Proceedings of the 2014 Winter Simulation Conference
A. Tolk, S. Y. Diallo, I. O. Ryzhov, L. Yilmaz, S. Buckley, and J. A. Miller, eds.

UNIFORMLY EFFICIENT SIMULATION FOR TAIL PROBABILITIES OF GAUSSIAN RANDOM FIELDS

Gongjun Xu

School of Statistics
University of Minnesota
Minneapolis, MN 55455, USA

ABSTRACT

In this paper, we consider rare-event simulation of the tail probabilities of Gaussian random fields. In particular, we design importance sampling estimators that are uniformly efficient for a family of Gaussian random fields with different mean and variance functions.

1 INTRODUCTION

Consider a mean zero Gaussian random field $\{f(t) : t \in T\}$ with unit variance living on a d-dimensional compact set $T \subset R^d$, that is, for every finite subset of $\{t_1, ..., t_n\} \subset T$, $(f(t_1), ..., f(t_n))$ is a mean zero multivariate Gaussian random vector with identity covariance matrix. Let $\mu(t) \in R$ and $\sigma(t) \in (0, \infty)$ be (deterministic) continuous functions. We are interested in the probability

$$w_{\sigma,\mu}(b) = P\left(\sup_{t\in T}\{\sigma(t)f(t) + \mu(t)\} > b\right), \text{ as } b \to \infty.$$

The mean and variance functions $\mu(t)$ and $\sigma^2(t)$ are unspecified and known only to be in certain ranges. In particular, we assume in this paper that for all $t \in T$, $\mu(t) \in [\mu_l, \mu_u]$ and $\sigma^2(t) \in [\sigma_l^2, \sigma_u^2]$.

The extremes of Gaussian random fields have wide applications in finance, spatial analysis, physical oceanography, and many other disciplines (Adler, Müller, and Rozovskiĭ 1996, Adler, Taylor, and Worsley 2010). Tail probabilities of the extremes have been extensively studied in the literature, with its focus mostly on the development of approximations and bounds for the suprema (Borell 1975, Tsirelson, Ibragimov, and Sudakov 1976, Piterbarg 1996, Sun 1993, Azais and Wschebor 2008, Adler and Taylor 2007). Tail probabilities of other convex functions of Gaussian random process have also been studied. For instance, Liu (2012) and Liu and Xu (2012b) derived the asymptotic approximations of the tail probabilities of the exponential integrals of Gaussian random fields; see also Liu and Xu (2013).

Most of the sharp theoretical approximations developed in the literature hold only for constant variance fields, which also need certain smoothness conditions of the Gaussian random fields (Adler and Taylor 2007, Adler, Blanchet, and Liu 2012). For the case of less smooth fields, the approximations involve the unknown Pickands' constants (Piterbarg 1996). Therefore, to evaluate the tail probabilities, rare-event simulation serves as an appealing alternative from a computational point of view. In particular, the design and the analysis do not require very sharp approximations of the tail probabilities. Importance sampling based efficient simulation procedures have been proposed in the literature to estimate the tail probabilities. Numerical methods for rare-event analysis of the suprema are studied in Adler, Blanchet, and Liu (2008) and more thoroughly in Adler, Blanchet, and Liu (2012); see also Azaïs and Wschebor (2009), Li and Liu (2013). Simulation study for the exponential integrals of the Gaussian random fields has been studied in Liu and Xu (2012a), Liu and Xu (2013), Liu and Xu (2014).

978-1-4799-7486-3/14/$31.00 ©2014 IEEE

To design asymptotically efficient importance sampling estimator, one needs to construct a change of measure that is tailored to a specific event and such a measure usually depends on the Gaussian random fields mean and variance functions. However, in applications, one is often interested in estimating many probabilities for a certain range of mean and variance parameter values. For instance, portfolio credit risk management may require the estimation of the tail probabilities of extremes for a family of Gaussian random fields. These problems motivate our study. We focus on the problem of simultaneous efficient estimation of $w_{\sigma,\mu}(b)$ for all possible $\mu(t) \in [\mu_l, \mu_u]$ and $\sigma^2(t) \in [\sigma_l^2, \sigma_u^2]$.

The remainder of the paper is organized as follows. In Section 2 we introduce some notions of efficiency and computational complexity under the setting of rare-event simulation. Section 3 provides the construction of our importance sampling estimator and shows the main properties of our algorithm. Some numerical simulations are conducted in Section 4 and detailed proofs of our main theorems are given in Section 5.

2 Efficiency of Rare-event Simulation and Importance Sampling

We first introduce some general notions of rare-event simulations. Given that the tail probability $w_{\sigma,\mu}(b)$ converges to zero, it is more meaningful to consider the relative error of a Monte Carlo estimator L_b with respect to $w_{\sigma,\mu}(b)$. This is because a trivial estimator $L_b^* \equiv 0$ has an error $|L_b^* - w_{\sigma,\mu}(b)| = w_{\sigma,\mu}(b) \to 0$. In the literature of rare-event simulation, one usually employs the concept of *weak efficiency* or *logarithmic efficiency* as an efficiency criterion (Asmussen and Glynn 2007).

