TAIL APPROXIMATIONS OF INTEGRALS OF GAUSSIAN RANDOM FIELDS

BY JINGCHEN LIU[1]

Columbia University

This paper develops asymptotic approximations of $P(\int_T e^{f(t)} dt > b)$ as $b \to \infty$ for a homogeneous smooth Gaussian random field, f, living on a compact d-dimensional Jordan measurable set T. The integral of an exponent of a Gaussian random field is an important random variable for many generic models in spatial point processes, portfolio risk analysis, asset pricing and so forth.

The analysis technique consists of two steps: 1. evaluate the tail probability $P(\int_\Xi e^{f(t)} dt > b)$ over a small domain Ξ depending on b, where $\operatorname{mes}(\Xi) \to 0$ as $b \to \infty$ and $\operatorname{mes}(\cdot)$ is the Lebesgue measure; 2. with Ξ appropriately chosen, we show that $P(\int_T e^{f(t)} dt > b) = (1 + o(1)) \operatorname{mes}(T) \times \operatorname{mes}^{-1}(\Xi) P(\int_\Xi e^{f(t)} dt > b)$.

1. Introduction. We consider a Gaussian random field living on a d-dimensional domain $T \subset R^d$, $\{f(t) : t \in T\}$. For every finite subset $\{t_1, \ldots, t_n\} \subset T$, $(f(t_1), \ldots, f(t_n))$ is a multivariate Gaussian random vector. The quantity of interest is

$$P\left(\int_T e^{f(t)} dt > b \right)$$

as $b \to \infty$.

The motivations of the study of $\int_T e^{f(t)} dt$ are from multiple sources. We will present a few of them. Consider a point process on R^d associated with a Poisson random measure $\{N_A\}_{A \in \mathcal{B}}$ with intensity $\lambda(t)$, where $\mathcal{B}$ represents the Borel sets of R^d. One important task in spatial modeling is to build in dependence structures. A popular strategy is to let $f(t) = \log \lambda(t)$, which can take all values in R, and model $f(t)$ as a Gaussian random field. Then, $\int_A e^{f(t)} dt = E(N(A)|\lambda(\cdot))$ for all $A \in \mathcal{B}$. With the multivariate Gaussian structure, it is easy to include linear predictors in the intensity process. For instance, [19] models $f(t) = U(t) + W(t)$, where $U(t)$ is the observed (deterministic) covariate process and $W(t)$ is a stationary AR(1) process. Similar models can be found in [18, 23, 40] which are special cases of the Cox process [21, 22]. Such a modeling approach has been applied

Key words and phrases. Gaussian random field, extremes.

to many disciplines, a short list of which is as follows: astronomy, epidemiology, geography, ecology, material science and so forth.

In portfolio risk analysis, consider a portfolio consisting of equally weighted assets $(S_1, \ldots, S_n)$. One stylized model is that $(\log S_1, \ldots, \log S_n)$ is a multivariate normal random vector (cf. [7, 12, 25, 26, 30]). The value of the portfolio $S = \sum_{i=1}^{n} S_i$ is then the sum of correlated log-normal random variables. If one can represent each asset price by the value of a Gaussian random field at one location in T, that is, $\log S_i = f(t_i)$. As the portfolio size tends to infinity and the asset prices become more correlated, the limit of the unit share price of the portfolio is $\lim_{n \to \infty} S/n = \int_T e^{f(t)} \, dt$. For more general cases, such as unequally weighted portfolios, the integral is possibly with respect to some other measures instead of the Lebesgue measure.

In option pricing, if we let $S(t)$ be a geometric Brownian motion (cf. [14], Chapter 5 of [27], Chapter 3.2 of [29]), the payoff function of an Asian option (with expiration time T) is a function of $\int_0^T S(t) \, dt$. For instance, the payoff of an Asian call option with strike price K is $\max(\int_0^T S(t) \, dt - K, 0)$; the payoff of a digital Asian call option is $I(\int_0^T S(t) \, dt > K)$.

We want to emphasize that the extreme behavior of $\int_T e^{f(t)} \, dt$ connects closely to that of $\sup_T f(t)$. As we will show in Theorem 1, with the threshold u appropriately chosen according to b, the probabilities of events $\{\int_T e^{f(t)} \, dt > b\}$ and $\{\sup_T f(t) \, dt > u\}$ have asymptotically the same decaying rate. It suggests that these two events have substantial overlap with each other. Therefore, we will borrow the intuitions and existing results on the high excursion of the supremum of random fields for the analysis of $\int_T e^{f(t)} \, dt$.

There is a vast literature on the extremes of Gaussian random fields mostly focusing on the tail probabilities of $\sup_T f(t)$ and its associated geometry. The results contain general bounds on $P(\sup_T f(t) > b)$ as well as sharp asymptotic approximations as $b \to \infty$. A partial literature contains [13, 16, 17, 31–33, 35, 37]. Several methods have been introduced to obtain asymptotic approximations, each of which imposes different regularity conditions on the random fields. A few examples are given as follows. The double sum method [34] requires expansions of the covariance function and locally stationary structure. The Euler–Poincaré Characteristic of the excursion set $[\chi(A_b)]$ approximation uses the fact that $P(M > b) \approx E(\chi(A_b))$, which requires the random field to be at least twice differentiable [1, 5, 38]. The tube method [36] uses the Karhunen–Loève expansion and imposes differentiability assumptions on the covariance function (fast decaying eigenvalues). The Rice method [9–11] represents the distribution of M (density function) in an implicit form. Recently, the efficient simulation algorithms are explored by [2, 3]. These two papers provided computation schemes that run in polynomial time to compute the tail probabilities for all Hölder continuous Gaussian random fields and in constant time for twice differentiable and homogeneous fields. In addition, [4] studied the geometric properties of a high level excursion set

for infinitely divisible non-Gaussian fields as well as the conditional distributions of such properties given the high excursion.

The distribution of $\int e^{f(t)}\,dt$ for the special case that $f(t)$ is a Wiener process has been studied by [27, 39]. For other general functionals of Gaussian processes and multivariate Gaussian random vectors, the tail approximation of the finite sum of correlated log-normal random variables has been studied by [8]. The corresponding simulation is studied in [15]. The gap between the finite sums of log-normal r.v.'s and the integral of continuous fields is substantial in the aspects of both generality and techniques.

The basic strategy of the analysis consists of two steps. The first step is to partition the domain T into n small squares of equal size denoted by A_i, $i = 1, \ldots, n$, and develop asymptotic approximations for each $p_i = P(\int_{A_i} e^{f(t)}\,dt > b)$. The size of A_i will be chosen carefully such that it is valid to use Taylor's expansion on $f(t)$ to develop the asymptotic approximations of p_i. The second step is to show that $P(\int_T e^{f(t)}\,dt > b) = (1 + o(1)) \sum_{i=1}^n p_i$. This implies that when computing $P(\int_T e^{f(t)}\,dt > b)$, we can pretend that all the $\int_{A_i} e^{f(t)}\,dt$'s are independent, though they are truly highly dependent. The sizes of the A_i's need to be chosen carefully. If A_i is too large, Taylor's expansion may not be accurate; if A_i is too small, the dependence of the fields in different A_i's will be high and the second step approximation may not be true. Since the first step of the analysis requires Taylor's expansion of the field, we will need to impose certain conditions on the field, which will be given in Section 2.

This paper is organized as follows. In Section 2 we provide necessary background and the technical conditions on the Gaussian random field in context. The main theorem and its connection to asymptotic approximation of $P(\sup_T f(t) > b)$ are presented in Section 3. In addition, two important steps of the proof are given in the same section, which lay out the proof strategy. Sections 4 and 5 give the proofs of the two steps presented in Section 3. Detailed lemmas and their proofs are given in the Appendix.

2. Some useful existing results.

2.1. *Preliminaries and technical conditions for Gaussian random field.* Consider a homogeneous Gaussian random field, $f(t)$, living on a domain T. Denote the covariance function by

$$C(t - s) = \text{Cov}(f(s), f(t)).$$

Throughout this paper, we assume that the random field satisfies the following conditions:

(C1) f is homogeneous with $Ef(t) = 0$ and $Ef^2(t) = 1$.
(C2) f is almost surely at least three times continuously differentiable with respect to t.

(C3) T is a d-dimension Jordan measurable compact subset of R^d.

(C4) The Hessian matrix of $C(t)$ at the origin is $-I$, where I is a $d \times d$ identity matrix.

Condition (C1) imposes unit variance. We will later study $\int_T e^{\sigma f(t)} \, dt$ and treat σ as an extra parameter. Condition (C2) implies that $C(t)$ is at least 6 times differentiable. In addition, the first, third and fifth derivatives of $C(t)$ evaluated at the origin are zero. For any $\tilde{f}(t)$ such that $\Delta \tilde{C}(0) = \Sigma$ and $|\Sigma| > 0$, (C4) can always be achieved by an affine transformation on the domain T by letting $\tilde{f}(t) = f(\Sigma^{1/2}t)$ and

$$\int_T e^{\sigma \tilde{f}(t)} \, dt = \int_T e^{\sigma f(\Sigma^{1/2}t)} \, dt = |\Sigma|^{-1/2} \int_{\{s \,:\, \Sigma^{-1/2}s \in T\}} e^{\sigma f(s)} \, ds,$$

where for a symmetric matrix Σ we let $\Sigma^{1/2}$ be a symmetric matrix such that $\Sigma^{1/2}\Sigma^{1/2} = \Sigma$.

For $\sigma > 0$, let

$$(1) \qquad\qquad\qquad I_\sigma(A) = \int_A e^{\sigma f(t)} \, dt$$

for the Jordan measurable set $A \subset T$. Of interest is

$$P\big(I_\sigma(T) > b\big)$$

as $b \to \infty$. Equivalently, we may consider that the variance of f is σ^2. However, it is notionally simpler to focus on a unit variance field and treat σ as a scale parameter.

We adopt the following notation. Let "∂" and "Δ" denote the gradient and Hessian matrix with respect to t, and "∂^2" denote the vector of second derivatives with respect to t. The difference between "Δ" and "∂^2" is that, for a specific t, $\Delta f(t)$ is a $d \times d$ symmetric matrix whose upper triangle entries are the elements of $\partial^2 f(t)$ which is a $(d(d+1)/2)$-dimensional vector. Let ∂_j denote the partial derivative with respect to the jth component of $t = (t_1, \ldots, t_d)$. We use similar notation for higher order derivatives. For b large enough, let u be the unique solution to

$$(2\pi/\sigma)^{d/2} u^{-d/2} e^{\sigma u} = b.$$

The uniqueness of u is immediate by noting that the left-hand side is monotone increasing with u for all $u > d/(2\sigma)$. In addition, we use the following notation and changes of variables:

$$\mu_1(t) = -(\partial_1 C(t), \ldots, \partial_d C(t)),$$

$$\mu_2(t) = \big(\partial_{ii}^2 C(t), i = 1, \ldots, d; \partial_{ij}^2 C(t), i = 1, \ldots, d, j = i+1, \ldots, d\big),$$

$$\mu_{02}^\top = \mu_{20} = \mu_2(0), \qquad f(0) = u - w, \qquad \partial f(0) = y,$$

$$\partial^2 f(0) = u\mu_{02} + z, \qquad \Delta f(0) = -uI + \mathbf{z}.$$

The vector μ_{20} contains the spectral moments of order two. Similar to $\Delta f(0)$ and $\partial^2 f(0)$, $\mathbf{z}$ is a symmetric matrix whose entries consist of elements in z. We create different notation because we will treat the second derivative of f as a matrix when doing Taylor's expansion and as a vector when doing integration. As stated in condition (C4), we have $\Delta C(0) = -I$. Equivalently, $\partial f(0)$ is a vector of independent unit variance Gaussian r.v.'s. We plan to show that in order to have $\int_T e^{f(t)} dt > b$, $\sup_T f(t)$ needs to reach a level around u. The distance between $f(0)$ and u is denoted by w. In addition, since $(f(0), \partial^2 f(0))$ is jointly independent of $\partial f(0)$, the distribution of $\partial f(0)$ is unaffected even if $f(0)$ reaches a high level. Further, the covariance between $f(0)$ and $\partial^2 f(0)$ is μ_{20}. Given $f(0) = u$, the conditional expectation of $\partial^2 f(0)$ is $u\mu_{02}$. The distance between $\partial^2 f(0)$ and this conditional expectation is denoted by vector z.

A well-known result (see, e.g., Chapter 5.5 in [5]) is that the joint distribution of $(f(0), \partial^2 f(0), \partial f(0), f(t))$ is multivariate normal with mean zero and variance

$$
\begin{pmatrix}
1 & \mu_{20} & 0 & C(t) \\
\mu_{02} & \mu_{22} & 0 & \mu_2^\top(t) \\
0 & 0 & I & \mu_1^\top(t) \\
C(t) & \mu_2(t) & \mu_1(t) & 1
\end{pmatrix},
$$

where $\mu_1(t)$, $\mu_2(t)$ and $\mu_{20} = \mu_{02}^\top$ is defined previously. The matrix μ_{22} is a $d(d+1)/2$ by $d(d+1)/2$ positive definite matrix and contains the 4th-order spectral moments arranged in an appropriate order. Conditional on $f(0) = u - w$, $\partial f(0) = y$ and $\Delta f(0) = -uI + \mathbf{z}$, $f(t)$ is a continuous Gaussian random field with conditional expectation

$$
(2) \qquad E(t) = (u - w, u\mu_{20} + z^\top, y^\top) \begin{pmatrix} \Gamma^{-1} & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} C(t) \\ \mu_2(t) \\ \mu_1(t) \end{pmatrix},
$$

where

$$
(3) \qquad \Gamma = \begin{pmatrix} 1 & \mu_{20} \\ \mu_{02} & \mu_{22} \end{pmatrix}.
$$

Note that $u\mu_{20} + z$ is the vector version of $-uI + \mathbf{z}$. Therefore, conditional on $(f(0), \partial f(0)^\top, \partial^2 f(0)^\top) = (u - w, y^\top, u\mu_{20} + z^\top)$, we have representation

$$
f(t) = E(t) + g(t),
$$

where $g(t)$ is a Gaussian random field with mean zero and

$$
E(t) = E(f(t)|f(0) = u - w, \partial f(0) = y, \partial^2 f(0) = u\mu_{02} + z),
$$

whose form is given in (2). Since $C(t)$ is six times differentiable, $E(t)$ is at least four times differentiable. Using the form of $E(t)$ in (2) and Γ in (3), after some

tedious calculations, we have

$$E(0) = u - w, \qquad \partial E(0) = y,$$

(4)
$$\Delta E(0) = -uI + \mathbf{z}, \qquad \partial_{ijk}^3 E(0) = y^\top \partial_{ijk}\mu_1(0),$$

$$\partial_{ijkl}^4 E(0) = (u - w, u\mu_{20} + z^\top)\Gamma^{-1}\begin{pmatrix} \partial_{ijkl}C(0) \\ \partial_{ijkl}\mu_2(0) \end{pmatrix}.$$

