

Large Scale Structure of the Universe

Rogério Rosenfeld
IFT-UNESP/ICTP-SAIFR/LIneA

Dark Energy Survey & LSST



Five lectures on Cosmology and Large Scale Structure

Lecture I: The average Universe

Lecture II: Distances and thermal history

Lecture III: The perturbed Universe

→ Lecture IV: Theoretical challenges and surveys

Lecture V: Observational cosmology with LSS

Plan for Lecture IV:

IV.1 – Theoretical challenges for LSS

IV.2 – Large Galaxy Surveys

IV.1- Theoretical challenges for LSS

IV.1.1 – Models of the Universe

IV.1.2 – Linear matter power spectrum

IV.1.3 – Challenge I: nonlinear modelling of matter perturbations

IV.1.4 – Challenge II: baryonic physics

IV.1.5 – Challenge III: the matter to galaxy connection

IV.1.1 – Models of the Universe

Evolution of perturbations must be computed in a given cosmological model. Which one?

Simplest model: flat Λ CDM.

Characterized by 6 cosmological parameters:

$$\{A_s, n_s, h, \Omega_m, \Omega_b \text{ and } \Omega_v\}$$

Next simplest: flat w CDM.

+ w (constant dark energy equation of state)

Next-to-next simplest: flat $w_0 w_a$ CDM.

+ $w_0 + w_a$ (dark energy equation of state: $w(a) = w_0 + (1-a) w_a$)

and many other models with **more parameters!**

IV.1.2 – Linear Matter Power Spectrum

A realistic linear matter power is obtained using a Boltzmann solver code (see Julien's lectures).

$$\delta_{\text{ini}} \xrightarrow{\text{"D(t) A(k)"}} \delta(t)$$

The most used ones are CLASS and CAMB.

Core Cosmology Library (CCL) is a public code developed by the Dark Energy Science Collaboration (DESC) of the Legacy Survey of Space and Time (LSST) Project.

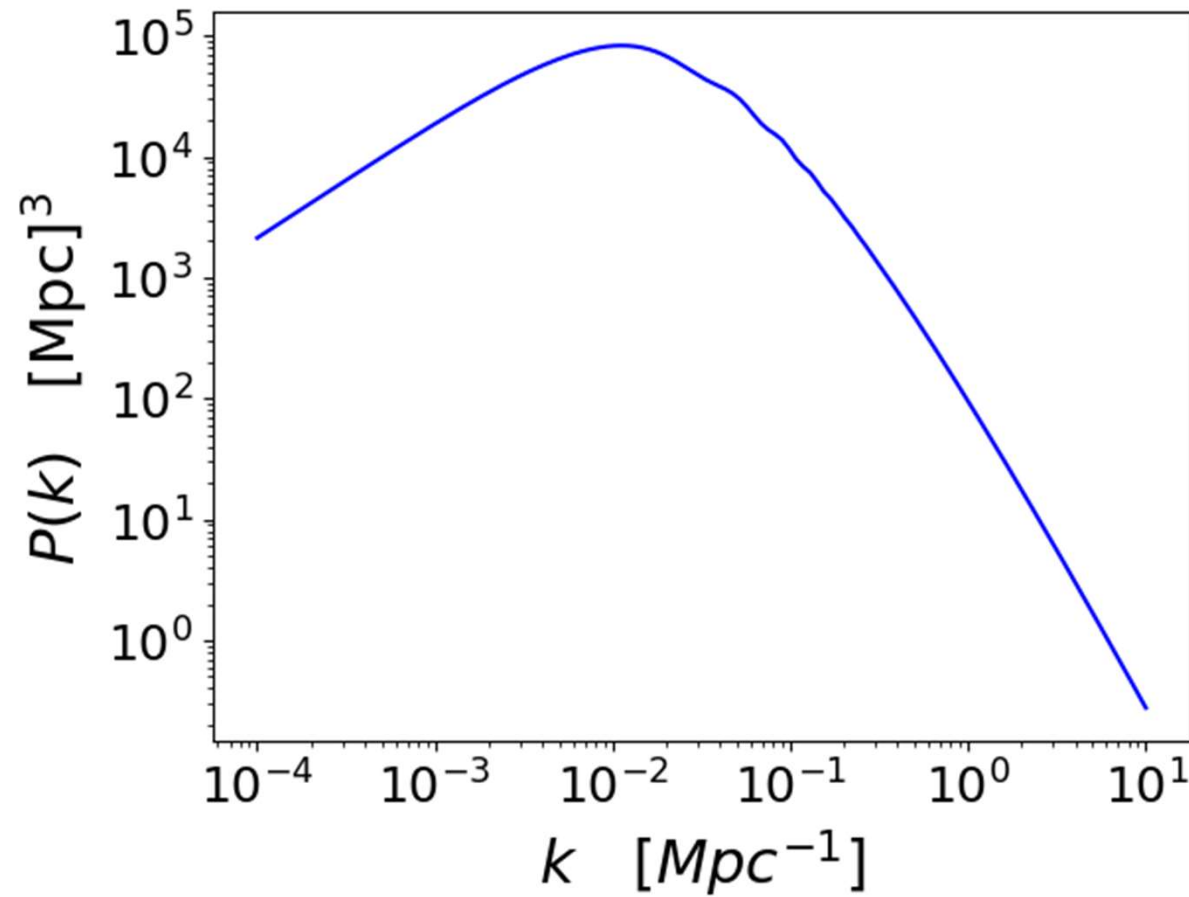
<https://ccl.readthedocs.io/en/latest/index.html>

<https://arxiv.org/abs/1812.05995>

See Google Colab notebook.

```
cosmo = ccl.Cosmology(Omega_c=0.27, Omega_b=0.045, h=0.67, A_s=2.1e-9,  
n_s=0.96,transfer_function='boltzmann_camb')
```

```
pk_lin = ccl.linear_matter_power(cosmo, k, a)
```



The linear power spectrum is a good approximation for small perturbations.
For CMB, $\delta \sim 10^{-5}$ and Julien is happy.

The linear power spectrum is **not** valid when perturbations are large at late times.

There are three main challenges to model the power spectrum at late times:

- . Nonlinear GR effects
- . Baryon physics
- . Connection between dark matter and galaxies

We will briefly touch on these challenges. I'll show references to tools that deal with them.

Many new physics models result in modifications of the power spectrum at small scales:
fuzzy dark matter, warm dark matter, modified gravity, etc (see Mustafa's lectures)
It is important to understand theoretically the power spectrum at small scales,

IV.1.3 – Challenge I: nonlinear modelling of matter perturbations

We have so far studied the linearized Newtonian evolution of perturbations.

CLASS and CAMB use full linearized GR.

But linearization approximation breaks down for large perturbations – small scales.

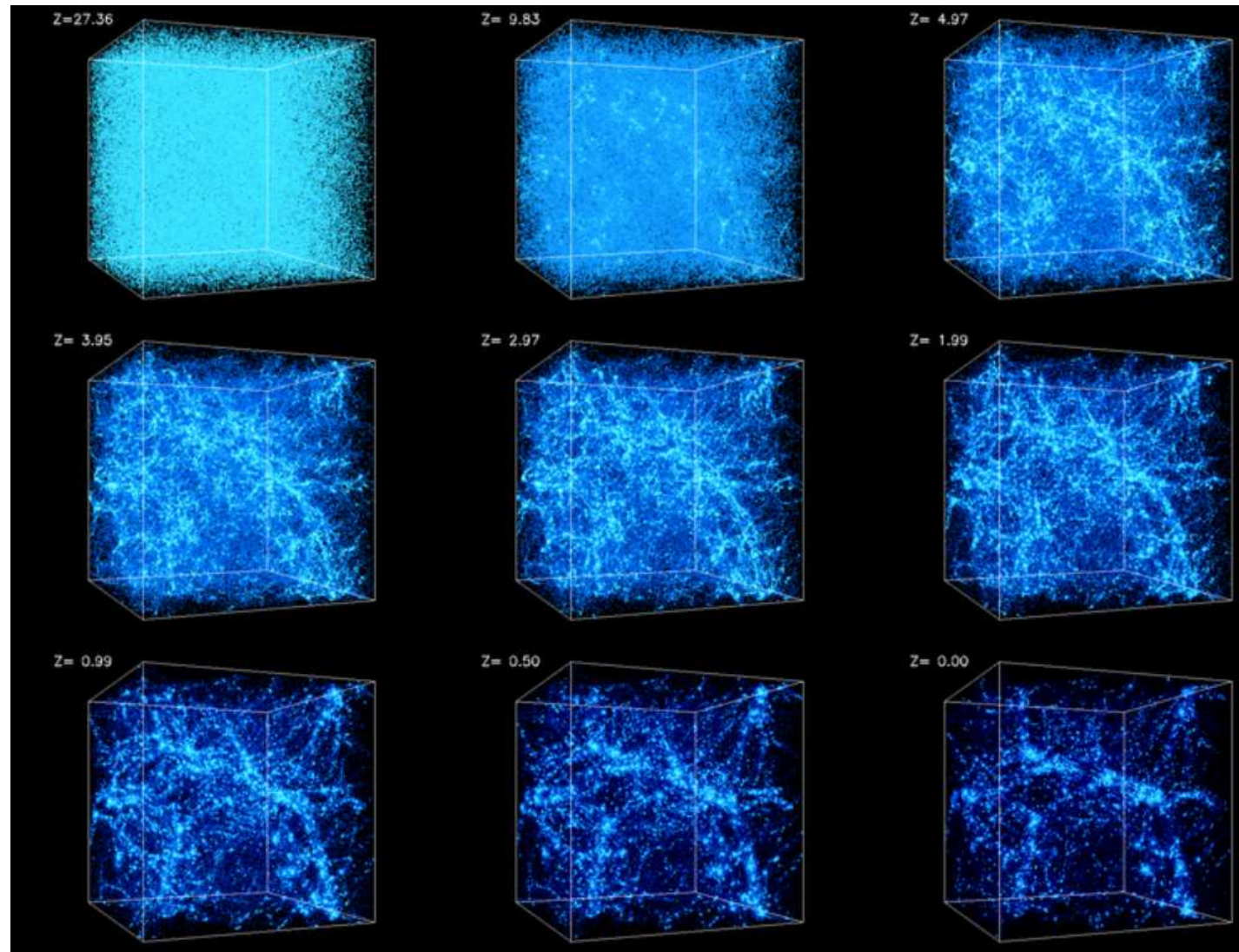
GR is highly nonlinear.

