

CMB Lensing Power Spectrum without Noise Bias

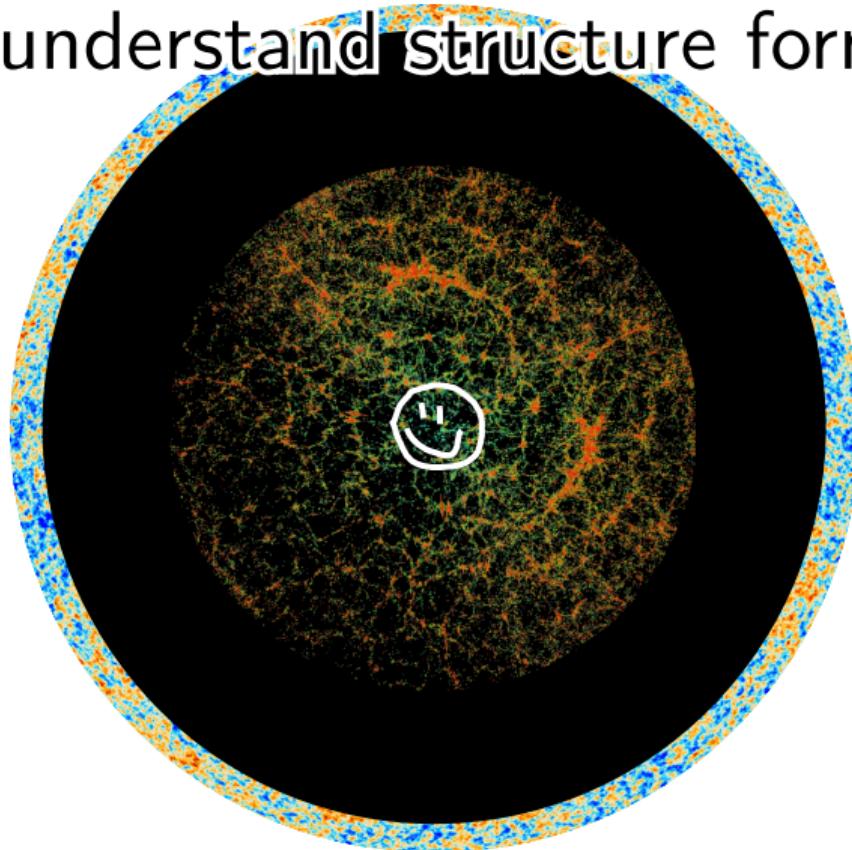
Delon Shen

Stanford Cosmology Seminar — April 1, 2024
[arxiv:2402.04309](https://arxiv.org/abs/2402.04309) with Manu Schaan and Simone Ferraro

Initial density perturbation in matter collapses gravitationally to form cosmic structure

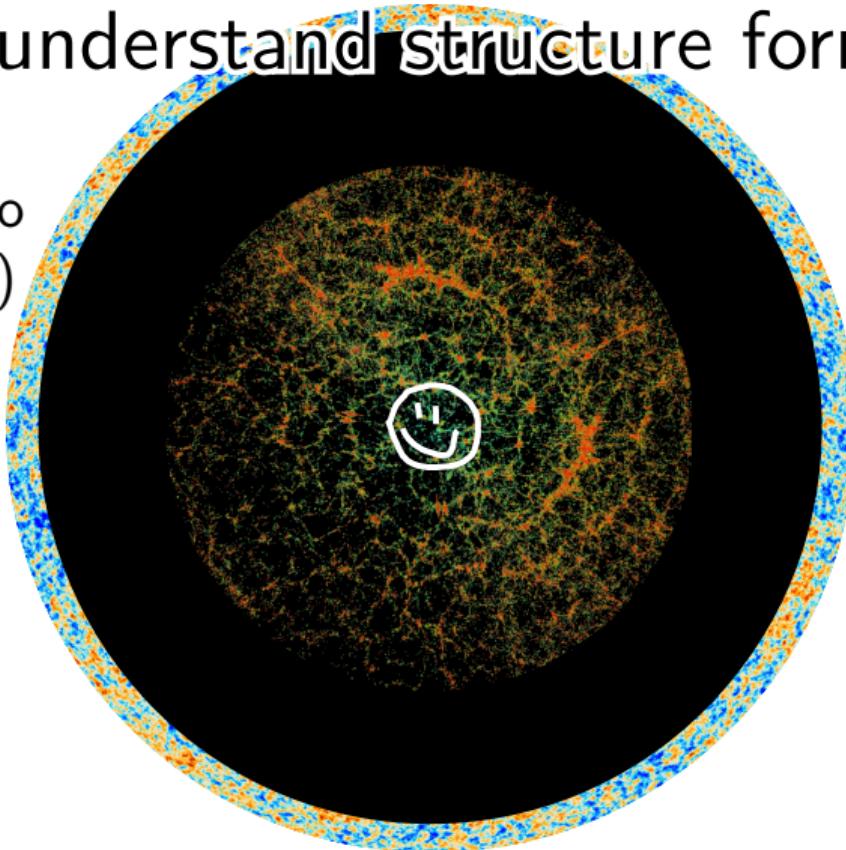
The dynamics of structure formation (e.g. how clumpy matter is over time) carries a lot of information about physics ($\sum m_\nu$, *nature of dark energy*, *properties of dark matter*, . . .)

Do we understand structure formation?



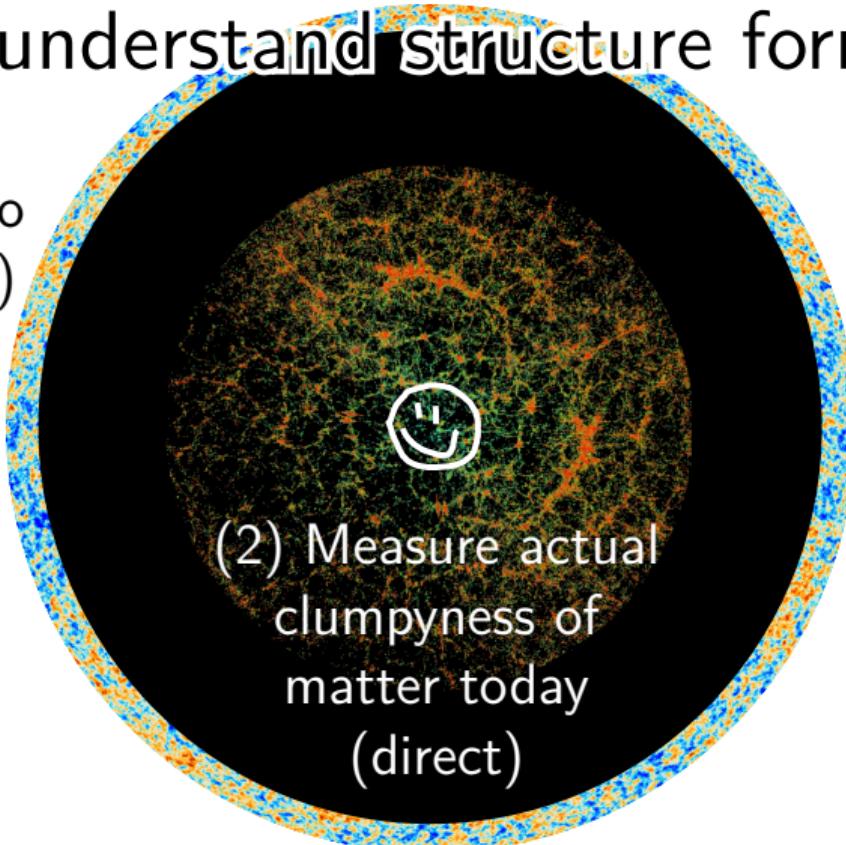
Do we understand structure formation?

(1) Fit Λ CDM to
CMB ($z \approx 1100$)
then predict
clumpiness of
matter today
(indirect)

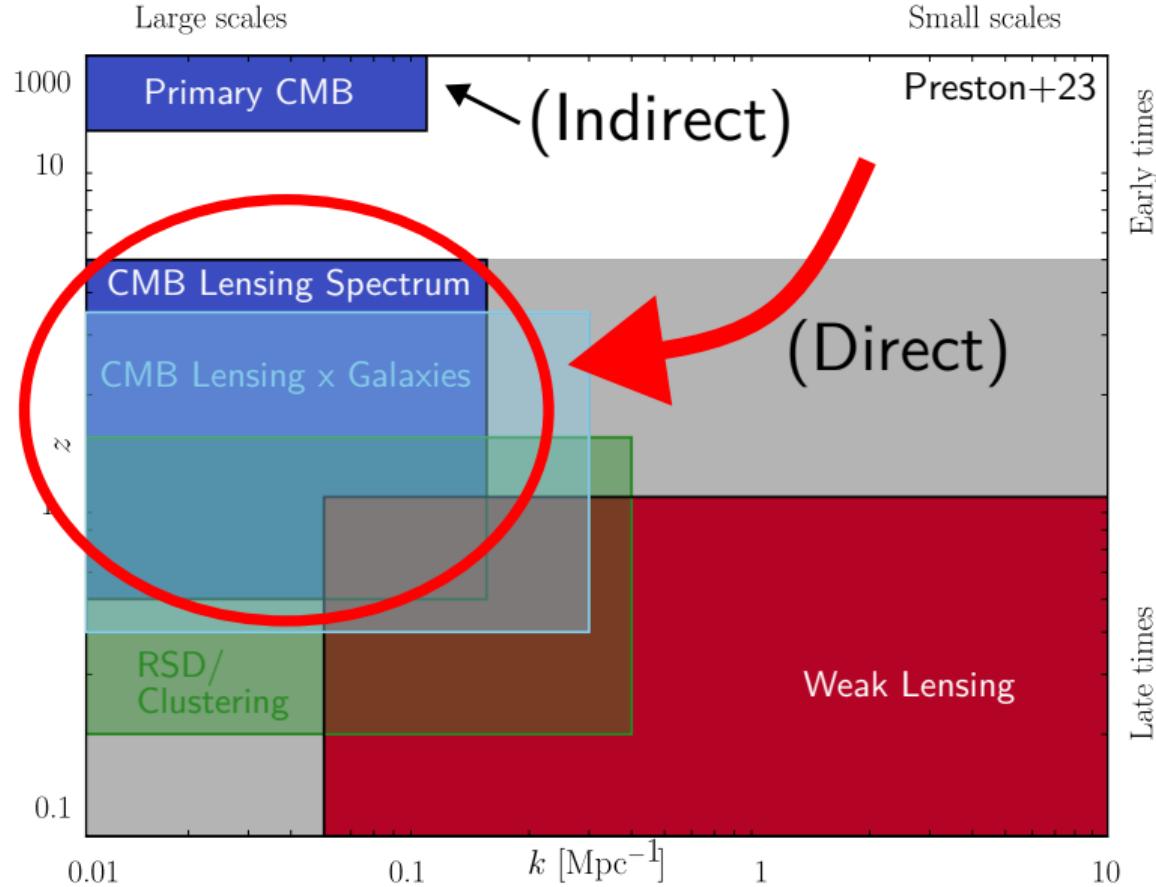


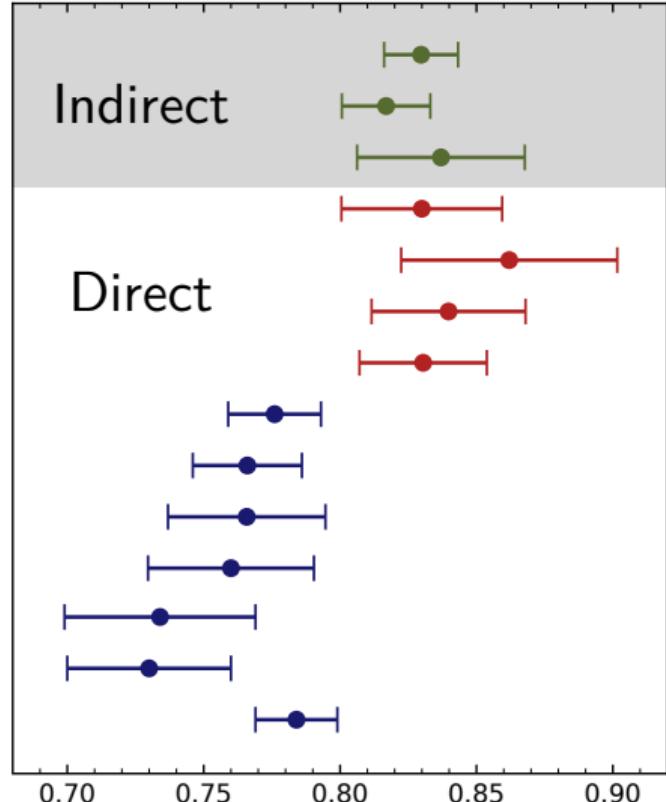
Do we understand structure formation?

(1) Fit Λ CDM to
CMB ($z \approx 1100$)
then predict
clumpiness of
matter today
(indirect)



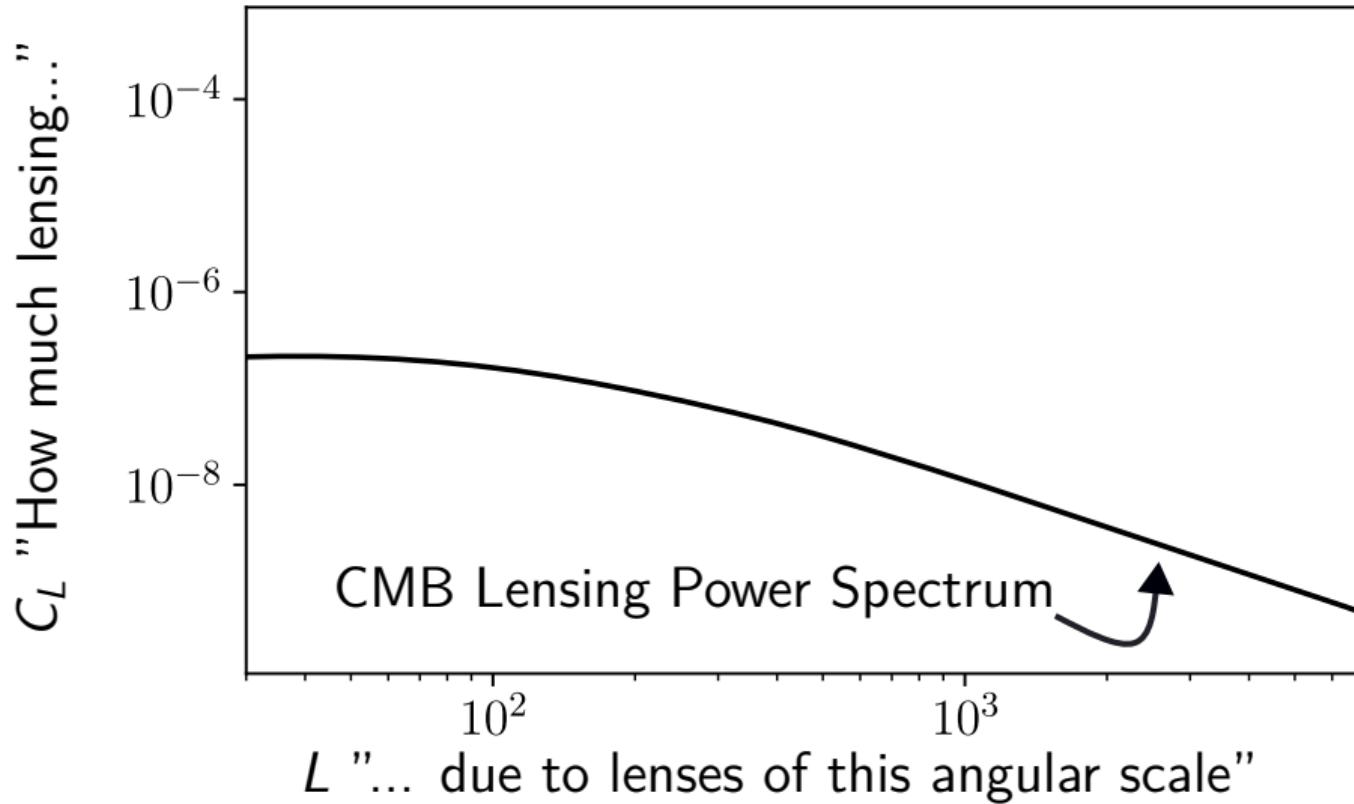
CMB lensing
probes different
scales and
redshifts from
other direct
probes!

