the angular momentum per unit mass of the dark matter and baryons are the same. If we know the specific angular momentum of the baryons and assume a profile then we get a size.

TIDAL TORQUES

- Why would a halo have nonzero angular momentum?
- The gravitational field around the collapsing region is likely to not be spherically symmetric. If there is a large mass concentration in one direction that will cause a tidal force over the collapsing region.
- At the turn around radius the region will experience the maximum torque as the lever arm is longest. This will still continue at collapse happens, but get weaker as the halo gets smaller.

SPIN

- We can define a dimensionless spin for the halo. The was first done in terms of the binding energy by Peebles in 1969

$$\lambda = \frac{J\sqrt{|E|}}{GM^{5/2}}$$

- but a more useful definition is just in terms of the relevant quantities and is often used now days

$$\lambda = \frac{J}{\sqrt{2}M_h R_h V_h}$$

- the two are the same if the density profile is SIS.

GALAXY SIZE

- So basically the idea is that original spin of the dark matter and the baryons is the same (at turnaround say).
- The baryons collapse to some much smaller size R_d, because of cooling, until angular momentum conservation halts the collapse. So if $\lambda_b = \lambda_{dm}$ then

$$R_d \approx \lambda_{dm} R_h$$

- One can try and make this more accurate by taking into account the density profile of the dark matter and the disk, the reaction of the dark matter contraction of the gas, etc.

ANALYTIC MODELS

- One can even then try and make a fully analytic model, by starting with angular momentum to get the gas surface density and then having a formula for how gas surface density is converted into stars.
- That gives the star formation rate in your disk which with an analytic model for feedback can give you the total stellar mass as a function of radius in your disk.
- Note that in this type of model because there are no mergers no spheroids are formed. This type of model can only be used for disk galaxies.

SEMI-ANALYTIC MODELS

- The idea of semi-analytic models is to combine the conceptual simplicity of analytic models with the stochastic merger history of halos.
- This can give a full model of galaxy formation, that is computationally cheap. The main limitation is that you can only get out what you put in.

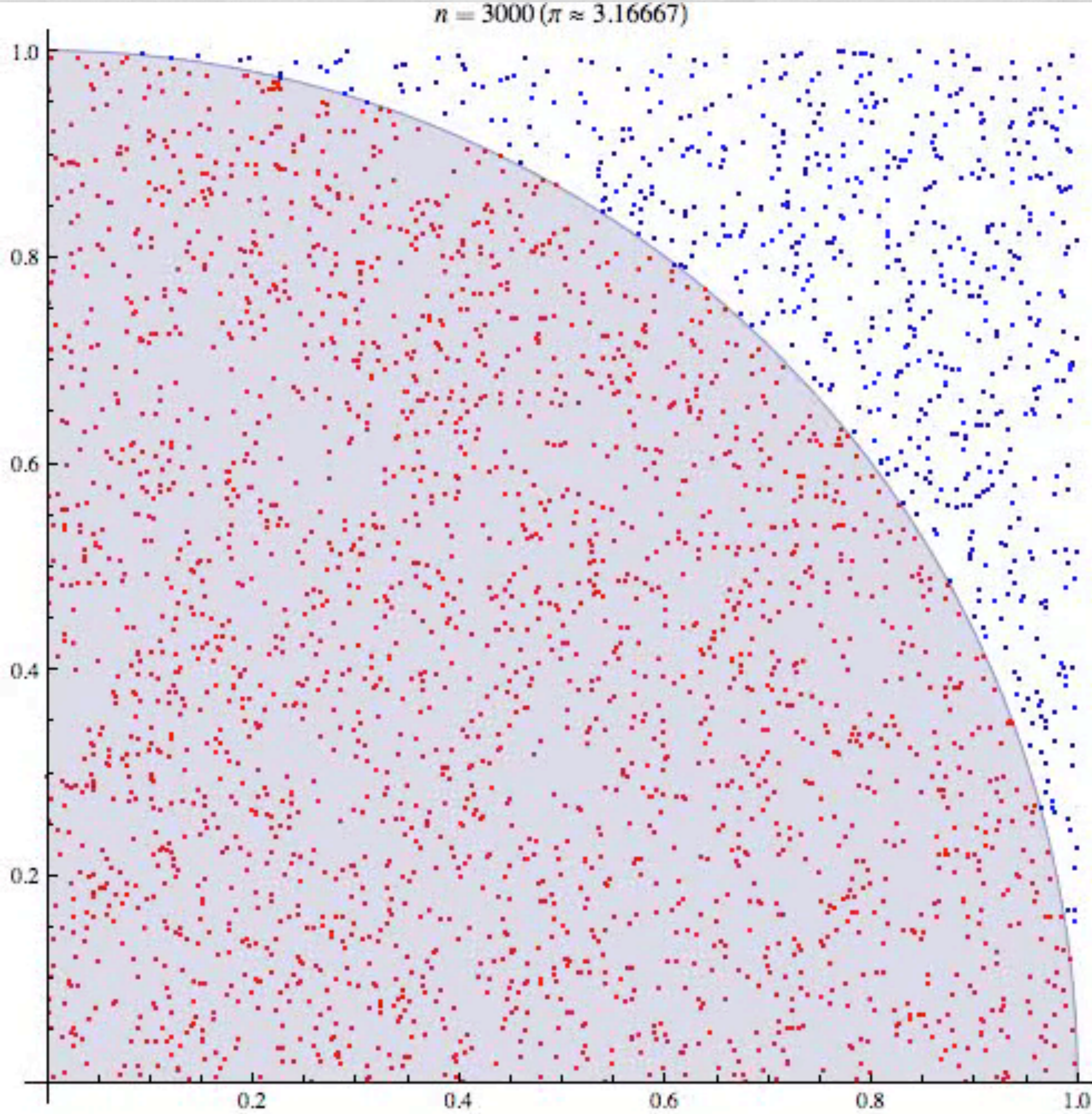
MONTE-CARLO

- The semi-analytic method should more appropriately be called a Monte-Carlo technique, one of the most powerful techniques in computer science.
- In Monte-Carlo one reverses the normal way of thinking about a physics problem. Instead of starting with initial conditions, solving some equations and then adding errors to get a probability distribution one starts with a probability distribution.

