Minkowski functionals for isotropic random fields in the Euclidean space and the sphere

Kuriki, Satoshi Inst. Statist. Math. (ISM) and SOKENDAI

Statistical Analysis of Random Fields in Cosmology KEK-Cosmo 2024, Mon 4 March 2024

Joint work with T. Matsubara and C. Hikage

 I. MFs for isotropic random fields: The Euclidean space case Matsubara & K (2021), Matsubara, Hikage & K (2022), K and Matsubara (2023)

II. MFs for isotropic random fields: The sphere case

I. MFs for isotropic random fields: The Euclidean space case Matsubara & K (2021), Matsubara, Hikage & K (2022), K and Matsubara (2023)

II. MFs for isotropic random fields: The sphere case

Excursion set for a smooth isotropic random field

► Isotropic random field X(t), $t \in T \subset \mathbb{R}^n$: $\forall T'(\text{finite set}) \subset T$,

$$\left\{X(t)\right\}_{t\in T'} \stackrel{d}{=} \left\{X(Pt+b)\right\}_{t\in T'}, \; \forall (P,b) \in O(n) \times \mathbb{R}^n$$

• Excursion set is the sup-level set of a function X(t):

$$T_v = \{t \in T \mid X(t) \ge v\} = X^{-1}([v, \infty))$$



Left: Isotropic random field, Right: Its excursion set

Minkowski functional (MF) and Lipschitz-Killing curvature

• Let $M \subset \mathbb{R}^n$ be a closed set. Tube about M with radius ρ :

 $\operatorname{Tube}(M,\rho) = \left\{ x \in \mathbb{R}^n \mid \operatorname{dist}(x,M) \le \rho \right\}$



Definitions of Minkowski functional V_j(·) and Lipschitz-Killing curvature L_j(·):
 For small ρ > 0, and ω_d = π^{d/2}/Γ(d/2 + 1),

$$\operatorname{Vol}_{n}(\operatorname{Tube}(M,\rho)) = \sum_{j=0}^{n} \omega_{n-j} \rho^{n-j} \mathcal{L}_{j}(M) = \sum_{j=0}^{n} \rho^{j} \binom{n}{j} \mathcal{V}_{j}(M)$$

 $(\mathcal{L}_j(\cdot) \text{ is defined independently of the ambient space})$

When the dimension n is 2

▶ When
$$n = 2$$
,

$$\mathcal{L}_{2}(M) = \mathcal{V}_{0}(M) = \int_{M} \mathrm{d}x = \operatorname{Area}(M)$$
$$\mathcal{L}_{1}(M) = \mathcal{V}_{1}(M) = \frac{1}{2} \int_{\partial M} \mathrm{d}x = \frac{1}{2} \operatorname{Length}(\partial M)$$
$$\mathcal{L}_{0}(M) = \frac{1}{\pi} \mathcal{V}_{2}(M) = \chi(M) \quad (\text{Euler characteristics})$$

For general n and i, L_i(M) is represented as an integral of the curvature measure over ∂M.

MF of the excursion set T_v as a goodness-of-fit statistic

▶ The Minkowski functional $\mathcal{V}_i(T_v)$ of the excursion set T_v can be used as a statistic for testing goodness-of-fit (e.g., applications in cosmology)





non-Gaussian

- observed Euler characteristic $\chi(T_v)$
- $\mathbb{E}[\chi(T_v)]$ under the assumption of Gaussianity
- Our purpose: To obtain the formulas $\mathbb{E}[\chi(T_v)]$ (and $\mathbb{E}[\mathcal{L}_k(T_v)]$) under non-Gaussian distributions

Kinematic Formula for isotropic random fields

Fact (Kinematic Formula in \mathbb{R}^n) For $A, B \subset \mathbb{R}^n$ and for $gB = \{Px + b \mid x \in B\}$ (g: rigid motion),

$$\int \chi(A \cap gB) \mathrm{d}g \propto \sum_{d=0}^{n} \mathcal{L}_{d}(A) \, \mathcal{L}_{n-d}(B).$$

▶ Let A := T, $gB := \{t \in \mathbb{R}^n \mid X(t) \ge x\}$, then $A \cap gB = T_x$. Proposition (Kinematic Formula for isotropic random fields) When X(t), $t \in T$, is a smooth isotropic random field,

$$\mathbb{E}[\chi(\mathbf{T}_{x})] = \mathbb{E}[\mathcal{L}_{0}(T_{x})] = \sum_{d=0}^{n} \mathcal{L}_{d}(\mathbf{T}) \Xi_{d}(x),$$

where $\Xi_d(x)$ is the Euler characteristic density.

