Exact occupation time distribution in a non-Markovian sequence and its relation to spin glass models

Satya N. Majumdar and David S. Dean
Laboratoire de Physique Quantique (UMR C5626 du CNRS), Université Paul Sabatier, 31062 Toulouse Cedex, France
(Received 9 July 2002; published 7 October 2002)

We compute exactly the distribution of the occupation time in a discrete *non-Markovian* toy sequence that appears in various physical contexts such as the diffusion processes and Ising spin glass chains. The non-Markovian property makes the results nontrivial even for this toy sequence. The distribution is shown to have non-Gaussian tails characterized by a nontrivial large deviation function that is computed explicitly. An exact mapping of this sequence to an Ising spin glass chain via a gauge transformation raises an interesting question for a generic finite sized spin glass model; at a given temperature, what is the distribution (over disorder) of the thermally averaged number of spins that are aligned to their local fields? We show that this distribution remains nontrivial even at infinite temperature and can be computed explicitly in few cases such as in the Sherrington-Kirkpatrick model with Gaussian disorder.

DOI: 10.1103/PhysRevE.66.041102 PACS number(s): 05.40.−a, 02.50.−r, 75.50.Lk

I. INTRODUCTION

The occupation time T of a stochastic process $x(t)$ is simply the time that the process spends above its mean value (say 0) when observed over the period $[0,t]$,

$$T=\int_0^t \theta(x(t'))dt', \tag{1}$$

where $\theta(x)$ is the Heaviside step function and we assume, for simplicity, that the process starts at $x(0)=0$. Since the seminal work of Lévy [1] who computed the exact probability distribution of T in the case when $x(t)$ is just an ordinary Brownian motion, there has been a lot of interest in the mathematics community to study the occupation time for various processes [2,3]. Recently the study of the occupation time has seen a revival in the physics community in the context of nonequilibrium systems [4,5] due to its potential applications in a wide range of physical systems which include, amongst others, optical imaging [6], analysis of the morphology of growing surfaces [7] and analysis of the fluorescence intermittency emitting from colloidal semiconductor dots [8].

The occupation time T is clearly a random variable. Its probability distribution $P(T,t)$ evidently depends on the window size t. It turns out that quite generically there are essentially two types of asymptotic behaviors of this distribution $P(T,t)$ depending on whether the underlying stochastic process $x(t)$ is *nonstationary* or *stationary*. A nonstationary process is that where the two-time correlation function $C(t,t')=\langle x(t)x(t')\rangle$ depends on both times t and t'. An example is the ordinary Brownian motion, $dx/dt=\eta(t)$, where $\eta(t)$ is a Gaussian white noise with $\langle \eta(t)\rangle=0$ and $\langle \eta(t)\eta(t')\rangle=\delta(t-t')$. In this case, $C(t,t')=\min(t,t')$. In a stationary process, on the other hand, the two-time correlation function depends only on the time difference, $C(t,t')=C(|t-t'|)$. A simple example of a stationary process is the Orstein-Uhlenbeck process, $dx/dt=-\lambda x+\eta(t)$, where a particle moves in a parabolic potential in presence of external thermal noise. In this case, the particle reaches equilibrium at long times when the two-time correlation simply becomes, $C(t,t')=\exp[-\lambda|t-t'|]$.

In the nonstationary case, one expects that in the asymptotic limit $t\to\infty$, $T\to\infty$ but keeping the ratio $r=T/t$ fixed, the distribution $P(T,t)$ has the generic scaling behavior,

$$P(T,t)\sim \frac{1}{t}f\left(\frac{T}{t}\right), \tag{2}$$

where the scaling function $f(r)$ has nonzero support only in the range $r\in[0,1]$. For example, in the case of ordinary Brownian motion, the scaling function $f(r)$ can be computed exactly [1], $f(r)=1/\pi\sqrt{r(1-r)}$. This is known as the Arc-Sine law of Lévy since the cumulative distribution has an arcsine form, $\int_0^r f(r')dr'=2\sin^{-1}(\sqrt{r})/\pi$. Note that for the Brownian case the scaling actually holds for all t and T. The analytical calculation of this scaling function $f(r)$ is, however, nontrivial even for this simple Brownian case. Following the work of Lévy, there have been various generalizations of this ArcSine law. For example, the scaling function $f(r)$ has been computed exactly for the so called Lévy processes [3], and recently for a more general class of renewal processes [9]. The occupation time distribution has also been studied recently for a Brownian particle moving in a random sinai type potential and the corresponding scaling function $f(r)$ has been computed exactly [13].

For stationary processes, on the other hand, the distribution $P(T,t)$ is expected to have the following generic asymptotic behavior in the appropriate scaling limit $T\to\infty$, $t\to\infty$ with the ratio $r=T/t$ fixed [10],

$$P(T,t)\sim e^{-t\Phi(T/t)}, \tag{3}$$

where $\Phi(r)$ is a large deviation function with, in general, non-Gaussian tails [10]. For example, for the Ornstein-Uhlenbeck stationary process discussed in the previous para-

graph, the function $\Phi(r)$ has recently been computed exactly by utilizing a mapping to a quantum mechanical path integral problem [10].

The calculation of either the scaling function $f(r)$ for nonstationary processes or the large deviation function $\Phi(r)$ for stationary processes is a challenging theoretical problem. So far, exact results exist only for Markov processes, where the value of the process $x(t)$ at time t depends only on its value at the previous time step, say at $t-\Delta t$, where Δt is an infinitesimal time step, but is completely independent on the previous history of the process. For example, the ordinary Brownian motion and the Orstein-Uhlenbeck processes are both Markovian. On the other hand, most processes in Nature are non-Markovian and the Markov processes are more of exceptions rather than rules. Non-Markov processes are known to be notoriously difficult for the analytical calculation of even simpler quantities such as persistence, i.e., the probability that the process does not change sign up to time t [11]. Naturally the analytical calculation of the occupation time distribution for non-Markovian processes is even more difficult.

For a certain class of ''smooth'' non-Markovian processes such as the diffusion equation, it is possible to compute the occupation time distribution [4] using the independent interval approximation (IIA) that assumes that the intervals between successive zero crossings are statistically independent [12]. But these IIA results are only approximate. To our knowledge, there exists no exact result for the occupation time distribution for a non-Markov process, either stationary or nonstationary. In this paper, we obtain an exact analytical result for the occupation time distribution for a stationary non-Markovian process. To be more precise, we actually study the occupation time distribution of a discrete stationary non-Markovian sequence and not a continuous stochastic process. Nevertheless the asymptotic behavior as given by Eq. (3) still remains true and we compute analytically the corresponding large deviation function $\Phi(r)$.

Recently the importance of studying the statistical properties such as the persistence and the distribution of the number of zeros of *discrete stochastic sequences*, as opposed to the more traditional *continuous stochastic processes*, has been emphasized in a number of articles [14–16]. There are two principal reasons for studying a stochastic sequence. First, in various experiments and numerical simulations, even though the underlying physical process is continuous in time, in practice one actually measures the events only at discrete time points. The result of this discretization can lead to subtle and important differences between the ''true'' properties of the process and the ''measured'' properties [14]. To estimate these differences, it is important to study the properties of a discrete sequence. The second reason follows from the observation [15] that many processes in Nature such as weather records are stationary under translations in time only by an integer multiple of a basic period. For example, the seasons repeat typically every one year. For such processes, it was observed in Ref. [15] that the persistence of the underlying continuous process coincides with that of the discrete sequence obtained from the measurement of the process only at times that are integer multiples of the basic period.

