# Cosmology before noon
# (large-scale structure at $2 < z < 6$)

Martin White (UCB/LBNL)

w/ Emanuele Castorina (Milan), Stephen Chen (UCB), Joe DeRose (Stanford), Simone Ferraro (LBNL), Nick Kokron (Stanford), Chirag Modi (CCA), Noah Sailer (UCB), Zvonimir Vlah (Cambridge), Risa Wechsler (Stanford), Mike Wilson (LBNL)

10 Feb 2021

# Outline

- Cosmology you can do before (cosmic) noon.
- (Future) Observational opportunities.
  - Return on current investments.
  - New directions.
- Theoretical frameworks.
  - Like the CMB, only better ...

# Next-generation science drivers

- In cosmology now we have a "standard model", based on General Relativity, inflation, dark matter (DM) and dark energy (DE).
- The model is stunningly successful, but completely phenomenological.
- We don't have a $1^{\mathrm{st}}$ principles understanding of much of the model.
  - ... or even a $2^{\mathrm{nd}}$ or a $3^{\mathrm{rd}}$ ...
- Need to test each piece to see what are only approximations, or perhaps what's "wrong" (test GR, inflation, DM and DE).

In the absence of a clear signal of new physics currently ... I will consider high-precision tests of the SM with a focus on large-scale structure (LSS; where some "tensions" have arisen)

# The "LSS program"

Probe metric, particle content and **both** epochs of accelerated expansion ... with high precision

- Expansion history and curvature
- Primordial non-Gaussianity ($f_{NL}^{\rm loc}$, $f_{NL}^{\rm eq}$, $f_{NL}^{\rm orth}$)
- Primordial or induced features, running of $n_s$
- Dark energy during MD
- DM interactions, light relics ($N_{\rm eff}$) and neutrinos

# Maximizing S/N

I want to maximize the S/N for new, BSM, physics

- There are many possible extensions to our SM ($\Lambda$CDM+GR).
- To my mind none are more compelling than others.
- If theory can't give us guidance, maybe phenomenology can?
    1. Work where inference is clean.
    2. Look where we haven't looked before (frontier!).
    3. If you don't know how to maximize $S$, then minimize $N$!

Push to higher redshift, in the epochs before cosmic noon!

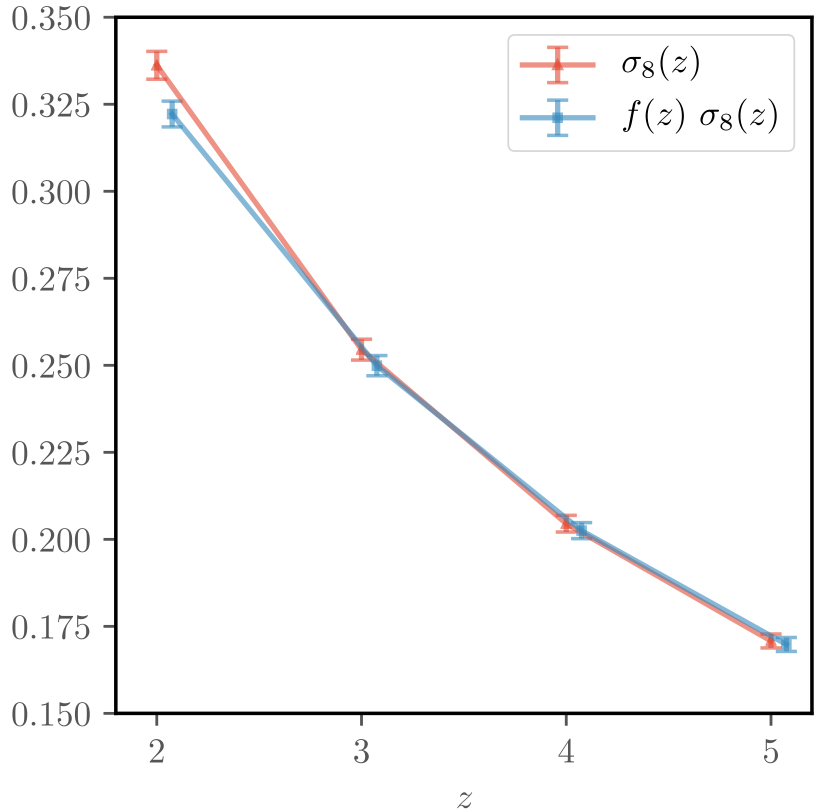
# Advantages of high $z$

Moving to higher $z$ gives us four simultaneous advantages:

1. Wide $z$ range leads to rotated degeneracy directions.
2. Larger volume.
   - More than $3\times$ as many "linear" modes in the $2 < z < 6$ Universe than $z < 2$.
   - Large volume $\Rightarrow$ small errors at "low" $k$, increased dynamic range to break degeneracies.
3. More linearity and correlation with ICs.
   - Get "unprocessed" information from the early Universe.
4. High precision theory.
   - Low $k$ modes are under good "theoretical control" using PT, little need for "nuisance parameter marginalization".
   - Everyone loves PT when you can use it – QED, Fermi liquids, CMB, ... LSS!
   - Theory becoming very advanced: lots of cross-fertilization with GR, CM and theory colleagues.

LSS at high-$z$ offers many of the advantages of CMB anisotropy!

# One example: growth rate



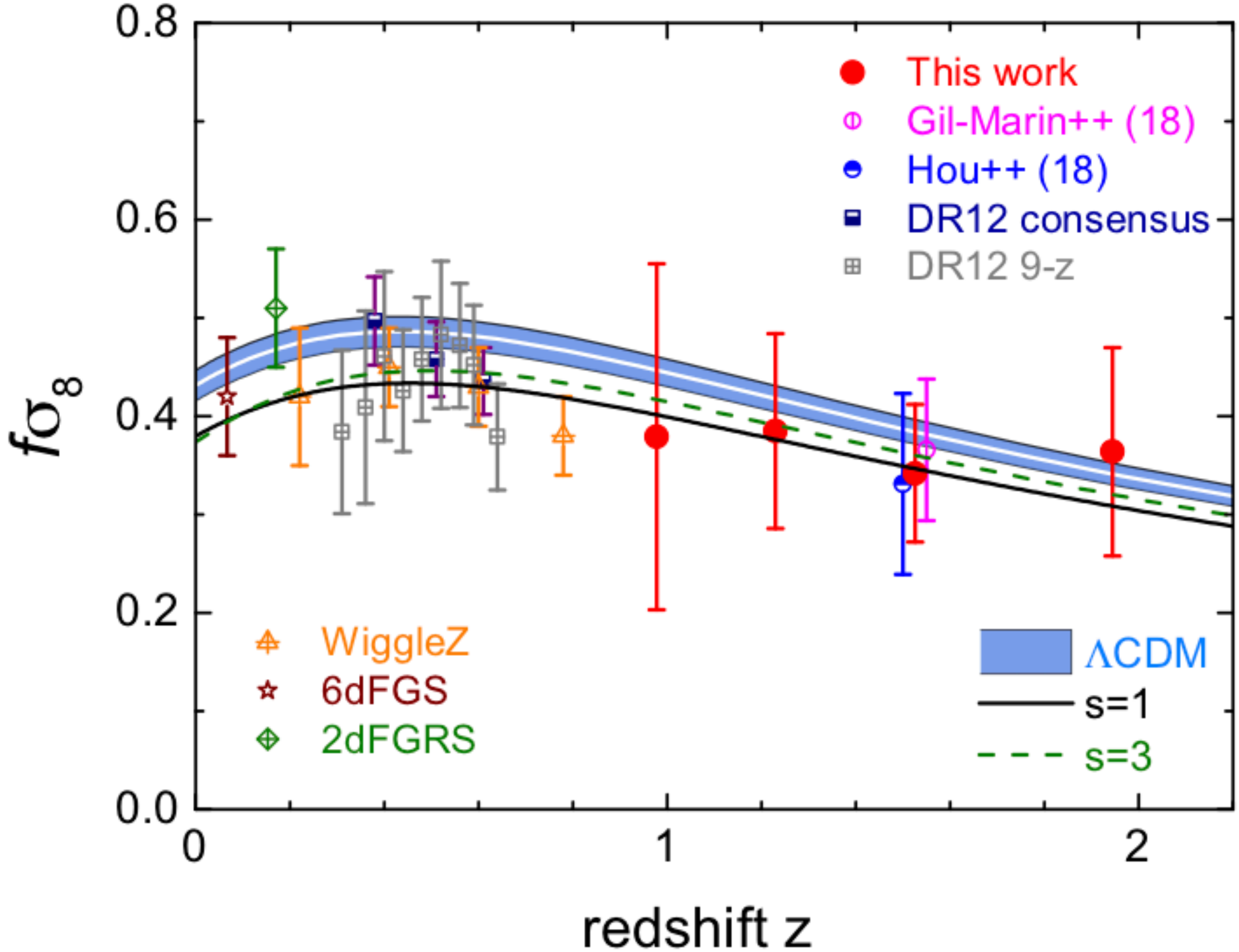

- Between $z \simeq 10^3$ and today, fluctuations grow by $\sim 10^3$.
- $\Lambda$CDM predicts growth very precisely.
- Marginalizing over unknown parameters, growth is predicted to 1.1% vs. $z$ (dominated by $m_\nu$ uncertainty).

