Weakly non-Gaussian formulas of cosmological random fields

Based on:
T.M., PRD, **101**, 043532 (2020)
Kuriki, T.M., arXiv:2011.04593 (mathematical paper, still submitted)
T.M., Kuriki, PRD, **104**, 103522 (2022)
T.M., Hikage, Kuriki, PRD, **105**, id.023527 (2022)

KEK
Takahiko Matsubara

5/10/2022, 17:00-18:00 (JST)

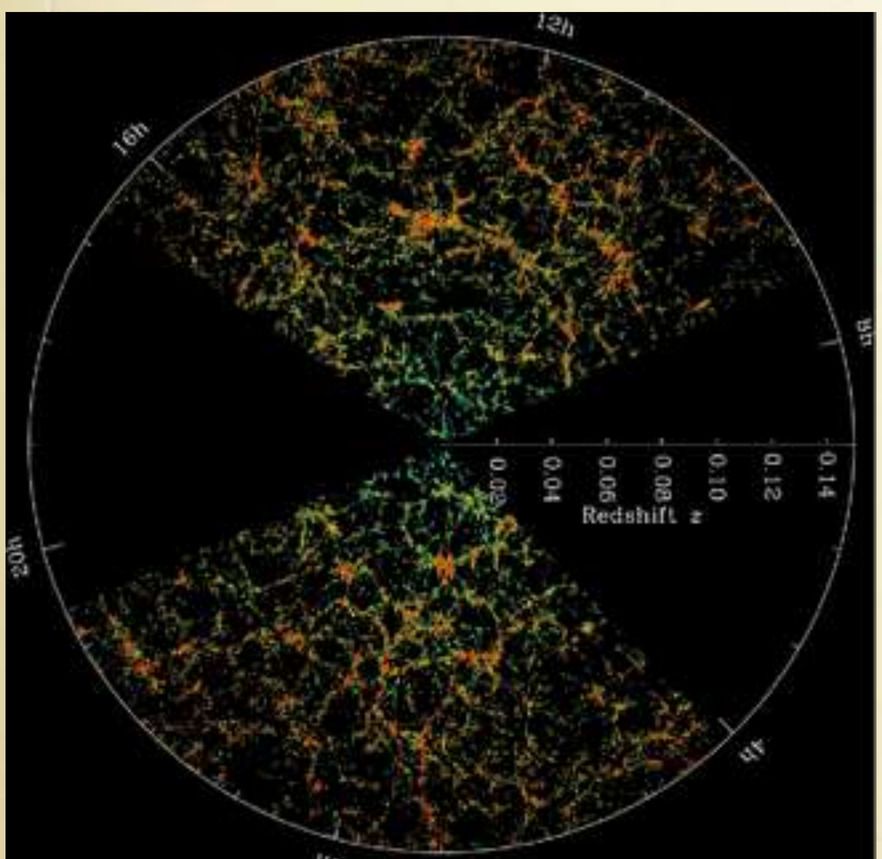
IAU-IAA Astrostats and Astroinfo seminar

Cosmological observables

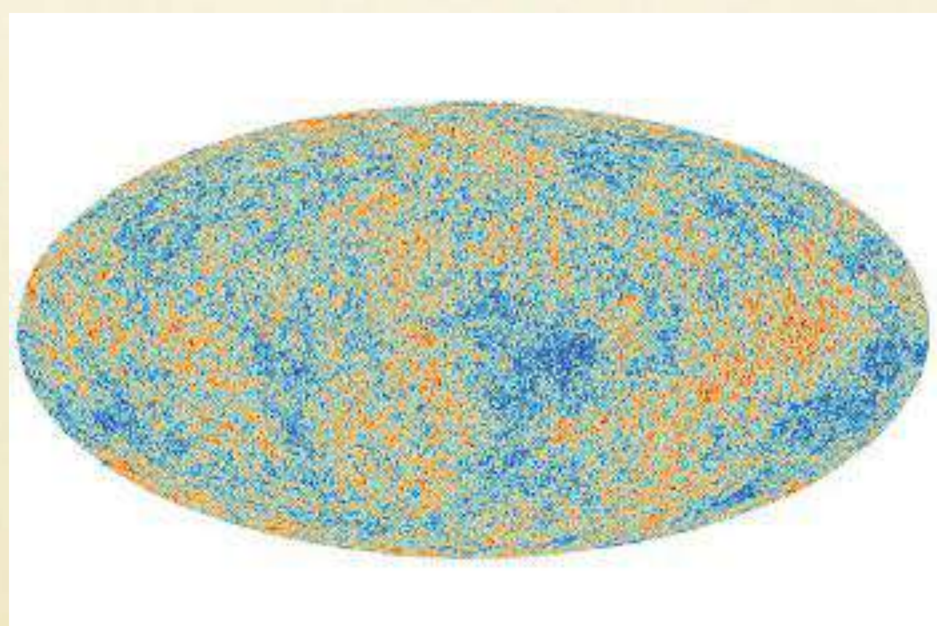
- Cosmology: studying the Universe as a whole
  - Not individual objects, but an overall tendency to study the properties of the Universe (space-time)
  - Thus, cosmological observables are necessarily statistical quantities
- Cosmological random fields
  - In recent cosmology, valuable information is provided by spatial distributions of various observables

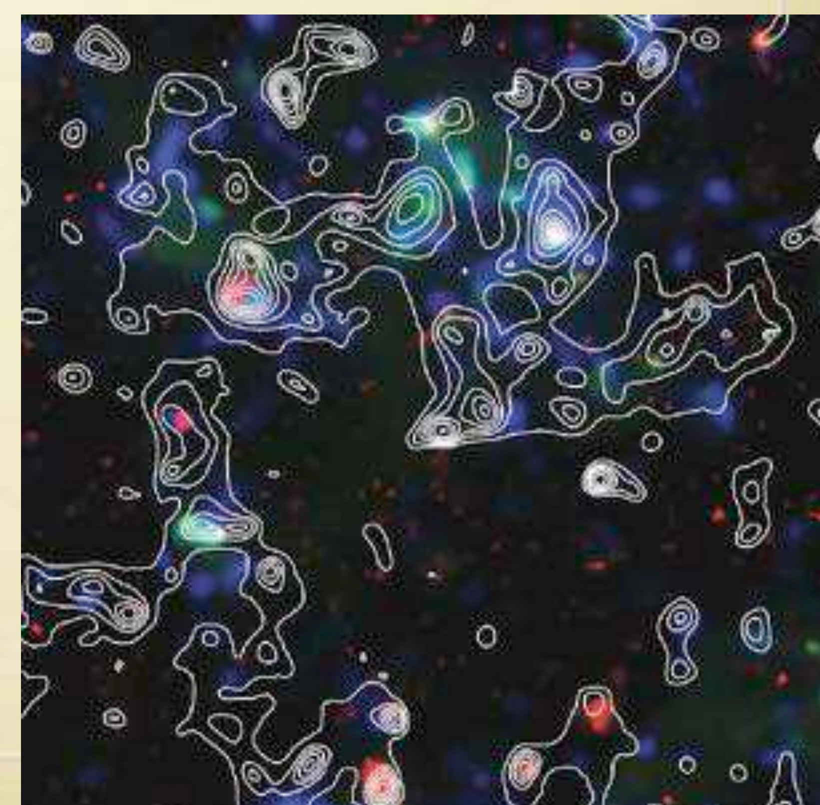
Cosmological random fields

- Cosmic Microwave Background :
    - Anisotropy of temperature & polarization pattern
- Large-scale Structure of the Universe :
    - Spatial distributions of galaxies, shapes of galaxies, weak lensing field, etc.



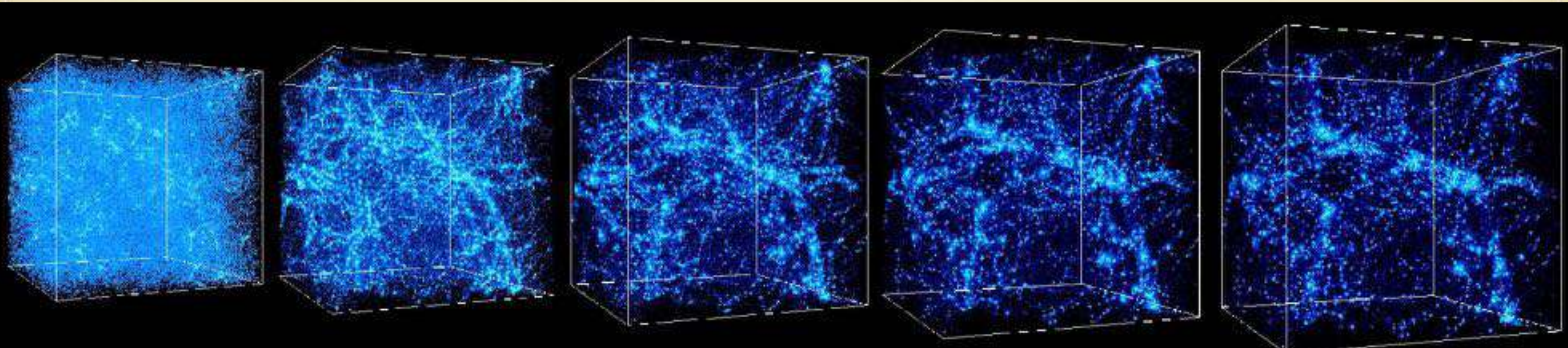
Large-scale structure of the Universe


Cosmic microwave background

Weak lensing field


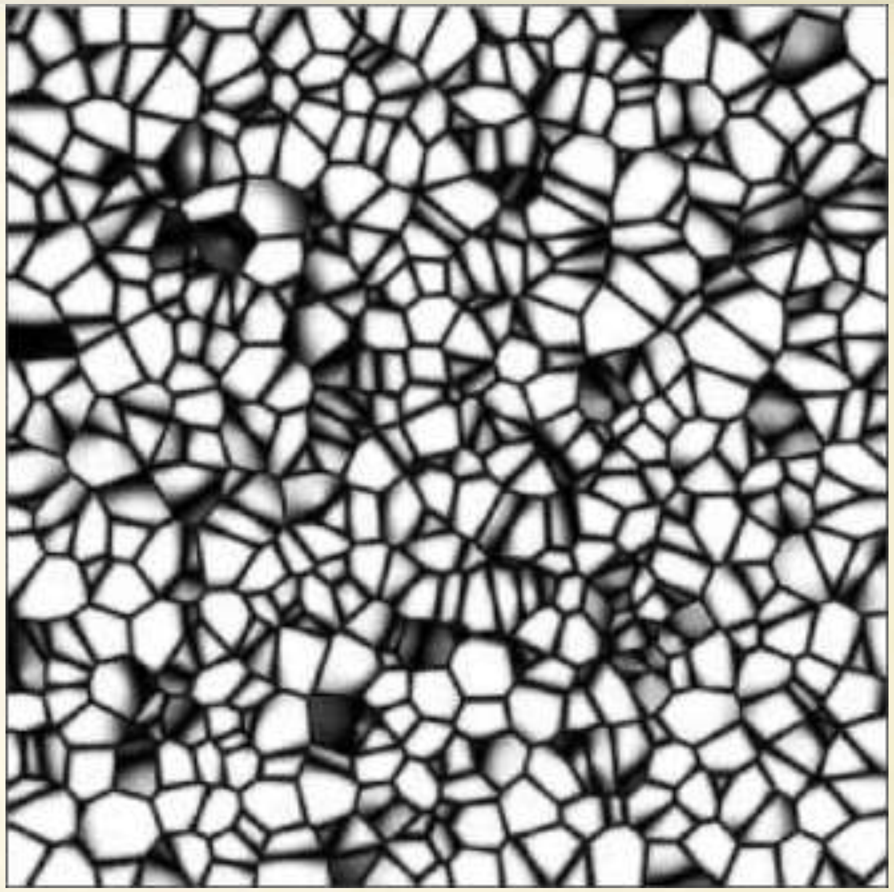
Random Gaussian field

- Initial density fluctuations of the Universe
  - Source of everything
  - Very small fluctuations
  - (It is believed that) Quantum fluctuations in the inflationary phase of the Universe
  - “Random Gaussian field” with very good accuracy



