Accurate estimation of the survival probability for trapping in two dimensions

Lazaros K. Gallos and Panos Argyrakis
Department of Physics, University of Thessaloniki, GR-54006 Thessaloniki, Greece
(Received 29 May 2001; published 22 October 2001)

In this work we study the mean survival probability $\Phi(n,c)$ of random walks on a two-dimensional lattice in the presence of traps of concentration c, as a function of the number of steps n. The computation of this quantity is performed indirectly by using the distribution of the number of sites visited S_n. In order to achieve an accurate description of this distribution we use a combination of numerical techniques. The method allows an accurate calculation of Φ down to very small values (of the order of 10^{-100}, for example), which is not possible via direct simulations. The survival probability is analyzed in terms of an asymptotic expansion, following the results of Donsker and Varadhan [Commun. Pure Appl. Math. **28**, 525 (1975); 32, 721 (1979)], and by using the outcome of a scaling ansatz, as described in our earlier work.

DOI: 10.1103/PhysRevE.64.051111 PACS number(s): 05.40.Fb, 72.10.−d

I. INTRODUCTION

The trapping problem is one of the most puzzling and most studied problems in the diffusion literature [1–16]. The problem itself can be very simply formulated. We consider a homogenous space in d dimensions, either continuous or discrete. In this space we randomly distribute immobile traps with a fixed known concentration c. These traps are typically spheres or occupy one lattice site each. Particles, originating at random positions, start moving in this environment by performing Brownian motion (in continuous space) or random walk (in the case of a lattice). Multiple occupancy is allowed for the particles and any number of particles can be on the same site simultaneously. When a particle meets a trap its motion stops and the moving particle gets annihilated on this trap. An infinite number of walkers can be absorbed on the same trap. This process corresponds to infinitely deep traps.

The quantity of interest in this work is the survival probability $\Phi(n,c)$, which is the average probability of a random walker to have survived, i.e., not encounter a trap, after n steps in a space containing a trap concentration equal to c. The problem described is only one of many possible variations that include moving traps, shallow traps that allow for detrapping, traps that can host one particle only, funnels encircling the traps, excluded volume effects for the particles, inhomogeneous (for example, fractal) space and many more, which are dictated from the physical system examined. However, even the simplest model is not easily amenable to mathematical handling and there is a large amount of literature on this topic. Exact solutions have been found only in special cases and there is no complete general solution.

There are many physical motivations for studying such a process. Almost any system comprised of two distinct entities and having random walkers moving in it can be described by one of the many variations of the trapping problem. Characteristic examples include compounds of two stoichiometrically different species, almost any material with defects or impurities, trapping of excitons, the kinetics of a chemical reaction $A+B\rightarrow B$, the probability of finding a taxi in a city, etc.

The problem has been extensively studied both analytically and computationally. There is a plethora of results, which, however, are usually restricted in a narrow time and concentration regime. A very important and basic relation, which is known for a long time, can be deduced by very simple arguments. Suppose that the moving particle has survived an n-step random walk. During this time it has visited S_n different sites (some of them more than once), which means that none of these S_n sites can be a trap. If the trap concentration is equal to c, the probability of one site not to be a trap is $1-c$, and consequently the statistical weight of the above described walk on the survival probability is $(1-c)^{S_n}$. Obviously, the average survival probability at time n is exactly equal to

$$\Phi(n,c)=\langle(1-c)^{S_n}\rangle=\langle e^{-\lambda S_n}\rangle, \tag{1}$$

where $\lambda=-\ln(1-c)$, and the average is over all particle realizations of the n-step random walk. Equation (1) can also be expressed in the form

$$\Phi(n,c)=\sum_{s=2}^{n+1} p_n(s)e^{-\lambda s}, \tag{2}$$

where $p_n(s)$ represents the density distribution function of the S_n values for a fixed number of steps n. This distribution has been known for some time in one dimension [3,17], and has been used for accurate asymptotic evaluation of $\Phi(n,c)$ [7,8]; but the only available information for this distribution when $d\geqslant 2$ is the first and second moments [18–20], as well as the fact that it tends to a Gaussian distribution as the dimensionality tends to infinity [2].