In particular, in Ref. [15] a specific discrete sequence was obtained as a limiting case of the diffusion equation on a hierarchical lattice. This rather simple *toy* sequence, even though non-Markovian, had the remarkable property of being solvable for certain statistical properties such as the persistence [15] and the distribution of the number of zeros [16]. Furthermore, these exact results were rather nontrivial [15,16] even for this toy sequence. It is always important to have such a solvable non-Markovian toy model that can then be used as a benchmark to predict the possible expected behaviors of various observables in a more complex non-Markovian process. In this paper, we show that the occupation time distribution can also be computed exactly for this toy model and like other quantities such as the persistence, it is rather nontrivial even for this simple toy model.

We further make an interesting observation that the occupation time for this toy sequence is related to a specific physical observable in an Ising spin glass chain with nearest neighbor interactions. In a given sample of the spin glass chain, one can ask: what is the average (thermal) number of spins that are aligned to the direction of their local fields? This physical object is a random variable that fluctuates from one sample of disorder to another. A natural question is: what is the probability distribution (over disorder) of this thermal average? It turns out that this distribution is nontrivial even at infinite temperature. In fact, we show that at infinite temperature this distribution in the spin glass chain coincides exactly with the occupation time distribution of the toy sequence mentioned above. This connection is useful as it raises a general question for any spin glass model (and not just restricted to a chain): what is the probability distribution of the average (thermal) number of spins that are aligned to their local fields? In this paper, we show that the analytical computation of this distribution in the limit of infinite temperature, though still nontrivial, is tractable in few cases. In particular, we calculate analytically this infinite temperature distribution in the Sherrington-Kirkpatrick (SK) model of mean field spin glasses [17].

The layout of the paper is as follows. In Sec. II, we define the toy sequence, recall some of its properties and known results and then compute the occupation time distribution exactly. In Sec. III, we establish the connection to a spin glass chain and raise the general question regarding the distribution of the average number of spins aligned to their local fields in a generic spin glass model. In Sec. IV, we compute this distribution analytically in the SK model at infinite temperature and show that it is nontrivial even at infinite temperature. Finally we conclude in Sec. V with a summary and some open questions.

II. THE TOY SEQUENCE AND ITS EXACT OCCUPATION TIME DISTRIBUTION

The toy sequence we study in this section was originally derived as a limiting case of the diffusion process on a hierarchical lattice [15]. This is a sequence $\{\psi_i\}$ of correlated random variables constructed via the following rule,

$$\psi_i = \phi_i + \phi_{i-1}, \quad i = 1,2,\ldots,N, \tag{4}$$

where $\phi(i)$'s are independent and identically distributed random variables, each drawn from the same symmetric continuous distribution $\rho(\phi)$. Note that even though $\phi(i)$'s are uncorrelated, the variables ψ_i's are correlated. The two point correlation function, $C_{i,j}=\langle \psi_i \psi_j \rangle$ can be easily computed from Eq. (4),

$$C_{i,j}=\sigma^2[2\,\delta_{i,j}+\delta_{i-1,j}+\delta_{i,j-1}], \tag{5}$$

where $\delta_{i,j}$ is the Kronecker delta function and $\sigma^2 = \int_{-\infty}^{\infty} \phi^2 \rho(\phi) d\phi$ which we assume to be finite. Thus the sequence $\{\psi_i\}$ has only nearest neighbor correlation. Also note that for large sequence size N, the sequence is *stationary* since $C_{i,j}$ depends only on the difference $|i-j|$, and not individually on i or j. The sequence $\{\psi_i\}$ is also *non-Markovian*. To see this, one can try to express a specific member of the sequence, say ψ_i, only in terms of other members of the sequence [15]. This can be easily done using Eq. (4) and one gets for any $i \geqslant 2$,

$$\psi_i=\sum_{k=1}^{i-1} (-1)^{k-1} \psi_{i-k} + \phi_i + (-1)^{i-1} \phi_0 . \tag{6}$$

This relation clearly demonstrates the history dependence of the sequence in the sense that ψ_i depends not just only on the previous member ψ_{i-1} (as would have been in the Markov case), but on the whole history of the sequence preceding ψ_i.

We now turn to the exact computation of the occupation time distribution for the sequence in Eq. (4). The occupation time R in this case is simply the number of ψ_i's that are positive out of the total number N and is given by the discrete counterpart of Eq. (1),

$$R=\sum_{i=1}^{N} \theta(\psi_i). \tag{7}$$

Clearly R is a random variable over the range $0 \leqslant R \leqslant N$. Let us denote its probability distribution by $P(R,N)$ that is formally given by

$$P(R,N)=\int \delta\left[R-\sum_{i=1}^{N} \theta(\phi_{i-1}+\phi_i) \right] \prod_i \rho(\phi_i) d\phi_i . \tag{8}$$

Analogous to the asymptotic behavior in Eq. (3) for continuous stationary processes, we will show that in the appropriate scaling limit $R \to \infty$, $N \to \infty$ but keeping the ratio $r=R/N$ fixed, the distribution $P(R,N)$ has the scaling behavior

$$P(R,N) \sim e^{-N\Phi(R/N)}, \tag{9}$$

where the large deviation function $\Phi(r)$ can be computed analytically. Note also that since $\rho(\phi)$ is symmetric around the origin, the number of positive members of the sequence must have the same distribution as the number of negative members, i.e., $P(R,N)=P(N-R,N)$. Consequently, we must have $\Phi(r)=\Phi(1-r)$, i.e., the large deviation function over the allowed range $0 \leqslant r \leqslant 1$ must be symmetric around $r=1/2$.

To compute the distribution $P(R,N)$ we use a transfer matrix method that has already been used successfully to calculate other quantities for this sequence such as the persistence [15] and the distribution of the number of sign changes [16]. To start with, we define $Q^{\pm}_{R,N}(\phi_0)$ denoting respectively the joint probability that the first member of the sequence ψ_1 is positive (negative) and that the sequence of size N has a total R number of positive members, given the value of ϕ_0. Let us also define $Q_{R,N}(\phi_0)=Q^{+}_{R,N}(\phi_0) + Q^{-}_{R,N}(\phi_0)$ that denotes the probability of having R positive members in a sequence of size N, given ϕ_0. The required occupation time distribution is then given by,

$$P(R,N)=\int_{-\infty}^{\infty} Q_{R,N}(\phi_0) \rho(\phi_0) d\phi_0 . \tag{10}$$

The reason for this small detour is simply that one can write quite easily a recursion relation for the joint probabilities $Q^{\pm}_{R,N}(\phi_0)$. However it is not easy to write a recursion directly for the distribution $P(R,N)$. The probabilities $Q^{\pm}_{R,N}(\phi_0)$ satisfy the following recursion relations:

$$Q^{+}_{R,N}(\phi_0)=\int_{-\phi_0}^{\infty} d\phi_1 \rho(\phi_1) Q_{R-1,N-1}(\phi_1)$$

$$Q^{-}_{R,N}(\phi_0)=\int_{-\infty}^{-\phi_0} d\phi_1 \rho(\phi_1) Q_{R,N-1}(\phi_1). \tag{11}$$

The above recursion relations are valid for all $0 \leqslant R \leqslant N$ and $N \geqslant 1$ with the initial conditions $Q^{+}_{0,0}(\phi_0)=1$ and $Q^{-}_{0,0}(\phi_0)=0$.

