Distribution of extreme first passage times of diffusion

Sean D. Lawley∗

October 29, 2019

Abstract

Many events in biology are triggered when a diffusing searcher finds a target, which is called a first passage time (FPT). The overwhelming majority of FPT studies have analyzed the time it takes a single searcher to find a target. However, the more relevant timescale in many biological systems is the time it takes the fastest searcher(s) out of many searchers to find a target, which is called an extreme FPT. In this paper, we apply extreme value theory to find a tractable approximation for the full probability distribution of extreme FPTs of diffusion. This approximation can be easily applied in many diverse scenarios, as it depends on only a few properties of the short time behavior of the survival probability of a single FPT. We find this distribution by proving that a careful rescaling of extreme FPTs converges in distribution as the number of searchers grows. This limiting distribution is a type of Gumbel distribution and involves the LambertW function. This analysis yields new explicit formulas for approximations of statistics of extreme FPTs (mean, variance, moments, etc.) which are highly accurate and are accompanied by rigorous error estimates.

1 Introduction

Events in biological systems are often triggered when a diffusing searcher finds a target [1, 2, 3, 4, 5]. Examples range from the initiation of the immune response when a searching T cell finds a cognate antigen [6], to the triggering of calcium release by diffusing IP$_3$ molecules that reach IP$_3$ receptors [7], to gene activation by the arrival of a diffusing transcription factor to a certain gene [8], to animals foraging for food [9, 10]. In such systems, the activation timescale is determined by the first passage time (FPT) of a searcher to a target.

The vast majority of FPT studies have focused on the time it takes a given single searcher to find a target. However, several recent works and commentaries have shown that the relevant timescale in many systems is actually the time it
takes the fastest searcher(s) to find a target out of a large group of searchers \[11, 12, 13, 14, 15, 16, 17, 18, 19, 20\]. For example, roughly \(N = 10^8\) sperm cells search for an egg in human reproduction, but fertilization occurs as soon as a single sperm cell finds the egg \[21, 22, 23, 24\]. Similarly, a calcium-induced calcium release in a dendritic spine occurs when the two fastest calcium ions out of roughly \(N = 10^3\) calcium ions find small Ryanodyne receptors at the base of the spine \[25\].

Importantly, the time it takes the fastest searcher(s) out of many searchers to find a target is much less than the time it takes a given single searcher to find a target. In fact, some have postulated that this is a general mechanism that operates across many biological systems and refer to it as the *redundancy principle* \[11\]. In particular, it is claimed that many seemingly redundant copies of a searcher (molecule, protein, cell, animal, etc.) are not superfluous, but rather have the specific functions of accelerating activation rates and ensuring successful search processes \[11\]. That is, the apparently “extra” copies are in fact necessary for biological function.

To investigate how the number of searchers affects the time it takes the fastest searcher(s) to find a target, consider \(N \gg 1\) independent and identical diffusive searchers. Let \(\tau_1, \ldots, \tau_N\) be their independent and identically distributed (iid) FPTs to reach some target. The first time one of these searchers finds the target is

\[
T_N := \min\{\tau_1, \ldots, \tau_N\}. 
\]

(1)

More generally, the \(k\)th fastest searcher finds the target at time

\[
T_{k,N} := \min \left\{ \{\tau_1, \ldots, \tau_N\} \setminus \bigcup_{j=1}^{k-1} \{T_{j,N}\} \right\}, \quad k \in \{1, \ldots, N\},
\]

(2)

where \(T_{1,N} := T_N\).

While the distribution and statistics of a single FPT, \(\tau_1\), are well understood in a variety of scenarios \[20, 27, 28, 29, 30\], studying the so-called *extreme* FPTs, \(T_{k,N}\), is notoriously difficult, both analytically and numerically \[31, 32, 12, 11, 33, 34\]. The essential difficulty is that extreme FPTs depend on very rare events. Indeed, while a typical searcher tends to wander around before finding the target, the fastest searchers move almost deterministically along the shortest geodesic path to the target \[31\]. This phenomenon is illustrated in Figure 1.

In this paper, we apply the theory of extreme statistics to find a tractable approximation for the full probability distribution of extreme FPTs of diffusion. This approximation can be applied in many scenarios as it depends on only a few properties of the short time behavior of the survival probability of a single FPT. We find this distribution by proving that a careful rescaling of extreme FPTs converges in distribution as the number of searchers grows. This limiting distribution is a type of Gumbel distribution and involves the so-called LambertW function (defined as the inverse of \(f(z) = ze^z\) \[35\]). This analysis yields new explicit formulas for statistics of extreme FPTs (mean, variance, moments, etc.). These formulas are highly accurate and are accompanied by
rigorous error estimates. Further, these formulas confirm and explain a conjecture that extreme FPT statistics can be approximated by a certain infinite series involving iterated logarithms [36].

The rest of the paper is organized as follows. We first summarize our main results in section 2. In section 3 we develop and state our precise mathematical results in more detail. We then illustrate these general results in a few examples in section 4. In the Discussion section, we describe relations to prior work and discuss applications of the theory. Finally, we collect all the mathematical proofs in an appendix.

2 Main results

Let \( \{\tau_n\}_{n \geq 1} \) be an iid sequence of FPTs with survival probability

\[
S(t) := \mathbb{P}(\tau_1 > t).
\]

Assume that \( S(t) \) has the short time behavior,

\[
1 - S(t) \sim A p e^{-C/t} \quad \text{as } t \to 0+,
\]

for some constants \( A > 0, C > 0, \) and \( p \in \mathbb{R} \). Throughout this work,

"\( f \sim g \)" means \( f/g \to 1 \).

We emphasize that [3] is a generic behavior for diffusion processes that holds in many diverse scenarios (see the Discussion section for more details).

Letting \( T_N \) denote the fastest FPT in [1], we prove the following convergence in distribution (Theorem [1]),

\[
\frac{T_N - b_N}{a_N} \to_d X \quad \text{as } N \to \infty,
\]

3
where $X$ has a standard Gumbel distribution, $\mathbb{P}(X > x) = \exp(-e^x)$, and

\begin{equation}
\begin{aligned}
a_N &= \frac{b_N}{\ln(AN)}, & b_N &= \frac{C}{\ln(AN)}, & \text{if } p = 0, \\
a_N &= \frac{b_N}{p(1 + W)}, & b_N &= \frac{C}{pW}, & \text{if } p \neq 0,
\end{aligned}
\end{equation}

(5)

and

\[ W = \begin{cases} 
W_0((C/p)(AN)^{1/p}) & \text{if } p > 0, \\
W_1((C/p)(AN)^{1/p}) & \text{if } p < 0,
\end{cases} \]

where $W_0(z)$ denotes the principal branch of the LambertW function and $W_{-1}(z)$ denotes the lower branch \[35\]. The LambertW function is a fairly standard function that is included in most modern computational software (it is sometimes called the product logarithm or the omega function). We note that Theorem 5 below gives alternative formulas for $a_N$ and $b_N$ which avoid the LambertW function.