The main analytical approaches to this problem include the Rosenstock approximation [5], where the mean value in Eq. (1) is replaced by its typical value, yielding

$$\Phi(n,c)=e^{-\lambda\langle S_n\rangle}, \tag{3}$$

the improvement of this approximation with the use of cumulant expressions [6], and mainly the work of Donsker and Varadhan [4] that resulted in an exact asymptotic expression for $\Phi(n,c)$

$$\lim_{n\to\infty}\Phi(n,c)=\exp(-k_d\lambda^{2/(2+d)}n^{d/(d+2)}). \tag{4}$$

In this expression, k_d is a positive constant depending on the dimensionality and the structure of the lattice. A detailed summary of all the above results can be found in Ref. [1].

The asymptotic expression in Eq. (4), though, has raised a lot of discussion because it is in contrast with the intuitively expected mean-field simple exponential decay, and because it has been impossible up to now to calculate by simulation methods the crossover time, which cannot be predicted by the theory, especially in the case of two dimensions. Recently, Bunde *et al.* [15] have proved that any direct Monte Carlo trapping simulation results in a simple exponential decay, because of the finite lattice size. Additionally, we recently presented a scaling ansatz using an indirect simulation method [16], which clearly showed the crossover regime from the Rosenstock approximation to the Donsker-Varadhan result.

The asymptotic behavior of the survival probability can also be deduced by heuristic arguments [11,12], which depict that the main mechanism dominating the Φ behavior is the combination of large trap-free regions with random walkers that are restricted in these regions. For shorter times, the important factor is mainly walks that are very compact, i.e., result in very small S_n values, even for large n. The relative probability of occurrence for these walks is very low, but can become important as n grows larger or for small values of c. By inspecting Eq. (2) we can see that the survival probability is the result of two fundamental terms; it is the sum of the points derived after multiplying two functions, one exponential and one—loosely speaking—Gaussian shaped. The main contribution in this sum comes from the left wing of the $p_n(s)$ distribution, because of the exponential factor that decays rapidly and only weighs the small values of s. The number of terms needed in Eq. (2) for acquiring a given accuracy for $\Phi(n,c)$, say 95%, depends both on the number of steps n and the trap concentration c. For smaller n and lower c we need a larger portion of the distribution. For example, if we fix $n=500$, we need around 40 terms when $c=0.9$ and 220 terms when $c=0.01$, in order to converge to a 95% accuracy. These terms correspond to the 8% and 44% portion of the $p_n(s)$ distribution in the s axis. Similarly, if we fix $c=0.9$ when $n=100$ we need 16 terms in Eq. (2), which corresponds to 16% of the distribution and is twice the corresponding value for $n=500$. Moreover, in all these cases more than 50% in the calculation of $\Phi(n,c)$ is contributed by the first 10–20 terms. It is thus clear that the regime around the peak of the distribution has practically a small contribution in the sum when the number of steps n and the concentration c are not too small. Upon lowering c the contribution of the peak becomes progressively more important, but the left-wing contribution still remains significant and cannot be ignored. Thus, under these conditions it is crucial to know in detail the distribution for small and intermediate values of S_n.

The factors described above (trap-free regions and compact walks) are both taken explicitly into account in Eq. (2). The large trap-free regions contribute through the term $e^{-\lambda s}$ for large values of s, and the *compact* walks through the small-s wing of the distribution $p_n(s)$. Since it is practically impossible to directly simulate the $p_n(s)$ distribution at its left wing, because of the extremely low probabilities associated with it, in this work we use an indirect method for this calculation.

II. METHOD

In this paper we use an indirect method of calculating the survival probability Φ, i.e., we first calculate the probability distribution $p_n(s)$ of S_n and then multiply it by $e^{-\lambda s}$ [Eq. (2)]. It is fairly straightforward to construct a Monte Carlo simulation in order to compute $p_n(s)$. A particle starts at the origin ($n=0$ and $S_0=1$) and performs a normal random walk on a lattice with all directions having the same probability. One has to simply record the number of sites visited after n steps. Upon repeating the same procedure many times (called different realizations of the walk) we construct a histogram, which when divided by the total number of walks simulated is a quite accurate approximation of $p_n(s)$, especially when the number of realizations is quite large.

