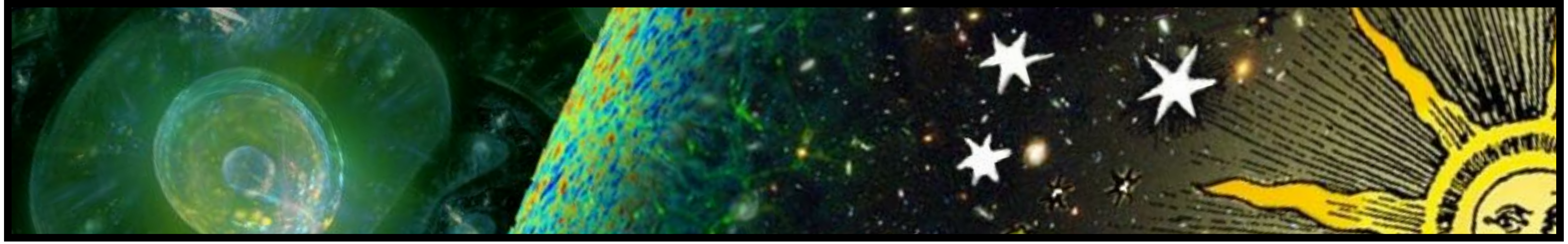
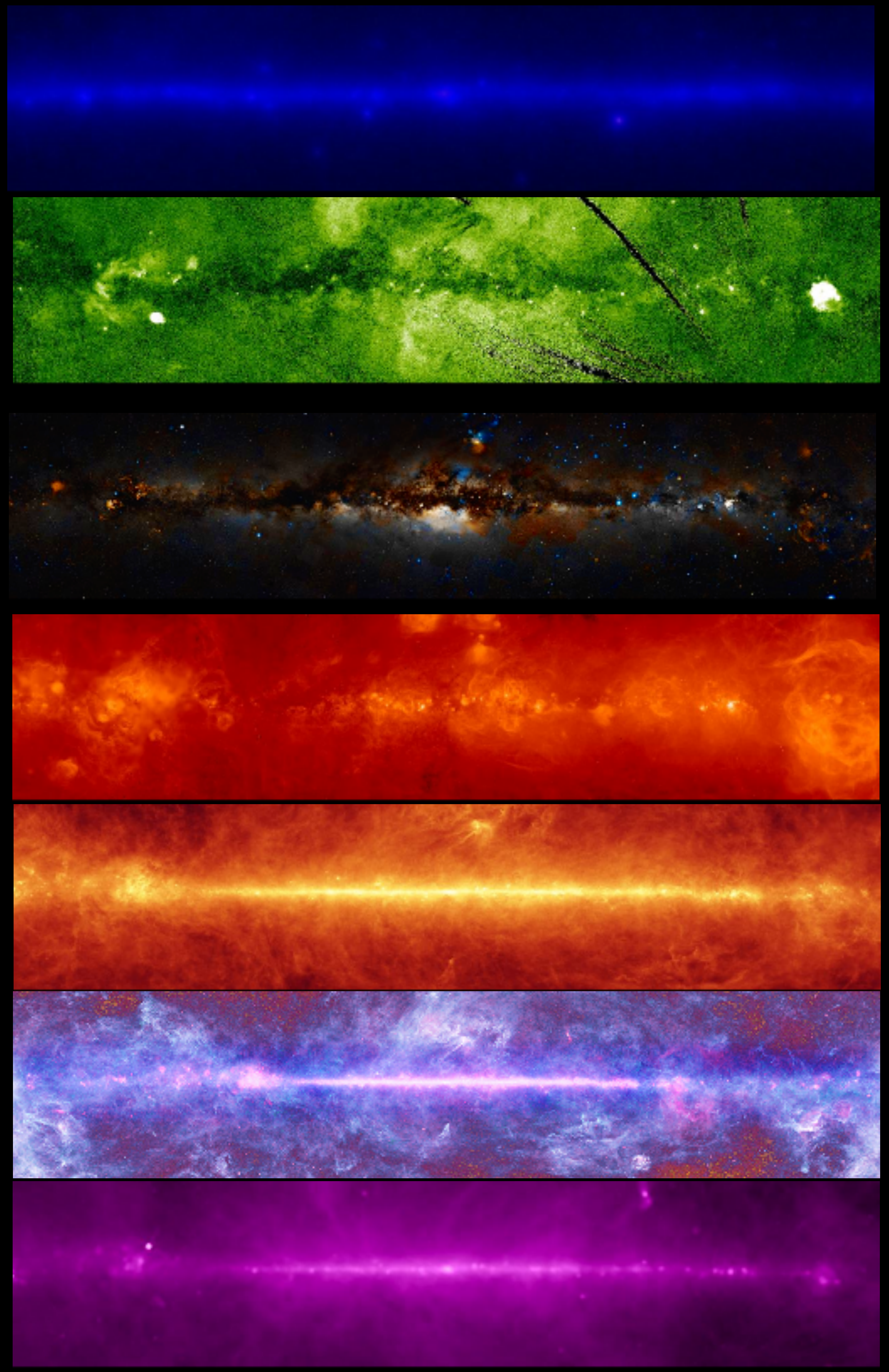


Prospects for understanding the physics of the Universe

Hiranya V. Peiris

UCL and Oskar Klein Centre Stockholm

The era of surveys



Gamma Ray (Fermi)

X Ray (ROSAT)

Optical (DSS)

H-alpha (WHAM/SHASSA/VTSS)

Far Infrared (IRAS)

Microwave (Planck)

Radio (Haslam)

Gravitational Waves (LIGO)

CREDIT: CHROMOSCOPE / CARDIFF

Highlights of era of precision cosmology

- *Determined the basic cosmological model (including measuring the age and composition of the Universe).*

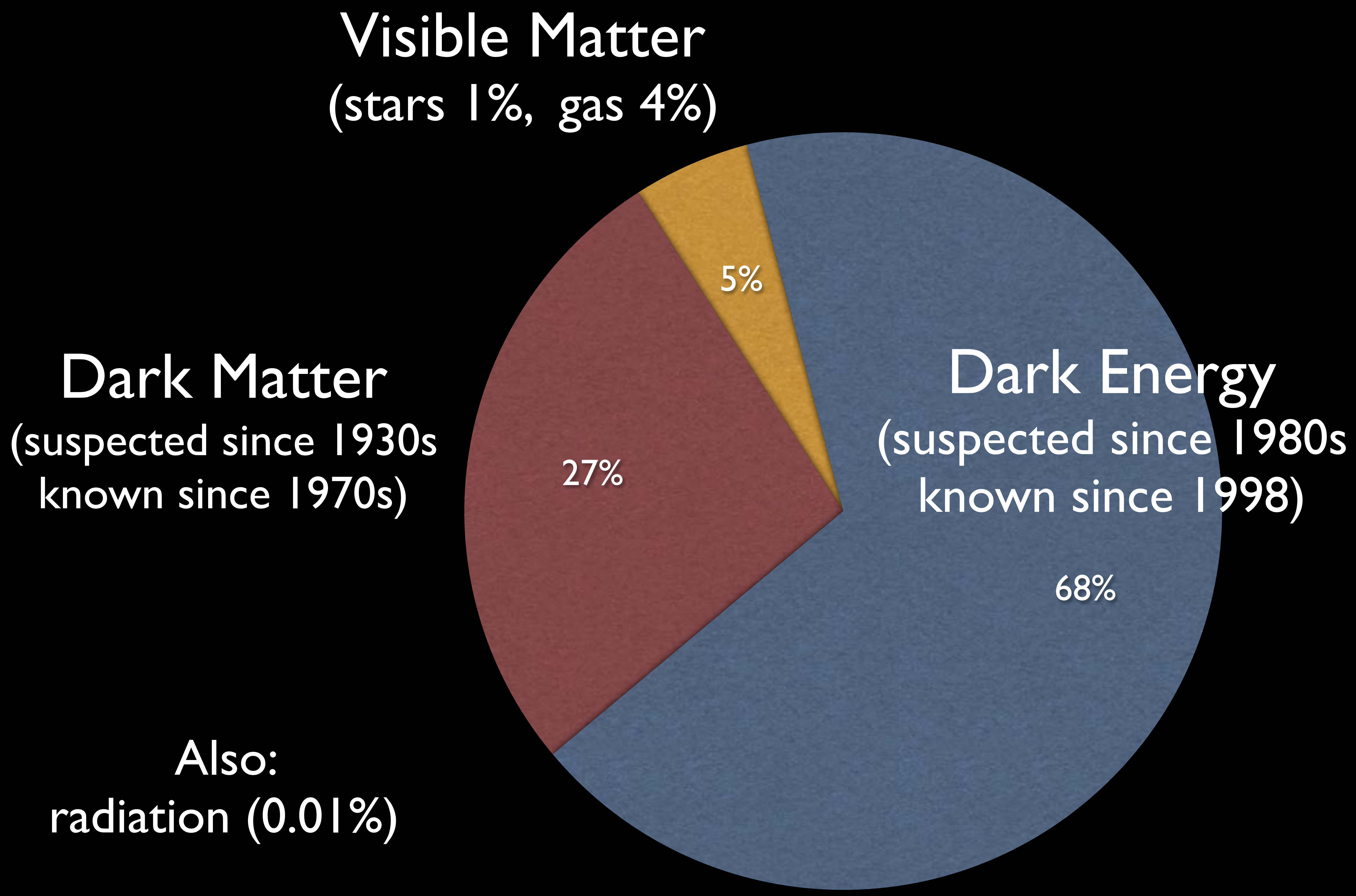
- *Found strong evidence for the quantum origin of cosmic structure.*

- *Measurements of* ***cosmic microwave background (CMB)*** *and* ***supernovae Ia*** *form cornerstones of this achievement.*

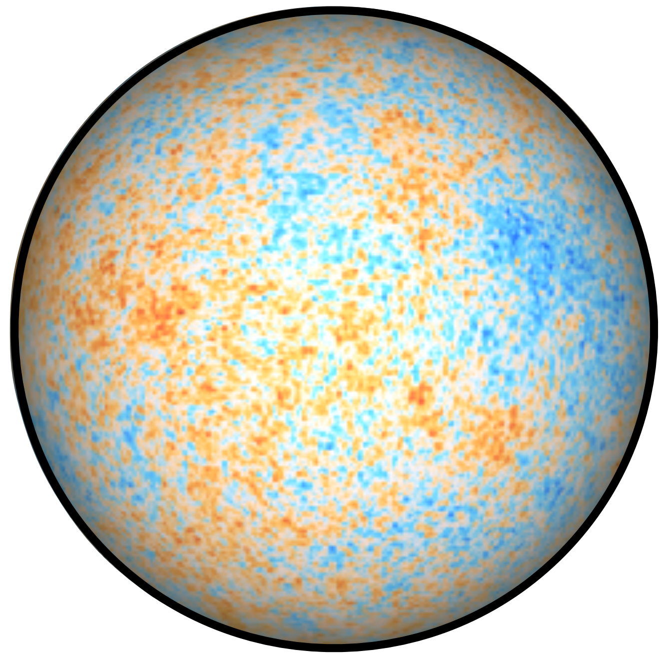
- ***Now + future: progress will come through multiple complementary probes.***

- ***Major theoretical questions remain unanswered.***

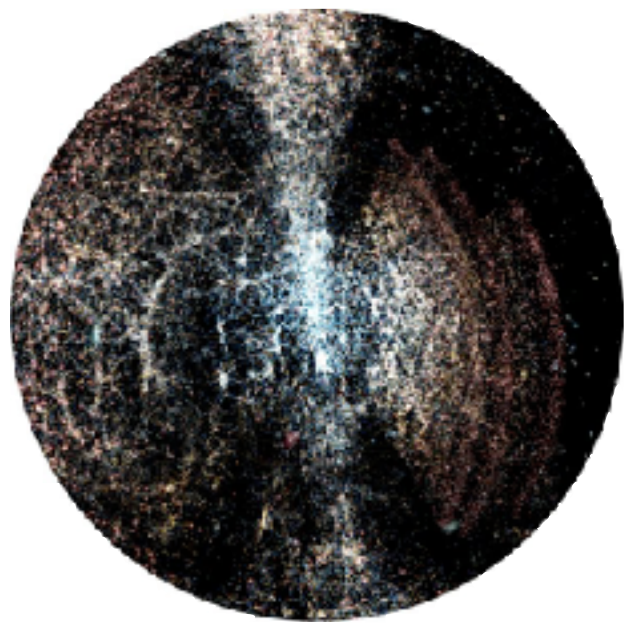
What is Dark Matter? Dark Energy?



Electromagnetic cosmological probes in the next decade



Cosmic Microwave Background



Large Scale Structure

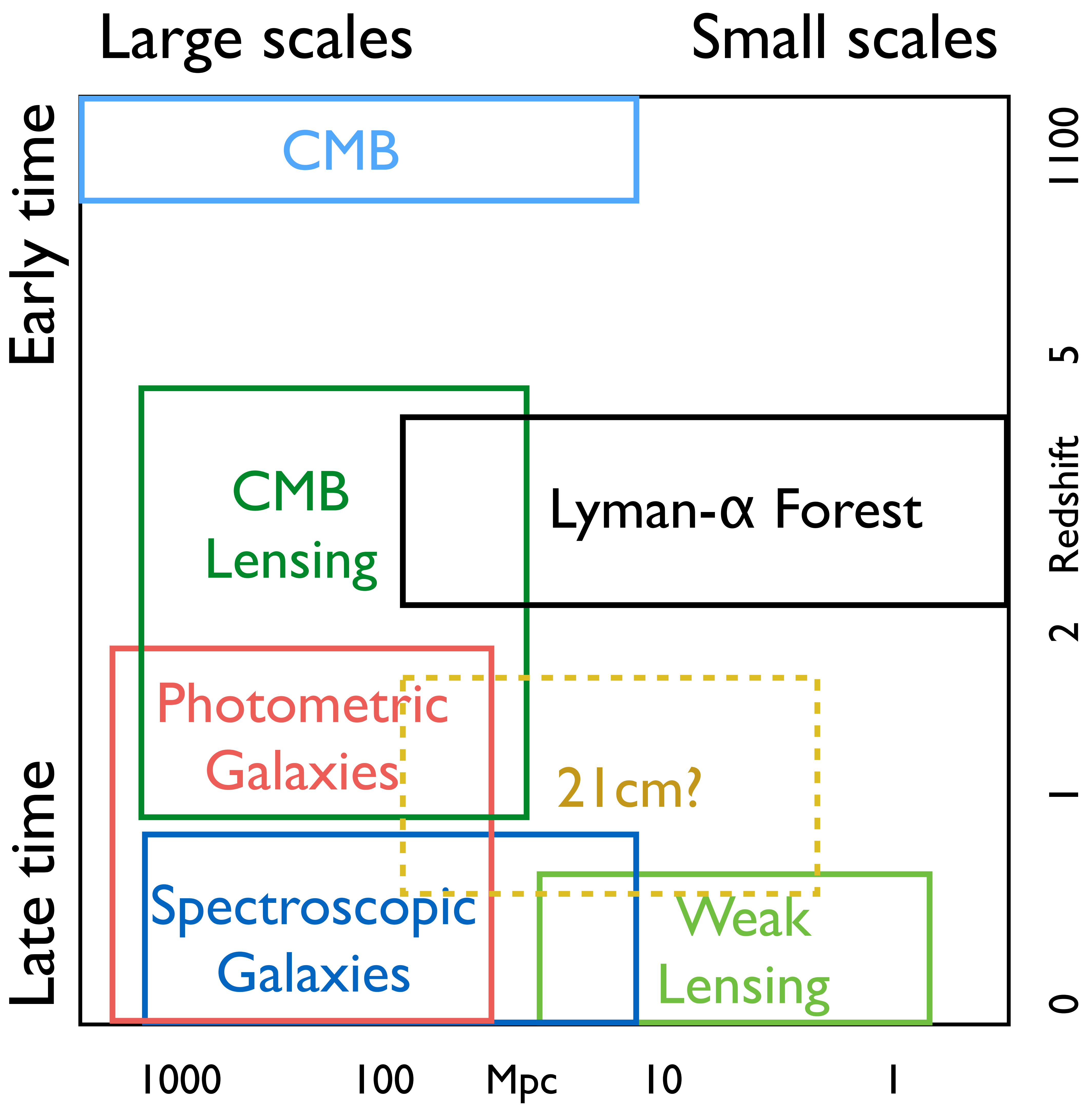


Figure: Andreu Font-Ribera

Legacy Survey of Space and Time

VERA C. RUBIN OBSERVATORY

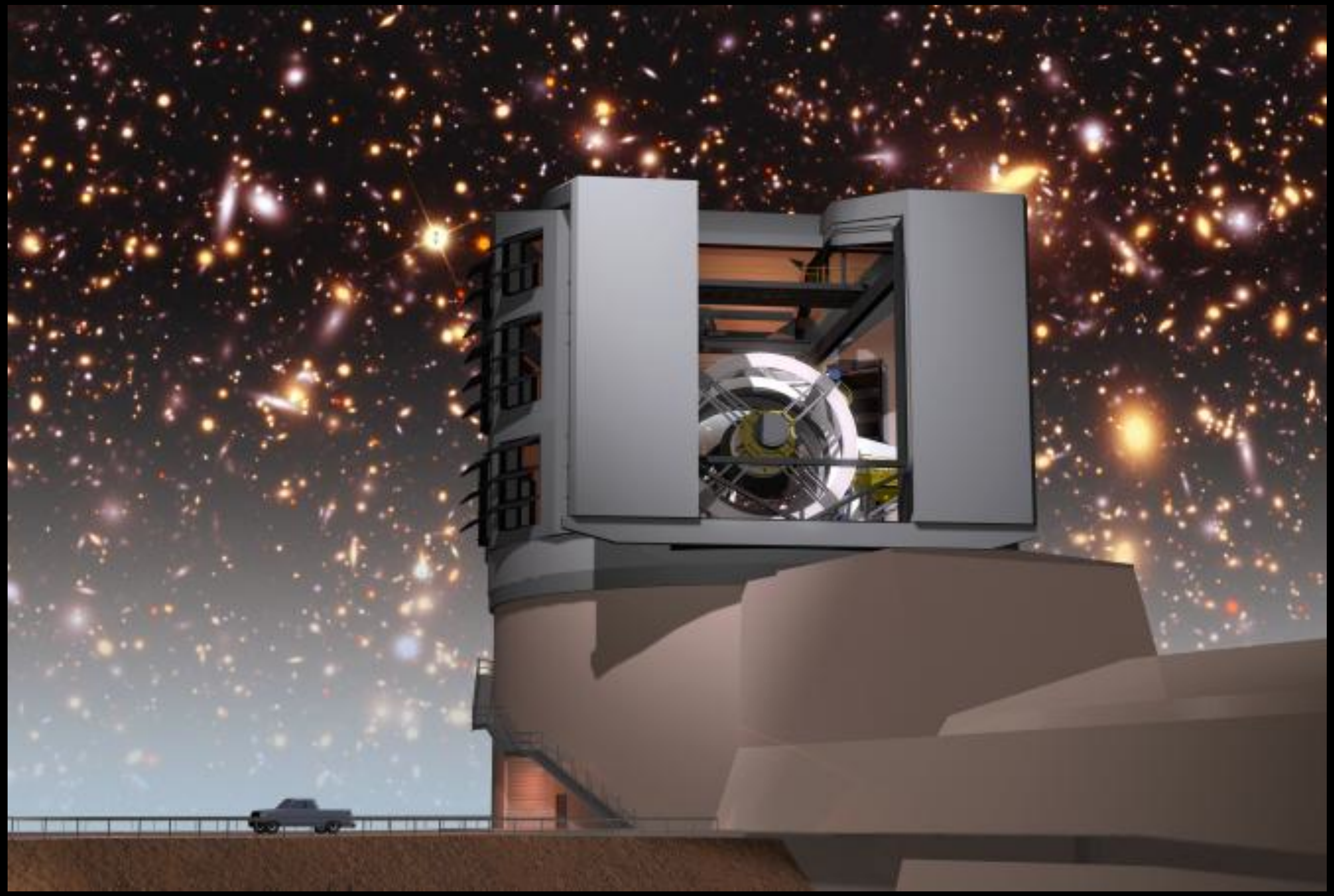
Rubin's Legacy Survey of Space and Time (LSST)

- Wide (half-sky), deep (24-27 mag), fast (every ~3 days) images
- Beginning in 2024, LSST will survey the Southern sky for 10 years
- Expand space-time volume 1000 times over current surveys

VERA C. RUBIN
OBSERVATORY

LSST
Legacy Survey of Space and Time

LSST: survey of 18,000 sq deg
(half the sky)



37 billion objects in space and time
30 trillion measurements
60 PB raw data (20 TB/night)

Dark matter-Dark energy

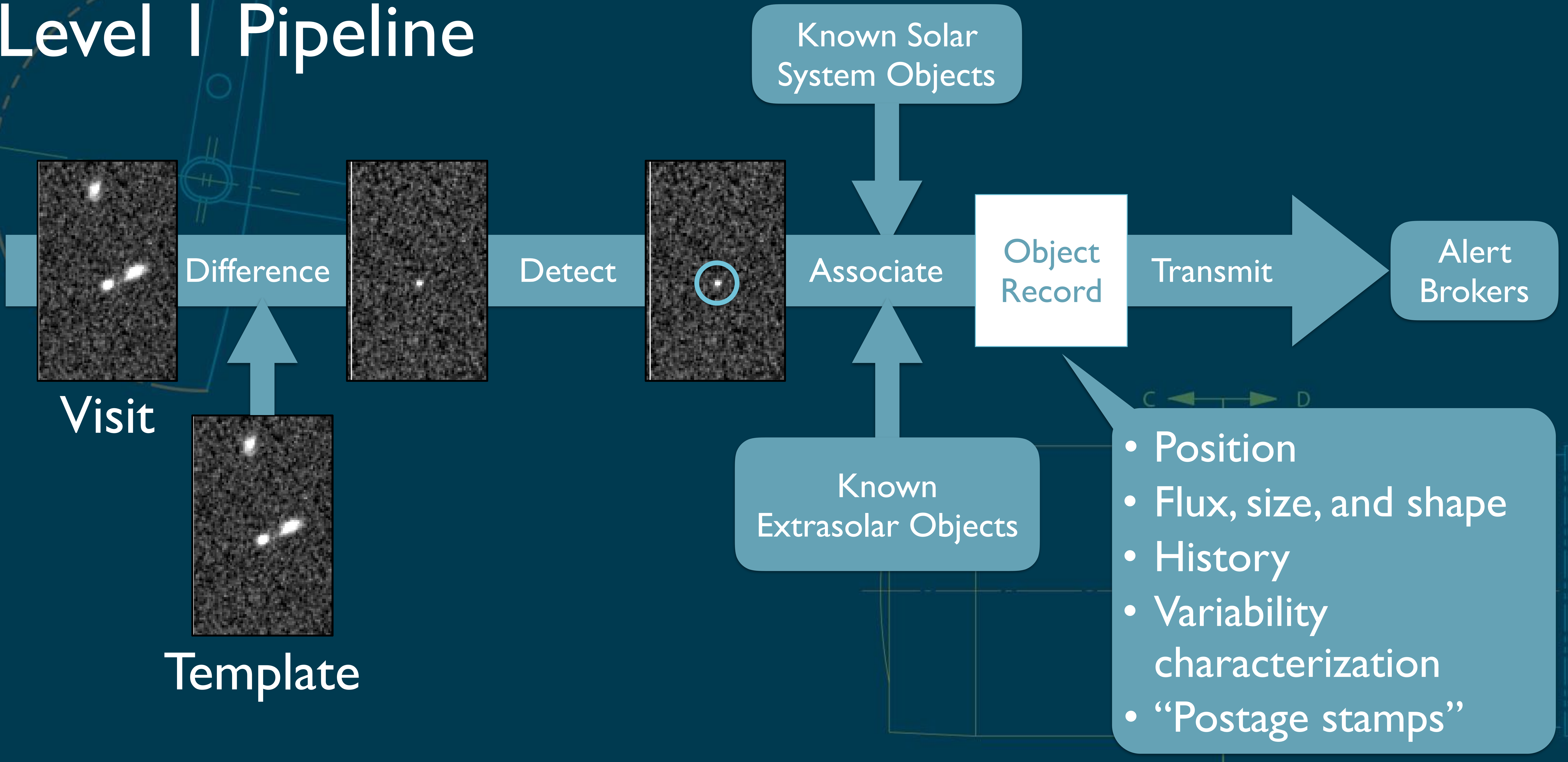
Solar system inventory

"Movie of the Universe"

