SADDLEPOINT APPROXIMATIONS AND BOUNDARY CROSSING
PROBABILITIES FOR RANDOM FIELDS AND THEIR APPLICATIONS

by

Tze Leung Lai

Technical Report No. 2005-16
July 2005

Department of Statistics
STANFORD UNIVERSITY
Stanford, California 94305-4065
SADDLEPOINT APPROXIMATIONS AND BOUNDARY CROSSING PROBABILITIES FOR RANDOM FIELDS AND THEIR APPLICATIONS

by

Tze Leung Lai
Department of Statistics
Stanford University

Technical Report No. 2005-16
July 2005

This research was supported in part by the
National Science Foundation grant DMS 0305749

Department of Statistics
STANFORD UNIVERSITY
Stanford, California 94305-4065

http://www-stat.stanford.edu
Saddlepoint Approximations and Boundary Crossing Probabilities for Random Fields and Their Applications

Tze Leung Lai

Abstract

A number of classical results on boundary crossing probabilities of Brownian motion and random walks are extended to certain classes of random fields, which include sums of independent random variables with multidimensional indices, multivariate empirical processes, and scan statistics in change-point detection as special cases. Key ingredients in these extensions are saddlepoint approximations for tail probabilities and geometric integration over tubular neighborhoods of extremal manifolds related to the large or moderate deviations theory of these fields.
1. Introduction

The term "saddlepoint approximation" for the tail probability, or for the density function at a quantile $x$ in the tail of the distribution of a sample mean or the ratio of two sample means was introduced by Daniels [11], who applied the method of steepest descent to the moment generating function $e^{\psi}$, leading to the saddlepoint of $\psi(z) - zx$ that appears in the approximation. Closely related results were derived earlier by Esscher [12] and Cramér [10] via an alternative approach based on change of measures and Edgeworth expansions. Extensions of these results to multidimensional sample mean vectors and to conditional distributions were developed by Borovkov and Rogozin [6], Barndorff-Nielsen and Cox [4], Skovgaard [30] and others; see the monograph by Jensen [19] that also considers extensions to Markov chains and to likelihood ratio statistics in exponential families.

Instead of the tail probability of a sample mean $\bar{X}_n = \frac{S_n}{n}$, consider the probability that $\max_{1 \leq k \leq n} S_k$ exceeds some threshold $c_n$. More generally, for random walks taking values in $\mathbb{R}^d$, consider the probability that $\max_{1 \leq k \leq n} kg(S_k/k)$ exceeds $c_n$, where $g : \mathbb{R}^d \to \mathbb{R}$. Section 2 reviews some recent developments in which saddlepoint approximations for the tail distribution of $S_k/k$ are used in conjunction with geometric integration theory to derive asymptotic formulas for the boundary crossing probability $P\{kg(S_k/k) > c_n \text{ for some } 1 \leq k \leq n\}$, with the threshold $c_n$ chosen in the "large deviation" regime under which the probability is of the exponential order $O(e^{-\rho n})$ for some $\rho > 0$. Applications of these results to sequential analysis and change-point detection in stochastic dynamical systems are also given in this connection. Section 3 considers extensions of the approach to (i) the case of $c_n$ in the "moderate deviation" regime for which the probability approaches 0 at an algebraic (instead of exponential rate), (ii) continuous-time processes (instead of discrete-time $S_n$), and (iii) multi-dimensional time for random fields (instead of random processes). The basic idea is like that in Section 2 which first develops sharp local approximations for the probabilities at different sites of the random field and then sums up these "generalized saddlepoint approximations" via geometric integration theory. Some concluding remarks are given in Section 4.

2. Asymptotic formulas for boundary crossing probabilities via geometric integration of saddlepoint approximations and related applications

Let $X_1, X_2, \ldots$ be i.i.d. $d$-dimensional nonlattice random vectors whose common moment generating function is finite in some neighborhood of the origin. Let $S_n = X_1 + \ldots + X_n$, $\mu_0 = EX_1$ and $\Theta = \{\theta : Ee^{\theta'X} < \infty\}$, where $'$ denotes transpose. Assume that Cov($X_1$)
is positive definite. Let \( \psi(\theta) = \log(Ee^{\theta'X}) \) denote the cumulant generating function of \( X_1 \). Let \( \Lambda \) be the closure of \( \nabla \psi(\Theta) \) and let \( \Lambda^0 \) be its interior. Denote the boundary of \( \Lambda \) by \( \partial\Lambda = \Lambda - \Lambda^0 \). Then \( \nabla \psi \) is a diffeomorphism from \( \Theta^0 \) onto \( \Lambda^0 \). Let \( \theta_\mu = (\nabla \psi)^{-1}(\mu) \). For \( \mu \in \Lambda^0 \), define
\[
\phi(\mu) = \sup_{\theta \in \Theta} \{ \theta'\mu - \psi(\theta) \} = \theta'_\mu \mu - \psi(\theta_\mu).
\] (2.1)

