MONTE CARLO FOR LARGE CREDIT PORTFOLIOS WITH POTENTIALLY HIGH CORRELATIONS

Jose H. Blanchet

Jingchen Liu

Department of Industrial Engineering & Operations Research
Columbia University
New York, NY

Department of Statistics
Columbia University
New York, NY

Xuan Yang

Department of Statistics
Columbia University
New York, NY

ABSTRACT

In this paper we develop efficient Monte Carlo methods for large credit portfolios. We assume the default indicators admit a Gaussian copula. Therefore, we are able to embed the default correlations into a continuous Gaussian random field, which is capable of incorporating an infinite size portfolio and potentially highly correlated defaults. We are particularly interested in estimating the expectations, such as the expected number of defaults given that there is at least one default and the expected loss given at least one default. All these quantities turn out to be closely related to the geometric structure of the random field. We will heavily employ random field techniques to construct importance sampling based estimators and provide rigorous efficiency analysis.

1   INTRODUCTION

1.1   Modeling credit defaults using latent random field structure

Consider a credit portfolio of size n. Let Y_k be the default indicator (1 for default and 0 otherwise) for the k-th obligor, for $k = 1, ..., n$. We further let the vector $Y = (Y_1, ..., Y_n)$ admit a Gaussian copula. Equivalently, Y has the following representation. There exists a latent multivariate Gaussian random vector $X = (X_1, ..., X_n) \sim N(\mu, \Sigma)$ and a threshold u such that $Y_k = I(X_k > u)$ for all $k = 1, ..., n$. Of interest in this paper are the probability of observing at least one default, namely,

$$w_n(u) = P\left(\sum_{k=1}^{n} Y_k > 0\right) = P\left(\sup_{k=1}^{n} X_k > u\right),$$

and also associated conditional expectations, such as the expected number of defaults

$$N_n(u) = E\left(\sum_{k=1}^{n} Y_k \,\Big|\, \sum_{k=1}^{n} Y_k > 0\right) = E\left(\sum_{k=1}^{n} I(X_k > u) \,\Big|\, \sup_{k=1}^{n} X_k > u\right).$$

In addition, we attach a deterministic amount of loss, l_k, to each default and consider the expected total loss conditional upon at least one default,

$$l_n(u) = E\left(\sum_{k=1}^{n} l_k Y_k \,\Big|\, \sum_{k=1}^{n} Y_k > 0\right) = E\left(\sum_{k=1}^{n} l_k I(X_k > u) \,\Big|\, \sup_{k=1}^{n} X_k > u\right).$$

We are interested in designing algorithms that are applicable to large portfolios (i.e. arbitrarily large n) and to situations in which $w_n(u) \approx 0$. In order to accommodate this situation we embed the random vector $(X_1, ..., X_n)$ inside a Gaussian random field, f, living in a continuous domain T. More precisely, we assume that there exists a subset $\{t_1, ..., t_n\} \subset T$ such that $(f(t_1), ..., f(t_n))$ has the same distribution as $(X_1, ..., X_n)$. Further, we denote the mean and covariance function

of f by

$$\mu(t) = Ef(t), \qquad C(s,t) = Cov(f(s), f(t)).$$

With a random field representation, studying $w_n(u)$, $N_n(u)$, and $l_n(u)$ involves investigating the behavior of the random field on the finite set. One of the advantages of embedding a finite multivariate Gaussian random vector inside a continuous field is that the latter object is capable of incorporating large portfolios and also potentially highly correlated defaults. Motivated by this situation we shall study suitable continuous analogues to $w_n(u)$ and $N_n(u)$. The algorithms that we shall propose have straightforward adaptations to the discrete case. The complexity properties that we shall derive carry over uniformly in n assuming that the embedding is in place. For a thorough study of extremes of multivariate Gaussian random variable and discrete Gaussian random field, see Asmussen and Rojas-Nandayapa (2008), Adler, Blanchet, and Liu (2008).

Let us turn to our continuous formulation. We impose the global Hölder condition, that is there exist κ, δ, $\beta > 0$ such that

$$E(f(s) - f(t))^2 \leq \kappa |t - s|^{2\beta}, \tag{1}$$

for all $|t_1 - t_2|, |s_1 - s_2| < \delta$. Typically, coefficient β is in the interval $(0, 1]$. In addition, we assume that $f(t)$ is almost surely continuous (with respect to t). Note that the almost sure continuity is genuinely one additional condition and cannot be implied from continuity conditions of $\mu(t)$ and $C(s,t)$.

Typically, one chooses T to be a d dimensional compact subset of $\mathbb{R}^d$. In fact, for most illustrative examples, T is chosen to be $[0,1]^d$, even though the theory developed in this paper is able to tackle much more complicated sets. There are two advantages of embedding a portfolio default pattern into a continuous random field. First, being a subset of $\mathbb{R}^d$ with non-empty interior, we can essentially embed an infinite dimension multivariate Gaussian random vector. Therefore, a random field living on a continuous domain is capable of modeling a large size portfolio. Second, due to the global Hölder condition and the continuity assumption of f, as $s \to t$, $f(s)$ and $f(t)$ tend to be "perfectly" correlated. Therefore, this embedding provides means to model highly correlated default patterns.