These recursion relations in Eq. (11) are easy to follow. Consider first the relation for $Q^{+}_{R,N}(\phi_0)$. In order for the first member ψ_1 to be positive, it follows from the definition, $\psi_1=\phi_1+\phi_0$, that $\phi_1 > -\phi_0$ for a given ϕ_0. This explains the integration range on the right hand side of the first line in Eq. (11). Also once the first member is positive, in order to have a total R positive members, we need to ensure that the rest of the chain of size $N-1$ (excluding the first member) has exactly $R-1$ positive members. The probability of this latter event, for a given ϕ_1, is simply $Q_{R,N}(\phi_1)$. This explains the integrand on the right hand side of Eq. (11). The second line of Eq. (11) can be understood following a similar line of reasoning. Note that the recursion relations in Eq. (11) also satisfy the one sided boundary conditions, $Q^{+}_{R,N}(-\infty)=0$ and $Q^{-}_{R,N}(\infty)=0$. The first condition follows from the fact that if $\phi_0 \to -\infty$, then the first member of the sequence $\psi_1=\phi_1+\phi_0$ can be positive only with a vanishing probability. On the other hand if $\phi_0 \to \infty$, then ψ_1 can be negative only with probability zero thus giving rise to the second condition. Note however that the values at the other boundaries, namely, $Q^{+}_{R,N}(\infty)$ and $Q^{-}_{R,N}(-\infty)$ are unspecified.

We define the generating functions as follows:

$$\tilde{Q}_N^{\pm}(\phi_0,y)=\sum_{R=0}^{\infty} Q_{R,N}^{\pm}(\phi_0)y^R, \tag{12}$$

with the understanding that $Q_{R,N}^{\pm}(\phi_0)=0$ for $R>N$ since R can take values only in the range $0\leqslant R\leqslant N$. We also define $\tilde{Q}_N(\phi_0,y)=\tilde{Q}_N^{+}(\phi_0,y)+\tilde{Q}_N^{-}(\phi_0,y)$. Using Eq. (11), it is easy to see that the generating functions satisfy the recursions

$$\tilde{Q}_N^{+}(\phi_0,y)=y\int_{-\phi_0}^{\infty} d\phi_1\rho(\phi_1)\tilde{Q}_{N-1}(\phi_1,y)$$

$$\tilde{Q}_N^{-}(\phi_0,y)=\int_{-\infty}^{-\phi_0} d\phi_1\rho(\phi_1)\tilde{Q}_{N-1}(\phi_1,y), \tag{13}$$

with the boundary conditions $\tilde{Q}_N^{+}(-\infty,y)=0$ and $\tilde{Q}_N^{-}(\infty,y)=0$ for all $y\geqslant 0$. These generating functions also satisfy the condition $\tilde{Q}_0(\phi_0,y)=1$ for all y. The following step is to differentiate the recursion relations in Eq. (13) with respect to ϕ_0 which gives

$$\frac{\partial\tilde{Q}_N^{+}(\phi_0,y)}{\partial\phi_0}=y\rho(-\phi_0)\tilde{Q}_{N-1}(-\phi_0,y)$$

$$\frac{\partial\tilde{Q}_N^{-}(\phi_0,y)}{\partial\phi_0}=-\rho(-\phi_0)\tilde{Q}_{N-1}(-\phi_0,y). \tag{14}$$

Further simplifications can be made by using the symmetry $\rho(-\phi_0)=\rho(\phi_0)$ and by making a change of variable from ϕ_0 to $u(\phi_0)=\int_0^{\phi_0}\rho(\phi)d\phi$. Note that since $\rho(\phi)$ is symmetric around the origin, $\phi_0\to-\phi_0$ corresponds to $u\to-u$. Thus $u(\phi_0)$ is a monotonic function of ϕ_0. Note further that as $\phi_0\to\pm\infty$, $u\to\pm 1/2$, where we have again used the fact that $\rho(\phi_0)$ is symmetric around the origin. Let us also write, $\tilde{Q}_N^{\pm}(\phi_0,y)=S_N^{\pm}(u,y)$ and $\tilde{Q}_N(\phi_0,y)=S_N(u,y)$, where $S_N(u,y)=S_N^{+}(u,y)+S_N^{-}(u,y)$. Then the relations in Eq. (14) simplify to

$$\frac{\partial S_N^{+}(u,y)}{\partial u}=yS_{N-1}(-u,y)$$

$$\frac{\partial S_N^{-}(u,y)}{\partial u}=-S_{N-1}(-u,y), \tag{15}$$

which are valid over $-1/2\leqslant u\leqslant 1/2$. In terms of the variable u, the boundary conditions $\tilde{Q}_N^{+}(-\infty,y)=0$ and $\tilde{Q}_N^{-}(\infty,y)=0$ translate to $S_N^{+}(-1/2,y)=0$ and $S_N^{-}(1/2,y)=0$ for all $y\geqslant 0$. Note also the interesting fact that the distribution $\rho(\phi)$ has completely disappeared in Eq. (15). The consequence of this, as we will see later, is that occupation time distribution $P(R,N)$ is completely universal, i.e., independent of the distribution $\rho(\phi)$ as long as it is symmetric and continuous.

The recursion relations in Eq. (15), though much simplified, are still nontrivial since they are nonlocal in u. We employ the following technique of separation of variables, $S_N^{\pm}(u,y)=\lambda^{-N}f^{\pm}(u)$, where we have suppressed the y dependence for convenience of notations. Substituting this form in Eq. (15), we get a nonlocal eigenvalue equation

$$\frac{df^{+}}{du}=y\lambda[f^{+}(-u)+f^{-}(-u)]$$

$$\frac{df^{-}}{du}=-\lambda[f^{+}(-u)+f^{-}(-u)], \tag{16}$$

where the eigenvalue λ is yet to be determined. We also have the boundary conditions, $f^{+}(-1/2)=0$ and $f^{-}(1/2)=0$. It is easy to see from Eq. (16) that the sum, $f(u)=f^{+}(u)+f^{-}(u)$ satisfies the nonlocal first order equation, $f'(u)=-\omega f(-u)$, where $\omega=\lambda(1-y)$. Differentiating this equation once more, we get a local second order equation, $f''(u)=-\omega^2 f(u)$ whose most general solution is given by $f(u)=A\cos(\omega u)+B\sin(\omega u)$, where A, B are arbitrary constants. One further notices that this general solution will also satisfy the first order non-local equation $f'(u)=-\omega f(-u)$ provided $B=-A$. Thus we arrive at the solution, $f(u)=A[\cos(\omega u)-\sin(\omega u)]$. Substituting this solution on the right hand side of the first line in Eq. (16) and solving the resulting equation using the boundary condition $f^{+}(-1/2)=0$, we get

$$f^{+}(u)=\frac{A\lambda y}{\omega}[\sin(\omega u)-\cos(\omega u)+\sin(\omega/2)+\cos(\omega/2)]. \tag{17}$$

