Theoretical questions inspired by next-generation surveys

Martin White

(UC Berkeley, LBNL, Higgs Center)

June 2024

Outline

- Opportunities
  - Context – DESI-II and Spec-S5.
- Targets and key numbers.
  - Differences with the low-z Universe.
- Challenges
  - The bias expansion and stochastic terms.
  - Radiative transfer.
  - Covariances.
  - The role of “beyond 2-point” statistics.

Context

The situation in the US is currently very positive for "survey cosmology", particularly spectroscopic surveys.

- DESI-ext is under "light" review at present.
- DESI-II is almost certain to go ahead (2029-2034)
  - P5: "Support DESI-II for cosmic evolution"
  - AD of the Office of Science: "DOE will work with the DESI Collaboration to carefully decide a scope, schedule and cost envelope for the DESI-II upgrade"
- Spec-S5 has a high likelihood of proceeding (2034-?)
- Spec-S5 will map
  - 62M galaxies at $z > 2$ for use in inflation and dark energy
  - 130M galaxies at $z < 2$ for use in late-time cosmology

> Spec-S5, holds great promise to advance our understanding and reach key theoretical benchmarks in several areas: inflationary physics via the statistical properties of primordial fluctuations, late-time cosmic acceleration, light relics, neutrino masses, and dark matter – *P5 report*.

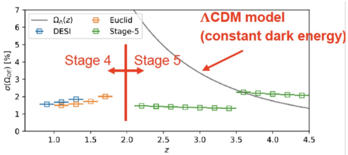
Facilities Assessment – SpecS5

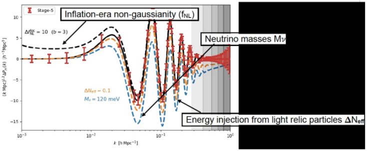
- Scientific assessment: ***"absolutely central"***

  The 2023 P5 report said Spec-S5 will "advance our understanding and reach key theoretical benchmarks in several areas" relevant to our understanding of the universe. Spec-S5 would maintain U.S. and DOE global leadership in the cosmic frontier through the 2030s.

- Technical readiness: ***"ready to initiate construction"***

  Given the recent design and R&D progress we rate Spec-S5 as ready to initiate construction within the next decade, assuming success of the near-term R&D program, and once formal agreement is reached with NOIRlab and NSF for use of the Mayall and Blanco telescopes.





Tracers of LSS at $2 < z < 6$

- CMB lensing (plus tSZ, kSZ, ...)
  - A natural "by-product" of next generation CMB surveys is high fidelity CMB lensing maps – probing the matter back to $z \simeq 1100$.
  - A lot of CMB lensing theory looks an awful lot like LPT. The "standard" method of computing lensing for the C_ℓs is simply the Zeldovich approximation in another context.
    - Several people have noticed this (Emanuele, Ravi, ...) but so far nobody has found a good way to exploit it.
- Galaxies: LBGs and LAEs
  - Dropout, or Lyman Break Galaxy (LBG) selection targets the steep 912Å break in an otherwise 'flat' spectrum.
  - Selects massive, actively star-forming galaxies.
  - A fraction of these objects have bright emission lines (LAEs). LAEs tend to be lower bias, younger (with little dust).
- The IGM ...

Lots of volume!

For a highly biased sample (neglect RSD)

$$\frac{\Delta P(k)}{P(k)} \approx \sqrt{\frac{2}{N_k}} \left[1 + \frac{1}{\bar{n}P}\right] = \frac{2\pi}{\sqrt{Vk^3\ \Delta \ln k}} \left[1 + \frac{1}{\bar{n}P}\right]$$

For 18K sq.deg. from $3.0 < z < 3.5$ we have $V = 34.5\, h^{-3}\mathrm{Gpc}^3$.

