

# Extrema of rescaled locally stationary Gaussian fields on manifolds

WANLI QIAO<sup>1</sup> and WOLFGANG POLONIK<sup>2</sup>

<sup>1</sup>Department of Statistics, George Mason University, 4400 University Drive, MS 4A7, Fairfax, VA 22030, USA. E-mail: wpolonik@ucdavis.edu

<sup>2</sup>Department of Statistics, University of California, One Shields Ave., Davis, CA 95616-8705, USA

Given a class of centered Gaussian random fields  $\{X_h(s), s \in \mathbb{R}^n, h \in (0, 1]\}$ , define the rescaled fields  $\{Z_h(t) = X_h(h^{-1}t), t \in \mathcal{M}\}$ , where  $\mathcal{M}$  is a compact Riemannian manifold. Under the assumption that the fields  $Z_h(t)$  satisfy a local stationary condition, we study the limit behavior of the extreme values of these rescaled Gaussian random fields, as  $h$  tends to zero. Our main result can be considered as a generalization of a classical result of Bickel and Rosenblatt (*Ann. Statist.* **1** (1973) 1071–1095), and also of results by Mikhaleva and Piterbarg (*Theory Probab. Appl.* **41** (1997) 367–379).

*Keywords:* extreme values; local stationarity; triangulation of manifolds

## 1. Introduction

In this work, we are studying the following problem. Let  $\{X_h(s), s \in \mathbb{R}^n\}$  be a class of Gaussian random fields indexed by  $h \in (0, 1] \subset \mathbb{R}$ . Define a rescaled random field via

$$Z_h(t) = X_h\left(\frac{t}{h}\right), \quad t \in \mathbf{E} \subset \mathbb{R}^n.$$

In the case of  $\mathbf{E}$  being a manifold, we are studying the extreme value behavior of  $Z_h(t)$  by seeking quantities  $a_h, b_h \in \mathbb{R}$  such that for fixed  $u \in \mathbb{R}$ ,

$$\lim_{h \rightarrow 0} \mathbb{P}\left(a_h \left(\sup_{t \in \mathbf{E}} Z_h(t) - b_h\right) \leq u\right)$$

is non-degenerate.

To our knowledge, existing literature on extreme value behavior and excursion probabilities considers rescaled Gaussian fields only with  $\mathbf{E} \subset \mathbb{R}^n$  being a compact interval or a hyper-cube, but not a manifold – see Pickands [37], Berman [6,7], Leadbetter *et al.* [34], Seleznev [44–46], Hüsler [29,30], Hüsler *et al.* [31], Tan [49] and Tan *et al.* [50]. In cases where the index set is a compact manifold, the random field is not rescaled – see Piterbarg [40], Mikhaleva and Piterbarg [36], Piterbarg and Stamatovich [39], and Cheng [16]. In contrast to that, we consider the case of rescaled fields with  $\mathbf{E}$  being a manifold. We will impose a local stationarity condition on the class of fields. This condition generalizes the notion of local stationarity used by Mikhaleva and Piterbarg [36]. The consideration of local stationarity in the context of excursion probabilities goes back to Berman [5], who considered one-dimensional processes.

Extreme value distributions of rescaled processes and fields also play an important role in the statistics literature. For instance, Bickel and Rosenblatt [10] consider a kernel density estimator  $\widehat{f}_n$  based on a sample of  $n$  independent observations from  $f$ . They derive the asymptotic distribution of the quantity  $\sup_{t \in [0,1]} |\widehat{f}_n(t) - f(t)|$ . The approach underlying their theoretical derivations is to first use strong approximations of  $\{\widehat{f}_n(t) - f(t), t \in [0, 1]\}$  by Gaussian processes, and then to derive a result of asymptotic excursion probability for the approximating Gaussian processes, where  $\mathbf{E} = [0, 1]$ . Here,  $h$  turns out to be the bandwidth used in the definition of the kernel density estimator. The case of a multivariate kernel density estimator was treated later in Rosenblatt [43] by using a similar approach. See Konakov and Piterbarg [33] for similar work in the regression case. One should perhaps note that the set-up in this work is framed slightly differently by considering processes and fields of the form  $\{X_h(s), s \in \mathcal{S}_h\}$ , where  $\mathcal{S}_h = \{s : s = h^{-1}t, t \in [0, 1]^d\}$ . Clearly, this set-up can easily be recast by using a rescaled process or field over a fixed domain. A similar comment applies to the probabilistic literature on rescaled Gaussian fields on intervals or hyper-cubes mentioned above. In contrast to the existing literature, we chose to rescale the fields itself, rather than the index sets. This seems to be more natural in the case of the index set being a manifold, and corresponding technical conditions tend to be easier to interpret.

Rescaled Gaussian processes have also been considered in the non-parametric Bayes literature. For instance, van der Vaart and van Zanten [52] explore the behavior of small ball probabilities for rescaled Gaussian processes of the form  $\{Z(h^{-1}t), t \in [0, 1]\}$  for a given Gaussian process  $Z(t), t \geq 0$ . Here  $h$  either tends to zero or to infinity, corresponding to a roughening or a smoothing effect on the sample paths, respectively. The importance of extrema of Gaussian processes and fields in the field of statistics is further underlined by recent work of Chernozhukov *et al.* [17]. See Sharpnack and Arias-Castro [47] for yet another application of extreme value distributions of Gaussian random fields in statistics. For some more general relevant references to Gaussian fields and extremes, we refer to the books by Adler and Taylor [1], and Azaïs and Wschebor [2].

Our studies are in part motivated by recent interest in statistical inference for manifolds. For instance, considering ridge lines of densities in  $\mathbb{R}^2$  (a manifold), Qiao and Polonik [41] derive a distributional result for the deviation (measured in supremum distance) of an estimator of the ridge line from the truth. Because of the pointwise asymptotic normality of the estimators under consideration and the use of the supremum distance, extreme value results for rescaled Gaussian fields on manifolds come into play there. In fact, the probabilistic result underlying the derivations in Qiao and Polonik [41] follows from the main result proven here.

There exists other statistical work dealing with ridges and related manifold type objects, and even though this work is not directly related to extreme value theory, we provide some references for the convenience of the interested reader. Ridges are considered in Hall *et al.* [27], Genovese *et al.* [23,25], and Chen *et al.* [12–14]. Further manifold-type objects studied in the statistical literature include level curves (Lindgren *et al.* [35], Cuevas *et al.* [19], Chen *et al.* [15], Qiao and Polonik [42]), boundaries of the support of a probability density function (Cuevas *et al.* [20], Biau *et al.* [8]), and integral curves (Koltchinskii *et al.* [32], Qiao and Polonik [41]). In corresponding real world problems, these objects correspond to various types of geometric objects, such as fault lines (geoscience), the cosmic web (astronomy), fiber tracts (neuroscience), or blood vessels (medical imaging).

The outline of the paper is as follows. Section 2 first introduces and discusses the notion of local stationarity used here, and then formalizes and discusses the main result. Its proof is presented in Section 3, and in Section 4 we collect some miscellaneous results and definitions that are used in the paper.

## 2. Main result

First, we define the notion of local stationarity used in the main result below. Let  $\|\cdot\|$  denote Euclidean norm, and for a square integrable field  $Z_h(t)$ , we let  $r_h(t_1, t_2)$  denote the covariance function of  $Z_h(t)$ .

**Definition 2.1 (Local equi- $(\alpha, D_t^h)$ -stationarity).** Let  $\{Z_h(t), t \in \mathcal{S} \subset \mathbb{R}^n, h \in \mathbb{H}\}$  be a class of non-homogeneous random fields, where  $\mathbb{H}$  is an index set. We say that this class is locally equi- $(\alpha, D_t^h)$ -stationary, if the following two conditions hold. For any  $s \in \mathcal{S}$  and  $h \in \mathbb{H}$ , there exists a non-degenerate matrix  $D_s^h$  such that

$$(i) \quad \frac{h^\alpha [1 - r_h(t_1, t_2)]}{\|D_s^h(t_1 - t_2)\|^\alpha} \rightarrow 1 \quad \text{as} \quad \frac{\max\{\|t_1 - s\|, \|t_2 - s\|\}}{h} \rightarrow 0, \tag{2.1}$$

uniformly in  $h \in \mathbb{H}$  and  $s \in \mathcal{S}$ , and

$$(ii) \quad 0 < \inf_{\substack{h \in \mathbb{H}, s \in \mathcal{S} \\ t \in \mathbb{R}^n \setminus \{0\}}} \frac{\|D_s^h t\|}{\|t\|} \leq \sup_{\substack{h \in \mathbb{H}, s \in \mathcal{S} \\ t \in \mathbb{R}^n \setminus \{0\}}} \frac{\|D_s^h t\|}{\|t\|} < \infty. \tag{2.2}$$

*Remarks.* (i) If this definition is applied for one fixed  $h_0 > 0$ , i.e., if  $\mathbb{H} = \{h_0\}$ , then this gives the definition of local  $(\alpha, D_t^{h_0})$ -stationarity as being used in Mikhaleva and Piterbarg [36], for instance. In other words, the above definition requires the class of fields to be locally  $(\alpha, D_t^h)$ -stationary *uniformly* in  $h \in \mathbb{H}$ .

(ii) Observe that Definition 2.1 implies that  $\text{Var}(Z_h(t)) = 1$ .

*Examples of local equi- $(\alpha, D_t^h)$ -stationary fields:* (i) Let  $\nabla^2 f$  be the Hessian matrix of a twice differentiable function  $f$ , and let  $\text{vech}$  denote the half-vectorization operator. Qiao and Polonik [41] consider Gaussian random fields indexed by a 1-dimensional manifold  $\mathcal{M}$  of the form

$$Z_h(t) = b(t)^T \int_{\mathbb{R}^2} \text{vech} \nabla^2 K \left( \frac{t}{h} - s \right) dW(s), \quad t \in \mathcal{M}, h \in (0, 1], \tag{2.3}$$

and show that this class of random fields is locally equi- $(\alpha, D_t^h)$ -stationary with  $n = \alpha = 2$ . Here  $W$  is a 2-dimensional Wiener process,  $b : \mathcal{M} \rightarrow \mathbb{R}^3$  is a smooth vector field, and  $K : \mathbb{R}^2 \mapsto \mathbb{R}^+$  is a smooth kernel function with the unit ball in  $\mathbb{R}^2$  as its support. Note that the appearance of the ratio  $h^{-1}t$  in the above random fields provides a motivation to write  $Z_h(t) = X_h(h^{-1}t)$ , where  $X_h(u)$  is a Gaussian random field defined on the rescaled manifold  $h^{-1}\mathcal{M}$ . This is the point of view taken in Qiao and Polonik [41].

(ii) Let  $\mathbb{S}^{n-1} = \{x \in \mathbb{R}^n : \|x\| = 1\}$  be the  $(n - 1)$ -dimensional unit sphere. For any  $t_1, t_2 \in \mathbb{S}^{n-1}$  let  $d(t_1, t_2) = \arccos\langle t_1, t_2 \rangle$  be the spherical distance between  $t_1$  and  $t_2$ . Let  $\{Z_h(t), t \in \mathbb{S}^{n-1}, h \in (0, 1]\}$  be a class of Gaussian fields with covariance functions

$$r_h(t_1, t_2) = \exp\left[-\left|g_h\left(\frac{t_1 + t_2}{\|t_1 + t_2\|}\right)d(t_1, t_2)\right|^\alpha\right], \tag{2.4}$$

where  $0 < \alpha \leq 1$ ,  $(t_1 + t_2)\|t_1 + t_2\|^{-1}$  is the midpoint between  $t_1$  and  $t_2$  on  $\mathbb{S}^{n-1}$ , and  $g_h(t), t \in \mathbb{S}^{n-1}$ , is a class of equi-continuous functions such that  $0 < \inf_{0 < h \leq 1, t \in \mathbb{S}^{n-1}} h g_h(t) \leq \sup_{0 < h \leq 1, t \in \mathbb{S}^{n-1}} h g_h(t) < \infty$ . In this setting, the field  $Z_h(t)$  is locally equi- $(\alpha, D_t^h)$ -stationary with  $D_t^h = h g_h(t) \mathbf{I}_n$ , where  $\mathbf{I}_n$  denotes the  $n \times n$  identity matrix. The verification of this fact is provided in Section 4. This example can be viewed as a non-stationary generalization of powered exponential covariance functions (also called stable covariance functions in Chilès and Delfiner [18]), obtained by introducing an additional index parameter. In principle, the generalization by including  $g_h$  as above, can be applied to similar isotropic stationary random fields on  $\mathbb{S}^{n-1}$  with covariance function only depending on  $d(t_1, t_2)$ . Examples of such covariance functions can be found in Huang *et al.* [28] and Gneiting [26], for instance.