The other function $f^{-}(u)$ then follows from the relation, $f^{-}(u)=f(u)-f^{+}(u)$, where $f(u)=A[\cos(\omega u)-\sin(\omega u)]$ and $f^{+}(u)$ is given by Eq. (17). The function $f^{-}(u)$ still has to satisfy the boundary condition $f^{-}(1/2)=0$. In fact, this condition determines the eigenvalue λ and we get $\tan(\omega/2)=(1-y)/(1+y)$, where $\omega=\lambda(1-y)$. For large N, only the smallest eigenvalue λ will dominate which is given by

$$\lambda=\frac{2}{(1-y)}\tan^{-1}\left(\frac{1-y}{1+y}\right). \tag{18}$$

Using the exact $\lambda(y)$ from Eq. (18), we are now ready to compute the large N behavior of the occupation time distribution $P(R,N)$. In Eq. (10), after making a change of variable $\phi_0\to u$, we find the generating function, $\Sigma_M P(R,N)y^R=\int_{-1/2}^{1/2}S_N(u,y)du$. We substitute the large N behavior $S_N(u,y)\approx\lambda^{-N}f(u)$ and carry out the integration using the exact expression of $f(u)$ to obtain the following exact large N result,

$$\sum_{M=0}^{\infty} P(R,N)y^R\approx\frac{2A}{\omega}\sin(\omega/2)[\lambda(y)]^{-N}, \tag{19}$$

where $\lambda(y)$ is given by Eq. (18). By inverting the generating function and carrying out a standard steepest decent analysis for large N, large R but keeping the ratio $r=R/N$ fixed, we get the desired result, $P(R,N)\sim\exp[-N\Phi(R/N)]$, where the large deviation function $\Phi(r)$ is given by the exact formula

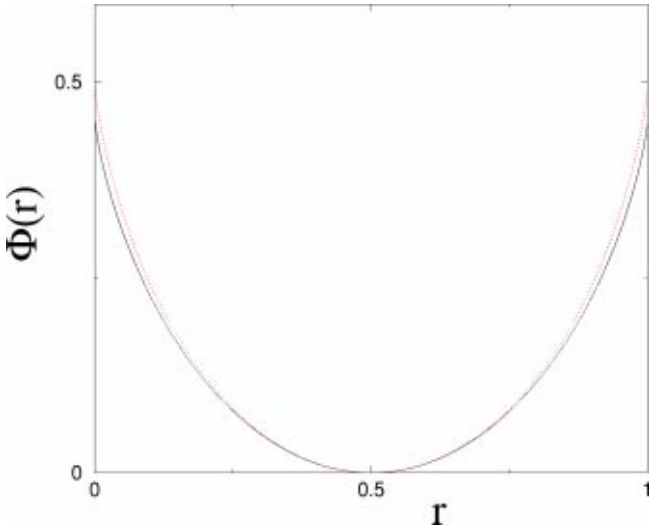

FIG. 1. The large deviation function $\Phi(r)$ plotted against r. The solid line corresponds to the large deviation function for the 1-d sequence and is obtained from Eq. (20) using Mathematica. The dotted line corresponds to that of the SK model obtained using Mathematica in Eq. (35) after the shift $\Phi(r)=\Theta(2r-1)$.

$$\Phi(r)=\max_{y}\left\{\ln\left[\frac{2y^{r}}{(1-y)}\tan^{-1}\left(\frac{1-y}{1+y}\right)\right]\right\}. \tag{20}$$

We first note that the function $Y(y,r)=2y^{r}\tan^{-1}[(1-y)/(1+y)]/(1-y)$ inside the ''ln'' in Eq. (20) is invariant under the transformation $y\to 1/y$ and $r\to 1-r$, i.e., $Y(y,r)=Y(1/y,1-r)$. This obviously indicates that $\Phi(r)=\Phi(1-r)$ as expected. Determining $\Phi(r)$ in closed form seems difficult, though it can be obtained quite trivially using Mathematica, as displayed by the solid line in Fig. 1.

It is however easy and instructive to obtain analytical expressions of $\Phi(r)$ in the regimes near $r=0$ and $r=1/2$. It turns out that these limits correspond, respectively, to $y\to 0$ and $y\to 1$ in the function $Y(y,r)$. Keeping r fixed we expand $Y(y,r)$ for small y and near $y\to 1$ in a Taylor series and then take the logarithm and maximize to obtain the following limiting behaviors:

$$\Phi(r)=\ln\left(\frac{\pi}{2}\right)+r\ln\left(\frac{r\pi}{(4-\pi)e}\right)+\cdots r\to 0$$
$$=\frac{6}{5}\left(r-\frac{1}{2}\right)^{2}+\cdots, r\to 1/2. \tag{21}$$

These limiting forms have interesting physical implications. Consider the limit $r\to 0$ or equivalently $R\to 0$. Note that $P(0,N)=P(N,N)\sim\exp[-\Phi(0)N]$ is just the probability that all the members are either negative or positive up to length N which is precisely the persistence of the sequence. The persistence for this sequence was earlier computed in Ref. [15] and it was found to decay for large N as $\exp[-\theta N]$ with the persistence exponent $\theta=\ln(\pi/2)$. Thus the limiting form of $\Phi(r)$ as $r\to 0$ in Eq. (21) is consistent with the persistence exponent, $\theta=\Phi(0)=\ln(\pi/2)$.

The other limit $r\to 1/2$ is also interesting and can be derived independently from a central limit theorem. To see this we find from Eq. (7) that $R-\langle R\rangle=\Sigma_{i=1}^{N}(x_i-\langle x_i\rangle)$, where $\langle R\rangle=N/2$, $x_i=\theta(\psi_i)$, and $\langle x_i\rangle=1/2$. In general the summands $(x_i-\langle x_i\rangle)$ are of course highly correlated and one cannot employ the central limit theorem to evaluate the sum. However, one can do so in the limit when $M\to\langle M\rangle$ when the variables $(x_i-\langle x_i\rangle)$ become only weakly correlated. Then the central limit theorem predicts a Gaussian distribution for the sum, $P(R,N)\sim\exp[-(R-N/2)^2/2\sigma_N^2]$, where $\sigma_N^2=\langle(R-N/2)^2\rangle$ is the variance. One can calculate this variance independently by computing the correlation functions $\langle(x_i-1/2)(x_j-1/2)\rangle$, where $x_i=\theta(\psi_i)$. It is shown in the Appendix that for large N, $\sigma_N^2=5N/12$. Hence the central limit theorem predicts that in the limit $R\to 1/2$, $P(R,N)\sim\exp[-6N(r-1/2)^2/5]$ thus yielding exactly the same limiting form of $\Phi(r)$ for $r\to 1/2$ as in Eq. (21).

Thus the occupation time distribution $P(R,N)$, though Gaussian near the mean value $R=N/2$, becomes non-Gaussian as $r=R/N$ deviates away from its mean and approaches the tails $r\to 0$ or $r\to 1$. This crossover from Gaussian behavior near $r=1/2$ to non-Gaussian behavior near $r\to 0,1$ is characterized by the large deviation function $\Phi(r)$ changing from a quadratic function near $r=1/2$ to nonquadratic behavior near $r\to 0,1$ as given by Eq. (21). We conclude this section by noting the important fact that the large deviation function $\Phi(r)$ in Eq. (20) and in fact the full occupation time distribution $P(R,N)$ is completely universal, i.e., independent of the distribution $\rho(\phi)$ [as long as $\rho(\phi)$ is symmetric and continuous]. Moreover, this universality holds for any arbitrary N and not just asymptotically for large N.

III. RELATION TO SPIN GLASS MODELS

We start this section by raising a physical question for a general spin glass model defined on a finite lattice of N sites: What is the distribution (over disorder) of the thermally averaged number of spins that are aligned to their local fields? This distribution depends on the temperature and on the system size N. It turns out that the distribution remains nontrivial even in the infinite temperature limit. In fact, for a nearest neighbor Ising spin glass chain, we show that this infinite temperature limiting distribution is precisely that of the occupation time distribution of the toy sequence computed in the previous section. The infinite temperature limit, though nontrivial, is tractable in few other cases such as the SK model of Ising spin glass, which will be discussed in detail in the following section.

