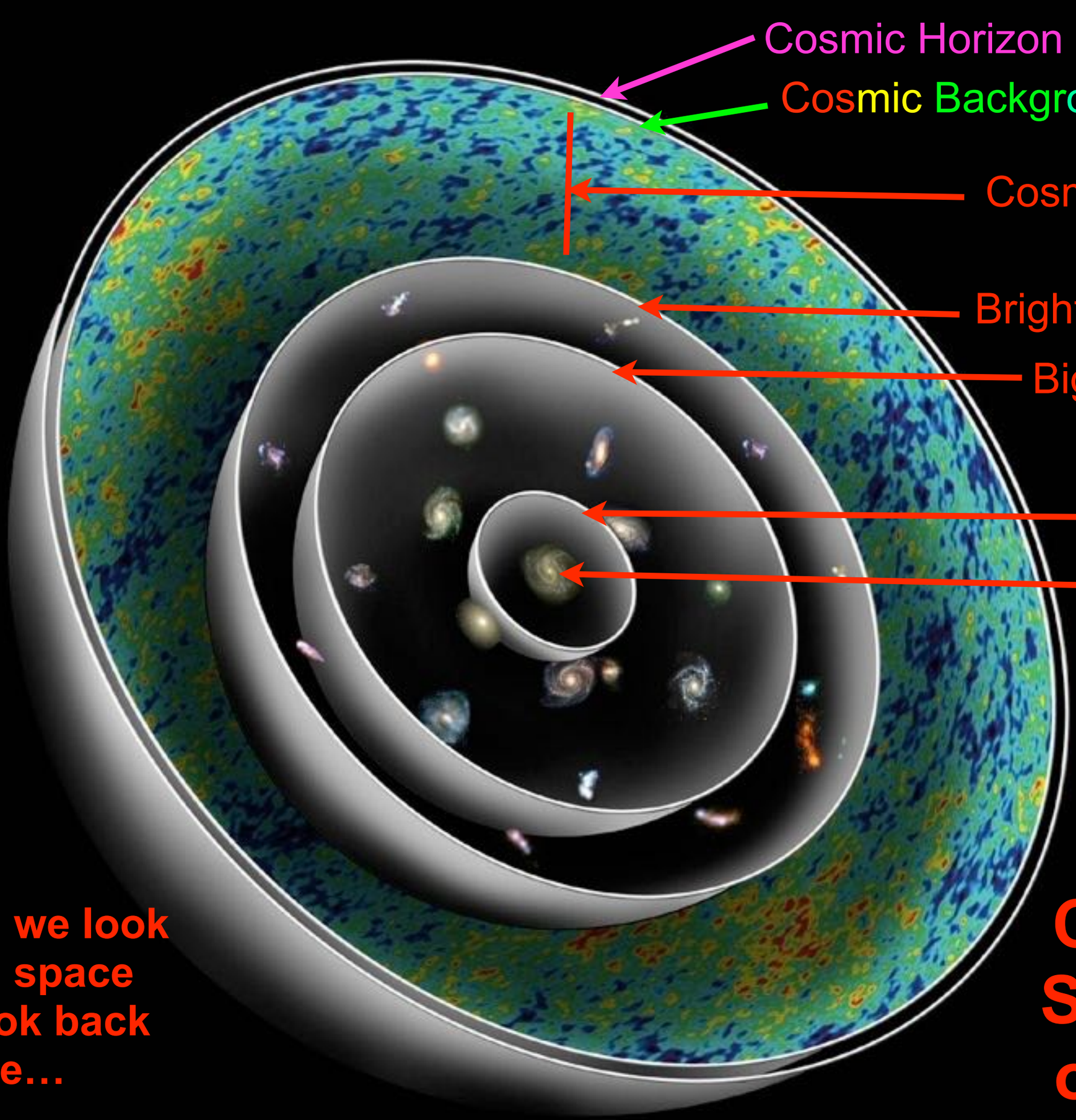
IS&T International Symposium on
Electronic Imaging 2018
SCIENCE AND TECHNOLOGY

30 January 2018 · Burlingame, California USA

Keynote: Imaging and Astronomy

Computer vision and deep learning applied to simulations and imaging of galaxies and the evolving universe

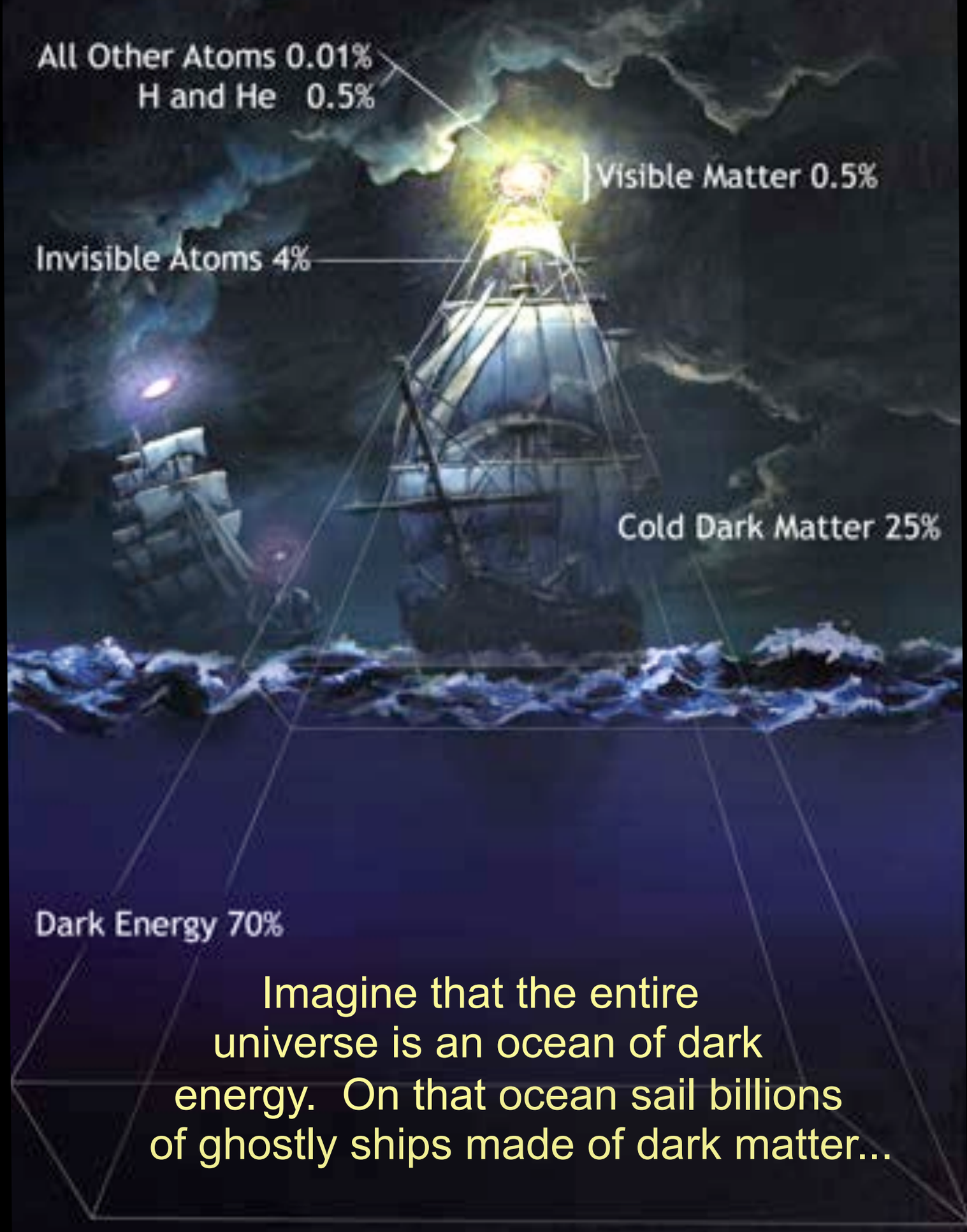
Joel Primack
University of California, Santa Cruz



When we look out in space we look back in time…

Cosmic Spheres of Time

Dark Matter Ships on a Dark Energy Ocean

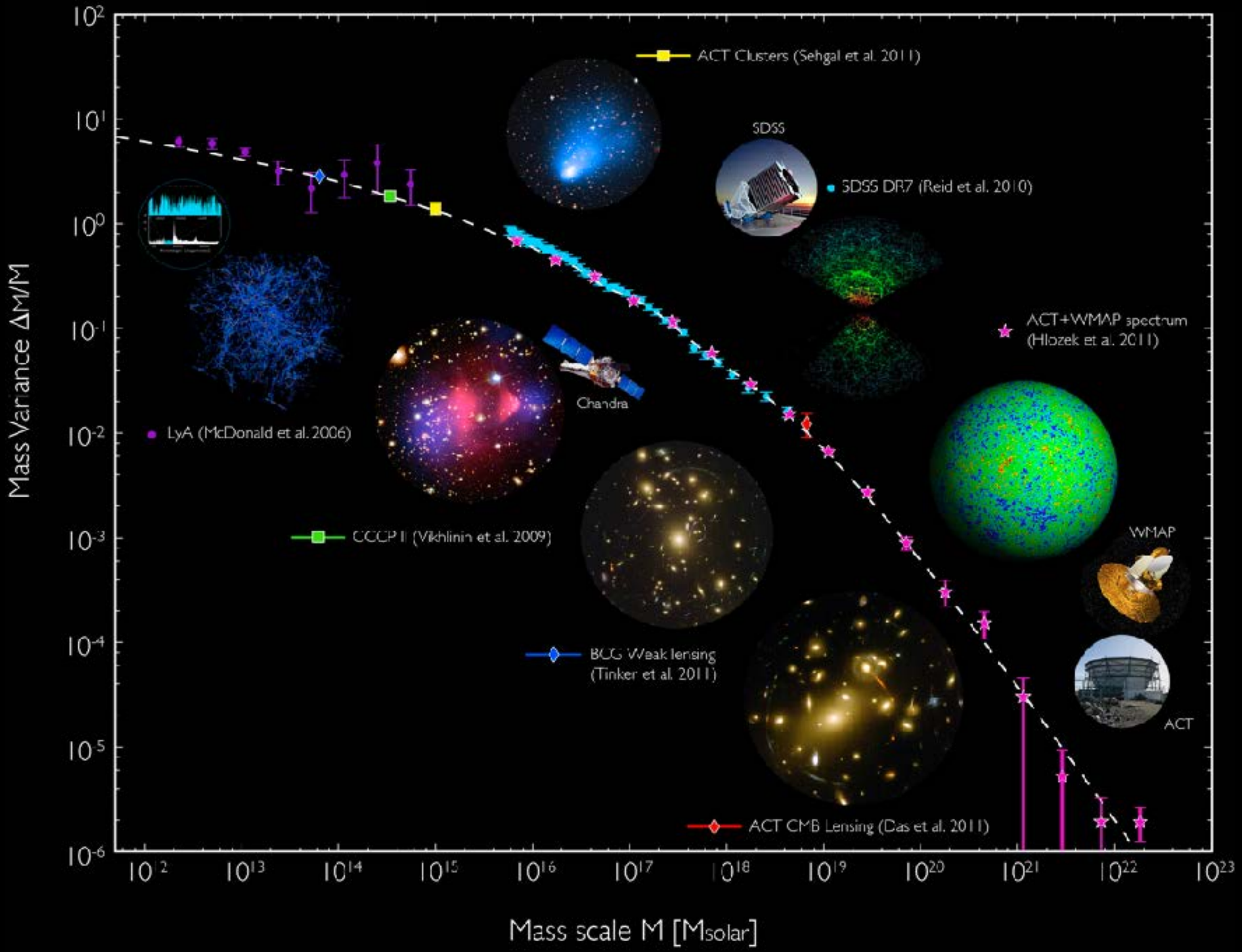


Matter and Energy Content of the Universe

ΛCDM

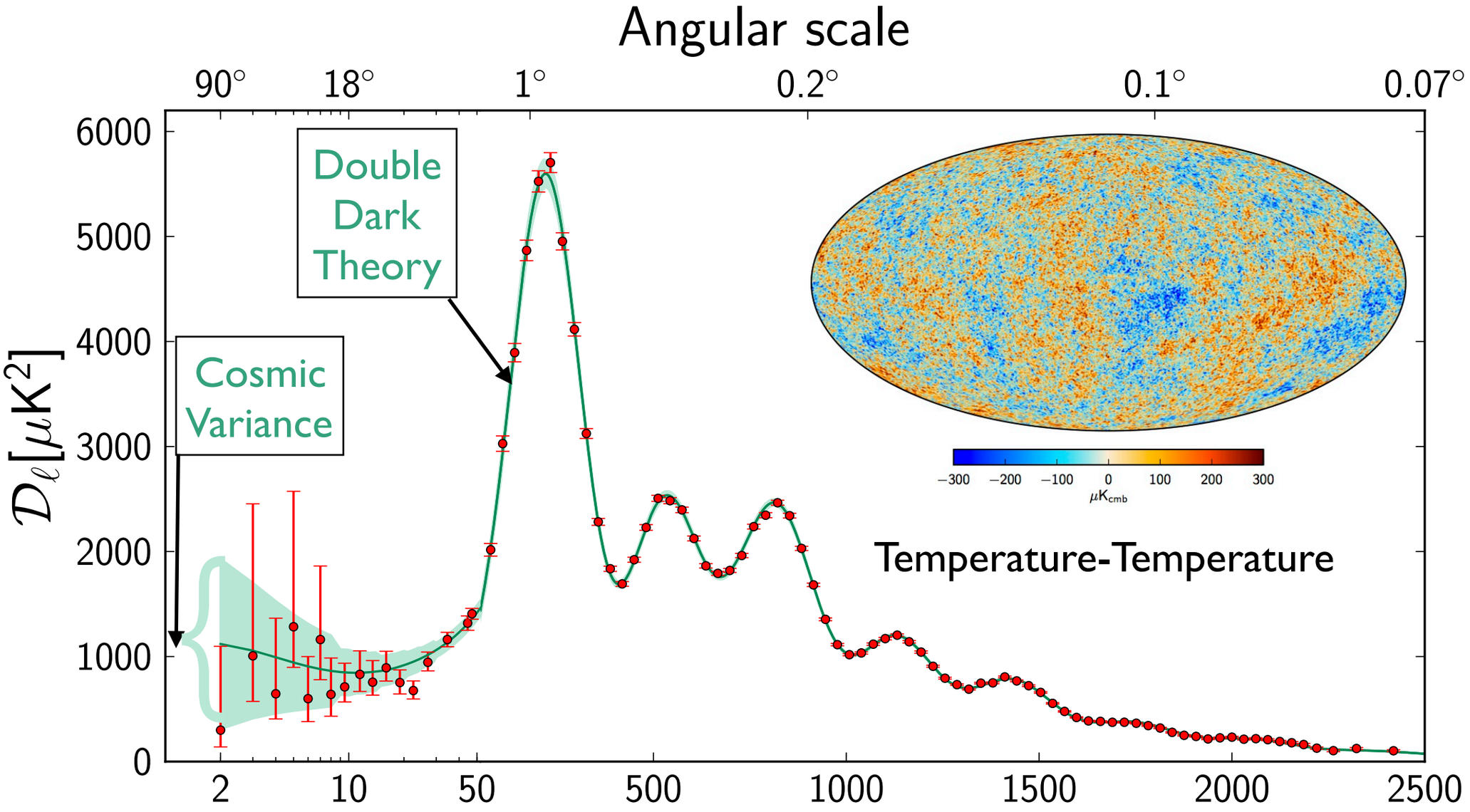
Double Dark Theory

Matter Distribution Agrees with Double Dark Theory!

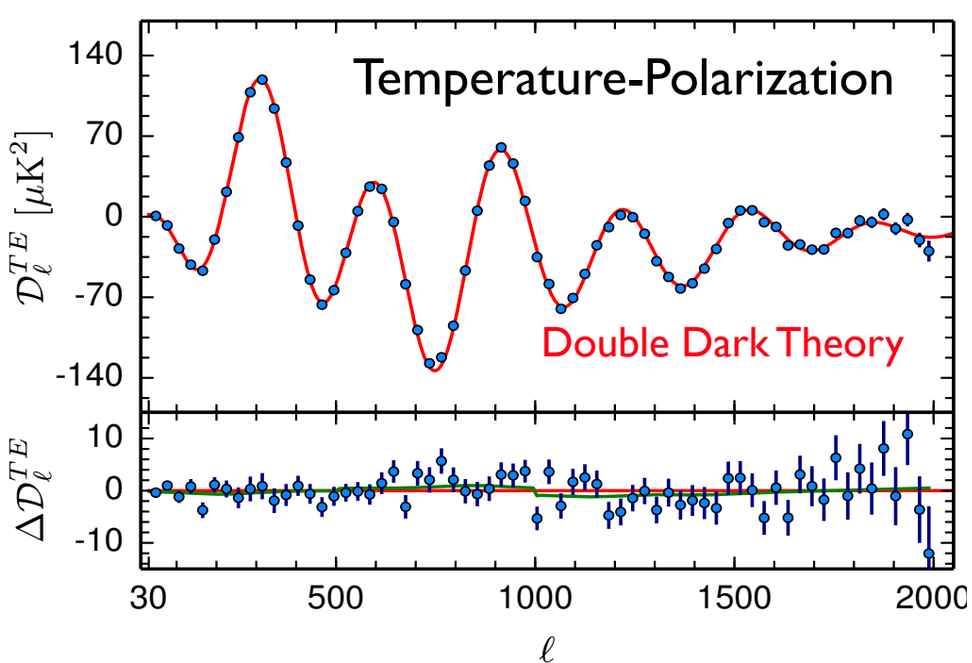


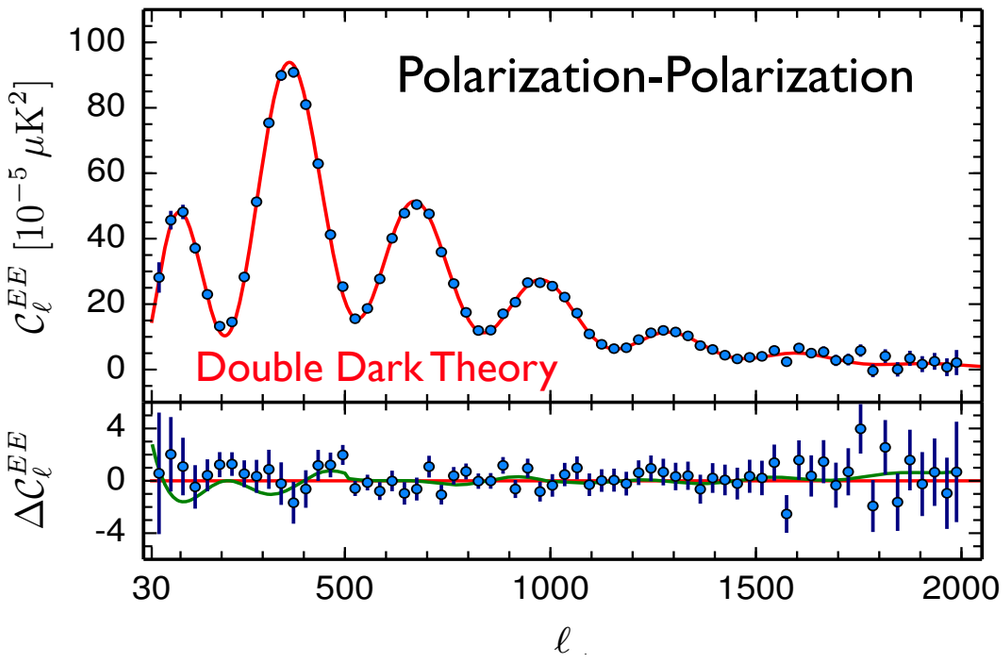
European Space Agency PLANCK Satellite Data

Released February 9, 2015



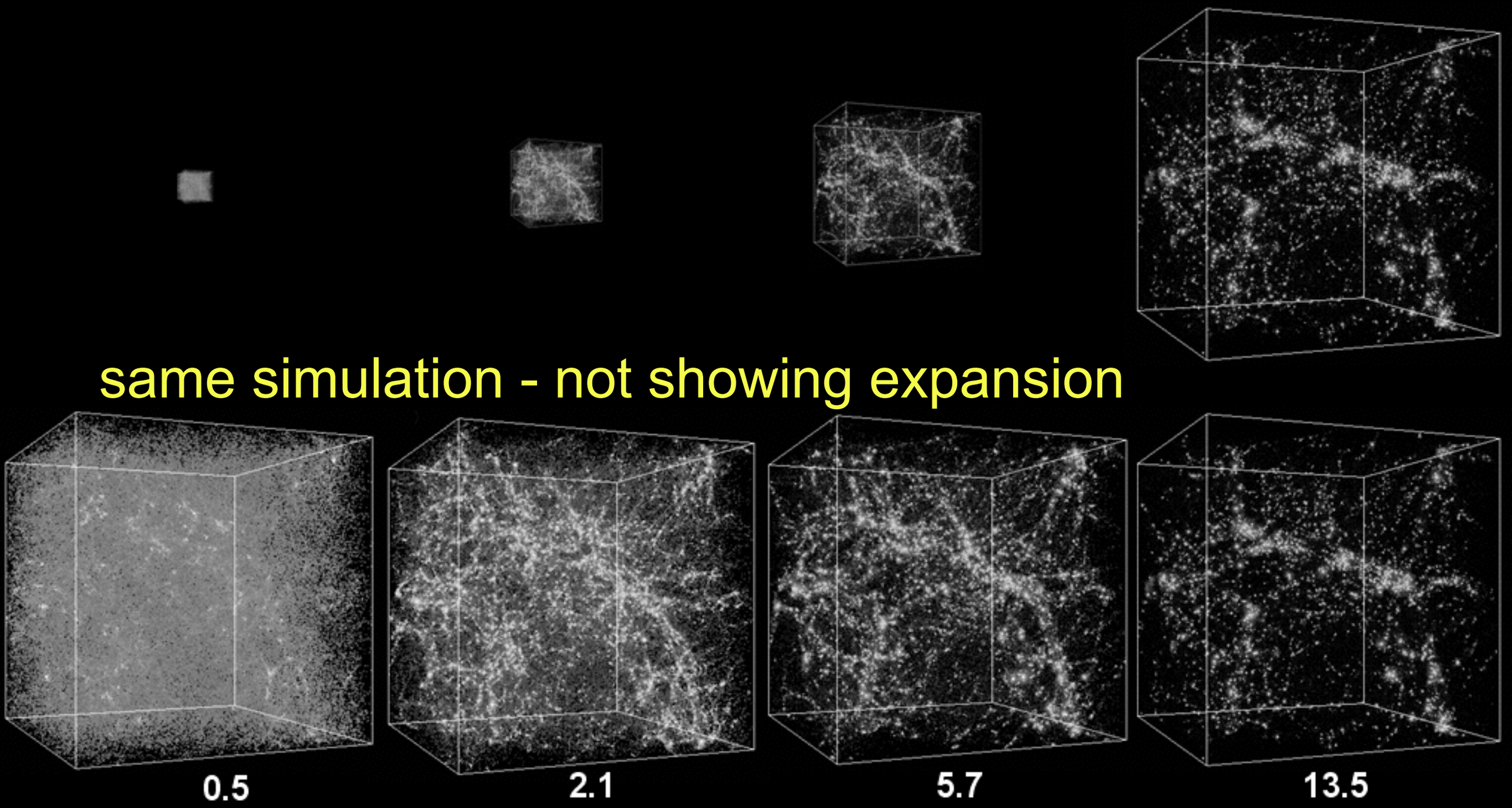
Agrees with Double Dark Theory!





"QUARKS. NEUTRINOS. MESONS. ALL THOSE DAMN PARTICLES YOU CAN'T SEE. THAT'S WHAT DROVE ME TO DRINK. BUT NOW I CAN SEE THEM!"

dark matter simulation - expanding with the universe



Andrey Kravtsov

CONSTRAINED LOCAL UNIVERSE SIMULATION

Stefan Gottloeber, Anatoly Klypin, Joel Primack
Visualization: Chris Henze (NASA Ames)

Aquarius Simulation

Volker Springel

Milky Way
100,000 light years

Milky Way Dark Matter Halo
1.5 million light years

Bolshoi Cosmological Simulation

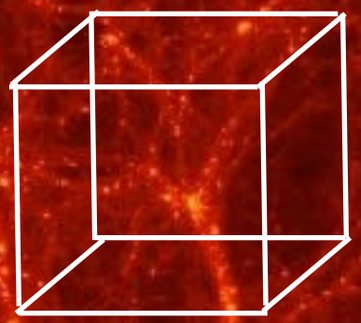
Anatoly Klypin & Joel Primack

8.6×10^9 particles 1/h kpc resolution

Pleiades Supercomputer at NASA Ames Research Center

1 Billion Light Years

100 Million Light Years



1 Billion Light Years

How the Halo of the Big Cluster Formed

100 Million Light Years

Bolshoi-Planck Cosmological Simulation

Merger Tree of a Large Halo

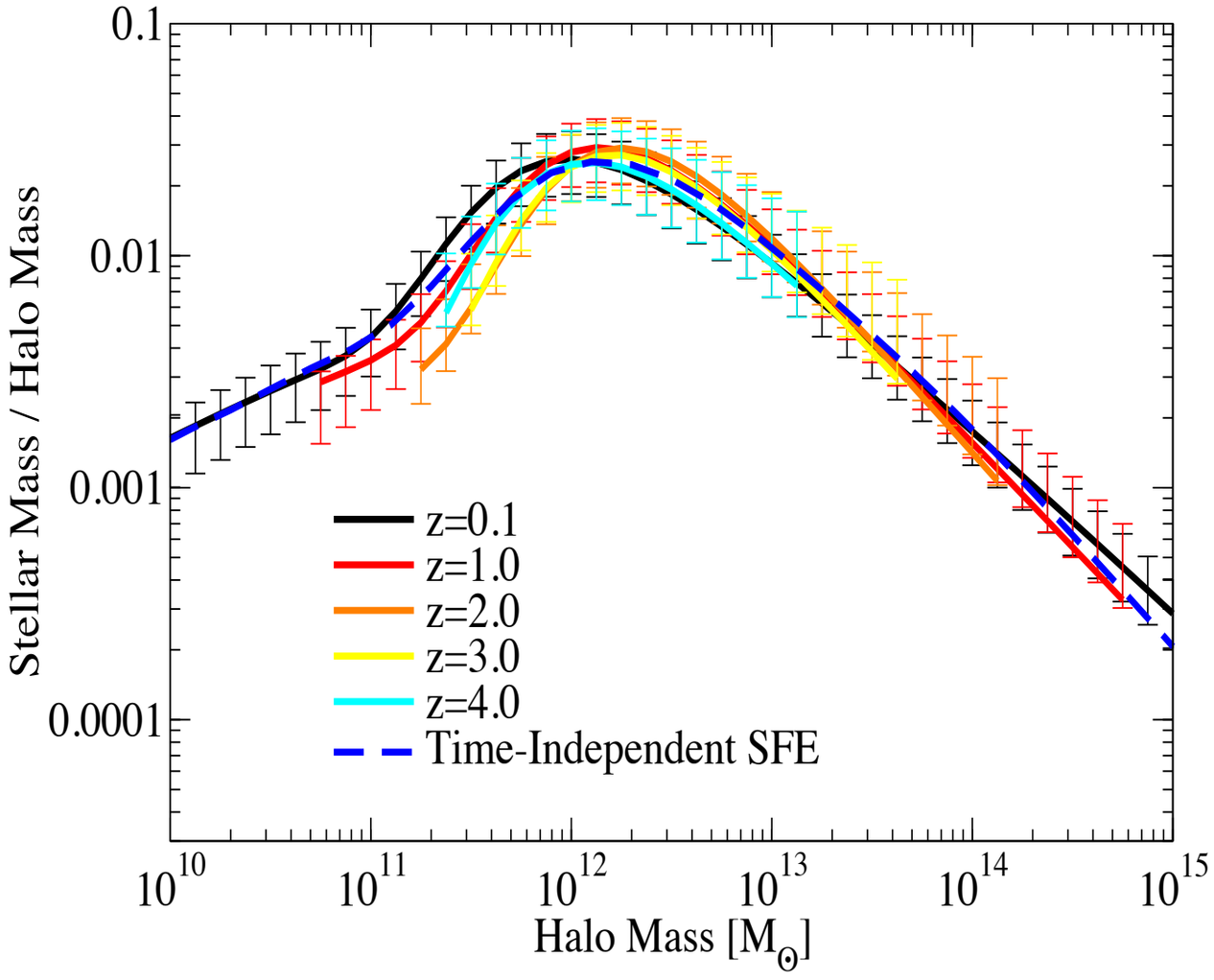
with Peter Behroozi & Christoph Lee

Structure Formation Methodology

- Starting from the Big Bang, we simulate the evolution of a representative part of the universe according to the Double Dark theory to see if the end result matches what astronomers actually observe.
- On the large scale the simulations produce a universe just like the one we observe. We're always looking for new phenomena to predict — every one of which tests the theory!
- But the way individual galaxies form is only partly understood because it depends on the interactions of the ordinary atomic matter as well as the dark matter and dark energy to form stars and super-massive black holes. We need help from observations.

