Available online at www.sciencedirect.com

ScienceDirect

Stochastic Processes and their Applications 133 (2021) 166–192

stochastic processes and their applications

www.elsevier.com/locate/spa

Extremes of locally stationary Gaussian and chi fields on manifolds

Wanli Qiao

Department of Statistics, George Mason University, 4400 University Drive, MS 4A7, Fairfax, VA 22030, USA

Received 9 January 2020; received in revised form 23 November 2020; accepted 24 November 2020
Available online 3 December 2020

Abstract

Depending on a parameter $h \in (0, 1]$, let $\{X_h(\boldsymbol{t}), \boldsymbol{t} \in \mathcal{M}_h\}$ be a class of centered Gaussian fields indexed by compact manifolds $\mathcal{M}_h$ with positive reach. For locally stationary Gaussian fields X_h, we study the asymptotic excursion probabilities of X_h on $\mathcal{M}_h$. Two cases are considered: (i) h is fixed and (ii) $h \to 0$. These results are also extended to obtain the limit behaviors of the extremes of locally stationary χ-fields on manifolds.

MSC: primary 60G70; 60G15

Keywords: Local stationarity; Excursion probabilities; Gaussian fields; Chi-fields; Voronoi diagrams; Positive reach

1. Introduction

We study the following two related problems in this paper.
(i) Let $\{X(\boldsymbol{t}), \boldsymbol{t} \in \mathcal{M}\}$ be a centered Gaussian field indexed on a compact submanifold $\mathcal{M}$ of $\mathbb{R}^n$. For $X(\boldsymbol{t})$ satisfying a local stationarity condition (see Definition 2.1), we derive the asymptotic form of the excursion probability

$$\mathbb{P}\left(\sup_{\boldsymbol{t}\in\mathcal{M}} X(\boldsymbol{t}) > u\right), \text{ as } u \to \infty. \tag{1.1}$$

(ii) Let $\{X_h(\boldsymbol{t}), \boldsymbol{t} \in \mathcal{M}_h\}_{h\in(0,1]}$ be a class of centered Gaussian fields, where $\mathcal{M}_h$ is a compact submanifold of $\mathbb{R}^n$ for each $h \in (0, 1]$. Suppose that we have the structure $\mathcal{M}_h = \mathcal{M}_{h,1} \times \mathcal{M}_{h,2}$ such that $\boldsymbol{t} = (\boldsymbol{t}_{(1)}^T, \boldsymbol{t}_{(2)}^T)^T \in \mathcal{M}_h$ means $\boldsymbol{t}_{(1)} \in \mathcal{M}_{h,1}$ and $\boldsymbol{t}_{(2)} \in \mathcal{M}_{h,2}$, where we allow $\mathcal{M}_{h,2}$ to be a null set. The Gaussian fields $X_h(\boldsymbol{t})$ we consider have a rescaled form $X_h(\boldsymbol{t}) =$

E-mail address: wqiao@gmu.edu.

https://doi.org/10.1016/j.spa.2020.11.006
0304-4149/

$\overline{X}_h(\boldsymbol{t}_{(1)}/h, \boldsymbol{t}_{(2)}), \boldsymbol{t} \in \mathcal{M}_h$ for some $\overline{X}_h$ satisfying the local stationarity condition. We derive the following limit result

$$\lim_{h\to 0} \mathbb{P}\left(a_h\left(\sup_{\boldsymbol{t}\in\mathcal{M}_h} X_h(\boldsymbol{t}) - b_h\right) \leq z\right) = e^{-e^{-z}}, \tag{1.2}$$

for some $a_h, b_h \in \mathbb{R}_+$ and any fixed $z \in \mathbb{R}$.

While there is a large amount of literature on excursion probabilities of Gaussian processes or fields (see, e.g., Adler and Taylor [1], and Azaïs and Wschebor [3]), most of the existing work only considers index sets $\mathcal{M}$ (or $\mathcal{M}_h$) of dimension n (the same as the ambient Euclidean space), while we focus on Gaussian fields indexed by manifolds that can be low-dimensional.

For problem (i), some relevant results can be found in Mikhaleva and Piterbarg [30], Piterbarg and Stamatovich [35], and Cheng [12]. Compared with these works, the framework of our result is more general in the following aspects: First of all, Cheng [12] studies the excursion probabilities of *locally isotropic* Gaussian random fields on manifolds, where local isotropy means the variance between two local points only depends on their (geodesic) distance, while we consider *locally stationary* Gaussian fields, for which not only the distance between the points but also their locations are involved in the variance. Furthermore, in Mikhaleva and Piterbarg [30] and Piterbarg and Stamatovich [35], the Gaussian fields are assumed to be indexed by $\mathbb{R}^n$, while we only require the index sets to be the manifolds. As pointed out in Cheng [12], it is not clear whether one can always find a Gaussian field indexed by $\mathbb{R}^n$ whose restriction on $\mathcal{M}$ is $X(\boldsymbol{t})$. Also see Cheng and Xiao [13] for some further arguments on this point. In addition, all the above works assume that the manifolds are smooth (C^∞), while we consider a much larger class of manifolds (only satisfying a *positive reach* condition). In fact, the properties of positive reach play a critical role in the geometric construction in our proofs.

For problem (ii), the study in Qiao and Polonik [37] corresponds to a special case of (1.2) when $\mathcal{M}_h \equiv \mathcal{M}$ for some manifold $\mathcal{M}$ independent of h, and $\mathcal{M}_{h,2} = \emptyset$. They use some ideas from Mikhaleva and Piterbarg [30] and also assume that X_h is indexed by a neighborhood of higher dimensions around $\mathcal{M}$, while we only need X_h to be indexed by the manifolds $\mathcal{M}_h$, by making use of the result developed for problem (i). This weaker requirement for the Gaussian fields finds broader applications when the Gaussian fields are observable or can be approximated only on low-dimensional manifolds. See (1.7) for example. At a more technical level, we use Voronoi diagrams to construct partitions to the index sets, as one of the major building blocks in the proof to utilize the classical double-sum method (Pickands [32]). See Sections 4.1 and 4.3 . This strategy is different from what is used in Qiao and Polonik [37], where they adopt Delaunay triangulations for the partitions. When extended from $\mathbb{R}^n$ to low-dimensional submanifolds, the construction of Delaunay triangulations becomes nontrivial and needs a particular algorithm. See Chapter 7 of Boissonnat, Chazal and Yvinec [9]. By contrast, the construction of Voronoi diagrams on manifolds is straightforward. We expect that the approach based on Voronoi diagrams can be generalized to study the extreme value distributions of Gaussian fields indexed by more sophisticated sets such as stratified spaces (i.e. sets with manifolds of different dimensions glued together). Furthermore, by using the assumed structure of $\mathcal{M}_h$, only rescaling the parameter $\boldsymbol{t}_{(1)}$ allows us to apply (1.2) to get asymptotic extreme value distributions of χ-fields on manifolds, which in fact is one of the motivations of this work, as described below.

Let $\{\boldsymbol{X}(\boldsymbol{s}), \boldsymbol{s} \in \mathcal{M}\}$ be a p-dimensional Gaussian vector field, where $\boldsymbol{X} = (X_1, \ldots, X_p)^T$ has zero mean and identity variance–covariance matrix. Note that we have suppressed the possible dependence of $\boldsymbol{X}$ and $\mathcal{M}$ on h. Define

$$\chi(\boldsymbol{s}) = [X_1^2(\boldsymbol{s}) + \cdots + X_p^2(\boldsymbol{s})]^{1/2}, \ \boldsymbol{s} \in \mathcal{M}, \tag{1.3}$$

which is called a χ-field, where we allow the components $X_i(\boldsymbol{s}_i)$ and $X_j(\boldsymbol{s}_j)$ to be dependent, if $\boldsymbol{s}_i \neq \boldsymbol{s}_j$. Let $\mathbb{S}^{p-1} = \{x \in \mathbb{R}^p : \ \|x\| = 1\}$ be the $(p-1)$-dimensional unit sphere. Using the property of Euclidean norm, we have

$$\sup_{\boldsymbol{s}\in\mathcal{M}} \chi(\boldsymbol{s}) = \sup_{\boldsymbol{s}\in\mathcal{M},\boldsymbol{v}\in\mathbb{S}^{p-1}} Y_h(\boldsymbol{s}, \boldsymbol{v}), \tag{1.4}$$

where $\boldsymbol{v} = (v_1, \ldots, v_p) \in \mathbb{R}^p$ and

$$Y(\boldsymbol{s}, \boldsymbol{v}) = X_1(\boldsymbol{s})v_1 + \cdots + X_p(\boldsymbol{s})v_p, \ \boldsymbol{s} \times \boldsymbol{v} \in \mathcal{M} \times \mathbb{S}^{p-1}.$$

Note that $Y(\boldsymbol{s}, \boldsymbol{v})$ is a zero-mean and unit-variance Gaussian field on $\mathcal{M} \times \mathbb{S}^{p-1}$. Using the relation in (1.4) and by applying the results in (1.1) and (1.2), we can study the asymptotic excursion probabilities of $\sup_{\boldsymbol{s}\in\mathcal{M}} \chi(\boldsymbol{s})$, and obtain a result in the form of

$$\lim_{h\to 0} \mathbb{P}\left(a_h \left(\sup_{\boldsymbol{s}\in\mathcal{M}} \chi(\boldsymbol{s}/h) - b_h \right) \leq z \right) = e^{-e^{-z}}. \tag{1.5}$$

The result in (1.5) (see Corollary 3.1) has the following two interesting applications. We consider a vector-valued signal plus noise model

$$\widehat{\boldsymbol{f}}_h(\boldsymbol{s}) = \boldsymbol{f}(\boldsymbol{s}) + \boldsymbol{X}(\boldsymbol{s}/h), \ \boldsymbol{s} \in \mathcal{M}, \tag{1.6}$$

where $\boldsymbol{f}(\boldsymbol{s})$ is a p-dimensional signal, $\boldsymbol{X}(\boldsymbol{s})$ is the noise modeled by the Gaussian vector field considered above. We assume that only $\widehat{\boldsymbol{f}}_h(\boldsymbol{s})$ is directly observable. Given $\alpha \in (0, 1)$, let z_α be such that $\exp(-\exp(-z_\alpha)) = 1 - \alpha$.

(a) Suppose that $\mathcal{M}$ is known, and the inference for the signal $\boldsymbol{f}(\boldsymbol{s})$ is of interest. We have the following asymptotic $(1-\alpha)$ confidence tube for $\boldsymbol{f}(\boldsymbol{s})$:

$$\mathcal{G}_h(\boldsymbol{s}) := \left\{ \boldsymbol{g} \in \mathbb{R}^p : \ a_h \left(\|\widehat{\boldsymbol{f}}_h(\boldsymbol{s}) - \boldsymbol{g}\| - b_h \right) \leq z_\alpha \right\}, \ \boldsymbol{s} \in \mathcal{M}. \tag{1.7}$$

In other words, $\mathbb{P}(\boldsymbol{f}(\boldsymbol{s}) \in \mathcal{G}_h(\boldsymbol{s}), \ \forall \boldsymbol{s} \in \mathcal{M}) \to 1 - \alpha$, as $h \to 0$.

(b) Suppose that the manifold $\mathcal{M}$ is unknown but implicitly defined by $\mathcal{M} = \{\boldsymbol{s} \in \mathcal{A} : \boldsymbol{f}(\boldsymbol{s}) = \boldsymbol{g}_0\}$, where $\mathcal{A} \subset \mathbb{R}^n$ is a known neighborhood of $\mathcal{M}$ (say, a unit cube), and $\boldsymbol{g}_0$ is a known p-dimensional vector so that $\mathcal{M}$ is the intersection of multiple level sets. Suppose that $\widehat{\boldsymbol{f}}_h(\boldsymbol{s})$ is observable on $\mathcal{A}$, and the inference for the manifold $\mathcal{M}$ is of interest. We have the following asymptotic $(1-\alpha)$ confidence region for $\mathcal{M}$:

$$\mathcal{F}_h := \left\{ \boldsymbol{s} \in \mathcal{A} : \ a_h \left(\|\widehat{\boldsymbol{f}}_h(\boldsymbol{s}) - \boldsymbol{g}_0\| - b_h \right) \leq z_\alpha \right\}. \tag{1.8}$$

That is, $\mathbb{P}(\mathcal{M} \subset \mathcal{F}_h) \to 1 - \alpha$, as $h \to 0$. See Remark 3.3(b) for more details.

In statistics the suprema of empirical processes can be approximated by the suprema of Gaussian processes or fields under regularity assumptions (see Chernozhukov et al. [14]). Applying results in (a) and (b) to the approximating Gaussian fields, one can study the statistical inference for a large class of objects including functions and geometric features (low-dimensional manifolds). In a form similar to (1.7), confidence bands for density functions are given in Bickel and Rosenblatt [7] and Rosenblatt [39]. Similar work for regression functions can be found in Konakov and Piterbarg [22]. We note that in these examples the study of the suprema of the approximating Gaussian processes or fields focuses on $\mathcal{M}$ being compact intervals or hypercubes. We expect that our result (1.7) is useful in studying functions supported on more general (low-dimensional) manifolds, especially in the context of *manifold* *learning*, which usually assumes that data lie on low-dimensional manifolds embedded in high-dimensional space. The result (1.8) is useful to infer the location of the manifolds. In fact, the results proved in this work provide the probabilistic foundation to our companion work Qiao [36], where the confidence regions for density ridges are obtained. Ridges are low-dimensional geometric features (manifolds) that generalize the concepts of local modes, and have been applied to model filamentary structures such as the Cosmic Web and road systems. See Qiao and Polonik [38] for a similar application for the construction of confidence regions for level sets.

The study of the asymptotic extreme value behaviors of χ-processes and fields has drawn quite some interest recently. To our best knowledge, the study in the existing literature has only focused on χ-processes and fields indexed by intervals or hypercubes, but not low-dimensional manifolds. See, for example, Albin et al. [2], Bai [4], Hashorva and Ji [20], Ji et al. [21], Konstantinides et al. [23], Lindgren [26], Ling and Tan [27], Liu and Ji [28,29], Piterbarg [33,34], Tan and Hashorva [41,42], Tan and Wu [43]. Also it is worth mentioning that it is often assumed that $X_1, \ldots, X_r$ are independent copies of a Gaussian process or field X in the literature, while the cross-dependence among $X_1, \ldots, X_r$ is allowed under certain constraints in this work. The cross-dependence structures of multivariate random fields have been important objects to study in multivariate geostatistics (see Genton and Kleiber [18]). Also see Zhou and Xiao [44] for the study of the excursion probability of a bivariate Gaussian random field over $\mathbb{R}^n$ with cross-dependence.

The paper is organized as follows: In Section 2 we introduce the concepts that we use in this paper to characterize the manifolds (positive reach) and the Gaussian fields (local stationarity). Then the result for (1.1) (called the unscaled case) is formulated in Theorem 2.1, As an application, a similar result for the χ-fields in presented in Corollary 2.2. In Section 3 we give the result (1.2) (called the rescaled case) in Theorem 3.1 and its χ-fields extension in Corollary 3.1. All the proofs are presented in Section 4, and the Appendix contains some miscellaneous results used in the paper, as well as a collection of concepts and facts related to manifolds and geometric integration theory.

2. Extremes of unscaled Gaussian and χ fields on manifolds

We consider a centered Gaussian field $X(\boldsymbol{t}),\ \ \boldsymbol{t} \in \mathcal{M}$, where $\mathcal{M}$ is an r-dimensional submanifold of $\mathbb{R}^n$ ($1 \le r \le n$). Let $r_X(\boldsymbol{t}_1, \boldsymbol{t}_2) = \mathrm{Cov}(X(\boldsymbol{t}_1), X(\boldsymbol{t}_2))$ for any $\boldsymbol{t}_1, \boldsymbol{t}_2 \in \mathcal{M}$. We first review some existing concepts in the literature that we need to characterize the covariance r_X of the Gaussian field X and the manifold $\mathcal{M}$.

For a positive integer $k \le n$, let $E = \{e_1, \ldots, e_k\}$ be a collection of positive integers such that $n = e_1 + \cdots + e_k$, and let $\boldsymbol{\alpha} = \{\alpha_1, \ldots, \alpha_k\}$ be a collection of positive numbers. Then the pair $(E, \boldsymbol{\alpha})$ is called a structure. Let $\|\cdot\|$ denote the Euclidean norm. Denote $E(0) = 0$ and $E(i) = e_1 + \cdots + e_i$, $i = 1, \ldots, k$. For any $\boldsymbol{t} = (t_1, \ldots, t_n)^T \in \mathbb{R}^n$, its structure module is denoted by $|\boldsymbol{t}|_{E,\alpha} = \sum_{i=1}^{k} \|\boldsymbol{t}_{(i)}\|^{\alpha_i}$, where $\boldsymbol{t}_{(i)} = (t_{E(i-1)+1}, \ldots, t_{E(i)})^T$. This notation has been used, e.g., in Chapter 2 of Piterbarg [34].

Suppose that $\alpha_i \le 2, i = 1, \ldots, k$, and consider a Gaussian field $W(\boldsymbol{t}), \boldsymbol{t} \in \mathbb{R}^n$, with continuous trajectories such that $\mathbb{E}W(\boldsymbol{t}) = -|\boldsymbol{t}|_{E,\boldsymbol{\alpha}}$ and $\mathrm{Cov}(W(\boldsymbol{t}), W(\boldsymbol{s})) = |\boldsymbol{t}|_{E,\boldsymbol{\alpha}} + |\boldsymbol{s}|_{E,\boldsymbol{\alpha}} - |\boldsymbol{t} - \boldsymbol{s}|_{E,\boldsymbol{\alpha}}$. It is known that such a field exists (see page 98, Piterbarg [34]). For any measurable subset $\mathcal{T} \subset \mathbb{R}^n$ define

$$H_{E,\boldsymbol{\alpha}}(\mathcal{T}) = \mathbb{E}\exp\Big(\sup_{t\in\mathcal{T}} W(\boldsymbol{t})\Big). \tag{2.1}$$

For any $T > 0$, denote $[0, T]^n = \{\boldsymbol{t} \in \mathbb{R}^n : t_i \in [0, T]\}$. The generalized Pickands' constant is defined as

$$H_{E,\boldsymbol{\alpha}} = \lim_{T\to\infty} \frac{H_{E,\boldsymbol{\alpha}}([0, T]^n)}{T^n},$$

which is a positive finite number. When $k = 1$, $E = \{1\}$ and $\boldsymbol{\alpha} = \alpha \in (0, 2]$, we denote $H_{E,\boldsymbol{\alpha}} = H_\alpha$. We use the following local stationarity concept (see Definition 7.1 in Piterbarg [34]).

