Cosmology from the Topology of Large Scale Structure

Stephen Appleby
Asia Pacific Center for Theoretical Physics
31st October, 2024

11th KIAS Workshop on Cosmology and Structure Formation
Gyeongju, Korea

In collaboration with :
Changbom Park, Pravabati Chingangbam, Sungwook E. Hong, Juhan Kim, Ho Seong Hwang, Priya Goyal, Joby Kochappan, Motonari Tonegawa

OVERVIEW

1. Minkowski Functionals — Introduction

2. Information content of Minkowski Functional curves

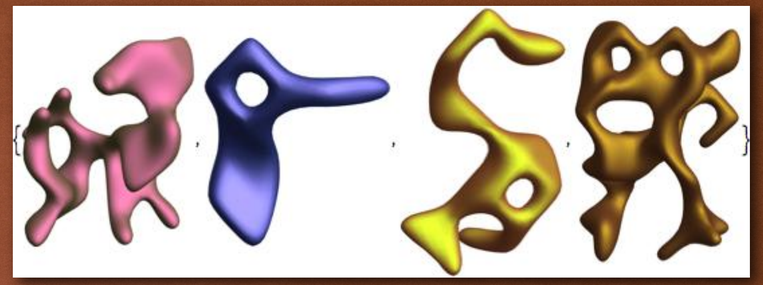
3. Systematic effects in the non-linear regime

4. Future modeling — simulations, approximations

MINKOWSKI FUNCTIONALS

- The Minkowski functionals describe the morphology and topology of a body or set of bodies
- In D dimensions, there are D+1 Minkowski functionals. They form a complete set of valuations.

$$K(A) + K(B) = K(A \cap B) + K(A \cup B)$$

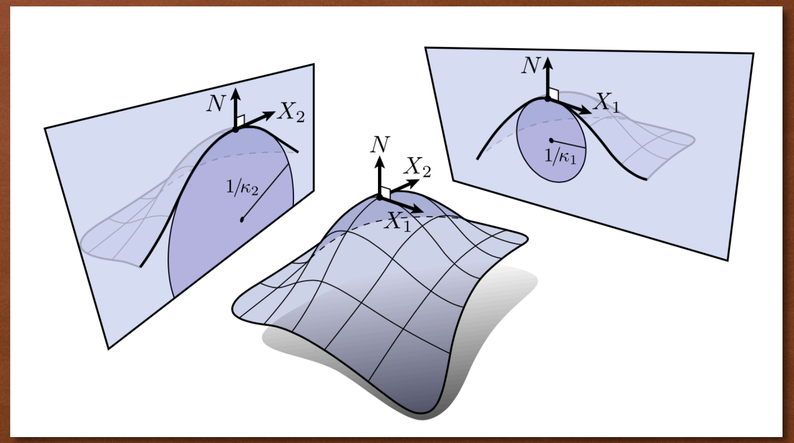


Volume $$W_0 = \frac{1}{V}\int_Q dV$$

Surface Area $$W_1 = \frac{1}{6V}\int_{\partial Q} dA$$

Mean Curvature $$W_2 = \frac{1}{3\pi V}\int_{\partial Q} G_2 dA$$

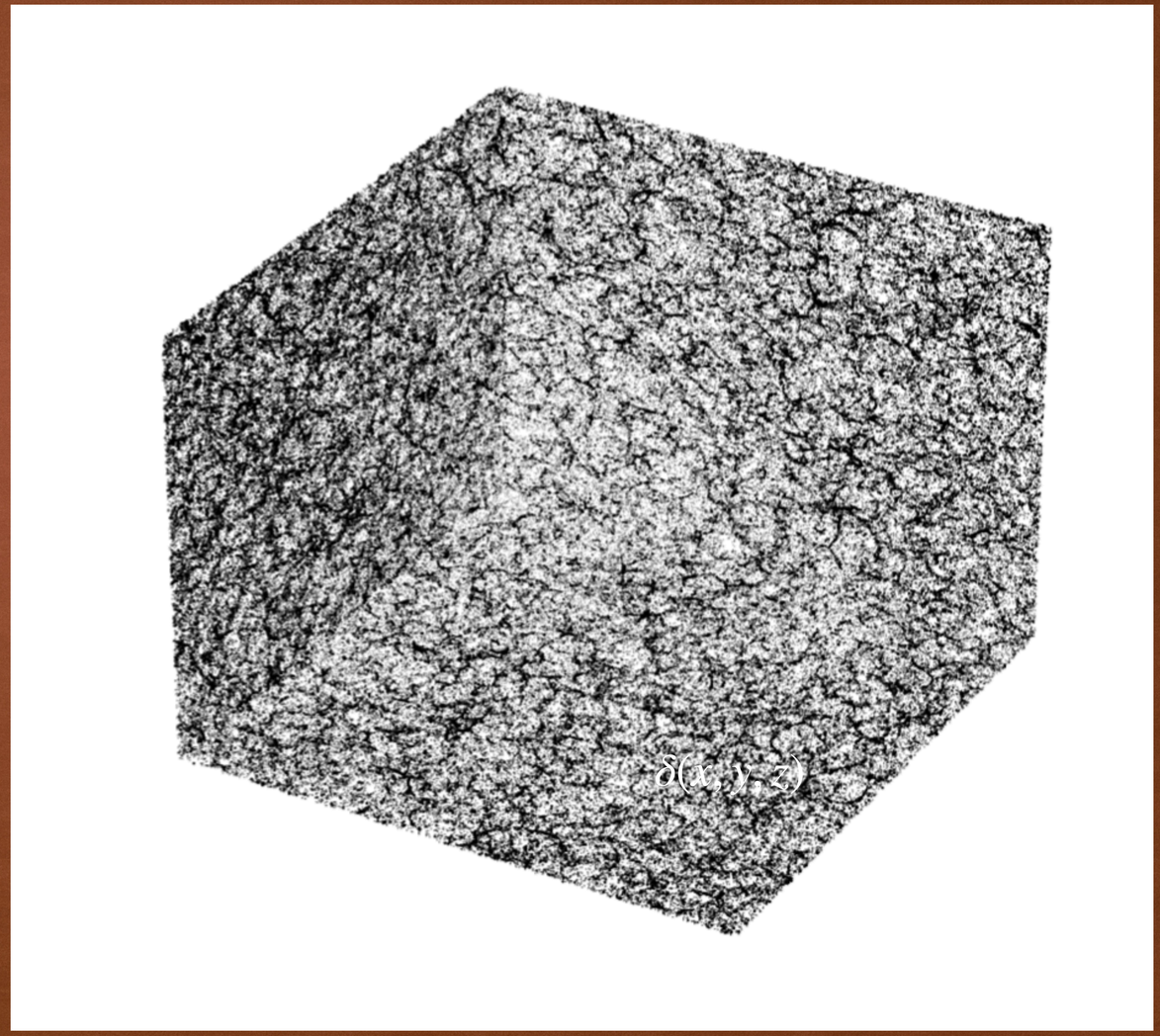
Gaussian Curvature $$W_3 = \frac{1}{4\pi^2 V}\int_{\partial Q} G_3 dA$$



$$G_1 = 1\,,\, G_2 = \frac{1}{2}\left(\frac{1}{\kappa_1} + \frac{1}{\kappa_2}\right)\,,\, G_3 = \frac{1}{\kappa_1 \kappa_2}$$

