IOP PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND THEORETICAL

J. Phys. A: Math. Theor. **43** (2010) 045208 (28pp) doi:10.1088/1751-8113/43/4/045208

The linking number and the writhe of uniform random walks and polygons in confined spaces

E Panagiotou[1], K C Millett[2] and S Lambropoulou[1]

[1] Department of Mathematics, National Technical University of Athens, Zografou Campus, GR 15780, Athens, Greece
[2] Department of Mathematics, University of California, Santa Barbara, CA 93106, USA

E-mail: panagiotou@math.ntua.gr, millett@math.ucsb.edu and sofia@math.ntua.gr

Received 13 July 2009, in final form 5 November 2009
Published 4 January 2010
Online at stacks.iop.org/JPhysA/43/045208

Abstract
Random walks and polygons are used to model polymers. In this paper we consider the extension of the writhe, self-linking number and linking number to open chains. We then study the average writhe, self-linking and linking number of random walks and polygons over the space of configurations as a function of their length. We show that the mean squared linking number, the mean squared writhe and the mean squared self-linking number of oriented uniform random walks or polygons of length n, in a convex confined space, are of the form $O(n^2)$. Moreover, for a fixed simple closed curve in a convex confined space, we prove that the mean absolute value of the linking number between this curve and a uniform random walk or polygon of n edges is of the form $O(\sqrt{n})$. Our numerical studies confirm those results. They also indicate that the mean absolute linking number between any two oriented uniform random walks or polygons, of n edges each, is of the form $O(n)$. Equilateral random walks and polygons are used to model polymers in θ-conditions. We use numerical simulations to investigate how the self-linking and linking number of equilateral random walks scale with their length.

PACS numbers: 02.10.Kn, 82.35.−x

1. Introduction

A polymer melt may consist of ring polymers (closed chains), linear polymers (open chains) or a mixed collection of ring and linear polymers. Polymer chains are long flexible molecules that impose spatial constraints on each other because they cannot intersect (de Gennes 1979, Rubinstein and Colby 2006). These spatial constraints, called *entanglements*, affect the conformation and motion of the chains in a polymer melt and have been studied using different models of *entanglement effects* in polymers (Orlandini *et al* 2000, Orlandini and Whittington

2004, 2007, Tzoumanekas and Theodorou 2006). However, a clear expression of exactly what one means by *entanglement* and how one can quantify the extent of its presence remains an elusive goal.

In the mathematical study of polymers, we usually consider conformations of chains with no interaction between monomers that are far apart along the chain, even if they approach each other in space. Such chains are called *ideal chains* (Rubinstein and Colby 2006). This situation has never been completely realized for real chains, but there are several types of polymeric systems with nearly ideal chains. At a special intermediate temperature, called the θ-temperature, chains are nearly in ideal conformations, because the attractive and repulsive parts of monomer–monomer interactions cancel each other. Even more importantly, linear polymer melts and concentrated solutions have practically ideal chain conformations because the interactions between monomers are almost completely screened by the surrounding chains. The conformation of an ideal chain with no interactions between monomers is the essential starting point for most models in polymer physics. Every possible conformation of an ideal chain can be mapped onto a random walk. Hence, random walk and ideal walk statistics are similar.

By Diao *et al* (1993), Diao (1995), Pippenger (1989), Sumners and Whittington (1988), we know that the probability that a polygon or open chain with n edges in the cubic lattice and 3-space is unknotted goes to zero as n goes to infinity. This result confirms the Frisch–Wasserman–Delbruck conjecture that long ring polymers in dilute solution will be knotted with high probability. A stronger theorem is that the probability that any specific knot type appears as a summand of a random walk or polygon goes to 1 as n goes to infinity. However, the probability of forming a knot with a given type goes to zero as n goes to infinity (Whittington 1992). The knot probability of polymer molecules also depends on the extent to which the molecule is geometrically confined. This has been studied by Arsuaga *et al* (2007) and Tesi *et al* (1994). For instance, DNA molecules confined to viral capsids have a much higher probability of being knotted. Moreover, the distribution of knot types is different from the distribution of DNA in solution (Arsuaga *et al* 2002, Weber *et al* 2006).