This approach works very well, but unfortunately the number of possible walks increases very rapidly with n and an enormous amount of computer resources would be needed even for $n=100$ (in which case there are $4^{100}\sim10^{60}$ possible random walks). In order to overcome this difficulty we use a method that allows us to compute $p_n(s)$ with a high accuracy, especially in the small S_n region. Our method consists of three basic steps.

(1) Describe the shape of the distribution around the maximum via direct computer simulations; (2) calculate exactly the left-wing of the distribution; and (3) interpolate between these two regimes.

Step 1. The first step is quite straightforward. We performed 10^{10} independent realizations of the walk and computed the probability distribution of S_n, as described above. In this way, the minimum value of S_n can be 10^{-10} and due to statistics considerations we can claim an accuracy of about 10^{-9}.

Step 2. The second step is an exact calculation of $p_n(s)$, starting from $s=2$ and increasing s as much as possible. We first count the number $g_n(s)$ of all possible n-step walks resulting in s sites visited. The probability $p_n(s)$ is then simply found by normalizing by the total number of walks 4^n, i.e., $p_n(s)=g_n(s)/4^n$.

The procedure for counting $g_n(s)$ is as follows: we construct a Markov chain starting from the point (0,0). At the first step four branches stem from this point, which correspond to the four neighbors. At step n, every point that was created during the $n-1$ step gives rise to four more branches. This is the standard Markov chain for a normal two-dimensional random walk. However, the number of the

walks increases exponentially and after a few steps there are not enough computer resources to follow the evolution of the system. In order to partially overcome this difficulty we fix a value $s_{\max}$. We continuously monitor the number of sites visited along each path. Whenever a path exceeds the value of $s_{\max}$ we cut this branch and thus all the portion of the tree, which would be generated by it. In this way, we have been able to count $g_n(s)$ up to $s=11$.

In practice, this counting was achieved by implementing the following algorithm. Each path of the chain is represented in the computer memory by a linked list, and each node of the list includes the two coordinates of the corresponding point and a pointer to the next node. Initially, there is only one path including just one node with coordinates (0,0) that of, course represents the origin of the lattice. After a step is made each path gives rise to four more paths that are exactly the same as the original except for the addition of one of the nearest neighbors. In order to improve the efficiency of the algorithm each list includes only once the coordinates of the sites visited. This means that we are only interested to find which sites have been visited until now, which is the definition of S_n, and not for the path itself. We also need to know the current site of the lattice where the particle resides and this information is held separately for each path. A check is also made so that when two lists include exactly the same set of sites and the current site is the same, one of them is destroyed and a variable, denoting how many times this particular set has occurred, is increased. We repeat the same procedure for n steps and count the number of sites included in each path. Whenever the value of S_n corresponding to a particular path surpasses the value $s_{\max}$ we destroy this list and do not deal with it anymore. In this way, we can have the exact distribution $p_n(s)$ for values up to $s=s_{\max}$. Due to the symmetry of the problem (all four directions are equiprobable and the same random walks are repeated in all directions, only rotated 90°) we can follow only one of the four initial branches at the first step and then multiply our result by 4.

In Fig. 1 we present results for the probability distribution $p_n(s)$ after $n=10$, 20, and 50 steps. We compare the Monte Carlo data of 10^{10} realizations to those obtained by the method just described and $s_{\max}=11$. It is obvious that the two data sets are in excellent agreement with each other, verifying thus the validity and proper implementation of the proposed algorithm.

Step 3. The final step for approximating the distribution is to interpolate between the exact small-s region computed in step 2 and the intermediate region, which was derived by the Monte Carlo procedure. This was achieved by using the polynomial interpolation routine in Ref. [21].