Mapping the Milky Way

Slide adapted from Ian Shipsey

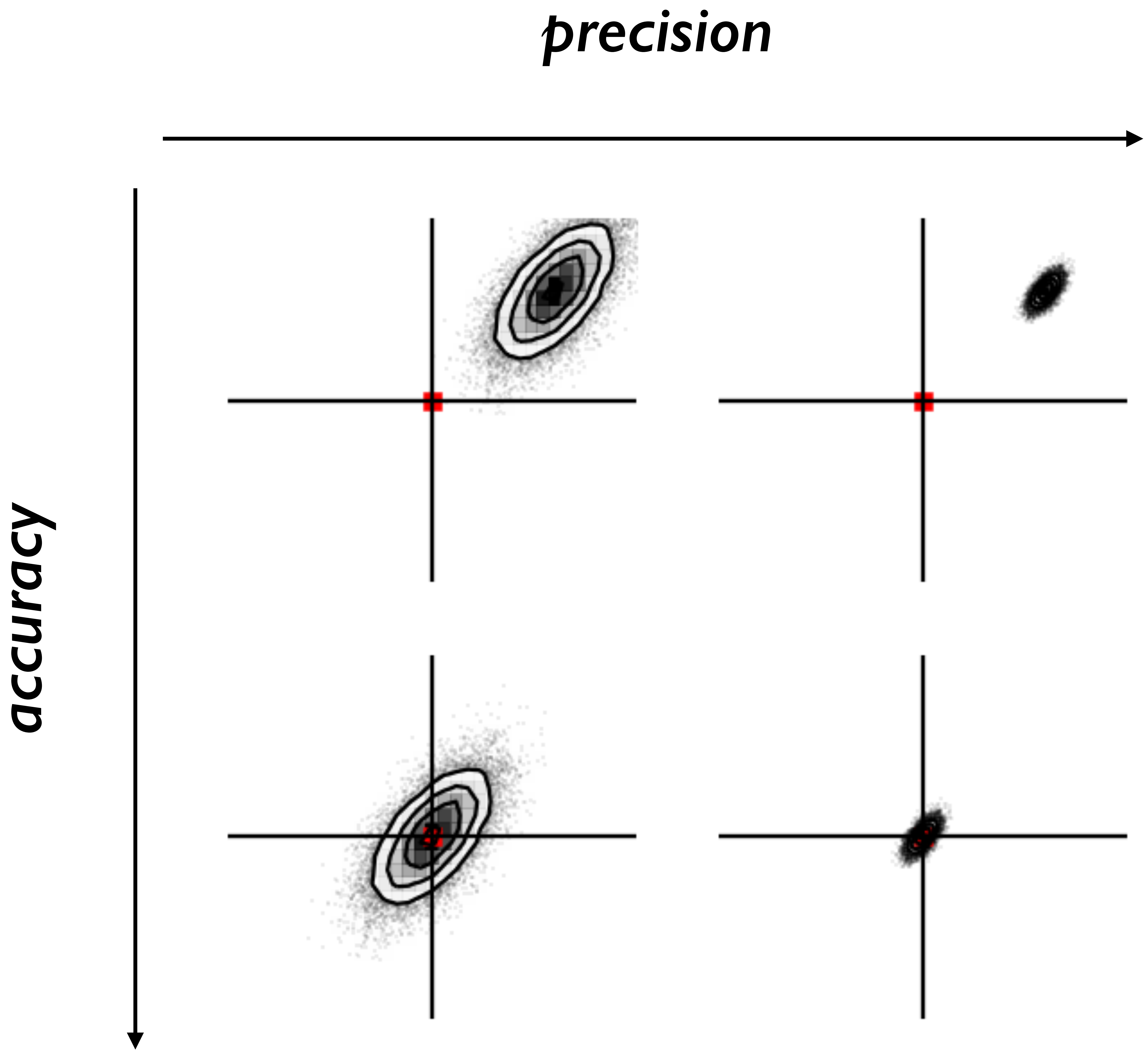
Level 1 Pipeline



0 seconds

60 seconds

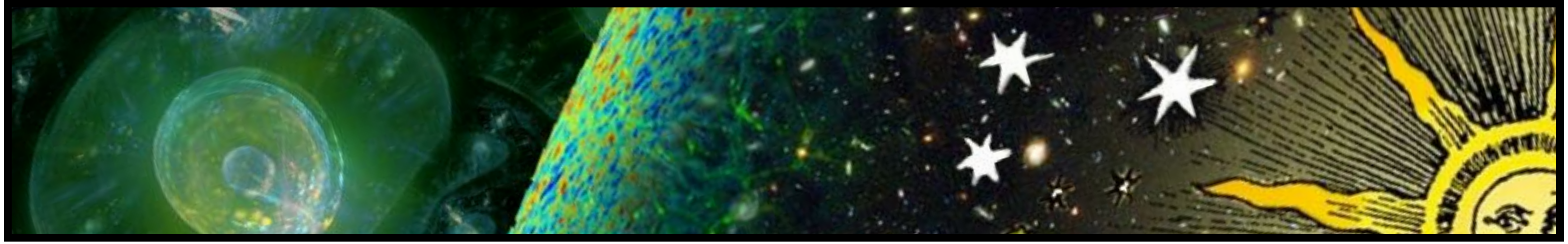
Meredith Rawls • @merrdiff

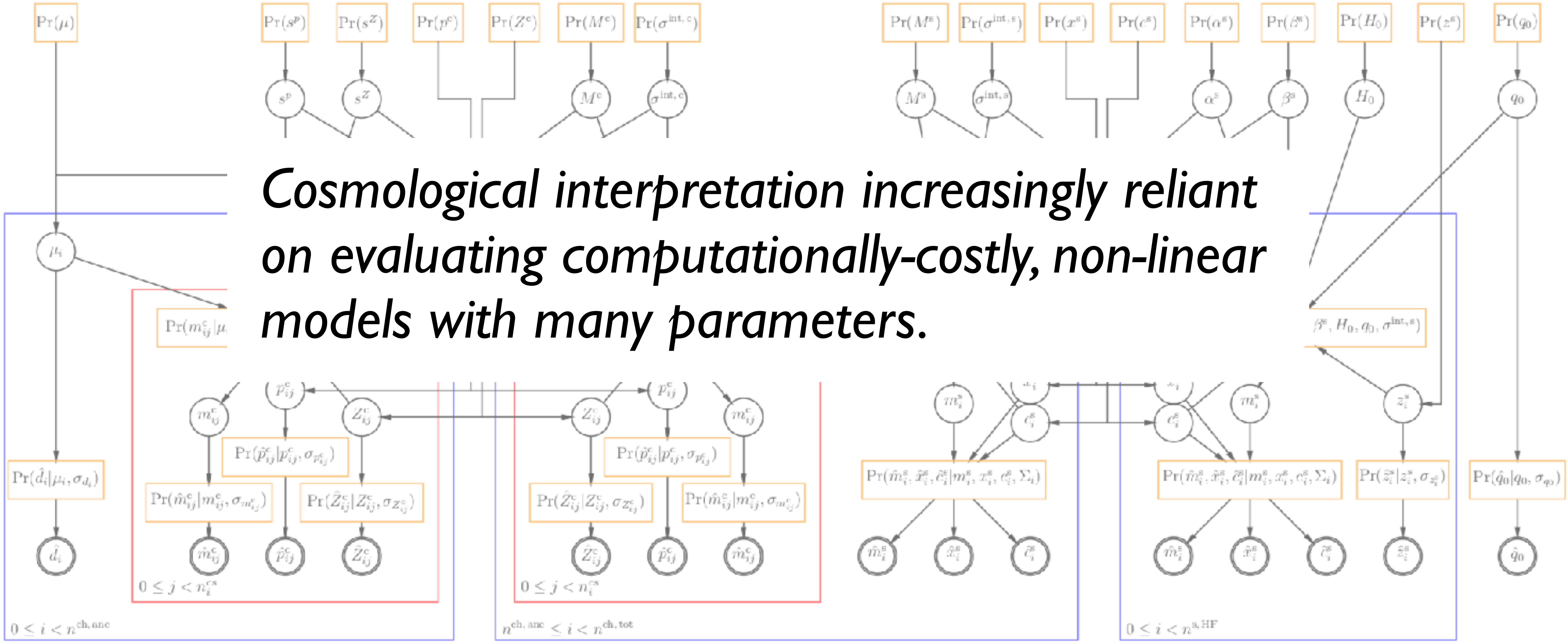

precision
accuracy

How should we compare

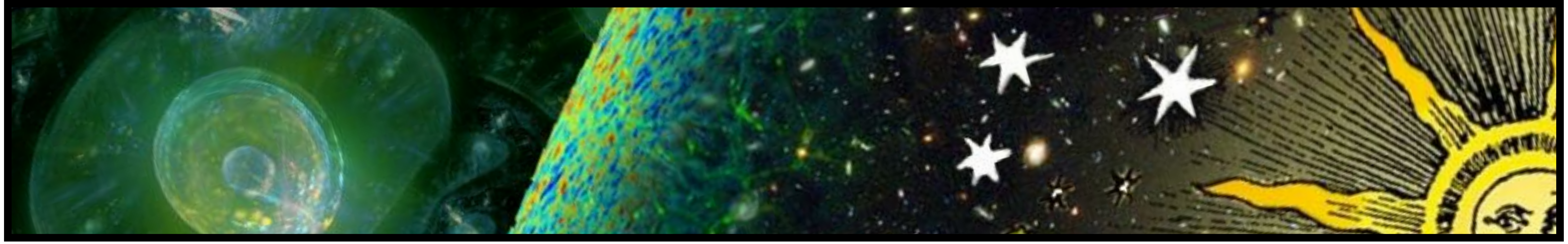
vs

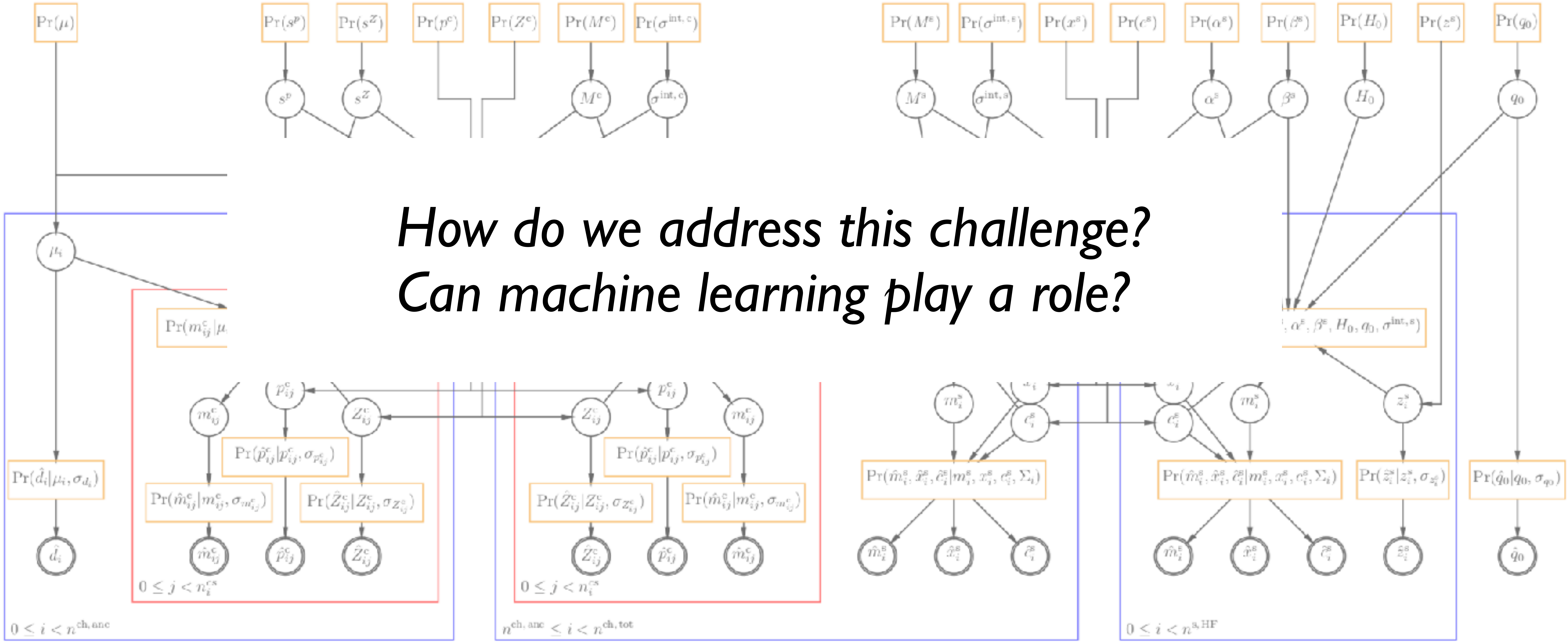
Theory **Data?**





Cosmological interpretation increasingly reliant on evaluating computationally-costly, non-linear models with many parameters.





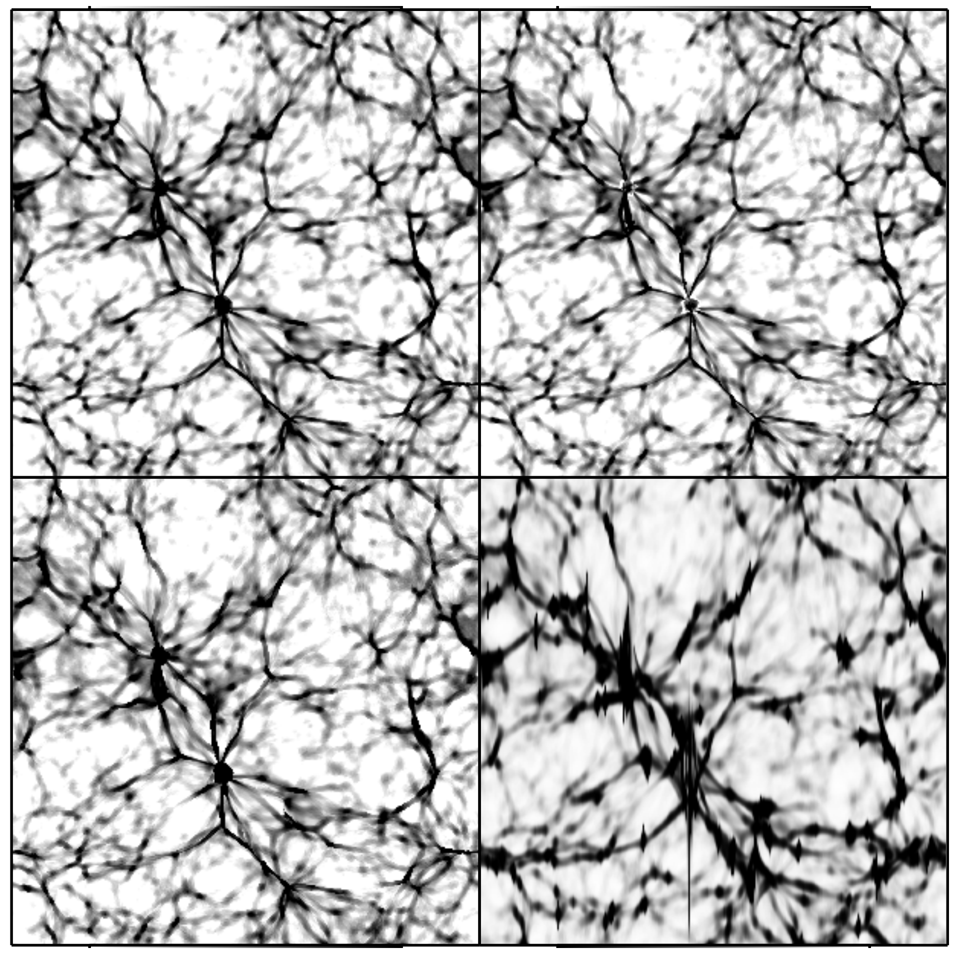
How do we address this challenge?
Can machine learning play a role?

Desiderata for solving cosmological modelling challenges with machine learning

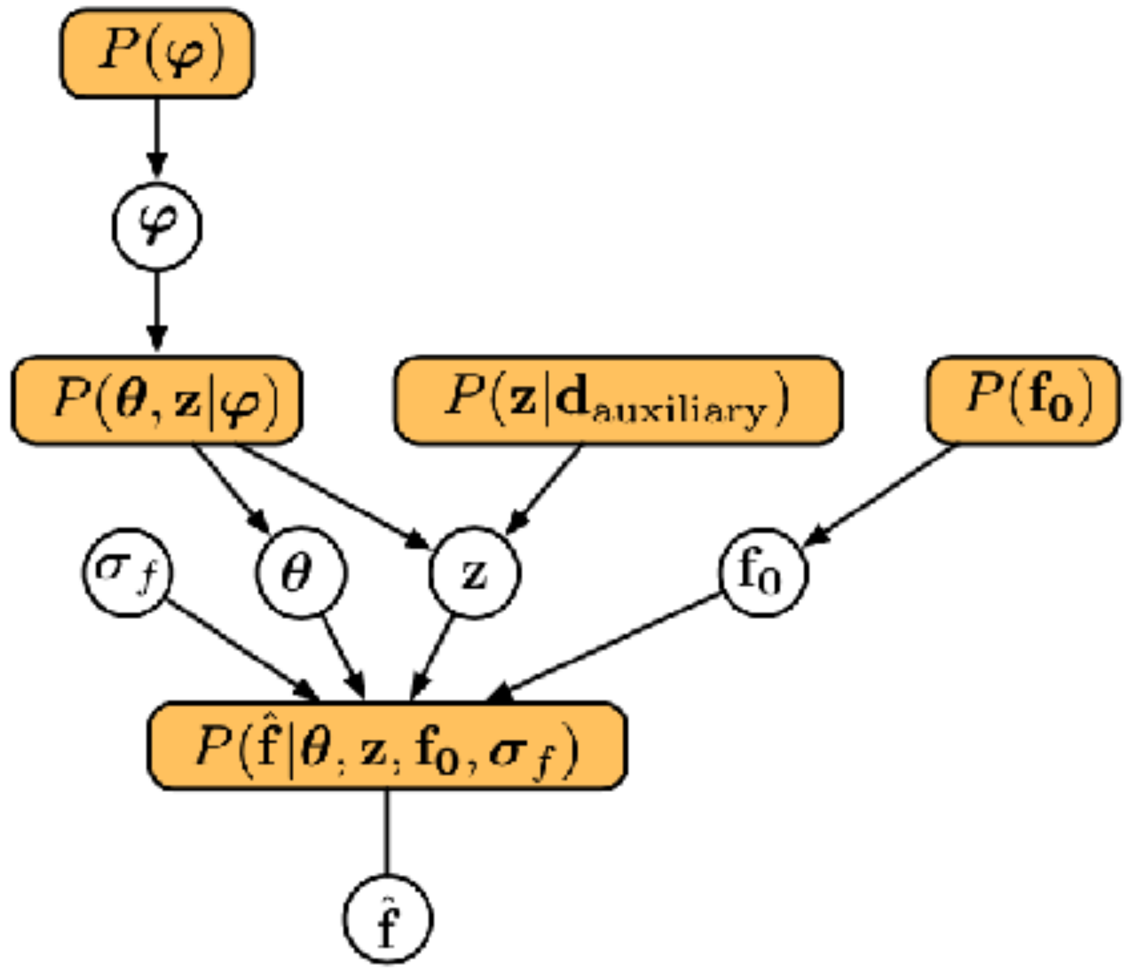
(i) interpretability: *account for why ML system reaches particular decision or prediction;*

(ii) explainability: *map this account onto existing knowledge in relevant science domain.*

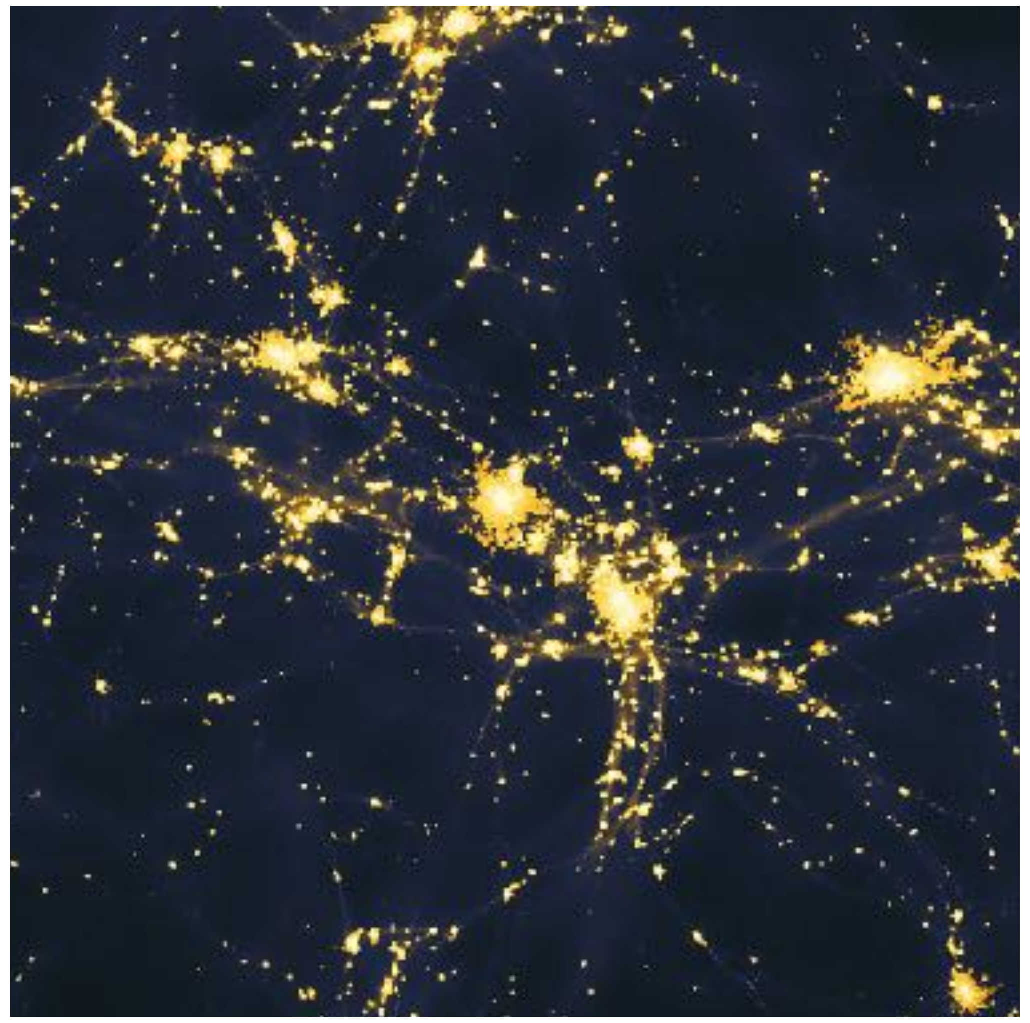
- *Currently challenging because of "black box" nature of powerful ML architectures.*



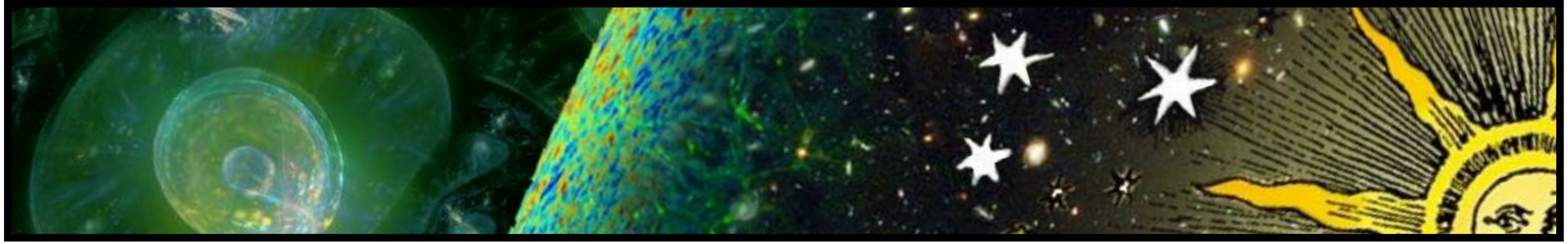
1. ML-accelerated emulation of observations



2. high dimensional cosmological inference with ML-accelerated parts



3. AI-enabled knowledge extraction about cosmological structure formation



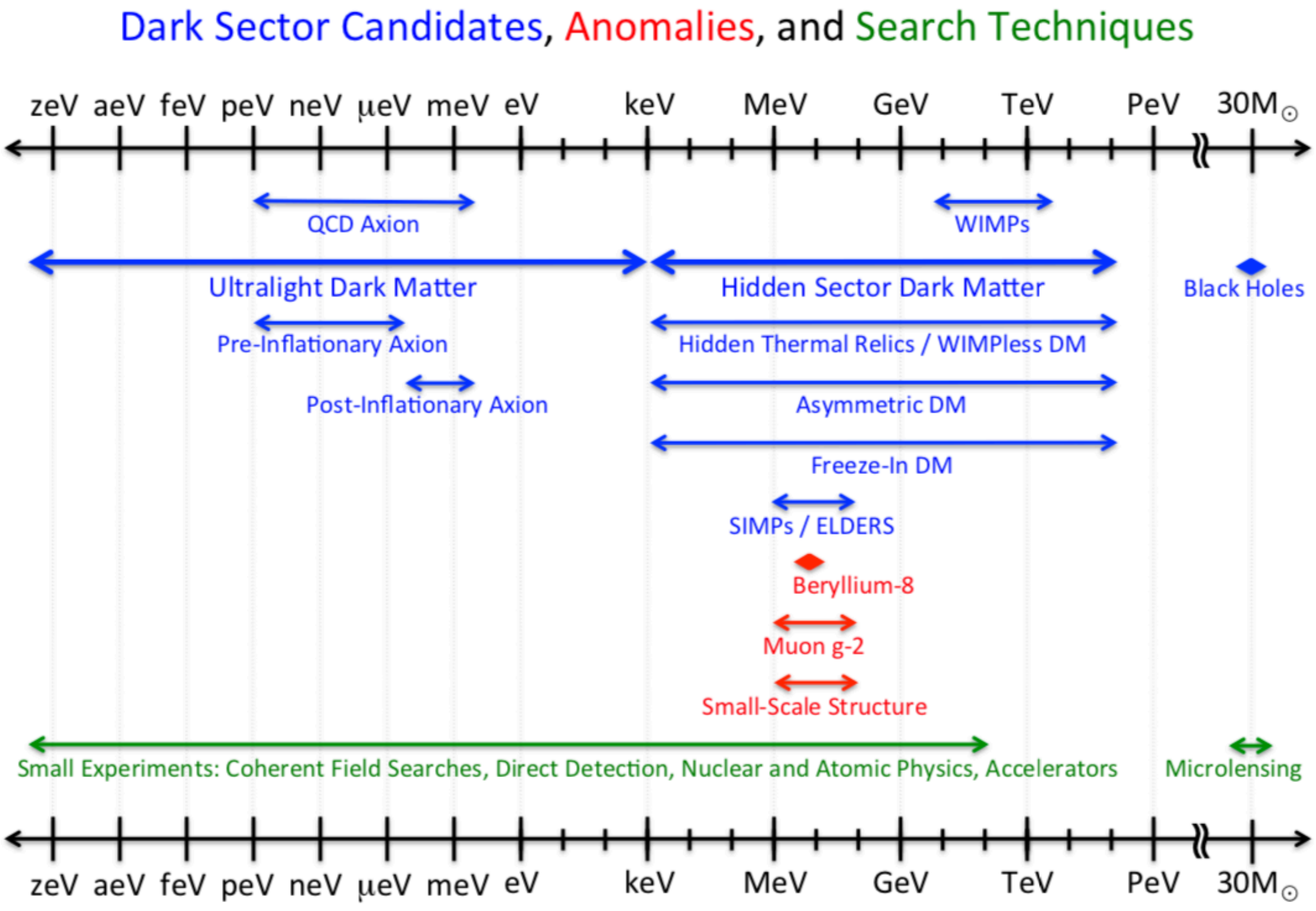
Efficient emulation of cosmological simulations

Keir Rogers
(Dunlap/Toronto)

Cora Dvorkin
(Harvard)

With: Andrew Pontzen, Simeon Bird, Andreu Font-Ribera, Licia Verde

What does the dark matter consist of?



