A Null test framework for B-mode measurements with POLARBEAR

Yuji CHINONE（茅根裕司）

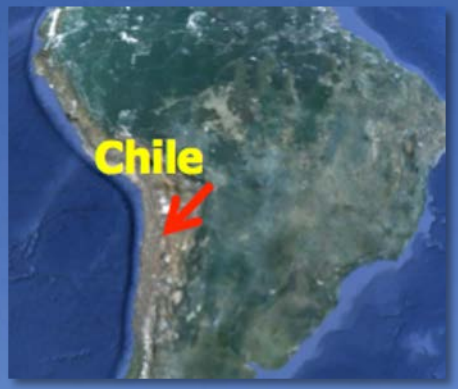
University of Tokyo

20:35-21:00, Wed, 2nd Dec 2020

POLARBEAR since 2012 (deployed in 2011)

ended? on 2017

5,200m @ Atacama Desert, Chile

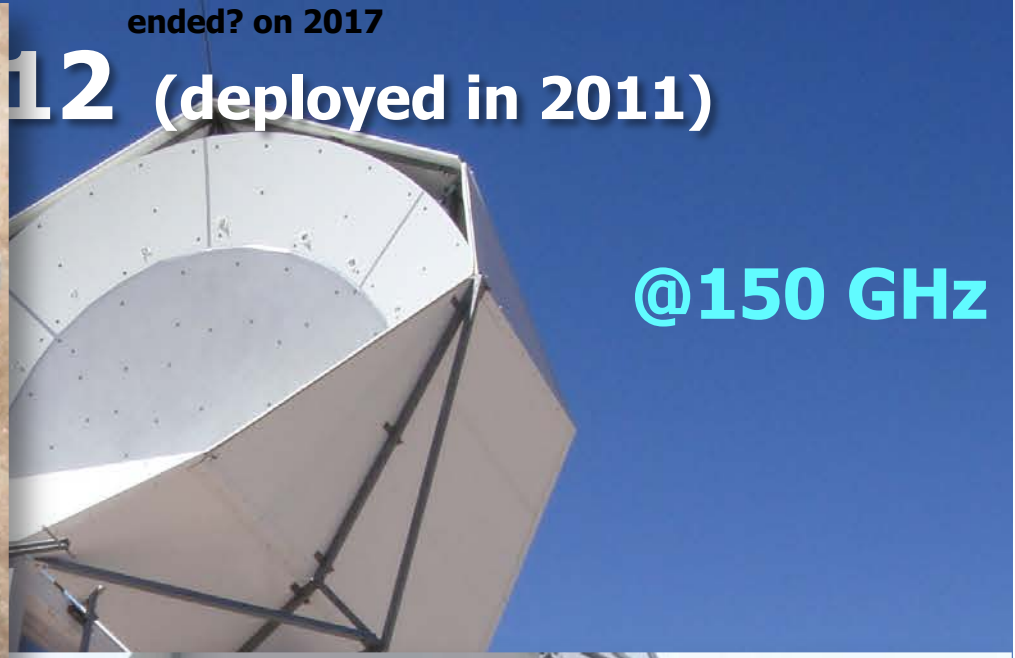


@150 GHz

ended? on 2017

12 (deployed in 2011)

@150 GHz



REMINDER: Our Scientific Publications

1. **Cross spectrum BTW lensing deflection w/ B mode x Cosmic Infrared Background (2014PhRvL.112m1302A)**
2. **First Power spectrum from lensing deflection w/ B mode (2014PhRvL.113b1301A)**
3. **First CMB lensing B-mode auto power spectrum (2014ApJ...794..171P)**
4. **Cosmic birefringence & primordial magnetic field (2015PhRvD..92l3509A)**
5. **Second CMB lensing B-mode auto power spectrum (2017ApJ...848..121P)**
6. **Cross spectrum BTW lensing by polarization x Subaru Hyper Suprime-Cam (2019ApJ...882...62N)**
7. **Cross spectrum BTW lensing by polarization x Herschel-ATLAS (2019ApJ...886...38A)**
8. **Degree scale CMB B-mode auto power spectrum (2020ApJ...897...55P)**
9. **Second Power spectrum from lensing deflection w/ B mode (2020ApJ...893...85F)**
10. **Internal CMB delensing (2020PhRvL.124m1301A)**
11. **Sub-Degree scale CMB E-mode auto power spectrum (2020arXiv200506168A)**

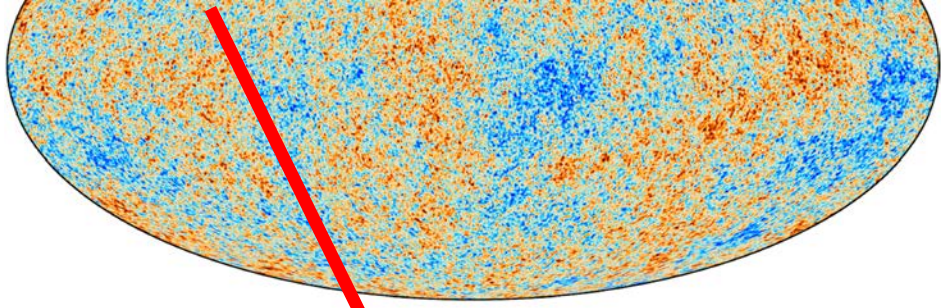
deep survey/wide survey

Introduction

- POLARBEAR performed **a blind analysis** for all the CMB science results
- **A critical framework is a "null test"**
- Developed the "**null test framework**" to **improve** & **validate our data/analysis** *iteratively before looking at final spectra*
  - I know (almost?) all the modern CMB experiments perform a null test (="jackknife test")
- For POLARBEAR, it's NOT just a **validation test**, BUT it's a **driving force** of our **data analysis**

Fukuwarai:
While blindfolded,
the players attempt to place the features onto the face in the correct positions.

Sky



$N_{\rm eff}$,
r,
$A_{\rm lens}$,
$M_\nu, \ldots$

Parameter estimate

Sky

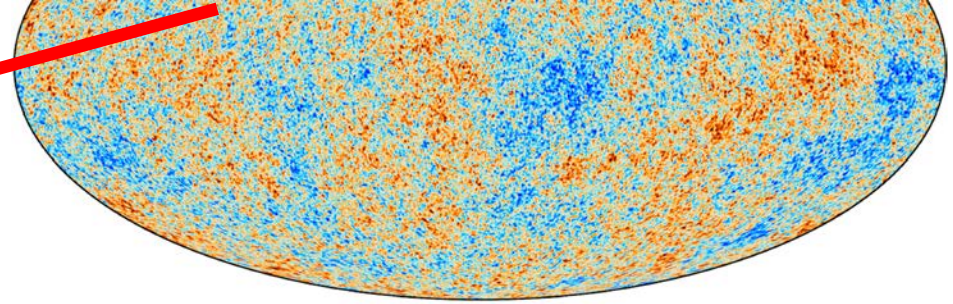
Measurement

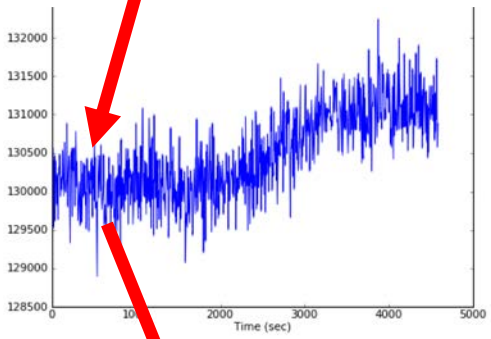
TOD
(Time-order data)

