ON THE CONDITIONAL DISTRIBUTIONS AND THE EFFICIENT SIMULATIONS OF EXPONENTIAL INTEGRALS OF GAUSSIAN RANDOM FIELDS

BY JINGCHEN LIU[1] AND GONGJUN XU

Columbia University and University of Minnesota

In this paper, we consider the extreme behavior of a Gaussian random field $f(t)$ living on a compact set T. In particular, we are interested in tail events associated with the integral $\int_T e^{f(t)} \, dt$. We construct a (non-Gaussian) random field whose distribution can be explicitly stated. This field approximates the conditional Gaussian random field f (given that $\int_T e^{f(t)} \, dt$ exceeds a large value) in total variation. Based on this approximation, we show that the tail event of $\int_T e^{f(t)} \, dt$ is asymptotically equivalent to the tail event of $\sup_T \gamma(t)$ where $\gamma(t)$ is a Gaussian process and it is an affine function of $f(t)$ and its derivative field. In addition to the asymptotic description of the conditional field, we construct an efficient Monte Carlo estimator that runs in polynomial time of $\log b$ to compute the probability $P(\int_T e^{f(t)} \, dt > b)$ with a prescribed relative accuracy.

1. Introduction. Consider a Gaussian random field $\{f(t) : t \in T\}$ living on a d-dimensional domain $T \subset R^d$ with zero mean and unit variance, that is, for every finite subset $\{t_1, \ldots, t_n\} \subset T$, $(f(t_1), \ldots, f(t_n))$ is a mean zero multivariate Gaussian random vector. Let $\mu(t)$ be a (deterministic) function and $\sigma \in (0, \infty)$ be a scale factor. Define

$$(1.1) \qquad \mathcal{I}(T) \triangleq \int_T e^{\sigma f(t) + \mu(t)} \, dt.$$

In this paper, we develop a precise asymptotic description of the conditional distribution of f given that $\mathcal{I}(T)$ exceeds a large value b, that is, $P(\cdot | \mathcal{I}(T) > b)$. In particular, we provide a tractable total variation approximation (in the sample path space) for such conditional random fields based on a change of measure technique. In addition to the asymptotic descriptions, we design efficient Monte Carlo estimators that run in polynomial time of $\log b$ for computing the tail probabilities

$$(1.2) \qquad v(b) = P\big(\mathcal{I}(T) > b\big) = P\bigg(\int_T e^{\sigma f(t) + \mu(t)} \, dt > b\bigg)$$

with a prescribed relative accuracy.

1.1. *The literature.* In the probability literature, the extreme behaviors of Gaussian random fields have been studied extensively. The results range from general bounds to sharp asymptotic approximations. An incomplete list of works includes [15, 20, 23, 35, 37, 39, 46, 50, 52]. A few lines of investigations on the supremum norm are given as follows. Assuming locally stationary structure, the double-sum method [49] provides the exact asymptotic approximation of $\sup_T f(t)$ over a compact set T, which is allowed to grow as the threshold tends to infinity. For almost surely at least twice differentiable fields, the authors of [2, 5, 53] derive the analytic form of the expected Euler–Poincaré characteristics of the excursion set $[\chi(A_b)]$ which serves as a good approximation of the tail probability of the supremum. The tube method [51] takes advantage of the Karhune–Loève expansion and Weyl's formula. A recent related work along this line is given by [48]. The Rice method [11–13] provides an implicit description of $\sup_T f(t)$. Change of measure based rare-event simulations are studied in [3]. The discussions also go beyond the Gaussian fields. For instance, [36] discusses the situations of Gaussian process with random variances. See also [4] for discussions on non-Gaussian cases. The distribution of $\mathcal{I}(T)$ is studied in the literature when $f(t)$ is a Brownian motion [29, 56]. Recently, [42, 43] derive the asymptotic approximations of $P(\mathcal{I}(T) > b)$ as $b \to \infty$ for three times differentiable and homogeneous Gaussian random fields.

Besides the tail probability approximations, rigorous analysis of the conditional distributions of stochastic processes given the occurrence of rare events is also an important topic. In the classic large deviations analysis for light-tailed stochastic systems, the sample path(s) that admits the highest probability (the most likely sample path) under the conditional distribution given the occurrence of a rare event is central to the entire analysis in terms of determining the appropriate exponential change of measure, developing approximations of the tail probabilities and designing efficient simulation algorithms; see, for instance, standard textbook [30]. For heavy-tailed systems, the conditional distributions and the most likely paths, which typically admit the so-called "one-big-jump" principle, are also intensively studied [8, 9, 17]. These results not only provide intuitive and qualitative descriptions of the conditional distribution, but also shed light on the design of rare-event simulation algorithms [16–18]—the best importance sampling estimator of the rare-event probability uses a change of measure corresponding to the interesting conditional distribution. In addition, the conditional distribution (or the conditional expectations) is also of practical interest. For instance, in risk management, the conditional expected loss given some rare/disastrous event is an important risk measure and stress test.

In the literature of Gaussian random fields, the exact Slepian models [conditional field given a local maximum or level crossing of $f(t)$] are studied intensively for twice differentiable fields. For instance, Leadbetter, Lindgren and Rootzén [38] give the Slepian model conditioning on an upcrossing of level u at time zero. Lindgren [40] treats conditioning on a local maximum of height u at time zero.

The first rigorous treatment of Slepian models for nonstationary processes is given by Lindgren [41]. Grigoriu [34] extends the results of Leadbetter, Lindgren and Rootzén [38] for level crossings to the general nonstationary case. This work is followed up by Gadrich and Adler [32]. In the later analysis, we will set an asymptotic equivalence between the conditional distribution given $\{\mathcal{I}(T) > b\}$ and that given the high excursion of the supremem of f. The latter can be characterized by the Slepain model.

1.2. *Contributions*.   In this paper, we pursue along this line for the extreme behaviors of Gaussian processes and begin to describe the conditional distribution of f given the occurrence of the event $\{\mathcal{I}(T) > b\}$. In particular, we provide both quantitative and qualitative descriptions of this conditional distribution. Furthermore, from a computational point of view, we construct a Monte Carlo estimator that takes a polynomial computational cost (in $\log b$) to estimate $v(b)$ for a prescribed relative accuracy.

Central to the analysis is the construction of a change of measure on the space $C(T)$ (continuous functions living on T). The application of the change of measure ideas is common in the study of large deviations analysis for the light-tailed stochastic systems. However, it is not at all standard in the study of Gaussian random fields. The proposed change of measure is not of a classical exponential-tilting form. This measure has several features that are appealing both theoretically and computationally. First, we show that the change of measure denoted by Q approximates the conditional measure $P(\cdot|\mathcal{I}(T) > b)$ in total variation as $b \to \infty$. Second, the measure Q is analytically tractable in the sense that the distribution of f under Q has a closed form representation and the Radon–Nikodym derivative dQ/dP takes the form of a d-dimensional integral. This tractability property has useful consequences. From a methodological point of view, the measure Q provides a very precise description of the mechanism that drives the rare event $\{\mathcal{I}(T) > b\}$. This result allows us to directly use the intuitive mechanism to provide functional probabilistic descriptions that emphasize the most important elements that are present in the interesting rare events. More technically, the analytical computations associated with the measure Q are easy (compared to the conditional measure), and the expectation $E^Q[\cdot]$ is theoretically much more tractable than $E[\cdot|\mathcal{I}(T) > b]$. Based on this result, we show that the tail event $\{\mathcal{I}(T) > b\}$ is asymptotically equivalent to the tail event of $\sup_T \gamma(t)$ where $\gamma(t)$ is an affine function of $f(t)$ and its derivative field $\partial^2 f(t)$ and $\gamma(t)$ implicitly depends on b. Thus, one can further characterize the conditional measure by means of the results on the Slepian model mentioned earlier.

Another contribution of this paper lies in the numerical evaluation of $v(b)$. The importance sampling algorithm associated with the proposed change of measure yields an efficient estimator for computing $v(b)$. An important issue concerns the implementation of the Monte Carlo method. The processes considered in this paper are continuous while computers can only represent discrete objects. Inevitably, we

will introduce a suitable discretization scheme and use discrete (random) objects to approximate the continuous processes. A naturally raised issue lies in the control of the approximation error relative to the probability $v(b)$. We will perform careful analysis and report the overall computational complexity of the proposed Monte Carlo estimators.

A key requirement of the current analysis is the twice differentiability of f. Our change of measure is written explicitly in the form of f, ∂f and $\partial^2 f$. A very interesting future study would be developing parallel results for nondifferentiable fields. The technical challenges are two-fold. First, there is lack of asymptotic analysis for the exponential integral of general nondifferentiable fields. To the author's best knowledge, the behavior of $\mathcal{I}(T)$ for nondifferentiable processes is investigated only when f is a Brownian motion whose techniques cannot be extend to general cases [29, 56]. In addition, there is a lack of descriptive tools (such as derivatives and the Palm model) for nondifferentiable processes. This also leads to difficulties in describing the Slepian model for level crossing. To the author's best knowledge, analytic description of Slepian models for excursion of $\sup_T f(t)$ are available only for twice differentiable fields. Despite of the smoothness limitation, the current analysis has important applications the details of which will be presented in the following section.

The rest of this paper is organized as follows. Two applications of this work are given in Section 2. In Section 3, we present the main results including the change of measure, the approximation of $P(\cdot | \mathcal{I}(T) > b)$ and the efficient Monte Carlo estimator of $v(b)$. Proofs of the theorems are given in Sections 4–7. A supplemental material [45] is provided including all the supporting lemmas.

2. Applications. The integral of exponential functions of Gaussian random fields plays an important role in many probability models. We present two such models for which the conditional distribution is of interest and the underlying random fields are differentiable.

2.1. *Spatial point process.* In spatial point process modeling, let $\lambda(t)$ be the intensity of a Poisson point process on T, denoted by $\{N(A) : A \subset T\}$. In order to build in spatial dependence structure and to account for overdispersion, the log-intensity is typically modeled as a Gaussian random field, that is, $\log \lambda(t) = f(t) + \mu(t)$ and then $E[N(A)|\lambda(\cdot)] = \int_A e^{f(t)+\mu(t)} \, dt$, where $\mu(t)$ is the mean function, and $f(t)$ is a zero-mean Gaussian process. For instance, Chan and Ledolter [22] consider the time series setting in which T is a one-dimensional interval, $\mu(t)$ is modeled as the observed covariate process and $f(t)$ is an autoregressive process; see [21, 24–26, 57] for more examples in high-dimensional domains.

For the purpose of illustration, we consider a very concrete case that the point process $N(\cdot)$ represents the spatial distribution of asthma cases over a geographical domain T. The latent intensity $\lambda(t)$ [or equivalently $f(t)$] represents the unobserved (and appropriately transformed) metric of the pollution severity at location t. The mean function can be written as a linear combination of the observed

covariates that may affect the pollution level, that is, $\mu(t) = \beta^\top x(t)$ is treated as a deterministic function. It is well understood that $\lambda(t)$ is a smooth function of the spatial parameter t at the macro level as the atmosphere mixes well; see, for example, [1]. One natural question in epidemiology is the following: upon observing an unusually high number of asthma cases, what is their geographical distribution, that is, the conditional distribution of the point process $N(\cdot)$ given that $N(T) > b$ for some large b?

First of all, Liu and Xu [43] show that $P(N(T) > b) \sim P(\mathcal{I}(T) > b)$ as $b \to \infty$. Following the same derivations, it is not difficult to establish the following convergence:

$$P\big(\cdot|N(T) > b\big) - P\big(\cdot|\mathcal{I}(T) > b\big) \to 0 \qquad \text{in total variation as } b \to \infty.$$

The total count $N(T)$ is a Poisson random variable with mean $\mathcal{I}(T)$. Intuitively speaking, the tail of the integral is similar to a lognormal random variable and thus is heavy-tailed. Its overshoot over level b is $O_p(b/\log b)$. However, a Poisson random variable with mean $\mathcal{I}(T) \sim b$ has standard deviation $\sqrt{b} \ll b/\log b$. Thus, a large number of $N(T)$ is mainly caused by a large value of $\mathcal{I}(T)$. The symmetric difference of the two sets $\{N(T) > b\}$ and $\{\mathcal{I}(T) > b\}$ vanishes, and the probability law of the entire system conditional upon observing that $N(T) > b$ is asymptotically the same as that given $\mathcal{I}(T) > b$. Therefore, the conditional distribution of $N(\cdot)$ given $N(T) > b$ is asymptotically another doubly-stochastic Poisson process whose intensity is $\lambda(t) = e^{\mu(t) + f(t)}$ where $f(t)$ follows the conditional distribution of $P(f \in \cdot|\mathcal{I}(T) > b)$.

Based on the main results presented momentarily, a qualitative description of the conditional distribution of $N(\cdot)$ is as follows. Given $N(T) > b$, the overshoot is of order $O_p(b/\log b)$, that is, $N(T) = b + O_p(b/\log b)$. The locations of the points are i.i.d. samples approximately following a d-dimensional multivariate Gaussian distribution with mean $\tau \in T$ and variance $\Sigma/\log b$ where Σ depends on the spectral moments of f. The distribution of τ is uniform over T if $\mu(t)$ is a constant; if $\mu(t)$ is not constant, τ has a density $l(t)$ presented in (3.13).

2.2. Financial application.

The exponential integral can be considered as a generalization of the sum of dependent lognormal random variables that has been studied intensively from different aspects in the applied probability literature (see [7, 10, 14, 27, 28, 31, 33]). In portfolio risk analysis, consider a portfolio of n assets $S_1, \ldots, S_n$. The asset prices are usually modeled as log-normal random variables. That is, let $X_i = \log S_i$ and $(X_1, \ldots, X_n)$ follows a multivariate normal distribution. The total portfolio value $S = \sum_{i=1}^n w_i S_i$ is the weighted sum of dependent log-normal random variables.

An important question is the behavior of this sum when the portfolio size becomes large and the assets are highly correlated. One may employ a latent space

approach used in the literature of social networks. More specifically, we construct a Gaussian process $\{f(t) : t \in T\}$ and map each asset i to a latent variable $t_i \in T$, that is, $\log S_i = f(t_i)$. Then the log-asset prices fall into a subset of the continuous Gaussian process. Furthermore, we construct a (deterministic) function $w(t)$ so that $w(t_i) = w_i$. Then, the unit share value of the portfolio is $\frac{1}{n} \sum w_i S_i = \frac{1}{n} \sum w(t_i) e^{f(t_i)}$. See [19, 43] for detailed discussions on the random field representations of large portfolios.

In the asymptotic regime that $n \to \infty$ and the correlations among the asset prices become close to one, the subset $\{t_i\}$ becomes dense in T. Ultimately, we obtain the limit

$$\frac{1}{n} \sum_{i=1}^{n} w_i S_i \to \int_T w(t) e^{f(t)} h(t) \, dt,$$

where $h(t)$ is the limiting spatial distribution of $\{t_i\}$ in T. Let $\mu(t) = \log w(t) + \log h(t)$. Then the (limiting) unit share price is $\mathcal{I}(T) = \int_T e^{f(t)+\mu(t)} \, dt$.

The current study provides an asymptotic description of the performance of each asset given the occurrence of the tail event $\mathcal{I}(T) > b$. This is of great importance in the study of the so-called *stress test* that evaluates the impact of shocks on and the vulnerability of a system. For instance, consider that another investor holds a different portfolio that has a substantial overlap with the current one, or it has exactly the same collection of assets but with different weights. Thus, this second portfolio corresponds to a different mean function $\mu'(t)$. The stress test investigates the performance of this second portfolio on the condition that a rare event has occurred to the first, that is,

$$P\left(\int_T e^{f(t)+\mu'(t)} \, dt \in \cdot \,\Big|\, \int_T e^{f(t)+\mu(t)} \, dt > b \right).$$

To characterize the above distribution, we need a precise description of the conditional measure $P(f \in \cdot \,|\, \int_T e^{f(t)+\mu(t)} \, dt > b)$.

3. Main results.

3.1. *Problem setting and notation.* Throughout this discussion, we consider a homogeneous Gaussian random field $\{f(t) : t \in T\}$ living on a domain $T \subset R^d$. Let the covariance function be

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

We impose the following assumptions:

(C1) f is stationary with $Ef(t) = 0$ and $Ef^2(t) = 1$.
(C2) f is almost surely at least two times differentiable with respect to t.
(C3) T is a d-dimensional compact set of R^d with piecewise smooth boundary.

(C4) The Hessian matrix of $C(t)$ at the origin is standardized to be $-I$, where I is the $d \times d$ identity matrix. In addition, $C(t)$ has the following expansion when t is close to 0

$$(3.1) \qquad C(t) = 1 - \tfrac{1}{2}t^\top t + C_4(t) + R_C(t),$$

where $C_4(t) = \frac{1}{24} \sum_{ijkl} \partial^4_{ijkl} C(0) t_i t_j t_k t_l$ and $R_C(t) = O(|t|^{4+\delta_0})$ for some $\delta_0 > 0$.

(C5) For each $t \in R^d$, the function $C(\lambda t)$ is a nonincreasing function of $\lambda \in R^+$.

(C6) The mean function $\mu(t)$ falls into either of the two cases:

(a) $\mu(t) \equiv 0$;

(b) the maximum of $\mu(t)$ is unique and is attained in the interior of T and $\mu(t + \varepsilon) - \mu(t) = \varepsilon^\top \partial \mu(t) + \varepsilon^\top \Delta \mu(t) \varepsilon + O(|\varepsilon|^{2+\delta_0})$ as $\varepsilon \to 0$.