*Further notation.* For an  $n \times r$  matrix  $G$  with  $r \leq n$ , we denote by  $\|G\|_r^2$  the sum of squares of all minor determinants of order  $r$ . Let further  $H_\alpha^{(r)}$  denote generalized Pickands constant of a Gaussian field (see Section 4 for a definition). At each  $u \in \mathcal{M}$ , let  $T_u \mathcal{M}$  denote the tangent space at  $u$  to  $\mathcal{M}$ . For an  $r$ -dimensional smooth manifold  $\mathcal{M}$  embedded in  $\mathbb{R}^n$ , let  $\Delta(\mathcal{M})$  denote the reach of  $\mathcal{M}$  (Federer [22]).  $\Delta(\mathcal{M})$  is the largest  $\lambda \geq 0$  such that each point in  $\mathcal{M} \oplus \lambda$  has a unique projection onto  $\mathcal{M}$ , where  $\mathcal{M} \oplus \lambda$  denotes the  $\lambda$ -enlarged set of  $\mathcal{M}$ , that is, the union of all open balls of radius  $\lambda$  with midpoints in  $\mathcal{M}$ . A compact smooth ( $C^2$ ) manifold embedded in an Euclidean space has a positive reach, for example, see Thäle [51]. A positive  $\Delta(\mathcal{M})$  indicates a ‘bounded curvature’ of  $\mathcal{M}$ . As indicated in Lemma 3 of Genovese *et al.* [24], on a manifold with a positive reach, a small Euclidean distance implies a small geodesic distance.

Now we state our main theorem, which generalizes Theorem 2 in Piterburg and Stamatovich [39] and Theorem A1 in Bickel and Rosenblatt [10].

**Theorem 2.1.** *Let  $\mathcal{H} \subset \mathbb{R}^n$  be a compact set. Let  $\{Z_h(t), t \in \mathcal{H}, 0 < h \leq 1\}$  be a class of Gaussian centered locally equi- $(\alpha, D_t^h)$ -stationary fields with  $0 < \alpha \leq 2$ , and all components of  $D_t^h$  continuous in both  $h$  and  $t$ . Let  $\mathcal{M} \subset \mathcal{H}$  be an  $r$ -dimensional compact Riemannian manifold with  $\Delta(\mathcal{M}) > 0$ . Suppose that  $D_t^h, t \in \mathcal{H}$ , uniformly converges, as  $h \downarrow 0$ , to a matrix field  $D_t, t \in \mathcal{H}$ , with continuous components. For  $u > 0$ , let*

$$Q(u) = \sup_{0 < h \leq 1} \left\{ |r_h(t + s, s)| : t + s \in \mathcal{M}, s \in \mathcal{M}, \|t\| > hu \right\}, \tag{2.5}$$

where  $r_h$  denotes the covariance function of  $Z_h(t)$ . Suppose that, for any  $u > 0$ , there exists  $\eta > 0$  such that

$$Q(u) < \eta < 1. \tag{2.6}$$

Furthermore, let  $u_0 > 0$  be such that for a function  $v(\cdot)$  and for  $u > u_0$ , we have

$$Q(u)|(\log u)^{2r/\alpha}| \leq v(u), \tag{2.7}$$

where  $v$  is a monotonically decreasing, such that, for some  $p > 0$ ,  $v(u^p) = O(v(u)) = o(1)$  and  $v(u)u^p \rightarrow \infty$  as  $u \rightarrow \infty$ . Let

$$\begin{aligned} \beta_h &= \left(2r \log \frac{1}{h}\right)^{\frac{1}{2}} \\ &+ \left(2r \log \frac{1}{h}\right)^{-\frac{1}{2}} \left[ \left(\frac{r}{\alpha} - \frac{1}{2}\right) \log \log \frac{1}{h} + \log \left\{ \frac{(2r)^{\frac{r}{\alpha} - \frac{1}{2}}}{\sqrt{2\pi}} H_\alpha^{(r)} I(\mathcal{M}) \right\} \right], \end{aligned} \tag{2.8}$$

where  $I(\mathcal{M}) = \int_{\mathcal{M}} \|D_s M_s\|_r ds$  with  $M_s$  an  $n \times r$  matrix with orthonormal columns spanning  $T_s \mathcal{M}$ . Then

$$\lim_{h \rightarrow 0} \mathbb{P} \left\{ \sqrt{2r \log \frac{1}{h}} \left( \sup_{t \in \mathcal{M}} |Z_h(t)| - \beta_h \right) \leq z \right\} = \exp\{-2 \exp\{-z\}\}. \tag{2.9}$$

Remarks.

1. Since our result is asymptotic for  $h \rightarrow 0$ , the quantity  $Q(u)$  in the theorem can be replaced with

$$Q_{h_0}(u) = \sup_{0 < h \leq h_0} \{ |r_h(t + s, s)| : t + s \in \mathcal{M}, s \in \mathcal{M}, \|t\| > hu \}, \tag{2.10}$$

for some (small)  $h_0 > 0$ .

2. The function  $v(u) = (\log u)^{-\beta}$ , for  $\beta > 0$ , provides an example of a function satisfying the properties required in the theorem.
3. Qiao and Polonik [41] apply a special case of the above theorem to the random fields  $Z_h(t)$  given in (2.3). There,  $r = 1$ ,  $n = 2$ ,  $\alpha = 2$ , and  $Q(u) = 0$  for  $u > u_0$ . This property of  $Q(u)$  in particular implies that the function  $v(u) = (\log u)^{-\beta}$ ,  $\beta > 0$ , works in this case. The fact that  $Q(u)$  has this particular property follows from the assumption that the support of  $K$  (and of its second order partial derivatives) is bounded. This implies that the covariances of  $Z_h(x_1)$  and  $Z_h(x_2)$  become zero once the distance  $\|x_1 - x_2\|$  exceeds a certain threshold.
4. Suppose that in the above example with covariance function (2.4), we further assume that  $\lim_{h \rightarrow 0} h g_h(t) = g(t)$ , with  $g(t)$  a continuous and bounded function on  $\mathbb{S}^{n-1}$ . Then conditions (2.2), (2.6) and (2.7) all hold (see Section 4), and thus we have (2.9) with

$$\begin{aligned} \beta_h &= \left(2(n-1) \log \frac{1}{h}\right)^{\frac{1}{2}} + \left(2(n-1) \log \frac{1}{h}\right)^{-\frac{1}{2}} \\ &\times \left[ \left(\frac{n-1}{\alpha} - \frac{1}{2}\right) \log \log \frac{1}{h} + \log \left\{ \frac{n(2n-2)^{(n-1)/\alpha - 1/2}}{\sqrt{2\pi}} H_\alpha^{(r)} \int_{\mathbb{S}^{n-1}} |g(s)| ds \right\} \right]. \end{aligned}$$

### 3. Proof of Theorem 2.1

We need some more notation. Given a set  $U$  and a metric  $d_U$  on  $U$ , a set  $S \subset U$  is an  $\varepsilon$ -net, if, (i) for any  $u \in U$ , we have  $\inf_{s \in S} d_U(s, u) \leq \varepsilon$ , and (ii)  $d_U(s, t) \geq \varepsilon$  for any  $s, t \in S$ . Let  $\phi$  and  $\Phi$  denote the standard normal density and cumulative distribution function, respectively, and let  $\bar{\Phi}(u) = 1 - \Phi(u)$  and  $\Psi(u) = u^{-1}\phi(u)$ . We also let  $V_r$  denote  $r$ -dimensional Hausdorff measure. With  $\beta_h$  in (2.8) let

$$\theta = \beta_h + \frac{1}{\sqrt{2r \log(1/h)}}z. \tag{3.1}$$

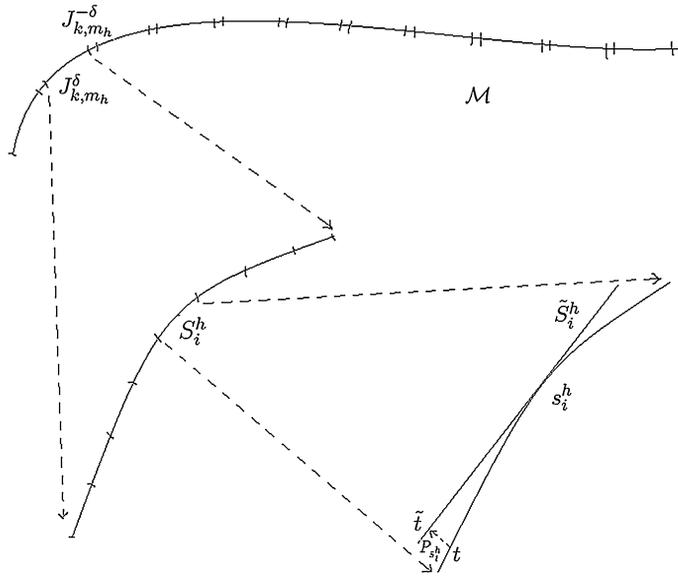
Note that  $\theta$  depends on  $h$  and  $z$ , but for simplicity we omit this dependence in our notation. With this notation, we can rewrite (2.9) as

$$\lim_{h \rightarrow 0} \mathbb{P} \left\{ \sup_{t \in \mathcal{M}} |Z_h(t)| \leq \theta \right\} = \exp\{-2 \exp\{-z\}\}.$$

The proof of the theorem is constructing various approximations to  $\sup_{t \in \mathcal{M}} |Z_h(t)|$  that will facilitate the control of the probabilities on the left-hand side. Essentially, the field  $Z_h(t)$  that lives on the manifold  $\mathcal{M}$  is linearized by first approximating the manifold locally via tangent planes, and then defining an approximating field on these tangent planes. This approach is typical for deriving extreme value results for such fields (e.g., see Hüsler *et al.* [31]). We begin with some preparations, thereby outlining the main ideas of the proof. The geometric picture underlying the idea of the proof is indicated in Figure 3.1.

(i) *Partitioning  $\mathcal{M}$* : We partition the manifold  $\mathcal{M}$  as follows. Suppose that  $V_r(\mathcal{M}) = \ell$ . For a fixed  $\ell^* < \ell$  and any  $0 < h \leq 1$ , there exists an  $(h\ell^*)$ -net on  $\mathcal{M}$  with respect to geodesic distance with cardinality of  $O((h\ell^*)^{-r})$ . A Delaunay triangulation using the  $(h\ell^*)$ -net results in a partition of  $\mathcal{M}$  into  $m_h = O((h\ell^*)^{-r})$  disjoint pieces  $\{J_{k,m_h} : k = 1, 2, \dots, m_h\}$ . The construction is such that  $\max_{k=1, \dots, m_h} V_r(J_{k,m_h})$ , the norm of this partition, is  $h^r O(\ell^{*r})$ , where the  $O$ -term is uniform in  $h$ . It is known that for compact Riemannian manifolds such  $(h\ell^*)$ -nets and Delaunay triangulations exist for any  $r \in \mathbb{Z}^+$  with  $r < n$ , and for  $\ell^*$  small enough (see, e.g., de Laet [21]). (In the case of  $r = 1$ , the construction just described simply amounts to choosing all the  $O((h\ell^*)^{-1})$  many sets  $J_{k,m_h}$  as pieces on the curve  $\mathcal{M}$ , which has length at most  $h\ell^*$ .) One should point out that while  $\ell^*$  has to be chosen sufficiently small, it is a constant not depending on  $h$ . In particular, this means that it does not tend to zero in this work.

(ii) *‘Small blocks–large blocks’ approach*: For sufficiently small  $\delta > 0$ , let  $\mathcal{B}^{-h\delta} \subset \mathcal{M}$  be the  $(h\delta)$ -enlarged neighborhood (using geodesic distance) of the union of the boundaries of all  $J_{k,m_h}$ . The minus sign in the superscript indicates that this set will be ‘cut out’ in the below construction. We obtain  $J_{k,m_h}^\delta = J_{k,m_h} \setminus \mathcal{B}^{-h\delta}$  (‘large blocks’) and  $J_{k,m_h}^{-\delta} = J_{k,m_h} \setminus J_{k,m_h}^\delta$  (‘small blocks’) for  $1 \leq k \leq m_h$ . Geometrically, we envision  $J_{k,m_h}^{-\delta}$  as a small tube along the boundaries of  $J_{k,m_h}$  (lying inside  $J_{k,m_h}$ ), and  $J_{k,m_h}^\delta$  is the set that remains when  $J_{k,m_h}^{-\delta}$  is cut out of  $J_{k,m_h}$ . We have  $\max_{k=1, \dots, m_h} V_r(J_{k,m_h}^{-\delta}) = h^r O(\delta)$ , where the  $O$ -term is uniform in  $h$ . The construction of the partition is such that the boundaries of the projections of all the sets  $J_{k,m_h}$ ,  $J_{k,m_h}^\delta$  and  $J_{k,m_h}^{-\delta}$  onto the local tangent planes are null sets, and thus Jordan measurable. This will be used below.



**Figure 3.1.** This figure visualizes some of the definitions introduced here in the case  $r = 1$  and  $n = 2$ .

Let  $B_h(A) = \{\sup_{t \in A} |Z_h(t)| \geq \theta\}$ , and denote  $p_h(A) = \mathbb{P}(B_h(A))$ . Also denote  $\mathcal{J}_{m_h}^\delta = \bigcup_{k \leq m_h} J_{k,m_h}^\delta$ . Approximating  $\mathcal{M}$  by  $\mathcal{J}_{m_h}^\delta$  leads to the approximation of  $p_h(\mathcal{M})$  by  $p_h(\mathcal{J}_{m_h}^\delta)$ . The volume of  $\bigcup_{k \leq m_h} J_{k,m_h}^{-\delta}$ , that is, the difference between the volumes of  $\mathcal{M}$  and  $\mathcal{J}_{m_h}^\delta$ , is of the order  $O(\delta)$  uniformly in  $h$ . The order of the difference  $p_h(\mathcal{M}) - p_h(\mathcal{J}_{m_h}^\delta)$  turns out to be of the same. Thus we have to choose  $\delta$  small enough.