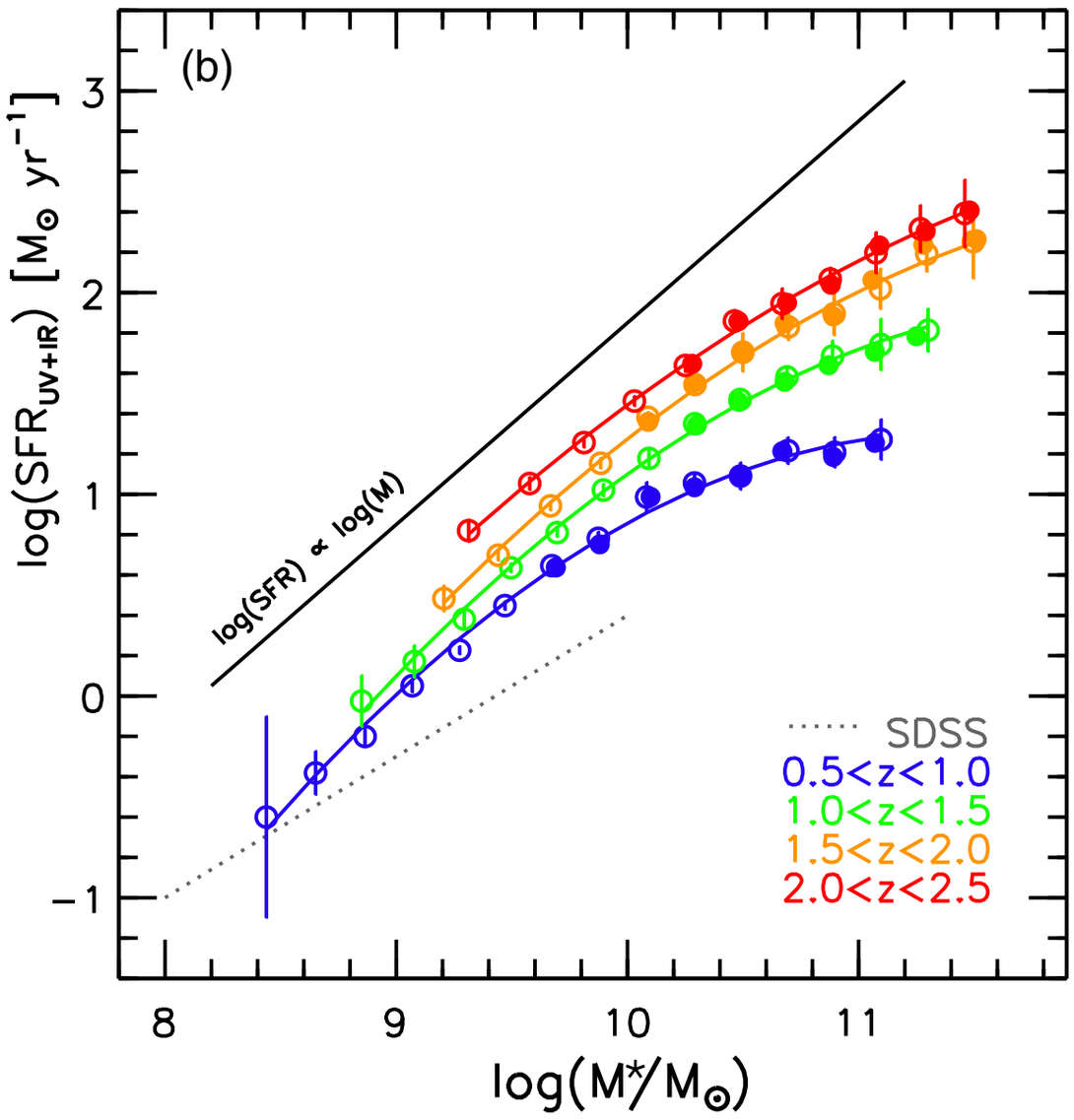
Two Key Discoveries About Galaxies

Galaxy Stellar Mass - Halo Mass Relation



The stellar mass to halo mass ratio at multiple redshifts as derived from observations compared to the Bolshoi cosmological simulation. Error bars show 1σ uncertainties. A time-independent Star Formation Efficiency predicts a roughly **time-independent stellar mass to halo mass (SMHM) relationship**. (Behroozi, Wechsler, Conroy, ApJL 2013)

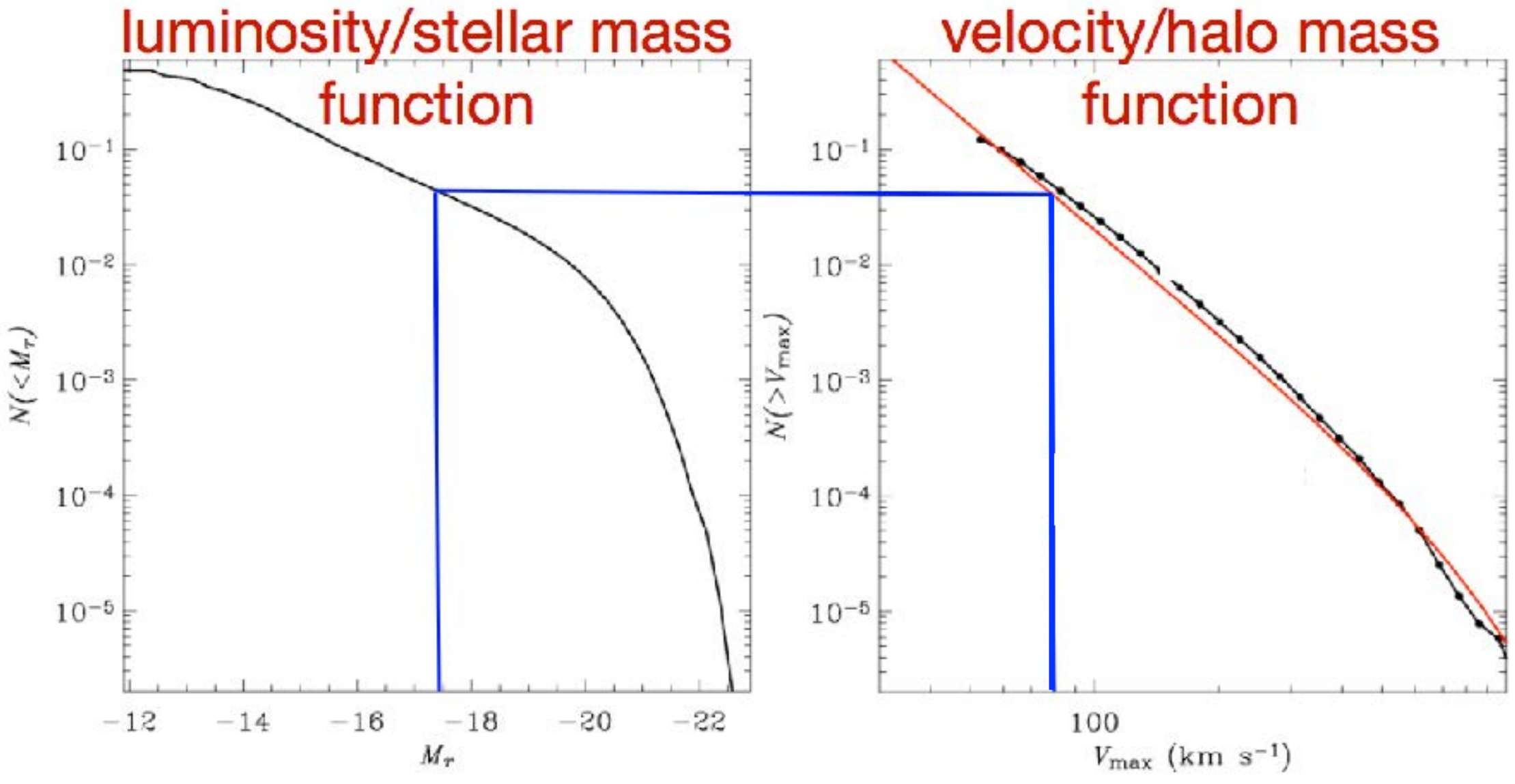
Star-forming Galaxies Lie on a "Main Sequence"



Just as the properties of hydrogen-burning stars are controlled by their mass, **the galaxy star formation rate (SFR) is approximately proportional to the stellar mass**, with ~0.3 dex dispersion and with the proportionality constant increasing with redshift up to about $z = 2.5$. (Whitaker et al. ApJ 2014)

Constraining the Galaxy Halo Connection: Star Formation Histories, Galaxy Mergers, and Structural Properties, by Aldo Rodriguez-Puebla, Joel Primack, Vladimir Avila-Reese, and Sandra Faber MNRAS 470, 651 (2017)

We use results from the Bolshoi-Planck simulation (Aldo Rodriguez-Puebla, Peter Behroozi, Joel Primack, Anatoly Klypin, Christoph Lee, Doug Hellinger 2016, MNRAS 462, 893), including halo and subhalo abundance as a function of redshift and median halo mass growth for halos of given M_{vir} at z = 0. Our semi-empirical approach uses **SubHalo Abundance Matching (SHAM)**, which matches the cumulative galaxy stellar mass function (GSMF) to the cumulative stellar mass function to correlate galaxy stellar mass with (sub)halo mass.



Assumption: every halo hosts a galaxy

SubHalo Abundance Matching (SHAM) Predicts Observed Galaxy Correlation Functions

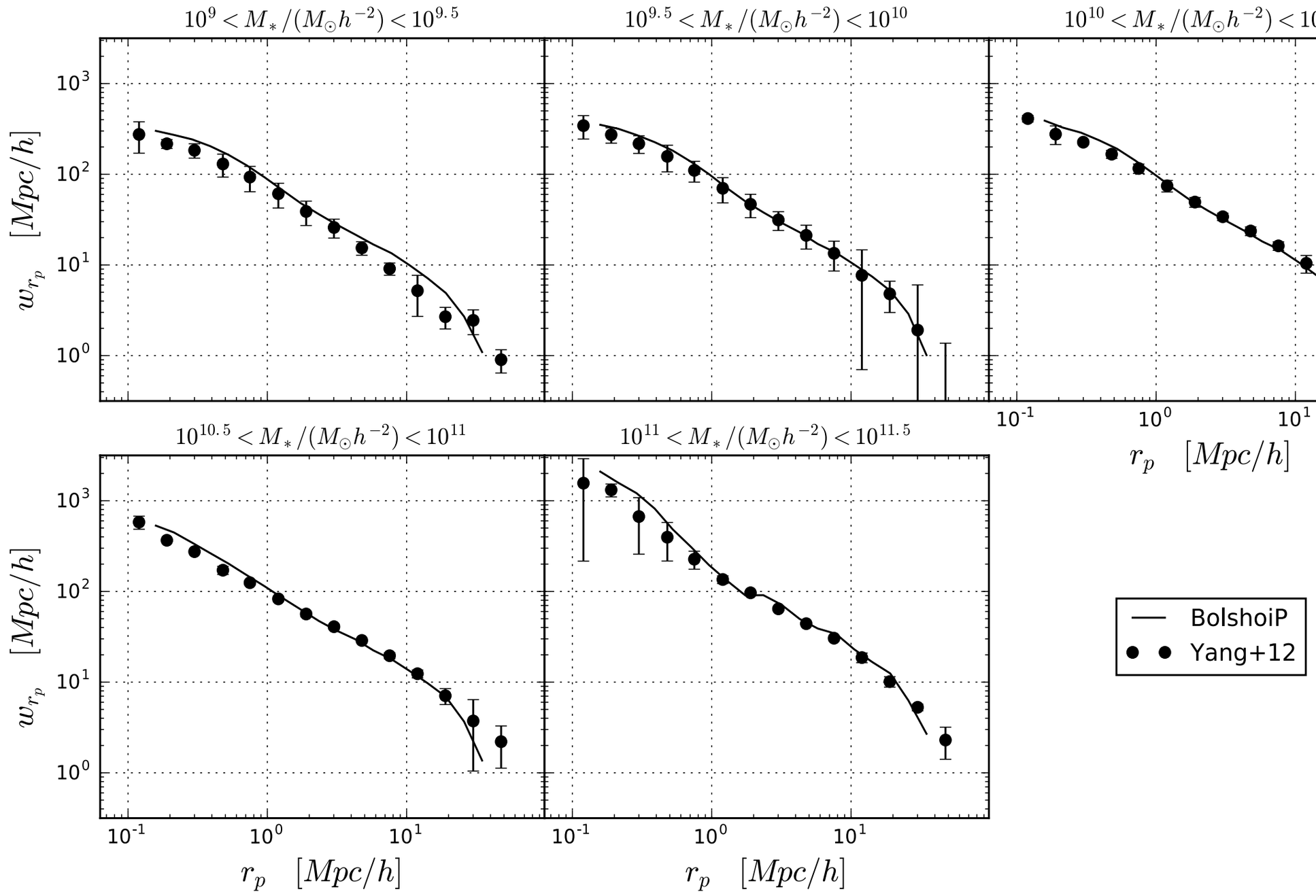
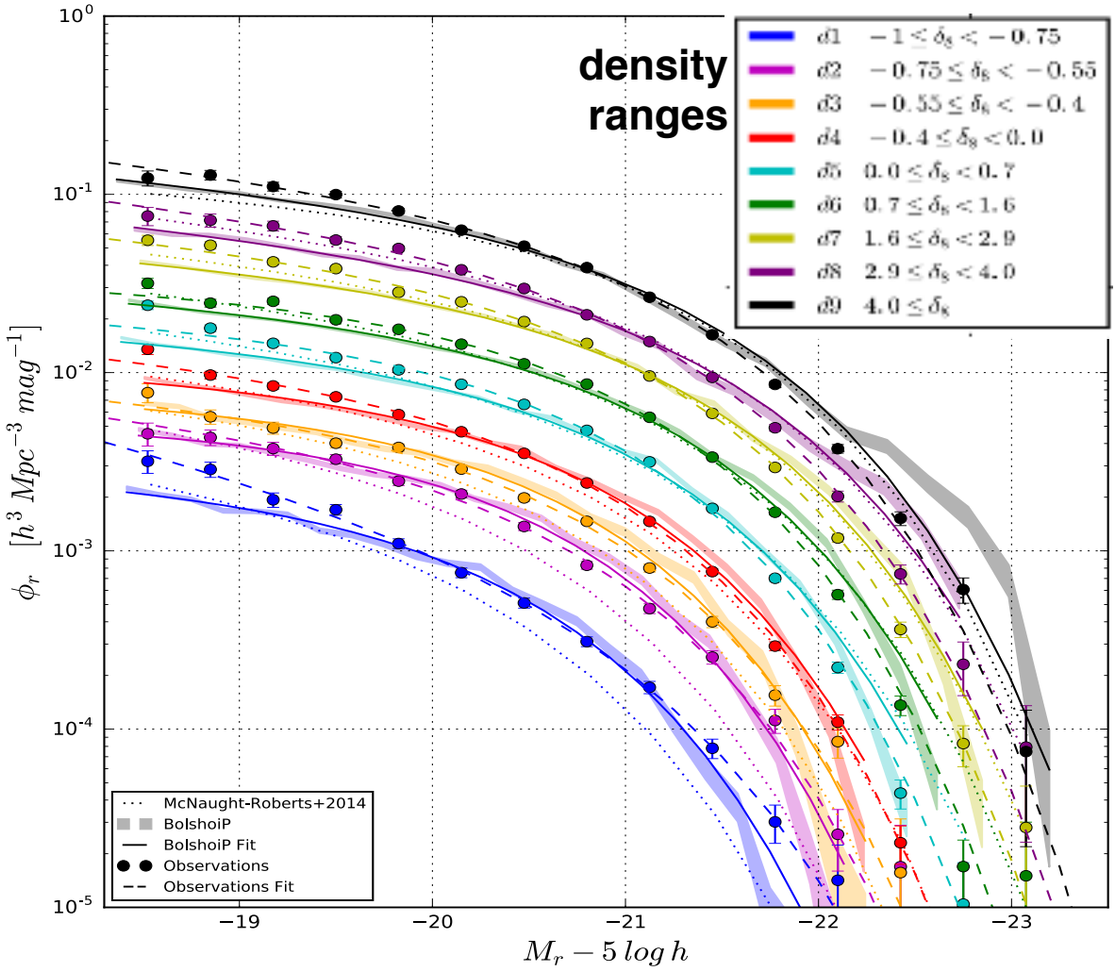


Figure 5. Two-point correlation function in five stellar mass bins. The solid lines show the predicted two-point correlation based on our stellar mass-to-$V_{\rm max}$ relation from SHAM, while the circles with error bars show the same but for SDSS DR7 (Yang et al. 2012).

Does the Galaxy-Halo Connection Vary with Environment? **R. Dragomir, A. Rodriguez-Puebla, J. Primack, C. Lee, MNRAS 2018**

SubHalo Abundance Matching (SHAM) Predicts Density Dependence of Galaxy Luminosity and Mass

r-Band Luminosity Function



Stellar Mass Function

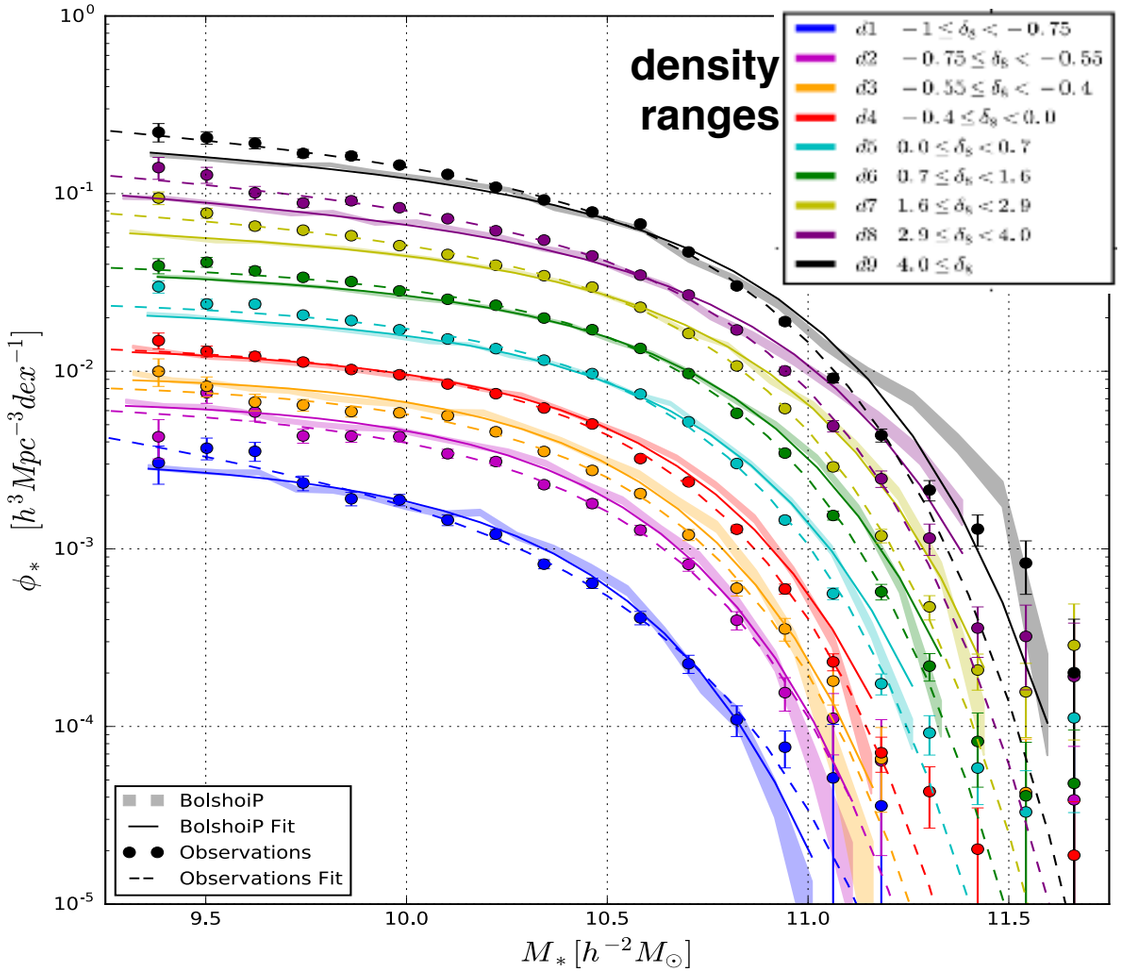


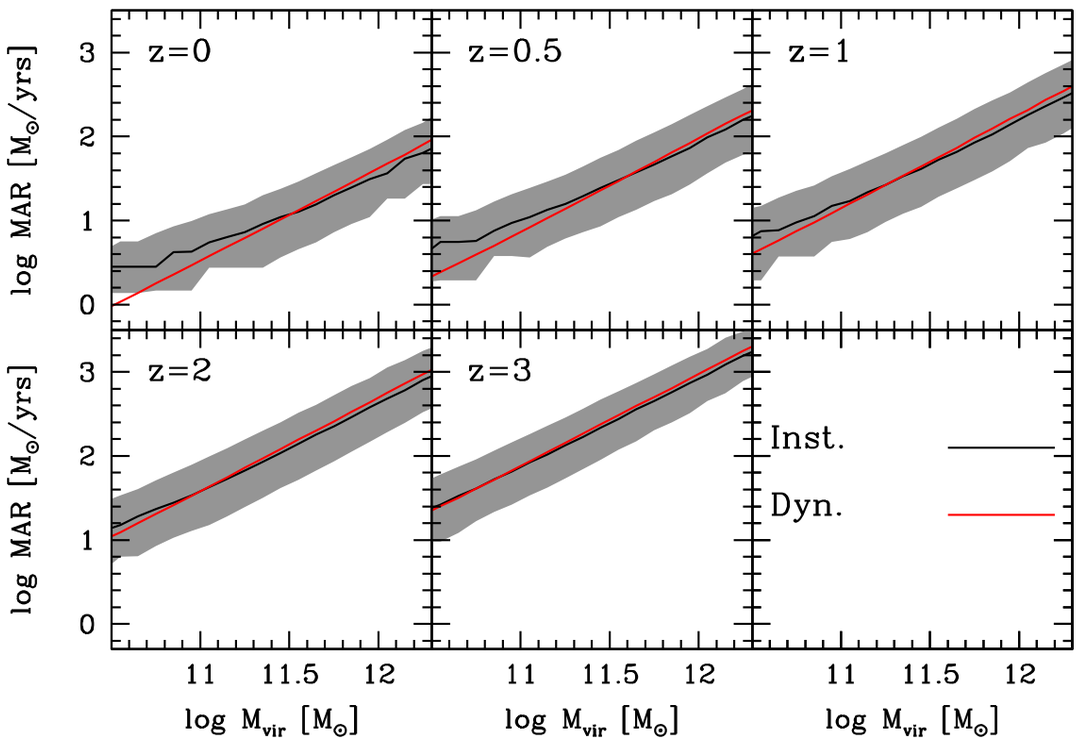
Figure 7. Left Panel: Comparison between the observed $r-$band GLF with environmental density in spheres of 8 h^{-1}Mpc, filled circles with error bars, and the ones predicted based on the BolshoiP simulation from SHAM, shaded regions. The dashed lines show the best fitting Schechter functions to the r-band GLFs from the GAMA survey (McNaught-Roberts et al. 2014). **Right Panel:** Similar to the left panel but for the GSMF with environmental density. Here again the dashed lines are the best fitting Schechter functions.

Does the Galaxy-Halo Connection Vary with Environment? **R. Dragomir, A. Rodriguez-Puebla, J. Primack, C. Lee, MNRAS 2018**

Main Sequence Star Formation Reflects Halo Mass Accretion

by Aldo Rodríguez-Puebla, Joel Primack, Peter Behroozi, Sandra Faber **MNRAS 2016**

Halo mass accretion rates z=0 to 3



$$\frac{dM_*}{dt} = \frac{\partial M_*(M_{\rm vir}(t),z)}{\partial M_{\rm vir}}\frac{dM_{\rm vir}}{dt} + \frac{\partial M_*(M_{\rm vir}(t),z)}{\partial z}\frac{dz}{dt}$$

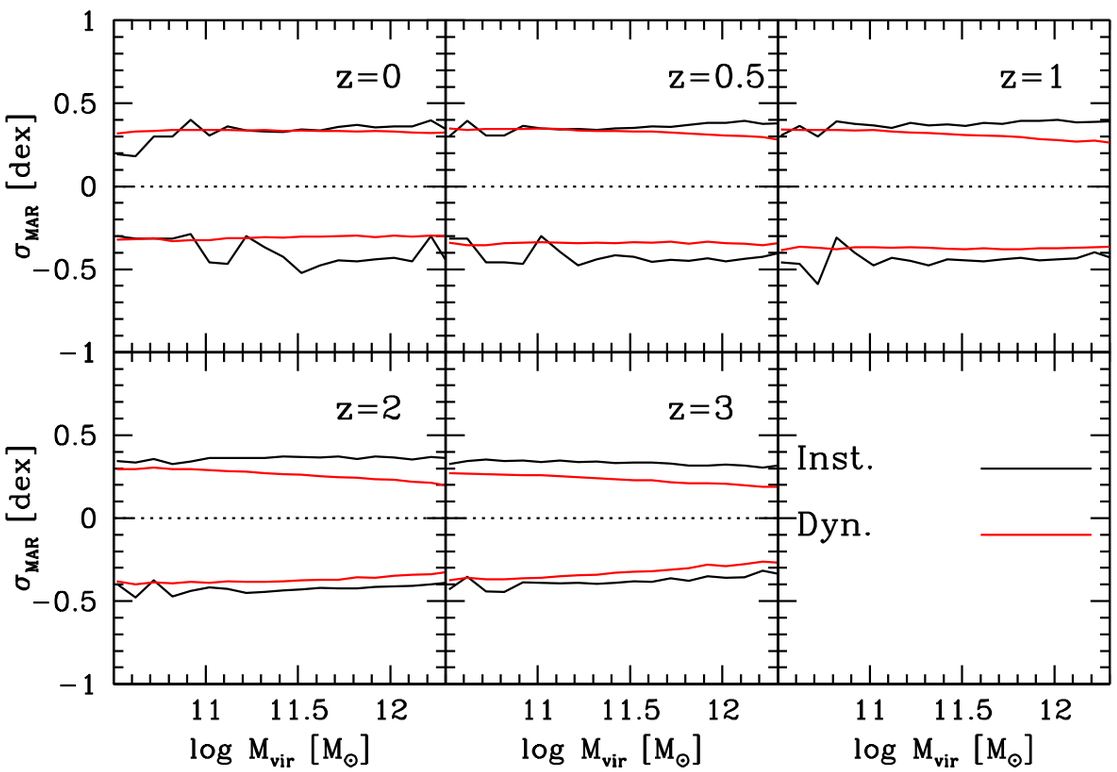
but if the M_*–M_{vir} relation is independent of redshift then the stellar mass of a central galaxy formed in a halo of mass $M_{vir}(t)$ is $M_* = M_*(M_{vir}(t))$ and the second term vanishes.

star formation rate $= \frac{dM_*}{dt} = \frac{\partial M_*(M_{\rm vir}(t),z)}{\partial M_{\rm vir}}\frac{dM_{\rm vir}}{dt}$

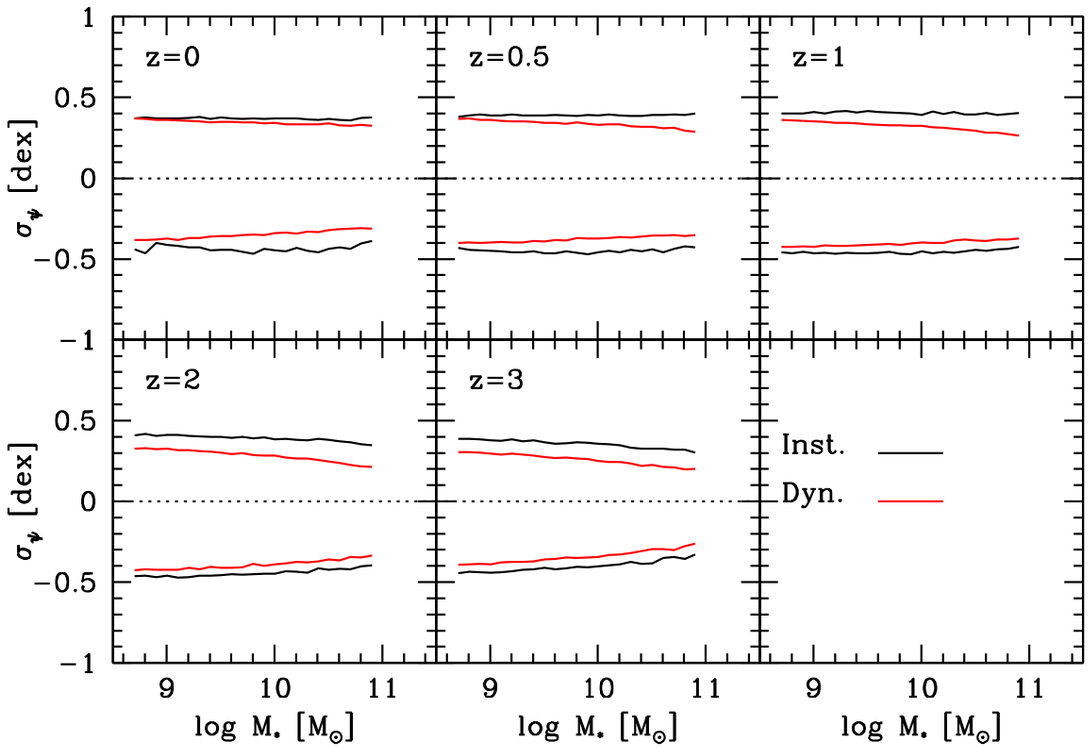
SMHM slope **halo mass accretion rate**

We call this **Stellar-Halo Accretion Rate Coevolution (SHARC)** if true halo-by-halo for star-forming galaxies.

