ORIGINAL ARTICLE

Success probability for selectively neutral invading species in the line model with a random fitness landscape

Suzan Farhang-Sardroodi[1] | Natalia L. Komarova[2] |
Marcus Michelen[3] | Robin Pemantle[4]

[1] Department of Mathematics, Ryerson University, Toronto, Ontario, Canada

[2] Department of Mathematics, University of California Irvine, Irvine, California, USA

[3] Department of Mathematics, Statistics and Computer Science, University of Illinois at Chicago, Chicago, Illinois, USA

[4] Department of Mathematics, University of Pennsylvania, Philadelphia, Pennsylvania, USA

Correspondence
Marcus Michelen, Department of Mathematics, Statistics and Computer Science, University of Illinois at Chicago, 851 S. Morgan Street Chicago, IL 60607-7045.
Email: michelen.math@gmail.com

Funding information
NSF, Grant/Award Numbers: DMS-1812601, DMS-1612674

Abstract

We consider a spatial (line) model for invasion of a population by a single mutant with a stochastically selectively neutral fitness landscape, independent from the fitness landscape for nonmutants. This model is similar to those considered earlier. We show that the probability of mutant fixation in a population of size N, starting from a single mutant, is greater than $1/N$, which would be the case if there were no variation in fitness whatsoever. In the small variation regime, we recover precise asymptotics for the success probability of the mutant. This demonstrates that the introduction of randomness provides an advantage to minority mutations in this model, and shows that the advantage increases with the system size. We further demonstrate that the mutants have an advantage in this setting only because they are better at exploiting unusually favorable environments when they arise, and not because they are any better at exploiting pockets of favorability in an environment that is selectively neutral overall.

KEYWORDS
birth-death process, random environment, RWRE

1 | INTRODUCTION

Evolution in random environments has attracted attention of ecologists and mathematical biologists for a long time. Consider direct competition dynamics between two types of organisms whose reproduction and death rates may be different in different spatial locations. It is clear that organisms with larger reproduction rates and lower death rates are more likely to rise from low numbers and eventually replace their slow reproducing, rapidly dying counterparts. The situation becomes more complicated if the environment consists of different patches, where different types enjoy evolutionary advantage while others are suppressed. Depending on the properties of this patchy environment, the reproduction and death rates of the organisms, and the details of the evolutionary process, a number of outcomes can be observed, see, e.g., Refs. 3–6 and also Modern Coexistence Theory.[7]

From early works of Haldane,[8] Fisher,[9] and Wright[10] almost 100 years ago, an important focus of many theoretical studies of evolution has been the probability and timing of mutant fixation, see also Kimura's studies of neutral evolution.[11,12] The general setting assumes the coexistence of different variants of an organism in a population, one of which is referred to as the "wild type" (or "normal") and the other(s) as "mutants" (or variants). Mutations may or may not confer selective advantage or disadvantage to an organism. In general, the term "neutral' in evolutionary theory refers to the type of variants that, although different from the wild type, is neither advantageous nor disadvantageous, that is, it does not experience a positive or negative selection pressure.

Mutant evolution in *random environments* became a topic of mathematical investigation around 1960s. Many early papers studied temporal fluctuations of the environment. For example, in Ref. 13, it was assumed that while the wild types had constant numbers of offspring, mutants' numbers of offspring were randomly changing every time step (but had the same mean as the wild types' offspring numbers). It was found that despite having the same mean number of offspring, the mutants behaved as if they were disadvantageous. References 14,15 studied a more general setting, where the division rates of both wild types and mutants were affected by the environmental changes. It was found that, surprisingly, the mutants behaved as if they were advantageous, despite having the same mean division rate, but only if the mutants were initially a minority. A similar result was found by Refs. 16, 17. Many results have been obtained in the framework of the Modern Coexistence Theory in ecology, e.g., regarding the instantaneous rate of increase of a rare species.[18–20] It was shown analytically by Refs. 3, 21, 22 that temporal randomness in division rates leads to a positive rate of increase of a minority mutant. Another set of analytical results concerns extinction times.[23–25]

In contrast to temporal variations, spatial environmental variations are associated with fitness differences that characterize different spatial locations (and do not change in time). For example, one can consider a stylized model where light conditions differ in different locations, and therefore growth and reproduction properties of plants may differ spot to spot. Let us suppose that the wild type plant needs high light to grow, but a mutant prefers shade. Then spots characterized by strong lighting conditions will result in an increase in wild type growth rate and a decrease in mutant growth rate. What can we say about the mutant fixation probability if the "high light" and "low light" spots are distributed with equal likelihood? In this example, the fitness values of wild type and mutant organisms are anticorrelated, that is, in a given spot, if a wild type plant has an elevated fitness value, a mutant will have a reduced fitness value. Different scenarios are possible, including the case where fitness values of wild type and mutant organisms are uncorrelated; this would correspond to a situation where the growth properties of wild type and mutant plants are determined by different and uncorrelated environmental factors, such as light and nutrients.

Two important examples of biological systems where evolution takes place in the presence of spatial randomness, are biofilms and tumors. Biofilms are collectives of microorganisms, such as bacteria or fungi, that coexist on surfaces within a slimy extracellular matrix. Evolutionary dynamics of these microorganisms take place in an environment characterized by significant heterogeneities, both in physical and chemical parameters, such as heterogeneities in the interstitial fluid velocity, gradients in the distribution of nutrients, and other metabolic substrates/products.[26,27] It has been suggested[28] that different organisms may respond differently to these diverse environmental stimuli, giving rise to evolutionary codynamics that can be modeled by using models similar to those studied here. The second example is evolution in cancerous populations, where the presence of highly heterogeneous environments has been documented, see, e.g., Refs. 29, 30. Cancerous cells in different locations across a tumor are exposed to different concentrations of oxygen, nutrients, immune signaling molecules, inflammatory mediators, and other nonmalignant cells that comprise the tumor microenvironment. Understanding tumor evolution under these spatially heterogeneous conditions is essential for understanding and combating long-standing challenges in oncology such as drug resistance in tumors. It also presents opportunities for creating new therapeutic strategies.[31]

In the literature, several modeling approaches have been used to study spatial randomness. In one class of models, agents are placed on a random network, where different vertices have different degrees; the nodes' fitness values are based on their numbers of interactions, making some vertices more advantageous than others. These types of settings have been used, e.g., in the context of the game theory/cooperation (e.g., Refs. 32–37). Another class of models is a finite island model, where agents are placed in patches (characterized by environmental differences) and a certain degree of patch-to-patch migration is assumed. Mutant fixation probability has been studied in the high migration rate[38] and the low migration rate[39] limit. Mutant fixation probability in the problem with two patches has been solved analytically in Ref. 40, where it was assumed that the mutation is advantageous in one patch and deleterious in the other patch. An extension to a multiple patch model was provided in Ref. 41, who investigated the accuracy of various approximations for mutant fixation probability. The role of spatially variable environments has been also addressed by the Modern Coexistence Theory, see, e.g., studies of species coexistence in Ref. 42.

In the recent papers[1,2] we studied the dynamics of mutant fixation in a model that is a generalization of the classical Moran model[43] and includes spatial randomness. We assumed that the population of organisms (or agents) remains constant and birth/death updates are performed with rules governed by the organisms' fitness parameters (birth and/or death rates). Interactions of replacing dead organisms by offspring of others happen along edges of a network that defines "neighborhoods." For example, in a model characterized by agents on a complete graph, every agent is in the neighborhood of everyone else, and therefore a dead organism can be replaced by offspring of any other agent. On the other hand, on a circular graph, each agent has exactly two neighbors. It was assumed that, for each realization of the evolutionary competition process, for each of the N sites, the birth and/or death rates of both types were assigned by randomly drawing the same distributions of values. Then, the probability of mutant fixation, starting a given initial location of mutant agents among the N spots, was calculated. Finally, this probability was averaged over all realizations of the fitness values. It was found that, somewhat surprisingly, the mutants showed an advantage compared to the normal types, as long as their initial number was smaller than a half. This result can be obtained for particular (relatively small) numbers of N, but no asymptotic results for large values of N were obtained analytically. It was observed, however, that the effect of randomness to "favor" minority mutant increased with the system size.

In this paper, we focus on the asymptotic behavior of the fixation probability of mutants in the presence of spatial randomness. We consider a spatial model similar to that used in Refs. 1, 2. It is a spatial (1D) version of the Moran process (see, e.g., Ref. 44) where spatial variations in the environment are implemented by random fitness values of wild type and mutant individuals at different sites. Models of this type (but without random fitness values) have been used previously to study cellular evolution in the context of cancerous transformation ([44,45]) and are relevant for describing, e.g., colonic crypts. To the best of our knowledge the results reported here are the first rigorous results for the problems of this kind.

2 | MODEL FORMULATION AND RESULTS

We consider the following model. The spatial environment consists of N sites, numbered $1, \dots, N$ arranged in a line with nearest neighbor edges. At each site there are two real parameters representing *fitness* values: a mutant fitness and a normal fitness, each chosen i.i.d. $1 \pm \delta$. These fitness values will remain fixed while the *state* of each site will change. Site 1 begins with state "mutant" and all other sites begin with the state "normal." The evolution proceeds in discrete time as follows: Replace each edge with two directed edges, one in each direction; at each time-step choose a directed edge (j, k) with $|j - k| = 1$ uniformly at random, and let ν_k and μ_k be the normal and mutant fitnesses of k; if j is mutant, then we set k to be mutant with probability $\mu_k / (\nu_k + \mu_k)$ and leave k unchanged with the remaining probability; similarly, if j is normal then we set k to be normal with probability $\nu_k / (\nu_k + \mu_k)$ and leave it unchanged otherwise.

This model may also be thought of as occurring in continuous time: Each directed edge is assigned an exponential clock of rate 1. When the clock edge u to v rings, u attempts to replace the type of v with its own type; if u is mutant, then the state of v is set to be mutant with $\frac{\mu_v}{\mu_v + \nu_v}$ and is unchanged with the remaining probability.

Because there are only finitely many sites and only two types, the process eventually fixates in one of two states: all mutants or all normal. We are interested in the probability of the event G of fixating in the state where all sites are mutants, and in particular how the probability that G occurs changes—after averaging over the random environment—as δ varies. More concretely, should more or less randomness help the mutant dominate?