Polymers in solution are flexible moving molecular conformations, so the open polymer chains can always be moved apart, at least at long time scales. Therefore, the definition of knotting or linking must refer to spatially fixed configurations, and, as a consequence, does not give rise to a topologically invariant concept. In contrast, for closed polymer chains, the concepts of knotting and linking are unchanged under continuous deformations that do not allow breakage of the chains or passage of one portion of a chain through another. To characterize the knotting of an open chain one can use the DMS method (Millett *et al* 2004, Millett and Sheldon 2005) to determine the spectrum of knotting arising from the distribution of knot types, created by closure of the open chain to the 'sphere at infinity'. In practice, one first determines the center of mass of the chain, next the radius of the smallest sphere (centered at the center of mass) containing the chain, and then randomly select closure points on the concentric sphere of 100 times this radius as a surrogate for the 'sphere at infinity'. The previous studies of this method have shown that the data generated are stable with respect to increasing radii so that, effectively, this provides the required data. One is then able to employ the methods of traditional knot or link theory (e.g. Kauffman (2001)) to study the topology of individual constituents or of the entire collection. The knot types of individual closures or of ring polymer chains can be analyzed using, for example, the Jones (1985) or the HOMFLYPT (Freyd *et al* 1985, Przytycki and Traczyk 1987, Ewing and Millett 1997) polynomial, which can distinguish the different knot and link types with high precision. Yet, even for collections of closed chains, the application of the Jones or HOMFLYPT invariants has been constrained by the computational complexity one encounters in the scientifically interesting range of scales

and by the present scarcity of theoretical results. Moreover, we know that there exist infinite families of collections of linked closed chains whose Jones polynomials are those of the unlinked collection.

A classical measure of entanglement is the Gauss linking integral which is a classical topological integer invariant in the case of pairs of ring polymers. As we mentioned before, polymers in solution are flexible moving molecules, so we will be interested in characterizations of their physical properties at fixed moments in time as well as in their time-averaged properties. For open or mixed pairs, the calculated quantity is a real number that is characteristic of the conformation, but changes continuously under continuous deformations of the constituent chains. Thus, the application of the Gauss linking integral to open polymer chains is very clearly not a topological invariant, but, rather, a quantity that depends on the specific geometry of the chains. This measure is very sensitive to the specific conformations that are analyzed (Agarwal *et al* 2004, Berger and Prior 2006, Laing and Sumners 2006, 2008). For pairs of 'frozen' open polymer chains, or for a mixed frozen pair, we will see that the Gauss linking integral can be applied to calculate an average linking number. In a similar manner, the Gauss linking integral can be applied to calculate the writhe or the self-linking number of a 'frozen' configuration of one open chain. It is true that a complicated tangle and a really untangled curve can have essentially the same writhe, but it takes special effort to construct untangled complicated looking curves with high absolute writhe. Exactly the same considerations apply for the linking number and the self-linking number. Indeed, computer experiments indicate that the linking number and the writhe are effective indirect measures of whatever one might call 'entanglement', especially in systems of 'random' filaments.

As far as the time-averaged properties of polymers in solution are concerned, we will study the averages of such properties over the entire space of possible configurations. The literature suggests that, from the perspective of experiments, theory and simulation, the average crossing number (ACN) has been the principal quantity employed by researchers (Arsuaga *et al* 2009, Arteca 1997, Arteca and Tapia 2000, Diao *et al* 2003, 2005, Edvinsson *et al* 2000, Freedman *et al* 1994, Reimann *et al* 2002, Sheng and Tsao 2002, Simon 2009). This is neither a topological nor a geometric characteristic, as it depends on the specific conformation and is changed under continuous deformations (Grassberger 2001). In some sense, the linking number is a strictly stronger measure of physical entanglement than the ACN. There has been a substantial body of research to make rigorous the relationship of the linking number with more scientifically intuitive concepts of entanglement (Arsuaga *et al* 2002, 2007a, 2007b, Barbi *et al* 2005, Bauer *et al* 1980, Buck *et al* 2008, Calugreanu 1961, Crick 1976, Edwards 1967, 1968, Fuller 1978, Hirshfeld 1997, Hoidn *et al* 2002, Holmes and Cozzarelli 2000, Iwata and Edwards 1989, Katritch *et al* 1996, Klenin and Langowski 2000, Kung and Kamien 2003, Lacher and Sumners 1991, Liu and Chan 2008, Mansfield 1994, Martinez-Robles *et al* 2009, McMillen and Goriely 2002, Micheletti *et al* 2006, Orlandini *et al* 1994, 2000, Orlandini and Whittington 2004, Ricca 2000, Rogen and Fain 2003, Stasiak *et al* 1998, Stump *et al* 1998, Vologodskii and Cozzarelli 1993, Vologodskii 1999, White and Bauer 1969).