Scatter of halo mass accretion rates



Implied scatter of star formation rates

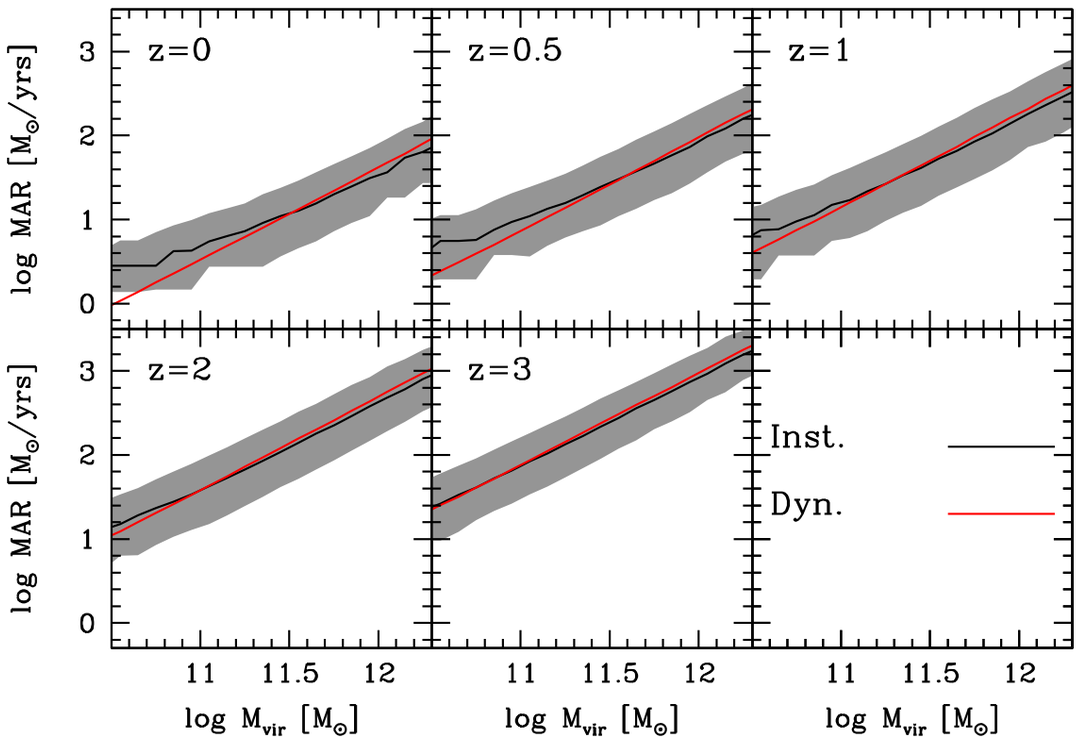


Consistent with observations!

Main Sequence Star Formation Reflects Halo Mass Accretion

by Aldo Rodríguez-Puebla, Joel Primack, Peter Behroozi, Sandra Faber **MNRAS 2016**

Halo mass accretion rates z=0 to 3



$$\frac{dM_*}{dt} = \frac{\partial M_*(M_{\rm vir}(t),z)}{\partial M_{\rm vir}}\frac{dM_{\rm vir}}{dt} + \frac{\partial M_*(M_{\rm vir}(t),z)}{\partial z}\frac{dz}{dt}$$

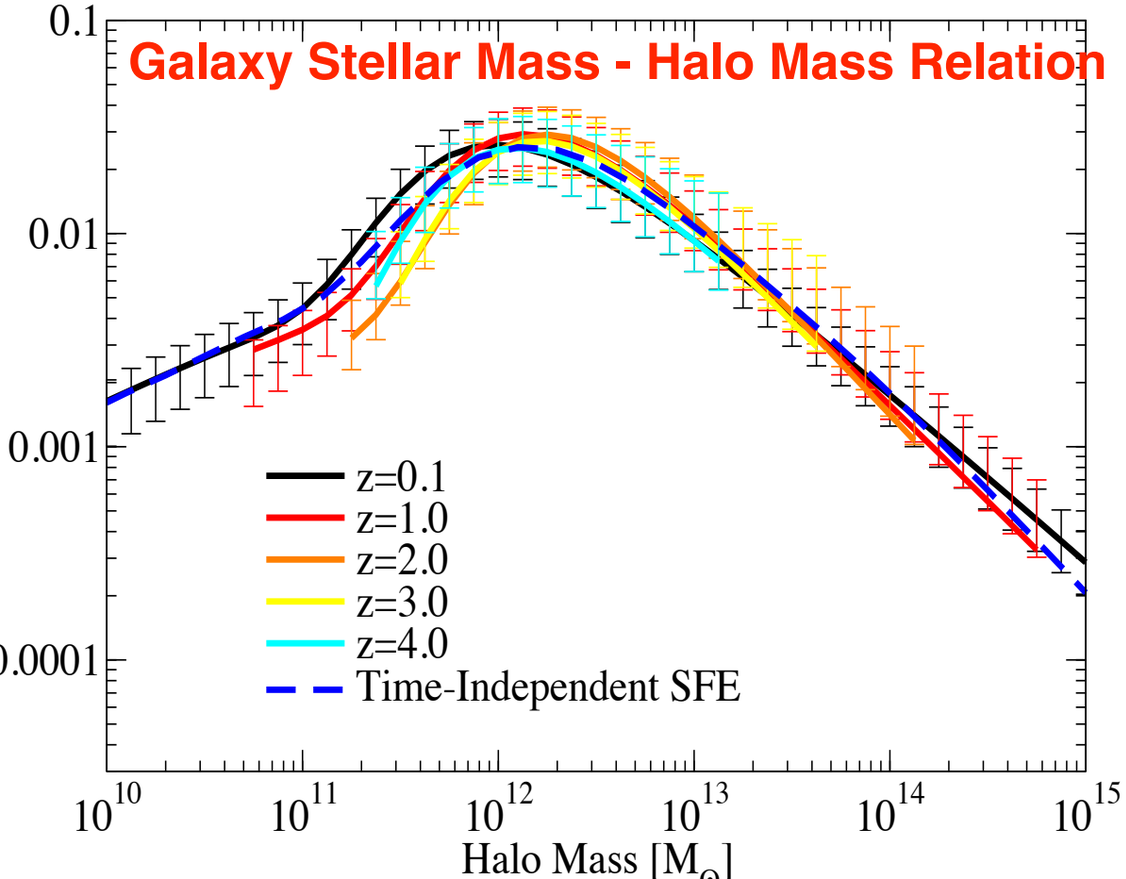
but if the M_*–M_{vir} relation is independent of redshift then the stellar mass of a central galaxy formed in a halo of mass $M_{vir}(t)$ is $M_* = M_*(M_{vir}(t))$ and the second term vanishes.

star formation rate $= \frac{dM_*}{dt} = \frac{\partial M_*(M_{\rm vir}(t),z)}{\partial M_{\rm vir}}\frac{dM_{\rm vir}}{dt}$

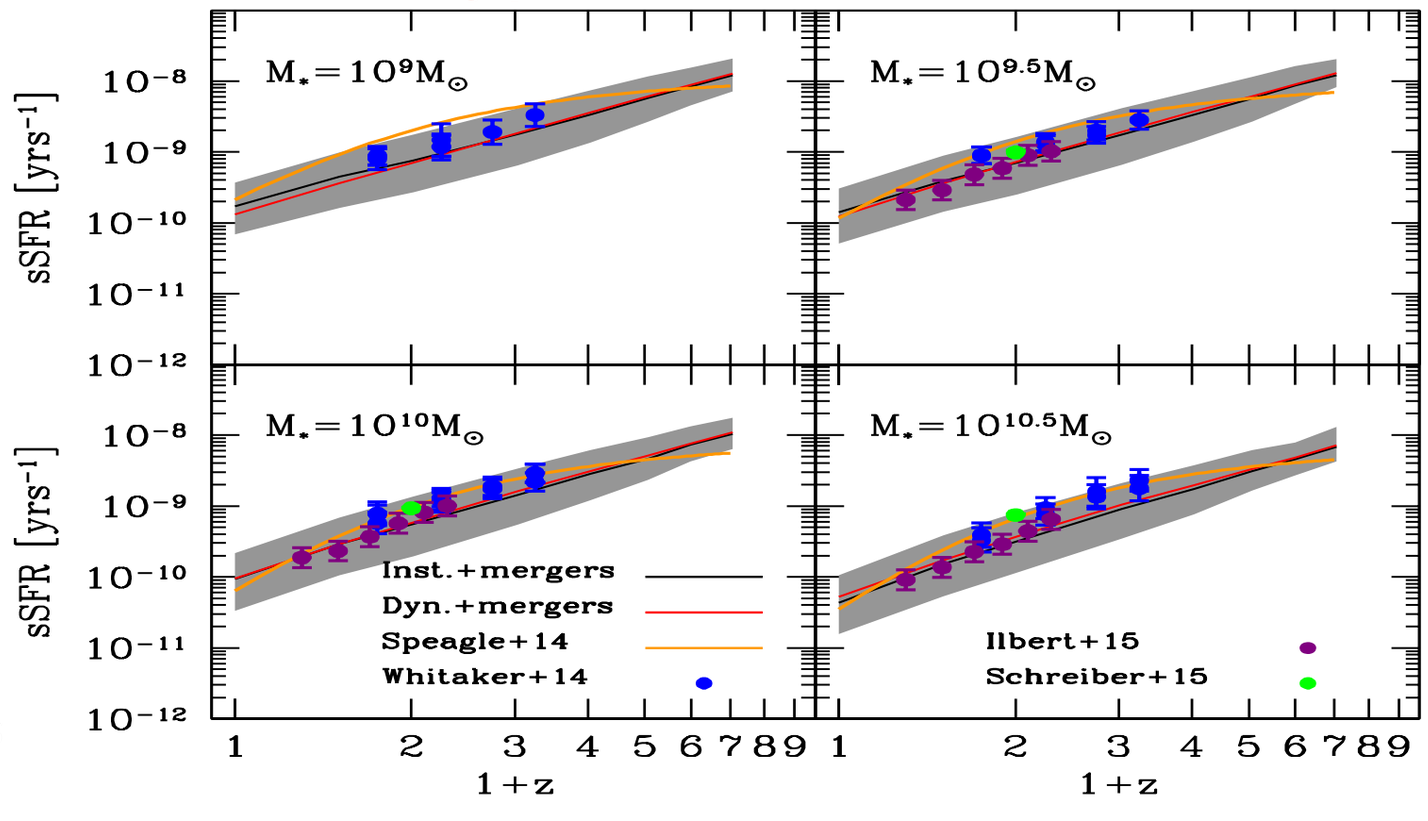
SMHM slope **halo mass accretion rate**

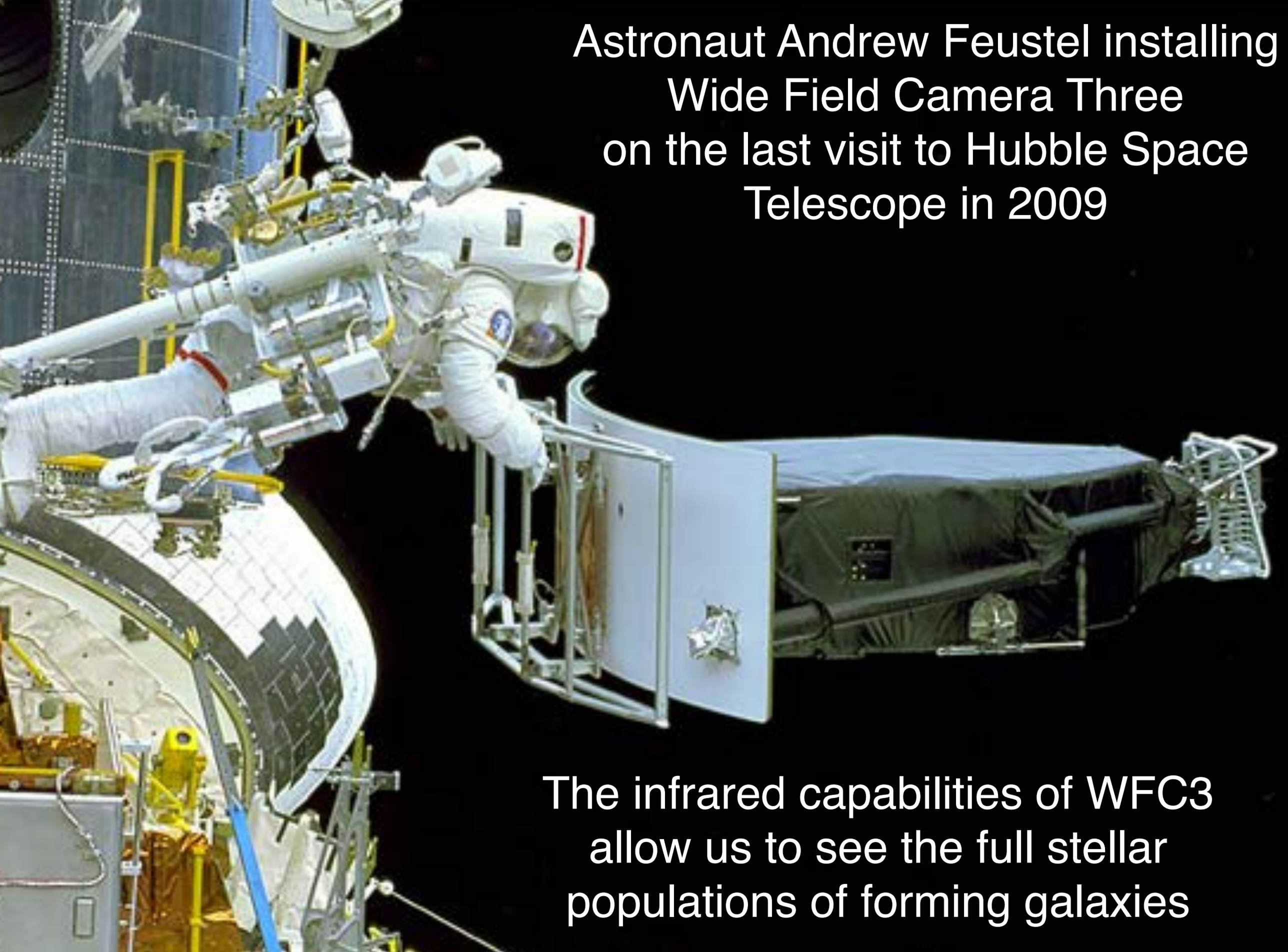
We call this **Stellar-Halo Accretion Rate Coevolution (SHARC)** if true halo-by-halo for star-forming galaxies.

Galaxy Stellar Mass - Halo Mass Relation



SHARC correctly predicts star formation rates to z ~ 4



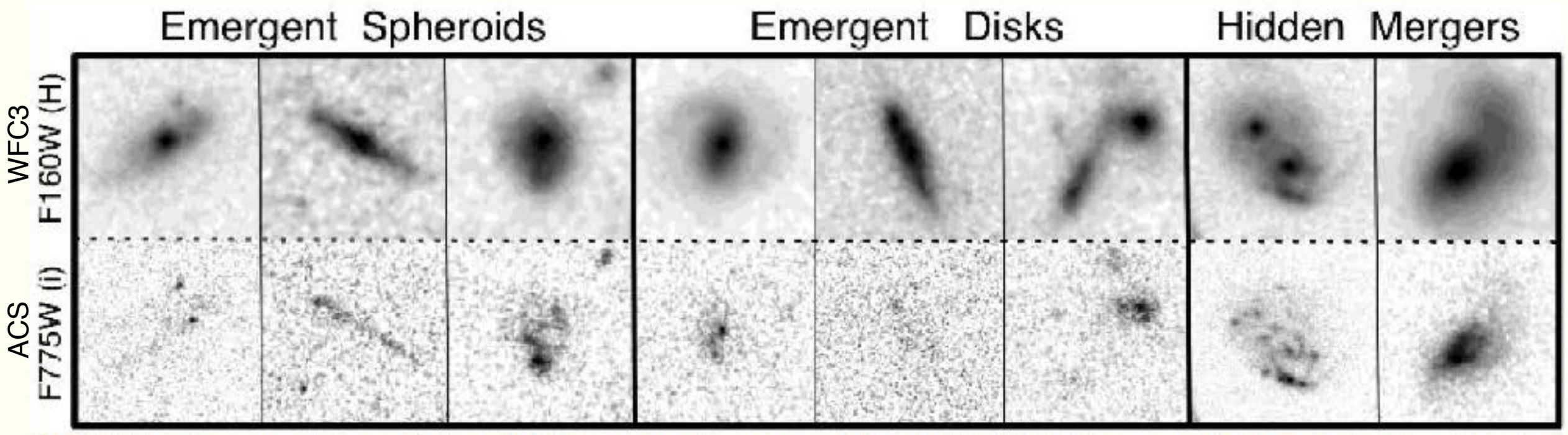


Astronaut Andrew Feustel installing
Wide Field Camera Three
on the last visit to Hubble Space
Telescope in 2009

The infrared capabilities of WFC3
allow us to see the full stellar
populations of forming galaxies

The CANDELS Survey

candels.ucolick.org



CANDELS: A Cosmic Odyssey

(blue 0.4 μm)(1+z) = 1.6 μm @ z = 3 (11.5 Gyr ago)
(red 0.7 μm)(1+z) = 1.6 μm @ z = 2.3 (10 Gyr ago)

CANDELS is a powerful imaging survey of the distant Universe being carried out with two cameras on board the Hubble Space Telescope.

- CANDELS is the largest project in the history of Hubble, with 902 assigned orbits of observing time. This is the equivalent of four months of Hubble time if executed consecutively, but in practice CANDELS will take three years to complete (2010-2013).
- The core of CANDELS is the revolutionary near-infrared WFC3 camera, installed on Hubble in May 2009. WFC3 is sensitive to longer, redder wavelengths, which permits it to follow the stretching of lightwaves caused by the expanding Universe. This enables CANDELS to detect and measure objects much farther out in space and nearer to the Big Bang than before. CANDELS also uses the visible-light ACS camera, and together the two cameras give unprecedented panchromatic coverage of galaxies from optical wavelengths to the near-IR.
- CANDELS will exploit this new lookback power to construct a "cosmic movie" of galaxy evolution that follows the life histories of galaxies from infancy to the present time. This work will cap Hubble's revolutionary series of discoveries on cosmic evolution and bequeath a legacy of precious data to future generations of astronomers.

Redshift

9 6 4 3 2 1 0.5 0.2 0

CANDELS

Star-formation rate
(solar masses/year/megaparsecs^3 $\times 10^{-2}$)

10
5
0

Cosmic dawn

Cosmic high noon

0 1 2 3 4 5 6 7 8 9 10 11 12 13

Time since Big Bang (billions of years)

Big Bang

Now 13.7

Most astronomers used to think

(1) that galaxies form as disks,

(2) that forming galaxies are pretty smooth, and

(3) that galaxies generally grow in radius as they grow in mass.

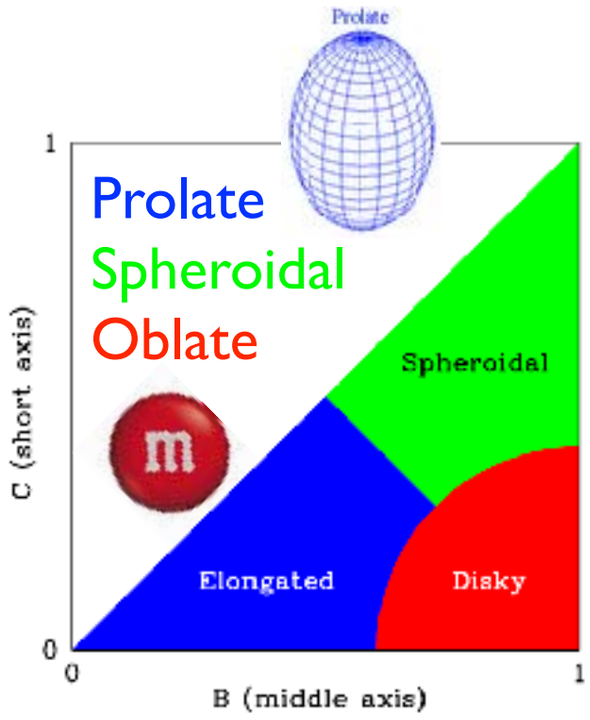
But CANDELS and other HST observations show that all these assumptions were wrong!

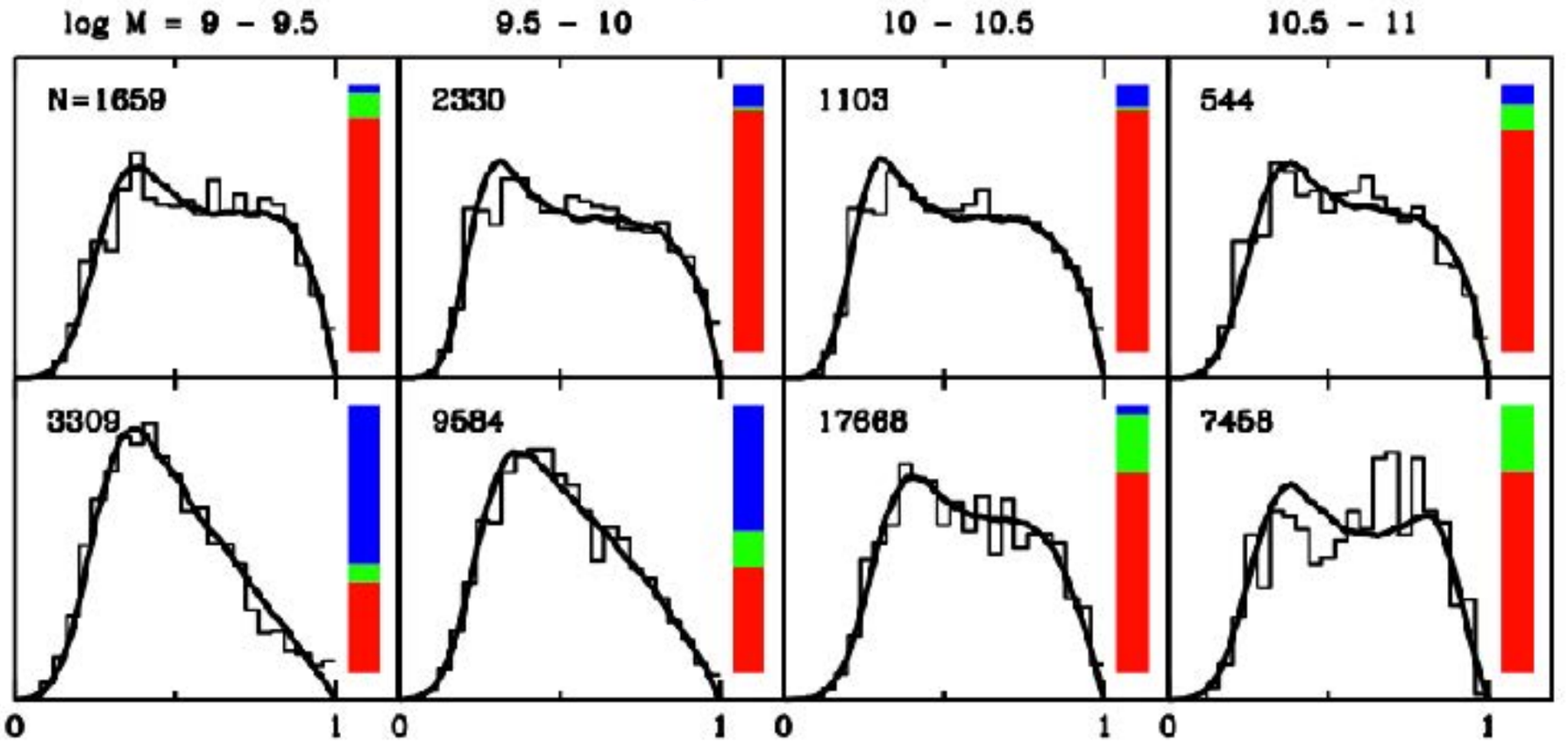
(1) The majority of star-forming galaxies at $z > 1$ apparently have mostly elongated (prolate) stellar distributions rather than disks or spheroids, and our simulations may explain why.

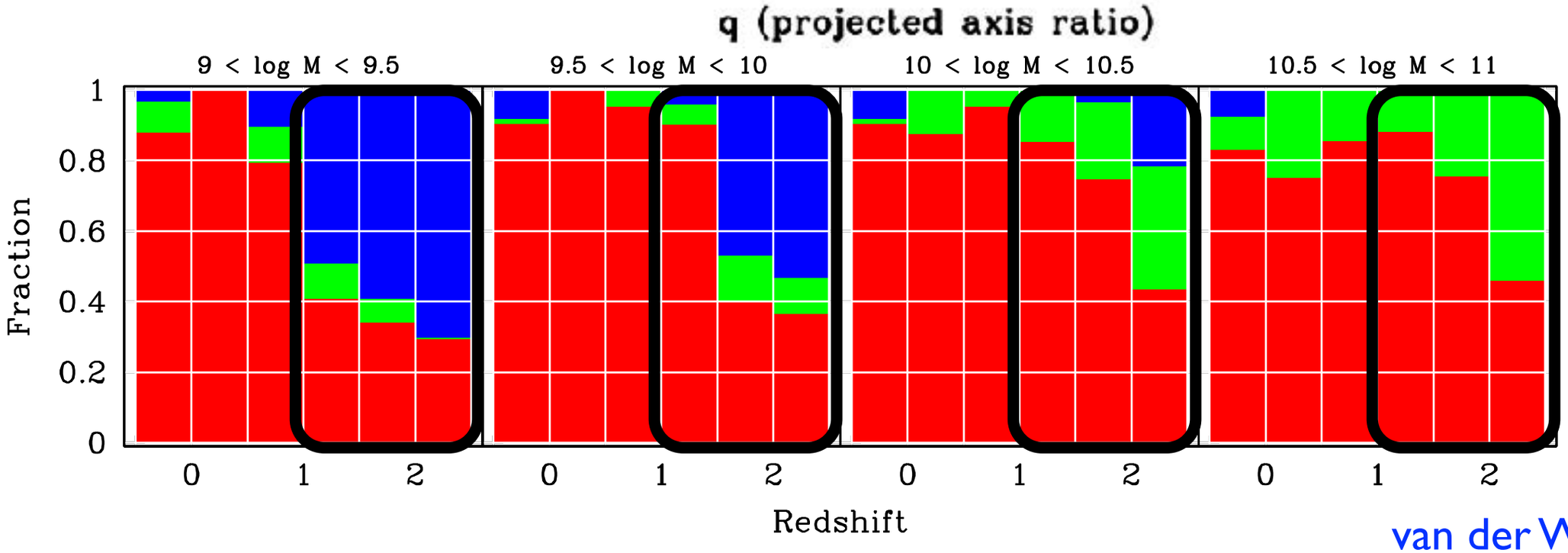
(2) A large fraction of star-forming galaxies at redshifts $1 < z < 3$ are found to have massive stellar clumps; these originate from phenomena including mergers and disk instabilities in our simulations.

(3) These phenomena also help to create compact stellar spheroidal galaxies ("nuggets") through galaxy compaction (rapid inflow of gas to galaxy centers, where it forms stars).