In the setting of a continuous field, we will use $t \in T$ for the index and therefore write $Y(t) = I(f(t) > u)$ as the default indicator. The analogues of the interesting quantities are,

$$w(u) = P\left(\sup_T f(t) > u\right), \qquad \text{and } N(u) = E\left(\int_T Y(t)dt \,\Big|\, \sup_T f(t) > u\right). \tag{2}$$

$w(u)$ corresponds to the probability of observing one default ($w_n(u)$) and $N(u)$ corresponds to the conditional average number of defaults ($N_n(u)/n$). We introduce additional notations. We define $A_u = \{t \in T : f(t) > u\}$ and let $m_{Leb}(\cdot)$ be the Lebesgue measure. Then, we have

$$N(u) = E\left(m_{Leb}(A_u) \,\Big|\, \sup_T f(t) > u\right). \tag{3}$$

In addition, the expected number of default can be viewed as a special case of the "expected loss" with unit loss associated to each default. To define the "expected loss", we introduce a continuous version of the l_i's. We define a measure, $L(\cdot)$, on T with density function so that

$$L(A) = \int_A loss(t)dt.$$

$loss(t)$ is the loss associated with default $Y(t)$. Then, we define the continuous analogue of the expected loss as

$$l(u) = E\left(\int loss(t)Y(t)dt \,\Big|\, \sup_T f(t) > u\right) = E\left(L(A_u) \,\Big|\, \sup_T f(t) > u\right). \tag{4}$$

For technical convenience, we assume that there exists $\delta_0 > 0$ such that $loss(t) \in (\delta_0^{-1}, \delta_0)$ for all $t \in T$.

In this paper, we consider the case that the default probability is small. Therefore, throughout the discussion, we are in the asymptotic regime with $u \to \infty$ and $\mu(t)$, $C(s,t)$ fixed.

The extremes of Gaussian random field have been extensively studied in literature, with special focus on approximations and bounds for the suprema (Piterbarg (1996), Sun (1993), Azais and Wschebor (2008), Borell (1975), Tsirelson, Ibragimov, and Sudakov (1976)). Also, there are closed form results on the excursion set A_u such as Adler and Taylor (2007). Numerical methods for rare event analysis are studied in Adler, Blanchet, and Liu (2008) and more thoroughly in Adler, Blanchet, and Liu (2010).

This paper is organized as follows. In Section 2, we provide the technical conditions and useful existing results on random fields, rare-event simulation, and importance sampling. In Section 3, we introduce the importance sampling

algorithm, its implementation and efficiency analysis. In Section 4, we provide numerical examples on the efficiency of this proposed algorithm.

2 PROBLEM SETTING AND PRELIMINARIES

Based on the discussion in the previous section, it suffices to study the behavior of a continuous random field. To proceed, we specify some technical conditions of the random field in context:

1. The parameter set T is a d dimensional compact subset of $\mathbb{R}^d$ with a non-empty interior. In addition, there exists a constant $\kappa > 0$, such that for any $t \in T$ and ε small enough,

    $$m_{Leb}(B(t, \varepsilon)) \geq \kappa \varepsilon^d,$$

    where $B(t, \varepsilon) = \{s \in T : |s - t| < \varepsilon\}$ and $|\cdot|$ denotes the L_1 norm.
2. There exist functions $\sigma_i : T \to \mathbb{R}$, for $i = 1, ..., m$, such that the random field has expansion,

    $$f(t) = \mu(t) + \sum_{i=1}^{m} \sigma_i(t) Z_i, \tag{5}$$

    where $Z_1, ..., Z_m$ are i.i.d. standard Gaussian random variables. In addition, we assume that $\mu(t)$ and $\sigma_i(t)$'s are uniformly Hölder continuous in T, that is, for some $\beta \in (0, 1]$

    $$|\mu(s) - \mu(t)| \leq \kappa |s - t|^\beta, \qquad \sup_i |\sigma_i(s) - \sigma_i(t)| < \kappa |s - t|^\beta,$$

    for all $|s - t| < \delta$. Lastly, let $\sigma^2(t) = \sum_{i=1}^{n} \sigma_i^2(t)$ and assume that there exists $\varepsilon_0 > 0$ such that

    $$\varepsilon_0 < \inf_T \sigma(t) \leq \sup_T \sigma(t) < \varepsilon_0^{-1}.$$

Condition 1 is considered as very mild and for technical purposes only. Condition 2 is an appealing feature from the aspects of both practical modeling and theoretical development. From the modeling point of view, $Z_1, ..., Z_m$ can be regarded as independent factors that influence the market. Each factor Z_i has different impact on each different obligor. The magnitude of this impact on the default indicator $Y(t)$ is characterized by $\sigma_i(t)$. Thanks to the Hölder continuity of $\sigma_i(t)$, two obligors tend to share similar impact from the same market factor if their spatial indices $t \in T$ are close to each other. Such a structure has been employed by Glasserman, Kang, and Shahabuddin (2007), Glasserman, Kang, and Shahabuddin (2008) to model credit portfolios. The functional form of $\sigma_i(t)$ can either be from *a priori* knowledge or based on statistical analysis of the real data (e.g. principle component type of analyses). Throughout the current discussion, we assume that the $\sigma_i(t)$'s are known in closed forms.