Several approaches have been developed to study the evolution of perturbations at the nonlinear level. I just mention a few in the following slides.

- N-body simulations

Newtonian physics usually.

Example on the right from
<http://cosmicweb.uchicago.edu/filaments.html>
DM-only



- Fits to N-body simulations

Widely used: **Halofit**, validated on Λ CDM

<https://arxiv.org/abs/1208.2701>

Based on 16 simulations. Claim 5% precision in the power spectrum for $k \leq 1 h/\text{Mpc}$ at $0 \leq z \leq 10$

Nonlinear power spectra based on a halo model

Most used: **HMCode**

<https://arxiv.org/abs/2009.01858>

Claim 2.5% per cent across a range of cosmologies, scales $k < 10 h\text{Mpc}^{-1}$, and redshifts $z < 2$

For beyond Λ CDM cosmologies: **ReACT**
Halo model reaction to non- Λ CDM cosmologies
<https://arxiv.org/abs/2005.12184>

- Fast N-body simulations

COLA

COLA (COmoving Lagrangian Acceleration) method described in S. Tashev, M. Zaldarriaga, D. Eisenstein 2013 (arXiv:1301.0322).
<https://bitbucket.org/tashev/colacode/src/hg/>

FastPM: <https://arxiv.org/abs/1603.00476>

- Lognormal realizations (really fast simulations from input summary statistics (not N-body))

FLASK, GLASS

GLASS: Generator for Large Scale Structure
<https://arxiv.org/abs/2302.01942>

- Emulators

N-body simulations are computationally expensive.

They can't be used for fast generation of power spectra for different cosmological parameters.

Suites of simulations for different cosmological parameters can be used to build fast **emulators** for the nonlinear matter power spectrum.

Example: **EuclidEmulator2** - <https://arxiv.org/abs/2010.11288>

Percent-level accurate emulation supported in the eight-dimensional parameter space of $w_0 w_a \text{CDM} + \sum m_\nu$ models between redshift $z=0$ and $z=3$ for $0.01 \text{ h/Mpc} \leq k \leq 10 \text{ h/Mpc}$. Based on more than 250 high-resolution simulations.

- Analytical methods

Higher-order perturbation theory, effective field theory, ...
Limited to $k \sim 0.5 \text{ h/Mpc}$.

Large body of work. Codes such as CLASS-PT, FAST-PT,...

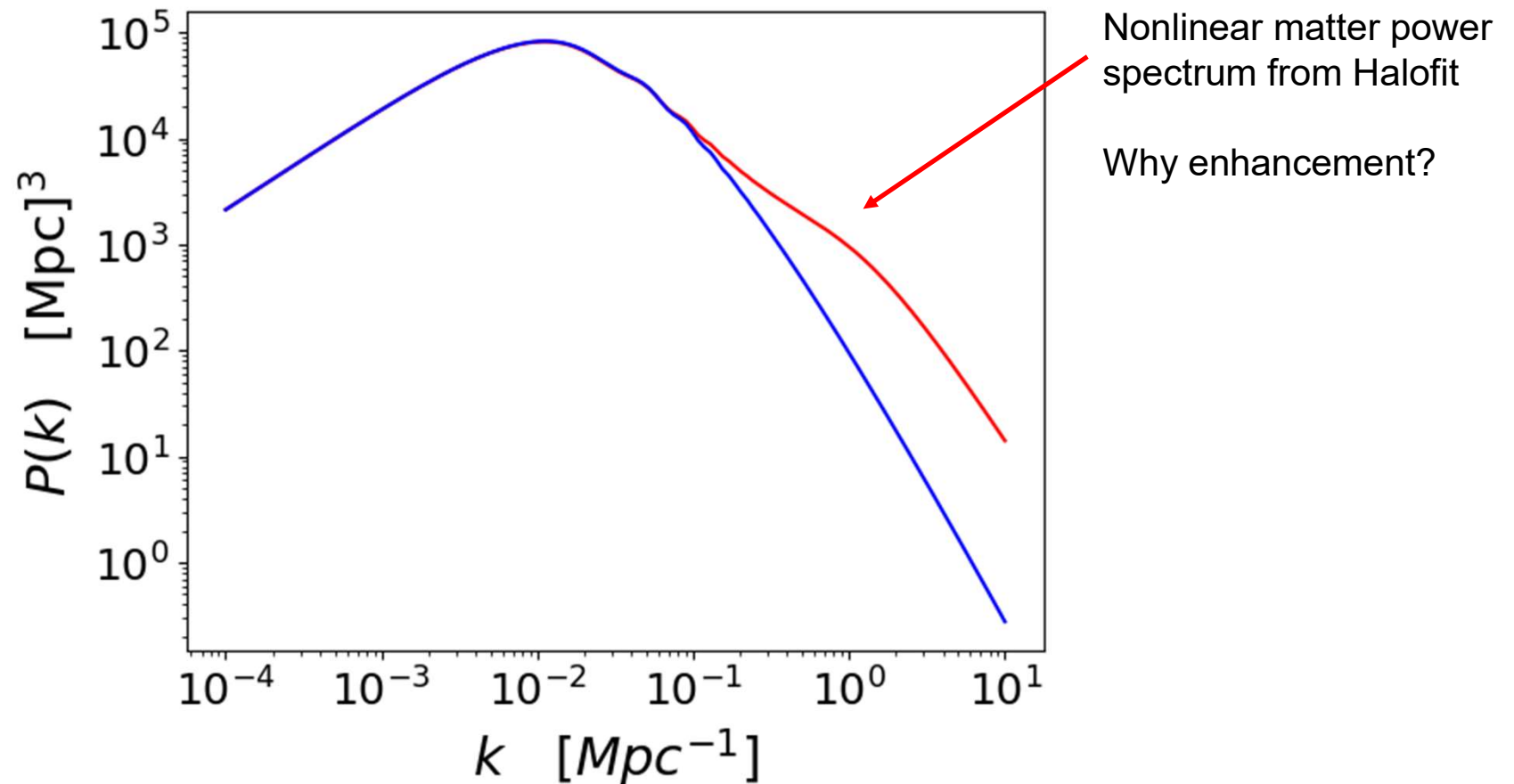
FAST-PT is a code to calculate quantities in cosmological perturbation theory at 1-loop (including, e.g., corrections to the matter power spectrum).

<https://github.com/JoeMcEwen/FAST-PT>

CLASS-PT is a modification of the CLASS code that computes the non-linear power spectra of dark matter and biased tracers in one-loop cosmological perturbation theory, for both Gaussian and non-Gaussian initial conditions.