The convergence in distribution in (4) means that if $N \gg 1$, then the distribution of the fastest FPT, $T_N$, is approximately Gumbel with shape parameter $b_N$ and scale parameter $a_N$. That is,

\[ \mathbb{P}(T_N > t) \approx \exp\left[ -\exp\left( \frac{t - b_N}{a_N} \right) \right] \quad \text{if } N \gg 1, \]

where $a_N, b_N$ are in (5). Note that essentially all the statistical information about a Gumbel distribution is immediately available (mean, median, mode, variance, moments, probability density function, etc., see Proposition 1 below). Therefore, this result provides all the statistical information for the fastest FPT (approximately for large $N$). For example, we prove that if $\mathbb{E}[T_N] < \infty$ for some $N \geq 1$, then (Theorem 6)

\[ \mathbb{E}[T_N] = b_N - \gamma a_N + o(a_N), \]

Variance($T_N$) = \[\frac{\pi^2}{6} a_N^2 + o(a_N^2)\],

where $\gamma \approx 0.5772$ is the Euler-Mascheroni constant and $f(N) = o(a_N^m)$ means $\lim_{N \to \infty} a_N^{-m} f(N) = 0$.

We prove similar results for the $k$th fastest FPT, $T_{k,N}$, defined in (2). In particular, we prove that the joint distribution of a rescaling of the $k$ fastest FPTs,

\[ \left( \frac{T_{1,N} - b_N}{a_N}, \ldots, \frac{T_{k,N} - b_N}{a_N} \right), \]

converges as $N \to \infty$ to a distribution that we give explicitly (Theorem 7). This result provides explicit approximations for statistics of $T_{k,N}$, including
(Theorem 8),

\[
E[T_{k,N}] = b_N + \psi(k)a_N + o(a_N) = E[T_N] + H_{k-1}a_N + o(a_N),
\]

\[
\text{Variance}(T_{k,N}) = \psi'(k)a_N^2 + o(a_N),
\]

where \( \psi(x) \) is the digamma function and \( H_{k-1} = \sum_{r=1}^{k-1} \frac{1}{r} \) is the \((k-1)\)-th harmonic number.

### 3 Mathematical analysis

#### 3.1 Fastest FPT

Let \( \{\tau_n\}_{n \geq 1} \) be an iid sequence of FPTs with survival probability \( S(t) := P(\tau_1 > t) \). Define the fastest FPT, \( T_N \), as in (1). Since the sequence \( \{\tau_n\}_{n \geq 1} \) is iid, it is immediate that the survival probability of \( T_N \) is

\[
P(T_N > t) = (P(\tau_1 > t))^N = (S(t))^N. \tag{6}
\]

While (6) is the exact distribution of \( T_N \), this formula is not particularly useful for understanding how the distribution depends on parameters or for calculating statistics of \( T_N \). Furthermore, the full survival probability \( S(t) \) of a single FPT is often unknown.

We thus seek a tractable approximation of (6) for large \( N \), which will thus depend only on the short time behavior of \( S(t) \). Now, (6) implies that the limiting distribution of \( T_N \) for large \( N \) is trivial,

\[
\lim_{N \to \infty} P(T_N > t) = \begin{cases} 1 & \text{if } t < t^*, \\ 0 & \text{if } t > t^*, \end{cases}
\]

where \( t^* := \inf\{t > 0 : S(t) < 1\} \). For nontrivial diffusion processes, we typically have \( t^* = 0 \). To ameliorate this problem, we study the distribution of \( T_N \) by finding a rescaling of \( T_N \) that has a nontrivial limiting distribution for large \( N \). Specifically, we find sequences \( \{a_N\}_{N \geq 1} \) and \( \{b_N\}_{N \geq 1} \) so that

\[
\frac{T_N - b_N}{a_N} \to_d X \quad \text{as } N \to \infty,
\]

for some random variable \( X \). In this paper, \( \to_d \) denotes convergence in distribution \( \text{[37]} \), which means

\[
P\left( \frac{T_N - b_N}{a_N} > x \right) = (S(a_Nx + b_N))^N \to G(x) \quad \text{as } N \to \infty \text{ for all } x \in \mathbb{R}, \tag{7}
\]

where \( G(x) = P(X > x) \).

Remarkably, the Fisher-Tippett-Gnedenko Theorem states that if (7) holds for a nondegenerate \( G \), then \( G \) must be either a Weibull, Frechet, or Gumbel distribution \( \text{[38]} \). This theorem is the cornerstone of extreme value theory, and
applies to the minimum or maximum of any sequence of iid random variables \[39, 40, 41\]. Since the limiting distribution must be one of these three types, this classical theorem is an extreme value analog of the central limit theorem.

We prove below that the typical short time behavior of \( S(t) \) ensures that \( G \) must be Gumbel. The following definition and proposition collects some facts about the Gumbel distribution.

**Definition.** A random variable \( X \) has a **Gumbel distribution** with location parameter \( b \in \mathbb{R} \) and scale parameter \( a > 0 \) if

\[
\mathbb{P}(X > x) = \exp \left( - \exp \left( \frac{x - b}{a} \right) \right), \quad \text{for all } x \in \mathbb{R}. \quad (8)
\]

If (8) holds, then we write

\( X =_d \text{Gumbel}(b,a) \).

**Proposition 1.** If \( X =_d \text{Gumbel}(b,a) \), then its survival probability is in (8), its probability density function is

\[
f_X(x) = \frac{1}{a} \exp \left[ \frac{x - b}{a} - \exp \left( \frac{x - b}{a} \right) \right], \quad x \in \mathbb{R},
\]

and its moment generating function is

\[
M_X(t) := \mathbb{E}[e^{tX}] = \Gamma(1 + at)e^{bt}, \quad t \in \mathbb{R},
\]

where \( \Gamma(\cdot) \) denotes the gamma function. Hence, the mean and variance are

\[
\mathbb{E}[X] = b - \gamma a, \quad \text{Variance}(X) = \frac{\pi^2}{6}a^2,
\]

where \( \gamma \approx 0.5772 \) is the Euler-Mascheroni constant. The mode and median are

\[
\text{Mode}(X) = b, \quad \text{Median}(X) = b + a \ln(\ln(2)) \approx b - 0.3665a.
\]

Now, it was recently proven that under very general assumptions, the survival probability of a single diffusive FPT has the following short time behavior,

\[
\lim_{t \to 0^+} t \ln(1 - S(t)) = -C < 0, \quad (9)
\]

where \( C = L^2/(4D) > 0 \) and \( D \) is a characteristic diffusivity and \( L \) is a certain geodesic distance \[34\], as long as the diffusive searchers cannot start arbitrarily close to the target. The next proposition shows that if (9) holds, then any nondegenerate limiting distribution \( G \) in (7) must be Gumbel.\footnote{Some authors define a Gumbel distribution slightly differently, by saying that \(-X\) has a Gumbel distribution with shape \(-b\) and scale \(a\) if (8) holds.}
Proposition 2. Let \( \{\tau_n\}_{n \geq 1} \) be an iid sequence of nonnegative random variables with \( S(t) := \mathbb{P}(\tau_1 > t) \), define \( T_N := \min\{\tau_1, \ldots, \tau_N\} \), and suppose (9) holds. If there exists sequences \( \{a_N\}_{N \geq 1} \) and \( \{b_N\}_{N \geq 1} \) with \( a_N > 0 \) and \( b_N \in \mathbb{R} \) so that
\[
\frac{T_N - b_N}{a_N} \xrightarrow{d} X \quad \text{as } N \to \infty,
\]
and \( X \) has a nondegenerate distribution, then \( X =_d \text{Gumbel}(b,a) \) for some \( b \in \mathbb{R} \) and \( a > 0 \).

The condition in (9) implies that
\[
S(t) = 1 - e^{-C/t + h(t)},
\]
where \( h(t) \) is some function satisfying \( th(t) \to 0 \) as \( t \to 0^+ \). The following proposition gives precise conditions on \( h(t) \) which yield rescalings \( \{a_N\}_{N \geq 1} \) and \( \{b_N\}_{N \geq 1} \) so that \( (T_N - b_N)/a_N \) converges in distribution to a Gumbel random variable.

