Long Wavelength Astronomy in SA

Jonathan Sievers (UKZN)

Come be Warm! (Feb 13-17 2017)

http://www.acru.ukzn.ac.za/~cosmosafari/

All pictures by Cynthia Chiang over 1 weekend

Come be Warm! (Feb 13-17 2017)

http://www.acru.ukzn.ac.za/~cosmosafari/

All pictures by Cynthia Chiang over 1 weekend

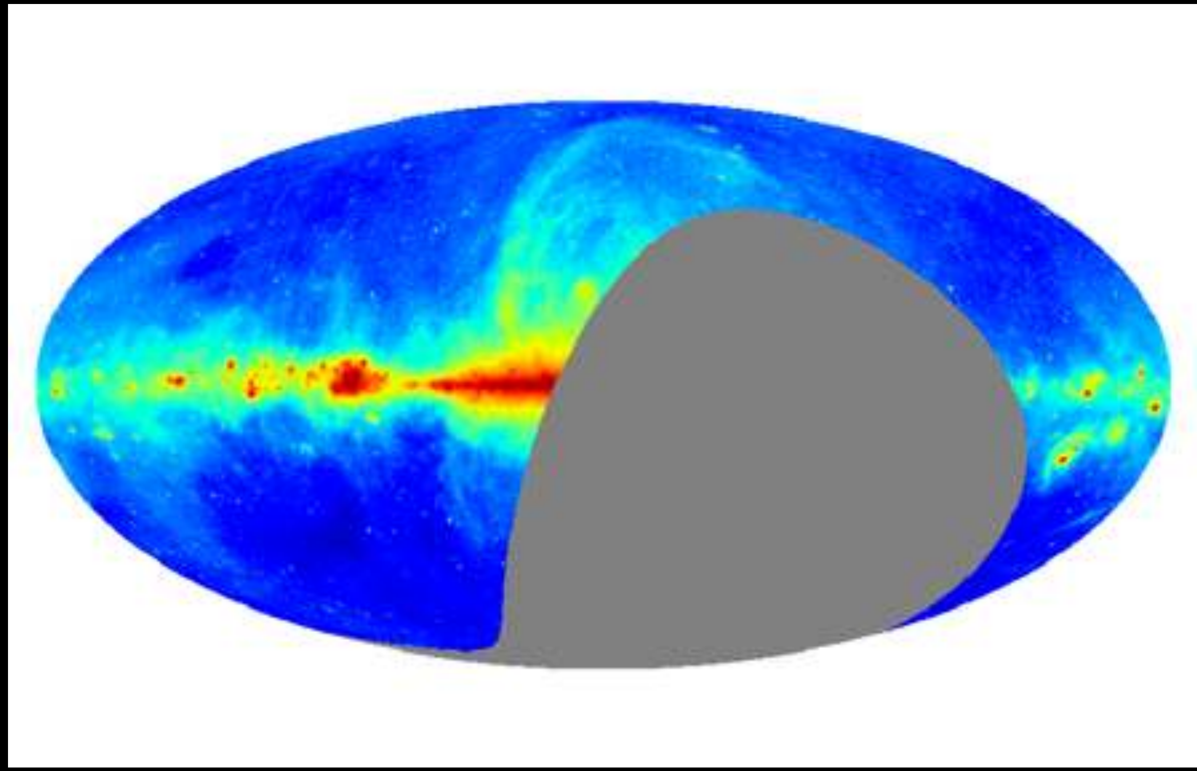
Not Just SKA

- Fantastic radio-quiet site has attracted other experiments
- Comes along with people working in the field
- Who bring pre-existing collaborations
- Leading to explosion of other datasets
- SA hosts C-BASS, PAPER/HERA, will host HERA
- SA people involved in ACT/SPT/SPIDER/Planck…
- These are all survey instruments/experiments - full datasets need to be analyzed

C-Band All-Sky Survey (C-BASS)

- Inflation probably gave rise to our universe. We know very little abou it.
- (Possible) handles - n_s, tensor-to-scalar (r), non-gaussianity
- CMB experiments going after r very hard, but foregrounds are big concern (see BICEP-2)
- Main polarized foregrounds are dust & synchrotron. Most expt's can do dust, but synchrotron more painful
- C-BASS surveys sky from California, SA at 5 GHz to provide synchrotron template. Single dish for large angle reconstruction.

C-BASS Data



- Northern system has single band @5 GHz
- Southern system has 128 channels, 100 Hz sampling rate
- Southern data rate ~20 GB/day
- Full survey will require simultaneous analysis of a few TB

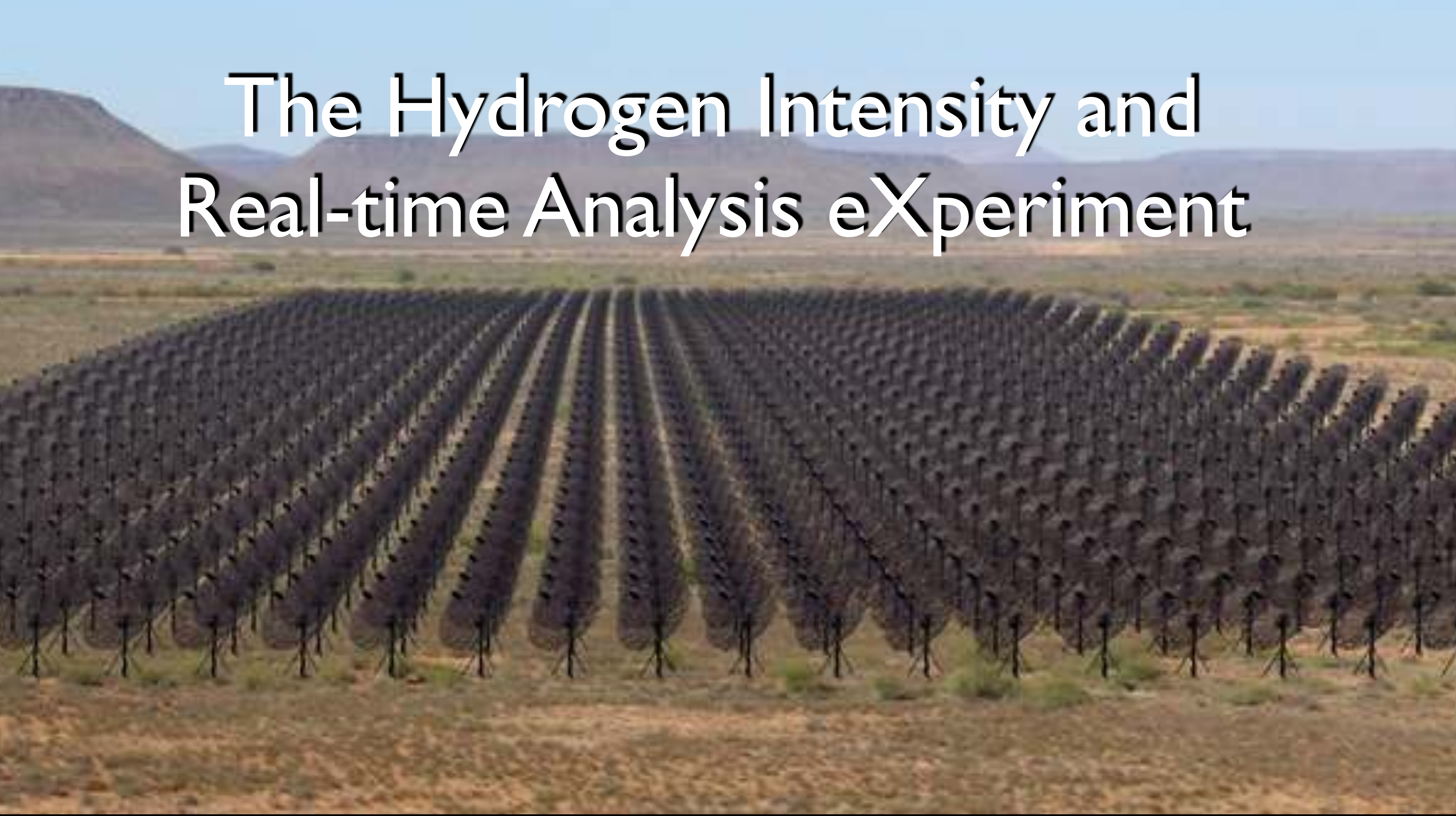
PAPER

- Precison Array for Probing the Epoch of Reionization (PAPER) has surveyed southern sky from SKA-SA site.
- PAPER placing limits on reionization from first stars.
- Staged in 32/64/128-element phases. Full dataset now complete.
- Native data rate @128 - $128^2/2*256$ channels @1 minute sampling, full pol = 25-50 GB/day

HERA

- HERA is follow-on to PAPER. Final plan ~350 antennas looking for EoR power spectrum (and maybe imaging?)
- Sited next to PAPER. Initial phase has already begun.
- Final native data rate - $350^2/2*1024$ channels, ~30 sec. sampling = 3-6 TB/day.

The Hydrogen Intensity and Real-time Analysis eXperiment



http://www.acru.ukzn.ac.za/~hirax/

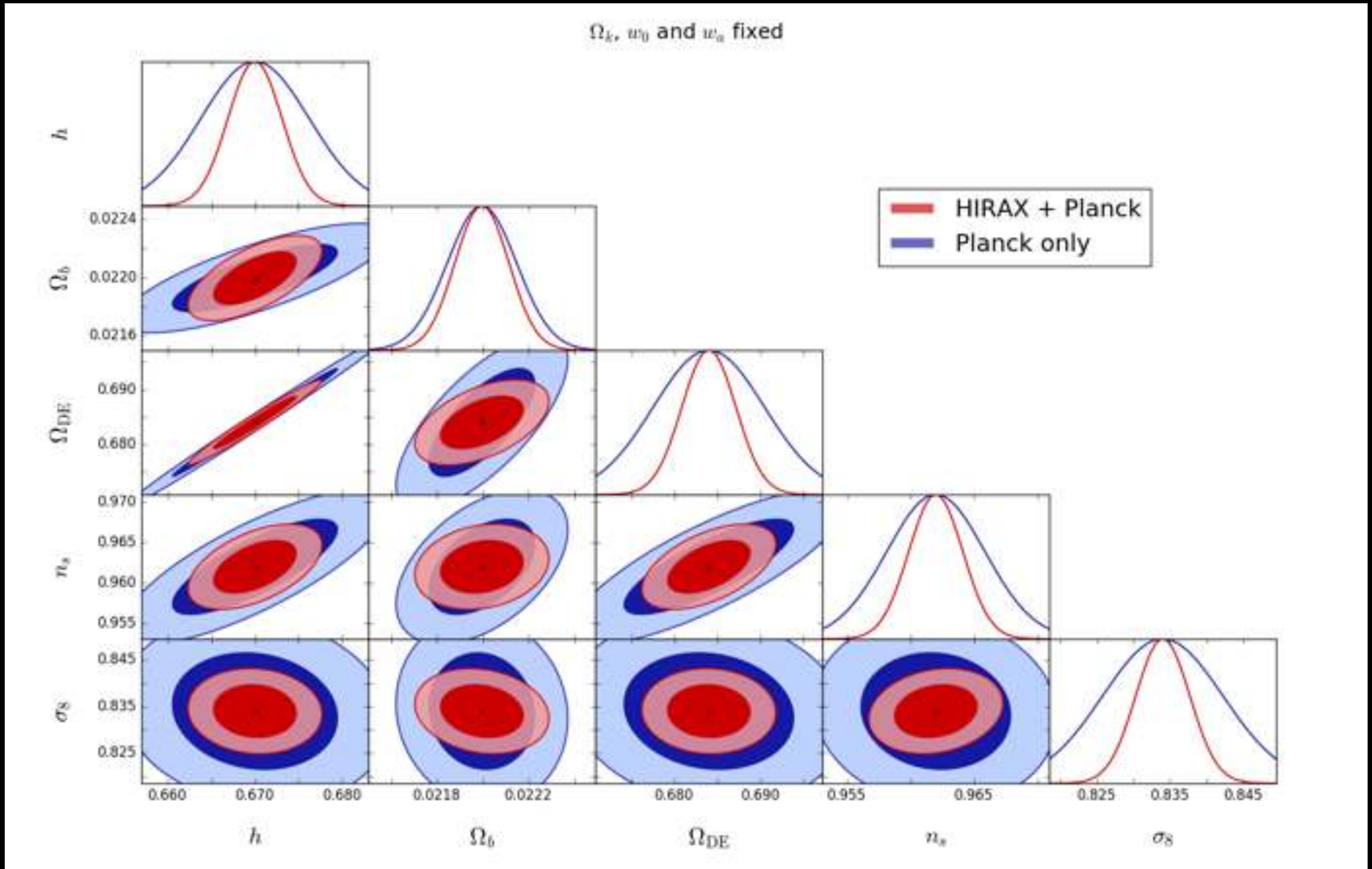
HIRAX - Intensity Mapping in South

- Hope to build 1024 dishes, have funding from UKZN, NRF for 128 in-hand.
- Main science goal is to do intensity mapping of BAOs in HI from z=0.8 to 2.5 (400-800 MHz).
- Sets design - small elements in compact array. BAO signal faint, so much collecting area.
- We will use 6m dishes on a regular array.
- Also fantastic for pulsars, FRBs, HI absorbers, galactic magnetism…

