



# Open data with DESI - Dark Energy Spectroscopic Instrument

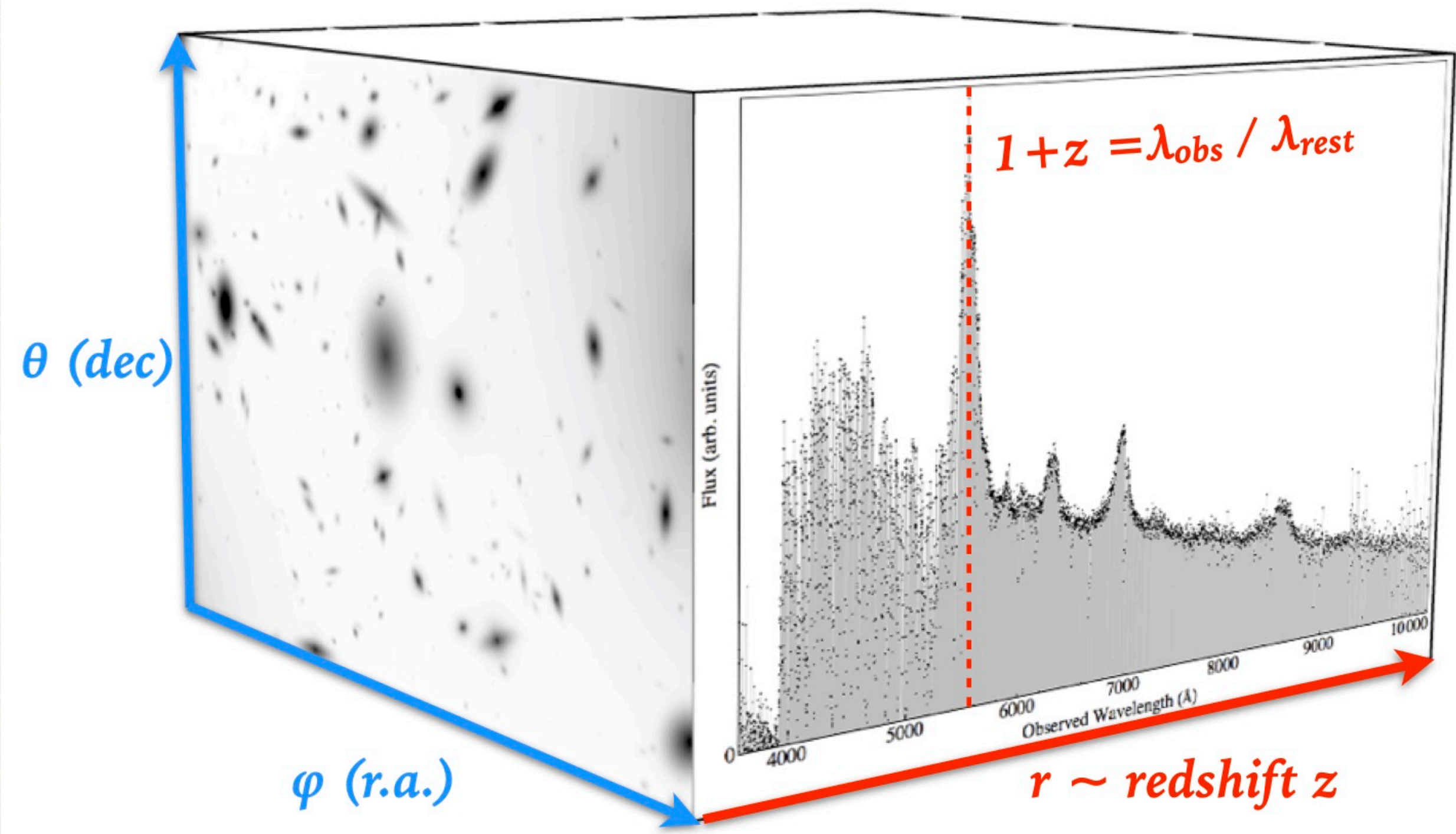
Eva-Maria Mueller

---

*14th May 2026*

