Application of Wavelet Based Statistics for Enhanced 21cm Parameter Constraints

I. Hothi, **E. Allys, B. Semelin, F. Boulanger**

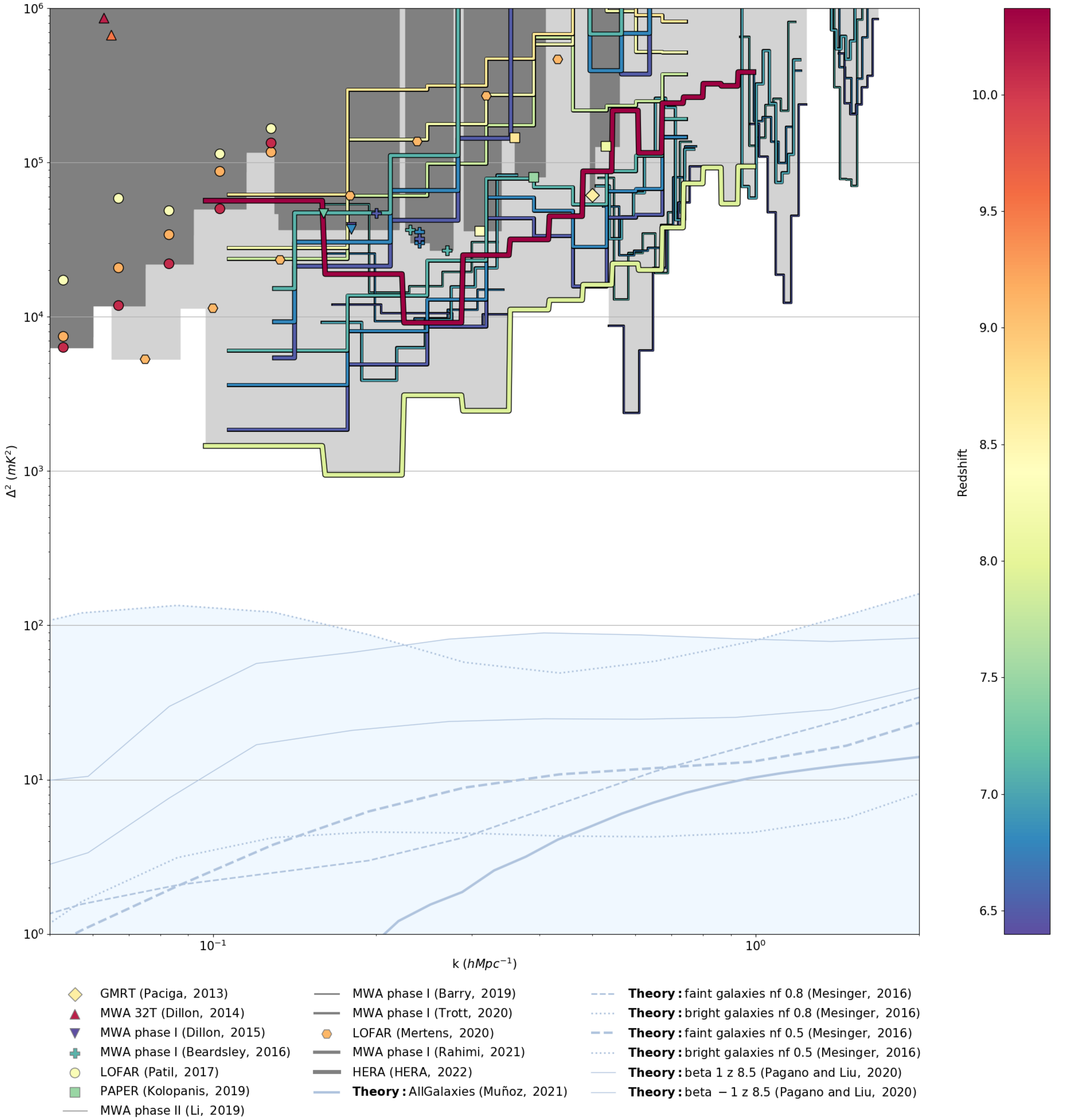
SCAN ME

LPENS
LABORATOIRE DE PHYSIQUE DE L'ÉCOLE NORMALE SUPÉRIEURE