In order to obtain the above identities, we need the following facts. The first, third and fifth derivatives of $C(t)$ evaluated at 0 are all zero. The first and second derivatives of $C(t)$ are contained in $\mu_1(t)$ and $\mu_2(t)$. We also need to use the fact that

$$\Gamma^{-1} = \begin{pmatrix} \dfrac{1}{1 - \mu_{20}\mu_{22}^{-1}\mu_{02}} & -\dfrac{\mu_{20}\mu_{22}^{-1}}{1 - \mu_{20}\mu_{22}^{-1}\mu_{02}} \\ -\dfrac{\mu_{22}^{-1}\mu_{02}}{1 - \mu_{20}\mu_{22}^{-1}\mu_{02}} & \mu_{22}^{-1} + \dfrac{\mu_{22}^{-1}\mu_{02}\mu_{20}\mu_{22}^{-1}}{1 - \mu_{20}\mu_{22}^{-1}\mu_{02}} \end{pmatrix}.$$

With the derivatives of $E(t)$, we can write

(5) $$E(t) = u - w + y^\top t + \tfrac{1}{2}t^\top(-uI + \mathbf{z})t + g_3(t) + g_4(t) + R(t).$$

If we let $t = (t_1, \ldots, t_d)$, then

(6) $$g_3(t) = \frac{1}{6}\sum_{i,j,k}\partial_{ijk}^3 E(0)t_it_jt_k, \qquad g_4(t) = \frac{1}{24}\sum_{i,j,k,l}\partial_{ijkl}^4 E(0)t_it_jt_kt_l,$$

and $R(t)$ is the remainder term of the Taylor expansion. The Taylor expansion of $E(t)$ is the same as $f(t)$ for the first two terms because $g(t)$ is of order $O(|t|^3)$. It is not hard to check that $\mathrm{Var}(g(t)) \le c|t|^6$ for some $c > 0$ and $|t|$ small enough.

2.2. *Some related existing results.* For the comparison with the high excursion of $\sup_T f(t)$, we cite one result for homogeneous random fields, which has been proved in more general settings in many different ways. See, for instance, [5, 9, 34]. This result is also useful for the proof of Theorem 1. For comparison purpose, we only present the result for the random fields discussed in this paper.

PROPOSITION 1. *Suppose Gaussian random field f satisfies condition* (C1)–(C4). *There exists a constant G such that*

$$P\Big(\sup_{t\in T} f(t) > u\Big) = (1 + o(1))G\,\mathrm{mes}(T)u^d P(f(0) > u)$$

as $u \to \infty$.

We also present one existing result on the tail probability approximation of the sum of correlated log-normal random variables which provides intuitions on the analysis of $\int_T e^{f(t)}\,dt$.

PROPOSITION 2.   *Let $X = (X_1, \ldots, X_n)$ be a multivariate Gaussian random variable with mean μ and covariance matrix Σ, with $\det(\Sigma) > 0$. Then,*

$$(7) \qquad P\left(\sum_{i=1}^{n} e^{X_i} > b\right) = (1 + o(1)) \sum_{i=1}^{n} P(e^{X_i} > b)$$

as $b \to \infty$.

The proof of this proposition can be found in [8, 28]. This result implies that the large value of $\sum_{i=1}^{n} e^{X_i}$ is largely caused by one of the X_i's being large. In the case that X_i's are independent, Proposition 2 is a simple corollary of the subexponentiality of log-normal distribution. Though the X_i's are correlated, asymptotically they are tail-independent. The result presented in the next section can be viewed as a natural generalization of Proposition 2. Nevertheless, the techniques are quite different from the following aspects. First, Proposition 2 requires Σ to be nondegenerated. For the continuous random fields, this is usually not true. As shown in the analysis, we indeed need to study the sum of random variables whose correlation converges to 1 when b tends to infinity. Second, the approximation in (7) is for a sum of a fixed number of random variables. The analysis of the continuous field usually needs to handle the situation that the number of random variables in a sum grows to infinity as $b \to \infty$. Last but not least, to obtain approximations for $P(\int_T e^{f(t)} dt > b)$, one usually needs to first obtain approximations for $P(\int_{\Xi_\varepsilon} e^{f(t)} dt > b)$ for some small domain $\Xi_\varepsilon \subset T$. We will address all these issues in later sections.

For notation convenience, we write $a_u = O(b_u)$ if there exists a constant $c > 0$ independent of everything such that $a_u \leq c b_u$ for all $u > 1$, and $a_u = o(b_u)$ if $a_u/b_u \to 0$ as $u \to \infty$ and the convergence is uniform in other quantities. We write $a_u = \Theta(b_u)$ if $a_u = O(b_u)$ and $b_u = O(a_u)$. In addition, we write $X_u = o_p(1)$ if $X_u \xrightarrow{p} 0$ as $u \to \infty$.

3. Main result.   The main theorem of this paper is stated as follows.

THEOREM 1.   *Let f be a Gaussian random field living on $T \subset R^d$ satisfying (C1)–(C4). Given $\sigma > 0$, for b large enough, u is the unique solution to equation*

$$(8) \qquad \left(\frac{2\pi}{\sigma}\right)^{d/2} u^{-d/2} e^{\sigma u} = b.$$

Then,

$$P\left(\int_T e^{\sigma f(t)} dt > b\right) = (1 + o(1)) H \operatorname{mes}(T) u^{d-1} \exp(-u^2/2)$$

as $b \to \infty$, *where* mes(T) *is the Lebesgue measure of* T,

$$
H = \frac{|\Gamma|^{-1/2} \det(\mu_{22})^{1/2} e^{((1/8)\mathbf{1}^\top \mu_{22}\mathbf{1} + (1/8)\sum_i \partial^4_{iiii} C(0))/\sigma^2}}{(2\pi)^{(d+1)(d+2)/4}}
$$
(9)
$$
\times \int_{R^{d(d+1)/2}} \exp\left\{ -\frac{1}{2}\left[B^\top B + \frac{(\mu_{20}\mu_{22}^{-1/2}B + \mu_{20}\mathbf{1}/(2\sigma))^2}{1 - \mu_{20}\mu_{22}^{-1}\mu_{02}} \right] \right\} dB,
$$

Γ *is defined in* (3), $\mu_{20}, \mu_{02}, \mu_{22}$ *are defined in the previous section and*

$$
\mathbf{1} = (\underbrace{1, \ldots, 1}_{d}, \underbrace{0, \ldots, 0}_{d(d-1)/2})^\top.
$$

REMARK 1. The integral in (9) is clearly in an analytic form. We write it as an integral because it arises naturally from the derivation.

COROLLARY 1. *Let f be a Gaussian random field living on $T \subset R^d$ satisfying* (C1)–(C4). *Adopting all the notation in Theorem* 1, *let $\tilde{b} = b(2\pi/\sigma)^{-d/2}$ and*

(10)
$$
\tilde{u} = \frac{\log \tilde{b}}{\sigma} + \frac{d}{2\sigma}\log\left(\frac{\log \tilde{b}}{\sigma}\right) + \left(\frac{d}{2}\right)^2 \frac{\log((\log \tilde{b})/\sigma)}{\sigma \log \tilde{b}}.
$$

Then,

$$
P\big(I_\sigma(T) > b\big) = \big(1 + o(1)\big) H \, \mathrm{mes}(T)\tilde{u}^{d-1} \exp(-\tilde{u}^2/2).
$$

PROOF. The result is immediate by the Taylor expansion on the left-hand side of equation (8) and note that $u - \tilde{u} = o(u^{-1})$. $\square$

As we see, the asymptotic tail decaying rates of $\sup_T f(t)$ and $\int_T e^{\sigma f(t)}\,dt$ take a very similar form. More precisely,

$$
P\left(\int_T e^{\sigma f(t)}\,dt > b \right) = \Theta(1) P\left(\sup_T f(t) > u \right)
$$

with u and b connected via (8). This fact suggests the following intuition on the tail probability of $I_\sigma(T)$. First, the event $\{I_\sigma(T) > b\}$ has substantial overlap with event $\{\sup_T f(t) > u\}$. It has been shown by many studies mentioned before that given u sufficiently large $\{\sup_T f(t) > u\}$ is mostly caused by just a single $f(t^*)$ being large for some $t^* \in T$. Put these two facts together, $\{I_\sigma(T) > b\}$ is mostly caused by $\{f(t^*) > u\}$, for some $t^* \in T$ not too close to the boundary of T. Therefore, conditional on $\{I_\sigma(T) > b\}$, the distribution of $f(t)$ is very similar to the distribution conditional on $\{\sup_T f(t) > u\}$. Of course, these two conditional distributions are not completely identical. The difference will be discussed momentarily. Now we perform some informal calculation to illustrate the shape of $f(t)$

given $f(t^*) = u$. Thanks to homogeneity, it is sufficient to study $t^* = 0$. Given $f(0) = u$,

$$E\big(f(t) | f(0) = u\big) = u C(t).$$

Since $C(t)$ is 6 times differentiable, $\partial C(0) = 0$ and $\Delta C(0) = -I$, we obtain $E(f(t)|f(0) = u) \approx u - u \cdot t^\top t / 2$. For the exact Slepian model of the random field given that f achieves a local maximum at t^* of level u, see [6]. Note that for b large,

$$\int_T e^{\sigma u - (1/2)\sigma u t^\top t}\, dt > b$$

is approximately equivalent to

$$(2\pi/\sigma)^{d/2} u^{-d/2} e^{\sigma u} > b.$$

In Theorem 1, this is exactly how u is defined. As shown in Figure 1, the three curves are $\exp\{E(f(t)|f(t^*) = u)\}$ for different t^*'s. Given that $\{\sup_T f(t) > u\}$, these three curves are equally likely to occur.

Second, as mentioned before, the conditional distributions of $f(t)$ are different given $\{I_\sigma(T) > b\}$ or $\{\sup_T f(t) > u\}$. This is why the two constants in Theorem 1 (H) and Proposition 1 (G) are different. The difference is due to the fact that the symmetric difference between $\{\sup_T f(t) > u\}$ and $\{\int_T e^{f(t)}\, dt > b\}$ is substantial though their overlap is significant too. Consider the following situation that contributes to the difference. $\sup_T f(t)$ is slightly less than u [e.g., by a magnitude of $O(u^{-1})$]. For this case, $I_\sigma(T)$ still has a large chance to be greater than b. For this
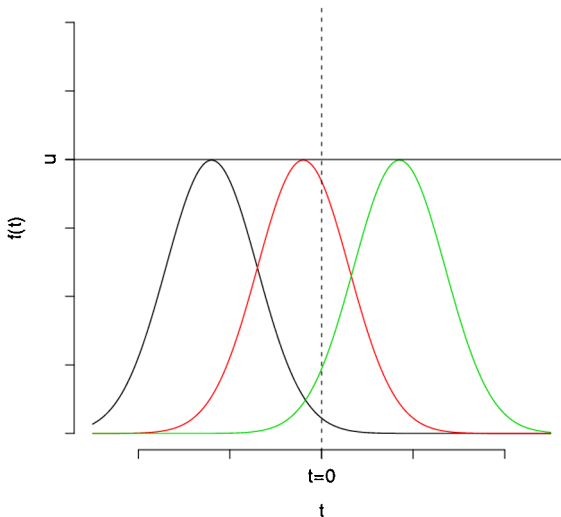

FIG. 1.   *One-dimensional example.*

sake we will need to consider the contribution of $\Delta f(0)$. As is shown in the technical proof, if $t^* = \arg\sup f(t) = 0$ and $\partial f(0) = 0$, then a sufficient and necessary condition for $I_\sigma(T) > b$ is that

$$f(0) + \frac{1}{2\sigma} \operatorname{Tr}(u^{-1}\Delta f(0) + I) > u + o(u^{-1}),$$

where Tr denotes the trace of a squared matrix. Note that conditional on $f(0) = u$, $E(\Delta f(0) | f(0) = u) = -uI$. Therefore, $\Delta f(0) + uI$ is of size $O(1)$. One well-known result is that the trace of a symmetric matrix is the sum of its eigenvalues. Let λ_i be the eigenvalues of $u^{-1}\Delta f(0) + I = \mathbf{z}/u$. Then, the sufficient and necessary condition is translated to $f(0) + \frac{1}{2\sigma}\sum_{i=1}^{d}\lambda_i > u$. This also suggests that, conditional on $I_\sigma(T) > b$, $w = f(0) - u$ is of size $O(u^{-1})$. This forms the intuition behind the proof of Theorem 2.

The proof of Theorem 1 consists of two steps presented in Sections 3.1 and 3.2, respectively. Each of the two steps is summarized as one theorem.

3.1. *Step* 1. Construct a cover of T, $\{A_1, \ldots, A_n\}$, such that $T \subset \bigcup_{i=1}^{n} A_i$. Each A_i is a closed square, $\operatorname{mes}(A_i \cap A_j) = 0$ for $i \neq j$. Because T is Jordan measurable, as $\sup_i \operatorname{mes}(A_i) \to 0$, $\operatorname{mes}(\bigcup_{i=1}^{n} A_i) - \operatorname{mes}(T) \to 0$. To simplify the analysis, we make each A_i of identical shape and let $A_i = \{t_i + s : s \in [0, a_b]^d\}$. The size of the partition n and choice of a_b depend on the threshold b. The first step analysis involves computing the integral $p_i \triangleq P(\int_{A_i} e^{\sigma f(t)}\, dt > b)$. Because f is homogeneous, it is sufficient to study p_1.