Definition 1 An estimator L_b is said to be *weakly efficient* or *logarithmic efficient* in estimating $w_{\sigma,\mu}(b)$ if $EL_b = w_{\sigma,\mu}(b)$ and

$$\limsup_{b\to\infty} \frac{Var(L_b)}{w_{\sigma,\mu}(b)^{2-\varepsilon}} = 0, \tag{1}$$

for all $\varepsilon > 0$.

Weak efficiency is a popular efficiency criterion in the rare-event simulation (Asmussen and Glynn 2007). Suppose that we want to estimate $w_{\sigma,\mu}(b)$ with certain relative accuracy, that is, to compute an estimator Z_b such that for some prescribed $\varepsilon, \delta > 0$,

$$P\left(\left|Z_b/w_{\sigma,\mu}(b) - 1\right| > \varepsilon\right) < \delta. \tag{2}$$

If a crude Monte Carlo simulation method is used, then it requires $n = O(\varepsilon^{-2}\delta^{-1}w_{\sigma,\mu}(b)^{-1})$ i.i.d. replicates. By the Borell-TIS lemma (Lemma 6), we know $w_{\sigma,\mu}(b) = \exp\{-(1+o(1))b^2/(2\sup_{t\in T}\sigma(t)^2)\}$. Therefore n has an exponential rate in b^2. Suppose that a weakly efficient estimator of $w_{\sigma,\mu}(b)$ has been obtained, denoted by L_b. Let $\{L_b^{(j)} : j = 1, ..., n\}$ be n 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 $Var^{1/2}(L_b)/n^{1/2}w_{\sigma,\mu}(b)$. A direct application of Chebyshev's inequality yields

$$P(|Z_b/w_{\sigma,\mu}(b) - 1| \geq \varepsilon) \leq \frac{Var(L_b)}{n\varepsilon^2 w_{\sigma,\mu}^2(b)}.$$

Thus, if L_b is a weakly efficient estimator, it suffices to simulate $n = o(\varepsilon^{-2}\delta^{-1}w_{\sigma,\mu}(b)^{-\varepsilon'})$ (for $\varepsilon' > 0$) i.i.d. replicates of L_b to achieve the accuracy in (2). Compared with the crude Monte Carlo simulation, weakly efficient estimators substantially reduce the computational cost.

To construct weakly efficient estimators, importance sampling is a commonly used method for the variance reduction. In particular, we have

$$w_{\sigma,\mu}(b) = E\left[1_{(\sup_{t\in T}\{\sigma(t)f(t)+\mu(t)\}>b)}\right] = E^Q\left[\frac{dP}{dQ}1_{(\sup_{t\in T}\{\sigma(t)f(t)+\mu(t)\}>b)}\right],$$

where $1_{(\cdot)}$ denotes the indicator function, Q is a probability measure such that dP/dQ is well defined on the set $\{\sup_{t\in T}\{\sigma(t)f(t)+\mu(t)\} > b\}$, and we use E and E^Q to denote the expectations under the measures P and Q, respectively. Then, the random variable defined by

$$L_{\sigma,\mu,b} = \frac{dP}{dQ}1_{(\sup_{t\in T}\{\sigma(t)f(t)+\mu(t)\}>b)} \tag{3}$$

is an unbiased estimator of $w_{\sigma,\mu}(b)$ under the measure Q. To have an efficient estimator, we want to choose $Q(\cdot)$ to be a good approximation of $P_b^*(\cdot) := P(\cdot|\sup_{t\in T}\{\sigma(t)f(t)+\mu(t)\} > b)$, the conditional probability distribution given $\sup_{t\in T}\{\sigma(t)f(t)+\mu(t)\} > b$.

The design of the new measure Q usually depends on the Gaussian random fields mean and variance functions, μ and σ. As a consequence, the designed measure Q that gives a weakly efficient estimator $L_{\sigma,\mu,b} = \frac{dP}{dQ}1_{(\sup_{t\in T}\{\sigma(t)f(t)+\mu(t)\}>b)}$ for $w_{\sigma,\mu}(b)$ may not be efficient any more for estimating $w_{\sigma',\mu'}(b)$, where $\sigma'(t)$ and $\mu'(t)$ are two different variance and mean functions. That is, the corresponding importance sampling estimator based on Q

$$L_{\sigma',\mu',b} := \frac{dP}{dQ}1_{(\sup_{t\in T}\{\sigma'(t)f(t)+\mu'(t)\}>b)}$$

may not be a weakly efficient estimator for $w_{\sigma',\mu'}(b)$. See Section 3.1.1 for more details.