# What is our data?

---



The optical sky is a data cube:  $\theta \times \varphi \times r$

# How is the data analysed?

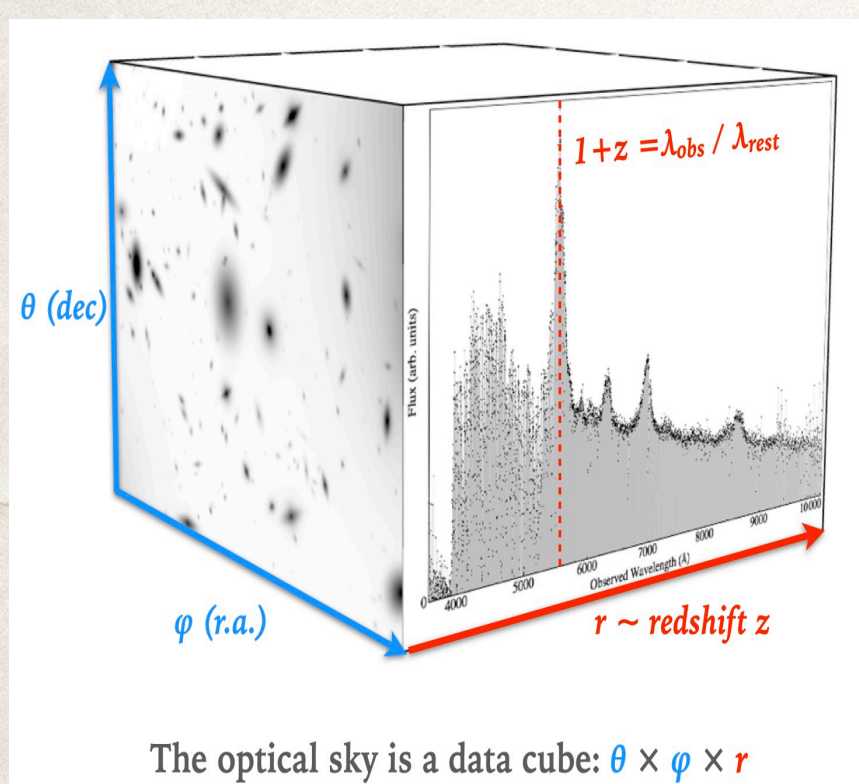
Spectra

