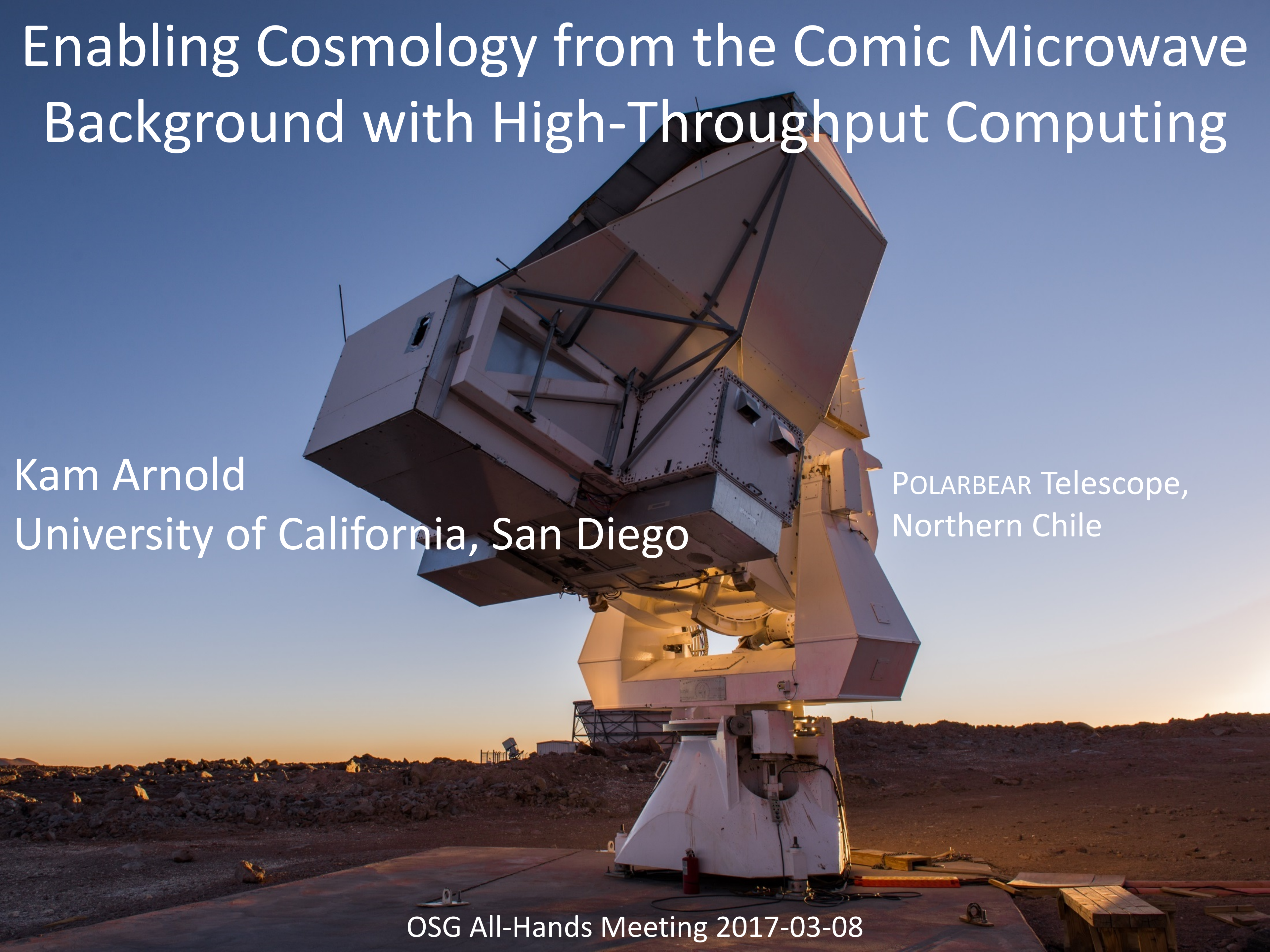
# Enabling Cosmology from the Comic Microwave Background with High-Throughput Computing

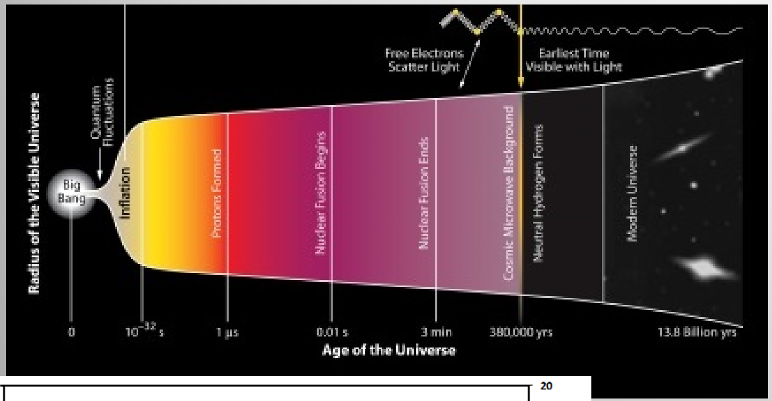
Kam Arnold
University of California, San Diego