Assuming 50% success for $m_{UV} < 24.5$ u-dropouts,
$\bar{n}P(k = 0.1) \simeq 3$, $\bar{n}P(k = 0.3) = 0.5$ and $\bar{n}P(k = 0.5) = 0.1$.
This implies:

$$\frac{\Delta P}{P} = 0.2\% \quad \text{at} \quad k \simeq 0.3\, h\,\mathrm{Mpc}^{-1} \quad \text{with} \quad \Delta \ln k = 0.1$$

and $< 1\%$ over more than 1.5 dex in scale (per $\Delta \ln k = 0.1$).

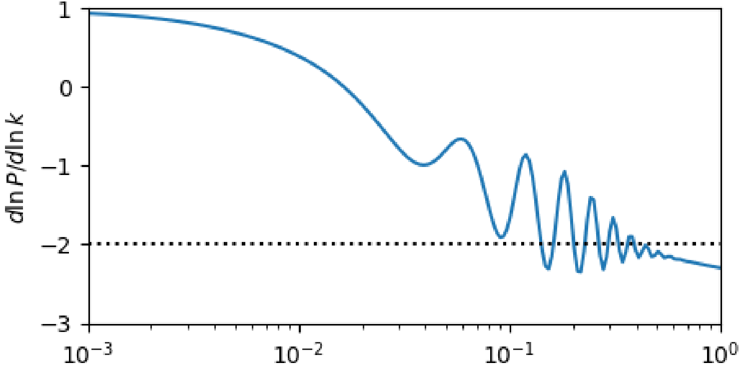
Implications of “long lever arm”

Are there “theoretical” implications for having small errors for modes with $k \ll k_{\rm nl}$?

- Longer dynamic range in scale where PT “works”
  - What does this imply about degeneracy breaking?
  - Does “EFT” work better, i.e. non-analytic terms important where CT are not?
  - More of our “field level” is linear where $P(k)$ is a sufficient statistic.
- Do we want/need 2-loop?
- The power spectrum slope is getting increasingly negative to high k. What does this imply?

Power spectrum flattening

The power-law slope, $d \ln P / d \ln k$, is approximately -2 near the non-linear scale at $z \approx 3 - 4$:



Implications of "long lever arm" & flat spectra

Consider a scaling cosmology, $P(k) \propto k^n$.
In real space, neglecting logs and writing $\kappa = (k/k_{\rm nl})$

$$\Delta^2(k) = \underbrace{\kappa^{n+3}}_{\rm tree} + \# \underbrace{\kappa^{2(n+3)}}_{\rm 1-loop} + \# \underbrace{\kappa^{n+5}}_{c.t.} + \# \underbrace{\kappa^{3(n+3)}}_{\rm 2-loop} + \cdots + \# \underbrace{\kappa^{7}}_{\rm stoch.} + \cdots$$

If we take $n = -2$

$$\underbrace{\kappa}_{\rm tree} + \# \underbrace{\kappa^2}_{\rm 1-loop} + \# \underbrace{\kappa^3}_{c.t.} + \# \underbrace{\kappa^3}_{\rm 2-loop} + \cdots$$

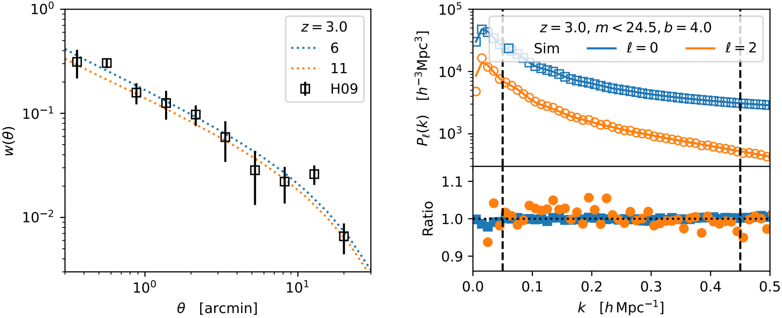
For $\kappa = 1/2$ only gaining powers of 2!
For $\kappa \simeq 0.1$ have tree+1-loop dominating.