Catalogs

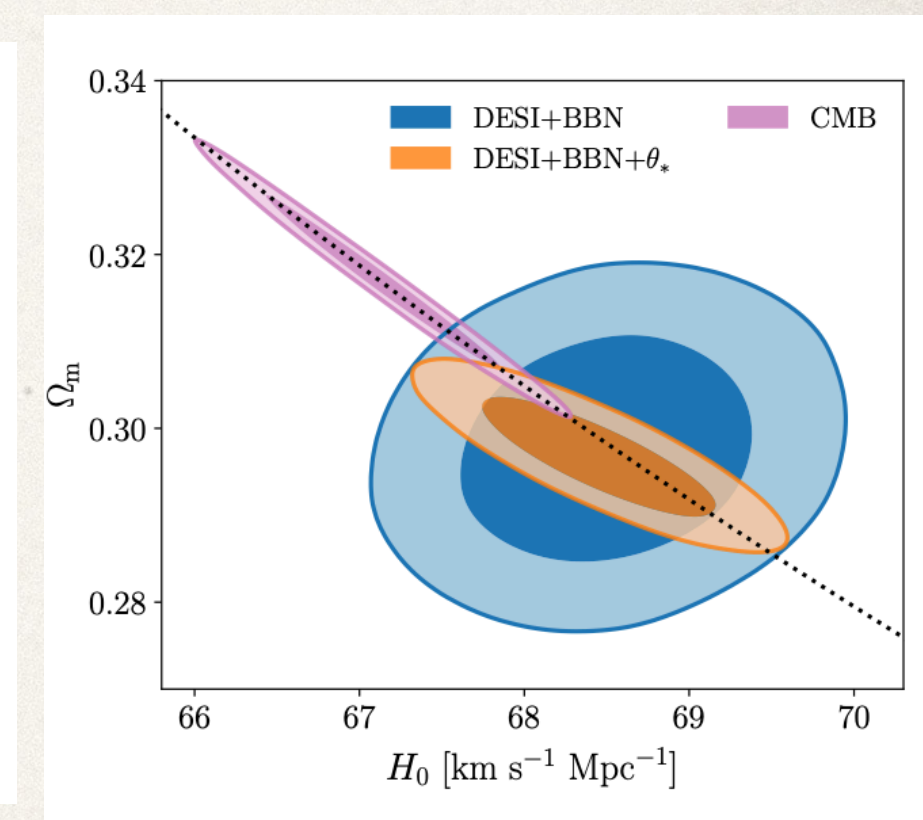
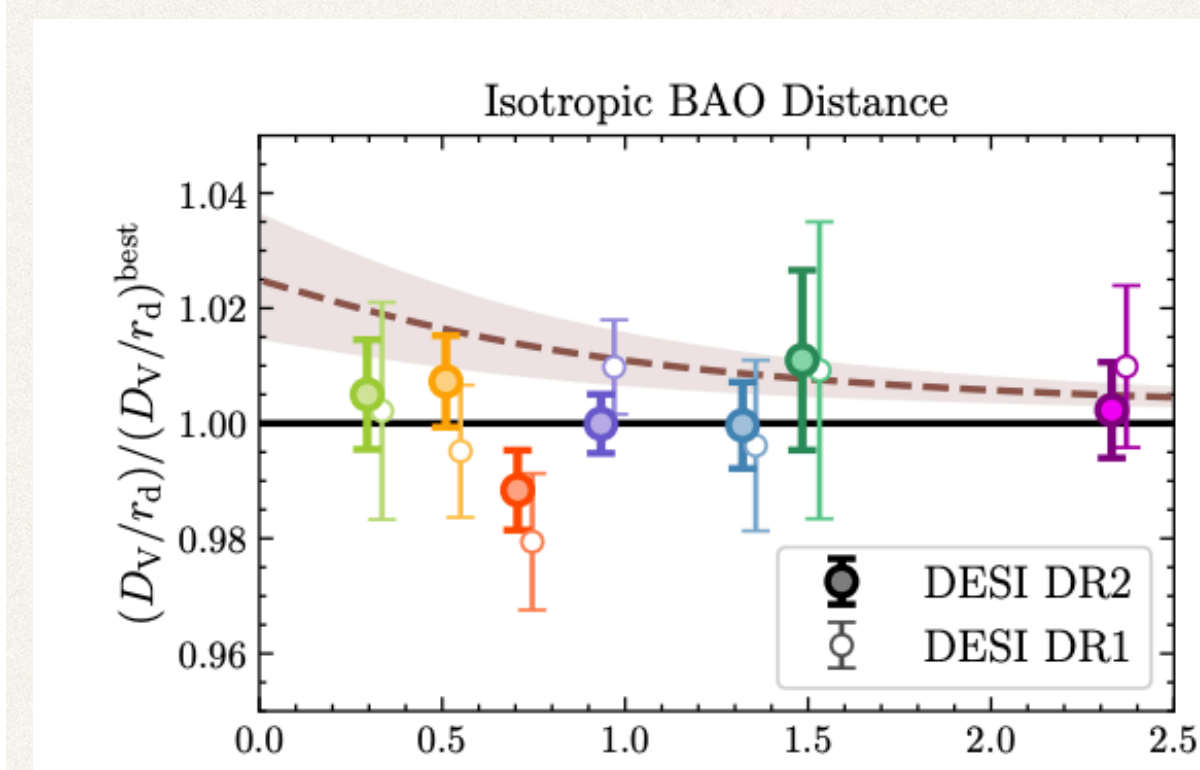
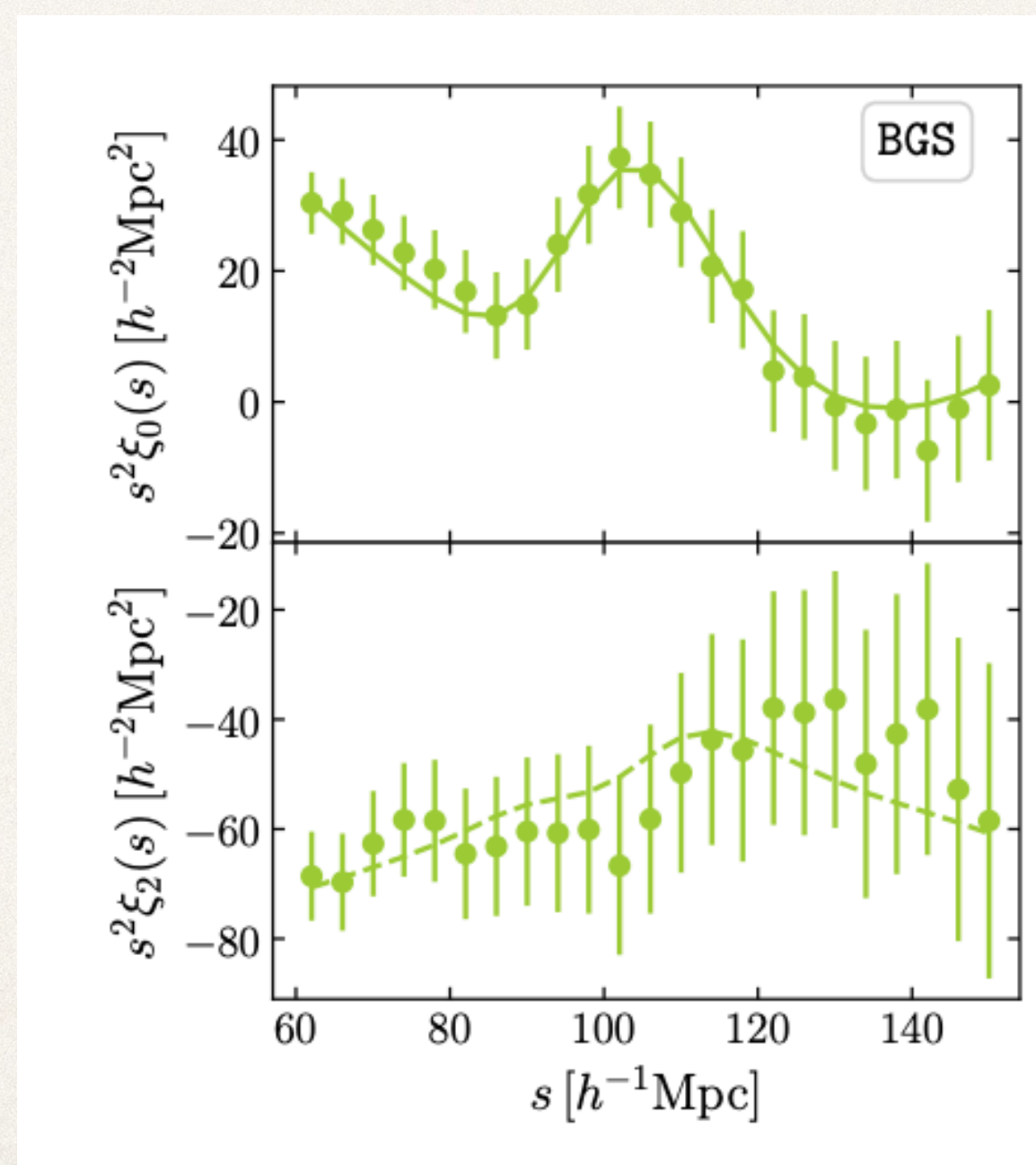
Compressed Statistics