Consider a spin glass model on a lattice of N sites defined by the Hamiltonian,

$$E=-\sum_{\langle i,j\rangle}J_{i,j}S_iS_j, \tag{22}$$

where S_i's are the spin variables (not necessarily Ising) and $J_{i,j}$ denotes the coupling between site i and site j. In the nearest neighbor model, the sum in Eq. (22) runs over nearest neighbor pairs. On the other hand, for long range mean

field models such as the SK model, the sum runs over all pairs of sites. The variables $J_{i,j}$'s are independent and each is drawn from the identical distribution $\rho(J)$, which we assume to be symmetric and continuous. Henceforth we will use the short hand notation $\vec{J}$ and $\vec{S}$ to denote, respectively, the set of couplings and the set of spins. Thus the J's have the joint distribution $Q[\vec{J}]d\vec{J}=\Pi_{i,j}\rho(J_{i,j})dJ_{i,j}$. The local field that a spin at site i sees is simply $h_i=\Sigma_j J_{i,j}S_j$. If the spin gets aligned to its local field, we must have $h_iS_i>0$. Hence the total number of spins $N_a[\vec{J},\vec{S}]$ in a given configuration that are aligned to their local fields can be formally written as,

$$N_a[\vec{J},\vec{S}]=\sum_i \theta[h_iS_i]=\sum_i \theta\left[S_i\sum_j J_{i,j}S_j\right]. \quad (23)$$

Evidently N_a is a random variable that depends on the couplings $\vec{J}$, as well as the spins $\vec{S}$. Let us first compute the thermal average of N_a over the spin configurations for a fixed quenched disorder $\vec{J}$,

$$\overline{N}_a(\vec{J})=\frac{1}{Z}\sum_{\vec{S}} N_a[\vec{J},\vec{S}]e^{-\beta E(\vec{S})}, \quad (24)$$

where $Z=\Sigma_{\vec{S}}e^{-\beta E(\vec{S})}$ is the partition function and β is the inverse temperature. This thermal average $\overline{N}_a(\vec{J})$ is a random variable that varies from one realization of disorder to another. We then ask: what is the probability distribution of this random variable (over disorder) at a given inverse temperature β? This probability distribution $\mathrm{Prob}(\overline{N}_a=R)=P_\beta(R,N)$ can be formally represented as

$$P_\beta(R,N)=\int \delta[R-\overline{N}_a(\vec{J})]Q[\vec{J}]d\vec{J}. \quad (25)$$

The analytical calculation of $P_\beta(R,N)$ at arbitrary β seems difficult. Let us, therefore, consider a simpler limit, namely, the limit of infinite temperature or equivalently $\beta\to 0$. In this limit, the thermal average in Eq. (24) becomes simple, $\overline{N}_a(\vec{J})=\Sigma_{\vec{S}}N_a[\vec{J},\vec{S}]/N_C$, where N_C is the total number of spin configurations. Thus all spin configurations are equally likely. However, the distribution $P_0(R,N)$ as in Eq. (25), even in this infinite temperature limit, is still nontrivial.

Let us now focus on Ising spins, where $S_i=\pm 1$. Here $N_C=2^N$, where N is total number of lattice sites. The Eq. (25), using Eq. (23), then becomes simpler for the Ising case, $P_0(R,N)=\int\delta[R-(1/2^N)\Sigma_{\vec{S}}\Sigma_{i=1}^N\theta(\Sigma_j J_{i,j}S_iS_j)]Q[\vec{J}]d\vec{J}$. The next step is to make a gauge transformation, $\phi_{i,j}=J_{i,j}S_iS_j$. Since the spins are Ising, i.e., $S_i=\pm 1$, $\phi_{i,j}$'s have the same distribution as the $J_{i,j}$'s. The advantage of this gauge transformation is that one can then do away with the configuration sum over the spins and we simply get

$$P_0(R,N)=\int \delta\left[R-\sum_{i=1}^N \theta\left(\sum_{j\neq i}\phi_{i,j}\right)\right]Q[\vec{\phi}]d\vec{\phi}. \quad (26)$$

Now consider the special case of a nearest neighbor Ising spin glass chain of size N, where $E=-\Sigma_i J_{i,i+1}S_iS_{i+1}$ with free boundary conditions. Various properties of this spin glass chain such as the statistics of the number of metastable states have been studied analytically by Li [18] and by Derrida and Gardner [19]. Recently it was also shown that the persistence in the toy sequence studied in this paper is the same as the average fraction of metastable spins in the Ising chain [15]. In the present context, we find the distribution $P_0(R,N)$ in Eq. (26) reduces to

$$P_0(R,N)=\int \delta\left[R-\sum_{i=1}^N \theta(\phi_{i-1}+\phi_i)\right]Q[\vec{\phi}]d\vec{\phi}. \quad (27)$$

Comparing Eqs. (27) and (8) one immediately finds $P_0(R,N)=P(R,N)$, where $P(R,N)$ is precisely the occupation time distribution that was computed in Sec. II. This thus establishes the promised link between the spin glass problem discussed in this section and the non-Markovian toy sequence discussed in Sec. II. The infinite temperature distribution (over disorder) of the thermally averaged number of locally aligned spins is identical to that of the occupation time distribution of the toy sequence discussed in Sec. II. From the exact results of $P(R,N)$ derived in Sec. II, one therefore knows the distribution $P_0(R,N)$ exactly as well.

A question naturally arises: Are there other solvable cases for $P_0(R,N)$ apart from the 1-d chain? In the following section we show that indeed the infinite range SK model is one such case where the distribution $P_0(R,N)$ can be computed analytically.

IV. THE SK MODEL

In this section we calculate the infinite temperature distribution $P_0(R,N)$ of the thermally averaged number of locally aligned spins in the infinite range SK model defined by the Hamiltonian in Eq. (22), where $\langle i,j\rangle$ runs over all pairs of the total number of N sites. The couplings $J_{i,j}$'s are independent of each other and we assume that each is drawn from a Gaussian distribution, $\rho(J)=\sqrt{N/2\pi}e^{-NJ^2/2}$. The choice $J\sim N^{-1/2}$ is necessary to ensure that the free energy is extensive in the large N limit. It is clear from Eq. (26) that if we define $\psi_i=\Sigma_{j\neq i}\phi_{i,j}$, where each of the $\phi_{i,j}$'s are independent Gaussian variables with the distribution $\rho(\phi)=\sqrt{N/2\pi}e^{-N\phi^2/2}$, then $R=\Sigma_{i=1}^N\theta(\psi_i)$. It turns out that for technical reasons it is easier to consider the variable $M=\Sigma_{i=1}^N\mathrm{sgn}(\psi_i)$, where $\mathrm{sgn}(x)=2\theta(x)-1$. Hence $M=2R-N$. In what follows we will compute the distribution $P_0(M,N)$ and derive the corresponding distribution of R using the simple shift $M=2R-N$.