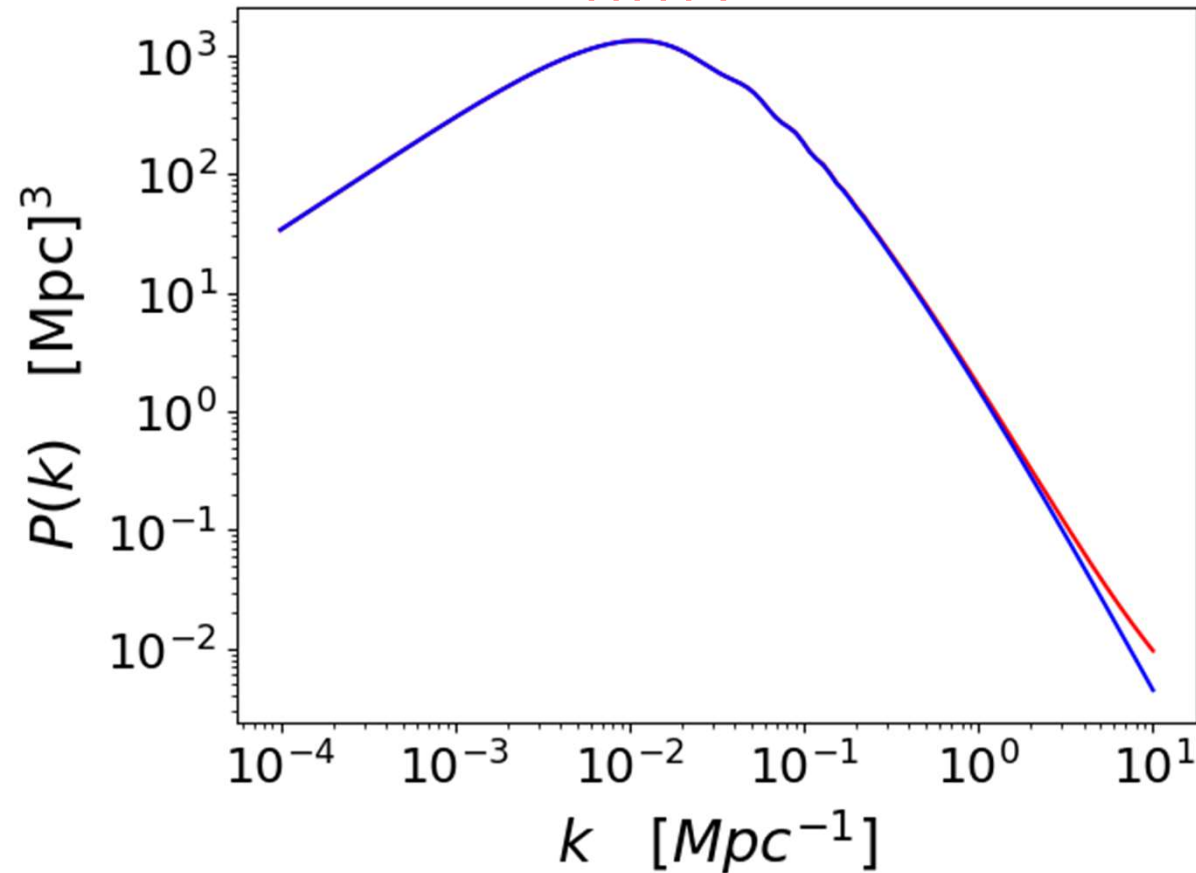
<https://github.com/Michalychforever/CLASS-PT>

Linear power spectrum is a good approximation (**at $z=0$**) for k up to $\sim 0.1 \text{ Mpc}^{-1}$



Linear power spectrum is a good approximation (**at $z=10$**) for k up to $\sim 5 \text{ Mpc}^{-1}$

WHY?



Bottom line:

For simple cosmological models there are already fast emulators.

Emulators are being developed for more complex models using neural networks.
Example: CosmoPower - <https://arxiv.org/abs/2106.03846>

Recent works on beyond- Λ CDM cosmologies with COLA – eg
“Enabling matter power spectrum emulation in beyond- Λ CDM cosmologies with COLA”
<https://arxiv.org/abs/2203.11120>
Hi-COLA: fast, approximate simulations of structure formation in Horndeski gravity
<https://arxiv.org/abs/2209.01666>

Nonlinear predictions for the power spectrum of Λ CDM and beyond are **fundamental** in order to use the full data set available in surveys (small scales - see next lecture)

IV.1.4 – Challenge II: baryonic physics

The Universe would be much simpler if we (baryons) were not around!

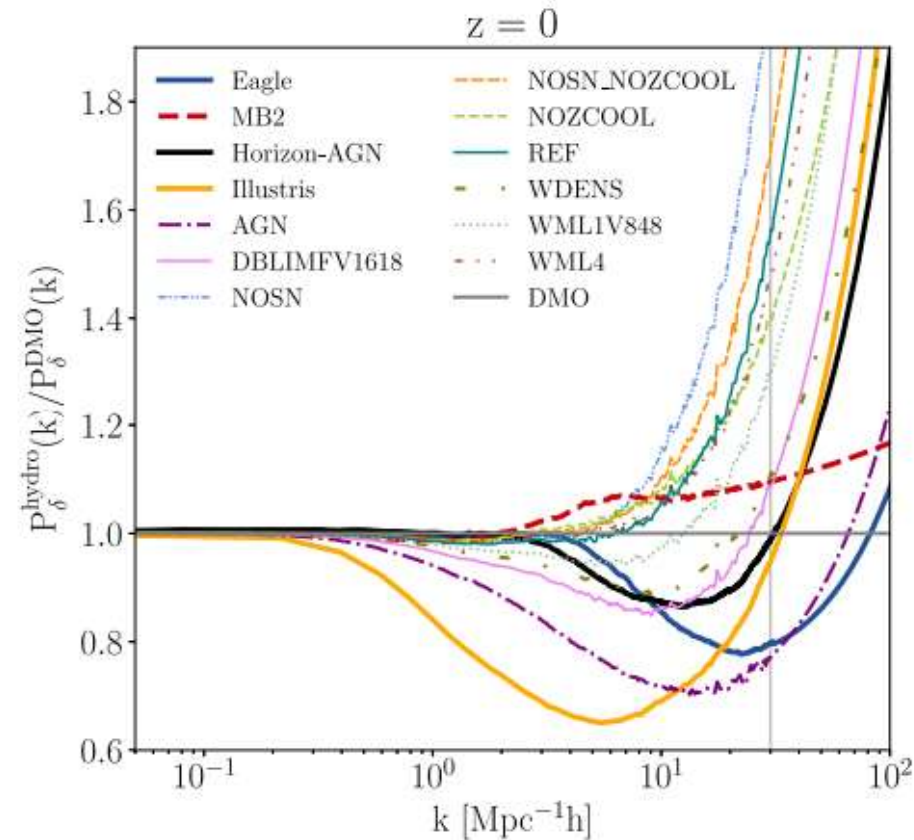
Baryons experience dissipative (EM) processes. They form stars that explodes and can halt the growth of perturbations at small scales together with AGNs: baryonic feedback.

This results in a suppression in the matter power spectrum at small scales.

These effects can be studied with detailed simulations that take into account these processes. Unfortunately there are many uncertainties in these simulations (many parameters). And they are very computer-intensive.

HMCode can model baryonic feedback in an empirical way.

Must study techniques to mitigate the baryonic uncertainties.



Uncertainties in the matter power spectrum from baryonic physics in 14 different simulations. From <https://arxiv.org/abs/1809.01146>