Prolate Galaxies Dominate at High Redshifts & Low Masses







van der Wel+2014

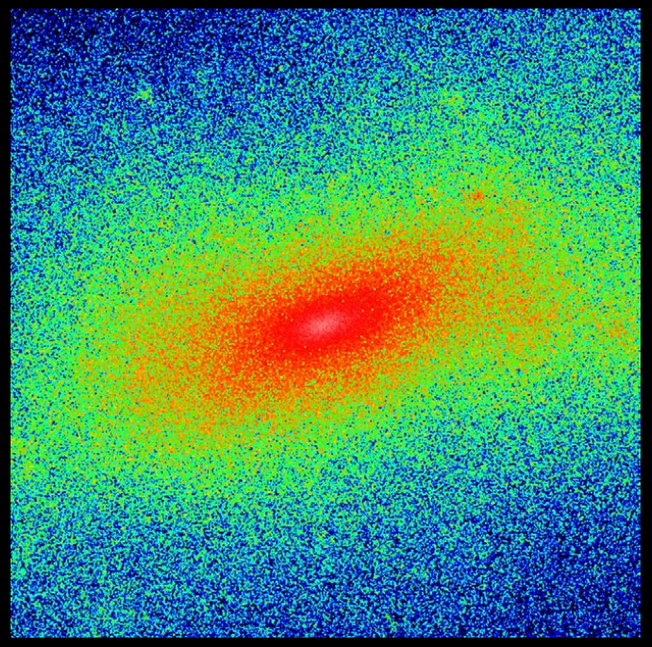
See also Morphological Survey of Galaxies z=1.5-3.6 Law, Steidel+ ApJ 2012

When Did Round Disk Galaxies Form? T. M. Takeuchi+ ApJ 2015

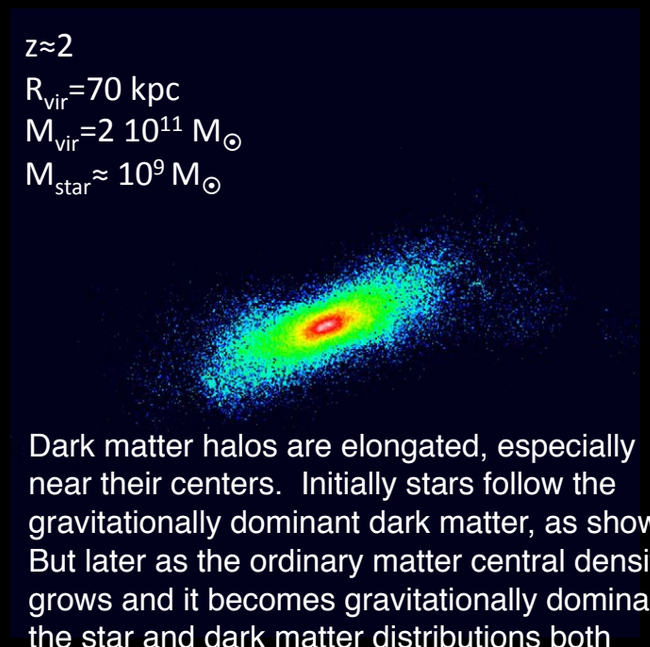
Our cosmological zoom-in simulations often produce elongated galaxies like observed ones. The elongated stellar distribution follows the elongated inner dark matter halo.

Prolate DM halo → elongated galaxy

DM **VELA28** **stars**

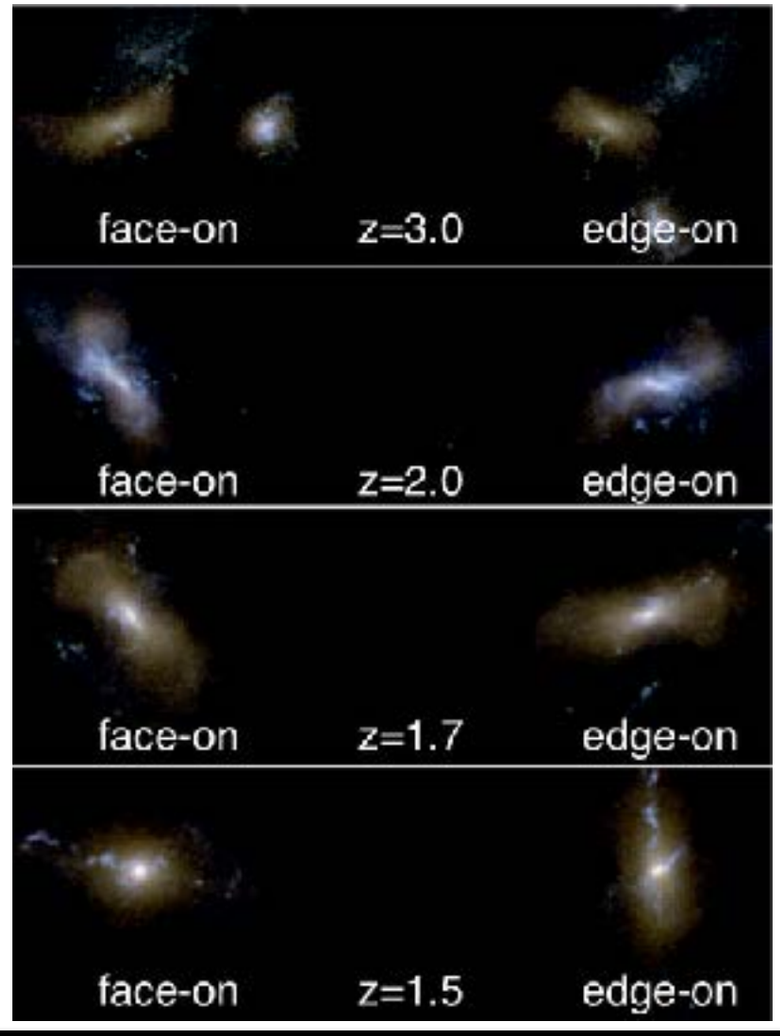


30 kpc

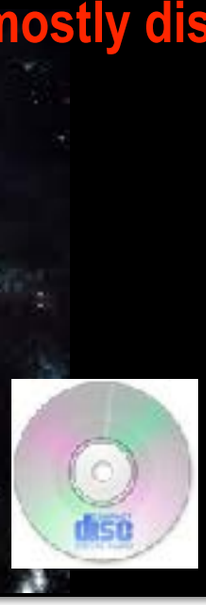


z≈2
R_{vir}=70 kpc
M_{vir}=2 10^{11} $M_\odot$
M_{star}≈ 10^9 $M_\odot$

Dark matter halos are elongated, especially near their centers. Initially stars follow the gravitationally dominant dark matter, as shown. But later as the ordinary matter central density grows and it becomes gravitationally dominant, the star and dark matter distributions both become disky — as observed by Hubble Space Telescope (van der Wel+ ApJL Sept 2014).



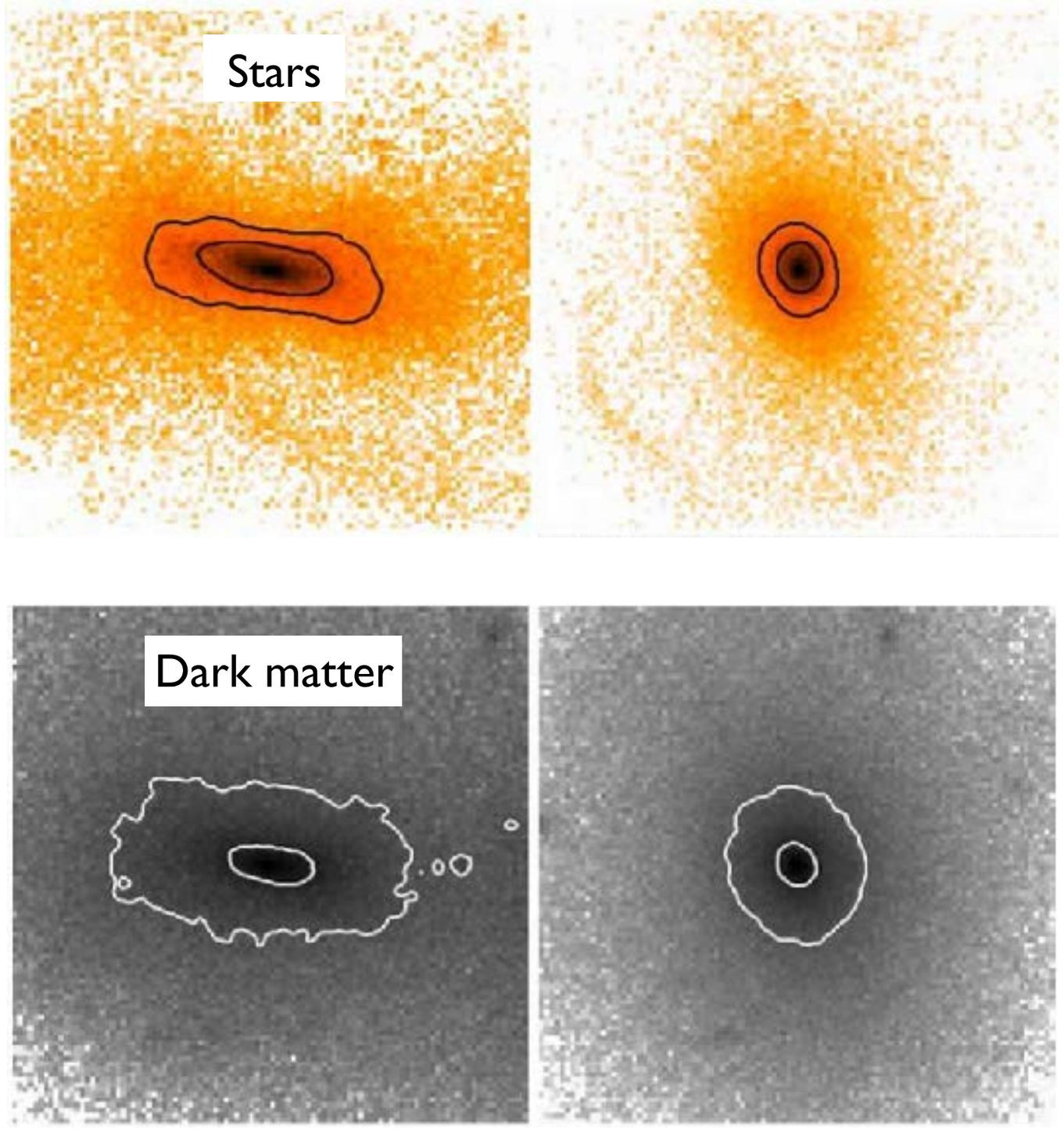
Nearby large galaxies are mostly disks and spheroids — but they start out looking more like pickles.

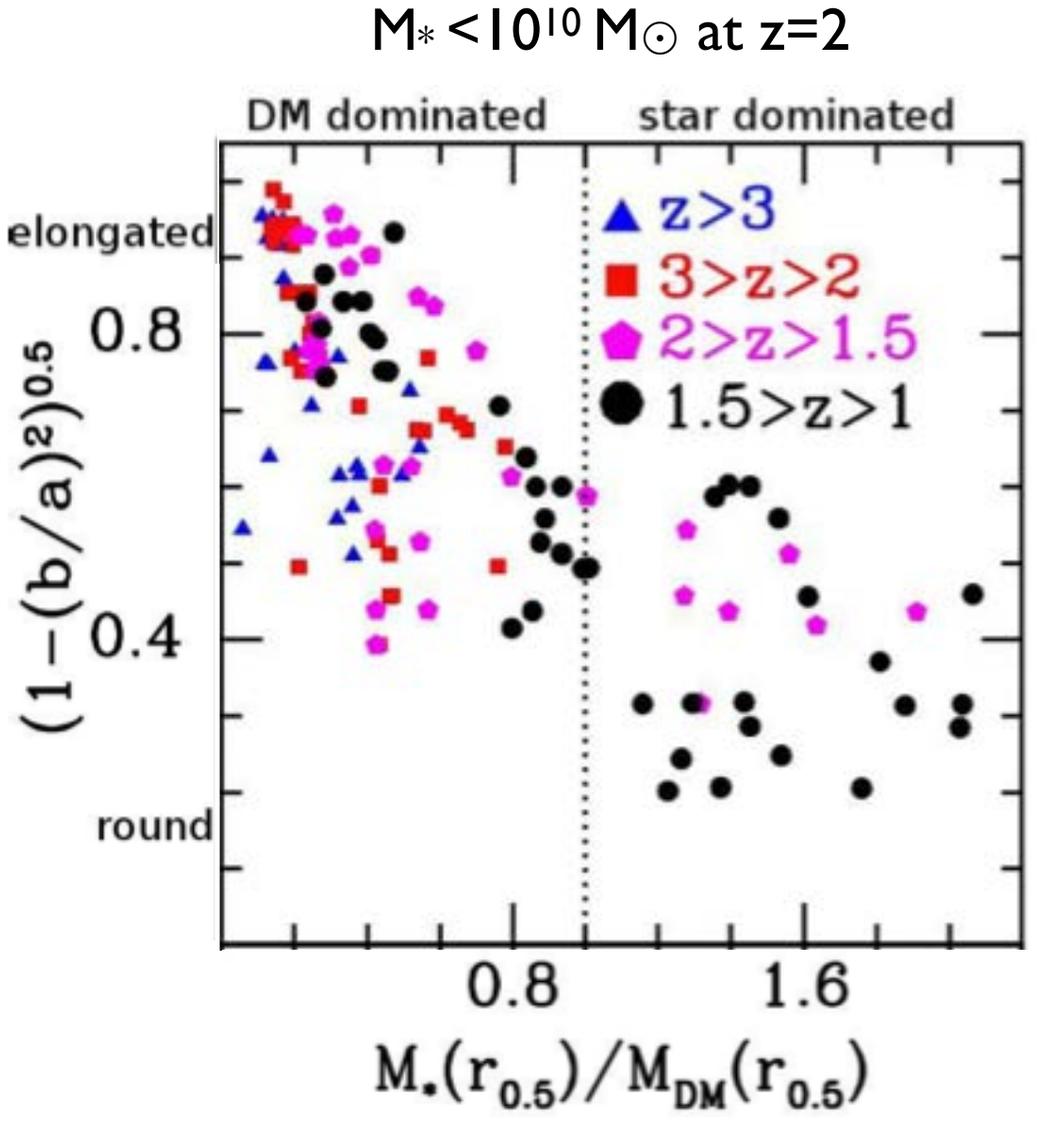


Formation of elongated galaxies with low masses at high redshift

Daniel Ceverino, Joel Primack and Avishai Dekel

MNRAS 2015



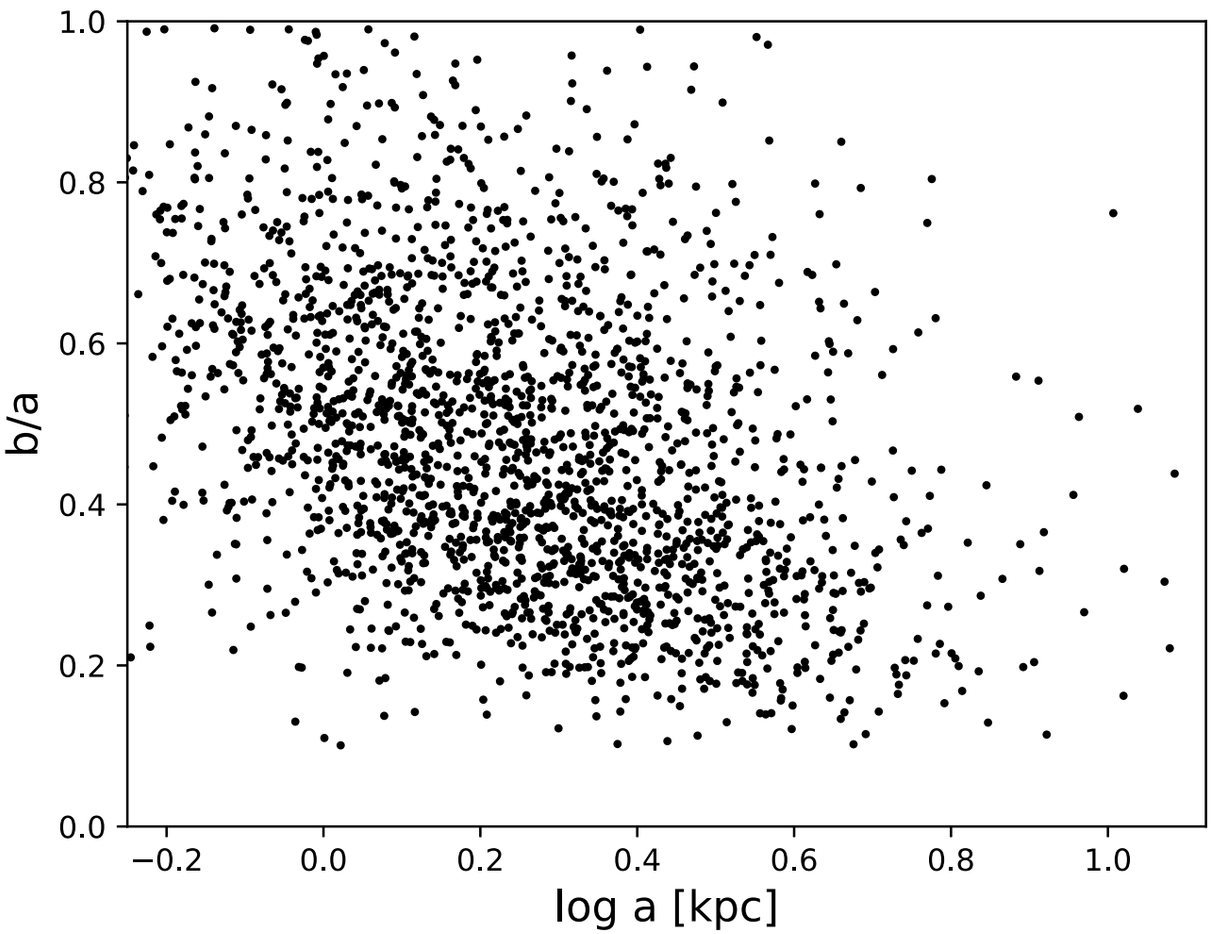


Also Tomassetti et al. 2016 MNRAS

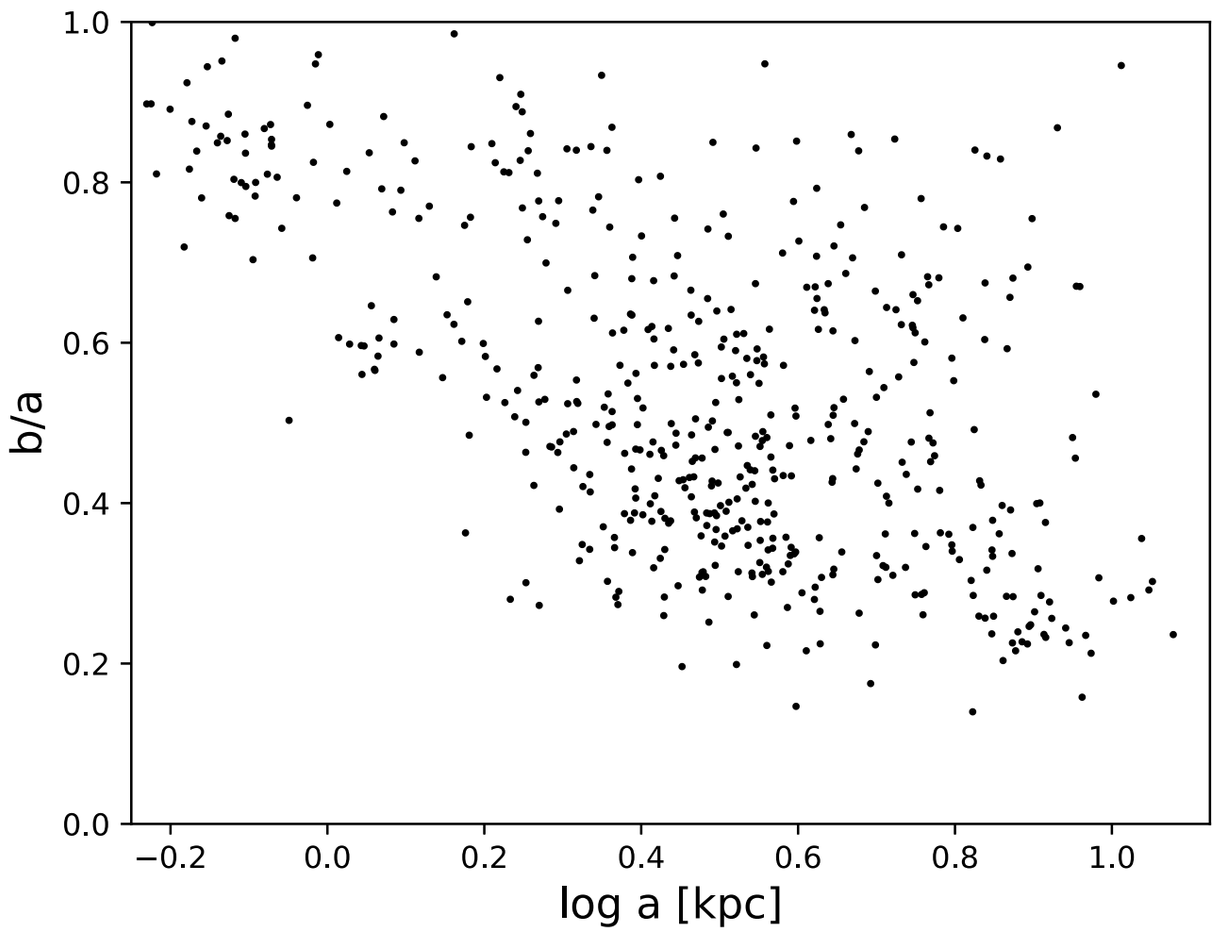
Simulated elongated galaxies are aligned with cosmic web filaments, become round after compaction (gas inflow fueling central starburst)

Compare Observed Galaxies with Simulated Galaxy Images

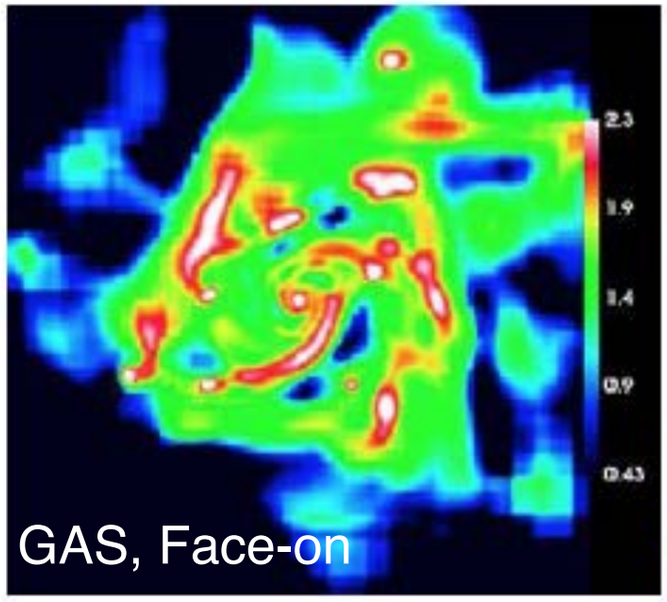
We convert our VELA galaxy simulations to realistic images using our Sunrise ray-tracing radiative transfer code to follow evolving starlight and its scattering and attenuation by dust, taking into account Hubble Space Telescope resolution (*CANDELization*). We observe the simulated galaxies with randomly located cameras, and analyze the images exactly like the real HST observations, using the GALFIT routine. The HST prolate galaxies are very similar to the simulated ones:

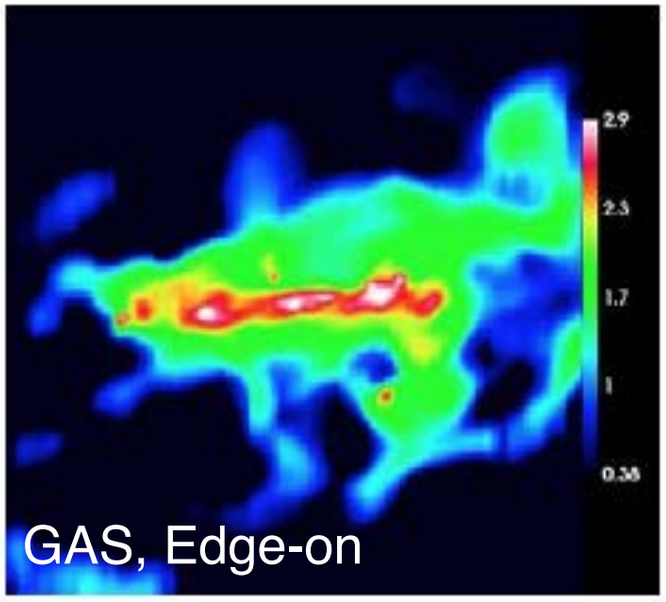


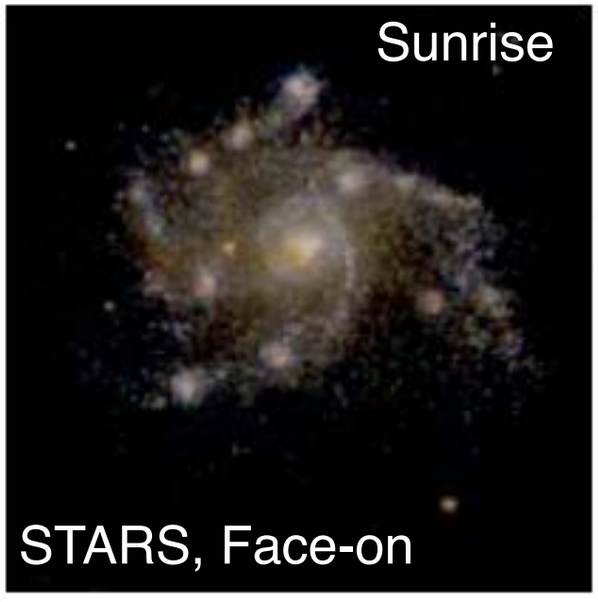
highest redshift, lowest mass CANDELS galaxies
~ 70% prolate

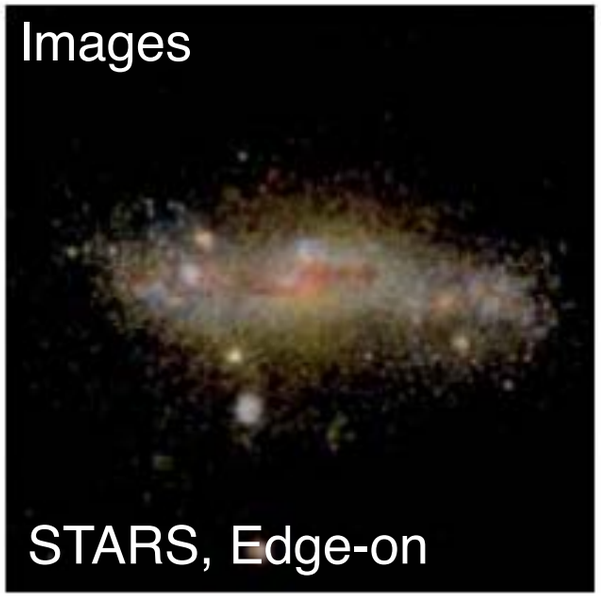


VELA simulated galaxies with prolate mass distributions
Zhang+2018 in prep.