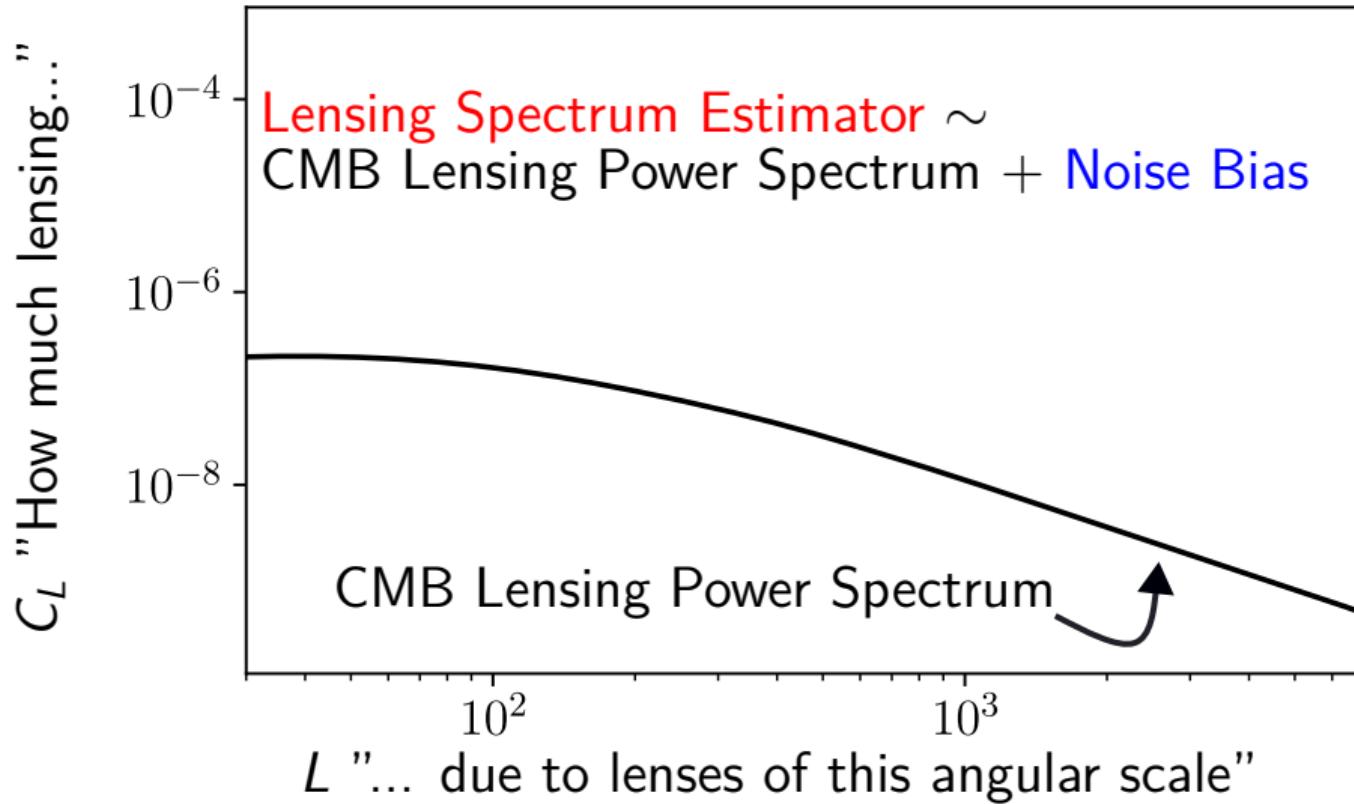


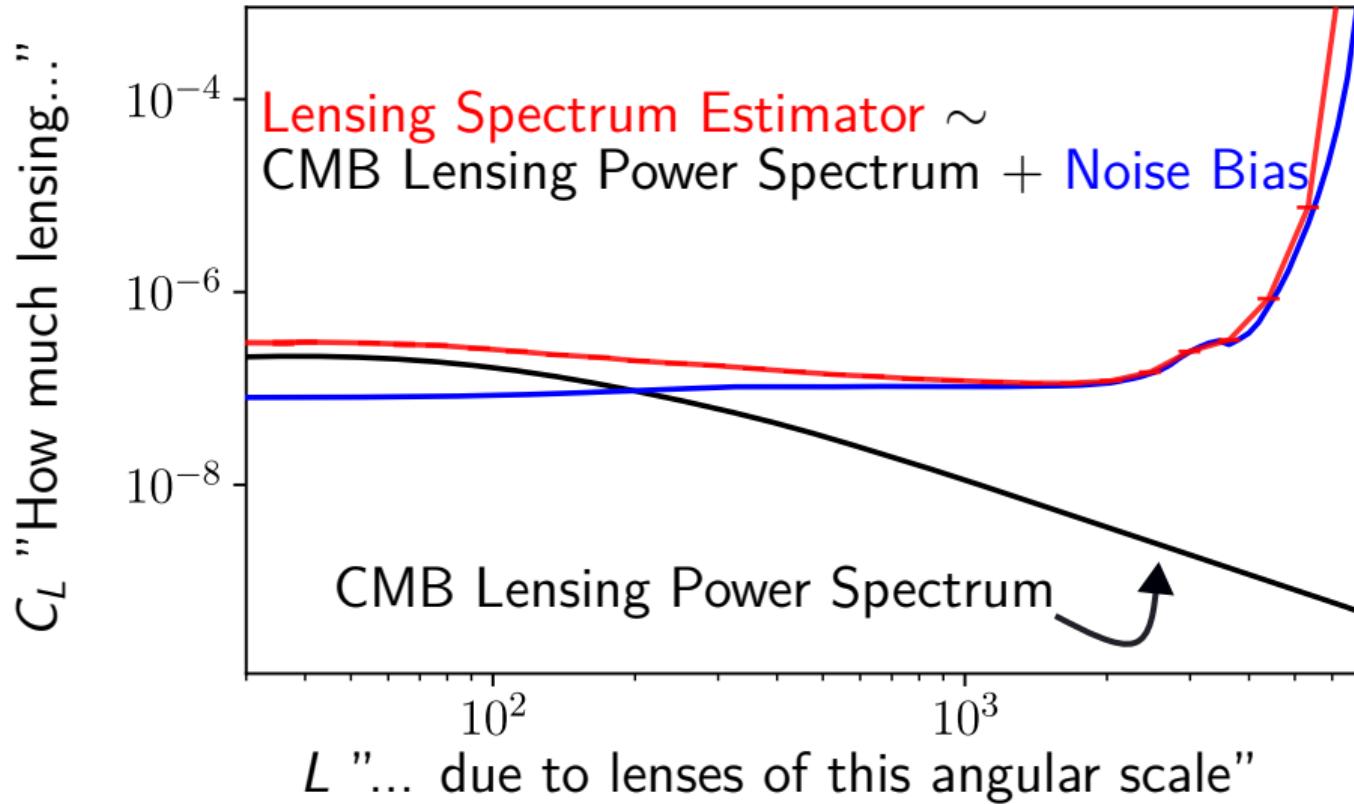


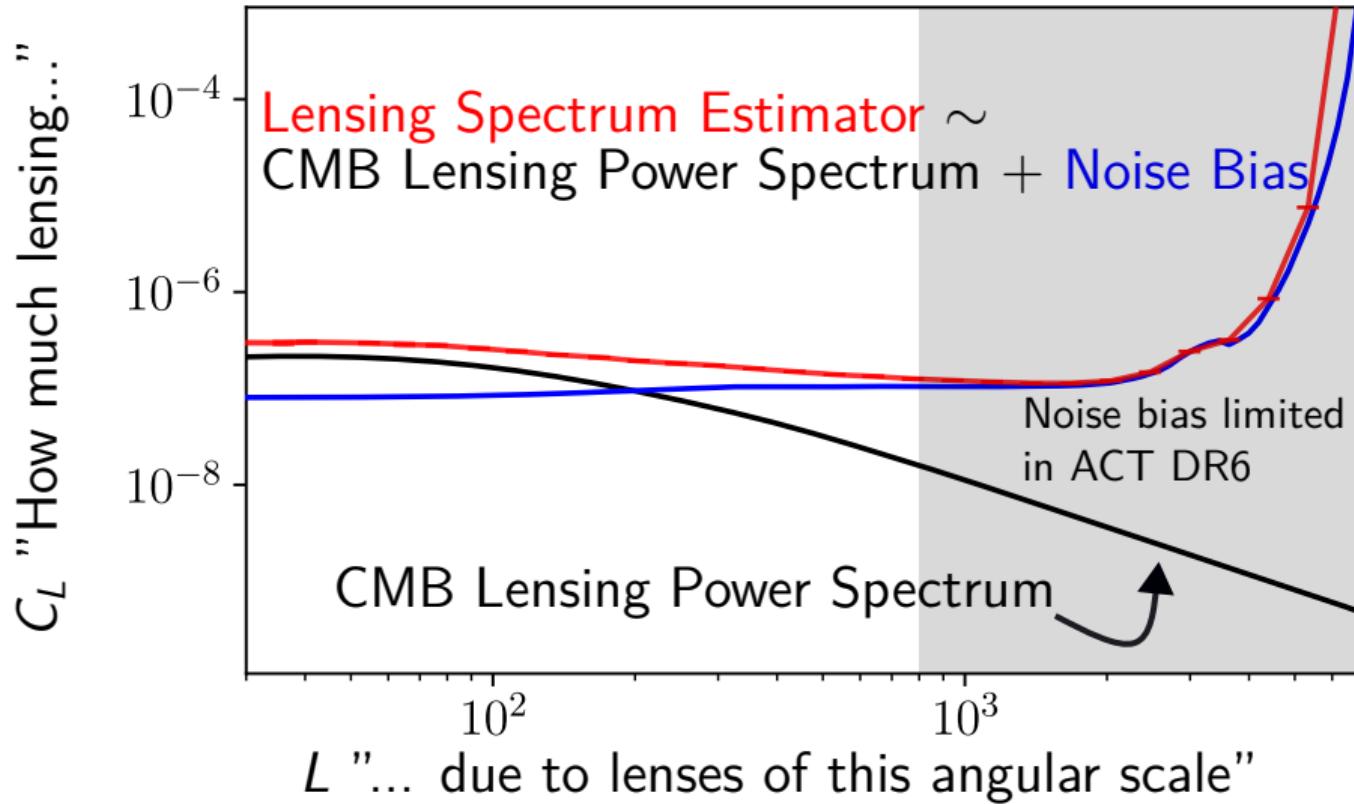
S8 "Matter Clumpiness"

Madhavacheril+23



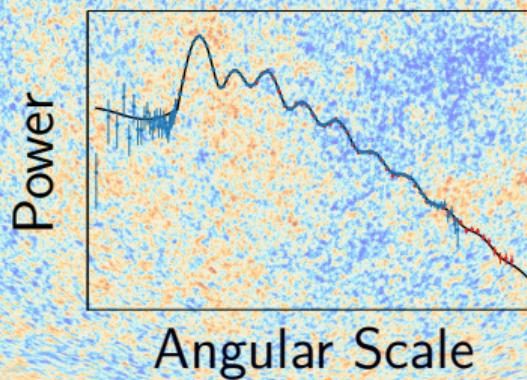


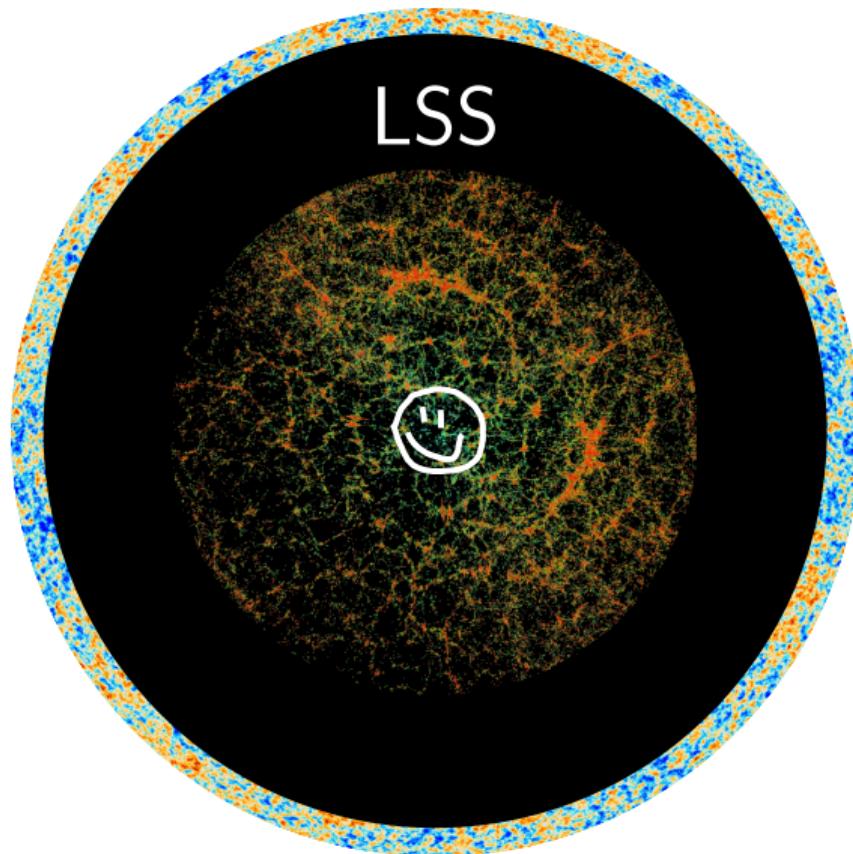




$T(\theta)$: Fluctuations

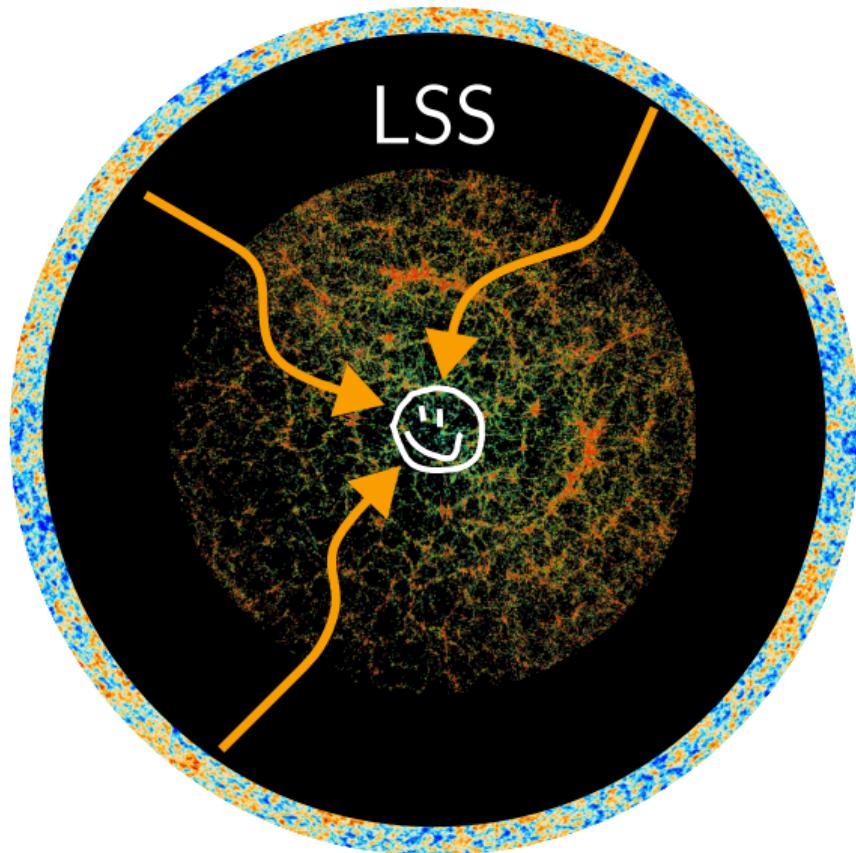
Statistically **Gaussian, Homogeneous, Isotropic**





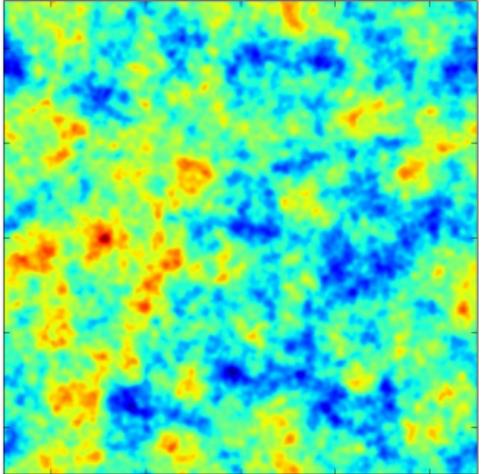
CMB

LSS

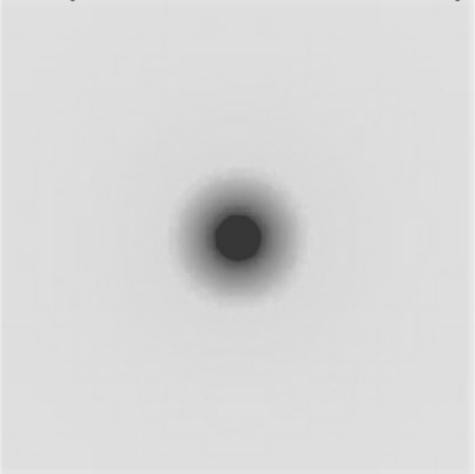


CMB
lensed by
LSS

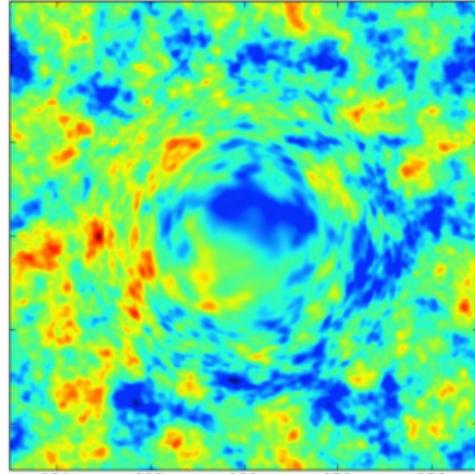
Unlensed CMB



$\nabla(\text{Lensing Potential})$



Lensed CMB

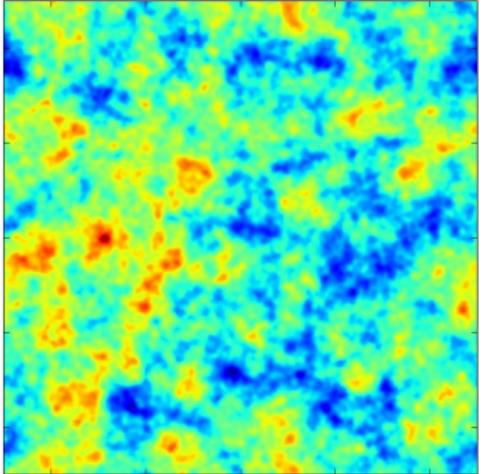


+

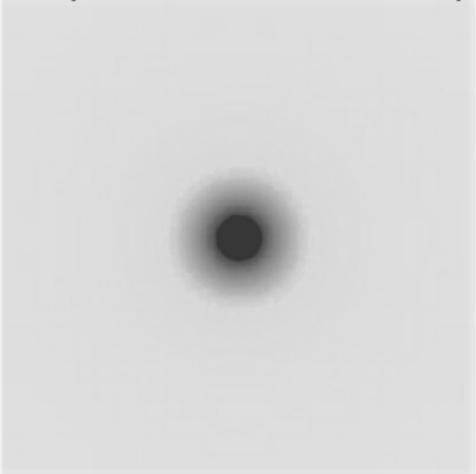
=

Sherwin

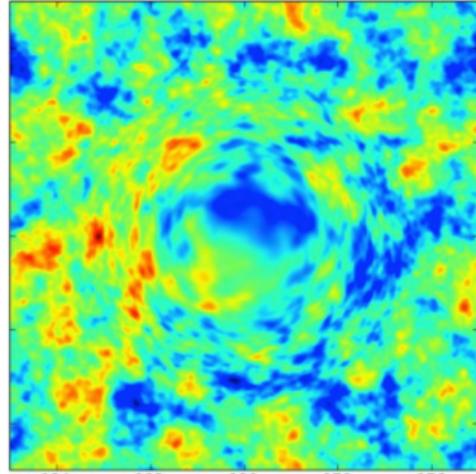
Unlensed CMB



$\nabla(\text{Lensing Potential})$



Lensed CMB



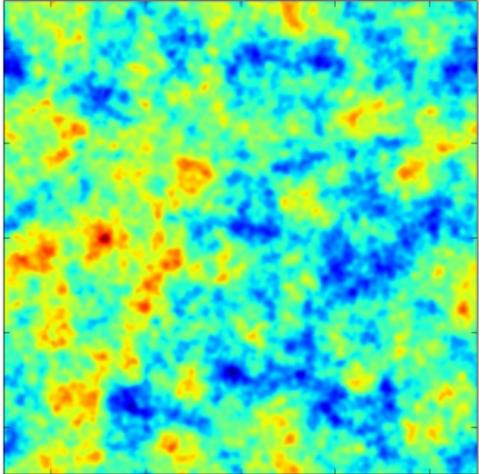
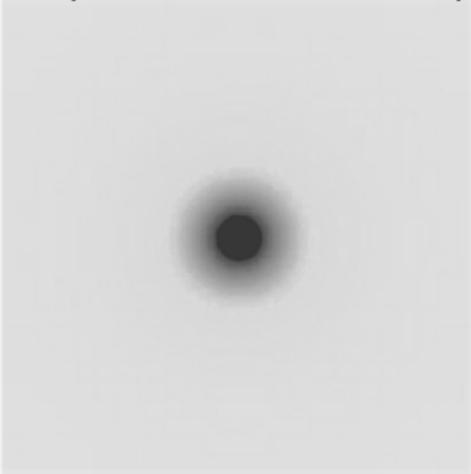
+