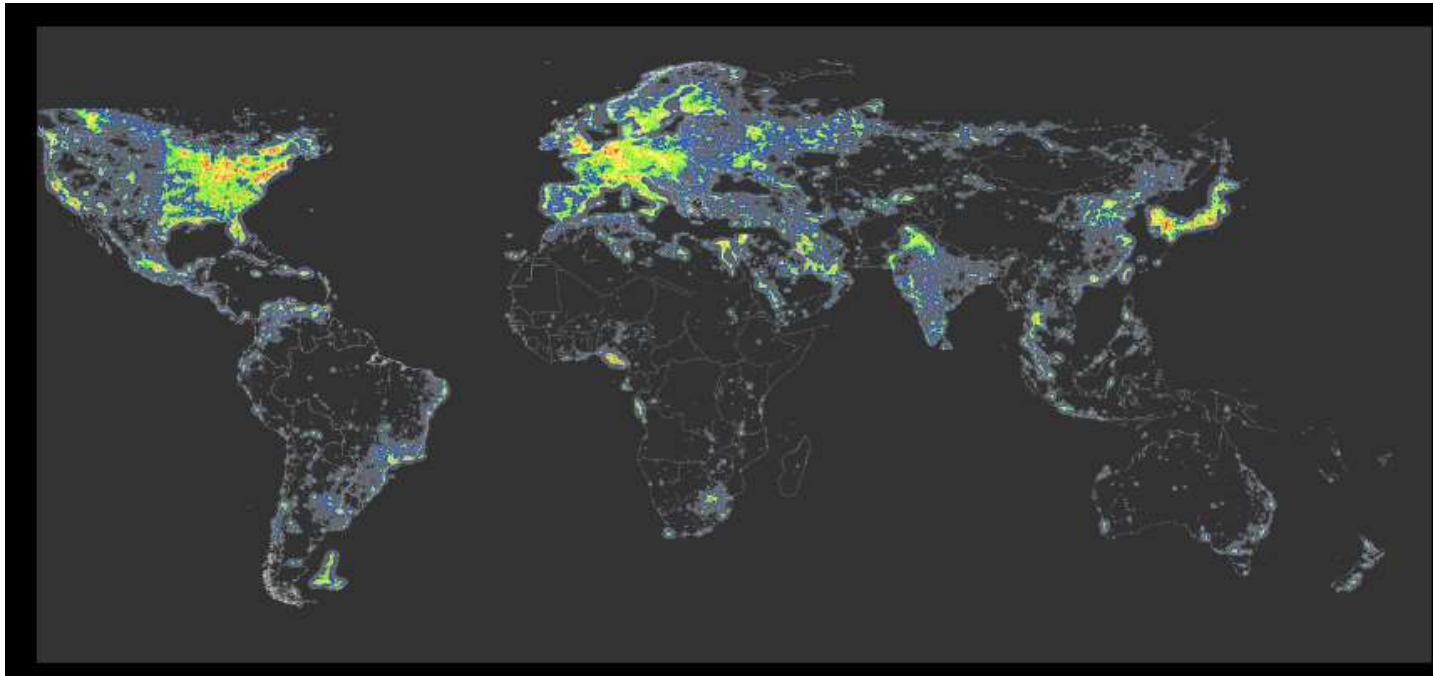
A method to include these uncertainties in the covariance matrix was proposed in Moreira et al. - <https://arxiv.org/abs/2104.01397>

IV.1.5 – Challenge III: the matter to galaxy connection

Theory so far predicts the growth of matter perturbations.

We observe galaxies.

Galaxies are biased tracers of the total matter distribution.



Galaxy bias: relation between galaxy distribution and total matter distribution.

$$\delta_{\text{galaxies}} \leftrightarrow \delta_{\text{matter}}$$

Galaxies form in regions where the matter density is “large enough”
(enough can be more precisely defined in simple models such as the spherical collapse).

Galaxy bias depend on redshift (since the Universe is smoother at high redshifts a larger bias is necessary).

Galaxy bias depend on the tracer (galaxy type, quasar, etc).

Simplest model: linear bias

$$\delta_{\text{galaxies}} = b_1(z) \delta_{\text{matter}}$$

Spoiler: for the Dark Energy Survey Y3 it was sufficient

In order to go beyond linear bias there are two main approaches:

- Phenomenological models:
based on a Halo Occupation Distribution (HOD) model combined with a halo mass function
- More “principled” models:
based on Effective Field Theory expansions in operators that are allowed by symmetries (such as rotation/galilean invariance, equivalence principle, etc). Lectures next week!

Three slides on EFT of galaxy bias: bias expansion (much more in Marko's lectures)

Expansion can be performed using Eulerian or Lagrangian coordinates.

Eulerian expansion in $\partial_i \partial_j \Phi$

$$\delta_g = \underbrace{b_1}_{\text{Linear bias}} \delta_m + \underbrace{b_2}_{\text{quadratic bias}} (\delta_m^2 - \langle \delta_m^2 \rangle) + \underbrace{b_{s^2}}_{\text{tidal bias}} (s^2 - \langle s^2 \rangle) + \underbrace{b_{\nabla^2}}_{\text{nonlocal bias}} \nabla^2 \delta_m + \underbrace{\varepsilon}_{\text{stochastic bias}}$$

Galaxy bias in the era of LSST: perturbative bias expansions -


<https://arxiv.org/abs/2307.03226>

"bias challenge paper"

Power spectrum in the Eulerian bias expansion:

$$P_{gg}(k) = \sum_{i,j} b_i b_j P_{ij}(k) + P_{SN}$$

Auto-correlation of
stochastic bias



$$b_i \in \{b_1, b_2, b_{s^2}, b_{\nabla^2}\}$$

$$\text{Example: } P_{1\nabla^2} \sim \langle \delta_m(\nabla^2 \delta_m) \rangle$$

A similar expansion can be made in Lagrangian coordinates.

The Eulerian expansion implemented in FAST-PT code

<https://github.com/JoeMcEwen/FAST-PT>.

The Lagrangian expansion implemented in Velocileptors code

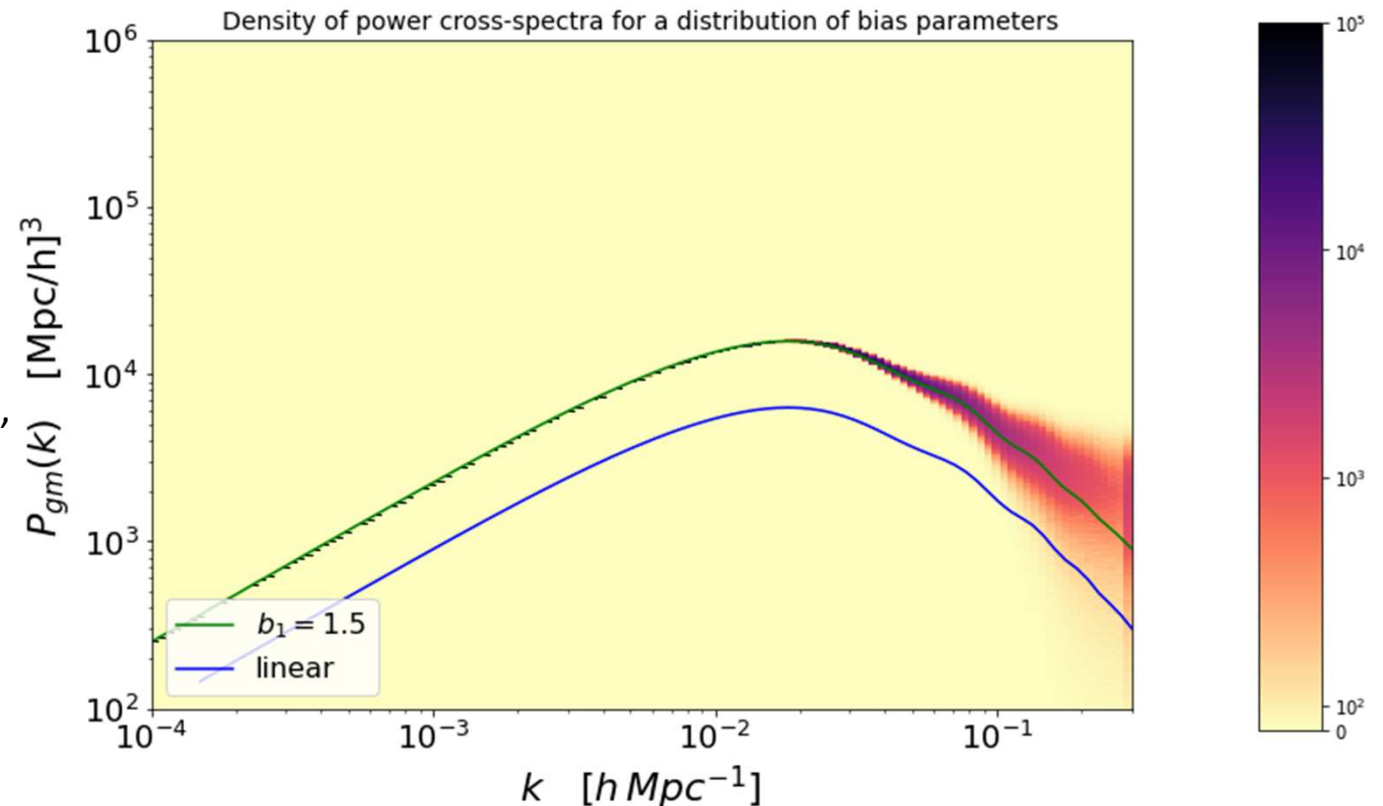
<https://github.com/sfschen/velocileptors>.