Constraining dark matter with cosmology

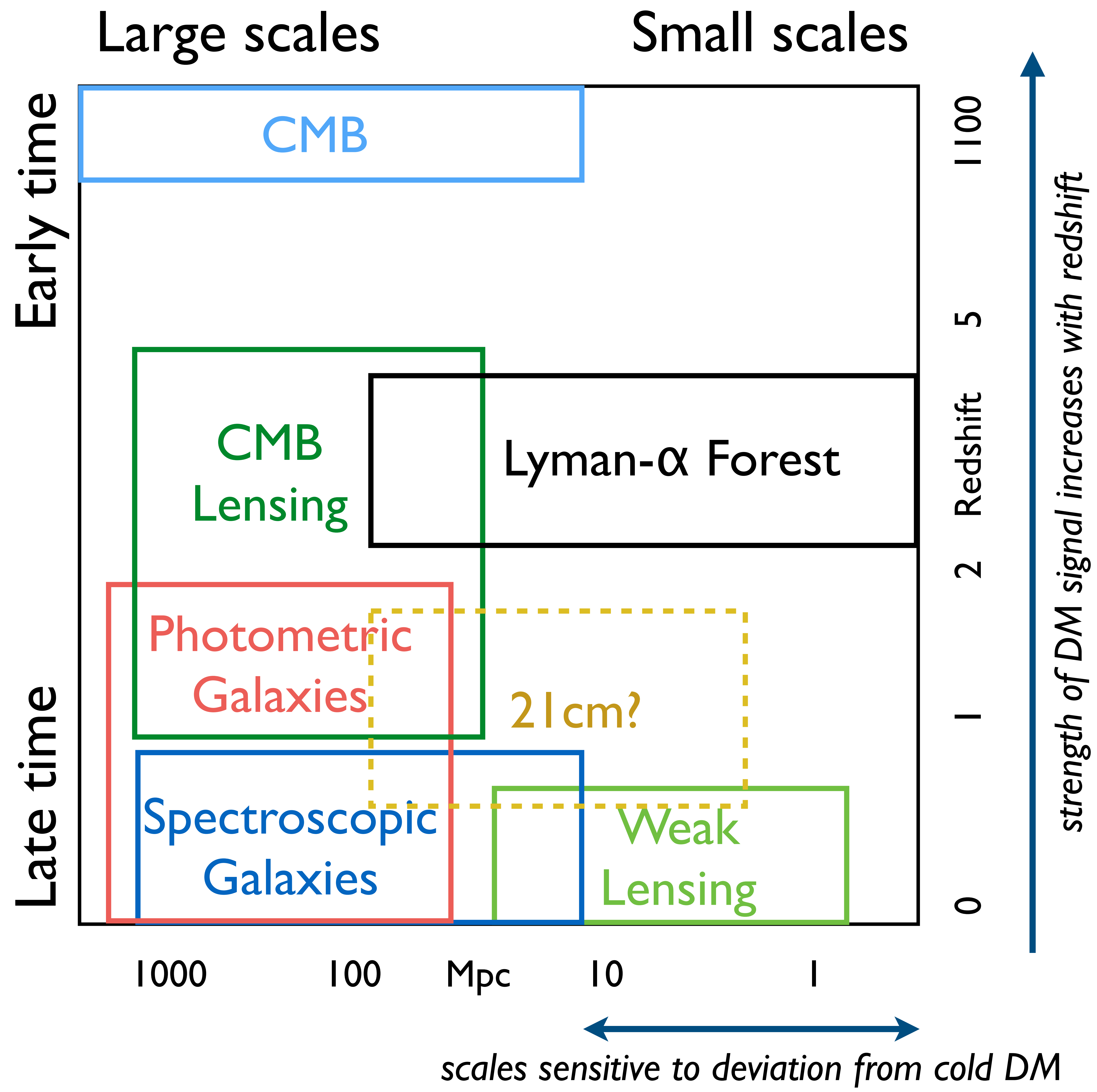
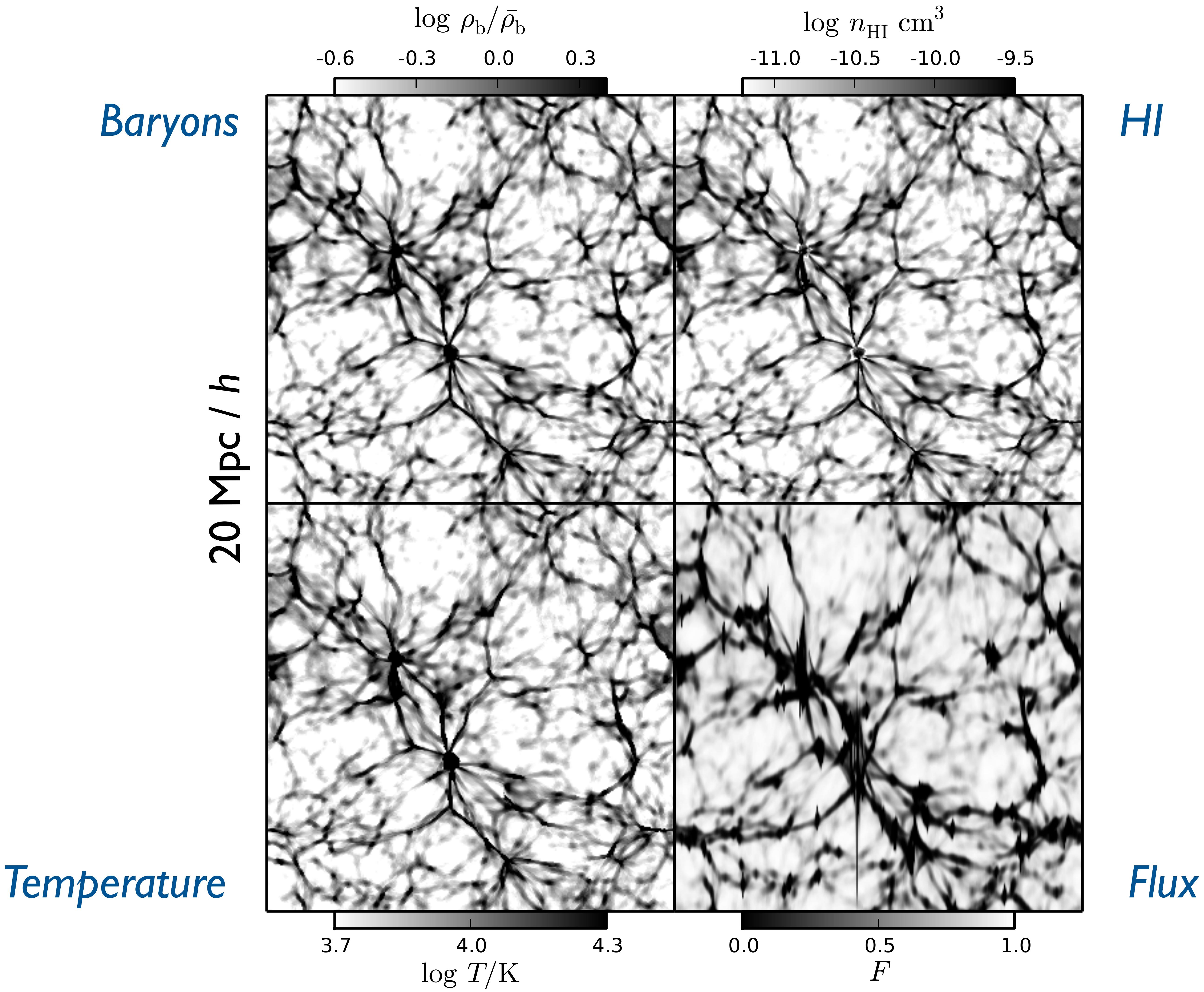


Figure: Andreu Font-Ribera

Lyman-alpha forest flux:
biased, redshift-space distorted tracer of cosmic density field



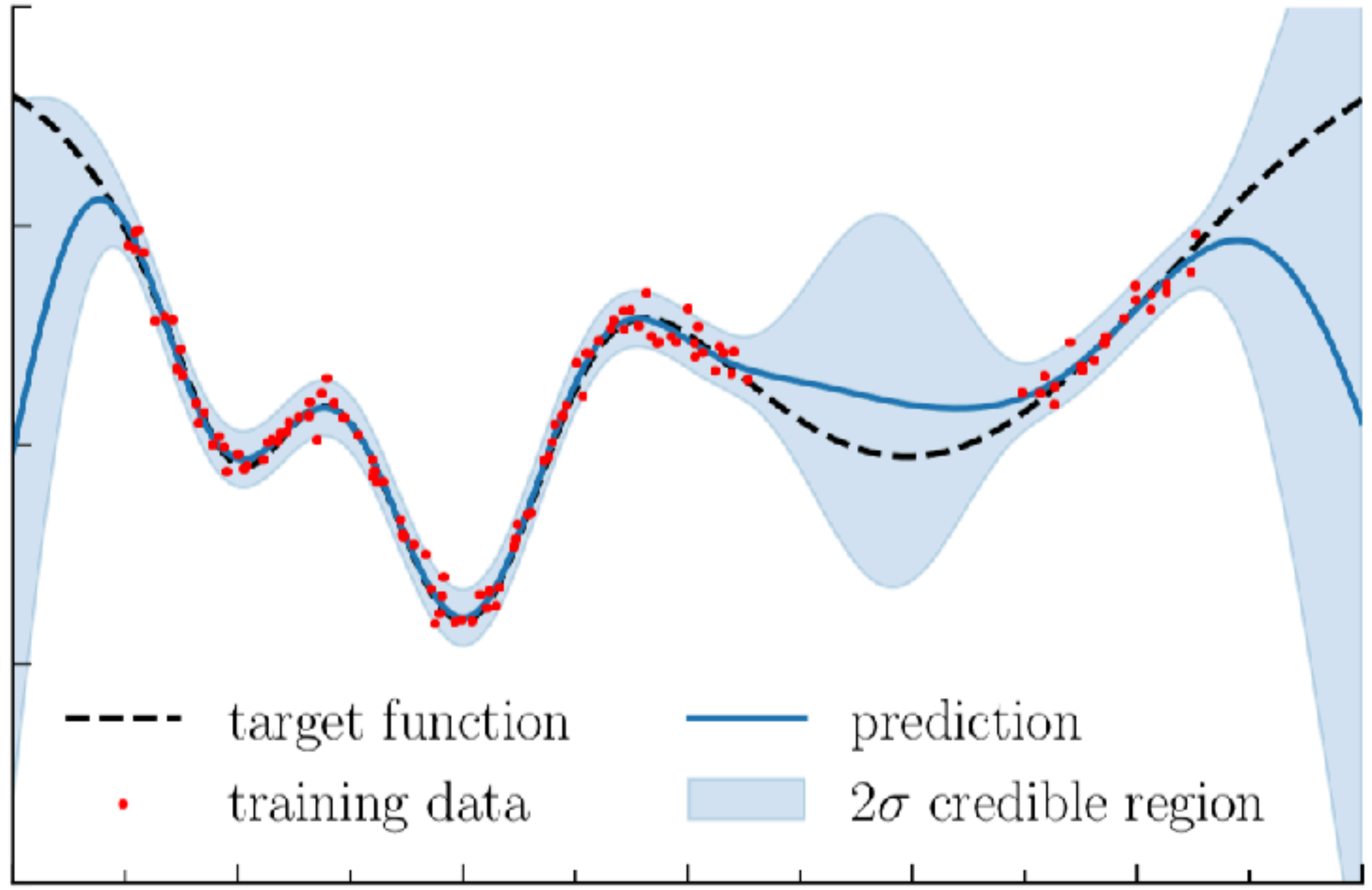
Accurate modelling requires intensive simulations:

~3000 CPU-hrs per parameter combination in 12-D parameter space

Figure: Lukić et al. (2015)

Gaussian process for emulating high-dimensional models

- *Smooth interpolation scheme that gives tight constraints where there are training points and broad constraints where there are none*



$$f(\mathbf{x}) \sim \mathcal{N}(0, K(\mathbf{x}, \mathbf{x}'; \theta))$$

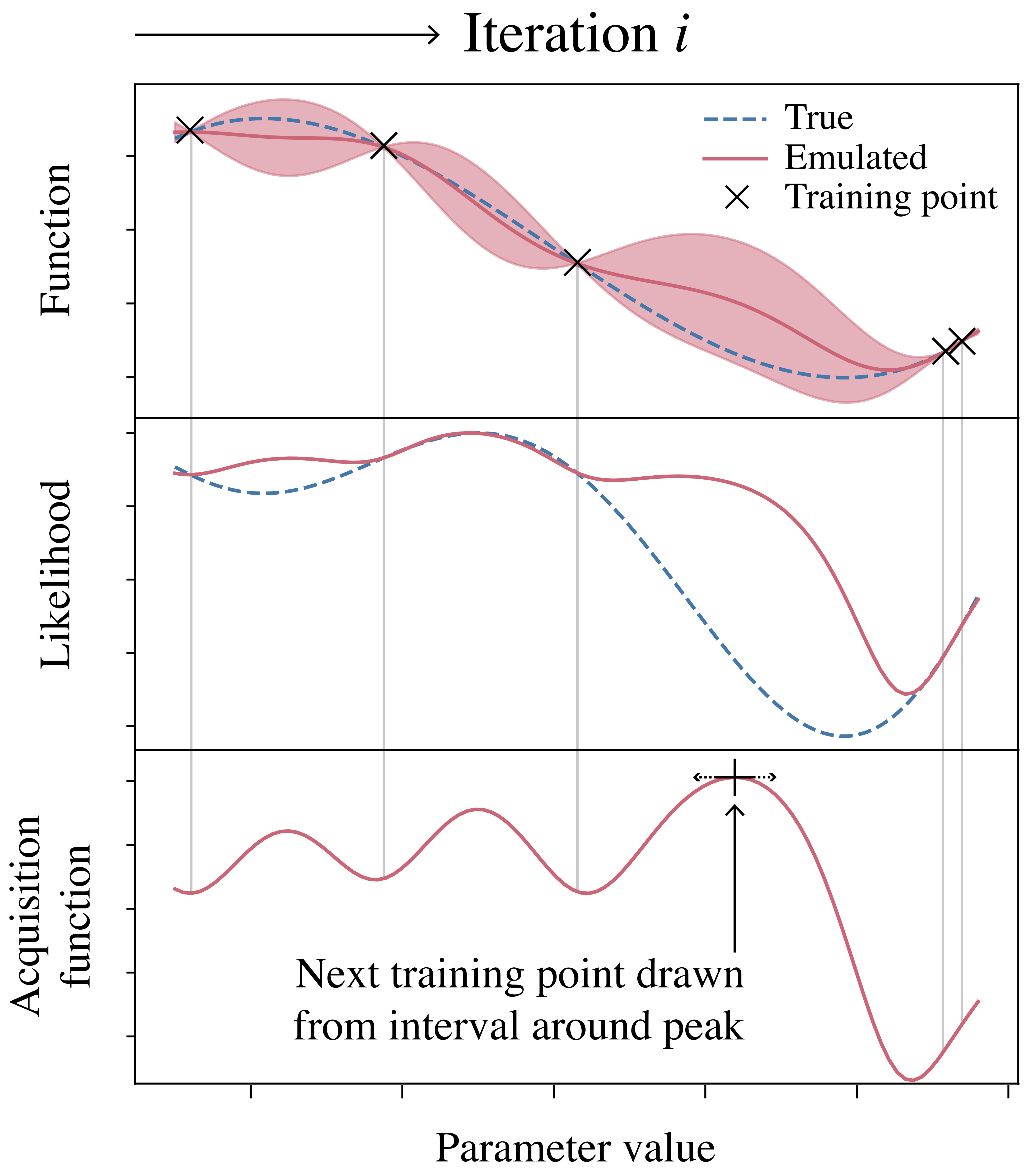
Simulation output

Simulation parameters

Kernel hyperparameters (covariance model)

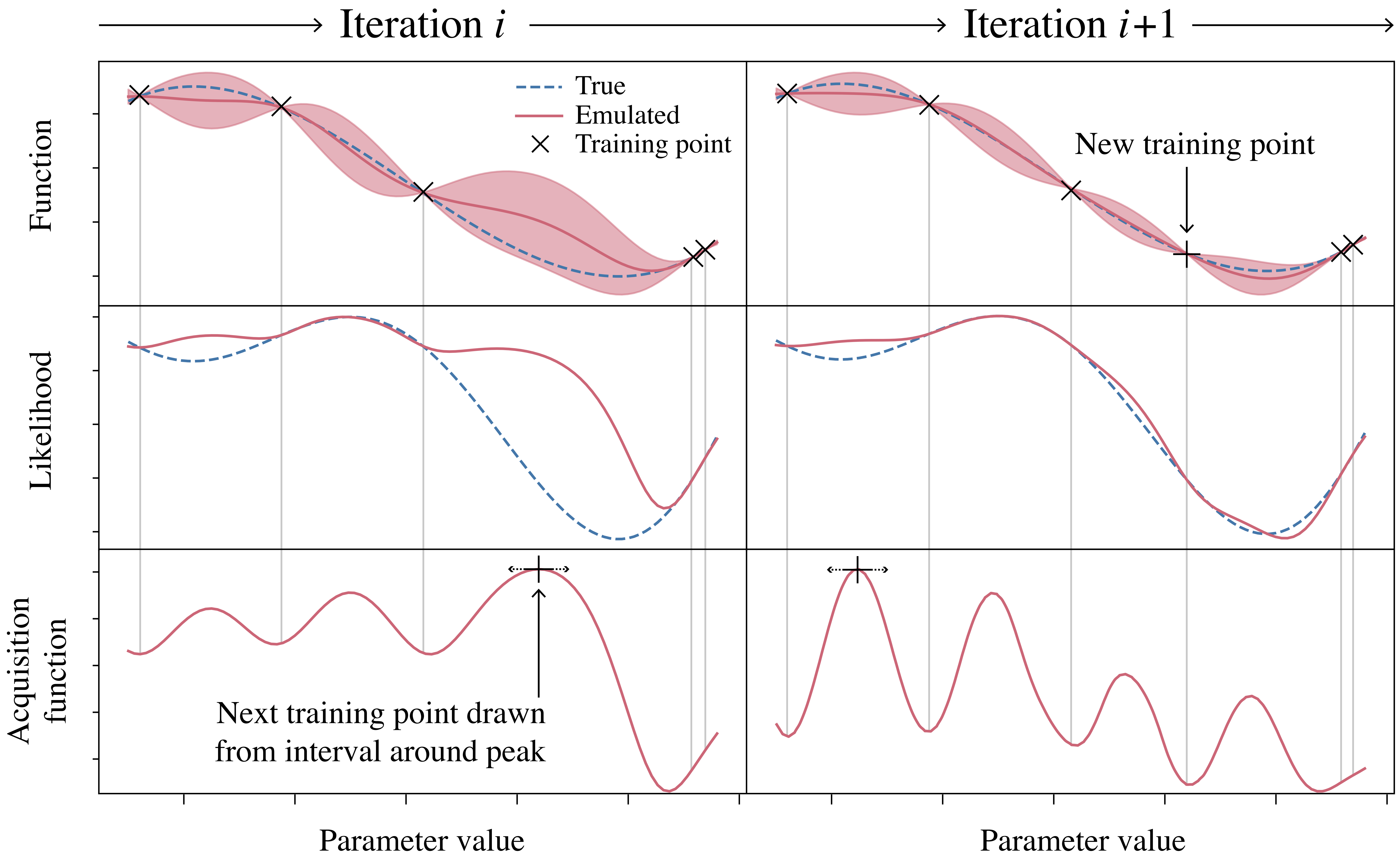
Figure: Leclercq (2018)

Key idea: active learning via Bayesian optimisation



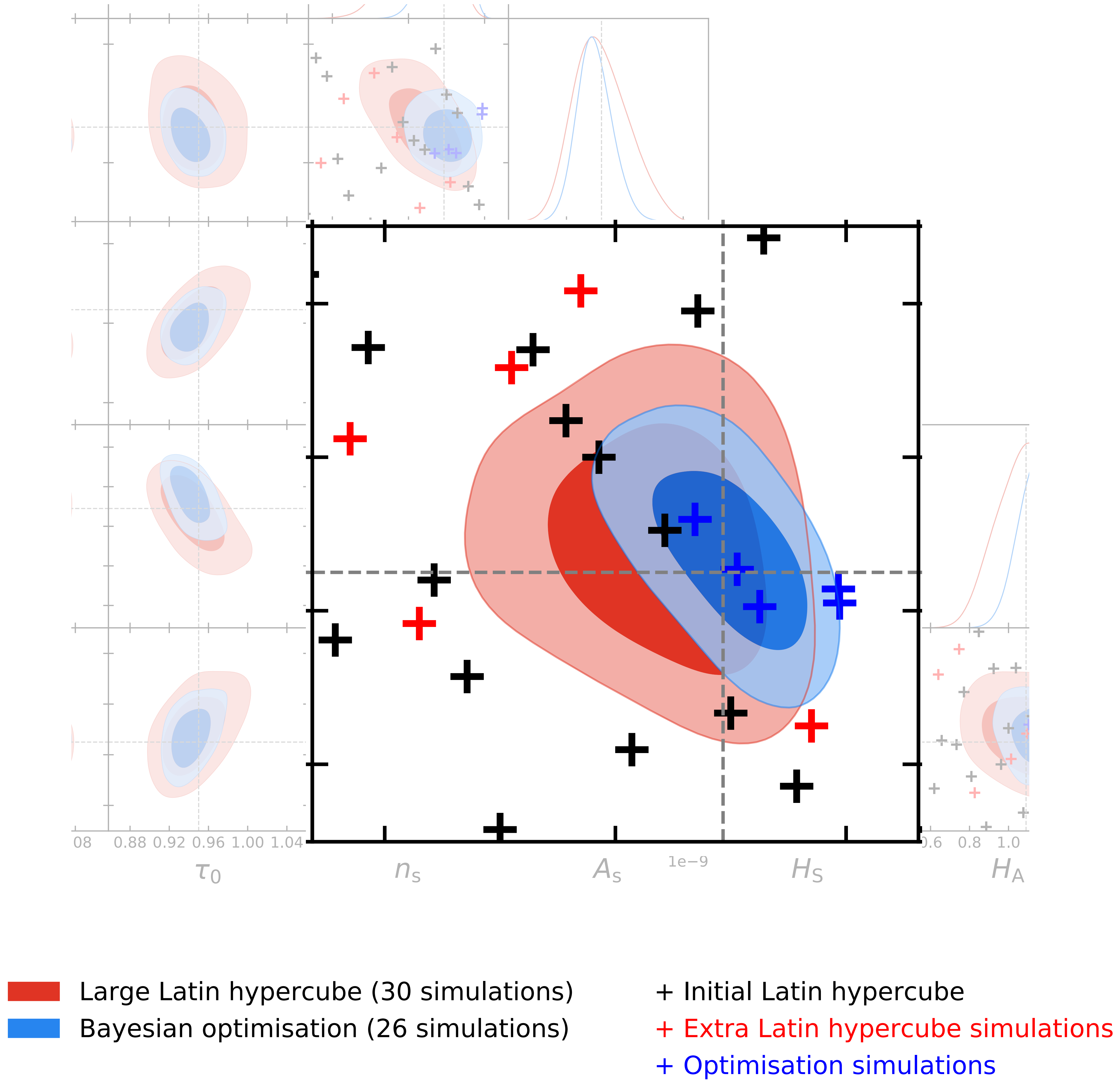
Rogers & Peiris (PRL, PRD, 2021a,b), Rogers, Peiris, Pontzen, Bird, Verde & Font-Ribera (JCAP, 2019), Leclercq (PRD, 2018)

