



Simulations & Data Challenges: Extragalactic Foregrounds

Zack Li
on behalf of Half Dome and WebSky teams

CMB-S4 Collaboration Meeting
April 3-6, 2023



Science goals for extragalactic foreground simulations



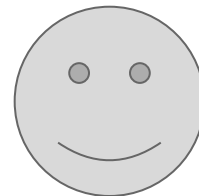
Validate data
analysis pipelines



Estimate
covariances



Emulators, predictions for
observables beyond 2-pt,
nonlinear analyses



Training sets for machine
learning, component
separation

Goals for deliverables

Stage IV cosmology

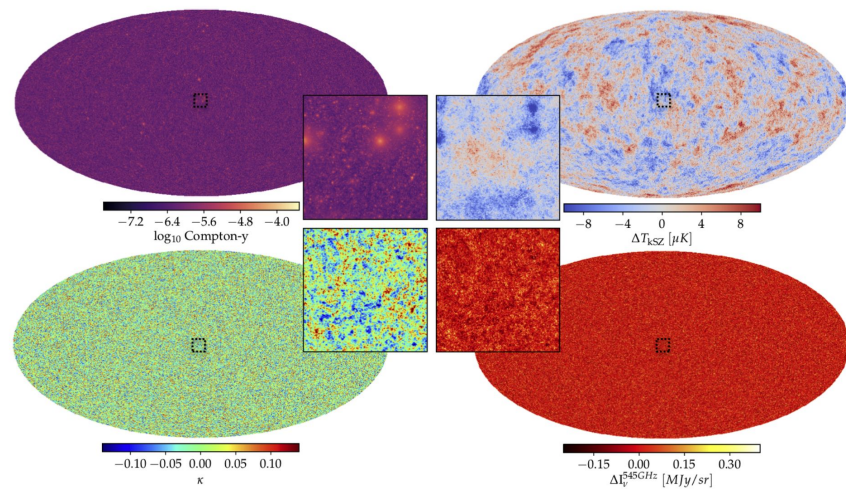
- Full-sky mocks of the millimeter sky
- Correlated with radio, IR, and optical galaxy surveys
- Agreement with existing data, particularly for cross-correlation
- Convenient PySM interface, $O(1000)$ realizations

Extensions and further science

- Beyond Λ CDM mocks with fNL, other non-gaussianity
- Multiple astrophysical models, as different as possible

DC-0 Status: WebSky 0.4

- Minor updates to CIB and higher nside, relative to Stein et al. 2020
- PySM interface for accessing the WebSky models, with interpolation for bandpass integration



HalfDome Team



Yici Zhong



Adrian Bayer



Yu Feng



Jia Liu



Joe DeRose



Zack Li



Giuseppe Puglisi



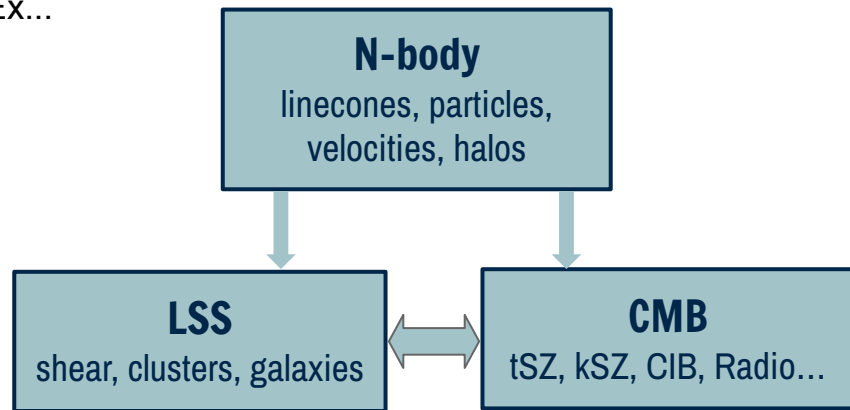
Mat Madhavacheril



Marcelo Alvarez

HalfDome Correlated Simulations: Philosophy

- Codes written with **future upgrades** in mind
 - Empirical, flexible astrophysical models to be regularly updated as observations improve
 - Codes in modules that can be updated/swapped independently
 - Validation tests for quality control when updating models/codes
- Designed for **cross-correlation** of **Stage IV** surveys
 - CMB surveys: CMB-S4, Simons Observatory, LiteBIRD...
 - LSS surveys: DESI, PFS, LSST, Roman, SPHEREx...
 - Aimed at production runs (1000s simulations)
 - Adopt fast codes (fastPM, rfof, CLF...)
 - Aggressive downsampling schemes



HalfDome Simulation Overview

- Simulation configuration:
 - $f_{\text{sky}} = 1$
 - $M_{\text{min}} = 10^{12} M_{\text{sun}}/h$
 - Box size = 5Gpc/h
 - $N_{\text{particle}} = 8192^3$
- Healpix and CAR mocks:
 - LSS: clusters, galaxies, shear, convergence
 - CMB: tSZ, kSZ, CIB, radio, CMB lensing
- Release ~ Fall 2023

Correlated Simulation Challenges

Assuming 1000 simulations*

1. Optical-CMB connections

AGN-Radio galaxies; SFR-CIB

Adding other observables (21cm, LIM, x-ray)?