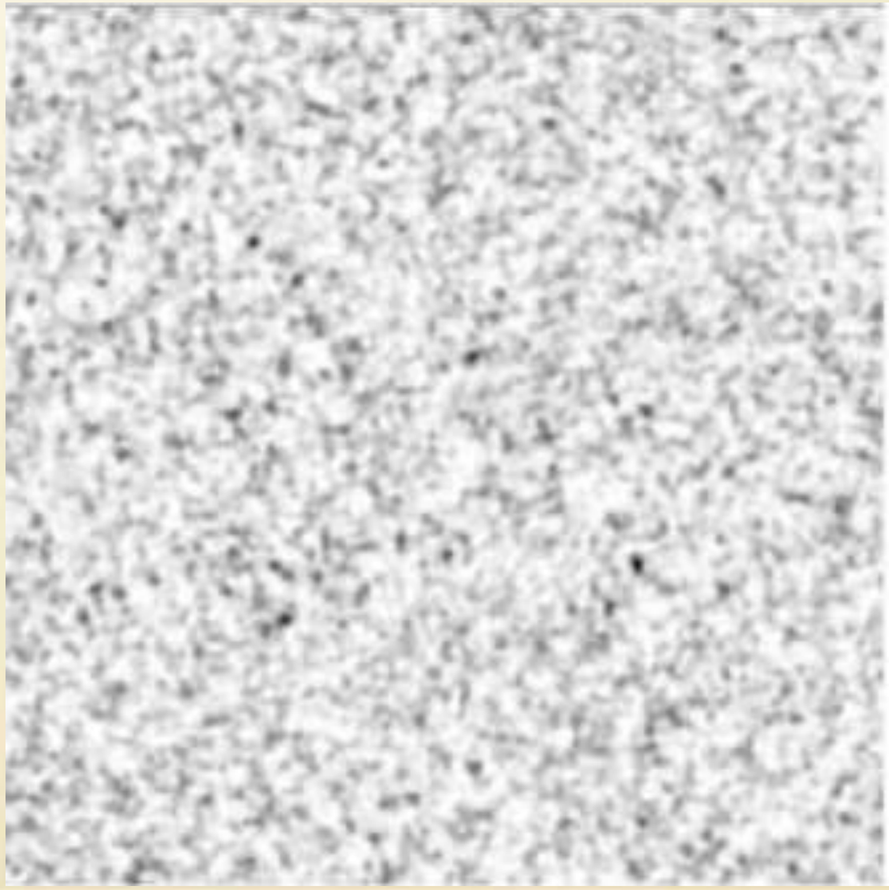
Visualizations by Andrey Kravtsov

Random Gaussian Field



Voronoi Tessellation

Lognormal mapping of RGF



Quartic mapping of RGF

Random Gaussian field

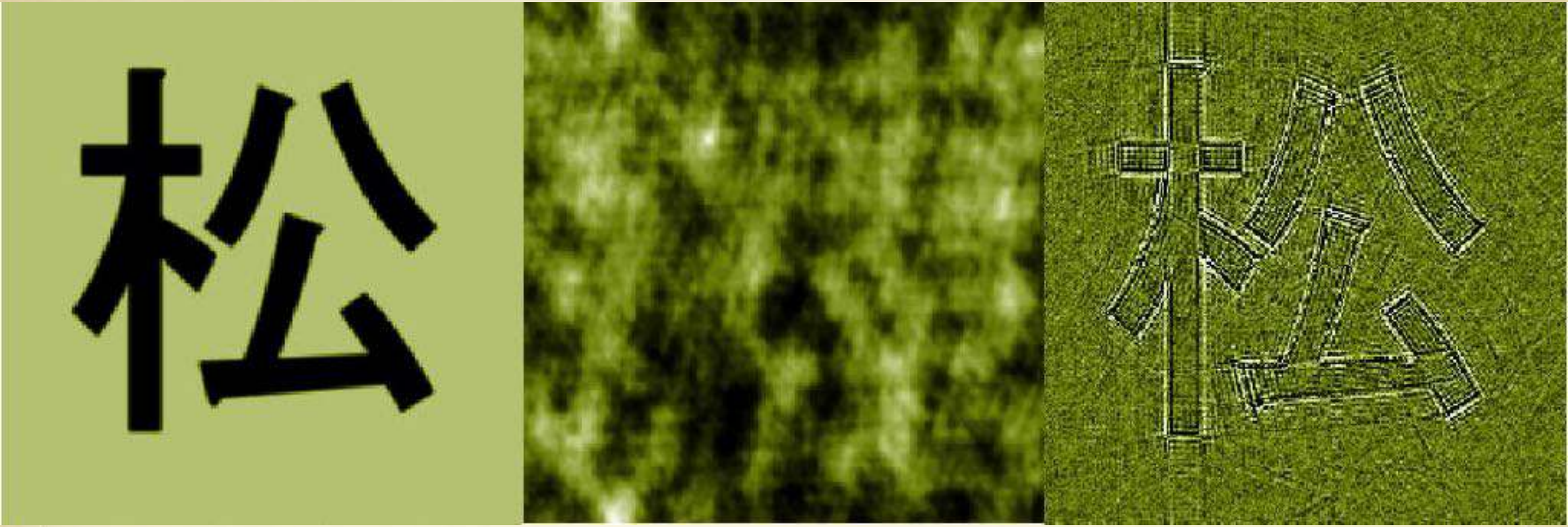
- All the statistical properties are characterized by the power spectrum

$$P(k) \propto \langle |\delta(\boldsymbol{k})|^2 \rangle$$

- Higher-order (n > 2) spectra (higher-order correlation functions) are all zero
- Central limit theorem
  - Sum of many independent variables with the same mean and variance tends toward a normal distribution (Gaussian distribution)
  -

Random Gaussian field

- Analysis of cosmological random field
  - The power spectrum (2-point statistics) is quite common
  - The power spectrum exhausts statistical information for the random Gaussian field
  - However, cosmological random fields are generally NOT random Gaussian fields
    - Nonlinear processes such as gravitational evolutions & structure formation
    - Non-Gaussianity in the initial fluctuations
- Necessities for analysis methods of random fields beyond the power spectrum



Identical power spectrum

Information of the power spectrum is erased (flat spectrum)

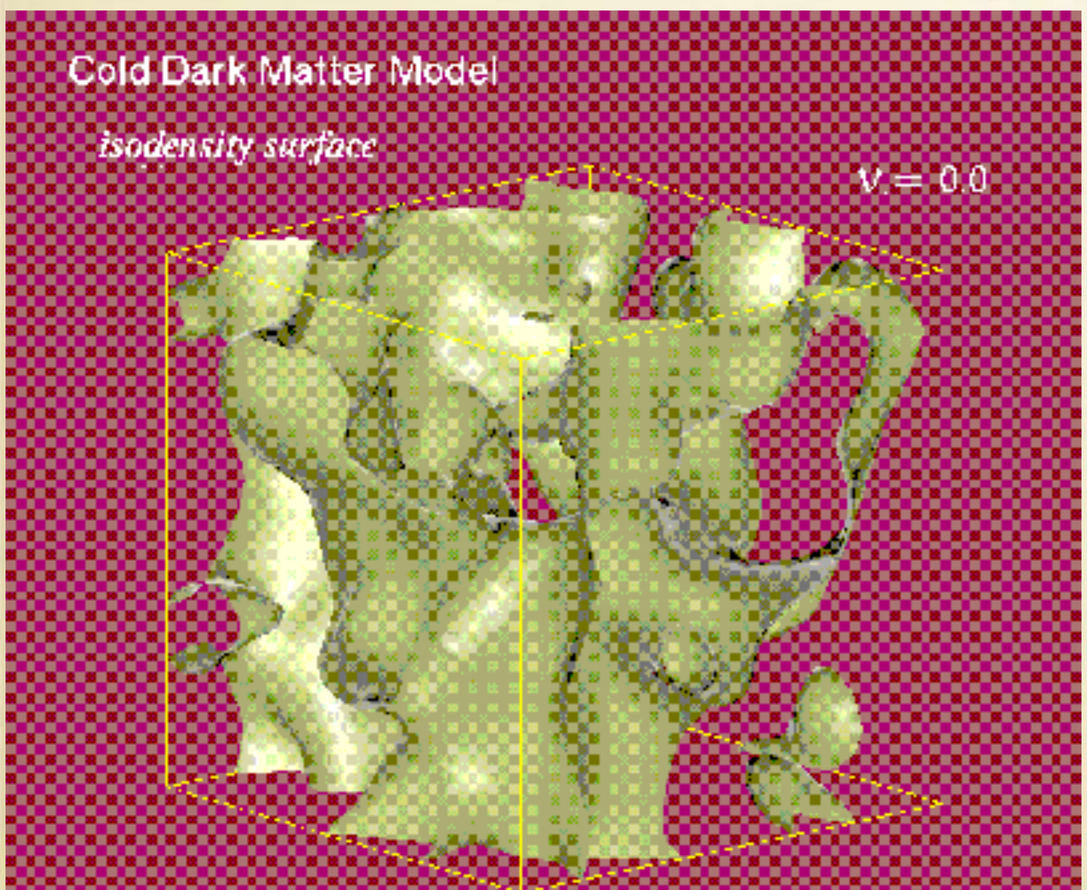
Statistical analysis of random fields

- Statistical methods beyond the power spectrum: Not unique
    - Higher-order extensions of the two-point statistics
        - Higher-order spectra, higher-order correlation functions
        - Higher-order moments (Skewness, Kurtosis,...)
    - Geometrical statistics
        - Genus statistic (Euler characteristic)
        - Minkowski functionals
        - Peak statistics, Statistics of critical points
        - Betti number, persistent homology (Topological data analysis)
    - Application of deep learning, etc.

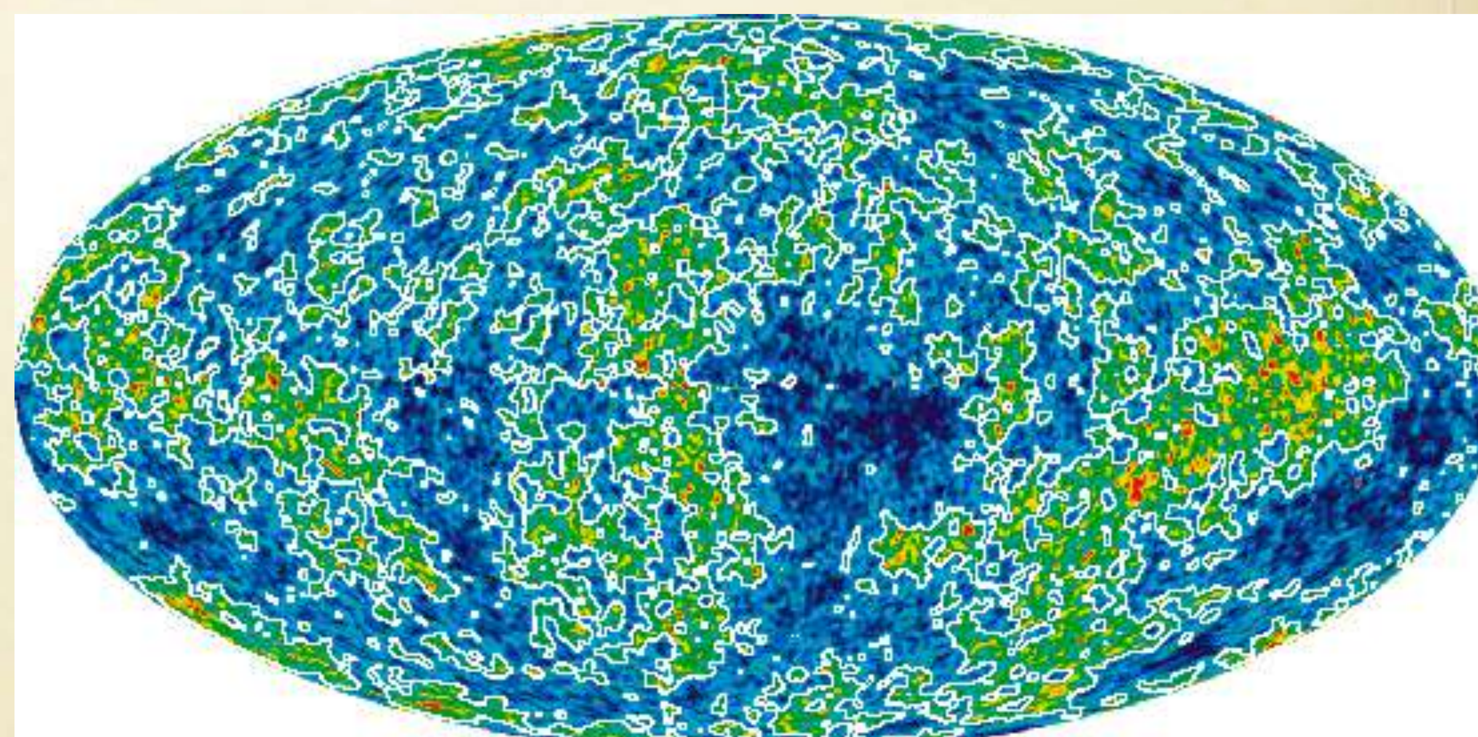
Minkowski functionals

What is the Minkowski functionals?

