
# Delensing, Lensing, and Neural Networks — a perspective from the CMB

W.L. Kimmy Wu
KICP Fellow
University of Chicago
Kavli Institute for Cosmological Physics

Dec 14, 2018
LBL INPA Seminar

# Outline

- Background
- Delensing of BICEP/Keck B-mode maps, with Planck + SPTpol E-mode maps
- Lensing reconstruction from SPTpol
- Convolutional Neural Networks on CMB lensing
- Summary and outlook

South Pole Telescope


Alessandro Manzotti
IAP

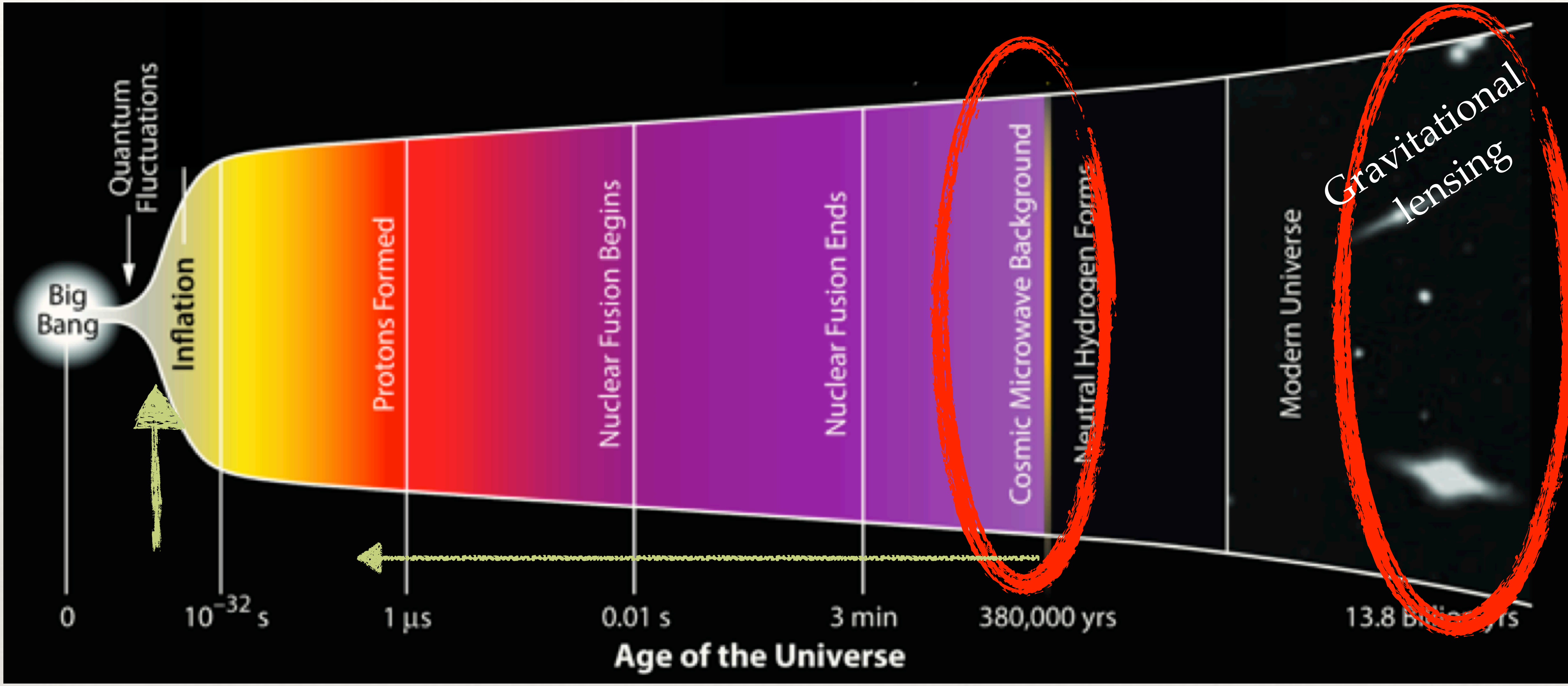
Monica Mocanu
UiO, Norway


Joao Caldeira
Fermilab



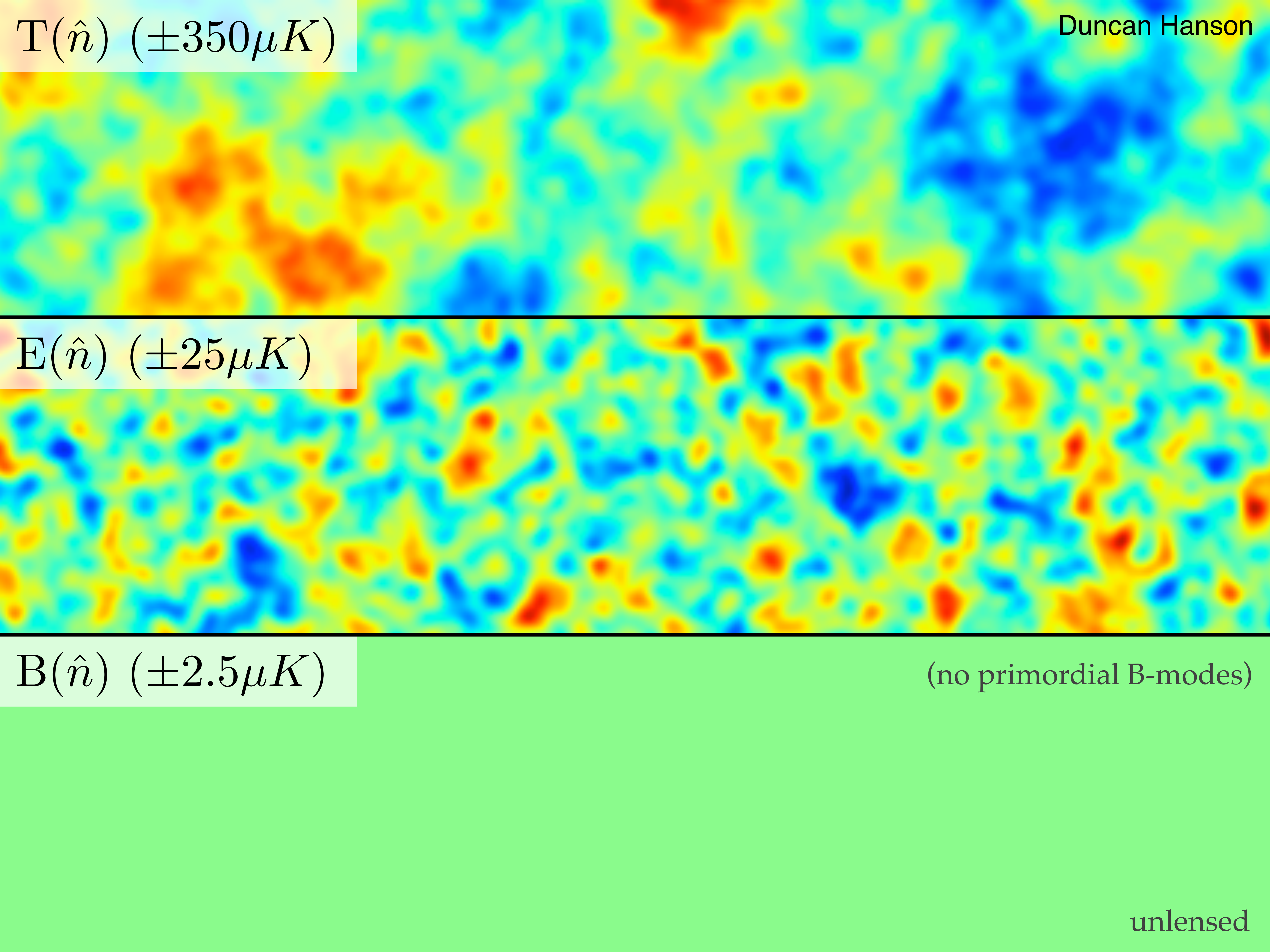
BICEP / Keck Array

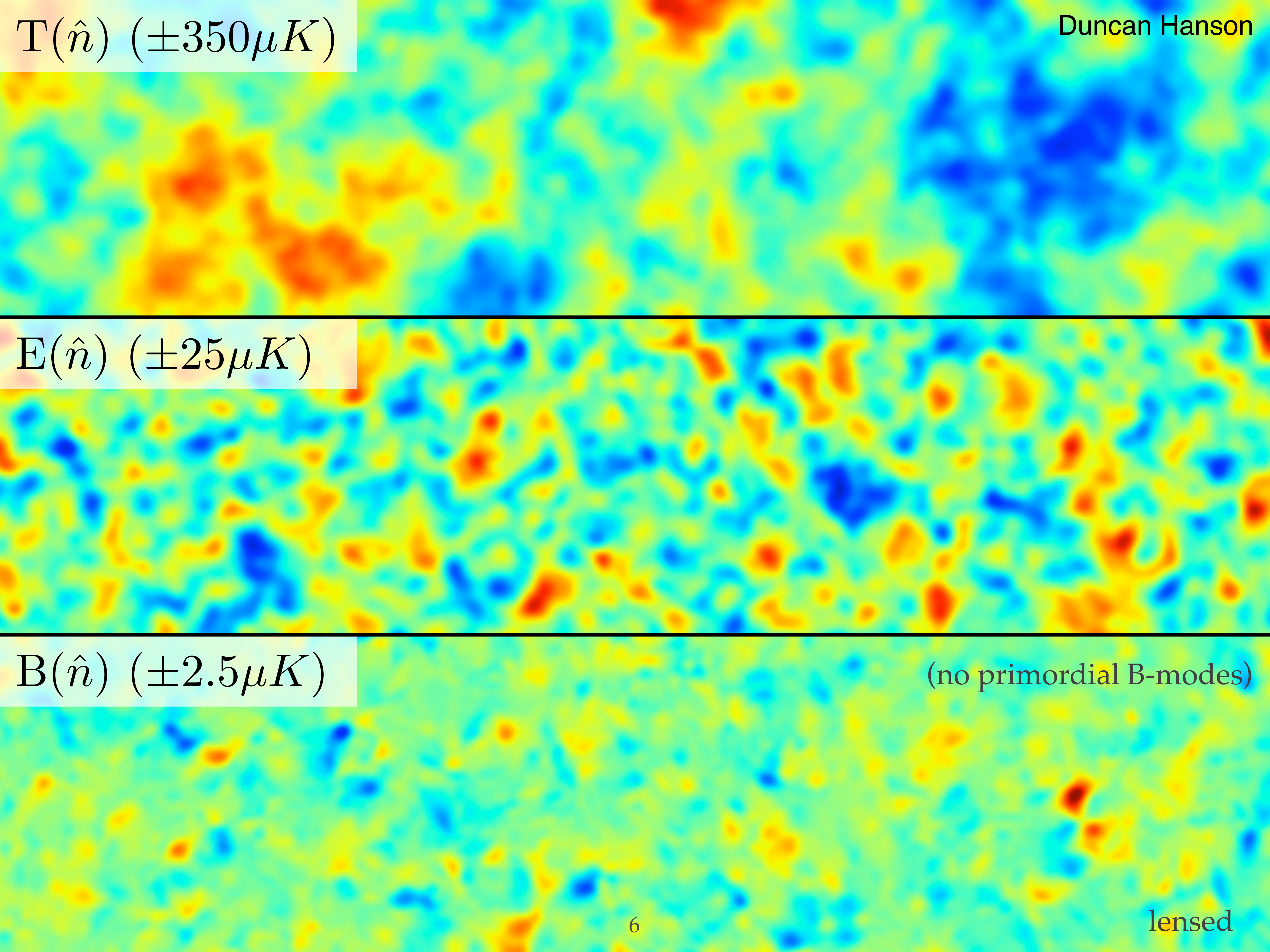
# The CMB is sensitive to our standard parameters and more!

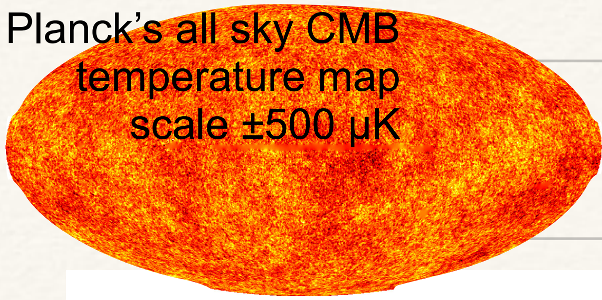



Background Figure: BICEP/Keck collaboration

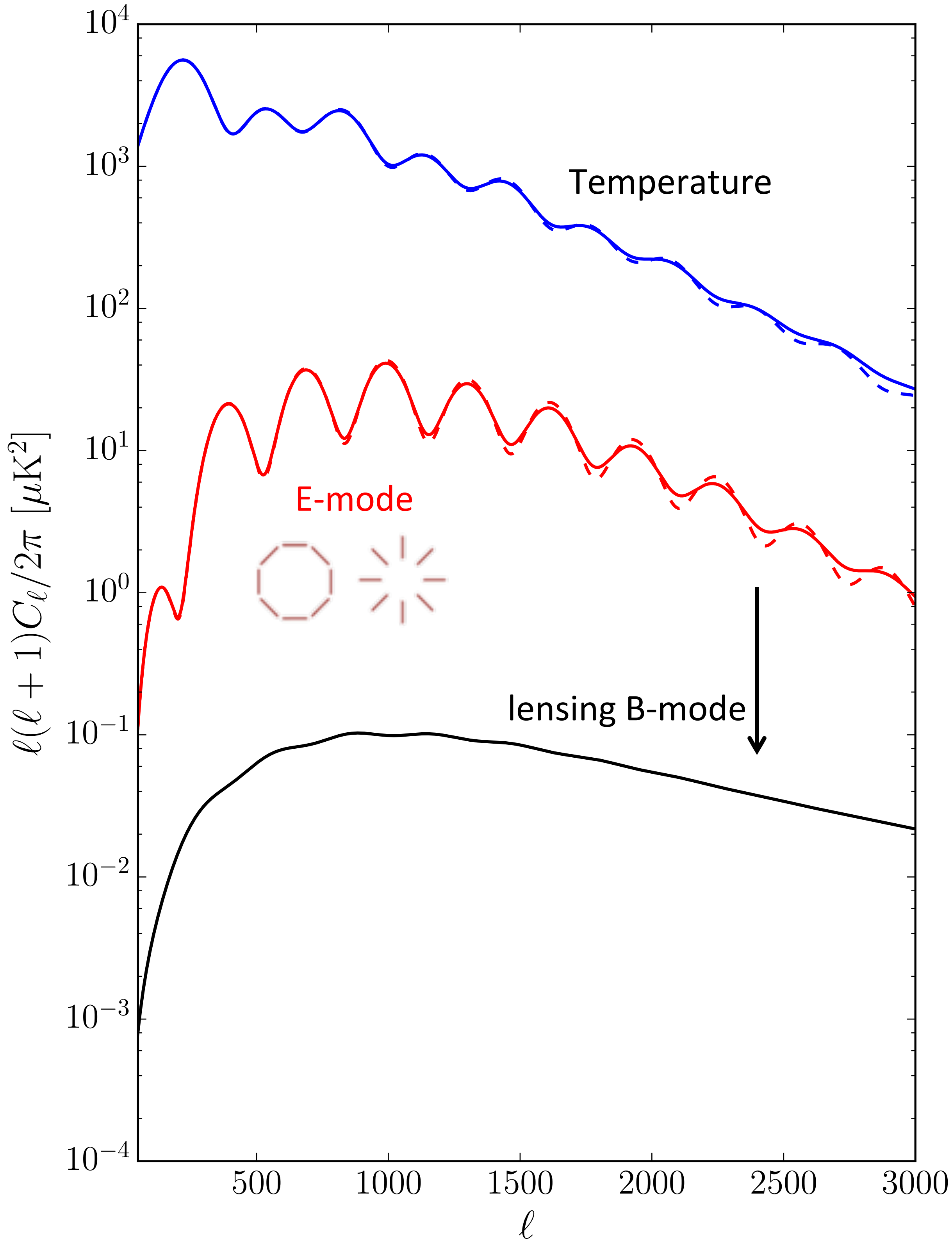
- CMB at recombination is sensitive to photon, matter, and baryon densities, initial conditions (A_s, n_s)
- And more! Primordial gravitational waves, dark radiation, …

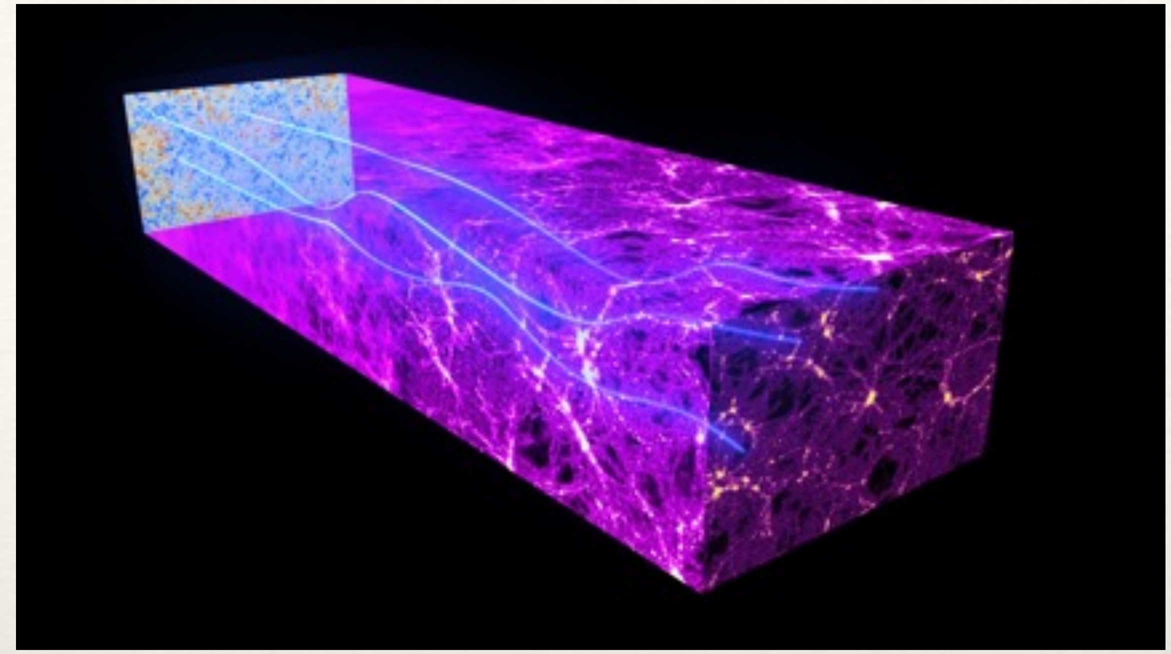

$T(\hat{n})$ $(\pm 350\mu K)$
Duncan Hanson
$E(\hat{n})$ $(\pm 25\mu K)$
$B(\hat{n})$ $(\pm 2.5\mu K)$
(no primordial B-modes)
unlensed

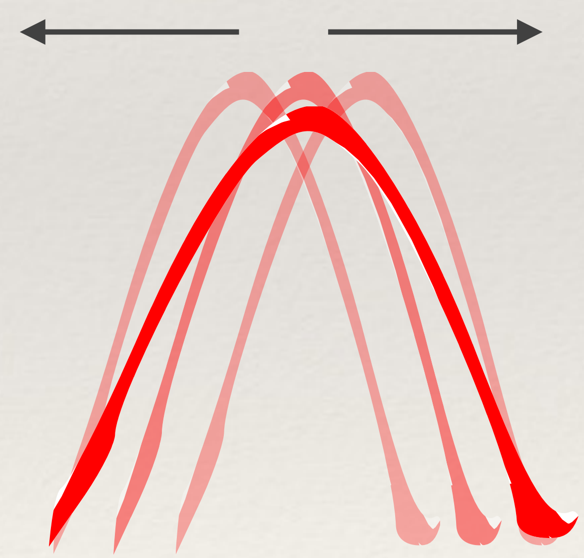

$T(\hat{n})$ $(\pm 350\mu K)$
Duncan Hanson
$E(\hat{n})$ $(\pm 25\mu K)$
$B(\hat{n})$ $(\pm 2.5\mu K)$
(no primordial B-modes)
lensed



