Unbiased diffusion of Brownian particles on disordered correlated potentials

Raúl Salgado-Garcia

E-mail: `raulsg@uaem.mx`
Centro de Investigación en Ciencias, Universidad Autónoma del Estado de Morelos. Avenida Universidad 1001, colonia Chamilpa, C.P. 62209, Cuernavaca Morelos, Mexico.

Cesar Maldonado

Centro de Modelamiento Matemático, Universidad de Chile, Beauchef 851, Edificio Norte, Piso 7, Santiago, Chile.

Abstract.
    In this work we study the diffusion of non-interacting overdamped particles, moving on unbiased disordered correlated potentials, subjected to Gaussian white noise. We obtain an exact expression for the diffusion coefficient which allows us to prove that the unbiased diffusion of overdamped particles on a random polymer does not depend on the correlations of the disordered potentials. This universal behavior of the unbiased diffusivity is a direct consequence of the validity of the Einstein relation and the decay of correlations of the random polymer. We test the independence on correlations of the diffusion coefficient for correlated polymers produced by two different stochastic processes, a one-step Markov chain and the expansion-modification system. Within the accuracy of our simulations, we found that the numerically obtained diffusion coefficient for these systems agree with the analytically calculated ones, confirming our predictions.

1. Introduction

The diffusion of particles on one-dimensional (1D) random media has been the subject of intense research. Since the introduction of the Sinai's model it was recognized that disordered systems can exhibit normal and anomalous diffusion as well as normal and anomalous drift [1]. Several models of diffusion in disordered media have been introduced since then [2–4] in order to understand the underlying mechanisms leading to such behaviors. Particularly some simple deterministic models for transport in disordered potentials have been studied in order to understand the origin of the transport properties from the very deterministic dynamics [5–9]. A model for transport in random polymers that takes into account the pair correlations between monomers (to which we will refer to as the "particle-polymer model") has been introduced by the authors in Ref. [9]. In such a work it was shown that the deterministic biased diffusion is influenced by the presence of the correlations in the polymer. Indeed, in that paper, the authors gave an exact expression for the particle current and the diffusion coefficient in terms of the first two moments of the "crossing times" and its corresponding pair correlation function [9], a result which is a consequence of the central limit theorem [10–12]. Based on the same arguments, it has been shown that when the particles along the polymer are placed at a finite temperature and subjected to a constant driving force, it is still possible to obtain an exact expression for the drift and diffusion coefficients in terms of the first and the second moments of the first passage time (FPT) as well as its pair correlation function [13]. In this work we will show that, for the particle-polymer model, the diffusion coefficient does not depend on the correlations between monomers along the polymer when the driving force is zero. The latter is a result that follows straightforwardly from the validity of the Einstein relation in disordered potentials at zero driving force.

In consequence the paper is organized as follows. In Section 2 we introduce the model as well as some general results concerning the transport process previously established. Then in Section 3 we give the exact expression for the diffusion coefficient in disordered polymers for zero driving force and introduce a simple model for the potential to illustrate the independence of the diffusivity on the correlations. First we use a Markov chain to produce exponentially correlated polymers which are used to induce a correlated potential. Then, we perform Langevin dynamics simulations to estimate the diffusion coefficient of overdamped particles moving on such polymers. We also perform analogous numerical experiments in the case in which the polymers are produced by means of the expansion-modification system, which is known to have polynomially decaying correlations. Finally in Section 4 we give a brief discussion and the main conclusions of our work. Two appendices are given containing detailed calculations of our main results.

2. The model

We will consider an ensemble of Brownian particles with overdamped dynamics moving on a 1D disordered potential $V(x)$ subjected to an external force F. The equation of motion of any of these particles is given by the stochastic differential equation,

$$\gamma dX_t = (f(X_t) + F)dt + \varrho_0 dW_t \qquad (1)$$

where X_t stands for the position of the particle and W_t is a standard Wiener process. The constants F and γ stand for the driving force and the friction coefficient

respectively. The constant ϱ_0 represents the square root of the noise intensity (or the thermal fluctuations), which according to the fluctuation-dissipation theorem satisfy the relation $\varrho_0^2 = 2\gamma\beta^{-1}$. Here β denotes, as usual, the inverse temperature times the Boltzmann constant, $\beta = 1/k_B T$. The function $f(x)$ represents minus the gradient of the potential $V(x)$ that the particle feels due to its interaction with the substrate where the motion occurs. As in Ref. [13], we model the substrate (or the polymer) as a chain of "unit cells" of constant length L. The unit cells represent the monomers comprising the polymer. We denote by $\mathcal{A}$ the set of possible monomer types, a set that is assumed to be finite or countable infinite. The polymer is represented by a bi-infinite symbolic sequence $\mathbf{a} := (\ldots, a_{-1}, a_0, a_1, \ldots)$, where $a_j \in \mathcal{A}$ stands for the monomer type located at the j-th cell, for all $j \in \mathbb{Z}$. The set of possible random polymers will be denoted by $\mathcal{A}^{\mathbb{Z}}$ according to the conventional notation in symbolic dynamics [14]. As in Refs. [9, 13], we assume that the disordered potential $V(x)$ is the result of the interaction of a particle with the random polymer.

Let $x \in \mathbb{R}$ denote the particle position along the substrate $\mathbf{a} \in \mathcal{A}^{\mathbb{Z}}$. A particle located at x feels a (random) potential that is not only function the particle position x, but also of the substrate, i.e. $V(x) = \psi(x, \mathbf{a})$. This is a consequence of the fact that the particle might interact with the whole polymer and not only with the closest monomer. If $x = 0$ we assume that the particle is located at the beginning of the 0-th monomer a_0. In Ref. [13] it was shown that the random potential $\psi(x, \mathbf{a})$ has the following property,

$$\psi(x + nL, \mathbf{a}) = \psi[x, \sigma^n(\mathbf{a})]. \tag{2}$$

where $\sigma : \mathcal{A}^{\mathbb{Z}} \to \mathcal{A}^{\mathbb{Z}}$ is the *shift mapping* defined as follows: if $\mathbf{a}, \mathbf{b} \in \mathcal{A}^{\mathbb{Z}}$ are such that $\sigma(\mathbf{a}) = \mathbf{b}$, then $b_i = a_{i+1}$ for all $i \in \mathbb{Z}$. As in Refs. [9, 13] we assume that the substrate is generated by some stochastic process through a stationary σ-invariant probability measure ν on $\mathcal{A}^{\mathbb{Z}}$. This hypothesis is equivalent to say that the statistical properties of the polymer are translationally invariant.

In Ref. [13] it has been shown that the particle flux J_{eff} and the diffusion coefficient D_{eff} can be written exactly as follows,

$$J_{\text{eff}} := \lim_{t \to \infty} \frac{\langle\langle X_t \rangle\rangle}{t} = \frac{L}{\langle\langle \tau(0 \to L) \rangle\rangle}, \tag{3}$$

$$D_{\text{eff}} := \lim_{t \to \infty} \frac{\text{Var}(X_t)}{2t} = \frac{L^2 \varrho_\tau^2}{2\langle\langle \tau(0 \to L) \rangle\rangle^3}. \tag{4}$$

where $\tau(0 \to L)$ is a random variable, called the *first passage time* (FPT), defined as the time that the particle spends to reach, for the first time, the position $X_\tau = L$ from the initial condition $X_0 = 0$. The quantity ϱ_τ^2 is defined in terms of the variance of the FPT and its autocorrelation function as,

$$\varrho_\tau^2 := \langle\langle \tau^2(0 \to L) \rangle\rangle - \langle\langle \tau(0 \to L) \rangle\rangle^2 + 2\sum_{m=1}^{\infty} C_\tau(m).$$

where

$$C_\tau(m) := \langle\langle \tau(0 \to L)\tau\big(mL \to (m+1)L\big) \rangle\rangle - \langle\langle \tau(0 \to L) \rangle\rangle^2.$$

It is important to stress that the notation $\langle\langle \cdot \rangle\rangle$ denotes a double average, one with respect to the noise and the other with respect to the disorder. In Ref. [13] it was pointed out that is important to distinguish between these two different averages. If we put an ensemble of non-interacting Brownian particles over a (fixed) random

polymer $\mathbf{a}$ we denote the average over the ensemble of particles as $\langle\cdot\rangle_\mathrm{n}$. This average will be referred to as the *average with respect to the noise* (w.r.n.). Once a certain observable has been averaged with respect to the noise, it still depends on the specific realization of the polymer $\mathbf{a}$. Thus we need to perform a second average which should be carried out over an ensemble of different realizations of the random polymer. This average is taken with respect to a stationary measure ν which characterizes the process producing the random polymers. This average will be referred to as the *average with respect to the polymers* (w.r.p.) and will be denoted by $\langle\cdot\rangle_\mathrm{p}$. Then, the double average $\langle\langle\cdot\rangle\rangle$ is actually achieved first by taking the average w.r.n. and after that by taking the average w.r.p., i.e., $\langle\langle\cdot\rangle\rangle = \langle\langle\cdot\rangle_\mathrm{n}\rangle_\mathrm{p}$. Additionally, we will use the following notation: $\mathrm{Var}_\mathrm{n}(\mathcal{O}) := \langle\mathcal{O}^2\rangle_\mathrm{n} - \langle\mathcal{O}\rangle_\mathrm{n}^2$ and $\mathrm{Var}_\mathrm{p}(\mathcal{O}) := \langle\mathcal{O}^2\rangle_\mathrm{p} - \langle\mathcal{O}\rangle_\mathrm{p}^2$ denote the variance of the observable $\mathcal{O}$ w.r.n. and w.r.p. respectively. Along this line, $\mathrm{Var}(\mathcal{O})$ therefore denote the variance of the observable $\mathcal{O}$ with respect to both, the noise and the polymer ensemble, i.e., $\mathrm{Var}(\mathcal{O}) := \langle\langle\mathcal{O}^2\rangle\rangle - \langle\langle\mathcal{O}\rangle\rangle^2$.