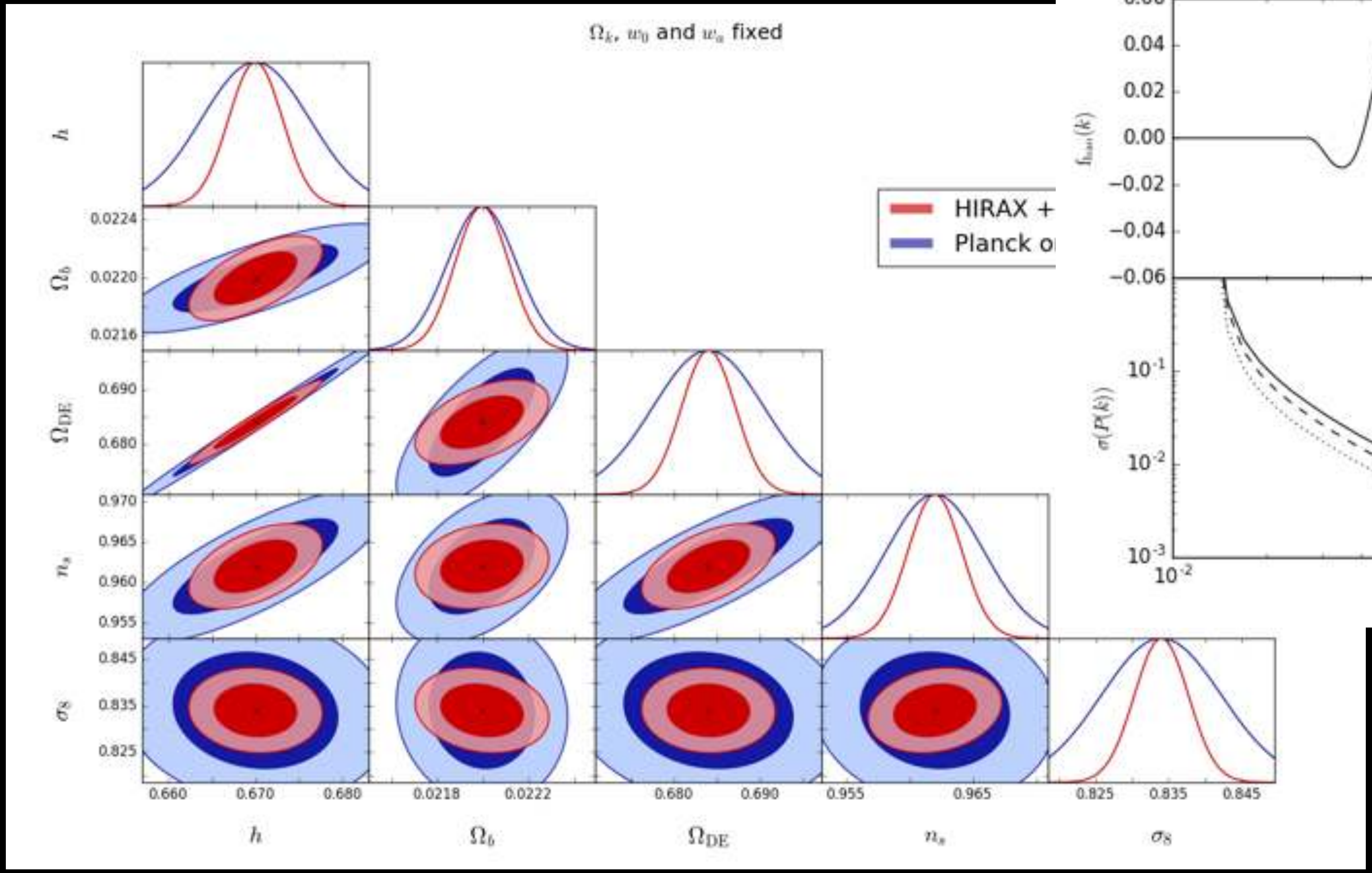
HIRAX - Intensity Mapping in South

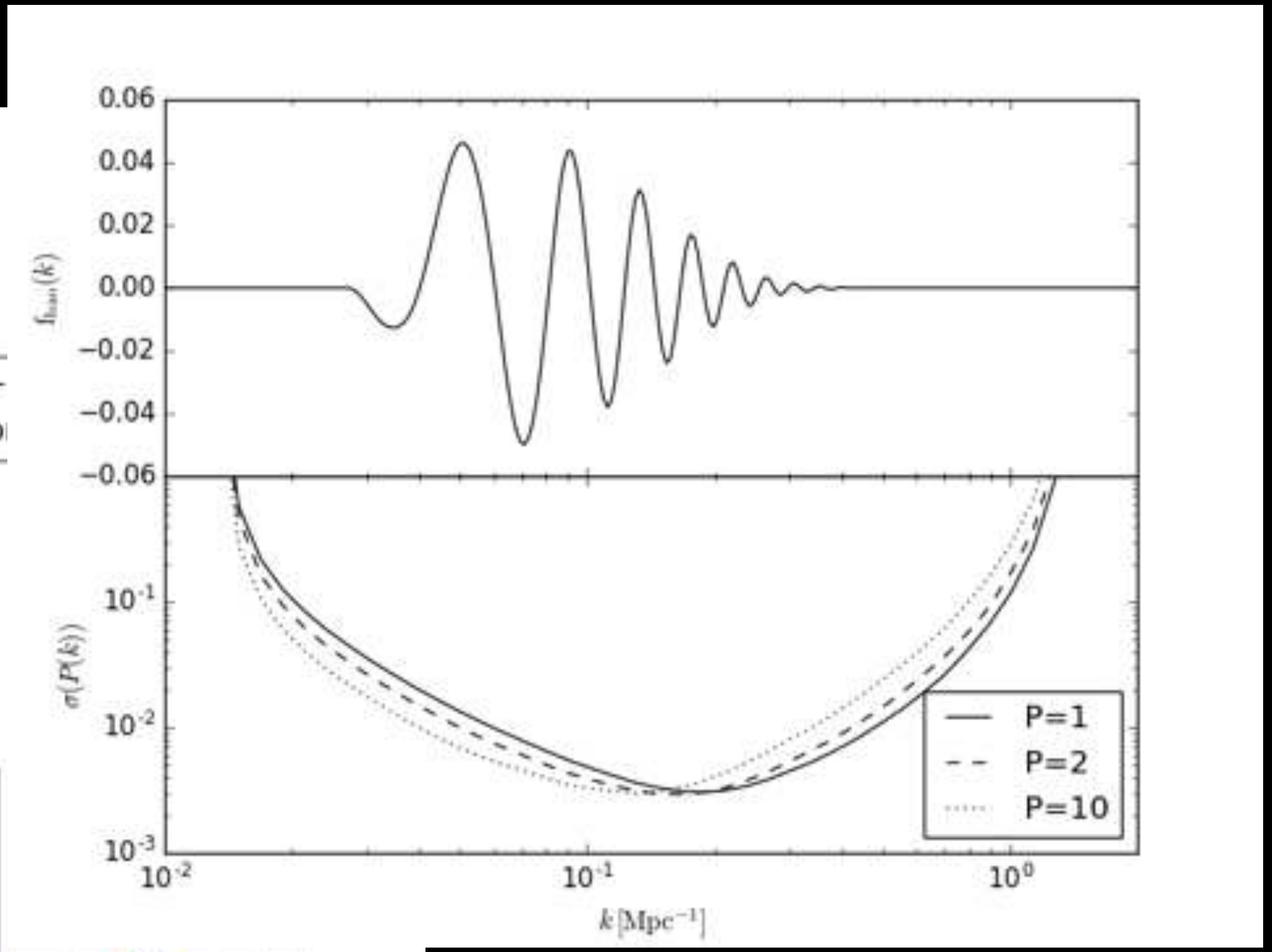
- Hope to buil… 28 in-hand.
- Main science… z=0.8 to 2.5 …
- Sets design - … so much collect…
- We will use 6…
- Also fantastic… …

Preliminary Forecast (Amadeus Witzemann and Mario Santos)



Preliminary Forecast (Amadeus Witzemann and Mario Santos)





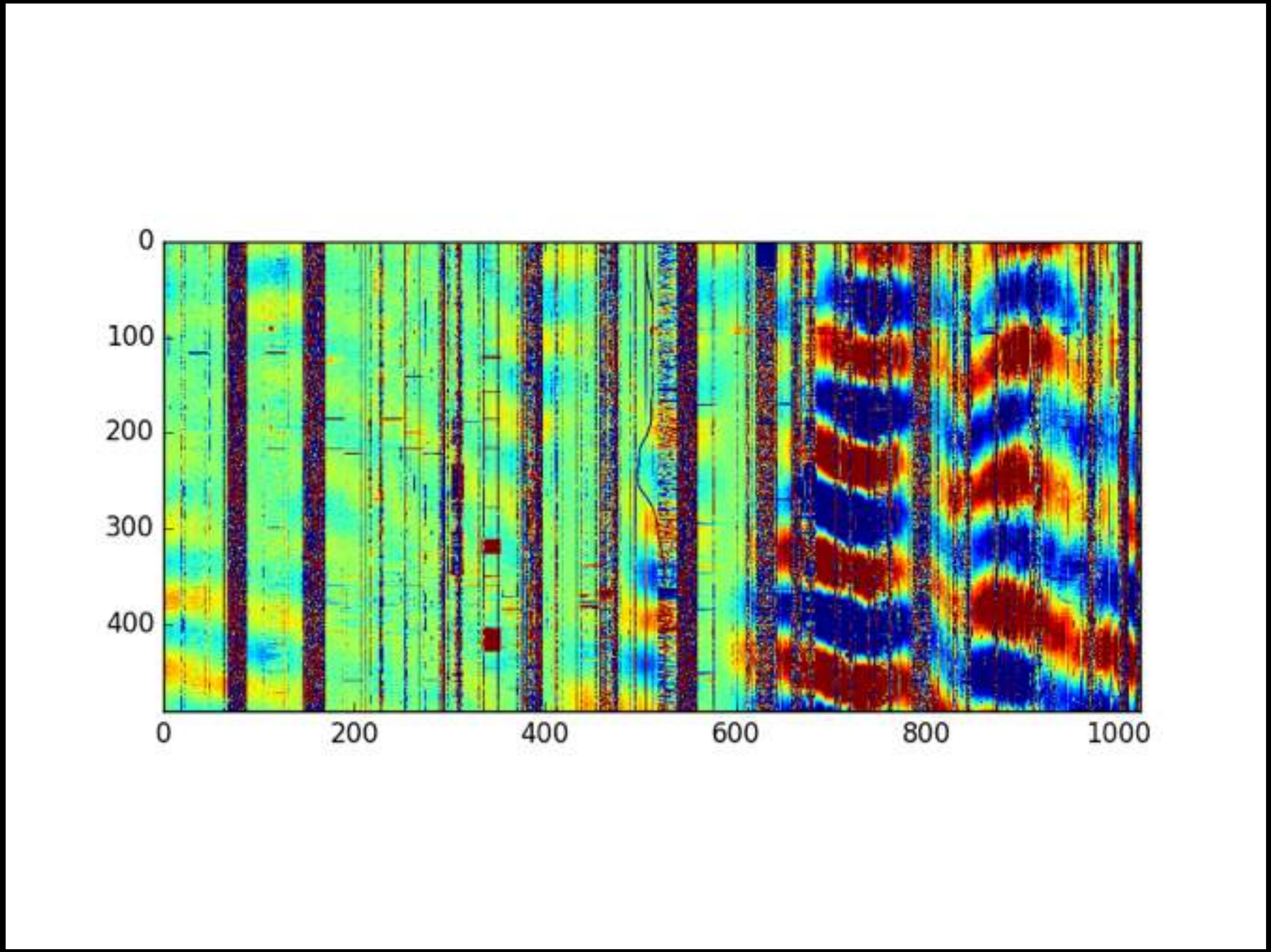
Status/Data Rate

- Building prototype at HartRAO outside of Johannesburg
- Plan there for 8-elements. Digitizing/channelizing on one FPGA board (Matt Dobbs/McGill), correlation on one GPU (Vanderlinde/Toronto)
- Single board/GPU combination has similar capability to pre-upgrade VLA
- Native data rate from BAOs - $1024^2/2$, 1024 channels, ~30s samples = 25-50 TB/day. Might need to process up to several PB simultaneously.

Fringe!

- Low frequency on right side of plot. Time increasing down. Ripples on right side are from an astronomical source.

- Vertical columns of stuff is RFI - we would lose these frequencies from HartRAO.



Redundancy

- PAPER, HERA, HIRAX (and other experiments!) are regularly spaced.

- For a perfect array, same baselines should have same values.

- Takes # of distinct baselines from $n(n-1)/2$ to $\sim 2n$

- Ignoring redundancy - PAPER OK, HERA hard, HIRAX almost impossible. @200 MB/s takes 2 years to read one year of HIRAX.

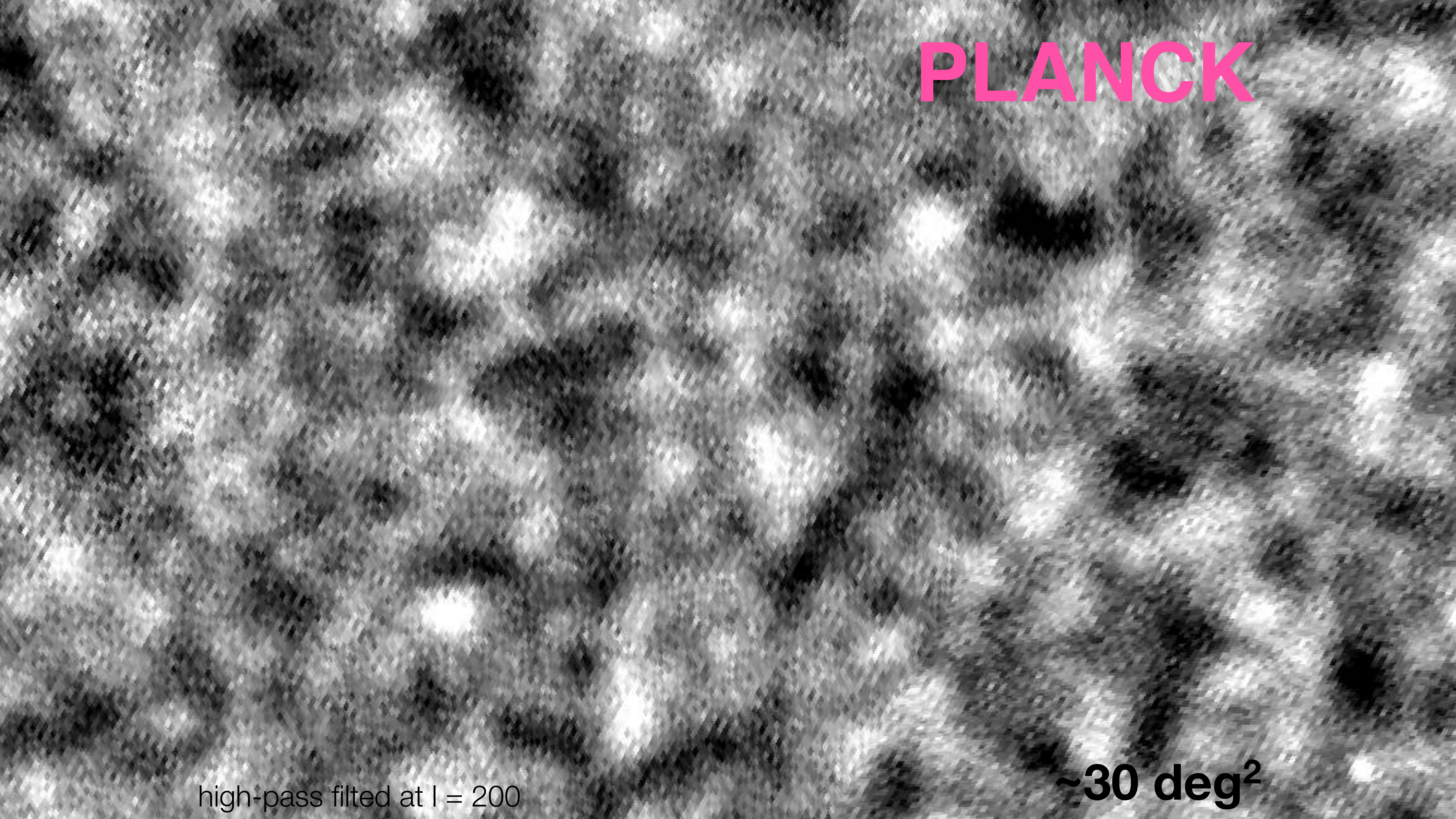
- W/redundancy HERA goes to 30-60 GB/day, HIRAX to 100-200 GB/day. Much more tractable.

