CMB-S4 Preliminary Baseline Design Validation - Measurement to Science: Galaxy Clusters

Sources working group

Srinivasan **Raghunathan**

10 March 2021

Work done with: *Marcelo* **Alvarez**, *Nick* **Battaglia**, *Gil* **Holder**, *Elena* **Pierpaoli** and *Nathan* **Whitehorn**.

Also previously: **J.** *Colin* **Hill** and *Mat* **Madhavacheril**.

CMB-S4

Overview / baseline setup

- **Overview:**
  - Predict the number of expected clusters for CMB-S4 (S4) experiment.
  - Use high-z clusters to be detected by S4 to understand the properties of intracluster medium (+ constrain cosmology).

- **Two surveys are considered:**
  - S4-Wide (f_{sky} = 65 %) - To be located in Chile.
  - S4-Ultra deep (f_{sky} = 3 %) - Delensing LAT to be located at the South Pole.

- **Instrument specs:**
  - **S4-Wide:** https://cmb-s4.org/wiki/index.php/Expected_Survey_Performance_for_Science_Forecasting#Instrument_Definition
  - **S4-Ultra deep (V3R025):** https://cmb-s4.uchicago.edu/wiki/index.php/Delensing_sensitivity_-_updated_sensitivities,_beams,_TT_noise

- **Inputs:**
  - Multi-band (30, 40, 95, 145, 225, 278) GHz simulated maps of S4-wide and S4-Ultra deep surveys.
  - Combine simulated maps using internal linear combination.
  - **Cluster:** Generalised NFW profile (Arnaud et al. 2009, arXiv: 0910.1234; Nagai et al. 2007, arXiv: 0703661).
  - **Simulations:** CMB, galactic and extragalactic foregrounds, and noise (described in upcoming slides).

Overview / baseline setup

- **Signal-to-noise threshold:** 5σ.
- **Observable:**

CMB-S4 shall detect (at 5σ) all galaxy clusters with an integrated Compton $Y_{SZ} \geq XX$ at $z \geq 1.5$ over the large area survey footprint (f_{sky} = 65%). Furthermore, it shall detect (at 5 sigma) all galaxy clusters with an integrated Compton $Y_{SZ} \geq YY$ at $z \geq 1.5$ over the de-lensing survey footprint (f_{sky} = 3%).

Thermal Sunyaev-Zeldovich (SZ) effect

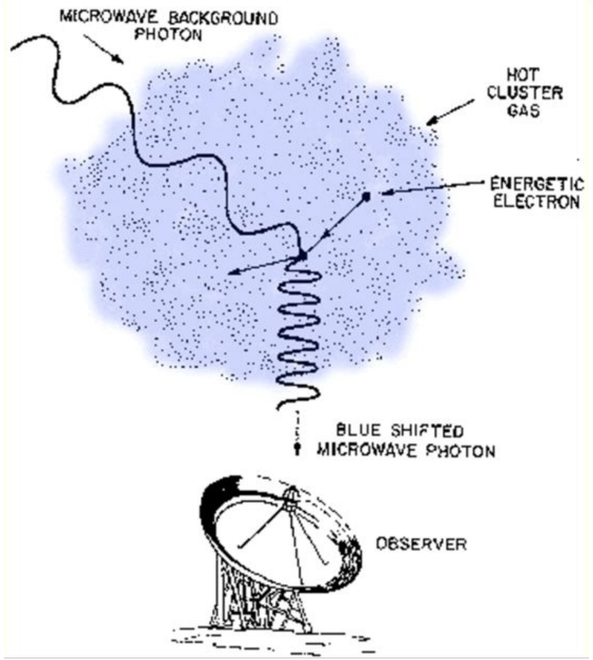
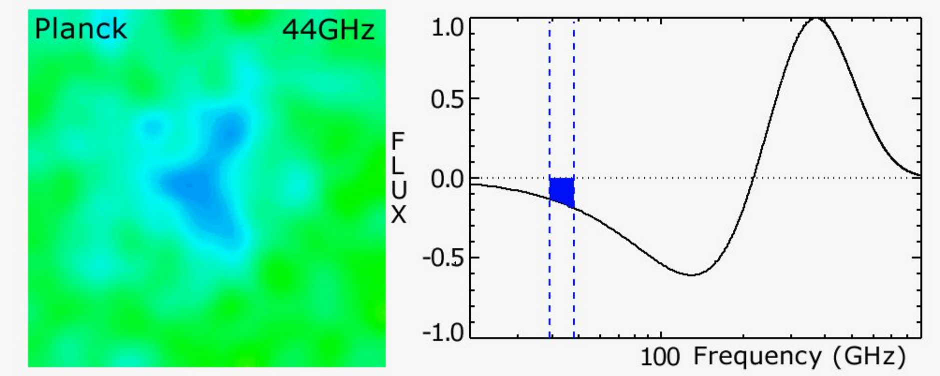


Image: L. Van Speybroeck

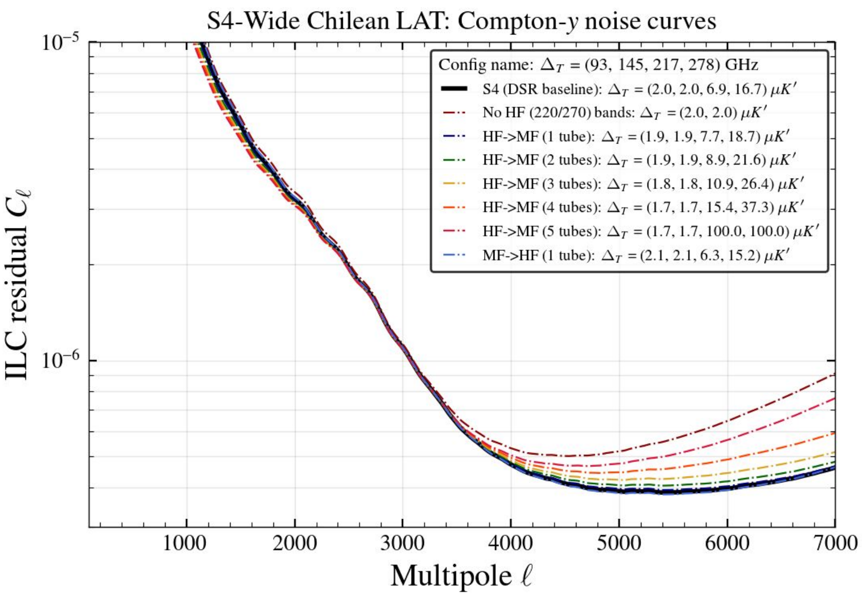
- Galaxy clusters contain hot gas (free electrons).
- CMB photons, that pass through clusters of galaxies, are inverse Compton scattered by free electrons in the intracluster medium (ICM).
- Used for blind detections of clusters in CMB surveys.
- ***SZ effect is redshift independent and hence allows us detect distant clusters.***



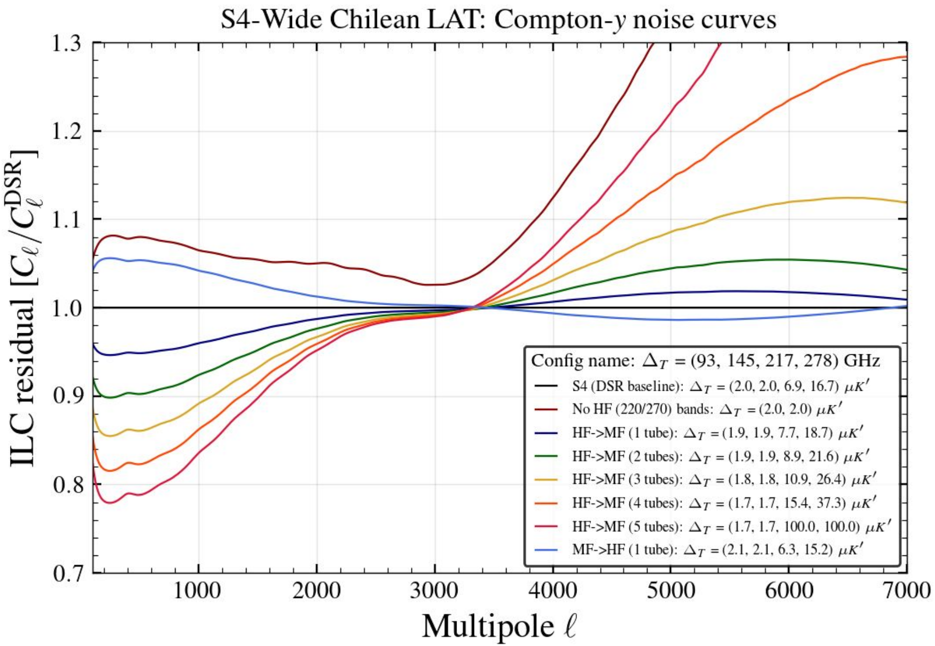
Frequency dependence of thermal SZ (Image: ESA)

