

**Molecular spiders on a plane**Tibor Antal<sup>1</sup> and P. L. Krapivsky<sup>2</sup><sup>1</sup>*School of Mathematics, Edinburgh University, Edinburgh, EH9 3JZ, United Kingdom*<sup>2</sup>*Department of Physics, Boston University, Boston, Massachusetts 02215, USA*

(Received 1 May 2012; published 29 June 2012)

Synthetic biomolecular spiders with “legs” made of single-stranded segments of DNA can move on a surface covered by single-stranded segments of DNA called substrates when the substrate DNA is complementary to the leg DNA. If the motion of a spider does not affect the substrates, the spider behaves asymptotically as a random walk. We study the diffusion coefficient and the number of visited sites for spiders moving on the square lattice with a substrate in each lattice site. The spider’s legs hop to nearest-neighbor sites with the constraint that the distance between any two legs cannot exceed a maximal span. We establish analytic results for bipedal spiders, and investigate multileg spiders numerically. In experimental realizations legs usually convert substrates into products (visited sites). The binding of legs to products is weaker, so the hopping rate from the substrates is smaller. This makes the problem non-Markovian and we investigate it numerically. We demonstrate the emergence of a counterintuitive behavior—the more spiders are slowed down on unvisited sites, the more motile they become.

DOI: [10.1103/PhysRevE.85.061927](https://doi.org/10.1103/PhysRevE.85.061927)

PACS number(s): 87.15.Vv, 02.50.Ey, 05.40.–a, 82.39.Fk

**I. INTRODUCTION**

Recent advances in DNA nanofabrication technology (see [1–4]) have led to the construction of multipedal walking molecular devices. The first was a bipedal object [3] walking on a one-dimensional path by DNA set strands with nucleic acid domains complementary to molecular imprints on the device legs and the substrate. Since then, several other similar bipedal DNA walkers have been synthesized (see [4–7] and a review [8]).

A different molecular design has been implemented in Refs. [9,10]. The resulting objects, known as molecular spiders, usually have many legs. Each leg (a short single strand of DNA) can bind to the substrate through the Watson-Crick base pair formation. A bound leg can either detach from the substrate without modifying it, or it can catalyze the cleavage of the substrate creating two product strands. The lower product remains bound to the surface, while the upper product is free to float away in solution. There is a residual binding of the leg to the remaining product, but it is weaker than the leg-substrate binding. Utilizing the effect of the spider’s motion on the molecular tracks it is possible to design environments where molecular spiders demonstrate some basic robotic behaviors [10].

When a spider is released on a surface coated with oligonucleotide substrates, it can cleave thousands of substrates before eventually detaching. The small size of spiders makes experimental observation of their motion very challenging. Atomic force microscopy imaging and single-molecule fluorescence studies have been successful to a certain degree [10], yet neither the details of the spider’s gait nor the individual paths of spiders have been resolved with sufficient certainty. Perspectives and challenges of the experimental work are surveyed in Ref. [11].

A number of modeling studies describing the motion of molecular spiders have been recently carried out. The motion of a single spider on a one-dimensional track has been investigated in Refs. [12,13]. The first article [12] ignores the difference between the substrate and the product and

makes a number of simplifying assumptions about the gait of the spiders. There are no limitations on the number of legs, however. The chief result of Ref. [12] is that the spider (which is a complicated self-interacting multileg object) can be replaced by a particle characterized by a single number, the diffusion coefficient; for the simplest gaits, the diffusion coefficient was analytically computed. In Ref. [13] we mainly considered bipedal spiders, but took into account that spiders affect the substrate (turn it into products). Despite the non-Markovian nature of the problem, the coarse-grained behavior turned out to be surprisingly simple, namely, the difference in residence times on the substrate and the product leads to the effective bias into the unvisited region.

Recent papers [14–17] utilized more detailed and complicated models mimicking the gait of spiders moving on one-dimensional tracks, the possibility of the detachment, etc. These studies numerically confirm the tendency of spiders to move into the unvisited region leaving behind the trail of the product. This key feature was observed experimentally [9,10] and proved theoretically [13] in the realm of simple models. An interesting new feature noticed in Ref. [16] is the emergence of a superdiffusive growth of the mean-square displacement  $\langle x^2(t) \rangle \sim t^\alpha$  with  $1 < \alpha < 2$ , which holds on a surprisingly long time span; eventually, the superdiffusive growth crosses over to the diffusive growth. Several rigorous results concerning the asymptotic behaviors (limit theorems, transience, recurrence, and rate of escape) of molecular spiders have been established in Refs. [18,19]. In Refs. [20,21] the motion of spiders in random environments has been studied.

As in our previous work [12,13], throughout this paper we will assume an idealized gait—the goal is not to mimic the complicated (and poorly known) gait of molecular spiders, but to qualitatively understand spiders’ macroscopic characteristics in the realm of simple models [22]. Previous theoretical analyses [12–17] have been focused on the motion of molecular spiders on one-dimensional tracks, while the goal of this work is to study a single spider moving on a two-dimensional lattice. If not stated otherwise, we tacitly

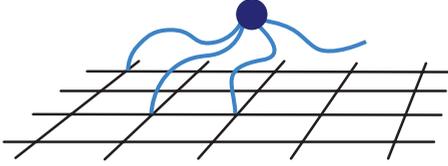


FIG. 1. (Color online) A molecular spider with four legs moving on the square lattice. The distances between any two legs should not exceed a threshold value (the maximal span  $S$ ). As long as this constraint is obeyed, each leg can hop to the empty (not occupied by another leg) nearest-neighbor site. As long as the rate of leg attachment greatly exceeds the rate of leg detachment, all legs are attached most of the time, and for relatively short time intervals one of the legs is detached, as illustrated above.

assume that the rate of attachment greatly exceeds the rate of detachment. In this case, the relative time when one leg is detached (this situation is illustrated in Fig. 1) is negligible and hence the possibility that two or more legs are detached simultaneously can be disregarded.

The following properties of molecular spiders will *always* be assumed: (1) hopping: when a leg detaches, it reattaches to a *neighboring* site; (2) exclusion: two legs cannot be attached to the same site; and (3) constraint: the distance between *any* two legs does not exceed a certain maximal span.

The restriction to the nearest-neighbor hopping can be relaxed (long-distance gaits have been probed in the one-dimensional setting [12]); the exclusion is of course the fundamental feature. The last property is the simplest constraint that assures the compactness of the spider. For one-dimensional spiders several types of constraint were considered, and the diffusion coefficients of these spiders were exactly obtained in most cases, due to mappings to exclusion processes [12].

The two-dimensional case is particularly important in current experiments. The actual situation is rather complicated, e.g., in some experiments there are a few layers of the substrate and hence a quasi-two-dimensional setting seems more appropriate; additionally, the substrates do not form a perfect square lattice. Nevertheless, we shall assume that a spider with aforementioned simple gait is placed on a square lattice [23]. We emphasize that the constraint regarding the separations between the legs roughly describes real molecular spiders. One realization of the constraint which is convenient for the numerical implementation defines the distance between points  $(x_1, y_1)$  and  $(x_2, y_2)$  via  $\max(|x_1 - x_2|, |y_1 - y_2|)$ . The neighborhood in this metric is geometrically a square, and it is often called the von Neumann neighborhood. For example, the simplest “von Neumann” spider is the bipedal spider with legs separated by distance  $S = 1$  at most. There are four possible configurations: horizontal, vertical, and two diagonal,

$$\begin{matrix} \circ & \circ & \bullet & \circ & \circ & \bullet & \bullet & \circ \\ \bullet & \bullet & \bullet & \circ & \bullet & \circ & \circ & \bullet \end{matrix}, \quad (1)$$

where  $\bullet$  represents a leg, and  $\circ$  an empty site. We found that the diffusion coefficient [24] of this spider is equal to  $1/4$ . Generally von Neumann spiders are more amenable to analysis, and we study them as well as more realistic Euclidean spiders.