- Minkowski functionals:
    - d+1 functionals in d-dimensional space
    - calculated from isocontours of random fields

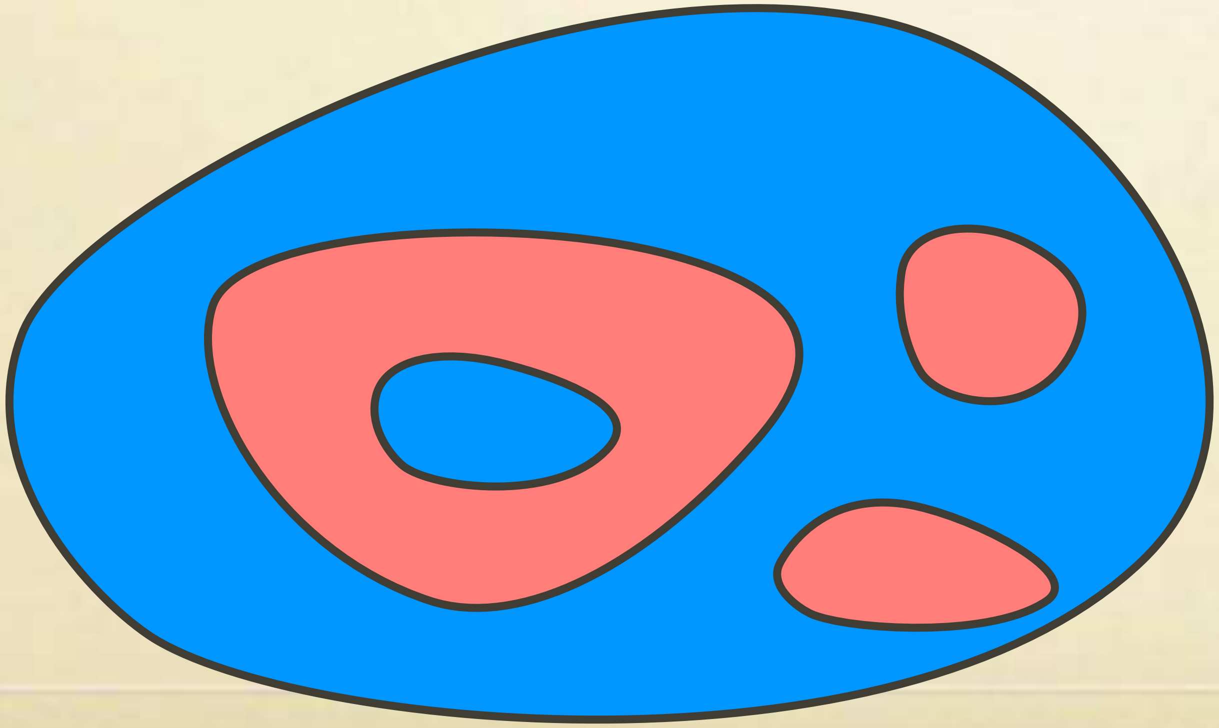


Isodensity surfaces in the large-scale structure


Isotemperature lines in CMB

Minkowski functionals in 2D

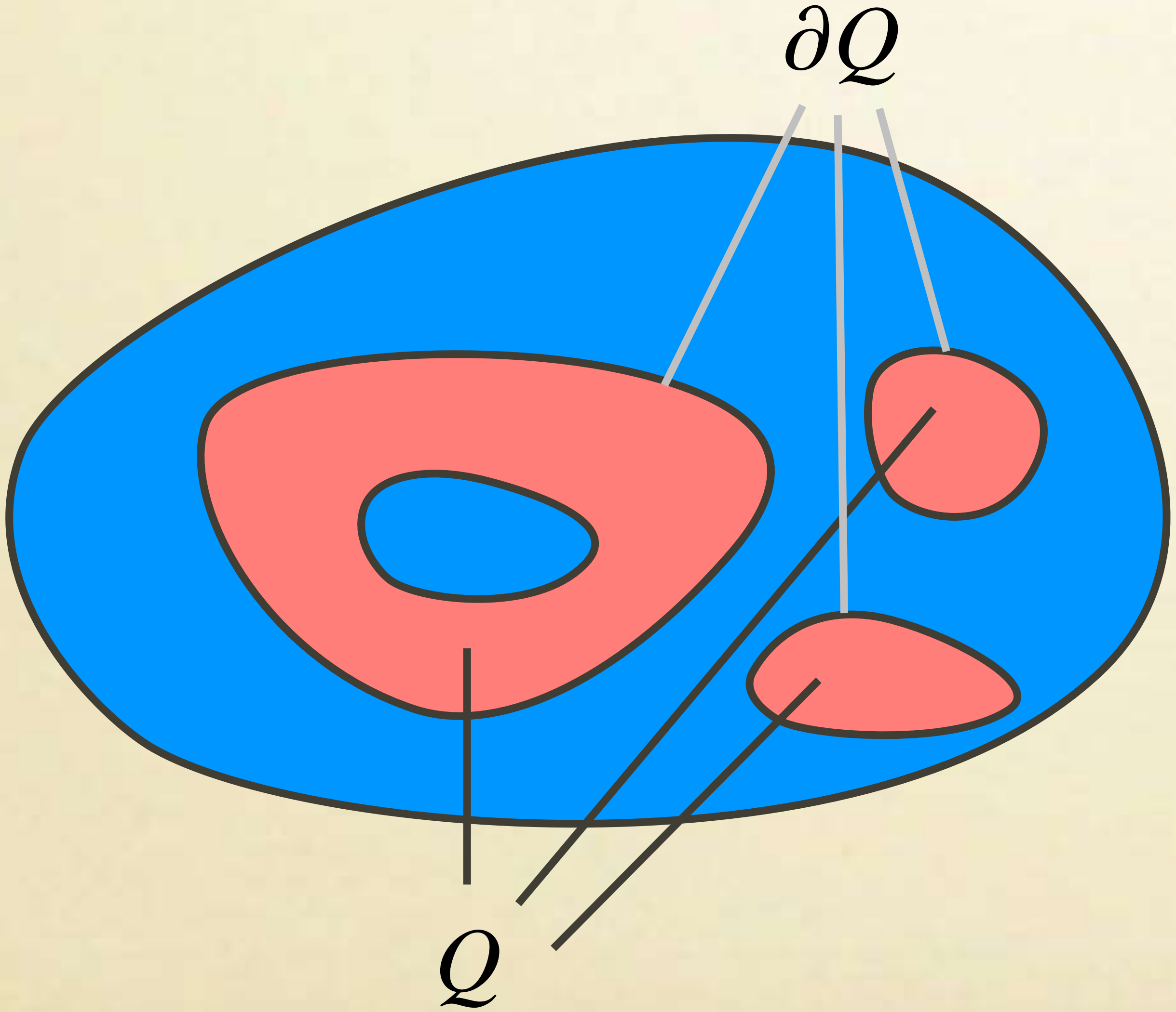
- Minkowski functionals in 2D $V_k^{(2)}, \quad (k = 0, 1, 2)$
  - d+1=3 functionals in d=2 dimensions
    - k=0: Area of thresholded regions $\propto V_0^{(2)}$
    - k=1: Length of isocontours $\propto V_1^{(2)}$
    - k=2: Euler number (#hot region - #cold region) $\propto V_2^{(2)}$



ex.) $V_2^{(2)} \propto 3 - 2 = +1$

Minkowski functionals in 2D

$$Q = \{(x,y) | f(x,y) \geq \nu\sigma\}, \quad \sigma = \langle f^2 \rangle^{1/2}$$



$$V_0(\nu) = \int_Q da$$

$$V_1(\nu) = \frac{1}{4}\int_{\partial Q} d\ell$$

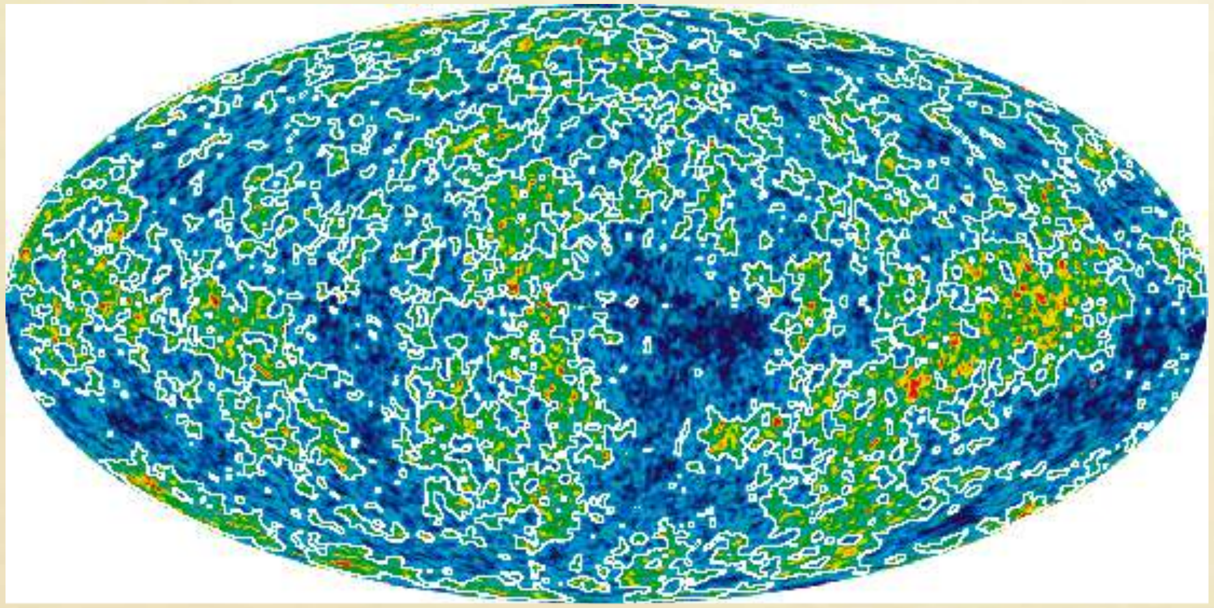
$$V_2(\nu) = \frac{1}{2\pi}\int_{\partial Q} \kappa d\ell$$

$\kappa = \dfrac{1}{R}$: geodesic curvature of ∂Q

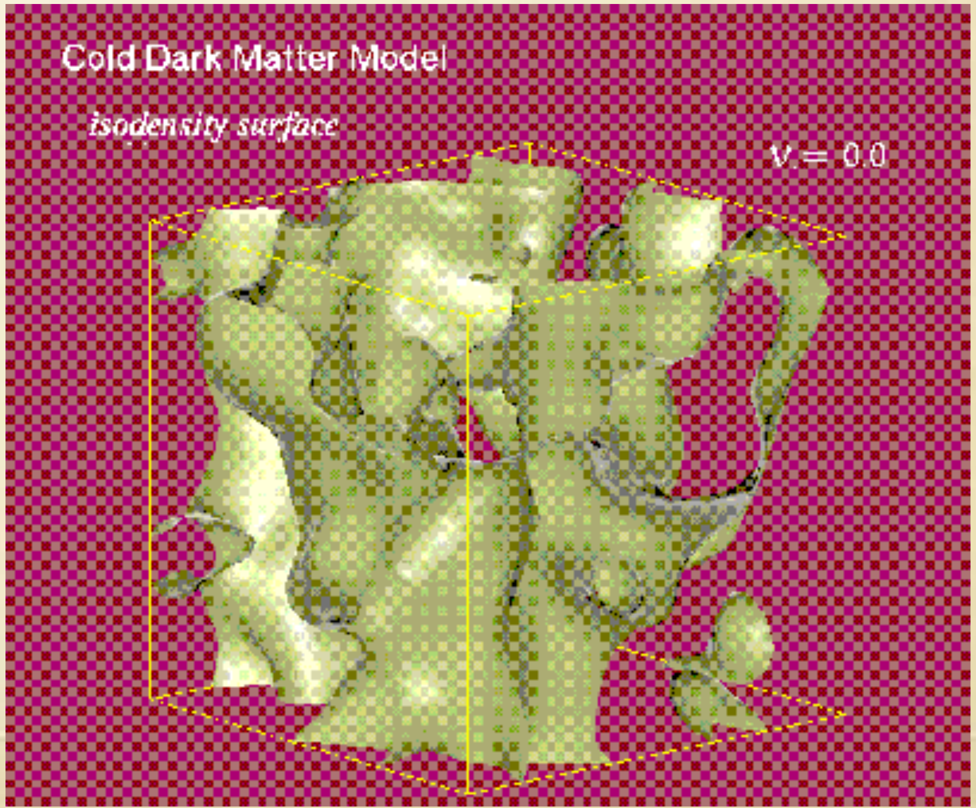
Minkowski functionals in 2D and 3D

d	V_0	V_1	V_2	V_3
2	Area	Total circumference	Euler characteristic	...
3	Volume	Surface area	Total mean curvature	Euler characteristic

