

Volume of Tubes and Distribution of the Maxima of Gaussian Random Fields

Satoshi Kuriki and Akimichi Takemura

1. Introduction

1.1. Gaussian random fields on manifolds. Let $X(p)$, $p \in \widetilde{M}$, be a real-valued random field on an orientable manifold \widetilde{M} . In this article, we consider a Gaussian random field with a smooth sample path and one-dimensional standard normal marginals (i.e., $X(p) \sim N(0, 1)$ for each $p \in \widetilde{M}$). Let $r(p, q) = \text{Cov}(X(p), X(q))$ denote the covariance function. If $r(p, q)$ is sufficiently smooth in a neighborhood of $p = q$ for each p , then the sample path is smooth with probability one and we can differentiate $X(p)$ to obtain Gaussian random fields ∇X , $\nabla^2 X$. We do not consider nonsmooth fields, such as Brownian motions or Ornstein-Uhlenbeck processes. Let $t = (t^i)$, $s = (s^i)$ denote local coordinate systems around p and q , respectively, and let $X_i(p) = \partial X(p)/\partial t^i$. A metric

$$(1.1) \quad g_{ij}(p) = \text{Cov}(X_i(p), X_j(p)) = \left. \frac{\partial^2 r(p, q)}{\partial t^i \partial s^j} \right|_{p=q}$$

is defined at each $p \in \widetilde{M}$. This metric is based on the following consideration.

Let ξ_1, ξ_2, \dots be independent standard normal $N(0, 1)$ random variables and let $\{\psi^k\}_{k \geq 1}$ denote an orthonormal basis of a reproducing kernel Hilbert space (RKHS) associated with the covariance function $r(p, q)$. Then $X(p)$ is expressed as ([2], Theorem 3.7)

$$(1.2) \quad X(p) = \sum_{k=1}^{\infty} \xi_k \psi^k(p),$$

where $\sum_{k=1}^{\infty} \psi^k(p)^2 = 1$ because we have assumed that the variance is equal to one. If the map $p \mapsto \psi(p)$ is a smooth injection from \widetilde{M} to ℓ^2 , then a metric on \widetilde{M} is induced from the inner product of ℓ^2 as $g_{ij}(p) = \sum_{k=1}^{\infty} (\partial \psi^k(p)/\partial t^i)(\partial \psi^k(p)/\partial t^j)$, which is equal to (1.1).

Now, let M be a compact submanifold of \widetilde{M} . In this article, we investigate an approximation to the distribution (upper tail probability) of the maximum of $X(p)$ over M ,

$$(1.3) \quad P\left(\max_{p \in M} X(p) \geq c\right)$$

This article originally appeared in Japanese in *Sūgaku* **60** (2) (2008), 134–155.
2000 *Mathematics Subject Classification*. Primary 60G60; Secondary 53C65.

when c is large. Our approach is to approximate (1.3) using integral-geometric quantities of the submanifold M .

We refer to (1.2) as a Karhunen-Loève (KL) expansion, although it is a generalization of the classical Karhunen-Loève expansion ([2]).

1.2. Tube method and Euler characteristic heuristic. Two methods are known to be applicable for approximating (1.3). One is the tube method and the other is the Euler characteristic heuristic. We give expositions of these methods based on works by the present authors ([26], [29], [42], [43]), J. Taylor and R. Adler ([44], [5]), and Taylor, et al. [45].

Section 2 presents an exposition of the tube method. A tube means a tubal neighborhood on the unit sphere in a Euclidean space. H. Hotelling [14] pointed out the relation between the distribution of the likelihood ratio test for a nonlinear regression model and the volume of a tube about a curve on the unit sphere and computed significance levels of the likelihood ratio test by explicitly computing the volume of the tube. H. Weyl [46] generalized the result to a general dimension. Their result is now known as the Hotelling-Weyl theorem. Although the Hotelling-Weyl theorem played an important role in the development of differential geometry ([23], [12]), it has been forgotten in statistics for a long time. More recently, Knowles and Siegmund [22] and Sun [38], [39] revived Hotelling's method and applied the result to some problem by pointing out that the probability (1.3) can be reduced to the volume of a tube when the KL-expansion (1.2) is a finite sum. This is the tube method.

We provide a detailed exposition of the Euler characteristic heuristic in Section 3. If the sample paths of a one-dimensional stochastic process $X(t)$, $t \in T \subset \mathbb{R}$ are smooth, then we can evaluate the expected number of upcrossings, that is, the number of times the graph $(t, X(t))$ crosses the horizontal line $X(t) = c$ from below ([21], [7]). This expected value for a large c has been traditionally used as an approximation to the tail probability $P(\max_{t \in T} X(t) \geq c)$ of the maximum of the stochastic process in the field of signal processing and other fields. As a generalization of this method to the general dimension, the Euler characteristic heuristic was proposed by Adler and A.M. Hasofer (e.g., [4], [13], [1]) and further developed as a practical method by Worsley (e.g., [48], [49]).

The relation between the tube method and the Euler characteristic heuristic was initially not understood ([3]). However, the authors ([26], [42]) proved that these two methods are equivalent for Gaussian random fields with a finite KL-expansion. The Euler characteristic heuristic is more general than the tube method. However, the tube method has many applications to multivariate analysis ([29], [30], [24]) and it presents more concrete geometric pictures in terms of the tubes. Furthermore, the evaluation of the error term by the tube method can be extended to the Euler characteristic heuristic. Therefore, we present separate expositions of the two methods. The results on the Euler characteristic heuristic including the most recent ones are extensively reviewed in the new book by Adler and Taylor [5].

In statistics, there is a strong need to evaluate the distribution of the maximum of a random field. Suppose that $X(p)$ is a test statistic for each p . Then, (1.3) corresponds to the adjustment of the p -value due to multiple testing (multiple comparisons). In hypothesis-testing problems, where the parameter under an alternative hypothesis is restricted to a cone, the asymptotic null distribution of the log

likelihood ratio statistic is the same as the distribution of the maximum of a Gaussian random field with standard normal marginals, and they can be treated by the tube method ([41], [32]). More generally, in nonregular parametric statistical models, which are called singular models (or locally conic models), the asymptotic null distribution of the log likelihood ratio statistic is the same as the distribution of the maximum of a Gaussian random field ([8], [11]). Note that Hotelling's nonlinear regression model is a typical singular statistical model. For these hypothesis-testing problems, we need the upper tail probability (such as the upper 5 percentile) and the approximation for large c is very useful.

In Section 4 we present two applications of the tube method to statistics. For other applications, see [30], [11], [25]. Other topics closely related to the tube method and the Euler characteristic heuristic include improvement of joint confidence sets ([34], [18], [40]) and the discrete tube method ([35], [36], [37]).

2. Tube method

2.1. Volume of tubes and distribution of the maxima. The standard form of our problem for the case that the KL-expansion is finite, containing n terms, is written as follows. Let $\mathbb{S}^{n-1} = \mathbb{S}(\mathbb{R}^n)$ be the unit sphere in \mathbb{R}^n and let $M \subset \mathbb{S}^{n-1}$ be a closed subset of \mathbb{S}^{n-1} . Let the elements of $\xi = (\xi_1, \dots, \xi_n)$ be independent and standard normal random variables. (We write this as $\xi \sim N_n(0, I_n)$.) $\langle \cdot, \cdot \rangle$ denotes the standard inner product of \mathbb{R}^n . Our problem is to find the distribution of the maximum of the Gaussian random field $X(p) = \langle \xi, p \rangle$, $p \in M$:

$$(2.1) \quad P\left(\max_{p \in M} \langle \xi, p \rangle \geq c\right).$$

This corresponds to setting $\psi(M)$ and $\psi(\widetilde{M})$ to M and \mathbb{S}^{n-1} in Section 1.

The tube (spherical tube) of radius θ about M is defined to be the set of points on \mathbb{S}^{n-1} whose great-circle distance to M is less than or equal to θ :

$$M_\theta = \left\{ q \in \mathbb{S}^{n-1} \mid \text{dist}(q, M) \leq \theta \right\}, \quad \text{dist}(q, M) = \min_{p \in M} \cos^{-1} \langle q, p \rangle.$$

For an n -dimensional standard normal random vector $\xi \sim N_n(0, I_n)$, its "length" $\|\xi\|$ and its "direction" $\zeta = \xi/\|\xi\|$ are independently distributed and the distribution of ζ is the uniform distribution over the unit sphere $\text{Unif}(\mathbb{S}^{n-1})$. Hence,

$$(2.2) \quad \begin{aligned} P\left(\max_{p \in M} \langle \xi, p \rangle \geq c\right) &= E \left[P\left(\max_{p \in M} \langle \zeta, p \rangle \geq \frac{c}{\|\xi\|} \mid \|\xi\|\right) \right] \\ &= E \left[P\left(\text{dist}(\zeta, M) \leq \cos^{-1}\left(\frac{c}{\|\xi\|}\right) \mid \|\xi\|\right) \right] \\ &= \frac{1}{\text{Vol}(\mathbb{S}^{n-1})} E \left[\text{Vol}\left(M_{\cos^{-1}(c/\|\xi\|)}\right) \right], \end{aligned}$$

where $\text{Vol}(\cdot)$ is the $(n-1)$ -dimensional volume. If the volume of the tube $\text{Vol}(M_\theta)$ can be evaluated for every θ , then we can integrate it once (that is, we can take the expected value with respect to $\|\xi\|$) to obtain the tail probability of the maximum (2.1). Furthermore, we know that the square of $\|\xi\|$ has the chi-square distribution with n degrees of freedom. By writing down this expected value and after some calculations, we obtain

$$\frac{P(\max_{p \in M} \langle \xi, p \rangle \geq c)}{c^n e^{-c^2/2}} = \frac{1}{2(2\pi)^{n/2}} \int_0^\infty \text{Vol}(M_{\cos^{-1}(1/\sqrt{\eta+1})})(\eta+1)^{n/2-1} e^{-c^2\eta/2} d\eta.$$

We see that $\text{Vol}(M_\theta)$ and the probability (2.1) are effectively related by the Laplace transformation and its inversion, and hence they are in a one-to-one relation. Furthermore, as $\theta \downarrow 0$, the asymptotic behavior of $\text{Vol}(M_\theta)$ corresponds to the asymptotic behavior of the upper tail probability (2.1) as $c \uparrow \infty$. As shown in the next subsection, it is easy to evaluate $\text{Vol}(M_\theta)$ for small θ . The basic idea of the tube method is to utilize this information to approximate the upper tail probability of the maximum as $c \uparrow \infty$.

2.2. Evaluation of the volume of a tube. We make some regularity assumptions on M for evaluating the volume of the tube M_θ . For applications to multivariate analysis, we need to consider some complicated M 's, which are defined in terms of linear inequalities or the nonnegative definite constraints of matrices ([27], [24]). With these applications in mind we define “piecewise smooth manifolds,” which are a generalization of “manifolds with boundaries” ([42]).

ASSUMPTION 2.1. Let M be a closed (i.e., without a boundary) m -dimensional C^2 -manifold or a piecewise smooth m -dimensional C^2 -manifold. Here, M is called a piecewise smooth m -dimensional C^r -manifold if

(i) M is written as a partition $M = \bigsqcup_{d=0}^m \partial M_d$ ($\partial M_m \neq \emptyset$), such that each component ∂M_d is a d -dimensional C^r -manifold consisting of finite connected components, and

(ii) M is C^r diffeomorphic to a convex cone $\mathbb{R}^d \times K$ in a neighborhood of each $p \in \partial M_d$, where $K \subset \mathbb{R}^{m-d}$ is a proper $(m-d)$ -dimensional closed convex cone.