The rest of this paper is organized as follows. Section II is devoted to von Neumann spiders. In Sec. III we present main results for bipedal Euclidean spiders; the detailed derivations

are given in the Appendix. An analysis of molecular spiders simplifies when the maximal span increases, and in Sec. IV we describe corresponding asymptotic behaviors. In the following sections we relax some of the assumptions about the spider gait, the influence of the spider’s motion on the environment, etc. In Sec. V we consider the influence of memory. We model the difference between the product and the substrate by postulating that the leg spends (on average) more time at newly visited sites, i.e., on the substrates. This slowdown in comparison with the motion on the products leads to faster (covering more unvisited sites) spiders; the reason for this phenomenon is an effective bias towards unvisited sites. This behavior has been observed and explained in one dimension [13], and it continues to hold in two dimensions. In Sec. VI we investigate what happens when the attachment rate is finite. We compute the mean time the spider spends on the surface and show that the probability to remain attached decays exponentially if the attachment rate greatly exceeds the detachment rate. We summarize our findings in Sec. VII.

## II. VON NEUMANN SPIDERS

A spider is quantified by lattice points  $\mathbf{r}_i = (x_i, y_i)$  with  $i = 1, \dots, L$  which describe the positions of its legs. We assume that the spider has maximal span  $S$ , so the distance between any two legs is  $\leq S$ . Each leg, when allowed, hops to neighboring sites (up, down, left, or right) at rate 1 in each direction. In this section we use the metric which assigns the distance between any two legs  $\mathbf{r}_i$  and  $\mathbf{r}_j$  according to the rule

$$|\mathbf{r}_i - \mathbf{r}_j|_\infty = \max(|x_i - x_j|, |y_i - y_j|).$$

In this metric, the neighborhood around the origin, i.e., the disk  $|\mathbf{r}|_\infty \leq S$ , is the square in Euclidean metric [see the example in Eq. (1)]; such a neighborhood is often called the von Neumann neighborhood.

A spider with maximal span  $S$  between the legs should therefore occupy the square with  $(S + 1)^2$  lattice sites; after shifting, this square becomes

$$\{(x, y) : x = 0, \dots, S; y = 0, \dots, S\}. \quad (2)$$

This spider can therefore have at most  $(S + 1)^2$  legs. Only spiders with

$$L \leq S(S + 1) \quad (3)$$

legs are mobile. More precisely, spiders with more legs than the above upper limit,  $S(S + 1) < L < (S + 1)^2$ , have a few legs which can move, but each such spider forever remains within its surrounding square (2), provided that we ignore multiple legs being detached simultaneously.

Let us first calculate the total number of possible configurations of the legs. To avoid multiple counting of configurations which are obtained by translation, we use the convention that there must be a leg both in the bottom row ( $y = 0$ ) and in the leftmost column ( $x = 0$ ), as in the example in Eq. (1). With this convention the total number of configurations is

$$\mathcal{C}(L, S) = \binom{(S + 1)^2}{L} - 2 \binom{S(S + 1)}{L} + \binom{S^2}{L}. \quad (4)$$

To establish (4) let us ignore for a moment the aforementioned convention restricting the position of one leg. The number

of such unrestricted configurations of a spider with  $L$  legs inside the  $m \times n$  rectangular is equal to  $\Omega(m, n) = \binom{mn}{L}$ . To obtain  $\mathcal{C}(L, S)$  we take the number of configurations without restriction  $\Omega(S+1, S+1)$ , subtract the number of configurations without a leg in the bottom row  $\Omega(S+1, S)$  or in the leftmost column  $\Omega(S, S+1)$ , and then add the number of configurations without a leg in both the bottom row and the leftmost column  $\Omega(S, S)$ , since we subtracted it twice. This leads to Eq. (4).

We are interested in the properties of the center of mass of the legs

$$\mathbf{R} = \frac{1}{L} \sum_{j=1}^L \mathbf{r}_j. \quad (5)$$

Due to the symmetry of the hopping rules, the mean position of the spider does not change  $\langle \mathbf{R} \rangle = 0$ , and we are interested in the variance  $\langle \mathbf{R}^2 \rangle$  of the center of mass.

### A. Bipedal spiders

Here we show that the bipedal spider with arbitrary maximal span  $S$  has the diffusion coefficient

$$D = \frac{1}{2} \left( 1 - \frac{1}{2S} \right). \quad (6)$$

The bipedal spider with  $S = 1$  corresponds to the bipedal spider with  $\ell = \sqrt{2}$  in the Euclidean version, and from Eq. (6) we indeed recover the already known result  $D = 1/4$ .

To derive (6) we start by noticing that bipedal spiders are ‘‘completely symmetric’’: In each configuration the spider hops at the same rate in each direction. For completely symmetric spiders on the square lattice, a remarkably simple formula

$$\langle \mathbf{R}^2 \rangle = \frac{\omega t}{L^2} = 4Dt, \quad \text{with} \quad D = \frac{\omega}{4L^2} \quad (7)$$

for the mean square of the center of mass, and respectively, for the diffusion coefficient of the spider, was established in Ref. [12]. In Eq. (7) we denote by  $\omega$  the total rate the spider hops, which is averaged over all stationary states. In general, for spiders with any number of legs and with symmetric hopping rates, the transitions between any two connected configurations occur at the same rates and hence all configurations have the same stationary probability  $1/\mathcal{C}$ . Therefore the total rate which is averaged over all stationary states, in short the stationary average rate, is equal to

$$\omega = \frac{1}{4\mathcal{C}} \sum_{j=1}^{\mathcal{C}} \omega_j, \quad (8)$$

where  $\omega_j$  is the total hopping rate from configuration  $j$ . Consequently, (7) can be rewritten as

$$D = \frac{1}{4L^2\mathcal{C}} \sum_{j=1}^{\mathcal{C}} \omega_j. \quad (9)$$

(Note that for the symmetric random walk on the square lattice  $L = 1$  and  $\omega = 4$ , hence the diffusion coefficient is one.)

Let us first calculate the average hopping rate to a given arbitrary direction, say to the right. Notice that in most configurations both legs can jump to the right, but in some,

only one of them. It is easier to enumerate the number of these later configurations, which is where one leg is blocked and cannot hop to the right. There is only one configuration where the two legs are next to each other in the same row  $\bullet\bullet$ , due to the convention that a leg is needed both in the bottom row and the leftmost column. In this configuration only the right leg can hop to the right. There are  $S+1$  configurations where one leg is in the bottom left corner  $(0, 0)$ , and the other is at maximal distance  $x = S$ , hence only the left leg can hop to the right.

There are further  $S$  configurations where the right leg cannot hop: one leg is in the bottom right corner  $(S, 0)$ , and the other leg is in the leftmost column, and in row  $y = 1, \dots, S$ . Hence all together there are  $1 + (S+1) + S = 2(S+1)$  configurations where only one leg can hop to the right, and  $\mathcal{C} - 2(S+1)$  where both can. Now using the fact that the average rate is the same for all four directions, Eq. (8) leads to

$$\omega = \frac{2(S+1) + 2[\mathcal{C} - 2(S+1)]}{\mathcal{C}} = 2 - \frac{1}{S}, \quad (10)$$

where we have taken into account that the number of configurations for bipedal spiders is  $\mathcal{C} = 2S(S+1)$ , as follows from Eq. (4). Substituting (10) into Eq. (7) we arrive at the diffusion coefficient (6).

### B. Multipedal spiders

It is much more challenging to compute the diffusion coefficient for multileg spiders. The chief reason is that spiders with more than two legs are not completely symmetric and hence one cannot use (7). To show the lack of symmetry it suffices to provide a configuration where the left and right hopping rates of the spider are different. Consider for example a tripod with two legs being in the same column and the third leg being at the maximal distance  $S$  in the  $x$  direction from both of the other legs, as illustrated here:

$$\begin{array}{c} \bullet \leftarrow S \rightarrow \bullet \\ \bullet \end{array} \quad (11)$$

This spider can hop to the right at rate two and to the left at rate one. (More precisely, both legs on the left can hop to the right, but not to the left; the leg on the right can hop to the left, but not to the right.) Similar configurations can be easily constructed for any spiders with  $L \geq 3$  legs.