The main subject of our study concerns uniform random chains confined in a symmetric convex region in $\mathbb{R}^3$. The observation that confined geometries have substantial impact on the structure of the polymer is an attractive target of the study of the local absolute linking (Arsuaga *et al* 2009, Micheletti *et al* 2006, Orlandini and Whittington 2004, Sheng and Tsao 2002, Tesi *et al* 1994). It is an effect that is captured locally by our measures of the absolute linking number of open chains or of a single chain. Developing analytical results on the complexity of knots formed by polymer chains in confined volumes is a very difficult problem. The uniform random polygon (URP) model (Millett 2000) may provide clues about showing some of these analytical results (Arsuaga *et al* 2007a, 2007b, 2009). Each edge of

the random open chain or polygon is defined by a pair of random points in the convex region with respect to the uniform distribution. Throughout this paper we will refer to open closed uniform random chains as uniform random walks and uniform random polygons, respectively. We will consider uniform random walks or polygons confined in the unit cube. Inspired by the effort of Arsuaga, Blackstone, Diao, Karadayi, and Saito in 'The linking of uniform random polygons in confined spaces' (Arsuaga *et al* 2007a) to forging a rigorous connection between physical entanglement and topological invariants, this manuscript reports on two innovations that directly pertain to applications to polymers (in contrast to the theory of topological knots or links).

First, we study *the mean absolute linking number, the mean absolute writhe and the mean absolute self-linking number* numerically and provide analytical results that will help us understand more about the scaling of those quantities (theorem 3.9). We point out that the results in Arsuaga *et al* (2007a) concern theoretical results and simulations to study the mean squared linking number. We very often employ the geometric mean of a quantity rather than the arithmetic mean. For example, in the study of polymers, the most frequent measure is the mean squared radius of gyration. Additionally, one has the mean squared end-to-end distance and the mean squared diameter (Rubinstein and Colby 2006). Statistics, however, tells us that while such mean squared quantities, that are geometric means, are more accessible to theoretical estimation, the mean absolute linking number corresponding to an arithmetic mean is a quantity that is not biased by small incidences of large linking numbers and, therefore, is a truer measure of the complexity intrinsic to the data. Yet it is much harder to estimate the mean absolute linking number. The second innovation is the *extension and application of the absolute linking number from closed polymer chains to open polymer chains*. We also study both numerically and analytically the mean squared writhe of open chains (theorem 3.1), the mean squared linking number of open chains (theorem 3.6), the mean squared self-linking number of open chains (theorem 3.7) and the mean absolute linking number between an open chain and a fixed closed curve (theorem 3.9).

More precisely, the paper is organized as follows. In section 2 we study the scaling of the mean squared writhe, the mean squared linking number and the mean squared self-linking number of oriented uniform random walks and polygons in confined space with respect to their length. Next, we study the scaling of the mean absolute linking number of an oriented uniform random walk or polygon and a simple closed curve, both contained in a unit cube, with respect to the length of the random walk or the polygon, respectively. In section 3 we present the results of our numerical simulations which confirm the analytical results presented in section 2. Although theoretical results about the absolute linking number between two uniform random walks or polygons in confined space appear difficult to acquire, we are able to provide numerical results in section 3. Also it is more difficult to provide analytical results for equilateral random walks or polygons than for uniform random walks or polygons. An equilateral random walk is an ideal chain composed of freely jointed segments of equal length in which the individual segments have no thickness (Diao *et al* 2003, 2005, Dobay *et al* 2003). In this direction, we give in section 3 numerical estimations of the scaling of the mean absolute linking number and the mean absolute self-linking number of equilateral random walks.

2. Measures of entanglement of open curves

In a generic orthogonal projection of two oriented polygonal chains, each crossing is of one of the types shown in figure 1. By convention, we assign +1 and −1 to a crossing of the first type and second type, respectively.

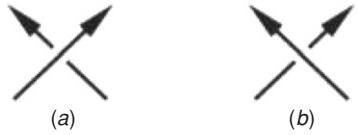

Figure 1. (*a*) +1 crossing and (*b*) −1 crossing.