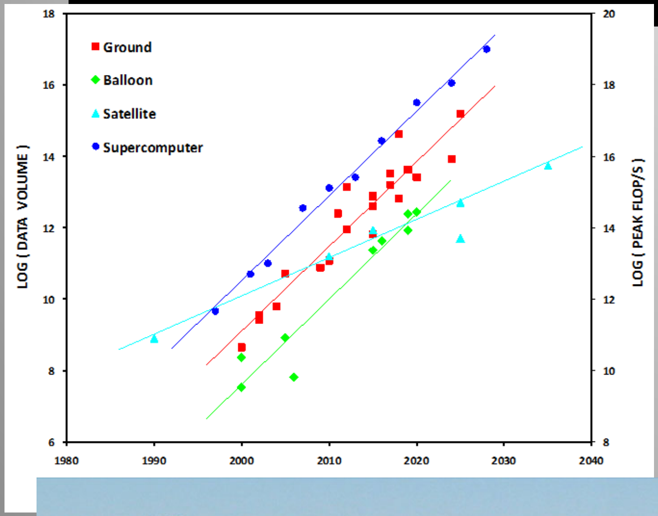


POLARBEAR Telescope, Northern Chile

OSG All-Hands Meeting 2017-03-08

# Outline



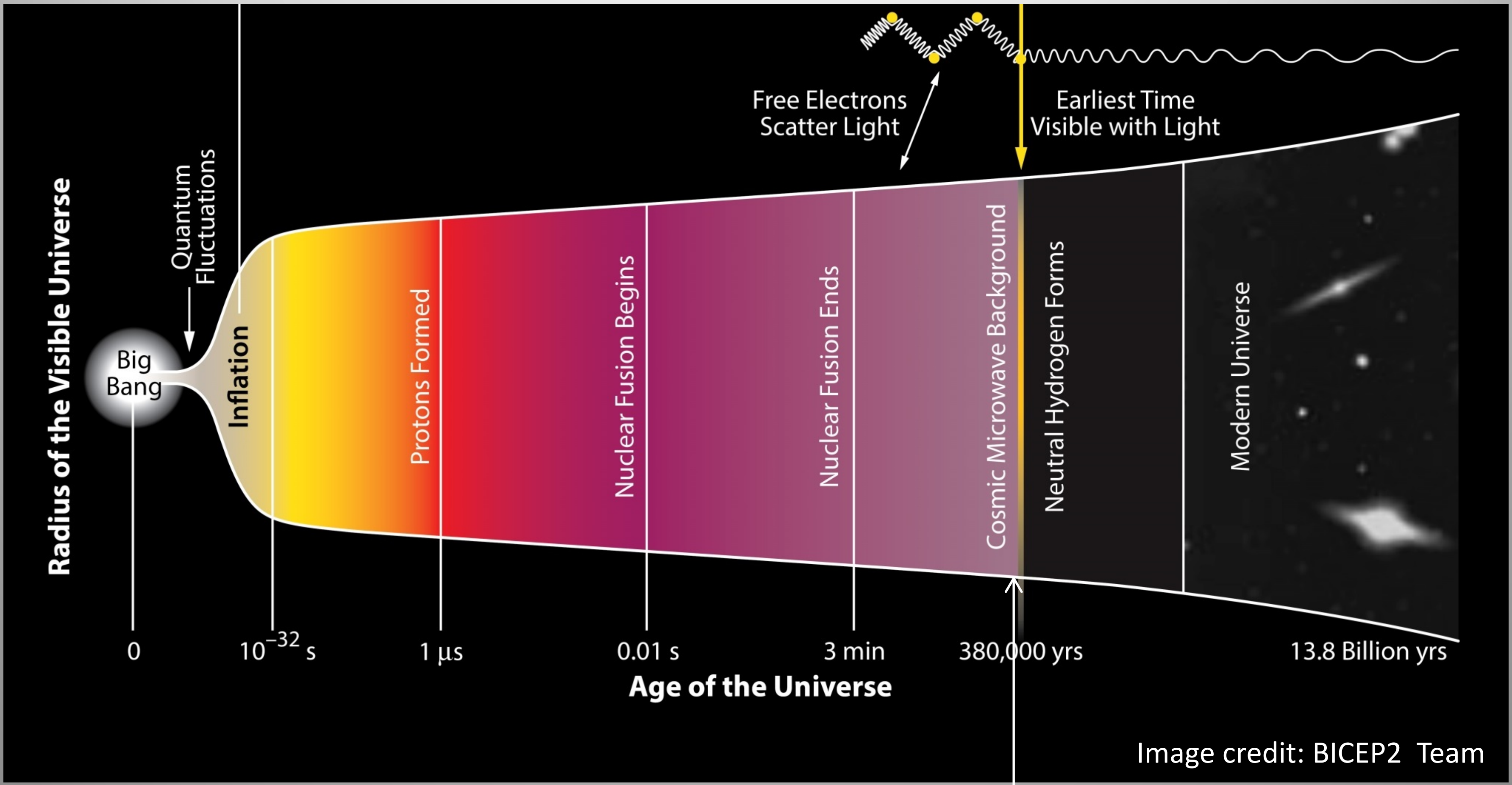


- Scientific Motivation
- The Computational Problem: what's a good fit to the OSG
- Looking forward to experiments in the next decade