We define a set of notation constantly used in the later development and provide some basic calculations. Let P_b^* be the conditional measure given $\{\mathcal{I}(T) > b\}$, that is,

$$P_b^*(f(\cdot) \in A) = P(f(\cdot) \in A | \mathcal{I}(T) > b).$$

Let "∂" denote the gradient and "Δ" denote the Hessian matrix with respect to t. The notation "∂^2" is used to denote the vector of second derivatives. The difference between $\partial^2 f(t)$ and $\Delta f(t)$ is that $\Delta f(t)$ is a $d \times d$ symmetric matrix whose diagonal and upper triangle consist of elements of $\partial^2 f(t)$. Furthermore, let $\partial_j f(t)$ be the partial derivative with respect to the jth element of t. Finally, we define the following vectors:

$$(3.2) \qquad \begin{aligned} \mu_1(t) &= -(\partial_1 C(t), \ldots, \partial_d C(t)), \\ \mu_2(t) &= (\partial^2_{ii} C(t), i = 1, \ldots, d; \partial^2_{ij} C(t), \\ &\qquad i = 1, \ldots, d-1, j = i+1, \ldots, d), \\ \mu_{02}^\top &= \mu_{20} = \mu_2(0). \end{aligned}$$

Suppose $0 \in T$. It is well known that $(f(0), \partial^2 f(0), \partial f(0), f(t))$ is a multivariate Gaussian random vector with mean zero and covariance matrix (cf. Chapter 5.5 of [5])

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

where the matrix μ_{22} is a $d(d+1)/2$-dimensional positive definite matrix and contains the 4th order spectral moments arranged in an appropriate order according to the order of elements in $\partial^2 f(0)$. Let $h(x, y, z)$ be the density function of

$(f(t), \partial f(t), \partial^2 f(t))$ evaluated at (x, y, z). Then, simple calculation yields that

(3.3)
$$h(x, y, z)$$
$$= \frac{\det(\Gamma)^{-1/2}}{(2\pi)^{(d+1)(d+2)/4}} e^{-(1/2)[y^\top y + (x - \mu_{20}\mu_{22}^{-1}z)^2/(1 - \mu_{20}\mu_{22}^{-1}\mu_{02}) + z^\top \mu_{22}^{-1}z]},$$

where $\det(\cdot)$ is the determinant of a matrix and

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

We define u as a function of b such that

(3.4)
$$\left(\frac{2\pi}{\sigma}\right)^{d/2} u^{-d/2} e^{\sigma u} = b.$$

Note that the above equation generally has two solutions: one is approximately $\sigma^{-1}\log b$, and the other is close to zero as $b \to \infty$. We choose u to be the one close to $\sigma^{-1}\log b$. For $\mu(t)$ and σ appearing in (1.1), we define

(3.5)
$$\mu_\sigma(t) = \mu(t)/\sigma, \qquad u_t = u - \mu_\sigma(t).$$

Approximately, u_t is the level that $f(t)$ needs to reach so that $\mathcal{I}(T) > b$. Furthermore, we need the following spatially varying set:

(3.6)
$$A_t = \{f(\cdot) \in C(T) : \alpha_t > u_t - \eta u_t^{-1}\},$$

where $\eta > 0$ is a tuning parameter that will be eventually sent to zero as $b \to \infty$ and α_t is a function of $f(t)$ and its derivative fields taking the form of

(3.7)
$$\alpha_t = f(t) + \frac{|\partial f(t)|^2}{2u_t} + \frac{\mathbf{1}^\top \bar{f}_t''}{2\sigma u_t} + \frac{B_t}{u_t}.$$

In the above equation (3.7), $\bar{f}_t''$ is defined as [with the notation in (3.2)]

(3.8)
$$\bar{f}_t'' = \partial^2 f(t) - u_t \mu_{02}.$$

The term B_t is a deterministic function depending only on $C(t)$, $\mu(t)$ and σ,

(3.9)
$$B_t = \frac{\mathbf{1}^\top \partial^2 \mu_\sigma(t) + d \times \mu_\sigma(t)}{2\sigma} + \frac{1}{8\sigma^2}\sum_i \partial_{iiii}^4 C(0) + |\partial \mu_\sigma(t)|^2,$$

where d is the dimension of T, and $\mathbf{1} = (\underbrace{1, \ldots, 1}_{d}, \underbrace{0, \ldots, 0}_{d(d-1)/2})^\top$. Note that $\alpha_t \approx f(t)$. Thus on the set A_t, $f(t) \approx \alpha_t > u_t - O(u^{-1})$. Together with the fact that $E[\partial^2 f(t)|f(t) = u_t] = u_t \mu_{02}$, $\bar{f}_t''$ is the standardized second derivative of f given that $f(t) = u_t$. In Section 3.2, we will show that the event $\{\mathcal{I}(T) > b\}$ is approximately $\bigcup_{t \in T} A_t$.

For notational 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 $a_u \sim b_u$ if $a_u/b_u \to 1$ as $u \to \infty$.

REMARK 1.   Condition C1 assumes unit variance. We treat the standard deviation σ as an additional parameter and consider $\int e^{\mu(t)+\sigma f(t)} \, dt$. Condition C2 implies that $C(t)$ is at least 4 times differentiable and the first and third derivatives at the origin are all zero. Differentiability is a crucial assumption in this analysis. Condition C3 restricts the results to finite horizon. Condition C4 assumes the Hessian matrix is standardized to be $-I$, which is to simplify notation. For any Gaussian process $g(t)$ with covariance function $C_g(t)$ and $\Delta C_g(0) = -\Sigma$ and $\det(\Sigma) > 0$, identity Hessian matrix can be obtained by an affine transformation by letting $g(t) = f(\Sigma^{1/2}t)$ and

$$\int_T e^{\mu(t)+\sigma g(t)} \, dt = \det(\Sigma^{-1/2}) \int_{\{s : \Sigma^{-1/2}s \in T\}} e^{\mu(\Sigma^{-1/2}s)+\sigma f(s)} \, ds.$$

Condition C5 is imposed for technical reasons so that we are able to localize the integration. For condition C6, we assume that $\mu(t)$ either is a constant or attains its global maximum at one place. If $\mu(t)$ has multiple (finitely many) maxima, the techniques developed in this paper still apply, but the derivations will be more tedious. Therefore, we stick to the uni-mode case.

REMARK 2.   The setting in (1.2) incorporates the case in which the integral is with respect to other measures with smooth densities. Then, if $\nu(dt) = \kappa(t) \, dt$, we will have that

$$\int_A e^{\mu(t)+\sigma f(t)} \nu(dt) = \int_A e^{\mu(t)+\log\kappa(t)+\sigma f(t)} \, dt,$$

which shows that the density can be absorbed by the mean function.

3.2. *Approximation of the conditional distribution.*   In this subsection, we propose a change of measure Q on the sample path space $C(T)$ that approximates P_b^* in total variation. Let P be the original measure. The measure Q is defined such that P and Q are mutually absolutely continuous. We define the measure Q under two different scenarios: $\mu(t)$ is not a constant and $\mu(t) \equiv 0$. Note that the measure Q obviously will depend on b. To simplify the notation, we omit the index b in Q whenever there is no ambiguity.

The measure Q takes a mixture form of three measures, which are weighted by $(1 - \rho_1 - \rho_2)$, ρ_1 and ρ_2, respectively (a natural constraint is that ρ_1, ρ_2 and

$1 - \rho_1 - \rho_2 \in [0, 1]$). We define Q through the Radon–Nikodym derivative

$$
\begin{aligned}
(3.10) \quad \frac{dQ}{dP} &= (1 - \rho_1 - \rho_2) \int_T l(t) \cdot \mathrm{LR}(t) \, dt + \rho_1 \int_T l(t) \cdot \mathrm{LR}_1(t) \, dt \\
&\quad + \rho_2 \int_T \frac{\mathrm{LR}_2(t)}{\mathrm{mes}(T)} \, dt,
\end{aligned}
$$

where ρ_1, ρ_2 will be eventually sent to 0 as b goes to infinity at the rate $(\log \log b)^{-1}$, $\mathrm{mes}(T)$ is the Lebesgue measure of T and

$$
\begin{aligned}
\mathrm{LR}(t) &= \frac{h_{0,t}(f(t), \partial f(t), \partial^2 f(t))}{h(f(t), \partial f(t), \partial^2 f(t))}, \\
(3.11) \quad \mathrm{LR}_1(t) &= \frac{h_{1,t}(f(t), \partial f(t), \partial^2 f(t))}{h(f(t), \partial f(t), \partial^2 f(t))}, \\
\mathrm{LR}_2(t) &= \frac{1/\sqrt{2\pi} \, e^{-(1/2)(f(t) - u_t)^2}}{(1/\sqrt{2\pi}) e^{-(1/2) f(t)^2}}.
\end{aligned}
$$

The density $h(f(t), \partial f(t), \partial^2 f(t))$ is defined in (3.3), $l(t)$ is a density function on T, $h_{0,t}$ and $h_{1,t}$ are two density functions. Before presenting the specific forms of $l(t)$, $h_{0,t}$ and $h_{1,t}$, we would like to provide an intuitive explanation of dQ/dP from a simulation point of view. One can generate $f(t)$ under the measure Q via the following steps:

(1) Generate $\iota \sim \mathrm{Bernoulli}(\rho_2)$.
(2) If $\iota = 1$, then:
  (a) generate τ uniformly from the index set T, that is, $\tau \sim \mathrm{Unif}(T)$;
  (b) given the realized τ, generate $f(\tau) \sim N(u_\tau, 1)$;
  (c) given $(\tau, f(\tau))$, simulate $\{f(t) : t \neq \tau\}$ from the original conditional distribution under P.
(3) If $\iota = 0$:
  (a) simulate a random variable τ following the density function $l(t)$;
  (b) given the realized τ, simulate $f(\tau) = x$, $\partial f(\tau) = y$, $\partial^2 f(\tau) = z$ from density function

$$
(3.12) \quad h_{\mathrm{all}}(x, y, z) = \frac{1 - \rho_1 - \rho_2}{1 - \rho_2} h_{0,\tau}(x, y, z) + \frac{\rho_1}{1 - \rho_2} h_{1,\tau}(x, y, z);
$$

  (c) given $(\tau, f(\tau), \partial f(\tau), \partial^2 f(\tau))$, simulate $\{f(t) : t \neq \tau\}$ from the original conditional distribution under P.

Thus, τ is a random index at which we twist the distribution of f and its derivatives. The likelihood ratio at a specific location τ is given by $\mathrm{LR}(\tau)$, $\mathrm{LR}_1(\tau)$ or $\mathrm{LR}_2(\tau)$ depending on the mixture component. The distribution of the rest of the field $\{f(t) : t \neq \tau\}$ given $(f(\tau), \partial f(\tau), \partial^2 f(\tau))$ is the same as that under P. It

is not hard to verify that the above simulation procedure is consistent with the Radon–Nikodym derivative in (3.10).

We now provide the specific forms of the functions defining Q. We first consider the situation when $\mu(t) \neq 0$. By condition C6, $\mu(t)$ admits its unique maximum at $t_* = \arg\sup_{t \in T} \mu(t)$ in the interior of T. Furthermore, the Hessian matrix $\Delta \mu_\sigma(t_*)$ is negative definite. The function $l(t)$ is a density on T such that for $t \in T$

$$(3.13) \quad l(t) = (1 + o(1)) \det(-\Delta \mu_\sigma(t_*))^{1/2} \left(\frac{u_{t_*}}{2\pi} \right)^{d/2} e^{(u_{t_*}/2)(t - t_*)^\top \Delta \mu_\sigma(t_*)(t - t_*)},$$

which is approximately a Gaussian density centered around t_*. As $l(t)$ is defined on a compact set t, the $o(1)$ term goes to zero as b tends to infinity. It is introduced to correct for the integral of $l(t)$ outside the region T that is exponentially small and does not affect the current analysis. The functions $h_{0,t}$ and $h_{1,t}$ are density functions on the vector space where $(f(t), \partial f(t), \partial^2 f(t))$ lives, and they are defined as follows (we will explain the following complicated functions momentarily):

$$h_{0,t}(f(t), \partial f(t), \partial^2 f(t))$$
$$= \mathbb{I}_{A_t} \times H_\lambda \times u_t \times e^{-\lambda u_t (f(t) + (\mathbf{1}^\top \bar{f}''_t/(2\sigma u_t)) + B_t/u_t - u_t)} \times e^{-|\partial f(t)|^2/2}$$
$$\times \exp\left\{ -\frac{1}{2} \left[\frac{|\mu_{20} \mu_{22}^{-1} \bar{f}''_t|^2}{1 - \mu_{20} \mu_{22}^{-1} \mu_{02}} + \left| \mu_{22}^{-1/2} \bar{f}''_t - \frac{\mu_{22}^{1/2} \mathbf{1}}{2\sigma} \right|^2 \right] \right\},$$

$$h_{1,t}(f(t), \partial f(t), \partial^2 f(t))$$
$$= \mathbb{I}_{A_t^c} \times H_{\lambda_1} \times u_t \times e^{\lambda_1 u_t (f(t) + (\mathbf{1}^\top \bar{f}''_t/(2\sigma u_t)) + B_t/u_t - u_t)} \times e^{-|\partial f(t)|^2/2}$$
$$\times \exp\left\{ -\frac{1}{2} \left[\frac{|\mu_{20} \mu_{22}^{-1} \bar{f}''_t|^2}{1 - \mu_{20} \mu_{22}^{-1} \mu_{02}} + \left| \mu_{22}^{-1/2} \bar{f}''_t - \frac{\mu_{22}^{1/2} \mathbf{1}}{2\sigma} \right|^2 \right] \right\},$$

where $\mathbb{I}$ is the indicator function, $A_t = \{f(\cdot) : f(t) + \frac{|\partial f(t)|^2}{2u_t} + \frac{\mathbf{1}^\top \bar{f}''_t}{2\sigma u_t} + \frac{B_t}{u_t} > u_t - \eta/u_t\}$ is defined as in (3.6), $\bar{f}''_t$ is defined as in (3.8), $\lambda < 1$ is positive and it will be sent to 1 as b goes to infinity, λ_1 is a fixed positive constant (e.g., $\lambda_1 = 1$) and the normalizing constants are defined as

$$H_\lambda = \frac{e^{-\lambda \eta} (1 - \lambda)^{d/2} \lambda}{(2\pi)^{d/2}}$$

$$(3.14) \quad \times \left[\int_{R^{d(d+1)/2}} e^{-(1/2)[|\mu_{20} \mu_{22}^{-1} z|^2/(1 - \mu_{20} \mu_{22}^{-1} \mu_{02}) + |\mu_{22}^{-1/2} z - \mu_{22}^{1/2} \mathbf{1}/(2\sigma)|^2]} \, dz \right]^{-1},$$

$$H_{\lambda_1} = \frac{e^{\lambda_1 \eta} (1 + \lambda_1)^{d/2} \lambda_1}{(2\pi)^{d/2}}$$

$$\times \left[\int_{R^{d(d+1)/2}} e^{-(1/2)[|\mu_{20} \mu_{22}^{-1} z|^2/(1 - \mu_{20} \mu_{22}^{-1} \mu_{02}) + |\mu_{22}^{-1/2} z - \mu_{22}^{1/2} \mathbf{1}/(2\sigma)|^2]} \, dz \right]^{-1}.$$

The constants H_λ and H_{λ_1} ensure that $h_{0,t}$ and $h_{1,t}$ are properly normalized densities.

Understanding the measure Q. The measure Q is designed such that the distribution of f under the measure Q is approximately the conditional distribution of f given $\mathcal{I}(T) > b$. The two terms corresponding to the probabilities ρ_1 and ρ_2 are included to ensure the absolute continuity and to control the tail of the likelihood ratio. Thus, ρ_1 and ρ_2 will be sent to zero eventually.

We now provide an explanation of the leading term corresponding to the probability $1 - \rho_1 - \rho_2$. To understand $h_{0,t}$, we use the notation α_t in (3.7) and rewrite the density function as

$$h_{0,t}\big(f(t), \partial f(t), \partial^2 f(t)\big)$$

$$\propto \mathbb{I}_{A_t} \exp\{-\lambda u_t(\alpha_t - u_t)\} \times \exp\left\{-\frac{1-\lambda}{2}|\partial f(t)|^2\right\}$$

$$\times \exp\left\{-\frac{1}{2}\left[\frac{|\mu_{20}\mu_{22}^{-1}\bar{f}_t''|^2}{1 - \mu_{20}\mu_{22}^{-1}\mu_{02}} + \left|\mu_{22}^{-1/2}\bar{f}_t'' - \frac{\mu_{22}^{1/2}\mathbf{1}}{2\sigma}\right|^2\right]\right\},$$

which factorizes into three pieces consisting of α_t, $\partial f(t)$ and $\bar{f}_t''$, respectively. We consider the change of variables from $(f(t), \partial f(t), \partial^2 f(t))$ to $(\alpha_t, \partial f(t), \bar{f}_t'')$. Then, under the distribution $h_{0,t}$, the random vectors α_t, $\partial f(t)$ and $\bar{f}_t''$ are independent. Note that $h_{0,t}$ is defined on the set $A_t = \{\alpha_t > u_t - \eta u_t^{-1}\}$ where η will be send to zero eventually. Then, $\alpha_t - u_t$ is approximately an exponential random variable with rate λu_t; $\partial f(t)$, and $\bar{f}_t''$ are two independent Gaussian random vectors. The density $h_{1,t}$ has a similar interpretation. The only difference is that $h_{1,t}$ is defined on the set $\{\alpha_t - u_t < -\eta u_t^{-1}\}$ and $u_t - \alpha_t$ follows approximately an exponential distribution. For the last piece corresponding to ρ_2, the density is simply an exponential tilting of $f(t)$.

Under the dominating mixture component, to generate an $f(t)$ from Q, a random index τ is first sampled from T following density $l(t)$, then $(f(\tau), \partial f(\tau), \partial^2 f(\tau))$ is sampled according to $h_{0,\tau}$. This implies that the large value of the integral $\int_T e^{\mu(t)+\sigma f(t)}\, dt$ is mostly caused by the fact that the field reaches a high level at τ; more precisely, α_τ reaches a high level of u_τ (with an exponential overshoot of rate λu_τ). Therefore, the random index τ localizes the position where the field α_t goes very high. The distribution of τ given as in (3.13) is very concentrated around t_*. This suggests that the maximum of α_t [or $f(t)$] is attained within $O_p(u^{-1/2})$ distance from t_*.

We now consider the case where $\mu(t) \equiv 0$. We choose $l(t)$ to be the uniform distribution over set T and have that

$$(3.15) \quad \begin{aligned} \frac{dQ}{dP} &= (1 - \rho_1 - \rho_2)\int_T \frac{\text{LR}(t)}{\text{mes}(T)}\, dt + \rho_1 \int_T \frac{\text{LR}_1(t)}{\text{mes}(T)}\, dt \\ &\quad + \rho_2 \int_T \frac{\text{LR}_2(t)}{\text{mes}(T)}\, dt, \end{aligned}$$

where $\mathrm{mes}(\cdot)$ is the Lebesgue measure. The following theorem states that Q is a good approximation of P_b^* with appropriate choice of the tuning parameters.

THEOREM 3. *Consider a Gaussian random field* $\{f(t) : t \in T\}$ *living on a domain* T *satisfying conditions* C1–C6. *If we choose the parameters defining the change of measure* $\eta = \rho_1 = \rho_2 = 1 - \lambda = (\log \log b)^{-1}$, *then we have the following approximation*:

$$\lim_{b \to \infty} \sup_{A \in \mathcal{F}} |Q(A) - P_b^*(A)| = 0,$$

where $\mathcal{F}$ *is the* σ*-field where the measures are defined.*

REMARK 4. Theorem 3 is the central result of this paper. We present its detailed proof. The technical developments of other theorems are all based on that of Theorem 3. Therefore, we only layout their key steps and the major differences from that of Theorem 3.

REMARK 5. The measure Q in the limit of the above theorem obviously depends on the tuning parameters (η, ρ_1, ρ_2, and λ) and the level b. To simplify the notation, we omit the indices of those parameters when there is no ambiguity.

REMARK 6. The measure corresponding to the last mixture component in (3.10), $\int_T \frac{\mathrm{LR}_2(t)}{\mathrm{mes}(T)} \, dt$, has been employed by [43] to develop approximations for $v(b)$. We emphasize that the measure constructed in this paper is substantially different. In fact, the measure corresponding to $\mathrm{LR}_2(t)$ does not appear in the main proof. We included it to control the tail of the likelihood ratio in one lemma.

To illustrate the application of the measure Q, we provide a further characterization of the conditional distribution P_b^* by presenting another approximation result which is easier to understand at an intuitive level. Let

$$\gamma_u(t) = f(t) + \frac{\mathbf{1}^\top \bar{f}_t''}{2\sigma u_t} + \frac{B_t}{u_t} + \mu_\sigma(t), \qquad \beta_u(T) = \sup_{t \in T} \gamma_u(t),$$

(3.16)
$$\tilde{P}_b(f(\cdot) \in A) = P(f(\cdot) \in A | \beta_u(T) > u).$$

The process $\gamma_u(t)$ is slightly different than α_t. The following theorem states that the measure Q also approximates the distribution $\tilde{P}_b$ in total variation for b large.