The boundary and the set of inner points of M are $\partial M = \bigsqcup_{d=0}^{m-1} \partial M_d$ and $\text{int } M = \partial M_m$, respectively.

In (ii) of Assumption 2.1, we require K to be a convex cone. Therefore M possesses a convex tangent cone $S_p M$ at each p . Since $M \subset \mathbb{S}^{n-1}$, we have $S_p M \subset T_p \mathbb{S}^{n-1}$, where $T_p \mathbb{S}^{n-1} = \{v \in \mathbb{R}^n \mid \langle v, p \rangle = 0\}$ denotes the tangent space of \mathbb{S}^{n-1} at p .

Let

$$N_p M = \{v \in T_p \mathbb{S}^{n-1} \mid \langle v, u \rangle \leq 0, \forall u \in S_p M\}$$

denote the dual cone (normal cone) of the tangent cone $S_p M$ within the tangent space $T_p \mathbb{S}^{n-1}$. Note that $T_p \mathbb{S}^{n-1}$ is written as a direct sum $T_p \mathbb{S}^{n-1} = S_p M \oplus N_p M$.

Let q be a point of \mathbb{S}^{n-1} outside M . The point $p = \text{pr}(q)$ attaining $\min_{p \in M} \text{dist}(q, p)$ is called a “projection of q onto M .” For each $q \in \mathbb{S}^{n-1} \setminus M$, let $p = \text{pr}(q)$, $\psi = \text{dist}(q, p)$, and $v = (q - p \cos \psi) / \sin \psi$. Then q is written as

$$q = p \cos \psi + v \sin \psi, \quad v \in \mathbb{S}(N_p M) \quad (\text{where } \mathbb{S}(N_p M) = N_p M \cap \mathbb{S}^{n-1}).$$

If q is close to M , then $\text{pr}(q)$ is unique. However, if q is far away from M , then there may be two distinct points $p, p' \in M$ minimizing the distance from q (they are equidistant from q). The maximum of the distance guaranteeing the uniqueness of all the projections is called the critical radius of the tube.

DEFINITION 2.2 (Critical radius). We say that the tube M_θ does not have a self-intersection if every point $q \in M_\theta \setminus M$ is uniquely written as

$$q = p \cos \psi + v \sin \psi, \quad p \in M, \quad v \in \mathbb{S}(N_p M), \quad \psi \in (0, \theta].$$

The supremum of the radius θ such that M_θ does not have a self-intersection,

$$\theta_c = \sup\{\theta \geq 0 \mid M_\theta \text{ does not have a self-intersection}\},$$

is the critical radius (reach) of M (Figure 1).

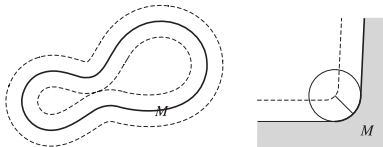


FIGURE 1. Tubes with a radius equal to the critical radius

THEOREM 2.3 ([43]). *If $M \subset \mathbb{S}^{n-1}$ satisfies Assumption 2.1, then its critical radius θ_c is positive.*

From Definition 2.2, if $\theta < \theta_c$, then M_θ does not have a self-intersection and $q \in M_\theta \setminus M$ is uniquely written in terms of the triplet (p, v, ψ) . The elements of this triplet are called the tube coordinates (Fermi coordinates) of q (Figure 2).

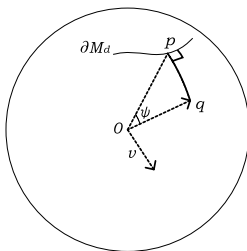


FIGURE 2. Tube coordinates (Fermi coordinates)

The Jacobian of the transformation $q \leftrightarrow (p, v, \psi)$ is given as follows.

LEMMA 2.4 ([28], [42]). *Let $p = \text{pr}(q) \in \partial M_d$ denote the projection of $q \in M_\theta \setminus M$ onto M . Then*

$$(2.3) \quad dS^{n-1}(q) = \det(I_d \cos \psi + H_p(v) \sin \psi) \sin^{n-d-2} \psi \, dM_d(p) \, d\psi \, dS_p^{n-d-2}(v).$$

Here, $dS^{n-1}(q)$ is the volume element of \mathbb{S}^{n-1} at q , $dM_d(p)$ is the volume element of ∂M_d at p , $dS_p^{n-d-2}(v)$ is the volume element of $\mathbb{S}(N_p M)$ at v . A $d \times d$ matrix $H_p(v)$ is the v -component of the second fundamental form of ∂M_d at p and its (i, j) -element $\langle -\partial^2 \phi / \partial t^i \partial t^j, v \rangle|_p$ is given in terms of a normal coordinate system $p = \phi(t)$, $t = (t^i)_{1 \leq i \leq d}$ at p .

We obtain the volume of a tube with radius θ smaller than θ_c by integrating (2.3) with respect to $dS^{n-1}(q)$ over the region $p \in \partial M_d$, $v \in \mathbb{S}(N_p M)$, $0 \leq \psi \leq \theta$ and then summing the terms over all values of d .

THEOREM 2.5 (Volume of a tube, [42]). For $0 \leq \theta \leq \max(\theta_c, \pi/2)$,

$$(2.4) \quad \begin{aligned} \text{Vol}(M_\theta) &= \sum_{d=0}^m \int_{\partial M_d} dM_d(p) \int_{\mathbb{S}(N_p M)} dS_p^{n-d-2}(v) \int_0^\theta d\psi \\ &\quad \times \det(I_d \cos \psi + H_p(v) \sin \psi) \sin^{n-d-2} \psi \\ &= \Omega_n \sum_{e=0}^m w_{m+1-e} \bar{B}_{(m+1-e)/2, (n-m-1+e)/2}(\cos^2 \theta). \end{aligned}$$

Here, $\Omega_n = \text{Vol}(\mathbb{S}^{n-1}) = 2\pi^{n/2}/\Gamma(n/2)$,

$$\begin{aligned} w_{m+1-e} &= \frac{1}{\Omega_{m+1-e} \Omega_{n-m-1+e}} \\ &\quad \times \sum_{d=m-e}^m \int_{\partial M_d} dM_d(p) \int_{\mathbb{S}(N_p M)} dS_p^{n-d-2}(v) \text{tr}_{d-m+e} H_p(v), \end{aligned}$$

tr_l is the l -th elementary symmetric function of the eigenvalues of a symmetric matrix (e.g., $\text{tr}_0 A = 1$, $\text{tr}_1 A = \text{tr} A$, $\text{tr}_d A = \det A$), and

$$\bar{B}_{a,b}(c) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_c^1 (1-t)^{a-1} t^{b-1} dt$$

is the upper probability of the beta distribution with parameter (a, b) .

The second equality in (2.4) follows from the expansion formula for the determinant $\det(I_d + A) = \sum_{l=0}^d \text{tr}_l A$.

The above formula gives the correct volume only for small radius θ . In the tube method, we pretend that the volume of the tube (2.4) is correct for all values of θ and plug it in (2.2). In (2.4), let $\theta := \cos^{-1}(c/\|\xi\|)$ and take the expected value with respect to $\|\xi\|^2 \sim \chi_n^2$. Then we obtain the following approximation.

THEOREM 2.6 (Approximation by the tube method, [42]). The tube formula approximation to the upper tail probability (2.1) of the maximum of a Gaussian random field is given by

$$(2.5) \quad \begin{aligned} \hat{P}\left(\max_{p \in M} \langle \xi, p \rangle \geq c\right) &= \frac{1}{(2\pi)^{n/2}} \sum_{d=0}^m \int_{\partial M_d} dM_d(p) \int_{\mathbb{S}(N_p M)} dS_p^{n-d-2}(v) \\ &\quad \times \int_c^\infty dr \int_0^\infty ds e^{-(r^2+s^2)/2} \det(rI_d + sH_p(v)) s^{n-d-2} \\ &= \sum_{e=0}^m w_{m+1-e} \bar{G}_{m+1-e}(c^2), \end{aligned}$$

where

$$\bar{G}_\nu(c) = \frac{1}{2^{\nu/2} \Gamma(\nu/2)} \int_c^\infty t^{\nu/2-1} e^{-t/2} dt$$

is the upper probability of the chi-square distribution with ν degrees of freedom. (We have used the notation \hat{P} as an approximation to a probability measure P .)

We will show in several steps that the above tube method approximation is indeed a good approximation.

2.3. Evaluation of the critical radius. Here we present a method to evaluate the critical radius θ_c . It also yields an exact tube formula for the volume of a tube. The following argument is given in a more general framework by Taylor, et al. [45].

For each $p \in M$ we want to characterize the set of points q such that p is their projection:

$$\text{pr}^{-1}(\{p\}) = \{q \in \mathbb{S}^{n-1} \mid p = \text{pr}(q)\}.$$

Recall that q can be written as $q = p \cos \psi + v \sin \psi$ ($v \in \mathbb{S}(N_p M)$). Since p is the closest point to q in M , we have $\text{dist}(q, p') > \text{dist}(q, p)$, $\forall p' \in M \setminus \{p\}$. This can be rewritten as

$$\cot \psi > \frac{\langle v, p' \rangle}{1 - \langle p, p' \rangle}, \quad \forall p' \in M \setminus \{p\}.$$

Therefore, if $\text{pr}(q) = p$, then $\psi = \cos^{-1}\langle q, p \rangle$ satisfies

$$\psi \leq \theta_c(p, v), \quad \text{where} \quad \cot \theta_c(p, v) = \sup_{p' \in M \setminus \{p\}} \frac{\langle v, p' \rangle}{1 - \langle p, p' \rangle}.$$

Conversely, if $\psi < \theta_c(p, v)$, then the projection of $q = p \cos \psi + v \sin \psi$ is $\text{pr}(q) = p$. It follows that

$$\begin{aligned} \left\{ p \cos \psi + v \sin \psi \mid v \in \mathbb{S}(N_p M), 0 \leq \psi < \theta_c(p, v) \right\} &\subset \text{pr}^{-1}(\{p\}) \\ &\subset \left\{ p \cos \psi + v \sin \psi \mid v \in \mathbb{S}(N_p M), 0 \leq \psi \leq \theta_c(p, v) \right\}. \end{aligned}$$

Since the difference of the left-hand side and the right-hand side is a null set (a set of zero spherical volume), by restricting the integral with respect to ψ in (2.4) to the interval $\psi < \theta_c(p, v)$, the self-intersection of the integration is avoided even for a tube with radius greater than the critical radius θ_c .

THEOREM 2.7 (Exact tube formula, [43]).

$$\begin{aligned} \text{Vol}(M_\theta) &= \sum_{d=0}^m \int_{\partial M_d} dM_d(p) \int_{\mathbb{S}(N_p M)} dS_p^{n-d-2}(v) \int_0^\theta d\psi \\ &\quad \times \mathbf{1}_{\{\psi < \theta_c(p, v)\}} \det(I_d \cos \psi + H_p(v) \sin \psi) \sin^{n-d-2} \psi, \end{aligned}$$

where $\mathbf{1}_{\{\cdot\}}$ is the indicator function taking values 0 or 1.