Key idea: active learning via Bayesian optimisation



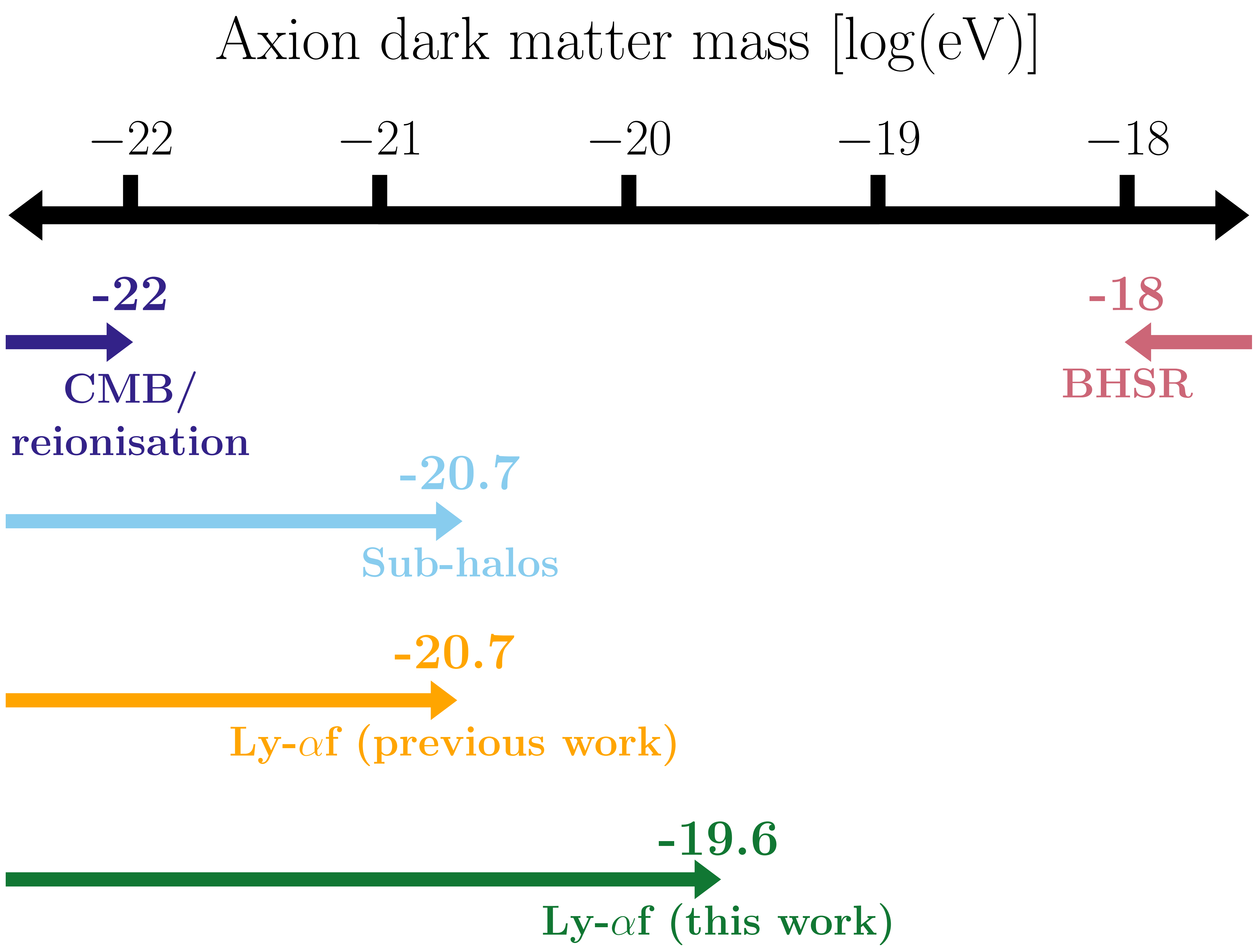
Rogers & Peiris (PRL, PRD, 2021a,b), Rogers, Peiris, Pontzen, Bird, Verde & Font-Ribera (JCAP, 2019), Leclercq (PRD, 2018)

Key idea: active learning via Bayesian optimisation



Rogers & Peiris (PRL, PRD, 2021a,b), Rogers, Peiris, Pontzen, Bird, Verde & Font-Ribera (JCAP, 2019), Leclercq (PRD, 2018)

“Canonical” 10^{-22} - 10^{-21} eV ULA dark matter strongly disfavoured

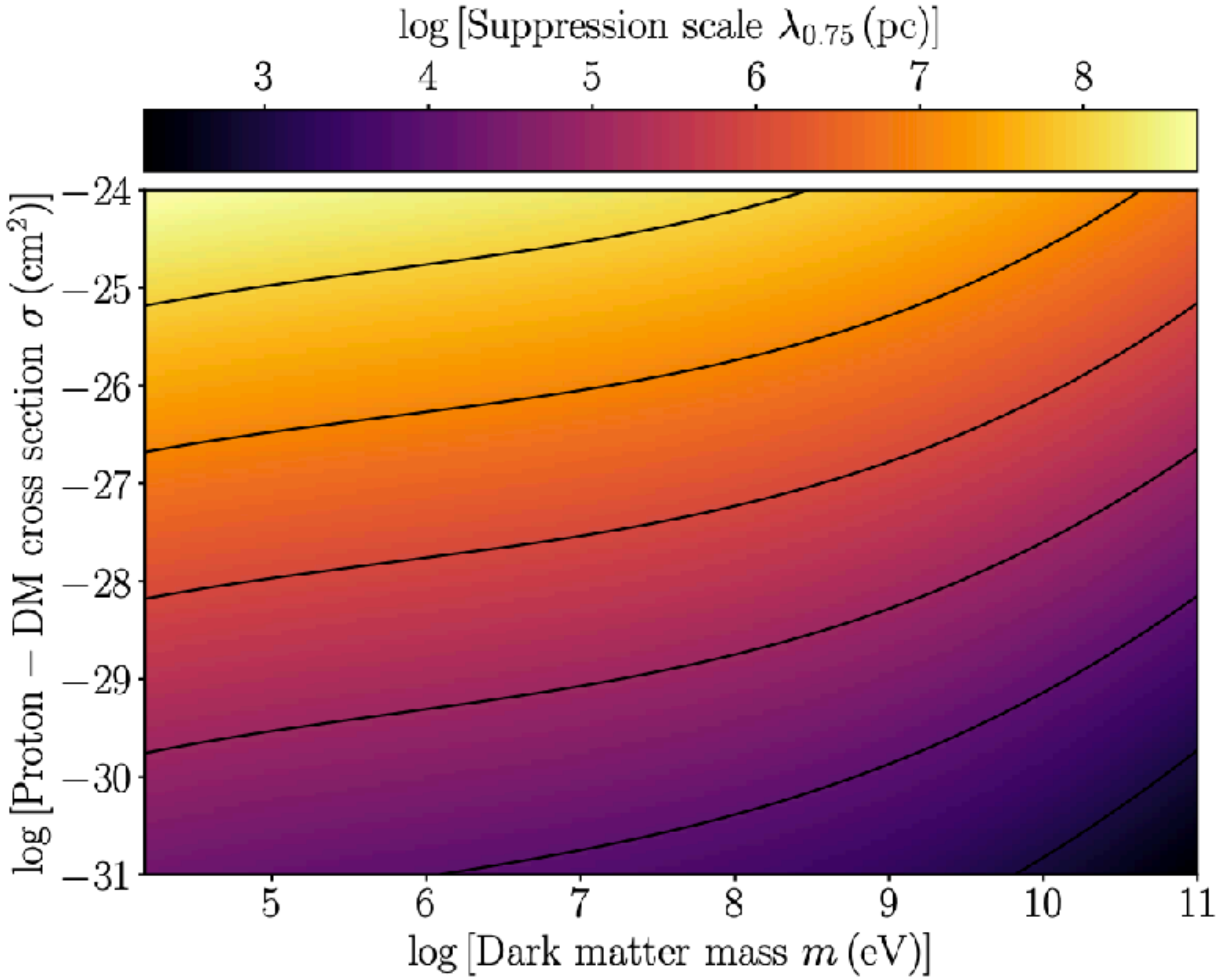


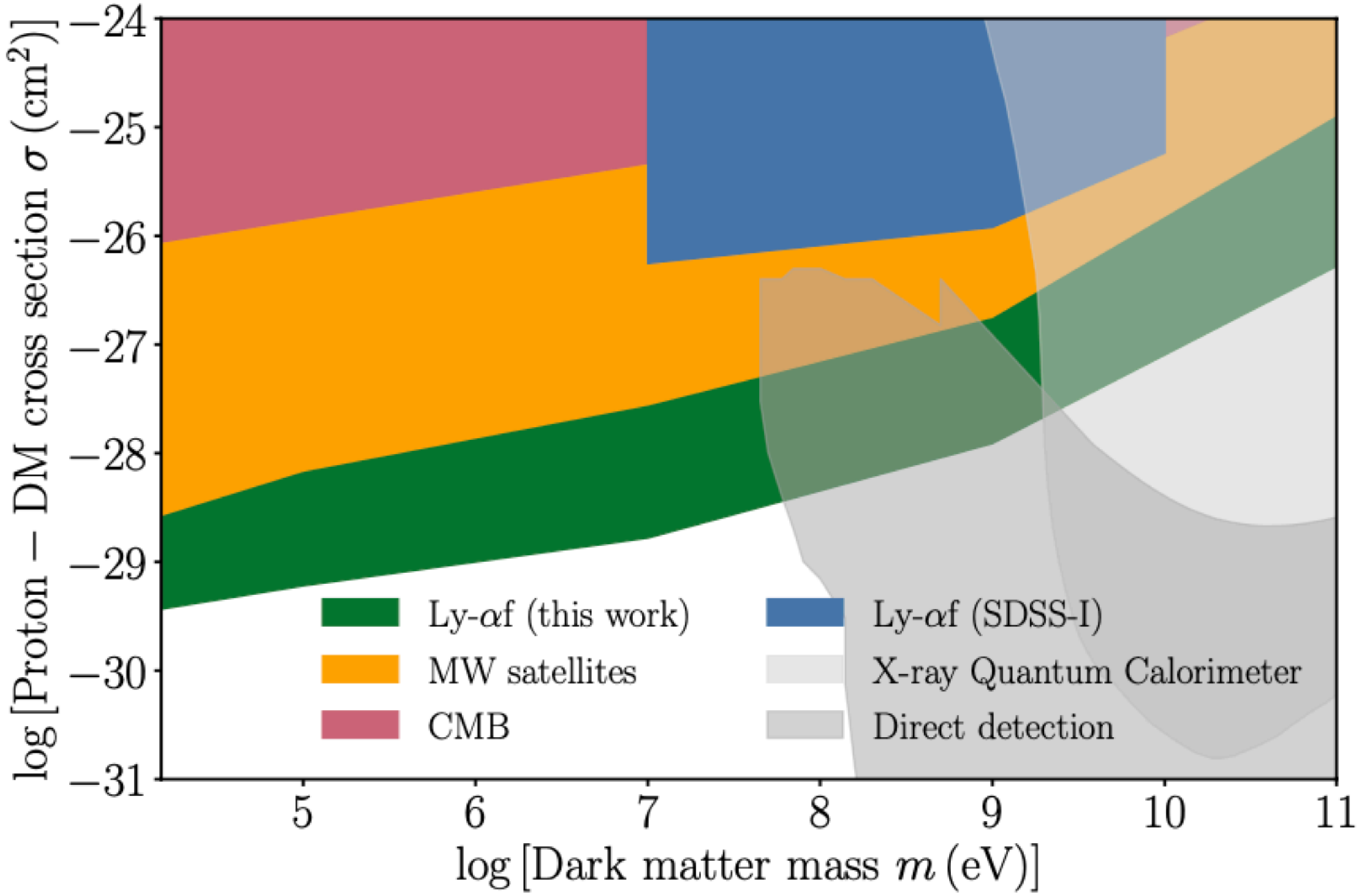
$$m_a > 2 \times 10^{-20}\,\mathrm{eV}$$

Improved bound by ~ order of magnitude

Rogers & Peiris (2021a,b, Phys. Rev. Lett; Phys. Rev. D)

New Lya limits on light dark matter – proton cross section



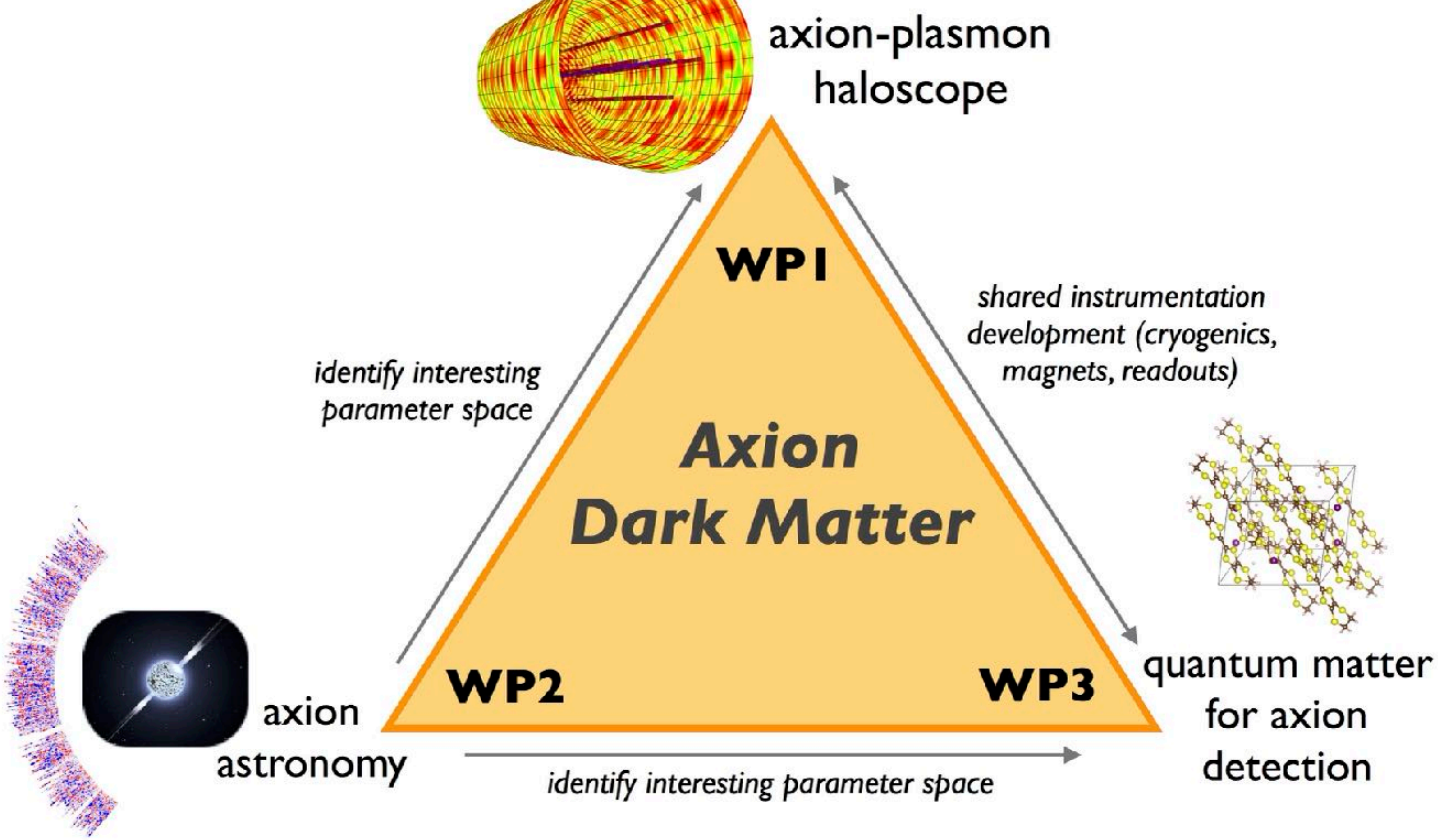


Strongest limits to-date on velocity-independent dark matter (DM) – proton cross section σ for DM masses m = 10 keV to 100 GeV

ROGERS, DVORKIN & PEIRIS (2022, PHYS.REV.LETT.)

AxionDM @ Stockholm

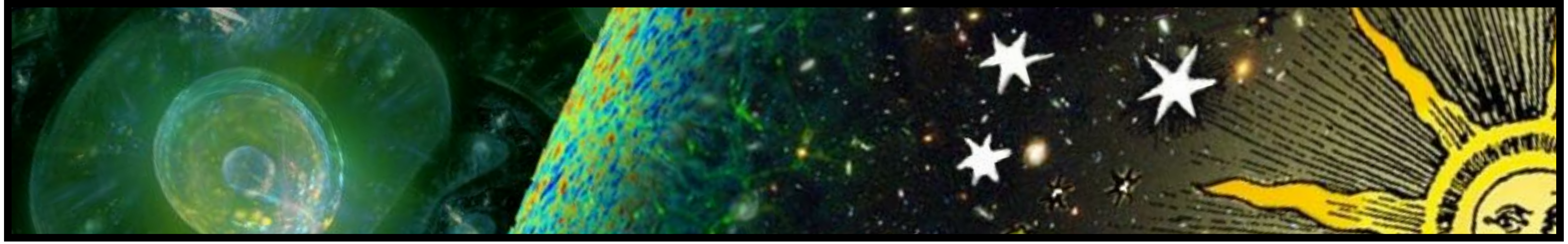
Detecting Axion Dark Matter In The Sky And In The Lab



European Research Council
Established by the European Commission

Peiris (PI) + Bonetti, Conrad, Gudmundsson, Marsh, Wilczek

VETENSKAPSRÅDET
THE SWEDISH RESEARCH COUNCIL



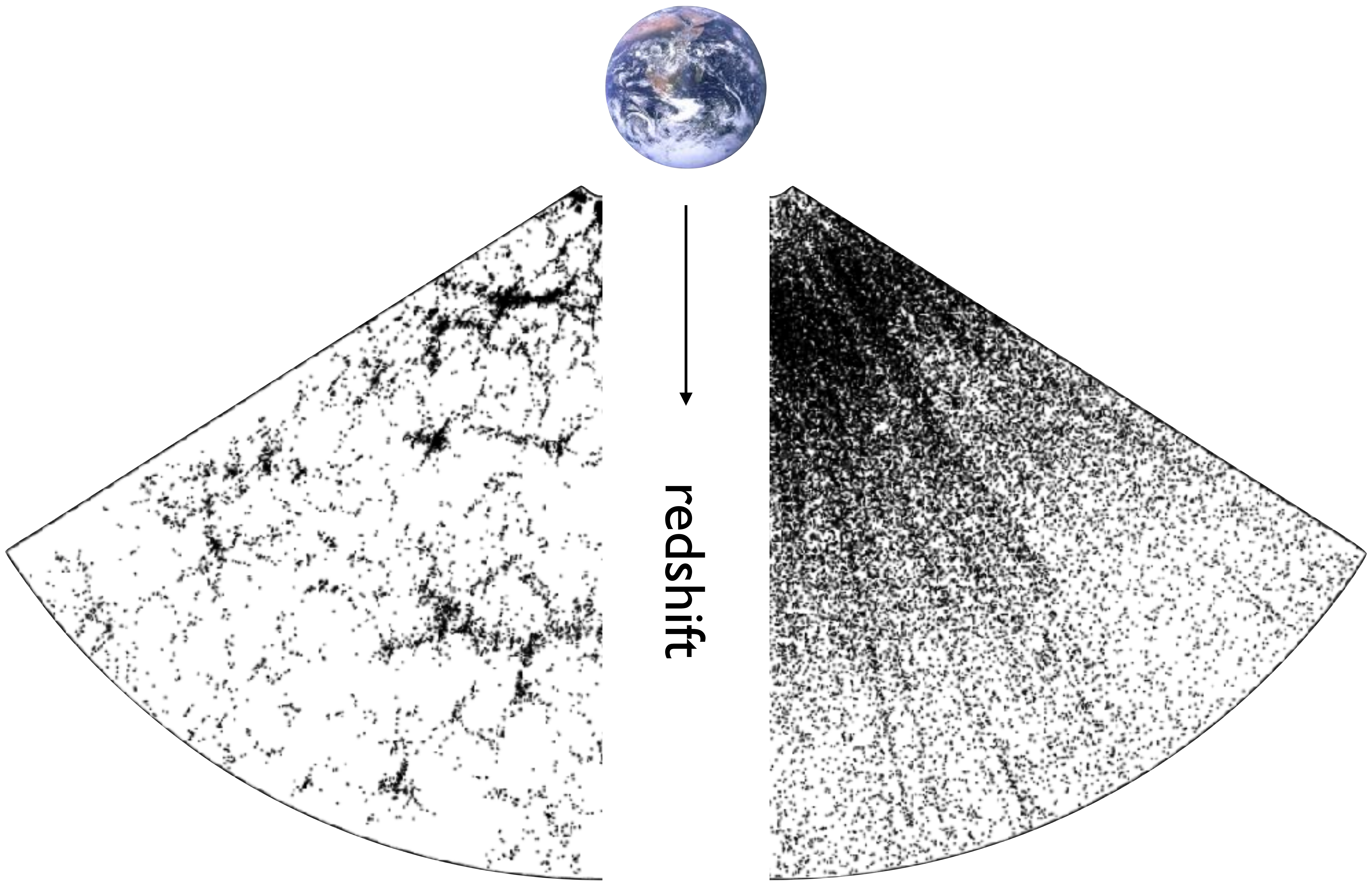
Bayesian hierarchical models with machine learning components

Justin Alsing
(OKC/Stockholm)

Boris Leistedt
(Imperial College London)

With: Joel Leja, Daniel Mortlock, Sinan Deger, Tassia Ferreira, George Efstathiou

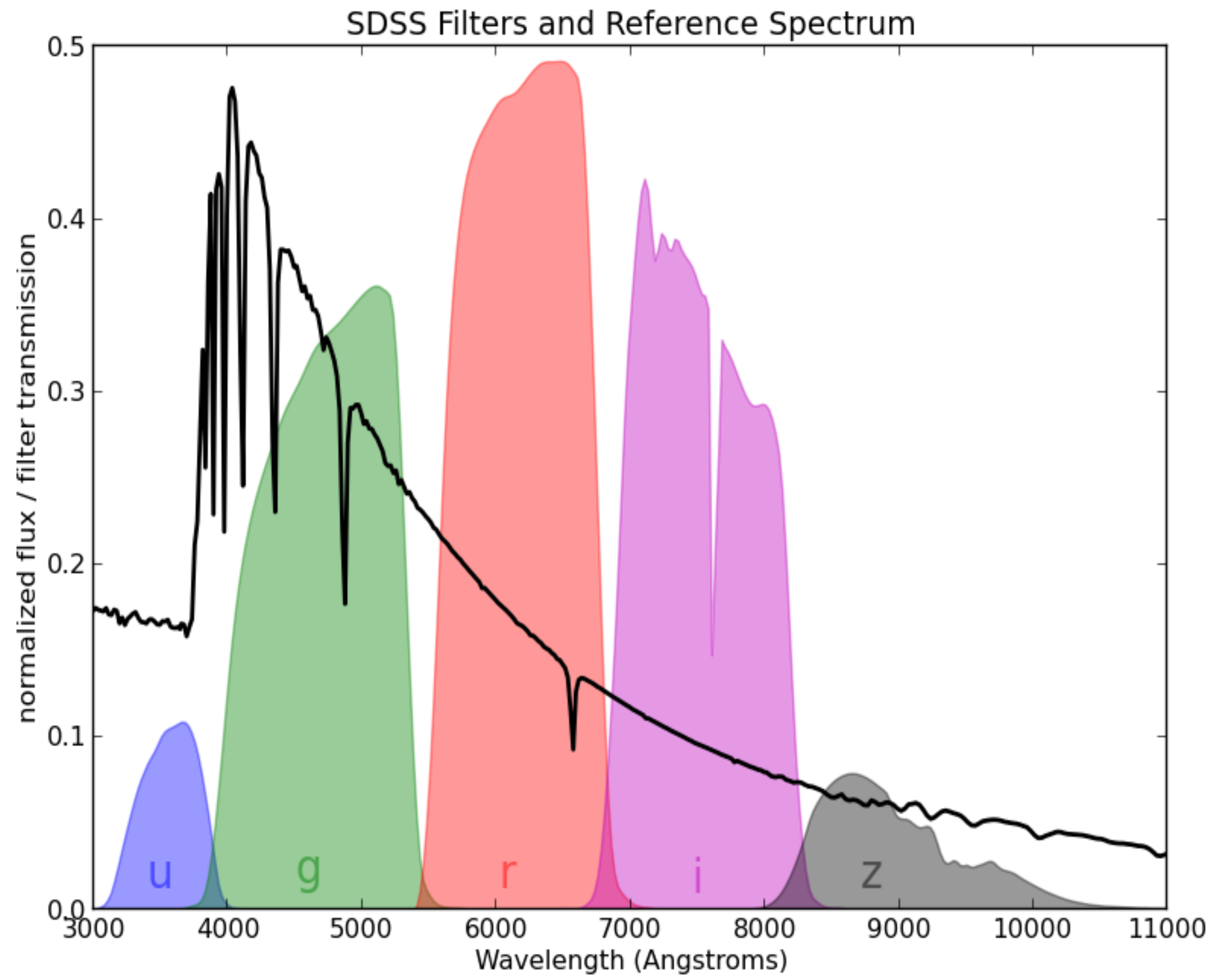
Observational frontier with galaxy surveys