In applications, one is often interested in estimating many probabilities for a certain range of mean and variance parameter values. This motivates us to construct an estimator L that is weakly efficient for a family of functions μ and σ. In particular, in this paper we consider μ and σ satisfying the following condition:

C1. Functions μ and σ are differentiable on the compact set T. For all $t \in T$, $\mu(t) \in [\mu_l, \mu_u]$ and $\sigma^2(t) \in [\sigma_l^2, \sigma_u^2]$.

We introduce the following uniform efficiency criterion.

Definition 2 For all μ and σ satisfying condition C1, we say

$$L_{\sigma,\mu,b} = \frac{dP}{dQ}1_{(\sup_{t\in T}\{\sigma(t)f(t)+\mu(t)\}>b)}$$

is uniformly weakly efficient if $EL_{\sigma,\mu,b} = w_{\sigma,\mu}(b)$ and

$$\limsup_{b\to\infty}\frac{Var(L_{\sigma,\mu,b})}{w_{\sigma,\mu}(b)^{2-\varepsilon}} = 0 \text{ for all } \varepsilon > 0 \tag{4}$$

uniformly with respect to μ and σ.

In the literature, a similar uniform efficiency definition has been proposed in Glasserman and Juneja (2008) to design an algorithm that is asymptotically efficient uniformly for a family of probability sets when estimating the tail probabilities of sums of light tailed random variables. The random variable parameters are assumed known in their case.

3 Uniform Efficient Simulation Algorithm

3.1 Discrete case

We start with the case when T contains finite points. We assume $T := \{t_1, \cdots, t_M\}$. We illustrate in Section 3.1.1 that a weakly efficient estimator may not be uniformly efficient. We then propose our procedure in Section 3.1.2.

3.1.1 A non-uniformly efficient estimator

For known μ and σ, Adler, Blanchet, and Liu (2012) proposed the following simulation procedure.

1. Simulate a random variable $\tau \in \{t_1, \cdots, t_M\}$ according to the following probability measure:
$$P(\tau_b = t_i) = \frac{P(\sigma(t_i)f(t_i) + \mu(t_i) > b)}{\sum_{j=1}^{M} P(\sigma(t_j)f(t_j) + \mu(t_j) > b)}. \tag{5}$$
2. Given the realized τ, simulate $f(\tau)$ conditional on $\sigma(\tau)f(\tau) + \mu(\tau) > b$.
3. Given $(\tau, f(\tau))$, simulate the rest $\{f(t) : t \neq \tau, t \in T\}$ from the original conditional distribution under P.

Let Q^* be the measure with respect to the above sampling procedure. We have

$$\frac{dQ^*}{dP} = \frac{\sum_{i=1}^{M} I_{(\sigma(t_i)f(t_i)+\mu(t_i)>b)}}{\sum_{i=1}^{M} P(\sigma(t_i)f(t_i) + \mu(t_i) > b)},$$

Adler, Blanchet, and Liu (2012) shows that $L_{\sigma,\mu,b} = \frac{dP}{dQ^*} 1_{(\sup_{t\in T}\{\sigma(t)f(t)+\mu(t)\}>b)}$ is an efficient estimator for $w_{\sigma,\mu}(b)$.

For different mean and variance functions, μ' and σ', we show that

$$L_{\sigma',\mu',b} := \frac{dP}{dQ^*} 1_{(\sup_{t\in T}\{\sigma'(t)f(t)+\mu'(t)\}>b)}$$

may not be an weakly efficient estimator for $w_{\sigma',\mu'}(b)$. For simplicity we first consider a special case when $\mu' = \mu$ and the variance function σ' satisfying $\sigma'(t) \leq \sigma(t)$ and $\max_{t_i\in T} \sigma'(t_i) < \max_{t_i\in T} \sigma(t_i)$. Then, under measure Q^*, if $\max_{t_i\in T}\sigma'(t_i)f(t_i) + \mu(t_i) > b$, $\sigma(t_i)f(t_i) + \mu(t_i) > b$ always happens and the change of measure is well defined. We have

$$\begin{aligned}
& E^Q\left[\left(\frac{dP}{dQ^*}\right)^2; \max_{t_i\in T}\sigma'(t_i)f(t_i) + \mu(t_i) > b\right] \\
= \; & E^Q\left[\left(\frac{\sum_{i=1}^{M} P(\sigma(t_i)f(t_i) + \mu(t_i) > b)}{\sum_{i=1}^{M} I_{(\sigma(t_i)f(t_i)+\mu(t_i)>b)}}\right)^2; \max_{t_i\in T}\sigma'(t_i)f(t_i) + \mu(t_i) > b\right] \\
= \; & E\left[\frac{\sum_{i=1}^{M} P(\sigma(t_i)f(t_i) + \mu(t_i) > b)}{\sum_{i=1}^{M} I_{(\sigma(t_i)f(t_i)+\mu(t_i)>b)}}; \max_{t_i\in T}\sigma'(t_i)f(t_i) + \mu(t_i) > b\right] \\
\geq \; & \frac{1}{M}\left(\sum_{i=1}^{M} P(\sigma(t_i)f(t_i) + \mu(t_i) > b)\right) \times w_{\sigma',\mu}(b) \\
\geq \; & \frac{1}{M}\max_{t_i\in T} P(\sigma(t_i)f(t_i) + \mu(t_i) > b) \times w_{\sigma',\mu}(b) \\
= \; & \exp\left\{-(1+o(1))\frac{b^2}{2\max_{t_i\in T}\sigma(t_i)^2} - (1+o(1))\frac{b^2}{2\max_{t_i\in T}\sigma'(t_i)^2}\right\},
\end{aligned}$$