The basic strategy to approximate p_1 is as follows. Because f is at least three times differentiable, the first and second derivatives are almost surely well defined. Without loss of generality, we assume that $0 \in A_1$. Conditional on $(f(0), \partial f(0), \Delta f(0))$, $f(t) = f(0) + \partial f(0)^\top t + \frac{1}{2}t^\top \Delta f(0)t + g_3(t) + g_4(t) + R(t) + g(t)$, where $g(t)$ is a Gaussian field with mean zero and variance of order $O(|t|^6)$. Then,

$$
\begin{aligned}
p_1 &= P\left(\int_{A_1} e^{\sigma f(t)}\, dt > b\right) \\
&= \int h(w, y, z) \\
&\quad \times P\left(\int_{A_1} e^{\sigma[u - w + y^\top t + (1/2)t^\top(-uI + \mathbf{z})t + g_3(t) + g_4(t) + R(t) + g(t)]}\, dt \right. \\
&\qquad \left. > b\right) dw\, dy\, dz,
\end{aligned}
$$

(11)

where $h(w, y, z)$ is the density function of $(f(0), \partial f(0), \Delta f(0))$ evaluated at $(u - w, y, u\mu_{20} + z)$, which is a multivariate Gaussian random vector. Let u be defined in (8). There exists a $\delta > 0$ (small enough) such that if we let A_i's be squares of size $\varepsilon = O(1)u^{-1/2+\delta}$ and, hence, $n = O(1)\operatorname{mes}(T)u^{d/2-d\delta}$, the asymptotics

of p_i can be derived by repeatedly using Taylor's expansion and evaluating the integral on the right-hand side of (11). The main result of this step is presented as follows. It establishes a similar result to that of Theorem 1 but within a much smaller domain.

THEOREM 2. *Let f be a Gaussian random field living in T satisfying conditions* (C1)–(C4). *Let $A_1 = \Xi_\varepsilon = \{t : |t|_\infty < \varepsilon\}$, where $|t|_\infty = \max_i |t_i|$. Let u and H be defined in Theorem 1. Without loss of generality, assume $\Xi_\varepsilon \subset T$ with $\varepsilon = \kappa u^{\delta - 1/2}$ for some δ small enough and u large enough. Then, for any $\kappa > 0$*

$$p_1 = p(\Xi_\varepsilon) = P\left(\int_{\Xi_\varepsilon} e^{\sigma f(t)}\, dt > b\right) = (1 + o(1)) H \operatorname{mes}(\Xi_\varepsilon) u^{d-1} e^{-u^2/2}$$

as $b \to \infty$.

The proof of this theorem is in Section 4. We will then choose each A_i to be of the same shape as Ξ_ε. Then, all the p_i's are identical.

3.2. *Step* 2. The second step is to show that with the particular choice of A_i in the first step, $P(\int_T e^{\sigma f(t)}\, dt > b) = (1 + o(1)) \sum_{i=1}^n p_i$. We first present the main result of the second step.

THEOREM 3. *Let f be a Gaussian random field satisfying conditions* (C1)–(C4) *and ε be chosen in Theorem 2. Let $\mathbf{k} \in \mathbb{Z}^d$ and $\Xi_{\varepsilon,\mathbf{k}} = 2\mathbf{k}\varepsilon + \Xi_\varepsilon$. Further, let $\mathcal{C}^- = \{\mathbf{k} : \Xi_{\varepsilon,\mathbf{k}} \subset T\}$ and $\mathcal{C}^+ = \{\mathbf{k} : \Xi_{\varepsilon,\mathbf{k}} \cap T \neq \varnothing\}$, then*

$$P\left(I_\sigma\left(\bigcup_{\mathbf{k}\in\mathcal{C}^+} \Xi_{\varepsilon,\mathbf{k}}\right) > b\right) = (1 + o(1)) \sum_{\mathbf{k}\in\mathcal{C}^+} P\left(I_\sigma(\Xi_{\varepsilon,\mathbf{k}}) > b\right)$$

and

$$P\left(I_\sigma\left(\bigcup_{\mathbf{k}\in\mathcal{C}^-} \Xi_{\varepsilon,\mathbf{k}}\right) > b\right) = (1 + o(1)) \sum_{\mathbf{k}\in\mathcal{C}^-} P\left(I_\sigma(\Xi_{\varepsilon,\mathbf{k}}) > b\right).$$

We consider

$$I_\sigma\left(\bigcup_{\mathbf{k}\in\mathcal{C}^+} \Xi_{\varepsilon,\mathbf{k}}\right) = \sum_{\mathbf{k}\in\mathcal{C}^+} I_\sigma(\Xi_{\varepsilon,\mathbf{k}})$$

as a sum of finitely many dependently and identically distributed random variables. The conclusion of the above theorem implies that the tail distribution of the sum of these dependent variables exhibits the so-called "one big jump" feature—the high excursion of the sum is mainly caused by just one component being large. This result is similar to that of the sum of correlated log-normal r.v.'s. Nevertheless, the gap between the analyses of finite sum and integral is substantial because the

correlation between fields in adjacent squares tends to 1. For finite sums, the correlation is always bounded away from 1. The key step in the proof of Theorem 3 is that the ε defined in Theorem 2, though tends to zero as $b \to \infty$, is large enough such that the one-big-jump principle still applies. We will connect the event of high excursion of $I_\sigma (\bigcup_{\mathbf{k} \in \mathcal{C}^+} \Xi_{\varepsilon,\mathbf{k}})$ to the high excursion of $\sup_{\bigcup_{\mathbf{k} \in \mathcal{C}^+} \Xi_{\varepsilon,\mathbf{k}}} f(t)$ and apply existing results on the bound on the supremum of Gaussian random fields. A short list of recent related literature on the "one-big-jump" principle and multivariate Gaussian random variables is [8, 24, 28].

With the preparation of the two steps, we are ready to present the proof for Theorem 1.

PROOF OF THEOREM 1.   From Theorem 2,

$$\sum_{\mathbf{k} \in \mathcal{C}^+} P\big(I_\sigma (\Xi_{\varepsilon,\mathbf{k}}) > b\big) = (1 + o(1)) H \operatorname{mes}\bigg(\bigcup_{\mathbf{k} \in \mathcal{C}^+} \Xi_{\varepsilon,\mathbf{k}} \bigg) u^{d-1} e^{-u^2/2},$$

$$\sum_{\mathbf{k} \in \mathcal{C}^-} P\big(I_\sigma (\Xi_{\varepsilon,\mathbf{k}}) > b\big) = (1 + o(1)) H \operatorname{mes}\bigg(\bigcup_{\mathbf{k} \in \mathcal{C}^-} \Xi_{\varepsilon,\mathbf{k}} \bigg) u^{d-1} e^{-u^2/2}.$$

Therefore, thanks to Theorem 3,

$$P\big(I_\sigma (T) > b\big) \geq P\bigg(I_\sigma \bigg(\bigcup_{\mathbf{k} \in \mathcal{C}^-} \Xi_{\varepsilon,\mathbf{k}} \bigg) > b \bigg)$$

$$\geq (1 + o(1)) H \operatorname{mes}\bigg(\bigcup_{\mathbf{k} \in \mathcal{C}^-} \Xi_{\varepsilon,\mathbf{k}} \bigg) u^{d-1} e^{-u^2/2};$$

similarly,

$$P\big(I_\sigma (T) > b\big) \leq P\bigg(I_\sigma \bigg(\bigcup_{\mathbf{k} \in \mathcal{C}^+} \Xi_{\varepsilon,\mathbf{k}} \bigg) > b \bigg)$$

$$\leq (1 + o(1)) H \operatorname{mes}\bigg(\bigcup_{\mathbf{k} \in \mathcal{C}^+} \Xi_{\varepsilon,\mathbf{k}} \bigg) u^{d-1} e^{-u^2/2}.$$

Jordan measurability of T implies that

$$\operatorname{mes}\bigg(\bigcup_{\mathbf{k} \in \mathcal{C}^+} \Xi_{\varepsilon,\mathbf{k}} \bigg) - \operatorname{mes}\bigg(\bigcup_{\mathbf{k} \in \mathcal{C}^-} \Xi_{\varepsilon,\mathbf{k}} \bigg) \to 0+.$$

Therefore,

$$P\big(I_\sigma (T) > b\big) = (1 + o(1)) H \operatorname{mes}(T) u^{d-1} e^{-u^2/2}. \qquad \square$$

4. Proof for Theorem 2.   In this section we present the proof of Theorem 2. We arrange all the lemmas and their proofs in the Appendix.

PROOF OF THEOREM 2.   We evaluate the probability by conditioning on $(f(0), \partial f(0), \partial^2 f(0))$,

$$p(\Xi_\varepsilon) = P\left(\int_{\Xi_\varepsilon} e^{\sigma f(t)} \, dt > b\right)$$

$$= \int_{\mathcal{R}} h(w, y, z)$$

(12)
$$\times P\left(\int_{\Xi_\varepsilon} e^{\sigma f(t)} \, dt > b \,\Big|\, f(0) = u - w, \partial f(0) = y,\right.$$

$$\left. \partial^2 f(0) = u\mu_{02} + z\right) dw \, dy \, dz$$

$$= \int_{\mathcal{R}} h(w, y, z) P\left(\int_{\Xi_\varepsilon} e^{\sigma E(t) + \sigma g(t)} \, dt > b\right) dw \, dy \, dz,$$

where $\mathcal{R} = R^{(d+1)(d+2)/2}$ and $h(w, y, z)$ is the density function of $(f(0), \partial f(0), \partial^2 f(0))$ evaluated at $(u - w, y, u\mu_{02} + z)$. Now we take a closer look at the integrand inside the above integral. Conditional on $f(0) = u - w, \partial f(0) = y, \partial^2 f(0) = u\mu_{02} + z$,

$$I_\sigma(\Xi_\varepsilon) = \int_{|t|_\infty < \varepsilon} e^{\sigma E(t) + \sigma g(t)} \, dt$$

$$= \int_{|t|_\infty < \varepsilon} \exp\left\{\sigma\left[u - w + y^\top t + \frac{1}{2}t^\top(-uI + \mathbf{z})t\right.\right.$$

$$\left.\left. + g_3(t) + g_4(t) + R(t) + g(t)\right]\right\} dt$$

$$= \det(uI - \mathbf{z})^{-1/2}$$

$$\times \int_{|(uI - \mathbf{z})^{-1/2}t|_\infty < \varepsilon} e^{\sigma\{u - w + (1/2)y^\top(uI - \mathbf{z})^{-1}y\}}$$

$$\times \exp\left\{\sigma\left[-\frac{1}{2}(t - (uI - \mathbf{z})^{-1/2}y)^\top(t - (uI - \mathbf{z})^{-1/2}y)\right.\right.$$

$$+ g_3((uI - \mathbf{z})^{-1/2}t) + g_4((uI - \mathbf{z})^{-1/2}t)$$

$$\left.\left. + R((uI - \mathbf{z})^{-1/2}t) + g((uI - \mathbf{z})^{-1/2}t)\right]\right\} dt.$$

For the second equality, we plugged in (5). For the last step, we first change the variable from t to $(uI - \mathbf{z})^{1/2}t$ and then write the exponent in a quadratic form

of t. We write the term inside the exponent without the factor σ as

$$
\begin{aligned}
J(t) = {} & -\tfrac{1}{2}\bigl(t - (uI - \mathbf{z})^{-1/2}y\bigr)^{\top}\bigl(t - (uI - \mathbf{z})^{-1/2}y\bigr) + g_3\bigl((uI - \mathbf{z})^{-1/2}t\bigr) \\
& + g_4\bigl((uI - \mathbf{z})^{-1/2}t\bigr) + R\bigl((uI - \mathbf{z})^{-1/2}t\bigr),
\end{aligned}
$$
(13)

which is asymptotically a quadratic form. But, as is shown later, g_3 and g_4 terms do play a role in the calculation. Also, it is useful to keep in mind that $J(t)$ depends on y and z. Hence, we can write

$$
\begin{aligned}
I_\sigma(\Xi_\varepsilon) &= \int_{|t|_\infty < \varepsilon} e^{\sigma f(t)}\, dt \\
&= \det(uI - \mathbf{z})^{-1/2} e^{\sigma\{u - w + (1/2)y^{\top}(uI - \mathbf{z})^{-1}y\}} \\
&\quad \times \int_{|(uI - \mathbf{z})^{-1/2}t|_\infty < \varepsilon} e^{\sigma J(t) + \sigma g((uI - \mathbf{z})^{-1/2}t)}\, dt.
\end{aligned}
$$

Let

$$
e^{H_0} = \int_{R^d} e^{-(\sigma/2)t^{\top}t}\, dt = (2\pi/\sigma)^{d/2}.
$$
(14)

Let u solve

$$
u^{-d/2} e^{\sigma u + H_0} = b.
$$
(15)

Then,

$$
\int_{\Xi_\varepsilon} e^{\sigma f(t)}\, dt > b,
$$

if and only if

$$
\begin{aligned}
&\det(uI - \mathbf{z})^{-1/2} e^{\sigma\{u - w + (1/2)y^{\top}(uI - \mathbf{z})^{-1}y\}} \int_{|(uI - \mathbf{z})^{-1/2}t|_\infty < \varepsilon} e^{\sigma J(t) + \sigma g((uI - \mathbf{z})^{-1/2}t)}\, dt \\
&\quad > u^{-d/2} e^{\sigma u + H_0}.
\end{aligned}
$$

We take the logarithm on both sides and rewrite the above inequality and have

$$
\begin{aligned}
(16)\qquad 0 < {} & \frac{\sigma}{2} y^{\top}(uI - \mathbf{z})^{-1}y - \sigma w - \frac{1}{2}\log\det(I - u^{-1}\mathbf{z}) \\
& + \log \int_{|(uI - \mathbf{z})^{-1/2}t|_\infty < \varepsilon} e^{\sigma J(t)}\, dt - H_0 \\
& + \log E \exp\bigl(\sigma g((uI - \mathbf{z})^{-1/2}S)\bigr) \\
(17)\qquad = {} & A(w, y, z) + \log E \exp\bigl(\sigma g((uI - \mathbf{z})^{-1/2}S)\bigr),
\end{aligned}
$$

where S is a random variable on the region that $|(uI - \mathbf{z})^{-1/2}S|_\infty \leq \varepsilon$ with density proportional to $e^{\sigma J(s)}$ and

(18)
$$A(w, y, z) = \frac{\sigma}{2}y^\top (uI - \mathbf{z})^{-1}y - \sigma w - \frac{1}{2}\log\det(I - u^{-1}\mathbf{z})$$
$$+ \log\int_{|(uI-\mathbf{z})^{-1/2}t|<\varepsilon} e^{\sigma J(t)}\,dt - H_0.$$

Thanks to Lemma 1, we only need to consider the set that

(19)
$$\mathcal{L} = \{|f(0) - u| \leq u^{2\delta+\varepsilon_0},$$
$$|\partial f(0)| \leq u^{1/2+\delta+\varepsilon_0}, |\partial^2 f(0) - u\mu_{20}| \leq u^{1/2+\varepsilon_0}\}.$$

Also, by abusing notation, we write

(20)
$$\mathcal{L} = \{|w| \leq u^{2\delta+\varepsilon_0}, |y| \leq u^{1/2+\delta+\varepsilon_0}, |z| \leq u^{1/2+\varepsilon_0}\}.$$