By the same derivation as that in Theorem 2.6, we have the following result.

THEOREM 2.8 (Distribution of the maximum of a Gaussian field, [43]).

(2.6)

$$\begin{aligned} P\left(\max_{p \in M} \langle \xi, p \rangle \geq c\right) &= \frac{1}{(2\pi)^{n/2}} \sum_{d=0}^m \int_{\partial M_d} dM_d(p) \int_{\mathbb{S}(N_p M)} dS_p^{n-d-2}(v) \int_c^\infty dr \int_0^\infty ds \\ &\quad \times \mathbf{1}_{\{r > s \cot \theta_c(p, v)\}} e^{-(r^2+s^2)/2} \det(rI_d + sH_p(v)) s^{n-d-2}. \end{aligned}$$

We can give the following interpretation of $\theta_c(p, v)$. Consider an arc starting from $p \in M$ in the direction $v \in \mathbb{S}(N_p M)$ along a great circle (i.e., the exponential map). If the length of the arc is smaller than $\theta_c(p, v)$, then the arc does not intersect any other arc starting from $p' \in M$ in the direction $v' \in \mathbb{S}(N_{p'} M)$. Therefore, a tube with radius smaller than $\theta_c = \inf_{p \in M, v \in \mathbb{S}(N_p M)} \theta_c(p, v)$ does not have a self-intersection. In addition, a tube with radius greater than θ_c has at least one point

of self-intersection. This implies that θ_c is the critical radius. From this argument, we have the following theorem.

THEOREM 2.9 ([11], Section 3.4).

$$\cot \theta_c = \sup_{p \in M, q \in M \setminus \{p\}} \frac{\max_{v \in \mathbb{S}(N_p M)} \langle v, q \rangle}{1 - \langle p, q \rangle}.$$

Note that $\max_{v \in \mathbb{S}(N_p M)} \langle v, q \rangle$ is the length of the orthogonal projection of q in \mathbb{R}^n onto the cone $N_p M \subset \mathbb{R}^n$.

2.4. Bounding the error in the tube method. The error $\Delta P(c)$ of the tube formula is given by the difference between (2.5) and (2.6):

$$\begin{aligned} \Delta P(c) &= \widehat{P}(\max_{p \in M} \langle \xi, p \rangle \geq c) - P(\max_{p \in M} \langle \xi, p \rangle \geq c) \\ &= -\frac{1}{(2\pi)^{n/2}} \sum_{d=0}^m \int_{\partial M_d} dM_d(p) \int_{\mathbb{S}(N_p M)} dS_p^{n-d-2}(v) \int_c^\infty dr \int_0^\infty ds \\ &\quad \times \mathbf{1}_{\{c \leq r \leq s \cot \theta_c(p, v)\}} e^{-(r^2+s^2)/2} \det(rI_d + sH_p(v)) s^{n-d-2}. \end{aligned}$$

We expand $\det(rI_d + sH_p(v))$ as $\det(rI_d + sH_p(v)) = \sum_{l=0}^d r^{d-l} s^l \text{tr}_l H_p(v)$ and consider bounding each term in terms of the eigenvalues of $H_p(v)$ from above. We also bound the indicator function of the region for integration as

$$\mathbf{1}_{\{c \leq r \leq s \cot \theta_c(p, v)\}} \leq \mathbf{1}_{\{c \leq r\}} \mathbf{1}_{\{c \tan \theta_c \leq s\}}$$

and integrate with respect to r, s . Then we have the following upper bound for the error.

THEOREM 2.10 ([29]). *There exists a constant K such that*

$$|\Delta P(c)| \leq K \bar{G}_n((1 + \tan^2 \theta_c) c^2) \quad (= O(c^{n-2} e^{-(1+\tan^2 \theta_c) c^2/2}), \quad c \rightarrow \infty).$$

The larger the critical radius θ_c is, the smaller the error term becomes. The terms in the tube method approximation (2.5) in Theorem 2.6 are of the order $\bar{G}_\nu(c^2) = O(c^{\nu-2} e^{-c^2/2})$ as $c \rightarrow \infty$. Therefore, as long as $\theta_c > 0$, the error of the tube method is exponentially smaller than the terms in the approximation.

2.5. The case of a closed manifold. As the last topic of the tube method, we consider the case that the manifold M is an m -dimensional closed manifold (i.e., a manifold without a boundary). The tube approximation is written as

$$\begin{aligned} \widehat{P}(\max_{p \in M} \langle \xi, p \rangle \geq c) &= \frac{1}{(2\pi)^{n/2}} \int_M dM(p) \int_{\mathbb{S}(N_p M)} dS_p^{n-m-2}(v) \\ &\quad \times \int_c^\infty dr \int_0^\infty ds e^{-(r^2+s^2)/2} \det(rI_m + sH_p(v)) s^{n-m-2}, \end{aligned}$$

where $dM(p)$ is the volume element of M at p .

Since $N_p M$ is a linear space, the integration (with respect to v) of a product of an odd number of elements of the second fundamental form is zero by symmetry. We now consider integration with respect to v and s . Note that

$$\frac{1}{(2\pi)^{(n-m-1)/2}} dS_p^{n-m-2}(v) e^{-s^2/2} s^{n-m-2} ds \quad (v \in \mathbb{S}(N_p M), \quad s \in (0, \infty))$$

is a probability measure. Under this measure, the random variable sv has the standard normal distribution over the linear space $N_p M$. Since $sH_p(v) = H_p(sv)$, the integral of the integrand with respect to s, v can be expressed as the expected value with respect to the random variable sv .

In the following we consider a covariant tensor field on M . For simplicity, we consider a field expressed in terms of an orthonormal frame.

We refer to a covariant tensor $A = (a_{i_1 \dots i_k; j_1 \dots j_k})$ of even rank $2k$ as a “double form” of type (k, k) if A is antisymmetric in the first half of the indices i_1, \dots, i_k as well as in the second half of the indices j_1, \dots, j_k . We denote the set of these tensors by $D^{k,k} = D^{k,k}(M)$ ([12]).

For two double forms $A = (a_{i_1 \dots i_k; j_1 \dots j_k}) \in D^{k,k}$, $B = (b_{i_1 \dots i_l; j_1 \dots j_l}) \in D^{l,l}$, their product $AB = C = (c_{i_1 \dots i_{k+l}; j_1 \dots j_{k+l}}) \in D^{k+l, k+l}$ is defined as follows:

$$c_{i_1 \dots i_{k+l}; j_1 \dots j_{k+l}} = \frac{1}{k! l!} \sum_{\pi, \sigma \in S_{k+l}} \operatorname{sgn}(\pi) \operatorname{sgn}(\sigma) \\ \times a_{i_{\pi(1)} \dots i_{\pi(k)}; j_{\sigma(1)} \dots j_{\sigma(k)}} b_{i_{\pi(k+1)} \dots i_{\pi(k+l)}; j_{\sigma(k+1)} \dots j_{\sigma(k+l)}}$$

where S_{k+l} is the set of permutations of $\{1, \dots, k+l\}$. The set of all double forms $\bigoplus_{k \geq 0} D^{k,k}$ constitutes a commutative algebra.

Any covariant tensor $A = (a_{ij})$ of rank 2 can be regarded as a double form of type $(1, 1)$. The following relations hold for double forms: $(A^2)_{i_1 i_2; j_1 j_2} = 2(a_{i_1 j_1} a_{i_2 j_2} - a_{i_1 j_2} a_{i_2 j_1})$, $(AB)_{i_1 i_2; j_1 j_2} = a_{i_1 j_1} b_{i_2 j_2} - a_{i_1 j_2} b_{i_2 j_1} - a_{i_2 j_1} b_{i_1 j_2} + a_{i_2 j_2} b_{i_1 j_1}$. Note also that the curvature tensor is a double form of type $(2, 2)$.

Define the trace of $A \in D^{k,k}$ by

$$\operatorname{Tr} A = \operatorname{Tr}(a_{i_1 \dots i_k; j_1 \dots j_k}) = \frac{1}{k!} \sum_{i_1, \dots, i_k=1}^m a_{i_1 \dots i_k; i_1 \dots i_k}.$$

Then the following formulas hold ([9], [44]).

(i) Let $A = (a_{ij})_{m \times m}$ be a double form of type $(1, 1)$ and regard it as a square matrix. Then its determinant can be written as

$$\det A = \frac{1}{m!} \operatorname{Tr}(A^m).$$

(ii) Let $A = (a_{i_1 \dots i_k; j_1 \dots j_k}) \in D^{k,k}$ ($1 \leq i_1, \dots, j_k \leq m$) and $I = (\delta_{i;j}) \in D^{1,1}$ ($m \times m$ identity matrix). Then

$$\operatorname{Tr}(AI^j) = \frac{(m-k)!}{(m-k-j)!} \operatorname{Tr} A.$$

In particular,

$$\operatorname{Tr}(I^j) = \frac{m!}{(m-j)!}.$$

Let V be the standard normal random vector over the linear space $N_p M$. We regard the V component of the second fundamental form $H(V) = (H_{ij}(V))$ at point p as a double form of type $(1, 1)$. The relation (the Gauss equation) between the curvature tensor and the second fundamental form is given by

$$-R_{ijkl} = E[H_{ik}(V)H_{jl}(V) - H_{il}(V)H_{jk}(V)] + (\delta_{ik}\delta_{jl} - \delta_{il}\delta_{jk})$$

([46]). This relation can be simply written as $-R = (1/2)(E[H(V)^2] + I^2)$ in terms of the double form. From the properties of moments of the Gaussian distribution, we have the following lemma.

LEMMA 2.11 ([44]).

$$E[H(V)^{2j}] = \frac{(2j)!}{j!2^j}(-2R - I^2)^j.$$

Note that the case $j = 1$ is the Gauss equation. In the following we abbreviate $H(V)$ as H :

$$\begin{aligned} (2.7) \quad E[\det(rI_m + H)] &= \frac{1}{m!} E[\mathrm{Tr}(rI + H)^m] \\ &= \frac{1}{m!} \sum_{j=0}^{\lfloor m/2 \rfloor} \binom{m}{2j} r^{m-2j} E[\mathrm{Tr}(I^{m-2j} H^{2j})] \\ &= \sum_{j=0}^{\lfloor m/2 \rfloor} \frac{1}{(2j)!} r^{m-2j} E[\mathrm{Tr}(H^{2j})] \\ &= \sum_{j=0}^{\lfloor m/2 \rfloor} \frac{(-1)^j}{j!} r^{m-2j} \mathrm{Tr}\left(R + \frac{1}{2}I^2\right)^j. \end{aligned}$$

Integrating both sides with respect to

$$(2.8) \quad \frac{1}{(2\pi)^{(m+1)/2}} \int_M dM(p) \int_c^\infty e^{-r^2/2} dr,$$

we obtain the following theorem.