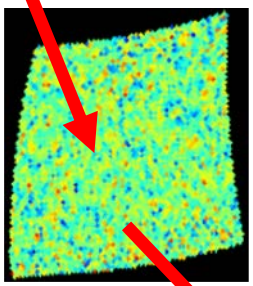
Map-making

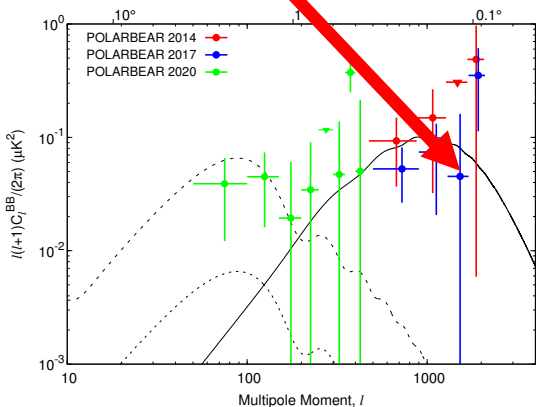
Power spectrum

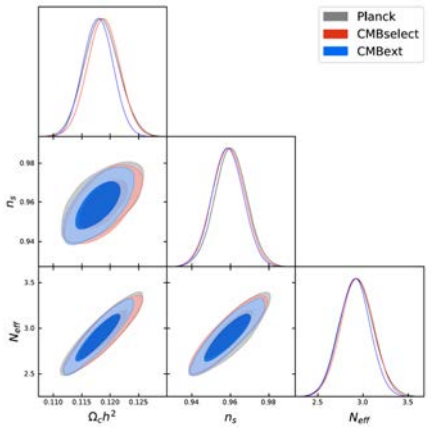
Parameter estimate











$N_{\rm eff}$,
r,
$A_{\rm lens}$,
$M_\nu, \ldots$

Sky

Measurement

Raw TOD

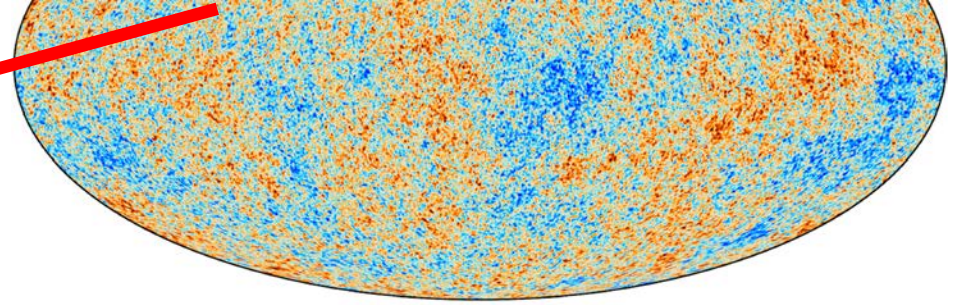
Calibration

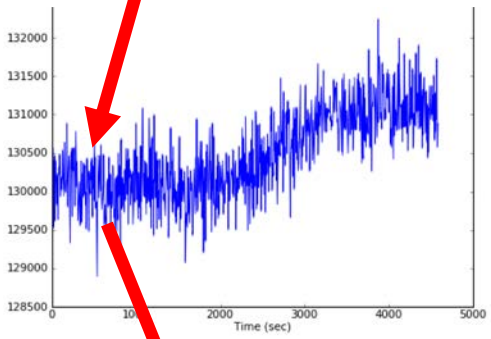
TOD processing

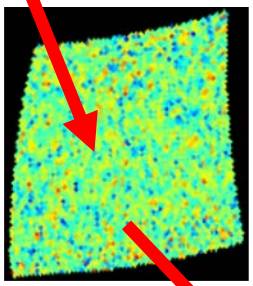
Map-making

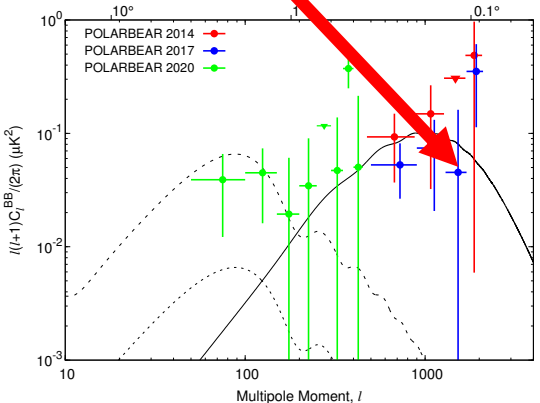
Power spectrum

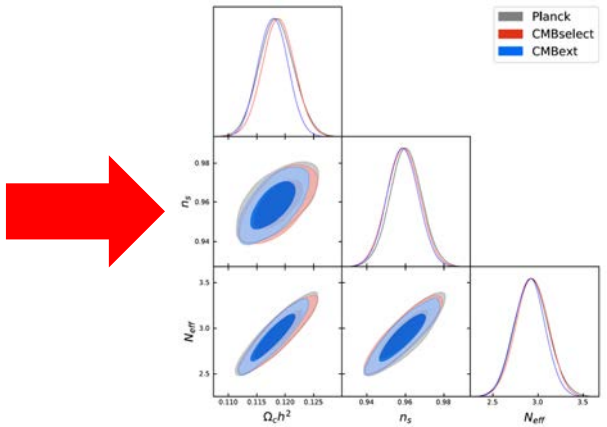
Parameter estimate











$N_{\rm eff}$,
r,
$A_{\rm lens}$,
$M_\nu, \ldots$

Sky

Measurement

Raw TOD

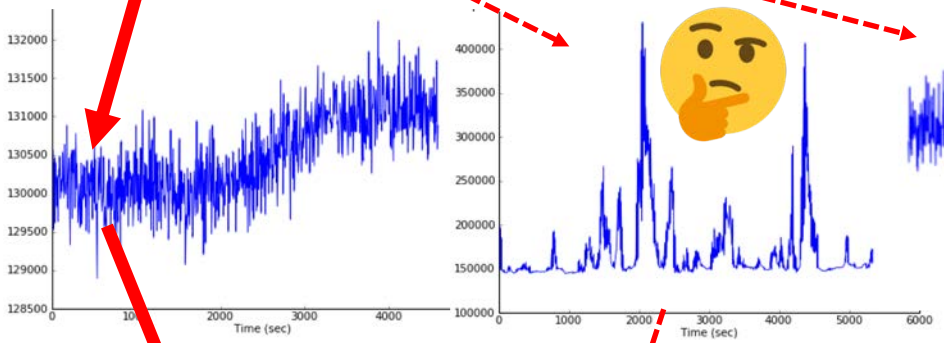
Calibration

TOD processing

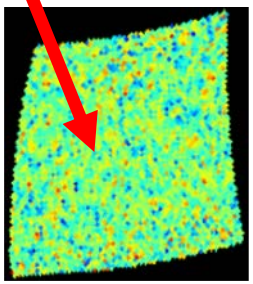
Map-making

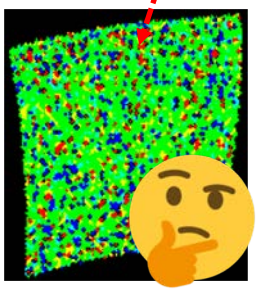
Power spectrum

Parameter estimate

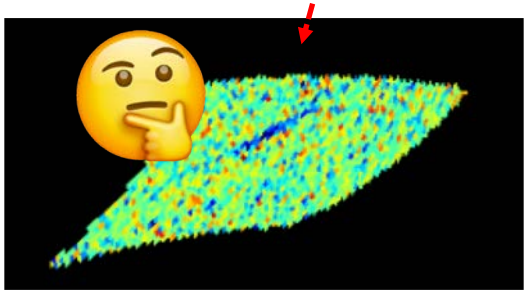


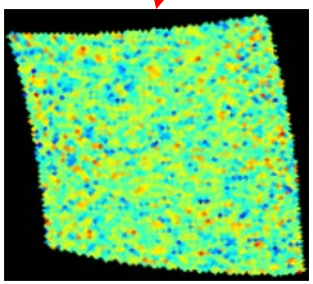
CR glitch,
Sidelobe pickup,
...





noisy,
looking
at clouds,
...





Sky

Measurem

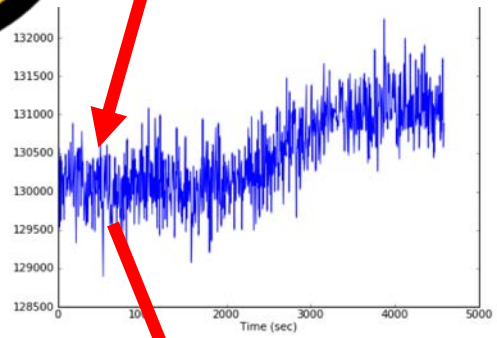
Raw TOD

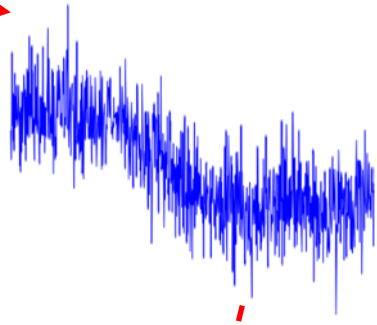
Calibration

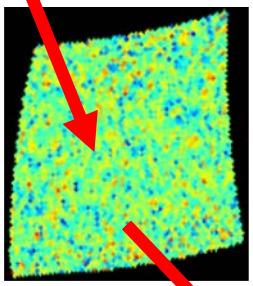
TOD processing

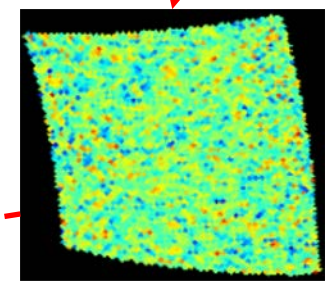
Map-making

Data selection

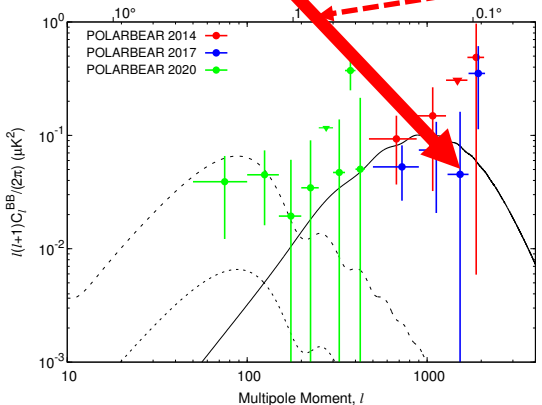


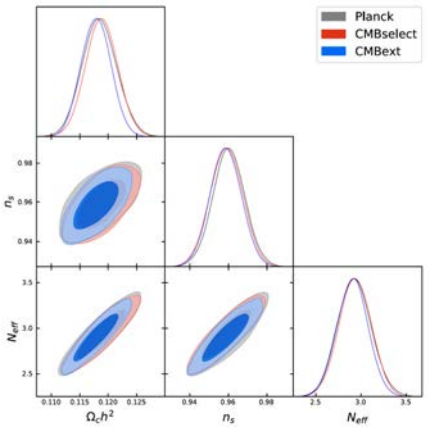






Power spectrum





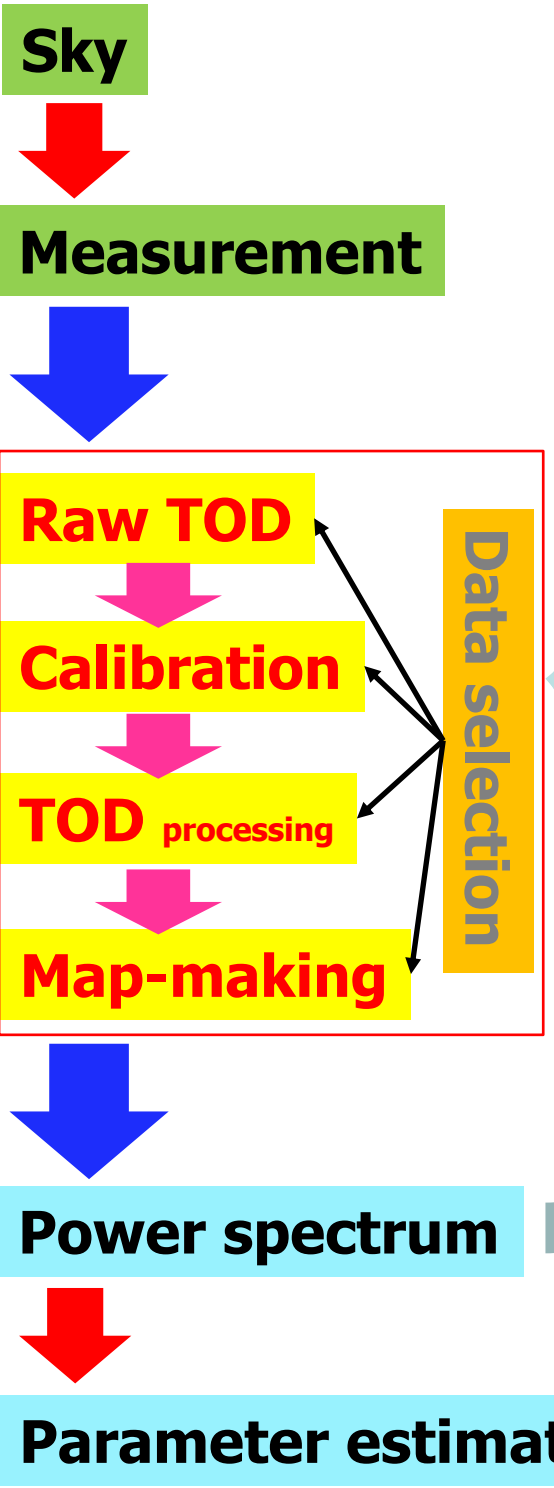
Parameter estimate

$N_{\rm eff}$,

r,

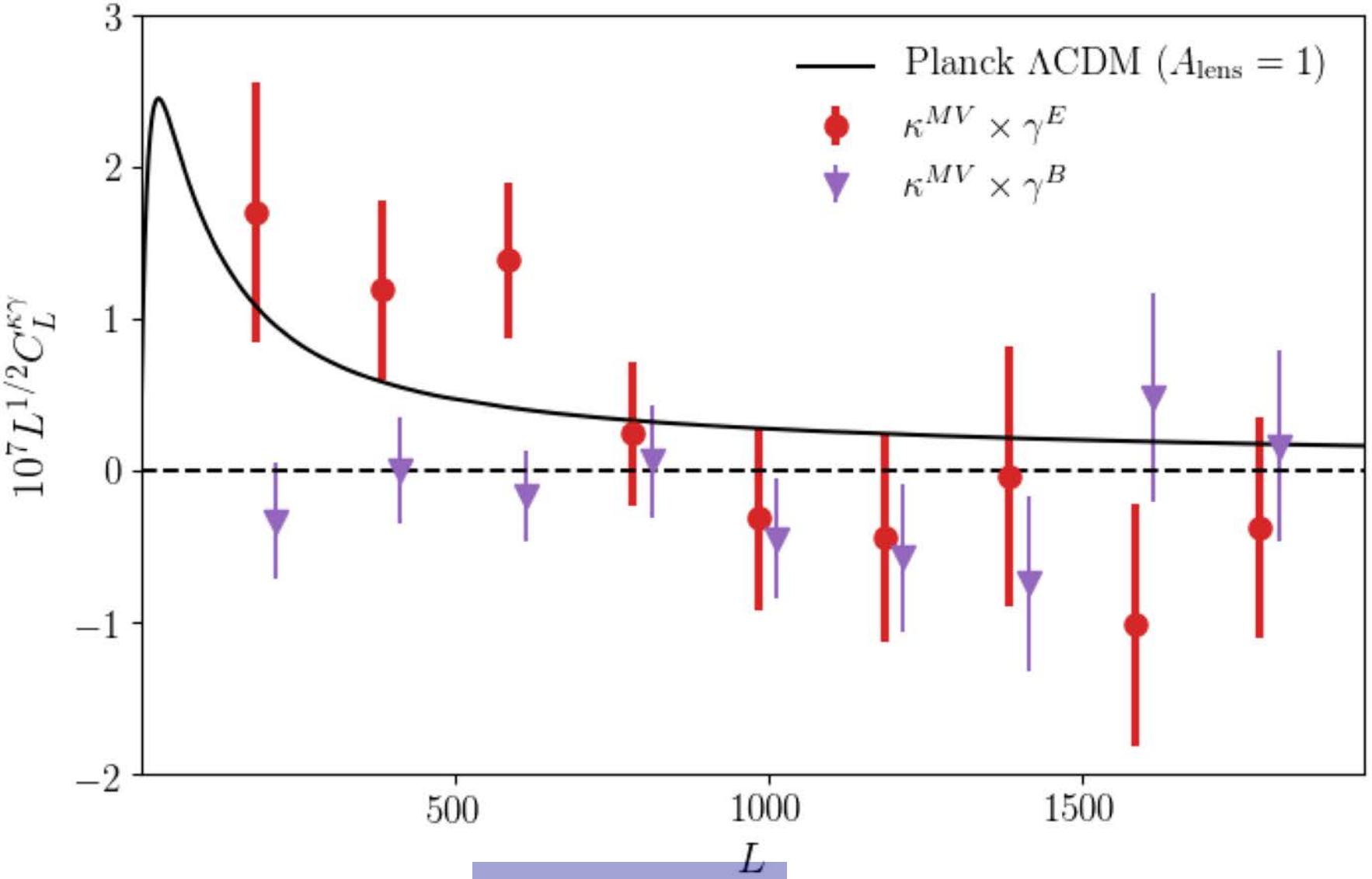
$A_{\rm lens}$,

$M_\nu, \ldots$



- Need "iteration" & "modification" of analysis
  - but how to stop them
  - wondering whether to judge it **based on (looking at) power spectrum / parameters** is **reasonable?**
- Should we worry that someone change their analysis
  - **so that** the result looks consistent w/ previous results and/or expectation, e.g. **LCDM**?,
  - or **so that someone can make a great discovery?** **(i.e. confirmation bias)**

Higher Lensing Amplitude by POLARBEAR x HSC



- Resulting amp. of $A_{\rm lens}$=**1.70±0.48**, rejecting the null hypothesis at **3.5σ**
  - This is the first measurement achieved by polarization-dominated CMB lensing
- Prefers higher lensing amplitude? ***Any new physics?*** ***or wrong analysis?***

Curiosities on Lensing Amplitude?