=

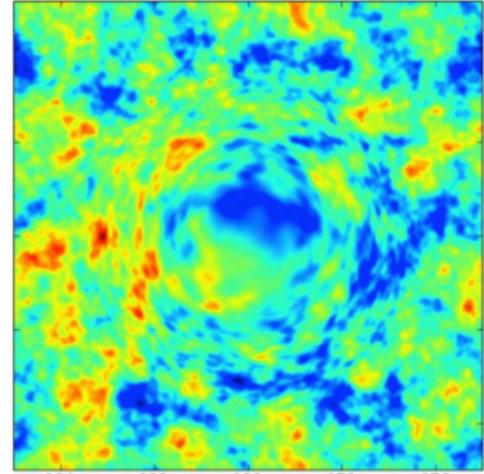
Sherwin

$$T^{\text{lensed}}(\hat{n}) = T^{\text{unlensed}}(\hat{n} + \nabla\{\text{Lensing Potential}(\hat{n})\})$$

Unlensed CMB

 $\nabla(\text{Lensing Potential})$ 

Lensed CMB



+

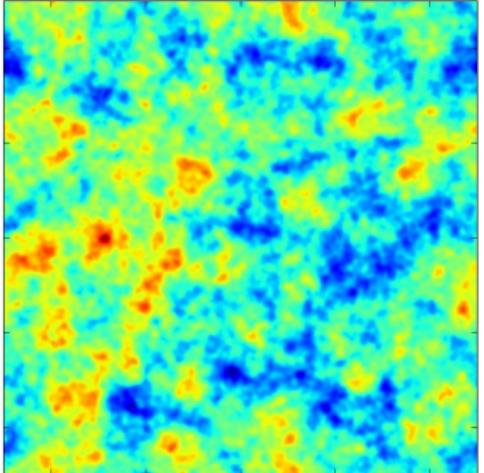
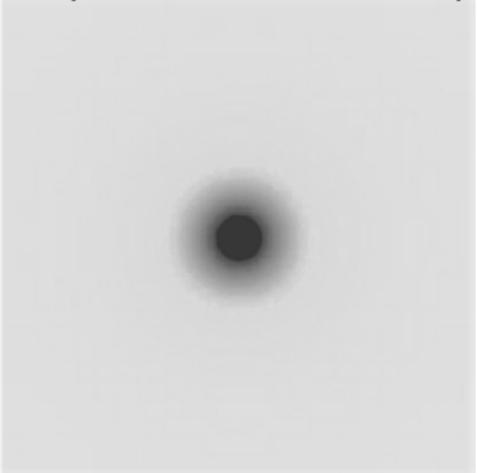
=

Sherwin

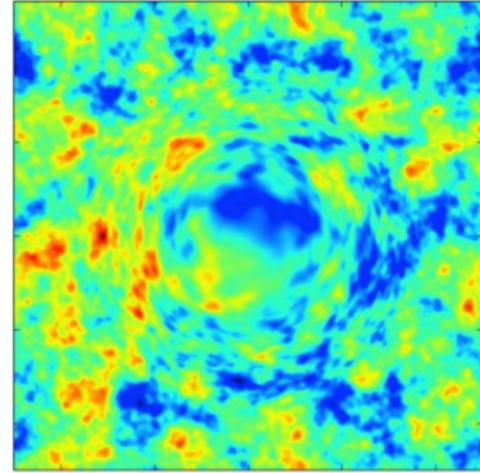
$$T^{\text{lensed}}(\hat{\mathbf{n}}) = T^{\text{unlensed}}(\hat{\mathbf{n}} + \nabla\{\text{Lensing Potential}(\hat{\mathbf{n}})\})$$

$$\text{Lensing Potential}(\hat{\mathbf{n}}) \sim \int_{\text{Line of Sight}} (\text{Redshift Kernel}) \times (\text{Matter Fluctuations})$$

Unlensed CMB

 $\nabla(\text{Lensing Potential})$ 

Lensed CMB



+

=

Sherwin

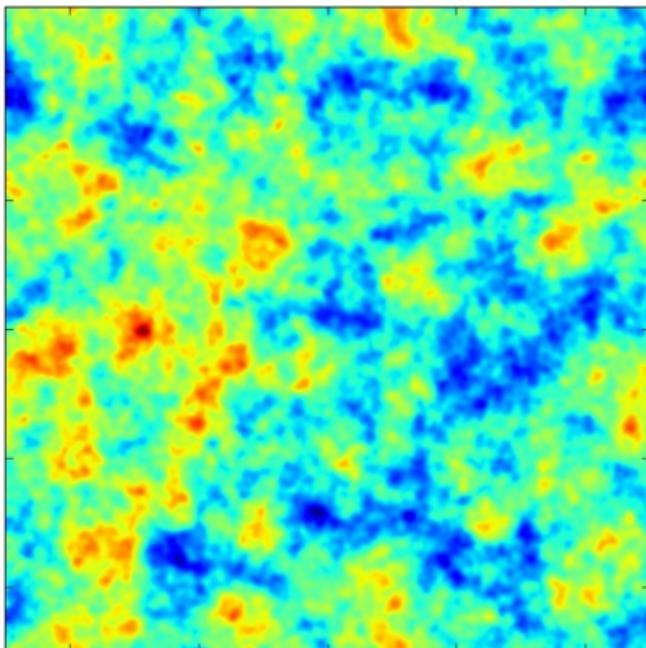
$$T^{\text{lensed}}(\hat{\mathbf{n}}) = T^{\text{unlensed}}(\hat{\mathbf{n}} + \nabla\{\text{Lensing Potential}(\hat{\mathbf{n}})\})$$

$$\text{Lensing Potential}(\hat{\mathbf{n}}) \sim \int_{\text{Line of Sight}} (\text{Redshift Kernel}) \times (\text{Matter Fluctuations})$$

$$\text{Lensing Power Spectrum} \sim \iint (\text{Redshift Kernel})^2 \times (\text{Matter Power Spectrum})$$

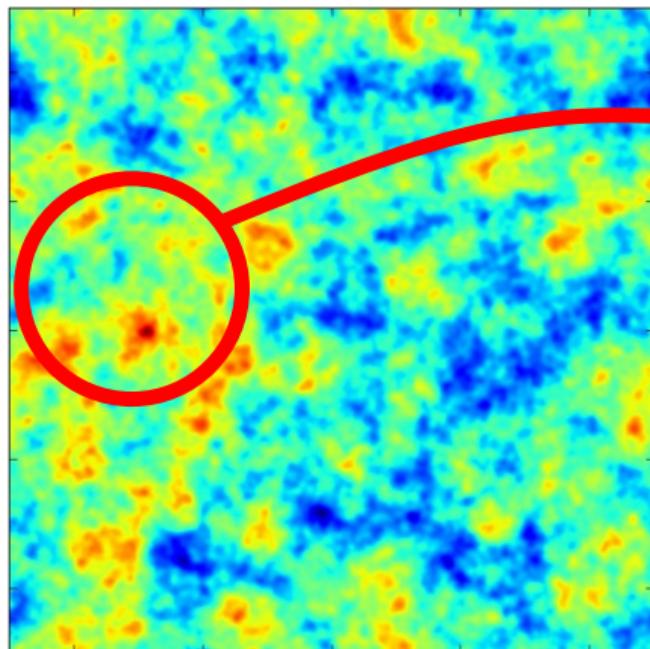
We can estimate lensing potential since
lensing breaks symmetry of CMB