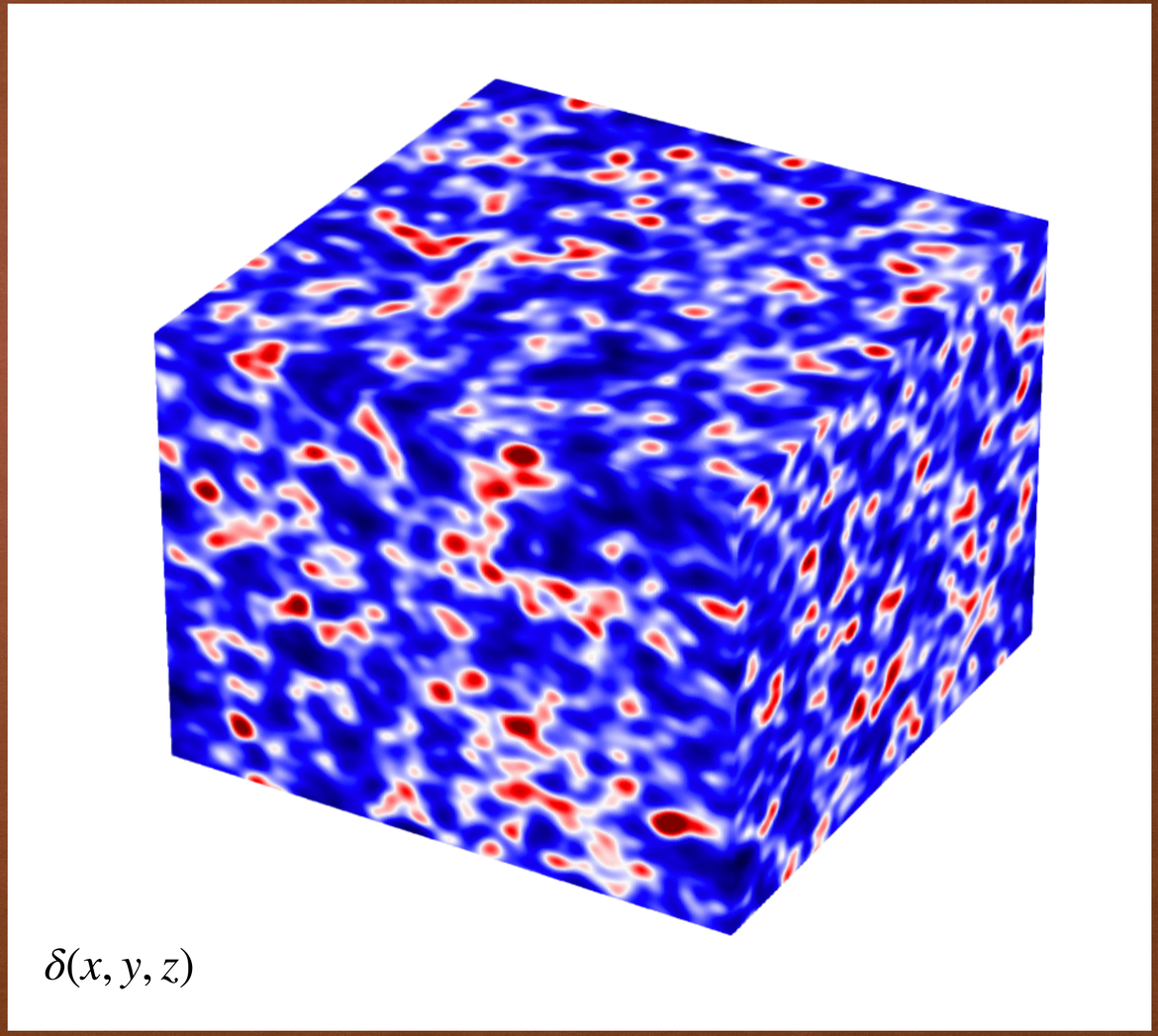
MINKOWSKI FUNCTIONALS

- Generate a continuous density field from a point distribution by smoothing
- Create a boundary of constant field value (in three dimensions, a surface)
- Measure the morphology/topology of the surfaces



MINKOWSKI FUNCTIONALS

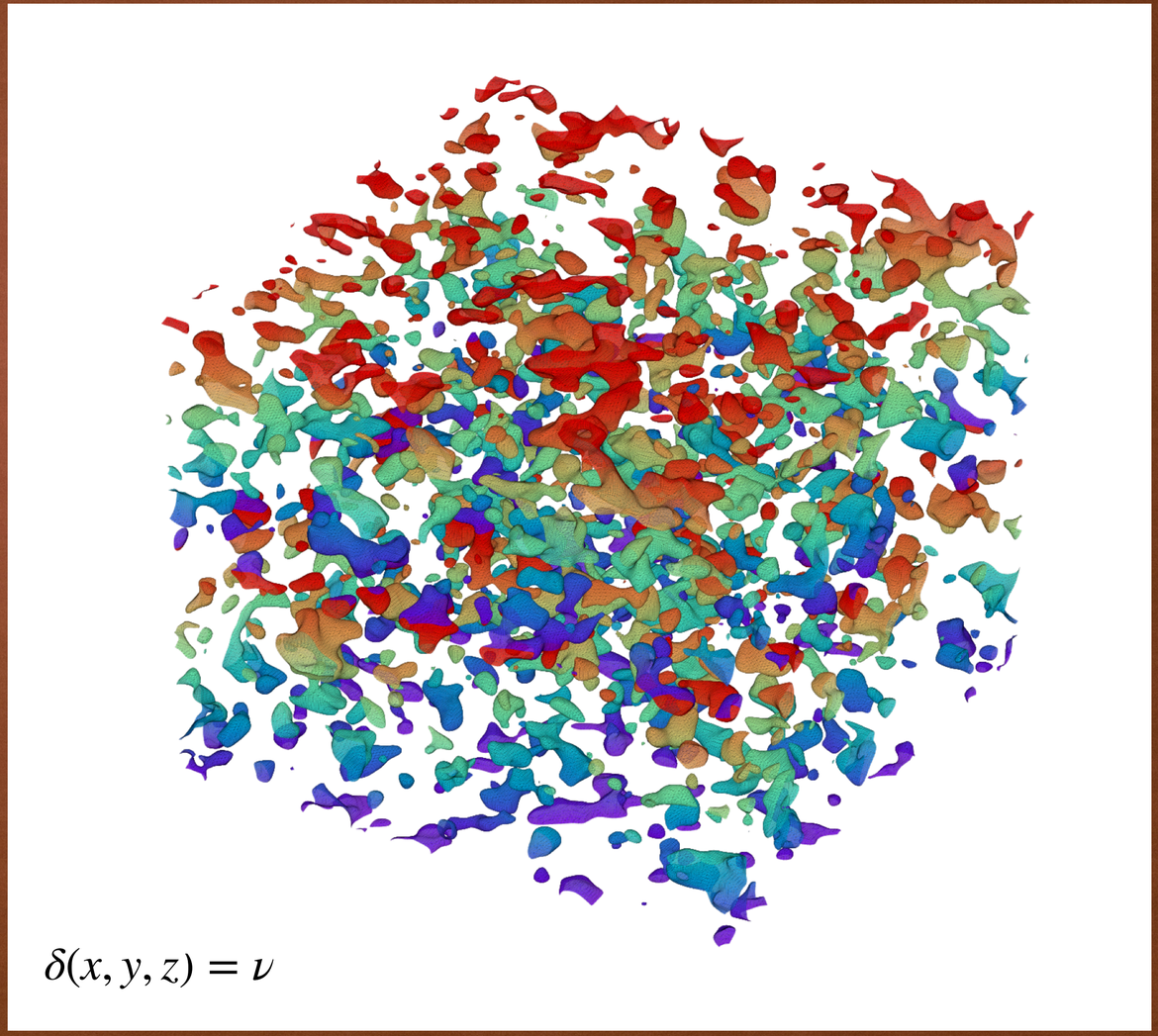
- Generate a continuous density field from a point distribution by smoothing
- Create a boundary of constant field value (in three dimensions, a surface)
- Measure the morphology/topology of the surfaces