S4-Wide: Band allocation for cluster survey

- Black curve is for S4-Wide DSR configuration and it is the most optimal configuration that we can have given the number of detectors. *Colin Hill arrived at similar results initially and independently.*
- ***CIB sets a floor and reducing the noise levels of MF at the expense of HF detectors does not help. The increase is noise level is not significant when we move 1 or 2 tubes.***



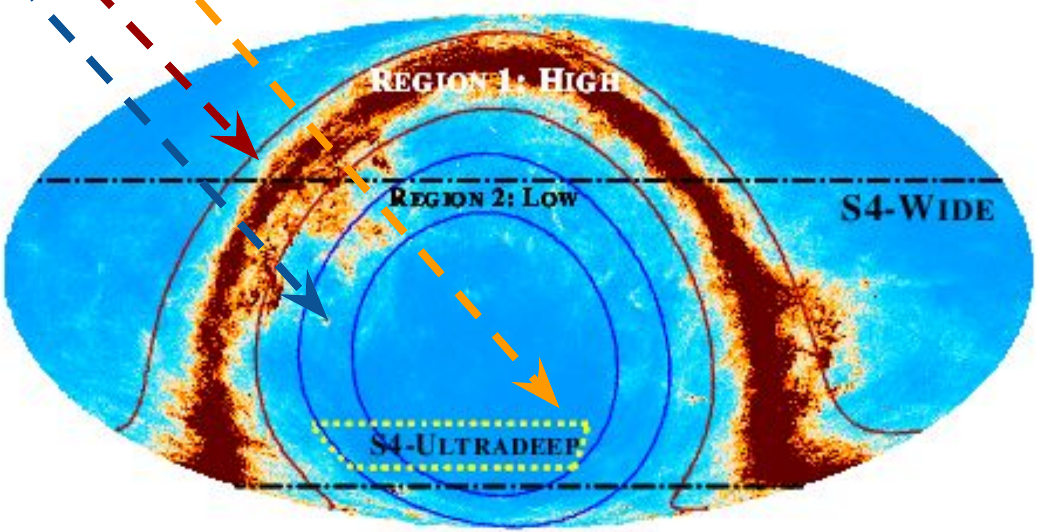
S4-Wide absolute Compton-y noise levels



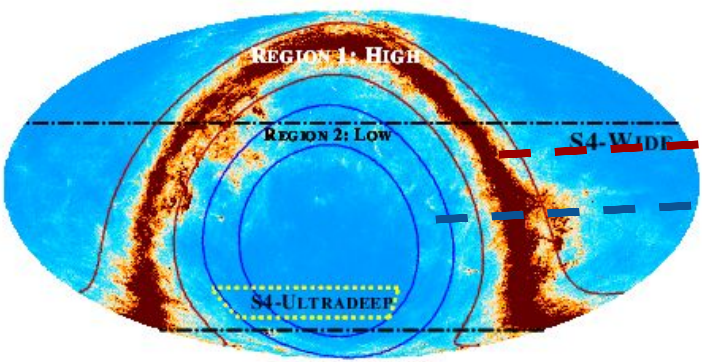
S4-Wide Compton-y noise compared to DSR configuration

End-to-end testing: Galactic foregrounds

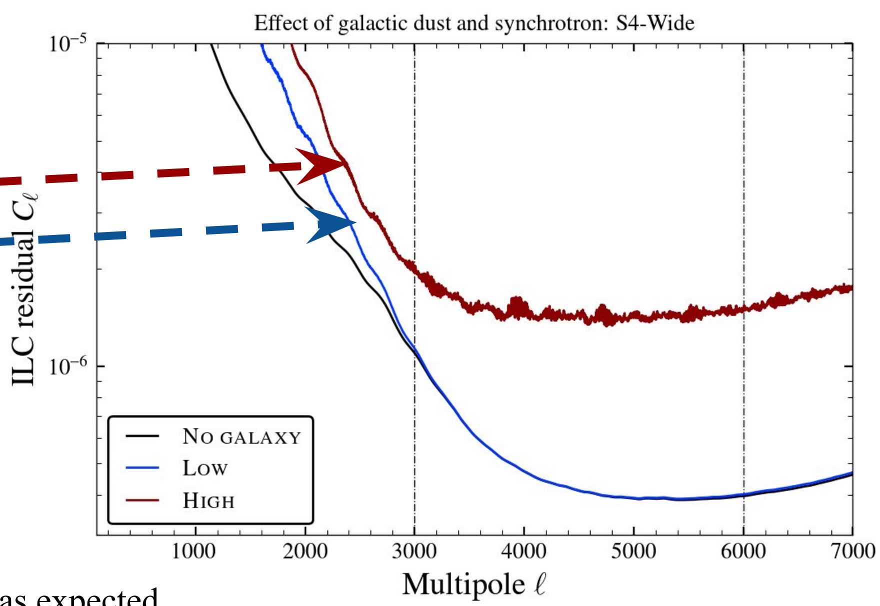
- **Galactic foregrounds:** Position dependent foregrounds added using pySM3 simulations from the data management (DM) group.
  - **S4-Wide:** Decomposed footprint into regions with low and high galactic emission.
    - High: $f_{sky} = 0.15$
    - Low: $f_{sky} = 0.5$
  - **S4-Ultra deep:** Assumed to have zero galactic emission.
  - *More about pySM3 in DM session.*



End-to-end testing: Galactic foregrounds



Galactic dust (pySM3) at 145 GHz



- Adding galactic emission increases the large-scale noise, as expected.
- At cluster-scales (highlighted band), the noise curves for "Region 2: Low" is not different from baseline case (with no galactic emission).
- When we are looking right through the plane of our galaxy, the noise increases by >x2 on small scales. This will degrade the SNR of clusters.
- **We decompose the S4-Wide (f_{sky} = 0.65) footprint into two assuming:**
  - ***f_{sky} = 0.15 will have galactic emission similar to "High".***
  - ***f_{sky} = 0.5 will have galactic emission similar to "Low".***

We also tried dividing the footprint into smaller sub-fields and did not any significant difference in results.

End-to-end testing: Extragalactic foregrounds

- **Extragalactic foregrounds:**
  - **kSZ:** Tried two different approaches: Uncorrelated or correlated to cluster. In both cases, we do not except any bias. kSZ only adds extra variance.
    - Uncorrelated: kSZ signals are Gaussian realisations of an underlying spectrum measured by the South Pole Telescope (SPT: Reichardt et al. 2020, arXiv: 2002.06197).
    - Correlated: kSZ signals are injected using Websky simulations (Stein et al. 2020, arXiv: 2001.08787).

End-to-end testing: Extragalactic foregrounds

- **Extragalactic foregrounds:**
  - **kSZ:** Tried two different approaches: Uncorrelated or correlated to cluster. In both cases, we do not except any bias. kSZ only adds extra variance.
    - Uncorrelated: kSZ signals are Gaussian realisations of an underlying spectrum measured by the South Pole Telescope (SPT: Reichardt et al. 2020, arXiv: 2002.06197).
    - Correlated: kSZ signals are injected using Websky simulations (Stein et al. 2020, arXiv: 2001.08787).