We can estimate lensing potential since
lensing breaks symmetry of CMB

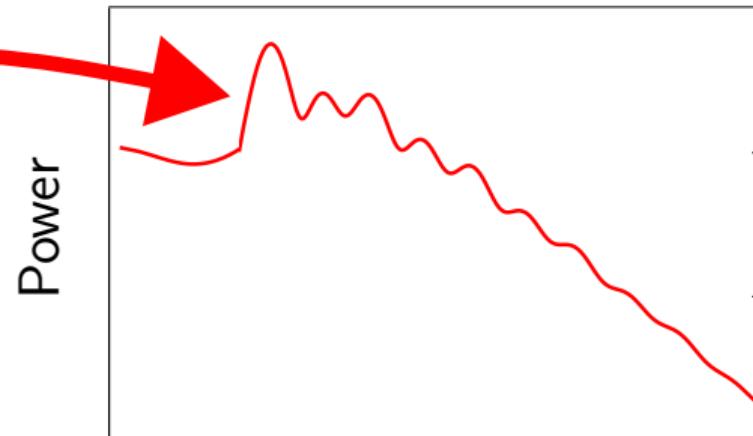


Unlensed CMB: **Statistically Homogeneous**

We can estimate lensing potential since
lensing breaks symmetry of CMB

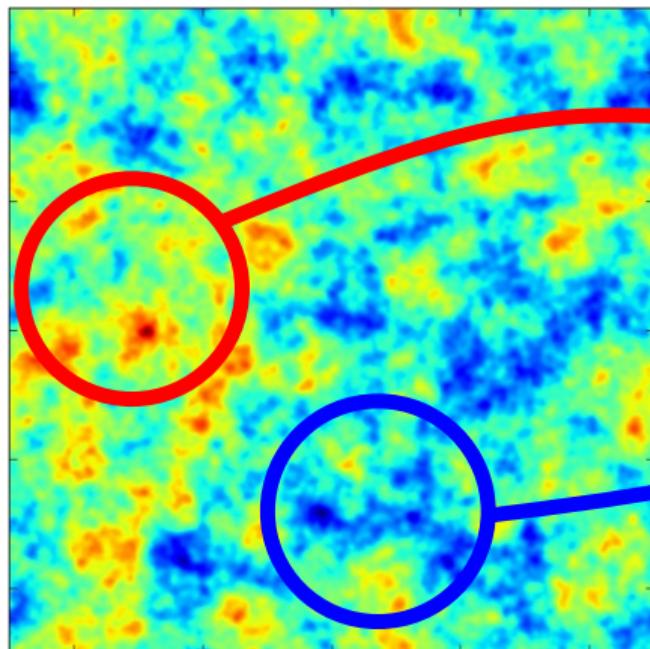


Unlensed CMB: **Statistically Homogeneous**

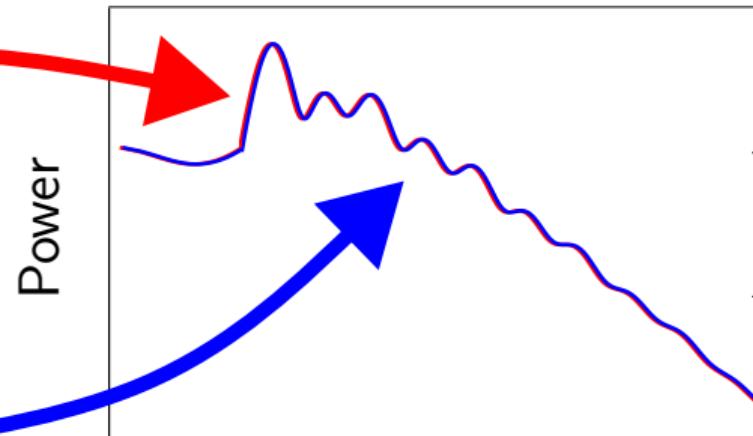


Angular Scale

We can estimate lensing potential since
lensing breaks symmetry of CMB

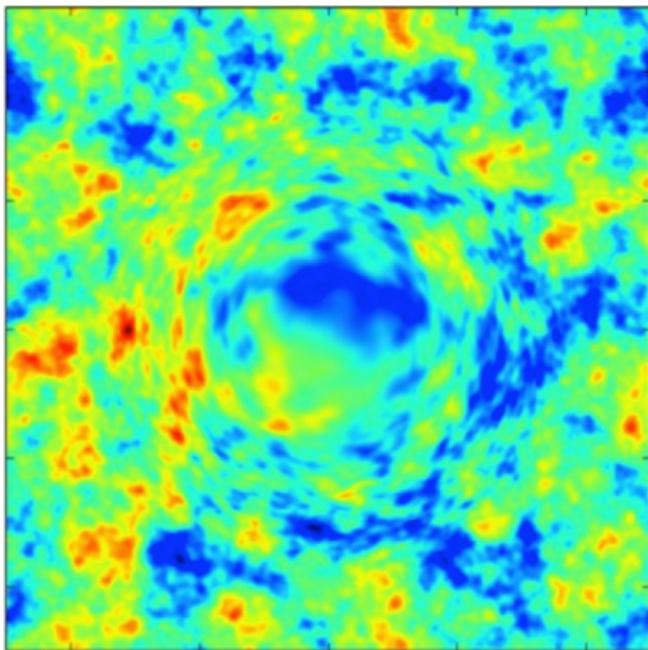


Unlensed CMB: **Statistically Homogeneous**



Angular Scale

We can estimate lensing potential since
lensing breaks symmetry of CMB



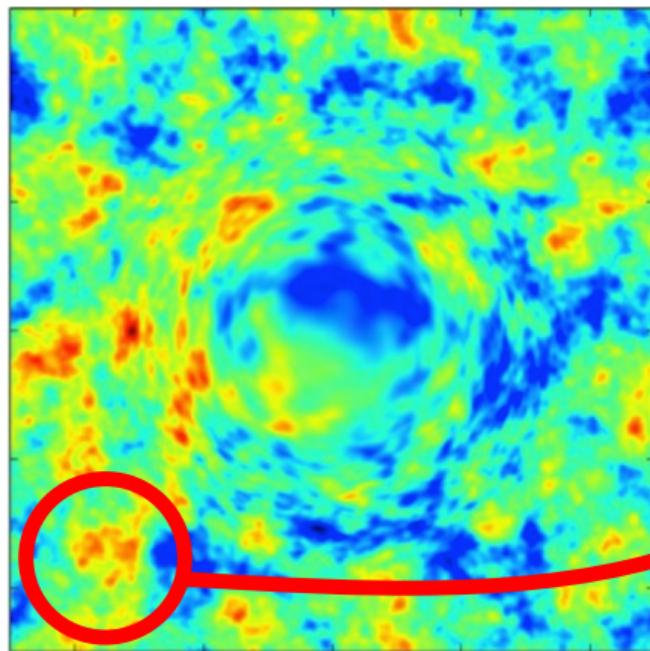
Power

Lensed CMB: **Statistically Homogeneous**

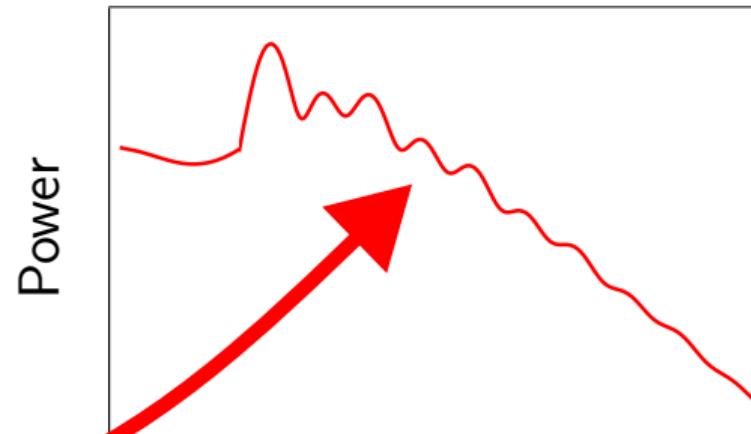


Angular Scale

We can estimate lensing potential since
lensing breaks symmetry of CMB

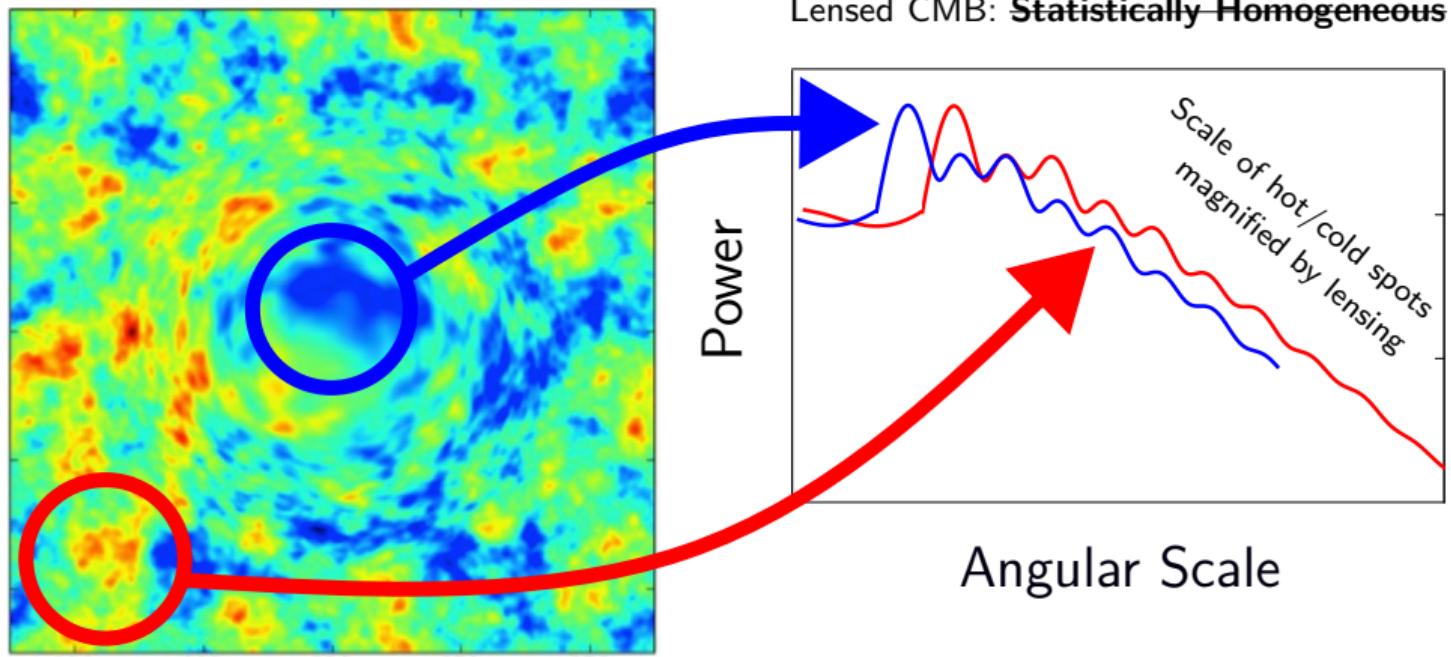


Lensed CMB: **Statistically Homogeneous**



Angular Scale

We can estimate lensing potential since
lensing breaks symmetry of CMB



For a statistically **homogeneous** field like the unlensed CMB different Fourier modes are statistically independent:

$$\langle T_{\ell}^{\text{unlensed}} T_{L-\ell}^{\text{unlensed}} \rangle = 0$$

For a statistically **homogeneous** field like the unlensed CMB different Fourier modes are statistically independent:

$$\langle T_{\ell}^{\text{unlensed}} T_{L-\ell}^{\text{unlensed}} \rangle = 0$$

Lensing of the CMB breaks this symmetry by inducing correlations in our lensed CMB:

$$\langle T_{\ell} T_{L-\ell} \rangle \sim \kappa_L$$

$$(\kappa \equiv -\nabla^2(\text{Lensing Potential})/2)$$

For a statistically **homogeneous** field like the unlensed CMB different Fourier modes are statistically independent:

$$\langle T_{\ell}^{\text{unlensed}} T_{L-\ell}^{\text{unlensed}} \rangle = 0$$

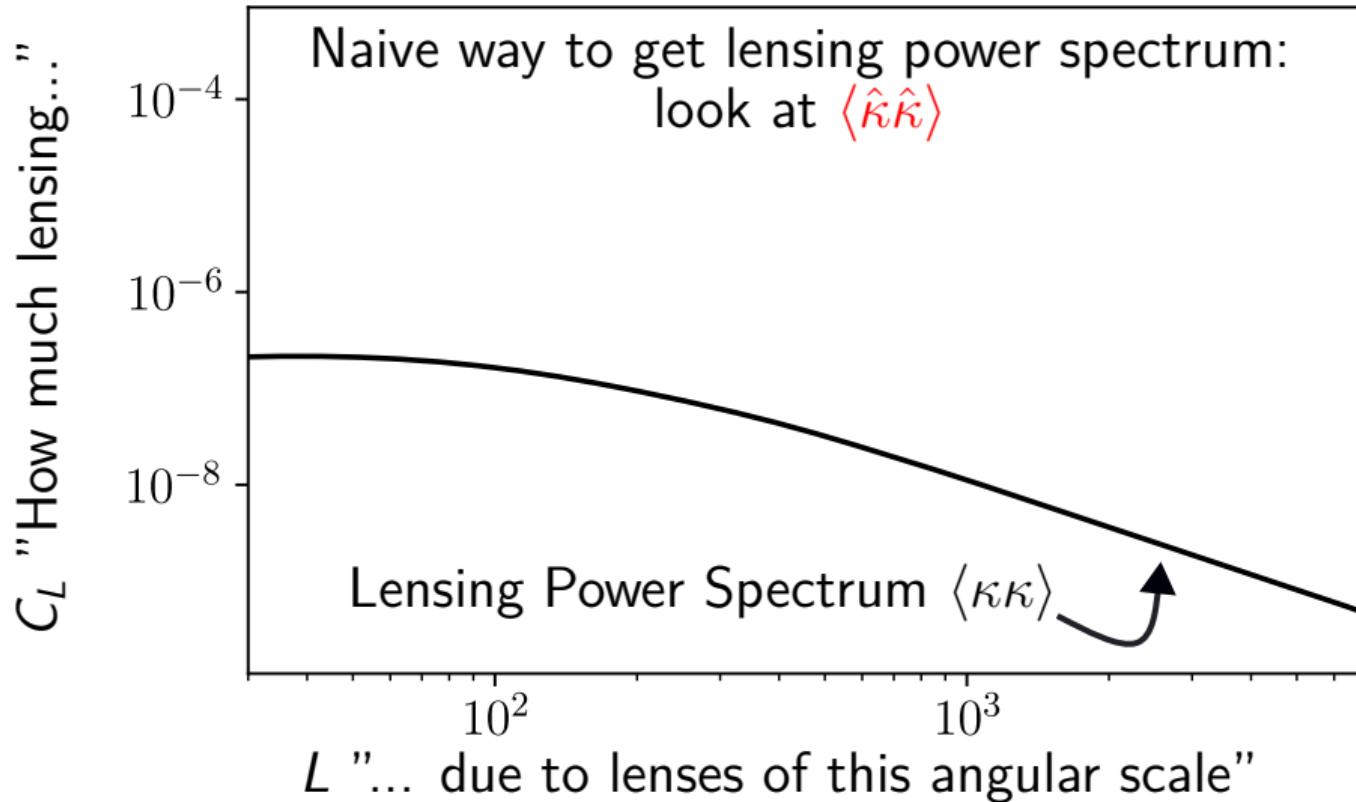
Lensing of the CMB breaks this symmetry by inducing correlations in our lensed CMB:

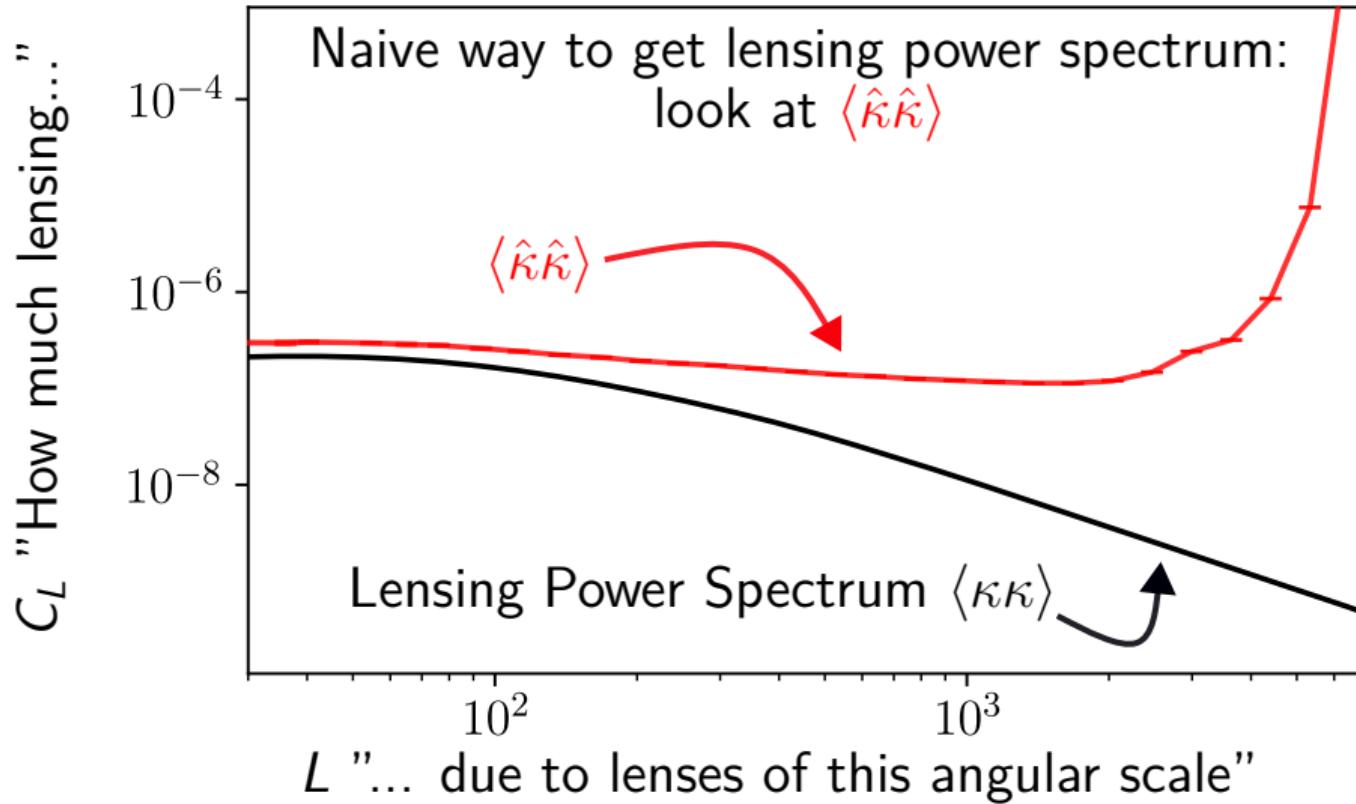
$$\langle T_{\ell} T_{L-\ell} \rangle \sim \kappa_L$$

$$(\kappa \equiv -\nabla^2(\text{Lensing Potential})/2)$$

So correlations that we do see in our map give us information about the lensing allowing us to build an **quadratic estimator** (QE) of κ out of these correlations.

$$\hat{\kappa}_L \sim \int_{\ell} T_{\ell} T_{L-\ell}$$





What happened?

What happened?

$$\hat{\kappa}_{\mathbf{L}} \sim \int_{\ell} T_{\ell} T_{\mathbf{L}-\ell} \Rightarrow \langle \hat{\kappa}_{\mathbf{L}} \hat{\kappa}_{\mathbf{L}}^* \rangle \sim \int_{\ell, \ell'} \langle T_{\ell} T_{\mathbf{L}-\ell} T_{-\ell'} T_{-\mathbf{L}+\ell'} \rangle$$

What happened?

$$\hat{\kappa}_L \sim \int_\ell T_\ell T_{L-\ell} \Rightarrow \langle \hat{\kappa}_L \hat{\kappa}_L^* \rangle \sim \int_{\ell, \ell'} \langle T_\ell T_{L-\ell} T_{-\ell'} T_{-L+\ell'} \rangle$$

The lensing contribution to $\langle TTTT \rangle$ is small because **lensing is a perturbation of the Gaussian random field T^{unlensed}**

What happened?

$$\hat{\kappa}_L \sim \int_\ell T_\ell T_{L-\ell} \Rightarrow \langle \hat{\kappa}_L \hat{\kappa}_L^* \rangle \sim \int_{\ell, \ell'} \langle T_\ell T_{L-\ell} T_{-\ell'} T_{-L+\ell'} \rangle$$

The lensing contribution to $\langle TTTT \rangle$ is small because **lensing is a perturbation of the Gaussian random field T^{unlensed}**

$$\langle TTTT \rangle \sim \underbrace{\text{Large Gaussian Contribution}}_{\text{Present even if no lensing}} + \text{Lensing Term} + \dots$$

$$\Rightarrow \langle \hat{\kappa} \hat{\kappa} \rangle \sim \text{Gaussian Bias} + \text{Lensing Power Spectrum} + \dots$$

Furthermore, there's extra stuff in our temperature maps

1. Foregrounds (CIB, SZ, . . .)
2. Detector noise (**for this talk, will focus on this**)

That are added to our maps:

$$T_\ell = T_\ell^{\text{Lensed}} + N_\ell^{\text{Detector}}$$

¹Generically N^{Detector} is a inhomogeneous non-Gaussian field so has contributions beyond the “Gaussian noise”, but for upcoming wide-field CMB maps, this detector noise can be expanded around a homogeneous Gaussian field.

Furthermore, there's extra stuff in our temperature maps

1. Foregrounds (CIB, SZ, ...)
2. Detector noise (**for this talk, will focus on this**)

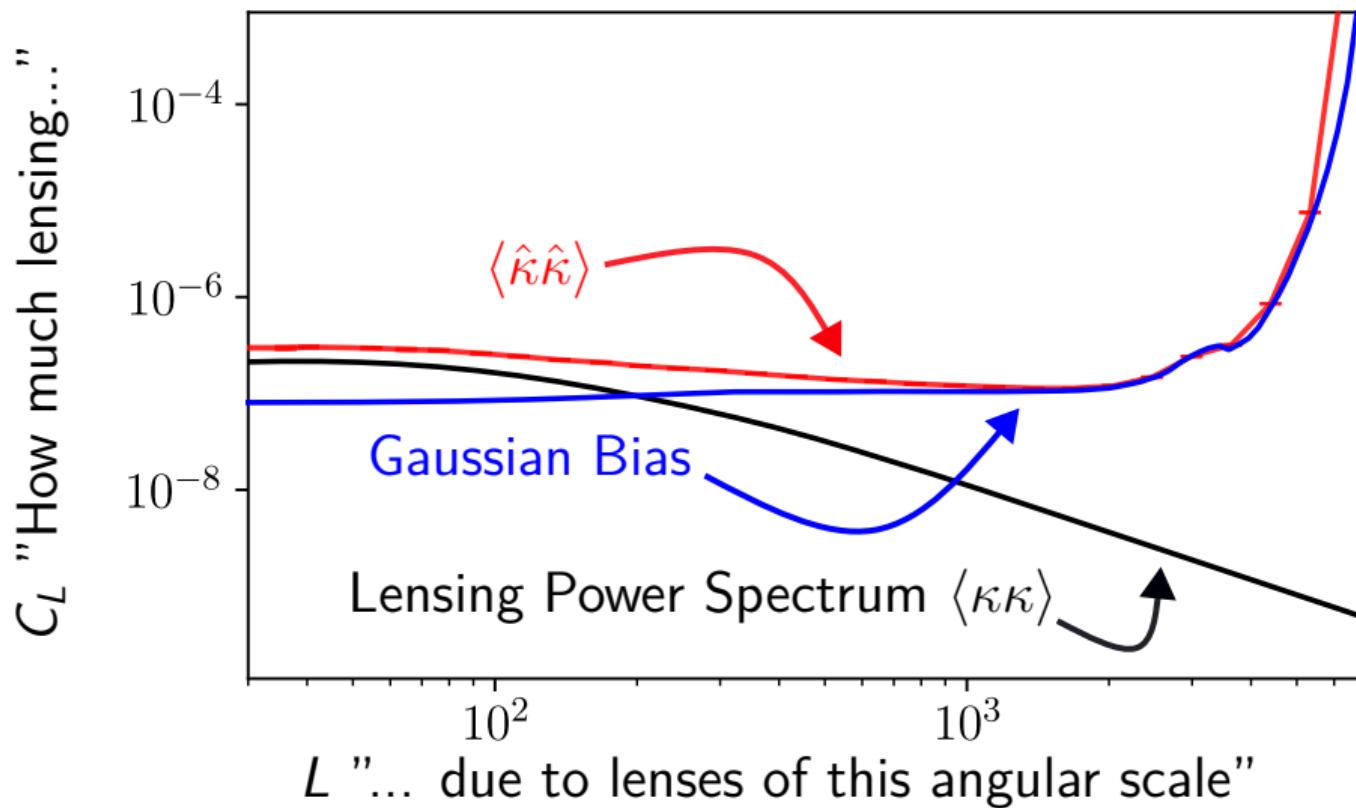
That are added to our maps:

$$T_\ell = T_\ell^{\text{Lensed}} + N_\ell^{\text{Detector}}$$

This changes $\langle TTTT \rangle$ by also contributing to the Gaussian bias¹:

$$\Rightarrow \langle \hat{\kappa} \hat{\kappa} \rangle \sim \text{Noise Bias} + \text{Lensing Power Spectrum} + \dots$$

¹Generically N^{Detector} is a inhomogeneous non-Gaussian field so has contributions beyond the "Gaussian noise", but for upcoming wide-field CMB maps, this detector noise can be expanded around a homogeneous Gaussian field.