$\delta(x, y, z)$

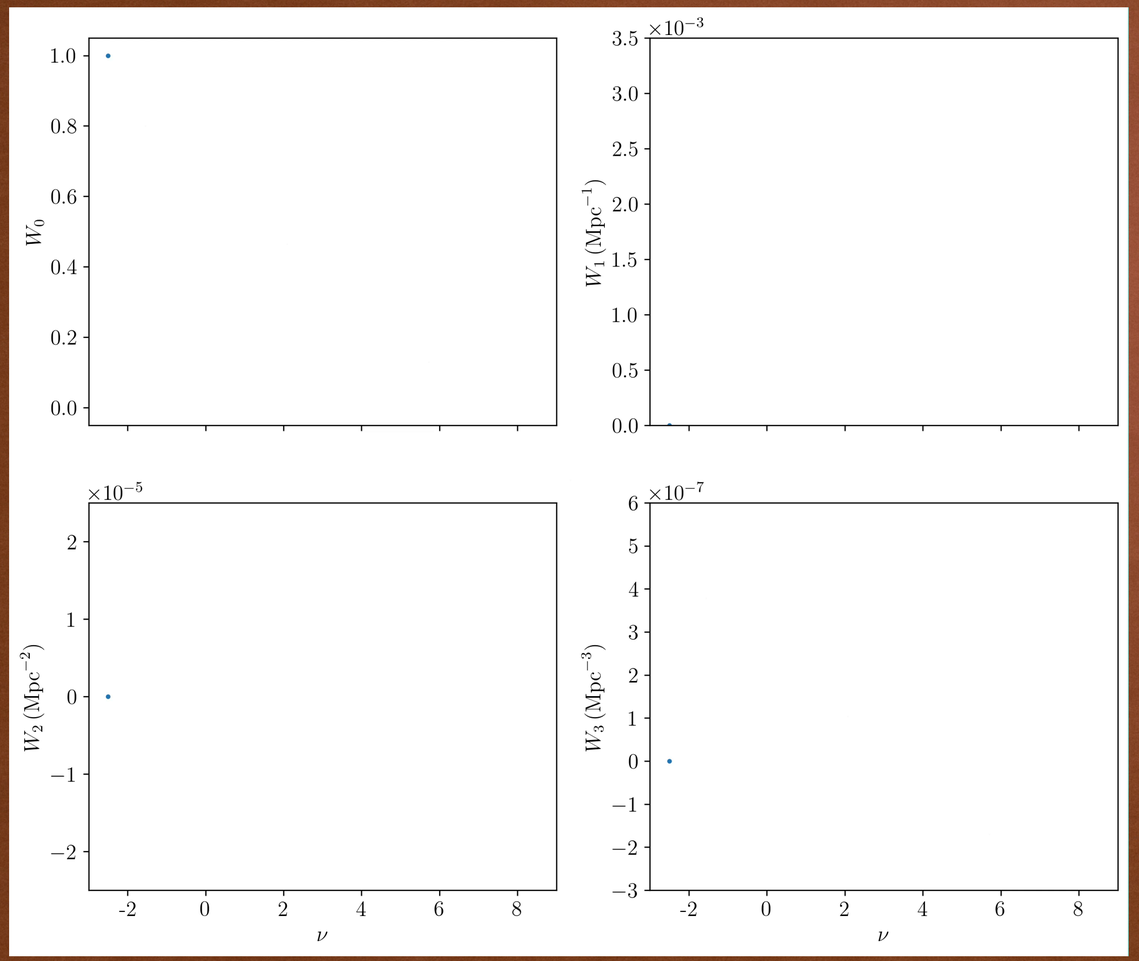
MINKOWSKI FUNCTIONALS

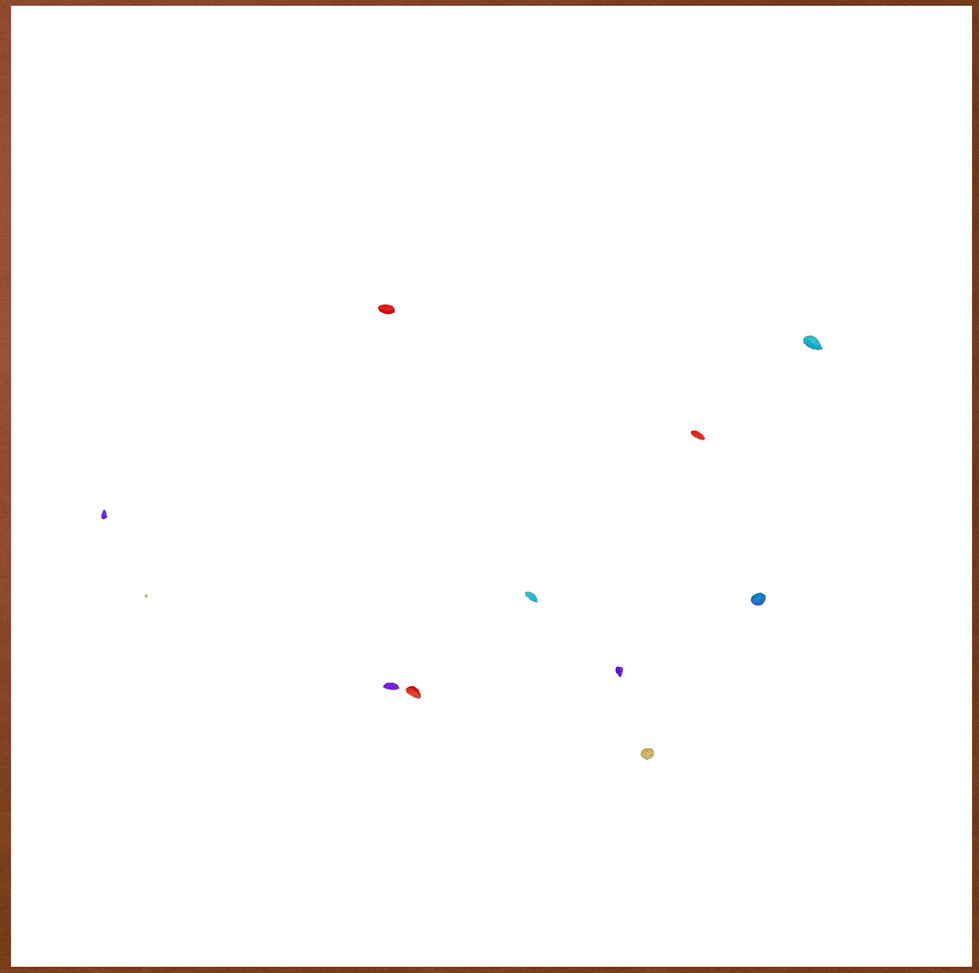
- Generate a continuous density field from a point distribution by smoothing
- Create a boundary of constant field value (in three dimensions, a surface)
- Measure the morphology/topology of the surfaces



MINKOWSKI FUNCTIONALS

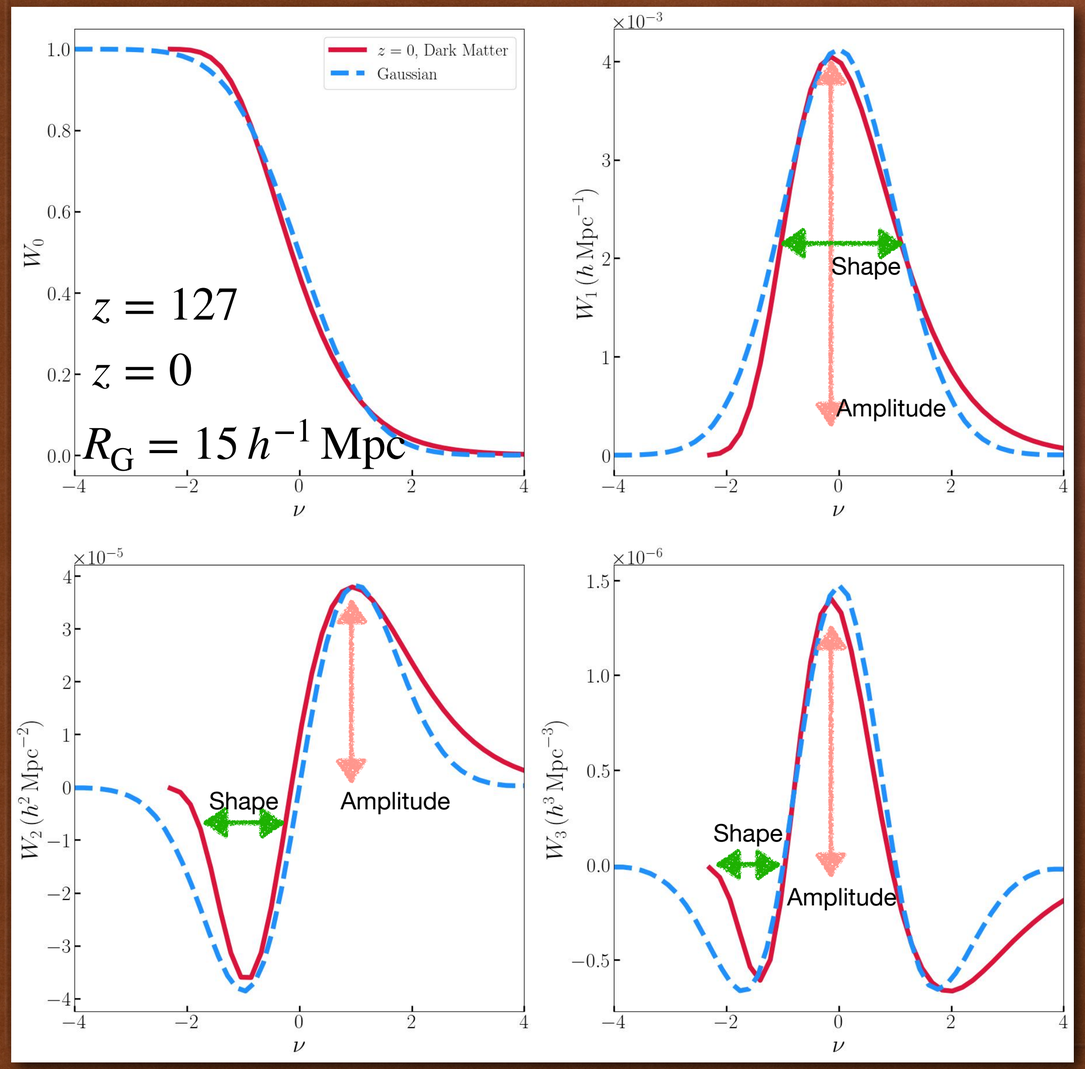
- Create a boundary of constant field value (in three dimensions, a surface)
- Measure the morphology and topology of the excursion sets





INFORMATION CONTENT

- What information is contained within the Minkowski Functional curves? We require a model of W_k as a function of cosmology.



$$W_k(\nu) = \alpha_k \left(\frac{\sigma_1}{\sqrt{3}\sigma_0}\right)^k e^{-\nu^2/2} \times$$