$d = 2$



$d = 3$



Minkowski functionals in general dimensions $V_k^{(d)}, \quad (k = 1, 2, \dots, d)$

- Minkowski functionals in d dimensions

$$V_0^{(d)}(\nu) = \int_Q d^d x$$

$$V_k^{(d)}(\nu) = \frac{1}{\omega_k d} \int_{\partial Q} d^{d-1}x\, K_{k-1}^{(d)}(\nu, \boldsymbol{x})$$

$$K_{k-1}^{(d)}(\nu, \boldsymbol{x}) = \frac{1}{{}_{d-1}C_{k-1}} \sum_{\mathrm{perm}.\sigma} \frac{1}{R_{\sigma(1)} R_{\sigma(2)} \cdots R_{\sigma(k-1)}}$$

$R_1, R_2, \dots, R_{d-1}$: principal curvatures of ∂Q

$\omega_k = \dfrac{\pi^{k/2}}{\Gamma(k/2+1)}$: volume of unit sphere in k dimensions

Minkowski functionals as functions of threshold

$$Q = \{\boldsymbol{x} | f(\boldsymbol{x}) \geq \nu\sigma\} \;\Rightarrow\; V_k^{(d)}(\nu)$$

$$\nu \;:\; \text{threshold}$$

- WMAP.mp4

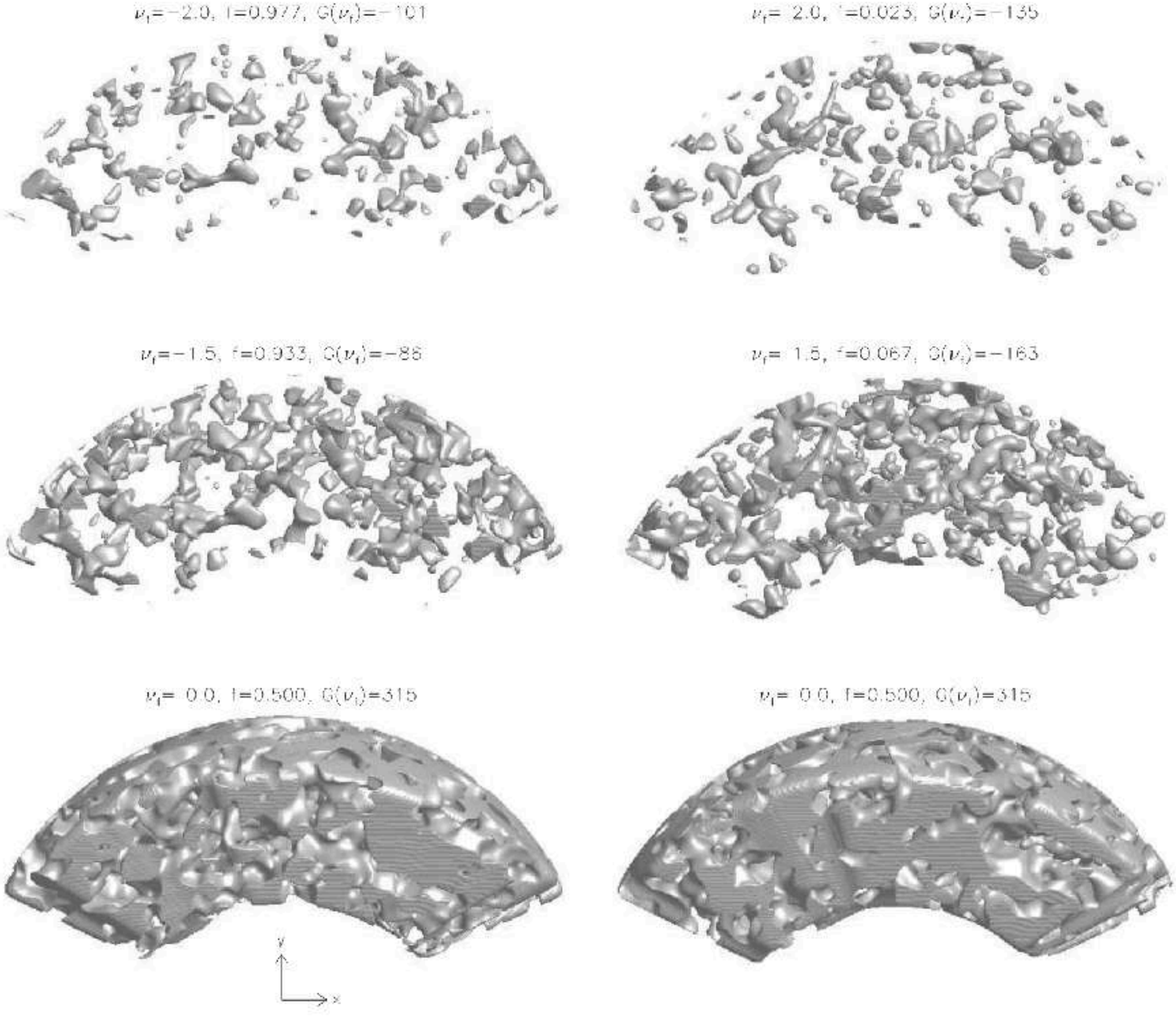
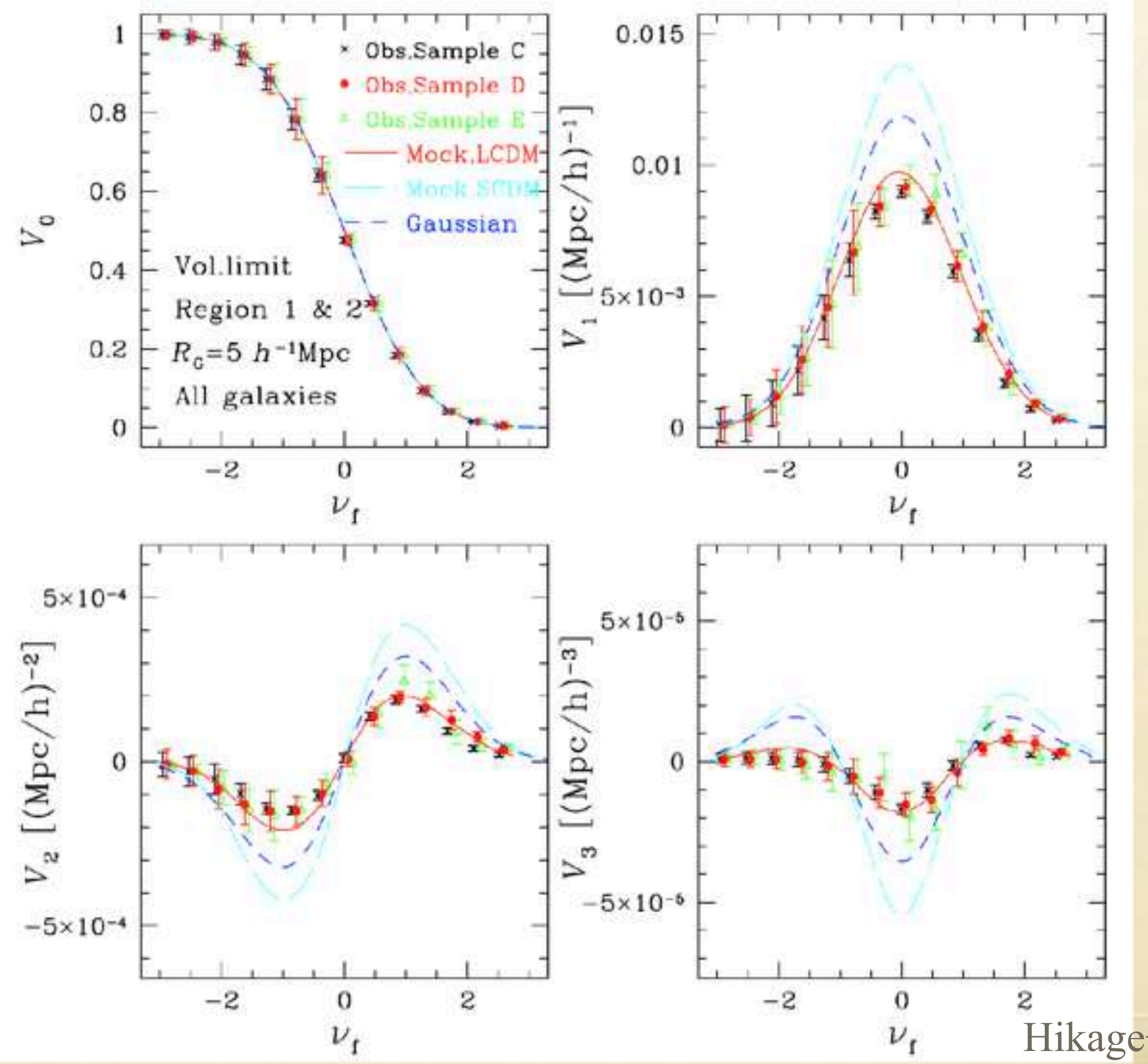


FIG. 7.— Three-dimensional view of the galaxy number density field of a mock sample smoothed with $R_G = 22\,h^{-1}$Mpc.

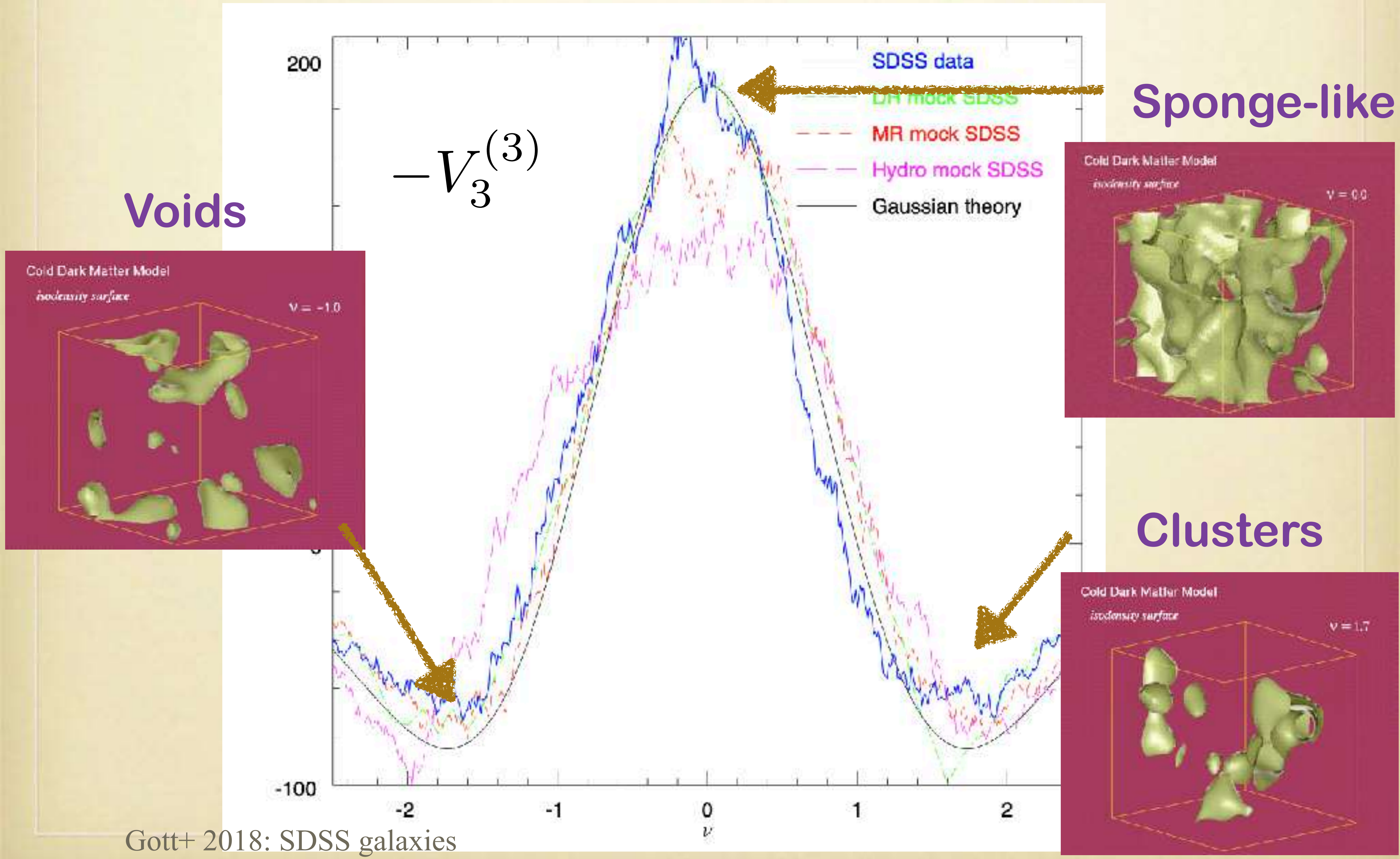
Minkowski functionals as functions of threshold