On small scales, even a tiny misestimate of Gaussian bias leads to a huge bias in estimated $\langle \kappa \kappa \rangle$

A key challenge in measuring the CMB lensing spectrum is estimating and subsequently **removing this Gaussian bias**².

²Stated more generally, we are trying to extract the non-Gaussian component of the 4-point function. This is a problem that appears generally in cosmology and the method we propose in principle is applicable to other areas where optimal estimation of the connected trispectrum/4-point function is of interest such as large scale structure.

A quick reminder: For a 1-D Gaussian random variable X with mean 0 and variance σ^2

$$\langle X \rangle = 0 \quad \langle X^2 \rangle = \sigma^2$$

1D analogue of
power spectrum

$$\langle X^4 \rangle = 3\sigma^4$$

A toy model for estimating CMB lensing spectra is estimating $\mathcal{K} \ll \sigma^4$ for a **nearly Gaussian** variable X

$$\langle X \rangle = 0 \quad \langle X^2 \rangle = \sigma^2$$

1D analogue of
power spectrum
and
CMB lensing
spectrum

$$\langle X^4 \rangle = 3\sigma^4 + \mathcal{K}$$

“ $\langle TTTT \rangle$ ” “ Noise Bias”

The **most naive** way to remove to
remove this Gaussian bias from $\langle X^4 \rangle$

The **most naive** way to remove to
remove this Gaussian bias from $\langle X^4 \rangle$

1. Assume some theoretical model for the variance σ^2

The **most naive** way to remove to remove this Gaussian bias from $\langle X^4 \rangle$

1. Assume some theoretical model for the variance σ^2
2. Subtract it from the estimated 4-point function from data $\widehat{\langle X^4 \rangle}$.

$$\hat{\mathcal{K}}_{\text{naive}} = \widehat{\langle X^4 \rangle} - 3\sigma^4$$

The **most naive** way to remove to remove this Gaussian bias from $\langle X^4 \rangle$

1. Assume some theoretical model for the variance σ^2
2. Subtract it from the estimated 4-point function from data $\widehat{\langle X^4 \rangle}$.

$$\hat{\mathcal{K}}_{\text{naive}} = \widehat{\langle X^4 \rangle} - 3\sigma^4$$

Features: (1) noisy relative to minimum variance estimator and (2) maximally sensitivity to theoretical mismodelling

The **optimal** way to remove to remove
this Gaussian bias from $\langle X^4 \rangle$

The **optimal** way to remove this Gaussian bias from $\langle X^4 \rangle$

1. Assume some theoretical model for **the variance** σ^2

The **optimal** way to remove this Gaussian bias from $\langle X^4 \rangle$

1. Assume some theoretical model for **the variance** σ^2
2. Derive the optimal way to combine theoretical model with the variance **estimated from data** to estimate the Gaussian bias

The **optimal** way to remove this Gaussian bias from $\langle X^4 \rangle$

1. Assume some theoretical model for **the variance** σ^2
2. Derive the optimal way to combine theoretical model with the variance **estimated from data** to estimate the Gaussian bias
3. Subtract it from the estimated 4-point function from data $\widehat{\langle X^4 \rangle}$.

$$\hat{\mathcal{K}}_{\text{opt}} = \widehat{\langle X^4 \rangle} - 3[\hat{\sigma}_{\text{opt}}(\sigma^2, \text{data})]^4$$

The **optimal** way to remove this Gaussian bias from $\langle X^4 \rangle$

1. Assume some theoretical model for **the variance** σ^2
2. Derive the optimal way to combine theoretical model with the variance **estimated from data** to estimate the Gaussian bias
3. Subtract it from the estimated 4-point function from data $\widehat{\langle X^4 \rangle}$.

$$\hat{\mathcal{K}}_{\text{opt}} = \widehat{\langle X^4 \rangle} - 3[\hat{\sigma}_{\text{opt}}(\sigma^2, \text{data})]^4$$

Features: (1) is the minimum variance estimator but (2) still has some sensitivity to theoretical mismodelling

The optimal way to remove the Gaussian bias from the data is to move this Gaussian bias from the data to the theoretical model $\langle X^4 \rangle$.

The standard way CMB Lensing Spectrum estimation is done "RDN(0)"

1. Assume some theoretical model for the variance σ^2
2. Derive the optimal way to compare the theoretical model with the variance estimated from the data to estimate the Gaussian bias
3. Use the bias from the estimated 4-point function ratio with data $\langle X^4 \rangle$.

$$\hat{E}_{\text{opt}} = \frac{\partial E}{\partial \sigma^2} \Big|_{\sigma^2=\sigma^2_{\text{opt}}} = \frac{\partial}{\partial \sigma^2} [\sigma^2_{\text{opt}}(\sigma^2, \text{data})]^4$$

Features: (1) the minimum variance estimator, but (2) still has some sensitivity to theoretical model choice.

Limitations of the Standard RDN⁽⁰⁾

- ▶ Relies on **computationaly expensive simulations**
- ▶ Estimated noise bias still **sensitive to small errors simulation**
 - ▶ **Small error** in noise bias is still **large error** in CMB lensing power spectrum
- ▶ In ACT DR6 (Qu+23), they could not simulate the complex instrument noise well enough to have unbiased measurements of the lensing spectrum at $L \sim 800$

Origin of the Gaussian bias

Recall that

$$\langle \hat{\kappa}_L \hat{\kappa}_L^* \rangle \sim \iint_{\ell, \ell'} \langle T_\ell T_{L-\ell} T_{-\ell'} T_{-L+\ell'} \rangle$$

Origin of the Gaussian bias

Recall that

$$\langle \hat{\kappa}_L \hat{\kappa}_L^* \rangle \sim \iint_{\ell, \ell'} \langle T_\ell T_{L-\ell} T_{-\ell'} T_{-L+\ell'} \rangle$$

- ▶ Gaussian bias from the 0th order contribution to $\langle TTTT \rangle$

Origin of the Gaussian bias

Recall that

$$\langle \hat{\kappa}_L \hat{\kappa}_L^* \rangle \sim \iint_{\ell, \ell'} \langle T_\ell T_{L-\ell} T_{-\ell'} T_{-L+\ell'} \rangle$$

- ▶ Gaussian bias from the 0th order contribution to $\langle TTTT \rangle$
- ▶ So in the integral, the Gaussian bias contributions come from terms where

$$\langle T_\ell^{\text{unlensed}} T_{L-\ell}^{\text{unlensed}} T_{-\ell'}^{\text{unlensed}} T_{-L+\ell'}^{\text{unlensed}} \rangle \neq 0 \quad (*)$$

Origin of the Gaussian bias

Recall that

$$\langle \hat{\kappa}_L \hat{\kappa}_L^* \rangle \sim \iint_{\ell, \ell'} \langle T_\ell T_{L-\ell} T_{-\ell'} T_{-L+\ell'} \rangle$$

- ▶ Gaussian bias from the 0th order contribution to $\langle TTTT \rangle$
- ▶ So in the integral, the Gaussian bias contributions come from terms where

$$\langle T_\ell^{\text{unlensed}} T_{L-\ell}^{\text{unlensed}} T_{-\ell'}^{\text{unlensed}} T_{-L+\ell'}^{\text{unlensed}} \rangle \neq 0 \quad (*)$$

- ▶ Statistical homogeneity implies that $\langle T_\ell^{\text{unlensed}} T_{\ell'}^{\text{unlensed}} \rangle \sim \delta^{(D)}(\ell + \ell')$

Origin of the Gaussian bias

Recall that

$$\langle \hat{\kappa}_L \hat{\kappa}_L^* \rangle \sim \iint_{\ell, \ell'} \langle T_\ell T_{L-\ell} T_{-\ell'} T_{-L+\ell'} \rangle$$

- ▶ Gaussian bias from the 0th order contribution to $\langle TTTT \rangle$
- ▶ So in the integral, the Gaussian bias contributions come from terms where

$$\langle T_\ell^{\text{unlensed}} T_{L-\ell}^{\text{unlensed}} T_{-\ell'}^{\text{unlensed}} T_{-L+\ell'}^{\text{unlensed}} \rangle \neq 0 \quad (*)$$

- ▶ Statistical homogeneity implies that $\langle T_\ell^{\text{unlensed}} T_{\ell'}^{\text{unlensed}} \rangle \sim \delta^{(D)}(\ell + \ell')$
- ▶ Thus, by Eq. (*), the Gaussian bias comes from terms in $\int_{\ell, \ell'} \boxed{\ell = \ell'}$ where

Our Method: Ignore $\ell = \ell'$ terms

$$\text{Our estimator of } \langle \kappa \kappa \rangle \sim \iint_{\ell \neq \ell'} \langle T_\ell T_{\ell-\ell} T_{-\ell'} T_{-\ell+\ell'} \rangle$$

$$(\text{for practical purposes}) \sim \underbrace{\iint_{\ell, \ell'} \langle T_\ell T_{\ell-\ell} T_{-\ell'} T_{-\ell+\ell'} \rangle}_{\text{The standard } \langle \hat{\kappa} \hat{\kappa} \rangle} - \underbrace{\int_{\ell} \langle T_\ell T_{\ell-\ell} T_{-\ell} T_{-\ell+\ell} \rangle}_{\ell = \ell' \text{ terms}}$$

Both components can be computed efficiently using FFT!

Our Method: Ignore the $\ell = \ell'$ terms

Essentially we avoid the Gaussian bias by throwing away a small subset of our data

(for practical purposes)

The step involves calculating the difference between the full covariance matrix and the $\ell = \ell'$ terms.

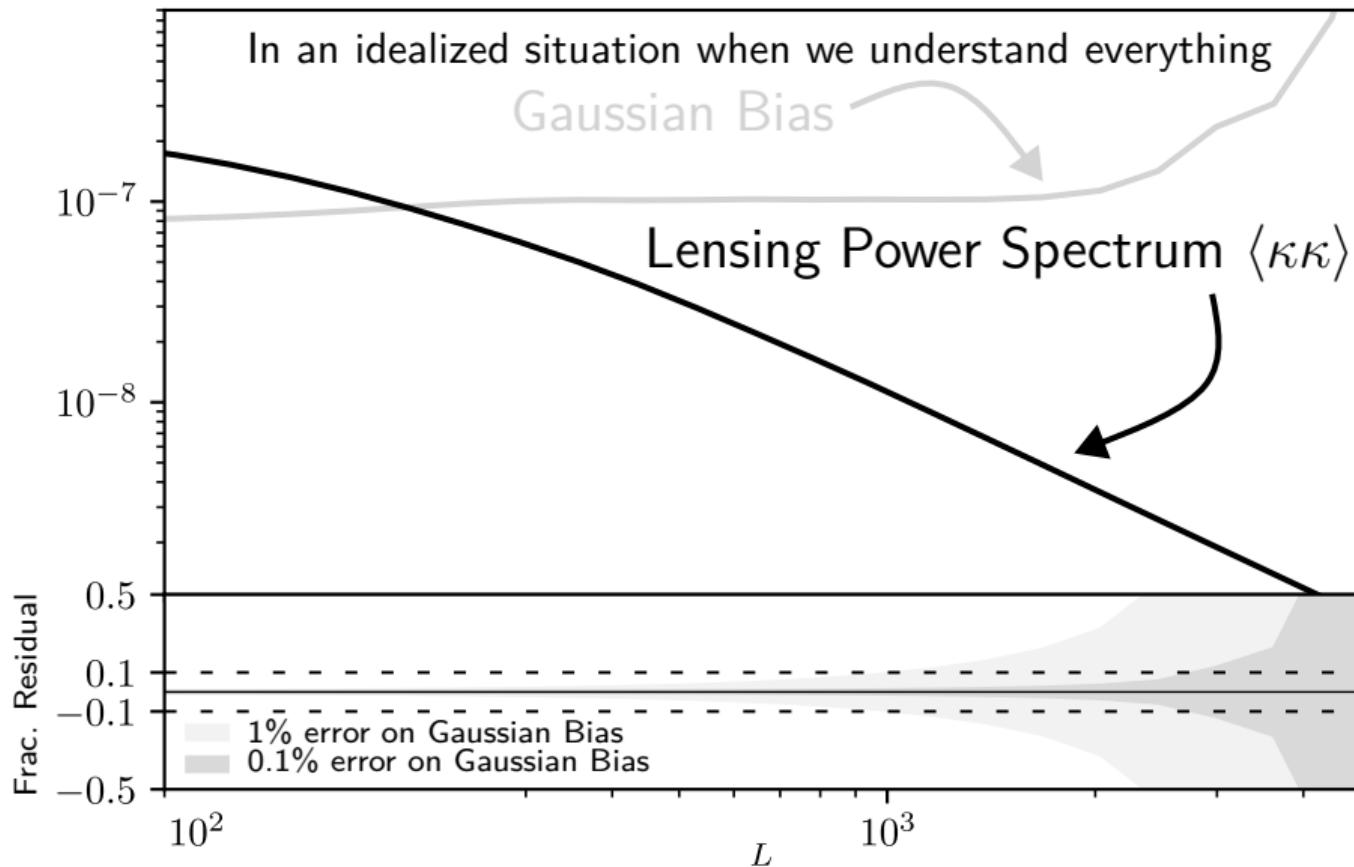
Both components can be computed efficiently using FFT!

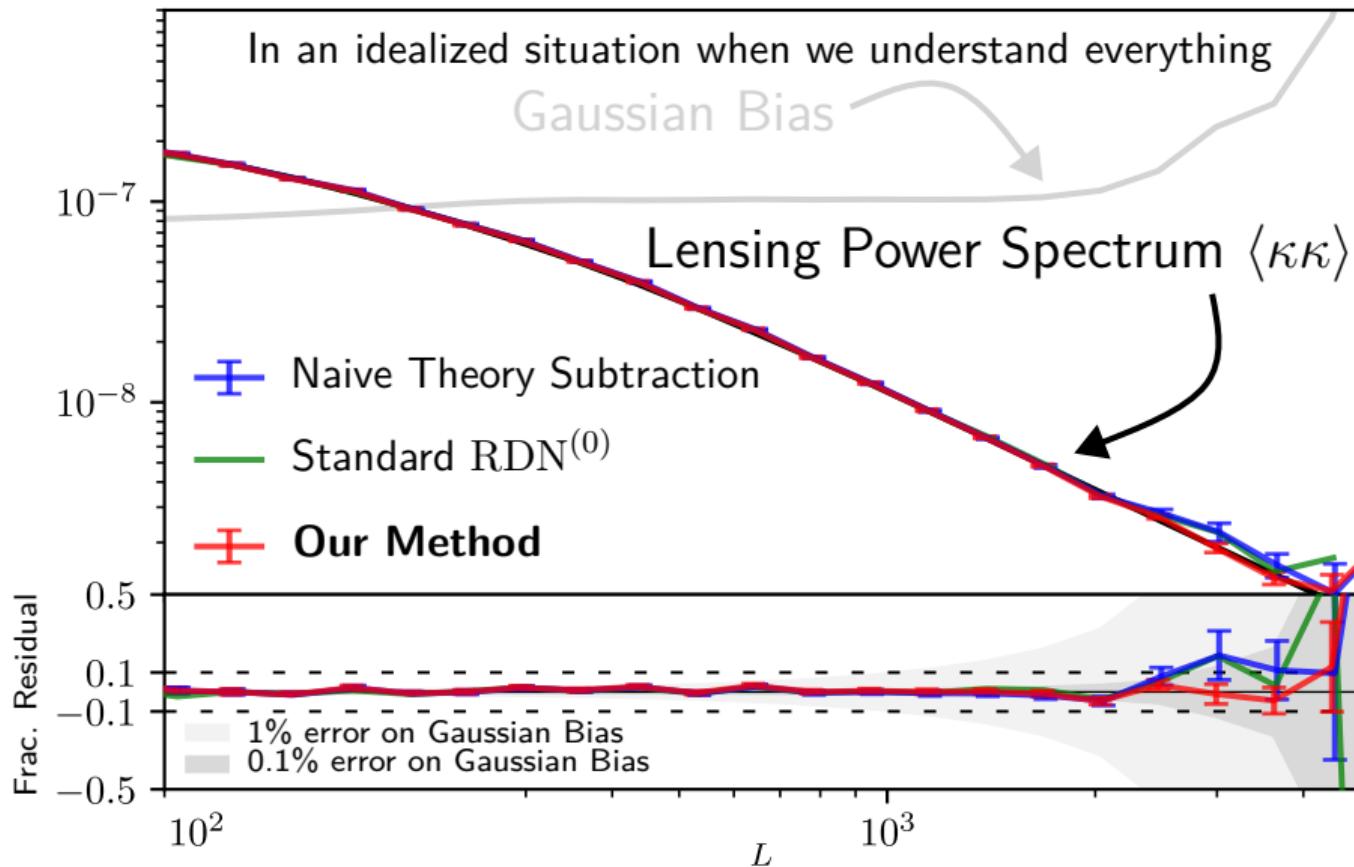
Features of our method

- ▶ Relies **only on data** thus
 1. Very fast to compute relative to standard RDN⁽⁰⁾
 2. Completely insensitive to errors in simulations