The matrix method described in Ref. [12] can be generalized to arbitrary dimension and in principle it allows one to analytically determine the diffusion coefficient for any spiders with sufficiently small number of legs. However, even in the simplest examples the exact calculations are rather cumbersome. For instance, even for the simplest  $S = 2$  tripod on the square lattice the number of configurations is equal to 48 [see Eq. (4)], so the computation of  $D$  leads to the necessity to diagonalize a  $48 \times 48$  matrix; obtaining this matrix is very laborious.

We performed simulations for the simplest spiders to measure their diffusion coefficients. This quantity is relatively easy to measure by probing the asymptotic of the mean-square displacement. It turns out that the correction to the true

TABLE I. Simulation results for the diffusion coefficients  $D$  for spiders with  $L$  legs and constraint  $S$ . For the random walk ( $L = 1$ ) the diffusion coefficient is  $D = 1$ ; for bipedal spiders ( $L = 2$ ) we have exact results given by Eq. (6).

	$S = 1$	$S = 2$	$S = 3$
$L = 1$	1	1	1
$L = 2$	0.25	0.375	0.417
$L = 3$		0.191	0.241
$L = 4$		0.0972	0.152
$L = 5$		0.0464	0.104

asymptotic decays is rather small, namely,

$$\langle \mathbf{R}^2 \rangle \approx 4Dt \left( 1 + \frac{a}{t} \right) \quad (12)$$

for multipedal spiders. (Similar corrections were observed in one dimension.) Note that there are no corrections at all for continuous time random walks (the monopodal spiders). The measured values of  $D$  are summarized in Table I.

To detect the motion of a single spider is still experimentally impossible, while various techniques allow one to count the total number of visited sites [9]. Since numerous spiders are usually released [9], dividing the actually observed total number of visited sites by the number of spiders makes the *mean* number of sites visited by a spider accessible. In the limit when the density of spiders is very low, the same result will emerge if we take a single spider and average the number of different visited sites over many realizations.

For a symmetric random walk which hops at rate one in each direction (so its diffusion coefficient is  $D = 1$ ), the mean number of different sites visited by the random walk scales as [25,26]

$$\langle N \rangle = \frac{At}{\ln Bt} + \mathcal{O}\left[\frac{t}{(\ln t)^3}\right] \quad (13)$$

in the large time limit. The amplitudes are  $A = 4\pi$  and  $B = 32 \exp(C_E - 1) = 20.97 \dots$  where  $C_E = 0.577 215 \dots$  denotes Euler's constant. (In Ref. [25], the random walk with diffusion coefficient  $D = 1/4$  has been analyzed; to recast the prediction of [25] to our setting, where  $D = 1$ , we rescaled time by a factor 4.)

We simulated the motion of different spiders and we obtained the same asymptotic behavior (13), with  $A$  and  $B$  depending on the number of legs  $L$ , and the constraint  $S$ . We are mainly interested in the leading order behavior  $A$ , however, fitting also the next correction  $B$  is unavoidable in order to get an estimate for  $A$ , due to the large subleading corrections. The results for the coefficient  $A$  are summarized in Table II.

One can see from Tables I and II that, according to simulations,  $D(L, S)$  and  $A(L, S)$  are monotonically decreasing functions of  $L$ , and monotonically increasing functions of  $S$  for multipedal spiders. The simplest conjecture is that asymptotically the spider is indistinguishable from the random walk on the square lattice. Mathematically, this would imply that  $A = 4\pi D$ . Simulation results for the diffusion coefficient show that  $A$  is slightly different than that.

TABLE II. Simulation results for the amplitude  $A$  in the asymptotic law for the number of visited sites  $\langle N \rangle = At / \ln t$ . Spiders have  $L$  legs and constraint  $S$ . For the random walk  $L = 1$  and the amplitude is  $A = 4\pi$ .

	$S = 1$	$S = 2$	$S = 3$
$L = 1$	12.57	12.57	12.57
$L = 2$	3.21	4.83	5.35
$L = 3$		2.38	3.20
$L = 4$		1.30	2.08
$L = 5$		0.627	1.44

### III. EUCLIDEAN SPIDERS

In this section we consider the more realistic Euclidean spiders. Thus we use the standard Euclidean metric to measure the distance between the legs:

$$|\mathbf{r}_i - \mathbf{r}_j| = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}.$$

The simplest Euclidean spider is bipedal with legs separated by maximal distance  $\ell = \sqrt{2}$ . This spider is identical to the simplest  $L = 2$ ,  $S = 1$  von Neumann spider; it is characterized by four configurations (1), and it has diffusion coefficient  $D = 1/4$ .

For the bipedal spider with maximal separation  $\ell = 2$ , there are six different configurations—four configurations (1) and two additional configurations, the horizontal configuration  $\bullet \circ \bullet$  and its vertical cousin. The computation of the diffusion coefficient gives  $D = 1/4$  [see Eq. (A1) in the Appendix], so this spider has the same diffusion coefficient as the previous one. This surprising result is a coincidence rather than a rule. We computed diffusion coefficients for bipedal spiders with many other maximal distances. For  $\ell = \sqrt{5}$ , four new configurations (with legs separated by the move of a knight in chess) appear. Next, change occurs for  $\ell = 2\sqrt{2}$  when two new “long” diagonal configurations arise. Then for  $\ell = 3$ , the horizontal configuration  $\bullet \circ \circ \bullet$  and analogous vertical configuration become possible. Varying  $\ell$  up to  $\sqrt{50}$ , the total number of distinct allowed configurations  $\mathcal{C}$  and the diffusion coefficient exhibit the behaviors summarized in Table III.

The diffusion coefficient tends to increase with  $\ell$ , yet its behavior is somewhat erratic and it can occasionally decrease ( $\frac{5}{14} < \frac{3}{8}, \frac{19}{48} < \frac{9}{22}, \frac{17}{40} < \frac{29}{68}, \frac{49}{112} < \frac{4}{9}, \frac{33}{74} < \frac{65}{144}$ , etc.). The last column reveals remarkable hidden regularities—the quantity  $(1/2 - D)|\mathcal{C}|$  is always half-integer with equilibrium patches of increasing length punctuated by upward jumps by  $1/2$ . These intriguing observations are explained by the neat general formula

$$D = \frac{1}{2} \left[ 1 - \frac{[\ell] + 1}{\mathcal{C}} \right], \quad (14)$$

where  $[\ell]$  is the integer part of  $\ell$ .

The derivation of Eq. (14) is somewhat lengthy (Appendix), but it is just an application of general formula (9) for the diffusion coefficient. The dependence of  $\mathcal{C}$  and rates on  $\ell$  is nontrivial and cannot be deduced analytically for an arbitrary  $\ell$ . Indeed, the problem of counting the total number of configurations is equivalent to the problem of computing  $\mathcal{N}_\ell$  which gives the total number of lattice sites within the disk of

TABLE III. Total number of configurations and diffusion coefficients for Euclidean spiders.