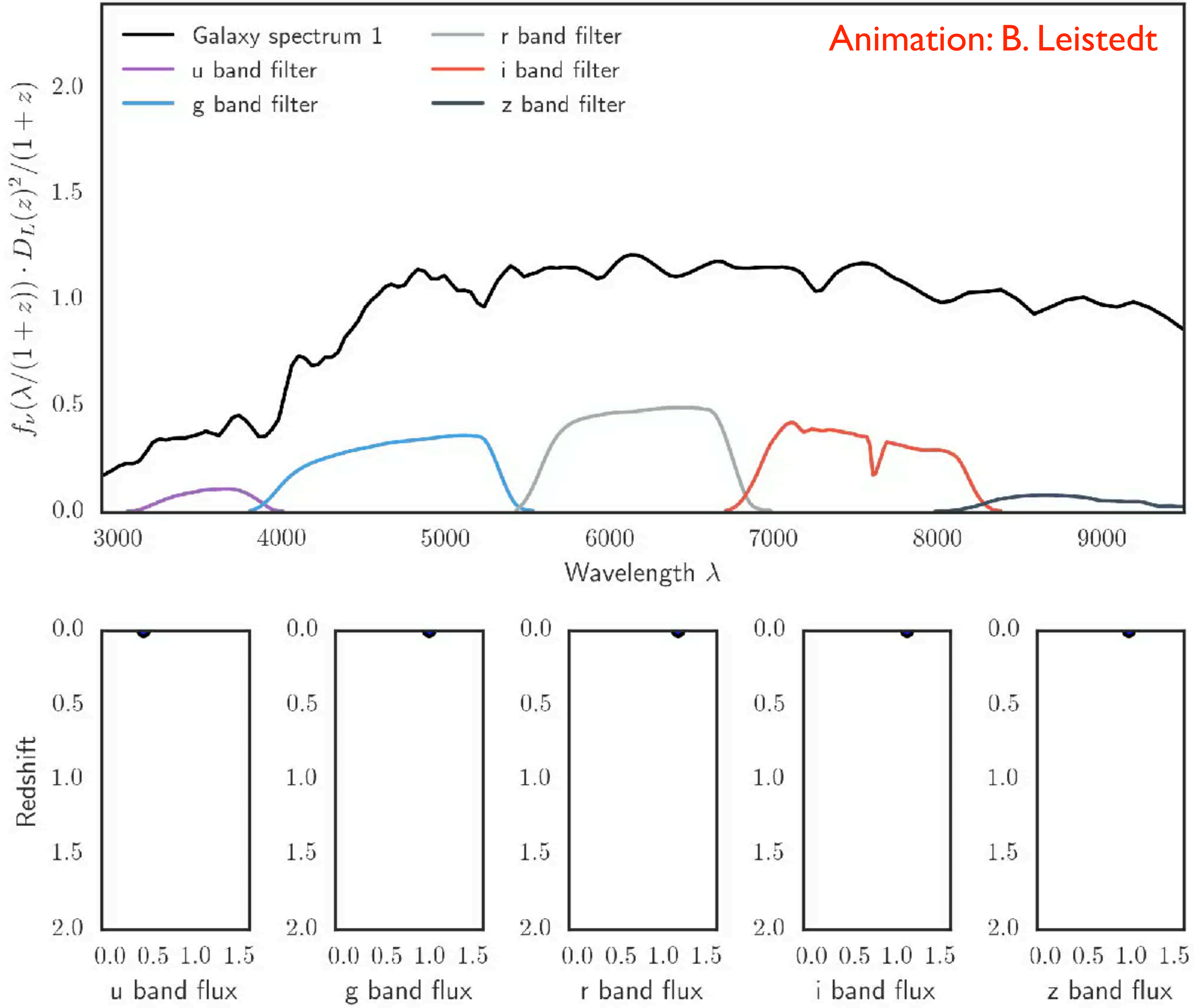
<u>Spectroscopic</u>
DESI (ground)

<u>Photometric</u>
LSST (ground), Euclid (space), Roman (space)

Spectroscopic vs photometric samples

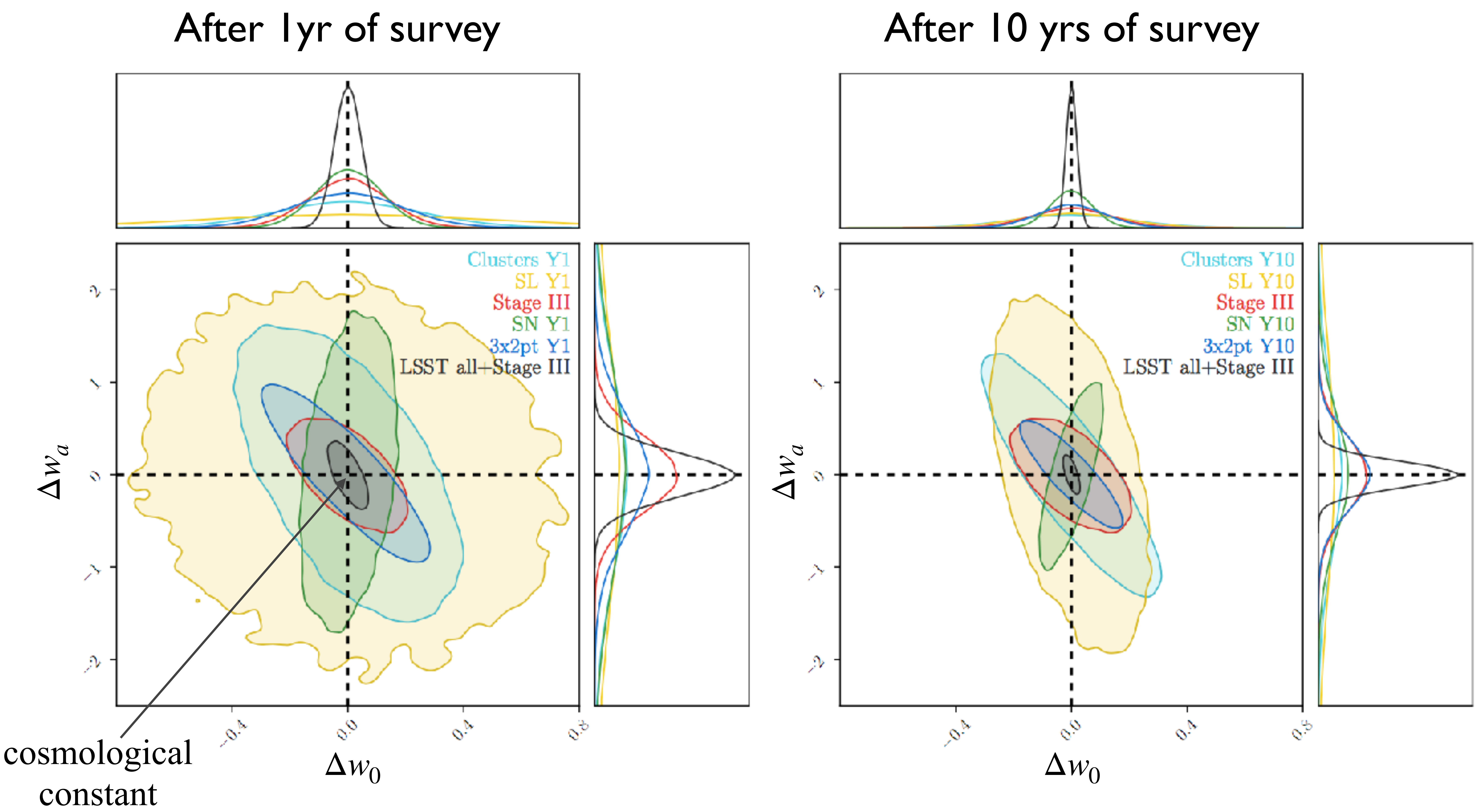


Photometric catalogues require ***redshift estimation***



Animation: B. Leistedt
Galaxy spectrum 1
u band filter
g band filter
r band filter
i band filter
z band filter
$f_\nu(\lambda/(1+z)) \cdot D_L(z)^2/(1+z)$
Wavelength λ
Redshift
u band flux
g band flux
r band flux
i band flux
z band flux

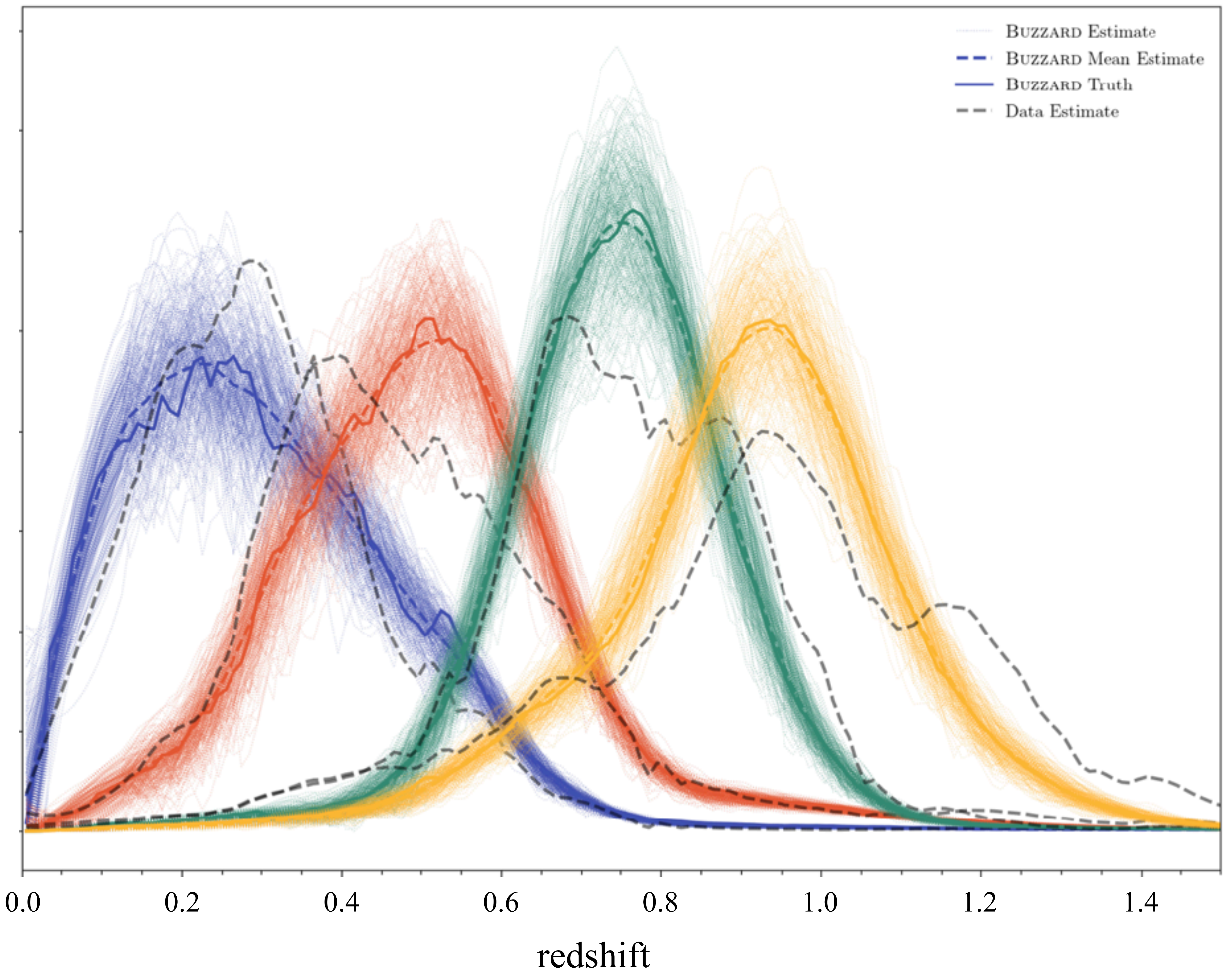
LSST and Dark Energy Science



Measuring if / how dark energy evolves with time

Forecasts: LSST DESC Collaboration

N(z): redshift distribution inference is challenging

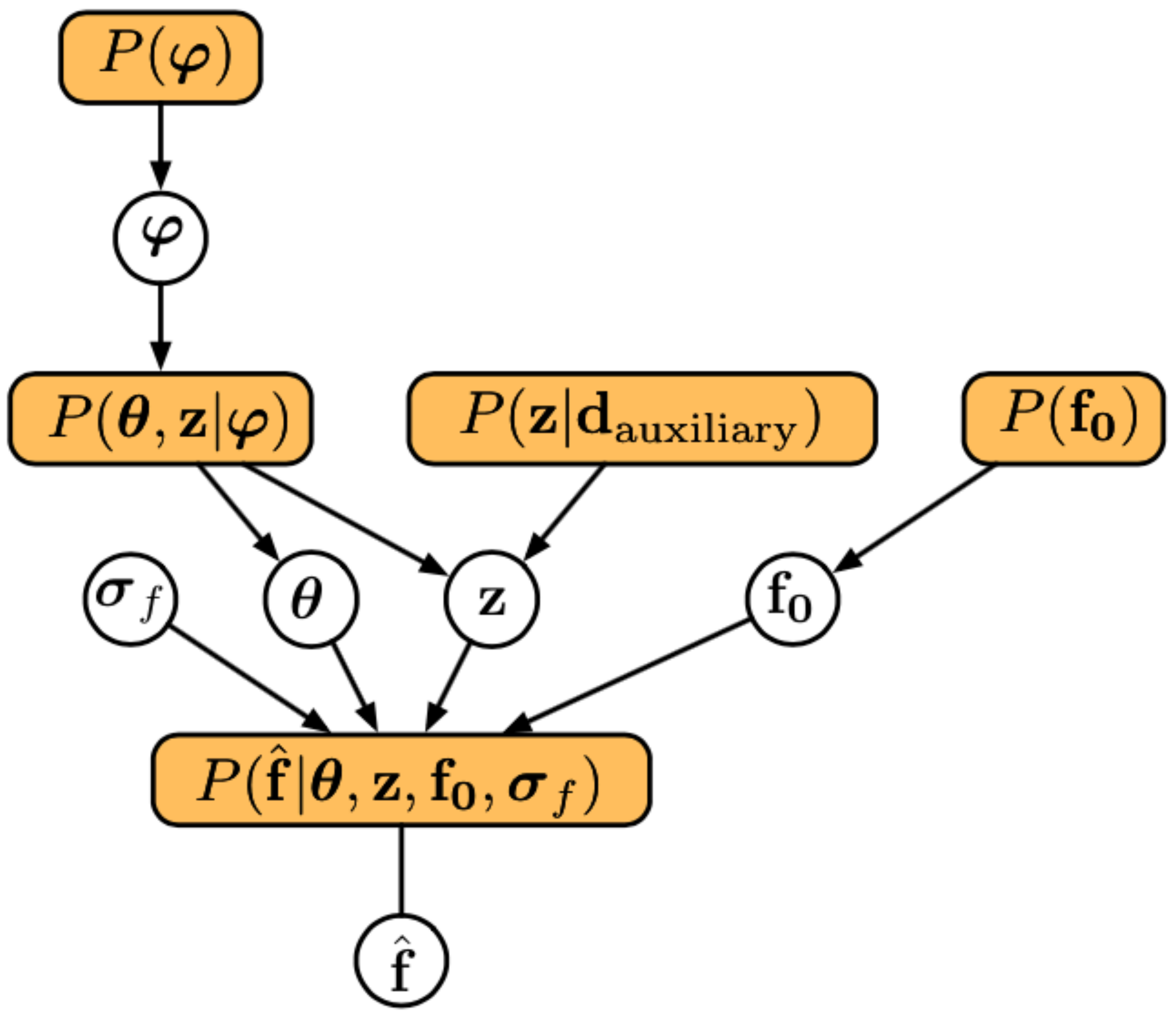


- *Spectroscopic training / calibration samples are:*
  - *not representative of photometric catalogues (due to brighter flux limits and population evolution)*
  - *heterogeneous and contain difficult-to-model selection effects*
- *Introduces biases which are difficult to mitigate at required precision*

Figure: Myles et al (DES Collaboration 2021)

Redshift distribution inference for static cosmology

- ***Key idea:*** *high-dimensional Bayesian hierarchical model with machine-learned parts.*



- *Neural network emulation of FSPS population synthesis model, describing realistic galaxy populations (replace templates).*

- *Flexible NN-parameterised probability density models (e.g. normalising flows) to describe population prior and selection effects.*

ALSING, PEIRIS, LEJA, HAHN, TOJEIRO, MORTLOCK, LEISTEDT, JOHNSON, CONROY (APJS, 2020); LEISTEDT, MORTLOCK AND PEIRIS (MNRAS, 2016)

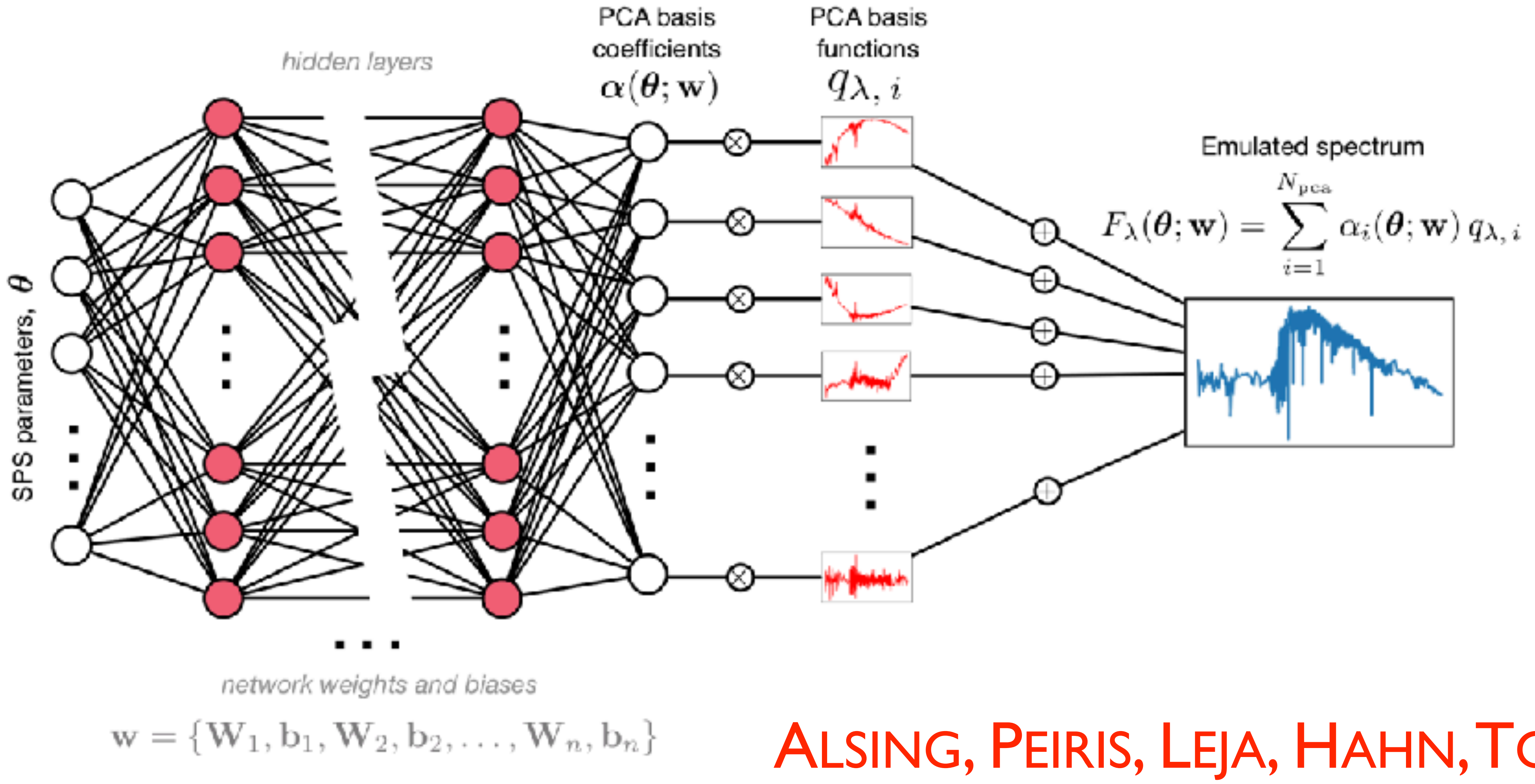
Emulating stellar population synthesis (SPS) models

- *SPS models (e.g. FSPS, Charlie Conroy and collaborators) are fast (<1 sec) but use cases require* ***large numbers of model evaluations****.*

- *Stage IV galaxy survey catalog sim ~ 10^{10} SPS evaluations*

- *Leja et al (2019) analysis of 60,000 galaxies under 14-parameter SPS model cost1.5 million CPU-hrs.*

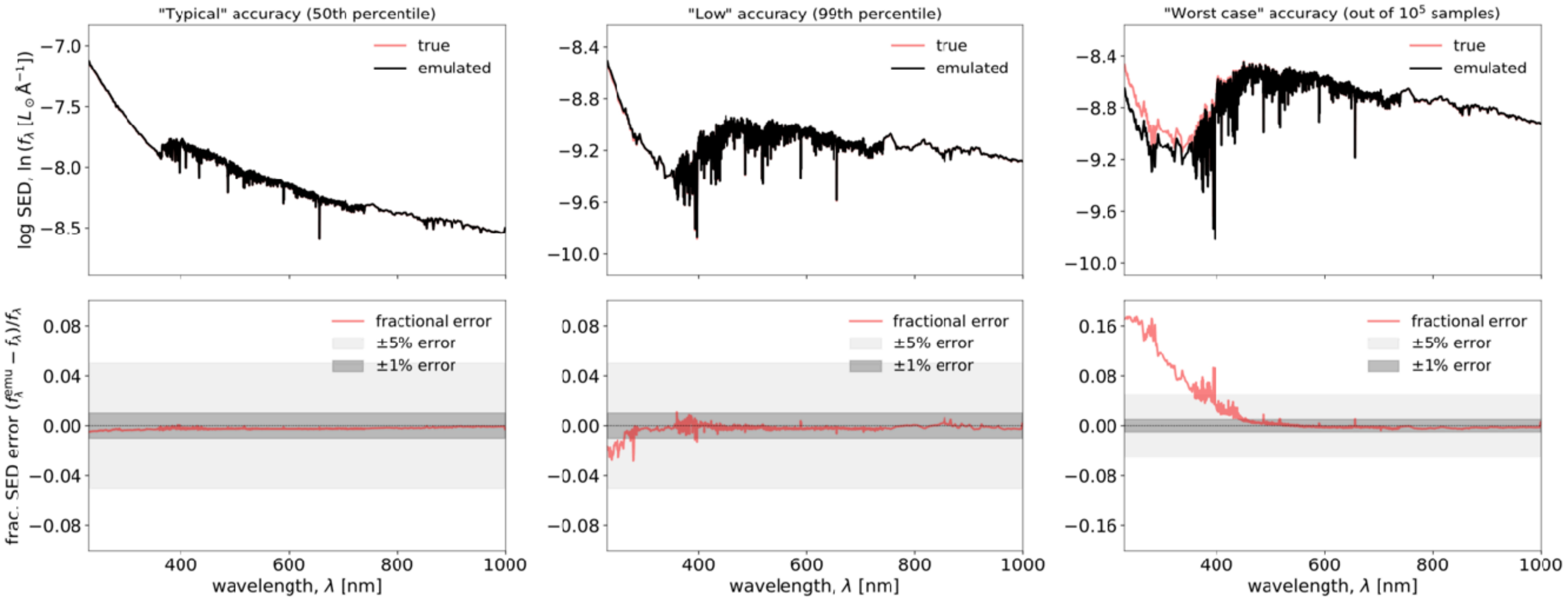
- *Can generate training sets of ~10^5* ***enabling neural network emulators.***



SPECULATOR SPS emulator

ALSING, PEIRIS, LEJA, HAHN, TOJEIRO, MORTLOCK, LEISTEDT, JOHNSON, CONROY (APJS, 2020)

Example: DESI Bright Galaxy Survey SEDs



- *Accuracy <1% over the 8-parameter FSPS model for >99% of SEDs*
- *Generating 10^6 SEDs takes 2s on Tesla K80 GPU (Speedup 10^5 over FSPS on CPU); inference under SPS models can make use of gradients*

ALSING ET AL. (APJS, 2020)

Forward modelling for n(z)

n(z): integral over **selection** **x** **data model** **x** **population model**

$$n(z) \equiv P(z|S)$$
$$= \frac{1}{P(S)} \int \left[\iint P(S|\hat{\mathbf{f}}, \theta, z) P(\hat{\mathbf{f}}|\theta, z, \sigma) P(\sigma) d\hat{\mathbf{f}} d\sigma \right] P(\theta, z) d\theta$$

Forward modelling for n(z)

n(z): integral over **selection x data model x population model**

$$n(z) \equiv P(z|S)$$
$$= \frac{1}{P(S)} \int \left[\iint P(S|\hat{\mathbf{f}}, \theta, z) P(\hat{\mathbf{f}}|\theta, z, \sigma) P(\sigma) d\hat{\mathbf{f}} d\sigma \right] P(\theta, z) d\theta$$

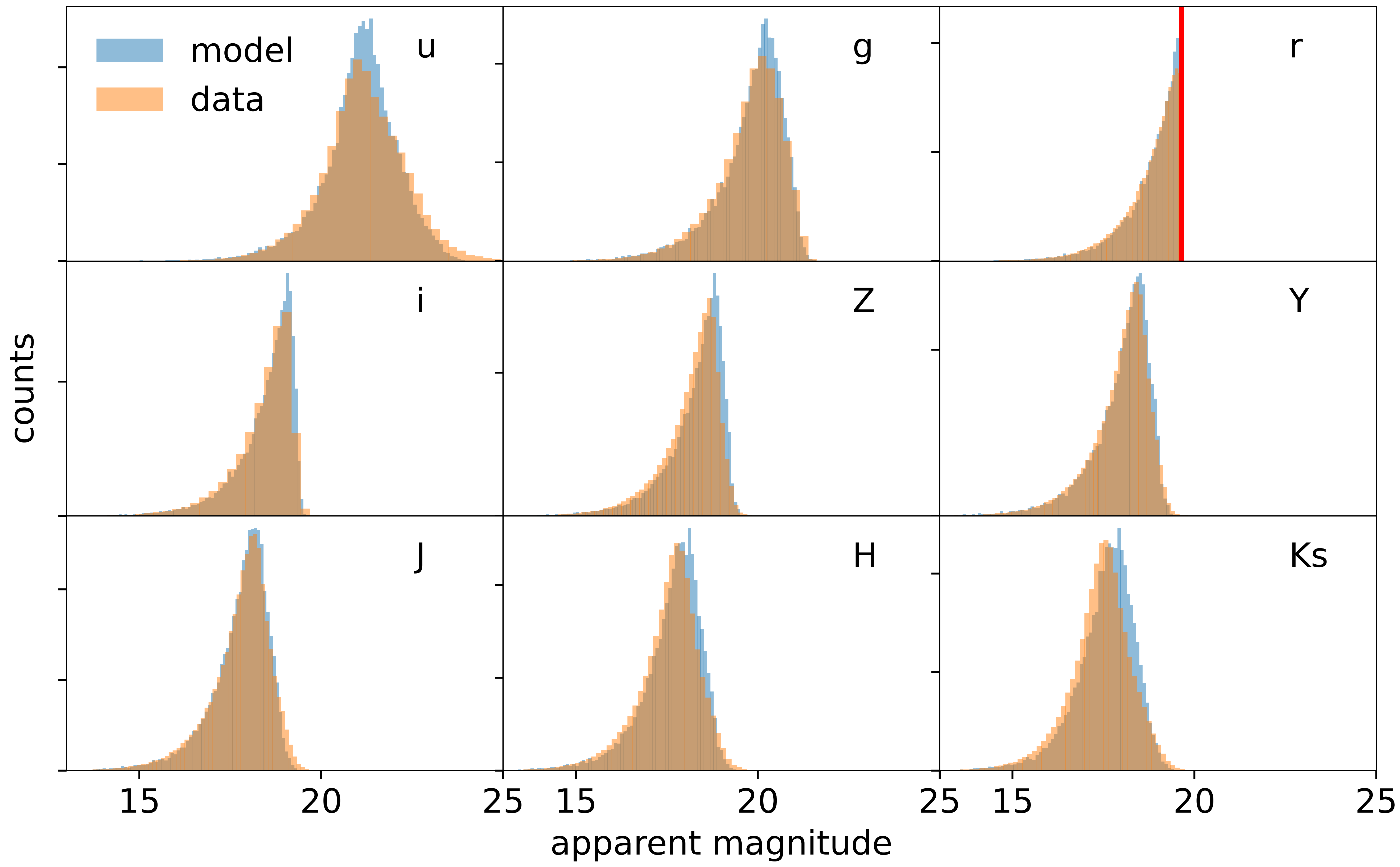
Advantages:

- *Does not rely on spectroscopic redshift calibration*
- *Auxiliary data (spec-z, extra surveys) can be included seamlessly (extended data vector or extra priors for objects with extra information)*
- *Connects cosmology with galaxy evolution*

"Turns photo-z back into an astrophysics problem" — Justin Alsing

Broadband data: does it work?

