CMB x LSS Simulations

"Synergies of Large Scale Structure Surveys with CMB-S4"
CMB-S4 2021 Summer Collaboration Meeting

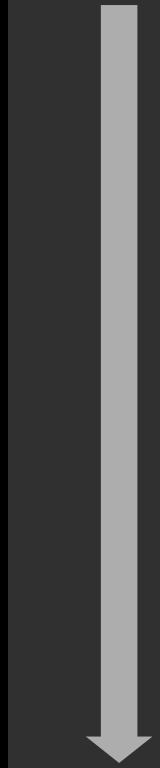
Jia Liu
(UC Berkeley → Kavli IPMU)

with inputs from Marcelo Alvarez, Yuuki Omori, Francois Lanusse

Why Cross-Correlations?

See all other talks in this session

Why Simulations?



Survey pipeline:
map → statistics → cosmology

Systematics — instrumentation
(beam, bandpass, electronics...)

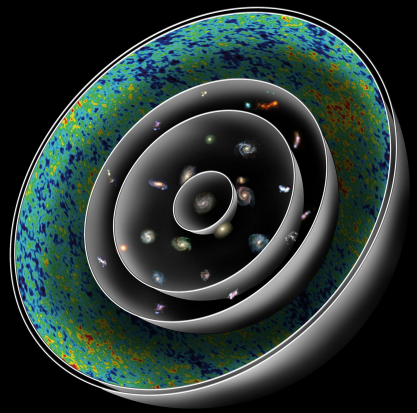
Astrophysics
(baryons, intrinsic alignments...)

Covariance

Modeling

More Stringent

Why **Correlated** Simulations?



Survey pipeline:
e.g. 6x2 analysis, CMB FGs x LSS

~~**Systematics — instrumentation**~~
~~(beam, bandpass, electronics...)~~

Astrophysics
(baryons, intrinsic alignments...)

Covariance between CMB x LSS

Modeling beyond the linear theory

More Stringent

Typical CMB simulations

- 2D maps
- Single cosmology
- Gravity-only / approx. methods
- Full-sky / curved
- 1-10 Gpc boxes
- Painted observables

Examples: Sehgal+2010, Stein+2020 (Websky)

Typical LSS simulations

- 2D (lensing) and 3D (galaxies)
- More cosmologies
- Gravity-only, some hydro
- Often flat-sky
- $\lesssim$Gpc box
- HOD / Semi-analytic models

Examples: LSST DC2, Euclid Flagship, DES Buzzard, Magneticum, Aemulus, BAHAMAS, SLICS, Quijote, MassiveNuS, and many many more..

See a list of public simulations compiled by Joanne Cohn https://w.astro.berkeley.edu/~jcohn/mocklist.html

Typical CMB simulations

- 2D maps
- Single cosmology
- Gravity-only / approx. methods
- Full-sky / curved
- 1-10 Gpc boxes
- Painted observables

Examples: Sehgal+2010, Stein+2020 (Websky)

Typical LSS simulations

- 2D (lensing) and 3D (galaxies)
- More cosmologies
- Gravity-only, some hydro
- Often flat-sky
- $\lesssim$Gpc box
- HOD / Semi-analytic models

Examples: LSST DC2, Euclid Flagship, DES Buzzard, Magneticum, Aemulus, BAHAMAS, SLICS, Quijote, MassiveNuS, and many many more..

