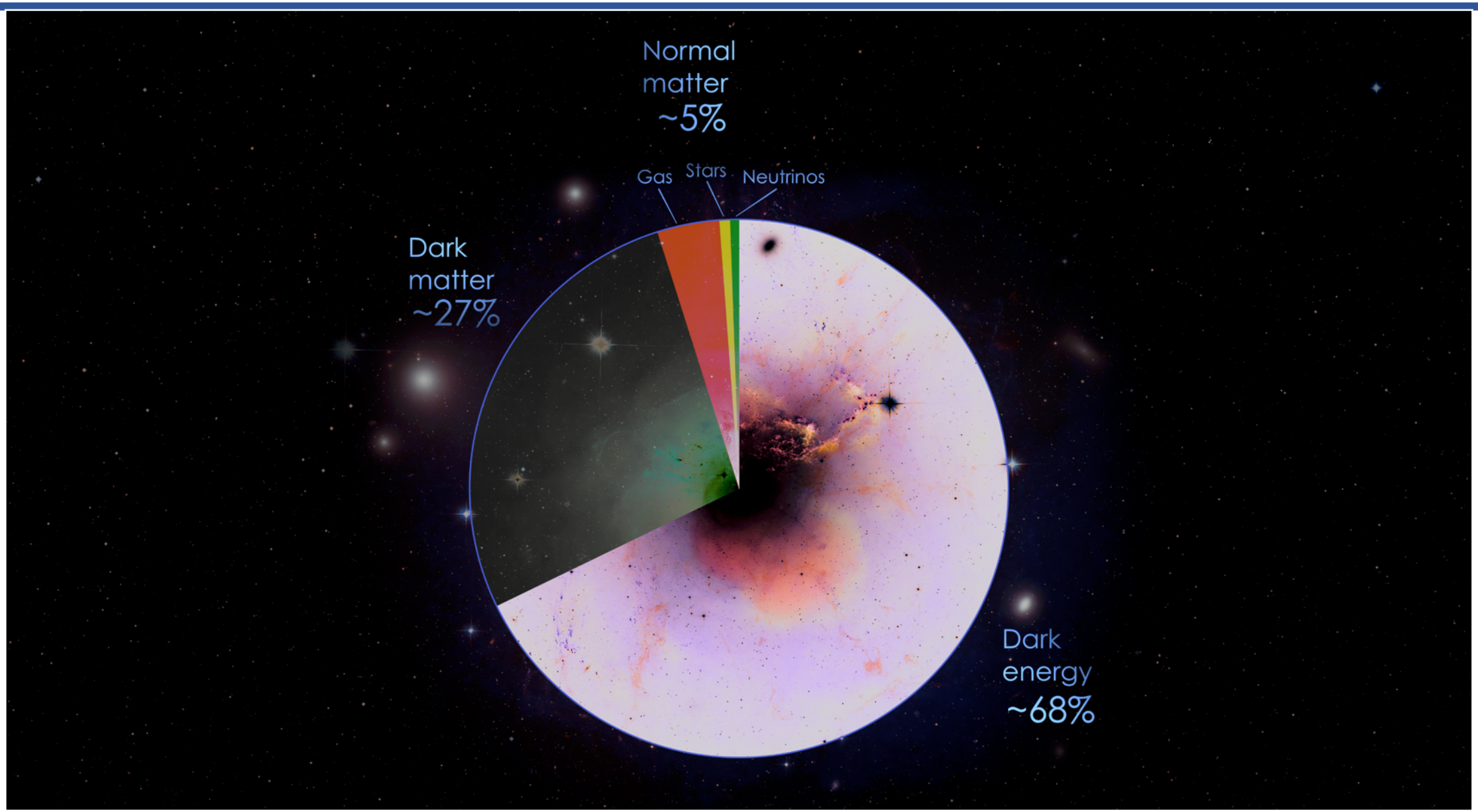
Enabling Dark Energy Research with Data Science

Alina Kiessling

NASA Jet Propulsion Laboratory

California Institute of Technology

The Universe as a Pie Chart



Dark Energy

Dark Energy affects the:

Expansion history of the Universe

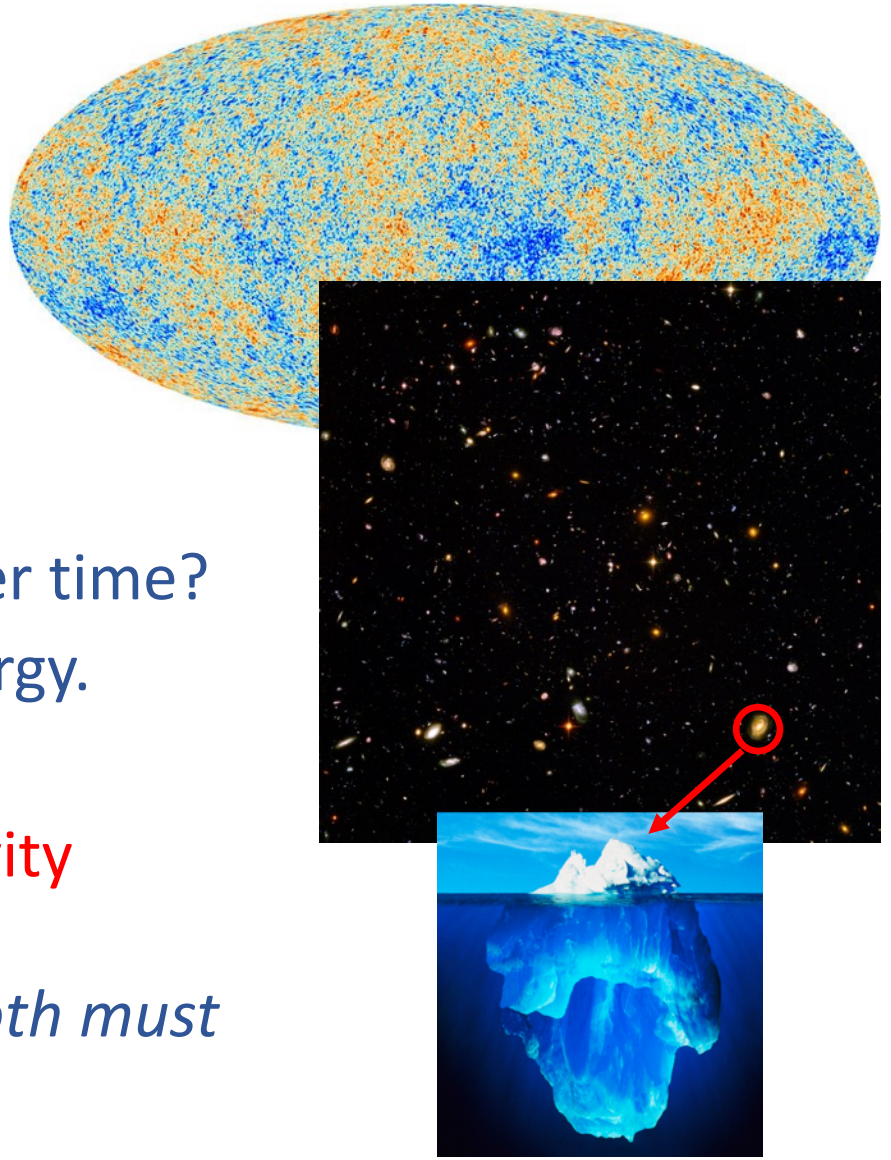
- How fast did the Universe expand?
- Also known as the geometry of the Universe.

