On correlations between certain random variables associated with first passage Brownian motion

Michael J Kearney and Andrew J Pye

*Faculty of Engineering and Physical Sciences, University of Surrey,
Guildford, Surrey, GU2 7XH, UK*

Richard J Martin

*Department of Mathematics, Imperial College London,
South Kensington, London, SW7 2AZ, UK*

(February 2014)

Abstract

We analyse how the area swept out by a Brownian motion up to its first passage time correlates with the first passage time itself, obtaining several exact results in the process. Additionally, we analyse the relationship between the time average of a Brownian motion during a first passage and the maximum value attained. The results, which find various applications, are in excellent agreement with simulations.

PACS: 02.50.-r, 05.40.-a

Email: m.j.kearney@surrey.ac.uk

1. Introduction

Consider a one-dimensional Brownian motion with negative drift,

$$x(t) = x_0 - \eta t + \sigma W(t); \quad \eta \geq 0 \quad (1)$$

where $W(t)$ is the standard Wiener process with $W(0) \equiv 0$ and $x_0 > 0$ is the initial condition. The first passage process is the evolution of (1) up to the random time T at which the path $x(t)$ first reaches zero; see figure 1. In this paper we are primarily interested in the integral (functional) of the process $A = \int_0^T x(t) dt$, itself a random variable corresponding to the area swept out by the path till the first passage time. This area variable, whose probability density $P_\eta(A, x_0)$ and associated properties have been analysed in some detail (see [1, 2] and references therein), has recently been found to arise in the study of intermittent burst processes in solar physics [3], non-orientated animal movement patterns [4], DNA breathing dynamics [5] and evolutionary dynamics in the context of colonising new territories [6]. The latter invokes certain discrete lattice models, in turn making a natural link to earlier work on compact directed percolation [7], the dynamics of avalanche processes [8] and queueing phenomena such as traffic jams [9].

Here we provide fresh insights into such problems by analysing the *joint* probability density $P_\eta(A, T, x_0)$. This allows us to study the correlation between the variables A and T in terms of a generalised correlation coefficient $R_\eta(A, T^k)$. In the zero-drift limit $\eta \rightarrow 0$ this is calculated explicitly for all $k \geq 1$. Moreover, by

conditioning the motion so that the first passage time is fixed, certain results for so-called Brownian excursions [10, 11] are generalized to an arbitrary initial value. As well as the aforementioned examples such excursions arise in the study of, for example, planar random loops [12] and the maximal relative height for fluctuating interfaces [13].

Knowledge of the joint probability density also allows us to study the *time* average $\bar{x} \equiv AT^{-1} = \frac{1}{T} \int_0^T x(t) dt$ of a *given* path up to its first passage time (see figure 1). This quantity (it should be noted that $\bar{x}$ is a random variable; each path has its own associated value) has a natural interpretation in many of the above problems. It is known empirically that when the drift is weak the correlation between $\bar{x}$ and the maximum value x_M during a first passage (see figure 1) is strong and, although a calculation of the correlation coefficient $R_\eta(AT^{-1}, x_M)$ remains elusive, several results are obtained which shed light on this. In particular, we show that the ensemble averaged moments $\langle AT^{-1} \rangle$ and $\langle x_M \rangle$ both diverge logarithmically in the zero-drift limit but in such a way that $\lim_{\eta \rightarrow 0} \langle 2AT^{-1} - x_M \rangle / x_0 = c$, where $c = 0.11593\dots$ is a universal constant.

Throughout the paper, comparison of the theory is made against the results of simulations. Quantities derived from simulations are based, for a given set of parameters, on sample sets of 10^7 realisations, or multiple sets thereof where needed to reduce statistical errors (which are most prevalent when the drift is small).

2. General theory

The variables A and $\bar{x}$ are examples of Brownian functionals; there are many others which find widespread application in the physical sciences and elsewhere, see e.g. [14] and references therein. To study them we adapt and extend the methods outlined in [1] and reviewed in [14]. The fundamental quantity we focus on is the joint probability density, $P_\eta(A, T, x_0)$, whose double Laplace transform may be written as,

$$\tilde{P}_\eta(s, p, x_0) \equiv \int_0^\infty \int_0^\infty P_\eta(A, T, x_0) e^{-sA} e^{-pT} dA dT = \left\langle e^{-sA - pT} \right\rangle \quad (2)$$

where $\langle \dots \rangle$ denotes the ensemble average over all possible paths of the process $x(t)$ till their first-passage time. In turn, this may be written in the form,

$$\tilde{P}_\eta(s, p, x_0) = \left\langle \exp \left(- \int_0^T (sx(\tau) + p) d\tau \right) \right\rangle. \quad (3)$$