THEOREM 7. *Consider a Gaussian random field* $\{f(t) : t \in T\}$ *living on a domain* T *satisfying conditions* C1–C6. *With the same choice of tuning parameters as in Theorem 3, that is,* $\eta = \rho_1 = \rho_2 = 1 - \lambda = (\log \log b)^{-1}$, Q *approximates* $\tilde{P}_b$ *in total variation, that is,*

$$\lim_{b \to \infty} \sup_{A \in \mathcal{F}} |Q(A) - \tilde{P}_b(A)| = 0.$$

3.3. *Some implications of the theorems.* The results of Theorems 3 and 7 provide both qualitative and quantitative descriptions of P_b^*. From a qualitative point of view, Theorems 3 and 7 suggest that

$$(3.17) \qquad \sup_{A \in \mathcal{F}} |P_b^*(A) - \tilde{P}_b(A)| \to 0$$

as $b \to \infty$. Note that $\gamma_u(t)$ itself is a Gaussian process. Thus, the above convergence result connects the tail events of exponential integrals to those of the supremum of another Gaussian random field that is a linear combination of f and its derivative field. We set up this connection mainly because the distribution of Gaussian random fields conditional on level crossing (also known as the Slepian model) is very well studied for smooth processes [32]. For the purpose of illustration, we cite one result in Chapter 6.2 of [6] when $\gamma_u(t)$ is stationary and twice differentiable. Let covariance function of $\gamma_u(t)$ be $C_\gamma(t)$. Conditional on $\gamma_u(t)$ achieving a local maximum at location t^* at level x, we have the following closed form representation of the conditional field:

$$(3.18) \qquad \gamma_u(t^* + t) = x C_\gamma(t) - W_x \beta(t) + g(t),$$

where

$$\beta(t) = \begin{pmatrix} 1 & \mu_{20}^\gamma \\ \mu_{02}^\gamma & \mu_{22}^\gamma \end{pmatrix}^{-1} \mu_2^{\gamma \top}(t),$$

μ_{ij}^γ's are the spectral moments of $C_\gamma(t)$, W_x is a $d(d+1)/2$ dimensional random vector whose density can be explicitly written down and $g(t)$ is a mean zero Gaussian process whose covariance function is also in a closed form; see [6] for the specific forms. If we set $x > u \to \infty$, the local maximum is asymptotically the global maximum. Furthermore, thanks to stationarity, the distribution of t^* is asymptotically uniform over T. The overshoot $x - u$ is asymptotically an exponential random variable. Thus, the conditional field $\gamma_u(t)$ can be written down explicitly through representation (3.18), the overshoot distribution and the distribution of t^*. Furthermore, the conditional distribution of $f(t)$ can be implied by (3.16) and conditional normal calculations.

From a quantitative point of view, Theorem 3 implies that for any bounded function $\Xi : C(T) \to R$ the conditional expectation $E[\Xi(f)|\mathcal{I}(T) > b]$ can be approximated by $E^Q[\Xi(f)]$, more precisely,

$$(3.19) \qquad E[\Xi(f)|\mathcal{I}(T) > b] - E^Q[\Xi(f)] \to 0$$

as $b \to \infty$. The expectation $E^Q[\Xi(f)]$ is much easier to compute (both analytically and numerically) via the following identity:

$$(3.20) \qquad E^Q[\Xi(f)] = E^Q[E[\Xi(f)|\iota, \tau, f(\tau), \partial f(\tau), \partial^2 f(\tau)]].$$

Note that the inner expectation is under the measure P in that the conditional distribution of f given $(f(\tau), \partial f(\tau), \partial^2 f(\tau))$ under Q is the same as that under P. Furthermore, conditional on $(f(\tau), \partial f(\tau), \partial^2 f(\tau))$, the process $f(t)$ is also a Gaussian process and has the expansion

$$f(t) = f(\tau) + \partial f(\tau)^\top (t - \tau) + \tfrac{1}{2}(t - \tau)^\top \Delta f(\tau)(t - \tau) + o(|t - \tau|^2).$$

These results provide sufficient tools to evaluate the conditional expectation

$$E\big[\Xi(f)|\iota, \tau, f(\tau), \partial f(\tau), \partial^2 f(\tau)\big].$$

Once the above expectation has been evaluated, we may proceed to the outer expectation in (3.20). Note that the inner expectation is a function of $(\iota, \tau, f(\tau), \partial f(\tau), \partial^2 f(\tau))$, the joint distribution of which is in a closed form. Thus, evaluating the outer expectation is usually an easier task. In fact, the proof of Theorem 3 is an exercise of the above strategy by considering that $\Xi(f) = (dP/dQ)^2$.

REMARK 8.   According to the detailed proof of Theorem 3, the approximation (3.19) is applicable to all the functions such that $\sup_b E[\Xi^2(f)|\mathcal{I}(T) > b] < \infty$. To see that, we need to change the statement and the proof of Lemma 13 presented in Section 4.

3.4. *Efficient rare-event simulation for $\mathcal{I}(T)$.*   In the preceding subsection we constructed a change of measure that asymptotically approximates the conditional distribution of f given $\mathcal{I}(T) > b$. In this section, we construct an efficient importance sampling estimator based on this change of measure to compute $v(b)$ as $b \to \infty$. We evaluate the overall computation efficiency using a concept that has its root in the general theory of computation in both continuous and discrete settings [47, 54]. In particular, completely analogous notions in the setting of complexity theory of continuous problems lead to the notion of tractability of a computational problem [55].

DEFINITION 9.   A Monte Carlo estimator is said to be a fully polynomial randomized approximation scheme (FPRAS) for estimating $v(b)$ if, for some q_1, q_2 and $d > 0$, it outputs an averaged estimator that is guaranteed to have at most $\varepsilon > 0$ relative error with confidence at least $1 - \delta \in (0, 1)$ in $O(\varepsilon^{-q_1}\delta^{-q_2}|\log v(b)|^d)$ function evaluations.

Equivalently, one needs to compute an estimator Z_b with complexity $O(\varepsilon^{-q_1}\delta^{-q_2}|\log v(b)|^d)$ such that

$$(3.21) \qquad\qquad P\big(|Z_b/v(b) - 1| > \varepsilon\big) < \delta.$$

In the literature of rare-event simulations, an estimator L_b is said to be *strongly efficient* in estimating $v(b)$ if $EL_b = v(b)$ and $\sup_b \operatorname{Var} L_b/v^2(b) < \infty$. Suppose

that a strongly efficient estimator L_b has been obtained. Let $\{L_b^{(j)} : j = 1, \ldots, n\}$ be i.i.d. copies of L_b. The averaged estimator

$$Z_b = \frac{1}{n} \sum_{j=1}^{n} L_b^{(j)}$$

has a *relative* mean squared error equal to $\sqrt{E(Z_b/v(b) - 1)^2} = \sqrt{\mathrm{Var}(L_b)} \times n^{-1/2} v(b)^{-1}$. A simple consequence of Chebyshev's inequlity yields

$$P\big(|Z_b/v(b) - 1| \geq \varepsilon\big) \leq \frac{\mathrm{Var}(L_b)}{\varepsilon^2 n v^2(b)}.$$

Thus, it suffices to simulate $n = O(\varepsilon^{-2}\delta^{-1})$ i.i.d. replicates of L_b to achieve the accuracy in (3.21).

The so-called importance sampling is based on the identity $P(A) = E^Q[\mathbb{I}_A \, dP/dQ]$. The random variable $\mathbb{I}_A \, dP/dQ$ is an unbiased estimator of $P(A)$. It is well known that if one chooses $Q(\cdot) = P(\cdot|A)$, then $\mathbb{I}_A \, dP/dQ$ has zero variance. The measure Q created in the previous subsection is a good approximation of P_b^*, and thus it naturally leads an estimator for $v(b)$ with small variance.

In addition to the variance control, another issue is that the random fields considered in this paper are continuous objects. A computer can only perform discrete simulations. Thus we must use a discrete object approximating the continuous field to implement the algorithms. The bias caused by the discretization must be well controlled relative to $v(b)$. In addition, the complexity of generating one such discrete object should also be considered in order to control the overall computational complexity to achieve an FPRAS.

We create a regular lattice covering T. Define

$$G_{N,d} = \left\{ \left(\frac{i_1}{N}, \frac{i_2}{N}, \ldots, \frac{i_d}{N} \right) : i_1, \ldots, i_d \in \mathbb{Z} \right\}.$$

For each $t = (t^1, \ldots, t^d) \in G_{N,d}$, define

$$T_N(t) = \{(s^1, \ldots, s^d) \in T : s^j \in (t^j - 1/N, t^j] \text{ for } j = 1, \ldots, d\}$$

that is, the $\frac{1}{N}$-cube intersected with T and cornered at t. Furthermore, let

(3.22) $T_N = \{t \in G_{N,d} : T_N(t) \neq \varnothing\}.$

Since T is compact, T_N is a finite set. We enumerate the elements in $T_N = \{t_1, \ldots, t_M\}$, where $M = O(N^d)$. We further define

$$X = (X_1, \ldots, X_M)^\top \triangleq \big(f(t_1), \ldots, f(t_M)\big)^\top$$

and use

$$v_M(b) = P\big(\mathcal{I}_M(T) > b\big)$$

as an approximation of $v(b)$ where

$$(3.23) \qquad \mathcal{I}_M(T) = \sum_{i=1}^{M} \text{mes}(T_N(t_i)) \times e^{\sigma X_i + \mu(t_i)}.$$

We have the following theorem to control the bias.

THEOREM 10. *Consider a Gaussian random field f satisfying conditions in Theorem* 3. *For any $\varepsilon_0 > 0$, there exists κ_0 such that for any $\varepsilon \in (0, 1)$, if $N \geq \kappa_0 \varepsilon^{-1-\varepsilon_0} (\log b)^{2+\varepsilon_0}$, then for $b > 2$*

$$\frac{|v_M(b) - v(b)|}{v(b)} < \varepsilon.$$

We estimate $v_M(b)$ using a discrete version of the change of measure proposed in the previous section. The specific algorithm is given as follows:

(1) Generate a random indicator $\iota \sim \text{Bernoulli}(\rho_2)$. If $\iota = 1$, then:
  (a) generate ι uniformly from $\{1, \ldots, M\}$;
  (b) generate $X_\iota \sim N(u_{t_\iota}, 1)$;
  (c) given (t_ι, X_ι), simulate the joint field $(f(t), \partial f(t), \partial^2 f(t))$ on the lattice $T_N \setminus \{t_\iota\}$ from the original conditional distribution under P.
(2) If $\iota = 0$:
  (a) if $\mu(t)$ is not constant, simulate a random index ι proportional to $l(t_\iota)$, that is, $P(\iota = i) = l(t_i)/\kappa$ and $\kappa = \sum_{i=1}^{M} l(t_i)$; if $\mu(t) \equiv 0$, then ι is simulated uniformly over $\{1, \ldots, M\}$;
  (b) given the realized ι, simulate $f(t_\iota) = X_\iota = x$, $\partial f(t_\iota) = y$, $\partial^2 f(t_\iota) = z$ from density function

$$h_{\text{all}}(x, y, z) = \frac{1 - \rho_1 - \rho_2}{1 - \rho_2} h_{0, t_\iota}(x, y, z) + \frac{\rho_1}{1 - \rho_2} h_{1, t_\iota}(x, y, z);$$

  (c) given $(t_\iota, f(t_\iota), \partial f(t_\iota), \partial^2 f(t_\iota))$, simulate the joint field $(f(t), \partial f(t), \partial^2 f(t))$ on the lattice $T_N \setminus \{t_\iota\}$ from the original conditional distribution under P.
(3) Output

$$(3.24) \qquad \tilde{L}_b = \mathbb{I}_{\{\mathcal{I}_M(T) > b\}} \bigg/ \bigg(\frac{1 - \rho_1 - \rho_2}{\kappa} \sum_{i=1}^{M} l(t_i) \, \text{LR}(t_i) + \frac{\rho_1}{\kappa} \sum_{i=1}^{M} l(t_i) \, \text{LR}_1(t_i)$$

$$+ \rho_2 \sum_{i=1}^{M} \frac{\text{LR}_2(t_i)}{M} \bigg).$$

Let Q_M be the measure induced by the above simulation scheme. Then it is not hard to verify that $\tilde{L}_b = \mathbb{I}_{\{\mathcal{I}_M(T) > b\}} \, dP/dQ_M$, and thus $\tilde{L}_b$ is an unbiased estimator of $v_M(b)$. The next theorem states the strong efficiency of the above algorithm.

THEOREM 11. *Suppose f is a Gaussian random field satisfying conditions in Theorem 3. If N is chosen as in Theorem 10 and all the other parameters are chosen as in Theorem 3, then there exists some constant $\kappa_1 > 0$ such that*

$$\sup_{b>1} \frac{E^{Q_M} \tilde{L}_b^2}{v_M^2(b)} \leq \kappa_1.$$

Let Z_b be the average of n i.i.d. copies of $\tilde{L}_b$. According to the results in Theorem 10, we have that

$$\left| \frac{Z_b}{v(b)} - 1 \right| \leq \left| \frac{Z_b}{v_M(b)} (v_M(b)/v(b) - 1) \right| + \left| \frac{Z_b}{v_M(b)} - 1 \right|$$

$$\leq \varepsilon \left| \frac{Z_b}{v_M(b)} \right| + \left| \frac{Z_b}{v_M(b)} - 1 \right|.$$

The results of Theorem 11 indicate that

$$P\big(\big|Z_b/v_M(b) - 1\big| \geq \varepsilon\big) \leq \frac{\kappa_1}{\varepsilon^2 n}.$$

If we choose $n = \kappa_1 \varepsilon^{-2} \delta^{-1}$, then

$$P\big(\big|Z_b/v(b) - 1\big| \geq 3\varepsilon\big) \leq \delta.$$

Thus, the accuracy level as in (3.21) has been achieved. Note that simulating one $\tilde{L}_b$ consists of generating a multivariate Gaussian random vector of dimension $M \times (d+1)(d+2)/2 = O(N^d) = O((\log b)^{(2+\varepsilon_0)d} \varepsilon^{-(1+\varepsilon_0)d})$. The complexity of generating such a vector is at the most $O(N^3)$. Thus the overall complexity is $O(\varepsilon^{-2-(3+3\varepsilon_0)d} \delta^{-1} (\log b)^{(6+3\varepsilon_0)d})$. The proposed estimator in (3.24) is a FPRAS.

REMARK 12. The proposed algorithm can also be used to compute conditional expectations via the representation $E[\Xi(f)|\mathcal{I}(T) > b] = E[\Xi(f); \mathcal{I}(T) > b]/v(b)$, where $E[\Xi(f); \mathcal{I}(T) > b]$ can be estimated by $\Xi(f) \, dP/dQ_M$ and $v(b)$ can be estimated by $\mathbb{I}_{\{\mathcal{I}(T)>b\}} \, dP/dQ_M$.

4. Proof of Theorem 3. We use the following simple yet powerful lemma to prove Theorem 3.

LEMMA 13. *Let Q_0 and Q_1 be probability measures defined on the same σ-field $\mathcal{F}$ such that $dQ_1 = r^{-1} dQ_0$ for a positive random variable r. Suppose that for some $\varepsilon > 0$, $E^{Q_1}[r^2] = E^{Q_0}[r] \leq 1 + \varepsilon$. Then*

$$\sup_{|X| \leq 1} \big|E^{Q_1}(X) - E^{Q_0}(X)\big| \leq \varepsilon^{1/2}.$$

PROOF.

$$\begin{aligned}
\big|E^{Q_1}(X) - E^{Q_0}(X)\big| &= \big|E^{Q_1}[(1-r)X]\big| \\
&\le E^{Q_1}|r-1| \le \big[E^{Q_1}(r-1)^2\big]^{1/2} \\
&= \big(E^{Q_1}[r^2]-1\big)^{1/2} \le \varepsilon^{1/2}. \qquad \square
\end{aligned}$$

We also need the following approximations for the tail probability $v(b)$. This proposition is an extension of Theorem 3.4 and Corollary 3.5 in [43]. We layout the key steps of its proof in the supplemental material [45].

PROPOSITION 14. *Consider a Gaussian random field $\{f(t) : t \in T\}$ living on a domain T satisfying conditions* C1–C6. *If $\mu(t)$ has one unique maximum in T denoted by t_*, then*

$$v(b) \sim (2\pi)^{d/2} \det\big(-\Delta\mu_\sigma(t_*)\big)^{-1/2} G(t_*) \cdot u^{d/2-1} \exp\left\{-\frac{(u-\mu_\sigma(t_*))^2}{2}\right\},$$

where u is as defined in (3.4), *and $G(t)$ is defined as*

$$\frac{\det(\Gamma)^{-1/2}}{(2\pi)^{(d+1)(d+2)/4}} e^{\mathbf{1}^T \mu_{22}\mathbf{1}/(8\sigma^2)+B_t}$$

$$\times \int_{R^{d(d+1)/2}} \exp\left\{-\frac{1}{2}\left[\frac{|\mu_{20}\mu_{22}^{-1}z|^2}{1-\mu_{20}\mu_{22}^{-1}\mu_{02}} + \left|\mu_{22}^{-1/2}z - \frac{\mu_{22}^{1/2}\mathbf{1}}{2\sigma}\right|^2\right]\right\} dz.$$

If $\mu(t) \equiv 0$, $G(t)$ is a constant denoted by G. Then

$$v(b) \sim \operatorname{mes}(T)G \cdot u^{d-1} e^{-u^2/2}.$$

4.1. *Case* 1: *$\mu(t)$ is not a constant.* To make the proof smooth, we arrange the statement of the rest supporting lemmas in the Appendix. We start the proof of Theorem 3 when $\mu(t)$ is not a constant. Note that

$$E^Q\left[\left(\frac{dP_b^*}{dQ}\right)^2\right] = v(b)^{-2} E^Q\left[\left(\frac{dP}{dQ}\right)^2; \mathcal{I}(T) > b\right].$$

Thanks to Lemma 13, we only need to show that for any $\varepsilon > 0$ there exists b_0 such that for all $b > b_0$

$$E^Q\left[\left(\frac{dP}{dQ}\right)^2; \mathcal{I}(T) > b\right] = E^Q\left[E_{\iota,\tau}^Q\left[\left(\frac{dP}{dQ}\right)^2; \mathcal{I}(T) > b\right]\right] \le (1+\varepsilon)v(b)^2,$$

where we use the notation $E_{\iota,\tau}^Q[\cdot] = E^Q[\cdot|\iota,\tau]$ to denote the conditional expectation given ι and τ. $\tau \in T$ is the random index described as in the simulation scheme

admitting a density function $l(t)$ if $\iota = 0$ and $\mathrm{mes}^{-1}(T)\mathbb{I}_T(t)$ if $\iota = 1$. Note that

$$
E^Q_{\iota,\tau}\left[\left(\frac{dP}{dQ}\right)^2; \mathcal{I}(T) > b\right]
$$

$$
= E^Q_{\iota,\tau}\left[E^Q_{\iota,\tau}\left[\left(\frac{dP}{dQ}\right)^2; \mathcal{I}(T) > b \Big| f(\tau), \partial f(\tau), \partial^2 f(\tau)\right]\right].
$$

For the rest of the proof, we mostly focus on the conditional expectation

$$
E^Q_{\iota,\tau}\left[\left(\frac{dP}{dQ}\right)^2; \mathcal{I}(T) > b \Big| f(\tau), \partial f(\tau), \partial^2 f(\tau)\right].
$$

The rest of the discussion is conditional on ι and τ. To simplify notation, for a given τ, we define

$$
f_*(t) = f(t) - u_\tau C(t - \tau).
$$

On the set $\{\mathcal{I}(T) > b\}$, $f(\tau)$ reaches a level u_τ, and $E[f(t)|f(\tau) = u_\tau] = u_\tau C(t - \tau)$. Thus, $f_*(t)$ is the field with the conditional expectation removed. From now on, we work with this shifted field $f_*(t)$. Correspondingly, we have

$$
\partial f_*(t) = \partial f(t) - u_\tau \partial C(t - \tau), \qquad \partial^2 f_*(t) = \partial^2 f(t) - u_\tau \partial^2 C(t - \tau).
$$