where we used the approximation that

$$w_{\sigma',\mu}(b) = \exp\left\{-(1+o(1))\frac{b^2}{2\max_{t_i\in T}\sigma'(t_i)^2}\right\}.$$

Then under the assumption that $\max_{t_i\in T}\sigma'(t_i) < \max_{t_i\in T}\sigma(t_i)$, we know for small ε,

$$\frac{E^Q\left[\left(\frac{dP}{dQ^*}\right)^2; \max_{t_i\in T}\sigma'(t_i)f(t_i)+\mu(t_i) > b\right]}{w_{\sigma',\mu}(b)^{2-\varepsilon}} \to \infty.$$

Therefore, the estimator $L_{\sigma',\mu,b} = \frac{dP}{dQ^*}1_{(\sup_{t\in T}\{\sigma'(t)f(t)+\mu(t)\}>b)}$ is not weakly efficient for estimating $w_{\sigma',\mu}(b)$.

On the other hand, for the case when $\max_{t_i\in T}\sigma'(t_i) > \max_{t_i\in T}\sigma(t_i)$, we can see that the change of measure may not be well defined since if $\max_{t_i\in T}\sigma'(t_i)f(t_i)+\mu(t_i) > b$, $\max_{t_i\in T}\sigma(t_i)f(t_i)+\mu(t_i) > b$ may not happen. Since for variance function such that $\sigma(t)\in[\sigma_l,\sigma_u], \forall t\in T$, we can always construct σ' satisfying one of the two cases consider above. Then we have $L_{\sigma',\mu',b} := \frac{dP}{dQ^*}1_{(\sup_{t\in T}\{\sigma'(t)f(t)+\mu'(t)\}>b)}$ is not an uniformly efficient estimator for all mean and variance functions.

3.1.2 Proposed method

In this section we propose a new change of measure which gives an uniformly efficient estimator. We describe the new measure Q in two ways. First, we specify the simulations of f from Q and then provide its Radon-Nikodym derivative with respect to P. Under the measure Q, $f(t)$ is generated according to the following algorithm:

Algorithm 1 For the discrete set $T = \{t_1,\cdots,t_M\}$, the algorithm is as follows:

1. Simulate a random variable ς with respect to some positive continuous density function g on $[\sigma_l, \sigma_u+\delta_b]$. Here we take $\delta_b = ab^{-1}$ for some $a > 0$.
2. Simulate a random variable τ uniformly over $T = \{t_1,\cdots,t_M\}$.
3. Given the realized ς and τ, simulate $f(\tau)$ conditional on $\varsigma f(\tau)+\nu > b$. Here ν is chosen as μ_u.
4. Given $(\tau, f(\tau))$, simulate the Gaussian process $\{f(t): t\neq\tau, t\in T\}$ from the original conditional distribution under P.

For the measure Q defined above, it is not hard to verify that P and Q are mutually absolutely continuous with the Radon-Nikodym derivative being

$$\frac{dQ}{dP} = \int_{\sigma_l}^{\sigma_u+\delta_b}\frac{\sum_{i=1}^M I(\varsigma f(t_i)+\nu > b)}{MP(\varsigma f(t_i)+\nu > b)}g(\varsigma)d\varsigma.$$

This gives importance sampling estimator

$$L_{\sigma,\mu,N}(b) = \left(\int_{\sigma_l}^{\sigma_u+\delta_b}\frac{\sum_{i=1}^M I(\varsigma f(t_i)+\nu > b)}{MP(\varsigma f(t_i)+\nu > b)}g(\varsigma)d\varsigma\right)^{-1} I_{(\sup_{i:t_i\in T_N}\sigma(t_i)f(t_i)+\mu(t_i)>b)}.$$