Correlated sims need to accommodate **BOTH** worlds

Public CMB x LSS Simulations

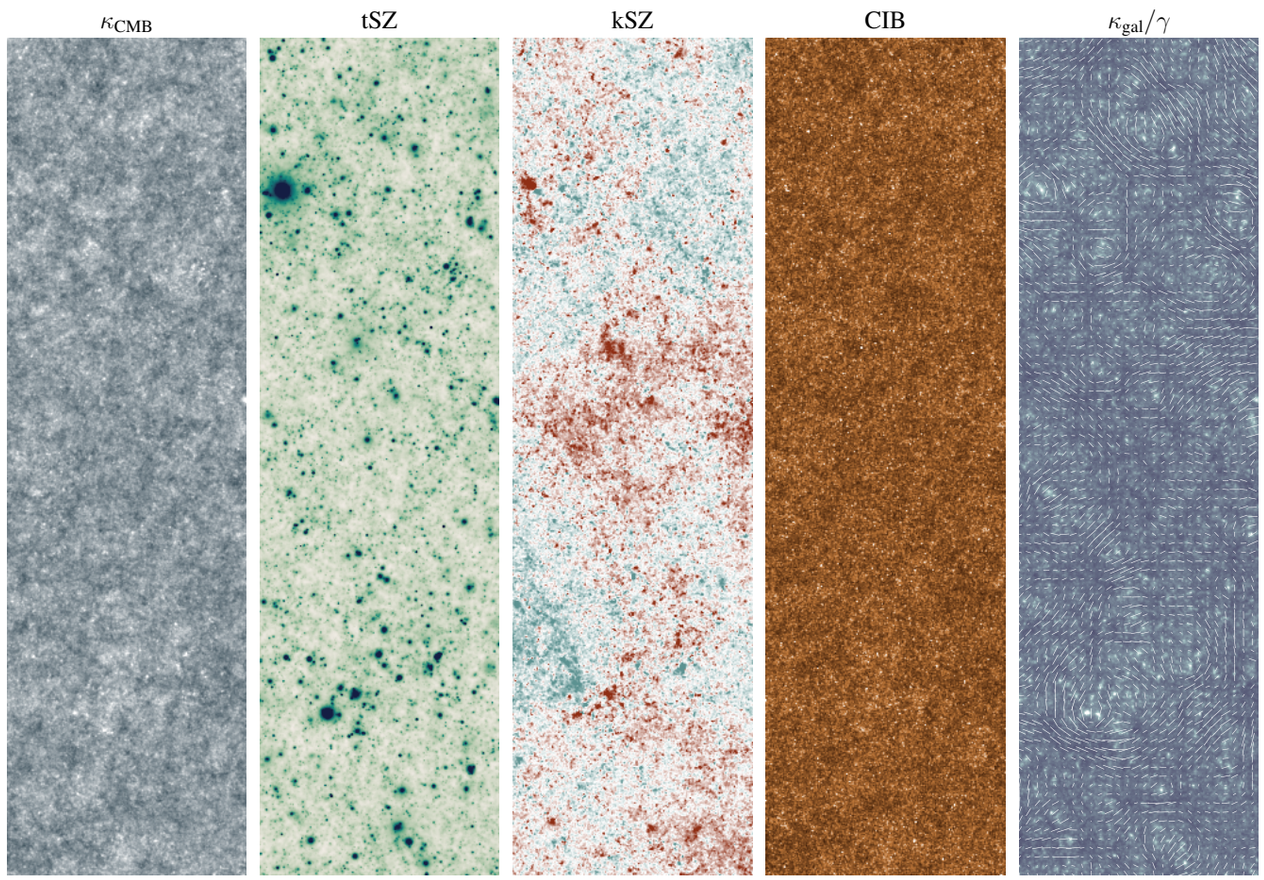
- From LSS experts:
  - **Takahashi+2017** [κ^{gal}, κ^{CMB}]: 108 maps x full-sky, WMAP9
  - **BAHAMAS** (McCarthy+2018) [κ^{gal}, κ^{CMB}, tSZ]: 25 maps x 25 deg^2, Planck15, WMAP9
  - **MassiveNuS** (Liu+2018) [κ^{gal}, κ^{CMB}]: 10,000 maps x 12.25 deg^2, 100 cosmologies (M_ν, A_s, Ω_m)
- From CMB experts:
  - **MillimeterDL** (Han+2021) [κ^{gal}, κ^{CMB}, tSZ, kSZ, CIB, Radio, lensed CMB]: 500 maps x full-sky, WMAP5
  - **MDPL2synsky** (Omori in prep) [δ, κ^{gal}, κ^{CMB}, tSZ, kSZ, CIB, Radio]: 1 map x full-sky, Planck15

MULTI DARK PLANCK 2 SYNTHETIC SKY SIMULATION

- Based on the MultiDark Planck 2 *N*-body simulation (Klypin 2016, $L = 1\ h^{-1}$Gpc box with 3840^3 particles), and associated halo catalogs.

- Galaxy evolution model (used for CIB/radio galaxies) from UniverseMachine (Behroozi 2018) and TRINITY (Zhang 2021).

- 1 full-sky realization (all the components provided as $N_{\rm side} = 8192$ maps or catalogs).



- Ray-tracing computed on $N_{\rm side} = 16384$. All the CMB secondaries are lensed using the lensing field.