We further define the following notation:

$$
w = f_*(\tau), \qquad y = \partial f_*(\tau), \qquad z = \partial^2 f_*(\tau), \qquad \mathbf{z} = \Delta f_*(\tau),
$$

$$
(4.1) \quad \tilde{y} = \partial f_*(\tau) + \partial \mu_\sigma(\tau), \qquad \tilde{\mathbf{z}} = \Delta f_*(\tau) + \mu_\sigma(\tau)I + \Delta \mu_\sigma(\tau),
$$

$$
w_t = f_*(t), \qquad y_t = \partial f_*(t), \qquad z_t = \partial^2 f_*(t), \qquad \bar{z}_t = \partial^2 f_*(t) - u_t \mu_{02}.
$$

Under the measure Q and a given τ, if $\iota = 0$, (w, y, z) has density function

$$
(4.2) \quad h^*_{\mathrm{all}}(w, y, z) = \frac{1 - \rho_1 - \rho_2}{1 - \rho_2} h^*_{0,\tau}(w, y, z) + \frac{\rho_1}{1 - \rho_2} h^*_{1,\tau}(w, y, z);
$$

if $\iota = 1$, then (w, y, z) follows density $h^*_\tau(w, y, z)$. The forms of the densities can be derived from $h_{0,t}$, $h_{1,t}$ and h. In particular, their expressions are given as follows:

$$
h^*_{0,\tau}(w, y, z) \propto \mathbb{I}_{A_\tau} \times \exp\left\{-\lambda u_\tau\left(w + \frac{\mathbf{1}^\top z}{2\sigma u_\tau} + \frac{B_\tau}{u_\tau}\right) - \frac{1}{2}|y|^2\right\}
$$

$$
\times \exp\left\{-\frac{1}{2}\left[\frac{|\mu_{20}\mu_{22}^{-1}z|^2}{1 - \mu_{20}\mu_{22}^{-1}\mu_{02}} + \left|\mu_{22}^{-1/2}z - \frac{\mu_{22}^{1/2}\mathbf{1}}{2\sigma}\right|^2\right]\right\},
$$

$$
h^*_{1,\tau}(w, y, z) \propto \mathbb{I}_{A^c_\tau} \times \exp\left\{\lambda_1 u_\tau\left(w + \frac{\mathbf{1}^\top z}{2\sigma u_\tau} + \frac{B_\tau}{u_\tau}\right) - \frac{1}{2}|y|^2\right\}
$$

$$
\times \exp\left\{-\frac{1}{2}\left[\frac{|\mu_{20}\mu_{22}^{-1}z|^2}{1 - \mu_{20}\mu_{22}^{-1}\mu_{02}} + \left|\mu_{22}^{-1/2}z - \frac{\mu_{22}^{1/2}\mathbf{1}}{2\sigma}\right|^2\right]\right\},
$$

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

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

and $A_\tau = \{w + \frac{y^\top y}{2u_\tau} + \frac{\mathbf{1}^\top z}{2\sigma u_\tau} + \frac{B_\tau}{u_\tau} > -\eta u_\tau^{-1}\}$ is defined as in (3.6).

In the next step, we will compute dQ/dP in the form of $f_*(t)$. Basically, we replace $f(t)$ by $f_*(t) + u_\tau C(t - \tau)$, $\partial f(t)$ by $y_t + u_\tau \partial C(t - \tau)$, $\partial^2 f(t)$ by $z_t + u_\tau \partial^2 C(t - \tau)$ and $\bar{f}_t'' = \partial^2 f(t) - u_t\mu_{02}$ by $\bar{z}_t + u_\tau \partial^2 C(t - \tau)$. For the likelihood ratio terms LR and LR_1 in (3.11), note that the $|\partial f(t)|^2$ terms in $h_{0,t}$ and $h_{1,t}$ cancel with those in $h(f(t), \partial f(t), \partial^2 f(t))$, that is,

$$LR(t) = \mathbb{I}_{A_t} \cdot H_\lambda \cdot u_t \exp\left\{-\lambda u_t\left(f(t) + \frac{\mathbf{1}^\top \bar{f}_t''}{2\sigma u_t} + \frac{B_t}{u_t} - u_t\right)\right.$$

$$\left. -\frac{1}{2}\left[\frac{|\mu_{20}\mu_{22}^{-1}\bar{f}_t''|^2}{1 - \mu_{20}\mu_{22}^{-1}\mu_{02}} + \left|\mu_{22}^{-1/2}\bar{f}_t'' - \frac{\mu_{22}^{1/2}\mathbf{1}}{2\sigma}\right|^2\right]\right\}$$

$$\bigg/ \left(\frac{\det(\Gamma)^{-1/2}}{(2\pi)^{(d+1)(d+2)/4}}\right.$$

$$\left. \times e^{-(1/2)[(f(t) - \mu_{20}\mu_{22}^{-1}\partial^2 f(t))^2/(1 - \mu_{20}\mu_{22}^{-1}\mu_{02}) + \partial^2 f(t)^\top \mu_{22}^{-1}\partial^2 f(t)]}\right).$$

We insert the notation in (4.1) and obtain that

$$LR(t) = \mathbb{I}_{A_t} \cdot u_t H_\lambda \exp\left\{-\lambda u_t\left(w_t + u_\tau C(t - \tau) + \frac{\mathbf{1}^\top(\bar{z}_t + \mu_2(t - \tau)u_\tau)}{2\sigma u_t}\right.\right.$$

$$\left.\left. + \frac{B_t}{u_t} - u_t\right)\right\}$$

$$(4.3) \qquad \times \exp\left\{-\frac{1}{2}\left[\frac{|\mu_{20}\mu_{22}^{-1}(\bar{z}_t + \mu_2(t - \tau)u_\tau)|^2}{1 - \mu_{20}\mu_{22}^{-1}\mu_{02}}\right.\right.$$

$$\left.\left. + \left|\mu_{22}^{-1/2}(\bar{z}_t + \mu_2(t - \tau)u_\tau) - \frac{\mu_{22}^{1/2}\mathbf{1}}{2\sigma}\right|^2\right]\right\}$$

$$\times h_{x,z}^{-1}(w_t + u_\tau C(t - \tau), z_t + u_\tau \partial^2 C(t - \tau)),$$

where

$$(4.4) \quad h_{x,z}(x, z) = \frac{\det(\Gamma)^{-1/2}}{(2\pi)^{(d+1)(d+2)/4}} e^{-(1/2)[(x - \mu_{20}\mu_{22}^{-1}z)^2/(1 - \mu_{20}\mu_{22}^{-1}\mu_{02}) + z^\top \mu_{22}^{-1}z]},$$

which is the function $h(x, y, z)$ with the $|y|^2$ term removed. Similarly, we have that

$$\mathrm{LR}_1(t) = \mathbb{I}_{A_t^c} \cdot u_t H_{\lambda_1} \exp\left\{ \lambda_1 u_t \left(w_t + u_\tau C(t - \tau) + \frac{\mathbf{1}^\top(\bar{z}_t + \mu_2(t - \tau)u_\tau)}{2\sigma u_t} \right.\right.$$

$$\left.\left. + \frac{B_t}{u_t} - u_t \right) \right\}$$

$$(4.5) \qquad \times \exp\left\{ -\frac{1}{2}\left[\frac{|\mu_{20}\mu_{22}^{-1}(\bar{z}_t + \mu_2(t - \tau)u_\tau)|^2}{1 - \mu_{20}\mu_{22}^{-1}\mu_{02}} \right.\right.$$

$$\left.\left. + \left| \mu_{22}^{-1/2}(\bar{z}_t + \mu_2(t - \tau)u_\tau) - \frac{\mu_{22}^{1/2}\mathbf{1}}{2\sigma} \right|^2 \right] \right\}$$

$$\times h_{x,z}^{-1}(w_t + u_\tau C(t - \tau), z_t + u_\tau \partial^2 C(t - \tau)).$$

With the analytic forms (4.3) and (4.5), we proceed to the likelihood ratio in (3.10)

$$(4.6) \qquad \frac{dQ}{dP} = (1 - \rho_1 - \rho_2)K + \rho_1 K_1 + \rho_2 K_2,$$

where

$$K = \int_{A^*} l(t)\,\mathrm{LR}(t)\,dt, \qquad K_1 = \int_{(A^*)^c} l(t)\,\mathrm{LR}_1(t)\,dt,$$

$$K_2 = \int_T \frac{e^{-(1/2)u_t^2 + u_t w_t + u_t u_\tau C(t - \tau)}}{\mathrm{mes}(T)}\,dt.$$

The set A^* [depending on the sample path $f_*(t)$] is defined as

$$\left\{ t : w_t + C(t - \tau)u_\tau + \frac{|y_t + u_\tau \cdot \partial C(t - \tau)|^2}{2u_t} + \frac{\mathbf{1}^\top(\bar{z}_t + u_\tau \mu_2(t - \tau))}{2\sigma u_t} + \frac{B_t}{u_t} \right.$$

$$\left. > u_t - \frac{\eta}{u_t} \right\}.$$

We may equivalently define $A^* = \{t : f \in A_t\}$. Note that $\mathrm{LR}(t) = 0$ if $f \notin A_t$. Thus, the integral K is on the set A^*, and K_1 is on the complement of A^*.

Based on the above results, we have that

$$E^Q\left[\left(\frac{dP}{dQ}\right)^2; \mathcal{I}(T) > b \right]$$

$$(4.7) \qquad \leq E^Q\left\{ E_{t,\tau}^Q\left[\frac{1}{[(1 - \rho_1 - \rho_2)K + \rho_1 K_1]^2}; \mathcal{I}(T) > b \right] \right\}$$

$$\leq E^Q\left\{ E_{t,\tau}^Q\left[\frac{1}{[(1 - \rho_1 - \rho_2)K]^2}; \mathcal{I}(T) > b, \mathcal{A}_\tau \geq 0 \right] \right\}$$

$$+ E^Q\left\{ E_{t,\tau}^Q\left[\frac{1}{[(1 - \rho_1 - \rho_2)K + \rho_1 K_1]^2}; \mathcal{I}(T) > b, \mathcal{A}_\tau < 0 \right] \right\},$$

where

$$(4.8) \qquad \mathcal{A}_\tau = w + \frac{y^\top y}{2u_\tau} + \frac{\mathbf{1}^\top z}{2\sigma u_\tau} + \frac{B_\tau}{u_\tau}.$$

Note that the term K_2 is not used in the main analysis. In fact, K_2 is only used in Lemma 17 for the purpose of localization that will be presented later. The rest of the analysis consists of three main parts.

Part 1. Conditional on $(\iota, \tau, f_*(\tau), \partial f_*(\tau), \partial^2 f_*(\tau))$, we study the event

$$(4.9) \qquad \mathcal{E}_b = \{\mathcal{I}(T) > b\},$$

and write the occurrence of this event almost as a deterministic function of $f_*(\tau)$, $\partial f_*(\tau)$ and $\partial^2 f_*(\tau)$, equivalently, (w, y, z).

Part 2. Conditional on $(\iota, \tau, f_*(\tau), \partial f_*(\tau), \partial^2 f_*(\tau))$, we express K and K_1 as functions of $f_*(\tau)$, $\partial f_*(\tau)$, $\partial^2 f_*(\tau)$ with small correction terms.

Part 3. We combine the results from the first two parts and obtain an approximation of (4.7).

All the subsequent derivations are conditional on ι and τ.

4.1.1. *Preliminary calculations.* To proceed, we provide the Taylor expansions for $f_*(t)$, $C(t)$ and $\mu(t)$:

- Expansion of $f_*(t)$ given $(f_*(\tau), \partial f_*(\tau), \partial^2 f_*(\tau))$. Let $t - \tau = ((t-\tau)_1, \ldots, (t-\tau)_d)$. Conditional on $(f_*(\tau), \partial f_*(\tau), \partial^2 f_*(\tau))$, we first expand the random function

$$f_*(t) = E[f_*(t)|f_*(\tau), \partial f_*(\tau), \partial^2 f_*(\tau)] + g(t-\tau)$$

$$(4.10) \qquad = f_*(\tau) + \partial f_*(\tau)^\top (t-\tau) + \tfrac{1}{2}(t-\tau)^\top \Delta f_*(\tau)(t-\tau)$$

$$+ R_f(t-\tau) + g(t-\tau),$$

where

$$R_f(t-\tau) = O\big(|t|^{2+\delta_0}\big(|f_*(\tau)| + |\partial f_*(\tau)| + |\partial^2 f_*(\tau)|\big)\big)$$

is the remainder term of the Taylor expansion of $E[f_*(t)|f_*(\tau), \partial f_*(\tau), \partial^2 f_*(\tau)]$. $g(t)$ is a mean zero Gaussian random field such that $Eg^2(t) = O(|t|^{4+\delta_0})$ as $t \to 0$. In addition, the distribution of $g(t)$ is independent of ι, τ, $f_*(\tau)$, $\partial f_*(\tau)$ and $\partial^2 f_*(\tau)$.

- Expansion of $C(t)$:

$$(4.11) \qquad C(t) = 1 - \tfrac{1}{2}t^\top t + C_4(t) + R_C(t),$$

where $C_4(t) = \frac{1}{24}\sum_{ijkl}\partial^4_{ijkl}C(0)t_i t_j t_k t_l$ and $R_C(t) = O(|t|^{4+\delta_0})$.

- Expansion of $\mu(t)$:

$$(4.12) \quad \mu(t) = \mu(\tau) + \partial\mu(\tau)^\top (t-\tau) + \tfrac{1}{2}(t-\tau)^\top \Delta\mu(\tau)(t-\tau) + R_\mu(t-\tau),$$

where $R_\mu(t-\tau) = O(|t-\tau|^{2+\delta_0})$.

We write

$$R(t) = R_f(t) + u_\tau R_C(t) + R_\mu(t)/\sigma$$

to denote all the remainder terms.

Choose small constants ϵ and δ such that $0 < \epsilon \ll \delta \ll \delta_0$. By writing

$$x \ll y,$$

we mean that x/y is chosen sufficiently small, but x/y does not change with b. Let

(4.13)
$$\mathcal{L} = \Big\{ |\tau - t_*| < u^{-1/2+\epsilon}, |w| \le u^{1/2+\epsilon}, |y| < u^\epsilon, |z| < u^\epsilon,$$
$$\sup_{|t-\tau|<u^{-1+\delta}} |z_t - z| < u^{-\epsilon}, \quad \sup_{|t-\tau|<u^{-1+\delta}} |g(t)| < u^{-1-\delta} \Big\}.$$

By Lemma 17 whose proof uses the last component $LR_2(t)$, we have that

$$E^Q\left[\left(\frac{dP}{dQ}\right)^2 ; \mathcal{E}_b, \mathcal{L}^c \right] = o(1)v^2(b).$$

Therefore we only need to consider the second moment on the set $\mathcal{L}$, that is,

(4.14)
$$E^Q\left[\left(\frac{dP}{dQ}\right)^2 ; \mathcal{E}_b, \mathcal{L} \right]$$
$$\le E^Q\left[E^Q_{\iota,\tau}\left[\frac{1}{[(1-\rho_1-\rho_2)K]^2} ; \mathcal{E}_b, \mathcal{L}, \mathcal{A}_\tau > 0 \right] \right]$$
$$+ E^Q\left[E^Q_{\iota,\tau}\left[\frac{1}{[(1-\rho_1-\rho_2)K + \rho_1 K_1]^2} ; \mathcal{E}_b, \mathcal{L}, \mathcal{A}_\tau < 0 \right] \right],$$

where K and K_1 are given as in (4.6). We will focus on the terms on the right-hand side of (4.14) in the subsequent derivations. Now, we start to carry out each part of the program.

4.2. *Part* 1. All the derivations in this part are conditional on specific values of ι, τ, $f_*(\tau)$, $\partial f_*(\tau)$ and $\partial^2 f_*(\tau)$, equivalently, ι, τ, w, y and z. By definition,

$$\mathcal{I}(T) = \int_T e^{\sigma f_*(t) + \sigma u_\tau C(t-\tau) + \mu(t)} \, dt.$$

We insert the expansions in (4.10), (4.11) and (4.12) into the expression of $\mathcal{I}(T)$ and obtain that

$$\mathcal{I}(T) = \int_{t \in T} \exp\Big\{ \sigma\Big[w + y^\top(t-\tau) + \frac{1}{2}(t-\tau)^\top z(t-\tau) + R_f(t-\tau)$$

(4.15)
$$+ g(t-\tau) \Big] \Big\}$$

$$\times \exp\Big\{(\sigma u - \mu(\tau))$$

$$\times \Big(1 - \frac{1}{2}(t - \tau)^\top (t - \tau) + C_4(t - \tau) + R_C(t - \tau)\Big)\Big\}$$

$$\times \exp\Big\{\mu(\tau) + \partial\mu(\tau)^\top (t - \tau) + \frac{1}{2}(t - \tau)^\top \Delta\mu(\tau)(t - \tau)$$

$$+ R_\mu(t - \tau)\Big\} dt,$$

where the first row corresponds to the expansion of $w_t = f_*(t)$, and the second and third rows correspond to those of $C(t)$ and $\mu(t)$, respectively. We write the exponent inside the integral in a quadratic form of $(t - \tau)$ and obtain that

$$\mathcal{I}(T) = \exp\Big\{\sigma u + \sigma w + \frac{\sigma}{2}\tilde{y}^\top (uI - \tilde{z})^{-1}\tilde{y}\Big\}$$

(4.16)
$$\times \int_{t\in T} \exp\Big\{-\frac{\sigma}{2}\big(t - \tau - (uI - \tilde{z})^{-1}\tilde{y}\big)^\top (uI - \tilde{z})$$

$$\times \big(t - \tau - (uI - \tilde{z})^{-1}\tilde{y}\big)\Big\}$$

$$\times \exp\{\sigma u_\tau C_4(t - \tau) + \sigma R(t - \tau)\} \times \exp\{\sigma g(t - \tau)\} dt,$$

where $\tilde{y}$ and $\tilde{z}$ are defined as in (4.1). Let $a(s)$ and $b(s)$ be two generic positive functions. Then we have the representation of the following integral:

$$\int_T a(s)b(s)\, ds = E\big[b(S)\big]\int_T a(s)\, ds,$$

where S is a random variable taking values in T with density $a(s)/\int_T a(t)\, dt$. Using this representation and the change of variable that $s = (uI - \tilde{z})^{1/2}(t - \tau)$, we write the big integral in (4.16) as a product of expectations and a normalizing constant, and obtain that

$$\mathcal{I}(T) = \det(uI - \tilde{z})^{-1/2} \exp\Big\{\sigma u + \sigma w + \frac{\sigma}{2}\tilde{y}^\top (uI - \tilde{z})^{-1}\tilde{y}\Big\}$$

$$\times \int_{(uI - \mathbf{z})^{-1/2}s + \tau \in T} \exp\Big\{-\frac{\sigma}{2}\big(s - (uI - \tilde{z})^{-1/2}\tilde{y}\big)^\top$$

$$\times \big(s - (uI - \tilde{z})^{-1/2}\tilde{y}\big)\Big\} ds$$

$$\times E\big[e^{\sigma u_\tau C_4((uI - \tilde{z})^{-1/2}S) + \sigma R((uI - \tilde{z})^{-1/2}S)}\big] \times E\big[e^{\sigma g((uI - \tilde{z})^{-1/2}\tilde{S})}\big].$$

The two expectations in the above display are taken with respect to S and $\tilde{S}$ given the process $g(t)$. S is a random variable taking values in the set $\{s : (uI - \tilde{z})^{-1/2}s +$