Planck 2018 results. VI.

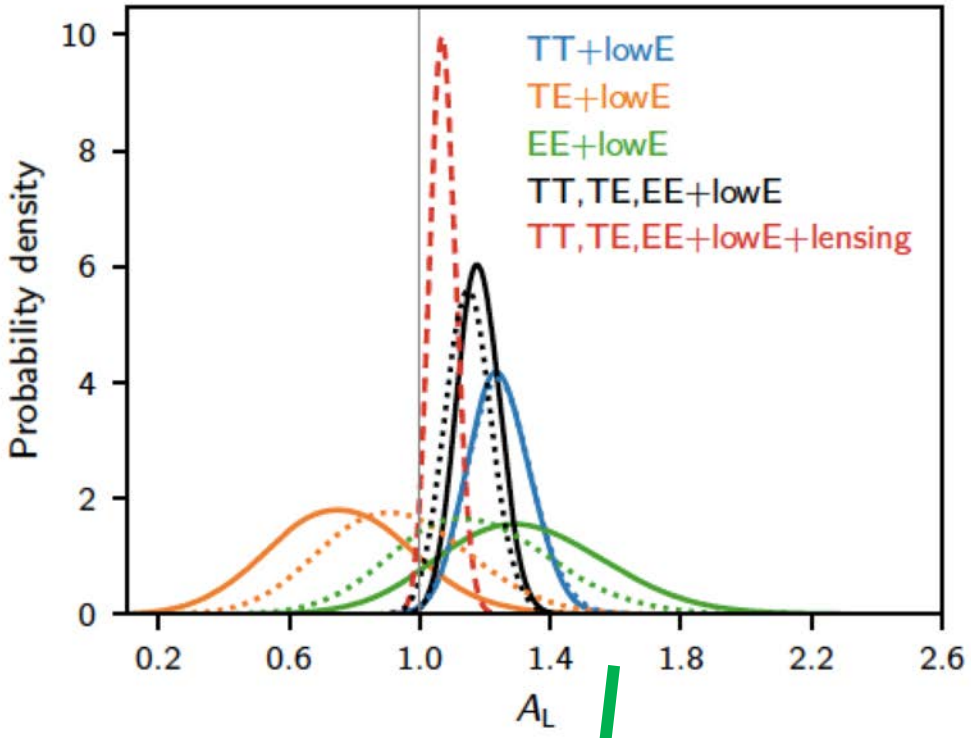
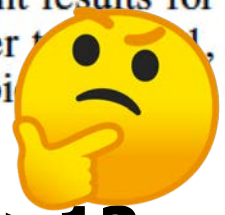


Fig. 23. Constraints on the value of the consistency parameter A_L, as a single-parameter extension to the base-ΛCDM model, using various combinations of *Planck* data. When only power spectrum data are used, $A_L > 1$ is favoured at about 3σ, but including the lensing reconstruction the result is consistent at 2σ with $A_L = 1$. The dotted lines show equivalent results for the CamSpec likelihood, which peak slightly nearer to 1, indicating some sensitivity of the A_L results to choices constructing the high-multipole likelihoods.

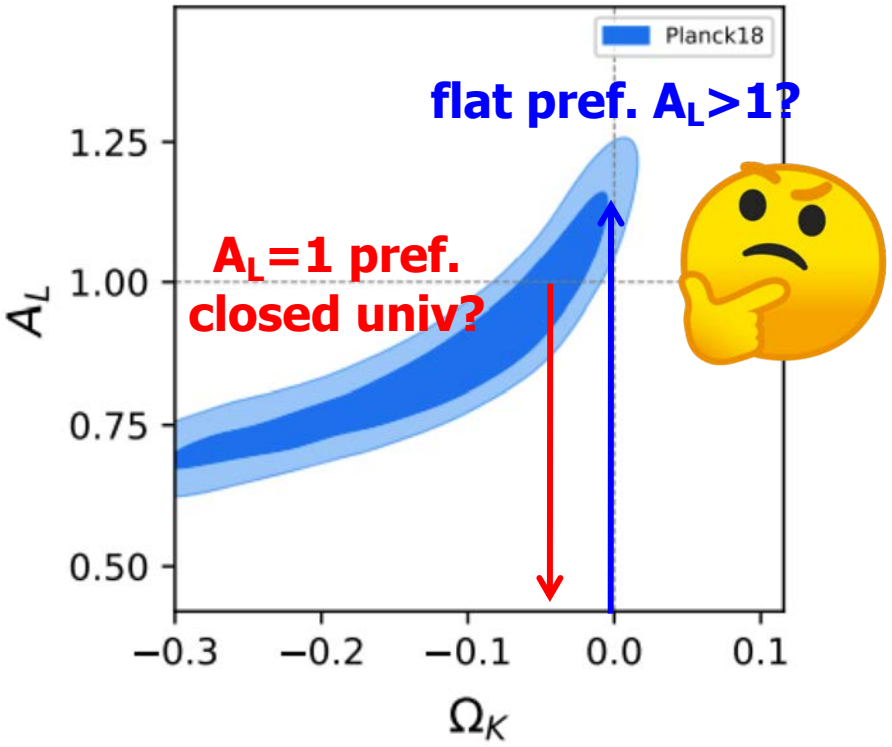


Planck might prefers A_L>1?

Planck evidence for a closed Universe and a possible crisis for cosmology

Eleonora Di Valentino, Alessandro Melchiorri ✉ & Joseph Silk

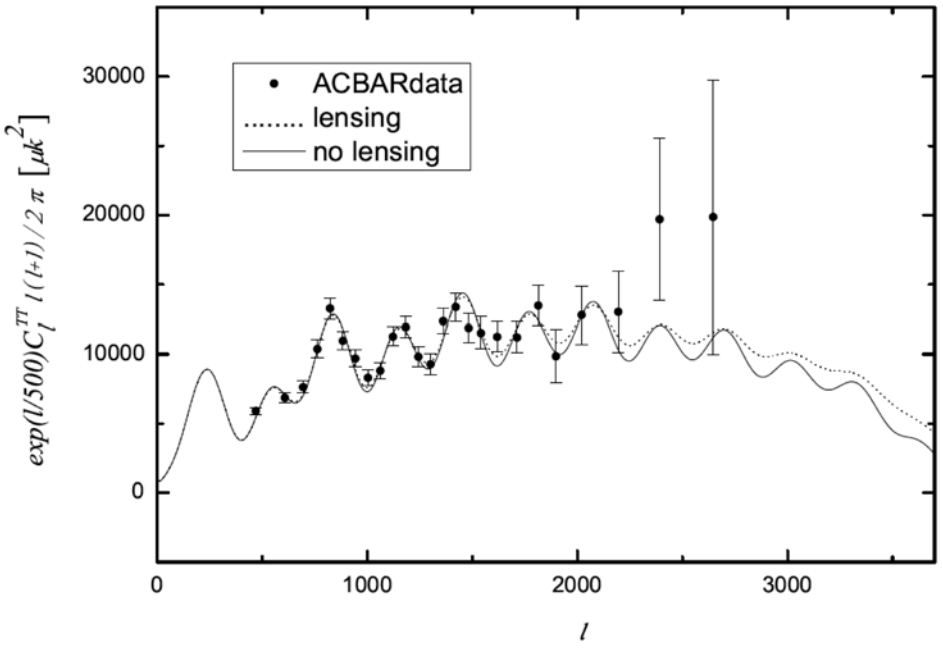
Nature Astronomy **4**, 196–203(2020) | Cite this article



Universe closed?

Curiosities on Lensing Amplitude?

c.f. historical result w/ ACBAR (2008)



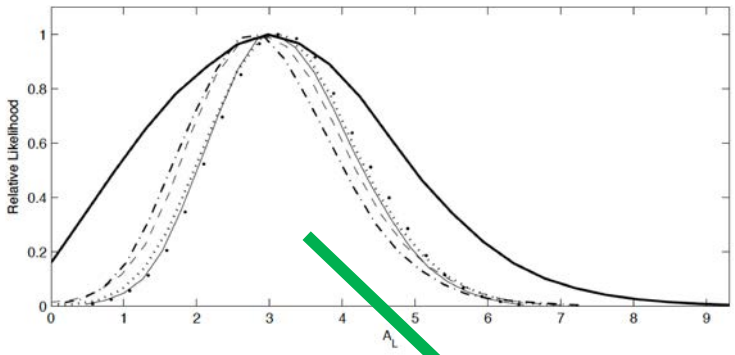


FIG. 3: Marginalized 1-D likelihood disribution for A_L for different datasets considered: WMAP3-alone (solid bold), WMAP3+ACBAR (dotted), WMAP3+"everything" (dotted bold), WMAP3+ACBAR+SZ1 (dotted-dash).

Modified Gravity prefers A_L>1?

Modified gravity

Since a very large number of models have been conceived here we use the parametrization of Daniel et al. 2008 ([41]), which is simple and easy to apply to several models. In this parameterization the gravitational slip is given by a function $\varpi(z)$ such that $\psi = (1+\varpi)\phi$ and is parameterized by a single parameter ϖ_0 defined as

varpi

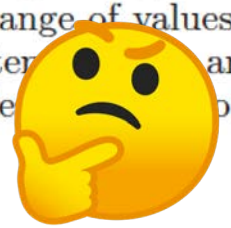
$$\varpi = \varpi_0 \frac{\Omega_\Lambda}{\Omega_m}(1+z)^{-3}. \quad (7)$$

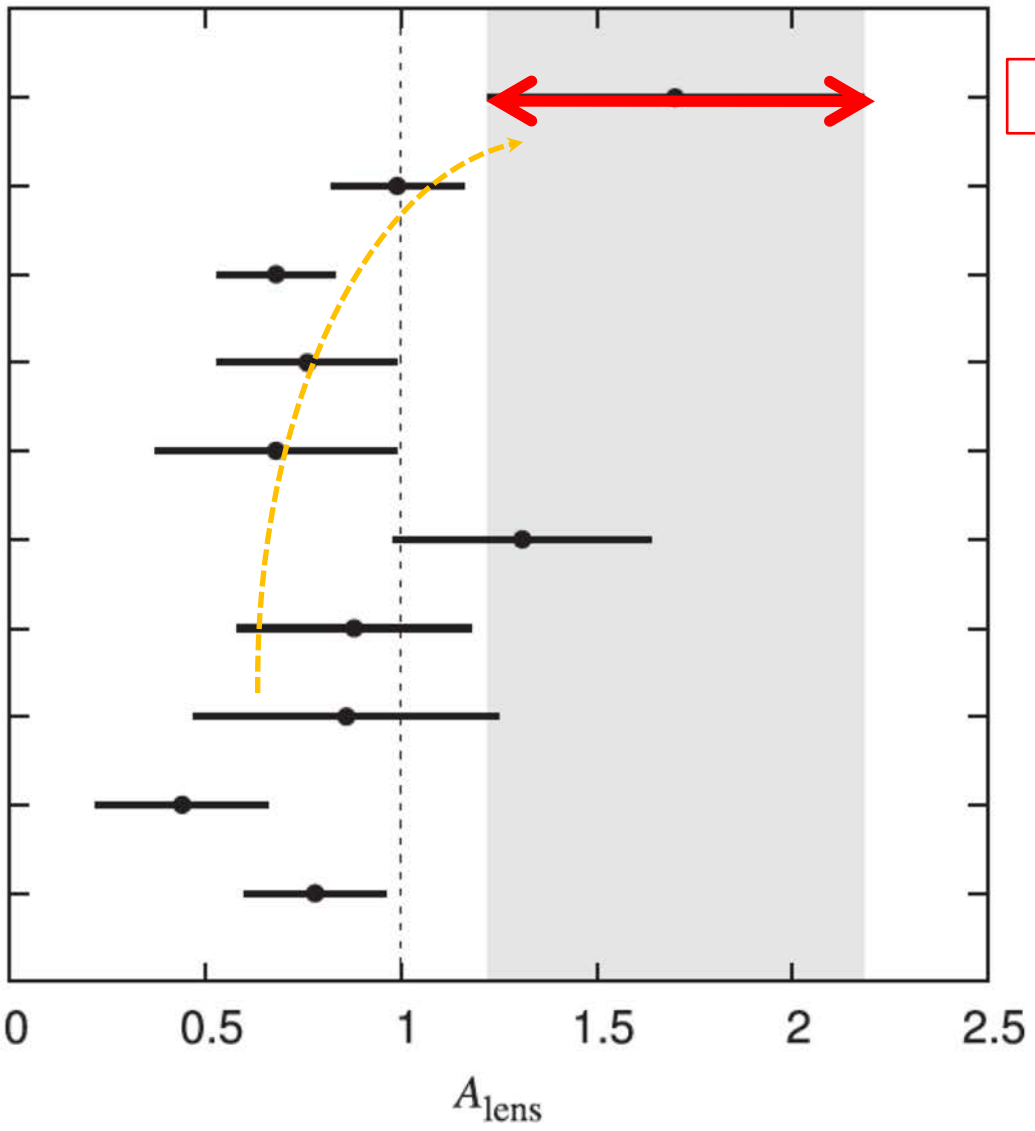
i.e. it starts to be relevant at dark energy (or modified gravity) appearance.