See also: Monday talk by Benedikt Riedel about South Pole Telescope & OSG

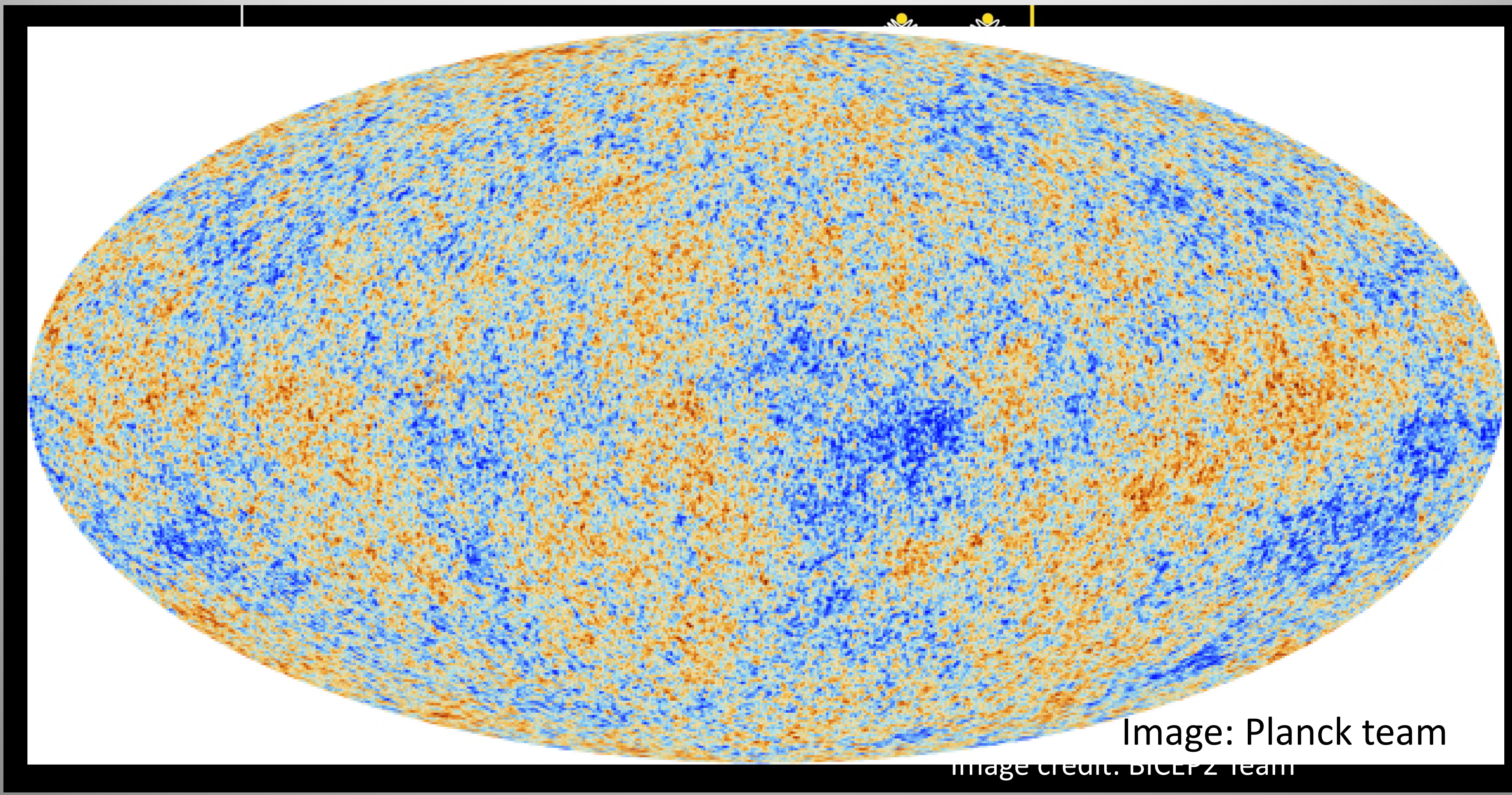
# SCIENTIFIC MOTIVATION

# Introduction to the CMB




Photon decoupling

# Introduction to the CMB

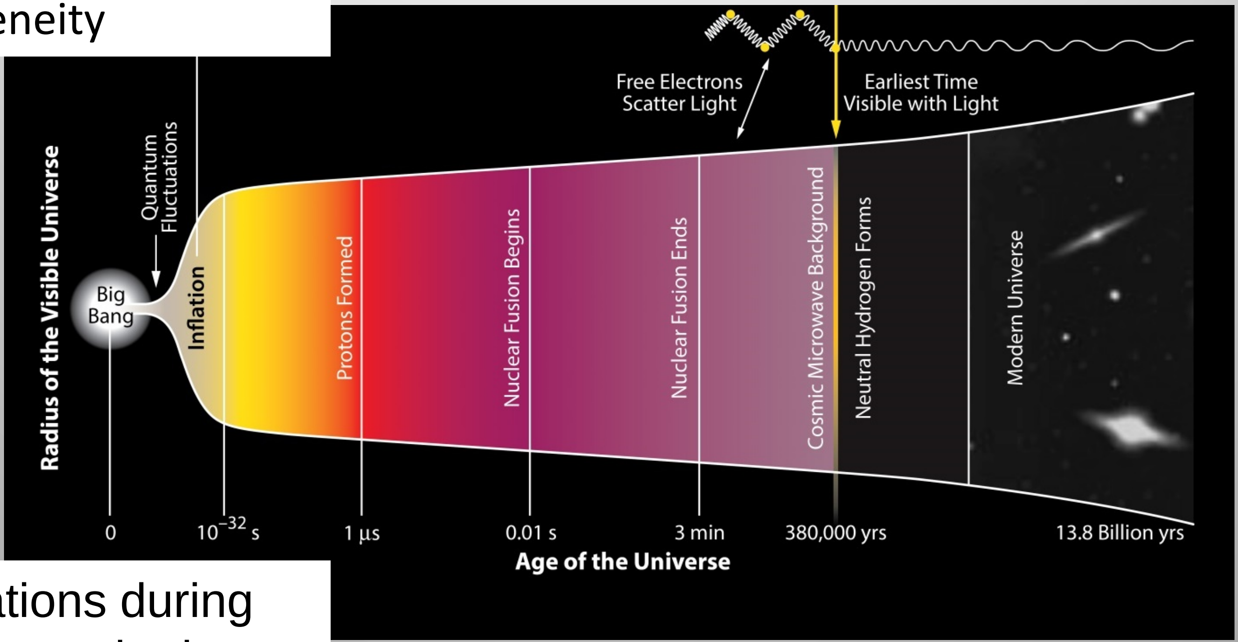



Color scale $\pm 2 \cdot 10^{-4} K$

$T_0 = 2.7K$

# Introduction to the CMB

Inflation does not predict complete homogeneity

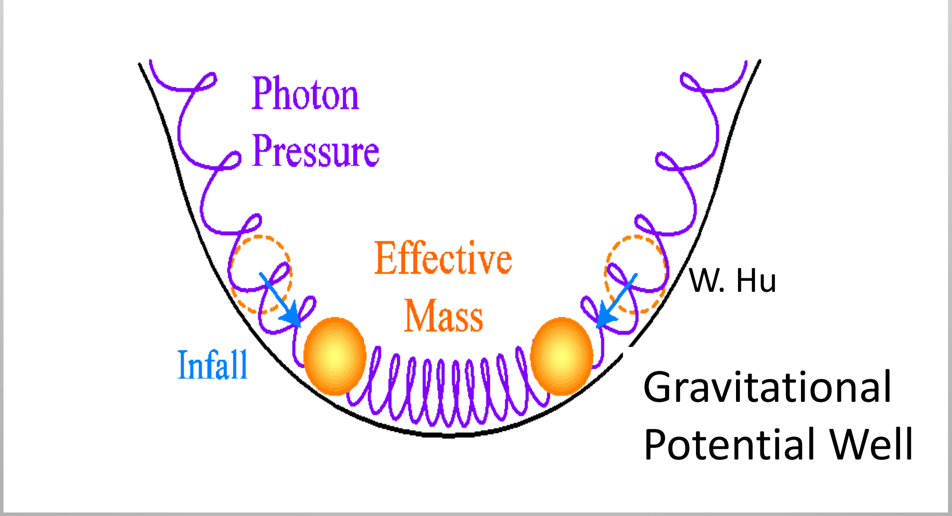



Quantum fluctuations during inflation imprint perturbations on the gravitational metric

metric perturbations separable into three types: ***scalar***, vector and ***tensor – inflationary gravitational waves***.