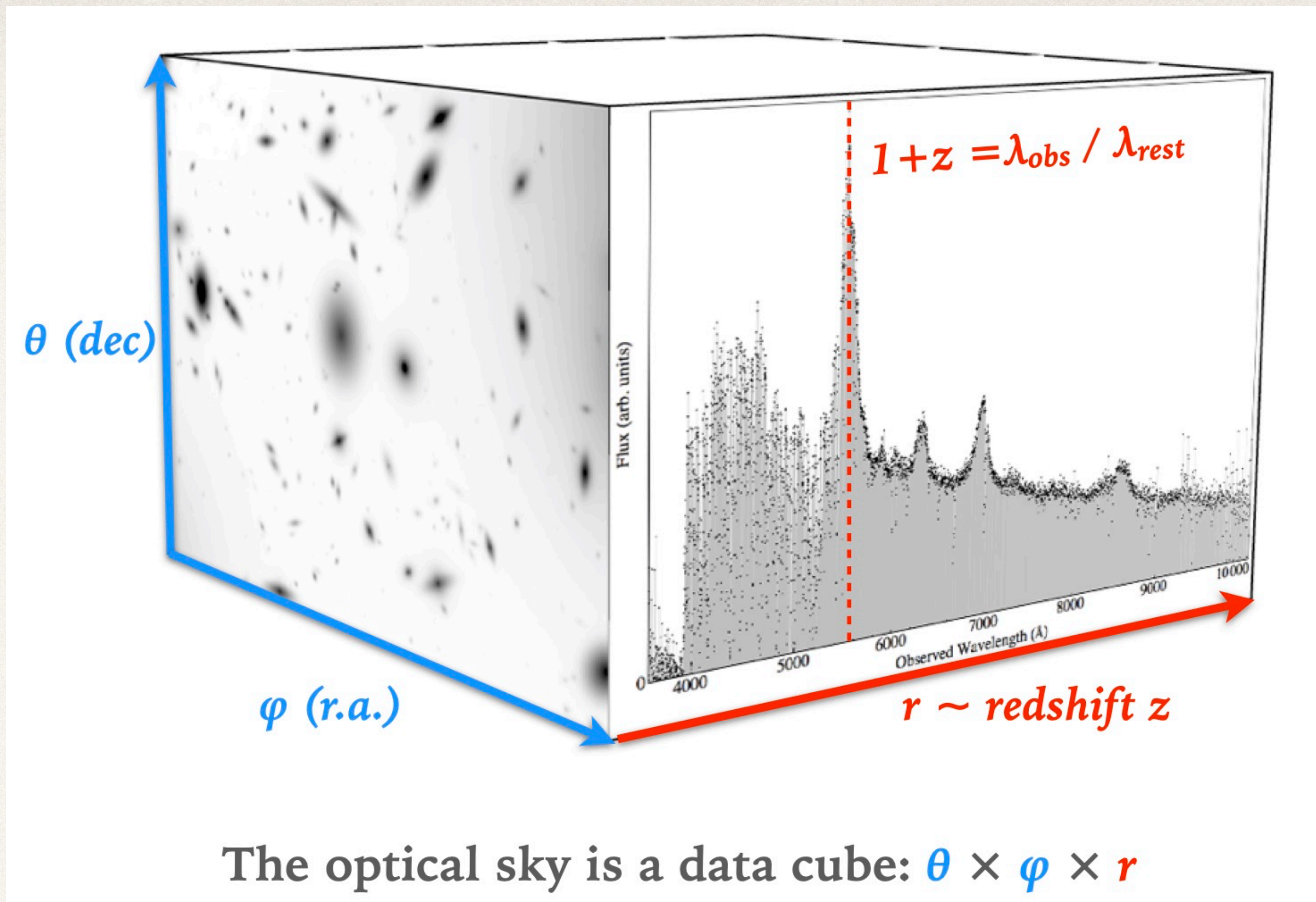
Likelihoods

Cosmological Interpretation



List of:  
Object,  $\theta$ ,  $\phi$ ,  $z$





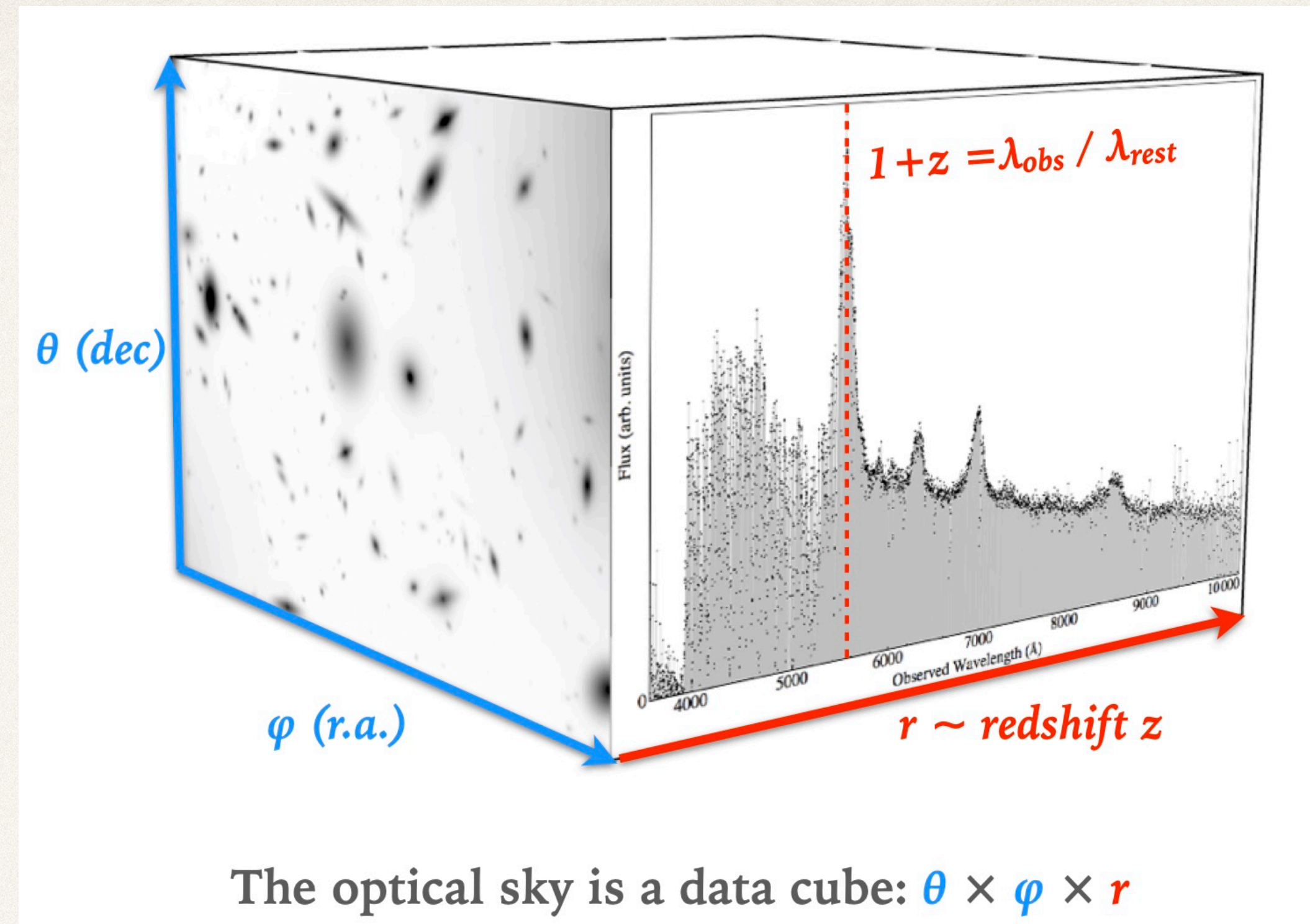
List of:  
Object,  $\theta$ ,  $\phi$ ,  $z$

# Data Products: Raw Spectra + Catalogs

Fit spectra to get redshift aka. distance

# Data Products: Value added Catalogs (VACs)

- ❖ Curated subsets of data with additional analysis beyond the core spectroscopic data processing pipeline outputs
- ❖ Many different types of VACs each accompanied with a publication



x velocity  
x mass  
x temperature ...