III. RESULTS

The method described above gives us the opportunity to study the behavior of $\Phi(n,c)$ for finite times with great accuracy and without having to face the problem of rare trap configurations. The distributions of S_n evaluated via this method are given in Fig. 2 for $n=100$, 200, and 500 steps. The points derived by the interpolation method are connected with the thick line. The right-most part of the distribution has not been calculated and no effort was made towards this direction because this information is not needed in our approach. For example, when $n=500$ and $c=0.1$ the contribution to the final result of $s=150$, where the maximum of the distribution is located, is less than 10^{-10}. The values of s, which are larger than $s=150$, contribute even less to the Φ value since $P_{500}(s)$ in this regime is smaller than $P_{500}(150)$ and can thus be neglected, as it has already been demonstrated at the end of Sec. I. Notice also that the error introduced by the interpolation step for $n=100$ is very small; only two points need to be computed by the interpolation routine and we can be pretty confident about their values. For $n=200$ steps the interpolated region is still narrow enough but for $n=500$ it is rather broad. It is normal that this regime grows as n increases since we cannot in-

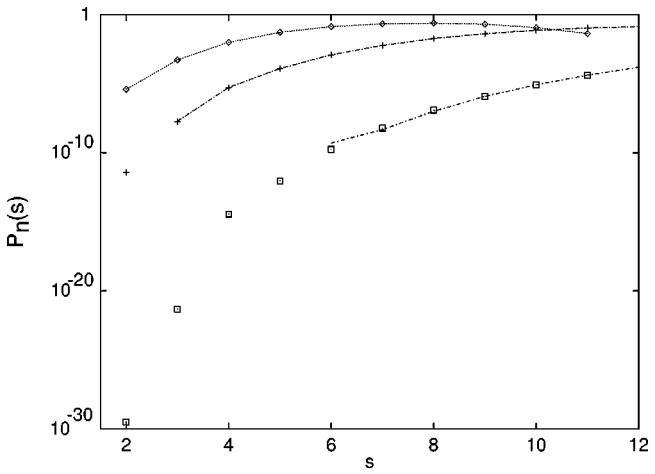

FIG. 1. Comparison of the exact enumeration technique to the Monte Carlo data. Here we plot $p_n(s)$ as a function of s. The lines correspond to the simulation data, while the points represent the exact result. We present the left wing of three distributions for (top to bottom) $n=10$, 20, and 50 steps.

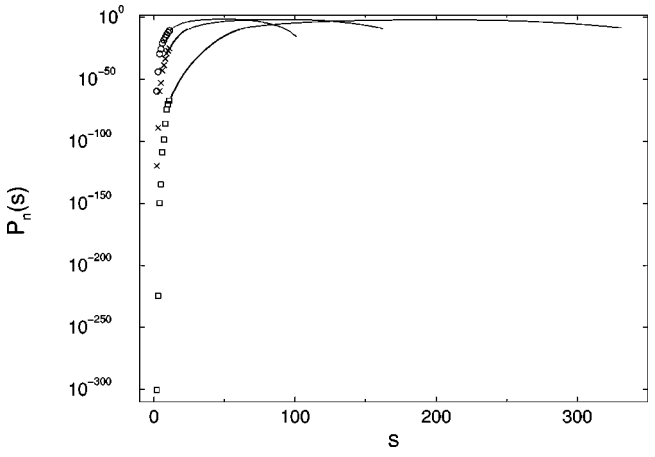

FIG. 2. Distributions of S_n calculated for $n=100$, 200, and 500 steps (top to bottom). The symbols are the results of the exact enumeration technique described in the text; the thin lines are the result of direct Monte Carlo simulations; and the thick lines represent the interpolated region.

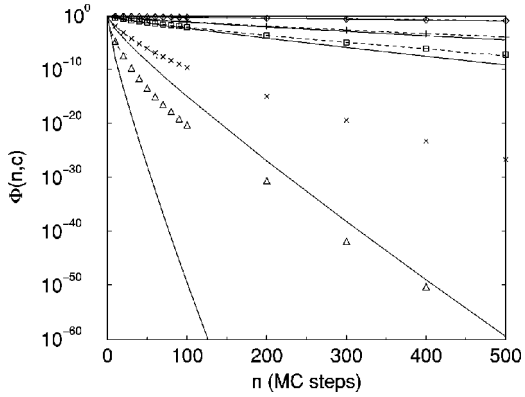

FIG. 3. Comparison of the survival probability $\Phi(n,c)$ calculated by this method (symbols), the Rosenstock approximation (solid line), and the Monte Carlo simulations (dashed line). Trap concentrations (top to bottom) $c=0.01$, 0.05, 0.1, 0.5, and 0.9.