If $\delta = 0$, there is no differential fitness and the fitness environment is deterministic. After k replacements, the mutants will always either be extinct or occupy some interval $1, \dots, X_k$. The process $\{X_k\}$ is a simple random walk stopped when it hits 0 or N, hence the probability that it stops at N is precisely $1/N$. Biologically this means that in the absence of any fitness differences between the wild type and mutant cells, the probability of any cell to fixate is the same and equals $1/N$. Note that if fixation probability is greater (smaller) than the initial share of the mutant, then this indicates the presence of positive (negative) selection acting in the system.

In the model considered here, when $\delta > 0$, the dynamics become more complicated. In fact they are the dynamics of a birth–death process in a random environment; equivalently, the dynamics may be thought of as a variant of the voter model where each site may be more or less susceptible to a given type. A similar model, but with circular boundary conditions (1 and N are neighbors), was analyzed in Ref. 1. There, it was proved for $4 \le N \le 8$ and empirically observed for much larger values of N that the probability of a mutant takeover is strictly greater than $1/N$, indicating the presence of positive selection for the mutant, although its fitness values are chosen the same distribution as those for the wild type cells. The goal of this paper is to establish the analogous
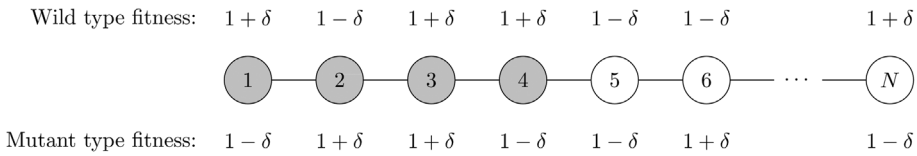
FIGURE 1 An instance of the model with mutant sites in gray and normal sites in white and both mutant and wild fitness types listed. The above state is identified with 4, because the mutant sites are $\{1, 2, 3, 4\}$

result rigorously for the line model and to give precise asymptotics for the annealed probability of a mutant takeover.

Let $(\Omega_N, \mathcal{F}_N, \mathbb{P}_N)$ be a probability space on which are defined independent Rademacher random variables (that is, ± 1 fair coin flips) $B_1, \dots, B_N$ and $B'_1, \dots, B'_N$, as well as rate 1 Poisson processes $\xi_t^{(i,j)}$ for $1 \leq i, j \leq N$ and $|i - j| = 1$, independent of the Rademacher variables and of each other. For $\delta \in (0, 1)$, the normal fitness at site k is the quantity $\mu_k := 1 + \delta B_k$ and the mutant fitness at site k is the quantity $\nu_k := 1 + \delta B'_k$; in this way, the model is defined simultaneously for all δ, although we will not do much to exploit this simultaneous coupling.

The states of the process are configurations where each site has a mutant (one) or normal cell (zero). Because we always consider the starting condition of having one mutant at site 1 and all others are normal, the collection of mutant cells is always some segment of sites $[1, k]$ and normal cells thereafter. Hence we can identify the state space with $\{0, 1, \dots, N\}$, with 0 corresponding to mutant extinction. Because we need only keep track of the right-most mutant to describe the state of the process, we first find the transition probabilities for the evolution of this right-most point. Figure 1 shows an instance of the model along with this identification.

At times corresponding to points of the Poisson process $\xi^{(i,j)}$, cell i tries to reproduce at site j. This only matters if $i = k$ or $j = k$, because otherwise sites i and j have the same state and no change in state can occur. Sampling only when the configuration changes yields a discrete time birth and death chain, absorbed at 0 and N, whose transition probabilities are easily characterized. Define the random quantities

$$\beta_k := \frac{\mu_k}{\nu_k + \mu_k}. \tag{1}$$

From state k the only relevant directed edges are $(k, k + 1)$ and $(k + 1, k)$ because these correspond to the mutant site k making $k + 1$ mutant and normal site $k + 1$ making k normal. Both attempted at rate 1 and succeeding with respective probabilities β_{k+1} and $1 - \beta_k$. Letting p_k denote the transition probability the right-most mutant being k to being $k + 1$, we have

$$p_k = \frac{\beta_{k+1}}{\beta_{k+1} + (1 - \beta_k)}. \tag{2}$$

We may now think of the evolution as occurring entirely on $\{0, 1, \dots, N\}$, where state k moves to step $k + 1$ with probability p_k and moves to $k - 1$ with probability $1 - p_k$.

Let $G = G(\delta)$ denote the event that the absorbing state N is reached before the absorbing state 0, under dynamics for the given δ. Our first result is an asymptotic expression for $\mathbb{P}_N(G(\delta))$ in the regime where $N \to \infty$ and $\delta\sqrt{N} \to c$.

Theorem 1 (Asymptotics when $\delta\sqrt{N} \to c$). *Fix $c > 0$ and suppose $N \to \infty$ and $\delta\sqrt{N} \to c$. Then,*

$$N\mathbb{P}_N(G(\delta)) \to g(c), \tag{3}$$

where

$$g(c) = \mathbb{E}\left[\frac{1}{\int_0^1 \exp(\sqrt{2c}B_s)\,ds}\right]$$

for a standard Brownian motion $\{B_s\}$. The function g is continuous and strictly increasing on $(0, \infty)$. It satisfies

$$g(c) - 1 \sim \frac{c^2}{6} \qquad as\ c \downarrow 0\,; \tag{4}$$

$$g(c) \sim \frac{c}{\sqrt{\pi}} \qquad as\ c \to \infty\,. \tag{5}$$

We note that continuity of g implies that for $\delta \ll N^{-1/2}$, then $\mathbb{P}_N(G(\delta)) \sim 1/N$ as in the $\delta = 0$ case. In the regime where $\delta \gg N^{-1/2}$ but still $\delta \ll (\log N)^{-\varepsilon}$, the asymptotic behavior of $\mathbb{P}_N(G(\delta))$ is as follows.

Theorem 2. *Assuming $\delta\sqrt{N} \to \infty$, suppose that there is an $\varepsilon > 0$ such that $\delta(\log N)^\varepsilon \to 0$. Then*

$$\mathbb{P}_N(G(\delta)) \sim \frac{\delta}{\sqrt{\pi N}}\,.$$

We do not expect this to hold if $\delta = \Theta(1)$ as $N \to \infty$ because without scaling, the graininess of the random walk may lead to a different constant than would be obtained by a Brownian approximation. Nevertheless, we believe the condition $\delta = o(\log N)^{-\varepsilon}$ to be unnecessary and we conjecture the following.

Conjecture 1. *If $\delta\sqrt{N} \to \infty$ and $\delta \to 0$ then*

$$\frac{\mathbb{P}_N(G(\delta))}{\delta N^{-1/2}} \to \frac{1}{\sqrt{\pi}}\,.$$

In any case, as $N \to \infty$ and in the absence of the requirement $\delta \to 0$,

$$0 < \frac{C_1\delta}{N^{1/2}} \le \mathbb{P}_N(G(\delta)) \le \frac{C_2\delta}{N^{1/2}}\,.$$

We interpret Theorems 1 and 2 as saying that the stochastic environment favors a minority invader. Indeed, in the absence of any randomness (that is, $\delta = 0$), the probability of neutral mutant fixation on a circle is given by $1/N$ (a result that can be demonstrated, e.g., by simple

symmetry considerations). Mutant fixation on a line model similar (but not identical) to the present one was studied by Ref. 44 and it was shown that it depends on the initial location of the mutant. It is the smallest for a mutant originally located at one of the ends of a line and increases toward the middle initial location, but never exceeds the value $1/N$. In the present model, in the absence of randomness, mutant fixation probability is given by $1/N$. Theorems 1 and 2 state that mutant fixation probability in the presence of randomness is greater than $1/N$, and that the quantity $N\mathbb{P}_N(G(\delta))$ increases with the system size (N) and with the amount of randomness (δ). In other words, despite having no explicit advantage, a mutant in the random environment gets fixated with a probability that is significantly larger than in the case of a nonrandom environment.

The following result shows that this effect is due to the minority taking advantage of the cases where the overall environment is more favorable, not environments where pockets favoring each type appear but are balanced against each other.

Theorem 3. *Let $N = 2k$ be an even integer and let Q_N denote $\mathbb{P}_N$ conditioned on $\sum_j B_j = \sum_j B'_j$. Then $NQ_N(G(\delta)) = 1$ for all N and all δ.*

To rephrase in biological terms, we note that among different realizations of wild type and mutant fitness values, there are cases where mutants experience an overall advantage ($\sum_j B_j < \sum_j B'_j$), an overall disadvantage ($\sum_j B_j > \sum_j B'_j$), or have a fitness configuration whose net sum is equal to that of the wild types, although locally mutants may experience positive or negative selection pressure (the case $\sum_j B_j = \sum_j B'_j$). Theorem 3 states that if we only consider the latter type of environments, mutants will behave exactly as expected in the absence of randomness. On the other hand, configurations with a net mutant advantage and disadvantage do not balance each other out and result in a positive selection pressure experienced by the mutant.

The outline of the remainder of the paper is as follows. In the next section, we show how the computation of $\mathbb{P}_N(G(\delta))$ reduces to computing an expectation of a functional of a random walk. From here, Theorems 1 and 2 can heuristically be inferred replacing the random walk with a corresponding Brownian motion via Donsker's Theorem. However, the only regime in which Donsker's Theorem applies is that of Theorem 1. Using this approach, we then verify in the case $\delta \sim cN^{-1/2}$ that the expectation commutes with the Brownian scaling limit. Section 4 computes the corresponding expectations for Brownian motion, based on results of Matsumoto, Yor, and others. Section 5 puts this together to prove Theorem 1. We also give the relatively brief proof of Theorem 3. Theorem 2 is proved in Section 6. This is proved in two stages, first when δ is required to decrease more rapidly than $(\log N)^{-1}$ and then when this is relaxed to $(\log N)^{-\varepsilon}$. The final section presents some numerical simulations and further questions.

3 | A SCALING RESULT

The following explicit formula for the probability of a birth and death process started at 1 to reach N before 0 is well known; we include its short proof for completeness.

Proposition 1. *In a birth and death process, let p_k be the probability of transition to $k + 1$ from k and let $q_k := 1 - p_k$ be the probability of transition to $k - 1$. Let Q_x denote the law of the process*

starting x and τ_a the hitting time at state a. Then

$$Q_1(\tau_N < \tau_0) = \frac{1}{\sum_{k=0}^{N-1} \prod_{j=1}^{k} \frac{q_j}{p_j}}. \tag{6}$$

Here, the first term of the sum is the empty product, equal to 1 by convention.