MONTE-CARLO

- This technique was developed by Ulam and van Neumann while working on the Manhattan project.
- The idea first came to Ulam while playing solitaire. Trying to solve a probability question that turned out to be hard he realized it would probably be easier to just deal 100 hands and see what happens.
- This is the basic idea of Monte-Carlo, you determine what will happen by running many realizations of your problem. As the number of realizations gets large, you get very accurate results.

A simple illustration is a way to calculate π. Just draw a circle on a square. Then drop things randomly on the square and count the number in the circle compared to the total. That number becomes very close to π as the number of points gets large.

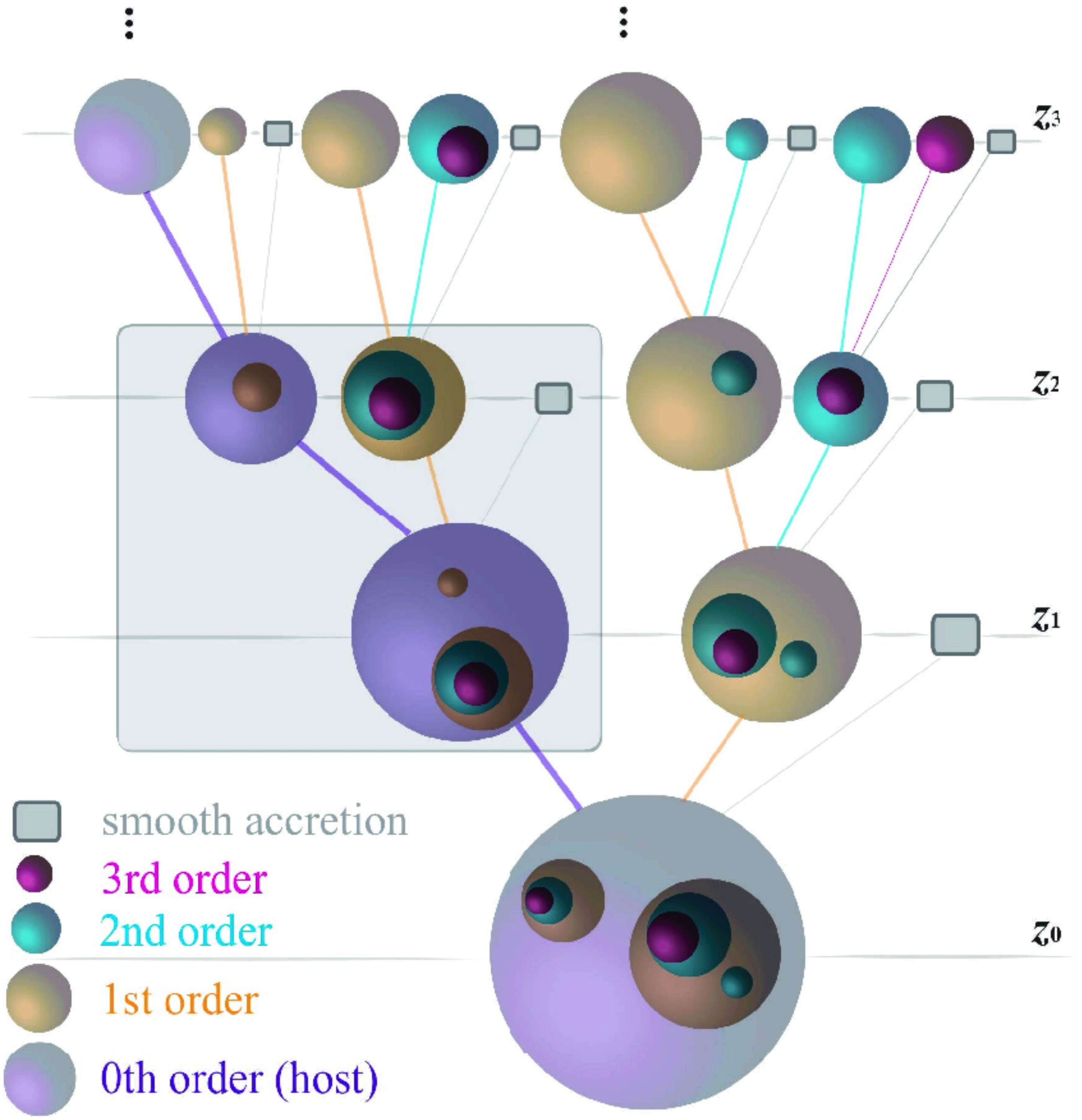


SEMI-ANALYTIC MODELS

- In semi-analytic models the basic randomness comes from the varying merger histories of galaxies. This is not included in analytic models.
- This stochasticity means that two halos of the same mass and angular momentum today can have differing galaxies inside them.
- There are a lot more steps though then just adding merger histories.

SEMI-ANALYTIC RECIPE

1. Merger Histories
2. Gas Cooling
3. Gas Disk Sizes and Surface Densities
4. Star Formation
5. Supernova Feedback
6. Supermassive Black Hole Formation and AGN Feedback



A dark matter merger tree can come from N-body simulations or from extended Press-Schechter. Now days there is no reason to use EPS. Besides the merger history there is also the issue of the location of the merged structure in the larger structure. This may differ from N-body because stars will change the dynamics.

MERGER TREES

- While merger tress seem straightforward there is actually a lot of details in the implementation.
    - When objects are close to merging they tend to mess up your algorithm for halo finding.
    - One only saves a finite number of time steps so the course time resolution effects what and when you call something a merger.
    - While a merger tree is a nice picture in reality things merge then come apart then merge then come apart. Things are pretty messy.
- Be aware that merger tress may vary based on definition and implementation.