Definition 2.1 (*Local-*$(E, \boldsymbol{\alpha}, D_{\boldsymbol{t}})$*-Stationarity*)**.** Let $\{Z(\boldsymbol{t}), \boldsymbol{t} \in \mathcal{M}\}$ be a Gaussian random field with covariance function r_Z, indexed on a submanifold $\mathcal{M}$ of $\mathbb{R}^n$. Z is said to be locally-$(E, \boldsymbol{\alpha}, D_{\boldsymbol{t}})$-stationary on $\mathcal{M}$, if for every $\boldsymbol{t} \in \mathcal{M}$ there exists an $n \times n$ nonsingular matrix $D_{\boldsymbol{t}}$ such that

$$r_Z(\boldsymbol{t}_1, \boldsymbol{t}_2) = 1 - |D_{\boldsymbol{t}}(\boldsymbol{t}_1 - \boldsymbol{t}_2)|_{E,\boldsymbol{\alpha}}(1 + o(1)), \tag{2.2}$$

as $\max\{\|\boldsymbol{t} - \boldsymbol{t}_1\|, \|\boldsymbol{t} - \boldsymbol{t}_2\|\} \to 0$ for $\boldsymbol{t}_1, \boldsymbol{t}_2 \in \mathcal{M}$, where the $o(1)$-term is uniform in $\boldsymbol{t} \in \mathcal{M}$.

Positive reach: We use the concept of reach to characterize the manifold $\mathcal{M}$. For a set $A \subset \mathbb{R}^n$ and a point $\boldsymbol{x} \in \mathbb{R}^n$, let $d(\boldsymbol{x}, A) = \inf\{\|\boldsymbol{x} - \boldsymbol{y}\| : \boldsymbol{y} \in A\}$ be the distance from $\boldsymbol{x}$ to A. The normal projection onto A is defined as $\pi_A(\boldsymbol{x}) = \{\boldsymbol{y} \in A : \|\boldsymbol{x} - \boldsymbol{y}\| = d(\boldsymbol{x}, A)\}$. For $\delta > 0$, let $\mathcal{B}(\boldsymbol{x}, \delta) = \{\boldsymbol{y} \in \mathbb{R}^n : \|\boldsymbol{x} - \boldsymbol{y}\| \le \delta\}$ be the closed ball centered at $\boldsymbol{x}$ with radius δ, and $\mathcal{B}^\circ(\boldsymbol{x}, \delta)$ be its interior. The reach of A, denoted by $\Delta(A)$, is defined as the largest $\delta > 0$ such that for each point $\boldsymbol{x} \in \cup_{\boldsymbol{y}\in A}\mathcal{B}(\boldsymbol{y}, \delta)$, $\pi_A(\boldsymbol{x})$ consists of a single point. See Federer [17]. The reach of a manifold is also called condition number (see Niyogi et al. [31]). A closed submanifold of $\mathbb{R}^n$ has positive reach if and only if it is $C^{1,1}$ (see Scholtes, [40]). Here a $C^{1,1}$ manifold by definition is a C^1 manifold equipped with a class of atlases whose transition maps have Lipschitz continuous first derivatives. The concept of positive reach is also closely related to "r-convexity" and "rolling conditions" (Cuevas et al. [15]).

Suppose that the structure $(E, \boldsymbol{\alpha})$ is given. Let $R = \{r_1, \ldots, r_k\}$ be a collection of positive integers such that $r_i \le e_i$, $i = 1, \ldots, k$, for which we denote $R \le E$. Let $r = r_1 + \cdots + r_k$. We impose the following assumptions on the manifold $\mathcal{M}$ and the Gaussian field $X(\boldsymbol{t})$, $\boldsymbol{t} \in \mathcal{M}$:

(A1) For $R \le E$, we assume that $\mathcal{M} = \mathcal{M}_1 \times \cdots \times \mathcal{M}_k$, where for $i = 1, \ldots, k$, $\mathcal{M}_i$ is an r_i-dimensional compact submanifold of $\mathbb{R}^{e_i}$ with positive reach.

(A2) Let $D_{\boldsymbol{t}} = \text{diag}(D_{1,\boldsymbol{t}}, \ldots, D_{k,\boldsymbol{t}})$ be a block diagonal matrix, where the dimension of $D_{i,\boldsymbol{t}}$ is $e_i \times e_i$, and the matrix-valued function $D_{i,\boldsymbol{t}}$ is continuous in $\boldsymbol{t} \in \mathcal{M}$, for $i = 1, \ldots, k$. For $0 < \alpha_1, \ldots, \alpha_k \le 2$, we assume that the Gaussian field $X(\boldsymbol{t})$ on $\mathcal{M}$ has zero mean and is locally-$(E, \boldsymbol{\alpha}, D_{\boldsymbol{t}})$-stationary.

Remark 2.1. With the condition in (A1), we have the following expression for $|D_{\boldsymbol{t}}(\boldsymbol{t}_1 - \boldsymbol{t}_2)|_{E,\boldsymbol{\alpha}}$ in (2.2).

$$|D_{\boldsymbol{t}}(\boldsymbol{t}_1 - \boldsymbol{t}_2)|_{E,\boldsymbol{\alpha}} = \sum_{j=1}^{k} \|D_{j,\boldsymbol{t}}(\boldsymbol{t}_{1,(j)} - \boldsymbol{t}_{2,(j)})\|^{\alpha_i},$$

where we denote $\boldsymbol{t}_i = (\boldsymbol{t}_{i,(1)}, \ldots, \boldsymbol{t}_{i,(k)})^T$, $i = 1, 2$. Note that the local stationarity condition for the Gaussian field is given using the structure $(E, \boldsymbol{\alpha})$ for $\mathbb{R}^n$. The structural assumptions on $\mathcal{M}$ and $D_{\boldsymbol{t}}$ in (A1) and (A2) are used to guarantee that a similar structure $(R, \boldsymbol{\alpha})$ can be found when the local stationarity of the Gaussian field is expressed on a low-dimensional manifold, which locally resembles $\mathbb{R}^r$. Note that, however, in the special case of $k = 1$ we do not have these structural constraints for $\mathcal{M}$ and D_t any more.

Some notation: Let $1 \leq m \leq n$. For an $n \times m$ matrix G, let $\|G\|_m^2$ be the sum of squares of all minor determinants of order m. For $m \geq 0$, let $\mathcal{H}_m$ be the m-dimensional normalized Hausdorff measure (see Definition A.1 in the appendix). It coincides with the m-dimensional Lebesgue measure for Lebesgue measurable sets when m is a positive integer. For a C^1 manifold M, at each $u \in M$, let $T_u M$ denote the tangent space of M at u. Let ϕ and Φ denote the standard normal density and cumulative distribution function, respectively, and let $\bar{\Phi}(u) = 1 - \Phi(u)$ and $\Psi(u) = u^{-1}\phi(u)$. Recall that $\boldsymbol{t} = (\boldsymbol{t}_{(1)}^T, \ldots, \boldsymbol{t}_{(k)}^T)^T$. The following is a result for the asymptotic behavior of the excursion probability of X on the manifold $\mathcal{M}$.

Theorem 2.1. *For a Gaussian field $X(\boldsymbol{t})$, $\boldsymbol{t} \in \mathcal{M}$ satisfying assumptions (A1) and (A2), if $r_X(\boldsymbol{t}, \boldsymbol{s}) < 1$ for all $\boldsymbol{t}, \boldsymbol{s}$ from $\mathcal{M}$, $\boldsymbol{t} \neq \boldsymbol{s}$, then*

$$\mathbb{P}\left(\sup_{\boldsymbol{t}\in\mathcal{M}} X(\boldsymbol{t}) > u\right) = H_{R,\boldsymbol{\alpha}} \int_{\mathcal{M}} \prod_{j=1}^{k} \|D_{j,\boldsymbol{t}} P_{j,\boldsymbol{t}_{(j)}}\|_{r_j} d\mathcal{H}_r(\boldsymbol{t}) \prod_{i=1}^{k} u^{2r_i/\alpha_i} \Psi(u)(1 + o(1)), \quad (2.3)$$

as $u \to \infty$, where $P_{j,\boldsymbol{t}_{(j)}}$ is an $e_j \times r_j$ matrix whose columns are orthonormal and span the tangent space $T_{\boldsymbol{t}_{(j)}}\mathcal{M}_j$.

Remark 2.2.

a. The factorization lemma (Lemma 6.4, Piterbarg [34]) implies that $H_{R,\alpha} = \prod_{i=1}^k H_{r_i,\alpha_i}$, where in the notation we do not distinguish between r_i (or α_i) and $\{r_i\}$ (or $\{\alpha_i\}$).
b. An equivalent expression of the integrand in (2.3) is given by (see (4.2) in the proof)

$$\prod_{j=1}^{k} \|D_{j,\boldsymbol{t}} P_{j,\boldsymbol{t}_{(j)}}\|_{r_j} = \|D_{\boldsymbol{t}} P_{\boldsymbol{t}}\|_r = \sqrt{\det(P_{\boldsymbol{t}}^T D_{\boldsymbol{t}}^T D_{\boldsymbol{t}} P_{\boldsymbol{t}})}, \quad (2.4)$$

where $P_{\boldsymbol{t}} = \operatorname{diag}(P_{1,\boldsymbol{t}_{(1)}}, \ldots, P_{k,\boldsymbol{t}_{(k)}})$, whose columns form a basis of the tangent space $T_{\boldsymbol{t}}\mathcal{M}$. The quantity in (2.4) is invariant if we choose a different basis of $T_{\boldsymbol{t}}\mathcal{M}$ for the projection matrix $P_{\boldsymbol{t}}$. See (4.3) in the proof. Here $P_{\boldsymbol{t}}^T D_{\boldsymbol{t}}^T D_{\boldsymbol{t}} P_{\boldsymbol{t}}$ is a Gramian matrix of the column vectors of $D_{\boldsymbol{t}} P_{\boldsymbol{t}}$, and $\|D_{\boldsymbol{t}} P_{\boldsymbol{t}}\|_r$ is the r-dimensional volume of the parallelotope formed by these vectors. Heuristically, $\|D_{\boldsymbol{t}} P_{\boldsymbol{t}}\|_r$ reflects the local variability of the Gaussian field $X(\boldsymbol{t})$ when projected to the tangent space $T_{\boldsymbol{t}}\mathcal{M}$, and is independent of the choice of the local coordinate system represented by $P_{\boldsymbol{t}}$.
c. When $E = R$, (2.3) becomes

$$\mathbb{P}\left(\sup_{\boldsymbol{t}\in\mathcal{M}} X(\boldsymbol{t}) > u\right) = H_{E,\boldsymbol{\alpha}} \int_{\mathcal{M}} |\det(D_{\boldsymbol{t}})| d\mathcal{H}_n(\boldsymbol{t}) \prod_{i=1}^{k} u^{2e_i/\alpha_i} \Psi(u)(1 + o(1)), \quad (2.5)$$

as $u \to \infty$, which is consistent with Theorem 7.1 in Piterbarg [34].

When $D_{\boldsymbol{t}} \equiv c\boldsymbol{I}$ for some constant $c \neq 0$ and $k = 1$ such that $\boldsymbol{\alpha} = \alpha$, the local stationarity condition of $X(\boldsymbol{t})$ used in Theorem 2.1 becomes the following local isotropy condition.

$$r_X(\boldsymbol{t}_1, \boldsymbol{t}_2) = 1 - |c|^\alpha \|\boldsymbol{t}_1 - \boldsymbol{t}_2\|^\alpha (1 + o(1)), \text{ as } \|\boldsymbol{t}_1 - \boldsymbol{t}_2\| \to 0. \quad (2.6)$$

We give the explicit form of the asymptotic excursion probability for this case in the following corollary, as an immediate result of Theorem 2.1.

Corollary 2.1. *When the assumptions in Theorem 2.1 hold with $D_t \equiv c\boldsymbol{I}$ for some constant $c \neq 0$ and $k = 1$ such that $\boldsymbol{\alpha} = \alpha$, we have*

$$\mathbb{P}\left(\sup_{t\in\mathcal{M}} X(\boldsymbol{t}) > u\right) = H_{R,\alpha}|c|^r \mathcal{H}_r(\mathcal{M}) u^{2r/\alpha}\,\Psi(u)(1+o(1)),\ \ as\ u\to\infty. \tag{2.7}$$

Remark 2.3. The case of $\mathcal{M} = \mathbb{S}^r$ when $n = r+1$ is of special interest in some applications. Let $d_{\mathbb{S}^r}(\boldsymbol{t}_1, \boldsymbol{t}_2) = \arccos(\boldsymbol{t}_1^T\boldsymbol{t}_2)$ be the spherical distance between $\boldsymbol{t}_1, \boldsymbol{t}_2 \in \mathbb{S}^r$. It is easy to see that the local isotropy condition (2.6) for $\mathcal{M} = \mathbb{S}^r$ can be equivalently written as

$$r_X(\boldsymbol{t}_1, \boldsymbol{t}_2) = 1 - |c|^\alpha[d_{\mathbb{S}^r}(\boldsymbol{t}_1, \boldsymbol{t}_2)]^\alpha(1+o(1)),\ \text{as}\ \|\boldsymbol{t}_1 - \boldsymbol{t}_2\| \to 0.$$

Correspondingly (2.7) becomes

$$\mathbb{P}\left(\sup_{t\in\mathbb{S}^r} X(\boldsymbol{t}) > u\right) = H_{R,\alpha}|c|^r \mathcal{H}_r(\mathbb{S}^r) u^{2r/\alpha}\,\Psi(u)(1+o(1)),\ \text{as}\ u\to\infty. \tag{2.8}$$

It is known that $\mathcal{H}_r(\mathbb{S}^r) = \frac{2\pi^{\frac{r+1}{2}}}{\Gamma(\frac{r+1}{2})}$, where Γ is the gamma function. This result is consistent with Theorem 2.4 in Cheng and Xiao (2016).

We will apply Theorem 2.1 to study the excursion probabilities of χ-fields indexed by manifolds. Let $\{\boldsymbol{X}(\boldsymbol{s}), \boldsymbol{s} \in \mathcal{L}\}$ be a centered p-dimensional ($p \geq 2$) Gaussian vector field, where $\boldsymbol{X} = (X_1, \ldots, X_p)^T$ with $\mathrm{Var}(X_i) = 1$, $i = 1, \ldots, p$, and $\mathcal{L}$ is an m-dimensional submanifold of $\mathbb{R}^n$ ($1 \leq m \leq n$). We consider the asymptotics of

$$\mathbb{P}\left(\sup_{s\in\mathcal{L}} \|\boldsymbol{X}(\boldsymbol{s})\| > u\right),\quad \text{as}\ u\to\infty. \tag{2.9}$$

Let $\boldsymbol{v} = (v_1, \ldots, v_p)^T \in \mathbb{R}^p$, $\boldsymbol{t} = (\boldsymbol{s}^T, \boldsymbol{v}^T)^T \in \mathbb{R}^{n+p}$, and

$$Y(\boldsymbol{t}) = Y(\boldsymbol{s}, \boldsymbol{v}) = X_1(\boldsymbol{s})v_1 + \cdots + X_p(\boldsymbol{s})v_p. \tag{2.10}$$

Due to the relation in (1.4), it is clear that (2.9) is equivalent to

$$\mathbb{P}\left(\sup_{t\in\mathcal{L}\times\mathbb{S}^{p-1}} Y(\boldsymbol{t}) > u\right),\quad \text{as}\ u\to\infty. \tag{2.11}$$

To study (2.9) through (2.11), we directly impose an assumption on the covariance function r_Y of Y, which we find convenient because it allows us to encode the possible cross-dependence structure among $X_1, \ldots, X_r$ into r_Y. See example (ii) below. For $i = 1, 2$, denote $\boldsymbol{t}_i = (\boldsymbol{s}_i^T, \boldsymbol{v}_i^T)^T$, where $\boldsymbol{v}_i^T = (v_{i,1}, \ldots, v_{i,p})$. Let $r_Y(\boldsymbol{t}_1, \boldsymbol{t}_2) = \mathrm{Cov}(Y(\boldsymbol{t}_1), Y(\boldsymbol{t}_2))$. Then notice that

$$\begin{aligned} r_Y(\boldsymbol{t}_1, \boldsymbol{t}_2) =& \sum_{i=1}^p\sum_{j=1}^p \mathrm{Cov}(X_i(\boldsymbol{s}_1), X_j(\boldsymbol{s}_2))v_{1,i}v_{2,j} \\ =& \boldsymbol{v}_1^T\boldsymbol{v}_2 - \sum_{i=1}^p\sum_{j=1}^p[\delta_{ij} - \mathrm{Cov}(X_i(\boldsymbol{s}_1), X_j(\boldsymbol{s}_2))]v_{1,i}v_{2,j} \\ =& 1 - \frac{1}{2}\|\boldsymbol{v}_1 - \boldsymbol{v}_2\|^2 - \sum_{i=1}^p\sum_{j=1}^p[\delta_{ij} - \mathrm{Cov}(X_i(\boldsymbol{s}_1), X_j(\boldsymbol{s}_2))]v_{1,i}v_{2,j}, \end{aligned} \tag{2.12}$$

where $\delta_{ij} = \mathbf{1}(i = j)$ is the Kronecker delta. The structure in (2.12) suggests the following assumption on $r_Y(\boldsymbol{t}_1, \boldsymbol{t}_2)$.

(A3) We assume that $Y(\boldsymbol{t})$ given in (2.10) is a local-$(E, \boldsymbol{\alpha}, D_{\boldsymbol{t}})$-stationary Gaussian field on $\mathcal{L} \times \mathbb{S}^{p-1}$ with $D_{\boldsymbol{t}} = \mathrm{diag}(B_{\boldsymbol{t}}, \frac{1}{\sqrt{2}}\boldsymbol{I}_p)$, where $B_{\boldsymbol{t}}$ is a nonsingular $n \times n$ matrix for all $\boldsymbol{t} \in \mathcal{L} \times \mathbb{S}^{p-1}$, $E = \{n, p\}$ and $\boldsymbol{\alpha} = \{\alpha, 2\}$, for some $0 < \alpha \le 2$. We assume that the matrix-valued function $B_{\boldsymbol{t}}$ is continuous in $\boldsymbol{t} \in \mathcal{L} \times \mathbb{S}^{p-1}$.

Remark 2.4. Note that assumption (A3) implies that for $\boldsymbol{s} \in \mathcal{L}$ and $1 \le i, j \le p$,

$$\mathrm{Cov}(X_i(\boldsymbol{s}), X_j(\boldsymbol{s})) = \begin{cases} 0 & i \neq j \\ 1 & i = j. \end{cases}$$

In other words, we are considering a Gaussian vector field $\boldsymbol{X}(\boldsymbol{s})$ whose variance–covariance matrix at any point $\boldsymbol{s} \in \mathcal{L}$ has been standardized. However, cross-dependence between $X_i(\boldsymbol{s}_i)$ and $X_j(\boldsymbol{s}_j)$ is still possible under assumption (A3) for $\boldsymbol{s}_i, \boldsymbol{s}_j \in \mathcal{L}$, $\boldsymbol{s}_i \neq \boldsymbol{s}_j$ and $i \neq j$.