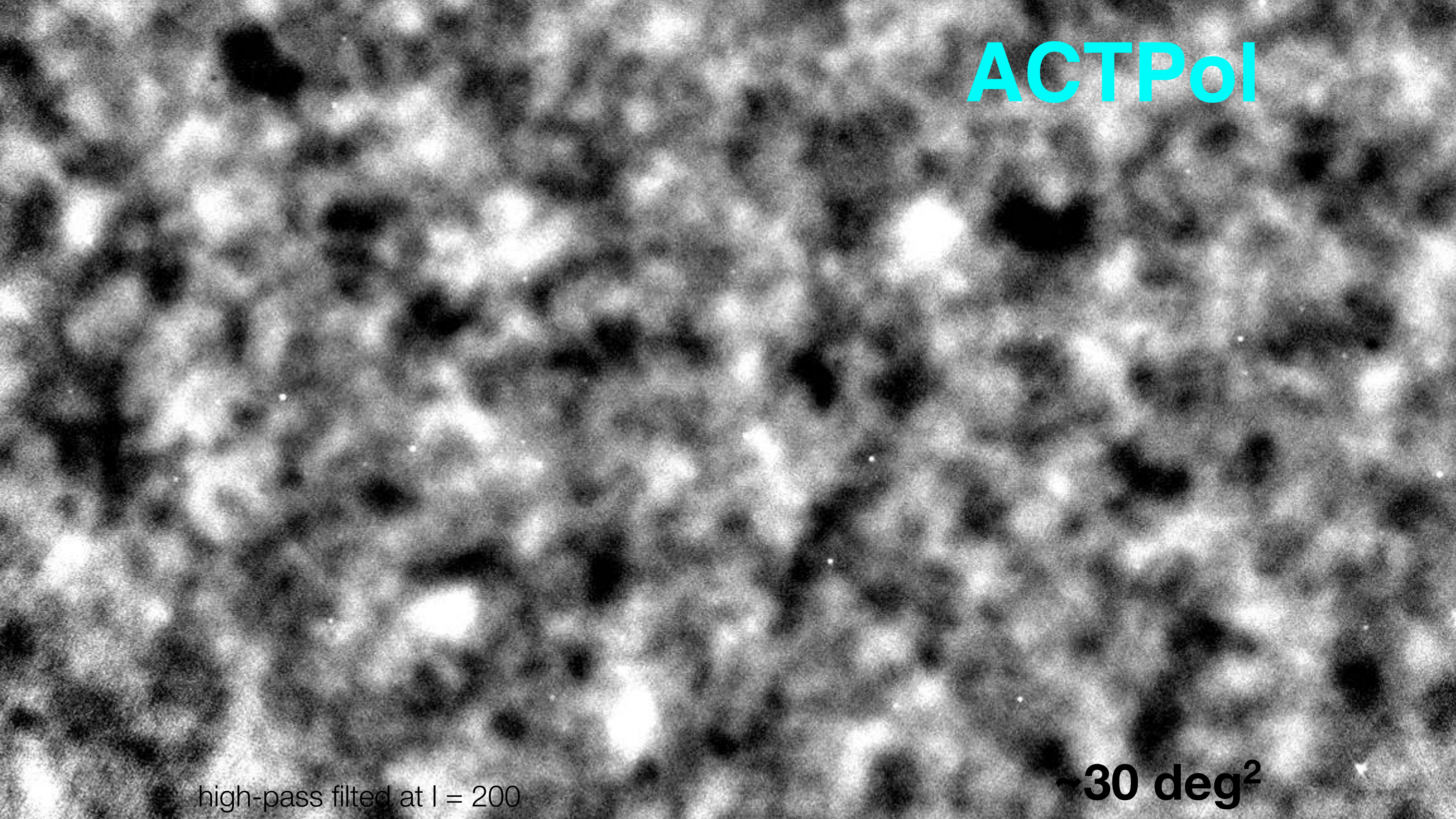
Unfortunately…

- Physical arrays are never perfectly redundant, data must be calibrated before combining
- Large foregrounds (10^3-10^4 times signal) can leak in. Naive combining will destroy usefulness of data
- Possible solution - quasi-redundant calibration/saving of eigenmodes. At realistic positional accuracy/foreground strengths inflates data volume by factor of 3.
- Inspired by omnical (Liu et al.) redundant calibration. Quasi-redundant assumes sky described as point sources plus gaussian random field described by power spectrum. Optimal under those assumptions.
- Testing out reference implementation - stay tuned for paper/pipeline.

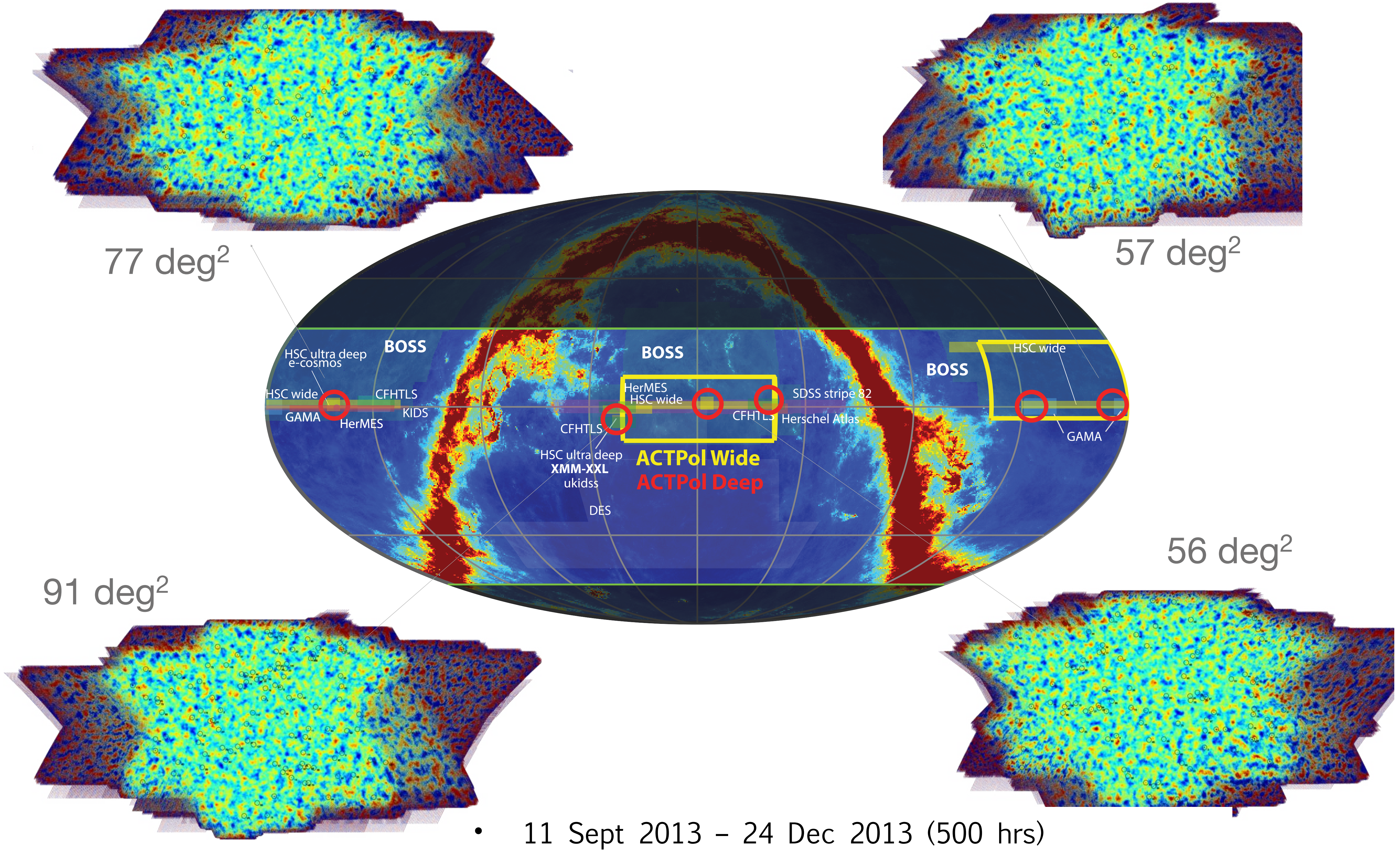
The Atacama Cosmology Telescope

- 5200 m (high)
- Desert (dry)
- Latitude -23°

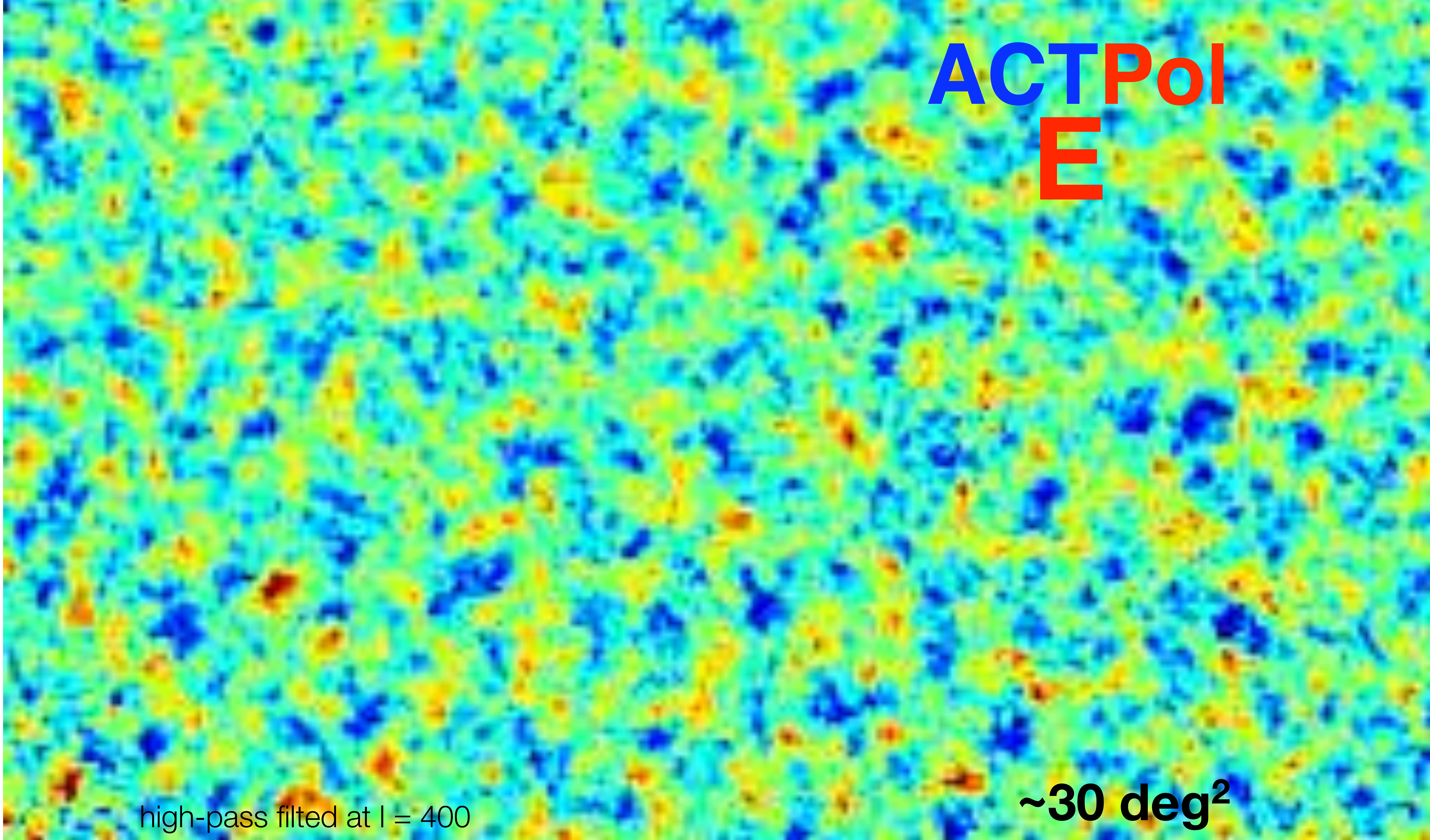

PLANCK
~30 deg^2
high-pass filted at l = 200


ACTPol
high-pass filted at l = 200
~30 deg²

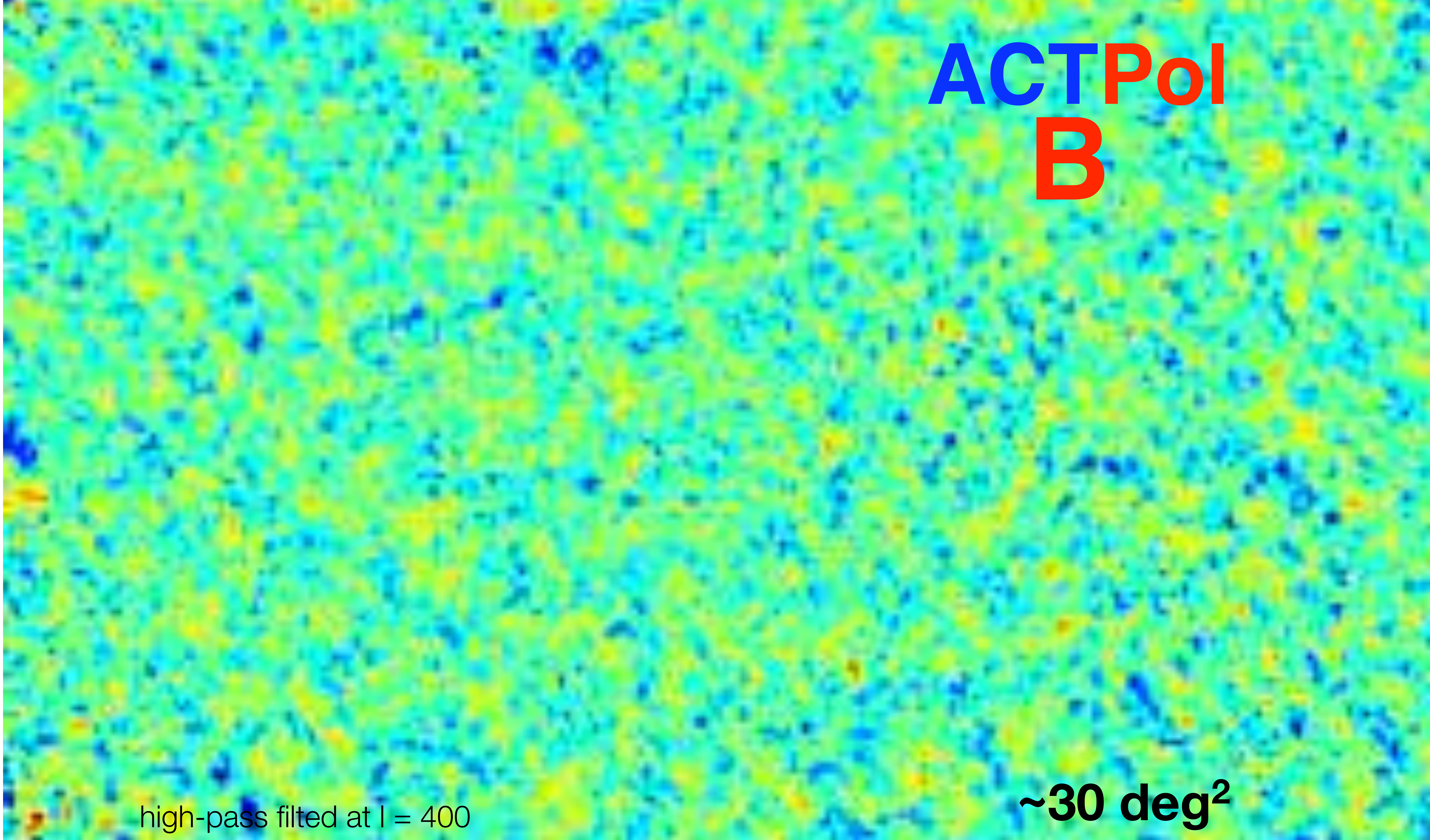
ACTPol Survey 2013



- 11 Sept 2013 – 24 Dec 2013 (500 hrs)
- Only one 150 GHz array installed then
- Four ~ 70 deg^2 patches