Initiative Physique des Infinis

LERMA l'Observatoire de Paris | PSL

The Spherically-Averaged 3D Power spectrum

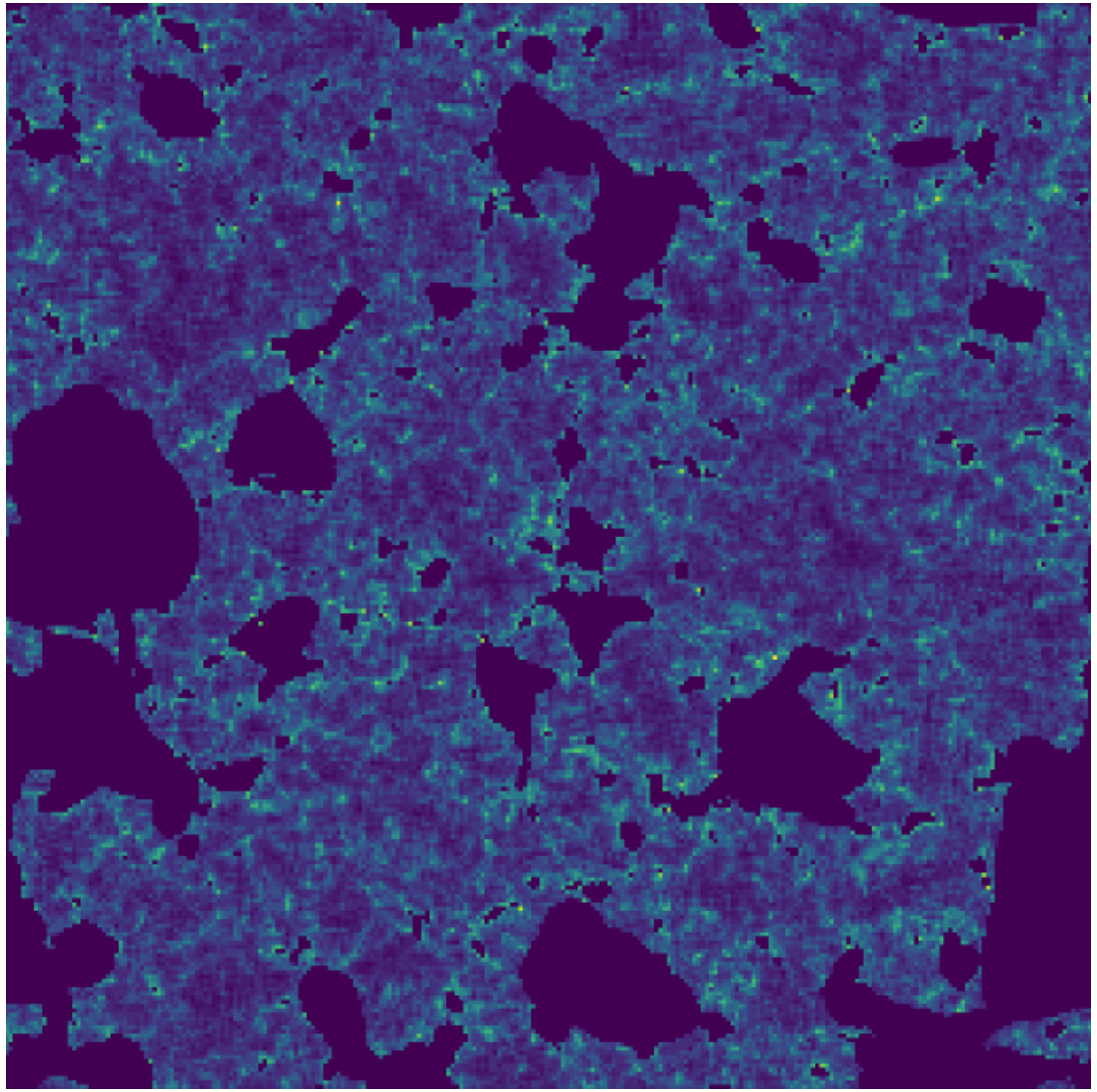


https://github.com/EoRImaging/eor_limits/tree/main

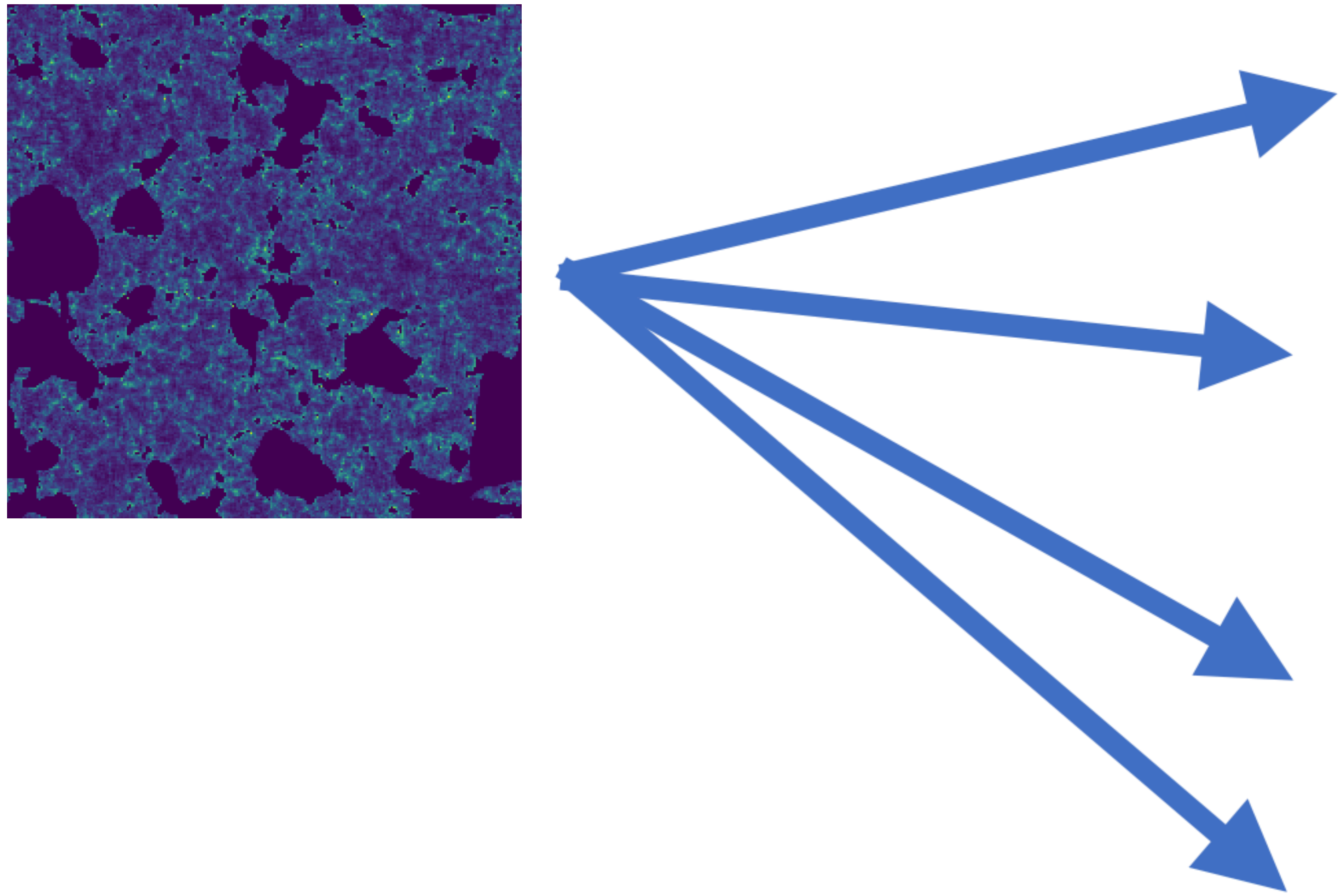
The need for 3D?