Using well developed procedures [14] one can, based on the Feynman-Kac formalism, express (3) in terms of a path integral and hence, using standard manipulations, show that $\tilde{P}_\eta(s, p, x_0)$ satisfies a backward Fokker-Planck equation;

$$\frac{\sigma^2}{2} \frac{\partial^2 \tilde{P}_\eta}{\partial x_0^2} - \eta \frac{\partial \tilde{P}_\eta}{\partial x_0} - (sx_0 + p) \tilde{P}_\eta = 0. \quad (4)$$

The relevant boundary conditions are $\tilde{P}_\eta(s, p, x_0 = 0) = 1$ and $\tilde{P}_\eta(s, p, x_0 \rightarrow \infty) = 0$.

Through the use of the integrating factor $e^{\eta x_0 / \sigma^2}$ one can eliminate the first derivative term and reduce (4) to an equation of Airy's form, whereupon the full solution is readily obtained as,

$$\tilde{P}_\eta(s, p, x_0) = e^{\eta x_0 / \sigma^2} \frac{\text{Ai}\left(\frac{\eta^2 + 2\sigma^2 p + 2\sigma^2 s x_0}{2^{2/3} s^{2/3} \sigma^{8/3}}\right)}{\text{Ai}\left(\frac{\eta^2 + 2\sigma^2 p}{2^{2/3} s^{2/3} \sigma^{8/3}}\right)} \quad (5)$$

where $\text{Ai}(z)$ is the standard Airy function which obeys the second order differential equation $\text{Ai}''(z) = z\text{Ai}(z)$ [15; Sect. 10.4.1]. The factor in the denominator of (5) (which is independent of x_0) ensures the correct overall behaviour at $x_0 = 0$.

The full Laplace inversion of (5) with respect to both s and p to obtain $P_\eta(A, T, x_0)$ is technically demanding but is not required in what follows. However, even without knowing the full solution, by examining the structure of (5) one can see that,

$$\tilde{P}_\eta(s, p, x_0) = e^{\eta x_0 / \sigma^2} \tilde{P}_0\left(s, p + \frac{\eta^2}{2\sigma^2}, x_0\right) \quad (6)$$

whereupon it follows that,

$$P_\eta(A, T, x_0) = e^{\eta x_0 / \sigma^2} e^{-\eta^2 T / 2\sigma^2} P_0(A, T, x_0). \quad (7)$$

This factorization is fundamentally a consequence of a theorem due to Girsanov [16].

It has important consequences, as will become clear later.

Using the well-known asymptotic form $\text{Ai}(z) \sim \frac{1}{2}\pi^{-1/2}z^{-1/4}e^{-2z^{3/2}/3}$ as $z \rightarrow \infty$

[15; Sect. 10.4.59], the leading term of $\tilde{P}_\eta(s, p, x_0)$ as $s \rightarrow 0$ tells us that,

$$\tilde{P}_\eta(s \rightarrow 0, p, x_0) \equiv \langle e^{-pT} \rangle = e^{\eta x_0 / \sigma^2} e^{-(\eta^2 + 2\sigma^2 p)^{1/2} x_0 / \sigma^2}. \quad (8)$$

This is simply the Laplace transform of the first passage time probability density

$P_\eta(T, x_0)$ and it may be inverted to give the well-known result [17],

$$P_\eta(T, x_0) = \frac{x_0}{\sqrt{2\pi\sigma^2}} \frac{1}{T^{3/2}} \exp\left\{-\frac{(x_0 - \eta T)^2}{2\sigma^2 T}\right\}. \quad (9)$$

For $\eta \geq 0$ the first passage time is finite with probability one; the process is null-recurrent when $\eta = 0$ and in the zero-drift limit the moments $\langle T^k \rangle$ diverge for $k \geq 1$.

This is a consequence of the fact that when the drift is small, paths of long duration become dominant. The mean and the variance are easily calculated;

$$\langle T \rangle = \frac{x_0}{\eta}; \quad \text{Var}(T) = \frac{x_0 \sigma^2}{\eta^3} \quad (10)$$

and, with a little more effort, all the integer moments (which can be expressed in terms of a modified Bessel function of order $k - \frac{1}{2}$ [1]) can be evaluated as a finite sum using a known Bessel function identity [15; Sect. 10.2.15];

$$\langle T^k \rangle = \sum_{j=1}^k \frac{(2k-1-j)!}{(j-1)!(k-j)!2^{k-j}} \frac{x_0^j \sigma^{2k-2j}}{\eta^{2k-j}}. \quad (11)$$

This expression is needed at various points in the analysis which follows.