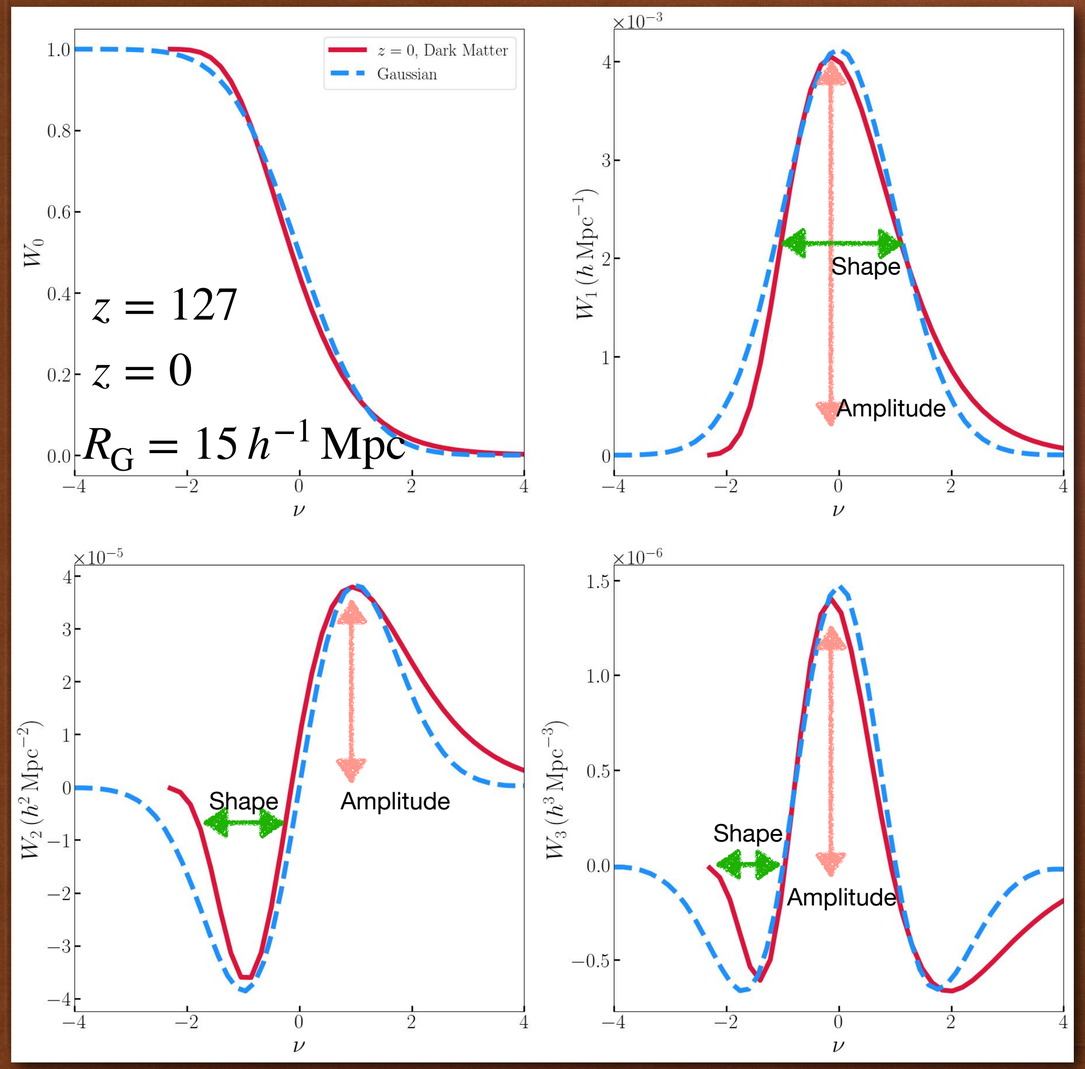
(Matsubara 2000)

$$\left(H_{k-1}(\nu) + \left[\frac{1}{6}S^{(0)}H_{k+2}(\nu) + \frac{k}{3}S^{(1)}H_k(\nu) + \frac{k(k-1)}{6}S^{(2)}H_{k-2}(\nu)\right]\sigma_0 + \mathcal{O}(\sigma_0^2)\right)$$

$$\sigma_0^2 = \langle\delta^2\rangle \qquad S^{(0)} \sim \frac{\langle\delta^3\rangle}{\sigma_0^4} \qquad H_0 = 1, \quad H_1 = \nu, \quad H_2 = 1 - \nu^2$$

INFORMATION CONTENT

- What information is contained within the Minkowski Functional curves? We require a model of W_k as a function of cosmology.



$$W_k(\nu) = \alpha_k \left(\frac{\sigma_1}{\sqrt{3}\sigma_0}\right)^k e^{-\nu^2/2} \times \left(H_{k-1}(\nu) + \left[\frac{1}{6}S^{(0)}H_{k+2}(\nu) + \frac{k}{3}S^{(1)}H_k(\nu) + \frac{k(k-1)}{6}S^{(2)}H_{k-2}(\nu)\right]\sigma_0 + \mathcal{O}(\sigma_0^2)\right)$$

(Matsubara 2000)

- To avoid dominant effects due to shot noise, we must smooth over (at least) the mean separation of the point distribution.

$$\bar{n} \to \bar{r} = \bar{n}^{-1/3}$$

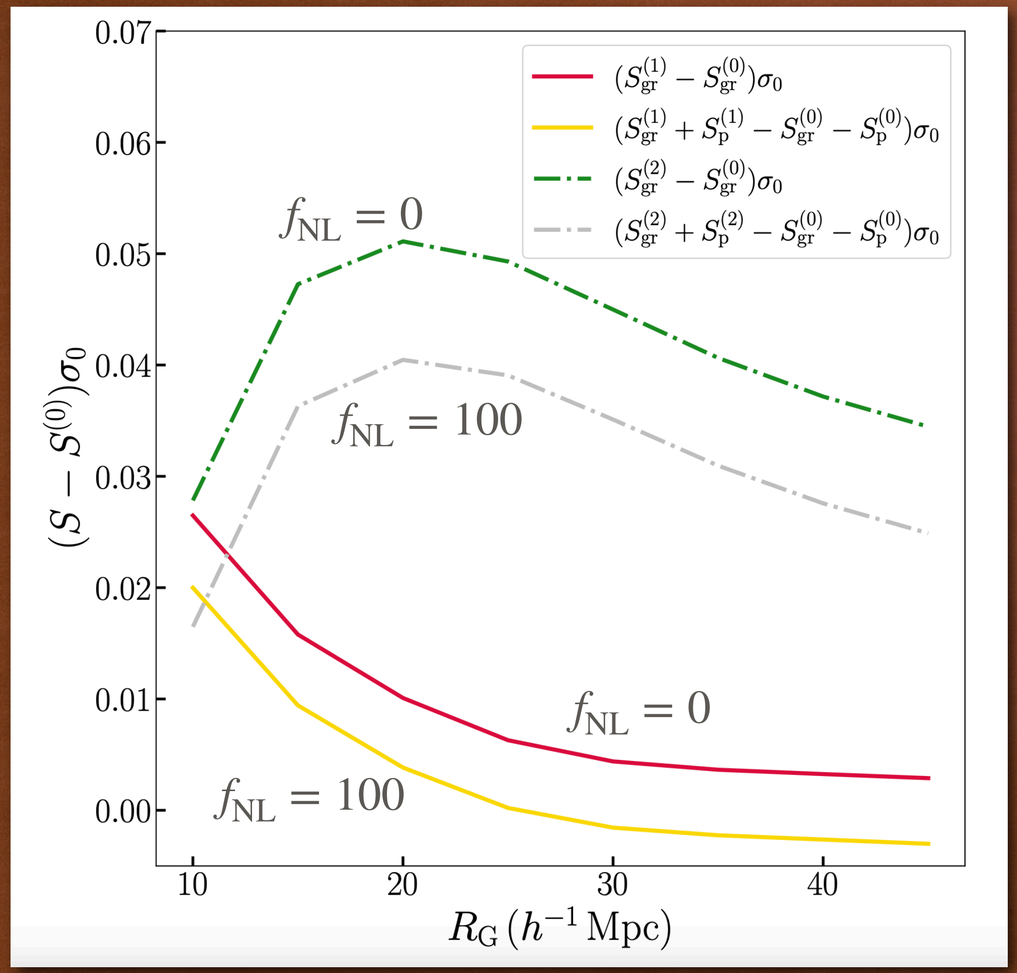
- The next generation of galaxy catalogs will have number densities corresponding to $\bar{r} \sim 10\,h^{-1}\,\mathrm{Mpc}$, so we must model the MF curves down to these scales.

$$\sigma_0^2 = \langle\delta^2\rangle \qquad S^{(0)} \sim \frac{\langle\delta^3\rangle}{\sigma_0^4} \qquad H_0 = 1, \quad H_1 = \nu, \quad H_2 = 1 - \nu^2$$