We show the efficiency of the proposed importance sampling procedure. Note that under Q, if $\max_{t_i\in T}\sigma(t_i)f(t_i)+\mu(t_i) > b$, $\varsigma f(t_i)+\nu > b$ holds for all $\varsigma > \max_{t_i\in T}\sigma(t_i)$ and therefore the change of

measure is well defined. For all μ and σ satisfying C1, we have

$$\begin{aligned}
& E^Q[L_{\sigma,\mu,N}(b)^2] \\
= \; & E^Q\left[\left(\int_{\sigma_l}^{\sigma_u+\delta_b} \frac{\sum_{i=1}^M I(\varsigma f(t_i)+\nu>b)}{MP(\varsigma f(t_i)+\nu>b)} g(\varsigma)d\varsigma\right)^{-2}; \max_{t_i\in T}\sigma(t_i)f(t_i)+\mu(t_i)>b\right] \\
\leq \; & M^2E^Q\left[\left(\int_{\max_{t_i\in T}\sigma(t_i)}^{\max_{t_i\in T}\sigma(t_i)+\delta_b} \frac{g(\varsigma)}{P(\varsigma f(t_i)+\nu>b)} d\varsigma\right)^{-2}; \max_{t_i\in T}\sigma(t_i)f(t_i)+\mu(t_i)>b\right] \\
= \; & O(1)M^2E^Q\left[\left(\int_{\max_{t_i\in T}\sigma(t_i)}^{\max_{t_i\in T}\sigma(t_i)+\delta_b} be^{\frac{(b-\nu)^2}{2\varsigma^2}} g(\varsigma)d\varsigma\right)^{-2}; \max_{t_i\in T}\sigma(t_i)f(t_i)+\mu(t_i)>b\right] \\
= \; & O(1)M^2e^{-(1+o(1))\frac{(b-\nu)^2}{\max_{t_i\in T}\sigma(t_i)^2}} Q(\max_{t_i\in T}\sigma(t_i)f(t_i)+\mu(t_i)>b) \\
= \; & O(1)M^2e^{-(1+o(1))\frac{(b-\nu)^2}{\max_{t_i\in T}\sigma(t_i)^2}}.
\end{aligned}$$

Therefore, we have for all ε

$$\limsup_{b\to\infty} \frac{E^Q[L^2; \max_{t_i\in T}\sigma(t_i)f(t_i)+\mu(t_i)>b]}{w_{\sigma,\mu}(b)^{2-\varepsilon}} = 0$$

holds uniformly for all μ and σ satisfying C1. This gives the uniformly weak efficiency.

Remark 2 From the above derivation, we can see that for any positive continuous density function g and any $\delta_b = ab^{-1}, a>0$, the importance sampling estimator is uniformly weakly efficient.

The parameter δ_b in the algorithm is introduced to bound the second moment of the importance sampling estimator. Otherwise, consider the case of constant variance $\sigma\in[\sigma_l,\sigma_u]$ and zero mean $\mu=0$. Then for σ taking the value of σ_u, the second moment of the corresponding estimator $L_{\sigma_u,N}(b)$ is lower bounded by

$$\begin{aligned}
& E^Q[L_{\sigma_u,N}(b)^2] \\
= \; & E\left[\left(\int_{\sigma_l}^{\sigma_u} \frac{\sum_{i=1}^M I(\varsigma f(t_i)>b)}{MP(\varsigma f(t_i)>b)} g(\varsigma)d\varsigma\right)^{-1}; \max_i \sigma_u f(t_i)>b\right] \\
\geq \; & P(\sigma_l f(0)>b)P(\max_i \sigma_u f(t_i)>b)E\left[\left(\int_{\sigma_l}^{\sigma_u} I(\max_i f(t_i)>b/\varsigma)g(\varsigma)d\varsigma\right)^{-1} \Big| \max_i f(t_i)>b/\sigma_u\right].
\end{aligned}$$

However, the conditional expectation cannot be controlled and we have the estimator $L_{\sigma_u,N}(b)$ is not efficient for $\sigma=\sigma_u$.

To achieve stronger efficiency results, we may choose g minimizing the variance function of the estimator. In addition, the parameter ν may also be randomly sampled from a distribution on $[\mu_l,\mu_u]$. We leave these issues for future study.

3.2 Continuous case

Direct simulation of a continuous random field is typically not a feasible task, and the change of measure proposed in the previous subsection is not directly applicable. Thus, we use a discrete object to approximate the continuous fields for the implementation. The bias caused by the discretization must be well controlled relative to $w_{\sigma,\mu}(b)$.

We create a regular lattice covering T in the following way. Let $G_{N,d}$ be a countable subset of R^d

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

That is, $G_{N,d}$ is a regular lattice on R^d. Furthermore, let

$$T_N = G_{N,d} \cap T, \tag{6}$$

which is the sub-lattice intersecting with T. Since T is compact, T_N is a finite set. We enumerate the elements in $T_N = \{t_1, \cdots, t_M\}$, where $M = O(N^d)$. Let

$$w_{\sigma,\mu,N}(b) = P\left(\sup_{t_i \in T_N} \sigma(t_i) f(t_i) + \mu(t_i) > b\right).$$

We use $w_{\sigma,\mu,N}(b)$ as a discrete approximation of $w_{\sigma,\mu}(b)$.

