Theory, Analysis, and Computing Summary

Lawrence Berkeley National Laboratory, November 15, 2017

Katrin Heitmann, Bhuv Jain, Martin White and
many contributions from people in the audience

Main Take-away Messages

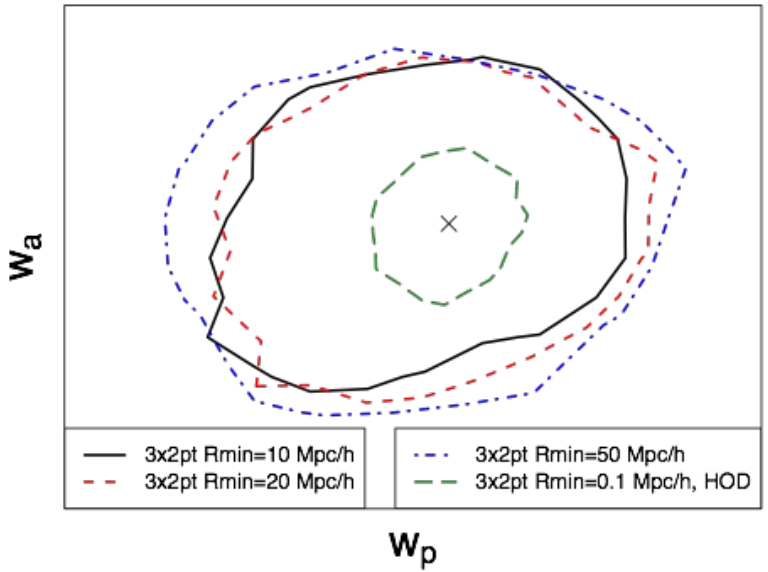
- A lot more science can be extracted from small scales, if we can model those scales appropriately
- Cross-correlations of different data sets also holds a wealth of information and constraining power
- Simulations and ancillary observations are crucial to extract the science
- Sharing simulations, tools, and data in a way that the community can easily access them, interact with them, and contribute to them are crucial!
- This will require investments in people to build infrastructure that is long-term sustainable!
- Every experiment is currently struggling to find enough analysis personpower --- for future surveys and multi-tracer analyses this problem will be highly exacerbated!

Gains from Small Scales: Dark Energy

Proof of principle

2-4x improvement in DE constraints for Stage IV experiments

Improvements due to self-calibrating nuisance parameters
from small-scale clustering and lensing



Krause & Eifler 2016

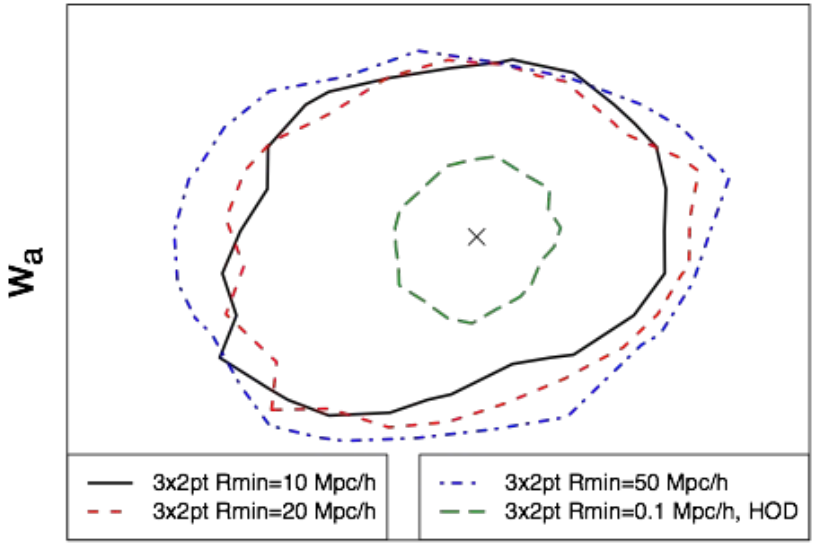
Slide from A. Hearin talk

We can surely do better, probably a lot

This exercise was for single-tracer probes

Far more information from multi-tracer cross-correlations

There is a **VAST** amount of additional constraining power
that we are simply throwing away if we only stick with single-tracer, quasi-linear probes



Slide from A. Hearin talk

Krause & Eifler 2016

Gains from small scales: beyond wCDM

Galaxies, galaxy clusters and other ‘small scale’ data can test gravity and dark matter interactions in a number of ways.