Features of our method

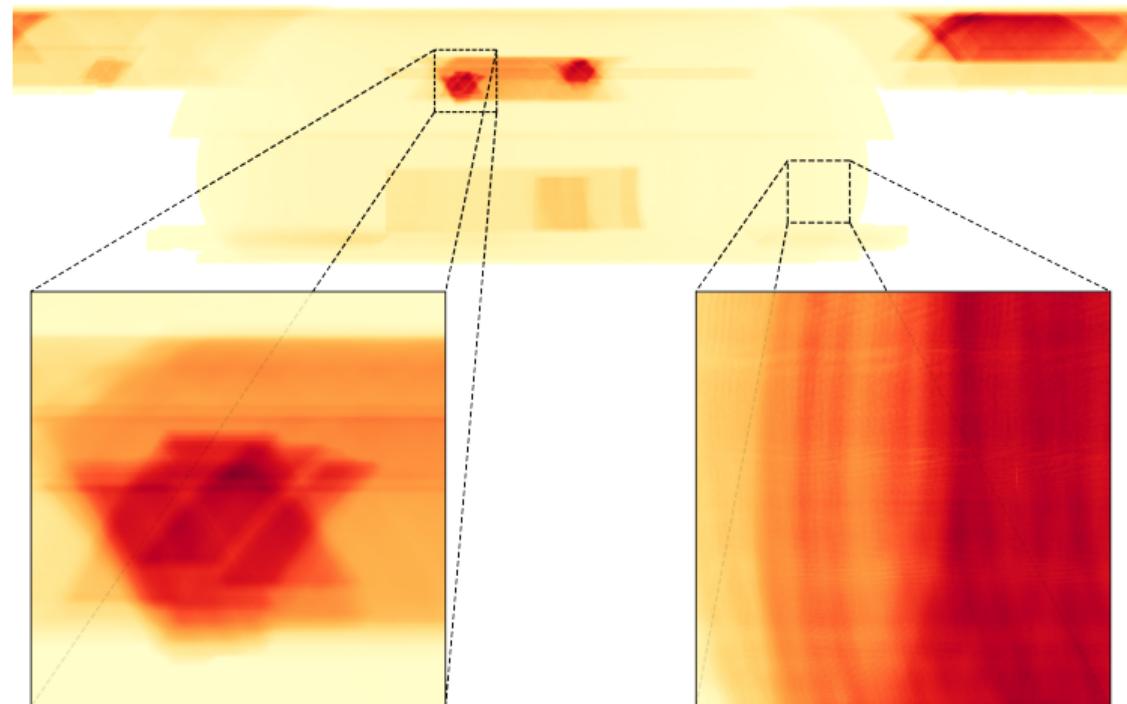
- ▶ Relies **only on data** thus
 1. Very fast to compute relative to standard RDN⁽⁰⁾
 2. Completely insensitive to errors in simulations
- ▶ One can reason using the toy model that the variance of **our method** is **asymptotically equivalent** to the variance of the **the optimal minimum variance estimator.**





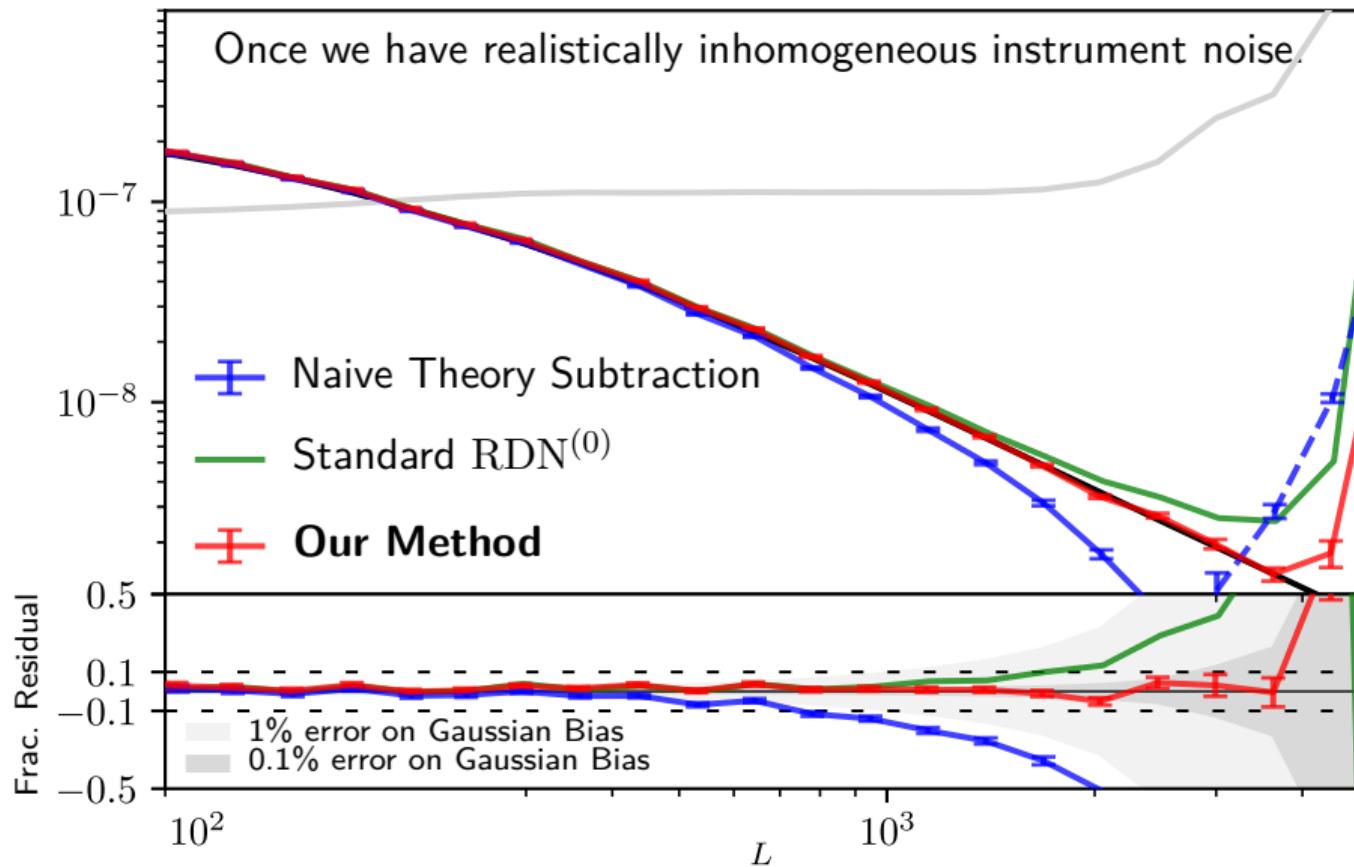
ACT DR5 IVar Map

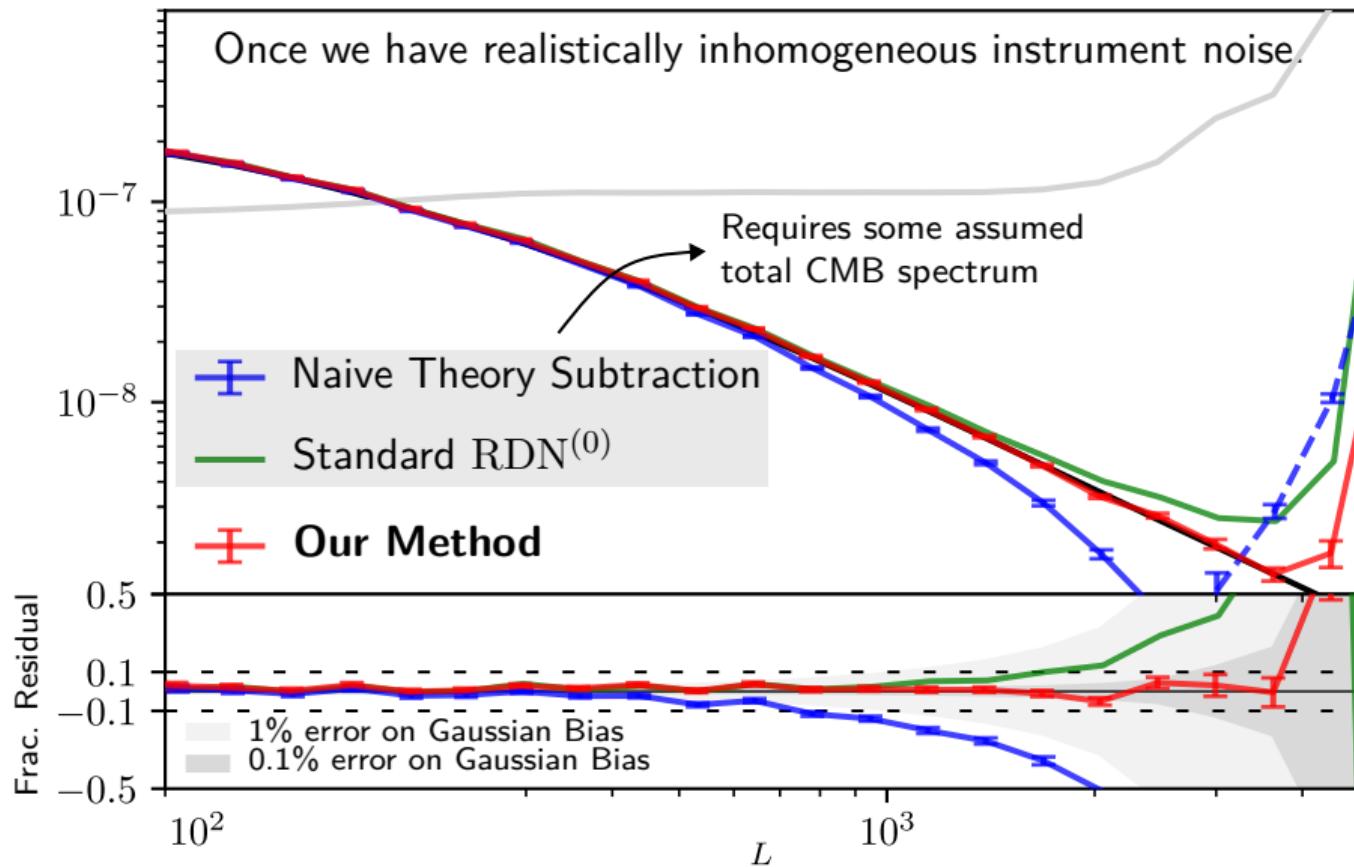
On real data, a key limitation for standard methods is modeling complex instrument noise sufficiently accurately.

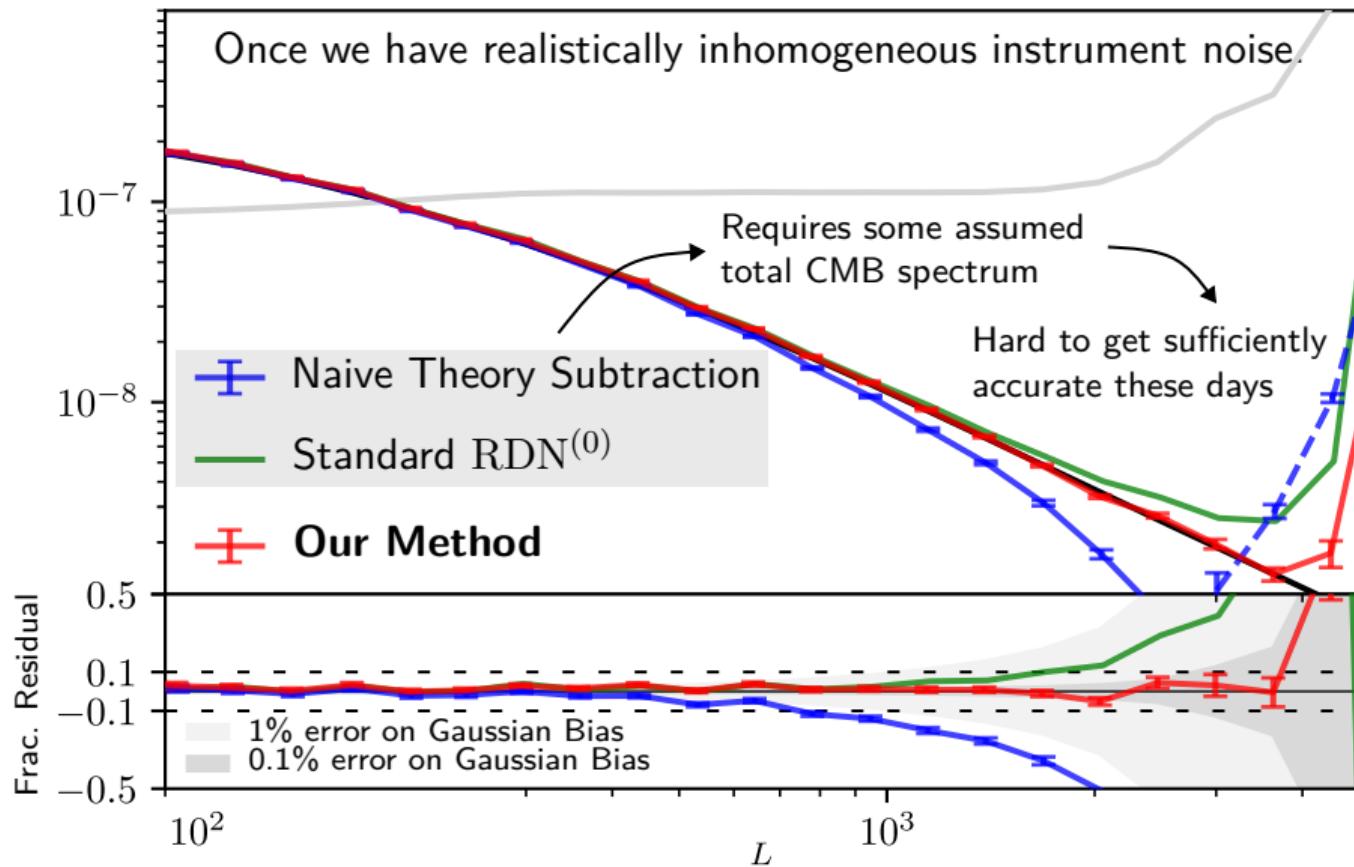


Extreme
Noise Scenario

Typical
Noise Scenario







Conclusion

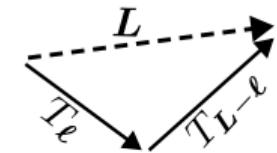
- ▶ We propose a novel estimator of the CMB lensing power spectrum which **relies only on data** making this estimator³
 1. Fast to compute
 2. Insensitive to errors in simulations and thus unbiased to small scales
- ▶ This estimator's noisiness is **asymptotically equivalent to the optimal estimator.**
- ▶ We showed this estimator is robust to the presence realistic complications like inhomogeneous instrument noise
 - ▶ Can also show our estimator is as robust to masking as standard methods, happy to talk about this if people interested!

³funnily enough, the two traits of our CMB lensing power spectrum estimator (fast and unbiased) are also traits that are desirable in any entity that holds power

Extra

Our Method

Recall $\hat{\kappa}_L \sim \int_\ell T_\ell T_{L-\ell} \sim \sum_\ell$



Our Method

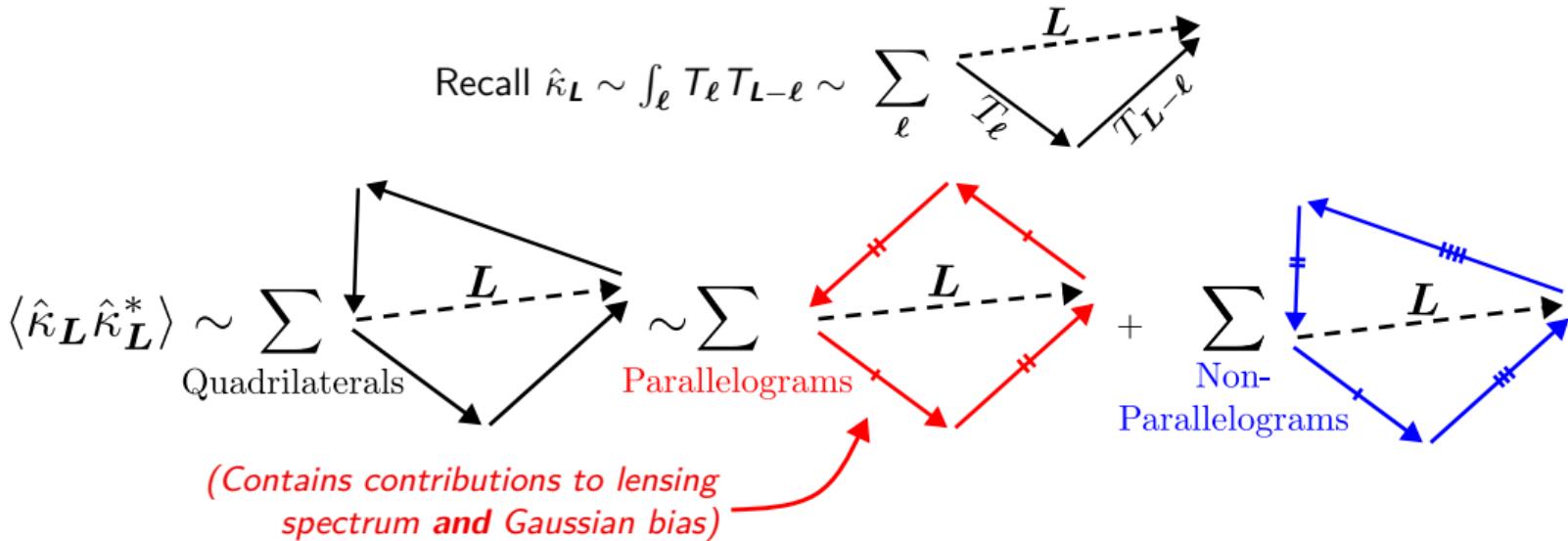
$$\text{Recall } \hat{\kappa}_L \sim \int_{\ell} T_{\ell} T_{L-\ell} \sim \sum_{\ell} \begin{array}{c} L \\ \nearrow T_{\ell} \searrow T_{L-\ell} \end{array}$$

$$\langle \hat{\kappa}_L \hat{\kappa}_L^* \rangle \sim \sum_{\text{Quadrilaterals}} \begin{array}{c} L \\ \nearrow \searrow \end{array}$$