• Expected MF except for $k \ge 1$:

$$\mathbb{E}[\mathcal{L}_k(T_x)] = \sum_{d=0}^{n-k} \frac{\Gamma(\frac{k+d+1}{2})\Gamma(\frac{1}{2})}{\Gamma(\frac{k+1}{2})\Gamma(\frac{d+1}{2})} \mathcal{L}_{k+d}(T) \Xi_d(x), \quad k = 1, \dots, n$$

• We only need to know the Euler characteristic density $\Xi_d(x)$.

EC density for the Gaussian random field

Proposition (Tomita, 1986, *PTP*, and many) Suppose that $Var(\nabla X(t)) = \gamma I$. When X(t) is Gaussian,

$$\Xi_d(x) = (\gamma/2\pi)^{d/2} \phi(x) H_{d-1}(x)$$

where $\phi(x)$: pdf of $\mathcal{N}(0,1)$, $H_j(x)$: Hermite polynomial Therefore.

$$\mathbb{E}[\chi(T_x)] = \sum_{d=0}^n \mathcal{L}_d(T) \, (\gamma/2\pi)^{d/2} \phi(x) H_{d-1}(x)$$

▶ When *X*(*t*) is not Gaussian?

Weakly non-Gaussian case

The non-Gaussianity is characterized by k-point correlation functions (k-th cumulant). In the applications to cosmology,

$$\operatorname{cum}(X(t_1), \dots, X(t_k)) = O(\nu^{k-2}), \ \nu \ll 1$$

e.g., Central limit random fields by Chamandy, et al. (2008):

$$X(t) = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} Z_{(i)}(t), \quad \nu = 1/\sqrt{N}.$$

Z_(i): zero mean i.i.d. isotropic non-Gaussian random field.
The k-point correlation function for the isotropic random field:

$$\operatorname{cum}(X(t_1),\ldots,X(t_k)) = \nu^{k-2}\kappa^{(k)}(\frac{1}{2}||t_1-t_2||^2,\frac{1}{2}||t_1-t_3||^2,\ldots,\frac{1}{2}||t_{k-1}-t_k||^2)$$

(A function of pairwise distances between t_1, \ldots, t_k)

Theorem

Under regularity conditions, as $\nu \rightarrow 0$ uniformly in x,

$$\Xi_{d,\nu}(x) = \left(\frac{\gamma}{2\pi}\right)^{d/2} \phi(x) \left(H_{d-1}(x) + \nu \Delta_{1,d}(x) + \nu^2 \Delta_{2,d}(x)\right) + o(\nu^{-2})$$

where

$$\Delta_{1,d}(x) = \frac{1}{2\gamma^2} \kappa_{11}(d)_2 H_{d-2}(x) - \frac{1}{2\gamma} \kappa_1 dH_d(x) + \frac{1}{6} \kappa_0 H_{d+2}(x)$$

$$\kappa_0 = \kappa^{(3)}(0,0,0), \ \kappa_1 = \frac{d\kappa^{(3)}(x,0,0)}{dx}|_{x=0}, \ \kappa_{11} = \frac{d^2 \kappa^{(3)}(x,y,0)}{dxdy}|_{x=y=0}$$