$\ell$	$ C $	$D$	$(1/2 - D)C$
$\sqrt{2}$	4	1/4	1
2	6	1/4	3/2
$\sqrt{5}$	10	7/20	3/2
$\sqrt{8}$	12	3/8	3/2
3	14	5/14	2
$\sqrt{10}$	18	7/18	2
$\sqrt{13}$	22	9/22	2
4	24	19/48	5/2
$\sqrt{17}$	28	23/56	5/2
$\sqrt{18}$	30	5/12	5/2
$\sqrt{20}$	34	29/68	5/2
5	40	17/40	3
$\sqrt{26}$	44	19/44	3
$\sqrt{29}$	48	7/16	3
$\sqrt{32}$	50	11/25	3
$\sqrt{34}$	54	4/9	3
6	56	49/112	7/2
$\sqrt{37}$	60	53/120	7/2
$\sqrt{40}$	64	57/128	7/2
$\sqrt{41}$	68	61/136	7/2
$\sqrt{45}$	72	65/144	7/2
7	74	33/74	4
$\sqrt{50}$	84	19/42	4
$\gg 1$	$\approx \pi \ell^2/2$	$\approx 1/2 - (\pi \ell)^{-1}$	$\approx \ell/2$

radius  $\ell$ ; more precisely,  $C = (\mathcal{N}_\ell - 1)/2$ . The investigation of  $\mathcal{N}_\ell$  constitutes the celebrated Gauss problem [27]. Of course,  $\mathcal{N}_\ell$  approximately grows as the area,  $\mathcal{N}_\ell \approx \pi \ell^2$ ; the deviation from this dominant growth law is extremely difficult to probe analytically. (The subleading asymptotic is unknown. More precisely, the proven upper bound on the growth of the subleading asymptotic is substantially weaker than the conjectural one; proving the conjectural asymptotic is known to be equivalent to proving the Riemann conjecture.) This subtlety is irrelevant as long as we are satisfied with the leading asymptotic. In our problem the case of large  $\ell$  is particularly simple as both legs essentially diffuse independently and therefore the diffusion coefficient of the center of mass is very close to the half of the diffusion coefficient of the random walk. Thus  $D_\ell \rightarrow 1/2$  when  $\ell \rightarrow \infty$ . Note that  $D < 1/2$  for any two-leg spider. Indeed, since  $\omega_j \leq 8$  (the equality occurs when each leg is allowed to hop to each of the four neighbors), the sum on the right-hand side of Eq. (9) cannot exceed  $8 \times C$  which proves  $D < 1/2$ . The derivation of Eq. (14) and a detailed description of the configurations for the threshold up to  $\ell = \sqrt{50}$  are given in the Appendix.

#### IV. SPIDERS WITH LARGE MAXIMAL SPAN

Interactions between legs (due to exclusion and the maximal span constraint) imply that molecular spiders are complicated self-interacting objects. Therefore it is very difficult to compute the dependence of the diffusion coefficient  $D(L, S)$  of a spider on the number of legs  $L$  and the maximal span  $S$ . The behavior of  $D(L, S)$  simplifies when the maximal span

$S$  becomes large. In the  $S \rightarrow \infty$  limit, typical separations between legs grow with time thereby making exclusion asymptotically negligible. Therefore in this limit we can treat legs as noninteracting random walkers. The diffusion coefficient of such spider is  $D(L, S = \infty) = 1/L$ . In this section we derive this result and then argue that the finite  $S$  correction has the  $1/S$  form.

##### A. Noninteracting legs

Consider a spider with  $L$  noninteracting legs. Each leg performs a random walk with hopping rates one in each direction. The mean and mean-square displacement for each leg read

$$\langle \mathbf{r}_j \rangle = 0, \quad \langle \mathbf{r}_j^2 \rangle = 4t, \quad j = 1, \dots, L, \quad (15)$$

where we have assumed that initially all legs are at the origin. (We continue to assume that the spider moves on the square lattice; generally on the cubic lattice in  $d$  dimensions the amplitude on the right-hand side of Eq. (15) is given by  $2d$ .) Using (15) one computes the variance of the center of mass (5) of the spider:

$$\langle \mathbf{R}^2 \rangle = L^{-2} \sum_{j=1}^L \langle \mathbf{r}_j^2 \rangle = L^{-1} 4t,$$

implying that the diffusion coefficient  $D(L)$  of the spider with  $L$  noninteracting legs (each performing the random walk with diffusion coefficient  $D \equiv 1$ ) is

$$D(L) = \frac{1}{L}. \quad (16)$$

This result is remarkably universal: It is valid in any dimension and it also does not depend on the lattice; the only requirement is the absence of bias.

##### B. Interacting legs

We now return to our original spiders which move on a lattice and obey two rules: (i) Two legs cannot be attached to the same site; and (ii) the distance between any two legs does not exceed  $S$ . In the  $S \rightarrow \infty$  limit the legs are asymptotically noninteracting. Therefore Eq. (16) implies that  $D(L, S = \infty) = 1/L$ .

The nontrivial task is to compute the leading correction which describes the deviation from Eq. (16) in the situation when the maximal span  $S$  is large, but finite. To guess the  $S$  dependence of the leading correction let us look at known *exact* results for spiders moving on the one-dimensional lattice. For the bipedal spider the diffusion coefficient is given by Ref. [12]

$$D(2, S) = \frac{1}{2} \left( 1 - \frac{1}{S} \right). \quad (17)$$

More generally in one dimension for multileg spiders with nearest-neighbor hopping and the constraint on the maximal span (spiders with global constraint in terminology of Ref. [12]), the diffusion coefficient reads [12]

$$D(L, S) = \frac{1}{L} \left( 1 - \frac{L-1}{S} \right). \quad (18)$$

This formula is valid for  $L = 1$  (the random walk) and all  $L \geq 2$ ,  $S \geq L - 1$ .

Equations (17) and (18) show that at least in one dimension the leading correction has the  $S^{-1}$  dependence on the maximal span. We now demonstrate the universality of this behavior by analyzing the bipedal spider in one dimension with a more complicated gate. Namely, let us assume two types of hops: the nearest-neighbor  $\pm 1$  hops occur with rate  $p$  each, and the next-nearest-neighbor  $\pm 2$  hops occur with rate  $q$  each. For a random walk with such gate, the diffusion coefficient is  $D = p + 4q$ , and since we always normalize this diffusion coefficient of the random walk to unity, we set  $1 = p + 4q$ . Using methods of Ref. [12], one obtains

$$D(2, S; p) = \frac{1}{2} \left( 1 - \frac{2-p}{S} \right), \quad (19)$$

showing that the  $S^{-1}$  behavior is universal, namely, it is insensitive to the details of the gate.

In two dimensions all known results also agree with the  $S^{-1}$  behavior of the leading correction [28]. The exact diffusion coefficient for the von Neumann bipedal spider (6) is in perfect agreement (no other corrections). For the Euclidean bipedal spider the behavior is more complicated [see (14)]; the asymptotic behavior is very simple:

$$D(2, \infty) - D(2, S) \simeq \frac{1}{\pi S}, \quad (20)$$

and it agrees with the  $S^{-1}$  asymptotic. Thus the exact results (6), (17)–(19), and the asymptotic (20) suggest the following conjectural asymptotic behavior:

$$D(L, \infty) - D(L, S) \simeq \frac{A_d(L)}{S} \quad \text{as } S \rightarrow \infty. \quad (21)$$

Intuitively, the  $1/S$  correction stems from the ratio of spider-leg configurations where the constraint is relevant (the distance of two or more legs is  $S$ ) to the total number of such configurations. The total number of configurations is essentially a volume of a domain of characteristic size  $S$ , while the number of configurations for which the constraint is relevant is the surface area of that domain, so the ratio is indeed proportional to  $1/S$ .

We have presented evidence in favor of the conjectural behavior (21) in one and two dimensions, but it is probably valid in arbitrary dimension  $d$ . The limiting diffusion coefficient  $D(L, \infty) = 1/L$  is universal, while the amplitude  $A_d(L)$  in the subleading term depends not only on the number of legs, but also on the spatial dimension  $d$ , and on the details of the gate. In all known examples the amplitude  $A_d(L)$  is positive, e.g., according to Eq. (18) the amplitude is  $A_1(L) = 1 - L^{-1}$  in one dimension. The positivity of  $A_d(L)$  is physically evident (the constraint on the maximal span makes spiders less motile), although it is not clear how to prove this positivity. Finally we note that simulation results for diffusion coefficients presented in Table I are in surprisingly good agreement with the conjectural asymptotic (21), namely, the quantity  $S[D(L, \infty) - D(L, S)]$  already changes very little when  $S$  increases from 2 to 3.

## V. NON-MARKOVIAN EFFECTS

In this section we continue to assume that the reattachment of a leg is instantaneous. In contrast to Secs. II and III, however, we take into account the effects of memory associated with previous visits of the legs. These effects are unavoidable in most experimental realizations—the first time a leg visits a site (an uncleaved substrate), the leg hops from this site only after it has cleaved it into a product. (The product is unaffected by future visits, so we only need to know if the site has been visited in the past or not.) Thus the motion of the synthetic molecular motor, the molecular spider, irreversibly changes the environment making the problem non-Markovian. Intriguingly, one natural molecular motor, a special protein called collagenase which moves along collagen fibrils, exhibits even stronger irreversible effect on its one-dimensional track and undergoes a biased diffusion [30–32].