Lemma 2 gives the form of $h(w, y, z)$. We plug in the results in Lemmas 1 and 2,

$$p(\Xi_\varepsilon) = \int_{\mathcal{R}} h(w, y, z)$$
$$\times P\big(A(w, y, z) + \log E\exp(\sigma g((uI - \mathbf{z})^{-1/2}S)) > 0\big)\,dw\,dy\,dz$$
$$= o(1)u^{-\alpha}e^{-u^2/2}$$
$$+ \int_{\mathcal{L}} h(w, y, z)$$
$$\times P\big(A(w, y, z) + \log E\exp(\sigma g((uI - \mathbf{z})^{-1/2}S)) > 0\big)\,dw\,dy\,dz$$
$$= o(1)u^{-\alpha}e^{-u^2/2}$$

(21)
$$+ \frac{1}{(2\pi)^{(d+1)(d+2)/4}}|\Gamma|^{-1/2}$$
$$\times \int_{\mathcal{L}} P\big(A(w, y, z) + \log E\exp(\sigma g((uI - \mathbf{z})^{-1/2}S)) > 0\big)$$
$$\times \exp\bigg\{-\bigg[\frac{1}{2}u^2 + \frac{u}{\sigma}A(w, y, z)$$
$$+ \frac{1}{2}y^\top\big(I - (I - \mathbf{z}/u)^{-1}\big)y$$
$$+ \frac{1}{2}\frac{(w + \mu_{20}\mu_{22}^{-1}z)^2}{1 - \mu_{20}\mu_{22}^{-1}\mu_{02}} + \frac{1}{2}z^\top\mu_{22}^{-1}z$$
$$+ \frac{u}{2\sigma}\log\det(I - u^{-1}\mathbf{z})$$
$$- \frac{u}{\sigma}\log\int_{|(uI-\mathbf{z})^{-1/2}t|<\varepsilon} e^{\sigma J(t)}\,dt + \frac{u}{\sigma}H_0\bigg]\bigg\}\,dw\,dy\,dz.$$

We define

$$\mathcal{I} = \frac{1}{2}u^2 + \frac{u}{\sigma}A(w, y, z) + \frac{1}{2}y^\top\big(I - (I - \mathbf{z}/u)^{-1}\big)y$$

(22)
$$+ \frac{1}{2}\frac{(w + \mu_{20}\mu_{22}^{-1}z)^2}{1 - \mu_{20}\mu_{22}^{-1}\mu_{02}} + \frac{1}{2}z^\top\mu_{22}^{-1}z + \frac{u}{2\sigma}\log\det(I - u^{-1}\mathbf{z})$$

$$- \frac{u}{\sigma}\log\int_{|(uI-\mathbf{z})^{-1/2}t|<\varepsilon} e^{\sigma J(t)}\,dt + \frac{u}{\sigma}H_0,$$

and proceed with some tedious algebra to write $\mathcal{I}$ in a friendly form for integration. First notice that

$$(I - u^{-1}\mathbf{z})^{-1} = \sum_{k=0}^{\infty} u^{-n}\mathbf{z}^n.$$

Plug this into the third term of $\mathcal{I}$ and obtain

$$\mathcal{I} = \frac{1}{2}u^2 + \frac{u}{\sigma}A(w, y, z) - (1 + O(|z|/u))\frac{u^{-1}}{2}y^\top\mathbf{z}y$$

$$+ \frac{1}{2}\frac{(w + \mu_{20}\mu_{22}^{-1}z)^2}{1 - \mu_{20}\mu_{22}^{-1}\mu_{02}} + \frac{1}{2}z^\top\mu_{22}^{-1}z + \frac{u}{2\sigma}\log\det(I - u^{-1}\mathbf{z})$$

$$- \frac{u}{\sigma}\log\int_{|(uI-\mathbf{z})^{-1/2}t|<\varepsilon} e^{\sigma J(t)}\,dt + \frac{u}{\sigma}H_0.$$

Situation 1 *of Lemma* 5. Adopt the notation in Lemmas 4 and 5. Note that according to the definition of Y in Lemma 4 that

$$Y = (y_i^2, i = 1, \ldots, d, 2y_i y_j, 1 \le i < j \le d)^\top,$$

$$\mathbf{1} = (\underbrace{1, \ldots, 1}_{d}, \underbrace{0, \ldots, 0}_{d(d-1)/2})^\top,$$

we obtain that

$$y^\top\mathbf{z}y = Y^\top z.$$

We plug in results of Lemmas 4 and 5. First, considering the first situation in Lemma 5, that is, $\mathcal{L}_1 = \mathcal{L} \cap \{|(uI - \mathbf{z})^{-1/2}y|_\infty \le \kappa u^\delta - u^{\delta/2}\}$, we have

$$\mathcal{I} = \frac{1}{2}u^2 + \frac{u}{\sigma}A(w, y, z) - (1 + O(|z|/u))\frac{u^{-1}}{2}Y^\top z$$

$$+ \frac{1}{2}\frac{(w + \mu_{20}\mu_{22}^{-1}z)^2}{1 - \mu_{20}\mu_{22}^{-1}\mu_{02}} + \frac{1}{2}z^\top\mu_{22}^{-1}z + \frac{u}{2\sigma}\log\det(I - u^{-1}\mathbf{z})$$

$$+ \frac{1}{8}(u^{-1}Y + \mathbf{1}/\sigma)^\top\mu_{22}(u^{-1}Y + \mathbf{1}/\sigma) - \frac{1}{8\sigma^2}\mathbf{1}^\top\mu_{22}\mathbf{1}$$

$$+ \frac{u}{\sigma} H_0 - \frac{u}{\sigma} H\big((uI - \mathbf{z})^{-1/2} y, (uI - \mathbf{z})^{1/2}, \varepsilon\big)$$

$$- \frac{1}{8\sigma^2} \sum_i \partial_{iiii} C(0) + o(1).$$

Also, it is useful to keep in mind that $\mathbf{1}$ is NOT a vector of 1's. The next step is to plug in the result of Lemma 3 and replace the $\log \det(I - u^{-1}\mathbf{z})$ term by

$$-u^{-1} \operatorname{Tr}(\mathbf{z}) + \tfrac{1}{2} u^{-2} \mathcal{E}^2(\mathbf{z}) + o(u^{-1})$$
$$= -u^{-1} \mathbf{1}^\top z + \tfrac{1}{2} u^{-2} \mathcal{E}^2(\mathbf{z}) + o(u^{-1})$$

and obtain

$$\mathcal{I} = \frac{1}{2} u^2 + \frac{u}{\sigma} A(w, y, z) - \big(1 + O(|z|/u)\big) \frac{u^{-1}}{2} Y^\top z$$

$$+ \frac{1}{2} \frac{(w + \mu_{20}\mu_{22}^{-1} z)^2}{1 - \mu_{20}\mu_{22}^{-1}\mu_{02}} + \frac{1}{2} z^\top \mu_{22}^{-1} z - \frac{1}{2\sigma} \mathbf{1}^\top z + \frac{1}{4\sigma u} \mathcal{E}^2(\mathbf{z})$$

$$+ \frac{1}{8}(u^{-1}Y + \mathbf{1}/\sigma)^\top \mu_{22}(u^{-1}Y + \mathbf{1}/\sigma) - \frac{1}{8\sigma^2} \mathbf{1}^\top \mu_{22}\mathbf{1}$$

$$+ \frac{u}{\sigma} H_0 - \frac{u}{\sigma} H\big((uI - \mathbf{z})^{-1/2} y, (uI - \mathbf{z})^{1/2}, \varepsilon\big)$$

$$- \frac{1}{8\sigma^2} \sum_i \partial_{iiii} C(0) + o(1).$$

Then, we group the terms $-\big(1 + O(|z|/u)\big)\frac{u^{-1}}{2} Y^\top z$ and $-\frac{1}{2\sigma}\mathbf{1}^\top z$ and leave the $O(|z|/u)$ to the end and have

$$\mathcal{I} = \frac{1}{2} u^2 + \frac{u}{\sigma} A(w, y, z) + \frac{1}{2} \frac{(w + \mu_{20}\mu_{22}^{-1} z)^2}{1 - \mu_{20}\mu_{22}^{-1}\mu_{02}}$$

$$+ \frac{1}{2} z^\top \mu_{22}^{-1} z - \frac{1}{2}(u^{-1}Y + \mathbf{1}/\sigma)^\top z$$

$$+ \frac{1}{8}(u^{-1}Y + \mathbf{1}/\sigma)^\top \mu_{22}(u^{-1}Y + \mathbf{1}/\sigma)$$

$$+ \frac{u}{\sigma} H_0 - \frac{u}{\sigma} H\big((uI - \mathbf{z})^{-1/2} y, (uI - \mathbf{z})^{1/2}, \varepsilon\big)$$

$$- \frac{1}{8\sigma^2} \mathbf{1}^\top \mu_{22}\mathbf{1} - \frac{1}{8\sigma^2} \sum_i \partial_{iiii} C(0) + o(1)$$

$$+ O(u^{-2}|z|^2|y|^2) + O(u^{-1}\mathcal{E}^2(\mathbf{z})).$$

Note that second and third lines in the above display is in fact in a quadratic form. We then have

$$
\mathcal{I} = \frac{1}{2}u^2 + \frac{u}{\sigma}A(w, y, z) + \frac{1}{2}\frac{(w + \mu_{20}\mu_{22}^{-1}z)^2}{1 - \mu_{20}\mu_{22}^{-1}\mu_{02}}
$$

$$
+ \frac{1}{2}\left[\mu_{22}^{-1/2}z - \frac{1}{2}\mu_{22}^{1/2}(u^{-1}Y + \mathbf{1}/\sigma)\right]^\top
$$

$$
\times \left[\mu_{22}^{-1/2}z - \frac{1}{2}\mu_{22}^{1/2}(u^{-1}Y + \mathbf{1}/\sigma)\right]
$$

$$
+ \frac{u}{\sigma}H_0 - \frac{u}{\sigma}H\left((uI - \mathbf{z})^{-1/2}y, (uI - \mathbf{z})^{1/2}, \varepsilon\right)
$$

$$
- \frac{1}{8\sigma^2}\mathbf{1}^\top\mu_{22}\mathbf{1} - \frac{1}{8\sigma^2}\sum_i \partial_{iiii}C(0) + o(1)
$$

$$
+ O(u^{-2}|z|^2|y|^2) + O(u^{-1}\mathcal{E}^2(\mathbf{z})).
$$

Now, consider another change of variable,

$$
\begin{aligned}
(23) \quad & A = A(w, y, z), \\
& B = \mu_{22}^{-1/2}z - \tfrac{1}{2}\mu_{22}^{1/2}(u^{-1}Y + \mathbf{1}/\sigma), \qquad y = y.
\end{aligned}
$$

Then, by noting that μ_{20} is a row vector in which the first d entries are -1's and the rest are 0's, we have

$$
w + \mu_{20}\mu_{22}^{-1}z = -\frac{A}{\sigma} + \mu_{20}\mu_{22}^{-1/2}B + \frac{1}{2\sigma}\mu_{20}\mathbf{1} + o(1).
$$

Therefore, we have

$$
\mathcal{I} = \frac{1}{2}u^2 + \frac{u}{\sigma}A + \frac{1}{2}B^\top B
$$

$$
+ \frac{1}{2}\frac{(-A/\sigma + \mu_{20}\mu_{22}^{-1/2}B + (1/(2\sigma))\mu_{20}\mathbf{1} + o(1))^2}{1 - \mu_{20}\mu_{22}^{-1}\mu_{02}}
$$

$$
+ \frac{u}{\sigma}H_0 - \frac{u}{\sigma}H\left((uI - \mathbf{z})^{-1/2}y, (uI - \mathbf{z})^{1/2}, \varepsilon\right)
$$

$$
- \frac{1}{8\sigma^2}\mathbf{1}^\top\mu_{22}\mathbf{1} - \frac{1}{8\sigma^2}\sum_i \partial_{iiii}C(0)
$$

$$
+ o(1) + O(u^{-2}|z|^2|y|^2)
$$

$$
+ O(u^{-1}\mathcal{E}^2(\mathbf{z})).
$$

We write $X_u = o_p(1)$ if $X_u \to 0$ in probability as $u \to \infty$. We insert the above form back to the integral in (21) and apply Lemma 7,

$$p(\Xi_\varepsilon) = o(1)u^{-\alpha}e^{-u^2/2}$$

$$+ \frac{1}{(2\pi)^{(d+1)(d+2)/4}}|\Gamma|^{-1/2}$$

$$\times \int_{\mathcal{L}} P(u \cdot A > o_p(1))$$

$$\times \exp\left\{-\left[\frac{1}{2}u^2 + \frac{u}{\sigma}A + \frac{1}{2}B^\top B\right.\right.$$

$$+ \frac{1}{2}\frac{(-A/\sigma + \mu_{20}\mu_{22}^{-1/2}B + 1/(2\sigma)\mu_{20}\mathbf{1} + o(1))^2}{1 - \mu_{20}\mu_{22}^{-1}\mu_{02}}$$

$$+ \frac{u}{\sigma}H_0 - \frac{u}{\sigma}H\big((uI - \mathbf{z})^{-1/2}y, (uI - \mathbf{z})^{1/2}, \varepsilon\big)$$

$$- \frac{1}{8\sigma^2}\mathbf{1}^\top\mu_{22}\mathbf{1} - \frac{1}{8\sigma^2}\sum_i \partial_{iiii}C(0) + o(1)$$

$$\left.\left. + O(u^{-2}|z|^2|y|^2) + O(u^{-1}\mathcal{E}^2(\mathbf{z}))\right]\right\} dw\,dy\,dz.$$

Note that Jacobian determinant is

$$\left|\det\left(\frac{\partial(w, z, y)}{\partial(A, B, y)}\right)\right| = \sigma^{-1}\det(\mu_{22})^{1/2}.$$

Note that when $|(uI - \mathbf{z})^{-1/2}y| \leq \kappa u^\delta - u^{\delta/2}$ (the first situation in Lemma 5),

$$uH_0 - uH\big((uI - \mathbf{z})^{-1/2}y, (uI - \mathbf{z})^{1/2}, \varepsilon\big) = o(1).$$

Then, with another change of variable, $A' = uA$, the integration on $\mathcal{L}_1$ is

$$\int_{\mathcal{L}_1} h(w, y, z)P\left(\int_{\Xi_\varepsilon} e^{\sigma E(t) + \sigma g(t)}\, dt > b\right) dw\,dy\,dz$$

$$= \frac{|\Gamma|^{-1/2}}{(2\pi)^{(d+1)(d+2)/4}}$$

$$\times \int_{\mathcal{L}_1} P(u \cdot A > o_p(1))$$

$$\times \exp\left\{-\left[\frac{1}{2}u^2 + \frac{u}{\sigma}A + \frac{1}{2}B^\top B\right.\right.$$