Asymptotic expansion of the Euler density II

and

$$\begin{split} \Delta_{2,d}(x) &= \left(-\frac{1}{6\gamma^3} (3\tilde{\kappa}_{111}^a + \tilde{\kappa}_{111}^d) + \frac{1}{8\gamma^4} (d-7)\kappa_{11}^2 \right) (d)_3 H_{d-3}(x) \\ &+ \left(\frac{1}{8\gamma^2} (\tilde{\kappa}_{11}^{aa} (d-2) + 4\tilde{\kappa}_{11}^a (d-1) \right) \\ &- \frac{1}{4\gamma^3} \kappa_1 \kappa_{11} (d-1) (d-4) \right) dH_{d-1}(x) \\ &+ \left(-\frac{1}{4\gamma} \tilde{\kappa}_1 + \frac{1}{24\gamma^2} (3\kappa_1^2 (d-2) + 2\kappa_0 \kappa_{11} (d-1)) \right) dH_{d+1}(x) \\ &+ \left(\frac{1}{24} \tilde{\kappa}_0 - \frac{1}{12\gamma} \kappa_0 \kappa_1 d \right) H_{d+3}(x) + \frac{1}{72} \kappa_0^2 H_{d+5}(x) \end{split}$$

$$\widetilde{\kappa}_0 = \kappa^{(4)}(0, 0, 0, 0, 0, 0), \ \widetilde{\kappa}_1 = \frac{\mathrm{d}\kappa^{(4)}(x_1, 0, 0, 0, 0, 0)}{\mathrm{d}x_1}|_{x_1=0}, \ \dots$$

Sketch of the proof

Kac-Rice formula (Morse theory) for the EC density:

 $\Xi_d(v) = \mathbb{E}[\chi(T_v)] = \mathbb{E}[\mathbb{1}(X(t) \ge v) \det(-\nabla^2 X(t))\delta(\nabla X(t))]$ (*)

 $\Xi_d(v)$ is independent of t (because X is isotropic).

Obtain the Edgeworth expansion of the pdf (or mgf) of

 $(X(t), \nabla X(t), \nabla^2 X(t)) \in \mathbb{R}^{1+d+d(d+1)/2}$

and evaluate $\Xi_d(v)$ by (*).

(Non)Detectable non-Gaussianity

Diagram of the derivatives:

$$\frac{\partial}{\partial x_{12}} \frac{\partial}{\partial x_{13}} \frac{\partial}{\partial x_{14}} \frac{\partial}{\partial x_{45}} \frac{\partial}{\partial x_{46}} \kappa^{(6)}(x_{12}, \dots, x_{56}) \Big|_{x=0}^{2} \frac{\partial}{\partial x_{10}} \frac{\partial}{\partial x_{13}} \frac{\partial}{\partial x_{14}} \frac{\partial}{\partial x_{23}} \frac{\partial}{\partial x_{45}} \frac{\partial}{\partial x_{46}} \kappa^{(6)}(x_{12}, \dots, x_{56}) \Big|_{x=0}^{2} \frac{\partial}{\partial x_{10}} \frac{\partial}{\partial x_{10}}$$

Theorem

The derivatives with loops do not appear in the formula for $\Xi_d(v)$.

In particular, the second derivatives don't appear

Chi-square random field

A weakly non-Gaussian random field when d.f. is large:

$$Y(t) = Y_N(t) = \frac{1}{\sqrt{2N}} \sum_{i=1}^N (Z_{(i)}(t)^2 - 1), \quad t \in T \subset \mathbb{R}^4,$$

 $Z_{(i)}(t)$: zero mean Gaussian s.t. $\mathbb{E}[Z_{(i)}(s)Z_{(i)}(t)]=e^{-\frac{1}{4}\|s-t\|^2}$, $i=1,2,\ldots$ i.i.d.

• The EC density for Y(t) is known:



EC density, N = 100 (--- true, ··· 0th approx, --1st approx, --2nd approx)



Difference from the true, N = 100(...0th approx, --1st approx, --2nd approx)

Comparison with simulator (Matsubara, Hikage & K, 2022)



Simulator (dot) and expansion formulas (solid line) R: Radius of smoothing kernel (h^{-1} Mpc) periodical boundary condition ($\partial T = \emptyset$)

Boundary correction

• When the dimension is n = 3, and $\partial T \neq \emptyset$,

 $\mathbb{E}[\chi(T_v)] = \operatorname{Vol}(T)\Xi_3(v) + \underbrace{\mathcal{L}_2(T)\Xi_2(v) + \mathcal{L}_1(T)\Xi_1(v) + \chi(T)\Xi_0(v)}_{\mathcal{L}_1(v)}$

boundary correction (contribution of ∂T)

• We use the same parameters for R = 20. T is a cubic.