THEOREM 2.12 (The case of a closed manifold (the first statement), [44]).

$$(2.9) \quad \widehat{P}\left(\max_{p \in M} \langle \xi, p \rangle \geq c\right) = \sum_{j=0}^{\lfloor m/2 \rfloor} w_{m+1-2j} \bar{G}_{m+1-2j}(c^2),$$

where

$$(2.10) \quad w_{m+1-2j} = \frac{(-1)^j \Gamma((m+1)/2 - j)}{\pi^{(m+1)/2} 2^{j+1} j!} \int_M \mathrm{Tr}\left(R + \frac{1}{2}I^2\right)^j dM(p).$$

We continue expanding for obtaining another expression:

$$\begin{aligned} \text{RHS of (2.7)} &= \sum_{j=0}^{\lfloor m/2 \rfloor} \frac{(-1)^j}{j!} r^{m-2j} \sum_{l=0}^j \binom{j}{l} \frac{1}{2^{j-l}} \mathrm{Tr}(R^l I^{2(j-l)}) \\ &= \sum_{j=0}^{\lfloor m/2 \rfloor} \frac{(-1)^j}{j! 2^j} r^{m-2j} \sum_{l=0}^j \binom{j}{l} 2^l \frac{(m-2l)!}{(m-2j)!} \mathrm{Tr} R^l \\ &\quad (\text{put } k = j - l) \\ &= \sum_{l=0}^{\lfloor m/2 \rfloor} \frac{(-1)^l}{l!} \mathrm{Tr} R^l \sum_{k=0}^{\lfloor (m-2l)/2 \rfloor} \frac{(-1)^k}{2^k k!} \frac{(m-2l)!}{(m-2l-2k)!} r^{m-2l-2k} \\ &= \sum_{l=0}^{\lfloor m/2 \rfloor} \frac{(-1)^l}{l!} \mathrm{Tr} R^l H_{m-2l}(r), \end{aligned}$$

where

$$H_n(r) = \sum_{k=0}^{\lfloor n/2 \rfloor} \frac{(-1)^k}{2^k k!} \frac{n!}{(n-2k)!} r^{n-2k}$$

is the Hermite polynomial of degree n . We again integrate with respect to (2.8).

THEOREM 2.13 (The case of a closed manifold (the second statement), [44]).
(2.11)

$$\widehat{P}\left(\max_{p \in M} \langle \xi, p \rangle \geq c\right) = \frac{1}{(2\pi)^{(m+1)/2}} \sum_{l=0}^{\lfloor m/2 \rfloor} \frac{(-1)^l}{l!} e^{-c^2/2} H_{m-2l-1}(c) \int_M \text{Tr } R^l dM(p),$$

where $e^{-c^2/2} H_{-1}(c) = \int_c^\infty e^{-r^2/2} dr$.

3. Euler characteristic heuristic

3.1. The excursion set. In this section, we survey the Euler characteristic heuristic, which is another method for approximating the distribution of the maximum of a random field with a smooth sample path. We explain the expression for the approximation formula, the equivalence of the Euler characteristic heuristic and the tube method applied to Gaussian random fields with a finite KL-expansion, and the bound for the approximating error, in that sequence.

Let \widetilde{M} be an orientable m -dimensional C^3 -manifold. Consider a random field $X(p)$, $p \in \widetilde{M}$, defined on \widetilde{M} with a sample path of C^2 -class. Assume that the random field has a variance, and that $\text{Corr}(X(p), X(q)) = 1 \Leftrightarrow p = q$. We do not assume the normality of the random field for a while. Let $M \subset \widetilde{M}$ be a piecewise smooth C^3 -submanifold in the sense of Assumption 2.1. We consider approximating the distribution of the maximum $\max_{p \in M} X(p)$ over M .

The set of indices $p \in M$ such that the value $X(p)$ is greater than or equal to a threshold c is called the excursion set

$$A_c = \{p \in M \mid X(p) \geq c\}.$$

From the definition of the excursion set, it follows that

$$P\left(\max_{p \in M} X(p) \geq c\right) = P(A_c \neq \emptyset) = E[1_{\{A_c \neq \emptyset\}}].$$

Here, $1_{\{\cdot\}}$ is a random variable taking the value 1 when the event occurs, and the value 0 otherwise. The Euler characteristic of A_c is denoted by $\chi(A_c)$. If we can assume the approximation

$$(3.1) \quad 1_{\{A_c \neq \emptyset\}} \approx \chi(A_c) \quad (\text{when } c \text{ is large}),$$

then, by taking the expectations of both sides, we have the approximation formula

$$P\left(\max_{p \in M} X(p) \geq c\right) = E[1_{\{A_c \neq \emptyset\}}] \approx E[\chi(A_c)] \quad (\text{when } c \text{ is large}).$$

This approximation is called the Euler characteristic heuristic. As we will see, it is simpler to evaluate the integral $E[\chi(A_c)]$ than $E[1_{\{A_c \neq \emptyset\}}]$. Thus, the above argument is useful.

Approximation (3.1) is based on the following intuitive arguments. If c is very large such that $c > \max_{p \in M} X(p)$, then $A_c = \emptyset$ and (3.1) holds trivially in the form $1_{\{A_c \neq \emptyset\}} = \chi(A_c) = 0$. In addition, unless the random field has some special structures, the point $p = p^*$ that attains the maximum should exist uniquely. If $c = X(p^*)$ exactly, then $A_c = \{p^*\}$. If $c < X(p^*)$ but c is very close to $X(p^*)$, then A_c should be a set containing p^* homeomorphic to a closed ball. For these regions, we have $1_{\{A_c \neq \emptyset\}} = \chi(A_c) = 1$, and (3.1) holds again. We cannot say anything more

in general when c is smaller. However, it is clear from the above observation that (3.1) holds in some form.

3.2. Morse's theorem and the approximation formula. As stated in Section 1, we consider the index set \widetilde{M} as an m -dimensional manifold endowed with the metric defined by the covariance matrix $g_{ij} = \text{Cov}(X_i, X_j)$ of the differential field $X_i = \partial X / \partial t^i$. In the following, we use local coordinates $(t^i)_{1 \leq i \leq m}$ of \widetilde{M} around $x \in \partial M_d$ such that (i) $t^{d+1} = \dots = t^m = 0$ represents ∂M_d , (ii) $(t^i)_{1 \leq i \leq d}$ forms local coordinates of ∂M_d around x , and (iii) at $x' \in \partial M_d$ in the neighborhood of x , $(\partial / \partial t^i)_{x'}$ ($i = d+1, \dots, m$) is orthogonal to $T_{x'} \partial M_d$.

We consider expressing the Euler characteristic of the excursion set by means of Morse's theorem. For this purpose, we provide a version of Morse's theorem for piecewise smooth manifolds.

The gradient of a function f on \widetilde{M} is written as $\nabla f = (\partial f / \partial t^i)_{1 \leq i \leq m}$ using local coordinates. The restriction of the function f to ∂M_d is denoted by $f|_{\partial M_d}$. Using the local coordinates of ∂M_d , the gradient $f|_{\partial M_d}$ and the Hesse matrix (Hesse form) are defined by $\nabla f|_{\partial M_d} = (\partial f|_{\partial M_d} / \partial t^i)_{1 \leq i \leq d}$, and $\nabla^2 f|_{\partial M_d} = (\partial^2 f|_{\partial M_d} / \partial t^i \partial t^j - \sum_{k=1}^d \Gamma_{ij}^k \partial f|_{\partial M_d} / \partial t^k)_{1 \leq i, j \leq d}$, respectively. Here, $\Gamma_{ij}^k = \sum_{l=1}^d \text{Cov}(X_{ij}, X_l) g^{lk}$ is the connection coefficient of ∂M_d . The points where the gradient becomes 0 are called the critical points.

DEFINITION 3.1 (Morse function). Let \widetilde{M} be an m -dimensional C^3 -manifold and let $M \subset \widetilde{M}$ be a piecewise smooth m -dimensional C^3 -submanifold. A C^2 -function f on \widetilde{M} is called the Morse function on M if

- (i) the critical points of f do not exist on the boundary $\partial M = \bigsqcup_{d < m} \partial M_d$ of M , and
- (ii) for every dimension $1 \leq d \leq m$, $f|_{\partial M_d}$ is not degenerate, that is, the Hesse matrix $\nabla^2 f|_{\partial M_d}$ is not degenerate on the critical points of $f|_{\partial M_d}$.

The set of all critical points of $f|_{\partial M_d}$ is denoted by Z_d . It can be proved that Z_d is a finite set if f is a Morse function on M . The number of negative eigenvalues of the Hesse matrix evaluated at the critical point $x^* \in Z_d$ of the function $f|_{\partial M_d}$ is called the index and denoted by $\text{index} \nabla^2 f|_{\partial M_d}(x^*)$.

THEOREM 3.2 (Morse's theorem for piecewise smooth manifolds, [10], [42]). *If a C^2 -function f on \widetilde{M} is a Morse function on M , then the following holds:*

$$\begin{aligned} & \chi(\{x \in M \mid f(x) \leq u\}) \\ &= \sum_{d=0}^m \sum_{k=0}^d (-1)^k \#\{x^* \in Z_d \mid \text{index} \nabla^2 f|_{\partial M_d}(x^*) = k, \\ & \quad - \nabla f(x^*) \in N_{x^*} M, f(x^*) \leq u\} \\ &= \sum_{d=0}^m \sum_{x^* \in Z_d} 1_{\{-\nabla f(x^*) \in N_{x^*} M, f(x^*) \leq u\}} \text{sgn} \det \nabla^2 f|_{\partial M_d}(x^*). \end{aligned}$$

Here, $N_x M = \{v \in T_x \widetilde{M} \mid \langle v, u \rangle \leq 0, \forall u \in S_x M\}$ is the normal cone of M at x , and sgn is the sign function taking values ± 1 according to the sign of the argument.

Morse's theorem for the manifold with boundaries (i.e., $\partial M_d = \emptyset$, $d \leq m-2$) is given in Morse and Cairns [33], Theorem 10.2. This theorem is regarded as

an extension of it. The critical point $x^* \in Z_d$ of $f|_{\partial M_d}$ is said to be an augmented critical point if $-\nabla f(x^*) \in N_{x^*}M$. Theorem 3.2 states that the Euler characteristic of a piecewise smooth manifold can be calculated by summing the indices of the augmented critical points.

Let $(T_x \partial M_d)^\perp$ be the orthogonal complement of $T_x \partial M_d$ in $T_x \widetilde{M}$. Since $\partial M_d \subset M$, we have $N_x M \subset (T_x \partial M_d)^\perp$. On the other hand, at the critical point $x^* \in Z_d$, we see that $\nabla f(x^*) \in (T_{x^*} \partial M_d)^\perp$. Let $\nabla_{\partial M_d}^\perp f(x)$ denote the orthogonal projection of $\nabla f(x) \in T_x \widetilde{M}$ onto $(T_x \partial M_d)^\perp$, that is,

$$\nabla_{\partial M_d}^\perp f(x) = (0, \dots, 0, \partial f / \partial t^{d+1}, \dots, \partial f / \partial t^m)|_x \in T_x \widetilde{M}.$$

Then the condition $-\nabla f(x^*) \in N_{x^*}M$ is equivalent to $-\nabla_{\partial M_d}^\perp f(x^*) \in N_{x^*}M$.

In order to obtain the Euler characteristic of the excursion set, we apply Theorem 3.2 by letting $f := -X$ and $u := -c$. For that purpose, $-X$ has to be a Morse function on M . Including this condition, Adler and Taylor [5] refer to the sufficient conditions under which the Euler characteristic heuristic is valid as the ‘‘suitable regularity’’ ([5], Theorem 11.3.1). They are too complicated to describe here in the general setting. We will give them later in the case of Gaussian random fields.