Let us denote by $T_1(\mathbf{a})$ the mean FPT w.r.n., i.e., $T_1(\mathbf{a}) := \langle\tau(0 \to L)\rangle_\mathrm{n}$. Then the drift and diffusion coefficients can be written as [13],

$$J_\mathrm{eff} = \frac{L}{\langle T_1\rangle_\mathrm{p}}, \tag{5}$$

$$D_\mathrm{eff} = D_\mathrm{noisy} + D_\mathrm{det}, \tag{6}$$

where D_noisy and D_det are referred to as the *noisy and deterministic parts* of D_eff respectively. These quantities are defined as follows

$$D_\mathrm{noisy} = \frac{L^2\langle\mathrm{Var}_\mathrm{n}(\tau(0 \to L))\rangle_\mathrm{p}}{2\langle T_1\rangle_\mathrm{p}^3}, \tag{7}$$

$$D_\mathrm{det} = \frac{L^2\mathrm{Var}_\mathrm{p}(T_1) + 2L^2\sum_{m=1}^\infty C_\tau(m)}{2\langle T_1\rangle_\mathrm{p}^3}. \tag{8}$$

The autocorrelation of the FPT can also be rewritten in terms of T_1, which gives

$$C_\tau(\ell) = \langle T_1(\mathbf{a})T_1\left[\sigma^\ell(\mathbf{a})\right]\rangle_\mathrm{p} - \langle T_1(\mathbf{a})\rangle_\mathrm{p}^2. \tag{9}$$

The reason for expressing the transport coefficients, J_eff and D_eff, in terms of the mean and the variance (with respect to the noise) of the FPT is that such quantities can be written exactly in terms of quadratures. Indeed, in Ref. [13] it was proved that the mean FPT w.r.n., T_1, can be written as,

$$T_1(\mathbf{a}) = \gamma\beta \sum_{m=1}^\infty e^{-m\beta FL}q_+(\mathbf{a})q_-[\sigma^{-m}(\mathbf{a})] + \gamma\beta I_0(\mathbf{a}), \tag{10}$$

and the variance of the FPT, $\mathrm{Var}_\mathrm{n}[\tau(0 \to L)]$, as

$$
\begin{aligned}
\mathrm{Var}_\mathrm{n}[\tau(0 \to L)] &= 2(\gamma\beta)^2\Bigg\{ \sum_{n=1}^\infty\sum_{m=1}^\infty\sum_{l=1}^\infty \Big(e^{-(n+m+l)\beta FL}q_+(\mathbf{a})q_+[\sigma^{-n}(\mathbf{a})] \\
&\quad \times\ q_-[\sigma^{-m-n}(\mathbf{a})]q_-[\sigma^{-n-l}(\mathbf{a})]\Big) \\
&\quad + 2\sum_{n=1}^\infty\sum_{m=1}^\infty e^{-(n+m)\beta FL}q_+(\mathbf{a})q_-[\sigma^{-n-m}(\mathbf{a})]I_0[\sigma^{-n}(\mathbf{a})] \\
&\quad + \sum_{n=1}^\infty e^{-n\beta FL}q_+(\mathbf{a})I_1[\sigma^{-n}(\mathbf{a})]
\end{aligned}
$$

$$+ \sum_{m=1}^{\infty} \sum_{l=1}^{\infty} e^{-(m+l)\beta FL} q_{-}[\sigma^{-m}(\mathbf{a})] q_{-}[\sigma^{-l}(\mathbf{a})] I_2(\mathbf{a})$$

$$+ 2 \sum_{m=1}^{\infty} e^{-m\beta FL} q_{-}[\sigma^{-m}(\mathbf{a})] I_3(\mathbf{a}) + I_4(\mathbf{a}) \Bigg\} \tag{11}$$

which defines completely the transport coefficients. Here, the functions $q_{+}, q_{-} : \mathcal{A}^{\mathbb{Z}} \to \mathbb{R}$ were defined as

$$q_{\pm}(\mathbf{a}) := \int_0^L dx \exp\big(\pm \beta[\psi(x,\mathbf{a}) - xF] \big), \tag{12}$$

and the functions $B_{\pm} : \mathbb{R} \times \mathcal{A}^{\mathbb{Z}} \to \mathbb{R}$ and $Q_{\pm} : \mathbb{R} \times \mathcal{A}^{\mathbb{Z}} \to \mathbb{R}$ were defined as,

$$Q_{\pm}(x,\mathbf{a}) := \int_0^x dy \exp\big(\pm \beta[\psi(y,\mathbf{a}) - yF] \big), \tag{13}$$

$$B_{\pm}(x,\mathbf{a}) := \exp\big(\pm \beta[\psi(x,\mathbf{a}) - xF] \big). \tag{14}$$

We also used the following short-hand notation for the involved integrals,

$$I_0(\mathbf{a}) = \int_0^L Q_{-}(x,\mathbf{a}) B_{+}(x,\mathbf{a}) dx,$$

$$I_1(\mathbf{a}) = \int_0^L [Q_{-}(x,\mathbf{a})]^2 B_{+}(x,\mathbf{a}) dx,$$

$$I_2(\mathbf{a}) = \int_0^L Q_{+}(x,\mathbf{a}) B_{+}(x,\mathbf{a}) dx,$$

$$I_3(\mathbf{a}) = \int_0^L B_{+}(x,\mathbf{a}) \int_0^x Q_{-}(u,\mathbf{a}) B_{+}(u,\mathbf{a}) du dx,$$

$$I_4(\mathbf{a}) = \int_0^L B_{+}(x,\mathbf{a}) \int_0^x B_{+}(u,\mathbf{a}) [Q_{-}(u,\mathbf{a})]^2 du dx. \tag{15}$$

In Ref. [13] it was shown that, for the case of non-correlated random polymers, the transport of particles in this system arises an interesting phenomenon which is the non-monotonic behavior of the diffusivity on the temperature. In this work we will study the effective diffusion coefficient of overdamped particles moving on correlated and uncorrelated potentials. Bellow we will show with certain generality that the unbiased diffusion does not depend on the correlations of the polymer beyond the number of monomers that the particle interacts with.

3. Unbiased diffusion on random polymers

A remarkable property that can be inferred from our formalism, is that the diffusivity in unbiased disordered potentials does not depend on the correlations of the substrate. This means that the particles in disordered potentials are unable to feel the pair correlations between monomers comprising the polymer. Consequently, the transport properties of a system of particles diffusing on correlated polymers are equal to those of a system of particles moving on an "equivalent" ensemble of uncorrelated polymers. Recently [15] this kind of behavior has been shown to occur in a similar model of diffusion on random potentials. In Ref. [15] Goychuk and Kharchenko calculated the diffusion coefficient of overdamped particles diffusing on a correlated random potential with a Gaussian distribution. Surprisingly their result does not depend on

the correlations, a conclusion that is reached by using the Einstein relation. Here we prove that for our model of random potentials the diffusion coefficient is independent on the correlations if the substrate is *ergodic* and has decay of correlations. The word ergodic in this context means that the polymer average can be taken as a time average if we think of the polymer as a discrete-time stochastic process.

Bouchaud and Georges in Ref. [4] have shown that the Einstein relation, $\mu = \beta D_{\text{eff}}(F = 0)$, remains valid for the class of disordered potentials we are considering. The constant μ stands for the mobility at zero bias, $\mu = \lim_{F \to 0} J(F)/F$. From the Einstein relation it is easy to obtain the diffusion coefficient since

$$D_{\text{eff}}(F = 0) = \frac{1}{\beta} \lim_{F \to 0} \frac{J_{\text{eff}}(F)}{F}.$$

In Appendix A we prove that the first moment of the FPT the divergence occurs as F^{-1}. This implies that the particle current vanishes, as expected, in the limit $F \to 0$. Actually the particle current behaves linearly in a small neighborhood of $F = 0$. This allows us to obtain the effective diffusion coefficient as follows

$$D_{\text{eff}}(F = 0) = \frac{L^2}{\gamma \beta Z_+ Z_-}. \tag{16}$$

Here the constants $Z_\pm$ are defined as

$$Z_\pm := \langle q_\pm(\mathbf{a}) \rangle_{\text{p}} = \left\langle \int_0^L \exp\left[\pm \beta \psi(x, \mathbf{a})\right] dx \right\rangle_{\text{p}}. \tag{17}$$

Notice that the effective diffusion coefficient depends only on the quantities $Z_\pm$ which are *local averages*, in the sense that the polymers involved in the averages are not shifted with respect to each other (as it occurs in the correlation function given in Eq. (9)). Thus, the expression (16) for the effective diffusion coefficient implies that, on *ergodic* polymers with decay of correlations, the diffusivity does not depend on the correlations between monomers at distant sites if the particle only interacts with the closest monomer to it (i.e., if the potential $\psi(x, \mathbf{a})$ only depends on one "coordinate" of $\mathbf{a}$). This means that particles on polymers with long range correlations exhibit the same diffusivity as the particles that moves on polymers without correlations but with the same one-dimensional marginals. More generally, the diffusivity of a particle moving on a random polymer will be influenced by the pair correlations between monomers as distant as the *range* of the potential induced by the interaction. Here, the *range* of a potential stands for the maximal distance between two monomers having a non vanishing interaction with the particle. This makes the averages with respect to the polymer ensemble to depend on the range of the interaction. In other words, if the range of the potential $\psi(x, \mathbf{a})$ is n, then, ψ can be seen as a function of n coordinates of $\mathbf{a}$, i.e., $\psi(x, \mathbf{a}) = \psi(y, a_j, a_{j+1}, \ldots, a_{j+n-1})$. Therefore, to perform the polymer averages $Z_\pm$ it is necessary to have at least the n-dimensional marginals of the probability measure ν defining the polymer ensemble i.e., we need to know the joint probability distribution of n monomers, $\mathbb{P}(a_j, a_{j+1}, \ldots, a_{j+n-1})$. With exception of the Bernoulli measure, the n-marginals cannot be decomposed, in general, as the product of the 1-dimensional marginals. In the following we give some examples to better illustrate these situations.