crease $s_{\max}$ further than 11 by using this algorithm. Therefore, we have not tried to go further in time than $n=500$ in order to retain our confidence in the accuracy of the data. From the figure it is seen that the portion of the curve, which is determined exactly, is smoothly connected to the portion that is determined by the Monte Carlo simulations. We believe that we have achieved a good representation of $p_n(s)$. As a simple test, we compared the first and second moments of the distribution and found them to agree with the theoretical results. For example, the mean value $\langle S_n \rangle$ was in complete agreement with the result of Henyey and Seshadri [19], which is considered the most accurate theoretical description.

Now, we can use Eq. (2) and calculate exactly the survival probability for any trap concentration c. We cannot, however, decrease c down to extremely small values because in this case the right wing of the distribution may become important. Results for certain values of c are shown in Fig. 3 and compared to the results of (a) the Rosenstock approximation and (b) direct Monte Carlo simulation of the trapping process. In the latter case we consider a lattice of 100×100 sites and assign a random trap configuration where the probability of a site being a trap is c. We perform a random walk and record the number of steps required before trapping. After a certain number of realizations, 10^9 in our case, we construct a histogram of these values and estimate thus how many random walkers have survived after n steps. We can clearly see from the figure that for small trap concentrations all the curves practically coincide although we can see small deviations for larger n. For large enough trap concentrations these deviations are much more prominent and start earlier in time. We can see that although the number of steps is not that large ($n<500$) the difference of the Rosenstock approximation data from the other two methods (which yield similar results) is important and the mean-field treatment fails. The Monte Carlo data follow the results of the method presented above, but we expect larger deviations when $\Phi(n,c)$ has very small values, as it was shown in [15].

An interesting question after acquiring accurate data is whether we can rescale them according to the Donsker-Varadhan expression. In other words, the question is when

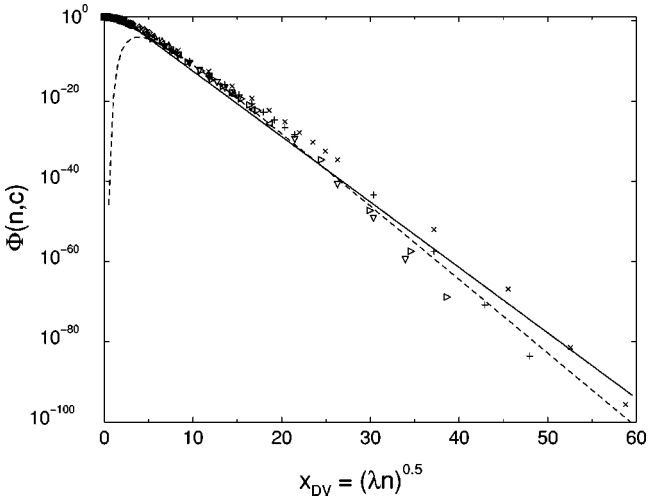

FIG. 4. Plot of the survival probability $\Phi(n,c)$ as a function of $x_{DV}=(\lambda n)^{1/2}$ for different trap concentrations, from $c=0.005$ to $c=0.99$. The solid and the dashed curves represent the two best fit lines of Eq. (5).

we plot Φ as a function of $x_{DV}=(\lambda n)^{1/2}$ whether the points will fall on the same curve and, if they do, which expression can describe this behavior. We try to represent $\Phi(n,c)$ in a form suggested by the asymptotic representation in $d=1$ [7,8]. For this purpose we will use $x_{DV}=(\lambda n)^{1/2}$ as the scaling variable, which is suggested by the Donsker-Varadhan expression. This plot is shown in Fig. 4 and we can see that the data fall roughly on the same curve, but a small dispersion is obvious in the plot and, moreover, this dispersion is systematic, in the sense that smaller concentrations yield slightly lower Φ values.