The function \( \phi \) is the convex dual of \( \psi \) and is also known as the *rate function* in large deviations theory. Let \( g : \Lambda \to \mathbb{R} \) and define the stopping time
\[
T_c = \inf\{k \geq n_0 : kg(S_k/k) > c\}.
\] (2.2)

To analyze \( P\{T_c \leq \alpha c\} \), first consider the case where the distribution function \( F \) of \( X_1 \) has a bounded continuous derivative so that \( S_n/n \) has the saddlepoint approximation
\[
P\{S_n/n \in d\mu\} = (1 + o(1))(n/2\pi)^{d/2}|\Sigma(\mu)|^{-1/2}e^{-n\phi(\mu)} \ d\mu,
\] (2.3)

where \( \Sigma(\mu) = \nabla^2 \psi(\theta)|_{\theta = \theta_\mu} \) and the \( o(1) \) term is uniform over compact subsets of \( \Lambda^0 \) (cf. [4], [6]). Let
\[
f(\mu) d\mu = P\{T_c \leq \alpha c, S_{T_c}/T_c \in d\mu\}
\]
\[
= \sum_{\delta c \leq n \leq \alpha c} P\{S_n/n \in d\mu\} I_{(ng(\mu)) > c}
\]
\[
\times P\{kg(S_k/k) < c \text{ for all } \delta c \leq k < n | S_n/n \in d\mu\}.
\] (2.4)

Note that \( P\{T_c \leq \alpha c\} = \int_{\mathbb{R}^d} f(\mu) d\mu \). Chan and Lai [7] obtained an asymptotic formula for the large deviation probability \( P\{T_c \leq \alpha c\} \), with \( n_0 \sim \delta c \) and \( 0 < \delta < \alpha \), by using Laplace's method to evaluate the integral \( \int_{\mathbb{R}^d} f(\mu) d\mu \) under the following regularity conditions on \( g \):

(A1) \( g \) is continuous on \( \Lambda^0 \) and there exists \( \epsilon_0 > 0 \) such that
\[
\sup_{\alpha^{-1} < g(\mu) < \delta^{-1} + \epsilon_0} g(\mu)/\phi(\mu) = r < \infty.
\]

(A2) \( M_\epsilon := \{ \mu : \alpha^{-1} < g(\mu) < \delta^{-1} + \epsilon \} \) and \( g(\mu)/\phi(\mu) = r \) is a \( q \)-dimensional oriented manifold for all \( 0 \leq \epsilon \leq \epsilon_0 \), where \( q \leq d \).

(A3) \( \liminf_{\mu \to \partial\Lambda} \phi(\mu) > (\delta r)^{-1} \), and there exists \( \epsilon_1 > 0 \) such that \( \phi(\mu) > (\delta r)^{-1} + \epsilon_1 \) if \( g(\mu) > \delta^{-1} + \epsilon_0 \).

(A4) \( g \) is twice continuously differentiable in some neighborhood of \( M_\epsilon \) and \( \sigma(\{\mu : g(\mu) = \delta^{-1} \text{ and } g(\mu)/\phi(\mu) = r\}) = 0 \), where \( \sigma \) is the volume element measure of \( M_\epsilon \).
Assumptions (A1)-(A3) imply that \( \sup_{\phi(\mu) > \alpha^{-1} \min(\delta^{-1}, g(\mu))/\phi(\mu)} \) can be attained on the \( q \)-dimensional manifold \( M_0 \). The first part of (A3) implies that there exists \( \epsilon^* > 0 \) such that

\[
M^* := \{ \mu : \alpha^{-1} \leq g(\mu) \leq \delta^{-1} + \epsilon^*, g(\mu)/\phi(\mu) = r \}
\]  

(2.5)

is a compact subset of \( \Lambda \); it clearly holds if \( \phi(\mu) \to \infty \) as \( \mu \to \partial \Lambda \), which is usually the case. For \( \mu \in M_0 \), let \( T M_0(\mu) \) denote the tangent space of \( M_0 \) at \( \mu \) and let \( T M_0^\perp(\mu) \) denote its orthogonal complement (i.e., \( T M_0^\perp(\mu) \) is the normal space of \( M_0 \) at \( \mu \)). Let \( \rho(\mu) = \phi(\mu) - g(\mu)/r \). By (A1) and (A3), \( \rho \) attains on \( M_0 \) its minimum value 0 over \( \{ \mu : \alpha^{-1} < g(\mu) < \delta^{-1} + \epsilon_0 \} \), and therefore

\[
\nabla \rho(\mu) = 0 \quad \text{and} \quad \nabla^2 \rho(\mu) \quad \text{is nonnegative definite for} \quad \mu \in M_0.
\]  

(2.6)

Let \( \Pi_\mu^\perp \) denote the \( d \times (d - q) \) matrix whose column vectors form an orthonormal basis of \( T M_0^\perp(\mu) \). Then the matrix \( \nabla^2_\mu \rho(\mu) := (\Pi_\mu^\perp)^T \nabla^2 \rho(\mu) \Pi_\mu^\perp \) is nonnegative definite for \( \mu \in M_0 \). Let \( | \cdot | \) denote the determinant of a nonnegative definite matrix, and assume also that