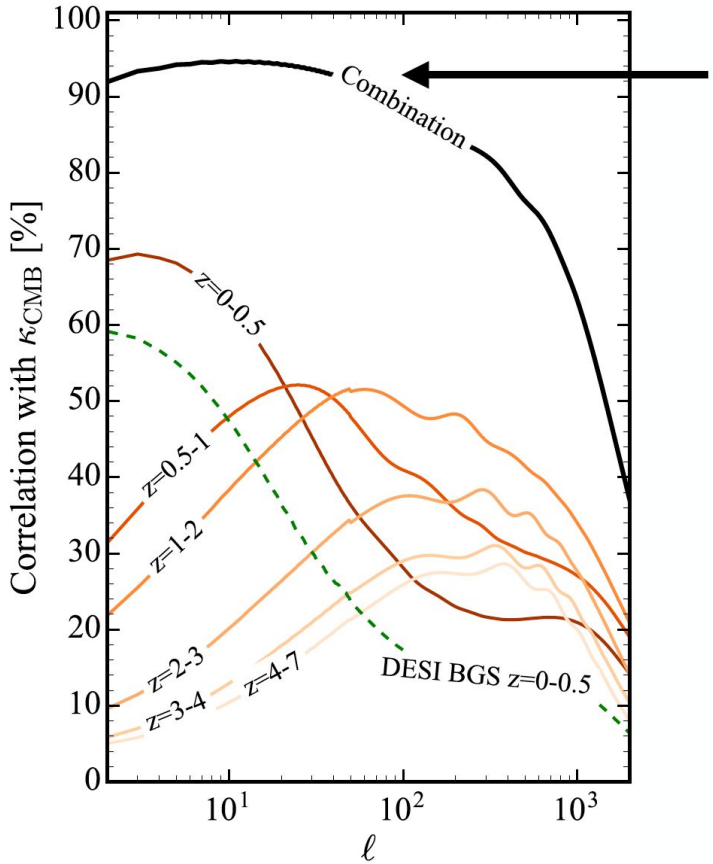
These tests involve both the light (multiwavelength) and the halo mass distribution.

To use DES, DESI and LSST data for novel probes, we need:

- Rigorous analysis software
- Simulations and mock catalogs designed for such tests
- Ancillary data
- The interaction of signals and selection effects is critical, requiring detailed simulations which model the joint distribution of observables with high accuracy (that can be validated).

CMBxLSS

… so they are highly correlated



Combined DESI+LSST sample more than 90% correlated with CMB-S4 lensing at $L<100$