(iii) *Refinement of the partition:* Let  $J$  denote one of the sets  $J_{k,m_h}$ ,  $J_{k,m_h}^\delta$  and  $J_{k,m_h}^{-\delta}$ . We can construct a cover  $\{S_i^h(J) \subset J, i = 1, \dots, N_h(J)\}$  of this set by using the same Delaunay triangulation technique as above, but of course based on a smaller mesh. As above, by controlling the mesh size, we can control the norm of the partition uniformly over  $h$ , because of the positive reach of  $\mathcal{M}$ .

The probabilities  $p_h(J_{k,m_h}^\delta)$  are approximated by  $\sum_{i=1}^{N_h} p_h(S_i^h)$ , with  $S_i^h (= S_i^h(J_{k,m_h}^\delta))$  the cover of  $J_{k,m_h}^\delta$  introduced above. It will turn out that, using Bonferroni's inequality, the approximation error can be bounded by a double sum (see (3.12)). To show this bound to be negligible as compared to the approximating sum  $\sum_{i=1}^{N_h} p_h(S_i^h)$ , we have to make sure that the volume of  $h^{-1}S_i^h$  is sufficiently small. It will turn out that each  $p_h(J_{k,m_h}^\delta)$  essentially behaves like the tail probability of a normal distribution.

(iv) *Projection onto tangent space of refined partition:* We approximate the small pieces  $S_i^h$  by their projections  $\tilde{S}_i^h$  onto a tangent space, and correspondingly approximate the probabilities  $p_h(S_i^h)$  by the probabilities of a transformed field over  $\tilde{S}_i^h$ . More precisely, we choose some point  $s_i^h$  on  $S_i^h$ , and project  $S_i^h$  orthogonally onto the tangent space of  $\mathcal{M}$  at the point  $s_i^h$ . We denote this projection map by  $P_{s_i^h}(\cdot)$  and we let  $\tilde{S}_i^h = P_{s_i^h}(S_i^h)$ , which, as indicated above, is

Jordan measurable by the construction. The error generated by this approximation is controlled by choosing the norm of the partition  $\{S_i^h\}$  to be sufficiently small.

In the following, we continue to often drop  $J_{k,m_h}^\delta$  in the notation and simply write  $S_i^h$  instead of  $S_i^h(J_{k,m_h}^\delta)$ , unless otherwise indicated. For further simplicity and generic discussion, we sometimes also omit the index  $i$  of  $s_i^h$ ,  $S_i^h$  and  $\tilde{S}_i^h$ .

(v) *Discretization:* The probabilities  $p_h(\tilde{S}_i^h)$  are approximated by replacing the supremum in  $p_h(\tilde{S}_i^h)$  by a maximum over a finite set, i.e., a set of ‘dense grid points’ on  $\tilde{S}_i^h$ . The accuracy of the approximation is controlled by choosing both  $h$  and the grid size (represented by  $\gamma$  as introduced below) to be sufficiently small. The construction of the dense grid is as follows:

Let  $\{M_{s^h}^j : j = 1, \dots, r\}$  be linearly independent, orthonormal vectors spanning the tangent space of  $\mathcal{M}$  at the point  $s^h$ , and let  $M_{s^h}$  denote the  $n \times r$  matrix with  $M_{s^h}^j$  as columns. For a given  $\gamma > 0$ , consider the (discrete) set  $\tilde{\Gamma}_{h\gamma\theta^{-2/\alpha}}(\tilde{S}^h) = \{u : u = s^h + \sum_{j=1}^r i_j h\gamma\theta^{-2/\alpha} M_{s^h}^j \in \tilde{S}^h, i_j \in \mathbb{Z}\}$  and let  $\Gamma_{h\gamma\theta^{-2/\alpha}}(S^h) = (P_{s^h})^{-1}(\tilde{\Gamma}_{h\gamma\theta^{-2/\alpha}}(\tilde{S}^h))$ , which is a subset of  $S^h$ . Note that the geodesic distance between any two adjacent points in  $\Gamma_{h\gamma\theta^{-2/\alpha}}(S^h)$  is still of the order  $O(h\gamma\theta^{-2/\alpha})$ , again due to the assumed positive reach of the manifold  $\mathcal{M}$ .

The union  $\mathbb{T}_h^\delta$  of all the sets  $\Gamma_{h\gamma\theta^{-2/\alpha}}(S^h)$  gives a set of dense points in  $\mathcal{J}_{m_h}^\delta$ . It will turn out that the probability  $1 - p_h(\mathbb{T}_h^\delta) = \mathbb{P}(\bigcap_k (B_h(\mathbb{T}_h^\delta \cap J_{k,m_h}^\delta))^c)$  can be approximated by assuming the events  $(B_h(\mathbb{T}_h^\delta \cap J_{k,m_h}^\delta))^c, k = 1, \dots, m_h$ , to be independent. To make sure this approximation is valid,  $\delta$  may be not too small, and  $\gamma$  may be not too small compared to  $h$ .

Putting everything together will then complete the proof.

**Details of the proof.** We now present the details by using the notation introduced above. Furthermore, for any set  $\mathcal{A}$  being one of the  $J$ ’s or their unions, let  $I_h(\mathcal{A}) = \int_{\mathcal{A}} \|D_s^h M_s\|_r ds$ . We split the proof into different parts in order to provide more structure. Note that the parts do not follow the same order of the steps outlined above.

*Part 1.* Recall the definition of the refined partition  $\{S_i^h, i = 1, \dots, N_h\}$  of  $J_{k,m_h}$  given in (iii). Here (and also in Part 2 of this proof) we require  $\theta$  to be large enough, but  $\theta$  is considered to be independent of  $h$ . We show that the corresponding derivations hold uniformly in  $h$  (as long as  $\theta$  is chosen to be large enough). This then will allow us to let  $\theta$  depend on  $h$  in the later parts of this proof. Specifically, we show that  $\sum_i p_h(S_i^h) \approx h^{-r} \theta^{2r/\alpha} \Psi(\theta) H_\alpha^{(r)} I_h(J_{k,m_h}^\delta)$  uniformly in  $h$  for  $\theta$  large enough, and that a similar approximation holds for  $J_{k,m_h}^\delta$  replaced by  $J_{k,m_h}$ . To this end, we will utilize the projections  $\tilde{S}_i^h$  of  $S_i^h$  onto the tangent space (see (iv)) as well as the approximation of  $\tilde{S}_i^h$  by a set of dense points introduced in (v).

The various asymptotic approximations in this step are similar to those in the proof of Theorem 1 in Mikhaleva and Piterbarg [36], but here we consider them in the uniform sense. Recall that the assumption of a positive reach of  $\mathcal{M}$  means that  $\mathcal{M}$  has a bounded curvature. Thus for any  $\varepsilon_1 > 0$ , there exists a constant  $\delta_1 > 0$  (not depending on  $h$ ) such that if the volumes of all  $S^h = S_i^h = S_i^h(J_{k,m_h}^\delta)$  are less than  $\delta_1$ , then we have

$$1 - \varepsilon_1 \leq \frac{V_r(\tilde{S}^h)}{V_r(S^h)} \leq 1 + \varepsilon_1, \tag{3.2}$$

where  $V_r(\cdot)$  is the  $r$ -dimensional Hausdorff measure. On  $\tilde{S}^h$  we consider the Gaussian field defined as  $\tilde{Z}_h(\tilde{t}) = Z_h(t)$ , with  $t \in S^h$  such that  $\tilde{t} = P_{S^h}(t) \in \tilde{S}^h$ . We will further omit the superscript  $h$  of  $S^h$  below. Due to local equi- $(\alpha, D_t^h)$ -stationarity of  $Z_h(t)$ , for any  $\varepsilon_2 > 0$ , the covariance function  $\tilde{r}_h(\tilde{t}_1, \tilde{t}_2)$  of the field  $\tilde{Z}_h(\tilde{t})$  satisfies

$$1 - (1 + \varepsilon_2/4) \|h^{-1} D_s^h(t_1 - t_2)\|^\alpha \leq \tilde{r}_h(\tilde{t}_1, \tilde{t}_2) \leq 1 - (1 - \varepsilon_2/4) \|h^{-1} D_s^h(t_1 - t_2)\|^\alpha$$

for all  $t_1, t_2 \in S^h$ , provided that  $V_r(S^h)h^{-r}$  is less than a certain threshold  $\delta_2$  depending only on  $\varepsilon_2$ . By possibly decreasing  $\delta_2$  further, we also have

$$1 - (1 + \varepsilon_2/2) \|h^{-1} D_s^h(\tilde{t}_1 - \tilde{t}_2)\|^\alpha \leq \tilde{r}_h(\tilde{t}_1, \tilde{t}_2) \leq 1 - (1 - \varepsilon_2/2) \|h^{-1} D_s^h(\tilde{t}_1 - \tilde{t}_2)\|^\alpha$$

for all  $\tilde{t}_1, \tilde{t}_2 \in \tilde{S}^h$ . Note that, due to the curvature of  $\mathcal{M}$  being bounded, these inequalities hold uniformly over all  $\tilde{S}^h$  under consideration.

On  $\tilde{S}^h$ , we introduce two homogeneous Gaussian fields  $Z_h^+(\tilde{t})$  and  $Z_h^-(\tilde{t})$  such that their covariance functions satisfy

$$\begin{aligned} r_h^+(\tilde{t}_1, \tilde{t}_2) &= 1 - (1 + \varepsilon_2) \|h^{-1} D_s^h(\tilde{t}_1 - \tilde{t}_2)\|^\alpha + o(\|h^{-1} D_s^h(\tilde{t}_1 - \tilde{t}_2)\|^\alpha) \quad \text{and} \\ r_h^-(\tilde{t}_1, \tilde{t}_2) &= 1 - (1 - \varepsilon_2) \|h^{-1} D_s^h(\tilde{t}_1 - \tilde{t}_2)\|^\alpha + o(\|h^{-1} D_s^h(\tilde{t}_1 - \tilde{t}_2)\|^\alpha), \end{aligned}$$

as  $\|h^{-1}(\tilde{t}_1 - \tilde{t}_2)\| = o(1)$ . Thus, if the volumes of all sets  $S^h$  under consideration are sufficiently small, then

$$r_h^+(\tilde{t}_1, \tilde{t}_2) \leq \tilde{r}_h(\tilde{t}_1, \tilde{t}_2) \leq r_h^-(\tilde{t}_1, \tilde{t}_2)$$

holds for all  $\tilde{t}_1, \tilde{t}_2 \in \tilde{S}^h$ . This can be achieved by possibly adjusting  $\delta_2$  from above. Slepian's comparison lemma (Slepian [48]) implies that

$$\begin{aligned} \mathbb{P}\left(\sup_{\tilde{t} \in \tilde{S}^h} Z_h^-(\tilde{t}) > \theta\right) &\leq \mathbb{P}\left(\sup_{\tilde{t} \in \tilde{S}^h} \tilde{Z}_h(\tilde{t}) > \theta\right) \\ &= \mathbb{P}\left(\sup_{t \in S^h} Z_h(t) > \theta\right) \\ &\leq \mathbb{P}\left(\sup_{\tilde{t} \in \tilde{S}^h} Z_h^+(\tilde{t}) > \theta\right), \end{aligned}$$

and that

$$\begin{aligned} \mathbb{P}\left(\max_{\tilde{t} \in \tilde{\Gamma}_{h\gamma\theta^{-2/\alpha}}(\tilde{S}^h)} Z_h^-(\tilde{t}) > \theta\right) &\leq \mathbb{P}\left(\max_{\tilde{t} \in \tilde{\Gamma}_{h\gamma\theta^{-2/\alpha}}(\tilde{S}^h)} \tilde{Z}_h(\tilde{t}) > \theta\right) \\ &= \mathbb{P}\left(\max_{t \in \Gamma_{h\gamma\theta^{-2/\alpha}}(S^h)} Z_h(t) > \theta\right) \\ &\leq \mathbb{P}\left(\max_{\tilde{t} \in \tilde{\Gamma}_{h\gamma\theta^{-2/\alpha}}(\tilde{S}^h)} Z_h^+(\tilde{t}) > \theta\right). \end{aligned} \tag{3.3}$$

For  $\tau \in \mathbb{R}^n$  such that  $h(1 + \varepsilon_2)^{-1/\alpha} (D_s^h)^{-1} \tau \in \tilde{S}^h$ , denote  $Y_h^+(\tau) = Z_h^+(h(1 + \varepsilon_2)^{-1/\alpha} (D_s^h)^{-1} \tau)$ . The covariance function  $r_{Y_h^+}(\tau_1, \tau_2)$  of  $Y_h^+(\tau)$  satisfies

$$\begin{aligned} r_{Y_h^+}(\tau_1, \tau_2) &= r_h^+(h(1 + \varepsilon_2)^{-1/\alpha} (D_s^h)^{-1} \tau_1, h(1 + \varepsilon_2)^{-1/\alpha} (D_s^h)^{-1} \tau_2) \\ &= 1 - (1 + \varepsilon_2) \|h^{-1} D_s^h (h(1 + \varepsilon_2)^{-1/\alpha} (D_s^h)^{-1} \tau_1 - h(1 + \varepsilon_2)^{-1/\alpha} (D_s^h)^{-1} \tau_2)\|^\alpha \\ &\quad + o(\|\tau_1 - \tau_2\|^\alpha) \\ &= 1 - \|\tau_1 - \tau_2\|^\alpha + o(\|\tau_1 - \tau_2\|^\alpha) \quad \text{as } \|\tau_1 - \tau_2\| \rightarrow 0. \end{aligned}$$