The existence of expansion (5). From the theoretical point of view, under very mild conditions of the covariance function (compactness and injectivity), Mercer's theorem (c.f. Riesz and Sz-Nagy 1955) guarantees that the random field has the Karhunen-Loève expansion, that is,

$$f(t) = \sum_{i=1}^{\infty} \lambda_i \varphi_i(t) Z_i,$$

where $\varphi_i(t)$'s are eigenfucntions corresponding to eigenvalues λ_i's. In addition, $\varphi_i(t)$'s are standardized orthogonal functions, such that,

$$\int_T \varphi_i(t) \varphi_j(t) dt = \delta_{ij}.$$

At least, one can view (5) as an approximation of the Karhunen-Loève expansion. In addition, Condition 2 does not require that $\sigma_i(t)$ to be orthogonal and is more flexible in some sense, although it does require a finite decomposition.

2.1 Performance measure of rare-event simulation

One of the main challenges in computing $w(u)$, $N(u)$, and $l(u)$ is that they all involves the event $\{\max_{t \in T} f(t) > u\}$ which is rare for large u. The goal of rare event simulation is precisely to cope with this situation. We now review some generic notions of efficiency in rare event estimation which we shall apply to our setting. Consider a sequence of sets B_u, indexed by the rarity parameter u, such that $w(u) = P(B_u) \to 0$ as $u \to \infty$. Usually, the rarity parameter u has

practical meaning. For instance, in the current situation, u is the high level above which the random fields will exceed thereby triggering a default. We want to compute/estimate $w(u)$. For a Monte Carlo estimator Z_u, it is more meaningful to consider the error of Z_u relative to $w(u)$. This is because a trivial estimator $Z_u^* \equiv 0$ has error $|Z_u^* - w(u)| = w(u) \to 0$. Therefore, one usually employs the concept of *weak efficiency* or *asymptotic optimality* as a performance measure, which is defined as follows.

Definition 1. *An estimator Z_u is said to be weakly efficient in estimating $w(u)$ if $EZ_u = w(u)$ and*

$$\lim_{u \to \infty} \frac{\log Var(Z_u)}{\log w^2(u)} = 1.$$

Asymptotic optimality is a popular efficiency criterion in rare event simulation (see Asmussen and Glynn (2007)). Let Z_u be an asymptotically optimal estimator for $w(u)$. It follows by using Chebyshev's inequality that in order to obtain an estimate which has ε relative error with respect to $w(u)$ with probability at least $1 - \delta$, one only needs to generate $n = o(\varepsilon^{-2} \delta^{-1} w^{-\varepsilon'}(u))$ independent replications of Z_u (for any $\varepsilon' > 0$) and take an average of them. In order to consider the total complexity, one also needs to take into account of the complexity in generating one Z_u. In the current context, the computation of a single Z_u can typically be done in constant time in u.

2.2 Importance sampling and computation of conditional expectations

In this paper, we use importance sampling to construct efficient estimators of rare-event probabilities. Importance sampling is based on the following basic identity. Consider two probability measures P and Q on some state space $\mathscr{X}$ with σ-algebra $\mathscr{F}$. If the Radon-Nikodym derivative $\frac{dP}{dQ}(\omega)$ is well defined on set $A \in \mathscr{F}$, then

$$P(A) = \int_A \frac{dP}{dQ}(\omega) Q(d\omega).$$

Note that if one chooses Q^* such that for each $B \in \mathscr{F}$,

$$Q^*(B) = P(B \cap A)/P(A),$$

then, $\frac{dP}{dQ^*} \equiv P(A)$ almost surely on the set A and therefore has zero variance. This implies that the best importance sampling distribution (with zero variance in estimating $P(A)$) is the conditional distribution given the event A occurs under measure P. Certainly, this zero variance estimator is not implementable, because the Radon-Nikodym derivative requires computing $P(A)$, which is the quantity of interest. Nevertheless, it provides general guidelines on how to construct efficient importance sampling estimators – a good approximation of the conditional distribution.

We now briefly explain how importance sampling can be used to estimate conditional expectations such as $N(u)$ and $l(u)$. Fix $K \in (0, \infty)$ and let $\chi(K, q)$ be the class of random variables X satisfying $0 \leq X \leq K$ with

$$E[X | B_u] = \Omega[1/\log(P(B_u))^q].^1 \tag{6}$$

Then, by noting that

$$\frac{E^Q(XL)}{E^Q(L)} = E[X | B_u] = \frac{E[X; B_u]}{P(B_u)}, \tag{7}$$

where $L = \frac{dP}{dQ} I_{B_u}$, it follows easily that an asymptotically optimal estimator can be obtained by constructing the natural estimator for $E[X | B_u]$; i.e. the ratio of the corresponding averaged importance sampling estimators suggested by the ratio in the left of (7).