  - **CIB:** Tried three different approaches: Uncorrelated or correlated to cluster.
    - Uncorrelated: CIB signals are Gaussian realisations of an underlying spectrum measured by SPT (George et al. 2015, arXiv: 1408.3161 and Reichardt et al. 2020, arXiv: 2002.06197). No bias expected.
      - CIB frequency decorrelations obtained from SPT x SPIRE (Viero et al. 2018, arXiv: 1810.10643).
    - Correlated: *Cluster correlated CIB can bias the tSZ signal reconstruction. They are injected using:*
      - MDPL2 simulations (Omori *in preparation*)
      - Websky simulations (Stein et al. 2020, arXiv: 2001.08787). *Websky simulations are also part of S4 DM working group.*
      - Spectra are scaled to match measurements by SPT.
      - ***Bias due to tSZ X CIB is <~0.3σ.***

CMB-S4

End-to-end testing: Extragalactic foregrounds

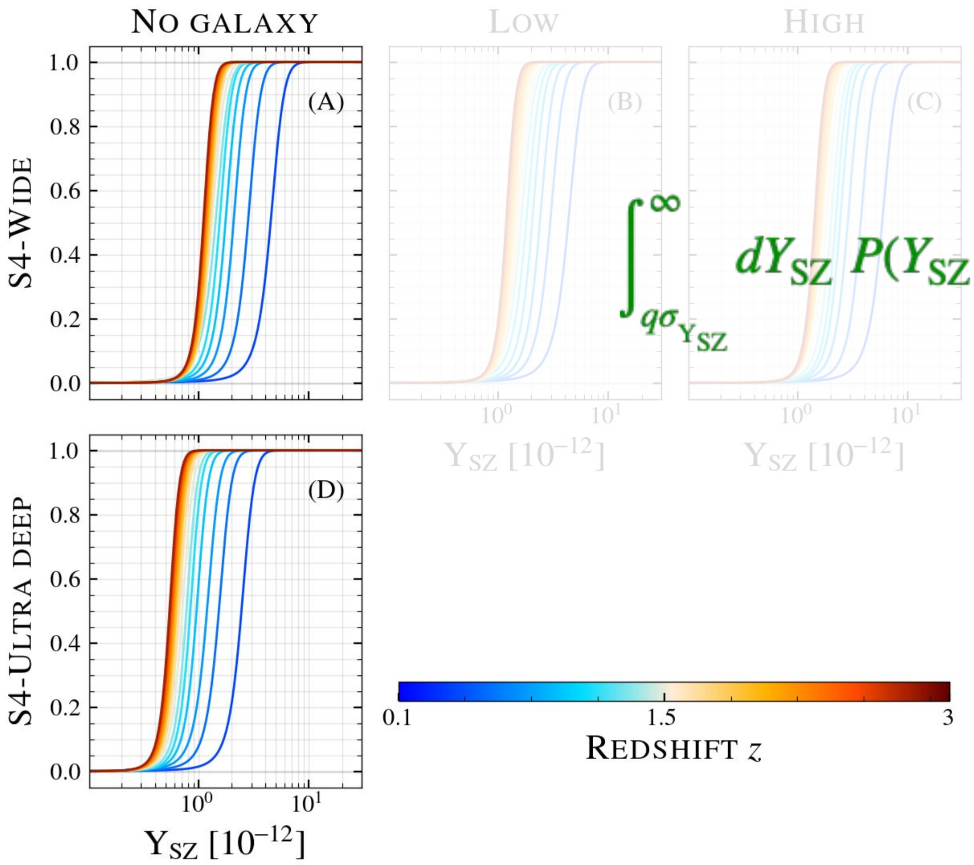
- **Extragalactic foregrounds:**
  - **kSZ:** Tried two different approaches: Uncorrelated or correlated to cluster. In both cases, we do not except any bias. kSZ only adds extra variance.
    - Uncorrelated: kSZ signals are Gaussian realisations of an underlying spectrum measured by the South Pole Telescope (SPT: Reichardt et al. 2020, arXiv: 2002.06197).
    - Correlated: kSZ signals are injected using Websky simulations (Stein et al. 2020, arXiv: 2001.08787).

  - **CIB:** Tried three different approaches: Uncorrelated or correlated to cluster.
    - Uncorrelated: CIB signals are Gaussian realisations of an underlying spectrum measured by SPT (George et al. 2015, arXiv: 1408.3161 and Reichardt et al. 2020, arXiv: 2002.06197). No bias expected.
      - CIB frequency decorrelations obtained from SPT x SPIRE (Viero et al. 2018, arXiv: 1810.10643).
    - Correlated: *Cluster correlated CIB can bias the tSZ signal reconstruction. They are injected using:*
      - MDPL2 simulations (Omori *in preparation*)
      - Websky simulations (Stein et al. 2020, arXiv: 2001.08787). *Websky simulations are also part of S4 DM working group.*
      - Spectra are scaled to match measurements by SPT.
      - ***Bias due to tSZ X CIB is <~0.3σ.***
  - **Radio galaxies:**
    - Uncorrelated to cluster: Radio galaxy signals are Gaussian realisations of an underlying spectrum corresponding to sources below masking threshold. These will only add extra variance. *(Also check Zack Li's talk in the parallel session.)*
    - Adding radio point sources >=0.5 mJy inside clusters leads to ~1σ bias. But this needs better modelling.

S4 cluster forecasts: End-to-end testing

- **Galactic foregrounds:**
  - Currently using pySM3 simulations (DM group).
- **Extragalactic foregrounds:**
  - CIB: Currently using Wesbky simulations (DM group) but also tried MDPL2 / Gaussian realisations.
  - Radio: Listen to Zack Li's talk in the parallel session.
- **CMB:**
  - Gaussian realisations of *Planck* 2015 `TT,TE,EE+lowP+lensing+ext` cosmology (*Planck* collaboration 2016 XIII, arXiv: 1502.01589).
- **Noise and instrument systematics:**
  - Isotropic - Instrumental + Band-band correlated atmospheric noise created using Gaussian realisations.
  - **Realistic noise simulations + systematics from DM group: *Pending.***

S4 cluster survey completeness



$$\int_{q\sigma_{Y_{SZ}}}^{\infty} dY_{SZ}\, P(Y_{SZ} \,|\, Y_{SZ}^{\rm true}) = 0.5\left(1 + {\rm erf}\left[\frac{Y_{SZ}^{\rm true} - q\sigma_{Y_{SZ}}}{\sqrt{2}\sigma_{Y_{SZ}}}\right]\right)$$

See also:

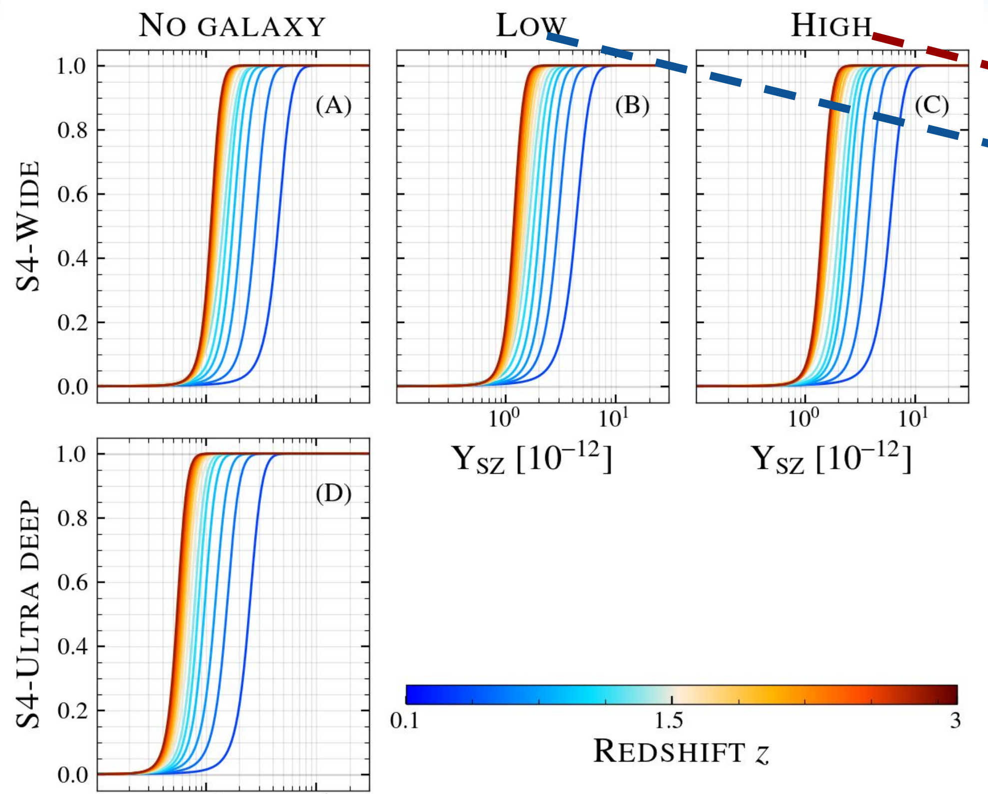
Planck **collaboration 2014 XX, arXiv: 1303.5080**

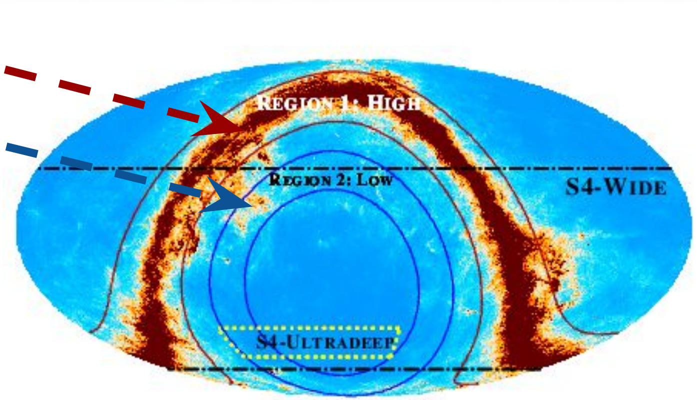
Planck **collaboration 2016 XXIV, arXiv: 1502.01597**

Alonso, Louis, Bull et al. 2016, arXiv: 1604.01382

CMB-S4

S4 cluster survey completeness



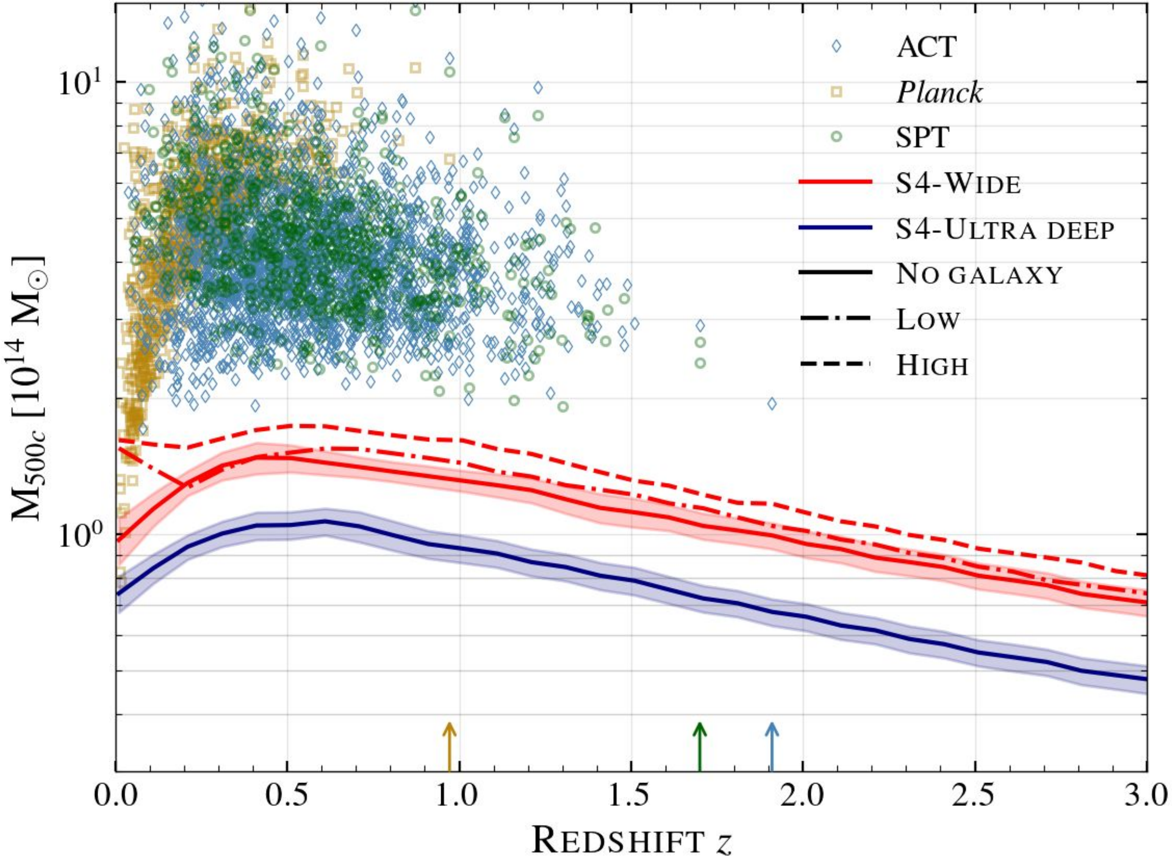


Flowdown text

CMB-S4 shall detect (at 5σ) all galaxy clusters with an integrated Compton $Y_{SZ} \geq 10^{-12}$ ***at*** $z \geq 1.5$ ***over the large area survey footprint*** **(*f_{sky}* = ~~65%~~ 50%)**. *Furthermore, it shall detect (at 5σ) all galaxy clusters with an integrated Compton* $Y_{SZ} \geq 5x10^{-13}$ ***at*** $z \geq 1.5$ ***over the de-lensing survey footprint (***f_{sky} *= 3%).*

CMB-S4

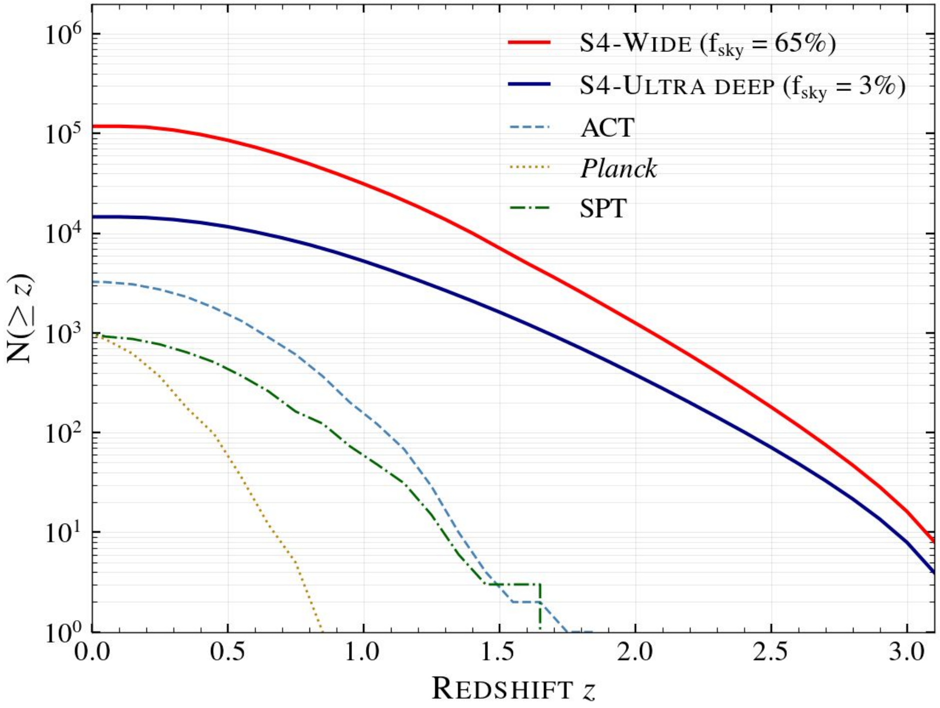
S4 cluster survey: Limiting mass vs redshift



CMB-S4

S4 cluster survey: Total objects

- **S4-Wide:** Contains clusters from low (f_{sky} = 0.5) + high (f_{sky} = 0.15) galactic emission regions. Removing high galactic emission region reduces ~18% objects.
- **High-z (z>=2) clusters:** S4-Wide → ~1000 clusters; S4-Ultra deep → ~200 clusters.