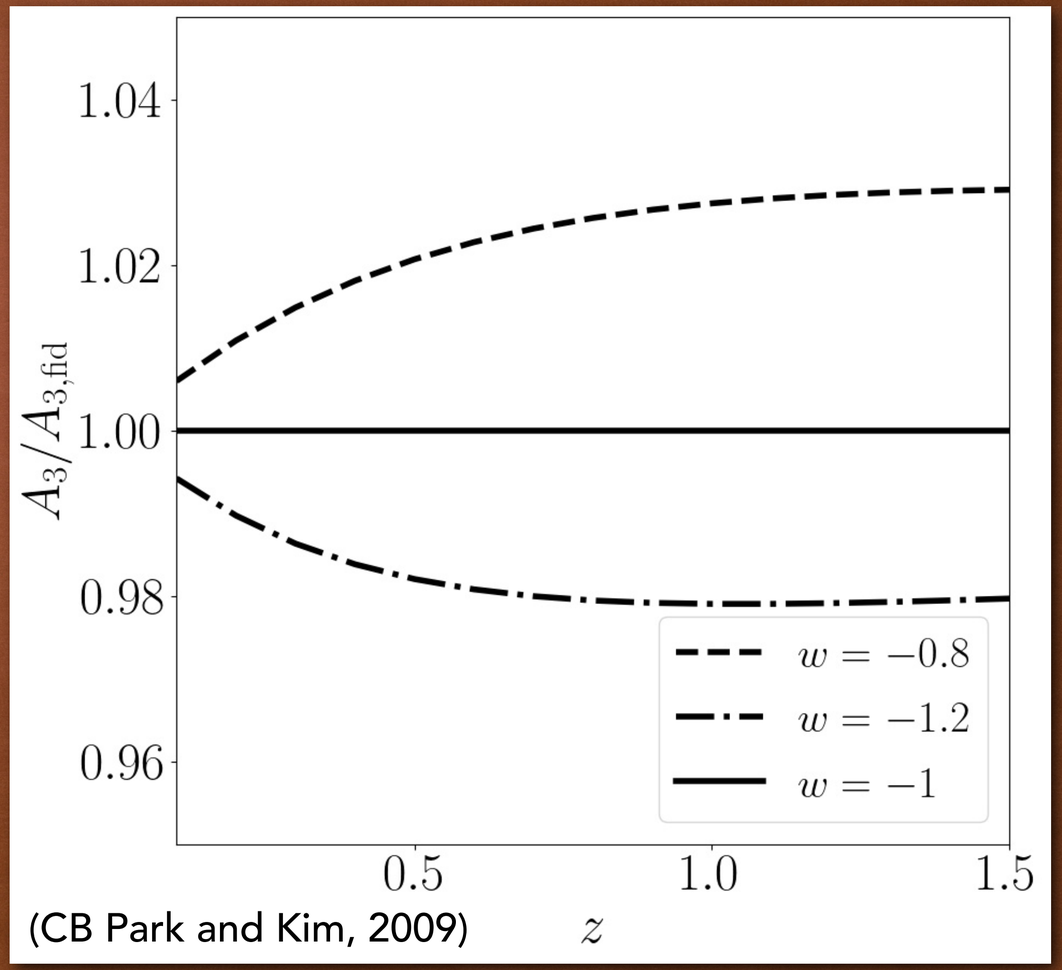
INFORMATION CONTENT

- We require a $\sim 1\,\%$ accurate measurement of the MF curves to do interesting science

Primordial Non-Gaussianity (local type)



Genus Amplitude W_3 as Standard Ruler

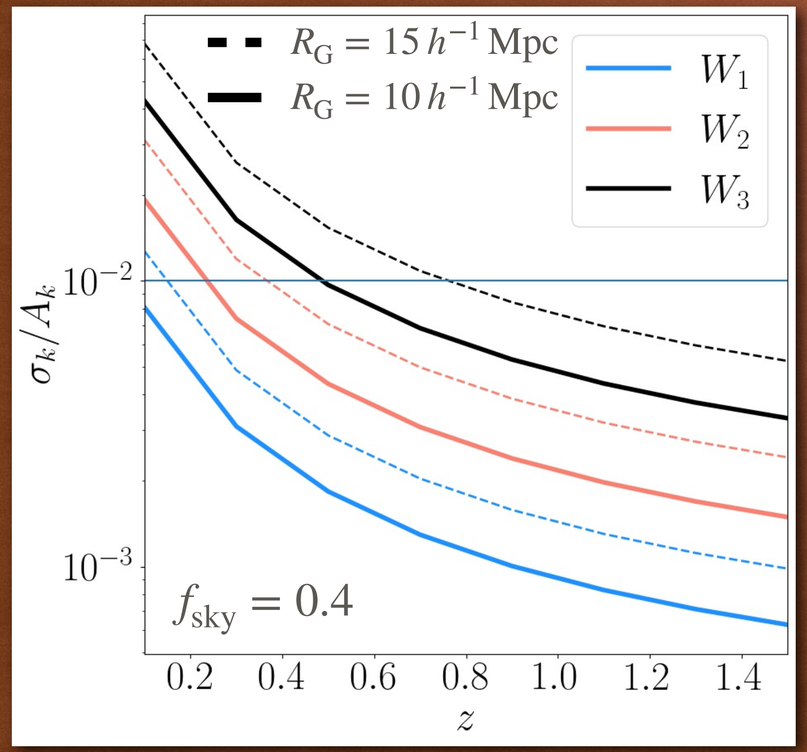


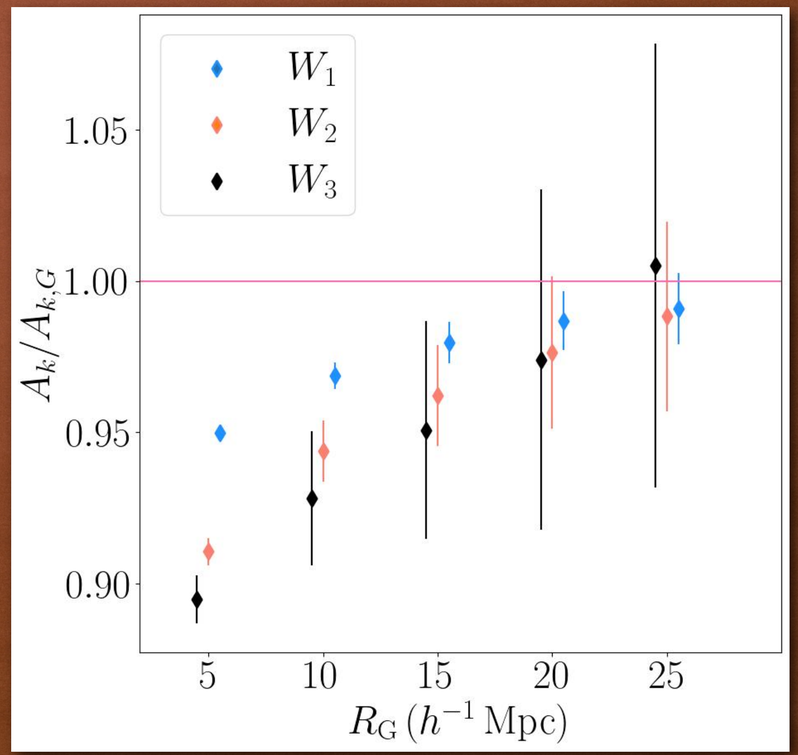
(CB Park and Kim, 2009)

STATISTICAL POWER VS BIAS

- Statistical power increases as we decrease the smoothing scale — more modes.
- Perturbation theory starts to fail at the scales of interest (between $\sim 10 - 20\,h^{-1}\,\mathrm{Mpc}$)

$$W_k(\nu) = \alpha_k \left(\frac{\sigma_1}{\sqrt{3}\sigma_0}\right)^k e^{-\nu^2/2} \times \left(H_{k-1}(\nu) + \left[\frac{1}{6}S^{(0)}H_{k+2}(\nu) + \frac{k}{3}S^{(1)}H_k(\nu) + \frac{k(k-1)}{6}S^{(2)}H_{k-2}(\nu)\right]\sigma_0 + \mathcal{O}(\sigma_0^2)\right)$$





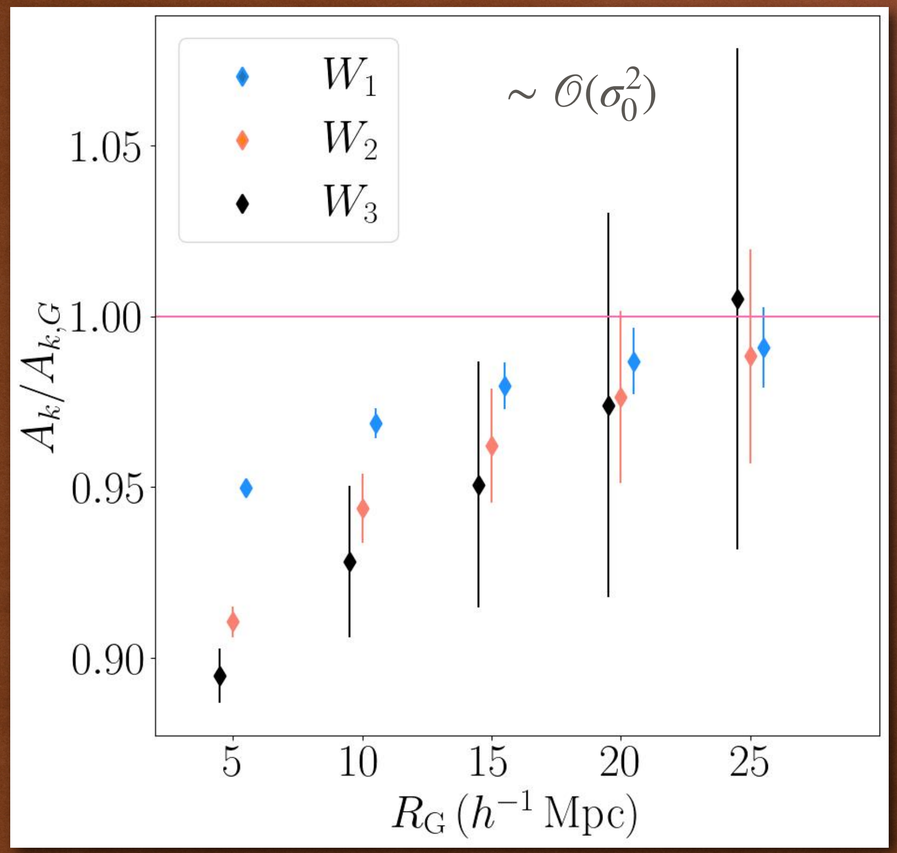
(Matsubara 2000)