$\tau \in T\}$ with density proportional to

$$(4.17) \qquad\qquad e^{-(\sigma/2)(s-(uI-\tilde{\mathbf{z}})^{-1/2}\tilde{y})^{\top}(s-(uI-\tilde{\mathbf{z}})^{-1/2}\tilde{y})},$$

and $\tilde{S}$ is a random variable taking values in the set $\{s : (uI - \tilde{\mathbf{z}})^{-1/2}s + \tau \in T\}$ with density proportional to

$$e^{-(\sigma/2)(s-(uI-\tilde{\mathbf{z}})^{-1/2}\tilde{y})^{\top}(s-(uI-\tilde{\mathbf{z}})^{-1/2}\tilde{y})+\sigma u_{\tau}C_4((uI-\tilde{\mathbf{z}})^{-1/2}s)+\sigma R((uI-\tilde{\mathbf{z}})^{-1/2}s)}.$$

Together with the definition of u that $(\frac{2\pi}{\sigma})^{d/2}u^{-d/2}e^{\sigma u} = b$, we obtain that $\mathcal{I}(T) > b$ if and only if

$$
\begin{aligned}
(4.18) \quad \mathcal{I}(T) &= \det(uI - \tilde{\mathbf{z}})^{-1/2}e^{\sigma u + \sigma w + (\sigma/2)\tilde{y}^{\top}(uI-\tilde{\mathbf{z}})^{-1}\tilde{y}} \\
&\quad \times \int_{(uI-\mathbf{z})^{-1/2}s+\tau \in T} e^{-(\sigma/2)(s-(uI-\tilde{\mathbf{z}})^{-1/2}\tilde{y})^{\top}(s-(uI-\tilde{\mathbf{z}})^{-1/2}\tilde{y})}\,ds \\
&\quad \times E\big[e^{\sigma u_{\tau}C_4((uI-\tilde{\mathbf{z}})^{-1/2}S)+\sigma R((uI-\tilde{\mathbf{z}})^{-1/2}S)}\big] \cdot e^{-u^{-1}\xi_u} \\
&> \left(\frac{2\pi}{\sigma}\right)^{d/2}u^{-d/2}e^{\sigma u},
\end{aligned}
$$

where

$$(4.19) \qquad\qquad \xi_u = -u\log\{E\exp[\sigma g((uI - \tilde{\mathbf{z}})^{-1/2}\tilde{S})]\}.$$

We take log on both sides, and plug in the result of Lemma 20 that handles the big expectation term in (4.18). Then inequality (4.18) is equivalent to

$$
\begin{aligned}
(4.20) \quad w &+ \frac{\tilde{y}^{\top}(uI-\tilde{\mathbf{z}})^{-1}\tilde{y}}{2} - \frac{\log\det(I - \tilde{\mathbf{z}}/u)}{2\sigma} + \frac{\sum_i \partial^4_{iiii}C(0)}{8\sigma^2 u} \\
&> \frac{\xi_u}{u\sigma} + \frac{o(|w|+|y|+|z|+1)}{u^{1+\delta_0/4}}.
\end{aligned}
$$

On the set $\mathcal{L}$, we further simplify (4.20) using the following facts (see Lemma 21):

$$\partial\mu_{\sigma}(\tau) = O(u^{-1/2+\epsilon}),$$

$$
\begin{aligned}
\log\det\left(I - \frac{\tilde{\mathbf{z}}}{u}\right) &= -\frac{1}{u}\mathrm{Tr}(\tilde{\mathbf{z}}) + o(u^{-1-\delta_0/4}) \\
&= -\frac{\mathbf{1}^{\top}(z + \partial^2\mu_{\sigma}(\tau)) + d \cdot \mu_{\sigma}(\tau)}{u} + o(u^{-1-\delta_0/4}),
\end{aligned}
$$

where Tr is the trace of a matrix. Therefore, on the set $\mathcal{L}$, (4.20) is equivalent to

$$
\begin{aligned}
w &+ \frac{y^{\top}y}{2u} + \frac{\mathbf{1}^{\top}(z + \partial^2\mu_{\sigma}(\tau)) + d \cdot \mu_{\sigma}(\tau)}{2\sigma u} + \frac{\sum_i \partial^4_{iiii}C(0)}{8\sigma^2 u} \\
&> \frac{\xi_u}{u\sigma} + \frac{o(|w|+|y|+|z|+1)}{u^{1+\delta_0/4}},
\end{aligned}
$$

and further, equivalently (by replacing u with u_τ),

$$w + \frac{y^\top y}{2u_\tau} + \frac{\mathbf{1}^\top (z + \partial^2 \mu_\sigma(\tau)) + d \cdot \mu_\sigma(\tau)}{2\sigma u_\tau} + \frac{\sum_i \partial^4_{iiii} C(0)}{8\sigma^2 u_\tau}$$

$$> \frac{\xi_u}{u\sigma} + \frac{o(|w| + |y| + |z| + 1)}{u^{1+\delta_0/4}}.$$

Using the notation defined as in (3.9) and (4.8), $\mathcal{I}(T) > b$ is equivalent to

$$\mathcal{A}_\tau + \frac{o(|w| + |y| + |z| + 1)}{u^{1+\delta_0/4}} > \frac{\xi_u}{u\sigma},$$

where $\mathcal{A}_\tau$ is defined as in (4.8). Furthermore, with $\epsilon \ll \delta_0$ and on the set $\mathcal{L}$, $o(|y| + |z|)/u^{-1-\delta_0/4} = o(u^{-1-\delta_0/8})$. For the above inequality, we absorb $o(wu^{-1-\delta_0/4})$ into $\mathcal{A}_\tau$ and rewrite it as

$$\mathcal{A}_\tau > (1 + o(u^{-1-\delta_0/4})) \left[\frac{\xi_u}{\sigma u} + o(u^{-1-\delta_0/8}) \right].$$

4.3. *Part* 2.   In part 2, we first consider $(1 - \rho_1 - \rho_2)K$ in the first expectation of (4.7) (which is on the set $\{\mathcal{A}_\tau \geq 0\}$) and then $(1 - \rho_1 - \rho_2)K + \rho_1 K_1$ in the second expectation of (4.7).

Part 2.1: *The analysis of K when $\mathcal{A}_\tau \geq 0$.*   Similarly to part 1, all the derivations are conditional on (ι, τ, w, y, z). We now proceed to the second part of the proof. More precisely, we simplify the term K defined as in (4.6), and write it as a deterministic function of (w, y, z) with a small correction term. Recall that

$$K = \int_{A^*} l(t) u_t H_\lambda \exp \left\{ -\lambda u_t \left(w_t + u_\tau C(t - \tau) + \frac{\mathbf{1}^\top (\bar{z}_t + \mu_2(t - \tau)u_\tau)}{2\sigma u_t} \right. \right.$$

$$\left. \left. + \frac{B_t}{u_t} - u_t \right) \right\}$$

$$\times \exp \left\{ -\frac{1}{2} \left[\frac{|\mu_{20}\mu_{22}^{-1}(\bar{z}_t + \mu_2(t - \tau)u_\tau)|^2}{1 - \mu_{20}\mu_{22}^{-1}\mu_{02}} \right. \right.$$

$$\left. \left. + \left| \mu_{22}^{-1/2}(\bar{z}_t + \mu_2(t - \tau)u_\tau) - \frac{\mu_{22}^{1/2}\mathbf{1}}{2\sigma} \right|^2 \right] \right\}$$

$$\times h_{x,z}^{-1}(w_t + u_\tau C(t - \tau), z_t + u_\tau \partial^2 C(t - \tau)) \, dt.$$

We plug in the forms of $h_{x,z}$ and $l(t)$ that are defined in (4.4) and (3.13) and obtain that

$$K = (2\pi)^{(d+1)(d+2)/4 - d/2} \det(\Gamma)^{1/2} \cdot \det(-\Delta\mu_\sigma(t_*))^{1/2} u_{t_*}^{d/2} H_\lambda$$

$$\times \int_{A^*} \exp \left\{ \frac{u_{t_*} \cdot (t - t_*)^\top \Delta\mu_\sigma(t_*)(t - t_*)}{2} \right\}$$

$$\times u_t$$

$$\times \exp\left\{-\lambda u_t\left(w_t + u_\tau C(t-\tau) + \frac{\mathbf{1}^\top(\bar{z}_t + \mu_2(t-\tau)u_\tau)}{2\sigma u_t} + \frac{B_t}{u_t} - u_t\right)\right\}$$

$$\times \exp\left\{-\frac{1}{2}\left[\frac{|\mu_{20}\mu_{22}^{-1}(\bar{z}_t + \mu_2(t-\tau)u_\tau)|^2}{1 - \mu_{20}\mu_{22}^{-1}\mu_{02}}\right.\right.$$

$$\left.\left.+ \left|\mu_{22}^{-1/2}(\bar{z}_t + \mu_2(t-\tau)u_\tau) - \frac{\mu_{22}^{1/2}\mathbf{1}}{2\sigma}\right|^2\right]\right\}$$

$$\times \exp\left\{\frac{1}{2}\left[\frac{(w_t + u_\tau C(t-\tau) - \mu_{20}\mu_{22}^{-1}(z_t + \mu_2(t-\tau)u_\tau))^2}{1 - \mu_{20}\mu_{22}^{-1}\mu_{02}}\right.\right.$$

$$\left.\left.+ (z_t + \mu_2(t-\tau)u_\tau)^\top \mu_{22}^{-1}(z_t + \mu_2(t-\tau)u_\tau)\right]\right\} dt.$$

For some δ' such that $\epsilon < \delta' < \delta$, where ϵ, δ are the parameters we used to define $\mathcal{L}$, we further restrict the integration region by defining

$$\mathcal{I}_2 = \int_{A^*, |t-\tau| < u^{-1+\delta'}} \exp\left\{\frac{u_{t_*}(t-t_*)^\top \Delta\mu_\sigma(t_*)(t-t_*)}{2}\right\}$$

$$\times u_t \times \exp\left\{-\lambda u_t\left(w_t + u_\tau C(t-\tau)\right.\right.$$

$$\left.\left.+ \frac{\mathbf{1}^\top(\bar{z}_t + \mu_2(t-\tau)u_\tau)}{2\sigma u_t} + \frac{B_t}{u_t} - u_t\right)\right\}$$

(4.21) $$\times \exp\left\{-\frac{1}{2}\left[\frac{|\mu_{20}\mu_{22}^{-1}(\bar{z}_t + \mu_2(t-\tau)u_\tau)|^2}{1 - \mu_{20}\mu_{22}^{-1}\mu_{02}}\right.\right.$$

$$\left.\left.+ \left|\mu_{22}^{-1/2}(\bar{z}_t + \mu_2(t-\tau)u_\tau) - \frac{\mu_{22}^{1/2}\mathbf{1}}{2\sigma}\right|^2\right]\right\}$$

$$\times \exp\left\{\frac{1}{2}\left[\frac{(w_t + u_\tau C(t-\tau) - \mu_{20}\mu_{22}^{-1}(z_t + \mu_2(t-\tau)u_\tau))^2}{1 - \mu_{20}\mu_{22}^{-1}\mu_{02}}\right.\right.$$

$$\left.\left.+ (z_t + \mu_2(t-\tau)u_\tau)^\top \mu_{22}^{-1}(z_t + \mu_2(t-\tau)u_\tau)\right]\right\} dt.$$

Thus

$$K \geq (2\pi)^{(d+1)(d+2)/4 - d/2} \det(\Gamma)^{1/2}$$

$$\times \det(-\Delta\mu_\sigma(t_*))^{1/2} u_{t_*}^{d/2} H_\lambda \cdot \mathcal{I}_2.$$

For the rest of part 2.1, we focus on $\mathcal{I}_2$. With some tedious algebra, Lemma 22 writes $\mathcal{I}_2$ in a more manageable form; that is, $\mathcal{I}_2$ equals

$$
\int_{A^*,\,|t-\tau|<u^{-1+\delta'}} \exp\left\{ \frac{u_{t_*}(t-t_*)^\top \Delta\mu_\sigma(t_*)(t-t_*)}{2} + \frac{u_t^2}{2} \right\} \times u_t
$$

$$
\times \exp\left\{ (1-\lambda)u_t\left[w_t + u_\tau C(t-\tau) - u_t \right] \right.
$$

$$
\left. + \frac{(1-\lambda)}{2\sigma}\mathbf{1}^\top(z_t - \mu_{02}u_t + \mu_2(t-\tau)u_\tau) - \lambda B_t - \frac{\mathbf{1}^\top\mu_{22}\mathbf{1}}{8\sigma^2} \right\}
$$

(4.22)

$$
\times \exp\Big\{ \big((w_t + u_\tau C(t-\tau) - u_t)^2
$$

$$
- 2(w_t + u_\tau C(t-\tau) - u_t)\mu_{20}\mu_{22}^{-1}(z_t - \mu_{02}u_t + \mu_2(t-\tau)u_\tau)\big)
$$

$$
/(2(1 - \mu_{20}\mu_{22}^{-1}\mu_{02}))\Big\}\,dt.
$$

Lemma 23 implies that $\{|t-\tau| < u^{-1+\delta'}\} \subset A^*$. Thus, on the set $\{\mathcal{A}_\tau > 0\}$, we have $A^* \cap \{|t-\tau| < u^{-1+\delta'}\} = \{|t-\tau| < u^{-1+\delta'}\}$ and we can remove A^* from the integration region of $\mathcal{I}_2$. In addition, on the set $\mathcal{L}$ and $|t-\tau| < u^{-1+\delta'}$, we have that

$$
u_\tau - u_t C(t-\tau) = O(u^{-1+2\delta'}), \qquad \mu_2(t-\tau) = \mu_{20} + O(|t-\tau|^2),
$$

$$
|u_\tau\mu_2(t-\tau) - u_t\mu_{20}| = O(u^{-1+2\delta'}), \qquad (u_\tau - u_t C(t-\tau))|z_t| = o(1).
$$

We insert the above estimates to (4.22). Together with the fact that

$$
\exp\left\{ \frac{u_{t_*}(t-t_*)^\top \Delta\mu_\sigma(t_*)(t-t_*)}{2} + \frac{u_t^2}{2} \right\} = (1+o(1))\exp\left\{ \frac{1}{2}u_{t_*}^2 \right\},
$$

we have that

$$
\mathcal{I}_2 \sim u \times \exp\left\{ \frac{1}{2}u_{t_*}^2 - \lambda B_{t_*} - \frac{\mathbf{1}^\top\mu_{22}\mathbf{1}}{8\sigma^2} \right\}
$$

$$
\times \int_{|t-\tau|<u^{-1+\delta'}} \exp\Big\{ (1-\lambda)
$$

$$
\times u_t\big[w_t + u_\tau(C(t-\tau) - 1) + (\mu_\sigma(t) - \mu_\sigma(\tau)) \big]
$$

$$
+ (1-\lambda)\frac{\mathbf{1}^\top z}{2\sigma} + \frac{w_t^2 - 2w_t\mu_{20}\mu_{22}^{-1}z_t + o(1)w_t}{2(1 - \mu_{20}\mu_{22}^{-1}\mu_{02})} \Big\}\,dt.
$$

Further, we have that

$$
w_t^2 - 2w_t\mu_{20}\mu_{22}^{-1}z_t + o(1)w_t = o(1) + u \cdot w \cdot O(u^{-1/2+\delta'}).
$$

Let $\zeta_u = O(u^{-1/2+\delta'})$, and we simplify $\mathcal{I}_2$ to

$$\mathcal{I}_2 \sim u \times \exp\left\{\frac{1}{2}u_{t_*}^2 - \lambda B_{t_*} - \frac{\mathbf{1}^\top \mu_{22}\mathbf{1}}{8\sigma^2}\right\}$$

$$\times \int_{|t-\tau|<u^{-1+\delta'}} \exp\left\{(1-\lambda)(u_\tau + \zeta_u)[\zeta_u w + w_t + u_\tau(C(t-\tau)-1)\right.$$

$$\left. + (\mu_\sigma(t) - \mu_\sigma(\tau))] + (1-\lambda)\frac{\mathbf{1}^\top z}{2\sigma}\right\} dt.$$

In what follows, we insert the expansions in (4.10), (4.11) and (4.12) into the expression of $\mathcal{I}_2$ and write the exponent as a quadratic function of $t-\tau$, and we obtain that on the set $\mathcal{L}$

$$\mathcal{I}_2 \sim u \times \exp\left\{\frac{1}{2}u_{t_*}^2 - \lambda B_{t_*} - \frac{\mathbf{1}^\top \mu_{22}\mathbf{1}}{8\sigma^2}\right\}$$

$$\times \exp\left\{(1-\lambda)(u_\tau + \zeta_u)\left((1+\zeta_u)w + \frac{1}{2}\tilde{y}^\top(uI - \tilde{\mathbf{z}})^{-1}\tilde{y} + \frac{\mathbf{1}^\top z}{2\sigma u_\tau}\right)\right\}$$

(4.23)

$$\times \int_{|t-\tau|<u^{-1+\delta'}} e^{-(1/2)(1-\lambda)(u_\tau+\zeta_u)(t-\tau-(uI-\tilde{\mathbf{z}})^{-1}\tilde{y})^\top(uI-\tilde{\mathbf{z}})(t-\tau-(uI-\tilde{\mathbf{z}})^{-1}\tilde{y})}$$

$$\times e^{(1-\lambda)(u_\tau+\zeta_u)[u_\tau C_4(t-\tau)+R(t-\tau)+g(t-\tau)]} dt,$$

where we recall that $\tilde{y} = y + \partial\mu_\sigma(\tau)$ and $\tilde{\mathbf{z}} = \mathbf{z} + u_\sigma(\tau)I + \Delta\mu_\sigma(\tau)$. This derivation is very similar to that from (4.15) to (4.16). In the last row of the above display, on the set $\mathcal{L}$ and $|t - \tau| < u^{-1+\delta'}$,

$$u^2 C_4(t - \tau) + u R(t - \tau) = o(1).$$

Therefore, they can be ignored. We consider the change of variable that

$$s = (1 - \lambda)^{1/2}(u_\tau + \zeta_u)^{1/2}(uI - \tilde{\mathbf{z}})^{1/2}(t - \tau)$$

and obtain that $\mathcal{I}_2$ equals (with the terms C_4 and R removed)

$$\mathcal{I}_2 \sim (1-\lambda)^{-d/2}u^{-d+1}\exp\left\{\frac{1}{2}u_{t_*}^2 - \lambda B_{t_*} - \frac{\mathbf{1}^\top \mu_{22}\mathbf{1}}{8\sigma^2}\right\}$$

$$\times \exp\left\{(1-\lambda)(u_\tau + \zeta_u)\left((1+\zeta_u)w + \frac{1}{2}\tilde{y}^\top(uI - \tilde{\mathbf{z}})^{-1}\tilde{y} + \frac{\mathbf{1}^\top z}{2\sigma u}\right)\right\}$$

(4.24)

$$\times \int_{s\in\mathcal{S}_u} e^{-(1/2)|s-(1-\lambda)^{1/2}(u_\tau+\zeta_u)^{1/2}(uI-\tilde{\mathbf{z}})^{-1/2}\tilde{y}|^2} ds$$

$$\times E\left[e^{(1-\lambda)(u_\tau+\zeta_u)g((1-\lambda)^{-1/2}(u_\tau+\zeta_u)^{-1/2}(uI-\tilde{\mathbf{z}})^{-1/2}S')}\right],$$

where $\mathcal{S}_u = \{s : |(1-\lambda)^{-1/2}(u_\tau + \zeta_u)^{-1/2}(uI - \tilde{\mathbf{z}})^{-1/2}s| < u^{-1+\delta'}\}$, and S' is a random variable taking values on the set $\mathcal{S}_u$ with density proportional to

$$e^{-(1/2)|s-(1-\lambda)^{1/2}(u_\tau+\zeta_u)^{1/2}(uI-\tilde{\mathbf{z}})^{-1/2}\tilde{y}|^2}.$$