Orange : $\mathbb{E}[\chi(T_x)]$ with boundary corrections Blue : without boundary corrections

I. MFs for isotropic random fields: The Euclidean space case Matsubara & K (2021), Matsubara, Hikage & K (2022), K and Matsubara (2023)

II. MFs for isotropic random fields: The sphere case

Orthogonally invariant random field on the sphere

- In the second half of the talk, we deal with a random field X(·) on the n-dim sphere Sⁿ_R with radius R.
- Instead of the isotropy, assume that X(t) is orthogonally invariant: ∀T'(finite set) ⊂ T,

$$\left\{X(t)\right\}_{t\in T'} \stackrel{d}{=} \left\{X(Pt)\right\}_{t\in T'}, \; \forall P \in O(n)$$

• Excursion set is the sup-level set of a function X(t):

$$T_v = \{t \in T \mid X(t) \ge v\} = X^{-1}([v, \infty))$$

• We consider the MFs of T_v . How does R affect?

Covariance and *k*-point correlation functions

▶ Distance: For
$$s, t \in \mathbb{S}_R^n$$
,

 $\operatorname{dist}_R(s,t) = \mathsf{the great circle distance between } s \mathsf{ and } t$



$$\begin{split} \operatorname{Cov}(X(s),X(t)) &= \mathsf{A} \text{ function of } \operatorname{dist}_R(s,t) \\ &= \rho \big(\tfrac{1}{2} \operatorname{dist}_R(s,t)^2 \big), \quad \mathsf{say}, \end{split}$$



$$\operatorname{cum}(X(s), X(t), X(u)) = \kappa^{(3)} \left(\frac{1}{2} \operatorname{dist}_R(s, t)^2, \frac{1}{2} \operatorname{dist}_R(s, u)^2, \frac{1}{2} \operatorname{dist}_R(t, u)^2\right)$$

Great circle distance



$$\frac{1}{2}\|s-t\| = R\sin\theta = R\sin(\frac{1}{2R}\mathrm{dist}(s,t))$$
 or
$$\frac{1}{2}\mathrm{dist}(s,t)^2 = x + \frac{1}{6R^2}x^2 + \cdots$$
 where

$$x = \frac{1}{2} \|s - t\|^2$$

Extended random field $\widetilde{X}(\cdot)$

• Construct an isotropic random field $\widetilde{X}(\cdot)$ on \mathbb{R}^{n+1} such that its restriction on \mathbb{S}^n_R satisfies

 $(\widetilde{X}|_{\mathbb{S}^n_R},\nabla\widetilde{X}|_{\mathbb{S}^n_R},\nabla^2\widetilde{X}|_{\mathbb{S}^n_R})(t) \stackrel{d}{=} (X,\nabla X,\nabla^2 X)(t), \quad t\in \mathbb{S}^n_R$

where ∇ is the gradient on the tangent space $T_t \mathbb{S}_R^n$. Noting that

$$\operatorname{cum}(X(t_1), X(t_2), \ldots) = \kappa^{(k)} \left(\frac{1}{2} \operatorname{dist}_R(t_1, t_2)^2, \ldots \right), \quad t_i \in \mathbb{S}_R^n$$

we define

$$\operatorname{cum}(\widetilde{X}(t_1), \widetilde{X}(t_2), \ldots) = \kappa^{(k)} \left(x_{12} + \frac{1}{6R^2} x_{12}^2 + \cdots, \ldots \right)$$

where

$$x_{ij} = \frac{1}{2} ||t_i - t_j||^2, \quad t_i \in \mathbb{R}^{n+1}$$

k-point correlations and there derivatives

- We apply the theorem for the Euclidean space to X with the index set T ⊂ Sⁿ_B ⊂ ℝⁿ⁺¹.
- We evaluate the parameters γ , κ_0 , κ_1, \ldots , $\widetilde{\kappa}_0, \ldots$ that construct $\Xi_{d,\nu}(x)$ as

$$\begin{split} \gamma &= \frac{\mathrm{d}}{\mathrm{d}x} \rho(x + \frac{1}{6R^2} x^2 + \cdots) \big|_{x=0} = \rho'(0) = \frac{\gamma}{\mathrm{Euclid}} \\ \kappa_1 &= \frac{\partial}{\partial x_{12}} \kappa^{(3)} (x_{12} + \frac{1}{6R^2} x_{12}^2 + \cdots, 0, 0) \big|_{x=0} \\ &= \frac{\partial}{\partial x_{12}} \kappa^{(3)} (x_{12}, 0, 0) \big|_{x=0} = \frac{\kappa_1}{\mathrm{Euclid}} \end{split}$$

etc.