An application of Lemma 4.2 gives that, for any  $\varepsilon_3 > 0$  and  $\theta$  large enough, we have uniformly in  $h \in (0, 1]$ , with  $\tilde{H}_{\alpha,r}(\gamma) = \gamma^{-r} H_\alpha^{(r)}(\gamma)$  that

$$\begin{aligned} &\frac{\mathbb{P}(\max_{\tilde{t} \in \tilde{\Gamma}_{h\gamma\theta^{-2/\alpha}}(\tilde{S}^h)} Z_h^+(\tilde{t}) > \theta)}{\theta^{2r/\alpha} \Psi(\theta)} \\ &= \frac{\mathbb{P}(\max_{\tau \in h^{-1}(1+\varepsilon_2)^{1/\alpha} D_s^h \tilde{\Gamma}_{h\gamma\theta^{-2/\alpha}}(\tilde{S}^h)} Z_h^+(h(1 + \varepsilon_2)^{-1/\alpha} (D_s^h)^{-1} \tau) > \theta)}{\theta^{2r/\alpha} \Psi(\theta)} \\ &= \frac{\mathbb{P}(\max_{\tau \in (1+\varepsilon_2)^{1/\alpha} D_s^h \tilde{\Gamma}_{\gamma\theta^{-2/\alpha}}(h^{-1}\tilde{S}^h)} Y_h^+(\tau) > \theta)}{\theta^{2r/\alpha} \Psi(\theta)} \tag{3.4} \\ &\leq \tilde{H}_{\alpha,r}(\gamma) \frac{(1 + \varepsilon_3)(1 + \varepsilon_2)^{r/\alpha}}{h^r} V_r(D_s^h \tilde{S}^h) \\ &= (1 + \varepsilon_2)^{r/\alpha} (1 + \varepsilon_3) \tilde{H}_{\alpha,r}(\gamma) \|D_s^h M_s\|_r V_r(\tilde{S}^h). \end{aligned}$$

Similarly, by defining  $Y_h^-(\tau) = Z_h^-((1 - \varepsilon_2)^{-1/\alpha} (D_s^h)^{-1} \tau)$ , we get

$$\frac{\mathbb{P}(\max_{\tilde{t} \in \tilde{\Gamma}_{h\gamma\theta^{-2/\alpha}}(\tilde{S}^h)} Z_h^-(\tilde{t}) > \theta)}{\theta^{2r/\alpha} \Psi(\theta)} \geq (1 - \varepsilon_2)^{r/\alpha} (1 - \varepsilon_3) \tilde{H}_{\alpha,r}(\gamma) \|D_s^h M_s\|_r V_r(\tilde{S}^h). \tag{3.5}$$

Combining (3.2), (3.3), (3.4) and (3.5), we obtain, for  $V_r(h^{-1} S^h)$  small enough and  $\theta$  large enough, that for any  $\varepsilon > 0$ ,

$$\begin{aligned} &\left(1 - \frac{\varepsilon}{4}\right) h^{-r} \tilde{H}_{\alpha,r}(\gamma) \|D_s^h M_s\|_r V_r(S^h) \\ &\leq \frac{\mathbb{P}(\max_{t \in \Gamma_{h\gamma\theta^{-2/\alpha}}(S^h)} Z_h(t) > \theta)}{\theta^{2r/\alpha} \Psi(\theta)} \\ &\leq \left(1 - \frac{\varepsilon}{4}\right) h^{-r} \tilde{H}_{\alpha,r}(\gamma) \|D_s^h M_s\|_r V_r(S^h). \end{aligned}$$

Since Lemma 4.1 says that  $\tilde{H}_{\alpha,r}(\gamma)(H_\alpha^{(r)})^{-1} \rightarrow 1$  as  $\gamma \rightarrow 0$ , we further have for  $\gamma$  sufficiently small that

$$\begin{aligned} & h^{-r} \left(1 - \frac{\varepsilon}{2}\right) H_\alpha^{(r)} \|D_s^h M_s\|_r V_r(S^h) \\ & \leq \frac{\mathbb{P}(\max_{t \in \Gamma_{h\gamma\theta^{-2/\alpha}}(S^h)} Z_h(t) > \theta)}{\theta^{2r/\alpha} \Psi(\theta)} \\ & \leq h^{-r} \left(1 - \frac{\varepsilon}{2}\right) H_\alpha^{(r)} \|D_s^h M_s\|_r V_r(S^h). \end{aligned} \tag{3.6}$$

This, in fact, holds for any  $S^h = S_i^h$ . We now want to sum over  $i$ . Recall that the collection  $\{S_i^h, i = 1, \dots, N_h\}$  defines a decomposition of  $J_{k,m_h}^\delta$  into small sets, and observe that thus  $\sum_{i=1}^{N_h} (\|D_{S_i^h}^h M_{S_i^h}\|_r V_r(S_i^h))$  can be interpreted as a Riemann sum. Namely, for any  $\varepsilon > 0$ , there exists  $\delta_3 > 0$  such that, for  $\max_{i=1, \dots, N_h} V_r(S_i^h) < \delta_3$ , we have

$$(1 - \varepsilon) I_h(J_{k,m_h}^\delta) \leq \sum_{i=1}^{N_h} \|D_{S_i^h}^h M_{S_i^h}\|_r V_r(S_i^h) \leq (1 + \varepsilon) I_h(J_{k,m_h}^\delta). \tag{3.7}$$

Here recall that  $I_h(J_{k,m_h}^\delta) = \int_{J_{k,m_h}^\delta} \|D_s^h M_s\|_r ds$ . The selection of  $\delta_3$  only depends on  $\varepsilon$ , and the uniformity in  $h$  comes from the fact that for any  $t_1, t_2 \in \mathcal{H}$ ,  $\|D_{t_1}^h - D_{t_2}^h\|_n = \|D_{t_1} - D_{t_2}\|_n + o(1)$  as  $h \rightarrow 0$ , and that  $D_t$  is continuous in  $t \in \mathcal{H}$ .

It follows from (3.6) and (3.7) that for both  $\gamma > 0$  and  $\sup_{0 < h \leq 1} \max_{i=1, \dots, N_h} V_r(h^{-1} S_i^h)$  sufficiently small, and  $\theta$  sufficiently large, we have that for any  $\varepsilon > 0$ ,

$$\begin{aligned} (1 - \varepsilon) h^{-r} H_\alpha^{(r)} I_h(J_{k,m_h}^\delta) & \leq \frac{\sum_{i=1}^{N_h} \mathbb{P}(\max_{t \in \Gamma_{h\gamma\theta^{-2/\alpha}}(S_i^h)} Z_h(t) > \theta)}{\theta^{2r/\alpha} \Psi(\theta)} \\ & \leq (1 + \varepsilon) h^{-r} H_\alpha^{(r)} I_h(J_{k,m_h}^\delta). \end{aligned} \tag{3.8}$$

Since the distribution of  $Z_h$  is symmetric, we also have

$$\begin{aligned} (1 - \varepsilon) h^{-r} H_\alpha^{(r)} I_h(J_{k,m_h}^\delta) & \leq \frac{\sum_{i=1}^{N_h} \mathbb{P}(\min_{t \in \Gamma_{h\gamma\theta^{-2/\alpha}}(S_i^h)} Z_h(t) < -\theta)}{\theta^{2r/\alpha} \Psi(\theta)} \\ & \leq (1 + \varepsilon) h^{-r} H_\alpha^{(r)} I_h(J_{k,m_h}^\delta). \end{aligned} \tag{3.9}$$

We emphasize that these inequalities hold when the norm of the partition falls below a certain threshold that is independent of the choice of  $h$ . Following a similar procedure as above, we see that (3.8) and (3.9) continue to hold (for  $\max_{i=1, \dots, N_h} V_r(h^{-1} S_i^h)$  sufficiently small, and  $\theta$  large enough), if  $\max_{t \in \Gamma_{h\gamma\theta^{-2/\alpha}}(S_i^h)} Z_h(t)$  in (3.8) is replaced by  $\sup_{t \in S_i^h} Z_h(t)$ , and similarly,  $\min_{t \in \Gamma_{h\gamma\theta^{-2/\alpha}}(S_i^h)} Z_h(t)$  in (3.9) is replaced by  $\inf_{t \in S_i^h} Z_h(t)$ . Moreover, if we consider  $S_i^h(J_{k,m_h})$

and  $S_i^h(J_{k,m_h}^{-\delta})$  instead of  $S_i^h (= S_i^h(J_{k,m_h}^\delta))$ , these inequalities continue to hold. In particular, for  $J_{k,m_h}$  we obtain that as  $\theta \rightarrow \infty$ ,

$$\frac{\sum_{i=1}^{N_h(J_{k,m_h})} \mathbb{P}(\sup_{t \in S_i^h(J_{k,m_h})} Z_h(t) > \theta)}{\theta^{2r/\alpha} \Psi(\theta)} = (1 + o(1)) h^{-r} H_\alpha^{(r)} I_h(J_{k,m_h}), \tag{3.10}$$

where the  $o(1)$ -term is uniform in both  $k \in \{1, \dots, m_h\}$  and  $h \in (0, 1]$ .

*Part 2.* Here we show that uniformly in  $h$ , as  $\theta \rightarrow \infty$ ,

$$\sum_{k \leq m_h} \mathbb{P}\left(\sup_{t \in J_{k,m_h}} Z_h(t) > \theta\right) = (1 + o(1)) h^{-r} \theta^{2r/\alpha} \Psi(\theta) H_\alpha^{(r)} I_h(\mathcal{M}). \tag{3.11}$$

Again we will use the various approximations introduced at the beginning of the proof. Let  $\{S_i^h : i = 1, \dots, N_h\}$  denote the partition of  $J_{k,m_h}$  constructed in (iii). This partition consists of closed, non-overlapping subsets, i.e., their interiors are disjoint. Furthermore, letting  $B_i = \{\sup_{t \in S_i^h} Z_h(t) > \theta\}$ , we trivially have

$$\mathbb{P}\left(\sup_{t \in J_{k,m_h}} Z_h(t) > \theta\right) = \mathbb{P}\left(\bigcup_{i=1}^{N_h} B_i\right).$$

We now use that

$$\sum_{i=1}^{N_h} \mathbb{P}(B_i) - \sum_{1 \leq i < j \leq N_h} \mathbb{P}(B_i \cap B_j) \leq \mathbb{P}\left(\bigcup_{i=1}^{N_h} B_i\right) \leq \sum_{i=1}^{N_h} \mathbb{P}(B_i), \tag{3.12}$$

and show that the double sum on the left-hand side is negligible as compared to the (simple) sum, so that we essentially have upper and lower bounds for  $\mathbb{P}(\bigcup_{i=1}^{N_h} B_i)$  in terms of  $\sum_{i=1}^{N_h} \mathbb{P}(B_i)$ . To see this, first observe that by using (3.10), we obtain for  $\max_{i=1, \dots, N_h} V_r(h^{-1} S_i^h)$  small enough that, as  $\theta \rightarrow \infty$ ,

$$\sum_{i=1}^{N_h} \mathbb{P}(B_i) = O(\theta^{2r/\alpha} \Psi(\theta)), \tag{3.13}$$

where the order of the right-hand side is exact by (3.10). We thus only need to show that  $\sum_{1 \leq i < j \leq N_h} \mathbb{P}(B_i \cap B_j) = o(\theta^{2r/\alpha} \Psi(\theta))$  as  $\theta \rightarrow \infty$ . Our proof generalizes the result in Mikhal'eva and Piterbarg [36] by considering uniformity in  $h$  in a rescaled case. It will turn out that we obtain the desired result if the norm of the partition given by  $S_i^h$  can be chosen sufficiently small, uniformly in  $h$ , which is feasible, as discussed at the beginning of the proof.