(A5) \( \inf_{\mu \in M_0} | \nabla^2_\mu \rho(\mu) | > 0 \), \quad \text{with} \quad \rho = \phi - g/r,

setting \( | \nabla^2_\mu \rho(\mu) | = 1 \) in the case \( d - q = 0 \). Under these assumptions and making use of (2.3) and (2.4), Chan and Lai [7] showed that

\[
P\{ T_c \leq \alpha c \} = \int_{\mathbb{R}^d} f(\mu) \, d\mu \sim \int_{U_{c^{-1/2} \log c}} f(\mu) \, d\mu,
\]  

(2.7)

where \( U_\eta \) is a tubular neighborhood of \( M_0 \) with radius \( \eta \), i.e., \( U_\eta = \{ y + z : y \in M_0, z \in T M_{0}^\perp(y) \ \text{and} \ ||z|| \leq \eta \} \).

The volumes of tubes around smooth closed curves and more general manifolds, first derived by Hotelling [16] and Weyl [31], have played a prominent role in the statistical theory of nonlinear regression models, cf. [20], [24], [28]. In these papers, the tube is a union of disks of radius \( \eta \) around all points of the manifold, and Weyl's tube formula is used to express its volume as a polynomial of degree \([q/2]\) in \( \eta \). The integral in (2.7) uses a slightly different formulation of the tubular neighborhood and certain differential geometric results on the so-called "infinitesimal change of volume function" (cf. [14]), from which it follows that as \( \eta := c^{-1/2} \log c \to 0 \),

\[
\int_{U_\eta} f(\mu) \, d\mu \sim \int_{M_0} \left\{ \int_{z \in T M_{0}^\perp(y), ||z|| \leq \eta} f(y + z) \, dz \right\} \, d\sigma(y).
\]  

(2.8)
The inner integral in (2.8) can be evaluated asymptotically by making use of (2.3) and (2.4), leading to the following asymptotic formula when \( n_0 \sim \delta c, \alpha > \delta \) and \( g(\mu_0) < \alpha^{-1} \):

\[
P \{ T_c \leq \alpha c \} \sim (c/2\pi r)^{q/2} e^{-c/r} \int_{M_0} \gamma(\mu) (\phi(\mu))^{-q/(2+1)} |\Sigma(\mu)|^{-1/2} |\nabla^2 \rho(\mu)|^{-1/2} d\sigma(\mu),
\]

(2.9)

where \( \nabla^2 \rho \) is given in (A5) and \( \gamma(\mu) = \int_0^\infty e^{-y} P \{ \min_{n \geq 1} S_n^{(\mu)} > y \} dy \); the integrand in \( \gamma(\mu) \) can be derived as the limit of

\[
e^{-y\psi(\mu)/g(\mu)} P\{(n-j)g(S_{n-j}/(n-j)) < c \text{ for all } 1 \leq j < c^{1/4}|S_n/n = \mu\},
\]

for which a standard time reversal argument yields i.i.d. \( X_i^{(\mu)} \) with \( P\{ X_i^{(\mu)} \in dx \} = e^{\theta \mu_x - \psi(\theta)} dF(x) \), leading to the \( S_i^{(\mu)} = \sum_{i=1}^n \{ \theta \mu_i X_i^{(\mu)} - \psi(\theta) \} \) in \( \gamma(\mu) \).

The preceding analysis has assumed that \( X_1 \) has a bounded continuous density function. Chan and Lai [7] also showed that the asymptotic formula (2.9) still holds under the much weaker assumption that \( X_1 \) be nonlattice. By partitioning \( \Lambda \) into suitably small cubes, [7] uses exponential tilting and a local limit theorem to modify the preceding analysis, replacing \( \in d\mu \) above by \( \in I_\mu \), where \( I_\mu \) denotes a cube centered at \( \mu \). Applications of (2.9) and its variants to sequential tests of composite hypotheses in multiparameter exponential families are also given in [7].

Moving averages play a fundamental role in the problem of detection, and diagnosis of abrupt changes in a stochastic system on the basis of sequential observations has many applications, including fault detection and diagnosis in complex dynamical systems and industrial processes, integrity monitoring of navigation systems, and radar and sonar signal processing (cf. [5], [13], [22], [24]). Lai [22], [23] developed a class of efficient sequential change-point detection procedures which are based on certain nonlinear functions of moving averages with different window sizes and which have stopping rules of the form

\[
T_c = \inf \left\{ n : \max_{k \in J(c)} k g(S_n(k)/k) > c \right\}.
\]

(2.10)

Here \( S_n \) is a \( d \)-dimensional random walk, \( S_n(k) = S_n - S_{n-k} \) so that \( S_n(k)/k \) is a moving average, \( g : \mathbb{R}^d \to \mathbb{R} \) and \( J_n, J(c) \) are subsets of \( \{1, 2, \ldots\} \). Closely related to \( T_c \) in (2.10) used for sequential detection is the scan statistic

\[
M_n = \max_{k \in J_n, k \leq j \leq n} k g(S_j(k)/k)
\]