Proposition 3. Let \( \{\tau_n\}_{n \geq 1} \) be an iid sequence of nonnegative random variables with \( S(t) := \mathbb{P}(\tau_1 > t) \), define \( T_N := \min\{\tau_1, \ldots, \tau_N\} \), and assume
\[
1 - S(t) \sim 1 - S_0(t) \quad \text{as } t \to 0^+,
\]
where
\[
S_0(t) = 1 - e^{-C/t + h(t)}, \quad \text{if } t > 0,
\]
for some constant \( C > 0 \) and some function \( h(t) \) that is twice-continuously differentiable for \( t > 0 \) and satisfies
\[
\lim_{t \to 0^+} t^2 h''(t) = 0. \tag{10}
\]
Then
\[
\frac{T_N - b_N}{a_N} \xrightarrow{d} X =_d \text{Gumbel}(0,1) \quad \text{as } N \to \infty,
\]
where
\[
a_N := \frac{-1}{NS'_0(b_N)} > 0, \quad b_N := S_0^{-1}(1 - 1/N) > 0, \quad N \geq 1. \tag{11}
\]

As we will see, it is typically the case that \( h(t) = \ln(At^p) \), which clearly satisfies (10). In this case, we work out the rescalings \( \{a_N\}_{N \geq 1} \) and \( \{b_N\}_{N \geq 1} \).

Theorem 4. Let \( \{\tau_n\}_{n \geq 1} \) be an iid sequence of nonnegative random variables with \( S(t) := \mathbb{P}(\tau_1 > t) \), define \( T_N := \min\{\tau_1, \ldots, \tau_N\} \), and assume there exists constants \( C > 0 \), \( A > 0 \), and \( p \in \mathbb{R} \) so that
\[
1 - S(t) \sim At^p e^{-C/t} \quad \text{as } t \to 0^+.
\]
Then

\[ \frac{T_N - b_N}{a_N} \to_d X = \text{Gumbel}(0, 1) \quad \text{as} \quad N \to \infty, \quad (12) \]

where

\[ a_N = \frac{b_N}{\ln(AN)}, \quad b_N = \frac{C}{\ln(AN)}, \quad \text{if} \quad p = 0, \]

\[ a_N = \frac{b_N}{p(1 + W)}, \quad b_N = \frac{C}{pW}, \quad \text{if} \quad p \neq 0, \quad (13) \]

and

\[ W = \begin{cases} W_0((C/p)(AN)^{1/p}) & \text{if} \quad p > 0, \\ W_1((C/p)(AN)^{1/p}) & \text{if} \quad p < 0, \end{cases} \quad (14) \]

where \( W_0(z) \) denotes the principal branch of the LambertW function and \( W_1(z) \) denotes the lower branch [35].

If the convergence in distribution in (12) holds for some rescalings \( \{a_N\}_{N \geq 1} \) and \( \{b_N\}_{N \geq 1} \), then we also have that

\[ \frac{T_N - b'_N}{a'_N} \to_d X = \text{Gumbel}(0, 1) \quad \text{as} \quad N \to \infty, \quad (15) \]

for any rescalings \( \{a'_N\}_{N \geq 1} \) and \( \{b'_N\}_{N \geq 1} \) that satisfy [42]

\[ \lim_{N \to \infty} a'_N = 1, \quad \lim_{N \to \infty} \frac{b'_N - b_N}{a_N} = 0. \quad (16) \]

The following theorem gives rescalings which avoid the LambertW functions used in Theorem 4 and are valid for any \( p \in \mathbb{R} \).

**Theorem 5.** Under the assumptions of Theorem 4, we have that

\[ \frac{T_N - b'_N}{a'_N} \to_d X = \text{Gumbel}(0, 1) \quad \text{as} \quad N \to \infty, \]

where

\[ a'_N = \frac{C}{(\ln N)^2}, \quad b'_N = \frac{C}{\ln N} + \frac{Cp\ln(\ln(N))}{(\ln N)^2} - \frac{C\ln(AC^p)}{(\ln N)^2}. \quad (17) \]

The conclusions of Propositions 2-3 and Theorems 4-5 concern convergence in distribution. In general, convergence in distribution does not imply moment convergence [37]. That is, \( X_N \to_d X \) does not necessarily imply \( \mathbb{E}[(X_N)^m] \to \mathbb{E}[X^m] \) for \( m > 0 \). However, Pickands proved that convergence in distribution does imply moment convergence for extreme values [43].
Theorem 6. Under the assumptions of Theorem 4 with \( \{a_N\}_{N \geq 1} \) and \( \{b_N\}_{N \geq 1} \) given by either (13) or (17), assume further that \( \mathbb{E}[T_N] < \infty \) for some \( N \geq 1 \). Then for each moment \( m \in (0, \infty) \), we have that
\[
\mathbb{E} \left[ \left( \frac{T_N - b_N}{a_N} \right)^m \right] \to \mathbb{E}[X^m] \text{ as } N \to \infty, \quad \text{where } X =_{d} \text{Gumbel}(0, 1). 
\]
Therefore,
\[
\mathbb{E}[(T_N - b_N)^m] = a_N^m \mathbb{E}[X^m] + o(a_N^m),
\]
where \( f(N) = o(a_N^m) \) means \( \lim_{N \to \infty} a_N^{-m} f(N) = 0 \). Further, if \( m > 0 \) is an integer, then \( \mathbb{E}[X^m] \) can be calculated explicitly by Proposition 1. For example, we have that
\[
\mathbb{E}[T_N] = b_N - \gamma a_N + o(a_N),
\]
\[
\text{Variance}(T_N) = \frac{\pi^2}{6} a_N^2 + o(a_N^2),
\]
where \( \gamma \approx 0.5772 \) is the Euler-Mascheroni constant.

3.2 \( k \)th fastest FPT

We now extend the results in the previous subsection on the fastest FPT to the \( k \)th fastest FPT,
\[
T_{k,N} := \min \left\{ \{\tau_1, \ldots, \tau_N\} \setminus \cup_{j=1}^{k-1} \{T_{j,N}\} \right\}, \quad k \in \{1, \ldots, N\},
\]
where \( T_{1,N} := T_N \).

Theorem 7. Under the assumptions of Theorem 4 with \( \{a_N\}_{N \geq 1} \) and \( \{b_N\}_{N \geq 1} \) given by either (13) or (17), we have that for each fixed \( k \geq 1 \),
\[
\frac{T_{k,N} - b_N}{a_N} \to_{d} X_k \text{ as } N \to \infty,
\]
where \( X_k \) has the probability density function,
\[
f_{X_k}(x) = \exp(kx - e^x) \frac{1}{(k-1)!}, \quad \text{for all } x \in \mathbb{R}. \tag{19}
\]
Furthermore, for each fixed \( k \geq 1 \), we have the following convergence in distribution for the joint random variables,
\[
\left( \frac{T_{1,N} - b_N}{a_N}, \ldots, \frac{T_{k,N} - b_N}{a_N} \right) \to_{d} X^{(k)} = (X_1, \ldots, X_k) \in \mathbb{R}^k \text{ as } N \to \infty,
\]
where the joint probability density function of \( X^{(k)} \in \mathbb{R}^k \) is
\[
f_{X^{(k)}}(x_1, \ldots, x_k) = \begin{cases} 
\exp(-e^{x_k}) \prod_{r=1}^{k} e^{x_r} & \text{if } x_1 \leq \cdots \leq x_k, \\
0 & \text{otherwise}.
\end{cases}
\]
The following theorem ensures the convergence of the moments of the $k$th fastest FPT.