Growth of structures

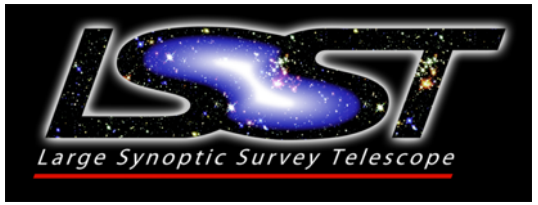
- How do dark matter structures evolve and grow over time?
- Attractive gravity competes with repulsive dark energy.

If Einstein's General Relativity is wrong, modified gravity theories could explain the accelerating expansion.

This would change the effects above differently, so *both must be measured*!



Example Upcoming Dark Energy Missions

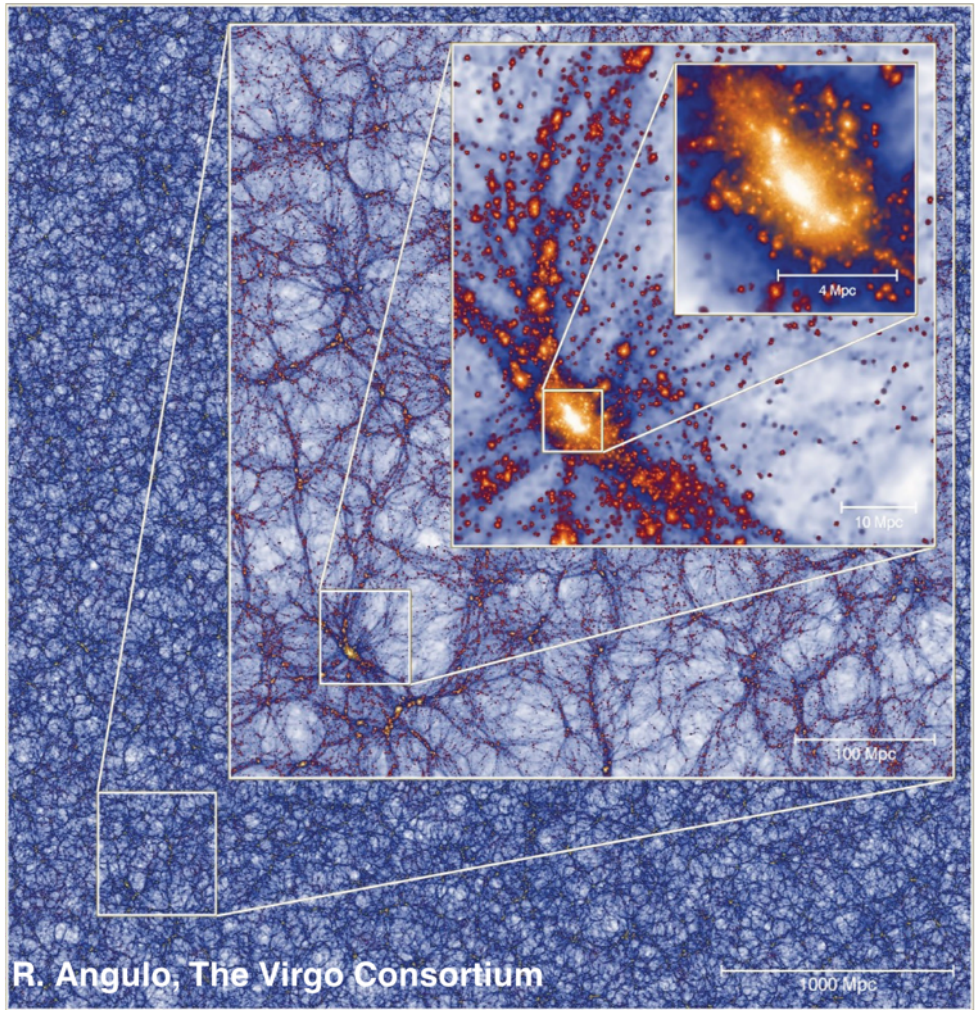


Proposed lifetime	2022 - 2032	2022 - 2028	2025 - 2031
Mirror size (m)	6.5 (effective diameter)	1.2	2.4
Survey size (sq deg)	20,000	15,000	2,227
Median z (WL)	0.9	0.9	1.2
Depth (AB mag)	~27.5	~24.5	~27
FoV (sq deg)	9.6	0.5 (Vis) 0.5 (NIR)	0.28
Filters	u-g-r-i-z-y	Y-J-H-Vis	Y-J-H-F184
Cosmological probes	WL, LSS, SN, SL, Cl	WL, LSS, SL, Cl,	WL, LSS, SN, SL, Cl

Dark Energy and Simulations

- Simulations are essential to the success of upcoming cosmological surveys.
  - Pre-launch, simulations provide a data set with known parameters.
  - Post-launch, simulations are integral to the analysis and interpretation of the observational data.