Is the prediction (theory) right?

[Along the way test gravity model, expansion history, contents, ...]

# Growth rate

We are far from making a 1% test ... (Zhao+19)




Getting more volume (higher redshift) would help ...

# What probes of the $2 < z < 6$ Universe will we have?

Continuous advances in detector technology and experimental techniques are pushing us into a new regime, enabling mapping of large-scale structure in the redshift window $2 < z < 6$ using both relativistic and non-relativistic tracers ...

# CMB = lensing at high $z$

We are witnessing a rapid scaling up of CMB experimental sensitivity as we move into the era of million-detector instruments!

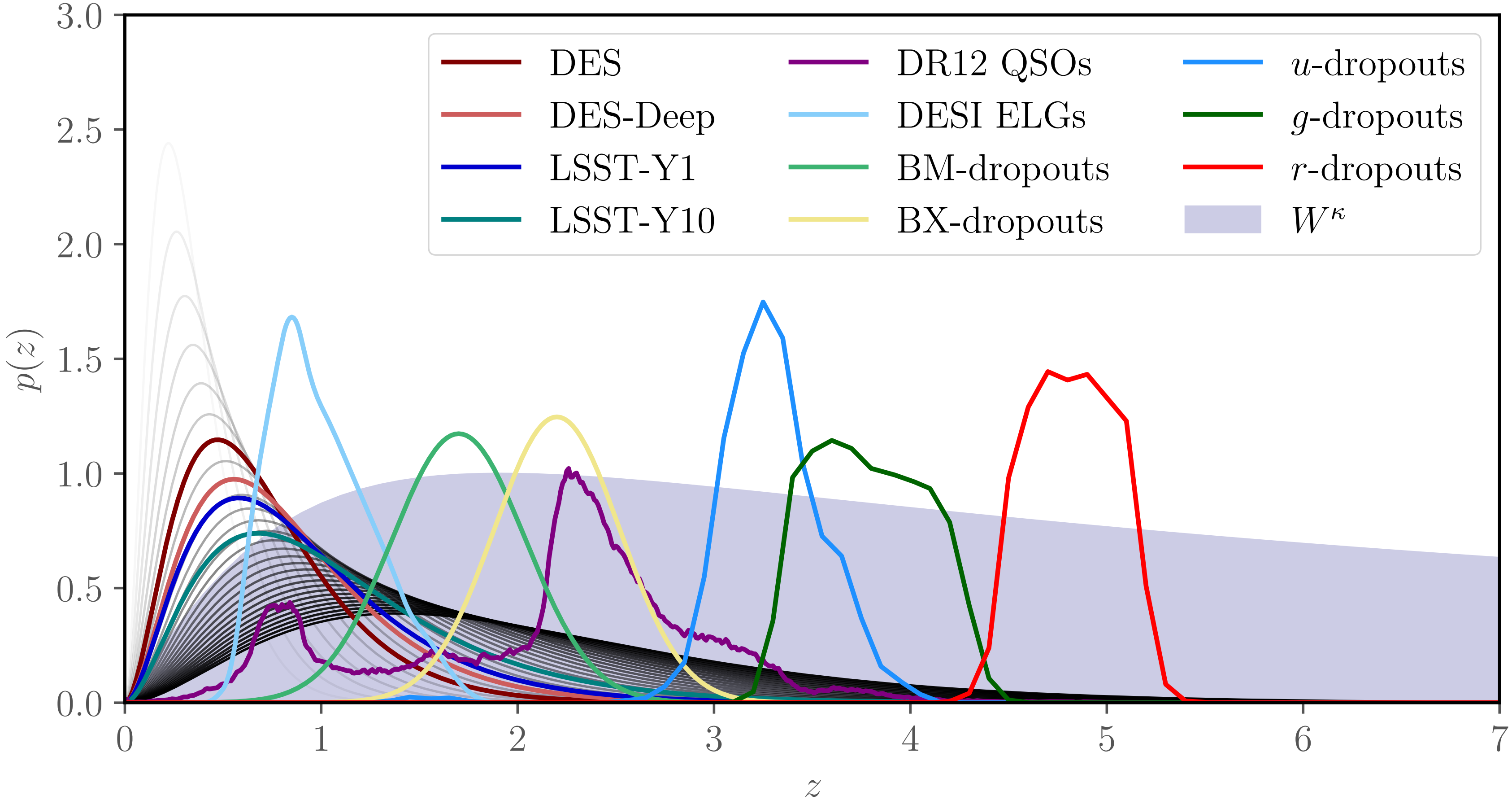
- A natural "by-product" of next generation CMB surveys to constrain primordial gravitational waves is high fidelity CMB lensing maps – probing the matter back to $z \simeq 1100$.
- Lensing is sensitive to mass, not light, and by using a relativistic tracer it gives access to the Weyl potential.
- But lensing is projected ...
- ... want to do cross-correlation with samples of known redshift.
- Lensing + galaxy surveys offer redshift specificity, higher S/N and lower systematics. Natural synergies: total greater than sum of the parts!

The promise of cross-correlations is that they enable new science as well as increased robustness of the core science of each project!

# Tracers of LSS at $2 < z < 6$

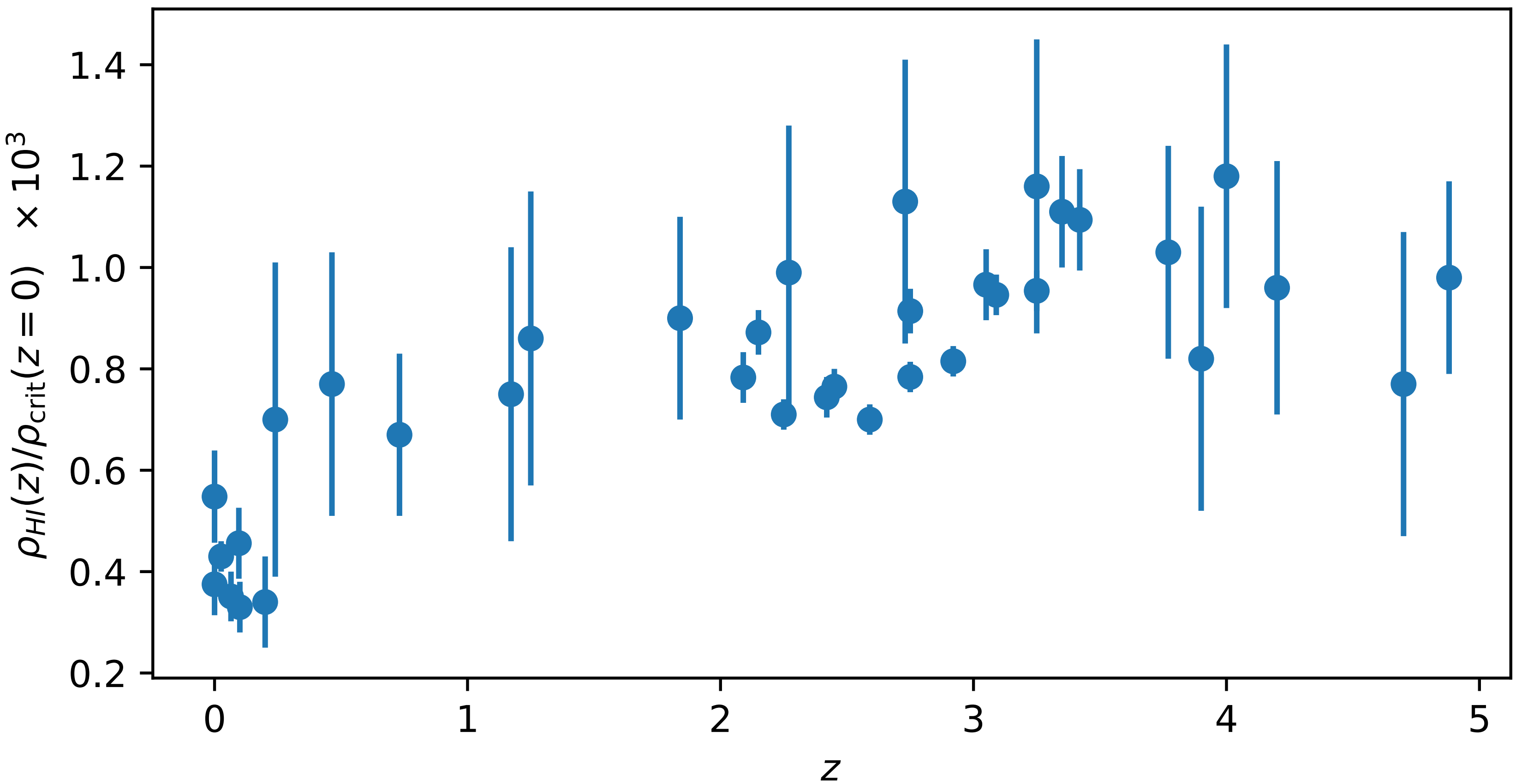
- There are lots of galaxies at high $z$, and we have pretty efficient ways of selecting them.
  - Dropout, or Lyman Break Galaxy (LBG) selection targets the steep break in an otherwise shallow $F_\nu$ spectrum bluewards of 912Å.
  - These objects have been extensively studied (for decades!).
  - Selects massive, actively star-forming galaxies – and a similar population over a wide redshift range.
  - LBGs lie on the main sequence of star formation and UV luminosity is approximately proportional to stellar mass.
  - A fraction of these objects have bright emission lines (LAEs).
- BBN $\Rightarrow$ there's lots of Hydrogen as well!
  - Hyperfine (mag. dip.) transition of H I ($p + e$ spin-spin coup.)
  - Very rare transition per atom ($\propto \mu^2/\lambda^3$); little absorption or confusion (no line at 710 MHz!).
  - Basically measuring "DLAs" or "HCD systems" ...