**Theorem 8.** Under the assumptions of Theorem 6 with $\{a_N\}_{N \geq 1}$ and $\{b_N\}_{N \geq 1}$ given by either (13) or (17), we have that for each moment $m \in (0, \infty)$,

$$
E \left[ \left( \frac{T_{k,N} - b_N}{a_N} \right)^m \right] \rightarrow E[X_k^m] \quad \text{as } N \rightarrow \infty,
$$

(21)

where $X_k$ has the probability density function in (19). Therefore,

$$
E[(T_{k,N} - b_N)^m] = a_N^m E[X_k^m] + o(a_N^m).
$$

Further, if $m > 0$ is an integer, then $E[X_k^m]$ can be explicitly calculated. In particular,

$$
E[T_{k,N}] = b_N + \psi(k)a_N + o(a_N) = E[T_{1,N}] + H_{k-1}a_N + o(a_N),
$$

$$
\text{Variance}(T_{k,N}) = \psi'(k)a_N^2 + o(a_N^2),
$$

where $\psi(x)$ is the digamma function and $H_{k-1} = \sum_{r=1}^{k-1} \frac{1}{r}$ is the $(k-1)$-th harmonic number.

### 4 Numerical examples

We now apply our results to three specific examples.

#### 4.1 One dimension

First consider the case of $N \geq 1$ independent searchers diffusing in one space dimension with diffusivity $D > 0$. Suppose the searchers each start at $L > 0$ and let $\tau_n$ be the first time the $n$th searcher reaches the origin. In this case,

$$
S(t) = \mathbb{P}(\tau_1 > t) = 1 - \text{erfc} \left( \frac{L}{\sqrt{4Dt}} \right), \quad t > 0,
$$

and thus

$$
1 - S(t) \sim \sqrt{\frac{4Dt}{\pi L^2}} e^{-L^2/(4Dt)} \quad \text{as } t \rightarrow 0^+.
$$

Therefore, Theorems 4-8 hold with

$$
A = \sqrt{\frac{4D}{\pi L^2}}, \quad p = \frac{1}{2}, \quad C = \frac{L^2}{4D}.
$$

In particular,

$$
\frac{T_N - b_N}{a_N} \rightarrow_d X = \text{Gumbel}(0,1) \quad \text{as } N \rightarrow \infty,
$$

(21)
where \( \{a_N\}_{N \geq 1} \) and \( \{b_N\}_{N \geq 1} \) are given by either (13) or (17). Hence, the distribution of \( T_N \) is approximately Gumbel\((b_N, a_N)\).

In the left panel of Figure 2, we plot the error of various approximations of the mean fastest FPT, \( \text{E}[T_N] \), as functions of \( N \). Specifically, we plot the relative error,

\[
\left| \frac{\text{E}[T_N] - T_N}{\text{E}[T_N]} \right|,
\]

where \( T_N \) is an approximation of \( \text{E}[T_N] \). The value of \( \text{E}[T_N] \) used in (22) is calculated by numerical approximation of the following integral,

\[
\text{E}[T_N] = \int_0^\infty (S(t))^N \, dt.
\]

The red dotted curve in the left panel of Figure 2 is the error (22) for the approximation \( T_N = L^2/(4D \ln N) \) (this approximation dates back to [31]). The blue dashed curve is for the approximation \( T_N = b'_N - \gamma a'_N \) where \( \{a'_N\}_{N \geq 1} \) and \( \{b'_N\}_{N \geq 1} \) are given by (17). The black solid curve is for the approximation \( T_N = b_N - \gamma a_N \) where \( \{a_N\}_{N \geq 1} \) and \( \{b_N\}_{N \geq 1} \) are given by (13). This figure shows that our approximations, \( b_N - \gamma a_N \) and \( b'_N - \gamma a'_N \), to the mean fastest FPT are much more accurate than \( L^2/(4D \ln N) \), and \( b_N - \gamma a_N \) is more accurate than \( b'_N - \gamma a'_N \).

In the left panel of Figure 3 we illustrate the convergence in distribution of \( (T_N - b_N)/a_N \) to \( X = \text{d} \) Gumbel\((0, 1)\) where \( \{a_N\}_{N \geq 1} \) and \( \{b_N\}_{N \geq 1} \) are given by (13). Specifically, we plot the probability density function of \( (T_N - b_N)/a_N \) for \( N \in \{10^2, 10^4, 10^6\} \), which approaches the density of \( X \) (namely, \( f_X(x) = \exp(x - e^x) \)) as \( N \) increases. In both Figures 2 and 3 we take \( L = D = 1 \).
4.2 Partial absorption

Consider the example in the previous subsection, but now suppose the target at the origin is partially absorbing. This means that when a searcher hits the target, it is either absorbed or reflected, and the probabilities of these two events are described by a parameter $\kappa > 0$ called the reactivity or absorption rate \[44\]. Mathematically, this means the Fokker-Planck equation describing the probability density for a searcher’s position has a Robin boundary condition at the origin involving the parameter $\kappa > 0 \[45\].

In this case, the survival probability for a single searcher is \[46\]

$$S(t) = \mathbb{P}(\tau_1 > t) = 1 - \text{erfc} \left( \frac{L}{\sqrt{4Dt}} \right) + e^{-\frac{\kappa t + L}{\kappa}} \text{erfc} \left( \frac{2\kappa t + L}{\sqrt{4Dt}} \right), \quad t > 0,$$

and thus

$$1 - S(t) \sim \frac{4}{\sqrt{\pi}} \frac{\kappa L}{D} \left( \frac{Dt}{L^2} \right)^{3/2} e^{-L^2/(4Dt)} \quad \text{as} \quad t \to 0^+.$$ 

Therefore, Theorems \[18\] hold with

$$A = \frac{4}{\sqrt{\pi}} \frac{\kappa L}{D} \left( \frac{D}{L^2} \right)^{3/2}, \quad p = \frac{3}{2}, \quad C = \frac{L^2}{4D}.$$

The middle panel of Figure 2 shows the relative error \(22\) of approximations to $\mathbb{E}[T_N]$ in this case of a partially absorbing target. The red dotted curve is again the error for the approximation $T_N = L^2/(4D \ln N)$ (this approximation was recently found and proven to have the correct large $N$ asymptotics \[34\]). The blue dashed and black solid curves again correspond respectively to $T_N = b'_N - \gamma a'_N$ and $T_N = b_N - \gamma a_N$ where $a'_N, b'_N$ are in \[17\] and $a_N, b_N$ are in \[13\].
Again, \(b_N - \gamma a_N\) and \(b'_N - \gamma a'_N\) are much more accurate than \(L^2/(4D \ln N)\), and \(b_N - \gamma a_N\) is more accurate than \(b'_N - \gamma a'_N\).