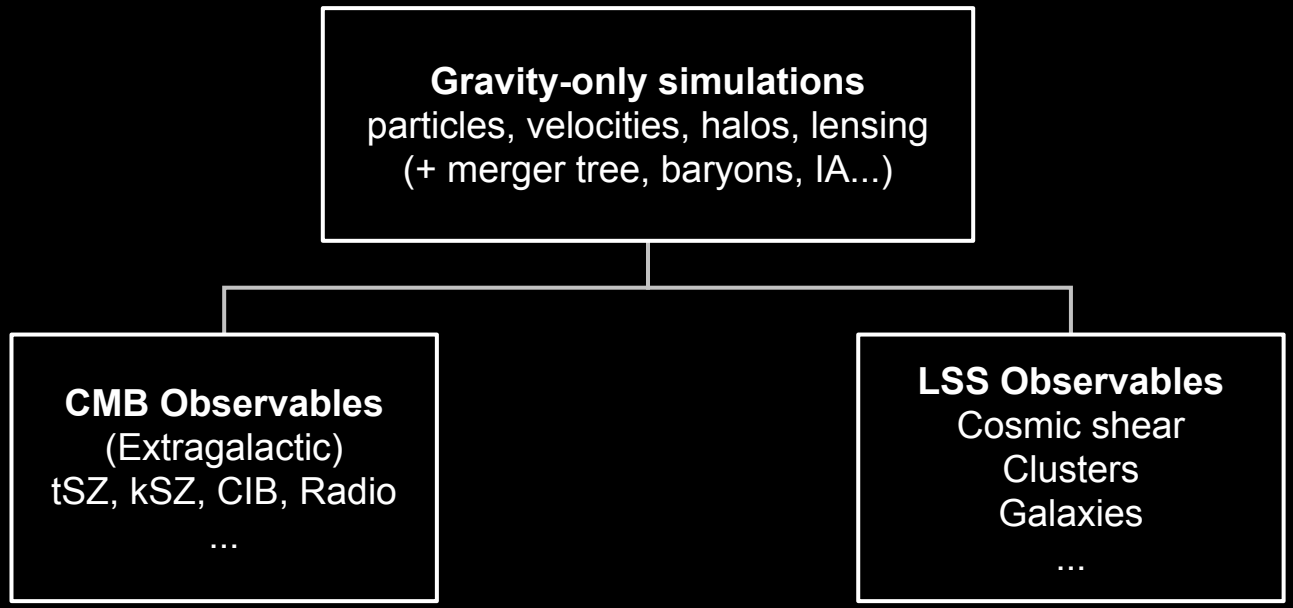
- Accuracy of the simulation tested for various types of analyses (high-ℓ *TT,* 6×2pt, tSZ cross-correlation, component separation, delensing etc).

- Plans to implement various galaxy types via HOD.