Since we are eventually interested in the limit $N\to\infty$, $M\to\infty$ but keeping the ratio $m=M/N$ fixed, we set $M=mN$ and write

$$P_0(m,N)=\left\langle \delta\left(\sum_{i=1}^N \mathrm{sgn}(\psi_i)-mN\right)\right\rangle_\psi$$
$$=\int_{-\infty}^{\infty}\frac{d\mu}{2\pi}e^{-i\mu mN}\left\langle e^{i\mu\sum_{i=1}^N \mathrm{sgn}(\psi_i)}\right\rangle_\psi, \quad (28)$$

where we have used the representation of the delta function, $\delta(x)=\int_{-\infty}^{\infty}e^{i\mu x}d\mu/2\pi$ and $\langle\rangle_{\psi}$ denotes the expectation over the distributions of ψ_i's. Using the identity, $\langle e^{i\mu\ \mathrm{sgn}(y)}\rangle_y = \Sigma_{\sigma=-1,1}e^{i\mu\sigma}\langle\theta(y\sigma)\rangle_y$, one can rewrite Eq. (28) as

$$P_0(m,N)=\int_{-\infty}^{\infty}\frac{d\mu}{2\pi}e^{-i\mu mN}\sum_{\{\sigma_i=-1,1\}}\left\langle\prod_i e^{i\mu\sigma_i}\theta(\psi_i\sigma_i)\right\rangle_{\psi}. \tag{29}$$

We use the representation, $\theta(x)=\int_0^{\infty}dl\int_{-\infty}^{\infty}(d\lambda/2\pi)e^{i\lambda(x-l)}$ in Eq. (29), make the transformation $\lambda_i\sigma_i\to\lambda_i$ and $\mu\to-\mu$ and then sum over the σ_i variables to obtain

$$P_0(m,N)=\int_{-\infty}^{\infty}\frac{d\mu}{2\pi}e^{i\mu mN}\prod_i\left\langle\left[\int_0^{\infty}dl_i 2\cos(\mu+\lambda_i l_i)\times\int_{-\infty}^{\infty}\frac{d\lambda_i}{2\pi}e^{i\lambda_i\psi_i}\right]\right\rangle_{\psi}. \tag{30}$$

The following step is to evaluate the expectation value $\langle\Pi_i e^{i\lambda_i\psi_i}\rangle_{\psi}$. Using $\psi_i=\Sigma_{j\neq i}\phi_{i,j}$, we note that $\Pi_i e^{i\lambda_i\psi_i}=\Pi_{i<j}e^{i(\lambda_i+\lambda_j)\phi_{i,j}}$. Using the Gaussian distribution $\rho(\phi_{i,j})$, one can easily evaluate the expectation value to finally obtain, $\langle\Pi_i e^{i\lambda_i\psi_i}\rangle_{\psi}=\exp[-\Sigma_{i,j}(\lambda_i+\lambda_j)^2/4N]$. We expand the sum $\Sigma_{ij}(\lambda_i+\lambda_j)^2=2N\Sigma_i\lambda_i^2+2(\Sigma_i\lambda_i)^2$ and use a Hubbard-Stratonovich transformation $\exp[-(\Sigma_i\lambda_i)^2/2N]=\sqrt{(N/2\pi)}\int_{-\infty}^{\infty}dz e^{iz\Sigma_i\lambda_i-Nz^2/2}$ to finally write the expected value of the product as

$$\left\langle\prod_i e^{i\lambda_i\psi_i}\right\rangle_{\psi}=\sqrt{\frac{N}{2\pi}}\int_{-\infty}^{\infty}dz e^{-Nz^2/2+iz\sum_i\lambda_i-\sum_i\lambda_i^2/2}. \tag{31}$$

We then substitute Eq. (31) in Eq. (30) and carry out the Gaussian integrations over the variables l_i's and λ_i's to get

$$P_0(m,N)=\sqrt{\frac{N}{2\pi}}\int_{-\infty}^{\infty}\frac{d\mu}{2\pi}e^{i\mu mN}\int_{-\infty}^{\infty}dz e^{-Nz^2/2}[A(z,\mu)]^N, \tag{32}$$

where the function $A(z,\mu)$ is given by

$$A(z,\mu)=\frac{1}{2}\left[e^{i\mu}\mathrm{erfc}\left(\frac{z}{\sqrt{2}}\right)+e^{-i\mu}\mathrm{erfc}\left(-\frac{z}{\sqrt{2}}\right)\right], \tag{33}$$

with $\mathrm{erfc}(x)=(2/\sqrt{\pi})\int_x^{\infty}e^{-u^2}du$ being the complementary error function.

We expand the right hand side of Eq. (33) in a binomial series, substitute the resulting series in Eq. (32) and carry out the integration with respect to μ to obtain

$$P_0(m,N)=\frac{1}{2^N}\sqrt{\frac{N}{2\pi}}\int_{-\infty}^{\infty}dz e^{-Nz^2/2}\binom{N}{(1-m)N/2}\times[\mathrm{erfc}(z/\sqrt{2})]^{(1-m)N/2}[\mathrm{erfc}(-z/\sqrt{2})]^{(1+m)N/2}. \tag{34}$$

Keeping m fixed we then use the Stirling's formula to approximate the combinatorial factor in Eq. (34) for large N and then use the steepest descent method to evaluate the integral in Eq. (34) for large N. This gives, ignoring pre-exponential factors, a similar asymptotic behavior as in the 1-d case, $P_0(m,N)\sim e^{-N\Theta(m)}$, where the large deviation function $\Theta(m)$ in this case is given by

$$\Theta(m)=\frac{1}{2}\min_z\left[z^2+\sum_{\sigma=-1,1}(1-m\sigma)\ln\left(\frac{(1-m\sigma)}{2\mathrm{erfc}(z\sigma/\sqrt{2})}\right)\right]. \tag{35}$$

In terms of the original variable $r=(1+m)/2$, the distribution is then given by $P_0(r,N)=P_0(m=2r-1,N)/2\sim\exp[-N\Phi(r)]$ with $\Phi(r)=\Theta(2r-1)$, where $\Theta(x)$ is given exactly by Eq. (35). The function $\Phi(r)$ is symmetric around $r=1/2$ since $\Theta(m)$ in Eq. (35) is symmetric around $m=0$. As in the 1-d case, it seems difficult to obtain a closed form expression of $\Phi(r)$. However, it can be easily evaluated from Eq. (35) using Mathematica as displayed by the dotted line in Fig. 1. Moreover, similar to the 1-d case, one can evaluate $\Phi(r)$ analytically near $r=1/2$, as well as near the tail regions $r\to0,1$. Omitting the details of algebra, we find

$$\Phi(r)=a+r\ln\left(\frac{br}{e}\right)+\cdots r\to0$$
$$=\frac{2\pi}{\pi+2}\left(r-\frac{1}{2}\right)^2+\cdots r\to1/2, \tag{36}$$

where $a=\ln(2)+z_0^2/2-\ln[\mathrm{erfc}(z_0/\sqrt{2})]$, $b=\mathrm{erfc}(z_0/\sqrt{2})/\mathrm{erfc}(-z_0/\sqrt{2})$, and z_0 is the root of the equation

$$z_0+\sqrt{\frac{2}{\pi}}\frac{e^{-z_0^2/2}}{\mathrm{erfc}(z_0/\sqrt{2})}=0. \tag{37}$$

Solving Eq. (37) numerically yields $z_0=-0.506\,054\ldots$ which gives $a=0.493\,919\ldots$ and $b=2.263\,61\ldots$.