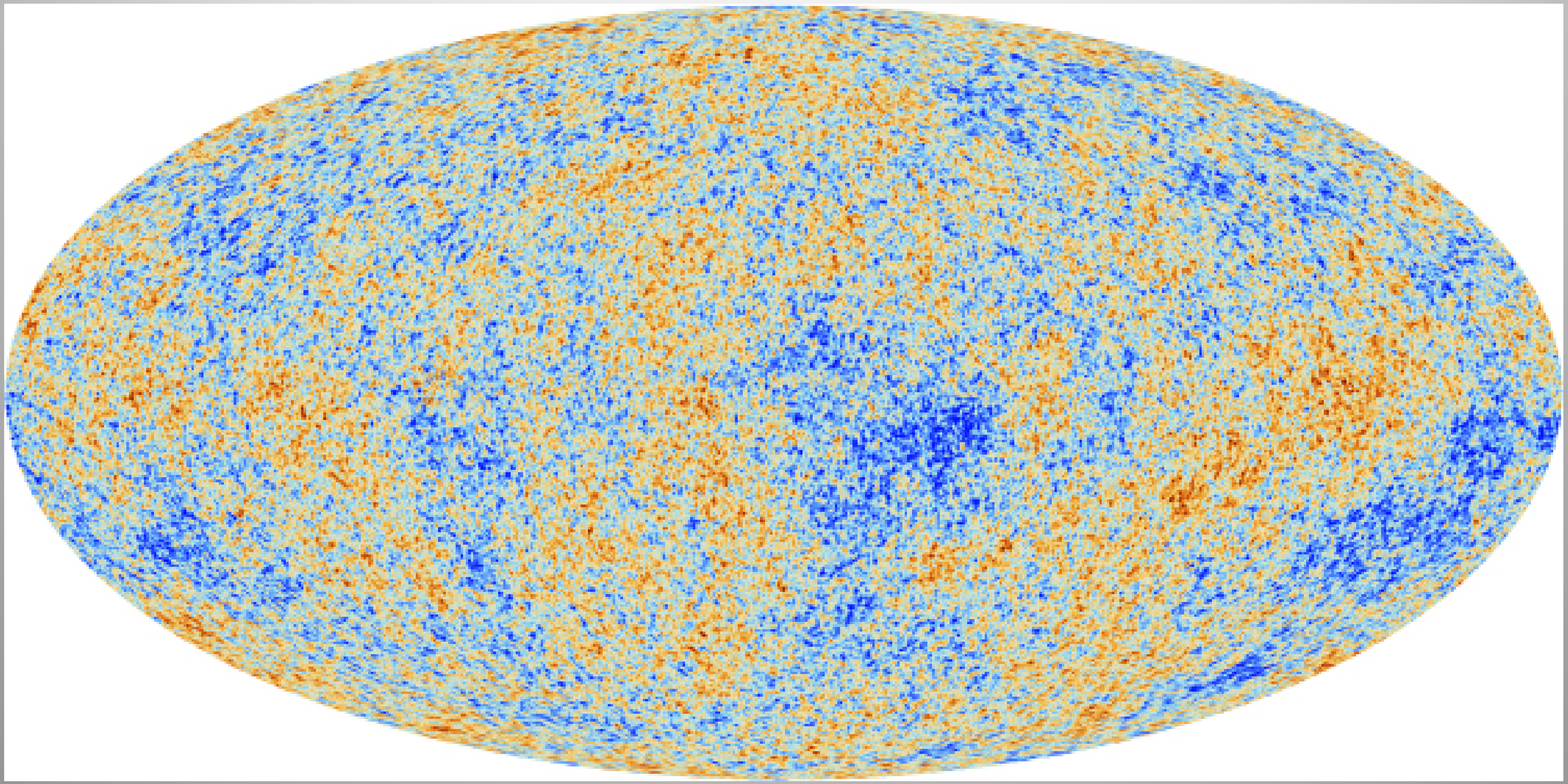
# Fluid Oscillations in Pressure & Density

## Photon-Electron-Nucleon Fluid




Period of oscillation ∝ spatial extent of potential well

# Map of CMB Temperature



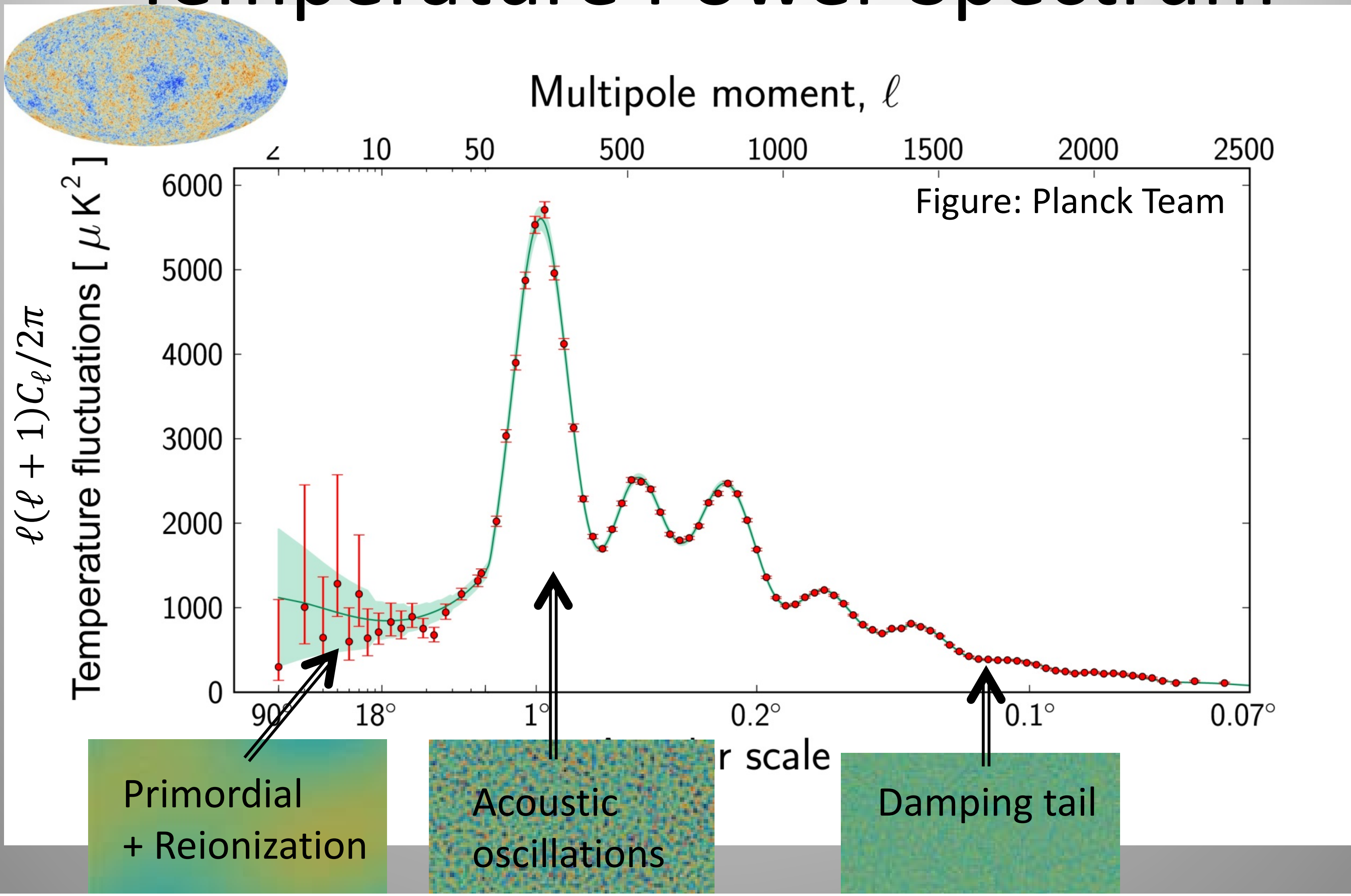
Decompose the temperature map into spherical harmonics