(24)
$$+ \frac{1}{2}\frac{(-A/\sigma + \mu_{20}\mu_{22}^{-1/2}B + 1/(2\sigma)\mu_{20}\mathbf{1} + o(1))^2}{1 - \mu_{20}\mu_{22}^{-1}\mu_{02}}$$

$$+ \frac{u}{\sigma} H_0 - \frac{u}{\sigma} H\big((uI - \mathbf{z})^{-1/2} y, (uI - \mathbf{z})^{1/2}, \varepsilon\big)$$

$$- \frac{1}{8\sigma^2} \mathbf{1}^\top \mu_{22} \mathbf{1} - \frac{1}{8\sigma^2} \sum_i \partial_{iiii} C(0) + o(1)$$

$$+ o(u^{-2}|z|^2|y|^2) + O(u^{-1}\mathcal{E}^2(\mathbf{z}))\bigg]\bigg\} \, dw \, dy \, dz$$

$$= \frac{\sigma^{-1}|\Gamma|^{-1/2}}{(2\pi)^{(d+1)(d+2)/4}} \det(\mu_{22})^{1/2}$$

$$\times e^{-(1/2)u^2 + (1/(8\sigma^2))\mathbf{1}^\top \mu_{22}\mathbf{1} + (1/(8\sigma^2))\sum_i \partial_{iiii} C(0)}$$

$$\times \int_{\mathcal{L}_1} P\big(u \cdot A > o_p(1)\big)$$

$$\times \exp\bigg\{-\bigg[\frac{u}{\sigma} \cdot A + \frac{B^\top B}{2}$$

$$+ \frac{(-A/\sigma + \mu_{20}\mu_{22}^{-1/2}B + 1/(2\sigma)\mu_{20}\mathbf{1} + o(1))^2}{2(1 - \mu_{20}\mu_{22}^{-1}\mu_{02})}$$

$$+ o(1) + O\big(u^{-2}|z|^2|y|^2 + u^{-1}\mathcal{E}^2(\mathbf{z})\big)\bigg]\bigg\} \, dA \, dB \, dy$$

$$= \frac{\sigma^{-1}|\Gamma|^{-1/2}}{(2\pi)^{(d+1)(d+2)/4}} \det(\mu_{22})^{1/2} u^{-1}$$

$$\times e^{-(1/2)u^2 + (1/(8\sigma^2))\mathbf{1}^\top \mu_{22}\mathbf{1} + (1/(8\sigma^2))\sum_i \partial_{iiii} C(0)}$$

$$\times \int_{\mathcal{L}_1} P\big(A' > o_p(1)\big)$$

$$\times \exp\bigg\{-\bigg[\frac{A'}{\sigma} + \frac{B^\top B}{2}$$

$$+ \frac{(-A'/(\sigma u) + \mu_{20}\mu_{22}^{-1/2}B + 1/(2\sigma)\mu_{20}\mathbf{1} + o(1))^2}{2(1 - \mu_{20}\mu_{22}^{-1}\mu_{02})}$$

$$+ o(1) + O\big(u^{-2}|z|^2|y|^2 + u^{-1}\mathcal{E}^2(\mathbf{z})\big)\bigg]\bigg\} \, dA' \, dB \, dy.$$

The second equality is a change of variable from (w, y, z) to (A, B, y). The third equality is a change of variable from (A, B, y) to (A', B, y). Note that $P(A' > o_p(1)) \to I(A' > 0)$ as $u \to \infty$. In addition, on the set $\mathcal{L}$,

$$O\big(u^{-2}|z|^2|y|^2 + u^{-1}\mathcal{E}^2(\mathbf{z})\big) = O(u^{-1+2\delta+2\varepsilon_0}|z|^2).$$

By choosing δ and ε_0 small enough, when $|B| < u^{1/4}$, $u^{-1+2\delta+2\varepsilon_0}|z|^2 = o(1)$; $|B| > u^{1/4}$, $|B| = \Theta(|z|)$, therefore,

$$\frac{B^\top B}{2} + O(u^{-1+2\delta+2\varepsilon_0}|z|^2) = (1 + o(1))\frac{B^\top B}{2}.$$

The integrant in (24) has the following bound, for $A' > 0$

$$P(A' > o_p(1))$$

$$\times \exp\left\{-\left[\frac{A'}{\sigma} + \frac{B^\top B}{2}\right.\right.$$

$$+ \frac{(-A'/(\sigma u) + \mu_{20}\mu_{22}^{-1/2}B + (1/(2\sigma))\mu_{20}\mathbf{1} + o(1))^2}{2(1 - \mu_{20}\mu_{22}^{-1}\mu_{02})}$$

$$\left.\left. + o(1) + O(u^{-2}|z|^2|y|^2 + u^{-1}\mathcal{E}^2(\mathbf{z}))\right]\right\}$$

$$\leq 2\exp\left\{-\frac{1}{\delta'}\left[\frac{A'}{\sigma} + \frac{B^\top B}{2} + \frac{(-A'/(\sigma u) + \mu_{20}\mu_{22}^{-1/2}B + (1/(2\sigma))\mu_{20}\mathbf{1})^2}{2(1 - \mu_{20}\mu_{22}^{-1}\mu_{02})}\right]\right\}$$

for δ' small enough. Note that the $o_p(1)$ is in fact $-u \log E \exp(g((uI - \mathbf{z})^{-1/2}S))$. Thanks to the result of Lemma 7, the integral of the left-hand side of the above display in the region $A' < 0$ is $o(1)$. By dominated convergence theorem, (24) equals

$$(1 + o(1))\frac{\sigma^{-1}|\Gamma|^{-1/2}}{(2\pi)^{(d+1)(d+2)/4}}$$

$$\times \det(\mu_{22})^{1/2}u^{-1}e^{-(1/2)u^2 + (1/(8\sigma^2))\mathbf{1}^\top\mu_{22}\mathbf{1} + (1/(8\sigma^2))\sum_i \partial_{iiii}C(0)}$$

$$(25) \quad \times \int_{A'>0, |y|_\infty < \kappa u^{\delta+1/2} - u^{\delta/2+1/2}} \exp\left\{-\left[A'/\sigma + \frac{B^\top B}{2}\right.\right.$$

$$\left.\left. + \frac{(\mu_{20}\mu_{22}^{-1/2}B + (1/(2\sigma))\mu_{20}\mathbf{1})^2}{2(1 - \mu_{20}\mu_{22}^{-1}\mu_{02})}\right]\right\} dA' \, dB \, dy$$

$$= (1 + o(1))H \operatorname{mes}(u\,\Xi_\varepsilon)u^{-1}e^{-(1/2)u^2},$$

where H is defined in (9). The above display is obtained by the fact that $\operatorname{mes}(u\,\Xi_\varepsilon) = u^d \operatorname{mes}(\Xi_\varepsilon) = \kappa^d u^{d/2+d\delta}$.

Situations 2 and 3 of Lemma 5. For the second situation in Lemma 5, let $\mathcal{L}_2 = \mathcal{L} \cap \{\kappa u^\delta - u^{\delta/2} < |(uI - \mathbf{z})^{-1/2}y|_\infty \leq (1 + \varepsilon_1)\kappa u^\delta\}$ and there exists $c_1 > 0$ such

that

$$\int_{\mathcal{L}_2} P\big(u \cdot A > o_p(1)\big)$$

$$\times \exp\Bigg\{-\Bigg[\frac{1}{2}u^2 + \frac{u}{\sigma}A + \frac{1}{2}B^\top B$$

$$+ \frac{1}{2}\frac{(-A/\sigma + \mu_{20}\mu_{22}^{-1/2}B + (1/(2\sigma))\mu_{20}\mathbf{1} + o(1))^2}{1 - \mu_{20}\mu_{22}^{-1}\mu_{02}}$$

(26)
$$+ \frac{u}{\sigma}H_0 - \frac{u}{\sigma}H\big((uI - \mathbf{z})^{-1/2}y, (uI - \mathbf{z})^{1/2}, \varepsilon\big)$$

$$- \frac{1}{8\sigma^2}\mathbf{1}^\top\mu_{22}\mathbf{1} - \frac{1}{8\sigma^2}\sum_i \partial_{iiii}C(0) + o(1)$$

$$+ o(u^{-2}|z|^2|y|^2) + O(u^{-1}\mathcal{E}^2(\mathbf{z}))\Bigg]\Bigg\}\,dw\,dy\,dz$$

$$\le \big(c_1\varepsilon_1^d + o(1)\big)H \operatorname{mes}(u\,\Xi_\varepsilon)u^{-1}e^{-(1/2)u^2}.$$

For the third situation, $\mathcal{L}_3 = \mathcal{L} \cap \{(1 + \varepsilon_1)\kappa u^\delta < |(uI - \mathbf{z})^{-1/2}y| \le u^{\delta+\varepsilon_0}\}$,

$$\int_{\mathcal{L}_3} P\big(u \cdot A > o_p(1)\big)$$

$$\times \exp\Bigg\{-\Bigg[\frac{1}{2}u^2 + \frac{u}{\sigma}A + \frac{1}{2}B^\top B$$

$$+ \frac{1}{2}\frac{(-A/\sigma + \mu_{20}\mu_{22}^{-1/2}B + (1/(2\sigma))\mu_{20}\mathbf{1} + o(1))^2}{1 - \mu_{20}\mu_{22}^{-1}\mu_{02}}$$

$$+ \frac{u}{\sigma}H_0 - \frac{u}{\sigma}H\big((uI - \mathbf{z})^{-1/2}y, (uI - \mathbf{z})^{1/2}, \varepsilon\big)$$

(27)
$$- \frac{1}{8\sigma^2}\mathbf{1}^\top\mu_{22}\mathbf{1} - \frac{1}{8\sigma^2}\sum_i \partial_{iiii}C(0) + o(1)$$

$$+ o(u^{-2}|z|^2|y|^2) + O(u^{-1}\mathcal{E}^2(\mathbf{z}))\Bigg]\Bigg\}\,dw\,dy\,dz$$

$$\le O(1)\Big(\frac{1}{2}\Big)^{u/\sigma}u^{-1}u^{(1/2+\delta+\varepsilon_0)d}e^{-(1/2)u^2}$$

$$= o(1)\operatorname{mes}(u\,\Xi_\varepsilon)u^{-1}e^{-(1/2)u^2}.$$

We put (25), (26) and (27) together and conclude the proof. $\quad\square$

5. Proof for Theorem 3. Similar to Section 4, we arrange all the lemmas and their proofs in the Appendix.

PROOF OF THEOREM 3. Since the proofs for $\mathcal{C}^+$ and $\mathcal{C}^-$ are complete analogue, we only provide the proof for $\mathcal{C}^+$. We prove for the asymptotics by providing bounds from both sides. We first discuss the easy case: the lower bound. Note that

$$P\left(I_\sigma\left(\bigcup_{\mathbf{k}\in\mathcal{C}^+}\Xi_{\varepsilon,\mathbf{k}}\right)>b\right)$$

$$\geq P\left(\max_{\mathbf{k}\in\mathcal{C}^+}I_\sigma(\Xi_{\varepsilon,\mathbf{k}})>b\right)$$

$$\geq \sum_{\mathbf{k}\in\mathcal{C}^+}P\left(I_\sigma(\Xi_{\varepsilon,\mathbf{k}})>b\right)-\sum_{\mathbf{k}\neq\mathbf{k}'}P\left(I_\sigma(\Xi_{\varepsilon,\mathbf{k}})>b,I_\sigma(\Xi_{\varepsilon,\mathbf{k}'})>b\right).$$

Thanks to Lemma 8,

$$P\left(I_\sigma\left(\bigcup_{\mathbf{k}\in\mathcal{C}^+}\Xi_{\varepsilon,\mathbf{k}}\right)>b\right)\geq(1+o(1))\sum_{\mathbf{k}\in\mathcal{C}^+}P\left(I_\sigma(\Xi_{\varepsilon,\mathbf{k}})>b\right).$$

The rest of the proof is to establish the asymptotic upper bound. To simplify our writing, we let

$$(28)\quad \mathcal{A}=I_\sigma(\Xi_\varepsilon),\qquad \mathcal{B}=I_\sigma\left(\sup_{\mathbf{k}\in\mathcal{N}}\Xi_{\varepsilon,\mathbf{k}}\right),\qquad \mathcal{D}=I_\sigma\left(\bigcup_{\mathbf{k}'\in\mathcal{C}^+\setminus[\{\mathbf{0}\}\cup\mathcal{N}]}\Xi_{\varepsilon,\mathbf{k}'}\right),$$

where $\mathcal{N}$ is the set of neighbors of Ξ_ε, that is, $\mathbf{k}\in\mathcal{N}$ if and only if

$$\inf_{s\in\Xi_\varepsilon,t\in\Xi_{\varepsilon,\mathbf{k}}}|s-t|=0.$$

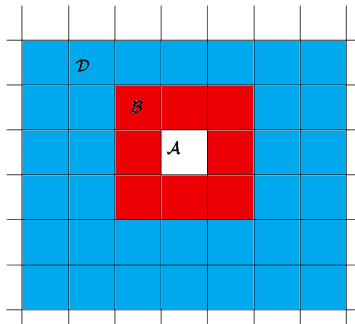An illustration of $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{D}$ is given in Figure 2.