3.1. Potential of range $n = 1$

In order to exemplify the independence on the correlations of the diffusion coefficient at zero tilt strength we use the potential model introduced in Ref. [13]. The potential model $V(x)$ used in such a work is of range 1 since the potential only depends on the monomer at which the particle is located. Explicitly the potential is defined as follows. If x represents the particle position and we write as $x = y + nL$, then

$$V(x) = \psi(x, a_n) = \begin{cases} a_n y & \text{if} \quad 0 \leq y < L/2 \\ a_n(L - y) & \text{if} \quad L/2 \leq y < L. \end{cases} \tag{18}$$

Here $n \in \mathbb{Z}$ stands for the index of the unit cell at which the particle is located and y is the relative position of the particle along such unit cell. The varible a_n is a random variable representing the monomer type at the nth unit cell. This potential model is shown schematically in Fig. 1. We observe that such potential is symmetric on every unit cell with a maximum located at $y = 1/2$. The height of the potential assumed to be random taking values from a finite set. Since the height of the potential is given by $a_n L/2$ we can assume that a_n represents the random variable, for every $n \in \mathbb{Z}$, which can take values from a set $\mathcal{A} := \{f_j : 1 \leq j \leq k\}$.
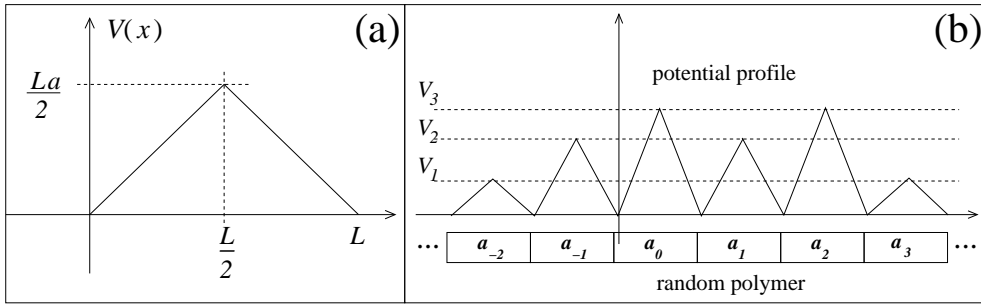
Figure 1. Schematic representation of the potential model. (a) The potential profile on the 0th unit cell. (b) A realization of the random potential with a few unit cells. In this case every monomer along the chain can be taken among three possible monomers ($k = 3$) with heights $V_1 = f_1 L/2$, $V_2 = f_2 L/2$ and $V_3 = f_3 L/2$. To perform analytical calculations as well as numerical simulations we have taken the values $f_1 = 0.8$, $f_2 = 4.2$, and $f_3 = 9.0$ (see text).

We should stress here that the sequence $\mathbf{a} := (\ldots, a_{-1}, a_0, a_1, \ldots) \in \mathcal{A}^{\mathbb{Z}}$ represents the polymer (with a_n the corresponding monomers for $n \in \mathbb{Z}$). The values f_j represent the "slopes" that the potential can take and, in some way, stand for the possible monomer types from which the polymer is built up. It is clear that the proposed potential depends only on one monomer, i.e., $V(x) = \psi(y, a_n)$ if $x = y + nL$. Since we are considering the Bernoulli measure on $\mathcal{A}^{\mathbb{Z}}$, we only need to specify the probability that a given monomer a_n equals a monomer type f_j for $1 \leq j \leq k$. In other words, we need to specify the 1-dimensional marginal, i.e., $\nu(a_n = f_j) =: p(f_j)$ to perform the polymer average. With these quantities we can state explicitly how to take the average with respect to the polymer ensemble. If $h : \mathcal{A} \to \mathbb{R}$, then we have

$$\langle h(a) \rangle_{\mathrm{p}} = \sum_{j=1}^{k} h(f_j) p(f_j). \tag{19}$$

Notice that this average does not depend on n, which reflects the fact that the chosen measure is translationally invariant (or shift-invariant). With this potential model we have that the integrals Z_+ and Z_- can be done exactly, giving

$$
\begin{aligned}
Z_\pm &= 2 \left\langle \frac{\exp\left(\pm\beta aL/2\right) - 1}{\pm\beta a} \right\rangle_{\mathrm{p}} \\
&= 2 \sum_{j=1}^{k} \frac{\exp\left(\pm\beta f_j L/2\right) - 1}{\pm\beta f_j} p(f_j).
\end{aligned}
\tag{20}
$$

The presence of correlations does not change the above average. This is because the function to be averaged only depends on one coordinate of $\mathbf{a} \in \mathcal{A}^{\mathbb{Z}}$. Thus, to perform the average over the polymer ensemble we only require the 1-dimensional marginals. Thus if two stochastic processes producing the (random) substrate have identical 1-dimensional marginals then they will exhibit the same diffusion coefficient at all temperatures. To confirm that it is the case, we simulate the Langevin dynamics of 10 000 overdamped particles diffusing on several random polymers. Then we calculate the diffusion coefficient from the exact expression given in Eq. (16) and compare it with the diffusivity obtained from the simulations. To perform the referred numerical simulations first we built the polymers using a set of $k = 3$ monomer types. The slopes defining the corresponding monomer types are chosen as $f_1 = 0.8$, $f_2 = 4.2$ and $f_3 = 9$. The set of possible monomer types will be represented by $\mathcal{A} := \{1, 2, 3\}$. Thus, a realization of the polymer $\mathbf{a} = (\ldots, a_{-1}, a_0, a_1, \ldots)$ has an associated sequence of potential profiles with "slopes" $(\ldots, f_{a_{-1}}, f_{a_0}, f_{a_1}, \ldots)$. In our numerical simulations we only vary the way in which the polymer is built up. For numerical reasons we fix the parameters L and γ to one throughout the rest of this work.

In Fig. 2 we observe the diffusion coefficient, as a function of the temperature, for particles moving on random polymers built up by means of a Bernoulli process and a Markov chain. On one hand, the Bernoulli process is defined uniquely by the 1-dimensional marginals, $p_j = \nu(a_n = f_j)$ for $j = 1, 2, 3$. Notice that the latter does no depend on the lattice site n, which means that the measure is shift invariant. For our numerical simulations we use the equidistribution for the monomer types, i.e., $p_1 = p_2 = p_3 = 1/3$. This means that the polymer is built up at random as follows: at each unit cell we chose at random a monomer type from the set $\mathcal{A} := \{1, 2, 3\}$ with the same probability. Such a choice is independent of the monomers chosen in the rest of the chain. On the other hand, the Markov process that we use to built up the polymers is defined through the stochastic matrix $P : \mathcal{A} \times \mathcal{A} \to [0, 1]$ given by

$$
P = \begin{pmatrix} 0 & p & q \\ q & 0 & p \\ p & q & 0 \end{pmatrix},
\tag{21}
$$

where p and q are parameters such that $q = 1 - p$. It is easy to see that this matrix is doubly stochastic and the unique invariant probability vector $\pi = \pi P$ is given by $\pi = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$. It is known that the correlations for any Markov process decay exponentially fast [16]. Contrary to the Bernoulli process, the monomer type that is chosen at random at a given cell depends on the choice already made at an adjacent unit cell. Notice that although the Bernoulli and the Markov processes are different in nature as explained above, they have the same 1-dimensional marginals. This means that the probability that a given monomer type appears along the polymer is the same regardless the process (Markovian or Bernoulli). This is a consequence
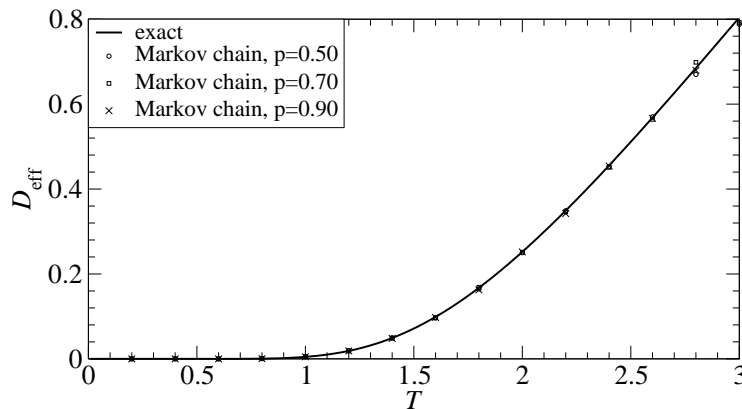
Figure 2. Effective diffusion coefficient for overdamped particles on "Markov" and "Bernoulli" polymers. We simulate the dynamics of 10000 overdamped particles on random polymers built up by means of a Markov chain. The stochastic matrix corresponding to such a Markov process is given in Eq. (21) having a parameter p. We calculated the diffusion coefficient from the simulations for three different cases $p = 0.50$, $p = 0.70$ and $p = 0.90$ of the Markov chain. We also calculate the diffusion coefficient for the case of the Bernoulli process. We observe that the diffusion coefficients for the Markov chain do not vary with the parameter p and coincide with the diffusion coefficient obtained from the Bernoulli process. Moreover, these diffusion coefficients are consistent with the values predicted by Eq. (16). This gives an example of the fact that the diffusion coefficient (in absence of driving force) is independent on the correlations present in the polymers.

of the fact the components of the invariant probability vector $\pi = (1/3, 1/3, 1/3)$ of the stochastic matrix P, actually correspond to the 1-dimensional marginals of the stationary Markov chain. In Fig. 2 we can appreciate that, within the accuracy of our numerical simulations, the diffusion coefficient for different realizations of the polymers coincide, regardless if they are built up by means of the stationary Markov chain or by means of the Bernoulli process. Moreover, these numerical experiments fits satisfactorily (within the accuracy of our simulations) to the exact curve calculated from Eq. (16).