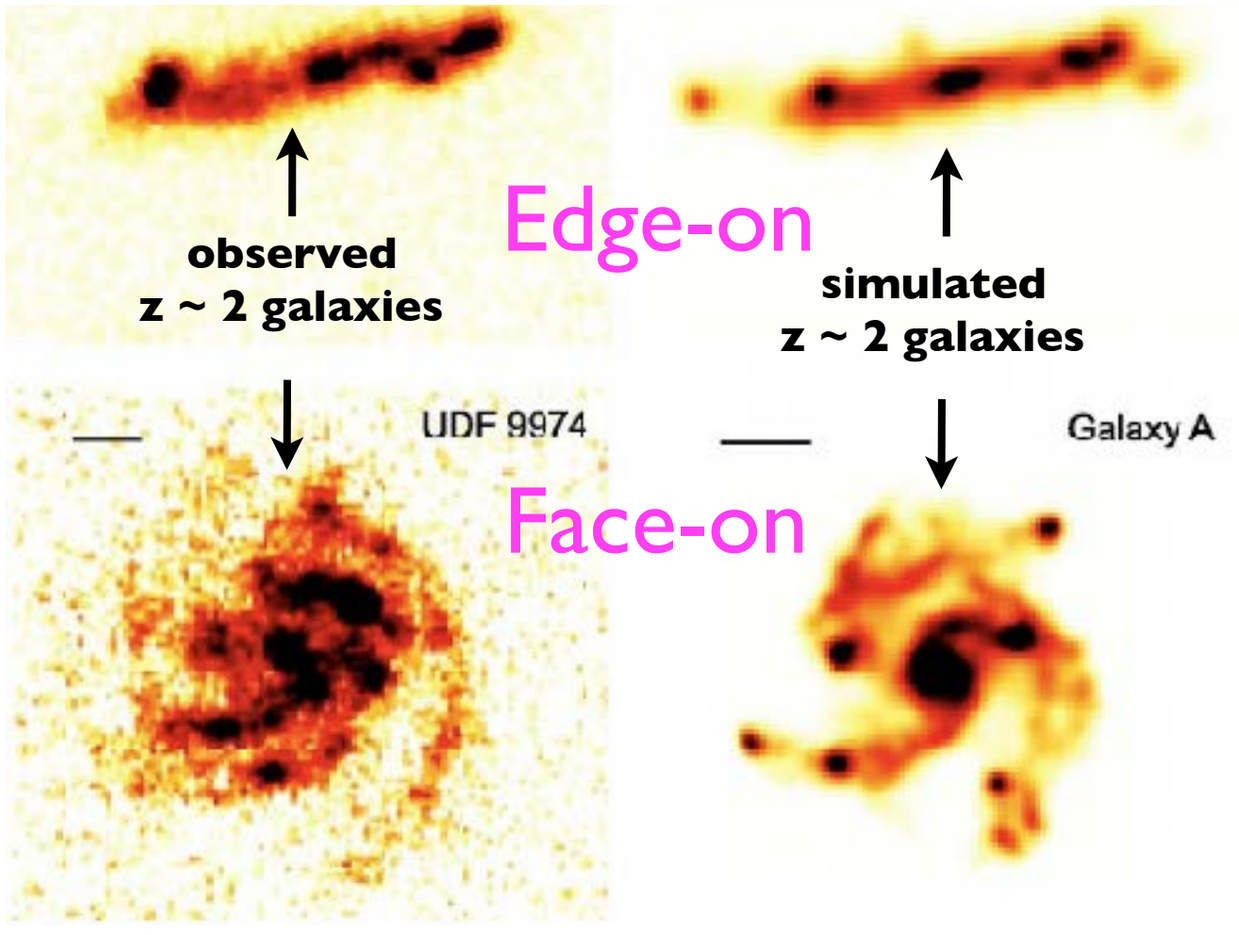




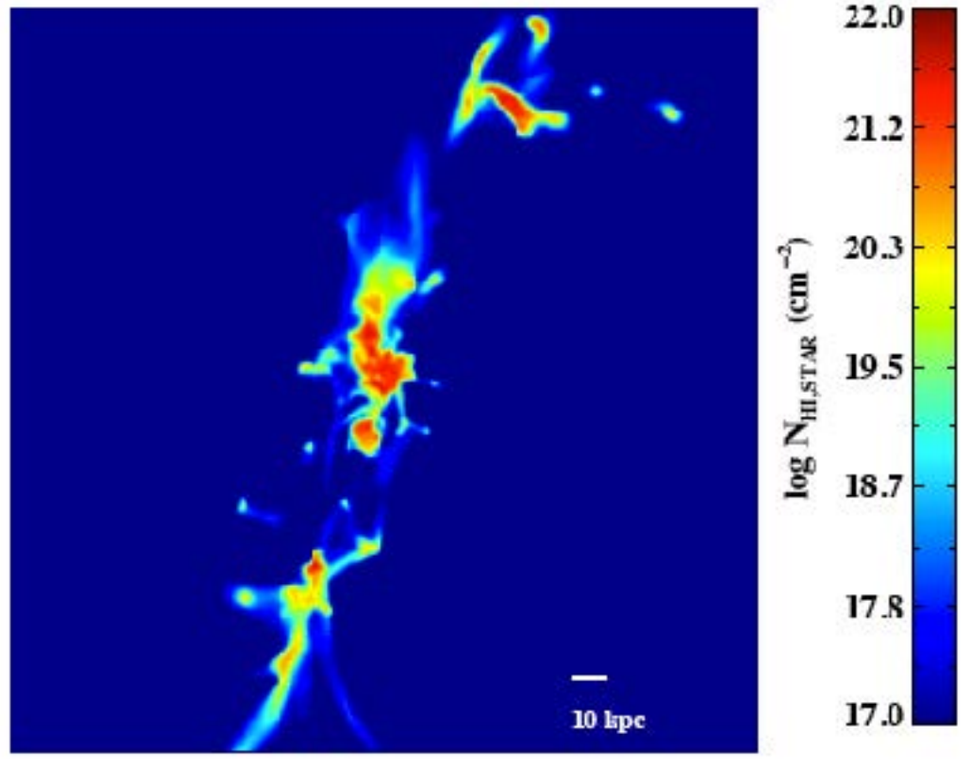


Simulated Clumpy Galaxies Compared with Observed Galaxies

Figure 1: Violently unstable disks in $\sim 10^{11}\mathrm{M}_\odot$ halos with $\sim 10^{9}\mathrm{M}_\odot$ clumps at $z = 2.3$: (a) face-on, (b) edge-on (Ceverino et al. 2009, resolution 70 pc, images 10 kpc across). RGB color images of the same simulated galaxy through dust using *Sunrise*: (c) face-on, (d) edge-on, illustrating how the clumps can be reddened and obscured when viewed edge-on.



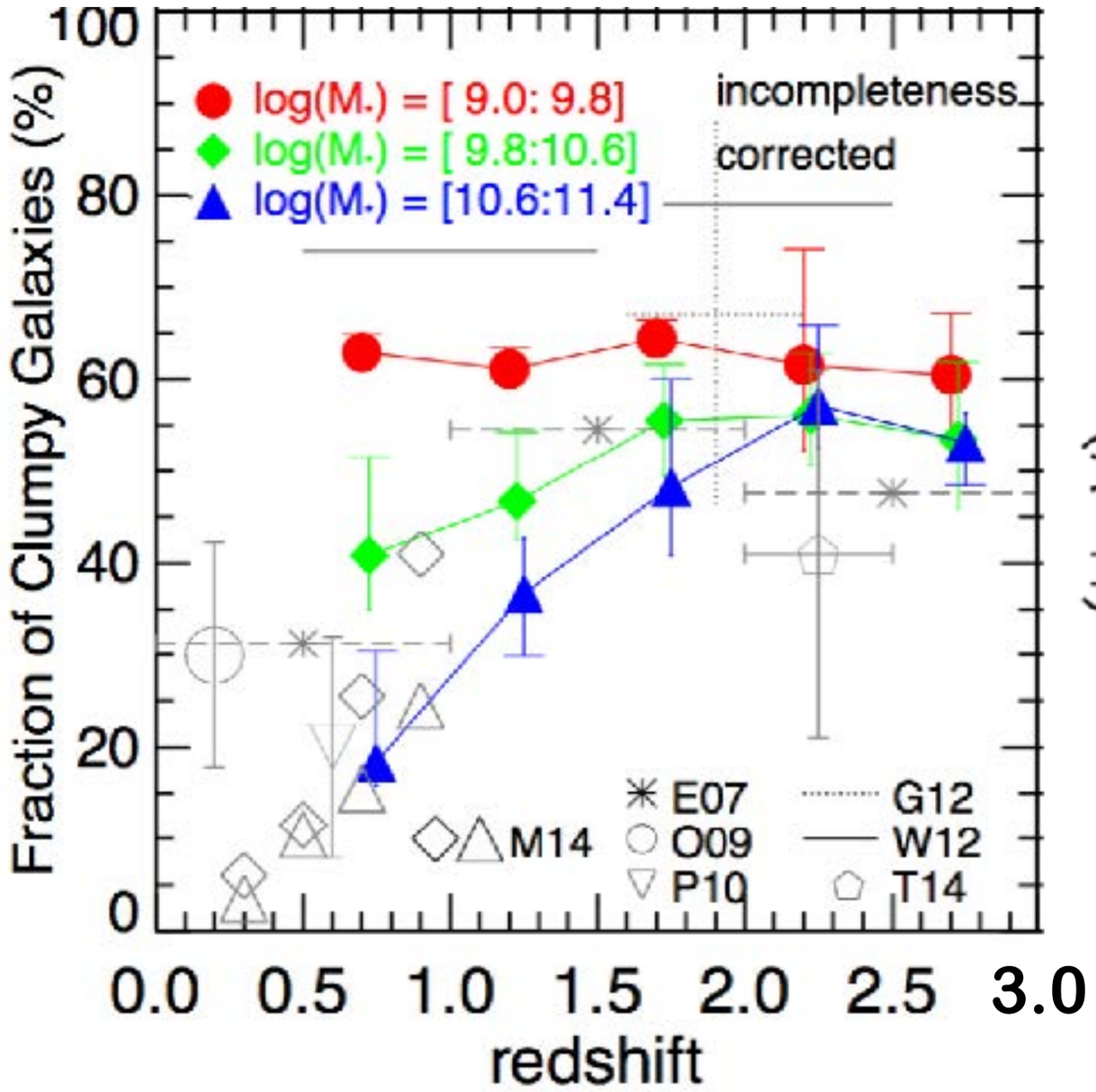
Ly alpha blobs from same simulation



Fumagalli, Prochaska, Kasen, Dekel, Ceverino, & Primack 2011

About 60% of star-forming galaxies are clumpy at z ~ 2.5.

The evolution of the clump fraction is mass-dependent.



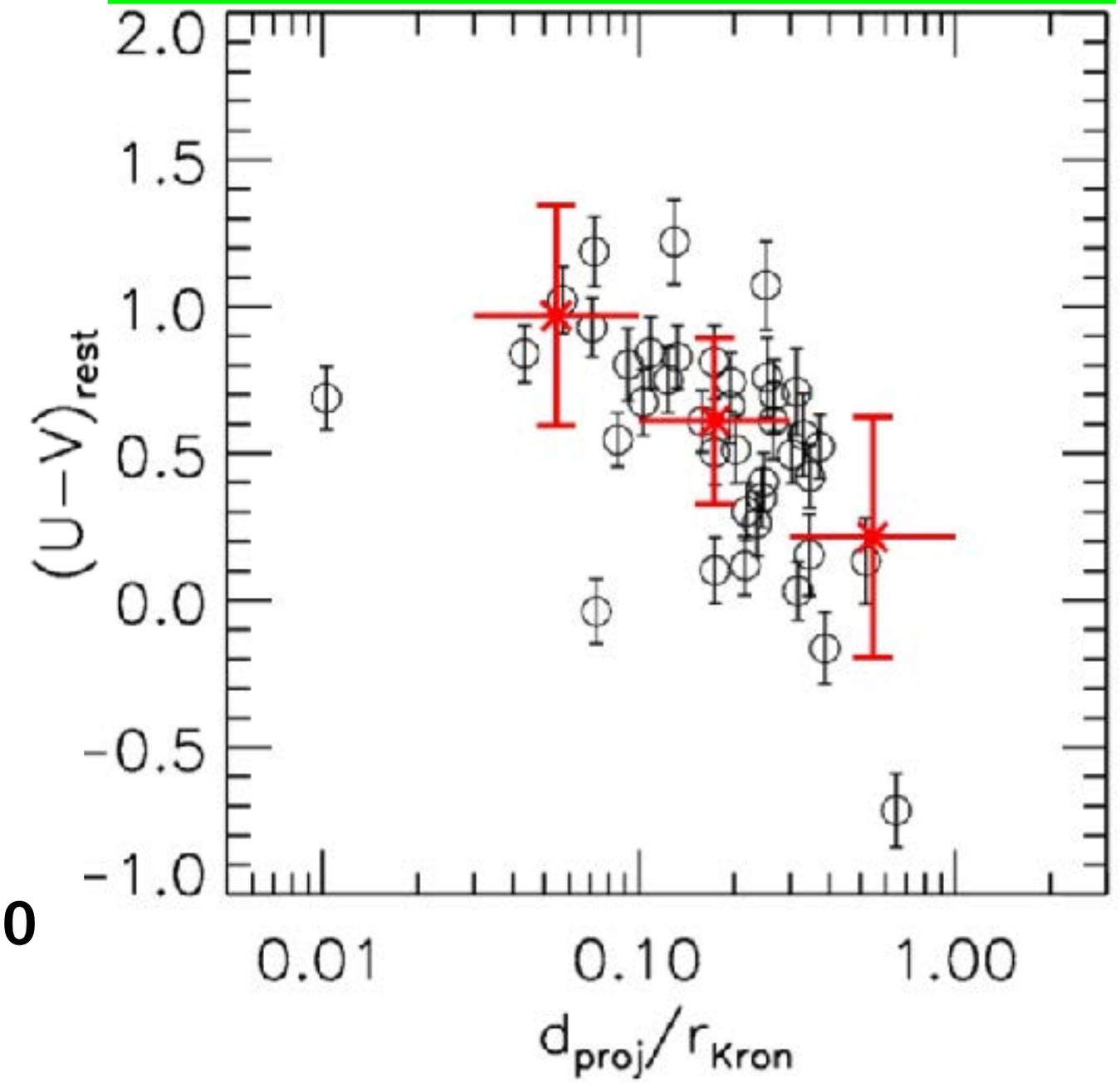
Yicheng Guo+2015

Clumps have radial variation of their UV-optical colors:

- outer clumps are bluer &

- central clumps are redder,

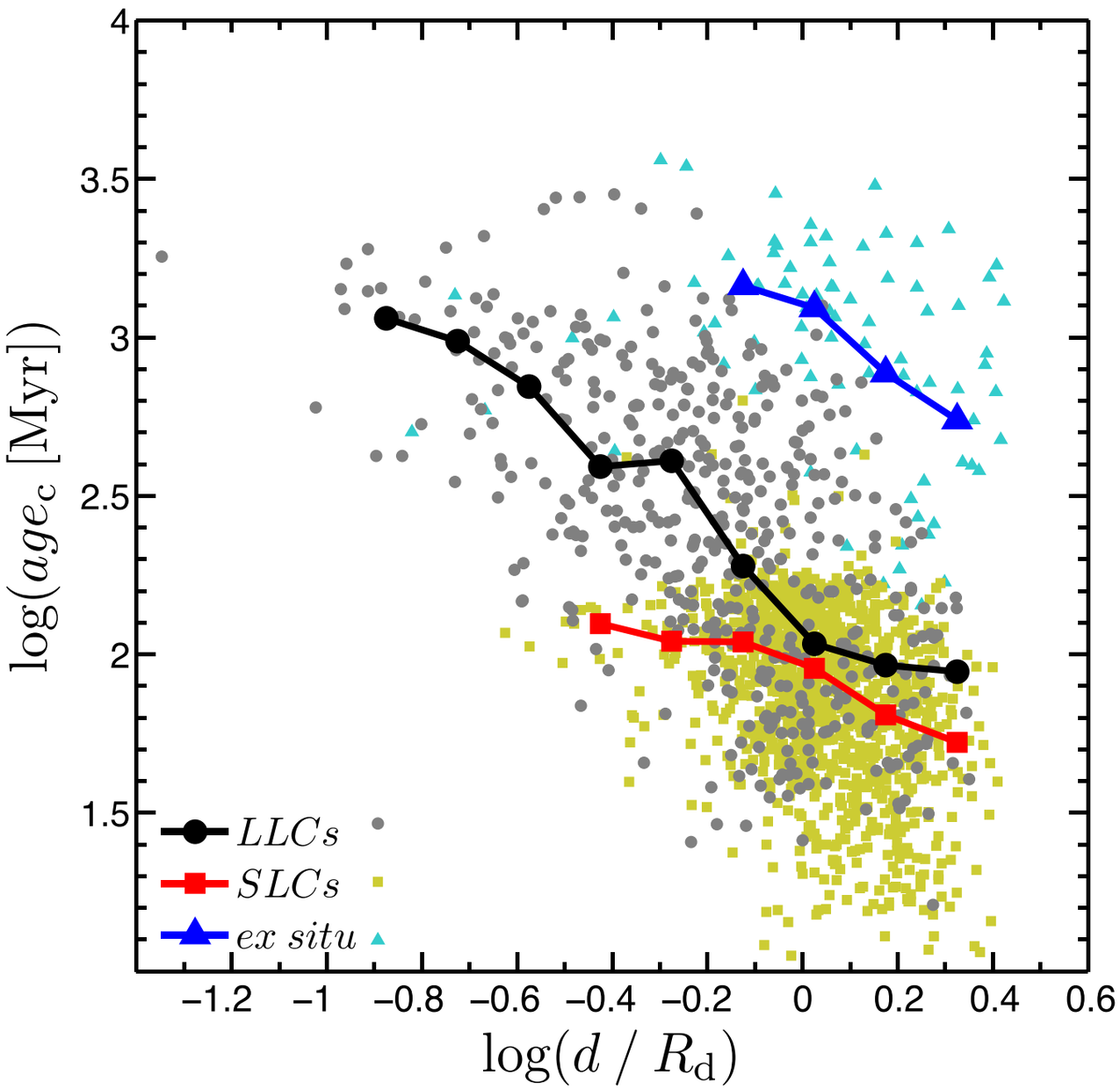
as clump radial migration predicts.



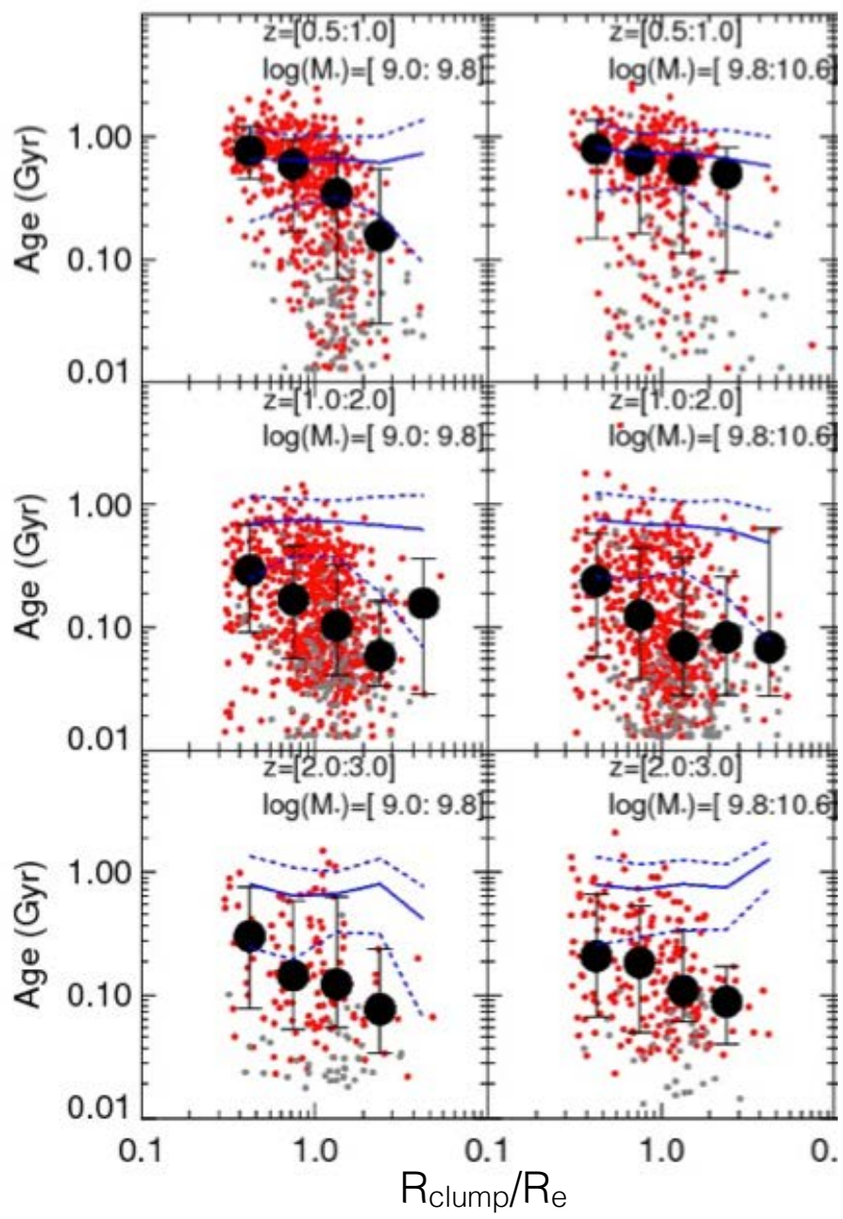
Yicheng Guo+2012

Simulated long lived clumps (LLCs) have age decreasing with radius because of clump radial migration.

Observed clumps have age decreasing with radius, different from the underlying disk, as clump radial migration predicts.

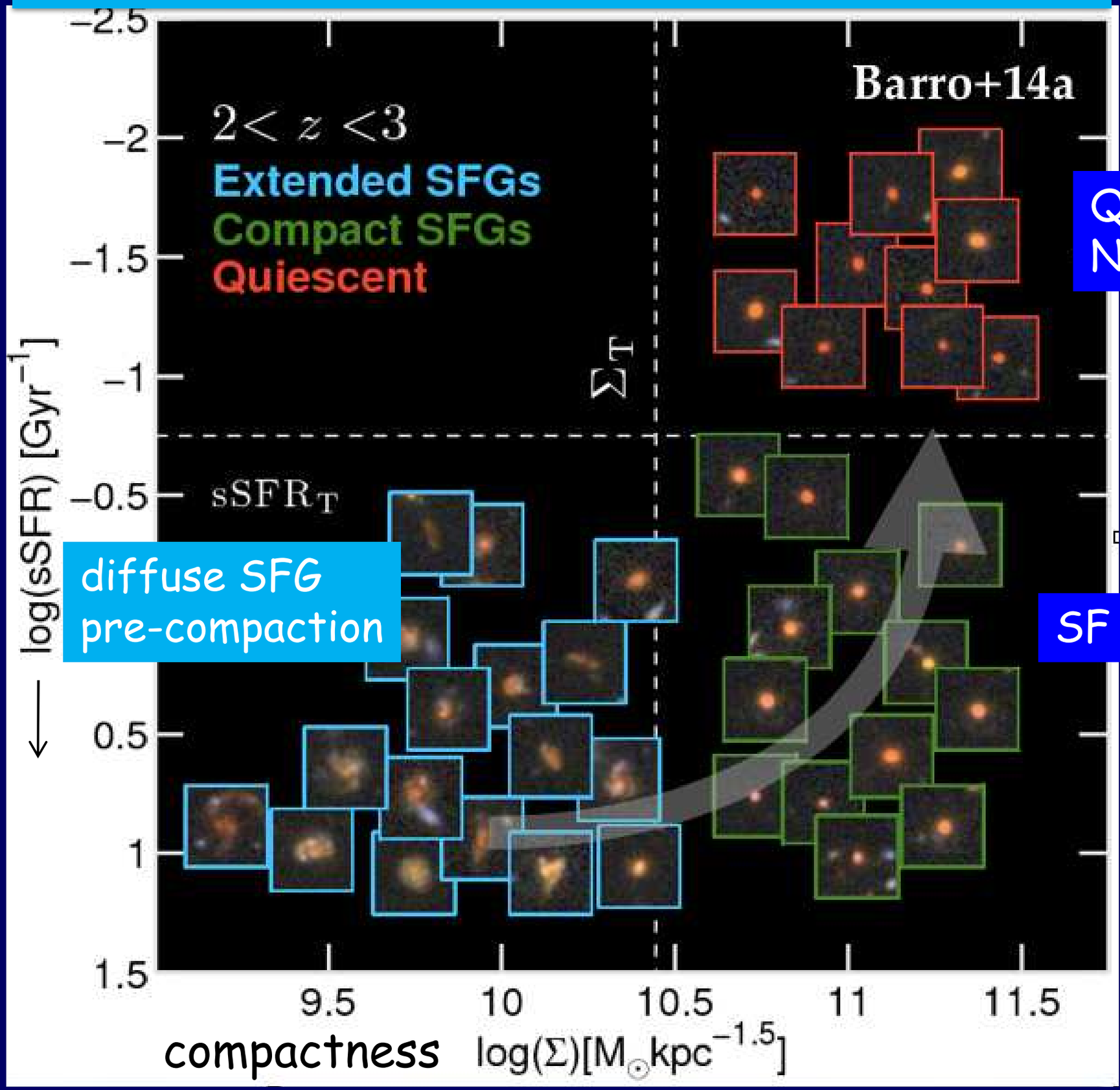


Nir Mandelker+2017



Yicheng Guo+2017

The Fast Track of Galaxy Evolution



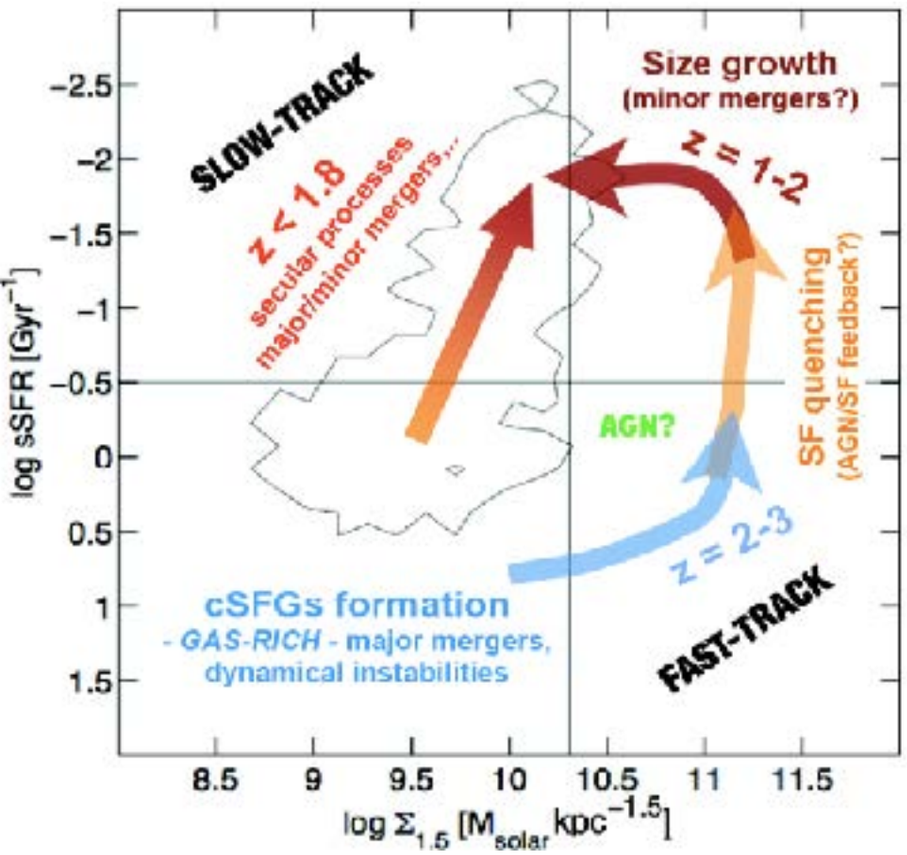
Quenched Nuggets

diffuse SFG pre-compaction

SF Nuggets

Ceverino+ RP simulations analyzed by Zolotov, Dekel, Tweed, Mandelker, Ceverino, & Primack MNRAS 2015