FIG. 2. *Illustration of $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{D}$.*

Further, let

$$(29) \quad b_0 = u^{-1-d/2}b, \qquad b - b_0 = (1 - u^{-1-d/2})b = e^{-(1+o(1))u^{-1-d/2}}b,$$

and u_0 solves

$$u_0^{-d/2} e^{\sigma u_0 + H_0} = b_0,$$

and there exists $c_0 > 0$ such that $u_0 > u - c_0 \log u$. The first step in developing the upper bound is to use the following inequality

$$
\begin{aligned}
P(\mathcal{A} + \mathcal{B} + \mathcal{D} > b) & \\
& \leq P(\mathcal{A} > b - b_0) + P(\mathcal{A} \leq b_0, \mathcal{A} + \mathcal{B} + \mathcal{D} > b) \\
(30) \qquad & \quad + P(b_0 < \mathcal{A} \leq b - b_0, \mathcal{A} + \mathcal{B} + \mathcal{D} > b) \\
& \leq P(\mathcal{A} > b - b_0) + P(\mathcal{B} + \mathcal{D} > b - b_0) \\
& \quad + P(\mathcal{A} > b_0, \mathcal{B} + \mathcal{D} > b_0, \mathcal{A} + \mathcal{B} + \mathcal{D} > b).
\end{aligned}
$$

From Theorem 2,

$$P(\mathcal{A} > b - b_0) = (1 + o(1))P(\mathcal{A} > b).$$

The next step is to show that the last term in (30) is ignorable. Note that

$$
\begin{aligned}
P(\mathcal{A} > b_0, & \mathcal{B} + \mathcal{D} > b_0, \mathcal{A} + \mathcal{B} + \mathcal{D} > b) \\
& = P(\mathcal{A} + \mathcal{B} > b - b_0, \mathcal{A} > b_0, \mathcal{B} + \mathcal{D} > b_0, \mathcal{A} + \mathcal{B} + \mathcal{D} > b) \\
& \quad + P(\mathcal{D} > b - b_0, \mathcal{A} > b_0, \mathcal{B} + \mathcal{D} > b_0, \mathcal{A} + \mathcal{B} + \mathcal{D} > b) \\
& \quad + P(\mathcal{A} + \mathcal{B} > b_0, \mathcal{D} > b_0, \mathcal{A} > b_0, \mathcal{B} + \mathcal{D} > b_0, \mathcal{A} + \mathcal{B} + \mathcal{D} > b) \\
& \leq P(\mathcal{A} + \mathcal{B} > b - b_0, \mathcal{A} > b_0, \mathcal{B} + \mathcal{D} > b_0, \mathcal{A} + \mathcal{B} + \mathcal{D} > b) \\
& \quad + 2P(\mathcal{D} > b_0, \mathcal{A} > b_0) \\
& = o(P(\mathcal{A} > b)).
\end{aligned}
$$

The last step is due to Lemmas 9 and 10. By noting that $\#(\mathcal{C}^+)u^{-1-d/2} = o(u^{-1})$, the conclusion of the theorem is immediate by induction, where $\#(\cdot)$ is the cardinality of a set. $\square$

APPENDIX: LEMMAS IN SECTIONS 4 AND 5

Lemma 1 isolated the dominating event so that we will be in good shape to use Taylor's expansion.

LEMMA 1.   *There exist $\varepsilon_0, \delta > 0$ small enough and κ large. Let $\varepsilon = \kappa u^{-1/2+\delta}$ such that for any $\alpha > 0$,*

$$P\left(|f(0) - u| > u^{2\delta+\varepsilon_0} \text{ or } |\partial f(0)| > u^{1/2+\delta+\varepsilon_0} \text{ or}\right.$$

$$\left.|\partial^2 f(0) - u\mu_{20}| > u^{1/2+\varepsilon_0}, \int \Xi_\varepsilon e^{f(t)} dt > b\right)$$

$$= o(1) u^{-\alpha} e^{-(1/2)u^2}.$$

PROOF.   Note that there exists c_1 such that $\sigma u \leq \log b + c_1 \log\log b$. Let $\sigma \tilde{u} = \log(b)$. Since we only consider the case that u is large, we always have $\text{mes}(\Xi_\varepsilon) < 1$:

$$P\left(f(0) < u - u^{2\delta+\varepsilon_0}, \int_{\Xi_\varepsilon} e^{f(t)} dt > b\right)$$

$$\leq P\left(f(0) < u - u^{2\delta+\varepsilon_0}, \sup f(t) > \tilde{u}\right)$$

$$\leq C P\left(f(0) < u - u^{2\delta+\varepsilon_0} | \sup f(t) > \tilde{u}\right) \tilde{u}^{d-1} e^{-\tilde{u}^2/2}.$$

The last inequality is an application of Proposition 1. Because for any $u' > \tilde{u}$, for some $\varepsilon_1 > 0$,

$$\inf_{t \in \Xi_\varepsilon} E\left(f(t) | \sup_{\Xi_\varepsilon} f(t) = u'\right)$$

$$\geq u' \inf_{t \in \Xi_\varepsilon} C(t) \geq u - \kappa^2 \varepsilon_1 u^{2\delta}(1 + o(1))$$

and

$$\sup_{t \in (-\varepsilon, \varepsilon)} \text{Var}\left(f(t) | \sup_{\Xi_\varepsilon} f(t) = u'\right) = O(\varepsilon^2) = O(u^{-1+2\delta}),$$

one can choose κ large enough such that

$$P\left(f(0) < u - u^{2\delta+\varepsilon_0} | \sup_{\Xi_\varepsilon} f(t) > \tilde{u}\right) = O(1) \exp(-u^{1+\varepsilon_0}/c_2).$$

Therefore,

$$P\left(f(0) < u - u^{2\delta+\varepsilon_0}, \int_{\Xi_\varepsilon} e^{f(t)} dt > b\right) = o(1) u^{-\alpha} e^{-u^2/2}$$

for all $\alpha > 0$. Also, $P(f(0) > u + u^{2\delta+\varepsilon_0}) = o(1) u^{-\alpha} e^{-(1/2)u^2}$. Hence,

$$P\left(|f(0) - u| > u^{2\delta+\varepsilon_0}, \int_{\Xi_\varepsilon} e^{f(t)} dt > b\right) = o(1) u^{-\alpha} e^{-(1/2)u^2}.$$

Similarly, we have

$$P\left(|f(0) - u| < u^{2\delta+\varepsilon_0}, |\partial f(0)| > u^{1/2+\delta+\varepsilon_0}, \int_{\Xi_\varepsilon} e^{f(t)}\, dt > b\right)$$

$$= o(1)u^{-\alpha}e^{-(1/2)u^2},$$

$$P\left(|f(0) - u| < u^{2\delta+\varepsilon_0}, |\partial^2 f(0) - u\mu_{20}| > u^{1/2+\varepsilon_0}, \int_{\Xi_\varepsilon} e^{f(t)}\, dt > b\right)$$

$$= o(1)u^{-\alpha}e^{-(1/2)u^2}.$$

The above two displays are immediate by noting that $(f(0), \partial f(0), \partial^2 f(0))$ is a multivariate Gaussian random vector. In addition, $(f(0), \partial^2 f(0))$ is independent of $\partial f(0)$ and the covariance between $f(0)$ and $\partial^2 f(0)$ is μ_{02}. $\quad\square$

LEMMA 2. *Let $h(w, y, z)$ be the density of $(f(0), \partial f(0), \partial^2 f(0))$. Then,*

$$h(w, y, z)$$

$$= \frac{1}{(2\pi)^{(d+1)(d+2)/4}}|\Gamma|^{-1/2}$$

$$\times \exp\left\{-\left[\frac{1}{2}u^2 + \frac{u}{\sigma}A(w, y, z) + \frac{1}{2}y^\top(I - (I - \mathbf{z}/u)^{-1})y\right.\right.$$

$$+ \frac{1}{2}\frac{(w + \mu_{20}\mu_{22}^{-1}z)^2}{1 - \mu_{20}\mu_{22}^{-1}\mu_{02}} + \frac{1}{2}z^\top\mu_{22}^{-1}z + \frac{u}{2\sigma}\log\det(I - u^{-1}\mathbf{z})$$

$$\left.\left.- \frac{u}{\sigma}\log\int_{|(uI-\mathbf{z})^{-1/2}t|<\varepsilon} e^{\sigma J(t)}\, dt + uH_0/\sigma\right]\right\}.$$

In addition,

$$(31) \qquad \Gamma^{-1} = \begin{pmatrix} \dfrac{1}{1 - \mu_{20}\mu_{22}^{-1}\mu_{02}} & -\dfrac{\mu_{20}\mu_{22}^{-1}}{1 - \mu_{20}\mu_{22}^{-1}\mu_{02}} \\ -\dfrac{\mu_{22}^{-1}\mu_{02}}{1 - \mu_{20}\mu_{22}^{-1}\mu_{02}} & \mu_{22}^{-1} + \dfrac{\mu_{22}^{-1}\mu_{02}\mu_{20}\mu_{22}^{-1}}{1 - \mu_{20}\mu_{22}^{-1}\mu_{02}} \end{pmatrix}.$$

PROOF. The form of Γ^{-1} in (31) is a result from linear algebra. The form of Γ^{-1} is direct application of the block matrix inverse from linear algebra. Note that

$$h(w, y, z) = \frac{1}{(2\pi)^{(d+1)(d+2)/4}}|\Gamma|^{-1/2}$$

$$\times \exp\left\{-\frac{1}{2}(u - w, z^\top + u\mu_{20}, y^\top)\begin{pmatrix}\Gamma^{-1} & 0 \\ 0 & I\end{pmatrix}\begin{pmatrix}u - w \\ z + u\mu_{02} \\ y\end{pmatrix}\right\}.$$

By plugging in the form of Γ^{-1}, we have

$$(u - w, z^\top + u\mu_{20}, y^\top)\begin{pmatrix} \Gamma^{-1} & 0 \\ 0 & I \end{pmatrix}\begin{pmatrix} u - w \\ z + u\mu_{02} \\ y \end{pmatrix}$$

$$= y^\top y + \frac{(u - w)^2}{1 - \mu_{20}\mu_{22}^{-1}\mu_{02}}$$

$$+ (z + u\mu_{02})^\top\left[\mu_{22}^{-1} + \frac{\mu_{22}^{-1}\mu_{02}\mu_{20}\mu_{22}^{-1}}{1 - \mu_{20}\mu_{22}^{-1}\mu_{02}}\right](z + u\mu_{02})$$

$$- 2(u - w)\frac{\mu_{20}\mu_{22}^{-1}}{1 - \mu_{20}\mu_{22}^{-1}\mu_{02}}(z + u\mu_{02})$$

$$= u^2 + y^\top y - 2wu + \frac{w^2}{1 - \mu_{20}\mu_{22}^{-1}\mu_{02}}$$

$$+ z^\top\left[\mu_{22}^{-1} + \frac{\mu_{22}^{-1}\mu_{02}\mu_{20}\mu_{22}^{-1}}{1 - \mu_{20}\mu_{22}^{-1}\mu_{02}}\right]z + 2w\frac{\mu_{20}\mu_{22}^{-1}}{1 - \mu_{20}\mu_{22}^{-1}\mu_{02}}z$$

$$= u^2 + \frac{2u}{\sigma}A(w, y, z) + y^\top(I - (uI - \mathbf{z})^{-1})y + \frac{(w + \mu_{20}\mu_{22}^{-1}z)^2}{1 - \mu_{20}\mu_{22}^{-1}\mu_{02}}$$

$$+ z^\top\mu_{22}^{-1}z + \frac{u}{\sigma}\log\det(I - u^{-1}\mathbf{z})$$

$$- \frac{2u}{\sigma}\log\int_{|(uI-\mathbf{z})^{-1/2}t| < \varepsilon} e^{J(t)}\,dt + \frac{2u}{\sigma}H_0.$$

Therefore, we conclude the proof. $\square$

LEMMA 3.

$$\log(\det(I - u^{-1}\mathbf{z})) = -u^{-1}\operatorname{Tr}(\mathbf{z}) + \tfrac{1}{2}u^{-2}\mathcal{E}^2(\mathbf{z}) + o(u^{-2}),$$

where Tr is the trace of a matrix, $\mathcal{E}^2(\mathbf{z}) = \sum_{i=1}^d \lambda_i^2$, and λ_i's are the eigenvalues of $\mathbf{z}$.

PROOF.   The result is immediate by noting that

$$\det(I - u^{-1}\mathbf{z}) = \prod_{i=1}^d (1 - \lambda_i/u)$$

and $\operatorname{Tr}(\mathbf{z}) = \sum_{i=1}^d \lambda_i$. $\square$

LEMMA 4.   *Let $y = (y_1, \ldots, y_d)^\top$ and $X \sim N(y/\sqrt{u}, I/\sqrt{\sigma})$. Then, on the set $\mathcal{L}$ defined in* (20),

$$
\begin{aligned}
E\big(g_3(X/\sqrt{u}) &+ g_4(X/\sqrt{u})\big) \\
&= -\frac{u^{-1}}{8}(u^{-1}Y + \mathbf{1}/\sigma)^\top \mu_{22}(u^{-1}Y + \mathbf{1}/\sigma) \\
&\quad + \frac{u^{-1}}{8\sigma^2}\mathbf{1}^\top \mu_{22}\mathbf{1} + \frac{u^{-1}}{8\sigma^2}\sum_i \partial_{iiii}C(0) + o(u^{-1}),
\end{aligned}
$$

where

$$
\begin{aligned}
Y &= (y_i^2, i = 1, \ldots, d, 2y_i y_j, 1 \le i < j \le d)^\top, \\
\mathbf{1} &= (\underbrace{1, \ldots, 1}_{d}, \underbrace{0, \ldots, 0}_{d(d-1)/2})^\top.
\end{aligned}
$$

PROOF.   Using the derivatives in (4), we have that

$$
\partial_{ijk}E(0) = -\sum_{l=1}^{d} \partial_{ijkl}^4 C(0) y_l,
$$

$$
\partial_{ijk}^4 E(0) = (u + O(|z| + |w|))\partial_{ijkl}C(0).
$$

We plug this into the definition of g_3 and g_4 in (6) and obtain, on the set $\mathcal{L}$,

$$
\begin{aligned}
E\big(g_3(X/\sqrt{u}) &+ g_4(X/\sqrt{u})\big) \\
&= -\frac{1}{6}u^{-3/2}\sum_{ijkl} \partial_{ijkl}^4 C(0) E(X_i X_j X_k y_l) \\
&\quad + \frac{u^{-1}}{24}\sum_{ijkl} \partial_{ijkl}^4 C(0) E(X_i X_j X_k X_l) + o(u^{-1}) \\
&= -\frac{1}{8}u^{-3/2}\sum_{ijkl} \partial_{ijkl}^4 C(0) E(X_i X_j X_k y_l) \\
&\quad + \frac{u^{-1}}{24}\sum_{ijkl} \partial_{ijkl}^4 C(0) E\big(X_i X_j X_k(X_l - y_l/\sqrt{u})\big) + o(u^{-1}) \\
&= -\frac{1}{8}u^{-3}\sum_{ijkl} \partial_{ijkl}^4 C(0) y_i y_j y_k y_l - \frac{3}{8\sigma}u^{-2}\sum_{il} y_i y_l \sum_j \partial_{iljj}^4 C(0) \\
&\quad + \frac{u^{-2}}{8\sigma}\sum_{ij} y_i y_j \sum_l \partial_{ijll}^4 C(0) + \frac{3u^{-1}}{24\sigma^2}\sum_i \partial_{iiii}^4 C(0) + o(u^{-1})
\end{aligned}
$$

$$= -\frac{1}{8}u^{-3}\sum_{ijkl}\partial^4_{ijkl}C(0)y_iy_jy_ky_l - \frac{1}{4\sigma}u^{-2}\sum_{ij}y_iy_j\sum_l\partial^4_{ijll}C(0)$$

$$+ \frac{u^{-1}}{8\sigma^2}\sum_i\partial^4_{iiii}C(0) + o(u^{-1}).$$

This last step is true because $\sum_{il}y_iy_l\sum_j\partial^4_{iljj}C(0) = \sum_{ij}y_iy_j\sum_l\partial^4_{ijll}C(0)$, which is just a change of index. Then, with the definition of Y and $\mathbf{1}$ in the statement of this lemma (note that $\mathbf{1}$ is NOT a vector of 1's), we have

$$E\big(g_3(X/\sqrt{u}) + g_4(X/\sqrt{u})\big)$$

$$= -\frac{u^{-3}}{8}Y^\top\mu_{22}Y - \frac{u^{-2}}{4\sigma}Y^\top\mu_{22}\mathbf{1} + \frac{u^{-1}}{8\sigma^2}\sum_i\partial^4_{iiii}C(0) + o(1)$$

$$= -\frac{u^{-1}}{8}(u^{-1}Y + \mathbf{1}/\sigma)^\top\mu_{22}(u^{-1}Y + \mathbf{1}/\sigma) + \frac{u^{-1}}{8\sigma^2}\mathbf{1}^\top\mu_{22}\mathbf{1}$$

$$+ \frac{u^{-1}}{8\sigma^2}\sum_i\partial^4_{iiii}C(0) + o(1). \qquad \square$$

LEMMA 5. *Let $J(t)$ be defined in* (13). *Then, on the set $\mathcal{L}$ the approximations of $\int_{|(uI-\mathbf{z})^{-1/2}t|_\infty<\varepsilon}e^{\sigma J(t)}\,dt$ under different situations are as follows:*