Proof. Consider the network in which the resistance between k and $k+1$ is $\prod_{j=1}^{k}(q_j/p_j)$. Note then that the random walk on this network is equivalent to that described in the proposition. The expression for Q_1 is the ratio of the conductance 1 to N to that plus the conductance 1 to 0. ■

We now show that the denominator is close to a functional of a random walk, which is close to a functional of a Brownian motion, and that these approximations are good enough to pass expectations to the limit.

For $1 \le k \le N-1$, denote

$$X_k := \log \frac{q_k}{p_k} = \log \frac{1-\beta_k}{\beta_{k+1}},$$

$$\widetilde{X}_k := \log \frac{1-\beta_{k+1}}{\beta_{k+1}} = \log \frac{\nu_{k+1}}{\mu_{k+1}},$$

with partial sums $S_k := \sum_{j=1}^{k} X_j$ and likewise for $\widetilde{S}$. Define $S_0 = \widetilde{S}_0 = 0$. The definition of X_k is chosen so that Equation (6) becomes

$$\mathbb{P}_N(G(\delta)) = \mathbb{E} \frac{1}{\sum_{k=0}^{N-1} \exp(S_k)}. \tag{7}$$

On the other hand $\widetilde{X}_k$ are chosen so that $\{\widetilde{S}_k\}$ is precisely a simple random walk on the lattice $\delta'\mathbb{Z}$, with holding probability $1/2$, where

$$\delta' := \log \frac{1+\delta}{1-\delta} = 2\delta + O(\delta^2)$$

For $\delta \le 1 - \varepsilon$, μ_j and ν_j are uniformly bounded away 0 and 1 and so β_j is bounded away 1 as well. Thus,

$$|\widetilde{S}_k - S_k| = |\log(1-\beta_{k+1}) - \log(1-\beta_1)| \le C_\varepsilon \delta \tag{8}$$

as long as $\delta \le 1 - \varepsilon$ by applying Taylor's theorem with remainder to $\log(1-\beta_{k+1})$ and $\log(1-\beta_1)$ as a function of δ. Donsker's theorem then gives

$$(S_{\lfloor tN \rfloor})_{t \in (0,1)} \xrightarrow{N \to \infty} \sqrt{2} \cdot c \cdot (B(t))_{t \in (0,1)} \tag{9}$$

in the càdlàg topology whenever $\delta\sqrt{N} \to c$, where $B(t)$ is Brownian motion.

In a moment we will show the following:

Lemma 1. *Suppose δ and N vary so that $\delta\sqrt{N}$ remains bounded away zero and infinity. Then the random variables $\{NQ_1(\tau_N < \tau_0)\}$ are uniformly integrable. Further, if $\delta\sqrt{N} \to c$ then*

$$N\mathbb{P}_N(G(\delta)) \to \mathbb{E}\left[\frac{1}{\int_0^1 \exp(\sqrt{2c}B_s)\,ds}\right],$$

where (B_s) is standard Brownian motion.

The second half of Lemma 1 is the first part of Theorem 1 and follows uniform integrability together with (6) and (9). This is because convergence of means follows uniform integrability together with convergence in distribution.

Proof of Lemma 1. A consequence of (8) is that

$$\left|\frac{\sum_{j=0}^{N-1} \exp(\widetilde{S}_j)}{\sum_{j=0}^{N-1} \exp(S_j)} - 1\right| = o(1).$$

It therefore suffices to show that the variables

$$\frac{N}{\sum_{j=0}^{N-1} \exp(\widetilde{S}_j)}$$

are uniformly integrable.

For a simple random walk, the reflection principle gives

$$\mathbb{P}[\min_{j \le n} S_j \le -r] = \mathbb{P}[S_n = -r] + 2\mathbb{P}[S_n < -r]$$

$$\le 2\mathbb{P}[S_n \le -r]$$

$$\le 2\exp\left(-2r^2/n\right),$$

where the last bound is by Hoeffding's inequality. Because $\{\widetilde{S}_k\}$ is a simple random walk scaled by δ' and holding with probability $1/2$, for all $k \in \mathbb{Z}^+$ and $t > 0$,

$$\mathbb{P}[\min_{j \le k} \widetilde{S}_j < -t] \le 2\exp\left(-\frac{2t^2}{\left(\log\left(\frac{1+\delta}{1-\delta}\right)\right)^2 k}\right). \tag{10}$$

Applying (10) shows that

$$\mathbb{P}[\min_{k \le N\varepsilon} \widetilde{S}_k \le -1] \le 2\exp\left(-B/\varepsilon\right) \tag{11}$$

for B depending continuously on $\delta\sqrt{N}$. Thus for $\varepsilon \in (0,1]$

$$\frac{N}{\sum_{j=0}^{N-1}\exp(\widetilde{S}_j)} \leq \frac{N}{\sum_{j=0}^{\varepsilon N}\exp(\widetilde{S}_j)} \leq \frac{1}{\varepsilon\exp\left(\min_{j\leq\varepsilon N}\widetilde{S}_j\right)}.$$

Thus, for x large enough and picking $\varepsilon = e/x$, we have

$$\mathbb{P}\left[\frac{N}{\sum_{j=0}^{N-1}\exp(\widetilde{S}_j)} > x\right] \leq \mathbb{P}\left[\frac{1}{\varepsilon\exp\left(\min_{j\leq\varepsilon N}\widetilde{S}_j\right)} > x\right] = \mathbb{P}\left[\min_{j\leq(eN/x)}\widetilde{S}_j < -1\right].$$

We have, for each $K \geq e$,

$$\mathbb{E}\left[\frac{N}{\sum_{j=0}^{N-1}\exp(\widetilde{S}_j)}\mathbf{1}\left\{\frac{N}{\sum_{j=0}^{N-1}\exp(\widetilde{S}_j)} > K\right\}\right] = \int_{x\geq K}\mathbb{P}\left[\frac{N}{\sum_{j=0}^{N-1}\exp(\widetilde{S}_j)} > x\right]dx$$

$$\leq \int_{x\geq K}\mathbb{P}\left[\min_{j\leq(eN/x)}\widetilde{S}_j < -1\right]dx$$

$$\leq \int_{x\geq K}2\exp(-Cx/e)\,dx.$$

This inequality holds for all N and converges to zero as $K \to \infty$, thereby showing uniform integrability.　∎

4　│　EVALUATION OF THE BROWNIAN INTEGRAL

Define the following functions of Brownian motion:

$$A_\alpha(t) := \int_0^t e^{\alpha B_s}\,ds\,;\tag{12}$$

$$m_\alpha(t) := \mathbb{E}A_\alpha(t)^{-1}.\tag{13}$$

In this notation, Lemma 1 proves the first statement of Theorem 1 with

$$g(c) := m_{c\sqrt{2}}(1).\tag{14}$$

To finish the proof of Theorem 1, it remains to evaluate (14). Expectations such as the one in (13) have been well studied.

Proposition 2 Ref. [52]. *Let $\{B_t : t \geq 0\}$ be a standard Brownian motion and let $A(t) := \int_0^t e^{2B_s}\, ds$. Then,*

$$\mathbb{E}\left[A_2(t)^{-1}|B_t = x\right] = \frac{xe^{-x}}{t\sinh x} \quad \text{if } x \neq 0 \, ; \tag{15}$$

$$\mathbb{E}\left[A_2(t)^{-1}|B_t = 0\right] = t^{-1} \, ; \tag{16}$$

$$m_2(t) \sim \sqrt{\frac{2}{\pi t}} \quad \text{as } t \to \infty \, . \tag{17}$$

The next lemma uses Brownian scaling to transfer these results to the -1 moment of $\int_0^1 e^{\alpha B_s}\, ds$.

Lemma 2. *For $\alpha, \nu, t > 0$,*

$$m_\alpha(t) = \frac{\alpha^2}{\nu^2} m_\nu\left(\frac{\alpha^2}{\nu^2} t\right). \tag{18}$$

It follows that

$$m_\alpha(1) \sim \frac{\alpha}{\sqrt{2\pi}} \quad \text{as } \alpha \to \infty \, . \tag{19}$$

Both proofs are straightforward although somewhat technical, and so we defer them to Appendix A.

5 | PROOFS OF THEOREMS 1 AND 3

Proof of Theorem 1. We have already evaluated g. Continuity and strict monotonicity will follow computing the second derivative of $\phi(x) = xe^{-x}/\sinh(x)$ explicitly. The estimate (5) follows immediately (14) and (19). It remains to prove (4), that is, to estimate $g(c) = m_{c\sqrt{2}}(1)$ near $c = 0$. Integrating (15) gives

$$t\, m_2(t) = \int \frac{xe^{-x}}{\sinh(x)}\, dN(0, t)(x).$$

Plugging in $xe^{-x}/\sinh(x) = 1 - x + x^2/3 + O(x^3)$ gives

$$t\, m_2(t) = 1 + \frac{t}{3} + O(t^{3/2})$$

as $t \downarrow 0$. Using (18) with $\alpha = c\sqrt{2}$ and $\nu = 2$ then gives

$$g(c) = \frac{c^2}{2} m_2\left(\frac{c^2}{2}\right) = 1 + \frac{c^2 + o(1)}{6},$$

proving (4). ∎

Proof. We show that $\phi''(x) > 0$ for $x \neq 0$. Computing, $\phi''(x) = e^{-x}(\sinh(x))^{-3}h(x)$ where

$$h(x) = x + 1 + (x - 1)e^{2x}.$$

Because we have taken out a factor of the same sign as x, we need to show that h is positive on $(0, \infty)$ and negative on $(-\infty, 0)$. Verifying first that $h(0) = h'(0) = 0$, the proof is concluded by observing that $h'(x)$ has a unique minimum at $x = 0$, because $h''(x) = 4xe^{2x}$ has the same sign as x. ∎