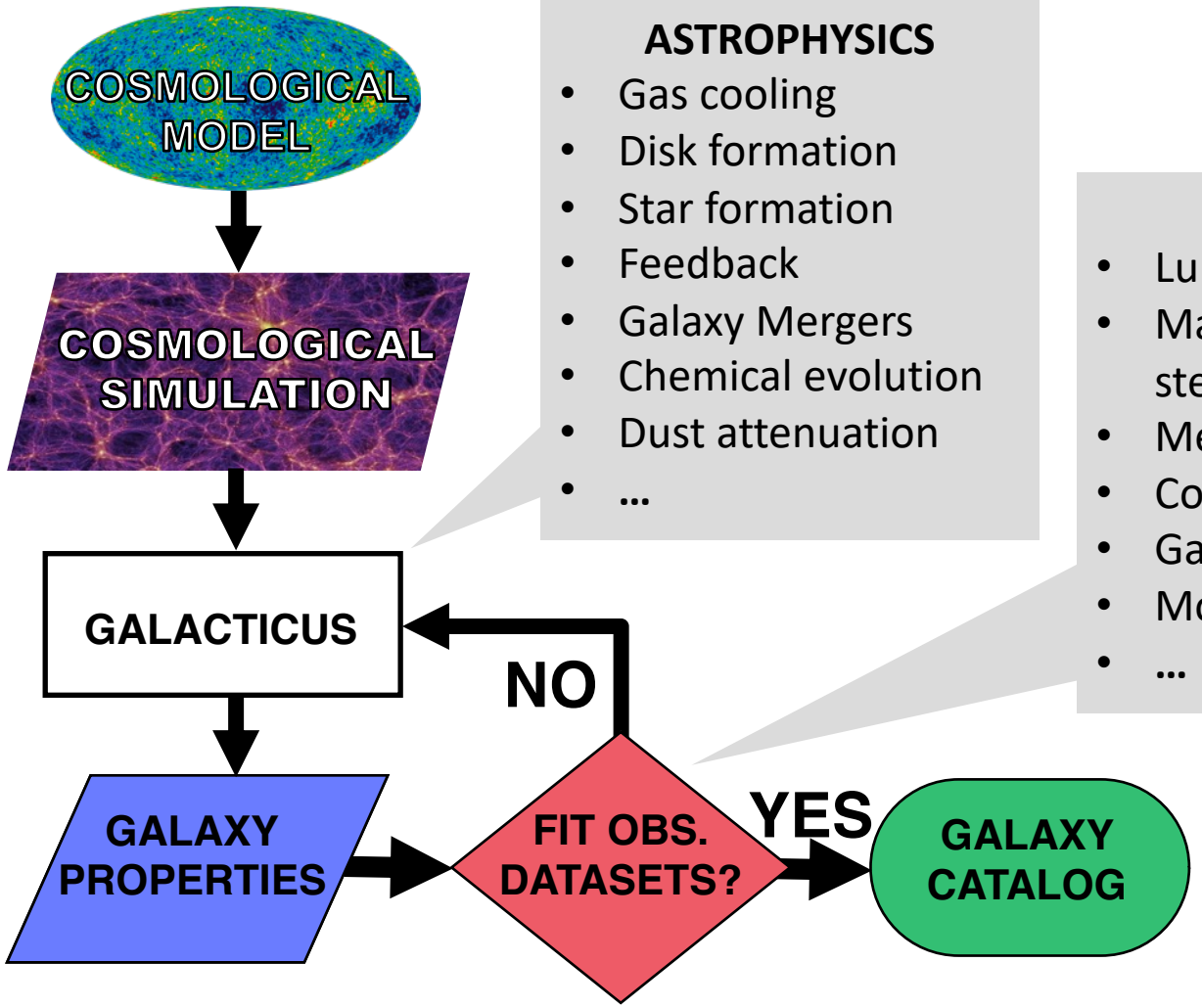
- Simulations are used to:
  - Develop and test analysis techniques.
  - Characterize effects caused by the telescope, instruments, and astrophysical systematics.
  - Generate covariance matrices to undertake parameter forecasting and data analysis.



Challenge: Adding Galaxies to Simulations

JPL Data Science Pilot Project: Enabling Scientific Analysis with Data Science

PI: Jeff Jewell; **Co-Is**: Andrew Benson (Carnegie), Luca Cinquini, Ashish Mahabal, Alex Merson, Alina Kiessling



Credit: Andrew Benson (Carnegie Observatories) bitbucket.org/galacticusdev/galacticus

- **Simulating galaxy formation and evolution is a complex process.**
- "Semi-analytical model": solves coupled sets of differential equations governing astrophysical processes.
- OpenMP/MPI parallelized.
- Open source.

Containerization & Deployment

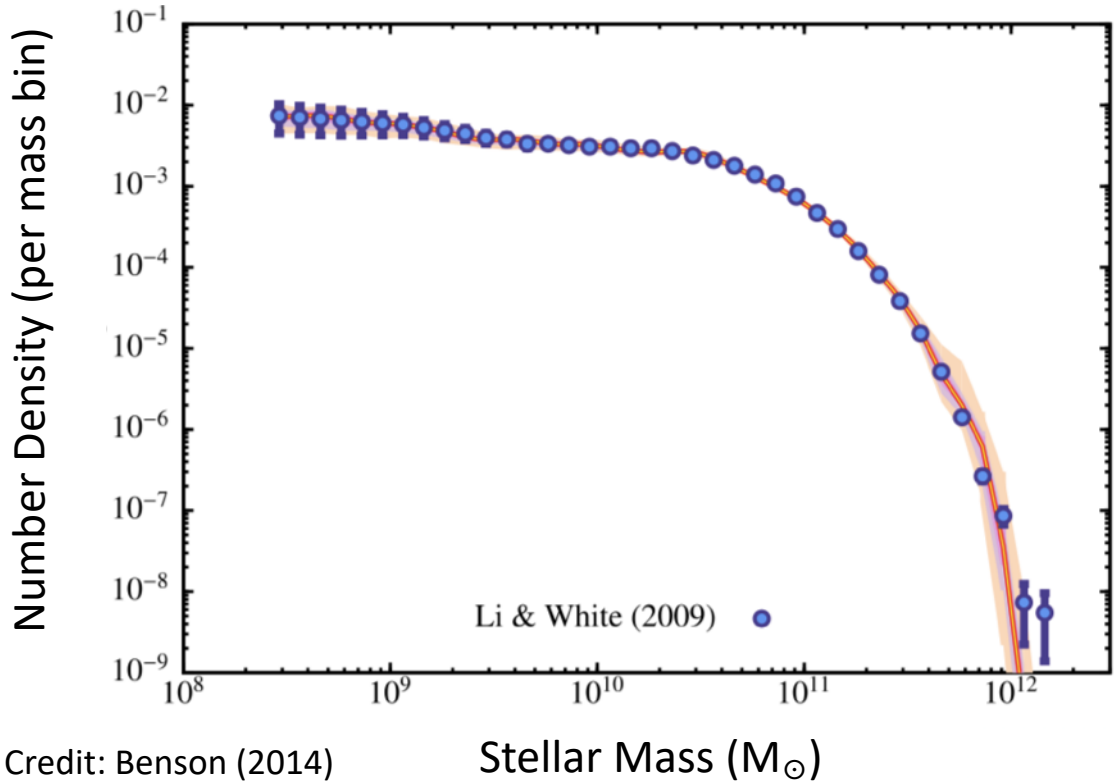
- Containerized version has been developed (needs parallelizing).
- To be deployed with scalable containerization environment (e.g. Docker Swarm, Kubernetes) on cloud service (e.g. AWS, Google Cloud Platform).