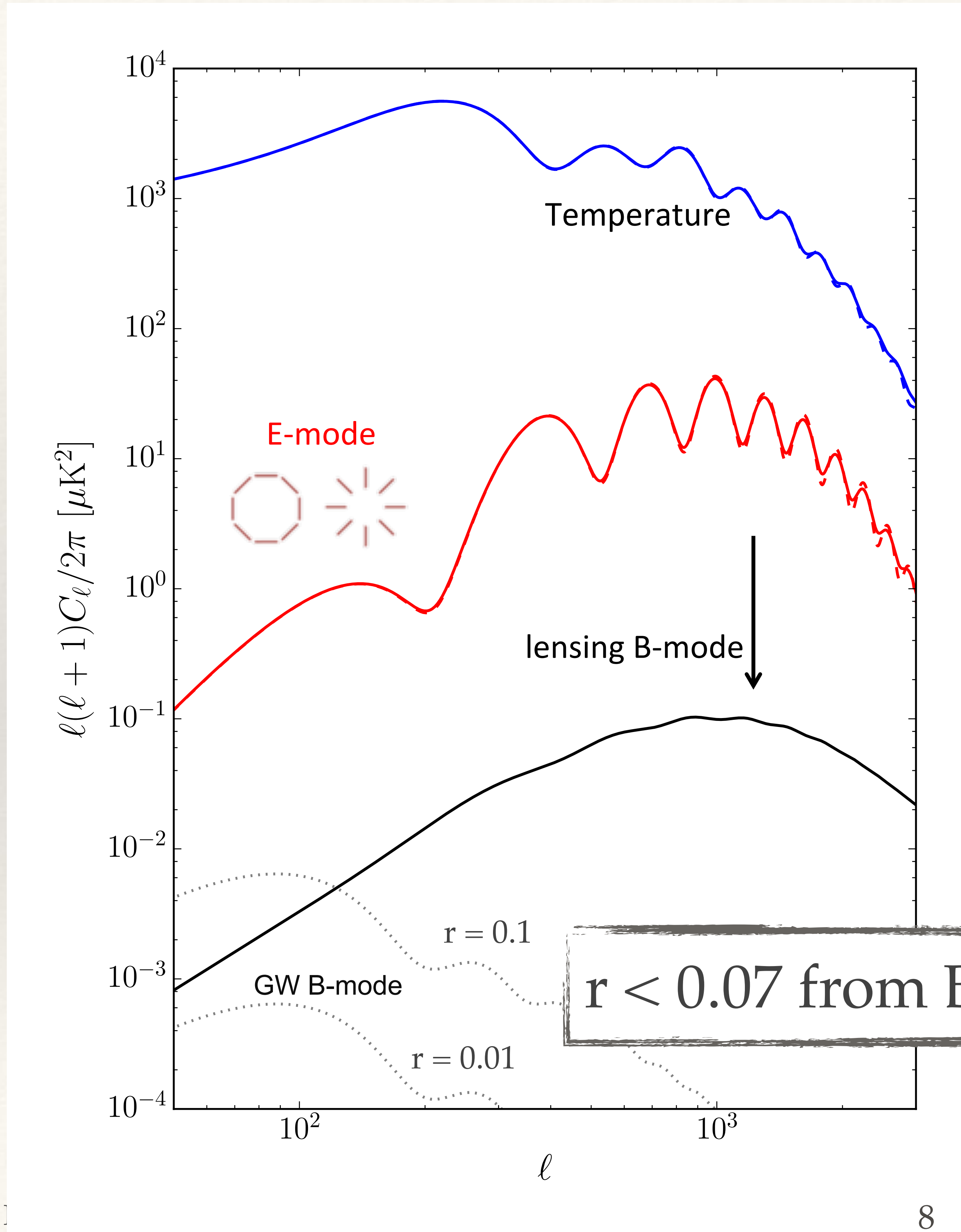

# Effect of lensing on the spectra



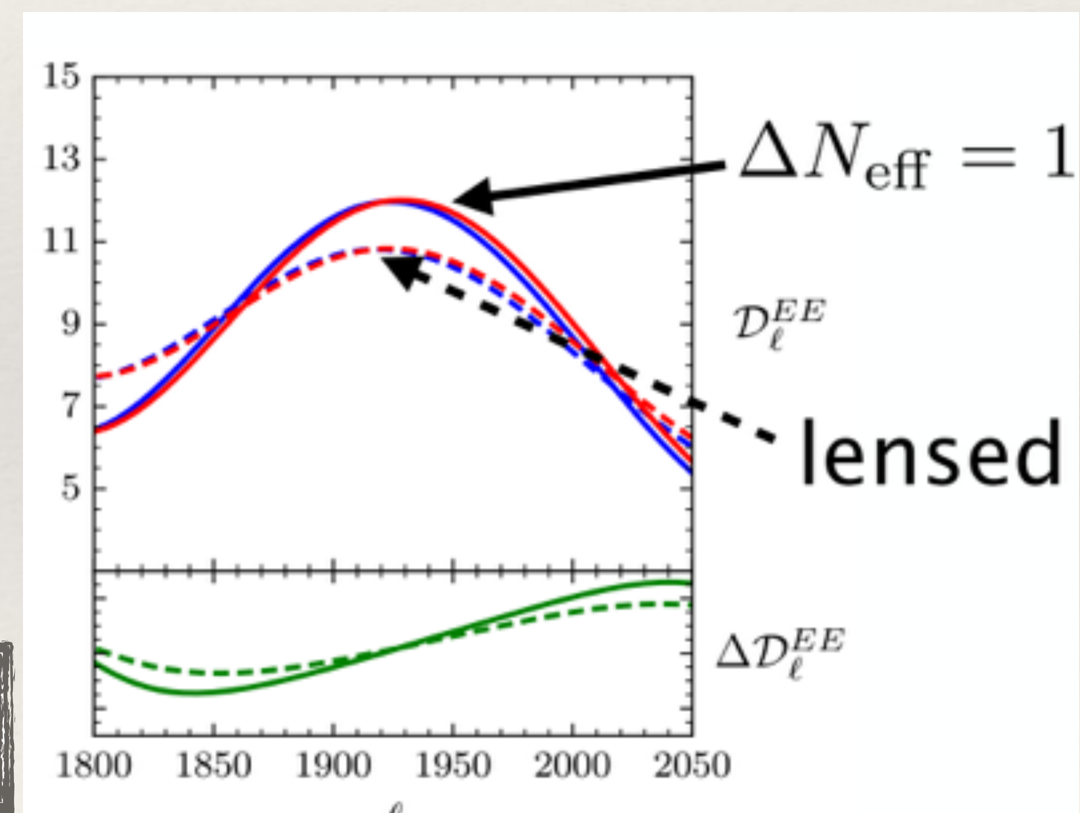


# Are the effect of lensing limiting our parameter constraints?



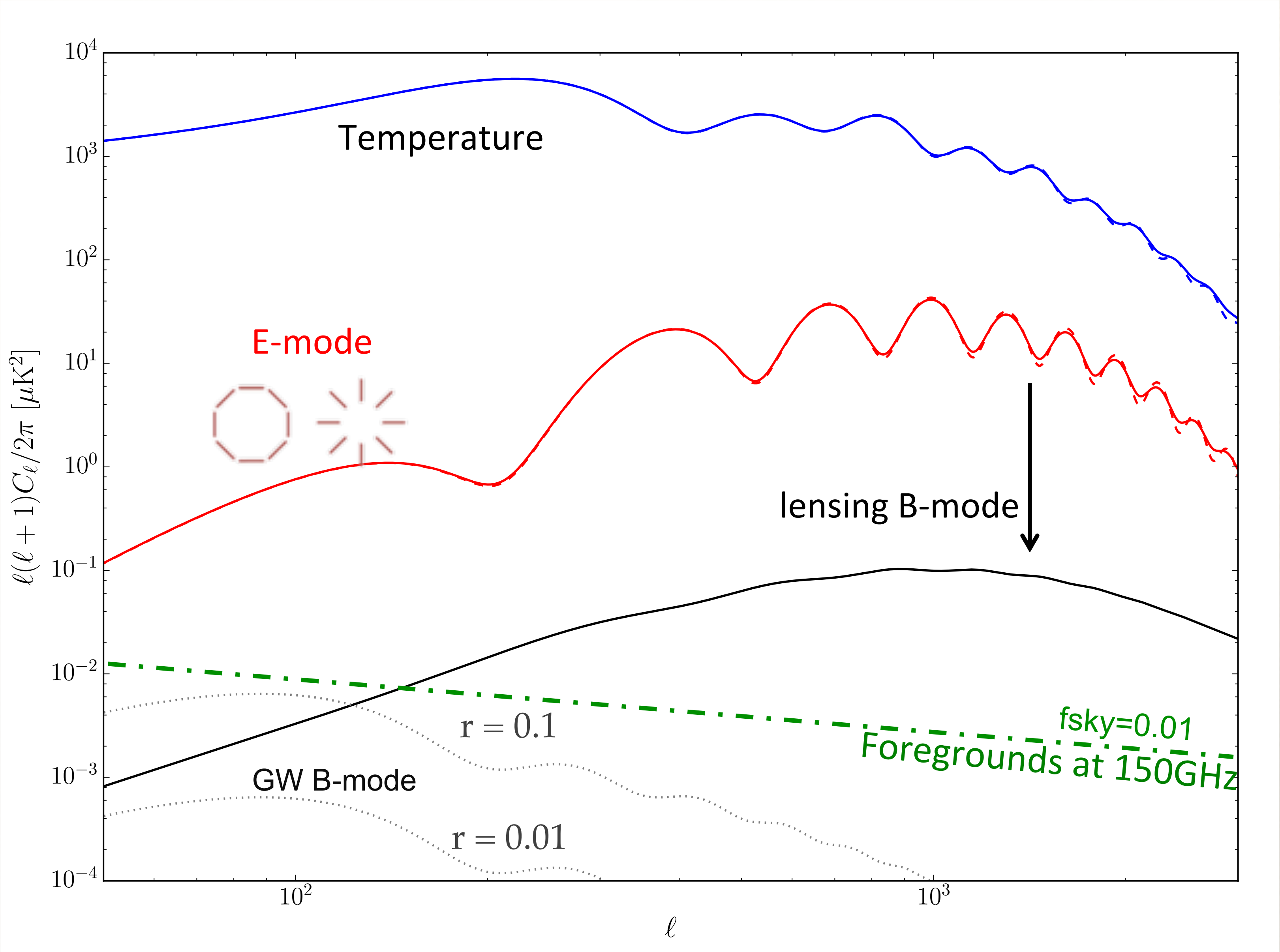

r < 0.07 from BK15

- For primordial gravitational wave (in ~2-3 years)
- For $N_{eff}$ (in ~5+? years)

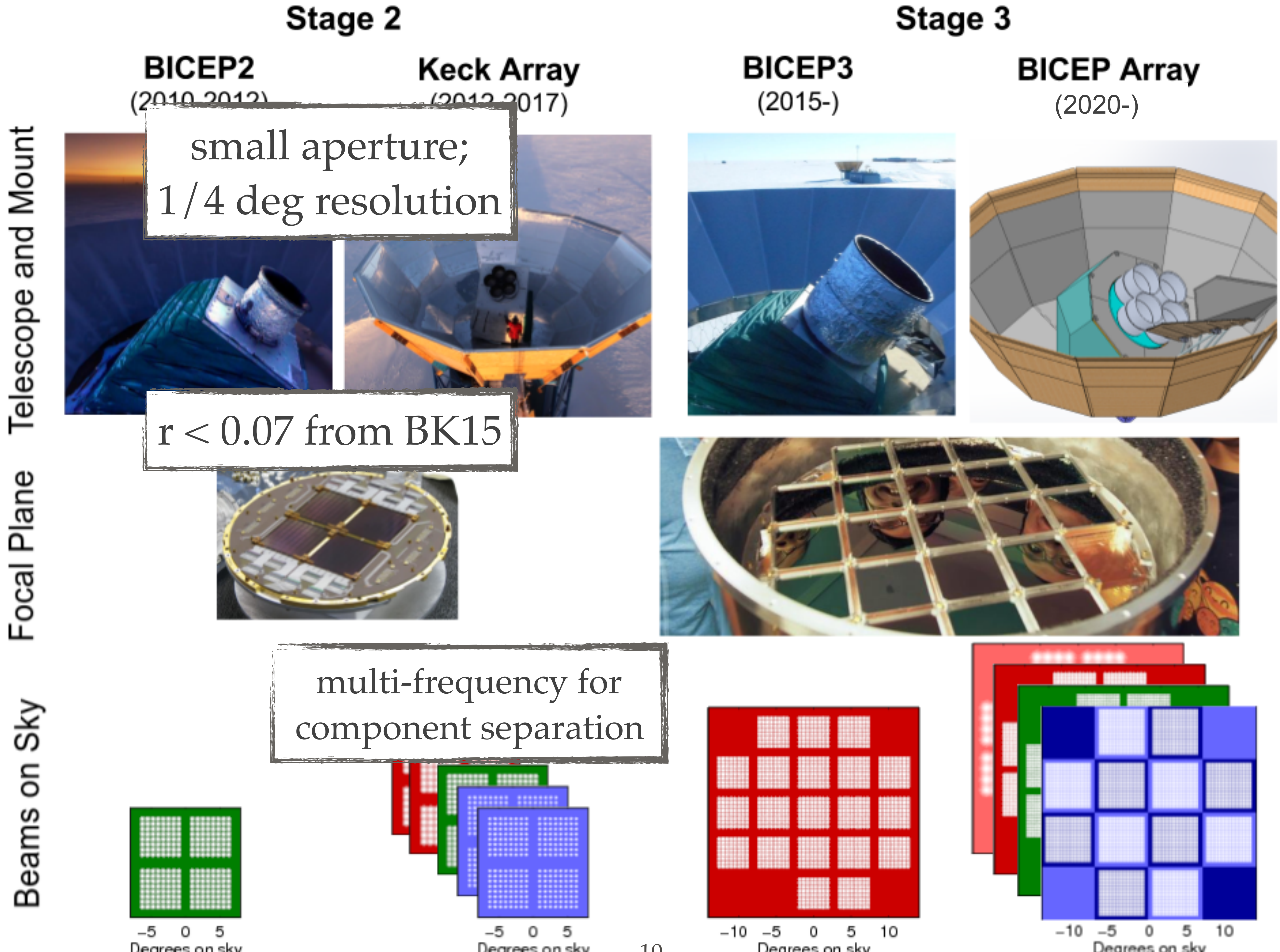



(Green et al: 1508.06342)

# Delensing for $r$ : a BICEP/Keck example




- We can fit lensing model + $r$ simultaneously, but limited by sample variance of lensing
- **Delensing** B-modes: using the ***realization-specific*** lensing B-mode sky to reduce lensing sample variance
- Especially important if observing a small sky patch




small aperture;
1/4 deg resolution

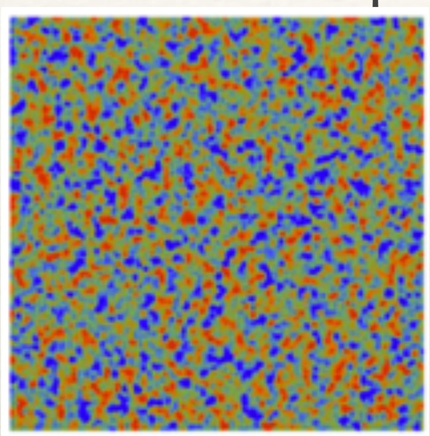
r < 0.07 from BK15

multi-frequency for
component separation

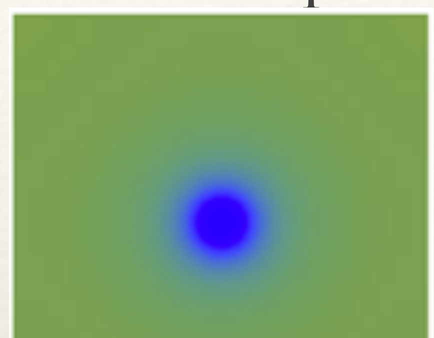
# Delensing: the idea

1. Use Phi tracer and lensed E map to get estimate of lensing B modes

lensed E map



Phi map



weighted integral of gravitational potential between us and CMB along line of sight

B template

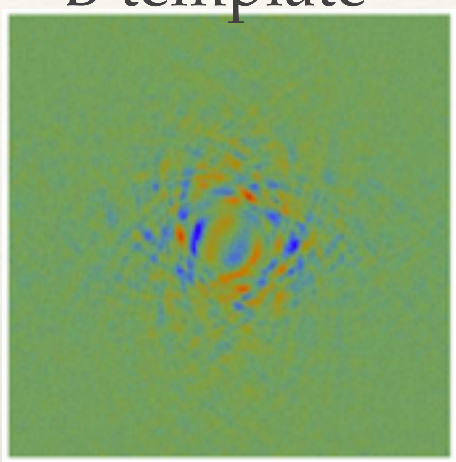
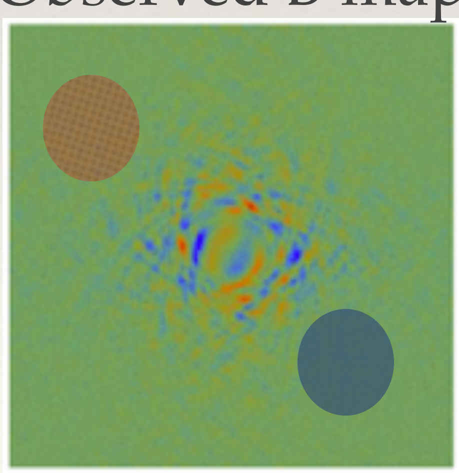


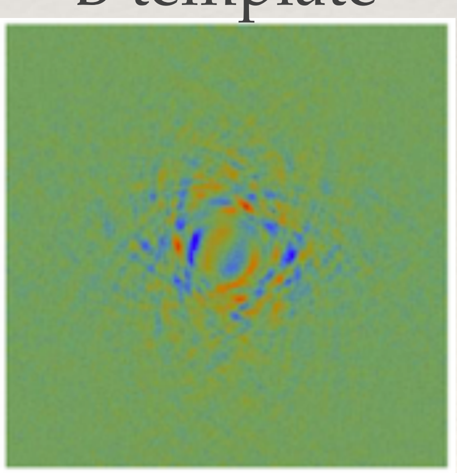
Image credit: Hu & Okamoto (2002)

2. Cross-correlate the lensing B template with observed B mode map to quantify how much lensing B modes are in the observed map

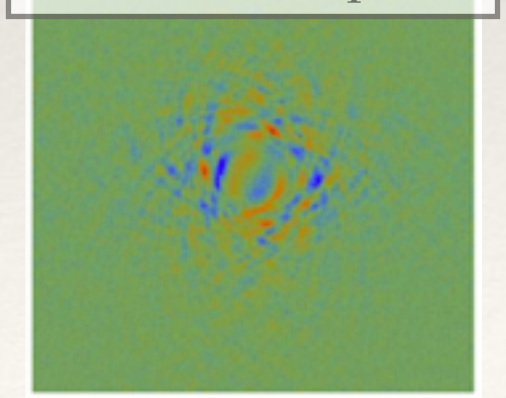
Observed B map



B template

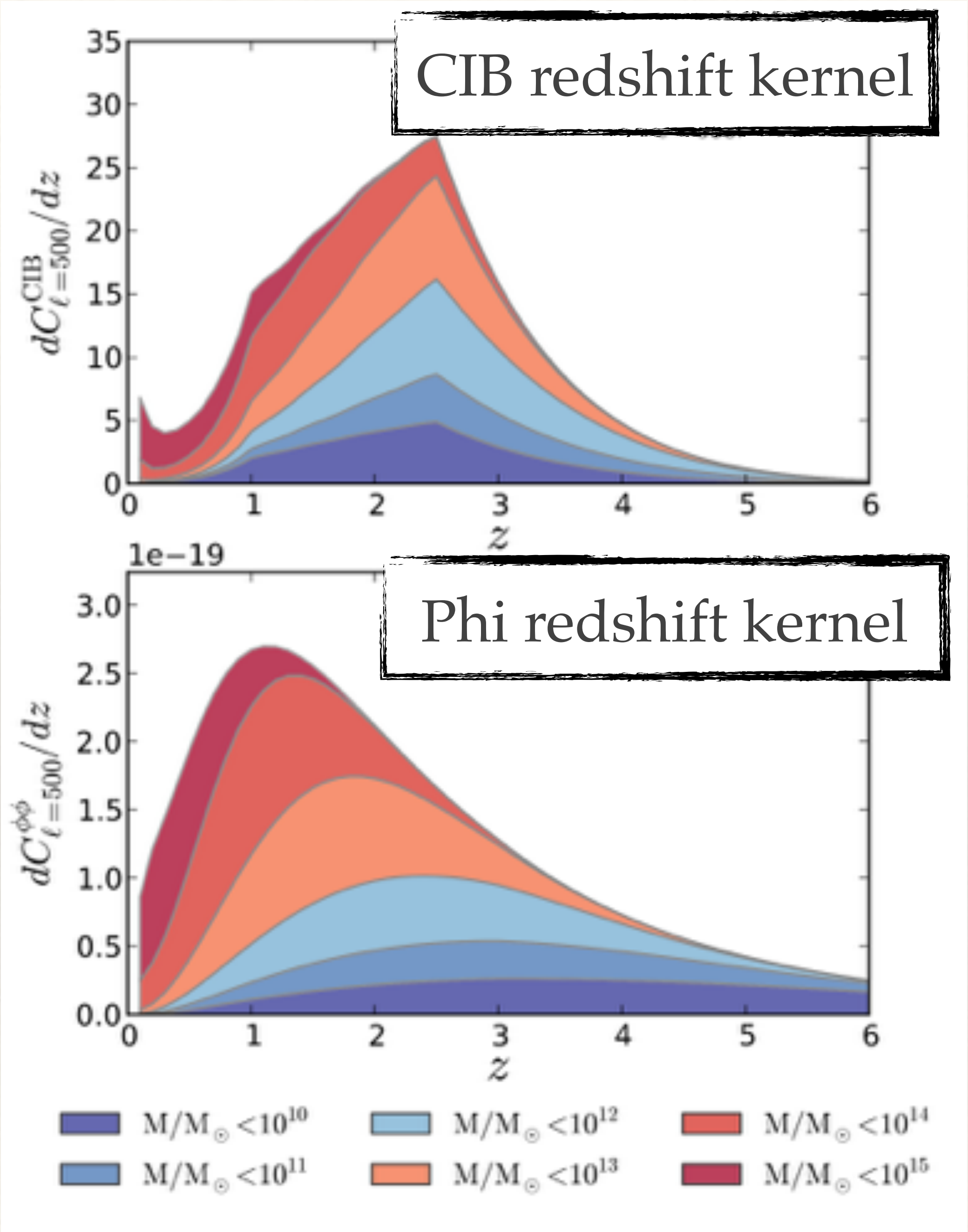


Common structure in both maps!