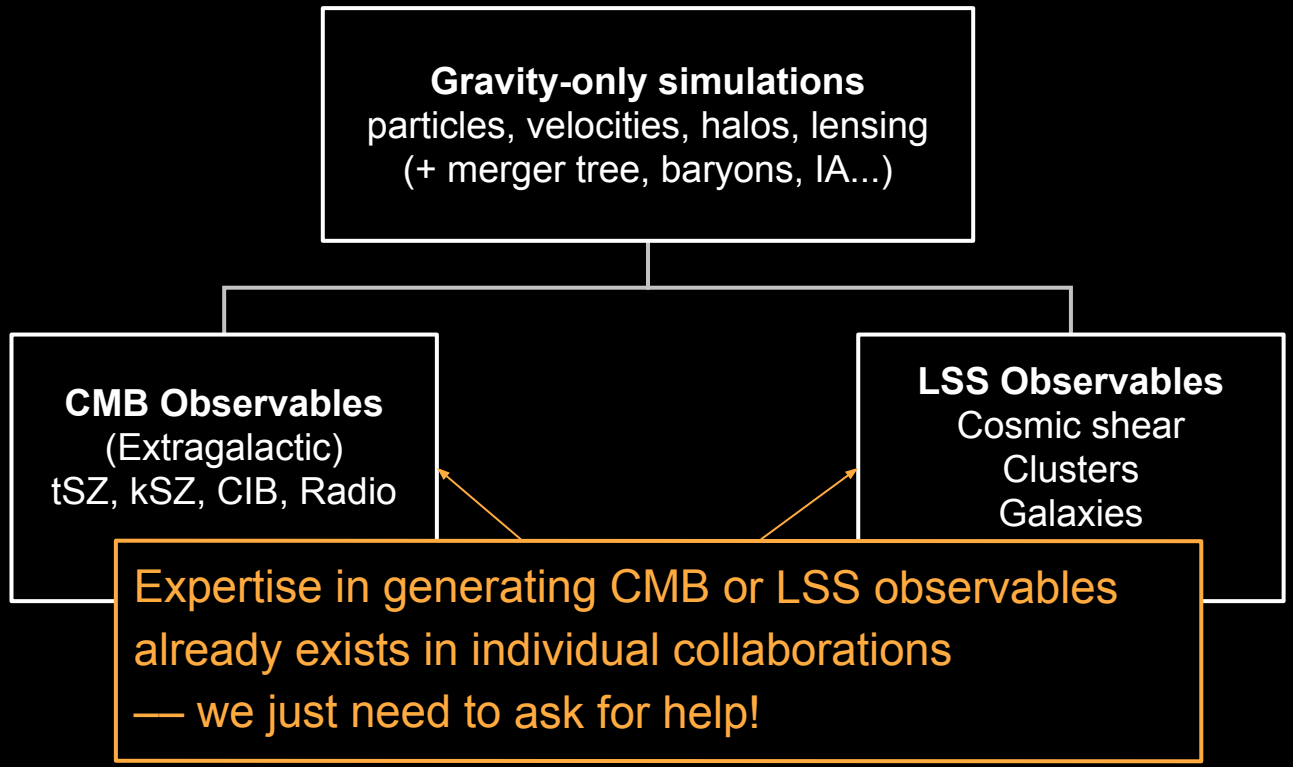
Observable	Redshift range
CMB lensing	$0 < z < 1089$
tSZ	$0 < z < 3.0$
kSZ	$0 < z < 3.0$
radio	$0 < z < 4.0$
CIB	$0 < z < 8.6$
δ (lensed)	$0 < z < 8.6$
$\kappa_{\rm g}, \gamma_1, \gamma_2$	$0 < z < 8.6$

Coming soon https://yomori.github.io/mdpl2synsky/index.html

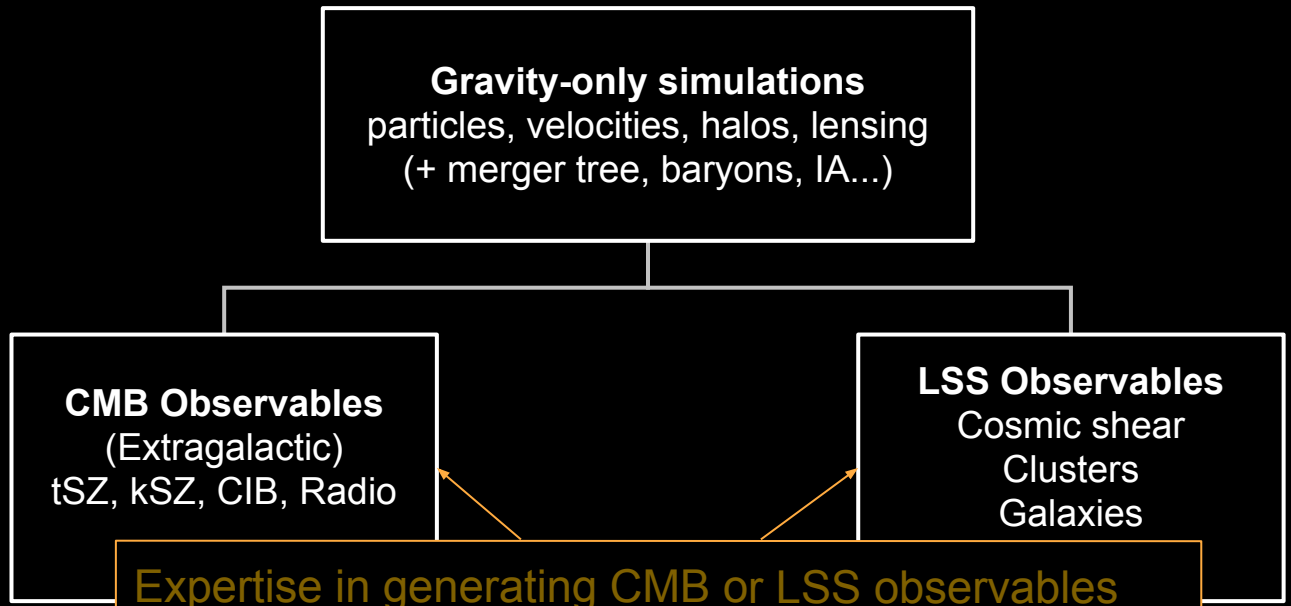
Roadmap to Stage 4 Correlated Simulations (**my 2 cents**)