# Data Products: Large Scale Structure Catalogs

---

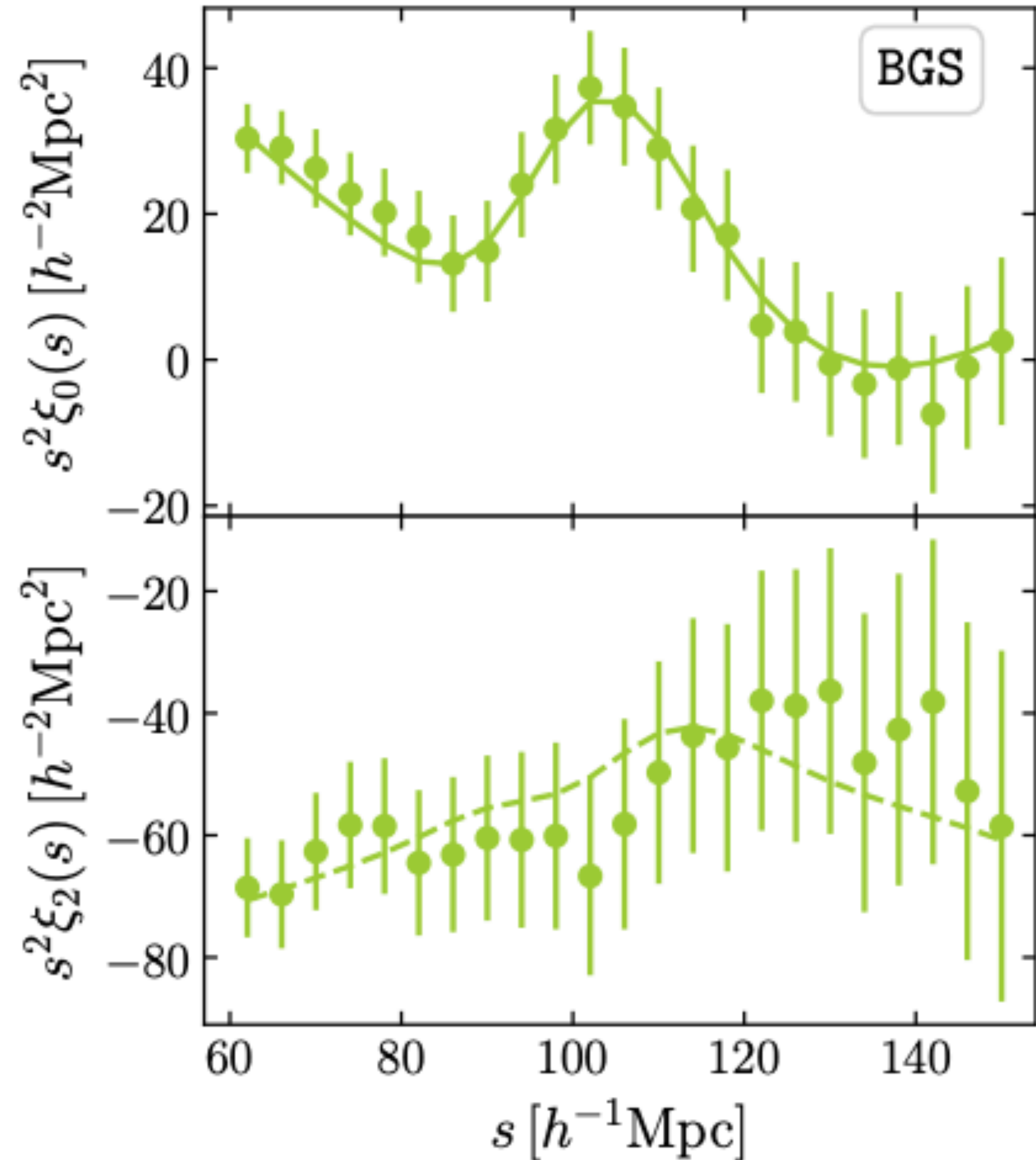
---

- ❖ “Cosmology-ready” catalogs
- ❖ Data “cleaned” of systematics
- ❖ Cosmology VAC