3.2. Potentials of range $n = 2$

In order to test the independence on long-range correlations of the diffusion coefficient we simulate the dynamics of overdamped particles on random polymers built up by means of the expansion modification (E-M) process. The E-M system is a stochastic process that was introduced as a simple model exhibiting spatial $1/f$ noise [17, 18] and, in particular, it was used to understand the long-range correlation present in DNA sequences [19, 20]. Such a system is defined by means of two fundamental processes as follows. Consider a "seed" (a symbol in the binary alphabet $\{0, 1\}$). The seed is expanded with a probability p and it is modified (it is replaced by the complementary symbol) with probability $1 - p$. In other words, if $a \in \{0, 1\}$ then, under the E-M process this symbol is subjected to the transformation,

$$a \mapsto \begin{cases} \overline{a} & \text{with probability } 1 - p \\ aa & \text{with probability } p. \end{cases}$$

and these transformations are extended coordinate-wise to words of arbitrary length in order to define the infinitely iterated E-M process. Here $\bar{a}$ stands for the complementary symbol, i.e., $\bar{1} = 0$ and $\bar{0} = 1$. This process allows the seed to grow in such a way that the resulting string of symbols becomes infinite with probability one. This stochastic process has been studied within the context of symbolic dynamics in Ref. [21]. In that work it was rigorously proved that the E-M system has a unique stationary measure having polynomial decay of correlations for an open set of values of the parameter p. Specifically, they proved that the pair correlation function $C(\ell)$ has a behavior of the form $C(\ell) \asymp \ell^\theta$ where θ is an exponent depending on the parameter p, i.e. $\theta = \theta(p)$. The authors of Ref. [21] also conjectured that this behavior for the correlation functions occurs for almost all the values of the expansion probability and gave an explicit formula for the corresponding exponent $\theta(p)$,

$$\theta(p) = \frac{\log(1 + p) - \log(|2p - 1|) - \log(|3p - 1|)}{\log(1 + p)}. \tag{22}$$

The treatment of the E-M system by means of symbolic dynamics allows us to obtain exact formulas for the n-dimensional marginals defining the (σ-invariant) stationary probability measure [21]. Indeed, the n-dimensional marginals are obtained as the stationary probability vectors of a hierarchy of stochastic matrices of size $2^n \times 2^n$. Every stochastic matrix in such a hierarchy gives the "transition probability" from one word of size n to another one under the action of expansions and modifications on the polymer. Following the treatment of Ref. [21], we prove in Appendix B that the stochastic matrices for the cases $n = 1$ and $n = 2$ are given by,

$$M_1 = \frac{1}{1 + p} \begin{pmatrix} 2p & 1 - p \\ 1 - p & 2p \end{pmatrix}, \tag{23}$$

$$M_2 = \frac{1}{1 + p} \begin{pmatrix} 2p^2 + pq & pq & pq & q^2 \\ p^2 + 2pq & p^2 & q^2 & pq \\ pq & q^2 & p^2 & p^2 + 2pq \\ q^2 & pq & pq & 2p^2 + pq \end{pmatrix}. \tag{24}$$

The respective invariant probability vectors $v_1 : \mathcal{A} \to [0, 1]$ and $v_2 : \mathcal{A}^2 \to [0, 1]$ are given by

$$v_1 = \left(\frac{1}{2}, \frac{1}{2} \right), \tag{25}$$

$$v_2 = \left(\frac{3 - 2p}{10 - 8p}, \frac{1 - p}{5 - 4p}, \frac{1 - p}{5 - 4p}, \frac{3 - 2p}{10 - 8p} \right), \tag{26}$$

which are obtained as the left eigenvectors associated to the largest eigenvalues of M_1 and M_2 respectively. With these quantities we can evaluate the average w.r.p. of any function on $\mathcal{A}^{\mathbb{Z}}$ that depends on one or two coordinates of the polymer. Specifically, if we have a function depending on one coordinate of the polymer, $g : \mathcal{A} \to \mathbb{R}$, then the average w.r.p. is defined by

$$\langle g(a) \rangle_{\mathrm{P}} = \sum_{a=0}^{1} g(a) v_1(a). \tag{27}$$

On the other hand if we have a function depending on two coordinates of the polymer, $h : \mathcal{A}^2 \to \mathbb{R}$, then the average w.r.p. should be achieved as follows,

$$\langle h(a, b) \rangle_{\mathrm{P}} = \sum_{a=0}^{1} \sum_{b=0}^{1} h(a, b) v_2(a, b), \tag{28}$$
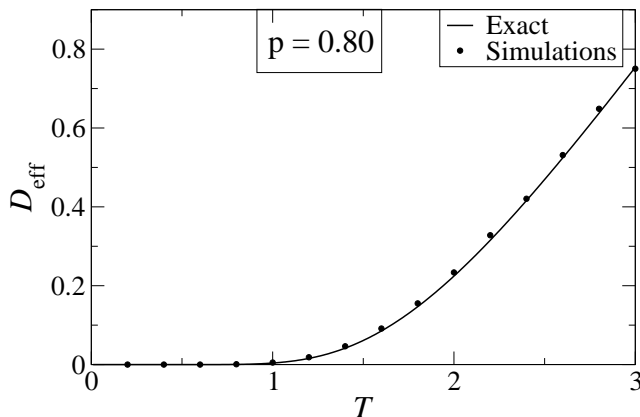
Figure 3. Effective diffusion coefficient of overdamped particles moving on E-M polymers with expansion probability $p = 0.80$. In this figure we display the effective diffusion coefficient estimated from the Langevin dynamic simulations (filled circles) and the theoretical prediction from Eq. (16) (solid line). This figure clearly shows that the effective diffusion coefficient obtained from the simulations is consistent with the theoretical prediction, at least within the accuracy of our numerical experiments.

a summation that will be represented alternatively as

$$\langle h(\mathbf{a}) \rangle_{\mathrm{p}} = \sum_{\mathbf{a} \in \mathcal{A}^2} h(\mathbf{a}) v_2(\mathbf{a}). \tag{29}$$

Now let us define a toy model for the potential induced by the particle-polymer interaction. We will assume that the particle interacts with two monomers when it is placed on the polymer. This interaction potential will depend on two coordinates of the random polymer. We will assume additionally that the induced potential ψ is piece-wise linear. Given a particle placed at the nth unit cell on the polymer $\mathbf{a}$, we write the particle position x as $x = y + nL$ with $y \in [0,1)$. Remember that in this case the polymer is built up from a binary "alphabet" $\mathcal{A} = \{0, 1\}$ (since we have only two symbols in the expansion-modification process). Then the potential is defined as

$$\psi(x, a_n, a_{n+1}) = \begin{cases} \varphi(a_n, a_{n+1}) y & \text{if} \quad 0 \leq y < L/2 \\ \varphi(a_n, a_{n+1})(L - y) & \text{if} \quad L/2 \leq y < L. \end{cases} \tag{30}$$

The function $\varphi : \mathcal{A}^2 \times \mathcal{A}^2 \to \mathbb{R}$ appearing in the equation above is such that,

$$\varphi(a, b) = \begin{cases} f_1 & \text{if} \quad a = 0 \text{ and } b = 0 \\ f_2 & \text{if} \quad (a = 0 \text{ and } b = 1) \quad \text{or} \quad (a = 1 \text{ and } b = 0) \\ f_3 & \text{if} \quad a = 1 \text{ and } b = 1. \end{cases} \tag{31}$$

The values f_i for $i = 1, 2, 3$ are chosen as in the preceding example, $f_1 = 0.80$, $f_2 = 4.2$ and $f_3 = 9.0$. With this model we have that the resulting potential on a given unit cell depends not only on the monomer type located in such a unit cell but also depends on the right adjacent monomer type. Note that the potential profile defined by (30) is the analogous to the potential profile defined by Eq. (18). The difference is that in the present model the slopes are defined by two coordinates of the polymer, instead of one as in the potential model of Eq. (18). With these definition we can now evaluate the quantities Z_+ and Z_- defined in Eq. (17) which allows us to obtain an expression
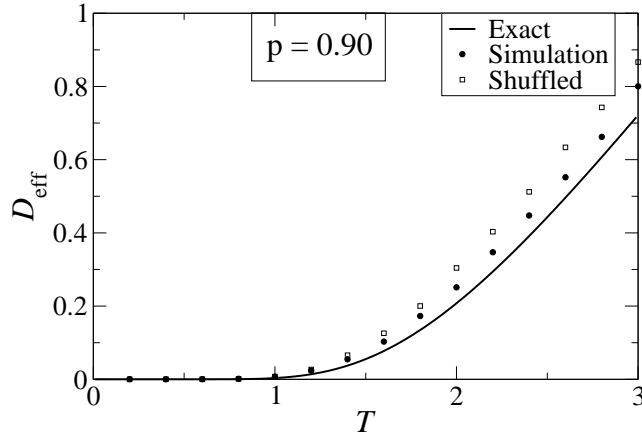
Figure 4. Effective diffusion coefficient of overdamped particles moving on E-M polymers with expansion probability $p = 0.90$. In this figure we display the effective diffusion coefficient estimated from the Langevin dynamic simulations (filled circles) and the theoretical prediction from Eq. (16) (solid line). We should notice a clear (non-random) deviation of the simulations with respect to the theoretical curve. In this figure we also display the effective diffusion coefficient of particle on the polymers produced by the E-M process but after shuffling the symbols. The shuffling process "destroy" the correlations and, as we can see, this diffusion coefficient differs from both, the theoretical result and the diffusivity originally calculated in the correlated polymer. This result evidence the fact that the E-M process has not reached its stationary state and the fact that the diffusion process in long-range correlated polymers undergoes a long transient resulting possible in a very slow convergence of the diffusion coefficient.