2+1 Statistics

2+1 Statistics



Consider a 2+1 statistic with the tools we have

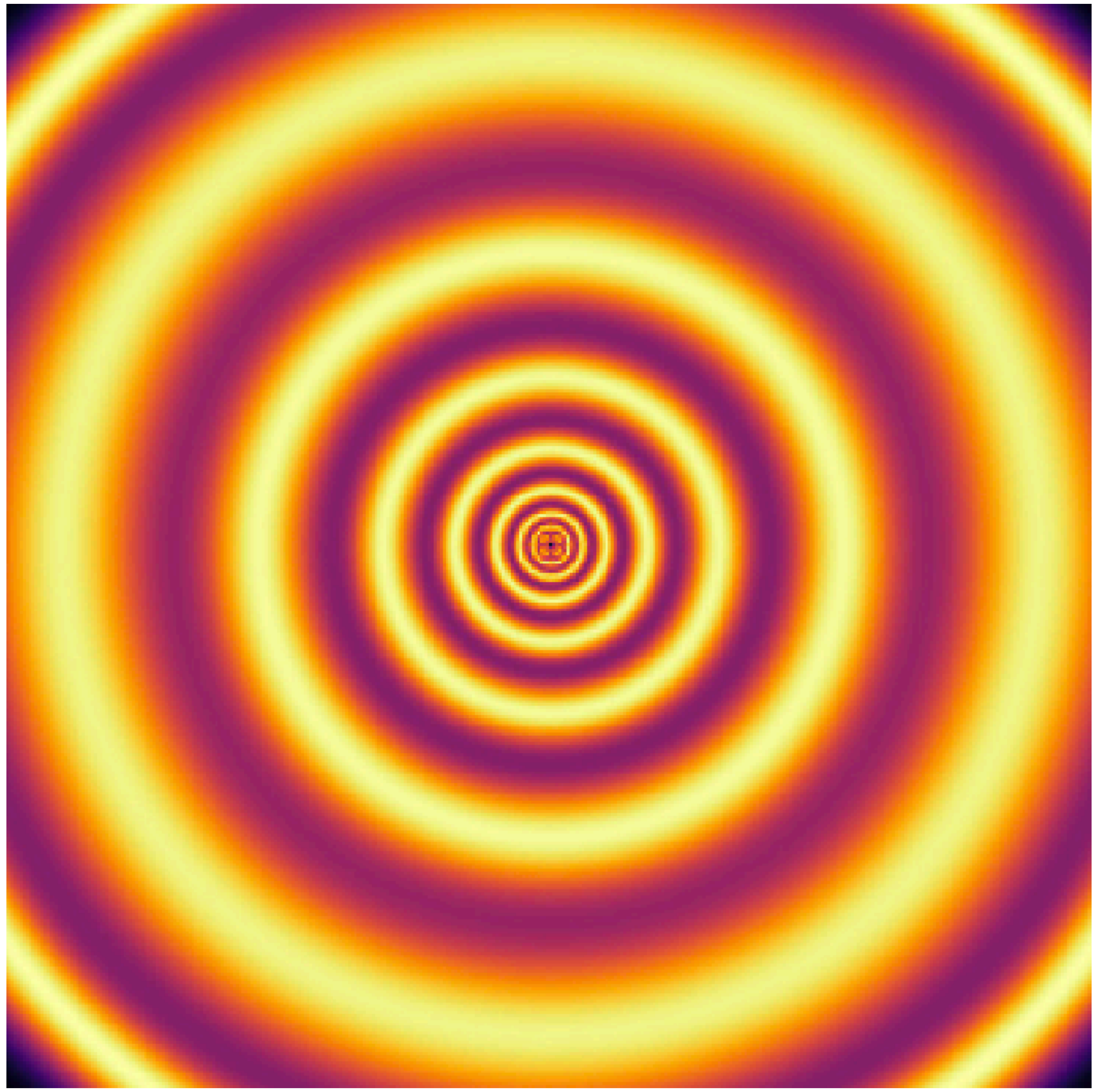


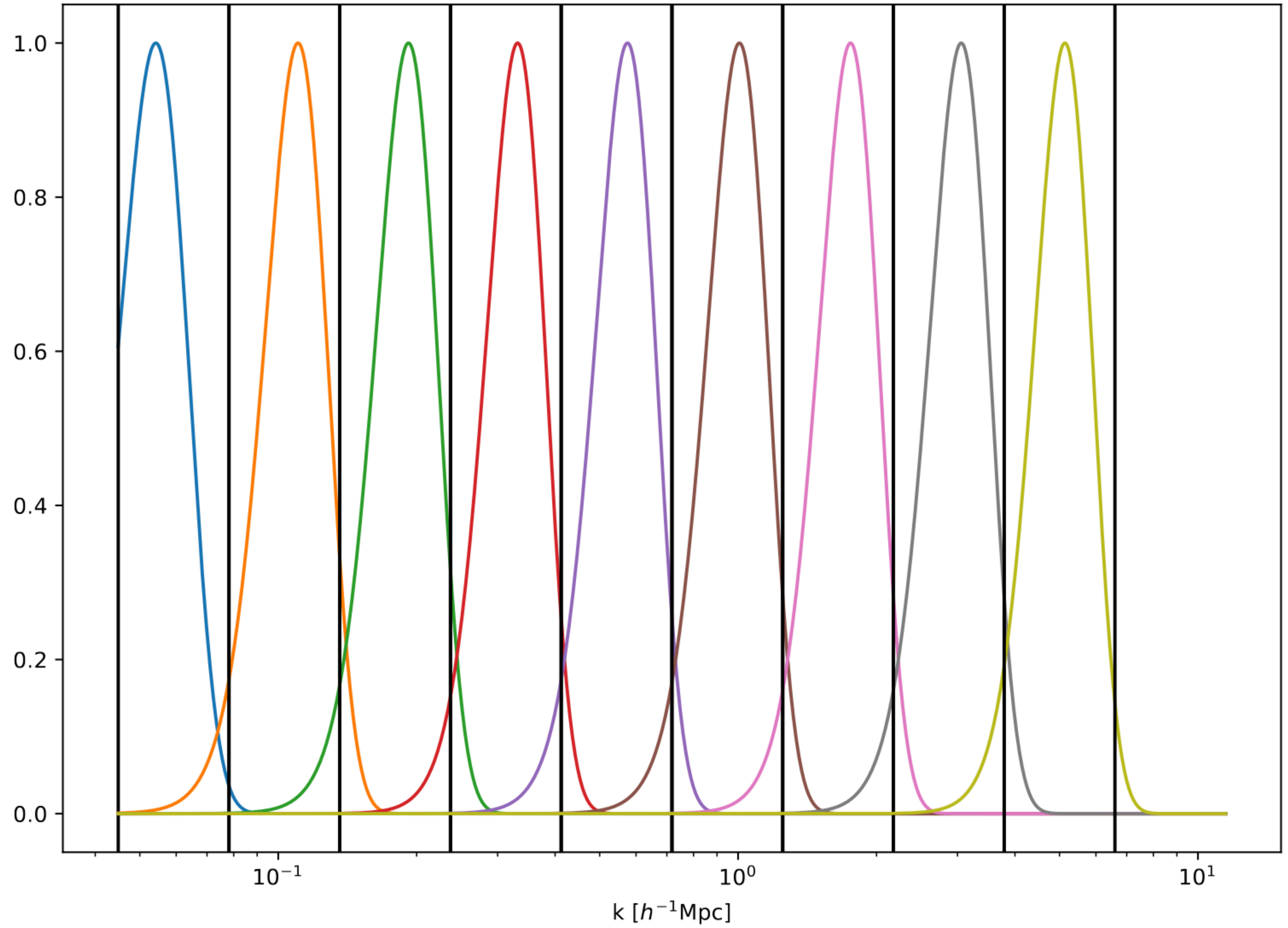
Coefficient
1
2
3
4

Consider a 2+1 statistic with the tools we have

Power Spectrum

- The power spectrum can be written as: $\int_{k_{min}}^{k_{max}} |\tilde{I}(\vec{k}) \cdot W_k(\vec{k})|^2 d\vec{k}$
- For our binning of the 2D power spectrum, they are concentric circles:





Wavelet Moments

- Definition of a wavelet moment: $\int_{\mathbb{R}^3} |I(\overrightarrow{x}) * \psi_j(\overrightarrow{x})|^q d\overrightarrow{x}$
  - $q > 0$
- When using wavelet statistics, an important question arises: what is the best wavelet to have? What probes the spatial resolutions of interest?
- Let's use the inverse Fourier transform of the binning function as our wavelet:

$$M_q(i) = \int_{\mathbb{R}^3} |I(\overrightarrow{x}) * \tilde{W}_i(\overrightarrow{x})|^q d\overrightarrow{x}$$

- In this work we use q = 1 and q = 2.

Wavelet Moments

- In this work we use q = 1 and q = 2.
- To decorrelate the two moments, we normalise as:

$$\bar{M}_1(i) = \frac{M_1(i)}{\sqrt{M_2(i)}}$$

Wavelet Scattering Transforms

- Wavelet transforms are a mathematical tool that allows for localised representation of data by decomposing it into coefficients that describe different scales and positions within the data.
- Wavelet Scattering Transforms are constructed by performing a series of wavelet transforms and the application of the modulus operator, resulting in the generation of a collection of scattering coefficients.
- The coefficients are constructed layer by layer, and we consider only the coefficients of the first two layers.