Roadmap to Stage 4 Correlated Simulations (my 2 cents)



We can either (1) tile small/high-res boxes or (2) run a full-sky light cone with fast gravity-only codes (e.g. FastPM, COLA)



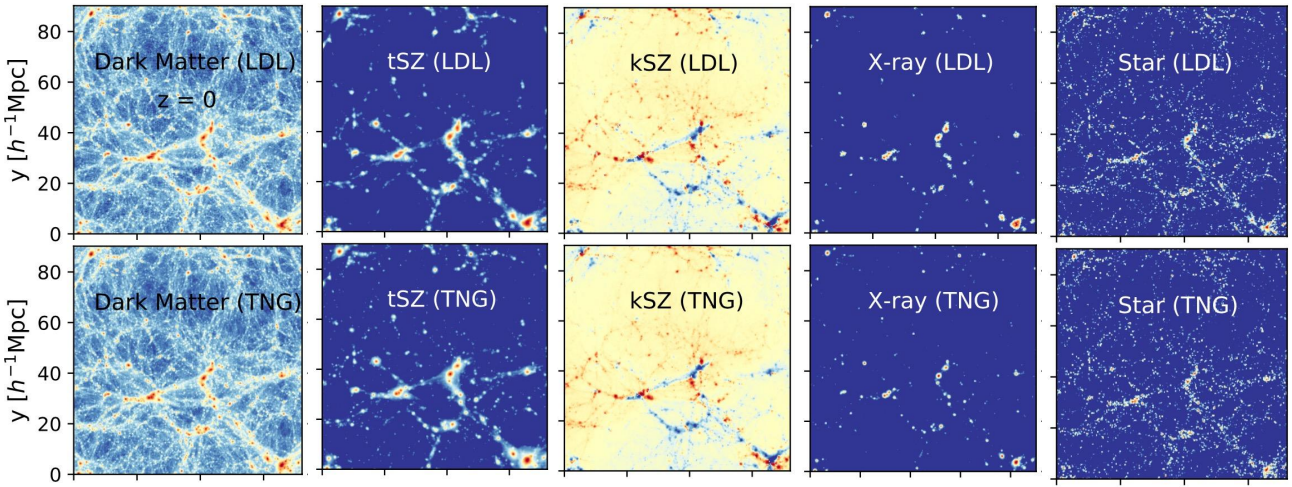
Expertise in generating CMB or LSS observables already exists in individual collaborations
— we just need to ask for help!

FastPM for the Gravity-only Simulations

Code: A particle-mesh gravity solver, with Potential Gradient Descent to improve small-scale clustering

Performance: 10^7 faster than hydro simulation; 10^4 faster than tree-PM N-body simulation; x10 slower than Websky (2LPT+peakpatch); used by the DESI collaboration

Opportunities: Lagrangian Deep Learning (astrophysics+painting); MADlens (differentiable lensing); flowPM (tensorflow-based/differentiability)



FastPM: Feng+2016
PGD: Dai+2018
LDL (*left img*): Dai & Seljak 2020
MADLens: Böhm+2020
FlowPM: Modi+2020

Challenges

- **Computing time** (for 10-100 correlated simulations)
  - A few x 10 million CPU-hours (need a large fraction of a supercomputer at once)
- **Storage (~PB)**
  - ~50TB per model (can be smaller if only keeping maps and catalogues)
- **Joint requirements and validation**
  - With limited resources, what are the key requirements?
  - Validation can take a long time
- **Maintenance**
  - Documentation
  - Data release
  - Future upgrades
- **Personnel**
  - Training and acknowledging the simulation scientists

Also see "Report from the Tri-Agency Cosmological Simulation Task Force" by Battaglia+2020

Interested & want to get involved?

- PhD projects
- Postdoc positions (to be announced in the fall)
- or simply share thoughts!

contact:
jialiu@berkeley.edu (email address expires on 8/31/2021)

https://liuxx479.github.io/ (for updated contact later)