GAS COOLING

- In each halo the gas is assumed to be heated to the virial temperature (the temperature of the dark matter in the halo) with a density profile that differs from code to code.
- Then the cooling time argument is used, but now as a function of density. So at some radius, called the cooling radius the gas is low enough density that it doesn't cool. The rest of the gas cools and falls to the disk in a free fall time.

GAS SURFACE DENSITIES

- Specific angular momentum conservation is used to get a size and therefore surface density profile for the gas disk.
- A star formation rate is determined based on the observation that star formation rate is proportional to gas surface density and a star formation time scale.

STAR FORMATION

- Since we have a merging history star formation can also depend on mergers.
- Starbursts can happen when major mergers occur. This is just some quicker conversion of gas into stars.
- Some models have even tried only allowing star formation if there is a minor merger to show these happen enough to explain star formation.

SPHEROID FORMATION

- From the merger history we also know when galaxies merge.
- Some mass fraction is chosen to convert disk galaxies into spheroids.
- Later accretion can reform a disk. In this way galaxies get a range of disk to bulge ratios.
- When galaxies fall into another halo it assumed they stop accreting gas. In this way the satellite galaxies in a halo will be gas poor and eventually end up redder.

SUPERNOVA FEEDBACK

- When stars are formed the most massive ones will live shortly and then explode as supernova.
- The energy will go into the gas heating it. In a SAM the hot gas is either returned to the hot gas in the halo or may be ejected from the halo completely.
- Numerical simulations have a very hard time making this work in practice, but in the SAM this is not a problem.

METAL PRODUCTION

- When the supernova go off they also create metals (everything with atomic number greater than 7).
- Some of the metals may be ejected with the hot gas, some may be mixed with the cold gas.
- The metallicity of the gas will change the cooling rate as we say in the cooling curve diagram.

AGN FEEDBACK

- Since all attempts with just supernova failed to really work, it was realized that something else was needed.
- The basic problem is that many elliptical galaxies show no sign of star formation for many Gyr.
- In the models once star formation ends there will always be some new gas accreted (or returned from older stars) and there is no way to stop it from forming stars.

AGN FEEDBACK

- However we know massive galaxies (maybe all galaxies) have active nuclei too. These emit tremendous energy and we can see them in old elliptical galaxies too.
- So about 10 years ago everyone started including this in their models to make the feedback work.
- This feedback is very strong and does a good job of stopping star formation once galaxies become massive enough.
- It also can explain the black hole - bulge mass relation.

SUMMARY

- The semi-analytic technique allows for large numbers of galaxies to be modeled and compared to data.
- It gives insight into what physical processes are important in galaxy formation. And allows for quick testing of hypothesis.
- The technique suffers from allowing one to guess behavior that may be unphysical and using formula that may be incorrect.
- It is strongest when used in combination with numerical simulations.

COSMOLOGICAL HYDRODYAMICS

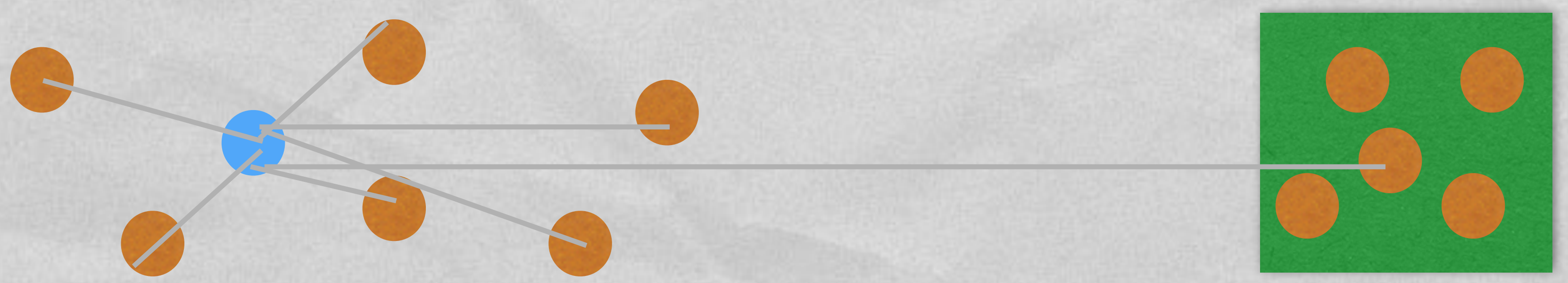
- A cosmological hydrodynamical code must do several things:
    1. Gravity Solver
    2. Gas Dynamics
    3. Photo-ionizing Background
    4. Gas Cooling
    5. Star Formation
    6. Feedback

N-BODY

- We want to solve the gravitational forces in an expanding universe. While formally this should require GR in practice one uses comoving coordinates that take care of the Universe's expansion and then solves Newtonian gravity on top of that.
- Even the largest simulations will only attempt to model a small part of the Universe, but our simulation will run into problems if there is a hard edge. So what is done is that the simulation is usually a periodic box so that particles on the right most edge feel the gravity of the left most edge.

N-BODY

- Gravity is usually solved using the N-body method, where mass is represented by particles.
- Direct force summation goes as N^2, this is bad. Being clever can reduce this to NlogN. Note that for a billion particles that is the difference between 10^{18} and 9×10^{9} or $\sim10^{8}$.
- This can be done by using the particles at short scales, but summing the particles at large scales.



N-BODY

- This is called the tree method, one uses the particles on the same branch, but just the overall mass for other branches.
- Alternatively one can convert particles into density on a grid (called PM for particle-mesh) and then Fourier transform the grid. In this way Poisson's equation becomes

$$k^2\Phi(k) = 4\pi G\rho(k)$$

- This makes for much faster computation, but of course can't resolve scales smaller than the grid.
- Now days the most popular technique is to combine the two so one has PM on large scales and tree on small scales.