3. Main results

In figure 2, the results of simulations for a particular value of the drift are presented showing the correlation between A and T . It is clear visually that the correlation is strong and one would like to understand this in detail. To this end, one can examine the correlator $\langle Ae^{-pT} \rangle$ which is given by;

$$\begin{aligned} \langle Ae^{-pT} \rangle &\equiv -\lim_{s \rightarrow 0} \frac{\partial}{\partial s} \tilde{P}_\eta(s, p, x_0) \\ &= e^{\eta x_0 / \sigma^2} e^{-(\eta^2 + 2\sigma^2 p)^{1/2} x_0 / \sigma^2} \left[\frac{x_0 \sigma^2}{2(\eta^2 + 2\sigma^2 p)} + \frac{x_0^2}{2(\eta^2 + 2\sigma^2 p)^{1/2}} \right]. \end{aligned} \quad (12)$$

It is a straightforward exercise to obtain (12) from (5) using the *full* asymptotic expansions of $\text{Ai}(z)$ and its derivative $\text{Ai}'(z)$ as $z \rightarrow \infty$ [15; Sect. 10.4.59 and Sect. 10.4.61], but the calculation is somewhat tedious. A faster approach is to differentiate each term of (4) with respect to s , then set $s = 0$ before solving the resulting

equation (which has a driving term involving $\tilde{P}_\eta(s=0, p, x_0)$, which is given by (8)).

Similarly, the correlator $\langle A^2 e^{-pT} \rangle$ may be determined using an obvious extension involving the second derivative with respect to s ;

$$\begin{aligned} \langle A^2 e^{-pT} \rangle = & e^{\eta x_0 / \sigma^2} e^{-(\eta^2 + 2\sigma^2 p)^{1/2} x_0 / \sigma^2} \times \\ & \left[\frac{5x_0^6 \sigma^6}{4(\eta^2 + 2\sigma^2 p)^{5/2}} + \frac{5x_0^2 \sigma^4}{4(\eta^2 + 2\sigma^2 p)^2} \right. \\ & \left. + \frac{5x_0^3 \sigma^2}{6(\eta^2 + 2\sigma^2 p)^{3/2}} + \frac{x_0^4}{4(\eta^2 + 2\sigma^2 p)} \right] \end{aligned} \quad (13)$$

and so on for higher orders as required. The structure of (12) and (13) reflect the structure of (7) and the mean and variance of the area variable may be derived immediately by setting $p = 0$ [2];

$$\langle A \rangle = \frac{x_0 \sigma^2}{2\eta^2} + \frac{x_0^2}{2\eta}; \quad \text{Var}(A) = \frac{5x_0 \sigma^6}{4\eta^5} + \frac{x_0^2 \sigma^4}{\eta^4} + \frac{x_0^3 \sigma^2}{3\eta^3}. \quad (14)$$

Not unexpectedly, given the behaviour of (10), these quantities also diverge in the limit of zero-drift.

The correlation coefficient $R(X, Y) \equiv \text{Cov}(X, Y) / \sqrt{\text{Var}(X)\text{Var}(Y)}$, where $\text{Cov}(X, Y) \equiv \langle XY \rangle - \langle X \rangle \langle Y \rangle$, is a standard measure of the *linear* dependence between two random variables, and perfect positive correlation implies $R = 1$. In the present context the natural quantity to study is $R_\eta(A, T)$; more generally, since an important

objective is to understand what knowledge of T tells us about A , it is instructive to study $R_\eta(A, T^k)$ as a function of k so as to ascertain the value of k for which the correlation is strongest. One could also consider $R_\eta(A^j, T)$ or even $R_\eta(A^j, T^k)$, but analytical tractability is more of an issue and the additional insights gained are comparatively modest. From (12) one obtains for $k = 1$,

$$\langle AT \rangle \equiv -\frac{\partial}{\partial p} \left\langle A e^{-pT} \right\rangle \Big|_{p=0} = \frac{x_0 \sigma^4}{\eta^4} + \frac{x_0^2 \sigma^2}{\eta^3} + \frac{x_0^3}{2\eta^2} \quad (15)$$

from which, in conjunction with (10) and (14), one may derive,

$$R_\eta(A, T) = \left(\frac{12 + 12\lambda + 3\lambda^2}{15 + 12\lambda + 4\lambda^2} \right)^{1/2}; \quad \lambda \equiv \frac{\eta x_0}{\sigma^2}. \quad (16)$$