Let  $U = \{(i, j) : B_i \text{ and } B_j \text{ are adjacent}\}$  and  $V = \{(i, j) : B_i \text{ and } B_j \text{ are not adjacent}\}$ , where ‘not adjacent’ means that their boundaries do not touch. Note that

$$\sum_{1 \leq i < j \leq N_h} \mathbb{P}(B_i \cap B_j) = \sum_{\substack{1 \leq i < j \leq N_h, \\ (i,j) \in U}} \mathbb{P}(B_i \cap B_j) + \sum_{\substack{1 \leq i < j \leq N_h, \\ (i,j) \in V}} \mathbb{P}(B_i \cap B_j). \tag{3.14}$$

In what follows we discuss the two sums on the right-hand side. First we consider the case of  $S_i^h, S_j^h$  being adjacent, that is,  $(i, j) \in U$ . The developments from Part 1 are here applied to  $S_i^h, S_j^h$  and  $S_i^h \cup S_j^h$ , respectively. We choose the points where the tangent spaces are placed to be the same for  $S_i^h, S_j^h$  and  $S_i^h \cup S_j^h$ , that is, we choose this point to lie on the boundary of both  $S_i^h$  and  $S_j^h$ . Simply denote this point as  $s$ . Then, by using the results from Part 1, for any  $\varepsilon > 0$ , when  $h^{-r} \max_{(i,j) \in U} V_r(S_i^h \cup S_j^h)$  is small enough and  $\theta$  is large enough, the bounds obtained as in Part 1 result in

$$\begin{aligned} \frac{\mathbb{P}(B_i \cap B_j)}{\theta^{2r/\alpha} \Psi(\theta)} &= \frac{\mathbb{P}(B_i) + \mathbb{P}(B_j) - \mathbb{P}(B_i \cup B_j)}{\theta^{2r/\alpha} \Psi(\theta)} \\ &\leq (1 + \varepsilon)h^{-r} H_\alpha^{(r)} \|D_s^h M_s\|_r V_r(S_i^h) + (1 + \varepsilon)h^{-r} H_\alpha^{(r)} \|D_s^h M_s\|_r V_r(S_j^h) \\ &\quad - (1 - \varepsilon)h^{-r} H_\alpha^{(r)} \|D_s^h M_s\|_r V_r(S_i^h \cup S_j^h) \\ &= 2\varepsilon h^{-r} H_\alpha^{(r)} \|D_s^h M_s\|_r [V_r(S_i^h) + V_r(S_j^h)]. \end{aligned}$$

Summing the right-hand side of this inequality over  $(i, j) \in U$ , results in a Riemann sum that approximates an integral over  $J_{k,m_h}$ . Since, by assumption,  $\lim_{h \rightarrow 0} D_t^h = D_t$  uniformly in  $t \in \mathcal{H}$ , and since the components of  $D_t$  are continuous and bounded in  $t \in \mathcal{H}$ , there exists  $c \in \mathbb{R}^+$  such that

$$\sup_{s \in \mathcal{M}, 0 < h \leq 1} \|D_s^h M_s\|_r \leq c. \tag{3.15}$$

Hence, noting that  $\varepsilon > 0$  is arbitrary, we have that as  $\max_{1 \leq i \leq N_h} V_r(h^{-1} S_i^h) \rightarrow 0$  and  $\theta \rightarrow \infty$ ,

$$\sum_{\substack{1 \leq i < j \leq N_h \\ (i,j) \in U}} \mathbb{P}(B_i \cap B_j) = o(\theta^{2r/\alpha} \Psi(\theta)). \tag{3.16}$$

Next, we proceed to consider the case that  $(i, j) \in V$ , i.e.,  $S_i^h, S_j^h$  are not adjacent on  $J_{k,m_h}$ . To find an upper bound for  $\mathbb{P}(B_i \cap B_j)$ , we first note that

$$\begin{aligned} \mathbb{P}(B_i \cap B_j) &= \mathbb{P}\left(\sup_{t \in S_i^h} Z_h(t) > \theta, \sup_{t \in S_j^h} Z_h(t) > \theta\right) \\ &\leq \mathbb{P}\left(\sup_{t \in S_i^h, s \in S_j^h} (Z_h(t) + Z_h(s)) > 2\theta\right). \end{aligned} \tag{3.17}$$

In order to further estimate this probability, we will use the following theorem from Borell [11].

**Theorem 3.1.** *Let  $\{X(t), t \in T\}$  be a real separable Gaussian process indexed by an arbitrary parameter set  $T$ , let*

$$\sigma^2 = \sup_{t \in T} \text{Var}X(t) < \infty, \quad m = \sup_{t \in T} \mathbb{E}X(t) < \infty,$$

and let the real number  $b$  be such that

$$\mathbb{P}\left(\sup_{t \in T} X(t) - \mathbb{E}X(t) \geq b\right) \leq \frac{1}{2}.$$

Then, for all  $x$

$$\mathbb{P}\left(\sup_{t \in T} X(t) > x\right) \leq 2\bar{\Phi}\left(\frac{x - m - b}{\sigma}\right).$$

We want to apply the above theorem to  $Z_h(t) + Z_h(s)$  with  $t \in S_i^h, s \in S_j^h$  and  $(i, j) \in V$  (see (3.17)), which requires some further preparations. First, observe that there exists a constant  $\zeta_1 > 0$  such that

$$\inf_{(i,j) \in V, t \in S_i^h, s \in S_j^h, 0 < h \leq 1} \|h^{-1}(t - s)\| > \zeta_1,$$

that is, the distance between any two nonadjacent elements of the partition (after rescaling) exceeds  $\zeta_1$  uniformly in  $h \in (0, 1]$ . This is due to the fact that the curvature of the manifold  $\mathcal{M}$  is bounded, and that  $V_r(h^{-1}S_j^h)$  is bounded away from zero uniformly in  $j$  and  $h$ . (See Lemma 3 of Genovese *et al.* [24] for more details underlying this argument.) The latter also implies that we can find a number  $N_0 > 0$  such that  $N_h$ , the number of sets  $S_i$ , satisfies  $N_h < N_0$  for all  $h$ . Assumption (2.6) implies that for  $\rho = \sup_{\|h^{-1}(t-s)\| \geq \zeta_1, 0 < h \leq 1} r_h(t, s)$ , we have  $\rho < 1$ , and thus we obtain

$$\sup_{0 < h \leq 1} \sup_{t \in S_i^h, s \in S_j^h} \text{Var}(Z_h(t) + Z_h(s)) \leq 2 + 2\rho$$

and

$$\sup_{0 < h \leq 1} \sup_{t \in S_i^h, s \in S_j^h} \mathbb{E}(Z_h(t) + Z_h(s)) = 0.$$

To see that the above theorem is applicable to  $Z_h(t) + Z_h(s)$  with  $t \in S_i^h$  and  $s \in S_j^h$ , it thus remains to show that there is a constant  $b$  with  $\mathbb{P}(\sup_{t \in S_i^h, s \in S_j^h} (Z_h(t) + Z_h(s)) > b) \leq \frac{1}{2}$ , and we do this now. First, note that

$$\begin{aligned} \mathbb{P}\left(\sup_{t \in S_i^h, s \in S_j^h} (Z_h(t) + Z_h(s)) > b\right) &\leq \mathbb{P}\left(\sup_{t \in J_{k,m_h}, s \in J_{k,m_h}} (Z_h(t) + Z_h(s)) > b\right) \\ &\leq \mathbb{P}\left(\sup_{t \in J_{k,m_h}} Z_h(t) > b/2\right). \end{aligned}$$

To bound the probability on the right-hand side further, observe that all the arguments in Part 1 hold uniformly in  $h$  as long as  $\theta$  is large enough. In other words, the conclusions there can

be restated by replacing  $\theta$  with  $x$  where  $x \rightarrow \infty$ . In particular, for any  $\varepsilon > 0$ , we can choose  $\max_{1 \leq i \leq N_h} V_r(h^{-1} S_i^h)$  small enough, such that

$$\mathbb{P}\left(\sup_{t \in J_{k,m_h}} Z_h(t) > x\right) \leq \sum_{i=1}^{N_h} \mathbb{P}\left(\sup_{t \in S_i^h} Z_h(t) > x\right) \leq (1 + \varepsilon)h^{-r} x^{2r/\alpha} \Psi(x) H_\alpha^{(r)} I_h(J_{k,m_h}),$$

for all  $1 \leq k \leq m_h$ . Hence, since  $x^{2r/\alpha} \Psi(x) \rightarrow 0$  as  $x \rightarrow \infty$ , we can find  $b$  such that  $\mathbb{P}(\sup_{t \in J_{k,m_h}} Z_h(t) > b/2) < 1/2$  for all  $1 \leq k \leq m_h$ , provided  $\max_{1 \leq i \leq N_h} V_r(h^{-1} S_i^h)$  is sufficiently small. Theorem 3.1 now gives (for  $\theta$  large enough) that

$$\mathbb{P}\left(\sup_{t \in S_i^h, s \in S_j^h} (Z_h(t) + Z_h(s)) > 2\theta\right) \leq 2\bar{\Phi}\left(\frac{\theta - b/2}{\sqrt{(1 + \rho)/2}}\right). \tag{3.18}$$

Since the total number of elements in the sum in (3.14) is bounded by  $N_h^2$ , it follows from (3.17) and (3.18) that uniformly in  $k$  (recall that the events  $B_i$  depend on  $k$ ), as  $\theta \rightarrow \infty$

$$\begin{aligned} \sum_{\substack{1 \leq i < j \leq N_h, \\ j-i > 1}} \mathbb{P}(B_i \cap B_j) &\leq 2N_h^2 \bar{\Phi}\left(\frac{\theta - b/2}{\sqrt{(1 + \rho)/2}}\right) \leq 2N_0^2 \bar{\Phi}\left(\frac{\theta - b/2}{\sqrt{(1 + \rho)/2}}\right) \\ &= o(\theta^{2r/\alpha} \Psi(\theta)), \end{aligned} \tag{3.19}$$

where we use the well-known fact that  $\lim_{u \rightarrow \infty} \frac{\bar{\Phi}(u)}{\Psi(u)} = 1$ .

Considering (3.13), (3.14), (3.16) and (3.19) and their corresponding conditions, we have

$$\mathbb{P}\left(\sup_{t \in J_{k,m_h}} Z_h(t) > \theta\right) = (1 + o(1)) \sum_{i=1}^{N_h} \mathbb{P}\left(\sup_{t \in S_i^h} Z_h(t) > \theta\right) \quad \text{as } \theta \rightarrow \infty, \tag{3.20}$$

where the  $o(1)$ -term is uniform in  $k$ . Combining (3.10) and (3.20), we have the asserted result (3.11) for  $\sup_{h \in (0,1]} \max_{1 \leq i \leq N_h} V_r(h^{-1} S_i^h)$  sufficiently small.

*Part 3.* Recall the definition of  $I(\mathcal{M})$  given right after (2.8). Using the expression of  $\theta$  in (3.1), we have, for any fixed  $z$  and as  $h \rightarrow 0$ , that

$$h^{-r} \theta^{2r/\alpha} \Psi(\theta) = \frac{h^{-r} \theta^{2r/\alpha - 1}}{\sqrt{2\pi}} \exp\left\{-\frac{\theta^2}{2}\right\} = \frac{\exp\{-z\}}{H_\alpha^{(r)} I(\mathcal{M})} (1 + o(1)) = O(1). \tag{3.21}$$

Observing that  $\max_{1 \leq k \leq m_h} V_r(h^{-1} J_{k,m_h}^{-\delta}) = O(\delta)$  (uniformly in  $h$ ), and using (3.21), we obtain for  $h$  small enough that

$$\begin{aligned} 0 &\leq \mathbb{P}\left(\sup_{t \in \mathcal{M}} Z_h(t) > \theta\right) - \mathbb{P}\left(\sup_{t \in \bigcup_{k \leq m_h} J_{k,m_h}^\delta} Z_h(t) > \theta\right) \\ &\leq \mathbb{P}\left(\sup_{t \in \mathcal{M} \setminus \bigcup_{k \leq m_h} J_{k,m_h}^\delta} Z_h(t) > \theta\right) \end{aligned}$$

$$\begin{aligned}
 &\leq \sum_{k=1}^{m_h} \mathbb{P} \left( \sup_{t \in J_{k,m_h}^{-\delta}} Z_h(t) > \theta \right) \\
 &\leq (1 + \varepsilon) h^{-r} \theta^{2r/\alpha} \Psi(\theta) H_\alpha^{(r)} \sum_{k=1}^{m_h} I_h(J_{k,m_h}^{-\delta}) \\
 &\leq O(\delta h^r) (1 + \varepsilon) h^{-r} H_\alpha^{(r)} c m_h \theta^{2r/\alpha} \Psi(\theta) \\
 &\leq O(\delta h^r) (1 + \varepsilon) H_\alpha^{(r)} c O(h^{-r}) \frac{\exp\{-z\}}{H_\alpha^{(r)} I(\mathcal{M})} \\
 &= O(\delta),
 \end{aligned}$$

where  $c$  is from (3.15). Similarly, we have

$$0 \leq \mathbb{P} \left( \inf_{t \in \mathcal{M}} Z_h(t) < -\theta \right) - \mathbb{P} \left( \inf_{t \in \bigcup_{k \leq m_h} J_{k,m_h}^\delta} Z_h(t) < -\theta \right) = O(\delta),$$

uniformly in  $0 < h \leq h_1$  for some  $h_1 > 0$ . Collecting what we have, we obtain that uniformly in  $0 < h \leq h_1$ ,

$$\mathbb{P} \left( \sup_{t \in \mathcal{M}} |Z_h(t)| \leq \theta \right) = \mathbb{P} \left( \sup_{t \in \bigcup_{k \leq m_h} J_{k,m_h}^\delta} |Z_h(t)| \leq \theta \right) + O(\delta) \tag{3.22}$$

and

$$\sum_{k=1}^{m_h} \mathbb{P} \left( \sup_{t \in J_{k,m_h}} |Z_h(t)| > \theta \right) = \sum_{k=1}^{m_h} \mathbb{P} \left( \sup_{t \in J_{k,m_h}^\delta} |Z_h(t)| > \theta \right) + O(\delta). \tag{3.23}$$

*Part 4.* Here we show that replacing  $\bigcup_{k \leq m_h} J_{k,m_h}^\delta$  by the dense ‘grid’  $\mathbb{T}_h^\delta$  (see (v)) leads to a negligible error in the corresponding extreme value probabilities.

We write  $\mathbb{T}_h^\delta = \bigcup_{1 \leq k \leq m_h} \Gamma_{h\gamma\theta^{-2/\alpha}}(J_{k,m}^\delta)$  as  $\{t_j, j = 1, \dots, N_h^*\}$ . Our assumptions ensure that  $N_h^* = O(\theta^{2r/\alpha} h^{-r} \gamma^{-r})$ . This behavior follows from the bounded curvature of the manifold  $\mathcal{M}$ , the construction of the triangulation, and the fact that  $V_r(\mathcal{M}) < \infty$ .