Under suitable regularity, the random field $-X$ becomes a Morse function on M with probability 1, and

$$\begin{aligned} \chi(A_c) &= \chi(\{p \in M \mid X(p) \geq c\}) \\ &= \sum_{d=0}^m \sum_{p^* \in Z_d} 1_{\{\nabla_{\partial M_d}^\perp X(p^*) \in N_{p^*}M, X(p^*) \geq c\}} \operatorname{sgn} \det(-\nabla^2 X|_{\partial M_d}(p^*)) \quad \text{a.s.} \end{aligned}$$

Let $\varepsilon \mathbb{D}^d \subset \mathbb{R}^d$ be a d -dimensional ball centered at the origin and with radius ε . Let

$$\delta_\varepsilon(p) = \frac{1_{\{\nabla X|_{\partial M_d}(p) \in \varepsilon \mathbb{D}^d\}}}{\operatorname{Vol}(\varepsilon \mathbb{D}^d)}.$$

As $\varepsilon \rightarrow 0$, the measure $\delta_\varepsilon(p) |\det \nabla^2 X|_{\partial M_d}(p)| \bigwedge_{i=1}^d dt^i|_p$ converges to the Dirac measure at the critical point p^* . Since the critical points are finite and discrete, the summation \sum_{p^*} can be replaced by the integration using δ_ε . Because of $\operatorname{sgn}(\det \nabla^2 X|_{\partial M_d}) |\det \nabla^2 X|_{\partial M_d}| = \det \nabla^2 X|_{\partial M_d}$, we have

$$\chi(A_c) = \sum_{d=0}^m \lim_{\varepsilon \rightarrow 0} \int_{\partial M_d} 1_{\{\nabla_{\partial M_d}^\perp X(p) \in N_p M, X(p) \geq c\}} \det(-\nabla^2 X|_{\partial M_d}(p)) \delta_\varepsilon(p) \bigwedge_{i=1}^d dt^i|_p.$$

(Note that ‘‘sgn’’ has vanished.) Consider taking the expectations of both sides. Let $\theta_{\nabla X|_{\partial M_d}(p)}(0)$ denote the density function of $\nabla X|_{\partial M_d}(p) = (X_i(p))_{1 \leq i \leq d}$ evaluated at 0. Then we have $\lim_{\varepsilon \rightarrow 0} E[\delta_\varepsilon(p)] = \theta_{\nabla X|_{\partial M_d}(p)}(0)$. Moreover,

$$\begin{aligned} E[\chi(A_c)] &= \sum_{d=0}^m \lim_{\varepsilon \rightarrow 0} \int_{\partial M_d} E \left[E[1_{\{\nabla_{\partial M_d}^\perp X(p) \in N_p M, X(p) \geq c\}} \right. \\ &\quad \left. \times \det(-\nabla^2 X|_{\partial M_d}(p)) |\nabla X|_{\partial M_d}(p)| \delta_\varepsilon(p) \right] \bigwedge_{i=1}^d dt^i|_p. \end{aligned}$$

Under suitable regularity, the integration and the limit are interchangeable and

(3.2)

$$E[\chi(A_c)] = \sum_{d=0}^m \int_{\partial M_d} E \left[1_{\{\nabla_{\partial M_d}^\perp X(p) \in N_p M, X(p) \geq c\}} \right. \\ \left. \times \det(-\nabla^2 X|_{\partial M_d}(p)) \mid \nabla X|_{\partial M_d}(p) = 0 \right] \theta_{\nabla X|_{\partial M_d}(p)}(0) \bigwedge_{i=1}^d dt^i|_p.$$

Note that thus far we did not assume the normality of the random field. Worsley [47] derived the expectation of the Euler characteristic heuristic for certain classes of the chi-square random field, the F random field, and the t random field. The authors [31] gave the approximations of the distributions of the largest eigenvalues of real random matrices following the Wishart distribution, the multivariate beta distribution, and the inverse Wishart distribution by considering the maxima of the quadratic forms of matrices.

3.3. The case of Gaussian random fields. From now on, we assume that $X(p)$, $p \in \widetilde{M}$, is a Gaussian random field with mean 0 and variance 1. Taking the differentials of both sides of $\text{Var}(X(p)) = 1$ yields $\text{Cov}(X(p), X_i(p)) = 0$, which implies that $X(p)$ and $\nabla X(p)$ are independent for each p . Furthermore, the regression residual of the second derivatives $-X_{ij}(p) = -\partial^2 X(p)/\partial t^i \partial t^j$ onto $X(p)$ and $\nabla X(p) = (X_i(p))_{1 \leq i \leq m}$ is

$$\begin{aligned} H_{ij}(p) &= -X_{ij}(p) - E[-X_{ij}(p) \mid X(p), \nabla X(p)] \\ &= -X_{ij}(p) - \{g_{ij}(p)X(p) - \Gamma_{ij}^k(p)X_k(p)\} \\ &= -(\nabla^2 X(p))_{ij} - g_{ij}(p)X(p). \end{aligned}$$

From the property of the normal distribution, $H(p) = (H_{ij}(p))$ is independent of $X(p)$ and $\nabla X(p)$ for each p . Since $\nabla X(p)$ follows the normal distribution with mean 0 and covariance matrix $(g_{ij}(p))_{1 \leq i, j \leq m}$, we have

$$\theta_{\nabla X(p)}(0) = \det(g_{ij}(p))^{-1/2} / (2\pi)^{m/2}.$$

Applying the same argument to the restriction $X|_{\partial M_d}$ of X to ∂M_d , we can rewrite (3.2) for the case of Gaussian random fields.

The suitable regularity for Gaussian random fields is given as follows ([5], Corollary 11.3.2, Theorem 12.4.2).

ASSUMPTION 3.3. Let $t = (t^i)_{1 \leq i \leq m}$ be a normal coordinate system of \widetilde{M} around $p \in M$. Put $X_i = \partial X / \partial t^i$ and $X_{ij} = \partial^2 X / \partial t^i \partial t^j$.

(i) For each $p \in M$, the joint distribution of $((X_i(p))_{1 \leq i \leq m}, (X_{ij}(p))_{1 \leq i \leq j \leq m})$ is not degenerate.

(ii) There exist constants $K, \alpha > 0$ such that for all $p \in M$ and q in its neighborhood, we have

$$\max_{1 \leq i, j \leq m} \text{Var}(X_{ij}(p) - X_{ij}(q)) \leq K |\log \|t - s\||^{-(1+\alpha)}.$$

(Here, we put the local coordinates of p, q as t, s .)

THEOREM 3.4. *Suppose that $X(p)$, $p \in \widetilde{M}$, is a Gaussian random field satisfying Assumption 3.3. Then,*

$$(3.3) \quad E[\chi(A_c)] = \sum_{d=1}^m \frac{1}{(2\pi)^{d/2}} \int_{\partial M_d} E \left[1_{\{X(p) \geq c, \nabla_{\partial M_d}^\perp X(p) \in N_p M\}} \right. \\ \left. \times \det(X(p)I_d + H(p)) \right] dM_d(p).$$

Here, $H(p) = (H_{ij}(p))_{d \times d}$ with $H_{ij}(p) = -(\nabla^2 X|_{\partial M_d})_{ij}(p) - \delta_{ij}X(p)$ is expressed in terms of a normal coordinate system of ∂M_d around p .

We consider Theorem 3.4 restricted to the Gaussian random field with a finite KL-expansion

$$(3.4) \quad X(p) = \langle \xi, p \rangle, \quad p \in M \subset \mathbb{S}^{n-1}, \quad \text{where } \xi = (\xi_i)_{1 \leq i \leq n} \sim N_n(0, I_n).$$

In a neighborhood of the point $p \in \partial M_d$, the points of ∂M_d are assumed to have a parameter representation in the form of $\phi(t) \in \partial M_d$, $t = (t^i)_{1 \leq i \leq d}$. Put $\phi_i = \partial\phi/\partial t^i$, $\phi_{ij} = \partial^2\phi/\partial t^i\partial t^j$, and so on. Moreover, let $\phi(t)$ be a normal coordinate system around p . Then $X_i(p) = \langle \xi, \phi_i \rangle|_p$ and $g_{ij}(p) = \text{Cov}(X_i(p), X_j(p)) = \langle \phi_i, \phi_j \rangle|_p = \delta_{ij}$. Therefore $X(p), X_1(p), \dots, X_d(p)$ are distributed independently according to the standard normal distribution. Moreover, since $X_{ij}(p) = \langle \xi, \phi_{ij} \rangle|_p$, $\Gamma_{ij}^k(p) = \text{Cov}(X_{ij}(p), X_k(p)) = \langle \phi_{ij}, \phi_k \rangle|_p = 0$, and $\langle \phi_{ij}, \phi \rangle|_p = -\delta_{ij}$, we observe that

$$H_{ij}(p) = \langle \xi, -\phi_{ij} - \delta_{ij}\phi \rangle|_p = \langle V(p), -\phi_{ij} \rangle|_p,$$

where

$$V(p) = \xi - \langle \xi, \phi \rangle|_p - \sum_{i=1}^d \langle \xi, \phi_i \rangle|_p \phi_i|_p.$$

$V(p)$ follows the multivariate standard normal distribution in the linear space $(T_p \partial M_d)^\perp$ independently of $X(p)$, $\nabla X|_{\partial M_d}(p) = (X_i(p))_{1 \leq i \leq d}$. Let $\phi_{d+1}, \dots, \phi_{n-1}$ form a basis of the complement subspace of $T_p \partial M_d$ in $T_p \mathbb{S}^{n-1}$. Then, $V(p) = \sum_{i=d+1}^{n-1} \langle \xi, \phi_i \rangle|_p \phi_i|_p$. This implies $V(p) = \nabla_{\partial M_d}^\perp X(p)$.

By letting $r = X(p)$, $s = \|V(p)\|$, and $v = V(p)/\|V(p)\|$, and writing the expectation in an integration form, $E[\chi(A_c)]$ in (3.3) is reduced to the tube method approximation formula (2.5). This is the equivalence between the tube method and the Euler characteristic heuristic.

However, differently from the tube method, the approximation formula by the Euler characteristic heuristic does not contain the dimension n of the ambient manifold. Therefore, it is applicable to Gaussian random fields without a finite KL-expansion.

REMARK 3.5 (The Chern-Gauss-Bonnet theorem). Let M be an m -dimensional closed manifold. Since the excursion set $A_{-\infty}$ with the threshold $c = -\infty$ is just the set of whole indices, (2.11) yields

$$E[\chi(A_{-\infty})] = \chi(M) = \begin{cases} \frac{(-1)^{m/2}}{(2\pi)^{m/2} (m/2)!} \int_M \text{Tr } R^{m/2} dM(p) & (m : \text{even}), \\ 0 & (m : \text{odd}). \end{cases}$$

3.4. Approximation error in the Euler characteristic heuristic. Finally, we briefly introduce the idea of Taylor, et al. [45] to bound the approximation error in the Euler characteristic heuristic for Gaussian random fields.