Following [41], we can easily approximate the relation between A_L and ϖ as

$$A_L(\varpi) = \left(\frac{G_\varpi(z=2)}{G_{\Lambda CDM}(z=2)}\right)^2 \left(\frac{2+\varpi}{2}\right)^2 \quad (8)$$

The difference in growth factors is evaluated at $z = 2$, since the lensing kernel peaks at that redshift. Larger values of ϖ_0 correspond to larger values of A_L. A value of $\varpi_0 \sim 1.5$ could produces very similar results on the CMB to $A_L \geq 1.5$ and thus bringing the signal inside the $1-\sigma$ cl. According to [41] this range of values of ϖ_0 is in agreement with the measured ter anisotropy cale odds with





2020 **pol**

POLARBEAR Polarization × HSC (this work)

SPT & Planck × DES Y1 (Omori et al. 2018)

Planck × KiDS-450 (Harnois-Déraps et al. 2017)

Planck × SDSS (Singh et al. 2017)

Planck × CFHTLenS (Harnois-Déraps et al. 2016)

Planck × RCSLenS (Harnois-Déraps et al. 2016)

SPT × DES SV (Kirk et al. 2016)

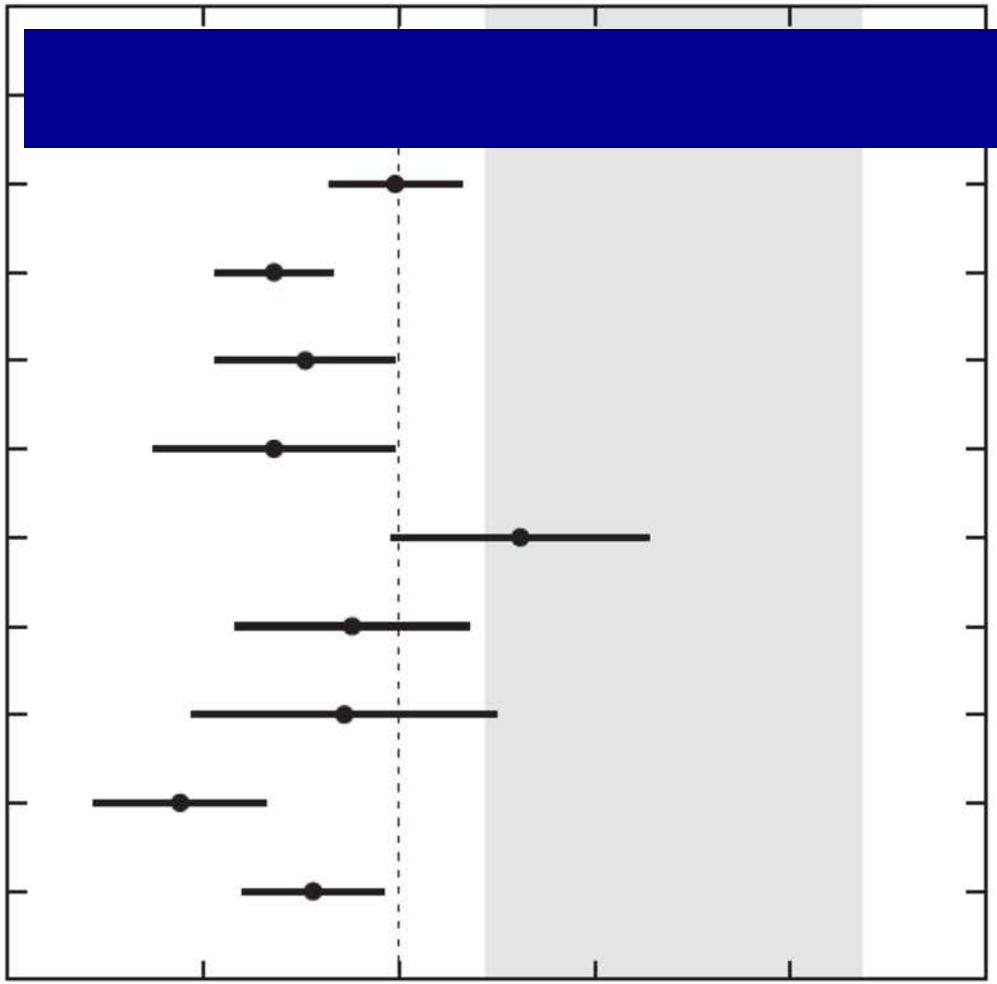
Planck × DES SV (Kirk et al. 2016)

Planck × CFHTLenS (Liu & Hill 2015)

ACT × CFHT Stripe 82 (Hand et al. 2015)

temp

- Our lensing amplitude is higher than unity → New physics!?
  - I know...
  - (but is consistent with the Planck prediction within 2σ (PTE=66%))
- Our lensing amplitude is higher than previous one → New discovery!?
  - I know...
  - (but is consistent because their amplitudes have a large variation of $A_{\rm lens}$=0.4-1.3)



pol

SPT & Planck × DES Y1 (Omori et al. 2018)
Planck × KiDS-450 (Harnois-Déraps et al. 2017)
Planck × SDSS (Singh et al. 2017)
Planck × CFHTLenS (Harnois-Déraps et al. 2016)
Planck × RCSLenS (Harnois-Déraps et al. 2016)
SPT × DES SV (Kirk et al. 2016)
Planck × DES SV (Kirk et al. 2016)
Planck × CFHTLenS (Liu & Hill 2015)
ACT × CFHT Stripe 82 (Hand et al. 2015)

temp

0

- **Any concern on SYSTEMATIC ERRORS?**
  - instrument systematics? foreground?
- **Any concern on CONFIRMATION BIAS?**
  - "Do you want or not want to report an inconsistent result w/ LCDM?"

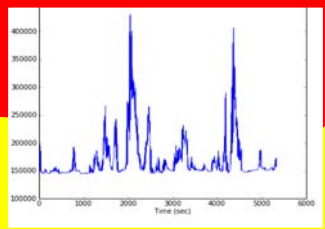
We had validated them before looking at the result w/ a "null test" and "systematic simulations"!?

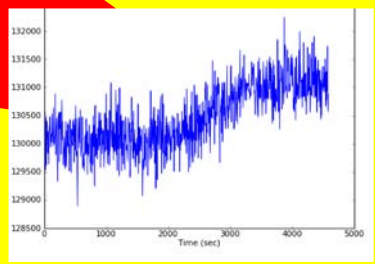
What Is “Null Test”?

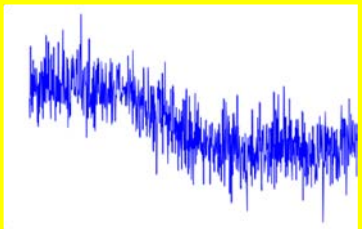
POLARBEAR

raw data ~10 TB

systematic due to bad weather





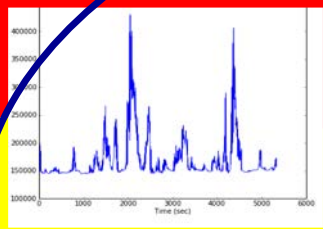


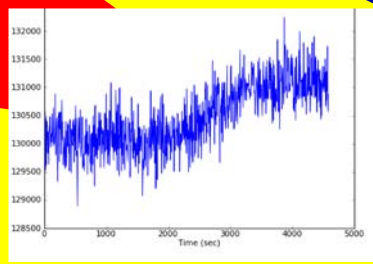
Want to reject systematic w/o looking at “result”

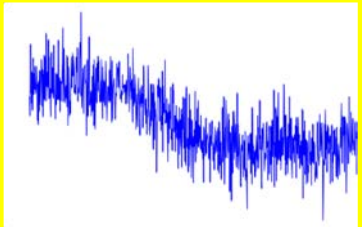
POLARBEAR

raw data ~10 TB

systematic due to bad weather





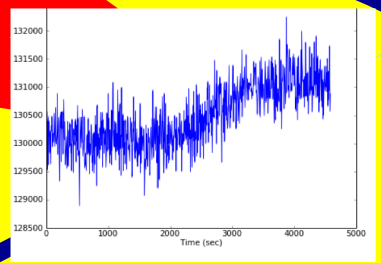


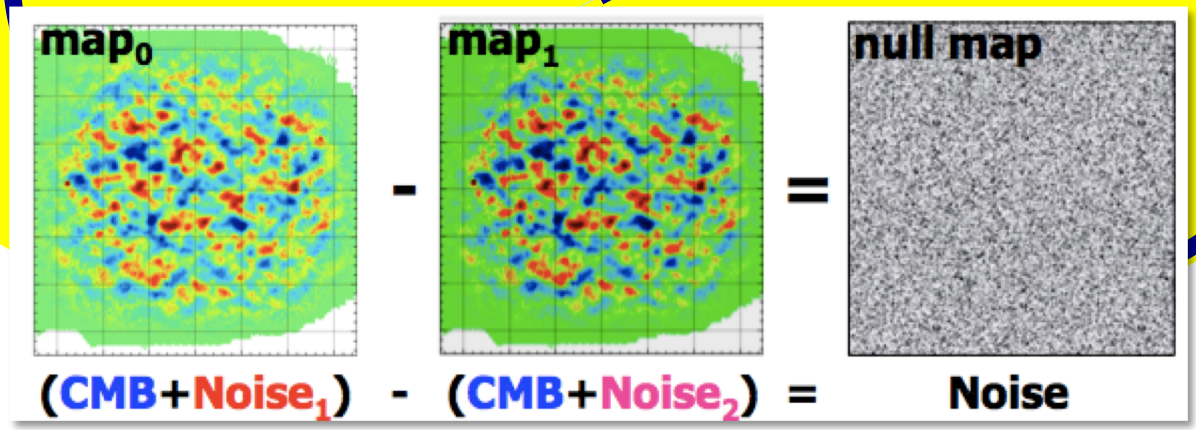
Applying data selection based on "how bad weather is?" to reject it!

POLARBEAR

raw data ~10 TB

systematic due to bad weather





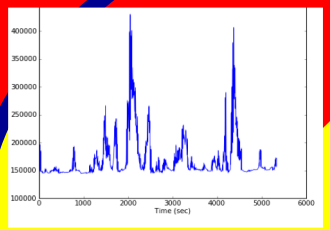
Not calculate spectrum itself from it,
but split them into two based on "how bad weather it is"
then take difference BTW two!

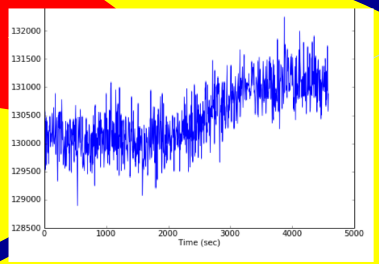
If nothing wrong is there, we see "null" signal