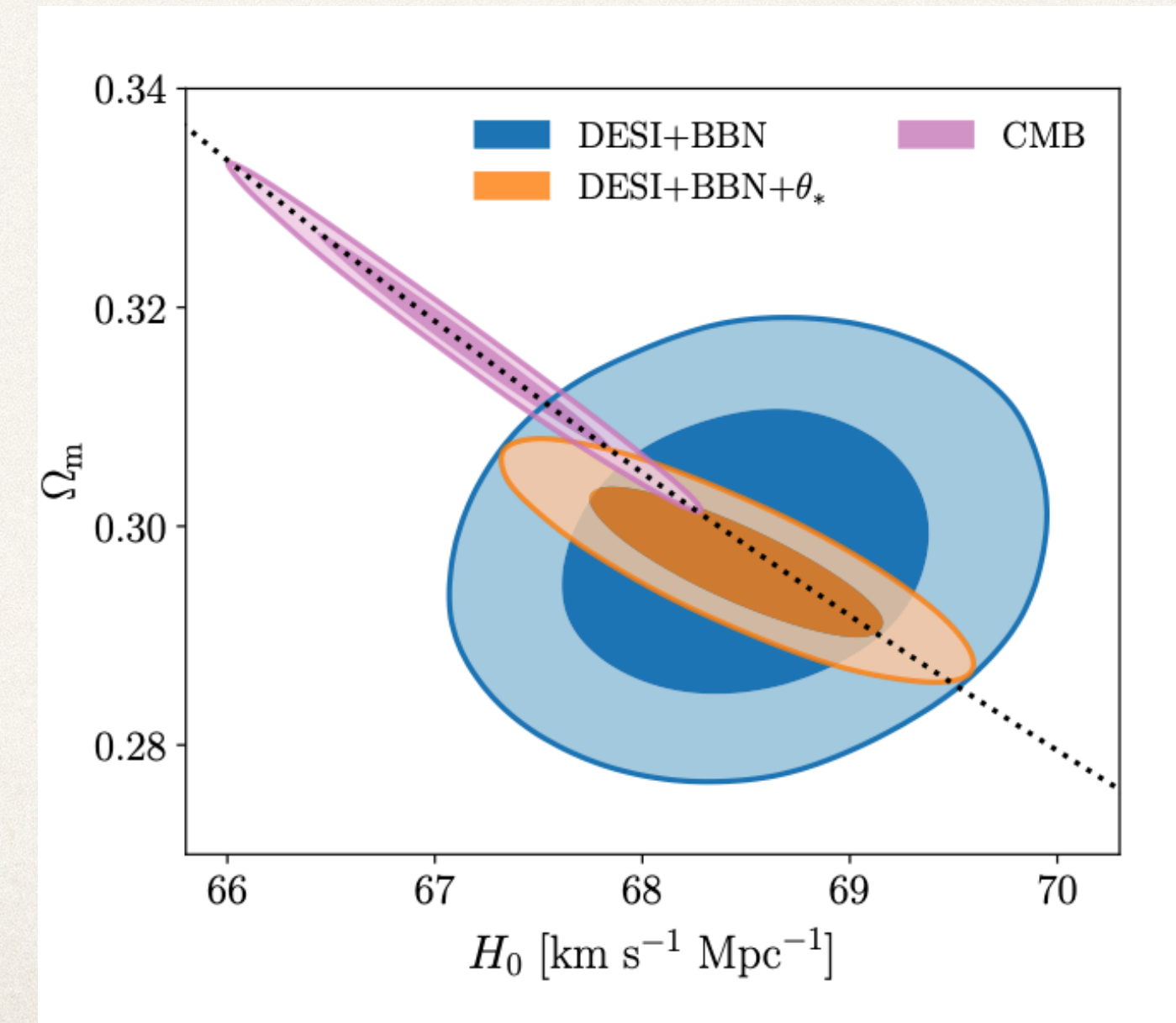
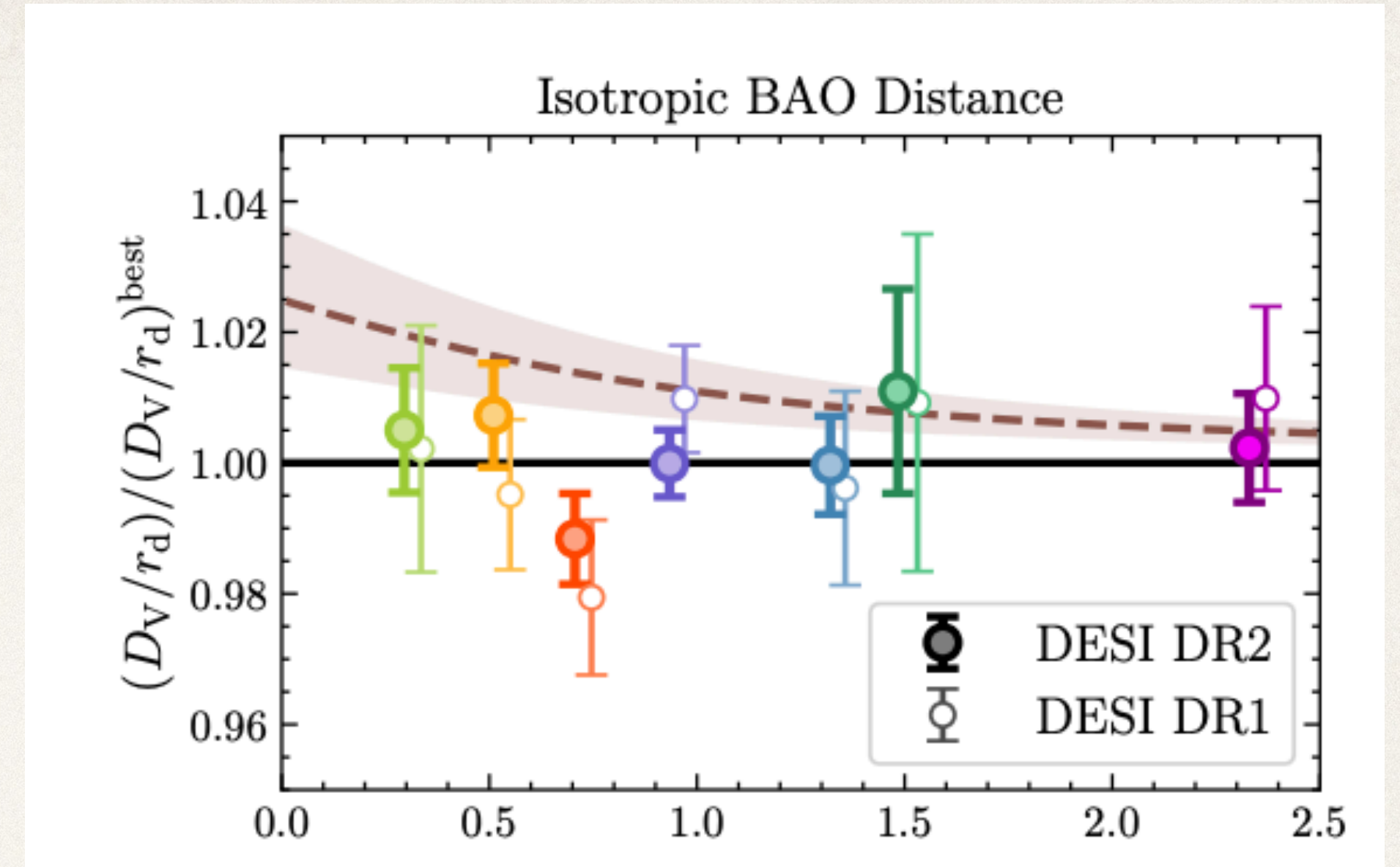
# Data Products: Compressed Statistics

- ❖ 100s of data bins instead of millions of objects
- ❖ Contains most of the information on the distribution of galaxies in the Universe
- ❖ Intermediate step processing the data