Corollary 2.2. *Let $\{\boldsymbol{X}(\boldsymbol{s}), \boldsymbol{s} \in \mathcal{L}\}$ be a Gaussian p-dimensional ($p \ge 2$) vector field with zero mean on a compact m-dimensional submanifold $\mathcal{L} \subset \mathbb{R}^n$ of positive reach, such that $\{Y(\boldsymbol{t}), \boldsymbol{t} \in \mathcal{L} \times \mathbb{S}^{p-1}\}$ in (2.10) satisfies assumption (A3). If $r_Y(\boldsymbol{t}_1, \boldsymbol{t}_2) < 1$ for all $\boldsymbol{t}_1, \boldsymbol{t}_2$ from $\mathcal{L} \times \mathbb{S}^{p-1}$, $\boldsymbol{t}_1 \neq \boldsymbol{t}_2$, then*

$$\mathbb{P}\left(\sup_{\boldsymbol{s}\in\mathcal{L}} \|\boldsymbol{X}(\boldsymbol{s})\| > u\right) = \frac{H_{m,\alpha}}{(2\pi)^{(p-1)/2}} \int_{\mathcal{L}\times\mathbb{S}^{p-1}} \|B_{\boldsymbol{t}} P_{\boldsymbol{s}}\|_m d\mathcal{H}_{m+p-1}(\boldsymbol{t}) u^{2m/\alpha+p-1} \Psi(u)(1+o(1)), \tag{2.13}$$

as $u \to \infty$, where $P_{\boldsymbol{s}}$ is an $n \times m$ matrix whose columns are orthonormal and span the tangent space $T_{\boldsymbol{s}}\mathcal{L}$.

Remark 2.5.

a. This corollary is a direct consequence of Theorem 2.1 using $R = (m, p-1)$. To see this, notice that $H_{R,\boldsymbol{\alpha}} = H_{m,\alpha} H_{p-1,2} = H_{m,\alpha}(\sqrt{\pi})^{-(p-1)}$, because of the factorization lemma (see Remark 2.2) and the well known fact $H_2 = (\pi)^{-1/2}$ (see page 31, Piterbarg [34]). Also notice that $\|\frac{1}{\sqrt{2}}\boldsymbol{I}_p P_{\boldsymbol{u}}\|_{p-1} = 2^{-(p-1)/2}$, where $P_{\boldsymbol{u}}$ is a $p \times (p-1)$ orthonormal matrix whose columns span the tangent space $T_{\boldsymbol{u}}\mathbb{S}^{p-1}$.
b. Even though the result in this corollary is stated for $p \ge 2$, it can be easily extended to the case $p = 1$. When $p = 1$, we write $\boldsymbol{X}(\boldsymbol{s}) = X(\boldsymbol{s}) \in \mathbb{R}$ and $\mathbb{S}^{p-1} = \{\pm 1\}$. Then using the same proof of this corollary, one can show that under the assumptions given in this corollary (in a broader sense such that $B_{\boldsymbol{t}} = B_{\boldsymbol{s}}$ only depends on $\boldsymbol{s} \in \mathcal{L}$, because $\mathbb{S}^{p-1}$ now is a discrete set), we have that as $u \to \infty$,
$$\mathbb{P}\left(\sup_{\boldsymbol{s}\in\mathcal{L}} |X(\boldsymbol{s})| > u\right) = 2H_{m,\alpha} \int_{\mathcal{L}} \|B_{\boldsymbol{s}} P_{\boldsymbol{s}}\|_m d\mathcal{H}_m(\boldsymbol{s}) u^{2m/\alpha} \Psi(u)(1+o(1)), \tag{2.14}$$
where the factor 2 on the right-hand side is the cardinality (i.e., the 0-dimensional Hausdorff measure) of the set $\mathbb{S}^0$.

Examples. Below we give two examples of Gaussian vector fields $\boldsymbol{X}$ that satisfy assumption (A3).

(i) Let $X_1(\boldsymbol{s}), \ldots, X_p(\boldsymbol{s})$ be i.i.d. copies of $\{X(\boldsymbol{s}), \ \boldsymbol{s} \in \mathcal{L}\}$, which is assumed to be locally-$(n, \alpha, B_{\boldsymbol{s}})$-stationary, where $0 < \alpha \le 2$, that is,

$$r_X(\boldsymbol{s}_1, \boldsymbol{s}_2) = 1 - \|B_{\boldsymbol{s}}(\boldsymbol{s}_1 - \boldsymbol{s}_2)\|^\alpha (1 + o(1)),$$

as $\max\{\|\boldsymbol{s}-\boldsymbol{s}_1\|, \|\boldsymbol{s}-\boldsymbol{s}_2\|\} \to 0$. In this case, (A3) is satisfied because

$$\begin{aligned} r_Y(\boldsymbol{t}_1,\boldsymbol{t}_2) =& r_X(\boldsymbol{s}_1,\boldsymbol{s}_2)\boldsymbol{v}_1^T\boldsymbol{v}_2 \\ =& 1-[\|B_s(\boldsymbol{s}_1-\boldsymbol{s}_2)\|^\alpha + \frac{1}{2}\|\boldsymbol{v}_1-\boldsymbol{v}_2\|^2](1+o(1)), \end{aligned}$$

as $\max\{\|\boldsymbol{t}-\boldsymbol{t}_1\|, \|\boldsymbol{t}-\boldsymbol{t}_2\|\} \to 0$. In other words, $Y(\boldsymbol{t})$ is locally-$(E,\boldsymbol{\alpha},D_{\boldsymbol{t}})$-stationary, where $D_{\boldsymbol{t}} = \mathrm{diag}(B_s, \frac{1}{\sqrt{2}}\boldsymbol{I}_p)$, $E=\{n,p\}$ and $\boldsymbol{\alpha}=\{\alpha,2\}$.

(ii) Consider $X_i(\boldsymbol{s})$ as a locally-$(n,2,(A_{\boldsymbol{s}}^{i,i})^{1/2})$ stationary field, where $A_{\boldsymbol{s}}^{i,i}$ are positive definite $n\times n$ matrices, for $i=1,\dots,p$. Also for $1\le i\ne j\le p$, suppose $\mathrm{Cov}(X_i(\boldsymbol{s}_1),X_j(\boldsymbol{s}_2)) = (\boldsymbol{s}_1-\boldsymbol{s}_2)^T A_{\boldsymbol{s}}^{i,j}(\boldsymbol{s}_1-\boldsymbol{s}_2)(1+o(1))$, as $\max\{\|\boldsymbol{s}-\boldsymbol{s}_1\|, \|\boldsymbol{s}-\boldsymbol{s}_2\|\} \to 0$, where $A_{\boldsymbol{s}}^{i,j}$ are $n\times n$ symmetric matrices. So overall for $1\le i\ne j\le p$ we may write

$$\mathrm{Cov}(X_i(\boldsymbol{s}_1),X_j(\boldsymbol{s}_2)) = \delta_{ij} - (\boldsymbol{s}_1-\boldsymbol{s}_2)^T A_{\boldsymbol{s}}^{i,j}(\boldsymbol{s}_1-\boldsymbol{s}_2)(1+o(1)),$$

as $\max\{\|\boldsymbol{s}-\boldsymbol{s}_1\|, \|\boldsymbol{s}-\boldsymbol{s}_2\|\} \to 0$. Using (2.12), we have

$$r_Y(\boldsymbol{t}_1,\boldsymbol{t}_2) = 1-\frac{1}{2}\|\boldsymbol{v}_1-\boldsymbol{v}_2\|^2 - (\boldsymbol{s}_1-\boldsymbol{s}_2)^T\left\{\sum_{i=1}^p\sum_{j=1}^p[v_iv_jA_{\boldsymbol{s}}^{i,j}]\right\}(\boldsymbol{s}_1-\boldsymbol{s}_2)(1+o(1)).$$

Let $A_{\boldsymbol{t}} = \sum_{i=1}^p\sum_{j=1}^p[v_iv_jA_{\boldsymbol{s}}^{i,j}]$. If $A_{\boldsymbol{t}}$ is positive definite, then (A3) is satisfied with $B_{\boldsymbol{t}} = (A_{\boldsymbol{t}})^{1/2}$, $E=n+p$ and $\boldsymbol{\alpha}=2$. The matrix $A_{\boldsymbol{t}}$ is positive definite under many possible conditions. For example, if for each i, $\lambda_{\min}(A_{\boldsymbol{t}}^{i,i}) > \sum_{j\ne i}|\lambda_{\min}(A_{\boldsymbol{t}}^{i,j})|$, where $\lambda_{\min}$ is the smallest eigenvalue of a matrix, then $A_{\boldsymbol{t}}$ is positive definite because for any $u\in\mathbb{R}^n$ with $\|u\|>0$ and any $\boldsymbol{v}\in\mathbb{S}_{r-1}$,

$$u^TA_{\boldsymbol{t}}u \ge \sum_{i=1}^p\sum_{j=1}^p\lambda_{\min}(A_{\boldsymbol{t}}^{i,j})v_iv_j\|u\|^2 = \boldsymbol{v}^T\Lambda_{\min}\boldsymbol{v}\|u\|^2 > 0,$$

where $\Lambda_{\min}$ is a matrix consisting of $\lambda_{\min}(A_{\boldsymbol{t}}^{i,j})$, which is positive definite.

3. Extremes of rescaled Gaussian and χ fields on manifolds

In this section, we consider a class of centered Gaussian fields $\{X_h(\boldsymbol{t}),\boldsymbol{t}\in\mathcal{M}_h\}_{h\in(0,h_0]}$ for some $0<h_0<1$, where $\mathcal{M}_h=\mathcal{M}_{h,1}\times\mathcal{M}_{h,2}$ are r-dimensional compact submanifolds of $\mathbb{R}^n$. We will develop the result in (1.2), where the index $\boldsymbol{t}$ is (partially) rescaled by multiplying h^{-1}. For simplicity of exposition, in the structure $(E,\boldsymbol{\alpha})$, we take $k=1$ or $k=2$. The case $k=1$ also corresponds to $\mathcal{M}_{h,2}=\emptyset$ for the case $k=2$. The results in this section can be generalized using the same structure $(E,\boldsymbol{\alpha})$ as in Section 2.

When $k=2$, we denote $K=\{1,2\}$ and have $\boldsymbol{\alpha}=\{\alpha_1,\alpha_2\}$, $E=\{e_1,e_2\}$ and $R=\{r_1,r_2\}$, where $1\le r_1\le e_1$, $1\le r_2\le e_2$, $r=r_1+r_2$, and $n=e_1+e_2$. For $\boldsymbol{t}=(\boldsymbol{t}_{(1)}^T,\boldsymbol{t}_{(2)}^T)^T\in\mathbb{R}^{e_1}\times\mathbb{R}^{e_2}=\mathbb{R}^n$, let $\xi_h:\mathbb{R}^n\to\mathbb{R}^n$ be a function such that $\xi_h(\boldsymbol{t})=(h\boldsymbol{t}_{(1)}^T,\boldsymbol{t}_{(2)}^T)^T$. For any $\boldsymbol{s}\in\mathcal{M}_h=\mathcal{M}_{h,1}\times\mathcal{M}_{h,2}$, let $D_{\boldsymbol{s},h}=\mathrm{diag}(D_{\boldsymbol{s},h}^{(1)},D_{\boldsymbol{s},h}^{(2)})$ be an $n\times n$ block diagonal matrix.

When $k=1$, we denote $K=\{1\}$ and have $\boldsymbol{\alpha}=\alpha=\alpha_1$, $r_2=0$, $E=n=e_1$ and $R=r=r_1$, where $1\le r\le n$. For $\boldsymbol{t}=\boldsymbol{t}_{(1)}\in\mathbb{R}^n$, let $\xi_h:\mathbb{R}^n\to\mathbb{R}^n$ be a function such that $\xi_h(\boldsymbol{t})=h\boldsymbol{t}$. For $\boldsymbol{s}\in\mathcal{M}_h=\mathcal{M}_{h,1}$, let $D_{\boldsymbol{s},h}=D_{\boldsymbol{s},h}^{(1)}$ be an $n\times n$ matrix.

We first give the following assumptions before formulating the main result of this section. Let ξ_h^{-1} be the inverse function of ξ_h. Denote $\overline{\mathcal{M}}_h=\xi_h^{-1}(\mathcal{M}_h)=\{\boldsymbol{t}:\xi_h(\boldsymbol{t})\in\mathcal{M}_h\}$. Let $\overline{X}_h(\boldsymbol{t})=X_h(\xi_h(\boldsymbol{t})),\boldsymbol{t}\in\overline{\mathcal{M}}_h$. Let $\bar{r}_h(\boldsymbol{t}_1,\boldsymbol{t}_2)$ be the covariance between $\overline{X}_h(\boldsymbol{t}_1)$ and $\overline{X}_h(\boldsymbol{t}_2)$, for $\boldsymbol{t}_1,\boldsymbol{t}_2\in\overline{\mathcal{M}}_h$.

(B1) For $i \in K$, assume that $\mathcal{M}_{h,i}$ is an r_i-dimensional compact submanifold of $\mathbb{R}^{e_i}$, with $\inf_{0<h\leq h_0} \Delta(\mathcal{M}_{h,i}) > 0$, and

$$0 < \inf_{0<h\leq h_0} \mathcal{H}_{r_i}(\mathcal{M}_{h,i}) \leq \sup_{0<h\leq h_0} \mathcal{H}_{r_i}(\mathcal{M}_{h,i}) < \infty, \; i \in K.$$

(B2) $\overline{X}_h(\boldsymbol{t})$ is locally-$(E, \boldsymbol{\alpha}, D_{\xi_h(\boldsymbol{t}),h})$-stationary in the following uniform sense: for $\boldsymbol{t}, \boldsymbol{t}_1, \boldsymbol{t}_2 \in \overline{\mathcal{M}}_h$, as $\max\{\|\boldsymbol{t} - \boldsymbol{t}_1\|, \|\boldsymbol{t} - \boldsymbol{t}_2\|\} \to 0$,

$$\bar{r}_h(\boldsymbol{t}_1, \boldsymbol{t}_2) = 1 - |D_{\xi_h(\boldsymbol{t}),h}(\boldsymbol{t}_1 - \boldsymbol{t}_2)|_{E,\boldsymbol{\alpha}}(1 + o(1)), \tag{3.1}$$

where the $o(1)$-term is uniform in $\boldsymbol{t} \in \overline{\mathcal{M}}_h$ and $0 < h \leq h_0$. Here for $i \in K$, the dimension of $D_{\boldsymbol{s},h}^{(i)}$ is $e_i \times e_i$, and the matrix-valued function $D_{\boldsymbol{s},h}^{(i)}$ of $\boldsymbol{s}$ has continuous components on $\mathcal{M}_h$. Also

$$0 < \inf_{0<h\leq h_0, \boldsymbol{s}\in\mathcal{M}_h} \lambda_{\min}([D_{\boldsymbol{s},h}^{(i)}]^T D_{\boldsymbol{s},h}^{(i)}) \leq \sup_{0<h\leq h_0, \boldsymbol{s}\in\mathcal{M}_h} \lambda_{\max}([D_{\boldsymbol{s},h}^{(i)}]^T D_{\boldsymbol{s},h}^{(i)}) < \infty, \; i \in K, \tag{3.2}$$

where $\lambda_{\min}$ and $\lambda_{\max}$ are the smallest and largest eigenvalues of symmetric matrices, respectively.

(B3) Suppose that, for any $x > 0$, there exists $\eta > 0$ such that $Q(x) < \eta < 1$, where

$$Q(x) = \sup_{0<h\leq h_0} \{|\bar{r}_h(\boldsymbol{t}, \boldsymbol{s})| : \boldsymbol{t}, \boldsymbol{s} \in \overline{\mathcal{M}}_h, \|\boldsymbol{t}_{(1)} - \boldsymbol{s}_{(1)}\| \geq x\}. \tag{3.3}$$

(B4) There exist $x_0 > 0$ and a function $v(\cdot)$ such that for $x > x_0$, we have

$$Q(x)\left|(\log x)^{2(r_1/\alpha_1 + r_2/\alpha_2)}\right| \leq v(x), \tag{3.4}$$

where v is monotonically decreasing, such that, for any $q > 0$, $v(x^q) = O(v(x)) = o(1)$ and $v(x)x^q \to \infty$ as $x \to \infty$.

Remark 3.1. Assumptions (B1)–(B3) extends their counterparts used in Theorem 2.1 to some forms that are uniform for the classes of Gaussian fields and manifolds. Assumption (B4) is analogous to the classical Berman condition used for proving extreme value distributions (see Berman [5]). An example of $v(x)$ in assumption (B4) is given by $v(x) = (\log x)^{-\beta}$, for some $\beta > 0$.

Theorem 3.1. *Suppose assumptions (B1)–(B4) hold. Let*

$$\beta_h = \left(2r_1 \log \frac{1}{h}\right)^{\frac{1}{2}} + \left(2r_1 \log \frac{1}{h}\right)^{-\frac{1}{2}} \times \left[\left(\frac{r_1}{\alpha_1} + \frac{r_2}{\alpha_2} - \frac{1}{2}\right) \log\log \frac{1}{h} + \log\left\{ \frac{(2r_1)^{\frac{r_1}{\alpha_1} + \frac{r_2}{\alpha_2} - \frac{1}{2}}}{\sqrt{2\pi}} H_{R,\boldsymbol{\alpha}} I_h(\mathcal{M}_h) \right\} \right], \tag{3.5}$$

where $I_h(\mathcal{M}_h) = \int_{\mathcal{M}_h} \|D_{\boldsymbol{t},h} P_{\boldsymbol{t}}\|_r d\mathcal{H}_r(\boldsymbol{t})$ with $P_{\boldsymbol{t}}$ an $n \times r$ matrix with orthonormal columns spanning $T_{\boldsymbol{t}}\mathcal{M}_h$. Then

$$\lim_{h\to 0} \mathbb{P}\left\{ \sqrt{2r_1 \log \tfrac{1}{h}} \left(\sup_{\boldsymbol{t}\in\mathcal{M}_h} X_h(\boldsymbol{t}) - \beta_h \right) \leq z \right\} = e^{-e^{-z}}. \tag{3.6}$$

Remark 3.2.

a. If there exists $\gamma > 0$ such that $I_h(\mathcal{M}_h) \to \gamma$ as $h \to 0$, then obviously γ can replace $I_h(\mathcal{M}_h)$ in the theorem. Also if $\mathcal{M}_h \equiv \mathcal{M}$ and $D_{t,h} \equiv D_t$ (i.e. they are independent of h), then $I_h(\mathcal{M}_h) = \int_{\mathcal{M}} \|D_t P_t\|_r d\mathcal{H}_r(\boldsymbol{t})$.
b. The case $k = 1$ corresponds to the scenario studied in Qiao and Polonik (2018). Compared with their result, here we only need X_h to be indexed by the manifold $\mathcal{M}_h$, instead of its neighborhood.
c. When $D_{t,h} \equiv c\boldsymbol{I}$, for all $\boldsymbol{t} \in \mathcal{M}_h$ for some constant $c \neq 0$ (also see Corollary 2.1), we have $I_h(\mathcal{M}_h) = |c|^r \mathcal{H}_r(\mathcal{M}_h)$.