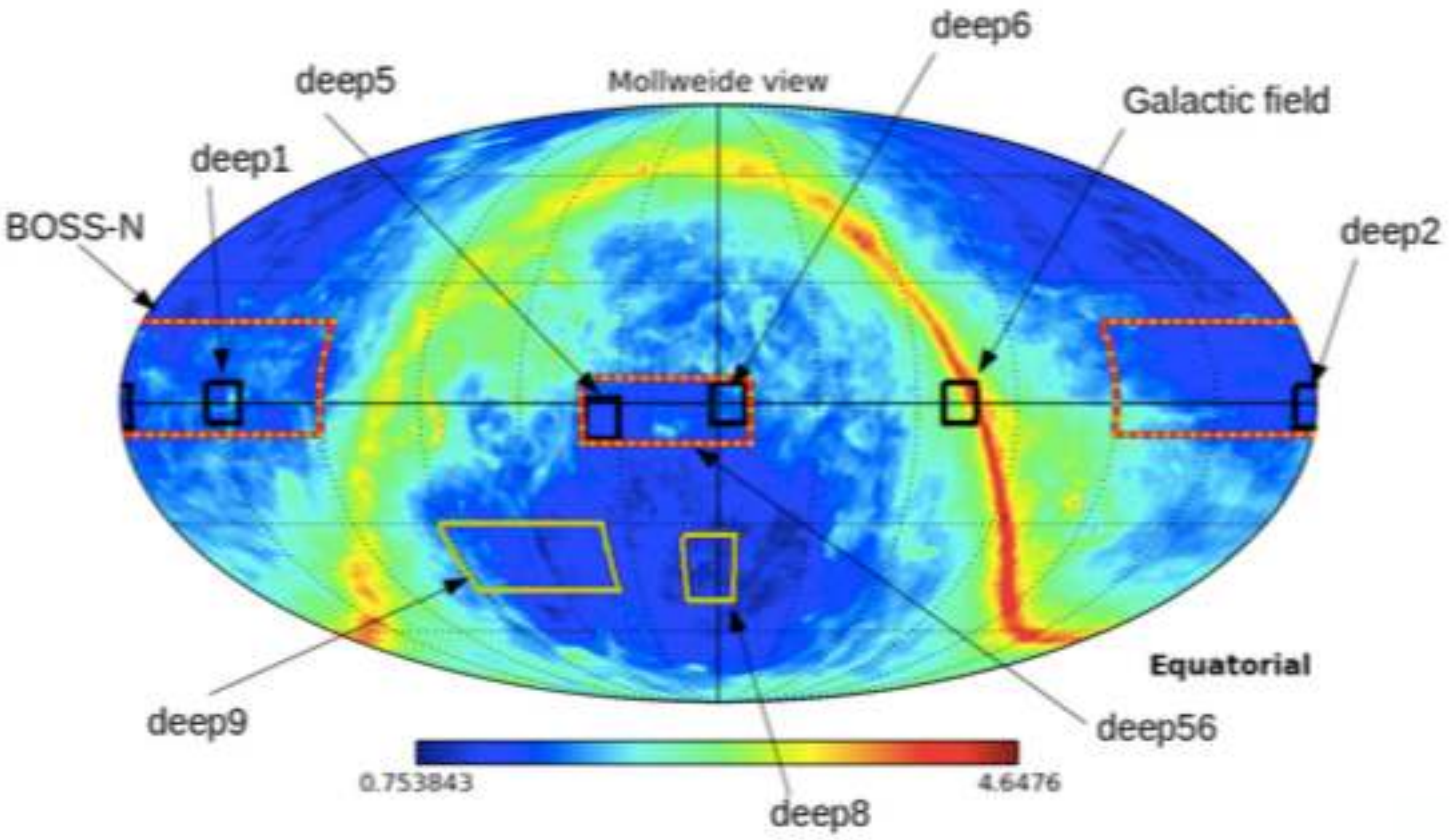
ACTPol
E
~30 deg^2
high-pass filted at l = 400


ACTPol
B
~30 deg^2
high-pass filted at l = 400

Full ACTPol survey

2014 & 2015

- Data already complete
- ~2000 square-degrees (BOSS N)
- ~600 square-degrees (deep 5-6)



The Near Future: ACTPol Data Analysis

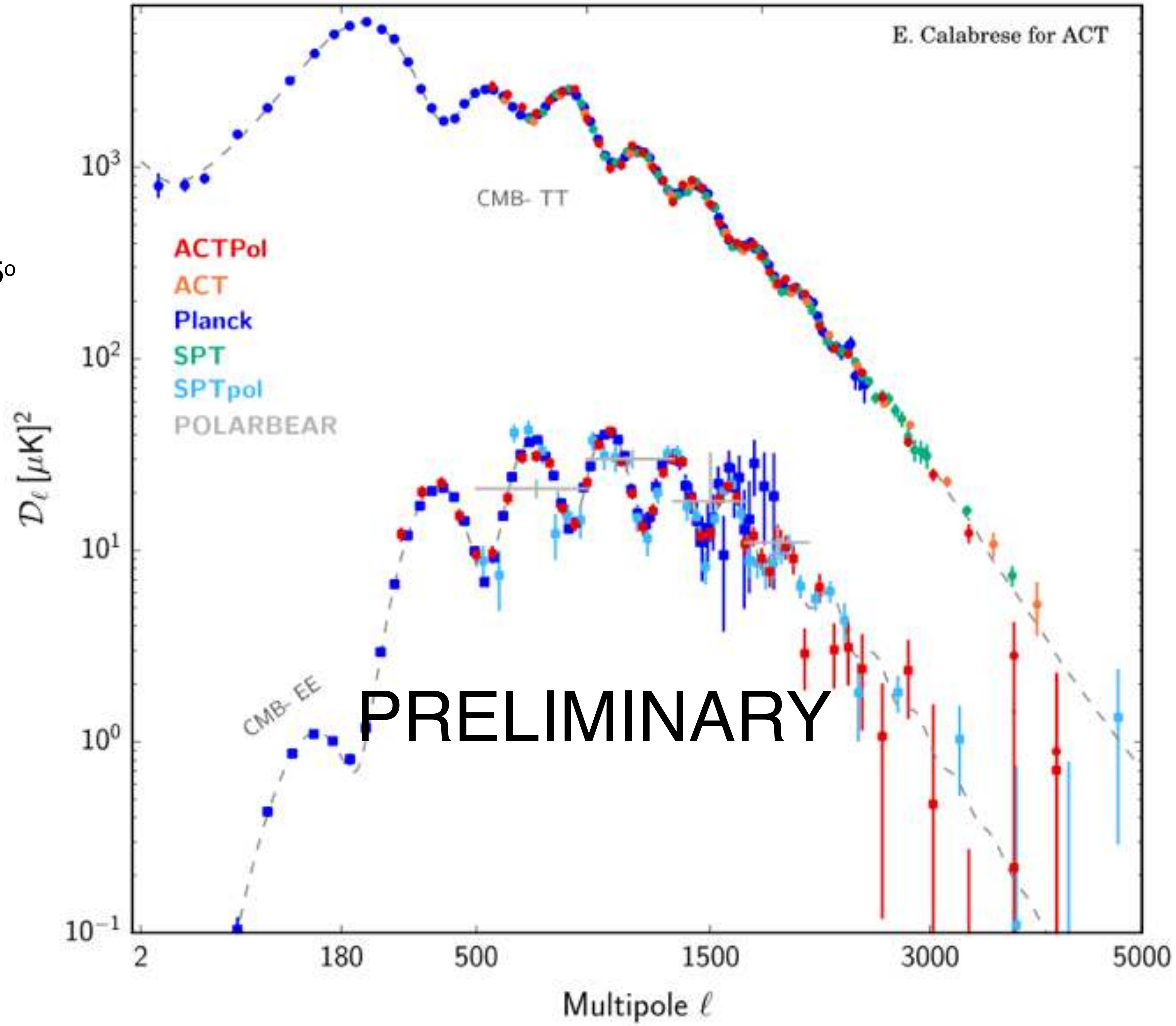
preliminary results

D56 Field
(12% of the ACTPol data)

~ 650 deg^2 , @ $\delta \sim -3^o$, RA ~ 15^o

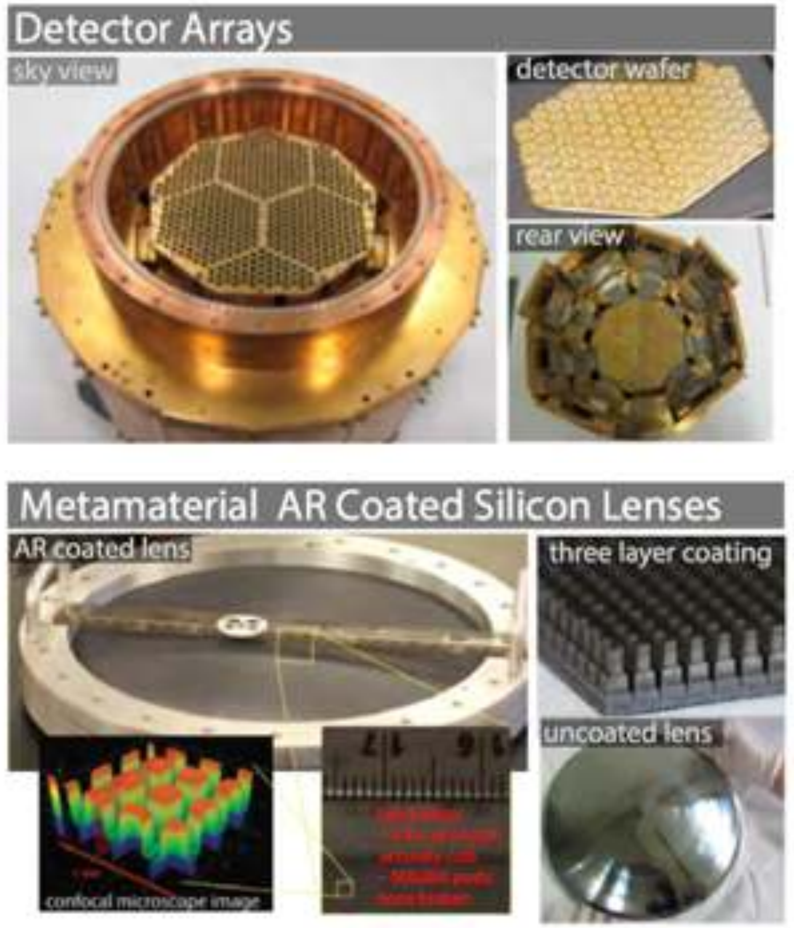
up next:

(1) lensing, SZ, and cross correlations with this map
(2) analyze the rest of 2014 and all of 2015