- Minkowski functionals in 3D (SDSS galaxies)



Hikage+ 2003: SDSS galaxies

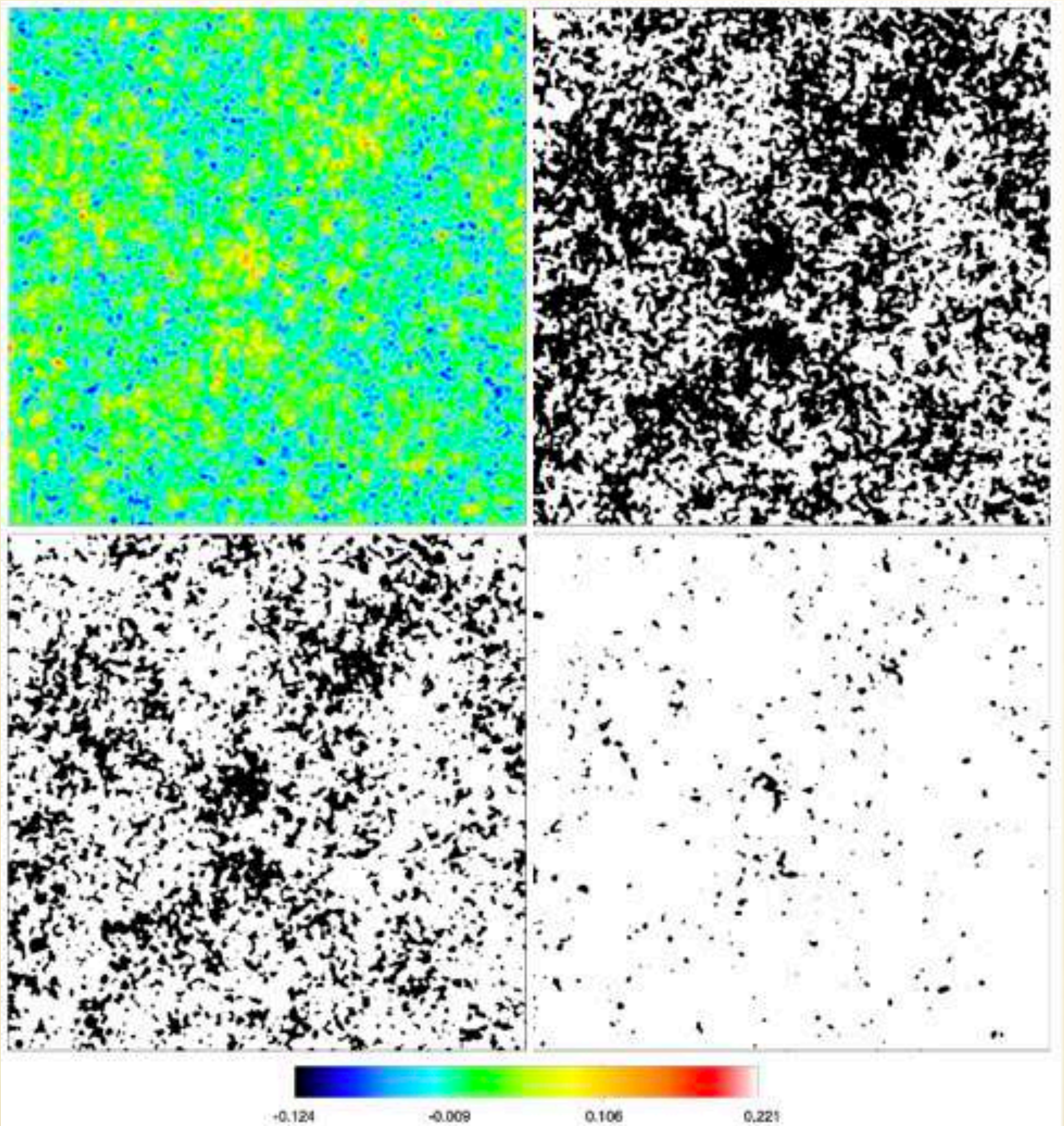
-V3: Topology of LSS (Genus curve)



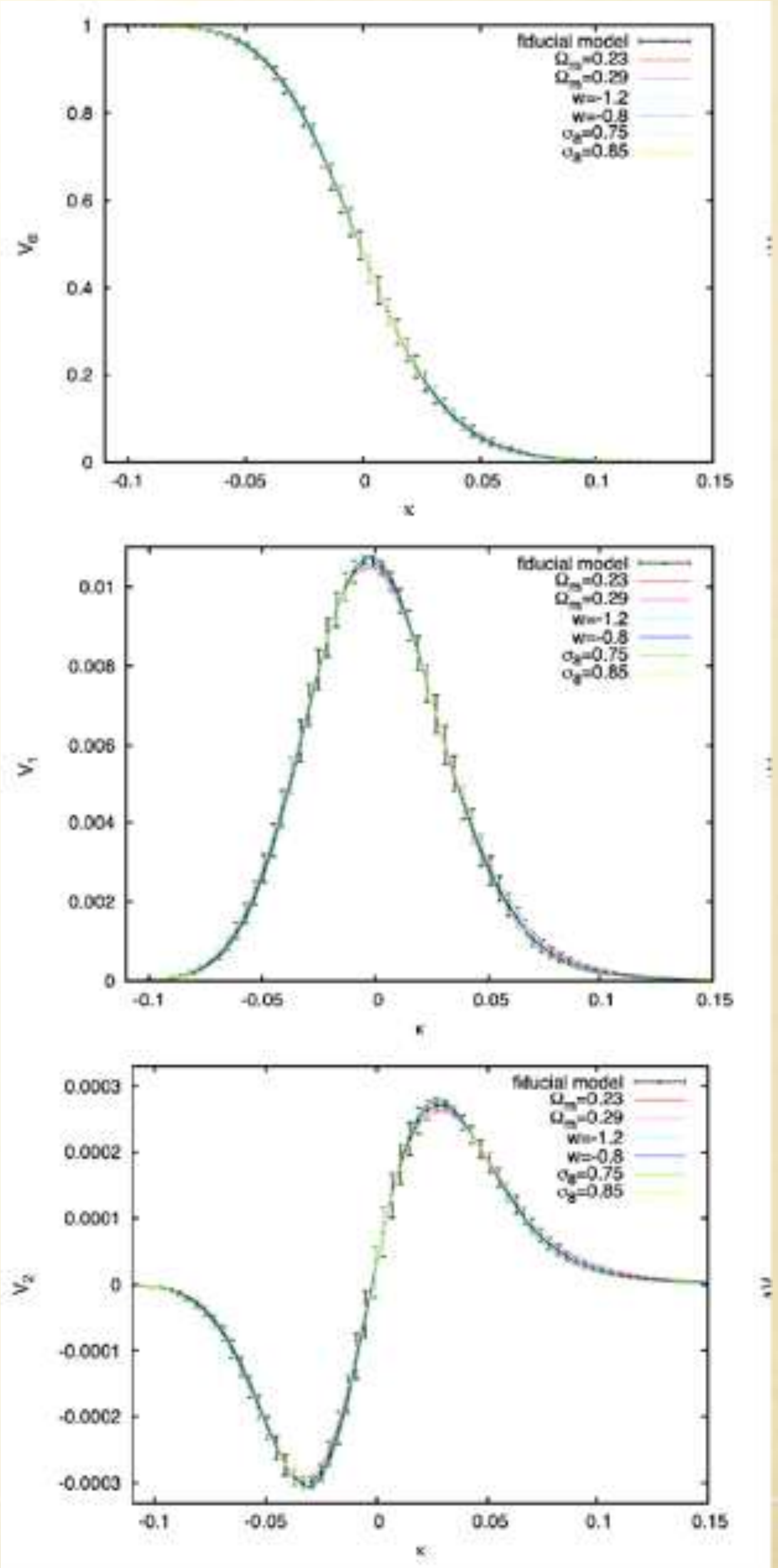
Gott+ 2018: SDSS galaxies

Application of MFs

- Weak gravitational lensing



Kratochvil+ 2011: Mock WL field



Minkowski functionals in Gaussian fields

- Analytic formula for Gaussian fields in general dimensions are known:
  - Tomita's formula (Tomita 1986)

$$V_k^{(d)}(\nu) = \frac{1}{(2\pi)^{(k+1)/2}} \frac{\omega_d}{\omega_{d-k}\omega_k} \left(\frac{\sigma_1}{\sqrt{d}\sigma} \right)^k e^{-\nu^2} H_{k-1}(\nu)$$

$$H_n(\nu) = e^{\nu^2/2} \left(-\frac{d}{d\nu} \right)^n e^{-\nu^2/2} \quad \text{(Hermite Polynomials)}$$

$$H_{-1}(\nu) \equiv e^{\nu^2/2} \int_\nu^\infty d\nu e^{-\nu^2/2} = \sqrt{\frac{\pi}{2}} \mathrm{erfc}\left(\frac{\nu}{\sqrt{2}} \right)$$

$$\sigma = \langle f^2 \rangle^{1/2}, \ \sigma_1 = \langle |\nabla f|^2 \rangle^{1/2}$$

$$\omega_k = \frac{\pi^{k/2}}{\Gamma(k/2+1)} \quad : \ \text{volume of unit sphere in } k \text{ dimensions}$$

Minkowski functionals in Gaussian fields

- MFs have universal shapes for Gaussian random fields as functions of threshold ν :

$$V_k^{(d)}(\nu) = \frac{1}{(2\pi)^{(k+1)/2}} \frac{\omega_d}{\omega_{d-k}\omega_k} \left(\frac{\sigma_1}{\sqrt{d}\sigma}\right)^k \underline{e^{-\nu^2} H_{k-1}(\nu)}$$

$V_1^{(3)}$ $V_2^{(3)}$ $V_3^{(3)}$

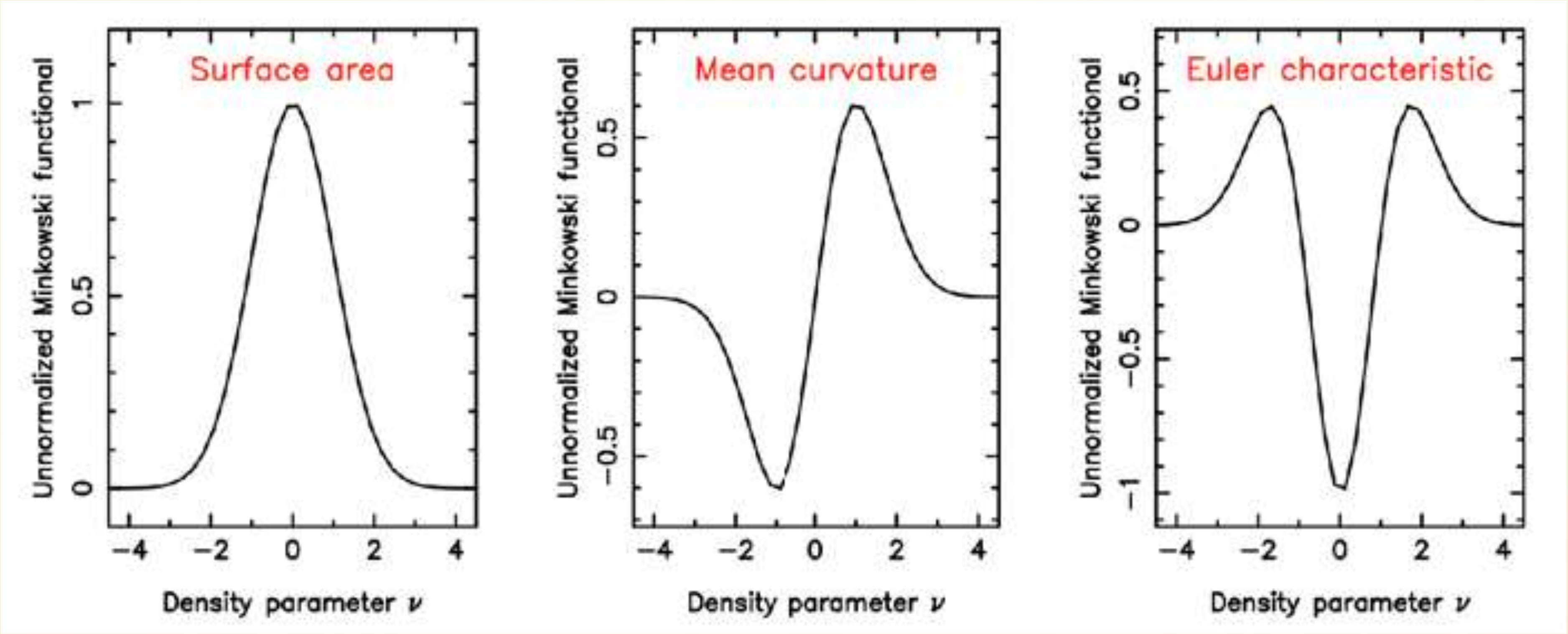


Fig:Blake+ 2018

Any deviation from these shapes: non-Gaussian signature!!