Next we consider the asymptotic extreme value distribution of rescaled χ-fields on manifolds. For some $0 < h_0 < 1$, let $\{\boldsymbol{X}_h(\boldsymbol{s}), \boldsymbol{s} \in \mathcal{L}_h\}_{h\in(0,h_0]}$ be a class of centered p-dimensional Gaussian random vector fields, where $\boldsymbol{X}_h = (X_{h,1}, \ldots, X_{h,p})^T$ and $\mathcal{L}_h$ are m-dimensional compact submanifolds of $\mathbb{R}^n$ $(1 \le m \le n)$. Let $\boldsymbol{v} = (v_1, \ldots, v_p)^T \in \mathbb{R}^p$ and $\boldsymbol{t} = (\boldsymbol{s}^T, \boldsymbol{v}^T)^T \in \mathbb{R}^{n+p}$. Let

$$X_h(\boldsymbol{t}) = X_h(\boldsymbol{s}, \boldsymbol{v}) = X_{h,1}(\boldsymbol{s})v_1 + \cdots + X_{h,p}(\boldsymbol{s})v_p, \;\; \boldsymbol{t} \in \mathcal{M}_h := \mathcal{L}_h \times \mathbb{S}^{p-1} \tag{3.7}$$

Using the property of Euclidean norm, we have

$$\sup_{\boldsymbol{s}\in\mathcal{L}_h} \|\boldsymbol{X}_h(\boldsymbol{s})\| = \sup_{\boldsymbol{t}\in\mathcal{M}_h} X_h(\boldsymbol{t}). \tag{3.8}$$

Corollary 3.1. *Suppose $p \ge 2$ and $\{X_h(\boldsymbol{t}),\ \boldsymbol{t} \in \mathcal{L}_h \times \mathbb{S}^{p-1}\}_{h\in(0,h_0]}$ in (3.7) satisfies assumptions (B1)–(B4) with $E = \{n, p\}$, $R = \{m, p-1\}$, $\boldsymbol{\alpha} = \{\alpha, 2\}$, and $D_{t,h} = diag(B_{t,h}, \frac{1}{\sqrt{2}}\boldsymbol{I}_p)$ where $B_{t,h}$ is a nonsingular $n \times n$ matrix. Let*

$$\beta_h = \Big(2m \log \frac{1}{h}\Big)^{\frac{1}{2}} + \Big(2m \log \frac{1}{h}\Big)^{-\frac{1}{2}} \Bigg[\Big(\frac{m}{\alpha} + \frac{p-2}{2}\Big) \log\log \frac{1}{h} + \log\Bigg\{\frac{(2m)^{\frac{m}{\alpha}+\frac{p-2}{2}}}{(\sqrt{2\pi})^p} H_{m,\alpha} I_h(\mathcal{M}_h)\Bigg\}\Bigg], \tag{3.9}$$

where $I_h(\mathcal{M}_h) = \int_{\mathcal{L}_h\times\mathbb{S}^{p-1}} \|B_{t,h}P_s\|_m d\mathcal{H}_{m+p-1}(\boldsymbol{t})$ with P_s an $n \times m$ matrix with orthonormal columns spanning $T_s\mathcal{L}_h$. Then

$$\lim_{h\to 0} \mathbb{P}\left\{\Big(2m \log \tfrac{1}{h}\Big)^{\frac{1}{2}} \left(\sup_{\boldsymbol{s}\in\mathcal{L}_h} \|\boldsymbol{X}_h(\boldsymbol{s})\| - \beta_h\right) \le z\right\} = e^{-e^{-z}}. \tag{3.10}$$

Remark 3.3.

a. The result in this corollary immediately follows from Theorem 3.1. See Remark 2.5(a) for some relevant calculation. Also, similar to Remark 2.5(b), the result in this corollary can be extended to the case $p = 1$ such that $\mathbb{S}^{p-1} = \{\pm 1\}$, for which (3.10) holds with

$$\beta_h = \Big(2m \log \frac{1}{h}\Big)^{\frac{1}{2}} + \Big(2m \log \frac{1}{h}\Big)^{-\frac{1}{2}} \Bigg[\Big(\frac{m}{\alpha} - \frac{1}{2}\Big) \log\log \frac{1}{h} + \log\Bigg\{\frac{(2m)^{\frac{m}{\alpha}-\frac{1}{2}}}{\sqrt{2\pi}} H_{m,\alpha} I_h(\mathcal{M}_h)\Bigg\}\Bigg],$$

where $I_h(\mathcal{M}_h) = 2\int_{\mathcal{L}_h} \|B_{s,h}P_s\|_m d\mathcal{H}_m(\boldsymbol{s})$.
b. In the introduction section, we briefly indicate two examples (a) and (b) as applications of this corollary using the signal plus noise model in (1.6), by taking $\mathcal{L}_h \equiv \mathcal{M}$ in this corollary. The application of this corollary to example (a) is straightforward. For example

(b), we consider $\mathcal{M}$ as (the intersection of) level sets with the form $\mathcal{M} = \{s \in \mathcal{A} : \boldsymbol{f}(s) = \boldsymbol{g}_0\}$, where $\mathcal{A} \subset \mathbb{R}^n$ is a neighborhood of $\mathcal{M}$ and $\boldsymbol{f}$ and $\boldsymbol{g}_0$ are p-dimensional vectors. We take $1 \leq p < n$. The positive reach condition of $\mathcal{M}$ required by this corollary is met when $\boldsymbol{f}$ satisfies the conditions given in Lemma 4.11 and Theorem 4.12 of Federer (1959), e.g., when the component functions of the Jacobian matrix of $\boldsymbol{f}$ are Lipschitz continuous and $\boldsymbol{g}_0$ corresponds to regular (i.e., non-critical) values of $\boldsymbol{f}$, meaning that the Jacobian matrix of $\boldsymbol{f}$ at every point of $\mathcal{M}$ has full rank. Also note that under this condition, $\mathcal{M}$ is an m-dimensional manifold by the constant-rank level set theorem (Theorem 5.12, Lee [24]), where $m = n - p$. If all the assumptions in this corollary are satisfied, then the validity of $\mathcal{F}_h$ in (1.8) as an asymptotic $(1 - \alpha)$ confidence region for $\mathcal{M}$ is simply the consequence of the equivalence of the following two events.

$$\mathcal{M} \subset \mathcal{F}_h \Longleftrightarrow a_h \left(\sup_{s \in \mathcal{M}} \|\widehat{\boldsymbol{f}}_h(s) - \boldsymbol{g}_0\| - b_h \right) \leq z_\alpha,$$

where we can take $a_h = (2m \log \frac{1}{h})^{\frac{1}{2}}$ and $b_h = \beta_h$ as in this corollary.

4. Proofs

4.1. Geometric construction for the proof of Theorem 2.1

The proof of Theorem 2.1 relies on some geometric construction on manifolds with positive reach, which we present first. Let M be an r-dimensional submanifold of $\mathbb{R}^n$. Suppose it has positive reach, i.e., $\Delta(M) > 0$. For $\varepsilon, \eta > 0$, a set of points Q on M is called an (ε, η)-sample, if

(i) ε-covering: for any $x \in M$, there exists $y \in Q$ such that $\|x - y\| \leq \varepsilon$;
(ii) η-packing: for any $x, y \in Q$, $\|x - y\| > \eta$.

For simplicity, we always use $\eta = \varepsilon$, and such an $(\varepsilon, \varepsilon)$-sample is called an ε-net. It is known that an ε-net always exists for any positive real number ε when M is bounded (see Lemma 5.2, Boissonnat, Chazal and Yvinec [9]). Let N_ε be the cardinality of this ε-net. Let

$$P_\varepsilon = \max\{n : \text{ there exists an } \varepsilon\text{-packing of } M \text{ of size } n\},$$
$$C_\varepsilon = \min\{n : \text{ there exists an } \varepsilon\text{-covering over } M \text{ of size } n\},$$

which are called the ε-packing and ε-covering numbers, respectively. It is known that (see Lemma 5.2 in Niyogi et al. [31])

$$P_{2\varepsilon} \leq C_\varepsilon \leq N_\varepsilon \leq P_\varepsilon.$$

Also it is given on page 431 of Niyogi et al. [31] that when $\varepsilon < \Delta(M)/2$

$$P_\varepsilon \leq \frac{\mathcal{H}_r(M)}{[\cos^r(\theta)]\varepsilon^r B_r},$$

where B_r is the volume of the unit r-ball, and $\theta = \arcsin(\varepsilon/2)$. This implies that $N_\varepsilon = O(\varepsilon^{-r})$, as $\varepsilon \to 0$, when $\mathcal{H}_r(M)$ is bounded.

Let $\{x_1, \ldots, x_{N_\varepsilon}\} \subset M$ be an ε-net. With this ε-net, we can construct a Voronoi diagram restricted on M consisting of N_ε Voronoi cells $V_1, \ldots, V_{N_\varepsilon}$, where $V_i = \{x \in M : \|x - x_i\| \leq \|x - x_j\|, \text{ for all } j \neq i\}$. The Voronoi diagram gives a partition of M, that is $M = \cup_{i=1}^{N_\varepsilon} V_i$.

Due to the definition of the ε-net, we have that

$$(\mathcal{B}(x_i, \varepsilon/2) \cap M) \subset V_i \subset (\mathcal{B}(x_i, \varepsilon) \cap M),\ i = 1, \ldots, N_\varepsilon.$$

In other words, the shape of all the Voronoi cells is always not very thin.

4.2. Proof of Theorem 2.1

We first give a lemma used in the proof of Theorem 2.1. Recall that a bounded subset of $\mathbb{R}^n$ is called Jordan measurable if its boundary has Lebesgue measure zero.

Lemma 4.1. *Suppose that the conditions in Theorem 2.1 hold. For a subset $V \subset \mathcal{M}$, suppose that there exist an open set $G \subset \mathbb{R}^r$ and a diffeomorphism $\psi : G \to V$, where the component functions of the Jacobian matrix J_ψ of ψ are uniformly continuous. For any subset $U \subset V$, if $\Omega := \psi^{-1}(U)$ is a compact Jordan set of positive r-dimensional Lebesgue measure, then as $u \to \infty$,*

$$\mathbb{P}\left(\sup_{t\in U} X(\boldsymbol{t}) > u\right) = H_{R,\boldsymbol{\alpha}} \int_U \prod_{j=1}^k \|D_{j,\boldsymbol{t}} P_{j,\boldsymbol{t}}\|_{r_j} d\mathcal{H}_r(\boldsymbol{t}) \prod_{i=1}^k u^{2r_i/\alpha_i}\, \Psi(u)(1 + o(1)). \tag{4.1}$$

Proof. Let $\widetilde{X} = X \circ \psi$, which is a Gaussian field indexed by $V \subset \mathbb{R}^r$. Consider $\widetilde{\boldsymbol{t}}, \widetilde{\boldsymbol{t}}_1, \widetilde{\boldsymbol{t}}_2 \in \Omega \subset V$ such that $\max\{\|\widetilde{\boldsymbol{t}} - \widetilde{\boldsymbol{t}}_1\|, \|\widetilde{\boldsymbol{t}} - \widetilde{\boldsymbol{t}}_2\|\} \to 0$. Since ψ is a diffeomorphism, we also have $\max\{\|\psi(\widetilde{\boldsymbol{t}}) - \psi(\widetilde{\boldsymbol{t}}_1)\|, \|\psi(\widetilde{\boldsymbol{t}}) - \psi(\widetilde{\boldsymbol{t}}_2)\|\} \to 0$. Using assumption (A1), we have

$$\begin{aligned}
\mathrm{Cov}(\widetilde{X}(\widetilde{\boldsymbol{t}}_1), \widetilde{X}(\widetilde{\boldsymbol{t}}_2)) =& \mathrm{Cov}(X(\psi(\widetilde{\boldsymbol{t}}_1)), X(\psi(\widetilde{\boldsymbol{t}}_2))) \\
=& 1 - |D_{\psi(\widetilde{\boldsymbol{t}})}(\psi(\widetilde{\boldsymbol{t}}_1) - \psi(\widetilde{\boldsymbol{t}}_2))|_{E,\boldsymbol{\alpha}}(1 + o(1)) \\
=& 1 - |D_{\psi(\widetilde{\boldsymbol{t}})} J_\psi(\widetilde{\boldsymbol{t}})(\widetilde{\boldsymbol{t}}_1 - \widetilde{\boldsymbol{t}}_2)|_{E,\boldsymbol{\alpha}}(1 + o(1)),
\end{aligned}$$

where in the last step we have used a Taylor expansion. Note that the above $o(1)$-term is uniform in $\widetilde{\boldsymbol{t}} \in \Omega$ due to the definition of the local-$(E, \boldsymbol{\alpha}, D_{\boldsymbol{t}})$-stationarity given in Definition 2.1, and the uniform continuity of J_ψ assumed in this lemma. Since the columns of the Jacobian matrix J_ψ span the tangent space $T_{\psi(\widetilde{\boldsymbol{t}})}\mathcal{M}$, and the matrix $D_{\psi(\widetilde{\boldsymbol{t}})}$ is assumed to be nonsingular, the matrix $D_{\psi(\widetilde{\boldsymbol{t}})} J_\psi(\widetilde{\boldsymbol{t}})$ is of full rank, and therefore

$$A(\widetilde{\boldsymbol{t}}) := [J_\psi(\widetilde{\boldsymbol{t}})]^T [D_{\psi(\widetilde{\boldsymbol{t}})}]^T D_{\psi(\widetilde{\boldsymbol{t}})} J_\psi(\widetilde{\boldsymbol{t}})$$

is positive definite. Also note that $A(\widetilde{\boldsymbol{t}})$ is a block diagonal matrix, where the diagonal blocks have dimension $r_i \times r_i$, $i = 1, \ldots, k$. Let $A(\widetilde{\boldsymbol{t}})^{1/2}$ be the principal square root matrix of $A(\widetilde{\boldsymbol{t}})$. We have that

$$\mathrm{Cov}(\widetilde{X}(\widetilde{\boldsymbol{t}}_1), \widetilde{X}(\widetilde{\boldsymbol{t}}_2)) = 1 - |A(\widetilde{\boldsymbol{t}})^{1/2}(\widetilde{\boldsymbol{t}}_1 - \widetilde{\boldsymbol{t}}_2)|_{R,\boldsymbol{\alpha}}(1 + o(1)).$$

Using Theorem 7.1 in Piterbarg [34], we obtain that as $u \to \infty$,

$$\mathbb{P}\left(\sup_{\widetilde{\boldsymbol{t}}\in\Omega} \widetilde{X}(\widetilde{\boldsymbol{t}}) > u\right) = H_{R,\boldsymbol{\alpha}} \int_\Omega \det[A(\widetilde{\boldsymbol{t}})^{1/2}] d\mathcal{H}_r(\widetilde{\boldsymbol{t}}) \prod_{i=1}^k u^{2r_i/\alpha_i}\, \Psi(u)(1 + o(1)).$$

By using the change of variables formula (see Corollary A.1 in the appendix) and noticing that $\sup_{\widetilde{\boldsymbol{t}}\in\Omega} \widetilde{X}(\widetilde{\boldsymbol{t}}) = \sup_{\boldsymbol{t}\in U} X(\boldsymbol{t})$, we have

$$\mathbb{P}\left(\sup_{t\in U} X(\boldsymbol{t}) > u\right) = H_{R,\boldsymbol{\alpha}} \int_U \frac{\det[A(\psi^{-1}(\boldsymbol{t}))^{1/2}]}{\det[B(\psi^{-1}(\boldsymbol{t}))^{1/2}]} d\mathcal{H}_r(\boldsymbol{t}) \prod_{i=1}^k u^{2r_i/\alpha_i}\, \Psi(u)(1 + o(1)),$$

where $B(\psi^{-1}(\boldsymbol{t})) = [J_\psi(\psi^{-1}(\boldsymbol{t}))]^T J_\psi(\psi^{-1}(\boldsymbol{t}))$. Let $\{p_1(\boldsymbol{t}), \ldots, p_r(\boldsymbol{t})\}$ be an orthonormal basis of the tangent space $T_{\boldsymbol{t}}\mathcal{M}$ and write $P_{\boldsymbol{t}} = [p_1(\boldsymbol{t}), \ldots, p_r(\boldsymbol{t})]$. There exists an $r \times r$ nonsingular matrix $Q_{\boldsymbol{t}}$ such that $J_\psi(\psi^{-1}(\boldsymbol{t})) = P_{\boldsymbol{t}} Q_{\boldsymbol{t}}$. Hence

$$\frac{\det[A(\psi^{-1}(\boldsymbol{t}))^{1/2}]}{\det[B(\psi^{-1}(\boldsymbol{t}))^{1/2}]} = \frac{\det[Q_{\boldsymbol{t}}]\det[(P_{\boldsymbol{t}}^T D_{\boldsymbol{t}}^T D_{\boldsymbol{t}} P_{\boldsymbol{t}})^{1/2}]}{\det[Q_{\boldsymbol{t}}]} = \det[(P_{\boldsymbol{t}}^T D_{\boldsymbol{t}}^T D_{\boldsymbol{t}} P_{\boldsymbol{t}})^{1/2}].$$

Notice that $P_{\boldsymbol{t}} = \mathrm{diag}(P_{1,\boldsymbol{t}_{(1)}}, \ldots, P_{k,\boldsymbol{t}_{(k)}})$, where $P_{j,\boldsymbol{t}_{(j)}}$ is an $e_j \times r_j$ matrix whose columns are orthonormal and span $T_{\boldsymbol{t}_{(j)}}\mathcal{M}_j$, $j = 1, \ldots, k$. Then by the Cauchy–Binet formula (see Broida and Williamson [11], page 214), we have

$$\|D_{\boldsymbol{t}} P_{\boldsymbol{t}}\|_r = \det[P_{\boldsymbol{t}}^T D_{\boldsymbol{t}}^T D_{\boldsymbol{t}} P_{\boldsymbol{t}}]^{1/2} = \prod_{j=1}^{k} \det[(P_{j,\boldsymbol{t}_{(j)}}^T D_{j,\boldsymbol{t}}^T D_{j,\boldsymbol{t}} P_{j,\boldsymbol{t}_{(j)}})^{1/2}] = \prod_{j=1}^{k} \|D_{j,\boldsymbol{t}} P_{j,\boldsymbol{t}_{(j)}}\|_{r_j}. \tag{4.2}$$

Therefore we get (4.1). We can also show that the quantity in (4.2) is invariant if we choose a different orthonormal basis of $T_{\boldsymbol{t}}\mathcal{M}$, say $\{\widetilde{p}_1(\boldsymbol{t}), \ldots, \widetilde{p}_r(\boldsymbol{t})\}$. Let $\widetilde{P}_{\boldsymbol{t}} = [\widetilde{p}_1(\boldsymbol{t}), \ldots, \widetilde{p}_r(\boldsymbol{t})]$. Then there exists an $r \times r$ orthogonal matrix $W_{\boldsymbol{t}}$ such that $\widetilde{P}_{\boldsymbol{t}} = P_{\boldsymbol{t}} W_{\boldsymbol{t}}$. We have

$$\det[\widetilde{P}_{\boldsymbol{t}}^T D_{\boldsymbol{t}}^T D_{\boldsymbol{t}} \widetilde{P}_{\boldsymbol{t}}]^{1/2} = \det[W_{\boldsymbol{t}} W_{\boldsymbol{t}}^T]^{1/2} \det[P_{\boldsymbol{t}}^T D_{\boldsymbol{t}}^T D_{\boldsymbol{t}} P_{\boldsymbol{t}}]^{1/2} = \det[P_{\boldsymbol{t}}^T D_{\boldsymbol{t}}^T D_{\boldsymbol{t}} P_{\boldsymbol{t}}]^{1/2}. \quad \square \tag{4.3}$$

Proof of Theorem 2.1. For any $\boldsymbol{t} \in \mathcal{M}$, denote $\mathcal{C}_{\boldsymbol{t}} = \mathcal{B}^\circ(\boldsymbol{t}, \Delta(\mathcal{M})/2) \cap \mathcal{M}$ and let $\tau \equiv \tau_{\boldsymbol{t}} : \mathcal{C}_{\boldsymbol{t}} \to T_{\boldsymbol{t}}\mathcal{M}$ be the projection map to the tangent space $T_{\boldsymbol{t}}\mathcal{M}$, that is, τ is a restriction of the normal projection $\pi_{T_{\boldsymbol{t}}\mathcal{M}}$ to the set $\mathcal{C}_{\boldsymbol{t}}$. Let $\mathcal{D}_{\boldsymbol{t}}$ be the image of $\mathcal{C}_{\boldsymbol{t}}$ under τ, which is an open set in $T_{\boldsymbol{t}}\mathcal{M}$. It is known that τ is a diffeomorphism between $\mathcal{C}_{\boldsymbol{t}}$ and $\mathcal{D}_{\boldsymbol{t}}$ (see Lemma 5.4, Niyogi et al. [31]). It follows from the proof of Theorem B in Leobacher and Steinicke [25] that the Jacobian of τ^{-1}, denoted by $J_{\tau^{-1}}$, is locally Lipschitz continuous in the following uniform sense: for any $\widetilde{\boldsymbol{t}} \in \mathcal{D}_{\boldsymbol{t}}$, let $a > 0$ be such that $\mathcal{B}^\circ(\widetilde{\boldsymbol{t}}, a) \cap T_{\boldsymbol{t}}\mathcal{M} \subset \mathcal{D}_{\boldsymbol{t}}$; for any $\boldsymbol{\delta} \in T_{\boldsymbol{t}}\mathcal{M}$ such that $\|\boldsymbol{\delta}\| < \frac{1}{2}a$, there exists a constant $L > 0$ only depending on $\Delta(\mathcal{M})$ such that

$$\|J_{\tau^{-1}}(\widetilde{\boldsymbol{t}} + \boldsymbol{\delta}) - J_{\tau^{-1}}(\widetilde{\boldsymbol{t}})\|_{\mathrm{op}} \le L\|\boldsymbol{\delta}\|, \tag{4.4}$$

where $\|\cdot\|_{\mathrm{op}}$ is the operator norm of matrices. Therefore $J_{\tau^{-1}}$ is uniformly continuous when restricted to $\tau^{-1}(\mathcal{B}(\boldsymbol{t}, \epsilon) \cap \mathcal{M})$ for any $0 < \epsilon < \Delta(\mathcal{M})/2$. Suppose that $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_r\}$ is an orthonormal basis of $T_{\boldsymbol{t}}\mathcal{M}$. Let $\iota : T_{\boldsymbol{t}}\mathcal{M} \to \mathbb{R}^r$ be a map such that $\iota(\boldsymbol{y}) = (y_1, \ldots, y_r) \in \mathbb{R}^r$ for $\boldsymbol{y} = y_1\boldsymbol{e}_1 + \cdots y_r\boldsymbol{e}_r \in T_{\boldsymbol{t}}\mathcal{M}$. Then $\psi := \tau^{-1} \circ \iota^{-1}$ is the diffeomorphism we need to apply Lemma 4.1.