A hybrid EFT including results from N-body simulations has been recently proposed.
anzu , a code for the cosmology dependence of hybrid EFT spectra for clustering and lensing
modeling- **1% accurate up to $k = 0.6 \text{ h/Mpc}$**

Abdias Aires (+ N. Kokron + F. Andrade-Oliveira + V. Miranda + RR): study mitigation in bias modelling uncertainties using velocileptors

10^5 samples of bias vectors that obey the probability distributions:

- b_1 : fixed at 1.5;
- b_2 : uniform in $(-5,5)$;
- counterterm: normal centered at 0, with mean deviation of 50;
- shot-noise: normal centered at 3000, with mean deviation of 30%;
- b_s : uniform in $(-5,5)$;
- b_3 : uniform in $(-5,5)$



IV.2- Large Galaxy Surveys

IV.2.1 – Landscape of galaxy surveys

IV.2.2 – Dark Energy Survey

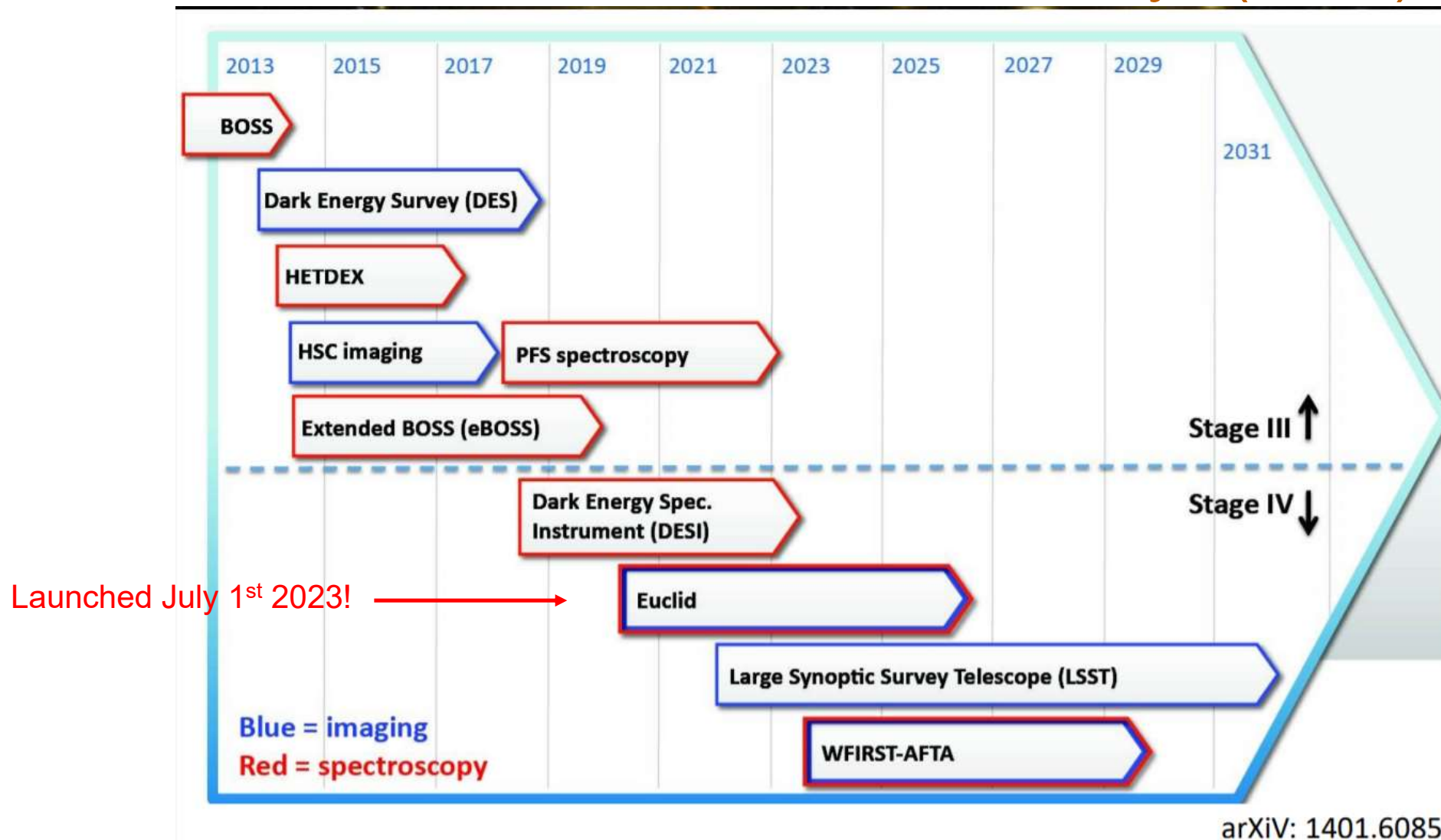
IV.2.3 – Legacy Survey of Space and Time

IV.2.1 – Landscape of galaxy surveys

Two main types of galaxy surveys:

- Spectroscopic: take spectra of galaxies
(good quality spectroscopic redshift vs smaller number of objects; no imaging)
- Photometric (imaging): take pictures of galaxies with different color filters
(fair quality photometric redshift vs larger number of objects; imaging)

An outdated schedule of surveys (2014)



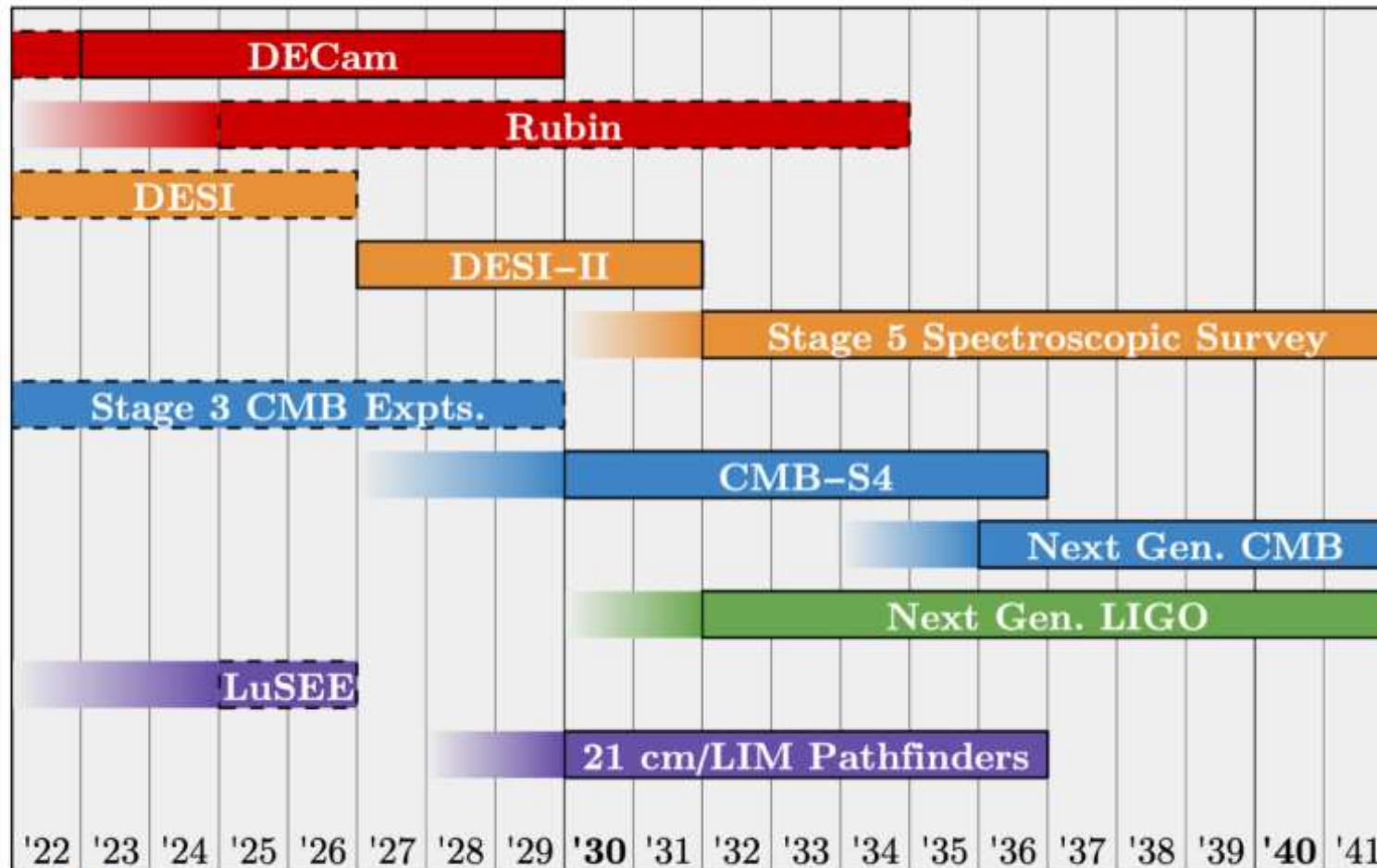


Figure 6-1. Current and potential future facilities probing cosmic acceleration that are or may be supported by DOE or NSF. Dashed boxes indicate fully-funded facilities. Facilities in red are optical imaging, in orange are optical spectroscopy, in blue are CMB, in green are gravitational waves, and in purple are radio/mm spectroscopy. The fade-in regions indicate commissioning periods, while the boxes indicate full survey observations.