Thus the limiting behaviors of $\Phi(r)$ near $r=1/2$ and $r\to0,1$ in the SK model in Eq. (36) are qualitatively similar to those in the 1-d case in Eq. (21). Using similar arguments as in Sec. I, one can understand the behavior near $r=1/2$ as a consequence of a central limit theorem that predicts a Gaussian behavior for $P_0(R,N)\sim\exp[-(R-N/2)^2/2\sigma_N^2]$, where $\sigma_N^2=\langle(R-N/2)^2\rangle$ is the variance. Comparing with Eq. (36) we find that the variance for large N is given exactly by, $\sigma_N^2=(\pi+2)N/4\pi$. This result for the variance can also be derived by a direct method as shown in the Appendix, thus providing an additional consistency check. The results for the statistics of R in the 1-d case and in the SK model can be jointly summarized as: The mean is always given by, $\langle R\rangle=N/2$ and the variance for large N is given by

$$\sigma_N^2=\frac{5}{12}N,\qquad 1\text{-}d$$
$$=\frac{\pi+2}{4\pi}N,\quad \mathrm{SK}. \tag{38}$$

The behavior in the tail region near $r \to 0$ (or equivalently near $r \to 1$) is also interesting. Let us, for simplicity, consider the $r \to 1$ limit where $R = N$. It is clear from the expression of $P_0(N,N)$ in Eq. (26) that the delta function will contribute only when each of the θ function inside the sum are satisfied, i.e., if $\Pi_i \theta(\Sigma_{j \neq i} \phi_{i,j}) = 1$. Thus evidently $P_0(N,N) = \langle \Pi_i \theta(\Sigma_{j \neq i} \phi_{i,j}) \rangle_\phi$. But this quantity is just the average number of *metastable* configurations in the spin glass. To see this clearly, consider again the spin glass Hamiltonian in Eq. (22). A spin configuration is called metastable if the energy required to flip any of the N spins is strictly positive. In other words, all the spins must be aligned to their local fields in a metastable configuration. Hence, for a fixed disorder, the fraction of metastable configurations (out of the total number of 2^N spin configurations) is given by, $f(\vec{J}) = 2^{-N} \Sigma_{\vec{S}} \Pi_i \theta(h_i s_i)$ where h_i's are the local fields. Finally, the average (over disorder) fraction of the metastable configurations is given by, $\langle f(\vec{J}) \rangle_J = \int f(\vec{J}) Q[\vec{J}] d\vec{J}$. Using once again the gauge transformation for Ising spins, $\phi_{i,j} = J_{i,j} S_i S_j$ one can easily express this average fraction as $\langle f(\vec{J}) \rangle_J = \langle \Pi_i \theta(\Sigma_{j \neq i} \phi_{i,j}) \rangle_\phi = P_0(N,N)$. Our results on the large deviation function near $r = 0,1$ in Eq. (36) indicates that $P_0(N,N) = P_0(0,N) \sim e^{-aN}$ for large N with $a = 0.493\,919 \ldots$. On the other hand, the average number of metastable configurations in the SK model was computed long ago by Tanaka and Edwards [20] and also by Bray and Moore [21] and this average is known to increase exponentially for large N as $\sim e^{\alpha N}$, where $\alpha = 0.1992$ [21]. Hence the average fraction scales as $\sim e^{\alpha N}/2^N = e^{-cN}$ with $c = \ln 2 - \alpha = 0.4939$ [21]. Thus the constant a in Eq. (36) is precisely the same as the constant c and hence the limiting behavior of our large deviation function near the tails $r = 0,1$ is completely consistent with the calculation of average number of metastable states.

Let us conclude this section with the following comment. In the case of the 1-d toy sequence, we found in Sec. II that the full occupation time distribution $P(R,N)$ and consequently the associated large deviation function is completely independent of the distribution $\rho(\phi)$. In the case of the SK model, we have derived the large deviation function for a specific form of the disorder distribution, namely the Gaussian form. Naturally the question arises as to how universal is this large deviation function as one changes the disorder distribution. Evidently for finite N the results in the SK case, unlike the 1-d case, will depend on the details of the distribution $\rho(J)$. However, due to the $1/\sqrt{N}$ scaling in the definition of the distribution of the $J_{i,j}$, the large N results including the large deviation are universal (upto rescaling by a constant factor), provided the variance of the $J_{i,j}$'s is finite. In the case of mean field spin glasses with power law or Lévy interactions [25], the variance of the $J_{i,j}$ is no longer finite and it would be interesting to study the occupation time distribution in this context.

V. SUMMARY AND CONCLUSION

The three main points of this paper are: (i) we have been able to derive exact results for the occupation time distribution of a non-Markovian process. In our case, the stochastic process is not continuous in time, but rather a discrete toy sequence. Nevertheless this toy sequence retains the non-Markovian property, which makes the results nontrivial. Besides the fact that exact results are always useful and instructive, this toy sequence also appears in various physical contexts such as diffusion process and spin glasses, thus extending the range of applications of our results. (ii) We also established an exact mapping of this sequence to an Ising spin glass chain using a gauge transformation. The occupation time distribution in the sequence then translates, via this mapping, into the distribution of the thermally averaged number of spins that are aligned to their local fields in the spin glass chain at infinite temperature when all spin configurations are equally likely. This observation raises an interesting question for any generic finite sized spin glass model: at a given temperature, what is the distribution of the thermally averaged number of locally aligned spins? Our exact results in one dimension show that this distribution remains nontrivial even at infinite temperature. (iii) We then were able to compute analytically this infinite temperature distribution in the SK model of spin glasses with Gaussian disorder and argued that for very large N, the associated large deviation function is again universal, i.e., independent of the precise form of the disorder distribution.

We leave open the possibility of computing this distribution at a finite temperature for any spin glass model. For example, it would be interesting to know how this distribution changes as one goes below the spin glass transition temperature.

The study of the number of metastable spins in various other spin glass models is an open question. We mention a few cases, where exact results along these lines may be possible as the average number of metastable states is calculable: the SK model in the presence of external fields [22], p-spin spin glass models [23], spin glasses on random graphs [24], mean field spin glasses with Lévy interactions [25], the Hopfield neural network model [26] and the Random Orthogonal Model [27]. The study of spin glass models on random graphs of fixed connectivity c are of particular interest as they interpolate between the one dimensional toy model studied here, at $c = 2$, and the SK model in the limit $c \to \infty$.

APPENDIX: DIRECT CALCULATION OF THE VARIANCE OF M

In this appendix we compute the variances of the occupation time both in the one dimensional toy sequence and in the SK model by a more direct method. These results are identical to those obtained from the limiting forms of the large deviation functions near $r = 1/2$.

We have in general,

$$\sigma_N^2 = \left\langle \left[\sum_i \left(\theta(\psi_i) - \frac{1}{2} \right)^2 \right] \right\rangle$$
$$= \frac{N}{4} + 2 \sum_{i<j} \left(\langle \theta(\psi_i)\, \theta(\psi_j) \rangle - \frac{1}{4} \right)$$
$$= \frac{N}{4} + \frac{1}{2} \sum_{i<j} \langle \mathrm{sgn}(\psi_i) \mathrm{sgn}(\psi_j), \qquad \text{(A1)}$$

where we have made use of the identities $\langle \theta(\psi_i)\rangle = 1/2$ for all i and $\mathrm{sgn}(x) = 2\theta(x) - 1$.

