Survival probabilities of history-dependent random walks

Uri Keshet[1] and Shahar Hod[2]

[1]*Physics Faculty, Weizmann Institute, Rehovot 76100, Israel*

[2]*The Racah Institute of Physics, The Hebrew University, Jerusalem 91904, Israel*

We analyze the dynamics of random walks with long-term memory (binary chains with long-range correlations) in the presence of an absorbing boundary. An analytically solvable model is presented, in which a dynamical phase transition occurs when the correlation strength parameter μ reaches a critical value μ_c. For strong positive correlations, $\mu > \mu_c$, the survival probability is asymptotically finite, whereas for $\mu < \mu_c$ it decays as a power law in time (chain length).

Dynamical systems with long-range spatial or temporal correlations are attracting considerable interest across many disciplines, lending applications to physical, biological, social, and economical sciences (see, e.g., [1–7] and references therein). Such systems are often analyzed using random walks, a fundamental concept of statistical physics. Random walks lend applications to numerous scientific fields (see, e.g., [8–13] and references therein). In particular, random walks in the presence of absorbing traps (boundaries) have been studied in recent years as models for various systems such as absorbing-state phase transitions [14,15], polymer adsorption [16], granular segregation [17], the spreading of an epidemic [18], and in the context of complex adaptive systems [19,20]. In this work, we analytically study random walks with an absorbing boundary, in which the jump probability is history-dependent, resulting in long-range correlations.

The statistical properties of data such as DNA strings, written texts, and financial data (e.g., stock market quotes) are known to significantly deviate from those of purely random sequences [21,7]. Such systems may be studied by mapping them onto a correlated sequence of symbols. Although the nature of the resulting sequence may depend upon the choice of mapping (see, e.g., [22]), the essential statistical properties of the original system are often preserved. By choosing a mapping of these systems onto two symbols [5], the problem is reduced to the exploration of correlated binary chains, which are equivalent to one-dimensional random walks with a constant step size. These binary chains have long-range correlations and often exhibit a superdiffusive nature, in which the variance grows asymptotically faster than the string length.

The preceding discussion motivates a study of random walks with a history-dependent jump probability, both with and without an absorbing boundary, as a means of facilitating our understanding of systems with long-range correlation. In Ref. [7], we presented a model for a history-dependent random walk. Although simple, this model features a dynamical phase transition between normal diffusion and superdiffusion as the correlation strength parameter reaches a critical value. In this work, we analyze random walks with long-range correlations in the presence of an absorbing boundary.

We begin by introducing a simple model that incorporates long-range correlations into an otherwise random sequence.

Consider a discrete string of binary symbols, $a_i \in \{0, 1\}$, in which the conditional probability of a given symbol (say, zero) occurring at the position $L+1$ is history-dependent, and is given by

$$p(k,L) = \frac{1}{2}\left(1 - \mu \frac{L - 2k}{L + L_0}\right), \quad (1)$$

where k is the number of such symbols (zeros) appearing in the preceding L bits. The correlation parameter μ, where $-1 < \mu < 1$, determines the strength of correlations in the system. The persistence condition $\mu > 0$ implies that a given symbol in the preceding sequence promotes the birth of a new identical symbol. In the antipersistence regime $\mu < 0$, on the other hand, each symbol inhibits the appearance of a new identical symbol. The parameter $L_0 > 1$ is a constant transient time. For $L \ll L_0$, the sequence is approximately random (uncorrelated), whereas for $L \gg L_0$, the effect of correlations takes over [23].

In this model, the conditional probability $p(k,L;\mu,L_0)$ depends on the number of zeros (or unities) in the preceding bits, and is independent of their arrangement. This allows one to obtain an *analytical* description of the system's dynamical behavior. As demonstrated in [7], this two-parameter model provides a good description of the observed statistical properties of various systems such as coarse-grained DNA strings, written texts, and financial data, when mapped onto a binary chain.