With (3.8), (3.15) and (3.21), we have

$$\mathbb{P} \left( \max_{t_j \in J_{k,m_h}^\delta} |Z_h(t_j)| > \theta \right) = O(h), \tag{3.24}$$

uniformly in  $k$  as  $h \rightarrow 0$ . In what follows, the maxima (or minima, respectively) taken over  $t_j$  run over  $j = 1, \dots, N_h^*$  (i.e., over all  $t_j \in \mathbb{T}_h^\delta$ ). This is not explicitly indicated in order to shorten the notation. We obtain that as  $h \rightarrow 0$ ,

$$\sum_{k=1}^{m_h} \log \left( 1 - \mathbb{P} \left( \max_{t_j \in J_{k,m_h}^\delta} |Z_h(t_j)| > \theta \right) \right) = (1 + o(1)) \sum_{k=1}^{m_h} \mathbb{P} \left( \max_{t_j \in J_{k,m_h}^\delta} |Z_h(t_j)| > \theta \right). \tag{3.25}$$

It follows from (3.8) and its analogous version with the maximum over the discrete set replaced by the supremum over  $t \in S_i^h$  (see discussion given below (3.9)), that for any  $\varepsilon > 0$ , and  $h, \gamma$  and the norm of partitions sufficiently small,

$$\begin{aligned} 0 &\leq \mathbb{P}\left(\sup_{t \in J_{k,m_h}^\delta} Z_h(t) > \theta\right) - \mathbb{P}\left(\max_{t_j \in J_{k,m_h}^\delta} Z_h(t_j) > \theta\right) \\ &\leq \sum_{i=1}^{N_h} \left[ \mathbb{P}\left(\sup_{t \in S_i^h} Z_h(t) > \theta\right) - \mathbb{P}\left(\max_{t_j \in S_i^h} Z_h(t_j) > \theta\right) \right] \\ &\leq \varepsilon h^{-r} \theta^{2r/\alpha} \Psi(\theta) H_\alpha^{(r)} I_h(J_{k,m_h}^\delta). \end{aligned}$$

Similarly, (3.9) and its corresponding ‘continuous’ version imply that, for  $h, \gamma$  and the norm of partitions sufficiently small, we have

$$0 \leq \mathbb{P}\left(\inf_{t \in J_{k,m_h}^\delta} Z_h(t) < -\theta\right) - \mathbb{P}\left(\min_{t_i \in J_{k,m_h}^\delta} Z_h(t_i) < -\theta\right) \leq \varepsilon h^{-r} \theta^{2r/\alpha} \Psi(\theta) H_\alpha^{(r)} I_h(J_{k,m_h}^\delta).$$

Consequently, if  $h, \gamma$ , and  $\max_{1 \leq k \leq m_h} V_r(h^{-1} J_{k,m_h}^\delta)$  are all small enough, we have

$$\begin{aligned} 0 &\leq \mathbb{P}\left(\sup_{t \in \bigcup_{k \leq m_h} J_{k,m_h}^\delta} |Z_h(t)| > \theta\right) - \mathbb{P}\left(\max_{t_j \in \bigcup_{k \leq m_h} J_{k,m_h}^\delta} |Z_h(t_j)| > \theta\right) \\ &\leq \sum_{k=1}^{m_h} \left[ \mathbb{P}\left(\sup_{t \in J_{k,m_h}^\delta} |Z_h(t)| > \theta\right) - \mathbb{P}\left(\max_{t_j \in J_{k,m_h}^\delta} |Z_h(t_j)| > \theta\right) \right] \\ &\leq \sum_{k=1}^{m_h} \left[ \mathbb{P}\left(\sup_{t \in J_{k,m_h}^\delta} Z_h(t) > \theta\right) + \mathbb{P}\left(\inf_{t \in J_{k,m_h}^\delta} Z_h(t) < -\theta\right) \right. \\ &\quad \left. - \mathbb{P}\left(\max_{t_j \in J_{k,m_h}^\delta} Z_h(t_j) > \theta\right) - \mathbb{P}\left(\min_{t_j \in J_{k,m_h}^\delta} Z_h(t_j) < -\theta\right) \right] \\ &\leq 2\varepsilon h^{-r} \theta^{2r/\alpha} \Psi(\theta) H_\alpha^{(r)} I_h\left(\bigcup_{k \leq m_h} J_{k,m_h}^\delta\right) \\ &\leq 2\varepsilon h^{-r} \theta^{2r/\alpha} \Psi(\theta) H_\alpha^{(r)} I_h(\mathcal{M}). \end{aligned} \tag{3.26}$$

Observe that by using (2.2) and the dominated convergence theorem, we have

$$\frac{I_h(\mathcal{M})}{I(\mathcal{M})} \rightarrow 1 \quad \text{as } h \rightarrow 0. \tag{3.27}$$

As a result of (3.21) and (3.27), we can write for  $\max_{1 \leq k \leq m_h} V_r(h^{-1} J_{k,m_h}^\delta)$  small enough that

$$\mathbb{P}\left(\sup_{t \in \bigcup_{k \leq m_h} J_{k,m_h}^\delta} |Z_h(t)| \leq \theta\right) = \mathbb{P}\left(\max_{t_j \in \bigcup_{k \leq m_h} J_{k,m_h}^\delta} |Z_h(t_j)| \leq \theta\right) + o(1) \quad (3.28)$$

and

$$\sum_{k=1}^{m_h} \mathbb{P}\left(\sup_{t \in J_{k,m_h}^\delta} |Z_h(t)| > \theta\right) = \sum_{k=1}^{m_h} \mathbb{P}\left(\max_{t_j \in J_{k,m_h}^\delta} |Z_h(t_j)| > \theta\right) + o(1), \quad (3.29)$$

as  $\gamma, h \rightarrow 0$ .

*Part 5.* Here we find an upper bound for the difference

$$\left| \mathbb{P}\left(\max_{t_j \in \bigcup_{k \leq m_h} J_{k,m_h}^\delta} |Z_h(t_j)| \leq \theta\right) - \prod_{k \leq m_h} \mathbb{P}\left(\max_{t_j \in J_{k,m_h}^\delta} |Z_h(t_j)| \leq \theta\right) \right|. \quad (3.30)$$

This step uses similar ideas as in the proof of Lemma 5.1 in Berman [4]. Define a probability measure  $\tilde{\mathbb{P}}$  such that, for any  $x_{t_j} \in \mathbb{R}$  with  $t_j \in \bigcup_{k \leq m_h} J_{k,m_h}^\delta$ ,

$$\tilde{\mathbb{P}}\left(Z_h(t_j) \leq x_{t_j}, t_j \in \bigcup_{k \leq m_h} J_{k,m_h}^\delta\right) = \prod_{k \leq m_h} \mathbb{P}(Z_h(t_j) \leq x_{t_j}, t_j \in J_{k,m_h}^\delta),$$

that is, under  $\tilde{\mathbb{P}}$  the vectors  $(Z_h(t_i) : t_i \in J_{k,m_h}^\delta \cap \mathbb{T}_h^\delta)$  and  $(Z_h(t_j) : t_j \in J_{k',m_h}^\delta \cap \mathbb{T}_h^\delta)$  are independent for  $k \neq k'$ . By Lemma 4.3, the difference in (3.30) can be bounded by

$$\begin{aligned} & 8 \sum_{\substack{k \leq m_h, k' \leq m_h, \\ k \neq k'}} \sum_{t_i \in J_{k,m_h}^\delta} \sum_{t_j \in J_{k',m_h}^\delta} \int_0^{|r_h(t_i, t_j)|} \phi(\theta, \theta, \lambda) d\lambda \\ &= 8 \sum_{\substack{k \leq m_h, k' \leq m_h, \\ k \neq k'}} \sum_{t_i \in J_{k,m_h}^\delta} \sum_{t_j \in J_{k',m_h}^\delta} \int_0^{|r_h(t_i, t_j)|} \frac{1}{2\pi(1-\lambda^2)^{1/2}} \exp\left(-\frac{\theta^2}{1+\lambda}\right) d\lambda \\ &\leq 8 \sum_{\substack{k \leq m_h, k' \leq m_h, \\ k \neq k'}} \sum_{t_i \in J_{k,m_h}^\delta} \sum_{t_j \in J_{k',m_h}^\delta} \frac{|r_h(t_i, t_j)|}{2\pi(1-(r_h(t_i, t_j))^2)^{1/2}} \exp\left(-\frac{\theta^2}{1+|r_h(t_i, t_j)|}\right). \end{aligned} \quad (3.31)$$

(Note that here the notation  $\{t_i \in J_{k,m_h}^\delta\}$  is a shortcut for  $\{t_i \in J_{k,m_h}^\delta \cap \mathbb{T}_h^\delta\}$ , and similarly for  $t_j$ .) For  $t_i \in J_{k,m_h}^\delta$  and  $t_j \in J_{k',m_h}^\delta$ , with  $k \neq k'$ , it follows from the boundedness of the curvature of the manifold  $\mathcal{M}$  that there exists  $\varsigma > 0$  such that  $\|h^{-1}(t_i - t_j)\| \geq \varsigma$ , uniformly for all  $0 < h \leq 1$ . (Similar arguments have been used above already.) Thus, we obtain from assumption (2.6) that there exists  $\eta > 0$ , dependent on  $\varsigma$ , with

$$|r_h(t_i, t_j)| < \eta < 1, \quad (3.32)$$

uniformly in  $t_i \in J_{k,m_h}^\delta$  and  $t_j \in J_{k',m_h}^\delta$ , with  $k \neq k'$  and  $0 < h \leq 1$ .

Let  $\omega$  be such that

$$0 < \omega < \frac{2}{(1 + \eta)} - 1.$$

In what follows, we take  $\gamma = v(h^{-1})^{1/3r}$ . We divide the triple sum in (3.31) into two parts by splitting up the pairs of indices  $i, j$  into two sets, according to whether  $\|t_i - t_j\| < h(N_h^*)^{\omega/r} \gamma \theta^{-2/\alpha}$  or not. For pairs satisfying this conditions, the number of summands in the triple sum is of the order  $O((N_h^*)^{\omega+1})$ , because there are a total of  $O(N_h^*)$  points and for each of them we have to consider at most  $O((N_h^*)^\omega)$  pairs. Taking into account (3.32), it follows that the sum in the first part of (3.31) is of the order

$$\begin{aligned} O\left((N_h^*)^{\omega+1} \exp\left\{-\frac{\theta^2}{1 + \eta}\right\}\right) &= O\left(\left(\frac{\theta^{2r/\alpha}}{h^r \gamma^r}\right)^{1+\omega} \exp\left\{-\frac{\theta^2}{1 + \eta}\right\}\right) \\ &= O\left(\left(\frac{(\log \frac{1}{h})^{r/\alpha}}{h^r \gamma^r}\right)^{1+\omega} \exp\left\{-\frac{2r \log \frac{1}{h}}{1 + \eta}\right\}\right) \\ &= O\left(h^{\frac{2r}{1+\eta} - r(1+\omega)} \left(\log \frac{1}{h}\right)^{\frac{(1+\omega)r}{\alpha}} \left(v\left(\frac{1}{h}\right)\right)^{-\frac{(1+\omega)\gamma}{3r}}\right), \end{aligned}$$

which tends to zero as  $h \rightarrow 0$ .

Now we consider the second part of (3.31), with pairs satisfying  $\|t_i - t_j\| \geq h(N_h^*)^{\omega/r} \gamma \theta^{-2/\alpha}$ . Noticing  $(1 + |r^h(t_i, t_j)|)^{-1} \geq 1 - |r^h(t_i, t_j)|$  and (3.32), we obtain the following bound:

$$8 \exp(-\theta^2) \sum_{\substack{k \leq m_h, k' \leq m_h, \\ k \neq k'}} \sum_{\substack{t_i \in J_{k, m_h}^\delta, t_j \in J_{k', m_h}^\delta, \\ \|t_i - t_j\| \geq h(N_h^*)^{\omega/r} \gamma \theta^{-2/\alpha}}} \frac{|r^h(t_i, t_j)|}{2\pi(1 - \eta^2)^{1/2}} \exp(\theta^2 |r^h(t_i, t_j)|). \tag{3.33}$$

By (2.7), and the fact that  $\theta^2 = O(\log h^{-1})$ , we have that

$$\sup_{\|t_i - t_j\| \geq h(N_h^*)^{\omega/r} \gamma \theta^{-2/\alpha}} \theta^2 |r^h(t_i, t_j)| \rightarrow 0 \quad \text{as } h \rightarrow 0.$$

Hence, (3.33) is of the order of

$$h^{2r} \sum_{\substack{k \leq m_h, k' \leq m_h, \\ k \neq k'}} \sum_{\substack{t_i \in J_{k, m_h}^\delta, t_j \in J_{k', m_h}^\delta, \\ \|t_i - t_j\| \geq h(N_h^*)^{\omega/r} \gamma \theta^{-2/\alpha}}} |r^h(t_i, t_j)|. \tag{3.34}$$

When  $h$  is sufficiently small, we have

$$\sup_{\|t_i - t_j\| \geq h(N_h^*)^{\omega/r} \gamma \theta^{-2/\alpha}} |r^h(t_i, t_j)| \leq \frac{v((N_h^*)^{\omega/r} \gamma \theta^{-2/\alpha})}{[\log((N_h^*)^{\omega/r} \gamma \theta^{-2/\alpha})]^{2r/\alpha}}. \tag{3.35}$$

Therefore, due to (2.7), (3.34) is of the order

$$\begin{aligned} & O\left(h^{2r} (N_h^*)^2 \frac{v((N_h^*)^{\omega/r} \gamma \theta^{-2/\alpha})}{[\log((N_h^*)^{\omega/r} \gamma \theta^{-2/\alpha})]^{2/\alpha}}\right) \\ &= O\left(\frac{(\log \frac{1}{h})^{2r/\alpha} v((N_h^*)^{\omega/r} \gamma \theta^{-2/\alpha})}{[\log(h^{-\omega} ((\log \frac{1}{h})^{1/\alpha} v(\frac{1}{h})^{-1/3r})^{\omega-1})]^{2r/\alpha} (v(\frac{1}{h}))^{2/3}}\right) \\ &= o(1) \quad \text{as } h \rightarrow 0. \end{aligned}$$

We have proved that the triple sum in (3.31) tends to zero as  $h$  goes to zero. Consequently, with this choice of  $\gamma$ , and as  $h \rightarrow 0$ ,

$$\mathbb{P}\left(\max_{t_j \in \bigcup_{k \leq m_h} J_{k,m_h}^\delta} |Z_h(t_j)| \leq \theta\right) = \prod_{k \leq m_h} \mathbb{P}\left(\max_{t_j \in J_{k,m_h}^\delta} |Z_h(t_j)| \leq \theta\right) + o(1), \quad (3.36)$$

where  $\delta > 0$  is fixed and small enough.