Mass calibration using CMB lensing

CMB lensing mass calibration: **Average mass of the sample**

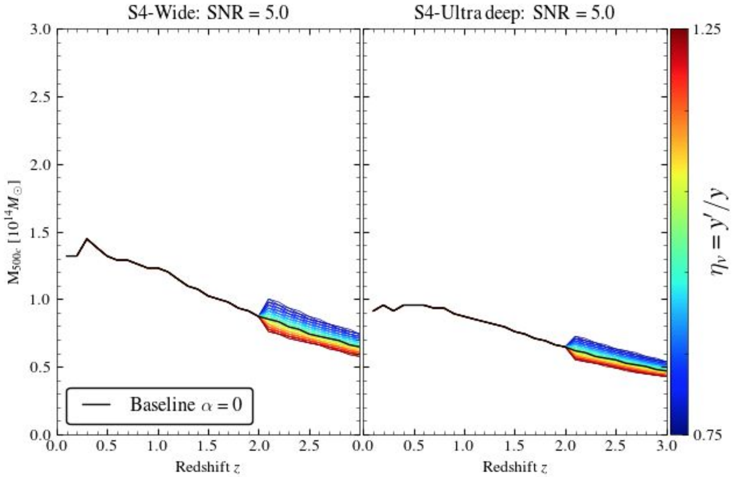
Survey	f_{sky}	Total clusters		$z \geq 2$ mass calibration with CMB lensing	
		All	$z \geq 2$	Temperature	Polarisation
S4-Wide	0.65	117k	850	23%	22%
S4-Wide (Low gal)	0.5	96k	727	25%	24%
S4-Ultra deep	0.03	14k	250	42%	22%

- For S4-Ultra deep, the lensing SNR is dominated by polarisation.
- Foregrounds (tSZ/kSZ) in temperature-based lensing reconstruction are mitigated using a inpainted-gradient QE (Raghunathan, Holder, Bartlett et al. 2019, arXiv: 1904.13392). Also see:
    - Madhvacheril & Hill 2018, arXiv: 1802.08230 for tSZ bias mitigation.
    - Fireslide by Kevin Levy (Friday session) for tSZ/kSZ bias mitigation.
- ***Combing T and P, we can measure the average mass of z≥2 clusters at 18-20 per cent level with S4-Wide or S4-Ultra deep.***

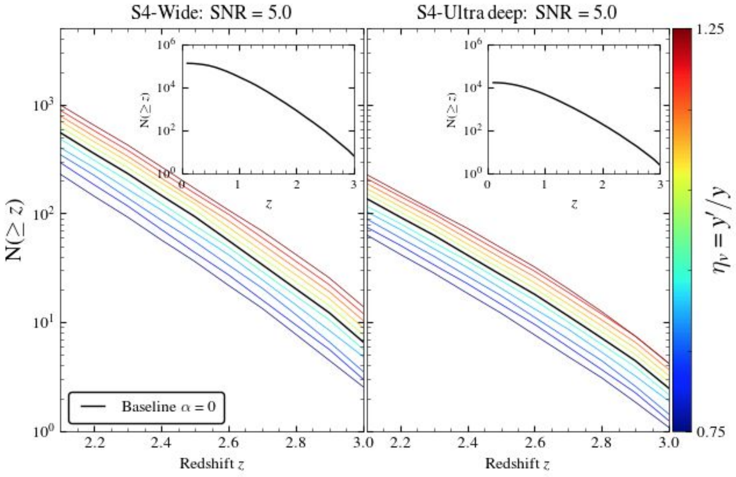
CMB-S4

Modifying the tSZ signal from high-z clusters

Cluster limiting mass vs redshift



Cumulative cluster counts



Enhancement

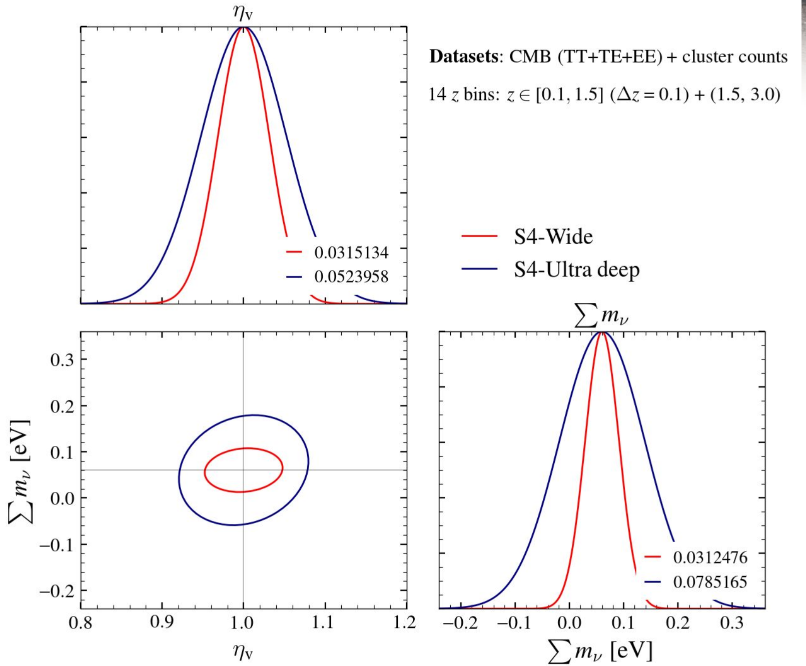
Suppression

$$y' = \eta_v y \text{ with } \eta_v \neq 1 \text{ for } z > 2$$

CMB-S4

Fisher forecasts

- **Datasets:** CMB TT/EE/TE + cluster counts.
- Redshift binning:
  - Binned low-z (z<1.5) clusters with Δz = 0.1.
  - Combined all clusters above z>=1.5 in one large z-bin.
- **Marginalised:** Ysz-M scaling relation / cosmological parameters.
- Not shown here:
  - η_v is degenerate with σ_8 and $1\text{-}b_{SZ}$ if we only pick clusters with z>2.
  - Lower σ_8 → lesser massive high-z clusters.
  - Degeneracies broken when including low-z information.



- ***Constraints on sum of neutrino masses:***
  - No τ prior: 31 (S4-Wide) and 78 (S4-Ultra deep) meV.
  - Planck $\sigma(\tau)$ = 0.007: 23 (S4-Wide) and 42 (S4-Ultra deep) meV.
  - LiteBIRD $\sigma(\tau)$ = 0.002: 14 (S4-Wide) and 34 (S4-Ultra deep) meV.

(See also: Louis & Alonso 2017, arXiv: 1609.03997; and Madhavacheril, Battaglia, Miyatake 2017, arXiv: 1708.07502.)

- ***Constraints on virialisation efficiency:*** τ prior does not affect the results much.
  - ~3.2% (S4-Wide) and 5% (S4-Ultra deep).
  - Including clusters from "high" galactic emission region improves this by ~13%.
  - Using Δz = 0.1 improves constraints by x3 but that is not realistic.