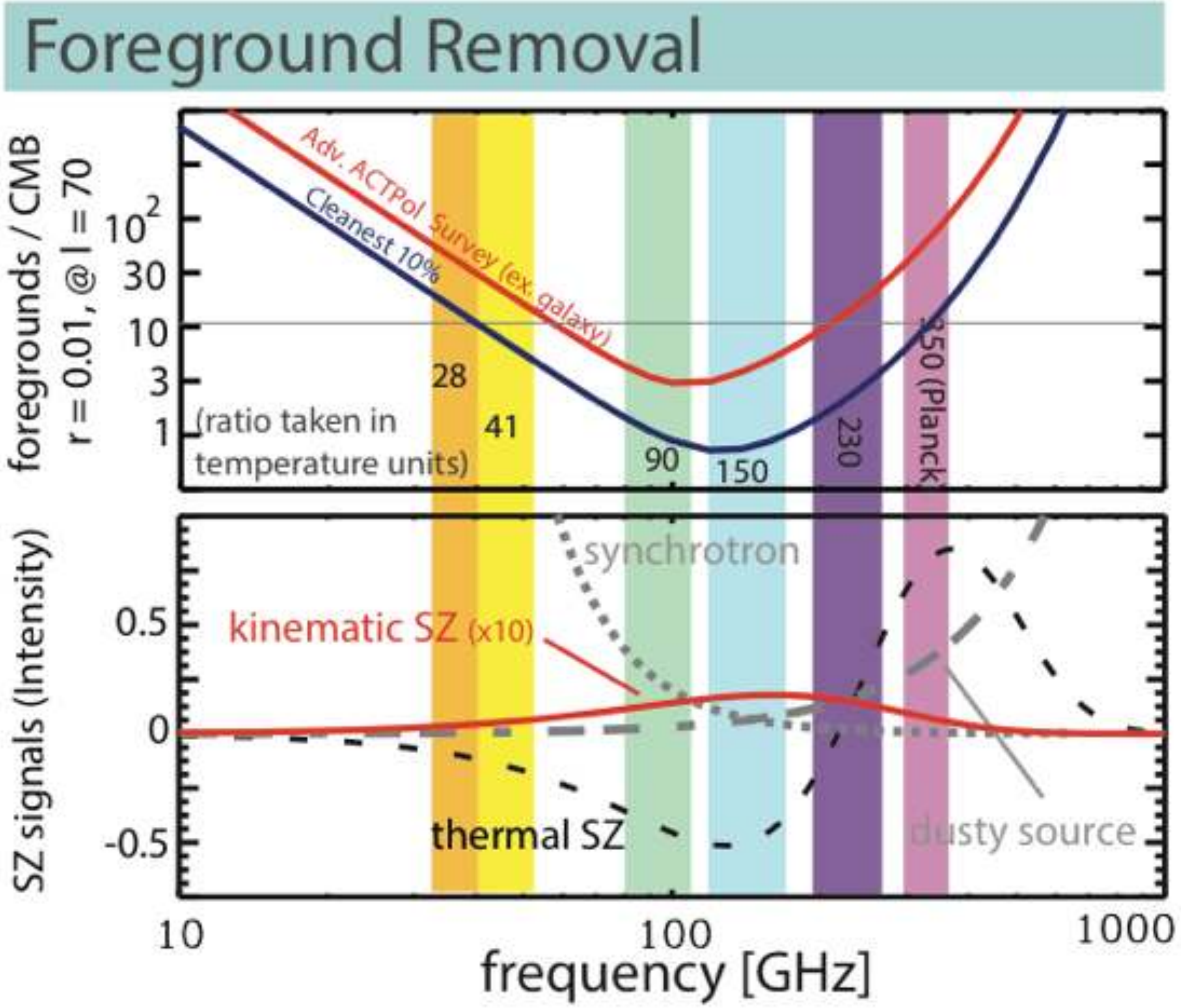
Advanced ACTPol

based on ACT technology



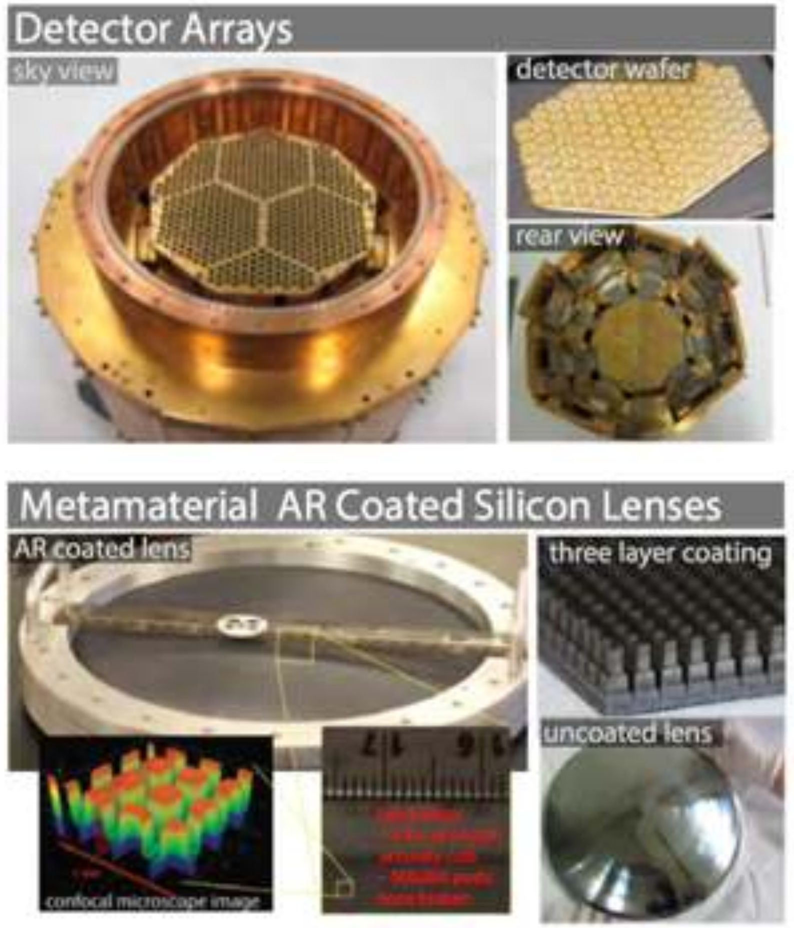
+ metamaterial Half Wave Plate

5 frequency bands



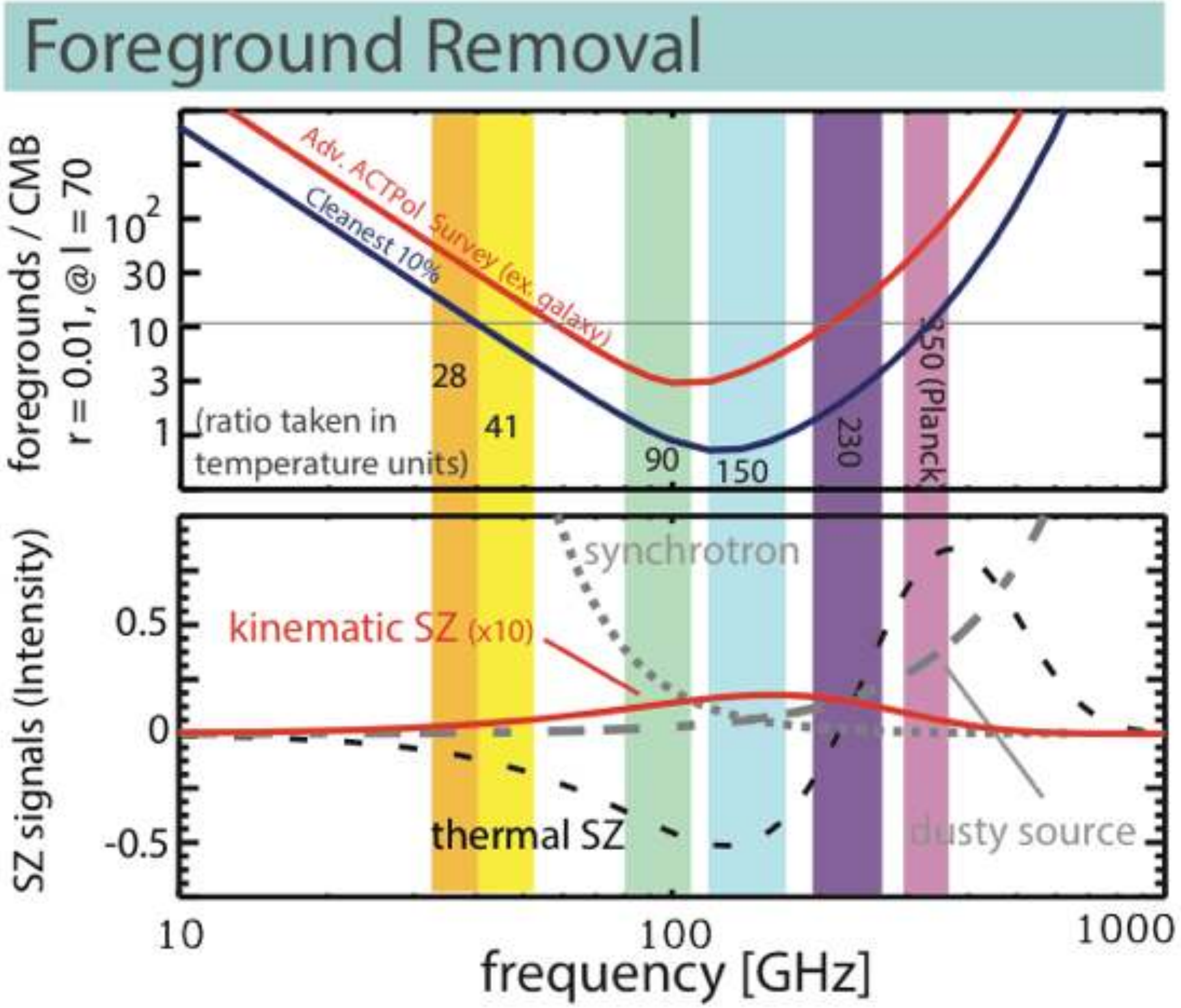
Advanced ACTPol

based on ACT technology



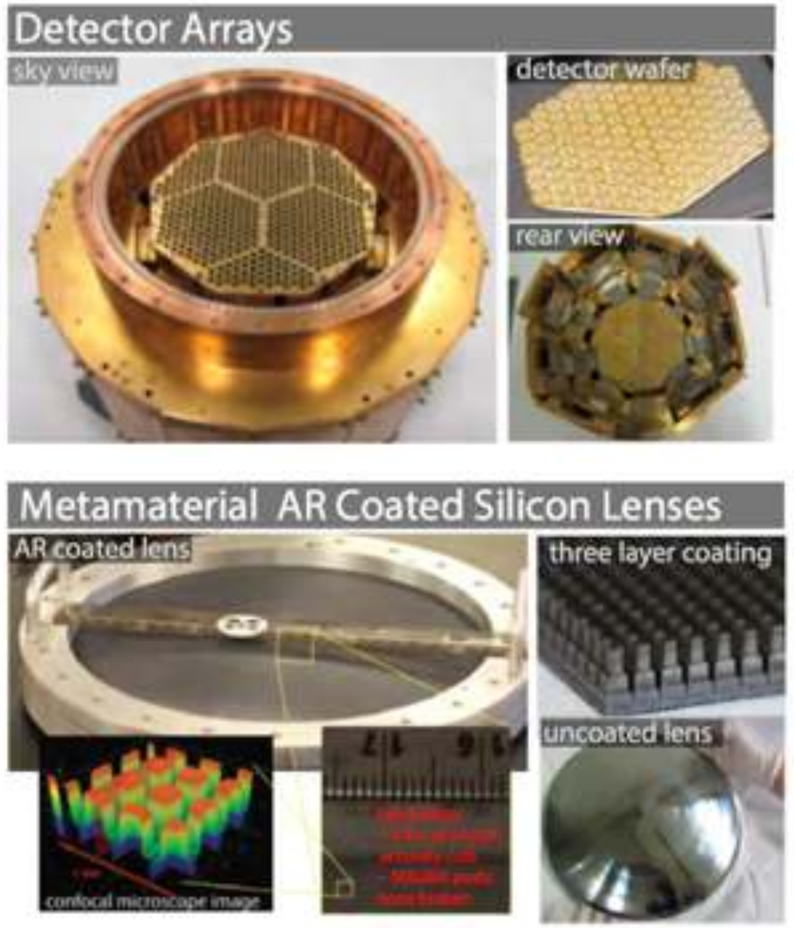
+ metamaterial
Half Wave Plate

5 frequency bands



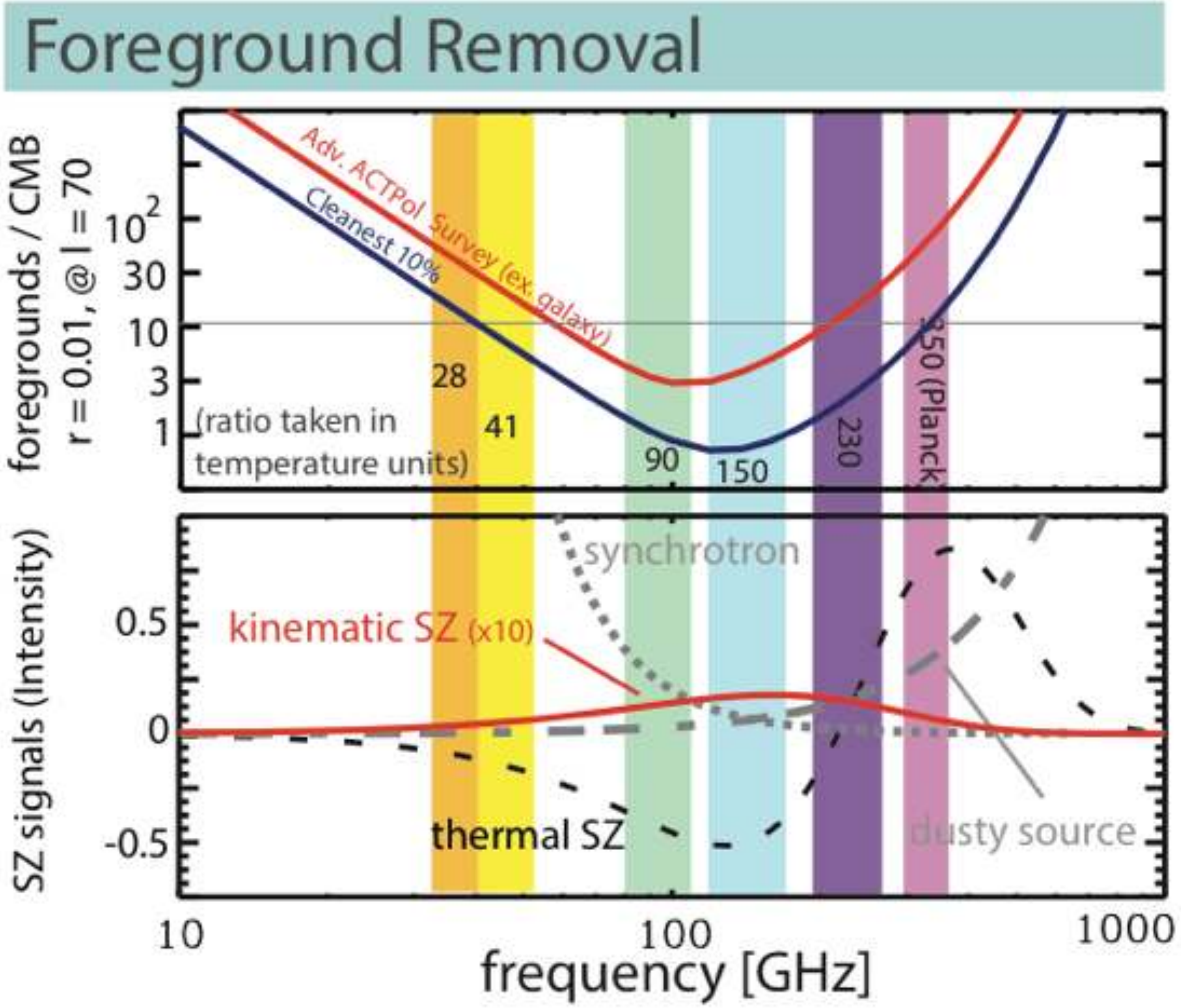
Advanced ACTPol

based on ACT technology



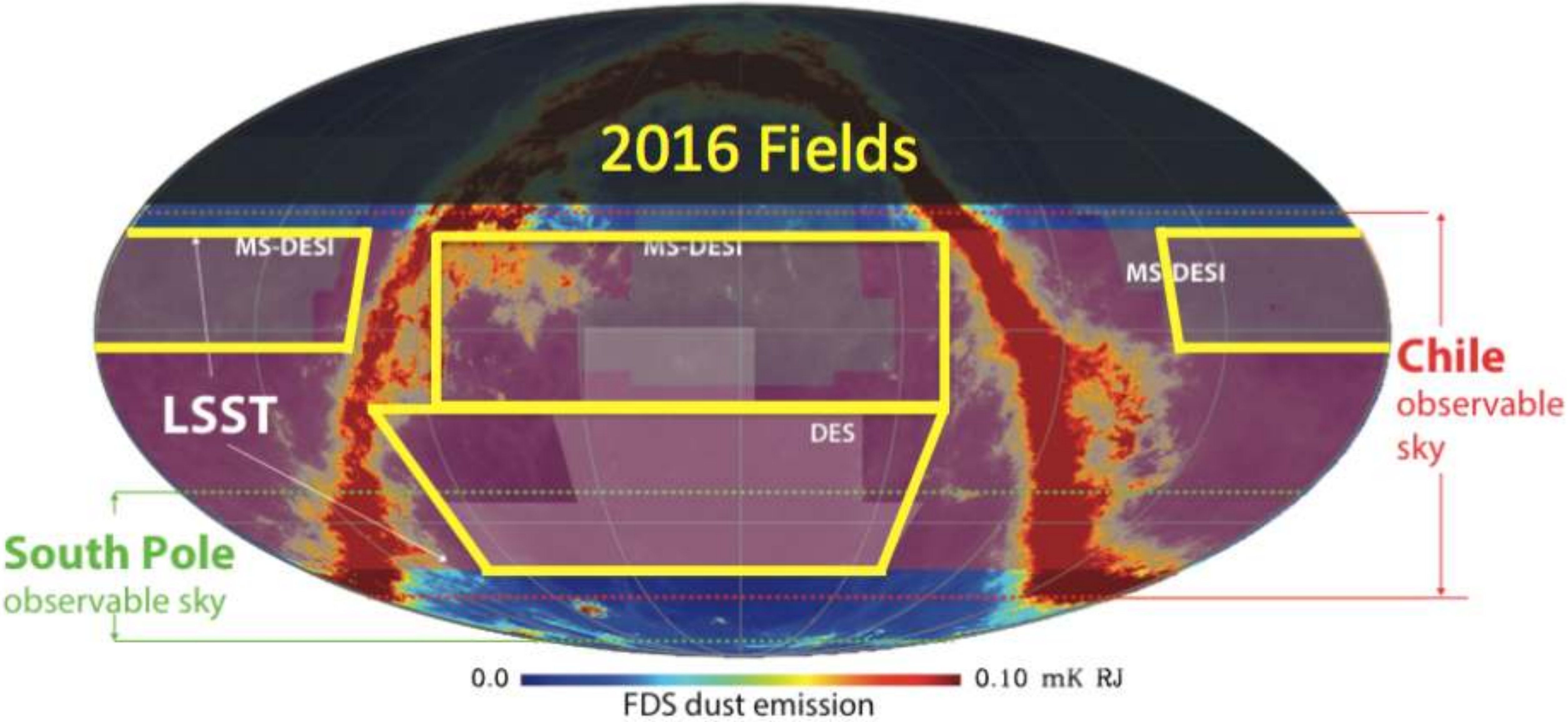
+ metamaterial Half Wave Plate

5 frequency bands



The Future Starts ~Now:

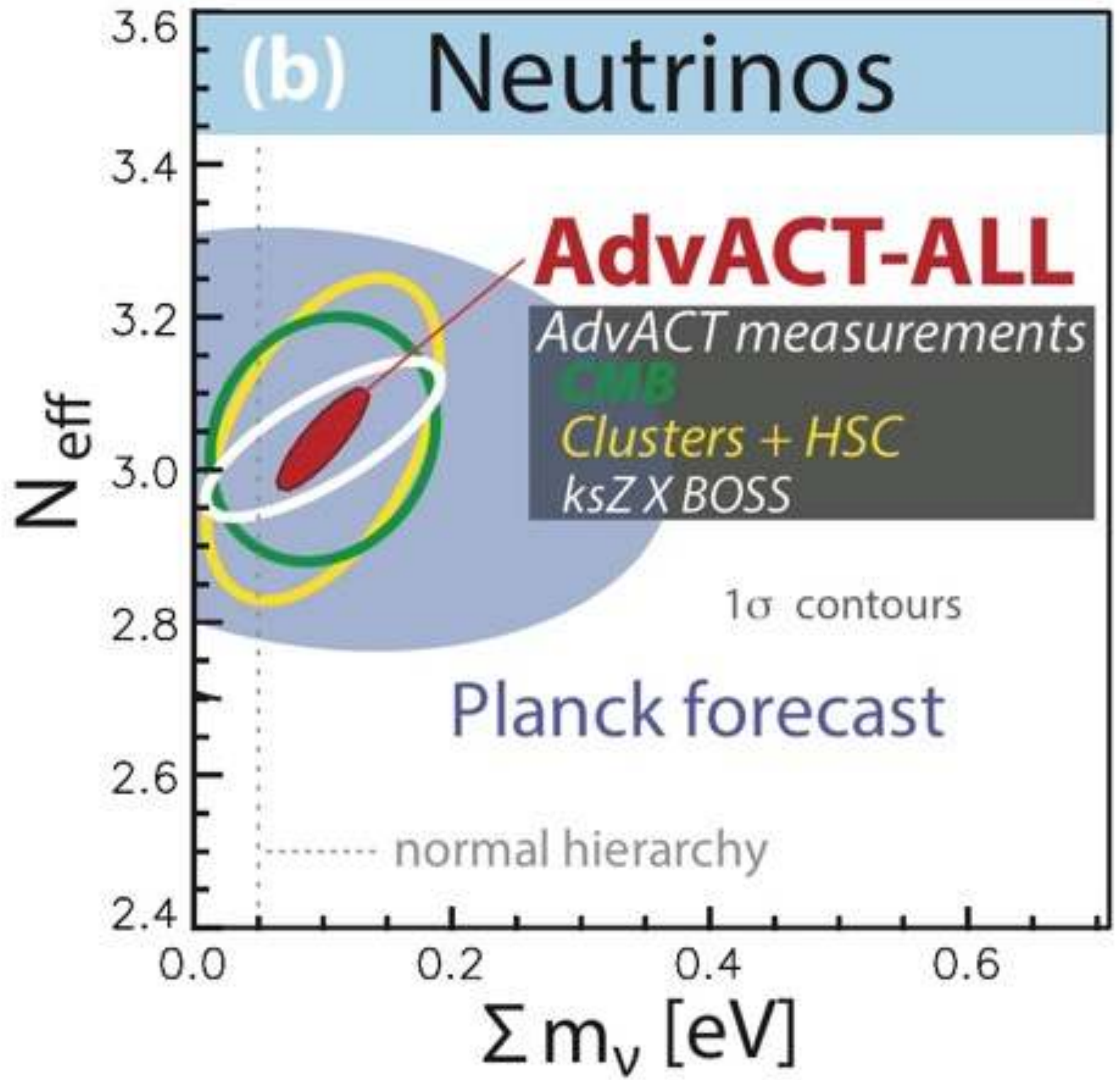
Advanced ACTPol Survey: 2016-2018

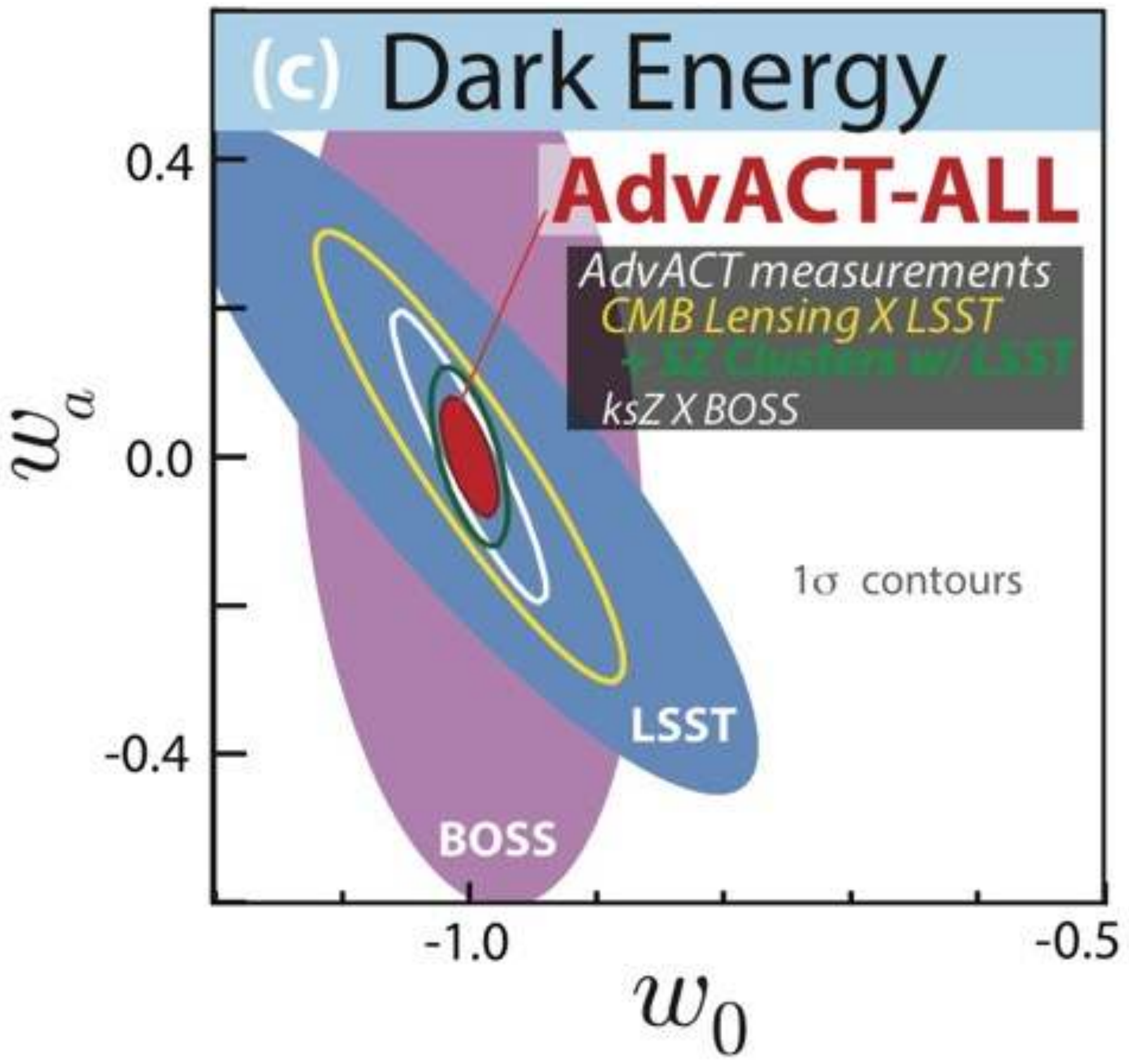


Cross Correlation Projections

Projected to improve Planck limit on Σm_ν by 10x!

Projected to improve LSST's DE FOM by 20x!





The Simons Observatory

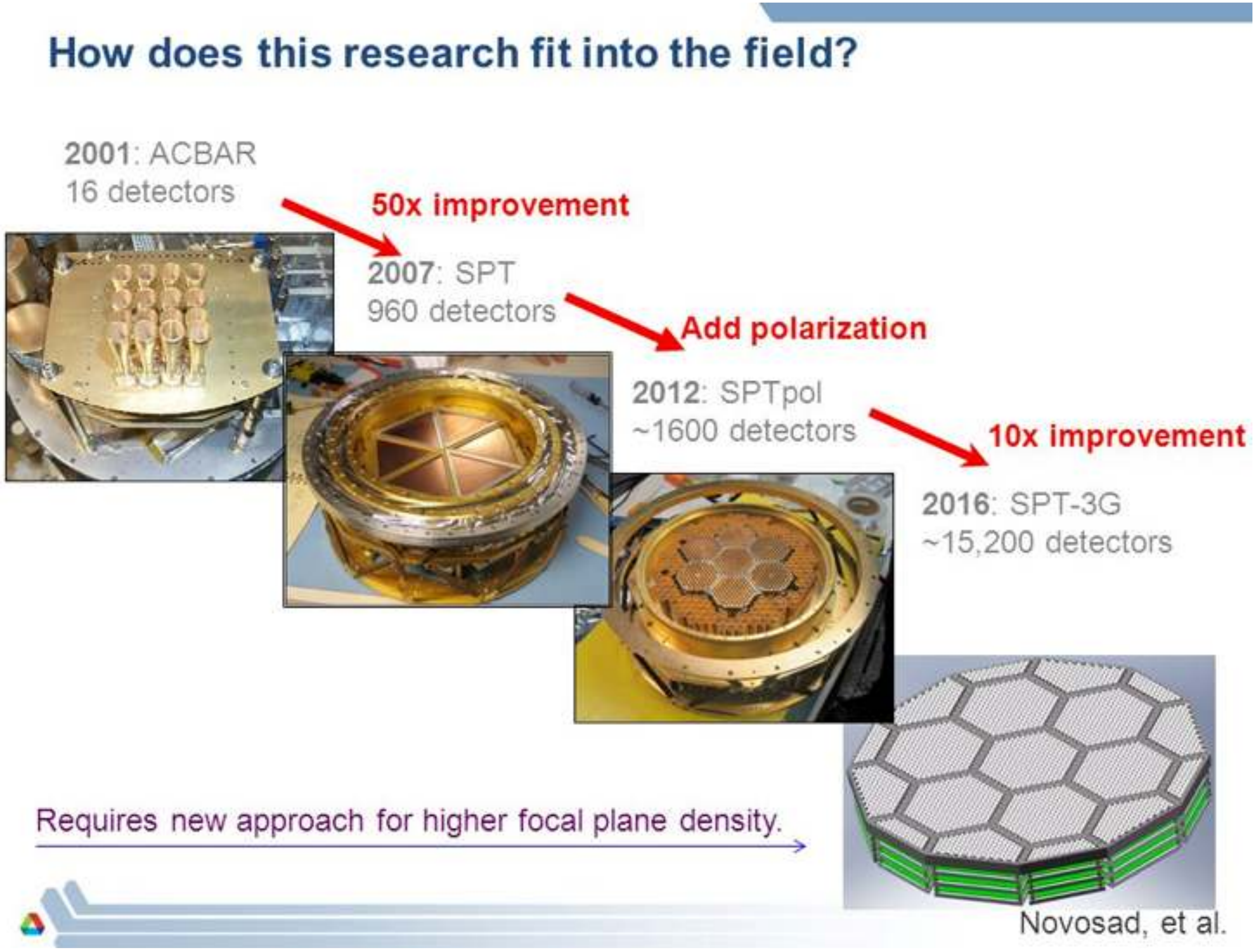
http://simonsobservatory.org

ALMA

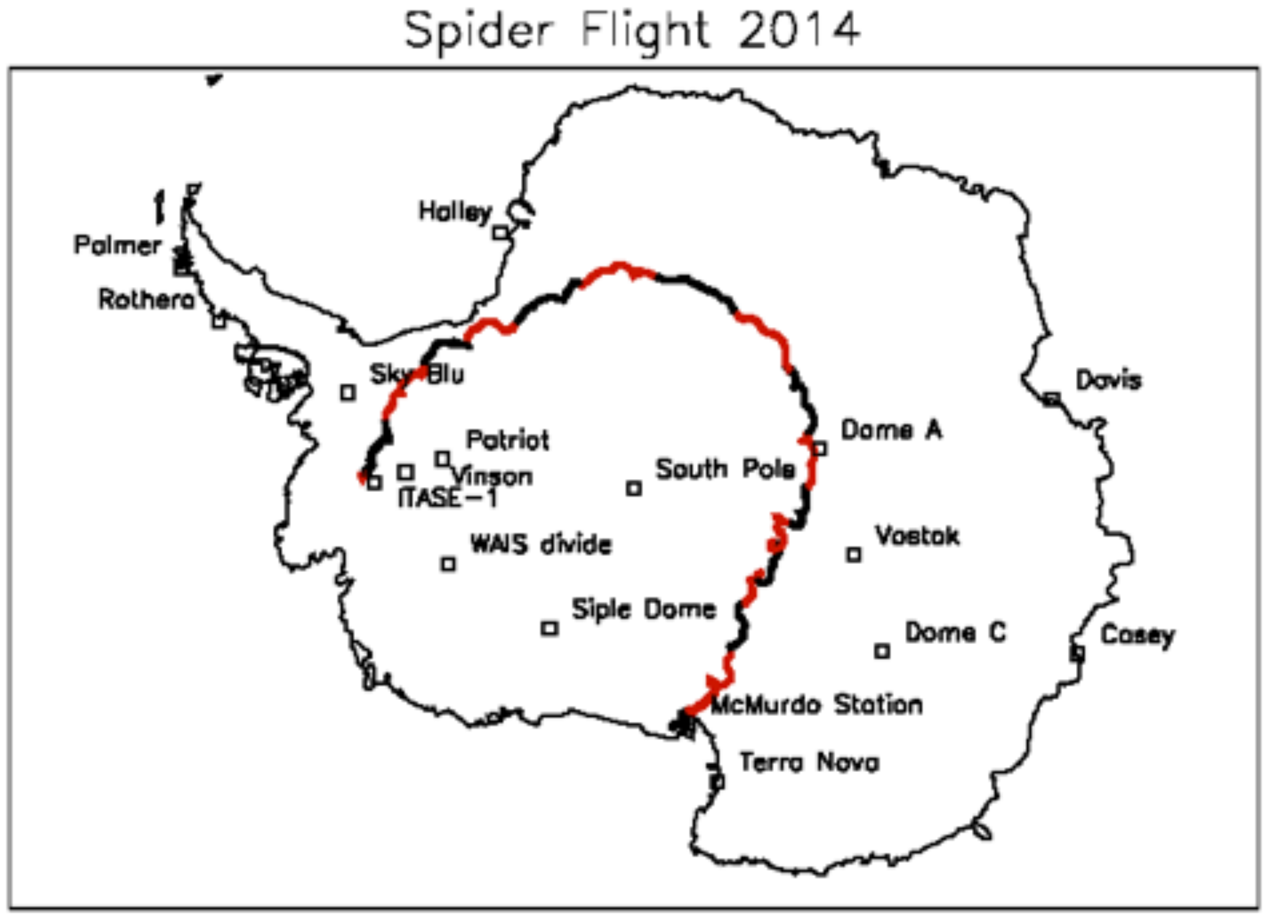
- A five year, $45M+ program to pursue key Cosmic Microwave Background science targets, and advance technology and infrastructure in preparation for CMB-S4.
- Merger of the ACT and POLARBEAR/Simons Array teams.
- Tentative plans include:
  - Major site infrastructure
  - Technology development (detectors, optics, cameras)
  - Demonstration of new high throughput telescopes.
  - CMB-S4 class receivers with partially filled focal planes.
  - Data analysis