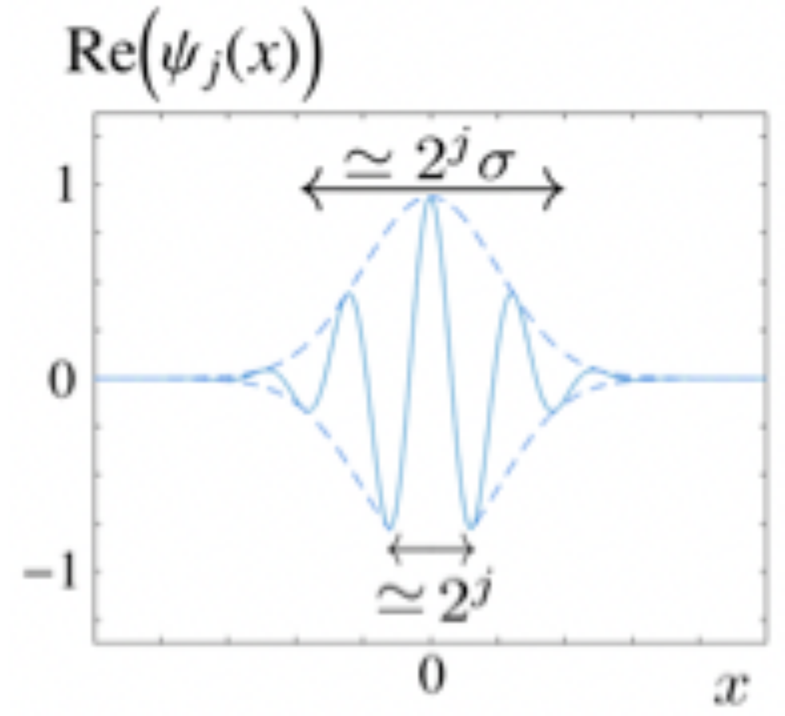
Wavelet Scattering Transforms

- The first layer is constructed by convolving the 2D field $I(\mathrm{x})$ with a family of wavelets ψ_{λ_1} and applying a modulus non-linearity: $S_1(\lambda_1) = \frac{1}{\mu_1} \int |I * \psi_{\lambda_1}|(\mathbf{x}) d^2\mathbf{x}$
- The second layer is constructed by convolving the field again with another family of wavelets ψ_{λ_2} and applying another modulus non-linearity, where $\lambda_1 > \lambda_2$:

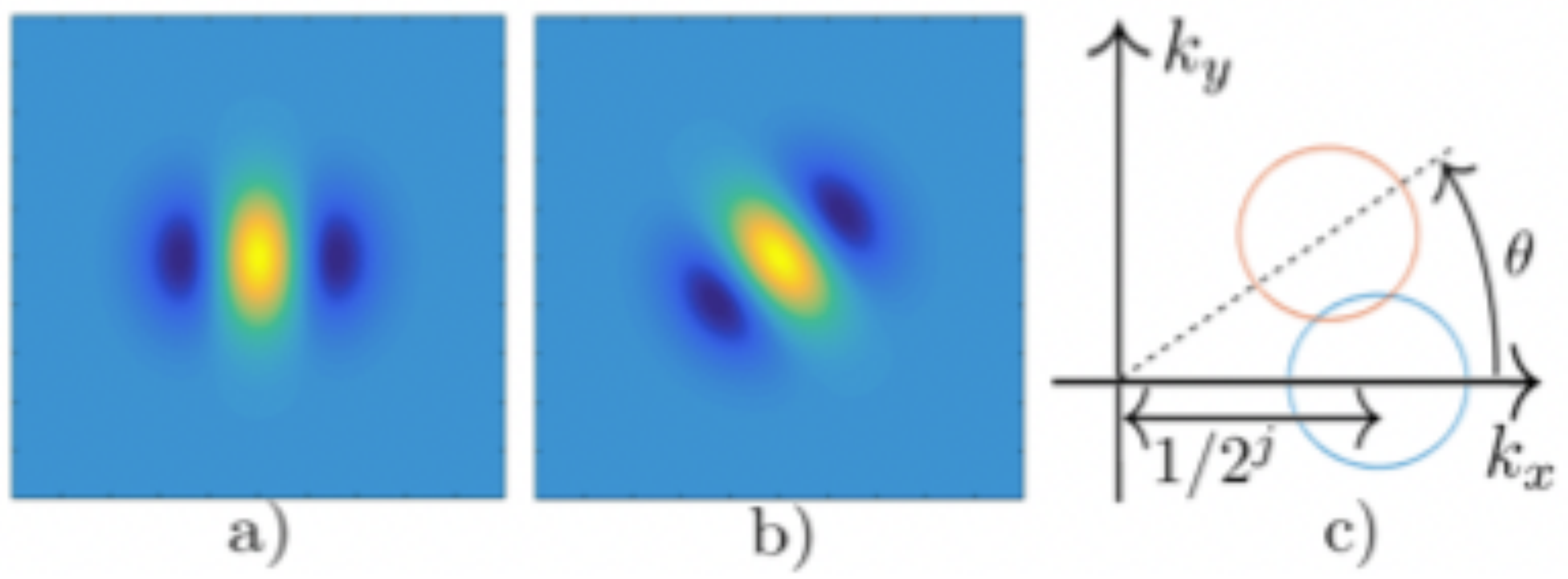
$$S_2(\lambda_1, \lambda_2) = \frac{1}{\mu_2} \int ||I * \psi_{\lambda_1}| * \psi_{\lambda_2}(\mathbf{x}) d^2\mathbf{x}$$

- To take into account the variability of S_2 due to the amplitude of the first wavelet convolution, we follow the usual normalisation by the first layer: $\bar{S}_2(\lambda_1, \lambda_2) = \frac{S_2(\lambda_1, \lambda_2)}{S_1(\lambda_1)}$

Wavelet Scattering Transforms



- For one application, we use the complex Morlet wavelets:
- Here, we dilate our wavelet on scales of 2^j: $\psi_{j,\theta}(\boldsymbol{x}) = 2^{-2j} \cdot \psi\left(2^{-j}\boldsymbol{r}_\theta^{-1}\boldsymbol{x}\right)$
- We then can rotate our wavelet, between 0 and π, we are probing a different region in Fourier space:



Wavelet Scattering Transforms

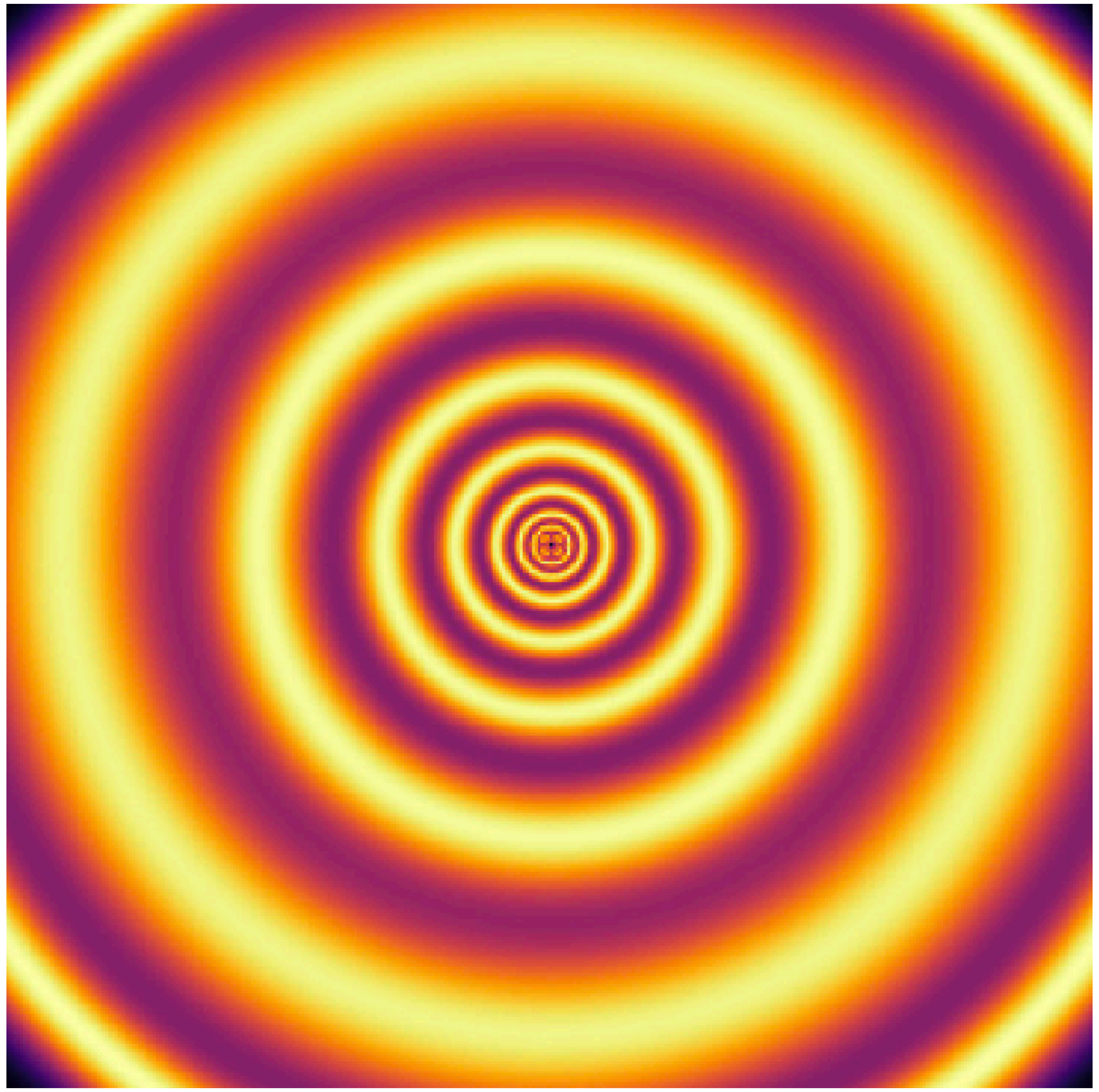
- For this family of wavelets with both angular and scalar dependence, we average over the angular dependence to retrieve completely isotropic features (RWST Allys+19).
- This averaging is on the logarithm of the coefficients:

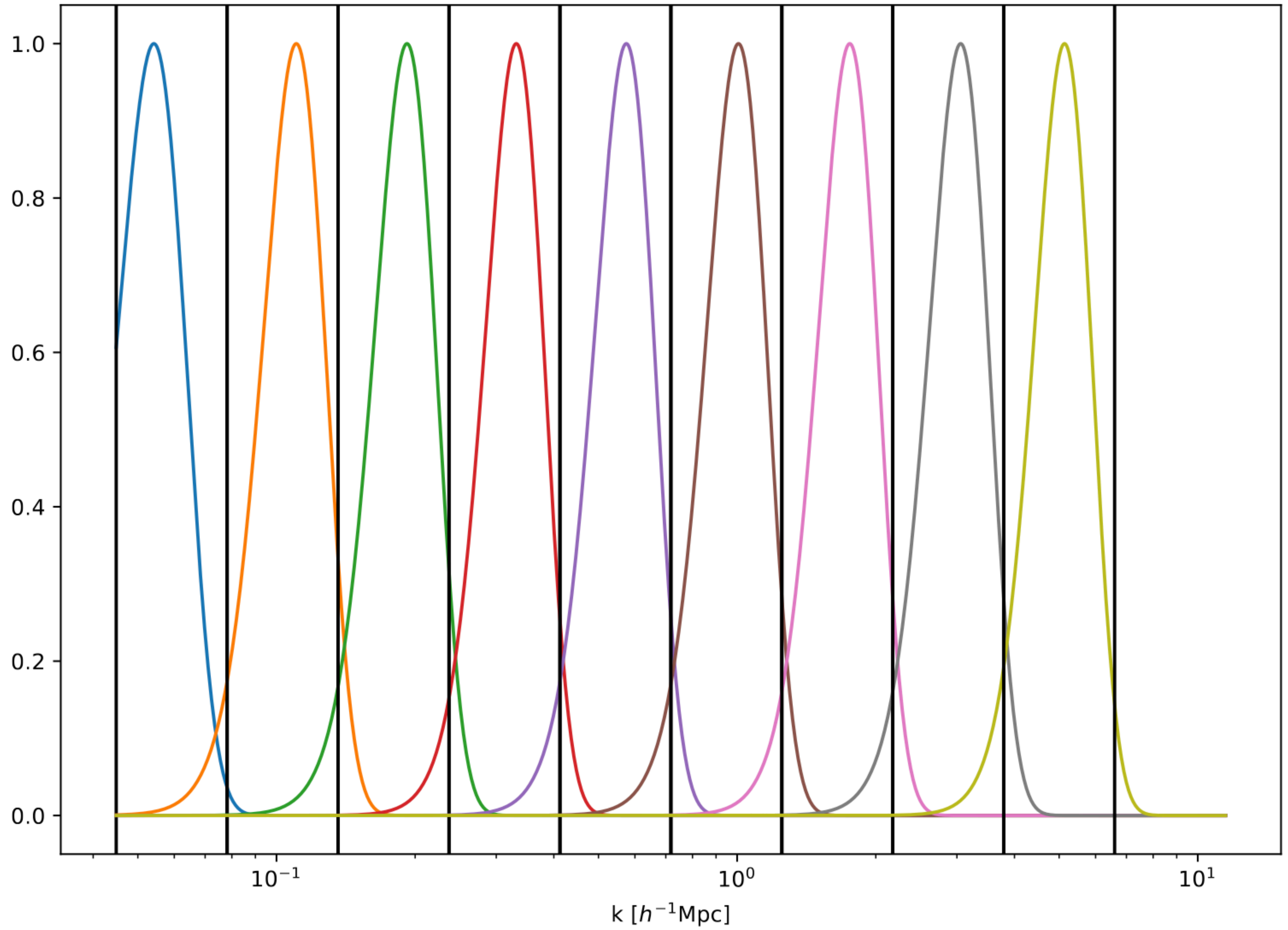
$$S_1^{iso} = \left\langle \log_2\left(S_1(j_1, \theta_1) \right) \right\rangle_{\theta_1}$$

$$\bar{S}_2^{iso} = \left\langle \log_2\left(\bar{S}_2(j_1, \theta_1, j_2, \theta_2) \right) \right\rangle_{\theta_1, \theta_2}$$

Wavelet Scattering Transforms

- For our second application, we use the same wavelets as the application of wavelet moments:





Summarising Our LoS information

- To summarise this statistic, we consider applying a continuous wavelet:

$$\psi(t) = e^{-\frac{t^2}{2}} \cos(5t)$$

- We dilate this wavelet by a factor of 2^{j_z}, to probe different scales.