$$T(\hat{n}) = \sum_{\ell,m} a_{\ell m} Y_{\ell m}(\hat{n})$$

Calculate the expectation value of the correlation between the coefficients

$$\langle a^*_{\ell m} a_{\ell' m'} \rangle = C_\ell^{TT} \delta_{mm'} \delta_{\ell\ell'}$$

# Temperature Power Spectrum

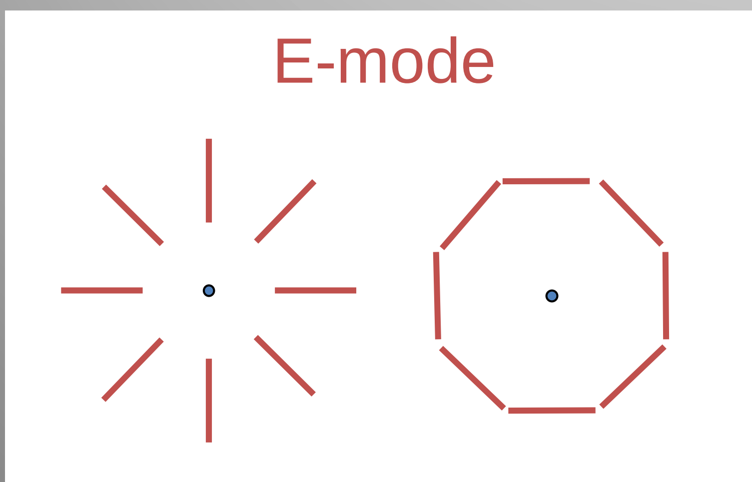


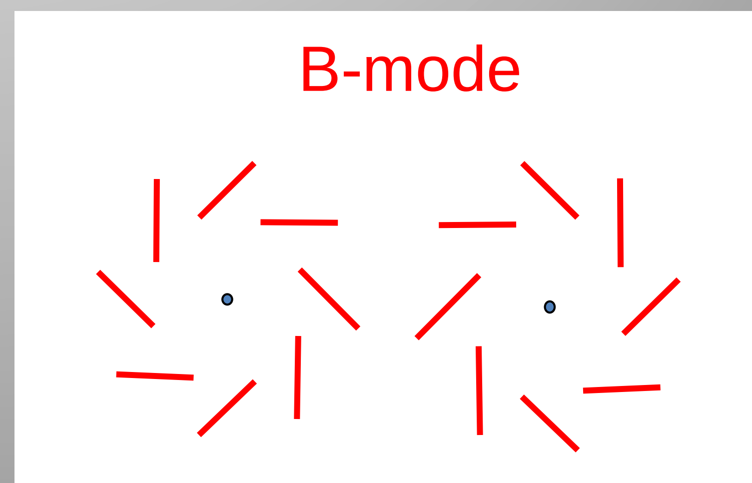

**Provides precision measurement of several cosmological parameters including dark matter content & geometry**

# Polarization of the CMB

Parity-symmetric "E-modes" are the dominant polarization in the CMB. They are created by the acoustic oscillation physics