Because the second derivatives do not appear, all of the parameters are equivalent for the Euclidean space case and the sphere case.

Expected MFs: the spherical case

Theorem

The expected MFs for the weakly non-Gaussian orthogonally invariant random field X(t), $t \in T$, on the sphere \mathbb{S}_{R}^{n} are

$$\mathbb{E}[\chi(T_x)] = \mathbb{E}[\mathcal{L}_0(T_x)] = \sum_{d=0}^n \mathcal{L}_d(T) \Xi_{d,\nu}(x),$$
$$\mathbb{E}[\mathcal{L}_k(T_x)] = \sum_{d=0}^{n-k} \frac{\Gamma(\frac{k+d+1}{2})\Gamma(\frac{1}{2})}{\Gamma(\frac{k+1}{2})\Gamma(\frac{d+1}{2})} \mathcal{L}_{k+d}(T) \Xi_{d,\nu}(x), \quad k = 1, \dots, n$$

The expression for the EC density $\Xi_{d,\nu}(x)$ is the same as in the Euclidean space case.

Remark

These formulas look exactly the same as in the Euclidean case. However, the T is a subset of the sphere and $\mathcal{L}_d(T)$ depends on R, and the definitions of ρ and $\kappa^{(k)}$ constructing $\Xi_{d,\nu}$ are different.

Corollary: $X(\cdot)$ is Gaussian and $T = \mathbb{S}_R^n$ (the whole sphere)

Recall that

$$\mathcal{L}_{j}(\mathbb{S}_{R}^{n}) = \begin{cases} 2R^{j} \binom{n}{j} \frac{\Omega_{n+1}}{\Omega_{n-j+1}} & (n-j: \text{even}), \\ 0 & (\text{otherwise}), \end{cases} \quad \Omega_{d} = \frac{2\pi^{d/2}}{\Gamma(d/2)} \end{cases}$$

Corollary

When $X(\cdot)$ is Gaussian and $T = \mathbb{S}^n_R$ (the whole sphere),

$$\begin{split} \mathbb{E}[\chi(T_x)] &= \frac{(\gamma R^2)^{n/2} \Omega_{n+1}}{(2\pi)^{n/2}} \phi(x) H_n(x; (\gamma R^2)^{-1}), \\ \text{where} \quad H_n(x; \delta) &= \sum_{k=0}^{[n/2]} \delta^{k/2} \frac{n!}{2^k k! (n-2k)!} H_{n-2k-1}(x) \end{split}$$

(Cheng and Xiao (2016); for n = 2, Schmalzing and Górski (1998))

An implication: When the random field is observed on the celestial sphere \mathbb{S}_1^2

Suppose that n = 2 and the random field on the sphere S²_R is observed on the celestial sphere S²₁:

 $\overline{X}(t) = X(Rt), \quad t \in \overline{T} = R^{-1}T \subset \mathbb{S}_1^n \text{ (unit sphere)}$

► Then,
$$\mathcal{L}_j(T) = R^j \mathcal{L}_j(\overline{T})$$
, $\gamma = \operatorname{Var}(\partial X(t)/\partial t^1) = \overline{\gamma}R^{-1}$ with $\overline{\gamma} = \operatorname{Var}(\partial \overline{X}(t)/\partial t^1)$,

$$\mathbb{E}[\chi(\overline{T}_x)] = \mathbb{E}[\chi(T_x)] = \sum_{j=0}^n (\gamma/2\pi)^{j/2} H_{j-1}(x) \mathcal{L}_j(T)$$
$$= \sum_{j=0}^n (\overline{\gamma}/2\pi)^{j/2} H_{j-1}(x) \mathcal{L}_j(\overline{T})$$

No information on *R* can be extracted from the MFs
► This is the case even when X(·) is weakly non-Gaussian.