In the one dimensional model only neighboring sites are correlated and hence one has

$$\sigma_N^2 = \frac{N}{4} + 2\sum_i \left(\langle \theta(\psi_i)\theta(\psi_{i+1})\rangle - \frac{1}{4} \right)$$

$$= -\frac{N}{4} + 2N\langle \theta(\psi_1)\theta(\psi_2)\rangle, \tag{A2}$$

where we have used the isotropy of the sites in the large N limit. One now has that

$$\langle \theta(\psi_1)\theta(\psi_2)\rangle = \int d\phi_0 d\phi_1 d\phi_2 \rho(\phi_0)\rho(\phi_1)\rho(\phi_2) \times \theta(\phi_0+\phi_1)\theta(\phi_1+\phi_2)$$

$$= \int_{-\infty}^{\infty} d\phi_0 \rho(\phi_0) \int_{-\phi_0}^{\infty} d\phi_1 \rho(\phi_1) \times \int_{-\phi_1}^{\infty} d\phi_2 \rho(\phi_2). \tag{A3}$$

We introduce the function $F(\phi) = \int_{-\phi}^{\infty} d\phi' \rho(\phi')$ and use the relations $\rho(\phi) = \rho(-\phi)$ and $dF/d\phi = \rho(\phi)$ to carry out the integration and thus obtain

$$\langle \theta(\psi_1)\theta(\psi_2)\rangle = \frac{1}{3} \tag{A4}$$

Putting this altogether gives the large N asymptotic result $\sigma_N^2 = 5N/12$. In fact using the generating function technique used in this paper, and hence taking into account the boundary terms exactly, one can show that $\sigma_N^2 = 5N/12 - 1/6$ for any N. Note that this result, for arbitrary N, is independent of the precise form of the distribution $\rho(\phi)$.

We now turn to the SK model where $\psi_i = \Sigma_{j\neq i}\phi_{i,j}$ and the random variables $\phi_{i,j}$'s are independent and identically distributed with the Gaussian distribution $\rho(\phi) = \sqrt{N/2\pi}e^{-N\phi^2/2}$. Clearly, one has $\langle \psi_i^2\rangle = (1 - 1/N)$ and $\langle \psi_i\psi_j\rangle = 1/N$. We use the well known identity that holds only for Gaussian random variables as follows:

$$\langle \mathrm{sgn}(X)\mathrm{sgn}(Y)\rangle = \frac{2}{\pi}\sin^{-1}\left(\frac{\langle XY\rangle}{\sqrt{\langle X^2\rangle\langle Y^2\rangle}} \right). \tag{A5}$$

Using this identity we get for $i \neq j$,

$$\langle \mathrm{sgn}(\psi_i)\mathrm{sgn}(\psi_j)\rangle = \frac{2}{\pi}\sin^{-1}\left(\frac{1}{N-1} \right). \tag{A6}$$

This yields

$$\sigma_N^2 = \frac{N}{4} + \frac{1}{2\pi}N(N-1)\,\sin^{-1}\left(\frac{1}{N-1} \right) \tag{A7}$$

which gives the result $\sigma_N^2 = N(\pi+2)/4\pi$ in the limit of large N. We note that the result in Eq. (A7) for finite N is valid only when the distribution of the $\phi_{i,j}$ is Gaussian. The finite N result for arbitrary distribution of the $\phi_{i,j}$ will depend in general on the details of the distribution and hence, in contrary to what happens in the one dimensional toy model, will not be universal. However, as argued in Sec. IV, the large N results including the result for the variance, i.e., $\sigma_N^2 = N(\pi + 2)/4\pi$ is universal as long as the variance of the $\phi_{i,j}$ is finite.

[1] P. Lévy, Compos. Math. **7**, 238 (1939).

[2] M. Kac, Trans. Am. Math. Soc. **65**, 1 (1949); D.A. Darling and M. Kac, Trans. Am. Math. Soc. **84**, 444 (1957); For a recent review of the occupation time in the mathematics literature see, S. Watanabe, Proc. Symp. Pure Math. **57**, 157 (1995). See, also, M. Yor, *Some Aspects of Brownian Motion* (Birkhauser, Zurich, 1992), Vol. 1.

[3] J. Lamperti, Trans. Am. Math. Soc. **88**, 380 (1958).

[4] I. Dornic and C. Godrèche, J. Phys. A **31**, 5413 (1998); T.J. Newman and Z. Toroczkai, Phys. Rev. E **58**, R2685 (1998).

[5] A. Dhar and S.N. Majumdar, Phys. Rev. E **59**, 6413 (1999); I. Dornic, A. Lamaitre, A. Baldassari, and H. Chaté, J. Phys. A **33**, 7499 (2000); T.J. Newman and W. Loinaz, Phys. Rev. Lett. **86**, 2712 (2001); G. De Smedt, C. Godrèche, and J.M. Luck, J. Phys. A **34**, 1247 (2001).

[6] G.H. Weiss and P.P. Calabrese, Physica A **234**, 443 (1996).

[7] Z. Toroczkai, T.J. Newman, and S. Das Sarma, Phys. Rev. E **60**, R1115 (1999).

[8] G. Messin, J.P. Hermier, E. Giacobino, P. Desbiolles, and M. Dahan (unpublished).

[9] C. Godrèche and J.M. Luck, J. Stat. Phys. **104**, 489 (2001).

[10] S.N. Majumdar and A.J. Bray, Phys. Rev. E **65**, 051112 (2002).

[11] For a recent review on persistence, see, S.N. Majumdar, Curr. Sci. **77**, 370 (1999).

[12] S.N. Majumdar, C. Sire, A.J. Bray, and S.J. Cornell, Phys. Rev. Lett. **77**, 2867 (1996); B. Derrida, V. Hakim, and R. Zeitak, *ibid.* **77**, 2871 (1996).

[13] S.N. Majumdar and A. Comtet, Phys. Rev. Lett. **89**, 060601 (2002).

[14] S.N. Majumdar, A.J. Bray, and G.C.M.A. Ehrhardt, Phys. Rev. E **64**, 015101 (2001); G.C.M.A. Ehrhardt and A.J. Bray, Phys. Rev. Lett. **88**, 070601 (2002); G.C.M.A. Ehrhardt, A.J. Bray, and S.N. Majumdar, Phys. Rev. E **65**, 041102 (2002).

[15] S.N. Majumdar and D. Dhar, Phys. Rev. E **64**, 046123 (2001).

[16] S.N. Majumdar, Phys. Rev. E **65**, 035104 (2002).

[17] D. Sherrington and S. Kirkpatrick, Phys. Rev. Lett. **35**, 1792 (1975).

[18] T. Li, Phys. Rev. B **24**, 6579 (1981).

[19] B. Derrida and E. Gardner, J. Phys. (Paris) **47**, 959 (1986).

[20] F. Tanaka and S.F. Edwards, J. Phys. F: Met. Phys. **13**, 2769 (1980).

[21] A.J. Bray and M.A. Moore, J. Phys. C **13**, L469 (1980); see, also, *ibid.* **14**, 1313 (1981).

[22] S.A. Roberts, J. Phys. C **14**, 3015 (1981); D.S. Dean, J. Phys. A **27**, L899 (1994).
[23] V.M. de Oliviera and J.F. Fontanari, J. Phys. A **30**, 8445 (1997).
[24] D.S. Dean, Eur. Phys. J. B **15**, 493 (2000); A. Lefèvre and D.S. Dean, *ibid.* **21**, 121 (2001).
[25] P. Cizeau and J.-P. Bouchaud, J. Phys. A **26**, L187 (1993).
[26] E. Gardner, J. Phys. A **19**, L1047 (1986); A. Treves and D.J. Amit, *ibid.* **21**, 3155 (1988).
[27] G. Parisi and M. Potters, J. Phys. A **28**, 5267 (1995).