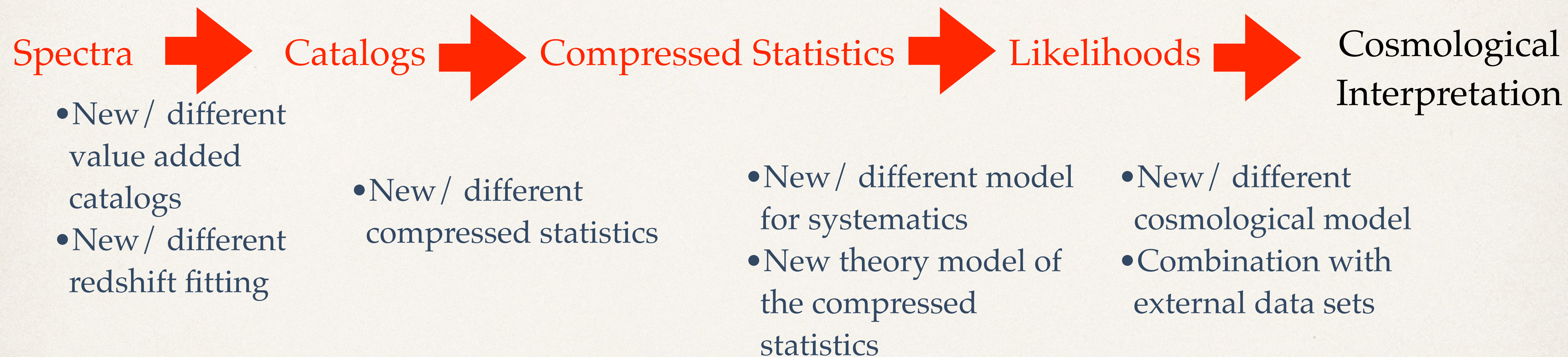
# Data Products: Likelihoods

- ❖ Sorry, we are mostly Bayesians (there is only one observable Universe)...
- ❖ A further compression step for speed



# How is the data used in the community?

---



Astronomy application:

- galaxy formation
- ...endless list



# DESI policies

---

- ❖ For key projects (papers): data and analysis pipeline must be made public
- ❖ Data availability statement for each publication
- ❖ Zenodo repo for all DESI publications

# Data Releases

---

- ❖ Website (Access, summary, information etc)
- ❖ Publication for each data release
- ❖ Github repository to recreate figures of the data release paper (notebook, processed data, figures)
- ❖ Github repo of all the software used to process the data
- ❖ Jupyter notebook tutorials on how to use the data

<https://data.desi.lbl.gov/doc/>

# DESI management team

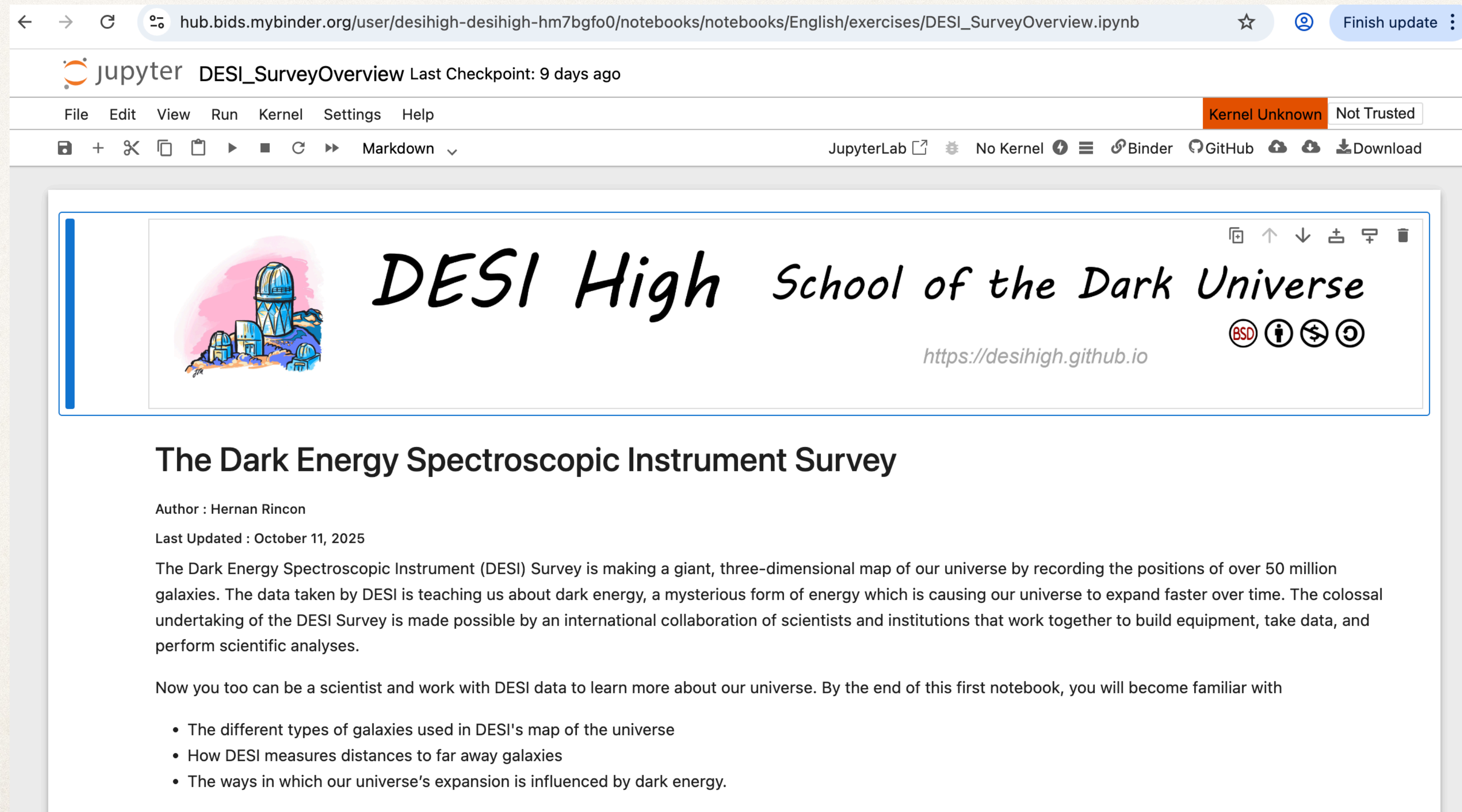
---

## Data System (WBS 1.8)

---

- Stephen Bailey ✉ [stephenbailey@lbl.gov](mailto:stephenbailey@lbl.gov) - Manager
  - David Kirkby ✉ [dkirkby@uci.edu](mailto:dkirkby@uci.edu) - Data System Scientist
  - Kyle Dawson ✉ [kdawson@astro.utah.edu](mailto:kdawson@astro.utah.edu) - Survey planning and operations
  - Robert Cahn ✉ [rncahn@lbl.gov](mailto:rncahn@lbl.gov) - Fiber Assignment Lead
  - Julien Guy ✉ [jguy@lbl.gov](mailto:jguy@lbl.gov) - Spectroscopic Pipeline Lead
  - Lado Samushia ✉ [lado@phys.ksu.edu](mailto:lado@phys.ksu.edu) - Large Scale Structure Catalog
  - Benjamin Weaver - Data Transfer, Archive & Distribution
  - Stephen Bailey ✉ [stephenbailey@lbl.gov](mailto:stephenbailey@lbl.gov) - Integration & Testing of Data Systems
  - Dustin Lang ✉ [dstn@cmu.edu](mailto:dstn@cmu.edu) - DESI Imaging Scientist
  - Adam Myers ✉ [geordiemyers@gmail.com](mailto:geordiemyers@gmail.com) - Target Selection Pipeline