Summary

- **Flowdown text:**

CMB-S4 shall detect (at 5σ) all galaxy clusters with an integrated Compton ***$Y_{SZ} \geq 10^{-12}$ at $z \geq 1.5$ over the large area survey footprint (f_{sky} = ~~65%~~ 50%)****. Furthermore, it shall detect (at 5σ) all galaxy clusters with an integrated Compton* ***$Y_{SZ} \geq 5\times10^{-13}$ at $z \geq 1.5$ over the de-lensing survey footprint (f_{sky} = 3%)****.*

- **Simulations used for validations:**
  - **Galactic foregrounds:** pySM3.
  - **Extragalactic foregrounds:** Simple Gaussian realisations / Websky / MDPL2. Radio sources inside clusters need to be properly modelled.
  - ***Noise / systematics:*** *Simulations from DM group yet to be integrated.*
- **3-5% constraints on $\sigma(\eta_v)$ if we fully understand physics of low-z clusters.**
- **Link to manuscript:** https://www.overleaf.com/read/xvwggqtwjkkz

DV: M2Science - Galaxy Clusters: Parallel

Convener: Nicholas Battaglia (Cornell University).

Date/time: 10 March 2021 at 3.30 p.m. eastern / 2.30 p.m. central / 12.30 p.m. pacific.

Talks:

- ***Optical Follow-Up*** by Lindsey Bleem (Argonne National Laboratory & KICP).
- ***DM Simulations*** by Reijo Keskitalo (Lawrence Berkeley National Laboratory & UC Berkeley).
- ***Source Modeling*** by Zack Li (Princeton University).
- **Discussion.**

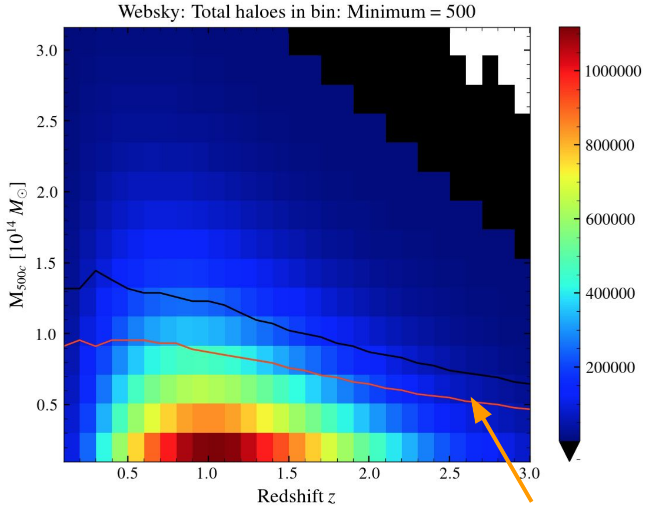
Backup slides

Effect of cluster correlated CIB/kSZ signals using websky simulations

Injecting CIB/kSZ signals from websky

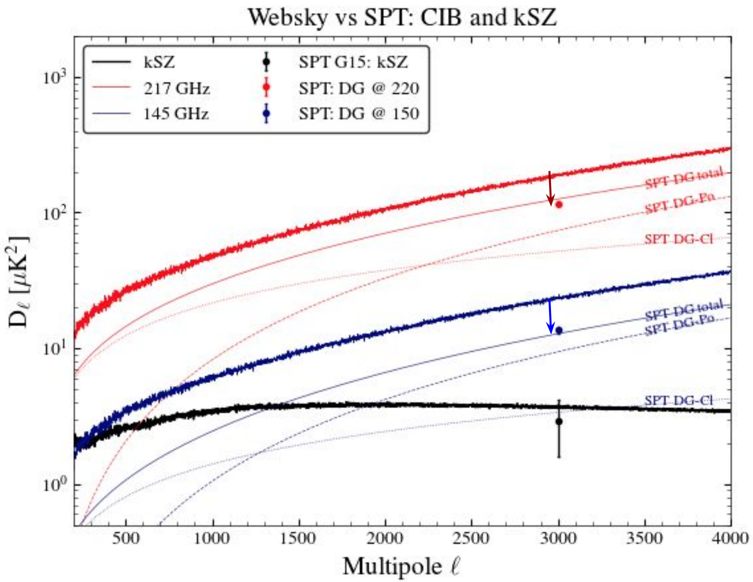
- Replaced Gaussian realisations of CIB/kSZ signals using websky simulations. (DG stands for dusty galaxies here or CIB.)
- Done by picking 2 deg x 2 deg cutouts from websky CIB/kSZ from
  - locations corresponding to cluster (M,z) under study. Number of websky haloes in each M,z bin is shown below in Figure (A).
  - random locations.
- Websky CIB signals were scaled down by ~0.75 to match measurements made by SPT (George et al. 2015 - G15: arXiv: 1408.3161), see Figure (B).
  - **Caveat:** This simple scaling does not make the clustering component of the websky DG signals identical to SPT, but it does not seem to matter much.

Figure (A): Black = <500 haloes; White = No haloes.

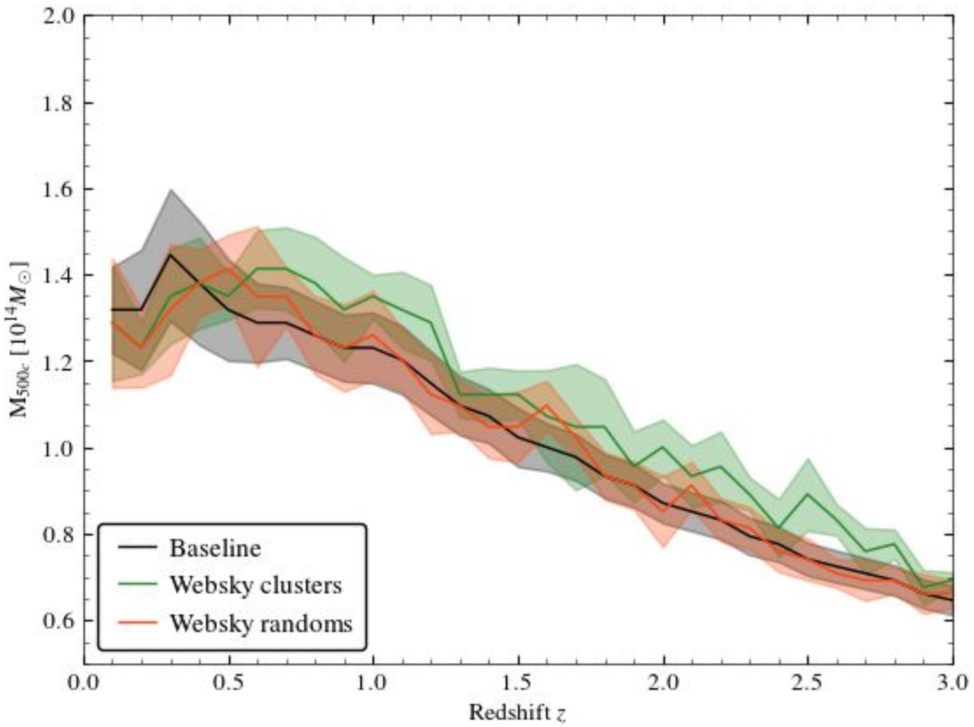


Limiting mass vs z curves for S4-wide (black) and S4-Ultra deep (red).

Figure (B)



Effect of CIB/kSZ signals from websky



Limiting mass vs z for S4-wide
(Have not checked this for S4-Ultra deep survey. But the results shown above should hold good for S4-Ultra deep too as their limiting masses are smaller than S4-wide.)

Black: Baseline results in which CIB/kSZ are Gaussian realisations and have no correlation with the cluster.

Red: Websky CIB/kSZ extracted from random locations: Not correlated with the cluster. Results match the baseline results.