We used the following form for the fit of the data for $\Phi(x_{DV})$:

$$\Phi(x_{DV})=A\exp\left(-k_2x_{DV}+\frac{a_1}{x_{DV}}\right). \tag{5}$$

We tried two different fits of the data. In the first one, we set $a_1\equiv0$ and took A and k_2 as free parameters. The results of this fit are $A=530.33$ and $k_2=3.756$. In the second fit, we set the value of $k_2\equiv4.26$, which is the precise value of the Donsker-Varadhan prediction, and we took A and a_2 as free parameters. The parameters were found to be $A=1.51\times10^9$ and $a_1=-63.35$. The figure shows that the first fit already describes the data well, and the second fit gives an equally good agreement with the data. The deviation from the data at small values of x are of no concern in view of the form of the ansatz [Eq. (5)], which is valid for large x only. The parameter values for the second fit are rather large and probably result by our fixing of the constant k_2. This may be an indication of the importance of a different prefactor that may also depend on x, or an indication of improper scaling variable. We also took further terms in the exponent of the form a_2/x_{DV}^2 etc. into account, but the fit did only slightly change when they are included.

In Ref. [16], we have indicated by using a scaling ansatz that the proper scaling variable for two dimensions should be

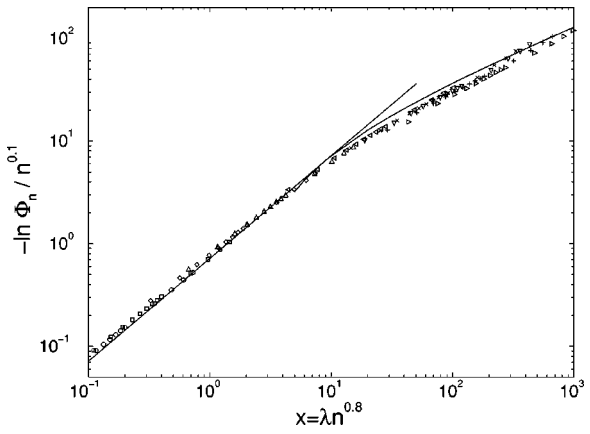

FIG. 5. Log-log plot of $-\ln\Phi(n,c)/n^{0.1}$ as a function of the scaling variable $x=\lambda n^{0.8}$. Different symbols correspond to different trap concentrations from $c=0.005$ up to $c=0.99$. The two lines are the ones predicted by Eq. (6).

$x=\lambda n^{0.8}$, and in this case the mean survival probability can take a form independent of the number of steps n and trap concentration c. This form is reproduced here for convenience:

$$\frac{-\ln\Phi(n,c)}{n^{0.1}}=\begin{cases} Ax, & x\leqslant x_c \\ Ax_c+k_d(x^{1/2}-x_c^{1/2}), & x>x_c \end{cases} \tag{6}$$

where A, k_d, and x_c are numerical constants with the approximate values $A=0.72$, $k_d=4.26$ and $x_c=8.76$. In Fig. 5 we plot the data acquired in this work rescaled according to Eq. (6). By comparing this figure to Fig. 3 of [16] we can see that, although the data were obtained by two completely different methods they obey the same behavior, verifying, thus the validity of the scaling variables used.

The important result of this method is that we can approximate the probability distribution $p_n(s)$ with a high accuracy and compute the survival probability $\Phi(n,c)$, a task that is impossible by a direct method. The attempt for a scaling analysis by using the variable $x_{DV}=(\lambda n)^{1/2}$, which is present in the Donsker-Varadhan expression, yielded a rather satisfactory convergence in a curve, which, however, is not good enough for one to be certain about the verification of the assumption that this is the proper scaling variable. On the contrary, much better accuracy can be obtained when the scaling variable is the one derived by our analysis in [16].