# Galaxies over the whole range




Wilson & White (2019)

# And lots of neutral Hydrogen



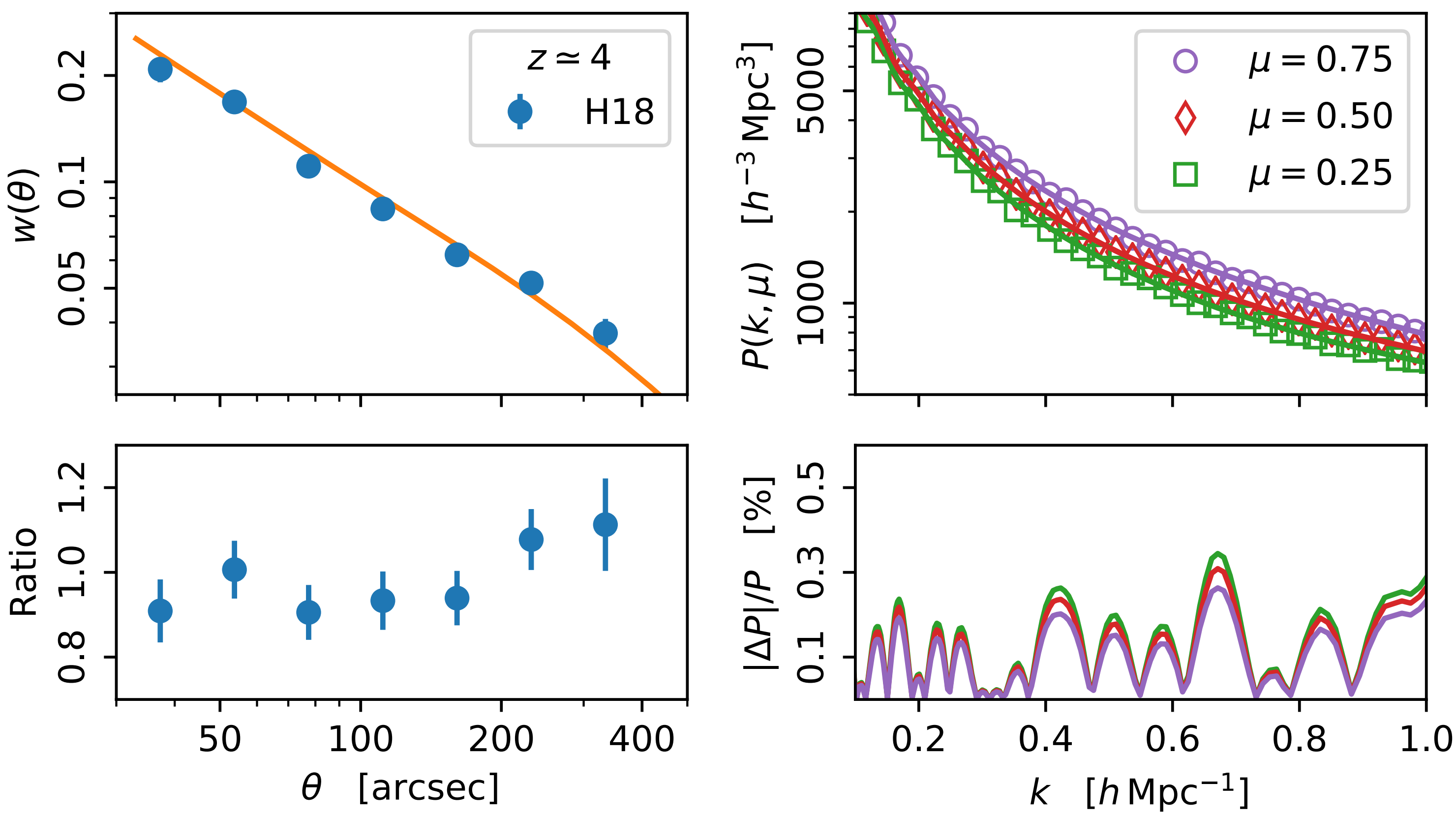

Compilation from Modi+21

# High precision

What could we do with such data?

# High precision

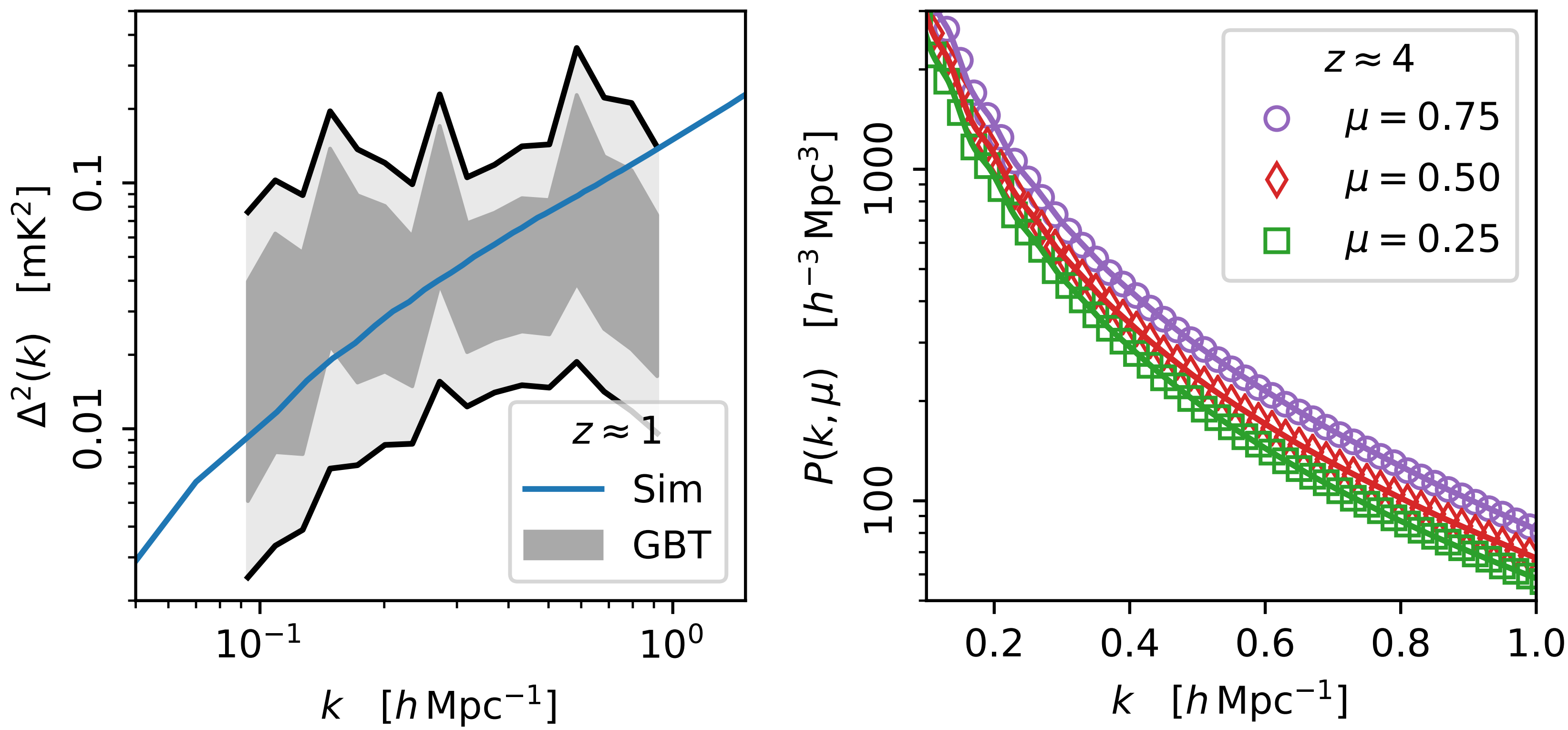
Out-of-the-box comparison of two, public, theory modeling codes