N-BODY

- However, at very small scales we want to avoid two body interactions because the particles in our simulation do not represent actual objects. Thus the two body interactions would be unphysical.
- This is done by introducing a smoothing scale called the gravitational softening that weakens gravity at small scales. This limits the force resolution in a simulation.

$$F_g = G\frac{m_1 m_2}{r^2 + \psi}$$

HYDRODYNAMICS

- As we’ve seen, solving hydrodynamics means solving the continuity and Euler’s equation.
- There are two general approaches to this.
  - Lagrangian - particles represent the fluid elements and move with the fluid.
  - Eulerian - there are fixed cells and the fluid moves through the cell walls.
- There are three main techniques used in numerical fluid dynamics.

HYDRODYNAMICS

1. Smooth Particle Hydrodynamics (SPH) - particles flow with the fluid, forces are calculated by mean quantities over a fixed number of particles (~32). In dense regions you get high resolution. This technique is known to not perform as well in certain nonlinear situations (shock capturing, Kelvin-Helmholtz instabilities).

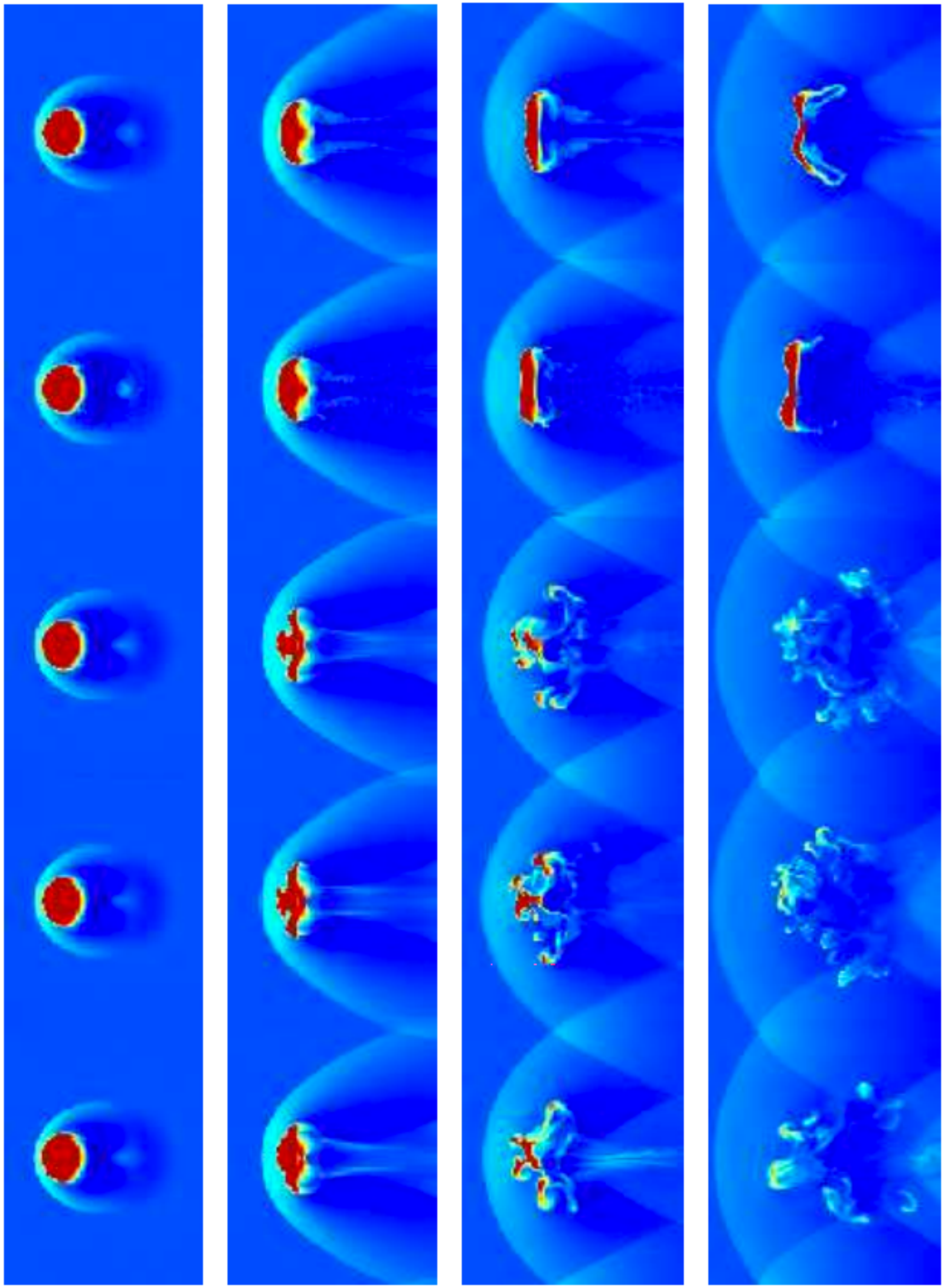
2. Adaptive Mesh Refinement (AMR) - this method employs a grid and the fluid moves across cell walls. The grid is refined in high density regions to achieve better resolution in those regions. One problem of this method is that you can't follow the time evolution of fluid elements unless tracer particles are added. This technique has difficulties when there are large scale fluid motions.

3. Moving Mesh (MM) - here a grid moves with the fluid with cells able to change volume to maintain roughly the same mass inside. This gives better velocity resolution.

SPH

- In SPH the particles carry quantities, but those quantities are not for an individual particle. They must be smoothed over some number of particles with a smoothing kernal.
- In ‘classic’ SPH the density is calculated first, then the thermal energy to get the pressure and the hydrodynamic acceleration.
- This however does not explicitly conserve energy. A variant that does called ‘entropy conserving’ SPH uses entropy as the evolved variable. This however has a side effect of creating an artificial pressure between hot and cold regions, making cold clumps resistant to disruption.

In SPH codes the cold clump remains after a few Kelvin-Helmholtz time scales. In the grid codes it is destroyed. This is blamed on an artificial surface tension created in SPH.



SPH

- This problem can be mitigated through a number of new techniques. One is to use a different kernel shape that uses more particles called SPHS, but requiring 10 times the particles increases the computational time.
- Pressure entropy SPH (PE-SPH) calculates the energy density and internal energy separately while still conserving entropy. Also including improvements in artificial viscosity goes a long way to addressing these problems.
- However, SPH remains problematic in may regimes.