for the diffusion coefficient through Eq. (16). Notice that the integral appearing in Eq. (17) is the same as in the case of potentials of range $n = 1$. Thus, it is easy to see that,

$$
\begin{aligned}
Z_\pm &= 2 \left\langle \frac{\exp\left(\pm\beta\varphi(\mathbf{a})L/2\right) - 1}{\pm\beta\varphi(\mathbf{a})} \right\rangle_{\mathrm{p}} \\
&= 2 \sum_{\mathbf{a}\in\mathcal{A}^2} \left[\frac{\exp\left(\pm\beta\varphi(\mathbf{a})L/2\right) - 1}{\pm\beta\varphi(\mathbf{a})} \right] v_2(\mathbf{a}).
\end{aligned}
\tag{32}
$$

In Fig. 3 we show the effective diffusion coefficient for particles along an E-M polymer with expansion probability $p = 0.80$. In this case the pair correlation function of the polymer is approximately given by $C(\ell) \approx \ell^{\theta(0.80)}$, where $\theta(0.80) \approx 1.296$, which implies that the correlations decay too slowly. To perform the numerical experiment (from which we estimate the effective diffusion coefficient) we simulated the Langevin dynamics of $35\,000$ particles, every particle moving on different realizations of the polymer, during a total time of 10^5 arb. units. We can appreciate in Fig. 3 that even in this situation, where the polymer has long range correlations, the diffusion coefficient obtained from the Langevin dynamics (filled circles) fits satisfactorily (within the accuracy of our simulations) with the diffusion coefficient predicted by Eq. (16) (solid line). This means, again, that the diffusivity does not depend on the large correlation present in the disordered substrate.

We now test our formula (16) for the effective diffusivity in a more dramatic situation. This is the case where the correlations decay much more slowly in

such a way that the correlation function behaves asymptotically as $C(\ell) \approx \ell^{\theta(0.90)}$ where $\theta(0.90) \approx 0.521$. In this case, a typical realization of the E-M polymer has large ordered domains which are responsible of the large correlations present in the (disordered) substrate. It is natural to expect, in the present case, large deviated results from the theoretical ones. This is, effectively, the case as we observe in Fig. 4. The effective diffusion coefficient obtained from Langevin dynamic simulations (filled circles) deviates moderately (but with a clear non random tendency) from the theoretical curve predicted by Eq. (16) (solid line). This behavior could be the result of two phenomena naturally present in the system. First we should remind that the E-M process relaxes slowly to its stationary state, especially for large values of the expansion probability p [21]. As it was stated in Ref. [21], we should be aware of this phenomenon since it could lead to distorted observations. On the other hand, another phenomenon that could also be responsible of the deviations observed in Fig. 4 is the fact that the diffusion of the particles is by itself a very slow process. Indeed, the effective diffusivity is a property of the whole polymer, and if the particles are unable to visit the whole polymer, the diffusion coefficient would have large statistical errors. To verify that these two phenomena are actually present in the numerical simulations we performed the following experiment. We toke the polymer produced by the E-M process (with $p = 0.90$) and we shuffled its monomers in order to "destroy" the large correlations on the substrate. This procedure was done by taking two monomers at random and interchanging its position along the polymer. We performed 10^6 "interchanges" on every polymer of 10^6 symbols long. Then we placed the particles on the shuffled polymers and performed the corresponding Langevin simulations. In Fig. 4 we display the diffusion coefficient calculated in this way (open squares). We appreciate that the resulting diffusion coefficient still deviates from the theoretical one. This means that the deviation of the diffusivity observed in the long range correlated polymers does not come from the correlations at all, but from the fact that the polymer itself, has not reached its stationary state. This phenomenon can be appreciated in Fig. 5. In such a figure we plot the relative frequency of occurrence of the word 00 for 50 realizations of the random polymer of 10^6 symbols long. We can observe that for the case $p = 0.80$ such relative frequency (filled circles) coincide in the average with the theoretical value (solid line), the two-dimensional marginal $v_2(00)$, predicted by Eq. (26). On the other hand, for the case $p = 0.90$, we can appreciate that the relative frequencies obtained from the simulations (stars) deviate largely from the theoretical prediction (dashed line) compared to the case $p = 0.80$. Indeed, the standard deviation of the relative frequency for the case $p = 0.80$ is $\Delta v \approx 0.001550471$ while for the case $p = 0.90$ we obtain $\Delta v \approx 0.02874428$ which differs in one order of magnitude. This behavior clearly affects the effective diffusion coefficient obtained from the simulations even in the case where the correlations are destroyed by shuffling the symbols.

4. Discussion and conclusions

In this work we have considered a model for unbiased diffusion of particles moving along disordered correlated polymers. We have proved that if the particle only interacts with the monomer closest to it, the diffusivity does not depend on the correlations between monomers along the polymer. In general we have proved that if the particle interacts with a few monomers on the polymer, it will depend only on the correlations of monomers as distant as the range of the interaction. This result was proved under the assumptions (i) that the Einstein relation is valid in this kind of
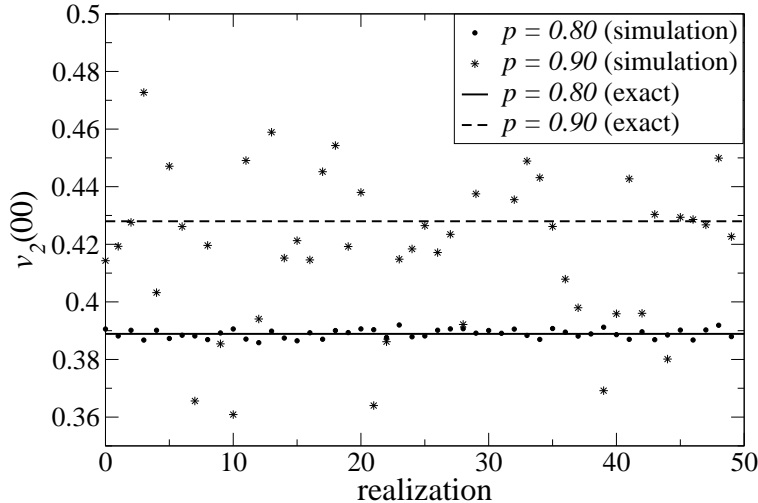
Figure 5. Two-dimensional marginal of the polymer corresponding to the word 00. We display the relative frequency of occurrence of the word 00 along every realization of the polymer. Every realization has 10^6 symbols and we produced 50 different realizations for the cases $p = 0.80$ and $p = 0.90$. We observe that in the case $p = 0.80$ the numerically obtained two-dimensional marginal $v_2(00)$ is consistent with the predicted by Eq. (26). We also appreciate that for the case $p = 0.90$ we have large statistical deviation from the theoretical value compared to the case $p = 0.90$. Actually, the standard deviation for $p = 0.90$ is an order of magnitude larger than the standard deviation of the case $p = 0.80$.

systems and (*ii*) that the polymer, regarded as a discrete-time stochastic process, is ergodic with decay of correlations. We illustrated the independence of the diffusivity on the polymer correlations by means of a simple model in two different cases. In the first case we produced the polymers by means of a three-states Markov process. Then these polymers were used to build up the potential felt by the particles. In this case we observed that the correlations does not alter the diffusivity and the corresponding effective diffusion coefficient that we obtained from the simulations was consistent with the exact result we derived. In the second case that we analyzed, the polymers were produced by means of the E-M process. The E-M process is known to exhibit long range correlations and the corresponding scaling laws for the correlation functions are known exactly. This feature of the E-M process allowed us to study the diffusivity in a system with a very slow decay of correlations. We could verify the independence on the correlations of the diffusivity; however we found that the effective diffusion coefficient seems to deviated moderately in the case in which the correlations decay as $C(\ell) \approx \ell^{\theta(0.90)}$ with $\theta(0.90) \approx 0.521$. We argued that this apparent inconsistence might be due to the fact that the E-M process relaxes slowly to its stationary state. Moreover, we observed that the numerically obtained two-dimensional marginal $v_2(00)$ (corresponding to the realizations of the polymers used for the numerical simulations) exhibits large statistical deviations on every realization of the polymer. This fact clearly affects the diffusivity along every polymer, resulting naturally in a deviation from the theoretical result. However it is also clear that the diffusion process by itself could be a very slow process since the particles should explore all the polymer in order to have an accurate estimation of the diffusivity. However, due to the existence

of large ordered domains in the polymer we could face a very slow convergence of the effective diffusion coefficient. It is therefore necessary a more detailed study of this class of systems particularly focusing in understanding how long is the transient and how does it behave. Knowing these properties could shed some light about the diffusion processes occurring in real systems in which the long-range correlations are naturally present, such as the DNA.