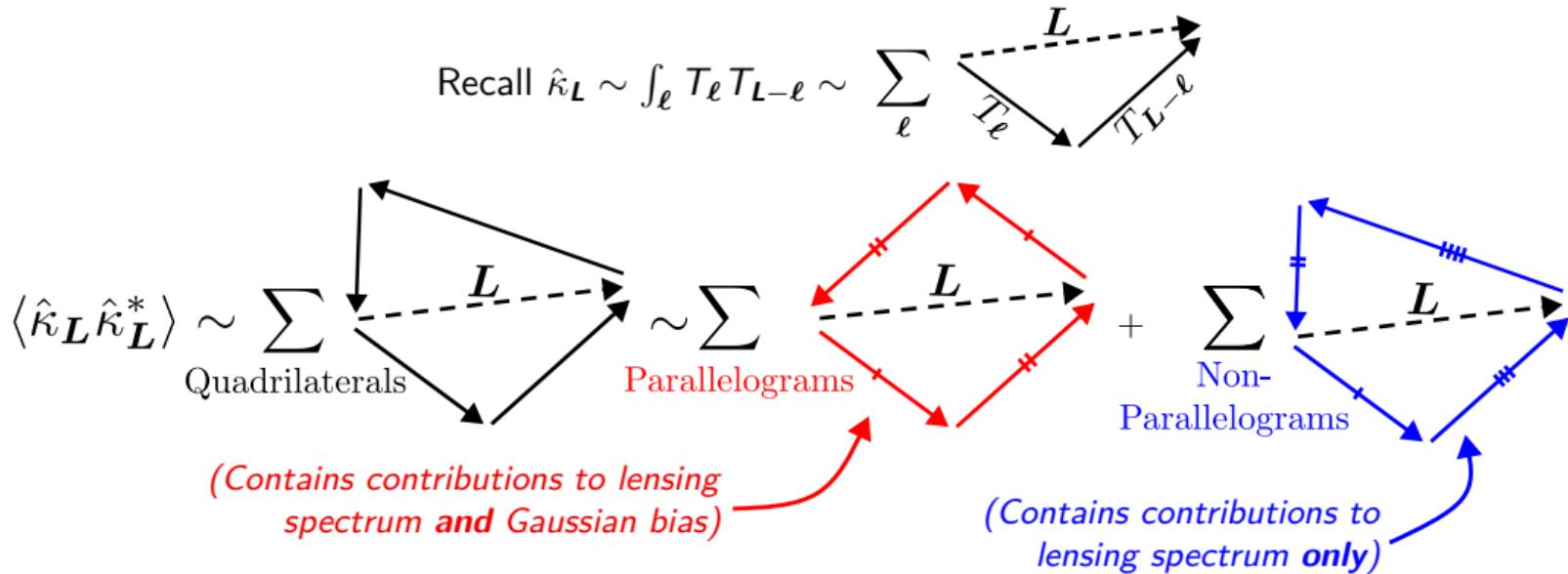
Our Method

$$\text{Recall } \hat{\kappa}_L \sim \int_{\ell} T_{\ell} T_{L-\ell} \sim \sum_{\ell} \begin{array}{c} L \\ \diagup \quad \diagdown \\ T_{\ell} \quad T_{L-\ell} \end{array}$$
$$\langle \hat{\kappa}_L \hat{\kappa}_L^* \rangle \sim \sum_{\text{Quadrilaterals}} \begin{array}{c} L \\ \diagup \quad \diagdown \\ \text{---} \quad \text{---} \end{array} \sim \sum_{\text{Parallelograms}} \begin{array}{c} L \\ \diagup \quad \diagdown \\ \textcolor{red}{\text{---}} \quad \textcolor{red}{\text{---}} \end{array} + \sum_{\text{Non-Parallelograms}} \begin{array}{c} L \\ \diagup \quad \diagdown \\ \textcolor{blue}{\text{---}} \quad \textcolor{blue}{\text{---}} \end{array}$$

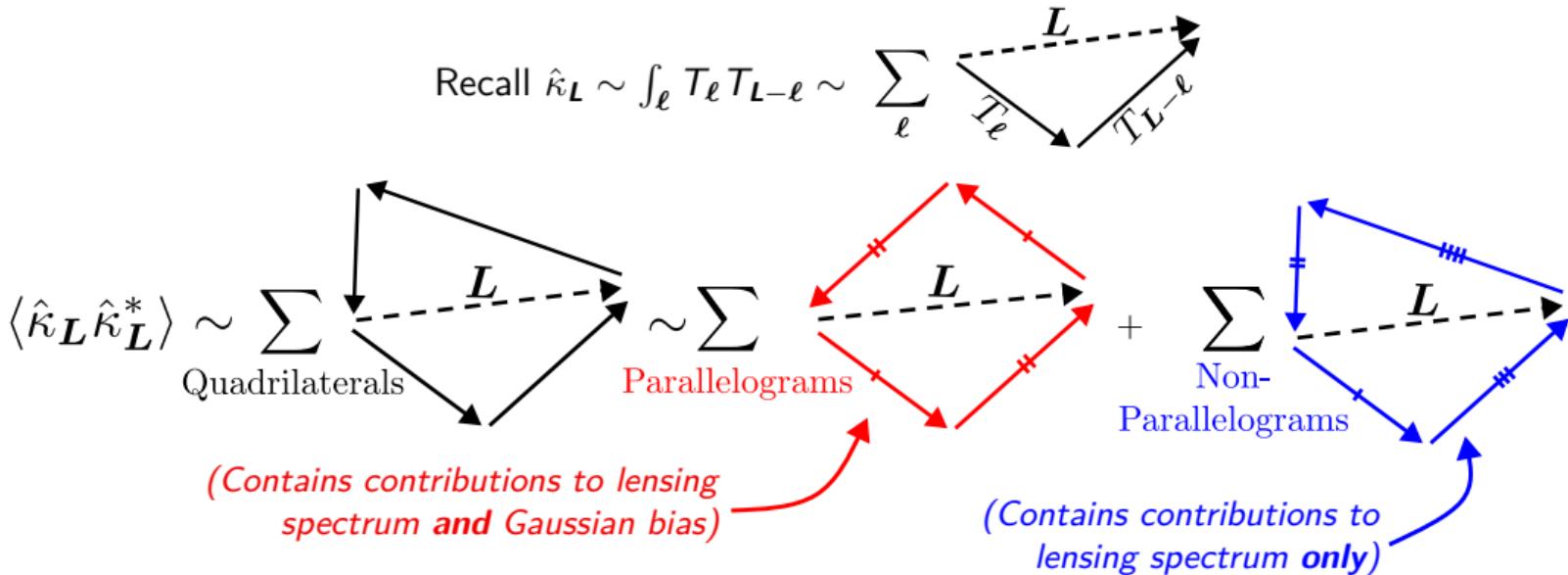
Our Method



Our Method

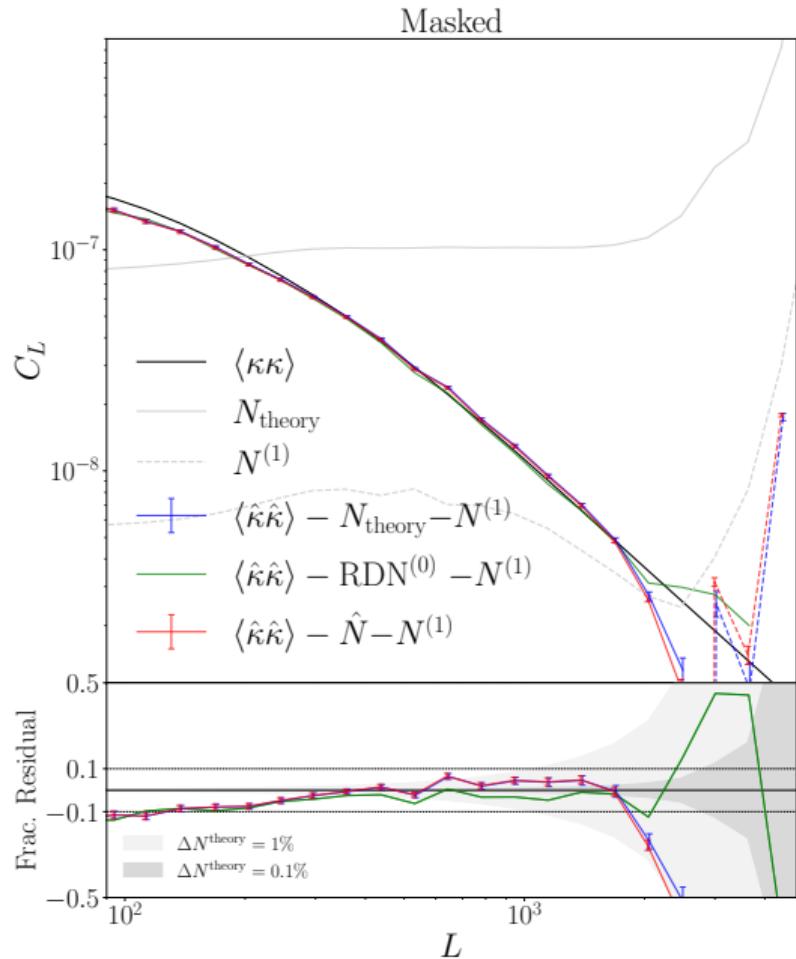


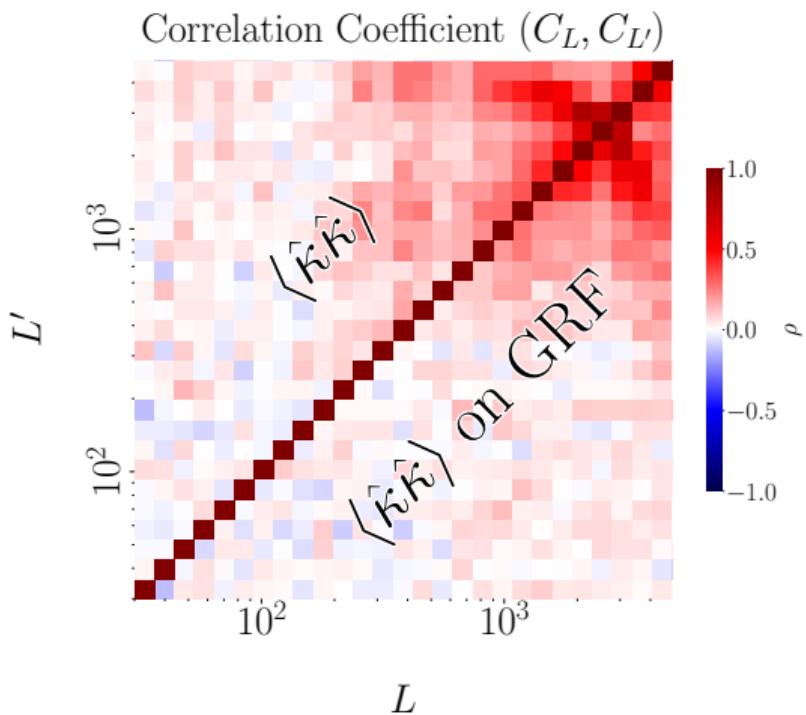
Our Method



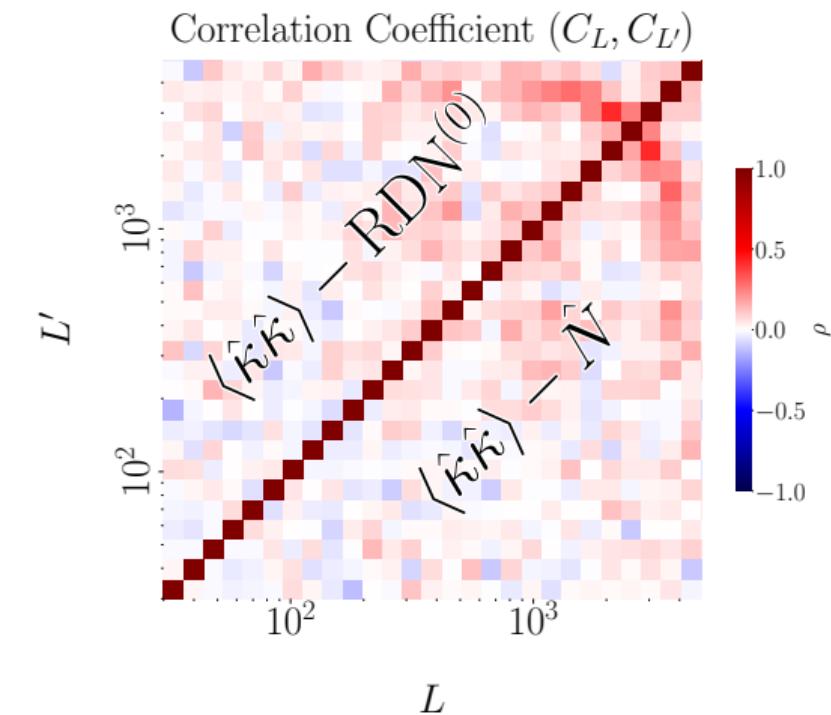
Idea: ignore all the quadrilaterals that **do** contribute to the Gaussian bias, the **parallelograms**

Masking introduces additional mode couplings and a mean-field that have to be handled. Here we show that our method is as good as the standard method in presence of masking! So usual methods to handle additional mode couplings should still work.

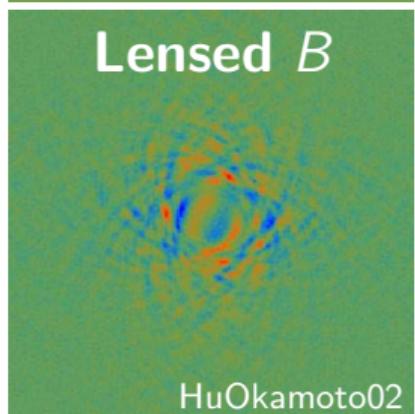
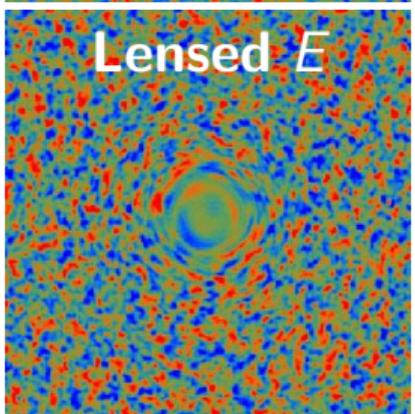
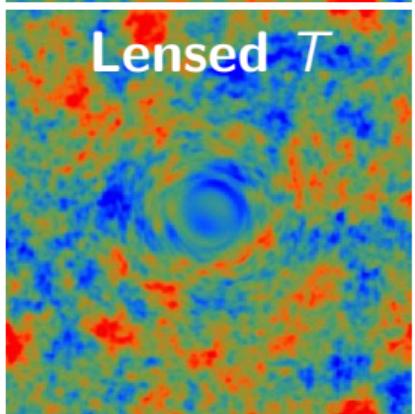
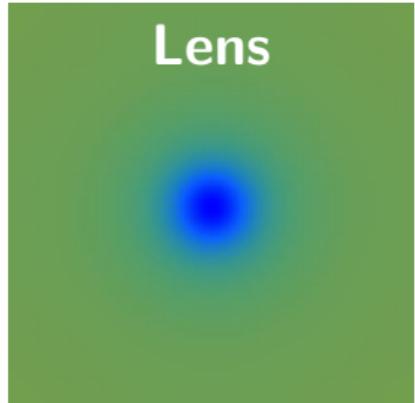
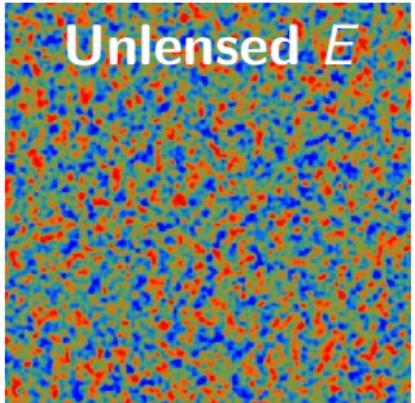
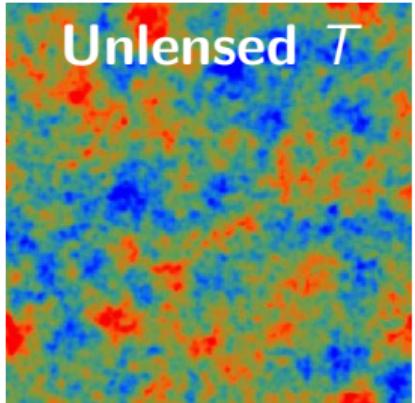




Non-trivial correlation structure due to Gaussian bias

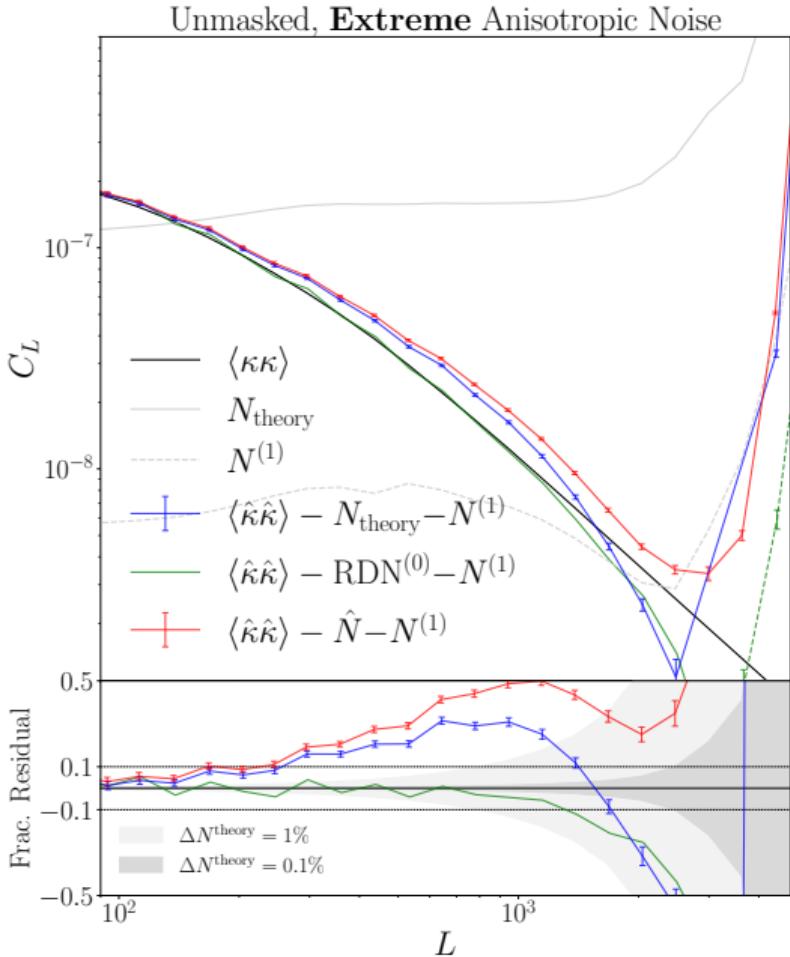


Methods to remove Gaussian bias should also remove non-trivial correlation structure



HuOkamoto02

In our method we implicitly assume that noise inhomogeneities can be expanded perturbatively around a homogeneous limit. However, once this is not true, for example in the extreme inhomogeneous noise scenario, our method breaks down while the standard RDN⁽⁰⁾, which does not make this assumption, still works. However **we do expect the assumption that noise inhomogeneities are small to hold for current and upcoming wide-field CMB maps.**



To intuitively understand the properties of (1) the current method and (2) the method we will propose to remove the Gaussian bias, lets consider a **toy model** for optimal trispectrum estimation⁴.