(1) *When $|(uI-\mathbf{z})^{-1/2}y|_\infty \le \kappa u^\delta - u^{\delta/2}$,*

$$\int_{|(uI-\mathbf{z})^{-1/2}t|_\infty<\varepsilon}e^{\sigma J(t)}\,dt$$

$$= \exp\big[\sigma E\big(g_3(X/\sqrt{u}) + g_4(X/\sqrt{u})\big)$$

$$+ H\big((uI-\mathbf{z})^{-1/2}y, (uI-\mathbf{z})^{1/2}, \varepsilon\big) + o(u^{-1})\big],$$

where

$$e^{H(y,\Sigma,\varepsilon)} = \int_{|\Sigma^{-1}t|<\varepsilon}e^{-(\sigma/2)(t-y)^\top(t-y)}\,dt,$$

and X is the random vector defined in Lemma 4. *In addition,*

$$H\big(0, (uI-\mathbf{z})^{1/2}, \varepsilon\big) - H_0 = o(u^{-1}).$$

(2) *For any $\varepsilon_1 > 0$, when $\kappa u^\delta - u^{\delta/2} \le |(uI-\mathbf{z})^{-1/2}y|_\infty \le (1+\varepsilon_1)\kappa u^\delta$,*

$$\int_{|(uI-\mathbf{z})^{-1/2}t|<\varepsilon}e^{\sigma J(t)}\,dt$$

$$\le \exp\big[\sigma E\big(g_3(X/\sqrt{u}) + g_4(X/\sqrt{u})\big)$$

$$+ H\big((uI-\mathbf{z})^{-1/2}y, (uI-\mathbf{z})^{1/2}, \varepsilon\big) + o(u^{-1})\big].$$

(3) *When* $(1 + \varepsilon_1)\kappa u^\delta < |(uI - \mathbf{z})^{-1/2}y|_\infty \le u^{\delta + \varepsilon_0}$

$$\int_{|(uI-\mathbf{z})^{-1/2}t|<\varepsilon} e^{\sigma J(t)}\, dt$$

$$\le \frac{1}{2}\exp[\sigma E(g_3(X/\sqrt{u}) + g_4(X/\sqrt{u})) + H_0 + o(u^{-1})].$$

PROOF.   Note that

$$\int_{|(uI-\mathbf{z})^{-1/2}t|<\varepsilon} e^{\sigma J(t)}\, dt$$

$$= e^{H((uI-\mathbf{z})^{-1/2}y,(uI-\mathbf{z})^{-1/2},\varepsilon)}$$

$$\times E\{\exp[\sigma g_3((uI - \mathbf{z})^{-1/2}X')$$

$$+ \sigma g_4((uI - \mathbf{z})^{-1/2}X') + \sigma R((uI - \mathbf{z})^{-1/2}X')];$$

$$|(uI - \mathbf{z})^{-1/2}X'| < \varepsilon\}.$$

Also, $(uI - \mathbf{z})^{-1/2}X' = (1 + O(z/u))X'/\sqrt{n}$ and

$$X' = X - y/\sqrt{u} + (uI - \mathbf{z})^{-1/2}y = X + O(u^{-3/2})|\mathbf{z}y|.$$

For the first situation, $|(uI - \mathbf{z})^{-1/2}y|_\infty \le \kappa u^\delta - u^{\delta/2}$ and $|z| < u^{1/2 + \varepsilon_0}$,

$$E\{\exp[\sigma g_3((uI - \mathbf{z})^{-1/2}X') + \sigma g_4((uI - \mathbf{z})^{-1/2}X') + \sigma R((uI - \mathbf{z})^{-1/2}X')];$$

$$|(uI - \mathbf{z})^{-1/2}X'| < \varepsilon\}$$

$$= E\{\exp[\sigma g_3(X'/\sqrt{u}) + \sigma g_4(X'/\sqrt{u}) + \sigma R(X'/\sqrt{u}) + o(u^{-1})];$$

$$|(uI - \mathbf{z})^{-1/2}X'| < \varepsilon\}$$

$$= E\{\exp[\sigma g_3(X'/\sqrt{u}) + \sigma g_4(X'/\sqrt{u}) + \sigma R(X'/\sqrt{u}) + o(u^{-1})]\}$$

$$= E\{\exp[\sigma g_3(X/\sqrt{u}) + \sigma g_4(X/\sqrt{u}) + \sigma R(X/\sqrt{u}) + o(u^{-1})]\}.$$

In addition, because $|R(t)| = O(u^{1/2+\delta}t^5)$, then

$$E\{\exp[\sigma g_3(X/\sqrt{u}) + \sigma g_4(X/\sqrt{u}) + \sigma R(X/\sqrt{u})]\}$$

$$= E\{\exp[\sigma g_3(X/\sqrt{u}) + \sigma g_4(X/\sqrt{u}) + o(u^{-1})]\}.$$

Further, $g_3(X/\sqrt{u}) = o(u^{-1/2+\delta})$ and $g_4(X/\sqrt{u}) = o(u^{-1/2+\delta})$, then by repeatedly using Talyor's expansion, we have

$$E\{\exp[\sigma g_3(X'/\sqrt{u}) + \sigma g_4(X'/\sqrt{u}) + \sigma R(X'/\sqrt{u})]; |(uI - \mathbf{z})^{-1/2}X'| < \varepsilon\}$$

$$= \exp\{E[\sigma g_3(X/\sqrt{u}) + \sigma g_4(X/\sqrt{u}) + o(u^{-1})]\}.$$

For the second situation, the inequality is immediate by noting that

$$E\{\exp[\sigma g_3(X'/\sqrt{u}) + \sigma g_4(X'/\sqrt{u}) + \sigma R(X'/\sqrt{u})]; |(uI - \mathbf{z})^{-1/2}X'| < \varepsilon\}$$

$$\le E\{\exp[\sigma g_3(X'/\sqrt{u}) + \sigma g_4(X'/\sqrt{u}) + \sigma R(X'/\sqrt{u})]\}.$$

For the third situation, note that the integral is not focusing on the dominating part, and the conclusion follows immediately.   □

The next lemma is known as the Borel-TIS lemma, which was proved independently by [16, 20].

LEMMA 6 (Borel-TIS).   *Let* $f(t), t \in \mathcal{U}, \mathcal{U}$ *is a parameter set, be a mean zero Gaussian random field.* f *is almost surely bounded on* $\mathcal{U}$. *Then,*

$$E\left(\sup_{\mathcal{U}} f(t)\right) < \infty$$

and

$$P\left(\max_{t \in \mathcal{U}} f(t) - E\left[\max_{t \in \mathcal{U}} f(t)\right] \geq b\right) \leq e^{-b^2/(2\sigma_{\mathcal{U}}^2)},$$

where

$$\sigma_{\mathcal{U}}^2 = \max_{t \in \mathcal{U}} \text{Var}[f(t)].$$

LEMMA 7.   *Let* $\log E \exp(g((uI - \mathbf{z})^{-1/2}S))$ *be defined in* (17), *then there exists a* $\lambda > 0$ *such that for all* $x > 0$

$$P\left(u^{-3/2-3\delta}\left|\log E \exp(\sigma g((uI - \mathbf{z})^{-1/2}S))\right| > x\right) \leq e^{-\lambda x^2}$$

for u sufficiently large.

PROOF.   Note that $g(t)$ is a mean zero Gaussian random field with $\text{Var}(g(t)) = O(|t|^6)$ and $|S| \leq \kappa u^\delta$. A direct application of the Borel-TIS lemma yields the result of this lemma.   □

LEMMA 8.   *For each* $\mathbf{k} \neq \mathbf{k}'$,

$$P\left(I_\sigma(\Xi_{\varepsilon,\mathbf{k}}) > b, I_\sigma(\Xi_{\varepsilon,\mathbf{k}'}) > b\right) = O(1)u^{d-1}e^{-u^2/2-\Theta(u)}.$$

PROOF.   Without loss of generality, we consider $\Xi_{\varepsilon,\mathbf{k}'} = \Xi_\varepsilon$. If Ξ_ε and $\Xi_{\varepsilon,\mathbf{k}'}$ are connected to each other,

$$\begin{aligned}
P\left(I_\sigma(\Xi_{\varepsilon,\mathbf{k}}) > b, I_\sigma(\Xi_\varepsilon) > b\right) &\leq P\left(I_\sigma(\Xi_\varepsilon \cup \Xi_{\varepsilon,\mathbf{k}}) > 2b\right) \\
&= O(1)m(\Xi_\varepsilon)u^{d-1}e^{-(1/2)(u+\log 2+o(1))^2} \\
&= O(1)u^{d-1}e^{-(1/2)u^2-(\log 2+o(1))u}.
\end{aligned}$$

The second step is an application of Theorem 2. If Ξ_ε and $\Xi_{\varepsilon,\mathbf{k}'}$ are not connected, that is, $\inf_{s\in\Xi_\varepsilon, t\in\Xi_{\varepsilon,\mathbf{k}'}} |s-t| \geq \varepsilon = \kappa u^{-1/2+\delta}$, then

$$
\begin{aligned}
P\big(&I_\sigma(\Xi_{\varepsilon,\mathbf{k}}) > b, I_\sigma(\Xi_\varepsilon) > b\big) \\
&\leq P\Big(\sup_{t\in\Xi_\varepsilon} f(t) > u - c\log u, \sup_{t\in\Xi_{\varepsilon,\mathbf{k}}} f(t) > u - c\log u\Big) \\
&\leq P\Big(\sup_{t\in\Xi_\varepsilon, s\in\Xi_{\varepsilon,\mathbf{k}}} f(t) + f(s) > 2u - 2c\log u\Big) \\
&\leq O(1)P\Big(Z > \frac{2u - 2c\log u + O(1)}{\sqrt{4 - \Theta(1)u^{-1+2\delta}}}\Big) \\
&= O(1)u^{d-1}e^{-(1/2)u^2 - \Theta(1)u^{1+2\delta}},
\end{aligned}
$$

where Z is a standard Gaussian random variable. The last inequality is an application of the Borel-TIS lemma in Lemma 6. $\square$

LEMMA 9.   *Let $\mathcal{A}$ and $\mathcal{D}$ be defined in* (28) *and b_0 be defined in* (29). *Then,*
$$
P(\mathcal{D} > b_0, \mathcal{A} > b_0) = o\big(P(\mathcal{A} > b)\big).
$$

PROOF.   Similar to the second case in the proof of Lemma 8, we have

$$
\begin{aligned}
P(D > b_0,&\, A > b_0) \\
&\leq P\Big(\sup_{\Xi_\varepsilon} f(t) > u - c_1\log u, \sup_{\bigcup_{\mathbf{k}\in C^+\setminus\{\mathbf{0},\mathbf{k}\}} \Xi_{\varepsilon,\mathbf{k}}} f(t) > u - c_1\log u\Big) \\
&\leq P\Big(\sup_{s\in\Xi_\varepsilon, t\in\bigcup_{\mathbf{k}\in C^+\setminus\{\mathbf{0},\mathbf{k}\}} \Xi_{\varepsilon,\mathbf{k}}} f(s) + f(t) > 2u - 2c_1\log u\Big) \\
&\leq u^\alpha e^{-(2u-2c_1\log u)^2/(2(4-2\kappa^2 u^{-1+2\delta}))} \leq e^{-u^2/2 - \Theta(1)u^{1+2\delta}}.
\end{aligned}
$$

The conclusion follows immediately. $\square$

LEMMA 10.   *Let $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{D}$ be defined in* (28) *and b_0 be defined in* (29). *Then,*
$$
P(\mathcal{A} + \mathcal{B} > b - b_0, \mathcal{A} > b_0, \mathcal{B} + \mathcal{D} > b_0) = o(1)P(\mathcal{A} > b).
$$

PROOF.   Note that there exists $c' > 0$ such that

$$
\begin{aligned}
P(\mathcal{A} + \mathcal{B} > &\, b - b_0, \mathcal{A} > b_0, \mathcal{B} + \mathcal{D} > b_0) \\
&\leq P\Big(\mathcal{A} + \mathcal{B} > b - b_0, \sup_{\Xi_\varepsilon} f(t) > u - c'\log u, \\
&\qquad\qquad \sup_{\bigcup_{\mathbf{k}\in\mathcal{C}^+\setminus\{\mathbf{0}\}} \Xi_{\varepsilon,\mathbf{k}}} f(t) > u - c'\log u\Big).
\end{aligned}
$$