2. Computing time: ~half billion CPU hours

Large memory need: quarter million cores simultaneously

3. Storage: ~10-100s PB

Raw: 500TB/run; aggressive downsampling needed->100TB per run

4. Cross-collaboration collaboration

Currently operate as a loose collaboration between people who are deeply embedded in DESI/LSST/SO/S4

Future access to multiple surveys' internal pipelines & proprietary data for future upgrades: N² MOUs??

Training and acknowledging simulation scientists

* A very conservative assumption..

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

A review of existing and upcoming extragalactic mocks

Sehgal <i>Sehgal et al. 2009, some updates in 2019 from Colin Hill</i>	One N-body realisation, this is a legacy dataset which remains useful for testing robustness to different astrophysics
WebSky <i>Stein et al. 2020, radio galaxies from Li et al. 2021</i>	Peak Patch simulations with halo-based response functions, widely used in Stage 3 analysis. Ongoing development for non-Gaussianities and line intensity mapping (CO and C II)
Agora <i>Omori 2022, LIM from Sato-Polito et al. 2023</i>	Halo-based models from a single MDPL2 halo realization shared with Universe Machine for galaxies, merger trees can provide more accurate galaxy physics
Half Dome <i>HD team in prep, Fall 2023</i>	Designed for Stage IV survey requirements, emphasis on empirical modeling and cross-correlations with galaxy surveys, $O(10-100)$ fastPM realizations for covariance