The probability $P(k, L+1)$ of finding k identical symbols (say, zeros) in a sequence of length $L+1$ follows from the evolution equation

$$P(k, L+1) = p(k-1, L)P(k-1, L) + [1 - p(k,L)]P(k,L). \quad (2)$$

Crossing to the continuous limit, one obtains the Fokker-Planck diffusion equation for the correlated process,

$$\frac{\partial P}{\partial L} = \frac{1}{2}\frac{\partial^2 P}{\partial x^2} - \frac{\mu}{L + L_0}\frac{\partial(xP)}{\partial x}, \quad (3)$$

where $x \equiv 2k - L$ is the distance from the origin in the corresponding random walk, and we have neglected high-order terms which are irrelevant for $1 \ll L_0 \ll L$. Solutions of Eq.

(3) under the initial condition $P(x, L=0) = \delta(x)$ were given in [7].

We introduce an absorbing boundary at $x=0$, by imposing the boundary condition $P(x \leq 0, L > 0) = 0$. The evolution equation (3) along with this boundary condition and the initial condition $P(x, L=0) = \delta(x)$ has a solution in the form

$$P(x,L) \propto \frac{(L+L_0)^\mu}{D(L)^{3/2}} x \Theta(x) \exp\left[-\frac{x^2}{2D(L)}\right], \quad (4)$$

where Θ is the Heaviside step function, and $D(L)$ is given by

$$D(L; \mu, L_0) \equiv \frac{L+L_0}{1-2\mu}\left[1 - \left(\frac{L_0}{L+L_0}\right)^{1-2\mu}\right]. \quad (5)$$

Equation (5) breaks down in the special case $\mu = \mu_c$, where $\mu_c \equiv 1/2$. In this case, $D(L)$ is given by

$$D(L; \mu = \mu_c, L_0) \equiv (L+L_0)\ln\left(\frac{L+L_0}{L_0}\right). \quad (6)$$

The first two moments of the distribution function of the survived walkers are given by $\langle x \rangle^2 = \pi D(L)/2$ and $\langle x^2 \rangle = 2D(L)$, where

$$\langle A \rangle \equiv \frac{\displaystyle\int_{-\infty}^{\infty} A(x) P(x,L)\,dx}{\displaystyle\int_{-\infty}^{\infty} P(x,L)\,dx}. \quad (7)$$

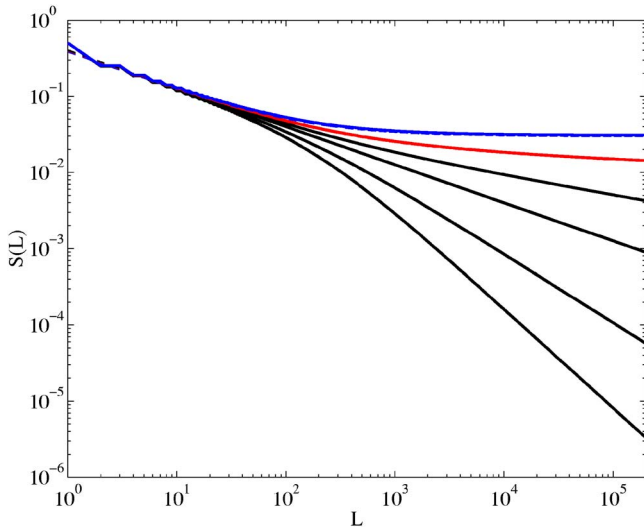The variance of the probability distribution $P(x,L)$ thus equals

FIG. 1. (Color online) Survival probability $S(L)$ as a function of sequence length L. Shown are numerical (solid lines) and analytical (dashed lines) results for $\mu = -0.8, -0.4, 0, 0.25, 0.5$, and 0.8 (from bottom to top). Here, $x_0 = 0$ (half the walkers survive the first step) and $L_0 = 100$. The analytical results agree with the numerical ones to better than 1%.

$$V(L; \mu, L_0) = \frac{4-\pi}{2} D(L). \quad (8)$$

This result implies that for $\mu < \mu_c$, the asymptotic variance scales linearly with the string length, whereas for $\mu > \mu_c$ it scales as $V \propto L^{2\mu}$. Hence, a history-dependent sequence with strong positive correlations ($\mu > \mu_c$) is characterized by a superdiffusion phase in which the variance grows asymptotically faster than L, both without [7] and with an absorbing boundary.