- Once we apply the Cosine Wavelets to each coefficient, we look to convert this information to a single number. To do this, we consider applying the ℓ^p-norm:

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{\frac{1}{p}}$$

Our Statistics

		Wavelet + Scaling
3D Spherically averaged PS	$\int_{k_{min}}^{k_{max}} \lvert \tilde{I}(\vec{k}) \cdot W_k(\vec{k}) \rvert^2 d\vec{k}$	3D Gaussian + Log10 binning
2+1 PS	$\int_{k_{min}}^{k_{max}} \lvert \tilde{I}(\vec{k}) \cdot W_k(\vec{k}) \rvert^2 d\vec{k}$	2D Gaussian + Log10 binning
Wavelet Moments	$M_{q=1+2}(i) = \int_{\mathbb{R}^2} \lvert I(\vec{x}) * \tilde{W}_i(\vec{x}) \rvert^q d\vec{x}$	2D Gaussian + Log10 binning
WST_m	$S_1^{iso} = \left\langle \log_2\left(S_1(j_1, \theta_1) \right) \right\rangle_{\theta_1}$ $S_2^{iso,1} = \left\langle \log_2\left(S_2(j_1, \theta_1, j_2, \theta_2) \right) \right\rangle_{\theta_1, \theta_2}$	Morlet + Dyadic
WST_w	$S_1(i) = \int \lvert I * \tilde{W}_i \rvert (\mathbf{x}) d^2\mathbf{x}$ $S_2(i_1, i_2) = \int \lvert \lvert I * \tilde{W}_{i_1} \rvert * \tilde{W}_{i_2} \rvert \mathbf{x} d^2\mathbf{x}$	2D Gaussian + Log10 binning

Simulation information

- We use 21cmFast for the simulation:
  - 200$h^{-1}Mpc$ (128x256x256)
  - 128 freq. channels at SKA resolution
  - Simulated between z = 8.82 (144.60 MHz) and z = 9.33 (137.46 MHz).
- We vary the following parameters:
  - T_{vir}: 50000 $\pm$ 5000 K
  - R_{max}= 15 $\pm$ 5 Mpc
  - ζ= 30 $\pm$ 5

Fisher set up

- We use 400 simulations for each parameter change and 600 simulations Fiducial.
- We are fully convergent (<10% err) after our 400 simulations.
- We take the evolution of each at apply the decomposition over scales of 2^{j_z}:
  - j_z = 1,2,3,4 with the ℓ^2-norm
  - j_z = 1,2 with the ℓ^1- and ℓ^2-norm
- These provide a good condition number, for the noiseless case, i.e., the condition number is below 10^7

Results

Our statistics will be denoted by $\bar{\phi}^{s}_{l:j}$

s **is the statistic**

l **is the summary used on the evolution**

j **are the scales that are summarised.**

For example, $\bar{\phi}^{WM}_{\ell^1,\ell^2:1,2}$ **represents the evolution-compressed wavelet moments, computed at** $j = 1$ **and** $j = 2$ **scales for both** ℓ^1 **and** ℓ^2 **norms.**

Results: Noiseless

- Wavelet Moments provides the most accurate constraints on astrophysical parameters compared to other methods
- 2+1 statistics outperform the 3D power spectrum
- Wavelets-based statistics outperform the power spectra statistics

Statistics (Results are $\log_{10}$)	T_{Vir}	R_{Max}	ζ
ϕ^{PS}_{3D}	7.42	2.01	1.50
$\bar{\phi}^{PS}_{\ell^2:1,2,3,4}$	7.25	2.05	1.37
$\bar{\phi}^{PS}_{\ell^1,\ell^2:1,2}$	7.13	2.04	1.28
$\bar{\phi}^{WM}_{\ell^2:1,2,3,4}$	5.45	0.93	−0.22
$\bar{\phi}^{WM}_{\ell^1,\ell^2:1,2}$	6.25	1.00	0.43
$\bar{\phi}^{WST_m}_{\ell^2:1,2,3,4}$	6.00	0.98	0.27
$\bar{\phi}^{WST_m}_{\ell^1,\ell^2:1,2}$	5.99	1.01	0.30
$\bar{\phi}^{WST_w}_{\ell^2:1,2,3,4}$	5.81	0.58	−0.07
$\bar{\phi}^{WST_w}_{\ell^1,\ell^2:1,2}$	5.79	0.60	−0.05

Table 5: We show the Cramer-Rao bounds for all of our summary statistics, in the case where we have no noise. The bound establishes a lower bound on the variance, i.e. the smallest uncertainty achievable for an unbiased estimate on a given parameter.