# CIB as a ϕ tracer

- ϕ: can reconstruct from CMB, but S/N rather low currently (Future will be better!)
- Cosmic infrared background (CIB) from dusty star-forming dusty galaxies with redshift distribution peaked between z~1 and 2.
- CMB lensing potential's redshift kernel peaks between 1 < z < 3
- Cross-correlation can be as high as ~80%

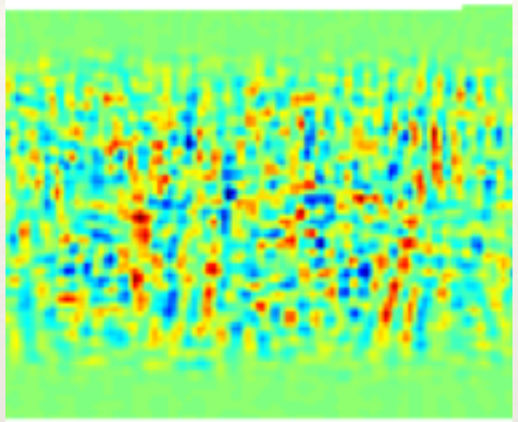



Planck 2013 XVIII

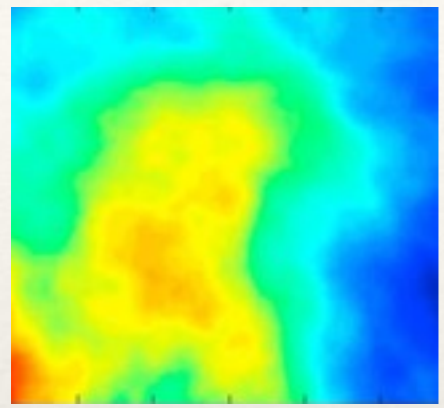
# Lensing template construction

1. Undeflect by $-\nabla(\phi)$

Lensed Q/U map



ϕ map

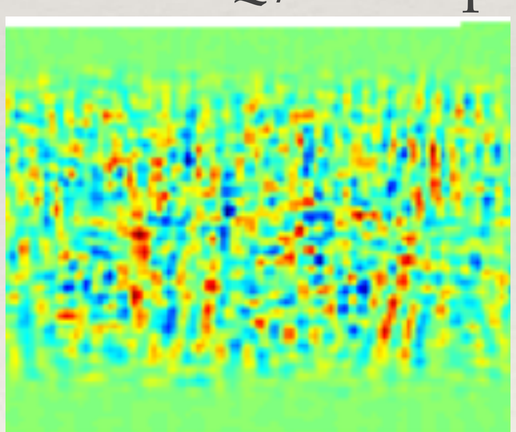


Q/U maps that has no lensing B-modes

(in ideal, noiseless case)

2. Difference the pre- and post-deflected map

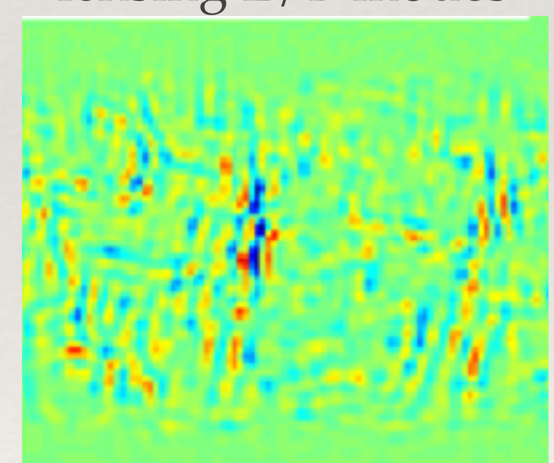
Lensed Q/U map



Q/U maps that has no lensing B-modes

(in ideal, noiseless case)

lensing template:
Q/U maps that contain lensing E/B modes



Feed the Q/U map through a B-estimator to get the power spectra as inputs to the multicomponent analysis.

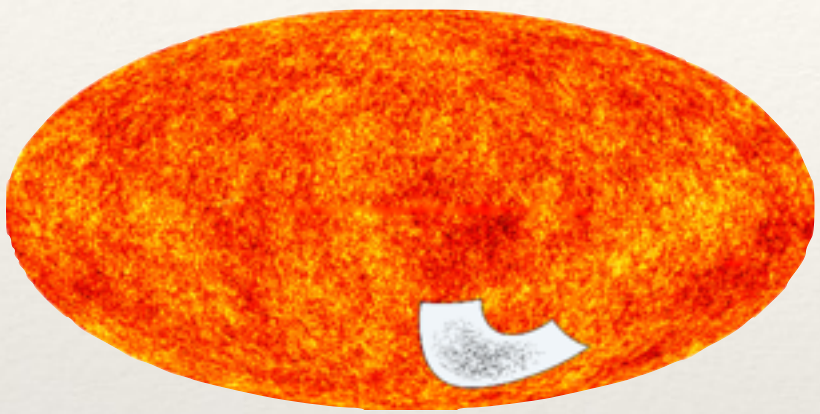
# Connecting delensing to σ(r)

BICEP/Keck analysis framework:

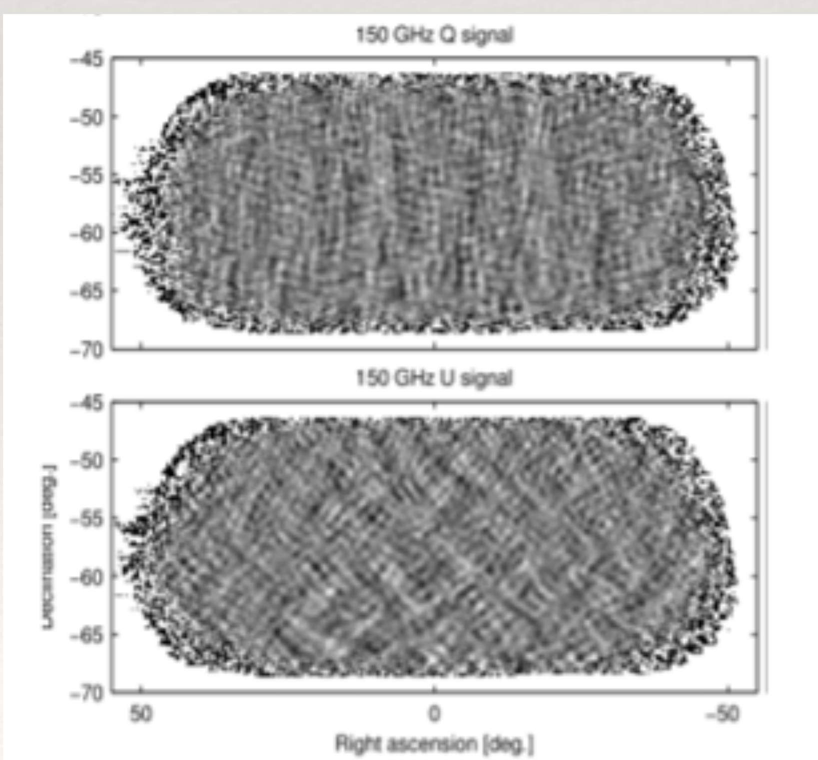
how is delensing incorporated

# BK multicomponent analysis (no delensing)

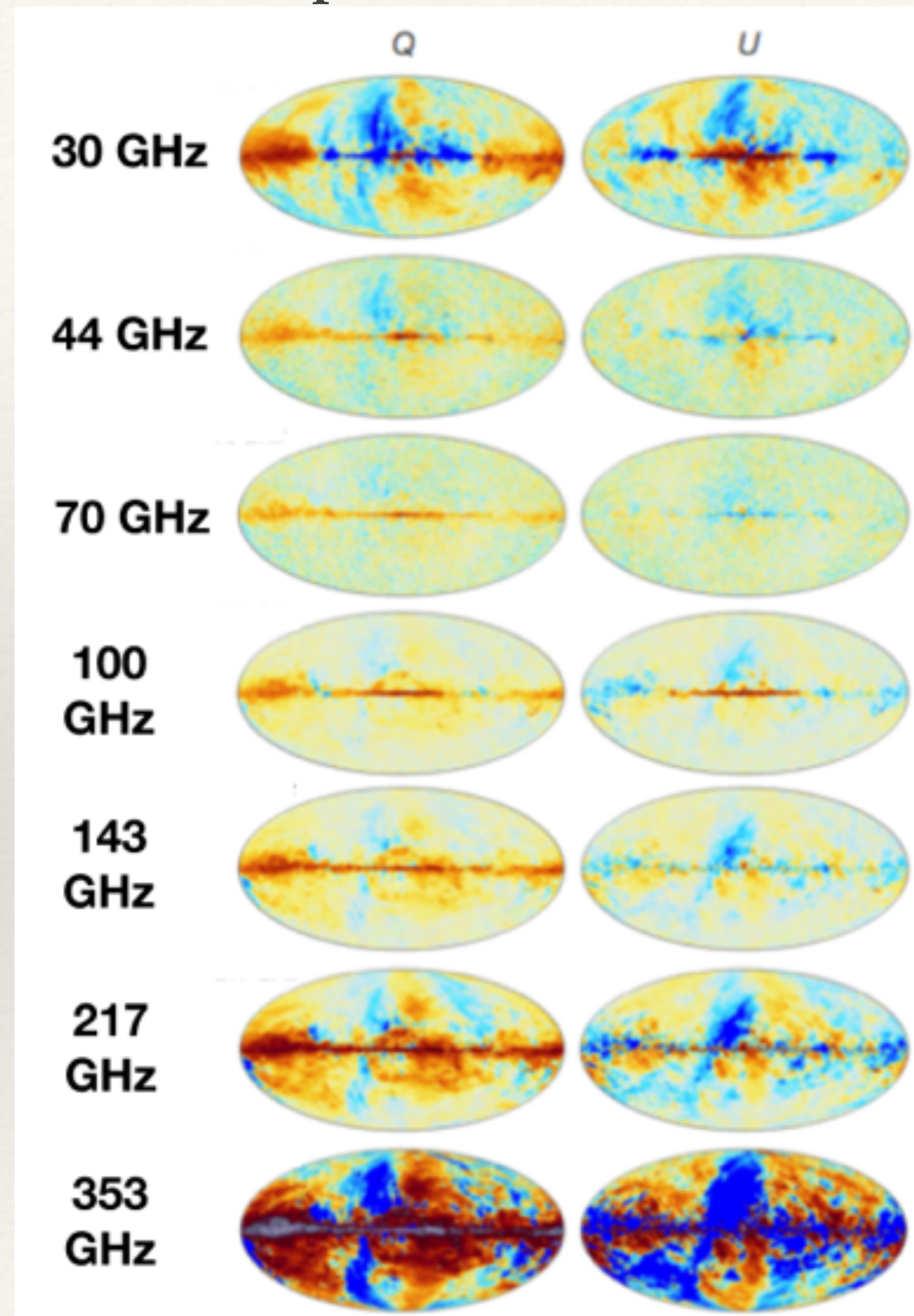
- Input maps to multicomponent analysis that extracts constraints on $r$



**Maps from BICEP/Keck (95/150GHz)**




**Maps from Planck**



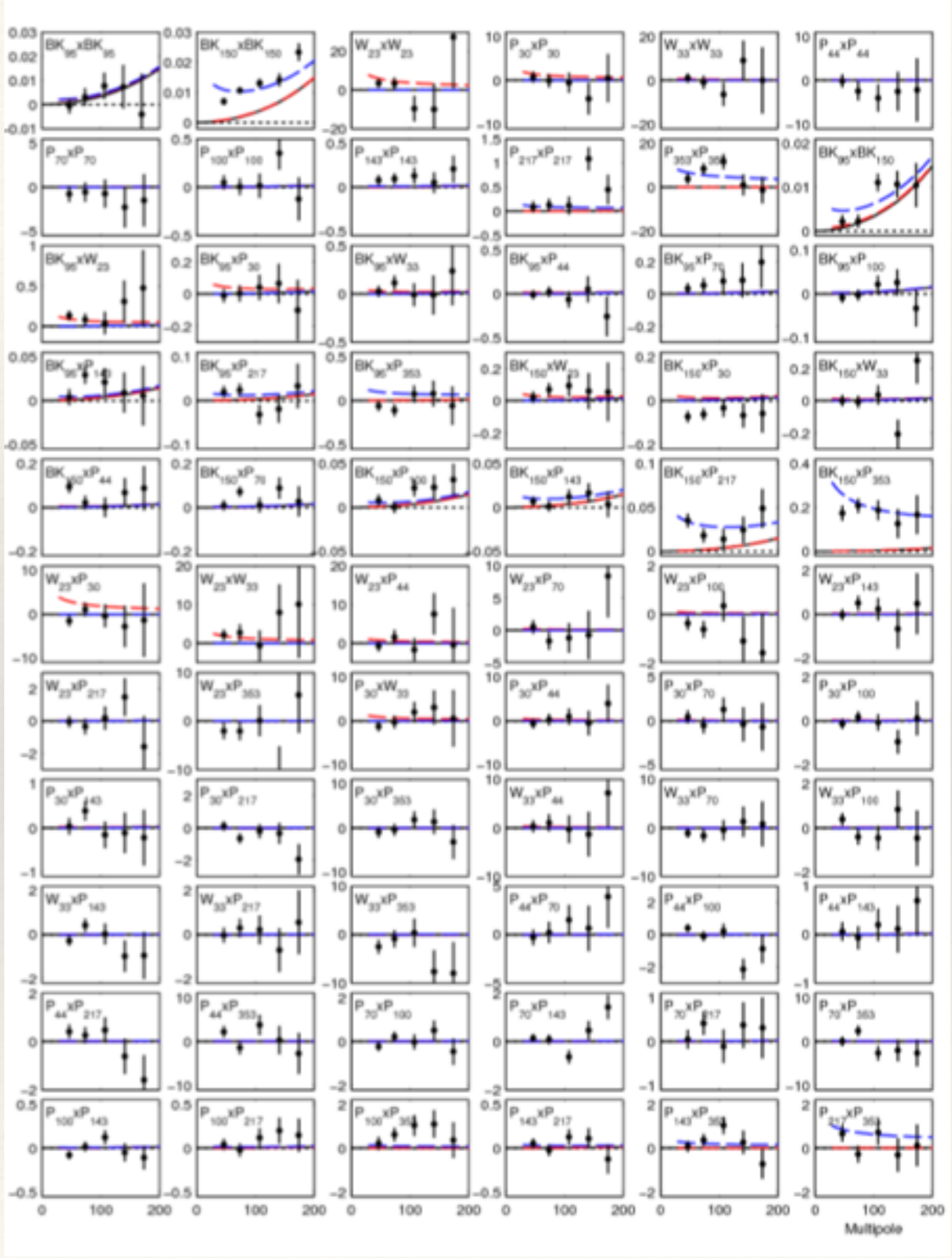

**+ WMAP**

# BK multicomponent analysis (no delensing)

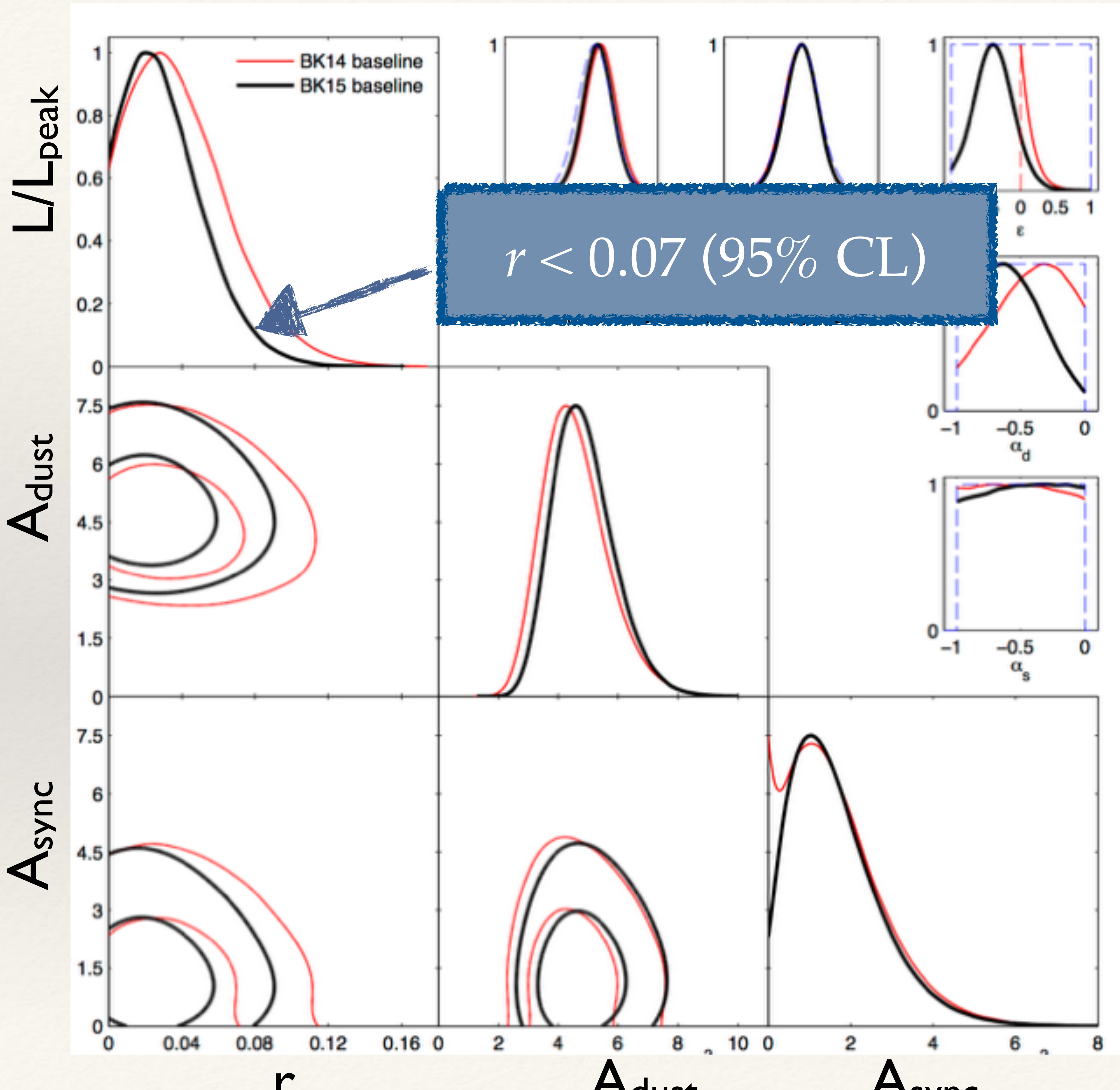
- Take the auto- and cross-spectra of the BICEP/Keck and WMAP/Planck maps
- To calculate the likelihood, compare the data bandpowers against the model expectation values of lensing BB, *r*, and 7 parameter foreground model:

  $A_{\rm dust}, \alpha_{\rm dust}, \beta_{\rm dust}, A_{\rm sync}, \alpha_{\rm sync}, \beta_{\rm sync}$

  dust/sync correlation



# BK15 constraints

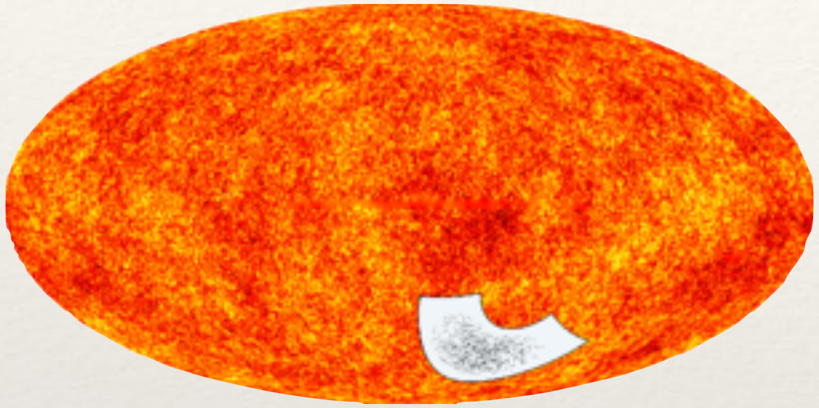



Allow Dust/
Sync correlation

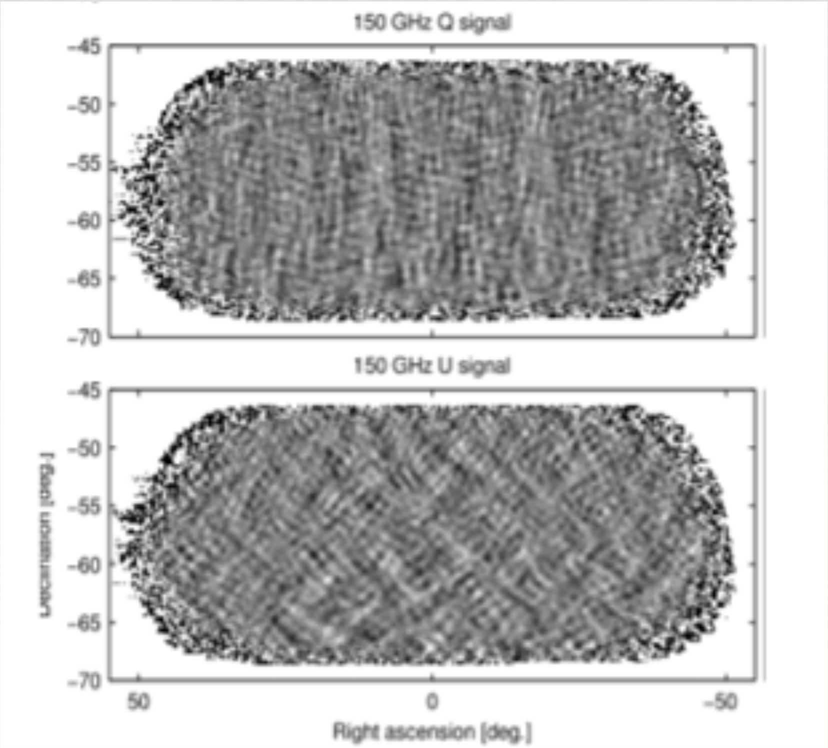
Allow generous
variation in spatial
spectral indices of
dust & sync

# BK multicomponent analysis (+ delensing)

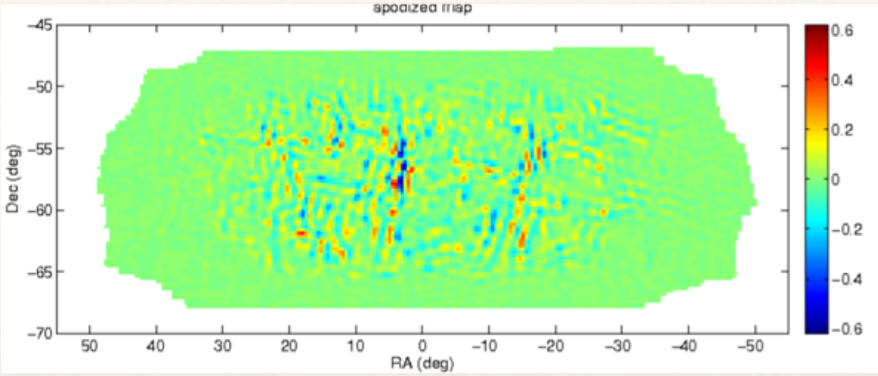
- Input maps to multicomponent analysis that extracts constraints on $r$