In the context of Gaussian random fields, we have that $B_u = \{\sup_T f(t) > u\}$ and one is very often interested in random variables X of the form $X = \Gamma(f)$, where $\Gamma : C(T) \to \mathbb{R}$ and $C(T)$ denotes the space of continuous functions on T. Endowing $C(T)$ with the uniform topology, consider functions Γ that are non-negative and bounded by a positive constant. An archetypical example, the total number of default $N(u) = E(m_{Leb}(A_u) | \sup_T f(t) > u)$, is known to satisfy (6) with $K = m_{Leb}(T)$. So it is with the total loss, $l(u) = E(L(A_u) | \sup_T f(t) > u)$ with $K = \varepsilon_0^{-1} m_{Leb}(T)$. There are many other examples of $\chi(K, q)$ with $K = m_{Leb}(T)$ which satisfy (6) for a suitable q, depending on the regularity properties of the field. In fact, if the mean and covariance properties of f are Hölder continuous, then, it is not difficult to see that q can be estimated. In turn, we have that an asymptotically optimal importance sampling algorithm for $w(u)$ would

---

[1] Given h and g positive, we shall use the familiar asymptotic notation $h(x) = O(g(x))$ if there is $c < \infty$ such that $h(x) \leq cg(x)$ for all x large enough; $h(x) = \Omega(g(x))$ if $h(x) \geq cg(x)$ if x is sufficiently large and $h(x) = o(g(x))$ as $x \to \infty$ if $h(x)/g(x) \to 0$ as $x \to \infty$; and $h(x) = \Theta(g(x))$ if $h(x) = O(g(x))$ and $h(x) = \Omega(g(x))$.

typically also yield a polynomial time algorithm for functional characteristics of the conditional field given high level excursions. Given the page limit, we restrict our discussion to the computation of $w(u)$. The estimators for $N(u)$ and $l(u)$ arise naturally once the estimator for $w(u)$ is in force.

3 EFFICIENT SIMULATION OF GAUSSIAN RANDOM FIELD

3.1 A change of measure on $C(T)$

In this section, we focus our attention on the efficient computation of the default probability

$$w(u) = P(\sup_T f(t) > u),$$

as $u \to \infty$. As discussed in the previous section, a change of measure that yields a small variance should be a good approximation of the conditional distribution of random field given the occurrence of the rare event. As discussed in Adler, Blanchet, and Liu (2008), Adler, Blanchet, and Liu (2010), the event $\{\sup_T f(t) > u\}$ for large number u is mostly caused by the fact that there exists one $\tau \in T$ where $f(\tau)$ is particularly large. Therefore, we will construct a change of measure in order to mimic such a feature. In addition, given $\{\sup_T f(t) > u\}$, not all $t \in T$ are equally likely to be large. An intuitive calculation gives that

$$P(f(t) > u | \sup_T f(t) > u) = \frac{P(f(t) > u)}{P(\sup_T f(t) > u)} \propto P(f(t) > u).$$

We interpret the above calculation as, given $\{\sup_T f(t) > u\}$, the probability that $f(t)$ takes a large value is proportional to $P(f(t) > u)$. Based on this, we propose a family of change of measures, Q_γ, indexed by parameter γ, such that the Random-Nykodym derivative is defined as follows,

$$L_\gamma^{-1} = \frac{dQ_\gamma}{dP} = \frac{m_{Leb}(A_\gamma)}{E\left(m_{Leb}(A_\gamma)\right)}, \tag{8}$$

where $A_\gamma = \{t : f(t) > \gamma\}$. It is not hard to verify that L_γ is a properly normalized Radon-Nikodym derivative with respect to $P(\cdot | \sup_T f(t) > \gamma)$. In addition, notice that

$$E\left(m_{Leb}(A_\gamma)\right) = E\left(\int_T I(f(t) > \gamma)dt\right) = \int_T P(f(t) > \gamma)dt.$$

Therefore, the computation of $\frac{dQ_\gamma}{dP}$ only involves pre-simulation evaluation of a finite dimensional integral.

In order to simulate a random field according the measure Q_γ, one follows the next three steps.

1. Simulate a random variable τ_γ according to

$$Q_\gamma(\tau_\gamma \in \cdot) = \frac{E(m_{Leb}(A_\gamma \cap \cdot))}{E(m_{Leb}(A_\gamma))}; \tag{9}$$

equivalently, τ_γ has density function, $p_\gamma(\cdot)$, satisfying

$$p_\gamma(t) = \frac{P(f(t) > \gamma)}{E(m_{Leb}(A_\gamma))}.$$

2. Conditional on τ_γ simulate $f(\tau_\gamma)$ from the conditional distribution of $f(\tau_\gamma)$ given that $f(\tau_\gamma) > \gamma$.
3. Simulate the rest of the field $\{f(t) : t \neq \tau_\gamma\}$ according the original distribution given the realized $f(\tau_\gamma)$.