POLARBEAR

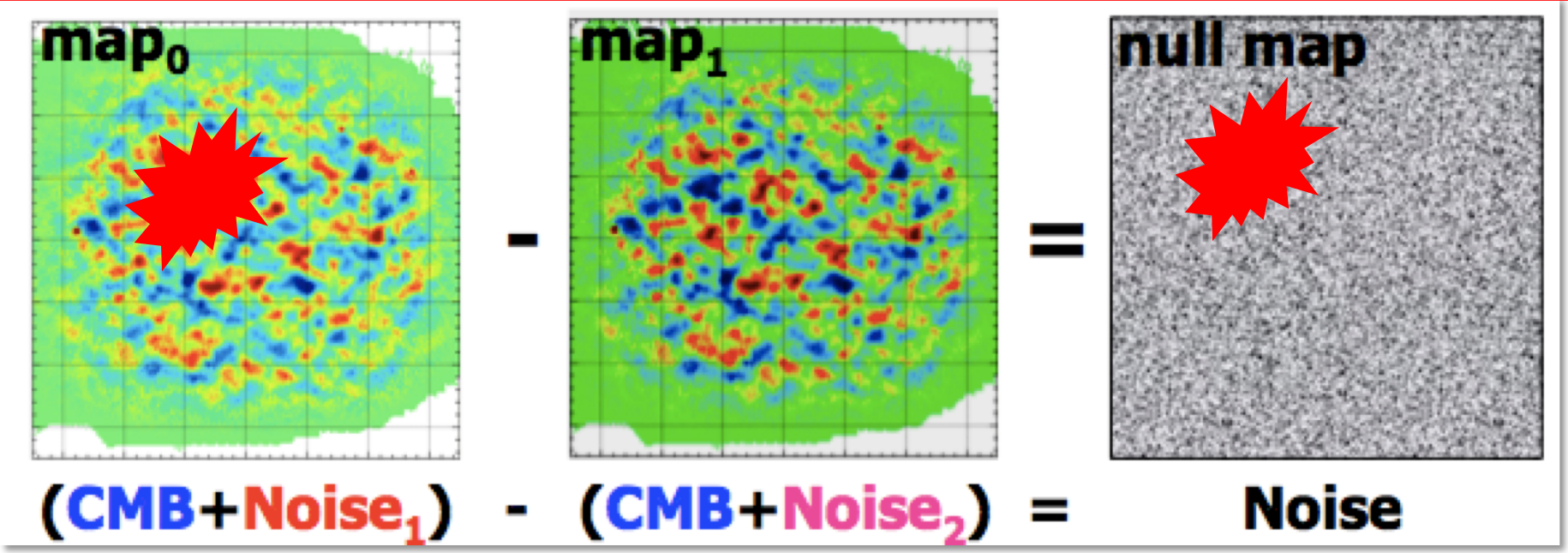
raw data ~10 TB

systematic due to bad weather



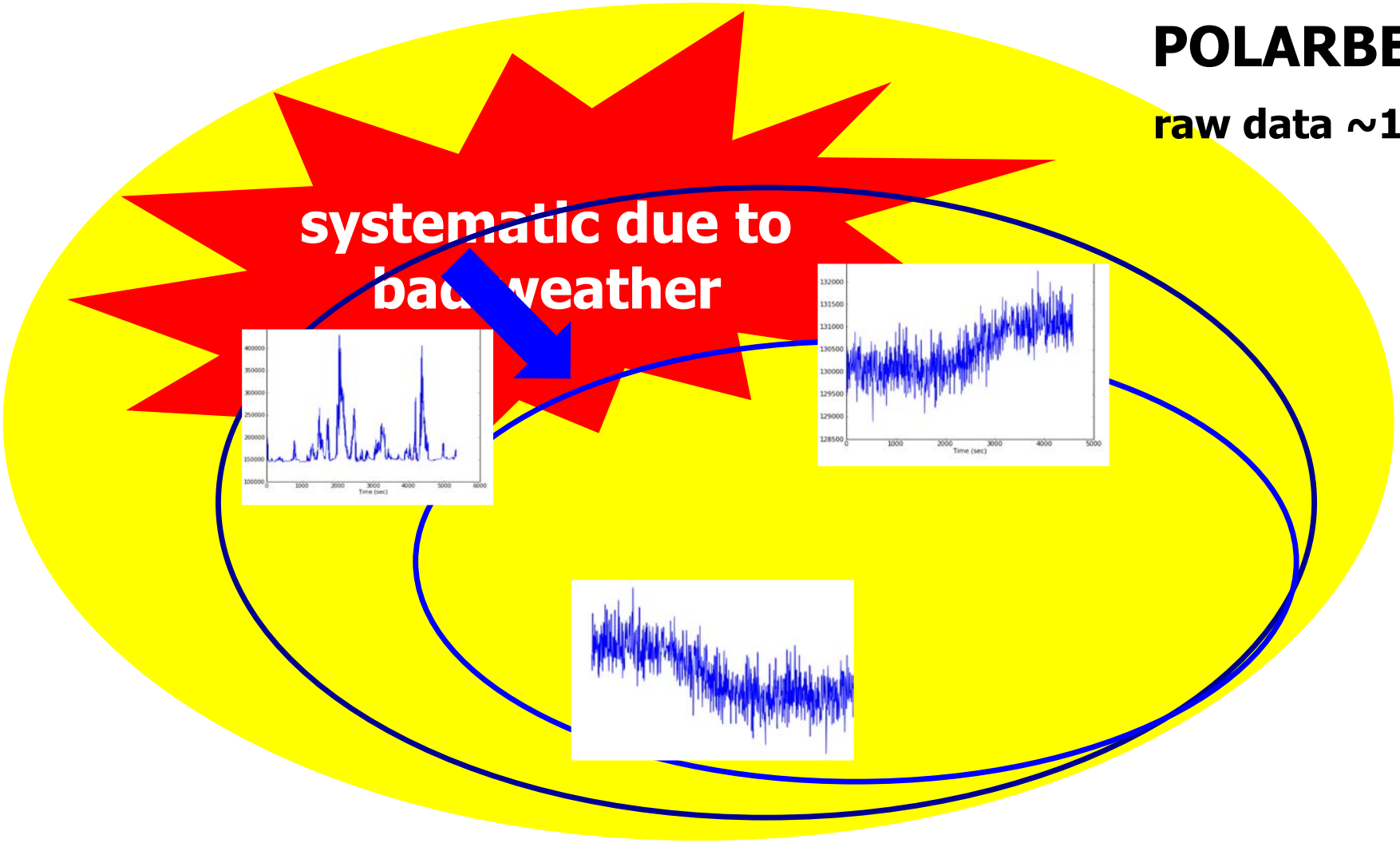


If something wrong is there, we see "non-null" signal



POLARBEAR

raw data ~10 TB



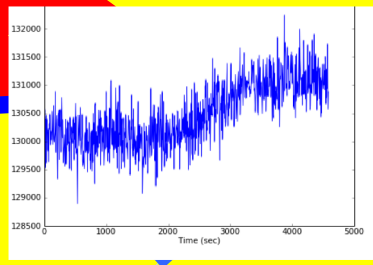
Let's tighten the selection…

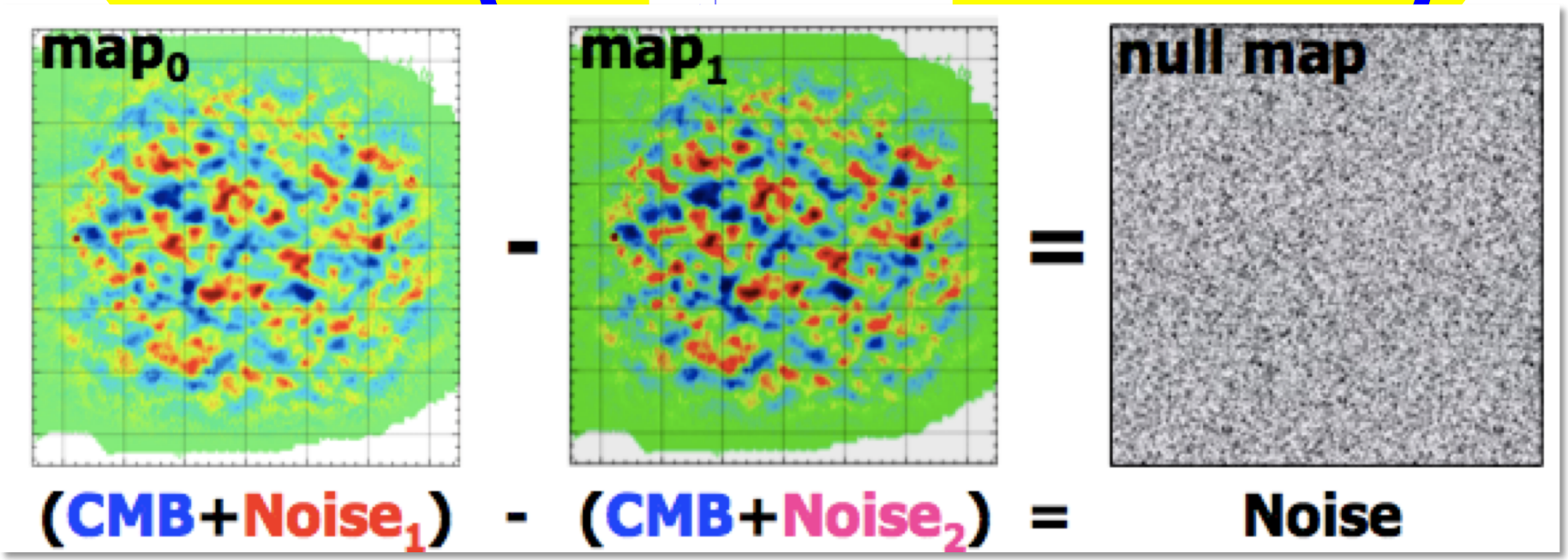
(or apply a new selection)