It suffices to show that the right-hand side of the above inequality is $o(1)P(\mathcal{A} > b)$ and also $o(1)P(\mathcal{A} + \mathcal{B} > b)$. In order to do so, we will borrow part of the

derivations in the proof of Theorem 2. Let u_* solve

$$(2\pi/\sigma)^{d/2} u_*^{-d/2} e^{\sigma u_*} = b - b_0.$$

Note that because $b_0 = u^{-1-d/2} b$, we have $|u - u_*| = o(u^{-1})$ and $e^{-u^2/2} = (1 + o(1)) e^{-u_*^2/2}$. By the results in (25), (26) and (27), we have

$$P\Big(\mathcal{A} + \mathcal{B} > b - b_0, \sup_{\Xi_\varepsilon} f(t) > u - c' \log u, \sup_{\bigcup_{\mathbf{k} \in \mathcal{C}^+ \setminus \{\mathbf{0}\}} \Xi_{\varepsilon,\mathbf{k}}} f(t) > u - c' \log u\Big)$$

$$= o(1) P(\mathcal{A} + \mathcal{B} > b - b_0)$$
$$+ (1 + o(1)) \sigma^{-1} \det(\Gamma)^{-1/2} \det(\mu_{22})^{1/2} u^{-1}$$
$$\times e^{-(1/2)u^2 + (1/(8\sigma^2))\mathbf{1}^\top \mu_{22} \mathbf{1} + (1/(8\sigma^2)) \sum_i \partial_{iiii} C(0)}$$
$$\times \int_{\{|y|_\infty \le 3\kappa u^{1/2+\delta}\}} P\Big(\mathcal{A}' > o_p(1), \sup_{\Xi_\varepsilon} E(t) + g(t) > u_* - c' \log u_*,$$
$$\sup_{\bigcup_{\mathbf{k} \in \mathcal{C}^+ \setminus \{\mathbf{0}\}} \Xi_{\varepsilon,\mathbf{k}}} E(t) + g(t) > u_* - c' \log u_*\Big)$$
$$\times \exp\Big\{-\Big[\frac{\mathcal{A}'}{\sigma} + \frac{\mathcal{B}^\top \mathcal{B}}{2}$$
$$+ \frac{(-\mathcal{A}'/(\sigma u) + \mu_{20} \mu_{22}^{-1/2} \mathcal{B} + (1/(2\sigma)) \mu_{20} \mathbf{1})^2}{2(1 - \mu_{20} \mu_{22}^{-1} \mu_{02})}\Big]\Big\} \, d\mathcal{A}' \, d\mathcal{B} \, dy.$$

Note that the only change in the above display from (25) is the probability inside the integral. In what follows, we show that it is almost always $o(1)$. Note that $\operatorname{Var}(g(t)) = O(|t|^6)$. Therefore, for any $f(0) < u + u^{\varepsilon_0}$ with $\varepsilon_0 < \delta/2$, if

(32)
$$\sup_{\Xi_\varepsilon} E(t) < u - c' \log u - \Theta(u^{-1}) \quad \text{or}$$
$$\sup_{\Xi_{3\varepsilon} \setminus \Xi_\varepsilon} E(t) < u - c' \log u - \Theta(u^{-1}),$$

then

$$P\Big(\sup_{\Xi_\varepsilon} E(t) + g(t) > u - c' \log u, \sup_{\bigcup_{\mathbf{k} \in \mathcal{C}^+ \setminus \{\mathbf{0}\}} \Xi_{\varepsilon,\mathbf{k}}} E(t) + g(t) > u - c' \log u\Big) = o(1).$$

This fact implies that $g(t)$ can be basically ignored. Therefore, it is useful to keep in mind that "$E(t) \approx f(t)$."

Since

$$P\Big(\sup_T f(t) > u + u^{-1+\varepsilon_0}\Big) = o(1) P(\mathcal{A} + \mathcal{B} > b - b_0),$$

we only need to consider the case that $\sup_T E(t) \le u + u^{-1+\varepsilon_0}$. Given the form

$$E(t) = u_* - w + y^\top t + \frac{1}{2} t^\top (-u_* I + \mathbf{z}) t + g_3(t) + g_4(t) + R(t),$$

which is asymptotically quadratic in $\Xi_{3\varepsilon}$, let $t^* = \arg\sup_{\Xi_{3\varepsilon}} E(t)$. On the set that $\sup_T E(t) \le u + u^{-1+\varepsilon_0}$, we have

$$\sup_{|t-t^*|>2u^{-1/2+\varepsilon_0/2}} E(t) < u - c' \log u - \Theta(u^{-1}).$$

Let $\partial \Xi_\varepsilon$ be the border of Ξ_ε. Then

(33)
$$\sup_{\Xi_\varepsilon} E(t) > u - c' \log u - \Theta(u^{-1}) \quad \text{and}$$

$$\sup_{\Xi_{3\varepsilon}\setminus\Xi_\varepsilon} E(t) > u - c' \log u - \Theta(u^{-1}),$$

only when $\inf_{t\in\partial\Xi_\varepsilon} |t - t^*| < u_*^{-1/2+\varepsilon_0}$. This implies that

$$\inf_{t\in\partial\Xi_\varepsilon} |t - (u_*I - \mathbf{z})^{-1}y| < u_*^{-1/2+\varepsilon_0}.$$

Therefore, $t^* = \arg\sup f(t)$ must be very closed to the boundary of Ξ_ε so as to have (33) hold.

Therefore, for all $\varepsilon_0 < \delta$

$$P\Big(\mathcal{A}+\mathcal{B} > b - b_0, \sup_{\Xi_\varepsilon} f(t) > u - c' \log u,$$

$$\sup_{\bigcup_{\mathbf{k}\in\mathcal{C}^+\setminus\{\mathbf{0}\}} \Xi_{\varepsilon,\mathbf{k}}} E(t) + g(t) > u - c' \log u\Big)$$

$$= o(1) P(\mathcal{A}+\mathcal{B} > b - b_0)$$

$$+ \big(1 + o(1)\big)\sigma^{-1} \det(\Gamma)^{-1/2} \det(\mu_{22})^{1/2} u^{-1}$$

$$\times e^{-(1/2)u^2 + (1/(8\sigma^2))\mathbf{1}^\top \mu_{22}\mathbf{1} + (1/(8\sigma^2))\sum_i \partial_{iiii}C(0)}$$

$$\times \int_{\inf_{t\in\partial\Xi_\varepsilon} |t-(u_*I-\mathbf{z})^{-1}y|<u_*^{-1/2+\varepsilon_0}} P\Big(A' > o_p(1),$$

$$\sup_{\Xi_\varepsilon} E(t) + g(t) > u - c' \log u,$$

$$\sup_{\bigcup_{\mathbf{k}\in\mathcal{C}^+\setminus\{\mathbf{0}\}} \Xi_{\varepsilon,\mathbf{k}}} E(t) + g(t) > u - c' \log u\Big)$$

$$\times \exp\Big\{-\Big[\frac{A'}{\sigma} + \frac{B^\top B}{2}$$

$$+ \frac{(-A'/(\sigma u) + \mu_{20}\mu_{22}^{-1/2}B + (1/(2\sigma))\mu_{20}\mathbf{1})^2}{2(1 - \mu_{20}\mu_{22}^{-1}\mu_{02})}\Big]\Big\} dA'\, dB\, dy$$

$$= o(1) P(\mathcal{A}+\mathcal{B} > b - b_0).$$

The last equation is because

$$\text{mes}\Big(\Big\{y: \inf_{t\in\partial\Xi_\varepsilon}|t-(u_*I-\mathbf{z})^{-1}y|<u_*^{-1/2+\varepsilon_0}\Big\}\Big)$$

$$= o\big(\text{mes}(\{y: |(uI-\mathbf{z})^{-1/2}y|_\infty \le \kappa u^\delta - u^{\delta/2}\})\big).$$

Hereby, we conclude the proof. □

REFERENCES

[1] ADLER, R. J. (1981). *The Geometry of Random Fields*. Wiley, Chichester. MR0611857

[2] ADLER, R. J., BLANCHET, J. H. and LIU, J. C. (2008). Efficient simulation for tail probabilities of Gaussian random fields. In *Proceedings of the 40th Winter Simulation Conference* (*Miami, Florida*) 328–336.

[3] ADLER, R. J., BLANCHET, J. H. and LIU, J. C. (2009). Efficient Monte Carlo for large excursions of Gaussian random fields. Preprint.

[4] ADLER, R. J., SAMORODNITSKY, G. and TAYLOR, J. E. (2009). High level excursion set geometry for non-Gaussian infinitely divisible random fields. Preprint.

[5] ADLER, R. J. and TAYLOR, J. E. (2007). *Random Fields and Geometry*. Springer, New York. MR2319516

[6] ADLER, R. J., TAYLOR, J. E. and WORSLEY, K. J. (2009). Applications of RANDOM FIELDS AND GEOMETRY: Foundations and case studies. Preprint.

[7] AHSAN, S. M. (1978). Portfolio selection in a lognormal securities market. *Journal of Economics* **38** 105–118.

[8] ASMUSSEN, S. and ROJAS-NANDAYAPA, L. (2008). Asymptotics of sums of lognormal random variables with Gaussian copula. *Statist. Probab. Lett.* **78** 2709–2714. MR2465111

[9] AZAÏS, J.-M. and WSCHEBOR, M. (2005). On the distribution of the maximum of a Gaussian field with *d* parameters. *Ann. Appl. Probab.* **15** 254–278. MR2115043

[10] AZAÏS, J.-M. and WSCHEBOR, M. (2008). A general expression for the distribution of the maximum of a Gaussian field and the approximation of the tail. *Stochastic Process. Appl.* **118** 1190–1218. MR2428714

[11] AZAÏS, J.-M. and WSCHEBOR, M. (2009). *Level Sets and Extrema of Random Processes and Fields*. Wiley, Hoboken, NJ. MR2478201

[12] BASAK, S. and SHAPIRO, A. (2001). Value-at-risk-based risk management: Optimal policies and asset prices. *Review of Financial Studies* **14** 371–405.

[13] BERMAN, S. M. (1985). An asymptotic formula for the distribution of the maximum of a Gaussian process with stationary increments. *J. Appl. Probab.* **22** 454–460. MR0789369

[14] BLACK, F. and SCHOLES, M. (1973). Pricing of options and corporate liabilities. *Journal of Political Economy* **81** 637–654.

[15] BLANCHET, J., JUNEJA, J. and ROJAS-NANDAYAPA, L. (2012). Efficient tail estimation for sums of correlated lognormals. *Ann. Oper. Res.* To appear.

[16] BORELL, C. (1975). The Brunn–Minkowski inequality in Gauss space. *Invent. Math.* **30** 207–216. MR0399402

[17] BORELL, C. (2003). The Ehrhard inequality. *C. R. Math. Acad. Sci. Paris* **337** 663–666. MR2030108

[18] CAMPBELL, M. J. (1994). Time-series regression for counts—an investigation into the relationship between sudden-infant-death-syndrome and environmental-temperature. *J. Roy. Statist. Soc. Ser. A* **157** 191–208.

[19] CHAN, K. S. and LEDOLTER, J. (1995). Monte Carlo EM estimation for time series models involving counts. *J. Amer. Statist. Assoc.* **90** 242–252. MR1325132

[20] CIREL'SON, B. S., IBRAGIMOV, I. A. and SUDAKOV, V. N. (1976). Norms of Gaussian sample functions. In *Proceedings of the Third Japan–USSR Symposium on Probability Theory* (*Tashkent*, 1975). *Lecture Notes in Math.* **550** 20–41. Springer, Berlin. MR0458556

[21] COX, D. R. (1955). Some statistical methods connected with series of events. *J. Roy. Statist. Soc. Ser. B.* **17** 129–157; discussion, 157–164. MR0092301

[22] COX, D. R. and ISHAM, V. (1980). *Point Processes*. Chapman and Hall, London. MR0598033

[23] DAVIS, R. A., DUNSMUIR, W. T. M. and WANG, Y. (2000). On autocorrelation in a Poisson regression model. *Biometrika* **87** 491–505. MR1789805

[24] DENISOV, D., DIEKER, A. B. and SHNEER, V. (2008). Large deviations for random walks under subexponentiality: The big-jump domain. *Ann. Probab.* **36** 1946–1991. MR2440928

[25] DEUTSCH, H.-P. (2004). *Derivatives and Internal Models. Finance and Capital Markets*, 3rd ed. Palgrave Macmillan, Hampshire.

[26] DUFFIE, D. and PAN, J. (1997). An overview of value at risk. *The Journal of Derivatives* **4** 7–49.

[27] DUFRESNE, D. (2001). The integral of geometric Brownian motion. *Adv. in Appl. Probab.* **33** 223–241. MR1825324

[28] FOSS, S. and RICHARD, A. (2010). On sums of conditionally independent subexponential random variables. *Math. Oper. Res.* **35** 102–119. MR2676758

[29] GLASSERMAN, P. (2004). *Monte Carlo Methods in Financial Engineering. Applications of Mathematics* (*New York*) **53**. Springer, New York. MR1999614

[30] GLASSERMAN, P., HEIDELBERGER, P. and SHAHABUDDIN, P. (2000). Variance reduction techniques for estimating value-at-risk. *Manag. Sci.* **46** 1349–1364.

[31] LANDAU, H. J. and SHEPP, L. A. (1970). On the supremum of a Gaussian process. *Sankhyā Ser. A* **32** 369–378. MR0286167

[32] LEDOUX, M. and TALAGRAND, M. (1991). *Probability in Banach Spaces*: Isoperimetry and Processes. *Ergebnisse der Mathematik und Ihrer Grenzgebiete* (3) [*Results in Mathematics and Related Areas* (3)] **23**. Springer, Berlin. MR1102015

[33] MARCUS, M. B. and SHEPP, L. A. (1970). Continuity of Gaussian processes. *Trans. Amer. Math. Soc.* **151** 377–391. MR0264749

[34] PITERBARG, V. I. (1996). *Asymptotic Methods in the Theory of Gaussian Processes and Fields. Translations of Mathematical Monographs* **148**. Amer. Math. Soc., Providence, RI. MR1361884

[35] SUDAKOV, V. N. and TSIRELSON, B. S. (1974). Extremal properties of half spaces for spherically invariant measures. *Zap. Nauchn. Sem. LOMI* **45** 75–82.

[36] SUN, J. (1993). Tail probabilities of the maxima of Gaussian random fields. *Ann. Probab.* **21** 34–71. MR1207215

[37] TALAGRAND, M. (1996). Majorizing measures: The generic chaining. *Ann. Probab.* **24** 1049–1103. MR1411488

[38] TAYLOR, J., TAKEMURA, A. and ADLER, R. J. (2005). Validity of the expected Euler characteristic heuristic. *Ann. Probab.* **33** 1362–1396. MR2150192

[39] YOR, M. (1992). On some exponential functionals of Brownian motion. *Adv. in Appl. Probab.* **24** 509–531. MR1174378

[40] ZEGER, S. L. (1988). A regression model for time series of counts. *Biometrika* **75** 621–629. MR0995107

COLUMBIA UNIVERSITY
1255 AMSTERDAM AVE.
NEW YORK, NEW YORK 10027
USA
E-MAIL: jcliu@stat.columbia.edu