Green: Websky CIB/kSZ extracted from locations corresponding to halo (M,z) under study: cluster correlated CIB/kSZ signals. Results are only slightly worse than black / red curves at high-z.

Summary: Cluster correlated CIB/kSZ signals degrade the high-z cluster SNR but not significantly. CIB is probably the main player here but we did not check that explicitly.

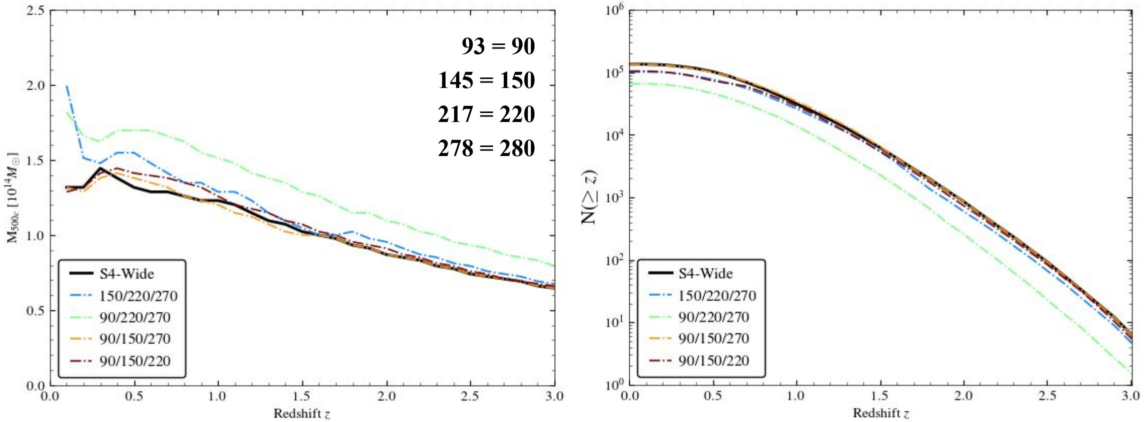
(Note: We have used limiting mass vs z as a proxy for cluster counts to analyse the results here although the latter is the actual science driver.)

Note (noisier websky curves compared to baseline setup in black):
In the baseline setup, the foreground signals introducing variance are the same for all clusters (M,z). This is not true when using websky, as the CIB/kSZ signals change for each cluster. This is why the curves with websky signals are noisier than baseline results in black.

Importance of individual bands for S4 cluster detection

Eliminating one band - Part I

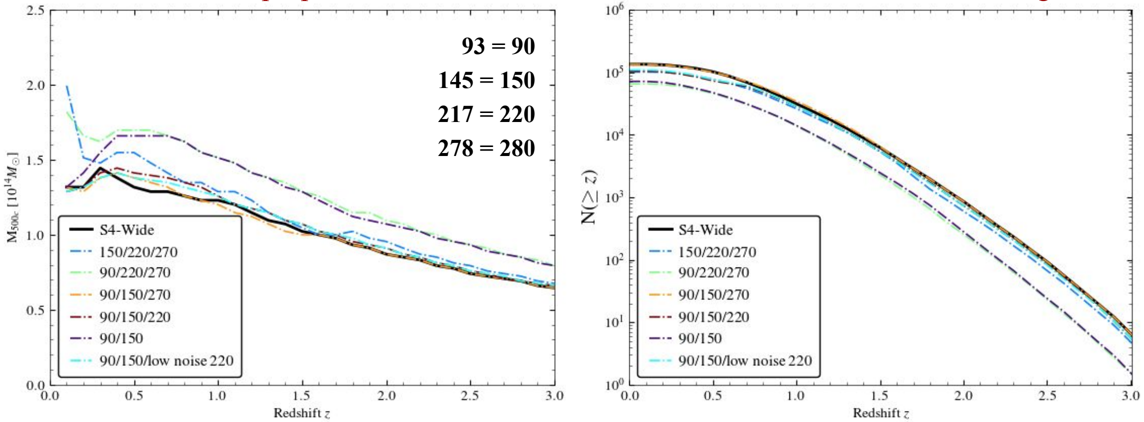
- We also checked the importance of individual frequency channels for cluster detection.
- Done by simply eliminating one band from the analysis. *Note that, unless specified otherwise, we do not reduce the noise of other channels when eliminating a band.*
- Simulations are simple Gaussian realisations which do not have proper CIB decorrelation between different bands. But, we investigated this with Websky as well (slide 9).



- **Black:** Baseline results with all four (93/145/217/278 GHz) bands.
- **Blue:** 93 GHz band has been removed and 145 GHz channel gets most of the weight now. Note that S4 93 and 145 GHz channels have similar white noise levels. However, the results w/o 93 are much better than results w/o 145 (green) because there is almost zero contamination from radio galaxy signals in these bands. At high-z, the 150 GHz beam is smaller than 93 GHz channel, but this should not be a significant effect. At low-z, the limiting mass increases because all the three (145/217/278) bands have a higher 1/f noise compared to 93 GHz.
- **Green:** w/o 145 GHz band: Results are significantly worse than everything else. 93 GHz gets most of the weight here but it has significant radio contamination and the 217/278 GHz bands are not helping much. Like mentioned above, 93 GHz beam is much worse than 145 GHz and this degrades the high-z end slightly.
- **Orange:** w/o 217 GHz band; **Red:** w/o 278 GHz band: The results do not differ significantly from baseline results. This is expected as these channels are mostly helping to handle the CIB variance. 278 GHz seem to be slightly more important than 217 GHz, probably because (1) CIB is less bright in 217 than in 278; (2) 278 GHz band is also probably helping a bit with tSZ while 217 has no tSZ. But, these remain less clear. We also tried removing both the high frequency bands - see next slide.

Eliminating one band - Part II

- We also checked the importance of individual frequency channels for cluster detection.
- Done by simply eliminating one band from the analysis. *Note that, unless specified otherwise, we do not reduce the noise of other channels when eliminating a band.*
- Simulations are simple Gaussian realisations which do not have proper CIB decorrelation between different bands. But, we investigated this with Websky as well (slide 9).

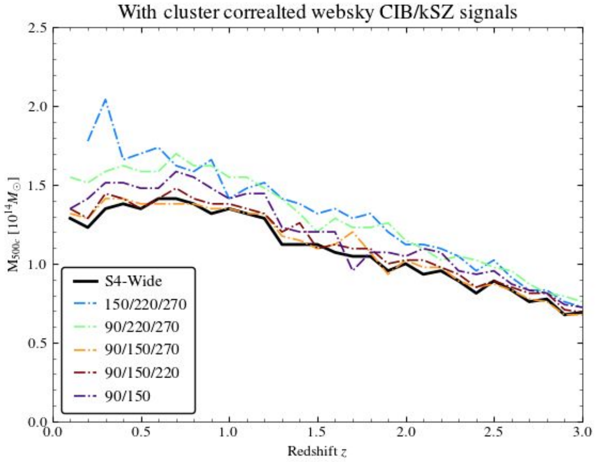


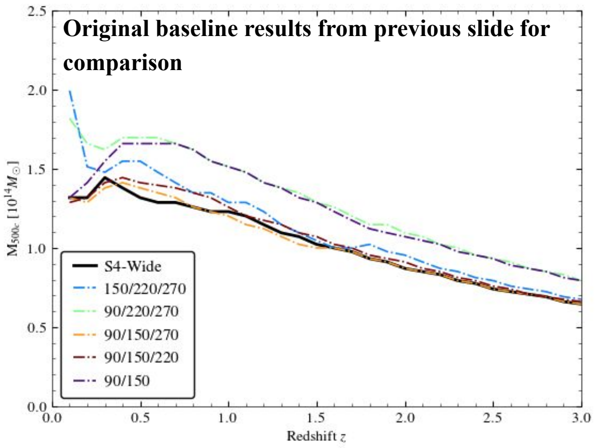
- **Purple:** Without both high (217/278 GHz) frequency bands. CIB becomes an issue here implying one or both of them are important to handle CIB.
- **Cyan:** Moved 278 GHz detectors to 217 GHz band → This reduces 217 GHz noise from 6.9 uK-arcmin to 6.3 uK-arcmin (although this simple inverse variance weighting may not work in practise). In any case, reducing the 217 GHz noise (cyan vs red) only has a marginal influence on the results.
- Cluster counts shown in the right panel seems to be in agreement with the limiting mass vs trends, as expected. Hence, we did not explore them further.