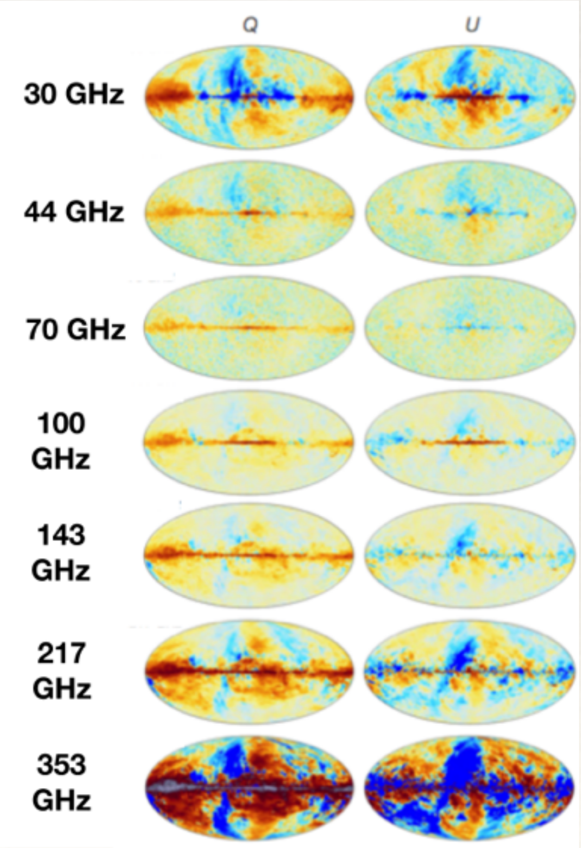
**Maps from BICEP/Keck (95/150GHz)**




+ lensing template

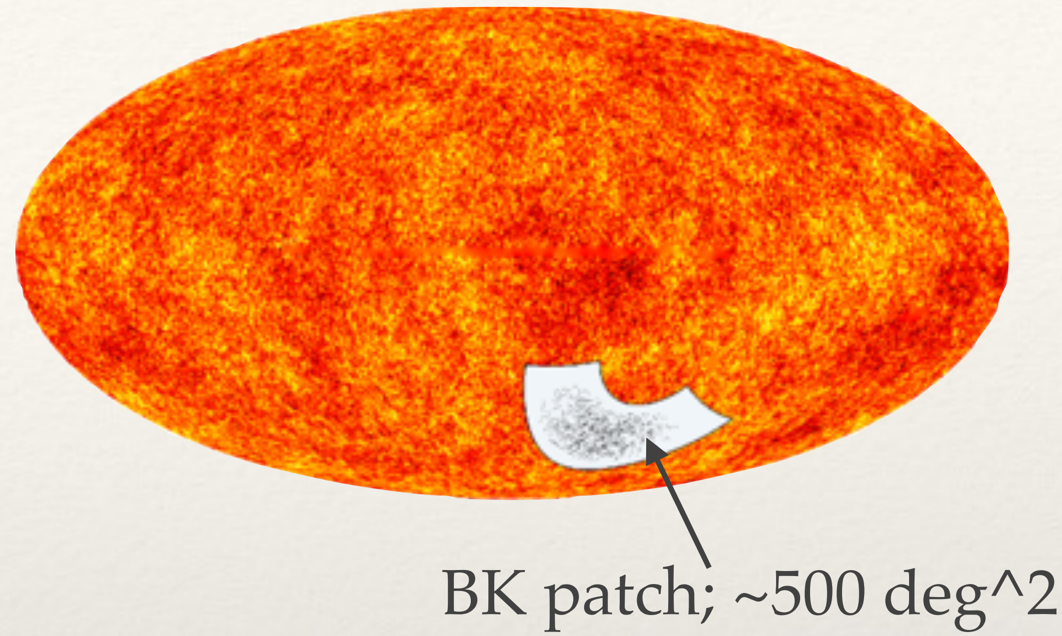


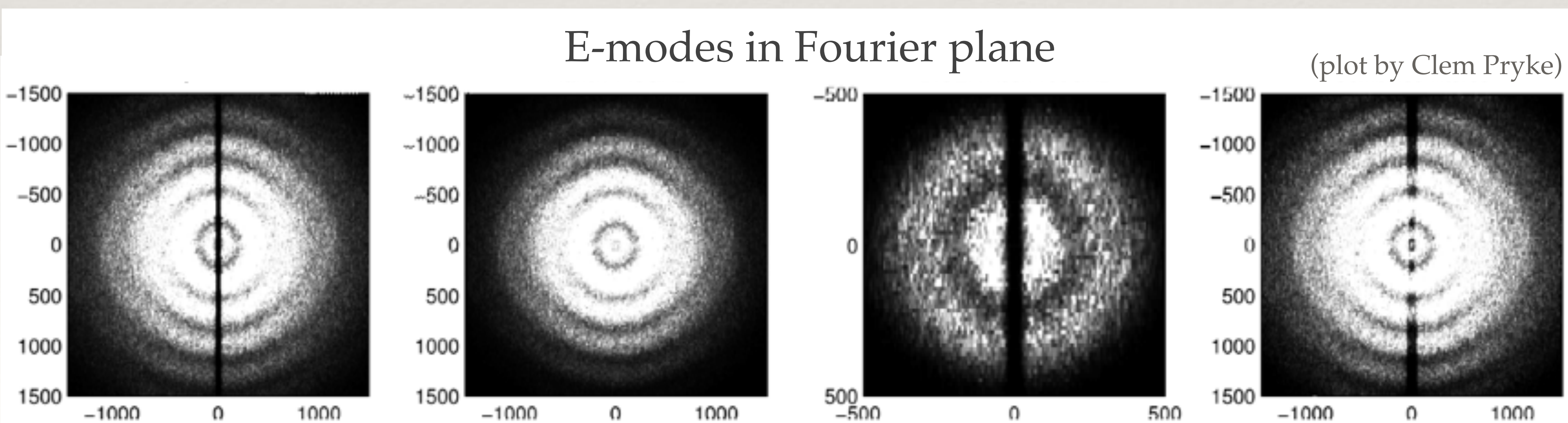

**Maps from Planck**



# Inputs to BK lensing template

- Phi tracer: Planck's CIB map
- Q/U maps: combination of BICEP/Keck, SPTpol, and Planck maps




SPTpol | Planck | BICEP/Keck (note different ell-range) | Combined

# Incorporating lensing template to likelihood

- Use same model: lensing BB, *r*, and 7 parameter foreground model:

  $A_{\text{dust}}, \alpha_{\text{dust}}, \beta_{\text{dust}}, A_{\text{sync}}, \alpha_{\text{sync}}, \beta_{\text{sync}}$ , and dust/sync correlation

- Adding the lensing template increases the total auto/cross BB spectra from 66 to 78

$$\mathcal{L}(\boldsymbol{\theta}|\boldsymbol{d}) \propto \frac{1}{\sqrt{|\mathbf{C}(\boldsymbol{\theta})|}} \exp\left(-\frac{1}{2}(\boldsymbol{d}-\boldsymbol{\mu}(\boldsymbol{\theta}))^{\dagger}\left[\mathbf{C}(\boldsymbol{\theta})\right]^{-1}(\boldsymbol{d}-\boldsymbol{\mu}(\boldsymbol{\theta}))\right)$$

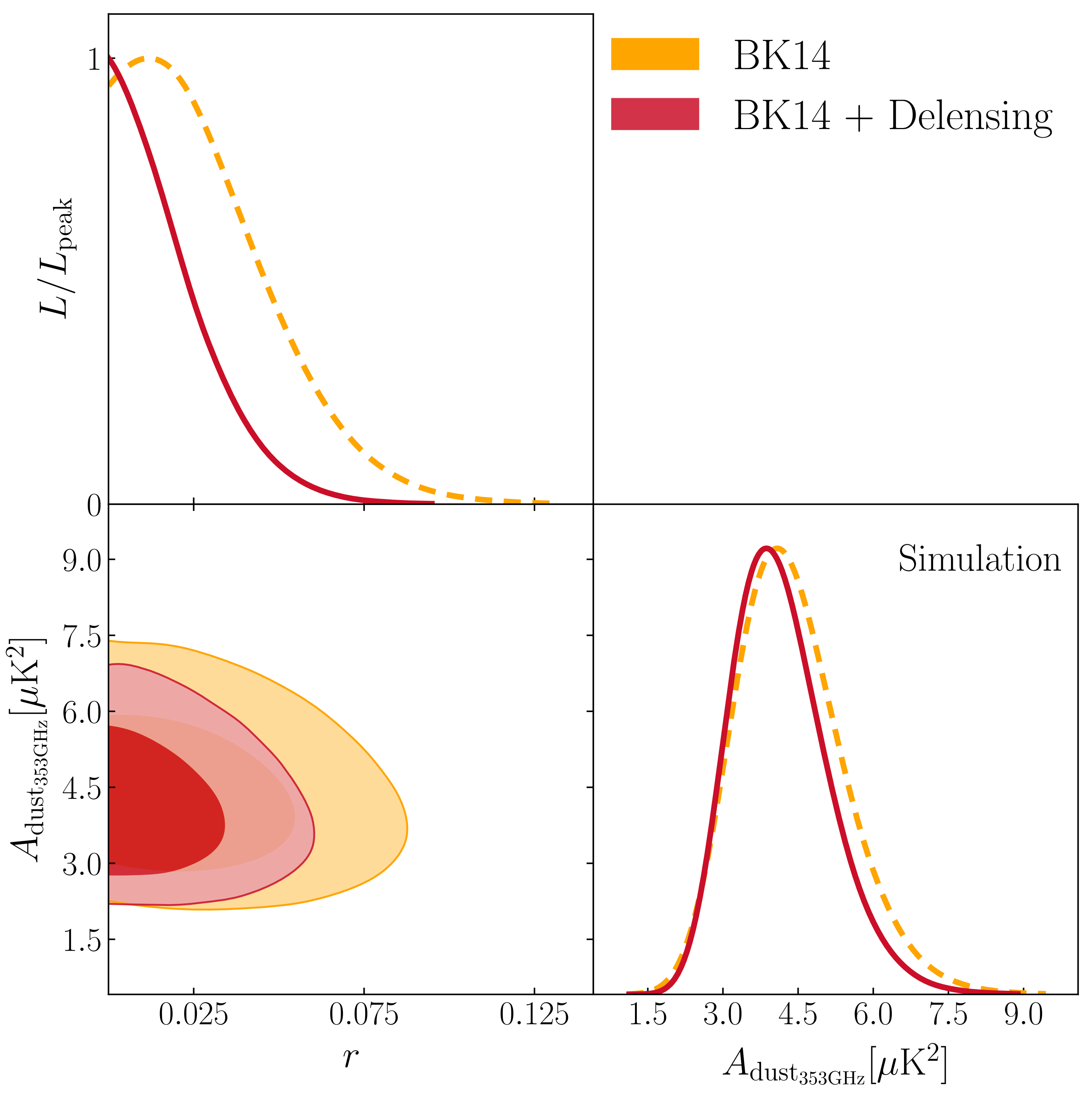
more data bandpowers

larger bandpower
covariance matrix

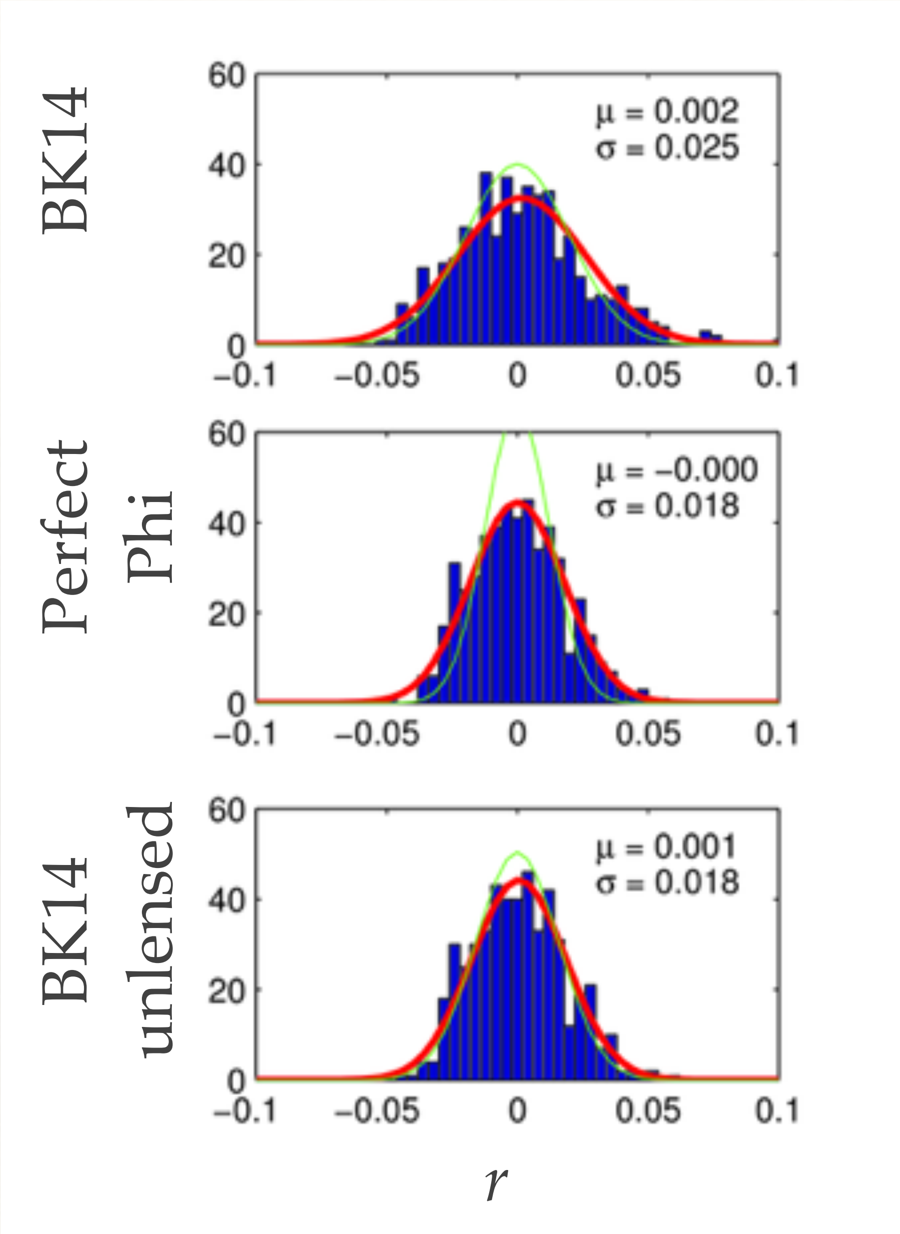
In the BK analysis, we use the HL likelihood.
The gaussian likelihood is for illustration purposes.

# Lensing template as input in multicomponent analysis

The covariance matrix that enters the likelihood has information of the covariance between the lensing BB spectrum and the observed BB spectrum —> reducing σ(r).



# How much do we improve σ(r)?




- With perfect ϕ map (no decorrelation, no noise), adding a lensing template to the BK14 data set improves σ(r) from 0.025 to 0.018
- Using CIB phi tracer to form the lensing template, σ(r) improves by ~10% from BK14

# Current limitation to delensing

- B mode variance is dominated by galactic foregrounds; even with perfect delensing we do not improve σ(r) very significantly
- CIB map we use has cross-correlation with underlying ϕ at 60-80%; need better cross-correlation to improve towards perfect delensing
- CMB-derived ϕ from next-generation CMB experiments can provide that!

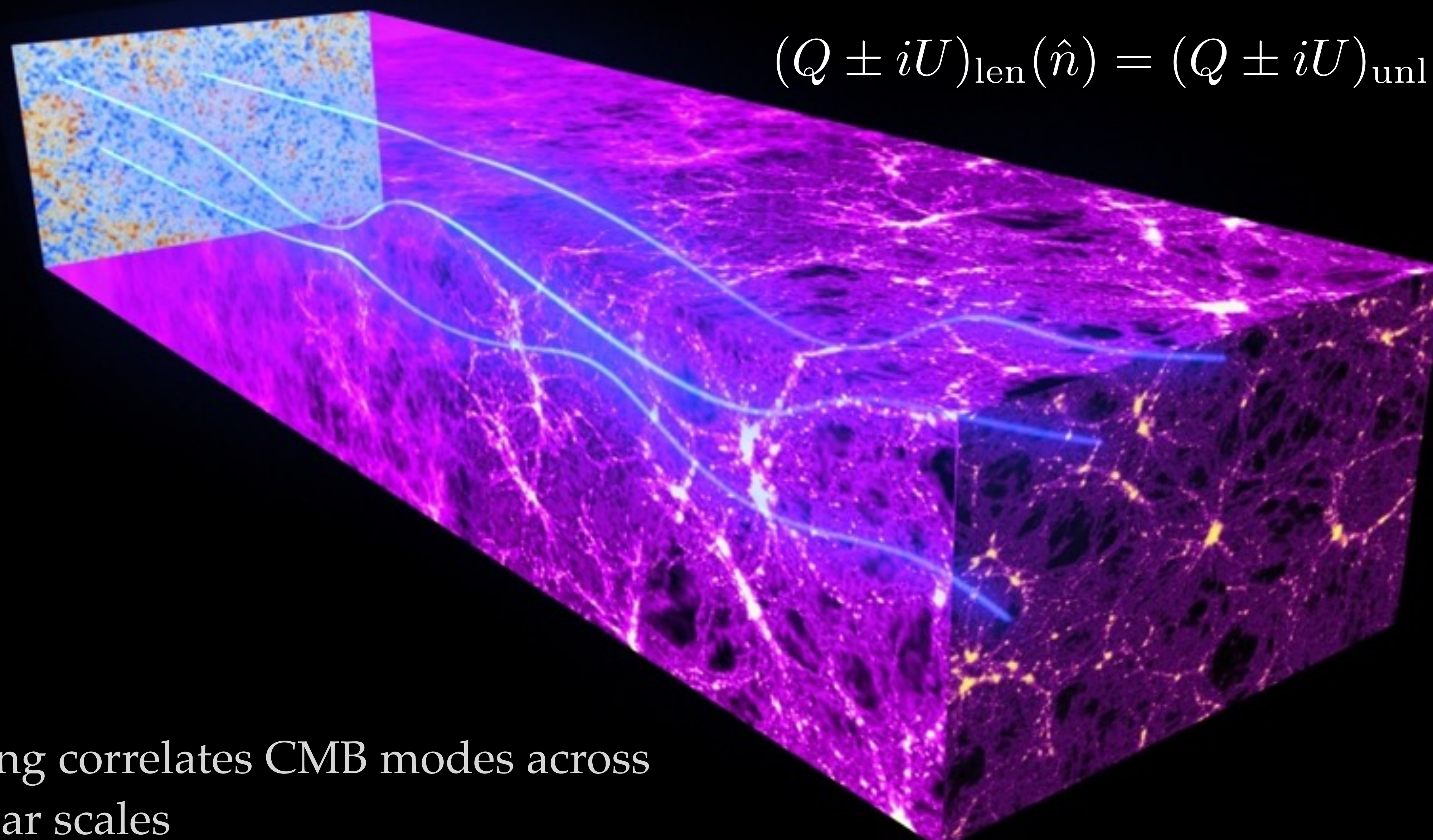
Key take-aways:

1) Incorporated delensing into a likelihood analysis for *r*;
2) Delensing reduces σ(r) by ~10% for the BK14 dataset.

# Getting a better ϕ tracer: CMB lensing reconstruction

$$T_{\rm len}(\hat{n}) = T_{\rm unl}(\hat{n} + \nabla\phi)$$

$$(Q \pm iU)_{\rm len}(\hat{n}) = (Q \pm iU)_{\rm unl}(\hat{n} + \nabla\phi)$$



Lensing correlates CMB modes across angular scales

These off-diagonal correlations ∝ ϕ(L); can use the correlations to measure ϕ!

Background image credit: ESA

# Lensing reconstruction: quadratic estimator

In equations, $\phi(\mathbf{L})$ can be estimated as follows:

$$\bar{\phi}^{XY}_{\mathbf{L}} = \frac{1}{R^{XY}_{\mathbf{L}}} \int d^2\boldsymbol{\ell}\, W^{XY}_{\boldsymbol{\ell},\boldsymbol{\ell}-\mathbf{L}}\, \bar{X}_{\boldsymbol{\ell}} \bar{Y}^{*}_{\boldsymbol{\ell}-\mathbf{L}}$$

Lensing potential

Normalization

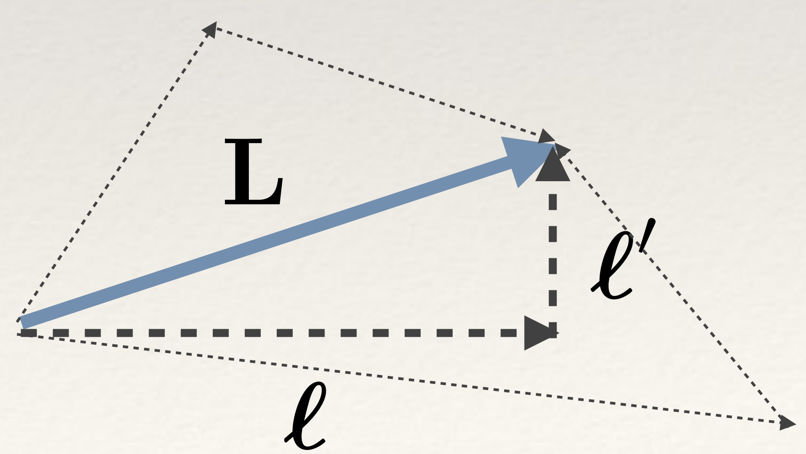
Weight function

Filtered CMB maps

X, Y = [T, E, B]

The estimated $\phi(\mathbf{L})$ is a weighted sum of the products of Fourier modes from X and Y for all the pairs of $\ell$ and $\ell'$ where $\mathbf{L} = \ell + \ell'$;

It picks out the correlations in the lensed CMB maps introduced by $\phi(\mathbf{L})$.