We choose $\epsilon < \Delta(\mathcal{M})/10$. Using the method in Section 4.1, we find an ϵ-net $\{\boldsymbol{t}_1, \ldots, \boldsymbol{t}_{N_\epsilon}\}$ for $\mathcal{M}$, and construct a partition of $\mathcal{M}$ with Voronoi cells $V_1, \ldots, V_{N_\epsilon}$, where $N_\epsilon = O(\epsilon^{-r})$. Since $V_i \subset (\mathcal{B}(\boldsymbol{t}_i, \epsilon) \cap \mathcal{M})$, $\tau \equiv \tau_{\boldsymbol{t}_i}$ is a diffeomorphism on V_i, $i = 1, \ldots, N_\epsilon$.

Using Lemma 4.1, we have that

$$\mathbb{P}\left(\sup_{\boldsymbol{t} \in V_i} X(\boldsymbol{t}) > u\right) = H_{R,\boldsymbol{\alpha}} \int_{V_i} \prod_{j=1}^{k} \|D_{j,\boldsymbol{t}} P_{j,\boldsymbol{t}}\|_{r_j} d\mathcal{H}_r(\boldsymbol{t}) \prod_{j=1}^{k} u^{2r_j/\alpha_j} \Psi(u)(1 + o(1)),$$

as $u \to \infty$, and hence

$$\sum_{i=1}^{N_\epsilon} \mathbb{P}\left(\sup_{\boldsymbol{t} \in V_i} X(\boldsymbol{t}) > u\right) = H_{R,\boldsymbol{\alpha}} \int_{\mathcal{M}} \prod_{j=1}^{k} \|D_{j,\boldsymbol{t}} P_{j,\boldsymbol{t}}\|_{r_j} d\mathcal{H}_r(\boldsymbol{t}) \prod_{j=1}^{k} u^{2r_j/\alpha_j} \Psi(u)(1 + o(1)). \tag{4.5}$$

We will apply the double-sum method (see Piterbarg [34]). Using the Bonferroni inequality, we have

$$\sum_{i=1}^{N_\epsilon} \mathbb{P}\left(\sup_{t\in V_i} X(\boldsymbol{t}) > u\right) - \sum_{i\neq j} \mathbb{P}\left(\sup_{t\in V_i} X(\boldsymbol{t}) > u,\ \sup_{t\in V_j} X(\boldsymbol{t}) > u\right)$$
$$\leq \mathbb{P}\left(\sup_{t\in \mathcal{M}} X(\boldsymbol{t}) > u\right) \leq \sum_{i=1}^{N_\epsilon} \mathbb{P}\left(\sup_{t\in V_i} X(\boldsymbol{t}) > u\right). \tag{4.6}$$

For any two subsets $\mathcal{A}_1, \mathcal{A}_2 \subset \mathbb{R}^n$, define

$$\begin{aligned} d_{\max}(\mathcal{A}_1, \mathcal{A}_2) &= \sup\{\|\boldsymbol{t}_1 - \boldsymbol{t}_2\| :\ \boldsymbol{t}_1 \in \mathcal{A}_1,\ \boldsymbol{t}_2 \in \mathcal{A}_2\},\\ d_{\min}(\mathcal{A}_1, \mathcal{A}_2) &= \inf\{\|\boldsymbol{t}_1 - \boldsymbol{t}_2\| :\ \boldsymbol{t}_1 \in \mathcal{A}_1,\ \boldsymbol{t}_2 \in \mathcal{A}_2\}. \end{aligned} \tag{4.7}$$

We divide the set of indices $S = \{(i, j) :\ 1 \leq i \neq j \leq N_\epsilon\}$ into S_1 and S_2, where $S_1 = \{(i, j) \in S :\ d_{\max}(V_i, V_j) \leq 5\epsilon\}$ and $S_2 = \{(i, j) \in S :\ d_{\max}(V_i, V_j)\} > 5\epsilon\}$. If $(i, j) \in S_1$, then there exists $\bar{\boldsymbol{t}} \in \mathcal{M}$ such that $(V_i \cup V_j) \subset (\mathcal{B}(\bar{\boldsymbol{t}}, 5\epsilon)\cap\mathcal{M}) \subset (\mathcal{B}^\circ(\bar{\boldsymbol{t}}, \Delta(\mathcal{M})/2)\cap\mathcal{M})$, due to the choice $\epsilon < \Delta(\mathcal{M})/10$. With the diffeomorphism ψ defined at the beginning of the proof, we apply Lemma 4.1, and have that as $u \to \infty$,

$$\begin{aligned} &\mathbb{P}\left(\sup_{t\in V_i} X(\boldsymbol{t}) > u,\ \sup_{t\in V_j} X(\boldsymbol{t}) > u\right)\\ =&\mathbb{P}\left(\sup_{t\in V_i} X(\boldsymbol{t}) > u\right) + \mathbb{P}\left(\sup_{t\in V_j} X(\boldsymbol{t}) > u\right) - \mathbb{P}\left(\sup_{t\in V_i\cup V_j} X(\boldsymbol{t}) > u\right)\\ =&o(1)H_{R,\boldsymbol{\alpha}} \int_{V_i\cup V_j} \prod_{j=1}^{k} \|D_{j,\boldsymbol{t}} P_{j,\boldsymbol{t}}\|_{r_j} d\mathcal{H}_r(\boldsymbol{t}) \prod_{j=1}^{k} u^{2r_j/\alpha_j}\, \Psi(u). \end{aligned}$$

Therefore as $u \to \infty$,

$$\sum_{(i,j)\in S_1} \mathbb{P}\left(\sup_{t\in V_i} X(\boldsymbol{t}) > u,\ \sup_{t\in V_j} X(\boldsymbol{t}) > u\right) = o\left(\prod_{i=1}^{k} u^{2r_i/\alpha_i}\, \Psi(u)\right). \tag{4.8}$$

Next we proceed to consider $(i, j) \in S_2$. Let $Y(\boldsymbol{t}, \boldsymbol{s}) = X(\boldsymbol{t}) + X(\boldsymbol{s})$. Note that

$$\mathbb{P}\left(\sup_{t\in V_i} X(\boldsymbol{t}) > u,\ \sup_{t\in V_j} X(\boldsymbol{t}) > u\right) \leq \mathbb{P}\left(\sup_{t\in V_i, s\in V_j} Y(\boldsymbol{t}, \boldsymbol{s}) > 2u\right). \tag{4.9}$$

In order to further bound the probability on the right-hand side, we will use the Borell inequality [10] (see Theorem D.1 in Piterbarg [34]). Notice that $d_{\min}(V_i, V_j) \geq d_{\max}(V_i, V_j) - 4\epsilon$, and hence

$$\min_{(i,j)\in S_2} d_{\min}(V_i, V_j) \geq \epsilon.$$

The assumption in the theorem guarantees that $\rho := \sup_{\|\boldsymbol{t}-\boldsymbol{s}\|\geq\epsilon} r_X(\boldsymbol{t}, \boldsymbol{s}) < 1$. This then yields that

$$\max_{(i,j)\in S_2} \sup_{(\boldsymbol{t},\boldsymbol{s})\in V_i\times V_j} \mathrm{Var}\,(Y(\boldsymbol{t}, \boldsymbol{s})) \leq 2 + 2\rho$$

and

$$\sup_{(i,j)\in S_2}\sup_{(\boldsymbol{t},\boldsymbol{s})\in V_i\times V_j}\mathbb{E}\,(Y(\boldsymbol{t},\boldsymbol{s}))=0.$$

Now it remains to show that $\mathbb{P}\left(\sup_{t\in V_i,s\in V_j}Y(\boldsymbol{t},\boldsymbol{s})>b\right)\leq 1/2$ for some constant b for all $(i,j)\in S_2$ in order to apply the Borell inequality to $Y(\boldsymbol{t},\boldsymbol{s})$. Such b exists because

$$\begin{aligned}\mathbb{P}\left(\sup_{t\in V_i,s\in V_j}Y(\boldsymbol{t},\boldsymbol{s})>u\right)&\leq\mathbb{P}\left(\sup_{t\in\mathcal{M},s\in\mathcal{M}}Y(\boldsymbol{t},\boldsymbol{s})>u\right)\leq\mathbb{P}\left(\sup_{t\in\mathcal{M}}X(\boldsymbol{t})>u/2\right)\\ \leq &H_{R,\boldsymbol{\alpha}}\int_{\mathcal{M}}\prod_{j=1}^{k}\|D_{j,\boldsymbol{t}}P_{j,\boldsymbol{t}}\|_{r_j}d\mathcal{H}_r(\boldsymbol{t})\prod_{j=1}^{k}\left(\frac{u}{2}\right)^{2r_j/\alpha_j}\Psi\left(\frac{u}{2}\right)(1+o(1)),\end{aligned}$$

which tends to zero as $u\to\infty$. The application of the Borell inequality now gives that

$$\mathbb{P}\left(\sup_{t\in V_i,s\in V_j}Y(\boldsymbol{t},\boldsymbol{s})>2u\right)\leq 2\bar{\Phi}\left(\frac{u-b/2}{\sqrt{(1+\rho)/2}}\right). \tag{4.10}$$

Also note that the cardinality $|S_2|\leq N_\epsilon^2\leq C\epsilon^{-2r}$, for some constant $C>0$. Using the well-known fact that $\bar{\Phi}(u)/\Psi(u)\to 1$ as $u\to\infty$, we have

$$\begin{aligned}\sum_{(i,j)\in S_2}\mathbb{P}\left(\sup_{t\in V_i}X(\boldsymbol{t})>u,\ \sup_{t\in V_j}X(\boldsymbol{t})>u\right)&\leq 2|S_2|\bar{\Phi}\left(\frac{u-b/2}{\sqrt{(1+\rho)/2}}\right)\\ &=2|S_2|\left(\frac{u-b/2}{\sqrt{(1+\rho)/2}}\right)^{-1}\phi\left(\frac{u-b/2}{\sqrt{(1+\rho)/2}}\right)(1+o(1))\\ &=o\left(\prod_{i=1}^{k}u^{2r_i/\alpha_i}\,\Psi(u)\right),\end{aligned} \tag{4.11}$$

as $u\to\infty$, where the last step follows from $\sqrt{(1+\rho)/2}<1$. Combining (4.5), (4.6), (4.8) and (4.11), we have the desired result. □

4.3. Geometric construction for the proof of Theorem 3.1

We first give some geometric construction used in the proof of Theorem 3.1. We focus on the case $k=2$ below. For $k=1$, only the geometric construction on $\mathcal{M}_{h,1}$ is needed.

(i) *Voronoi diagram on $\mathcal{M}_h$*: Let $\ell_1=\inf_{h\in(0,h_0]}\Delta(\mathcal{M}_{h,1})/2$. It is known from Section 4.1 that there exists an $(h\ell_1)$-net $\{\boldsymbol{s}_1,\ldots,\boldsymbol{s}_{m_h}\}$ on $\mathcal{M}_h$, where $m_h=O((h\ell_1)^{-r_1})$ is the cardinality of the net. With this $(h\ell_1)$-net and using the technique described in Section 4.1, we construct a Voronoi diagram restricted on $\mathcal{M}_{h,1}$. The collections of the cells are denoted by $\{J_{k,h}:k=1,\ldots,m_h\}$, which forms a partition of $\mathcal{M}_{h,1}$. Similarly for $\mathcal{M}_{h,2}$, with $\ell_2=\inf_{h\in(0,h_0]}\Delta(\mathcal{M}_{h,2})/2$, there exists an ℓ_2-net $\{\boldsymbol{u}_1,\ldots,\boldsymbol{u}_{n_h}\}$ on $\mathcal{M}_{h,2}$, where $n_h=O(\ell_2^{-r_2})$. The cells of the corresponding Voronoi diagram on $\mathcal{M}_{h,2}$ are denoted by $U_{1,h},\ldots,U_{n_h,h}$.

(ii) *Separation of Voronoi cells:* The construction of the Voronoi diagram restricted on $\mathcal{M}_{h,1}$ guarantees that each cell $J_{k,h}\supset(\mathcal{M}_{h,1}\cap\mathcal{B}(\boldsymbol{s}_k,(h\ell_1)/2))$. In other words, $J_{k,h}$ is not too thin. For $0<\delta<\ell_1/2$, let $\partial\mathcal{J}_h=\cup_{k=1}^{m_h}(\partial J_{k,h})$ be the union of all the boundaries of the cells. Let

$$\mathcal{B}^{h\delta}=\{x\in\mathcal{M}_h:d(x,\partial\mathcal{J}_h)\leq h\delta\},$$

which is the $(h\delta)$-enlarged neighborhood of $\partial\mathcal{J}_h$. We obtain $J^{\delta}_{k,h} = J_{k,h}\backslash\mathcal{B}^{h\delta}$ and $J^{-\delta}_{k,h} = J_{k,h}\backslash J^{\delta}_{k,h}$ for $1 \le k \le m_h$. The geometric construction ensures that if $k \neq k'$, $J^{\delta}_{k,h}$ and $J^{\delta}_{k',h}$ are separated by $\mathcal{B}^{h\delta}$, which is partitioned as $\{J^{-\delta}_{k,h},\ k = 1,\dots,m_h\}$. Furthermore it is clear from the definition of $\mathcal{B}^{h\delta}$ that with $d_{\min}$ defined in (4.7) we have

$$d_{\min}(J^{\delta}_{k,h}, J^{\delta}_{k',h}) \ge 2h\delta,\ \text{for } k \neq k'. \tag{4.12}$$

(iii) *Discretization:* We construct a dense grid on $\mathcal{M}_h$ as follows. Let $\Pi_{k,j} = (\Pi_{\boldsymbol{s}_k}, \Pi_{\boldsymbol{u}_j})$ be the projection map from $J_{k,h} \times U_{j,h}$ to the tangent space $T_{\boldsymbol{s}_k}\mathcal{M}_{h,1} \times T_{\boldsymbol{u}_j}\mathcal{M}_{h,2}$. Let the image of $J_{k,h} \times U_{j,h}$ be $\widetilde{J}_{k,h} \times \widetilde{U}_{j,h}$. The choice of ℓ_1 and ℓ_2 guarantees that $\Pi_{k,j}$ is a diffeomorphism. Let $\{M^i_{\boldsymbol{s}_k} : i = 1,\dots,r_1\}$ be orthonormal vectors spanning the tangent space $T_{\boldsymbol{s}_k}\mathcal{M}_{h,1}$. For any given $\gamma,\theta > 0$, consider the (discrete) set $\widetilde{\Xi}_{h\gamma\theta^{-2/\alpha_1}}(\widetilde{J}_{k,h}) = \{\boldsymbol{t} \in \widetilde{J}_{k,h} : \boldsymbol{t} = \boldsymbol{s}_k + (h\gamma\theta^{-2/\alpha_1})\sum_{i=1}^{r_1} e_i M^i_{\boldsymbol{s}_k}, e_i \in \mathbb{Z}\}$ and let $\Xi_{h\gamma\theta^{-2/\alpha_1}}(J_{k,h}) = \Pi^{-1}_{\boldsymbol{s}_k}(\widetilde{\Xi}_{h\gamma\theta^{-2/\alpha_1}}(\widetilde{J}_{k,h}))$, which is a subset of $J_{k,h}$. Similarly, let $\{M^i_{\boldsymbol{u}_j} : i = 1,\dots,r_2\}$ be orthonormal vectors spanning the tangent space $T_{\boldsymbol{u}_j}\mathcal{M}_{h,2}$ and we discretize $\widetilde{U}_{j,h}$ with $\widetilde{\Xi}_{\gamma\theta^{-2/\alpha_2}}(\widetilde{U}_{j,h}) = \{\boldsymbol{v} \in \widetilde{U}_{j,h} : \boldsymbol{v} = \boldsymbol{u}_j + \gamma\theta^{-2/\alpha_2}\sum_{i=1}^{r_2} e_i M^i_{\boldsymbol{u}_j}, e_i \in \mathbb{Z}\}$ and denote $\Xi_{\gamma\theta^{-2/\alpha_2}}(U_{j,h}) = \Pi^{-1}_{\boldsymbol{u}_j}(\widetilde{\Xi}_{\gamma\theta^{-2/\alpha_2}}(\widetilde{U}_{j,h}))$.