Implications of "special selections"

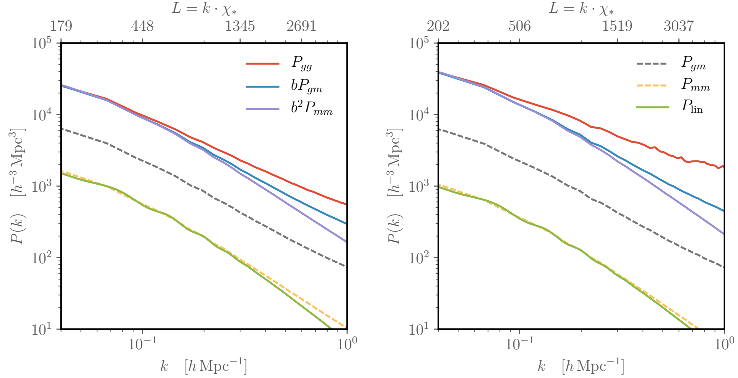
Redshifts > 2 are a long way away, and we're planning to select special sub-populations of objects.

- Within the context of PT, all objects are just biased tracers of the matter field. Special selection $\Rightarrow$ implications for bias.
- 'Break' in the halo mass function, $M_{h,\star}$, shifts to smaller masses at high z.
- Large d_L means even faint galaxies are very luminous, e.g. high $M_\star$ or SFR.
- Bias tends to be large, and therefore scale-dependent.
- Satellite fractions tend to be low (we're on the steeply falling part of the halo mass function), suggesting smaller FoG.
- Stochastic terms smaller than we're used to at lower z for fixed b (since $\bar{n} \propto \bar{\rho}/M_{h,\star}$).

Scale-dependent bias: LBGs

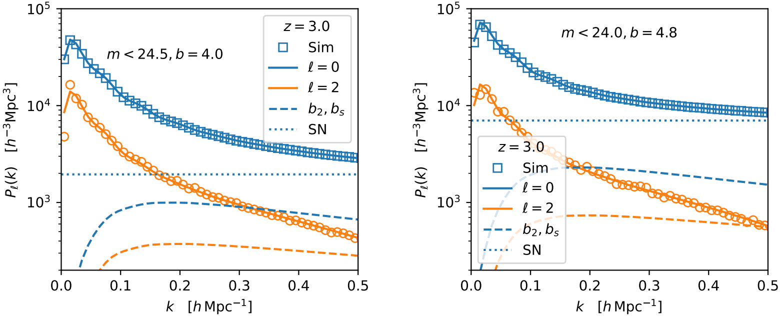


Scale-dependent bias: LBGs

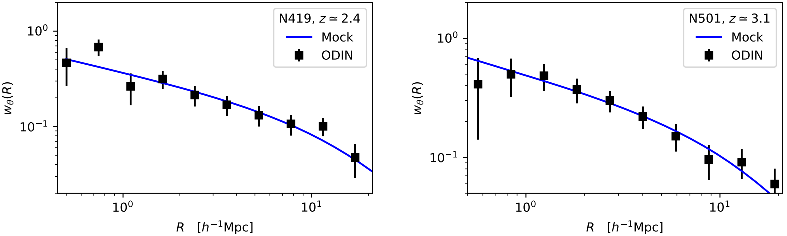


Wilson+19

Scale-dependent bias: LBGs



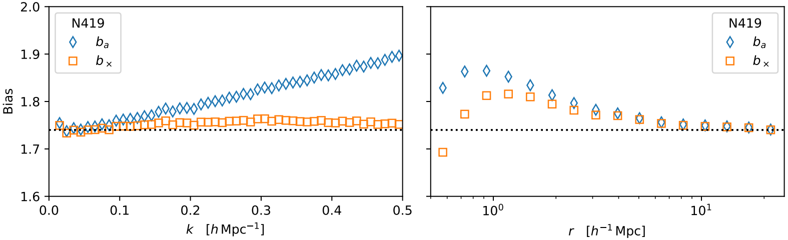
Lyman-α emitters (LAEs)



White+24

Scale-dependent bias: LAEs

For the LAEs, with lower bias, these effects are much reduced.