It is not hard to show that the field generated according to this three steps follows distribution of f under Q_γ (c.f. Adler, Blanchet, and Liu (2010)). Now, we need to choose the tuning parameter γ, which turns out to be crucial in the variance control. Intuitively, the random variable τ_γ indicates the location where a high excursion might occur. A natural and intuitive choice of γ would certainly be $\gamma = u$: it is the irregular behavior of the random field at location τ_γ that triggered the rare event. However, this choice turns out to be a greedy one. If let $\gamma = u$, the estimator of $w(u)$ takes form

$$L_u = \frac{E\left(m_{Leb}(A_u)\right)}{m_{Leb}(A_u)}.$$

Note that we used the fact that $Q_u(\sup_T f(t) > u) = 1$. Unfortunately, it is not hard to construct examples (with $d \geq 2$) showing that L_u has infinite variance under Q_u. An intuitive and empirical explanation to this phenomenon is that by choosing $\gamma = u$ the algorithm is too "greedy". It puts 100% bet on τ_γ for exceeding the target threshold u. Technically, it results in the denominator $m_{Leb}(A_u)$ in the definition of L_u potentially being small. Due to this concern, we propose to choose

$$\gamma_{a,u} = u - a/u. \tag{10}$$

for some $a > 0$. It turns out that the choice of a is rather flexible. The introduction of this "back-up" of magnitude a/u shows a certain amount of reservation that the random field may not be able to exceed level u at location $\tau_{\gamma_{a,u}}$. However, for a given a, it is not hard to derive from the tail probability of Gaussian distribution that the probability $Q_{\gamma_{a,u}}(f(\tau_{\gamma_{a,u}}) \leq u)$ is bounded away from zero and one as $u \to \infty$. Therefore, $f(\tau_{\gamma_{a,u}})$ still has large chance exceeding u though not always. It is helpful to view $\tau_{\gamma_{a,u}}$ as an index that localizes the excursion set A_u though it may not belong to A_u itself.

The change of measure Q_γ is also analytically appealing. The form of the L_γ only involves a deterministic integral, $\int_T P(f(t) > \gamma)dt$, and the volume of the excursion set $m_{Leb}(A_\gamma)$. There exists a substantial literature on the geometric properties of excursion sets. It includes rigorous results such as the Euler characteristic for constant variance twice differentiable fields (e.g. Taylor and Adler (2003)) as well as heuristics. We will later employ the existing bounds and asymptotic results to provide stochastic bounds on the distribution of $m_{Leb}(A_\gamma)$. On the other hand, the three step simulation procedure provides a vivid description of the distribution of the field under Q_γ. As will be shown in later analysis, we will compute the second moment of the estimator by first conditioning on $(\tau_{\gamma_{a,u}}, f(\tau_{\gamma_{a,u}}))$. Note that, conditional on $f(\tau_{\gamma_{a,u}})$ the distribution of $\{f(t) : t \neq \tau_{\gamma_{a,u}}\}$ is another Gaussian random field with computable means and covariance function. Therefore, the analyses of second moment of corresponding estimators are manageable. From now on, to simplify notation, we write $\tau_{\gamma_{a,u}}$ as τ and $Q_{\gamma_{a,u}}$ as Q whenever it does not cause confusion.

3.2 Monte Carlo for fields with finite expansion and efficiency analysis

Direct simulation of a continuous random field is typically not a feasible task on a discrete computer. However, with the structure in (5), one only has to simulate i.i.d. random variables $Z_1, ..., Z_n$ and take a linear combinations with $\sigma_i(t)$'s which are known in closed form. Similarly, simulating a random field under Q turns out to be feasible too. First, one simulate τ according to (9). Then, simulate $f(\tau) = \sum_{i=1}^n \sigma_i(\tau)Z_i$ given $f(\tau) > u - a/u$; and lastly, we simulate the rest field given $f(\tau)$. The last two steps can in fact be merged to one by simulating $Z_1, ..., Z_n$ given that $\sum_{i=1}^n \sigma_i(\tau)Z_i > u - a/u$, which only consists of simulating one truncated univariate normal random variable and a multivariate Gaussian random vector. Here, we want to emphasize that the simulation of a $N(0,1)$ conditional on its being greater than b arbitrarily large can be done efficiently using acceptance rejection method by proposing from $b + Exp(1)/b$. The acceptance probability in fact can be shown to converge to one as $b \to \infty$.

Having solved the implementation issue, we proceed to the efficiency analysis. We first focus on the estimator for $w(u)$. Note that, we intend to control the variance of

$$L_{w,\gamma_{a,u}} = \frac{E\left(m_{Leb}(A_{\gamma_{a,u}})\right)}{m_{Leb}(A_{\gamma_{a,u}})} I(\sup_T f(t) > u). \tag{11}$$

Then, the second moment of the estimator is

$$\begin{aligned}
E^Q\left(L_{w,\gamma_{a,u}}^2\right) &= E^2(m_{Leb}(A_{\gamma_{a,u}}))E^Q\left(\frac{1}{m_{Leb}^2(A_{\gamma_{a,u}})}; \sup_T f(t) > u\right) \\
&= E^2(m_{Leb}(A_{\gamma_{a,u}}))\int_T E^Q\left(\frac{1}{m_{Leb}^2(A_{\gamma_{a,u}})}; \sup_T f(t) > u \,\middle|\, \tau\right) p_{\gamma_{a,u}}(\tau)d\tau \\
&= \left(\int P(f(t) > A_{\gamma_{a,u}})dt\right)^2 \int_T E^Q\left(\frac{1}{m_{Leb}^2(A_{\gamma_{a,u}})}; \sup_T f(t) > u \,\middle|\, \tau\right) p_{\gamma_{a,u}}(\tau)d\tau.
\end{aligned} \tag{12}$$

Before, proceeding to the core analysis, we first introduce a few useful lemmas. The first one is known as B-TIS ineuqality (Borell (1975), Tsirelson, Ibragimov, and Sudakov (1976)).