Proof of Theorem 3. Extend the definition of X_k by reducing modulo N, thus $X_N := \log(1 - \beta_N) - \log \beta_1$ and so forth. This makes the sequence $\{X_k : k \geq 1\}$ periodic and shift invariant, that is, $(X_1, \dots, X_{N-1}, X_N) \overset{D}{=} (X_2, \dots, X_N, X_1)$. Observe also that $Q_N(S_N = 0) = 1$ because $S_N = \sum_{j=1}^{N} \log(1 - \beta_j) - \sum_{j=1}^{N} \log \beta_j$ and the multiset of N values of β_j is the same as the multiset of N values of $1 - \beta_j$. This implies that for each $k \geq 0$ we have $S_{k+N} = S_k$ and so

$$\sum_{j=0}^{N-1} \exp(S_j) = \sum_{j=0}^{N-1} \exp(S_{j+k}).$$

By shift invariance of $(X_1, \dots, X_N)$, we may shift this sequence k times to shift the sequence $(S_k, \dots, S_{k+N})$ to $(S_k - S_k, \dots, S_{k+N} - S_k)$. In particular, this shows that

$$\mathbb{E}\frac{1}{\sum_{j=0}^{N-1} \exp(S_j)} = \mathbb{E}\frac{1}{\sum_{j=0}^{N-1} \exp(S_{j+k})} = \mathbb{E}\frac{1}{\sum_{j=0}^{N-1} \exp(S_j - S_k)} = \mathbb{E}\frac{S_k}{\sum_{j=0}^{N-1} \exp(S_j)}.$$

Averaging over all $0 \leq k \leq N - 1$ shows

$$\mathbb{E}\frac{1}{\sum_{j=0}^{N-1} \exp(S_j)} = N^{-1} \sum_{k=0}^{N-1} \mathbb{E}\frac{S_k}{\sum_{j=0}^{N-1} \exp(S_j)} = N^{-1}$$

proving Theorem 3. ∎

6 | PROOF OF THEOREM 2

6.1 | KMT coupling and preliminaries

A key result for studying this larger regime δ is coupling of random walk to Brownian motion.

Lemma 3 (KMT coupling). *Let $\{X_k\}_{k \geq 0}$ be a simple random walk with i.i.d. increments ξ so that $\mathbb{E}[\xi] = 0, \mathbb{E}[\xi^2] = 1$ and $\mathbb{E}[e^{t|\xi|}] < \infty$ for t sufficiently small. Extend X_k to continuous time by*

defining $X_t = X_{\lfloor t \rfloor}$. Then there exists a constant C so that for all $T > e$ there is a coupling so that

$$\mathbb{P}\left[\sup_{t \in [0,T]} |X_t - B_t| \geq C \log(T) \right] \leq \frac{1}{T^2},$$

where B_t is standard Brownian motion.

Proof. This is stated as eq. (17) in Ref. 46. In fact one can replace T^{-2} by a stretched exponential; a slightly weaker result that may be quoted (Ref. 47 Theorem 7.1.1) is that for any $\alpha > 0$ there is a C_α such that

$$\mathbb{P}\left[\sup_{t \in [0,T]} |X_t - B_t| \geq C_\alpha \log(T) \right] \leq T^{-\alpha} . \tag{20}$$

■

Another basic lemma is the well-known uniform estimate for random walk hitting probabilities whose proof is standard and so we prove it in Appendix B.

Lemma 4. *Let $\{S_n\}_{n \geq 0}$ be a random walk whose i.i.d. increments $\{X_n\}$ satisfy the hypotheses of the KMT coupling. Then, if u ranges over $[M^\varepsilon, M^{1/2 - \varepsilon}]$ for some $\varepsilon \in (0, 1/2)$,*

$$\mathbb{P}(\max_{1 \leq j \leq M} S_j \leq u) \sim \sqrt{\frac{2}{\pi}} \frac{u}{M^{1/2}}$$

as $M \to \infty$, uniformly in u and the random walk.

We will prove Theorem 2 in two cases, both of which will make further use of the KMT coupling. Two relevant functionals of random walk will correspond to two functionals of Brownian motion, and we will need to show that the expectations in the random walk case are asymptotically equivalent to those in the Brownian motion case. For this, we require a few results to show that these functionals are sufficiently well behaved.

6.2 | Two Brownian functionals

The functionals of interest are

$$X_M := \frac{1}{\int_0^M \exp(B_s) \, ds},$$

$$Y_M := \frac{B_M^-}{\int_0^M \exp(2B_s) \, ds},$$

where we use the notation $Z^- := \max\{-Z, 0\}$.

Lemma 5. *The family of variables $\{\frac{X_M}{\mathbb{E}X_M}\}_{M \geq 1}$ is uniformly integrable.*

Proof. We first claim that $\mathbb{E}[X_M] = \Theta(M^{-1/2})$. Indeed, by Brownian scaling we have

$$\mathbb{E}X_M = \mathbb{E}\left(\int_0^M \exp(B_s)\,ds\right)^{-1} = \frac{1}{4}\mathbb{E}\left(\int_0^{M/4} \exp(2B_t)\,dt\right)^{-1} \sim \frac{1}{4}\cdot\sqrt{\frac{8}{\pi M}},$$

where the asymptotic relation is by Proposition 2.

We claim that there exists a universal $C > 0$ so that $\mathbb{P}[X_M \geq t] \leq Ct^{-2}M^{-1/2}$. For $t \leq M^{1/3}$, note that $\mathbb{P}[X_M \geq t]$ is at most the probability that both $\min_{0 \leq s \leq 1} B_s \leq -\log(t)$ and that after first hitting $-\log(t)$, B_s do not spend more than 1 unit of time above $-\log(t)$; this is because if $\min_{0 \leq s \leq 1} B_s > -\log t$ then

$$X_M \leq \frac{1}{\int_0^1 \exp(B_s)\,ds} < \frac{1}{\int_0^1 \exp(-\log t)\,ds} = t.$$

Let $\tau = \inf\{s \in [0,1] : B_s = -\log(t)\}$. The reflection principle gives

$$\mathbb{P}[\inf_{s \in [0,1]} B_s \leq -\log(t)] = 2\mathbb{P}[B_1 \leq -\log(t)] \leq C\exp\left(-\frac{1}{2}(\log(t))^2\right).$$

Conditioned on $\tau < \infty$, note that the probability B_s does not go above $-\log(t)$ for more than 1 unit of time is $\frac{2}{\pi}\arcsin(\sqrt{1/(M-\tau)}) = \Theta(M^{-1/2})$ by the strong Markov property together with Lévy's arcsin law. For all t we have

$$\mathbb{P}[X_M \geq t] \leq C\exp\left(-\frac{1}{2}(\log(t))^2\right)\frac{1}{\sqrt{M}}.$$

For all $t \geq 1$, there is some constant $C' > 0$ so that

$$\exp\left(-\frac{1}{2}(\log t)^2\right) \leq Ct^{-2}$$

and so for $t \geq 1$ we have

$$\mathbb{P}[X_m \geq t] = O(t^{-2}M^{-1/2}).$$

For K large enough, we may then bound

$$\mathbb{E}\left[\frac{X_M}{\mathbb{E}X_M}\mathbf{1}_{X_M/\mathbb{E}X_M \geq K}\right] \leq C\sqrt{M}\int_K^\infty t^{-2}M^{-1/2}\,dt \leq C/K.$$

Taking $K \to \infty$ completes the proof. ■

We find the asymptotics of the first two moments of the other functional; this calculation-heavy proof is done in Appendix B.

Lemma 6. *As $M \to \infty$ we have $\mathbb{E}Y_M \to 1$ and $\mathbb{E}Y_M^2 \sim 4\sqrt{2/\pi}\sqrt{M}$.*

Lastly, we show that the expectation of Y_M is not dominated by the contribution when Y_M is much larger than $\sqrt{M}$. Again, we defer the proof to Appendix B.

Lemma 7. *Let Y_M denote the random variable $\dfrac{B_M^-}{\int_0^M \exp(2B_s)\,ds}$. Then for all events E with $\mathbb{P}[E] \leq M^{-1/2-\varepsilon}$ for some $\varepsilon > 0$ we have $\mathbb{E}[Y_M \mathbf{1}_E] = o(1)$ as $M \to \infty$.*

6.3 | Medium-sized case: $N^{-1/2} \ll \delta \ll 1/\log N$

Lemma 8. *There exists a constant C' so that for δ in $[0, 1-\varepsilon]$ we have*

$$e^{-C'\delta \log(N)}\mathbb{E}\frac{\mathbf{1}_{E_n^c}}{N\int_0^1 \exp\left(\sqrt{N}\delta'B_t/\sqrt{2}\right)dt} - 1/N^2 \leq \mathbb{P}_N(G(\delta))$$

$$\leq e^{C'\delta \log(N)}\mathbb{E}\frac{\mathbf{1}_{E_n^c}}{N\int_0^1 \exp\left(\sqrt{N}\delta'B_t/\sqrt{2}\right)dt} + 1/N^2,$$

where E_N is an event with $\mathbb{P}(E_N) \leq N^{-2}$.

Proof. For $t \in [0, N]$, define $\widetilde{S}_t = \widetilde{S}_{\lfloor t \rfloor}$. Then note that

$$\sum_{j=0}^{N-1} \exp(\widetilde{S}_j) = \int_0^N \exp(\widetilde{S}_t)\,dt.$$

Let

$$p(N) = Q_1(\tau_N < \tau_0) = \frac{1}{\sum_{k=0}^{N-1}\exp(S_k)}$$

and recall that $\mathbb{P}_N(G(\delta)) = \mathbb{E}p(N)$. By (7) and (8), we have

$$\exp(-C_\varepsilon \delta) \leq p(N) \cdot \left(\int_0^N \exp(\widetilde{S}_t)\,dt\right) \leq \exp(C_\varepsilon \delta).$$

Note that $\widetilde{S}_k$ is a random walk whose increments have variance $\frac{(\delta')^2}{2}$. By Lemma 3, there exists a coupling so that

$$\mathbb{P}\left[\sup_{t\in[0,N]}\left|\frac{\sqrt{2}}{\delta'}\widetilde{S}_t - B_t\right| \geq C\log(N)\right] \leq \frac{1}{N^2}.$$

Letting E_N denote the event on the left-hand side, conditioned on the event E_N, we have

$$\exp\left(-C_\varepsilon\delta - \frac{C\delta'}{\sqrt{2}}\log(N)\right) \le p(N)\cdot\left(\int_0^N \exp(\delta' B_t/\sqrt{2})\,dt\right) \le \exp\left(C_\varepsilon\delta + \frac{C\delta'}{\sqrt{2}}\log(N)\right).$$