We denote the union of all the grid points by

$$\Gamma_{h,\gamma,\theta} = \cup_{k=1}^{m_h} \cup_{j=1}^{n_h} [\Xi_{h\gamma\theta^{-2/\alpha_1}}(J_{k,h}) \times \Xi_{\gamma\theta^{-2/\alpha_2}}(U_{j,h})] \tag{4.13}$$

$$= [\cup_{k=1}^{m_h} \Xi_{h\gamma\theta^{-2/\alpha_1}}(J_{k,h})] \times [\cup_{j=1}^{n_h} \Xi_{\gamma\theta^{-2/\alpha_2}}(U_{j,h})]. \tag{4.14}$$

For any discrete set $\mathcal{A}$, let $|\mathcal{A}|$ be the cardinality of $\mathcal{A}$. Denote $N^{(1)}_h = |\cup_{k=1}^{m_h} \Xi_{h\gamma\theta^{-2/\alpha_1}}(J_{k,h})|$. Then obviously as $h,\gamma \to 0$ and $\theta \to \infty$,

$$N^{(1)}_h = |\cup_{k=1}^{m_h} \widetilde{\Xi}_{h\gamma\theta^{-2/\alpha_1}}(\widetilde{J}_{k,h})| = O\left(\frac{\sum_{k=1}^{m_h}\mathcal{H}_{r_1}(\widetilde{J}_{k,h})}{(h\gamma\theta^{-2/\alpha_1})^{r_1}}\right) = O\left(\frac{\mathcal{H}_{r_1}(\mathcal{M}_{h,1})}{(h\gamma\theta^{-2/\alpha_1})^{r_1}}\right)$$
$$= O(\theta^{2r_1/\alpha_1}h^{-r_1}\gamma^{-r_1}).$$

Similarly, the cardinality of $\cup_{j=1}^{n_h}\Xi_{\gamma\theta^{-2/\alpha_2}}(U_{j,h})$ is given by

$$N^{(2)}_h := |\cup_{j=1}^{n_h} \Xi_{\gamma\theta^{-2/\alpha_2}}(U_{j,h})| = O(\theta^{2r_2/\alpha_2}\gamma^{-r_2}). \tag{4.15}$$

It is easy to see that $(\mathcal{J}^{\delta}_h \times \mathcal{M}_{h,2}) \cap \Gamma_{h,\gamma,\theta} = [\cup_{k=1}^{m_h}\Xi_{h\gamma\theta^{-2/\alpha_1}}(J^{\delta}_{k,h})] \times [\cup_{j=1}^{n_h}\Xi_{\gamma\theta^{-2/\alpha_2}}(U_{j,h})]$, and for $\delta > 0$ fixed and small,

$$N^{(1)}_{h,\delta} := |\cup_{k=1}^{m_h} \Xi_{h\gamma\theta^{-2/\alpha_1}}(J^{\delta}_{k,h})| = O(N^{(1)}_h) = O(\theta^{2r_1/\alpha_1}h^{-r_1}\gamma^{-r_1}). \tag{4.16}$$

4.4. *Proof of Theorem 3.1*

We focus on the case $k = 2$ in the proof. The proof for $k = 1$ (corresponding to $\mathcal{M}_{2,h} = \emptyset$ when $k = 2$) is omitted since it is similar and simpler. For a random process or field $X(\boldsymbol{t})$, $\boldsymbol{t} \in \mathcal{S} \subset \mathbb{R}^n$ and $\theta \in \mathbb{R}$, we denote

$$\mathbb{P}_X(\theta, \mathcal{S}) = \mathbb{P}(\sup_{t\in\mathcal{S}} X(t) \le \theta),$$
$$\mathbb{Q}_X(\theta, \mathcal{S}) = 1 - \mathbb{P}_X(\theta, \mathcal{S}).$$

With β_h in (3.9), let

$$\theta_{h,z} = \beta_h + \frac{1}{\sqrt{2r_1\log(1/h)}}z. \tag{4.17}$$

With this notation, we can rewrite (3.10) as

$$\lim_{h\to 0}\mathbb{P}_{X_h}(\theta_{h,z},\mathcal{M}_h)=e^{-e^{-z}}.$$

To prove Theorem 3.1, we need to establish a sequence of approximations using the above geometric construction, detailed in Lemmas 4.2–4.7 as follows.

Recall that $I_h(\mathcal{A})=\int_{\mathcal{A}}\|D_{\boldsymbol{t},h}P_{\boldsymbol{t}}\|_{r_1}d\mathcal{H}_r(\boldsymbol{t})$ for any Borel subset $\mathcal{A}\subset\mathcal{M}_h$. In the following lemmas we consider θ as a large number with $\theta=\theta_{h,z}$ as a special case in mind.

Lemma 4.2. *For any $\epsilon>0$, there exist $\theta_0>0$ such that for all $\theta\geq\theta_0$, $0<h\leq h_0$, and $J_k\in\{J_{k,h},J_{k,h}^{\delta},J_{k,h}^{-\delta}\}$ with $1\leq k\leq m_h$, we have for some $\epsilon_{k,h}$ with $|\epsilon_{k,h}|\leq\epsilon$,*

$$\frac{\mathbb{Q}_{X_h}(\theta,J_k\times\mathcal{M}_{h,2})}{\theta^{2(r_1/\alpha_1+r_2/\alpha_2)}\Psi(\theta)}=(1+\epsilon_{k,h})h^{-r_1}\;H_{R,\boldsymbol{\alpha}}I_h(J_k\times\mathcal{M}_{h,2}).\tag{4.18}$$

Proof. For $J_k\in\{J_{k,h},J_{k,h}^{\delta},J_{k,h}^{-\delta}\}$, denote $\overline{J}_k=\{\boldsymbol{t}_{(1)}/h:\boldsymbol{t}_{(1)}\in J_k\}$ such that $\overline{J}_k\times\mathcal{M}_{h,2}\subset\overline{\mathcal{M}}_h$. Then notice that $\overline{J}_k$ has a positive diameter and r_1-dimensional Hausdorff measure. Recall that $\xi_h(\boldsymbol{t})=(h\boldsymbol{t}_{(1)}^T,\boldsymbol{t}_{(2)}^T)^T$ for $\boldsymbol{t}=(\boldsymbol{t}_{(1)}^T,\boldsymbol{t}_{(2)}^T)^T\in\overline{J}_k\times\mathcal{M}_{h,2}$ and the Gaussian field $\overline{X}_h(\boldsymbol{t})=X_h(\xi_h(\boldsymbol{t}))$ is locally-$(E,\boldsymbol{\alpha},D_{\xi_h(\boldsymbol{t}),h})$-stationary on $\overline{J}_k\times\mathcal{M}_{h,2}$. Let $\overline{I}_h(\mathcal{A})=\int_{\mathcal{A}}\|D_{\xi_h(\boldsymbol{t}),h}P_{\boldsymbol{t}}\|_{r_1}d\mathcal{H}_r(\boldsymbol{t})$ for any Borel subset $\mathcal{A}\subset\overline{\mathcal{M}}_h$. Then using Theorem 2.1, we obtain that

$$\frac{\mathbb{Q}_{\overline{X}_h}(\theta,\overline{J}_k\times\mathcal{M}_{h,2})}{\theta^{2(r_1/\alpha_1+r_2/\alpha_2)}\Psi(\theta)}=H_{R,\boldsymbol{\alpha}}\overline{I}_h(\overline{J}_k\times\mathcal{M}_{h,2})(1+o(1)),$$

where the $o(1)$-term is uniform in $1\leq k\leq m_h$ and $0<h\leq h_0$, because of the uniformity in assumptions (B1)–(B3), as well as the fact the Jacobian matrix of the diffeomorphism τ we establish in the proof of Theorem 2.1 is locally Lipschitz continuous with the Lipschitz constant only depending on the reach of the manifold as shown in (4.4). Noticing that $\overline{I}_h(\overline{J}_k\times\mathcal{M}_{h,2})=h^{-r_1}I_h(J_k\times\mathcal{M}_{h,2})$, we then get the desired result. □

Lemma 4.3. *For any $\epsilon>0$, there exist $\gamma_0>0$, $\theta_0>0$ such that for all $\gamma\leq\gamma_0$, $\theta\geq\theta_0$, $0<h\leq h_0$, and $J_k\in\{J_{k,h},J_{k,h}^{\delta},J_{k,h}^{-\delta}\}$ with $1\leq k\leq m_h$, we have for some $\epsilon_{k,h}$ with $|\epsilon_{k,h}|\leq\epsilon$,*

$$\frac{\mathbb{Q}_{X_h}(\theta,(J_k\times\mathcal{M}_{h,2})\cap\Gamma_{h,\gamma,\theta})}{\theta^{2(r_1/\alpha_1+r_2/\alpha_2)}\Psi(\theta)}=(1+\epsilon_{k,h})h^{-r_1}\;\widetilde{H}_{R,\boldsymbol{\alpha}}(\gamma)I_h(J_k\times\mathcal{M}_{h,2}),\tag{4.19}$$

where $\widetilde{H}_{R,\boldsymbol{\alpha}}(\gamma)$ only depends on γ such that $\widetilde{H}_{R,\boldsymbol{\alpha}}(\gamma)\to H_{R,\boldsymbol{\alpha}}$ as $\gamma\to 0$.

Proof. The proof is similar to that of Lemma 4.2. The main difference is that, instead of applying Theorem 2.1, we use Lemma A.3 in the appendix. Note that in order to apply Lemma A.3, one needs to find a diffeomorphism $\psi_{k,j}$ between $J_k\times U_{j,h}$ and its preimage in $\mathbb{R}^r$, for each $k=1,\ldots,m_h$ and $j=1,\ldots,n_h$. This diffeomorphism is constructed in the same way as shown at the beginning of the proof of Theorem 2.1. By using Lemma A.3 and following the proofs of Theorem 2.1 and Lemma 4.3 we obtain the result stated in this lemma. □

Lemma 4.4. *For $\theta=\theta_{h,z}$ given in (4.17) with any fixed z, we have that as $h\to 0$,*

$$h^{-r_1}\theta^{2(r_1/\alpha_1+r_2/\alpha_2)}\Psi(\theta)=\frac{e^{-z}}{H_{R,\boldsymbol{\alpha}}I_h(\mathcal{M}_h)}(1+o(1))=O(1).\tag{4.20}$$

Proof. Observe that the first equality in (4.20) follows from a direct calculation using (4.17). Next we show (4.20) is bounded. Recall that $\|D_{t,h}P_t\|_r = [\det(P_t^T D_{t,h}^T D_{t,h} P_t)]^{1/2}$ (see (4.2)), where the columns of P_t are orthonormal and span the tangent space $T_t\mathcal{M}_h$. Notice that for any $\boldsymbol{a} \in \mathbb{S}^{r-1}$ we have $P_t\boldsymbol{a} \in \mathbb{S}^{n-1}$. Therefore

$$\begin{aligned}\lambda_{\max}(P_t^T D_{t,h}^T D_{t,h} P_t) &= \sup_{\boldsymbol{a}\in\mathbb{S}^{r-1}} \boldsymbol{a}^T(P_t^T D_{t,h}^T D_{t,h} P_t)\boldsymbol{a} \le \sup_{\boldsymbol{b}\in\mathbb{S}^{n-1}} \boldsymbol{b}^T(D_{t,h}^T D_{t,h})\boldsymbol{b}\\ &= \lambda_{\max}(D_{t,h}^T D_{t,h}).\end{aligned}$$

Similarly,

$$\begin{aligned}\lambda_{\min}(P_t^T D_{t,h}^T D_{t,h} P_t) &= \inf_{\boldsymbol{a}\in\mathbb{S}^{r-1}} \boldsymbol{a}^T(P_t^T D_{t,h}^T D_{t,h} P_t)\boldsymbol{a} \ge \inf_{\boldsymbol{b}\in\mathbb{S}^{n-1}} \boldsymbol{b}^T(D_{t,h}^T D_{t,h})\boldsymbol{b}\\ &= \lambda_{\min}(D_{t,h}^T D_{t,h}).\end{aligned}$$

It then follows from (4.2) that

$$[\lambda_{\min}(D_{t,h}^T D_{t,h})]^{r/2} \le \|D_{t,h}P_t\|_r \le [\lambda_{\max}(D_{t,h}^T D_{t,h})]^{r/2}. \tag{4.21}$$

The left-hand side in (4.20) is bounded because with assumption (B2) we have

$$\begin{aligned}0 &< \inf_{0<h\le h_0, t\in\mathcal{M}_h} [\lambda_{\min}(D_{t,h}^T D_{t,h})]^{r/2} \inf_{0<h\le h_0} \mathcal{H}_r(\mathcal{M}_h)\\ &\le \inf_{0<h\le h_0} I_h(\mathcal{M}_h) \le \sup_{0<h\le h_0} I_h(\mathcal{M}_h)\\ &\le \sup_{0<h\le h_0, t\in\mathcal{M}_h} [\lambda_{\max}(D_{t,h}^T D_{t,h})]^{r/2} \sup_{0<h\le h_0} \mathcal{H}_r(\mathcal{M}_h) < \infty. \quad \square\end{aligned}$$

Denote $\mathcal{J}_h^\delta = \bigcup_{k\le m_h} J_{k,h}^\delta$. Recall that $\mathcal{M}_h = \mathcal{M}_{h,1} \times \mathcal{M}_{h,2}$. Approximating $\mathcal{M}_h$ by $\mathcal{J}_h^\delta \times \mathcal{M}_{h,2}$ leads to the approximation of $\mathbb{Q}_{X_h}(\theta, \mathcal{M}_h)$ by $\mathbb{Q}_{X_h}(\theta, \mathcal{J}_h^\delta \times \mathcal{M}_{h,2})$. The volume of $\bigcup_{k\le m_h} J_{k,h}^{-\delta}$, i.e., the difference between the volumes of $\mathcal{M}$ and $\mathcal{J}_h^\delta$, is of the order $O(\delta)$ uniformly in h. As the next lemma shows, the order of the difference $\mathbb{Q}_{X_h}(\theta, \mathcal{M}_h) - \mathbb{Q}_{X_h}(\theta, \mathcal{J}_h^\delta \times \mathcal{M}_{h,2})$ turns out to be of the same order.

Lemma 4.5. *With $\theta = \theta_{h,z}$ given in (4.17), there exists a positive constant $C < \infty$ such that for all δ and h small enough,*

$$0 < \mathbb{P}_{X_h}(\theta, \mathcal{J}_h^\delta \times \mathcal{M}_{h,2}) - \mathbb{P}_{X_h}(\theta, \mathcal{M}_h) \le C\delta, \tag{4.22}$$

and

$$0 < \sum_{k=1}^{m_h} \mathbb{Q}_{X_h}(\theta, J_{k,h} \times \mathcal{M}_{h,2}) - \sum_{k=1}^{m_h} \mathbb{Q}_{X_h}(\theta, J_{k,h}^\delta \times \mathcal{M}_{h,2}) \le C\delta. \tag{4.23}$$

Proof. Using (3.2) and (4.21), we have that

$$\sup_{0<h\le h_0, t\in\mathcal{M}_h} \|D_{t,h}P_t\|_r \le \sup_{0<h\le h_0, t\in\mathcal{M}_h} [\lambda_{\max}(D_{t,h}^T D_{t,h})]^{r/2} =: C_1 < \infty. \tag{4.24}$$

Also note that there exists a positive constant $C_2 < \infty$ such that $\max_{1\le k\le m_h} \mathcal{H}_r(J_{k,h}^{-\delta} \times \mathcal{M}_h) \le C_2\delta h^{r_1}$ for all $h \in (0, h_0]$. Our construction of the partition of the $\mathcal{M}_h$ guarantees that there exists a positive constant $C_3 < \infty$ such that $m_h \le C_3 h^{-r_1}$. Therefore

$$\sum_{k=1}^{m_h} I_h(J_{k,h}^{-\delta} \times \mathcal{M}_{h,2}) \le m_h \sup_{0<h\le h_0, t\in\mathcal{M}_h} \|D_{t,h}P_t\|_r \max_{1\le k\le m_h} \mathcal{H}_r(J_{k,h}^{-\delta} \times \mathcal{M}_h) \le C_1C_2C_3\delta. \tag{4.25}$$

Using Lemma 4.2, for any $\epsilon > 0$, we have for h small enough that

$$
\begin{aligned}
0 &\le \mathbb{Q}_{X_h}(\theta_{h,z}, \mathcal{M}_h) - \mathbb{Q}_{X_h}(\theta_{h,z}, \mathcal{J}_h^{\delta} \times \mathcal{M}_{h,2}) \\
&\le \sum_{k=1}^{m_h} \mathbb{Q}_{X_h}(\theta_{h,z}, J_{k,h}^{-\delta} \times \mathcal{M}_{h,2}) \\
&\le (1+\epsilon) h^{-r_1} H_{R,\boldsymbol{\alpha}} \theta_{h,z}^{2(r_1/\alpha_1 + r_2/\alpha_2)} \Psi(\theta) \sum_{k=1}^{m_h} I_h(J_{k,h}^{-\delta} \times \mathcal{M}_{h,2}).
\end{aligned}
$$

Then (4.23) follows from Lemma 4.4 and (4.25). Also (4.22) holds because

$$
0 < \mathbb{P}_{X_h}(\theta, \mathcal{J}_h^{\delta} \times \mathcal{M}_{h,2}) - \mathbb{P}_{X_h}(\theta, \mathcal{M}_h) \le \sum_{k=1}^{m_h} \mathbb{Q}_{X_h}(\theta_{h,z}, J_{k,h}^{-\delta} \times \mathcal{M}_{h,2}). \quad \square
$$

With $\Gamma_{h,\gamma,\theta}$ given in (4.13), $(\mathcal{J}_h^{\delta} \times \mathcal{M}_{h,2}) \cap \Gamma_{h,\gamma,\theta}$ represents a set of grid points over $\mathcal{J}_h^{\delta} \times \mathcal{M}_{h,2}$. Next we show that the excursion probabilities over these two sets are close, by choosing both h and γ sufficiently small.

Lemma 4.6. *With $\theta = \theta_{h,z}$ given in (4.17), we have that*

$$
\mathbb{P}_{X_h}(\theta, \mathcal{J}_h^{\delta} \times \mathcal{M}_{h,2}) = \mathbb{P}_{X_h}(\theta, (\mathcal{J}_h^{\delta} \times \mathcal{M}_{h,2}) \cap \Gamma_{h,\gamma,\theta}) + o(1) \tag{4.26}
$$

and

$$
\sum_{k=1}^{m_h} \mathbb{Q}_{X_h}(\theta, J_{k,h}^{\delta} \times \mathcal{M}_{h,2}) = \sum_{k=1}^{m_h} \mathbb{Q}_{X_h}(\theta, (J_{k,h}^{\delta} \times \mathcal{M}_{h,2}) \cap \Gamma_{h,\gamma,\theta}) + o(1), \tag{4.27}
$$

as $\gamma, h \to 0$.