GRID

- The other main way to solve hydrodynamics is to discretize the fluid onto grid cells. One then computes the advection properties of the fluid across the cell boundaries.
- With a fixed grid the main problem is that high spatial resolution requires many many cells. Adaptive Mesh Refinement (AMR) solves this problem by splitting the cells in certain regions so that higher resolution is only achieved in the region of interest.

GRID

- Hydro is solved by what is called a high-order Godunov scheme. The Riemann problem is solved across cell faces, which yields a pressure and then fluid is moved across the cell face.
- If the fluid is assumed to be uniform in the cell this is a first-order Godunov scheme. Higher orders can be reached by interpolating the fluid properties in the cell called Piecewise Parabolic Method (PPM).
- This gives more accurate results, but requires more cells and thus lowers the effective resolution.

GRID

- Eulerian methods do a better job and shock capturing and surface boundaries and with instabilities in general. If those are an important part of your problem you probably want to use a grid code.
- SPH has the advantage that you can track a particle through time. This is nice for talking about histories. The history of some mass is much harder to discuss with grid codes, though tracer particles can be used to overcome this problem.

MOVING MESH

- A third technique which tries to combine the strengths of the other two is to have a moving or deformable mesh.
- In this method the Riemann problem is solved on a mesh, but when the fluid is moved based on those forces the mesh is also moved or remade around the new fluid densities.
- In this way the calculation is good at shock capturing ,etc. but also has the advantage of Lagrangian behavior following the mass elements as they evolve and generating higher resolution in denser environments. The Arepo and Gizmo codes use this technique.

NUMERICAL METHODS

- Even when using the same general technique there are still many choices to be made in how the fluid equations are solved.
- This means different codes using the same techniques may not give exactly the same answers.
- Take Away - Don’t believe something just because someone ran a simulation, understanding the numerical technique and its weaknesses can be important.

HEATING AND COOLING

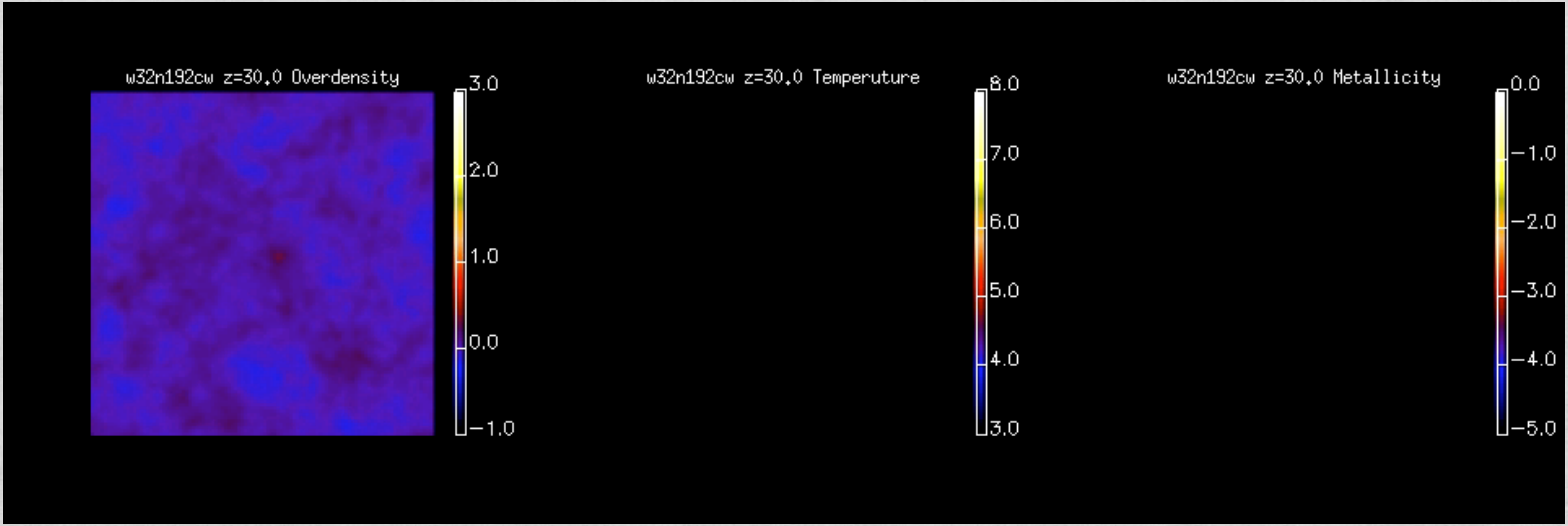
- The main difference between dark matter and baryons is that atoms can cool and be heated.
- Radiative cooling and heating play the major role besides the heating from expansion and shock heating.
- Conduction is usually not included in simulations in astrophysical plasmas its importance is poorly understood because magnetic fields can reduce the conduction rate by a factor of 10^{12}.

COOLING

- Radiative cooling must be included otherwise galaxies won't form.
- Metal line cooling is important which creates the difficulty that the rate of cooling depends on supernovas ejecting metals back into the gas.
- Stars form in gas that is at ~10K, and on scales of 0.1pc. Since these are very small scales cooling is often stopped at higher temperatures.
- People used to stop cooling at 10^4K, but now days most codes go to a few hundred K.

PHOTO-IONIZATION

- In cosmological simulations we know there is UV radiation that ionizes the Universe.
- This is usually put in by hand, because we don't exactly know the sources and including sources makes the calculation more expensive.
- Usually the fact that this background can be blocked in places is also ignored, instead the effect on gas with out shielding is only included.
- Simulations that study the epic of reionization must deal with the fact ionization propagates from sources. This can either be done in post processing or codes are just starting to try and include proper radiative transfer in simulations.



The issue of metal diffusion is very complex. We have to know how many heavy elements are produces in supernova, how they are ejected and mixed with the local medium and then how they are spread throughout the galaxy and into the intergalactic medium. Luckily we can observe the metallicity of intergalactic gas, so we can try and get this right.