Challenge: Calibrating the Galaxy Formation and Evolution Model

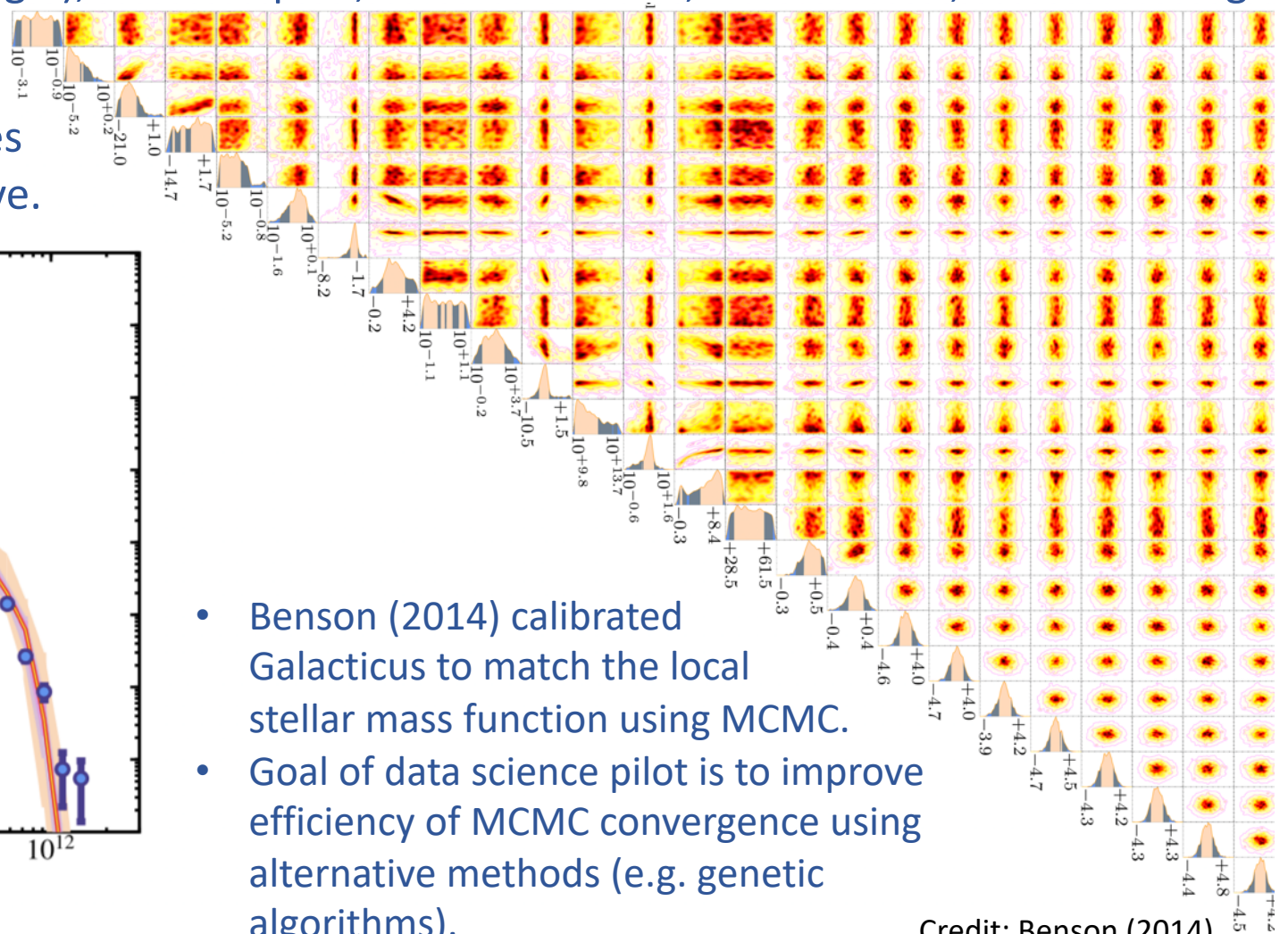
JPL Data Science Pilot Project: Enabling Scientific Analysis with Data Science

PI: Jeff Jewell; **Co-Is**: Andrew Benson (Carnegie), Luca Cinquini, Ashish Mahabal, Alex Merson, Alina Kiessling

- The model has 10s - 100s parameters!
- The model must be run 100s-1000s times – calibration is computationally expensive.



Credit: Benson (2014)



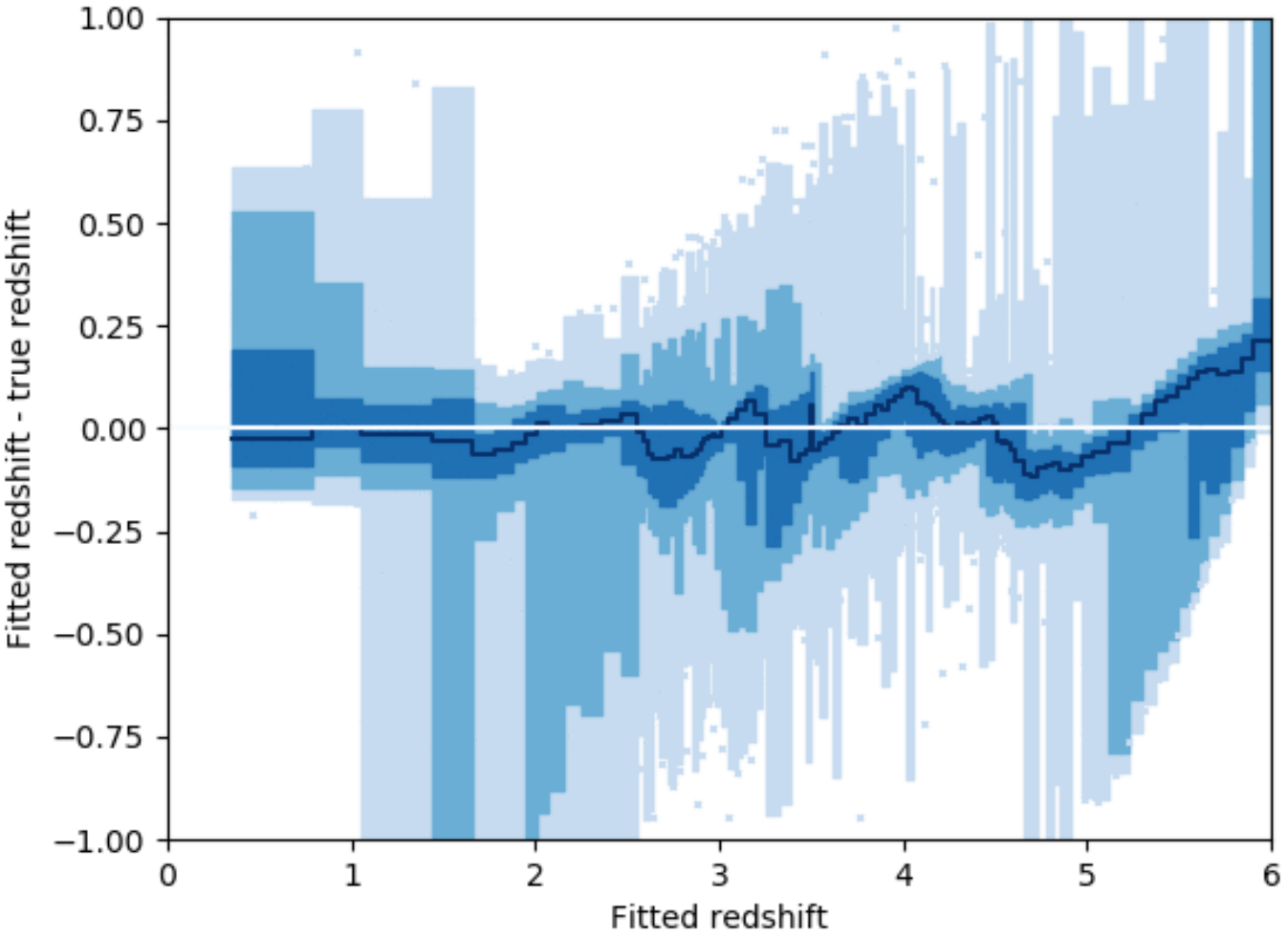
- Benson (2014) calibrated Galacticus to match the local stellar mass function using MCMC.
- Goal of data science pilot is to improve efficiency of MCMC convergence using alternative methods (e.g. genetic algorithms).