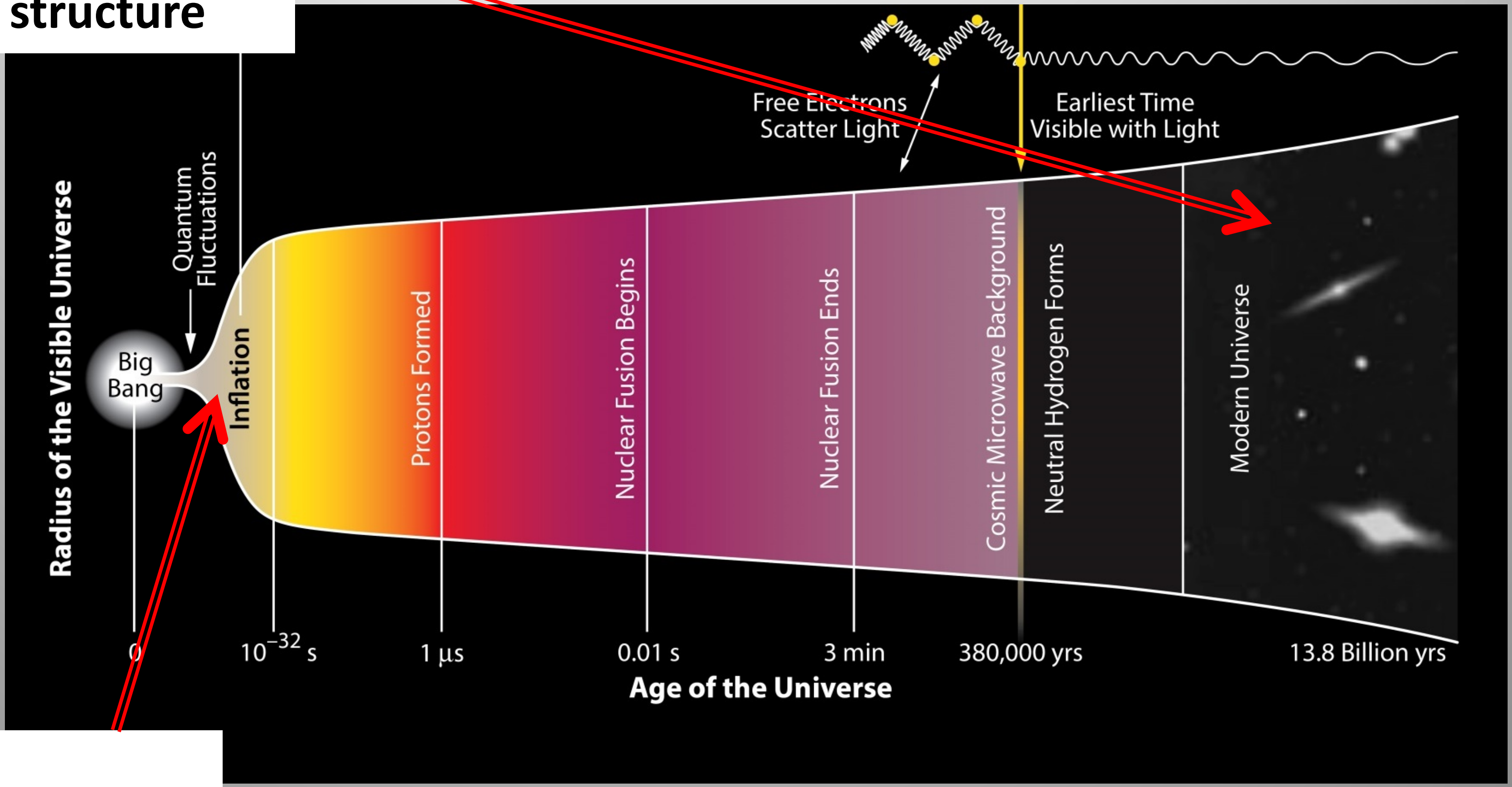
Parity anti-symmetric "B-modes" are evidence of physics other than the acoustic oscillations.



# B-mode Polarization: Two Sources

**Gravitational lensing of E-mode polarization by intervening structure**

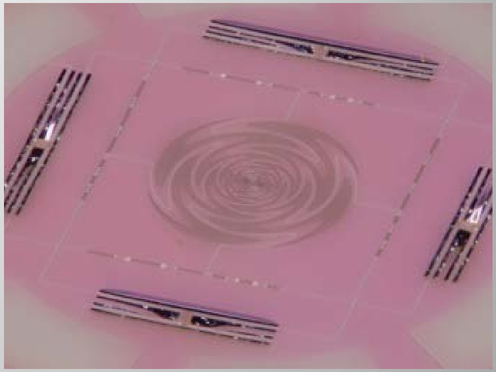



**Inflationary gravitational waves**

# THE COMPUTATIONAL PROBLEM

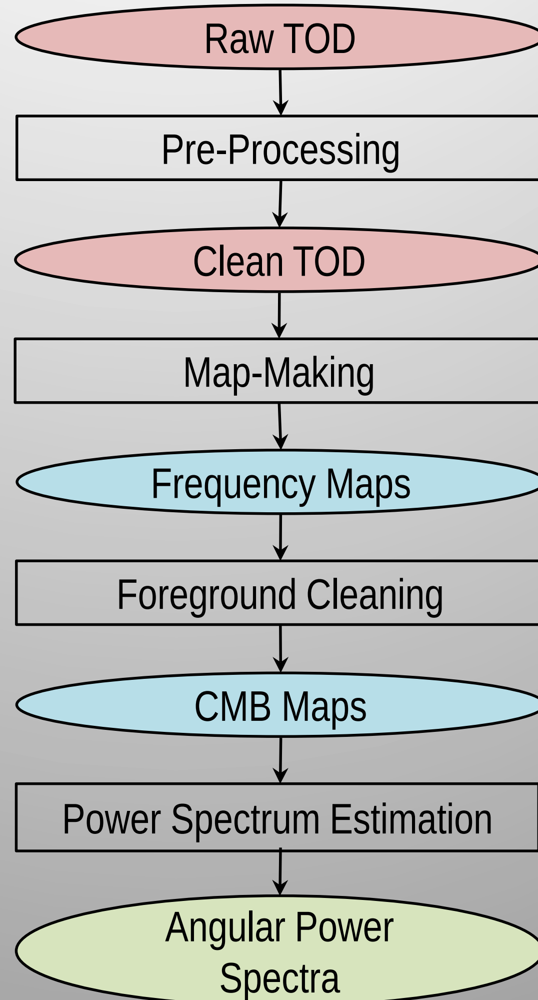
# The Measurement

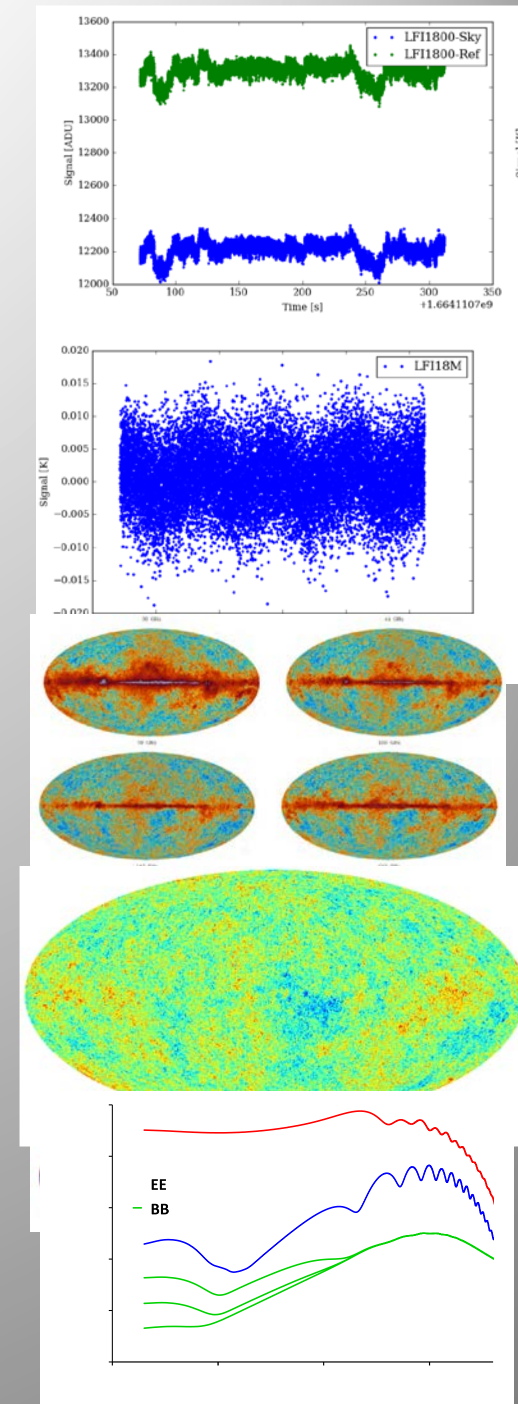
- Use detectors sensitive to polarization
- Use detectors with different spectral sensitivity
- Use a telescope with sufficient angular resolution
- Scan the telescope across the sky and measure detector response as a function of time $d_t$ (Time-Ordered Data, TOD)
- Repeat scans for years to measure the true sky signal while averaging away noise, atmosphere, etc.