Our main interest is the leading order behavior of two quantities: the mean square position  $\langle \mathbf{R}^2 \rangle \sim 4Dt$ , i.e., the diffusion coefficient, and the mean number of visited sites  $\langle N \rangle \sim At / \ln t$ . We start with the one-leg spider, the random walk, where the effect of memory is asymptotically negligible, yet the corrections to the leading behaviors are qualitatively similar to those which arise for multilegged spiders.

The results reported in this section are mostly numerical. In two dimensions, all simulations are made for von Neumann spiders. The convergence to the true asymptotic behaviors is slow, especially in two dimensions where it is logarithmically slow. Therefore to extract accurate numerical predictions for the diffusion coefficient  $D$  and the amplitude  $A$  we need to know functional forms of correction terms. We make use [in Eqs. (22)–(24)] of correction terms which provide good fit to the data. The functional forms of these leading correction terms are still lacking theoretical justification.

### A. Random walk, $L = 1$

We consider the random walk which hops symmetrically to nearest-neighbor sites and changes substrates into products. As in Ref. [13], the change of the substrates to the products (which occurs after the leg first visits the site) is modeled by postulating that the hopping rate to each neighbor is equal to 1 for the site which has not been visited before (the substrate), and to  $r$  for the site which has been visited in the past. Thus the hopping rate is determined by the state of the issuing site, but not by the state of the target site. Mathematically, the parameter  $r$  can be an arbitrary positive number,  $r > 0$ ; in experiments, the detachment from the product is easier, so  $0 < r < 1$  as we shall assume in the following.

According to simulations, the memory has no effect on a random walk in the leading order. Independently of the value of  $r$ , we obtained  $D = 1$  for the diffusion coefficient, and  $A = 4\pi$  for the amplitude. The same universality was numerically observed in one dimension [13]; for the amplitude  $A$ , the lack of dependence on  $r$  was established analytically [13].

To extract the leading order behavior from the simulation data one has to investigate higher order corrections as well. Our numerical simulations indicate that the mean-square position

has the following correction terms in the presence of memory:

$$\langle \mathbf{R}^2 \rangle \approx \begin{cases} 2Dt[1 - a_1(r)/\sqrt{t}] & : d = 1 \\ 4Dt[1 - a_2(r)/\ln t] & : d = 2, \end{cases} \quad (22)$$

while there are no corrections at all in the absence of memory  $r = 1$ ; that is,  $a_d(1) = 0$ . The large corrections in two dimensions are especially important in determining the diffusion coefficient. Fitting our data to Eq. (22) we got  $D = 1$  both in one and two dimensions. The amplitudes of the correction terms are positive,  $a_d(r) > 0$  when  $0 < r < 1$ , and monotonically decreasing functions of  $r$ . Therefore the memory slows down the random walk in the next to leading order. This behavior is understandable, as the random walk slows down at newly visited sites.

For the average number of sites  $\langle N \rangle$  visited during the time interval  $(0, t)$ , the corrections to the leading asymptotic are also more important (since they vanish much more slowly with time) in two dimensions. According to our simulations,

$$\langle N \rangle \approx \begin{cases} A_1(r) \sqrt{t}[1 - c_1(r)/\sqrt{t}] & : d = 1 \\ \frac{A_2(r)t}{\ln t} [1 - c_2(r)/\ln t] & : d = 2. \end{cases} \quad (23)$$

The correction amplitudes are again positive,  $c_d(r) > 0$  when  $0 < r < 1$ , and monotonically decreasing functions of  $r$ . Fitting our data to (23) we found that the amplitudes  $A_d(r)$  do not depend on  $r$  for the random walk. [Equation (23) also describes the average number of sites visited by a spider, and in that situation the amplitudes  $A_d(r)$  do depend on  $r$ .] Note that in two dimensions we found the same type of asymptotic behavior both with and without memory[ see Eq. (13)], and for an arbitrary number of legs.

### B. Multipedal molecular spiders, $L > 1$

In order to investigate the effects of memory we numerically studied the two simplest spiders: the bipedal  $L = 2$ ,  $S = 1$  spider, and the tripod  $L = 3$ ,  $S = 2$ .

We have seen in Sec. II that when memory effects are ignored, the corrections to the diffusion coefficient  $D$  quickly decay with time (12). This feature makes  $D$  easy to measure. In the presence of memory, however, a much slower decaying correction term appears in two dimensions:

$$\langle \mathbf{R}^2 \rangle \approx \begin{cases} 2D_1(r)t[1 - a_1(r)/\sqrt{t}] & : d = 1 \\ 4D_2(r)t[1 - a_2(r)/(\ln t)^2] & : d = 2. \end{cases} \quad (24)$$

In one dimension, the correction also decays slower with time in the presence of memory, yet it remains algebraic; the qualitative behavior is similar to the random walk with memory [see (22)]. The  $(\ln t)^{-2}$  type correction observed in two dimensions is unusual; it is rather slow of course, yet it is faster than for the random walk.

Using (24) and (23) we extract the diffusion coefficient  $D(r)$  and the amplitude  $A(r)$ ; their dependencies on  $r$  are displayed [29] in Figs. 2 and 3. Surprisingly, the slowdown of legs at new sites leads to larger diffusion coefficients and increases the average number of visited sites in the large time limit. The qualitative reason is that the slow leg at new sites keeps the other legs close to newly visited sites, which generates an effective bias toward new sites, and thereby it leads to

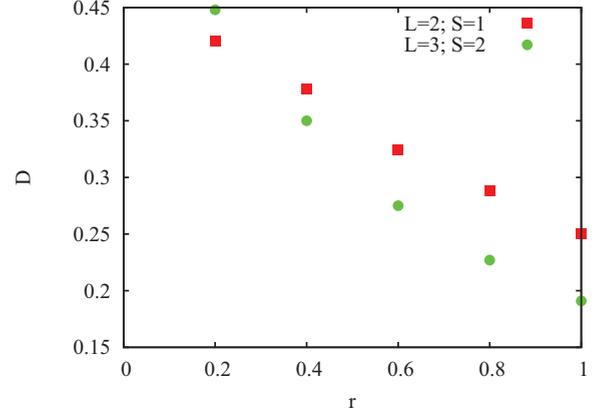


FIG. 2. (Color online) Diffusion coefficient  $D$  describing the asymptotic behavior of the mean-square position of a spider  $\langle \mathbf{R}^2 \rangle \sim 4Dt$  as a function of the memory parameter  $r$ . Simulation data are shown for the simplest bipedal spider with  $S = 1$ , and for the simplest tripod with  $S = 2$ . The no memory case corresponds to  $r = 1$ .

the increase in the number of visited sites. This effect has already been observed, and quantitatively understood, for one-dimensional spiders [13]. Although here we talk about the asymptotic behavior, the actual number of mean visited sites also becomes larger for smaller values of  $r$  after about a few hundred time units.

An interesting new feature of two-dimensional spiders is that the tripod is somewhat more sensitive to the slowdown at new sites than the bipedal spider, as  $D$  changes more rapidly with changing  $r$  (see Fig. 2). Another remark is that, as one can see in Figs. 2 and 3, with large enough slowdown ( $r \lesssim 0.6$ ) the tripod becomes faster than the bipedal spider without memory ( $r = 1$ ).

An intriguing feature of these two-dimensional spiders is that for sufficiently small  $r$  tripods become more motile than bipedal spiders (see Fig. 2). A possible reason for this effect is that the spider with more legs sticks to the domain of new sites more efficiently.