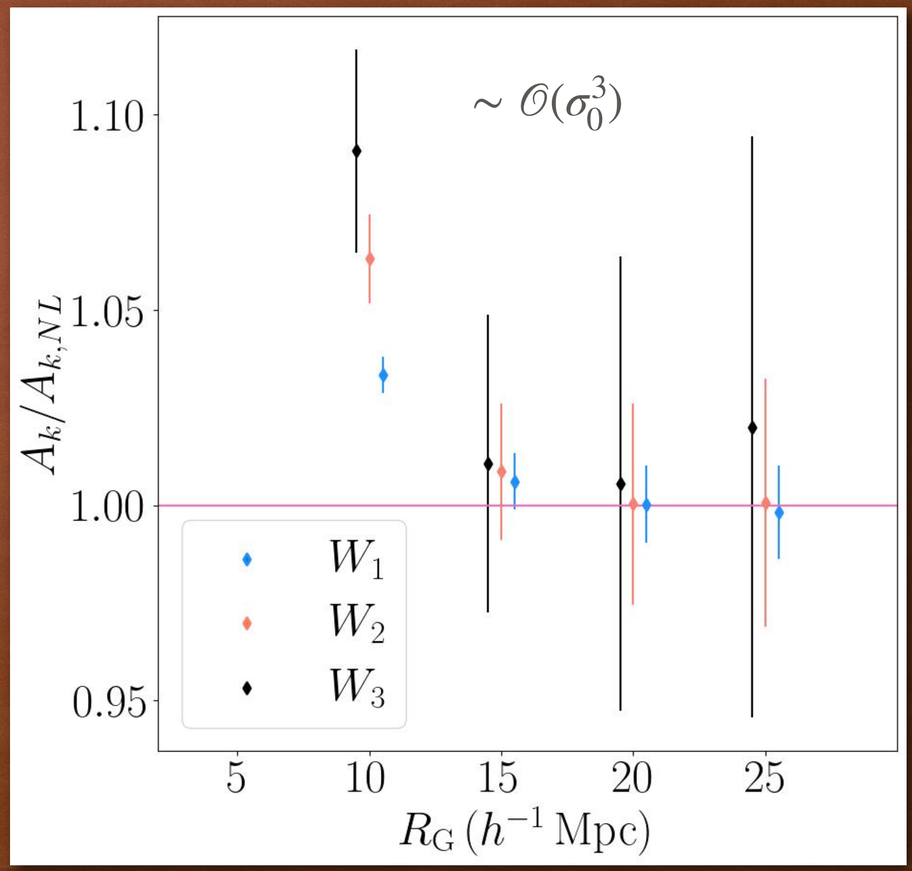
STATISTICAL POWER VS BIAS

- How to get rid of non-linear systematic bias? Go to higher order in perturbation theory…

$$A_k = \alpha_k \left(\frac{\sigma_1}{\sqrt{3}\sigma_0} \right)^k \longrightarrow A_k = \alpha_k \left(\frac{\sigma_1}{\sqrt{3}\sigma_0} \right)^k \left[1 + F\left(K^{(0)}, K^{(1)} \ldots, (S^{(0)})^2, \ldots \right) \times \sigma_0^2 \right]$$

(Matsubara et al. 2020)

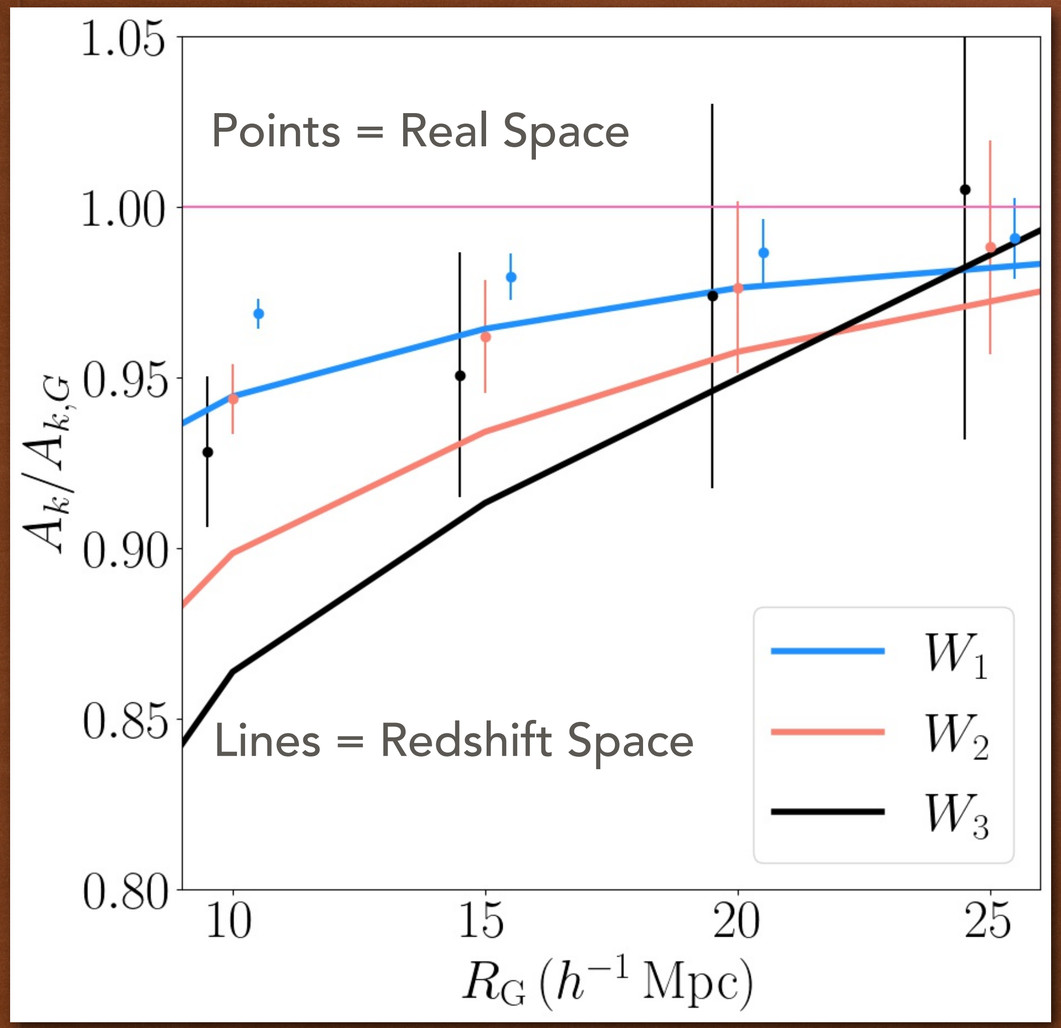




NON-LINEAR SYSTEMATICS – RSD

- Other non-Gaussian processes — Redshift Space Distortion (RSD)

(Matsubara 1996, Codis et al 2013)



Effect on cumulants $\sigma_0^2 = \langle\delta^2\rangle\,, \quad S^{(0)} = \langle\delta^3\rangle/\sigma_0^4 \quad \ldots$

$$\sigma_0^2 \sim \int d^3\mathbf{k}\, P(k)\, W^2(kR_G)$$ Real Space

$$\sigma_0^2 \sim \int d^3\mathbf{k}\, (1+f\mu^2)^2 P(k)\, W^2(kR_G)$$ Kaiser effect

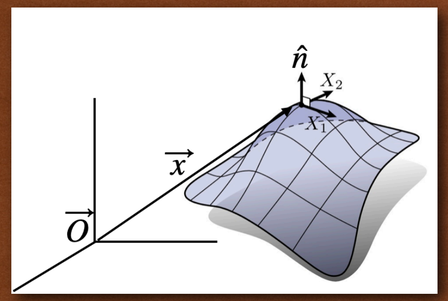
$$\sigma_0^2 \sim \int d^3\mathbf{k}\, (1+f\mu^2)^2 P(k)\, W^2(kR_G)\, \exp\left[-\sigma_v^2(z,R_G)\,\mu^2 k^2\right]$$ "FoG"

Non-linear effects larger in redshift space

NON-LINEAR SYSTEMATICS – RSD

- RSD is better controlled by considering “projections” of the Minkowski Functionals parallel and perpendicular to the line of sight :