Acknowledgments

Appendix A. Diffusion coefficient in the limit of zero tilt strength

In Ref. [13] it has been shown that in the polymer-particle interaction model for diffusion on random media the particle flux can be written exactly in terms of the mean FPT,

$$J_{\text{eff}} = \frac{L}{\langle T_1 \rangle_{\text{p}}}, \tag{A.1}$$

where the mean FPT w.r.n. T_1 is given by

$$T_1(\mathbf{a}) = \gamma\beta \sum_{m=1}^{\infty} e^{-m\beta FL} q_+(\mathbf{a}) q_- [\sigma^{-m}(\mathbf{a})] + \gamma\beta I_0(\mathbf{a}). \tag{A.2}$$

Our goal here is to prove that in the limit of $F = 0$ the mobility μ,

$$\mu = \lim_{F \to 0} \frac{J(F)}{F}$$

is well defined for any stationary measure ν defining the polymer ensemble. The only condition we impose is the decay of correlations even at a very slow rate. In other words, we impose a mild condition on the rate of mixing of the dynamical system, which in turn implies ergodicity [22]. By ergodicity we mean that the time-average of any regular observable $\mathcal{O} : \mathcal{A}^{\mathbb{Z}} \to \mathbb{R}$, can be interchanged by its space average with respect to the shift-invariant probability measure ν, i.e.,

$$\int_{\mathcal{A}^{\mathbb{Z}}} \mathcal{O} \, d\nu = \lim_{n \to \infty} \frac{1}{n} \sum_{j=0}^{n-1} \mathcal{O} \circ \sigma^j(\mathbf{a}) \quad \nu\text{-almost everywhere.}$$

Here our observables will be $q_{\pm}$ and T_1 which are non-negative and for any potential ψ bounded by above, they are also bounded, and furthermore are $\text{L}^1(\nu)$, i.e.

$$\langle T_1 \rangle_{\text{p}} = \int_{\mathcal{A}^{\mathbb{Z}}} T_1 d\nu = \int_{\mathcal{A}^{\mathbb{Z}}} |T_1| \, d\nu < \infty.$$

Next, we will obtain an expression for the asymptotic behavior of J_{eff}/F as F approaches to zero. In order to do this, let us first estimate bounds for $\langle T_1 \rangle_{\text{p}}$. We assume that the system $(\mathcal{A}^{\mathbb{Z}}, \sigma, \nu)$ satisfies that for every $f \in \text{L}^{\infty}(\nu)$ and $g \in \text{L}^1(\nu)$ there exists a decreasing-to-zero function $r(m)$ such that

$$\left| \int f \cdot g \circ \sigma^m d\nu - \int f d\nu \int g d\nu \right| \leq ||f||_{\infty} ||g||_1 \cdot r(m),$$

where $|| \cdot ||_\infty$ and $|| \cdot ||_1$ denote the $L^\infty(\nu)$ and $L^1(\nu)$ norms respectively. The rate function $r(m)$ will be precisely given below. Let us write $C := ||f||_\infty ||g||_1$, this is a constant since $q_\pm$ are essentially bounded and $L^1(\nu)$, and thus we are able to give the following bounds

$$\langle T_1 \rangle_{\mathrm{p}} \lessgtr \gamma\beta \sum_{m=1}^\infty e^{-m\beta FL} \left(\int q_+ d\nu \int q_- d\nu \pm Cr(m) \right) + \gamma\beta \int I_0 d\nu, \quad (\text{A.3})$$

$$\lessgtr \gamma\beta \left(\frac{\langle q_+ \rangle_{\mathrm{p}} \langle q_- \rangle_{\mathrm{p}}}{1 - e^{-\beta FL}} \pm C \sum_{m=1}^\infty e^{-m\beta FL} r(m) + \int I_0 d\nu \right). \quad (\text{A.4})$$

Which makes that the particle flux (Eq. A.1) be bounded by above and below as follows,

$$J_{\mathrm{eff}} \lessgtr \frac{L}{\gamma\beta \left(\frac{\langle q_+ \rangle_{\mathrm{p}} \langle q_- \rangle_{\mathrm{p}}}{1 - e^{-\beta FL}} \mp C \sum_{m=1}^\infty e^{-m\beta FL} r(m) + \int I_0 d\nu \right)}.$$

Next, let us consider the approximation as $F \to 0$ of the following quantity,

$$\lim_{F \to 0} F \left(\frac{\langle q_+ \rangle_{\mathrm{p}} \langle q_- \rangle_{\mathrm{p}}}{1 - e^{-\beta FL}} \pm C \sum_{m=1}^\infty e^{-m\beta FL} r(m) + \int I_0 d\nu \right).$$

The third part goes to zero since $\int I_0 d\nu$ is constant. By limit identities, the first part is given by

$$\lim_{F \to 0} \frac{F \cdot \langle q_+ \rangle_{\mathrm{p}} \langle q_- \rangle_{\mathrm{p}}}{1 - e^{-\beta FL}} = \frac{Z_+ Z_-}{\beta L}.$$

It remains to estimate the summation involving the rate function, for which we make use of the assumption on the decay of the correlations. Let us assume that the rate function $r(m) = m^{-\alpha}$ for any $\alpha > 0$. This implies that the sum

$$\sum_{m=1}^\infty e^{-m\beta FL} m^{-\alpha},$$

can be written as the polylogarithm function [23], defined as,

$$\mathrm{Li}_\alpha(x) := \sum_{k=1}^\infty \frac{x^k}{k^\alpha}. \quad (\text{A.5})$$

It is known that the polylogarithm function has the asymptotic behavior [23]

$$\mathrm{Li}_s(e^{-\xi}) \approx \Gamma(1-s)\xi^{s-1} \quad (\text{A.6})$$

for $|\xi| \ll 1$ and for non-integer s (in particular for $0 < s < 1$), being Γ the well known gamma function. If we consider a $F_0 > 0$ small enough such that $\beta F_0 L \ll 1$, for any fixed β and L, then, one may write for every $F \le F_0$,

$$FC \sum_{m=1}^\infty e^{-m\beta FL} m^{-\alpha} \approx CF \cdot \Gamma(1-\alpha)(\beta FL)^{\alpha-1},$$

for all $\alpha > 0$. In this way we have that

$$\lim_{F \to 0} FC \sum_1^\infty e^{-m\beta FL} m^{-\alpha} = \lim_{F \to 0} C \cdot \Gamma(1-\alpha)(\beta L)^{\alpha-1} F^\alpha \to 0,$$

for all $\alpha > 0$ and $\beta, L > 0$. Now, using the Einstein relation we see that the effective diffusion coefficient is given by

$$D_{\text{eff}} = \frac{1}{\beta} \lim_{F \to 0} \frac{J_{\text{eff}}}{F} = \frac{1}{\gamma\beta} \frac{L^2}{Z_+ Z_-},$$

which is the expression for the diffusivity anticipated in Eq. (16).

On one hand, the above treatment gave us the desired expression for the diffusivity when one takes firstly the average on the state space and after that the approximation when F goes to zero. On the other hand, one would expect that the expression (16) remains valid when one takes first the approximation $F \to 0$ and after that the average with respect to the polymer ensemble. Here we give a result that enables us to argue that this could be actually true. Nevertheless, we left for further research the mathematically rigorous treatment. The importance on this kind of result is that the Einstein relation should hold even if we do not perform the average over the polymer ensemble. This means in particular that the mean FPT T_1 should behave as F^{-1} for $F \to 0$ which in turn implies Eq. (16) as we will see below.

For the sake of convenience, we introduce a quantity of interest. Let $N \in \mathbb{N}$ and define the observable $g_N : \mathcal{A}^{\mathbb{Z}} \to \mathbb{R}$ as,

$$g_N(\mathbf{a}) := \frac{1}{N} \sum_{m=1}^{\infty} e^{-mz/N} f \circ \sigma^m(\mathbf{a}), \tag{A.7}$$

for some $z \in \mathbb{R}^+$. The discussion in the rest of this appendix is to justify the approximation

$$\sum_{m=1}^{\infty} e^{-mz/N} f \circ \sigma^m(\mathbf{a}) \approx \frac{N}{z} \int_{\mathcal{A}^{\mathbb{Z}}} f \, d\mu = \frac{N}{z} \langle f(\mathbf{a}) \rangle_{\text{p}},$$

for large N.

For the moment, let us go to expression for the mean FPT and see that the approximation above may be applied to the summations appearing in T_1 in Eq. (A.2) if we identify N with $1/F$,

$$T_1(\mathbf{a}) = \gamma\beta q_+(\mathbf{a}) \sum_{m=1}^{\infty} e^{-m\beta FL} q_- [\sigma^{-m}(\mathbf{a})] + \gamma\beta I_0(\mathbf{a}),$$

$$\approx \gamma\beta q_+(\mathbf{a}) \frac{1}{\beta FL} \langle q_-(\mathbf{a}) \rangle_{\text{p}} + \gamma\beta I_0(\mathbf{a}). \tag{A.8}$$

Notice that in the limit $F \to 0$ the mean FPT goes to infinity as F^{-1}. The term $\gamma\beta I_0(\mathbf{a})$ appearing in the above expression remains finite if $F = 0$, and therefore it can be neglected in further calculations. Then we have that the particle flux behaves asymptotically as

$$J_{\text{eff}} = \frac{L}{\langle T_1 \rangle_{\text{p}}} \approx \frac{FL^2}{\gamma} \frac{1}{\langle q_+(\mathbf{a}) \rangle_{\text{p}} \langle q_-(\mathbf{a}) \rangle_{\text{p}}},$$

when F is small enough. Now, using the Einstein relation we see that the effective diffusion coefficient is given by

$$D_{\text{eff}} = \frac{1}{\beta} \lim_{F \to 0} \frac{J_{\text{eff}}}{F} = \frac{L^2}{\gamma\beta} \frac{1}{Z_+ Z_-},$$

where $Z_{\pm}$ are given by Eq. (17) and which is the expression for the diffusivity anticipated in Eq. (16).

Next, in order to justify our claim, we give the following proposition.