Analytic formula of MFs in weakly non-Gaussian fields

- Second-order formula: This work

$$V_k^{(d)}(\nu) = \frac{1}{(2\pi)^{(k+1)/2}}\frac{\omega_d}{\omega_{d-k}\omega_k}\left(\frac{\sigma_1}{\sqrt{d}\sigma_0}\right)^k e^{-\nu^2/2}\Bigg[\!\!\Bigg[H_{k-1}(\nu) + \left[\frac{1}{6}S^{(0)}H_{k+2}(\nu) + \frac{k}{3}S^{(1)}H_k(\nu) + \frac{k(k-1)}{6}S^{(2)}H_{k-2}(\nu)\right]\sigma_0$$
$$+ \left\{\frac{1}{72}(S^{(0)})^2 H_{k+5}(\nu) + \left(\frac{1}{24}K^{(0)} + \frac{k}{18}S^{(0)}S^{(1)}\right)H_{k+3}(\nu) + k\left[\frac{1}{8}K^{(1)} + \frac{k-1}{36}S^{(0)}S^{(2)} + \frac{k-2}{18}(S^{(1)})^2\right]H_{k+1}(\nu)\right.$$
$$+ k\left[\frac{k-2}{16}K_1^{(2)} + \frac{k}{16}K_2^{(2)} + \frac{(k-1)(k-4)}{18}S^{(1)}S^{(2)}\right]H_{k-1}(\nu)$$
$$\left. + k(k-1)(k-2)\left[\frac{1}{24}K^{(3)} + \frac{k-7}{72}(S^{(2)})^2\right]H_{k-3}(\nu)\right\}\sigma_0^2 + O\left(\sigma_0^3\right)\Bigg]\!\!\Bigg]. \quad ($$

$$S^{(0)} = \frac{\langle f^3\rangle_c}{\sigma_0^4}, \quad S^{(1)} = \frac{3}{2}\frac{\langle f|\nabla f|^2\rangle_c}{\sigma_0^2\sigma_1^2},$$
$$S_1^{(2)} = \frac{-3d}{2(d-1)}\frac{\langle|\nabla f|^2\triangle f\rangle_c}{\sigma_1^4},$$

$$K^{(0)} = \frac{\langle f^4\rangle_c}{\sigma_0^6}, \quad K^{(1)} = 2\frac{\langle f^2|\nabla f|^2\rangle_c}{\sigma_0^4\sigma_1^2},$$
$$K_1^{(2)} = \frac{-2d}{(d+2)(d-1)}\frac{(d+2)\langle f|\nabla f|^2\triangle f\rangle_c + \langle|\nabla f|^4\rangle_c}{\sigma_0^2\sigma_1^4},$$
$$K_2^{(2)} = \frac{-2d}{(d+2)(d-1)}\frac{(d+2)\langle f|\nabla f|^2\triangle f\rangle_c + d\langle|\nabla f|^4\rangle_c}{\sigma_0^2\sigma_1^4},$$
$$K^{(3)} = \frac{2d^2}{(d-1)(d-2)}\frac{\langle|\nabla f|^2(\triangle f)^2\rangle_c - \langle|\nabla f|^2 f_{ij}f_{ij}\rangle_c}{\sigma_1^6}. \quad (7$$

- This formula is valid in general dimensions
- Consistent with all the previous formulas
- Derivation is much smarter than previous methods
  - useful mathematical identities, etc. For details, see the paper to come out soon!

Analytic formula of MFs in weakly non-Gaussian fields

- Some fragments of the proof (see the forthcoming paper)
  - Crofton's formula

$$V_k^{(d)}(\nu) = \frac{\omega_d}{\omega_{d-k}\omega_k} \int_{\mathcal{E}_k^{(d)}} d\mu_k(E)\chi^{(k)}(\mathcal{F}_\nu \cap E),$$

  - Density of Euler characteristic

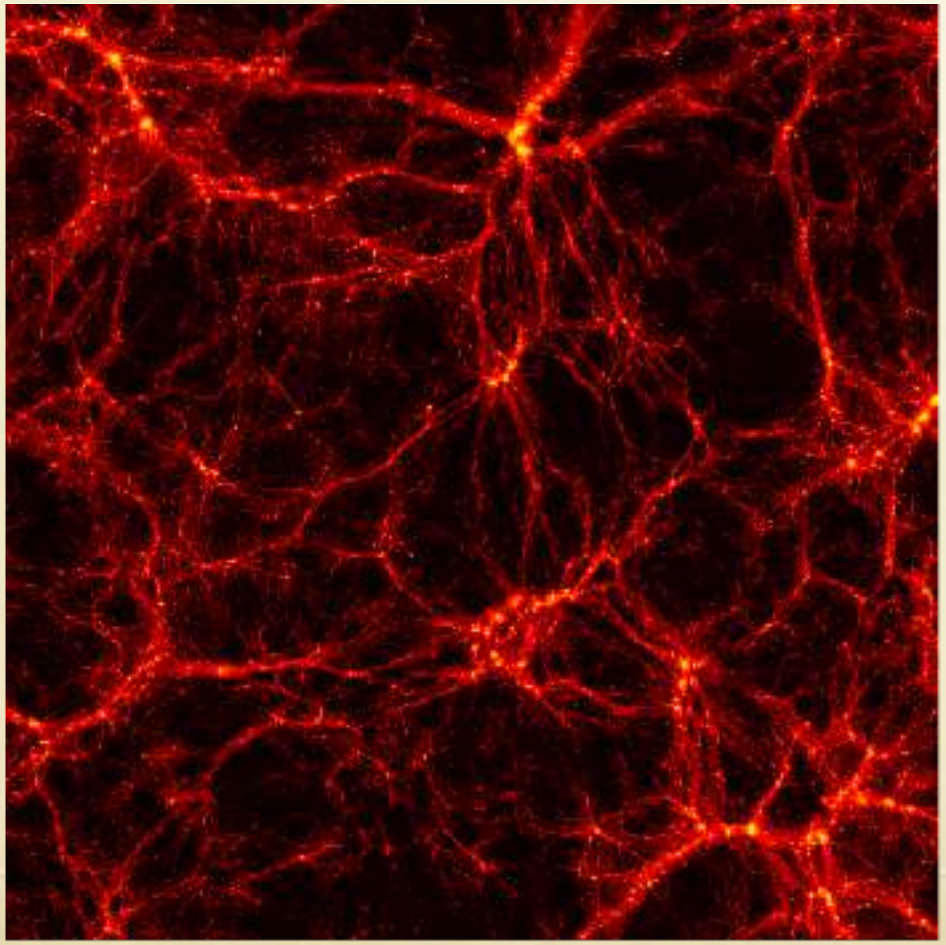
$$n_\chi^{(d)}(\nu) = (-1)^d \left(\frac{\sigma_2}{\sigma_1}\right)^d \left\langle \Theta(\alpha - \nu)\delta^d(\boldsymbol{\eta}) \det \zeta \right\rangle,$$

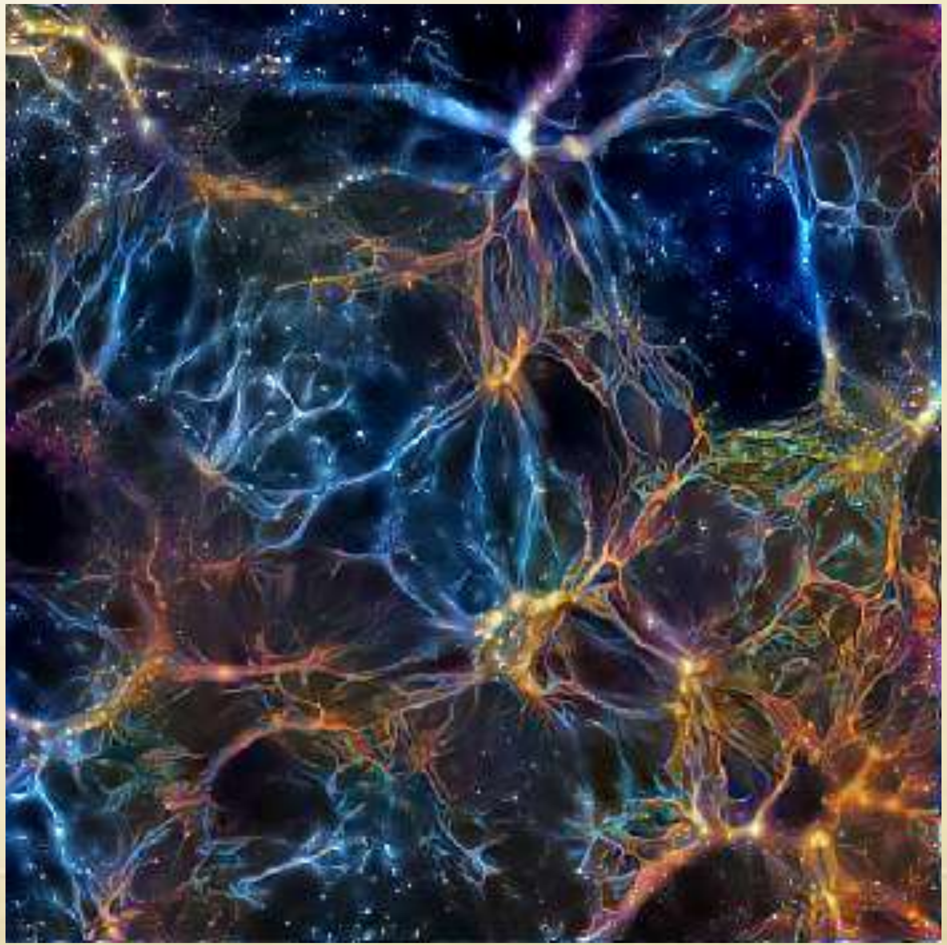
  - Multivariate Edgeworth expansion

$$\begin{aligned}\langle F \rangle = \langle F \rangle_{\mathrm{G}} &+ \frac{1}{6} C^{(3)}_{\mu_1\mu_2\mu_3} \left\langle \frac{\partial^3 F}{\partial X_{\mu_1} \partial X_{\mu_2} \partial X_{\mu_3}} \right\rangle_{\mathrm{G}} \sigma_0 \\ &+ \left[\frac{1}{24} C^{(4)}_{\mu_1\mu_2\mu_3\mu_4} \left\langle \frac{\partial^4 F}{\partial X_{\mu_1} \partial X_{\mu_2} \partial X_{\mu_3} \partial X_{\mu_4}} \right\rangle_{\mathrm{G}} \right. \\ &\left. + \frac{1}{72} C^{(3)}_{\mu_1\mu_2\mu_3} C^{(3)}_{\mu_4\mu_5\mu_6} \left\langle \frac{\partial^6 F}{\partial X_{\mu_1} \partial X_{\mu_2} \partial X_{\mu_3} \partial X_{\mu_4} \partial X_{\mu_5} \partial X_{\mu_6}} \right\rangle_{\mathrm{G}} \right] {\sigma_0}^2 \\ &+ \mathcal{O}\left({\sigma_0}^3\right). \qquad (1\end{aligned}$$

Comparisons with N-body simulations

- We calculate the Minkowski functionals in the numerical N-body simulations of the large-scale structure
- Quijote simulations:
  - publicly available N-body suites (Vilaescusa-Navarro et al. 2020)
  - 512^3 particles in $1(h^{-1}\text{Gpc})^3$ boxes × 15,000 realizations, etc.
- Theoretical predictions are calculated from the nonlinear perturbation theory of gravitational evolution