This is a function of the dimensionless parameter λ only and is plotted as such in figure 3; the agreement with the results of simulations is excellent. To calculate $R_\eta(A, T^k)$ for integer $k > 1$ one can use the natural generalization of (15) in conjunction with (11) and (14). As an illustration, $R_\eta(A, T^2)$ is also shown in figure 3 and again the agreement with the results of simulations is excellent. Of particular interest is the zero-drift limit and for any positive integer k ,

$$R_0(A, T^k) \equiv \lim_{\eta \rightarrow 0} R_\eta(A, T^k) = 4^k k! \sqrt{\frac{(2k-1)!}{10(4k-2)!}}. \quad (17)$$

This expression is independent of the parameters of the problem, i.e. it is purely geometric in origin, and therefore it relates fundamentally to the structure of Brownian motion itself. It should be noted that one cannot calculate $R_0(A, T^k)$ directly (i.e. without invoking such a limiting procedure) since the key quantities involved are infinite.

If one assumes that it is legitimate to analytically continue (17) to non-integer values of $k > 1$ using the representation $k! = \Gamma(k+1)$, the strongest correlation occurs when $k = 3/2$ with $R_0(A, T^{3/2}) = \sqrt{3\pi/10} \approx 0.971$, which agrees to high precision with the data shown in figure 3. This suggests that for paths of long duration the ‘typical’ area will scale as $T^{3/2}$, the argument being that such paths contribute most when the drift is small. Such a scaling can also be argued on dimensional grounds based on an assumed irrelevance of the initial condition as $T \rightarrow \infty$. From a different perspective, if we define $Z \equiv A - \beta_k T^k$ for some β_k , then the choice $\beta_k \equiv \text{Cov}(A, T^k) / \text{Var}(T^k)$ ensures that $\text{Cov}(Z, T^k) = 0$ and minimises $\text{Var}(Z)$. Using the above results one finds that in the zero-drift limit,

$$\beta_k = 2^{3k-2} \frac{k!(2k-1)!}{(4k-2)!} \left(\frac{\eta^{2k-3}}{\sigma^{2k-4}} \right) [1 + O(\eta)]. \quad (18)$$

Based on the idea that paths of long duration dominate when the drift is small, the defining equation $Z \equiv A - \beta_k T^k$ suggests that $\lim_{T \rightarrow \infty} A/T^k \approx \lim_{\eta \rightarrow 0} \beta_k$ in some ‘average’ sense. For $k < 3/2$ one has $\lim_{\eta \rightarrow 0} \beta_k = \infty$ and for $k > 3/2$ one has

$\lim_{\eta \rightarrow 0} \beta_k = 0$, whilst for $k = 3/2$ (after analytic continuation) one sees that

$\lim_{\eta \rightarrow 0} \beta_{3/2} = \sqrt{\pi/8} \sigma$. This implies that ‘typically’ $A \approx \sqrt{\pi/8} \sigma T^{3/2}$ as $T \rightarrow \infty$.

We now wish to make this analysis precise. To do so we start by calculating the quantity $\langle e^{-sA} \rangle_T$, where $\langle \dots \rangle_T$ denotes an ensemble average over paths subject to the constraint that the first-passage time T is *fixed*. The key defining expression is,

$$\langle e^{-sA} \rangle_T \equiv \frac{\int_0^\infty e^{-sA} P_\eta(A, T, x_0) dA}{P_\eta(T, x_0)} = \frac{\int_0^\infty e^{-sA} P_0(A, T, x_0) dA}{P_0(T, x_0)} \quad (19)$$

where the simplification occurring at the second step is a consequence of (7). It is immediately clear that $\langle e^{-sA} \rangle_T$ is *independent* of the drift, the formal explanation being that the probabilistic weight of any path of the process (1) conditioned to be fixed at both ends is independent of the drift. Technically, the set of such paths has measure zero but this is not an obstacle to obtaining meaningful results, e.g. as discussed in [10]. By rearranging (19) and taking the Laplace transform of both sides of the resulting equation with respect to the first passage time T , one can employ (2) and (5) to write down,

$$\int_0^\infty e^{-pT} \langle e^{-sA} \rangle_T P_0(T, x_0) dT = \tilde{P}_0(s, p, x_0) = \frac{\text{Ai}\left(\frac{2^{1/3} p + 2^{1/3} s x_0}{s^{2/3} \sigma^{2/3}}\right)}{\text{Ai}\left(\frac{2^{1/3} p}{s^{2/3} \sigma^{2/3}}\right)}. \quad (20)$$