STAR FORMATION

- Numerically star formation means replacing gas with a star particle (mass but no fluid dynamics).
- This is basically done by a density criteria and maybe some velocity and time constraints.
- Star particles are much more massive than individual stars and represent star clusters or multiple star clusters.

ISM

- The interstellar medium where stars form is not resolved in cosmological simulations. Stars form in giant molecular clouds that are of order 1pc in size.
- If gas is allowed to cool without limit this usually results in sever fragmentation and gas disks that fall apart.
- So instead extra pressure can be added to the ISM, supposedly this comes form stars or a multiphase medium.

BLACK HOLE GROWTH

- Finally many current simulations also include the growth of super massive black holes.
- These start as seed black holes with masses around 10^4 M$_\odot$. This is because we don't really know how these supermassive black holes get started and even if we did they would be to small to resolve.
- These black holes then usually grow according to Bondi accretion

$$\dot{M}_{Bondi} = \alpha \frac{4\pi G^2 M_{BH}^2 \rho}{(c_s^2 + v^2)^{3/2}}$$

- Finally black holes are merged when their separation is less than some value.

FEEDBACK

- The final ingredient to add is feedback and this is most of what has gone on in this field for the past twenty years.
- Even considering the same physics, like supernova, how to implement it can vary widely.
- Often tricks are used, like turning off cooling in the gas, or directly giving the gas momentum instead of energy.
- These are all made necessary because the actual scale feedback is occurring at is much much smaller than the scales that can be resolved in the simulation.

SUPERNOVA FEEDBACK

- In thermal feedback one just adds the energy from a supernova to the surrounding gas. However, because stars from in dense regions this energy is just radiated away.

- In blast wave feedback radiative cooling is shut off so that the gas can feel the high pressure and develop a large scale outflow.

- Kinetic feedback gives the gas momentum instead of energy which by construction forces it to move. Sometimes hydrodynamics is also shutoff so that the gas can move to large distances before hydro is turned back on.

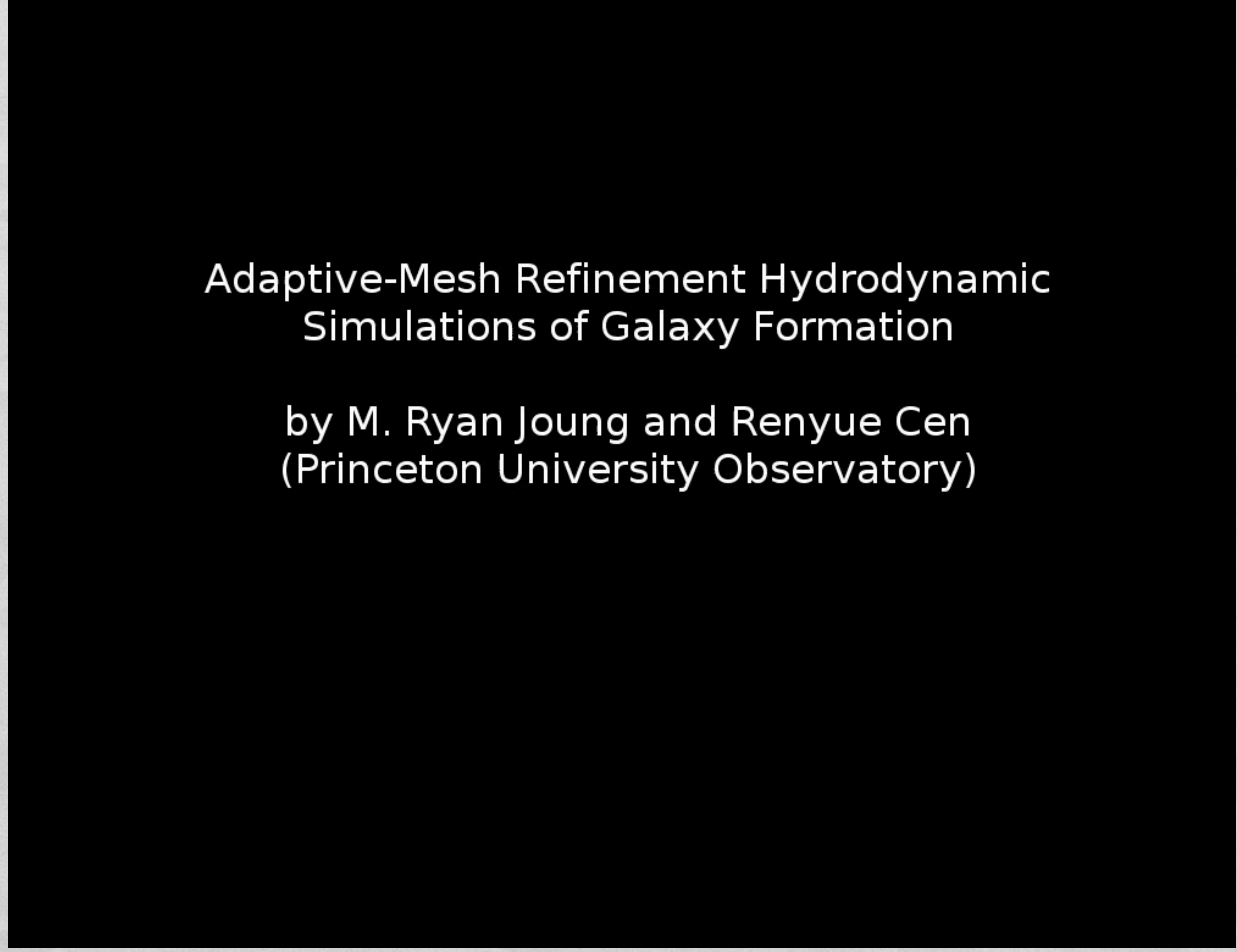
- Other attempts to make feedback work include giving the star particles a kick before they explode so that the feedback doesn't occur in dense regions and having winds from massive stars first heat the gas again making it less dense so that feedback is more effective.