Summary:

- 150 GHz band is the most important channel. 90 GHz is the next most important. Removing either of them degrades SNR.
- One or both of the high (217/278) frequency bands are needed to handle CIB. 278 seem to be a bit more important than 217 GHz, but it is not 100 per cent clear.
- **We will now repeat the above analysis using CIB from websky that include proper (or a better) CIB decorrelation between different bands. See next slide.**

Testing with websky CIB: Eliminating one band





- Left plot shows the results after swapping Gaussian realisations of CIB/kSZ with cluster correlated signals from websky.
- These results seem to be roughly in agreement with the original results shown in right for reference.
- There are some differences. For example, difference between the curves seem to be smaller at high-z in the left compared to right, but those differences are well within the error bar (not shown here).
- **Summary: The inference we obtained from baseline setup -- removing 150 / 90 GHz degrades SNR; and one or both high frequency bands are needed -- remains unaltered.**

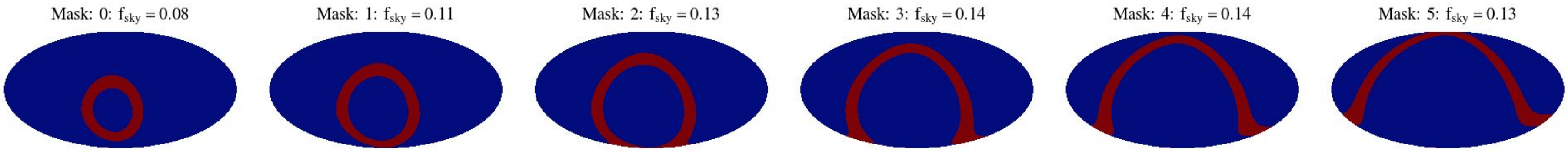
Note (noisier websky curves compared to baseline setup in black): In the baseline setup, the signals introducing variance are the same for all clusters (M,z). This is not true when using websky, as the CIB/kSZ signals change for each cluster. This is why the curves with websky signals are noisier than baseline results in black.

Inclusion of galactic dust and synchrotron signals

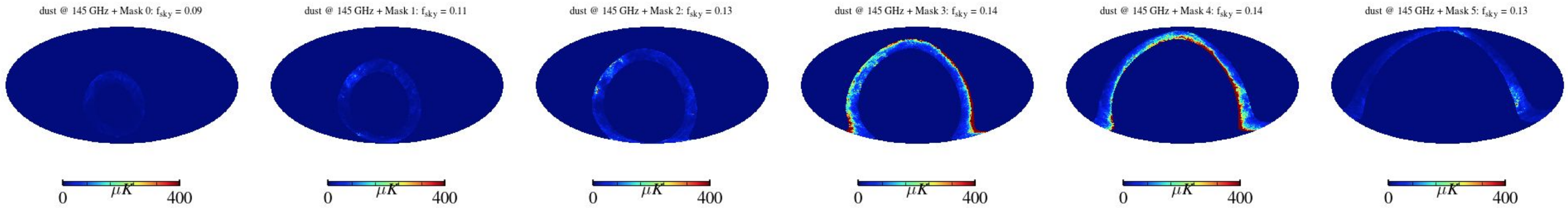
Adding galactic emission using pySM3 simulations

- In all the previous forecasts, we have ignored the effect of galactic dust/synchrotron signals.
- But given that S4-Wide (Chilean LAT) will survey a large fraction of sky (fsky>=0.6), assuming galactic emission does not affect us is not entirely true.
- This is fine for S4-Ultra deep (delensing LAT), though.
- In this exercise, we study the effect of galactic emission on S4-Wide cluster search.
- We do this by splitting the S4-Wide footprint into six different chunks (latitude = -60 to 30 degrees with delta_lat = 15 degrees) and assume that the galactic emission is close to constant within each of these subfields. See the images below.
- The signals come from pySM3 simulations picked from https://github.com/CMB-S4/s4mapbasedsims/tree/master/202002_foregrounds_extragalactic_cmb_tophat

S4-Wide subfields

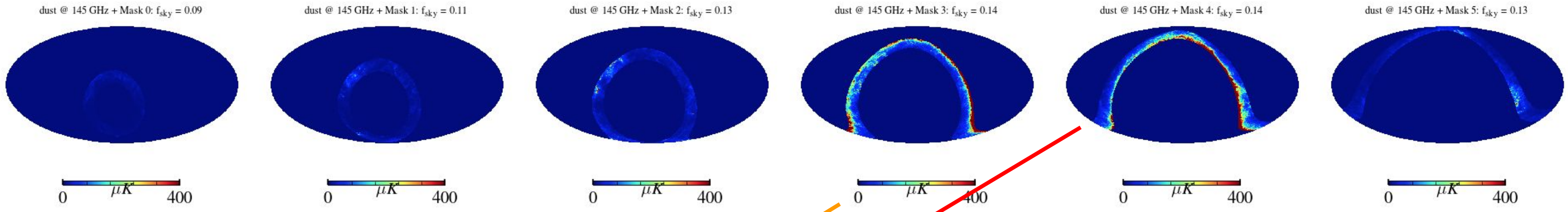


Galactic dust at 145 GHz in each S4-Wide subfields

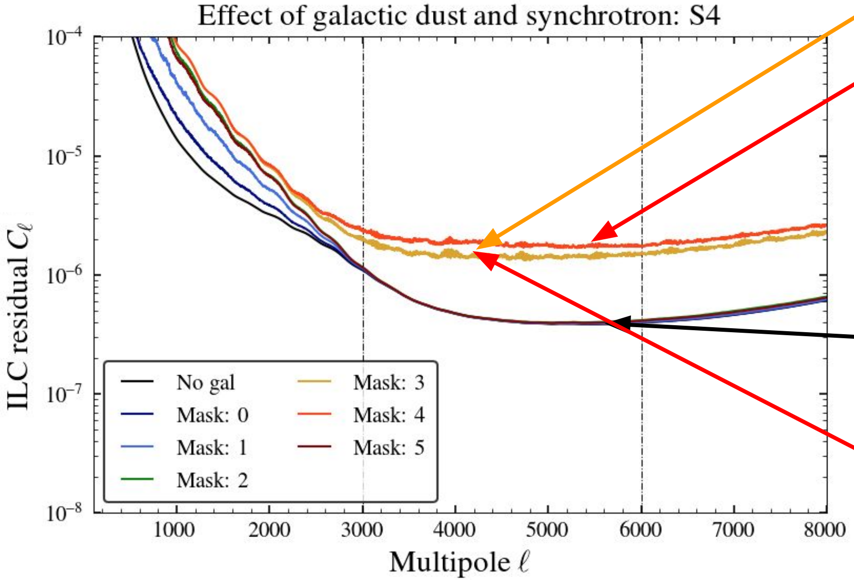


Adding galactic emission using pySM3 simulations

Galactic dust at 145 GHz in each S4-Wide subfields



Compton-y noise curves



Plot on the left compares the noise level of Compton-y maps in different subfields (with different levels of galactic emission) to the baseline S4 Chilean LAT shown in black (with no galaxy).

First thing we note that is that adding galactic emission increases the large-scale noise, as expected. The increase is consistent with the amount of emission in the above figure: mask 0 (blue) > mask 1 > … > mask 4 (red).

At cluster-scales (highlighted band), the noise curves for Mask: 0/1/2/5 are not different from baseline case (with no galactic emission).

When we are looking right through the plane of our galaxy, the noise increases by >x3 on small scales. This will degrade the SNR of clusters in these two subfields.