(2.11)

used in fixed-sample change-point problems arising in computational biology (cf. [3], [21]). Chan and Lai [8] have shown under conditions (A1)-(A5) on \( g \) that there exist \( q \in \{0, \ldots, d\} \)
and \( r > 0 \) (depending on \( g \)) such that
\[
e^{-c/r} (c/r)^{q/2} T_c \text{ has a limiting exponential distribution as } c \to \infty, \tag{2.12}
\]
\( M_n - r\{\log n + (q/2) \log \log n\} \text{ has a limiting Gumbel-type distribution as } n \to \infty. \tag{2.13}\]

In the case \( d = 1 \) and \( g(\mu) = \mu, \) (2.13) with \( q = 0 \) has been established by Iglehart [17] in the context of longest waiting times in a GI/G/1 queue and by Karlin et al. [21] in the context of high-scoring segments in a DNA or amino acid sequence. In the case \( d = 1, \) \( g(\mu) = \mu^2/2 \)
and standard normal \( X_i, \) (2.12) has been established by Siegmund and Venkatraman [27] in the context of generalized likelihood ratio control charts. The asymptotic theory in Chan and Lai [8] concerning (2.10) and (2.11) unifies these previous results and also leads to definitive solutions of a variety of change-point detection problems, giving in particular (i)

the extension of i.i.d. to Markov-dependent \( \xi_i \) (so that more general stochastic systems can be treated), and (ii) suitable choice of \( g \) and \( J_n \) or \( J(c) \) in (2.10) or (2.11) to achieve both statistical and computational efficiency.

A unified approach in Chan and Lai [8] to derive (2.12) and (2.13) is based on integrating saddlepoint approximations for Markov random walks with respect to certain measures over some tubular neighborhood of an extremal \( q \)-dimensional manifold in \( \mathbb{R}^d, \) which incorporates both the critical temporal and spatial components of the problem, yielding an asymptotic formula for the following boundary crossing probability:
\[
P\{\max_{\delta c \leq k \leq \alpha c} kg(S_n(k)/k) > c \text{ for some } n \leq \beta c\} \sim (c/r)^{q/2} ce^{-c/r}(\beta \zeta^{(1)}_{\alpha,\delta} - \zeta^{(2)}_{\alpha,\delta}), \tag{2.14}\]

where \( 0 < \delta < \alpha < \beta \) and \( \zeta^{(i)}_{\alpha,\delta} (i = 1, 2) \) are constants that can be expressed as integrals over the extremal manifold. To see how (2.12) can be derived from (2.14) in the case of i.i.d. \( X_i \) and \( J(c) = \{k : \delta c \leq k \leq \alpha c\}, \) let \( x > 0, \) \( m = (\zeta^{(1)}_{\alpha,\delta})^{-1}(c/r)^{-q/2} e^{c/r} x, \)
and partition the interval \([0, m]\) into \( K \sim m/(\beta c) \) disjoint intervals \( I_1, \ldots, I_K \) with equal length \( \beta c(1 + o(1)). \)

Then the events
\[
A_j = \left\{ \max_{\delta c \leq k \leq \alpha c} kg(S_n(k)/k) > c \text{ for some } n \in I_j \text{ and } n - k \in I_j \right\}
\]
are independent and have the same probability \( p_c. \) Letting \( \bar{A}_j \) denote the complement of \( A_j, \) it follows from (2.14) that
\[
P(\cap_{j=1}^K \bar{A}_j) = (1 - p_c)^K \sim e^{-K p_c} \to \exp\{-x[1 - (\beta \zeta^{(1)}_{\alpha,\delta})^{-1}\zeta^{(2)}_{\alpha,\delta}]/r\}
\]
as \( c \to \infty. \) Hence \( P(\cup_{j=1}^K A_j) \to 1 - \exp\{-x[1 - (\beta \zeta^{(1)}_{\alpha,\delta})^{-1}\zeta^{(2)}_{\alpha,\delta}]/r\}. \)

Let
\[
B_j = \left\{ \max_{\delta c \leq k \leq \alpha c} kg(S_n(k)/k)) > c \text{ for some } n \in I_{j+1} \text{ and } n - k \in I_j \right\}.
\]
Then by a similar argument involving a straightforward modification of (2.14), it can be shown that \( P(\bigcup_{j=1}^{K} B_j) \to 1 - \exp[-(\beta c(1)_{\alpha,\beta}^{(2)} x/\alpha)_{\alpha,\beta}^{(2)}] \) as \( c \to \infty \) for sufficiently large \( \beta \) (with \( \beta > \alpha \)). Taking \( \beta \) arbitrarily large, since
\[
P(\bigcup_{j=1}^{K} A_j) \leq P\{T_c \leq m\} \leq P(\bigcup_{j=1}^{K} A_j) + P(\bigcup_{j=1}^{K} B_j),
\]
it then follows that
\[
P\{T_c \leq m\} \to 1 - e^{-x} \quad \text{as} \quad c \to \infty.
\] (2.15)

Since \( P\{T_c \leq n\} = P\{M_n > c\} \), (2.13) follows from (2.12) by using (2.15) with \( c = r\{\log n + (q/2)\log\log n\} + t \).