Lemma 1 (Borell-TIS). *Let $f(\cdot)$ be a real-valued, separable, continuous Gaussian process. Suppose that*

$$\sigma^2(T) = \sup_T Var(f(t)) < \infty.$$

Let $m = E(\sup_T f(t))$. Then, for all b,

$$P\left(\sup_T f(t) > b\right) \leq 2\left(1 - \Phi\left(\frac{b-m}{\sigma(T)}\right)\right),\tag{13}$$

where $\Phi(\cdot)$ is the c.d.f. of the standard Gaussian distribution.

The next lemma uses the B-TIS inequality and one result by Dudley (1973) and improves the bound in the presence of entropy bounds. The proof is given in Chapter 4.1 in Adler and Taylor (2007).

Lemma 2. *Let f be a mean zero almost surely continuous Gaussian random field over T with entropy function N, where $N(\varepsilon)$ is the smallest number of d-balls that cover T and the intrinsic matric d is defined as follows,*

$$d(s,t) = \sqrt{E(f(s) - f(t))^2}.$$

If $N(\varepsilon) \leq K\varepsilon^{-\alpha}$, then for all sufficiently large u

$$P\left(\sup_T f(t) > u\right) \leq C_\alpha u^{\alpha+\eta} e^{-u^2/2\sigma^2(T)},$$

for every $\eta > 0$, where $C_\alpha = C(K, \alpha, \sigma_T^2)$ is a finite constant.

Now, we proceed to the statement of the core lemma, which provides a bound on the volume of the excursion set.

Lemma 3. *Consider a random field f living on T satisfying Conditions 1 and 2. There exists constants $\kappa^*, \kappa_1, \delta^* \in (0, \infty)$, such that for a given $t \in T$, any $z > 0$, and $y > \kappa_1 u^{2d/\beta}$*

$$P(m_{Leb}(A_{\gamma_{a,u}}) < y^{-1}; \sup_T f(t) > u | f(t) = \gamma_{a,u} + z/u) \leq \kappa^* \exp\left(-\delta^* \frac{a^2}{u^2} y^{\beta/d}\right)\tag{14}$$

Proof of Lemma 3. Without loss of generality, we proceed the proof by assuming that $t = 0$. This is simply for notational convenience. Using the decomposition in (2), the covariance function takes the form,

$$C(s,t) = \sum_{i=1}^n \sigma_i(s)\sigma_i(t).$$

To save some notation, we choose κ in Condition 2 large enough such that

$$|C(s_1, t_1) - C(s_2, t_2)| \leq \kappa|t_1 - t_2|^\beta + \kappa|s_1 - s_2|^\beta,$$

for all $|t_1 - t_2|, |s_1 - s_2| < \delta$. Conditional on $f(0) = \gamma_{a,u} + z/u$, $f(t)$ is equal in distribution to

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

where

$$E(t) = \mu(t) + \frac{C(0,t)}{\sigma^2(0)}(\gamma_{a,u} + z/u - \mu(0)),$$

and $g(t)$ is an a.s. continuous mean zero Gaussian random field with with covariance function

$$\Gamma(s,t) = C(s,t) - \sigma^{-1}(0)C(0,s)C(0,t).$$

By simple algebra, we have

$$\Gamma(s,t) \leq 2\kappa_g \sigma(t)\sigma(s)\left(|t|^\beta + |s|^\beta + |t-s|^\beta\right),$$

and

$$d_g(s,t) = \sqrt{E(g(s) - g(t))^2} \leq \kappa_g|s-t|^{\beta/2}.$$

For proof of the above statement on process g, see Lemma 6.9 in Adler, Blanchet, and Liu (2010). We prove the bound in (14) by two cases: $z \geq a$ and $z < a$. We start with the first case that $z \geq a$. For any $y > \kappa_1 u^{2d/\beta}$, we have that for some $\lambda > 0$ (using Condition 1 and simple algebraic manipulations),

$$P\left(m_{Leb}(A_{\gamma_{a,u}}) < y^{-1}; \sup_T f(t) > u | f(0) = \gamma_{a,u} + z/u\right) \leq P\left(\sup_{|t| \leq \lambda y^{-1/d}} E(t) + g(t) < \gamma_{a,u}\right).$$

We choose κ_1 large enough such that for all $y > \kappa_1 u^{2d/\beta}$

$$\sup_{|t| \leq \lambda y^{-1/d}} \left|\mu(t) - \mu(0)\frac{C(0,t)}{\sigma^2(0)}\right| < \frac{z}{4u}, \qquad \sup_{|t| \leq \lambda y^{-1/d}} \left|\frac{C(0,t)}{\sigma^2(0)} - 1\right|\gamma_{a,u} \leq \frac{z}{4u}.$$