The reciprocal Airy function $1/\text{Ai}(z)$ is analytic everywhere except for a series of simple poles which all lie on the negative real axis at $z = -\alpha_k$, with $\alpha_1 = 2.338\dots$, $\alpha_2 = 4.087\dots$ etc. [15; Table 10.13]. Accordingly, the formal inversion of (20) with respect to the variable p is straightforward and one finds in conjunction with (9) that,

$$\begin{aligned} \left\langle e^{-sA} \right\rangle_T &= \sqrt{2\pi} \left(\frac{s^{2/3} \sigma^{5/3} T^{3/2}}{2^{1/3} x_0} \right) e^{x_0^2 / 2\sigma^2 T} \\ &\times \sum_{k=1}^{\infty} \frac{\text{Ai}\left(\frac{2^{1/3} x_0 s^{1/3}}{\sigma^{2/3}} - \alpha_k\right)}{\text{Ai}'(-\alpha_k)} e^{-\alpha_k s^{2/3} \sigma^{2/3} T^{1/3}}. \end{aligned} \quad (21)$$

This is a generalization of a famous result for Brownian excursions, which may be recovered by taking the limit $x_0 \rightarrow 0$. Thus, introducing the dimensionless variable $y = s\sigma T^{3/2}$ one finds that,

$$\lim_{x_0 \rightarrow 0} \left\langle e^{-yA / \sigma T^{3/2}} \right\rangle_T = y \sqrt{2\pi} \sum_{k=1}^{\infty} e^{-\alpha_k y^{2/3} / 2^{1/3}}. \quad (22)$$

This is the Laplace transform of the so-called Airy distribution [10, 11]. Other derivations of this result may be found in [18, 19], and a direct derivation using path integral techniques is given in [10]. The derivation here has the virtue that the conditioning on paths of fixed duration is implemented directly, and there is no need to explicitly evaluate any path integral using the eigenfunctions and eigenvalues of its corresponding Hamiltonian operator.

As things stand, one cannot easily use (21) to obtain an expression for $\langle A \rangle_T$.

However, such an expression can be obtained from (20) and (12) as follows;

$$\begin{aligned} \int_0^\infty e^{-pT} \langle A \rangle_T P_0(T, x_0) dT &\equiv -\lim_{s \rightarrow 0} \frac{\partial}{\partial s} \tilde{P}_0(s, p, x_0) \\ &= e^{-(2p)^{1/2} x_0 / \sigma} \left[\frac{x_0}{2(2p)} + \frac{x_0^2}{2\sigma(2p)^{1/2}} \right]. \end{aligned} \quad (23)$$

The Laplace inversion of the right hand side of (23) can be carried out exactly using standard results [15; Sect. 29.3.83 and Sect. 29.3.84] and after suitable rearrangement in conjunction with (9) one has that,

$$\langle A \rangle_T = \frac{x_0 T}{2} + \sqrt{\frac{\pi}{8}} \sigma T^{3/2} e^{x_0^2 / 2T\sigma^2} \operatorname{erfc} \left(\frac{x_0}{\sqrt{2\sigma^2 T}} \right). \quad (24)$$

Comparison of (24) with the results of simulations is made in figure 2 and the agreement is excellent. Simulations carried out for different values of the drift confirm the fact that $\langle A \rangle_T$ is independent of the drift, despite the fact that the expected values $\langle T \rangle$ and $\langle A \rangle$ taken in isolation are strongly dependent on the drift.