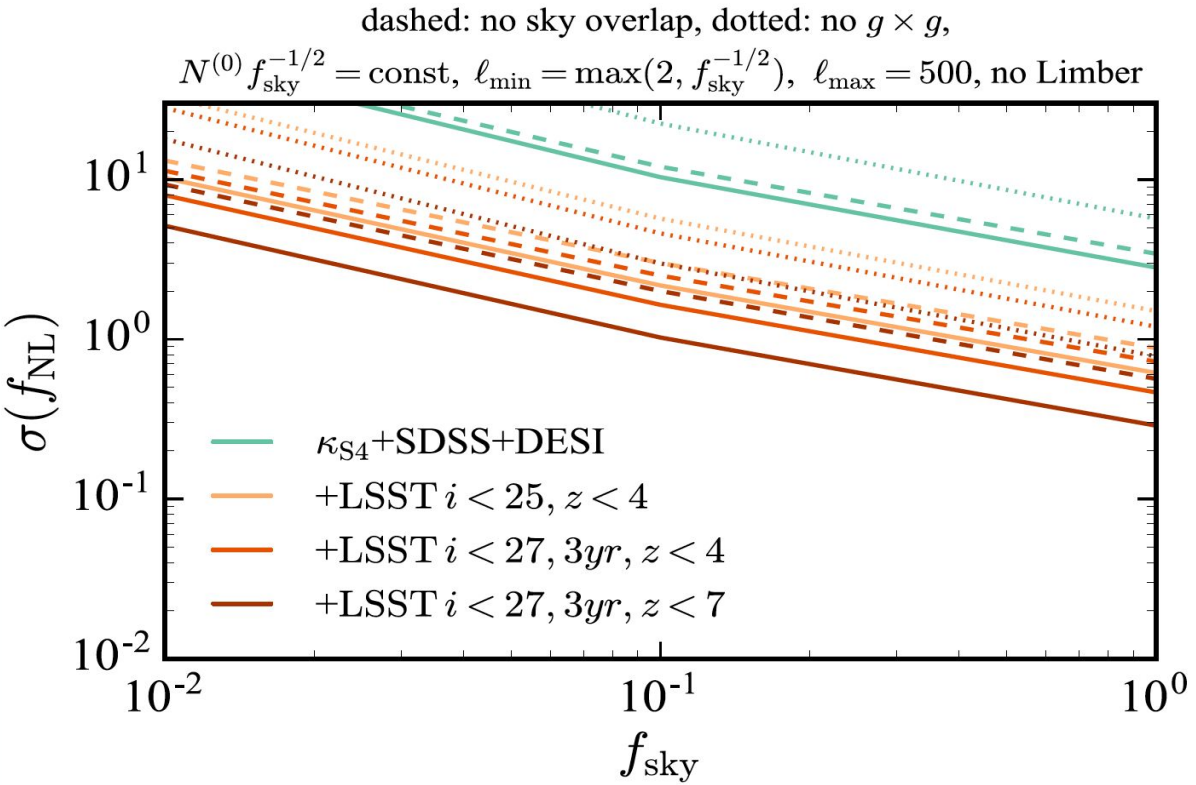
[includes shot noise and lensing noise]

Slide from U. Seljak talk

MS & Seljak (in prep.)

CMBxLSS

Local f_{NL} as function of sky fraction



Slide from U. Seljak talk

Dalal (2008), MS & Seljak (in prep.)

CMB-LSS: What is needed to make this happen?

Development of analysis framework (joint nuisance parameters, bias models, covariances, systematics models, scalable algorithms).

"Same sky" mocks with the signals properly correlated.

-- Need to investigate level of abstraction at which we can deal with foreground separation, timestream filtering, photometric calibration, etc. (would there be a need to "run the mission" over the maps?). How would we share these tools?

Explore requirements on galaxy and CMB surveys for joint analysis e.g. LSST photometric accuracy or interloper fractions to get low ell clustering for f_{NL}

Identification of teams that cross disparate communities to perform these analyses.

Validating mocks: for 1 survey

Each project has different sensitivities to non-idealities in the mocks and different quantities that they need.

Expertise and data to validate mock catalogs resides in the collaborations.

This suggests mock validation should be done at the project level, but it does not require that the input simulations be a project responsibility if there is another way to obtain them.

Selecting and curating validation data sets is a large amount of work, which is separate from the work required to scope and run simulations.

We should explore the best way to serve simulation data in a way that is easy for the surveys to ingest and process. Would a small investment of $ now, save $$ and time for each of the projects downstream?

Validating mocks: across surveys

Things become more interesting when we cross project lines:

-- For joint analysis, how do we produce the mocks?