WebSky 2: focus on non-Gaussianity, LIM



Nate Carlson



Martine Lokken



Tom Morrison



Doga Tolgay



Dongwoo Chung



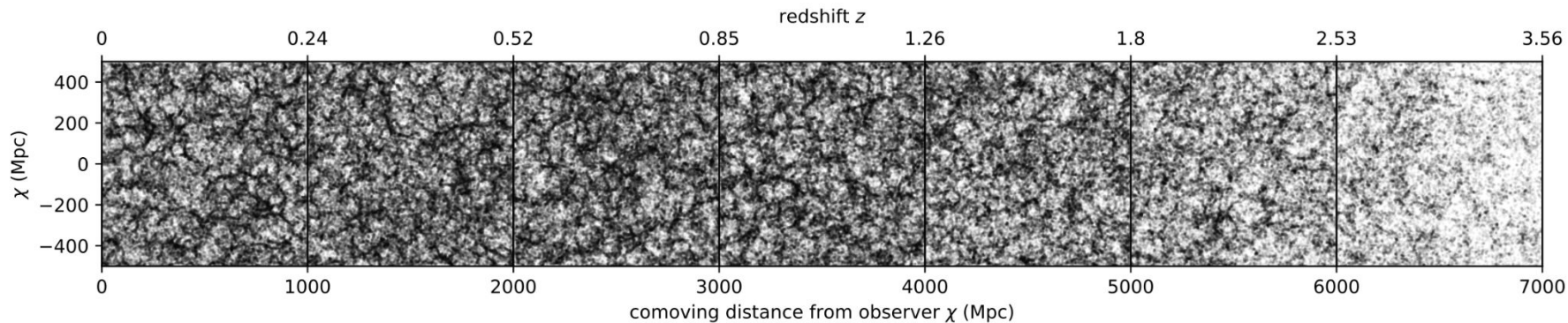
Zack Li



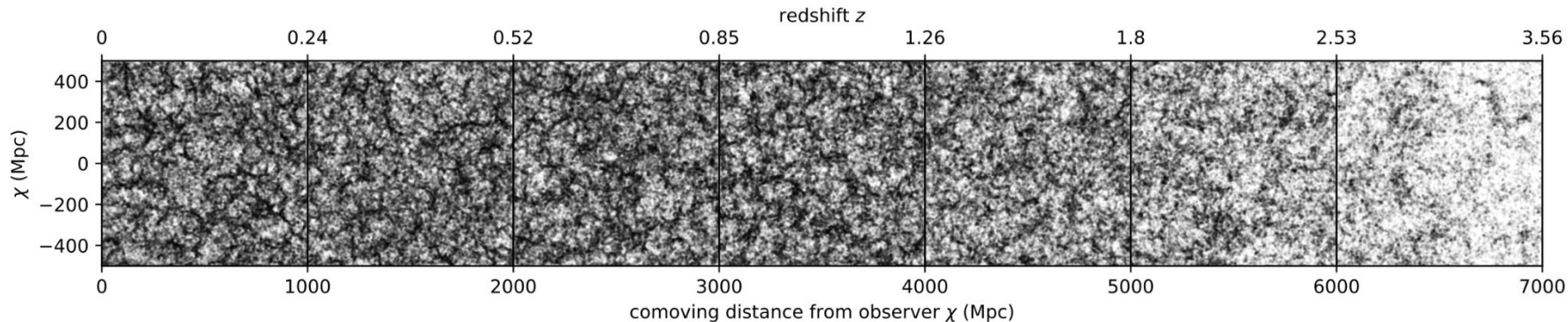
Dick Bond



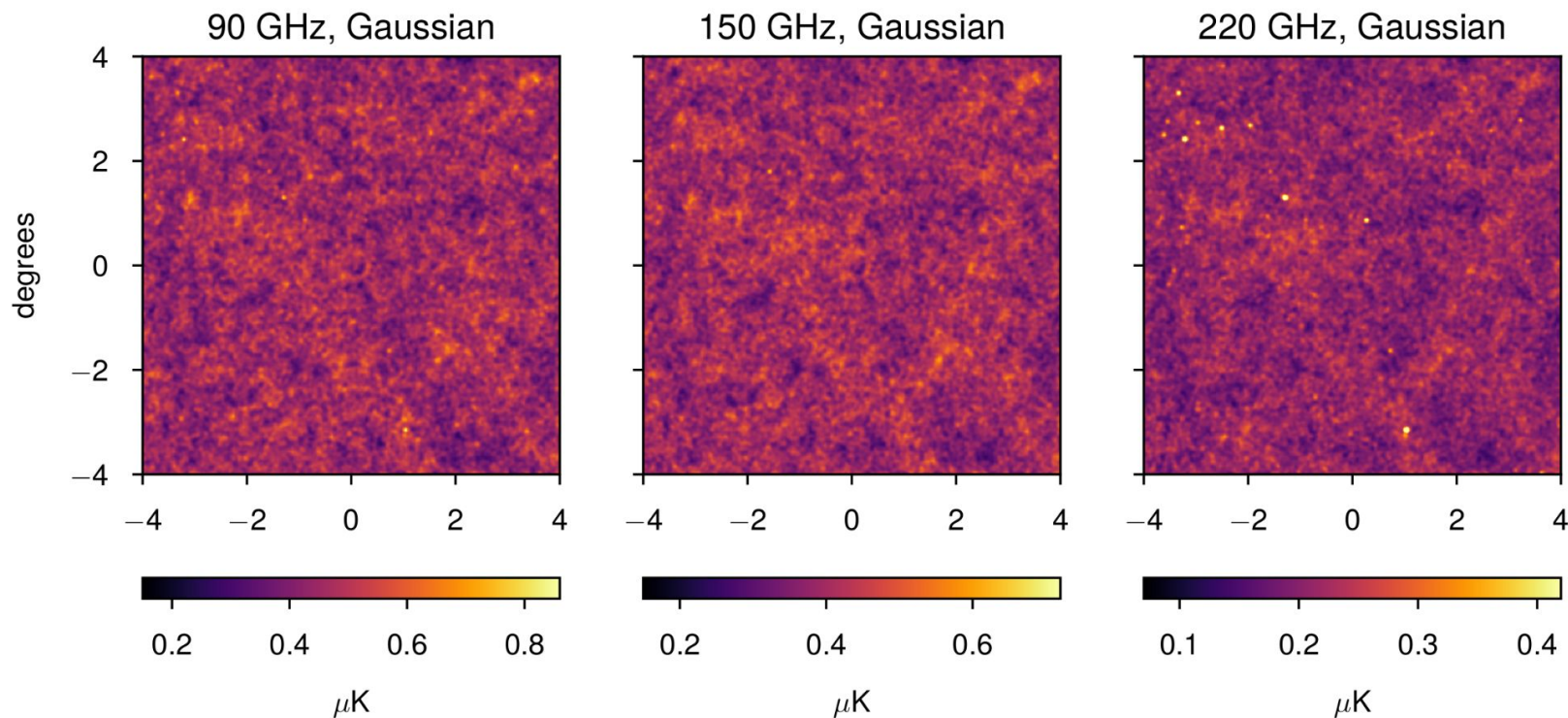
Gaussian Peak Patches



Non-Gaussian



Extragalactic CO lines seen through ACT bandpasses



Future and DC-1

- **Stay tuned for the Fall 2023 first release of Half Dome**, full-sky multiwavelength mocks designed with CMB-S4 and galaxy survey cross-correlations in mind
- WebSky 2 development will explore well-motivated forms of non-Gaussianity from inflation beyond fNL, as well as halo-based models for extragalactic CO and C II
- Having multiple millimeter sky simulations will be very useful for pipeline validation, particularly for machine learning methods which require training sets