Credit: Benson (2014)

Challenge: Determining Galaxy Redshifts

- Cosmological analyses rely on accurate knowledge of the distance to each observed galaxy.
  - Upcoming surveys like WFIRST, Euclid, and LSST will observe billions of galaxies in ~8 broadband filters spanning optical to near-IR wavelengths.
  - Methods to determine photometric redshifts are currently in development to improve their accuracy.

- Machine Learning is being explored as a method to improve the accuracy of photometric redshift determination.



Credit: Melanie Simet (JPL/UC Riverside)

Challenge: Determining Galaxy Redshifts

- Dimensionality reduction of galaxy color space using the Self-Organizing Map (SOM) has also made promising progress toward more accurate photometric redshift measurements.
- What is a SOM?

Credit: Shoubaneh Hemmati (JPL/Caltech)

Challenge: Determining Galaxy Redshifts

- Similar in one dimension, e.g. Subject

Credit: Shoubaneh Hemmati (JPL/Caltech)

Challenge: Determining Galaxy Redshifts

- Similar in another dimension, e.g. Style

Credit: Shoubaneh Hemmati (JPL/Caltech)

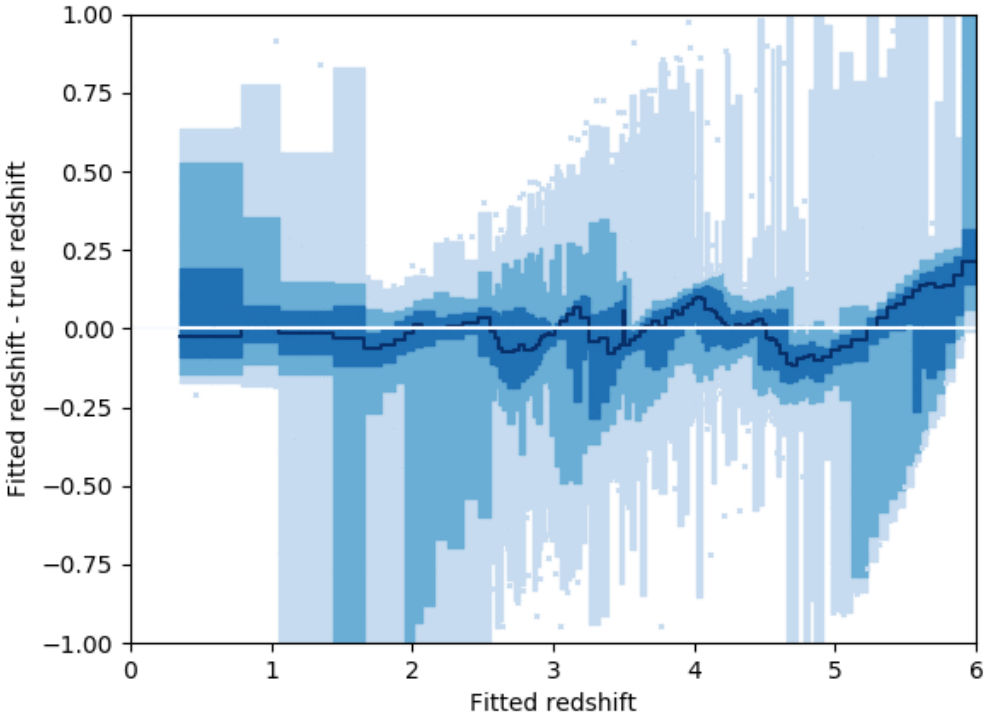
Challenge: Determining Galaxy Redshifts

- The SOM represents a high-dimensional data space in a topologically ordered way. Objects in similar parts of the high-dimensional space are grouped together in the low-dimensional representation.

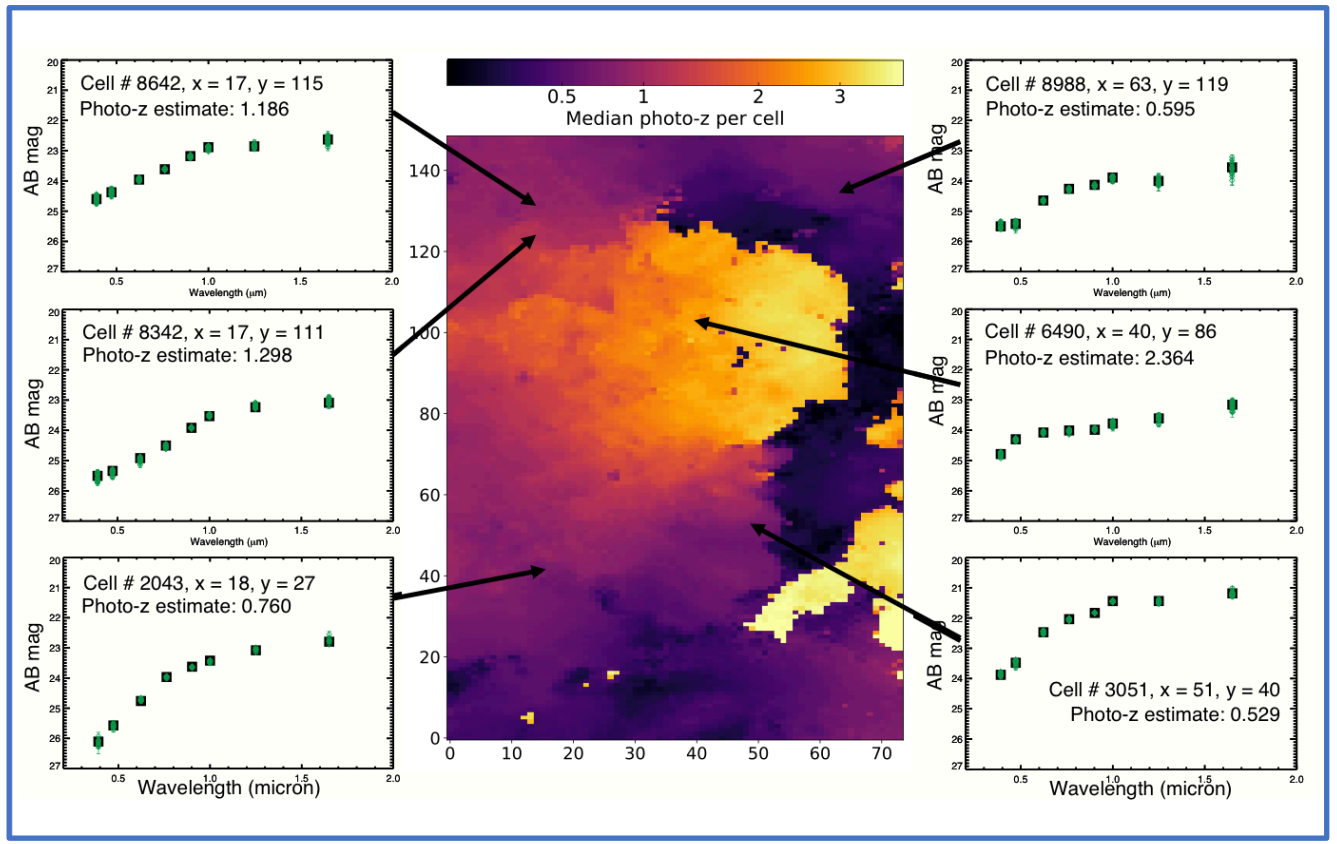
Credit: Shoubaneh Hemmati (JPL/Caltech)

Challenge: Determining Galaxy Redshifts

- There is real value in using more than one analysis method. Scientists can validate results where the methods agree and interesting insights can occur where the methods disagree.