The middle panel of Figure 3 illustrates the convergence in distribution of \((T_N - b_N)/a_N\) to \(X = \text{dGumbel}(0, 1)\) in this case of a partially absorbing target (again, for \(N \in \{10^2, 10^4, 10^6\}\) and where \(\{a_N\}_{N \geq 1}\) and \(\{b_N\}_{N \geq 1}\) are given by \((13)\)). In both Figures 2 and 3 we take \(L = D = \kappa = 1.43\) Three dimensions

Finally, consider the case where the \(N \geq 1\) independent searchers diffuse in thre
dimensional space, and let \(\tau_n\) be the first time the \(n\)th searcher leaves a sphere of radius \(L > 0\) centered at its starting location. In this case, the survival probability for a single searcher is \((13)\).

\[
S(t) = \mathbb{P}(\tau_1 > t) = 1 - 2\sqrt{\frac{L^2}{\pi D t}} \sum_{j=0}^{\infty} e^{-j(j+1/2)^2} L^2/(4Dt), \quad t > 0,
\]

and thus

\[
1 - S(t) \sim 2\sqrt{\frac{L^2}{\pi D t}} e^{-L^2/(4Dt)} \quad \text{as} \quad t \to 0 +.
\]

Therefore, Theorems 4-8 hold with

\[
A = 2 \sqrt{\frac{L^2}{\pi D}}, \quad p = -\frac{1}{2}, \quad C = \frac{L^2}{4D}.
\]

The right panel of Figure 2 shows the relative error \((22)\) of approximations to \(\mathbb{E}[T_N]\) in this three dimensional example. The red dotted curve is again the error for the approximation \(T_N = L^2/(4D \ln N)\) (this approximation was found in \((20)\)). The blue dashed and black solid curves again correspond respectively to \(T_N = b'_N - \gamma a'_N\) and \(T_N = b_N - \gamma a_N\) where \(a'_N, b'_N\) are in \((17)\) and \(a_N, b_N\) are in \((13)\) and \((14)\). Further, the right panel of Figure 3 illustrates the convergence in distribution of \((T_N - b_N)/a_N\) to \(X = \text{dGumbel}(0, 1)\) in this three dimensional example (again, for \(N \in \{10^2, 10^4, 10^6\}\) and where \(\{a_N\}_{N \geq 1}\) and \(\{b_N\}_{N \geq 1}\) are given by \((13)\)). In both Figures 2 and 3 we take \(L = D = 1.5\) Discussion

In this work, we found tractable approximations for the full probability distribution of extreme FPTs of diffusion. These approximate distributions depend on only three parameters describing the short time behavior of the survival probability of a single searcher, and we proved that these approximations are exact in the many searcher limit. We used our approximate distributions to derive new formulas for statistics of extreme FPTs and prove rigorous error estimates.

Extreme FPTs of diffusion were first studied in 1983 by Weiss, Shuler, and Lindenberg \((31)\), where they found approximations of \(\mathbb{E}[T_{k,N}]\) for large \(N\) in
various one dimensional problems. Statistics of extreme FPTs of diffusion in one dimensional or spherically symmetric domains were further studied in [32, 47, 36, 48, 49, 21]. Recently, approximate formulas for the moments of extreme FPTs of diffusion in more general two and three dimensional domains were derived in [50, 12, 33]. Even more recently, it was proven in significant generality that the $m$th moment of the $k$th fastest FPT has the large $N$ behavior,

$$\mathbb{E}[(T_{k,N})^m] \sim \left(\frac{L^2}{4D \ln N}\right)^m \text{ as } N \to \infty,$$

(23)

where $D$ is a characteristic diffusivity and $L$ is a certain geodesic distance [34].

The moment formulas derived in the present work agree with (23) to leading order, but are much more accurate for finite $N$. In addition, the moment formulas in the present work explain and confirm a remarkable conjecture in [47]. In that work, the authors conjectured that the mean fastest FPT to escape a ball of radius $L$ in dimension $d \geq 2$ has the following approximation,

$$\mathbb{E}[T_{1,N}] \approx \frac{L^2}{4D \ln N} \left[1 + \sum_{n=1}^{\infty} (\ln N)^{-n} \sum_{m=0}^{n} K_m(n)(\ln \ln N)^m\right],$$

(24)

for some unknown constants $\{\{K_m(n)\}_{m=0}^{n}\}_{n \geq 1}$ (some of which were estimated numerically). To derive (24) from our results, first note that the principal branch of the LambertW function has the following expansion for $z \gg 1$ [35],

$$W_0(z) = L_1 - L_2 + \sum_{i=0}^{\infty} \sum_{j=1}^{\infty} c_{ij} L_1^{-i-j} L_2^j,$$

(25)

where $L_1 = \ln z$, $L_2 = \ln \ln z$,

$$c_{ij} = \frac{(-1)^i}{j!} \left[\begin{array}{l} i + j \\ i + 1 \end{array}\right],$$

and $\left[\begin{array}{l} i+j \\ i+1 \end{array}\right]$ are non-negative Stirling numbers of the first kind. Similarly, the lower branch, $W_{-1}(z)$, has the expansion in (25) for $-1 \ll z < 0$ if $L_1 = \ln(-z)$ and $L_2 = \ln(-\ln(-z))$ [35]. Therefore, upon using the definitions in (13)-(14) and the expansion in (25), it follows that our formula $\mathbb{E}[T_{1,N}] \approx b_N - \gamma a_N$ is exactly of the form in the conjecture (24).

Finally, we emphasize that our results apply to any FPT problem where the survival probability $S(t) = \mathbb{P}(\tau_1 > t)$ of a single searcher satisfies

$$1 - S(t) \sim At^p e^{-C/t} \quad \text{as } t \to 0+,$$

(26)

for some constants $C > 0$, $A > 0$, and $p \in \mathbb{R}$. The behavior in (26) is very generic for diffusion processes and holds in many diverse scenarios. For example, reference [31] found (26) for one-dimensional drift-diffusion processes with a broad class of potential (drift) fields. Similarly, reference [36] found (26) for the first time a pure diffusion in dimension $d \geq 1$ moves any distance $L > 0$ from its
initial location (and referred to [26] as a “universal” form). Further, reference 50 formally derived (26) for a pure diffusion searching for an arbitrarily placed small target in a hyperspherical domain in dimension $d = 3$. It is also known that [26] holds for pure diffusion in dimension $d = 1$ with a partially absorbing target (see section 4.2 above). Further, it was proven in [34] that under very general conditions (including (i) diffusions in $\mathbb{R}^d$ with space-dependent diffusivities and drift fields and (ii) diffusions on $d$-dimensional smooth Riemannian manifolds that may contain reflecting obstacles), the survival probability satisfies

$$\lim_{t \to 0^+} t \ln(1 - S(t)) = -\frac{L^2}{4D} < 0,$$

(27)

where $D > 0$ is a characteristic diffusivity and $L > 0$ is a certain geodesic distance that depends on any space-dependence or anisotropy in the diffusivity (if the diffusivity is constant in space, then $L$ is merely the shortest distance from the starting location to the target). Therefore, if (26) holds in a particular problem, then (27) implies that $C = L^2/(4D)$, and thus the only parameters to be found are $A$ and $p$.

To illustrate how to apply our results in a particular FPT problem, note that reference [12] formally derived (26) for a pure diffusion searching for a small target on an otherwise reflecting boundary of an arbitrary domain in dimensions $d = 2$ or $d = 3$. Specifically, in dimension $d = 3$, the authors of [12] derived

$$1 - S(t) = P(\tau < t) \sim \frac{a^2}{L\sqrt{\pi D t}} e^{-L^2/(4Dt)} \text{ as } t \to 0^+,$$

(28)

where $D$ is the diffusivity, $a$ is the target radius, and $L$ is the distance from the initial searcher location to the target. Hence, if (28) is valid, then our results (namely, Theorems 4-8) hold with

$$A = \frac{a^2}{L\sqrt{\pi D}}, \quad p = -\frac{1}{2}, \quad C = \frac{L^2}{4D}.$$

In particular, our results immediately yield the approximate distribution and statistics of $T_{k,N}$ in this problem.