# Data Reduction

- Sequence of S/N-increasing data compressions via domain transformations:
  - Time samples: $\mathcal{N}_t$
  - Pixels: $\mathcal{N}_p \ll \mathcal{N}_t$
  - Multipoles: $\mathcal{N}_\ell \ll \mathcal{N}_p$
- Each domain exposes different systematics => iterative looping.
- Must propagate both data *and their covariance* for a sufficient statistic.



# Exact CMB Analysis

Model TOD $d_t$ (time-ordered data) as noise and sky-synchronous CMB:

$$d_t = n_t + P_{tp} s_p$$

Estimate the noise correlations from the (noise-dominated) data

$$N_{tt}'^{-1} = f(|t - t'|) \sim invFFT\left(\frac{1}{FFT(d_t)}\right)$$

*Analytically* maximize the likelihood of the map given the data and the noise covariance matrix $N$

$$m_p = \left(P^T N^{-1} P\right)^{-1} P^T N^{-1} d_t$$

Construct the pixel domain noise covariance matrix and iteratively maximize the likelihood of the CMB angular power spectrum given the map and its covariance matrix

# The Exact Analysis Challenge

Science goals drive us to observe more sky, at higher resolution, at more frequencies, in temperature and polarization.

| | BOOMERanG (2000) | Planck (2015) |
|---|---|---|
| Sky fraction | 5% | 100% |
| Resolution | 20’ | 5’ |
| Frequencies | 1 | 9 |
| Components | 1 | 3 |
| Pixels | $O(10^5)$ | $O(10^9)$ |
| Operations | $O(10^{15})$ | $O(10^{27})$ |

Exact methods are no longer computationally tractable.

# Approximate CMB Analysis

- Produce filtered, biased, **less computationally expensive** maps as a tool to get to power spectra
- Use Monte Carlo simulations to determine effect of imperfect mapmaking on power-spectrum estimation
  - **Simulations required anyway to understand the effect of instrument systematic uncertainties**
- Dominant cost is simulating & mapping time-domain data for Monte Carlo realizations: $O(\mathcal{N}_{mc}\mathcal{N}_t)$
  - Number of Monte-Carlo realizations needs to be $\mathcal{N}_{mc} \sim O\left(10^3\right)$

# Simulations of CMB Analysis

- Linear algorithms reduce calculation costs but I/O & communication costs become more significant
  - Need to read in entire $\mathcal{N}_t$ TOD to make a map
- But, experiment Monte Carlo simulations **create the TOD and then map it on-the-fly, never having to do the I/O** – these simulations are a good fit to the OSG.
- Another good fit: analyses that take existing maps or sub-maps as input
  - POLARBEAR Collaboration's first result (measuring lensing with the CMB): **post-map analysis done on the OSG, led here at UC San Diego**