AGN FEEDBACK

- Observations of AGN show large amounts of radiation, highly relativistic radio jets and gas outflows from the AGN. All of these should add energy to the gas, but realistically modeling any of it is still currently not possible.
- So AGN feedback is also implemented subgrid. It can be broadly classified as 'quasar mode' for radiative feedback or 'radio mode' for feedback from jets.
- Simulations try to implement both in simplistic ways. Note that one goal is to get the M_{BH}-σ relation.

ZOOM-IN

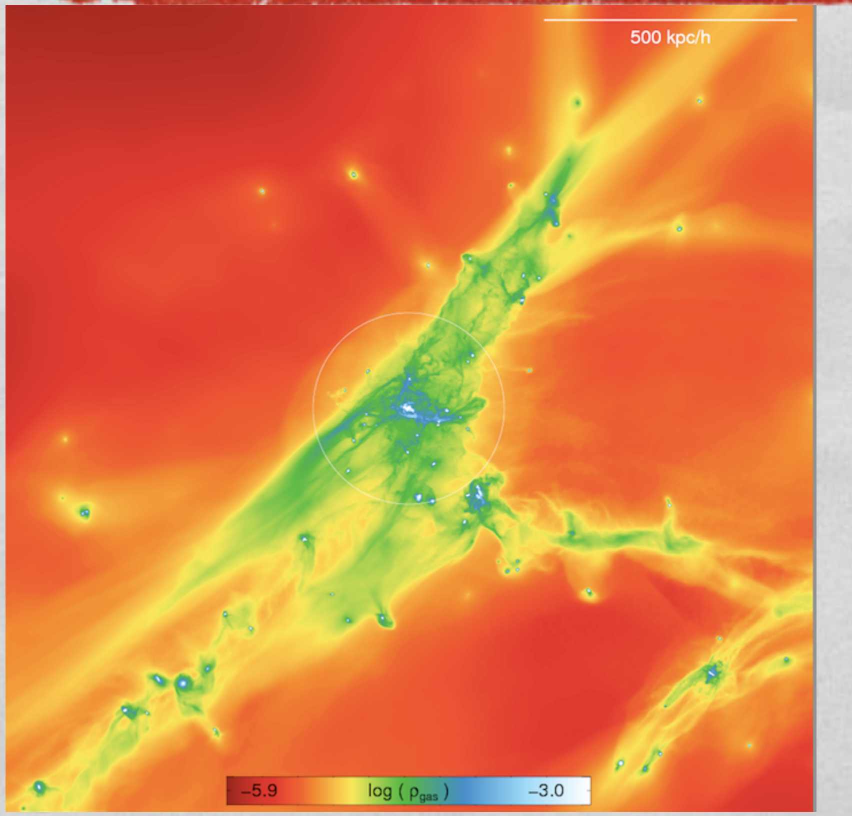
- The scales to be probed in galaxy formation are so vast that it is often best to do a zoom-in simulation.
- To do this, the gravity is calculated on cosmological scales, but the hydro is only calculated with resolution around one single galaxy.
- As long as galaxies evolve mostly independently this should be fine.



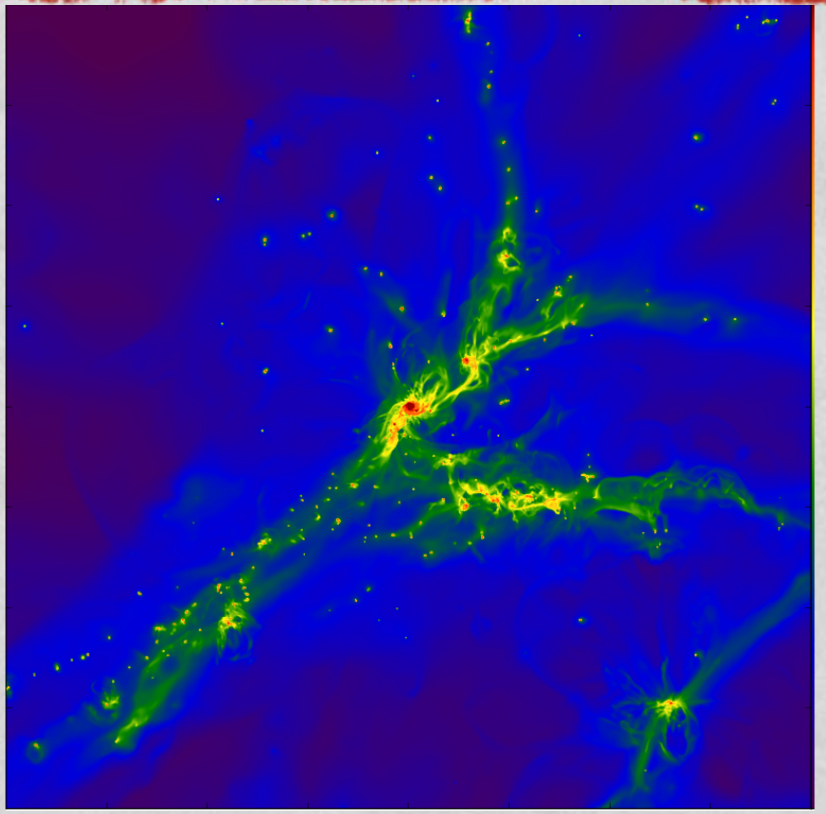
Adaptive-Mesh Refinement Hydrodynamic
Simulations of Galaxy Formation

by M. Ryan Joung and Renyue Cen
(Princeton University Observatory)