Therefore,

$$\sup_{|t| \leq \lambda y^{-1/d}} |E(t) - E(0)| \leq \frac{z}{2u},$$

where $E(0) = \gamma_{\alpha,u} + z/u$. Then, by Lemma 2,

$$P\left(m_{Leb}(A_{\gamma_{a,u}}) < y^{-1}; \sup_T f(t) > u | f(0) = \gamma_{a,u} + z/u\right)$$

$$\leq P\left(\sup_{|t| \leq \lambda y^{-1/d}} -g(t) > \frac{a}{2u}\right)$$

$$\leq \exp\left(-\delta^* \frac{a^2}{u^2} y^{\beta/d}\right),$$

for δ^* sufficiently small. The last inequality is due to Lemma 2. For the second case that $z \in (0, a)$, we let $t^* = \arg\sup_T f(t)$ and define a new field $\xi(s,t) = f(s) - f(t)$ living on $T \times T$. Then, for some $\lambda > 0$ (again using Condition 1)

$$P\left(m_{Leb}(A_{\gamma_{a,u}}) < y^{-1}; \sup_T f(t) > u | f(0) = \gamma_{a,u} + z/u\right)$$

$$\leq P\left(\sup_{|t - t^*| \leq \lambda y^{-1/d}} f(t) < \gamma_{a,u}; f(t^*) > u | f(0) = \gamma_{a,u} + z/u\right)$$

$$\leq P\left(\sup_{|s - t| \leq \lambda y^{-1/d}} \xi(s,t) > a/u | f(0) = \gamma_{a,u} + z/u\right).$$

Similarly, by choosing κ_1 sufficiently large, $\sup_{|s-t| \leq \lambda y^{-1/d}} |E(s) - E(t)| < \frac{a}{2u}$ for all $y > \kappa_1 u^{2d/\beta}$. In addition,

$$Var(\xi(s,t)) \leq E(\xi^2(s,t)) = d_g^2(s,t) \leq \kappa_g^2 |s - t|^\beta.$$

Then, by Lemma 2,

$$P\left(\sup_{|s - t| \leq \lambda y^{-1/d}} \xi(s,t) > a/u | f(0) = \gamma_{a,u} + z/u\right) \leq \exp\left(-\delta^* \frac{a^2}{u^2} y^{\beta/d}\right),$$

for δ^* sufficiently small. □

Using the core lemmas, we are ready to present the main theorem of this paper.

Theorem 4. *Let f be a Gaussian random field satisfying Conditions 1 and 2. Then, for any given $a > 0$, the estimator in (11) is asymptotically efficient in estimating $w(u)$.*

Proof of Theorem 4. Thanks to Lemma 3, for all $y > \kappa_1$,

$$P(m_{Leb}^{-1}(A_{\gamma_{a,u}}) u^{-2d/\beta} > y; \sup_T f(t) > u) \leq \exp\left(-\delta^* a^2 y^{\beta/d}\right).$$

By stochastic domination (Lemma 6.11 in Adler, Blanchet, and Liu (2010)), there exists $c > 0$ such that

$$E\left(\frac{1}{m^2_{Leb}(A_{\gamma_{a,u}})}; \sup_T f(t) > u \,\Big|\, \tau\right) \le cu^{4d/\beta}.$$

Then, according to (12), using $c' > 0$ to denote a generic positive constant, we obtain

$$
\begin{aligned}
E^Q\left(L^2_{w,\gamma_{a,u}}\right) &= \left(\int P(f(t) > \gamma_{a,u})dt\right)^2 \int_T E\left(\frac{1}{m^2_{Leb}(A_{\gamma_{a,u}})}; \sup_T f(t) > u \,\Big|\, \tau\right) p_{\gamma_{a,u}}(\tau)d\tau \\
&\le cm^2_{Leb}(T)\sup_T P^2(f(t) > \gamma_{a,u})u^{4d/\beta} \\
&\le c'm^2_{Leb}(T)u^{4d/\beta}\sup_T P^2(f(t) > u) \\
&\le c'm^2_{Leb}(T)u^{4d/\beta}P^2\left(\sup_T f(t) > u\right).
\end{aligned}
$$

The conclusion of weak efficiency is immediate by using B-TIS inequality to obtain that

$$\lim_{u\to\infty}\frac{1}{u^2}\log P^2\left(\sup_T f(t) > u\right) = -\frac{1}{2\sigma^2(T)}.$$

$\square$

4 NUMERICAL IMPLEMENTATION

In this section, we apply our algorithm to a few mean zero and constant variance fields.

To generate f from the Q, one needs the following steps:

1. Simulate τ. Since f has mean zero and constant variance, we simulate τ uniformly over the domain T with respect to Lebesgue measure.
2. Simulate $f(\tau)$. Under the original law, $f(\tau) \sim N(0, \sigma^2)$. We simulate $f(\tau)$ given that $f(\tau) > \gamma_{a,u} = u - a/u$.
3. Simulate $(Z_1, ..., Z_m)$ given that $\sum_{i=1}^m \sigma_i(\tau)Z_i = f(\tau)$, which is a multivariate Gaussian random vector.
4. Compute the estimator. We need $Em_{Leb}(A_{\gamma_{a,u}})$ and $m_{Leb}(A_{\gamma_{a,u}}(f))$. Then, output

$$\frac{Em_{Leb}(A_{\gamma_{a,u}})}{m_{Leb}(A_{\gamma_{a,u}}(f))}I(\sup_T f(t) > u).$$

For all the examples, we fix $a = 0.5$.