**Acknowledgments**

The author was supported by the National Science Foundation (Grant Nos. DMS-1814832 and DMS-1148230).

6 Appendix

In this appendix, we collect the proofs of the propositions and theorems of section 3.
Proof of Proposition 1. This proposition merely collects basic results on Gumbel random variables, all of which follow directly from [3].

Proof of Proposition 2. Since most results in extreme value theory are formulated in terms of the maximum of a set of random variables, define

\[ M_N := \max \{-\tau_1, \ldots, -\tau_N\} = -T_N, \]

and \( F(x) = \mathbb{P}(-\tau_1 < x) = S(-x) \). If there exists normalizing constants \( \{a_N\}_{N \geq 1} \) and \( \{b_N\}_{N \geq 1} \) so that \((M_N - b_N)/a_N\) converges in distribution as \( N \to \infty \) to a nontrivial random variable, then the distribution of that random variable can only be Frechet, Weibull, or Gumbel [38]. Since

\[ x^* := \sup \{x : F(x) < 1\} = 0 < \infty, \]

Theorem 1.2.1 in [40] ensures that the limiting distribution cannot be Frechet.

Furthermore, if the limiting distribution is Weibull, then Theorem 1.2.1 in [40] guarantees that there exists some \( \gamma < 0 \) so that

\[ \lim_{t \to 0^+} \frac{1 - F(-tx)}{1 - F(-t)} = \lim_{t \to 0^+} \frac{1 - S(tx)}{1 - S(t)} = x^{-1/\gamma} \quad \text{for all } x > 0. \quad (29) \]

Now, it follows directly from (9) that \( S(t) = 1 - e^{-C/t + h(t)} \) for some function \( h(t) \) satisfying

\[ \lim_{t \to 0^+} th(t) = 0. \quad (30) \]

Therefore, we claim that (29) is violated with, for example, \( x = 2 \). To see this, note that

\[ \lim_{t \to 0^+} \frac{1 - S(2t)}{1 - S(t)} = \lim_{t \to 0^+} e^{C/t + h(2t) - h(t)}. \]

By (30), we are assured that

\[ -\frac{C}{2t} \leq h(t) \leq \frac{C}{2t} \quad \text{for sufficiently small } t. \]

Hence,

\[ \lim_{t \to 0^+} e^{C/t + h(2t) - h(t)} \geq \lim_{t \to 0^+} e^{3C/(4t)} = +\infty, \]

which indeed violates (29). Therefore, if the limiting distribution is nondegenerate, it must be Gumbel.

Proof of Proposition 3. Using the assumptions on \( h(t) \) in [10], a direct calculation shows that

\[ \lim_{t \to 0^+} \frac{d}{dt} \left( \frac{1 - S_0(t)}{S_0'(t)} \right) = 0. \]
Therefore, Theorem 2.1.2 in [41] ensures that
\[
\lim_{N \to \infty} (S_0(a_N x + b_N))^N = \exp(-e^x), \quad \text{for all } x \in \mathbb{R},
\] (31)
for some choice of normalizing constants \(\{a_N, b_N\}\) \(N \geq 1\). Remark 1.1.9 in [40] implies we can take \(a_N\) and \(b_N\) as in [41].

Now, (31) is equivalent to
\[
\lim_{N \to \infty} N \ln(S_0(a_N x + b_N)) = -e^x, \quad \text{for all } x \in \mathbb{R}.
\] Hence, it must be the case that \(S_0(a_N x + b_N) \to 1\) as \(N \to \infty\), and thus a straightforward application of L'Hospital's rule gives
\[
-\ln(S_0(a_N x + b_N)) \sim 1 - S_0(a_N x + b_N) \quad \text{as } N \to \infty.
\] Therefore, (31) is equivalent to
\[
\lim_{N \to \infty} N(1 - S_0(a_N x + b_N)) = e^x, \quad \text{for all } x \in \mathbb{R}.
\] (32)
Now, \(1 - S_0(t) \sim 1 - S_0(t)\) as \(t \to 0^+\) by assumption. Hence, (32) holds with \(S_0\) replaced by \(S\), which then implies that (31) holds with \(S_0\) replaced by \(S\), which completes the proof.

**Proof of Theorem 4.** The theorem follows from Proposition 3 upon calculating \(\{a_N\}_{N \geq 1}\) and \(\{b_N\}_{N \geq 1}\) in [13] for \(h(t) = \ln(At^p)\) and using properties of the LambertW function [35].

**Proof of Theorem 5.** The theorem follows immediately from Theorem 4 and (15)-(16).

**Proof of Theorem 6.** By assumption, \(E[T_N] < \infty\) for some \(N \geq 1\). Hence, if \(m \in (0, 1)\), then \(E[(T_N)^m] \leq 1 + E[T_N] < \infty\). If \(m \geq 1\), then it is straightforward to check that (see the proof of Proposition 2 in [34])
\[
E[(T_{2^m-1,N})^m] < \infty.
\]
Since \(E[X^m] < \infty\), applying Theorem 2.1 in [33] completes the proof.

**Proof of Theorem 7.** The convergence in distribution in (18) and (20) follows immediately from Theorem 4 above and Theorem 3.5 in [39].

**Proof of Theorem 8.** While convergence in distribution does not necessarily imply convergence of moments, it does imply convergence of moments if the sequence of random variables is uniformly integrable [37]. Hence, it is sufficient to prove that
\[
\sup_N E\left[\left(\frac{T_{k,N} - b_N}{a_N}\right)^2\right] < \infty
\] (33)
since (33) ensures that \( \{ \frac{T_{k,N} - b_N}{a_N} \} \) is uniformly integrable [37].

By assumption, \( 1 - S(t) \sim Atpe^{-C/t} \) as \( t \to 0^+ \). Hence, there exists a \( \delta > 0 \) so that

\[
1 - A_1 t pe^{-C/t} \leq 1 - S(t) \leq 1 - A_0 t pe^{-C/t}, \quad \text{if } t \in (0, \delta],
\]

where \( 0 < A_0 < A < A_1 \). Define the survival probability

\[
S_+(t) = \begin{cases} 
1 & t \leq 0, \\
1 - A_0 t pe^{-C/t} & t \in (0, \delta], \\
S(t) & t > \delta.
\end{cases}
\]

Define \( S_-(t) \) similarly with \( A_0 \) replaced by \( A_1 \). Hence, \( S_-(t) \leq S(t) \leq S_+(t) \) for all \( t \in \mathbb{R} \). Let \( \{ U_n \} \) be an iid sequence of random variables, each with a uniform distribution on \([0, 1]\). Define

\[
\tau_n := S^{-1}(U_n), \\
\tau_n^- := S^{-1}(U_n), \\
\tau_n^+ := S^{-1}(U_n),
\]

and

\[
T_{k,N} := \min \left\{ \{ \tau_1, \ldots, \tau_N \} \cup \bigcup_{j=1}^{k-1} \{ T_{j,N} \} \right\}, \quad k \in \{1, \ldots, N\},
\]

\[
T_{k,N}^\pm := \min \left\{ \{ \tau_1^\pm, \ldots, \tau_N^\pm \} \cup \bigcup_{j=1}^{k-1} \{ T_{j,N}^\pm \} \right\}, \quad k \in \{1, \ldots, N\},
\]

where \( T_{1,N} := \min\{\tau_1, \ldots, \tau_N\} \) and \( T_{1,N}^\pm := \min\{\tau_1^\pm, \ldots, \tau_N^\pm\} \). By construction, we have that

\[
T_{k,N}^- \leq T_{k,N} \leq T_{k,N}^+ \quad \text{almost surely.}
\]