Proof. Lemmas 4.2 and 4.3 imply that for any $\epsilon > 0$, there exist $\gamma_0 > 0$ and $\theta_0 > 0$ such that for all $0 < h \le h_0$, $\gamma \le \gamma_0$ and $\theta \ge \theta_0$,

$$
\begin{aligned}
0 &\le \mathbb{Q}_{X_h}(\theta, J_{k,h}^{\delta} \times \mathcal{M}_{h,2}) - \mathbb{Q}_{X_h}(\theta, (J_{k,h}^{\delta} \times \mathcal{M}_{h,2}) \cap \Gamma_{h,\gamma,\theta}) \\
&\le \epsilon h^{-r_1} \theta^{2(r_1/\alpha_1 + r_2/\alpha_2)} \Psi(\theta) H_{R,\boldsymbol{\alpha}} I_h(J_{k,h}^{\delta} \times \mathcal{M}_{h,2}).
\end{aligned}
$$

As a result,

$$
\begin{aligned}
0 &\le \mathbb{Q}_{X_h}(\theta, \mathcal{J}_h^{\delta} \times \mathcal{M}_{h,2}) - \mathbb{Q}_{X_h}(\theta, (\mathcal{J}_h^{\delta} \times \mathcal{M}_{h,2}) \cap \Gamma_{h,\gamma,\theta}) \\
&\le \sum_{k=1}^{m_h} \Big[\mathbb{Q}_{X_h}(\theta, J_{k,h}^{\delta} \times \mathcal{M}_{h,2}) - \mathbb{Q}_{X_h}(\theta, (J_{k,h}^{\delta} \times \mathcal{M}_{h,2}) \cap \Gamma_{h,\gamma,\theta}) \Big] \\
&\le \epsilon h^{-r_1} \theta^{2(r_1/\alpha_1 + r_2/\alpha_2)} \Psi(\theta) H_{R,\boldsymbol{\alpha}} I_h\big(\mathcal{J}_h^{\delta} \times \mathcal{M}_{h,2}\big) \\
&\le \epsilon h^{-r_1} \theta^{2(r_1/\alpha_1 + r_2/\alpha_2)} \Psi(\theta) H_{R,\boldsymbol{\alpha}} I_h(\mathcal{M}_h).
\end{aligned}
$$

Then (4.26) and (4.27) immediately follow from (4.20). $\square$

Recall that $(\mathcal{J}_h^{\delta} \times \mathcal{M}_{h,2}) \cap \Gamma_{h,\gamma,\theta}$ gives a set of dense grid points in $\mathcal{J}_h^{\delta} \times \mathcal{M}_{h,2}$. For any $1 \le k \le m_h$, denote the set $T_k^{h,\gamma,\theta} = (J_{k,h}^{\delta} \times \mathcal{M}_{h,2}) \cap \Gamma_{h,\gamma,\theta}$. Define a probability measure $\widetilde{\mathbb{P}}$ such that under $\widetilde{\mathbb{P}}$ the vectors $(X_h(\boldsymbol{t}) : \boldsymbol{t} \in T_k^{h,\gamma,\theta})$ and $(X_h(\boldsymbol{t}') : \boldsymbol{t}' \in T_{k'}^{h,\gamma,\theta})$ are independent for $k \neq k'$. In other words, $\widetilde{\mathbb{P}}_{X_h}(\theta, (\mathcal{J}_h^{\delta} \times \mathcal{M}_{h,2}) \cap \Gamma_{h,\gamma,\theta}) = \prod_{k \le m_h} \mathbb{P}_{X_h}(\theta, (J_{k,h}^{\delta} \times \mathcal{M}_{h,2}) \cap \Gamma_{h,\gamma,\theta})$. As the next lemma shows, the probability $\mathbb{P}_{X_h}(\theta, (\mathcal{J}_h^{\delta} \times \mathcal{M}_{h,2}) \cap \Gamma_{h,\gamma,\theta})$ can be approximated by using the probability measure $\widetilde{\mathbb{P}}$, when h, γ and δ are small.

Lemma 4.7. *For $\delta > 0$ fixed and small enough, there exists $\gamma = \gamma(h) \to 0$ as $h \to 0$, such that with $\theta = \theta_{h,z}$ given in (4.17), we have*

$$\mathbb{P}_{X_h}(\theta, (\mathcal{J}_h^\delta \times \mathcal{M}_{h,2}) \cap \Gamma_{h,\gamma,\theta}) = \prod_{k \le m_h} \mathbb{P}_{X_h}(\theta, (J_{k,h}^\delta \times \mathcal{M}_{h,2}) \cap \Gamma_{h,\gamma,\theta}) + o(1). \tag{4.28}$$

Proof. Denote $\boldsymbol{t} = (\boldsymbol{t}_{(1)}^T, \boldsymbol{t}_{(2)}^T)^T$ and $\boldsymbol{t}' = (\boldsymbol{t}'^T_{(1)}, \boldsymbol{t}'^T_{(2)})^T$, where $\boldsymbol{t}_{(1)}, \boldsymbol{t}'_{(1)} \in \mathbb{R}^{e_1}$ and $\boldsymbol{t}_{(2)}, \boldsymbol{t}'_{(2)} \in \mathbb{R}^{e_2}$. For $\boldsymbol{t} \in T_k^{h,\gamma,\theta}$ and $\boldsymbol{t}' \in T_{k'}^{h,\gamma,\theta}$ with $k \neq k'$, we have $\boldsymbol{t}_{(1)} \in J_{k,h}^\delta$ and $\boldsymbol{t}'_{(1)} \in J_{k',h}^\delta$, and hence for all $0 < h \le h_0$, by (4.12) we have

$$\|\xi_h^{-1}(\boldsymbol{t}) - \xi_h^{-1}(\boldsymbol{t}')\| \ge \|(\boldsymbol{t}_{(1)} - \boldsymbol{t}'_{(1)})/h\| \ge (2h\delta)/h = 2\delta > 0.$$

Let $r_h(\boldsymbol{t}_1, \boldsymbol{t}_2)$ be the covariance between $X_h(\boldsymbol{t}_1)$ and $X_h(\boldsymbol{t}_2)$, for $\boldsymbol{t}_1, \boldsymbol{t}_2 \in \mathcal{M}_h$. By the definition of $\bar{r}_h$ we can write $r_h(\boldsymbol{t}_1, \boldsymbol{t}_2) = \bar{r}_h(\xi_h^{-1}(\boldsymbol{t}_1), \xi_h^{-1}(\boldsymbol{t}_2))$. Then assumption (B3) implies that there exists $\eta = \eta(\delta) > 0$, such that

$$\sup_{0<h\le h_0} \sup_{k \neq k'} \sup_{\boldsymbol{t} \in T_k^{h,\gamma,\theta}} \sup_{\boldsymbol{t}' \in T_{k'}^{h,\gamma,\theta}} |r_h(\boldsymbol{t}, \boldsymbol{t}')| < \eta < 1. \tag{4.29}$$

By Lemma 4.1 of Berman [6] (aslo see Lemma A4 of Bickel and Rosenblatt [7]), we have

$$\begin{aligned}
&\left|\mathbb{P}_{X_h}(\theta, (\mathcal{J}_h^\delta \times \mathcal{M}_{h,2}) \cap \Gamma_{h,\gamma,\theta}) - \widetilde{\mathbb{P}}_{X_h}(\theta, (\mathcal{J}_h^\delta \times \mathcal{M}_{h,2}) \cap \Gamma_{h,\gamma,\theta})\right| \\
\le& 8 \sum_{1 \le k \neq k' \le m_h} \sum_{\boldsymbol{t} \in T_k^{h,\gamma,\theta}} \sum_{\boldsymbol{t}' \in T_{k'}^{h,\gamma,\theta}} \int_0^{|r_h(\boldsymbol{t},\boldsymbol{t}')|} \frac{1}{2\pi(1-\lambda^2)^{1/2}} \exp\left(-\frac{\theta^2}{1+\lambda}\right) d\lambda \\
\le& \sum_{1 \le k \neq k' \le m_h} \sum_{\boldsymbol{t} \in T_k^{h,\gamma,\theta}} \sum_{\boldsymbol{t}' \in T_{k'}^{h,\gamma,\theta}} \zeta_h(\boldsymbol{t}, \boldsymbol{t}'),
\end{aligned} \tag{4.30}$$

where

$$\zeta_h(\boldsymbol{t}, \boldsymbol{t}') = \frac{4|r_h(\boldsymbol{t}, \boldsymbol{t}')|}{\pi(1-\eta^2)^{1/2}} \exp\left(-\frac{\theta^2}{1 + |r_h(\boldsymbol{t}, \boldsymbol{t}')|}\right).$$

We take $\gamma = [v(h^{-1})]^{(1/(3r_1+3r_2))}$. Let ω be such that $0 < \omega < \frac{2}{(1+\eta)} - 1$, and define

$$\mathcal{G}_{h,\gamma,\theta}^{(1)} = \{(\boldsymbol{t}, \boldsymbol{t}') \in T_k^{h,\gamma,\theta} \times T_{k'}^{h,\gamma,\theta} : \|\boldsymbol{t}_{(1)} - \boldsymbol{t}'_{(1)}\| < h(N_{h,\delta}^{(1)})^{\omega/r_1} \gamma \theta^{-2/\alpha_1}, 1 \le k \neq k' \le m_h\},$$
$$\mathcal{G}_{h,\gamma,\theta}^{(2)} = \{(\boldsymbol{t}, \boldsymbol{t}') \in T_k^{h,\gamma,\theta} \times T_{k'}^{h,\gamma,\theta} : \|\boldsymbol{t}_{(1)} - \boldsymbol{t}'_{(1)}\| \ge h(N_{h,\delta}^{(1)})^{\omega/r_1} \gamma \theta^{-2/\alpha_1}, 1 \le k \neq k' \le m_h\},$$

where $N_{h,\delta}^{(1)}$ is given in (4.16). Then the triple sum on the right-hand side of (4.30) can be written as

$$\sum_{(\boldsymbol{t},\boldsymbol{t}') \in \mathcal{G}_{h,\gamma,\theta}^{(1)}} \zeta_h(\boldsymbol{t}, \boldsymbol{t}') + \sum_{(\boldsymbol{t},\boldsymbol{t}') \in \mathcal{G}_{h,\gamma,\theta}^{(2)}} \zeta_h(\boldsymbol{t}, \boldsymbol{t}'). \tag{4.31}$$

Note that the cardinality of $\mathcal{G}_{h,\gamma,\theta}^{(1)}$ is of the order $O((N_{h,\delta}^{(1)})^{\omega+1}(N_h^{(2)})^2)$, where $N_h^{(2)}$ is given in (4.15). Hence for the first sum in (4.31) we have

$$\begin{aligned}
\sum_{(\boldsymbol{t},\boldsymbol{t}') \in \mathcal{G}_{h,\gamma,\theta}^{(1)}} \zeta_h(\boldsymbol{t}, \boldsymbol{t}') =& O\left((N_{h,\delta}^{(1)})^{\omega+1}(N_h^{(2)})^2 \exp\left\{-\frac{\theta^2}{1+\eta}\right\}\right) \\
=& O\left(\left(\frac{\theta^{2r_1/\alpha_1}}{h^{r_1}\gamma^{r_1}}\right)^{1+\omega} \frac{\theta^{4r_2/\alpha_2}}{\gamma^{2r_2}} \exp\left\{-\frac{\theta^2}{1+\eta}\right\}\right)
\end{aligned}$$

$$
\begin{aligned}
&= O\left(\left(\frac{(\log\frac{1}{h})^{r_1/\alpha_1+2r_2/[\alpha_2(1+\omega)]}}{h^{r_1}\gamma^{r_1+2r_2/(1+\omega)}}\right)^{1+\omega}\exp\left\{-\frac{2r_1\log\frac{1}{h}}{1+\eta}\right\}\right)\\
&= O\left(h^{\frac{2r_1}{1+\eta}-r_1(1+\omega)}\left(\log\frac{1}{h}\right)^{\frac{(1+\omega)r_1}{\alpha_1}+\frac{2r_2}{\alpha_2}}\left(v(\tfrac{1}{h})\right)^{-\frac{(1+\omega)r_1+2r_2}{3r_1+3r_2}}\right)\\
&= o(1) \quad \text{as} \quad h\to 0. \qquad (4.32)
\end{aligned}
$$

Now we consider the second sum in (4.31). Because $(1+|r_h(\boldsymbol{t},\boldsymbol{t}')|)^{-1}\geq 1-|r_h(\boldsymbol{t},\boldsymbol{t}')|$, we have

$$
\zeta_h(\boldsymbol{t},\boldsymbol{t}') \leq \frac{4|r_h(\boldsymbol{t},\boldsymbol{t}')|}{\pi(1-\eta^2)^{1/2}}\exp\left(-(1-|r_h(\boldsymbol{t},\boldsymbol{t}')|)\theta^2\right).
$$

Since $\theta^2 = O(\log\frac{1}{h})$, with Q given in (3.3) and by using (3.4) we have that for h small enough,

$$
\begin{aligned}
&\sup_{(\boldsymbol{t},\boldsymbol{t}')\in\mathcal{G}^{(2)}_{h,\gamma,\theta}} |r_h(\boldsymbol{t},\boldsymbol{t}')|\theta^2 \leq Q((N^{(1)}_{h,\delta})^{\omega/r_1}\gamma\theta^{-2/\alpha_1})\theta^2\\
&\leq \frac{v((N^{(1)}_{h,\delta})^{\omega/r_1}\gamma\theta^{-2/\alpha_1})\theta^2}{(\log((N^{(1)}_{h,\delta})^{\omega/r_1}\gamma\theta^{-2/\alpha_1}))^{2(r_1/\alpha_1+r_2/\alpha_2)}} \to 0.
\end{aligned}
$$

Also notice that $\exp(-\theta^2) = O(h^{2r_1})$, and hence $\exp\left(-(1-|r_h(\boldsymbol{t},\boldsymbol{t}')|)\theta^2\right) = O(h^{2r_1})$ uniformly in $(\boldsymbol{t},\boldsymbol{t}')\in\mathcal{G}^{(2)}_{h,\gamma,\theta}$. Consequently, when h is sufficiently small, there exists a constant $C>0$ such that

$$
\sup_{(\boldsymbol{t},\boldsymbol{t}')\in\mathcal{G}^{(2)}_{h,\gamma,\theta}} \zeta_h(\boldsymbol{t},\boldsymbol{t}') \leq Ch^{2r_1}\frac{v((N^{(1)}_{h,\delta})^{\omega/r_1}\gamma\theta^{-2/\alpha_1})}{[\log((N^{(1)}_{h,\delta})^{\omega/r_1}\gamma\theta^{-2/\alpha_1})]^{2r_1/\alpha_1+2r_2/\alpha_2}}. \qquad (4.33)
$$

Therefore it follows from (4.15) and (4.16) that

$$
\begin{aligned}
&\sum_{(\boldsymbol{t},\boldsymbol{t}')\in\mathcal{G}^{(2)}_{h,\gamma,\theta}} \zeta_h(\boldsymbol{t},\boldsymbol{t}')\\
&= O\left(h^{2r_1}(N^{(1)}_{h,\delta})^2(N^{(2)}_h)^2\frac{v((N^{(1)}_{h,\delta})^{\omega/r_1}\gamma\theta^{-2/\alpha_1})}{[\log((N^{(1)}_{h,\delta})^{\omega/r_1}\gamma\theta^{-2/\alpha_1})]^{2r_1/\alpha_1+2r_2/\alpha_2}}\right)\\
&= O\left(\frac{(\log\frac{1}{h})^{2r_1/\alpha_1+2r_2/\alpha_2}v((N^{(1)}_{h,\delta})^{\omega/r_1}\gamma\theta^{-2/\alpha_1})}{\left[\log\left(h^{-\omega}\left((\log\frac{1}{h})^{1/\alpha_1}v(\frac{1}{h})^{-1/(3r_1+3r_2)}\right)^{\omega-1}\right)\right]^{2r_1/\alpha_1+2r_2/\alpha_2}\left(v(\frac{1}{h})\right)^{2/3}}\right)\\
&= o(1) \quad \text{as} \quad h\to 0. \qquad (4.34)
\end{aligned}
$$

Combining (4.30), (4.32) and (4.34), we obtain (4.28). □

Proof of Theorem 3.1. We choose the same $\gamma=\gamma(h)$ as in Lemma 4.7, and use $\theta=\theta_{h,z}$ given in (4.17). For any arbitrarily small (fixed) $\delta>0$, by using (4.22), (4.26), and (4.28), we have that as $h \to 0$,

$$\begin{aligned}\mathbb{P}_{X_h}(\theta, \mathcal{M}_h) &= \prod_{k\le m_h} \mathbb{P}_{X_h}(\theta, (J^{\delta}_{k,h} \times \mathcal{M}_{h,2}) \cap \Gamma_{h,\gamma,\theta}) + o(1) + O(\delta)\\ &= \exp\Big\{\sum_{k\le m_h} \log\Big(1 - \mathbb{Q}_{X_h}(\theta, (J^{\delta}_{k,h} \times \mathcal{M}_{h,2}) \cap \Gamma_{h,\gamma,\theta})\Big)\Big\} + o(1) + O(\delta)\\ &= \exp\Big\{-(1 + o(1)) \sum_{k\le m_h} \mathbb{Q}_{X_h}(\theta, (J^{\delta}_{k,h} \times \mathcal{M}_{h,2}) \cap \Gamma_{h,\gamma,\theta})\Big\} + o(1) + O(\delta).\end{aligned}$$

Then by using (4.27), (4.23), and (4.18), we get

$$\mathbb{P}_{X_h}(\theta, \mathcal{M}_h) = \exp\big\{-(1+o(1))h^{-r_1}\, \theta^{2(r_1/\alpha_1+r_2/\alpha_2)}\, \Psi(\theta) H_{R,\boldsymbol{\alpha}} I_h(\mathcal{M}_h)+O(\delta)\big\}+o(1)+O(\delta).$$

The proof is completed by noticing (4.20). □

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The comments of an associate editor and two anonymous referees have led to significant improvements of the paper and are gratefully acknowledged. This work is partially supported by National Science Foundation, USA grant DMS 1821154.

Appendix

In this appendix, we collect some miscellaneous results that are straightforward extensions from some existing results in the literature, and have been used in our proofs, as well as some basic facts about manifolds and geometric integration theory.

A.1. Miscellaneous results

For an integer $\ell > 0$ and $\gamma > 0$, let $C(\ell, \gamma) = \{t\gamma : t \in [0, \ell]^n \cap \mathbb{Z}^n\}$. Given a structure $(E, \boldsymbol{\alpha})$, let $H_{E,\boldsymbol{\alpha}}(\ell, \gamma) = H_{E,\boldsymbol{\alpha}}(C(\ell, \gamma))$, where $H_{E,\boldsymbol{\alpha}}$ is defined in (2.1) for subsets of $\mathbb{R}^n$, and

$$H_{E,\boldsymbol{\alpha}}(\gamma) = \lim_{\ell\to\infty} \frac{H_{E,\boldsymbol{\alpha}}(\ell, \gamma)}{\ell^n}.$$

The existence of this limit follows from Pickands [32]. Using the factorization lemma (Lemma 6.4 of Piterbarg [34]) and Theorem B3 of Bickel and Rosenblatt [8], we have

Lemma A.1. $H_{E,\boldsymbol{\alpha}} = \lim_{\gamma\to 0} \frac{H_{E,\boldsymbol{\alpha}}(\gamma)}{\gamma^n}$.

Let $\Gamma_{E,\boldsymbol{\alpha}}(\gamma, u) = \{(x_1, \dots, x_k) \in \mathbb{R}^n : x_i = \gamma u^{-2/\alpha_i} \boldsymbol{\ell}_i, \boldsymbol{\ell}_i \in \mathbb{Z}^{e_i}, i = 1, \dots, k\}$. The following result extends Lemma 4.2 in Qiao and Polonik [37] from assuming a simple structure with $E = \{n\}$ and a scalar $0 < \boldsymbol{\alpha} = \alpha \le 2$ to a more general structure. The proof uses similar ideas and therefore is omitted. Also see Lemma 3 of Bickel and Rosenblatt [8], and Lemma 7.1 of Piterbarg [34].