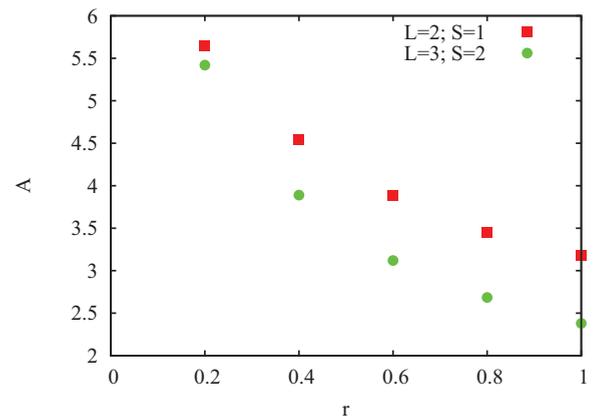


FIG. 3. (Color online) Amplitude  $A$  describing the asymptotic behavior of the number of visited sites as  $\langle N \rangle \sim At/\ln t$  as a function of the memory parameter  $r$ . Simulation data are shown for the simplest bipedal spider with  $S = 1$ , and for the simplest tripod with  $S = 2$ .

## VI. UNBINDING OF SPIDERS

All our previous analyses have relied on the assumption that the reattachment of a leg is instantaneous, so the process is controlled by detachment. This implies that spiders remain fully attached and never leave the surface. If the reattachment takes time, the problem becomes more complicated but not necessarily intractable—for molecular motors [33], for instance, some analyses allow complete detachment (unbinding) from cytoskeletal filaments [34–40].

In this section we treat a more restricted problem, namely, we compute the probability that a spider remains attached. We disregard memory and lattice effects and focus on the attachment-detachment process.

Consider a spider with  $L$  legs. Let  $\Pi_n(t)$  be the probability that  $n$  of its legs are attached. To simplify notation we set the detachment rate equal to unity; we shall disregard the difference between the substrate and the product. We denote the attachment rate by  $\lambda$ . The probabilities  $\Pi_n(t)$  with  $n = 1, \dots, L$  obey

$$\frac{d\Pi_n}{dt} = (n+1)\Pi_{n+1} - n\Pi_n - \lambda(L-n)\Pi_n + \lambda(L-n+1)\Pi_{n-1}, \quad (25)$$

where the terms on the right-hand side in the top line account for detachment and the terms in the bottom line describe reattachment. Equation (25) remains applicable to extreme probabilities  $\Pi_1$  and  $\Pi_L$  if we set

$$\Pi_0(t) \equiv 0, \quad \Pi_{L+1}(t) \equiv 0. \quad (26)$$

The latter relation is obvious (by definition, the spider has  $L$  legs); the former relation is actually the assumption that if all legs are detached, the spider leaves the surface and never reattaches to it. The initial condition is

$$\Pi_n(0) = \delta_{n,L} \quad (27)$$

if we imagine that the spider is initially fully attached.

An analysis of the initial-boundary value problems (25)–(27) is rather straightforward, so we only present one asymptotic result which is valid in the most interesting limit of quick reattachment,  $\lambda \gg 1$ . In this limit we found that, for any  $L$ , the probability  $\Pi(t) = \sum_{n=1}^L \Pi_n(t)$  that the spider remains attached at time  $t$  is exponential:

$$\Pi = \exp(-Lt/\lambda^{L-1}). \quad (28)$$

Equation (28) shows that the mean time for the spider to detach is

$$\langle t_L \rangle = \frac{\lambda^{L-1}}{L} \quad (29)$$

in the  $\lambda \gg 1$  limit.

If we were only interested in the mean detachment time, we could have determined it exactly without computing the probabilities  $\Pi_n(t)$ . Indeed, utilizing an exact solution for the adsorption time for the so-called one-step process [41], one finds an exact expression

$$\langle t_L \rangle = \sum_{a=1}^L \lambda^{L-a} \sum_{b=1}^a \frac{(L-b)!(b-1)!}{(a-b)!(L-a+b)!}, \quad (30)$$

which is a polynomial in  $\lambda$ . Keeping only the dominant  $a = 1$  term we see that (30) reduces to (29). Displaying a few more terms we obtain

$$\langle t_L \rangle = \lambda^{L-1} \frac{1}{L} + \lambda^{L-2} \left[ 1 + \frac{1}{L(L-1)} \right] + \lambda^{L-3} \left[ \frac{L-1}{2} + \frac{1}{L-1} + \frac{2}{L(L-1)(L-2)} \right] + \dots$$

The above analysis is “zero dimensional” as we ignored the lattice. For instance, consider a tripod in the configuration  $\bullet \bullet \bullet$  and imagine that one of its legs detaches. If an extreme leg detaches, its reattachment rate is clearly larger than the reattachment rate of the middle leg. In Eq. (25) this feature is ignored. Therefore combining detachment with diffusion makes the problem very complicated even in the absence of memory effects. A numerical analysis of models which take into account the possibility of the detachment has been undertaken in Refs. [14,15]. For the bipedal molecular spider, the one-dimensional version of this problem could be tractable as long as the memory effects are ignored.

## VII. SUMMARY

We investigated the motion of a single molecular spider on the square lattice. The limit when the motion of the spider does not affect the environment is tractable for bipedal spiders, while spiders with more than two legs remain very challenging for analytical work. For bipedal spiders we computed the diffusion coefficient with an arbitrary maximal span between the legs. Generally, the increase in the maximal span leads to the increase in the diffusion coefficient. This phenomenon strictly holds for one type of bipedal spiders (von Neumann), while for Euclidean bipedal spiders an increase of the span can sometimes, unexpectedly, decrease the diffusion coefficient.

We explored the behavior of von Neumann spiders with more than two legs by means of numerical simulations. In general we found that the increase of the number of legs makes spiders less motile. The increase of the maximal span  $S$  (maximal allowed distance between the legs), on the other hand, makes spiders more motile. In the infinite span limit the diffusion coefficient is reciprocal to the number of legs,  $D(L, S = \infty) = 1/L$ . We argued that the leading large  $S$  correction to the diffusion coefficient is proportional to  $1/S$ . We also considered the effect of spiders completely unbinding from the substrate, and we found that the time it takes grows exponentially with the number of legs. The reason is that a spider unbinds only if all of its legs simultaneously unbind.

In experimental realizations, the legs usually convert substrates into shorter products that have a lower affinity for the legs. Assuming that the substrate turns into the product after the first visit of a leg, we investigated this non-Markovian problem numerically. We showed that the long-time behavior is diffusive in character. More precisely, we demonstrated that the mean-square displacement grows as  $t$  and the total number of distinct visited sites grows as  $t/\ln t$ . The amplitudes are affected by memory. Furthermore, the non-Markovian nature of the problem leads to very large subleading corrections. For instance, the relative magnitude of the subleading correction to the mean-square displacement decays with time in a very slow  $(\ln t)^{-2}$  manner.

The most surprising influence of memory is that the more spiders are slowed down on unvisited sites, the more motile spiders become. An explanation of this very counterintuitive behavior is that the “stickiness” to unvisited sites generates an effective bias toward unvisited sites, which results in the increase of the visited area. For example, although without memory an  $L = 3$ ,  $S = 2$  spider is slower than an  $L = 2$ ,  $S = 1$  spider, with the memory effect the three-leg spider can visit more sites on average.

### ACKNOWLEDGMENTS

We thank Darko Stefanovic and Milan Stojanovic for helpful discussions and suggestions. P.L.K. acknowledges financial support from NSF Grant No. CCF-0829541.

### APPENDIX: EUCLIDEAN BIPEDAL SPIDERS

For small  $\ell$ , we can count various configurations and compute their rates by hand. Here we record what happens when  $\ell$  varies up to  $\sqrt{50}$ . Configurations that arise when  $\ell$  varies up to  $\sqrt{20}$  are depicted below.

#### 1. $\sqrt{2} \leq \ell < 2$

There are four different configurations. The rates are  $\omega_1 = \omega_2 = \omega_3 = \omega_4 = 4$  and hence (9) gives  $D = 1/4$ .

#### 2. $2 \leq \ell < \sqrt{5}$

Two additional configurations are  $\bullet \circ \bullet$  and its vertical cousin, so that  $\mathcal{C} = 6$ . We have

$$\omega(\bullet \bullet) = 6, \quad \omega(\bullet \circ \bullet) = 2, \quad \omega_{\text{diag}} = 8,$$

and therefore

$$D = \frac{1}{16} \frac{1}{6} [6 + 2 + 4] \times 2 = \frac{1}{4}. \quad (\text{A1})$$

Surprisingly, the diffusion coefficient is the same as for the previous two-leg spider.