Simulated galaxy population (encoding galaxy evolution calibrated to observations), combined with data model and selection cuts, should be able to predict redshift distribution.



Selection for GAMA survey

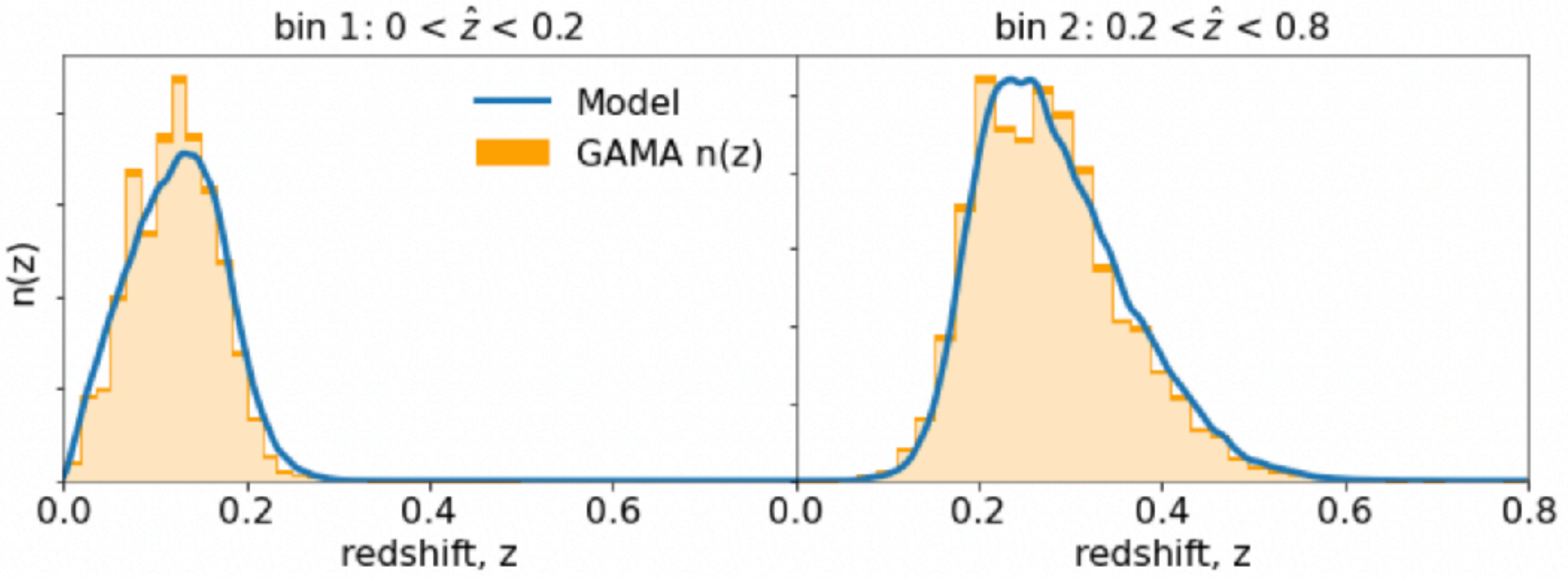
Validation: *two spectroscopic surveys with straightforward selection cuts*

1. GAMA (ugriZYJHKs): *r < 19.65, (J-Ks) > 0.01*
2. VVDS (UBVRI): *I < 22.5, star-galaxy separation done at level of spectra*

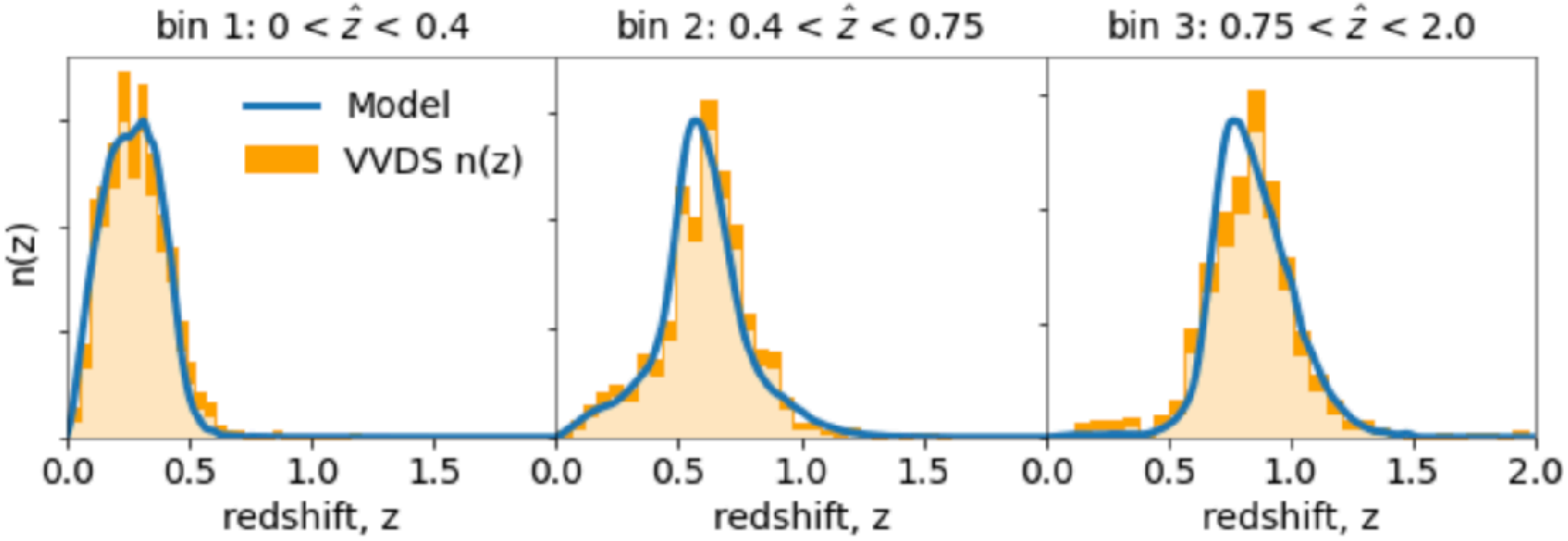
Alsing et al. (arXiv:2207.07673, ApJS accepted)

How good is the baseline model?

GAMA n(z) tomography
population model X data model X selection



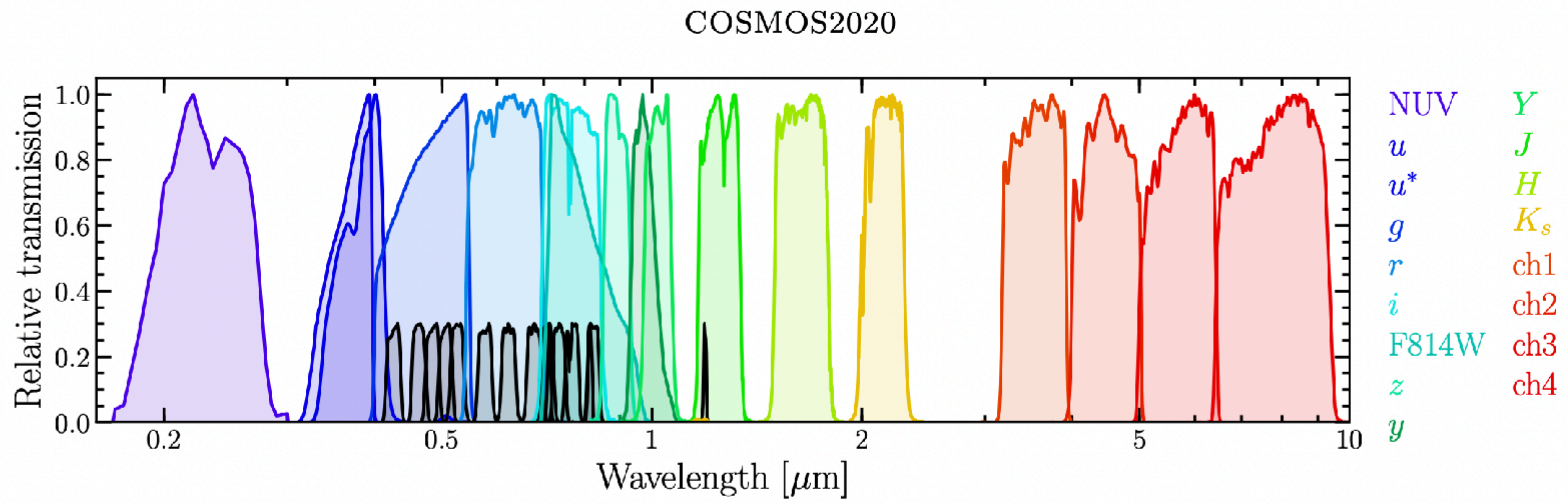
VVDS n(z) tomography
population model X data model X selection



Baseline model n(z) bias < 0.01 *<u>before parameter inference</u> (no data!)*

ALSING ET AL. (ARXIV:2207.07673, APJS ACCEPTED)

Narrow-band data: validation with COSMOS2020



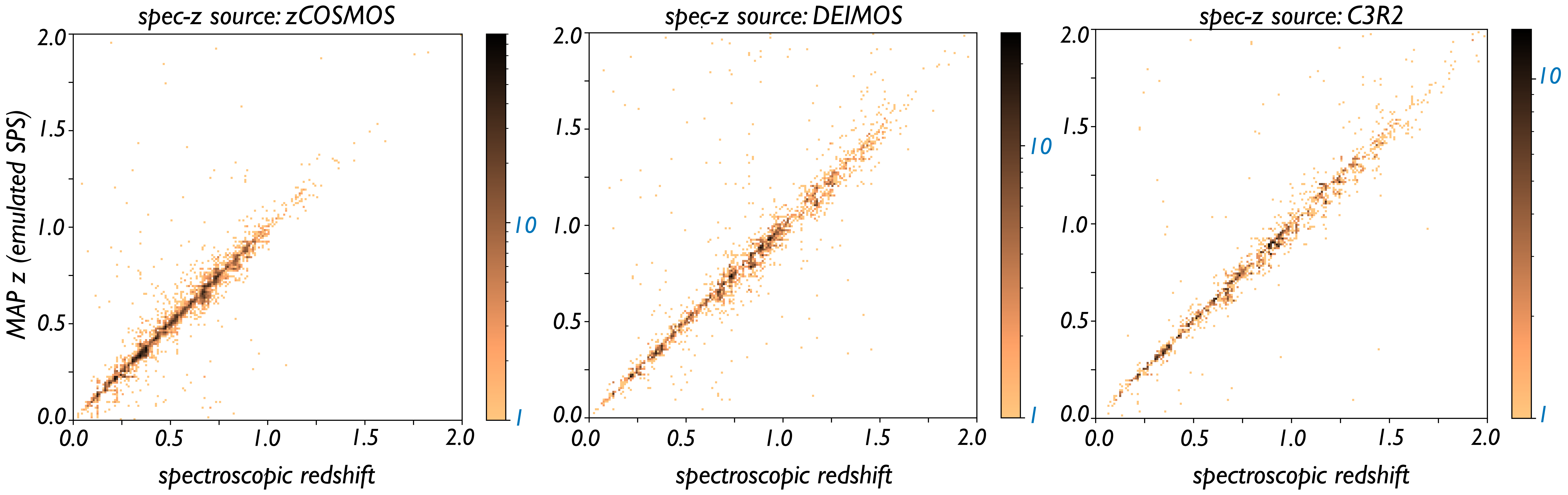
Photometric data: COSMOS2020 multiwavelength Farmer catalogue

Population model: Prospector-alpha emulators of both fluxes and emission lines

Data model: Optimization of zero-points per band and (broadband and emission line) hyperparameters

Weaver et al (2021), Leistedt et al. (arXiv:2207.07673, ApJS accepted)

Narrow-band data: validation with COSMOS2020



Representative of bright sources

Representative of colours

"Spec-z quality"

Photometric data: COSMOS2020 multiwavelength Farmer catalogue

Population model: Prospector-alpha emulators of both fluxes and emission lines

Data model: Optimization of zero-points per band and (broadband and emission line) hyperparameters

LEISTEDT ET AL. (ARXIV:2207.07673, APJS ACCEPTED)

Next steps!

- ***Hierarchical inference not scalable?***

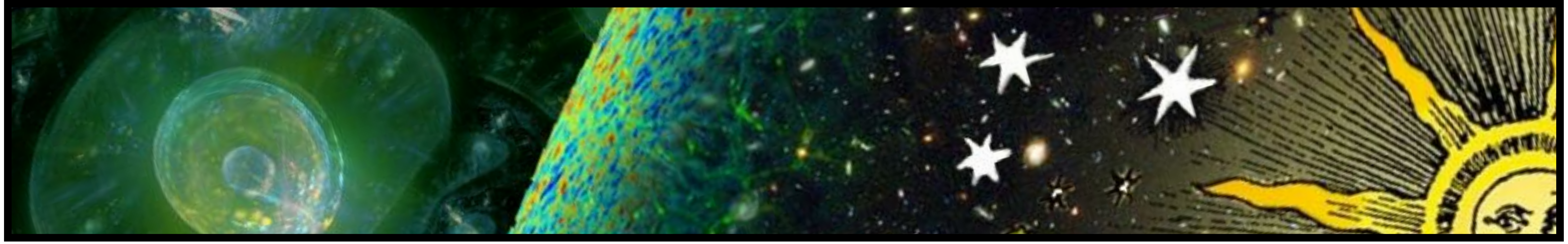
  Already made progress on simulation-based inference approach — advantage of not needing to explicitly model selection effects parametrically, only to forward model them in a simulation.

- ***Is the SPS population prior good enough for deeper data?***

  Improvements to population prior (star formation history and dust modelling) under way.

- ***How do we validate analyses of deeper data when little spectroscopy available?***

  Developing posterior predictive checks in colour/flux space (Bayesian "cross-validation")



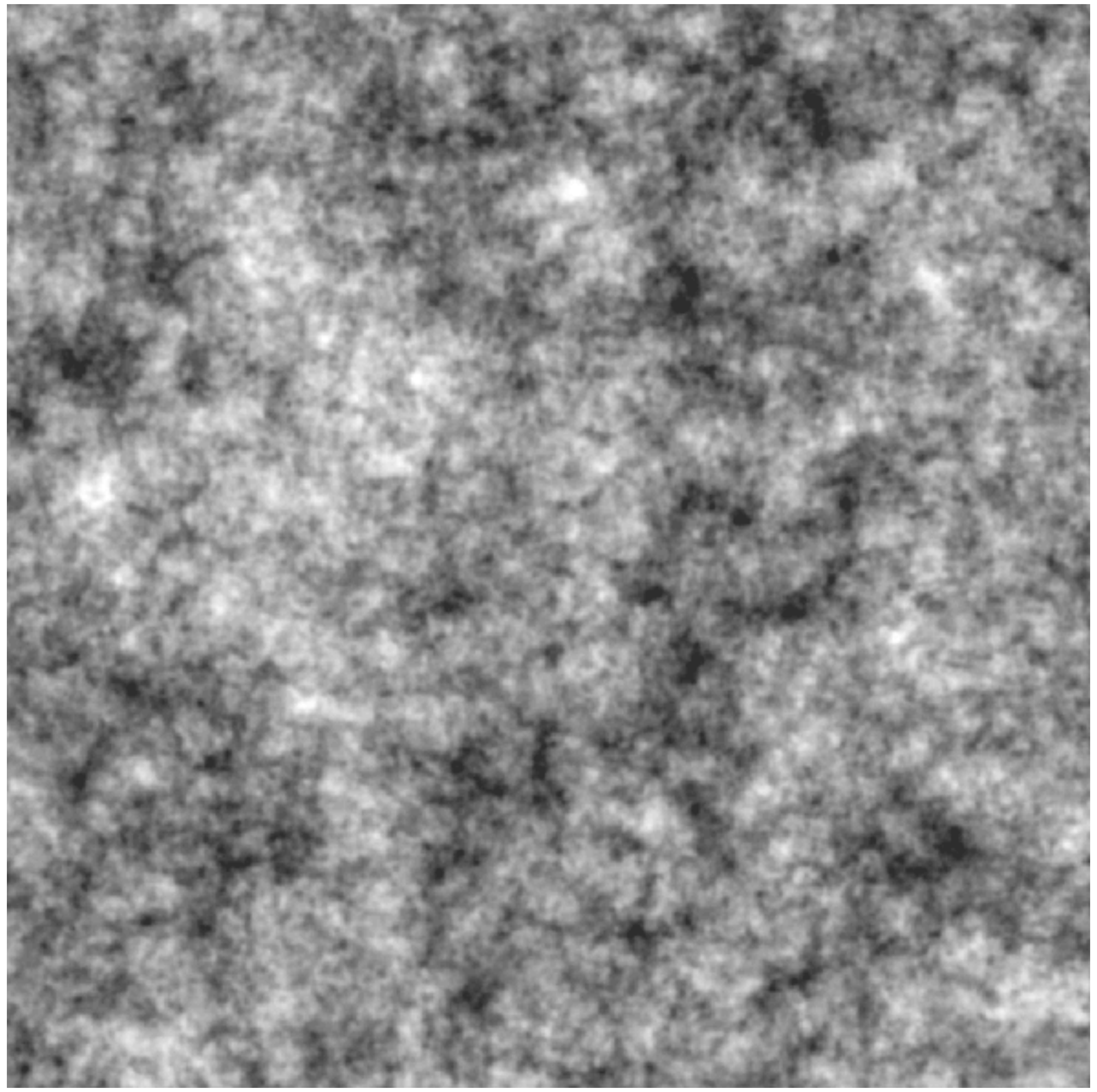
Knowledge extraction using deep learning

Luisa Lucie-Smith
(MPA/Garching)

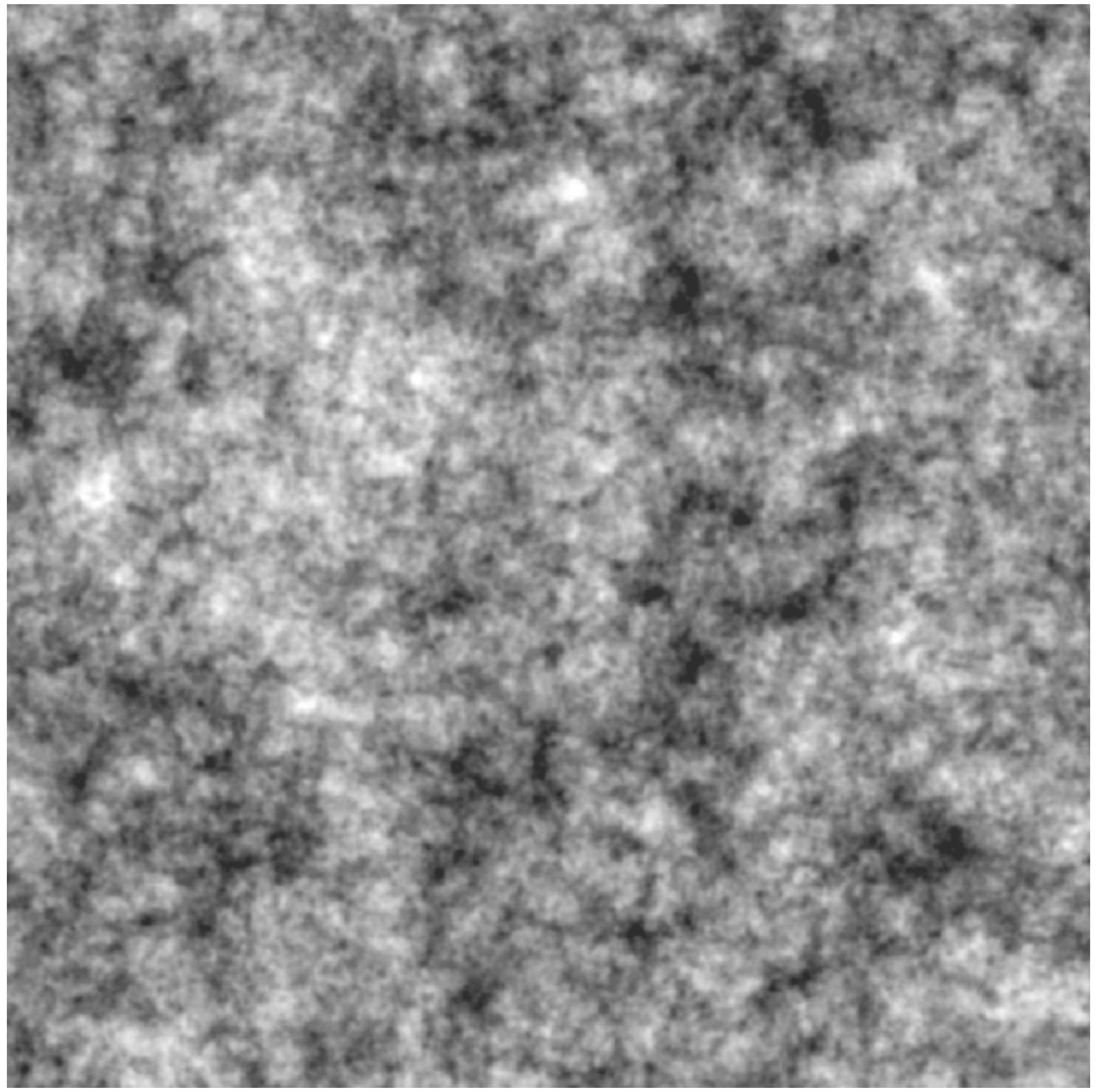
Andrew Pontzen
(UCL)

With: Brian Nord, Jeyan Thiyagalingam, Davide Piras, Lillian Guo

Understanding cosmological structure formation



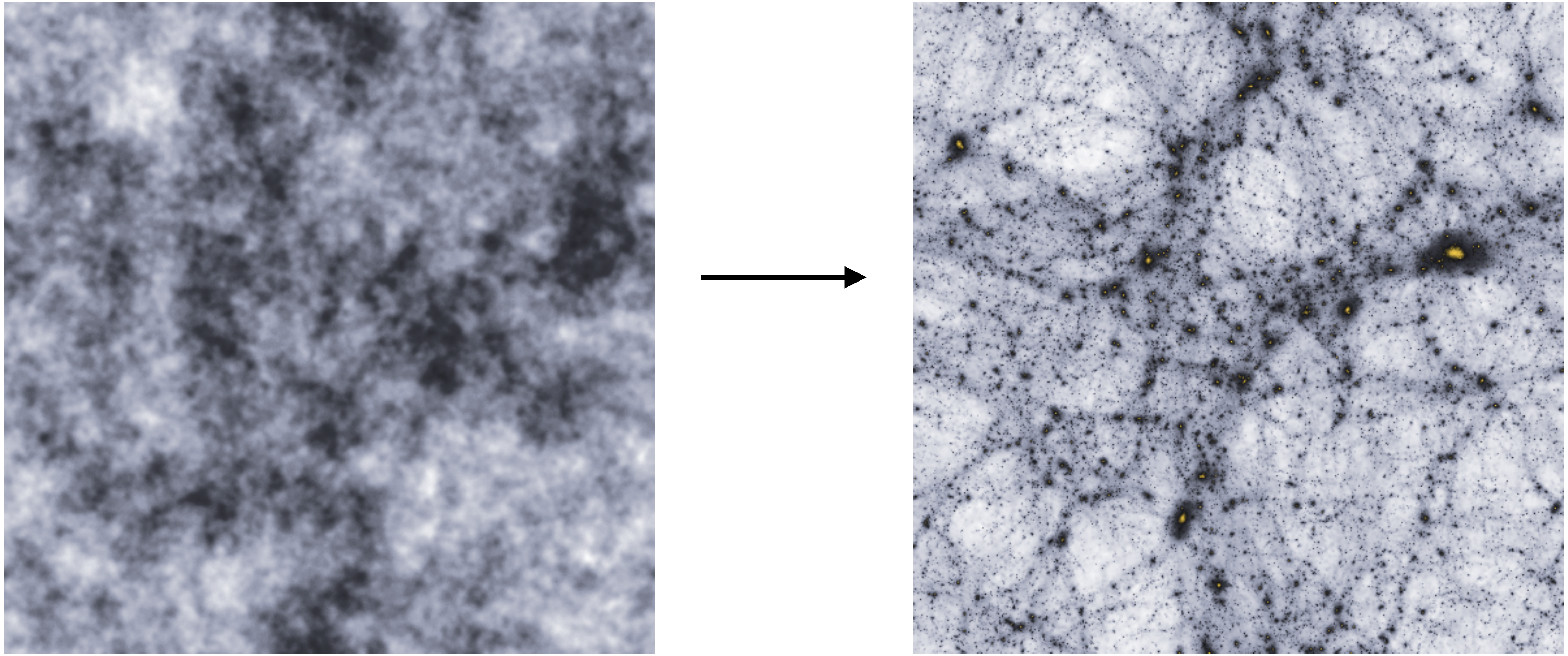
Perturbations in matter density at early times



Large-scale structure at late times

Movie: Pontzen et al. (2016)

Understanding cosmological structure formation



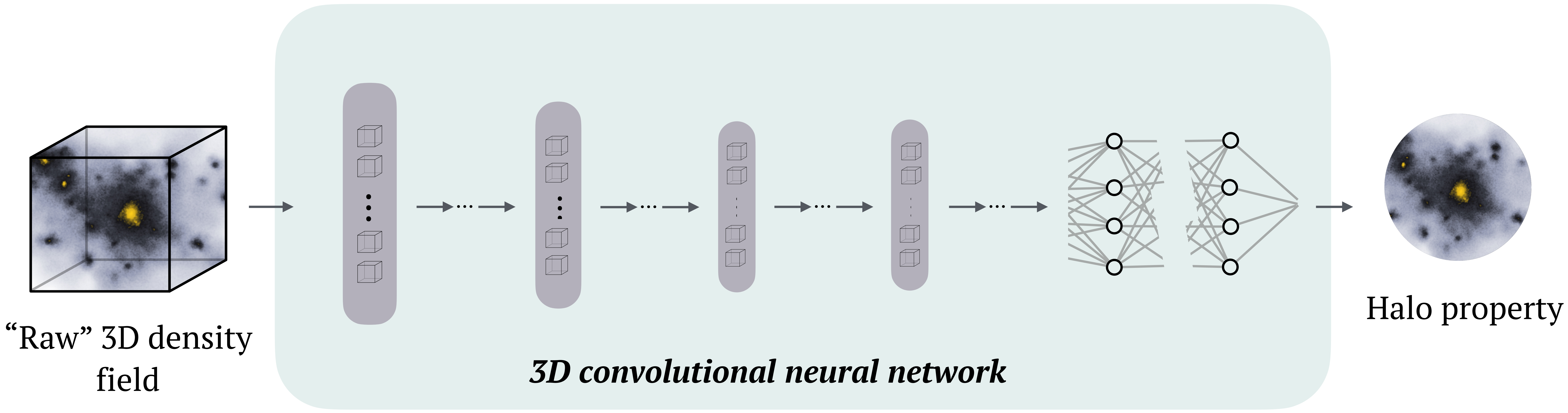
Law of gravity determines mapping
But does not give an *explanation* of mapping
(cf biochemistry vs biology)

“More is different”: emergent phenomena in cosmology

- Can we reliably access rich information in **cosmic web**?
- Can we understand **“mesoscale” phenomena** in structure formation?
- How do “universal” properties emerge?
- Can machine learning play a role in building accurate mesoscale models of complex phenomena?