White+24

Bias expansion

What implications does this have?

- High bias can be a boon or a curse
  - Higher S/N in the monopole, more sensitive to f_{NL}^{loc}, ...
  - Smaller RSD, so one important "protected by symmetry" signal is "lost"
- There's no problem, in principle, in going to higher order in the bias expansion – but we need to worry about degeneracies, projection effects, loss of constraining power.
  - Would we need simulation-based priors?
  - What measurements would we use to validate them?

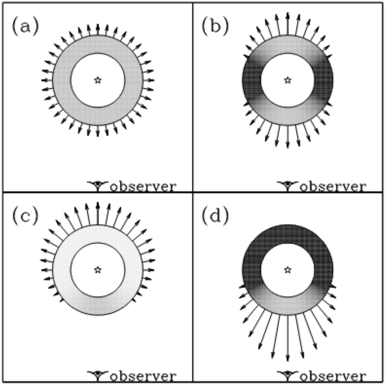
Line-of-sight and RT

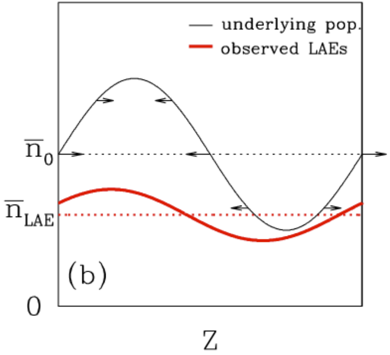
If we want low bias, that's probably faint galaxies so we need 'strong' lines to have decent redshift success rate.

- Lyα is a resonant line, so strongly affected by radiative transfer (RT).
- RT modulates the galaxy selection depending upon local density and (line-of-sight) velocity divergence.
  - These are key signals for us!
- How strongly is currently under debate
  - Zheng+11 argue for a large effect.
  - Behrens+18 claim that this is due to poor resolution in the older simulation. (If gas very dense where Lyα is emitted, random walks more in frequency space before leaving the galaxy.)
- This plays havoc with our ability to constrain some parameters (Ebina+24)
- We know the physics – can we "break this degeneracy" using other measurements?

LAE RT

In the presence of RT, how bright an LAE looks depends upon the local density and line-of-sight velocity 'environment'.





Zheng+11

Covariances

- For BOSS/eBOSS/DESI can (largely) get away with analytic covariances.
- Efficient ways of doing the mask convolution/window function?
- At higher z, fixed angle subtends larger scale and $k_{\rm nl}$ is larger. "Survey effects" more important.
- Redshift efficiency will be lower, more complex selection.
- Will we need mocks for Cov?
  - Are there tricks we can use to make them tractable given the scales are "linear"?
  - History of using PT at $z_{\rm obs}$ is mixed – but better than lognormal?
  - Do we need 'objects', or just fields?
  - Does this push us to data compression? Different for each kind of analysis?

Bispectrum, and beyond

- How does having a high bias impact the usefulness of the bispectrum for breaking degeneracies?
  - Qualitatively it suggests the need to use lower $k_{\rm max}$ making B less useful (relative to P).
  - Qualitative arguments are ... qualitative.
- Does having a longer lever arm to $k_{\rm nl}$, and hence many more "linear" configurations change anything? Subsets that are less sensitive to hard-to-model things?
- If our covariance is numerical, will we need data compression?
- Shot noise is low at $k \simeq 0.1\,h\,{\rm Mpc}^{-1}$ but quite large at $k_{\rm nl}$. Limitation for field-level methods?

A multi-tracer, multi-survey world

The next decade, in particular, will see us working in a multi-survey world.