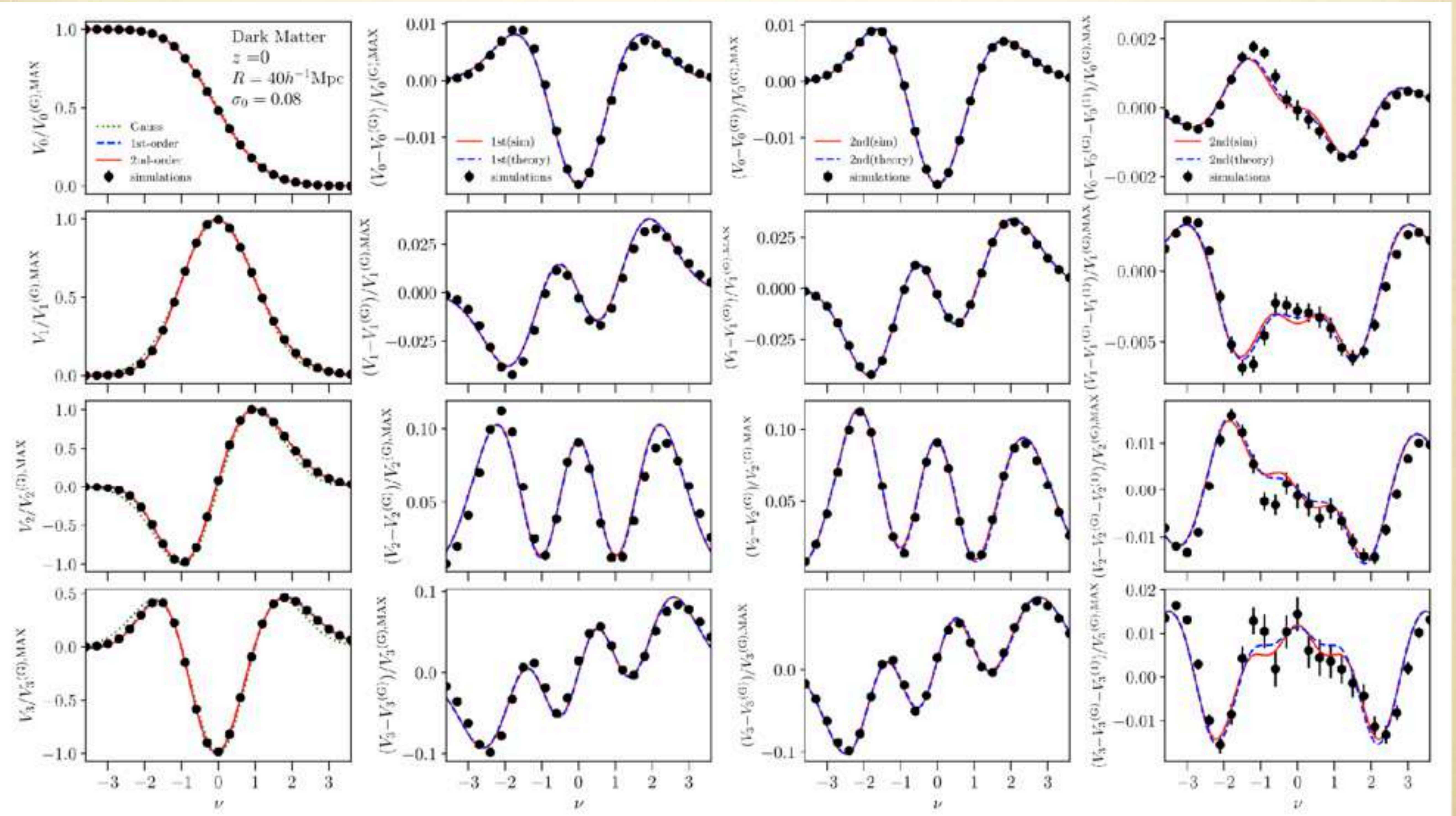


Comparisons with numerical N-body simulations

R [h^{-1}Mpc]	10	20	30	40
σ_0	0.385 0.3804 ± 0.0001	0.193 0.1899 ± 0.0001	0.121 0.1194 ± 0.0001	0.0845 0.08374 ± 0.0001
σ_1	0.0367 0.03652 ± 0.00001	0.0101 0.009918 ± 0.000004	0.00441 0.004352 ± 0.000003	0.00240 0.002371 ± 0.000002
$S^{(0)}$	3.56 3.762 ± 0.004	3.40 3.46 ± 0.01	3.33 3.36 ± 0.02	3.28 3.28 ± 0.05
$S^{(1)}$	3.63 3.932 ± 0.003	3.45 3.531 ± 0.006	3.36 3.41 ± 0.01	3.31 3.34 ± 0.03
$S^{(2)}$	3.66 4.499 ± 0.004	3.68 3.887 ± 0.006	3.71 3.81 ± 0.01	3.72 3.78 ± 0.03
$K^{(0)}$	23.2 26.66 ± 0.09	20.9 21.6 ± 0.2	19.9 20.2 ± 0.5	19.2 19 ± 1
$K^{(1)}$	23.8 28.77 ± 0.09	21.3 22.2 ± 0.1	20.2 20.6 ± 0.3	19.5 19.8 ± 0.8
$K_1^{(2)}$	30.6 41.7 ± 0.1	28.3 30.4 ± 0.2	27.2 28.1 ± 0.4	26.7 27.5 ± 0.8
$K_2^{(2)}$	18.8 26.7 ± 0.1	17.7 19.2 ± 0.1	17.3 18.0 ± 0.2	17.0 17.8 ± 0.6
$K^{(3)}$	25.1 45.5 ± 0.2	25.6 29.6 ± 0.2	26.1 27.6 ± 0.4	26.4 27 ± 1

Comparisons with numerical N-body simulations

- Minkowski functionals, R = 40 h^{-1}Mpc



Summary

- We derived an analytic formula of MFs in general dimensions with second-order non-Gaussianity
- The formula in 3D is applied to the large-scale structure of the Universe
- Results of perturbation theory agree well with numerical simulations for large smoothing radius
- Future:
    - Biasing & Redshift-space distortions
    - Constraining primordial non-Gaussianity and models of early universe by MFs?

Statistics of peaks

celebrated "BBKS" paper

Journal, **304**:15–61, 1986 May 1

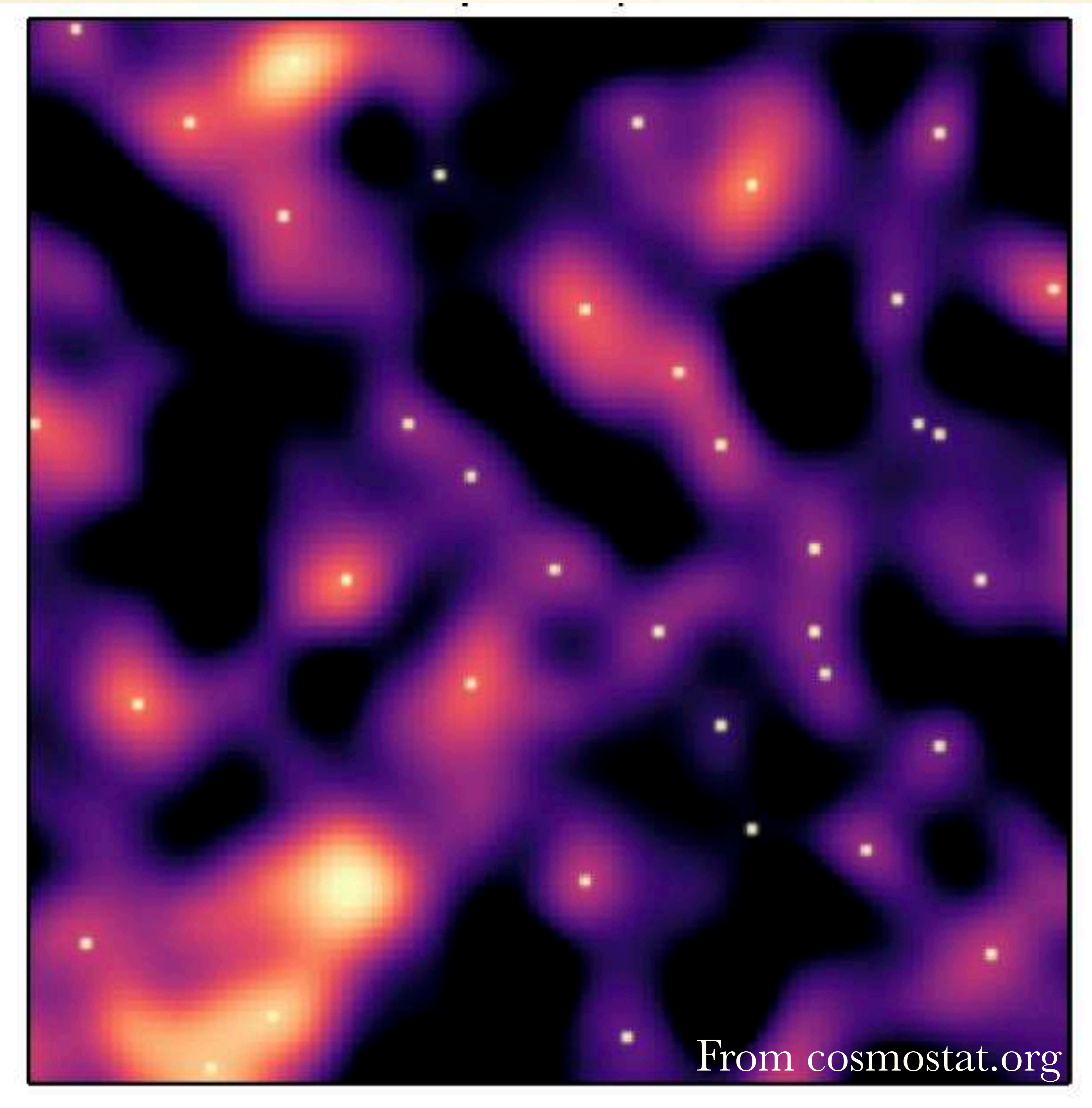
THE STATISTICS OF PEAKS OF GAUSSIAN RANDOM FIELDS

J. M. Bardeen[1]
Physics Department, University of Washington
J. R. Bond[1]
Physics Department, Stanford University
N. Kaiser[1]
Astronomy Department, University of California at Berkeley, and Institute of Astronomy, Cambridge University
AND
A. S. Szalay[1]
Astrophysics Group, Fermilab
Received 1985 July 25; accepted 1985 October 9

ABSTRACT

nological density fluctuations are often assumed to be Gaussian random fields. The local ma
elds are obvious sites for the formation of nonlinear structures. The statistical properties of th
used to predict the abundances and clustering properties of objects of various types. In this pa
(1) the number density of peaks of various heights $\nu\sigma_0$ above the rms σ_0; (2) the factor by wh
ensity is enhanced in large-scale overdense regions; (3) the n-point peak-peak correlation functio

Peak abundance



Peak abundance

- Derived formula for the peak abundance in weakly non-Gaussian fields in generally N dimensions

$$
\begin{aligned}
\bar{n}_{\rm pk}(\nu) &\equiv \left\langle n_{\rm pk}(\nu) \right\rangle \\
&= G_{00000} + \frac{\sigma_0}{6}\Big[G_{30000}S^{(0)} + 4\gamma G_{21000}S^{(1)} + 3\gamma^2 G_{12000}S_2^{(2)} \\
&\quad + \gamma^3 G_{03000}S_1^{(3)} + 4G_{10100}S^{(1)} + \frac{4(N-1)}{N}\gamma G_{01100}S^{(2)} \\
&\quad + 6\gamma^2 G_{10010}\left(S_2^{(2)} - S^{(2)}\right) + \frac{6}{N-1}\gamma^3 G_{01010}\left(NS_2^{(3)} - S_1^{(3)}\right) \\
&\quad + \frac{3(N-2)(N+2)^2}{2(N+4)}\gamma^3 G_{00001}\left(\frac{N+2}{3}S_1^{(3)} - NS_2^{(3)}\right)\Big], \qquad (25)
\end{aligned}
$$

where

$$
S^{(0)} \equiv \frac{\left\langle f^3 \right\rangle_{\rm c}}{{\sigma_0}^4}, \quad S^{(1)} \equiv -\frac{3}{4}\frac{\left\langle f^2 \triangle f \right\rangle_{\rm c}}{{\sigma_0}^2{\sigma_1}^2},
$$

$$
S^{(2)} \equiv -\frac{3N}{2(N-1)}\frac{\langle(\nabla f\cdot\nabla f)\triangle f\rangle_{\rm c}}{{\sigma_1}^4}, \quad S_2^{(2)} \equiv \frac{\left\langle f(\triangle f)^2 \right\rangle_{\rm c}}{{\sigma_1}^4},
$$