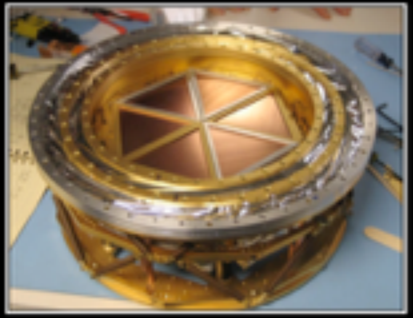


# The South Pole Telescope (SPT)

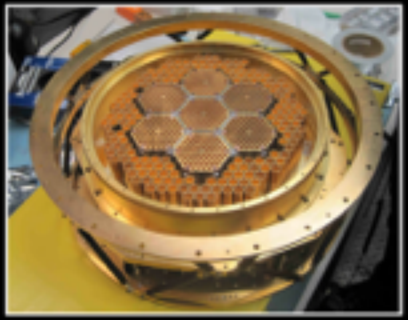
10-meter sub-mm quality wavelength telescope

95, 150, 220 GHz and
1.6, 1.2, 1.0 arcmin resolution

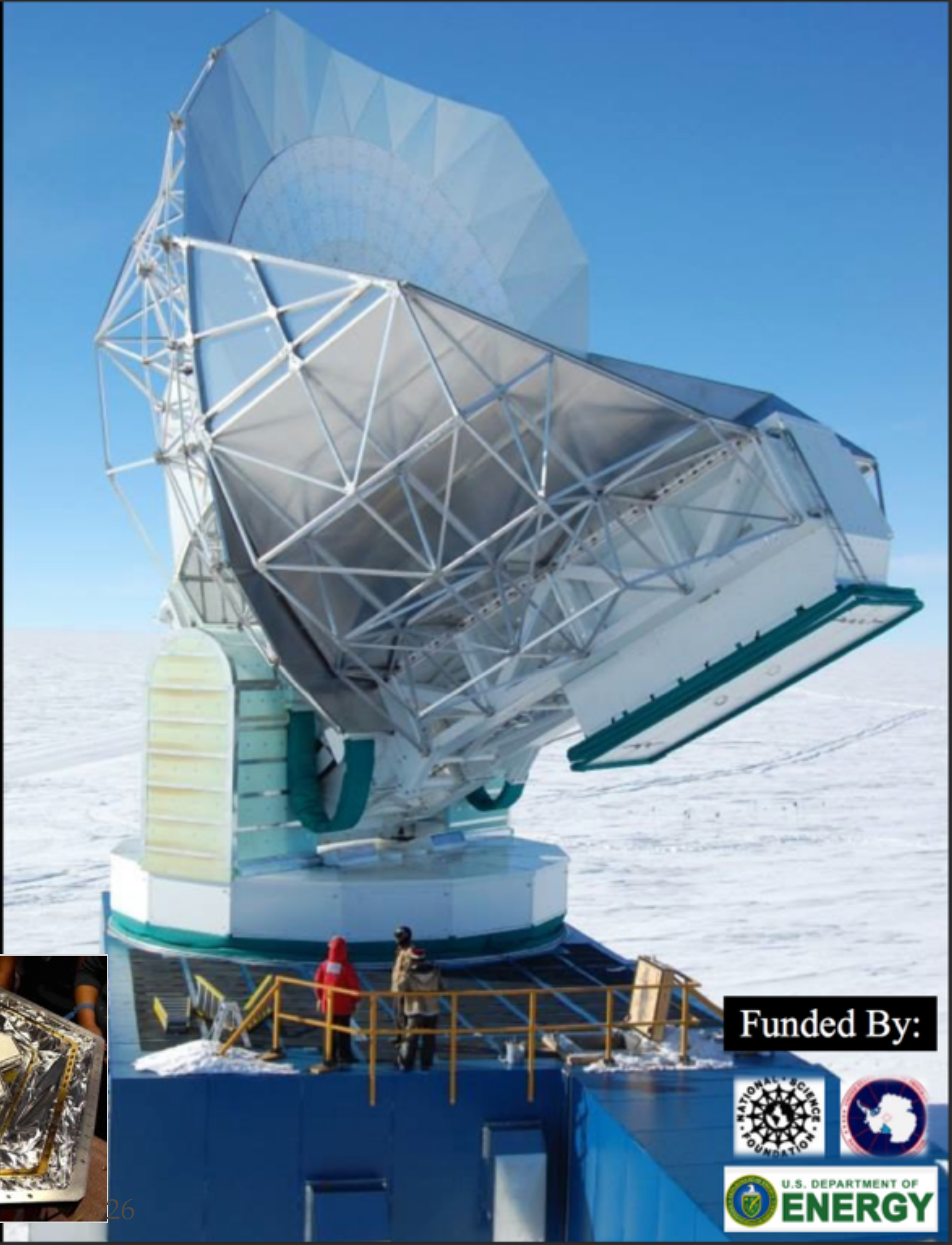
**2007: SPT-SZ**
960 detectors
95,150,220 GHz



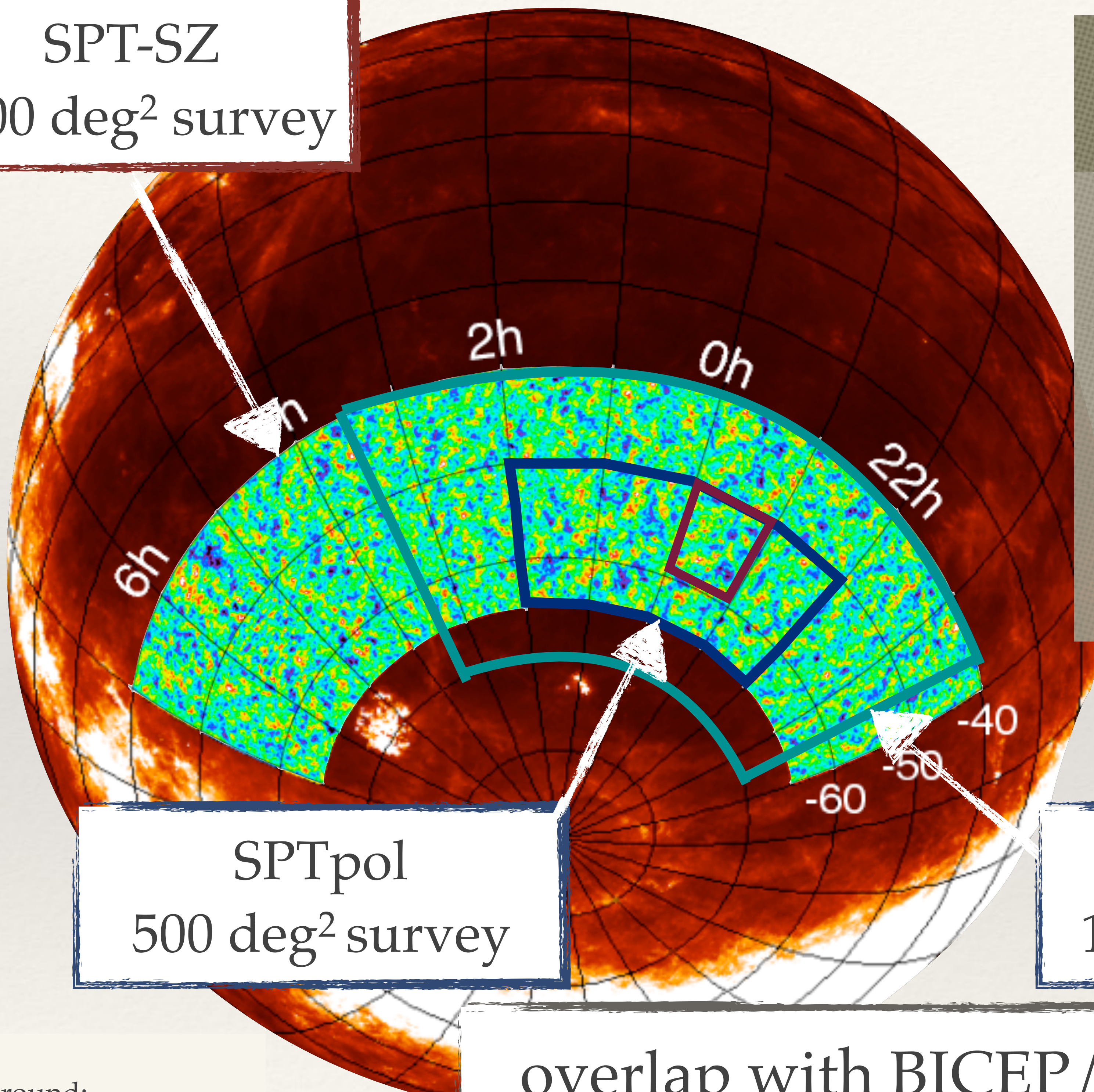
**2012: SPTpol**
1600 detectors
95,150 GHz
***+Polarization***



**2016: SPT-3G**
~16,000 detectors
95,150,220 GHz
***+Polarization***



# SPT surveys

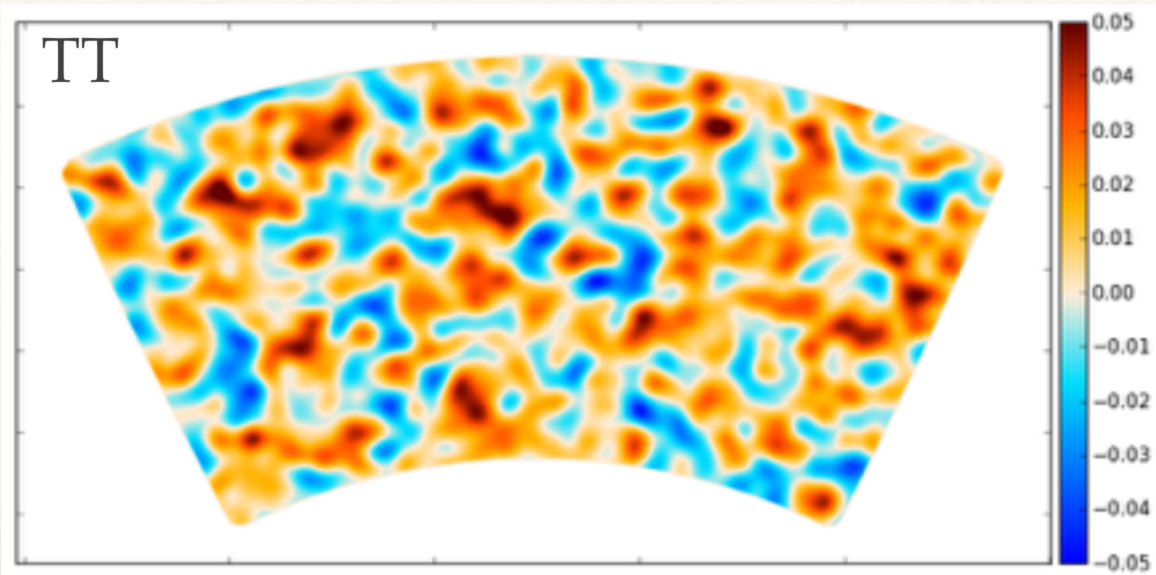



| receiver | sky area | noise @ 150GHz |
|---|---|---|
| SPT-SZ | 2500 deg² | 17 uK’ |
| SPTpol | 500 deg² | 7 uK’ |
| SPT-3G | 1500 deg² | 2.2 uK’ (projected) |

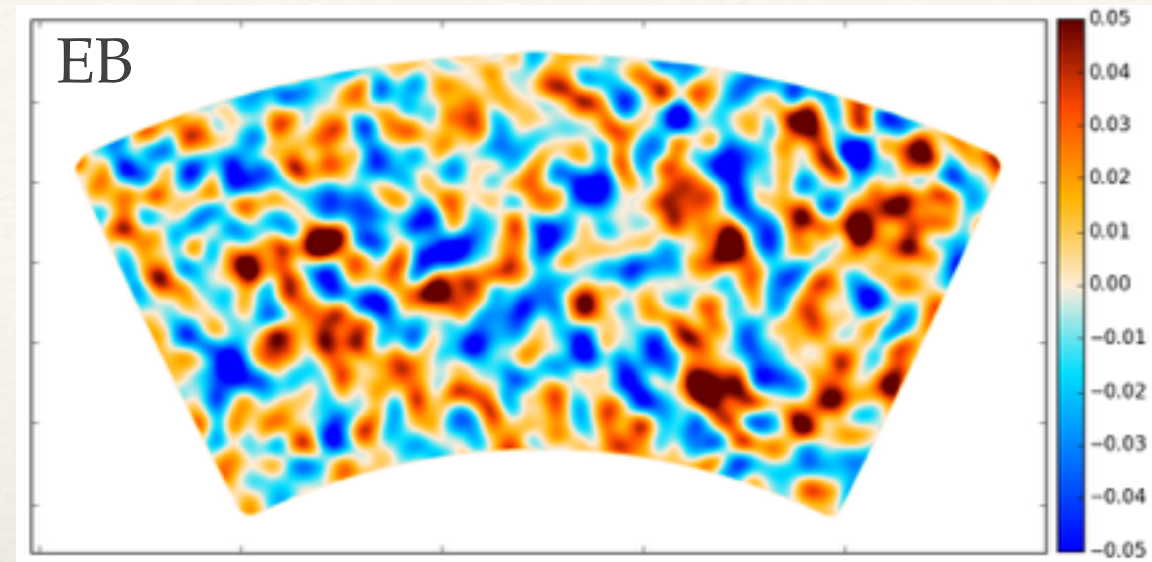
Background:
IRAS dust map
Schlegel et al 1998

# SPTpol lensing map

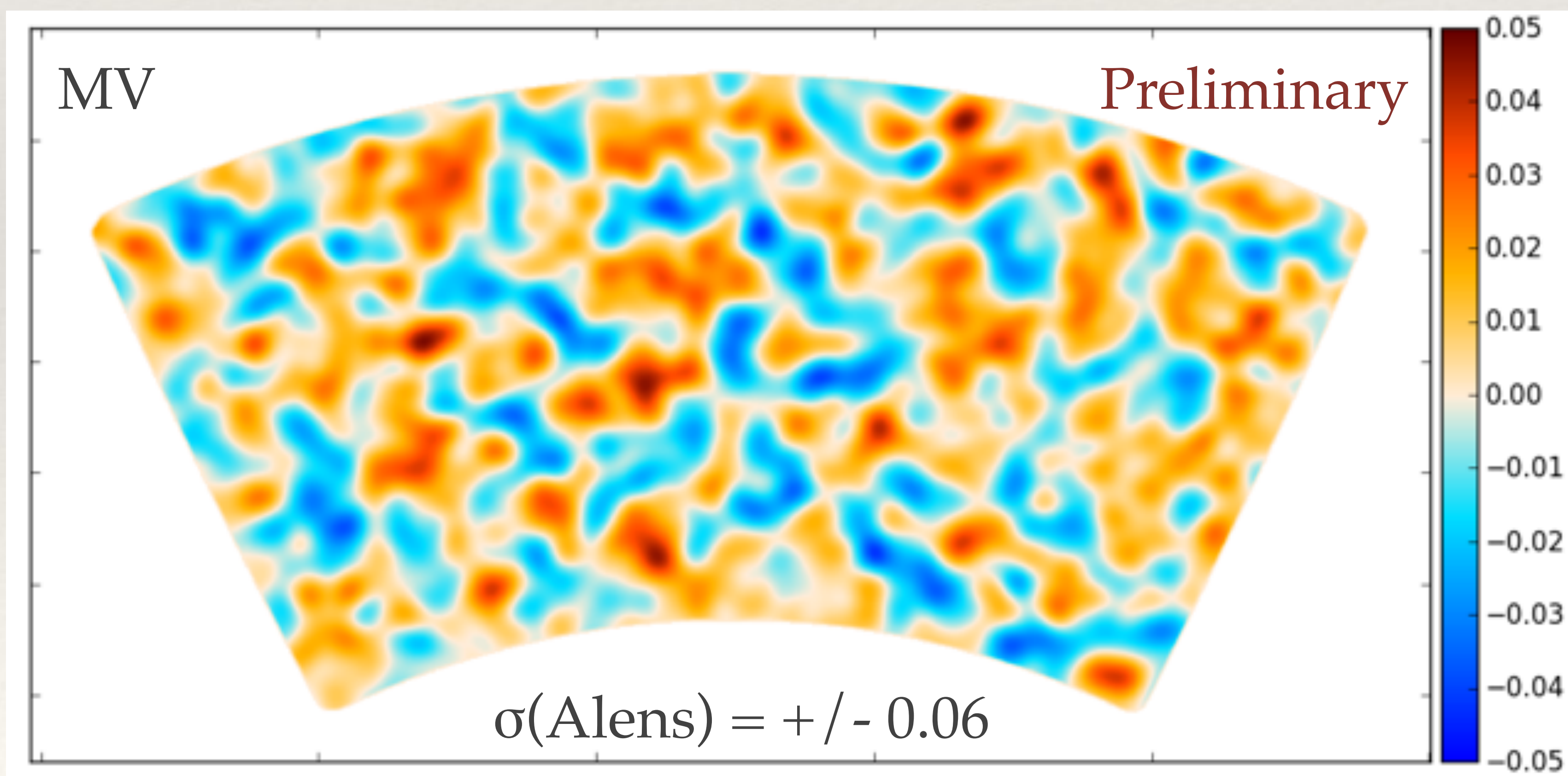
$$\bar{\phi}^{XY}_{\mathbf{L}} = \frac{1}{R^{XY}_{\mathbf{L}}} \int d^2\boldsymbol{\ell}\, W^{XY}_{\boldsymbol{\ell},\boldsymbol{\ell}-\mathbf{L}}\, \bar{X}_{\boldsymbol{\ell}} \bar{Y}^{*}_{\boldsymbol{\ell}-\mathbf{L}}$$




σ(Alens) = +/- 0.08



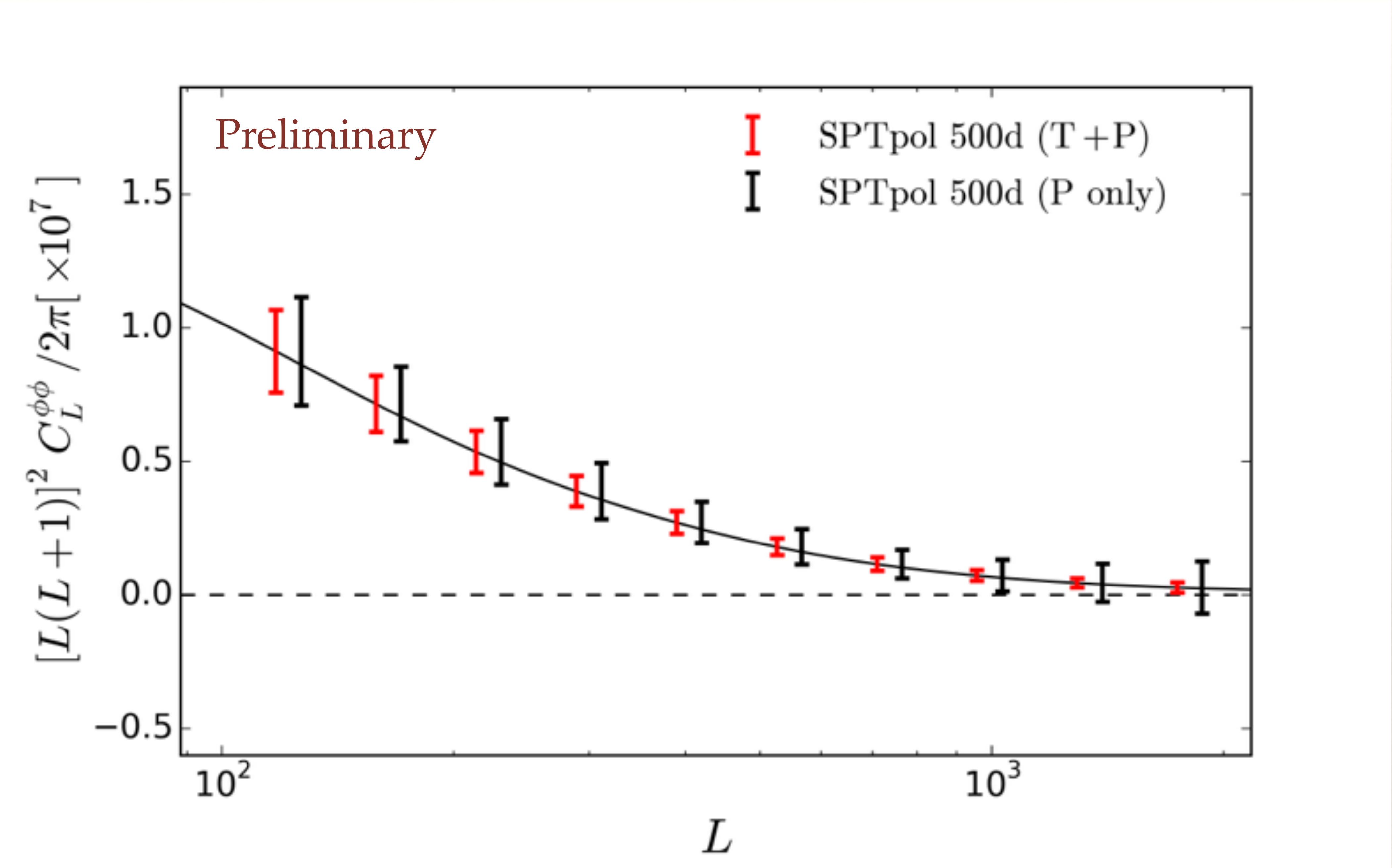

σ(Alens) = +/- 0.13




σ(Alens) = +/- 0.06

Work led by M. Mocanu

# Cosmology from the $Cl^{\phi\phi}$ spectrum

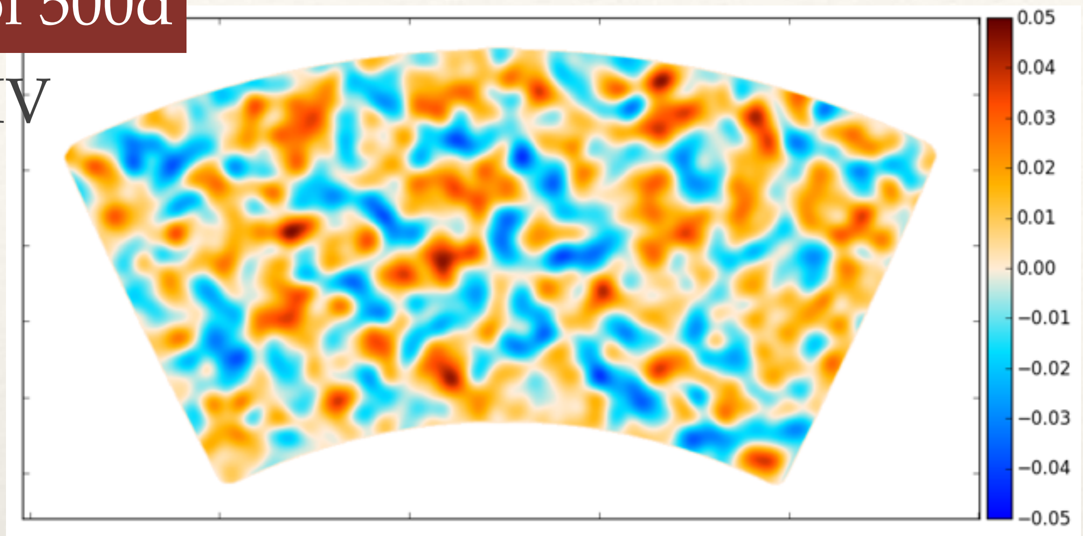



- neutrino mass
- $\sigma_8$ / $A_{lens}$ / $\Omega_M$ (cross-correlate with / compare against optical surveys)