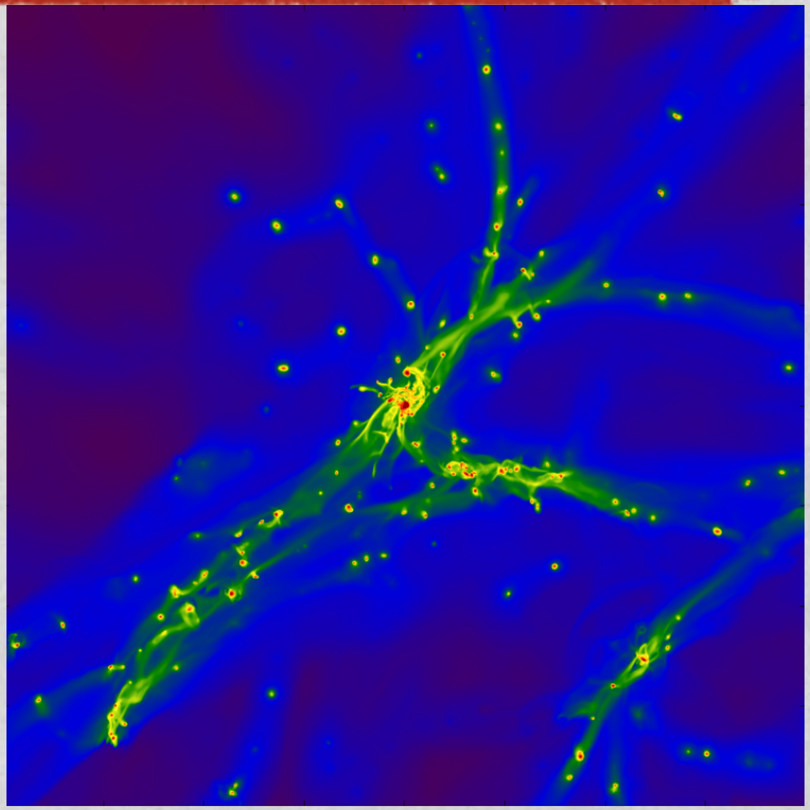
SCYLLA COMPARISON



MM-AREPO

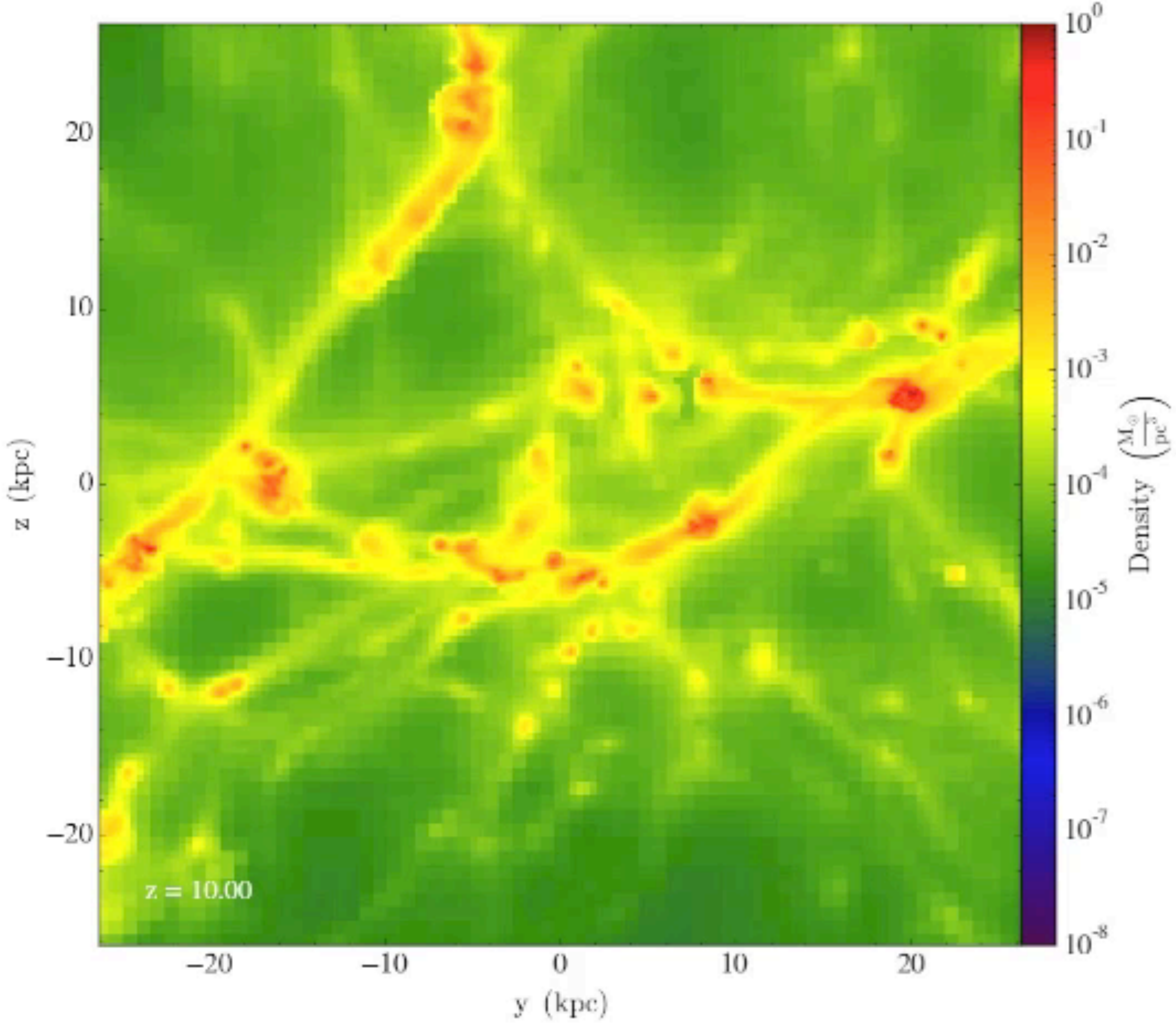


AMR-ENZO



AMR-RAMSES

In my own research we are trying to understand the differences that occur when using different codes. Our main goal is to separate the effects of different numerical techniques and different feedback prescriptions. It is clear that different codes give different results, but it takes a great deal of effort to understand why.


z = 10.00
z (kpc)
y (kpc)
20
10
0
−10
−20
−20
−10
0
10
20
10^0
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
10^{-8}
Density $\left(\frac{M_\odot}{pc^3}\right)$