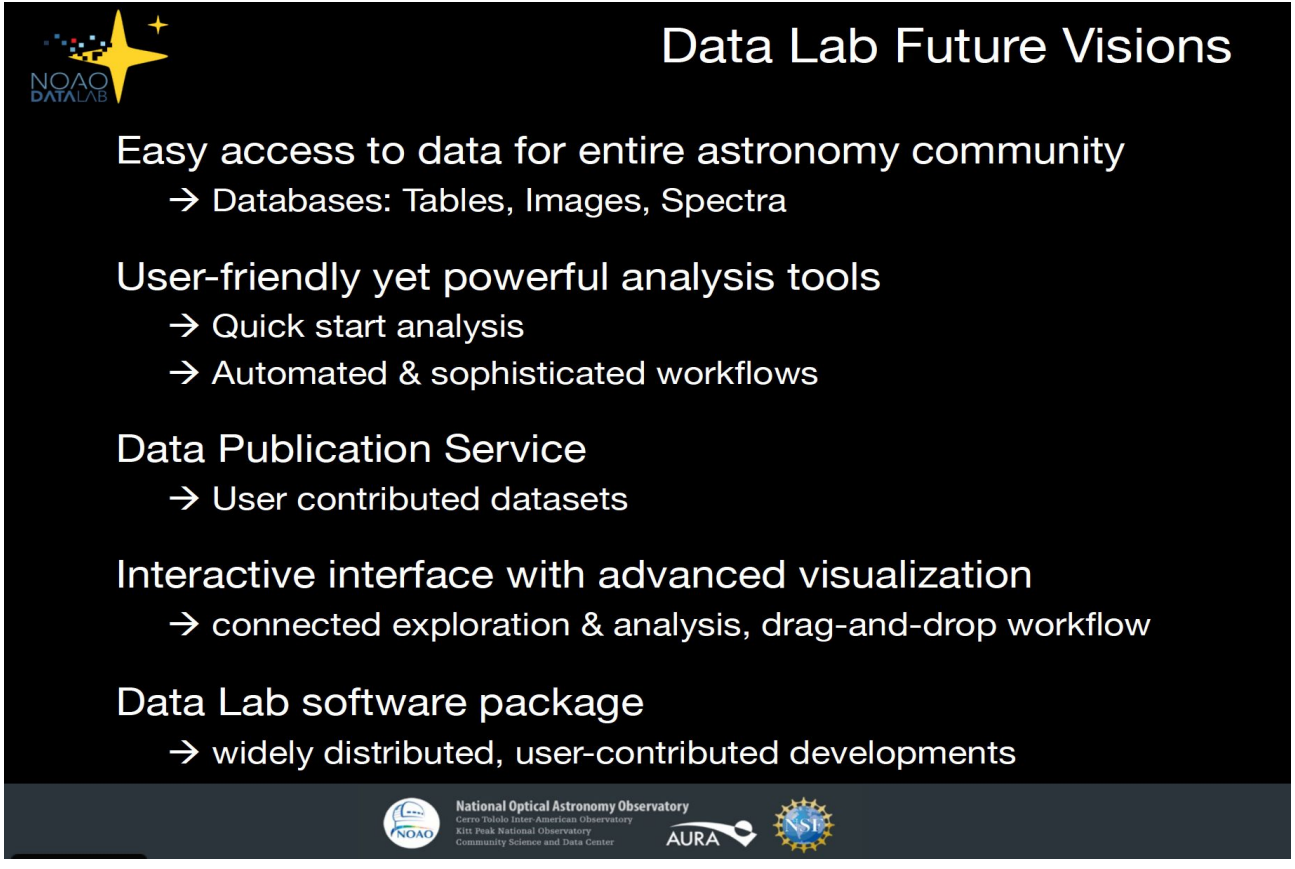
-- How do we capitalize on efforts which would cross project boundaries?

As part of the mock validation, one finds key degeneracies or unknowns. With these in hand, how do we “close the loop” to reduce these uncertainties?

-- Is there a mechanism for suggesting small “ancillary” projects in one project which would benefit the modeling for another project?

-- Can we create a mechanism for obtaining supplementary observations with non-DOE facilities that could reduce the modeling uncertainties in a planned analysis?

Sharing Data, Tools, and Simulations



Slide from S. Juneau

Main Take-away Messages

- A lot more science can be extracted from small scales, if we can model those scales appropriately
- Cross-correlations of different data sets also holds a wealth of information and constraining power
- Simulations and ancillary observations are crucial to extract the science
- Sharing simulations, tools, and data in a way that the community can easily access them, interact with them, and contribute to them are crucial!
- This will require investments in people to build infrastructure that is long-term sustainable!
- Every experiment is currently struggling to find enough analysis personpower --- for future surveys and multi-tracer analyses this problem will be highly exacerbated!