The typical spread (variance) of the data is also of interest. To that end, we can evaluate $\langle A^2 \rangle_T$ from (20) and (13) by differentiating twice with respect to s ;

$$\int_0^\infty e^{-pT} \langle A^2 \rangle_T P_0(T, x_0) dT = e^{-(2p)^{1/2} x_0 / \sigma} \times \left[\frac{5x_0 \sigma}{4(2p)^{5/2}} + \frac{5x_0^2}{4(2p)^2} + \frac{5x_0^3}{6\sigma(2p)^{3/2}} + \frac{x_0^4}{4\sigma^2(2p)} \right] \quad (25)$$

followed again by carrying out the Laplace inversion using standard results [15; Sect. 29.3.86 and Sect. 7.2.5];

$$\langle A^2 \rangle_T = \frac{5}{12} \sigma^2 T^3 + \frac{5}{12} x_0^2 T^2 - \frac{1}{12} \sqrt{\frac{2\pi}{\sigma^2}} x_0^3 T^{3/2} e^{x_0^2 / 2T\sigma^2} \operatorname{erfc}\left(\frac{x_0}{\sqrt{2\sigma^2 T}}\right). \quad (26)$$

The calculated variance $\langle A^2 \rangle_T - \langle A \rangle_T^2$ agrees well with the simulated data.

As regards limiting behaviour, both (24) and (26) reproduce the known results for Brownian excursions in the limit $x_0 \rightarrow 0$ [10, 11];

$$\lim_{x_0 \rightarrow 0} \langle A \rangle_T = \sqrt{\frac{\pi}{8}} \sigma T^{3/2}; \quad \lim_{x_0 \rightarrow 0} \langle A^2 \rangle_T = \frac{5}{12} \sigma^2 T^3. \quad (27)$$

Conversely, by considering the limiting behaviour of, say, (24) as $T \rightarrow \infty$,

$$\langle A \rangle_T = \sqrt{\frac{\pi}{8}} \sigma T^{3/2} \left[1 + \frac{x_0^2}{2\sigma^2 T} + O\left(\frac{1}{T^{3/2}}\right) \right] \quad (28)$$

one sees that the initial condition becomes unimportant once $T \gg x_0^2 / 2\sigma^2$ and in this regime the Brownian excursion results (27) are also recovered asymptotically, i.e. $\langle A \rangle_T \sim \sqrt{\pi/8}\sigma T^{3/2}$ as $T \rightarrow \infty$. This confirms the heuristic analysis presented above based on studying the correlation function $R_\eta(A, T^k)$ in the zero-drift limit.

We now change track. Following the discussion in the Introduction, in figure 4 the results of numerical simulations are shown which highlight the correlation between the time average of a *given* path till its first passage time, $\bar{x} \equiv AT^{-1}$, and the maximum value attained by the same path, x_M . It is clear that the correlation is very strong for these parameter values, which correspond to a situation where the drift is weak. To emphasize the point, in figure 5 we plot the correlation coefficient $R_\eta(AT^{-1}, x_M)$ as a function of the dimensionless parameter λ as defined in (16). A direct theoretical calculation of $R_\eta(AT^{-1}, x_M)$ is beyond the scope of the present work, but key insights may be obtained by analysing the first moments. Thus from (12) one can derive the important result,

$$\langle AT^{-1} \rangle \equiv \int_0^\infty \langle Ae^{-pT} \rangle dp = \frac{x_0}{2} \left[1 + \int_0^\infty \frac{e^{-u}}{u + \lambda} du \right] \quad (29)$$

which is the expected value of the time average $\bar{x}$. This depends upon the drift through the parameter λ and is plotted as such in figure 6, together with the results of simulations. The agreement is excellent. The dependence of (29) on the exponential integral has interesting implications which become apparent when considering the limiting behaviour;

$$\langle AT^{-1} \rangle = \frac{x_0}{2} \times \begin{cases} \log\left(\frac{1}{\lambda}\right) + 1 - \gamma + O(\lambda \log \lambda); & \lambda \rightarrow 0 \\ 1 + O\left(\frac{1}{\lambda}\right); & \lambda \rightarrow \infty \end{cases} \quad (30)$$

where $\gamma = 0.57721\dots$ is Euler's constant. The zero-drift limit of $\langle AT^{-1} \rangle$ is only weakly (logarithmically) divergent, unlike the ratio $\langle A \rangle / \langle T \rangle$ which is power law divergent. This is due to a cancelling effect resulting from the positive correlation between A and T . We wish to compare this with the behaviour of $\langle x_M \rangle$, for which it is known that [1],

$$\langle x_M \rangle = x_0 + x_0 \left(\frac{e^{2\lambda} - 1}{2\lambda} \right) \log\left(\frac{1}{1 - e^{-2\lambda}} \right) \quad (31)$$

which is also shown, together with the results of simulations, in figure 6. This quantity has the limiting behaviour,

$$\langle x_M \rangle = x_0 \times \begin{cases} \log\left(\frac{1}{\lambda}\right) + 1 - \log 2 + O(\lambda \log \lambda); & \lambda \rightarrow 0 \\ 1 + O\left(\frac{1}{\lambda}\right); & \lambda \rightarrow \infty. \end{cases} \quad (32)$$