Note that as we let both the time \( n \) and window size \( k \) vary, the moving sums \( S_n(k) \), \( n \geq k \), have a two-dimensional time index \( (k,n) \). Whereas the boundary crossing probability in (2.14) is in the large deviation regime and involves the moment generating function of \( X_1 \) via the rate function (2.1), the next section considers the case where the boundary crossing probability is in the moderate deviation regime, for which the asymptotic normality of the random field \( \{S_n(k), n \geq k\} \) plays a crucial role so that only the first two moments of \( X_1 \) are involved in the asymptotic formula.

3. Boundary crossing probabilities for asymptotically Gaussian random fields and some applications

To begin with, let \( \{W(t) : t \geq 0\} \) be Brownian motion and let \( T_c = \inf\{t \geq 0 : W(t) \geq b_c(t)\} \) be the first time when Brownian motion crosses a positive continuously differentiable boundary \( b_c \). Strassen [30], Jennen and Lerche [18] and others have shown that \( T_c \) has a density function \( p_c \) and that under certain additional conditions on \( b_c \), \( p_c \) has the “tangent approximation”
\[
p_c(t) = t^{-3/2}a_c(t)\varphi(b_c(t)/\sqrt{t}), \tag{3.1}
\]
where \( \varphi(x) = (2\pi)^{-1/2}e^{-x^2/2} \) is the standard normal density function and \( a_c(t) = b_c(t) - t b'_c(t) \).

Note that in the case of a linear boundary \( b_c(t) = \alpha + \beta t \) (with \( \alpha > 0 \) and \( \beta > 0 \)), the well known Bachelier–Lévy formula yields \( p_c(t) = t^{-3/2}a_c(t)\varphi(b_c(t)/\sqrt{t}) \), so (3.1) simply replaces \( \alpha \) by the intercept \( a_c(t) \) of the tangent line passing through \( (t,b_c(t)) \), and is therefore called a “tangent approximation”. For concave boundaries \( b_c(t) = b(t) \) that becomes infinite as \( t \to \infty \), one typically has \( b'(t) = o(b(t)/t) \), so one can replace \( a_c(t) \) in (3.1) by \( b(t) \). There is a close connection between this approximation to \( p_c(t) \) and the Kolmogorov–Erdős–Feller test, which yields for nondecreasing \( b(t)/\sqrt{t} \) the 0-1 dichotomy
\[
P\{W(t) < b(t) \text{ for all large } t\} = 1 \quad \text{(or 0)} \text{ if } I(b) < \infty \quad \text{(or } = \infty\text{)}, \tag{3.2}
\]
where \( I(b) = \int_1^\infty t^{-3/2} b(t) \varphi(b(t)/\sqrt{t}) \, dt < \infty \).

To extend (3.1) and (3.2) to much more general processes, Chan and Lai [9] use (i) moderate deviation approximations to marginal tail probabilities and (ii) weak convergence (to a limiting Gaussian process) of a certain conditional process given that the process attains a high level near the boundary at time \( t \). A key idea of their extension is to relax the requirement that the left hand side of (3.1) be a first exit density, regarding it instead as a “local” exit density at time \( t \) so that the probability that the process ever crosses the boundary within time interval \( D \) is asymptotically equal to the integral of the right hand side of (3.1) over \( D \). Not only does this avoid the technical assumptions that need to be imposed to ensure that the first exit time \( T_c \) indeed has a density function with respect to Lebesgue measure, but it also dispenses with the notion of having a totally ordered set \( D \) so that the “first” time of exit can be defined. This approach therefore can be applied to random fields (with multidimensional time that is not totally ordered).

Let \( \psi(c) = \varphi(c)/c = \left(2\pi c^2\right)^{-1/2} \exp(-c^2/2) \). For vectors \( t, u \in \mathbb{R}^d \), the relation \( t \leq u \) means \( t_i \leq u_i \) for all \( i \), and \( t < u \) means \( t_i < u_i \) for all \( i \). Let \( S^{d-1} \) denote the \((d-1)\) dimensional unit sphere, and \( \mathbb{Z}_+ \) (\( \mathbb{R}_+ \)) denote the set of positive integers (real numbers). For \( \zeta \geq 0 \), let \( I_{t,\zeta} = \prod_{i=1}^d [t_i, t_i + \zeta] \). For \( D \subset \mathbb{R}^d \) and \( \delta > 0 \), define \( [D]_{\delta} = \{t+u : t \in D, \|u\| < \delta\} \). Let \( 0 < \alpha \leq 2 \) and \( \{W_t(u) : u \in [0, \infty)^d\} \) be a continuous Gaussian random field such that

\[
W_t(0) = 0, \quad E[W_t(u)] = -\|u\|^{\alpha} r_t(u/\|u\|)/2, \quad \text{Cov}(W_t(u), W_t(v))
= \left(\|u\|^{\alpha} r_t(u/\|u\|) + \|v\|^{\alpha} r_t(v/\|v\|) - \|u - v\|^{\alpha} r_t((u - v)/\|u - v\|)\right)/2,
\]

where \( r_t : \mathbb{S}^{d-1} \to \mathbb{R}_+ \) is a continuous function satisfying

\[
\sup_{v \in \mathbb{S}^{d-1}} |r_t(v) - r_u(v)| \to 0 \quad \text{as} \quad u \to t.
\]