We estimate $w_{\sigma,\mu,N}(b)$ by importance sampling, which is based on the change of measure proposed in the discrete case. In particular we define Q_N and P_N as the discrete versions (on T_N) of Q and P respectively. Then dQ_N/dP_N takes the form:

$$\frac{dQ_N}{dP_N} = \int_{\sigma_l}^{\sigma_u+\delta_b} \frac{\sum_{i=1}^M I(\varsigma f(t_i) + \nu > b)}{MP(\varsigma f(t_i) + \nu > b)} g(\varsigma) d\varsigma. \tag{7}$$

Note that here M depends on N and goes to infinity as $N \to \infty$. This gives importance sampling estimator

$$L_{\sigma,\mu,N}(b) = \left(\int_{\sigma_l}^{\sigma_u+\delta_b} \frac{\sum_{i=1}^M I(\varsigma f(t_i) + \nu > b)}{MP(\varsigma f(t_i) + \nu > b)} g(\varsigma) d\varsigma\right)^{-1} I_{(\sup_{i:t_i \in T_N} \sigma(t_i) f(t_i) + \mu(t_i) > b)}.$$

We have the next theorem to control the bias of the estimator $L_{\sigma,\mu,N}(b)$.

Theorem 3 Suppose f is a Gaussian random field that is twice differentiable and the functions μ and σ satisfy condition C1. For any $\varepsilon_0 > 0$, there exists constant κ_0 such that for any $\varepsilon \in (0,1)$, if $N > \kappa_0 \varepsilon^{-1-\varepsilon_0} b^{2+\varepsilon_0}$, then for $b > 1$,

$$\frac{\left|w_{\sigma,\mu,N}(b) - w_{\sigma,\mu}(b)\right|}{w_{\sigma,\mu}(b)} < \varepsilon$$

uniformly for all μ and σ satisfying condition C1.

The next theorem controls the variance of the estimator $L_{\sigma,\mu,N}(b)$, whose proof follows from a similar argument as in the discrete case and we omit the details.

Theorem 4 Suppose f is a Gaussian random field that twice differentiable and the functions μ and σ satisfy condition C1. If N is chosen as in Theorem 3, then for any $\varepsilon' > 0$, we have

$$\limsup_{b \to \infty} \frac{E^{Q_N} L^2_{\sigma,\mu,b}}{w^{2-\varepsilon'}(b)} = 0$$

uniformly for all μ and σ satisfying C1.

We simulate n i.i.d. copies of $L_{\sigma,\mu,b}$ via Algorithm 1, $\{L^{(j)}_{\sigma,\mu,b} : j = 1, ..., n\}$, and the averaged estimator is

$$Z_b = \frac{1}{n} \sum_{j=1}^n L^{(j)}_{\sigma,\mu,b}.$$

From the discussion in Section 2 and Theorems 3 and 4, in order to achieve an ε relative error with probability at least $1-\delta$, we need to have $n = O(\varepsilon^{-2}\delta^{-1} w^{-\varepsilon'}(b))$ for any $\varepsilon' > 0$. This holds for all possible σ and μ satisfying condition C1.

Remark 5 In the main theorems, we assume the Gaussian random fields are twice differentiable. The proposed efficient simulation algorithm can be generalized for non-differentiable Gaussian random fields following similar mixture change of measures. A different discretization size may be chosen, and the complexity analysis may be very different. An adaptive discretization procedure as in Li and Liu (2013) may be applied here, and we leave this for future study.

4 Simulation

We use a simple example to illustrate the performance of the proposed method. Consider i.i.d. standard normal random variables $\{f(t_i), i = 1, \cdots, 100\}$. For simplicity, we take $\mu = 0$ and σ constant. We want to compute probabilities $P(\sigma \max_i f(t_i) > b)$ for $\sigma \in [0.3, 1]$ and $b = 3$. This is equivalent to simulating $P(\max_i f(t_i) > b)$ for all $b \in [3, 10]$. The following tables display the simulation results for $\sigma = 0.3, 0.6$ and 1.

The estimated tail probabilities $w_\sigma(b)$ along with the estimated standard deviations $sd^Q(L_{\sigma,b}) = \sqrt{Var^Q(L_{\sigma,b})}$ are shown in Table 1. All the results are based on 10^4 independent simulations. We also give the theoretical values of the tail probabilities. We can see for different σ values, the estimates are close to the true values.

In Table 2 we show the results from the algorithm in Section 3.1.1, where the change of measure is constructed based on $\sigma = 1$. Compared with the proposed method, we can see that the estimate is more efficient when σ value is equal to the design value 1 and less efficient for other σ values. In particular, when $\sigma = 0.3$, it gives 0 estimate value.

Table 1: Estimates of $w_\sigma(b)$, $sd^Q(L_{\sigma,b})$, and $sd^Q(L_{\sigma,b})/w_\sigma(b)$. All results are based on 10^4 independent simulations and thus the standard errors of the estimates are $sd^Q(L_{\sigma,b})/100$.