Recall that the approximation by the Euler characteristic heuristic is stated in terms of random variables, and hence it is independent of the number of random variables forming the Gaussian random field (i.e., the number n of terms in the KL-expansion). In the tube method, by restricting the range of the integral to

$$(3.5) \quad r > \sup_{q \in M \setminus \{p\}} \frac{\langle sv, q \rangle}{1 - \langle p, q \rangle},$$

the integration was completed without overlapping, and accordingly, the distribution of the maximum was obtained exactly. However, $sv \in N_p M$ is an $(n-d)$ -dimensional vector, and in this sense (3.5) contains n . In the following, by rewriting (3.5) into an expression in terms of random variables, we consider its analogue, which is applicable to the case of an infinite KL-expansion. For the sake of simplicity, we treat the case where M is an m -dimensional closed manifold.

We attempt to establish the correspondence with the finite-dimensional case. Let $X(p) = \langle \xi, p \rangle$, $p \in M$. From the consideration in the previous subsection, using the local coordinates $p = \phi(t)$, we have $r = X(p)$, $sv = \nabla_M^\perp X(p) = \xi - \langle \xi, \phi \rangle \phi|_p - \sum_{i=1}^m \langle \xi, \phi_i \rangle \phi_i|_p$, and $\langle sv, q \rangle = X(q) - E[X(q) | X(p)] - E[X(q) | \nabla X(p)]$. Moreover, $\langle p, q \rangle = E[X(p)X(q)]$. Summarizing the above, the random variable expression of (3.5) is written as

$$(3.6) \quad X(p) > \sup_{q \in M \setminus \{p\}} W_p(q), \quad W_p(q) = \frac{X(q) - E[X(q) | X(p), \nabla X(p)]}{1 - E[X(p)X(q)]}.$$

Note that $W_p(q)$ ($\forall q \neq p$) is independent of $X(p)$ because $X(q) - E[X(q) | X(p), \nabla X(p)]$ is independent of $X(p)$ for each q .

Actually, by taking the expectation for (3.3) within the restricted range (3.6), the exact distribution of the maximum is obtained under some additional regularity conditions. The approximation error of the Euler characteristic heuristic is written as follows:

$$(3.7) \quad \begin{aligned} \Delta P(c) &= E[A_c] - P\left(\max_{p \in M} X(p) \geq c\right) \\ &= -\frac{1}{(2\pi)^{m/2}} \int_M E\left[\mathbf{1}_{\{c \leq X(p) \leq \sup_{q \in M \setminus \{p\}} W_p(q)\}} \det(X(p)I_m + H(p))\right] dM(p). \end{aligned}$$

Noting the expansion $\det(X(p)I_m + H(p)) = \sum_{l=0}^m X(p)^{m-l} \text{tr}_l H(p)$ and

$$\mathbf{1}_{\{c \leq X(p) \leq \sup_{q \in M \setminus \{p\}} W_p(q)\}} \leq \mathbf{1}_{\{c \leq X(p)\}} \mathbf{1}_{\{c \leq \sup_{q \in M \setminus \{p\}} W_p(q)\}},$$

as well as the independence of the triplet $X(p)$, $\sup_{q \in M \setminus \{p\}} W_p(q)$, and $H(p)$, the absolute value of the right-hand side is bounded above by

$$\frac{1}{(2\pi)^{m/2}} \int_M E\left[\mathbf{1}_{\{c \leq X(p)\}} X(p)^{m-l}\right] E\left[\mathbf{1}_{\{c \leq \sup_{q \in M \setminus \{p\}} W_p(q)\}} |\text{tr}_l H(p)|\right] dM(p)$$

($l = 0, \dots, m$). Moreover, by Hölder's inequality, we have

$$E\left[\mathbf{1}_{\{c \leq \sup_{q \in M \setminus \{p\}} W_p(q)\}} |\text{tr}_l H(p)|\right] \leq E\left[|\text{tr}_l H(p)|^r\right]^{1/r} P\left(\sup_{q \in M \setminus \{p\}} W_p(q) \geq c\right)^{1/s}$$

($1/r + 1/s = 1$). Applying Borell's inequality ([2]) about the maximum of Gaussian random fields to the zero-mean Gaussian random field $W_p(q)$, $q \in M \setminus \{p\}$, we obtain the asymptotic upper bound

$$\lim_{c \rightarrow \infty} c^{-2} \log P\left(\sup_{q \in M \setminus \{p\}} W_p(q) \geq c\right) \leq -\frac{1}{2\sigma_c^2}, \quad \sigma_c^2 = \sup_{p \in M, q \in M \setminus \{p\}} \text{Var}(W_p(q)),$$

where

$$\text{Var}(W_p(q)) = \frac{1 - \text{Cov}(X(q), X(p))^2 - \text{Cov}(X(q), \nabla X(p)) \text{Var}(\nabla X(p))^{-1} \text{Cov}(\nabla X(p), X(q))}{(1 - \text{Cov}(X(q), X(p)))^2}.$$

In addition, since the marginal distribution of $X(p)$ is the standard normal, we have

$$\lim_{c \rightarrow \infty} c^{-2} \log E\left[1_{\{c \leq X(p)\}} X(p)^{m-l}\right] = -\frac{1}{2}.$$

Combining the above relations and adjusting r and s according to c as

$$r \rightarrow \infty, \quad c^{-2} \log E[|\text{tr}_l H(p)|^r]^{1/r} \rightarrow 0 \quad (c \rightarrow \infty)$$

(note that $s \rightarrow 1$), we can show that

$$(3.8) \quad \limsup_{c \rightarrow \infty} c^{-2} \log |\Delta P(c)| \leq -\frac{1}{2} \left(1 + \frac{1}{\sigma_c^2}\right).$$

In the setting of the finite KL-expansion (3.4), after some simple calculations we see that $\sigma_c^2 = \cot^2 \theta_c$. This is consistent with the exponential order of the approximating error of the tube method given in Theorem 2.10. As a counterpart of the critical radius, σ_c^2 is called the critical variance.

When the random field is non-Gaussian, it is possible to write down the upper bound for the approximating error $|\Delta P(c)|$ corresponding to (3.7). However, it seems difficult to give some asymptotic bounds in general settings such as those in (3.8). The asymptotic error bounds of the Euler characteristic heuristic for the chi-square random fields, the beta random fields, and the inverse chi-square random fields defined by the quadratic forms of real random matrices are given in [31].

4. Applications to statistics

4.1. The maximum of multilinear forms and the multiway analysis of variance. Suppose that each element of the k -way array $\Xi = (\xi_{j_1 \dots j_k})$, $j_i = 1, \dots, q_i$, $i = 1, \dots, k$, is distributed independently according to the standard normal distribution $N(0, 1)$. For $i = 1, \dots, k$, let $h_i = (h_{i1}, \dots, h_{iq_i}) \in \mathbb{S}^{q_i-1}$ be a q_i -dimensional unit coefficient vector, and define a k -linear form with respect to the h_i 's:

$$(4.1) \quad \langle \xi, h_1 \otimes \dots \otimes h_k \rangle = \sum_{j_1=1}^{q_1} \dots \sum_{j_k=1}^{q_k} \xi_{j_1 \dots j_k} h_{1j_1} \dots h_{kj_k}.$$

Here, \otimes denotes the Kronecker product and $\xi = (\xi_{11\dots 1}, \xi_{11\dots 2}, \dots, \xi_{q_1 q_2 \dots q_k})$ is the vector of the lexicographically ordered elements of Ξ .

The maximum of this multilinear form,

$$(4.2) \quad \max_{h_i \in \mathbb{S}^{q_i-1}, \forall i} \langle \xi, h_1 \otimes \dots \otimes h_k \rangle,$$

is an extension of the largest singular value of a matrix. In particular when $k = 2$, the maximum is the largest singular value of the $q_1 \times q_2$ random matrix consisting of independent standard normal random variables, and the distribution is reduced to that of the square root of the largest eigenvalue of the $q_2 \times q_2$ Wishart distribution $W_{q_2}(q_1, I_{q_2})$ with q_1 degrees of freedom.

The distribution of the maximum of the multilinear form (4.2) is required in the statistical testing for the hypothesis of no interaction in a multiway analysis of variance model. For testing the hypothesis of no interaction in the two-way data without replication $\{x_{ij}\}_{I \times J}$, we have to model the interaction. Johnson and Graybill [19] modeled the interaction term as a bilinear form of rank 1 as follows:

$$(4.3) \quad x_{ij} = \alpha_i + \beta_j + \phi u_i v_j + \varepsilon_{ij}, \quad i = 1, \dots, I, \quad j = 1, \dots, J.$$

Here, α_i , β_j , ϕ , u_i , and v_j are unknown parameters, and ε_{ij} is the error term distributed independently according to $N(0, \sigma^2)$. When σ^2 is known, the likelihood ratio test for testing the hypothesis of no interaction $H_0 : \phi = 0$ has the same distribution of the maximum (4.2) with $k = 2$, $q_1 = I - 1$, and $q_2 = J - 1$ under the null hypothesis.

For the three-way data $\{x_{ijk}\}_{I \times J \times K}$ without replication, Kawasaki and Miyakawa [20] proposed the following model as an extension to (4.3):

$$x_{ijk} = (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + \phi u_i v_j w_k + \varepsilon_{ijk}, \\ i = 1, \dots, I, \quad j = 1, \dots, J, \quad k = 1, \dots, K.$$

Under the null hypothesis of no three-factor interaction $H_0 : \phi = 0$, the null distribution of the likelihood ratio test is reduced to the distribution of the maximum (4.2) with $k = 3$, $q_1 = I - 1$, $q_2 = J - 1$, and $q_3 = K - 1$.

The multilinear form (4.1) can be regarded as a Gaussian random field $\langle \xi, h \rangle$, $h \in M$, with

$$M = \{h_1 \otimes \dots \otimes h_k \mid h_i \in \mathbb{S}^{q_i-1}, \quad i = 1, \dots, k\}$$

the index set, and ξ an $n = \prod_{i=1}^k q_i$ dimensional standard normal random vector. M is the direct product of the unit spheres $\mathbb{S}^{q_1-1} \otimes \dots \otimes \mathbb{S}^{q_k-1}$, and is regarded as a subset of \mathbb{S}^{n-1} . Hence, the tube method can be applied to our problem. Note that the dimension of M is $m = \sum_{i=1}^k q_i - k$. Noting that the degree of the map $(h_1, \dots, h_k) \mapsto h_1 \otimes \dots \otimes h_k$ is 2^{k-1} , we have

$$(4.4) \quad \text{Vol}(M) = 2^{-(k-1)} \prod_{i=1}^k \Omega_{q_i}, \quad \Omega_q = \text{Vol}(\mathbb{S}^{q-1}) = \frac{2\pi^{q/2}}{\Gamma(q/2)}.$$

Divide the indices $\{1, \dots, m\}$ into k partitions as follows:

$$A_1 = \{1, \dots, q_1-1\}, \quad A_2 = \{q_1, \dots, q_2-2\}, \quad \dots, \quad A_k = \{q_1 + \dots + q_{k-1} - k + 2, \dots, m\}.$$

When i and j belong to the same partition (i.e., there exists h such that $i, j \in A_h$), write $i \sim j$. When $i \sim j$ does not hold, write $i \not\sim j$. The curvature tensor of M in terms of a normal coordinate system is written as follows:

$$R_{ij;kl} = \begin{cases} -(\delta_{ik}\delta_{jl} - \delta_{il}\delta_{jk}) & (i \sim j), \\ 0 & (\text{otherwise}). \end{cases}$$

Recall that this is the double form of type (2, 2).