It can be shown that the following limit exists and is uniformly continuous:

\[
H(t) = \lim_{K \to \infty} K^{-d} H_K(t),
\]

where \( H_K(t) = \int_0^\infty e^y P \{\sup_{0 \leq u \leq K} W_t(u) > y\} \, dy \).

For \( c > 0 \), let \( X_c \) be a random field on some domain in \( \mathbb{R}^d \) such that \( E X_c(t) = 0 \) and \( E X_c^2(t) = 1 \) for all \( t \). Let \( D \) be such that \([D]_{\delta}\) is a subset of the domain of \( X_c \) for some \( \delta > 0 \) and all \( c \) large enough. Let \( \rho_c(t, u) = E[X_c(t)X_c(u)] \). Suppose there exist \( 0 < \alpha \leq 2 \) and a slowly varying function \( L \) such that as \( u \to 0 \),

\[
(C) \quad \rho_c(t, t + u) = 1 - (1 + o(1)) \|u\|^{\alpha} L(\|u\|) r_t(u/\|u\|)
\]
uniformly over \( t \in [D]_\delta \). Let \( \Delta_c = \min\{x > 0 : x^\alpha L(x) = (2c^2)^{-1}\} \). For example, if \( L(x) \equiv 1 \), then \( \Delta_c = (2c^2)^{-1/\alpha} \). Moreover, assume that the following conditions also hold uniformly over \( t \in [D]_\delta \), as \( c \to \infty \):

\[
\text{(B1)} \quad P\{X_c(t) > c - y/c\} \sim \psi(c - y/c)
\]

uniformly over positive, bounded values of \( y \), and for any \( a > 0 \) and positive integers \( m \),

\[
\text{(B2)} \quad \{c[X_c(t + ak\Delta_c) - X_c(t)] : 0 \leq k_i < m\}|X_c(t) = c - y/c \Rightarrow \{W_t(ak) : 0 \leq k_i < m\}
\]

as \( c \to \infty \), uniformly over positive, bounded values of \( y \), where we use \( \Rightarrow \) to denote weak convergence and "\( X_c(t) = c - y/c\)" to denote that the distribution is conditional on \( X_c(t) = c - y/c \). Whereas (B1) refers to the marginal distribution of \( X_c(t) \), saying that \( \{X_c(t) > c - y/c\} \) has probability like that of a standard normal, the joint distribution of \( X_c(\cdot) \) is assumed in (B2) to be asymptotically normal in the sense of weak convergence for local increments conditioned on \( X_c(t) = c - y/c \). Note that the same \( \alpha, L(\cdot) \) and \( r_t(\cdot) \) appear in (C) and in the mean and covariance functions (3.3) of the Gaussian field \( W_t(\cdot) \) in (B2). Under (C), (B1), (B2) and some additional regularity conditions, Chan and Lai [9] derived the following analog of (3.1) for asymptotically Gaussian random fields: As \( c \to \infty \) and \( \ell_c \to \infty \) such that \( \ell_c = o(\Delta_c^{-1}) \),

\[
P\{ \sup_{u \in I_t, \ell_c, \Delta_c} X_c(u) > c \} \sim \ell_c^d \psi(c) H(t), \quad (3.5)
\]

\[
P\{ \sup_{u \in I_t, \ell_c, \Delta_c} X_c(u) > c, \sup_{v \in B[I_t, \ell_c, \Delta_c]} X_c(v) > c \} = o(\ell_c^d \psi(c)), \quad (3.6)
\]

uniformly over \( t \in D \) and over subsets \( B \) of \([D]_\delta \) with bounded volume, where \( H(t) \) is defined in (3.4). Dividing (2.11) by \( (\ell_c \Delta_c)^d \), which is the volume of the cube \( I_t, \ell_c, \Delta_c \), yields an asymptotic boundary crossing "density" \( \Delta_c^{-d} \psi(c) H(t) \) of \( X_c \) at \( t \). By integrating this "density" over a bounded and Jordan measurable set \( D \), or more precisely, by summing (3.5) over the "tiles" \( I_t, \ell_c, \Delta_c \) of \( D \) and applying (3.6), we obtain

\[
P\{\sup_{t \in D} X_c(t) > c\} \sim \psi(c) \Delta_c^{-d} \int_D H(t) \, dt \quad \text{as} \quad c \to \infty. \quad (3.7)
\]