#### 3. $\sqrt{5} \leq \ell < 2\sqrt{2}$

Four more configurations become possible:  $\mathcal{C} = 10$ . The rates of these new knight configurations (Fig. 4) are  $\omega_{\text{knight}} = 4$ . The rates of the previous configurations are

$$\omega(\bullet \bullet) = \omega(\bullet \circ \bullet) = 6, \quad (\text{A2})$$

while the diagonal configurations reach the maximal possible rate  $\omega_{\text{diag}} = 8$ . Thus

$$D = \frac{1}{16} \frac{1}{10} [6 \times 4 + 8 \times 2 + 4 \times 4] = \frac{7}{20}.$$

#### 4. $2\sqrt{2} \leq \ell < 3$

Two new diagonal configurations (Fig. 4) of length  $2\sqrt{2}$  become possible; overall  $\mathcal{C} = 12$ . The rates remain the same [Eq. (A2)] for horizontal and vertical configurations of both kinds, and for short diagonal configurations ( $\omega_{\text{diag}} = 8$ ). The rates increase for knight configurations:  $\omega_{\text{knight}} = 6$ . Finally for the new diagonal configurations (Fig. 4) the rates are  $\omega_{\text{diag}}^{(2)} = 4$ .

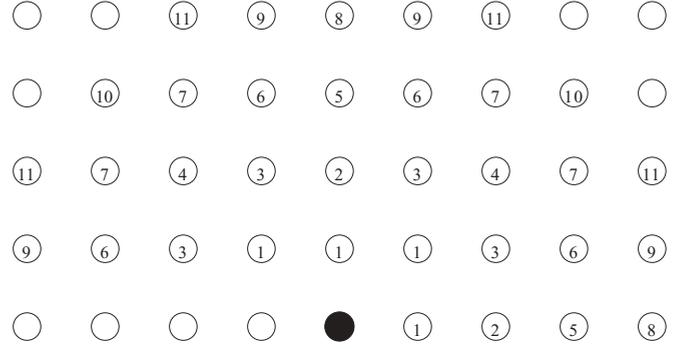


FIG. 4. Nonisomorphic configurations for  $\ell = \sqrt{20}$ . Configurations share the center site (filled disk) while another leg is in a labeled disk. The label counts the “bifurcation” event when such configurations first appear (the first bifurcation is identified with  $\ell = \sqrt{2}$ ). For instance, the simplest knight configurations appear in the third bifurcation event (when  $\ell$  passes through  $\sqrt{5}$ ) and hence the corresponding label is 3.

Thus

$$D = \frac{1}{16} \frac{1}{12} [6 \times 4 + 8 \times 2 + 6 \times 4 + 4 \times 2] = \frac{3}{8}.$$

Note that as the number of configurations from 6 to 10 to 12, the diffusion coefficient also gets larger.

#### 5. $3 \leq \ell < \sqrt{10}$

The rates of horizontal configurations are

$$\omega(\bullet \bullet) = 6, \quad \omega(\bullet \circ \bullet) = 8, \quad \omega(\bullet \circ \circ \bullet) = 2, \quad (\text{A3})$$

and similarly for the vertical one. The other rates are

$$\omega_{\text{diag}} = 8, \quad \omega_{\text{diag}}^{(2)} = 4, \quad \omega_{\text{knight}} = 6. \quad (\text{A4})$$

The diffusion coefficient  $D = 5/14$  is *smaller* than the diffusion coefficient  $D = 3/8$  that characterizes spiders with  $2\sqrt{2} \leq \ell < 3$ .

#### 6. $\sqrt{10} \leq \ell < \sqrt{13}$

Four new knight configurations obtained by hopping one lattice spacing in one direction and three lattice spacings in orthogonal direction arise. The rates which change compared to rates (A3) and (A4) are

$$\omega(\bullet \circ \circ \bullet) = 6, \quad \omega_{\text{knight}} = 8, \quad \omega_{\text{knight}}^{(2)} = 4.$$

The diffusion coefficient is  $D = 7/18$ .

#### 7. $\sqrt{13} \leq \ell < 4$

Four new knight configurations obtained by hopping two lattice spacings in one direction and three lattice spacings in orthogonal direction arise. The rates which change are

$$\omega_{\text{diag}}^{(2)} = 8, \quad \omega_{\text{knight}}^{(2)} = 6, \quad \omega_{\text{knight}}^{(3)} = 4.$$

The diffusion coefficient is  $D = 9/22$ .

#### 8. Large $\ell$

We now outline following bifurcations. We classify various configurations into linear, diagonal, and knight. Each linear

TABLE IV. The rates of the linear and diagonal configurations for Euclidean spiders in the  $5 \leq \ell \leq 7$  window.

$\ell$	1	2	3	4	5	6	7	(1,1)	(2,2)	(3,3)	(4,4)
5	6	8	8	8	2	0	0	8	8	8	0
$\sqrt{26}$	6	8	8	8	6	0	0	8	8	8	0
$\sqrt{29}$	6	8	8	8	6	0	0	8	8	8	0
$\sqrt{32}$	6	8	8	8	6	0	0	8	8	8	4
$\sqrt{34}$	6	8	8	8	6	0	0	8	8	8	4
6	6	8	8	8	8	2	0	8	8	8	4
$\sqrt{37}$	6	8	8	8	8	6	0	8	8	8	4
$\sqrt{40}$	6	8	8	8	8	6	0	8	8	8	4
$\sqrt{41}$	6	8	8	8	8	6	0	8	8	8	8
$\sqrt{45}$	6	8	8	8	8	6	0	8	8	8	8
7	6	8	8	8	8	8	2	8	8	8	8

configuration is either horizontal or vertical. There are also two kinds of diagonal configurations, while knight configurations have four different types. The rates of linear configurations and diagonal configurations [by  $(d, d)$  we denote the diagonal configuration of length  $d\sqrt{2}$ ] change with their size. Some of these rates are collected in Table IV (we present results in the window  $5 \leq \ell \leq 7$ ; overall we analyzed  $\ell \leq 11$ ).

In Table V we summarize how the rates of the knight configurations ( $nm$  denotes the knight configuration obtained by hopping  $n$  lattice spacings in one direction and  $m$  lattice spacings in another) vary with  $\ell$ .

We now give some explanations. The behavior of the quantity  $\Phi \equiv (1/2 - D)\mathcal{C}$  (see Table III) is easy to understand. First, using Eq. (9) we can rewrite  $\Phi$  as

$$\Phi = \frac{1}{2}\mathcal{C} - \frac{1}{16}\mathcal{R}, \quad \mathcal{R} = \sum_{j=1}^{\mathcal{C}} r_j. \quad (\text{A5})$$

TABLE V. The rates of the knight configurations for Euclidean spiders in the  $5 \leq \ell \leq 7$  window.