Over half the sky, within $3.5 < z < 4.5$ there are over a billion modes out to $k = 1\,h\,\mathrm{Mpc}^{-1}$!

# High precision

There's nothing special about galaxies here ... HI would work too!



# Models of large-scale structure (LSS)

How do people model measurements of large-scale structure?

- There are two broad classes of approaches to modeling LSS: numerical and analytical.
- Numerical approaches (simulations)
  - Techniques for solving systems of pure dark matter are well developed; though the combination of volume and resolution required by next-gen surveys is very demanding.
  - The best way to deal with the complexities of galaxy formation, hydrodynamics and multiple species is still an open research problem.
- I will discuss analytic approaches based on perturbation theory (PT) – which have seen a renaissance in recent years.
- Most practitioners use some combination ...
  - All N-body codes use PT for initial conditions.
  - N-body can be used to test PT for fiducial models.
  - New ideas for combining the two: "best of both worlds".

# A funny thing happened ...

- Cosmology is riding the Moore's law/big data revolution like many other fields.
- Even though computing/simulation is becoming a bigger component of the analysis toolkit, modern surveys are empowering theorists as never before ...
- We have the technology to survey very large volumes at larger distance (i.e. earlier times).
  - Fluctuations are linear, or quasi-linear ($\delta \lesssim 1$).
  - Such modes are under good "theoretical control" using PT.
  - We're now computing small corrections to "almost linear" quantities; a regime in which PT excels.
  - Bigger surveys demand higher precision: "almost" isn't good enough.

# PT: two flavors

Eulerian (standard)
Treat cold dark matter as a pressureless (perfect) fluid obeying

$$\begin{aligned} \partial_\tau \delta + \nabla \cdot [(1+\delta)\mathbf{v}] &= 0 \\ \partial_\tau \mathbf{v} + \mathcal{H}\mathbf{v} + \mathbf{v}\cdot\nabla\mathbf{v} &= -\nabla\Phi \end{aligned}$$

with the $\mathcal{H}v$ term being "Hubble drag" arising from the expansion of space.

Lagrangian
Treat cold dark matter as a collisionless system

$$\mathbf{x}(\mathbf{q}) = \mathbf{q} + \mathbf{\Psi}(\mathbf{q}, \tau)$$

with

$$\partial_\tau^2 \mathbf{\Psi} + \mathcal{H}\partial_\tau \mathbf{\Psi} = -\nabla\Phi\left(\mathbf{q} + \mathbf{\Psi}(\mathbf{q})\right)$$

then derive density from

$$1+\delta(\mathbf{x}) = \int d^3q\, \delta^{(D)}\left[\mathbf{x} - \mathbf{q} - \mathbf{\Psi}(\mathbf{q})\right]$$

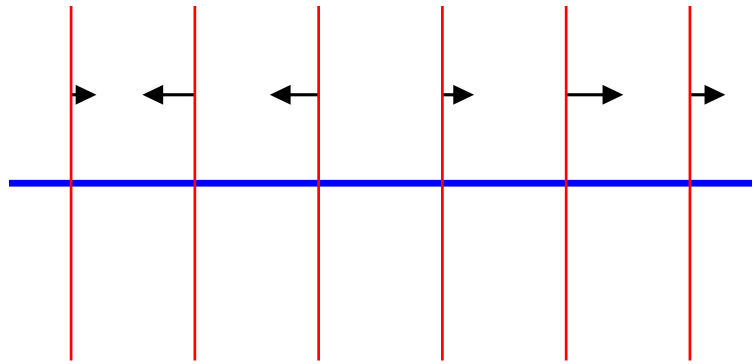
(Both derivable from the Vlasov equation)

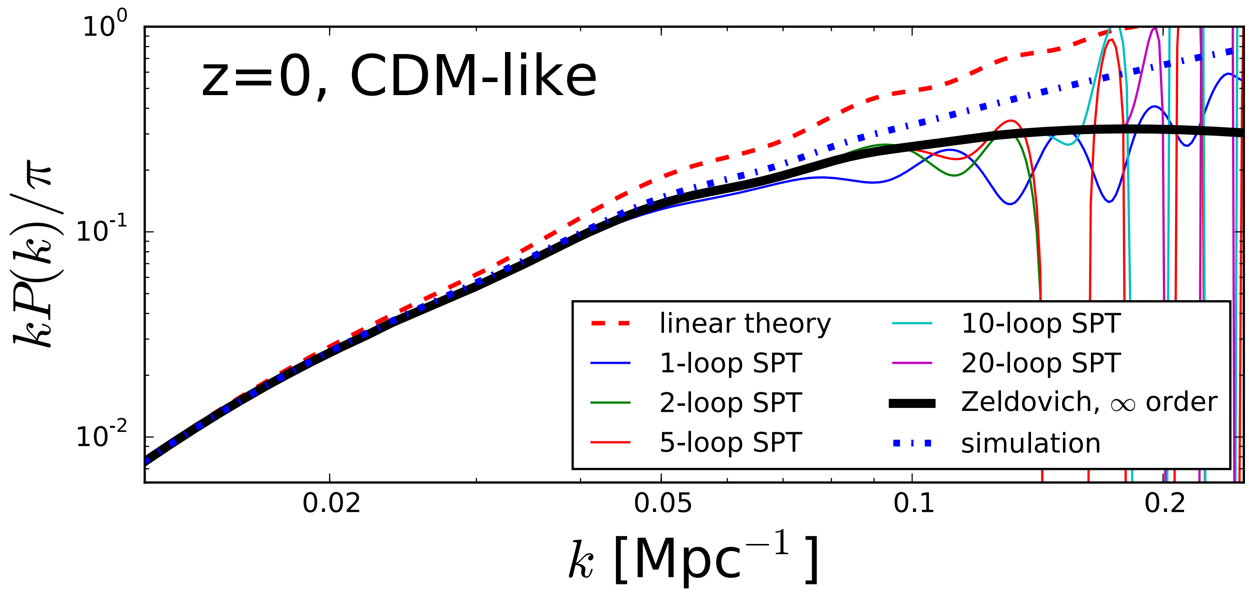
# A problem emerges

- These two approaches give the same predictions, order by order in perturbation theory.
- This sounds good, but actually ... **this indicates a problem!**
- Two frameworks for PT describe different systems:
  - pressureless fluid (Eulerian) and
  - collisionless fluid (Lagrangian),
- Think what happens when streams connect: shocks vs. caustics.

# A toy model

Consider a collection of uniform, parallel, 2D sheets of matter moving normal to the sheets under gravity.






McQuinn & White (2016)

# UV: effective field theory

This problem (and solution) is well known in many areas of physics!

- EOM are non-linear, so have "composite" terms like $v\delta$.
- Products in configuration space become convolutions in Fourier space.
- In Fourier space $\delta^{(2)}(k) \sim \int dk' \, \mathcal{K} \, \delta^{(1)}(k-k')\delta^{(1)}(k')$
- But $\delta^{(1)}(k')$ is not small for high $k'$: PT breaks down.
- Need to regularize and introduce counter terms.
  - In Eulerian PT the lowest order counter term looks like a pressure force.
  - Lagrangian PT looks like a multipole expansion of extended objects – how they respond to low-$k$ potentials and tides.