*Final part.* Collecting the above approximations we obtain the following. For  $\delta > 0$ ,  $\sup_{h \in (0,1]} \max_{1 \leq i \leq N_h} V_r(h^{-1} S_i^h)$  fixed and chosen small enough, and  $\gamma = v(h^{-1})^{1/3r}$ , we have that as  $h \rightarrow 0$ ,

$$\begin{aligned} & \mathbb{P}\left(\sup_{t \in \mathcal{M}} |Z_h(t)| \leq \theta\right) \\ & \stackrel{(3.22)}{=} \mathbb{P}\left(\sup_{t \in \bigcup_{k \leq m_h} J_{k,m_h}^\delta} |Z_h(t)| \leq \theta\right) + o(1) \\ & \stackrel{(3.28)}{=} \mathbb{P}\left(\max_{t_j \in \bigcup_{k \leq m_h} J_{k,m_h}^\delta} |Z_h(t_j)| \leq \theta\right) + o(1) \\ &= \mathbb{P}\left(\bigcap_{k \leq m_h} \left(\max_{t_j \in J_{k,m_h}^\delta} |Z_h(t_j)| \leq \theta\right)\right) + o(1) \\ & \stackrel{(3.36)}{=} \prod_{k \leq m_h} \mathbb{P}\left(\max_{t_j \in J_{k,m_h}^\delta} |Z_h(t_j)| \leq \theta\right) + o(1) \\ &= \exp\left\{\sum_{k \leq m_h} \log\left(1 - \mathbb{P}\left(\max_{t_j \in J_{k,m_h}^\delta} |Z_h(t_j)| > \theta\right)\right)\right\} + o(1) \\ & \stackrel{(3.25)}{=} \exp\left\{-(1 + o(1)) \sum_{k \leq m_h} \mathbb{P}\left(\max_{t_j \in J_{k,m_h}^\delta} |Z_h(t_j)| > \theta\right)\right\} + o(1) \\ & \stackrel{(3.29)}{=} \exp\left\{-(1 + o(1)) \left[\sum_{k \leq m_h} \mathbb{P}\left(\sup_{t \in J_{k,m_h}^\delta} |Z_h(t)| > \theta\right) - o(1)\right]\right\} + o(1) \end{aligned}$$

$$\stackrel{(3.23)}{=} \exp\left\{-2(1+o(1)) \sum_{k \leq m_h} \mathbb{P}\left(\sup_{t \in J_{k,m_h}} Z_h(t) > \theta\right)\right\} + o(1)$$

$$\stackrel{(3.11)}{=} \exp\{-2(1+o(1))h^{-r}\theta^{2r/\alpha}\Psi(\theta)H_\alpha^{(r)}I_h(\mathcal{M})\} + o(1).$$

This completes our proof by using (3.21) and (3.27). □

### 4. Miscellaneous

In this section, we collect some miscellaneous results and definitions that are needed above. We present them in a separate section in order to not interrupt the flow of the above proof.

*Definition of generalized Pickands constant* (following Piterbarg and Stamatovich [39]). For  $0 < \alpha \leq 2$ , let  $\chi_\alpha(t)$  be a continuous Gaussian field with  $\mathbb{E}\chi_\alpha(t) = -\|t\|^\alpha$  and  $\text{Cov}(\chi_\alpha(t), \chi_\alpha(s)) = \|t\|^\alpha + \|s\|^\alpha - \|t-s\|^\alpha$  where  $s, t \in \mathbb{R}^n$ . The existence of such a field  $\chi_\alpha(t)$  follows from Mikhaleva and Piterbarg [36].

For any compact set  $\mathcal{T} \subset \mathbb{R}^n$  define

$$H_\alpha(\mathcal{T}) = \mathbb{E} \exp\left(\sup_{t \in \mathcal{T}} \chi_\alpha(t)\right).$$

Let  $D$  be a non-degenerated  $n \times n$  matrix. For a set  $A \subset \mathbb{R}^n$ , let  $DA = \{Dx, x \in A\}$  denote the image of  $A$  under  $D$ . For any  $q > 0$ , we let

$$[0, q]^r = \{t : t_i \in [0, q], i = 1, \dots, r; t_i = 0, i = r + 1, \dots, n\}$$

denote a cube of dimension  $r$  generated by the first  $r$  coordinates in  $\mathbb{R}^n$ . Let

$$H_\alpha^{D\mathbb{R}^r} = \lim_{q \rightarrow \infty} \frac{H_\alpha(D[0, q]^r)}{\lambda_r(D[0, q]^r)},$$

where  $\lambda_r$  denotes Lebesgue measure in  $\mathbb{R}^r$ . It is known that  $H_\alpha^{D\mathbb{R}^r}$  exists, and  $0 < H_\alpha^{D\mathbb{R}^r} < \infty$  (see Beljaev and Piterbarg [3]). For  $D = I$ , the identity matrix, we write  $H_\alpha^{(r)} = H_\alpha^{I\mathbb{R}^r}$ . Since by definition the random field  $\chi_\alpha(\cdot)$  is isotropic,  $H_\alpha^{D\mathbb{R}^r} = H_\alpha^{(r)}$  for any orthogonal matrix  $D$ . The constant  $H_\alpha^{(n)}$  is the (generalized) *Pickands constant*.

Furthermore, for positive integers  $l$  and  $\gamma > 0$ , let

$$\begin{aligned} C^r(l, \gamma) &= \{t\gamma : t_i \in [0, l] \cap \mathbb{N}_0, i = 1, \dots, r; t_i = 0, i = r + 1, \dots, n\} \\ &= \gamma([0, l]^r \cap \mathbb{N}_0^n), \end{aligned}$$

and  $H_\alpha^{D, (r)}(l, \gamma) = H_\alpha(DC^r(l, \gamma))$ . Again, for orthogonal  $D$ , due to isotropy of  $\chi_\alpha(\cdot)$ , we write  $H_\alpha^{(r)}(l, \gamma) = H_\alpha^{D, (r)}(l, \gamma)$ . We let

$$H_\alpha^{(r)}(\gamma) = \lim_{l \rightarrow \infty} \frac{H_\alpha^{(r)}(l, \gamma)}{l^r},$$

assuming this limit exists, and for  $r = n$ , we simply write  $H_\alpha(l, \gamma)$  and  $H_\alpha(\gamma)$  instead of  $H_\alpha^{(n)}(l, \gamma)$  and  $H_\alpha^{(n)}(\gamma)$ , respectively. We have the following lemma from Bickel and Rosenblatt [9].

**Lemma 4.1.**  $H_\alpha^{(r)} = \lim_{\gamma \rightarrow 0} \frac{H_\alpha^{(r)}(\gamma)}{\gamma^r}$ .

In the following, we present further results for Gaussian fields that are used in the proof. Lemma 4.2 is an extension of Lemma A3 in Bickel and Rosenblatt [10], Lemma 3 and Lemma 5 of Bickel and Rosenblatt [9], and also of Lemma 2.5 in Pickands [38].

**Lemma 4.2.** Let  $X(t)$  be a centered homogeneous Gaussian field on  $\mathbb{R}^n$  with covariance function

$$r(t) = \mathbb{E}(X(t+s)X(s)) = 1 - \|t\|^\alpha + o(\|t\|^\alpha).$$

Let  $\mathcal{T}$  be a Jordan measurable set embedded in a  $r$ -dimensional linear space. For  $\gamma, x > 0$ , let  $\mathcal{G}(\mathcal{T}, \gamma, x)$  be a collection of points defining a mesh contained in  $\mathcal{T}$  with mesh size  $\gamma x^{-2/\alpha}$ . Assume that for  $\|t\|$  small enough,

$$\xi(\|t\|) := \inf_{0 < \|s\| \leq \|t\|} \|s\|^{-\alpha} (1 - r(s))/2 > 0. \tag{4.1}$$

Then

$$\lim_{x \rightarrow \infty} \frac{\mathbb{P}(\max\{X(t) : t \in \mathcal{G}(\mathcal{T}, \gamma, x)\} > x)}{x^{2r/\alpha} \Psi(x)} = V_r(\mathcal{T}) \frac{H_\alpha^{(r)}(\gamma)}{\gamma^r} \tag{4.2}$$

and

$$\lim_{x \rightarrow \infty} \frac{\mathbb{P}(\sup\{X(t) : t \in \mathcal{T}\} > x)}{x^{2r/\alpha} \Psi(x)} = V_r(\mathcal{T}) H_\alpha^{(r)}, \tag{4.3}$$

uniformly in  $\mathcal{T} \in \mathcal{E}_c$ , where  $\mathcal{E}_c$  is the collection of all  $r$ -dimensional Jordan measurable sets with diameter bounded by  $c < \infty$ . Similarly,

$$\lim_{x \rightarrow \infty} \frac{\mathbb{P}(\inf\{X(t) : t \in \mathcal{T}\} < -x)}{x^{2r/\alpha} \Psi(x)} = V_r(\mathcal{T}) H_\alpha^{(r)}, \tag{4.4}$$

uniformly in  $\mathcal{T} \in \mathcal{E}_c$ .

**Proof.** The results in Lemma 3 and Lemma 5 of Bickel and Rosenblatt [9] are similar but they are only given for two-dimensional squares. It is straightforward to generalize them to hyperrectangles and further to Jordan measurable sets. □

**Lemma 4.3 (Lemma 4.1 of Berman [4] as given in Lemma A4 of Bickel and Rosenblatt [10]).** *Let*

$$\phi(x, y, \rho) = \frac{1}{2\pi(1 - \rho^2)^{1/2}} \exp\left\{-\frac{x^2 - 2\rho xy + y^2}{2(1 - \rho^2)}\right\}.$$

*Let  $\Sigma_1 = \{r_{ij}\}, \Sigma_2 = \{s_{ij}\}$  be  $N \times N$  nonnegative semi-definite matrices with  $r_{ii} = s_{ii} = 1$  for all  $i$ . Let  $X = (X_1, \dots, X_N)$  be a mean 0 Gaussian vector with covariance matrices  $\Sigma_1$  or  $\Sigma_2$ , corresponding to probability measures  $\mathbb{P}_{\Sigma_1}$  and  $\mathbb{P}_{\Sigma_2}$ , respectively. Let  $u_1, \dots, u_N$  be nonnegative numbers and  $u = \min_j u_j$ . Then*

$$\left| \mathbb{P}_{\Sigma_1}[X_j \leq u_j, 1 \leq j \leq N] - \mathbb{P}_{\Sigma_2}[X_j \leq u_j, 1 \leq j \leq N] \right| \leq 4 \sum_{i,j} \left| \int_{s_{ij}}^{r_{ij}} \phi(u, u, \lambda) d\lambda \right|.$$

*Further discussion of Example (ii) (given below Definition 2.1).* Here we show the asserted local equi- $(\alpha, D_t^h)$ -stationarity of the fields given in Example (ii) above, and we also show that for this example the conditions of the main theorem can be verified.

*Local equi- $(\alpha, D_t^h)$ -stationarity.* We show that the random fields  $\{Z_h(t), t \in \mathbb{S}^{n-1}\}, h \in (0, 1]$ , with covariance function (2.4), are locally equi- $(\alpha, D_t^h)$ -stationary with

$$D_s^h = hg_h(s)\mathbf{I}_n.$$

First, note that  $d(t_1, t_2) = \arccos\langle t_1, t_2 \rangle = \arccos(1 - \frac{1}{2}\|t_1 - t_2\|^2)$  for  $t_1, t_2 \in \mathbb{S}^{n-1}$ . Using a Taylor series expansion, we have  $d(t_1, t_2) = \|t_1 - t_2\|(1 + o(1))$  as  $\|t_1 - t_2\| \rightarrow 0$ . For any  $h \in (0, 1]$  and  $s, t_1, t_2 \in \mathbb{S}^{n-1}$ , we have

$$\frac{h^\alpha [1 - r_h(t_1, t_2)]}{\|D_s^h(t_1 - t_2)\|^\alpha} = \frac{1 - \exp[-|g_h(\frac{t_1+t_2}{\|t_1+t_2\|}) \arccos\langle t_1, t_2 \rangle|^\alpha]}{|g_h(s)\|t_1 - t_2\|^\alpha}.$$

Therefore, as  $h^{-1} \max\{\|t_1 - s\|, \|t_2 - s\|\} \rightarrow 0$ ,

$$\frac{h^\alpha [1 - r_h(t_1, t_2)]}{\|D_s^h(t_1 - t_2)\|^\alpha} \approx \frac{|g_h(\frac{t_1+t_2}{\|t_1+t_2\|}) \arccos\langle t_1, t_2 \rangle|^\alpha}{|g_h(s)\|t_1 - t_2\|^\alpha} \approx \frac{|g_h(\frac{t_1+t_2}{\|t_1+t_2\|})|^\alpha}{|g_h(s)|^\alpha} \approx 1,$$

uniformly in  $h$  and  $s$ . This is (2.1).