When \( X_c \) is a stationary isotropic Gaussian random field, the asymptotic formula (3.7) is due to Qualls and Watanabe [26]. Chan and Lai [9] also applied (3.5) and (3.6) to derive an asymptotic approximation to the boundary crossing probability \( P\{X_c(t) > b_c(t) \text{ for some } t \in D_c\} \) by using the probabilities \( p_c(t) = P\{X_c(t) > b_c(t) \text{ for some } t \in I_{t, c}\} \) as building
blocks, where $\zeta_c \to 0$ is so chosen that $\sup_{t \in [D_c], b_c} \Delta b_c(t) = o(\zeta_c)$ as $c \to \infty$. Under certain regularity conditions on $D_c$ and $b_c$, they showed that as $c \to \infty$,

$$P\{X_c(t) > b_c(t) \text{ for some } t \in D_c\} \sim \int_{D_c} \psi(b_c(t))\Delta_{b_c(t)}^{-d} H(t) \, dt. \quad (3.8)$$

Consider the special case in which $b_c(t) = cb(t)$ for some positive function $b$ possessing continuous second derivatives on $[D]$, where $D$ is a compact Jordan measurable set. Let $b_D = \inf_{t \in D} b(t)$ and assume that $\mathcal{M} = \{t \in D : b(t) = b_D\}$ is a $q$-dimensional manifold (with boundary) such that $\sigma(\mathcal{M} \cap \partial D) = 0$. Letting $\{e_1(t), \ldots, e_{d-q}(t)\}$ be an orthonormal basis of $T\mathcal{M}(t)$, define the $d \times (d-q)$ matrix $A(t) = (e_i(t) \cdot e_{d-q}(t))$ and assume that $\nabla^2 b(t) := A'(t)\nabla^2 b(t) A(t)$ is a positive definite $q \times q$ matrix for all $t \in \mathcal{M}$. Then for $\alpha < 2$, applying Laplace's method to the integral in (3.8) and using geometric integration over a tubular neighborhood of $\mathcal{M}$ as in (2.8) and (2.9) yields

$$P\{X_c(t) > cb(t) \text{ for some } t \in D\} \sim \psi(cb_D)b_D^{2d/\alpha} \Delta_{d}^{-d} (2\pi/c^2b_D)^{(d-q)/2} \int_{\mathcal{M}} |\nabla^2 b(t)|^{-1/2} H(t) d\sigma(t) \text{ as } c \to \infty. \quad (3.9)$$

### 3.1. Applications to multidimensional empirical processes

Let $Y_1, Y_2, \ldots$ be i.i.d. $d$-dimensional random vectors with common distribution function $F$, and let $F_n(t) = n^{-1} \sum_{i=1}^{n} 1(Y_i \leq t)$, $t \in \mathbb{R}^d$, be the empirical distribution function of $Y_1, \ldots, Y_n$. Let $Z_n(t) = \sqrt{n}(F_n(t) - F(t))$ be the multivariate empirical process. The limiting distribution of $Z_n$ is that of a Gaussian sheet $Z^0$, for which Adler and Brown [1] proved that

$$K_d F \epsilon^2 \lesssim \epsilon \sup_t \{Z^0(t) > c\} \lesssim K_d \epsilon^2 \epsilon^{-2\epsilon^2}, \quad (3.10)$$

where $K_d$ is a constant depending only on $d$ and $K_d F$ is a constant depending on both $d$ and the distribution $F$. For the case $d = 2$ with independent components $Y_{1,1}$ and $Y_{1,2}$ of $Y_1$, $Z^0$ is a pinned Brownian sheet, for which Hogan and Siegmund [15] sharpened (3.10) into

$$P\{\sup_t \{Z^0(t) > c\} \sim (4 \log 2) \epsilon^2 \epsilon^{-2\epsilon^2} \text{ as } \epsilon \to \infty. \quad (3.11)$$

As a corollary of (3.9), Chan and Lai [9] showed that if the sample size $n_c$ increases to $\infty$ with $c$ such that $c = o(n_c^{1/\theta})$, then we can replace $Z^0(t)$ in (3.11) by $Z_{n_c}(t)$ and also extend the result to general $d$ and general distribution $F$ such that $F$ is continuously differentiable and $\partial F/\partial t_i > 0$ for $1 \leq i \leq d$. Specifically, letting $\mathcal{M} = \{t : F(t) = \frac{1}{2}\}$, they showed that as
\(c \to \infty,\)
\[
P\{\sup_t Z_{n,c}(t) > c\} \sim (8c^2)^{d-1}e^{-2c^2} \int_{\mathcal{M}} \|\nabla F(t)\|^{-1} \prod_{i=1}^{d} \frac{\partial F(t)}{\partial t_i} d\sigma(t),
\]
(3.12)
\[
P\{\sup_t |Z_{n,c}(t)| > c\} \sim 2(8c^2)^{d-1}e^{-2c^2} \int_{\mathcal{M}} \|\nabla F(t)\|^{-1} \prod_{i=1}^{d} \frac{\partial F(t)}{\partial t_i} d\sigma(t).
\]
(3.13)