THEOREM 4.1 ([29]). *Suppose that a_1, a_2, \dots, a_{2e} are distinct elements of $\{1, 2, \dots, m\}$. We denote the number of their pairings into e pairs such that two elements of a pair do not belong to the same partition, that is, the number of*

$$\{(a_1, a_2), \dots, (a_{2e-1}, a_{2e}) \mid a_1 < a_3 < \dots < a_{2e-1}, \\ a_{2l-1} < a_{2l}, a_{2l-1} \approx a_{2l}, \forall l = 1, \dots, e\}$$

by $n_k(q_1 - 1, \dots, q_k - 1; e)$.

(i) *The tube method approximation formula for the distribution of the maximum of the multilinear form is given by (2.9) with the coefficients*

$$w_{m+1-2e} = \frac{\pi^{(k-1)/2}}{\prod_{i=1}^k \Gamma(q_i/2)} \left(-\frac{1}{2}\right)^e \Gamma\left(\frac{1}{2}(m+1) - e\right) n_k(q_1 - 1, \dots, q_k - 1; e),$$

$e = 0, 1, \dots, [m/2]$.

(ii) *The critical radius of M is given by*

$$\theta_c = \cos^{-1} \sqrt{\frac{2k-2}{3k-2}}.$$

PROOF. Here, we will give a proof for (i) different from [29]. Put $I = (\delta_{i,j}) \in D^{1,1}$, $\tilde{R} = R + I^2/2$. We evaluate $\text{Tr}\tilde{R}^e$ in the sense of the double form. Noting that $(I^2/2)_{ij;kl} = \delta_{ik}\delta_{jl} - \delta_{il}\delta_{jk}$, we have

$$\tilde{R}_{ij;kl} = \begin{cases} \delta_{ik}\delta_{jl} - \delta_{il}\delta_{jk} & (i \approx j), \\ 0 & (\text{otherwise}). \end{cases}$$

Also, since

$$(\tilde{R}^e)_{i_1 \dots i_{2e}; j_1 \dots j_{2e}} = \frac{1}{(2^e)^2} \sum_{\pi, \sigma \in S_{2e}} \text{sgn}(\pi) \text{sgn}(\sigma) \\ \times \tilde{R}_{i_{\pi(1)} i_{\pi(2)}; j_{\sigma(1)} j_{\sigma(2)}} \cdots \tilde{R}_{i_{\pi(2e-1)} i_{\pi(2e)}; j_{\sigma(2e-1)} j_{\sigma(2e)}},$$

we have

$$\begin{aligned} \text{Tr}\tilde{R}^e &= \frac{1}{(2e)!} \sum_{i_1, \dots, i_{2e}} \frac{1}{(2^e)^2} \sum_{\pi \in S_{2e}} \sum_{\sigma \in S_{2e}} \text{sgn}(\pi) \text{sgn}(\sigma) \\ &\quad \times \tilde{R}_{i_{\pi(1)} i_{\pi(2)}; i_{\sigma(1)} i_{\sigma(2)}} \cdots \tilde{R}_{i_{\pi(2e-1)} i_{\pi(2e)}; i_{\sigma(2e-1)} i_{\sigma(2e)}} \\ &= \sum_{i_1 < \dots < i_{2e}} \sum_{\pi \in S_{2e}^*} \sum_{\sigma \in S_{2e}^*} \text{sgn}(\pi) \text{sgn}(\sigma) \\ &\quad \times \tilde{R}_{i_{\pi(1)} i_{\pi(2)}; i_{\sigma(1)} i_{\sigma(2)}} \cdots \tilde{R}_{i_{\pi(2e-1)} i_{\pi(2e)}; i_{\sigma(2e-1)} i_{\sigma(2e)}} \\ &\quad (S_{2e}^* = \{\pi \in S_{2e} \mid \pi(2l-1) < \pi(2l), l = 1, \dots, e\}) \\ &= \sum_{i_1 < \dots < i_{2e}} \sum_{\pi \in S_{2e}^*} \sum_{\sigma \in S_{2e}^*} \text{sgn}(\pi) \text{sgn}(\sigma) \delta_{\pi(1)\sigma(1)} \delta_{\pi(2)\sigma(2)} \mathbf{1}_{\{\pi(1) \approx \pi(2)\}} \\ &\quad \cdots \delta_{\pi(2e-1)\sigma(2e-1)} \delta_{\pi(2e)\sigma(2e)} \mathbf{1}_{\{\pi(2e-1) \approx \pi(2e)\}} \\ &= \sum_{i_1 < \dots < i_{2e}} \sum_{\pi \in S_{2e}^*} \mathbf{1}_{\{\pi(1) \approx \pi(2)\}} \cdots \mathbf{1}_{\{\pi(2e-1) \approx \pi(2e)\}} \\ &= e! n_k(q_1 - 1, \dots, q_k - 1; e). \end{aligned}$$

The proof is completed by substituting this and (4.4) into (2.10). \square

REMARK 4.2. When $k = 2$,

$$n_2(q_1 - 1, q_2 - 1; e) = \frac{(q_1 - 1)!(q_2 - 1)!}{(q_1 - 1 - e)!(q_2 - 1 - e)!}.$$

$n_k(d_1, \dots, d_k; e)$ for $k \geq 3$ can be evaluated by recurrence relations ([29], Lemma A.2).

4.2. Significance in the projection pursuit. Suppose that for N individuals, p -dimensional vector data $X_i \in \mathbb{R}^p$ ($i = 1, \dots, N$) are observed as independent samples. In such multidimensional data analysis (multivariate analysis), the projections of the data into lower-dimensional subspaces are often used in order to focus on the features that interest the data analyst.

In the principal component analysis or the canonical correlation analysis, subspaces are selected so that the variance of the projected data is large. In the projection pursuit, the subspaces that comprise large nonnormality (that is, the cumulants with degrees 3 or higher are large) are selected ([15]). Fast ICA, one of the independent component analysis techniques, is a similar method ([16]). The procedure is described as follows. Here, we restrict our attention to projection to a one-dimensional subspace.

(i) Let $h \in \mathbb{S}^{p-1}$ ($= \{h \in \mathbb{R}^p \mid \|h\| = 1\}$), be a directional vector. Make an orthogonal projection of the data onto the one-dimensional space with respect to the direction h as $Y_i = \langle h, X_i \rangle$ ($i = 1, \dots, N$). Moreover, standardize the data as $Z_i = (Y_i - \bar{Y})/s_Y$ ($i = 1, \dots, N$) with the sample mean \bar{Y} and the sample standard deviation s_Y .

(ii) Based on the standardized data, calculate the measure of nonnormality $I_N(h)$ (projection pursuit index). Here, as an example, we let

$$I_N(h) = \frac{1}{\sqrt{N}} \sum_{i=1}^N \left(e^{\theta Z_i - \theta^2/2} - 1 \right) / \sqrt{e^{\theta^2} - 1 - \theta^2 - \frac{\theta^4}{2}},$$

where $\theta \neq 0$ is a constant. Note that if Z_i follows the standard normal distribution, then $I_N(h) \sim N(0, 1)$ follows asymptotically for each $h \in \mathbb{S}^{p-1}$.

(iii) Find the direction $h^* = \arg \max_{h \in \mathbb{S}^{p-1}} I_N(h)$ that attains the maximum of the projection pursuit index numerically.

If the distribution of X_i is a multivariate Gaussian, then the joint distribution of Z_i ($i = 1, \dots, N$) is free from h , and hence the marginal distribution of $I_N(h)$ is irrelevant to h . However, because of stochastic fluctuations, a random continuous function $I_N(h)$ has the point h^* that attains its maximum. Therefore, it is important to assess whether the numerically searched direction h^* is caused by stochastic fluctuations and is not really meaningful. For the assessment, we can use the framework of the statistical testing hypotheses. Assuming the null hypothesis H_0 : “the distribution of X_i is the multivariate normal,” write the upper probability of the maximum of $I_N(h)$ under H_0 as

$$\bar{F}_N(c) = P\left(\max_{h \in \mathbb{S}^{p-1}} I_N(h) \geq c \mid H_0 \right).$$

We can assess the significance of the maximum by checking that the maximum value $I_N(h^*)$ is located in the upper tail of the distribution, that is, by ensuring that

$$\bar{F}_N(I_N(h^*)) \leq \alpha,$$

where α ($= 0.05$, for example) is the significance level of the test.

However, the distribution of $\max_{h \in \mathbb{S}^{p-1}} I_N(h)$ is too complicated to handle analytically. Thus, we adopt the asymptotic approximation as the sample size N goes to infinity. The following theorem follows from the standard central limit theorem for Banach-valued random variables ([6], [17]).

THEOREM 4.3. *The Banach space on \mathbb{S}^{p-1} is denoted by $C(\mathbb{S}^{p-1})$. Let X_i , $i = 1, 2, \dots$, be i.i.d. samples taken from the p -dimensional multivariate normal distribution $N_p(\mu, \Sigma)$ ($\det \Sigma > 0$). As $N \rightarrow \infty$, $I_N(\cdot)$ converges to a Gaussian random field $I(\cdot)$ in distribution on $C(\mathbb{S}^{p-1})$. Here, $E[I(h)] = 0$, $\text{Var}[I(h)] = 1$, and*

$$\text{Cov}(I(g), I(h)) = r(g, h) = \frac{e^{\theta^2 \langle g, h \rangle} - 1 - \theta^2 \langle g, h \rangle - \theta^4 \langle g, h \rangle^2 / 2}{e^{\theta^2} - 1 - \theta^2 - \theta^4 / 2}.$$

From the continuous mapping theorem,

$$\bar{F}_N(c) \rightarrow \bar{F}(c) = P\left(\max_{h \in \mathbb{S}^{p-1}} I(h) \geq c \mid H_0\right) \quad (N \rightarrow \infty)$$

holds. Thus, we have an approximation $\bar{F}_N(c) \approx \bar{F}(c)$ when N is large. $\bar{F}(c)$ is evaluated by the tube method and the Euler characteristic heuristic.

The idea stated here that the significance of the projection pursuit is measured by the tube method is due to Sun [38]. In [38], an approximation using the first two terms of (2.9) is proposed. In the case of the example treated here, we can derive all terms of (2.9) explicitly.

The random field $I(\cdot)$ is a spherically isotropic field with the covariance function depending only on the inner product $\langle g, h \rangle$ of two points $g, h \in \mathbb{S}^{p-1}$. Therefore, we only have to derive the metric and the curvature at an arbitrary point. Taking local coordinates around $(0, \dots, 0, 1)$, $g = (s^1, \dots, s^{p-1}, \sqrt{1 - \sum (s^i)^2})$ and $h = (t^1, \dots, t^{p-1}, \sqrt{1 - \sum (t^i)^2})$, we have $g_{ij} = \partial^2 r(g, h) / \partial s^i \partial t^j |_{s=t=0} = \rho(\theta) \delta_{ij}$ and $R_{ij;kl} = -\rho(\theta)(\delta_{ik}\delta_{jl} - \delta_{il}\delta_{jk})$, where

$$\rho(\theta) = \frac{(e^{\theta^2} - 1 - \theta^2) \theta^2}{e^{\theta^2} - 1 - \theta^2 - \theta^4 / 2}.$$

Noting that the dimension of the index set is $m = p - 1$, the volume of the index set is calculated as $\text{Vol}(M) = \rho(\theta)^{(p-1)/2} \Omega_p$. In addition, the curvature tensor is represented in terms of a normal coordinate system as $R_{ij;kl} = -\rho(\theta)^{-1}(\delta_{ik}\delta_{jl} - \delta_{il}\delta_{jk})$, that is, $R = (-1/2\rho(\theta))I^2$ as a double form. Substituting $\text{Vol}(M)$ and

$$\text{Tr} R^l = \frac{(-1)^l (p-1)!}{2^l \rho(\theta)^l (p-1-2l)!}$$

into (2.11), we have the Euler characteristic approximation for $\bar{F}(c)$.