Credit: Melanie Simet (JPL/UC Riverside)



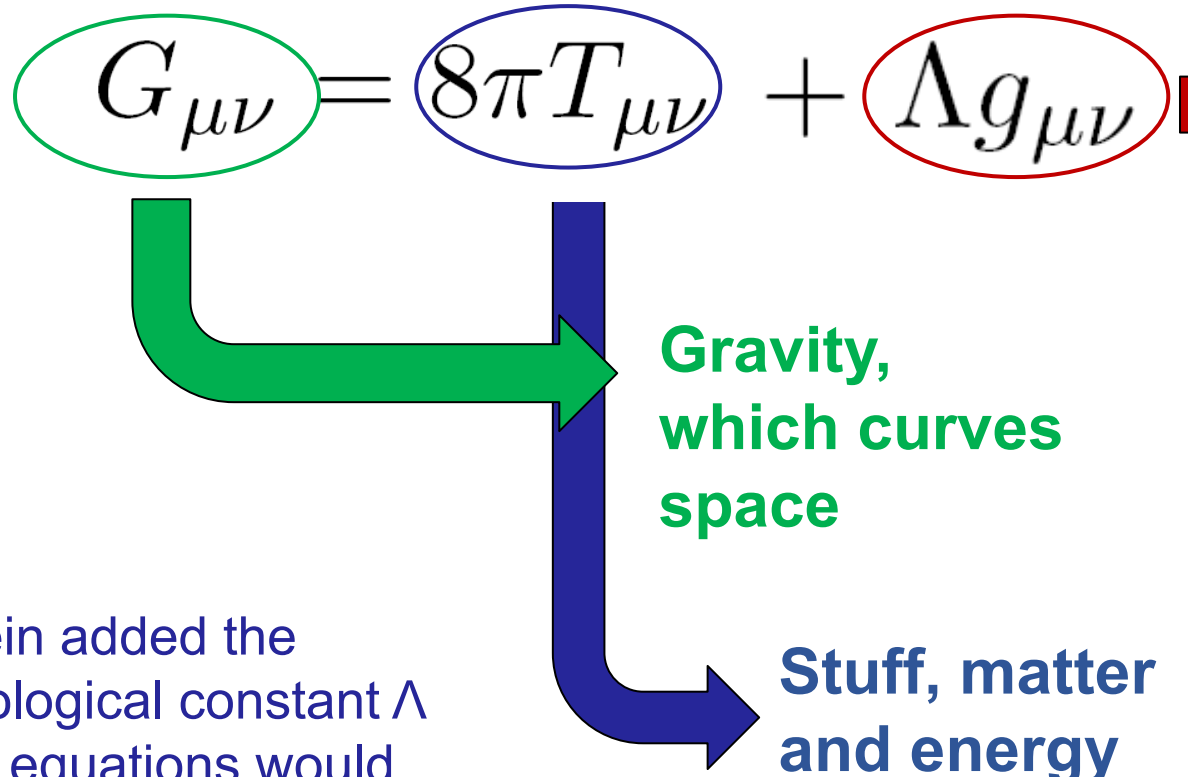
Credit: Dan Masters (JPL/Caltech); Masters et al. 2015

Summary

- Efforts to utilize data science techniques to enable the dark energy research are already underway.
- Dark energy research could benefit greatly from data science techniques:
  - Big-data techniques – upcoming cosmological surveys will generate huge simulated and observational data sets.
  - Machine learning – galaxy and cosmological properties can be challenging to determine accurately.
  - Dimensionality reduction – high-dimensional parameter spaces can be difficult to visualize and interpret.
  - Visualization – ground-breaking results need to be communicated effectively.

Backup slides

Einstein's Field Equation



Einstein added the cosmological constant Λ so his equations would work for a static universe (which was considered aesthetically pleasing)

The Universe is Expanding

Distance

Velocity

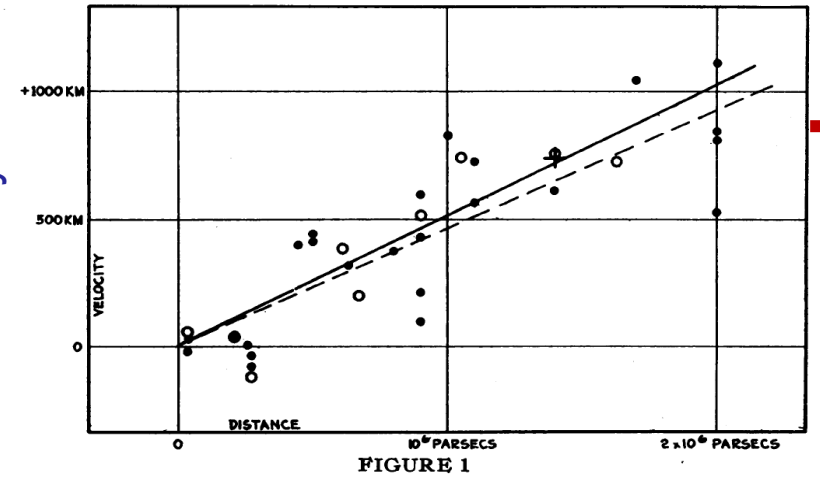


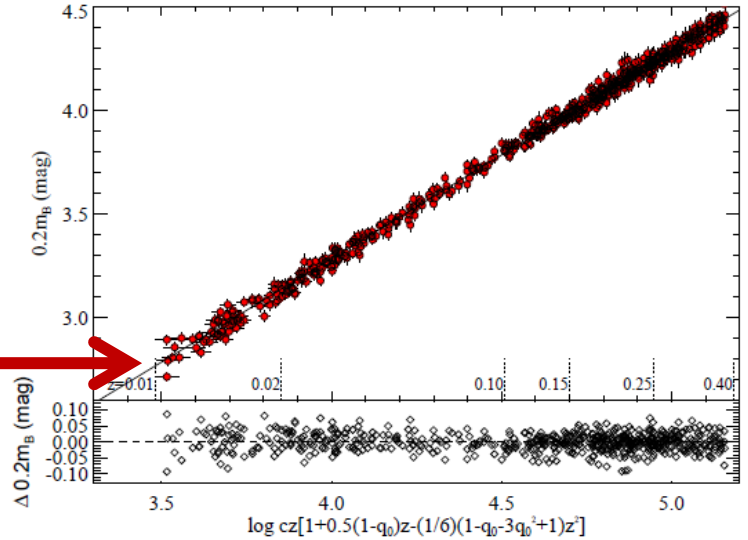
FIGURE 1

Velocity-Distance Relation among Extra-Galactic Nebulae.