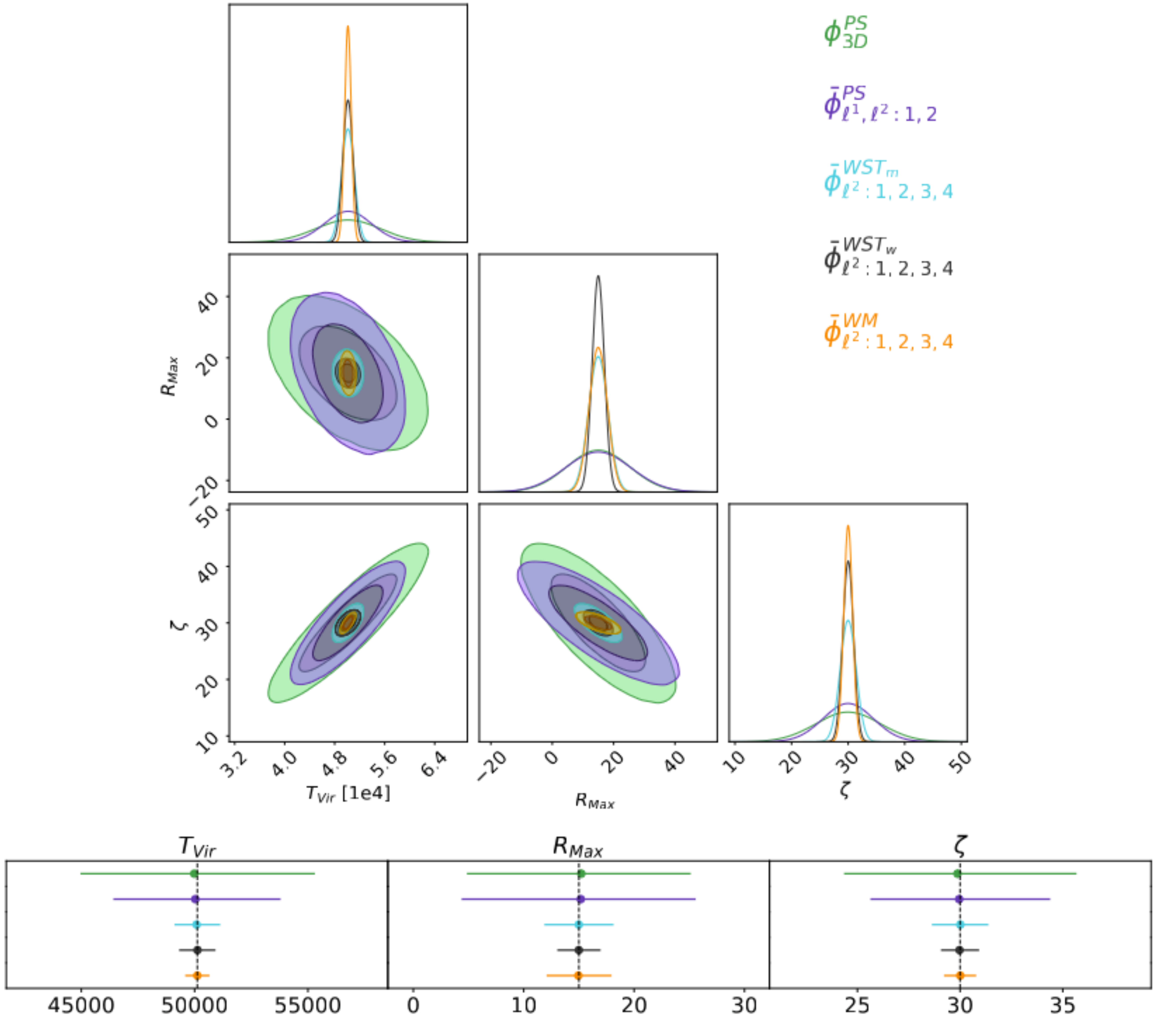


Fig. 4: Results from the Fisher analysis of the three different summary statistics of the 21cm signal, when 3 astrophysical parameter are varied, as we have noiseless data, considering cosmic variance as the only source of variance.
Top: The corner plot of our noiseless Fisher analysis, showing that $\bar{\phi}^{WM}_{\ell^2:1,2,3,4}$ providing the tightest contours. *Bottom:* The ±68% credibility intervals of our different astrophysical parameters, for each statistic. The ordering of the statistics is based on their performance, going from least constraining statistic (top) to most constraining (bottom).

Results: 100 hrs SKA nois

- In the high-noise case, WST_w provides the tightest constraints
- The dyadic scales of WST_m favour the more noise inflicted scales
- The 2+1 statistics, overall, produce the tightest constraints compared to the spherically averaged power spectrum

Statistics (Results are $\log_{10}$)	T_{Vir}	R_{Max}	ζ
ϕ^{PS}_{3D}	9.22	3.60	2.97
$\bar{\phi}^{PS}_{\ell^2:1,2,3,4}$	8.90	3.52	2.80
$\bar{\phi}^{PS}_{\ell^1,\ell^2:1,2}$	8.68	3.24	2.56
$\bar{\phi}^{WM}_{\ell^2:1,2,3,4}$	8.16	2.84	2.27
$\bar{\phi}^{WM}_{\ell^1,\ell^2:1,2}$	8.09	2.76	2.20
$\bar{\phi}^{WST_m}_{\ell^2:1,2,3,4}$	9.31	3.88	3.21
$\bar{\phi}^{WST_m}_{\ell^1,\ell^2:1,2}$	9.22	3.78	3.12
$\bar{\phi}^{WST_w}_{\ell^2:1,2,3,4}$	7.31	1.73	1.29
$\bar{\phi}^{WST_w}_{\ell^1,\ell^2:1,2}$	7.28	1.71	1.27

Table 6: We show the Cramer-Rao bounds for all of our summary statistics, in the case where we have 100 hours of SKA noise.