3.2. Applications to sums of random variables with multidimensional indices

Let \(Y_k, k \in \mathbb{Z}_+^d\), be i.i.d. random variables such that
\[
EY_k = 0, \quad EY_k^2 = 1 \quad \text{and} \quad E|Y_k|^3 < \infty.
\]
(3.14)

Let \(S_n = \sum_{k \leq n} Y_k\), where \(k \leq n\) denotes that \(k_i \leq n_i\) for \(1 \leq i \leq d\). Let \(|n| = \prod_{i=1}^{n} n_i\), \(\log n = (\log n_1, \ldots, \log n_d)\) and \(\exp(t) = (\exp(t_1), \ldots, \exp(t_d))\). Define \(X(\log n) = |n|^{-1/2} S_n\) and extend the domain of \(X\) to \([0, \infty)^d\) by defining \(X(t) = X(\log n)\) when \(\log n_i \leq t_i < \log(n_i + 1)\) for all \(i\). If \(t = \log n\) and \(t + u = \log m\) for some \(m, n \in \mathbb{Z}_+^d\), then
\[
1 - \rho(t, t + u) = 1 - \text{Cov}(|n|^{-1/2} S_n, |m|^{-1/2} S_m) = 1 - \exp\left(-\sum_i |u_i|/2\right) \sim \sum_i |u_i|/2
\]
as \(u \to 0\). Hence (C) holds with \(\alpha = 1\), \(L(x) \equiv 1\), \(r_t(u) = \sum_i |u_i|/2\), and therefore \(\Delta_c = (2c^2)^{-1}\). Let \(X_c = X, \rho_c = \rho\) and \(D_c\) be a Jordan measurable subset of \(\{t : \Sigma t_i \geq c^3\}\) such that \([D_c]_\beta \subset G \equiv \{t : t_i/\Sigma t_j \geq \epsilon\} \text{ for all } i\}.\) Chan and Lai [9] made use of (3.5) and (3.6) to show that as \(c \to \infty\)
\[
P\{\ \sup_{t \in D_c} X(t) > c\} \sim 2^{-d} c^{2d} \Psi(c) \text{ vol}(D_c).
\]
(3.15)

This result leads to an extension of the Kolmogorov–Erdős–Feller test (3.2) to the case of multidimensional time. Let \(\beta : \mathbb{Z}_+^d \to (0, \infty)\) be nondecreasing in the sense that \(\beta(m) \leq \beta(n)\) for all \(m \leq n\). We say that \(\beta\) is an upper (lower) class function if \(\sup\{|n| : |n|^{-1/2} S_n > \beta(n)\} < (\geq) \infty\) a.s. For \(\epsilon \geq 0\), let \(F_\epsilon = \{n \in \mathbb{Z}_+^d : \log n_i/\log |n| \geq \epsilon\} \text{ for all } i\};\) in particular \(F_0 = \mathbb{Z}_+^d\). Letting
\[
J_\epsilon = \sum_{n \in F_\epsilon} |n|^{-1} \beta^{2d-1}(n) e^{-\beta^2(n)/2},
\]
(3.16)
Chan and Lai [9] showed that (i) \(\beta\) is an upper class function if \(J_0 < \infty\), and (ii) \(\beta\) is a lower class function if \(J_\epsilon = \infty\) for some \(\epsilon > 0\).

4. Concluding remarks
We have shown how a number of classical results on boundary crossing probabilities for Brownian motion and random walks can be extended to certain classes of random fields. Key ingredients in these extensions are (a) saddlepoint approximations for tail probabilities or asymptotic “local densities”, and (b) geometric integration over tubular neighborhoods of extremal manifolds in applying Laplace’s method to integrate (or add) the local densities or saddlepoint approximations.

Our formulation of “asymptotically Gaussian” random fields bears some resemblance to Aldous’ [2] Poisson clumping heuristic, which involves i.i.d. clumps of high-level excursions of a stochastic process $X(t)$, with the stochastic structure of the clump determined by the conditional limiting process (like that in (B2)) of normalized local increments. Whereas the Poisson clumping heuristic only suggests an asymptotic approximation $P\{\sup_{t \in D} X(t) \leq c\}$ of the form $e^{-p_c}$ with $p_c \to 0$, our approach actually gives a rigorous derivation of an asymptotic formula for $p_c$. Instead of a single stochastic process $X(t)$, our formulation involves a family of random fields $X_c(t)$ with $EX_c(t) = 0$ and $\text{var}(X_c(t)) = 1$. It consists of two basic components: (i) a normal approximation to the probability of $X_c(t)$ exceeding some high level (depending on $c$) in (B1), and (ii) the weak convergence of the finite-dimensional distributions of the local increments conditioned on $X_c(t) = c - y/c$ in (B2). The covariance structure of the local increments given by condition (C) and the closely related mean and covariance functions (3.3) of the limiting Gaussian random field in (B2) play a key role in the asymptotic formulas (3.5), (3.7) and (3.8).

References