# Bias, peaks and EFT

## But what about galaxies?

- Write $\delta_{\rm gal}$ as a functional of the initial (long wavelength) density, velocity and potential fields: $\delta_{\rm gal}[\delta, \partial \mathbf{v}, \partial\partial\Phi, \cdots]$
- Coefficients of an expansion in e.g. $\delta$ are bias coefficients.

$$\delta_{\rm gal}(\mathbf{x}) = b_1\delta(\mathbf{x}) + b_2\delta^2(\mathbf{x}) + \cdots + \text{stochastic} + \cdots$$

- Bias coefficients incorporate our uncertainty about complicated galaxy formation physics in addition to UV effects.
  - Dark matter halo formation, merger history, ...
  - Chemistry and gas cooling.
  - Star formation, SNe, AGN
  - Thermal and kinetic feedback
  - Background radiation

# Bias, peaks and EFT

- **While the processes that form and shape galaxies and other objects are complex, all such objects arise from simple initial conditions acted upon by physical laws which obey well-known symmetries.**
- For non-relativistic tracers these are
  - the equivalence principle
  - translational, rotational and
  - Galilean invariance.
- This highly restricts the kinds of terms that can arise in a bias expansion, no matter how complex the history.

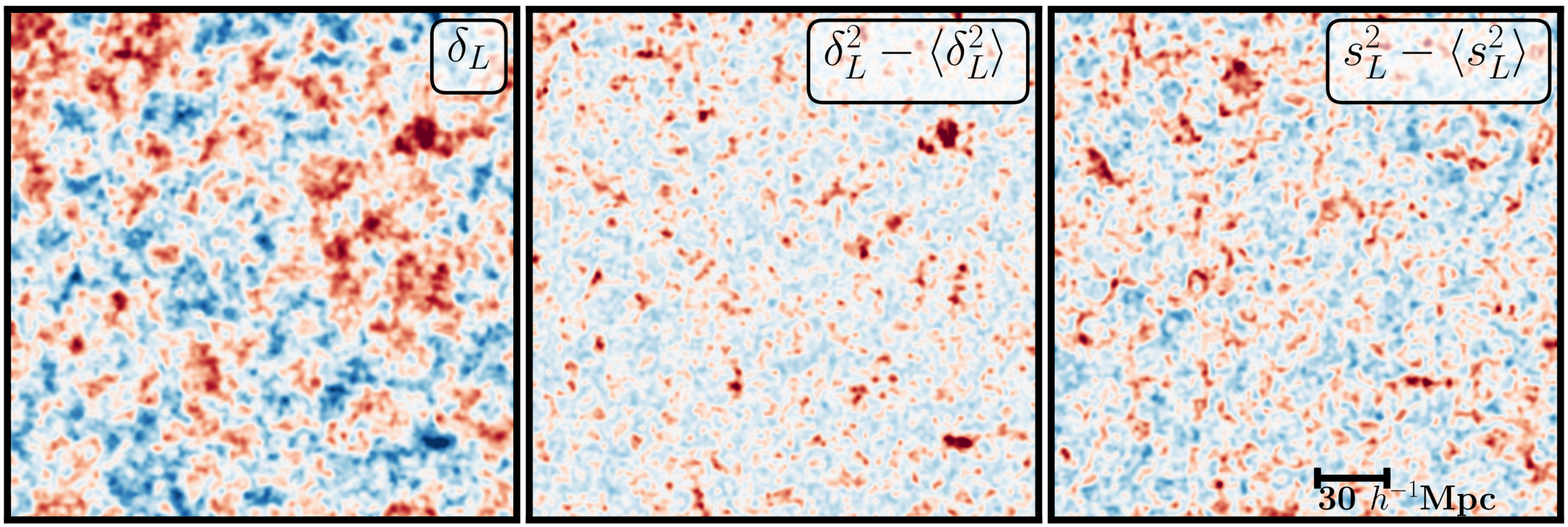
Symmetry arguments are extremely powerful for bias since we don't understand the small-scale physics of bias.

# Aside: Simulations and Symmetries

- We can simulate structure formation in a DM-only Universe pretty well.
- It's the baryonic component that is "hard"!
  - Don't understand cooling, star-formation, feedback, ...
  - Resort to parameterized models (when to stop adding parameters, how to test for numerical convergence?)
- Symmetries-based thinking is ubiquitous in PT studies and very powerful.
- PT folks and simulators are trying to solve the same problems ...
- Can we have the best of both worlds?
  - Use dynamics from N-body simulations, but the "galaxies" (symmetries-based bias technique) from perturbation theory [Modi+20].

# The procedure in pictures

Generate initial conditions as per usual ... from $\delta_L$ you can also compute $\delta_L^2$ and the shear field, $s_{ij}$:

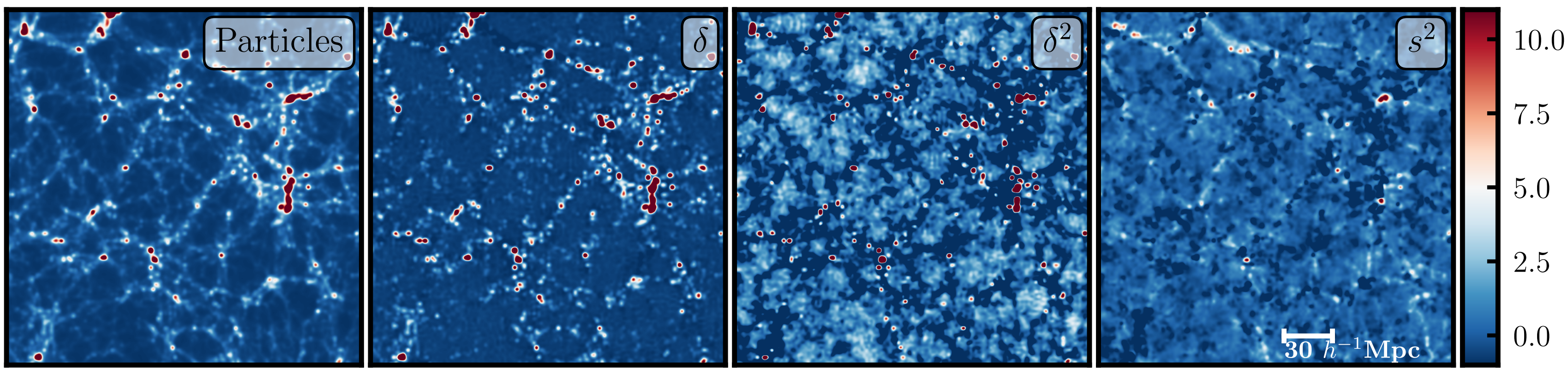



Each particle is assigned the $\delta_L$, ... at its initial position.

Kokron+21

# The procedure in pictures

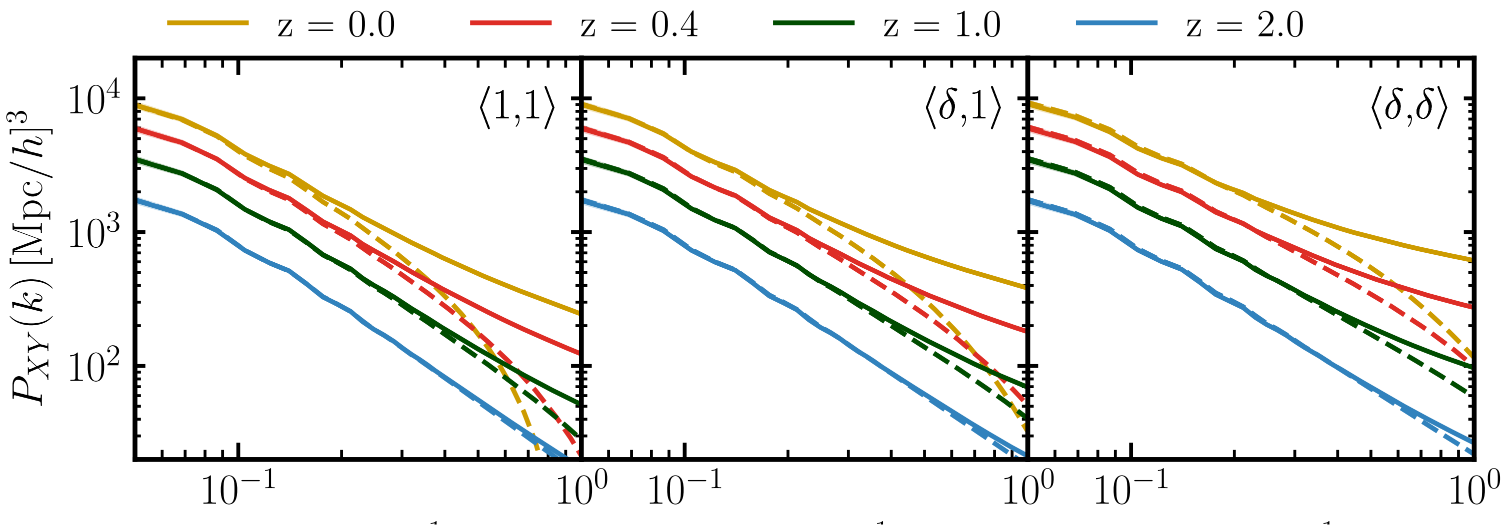
Advect the particles to their final positions using the full N-body dynamics (i.e. run the simulation), and bin using weights 1, $\delta_L$, $\delta_L^2$, etc.