Example 5 (one dimensional case). *Let*

$$f(t) = Z_1\cos(t) + Z_2\sin(t)$$

and $T = [0, 3/4]$ where Z_1 and Z_2 are i.i.d. standard Gaussian r.v. The tail probability of $\sup_T f(t)$ is know in closed form (Adler and Taylor (2007))

$$P(\sup_{t\in T} f(t) > u) = 1 - \Phi(u) + \frac{3}{8\pi}\exp(-u^2/2)$$

Example 6 (two dimensional case). *Let*

$$f(t_1, t_2) = Z_1\cos(t_1) + Z_2\sin(t_1) + Z_3\cos(t_2) + Z_4\sin(t_2),$$

where $(t_1, t_2) \in T = [0, 2\pi]^2$. Z_i's are i.i.d. standard Gaussian random variables. We compute $P(\sup_{t\in T} f(t) > u)$. This quantity also has closed form

$$P(\sup_{t\in T} f(t) > u) = \frac{\pi}{2}ue^{-u^2/4}(2\Phi(\frac{u}{\sqrt{2}}) - 1) + \exp(-u^2/2).$$

Table 1: Simulation results for Example 5.

u	True value	Est.	Std. Er.	Std. Er./Est.
3	2.68E-03	2.72E-03	2.3E-05	0.0086
5	7.32E-07	7.33E-07	6.4E-09	0.0087
10	3.06E-23	3.08E-23	2.7E-25	0.0089
20	1.93E-88	1.95E-88	1.6E-90	0.0081

All result are based on 10^4 independent simulations.

Table 2: Simulation results for Example 6.

u	True value	Est.	Std. Er.	Std. Er./Est.
5	8.55E-03	8.64E-03	6.6E-05	0.0076
10	1.23E-10	1.23E-10	1.0E-12	0.0082
20	6.59E-43	6.44E-43	5.4E-45	0.0083

All result are based on 10^4 independent simulations.

Table 3: Simulation results for Example 7.

u	Est.	Std. Er.	Std. Er./Est.
7	5.88E-03	1.3E-04	0.023
10	2.59E-06	5.2E-08	0.023
15	5.06E-15	7.0E-17	0.014
20	1.54E-27	1.5E-29	0.010
30	1.32E-63	6.8E-66	0.005

All result are based on 10^4 independent simulations.

Example 7 (there dimensional case). *Let*

$$f(t_1, t_2, t_3) = Z_1 \cos(t_1) + Z_2 \sin(t_1) + Z_3 \cos(t_2) + Z_4 \sin(t_2) + Z_5 \cos(t_3) + Z_6 \sin(t_3),$$

and $(t_1, t_2, t_3) \in T = [0, 2\pi]^3$. Z_i*'s are i.i.d. standard Gaussian. Table 3 shows the results.*

The simulation results are given in Tables 1, 2, and 3.

REFERENCES

Adler, R., J. Blanchet, and J. Liu. 2008. Efficient simulation for tail probabilities of gaussian random fields. In *Proceeding of Winter Simulation Conference*, ed. S. Mason, L. M. R.R. Hill, O. Rose, T. Jefferson, and J. W. Fowler, 328–336: IEEE.

Adler, R., J. Blanchet, and J. Liu. 2010. Efficient monte carlo for large excursions of gaussian random fields. *Preprint, available at http://arxiv.org/abs/1005.0812*.

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

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

Asmussen, S., and L. Rojas-Nandayapa. 2008. Asymptotics of sums of lognormal random variables with gaussian copula. *Stat. Probab. Lett.* 78:2709–2714.

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.

Borell, C. 1975. The Brunn-Minkowski inequality in Gauss space. *Inventiones Mathematicae*.

Dudley, R. 1973. Sample functions of the gaussian process. *Annals of Probability* 1 (1): 66–103.

Glasserman, P., W. Kang, and P. Shahabuddin. 2007. Large deviations in multifactor portfolio credit risk. *Mathematical Finance* 17 (3): 345–379.

Glasserman, P., W. M. Kang, and P. Shahabuddin. 2008. Fast simulation of multifactor portfolio credit risk. *Operations Research* 56 (5): 1200–1217.

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

Riesz, F., and B. Sz-Nagy. 1955. *Functional analysis*. New York: Frederick Ungar Publ. Co.

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

Taylor, J., and R. Adler. 2003. Euler characteristics for gaussian fields on manifolds. *The Annals of Probability* 31 (2): 533–563.

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

JOSE H. BLANCHET is an Assistant Professor in the Department of Industrial Engineering and Operations Research at Columbia University. Jose holds a M.Sc. in Engineering- Economic Systems and Operations Research and a Ph.D. in Management Science and Engineering, both from Stanford University. He also holds two B.Sc. degrees: one in Actuarial Science and another one in Applied Mathematics from ITAM (Mexico). Jose worked for two years as an analyst in Protego Financial Advisors, a leading investment bank in Mexico. He has research interests in applied probability, computational finance, performance engineering, queueing theory, risk management, rare-event analysis, statistical inference, stochastic modelling, and simulation.

JINGCHEN LIU is an Assistant Professor in the Department of Statistics at Columbia University. Jingchen received his Ph.D. in statistics from Harvard University in June, 2008.

XUAN YANG is a Ph.D. student in the Department of Statistics at Columbia University.