For a generic projection of two oriented curves l_1, l_2 to a plane defined by a vector $\boldsymbol{\xi} \in S^2$, the *linking number of a diagram*, denoted $lk_{\boldsymbol{\xi}}(l_1, l_2)$, is equal to one half of the algebraic sum of crossings between the projected curves. The *linking number* of two oriented curves is then equal to the average linking number of a diagram over all possible projection directions, i.e. $L(l_1, l_2) = 1/4\pi\left(\int_{\boldsymbol{\xi}\in S^2} lk_{\boldsymbol{\xi}}(l_1, l_2)\,\mathrm{d}S\right)$. This can be expressed by the Gauss linking integral for two oriented curves.

Definition 2.1. *The Gauss linking number of two oriented curves l_1 and l_2, whose arc-length parametrization is $\gamma_1(t), \gamma_2(s)$ respectively, is defined as a double integral over l_1 and l_2 (Gauss 1877):*

$$L(l_1, l_2) = \frac{1}{4\pi}\int_{[0,1]}\int_{[0,1]} \frac{(\dot\gamma_1(t), \dot\gamma_2(s), \gamma_1(t) - \gamma_2(s))}{|\gamma_1(t) - \gamma_2(s)|^3}\,\mathrm{d}t\,\mathrm{d}s, \tag{1}$$

where $(\dot\gamma_1(t), \dot\gamma_2(s), \gamma_1(t) - \gamma_2(s))$ is the triple product of $\dot\gamma_1(t)$, $\dot\gamma_2(s)$ and $\gamma_1(t) - \gamma_2(s)$.

Similarly, for the generic orthogonal projection of one oriented curve l to a plane defined by a vector $\boldsymbol{\xi} \in S^2$, we define the *writhe of a diagram*, denoted $\mathrm{Wr}_{\boldsymbol{\xi}}(l)$, to be equal to the algebraic sum of crossings of the projection of the curve with itself. Then the *writhe* of a curve is defined as the average writhe of a diagram of the curve over all possible projections, i.e. $\mathrm{Wr}(l) = 1/4\pi\left(\int_{\boldsymbol{\xi}\in S^2} \mathrm{Wr}_{\boldsymbol{\xi}}(l)\,\mathrm{d}S\right)$. This can be expressed as the Gauss linking integral over one curve.

Definition 2.2. *The writhe of an oriented curve l, whose arc-length parametrization is $\gamma(t)$, is defined by the Gauss linking integral over a curve*

$$\mathrm{Wr}(l) = \frac{1}{2\pi}\int_{[0,1]^*}\int_{[0,1]^*} \frac{(\dot\gamma(t), \dot\gamma(s), \gamma(t) - \gamma(s))}{|\gamma(t) - \gamma(s)|^3}\,\mathrm{d}t\,\mathrm{d}s, \tag{2}$$

where $[0, 1]^* \times [0, 1]^* = \{(x, y) \in [0, 1] \times [0, 1] | x \neq y\}$.

We define the *average crossing number* of a curve l, whose parametrization is $\gamma(t)$, to be the average sum of crossings in a generic orthogonal projection over all possible projection directions. It is expressed by a double integral over l:

$$\mathrm{ACN} = \frac{1}{2\pi}\int_{[0,1]}\int_{[0,1]} \frac{|(\dot\gamma(t), \dot\gamma(s), \gamma(t) - \gamma(s))|}{|\gamma(t) - \gamma(s)|^3}\,\mathrm{d}t\,\mathrm{d}s, \tag{3}$$

where $(\dot\gamma(t), \dot\gamma(s), \gamma(t) - \gamma(s))$ is the triple product of $\dot\gamma(t)$, $\dot\gamma(s)$ and $\gamma(t) - \gamma(s)$.

We observe that the geometrical meaning of the linking number and the writhe is the same for open or closed curves. The linking number of two curves is the average over all possible projection directions of half the algebraic sum of crossings between the two components in the projection of the curves. Similarly the writhe of a curve is the average over all projection directions of the algebraic sum of crossings in the projection of the curve.

In the case of oriented closed curves, the linking number is an integer topological invariant, i.e. it is invariant under isotopic moves of the curves. But in the case of oriented open curves,

the linking number is not topological invariant and it is not an integer. If the open curves are allowed to move continuously without intersecting each other or themselves in space, all the above measures are continuous functions in the space of configurations. Furthermore, as the endpoints of the curves move toward coincidence, the linking number, self-linking number or writhe tends to the values of those measures for the resulting closed knots or links.