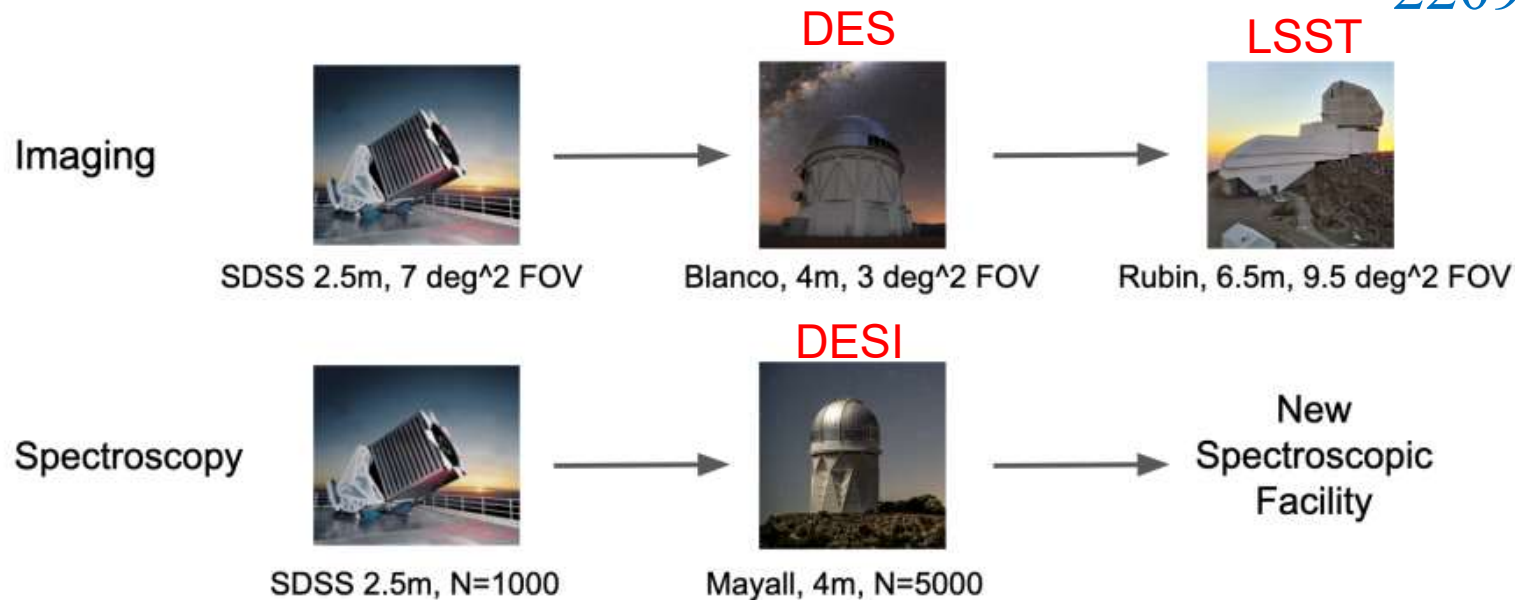
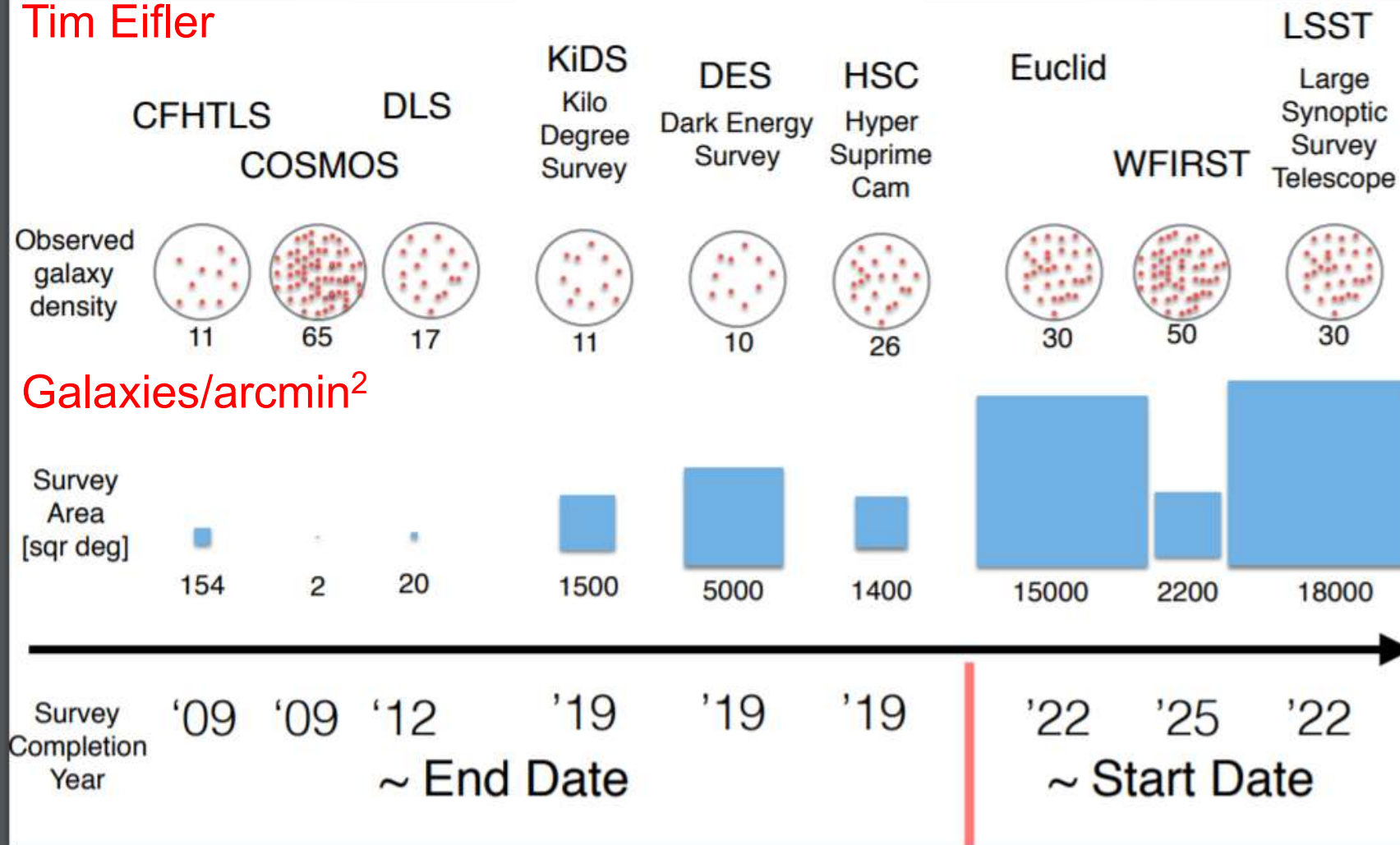


Figure 6-4. Summary of imaging and spectroscopic surveys and facilities, ongoing and planned, that are supported by DOE/NSF partnerships. The international ground and space-based landscape of optical wide-field surveys, ongoing and planned, is very rich but for clarity is not represented here. SDSS had both imaging and spectroscopic capabilities, the Blanco telescope was used to carry out the DES, and the Mayall is currently used for DESI. In the near future, the Rubin Observatory will begin LSST. A new spectroscopic facility would open up new scientific opportunities.

Photometric Dark Energy Surveys

Tim Eifler



In the following I will focus on two surveys that I'm a member of:

- Dark Energy Survey (DES)
- Rubin Observatory's Legacy Survey of Space and Time (LSST)

DES and LSST are photometric surveys and share the same techniques.