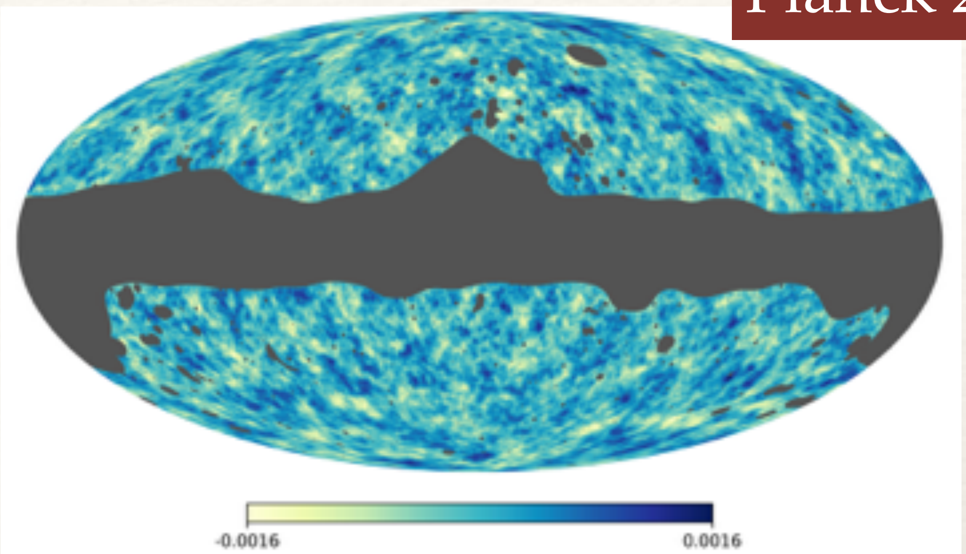
# Lensing map noise

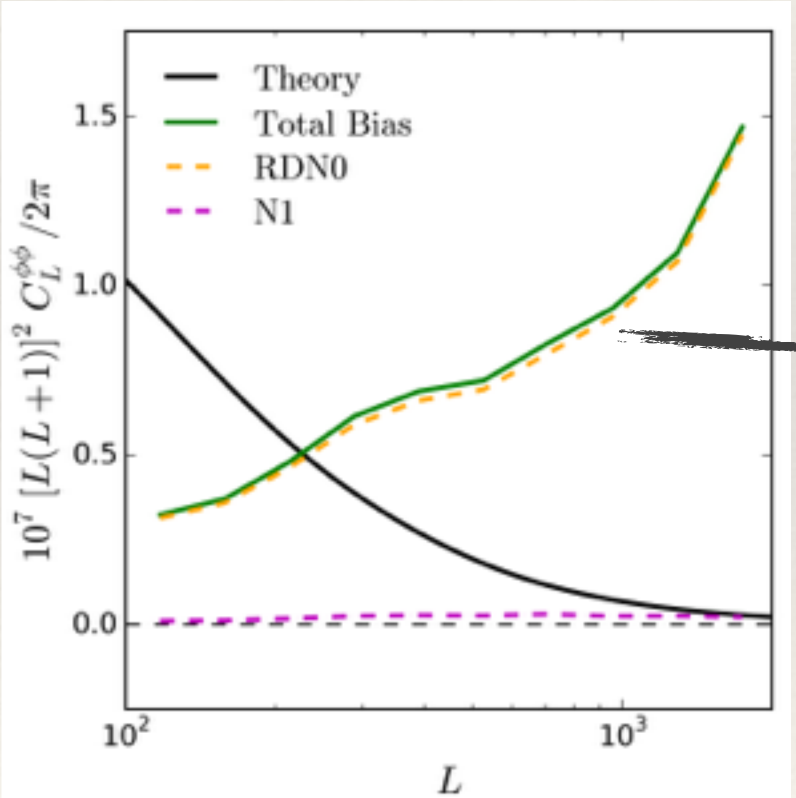
SPTpol 500d

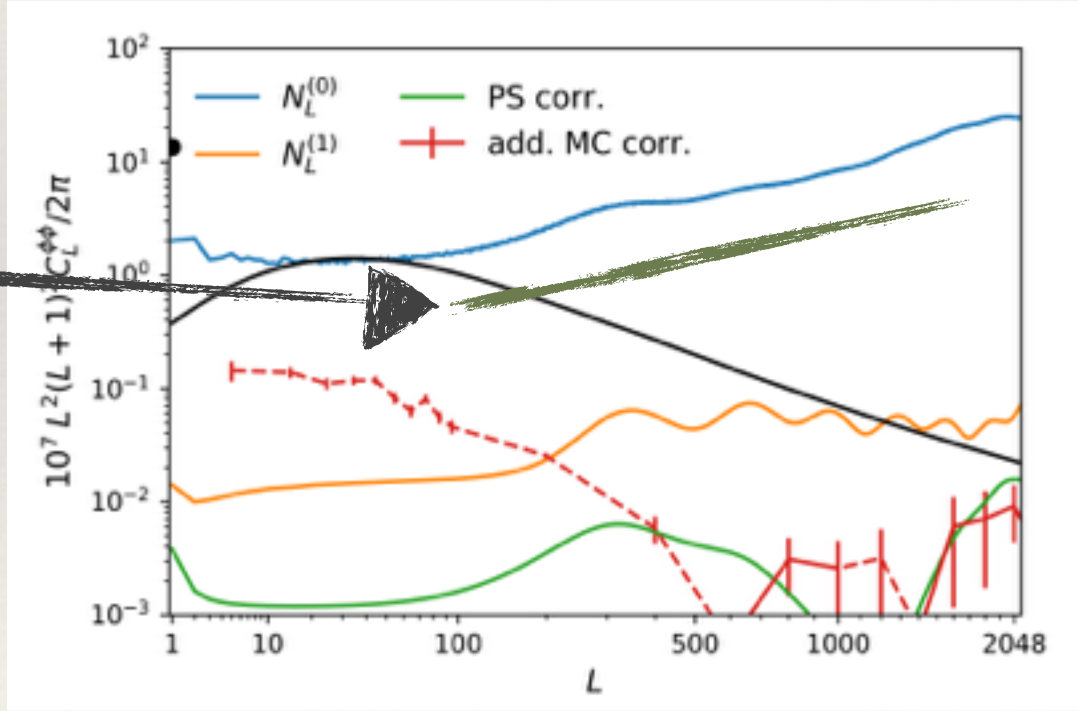
MV




Planck 2018




Planck 2018 results. VIII. Gravitational lensing

High S/N per lensing mode measurement in the SPTpol patch important for delensing

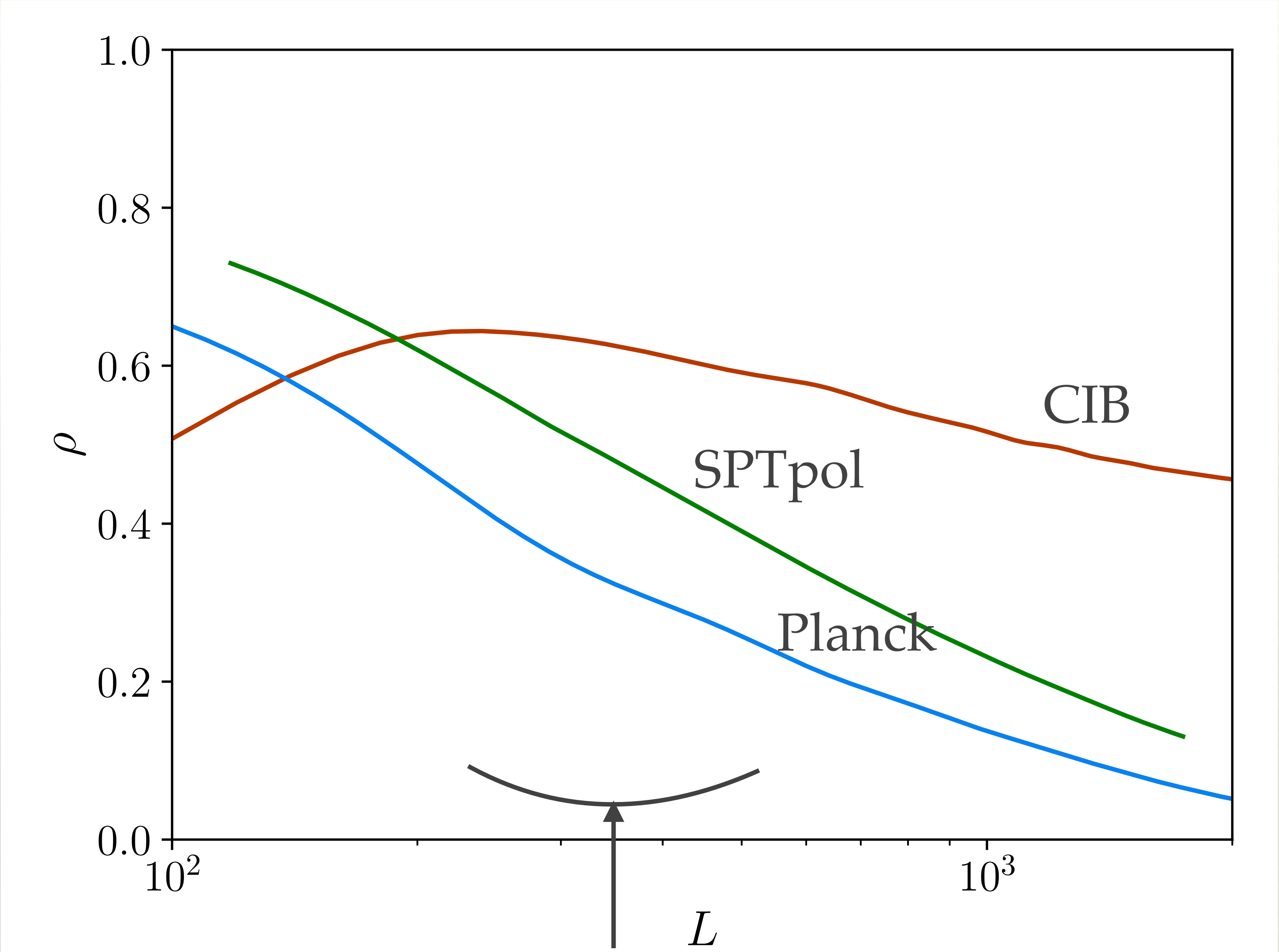
# Delensing efficiency

Cross-correlation of tracer and ϕ-field

$$\rho_\ell = \frac{C_l^{\text{tracer-}\phi}}{\sqrt{C_l^{\text{tracer-tracer}} C_l^{\phi\phi}}}$$

For CMB reconstructed ϕ

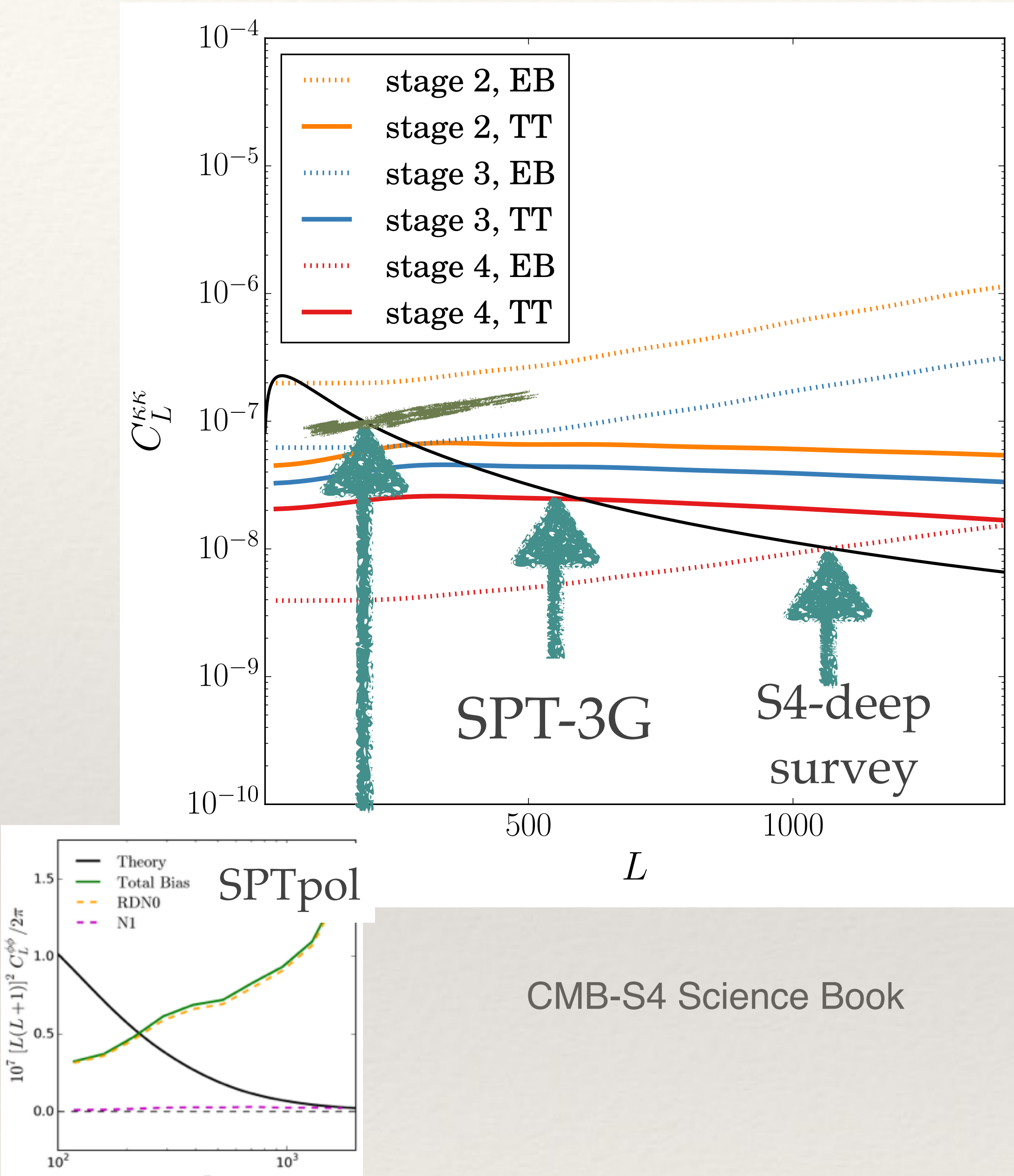
$$\rho_\ell = \sqrt{\frac{C_\ell^{\phi\phi}}{C_\ell^{\phi\phi} + N_\ell^{\phi\phi}}}$$



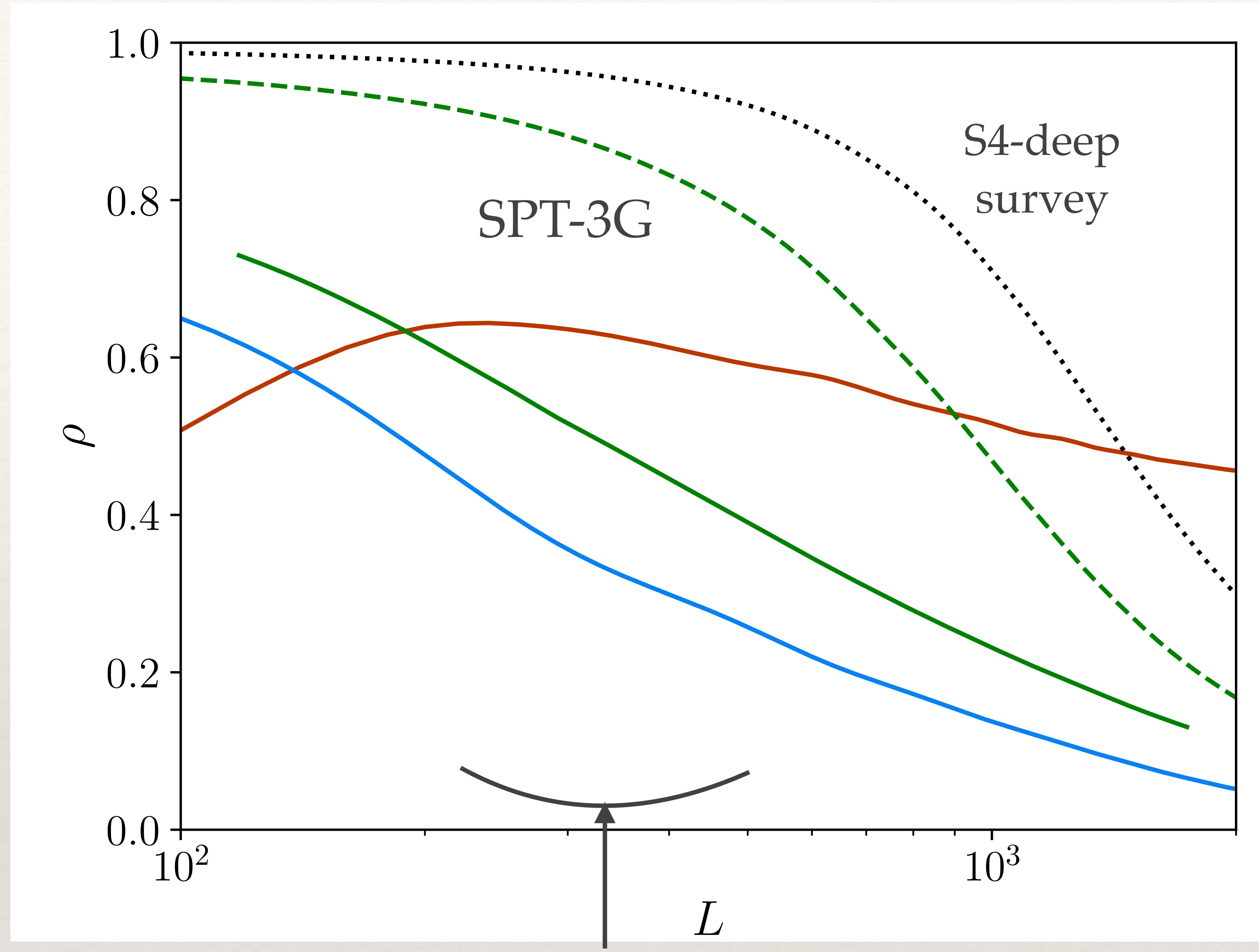

~scales of lenses that source most lensing B-modes

- In the limit that the E-mode noise is small, the correlation between the ϕ tracer and the underlying phi field determines how well the lensing B-modes are estimated —> delensing efficiency

# Forecasts (SPT-3G / CMB-S4)




CMB-S4 Science Book

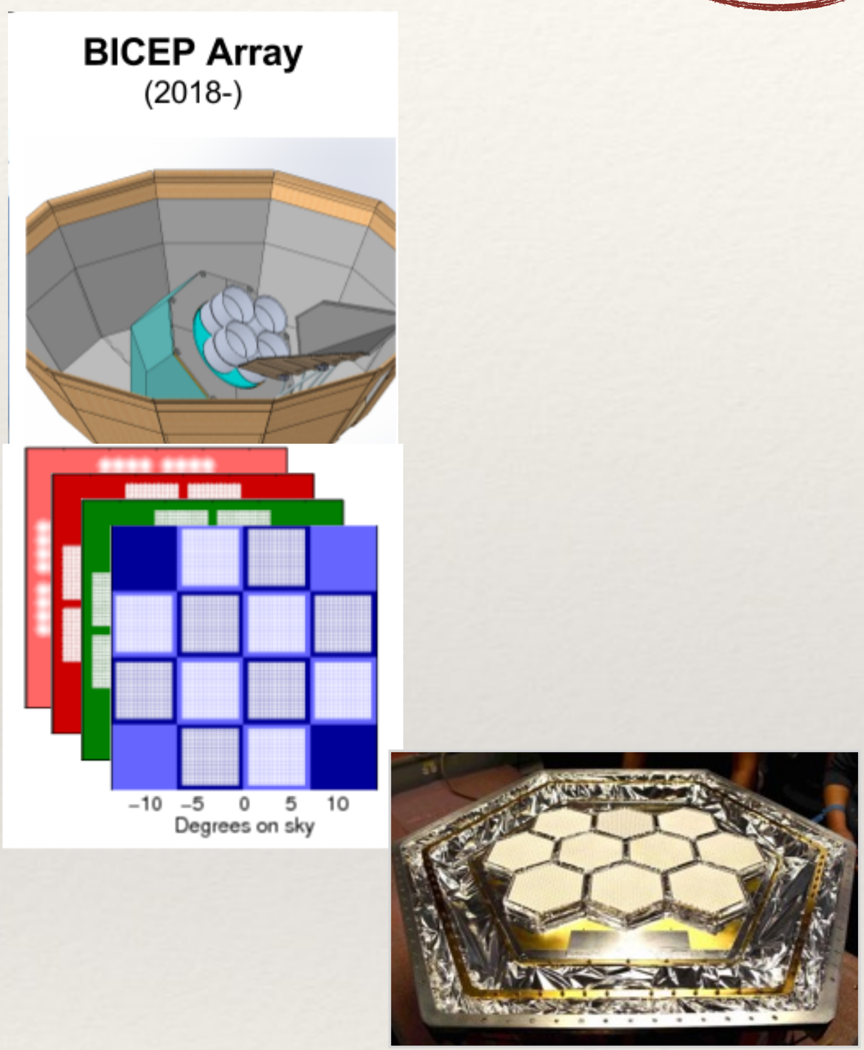


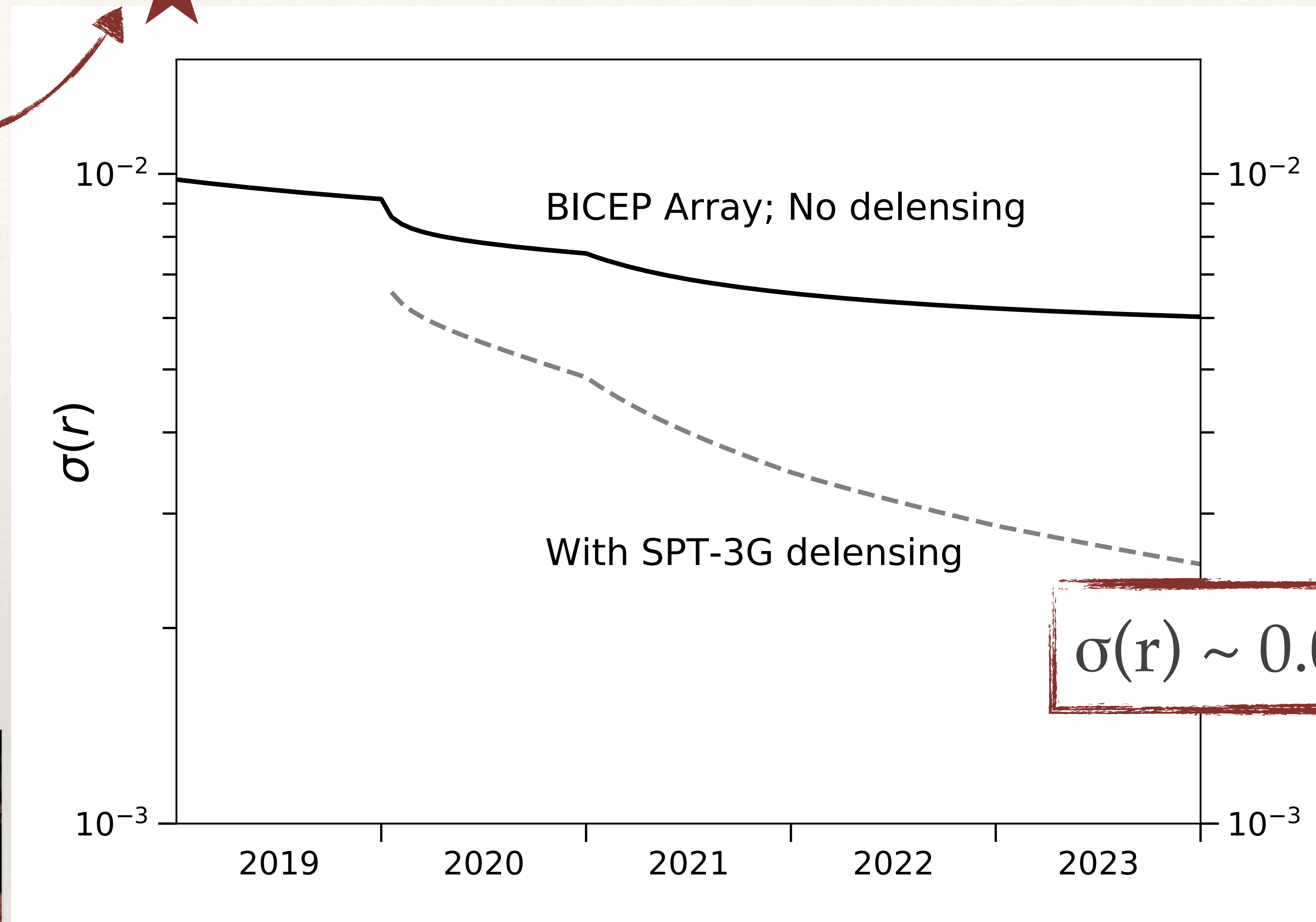

~scales of lenses that source most lensing B-modes

CMB reconstructed ϕ will soon be the best lensing potential tracer for B-mode delensing