We use κ to denote the last two terms of (4.24), that is,

(4.25)

$$\kappa = \int_{\mathcal{S}_u} e^{-(1/2)|s-(1-\lambda)^{1/2}(u_\tau+\zeta_u)^{1/2}(uI-\tilde{\mathbf{z}})^{-1/2}\tilde{y}|^2} ds$$

$$\times E\big[e^{(1-\lambda)(u_\tau+\zeta_u)g((1-\lambda)^{-1/2}(u_\tau+\zeta_u)^{-1/2}(uI-\tilde{\mathbf{z}})^{-1/2}S')}\big].$$

It is helpful to keep in mind that κ is approximately $(2\pi)^{d/2}$. We insert κ back to the expression of $\mathcal{I}_2$. Together with the fact that $\tilde{y}^\top (uI-\tilde{\mathbf{z}})^{-1}\tilde{y} = |\tilde{y}|^2/u + o(u^{-1})$, we have

(4.26)

$$\mathcal{I}_2 \sim \kappa(1-\lambda)^{-d/2}u^{-d+1}\exp\left\{\frac{1}{2}u_{t_*}^2 - \lambda B_{t_*} - \frac{\mathbf{1}^\top \mu_{22}\mathbf{1}}{8\sigma^2}\right\}$$

$$\times \exp\left\{(1-\lambda)(u_\tau+\zeta_u)\Big((1+\zeta_u)w + \frac{|\tilde{y}|^2}{2u_\tau} + \frac{\mathbf{1}^\top z}{2\sigma u_\tau}\Big)\right\}.$$

Thus, we have that on the set $\{\mathcal{A}_\tau > 0\}$,

$$K \geq (2\pi)^{(d+1)(d+2)/4-d/2}\det(\Gamma)^{1/2} \cdot \det(-\Delta\mu_\sigma(t_*))^{1/2}u_{t_*}^{d/2}H_\lambda \cdot \mathcal{I}_2$$

$$= (\kappa + o(1))(2\pi)^{(d+1)(d+2)/4-d/2}\det(\Gamma)^{1/2}$$

(4.27)

$$\times \det(-\Delta\mu_\sigma(t_*))^{1/2}H_\lambda \cdot (1-\lambda)^{-d/2}u^{-d/2+1}$$

$$\times \exp\left\{\frac{1}{2}u_{t_*}^2 - \lambda B_{t_*} - \frac{\mathbf{1}^\top \mu_{22}\mathbf{1}}{8\sigma^2}\right.$$

$$\left. + (1-\lambda)(u_\tau+\zeta_u)\Big((1+\zeta_u)w + \frac{|\tilde{y}|^2}{2u_\tau} + \frac{\mathbf{1}^\top z}{2\sigma u_\tau}\Big)\right\}.$$

We further insert the $\mathcal{A}_\tau$ defined in (4.8) into (4.27) and obtain that

$$K \geq (\kappa + o(1))(2\pi)^{(d+1)(d+2)/4-d/2}\det(\Gamma)^{1/2}$$

$$\times \det(-\Delta\mu_\sigma(t_*))^{1/2}H_\lambda \cdot (1-\lambda)^{-d/2}u^{-d/2+1}$$

(4.28)

$$\times \exp\left\{\frac{1}{2}u_{t_*}^2 - B_{t_*} - \frac{\mathbf{1}^\top \mu_{22}\mathbf{1}}{8\sigma^2} + (1-\lambda)u_\tau(1+o(1))\mathcal{A}_\tau\right.$$

$$\left. + (1-\lambda)\zeta_u \cdot (|\tilde{y}|^2 + |z|)\right\}.$$

Part 2.2: *The analysis of dP/dQ when $\mathcal{A}_\tau < 0$.* In this part, we focus mostly on the K_1 term, whose handling is very similar to that of K. Therefore, we only list out the key steps. For some large constant M, let

$$D = \{|t - \tau - (uI - \tilde{\mathbf{z}})^{-1}\tilde{y}| < Mu^{-1}\}$$

that is, the dominating region of the integral. We split the set $D = (A^* \cap D) \cup ((A^*)^c \cap D)$. There are two situations: $\mathrm{mes}((A^*)^c \cap D) > \mathrm{mes}(A^* \cap D)$ and

mes$((A^*)^c \cap D) \le$ mes$(A^* \cap D)$. For the first situation, the term K_1 is dominating; for the second situation, the term K (more precisely $\mathcal{I}_2$) is dominating.

To simplify K_1, we write it as

$$
\begin{aligned}
K_1 &= (2\pi)^{(d+1)(d+2)/4-d/2} \det(\Gamma)^{1/2} \cdot \det\big(-\Delta\mu_\sigma(t_*)\big)^{1/2} u_{t_*}^{d/2} H_{\lambda_1} \\
&\quad \times \left[\int_{(A^*)^c \cap D} + \cdots + \int_{(A^*)^c \cap D^c} \cdots \right] \\
&\triangleq (2\pi)^{(d+1)(d+2)/4-d/2} \det(\Gamma)^{1/2} \cdot \det\big(-\Delta\mu_\sigma(t_*)\big)^{1/2} u_{t_*}^{d/2} H_{\lambda_1} \\
&\quad \times [\mathcal{I}_{1,2} + \mathcal{I}_{1,3}].
\end{aligned}
$$

Note that the difference between K_1 and K is that the term "$-\lambda$" has been replaced by "λ_1." With exactly the same derivation for (4.22), we obtain that $\mathcal{I}_{1,2}$ equals [by replacing "$-\lambda$" in (4.22) by "λ_1"]

$$
\begin{aligned}
\int_{(A^*)^c \cap D} &\exp\left\{ \frac{u_{t_*}(t-t_*)^\top \Delta\mu_\sigma(t_*)(t-t_*)}{2} + \frac{1}{2}u_t^2 \right\} \times u_t \\
&\times \exp\Big\{ (1+\lambda_1)u_t\big[w_t + u_\tau C(t-\tau) - u_t\big] \\
&\qquad + \frac{(1+\lambda_1)}{2\sigma}\mathbf{1}^\top\big(z_t - \mu_{02}u_t + \mu_2(t-\tau)u_\tau\big) + \lambda_1 B_t - \frac{\mathbf{1}^\top \mu_{22}\mathbf{1}}{8\sigma^2} \Big\} \\
(4.29) \quad &\times \exp\big\{ \big((w_t + u_\tau C(t-\tau) - u_t)^2 \\
&\qquad - 2\big(w_t + (u_\tau C(t-\tau) - u_t)\big) \\
&\qquad \times \mu_{20}\mu_{22}^{-1}\big(z_t - \mu_{02}u_t + \mu_2(t-\tau)u_\tau\big) \\
&\qquad\qquad /\big(2(1 - \mu_{20}\mu_{22}^{-1}\mu_{02})\big)\big\} \, dt.
\end{aligned}
$$

With a very similar derivation as in part 2.1, in particular, the result in (4.23), we have that

$$
\begin{aligned}
\mathcal{I}_{1,2} &\sim u \times \exp\left\{ \frac{1}{2}u_{t_*}^2 + \lambda_1 B_{t_*} - \frac{\mathbf{1}^\top \mu_{22}\mathbf{1}}{8\sigma^2} \right\} \\
&\times \exp\left\{ (1+\lambda_1)(u_\tau + \zeta_u)\Big((1+\zeta_u)w + \frac{1}{2}\tilde{y}^\top(uI - \tilde{\mathbf{z}})^{-1}\tilde{y} + \frac{\mathbf{1}^\top z}{2\sigma u} \Big) \right\} \\
(4.30) \quad &\times \int_{(A^*)^c \cap D} \exp\Big\{ (1+\lambda_1)(u_\tau + \zeta_u)\Big[-\frac{1}{2}\big(t - \tau - (uI - \tilde{\mathbf{z}})^{-1}\tilde{y}\big)^\top(uI - \tilde{\mathbf{z}}) \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad \times \big(t - \tau - (uI - \tilde{\mathbf{z}})^{-1}\tilde{y}\big)\Big] \Big\} \\
&\times \exp\big\{ (1+\lambda_1)(u_\tau + \zeta_u)\big[u_\tau C_4(t-\tau) + R(t-\tau) + g(t-\tau)\big]\big\} \, dt.
\end{aligned}
$$

Furthermore, similarly to the results in (4.26), we have that

$$
\mathcal{I}_{1,2} \sim \kappa_{1,2}(1+\lambda_1)^{-d/2}u^{-d+1}e^{(1/2)u_{t_*}^2+\lambda_1 B_{t_*}-\mathbf{1}^\top \mu_{22}\mathbf{1}/(8\sigma^2)}
$$

(4.31)

$$
\times e^{(1+\lambda_1)(u_\tau+\zeta_u)((1+\zeta_u)w+(1/2)\tilde{y}^\top(uI-\tilde{\mathbf{z}})^{-1}\tilde{y}+\mathbf{1}^\top z/(2\sigma u_\tau))},
$$

where $\kappa_{1,2}$ is defined as

$$
\kappa_{1,2} = \int_{t_1(s)\in(A^*)^c\cap D} e^{-1/2|s-(1+\lambda_1)^{1/2}(u_\tau+\zeta_u)^{1/2}(uI-\tilde{\mathbf{z}})^{-1/2}\tilde{y}|^2}\,ds
$$

$$
\times E\big[e^{(1+\lambda_1)(u_\tau+\zeta_u)g((1+\lambda_1)^{-1/2}(u_\tau+\zeta_u)^{-1/2}(uI-\tilde{\mathbf{z}})^{-1/2}S_{1,2})}\big],
$$

the change of variable $t_1(s) = \tau + (1+\lambda_1)^{-1/2}(u_\tau+\zeta_u)^{-1/2}(uI-\tilde{\mathbf{z}})^{-1/2}s$ and $S_{1,2}$ is a random variable taking values in the set $\{s : t(s) \in (A^*)^c \cap D\}$ with an appropriately chosen density function similarly as in (4.24). In summary, the only difference between $\mathcal{I}_{1,2}$ and $\mathcal{I}_2$ lies in that the multiplier $-\lambda$ is replaced by λ_1.

We now proceed to providing a lower bound of $(1-\rho_1-\rho_2)K + \rho_1 K_1$. Note that

$$
\max\{\mathrm{mes}((A^*)^c \cap D), \mathrm{mes}(A^* \cap D)\} \geq \tfrac{1}{2}\mathrm{mes}(D).
$$

Therefore at least one of $(A^*)^c \cap D$ and $A^* \cap D$ is nonempty. If $\mathrm{mes}((A^*)^c \cap D) \geq \frac{1}{2}\mathrm{mes}(D)$, we have the bound

$$
(1-\rho_1-\rho_2)K + \rho_1 K_1 \geq \rho_1 K_1 \geq \Theta(1)\rho_1 u^{d/2}\mathcal{I}_{1,2}.
$$

Similarly, if $\mathrm{mes}(A^* \cap D) \geq \frac{1}{2}\mathrm{mes}(D)$, we have that

$$
(1-\rho_1-\rho_2)K + \rho_1 K_1 \geq \Theta(1)(1-\rho_1-\rho_2)u^{d/2}\mathcal{I}_2.
$$

We further split $\mathcal{I}_2$ in part 2.1 into two parts:

(4.32)
$$
\mathcal{I}_2 = \int_{A^*\cap D} \cdots\, dt + \int_{A^*\cap D^c} \cdots\, dt \triangleq \mathcal{I}_{2,1} + \mathcal{I}_{2,2}.
$$

Similarly to the derivation of $\mathcal{I}_{1,2}$, we have that

$$
\mathcal{I}_{2,1} \sim \kappa_{2,1}(1-\lambda)^{-d/2}u^{-d+1}e^{(1/2)u_{t_*}^2-\lambda B_{t_*}-\mathbf{1}^\top \mu_{22}\mathbf{1}/(8\sigma^2)}
$$

$$
\times e^{(1-\lambda)(u_\tau+\zeta_u)((1+\zeta_u)w+|\tilde{y}|^2/(2u_\tau)+\mathbf{1}^\top z/(2\sigma u_\tau))},
$$

where

(4.33)
$$
\kappa_{2,1} = \int_{t_2(s)\in A^*\cap D} e^{-1/2|s-(1-\lambda)^{1/2}(u_\tau+\zeta_u)^{1/2}(uI-\tilde{\mathbf{z}})^{-1/2}\tilde{y}|^2}\,ds
$$

$$
\times E\big[e^{(1-\lambda)(u_\tau+\zeta_u)g((1-\lambda)^{-1/2}(u_\tau+\zeta_u)^{-1/2}(uI-\tilde{\mathbf{z}})^{-1/2}S_{2,1})}\big].
$$

$S_{2,1}$ is a random variable taking values on the set $\{s : t_2(s) \in A^* \cap D\}$ with an appropriate density function similarly as in (4.24) and $t_2(s) = \tau + (1-\lambda)^{-1/2} \times (u_\tau+\zeta_u)^{-1/2}(uI-\tilde{\mathbf{z}})^{-1/2}s$.

Then combining the above results of $\mathcal{I}_{1,2}$ and $\mathcal{I}_{2,1}$, we have that for the case in which $\mathcal{A}_\tau < 0$

$$
\begin{aligned}
\rho_1 K_1 &+ (1 - \rho_1 - \rho_2) K \\
&\geq \Theta(1) u^{d/2} \big[\mathbb{I}_{C_1} \rho_1 \mathcal{I}_{1,2} + \mathbb{I}_{C_2} (1 - \rho_1 - \rho_2) \mathcal{I}_{2,1} \big] \\
&\geq \Theta(1) u^{-d/2+1} e^{(1/2) u_{t_*}^2} \\
&\quad \times \big[\mathbb{I}_{C_1} \cdot \rho_1 \kappa_{1,2} e^{(1+\lambda_1)(u_\tau + \zeta_u)((1+\zeta_u)w + |\tilde{y}|^2/(2u_\tau) + \mathbf{1}^\top z/(2\sigma u_\tau))} \\
&\quad\quad + \mathbb{I}_{C_2} \cdot (1 - \rho_1 - \rho_2)(1 - \lambda)^{-d/2} \kappa_{2,1} \\
&\quad\quad\quad \times e^{(1-\lambda)(u_\tau + \zeta_u)((1+\zeta_u)w + |\tilde{y}|^2/(2u_\tau) + \mathbf{1}^\top z/(2\sigma u_\tau))} \big],
\end{aligned}
$$

where $C_1 = \{ f(\cdot) : \mathrm{mes}((A^*)^c \cap D) \geq \mathrm{mes}(A^* \cap D) \}$ and $C_2 = C_1^c$. We further insert $\mathcal{A}_\tau$ defined in (4.8). Note that on the set $\{\mathcal{A}_\tau < 0\}$, $(1 + \lambda_1)\mathcal{A}_\tau < (1 - \lambda)\mathcal{A}_\tau$ and B_t is bounded away from zero and infinity. Then

$$
\begin{aligned}
(4.34) \quad
(1 - \rho_1 - \rho_2) K &+ \rho_1 K_1 \\
&\geq \Theta(1) u^{-d/2+1} e^{(1/2) u_{t_*}^2} \cdot e^{(1+\lambda_1)(1+\zeta_u) u_\tau \mathcal{A}_\tau + \zeta_u \cdot (|\tilde{y}|^2 + |z|)} \\
&\quad \times \big[\mathbb{I}_{C_1} \cdot \rho_1 \kappa_{1,2} + \mathbb{I}_{C_2} \cdot (1 - \rho_1 - \rho_2)(1 - \lambda)^{-d/2} \kappa_{2,1} \big].
\end{aligned}
$$

Part 3.   We now put together the results in parts 1 and 2 and obtain an approximation for (4.7). Recall that

$$
\begin{aligned}
(4.35) \quad
E^Q &\Big[\Big(\frac{dP}{dQ} \Big)^2 ; \mathcal{E}_b, \mathcal{L} \Big] \\
&\leq E^Q \Big[\frac{1}{[(1 - \rho_1 - \rho_2) K]^2} ; \mathcal{E}_b, \mathcal{L}, \mathcal{A}_\tau \geq 0 \Big] \\
&\quad + E^Q \Big[\frac{1}{[(1 - \rho_1 - \rho_2) K + \rho_1 K_1]^2} ; \mathcal{E}_b, \mathcal{L}, \mathcal{A}_\tau < 0 \Big].
\end{aligned}
$$

We consider the two terms on the right-hand side of the above display one by one. We start with the first term

$$
\begin{aligned}
(4.36) \quad
E^Q &\Big[\frac{1}{[(1 - \rho_1 - \rho_2) K]^2} ; \mathcal{E}_b, \mathcal{L}, \mathcal{A}_\tau \geq 0 \Big] \\
&= E^Q \Big[\frac{1}{[(1 - \rho_1 - \rho_2) K]^2} ; \mathcal{E}_b, \mathcal{L}, \mathcal{A}_\tau \geq 0, \iota = 0 \Big] \\
&\quad + E^Q \Big[\frac{1}{[(1 - \rho_1 - \rho_2) K]^2} ; \mathcal{E}_b, \mathcal{L}, \mathcal{A}_\tau \geq 0, \iota = 1 \Big].
\end{aligned}
$$

The index τ admits density $l(t)$ when $\iota = 0$ and τ is uniformly distributed over T if $\iota = 1$.