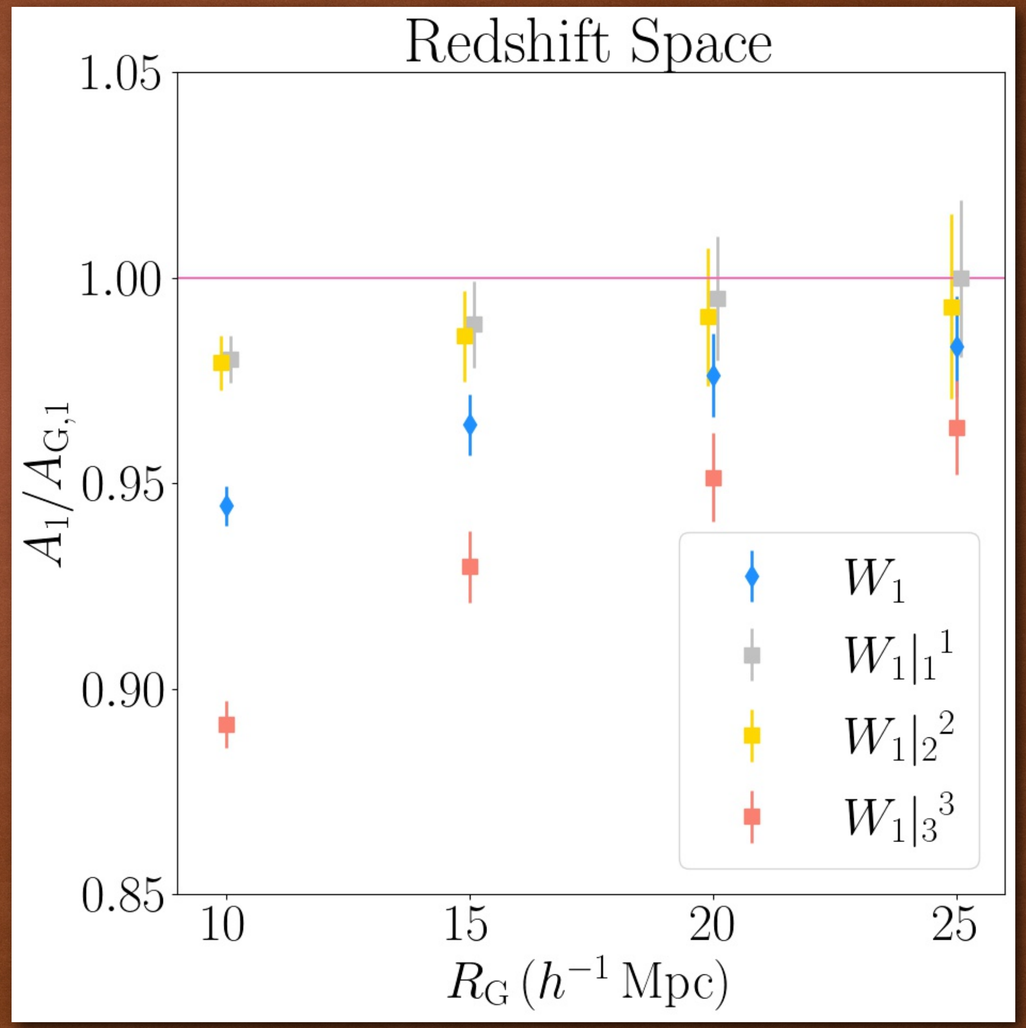
$$W_1 = \frac{1}{6V}\int_{\partial Q} dA$$



$$W_1|_i{}^j = \frac{1}{6V}\int_{\partial Q} \hat{n}_i \hat{n}^j dA$$

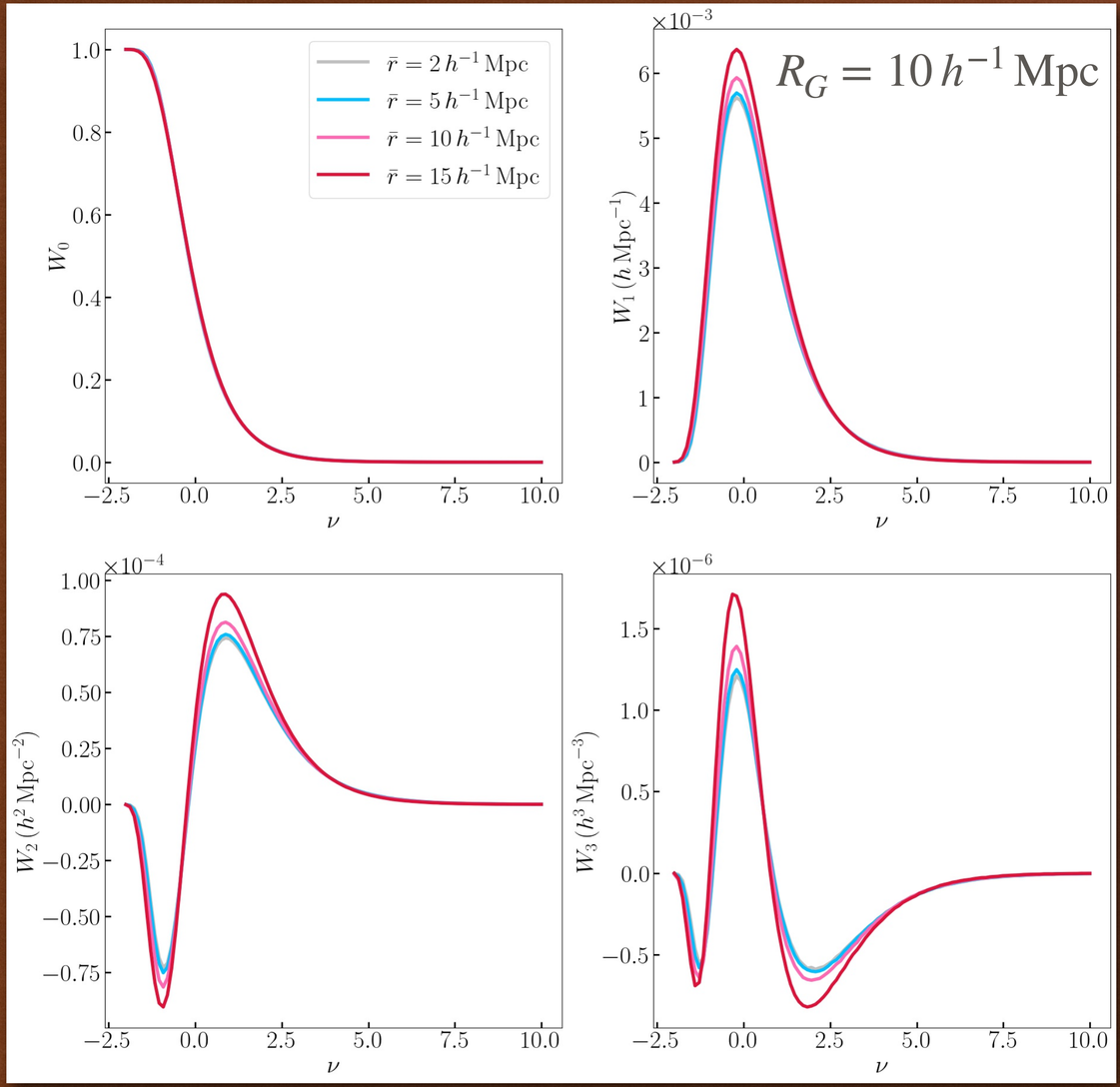
(Chingangbam et al. 2017)

- Non-linear redshift space distortion effects are mitigated perpendicular to the line of sight.



NON-LINEAR SYSTEMATICS – SHOT NOISE

- Other non-Gaussian processes — shot noise and galaxy bias



- We wish to extract as much information as possible from a galaxy catalog → smooth on the mean particle separation.
- Shot noise generates a significant increase in the amplitude of the MFs.
- Can be modeled — $\sigma_0^2 = \langle \delta^2 \rangle$ $S^{(0)} \sim \dfrac{\langle \delta^3 \rangle}{\sigma_0^4}$

$$P_{\mathrm{g}} \to b_1^2 P(k,z) + \frac{1}{\bar{n}}$$

$$B_{\mathrm{g}} = F(k_1, k_2, b_1, b_2, \dots) P(k_1) P(k_2) + \frac{1}{\bar{n}^2} + \frac{1}{\bar{n}} P(k_1) \dots$$

- But, if we get our distances wrong, we can incorrectly infer $\bar{n}$ and obtain spurious amplitudes and sources of non-Gaussianity.

NON-LINEAR MODELING

- Modeling the non-linear regime — use simulations to remove known standard model gravitational effects, only ‘excess’ signal remains?

- Example — ‘Data’ is a snapshot box (Planck cosmology) with a local type PNG initial condition $f_{\rm nl} = 100$. Measure the Minkowski Functionals, extract Hermite coefficients at $z = 0$ —

$$W_k^{(d)} \sim A_k^{(d)} e^{-\nu_V^2/2} \left[H_{k-1} + a_k^{(d)} H_k + b_k^{(d)} H_{k-2} + \dots \right]$$

- We measure the same Minkowski Functionals from ‘mock’ data: $z = 0$ snapshot boxes with $f_{\rm nl} = 0$.

$$W_k^{(m)} \sim A_k^{(m)} e^{-\nu_V^2/2} \left[H_{k-1} + a_k^{(m)} H_k + b_k^{(m)} H_{k-2} + \dots \right]$$