It is clear from figure 6 that $2\langle AT^{-1} \rangle \approx \langle x_M \rangle$, and by comparing (30) and (32) one sees that $\lim_{\eta \rightarrow 0} \langle 2AT^{-1} - x_M \rangle / x_0 = c$, where $c \equiv \log 2 - \gamma = 0.11593\dots$ is a universal constant. The estimator $AT^{-1} \triangleq (x_M + x_0 c) / 2$ is plotted in figure 4 and works well in the sense of providing a best-fit; a simple geometric interpretation is that paths of long duration are typically ‘triangular’ in shape, i.e. $A \approx x_M T / 2$ (see figure 1). It should be noted that for Brownian excursions the relationship $\langle A \rangle_T T^{-1} = \langle x_M \rangle_T / 2$ holds exactly, where $\langle x_M \rangle_T = \sqrt{\pi / 2} \sigma T^{1/2}$ is the expected value of the maximum [20].

Two further observations are worth making. First, given that the zero-drift limit is dominated by paths of long duration, the divergent behaviour of (30) confirms that $\bar{x} \equiv AT^{-1}$ is typically large for such paths. By way of contrast, the quantity

$$\langle AT^{-2} \rangle \equiv \int_0^\infty \int_p^\infty \langle A e^{-p'T} \rangle dp' dp = \frac{3\sigma^2}{4x_0} \left[1 + \frac{2\lambda}{3} + \frac{\lambda}{3} \int_0^\infty \frac{u e^{-u}}{u + \lambda} du \right] \quad (33)$$

is finite in the zero-drift limit, i.e. $\lim_{\eta \rightarrow 0} \langle AT^{-2} \rangle = 3\sigma^2 / 4x_0$. Thus, for paths of long duration, instances where AT^{-2} is large are unlikely, commensurate with the fact that $\langle A \rangle_T$ typically scales as $T^{3/2}$ as $T \rightarrow \infty$. Second, from (13), after considerable algebraic manipulation and simplification, one can derive,

$$\langle A^2 T^{-2} \rangle \equiv \int_0^\infty \int_p^\infty \langle A^2 e^{-p'T} \rangle dp' dp = \frac{x_0^2}{4} \left[\frac{5}{3\lambda} + 1 + \frac{2\lambda^2}{3} \int_0^\infty \frac{e^{-u}}{(u + \lambda)^3} du \right]. \quad (34)$$

In conjunction with (29) the variance of $\bar{x} \equiv AT^{-1}$ may thus be written down and in the zero-drift limit one has,

$$\text{Var}(AT^{-1}) = \frac{5x_0\sigma^2}{12\eta} \left[1 + O(\eta(\log \eta)^2) \right]. \quad (35)$$

Once more the divergence is relatively weak as a result of the positive correlation between A and T . As a check, we have compared (33) and (34) against the results of simulations and found them to be in good agreement.

4. Discussion

To briefly illustrate the application of the above results we consider two examples, namely compact directed percolation [7] and a Markovian queueing model of a traffic jam [9]. It is easily demonstrated that the Brownian motion approach is justified in a limiting sense in each case and we do not discuss that aspect any further here.

Regarding compact directed percolation, one can associate A with the cluster size and T with the cluster duration (or length). The zero-drift case $\eta = 0$ corresponds to the percolation threshold, whereupon (17) and (18) provide a novel characterisation of the critical clusters. The variable $\bar{x}$ has a natural interpretation as the average spatial width of a given cluster; in this context the first (29) and second (34) moments are of interest just below the percolation threshold, whilst the close correlation with the maximum width x_M is of particular note, suggesting as it does that the latter will typically be about twice the former for a given critical cluster.

Regarding the traffic jam model, one can associate A with the cumulative waiting time of all the vehicles involved in the jam at some time or other and T with the jam duration. The zero-drift case corresponds to the situation where the average rate of exit of vehicles from the front of the jam is precisely equal to the average rate of arrival of vehicles at the rear of the jam. For an initial jam length x_0 , one can interpret (24) as the expected cumulative waiting time (a cost variable) given that the jam duration is T . Independence of the drift parameter means the result is not influenced by, say, the average rate of arrival of vehicles at the rear of the jam; a counter-intuitive observation at first sight. The variable $\bar{x}$ corresponds to the time averaged number of cars in a given traffic jam over its lifetime and the previous observations apply equally with respect to how this quantity relates to the maximum instantaneous number of cars in the jam during its lifetime x_M . Given that roads have finite capacity, from an operational perspective it would be of interest to explore this further (the same applies to certain digital communication channels which have a finite buffer). Ideally this would entail a direct calculation of $R_\eta(AT^{-1}, x_M)$, as shown in figure 5, but even in the limit $\eta \rightarrow 0$ this remains an open challenge.