The survival probability $S(L) \equiv \int_0^\infty P(x,t)\,dx$ of the walkers is given, for $\mu \neq \mu_c$, by

$$S(L; \mu \neq \mu_c) \propto \left|\left(\frac{L+L_0}{L_0}\right)^{1-2\mu} - 1\right|^{-1/2}, \quad (9)$$

whereas for $\mu = \mu_c$ we find

$$S(L; \mu_c) \propto \ln^{-1/2}\left(\frac{L+L_0}{L_0}\right). \quad (10)$$

The survival probability thus changes its asymptotic ($L \gg L_0$) behavior at the phase transition value $\mu_c = 1/2$, and one identifies three qualitatively different regimes,

$$S(L \gg L_0) \propto \begin{cases} L^{-(1/2)+\mu}, & \mu < \mu_c \\ \ln^{-1/2}(L/L_0), & \mu = \mu_c \\ \text{const}, & \mu > \mu_c. \end{cases} \quad (11)$$

The normalization of the survival probability is sensitive to the discrete initial conditions. Since the chain is nearly random for $L \ll L_0$, the normalization may be approximated by equating $S(L)$, for $1 \ll L \ll L_0$, to the survival probability of a purely random, continuous walk,
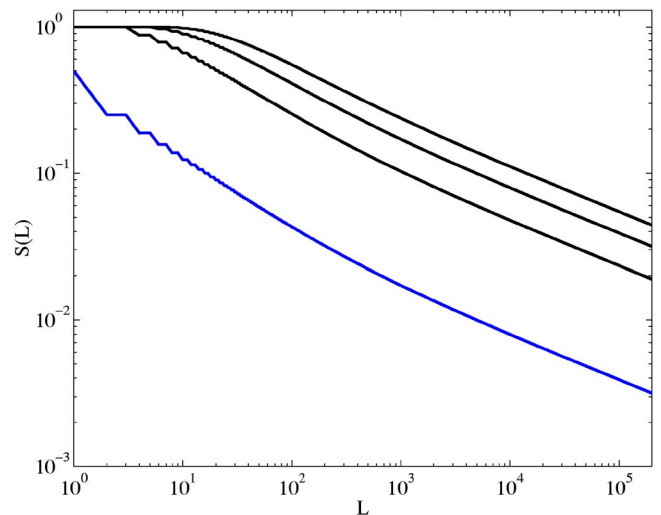

FIG. 2. (Color online) Survival probability S as a function of string length L for various boundary locations x_0. We present results for $x_0 = 0, -2, -4$, and -6 (from bottom to top), with $\mu = 0.2$ and $L_0 = 100$.

$$S_0(L \gg 1) \simeq 1 - 2 \int_{|x_0|}^{\infty} \frac{1}{\sqrt{2\pi L}} \exp\left(-\frac{x^2}{2L}\right) dx = \mathrm{erf}(|x_0|/\sqrt{2L}),$$

(12)

where x_0 is the distance between the absorbing boundary and the origin. We thus define

$$S(1 \ll L \ll L_0) \simeq \mathrm{erf}(|x_0|/\sqrt{2L}).$$

(13)

In order to confirm the analytical results, we perform numerical simulations of (discrete) binary sequences. Figure 1 displays the resulting survival probability $S(L)$ of correlated strings with various values of the correlation parameter μ. We find an excellent agreement between the analytical results [Eqs. (9), (10), and (13)] and the numerical ones.

The preceding results, starting from the distribution function in Eq. (4), are valid only when the absorbing boundary is placed at the origin. One would like to generalize the results for boundaries located at arbitrary locations x_0, under the boundary condition $P(x \le x_0, L > 0) = 0$. Note that the diffusion equation (3) is *not* invariant under the translation $x \to x + d$, where d is a constant. The generalization of our solution for $x_0 \ne 0$ is therefore nontrivial. Nevertheless, we have verified numerically that the asymptotic behavior of the survival probability given by Eq. (11) remains valid for arbitrary values of x_0, in all three regimes. The similar asymptotic behavior for different choices of x_0 is demonstrated numerically in Fig. 2.