(No need for halo or subhalo finding, merger trees, etc.)

# The procedure in pictures

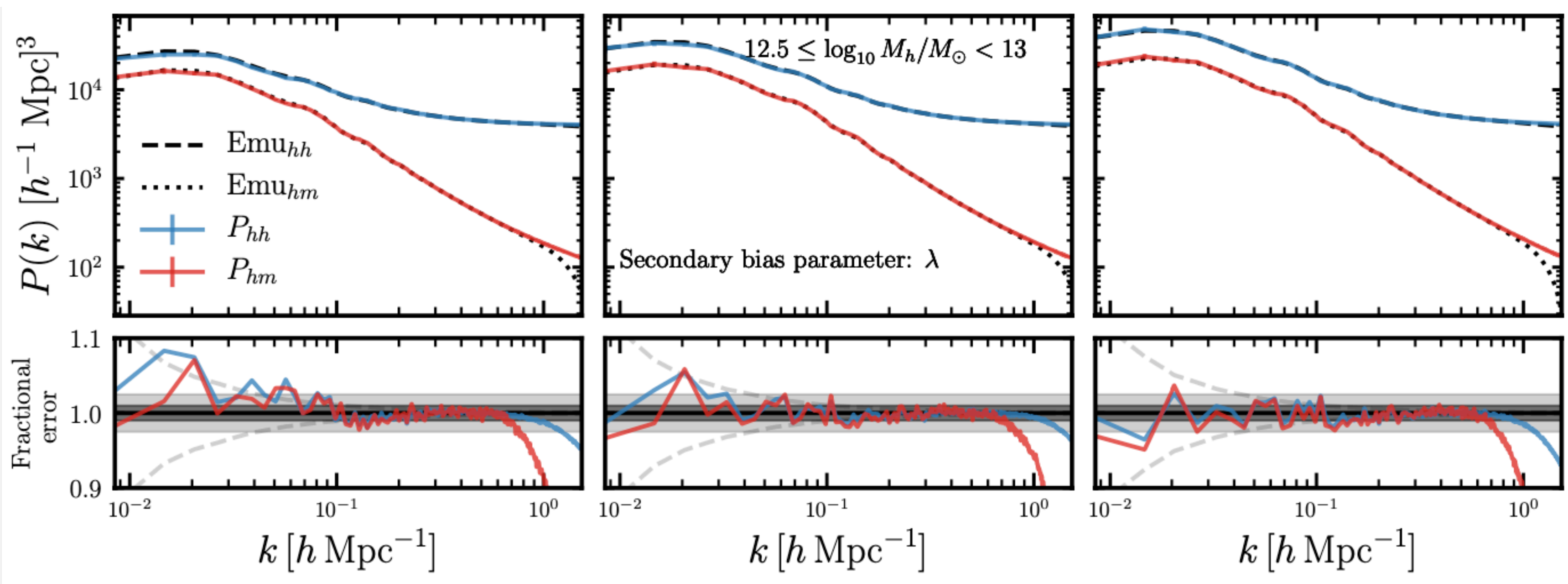
Take all of the cross-spectra, $P_{XY}(k)$ using standard FFT methods, e.g.




The power spectrum for any biased tracer, or the cross-spectrum between any two tracers, is a linear combination of these "basis spectra" (10 in all) with analytic "bias dependence": $\sum_{ij} b_i b_j P_{ij}$.

# The Aemulus emulator

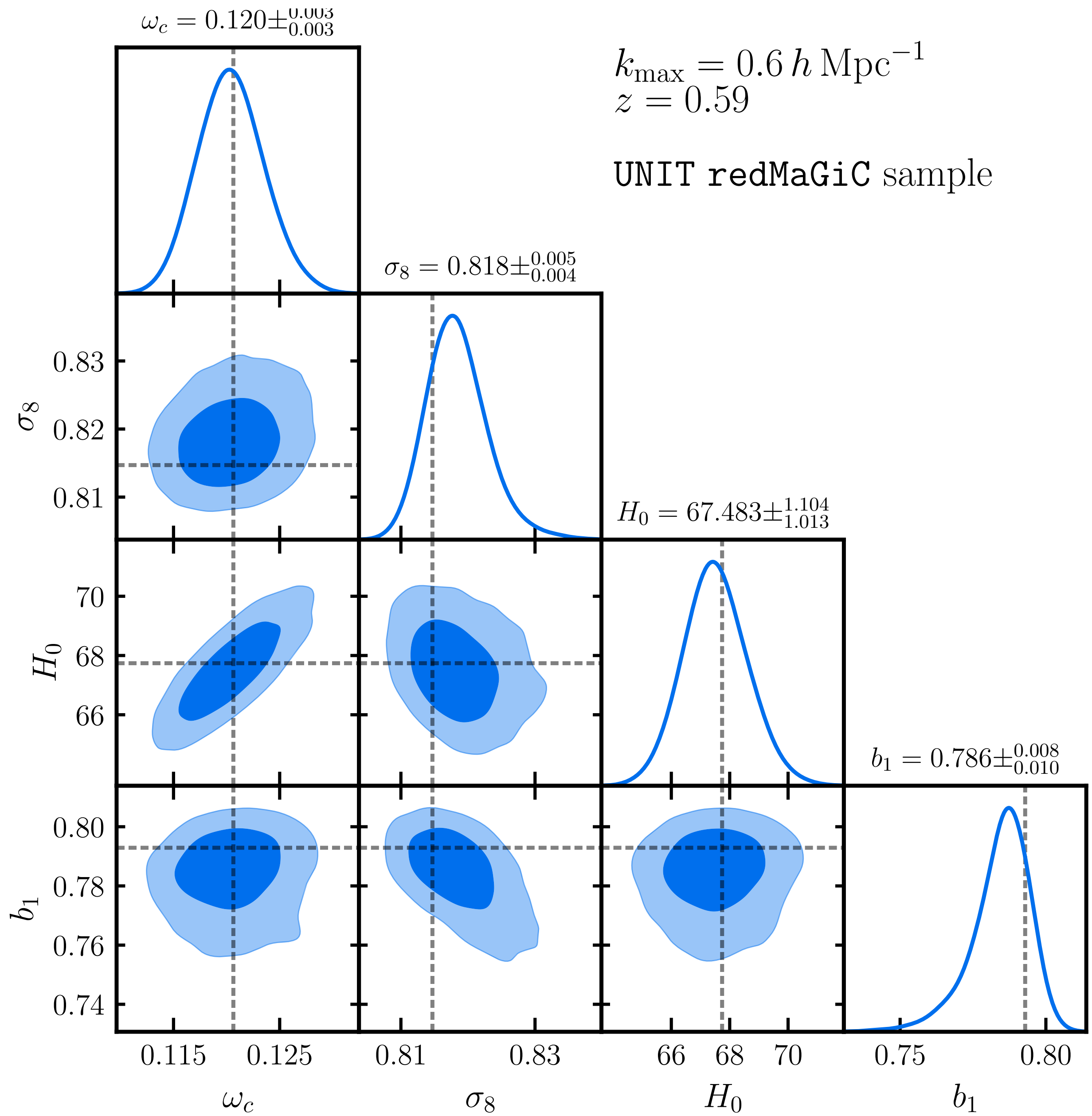
Can fit mock catalog data for "3 × 2pt analyses" to 1-2% even for samples with assembly bias and other complex selections and even including hydrodynamics.




Now we can simply "emulate" the basis spectra using standard techniques (no need to emulate the bias parameters – analytic)!

# Cosmology from a noise-free “RedMagic” sample

Cosmological constraints from a fit to noise-free mock data (from a <u>different simulation</u> at a <u>different $z$</u>!) returns unbiased constraints.