# EXPERIMENTS IN THE COMING DECADE

# $\mathcal{N}_t$: existing & future experiments & supercomputers

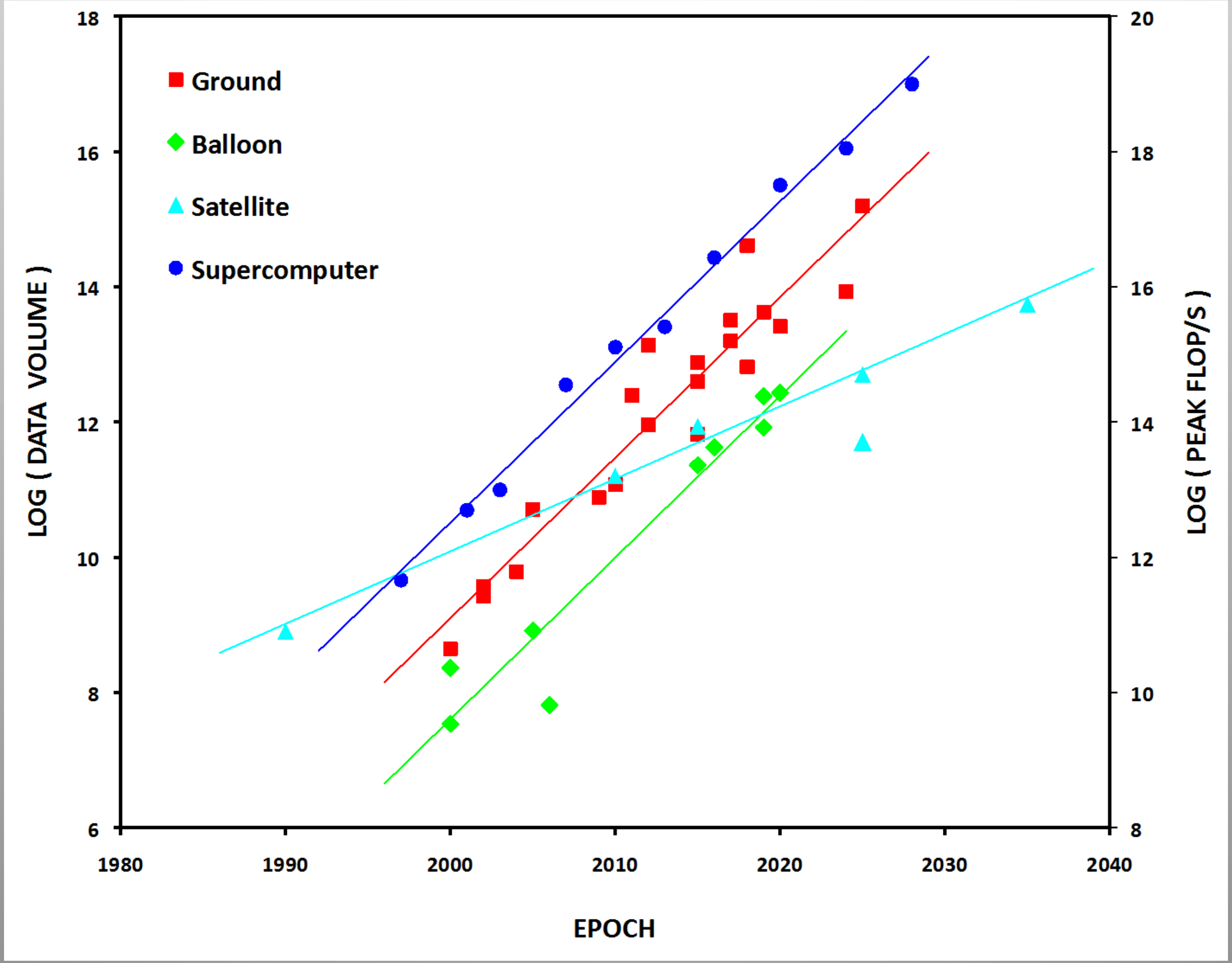



Figure: Julian Borrill, LBNL NERSC

# $\mathcal{N}_t$: existing & future experiments & supercomputers

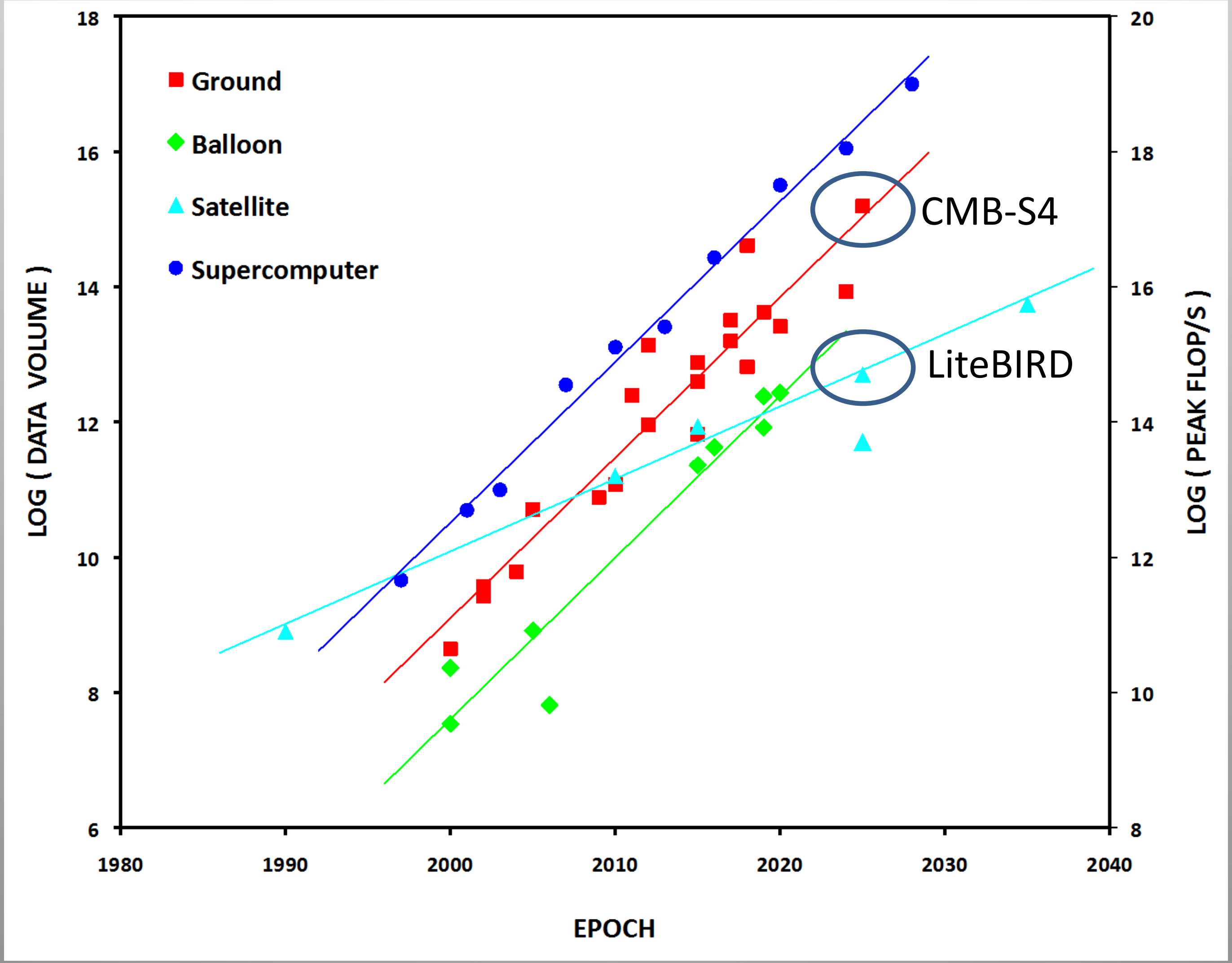



Figure: Julian Borrill, LBNL NERSC


- Currently in DOE/NSF joint concept definition - approaching critical decision 0.
  - Four community-wide conferences over the past 2 years to define project
  - Likely configuration: deploy $O(10^6)$ bolometric detectors spread over several telescopes at Chile and the South Pole.
- Two order of magnitude increase in $\mathcal{N}_t$ from today's experiments
- Requires lower floor for systematics residuals in data analysis
  - **More detailed & computationally expensive simulations (larger $\mathcal{N}_{mc}$).**
  - More complex mitigation in pre-processing.
- Hope to use OSG as an important tool for simulations, combined with NERSC-like resource
  - South Pole Telescope is leading the way here, Simons Array (currently deploying in Chile) has done some analysis on OSG, hopefully more
  - Joining forces in this effort for CMB-S4.

# Conclusion

- The Cosmic Microwave Background can give us information about cosmology and the fundamental physics of our universe
- The computational requirements are large, even with approximate methods, but individual instrument simulations can be packaged for OSG-like systems
- In the future, leveraging systems like OSG will be an important component of the analysis pipeline