Therefore, if \( 1_A \) denotes the indicator function on an event \( A \), then

\[
\mathbb{E} \left[ \left( \frac{T_{k,N} - b_N}{a_N} \right)^2 \right] = \mathbb{E} \left[ \left( \frac{T_{k,N} - b_N}{a_N} \right)^2 \mathbf{1}_{T_{k,N} > b_N} \right] + \mathbb{E} \left[ \left( \frac{T_{k,N} - b_N}{a_N} \right)^2 \mathbf{1}_{T_{k,N} < b_N} \right]
\]

\[
\leq \mathbb{E} \left[ \left( \frac{T_{k,N}^+ - b_N}{a_N} \right)^2 \mathbf{1}_{T_{k,N} > b_N} \right] + \mathbb{E} \left[ \left( \frac{T_{k,N}^- - b_N}{a_N} \right)^2 \mathbf{1}_{T_{k,N} < b_N} \right]
\]

\[
\leq \mathbb{E} \left[ \left( \frac{T_{k,N}^+ - b_N}{a_N} \right)^2 \right] + \mathbb{E} \left[ \left( \frac{T_{k,N}^- - b_N}{a_N} \right)^2 \right].
\]

Hence, it remains to show that

\[
\sup_N \mathbb{E} \left[ \left( \frac{T_{k,N}^+ - b_N}{a_N} \right)^2 \right] < \infty.
\]
Now,
\[
E\left[\left(\frac{T_{k,N}^\pm - b_N}{a_N}\right)^2\right] = \int_0^\infty P\left(\left(\frac{T_{k,N}^\pm - b_N}{a_N}\right)^2 > t\right) dt
\]
\[
= \int_0^\infty P(T_{k,N}^\pm - b_N < a_N\sqrt{t}) dt + \int_0^\infty P(b_N - T_{k,N}^\pm > a_N\sqrt{t}) dt
\]
\[
=: I_1 + I_2.
\]
Since \(T_{1,N}^\pm \leq T_{k,N}^\pm\) almost surely for any \(k \in \{1, \ldots, n\}\), we have that
\[
I_2 \leq \int_0^\infty P(b_N - T_{1,N}^\pm > a_N\sqrt{t}) dt \leq \int_0^\infty P\left(\left(\frac{T_{1,N}^\pm - b_N}{a_N}\right)^2 > t\right) dt
\]
\[
= E\left[\left(\frac{T_{1,N}^\pm - b_N}{a_N}\right)^2\right].
\]
Now, Theorem 6 implies that
\[
E\left[\left(\frac{T_{1,N}^\pm - b_N}{a_N}\right)^2\right] \to E[X^2] < \infty \text{ as } N \to \infty,
\]
where \(\{a_N^\pm\}_{N \geq 1}\) and \(\{b_N^\pm\}_{N \geq 1}\) are given by (13) with \(A\) replaced by \(A_0\) or \(A_1\).

Now, it is straightforward to check that there exists \(\alpha^\pm > 0\) and \(\beta^\pm \in \mathbb{R}\) so that
\[
\frac{a_N^\pm}{a_N} \to \alpha^\pm \text{ and } \frac{b_N^\pm - b_N}{a_N} \to \beta^\pm \text{ as } N \to \infty.
\]
Therefore, Proposition 1.1 and Remark 1 in [42] imply that \(E[\left(\frac{T_{1,N}^\pm - b_N}{a_N}\right)^2]\) converges to some finite constant as \(N \to \infty\). Hence,
\[
\sup_N I_2 < \infty.
\]

Moving to \(I_1\), note first that
\[
P(T_{k,N}^\pm > x) = P(T_{1,N}^\pm > x) + \sum_{j=1}^{k-1} P(T_{j,N}^\pm < x < T_{j+1,N}^\pm).
\]
Hence,
\[
I_1 = \int_0^\infty P(T_{1,N}^\pm > a_N\sqrt{t} + b_N) dt + \sum_{j=1}^{k-1} \int_0^\infty P(T_{j,N}^\pm < a_N\sqrt{t} + b_N < T_{j+1,N}^\pm) dt
\]
\[
=: I_3 + I_4.
\]
Now, \(I_3\) can be handled similarly to \(I_2\) to obtain
\[
\sup_N I_3 < \infty.
\]
Hence, it remains to show that \( \sup_N I_4 < \infty \). Now, since \( \{ \tau_{n}^\pm \}_{n \geq 1} \) are iid, it follows that

\[
P(T_{j,N}^+ < x < T_{j+1,N}^+) = \binom{N}{j} (1 - S_\pm(x))^j (S_\pm(x))^{N-j}, \text{ if } j \in \{1, \ldots, k - 1\}.
\]

Hence,

\[
I_4 = \sum_{j=1}^{k-1} \binom{N}{j} \int_0^\infty (1 - S_\pm(a_N \sqrt{t} + b_N))^j (S_\pm(a_N \sqrt{t} + b_N))^{N-j} dt.
\]

An application of Laplace’s method with a movable maximum (see, for example, [51] section 6.4) shows that each term in this sum is bounded in \( N \), and so the proof is complete.

**References**

[1] Tom Chou and Maria R. D’Orsogna. First passage problems in biology. In *First-Passage Phenomena and Their Applications*, pages 306–345. World Scientific, 2014.

[2] D Holcman and Z Schuss. Time scale of diffusion in molecular and cellular biology. *Journal of Physics A: Mathematical and Theoretical*, 47(17):173001, May 2014.

[3] Paul C. Bressloff and Jay M. Newby. Stochastic models of intracellular transport. *Reviews of Modern Physics*, 85(1):135–196, January 2013.

[4] D Holcman and Z Schuss. Control of flux by narrow passages and hidden targets in cellular biology. *Reports on Progress in Physics*, 76(7):074601, 2013.

[5] Z. Schuss, A. Singer, and David Holcman. The narrow escape problem for diffusion in cellular microdomains. *Proceedings of the National Academy of Sciences*, 104(41):16098–16103, 2007.

[6] Monica I Delgado, Michael J Ward, and Daniel Coombs. Conditional mean first passage times to small traps in a 3-D domain with a sticky boundary: Applications to T cell searching behavior in lymph nodes. *Multiscale Modeling & Simulation*, 13(4):1224–1258, 2015.

[7] Samuel S Wang, Adawia A Alousi, and Stuart H Thompson. The lifetime of inositol 1, 4, 5-trisphosphate in single cells. *The Journal of General Physiology*, 105(1):149–171, 1995.

[8] Daniel R Larson, Daniel Zenklusen, Bin Wu, Jeffrey A Chao, and Robert H Singer. Real-time observation of transcription initiation and elongation on an endogenous yeast gene. *Science*, 332(6028):475–478, 2011.
[9] HW McKenzie, MA Lewis, and EH Merrill. First passage time analysis of animal movement and insights into the functional response. *Bull Math Biol*, 71(1):107–129, 2009.

[10] V Kurella, J C Tzou, D Coombs, and M J Ward. Asymptotic analysis of first passage time problems inspired by ecology. *Bull Math Biol*, 77(1):83–125, 2015.

[11] Z. Schuss, K. Basnayake, and D. Holcman. Redundancy principle and the role of extreme statistics in molecular and cellular biology. *Physics of Life Reviews*, January 2019.