IV. SUMMARY

In this paper we have investigated the trapping problem in $d=2$ by characterizing the survival probability of particles diffusing in the presence of traps as a function of the number of steps and the concentration of the traps. The dimensionality $d=2$ is practically the most difficult to attack in problems of this nature. The theoretical treatment is more demanding than in other dimensions because this is the borderline dimension for recurrent random walks, i.e., for $d=1$ all the walks return at some time at their origin, while for $d=3$ only a certain percentage returns. For instance, the mean number of distinct sites visited by a random walk of n steps $\langle S_n\rangle$ contains logarithmic terms in $d=2$, where approximately [18]

$$\langle S_n\rangle=\frac{n\pi}{\ln(8n)}, \tag{7}$$

while for $d=1$ and $d\geqslant 3$ $\langle S_n\rangle$ follows a power law, and it is proportional to $n^{1/2}$ and n, respectively.

We deduced the survival probability $\Phi(n,c)$ from the probability density $p_n(s)$ of S_n, which was accurately obtained by numerical techniques, and compared it with an asymptotic expansion of the survival probability $\Phi(n,c)$. The comparison with the asymptotic expansion was not entirely satisfactory in that no unique determination of the relative terms was possible. When the same data are rescaled according to another combination of c and n, which was derived by using a completely different route, we can see that the same figure can be reproduced exactly for both methodologies and the data fall on the same curve (Fig. 5), where two characteristic regimes can be easily identified.

Of course, there is no direct practical application in achieving an accuracy for $\Phi(n,c)$ of the order of 10^{-100}. However, in order to correctly characterize the different regimes of behavior, to be able to reach the Donsker-Varadhan regime and to decide for a proper scaling variable it is required to have very accurate data. Direct simulation methods have failed to identify the different regimes, due to the limitations posed by the possible $\Phi(n,c)$ values that can be attained.

[1] F. den Hollander and G.H. Weiss, in *Contemporary Problems in Statistical Physics*, edited by G.H. Weiss (SIAM, Philadelphia, 1994).
[2] G.H. Weiss, *Aspects and Applications of the Random Walk* (North-Holland, Amsterdam, 1994).
[3] G.H. Weiss and R.J. Rubin, Adv. Chem. Phys. **52**, 363 (1983).
[4] M.D. Donsker and S.R.S. Varadhan, Commun. Pure Appl. Math. **28**, 525 (1975); **32**, 721 (1979).
[5] H.B. Rosenstock, J. Math. Phys. **11**, 487 (1970).
[6] G. Zumofen and A. Blumen, Chem. Phys. Lett. **88**, 63 (1982).
[7] J.K. Anlauf, Phys. Rev. Lett. **52**, 1845 (1984).
[8] K.W. Kehr, J.K. Anlauf, and R. Czech, in *Transport and Relaxation in Random Materials*, edited by J. Klafter, R.J. Rubin, and M.F. Shlesinger (World Scientific, Singapore, 1986).
[9] Th.M. Nieuwenhuizen, Phys. Rev. Lett. **62**, 357 (1989).
[10] T.C. Lubensky, Phys. Rev. A **30**, 2657 (1984).
[11] B.Ya. Balagurov and V.G. Vaks, Zh. Eksp. Teor. Fiz. **65** 1939 (1973) [Sov. Phys. JETP **38**, 968 (1974)].
[12] P. Grassberger and I. Procaccia, J. Chem. Phys. **77**, 6281 (1982).
[13] H. Meirovitch, Phys. Rev. A **32**, 3699 (1985).
[14] S. Havlin, M. Dishon, J.E. Kiefer, and G.H. Weiss, Phys. Rev. Lett. **53**, 407 (1984).
[15] A. Bunde, S. Havlin, J. Klafter, G. Graff, and A. Shehter,

Phys. Rev. Lett. **78**, 3338 (1997).
[16] L.K. Gallos, P. Argyrakis, and K.W. Kehr, Phys. Rev. E **63**, 021104 (2001).
[17] G.H. Weiss and R.J. Rubin, J. Stat. Phys. **14**, 333 (1976).
[18] E.W. Montroll and G.H. Weiss, J. Math. Phys. **6**, 167 (1965).
[19] F.S. Henyey and V. Seshadri, J. Chem. Phys. **76**, 5530 (1992).
[20] D.C. Torney, J. Stat. Phys. **44**, 49 (1986).
[21] W.H. Press, B.P. Flannery, S.A. Teukolsky, and W.T. Vetterling, *Numerical Recipes in* C (Cambridge University Press, Cambridge, 1988).