	Est.	$sd^Q(L_{\sigma,b})$	$sd^Q(L_{\sigma,b})$/Est.	Theoretical Value
$\sigma = 0.3$	7.55e-22	5.33e-21	7.05	7.62e-22
$\sigma = 0.6$	2.93e-05	1.33e-04	4.52	2.87e-05
$\sigma = 1$	1.26e-01	5.92e-01	4.69	1.26e-01

Table 2: Estimates based on the algorithm in Section 3.1.1.

	Est.	$sd^Q(\tilde{L}_b)$	$sd^Q(\tilde{L}_b)$/Est.
$\sigma = 0.3$	0	0	NA
$\sigma = 0.6$	1.35e-05	1.35e-03	1.00e+02
$\sigma = 1$	1.26e-01	2.32e-02	1.84e-01

5 Proof of Theorem 3

The following lemma is known as the Borell-TIS lemma, which is proved independently by Borell (1975) and Tsirelson, Ibragimov, and Sudakov (1976).

Lemma 6 (Borell-TIS) Let $f(t)$, $t \in \mathscr{U}$, $\mathscr{U}$ is a parameter set, be a mean zero Gaussian random field. f is almost surely bounded on $\mathscr{U}$. Then, $E[\sup_{\mathscr{U}} f(t)] < \infty$, and

$$P\left(\sup_{t\in\mathscr{U}} f(t) - E[\sup_{t\in\mathscr{U}} f(t)] \geq b\right) \leq \exp\left(-\frac{b^2}{2\sigma^2_{\mathscr{U}}}\right),$$

where $\sigma^2_{\mathscr{U}} = \sup_{t\in\mathscr{U}} \mathrm{Var}[f(t)]$.

The Borell-TIS lemma provides a general bound of the tail probability. In most cases, $E[\sup_t f(t)]$ is much smaller than b. Thus, for b that is sufficiently large, the tail probability can be further bounded by:

$$P\Big(\sup_{t\in\mathscr{U}} f(t) > b\Big) \leq \exp\left(-\frac{b^2}{4\sigma^2_{\mathscr{U}}}\right).$$

The following lemma provides an upper bound of the density function of $\sup_{t\in T} \sigma(t)f(t) + \mu(t)$, whose proof follows from Ehrhard's Inequality (Ehrhard 1983); see also Chapter 4 in Bogachev (1998).

Lemma 7 Under the conditions of Theorem 3, let $F'_{\sigma,\mu}(x)$ be the probability density function of $\sup_{t\in T}\sigma(t)f(t)+\mu(t)$. Then, $F'_{\sigma,\mu}(x)$ exists almost everywhere. Moreover, as x goes to infinity,

$$F'_{\sigma,\mu}(x) = (1+o(1))\Big\{\sup_{t\in T}\sigma(t)\Big\}^{-1} x w_{\sigma,\mu}(x). \tag{8}$$

For some $A>0$, let

$$\mathscr{L} = \Big\{\sup_{t\in T}|f(t)| \le A(1-b^{-2}\log\varepsilon)b, \quad \sup_{t\in T}|\partial f(t)| \le A(1-b^{-2}\log\varepsilon)b\Big\}. \tag{9}$$

Here ∂ denote the gradient and $|\cdot|$ is the L_1 norm. Note that $\partial f(t)$ is a d-dimensional Gaussian random field and the variance function of each component is bounded on T. Then by Lemma 6, there exists $\lambda>0$ such that for A sufficiently large (independent of u and ε)

$$\begin{aligned} P(\mathscr{L}^c) &\le P\Big(\sup_{t\in T}|f(t)| > A(1-b^{-2}\log\varepsilon)b\Big) + P\Big(\sup_{t\in T}|\partial f(t)| > A(1-b^{-2}\log\varepsilon)b\Big) \\ &\le \exp\Big(-\lambda A^2(1-b^{-2}\log\varepsilon)^2 b^2\Big) \\ &= o(1)\varepsilon w_{\sigma,\mu}(b). \end{aligned}$$

We need to control the estimation bias caused by the discretization. In particular, we have

$$\begin{aligned} |w_{\sigma,\mu}(b) - w_{\sigma,\mu,N}(b)| &= \left|P\left(\sup_{t\in T}\sigma(t)f(t)+\mu(t) > b\right) - P\left(\sup_{i:t_i\in T_N}\sigma(t_i)f(t_i)+\mu(t_i) > b\right)\right| \\ &\le P\left(\sup_{t\in T}\sigma(t)f(t)+\mu(t) > b, \sup_{i:t_i\in T_N}\sigma(t_i)f(t_i)+\mu(t_i) < b, \mathscr{L}\right) + o(1)\varepsilon w_{\sigma,\mu}(b). \end{aligned}$$

On the set $\mathscr{L}$, we know

$$\sup_{t\in T,|t-t_i|<N^{-1}}|\sigma(t)f(t)+\mu(t)-(\sigma(t_i)f(t_i)+\mu(t_i))| \le \frac{1}{N}\{O(1)\sup_{t\in T}|\partial f(t)| + O(1)\sup_{t\in T}|f(t)| + O(1)\}.$$