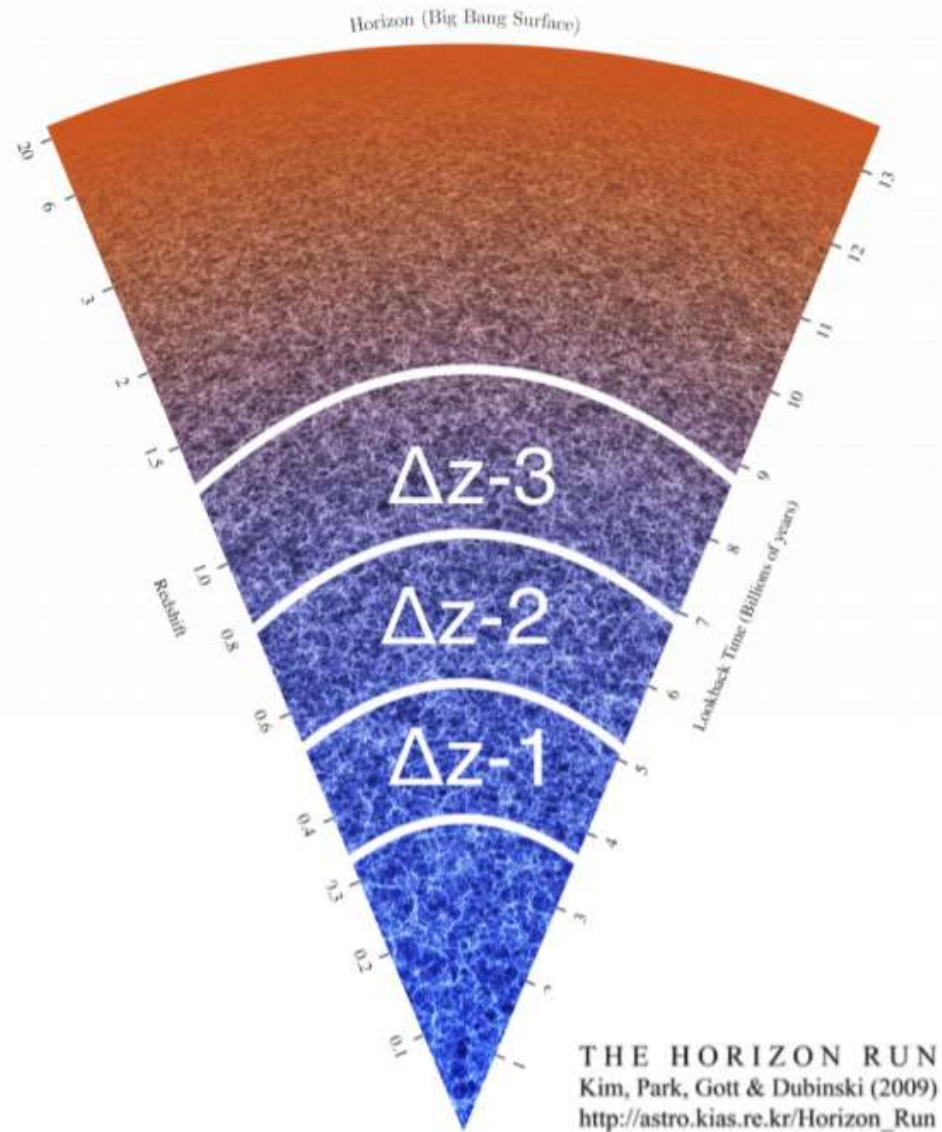
Photometric redshifts are less precise than spectroscopic ones.

3D information is sliced into redshift bins ("tomographic redshift bins").

Analyses are performed in 2D for each redshift bin.

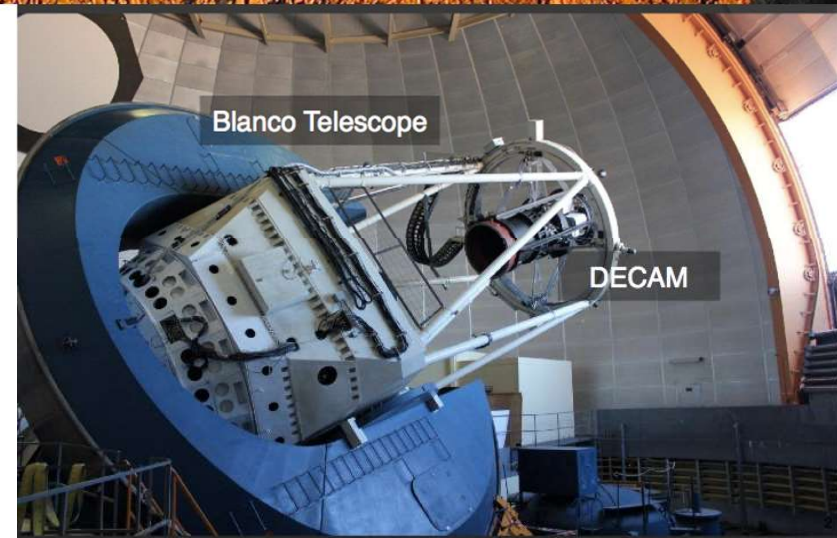
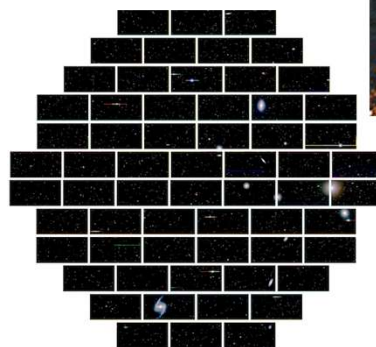
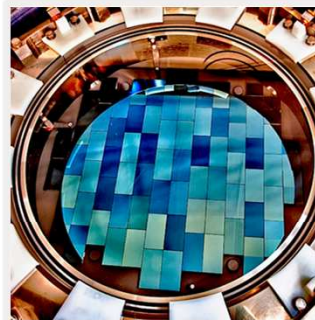
Angular correlation functions and/or angular power spectra.

Make
measurements
tomographically



IV.2.2 – Dark Energy Survey

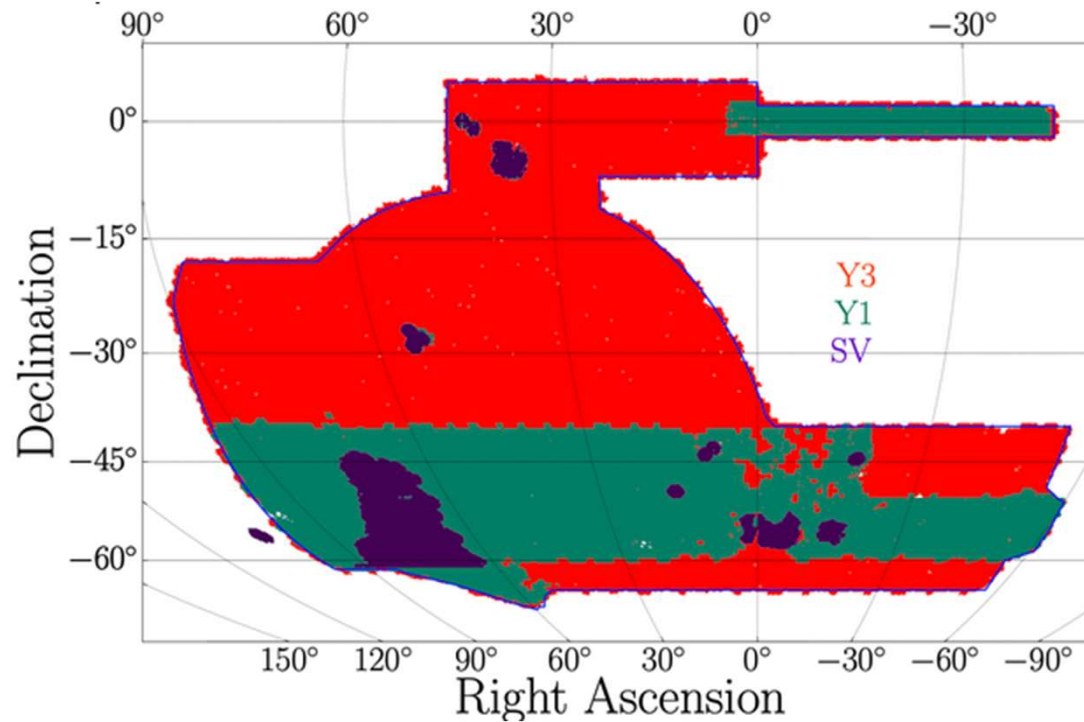
- >600 members, 25 institutions, 7 countries (only Brasil from LA)
- 570 Megapixel camera for the Blanco 4m telescope in Chile.
- Full survey, ~5.5Y. 2013-2019 (**Y3 2013-16**).
- Wide field: 5000 sq. deg. in 5 bands grizY. ~23 magnitude.
- DES Y3: Positions and shapes of > **100M galaxies**.
- DES Y6: final analysis under way!



Y3 Gold Catalog (Sevilla-Noarbe et al., 2011.03407)

- The data included in Y3 Gold spans 345 distinct nights of observations with at least one observation passing quality tests from 2013 August 15 to 2016 February 12
- Gold sample: selection of objects from multi-epoch images passing quality cuts, photometric calibration, masking, signal-to-noise >10 , objects brighter than $i>23$, star-galaxy separation, survey property maps, ...
Total of 319 million objects.