CREDIT: UCL / PONTZEN

Why convolutional neural networks?



Advantages:

- no *featurization:* CNN learns directly from "raw data"
- CNN learns which features of the raw data are relevant for halo property
- CNNs are able to effectively learn complicated highly non-linear mappings

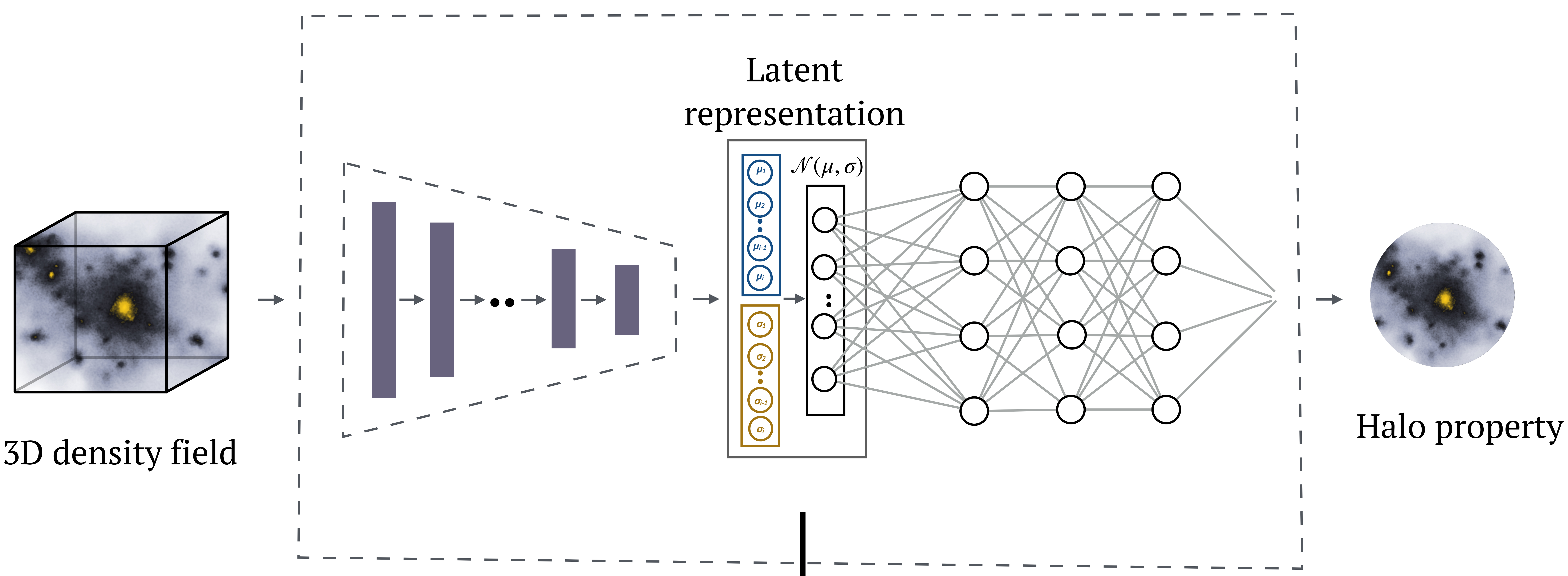
Disadvantages:

- DL algorithms are "black-box" algorithms, encoding features in very high-dimensional models.

How do we extract physical knowledge from a DL algorithm?

New framework for knowledge extraction using AI

Supervised variational encoder

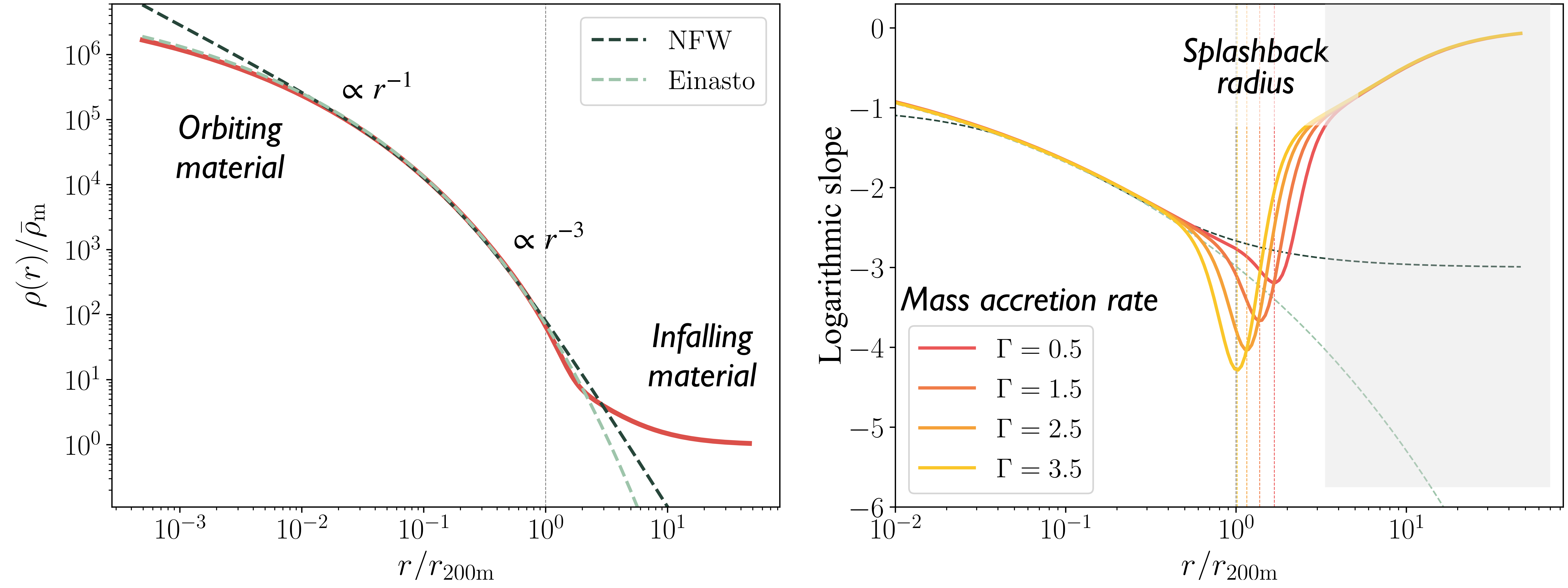


Latent variables encode most relevant aspects of 3D density field about halo property

Model compression to enable "explainable" AI

ITEN ET AL (PHYS. REV. LETT. 2020), LUCIE-SMITH, PEIRIS, PONTZEN, NORD, PIRAS (PHYS. REV. D, 2022)

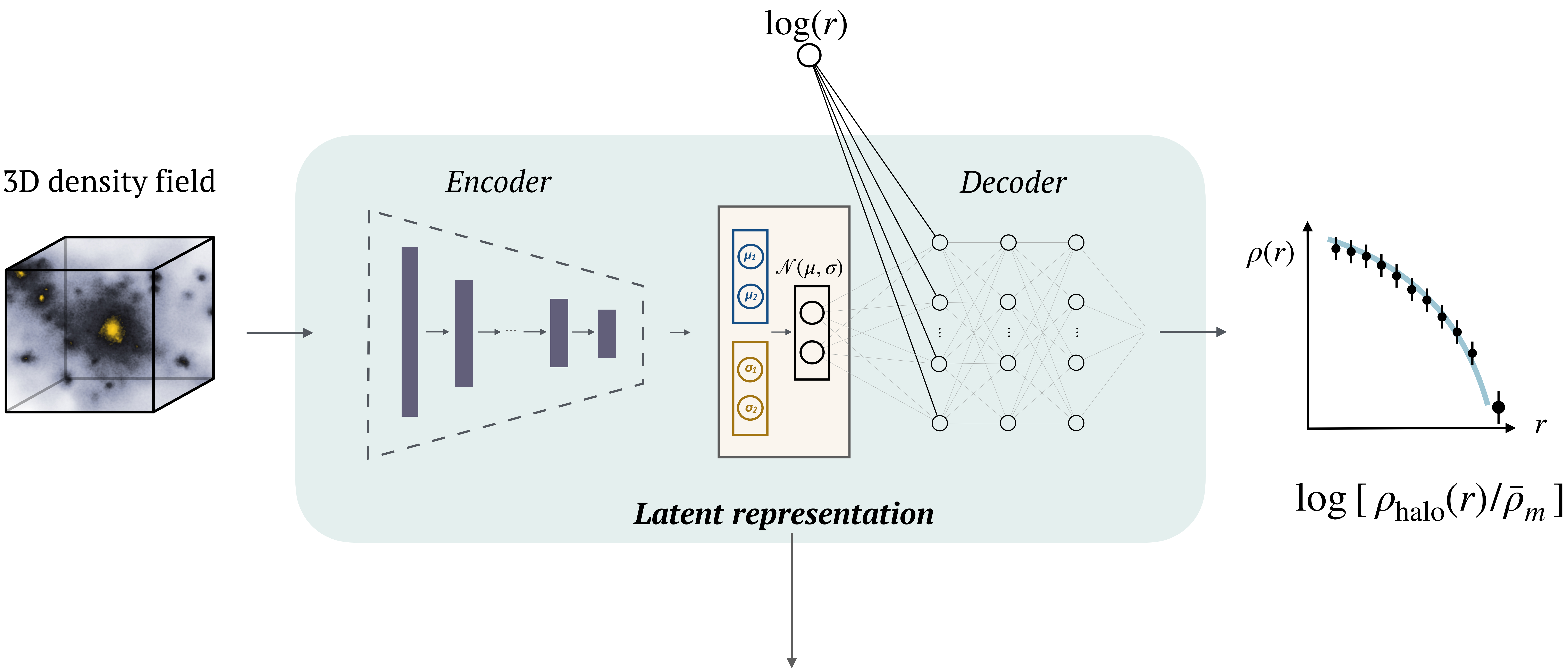
Case study: can neural networks discover the building blocks of dark matter halo profiles?



Existing physical models, based on empirical fitting functions, lack ***explainability***

Navarro, Frenk, White (1996, 1997); Einasto (1965); Diemer & Kravstov (2014); Adhikari et al. (2014); More et al. (2015)

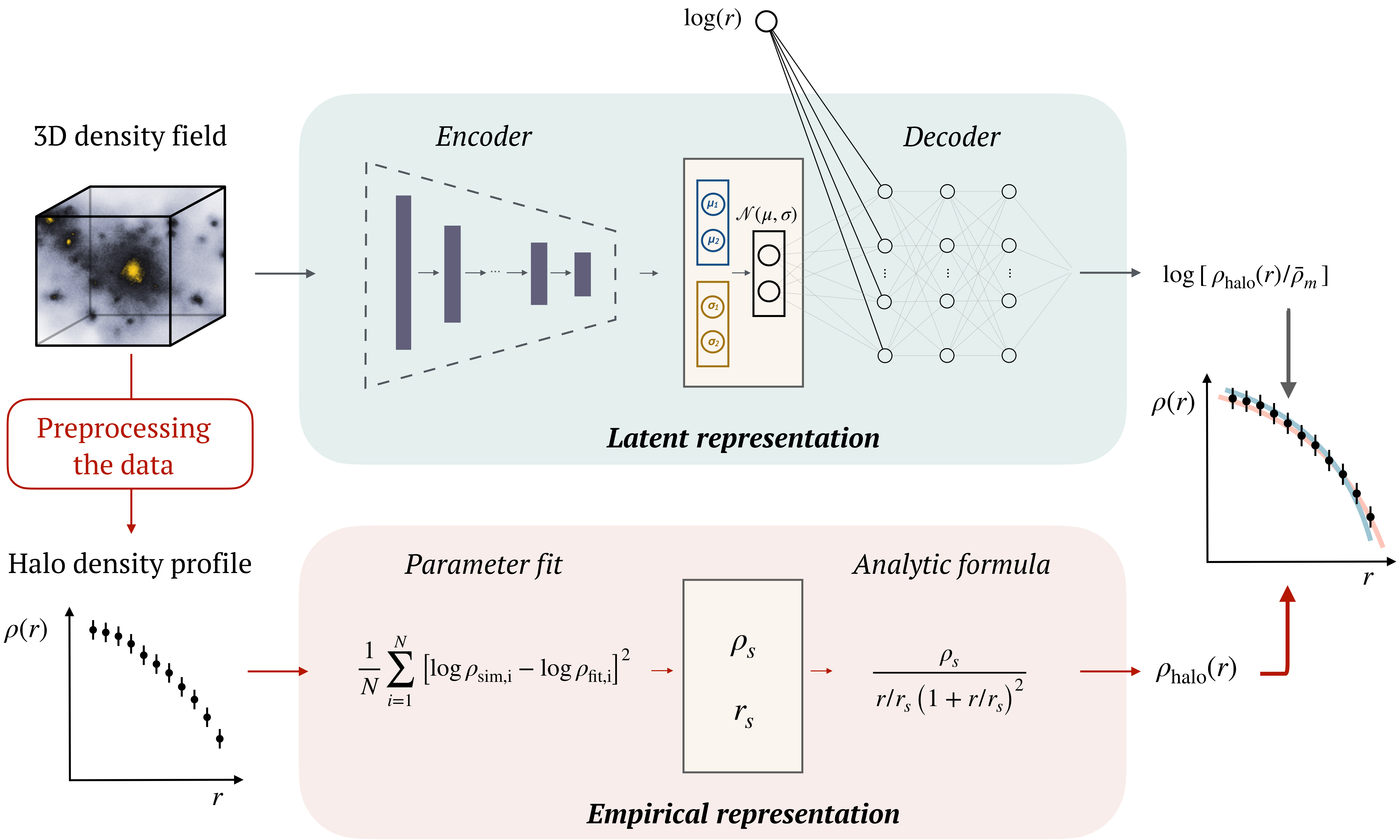
Designing an interpretable variational encoder for knowledge extraction



Latent representation retains all the information used by model to predict density profiles

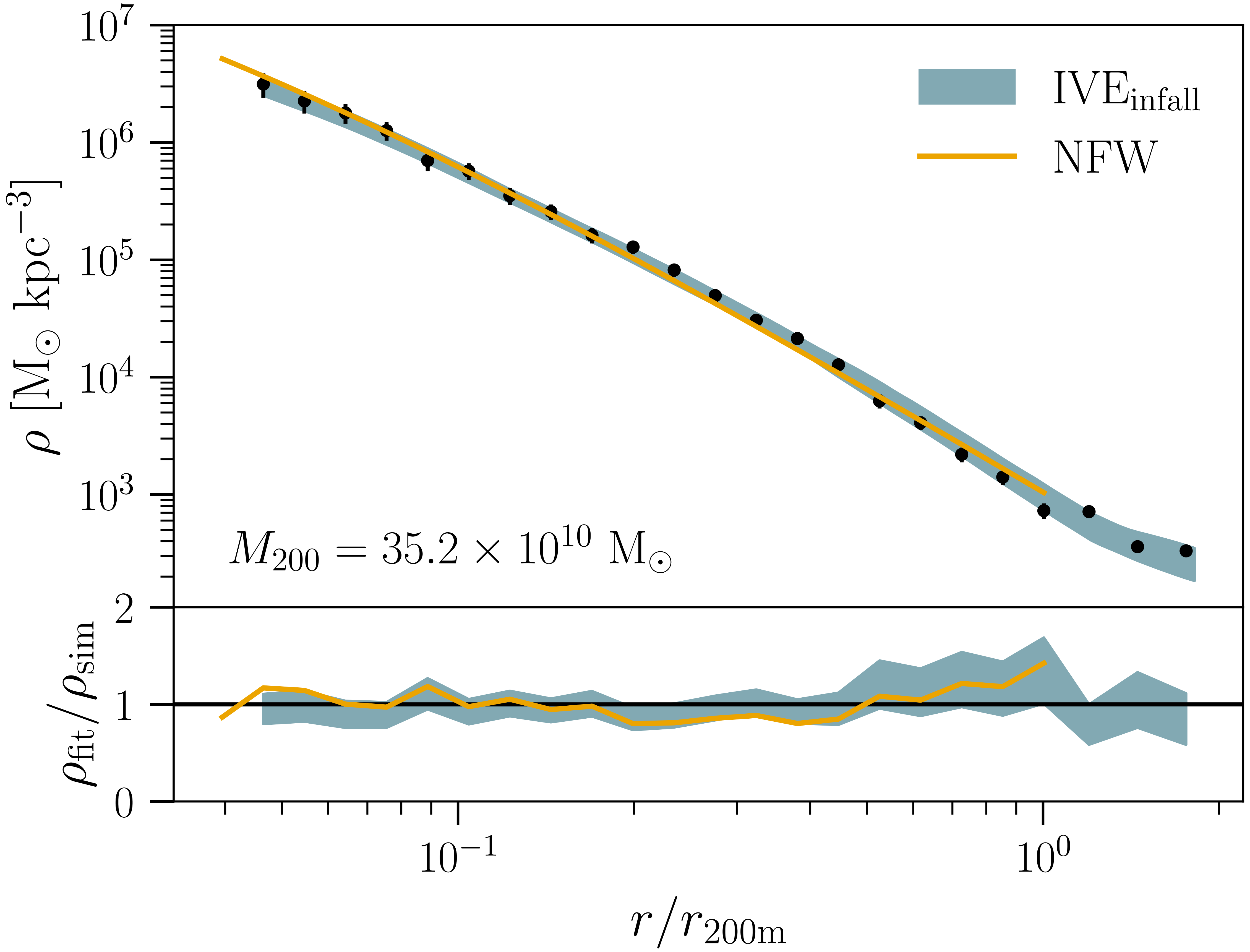
Lucie-Smith, Peiris, Pontzen, Nord + (2022)

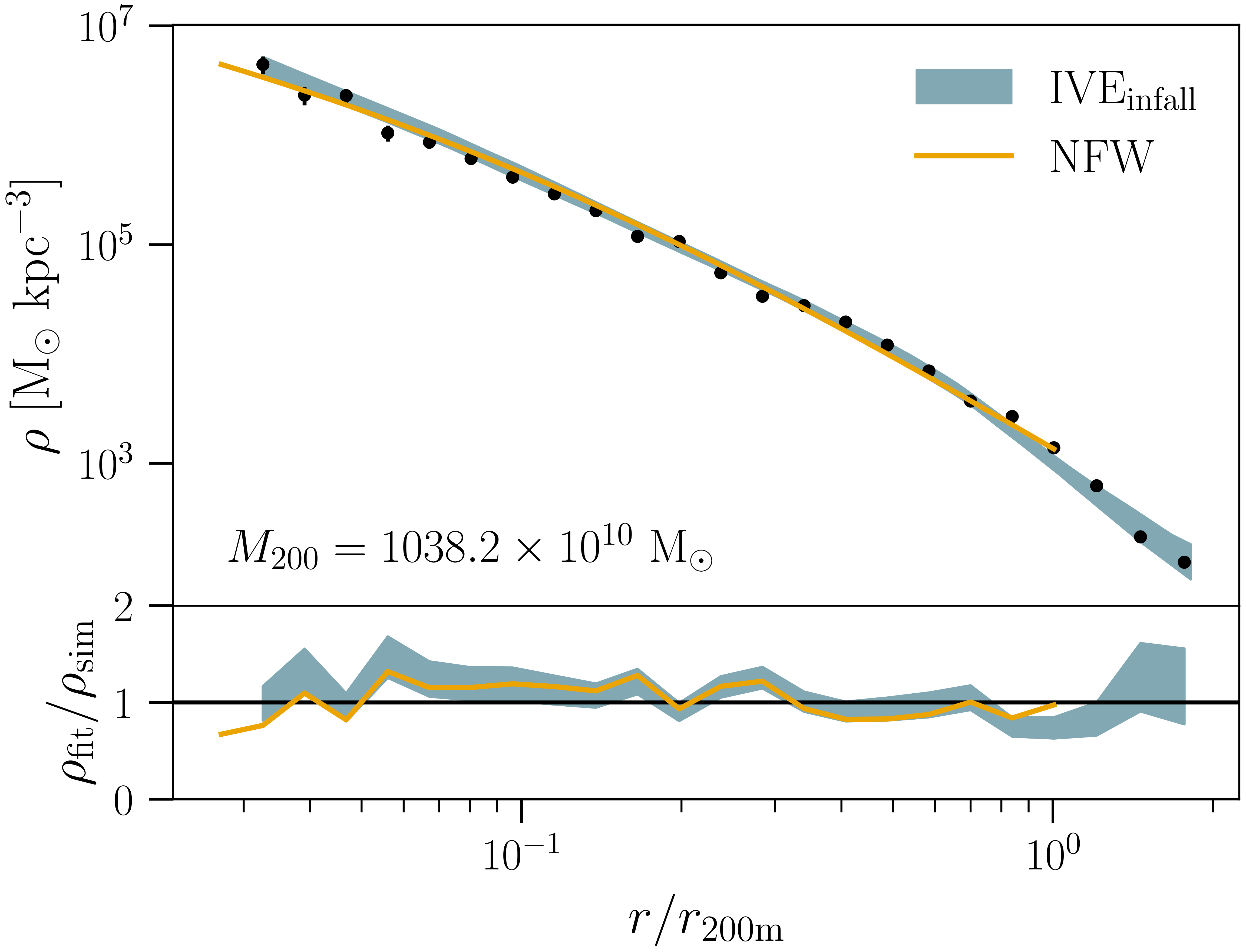
Case study: can neural networks discover the building blocks of dark matter halo profiles?