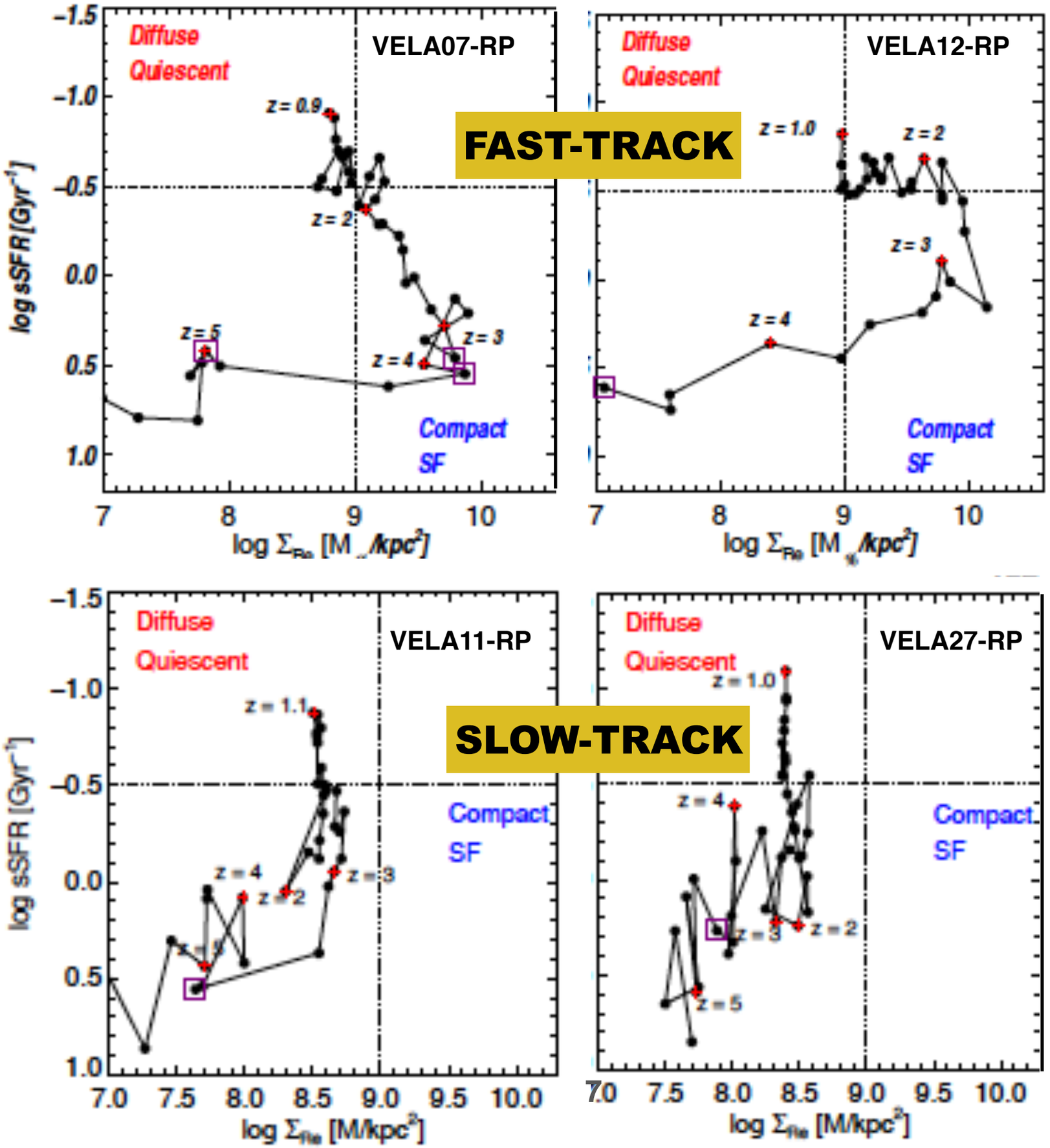
Barro+ (CANDELS) 2013



COMPACTION —>

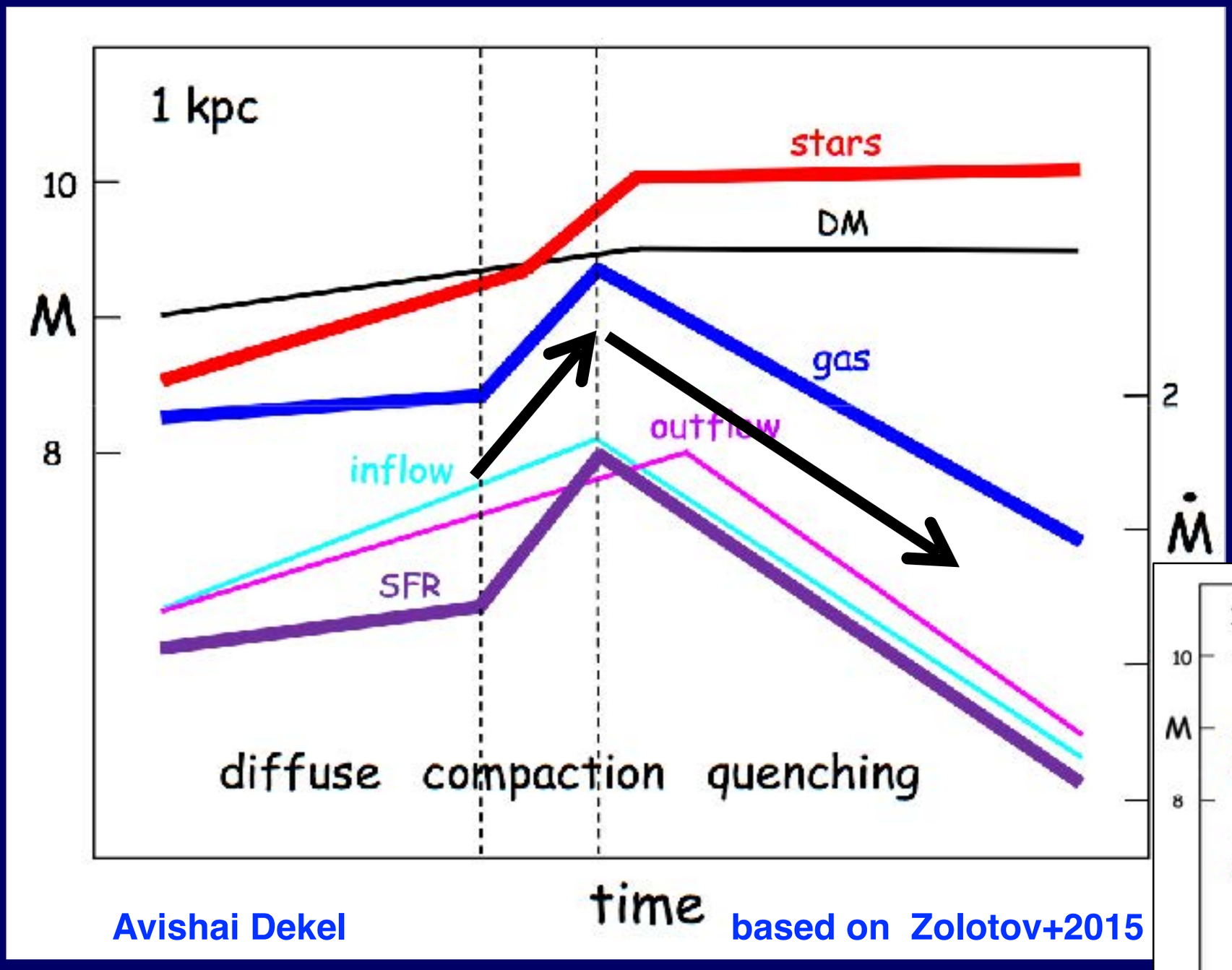
minor merger

major merger



Zolotov+2015

Compaction and Quenching in the Inner 1 kpc

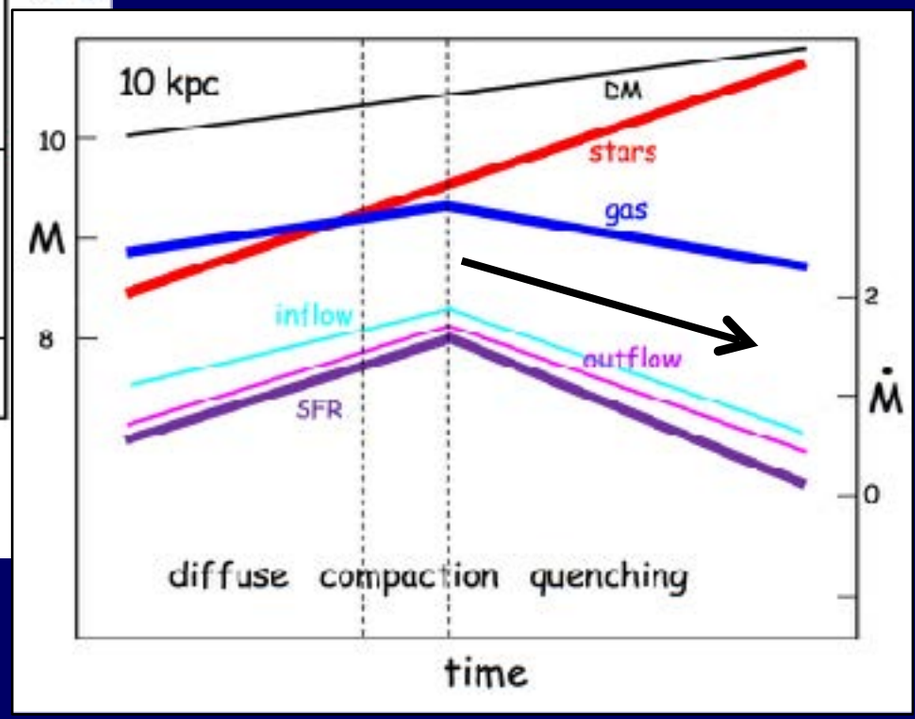


Avishai Dekel

based on Zolotov+2015

inner 1 kpc

Inner 10 kpc



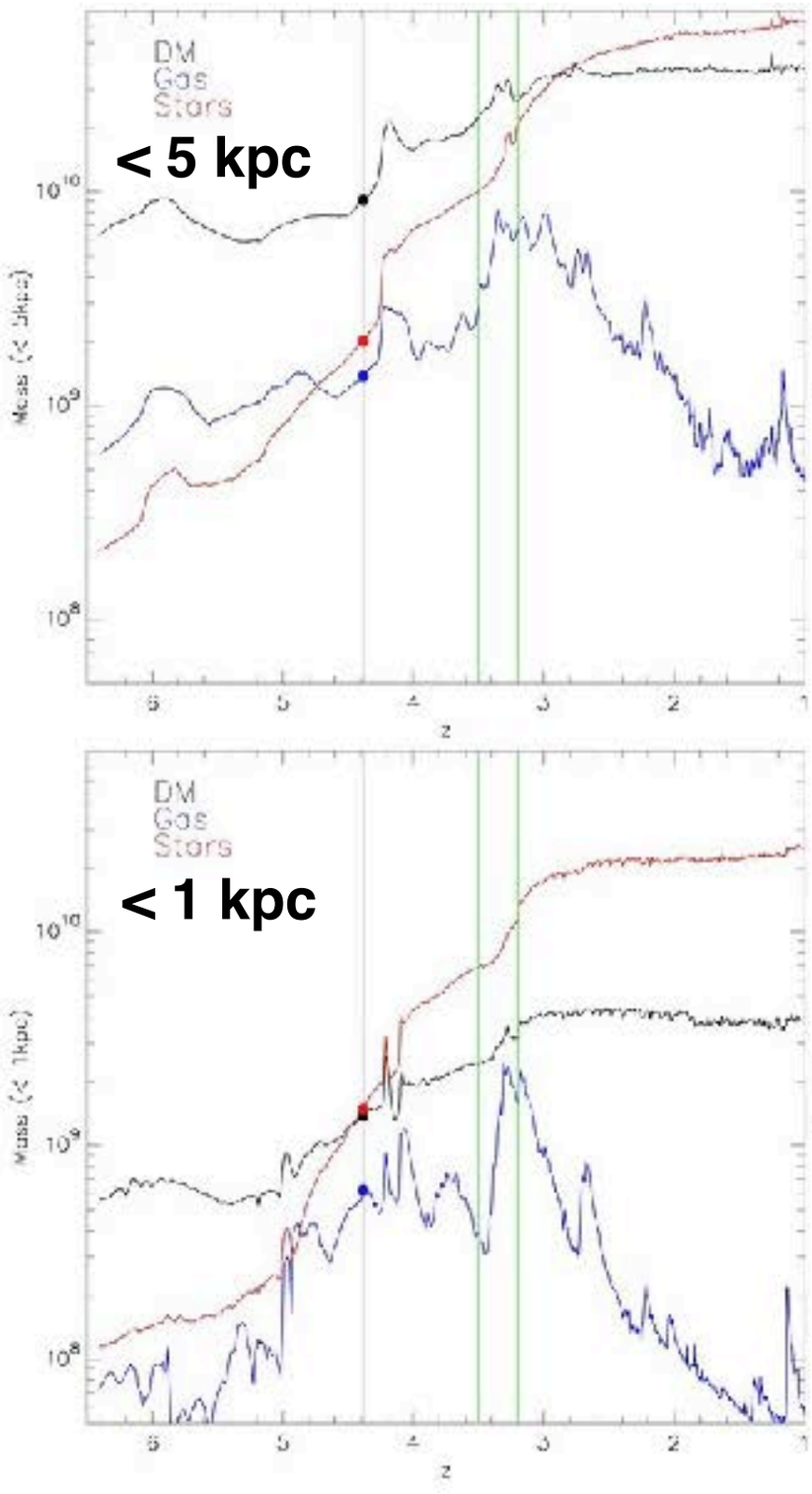
Gen 3 VELA07-RP Animations z = 4.4 to 2.3

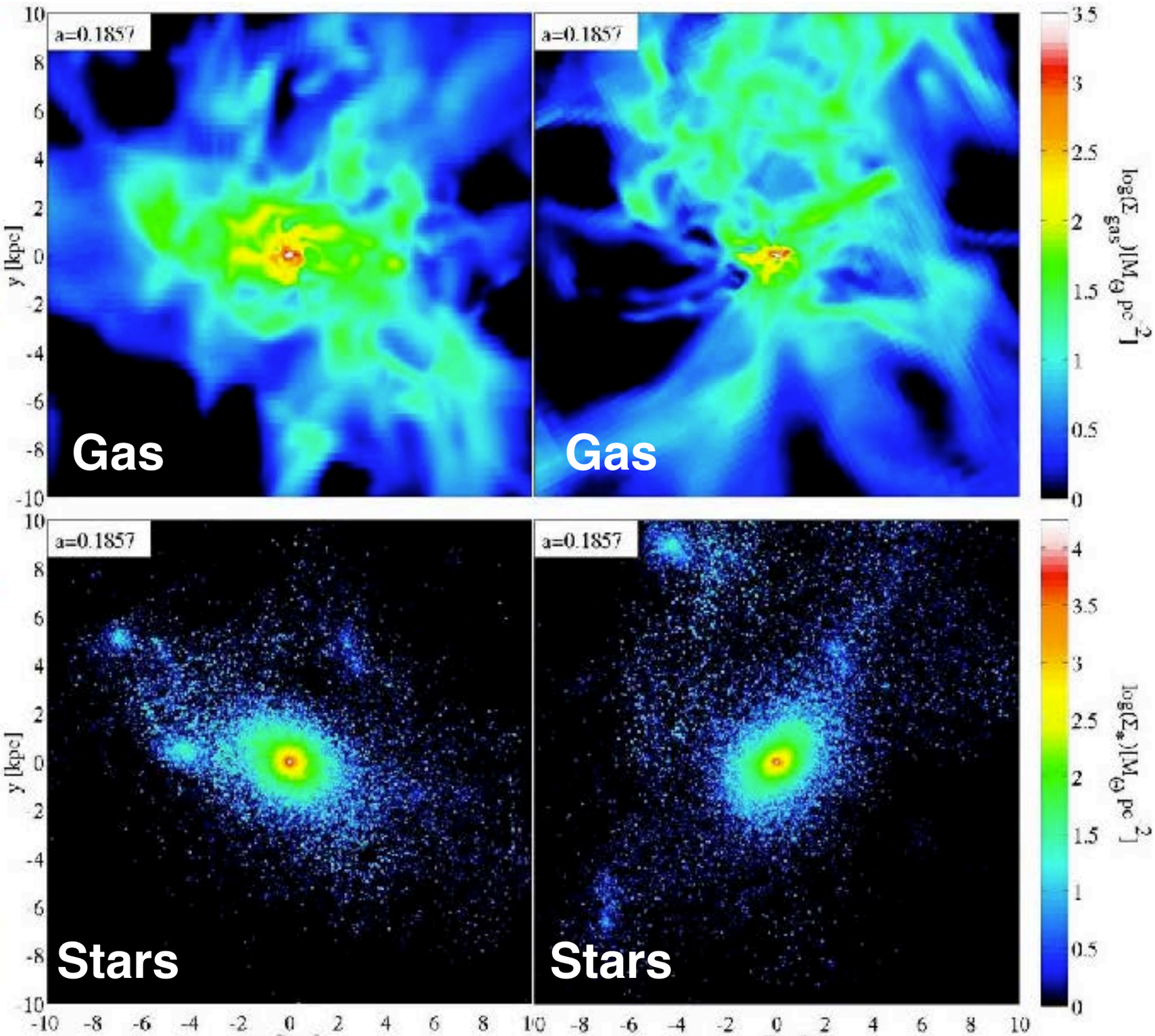
DM
Gas
Stars

Compaction

Face-on

Edge-on





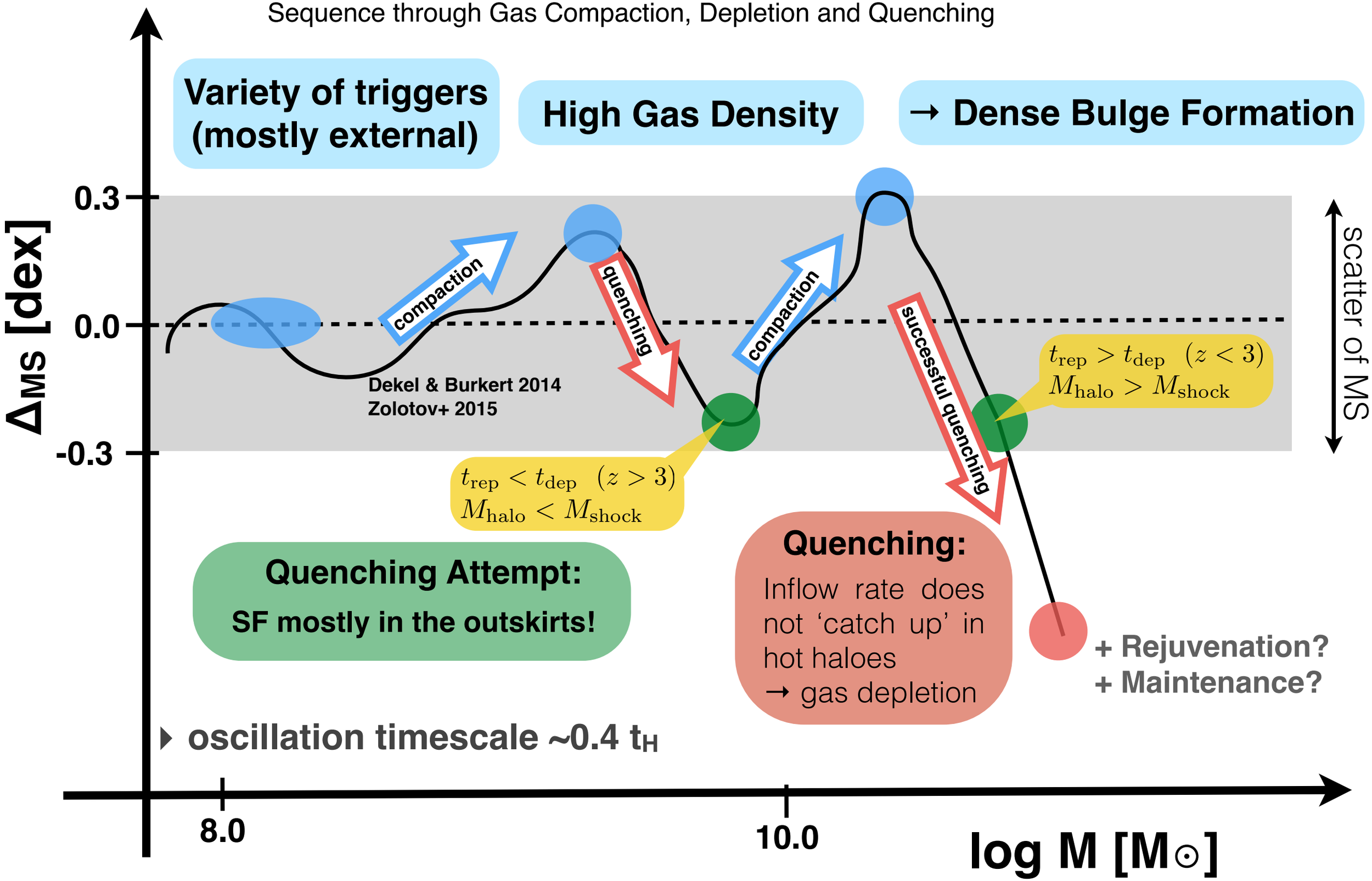
Evolution of Galaxies about the Star-Forming Main Sequence

Tacchella+2016b The Confinement of Star-Forming Galaxies into a Main Sequence through Gas Compaction, Depletion and Quenching

Variety of triggers (mostly external)

High Gas Density

→ Dense Bulge Formation



Quenching Attempt:

SF mostly in the outskirts!

Quenching:

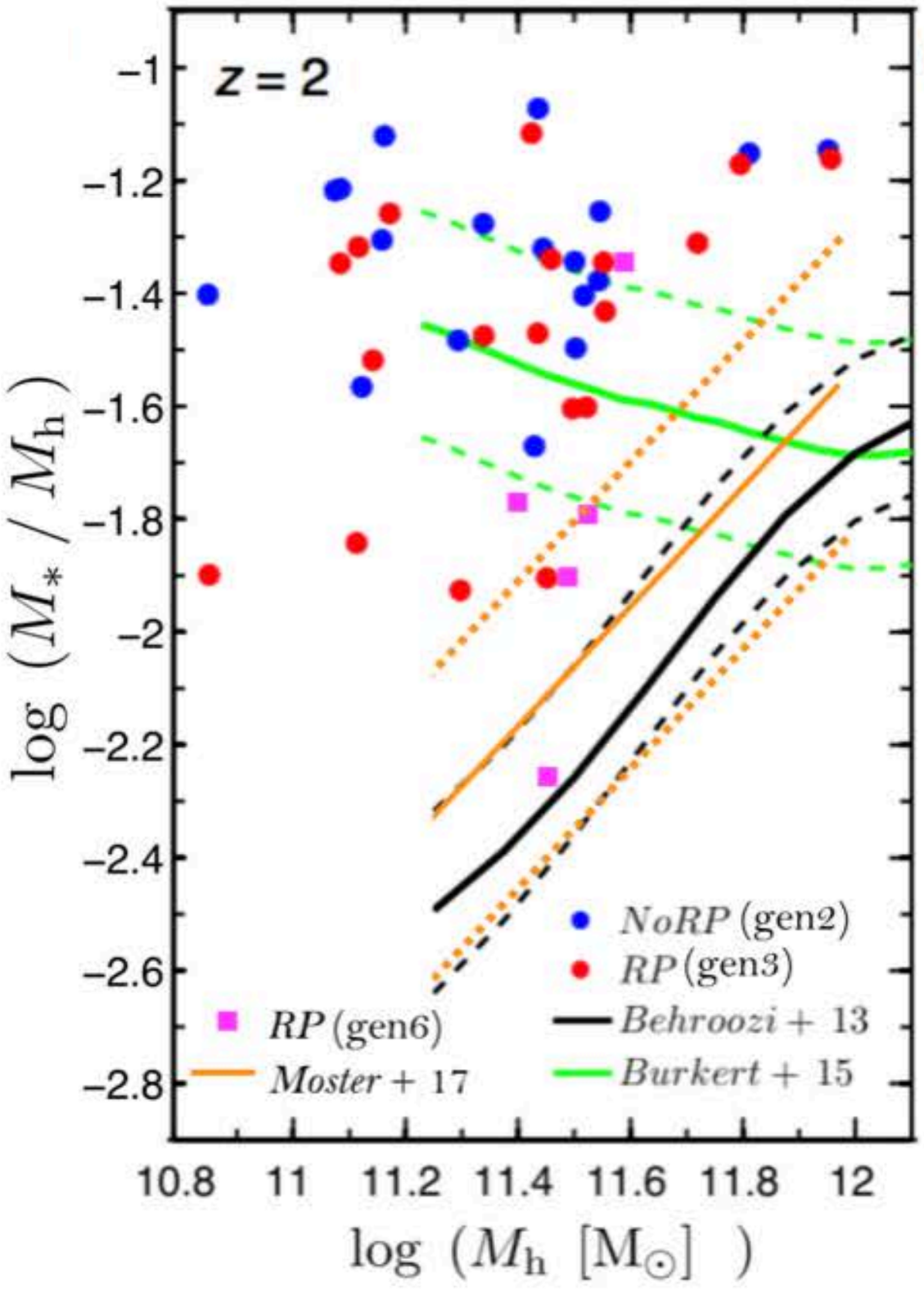
Inflow rate does not 'catch up' in hot haloes
→ gas depletion

+ Rejuvenation?
+ Maintenance?

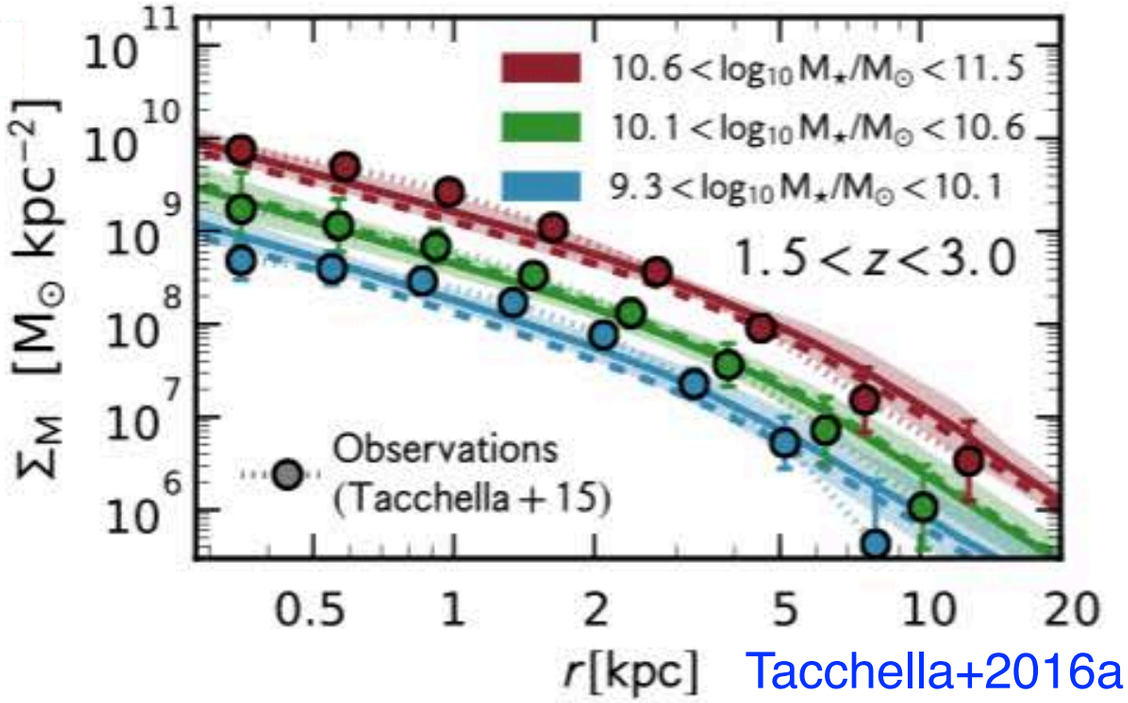
oscillation timescale ~0.4 t_H

Simulated and Observed Galaxies Are Very Similar

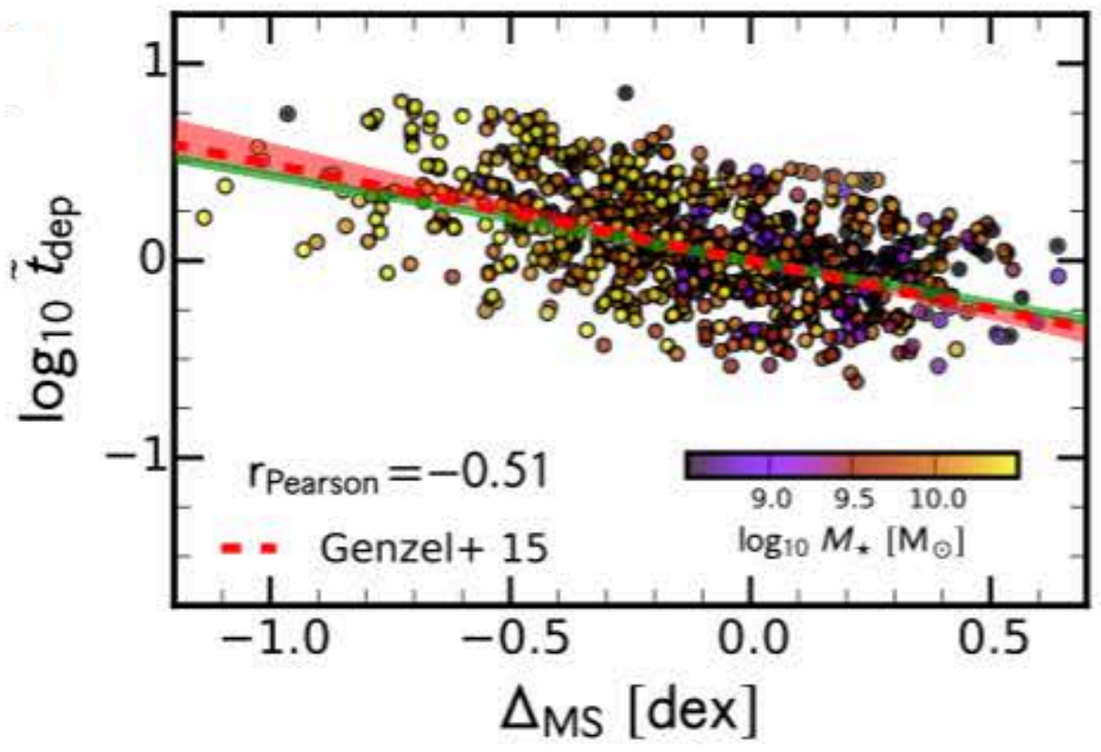
Stellar Mass/Halo Mass



Stellar Mass Profiles



Tacchella+2016a



Depletion Time vs. $\Delta_{\text{MainSequence}}$

Most astronomers used to think

(1) that galaxies form as disks,

(2) that forming galaxies are pretty smooth, and

(3) that galaxies generally grow in radius as they grow in mass.