Proposition 1 *Let $(\mathcal{A}^{\mathbb{Z}}, \sigma, \mathcal{B}, \nu)$ be a measure-preserving dynamical system, where $\sigma : \mathcal{A}^{\mathbb{Z}} \to \mathcal{A}^{\mathbb{Z}}$ is the shift mapping, $\mathcal{B}$ is a Borel sigma-algebra on $\mathcal{A}^{\mathbb{Z}}$ and ν is a shift-invariant probability measure satisfying decay of correlations at rate $O(m^{-\alpha})$ for some $\alpha > 0$. Let $f : \mathcal{A}^{\mathbb{Z}} \to \mathbb{R}$ be a square integrable function on the full shift $\mathcal{A}^{\mathbb{Z}}$, i.e.,*

$$\int_{\mathcal{A}^{\mathbb{Z}}} f^2 d\nu < \infty.$$

And let g_N be the observable defined in (A.7). Then, for every $\epsilon > 0$ there exists a sufficiently large N such that

$$\nu\left(\left\{\mathbf{a} \in \mathcal{A}^{\mathbb{Z}} : \left|g_N(\mathbf{a}) - \frac{1}{z}\int_{\mathcal{A}^{\mathbb{Z}}} f \, d\nu\right| \geq \epsilon\right\}\right) \leq \frac{C \cdot \Gamma(1-\alpha)z^{\alpha-2}}{N^{\alpha}\epsilon^2}.$$

This inequality states that the larger we take N the more probable is to find g_N around $(1/z)\int f d\nu$ in a ϵ neighborhood. The proof of this proposition use the decay of the auto-correlation function, for that reason the square integrable assumption on the observable f.

Proof: First, notice that

$$\int_{\mathcal{A}^{\mathbb{Z}}} g_N \, d\nu = \frac{1}{N}\sum_{m=1}^{\infty} e^{-mz/N}\int_{\mathcal{A}^{\mathbb{Z}}} f \circ \sigma^m d\nu,$$

$$= \frac{1}{N}\frac{1}{1-e^{-z/N}}\int_{\mathcal{A}^{\mathbb{Z}}} f d\nu \tag{A.9}$$

because ν is σ-invariant. For the sake of clarity let us denote by $\bar{f} := \int_{\mathcal{A}^{\mathbb{Z}}} f d\nu$ and let us define $K : \mathbb{N} \to \mathbb{R}$ as follows,

$$K(N) := \frac{1}{N}\frac{1}{1-e^{-z/N}}$$

and note that for every $z > 0$

$$\lim_{N\to\infty} K(N) = \frac{1}{z}.$$

Observe that for any $\epsilon > 0$ we have

$$\nu\left(\left\{\left|g_N(\mathbf{a}) - \frac{\bar{f}}{z}\right| \geq \epsilon\right\}\right) = \nu\left(\left\{\left|g_N(\mathbf{a}) - K(N)\bar{f} + K(N)\bar{f} - \frac{\bar{f}}{z}\right| \geq \epsilon\right\}\right),$$

$$\leq \nu\left(|g_N - K(N)\bar{f}| > \epsilon/2\right) + \nu\left(|K(N)\bar{f} - \bar{f}/z| > \epsilon/2\right).$$

For the second summand in the right-hand side of the inequality, there is no probability. So, there exists a $N^*(\epsilon)$ such that for all $N \geq N^*$, $|K(N)\bar{f} - \bar{f}/z| \leq \epsilon/2$, and so that part eventually vanishes. While for the first summand, by the Chebyshev inequality we have that for all $N \in \mathbb{N}$,

$$\nu\left(\left\{|g_N(\mathbf{a}) - K(N)\bar{f}| \geq \epsilon/2\right\}\right) \leq \frac{4\text{Var}(g_N)}{\epsilon^2}.$$

The variance of g_N can be rewritten in terms of the auto-correlation function,

$$\text{Var}(g_N) = \int g_N^2 d\nu - \left(\int g_N d\nu\right)^2 = \int g_N^2 d\nu - K^2(N)\bar{f}^2.$$

We need to calculate $\int g_N^2 d\nu$ and we obtain that

$$\int g_N^2 d\nu = \frac{1}{N^2}\sum_{m=1}^{\infty}\sum_{k=1}^{\infty} e^{-mz/N}e^{-kz/N}\int (f\circ\sigma^m)(f\circ\sigma^k) d\nu,$$

using the σ-invariance of ν we get,

$$\mathrm{Var}(g_N) = \frac{1}{N^2} \sum_{m=1}^{\infty} \sum_{k=1}^{\infty} e^{-mz/N} e^{-kz/N} \left[\int f \cdot f \circ \sigma^{m-k} \, d\nu - \left(\int f \, d\nu \right)^2 \right]$$

$$= \frac{1}{N^2} \sum_{m=1}^{\infty} \sum_{k=1}^{\infty} e^{-mz/N} e^{-kz/N} C_f(m-k). \tag{A.10}$$

where C_f stands for the auto-correlation function of the observable f, i.e.,

$$C_f(\ell) := \int f \cdot f \circ \sigma^{\ell} \, d\nu - \left(\int f \, d\nu \right)^2.$$

The expression for the variance of g_N given in Eq. (A.10) can rewritten by changing the summation order, as follows

$$\mathrm{Var}(g_N) = \frac{2}{N^2} \sum_{\ell=1}^{\infty} \sum_{n=1}^{\infty} e^{-2nz/N} e^{-\ell z/N} C_f(\ell) + \frac{1}{N^2} C_f(0) \sum_{n=1}^{\infty} e^{-2nz/N}.$$

$$= K(N/2) \frac{1}{N} \sum_{\ell=1}^{\infty} e^{-\ell z/N} C_f(\ell) + \frac{1}{N} C_f(0) \frac{K(N/2)}{2}.$$

Now, the assumption on the decay of correlations becomes cricial. Let us assume that the auto-correlation function for f behaves as $C_f(\ell) = C_0 \ell^{-\alpha}$ for any $\alpha > 0$ and some $C_0 > 0$. In this case we can write the variance of g_N as follows

$$\mathrm{Var}(g_N) = K(N/2) \frac{C_0}{N} \sum_{\ell=1}^{\infty} \ell^{-\alpha} e^{-\ell z/N} + \frac{1}{N} C_f(0) \frac{K(N/2)}{2}$$

$$= K(N/2) \frac{C_0}{N} \mathrm{Li}_\alpha(e^{-z/N}) + \frac{1}{N} C_f(0) \frac{K(N/2)}{2}.$$

In the second equation we have rewritten the expression for the variance of g_N using the polylogarithm function (A.5). From the approximation (A.6) we are able to write down an asymptotic expression for the variance of g_N for a sufficiently large N,

$$\mathrm{Var}(g_N) \approx K(N/2) \frac{C_0}{N} \Gamma(1-\alpha) \left(\frac{z}{N} \right)^{\alpha-1} + \frac{1}{N} C_f(0) \frac{K(N/2)}{2}.$$

Finally, since $K(N/2) \to 1/z$ as $N \to \infty$, we easily see that $\mathrm{Var}(g_N) \to 0$ in the same limit, and specifically

$$\mathrm{Var}(g_N) \approx C_0 \Gamma(1-\alpha) z^{\alpha-2} N^{-\alpha}.$$

The latter implies that we can chose an appropriate constant C such that for a sufficiently large N

$$\nu \left(\left\{ \left| g_N(\mathbf{a}) - \frac{\bar{f}}{z} \right| \geq \epsilon \right\} \right) \leq \frac{C \cdot \Gamma(1-\alpha) z^{\alpha-2}}{N^\alpha \epsilon^2}.$$

Which concludes the proof the proposition.

$$\square$$

Appendix B. The n-dimensional marginals for the expansion-modification system

The 1-dimensional marginals for the E-M system can be understood as the relative frequency of occurrence of symbols (monomers) of every kind along a typical realization of the symbolic sequence (the polymer) generated by the E-M process at the stationary state. For the sake of simplicity here we will call "polymer" to an infinite symbolic sequence (i.e., a point in the space $\{0,1\}^{\mathbb{Z}}$).

The processes generating the infinitely long polymer of the E-M system, give us the way to obtain the marginals as follows. First, let $\mathbf{a}_0 \in \{0,1\}^{\mathbb{Z}}$ be an "initial polymer" (of length 0) and allow it evolve, coordinatewise, according to the E-M processes defined as

$$x \mapsto \begin{cases} \overline{x} & \text{with probability } 1-p \\ xx & \text{with probability } p. \end{cases}$$

These rules generate a sequence of "polymers" of finite size $\mathbf{a}_0 \mapsto \mathbf{a}_1 \mapsto \mathbf{a}_2 \cdots$ that attain a stationary state for $t \to \infty$. The E-M system is actually a discrete-time Markov process [21]. Let us call $f_0(t)$ and $f_1(t)$ the fraction of 0's and 1's in the symbolic chain $\mathbf{a}_t$ at time $t \in \mathbb{N}$. This means that in any subsequence of $\mathbf{a}_t$ of n symbols we will find typically $nf_0(t)$ zeroes and $nf_1(t)$ ones. We will obtain a recurrence relation for f_0 and f_1. First notice that the chain $\mathbf{a}_{t+1}$ is generated from $\mathbf{a}_t$ by means of a fraction of p expansions and $1-p$ modifications. This means that a subsequence $\mathbf{b}$ of size n in $\mathbf{a}_t$ generate a sequence of size $n(1+p)$ in $\mathbf{a}_{t+1}$. The total symbols that are expanded in $\mathbf{b}$ is pn and the total symbols that are modified is $(1-p)n$. From the total symbols that are "expanded", a fraction f_0 of them are zeroes and a fraction f_1 of them are ones. Analogously from the total symbols that are "modified", a fraction f_0 of them are zeroes and a fraction f_1 of them are ones. This means that the number of symbols that are zeroes in the new chain has two contributions: the number of 0's that were expanded $2npf_0(t)$ plus the number of 1's that were modified $n(1-p)f_1(t)$ becoming zeroes in the new chain. An analogous statement follows immediately for $f_1(t)$. Since the size of the new chain is $(1+p)n$ we have that the fractions $f_0(t)$ and $f_1(t)$ obey the following recurrence relations,

$$f_0(t+1) = \frac{2pf_0(t) + (1-p)f_1(t)}{1+p} \tag{B.1}$$

$$f_1(t+1) = \frac{(1-p)f_0(t) + 2pf_1(t)}{1+p}. \tag{B.2}$$

Notice that this defines a one-step Markov chain with probability matrix M_1 : $\{0,1\} \times \{0,1\} \to [0,1]$,

$$M_1 = \frac{1}{1+p} \begin{pmatrix} 2p & 1-p \\ 1-p & 2p \end{pmatrix}. \tag{B.3}$$