⁴Originally presented in Smith+18

In Toy Model

Consider a **generic** weakly non-Gaussian random **variable** X with

- ▶ Zero mean
- ▶ Some assumed **variance** σ^2
- ▶ Small **kurtosis** $\mathcal{K} \ll \sigma^4$ we wish to estimate

$$\langle X^2 \rangle = \sigma^2$$

$$\langle X^4 \rangle = 3\sigma^4 + \mathcal{K}$$

In Toy Model

Consider a **generic** weakly non-Gaussian random **variable** X with

- ▶ Zero mean
- ▶ Some assumed **variance** σ^2
- ▶ Small **kurtosis** $\mathcal{K} \ll \sigma^4$ we wish to estimate

$$\langle X^2 \rangle = \sigma^2$$

$$\langle X^4 \rangle = 3\sigma^4 + \mathcal{K}$$

In CMB Lensing

Consider the **lensed CMB temperature map**, a weakly non-Gaussian random **field** with

- ▶ $\langle T \rangle = 0$
- ▶ Some assumed **power spectrum**
- ▶ Small **connected trispectrum** (induced by lensing) we wish to estimate

Let $\{x_i\}$ be N independent realizations of the random variable X . Generically, an estimator of the Kurtosis \mathcal{K} will look like

$$\hat{\mathcal{K}} = \frac{1}{N} \sum_{i=1}^N x_i^4 - \hat{\mathcal{G}} = (\text{Sample 4-pt function}) - (\text{Estimate of Gaussian Bias})$$

Let $\{x_i\}$ be N independent realizations of the random variable X . Generically, an estimator of the Kurtosis \mathcal{K} will look like

$$\hat{\mathcal{K}} = \frac{1}{N} \sum_{i=1}^N x_i^4 - \hat{\mathcal{G}} = (\text{Sample 4-pt function}) - (\text{Estimate of Gaussian Bias})$$

| Estimate of Gaussian Bias | $N \times \text{Var}(\hat{\mathcal{K}})$ | $\sigma^2 = \sigma_{\text{true}}^2 - \epsilon$ | In CMB Lensing |
|--|--|--|----------------|
| $\hat{\mathcal{G}}_{\text{naive}} = 3\sigma^4$ | | | |

Let $\{x_i\}$ be N independent realizations of the random variable X . Generically, an estimator of the Kurtosis \mathcal{K} will look like

$$\hat{\mathcal{K}} = \frac{1}{N} \sum_{i=1}^N x_i^4 - \hat{\mathcal{G}} = (\text{Sample 4-pt function}) - (\text{Estimate of Gaussian Bias})$$

| Estimate of Gaussian Bias | $N \times \text{Var}(\hat{\mathcal{K}})$ | $\sigma^2 = \sigma_{\text{true}}^2 - \epsilon$ | In CMB Lensing |
|--|--|--|----------------|
| $\hat{\mathcal{G}}_{\text{naive}} = 3\sigma^4$ | $96\sigma^8$ | | |

Let $\{x_i\}$ be N independent realizations of the random variable X . Generically, an estimator of the Kurtosis \mathcal{K} will look like

$$\hat{\mathcal{K}} = \frac{1}{N} \sum_{i=1}^N x_i^4 - \hat{\mathcal{G}} = (\text{Sample 4-pt function}) - (\text{Estimate of Gaussian Bias})$$

| Estimate of Gaussian Bias | $N \times \text{Var}(\hat{\mathcal{K}})$ | $\sigma^2 = \sigma_{\text{true}}^2 - \epsilon$ | In CMB Lensing |
|--|--|--|----------------|
| $\hat{\mathcal{G}}_{\text{naive}} = 3\sigma^4$ | $96\sigma^8$ | $\Delta\mathcal{K} \sim \epsilon$ | |

Let $\{x_i\}$ be N independent realizations of the random variable X . Generically, an estimator of the Kurtosis \mathcal{K} will look like

$$\hat{\mathcal{K}} = \frac{1}{N} \sum_{i=1}^N x_i^4 - \hat{\mathcal{G}} = (\text{Sample 4-pt function}) - (\text{Estimate of Gaussian Bias})$$

| Estimate of Gaussian Bias | $N \times \text{Var}(\hat{\mathcal{K}})$ | $\sigma^2 = \sigma_{\text{true}}^2 - \epsilon$ | In CMB Lensing |
|--|--|--|---------------------------|
| $\hat{\mathcal{G}}_{\text{naive}} = 3\sigma^4$ | $96\sigma^8$ | $\Delta\mathcal{K} \sim \epsilon$ | $N_{\text{theory}}^{(0)}$ |

Let $\{x_i\}$ be N independent realizations of the random variable X . Generically, an estimator of the Kurtosis \mathcal{K} will look like

$$\hat{\mathcal{K}} = \frac{1}{N} \sum_{i=1}^N x_i^4 - \hat{\mathcal{G}} = (\text{Sample 4-pt function}) - (\text{Estimate of Gaussian Bias})$$

| Estimate of Gaussian Bias | $N \times \text{Var}(\hat{\mathcal{K}})$ | $\sigma^2 = \sigma_{\text{true}}^2 - \epsilon$ | In CMB Lensing |
|---|--|--|---------------------------|
| $\hat{\mathcal{G}}_{\text{naive}} = 3\sigma^4$ | $96\sigma^8$ | $\Delta\mathcal{K} \sim \epsilon$ | $N_{\text{theory}}^{(0)}$ |
| $\hat{\mathcal{G}}_{\text{opt}} = \frac{6\sigma^2}{N} \sum_i x_i^2 - 3\sigma^4$ | $24\sigma^8$ | $\Delta\mathcal{K} \sim \epsilon^2$ | Current Standard |

Let $\{x_i\}$ be N independent realizations of the random variable X . Generically, an estimator of the Kurtosis \mathcal{K} will look like

$$\hat{\mathcal{K}} = \frac{1}{N} \sum_{i=1}^N x_i^4 - \hat{\mathcal{G}} = (\text{Sample 4-pt function}) - (\text{Estimate of Gaussian Bias})$$

| Estimate of Gaussian Bias | $N \times \text{Var}(\hat{\mathcal{K}})$ | $\sigma^2 = \sigma_{\text{true}}^2 - \epsilon$ | In CMB Lensing |
|---|--|--|---------------------------|
| $\hat{\mathcal{G}}_{\text{naive}} = 3\sigma^4$ | $96\sigma^8$ | $\Delta\mathcal{K} \sim \epsilon$ | $N_{\text{theory}}^{(0)}$ |
| $\hat{\mathcal{G}}_{\text{opt}} = \frac{6\sigma^2}{N} \sum_i x_i^2 - 3\sigma^4$ | $24\sigma^8$ | $\Delta\mathcal{K} \sim \epsilon^2$ | Current Standard |
| $\hat{\mathcal{G}}_{\text{alt}} = \frac{3}{N(N-1)} \sum_{i \neq j} x_i^2 x_j^2$ | | | |

Let $\{x_i\}$ be N independent realizations of the random variable X . Generically, an estimator of the Kurtosis \mathcal{K} will look like

$$\hat{\mathcal{K}} = \frac{1}{N} \sum_{i=1}^N x_i^4 - \hat{\mathcal{G}} = (\text{Sample 4-pt function}) - (\text{Estimate of Gaussian Bias})$$

| Estimate of Gaussian Bias | $N \times \text{Var}(\hat{\mathcal{K}})$ | $\sigma^2 = \sigma_{\text{true}}^2 - \epsilon$ | In CMB Lensing |
|---|--|--|---------------------------|
| $\hat{\mathcal{G}}_{\text{naive}} = 3\sigma^4$ | $96\sigma^8$ | $\Delta\mathcal{K} \sim \epsilon$ | $N_{\text{theory}}^{(0)}$ |
| $\hat{\mathcal{G}}_{\text{opt}} = \frac{6\sigma^2}{N} \sum_i x_i^2 - 3\sigma^4$ | $24\sigma^8$ | $\Delta\mathcal{K} \sim \epsilon^2$ | Current Standard |
| $\hat{\mathcal{G}}_{\text{alt}} = \frac{3}{N(N-1)} \sum_{i \neq j} x_i^2 x_j^2$ | $24\sigma^8 + O(1/N)$ | | |

Let $\{x_i\}$ be N independent realizations of the random variable X . Generically, an estimator of the Kurtosis \mathcal{K} will look like

$$\hat{\mathcal{K}} = \frac{1}{N} \sum_{i=1}^N x_i^4 - \hat{\mathcal{G}} = (\text{Sample 4-pt function}) - (\text{Estimate of Gaussian Bias})$$

| Estimate of Gaussian Bias | $N \times \text{Var}(\hat{\mathcal{K}})$ | $\sigma^2 = \sigma_{\text{true}}^2 - \epsilon$ | In CMB Lensing |
|---|--|--|---------------------------|
| $\hat{\mathcal{G}}_{\text{naive}} = 3\sigma^4$ | $96\sigma^8$ | $\Delta\mathcal{K} \sim \epsilon$ | $N_{\text{theory}}^{(0)}$ |
| $\hat{\mathcal{G}}_{\text{opt}} = \frac{6\sigma^2}{N} \sum_i x_i^2 - 3\sigma^4$ | $24\sigma^8$ | $\Delta\mathcal{K} \sim \epsilon^2$ | Current Standard |
| $\hat{\mathcal{G}}_{\text{alt}} = \frac{3}{N(N-1)} \sum_{i \neq j} x_i^2 x_j^2$ | $24\sigma^8 + O(1/N)$ | $\Delta\mathcal{K} = 0$ | |

Let $\{x_i\}$ be N independent realizations of the random variable X . Generically, an estimator of the Kurtosis \mathcal{K} will look like

$$\hat{\mathcal{K}} = \frac{1}{N} \sum_{i=1}^N x_i^4 - \hat{\mathcal{G}} = (\text{Sample 4-pt function}) - (\text{Estimate of Gaussian Bias})$$

| Estimate of Gaussian Bias | $N \times \text{Var}(\hat{\mathcal{K}})$ | $\sigma^2 = \sigma_{\text{true}}^2 - \epsilon$ | In CMB Lensing |
|---|--|--|---------------------------|
| $\hat{\mathcal{G}}_{\text{naive}} = 3\sigma^4$ | $96\sigma^8$ | $\Delta\mathcal{K} \sim \epsilon$ | $N_{\text{theory}}^{(0)}$ |
| $\hat{\mathcal{G}}_{\text{opt}} = \frac{6\sigma^2}{N} \sum_i x_i^2 - 3\sigma^4$ | $24\sigma^8$ | $\Delta\mathcal{K} \sim \epsilon^2$ | Current Standard |
| $\hat{\mathcal{G}}_{\text{alt}} = \frac{3}{N(N-1)} \sum_{i \neq j} x_i^2 x_j^2$ | $24\sigma^8 + O(1/N)$ | $\Delta\mathcal{K} = 0$ | Our method |

In Toy Model

The alternative estimator for the kurtosis $\hat{\mathcal{K}}_{\text{alt}}$ has equivalent performance to optimal **kurotsis** estimator for large N but is completely insensitive to mismodelling of **the variance** σ^2

In Toy Model

The alternative estimator for the kurtosis $\hat{\mathcal{K}}_{\text{alt}}$ has equivalent performance to optimal **kurotsis** estimator for large N but is completely insensitive to mismodelling of **the variance** σ^2

In CMB Lensing

Our estimator for the CMB lensing power spectrum $\langle \kappa \kappa \rangle$ asymptotically has equivalent performance to the optimal **trispectrum** estimator but is completely insensitive to mismodelling of **the observed CMB temperature power spectrum.**

In Toy Model

$$\hat{\mathcal{K}}_{\text{alt}} = \frac{1}{N} \sum_i x_i^4 - \frac{3}{N(N-1)} \sum_{i \neq j} x_i^2 x_j^2$$

Since x_i is statistically independent from x_j when $i \neq j$, the second term yields only the disconnected Gaussian bias.

In Toy Model

$$\hat{\mathcal{K}}_{\text{alt}} = \frac{1}{N} \sum_i x_i^4 - \frac{3}{N(N-1)} \sum_{i \neq j} x_i^2 x_j^2$$

Since x_i is statistically independent from x_j when $i \neq j$, the second term yields only the disconnected Gaussian bias.

In other words, instead of assuming some σ^2 , you estimate it directly from data.

In Toy Model

$$\hat{\mathcal{K}}_{\text{alt}} = \frac{1}{N} \sum_i x_i^4 - \frac{3}{N(N-1)} \sum_{i \neq j} x_i^2 x_j^2$$

Since x_i is statistically independent from x_j when $i \neq j$, the second term yields only the disconnected Gaussian bias.

In other words, instead of assuming some σ^2 , you estimate it directly from data.

In CMB Lensing

Recall that

$$\langle \hat{\kappa}_L \hat{\kappa}_L^* \rangle \sim \int_{\ell, \ell'} \langle T_\ell T_{L-\ell} T_{-\ell'} T_{-L+\ell'} \rangle$$

Our proposed estimator generalization of the toy model's \mathcal{K}_{alt}

$$\langle \kappa \kappa \rangle \sim \langle \hat{\kappa} \hat{\kappa} \rangle - \int_{\ell} \langle T_\ell T_{L-\ell} T_{-\ell} T_{-L+\ell} \rangle$$

Similar to the toy model case, the second term contains all the combination of (ℓ, ℓ') contributing to $\int_{\ell, \ell'}$ which contain a disconnected Gaussian bias.

What people usually do to estimate the Gaussian bias

In Toy Model

$$\hat{\mathcal{G}}_{\text{opt}} = \frac{6\sigma^2}{N} \sum_i x_i^2 - 3\sigma^4$$

What people usually do to estimate the Gaussian bias

In Toy Model

$$\hat{\mathcal{G}}_{\text{opt}} = \frac{6\sigma^2}{N} \sum_i x_i^2 - 3\sigma^4$$

The optimal way to combine
a **theoretical prediction**
for variance σ^2 and the
sample variance when
estimating the kurtosis

What people usually do to estimate the Gaussian bias

In Toy Model

$$\hat{\mathcal{G}}_{\text{opt}} = \frac{6\sigma^2}{N} \sum_i x_i^2 - 3\sigma^4$$

The optimal way to combine a **theoretical prediction for variance** σ^2 and the **sample variance** when estimating the kurtosis

In CMB Lensing

Let $\hat{\kappa}^{ds}$ be the QE using data for one temperature map and simulations for the other. Similar for $\hat{\kappa}^{ss'}$

$$\begin{aligned} \text{RDN}_L^{(0)} \equiv & \left\langle C_L(\hat{\kappa}^{ds}, \hat{\kappa}^{ds}) + C_L(\hat{\kappa}^{ds}, \hat{\kappa}^{sd}) \right. \\ & + C_L(\hat{\kappa}^{sd}, \hat{\kappa}^{ds}) + C_L(\hat{\kappa}^{sd}, \hat{\kappa}^{sd}) \\ & \left. - (C_L(\hat{\kappa}^{ss'}, \hat{\kappa}^{ss'}) + C_L(\hat{\kappa}^{ss'}, \hat{\kappa}^{s's})) \right\rangle_{s,s'} \end{aligned}$$

The optimal way to combine **simulations of CMB maps with realistic complications** (instruments noise, foregrounds, etc.) and **actual maps** when estimating the connected trispectrum

What people usually do to estimate the Gaussian bias

In Toy Model

$$\hat{\mathcal{G}}_{\text{opt}} = \frac{6\sigma^2}{N} \sum_i x_i^2 - 3\sigma^4$$

The optimal way to combine a **theoretical prediction for variance** σ^2 and the **sample variance** when estimating the kurtosis

In CMB Lensing

Let $\hat{\kappa}^{ds}$ be the QE using data for one temperature map and simulations for the other. Similar for $\hat{\kappa}^{ss'}$

$$\begin{aligned} \text{RDN}_L^{(0)} \equiv & \left\langle C_L(\hat{\kappa}^{ds}, \hat{\kappa}^{ds}) + C_L(\hat{\kappa}^{ds}, \hat{\kappa}^{sd}) \right. \\ & + C_L(\hat{\kappa}^{sd}, \hat{\kappa}^{ds}) + C_L(\hat{\kappa}^{sd}, \hat{\kappa}^{sd}) \\ & \left. - (C_L(\hat{\kappa}^{ss'}, \hat{\kappa}^{ss'}) + C_L(\hat{\kappa}^{ss'}, \hat{\kappa}^{s's})) \right\rangle_{s,s'} \end{aligned}$$

The optimal way to combine **simulations of CMB maps with realistic complications** (instruments noise, foregrounds, etc.) and **actual maps** when estimating the connected trispectrum

For modern data, simulations (1) expensive and (2) require modelling everything to exquisite accuracy