Hubble, 1929

Hubble Space Telescope

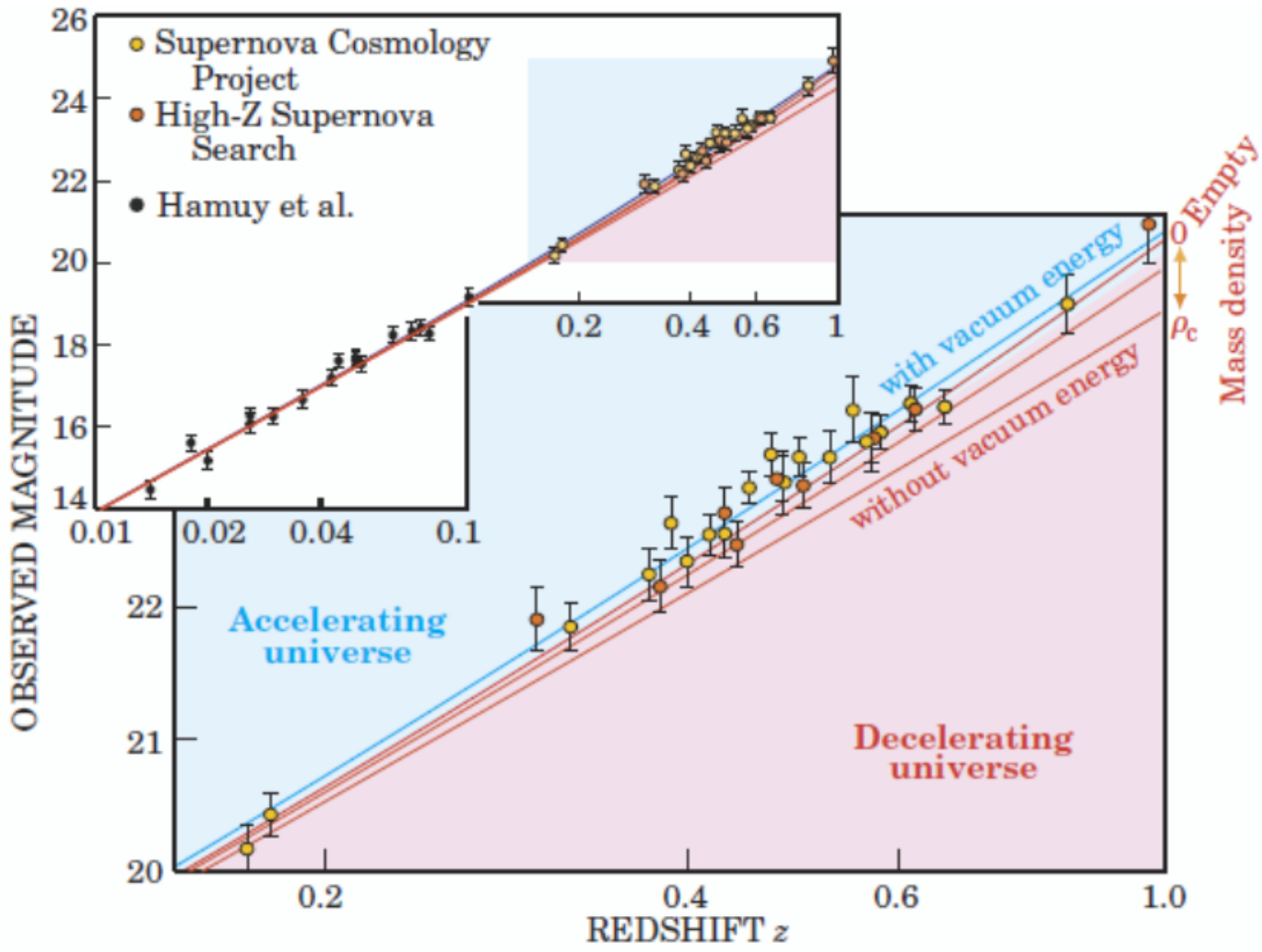
The farther away something is, the faster it is moving away, *in all directions!*



Riess et al, 2016

The Expansion is *Accelerating!*

Dimmer



Brighter

Close → Far Away

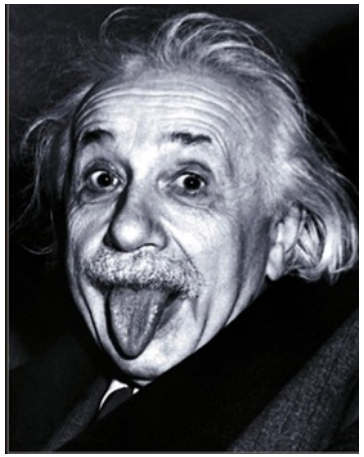
Einstein's Field Equation

$$G_{\mu\nu} = 8\pi T_{\mu\nu} + \Lambda g_{\mu\nu}$$

Gravity, which curves space

Stuff, matter and energy

Cosmological Constant, (maybe not a constant)



In 1998, when the accelerating expansion was discovered, Λ was reintroduced!

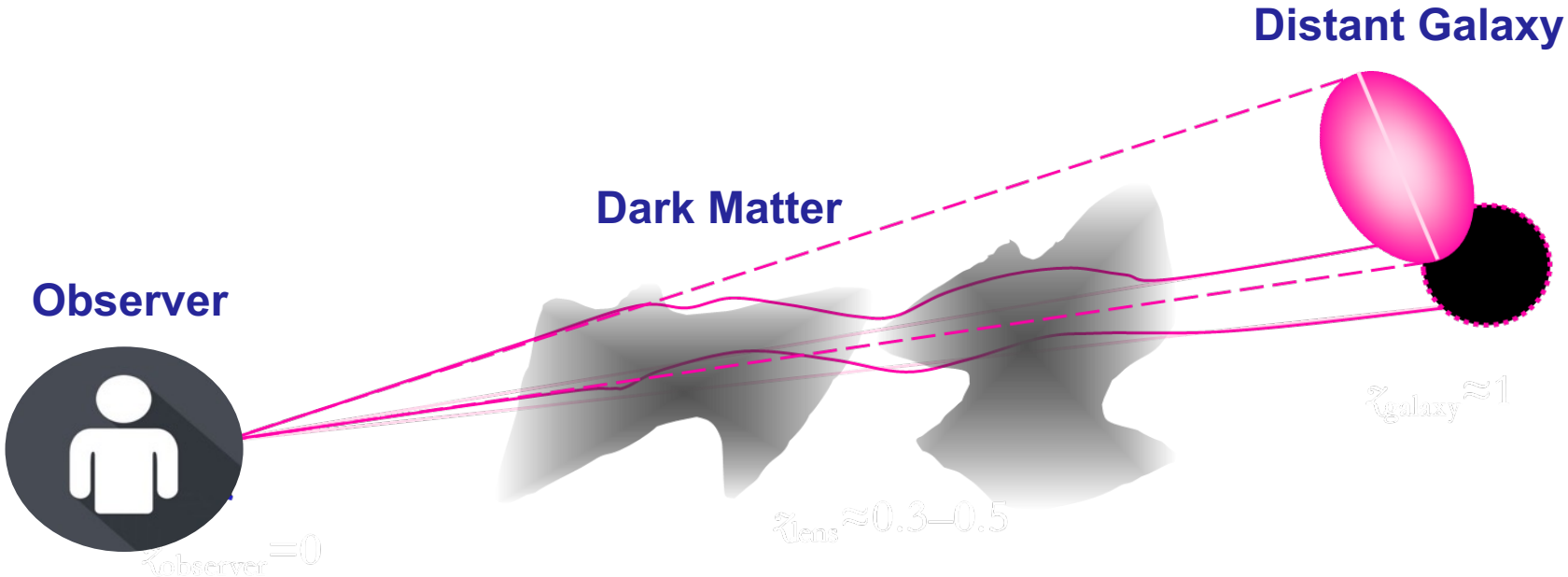
Einstein's 'biggest blunder' went on to win a Nobel Prize!