The above results can be readily generalized for a biased random walk with a moving boundary. For example, for the biased jump probability

$$p(k, L) = \frac{1}{2}\left(1 + q - \mu \frac{L - 2k}{L + L_0}\right),$$

(14)

the above results will hold if we apply the transformation $x \to x - x_c(L)$, with

$$x_c(L) \equiv q \frac{L + L_0}{1 - \mu}\left[1 - \left(\frac{L_0}{L + L_0}\right)^{1-\mu}\right].$$

(15)

The study of (correlated) random walks in the presence of an absorbing boundary may be relevant for the analysis of financial data such as stock market quotes. For example, it is interesting to study the survival probability $S(t)$ of an investor, defined as the probability to be in profit for all times preceding t. This provides a useful measure of the investment risk.

In summary, in this work we have analyzed the dynamics of random walks in which the jump probabilities are history-dependent, in the presence of an absorbing boundary. Using an analytically solvable model, we identify a dynamical phase transition characterizing the system's global behavior. For small values of the correlation strength ($\mu < \mu_c$), the survival probability decays as $S \propto L^{-(\mu_c - \mu)}$, whereas for $\mu > \mu_c$ the system is characterized by *finite* asymptotic survival probabilities.

[1] R. N. Mantegna and H. E. Stanley, Nature (London) **376**, 46 (1995).

[2] I. Kanter and D. A. Kessler, Phys. Rev. Lett. **74**, 4559 (1995).

[3] H. E. Stanley *et al.*, Physica A **224**, 302 (1996).

[4] A. Provata and Y. Almirantis, Physica A **247**, 482 (1997).

[5] O. V. Usatenko and V. A. Yampol'skii, Phys. Rev. Lett. **90**, 110601 (2003).

[6] A. C. C. Yang, S. S. Hseu, H. W. Yien, A. L. Goldberger, and C. K. Peng, Phys. Rev. Lett. **90**, 108103 (2003).

[7] S. Hod and U. Keshet, Phys. Rev. E **70**, 015104(R) (2004).

[8] M. N. Barber and B. W. Ninham, *Random and Restricted Walks* (Gordon and Breach, New York, 1970).

[9] N. G. van Kampen, *Stochastic Processes in Physics and Chemistry* (North-Holland, Amsterdam, 1992).

[10] R. Fernandez, J. Frohlich, and A. D. Sokal, *Random Walks, Critical Phenomena, and Triviality in Quantum Field Theory* (Springer-Verlag, Berlin, 1992).

[11] G. H. Weiss, *Aspects and Applications of the Random Walk* (North-Holland, Amsterdam, 1994).

[12] D. ben-Avraham and S. Havlin, *Diffusion and Reactions in Fractals and Disordered Systems* (Cambridge University Press, Cambridge, 2000).

[13] R. Dickman and D. ben-Avraham, Phys. Rev. E **64**, 020102(R) (2001).

[14] J. Marro and R. Dickman, *Nonequilibrium Phase Transitions in Lattice Models* (Cambridge University Press, Cambridge, 1999).

[15] H. Hinrichsen, Adv. Phys. **49**, 815 (2000).

[16] K. De'Bell and T. Lookman, Rev. Mod. Phys. **65**, 87 (1993).

[17] Z. Farkas and T. Fulop, J. Phys. A **34**, 3191 (2001).

[18] P. Grassberger, H. Chate, and G. Rousseau, Phys. Rev. E **55**, 2488 (1997).

[19] S. Hod and E. Nakar, Phys. Rev. Lett. **88**, 238702 (2002).

[20] S. Hod, Phys. Rev. Lett. **90**, 128701 (2003).

[21] A. Schenkel, J. Zhang, and Y. C. Zhang, Fractals **1**, 47 (1993).

[22] S. L. Narasimhan, J. A. Nathan, P. S. R. Krishna, and K. P. Murthy, e-print cond-mat/0409053.

[23] The introduction of the parameter L_0 is mainly motivated by the observed behavior of the variance of data such as DNA sequences, written texts, and financial data. These systems are characterized by normal diffusion for $L \ll L_0$ and by a super-diffusive behavior for $L \gg L_0$, where $L_0 \sim 10 - 100$.