Consider the first expectation in (4.36). Note that conditionally on τ and $\iota = 0$, on the set $\mathcal{L} \cap \{\mathcal{A}_\tau \geq 0\}$, (w, y, z) follows density $(1 - \rho_1 - \rho_2)h_{0,\tau}^*(w, y, z)/(1 - \rho_2)$ defined as in (4.2). Thus, according to (4.28), we have that the conditional expectation

$$E^Q\left[\frac{1}{(1 - \rho_1 - \rho_2)^2 K^2}; \mathcal{E}_b, \mathcal{L}, \mathcal{A}_\tau \geq 0 \Big| \iota = 0, \tau\right]$$

$$\leq (1 + o(1))\left[\frac{H_\lambda^{-1}\det(\Gamma)^{-1/2}\det(-\Delta\mu_\sigma(t_*))^{-1/2}}{(2\pi)^{(d+1)(d+2)/4 - d/2}}\right.$$

$$(4.37) \qquad\qquad \left. \times (1 - \lambda)^{d/2}u^{d/2 - 1}e^{-(1/2)u_{t_*}^2 + B_{t_*} + \mathbf{1}^\top\mu_{22}\mathbf{1}/(8\sigma^2)}\right]^2$$

$$\times \int_{\mathcal{A}_\tau > 0, \mathcal{L}} e^{-2(1-\lambda)u((1 + o(1))\mathcal{A}_\tau + o(|y|^2/(2u) + \mathbf{1}^\top z/(2\sigma u)))} \cdot \gamma_u(u\sigma\mathcal{A}_\tau)$$

$$\times \frac{1 - \rho_1 - \rho_2}{1 - \rho_2}h_{0,\tau}^*(w, y, z)\,dw\,dy\,dz,$$

where

$$\gamma_u(x) = E\left[\frac{1}{(1 - \rho_1 - \rho_2)^2\kappa^2};\right.$$

$$\left. x > (1 + o(u^{-1 - \delta_0/4}))[\xi_u + o(u^{-\delta_0/8})]\Big|\iota, \tau, w, y, z\right],$$

with the expectation taken with respect to the process $g(t)$. We insert the analytic form of $h_{0,\tau}^*(w, y, z)$ into (4.2) and obtain that

$$\int_{\mathcal{A}_\tau > 0, \mathcal{L}} e^{-2(1-\lambda)u((1 + o(1))\mathcal{A}_\tau + o(|y|^2/(2u) + \mathbf{1}^\top z/(2\sigma u)))} \cdot \gamma_u(u\sigma\mathcal{A}_\tau)$$

$$\times \frac{1 - \rho_1 - \rho_2}{1 - \rho_2}h_{0,\tau}^*(w, y, z)\,dw\,dy\,dz$$

$$= \frac{(1 - \rho_1 - \rho_2)H_\lambda \cdot u_\tau}{1 - \rho_2}$$

$$(4.38) \qquad \times \int_{\mathcal{A}_\tau > 0} \gamma_u(u\sigma\mathcal{A}_\tau)\exp\{-2(1 - \lambda + o(1))u\mathcal{A}_\tau + o(|z| + |y|^2)\}$$

$$\times \exp\left\{-\lambda u_\tau\mathcal{A}_\tau\right.$$

$$- \frac{1}{2}\left[\frac{|\mu_{20}\mu_{22}^{-1}z|^2}{1 - \mu_{20}\mu_{22}^{-1}\mu_{02}} + \left|\mu_{22}^{-1/2}z - \frac{\mu_{22}^{1/2}\mathbf{1}}{2\sigma}\right|^2\right]$$

$$\left. - \frac{1 - \lambda}{2}y^\top y\right\}d\mathcal{A}_\tau\,dy\,dz.$$

Thanks to the Borel–TIS inequality (Lemma 16), Lemma 19 and the definition of κ in (4.25), for $x > 0$, $\gamma_u(x)$ is bounded and as $b \to \infty$,

$$E\left[\frac{1}{\kappa^2}; x > (1 + o(u^{-1-\delta_0/4}))[\xi_u + o(u^{-\delta_0/8})]\right] \to (2\pi)^{-d}.$$

Thus, by the dominated convergence theorem and with H_λ defined as in (3.14), as $u \to \infty$, we have that

$$(4.38) \sim \frac{(2\pi)^{-d}}{(1 - \rho_1 - \rho_2)(1 - \rho_2)} \frac{e^{-\lambda\eta}\lambda}{2 - \lambda}.$$

We insert it back to (4.37) and obtain that

$$E^Q\left[\frac{1}{(1 - \rho_1 - \rho_2)^2 K^2}; \mathcal{E}_b, \mathcal{L}, \mathcal{A}_\tau \geq 0 \Big| \iota = 0, \tau\right]$$

$$\leq (1 + o(1)) \frac{(2\pi)^{-d}}{(1 - \rho_1 - \rho_2)(1 - \rho_2)} \frac{e^{-\lambda\eta}\lambda}{2 - \lambda}$$

(4.39)

$$\times \left[\frac{H_\lambda^{-1} \det(\Gamma)^{-1/2} \det(\Delta\mu_\sigma(t_*))^{-1/2}}{(2\pi)^{(d+1)(d+2)/4 - d/2}}\right.$$

$$\left. \times (1 - \lambda)^{d/2} u^{d/2 - 1} e^{-(1/2)u_{t_*}^2 + B_{t_*} + \mathbf{1}^\top \mu_{22} \mathbf{1}/(8\sigma^2)}\right]^2.$$

Using the asymptotic approximation of $v(b)$ given by Proposition 14, we obtain that

$$E^Q\left[\frac{1}{[(1 - \rho_1 - \rho_2)K]^2}; \mathcal{E}_b, \mathcal{L}, \mathcal{A}_\tau \geq 0, \iota = 0\right]$$

(4.40)

$$\leq \frac{1 + o(1)}{1 - \rho_1 - \rho_2} \frac{e^{\lambda\eta}}{\lambda(2 - \lambda)} v^2(b).$$

We choose

$$\rho_1 = \rho_2 = \eta = 1 - \lambda = 1/\log\log b \sim 1/\log u.$$

Then, the right-hand side of the above inequality is bounded by $(1 + \varepsilon)v^2(b)$ for b sufficiently large.

The handling of the second term of (4.36) is similar except that (w, y, z) follows density $h_\tau^*(w, y, z)$. Thus, we only mention the key steps. Note that

$$E^Q\left[\frac{1}{(1 - \rho_1 - \rho_2)^2 K^2}; \mathcal{E}_b, \mathcal{L}, \mathcal{A}_\tau \geq 0 \Big| \iota = 1, \tau\right]$$

$$= (1 + o(1)) \left[\frac{H_\lambda^{-1} \det(\Gamma)^{-1/2} \det(-\Delta\mu_\sigma(t_*))^{-1/2}}{(2\pi)^{(d+1)(d+2)/4 - d/2}}\right.$$

(4.41)

$$\left. \times (1 - \lambda)^{d/2} u^{d/2 - 1} e^{-(1/2)u_{t_*}^2 + B_{t_*} + \mathbf{1}^\top \mu_{22} \mathbf{1}/(8\sigma^2)}\right]^2$$

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

$$\times \int_{\mathcal{A}_\tau \geq 0, \mathcal{L}} \gamma_u(u\sigma \mathcal{A}_\tau)$$

$$\times \exp\left\{-2(1-\lambda)u\mathcal{A}_\tau \right.$$

$$-\frac{1+o(1)}{2}$$

$$\left. \times \left[y^\top y + \frac{|w - \mu_{20}\mu_{22}^{-1}z|^2}{1 - \mu_{20}\mu_{22}^{-1}\mu_{02}} + z^\top \mu_{22}^{-1}z\right]\right\} d\mathcal{A}_\tau \, dy \, dz$$

$$= O(1)(1-\lambda)^{-1}u^{-1} \cdot u^{d-2}e^{-u_{t*}^2}.$$

According to the asymptotic form of $v(b)$ and with $\rho_2 = 1 - \lambda = 1/\log\log b$, we have that

(4.42)
$$E^Q\left[\frac{1}{[(1-\rho_1-\rho_2)K]^2}; \mathcal{E}_b, \mathcal{L}, \mathcal{A}_\tau \geq 0, \iota = 1\right]$$
$$= O(1)\rho_2(1-\lambda)^{-1}u^{-1} \cdot u^{d-2}e^{-u_{t*}^2} = o(1)v^2(b).$$

Therefore, combining the results in (4.40) and (4.42), we have the first term in (4.35) is bounded by $(1+2\varepsilon)v^2(b)$.

The last step is to show that the second term of (4.35) is of a smaller order of $v^2(b)$. First, we split the expectation

(4.43)
$$E^Q\left[\frac{1}{[(1-\rho_1-\rho_2)K + \rho_1 K_1]^2}; \mathcal{E}_b, \mathcal{L}, \mathcal{A}_\tau < 0\right]$$
$$= E^Q\left[\frac{1}{[(1-\rho_1-\rho_2)K + \rho_1 K_1]^2}; \mathcal{E}_b, \mathcal{L}, \mathcal{A}_\tau < 0, \iota = 1\right]$$
$$+ E^Q\left[\frac{1}{[(1-\rho_1-\rho_2)K + \rho_1 K_1]^2}; \mathcal{E}_b, \mathcal{L}, -\eta/u_\tau < \mathcal{A}_\tau < 0, \iota = 0\right]$$
$$+ E^Q\left[\frac{1}{[(1-\rho_1-\rho_2)K + \rho_1 K_1]^2}; \mathcal{E}_b, \mathcal{L}, \mathcal{A}_\tau \leq -\eta/u_\tau, \iota = 0\right].$$

We study these three terms one by one. Let

(4.44)
$$\gamma_{1,u}(x) = E\left[\frac{1}{[\mathbb{I}_{C_1} \cdot \rho_1\kappa_{1,2} + \mathbb{I}_{C_2} \cdot (1-\rho_1-\rho_2)(1-\lambda)^{-d/2}\kappa_{2,1}]^2};\right.$$
$$\left. x > (1+o(u^{-1-\delta_0/4}))[\xi_u + o(u^{-\delta_0/8})]\Big| \iota, \tau, w, y, z\right].$$

We start with the first expectation in (4.43). Plugging in the lower bound for $(1 - \rho_1 - \rho_2)K + \rho_1 K_1$ derived in (4.34), we have

$$E^Q\left[\frac{1}{[(1 - \rho_1 - \rho_2)K + \rho_1 K_1]^2}; \mathcal{E}_b, \mathcal{L}, \mathcal{A}_\tau < 0 \Big| \iota = 1, \tau\right]$$

$$= O(1)u^{d-2}e^{-u_{t*}^2}$$

$$(4.45) \qquad \times \int_{\mathcal{A}_\tau < 0} \gamma_{1,u}(u\sigma \mathcal{A}_\tau)$$

$$\times \exp\Big\{-2(1 + \lambda_1)u\mathcal{A}_\tau$$

$$-\frac{1}{2}\Big[y^\top y + \frac{|w - \mu_{20}\mu_{22}^{-1}z|^2}{1 - \mu_{20}\mu_{22}^{-1}\mu_{02}} + z^\top \mu_{22}^{-1}z\Big]\Big\} d\mathcal{A}_\tau \, dy \, dz.$$

We deal with the $\gamma_{1,u}(u\sigma \mathcal{A}_\tau)$ term in the above integration. On the set $\mathcal{L}$, $u\sigma \mathcal{A}_\tau > -u^{3/2+\epsilon}$. By Lemma 24, for $-u^{3/2+\epsilon} < x < 0$, there exists a constant $\delta^* > 0$ such that

$$E\left[\frac{1}{\rho_1^2 \kappa_{1,2}^2}; x > (1 + o(u^{-1-\delta_0/4}))[\xi_u + o(u^{-\delta_0/8})]\Big| \iota, \tau, w, y, z, C_1\right]$$

$$= O(1)\rho_1^{-2}e^{u^{\delta^*}x}$$

and

$$E\left[\frac{1}{(1 - \rho_1 - \rho_2)^2 \kappa_{2,1}^2}; x > (1 + o(u^{-1-\delta_0/4}))[\xi_u + o(u^{-\delta_0/8})]\Big| \iota, \tau, w, y, z, C_2\right]$$

$$= O(1)(1 - \rho_1 - \rho_2)^{-2}(1 - \lambda)^{-d}e^{u^{\delta^*}x}.$$

Therefore, the above approximations and the dominated convergence theorem imply that conditionally on $\mathcal{L}$,

$$\int_{\mathcal{A}_\tau < 0} \gamma_{1,u}(u\sigma \mathcal{A}_\tau)e^{-2(1+\lambda_1)u\mathcal{A}_\tau} d\mathcal{A}_\tau = O(1) \cdot \max\{\rho_1^{-2}, (1 - \lambda)^{-2d}\} \cdot u^{-1-\delta^*}.$$

Thus, (4.45) equals

$$(4.45) = O(1)\max\{\rho_1^{-2}, (1 - \lambda)^{-2d}\} \cdot u^{-1-\delta^*} \cdot u^{d-2}e^{-u_{t*}^2}.$$

Taking expectation of the above equation with respect to ι and τ and choosing ρ_1, ρ_2 and $1 - \lambda$ be $(\log\log b)^{-1}$, we have

$$(4.46) \quad E\left[\frac{1}{[(1 - \rho_1 - \rho_2)K + \rho_1 K_1]^2}; \mathcal{E}_b, \mathcal{L}, \mathcal{A}_\tau < 0, \iota = 1\right] = o(1)v^2(b).$$

For the second term in (4.43), with the same bound of $\gamma_{1,u}$, we have

$$
E^Q\left[\frac{1}{[(1-\rho_1-\rho_2)K+\rho_1 K_1]^2}; \mathcal{E}_b, \mathcal{L}, -\eta/u_\tau < \mathcal{A}_\tau < 0 \Big| \iota = 0, \tau\right]
$$

$$
= O(1)u^{d-2}e^{-u_{t*}^2}
$$

$$
\times u_\tau \int_{-\eta/u_\tau < \mathcal{A}_\tau < 0} \gamma_{1,u}(u\sigma\mathcal{A}_\tau)e^{-2(1+\lambda_1)u\mathcal{A}_\tau}e^{-\lambda u_\tau \mathcal{A}_\tau}
$$

$$
\times \exp\left\{-\frac{1}{2}\left[\frac{|\mu_{20}\mu_{22}^{-1}z|^2}{1-\mu_{20}\mu_{22}^{-1}\mu_{02}}+\left|\mu_{22}^{-1/2}z-\frac{\mu_{22}^{1/2}\mathbf{1}}{2\sigma}\right|^2\right]\right.
$$

$$
\left. -\frac{1-\lambda}{2}y^\top y\right\} d\mathcal{A}_\tau\, dy\, dz
$$

$$
= O(1)\cdot \max\{\rho_1^{-2}, (1-\lambda)^{-2d}\}\cdot u^{-\delta^*}\cdot u^{d-2}e^{-u_{t*}^2}
$$

$$
= o(1)v^2(b),
$$

and similarly for the third term in (4.43),

$$
E^Q\left[\frac{1}{[(1-\rho_1-\rho_2)K+\rho_1 K_1]^2}; \mathcal{E}_b, \mathcal{L}, \mathcal{A}_\tau \le -\eta/u_\tau \Big| \iota = 0, \tau\right]
$$

$$
= O(1)\rho_1\cdot u^{d-2}e^{-u_{t*}^2}u_\tau
$$

$$
\times \int_{\mathcal{A}_\tau < -\eta/u_\tau} \gamma_{1,u}(u\sigma\mathcal{A}_\tau)
$$

$$
\times e^{-2(1+\lambda_1)u\mathcal{A}_\tau}
$$

$$
(4.47) \qquad \times \exp\left\{\lambda_1 u_\tau \mathcal{A}_\tau\right.
$$

$$
-\frac{1}{2}\left[\frac{|\mu_{20}\mu_{22}^{-1}z|^2}{1-\mu_{20}\mu_{22}^{-1}\mu_{02}}+\left|\mu_{22}^{-1/2}z-\frac{\mu_{22}^{1/2}\mathbf{1}}{2\sigma}\right|^2\right]
$$

$$
\left. -\frac{1+\lambda_1}{2}y^\top y\right\} d\mathcal{A}_\tau\, dy\, dz
$$

$$
= O(1)\rho_1\cdot \max\{\rho_1^{-2}, (1-\lambda)^{-2d}\}\cdot u^{-\delta^*}\cdot u^{d-2}e^{-u_{t*}^2}
$$

$$
= o(1)v^2(b).
$$

We put all the estimates in (4.40), (4.42), (4.46) and (4.47) back to (4.35). For any $\varepsilon > 0$, if we choose $\eta = \rho_1 = \rho_2 = 1 - \lambda = 1/\log\log b$, then for b sufficiently large we have that

$$
E^Q\left[\left(\frac{dP}{dQ}\right)^2; \mathcal{E}_b, \mathcal{L}\right] \le (1+3\varepsilon)v^2(b).
$$

We complete the proof of Theorem 3 for the case that $\mu(t) \neq 0$.

4.4. *Case* 2: *Constant mean function.*   The proof when $\mu(t) \equiv 0$ is very similar, except that we need to consider two situations: first, τ is not close to the boundary of T and otherwise. More precisely, for a given $\delta' > 0$ small enough, we consider the case when $\tau \in \{t : |t - \tau| \leq u^{-1/2+\delta'}\} \subset T$ and otherwise.

For the first situation, τ is "far away" from the boundary of T, which is the important case, the derivation is same as that of the case where $\mu(t)$ is not a constant. For the case in which τ is within $u^{-1/2+\delta'}$ distance from the boundary of T, the contribution of the boundary case is $o(v^2(b))$. An intuitive interpretation is that the important region of the integral $\mathcal{I}(T)$ might be cut off by the boundary of T. Therefore, in cases that τ is too close to the boundary, the tail $\mathcal{I}(T)$ is not heavier than that of the interior case. The rigorous analysis is basically repeating the parts 1, 2 and 3 on a truncated region. Therefore, we omit the details.

5. Proof of Theorem 7.   The proof of Theorem 7 is analogous to that of Theorem 3. According to Lemma 18, we focus on the set (for some small $\epsilon_0 > 0$)

$$(5.1) \qquad \mathcal{L}_* = \mathcal{L} \cap \Big\{ \sup_{|t-\tau|>2u^{-1/2+\epsilon}} g(t) - \epsilon_0 u |t|^2 < 0 \Big\}.$$

A similar three-part procedure is applied here.

In part 1, using the transformation from f to the process f_*, we have

$$\beta_u(T) = \sup_{t \in T} \Big\{ f(t) + \frac{\mathbf{1}^\top \bar{f}_t''}{2\sigma u_t} + \frac{B_t}{u_t} + \mu_\sigma(t) \Big\}$$

$$= \sup_{t \in T} \Big\{ f_*(t) + u_\tau C(t - \tau) + \frac{\mathbf{1}^\top (z_t - u_t \mu_{02} + u_\tau \mu_2(t - \tau))}{2\sigma u_t}$$

$$+ \frac{B_t}{u_t} + \mu_\sigma(t) \Big\}.$$

We insert the expansions in (4.10), (4.11) and (4.12) into the expression of $\beta_u(T)$ and obtain that $\beta_u(T)$ equals

$$\sup_{t \in T} \Big\{ w + y^\top (t - \tau) + \frac{1}{2}(t - \tau)^\top z(t - \tau) + R_f(t - \tau) + g(t - \tau)$$

$$+ u_\tau \Big(1 - \frac{1}{2}(t - \tau)^\top (t - \tau) + C_4(t - \tau) + R_C(t - \tau) \Big)$$

$$+ \mu_\sigma(\tau) + \partial \mu_\sigma(\tau)^\top (t - \tau) + \frac{1}{2}(t - \tau)^\top \Delta \mu_\sigma(\tau)(t - \tau) + \sigma^{-1} R_\mu(t - \tau)$$

$$+ \frac{\mathbf{1}^\top (z_t - u_t \mu_{02} + u_\tau \mu_2(t - \tau))}{2\sigma u_t} + \frac{B_t}{u_t} \Big\}$$

$$= \sup_{t \in T} \left\{ u + w + \frac{1}{2} \tilde{y}^\top (uI - \tilde{\mathbf{z}})^{-1} \tilde{y} \right.$$

$$- \frac{1}{2} \left(t - \tau - (uI - \tilde{\mathbf{z}})^{-1} \tilde{y} \right)^\top (uI - \tilde{\mathbf{z}}) \left(t - \tau - (uI - \tilde{\mathbf{z}})^{-1} \tilde{y} \right)$$

$$+ u_\tau C_4 (t - \tau) + R(t - \tau) + g(t - \tau)$$

$$\left. + \frac{\mathbf{1}^\top (z_t - u_t \mu_{02} + u_\tau \mu_2 (t - \tau))}{2\sigma u_t} + \frac{B_t}{u_t} \right\}.$$

Note that the above display is approximately a quadratic function of $t - \tau$ and is maximized approximately at $t - \tau = (uI - \tilde{\mathbf{z}})^{-1} \tilde{y}$. In addition, on the set $\mathcal{L}_*$, we have that $|\tau - t_*| < 2u^{-1/2+\epsilon}$ and thus $\tilde{y} = y + O(u^{-1/2+\epsilon})$. Therefore, on the set $\mathcal{L}_*$, we have the following approximation of $\beta_u(T)$:

$$\mathcal{A}_\tau + \inf_{|t-\tau|<2u^{-1/2+\epsilon}} g(t) \leq \beta_u(T) - u + u^{-1-\delta_0/4} o(|w| + |y| + |z|)$$

$$\leq \mathcal{A}_\tau + \sup_{|t-\tau|<2u^{-1/2+\epsilon}} g(t).$$

Thus, we obtain the same representation as in part 1 in the proof of Theorem 3.