POLARBEAR

raw data ~10 TB

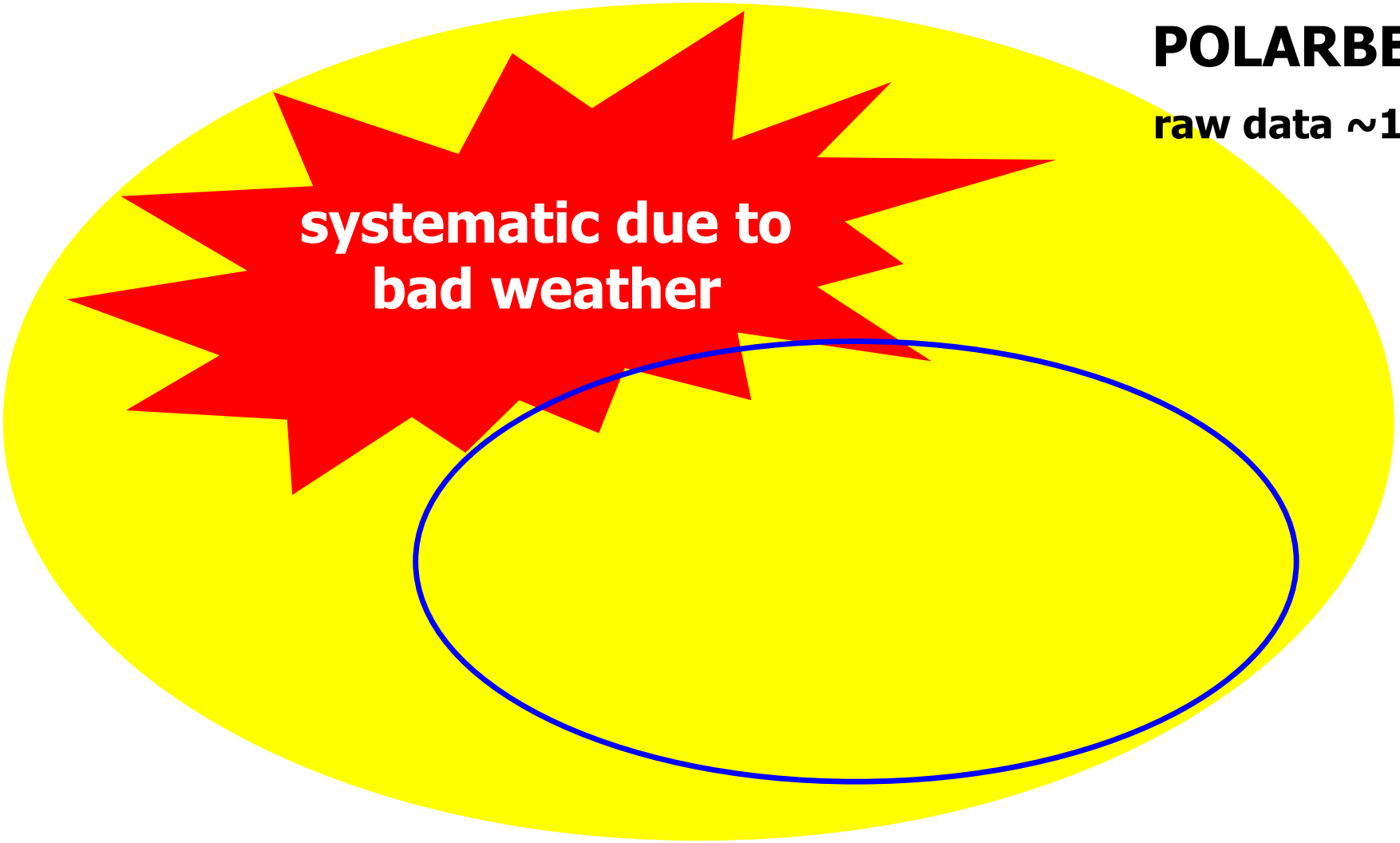
systematic due to bad weather





POLARBEAR

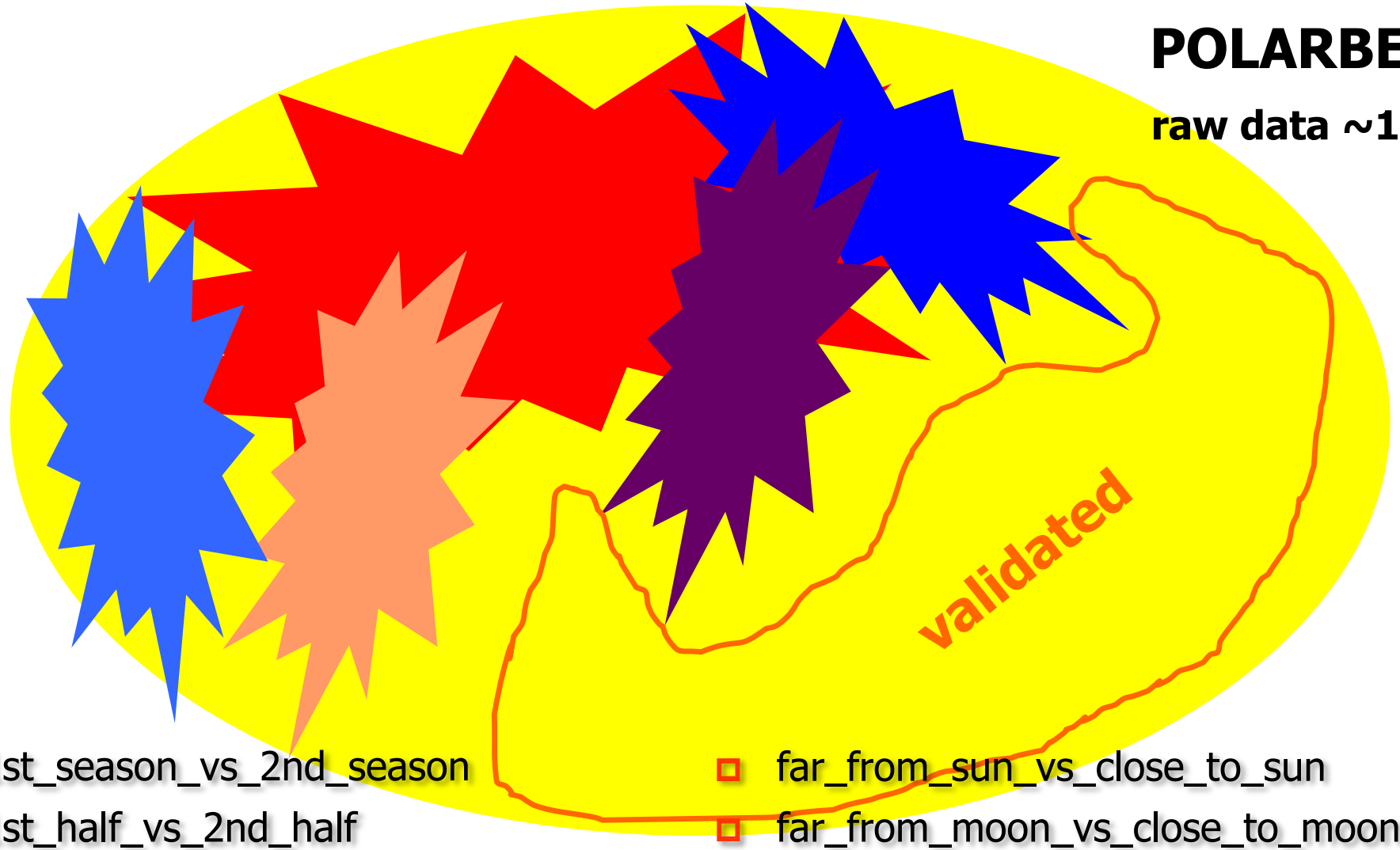
raw data ~10 TB



Successfully reject the systematic w/o looking at "result"!

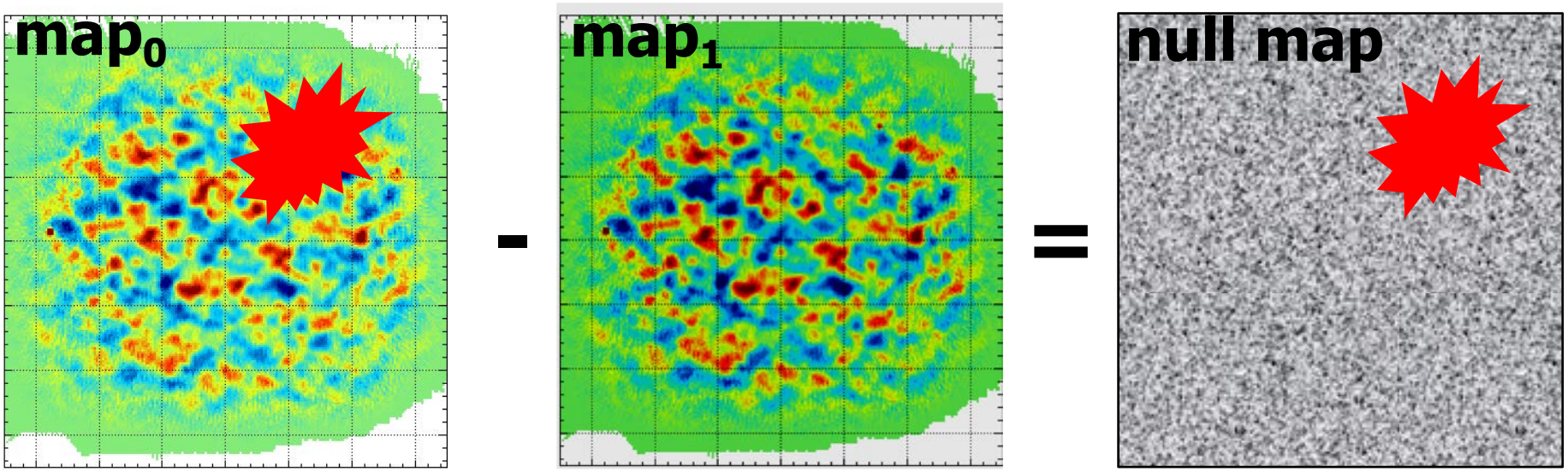
"split-data" is one of the common approaches of the blind analysis

Many Null Tests of POLARBEAR Data

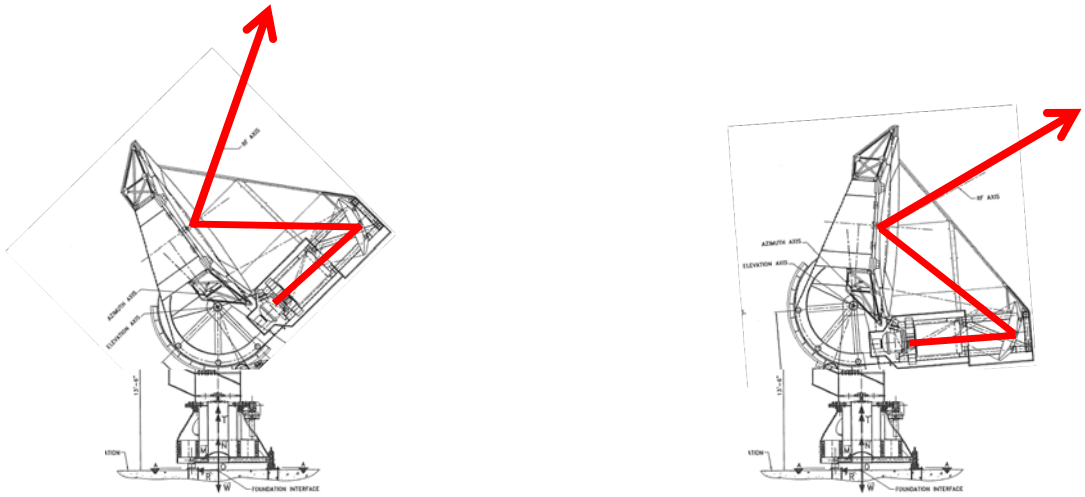


- 1st_season_vs_2nd_season
- 1st_half_vs_2nd_half
- high_gain_ces_vs_low_gain_ces
- **high_elevation_vs_low_elevation**
- rising_vs_setting
- high_pwv_vs_low_pwv
- far_from_sun_vs_close_to_sun
- far_from_moon_vs_close_to_moon
- sun_above_horizon_vs_sun_below_horizon
- left_going_scan_vs_right_going_scan
- q_pixels_vs_u_pixels
- left_side_pixels_vs_right_side_pixels

Found Problem: High Elevation vs. Low Elevation



(CMB+Noise$_1$) - (CMB+Noise$_2$) = Noise

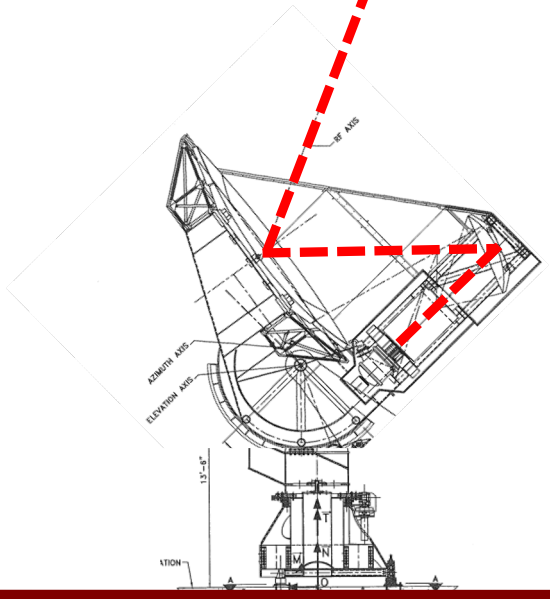


high elevation vs. low elevation

telescope

@high el

Mt. Toco
~300K



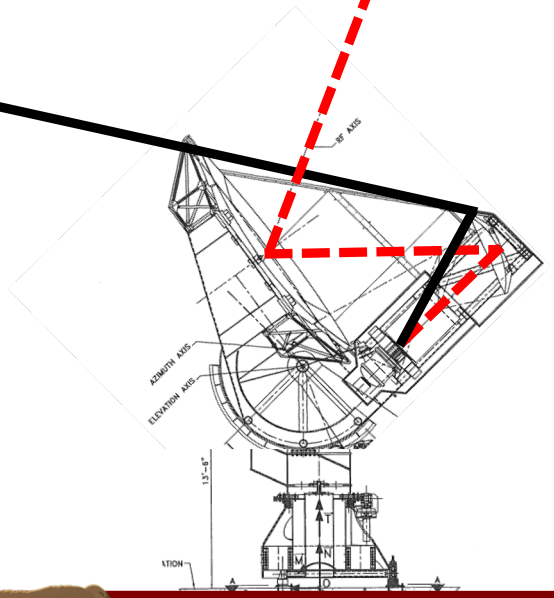
Stray Light Due to Far-sidelobe

telescope

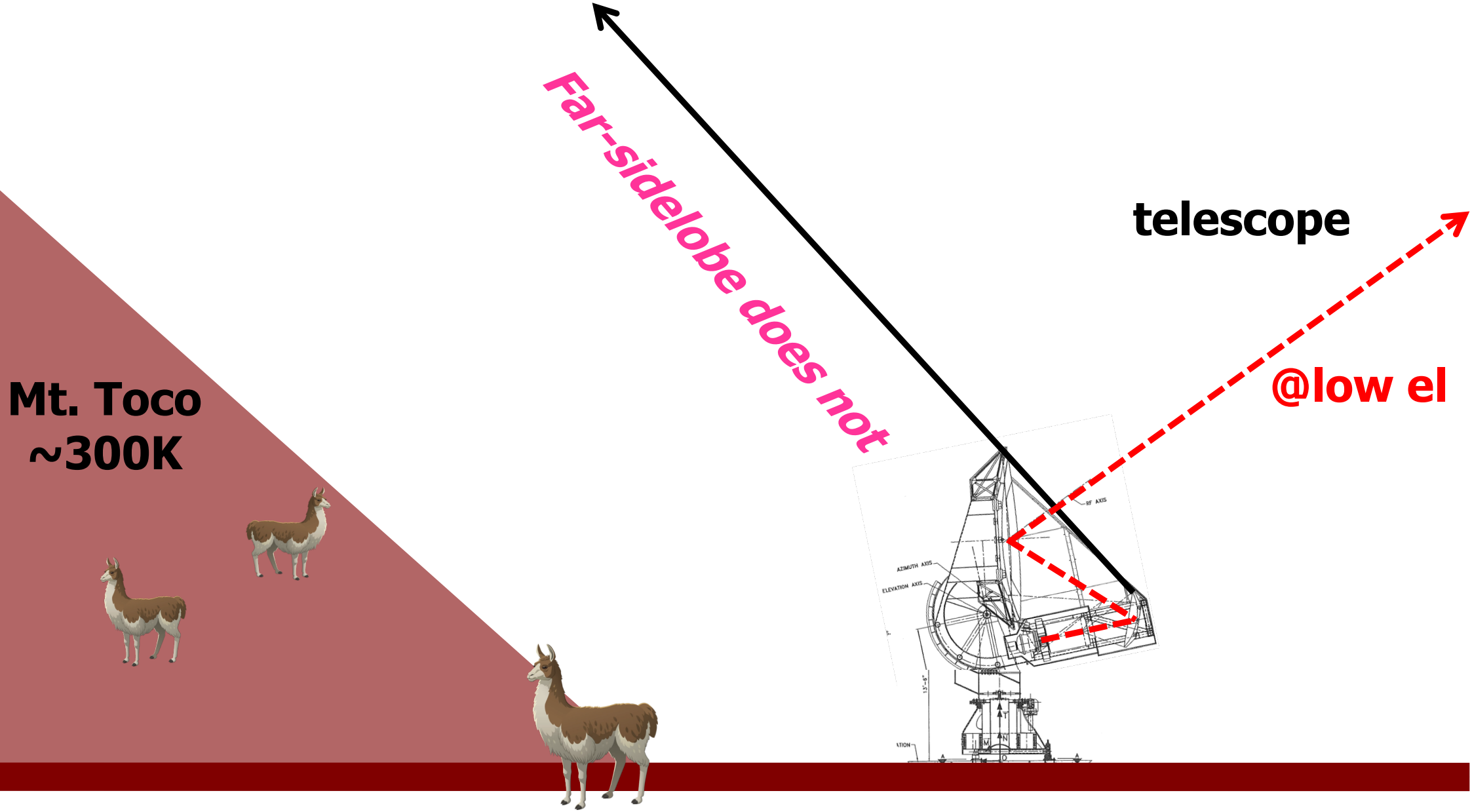
@high el

If the telescope has a back-lobe,
this could look at Mt.!?