Lemma A.2. *Given a structure* $(E, \boldsymbol{\alpha})$, *let* $X(\boldsymbol{t})$, $\boldsymbol{t} \in \mathbb{R}^n$, *be a centered homogeneous Gaussian field with covariance function* $r(\boldsymbol{t}) = \mathbb{E}(X(\boldsymbol{t}+\boldsymbol{s})X(\boldsymbol{s})) = 1-|\boldsymbol{t}|_{E,\boldsymbol{\alpha}}(1+(1))$, *as* $\boldsymbol{t} \to 0$. *Then there exists* $\delta_0 > 0$ *such that for any compact Jordan measurable set* A *of positive* n*-dimensional Lebesgue measure with diameter not exceeding* δ_0, *the following asymptotic behavior occurs:*

$$\mathbb{P}\left(\sup_{\boldsymbol{t}\in A_{\gamma,u}} X(\boldsymbol{t}) > u\right) = \frac{H_{E,\boldsymbol{\alpha}}(\gamma)}{\gamma^n}\mathcal{H}_n(A)\prod_{i=1}^{k} u^{2e_i/\alpha_i}\,\Psi(u)(1+o(1)),$$

as $u \to \infty$, *where* $A_{\gamma,u} = A \cap \Gamma_{E,\boldsymbol{\alpha}}(\gamma, u)$.

The next theorem is similar to Theorem 7.1 of Piterbarg [34], except that the supremum is over a dense grid. The proof is similar, where one needs to replace the role of Lemma 7.1 of Piterbarg [34] by our Lemma A.2.

Theorem A.1. *Let* $X(\boldsymbol{t})$, $\boldsymbol{t} \in A \subset \mathbb{R}^n$ *be a locally-*$(E, \boldsymbol{\alpha}, D_{\boldsymbol{t}})$*-stationary Gaussian field with zero mean, where* A *is a compact Jordan set of positive* n*-dimensional Lebesgue measure. Assume also that the matrix-valued function* $D_{\boldsymbol{t}}$ *is continuous in* $\boldsymbol{t}$ *and non-singular everywhere on* A. *Then if* $r_X(\boldsymbol{t}, \boldsymbol{s}) < 1$ *for all* $\boldsymbol{t}, \boldsymbol{s}$ *from* A, $\boldsymbol{t} \neq \boldsymbol{s}$, *the following asymptotic behavior occurs:*

$$\mathbb{P}\left(\sup_{\boldsymbol{t}\in A_{\gamma,u}} X(\boldsymbol{t}) > u\right) = \frac{H_{E,\boldsymbol{\alpha}}(\gamma)}{\gamma^n}\int_A |\det D_{\boldsymbol{t}}|d\boldsymbol{t}\prod_{i=1}^{k} u^{2e_i/\alpha_i}\,\Psi(u)(1+o(1)),$$

as $u \to \infty$, *where* $A_{\gamma,u} = A \cap \Gamma_{E,\boldsymbol{\alpha}}(\gamma, u)$.

The following lemma is analogous to Lemma 4.1 with the index set being a grid. The proof is also similar to that of Lemma 4.1, except that in the proof we use Theorem A.1 to replace the role of Theorem 7.1 of Piterbarg [34].

Lemma A.3. *Suppose that the conditions in Theorem 2.1 hold. For a subset* $V \subset \mathcal{M}$, *suppose that there exist an open set* $G \subset \mathbb{R}^r$ *and a diffeomorphism* $\psi : G \to V$, *where the component functions of the Jacobian matrix* J_ψ *of* ψ *are uniformly continuous. For any subset* $U \subset V$, *if* $\Omega := \psi^{-1}(U)$ *is a compact Jordan set of positive* r*-dimensional Lebesgue measure, then as* $u \to \infty$,

$$\mathbb{P}\left(\sup_{\boldsymbol{t}\in M_{\gamma,u}} X(\boldsymbol{t}) > u\right) = \frac{H_{R,\boldsymbol{\alpha}}(\gamma)}{\gamma^r}\int_U \prod_{j=1}^{k} \|D_{j,\boldsymbol{t}}P_{j,\boldsymbol{t}}\|_{r_j} d\mathcal{H}_r(\boldsymbol{t})\prod_{i=1}^{k} u^{2r_i/\alpha_i}\,\Psi(u)(1+o(1)), \tag{A.1}$$

where $M_{\gamma,u} = \psi(\Omega \cap \Gamma_{R,\boldsymbol{\alpha}}(\gamma, u))$.

A.2. Manifolds and geometric integration theory

We collect some basic facts about manifolds and geometric integration theory used in this paper for the convenience of the reader. There exist many texts on manifolds (see, e.g., Guillemin and Pollack [19]) and geometric measure theory (see, e.g., Evans and Gariepy [16]).

A map $\psi : \mathcal{A} \to \mathbb{R}^n$ defined on an arbitrary subset $\mathcal{A} \subset \mathbb{R}^r$ is said to be differentiable, if for each point $x \in \mathcal{A}$, around x there exist an open set $U \subset \mathbb{R}^r$ and a differentiable map $F : U \to \mathbb{R}^n$ such that F and ψ are equal on $U \cap \mathcal{A}$. For a subset $\mathcal{B} \subset \mathbb{R}^n$, the map $\psi : \mathcal{A} \to \mathcal{B}$ is called a diffeomorphism if ψ is one to one and onto, and if the inverse map $\psi^{-1} : \mathcal{B} \to \mathcal{A}$

is also differentiable. A subset $\mathcal{M} \subset \mathbb{R}^n$ is an r-dimensional submanifold, if for any point $\mathbf{t} \in \mathcal{M}$, there exists a neighborhood V of x in $\mathcal{M}$, an open set $G \subset \mathbb{R}^r$, and a diffeomorphism $\psi : G \to V$. Denote $\widetilde{\boldsymbol{t}} = \psi^{-1}(\boldsymbol{t}) \in G$ and let $J_\psi(\widetilde{\boldsymbol{t}})$ be the Jacobian matrix of ψ at $\widetilde{\boldsymbol{t}}$, whose dimension is $n \times r$. When $\psi : G \to V$ is a diffeomorphism, $J_\psi(\widetilde{\boldsymbol{t}})$ has full rank for all $\widetilde{\boldsymbol{t}} \in G$.

Consider $J_\psi(\widetilde{\boldsymbol{t}})$ as a linear map from $\mathbb{R}^r$ to $\mathbb{R}^n$. Define the tangent space of $\mathcal{M}$ at $\boldsymbol{t}$, also denoted by $T_{\boldsymbol{t}}\mathcal{M}$, to be the image of the map $J_\psi(\widetilde{\boldsymbol{t}}) : \mathbb{R}^r \to \mathbb{R}^n$, which is an r-dimensional subspace of $\mathbb{R}^n$. The r-dimensional normalized Hausdorff measure, denoted by $\mathcal{H}_r$, is defined as follows. Let diam(E) be the diameter of a subset E of Euclidean spaces.

Definition A.1. For any subset E of a Euclidean space, define

$$\mathcal{H}_r(E) = \omega_r \lim_{\delta \downarrow 0} \inf \Big\{ \sum_{i=1}^{\infty} (\mathrm{diam}(E_i))^r : \ E \subset \bigcup_{i=1}^{\infty} E_i, \ \mathrm{diam}(E_i) < \delta \Big\},$$

where $\omega_r = \frac{\pi^{r/2}}{\Gamma(\frac{r}{2}+1)}$ is the r-dimensional Lebesgue volume of the unit ball in $\mathbb{R}^r$.

It is known that for any positive integer r, the r-dimensional Hausdorff measure coincides with the r-dimensional Lebesgue measure on the class of Lebesgue measurable sets, and the 0-dimensional Hausdorff measure gives the cardinality of a discrete set. Let $\widetilde{\boldsymbol{t}} = (\widetilde{t}_1, \ldots, \widetilde{t}_r)^T$. Consider an r-dimensional infinitesimal rectangle $\mathcal{R}_{\widetilde{\boldsymbol{t}}} := [\widetilde{t}_1, \widetilde{t}_1 + d\widetilde{t}_1] \times \cdots \times [\widetilde{t}_r, \widetilde{t}_r + d\widetilde{t}_r]$. The corresponding image of this rectangle under $J_\psi(\widetilde{\boldsymbol{t}})$ is an r-dimensional parallelotope in the tangent space $T_{\boldsymbol{t}}\mathcal{M}$ with r-dimensional volume $\mathcal{H}_r(J_\psi(\widetilde{\boldsymbol{t}})(\mathcal{R}_{\widetilde{\boldsymbol{t}}})) = B_\psi(\widetilde{\boldsymbol{t}})\mathcal{H}_r(\mathcal{R}_{\widetilde{\boldsymbol{t}}}) = B_\psi(\widetilde{\boldsymbol{t}}) d\widetilde{t}_1 \times \cdots \times d\widetilde{t}_r$, where $B_\psi(\widetilde{\boldsymbol{t}}) = [\det(J_\psi(\widetilde{\boldsymbol{t}})^T J_\psi(\widetilde{\boldsymbol{t}}))]^{1/2}$ is called the Jacobian determinant, and it is the scale factor that reflects the local change of volume in the transformation.

Theorem A.2 (*Area Formula, See Theorem 2, Evans and Gariepy [16], Chapter 3.3*)**.** *Let G be an open subset of $\mathbb{R}^r$ and $\psi : G \to \mathbb{R}^n$ be a continuously differentiable map, $1 \le r \le n$. Then for any Lebesgue measurable subset $\Omega \subset G$ and non-negative Lebesgue measurable function $g : G \to \mathbb{R}$, we have*

$$\int_\Omega g(\widetilde{\boldsymbol{t}})[\det(J_\psi(\widetilde{\boldsymbol{t}})^T J_\psi(\widetilde{\boldsymbol{t}}))]^{1/2} d\mathcal{H}_r(\widetilde{\boldsymbol{t}}) = \int_{\mathbb{R}^n} \sum_{\widetilde{\boldsymbol{t}} \in \Omega \cap \psi^{-1}(\boldsymbol{t})} g(\widetilde{\boldsymbol{t}}) d\mathcal{H}_r(\boldsymbol{t}).$$

When $\psi : G \to V \subset \mathbb{R}^n$ is a diffeomorphism, for any $\boldsymbol{t} \in V$, the set $\psi^{-1}(\boldsymbol{t})$ only consists of a single point. Then the above area formula can be simplified into the following form.

Lemma A.4 (*Change of Variables*)**.** *Let G be an open subset of $\mathbb{R}^r$ and $\psi : G \to V \subset \mathbb{R}^n$ be a C^1-diffeomorphism, $1 \le r \le n$. Then for any Lebesgue measurable subset $\Omega \subset G$ and non-negative Lebesgue measurable function $g : G \to \mathbb{R}$, we have with $U = \psi(\Omega)$ that*

$$\int_\Omega g(\widetilde{\boldsymbol{t}})[\det(J_\psi(\widetilde{\boldsymbol{t}})^T J_\psi(\widetilde{\boldsymbol{t}}))]^{1/2} d\mathcal{H}_r(\widetilde{\boldsymbol{t}}) = \int_U g(\psi^{-1}(\boldsymbol{t})) d\mathcal{H}_r(\boldsymbol{t}). \tag{A.2}$$

Lemma A.4 is a typical change of variable formula. A slightly different form of this formula is given in the following corollary.

Corollary A.1. *Under the same assumptions as in Lemma A.4, if $f : G \to \mathbb{R}$ is a non-negative Lebesgue measurable function, we have*

$$\int_\Omega f(\widetilde{\boldsymbol{t}}) d\mathcal{H}_r(\widetilde{\boldsymbol{t}}) = \int_U \frac{f(\psi^{-1}(\boldsymbol{t}))}{[\det(J_\psi(\psi^{-1}(\boldsymbol{t}))^T J_\psi(\psi^{-1}(\boldsymbol{t})))]^{1/2}} d\mathcal{H}_r(\boldsymbol{t}). \tag{A.3}$$

Proof. This can be immediately obtained by taking $g(\widetilde{t}) = f(\widetilde{t})[\det(J_\psi(\widetilde{t})^T J_\psi(\widetilde{t}))]^{-1/2}$ in Lemma A.4. □

References

[1] R.J. Adler, J.E. Taylor, Random Fields and Geometry, Springer, New York, 2007.
[2] J.M.P. Albin, E. Hashorva, L. Ji, C. Ling, Extremes and limit theorems for difference of chi-type processes, ESAIM Probab. Stat. 20 (2016) 349–366.
[3] J.-M. Azaïs, M. Wschebor, Level Sets and Extrema of Random Processes and Fields, John Wiley & Sons, Hoboken, NJ, 2009.
[4] L. Bai, Extremes of locally-stationary chi-square processes on discrete grids, 2018, arXiv:1807.11687.
[5] S.M. Berman, Limit theorems for the maximum term in stationary sequences, Ann. Math. Statist. 35 (1964) 502–516.
[6] S.M. Berman, Asymptotic independence of the numbers of high and low level crossings of stationary Gaussian processes, Ann. Math. Statist. 42 (1971) 927–945.
[7] P. Bickel, M. Rosenblatt, On some global measures of the deviations of density function estimates, Ann. Statist. 1 (1973) 1071–1095.
[8] P. Bickel, M. Rosenblatt, Two-dimensional random fields, in: P.K. Krishnaiah (Ed.), Multivariate Analysis III, Academic Press, New York, 1973, pp. 3–15.
[9] J.-D. Boissonnat, F. Chazal, M. Yvinec, Geometric and Topological Inference, Cambridge University Press, New York, NY, 2018.
[10] C. Borell, The Brunn–Minkowski inequality in Gauss space, Invent. Math. 30 (1975) 207–216.
[11] J.G. Broida, S.G. Willamson, A Comprehensive Introduction to Linear Algebra, Addison-Wesley, 1989.
[12] D. Cheng, Excursion probabilities of isotropic and locally isotropic Gaussian random fields on manifolds, Extremes 20 (2017) 475–487.
[13] D. Cheng, Y. Xiao, Excursion probability of Gaussian random fields on sphere, Bernoulli 22 (2016) 1113–1130.
[14] V. Chernozhukov, D. Chetverikov, K. Kato, Gaussian approximation of suprema of empirical processes, Ann. Statist. 42 (2014) 1564–1597.
[15] A. Cuevas, R. Fraiman, B. Pateiro-López, On statistical properties of sets fulfilling rolling-type conditions, Adv. Appl. Probab. 44 (2012) 311–329.
[16] L.C. Evans, R.F. Gariepy, Measure Theory and Fine Properties of Functions, CRC Press, Boca Raton, FL, 1992.
[17] H. Federer, Curvature measures, Trans. Amer. Math. Soc. 93 (1959) 418–491.
[18] M.G. Genton, W. Kleiber, Cross-covariance functions for multivariate geostatistics, Statist. Sci. 30 (2015) 147–163.
[19] V. Guillemin, A. Pollack, Differential Topology, Prentice-Hall, Inc., Englewwood Cliffs, New Jersey, 1974.
[20] E. Hashorva, L. Ji, Piterbarg theorems for chi-processes with trend, Extremes 18 (2015) 37–64.
[21] L. Ji, P. Liu, S. Robert, Tail asymptotic behavior of the supremum of a class of chi-square processes, Statist. Probab. Lett. 154 (2019) 108551.
[22] V.D. Konakov, V.I. Piterbarg, On the convergence rate of maximal deviations distributions for kernel regression estimates, J. Multivariate Anal. 15 (1984) 279–294.
[23] D. Konstantinides, V. Piterbarg, S. Stamatovic, Gnedenko-type limit theorems for cyclostationary χ^2-processes, Lith. Math. J. 44 (2) (2004) 157–167.
[24] J.M. Lee, Introduction to Smooth Manifolds, second ed., Springer, New York, 2013.
[25] G. Leobacher, A. Steinicke, Existence, uniqueness and regularity of the projection onto differentiable manifolds, 2020, arXiv:1811.10578.
[26] G. Lindgren, Slepian models for χ^2-processes with dependent components with application to envelope upcrossings, J. Appl. Probab. 26 (1) (1989) 36–49.
[27] C. Ling, Z. Tan, On maxima of chi-processes over threshold dependent grids, Statistics 50 (3) (2016) 579–595.
[28] P. Liu, L. Ji, Extremes of chi-square processes with trend, Probab. Math. Statist. 36 (1) (2016).
[29] P. Liu, L. Ji, Extremes of locally stationary chi-square processes with trend, Stochastic Process. Appl. 127 (2) (2017) 497–525.
[30] T.L. Mikhaleva, V.I. Piterbarg, On the distribution of the maximum of a Gaussian field with constant variance on a smooth manifold, Theory Probab. Appl. 41 (1997) 367–379.
[31] P. Niyogi, S. Smale, S. Weinberger, Finding the homology of submanifolds with high confidence from random samples, Discrete Comput. Geom. 39 (2008) 419–441.

[32] J. Pickands, III, Upcrossing probabilities for stationary Gaussian processes, Trans. Amer. Math. Soc. 145 (1969) 51–73.
[33] V.I. Piterbarg, High excursion for nonstationary generalized chi-square processes, Stochastic Process. Appl. 53 (1994) 307–337.
[34] V.I. Piterbarg, Asymptotic Methods in the Theory of Gaussian Processes and Fields, in: Translations of Mathematical Monographs, vol. 148, American Mathematical Society, Providence, RI, 1996.
[35] V.I. Piterbarg, S. Stamatovich, On maximum of Gaussian non-centered fields indexed on smooth manifolds, in: N. Balakrishnan, I.A. Ibragimov, V.B. Nevzorov (Eds.), Asymptotic Methods in Probability and Statistics with Applications; Statistics for Industry and Technology, Birkhäuser, Boston, MA, 2001, pp. 189–203.
[36] W. Qiao, Asymptotic confidence regions for density ridges, Bernoulli (2021) arXiv:2004.11354, in press.
[37] W. Qiao, W. Polonik, Extrema of rescaled locally stationary Gaussian fields on manifolds, Bernoulli 24 (3) (2018) 1834–1859.
[38] W. Qiao, W. Polonik, Nonparametric confidence regions for level sets: statistical properties and geometry, Electron. J. Stat. 13 (1) (2019) 985–1030.
[39] M. Rosenblatt, On the maximal deviation of k-dimensional density estimates, Ann. Probab. 4 (1976) 1009–1015.
[40] S. Scholtes, On hypersurfaces of positive reach, alternating Steiner formulae and Hadwiger's Problem, 2013, arXiv:1304.4179.
[41] Z. Tan, E. Hashorva, Exact asymptotics and limit theorems for supremum of stationary χ-processes over a random interval, Stochastic Process. Appl. 123 (8) (2013) 2983–2998.
[42] Z. Tan, E. Hashorva, Limit theorems for extremes of strongly dependent cyclo-stationary χ-processes, Extremes 16 (2) (2013) 241–254.
[43] Z. Tan, C. Wu, Limit laws for the maxima of stationary chi-processes under random index, TEST 23 (4) (2014) 769–786.
[44] Y. Zhou, Y. Xiao, Tail asymptotics for the extremes of bivariate Gaussian random fields, Bernoulli 23 (3) (2017) 1566–1598.