Because $\delta' \le C'_\varepsilon\delta$ for some constant C'_ε, we can find a new constant C'' so that

$$\exp\left(-C''\delta\log(N)\right) \le p(N)\cdot\left(\int_0^N \exp(\delta' B_t/\sqrt{2})\,dt\right) \le \exp\left(C''\delta\log(N)\right)$$

conditioned on E_N. Because $p(N) \le 1$, and $\mathbb{P}[E_N] \le \frac{1}{N^2}$, the lemma follows taking expectations and Brownian scaling. ∎

Proof of Theorem 2 (for medium δ). Because $\delta\log(N) \to 0$, Lemmas 3 and 5 imply

$$\mathbb{E}[p(N)] \sim \frac{1}{N} m_{\sqrt{N}\delta'/\sqrt{2}}(1),$$

where the uniform integrability guaranteed by Lemma 5 implies that we may ignore the $\mathbf{1}_{E_n^c}$ term. Applying Lemma 2 completes the proof. ∎

6.4 | Large case: $\delta = o(1/(\log N)^\varepsilon)$

Let $r > 6$ be a real parameter to be chosen later and set $T := \lceil\delta^{-r}\rceil$. Note that $\exp(\delta^{-r}) \gg \exp((\log N)^{r\varepsilon})$ grows faster than any polynomial in N once $r\varepsilon > 1$. Also, we may assume that $\delta^{-r} = o(N^s)$ for any positive s because the medium case already covers the regime $\delta \le (\log N)^{-2}$, say, and in the complement of this case, certainly any negative power of δ grows more slowly than any power of N. To handle the case at hand, we first show that the main contribution to $\mathbb{P}_N(G(\delta))$ is the first T steps of the walk. This is handled in Lemma 9. Analyzing the resulting functional of the random walk up to T will be done in a similar manner to the "medium-sized case": the KMT coupling will be employed to compare the random walk to a Brownian motion, and the corresponding functionals of Brownian motion will be the same as those appearing in Lemmas 6 and 7.

Recalling the process $\{\widetilde{S}_k\}$ in Section 3, we denote $Z := \widetilde{S}_T$ and $A := \sum_{k=0}^{T-1}\exp(\widetilde{S}_k)$. For a real number or random variable X, we use the notation $X^+ := \max(X,0)$ for the positive part of X and $X^- := \max(-X,0)$ for the negative part of X.

Lemma 9. *Let $s > \varepsilon^{-1}$ where we recall $\delta = o(1/(\log N)^\varepsilon)$. Then*

$$\mathbb{P}_N(G(\delta)) = (1+o(1))\frac{2}{\delta'\sqrt{\pi N}}\left(\mathbb{E}\frac{Z^-}{A} + O\left(\mathbb{E}\frac{\delta^{-s}}{A}\right)\right) + O(e^{-\delta^{-s}}).$$

Proof. Recall (8) that $\{\widetilde{S}_n\}$ is uniformly close to $\{S_n\}$ as $\delta \to 0$, therefore (7) implies $\mathbb{P}_N(G(\delta)) \sim \mathbb{E}(A^{-1})$ which we will use instead of (7).

We show the asymptotic equality in the statement of the theorem as two inequalities:

$$\mathbb{P}_N(G(\delta)) \geq (1 + o(1)) \frac{2}{\delta' \sqrt{\pi N}} \left(\mathbb{E} \frac{Z^-}{A} + O\left(\mathbb{E} \frac{\delta^{-s}}{A} \right) \right); \tag{21}$$

$$\mathbb{P}_N(G(\delta)) \leq (1 + o(1)) \frac{2}{\delta' \sqrt{\pi N}} \left(\mathbb{E} \frac{Z^-}{A} + O\left(\mathbb{E} \frac{\delta^{-s}}{A} \right) \right) + e^{-\delta^{-s}}. \tag{22}$$

Choose r so that $r/2 - 1 > s > \varepsilon^{-1}$. Let G' denote the event $\{\max_{T \leq m \leq N} \widetilde{S}_m \leq -\delta^{-s}\}$. We condition on A and Z.

$$\mathbb{P}_N(G(\delta) \mid A, Z) = (1 + o(1)) \mathbb{E} \left[\frac{1}{\sum_{k=0}^{N-1} \exp(\widetilde{S}_k)} \,\middle|\, A, Z \right]$$

$$\geq (1 + o(1)) \mathbb{E} \left[\frac{\mathbf{1}_{G'}}{\sum_{k=0}^{N-1} \exp(\widetilde{S}_k)} \,\middle|\, A, Z \right]$$

$$\geq \frac{1}{A + Ne^{-\delta^{-s}}} \mathbb{P}_N(G' \mid A, Z). \tag{23}$$

Because $\widetilde{S}_m/(\delta'/\sqrt{2})$ is a random walk with centered increments of variance 1, Lemma 4 gives

$$\mathbb{P}_N(G' \mid A, Z) \sim \sqrt{\frac{2}{\pi(N - T)}} \cdot \frac{(-Z - \delta^{-s})^+}{\delta'/\sqrt{2}} \sim \frac{2(Z + \delta^{-s})^-}{\delta' \sqrt{\pi N}}.$$

Combining with (23) gives

$$\mathbb{P}_N(G(\delta) \mid A, Z) \geq (1 + o(1)) \frac{2}{\delta' \sqrt{\pi N}} \cdot \frac{(Z + \delta^{-s})^-}{A + Ne^{-\delta^{-s}}}.$$

The quantity A is at least 1, while $Ne^{-\delta^{-s}} \to 0$, therefore $A + Ne^{-\delta^{-s}} \sim A$. Further, write

$$(Z + \delta^{-s})^- = -Z\mathbf{1}_{Z \leq -\delta^{-s}} - \delta^{-s}\mathbf{1}_{Z \leq -\delta^{-s}} = -Z(\mathbf{1}_{Z \leq 0} - \mathbf{1}_{Z \in [-\delta^{-s}, 0]}) - \delta^{-s}\mathbf{1}_{Z \leq -\delta^{-s}} = Z^- + O(\delta^{-s}).$$

Taking unconditional expectations now gives (21).

For the reverse inequality, let G'' denote the event that $\{\max_{T \leq m \leq N} \widetilde{S}_m \leq \delta^{-s}\}$. On the complement of this event, at least one summand in the denominator of (7) is at least $\exp(\delta^{-s})$. Because $\mathbb{P}_N(G(\delta) \mid A, Z)$ is always at most $1/A$, we see that

$$\mathbb{P}_N(G(\delta) \mid A, Z) \leq \frac{1}{A} \mathbb{P}_N(G'' \mid A, Z) + \frac{1}{A} \mathbb{P}_N((G'')^c \mid A, Z)$$

$$\leq \frac{1}{A} \mathbb{P}_N(G'' \mid A, Z) + \exp(-\delta^{-s})$$

$$\leq (1 + o(1)) \sqrt{\frac{2}{\pi(N - T)}} \frac{(\delta^{-s} - Z)^+}{A\delta'/\sqrt{2}} + e^{-\delta^{-s}}.$$

Again, taking unconditional expectations finishes, yielding (22). ∎

We are now in a position to apply the KMT coupling to find the expectations that appear in Lemma 9.

Lemma 10.

$$\mathbb{E}\frac{Z^-}{A} \sim \delta^2\,; \tag{24}$$

$$\mathbb{E}\frac{1}{A} \sim \frac{\delta}{\sqrt{\pi T}}. \tag{25}$$

The proof is similar to the results of Section 6.3, and so we defer to the Appendix once more. While the second asymptotic equality follows the result of Section 6.3, we include a proof as well because the intermediate steps required for the first expression essentially prove it.

Proof of Theorem 2 (for large δ). Combining Lemma 9 with Lemma 10 gives

$$P_N(G(\delta)) = (1 + o(1))\frac{2}{\delta'\sqrt{\pi N}}\left(\mathbb{E}\frac{Z^-}{A} + O\left(\mathbb{E}\frac{\delta^{-s}}{A}\right)\right) + O(e^{-\delta^{-s}})$$

$$= (1 + o(1))\frac{1}{\delta\sqrt{\pi N}}\left(\delta^2 + O\left(\frac{\delta^{1-s}}{\sqrt{T}}\right)\right) + O(e^{-\delta^{-s}})$$

$$\sim \frac{\delta}{\sqrt{\pi N}},$$

where we used that $T = \Omega(\delta^{-r}) = \Omega(\delta^{-6})$ to show that $\delta^2 + O(\delta^{1-s}T^{-1/2}) \sim \delta^2$. ∎

7 | NUMERICAL SIMULATIONS AND FURTHER QUESTIONS, AND BIOLOGICAL APPLICATIONS

To double check the results of Theorem 1, we simulated the process for $N = 250$ and $c = 2$. Thus, $\delta = c/\sqrt{N} = 2/\sqrt{250} \approx 0.126$. Theorem 1 predicts that as $N \to \infty$, $N\mathbb{P}_N(G(\delta)) \to g(2)$. Numerically evaluating the integral defining $g(2)$ gives approximately 1.516. Our quick and dirty Monte Carlo simulation gives $N\mathbb{P}_N(G(\delta)) = 1.521 \pm 0.06$. We could have done more simulations to lower the standard error, but in fact because we ran simulations for $N = 10m$ for every $m \leq 25$, there is already greater accuracy. Figure 2 shows all of these data points, as well as similar data for $c = 3$ and $1 \leq m \leq 15$ (here $g(3) \approx 1.97$). The limits predicted by Theorem 1 are corroborated, or at least not contradicted, by the data.