Since we use the same change of measure, the analysis of the likelihood ratio is exactly the same as part 2 of Theorem 3. For part 3, we compute the second moment of dP/dQ on the set $\{\beta_u(T) > u\}$. This is also identical to the proof of Theorem 3. Thus, with the same choice of tuning parameters, we have that

$$E^Q \left[\left(\frac{dP}{dQ} \right)^2; \beta_u(T) > u \right] \leq (1 + \varepsilon) v^2(b).$$

Additionally, Lemma 18 provides an approximation that $P(\beta_u(T) > u) \sim v(b)$. Thus, we use Lemma 13 (presented at the beginning of Section 4) and complete the proof.

6. Proof of Theorem 10. For the bias control, we need the following result [44].

PROPOSITION 15. *Suppose that conditions* C1–C6 *are satisfied. Let* $F'(x)$ *be the probability density function of* $\log \mathcal{I}(T) = \log \int_T e^{\sigma f(t) + \mu(t)} \, dt$. *Then the following approximation holds as* $x \to \infty$:

$$F'(x) \sim \sigma^{-2} x \cdot v(e^x).$$

Thus, for any small ε,

$$(6.1) \qquad P\big(b < \mathcal{I}(T) < b(1 + \varepsilon / \log b) | \mathcal{I}(T) > b\big) = \Theta(\varepsilon).$$

Similar to the log-normal distribution, the overshoot of $\mathcal{I}(T)$ is $\Theta(b/\log b)$. Note that

$$\big| v_M(b) - v(b) \big| \leq P\big(\mathcal{I}(T) > b, \mathcal{I}_M(T) < b \big) + P\big(\mathcal{I}(T) < b, \mathcal{I}_M(T) > b \big).$$

Let

$$\mathcal{L}_\varepsilon = \Big\{ \sup_{t \in T} \big|\partial f(t)\big| \leq 2\big(1 - u^{-2}\log\varepsilon\big)u \Big\}.$$

Note that $\partial f(t)$ is a d-dimensional Gaussian process. Using Borel–TIS lemma, we obtain that

$$P\big(\mathcal{L}_\varepsilon^c\big) = o(1)\varepsilon \cdot v(b).$$

Therefore, it is sufficient to control $P(\mathcal{I}(T) > b, \mathcal{I}_M(T) < b, \mathcal{L}_\varepsilon)$ and $P(\mathcal{I}(T) < b, \mathcal{I}_M(T) > b, \mathcal{L}_\varepsilon)$.

By the definition of $\mathcal{I}_M$ in (3.23), there exists a constant $c_1 > 0$ such that

$$\Delta = \big|\mathcal{I}(T) - \mathcal{I}_M(T)\big| \leq \sum_{i=1}^{M} \bigg| \int_{T_N(t_i)} e^{\sigma f(t) + \mu(t)}\, dt - \mathrm{mes}\big(T_N(t_i)\big) \cdot e^{\sigma f(t_i) + \mu(t_i)} \bigg|$$

$$\leq c_1 \min\{\mathcal{I}_M(T), \mathcal{I}(T)\} \cdot \sup_{t \in T} \big|\partial f(t)\big|/N.$$

Then we have, on the set $\mathcal{L}_\varepsilon$, $\Delta \leq 2c_1 \min\{\mathcal{I}_M(T), \mathcal{I}(T)\}(1 - u^{-2}\log\varepsilon)u/N$, which implies that

$$P\big(\mathcal{I}(T) > b, \mathcal{I}_M(T) < b, \mathcal{L}_\varepsilon\big) \leq P\big(b < \mathcal{I}(T) < b\big(1 + 2\big(1 - u^{-2}\log\varepsilon\big)u/N\big)\big)$$

$$= O(1)\frac{u\big(1 - u^{-2}\log\varepsilon\big)\log b}{N} v(b).$$

The last step is due to the result of Proposition 15 and further (6.1). Thus, it is sufficient to choose $N = O(\varepsilon^{-1-\varepsilon_0} u^{2+\varepsilon_0})$ so that the above probability is bounded by $\varepsilon v(b)$. The bound of $P(\mathcal{I}(T) < b, \mathcal{I}_M(T) > b, \mathcal{L}_\varepsilon)$ is completely analogous.

7. Proof of Theorem 11. The proof of Theorem 11 is similar to that of Theorem 3. Therefore, we only lay out the key steps. The only difference is that we replace the integral by a finite sum over T_N. Recall that the proof of Theorem 3 consists of three parts: first, we write the event $\{\mathcal{I}_M(T) > b\}$ as a function of (w, y, z) (with an ignorable correction term); second, we write the likelihood ratio as a function of (w, y, z) (with an ignorable correction term); third, we integrate the likelihood ratio with respect to (ι, τ, w, y, z). For the current proof, we also have three similar parts.

Part 1. For the first step in the proof of Theorem 3, we write $\mathcal{I}(T) > b$ if and only if $\mathcal{A}_\tau + \frac{o(|w|+|y|+|z|+1)}{u^{1+\delta_0/4}} > u^{-1}\sigma^{-1}\xi_u$. With the current discretization size, as proved in Theorem 10,

$$\log\mathcal{I}(T) - \log\mathcal{I}_M(T) = o\big(u^{-1-\varepsilon_0/2}\big).$$

Thus, we reach the same result that $\mathcal{I}_M(T) > b$ if $\mathcal{A}_\tau + \frac{o(|w|+|y|+|z|)}{u^{1+\delta_0/4}} + o(u^{-1-\varepsilon_0/2}) > u^{-1}\sigma^{-1}\xi_u$.

Part 2. Consider the likelihood ratio

$$\frac{dQ}{dP} = \int_T \left[(1 - \rho_1 - \rho_2) l(t) \, \mathrm{LR}(t) + \rho_1 l(t) \, \mathrm{LR}_1(t) + \frac{\rho_2}{\mathrm{mes}(T)} \, \mathrm{LR}_2(t) \right] dt.$$

Under the discretization setup, we have

$$\frac{dQ_M}{dP} = \frac{1 - \rho_1 - \rho_2}{\kappa} \sum_{i=1}^M l(t_i) \, \mathrm{LR}(t_i) + \frac{\rho_1}{\kappa} \sum_{i=1}^M l(t_i) \, \mathrm{LR}_1(t_i) + \rho_2 \sum_{i=1}^M \frac{1}{M} \, \mathrm{LR}_2(t_i),$$

which is a discrete approximation of dQ/dP. In the proof of Theorem 3, after taking all the terms not consisting of t out of the integral [such as that in (4.23)], the discrete sum is essentially approximating the following integral:

$$\int_{|t-\tau| < u^{-1+\delta'}} e^{-((1-\lambda)(u_\tau + \zeta_u)/2)(t - \tau - (uI - \tilde{\mathbf{z}})^{-1}\tilde{y})^\top (uI - \tilde{\mathbf{z}})(t - \tau - (uI - \tilde{\mathbf{z}})^{-1}\tilde{y})} \, dt.$$

The above integral concentrates on a region of size $O(u^{-1})$. Given that we choose $N > u^2$, the discretized likelihood ratio in dQ_M/dP approximate dQ/dP up to a constant in the sense that

$$(7.1) \qquad\qquad\qquad \frac{dQ_M}{dP} = \Theta(1) \frac{dQ}{dP}.$$

Part 3. With the results of parts 1 and 2, the analysis of part 3 is completely analogous to part 3 in the proof of Theorem 3. Thus, we conclude that

$$E^{Q_M}(\tilde{L}_b^2) \leq \kappa_1 v(b)^2,$$

where the constant κ_1 depends on the $\Theta(1)$ in (7.1).

APPENDIX: THE LEMMAS

In this section, we state all the lemmas used in the previous sections. To facilitate reading, we move several lengthy proofs (Lemmas 17, 18, 20, 22, 23 and 24) to the supplemental materials [45], as those proofs are not particularly related to the proof of the theorems and mostly involve tedious elementary algebra.

The first lemma is known as the Borel–TIS lemma, which was proved independently by [20, 23].

LEMMA 16 (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 $E(\sup_{\mathcal{U}} f(t)) < \infty$, 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 $\sigma_{\mathcal{U}}^2 = \max_{t \in \mathcal{U}} \mathrm{Var}[f(t)]$.

LEMMA 17. *Conditionally on the set $\mathcal{L}$ as defined in* (4.13), *we have that*

$$E^Q\left[\left(\frac{dP}{dQ}\right)^2; \mathcal{I}(T) > b, \mathcal{L}^c\right] = o(1)v^2(b).$$

LEMMA 18. *On the set $\mathcal{L}_*$ as defined in* (5.1), *we have that for $k = 1$ and 2*

$$E^Q\left[\left(\frac{dP}{dQ}\right)^k; \beta_u(T) > u, \mathcal{L}_*^c\right] = o(1)P(\beta_u(T) > u)^k.$$

In addition, we have the approximation $P(\beta_u(T) > u) \sim v(b)$.

LEMMA 19. *Let ξ_u be as defined in* (4.19), *then there exist small constants $\delta^*, \lambda', \lambda'' > 0$ such that for all $x > 0$ and sufficiently large u*

$$P(|\xi_u| > x) \le e^{-\lambda' u^{\delta^*} x^2} + e^{-\lambda'' u^2}.$$

PROOF. For $\delta < \delta_0/10$, we split the expectation into two parts $\{|\tilde{S}| \le u^\delta\}$ and $\{|\tilde{S}| > u^\delta, \tau + (uI - \mathbf{z})^{-1/2}\tilde{S} \in T\}$. Note that $|S| \le \kappa u^\delta$ and $g(t)$ is a mean zero Gaussian random field with $\mathrm{Var}(g(t)) = O(|t|^{4+\delta_0})$. A direct application of the Borel–TIS inequality (Lemma 16) yields the result of this lemma. □

LEMMA 20. *Let S be a random variable taking values in $\{s : (uI - \mathbf{z})^{-1/2}s + \tau \in T\}$ with density proportional to*

$$e^{-(\sigma/2)(s-(uI-\mathbf{z})^{-1/2}\tilde{y})^\top(s-(uI-\mathbf{z})^{-1/2}\tilde{y})}.$$

If $|y| \le u^{1/2+\epsilon}$ and $|z| \le u^{1/2+\epsilon}$ and $\epsilon \ll \delta_0$, then

$$\log\{Ee^{\sigma(u-\mu_\sigma(\tau))C_4((uI-\mathbf{z})^{-(1/2)}S)+\sigma R((uI-\mathbf{z})^{-(1/2)}S)}\}$$

$$= \frac{1}{8\sigma u}\sum_i \partial_{iiii}^4 C(0) + \frac{o(|w| + |y| + |z| + 1)}{u^{1+\delta_0/4}},$$

where the expectation is taken with respect to S.

LEMMA 21.

$$\log(\det(I - u^{-1}\mathbf{z})) = -u^{-1}\mathrm{Tr}(\mathbf{z}) + \frac{1}{2}u^{-2}\mathbf{I}_2(\mathbf{z}) + o(u^{-2}),$$

where Tr is the trace of a matrix, $\mathbf{I}_2(\mathbf{z}) = \sum_{i=1}^d \lambda_i^2$, 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 $\mathrm{Tr}(\mathbf{z}) = \sum_{i=1}^d \lambda_i$. □

LEMMA 22. *On the set $\mathcal{L}$, $\mathcal{I}_2$ defined as in* (4.21) *can be written as*

$$\int_{A^*, |t-\tau| < u^{-1+\delta'}} \exp\left\{ \frac{u_{t_*}(t-t_*)^\top \Delta\mu_\sigma(t_*)(t-t_*)}{2} + \frac{u_t^2}{2} \right\} \times u_t$$

$$\times \exp\left\{ (1-\lambda)u_t[w_t + u_\tau C(t-\tau) - u_t] \right.$$

$$\left. + \frac{(1-\lambda)}{2\sigma}\mathbf{1}^\top(z_t - \mu_{02}u_t + \mu_2(t-\tau)u_\tau) - \lambda B_t - \frac{\mathbf{1}^\top \mu_{22}\mathbf{1}}{8\sigma^2} \right\}$$

$$\times \exp\left\{ ((w_t + u_\tau C(t-\tau) - u_t)^2 - 2(w_t + u_\tau C(t-\tau) - u_t) \right.$$

$$\times \mu_{20}\mu_{22}^{-1}(z_t - \mu_{02}u_t + \mu_2(t-\tau)u_\tau))$$

$$\left. /(2(1 - \mu_{20}\mu_{22}^{-1}\mu_{02})) \right\} dt.$$

LEMMA 23. *For $\eta = 1/\log\log b$, on the set $\mathcal{L}$, if $\mathcal{A}_\tau \geq 0$, then*

$$\{|t-\tau| \leq u^{-1+\delta'}\} \subseteq A^*.$$

LEMMA 24. *On the set $\mathcal{L}$, there exists some $\delta^* > 0$ such that for all $-u^{3/2+\epsilon} < x < 0$,*

$$E\left[\frac{1}{\rho_1^2 \kappa_{1,2}^2}; x > (1 + o(u^{-1-\delta_0/4}))[\xi_u + o(u^{-\delta_0/8})] \Big| \iota, \tau, w, y, z, C_1 \right]$$

$$= O(1)\rho_1^{-2} e^{u^{\delta^*}x},$$

$$E\left[\frac{1}{(1-\rho_1-\rho_2)^2 \kappa_{2,1}^2}; x > (1 + o(u^{-1-\delta_0/4}))[\xi_u + o(u^{-\delta_0/8})] \Big| \iota, \tau, w, y, z, C_2 \right]$$

$$= O(1)(1-\rho_1-\rho_2)^{-2}(1-\lambda)^{-d} e^{u^{\delta^*}x},$$

where $C_1 = \{\text{mes}(A^c \cap D) \geq \text{mes}(A \cap D)\}$ and $C_2 = \{\text{mes}(A^c \cap D) < \text{mes}(A \cap D)\}$.

SUPPLEMENTARY MATERIAL

Supplement to "On the conditional distributions and the efficient simulations of exponential integrals of gaussian random fields" (DOI: 10.1214/13-AAP960SUPP; .pdf). Proofs of Proposition 14 and Lemmas 17, 18, 20, 22, 23 and 24 are provided in the supplementary material.

REFERENCES

[1] ABERG, S. and GUTTORP, P. (2008). Distribution of the maximum in air pollution fields. *Environmetrics* **19** 183–208. MR2420185

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

[3] ADLER, R. J., BLANCHET, J. H. and LIU, J. (2012). Efficient Monte Carlo for high excursions of Gaussian random fields. *Ann*. *Appl*. *Probab*. **22** 1167–1214. MR2977989

[4] ADLER, R. J., SAMORODNITSKY, G. and TAYLOR, J. E. (2013). High level excursion set geometry for non-Gaussian infinitely divisible random fields. *Ann*. *Probab*. **41** 134–169. MR3059195

[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. *Zeitschrift für Nationalokonomie—Journal of Economics* **38** 105–118.

[8] ASMUSSEN, S. (2000). *Ruin Probabilities*. *Advanced Series on Statistical Science & Applied Probability* **2**. World Scientific, River Edge, NJ. MR1794582

[9] ASMUSSEN, S. and KLÜPPELBERG, C. (1996). Large deviations results for subexponential tails, with applications to insurance risk. *Stochastic Process*. *Appl*. **64** 103–125. MR1419495

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

[11] 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

[12] 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

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

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

[15] 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

[16] BLANCHET, J. and LIU, J. (2010). Efficient importance sampling in ruin problems for multidimensional regularly varying random walks. *J*. *Appl*. *Probab*. **47** 301–322. MR2668490

[17] BLANCHET, J. and LIU, J. (2012). Efficient simulation and conditional functional limit theorems for ruinous heavy-tailed random walks. *Stochastic Process*. *Appl*. **122** 2994–3031. MR2931349

[18] BLANCHET, J. H. and LIU, J. (2008). State-dependent importance sampling for regularly varying random walks. *Adv*. *in Appl*. *Probab*. **40** 1104–1128. MR2488534

[19] BLANCHET, J. H., LIU, J. and YANG, X. (2010). Monte Carlo for large credit portfolios with potentially high correlations. In *Proceedings of the 2010 Winter Simulation Conference*, Baltimore, MD.

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

[21] 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.

[22] 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

[23] 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

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

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

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

[27] DEUTSCH, H. P. (2004). *Derivatives and Internal Models*, 3rd ed. Palgrave Macmillan, Basingstoke, UK.

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

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

[30] DUPUIS, P. and ELLIS, R. S. (1997). *A Weak Convergence Approach to the Theory of Large Deviations*. Wiley, New York. MR1431744

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

[32] GADRICH, T. and ADLER, R. J. (1993). Slepian models for nonstationary Gaussian processes. *J. Appl. Probab.* **30** 98–111. MR1206355

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

[34] GRIGORIU, M. (1989). Reliability of Daniels systems subject to quasistatic and dynamic nonstationary Gaussian load processes. *Prob. Eng. Mech.* **4** 128–134.

[35] HÜSLER, J. (1990). Extreme values and high boundary crossings of locally stationary Gaussian processes. *Ann. Probab.* **18** 1141–1158. MR1062062

[36] HÜSLER, J., PITERBARG, V. and ZHANG, Y. (2011). Extremes of Gaussian processes with random variance. *Electron. J. Probab.* **16** 1254–1280. MR2827458

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

[38] LEADBETTER, M. R., LINDGREN, G. and ROOTZÉN, H. (1983). *Extremes and Related Properties of Random Sequences and Processes*. Springer, New York. MR0691492

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

[40] LINDGREN, G. (1970). Some properties of a normal process near a local maximum. *Ann. Math. Statist.* **41** 1870–1883. MR0272051

[41] LINDGREN, G. (1979). Prediction of level crossings for normal processes containing deterministic components. *Adv. in Appl. Probab.* **11** 93–117. MR0517553

[42] LIU, J. (2012). Tail approximations of integrals of Gaussian random fields. *Ann. Probab.* **40** 1069–1104. MR2962087

[43] LIU, J. and XU, G. (2012). Some asymptotic results of Gaussian random fields with varying mean functions and the associated processes. *Ann. Statist.* **40** 262–293. MR3014307

[44] LIU, J. and XU, G. (2013). On the density functions of integrals of Gaussian random fields. *Adv. in Appl. Probab.* **45** 398–424. MR3102457

[45] LIU, J. and XU, G. (2014). Supplement to "On the conditional distributions and the efficient simulations of exponential integrals of gaussian random fields." DOI:10.1214/13-AAP960SUPP.

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

[47] MITZENMACHER, M. and UPFAL, E. (2005). *Probability and Computing*: *Randomized Algorithms and Probabilistic Analysis*. Cambridge Univ. Press, Cambridge. MR2144605

[48] NARDI, Y., SIEGMUND, D. O. and YAKIR, B. (2008). The distribution of maxima of approximately Gaussian random fields. *Ann. Statist.* **36** 1375–1403. MR2418661

[49] 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

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

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

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

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

[54] TRAUB, J. F., WASILKOWSKI, G. W. and WOŹNIAKOWSKI, H. (1988). *Information-based Complexity*. Academic Press, Boston, MA. MR0958691

[55] WOŹNIAKOWSKI, H. (1997). Computational complexity of continuous problems. In *Nonlinear Dynamics*, *Chaotic and Complex Systems* (*Zakopane*, 1995) 283–295. Cambridge Univ. Press, Cambridge. MR1481726

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

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

DEPARTMENT OF STATISTICS
COLUMBIA UNIVERSITY
1255 AMSTERDAM AVENUE
NEW YORK, NEW YORK 10027
USA
E-MAIL: jcliu@stat.columbia.edu

SCHOOL OF STATISTICS
UNIVERSITY OF MINNESOTA
367 FORD HALL, 224 CHURCH ST. SE
MINNEAPOLIS, MINNESOTA 55455
USA
E-MAIL: xuxxx360@umn.edu