Together with Lemma 7, this implies that

$$|w_{\sigma,\mu}(b) - w_{\sigma,\mu,N}(b)| \le \frac{1}{N}\{O(1)\sup_{t\in T}|\partial f(t)| + O(1)\sup_{t\in T}|f(t)| + O(1)\} \times b w_{\sigma,\mu}(b) + o(1)\varepsilon w_{\sigma,\mu}(b).$$

Thus it is sufficient to choose $N = O(\varepsilon^{-1-\varepsilon_0}b^{2+\varepsilon_0})$ so that

$$|w_{\sigma,\mu}(b) - w_{\sigma,\mu,N}(b)| \le \varepsilon w_{\sigma,\mu}(b).$$

This completes our proof.

REFERENCES

Adler, R., J. Blanchet, and J. Liu. 2008, December. "Efficient Simulation for Tial Probabilities of Gaussian Random Fields". In *Proceedings of the 2008 Winter Simulation Conference*, edited by S. J. Mason, R. R. Hill, L. Moench, O. Rose, T. Jefferson, and J. W. Fowler, 328–336. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Adler, R., J. Blanchet, and J. Liu. 2012. "Efficient Monte Carlo for Large Excursions of Gaussian Random Fields". *Annals of Applied Probability* 22 (3): 1167–1214.

Adler, R., and J. Taylor. 2007. *Random fields and geometry*. Springer.

Adler, R. J., P. Müller, and B. L. Rozovskiĭ. 1996. *Stochastic modelling in physical oceanography*. Springer.

Adler, R. J., J. E. Taylor, and K. J. Worsley. 2010. *Applications of random fields and geometry: Foundations and case studies*. Available at http://webee. technion.ac.il/people/adler/hrf.pdf.

Asmussen, S., and P. Glynn. 2007. *Stochastic Simulation: Algorithms and Analysis*. New York, NY, USA: Springer.

Azais, J. M., and M. Wschebor. 2008. "A general expression for the distribution of the maximum of a Gaussian field and the approximation of the tail". *Stochastic Processes and Their Applications* 118 (7): 1190–1218.

Azaïs, J.-M., and M. Wschebor. 2009. *Level sets and extrema of random processes and fields*. John Wiley & Sons.

Bogachev, V. 1998. *Gaussian Measures*. Mathematical Surveys and Monographs Volume 62. American Mathematical Society.

Borell, C. 1975. "The Brunn-Minkowski inequality in Gauss space". *Inventiones Mathematicae* 30 (2): 207–216.

Ehrhard, A. 1983. "Symétrisation dans l'space de Gauss". *Math. Scand.* 53:281–301.

Glasserman, P., and S. Juneja. 2008. "Uniformly efficient importance sampling for the tail distribution of sums of random variables". *Mathematics of Operations Research* 33 (1): 36–50.

Li, X., and J. Liu. 2013. "Rare-event Simulation and Efficient Discretization for the Supremum of Gaussian Random Fields". *Technical Report (arXiv:1309.7365)*.

Liu, J. 2012. "Tail Approximations of Integrals of Gaussian Random Fields". *Annals of Probability* 40 (3): 1069–1104.

Liu, J., and G. Xu. 2012a. "Rare-event simulations for exponential integrals of smooth Gaussian processes". In *Proceedings of the Winter Simulation Conference*. Winter Simulation Conference.

Liu, J., and G. Xu. 2012b. "Some Asymptotic Results of Gaussian Random Fields with Varying Mean Functions and the Associated Processes". *Annals of Statistics* 40:262–293.

Liu, J., and G. Xu. 2013. "On the Conditional Distributions and the Efficient Simulations of Exponential Integrals of Gaussian Random Fields". *Annals of Applied Probability* to appear. Available at http://arxiv.org/abs/1204.5546.

Liu, J., and G. Xu. 2014. "Efficient Simulations for the Exponential Integrals of Hölder Continuous Gaussian Random Fields". *The ACM Transactions on Modeling and Computer Simulation* 24:9:1–9:24.

Piterbarg, V. I. 1996. *Asymptotic methods in the theory of Gaussian processes and fields*. Providence, R.I.: American Mathematical Society.

Sun, J. Y. 1993. "Tail probabilities of the maxima of gaussian random-fields". *Annals of Probability* 21 (1): 34–71.

Tsirelson, B., I. Ibragimov, and V. Sudakov. 1976. "Norms of Gaussian sample functions". *Proceedings of the Third Japan-USSR Symposium on Probability Theory (Tashkent, 1975)* 550:20–41.

AUTHOR BIOGRAPHIES

GONGJUN XU is an Assistant Professor in the School of Statistics at the University of Minnesota, Twin-Cities. Gongjun received his Ph.D. in statistics from Columbia University in May, 2013. His research interests include Monte Carlo methods and rare event simulation. His email address is xuxxx360@umn.edu.