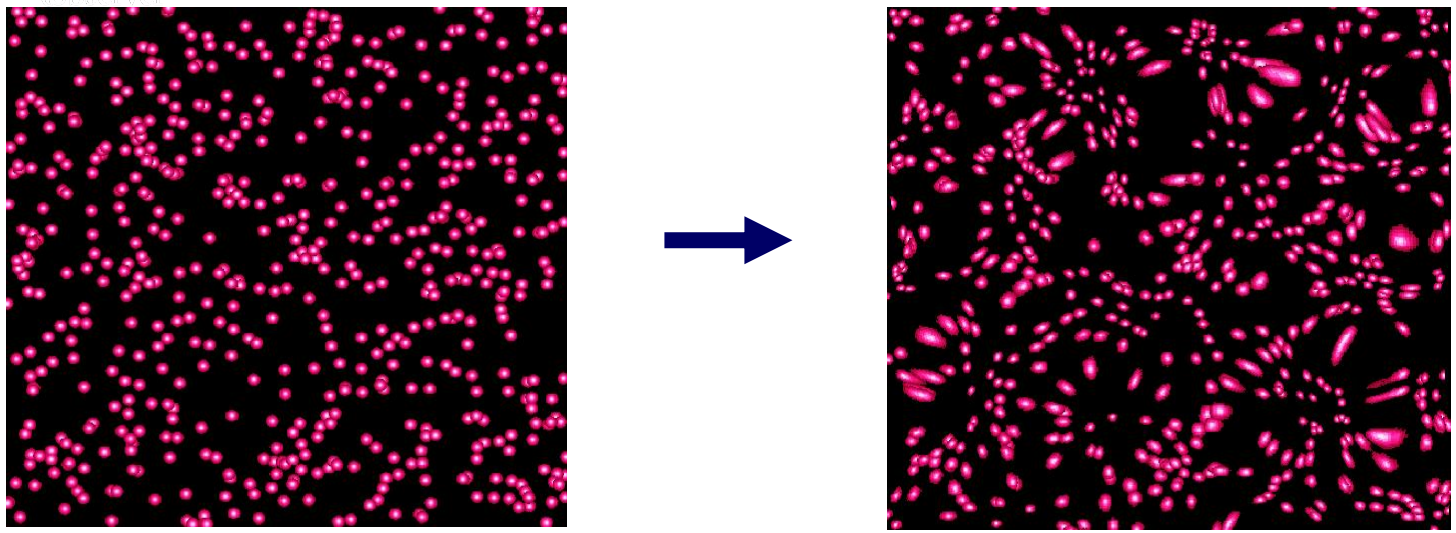
We could also be wrong about the left hand side (gravity). *We will test this* as we attempt to measure dark energy.

How do we investigate the Dark Universe?

- Gravitational Lensing

Probes both the **growth rate of structure** and the **expansion history** of the Universe





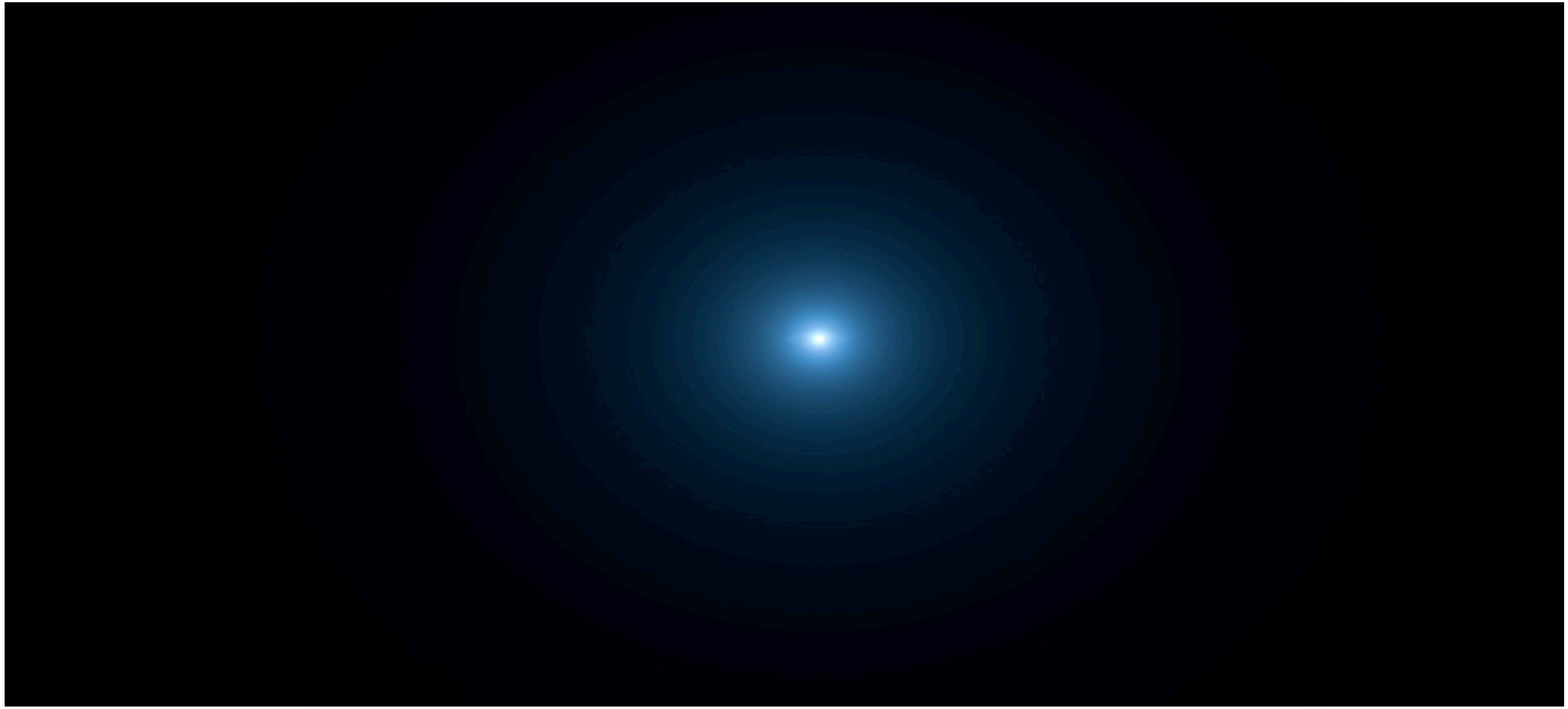
Unlensed

Lensed

How do we investigate the Dark Universe?

- Galaxy Clustering

Probes the **expansion history** of the Universe



Credit: WFIRST