# Aside: a new, hybrid technique

- Use dynamics from N-body simulations, but the "galaxies" (symmetries-based bias technique) from perturbation theory.
  - Use the strengths of each approach.
  - Systematic procedure and controlled approximations, well defined notions of convergence, ...
- Relaxes demands on N-body.
- Dramatically reduces the dimension of the emulation problem, automatically includes assembly bias and baryonic effects.
- This can be used to produce power spectra (as above) but it can also generate all of the polyspectra.
- Since it works at the field level, it can also be combined with new forward modeling techniques.

# Back to PT

Back to perturbation theory ...

# Velocileptors

- We have developed schemes to handle small-scale physics with a complete set of counterterms
  - Also handles complex FoG models and redshift errors!
- Do resummation of the long-wavelength displacements that are so important for getting the BAO peak "right".
- At 1-loop result is an integral over integrals of 1 or 2 powers of $P_{\rm lin}(k)$ times kernels.
- In redshift space there aren't many symmetries to exploit, so you need to deal with the multidimensional integral.
- Want to include and organize the terms in a way that makes the integrals accurate, stable and fast to evaluate.
- We use FFTs and hypergeometric functions, which leads to very efficient numerical evaluation.
- Can evaluate $P(\mathbf{k})$ at $\mathcal{O}(10^2)$ $k$-values in under 1 s on my 10 year old Mac laptop!

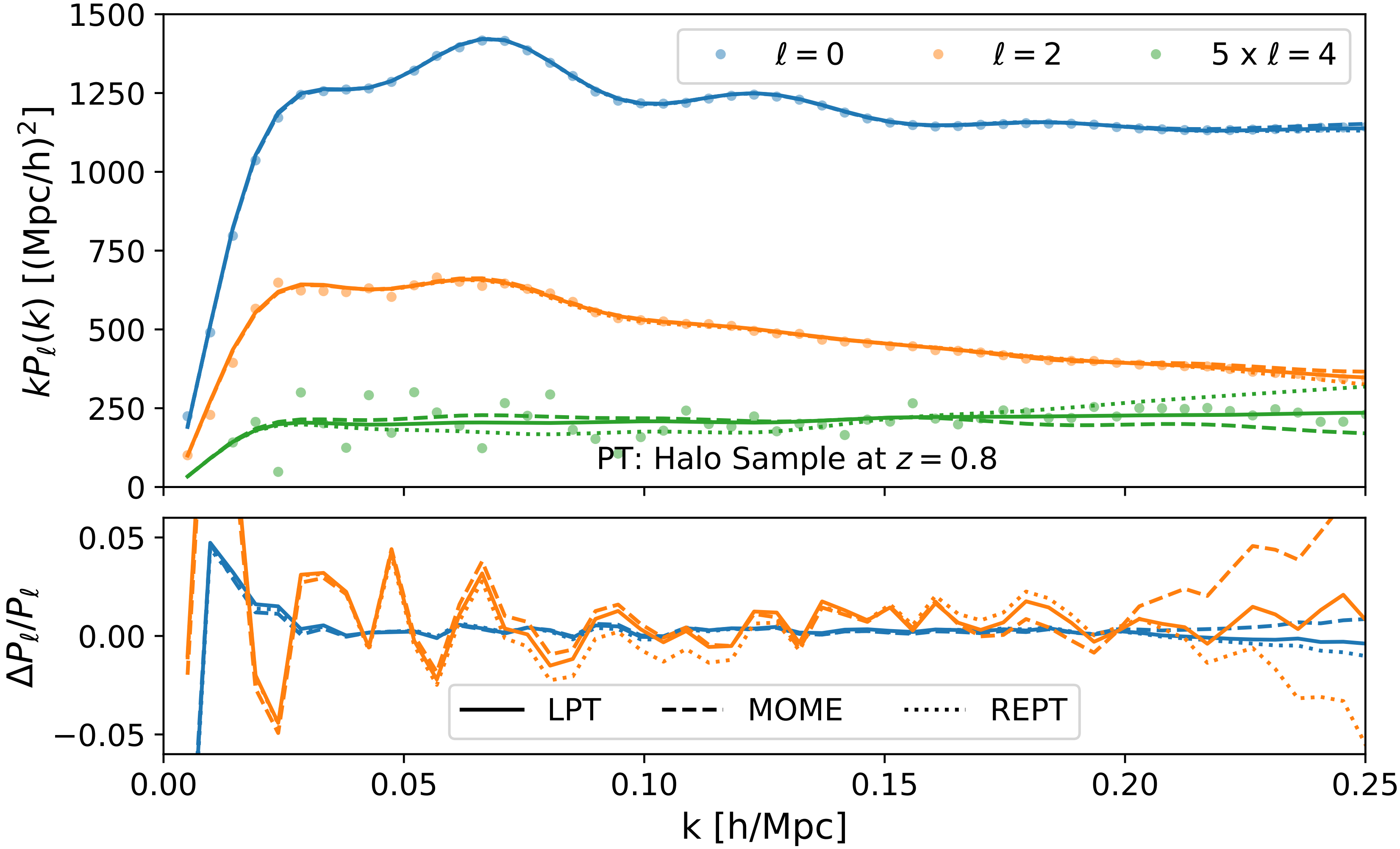
# Velocileptors

- We have a public, (pure) Python package for these models.
- Being used in a number of surveys and data analyses now.
- Many ways to combine velocities and densities in power spectra: direct PT expansion, moment expansion, Gaussian streaming model, Fourier streaming model.
- Available in both LPT and EPT variants (allowing cross-checks!)
- Fourier and configuration space, auto- and cross-spectra.
- Fast and "easy to use" (pip install, Jupyter notebooks, ...).

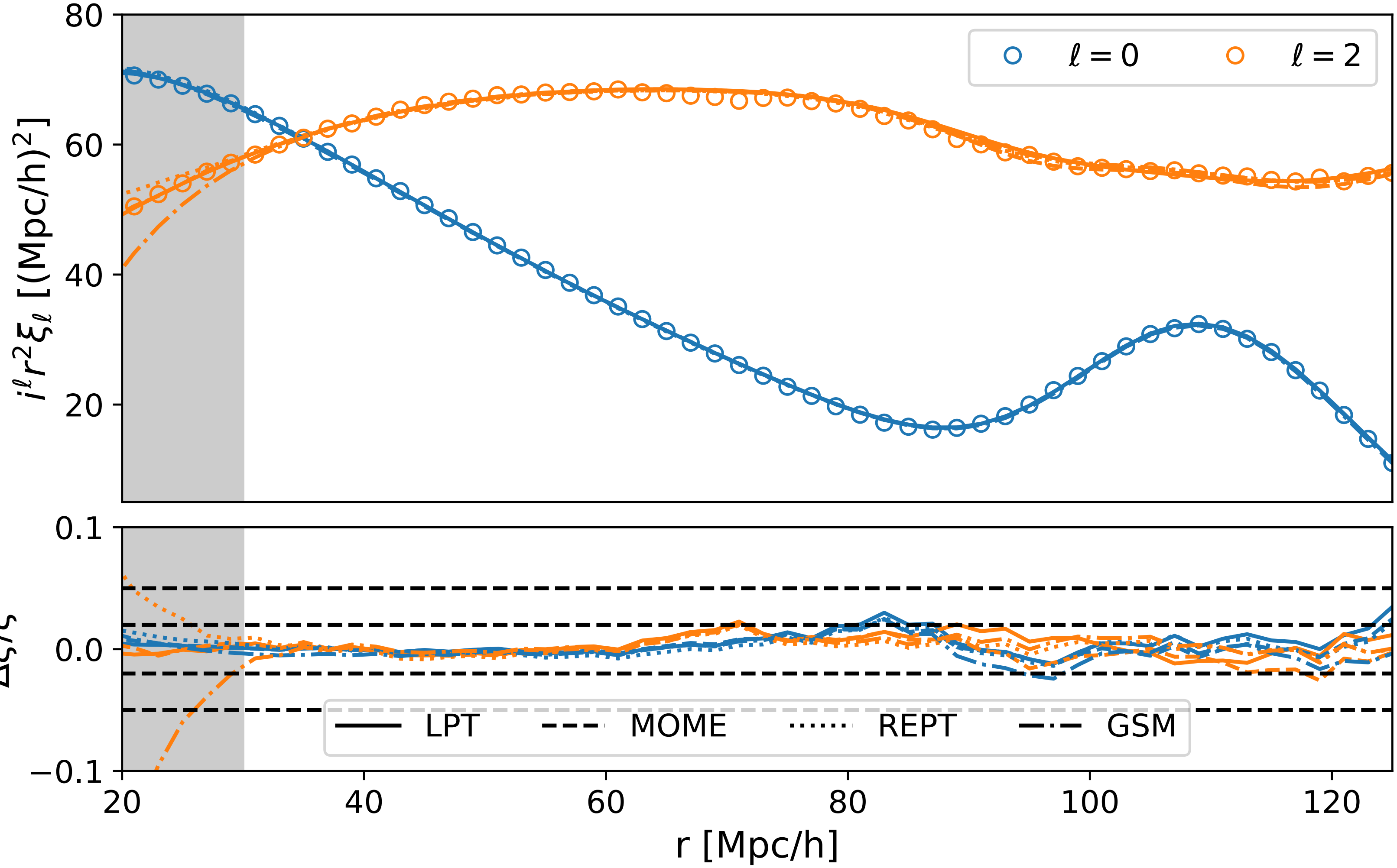
Vlah+(15, 16, 19),
Chen+(19a,b),
Modi+(20),
Chen+(20a,b,c)