Lucie-Smith, Peiris, Pontzen, Nord + (2022)

Examples of fits created by interpretable variational encoder

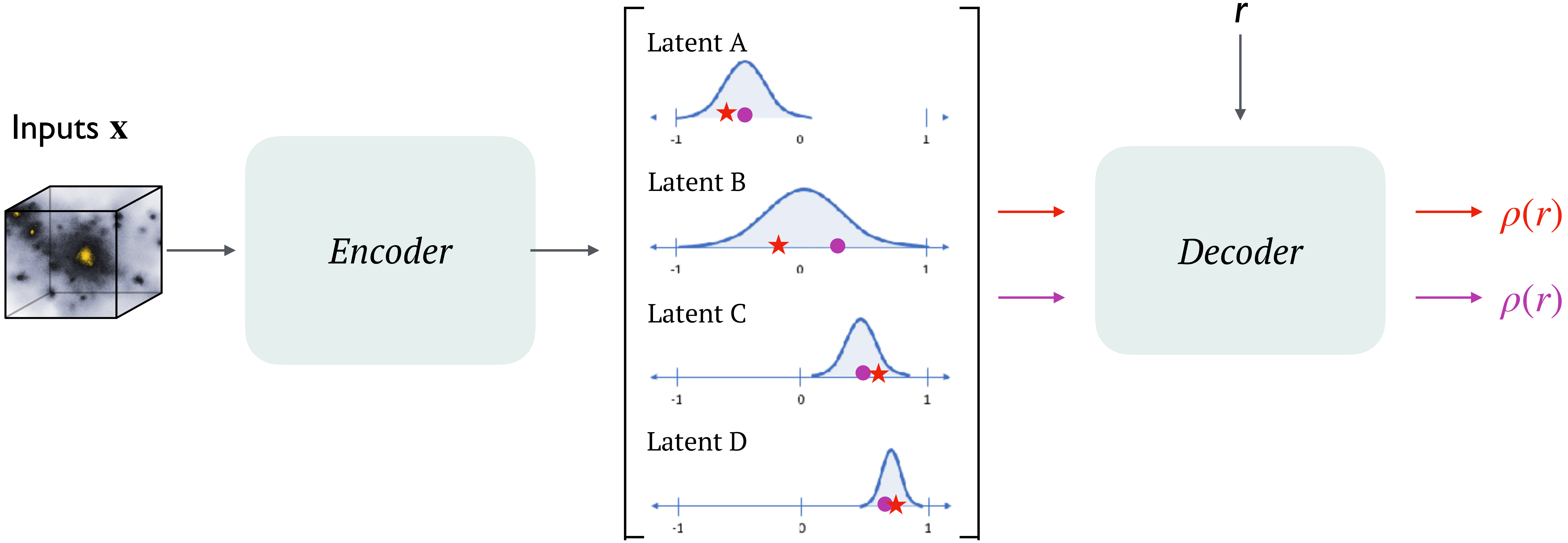




LUCIE-SMITH, PEIRIS, PONTZEN, NORD + (2022)

Desired latent representation properties for interpretability

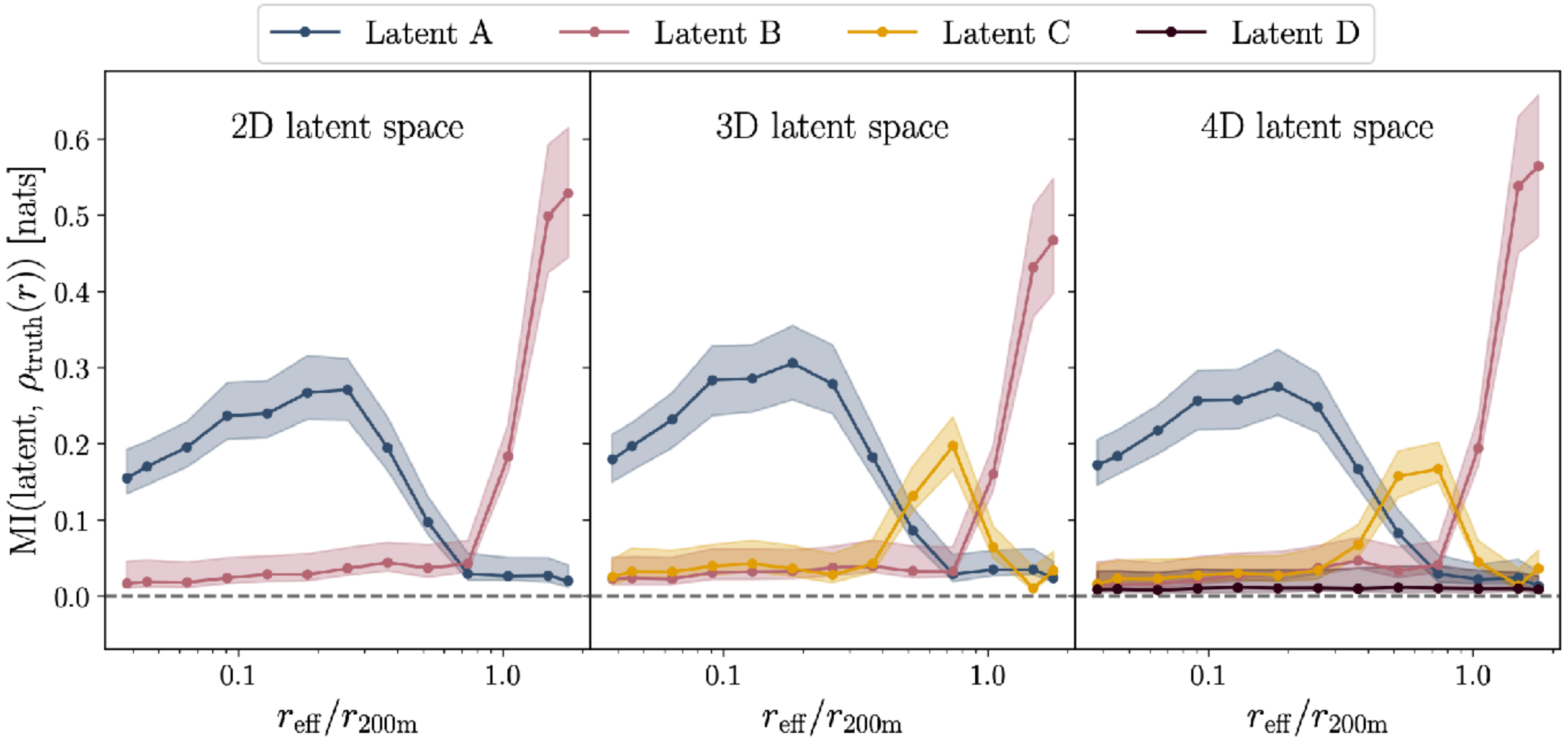
$$p(\mathbf{z}\,|\,\mathbf{x}) = \prod_{i=1}^{L} \mathcal{N}(\mu_i(\mathbf{x}), \sigma_i(\mathbf{x}))$$



- ***Interpretability*** *can be achieved if latent space is* ***disentangled**:* *independent factors of variation in profiles captured by different latents (“non-linear PCA”)*
- *Disentanglement encouraged via loss function optimised during training*
- *Degree of disentanglement measured using* ***mutual information*** *between latents*

Lucie-Smith, Peiris, Pontzen, Nord + (2022)

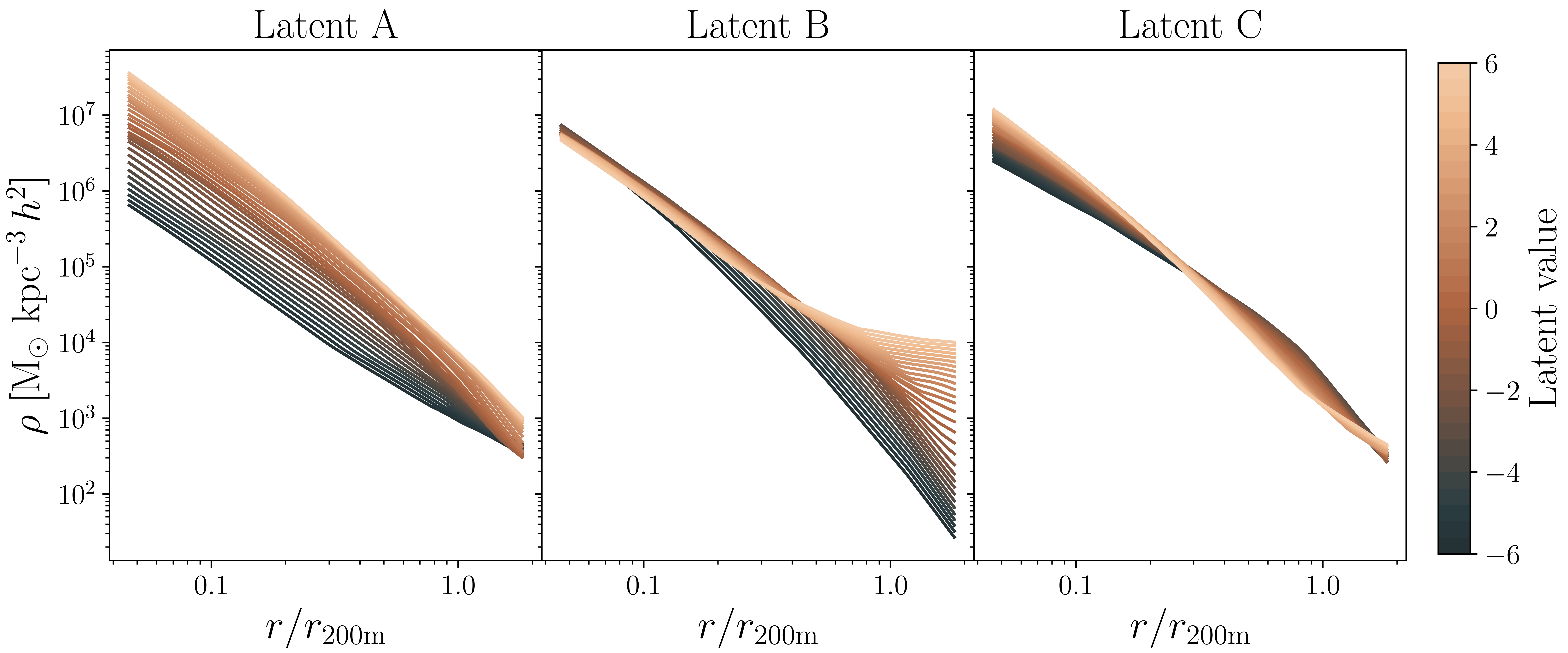
Interpreting the latent space using mutual information



***Explainability** achieved by evaluating mutual information between latents and ground truth density profile*

LUCIE-SMITH, PEIRIS, PONTZEN, NORD + (2022)

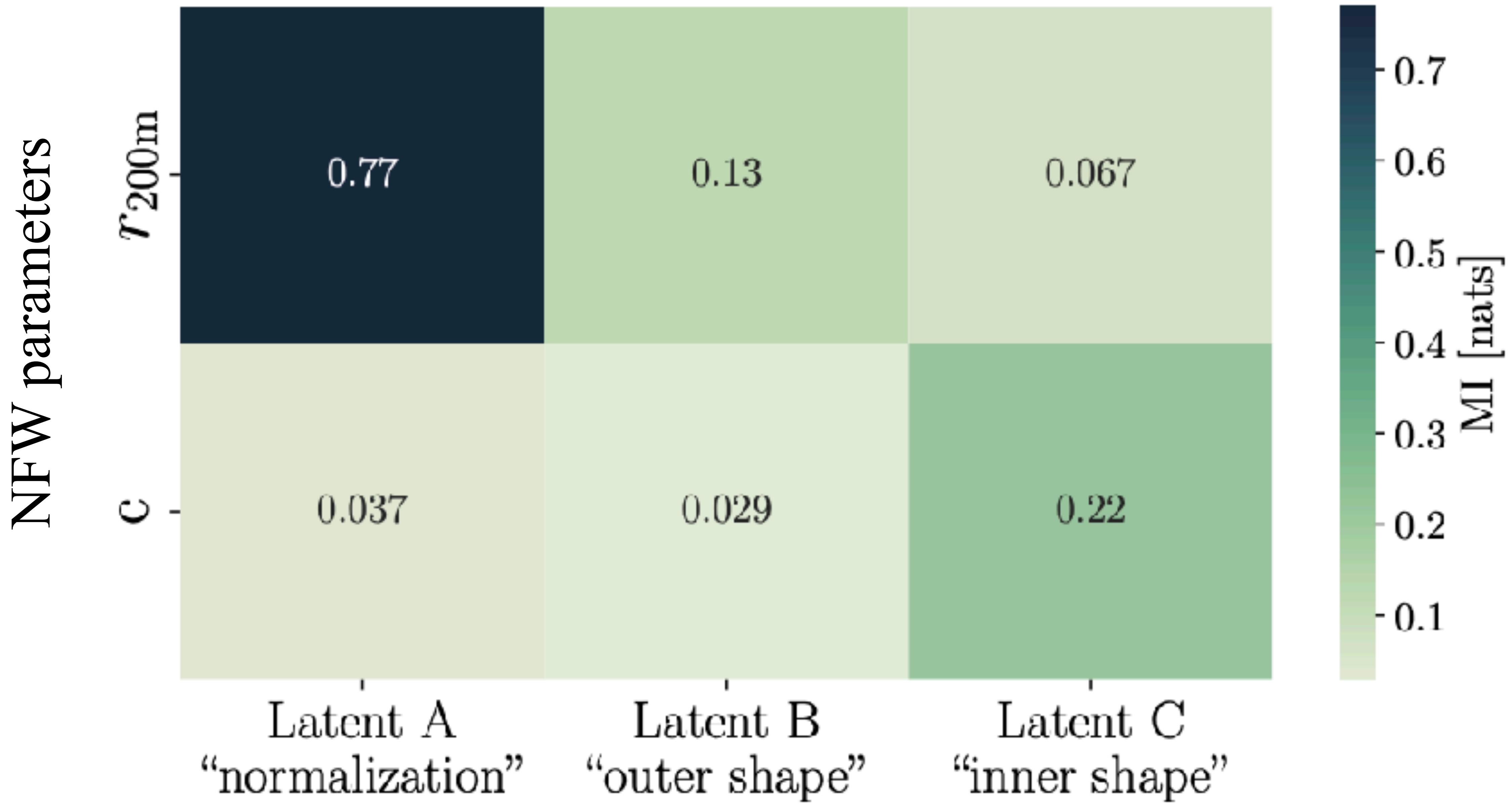
What has the machine learnt?

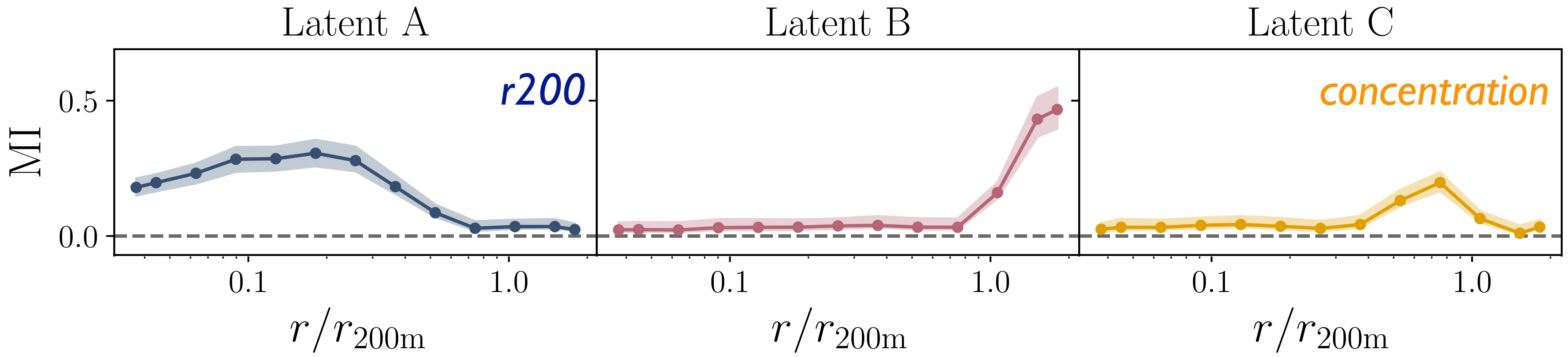


normalisation of profile　　*shape of outer component*　　*steepness of profile*

LUCIE-SMITH, PEIRIS, PONTZEN, NORD + (2022)

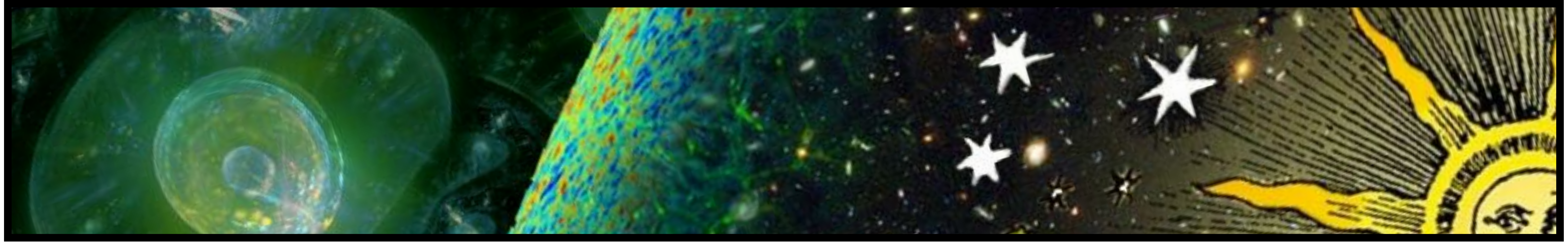
IVE re-discovers NFW parameters + additional "splashback" feature





IVE for model compression + mutual information for interpretability enabled ML-driven discoveries

LUCIE-SMITH, PEIRIS, PONTZEN, NORD + (2022)



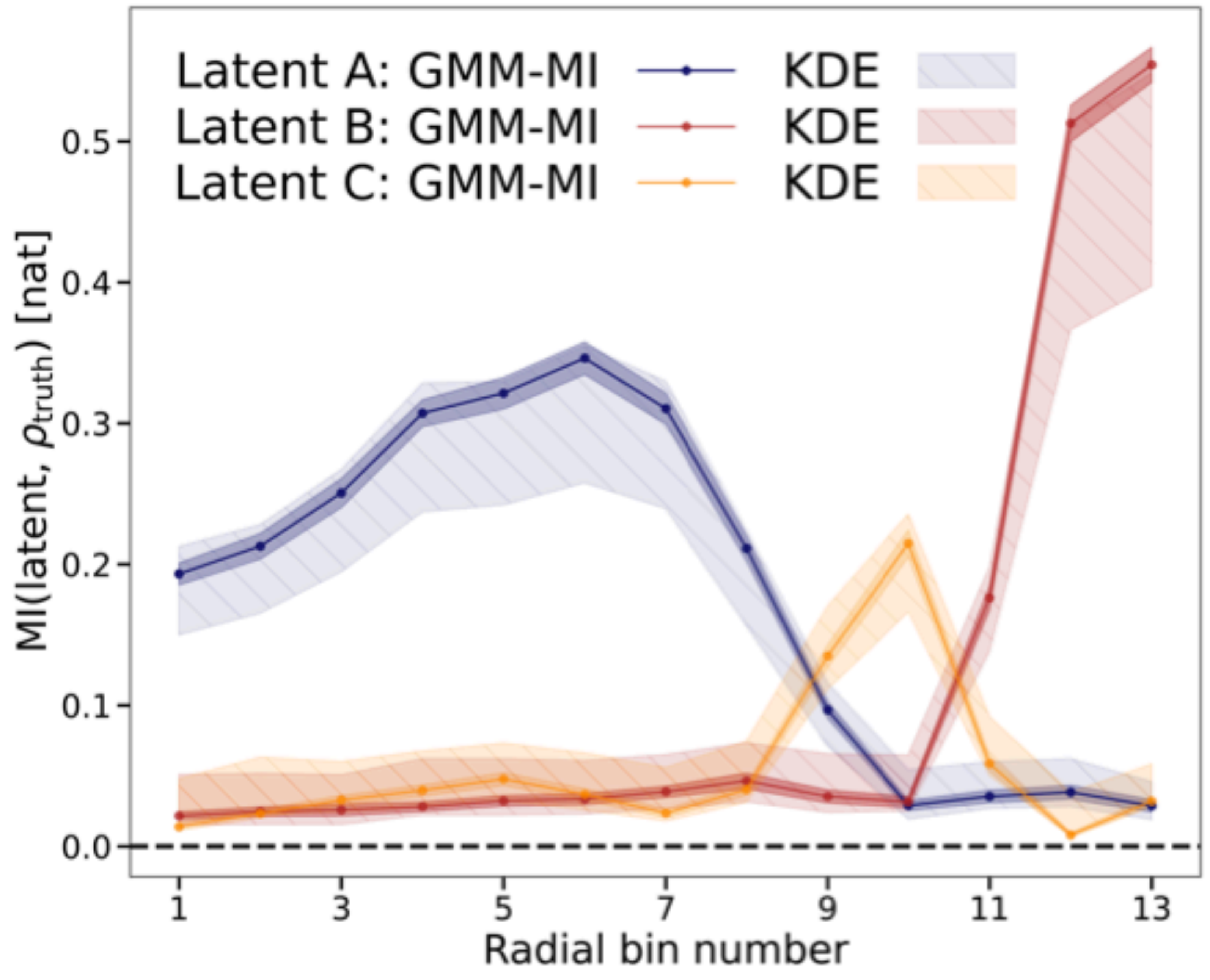
Estimating mutual information for deep-learning interpretability

Davide Piras
(UCL/Geneva)

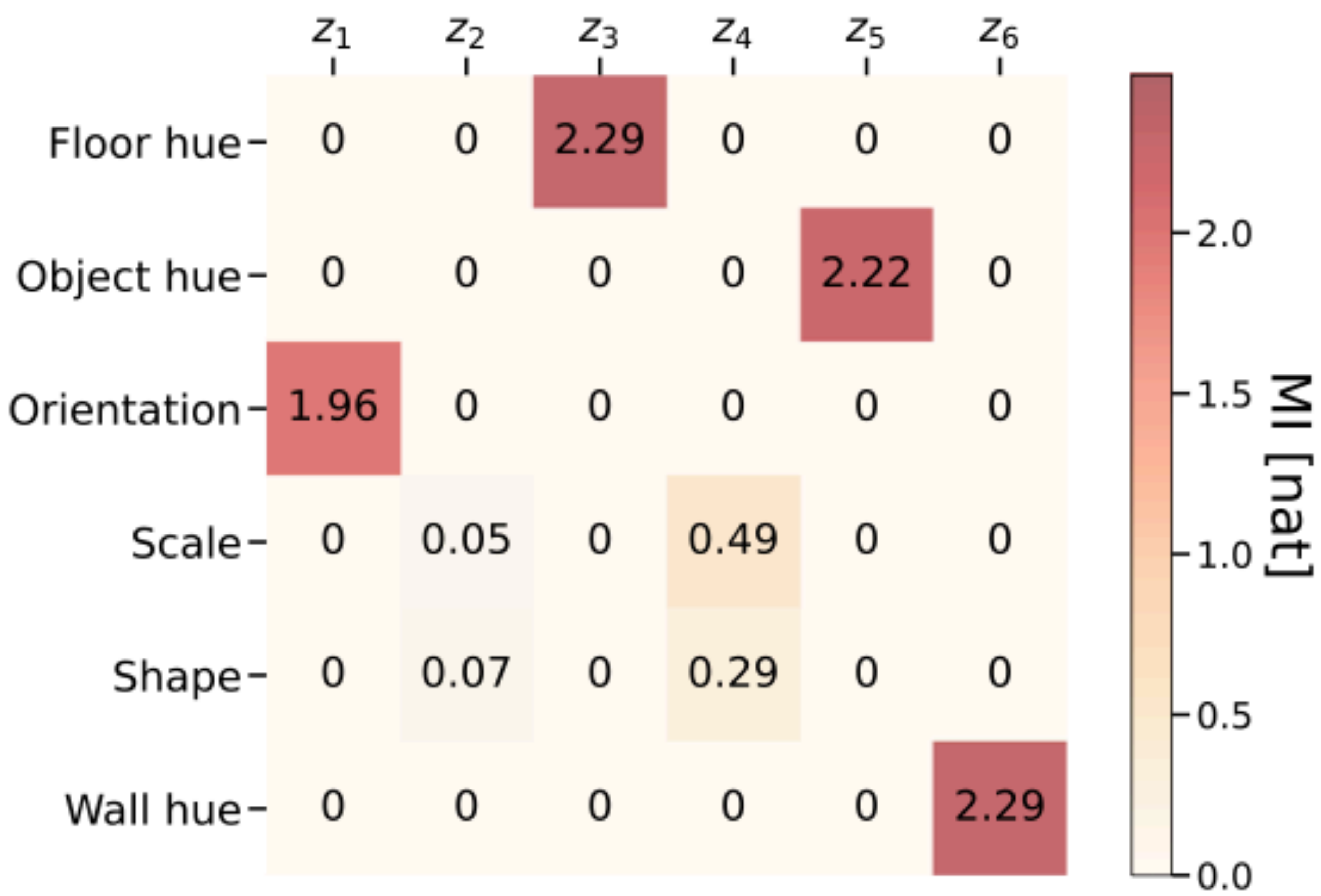
With: Andrew Pontzen, Luisa Lucie-Smith, Lillian Guo, Brian Nord

GMM-MI: accurate MI estimation with uncertainties

$$\mathrm{MI}(X,Y) = \int p(x,y) \log \frac{p(x,y)}{p(x)p(y)} \mathrm{d}x\mathrm{d}y$$



Continuous example: halo profiles



Categorical example: 3D Shapes

Density estimation based on Gaussian mixtures using samples from joint distribution;
Provides uncertainty due to finite sample size using bootstrap

https://github.com/dpiras/GMM-MI

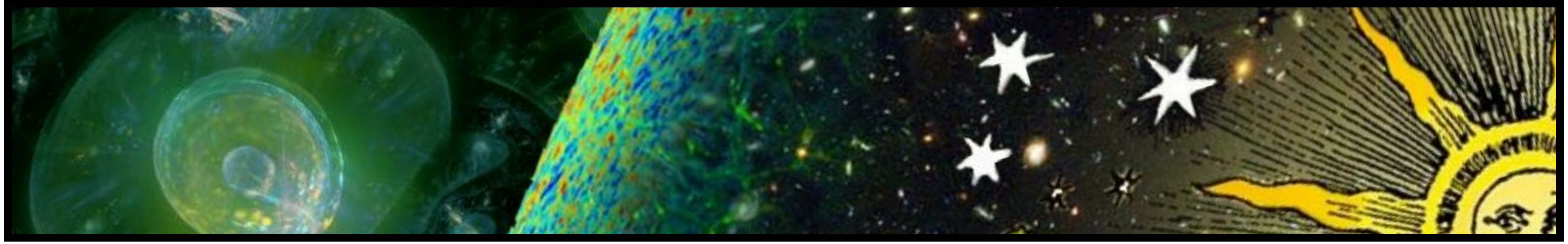
PIRAS, PEIRIS, PONTZEN, LUCIE-SMITH, GUO, NORD (SUBMITTED TO PRE, 2022)

erc
European Research Council
Established by the European Commission

CosmicExplorer: Exploring the Cosmos with the Vera Rubin Observatory

Aims: (i) AI-boosted modelling for cosmological analysis (ii) new cross-validation methods for diagnosis of systematics (iii) explainable AI to develop cosmic web as robust cosmological probe.

12 postdoc positions open at the Oskar Klein Centre Stockholm (okc.albanova.se)



Powerful methods available now to enable cosmology with complex, costly models.

Allows machines to take on the drudgery, leaving humans to focus on the physics.

Cosmoparticle, www.peneloperosecowley.com