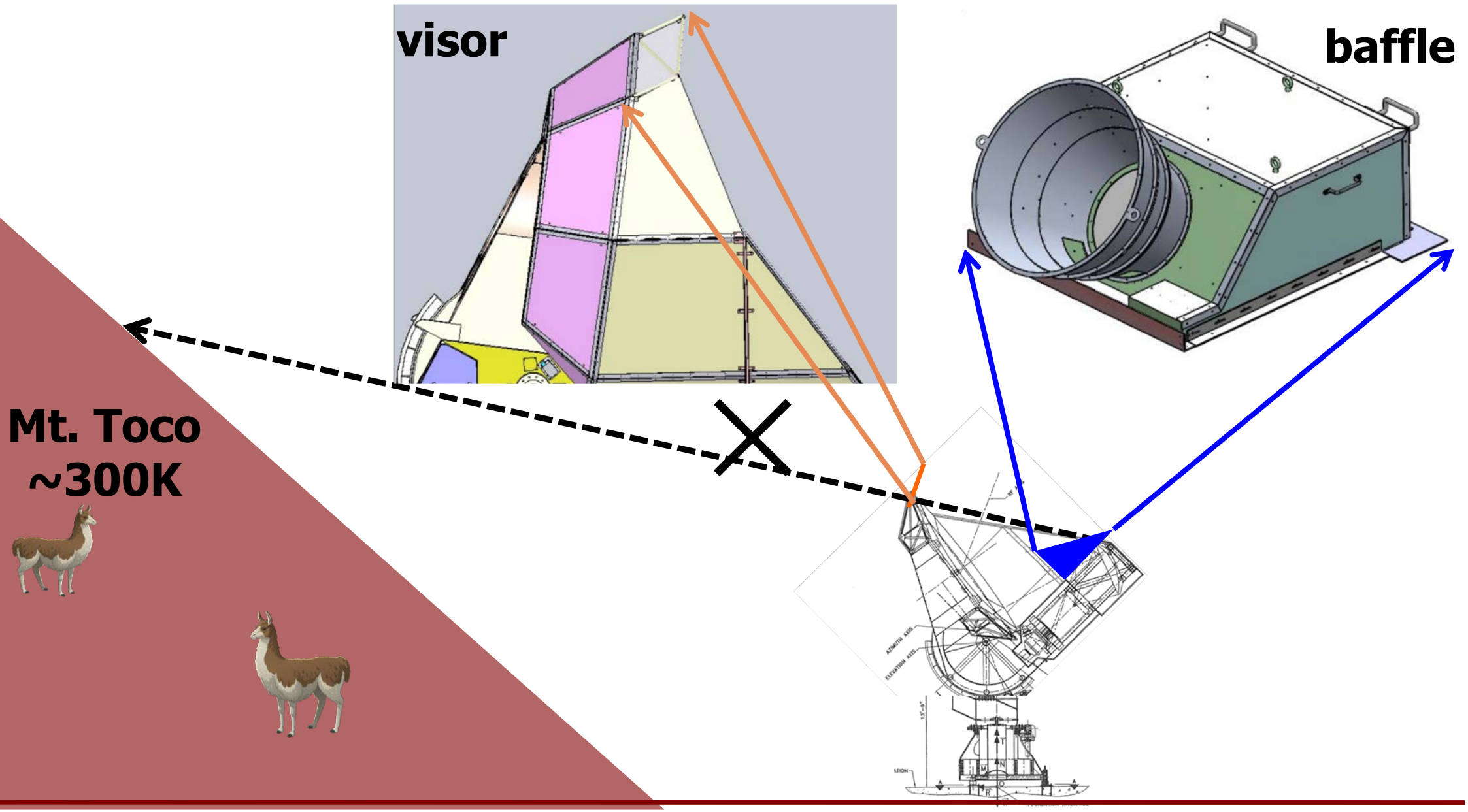
Mt. Toco
~300K



Stray Light Due to Far-sidelobe



This test could purely detect this systematic

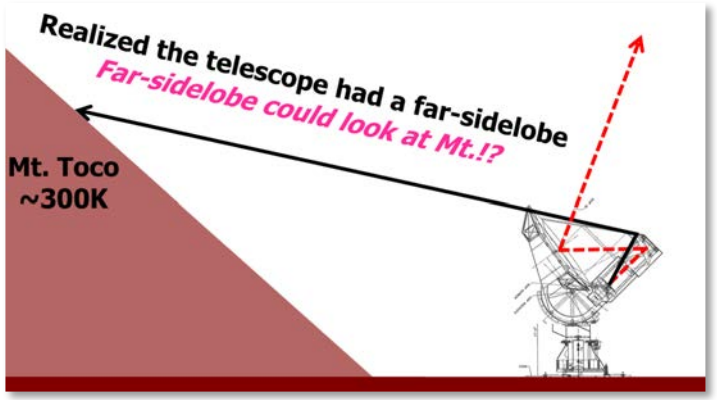


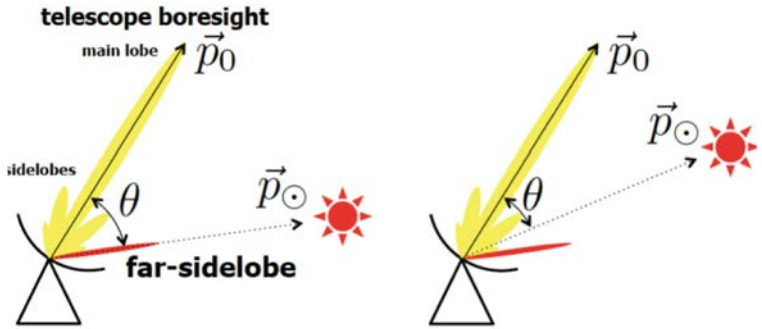
Also important because we can make an active feedback to instrument

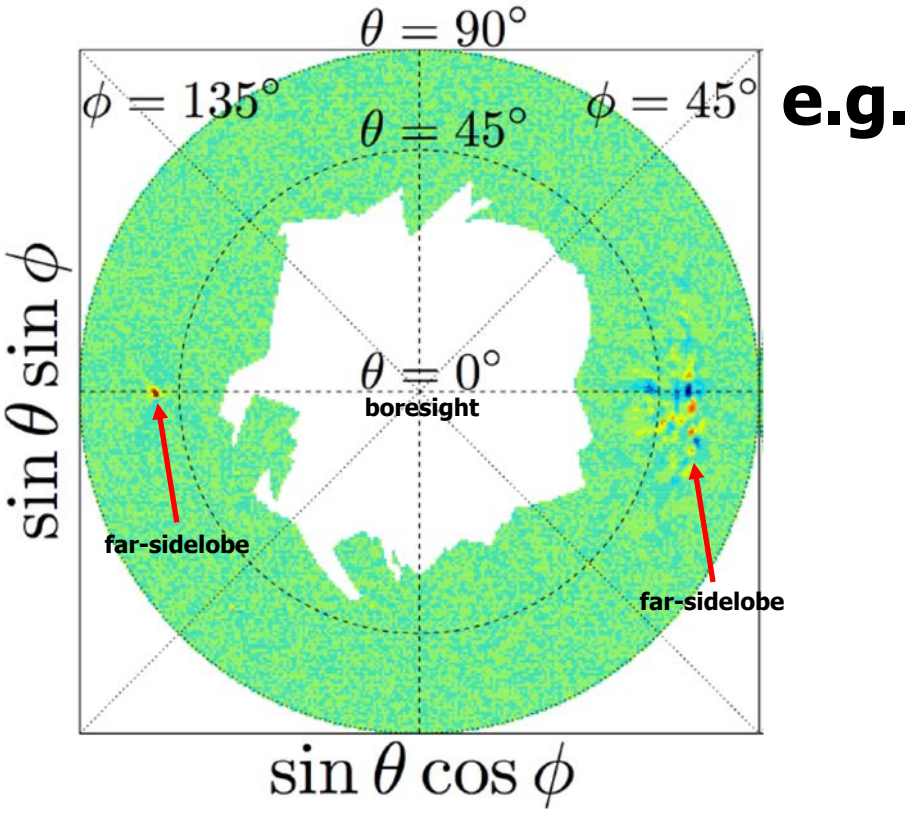
On 1st Japan-US session, Fred Matsuda showed baffle design for SO, w/ leaning a lot from this

How We Solved The Problem?

- Must mitigate it because this could bias our result
  - introduce a ground template subtraction to mitigate it
  - must validate it carefully
- First measured this back-lobe by scanning the Sun
  - making a map relative to Sun (so-called "Sun-centered" map)







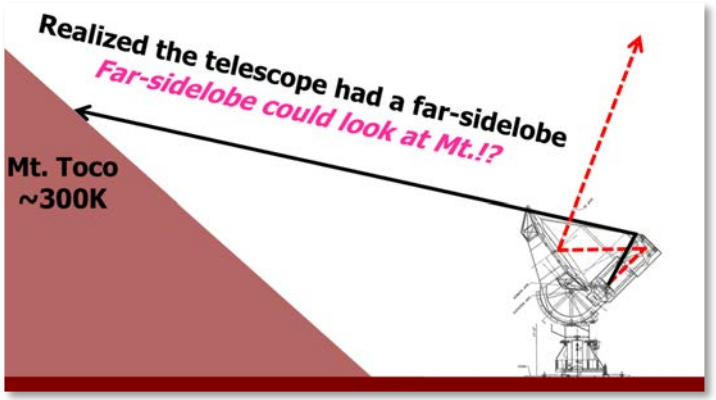
e.g.

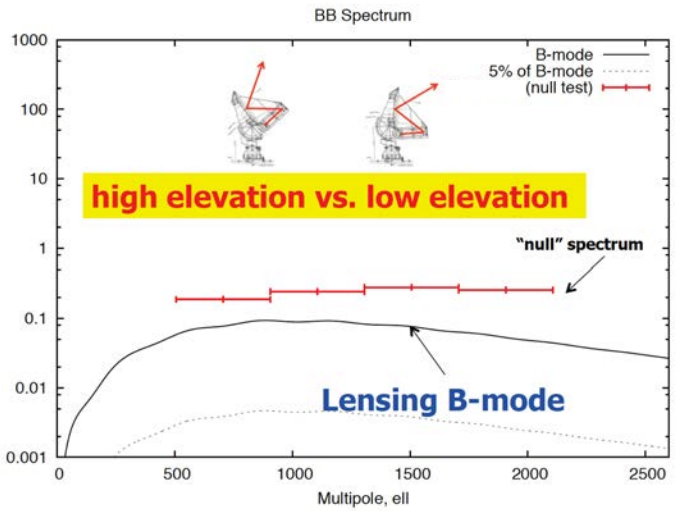
Sun-centered map(SCM)=far-sidelobe map (2π str.)
(Sorry, this is not PB's SCM but QUIET)

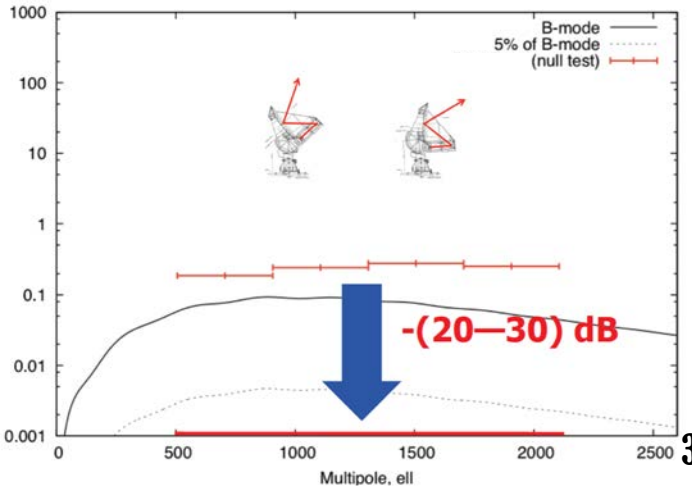
How We Solved The Problem?