BERKELEY LAB · CONICYT · HEISING-SIMONS FOUNDATION · NSF · Penn · PRINCETON UNIVERSITY · SIMONS FOUNDATION

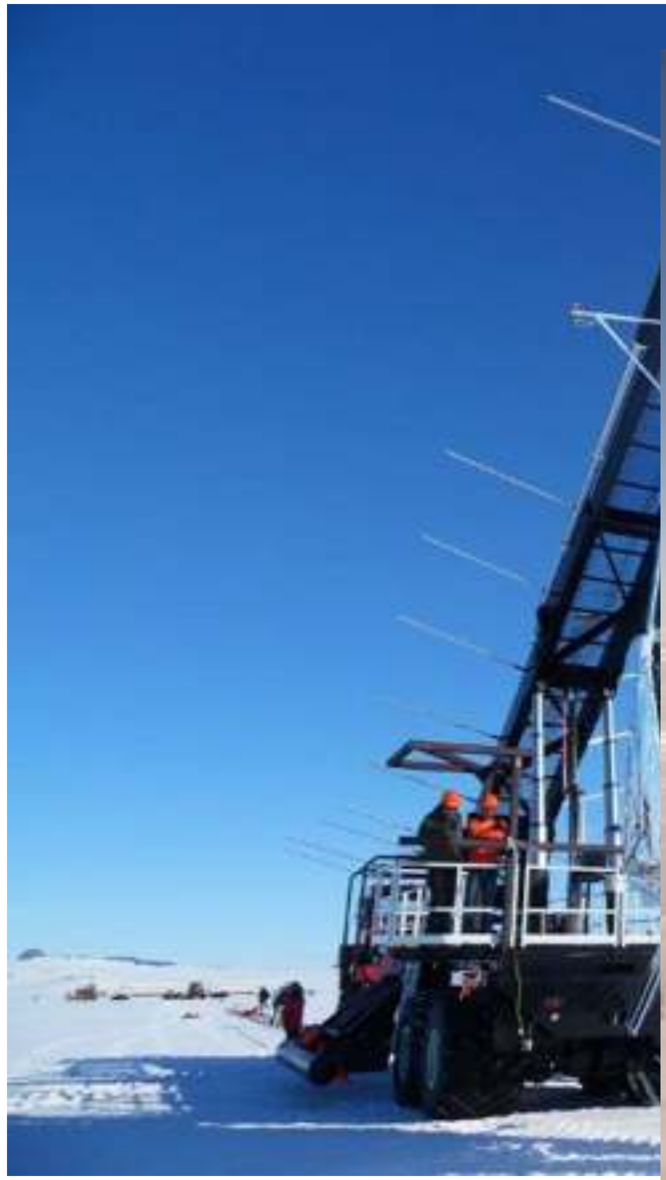
South Pole Telescope

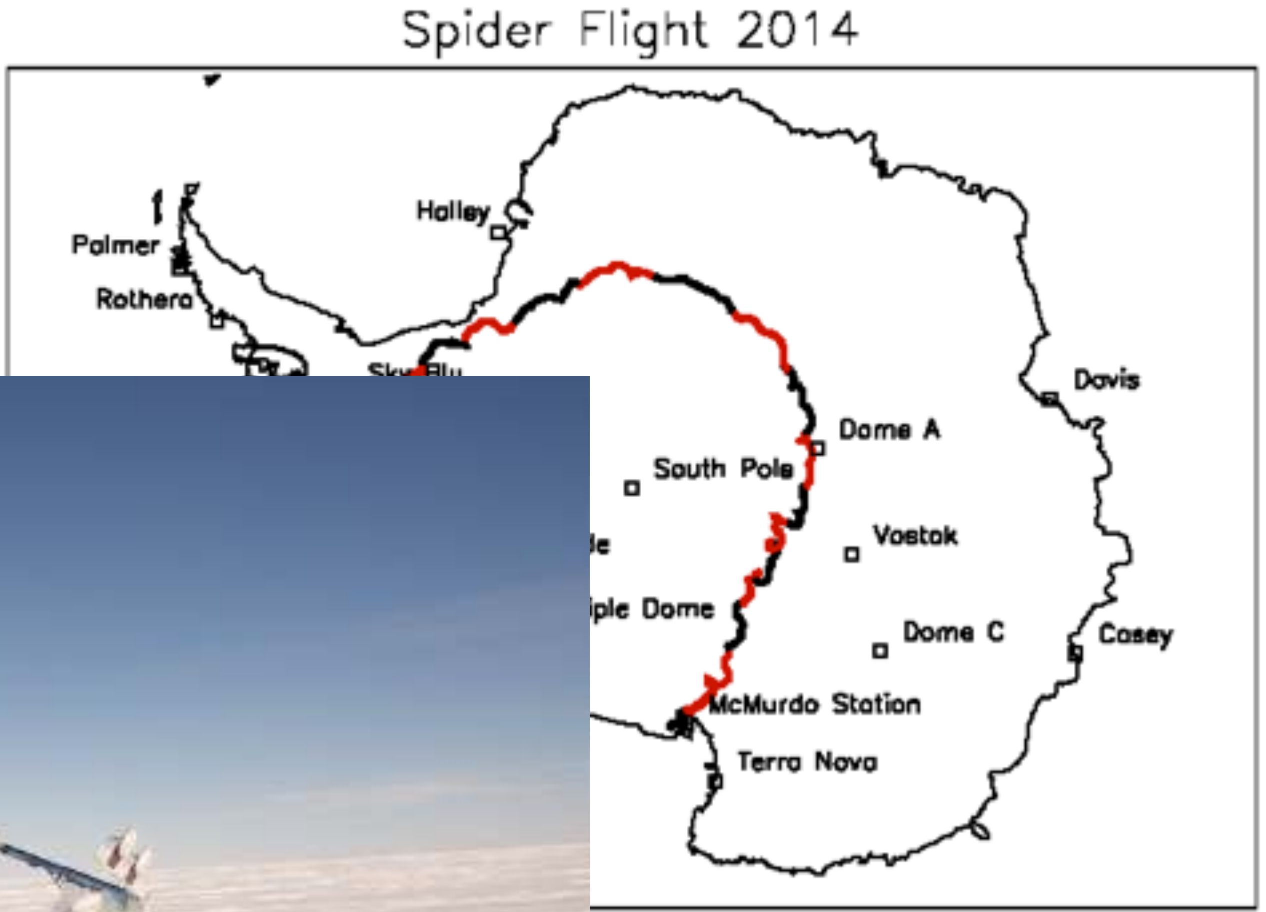


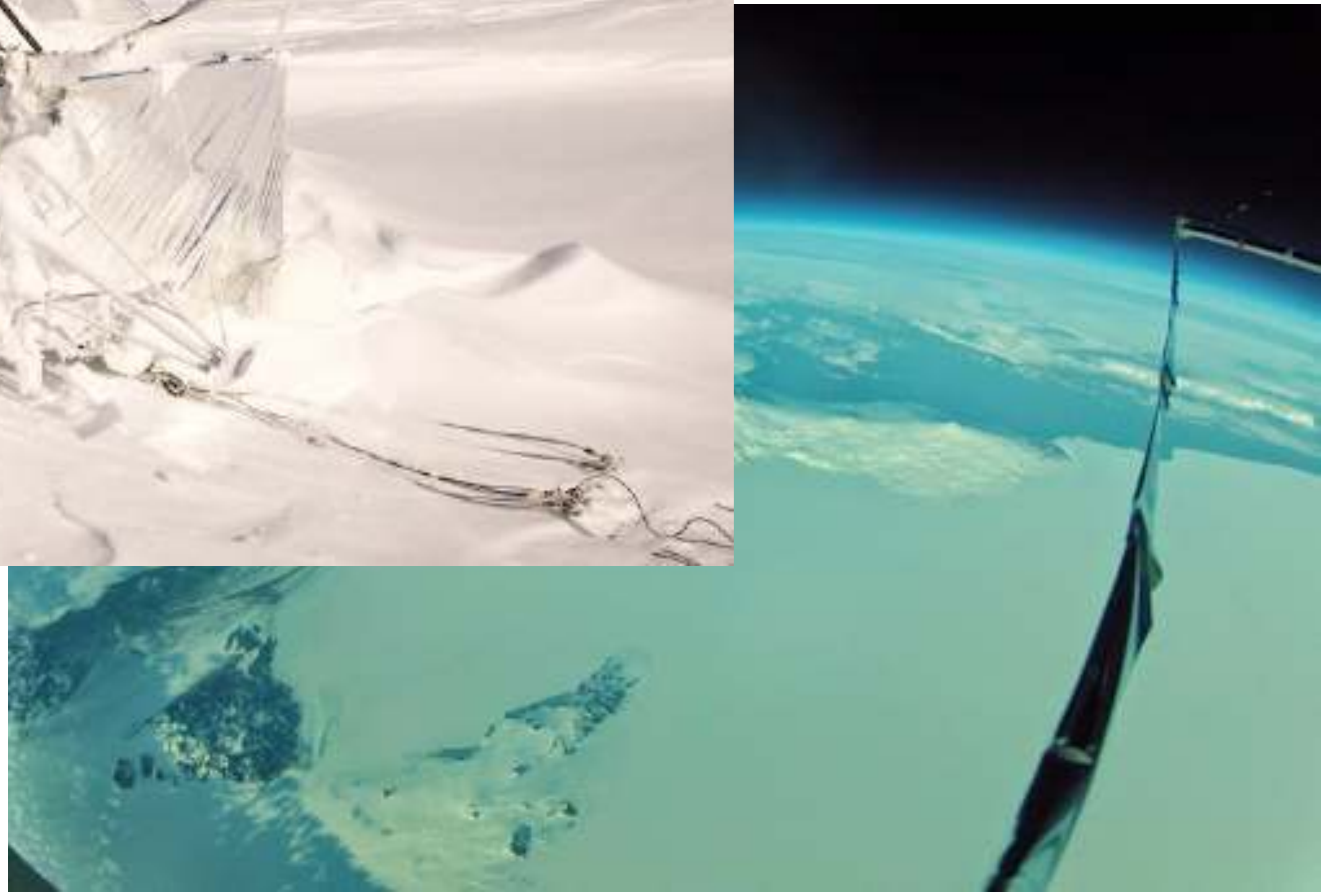
SPIDER - Searching for B-modes



SPIDER - Searching for B-modes







CMB Data Volumes

- SPIDER - few TB. Flight 2 will be the same
- Advanced ACT - 5500 detectors, 400 Hz = native rate of 700 GB/night
- SPT-3G - 15K detectors, 200 Hz = 1 TB/night
- Need heavy-duty computing to get data analyzed. Fortunately…

Hippo Cluster

UKZN Hippo Cluster w/ Richman Dlamini (HPC support staff)

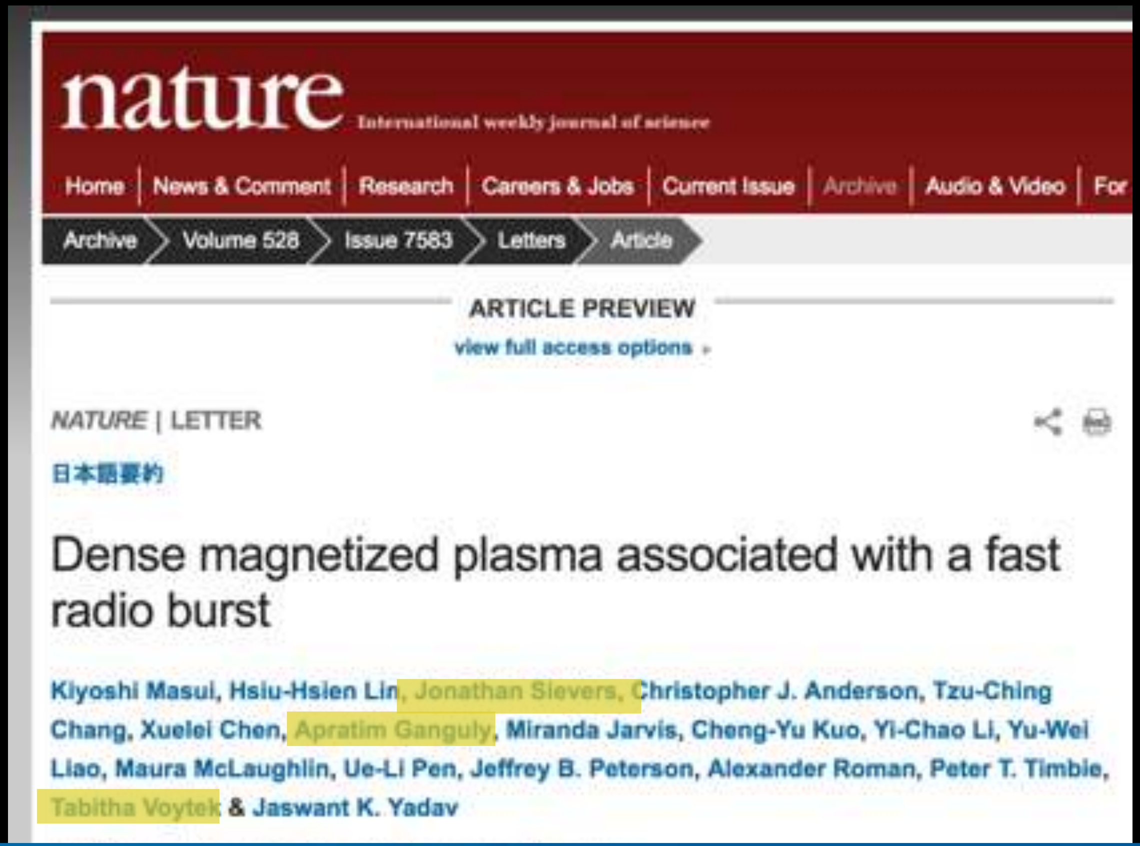
Hippo Hardware

- Compute: 1000 cores. Dual-socket E5-2680 v2 (10-core Ivy Bridge) 2.8 GHz CPUs, 64 GB/node RAM, 1 TB/node local disk.
- Storage: 2 lustre OSSs plus MDS, total 100 TB usable space.
- Gb networking to compute nodes, dual 10 Gb out of each OST. Full cluster can read from disk simultaneously at nearly full network speed.
- Management node (Dell box donated by ICS), one compute node set aside for login/devel.
- At installation, we believe hippo was largest HP system in southern hemisphere, fastest university cluster in SA.
- System has been very reliable - >99.5% uptime over last 12 months.