But CANDELS and other HST observations show that all these assumptions were wrong!

(1) The majority of star-forming galaxies at z > 1 apparently have mostly elongated (prolate) stellar distributions rather than disks or spheroids, and our simulations may explain why.

(2) A large fraction of star-forming galaxies at redshifts 1 < z < 3 are found to have massive stellar clumps; these originate from phenomena including mergers and disk instabilities in our simulations.

(3) These phenomena also help to create compact stellar spheroidal galaxies ("nuggets") through galaxy compaction (rapid inflow of gas to galaxy centers where it forms stars in our simulations).

Our VELA galaxy simulations have been sufficiently successful in explaining these phenomena that we want to use the mock image plus simulation metadata as a *deep learning training set*, to see if a trained DL can determine compaction status from observations.

A toy model of a neuron: “perceptron”

Frank Rosenblatt
1928-1971

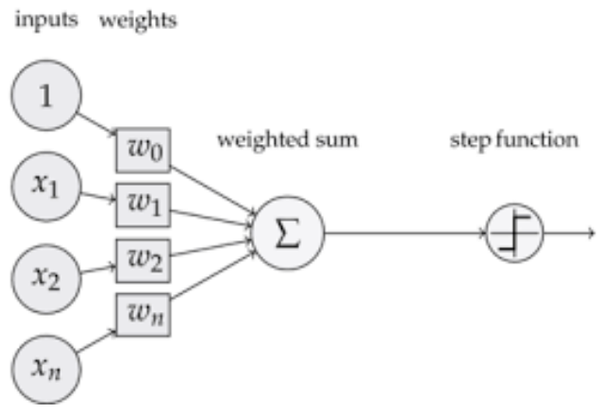
vs.

Marvin Minsky
1927-2016

Simplify the neuron to a sum over weighted inputs and a nonlinear activation function.

- no spikes
- no recurrence or feedback *
- no dynamics or state *
- no biophysics

$$y = f(\sum_i w_i x_i + b)$$



$$f(z) = max(0, z)$$

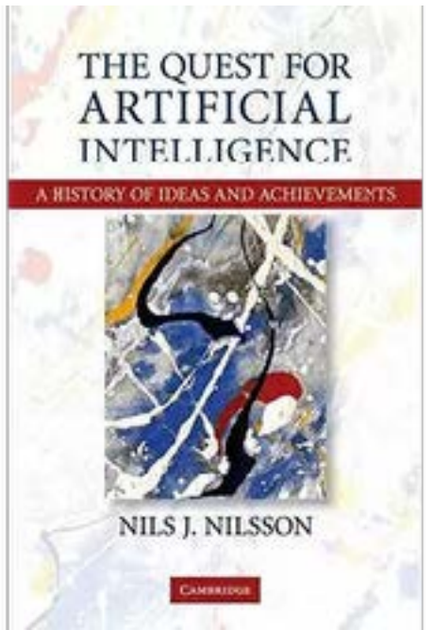
The perceptron: a probabilistic model for information storage and organization in the brain.
F Rosenblatt (1958)

“During the late 1950s and early 1960s … Rosenblatt and Minsky debated on the floors of scientific conferences the value of biologically inspired computation, Rosenblatt arguing that his neural networks could do almost anything and Minsky countering that they could do little.”

Web version of *The Quest for Artificial Intelligence* by Nils Nilsson, nicely covers Minsky and Rosenblatt (as well as a lot of other relevant AI material).

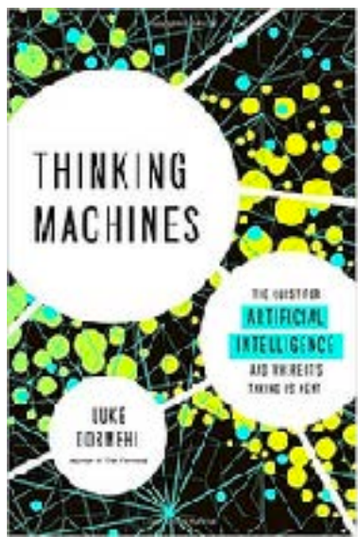
During the 1960s, neural net researchers employed various methods for changing a network's adjustable weights so that the entire network made appropriate output responses to a set of "training" inputs. For example, **Frank Rosenblatt** at Cornell adjusted weight values in the final layer of what he called the three-layer alpha-perceptron. But what stymied us all was how to change weights in more than one layer of multilayer networks…

That problem was solved in the mid-1980s by the invention of a technique called "back propagation" (**backprop** for short) introduced by David Rumelhart, **Geoffrey E. Hinton**, and Ronald J. Williams. In response to an error in the network's output, backprop makes small adjustments in all of the weights so as to reduce that error. It can be regarded as a hill-descending method – searching for low values of error over the landscape of weights. Backprop uses calculus to precompute the best set of weight changes. Starting in 2012, deep learning methods on powerful GPUs have outperformed all traditional AI methods.

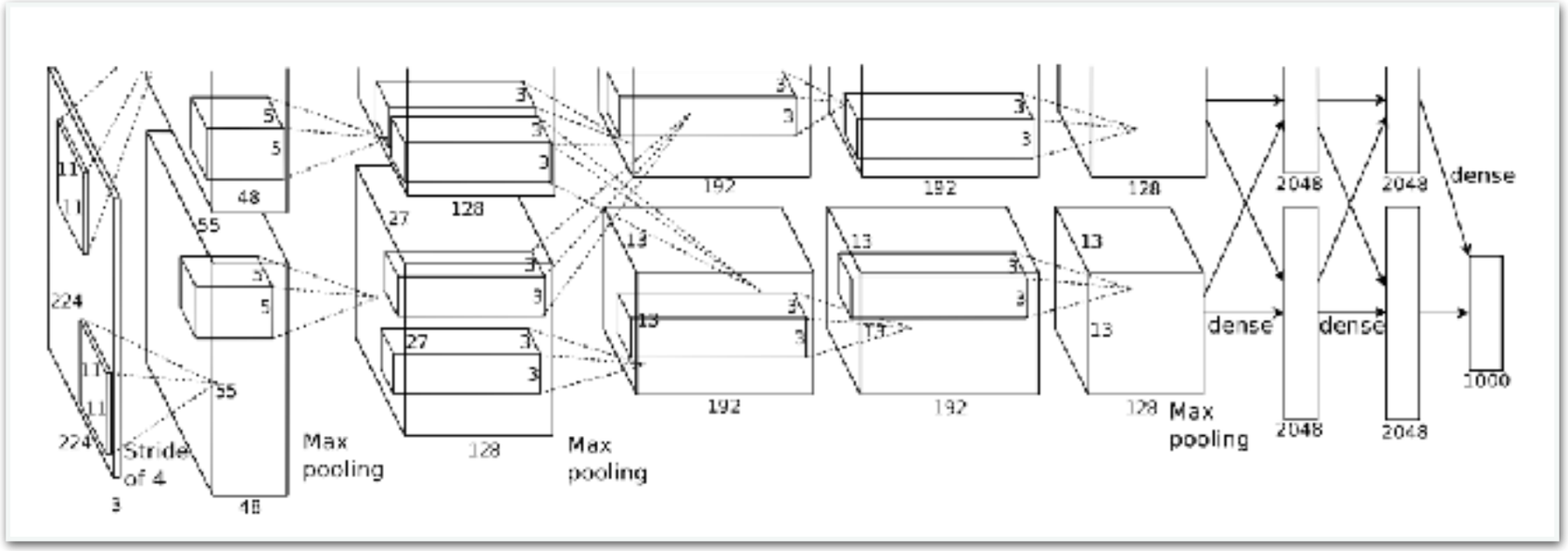


From *The Quest for Artificial Intelligence* by Nils Nilsson, Chapter 29.

See also



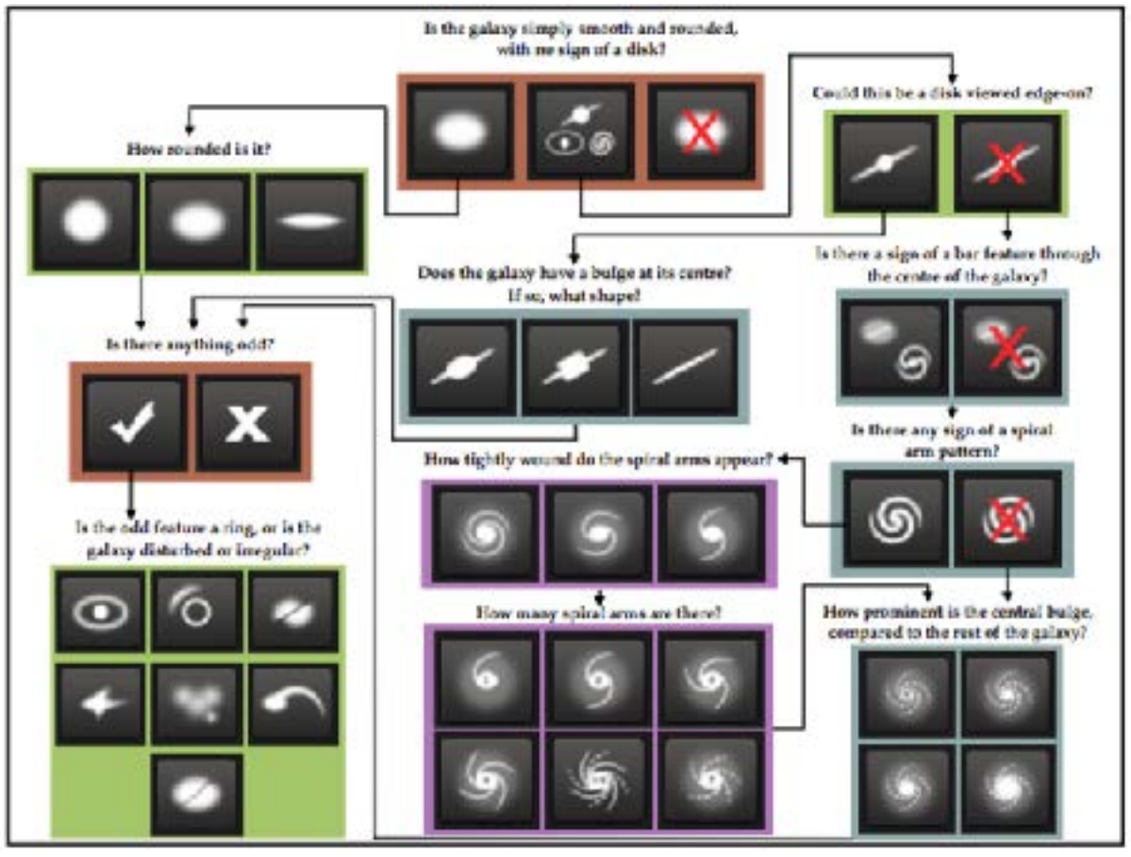
Deep convolutional neural networks



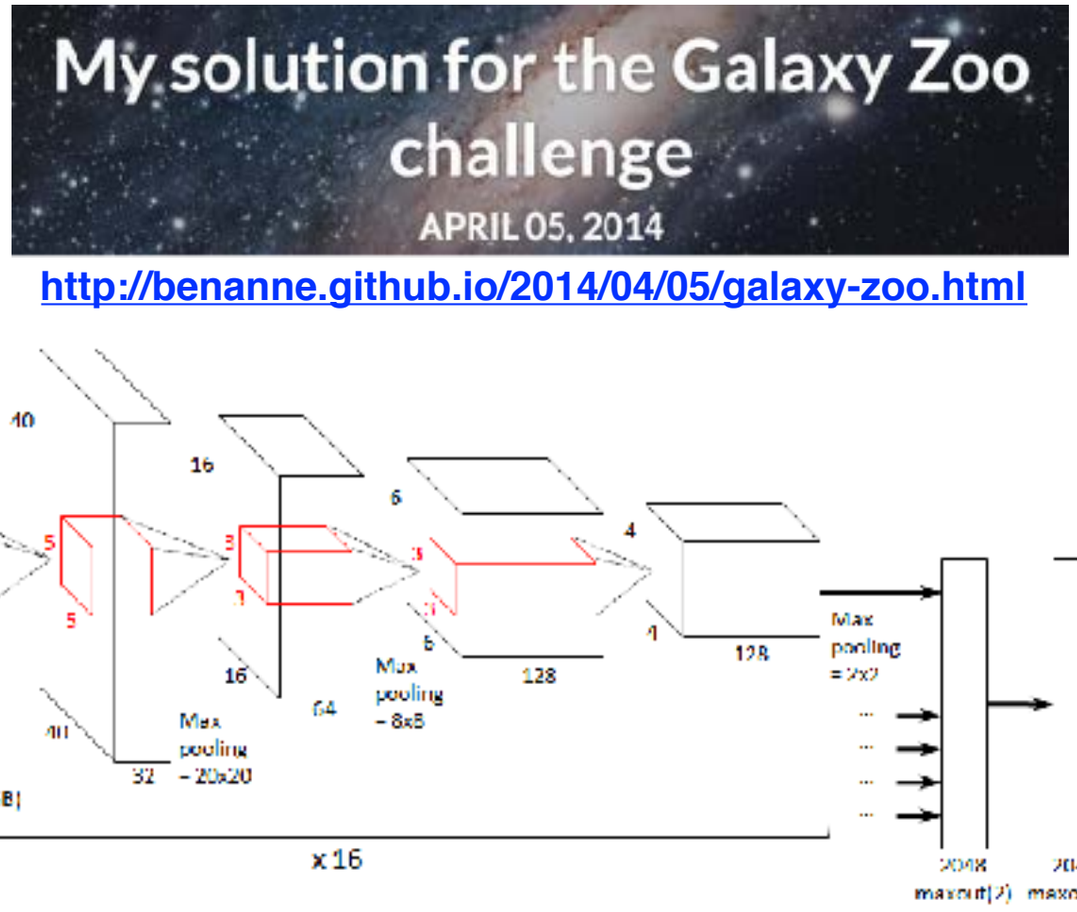
ImageNet Classification with Deep Convolutional Neural Networks
A Krizhevsky I Sutskever, G Hinton (2012)

- Multi-layer perceptron trained with back-propagation are ideas known since the 1980's.
- The success of deep learning in the past 6 years is due to more powerful computers (GPUs) and better code.

Sander Dieleman used a deep learning code to predict Galaxy Zoo's nearby galaxy image classifications with high accuracy, winning the 2014 Kaggle competition



The Galaxy Zoo 2 decision tree. Reproduced from fig.1 in Willett et al. (2013).



http://benanne.github.io/2014/04/05/galaxy-zoo.html

Krizhevsky-style diagram of the architecture of the best performing network.

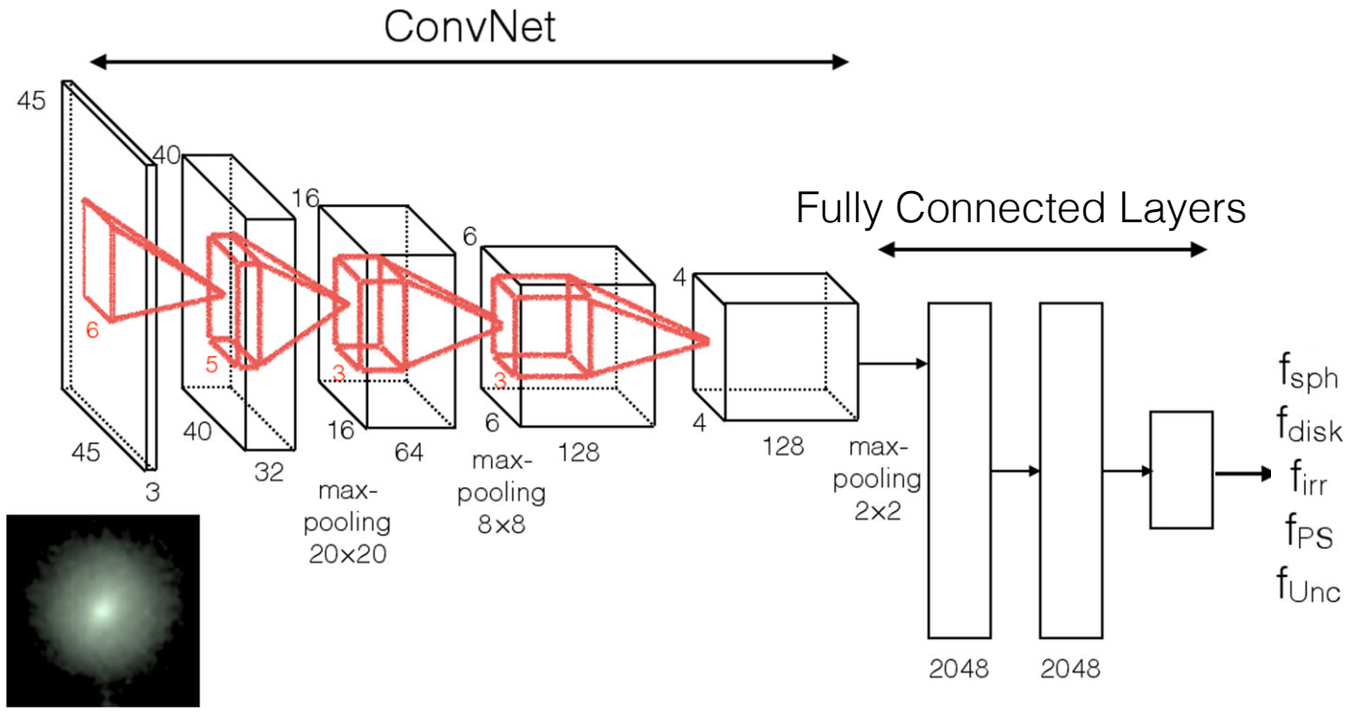
Dieleman, Willett, Dambre 2015, Rotation-invariant convolutional neural networks for galaxy morphology prediction, MNRAS

From the Abstract: We present a deep neural network model for galaxy morphology classification which exploits translational and rotational symmetry. For images with high agreement among the Galaxy Zoo participants, our model is able to reproduce their consensus with near-perfect accuracy (>99 per cent) for most questions.

Marc Huertas-Company used Dieleman's code to classify CANDELS galaxy images

H-C et al. 2015, Catalog of Visual-like Morphologies in 5 CANDELS Fields Using Deep Learning

In this work, we mimic human perception with deep learning using convolutional neural networks (ConvNets). The ConvNet is trained to reproduce the CANDELS visual morphological classification based on the efforts of 65 individual classifiers who contributed to the visual inspection of the ~8000 galaxies in the GOODS-S field. It was then applied to the other four CANDELS fields. The galaxy classification data was then released to the astronomical community.



Configuration of the Convolutional Neural Network used in this paper, based on the one used by Dieleman et al. (2015) on SDSS galaxies. It is made of 5 convolutional layers followed by 2 fully connected perceptron layers.

Following the approach in CANDELS, we associate five real numbers with each galaxy corresponding to the frequency at which expert classifiers flagged a galaxy as having a bulge, having a disk, presenting an irregularity, being compact or point-source, and being unclassifiable. Galaxy images are interpolated to a fixed size, rotated, and randomly perturbed before feeding the network to (i) avoid over-fitting and (ii) reach a comparable ratio of background versus galaxy pixels in all images. ConvNets are able to predict the votes of expert classifiers with a <10% bias and a ~10% scatter. This makes the classification almost equivalent to a visual-based classification. The training took 10 days on a GPU and the classification is performed at a rate of 1000 galaxies/hour.

H-C et al. 2016, Mass assembly and morphological transformations since z ~ 3 from CANDELS

We quantify the evolution of star-forming and quiescent galaxies as a function of morphology from z ~ 3 to the present. Our main results are: 1) At z ~ 2, 80% of the stellar mass density of star-forming galaxies is in irregular systems. However, by z ~ 0.5, irregular objects only dominate at stellar masses below $10^9 M_\odot$.

2) Quenching: We confirm that galaxies reaching a stellar mass $M_* \sim 10^{10.8} M_\odot$ tend to quench. Also, quenching implies the presence of a bulge: the abundance of massive red disks is negligible at all redshifts

Detecting wet compaction at high redshift with deep learning

Marc Huertas-Company, Joel Primack, Avishai Dekel, David Koo, et al. - in prep. 2018

ABSTRACT

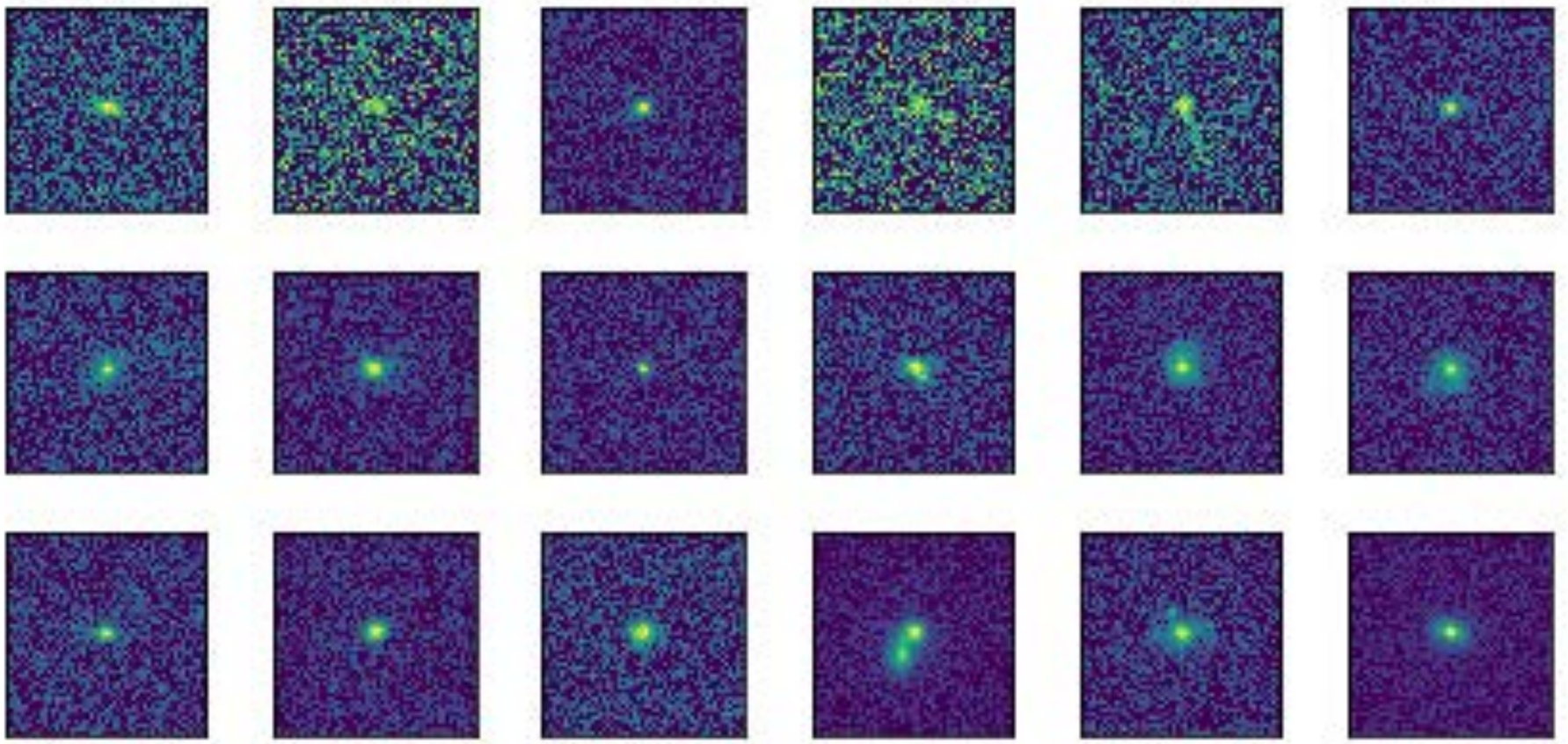
We explore a new approach to classify galaxy images from deep surveys oriented towards detecting astrophysical processes calibrated on cosmological hydrodynamic galaxy simulations. To illustrate the methodology we focus on wet compaction. Recent theoretical and observational works have suggested that compact bulges at high redshift might be formed through gas inflows (wet compaction events) before quenching. We train a simple Convolutional Neural Network (CNN) with mock *CANDELized* images from our VELA zoom-in simulations that are selected for being in a wet-compaction phase according to the assembly history extracted from the simulation. We show that the CNN is able to retrieve a galaxy in the compaction phase within a time window of ±0.3 Hubble times based only on the pixels distribution. We then use the trained network to classify real galaxies from the CANDELS survey into three classes (pre-compaction, compaction and post-compaction). We find that compaction typically occurs at a characteristic stellar mass of $\sim 10^{9.5-10}$ solar masses all redshifts, as in the VELA simulations. The galaxies that are experiencing compaction in the CANDELS redshift range ($1 < z < 3$) are therefore typically the progenitors of $\sim 10^{10.5}$ solar mass galaxies at $z \sim 0$, like the Milky Way. The presented technique can be generalized to other processes and could constitute an alternative way of classifying galaxies in the era of massive imaging surveys and cosmological simulations, to help improve the comparison between theory and observations.

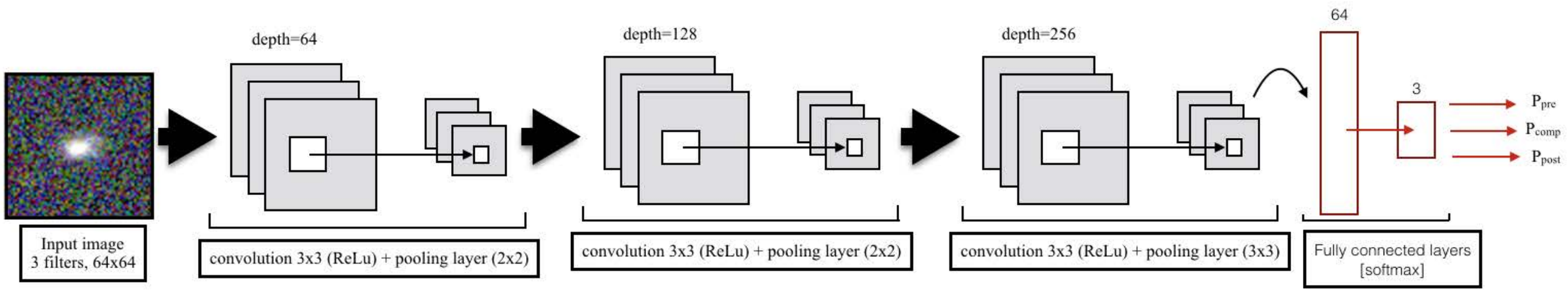
Examples of CANDELized simulated galaxy images

pre-compaction

compaction

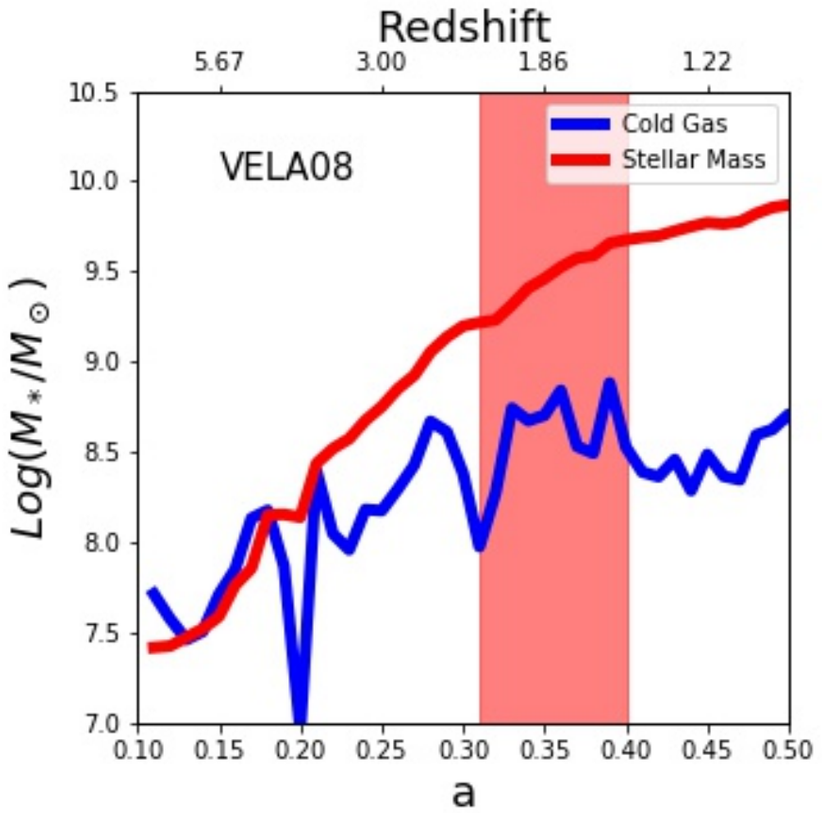
post-compaction

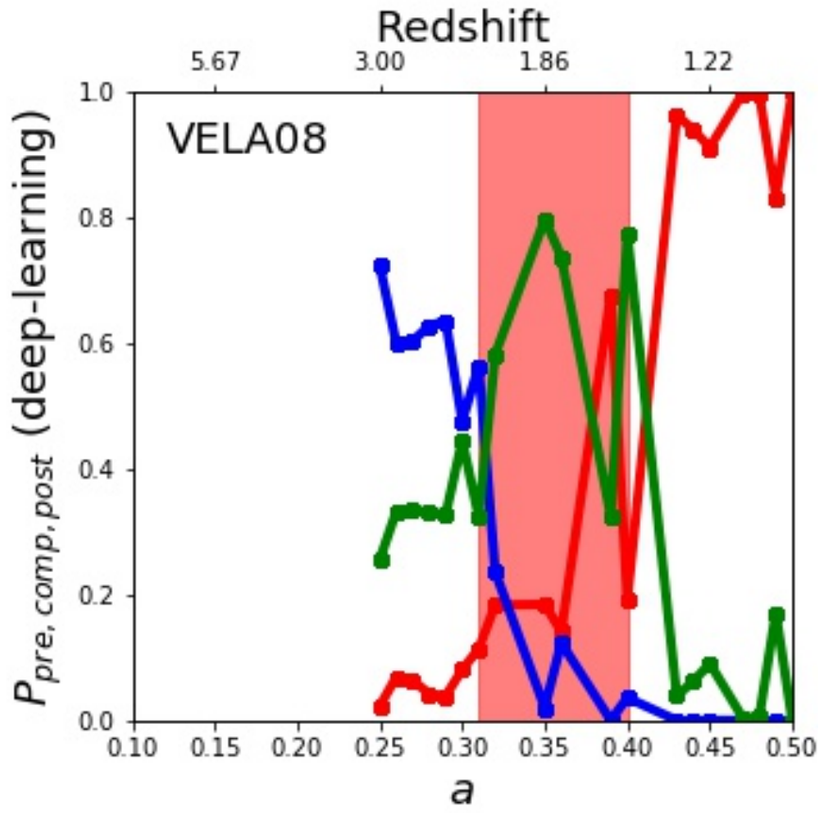




Architecture of the deep network used for classification in this work. The network is a standard and simple CNN configuration made of 3 convolutional layers followed by pooling and dropout.