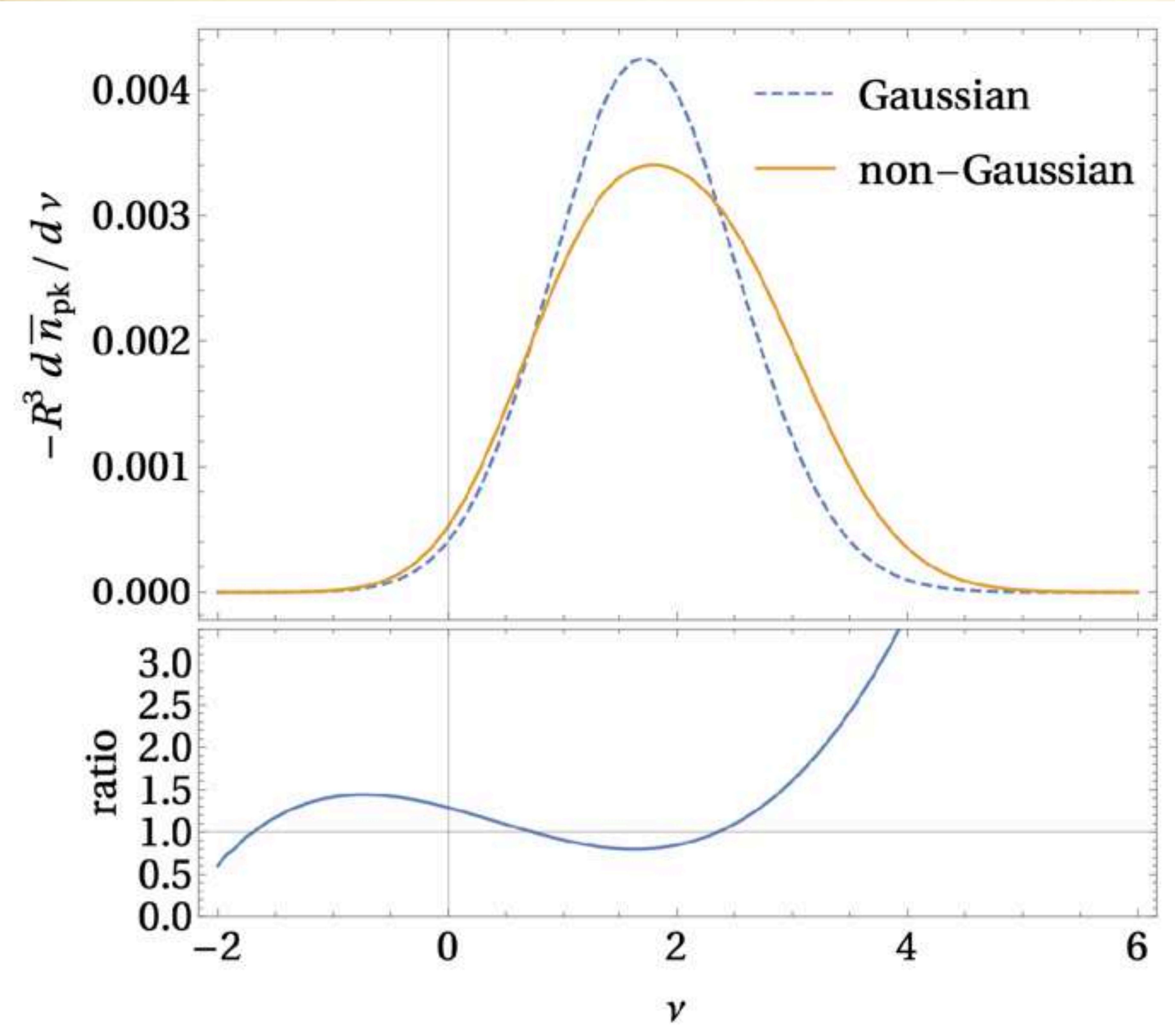
$$
S_1^{(3)} \equiv -\frac{{\sigma_0}^2}{{\sigma_1}^6}\left\langle (\triangle f)^3 \right\rangle_{\rm c}, \quad S_2^{(3)} \equiv -\frac{{\sigma_0}^2}{{\sigma_1}^6}\left\langle f_{ij}f_{ij}\triangle f \right\rangle_{\rm c}. \qquad (26)
$$

$$
\begin{aligned}
G_{ijklm}(\nu) &= \frac{1}{(2\pi)^{N/2}}\left(\frac{\sigma_2}{\sqrt{N}\sigma_1}\right)^N X_k \\
&\quad \times \int_0^\infty dx\, H_{i-1,j}(\nu,x)\, N(\nu,x)\, f_{lm}(x), \qquad (35)
\end{aligned}
$$

where

$$
\begin{aligned}
f_{lm}(x) &\equiv N^N \int d^{N-1}W\, \Theta(\lambda_N)\, \lambda_1\cdots\lambda_N\, F_{lm}(J_2,J_3) \\
&\quad \times \exp\left[-\frac{(N-1)(N+2)}{4}J_2\right]. \qquad (36)
\end{aligned}
$$

Example 1: Abundance of peaks in 3D density field

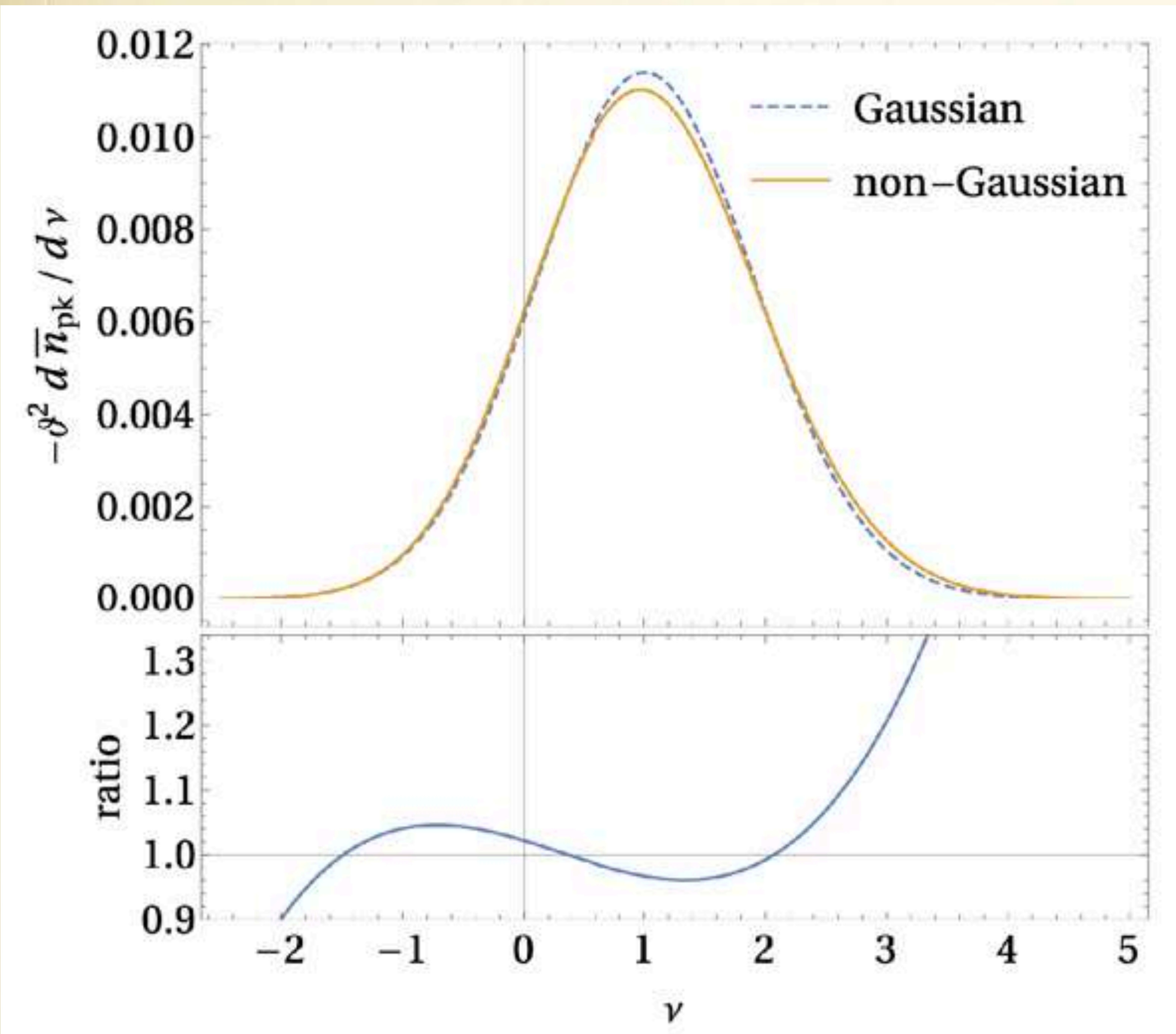


$$\delta = \frac{\rho - \bar{\rho}}{\bar{\rho}} = \nu\sigma$$

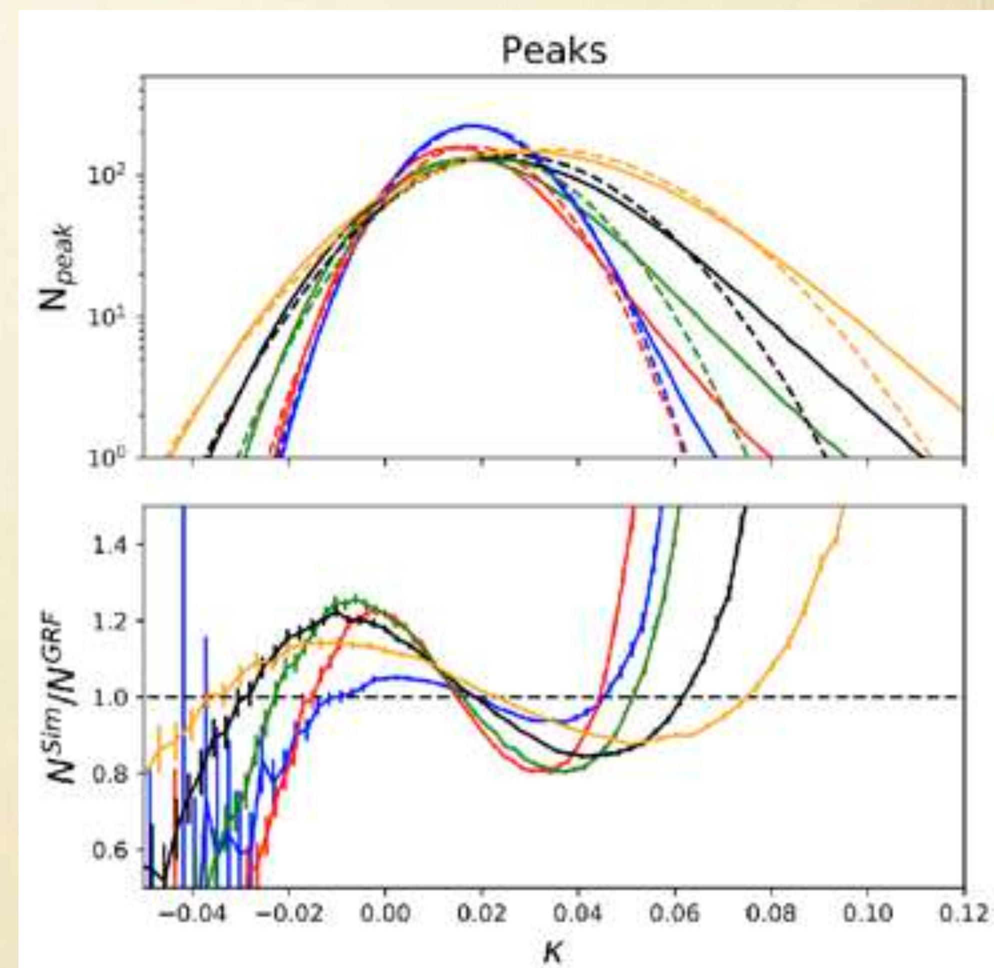
Example 2: Abundance of peaks in 2D weak lensing field

convergence $\kappa = \nu\sigma$

Theoretical prediction:



Weak lensing simulations:



Coulton+ 2019

Summary

- Statistics of peaks are important clues of cosmological structure formation
- Gaussian formulas of BBKS for the peak statistics are extended to include effects of weak non-Gaussianity
- Some examples are presented in the case of gravitational non-Gaussianity