Hippo has significantly strengthened international collaborations

Collaborating institutions include Princeton, Caltech, Oxford, UToronto, APC-Paris, NRL&IUCAA (India), PUC (Chile), Stockholm, UChicago, UBC, WVU, many others…

Research Outputs - Fast Radio Bursts



nature International weekly journal of science

Home | News & Comment | Research | Careers & Jobs | Current Issue | Archive | Audio & Video | For

Archive > Volume 528 > Issue 7583 > Letters > Article

ARTICLE PREVIEW

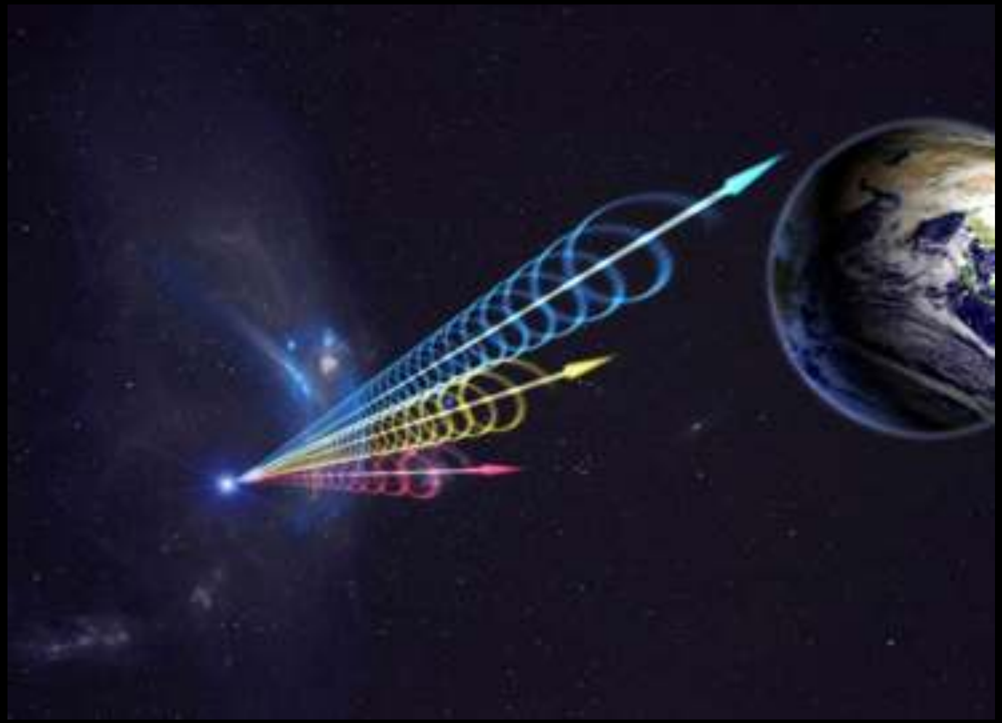
view full access options ›

NATURE | LETTER

日本語要約

Dense magnetized plasma associated with a fast radio burst

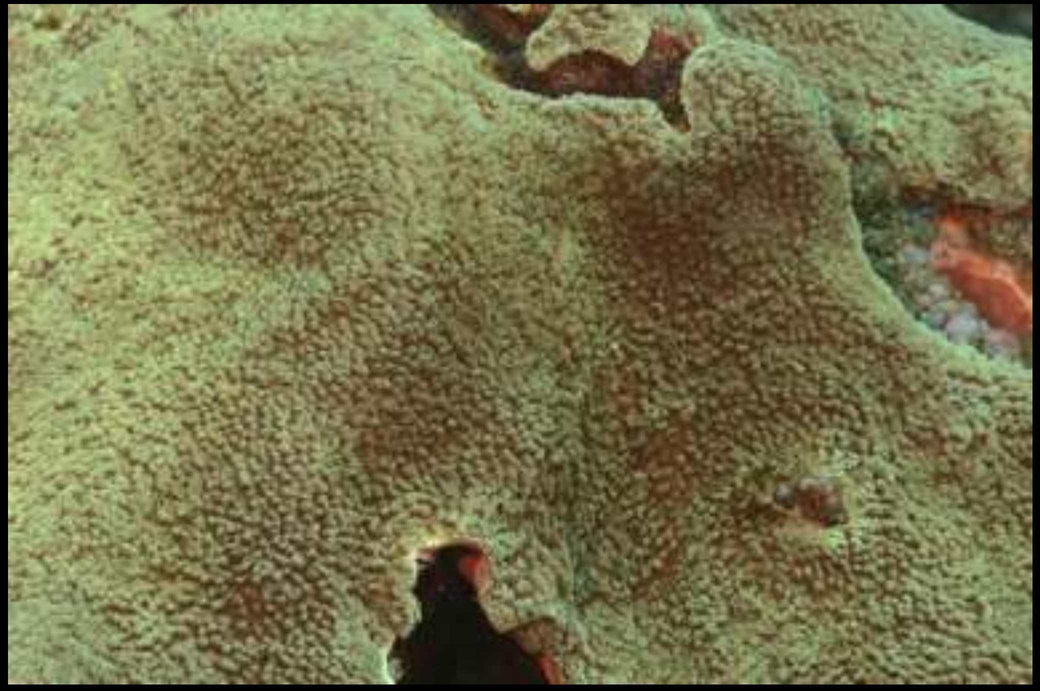
Kiyoshi Masui, Hsiu-Hsien Lin, Jonathan Sievers, Christopher J. Anderson, Tzu-Ching Chang, Xuelei Chen, Apratim Ganguly, Miranda Jarvis, Cheng-Yu Kuo, Yi-Chao Li, Yu-Wei Liao, Maura McLaughlin, Ue-Li Pen, Jeffrey B. Peterson, Alexander Roman, Peter T. Timbie, Tabitha Voytek & Jaswant K. Yadav



Fast Radio Burst analysis carried out on hippo. Results reported in Nature, written up over 100 times including Discover, Scientific American, Washington Post, many others…

RNA Transcriptomes

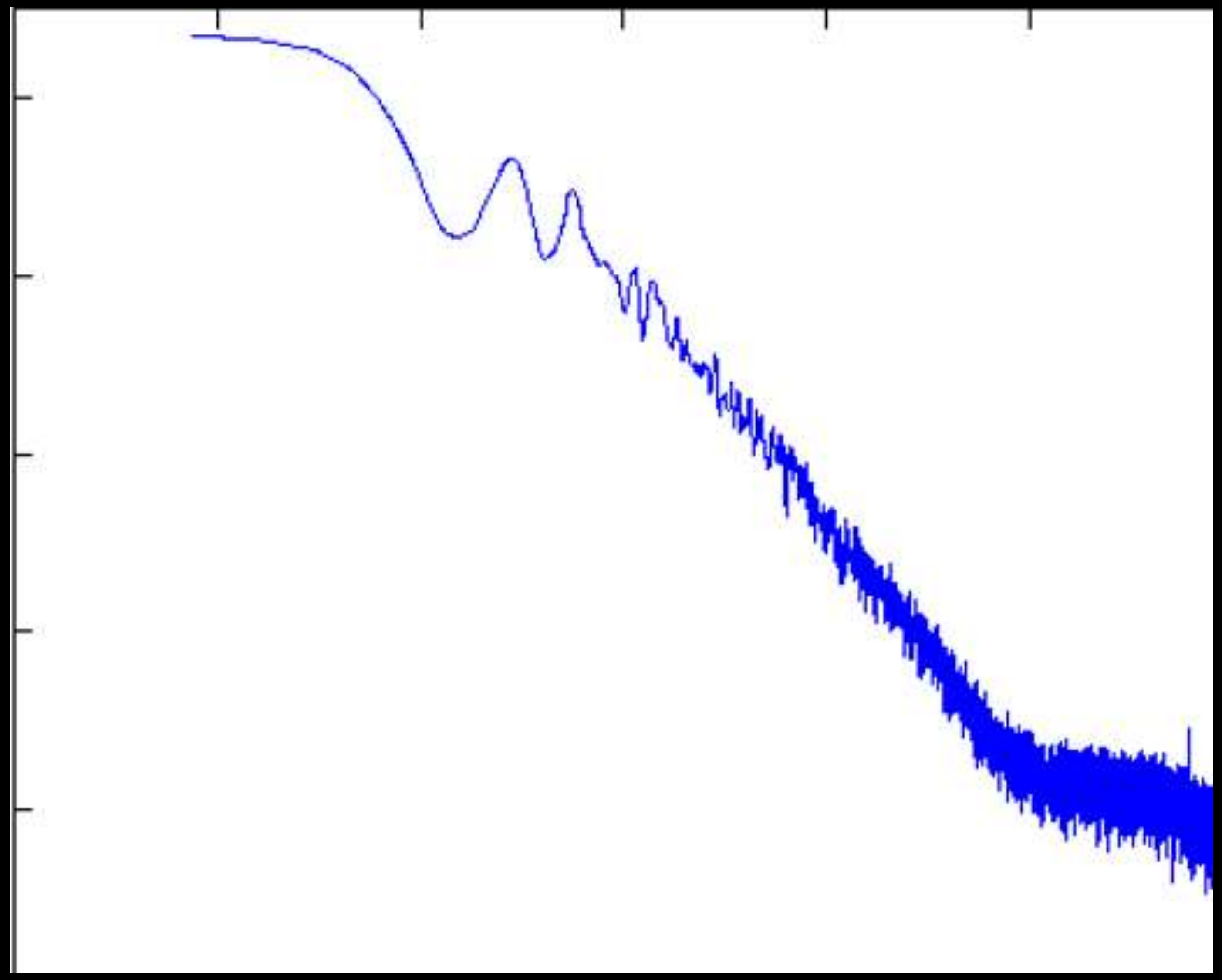
D. Pearton (SA Association for Marine Biological Research) has been assembling RNA transcriptomes of corals & fish, including dusky kob.
3 students to-date used results, will present in Hawaii next month with publication soon after.



Polymer Studies/Materials Science

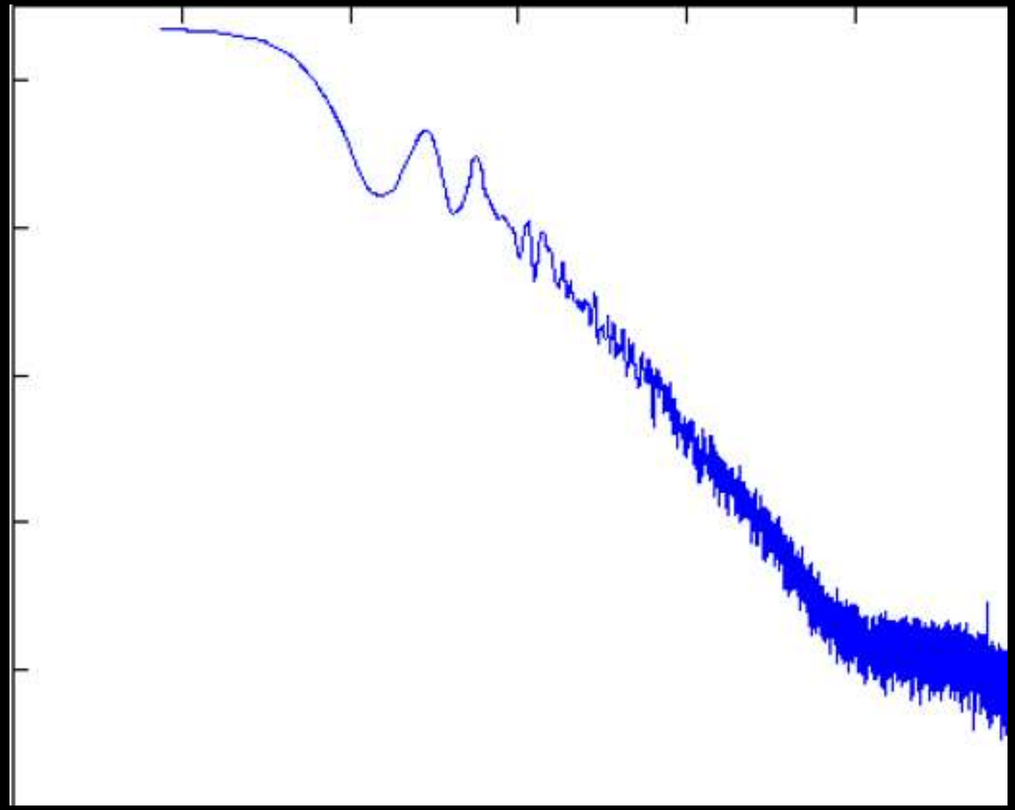
- G. Pellicane (PMB) & students/collaborators discovered unexpected behaviour of mixed linear/cyclic polymer chains.
- Work featured in PrE

Compression



- How much information is in contained in data?
- Shannon entropy - S (n_{bits})=$\sum -p_i \log_2(p_i)$
- Goal - write data in such a way to minimize S.
- If data are Gaussian, stationary, then power spectrum is complete statistical description.

Lossless Compression



- If power spectrum is known, can write down optimal predictor for data point using other data points.
- Residual from difference between data and its prediction is minimal variance estimator - compression of that gives best-possible results.
- Technique used widely, e.g. FLAC compressor for audio.
- Special but common case - power spectrum lowest at highest frequencies. Adjacent samples have nearly same expected value, so difference between adjacent samples has sqrt(2) times best possible estimate.
- Compressing difference gets within 0.5 bits/sample of optimal. Works if you find any variable where PS small at high-k.

Lossy Compression

- How much data should we keep? How do we know what to throw away?
- Think of data as a low-frequency/low-information signal plus white noise component. White noise takes up all the space.
- Keeping n bits per sample of white noise equivalent to adding second noisy component so data + component gives results quantized to our n bits.
- S.D. of noisy signal is ~$0.25/2^{nbit}$ relative to base σ. Variance goes like that squared, or $2^{-2(nbit+2)}$.
- 3 bits - increase in variance is 10^{-3}, or about 12 hours/year. 6 bits is 10 minutes/year

Are You Short on Disk Space?

- If you have a lot of data and are hurting on storage, simple procedure may help.
- Find a direction in your data where difference between adjacent samples (mostly) looks like white noise.
- Quantize difference at a few bits relative to σ (ignoring outliers).
- gzip/bzip2 the quantized difference.
- This procedure should apply to a wide variety of astronomical datasets.

Summary

- South Africa very attractive place for all sorts of radio experiments
- With lots of experiments comes lots of data.
- Computational facilities growing (c.f. Shazrene, Russ Taylor)
- You may wish to compress your data…