- Subtract the mock values, study the differences

$$\Delta a_k = a_k^{(d)} - a_k^{(m)}$$

$$\Delta b_k = b_k^{(d)} - b_k^{(m)}$$

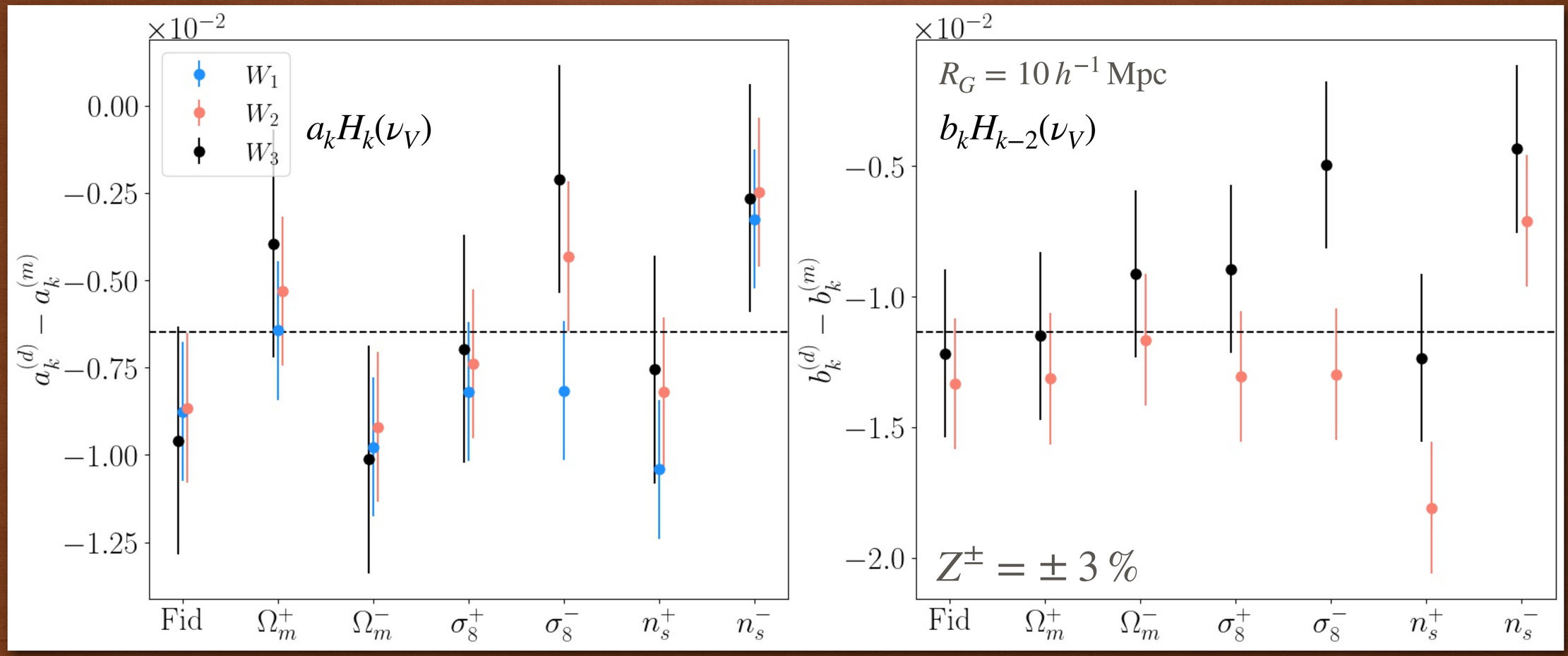


‘Subtract’ known, non-linear gravitational effects, are we left with primordial signal?

NON-LINEAR MODELING

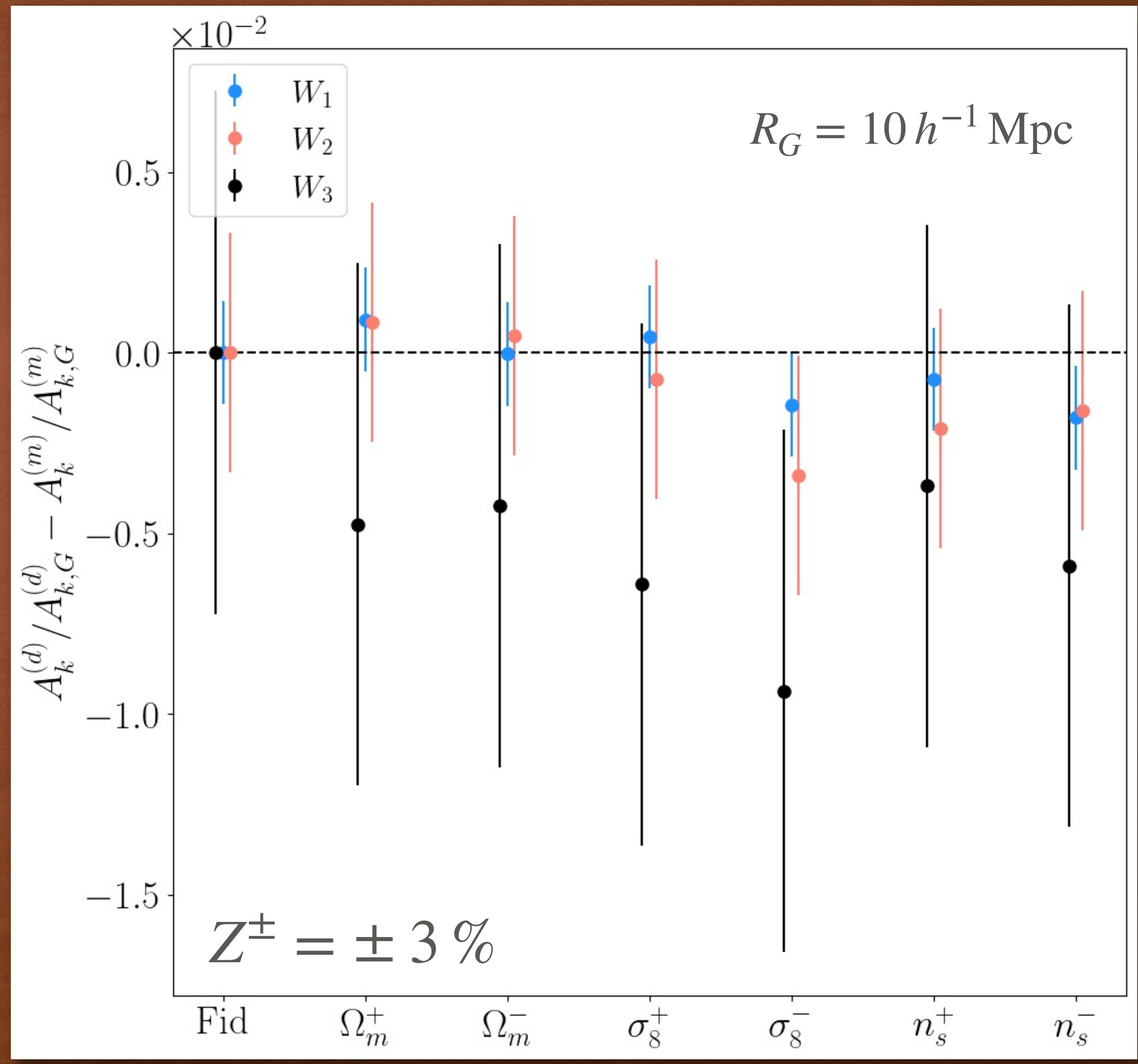
- Modeling the non-linear regime — use simulations to remove known standard model gravitational effects, only ‘excess’ signal remains?

- ‘Data’ is a snapshot box (Planck cosmology) with a local type PNG initial condition $f_{\rm nl} = 100$.



NON-LINEAR MODELING

- Modeling the non-linear regime — use simulations to remove known standard model gravitational effects, only ‘excess’ signal remains?
- ‘Data’ is a snapshot box (Planck cosmology).
- If we are treating the amplitude of the MF curves as a ’standard population’, we do not care about its absolute value.
- We can remove the non-linear contribution by subtraction, but the results rely on the mock data being close to the ‘truth’.
- For modest variations in cosmological parameters, the method seems robust.



NON-LINEAR MODELING

- The MF curves are modeled using the Edgeworth (or Gram-Charlier) expansion

$$\rho(\delta) = \frac{e^{-\delta^2/2\sigma^2}}{\sqrt{2\pi}\sigma} \sum_{N=0}^{\infty} \frac{1}{N!} \sum_{n_1=0}^{\infty} \cdots \sum_{n_N=0}^{\infty} \frac{\langle \delta^{n_1} \rangle}{n_1! \sigma_0^{n_1}} \cdots \frac{\langle \delta^{n_N} \rangle}{n_N! \sigma_0^{n_N}} H_{n_1+\ldots+n_N}(\delta/\sigma_0)$$

- Late time gravitational collapse generates mode couplings, and the cumulants cannot be written recursively:

$$\langle \delta^3 \rangle \sim \int d^3k_1 d^3k_2 F_2(\mathbf{k}_1, \mathbf{k}_2) P(k_1) P(k_2)$$

- The kernel $F(\mathbf{k}_1, \mathbf{k}_2)$ is a non-linear function of the modes; we cannot write $\langle \delta^n \rangle = a_n G\left(\langle \delta^{n-1} \rangle, \ldots\right)$.

- We can potentially simplify the mode couplings. For example, under the Zel'Dovich approximation

$$F_n(\mathbf{k}_1, \ldots \mathbf{k}_n) = \frac{1}{n!} \frac{\mathbf{k}_1 . \mathbf{k}}{k_1^2} \cdots \frac{\mathbf{k}_n . \mathbf{k}}{k_n^2}$$

(Bernardeau 2001)

$\mathbf{k} = \mathbf{k_1} + \ldots + \mathbf{k}_n$

CONCLUSIONS

- The Minkowski Functionals are simple summary statistics that can be extracted from current and future galaxy data.
- We can measure the two- and three-point cumulants and constrain cosmological parameters, when smoothing the galaxy point distributions over quasi-linear scales.
- The next generation of galaxy catalogs will provide data on $\lesssim 10\,h^{-1}\,\mathrm{Mpc}$ scales. Modeling of non-linear effects should be improved in this regime.
- We can use simulations to remove non-linear ‘artifacts’, or use approximations to construct the joint PDF of the density field and its derivatives.
- Cosmological scale hydrodynamical simulations will be required to study the effect of galaxy bias, baryons etc.