*Verification of conditions of Theorem 2.1.* For any  $h \in (0, 1]$ , and  $t_1, t_2 \in \mathbb{S}^{n-1}$  with  $\|t_1 - t_2\| > h\delta$ , we have

$$\begin{aligned} |r_h(t_1, t_2)| &= \exp\left[-\left|g_h\left(\frac{t_1 + t_2}{\|t_1 + t_2\|}\right) \arccos\left(1 - \frac{\|t_1 - t_2\|^2}{2}\right)\right|^\alpha\right] \\ &\leq \exp\left[-\left|g_h\left(\frac{t_1 + t_2}{\|t_1 + t_2\|}\right) \arccos\left(1 - \frac{h^2\delta^2}{2}\right)\right|^\alpha\right]. \end{aligned}$$

There exists  $h_0 > 0$  such that, for  $0 < h < h_0$ ,  $\arccos(1 - \frac{1}{2}h^2u^2) = hu(1 + o(1))$  uniformly in  $\delta$ . Consequently, we have for  $Q_{h_0}(u)$  as in (2.10) that

$$Q_{h_0}(u) \leq \exp \left\{ - \left| \left[ \inf_{0 < h < h_0, s \in \mathbb{S}^{n-1}} hg_h(s) \right] u(1 + o(1)) \right|^\alpha \right\}.$$

By using the assumed upper and lower bounds for  $hg_h(s)$ , it follows that conditions (2.2), (2.6) and (2.7) are satisfied (where  $Q(u)$  is replaced with  $Q_{h_0}(u)$  – see Remark 1 below Theorem 2.1). It is now straightforward to apply Theorem 2.1 to  $\{Z_h(t), t \in \mathbb{S}^{n-1}\}$ .

## Acknowledgements

Many of the basic ideas of this work have been developed while Wanli Qiao was a PhD student at the University of California, Davis.

This research was partially supported by the NSF Grant DMS 1107206.

The authors would like to thank the Associate Editor and the referees for careful reading of the manuscript and for insightful comments that lead to significant improvements. In particular, the formulation of the problem via rescaled fields, rather than rescaled index sets, was suggested by the Associate Editor.

## References

- [1] Adler, R.J. and Taylor, J.E. (2007). *Random Fields and Geometry. Springer Monographs in Mathematics*. New York: Springer. [MR2319516](#)
- [2] Azaïs, J.-M. and Wschebor, M. (2009). *Level Sets and Extrema of Random Processes and Fields*. Hoboken, NJ: Wiley. [MR2478201](#)
- [3] Beljaev, Ju.K. and Piterbarg, V.I. (1972). The asymptotic behavior of the average number of the A-points of upcrossings of a Gaussian field beyond a high level. *Dokl. Akad. Nauk SSSR* **203** 9–12. [MR0300334](#)
- [4] Berman, S.M. (1971). Asymptotic independence of the numbers of high and low level crossings of stationary Gaussian processes. *Ann. Math. Stat.* **42** 927–945. [MR0285047](#)
- [5] Berman, S.M. (1974). Sojourns and extremes of Gaussian processes. *Ann. Probab.* **2** 999–1026. [MR0372976](#)
- [6] Berman, S.M. (1982). Sojourns and extremes of stationary processes. *Ann. Probab.* **10** 1–46. [MR0637375](#)
- [7] Berman, S.M. (1992). *Sojourns and Extremes of Stochastic Processes. The Wadsworth & Brooks/Cole Statistics/Probability Series*. Pacific Grove, CA: Wadsworth & Brooks/Cole Advanced Books & Software. [MR1126464](#)
- [8] Biau, G., Cadre, B. and Pelletier, B. (2008). Exact rates in density support estimation. *J. Multivariate Anal.* **99** 2185–2207. [MR2463383](#)
- [9] Bickel, P. and Rosenblatt, M. (1973). Two-dimensional random fields. In *Multivariate Analysis, III (Proc. Third Internat. Sympos., Wright State Univ., Dayton, Ohio, 1972)* 3–15. New York: Academic Press. [MR0348832](#)
- [10] Bickel, P.J. and Rosenblatt, M. (1973). On some global measures of the deviations of density function estimates. *Ann. Statist.* **1** 1071–1095. [MR0348906](#)

- [11] Borell, C. (1975). The Brunn–Minkowski inequality in Gauss space. *Invent. Math.* **30** 207–216. [MR0399402](#)
- [12] Chen, Y.-C., Genovese, C.R. and Wasserman, L. (2013). Uncertainty measures and limiting distributions for filament estimation. Available at [arXiv:1312.2098](#).
- [13] Chen, Y.-C., Genovese, C.R. and Wasserman, L. (2014). Generalized mode and ridge estimation. Available at [arXiv:1406.1803](#).
- [14] Chen, Y.-C., Genovese, C.R. and Wasserman, L. (2015). Asymptotic theory for density ridges. *Ann. Statist.* **43** 1896–1928. [MR3375871](#)
- [15] Chen, Y.-C., Genovese, C.R. and Wasserman, L. (2016). Density level set: Asymptotics, inference, and visualization. Available at [arXiv:1504.05438](#).
- [16] Cheng, D. (2015). Excursion probabilities of isotropic and locally isotropic Gaussian random fields on manifolds. Available at [arXiv:1504.08047](#).
- [17] Chernozhukov, V., Chetverikov, D. and Kato, K. (2014). Gaussian approximation of suprema of empirical processes. *Ann. Statist.* **42** 1564–1597. [MR3262461](#)
- [18] Chilès, J.-P. and Delfiner, P. (1999). *Geostatistics: Modeling Spatial Uncertainty*. *Wiley Series in Probability and Statistics: Applied Probability and Statistics*. New York: Wiley. [MR1679557](#)
- [19] Cuevas, A., González-Manteiga, W. and Rodríguez-Casal, A. (2006). Plug-in estimation of general level sets. *Aust. N. Z. J. Stat.* **48** 7–19. [MR2234775](#)
- [20] Cuevas, A. and Rodríguez-Casal, A. (2004). On boundary estimation. *Adv. in Appl. Probab.* **36** 340–354. [MR2058139](#)
- [21] de Laat, D. (2011). Approximating manifolds by meshes: Asymptotic bounds in higher codimension. Master thesis, University of Groningen.
- [22] Federer, H. (1959). Curvature measures. *Trans. Amer. Math. Soc.* **93** 418–491. [MR0110078](#)
- [23] Genovese, C.R., Perone-Pacifico, M., Verdinelli, I. and Wasserman, L. (2012). The geometry of non-parametric filament estimation. *J. Amer. Statist. Assoc.* **107** 788–799. [MR2980085](#)
- [24] Genovese, C.R., Perone-Pacifico, M., Verdinelli, I. and Wasserman, L. (2012). Minimax manifold estimation. *J. Mach. Learn. Res.* **13** 1263–1291. [MR2930639](#)
- [25] Genovese, C.R., Perone-Pacifico, M., Verdinelli, I. and Wasserman, L. (2014). Nonparametric ridge estimation. *Ann. Statist.* **42** 1511–1545. [MR3262459](#)
- [26] Gneiting, T. (2013). Strictly and non-strictly positive definite functions on spheres. *Bernoulli* **19** 1327–1349. [MR3102554](#)
- [27] Hall, P., Qian, W. and Titterton, D.M. (1992). Ridge finding from noisy data. *J. Comput. Graph. Statist.* **1** 197–211. [MR1270818](#)
- [28] Huang, C., Zhang, H. and Robeson, S.M. (2011). On the validity of commonly used covariance and variogram functions on the sphere. *Math. Geosci.* **43** 721–733. [MR2824128](#)
- [29] Hüsler, J. (1990). Extreme values and high boundary crossings of locally stationary Gaussian processes. *Ann. Probab.* **18** 1141–1158. [MR1062062](#)
- [30] Hüsler, J. (1999). Extremes of Gaussian processes, on results of Piterbarg and Seleznev. *Statist. Probab. Lett.* **44** 251–258. [MR1711625](#)
- [31] Hüsler, J., Piterbarg, V. and Seleznev, O. (2003). On convergence of the uniform norms for Gaussian processes and linear approximation problems. *Ann. Appl. Probab.* **13** 1615–1653. [MR2023892](#)
- [32] Koltchinskii, V., Sakhanenko, L. and Cai, S. (2007). Integral curves of noisy vector fields and statistical problems in diffusion tensor imaging: Nonparametric kernel estimation and hypotheses testing. *Ann. Statist.* **35** 1576–1607. [MR2351098](#)
- [33] Konakov, V.D. and Piterbarg, V.I. (1984). On the convergence rate of maximal deviation distribution for kernel regression estimates. *J. Multivariate Anal.* **15** 279–294. [MR0768499](#)
- [34] Leadbetter, M.R., Lindgren, G. and Rootzén, H. (1983). *Extremes and Related Properties of Random Sequences and Processes*. *Springer Series in Statistics*. New York: Springer. [MR0691492](#)

- [35] Lindgren, G. and Rychlik, I. (1995). How reliable are contour curves? Confidence sets for level contours. *Bernoulli* **1** 301–319. [MR1369163](#)
- [36] Mikhaleva, T.L. and Piterbarg, V.I. (1997). On the distribution of the maximum of a Gaussian field with constant variance on a smooth manifold. *Theory Probab. Appl.* **41** 367–379.
- [37] Pickands, J. III (1969). Asymptotic properties of the maximum in a stationary Gaussian process. *Trans. Amer. Math. Soc.* **145** 75–86. [MR0250368](#)
- [38] Pickands, J. III (1969). Upcrossing probabilities for stationary Gaussian processes. *Trans. Amer. Math. Soc.* **145** 51–73. [MR0250367](#)
- [39] Piterbarg, V. and Stamatovich, S. (2001). On maximum of Gaussian non-centered fields indexed on smooth manifolds. In *Asymptotic Methods in Probability and Statistics with Applications (St. Petersburg, 1998)*. *Stat. Ind. Technol.* 189–203. Boston, MA: Birkhäuser. [MR1890326](#)
- [40] Piterbarg, V.I. (1996). *Asymptotic Methods in the Theory of Gaussian Processes and Fields. Translations of Mathematical Monographs* **148**. Providence, RI: Amer. Math. Soc. Translated from the Russian by V.V. Piterbarg. Revised by the author. [MR1361884](#)
- [41] Qiao, W. and Polonik, W. (2016). Theoretical analysis of nonparametric filament estimation. *Ann. Statist.* **44** 1269–1297. [MR3485960](#)
- [42] Qiao, W. and Polonik, W. (2016). Investigation of nonparametric confidence regions for density level sets. In preparation.
- [43] Rosenblatt, M. (1976). On the maximal deviation of  $k$ -dimensional density estimates. *Ann. Probab.* **4** 1009–1015. [MR0428580](#)
- [44] Seleznev, O. (1996). Large deviations in the piecewise linear approximation of Gaussian processes with stationary increments. *Adv. in Appl. Probab.* **28** 481–499. [MR1387887](#)
- [45] Seleznev, O. (2005). Asymptotic behavior of mean uniform norms for sequences of Gaussian processes and fields. *Extremes* **8** 161–169. [MR2275916](#)
- [46] Seleznev, O.V. (1991). Limit theorems for maxima and crossings of a sequence of Gaussian processes and approximation of random processes. *J. Appl. Probab.* **28** 17–32. [MR1090444](#)
- [47] Sharpnack, J. and Arias-Castro, E. (2016). Exact asymptotics for the scan statistic and fast alternatives. *Electron. J. Stat.* **10** 2641–2684. [MR3546971](#)
- [48] Slepian, D. (1962). The one-sided barrier problem for Gaussian noise. *Bell Syst. Tech. J.* **41** 463–501. [MR0133183](#)
- [49] Tan, Z. (2015). Limit laws on extremes of non-homogeneous Gaussian random fields. Available at [arXiv:1501.04422](#).
- [50] Tan, Z., Hashorva, E. and Peng, Z. (2012). Asymptotics of maxima of strongly dependent Gaussian processes. *J. Appl. Probab.* **49** 1106–1118. [MR3058991](#)
- [51] Thäle, C. (2008). 50 years sets with positive reach – A survey. *Surv. Math. Appl.* **3** 123–165. [MR2443192](#)
- [52] van der Vaart, A. and van Zanten, H. (2007). Bayesian inference with rescaled Gaussian process priors. *Electron. J. Stat.* **1** 433–448. [MR2357712](#)

Received August 2015 and revised October 2016