- Is this a strength of analytic models?
  - Principled way of combining datasets.
  - Improved degeneracy breaking.
  - Combination of volume and mass resolution <u>very</u> challenging for simulations. Tiny error bars don't help!
- Is this a place where simulations shine?
  - Some valuable statistics have support at high k, not likely to be covered by PT.
  - Forward modeling allows more sophisticated matching to surveys.
- This is likely to be an area where combining simulations and theory is critical!

Conclusions

We are in the midst of the "golden age of cosmological surveys".

- DESI, Euclid, Rubin, Roman, ... will keep us busy for some time!
- The prognosis for future spectroscopic surveys targeting "high z" looks good.
- This is an area where PT should shine!
- Long lever arms in scale and time where errors are small, but
  - $k^3 P$ is getting flatter.
  - we'll be selecting "special" objects.
- Future surveys represent a huge challenge to simulations – large volume, small halos, extreme precision – so want to think about combination of theory and sims.

The End!

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!

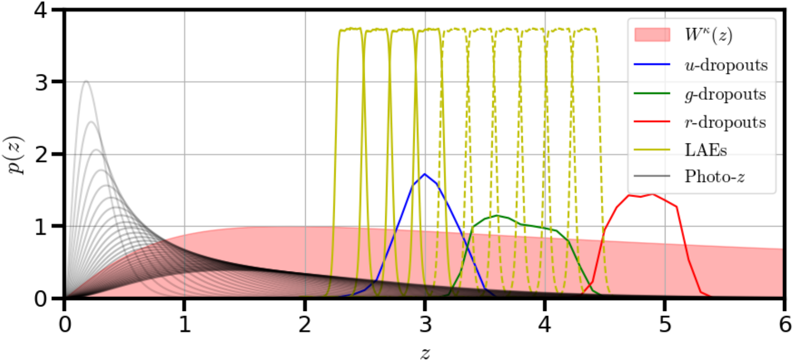
- A natural "by-product" of next generation CMB surveys to constrain primordial gravitational waves is high fidelity CMB lensing maps – probing the matter back to $z \simeq 1100$.
- This is a natural "multi-tracer", and one whose (Lagrangian) biases are all known ($= 0$)!
- A lot of CMB lensing theory looks an awful lot like LPT. The "standard" method of computing lensing for the C_ℓs is simply the Zeldovich approximation in another context.
  - Several people have noticed this (Emanuele, Ravi, ...) but so far nobody has found a good way to exploit it.

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_ν 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). The LAE population tends to be lower bias, younger and with low dust. If you can select them, they're easy to redshift!

Galaxies over the whole range



Ebina+24

Maximizing S/N

Motivation for high-z cosmology:

Want to maximize the S/N for new, BSM, physics

- There are many possible extensions to our SM (ΛCDM+GR).
- 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 ($z \simeq 2$)!

Tensions in the current model

- These tensions are the focus of a lot of effort in the field!
- They resist 'easy' solution.
  - I (for one) am pretty mystified as to what is going on!
- The evidence is not as robust as we'd like, but more data like this are coming very soon!
- They have only arisen as we've shrunk the error bars: "precision" cosmology.
  - 'Hubble tension' and 'growth tension' represent $\mathcal{O}(10\%)$ shifts in parameters.
  - Seeing such things at $> 5\,\sigma$ requires $\sigma \simeq 1 - 2\%$

Since the model is working "pretty well" any signatures of BSM physics or deviations from ΛCDM are likely to be subtle ...

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".

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

The "LSS program": planning for what comes next

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

- Expansion history and curvature (metric)
- 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
- etc.

Massive neutrinos

- Galaxies probe the $c + b$ field while lensing probes the matter.
- At linear level use $P_{cb}(k)$ for galaxies and $P_{cb,m}$ for galaxy-lensing cross-correlation.
  - Good to sub-percent level (e.g. Bayer+21)
- If care is taken with normal ordered bias operators, can use $P_{cb,m}$ in loops with corrections of order $f_\nu P_{\rm lin}^2 \ll 1$ and be correct even in the "transition regime" from clustered to free-streaming neutrinos.