Next we ran simulations to investigate Conjecture 1. Recall, the limit is known when $\delta \to 0$ as fast as any power $(\log N)^{-\varepsilon}$, whereas this should fail when δ remains constant; the conjecture covers the ground in between, which is clearly too slim to distinguish numerically. The best we could do was to hold δ constant, thus allowing $c := \delta\sqrt{N}$ to go to infinity. One might expect (3)
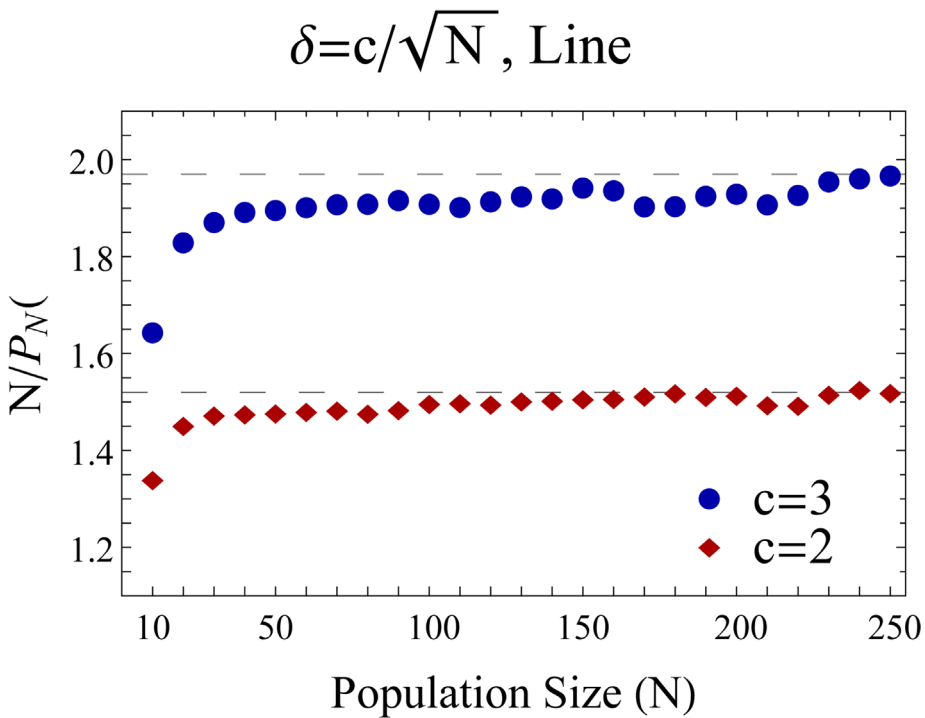
FIGURE 2 The average mutant fixation probability, $\langle P_N \rangle$ times N as a function of N for the line model. The fitness of both mutants and normals at different locations are drawn the same two-valued distribution function, where values $1 - \delta$ and $1 + \delta$ are equally likely. Two different values of δ are used: $2/\sqrt{N}$ (red points) and $3/\sqrt{N}$ (blue points). The points are based on stochastic simulations and each data point represents the average over 10^6 independent realizations

that $N\mathbb{P}_N(G(\delta))$ is well approximated by $g(c)$, leading to

$$\sqrt{\pi N}\mathbb{P}_N(G(\delta)) \approx \frac{g(\delta\sqrt{N})\sqrt{\pi}}{\sqrt{N}}, \tag{26}$$

which is asymptotic to δ by (5). Indeed, the data (red points in Figure 3) is a very good match for (26) (the blue curve in Figure 3), which can be seen to be asymptotic to 0.2.

Among the open questions on this model, one that looms large is whether these results or something similar can be transferred to the circular model. Between the line and circle model, neither seems inherently more compelling; however the fact that the birth and death chain reasoning holds only for the line model has prevented us understanding the situation on any other graphs or initial conditions. We enumerate some problems in what we expect to be increasing order of difficulty.

Problem 1. On a line segment graph, extend the model to the case where the initial configuration is something other than mutants in an interval containing an endpoint.
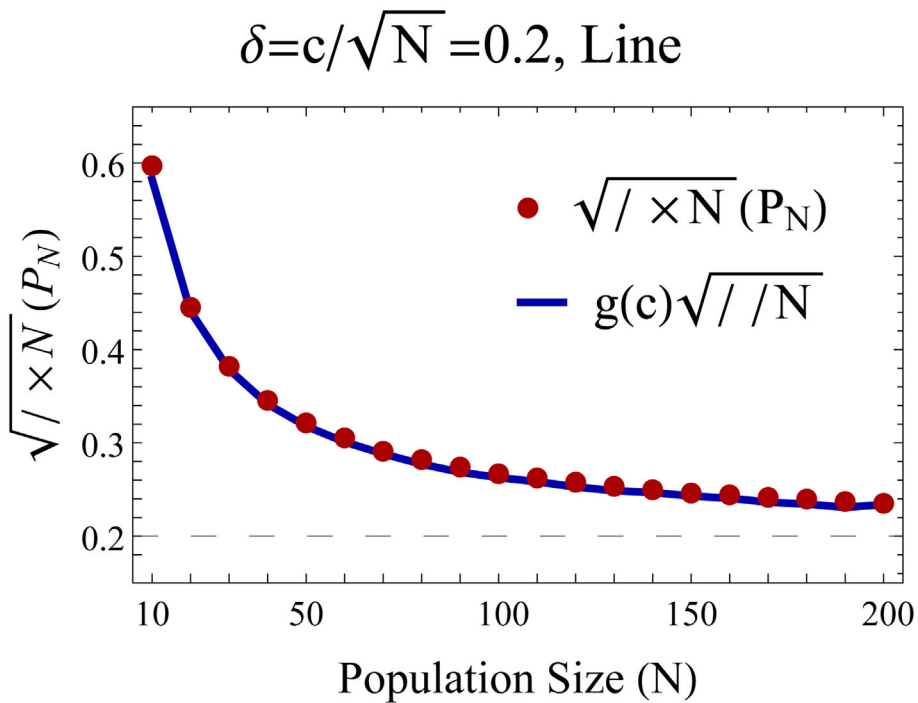
Problem 2. Extend the analysis to a circle.

FIGURE 3 Formula in Equation (26) ($g(\delta\sqrt{N})\sqrt{\pi/N}$, blue curve) is compared with the stochastic simulation results for $\sqrt{\pi \times N}\langle P_N \rangle$ (red points), plotted a functions of N, with $\delta = 0.2$

We were curious whether empirically, the circle appears to behave differently the line. Figure 4 shows the comparison. It appears that the limiting value of $N\mathbb{P}_N(G(\delta))$ for the circle is just a shade less than for the line. However, also, it appears that the value approaches the limit much faster for the circle, and perhaps with less sample variance. On a circle, starting with a single mutation, the interval set of mutant sites remains an interval, which can now grow and shrink at both ends rather than just on the right. These two growth processes are not independent, but may still be the reason we observe faster convergence and lesser variance.

Problem 3. Extend the analysis to any graph with a vertex of degree at least 3. The difficulty here is that the cluster of mutants can become disconnected.

The present study contributes to theoretical understanding of evolutionary processes that have important biomedical applications. The first step in cancer initiation is often a spread and (local) fixation of a neutral mutation, which by itself does not confer an explicit selective advantage to the cell, but serves as a springboard for further transformations. For example, mutations in the so-called tumor suppressor genes drive the progression of many cancers, including colorectal, breast, uterine, ovarian, lung, head and neck, pancreatic, and bladder cancer.[48,49] Tumor suppressor genes are sometimes compared with a brake pedal on a car, as they keep the cell's reproduction in check, preventing it dividing too quickly. An inactivation of a single copy of a tumor suppressor gene is often considered a "neural" mutation, because if the second copy is still active, an inactivation of a single copy of the gene does not result in any phenotypic changes. It is only when the second copy of a tumor suppressor gene is inactivated, the cell starts experiencing a
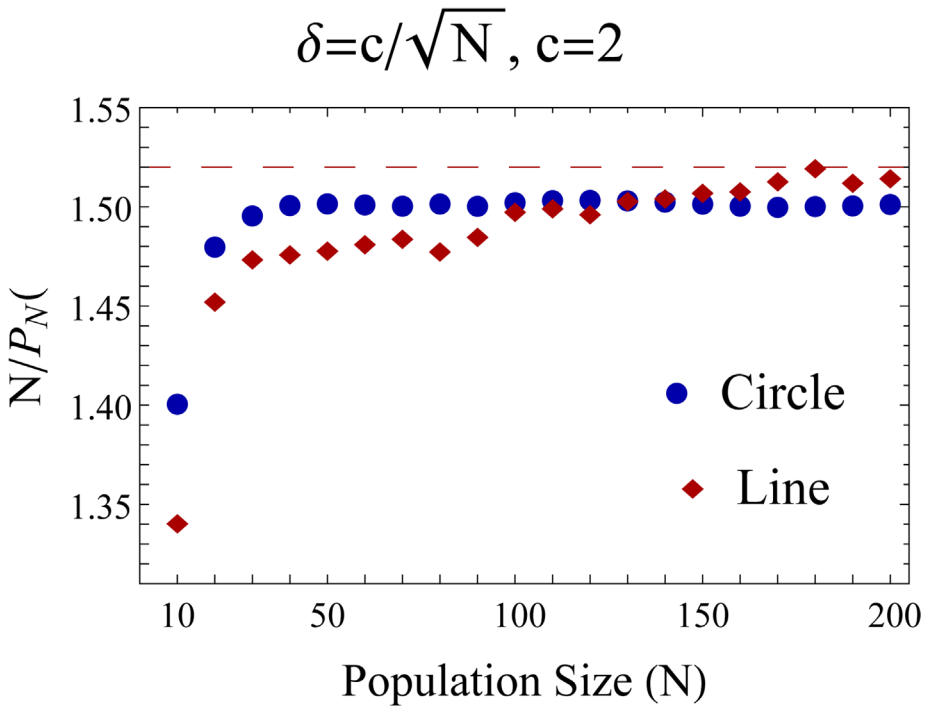
FIGURE 4 Comparison of the line and the circle model. The quantity $N\langle P_N \rangle$ is plotted as a function of N for the circle model (blue dots) and line model (red dots) with $\delta = 2/\sqrt{N}$. For each value of N, the average of 10^6 random simulations is presented

selective advantage (because the "brake" is "off"). The spread and local fixation of mutants with a single, selectively neutral, mutation inactivating the first copy of a tumor suppressor gene is the type of problem where our present results can be applied. For example, a plausible scenario for colorectal cancer initiation is fixation of a single-hit mutant, which comes to dominate a local compartment of colonic tissue (called a crypt). This could be followed eventually by the second mutation, which then leads to a local outgrowth and creation of a "dysplastic crypt" or a polyp. The first stage (the fixation of neutral, single-hit mutants) has been studied extensively in the context of tumor-suppressor gene inactivation (see, e.g., Refs. 50, 51) but not in the presence of environmental randomness. Results reported in this paper allow to account for the role of variability in tissue microenvironment, and suggest that single-mutant fixation is more likely than predicted by nonrandom models. Further models that include more realistic geometries, as well as heterogeneity of cell types (such as stem cells vs. differentiated cells) will require further mathematical efforts. The current article lays a foundation for such future efforts.

ACKNOWLEDGMENTS

REFERENCES
1. Farhang-Sardroodi S, Darooneh AH, Nikbakht M, Komarova NL, Kohandel M. The effect of spatial randomness on the average fixation time of mutants. *PLoS Comput. Biol.* 2017;13(11):e1005864.