References

- [1] Kearney M J and Majumdar S N 2005 *J. Phys. A: Math. Gen.* **38** pp. 4097-4104
- [2] Kearney M J, Majumdar S N and Martin R J 2007 *J. Phys. A: Math. Theor.* **40** pp. F863-F869

- [3] Watkins N W, Credington D, Sanchez R, Rosenberg S J and Chapman S C
2009 *Phys. Rev E* **79** 041124
- [4] Reynolds A M 2010 *J. Roy. Soc. Interface* **7** pp. 1753-1758
- [5] Bandyopadhyay M, Gupta S and Segal D 2011 *Phys. Rev. E* **83** 031905
- [6] Lavrentovich M O, Korolev K S and Nelson D R 2013 *Phys. Rev. E* **87**
102103
- [7] Kearney M J 2004 *J. Phys. A: Math. Gen.* **37** pp. 8421-8431
- [8] Stapleton M A and Christensen K 2006 *J. Phys. A: Math. Gen.* **39** pp. 9107–
9126
- [9] Kearney M J 2006 *Phys. Rev. E* **74** 061115
- [10] Majumdar S N and Comtet A 2005 *J. Stat. Phys.* **119** pp. 777-826
- [11] Janson S 2007 *Probab. Surv.* **4** pp. 80-145
- [12] Richard C 2004 *J. Phys. A: Math. Gen.* **37** pp. 4493-4500
- [13] Majumdar S N and Comtet A 2004 *Phys. Rev. Lett.* **92** 225501
- [14] Majumdar S N 2005 *Current Science* **89** pp. 2076-2092; arXiv:cond-
mat/0510064
- [15] Abramowitz M and Stegun I A 1972 *Handbook of mathematical functions*
(New York: Dover)
- [16] Revuz D and Yor M 1999 *Continuous Martingales and Brownian Motion*, 3rd
ed. (Berlin: Springer)
- [17] Redner S 2001 *A Guide to First-Passage Processes* (Cambridge: Cambridge
University Press).
- [18] Darling D A 1983 *Ann. Probab.* **11** pp. 803-806
- [19] Louchard G 1984 *J. Appl. Prob.* **21** pp. 479-499

- [20] Majumdar S N, Randon-Furling J, Kearney M J and Yor M 2008 *J. Phys. A: Math. Theor.* **41** 365005

Figure 1. An illustration of the first passage process. The lower dashed line indicates the time average value $\bar{x} \equiv AT^{-1}$ over the interval $[0, T]$; the upper dashed line indicates the maximum value x_M .

Figure 2. Simulations (triangles) showing the correlation between A and T , with $x_0 = 10$, $\sigma = 1$, and $\eta = 0.01$. The solid line is a theoretical calculation of $\langle A \rangle_T$; also shown are the corresponding numerical averages (diamonds).

Figure 3. Results based on simulations for correlation coefficient $R_\eta(A, T^k)$ as a function of $\lambda \equiv \eta x_0 / \sigma^2$ for $k=1$ (squares), $k=3/2$ (diamonds) and $k=2$ (triangles), with $x_0=10$ and $\sigma=1$. The two solid lines are theoretical calculations.

Figure 4. Simulations (triangles) showing the correlation between AT^{-1} and x_M , with $x_0 = 10$, $\sigma = 1$, and $\eta = 0.01$. The dotted line corresponds to the estimator discussed in the main text.

Figure 5. Results based on simulations (squares) for the correlation coefficient $R_\eta(AT^{-1}, x_M)$ as a function of $\lambda \equiv \eta x_0 / \sigma^2$ with $x_0 = 10$ and $\sigma = 1$. The dotted line is a guide to the eye.

Figure 6. Results based on simulations for $\langle x_M \rangle$ (squares), $\langle AT^{-1} \rangle$ (triangles) and $\langle 2AT^{-1} - x_M \rangle$ (diamonds) as a function of $\lambda \equiv \eta x_0 / \sigma^2$ with $x_0 = 10$ and $\sigma = 1$. The solid lines are theoretical calculations.