In the limit of $t \to \infty$ the fractions f_0 and f_1 tend to the unique invariant probability vector for M_1. It can be seen that such a probability vector is given by

$$v_1 = \left(\frac{1}{2}, \frac{1}{2} \right). \tag{B.4}$$

Next we can generalize the concept of fractions of zeroes and ones, to the *relative frequency* $f_t(\mathbf{w})$ with which a word $\mathbf{w} \in \{0,1\}^k$ of size k occurs along a the polymer

Table B1. Effective number of words of size two given by the expansion-modification process. This rules allows calculate the recurrence relations for the evolution of frequency of appearance of words of size two.

Process	Probability	Initial words			
		00 gives	01 gives	10 gives	11 gives
ee	p^2	00, 00	00, 01	11, 10	11, 11
em	$p(1-p)$	00, 01	00, 00	11, 11	11, 10
me	$p(1-p)$	10	11	00	01
mm	$(1-p)^2$	11	10	01	00

$\mathbf{a}_t \in \{0,1\}^{\mathbf{Z}}$ at time t.

$$f_t(\mathbf{w}) := \lim_{n\to\infty} \sum_{j=0}^{n-1} \mathbb{I}_{[\mathbf{w}]}\left(\sigma^j(\mathbf{a}_t)\right) \tag{B.5}$$

where $\mathbb{I}$ is the indicating function, which takes the value 1 if $\mathbf{a}_t$ belongs to the set $[\mathbf{w}]$ and take the value 0 if $\mathbf{a}_t$ does not belong to the set $[\mathbf{w}]$. The set $[\mathbf{w}]$ called *cylinder* is defined as the set of all infinite sequences $\mathbf{x} \in \{0,1\}^{\mathbb{Z}}$ with prefix $\mathbf{w}$, i.e.,

$$[\mathbf{w}] := \left\{\mathbf{x} \in \{0,1\}^{\mathbb{Z}} : x_0 = w_0, x_1 = w_1, \ldots, x_{k-1} = w_{k-1}\right\}.$$

In other words, we calculate the frequency that a word of size k appears in a polymer by counting the number times such a word appears in a "window" of size k which slides along the polymer. Let us analyze the case $k = 2$. To obtain the stationary 2-dimensional marginals, or equivalently, the stationary relative frequencies of words of size 2, we can obtain recurrence relations as in the 1-dimensional case. First notice that a word of size 2 generate words of sizes 2, 3 and 4 by means of expansions and modifications. We have four cases that should be considered. For example take the word $\mathbf{a} = 00$. This word can be transformed under the E-M process into $\mathbf{b} = 0000, 100, 001, 11$. Each case corresponds to one of the four possibilities of the process: ee, em, me and mm (e stands for "expansion" and m stands for "modification"). First notice that in the case of two expansions ($00 \overset{\text{ee}}{\to} 0000$) we obtain, by sliding a window of size two, three times the word 00. However, this fact does not mean that pure expansions has the effect of triplicate the frequency of appearance of 00. For example, the block 00000 transforms under pure expansion into the block 0000000000. A direct counting shows, by sliding a window of size two, that the word 00 appears four times before the transformation and nine times after such a process. In general, the number of times that the word 00 appears in the block resulting from pure expansions of a block consisting of n zeros, is $2n+1$. This means that the frequency of appearance of 00 is duplicated in the limit of an infinite polymer. In other words, the effective number of words that produces the word 00 under the process ee is two and are 00, 00. A similar analysis shows that, in order to count correctly the number of times a word appears under the processes em, me, and mm we should take into account certain "rules" which are summarized in Table B1. With these rules it is easy to see that the relative frequencies with which the words 00, 01, 10 and 11 appear after a process of expansions and modifications of an infinite polymer, are given by,

$$f_{t+1}(00) = \frac{1}{1+p}\left((2p^2 + pq)f_t(00) + (p^2 + 2pq)f_t(01) + pqf_t(10) + q^2 f_t(11)\right)$$

$$f_{t+1}(01) = \frac{1}{1+p}\bigg(pqf_t(00) + p^2 f_t(01) + q^2 f_t(10) + pqf_t(11)\bigg)$$

$$f_{t+1}(10) = \frac{1}{1+p}\bigg(pqf_t(00) + q^2 f_t(01) + p^2 f_t(10) + pqf_t(11)\bigg)$$

$$f_{t+1}(11) = \frac{1}{1+p}\bigg(q^2 f_t(00) + pqf_t(01) + (p^2 + 2pq)f_t(10) + (2p^2 + pq)f_t(11)\bigg). \quad \text{(B.6)}$$

The above recurrence relations can be recast into a more compact form by defining the probability vector p_2 (the time-dependent two-dimensional marginals) whose components are the relative frequencies at time t,

$$p_2(t) := \bigg(f_{t+1}(00),\ f_{t+1}(01),\ f_{t+1}(10),\ f_{t+1}(11)\bigg),$$

in terms of which we rewrite Eq. (B.6) as a matrix equation,

$$p_2(t+1) = p_2(t)M_2$$

where the matrix M_2 defined as

$$M_2 = \frac{1}{1+p}\begin{pmatrix} 2p^2 + pq & pq & pq & q^2 \\ p^2 + 2pq & p^2 & q^2 & pq \\ pq & q^2 & p^2 & p^2 + 2pq \\ q^2 & pq & pq & 2p^2 + pq \end{pmatrix}. \quad \text{(B.7)}$$

In the limit of $t \to \infty$, it has been proved in Ref [21] that the system reaches a stationary state. This means that the time-dependent probability vector $p_2(t)$ is well defined in the limit $t \to \infty$. The probability vector at infinity is the left eigenvector of M_2, which we denote by v_2, associated to the (largest) eigenvalue $\lambda = 1$,

$$v_2 = \lim_{t\to\infty} p_2(t),$$

and v_2 satisfies the equation

$$v_2 = v_2 M_2,$$

through which we calculated the 2-dimensional marginal given in Eq (26).

References

[1] Sinai Y G 1983 *Theory of Probability & Its Applications* **27** 256–268
[2] Haus J W and Kehr K W 1987 *Physics Reports* **150** 263–406
[3] Havlin S and Ben-Avraham D 1987 *Advances in Physics* **36** 695–798
[4] Bouchaud J P and Georges A 1990 *Physics reports* **195** 127–293
[5] Denisov S I, Denisova E S and Kantz H 2010 *The Eur. Phys. J. B* **76** 1–11
[6] Denisov S I, Kostur M, Denisova E S and Hänggi P 2007 *Phys. Rev. E* **75** 061123
[7] Denisov S I, Kostur M, Denisova E S and Hänggi P 2007 *Phys. Rev. E* **76** 031101
[8] Denisov S I and Kantz H 2010 *Phys. Rev. E* **81** 021117
[9] Salgado-García R and Maldonado C 2013 *Phys. Rev. E* **88** 062143
[10] Gnedenko B V and Kolmogorov A N 1968 *Limit Distributions for Sums of Independent Random Variables* vol 233 (Addison-Wesley Reading)
[11] Gouëzel S 2004 *Prob. Theor. and Rel. Fields* **128** 82–122
[12] Chazottes J R 2012 *arXiv preprint. arXiv:1201.3833*
[13] Salgado-García R 2014 *Phys. Rev. E* **90** 032105
[14] Lind D A 1995 *An Introduction to Symbolic Dynamics and Coding* (Cambridge University Press, Cambridge)
[15] Goychuk I and Kharchenko V O 2014 *Phys. Rev. Lett.* **113**(10) 100601 URL http://link.aps.org/doi/10.1103/PhysRevLett.113.100601

[16] Levin D A, Peres Y and Wilmer E L 2009 *Markov Chains and Mixing Times* (American Mathematical Society, Providence)

[17] Li W 1989 *Europhys. Lett.* **10** 395

[18] Li W 1991 *Phys. Rev. A* **43** 5240

[19] Li W and Kaneko K 1992 *Europhys. Lett.* **17** 655

[20] Li W, Marr T G and Kaneko K 1994 *Physica D* **75** 392–416

[21] Salgado-García R and Ugalde E 2013 *Journal of Statistical Physics* **153** 842–863

[22] Brin M and Stuck G 2002 *Introduction to dynamical systems* (Cambridge University Press, Cambridge)

[23] Wood D 1992 The computation of polylogarithms Tech. Rep. 15-92* University of Kent, Computing Laboratory University of Kent, Canterbury, UK URL http://www.cs.kent.ac.uk/pubs/1992/110

[24] Gouëzel S and Melbourne I 2014 *Electron. J. Probab.* **19** no. 93, 30

[25] Melbourne I and Terhesiu D 2014 *Ergodic Theory Dynam. Systems* **34** 893–918