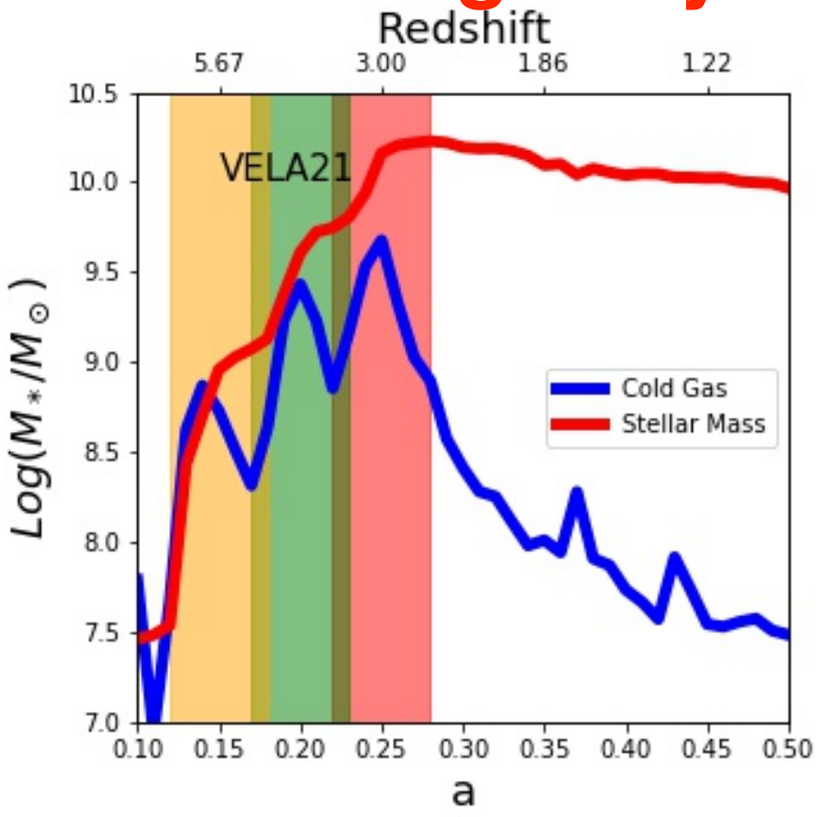
Simulated galaxy with single compaction event

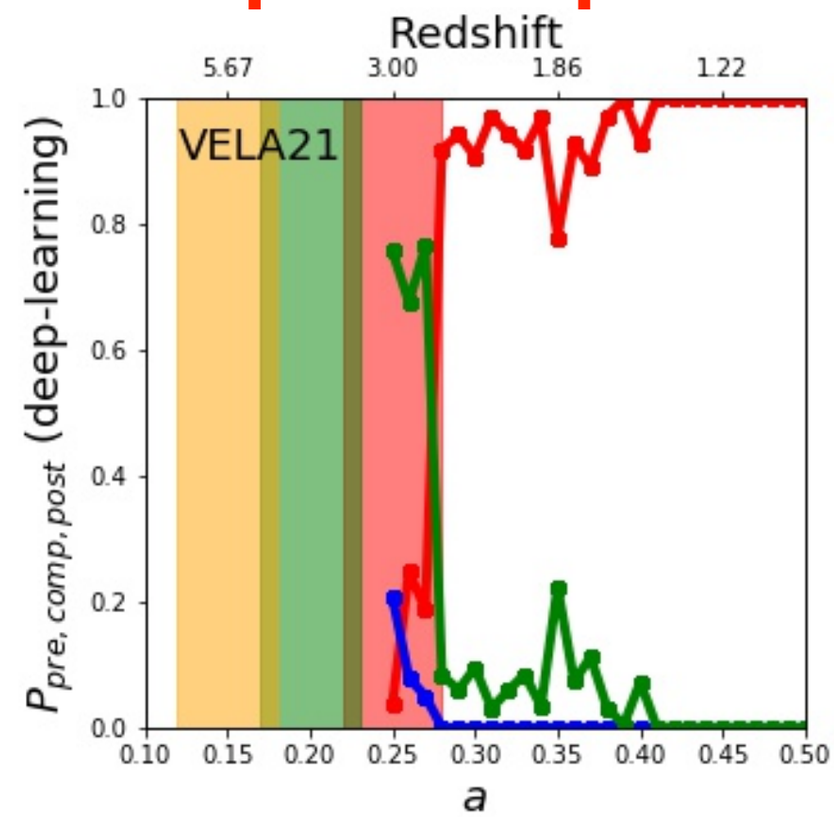




pre-compaction
compaction
post-compaction

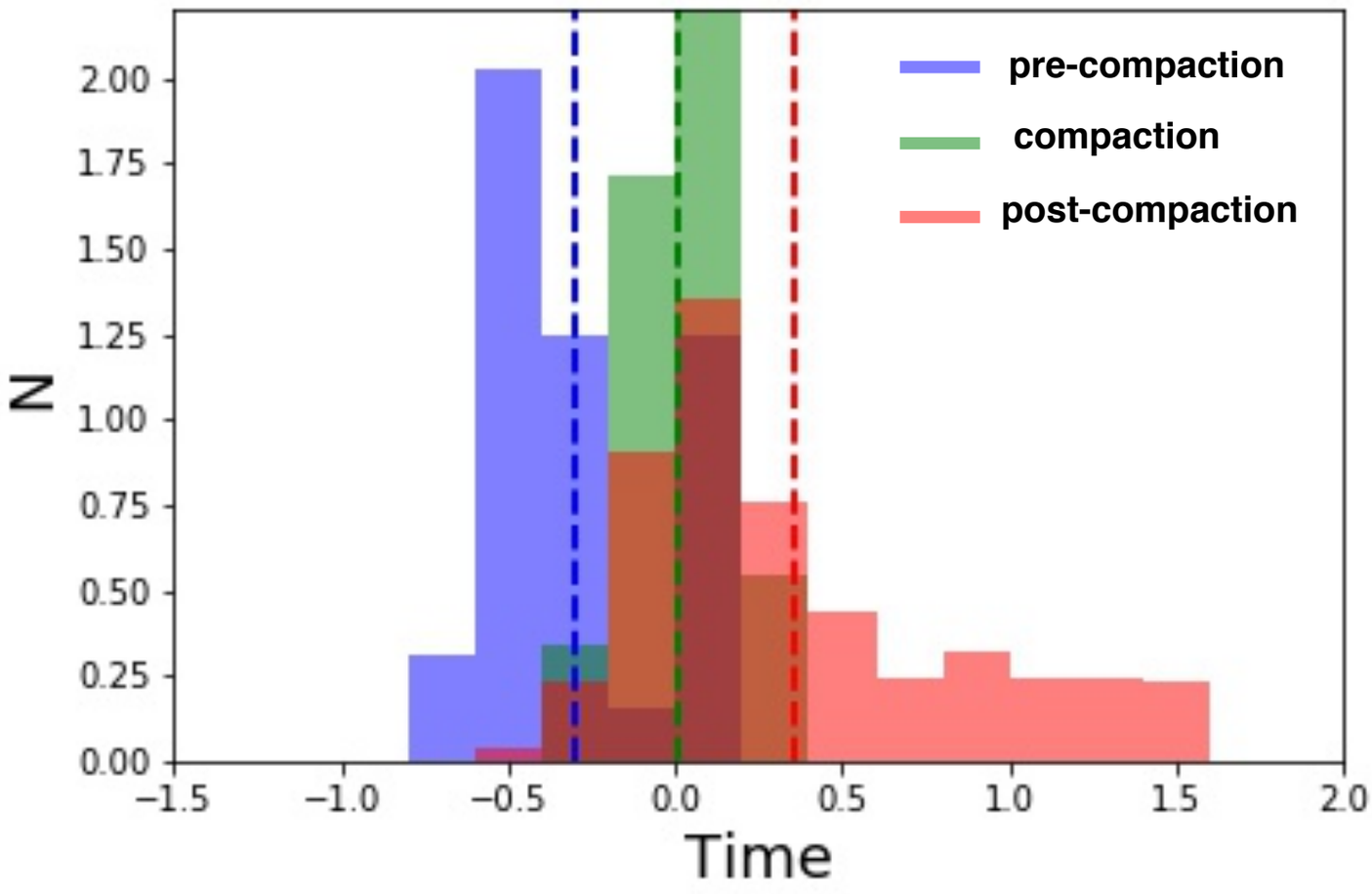
Simulated galaxy with multiple compaction events





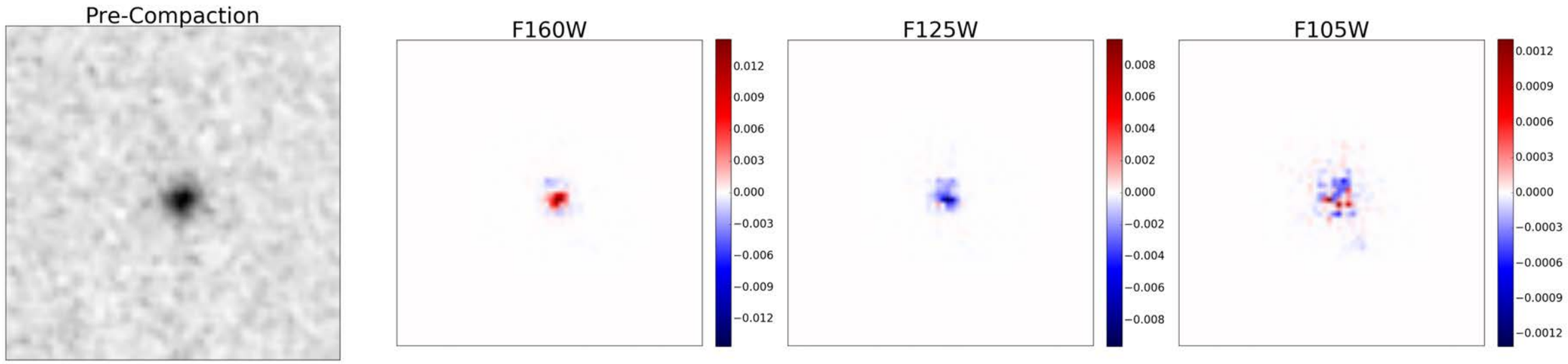
pre-compaction
compaction
post-compaction

Testing the Trained Deep Learning Code

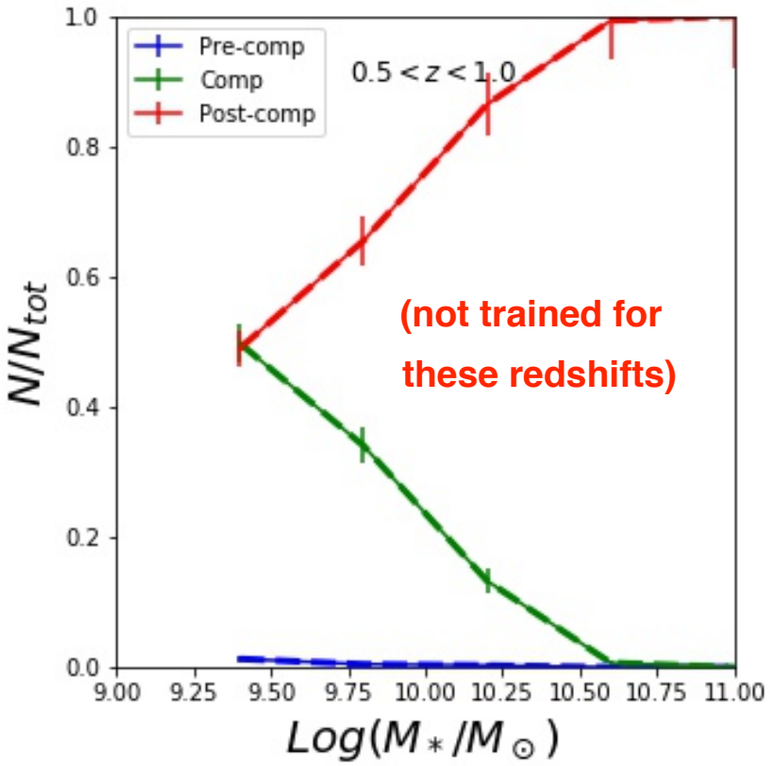


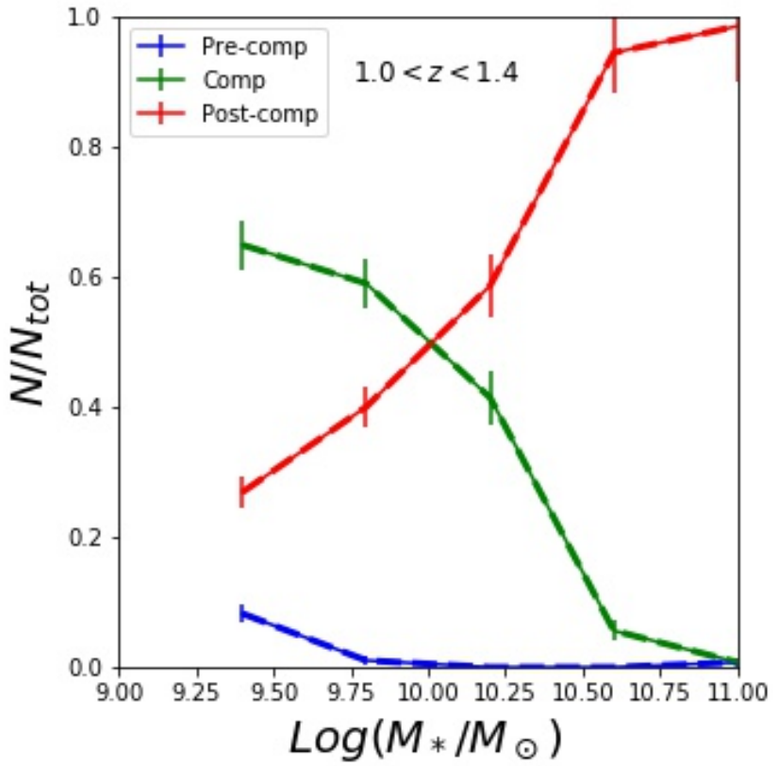
Observability of the compaction event with the calibrated classifier. The histograms show the distributions of time (relative to the Hubble time at the time of compaction). The dashed vertical lines show the average values for each class with the same color code. Despite some overlap, the classifier is able to establish temporal constraints on the different phases. Integrated gradient method shows that the classifier is using relevant pixels, not noise.

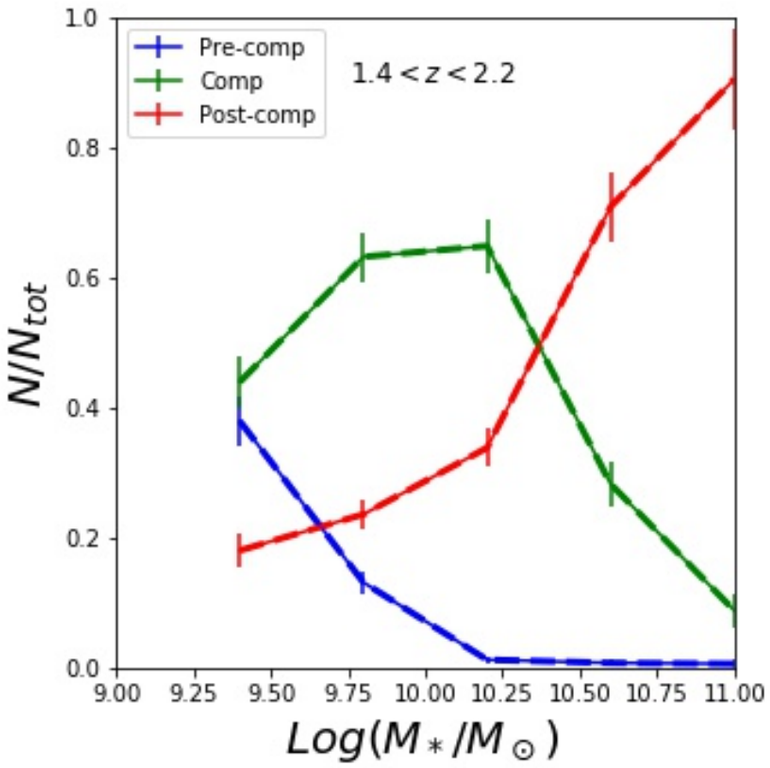
Integrated gradients output of the model. The left column is the original image and the other columns show the integrated gradients for the different wavelength filters. The network automatically detects the pixels belonging to the galaxy and used all of them to make the decisions.

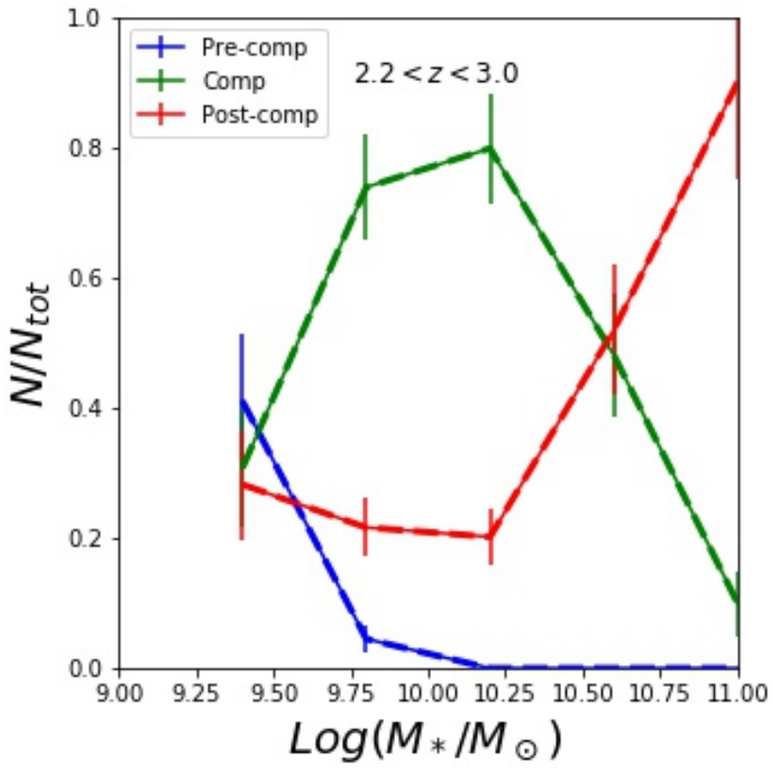


Applying the Trained Deep Learning Code to CANDELS Galaxies





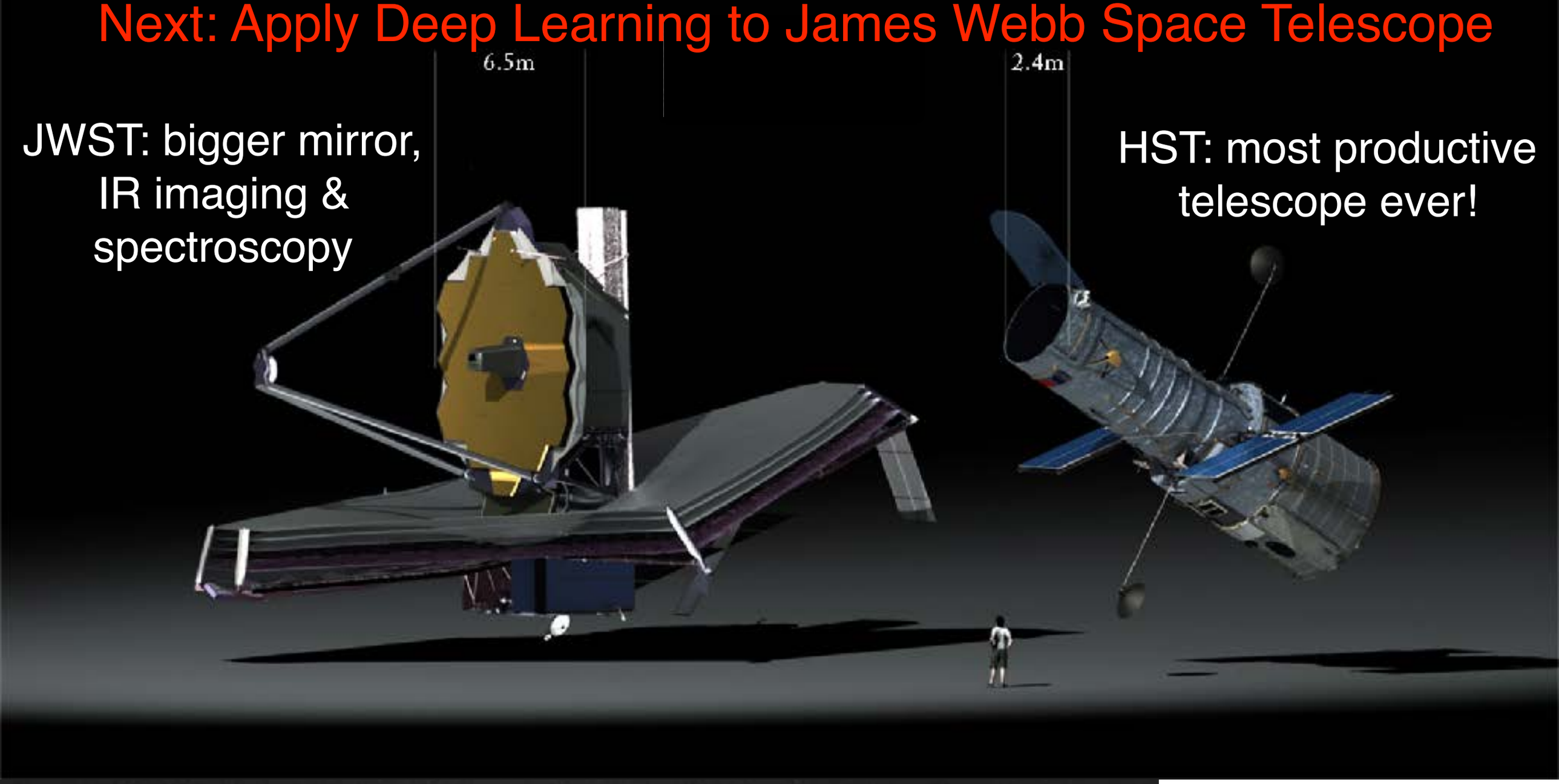




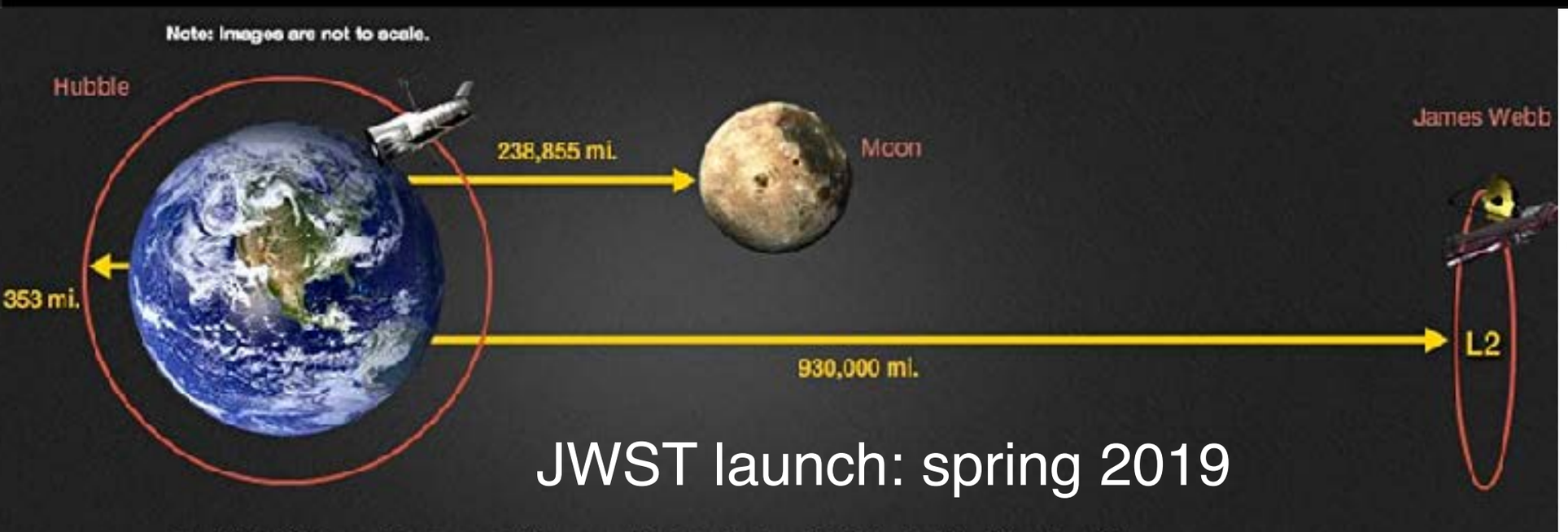
Stellar mass distributions of HST CANDELS galaxies in pre-compaction, compaction, and post-compaction phases in different redshift bins. The DL code correctly shows the temporal evolution. Galaxies in the compaction phase typically peak at stellar masses $10^{9.5–10}$ M_{sun} at all redshifts, as in the VELA simulations.

Next: Apply Deep Learning to James Webb Space Telescope

JWST: bigger mirror, IR imaging & spectroscopy



HST: most productive telescope ever!



JWST launch: spring 2019

MIRI: Mid Infrared Instrument, 5-28 μm
NIRCam: Near Infrared Camera, 0.6 - 5 μm
NIRISS: Near-InfraRed Imager/SSpectrograph
NIRSpec: Near Infrared Spectrograph, 0.6-5 μm

z = 1.5

stars

10 kpc

Computer vision and deep learning applied to simulations and imaging of galaxies and the evolving universe

Joel Primack

University of California, Santa Cruz

Large-scale simulations track the evolution of structure in the universe of dark energy and cold dark matter on scales of billions of light years

Cosmological zoom-in simulations model how individual galaxies evolve through the interaction of atomic matter, dark matter, and dark energy

Our VELA galaxy simulations agree with HST CANDELS observations that most galaxies start prolate, becoming spheroids or disks after compaction events

A deep learning code was trained with VELA galaxy images plus metadata describing whether they are pre-compaction, compaction, or post-compaction

The trained deep learning code was able to identify the compaction and post-compaction phases in CANDELized images

The trained deep learning code was also able to identify these phases in real HST CANDELS observations, finding that compaction occurred for stellar mass $10^{9.5\text{-}10}$ M_{sun}, as in the simulations — and James Webb Space Telescope will allow us to do even better