Summary

What we have done:

Expected MFs formulas are provided when the random field X(·) is isotropic (or, orthogonally invariant) and weakly non-Gaussian in the Euclidean space (or, on the sphere, respectively).

Take home messages:

- 1. The boundary effect of the MF is often substantial and cannot be ignored. The boundary correction is easy and should be incorporated always.
- 2. The expected MF for the sphere (=space of constant curvature) is almost same as the Euclidean space case.
- 3. The MFs of the celestial sphere data do not contain the information on the curvature R^{-1} .

References I

- Chamandy, N., Worsley, K. J., Taylor, J. and Gosselin, F. (2008). Tilted Euler characteristic densities for Central Limit random fields, with application to "bubbles", *Ann. Statist.*, **36** (5), 2471–2507.
- 2. Cheng, D. and Xiao, Y. (2016). Excursion probability of Gaussian random fields on sphere Bernoulli **22** (2), 1113–1130.
- 3. Kuriki, S. and Matsubara, T. (2023). Asymptotic expansion of the expected Minkowski functional for isotropic central limit random fields, *Advances in Applied Probability*, **55** (4) (Dec 2023).
- Matsubara, T. & Kuriki, S., Weakly non-Gaussian formula for the Minkowski functionals in general dimensions, Physical Review D, 104, 103522 (Nov 2021).
- Matsubara, T., Hikage, C. & Kuriki,S., Minkowski functionals and the nonlinear perturbation theory in the large-scale structure: second-order effects, Physical Review D, 105, 023527 (Jan 2022).

- 6. Schmalzing, J. and Górski, M. G. (1998). Minkowski functionals used in the morphological analysis of cosmic microwave background anisotropy maps, *Mon. Not. R. Astron. Soc.*, **297**, 355–365.
- Tomita, H. (1986). Curvature invariants of random interface generated by Gaussian fields, *Progress of Theoretical Physics*, **76** (4), 952–955

Assumption

(i)
$$t \mapsto X_N(t)$$
 is of C^2 a.s.
(ii) Recall that $t = (t^j)_{1 \le j \le n}$. There exists

$$\frac{\partial^8 \mathbb{E} [X_N(t_1) X_N(t_2) X_N(t_3) X_N(t_4)]}{\partial t_1^{i_1} \partial t_1^{j_1} \cdots \partial t_4^{i_4} \partial t_4^{j_4}} \quad \text{around } t_1 = t_2 = t_3 = t_4$$

(iii) For t fixed, $(X_N(t), \nabla X_N(t), \nabla^2 X_N(t))$ has a density p_N . p_N is bounded for some N. It has a moment of order $\binom{n+2}{2} + 1$.

Key identities on the Hermite polynomial

►
$$A \sim \text{GOE}(n)$$
, that is, $A = (a_{ij}) \in \text{Sym}(n)$, $a_{ii} \sim \mathcal{N}(0, 2)$,
 $a_{ij} \sim \mathcal{N}(0, 1)$ $(i < j)$. Then, $\mathbb{E}[e^{\text{tr}(\Theta A)}] = e^{\text{tr}(\Theta^2)}$

▶ For $B = (b_{ij}) \in Sym(n)$, define a matrix differential operator

$$(D_B)_{ij} = (1/2)(1+\delta_{ij})(\partial/\partial b_{ij}) \quad (i \le j)$$

Lemma

Let $A \sim \operatorname{GOE}(n)$. Let $m = \sum_{i=1}^{\ell} c_i$. Then,

$$(-1/2)^{m-\ell}(n)_m H_{n-m}(x) = \mathbb{E} \left[\operatorname{tr}(D_A^{c_1}) \cdots \operatorname{tr}(D_A^{c_\ell}) \det(xI_n + A) \right]$$
$$= \det(xI + D_\Theta) \left(e^{\operatorname{tr}(\Theta^2)} \operatorname{tr}(\Theta^{c_1}) \cdots \operatorname{tr}(\Theta^{c_\ell}) \right) \Big|_{\Theta = 0}$$

In particular, when $\ell = m = 0$,

$$H_n(x) = \mathbb{E}[\det(xI_n + A)] = \det(xI_n + D_{\Theta})e^{\operatorname{tr}(\Theta^2)}\Big|_{\Theta=0}$$