[12] K Basnayake, Z Schuss, and D Holcman. Asymptotic formulas for extreme statistics of escape times in 1, 2 and 3-dimensions. *J Nonlinear Sci*, 29(2):461–499, 2019.

[13] D Coombs. First among equals: Comment on “Redundancy principle and the role of extreme statistics in molecular and cellular biology” by Z. Schuss, K. Basnayake and D. Holcman. *Physics of life reviews*, 28:92–93, 2019.

[14] S Redner and B Meerson. Redundancy, extreme statistics and geometrical optics of brownian motion. comment on “Redundancy principle and the role of extreme statistics in molecular and cellular biology” by Z. Schuss et al. *Physics of life reviews*, 28:80–82, 2019.

[15] I M Sokolov. Extreme fluctuation dominance in biology: On the usefulness of wastefulness: Comment on “Redundancy principle and the role of extreme statistics in molecular and cellular biology” by Z. Schuss, K. Basnayake and D. Holcman. *Physics of life reviews*, 2019.

[16] D A Rusakov and L P Savtchenko. Extreme statistics may govern avalanche-type biological reactions: Comment on “Redundancy principle and the role of extreme statistics in molecular and cellular biology” by Z. Schuss, K. Basnayake, D. Holcman. *Physics of life reviews*, 2019.

[17] L M Martyushev. Minimal time, weibull distribution and maximum entropy production principle. comment on “Redundancy principle and the role of extreme statistics in molecular and cellular biology” by Z. Schuss et al. *Physics of life reviews*, 28:83–84, 2019.

[18] M V Tamm. Importance of extreme value statistics in biophysical contexts: Comment on “Redundancy principle and the role of extreme statistics in molecular and cellular biology.”. *Physics of life reviews*, 2019.

[19] K. Basnayake, A. Hubl, Z. Schuss, and D. Holcman. Extreme Narrow Escape: Shortest paths for the first particles among N to reach a target window. *Physics Letters A*, 382(48):3449–3454, December 2018.
[20] Claire Guerrier and David Holcman. The First 100 nm Inside the Pre-synaptic Terminal Where Calcium Diffusion Triggers Vesicular Release. *Frontiers in Synaptic Neuroscience*, 10, July 2018.

[21] B Meerson and S Redner. Mortality, redundancy, and diversity in stochastic search. *Phys Rev Lett*, 114(19):198101, 2015.

[22] Karine Reynaud, Zeev Schuss, Nathalie Rouach, and David Holcman. Why so many sperm cells? *Communicative & Integrative Biology*, 8(3):e1017156, May 2015.

[23] Peter W Barlow. Why so many sperm cells? Not only a possible means of mitigating the hazards inherent to human reproduction but also an indicator of an exaptation. *Communicative & Integrative Biology*, 9(4):e1204499, 2016.

[24] J Yang, I Kupka, Z Schuss, and D Holcman. Search for a small egg by spermatozoa in restricted geometries. *Journal of Mathematical Biology*, 73(2):423–446, 2016.

[25] Kanishka Basnayake, David Mazaud, Alexis Bemelmans, Nathalie Rouach, Eduard Korkotian, and David Holcman. Fast calcium transients in dendritic spines driven by extreme statistics. *PLOS Biology*, 17(6):e2006202, June 2019.

[26] O. Benichou, D. Grebenkov, P. Levitz, C. Loverdo, and R. Voituriez. Optimal Reaction Time for Surface-Mediated Diffusion. *Phys Rev Lett*, 105(15):150606, October 2010.

[27] Alexei F Cheviakov, Michael J Ward, and Ronny Straube. An asymptotic analysis of the mean first passage time for narrow escape problems: Part II: The sphere. *Multiscale Model Simul*, 8(3):836–870, 2010.

[28] P C Bressloff and S D Lawley. Escape from subcellular domains with randomly switching boundaries. *Multiscale Model Simul*, 13(4):1420–1445, 2015.

[29] A E Lindsay, A J Bernoff, and M J Ward. First passage statistics for the capture of a brownian particle by a structured spherical target with multiple surface traps. *Multiscale Model Simul*, 15(1):74–109, 2017.

[30] S D Lawley and C E Miles. Diffusive search for diffusing targets with fluctuating diffusivity and gating. *Journal of Nonlinear Science*, 2019. https://doi.org/10.1007/s00332-019-09564-1.

[31] G H Weiss, K E Shuler, and K Lindenberg. Order statistics for first passage times in diffusion processes. *J Stat Phys*, 31(2):255–278, 1983.

[32] S B Yuste and K Lindenberg. Order statistics for first passage times in one-dimensional diffusion processes. *J Stat Phys*, 85(3-4):501–512, 1996.
[33] S D Lawley and J B Madrid. A probabilistic approach to extreme statistics of brownian escape times in dimensions 1, 2, and 3. arXiv:1907.07515, 2019.

[34] SD Lawley. Universal formula for extreme first passage statistics of diffusion. arXiv:1909.09883, 2019.

[35] RM Corless, GH Gonnet, DEG Hare, DJ Jeffrey, and DE Knuth. On the LambertW function. Advances in Computational mathematics, 5(1):329–359, 1996.

[36] S B Yuste, L Acedo, and K Lindenberg. Order statistics for d-dimensional diffusion processes. Phys Rev E, 64(5):052102, 2001.

[37] P Billingsley. Convergence of probability measures. John Wiley & Sons, 2013.

[38] RA Fisher and LHC Tippett. Limiting forms of the frequency distribution of the largest or smallest member of a sample. In Mathematical Proceedings of the Cambridge Philosophical Society, volume 24, pages 180–190. Cambridge University Press, 1928.

[39] S Coles. An introduction to statistical modeling of extreme values, volume 208. Springer, 2001.

[40] L De Haan and A Ferreira. Extreme value theory: an introduction. Springer Science & Business Media, 2007.

[41] M Falk, J Hüsler, and RD Reiss. Laws of small numbers: extremes and rare events. Springer Science & Business Media, 2010.

[42] Z Peng and S Nadarajah. Convergence rates for the moments of extremes. Bulletin of the Korean Mathematical Society, 49(3):495–510, 2012.

[43] J Pickands. Moment convergence of sample extremes. The Annals of Mathematical Statistics, 39(3):881–889, 1968.

[44] D S Grebenkov. Partially reflected brownian motion: a stochastic approach to transport phenomena. Focus on probability theory, pages 135–169, 2006.

[45] S. D. Lawley and J. P. Keener. A new derivation of robin boundary conditions through homogenization of a stochastically switching boundary. SIAM J. Appl. Dyn. Syst., 14(4), 2015.

[46] AJ Bernoff and AE Lindsay. Numerical approximation of diffusive capture rates by planar and spherical surfaces with absorbing pores. SIAM J Appl Math, 78(1):266–290, 2018.

[47] SB Yuste and L Acedo. Diffusion of a set of random walkers in euclidean media. first passage times. J Phys A, 33(3):507, 2000.
[48] H van Beijeren. The uphill turtle race: on short time nucleation probabilities. *J Stat Phys*, 110(3-6):1397–1410, 2003.

[49] S Redner and B Meerson. First invader dynamics in diffusion-controlled absorption. *J Stat Mech*, 2014(6):P06019, 2014.

[50] S Ro and Y W Kim. Parallel random target searches in a confined space. *Phys Rev E*, 96(1):012143, 2017.

[51] CM Bender and SA Orszag. *Advanced mathematical methods for scientists and engineers I: Asymptotic methods and perturbation theory*. Springer Science & Business Media, 2013.