2. Farhang-Sardroodi S, Darooneh AH, Kohandel M, Komarova NL. Environmental spatial and temporal variability and its role in non-favoured mutant dynamics. *J R Soc Interface* 2019;16(157):20180781.

3. Chesson PL, Warner RR. Environmental variability promotes coexistence in lottery competitive systems. *Amer Nat.* 1981;117(6):923–943.

4. Pulliam HP. Sources, sinks, and population regulation. *Amer Nat.* 1988;132(5):652–661.

5. Hassell MP, Comins HN, May RM. Species coexistence and self-organizing spatial dynamics. *Nature* 1994;370(6487):290–292.

6. Hanski IA, Gaggiotti OE. *Ecology, Genetics and Evolution of Metapopulations*. New York: Academic Press, 2004.

7. Ellner SP, Snyder RE, Adler PB, Hooker G. An expanded modern coexistence theory for empirical applications. *Ecol Lett.* 2019;22(1):3–18.

8. Burdon J, Haldane S. A mathematical theory of natural and artificial selection, part v: selection and mutation. *Math Proc Cambridge Philos Soc.* 1927;23(7):838–844.

9. Fisher RA. The evolution of dominance in certain polymorphic species. *Amer Nat.* 1930;64(694):385–406.

10. Wright S. Evolution in mendelian populations. *Genetics* 1931;16(2):97.

11. Kimura M. Evolutionary rate at the molecular level. *Nature* 1968;217(5129):624–626.

12. Kimura M. The neutral theory of molecular evolution and the world view of the neutralists. *Genome* 1989;31(1):24–31.

13. Gillespie JH. Natural selection for variances in offspring numbers: a new evolutionary principle. *Amer Nat.* 1977;111(981):1010–1014.

14. Frank SA, Slatkin M. Evolution in a variable environment. *Amer Nat.* 1990;136(2):244–260.

15. Frank SA. Natural selection. I. Variable environments and uncertain returns on investment. *J Evol Biol.* 2011;24(11):2299–2309.

16. Melbinger A, Vergassola M. The impact of environmental fluctuations on evolutionary fitness functions. *Sci Rep.* 2015;5:15211.

17. Cvijović I, Good BH, Jerison ER, Desai MM. Fate of a mutation in a fluctuating environment. *Proc Natl Acad Sci*, 2015;112(36):E5021–E5028.

18. Chesson P. Multispecies competition in variable environments. *Theor Popul Biol.* 1994;45(3):227–276.

19. Chesson P. Mechanisms of maintenance of species diversity. *Annu Rev Ecol Evol Syst.* 2000;31(1):343–366.

20. Adler PB. Janneke HilleRisLambers, Jonathan M Levine. A niche for neutrality. *Ecol Lett.* 2007;10(2):95–104.

21. Meyer I, Shnerb NM. Noise-induced stabilization and fixation in fluctuating environment. *Sci Rep.* 2018;8(1):1–12.

22. Meyer I, Shnerb NM. Evolutionary dynamics in fluctuating environment. *Phys Rev Res.* 2020;2(2):023308.

23. Kessler D, Suweis S, Formentin M, Shnerb NM. Neutral dynamics with environmental noise: age-size statistics and species lifetimes. *Physical Review E* 2015;92(2):022722.

24. Hidalgo J, Suweis S, Maritan A. Species coexistence in a neutral dynamics with environmental noise. *J Theor Biol.* 2017;413:1–10.

25. Danino M, Shnerb NM. Fixation and absorption in a fluctuating environment. *J Theor Biol.* 2018;441:84–92.

26. Stewart PS, Franklin MJ. Physiological heterogeneity in biofilms. *Nat Rev Microbiol.* 2008;6(3):199–210.

27. Jayasinghe N, Franks A, Nevin KP, Mahadevan R. Metabolic modeling of spatial heterogeneity of biofilms in microbial fuel cells reveals substrate limitations in electrical current generation. *Biotechnol J.* 2014;9(10):1350–1361.

28. Boles BR, Thoendel M, Singh PK. Self-generated diversity produces "insurance effects" in biofilm communities. *Proc Natl Acad Sci.* 2004;101(47):16630–16635.

29. Li X-F, Carlin S, Urano M, Russell J, Ling CC, O'Donoghue JA. Visualization of hypoxia in microscopic tumors by immunofluorescent microscopy. *Cancer Res.* 2007;67(16):7646–7653.

30. Graves EE, Maity A, Le Q-T. The tumor microenvironment in non-small-cell lung cancer. *Semin. Radiat. Oncol.* 2010;20(3):156–163. https://doi.org/10.1016/j.semradonc.2010.01.003

31. Yuan Y. Spatial heterogeneity in the tumor microenvironment. *Cold Spring Harb Perspect Med.* 2016;6(8):a026583.

32. Santos FC, Pacheco JM. Scale-free networks provide a unifying framework for the emergence of cooperation. *Appl Phys Lett.* 2005;95(9):098104.

33. Santos FC, Rodrigues JF, Pacheco JM. Graph topology plays a determinant role in the evolution of cooperation. *Proc R Soc B Biol Sci* 2006;273(1582):51–55.

34. Santos FC, Pacheco JM, Lenaerts T. Cooperation prevails when individuals adjust their social ties. *PLoS Comput Biol* 2006;2(10):e140.

35. Santos FC, Santos MD, Pacheco JM. Social diversity promotes the emergence of cooperation in public goods games. *Nature* 2008;454(7201):213–216.

36. Tomassini M, Pestelacci E, Luthi L. Social dilemmas and cooperation in complex networks. *Int J Mod Phys C* 2007;18(07):1173–1185.

37. Maciejewski W, Fu F, Hauert C. Evolutionary game dynamics in populations with heterogenous structures. *PLoS Comput Biol.* 2014;10(4):e1003567.

38. Nagylaki T. The strong-migration limit in geographically structured populations. *J Math Biol.* 1980;9(2):101–114.

39. Tachida H, Iizuka M. Fixation probability in spatially changing environments. *Genet Res.* 1991;58(3):243–251.

40. Gavrilets S, Gibson N. Fixation probabilities in a spatially heterogeneous environment. *Popul Ecol.* 2002;44(2):51–58.

41. Whitlock MC, Gomulkiewicz R. Probability of fixation in a heterogeneous environment. *Genetics* 2005;171(3):1407–1417.

42. Chesson P. General theory of competitive coexistence in spatially-varying environments. *Theor Popul Biol.* 2000;58(3):211–237.

43. Patrick Alfred Pierce Moran. Random processes in genetics. *Math Proc Cambridge Philos Soc.* 1958;54(1):60–71.

44. Komarova NL. Spatial stochastic models for cancer initiation and progression. *Bull Math Biol.* 2006;68(7):1573–1599.

45. Michor F, Iwasa Y, Rajagopalan H, Lengauer C, Nowak MA. Linear model of colon cancer initiation. *Cell Cycle* 2004;3(3):356–360.

46. Aurzada F, S Dereich. Universality of the asymptotics of the one-sided exit problem for integrated processes. *Ann IHP Prob Stat.* 2013;49:236–251.

47. Lawler G, Limic V. *Random Walk: A Modern Introduction*. Cambridge: Cambridge University Press, 2010.

48. Sherr CJ. Principles of tumor suppression. *Cell* 2004;116(2):235–246.

49. Joyce C, Kasi A. Cancer, tumor-suppressor genes. In *StatPearls [Internet]*. StatPearls Publishing, 2018. https://europepmc.org/article/nbk/nbk532243

50. Nowak MA, Komarova NL, Sengupta A, et al.. The role of chromosomal instability in tumor initiation. *Proc Natl Acad Sci.* 2002;99(25):16226–16231.

51. Komarova NL, Sengupta A, Nowak MA. Mutation–selection networks of cancer initiation: tumor suppressor genes and chromosomal instability. *J Theor Biol.* 2003;223(4):433–450.

52. Matsumoto H, Yor M. Exponential functionals of Brownian motion, I: probability laws at fixed time. *Probab Surv.* 2005;2:312–347.

APPENDIX A: PROOFS SECTION 4

Proof of Proposition 2. The first two are proved as Matsumoto and Yor (Ref. 52, Proposition 5.9). We quickly derive the third by integrating (15). Letting $y := x/\sqrt{t}$,

$$m_2(t) = t^{-1} \int \frac{xe^{-x}}{\sinh(x)} \, dN(0,t)(x)$$

$$= t^{-1/2} \int \frac{ye^{-\sqrt{t}y}}{\sinh(\sqrt{t}y)} \, dN(0,1)(y).$$

As $t \to \infty$, the quantity $ye^{-\sqrt{t}y}/\sinh(\sqrt{t}y)$ converges pointwise to $2|y|\mathbf{1}_{y<0}$. Truncating, integrating , and taking limits gives

$$t^{1/2}m_2(t) \to \int_{-\infty}^{0} 2|y|\, dN(0,1)(y) = \mathbb{E}|N(0,1)| = \sqrt{\frac{2}{\pi}}.$$

∎

Proof of Lemma 2. Let $f_\alpha(x,t)$ denote the density of $A_\alpha(t)^{-1}$ at x. Let $W_t := (\alpha/\nu)B_{\nu^2t/\alpha^2}$. Then $\{W_t\}$ is also a standard Brownian motion and $\alpha B_t = \nu W_t$. Hence,

$$f_\alpha(x,t)dx = \mathbb{P}\left(\frac{1}{\int_0^t e^{\alpha B_s}\, ds} \in [x, x+dx]\right)$$

$$= \mathbb{P}\left(\frac{1}{(\nu/\alpha)^2 \int_0^{(\alpha/\nu)^2 t} e^{\nu W_u}\, du} \in [x, x+dx]\right)$$

$$= \mathbb{P}\left(\frac{1}{\int_0^{(\alpha/\nu)^2 t} e^{\nu W_u}\, du} \in [(\nu/\alpha)^2 x, (\nu/\alpha)^2 x + (\nu/\alpha)^2 dx]\right)$$

$$= \frac{\nu^2}{\alpha^2} f_\nu\left(\frac{\nu^2}{\alpha^2}x, \frac{\alpha^2}{\nu^2}t\right)dx. \tag{A1}$$