$\ell$	21	31	32	41	42	43	51	52	53	54	61	62	63
5	8	8	8	6	6	4	0	0	0	0	0	0	0
$\sqrt{26}$	8	8	8	8	6	4	4	0	0	0	0	0	0
$\sqrt{29}$	8	8	8	8	8	4	6	4	0	0	0	0	0
$\sqrt{32}$	8	8	8	8	8	6	6	4	0	0	0	0	0
$\sqrt{34}$	8	8	8	8	8	8	6	6	4	0	0	0	0
6	8	8	8	8	8	8	6	6	4	0	0	0	0
$\sqrt{37}$	8	8	8	8	8	8	8	6	4	0	4	0	0
$\sqrt{40}$	8	8	8	8	8	8	8	8	4	0	6	4	0
$\sqrt{41}$	8	8	8	8	8	8	8	8	6	4	6	4	0
$\sqrt{45}$	8	8	8	8	8	8	8	8	8	4	6	6	4
7	8	8	8	8	8	8	8	8	8	4	6	6	4

Now let us examine the increment of  $\Phi$  that occurs when new configurations are born. A direct counting gives

$$(\Delta\mathcal{C}, \Delta\mathcal{R})_{\text{diag}} = (2, 16), \quad (\text{A6a})$$

$$(\Delta\mathcal{C}, \Delta\mathcal{R})_{\text{linear}} = (2, 8), \quad (\text{A6b})$$

$$(\Delta\mathcal{C}, \Delta\mathcal{R})_{\text{knight}} = (4, 32). \quad (\text{A6c})$$

Plugging (A6b) and (A6c) into (A5) we find that

$$(\Delta\Phi)_{\text{linear}} = \frac{1}{2}, \quad (\Delta\Phi)_{\text{diag}} = (\Delta\Phi)_{\text{knight}} = 0. \quad (\text{A7})$$

Thus  $\Phi$  increases by  $1/2$  when a pair of linear configurations are born and does not change when configurations of other types are added. This explains why  $\Phi$  is half-integer and why the jumps in  $\Phi$  occur when  $\ell$  passes the integer value. Hence  $\Phi = (\lfloor \ell \rfloor + 1)/2$ , where  $\lfloor \ell \rfloor$  is the integer part of  $\ell$ ; that is, the largest integer not exceeding  $\ell$ . Plugging this into  $D = 1/2 - \Phi/\mathcal{C}$  we arrive at Eq. (14).

[1] M. N. Stojanovic and D. Stefanovic, *Nat. Biotechnol.* **21**, 1069 (2003).  
 [2] H. Yan, X. Zhang, Z. Chen, and N. C. Seeman, *Nature (London)* **415**, 62 (2002); N. C. Seeman, *Trends Biochem. Sci.* **30**, 119 (2005).  
 [3] W. B. Sherman and N. C. Seeman, *Nano Lett.* **4**, 1203 (2004).  
 [4] J.-S. Shin and N. A. Pierce, *J. Am. Chem. Soc.* **126**, 10834 (2004).  
 [5] J. Bath, S. J. Green, K. E. Allen, and A. J. Turberfield, *Small* **5**, 1513 (2009).  
 [6] T. Omabegho, R. Sha, and N. C. Seeman, *Science* **324**, 67 (2009).  
 [7] S. F. J. Wickham, M. Endo, Y. Katsuda, K. Hidaka, J. Bath, H. Sugiyama, and A. J. Turberfield, *Nat. Nanotechnol.* **6**, 166 (2011).  
 [8] E. R. Kay, D. A. Leigh, and F. Zerbetto, *Angew. Chem., Int. Ed.* **46**, 72 (2007).  
 [9] R. Pei, S. K. Taylor, D. Stefanovic, S. Rudchenko, T. E. Mitchell, and M. N. Stojanovic, *J. Am. Chem. Soc.* **128**, 12693 (2006).  
 [10] K. Lund *et al.*, *Nature (London)* **465**, 206 (2010).  
 [11] M. N. Stojanovic and D. Stefanovic, *J. Comput. Theor. Nanosci.* **3**, 434 (2011).  
 [12] T. Antal, P. L. Krapivsky, and K. Mallick, *J. Stat. Mech.* (2007) P08027.  
 [13] T. Antal and P. L. Krapivsky, *Phys. Rev. E* **76**, 021121 (2007).  
 [14] L. Samii, H. Linke, M. J. Zuckermann, and N. R. Forde, *Phys. Rev. E* **81**, 021106 (2010).  
 [15] L. Samii, G. A. Blab, E. H. C. Bromley, H. Linke, P. M. G. Curmi, M. J. Zuckermann, and N. R. Forde, *Phys. Rev. E* **84**, 031111 (2011).  
 [16] O. Semenov, M. J. Olah, and D. Stefanovic, *Phys. Rev. E* **83**, 021117 (2011).  
 [17] O. Semenov, M. J. Olah, and D. Stefanovic, in *DNA 17: Proceedings of The Seventeenth International Meeting on DNA Computing and Molecular Programming*, Lecture Notes in Computer Science Vol. 6397 (Springer-Verlag, Berlin, 2011), pp. 204–216.  
 [18] I. Ben-Ari, K. Boushaba, A. Matzavinos, and A. Roitershtein, *Bull. Math. Biol.* **73**, 1932 (2011).

- [19] C. Gallesco, S. Müller, and S. Popov, *ESAIM: Probab. Stat.* **15**, 390 (2011).
- [20] C. Gallesco, S. Müller, S. Popov, and M. Vachkovskaia, *ALEA-Latin American Journal of Probability and Mathematical Statistics* **8**, 129 (2011).
- [21] R. Juhász, *J. Stat. Mech.* (2007) P11015.
- [22] Mathematical models animal locomotion are reviewed by P. Holmes, R. J. Full, D. Koditschek, and J. Guckenheimer, *SIAM Rev.* **48**, 207 (2006).
- [23] The generalization to the triangular and honeycomb lattices should be straightforward while new behaviors can emerge on *random* planar lattices.
- [24] Throughout this paper we set both the lattice spacing and the hopping rate of a leg from the substrate to unity, so that for a one-leg spider (a random walk) the diffusion coefficient is  $D = 1$ . With our conventions, the diffusion coefficient of any spider is dimensionless.
- [25] F. van Wijland, S. Caser, and H. J. Hilhorst, *J. Phys. A* **30**, 507 (1997).
- [26] P. L. Krapivsky, S. Redner, and E. Ben-Naim, *A Kinetic View of Statistical Physics* (Cambridge University Press, Cambridge, 2010).
- [27] R. K. Guy, *Unsolved Problems in Number Theory* (Springer-Verlag, New York, 1994).
- [28] In this section we use the same notation  $S$  for the maximal span both for von Neumann and Euclidean spiders. In Sec. III and the Appendix we use notation  $\ell$  for the maximal span of Euclidean spiders. One reason is that, e.g., the simplest motile bipedal von Neumann spider has  $S = 1$  and it is identical to the simplest motile bipedal Euclidean spider which has  $\ell = \sqrt{2}$ .
- [29]  $D_2(r)$  and  $A_2(r)$  are actually displayed in Figs. 2 and 3. In Eqs. (22)–(24) we used an index to distinguish between dimensions  $d = 1$  and  $d = 2$ , but throughout the paper we are using the two-dimensional framework, so we suppress the index.
- [30] S. Saffarian, I. E. Collier, B. L. Marmer, E. L. Elson, and G. Goldberg, *Science* **306**, 108 (2004).
- [31] T. Antal and P. L. Krapivsky, *Phys. Rev. E* **72**, 046104 (2005).
- [32] A. Yu. Morozov, E. Pronina, A. B. Kolomeisky, and M. N. Artyomov, *Phys. Rev. E* **75**, 031910 (2007).
- [33] J. Howard, *Mechanics of Motor Proteins and the Cytoskeleton* (Sinauer Associates, Sunderland, MA, 2001).
- [34] A. Parmeggiani, T. Franosch, and E. Frey, *Phys. Rev. Lett.* **90**, 086601 (2003).
- [35] M. R. Evans, T. Hanney, and Y. Kafri, *Phys. Rev. E* **70**, 066124 (2004).
- [36] S. Klumpp, M. J. I. Müller, and R. Lipowsky, in *Traffic and Granular Flow '05*, edited by A. Schadschneider *et al.* (Springer-Verlag, Berlin, 2007), pp. 251–261.
- [37] S. Klumpp and R. Lipowsky, *Proc. Natl. Acad. Sci. USA* **102**, 17284 (2005).
- [38] M. J. I. Müller, S. Klumpp, and R. Lipowsky, *Proc. Natl. Acad. Sci. USA* **105**, 4609 (2008).
- [39] P. Greulich, A. Garai, K. Nishinari, A. Schadschneider, and D. Chowdhury, *Phys. Rev. E* **75**, 041905 (2007).
- [40] M. Mabilia, T. Reichenbach, H. Hinsch, T. Franosch, and E. Frey, *Banach Cent. Publ.* **80**, 101 (2008).
- [41] N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry* (Elsevier, Amsterdam, 2003).