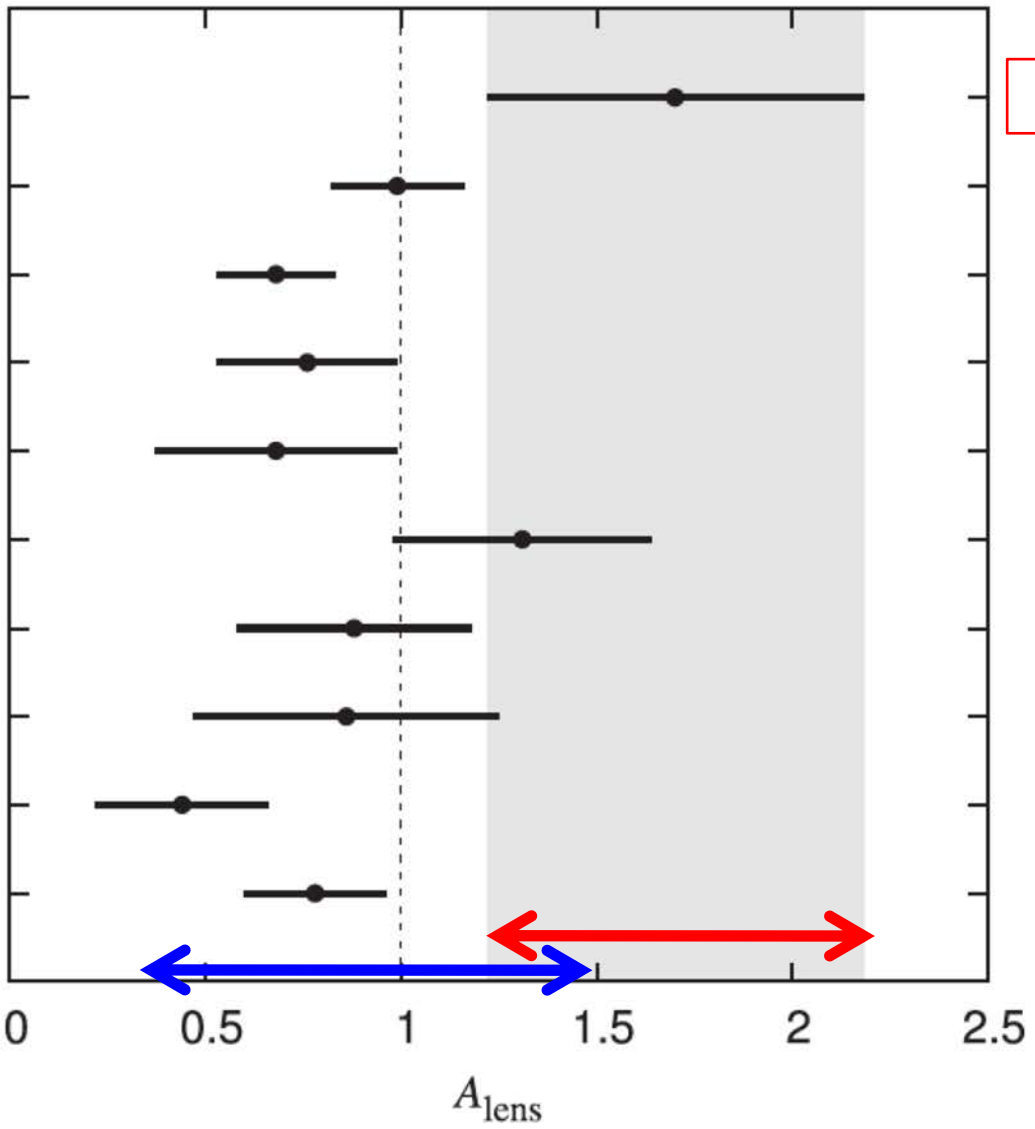
- Must mitigate it because this could bias our result
  - introduce a ground template subtraction to mitigate it
  - must validate it carefully
- First measured this back-lobe by scanning the Sun
  - making a map relative to Sun (so-called "Sun-centered" map)
- Simulated the full CMB obs. to reproduce the observed contamination by back-lobe's scanning ground map
  - Note we, of course, have NEVER looked at the contamination in "non-null" spectrum; we have looked at it in "null" spectrum
    - ✓ We didn't know how our central value should be during this analysis
- Successfully reproduced it in simulated "null" spectrum
- Demonstrated the ground template subtraction can mitigate it in simulated "non-null" and "null" spectrum as well as demonstrated it in "null" spectrum of data
- Concluded we validated & mitigated the contamination
  - NOT looking at the contamination in "non-null" spectrum of data!

A null test w/ corresponding systematics simulation is powerful to validate data/analysis w/o looking at result









pol

POLARBEAR Polarization × HSC (this work)

SPT & Planck × DES Y1 (Omori et al. 2018)

Planck × KiDS-450 (Harnois-Déraps et al. 2017)

Planck × SDSS (Singh et al. 2017)

Planck × CFHTLenS (Harnois-Déraps et al. 2016)

Planck × RCSLenS (Harnois-Déraps et al. 2016)

SPT × DES SV (Kirk et al. 2016)

Planck × DES SV (Kirk et al. 2016)

Planck × CFHTLenS (Liu & Hill 2015)

ACT × CFHT Stripe 82 (Hand et al. 2015)

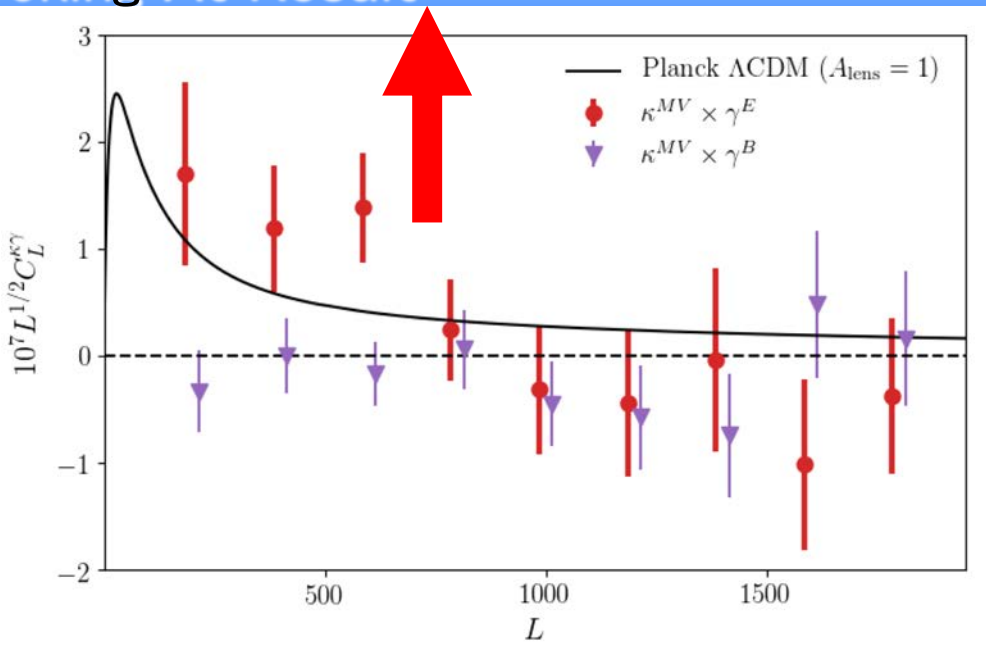
temp

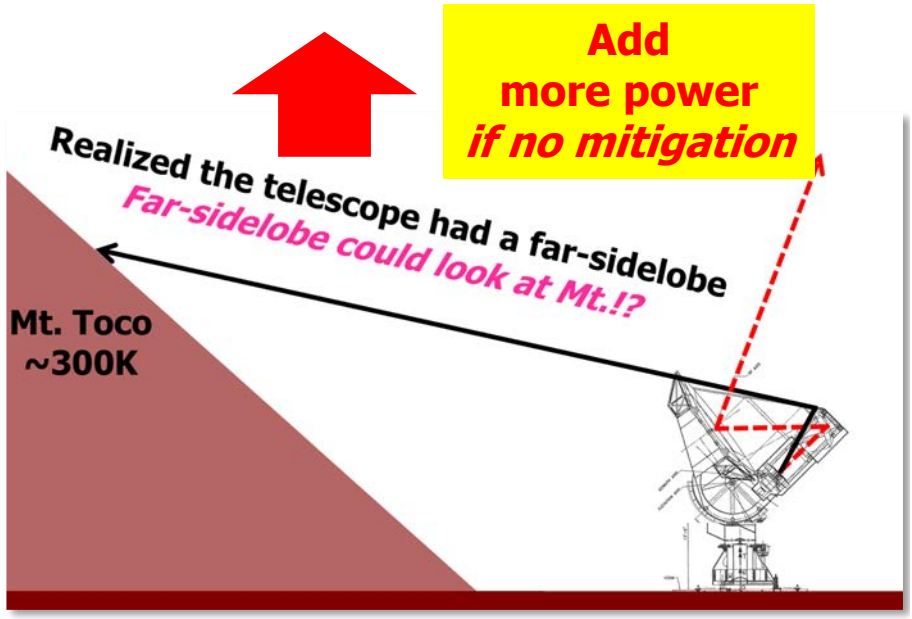
- **No systematic error/bias is observed;** Validated it before looking at result
- Our lensing amplitude **is consistent with the Planck within 2σ**
- It agrees with **the other cross-correlation analyses**

→No evidence to claim any new physics

Situation We Might Fall Into
If We Can't Find This Problem Before Looking At Result

- Assumed our true amplitude is 2σ away from LCDM (but we don't know it)
- If we don't know the problem & no mitigation is done, this could bias our amplitude of A_L
  - **If this bias is >1σ, our A_L becomes >3σ away from LCDM!?**
- If we see it before solving the problem
  - someone would search for the source
  - someone might be eager to confirm it's true because they might want a discovery!?
  - someone could try to just debias it **by filtering/DS**
- This might introduce the confirmation bias
  - we expect we can be free from that (because we're clever enough!?)
  - however **we can be free from that if we do not look at the result anyway**



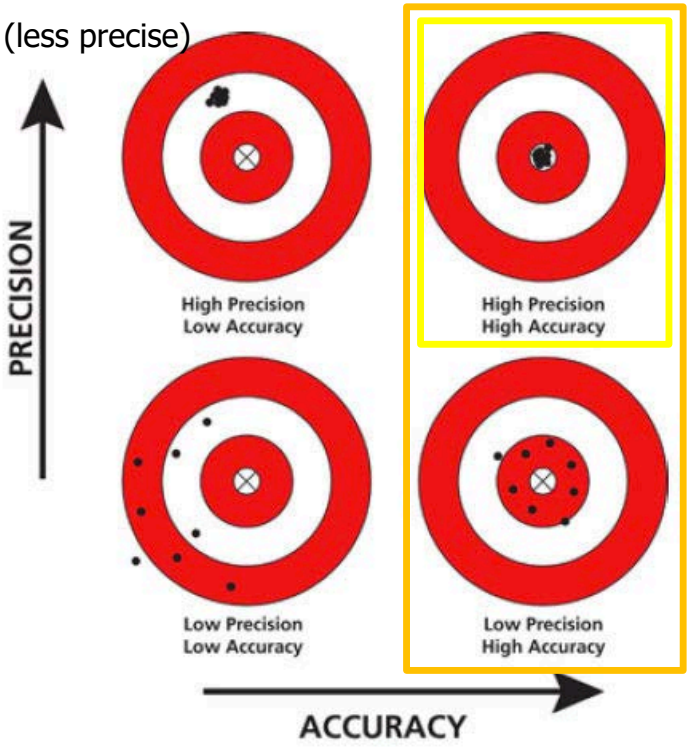


Important to try to validate data/analysis as best as possible before looking at result

if we don't know this

Feasibility Of Null Tests & Accuracy / Precision

- Computational cost by running various null tests iteratively to investigate data/analysis could be **expensive** (even a single data analysis procedure is not expensive because they need many MCs)
- Optimizing pipeline design & choice of analysis algorithm must be important
  - e.g. MPI/OpenMP on cluster/supercomputer; map-maker (**filter-bin**, MLE/MAP)
- If null tests are too expensive to run iteratively, we might miss chance to find unanticipated systematics, resulting in a more-systematics-concerned result (=less accurate)
  - The most optimal estimator is optimal to get the smallest stat. error bars (=more precise)
    - ✓ But it's sometimes not optimal to run again & again
  - Not perfectly optimal estimator may give larger stats. error bars (less precise)
    - ✓ But it's easy to run again & again to investigate unanticipated sys. errors
- We want a highly accurate & highly precise result, but we need the balance!
- ***Since CMB experiments are entering the systematic-dominated regime, we should try to achieve highly accurate result***



One-Minute Summary

- POLARBEAR performed **a blind analysis** for all the CMB science results
- **A critical framework is a "null test"** (& systematic simulations)
- POLARBEAR developed the "**null test framework**" to **improve** & **validate** the **data/analysis** *iteratively before looking at final spectra*
  - Must be useful to discover and investigate **unanticipated systematics** & it's critical to avoid **any confirmation bias**
- For POLARBEAR, it's NOT just a **validation test**, BUT it's a **driving force** of the **data/analysis;** ***Null test & systematic simulations are the two wheels of a cart***
- **Running many & various null tests** must be **useful**, but could be **expensive**; Need a great attention to its feasibility and accuracy/precision
- Since CMB experiments are entering the **systematic-dominated regime**, we should try to achieve **highly accurate result**