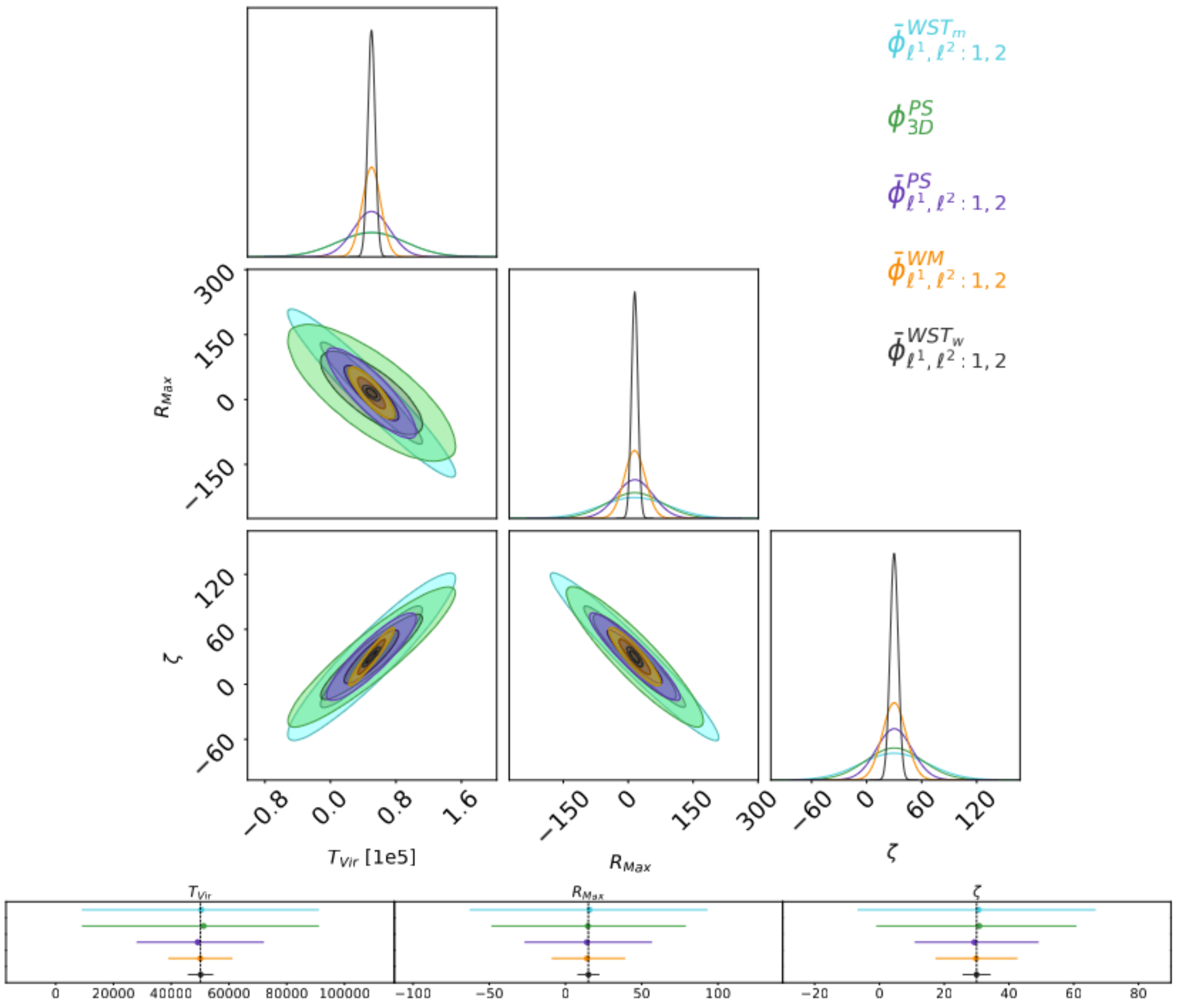
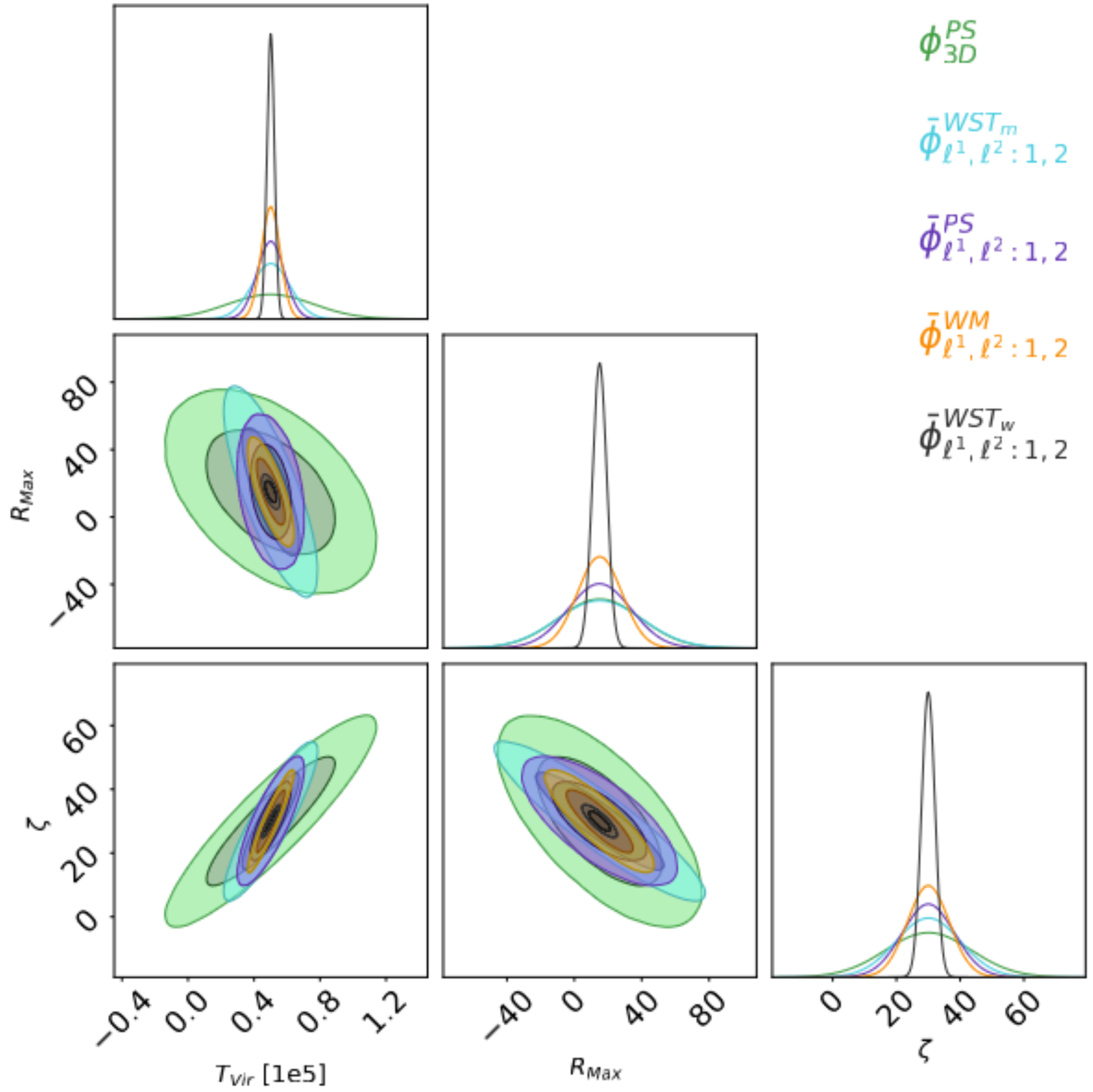
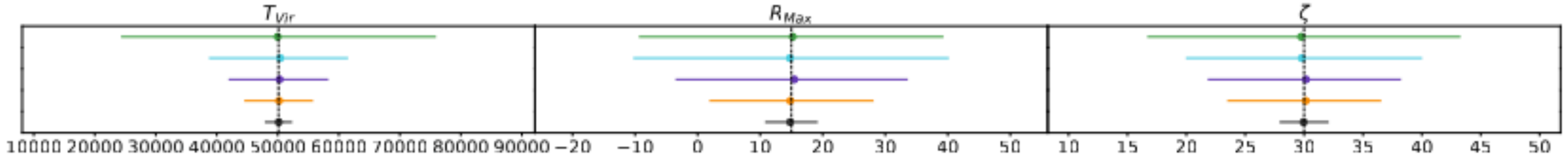


Fig. 5: The same as Fig. 4, but for 100 hours of SKA noise, where the noise is the dominant source of variance. We see now that WST_m is the worse performing statistic, and WST_w, with its with its evolution along the lightcone summarised by the ℓ^1-norm and ℓ^2-norm on scales $j_z = 1,2$, which utilises wavelets derived from the power spectra binning provides the tightest contours.

Results: 1000 hrs SKA noise

- In the lower noise case, all 2+1 statistics outperform the spherically-averaged power spectrum
- WST_w continues to produce produce the tightest constraints
- The Wavelet-based statistics using wavelets derived from PS binning produce the tightest constraints.





Statistics (Results are $\log_{10}$)	T_{Vir}	R_{Max}	ζ
ϕ_{3D}^{PS}	8.82	2.77	2.24
$\bar{\phi}_{\ell^2:1,2,3,4}^{PS}$	7.85	2.74	1.94
$\bar{\phi}_{\ell^1,\ell^2:1,2}^{PS}$	7.82	2.53	1.82
$\bar{\phi}_{\ell^2:1,2,3,4}^{WM}$	7.56	2.36	1.70
$\bar{\phi}_{\ell^1,\ell^2:1,2}^{WM}$	7.50	2.23	1.62
$\bar{\phi}_{\ell^2:1,2,3,4}^{WST_m}$	8.10	2.80	2.00
$\bar{\phi}_{\ell^1,\ell^2:1,2}^{WST_m}$	8.11	2.79	2.00
$\bar{\phi}_{\ell^2:1,2,3,4}^{WST_w}$	6.74	1.24	0.65
$\bar{\phi}_{\ell^1,\ell^2:1,2}^{WST_w}$	6.69	1.24	0.63

Table 7: We show the Cramer-Rao bounds for all of our summary statistics, in the case where we have 1000 hours of SKA noise.

Conclusion

- The 2+1 statistics provide tighter constraints compared to the 3D spherically averaged power spectrum
- For the noiseless case, Wavelet Moments provides the tightest constraints
- In the two noise cases, high and low noise, WST_w provides the tightest constraint.