DES footprint (Science Verification, Y1 and Y3)



Sevilla-Noarbe, Bechtol et al., 2011.03407

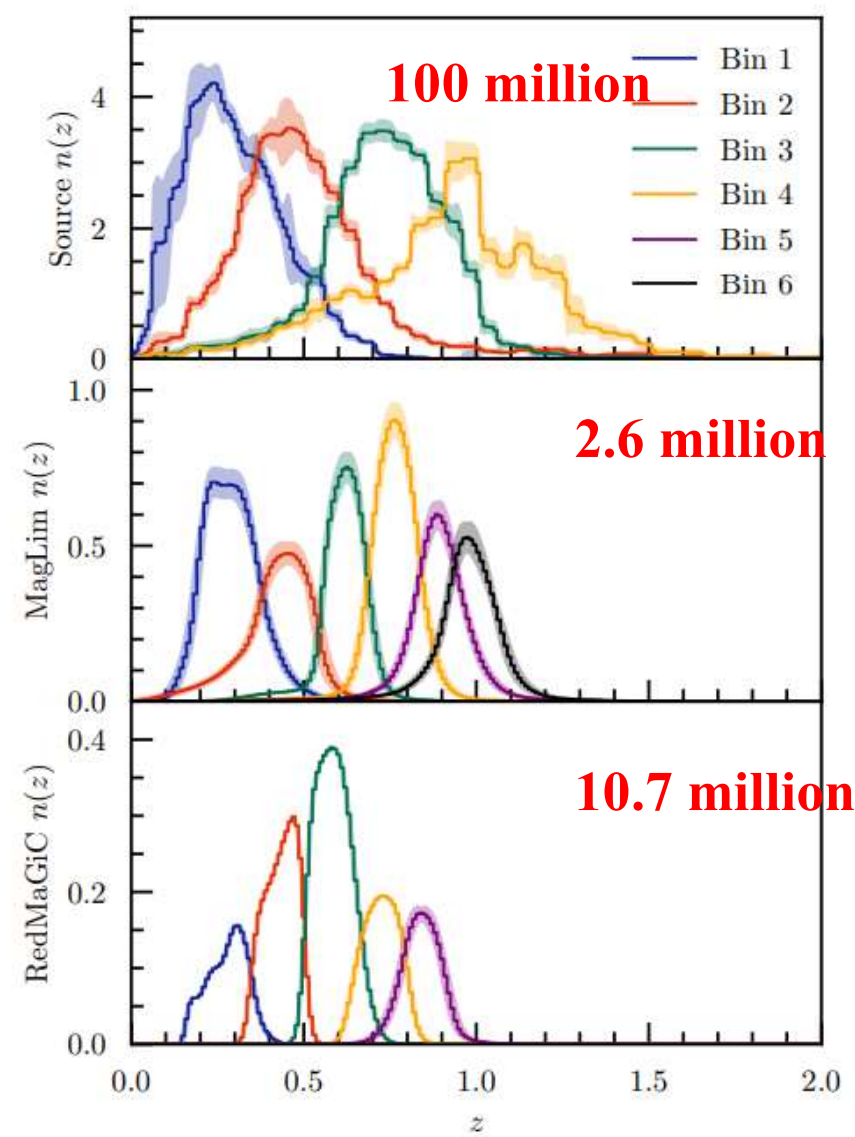
Science Verification: 139 deg²

Y1: 1786 deg² (1321 deg² for 3x2pt cosmological analyses)

Y3: 4946 deg² (4143 deg² for 3x2pt cosmological analyses)

- Gold sample is used to produce science-ready catalogues:
 1. Shear catalog (Metacalibration)
 2. Position catalogs (RedMagic and MagLim)
 3. BAO-optimized catalog

Redshift distributions of galaxies





Legacy Survey of Space and Time

LSST is a 10-year survey to be conducted at the
Vera Rubin Observatory in Chile (CTIO)

using the

Simonyi Survey Telescope – 8.4 meters primary mirror
9.6 deg² field of view

with the

LSSTCam

largest digital camera ever built (SLAC) – 3.2 Gigapixels
189 science CCDs
6 filters: ugrizy

Construction started in 2015

10 years of observations are planned – 2024 to 2034

LSST Science Collaborations

8 Science Collaborations



Active Galactic Nuclei



Dark Energy



Informatics and Statistics



Transients and Variable Stars



Stars, Milky Way, and Local Volume



Strong Lensing



Galaxies



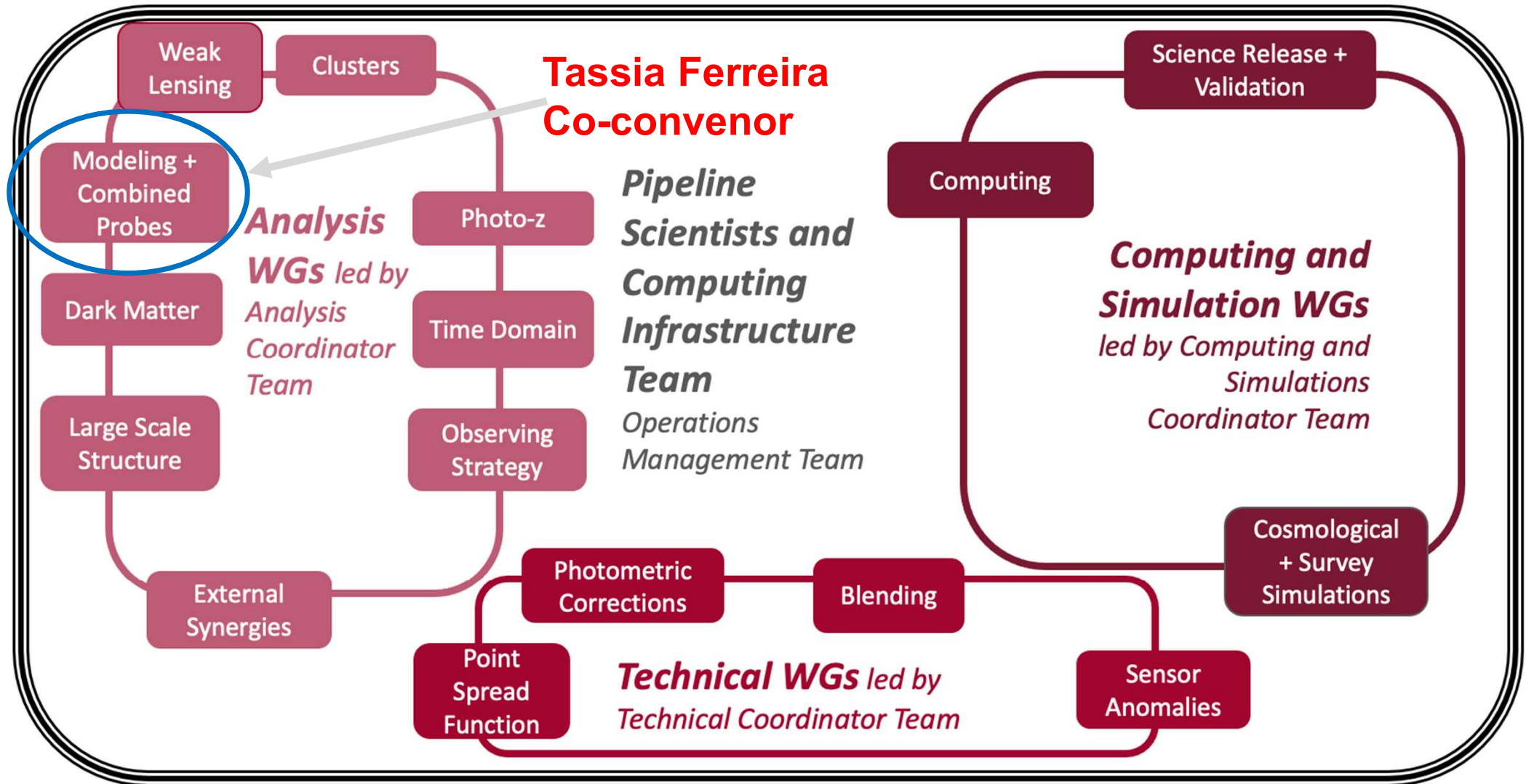
Solar System



1100+ members in 20+ countries,

DESC Management led by
Spokesperson Team

Tassia Ferreira
Co-convenor



Argentina, Brazil, Chile and Mexico participate in LSST

We are planning to organize a LSST@LA series starting next year.

Next lecture: learning about the Universe from photometric surveys.