# BICEP Array + SPT-3G

Current σ(r) = 0.02
(BK15)




σ(r) ~ 0.003

Assuming BK15 foreground model: $A_{\rm dust}, \alpha_{\rm dust}, \beta_{\rm dust}, A_{\rm sync}, \alpha_{\rm sync}, \beta_{\rm sync}$

σ(r) saturates without delensing even with the addition of 30/40 GHz and 220/280 GHz receivers in BICEP Array

Key takeaways:

1) CMB lensing map from SPTpol survey has S/N > 1 measurements for L < 250 modes;
2) Precise measurement of CMB lensing amplitude (~6%) and will provide relevant constraints for cosmological parameters
3) CMB reconstructed phi will soon be competitive for delensing.
4) BICEP Array + SPT-3G delensing is projected to give $\sigma(r) \sim 0.003$.

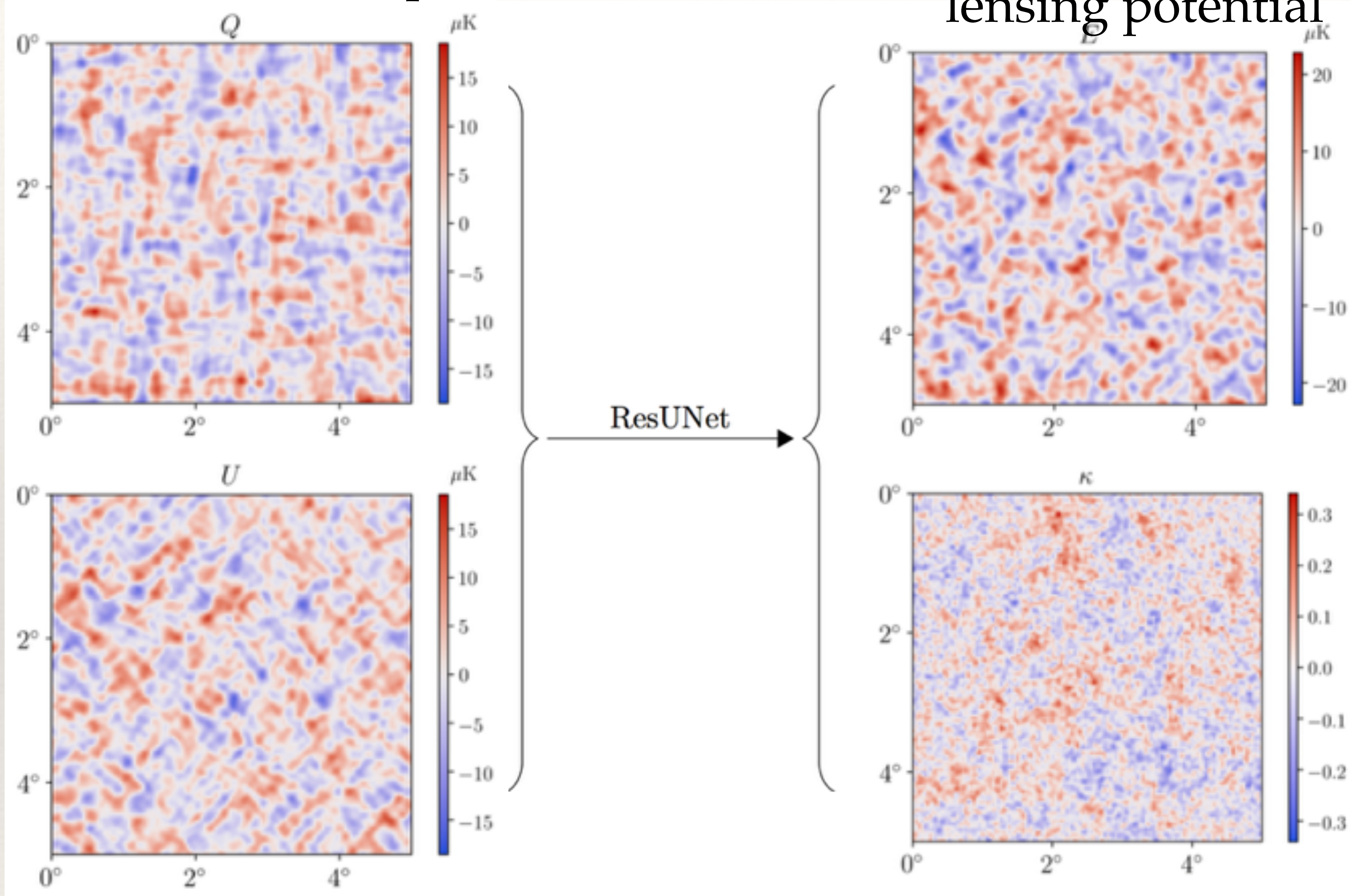
# Neural networks
# for CMB lensing reconstruction

(arXiv: 1810.01483)

# Why neural network?

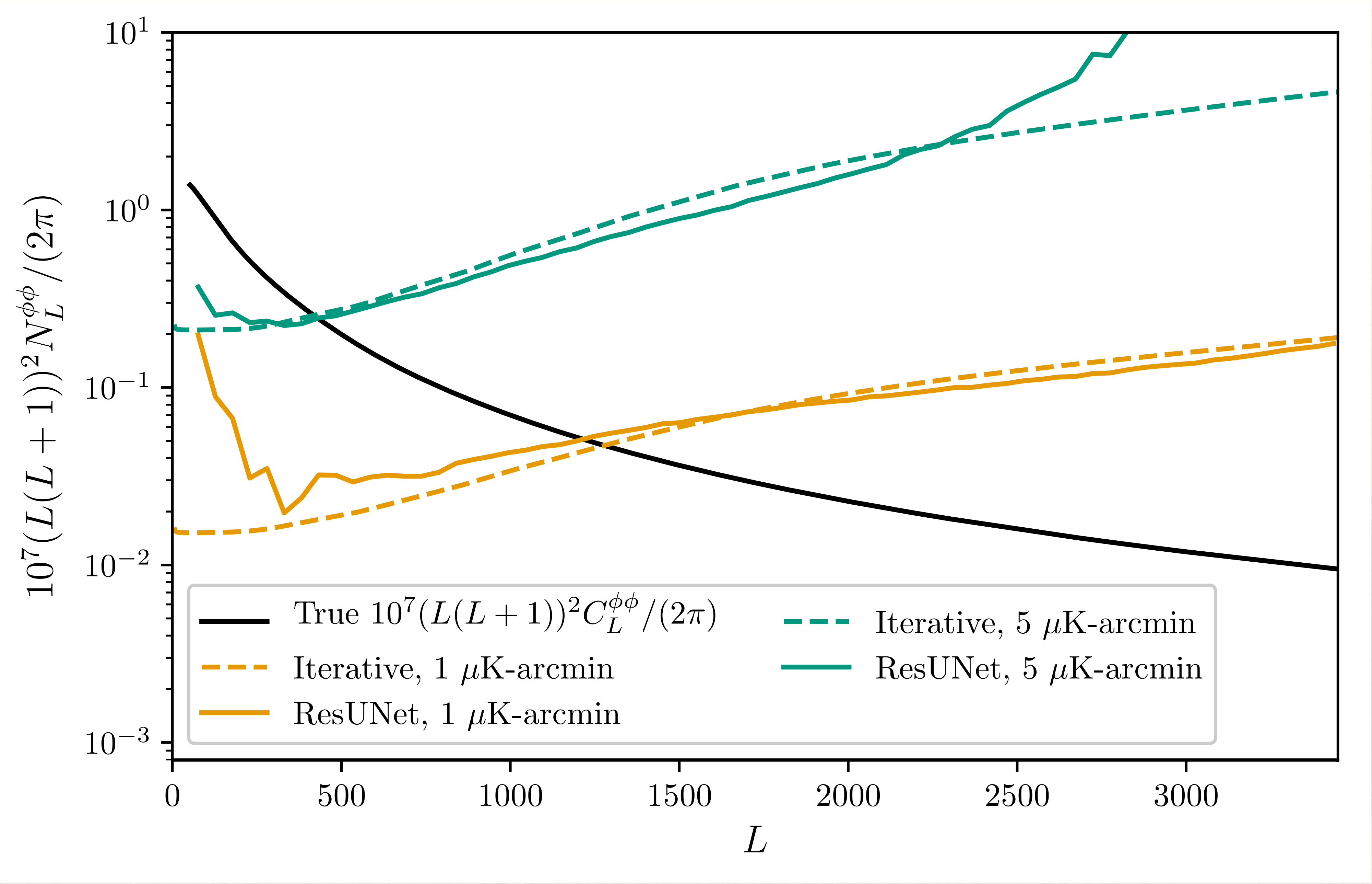
observables:
lensed CMB maps

underlying fields:
primordial CMB,
lensing potential



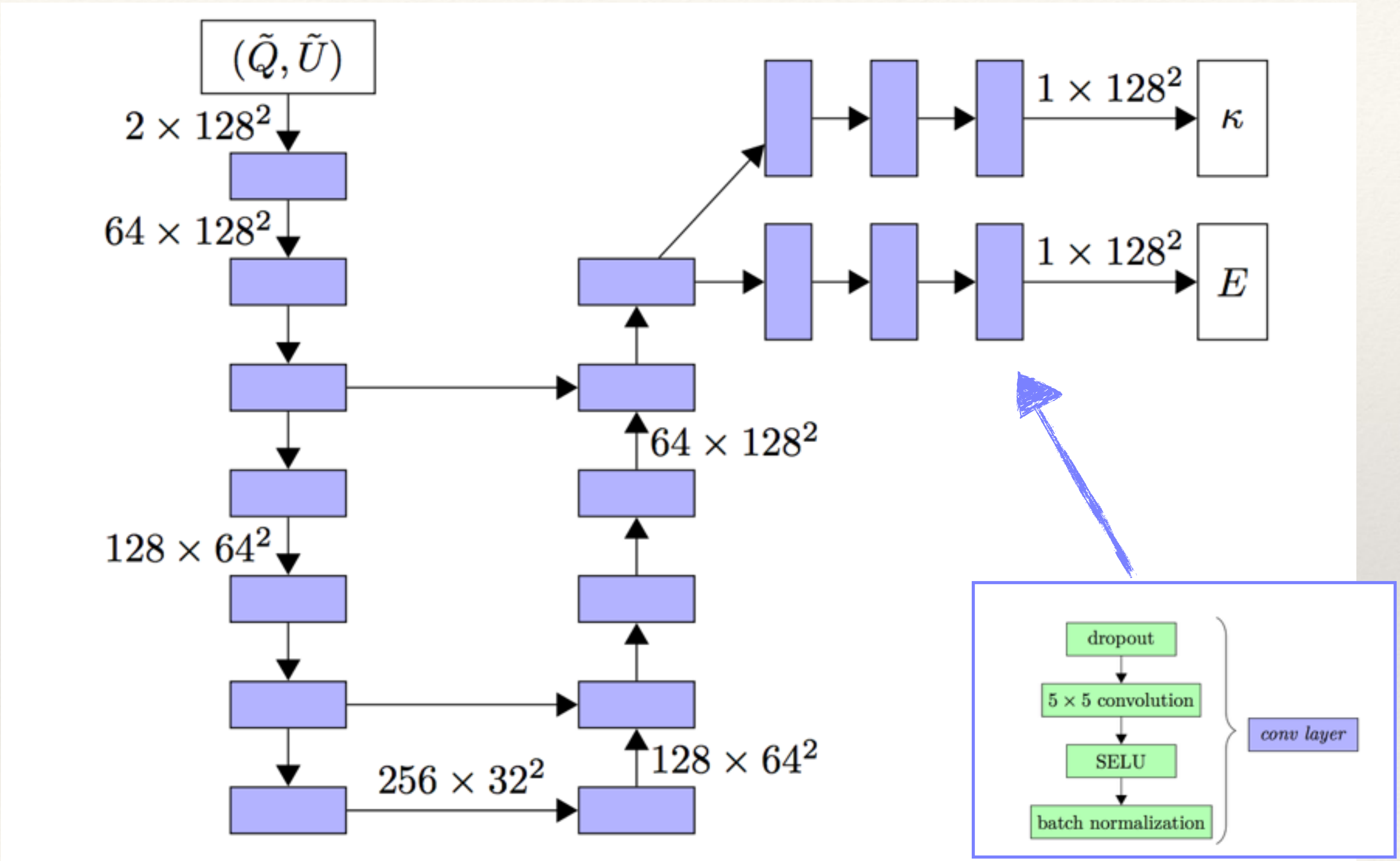

- network needs to perform transformation from one set of images to another; seem to be a good fit for neural networks
- Real need of beyond quadratic estimators to get optimal lensing reconstruction

# It works!




solid lines: neural net
dashed lines: maximum likelihood approx.
1uK-arcmin: ~CMB-S4 level noise

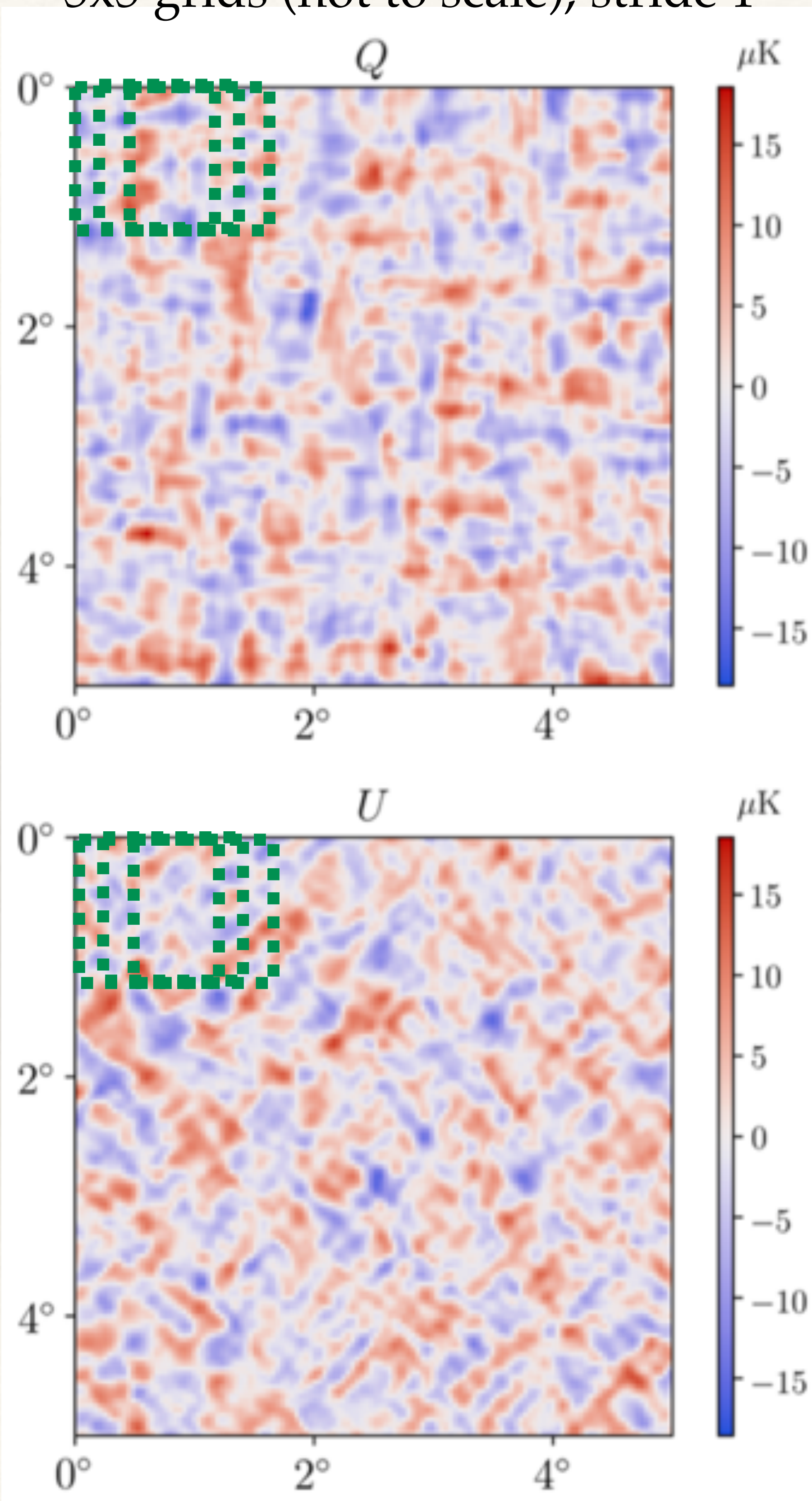
# Network architecture




- 11200 sets of Q/U, E/K sim maps; 80:10:10 training:validation:test sets

# Convolution layer

5x5 grids (not to scale), stride 1

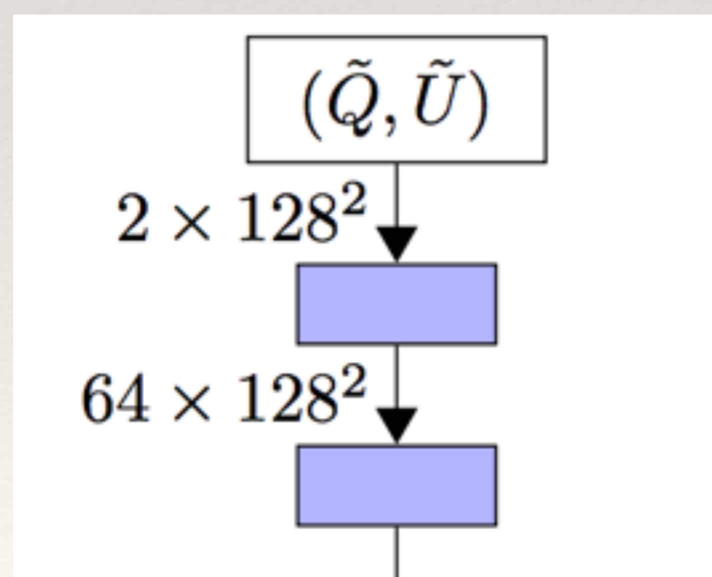



For each 2x5x5 grid's 50 inputs,
it outputs one value:

$$q = \sum_{j=1}^{2} \sum_{i=1}^{25} A_i^j p_i^j$$

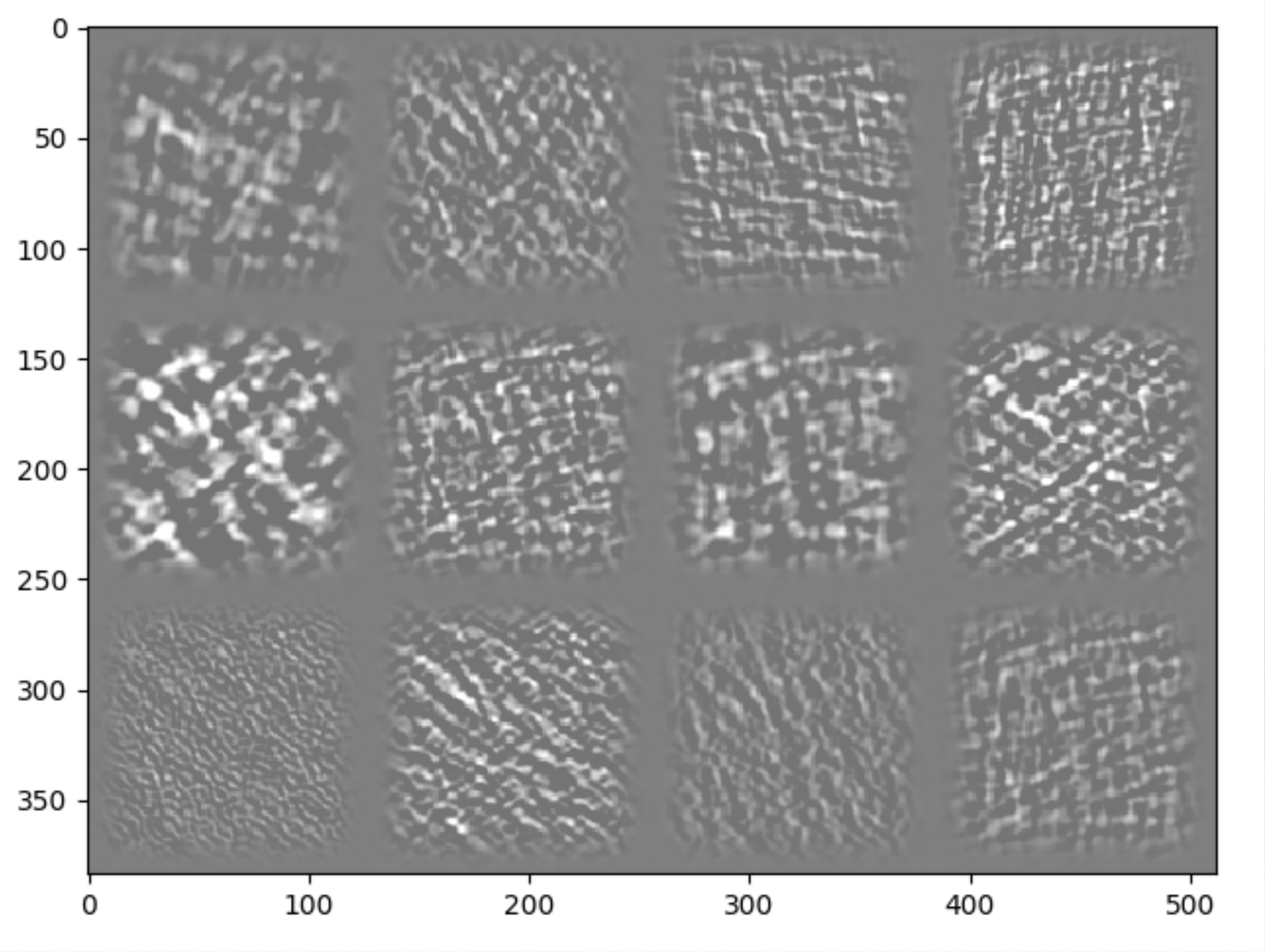
$A_i^j$ is what training determines.