# DESI High for High School Students



hub.bids.mybinder.org/user/desihigh-desihigh-hm7bgfo0/notebooks/notebooks/English/exercises/DESI\_SurveyOverview.ipynb


Finish update

jupyter DESI\_SurveyOverview Last Checkpoint: 9 days ago

File Edit View Run Kernel Settings Help

Kernel Unknown Not Trusted

JupyterLab No Kernel Binder GitHub Download



## DESI High School of the Dark Universe

<https://desihigh.github.io>

BSD i \$ ©

### The Dark Energy Spectroscopic Instrument Survey

Author : Hernan Rincon

Last Updated : October 11, 2025

The Dark Energy Spectroscopic Instrument (DESI) Survey is making a giant, three-dimensional map of our universe by recording the positions of over 50 million galaxies. The data taken by DESI is teaching us about dark energy, a mysterious form of energy which is causing our universe to expand faster over time. The colossal undertaking of the DESI Survey is made possible by an international collaboration of scientists and institutions that work together to build equipment, take data, and perform scientific analyses.

Now you too can be a scientist and work with DESI data to learn more about our universe. By the end of this first notebook, you will become familiar with

- The different types of galaxies used in DESI's map of the universe
- How DESI measures distances to far away galaxies
- The ways in which our universe's expansion is influenced by dark energy.

# DESI High for High School Students

The screenshot shows the DESI High website interface. At the top left is the DESI High logo. The navigation menu includes 'DESI High', 'FAQ', 'Teachers', 'Partners', 'About', and 'Code of Conduct'. On the right side of the header, there are buttons for 'Let's start!' (yellow), 'Forum' (white), and a search icon (AZ) with a dropdown arrow.

The main content area features two primary sections:

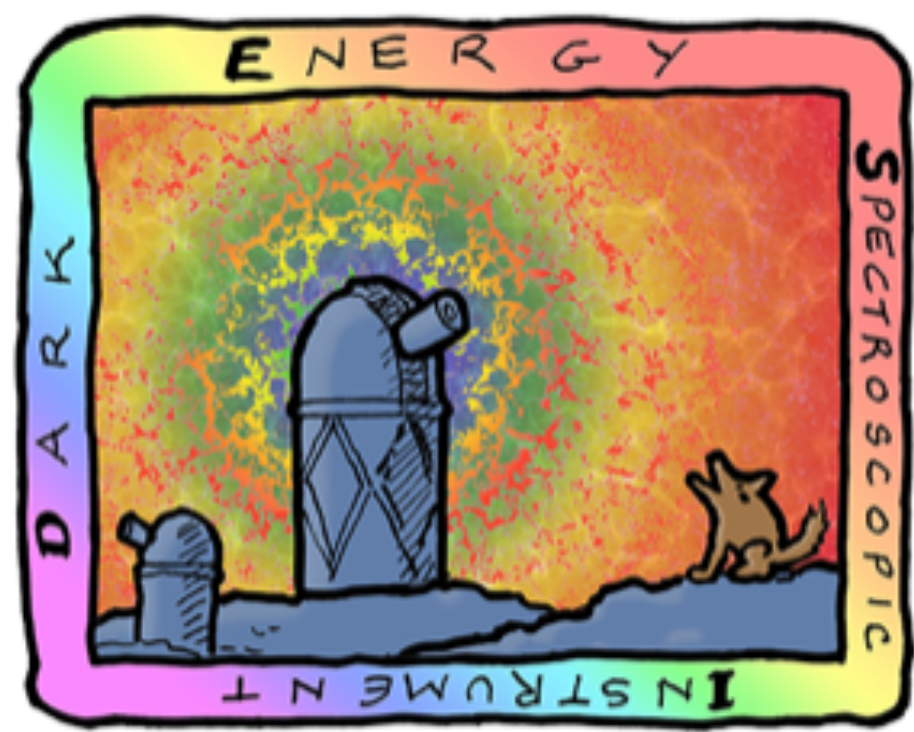
- The DESI Survey**: A green header section with the subtext 'Explore the DESI survey, its goals, and how it maps the universe.' Below this are two tabs: 'Exercises' (active) and 'Solutions'. A list of four items follows:
  - DESI Survey Overview (Easy)
  - Galaxy Spectra (Easy, marked with a red 'X')
  - Redshift and Distance (Medium)
  - Mapping the Universe (Medium)
- Physics 101**: A dark grey header section with the subtext 'Get a solid foundation in the physics concepts that underpin cosmology.' Below this are two tabs: 'Exercises' (active) and 'Solutions'. A list of two items follows:
  - Angular Coordinates on the Celestial Sphere (Easy)
  - Wave Physics (Easy, marked with a red 'X')

How can we avoid miss-usage of our data?

*Data processing is complex...*

Is a global platform for higher education resources using open data possible?

*Each collaboration or research field has resources...the problem is knowing that it exists  
and where to find it!*



# DARK ENERGY SPECTROSCOPIC INSTRUMENT

U.S. Department of Energy Office of Science

*Anyone should be able to reproduce key results!*

Guiding principle

---

# Key points

---

- ❖ Challenge: How to avoid miss-usage of our data?
- ❖ Future of open data in HE: Is a global platform for resources possible?
- ❖ Guiding principle: Anyone should be able to reproduce our key results