THEOREM 4.4.

$$(4.5) \quad \hat{P}\left(\max_{h \in \mathbb{S}^{p-1}} I(h) \geq c\right) = \frac{\rho(\theta)^{(p-1)/2}}{2^{p/2-1} \Gamma(p/2)} \sum_{l=0}^{\lfloor (p-1)/2 \rfloor} \frac{(p-1)!}{\rho(\theta)^l 2^l l! (p-1-2l)!} \times e^{-c^2/2} H_{p-2l-2}(c).$$

In [44], Section 6.3, an equivalent expression to (4.5) is given in another setting as the spherically isotropic random field explained here.

References

1. Adler, R. J. (1981). *The Geometry of Random Fields*. John Wiley and sons, Chichester. MR611857 (82h:60103)
2. Adler, R. J. (1990). *An Introduction to Continuity, Extrema and Related Topics for General Gaussian Processes*. Institute of Mathematical Statistics, Lecture Notes–Monograph Series, Vol. 12, IMS, Hayward, CA. MR1088478 (92g:60053)
3. Adler, R. J. (2000). On excursion sets, tube formulas and maxima of random fields. *Ann. Appl. Probab.*, **10**, 1–74. MR1765203 (2001g:60082)
4. Adler, R. J. and Hasofer, A. M. (1976). Level crossings for random fields. *Ann. Probab.*, **4**, 1–12. MR0405559 (53:9352)
5. Adler, R. J. and Taylor, J. E. (2007). *Random Fields and Geometry*. Springer, New York. MR2319516 (2008m:60090)
6. Araujo, A. and Giné, E. (1980). *The Central Limit Theorem for Real and Banach Valued Random Variables*. Wiley, New York. MR576407 (83e:60003)
7. Azaïs, J.-M. and Wschebor, M. (2005). On the distribution of the maximum of a Gaussian field with d parameters. *Ann. Appl. Probab.*, **15**, 254–278. MR2115043 (2005k:60117)
8. Dacunha-Castelle, D. and Gassiat, E. (1997). Testing in locally conic models, and application to mixture models. *ESAIM Probability and Statistics*, **1**, 285–317. MR1468112 (98h:62024)
9. Federer, H. (1959). Curvature measures. *Trans. Amer. Math. Soc.*, **93**, 418–491. MR0110078 (22:961)
10. Fu, J. H. G. (1989). Curvature measures and generalized Morse theory. *J. Differential Geometry*, **30**, 619–642. MR1021369 (90j:53094)
11. Fukumizu, K., Kuriki, S., Takeuchi, K., and Akahira, M. (2004). *Statistical Theory of Singular Models* (in Japanese), Iwanami, Tokyo.
12. Gray, A. (2004). *Tubes*, 2nd ed. Birkhäuser, Boston. MR2024928 (2004j:53001)
13. Hasofer, A. M. (1978). Upcrossings of random fields. *Adv. Appl. Probab., Supplement*, **10**, 14–21. MR0494452 (58:13313)
14. Hotelling, H. (1939). Tubes and spheres in n -spaces, and a class of statistical problems. *Amer. J. Math.*, **61**, 440–460. MR1507387
15. Huber, P. J. (1985). Projection pursuit. *Ann. Statist.*, **13**, 435–525. MR790553 (88b:62118)
16. Hyvärinen, A., Karhunen, J. and Oja, E. (2001). *Independent Component Analysis*. Wiley-Interscience, New York.
17. Jain, N. C. and Marcus, M. B. (1975). Central limit theorems for $C(S)$ -valued random variables. *J. Functional Anal.*, **19**, 216–231. MR0385994 (52:6853)
18. Johansen, S. and Johnstone, I. (1990). Hotelling’s theorem on the volume of tubes: Some illustrations in simultaneous inference and data analysis. *Ann. Statist.*, **18**, 652–684. MR1056331 (92d:62018)
19. Johnson, D. E. and Graybill, F. A. (1972). An analysis of a two-way model with interaction and no replication. *J. Amer. Statist. Assoc.*, **67**, 862–868. MR0400566 (53:4397)
20. Kawasaki, H. and Miyakawa, M. (1996). A test of three-factor interaction in a three-way layout without replication (in Japanese). *Hinshitu (Quality)*, **26**, 97–108.
21. Kedem, B. (1994). *Time Series Analysis by Higher Order Crossings*. IEEE Press, New York. MR1261636 (96a:62092)
22. Knowles, M. and Siegmund, D. (1989). On Hotelling’s approach to testing for a nonlinear parameter in regression. *Internat. Statist. Rev.*, **57**, 205–220.
23. Kobayashi, S. (1999). Differential geometry in Japan in the 1940s and 1950s. *Sūgaku Expositions*, **12**, 85–93. MR1692106
24. Kuriki, S. (2005). Asymptotic distribution of inequality restricted canonical correlation with application to tests for independence in ordered contingency tables. *J. Multivariate Anal.*, **94**, 420–449. MR2167923 (2006h:62060)
25. Kuriki, S. (2008). Statistical models and adjusting for the multiplicity of tests in the QTL analysis (in Japanese). In *Statistical Science in the 21 Century*, Vol. 2, Konishi, S. and Kunitomo, N. eds., Univ. Tokyo Press, Tokyo, 315–356.
26. Kuriki, S. and Takemura, A. (1999). Distribution of the maximum of Gaussian random field: Tube method and Euler characteristic method (in Japanese). *Proc. Inst. Statist. Math.*, **47**, 201–221. MR1720026 (2000h:60044)

27. Kuriki, S. and Takemura, A. (2000). Some geometry of the cone of nonnegative definite matrices and weights of associated $\bar{\chi}^2$ distribution. *Ann. Inst. Statist. Math.*, **52**, 1–14. MR1771476 (2001d:62060)
28. Kuriki, S. and Takemura, A. (2000). Shrinkage estimation towards a closed convex set with a smooth boundary. *J. Multivariate Anal.*, **75**, 79–111. MR1787403 (2001h:62034)
29. Kuriki, S. and Takemura, A. (2001). Tail probabilities of the maxima of multilinear forms and their applications. *Ann. Statist.*, **29**, 328–371. MR1863962 (2002h:62163)
30. Kuriki, S. and Takemura, A. (2002). Application of tube formula to distributional problems in multiway layouts. *Appl. Stoch. Models Bus. Ind.*, **18**, 245–257. MR1932640 (2003h:62030)
31. Kuriki, S. and Takemura, A. (2008). Euler characteristic heuristic for approximating the distribution of the largest eigenvalue of an orthogonally invariant random matrix. *J. Statist. Plann. Inference*, **138**, 3357–3378. MR2450081
32. Lin, Y. and Lindsay, B. G. (1997). Projections on cones, chi-bar squared distributions, and Weyl’s formula. *Statist. Probab. Lett.*, **32**, 367–376. MR1602207 (98m:62035)
33. Morse, M. and Cairns, S. S. (1969). *Critical Point Theory in Global Analysis and Differential Topology*. Academic Press, New York. MR0245046 (39:6358)
34. Naiman, D. Q. (1990). Volumes of tubular neighborhoods of spherical polyhedra and statistical inference. *Ann. Statist.*, **18**, 685–716. MR1056332 (91k:60024)
35. Naiman, D. Q. and Wynn, H. P. (1992). Inclusion-exclusion-Bonferroni identities and inequalities for discrete tube-like problems via Euler characteristics. *Ann. Statist.*, **20**, 43–76. MR1150334 (93b:60035)
36. Naiman, D. Q. and Wynn, H. P. (1997). Abstract tubes, improved inclusion-exclusion identities and inequalities and importance sampling. *Ann. Statist.*, **25**, 1954–1983. MR1474076 (99h:60034)
37. Ninomiya, Y. (2004). Construction of conservative test for change-point problem in two-dimensional random fields. *J. Multivariate Anal.*, **89**, 219–242. MR2063632 (2005a:62143)
38. Sun, J. (1991). Significance levels in exploratory projection pursuit. *Biometrika*, **78**, 759–769. MR1147012 (93k:62128)
39. Sun, J. (1993). Tail probabilities of the maxima of Gaussian random fields. *Ann. Probab.*, **21**, 34–71. MR1207215 (94e:60048)
40. Sun, J. and Loader, C. R. (1994). Simultaneous confidence bands for linear regression and smoothing. *Ann. Statist.*, **22**, 1328–1345. MR1311978 (95j:62022)
41. Takemura, A. and Kuriki, S. (1997). Weights of $\bar{\chi}^2$ distribution for smooth or piecewise smooth cone alternatives. *Ann. Statist.*, **25**, 2368–2387. MR1604465 (99c:62156)
42. Takemura, A. and Kuriki, S. (2002). Maximum of Gaussian field on piecewise smooth domain: Equivalence of tube method and Euler characteristic method. *Ann. Appl. Probab.*, **12**, 768–796. MR1910648 (2003d:60102)
43. Takemura, A. and Kuriki, S. (2003). Tail probability via the tube formula when the critical radius is zero. *Bernoulli*, **9**, 535–558. MR1997496 (2005g:60027)
44. Taylor, J. E. and Adler, R. (2003). Euler characteristics for Gaussian fields on manifolds. *Ann. Probab.*, **31**, 533–563. MR1964940 (2005g:58072)
45. Taylor, J. E., Takemura, A. and Adler, R. (2005). Validity of the expected Euler characteristic heuristic. *Ann. Probab.*, **33**, 1362–1396. MR2150192 (2006b:60073)
46. Weyl, H. (1939). On the volume of tubes. *Amer. J. Math.*, **61**, 461–472. MR1507388
47. Worsley, K. J. (1994). Local maxima and the expected Euler characteristic of excursion sets of χ^2 , F and t fields. *Adv. Appl. Probab.*, **26**, 13–42. MR1260300 (94i:60064)
48. Worsley, K. J. (1995). Estimating the number of peaks in a random field using the Hadwiger characteristic of excursion sets, with applications to medical images. *Ann. Statist.*, **23**, 640–669. MR1332586 (96b:60132)
49. Worsley, K. J. (1995). Boundary corrections for the expected Euler characteristic of excursion sets of random fields, with an application to astrophysics. *Adv. Appl. Probab.*, **27**, 943–959. MR1358902 (97b:60068)

INSTITUTE OF STATISTICAL MATHEMATICS, ROIS, 4-6-7 MINAMI-AZABU, MINATO-KU, TOKYO
106-8569, JAPAN

E-mail address: `kuriki@ism.ac.jp`

DEPARTMENT OF MATHEMATICAL INFORMATICS, UNIVERSITY OF TOKYO, 7-3-1 HONGO,
BUNKYO-KU, TOKYO 113-0033, JAPAN

E-mail address: `takemura@stat.t.u-tokyo.ac.jp`

Translated by SATOSHI KURKI and AKIMICHI TAKEMURA