There are 2x128$^2$ of these 5x5 grids
from one pair of Q/U maps.
It outputs a 1x128$^2$ image.

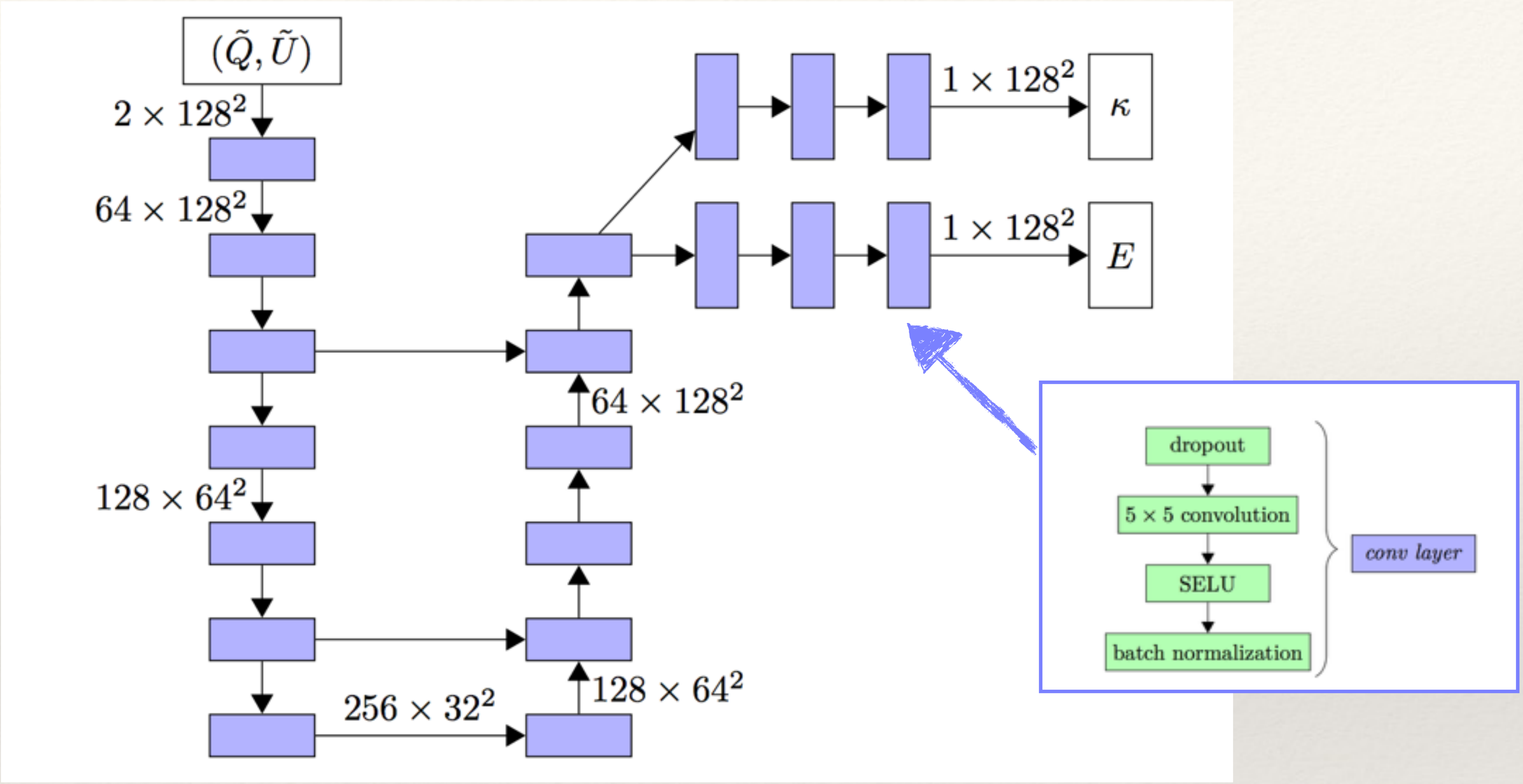



For this first step,
we create 64 versions
of the 1x128$^2$ image

# The first 12 of the 64 outputs at the first step

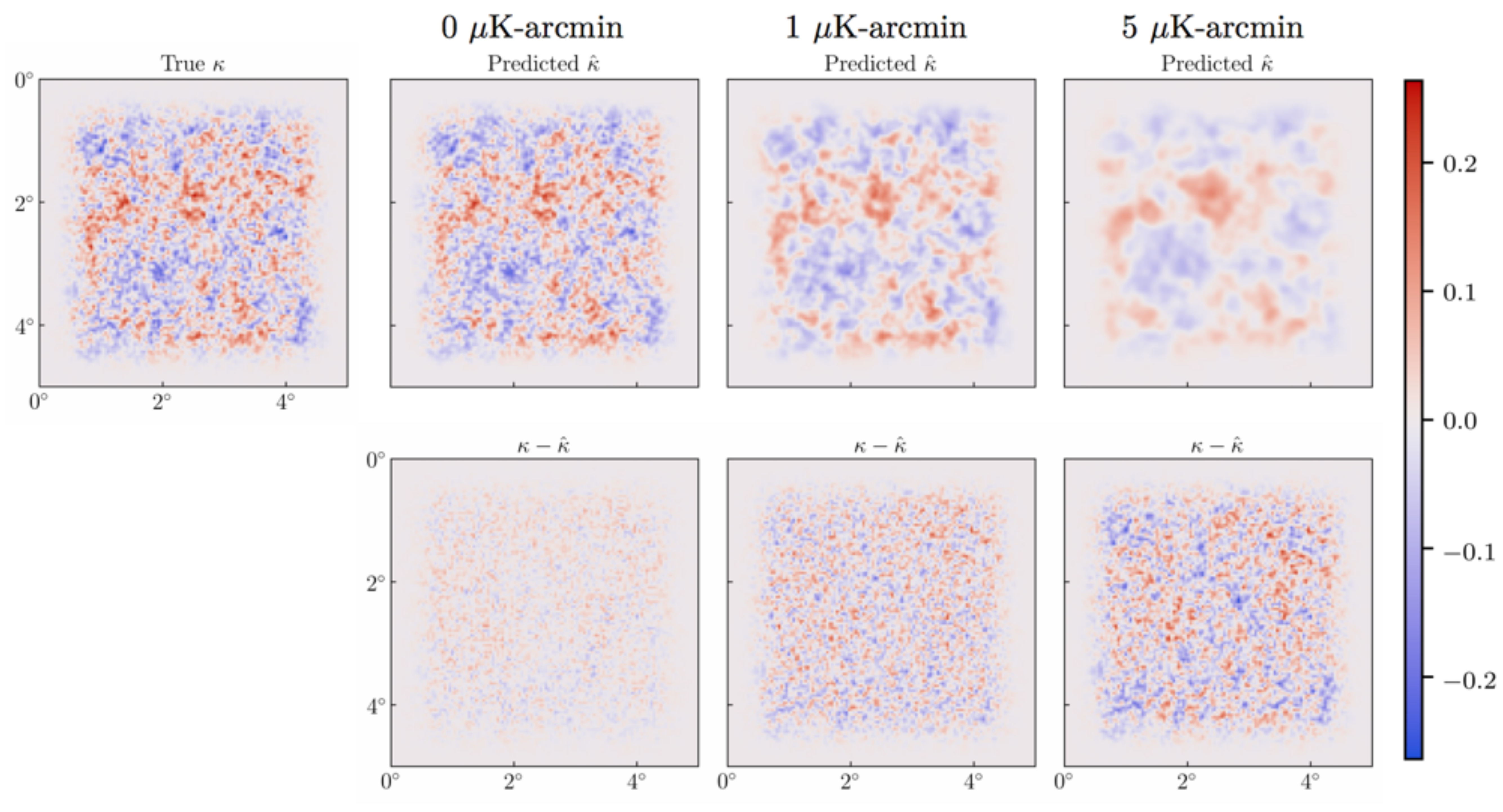


# Network architecture

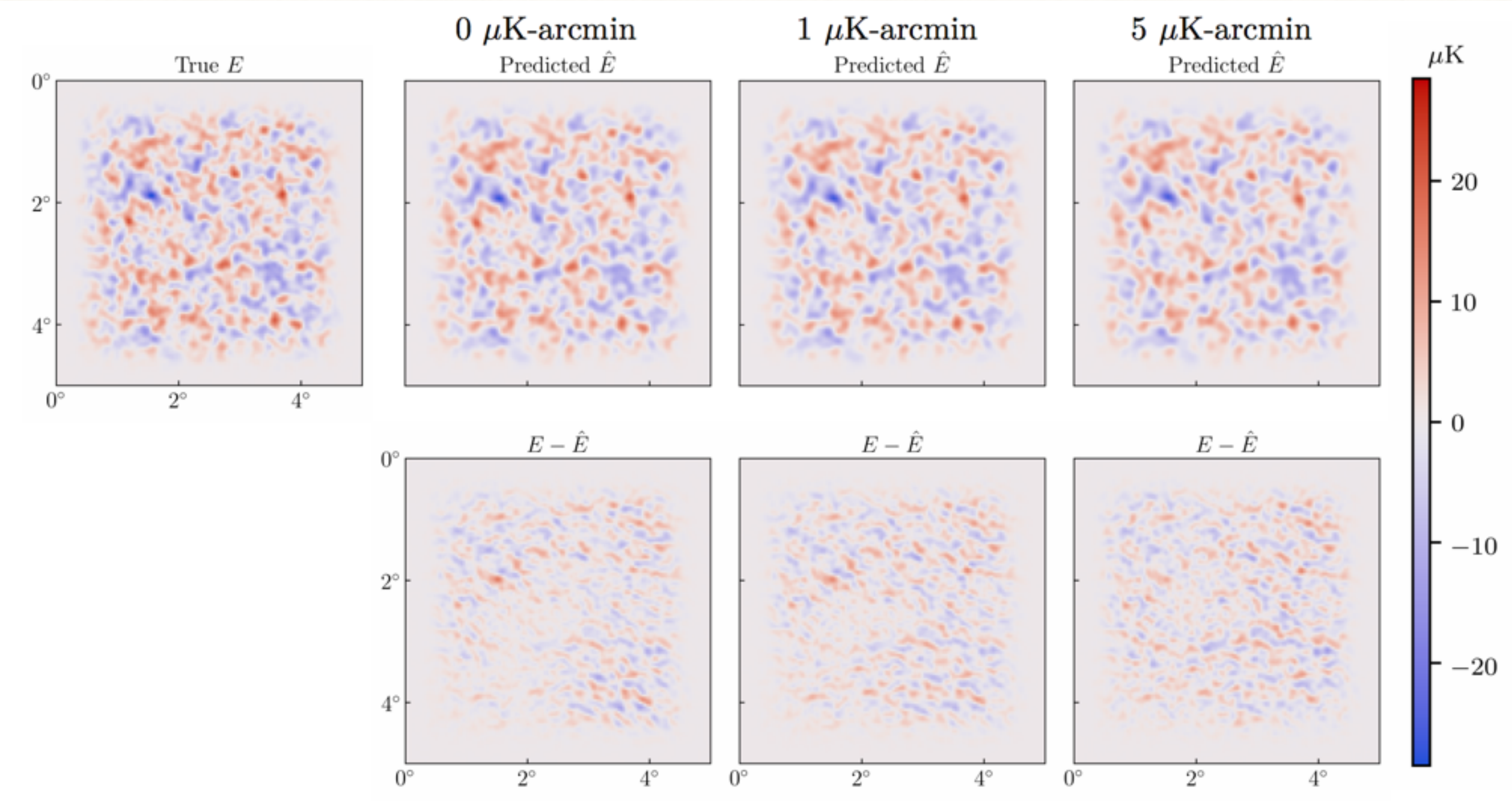



- The loss function is MSE (mean square error) between the output E/K and the true E/K; choose $A_i^j$ to minimize MSE.
- Residual UNet; "residual connections" at before each dimension changing step; "skip connections" across the layers with same dimensions

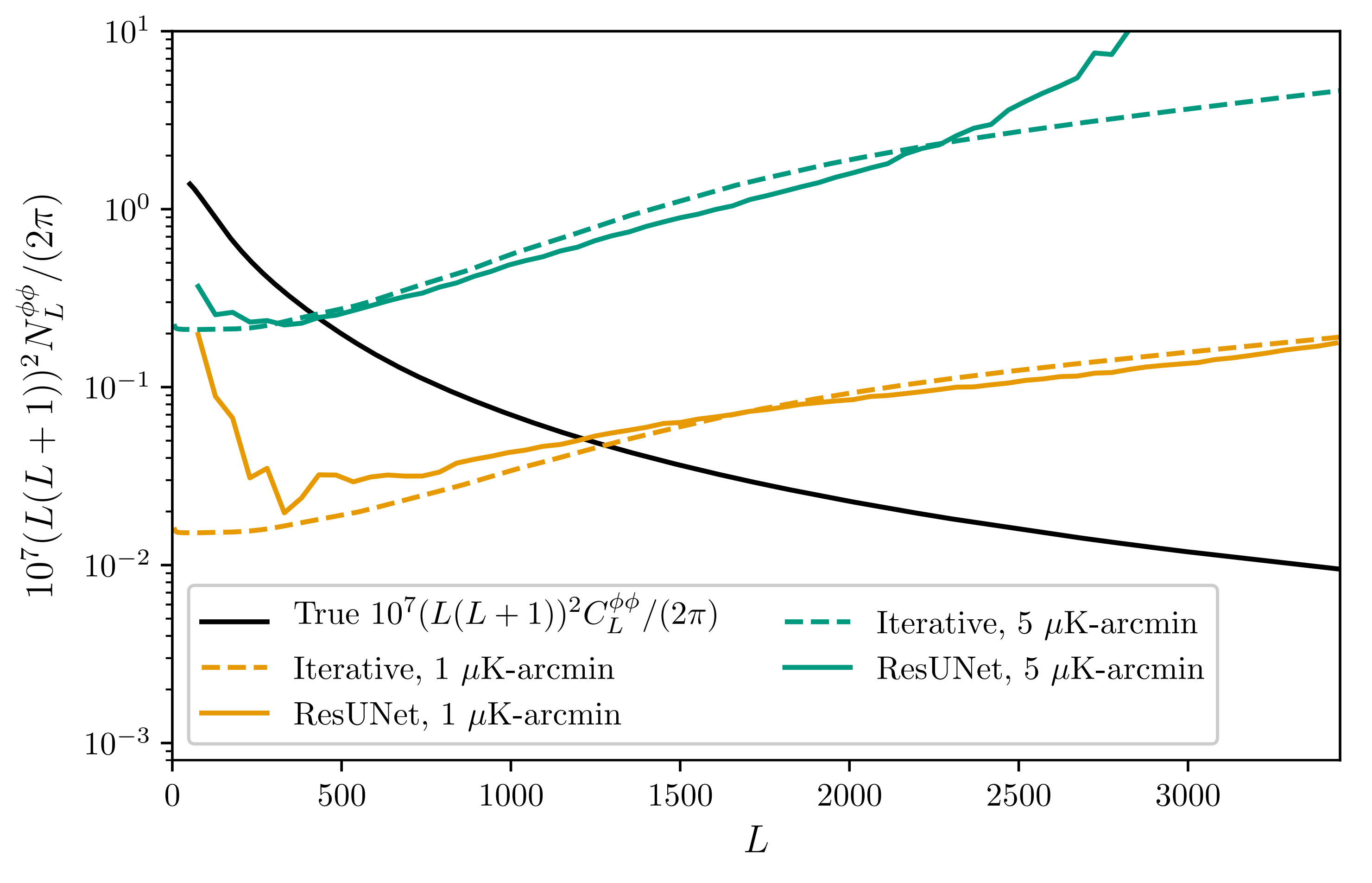
# Lensing field recovery



# Unlensed E-mode recovery



# Compare to "physics-ful" methods




solid lines: neural net
dashed lines: maximum likelihood approx.
1uK-arcmin: ~CMB-S4 level noise

- The NN approach doesn't completely recover the input
- We quantify that decorrelation as noise in our reconstruction
- Standard reconstruction has noise terms due to spurious correlations of random gaussian fields, etc.
- Reaching similar levels of noise as maximum likelihood methods!

# Tests

Is the network really sensitive to lensing??

- Null test
- Sensitive to differences in input $\Omega_M$

How sensitive are the outputs to the initialization randomness?

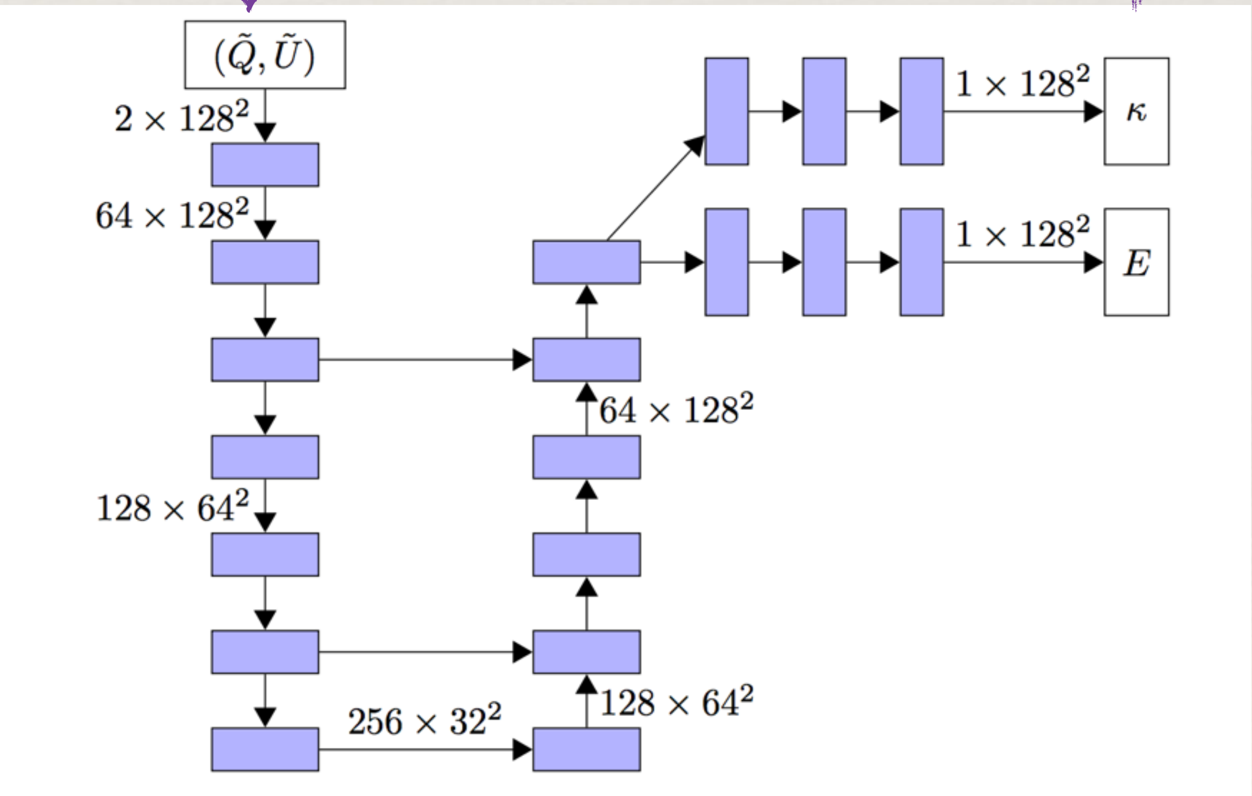
- Randomness due to initialization

Toy fit for cosmology

- Fit $\Omega_M$

1) maps that are not lensed
2) maps generated with more/less $\Omega_M$

?



Key takeaways:

1) The network $\kappa$ recovery's S/N is similar to maximum-likelihood methods;
2) Network is sensitive to changes in cosmology;
3) Lots to explore in understanding how the network extracts information.

# Summary

- Lensing variance will dominate σ(r) in the next few years; σ(r) is currently dominated by foreground and instrumental noise uncertainties.
  - First demonstration of σ(r) reduction underway.
- Lensing potential reconstructed from CMB maps will be competitive ϕ tracers for delensing B-modes by SPT-3G era.
  - SPTpol lensing potential is amongst the highest S/N per mode measurement to date.
- Future low-noise experiments will benefit from beyond quadratic estimator lensing/delensing. Neural network is a viable technique.

# Thank you for listening!