Consequently, changing variables to $\theta = (\nu^2/\alpha^2)x$,

$$m_\alpha(t) = \int_{-\infty}^{\infty} x f_\alpha(x,t)\, dx$$

$$= \frac{\nu^2}{\alpha^2} \int_{-\infty}^{\infty} x f_\nu\left(\frac{\nu^2}{\alpha^2}x, \frac{\alpha^2}{\nu^2}t\right)dx$$

$$= \frac{\alpha^2}{\nu^2} \int_{-\infty}^{\infty} \theta f_\nu\left(\theta, \frac{\alpha^2}{\nu^2}t\right)d\theta$$

$$= \frac{\alpha^2}{\nu^2} m_\nu\left(\frac{\alpha^2}{\nu^2}t\right),$$

proving (18). Set $\nu = 2$ and $t = 1$, plug into (17), and send α to infinity to obtain

$$m_\alpha(1) = \frac{\alpha^2}{4}m_2\left(\frac{\alpha^2}{4}\right) \sim \frac{\alpha^2}{4}\sqrt{\frac{2}{\pi\alpha^2/4}} = \frac{\alpha}{\sqrt{2\pi}},$$

proving (19). ∎

APPENDIX B: PROOFS SECTION 6

Proof of Lemma 4. For Brownian motion run to time M, the reflection principle gives

$$\mathbb{P}(\sup_{0 \le t \le M} B_t \le u) = 1 - 2\mathbb{P}_0(B_M \ge u),$$

which is asymptotic to $(2/\pi)^{1/2}uM^{-1/2}$ uniformly as u varies over the $(0, M^{1/2-\varepsilon}]$ for any $\varepsilon \in (0, 1/2)$. Pick $\alpha > 1/2$. By (20), one then has

$$\sqrt{\frac{2}{\pi}} \frac{u - C_\alpha \log M}{M^{1/2}} - M^{-\alpha} \le \mathbb{P}(\max_{1 \le j \le M} S_j \le u) \le \sqrt{\frac{2}{\pi}} \frac{u + C_\alpha \log M}{M^{1/2}} + M^{-\alpha}.$$

■

Proof of Lemma 6. By (15), we compute

$$\mathbb{E} \frac{B_M^-}{\int_0^M \exp(2B_s)\,ds} = \int_{\mathbb{R}} \frac{-x^2 e^{-x} \mathbf{1}_{x \le 0}}{M \sinh(x)}\,dN(0, M)(x)$$

$$= \int_{\mathbb{R}} \frac{-y^2 e^{-\sqrt{M}y} \mathbf{1}_{y \le 0}}{\sinh(\sqrt{M}y)}\,dN(0, 1)(y).$$

As $M \to \infty$, $e^{-\sqrt{M}y}/\sinh(\sqrt{M}y) \to -2 \cdot \mathbf{1}_{y < 0}$; truncating, integrating, and taking limits then gives

$$\mathbb{E} \frac{B_M^-}{\int_0^M \exp(2B_s)\,ds} \to \int_{-\infty}^0 2y^2 dN(0, 1)(y) = 1.$$

By differentiating eq. (5.7) of Ref. [52] twice with respect to λ, we have that

$$\mathbb{E}\left[\left(\int_0^M \exp(2B_s)\,ds\right)^{-2} \bigg| B_M = x\right] = \frac{e^{-2x}(x^2 \sinh(x) + Mx \cosh(x) - M \sinh(x))}{M^2 \sinh(x)^3}.$$

This implies that

$$\frac{\mathbb{E}[Y_M^2]}{\sqrt{M}} = \int_{\mathbb{R}} \frac{e^{-2x} \mathbf{1}_{x<0}(x^4 \sinh(x) + Mx^3 \cosh(x) - Mx^2 \sinh(x))}{M^{5/2} \sinh(x)^3}\,dN(0, M)(x)$$

$$= \int_{\mathbb{R}} \frac{e^{-2y\sqrt{M}} \mathbf{1}_{y<0}(M^2 y^4 \sinh(\sqrt{M}y) + M^{5/2} y^3 \cosh(\sqrt{M}y) - M^2 y^2 \sinh(\sqrt{M}y))}{M^{5/2} \sinh(\sqrt{M}y)^3}\,dN(0, 1)(y).$$

The integrand converges to $-4y^3\mathbf{1}_{y<0}$ as $M \to \infty$; truncating, integrating, and taking limits shows

$$\frac{\mathbb{E}[Y_M^2]}{\sqrt{M}} \to \int_{\mathbb{R}} -4y^3\mathbf{1}_{y<0}dN(0,1)(y) = 4\sqrt{\frac{2}{\pi}}.$$

∎

Proof of Lemma 7. Using Lemma 6 together with Chebyshev's inequality, we see that

$$\mathbb{P}[Y_M \geq t] \leq C\sqrt{M}t^{-2}$$

for some constant C. This implies that

$$\mathbb{E}[Y_M\mathbf{1}_{Y_M \geq M^{1/2+\varepsilon/2}}] = \int_{M^{1/2+\varepsilon/2}}^{\infty} \mathbb{P}[Y_M \geq t]\,dt \leq CM^{-\varepsilon/2} = o(1).$$

We then may write

$$\mathbb{E}[Y_M\mathbf{1}_E] = \mathbb{E}[Y_M\mathbf{1}_E\mathbf{1}_{Y_M \geq M^{1/2+\varepsilon/2}}] + \mathbb{E}[Y_M\mathbf{1}_E\mathbf{1}_{Y_M < M^{1/2+\varepsilon/2}}] \leq o(1) + M^{1/2+\varepsilon/2}\mathbb{P}[E] \to 0.$$

∎

Proof of Lemma 10. By Lemma 3, there exists a coupling of $\{\frac{\widetilde{S}_t}{\delta'/\sqrt{2}}\}_{0 \leq t \leq T}$ and $\{B_t\}_{0 \leq t \leq T}$ so that

$$\mathbb{P}\left[\sup_{t \in [0,T]}\left|\frac{\widetilde{S}_t}{\delta'/\sqrt{2}} - B_t\right| \geq C\log(T)\right] \leq \frac{1}{T^2}.$$

Let E denote the event in the above probability; conditioned on E^c, we have

$$A = \int_0^T \exp(\widetilde{S}_s)\,ds = \int_0^T \exp(\delta'B_s/\sqrt{2} + O(\delta'\log(T)))\,ds \sim \int_0^T \exp(\delta'B_s/\sqrt{2})\,ds,$$

where the last asymptotic equality follows $\delta'\log(T) = O(\delta\log\delta) \to 0$. This means that

$$\mathbb{E}[A^{-1}] = \mathbb{E}[A^{-1}\mathbf{1}_E] + \mathbb{E}[A^{-1}\mathbf{1}_{E^c}] = O(T^{-2}) + (1+o(1))\mathbb{E}\frac{\mathbf{1}_{E^c}}{\int_0^T \exp(\delta'B_s/\sqrt{2})\,ds}. \tag{B1}$$

By Brownian scaling,

$$\int_0^T \exp\left(\delta'B_s/\sqrt{2}\right)ds = \int_0^T \exp\left(2(B_{(\delta')^2s/2^3})\right)ds = 2^3(\delta')^{-2}\int_0^{T(\delta')^2/2^3} \exp(2B_s)\,ds$$

because $\delta' \sim 2\delta$.

By assumption, $T(\delta')^2 \to \infty$; further, the uniform integrability statement in Lemma 5 shows that

$$\mathbb{E}\left[\left(\int_0^{T(\delta')^2/2^3} \exp(2B_s)\,ds\right)^{-1}\mathbf{1}_{E^c}\right] \sim \mathbb{E}\left[\left(\int_0^{T(\delta')^2/2^3} \exp(2B_s)\,ds\right)^{-1}\right]$$

because the variables $\mathbf{1}_E X_M/\mathbb{E}[X_M]$ converge almost surely to 0 (because $\mathbb{P}(E) \to 0$) and are uniformly integrable. From here, (B1) then gives

$$\mathbb{E}[A^{-1}] = O(T^{-2}) + (1 + o(1)) \cdot \frac{(\delta')^2}{2^3} \cdot \sqrt{\frac{2}{\pi T}} \cdot \frac{2^{3/2}}{\delta'} = O(T^{-2}) + (1 + o(1))\frac{2^2}{\delta'}\frac{1}{\sqrt{\pi T}} \sim \frac{\delta}{\sqrt{\pi T}}.$$

Similarly,

$$\mathbb{E}[Z^-/A] = \mathbb{E}\left[\frac{Z^-}{A} \cdot \mathbf{1}_E\right] + \mathbb{E}\left[\frac{Z^-}{A} \cdot \mathbf{1}_{E^c}\right]$$

$$= O(T^{-2}) + \frac{\delta'}{\sqrt{2}}\mathbb{E}\left[\frac{B_T^-\mathbf{1}_{E^c}}{\int_0^T \exp\left(\delta'B_s/\sqrt{2}\right)ds}\right] + O\left(\delta\log(T)\mathbb{E}\left[\left(\int_0^T \exp(\delta'B_s/\sqrt{2})\,ds\right)^{-1}\right]\right)$$

$$= O\left(\frac{\log(T)}{\sqrt{T}}\right) + \frac{\delta'}{\sqrt{2}}\mathbb{E}\left[\frac{B_T^-\mathbf{1}_{E^c}}{\int_0^T \exp\left(\delta'B_s/\sqrt{2}\right)ds}\right].$$

Using the same Brownian scaling as in below (B1), note

$$\frac{B_T^-}{\int_0^T \exp(\delta'B_s/\sqrt{2})\,ds} \stackrel{d}{=} \frac{\delta'}{2^{3/2}}\frac{B_{(2^{3/2}/\delta')^2T}^-}{\int_0^{(2^{3/2}/\delta')^2T} \exp(2B_s)\,ds} = \frac{\delta'}{2^{3/2}}Y_M,$$

where we set $M = (2^{3/2}/\delta')^2T$. By Lemma 7 together with Lemma 6, we have

$$\mathbb{E}[Y_M\mathbf{1}_{E^c}] = \mathbb{E}[Y_M] + o(1) = 1 + o(1).$$

Combining the above equalities provides

$$\mathbb{E}[Z^-/A] = O\left(\frac{\log(T)}{\sqrt{T}}\right) + (1 + o(1)) \cdot \frac{\delta'}{\sqrt{2}} \cdot \frac{\delta'}{2^{3/2}} \sim \delta^2.$$

■