https://github.com/sfschen/velocileptors

# Power spectrum multipoles

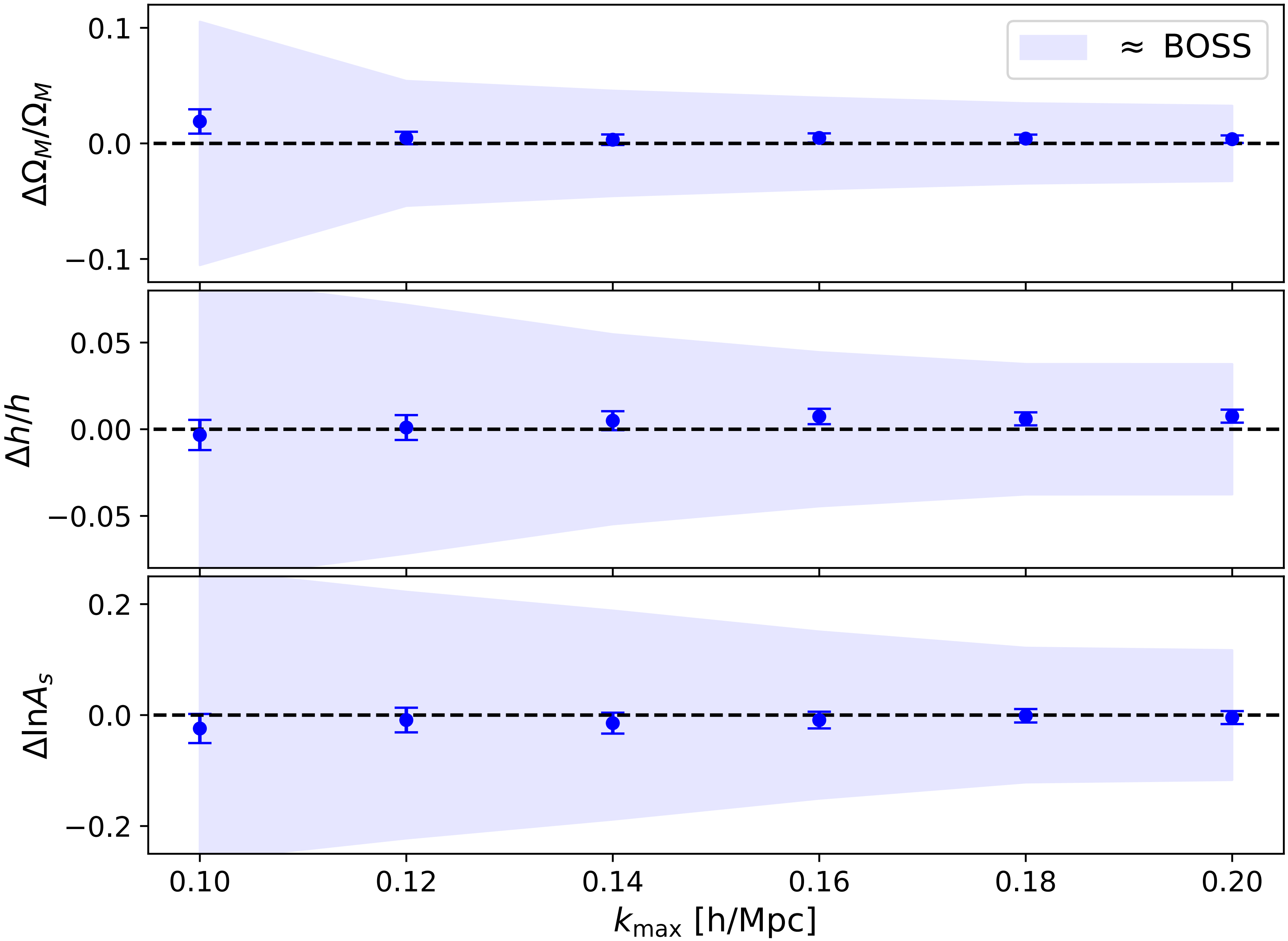


# Correlation function multipoles



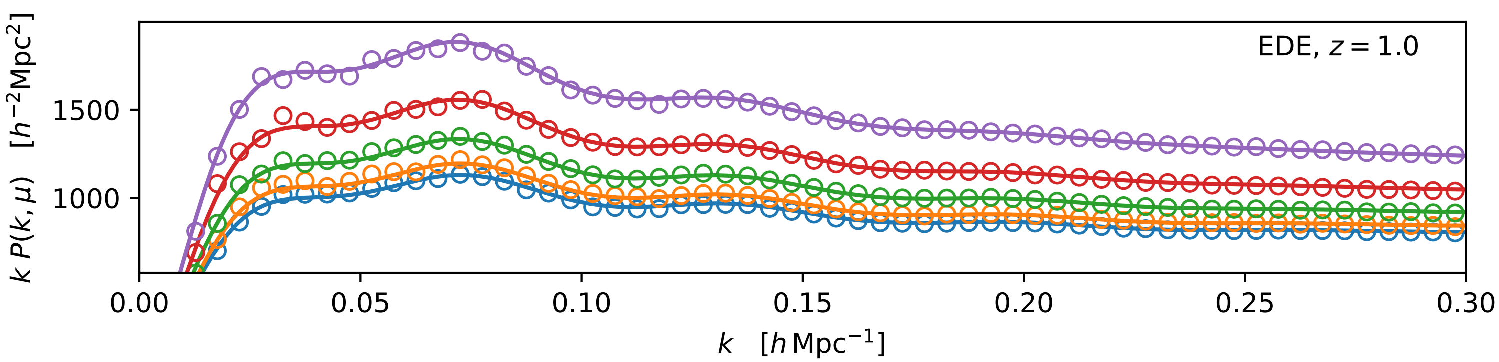
# PT blind challenge (Nishimichi+20)

Inferring parameters from fits to mock survey data:



# Features induced in $P(k)$ by expansion history

- Can also look at extended models: here a short period of DE domination at $z \simeq 10^4$ peaking at $\rho_{DE}/\rho_{\rm tot} = 10\%$.
- Different growth of modes inside and outside the horizon leads to an "induced feature" in the power spectrum – which we model very, very well with `velocileptors`.




- Forecasts indicate future surveys could measure the effects of EDE (or light degrees of freedom) to percent level at any time after the Universe was a thousand years old (Sailer+21).

# Conclusions I

- We are in the midst of the "golden age of cosmological surveys".
- There are **many** (quasi-)linear modes left to map!
- These will allow precisions tests of SM and GR, and improve constraints on parameters by substantial factors (or find something new!).
  - Already (several) percent-ish level constraints at lower $z$ are turning up much-discussed "tensions".
- If theory can't give us guidance, maybe phenomenology can?
  - Work where inference is clean.
  - Look where we haven't looked before.
  - If you don't know how to maximize $S$, then minimize $N$!
- The best observational approaches are still TBD.
  - Pilot programs and R&D

# Conclusions II

- Increasing survey power is driving a renaissance in analytic models of large-scale structure.
  - More perturbative modes at higher precision!
  - Form and techniques familiar from other areas of physics.
  - A few "cosmology" wrinkles.
- The models are well motivated and work well on current data.
  - Well motivated inference problem.
  - Allow us to forecast performance of future surveys reliably.
  - Survey optimization.
- Adding "beyond standard model" physics or new probes is an active area of research.
- The benefits and disadvantages of the different approaches is still not fully understood... after 50 years we still don't understand structure formation anywhere near as well as we'd like to!

*The End!*