Polymer chains are often modeled using open or closed polygonal curves. There exist several models of random walks or polygons that can be used and that are representative of the properties of different polymer melts. Since a polymer melt can take different formations in space through time, we are interested in a measure of complexity of polymer chains that will be independent of a specific configuration and characteristic of the configuration space and how it depends on the length of the chains. In this paper, we will focus our study on uniform random walks in a confined space as this provides a simplified model for our theoretical study and will have a similar behaviour to other polymer models (or will give us some insight into the study of more realistic models).

3. Uniform random walks and polygons in a confined space

The uniform random walks and polygons are modeled after the URP model, introduced by Millett (2000). In this model there are no fixed bond lengths and each coordinate of a vertex of the uniform random polygon contained in C^3, where $C = [0, 1]$ is drawn from a uniform distribution over [0, 1].

The following theorem has been proved by Arsuaga *et al* (2007).

Theorem 3.1. *The mean squared linking number between two oriented uniform random polygons X and Y of n edges, each contained in C^3, is $\frac{1}{2}n^2q$ where $q > 0$. A similar result holds if C^3 is replaced by a symmetric convex set in $\mathbb{R}^3$.*

The above result is independent of the orientation of the two uniform random polygons. Due to the weight squaring given to larger linking number, we propose that the mean of the absolute value of the linking number between oriented uniform random walks or polygons would be a more informative measure of the expected degree of linking.

3.1. The mean squared writhe of an oriented uniform random walk in a confined space

In this section, we study the scaling of the writhe of an oriented uniform random walk (or polygon) contained in C^3.

We are interested in defining the average squared writhe of an n-step uniform random walk or polygon in confined space, where the average is taken over the entire population of open or closed uniform random walks or polygons in the confined space C^3. We distribute vertices according to the uniform distribution on the cube. More explicitly, the space of configurations in this case is $\Omega = [0, 1]^{3(n+1)}\backslash N$ and $\Omega = [0, 1]^{3n}\backslash N$, respectively, where N is the set of singular configurations, i.e. when a walk or polygon intersects itself. Then N is a set of measure zero (Randell 1988a, 1988b, Calvo 1999).

The average writhe over the space of chains or polygons is zero as there is a sign balance occurring due to the mirror reflection involution on the space of configurations. This is why we choose to study the mean squared writhe of a uniform random walk of n edges in C^3.

Theorem 3.2. *The mean squared writhe of an oriented uniform random walk or polygon of n edges, each contained in C^3, is of the order of $O(n^2)$. Similar results hold if C^3 is replaced by a symmetric convex set in $\mathbb{R}^3$.*

Let us consider two (independent) oriented random edges l_1 and l_2 of an oriented uniform random polygon P_n and a fixed projection plane defined by a normal vector $\boldsymbol{\xi} \in S^2$. Since the end points of the edges are independent and are uniformly distributed in C^3, the probability that the projections of l_1 and l_2 intersect each other is a positive number which we will call $2p$. We define a random variable ϵ in the following way: $\epsilon = 0$ if the projection of l_1 and l_2 has no intersection, $\epsilon = -1$ if the projection of l_1 and l_2 has a negative intersection, and $\epsilon = 1$ if the projection of l_1 and l_2 has a positive intersection. Note that, in the case the projections of l_1 and l_2 intersect, ϵ is the sign of their crossing. Since the end points of the edges are independent and are uniformly distributed in C^3, we then see that $P(\epsilon = 1) = P(\epsilon = -1) = p$, $E[\epsilon] = 0$ and $\mathrm{Var}(\epsilon) = E[\epsilon^2] = 2p$.

We will need the following lemma, modeled after lemma 1 by Arsuaga *et al* (2007), concerning the case when there are four edges involved (some of them may be identical or they may have a common end point): l_1, l_2, l_1' and l_2'. Let ϵ_1 be the random number ϵ defined above between l_1 and l_1' and let ϵ_2 be the random number defined between l_2 and l_2'.

Lemma 3.3.

(1) If the end points of l_1, l_2, l_1' and l_2' are distinct, then $E[\epsilon_1\epsilon_2] = 0$ (this is the case when there are eight independent random points involved).

(2) If $l_1 = l_2$, and the end points of l_1, l_1' and l_2' are distinct (this reduces the case to where there are only three random edges with six independent points involved), then $E[\epsilon_1\epsilon_2] = 0$.

(3) If l_1 and l_1', or l_2 and l_2', have a common end point, then $E[\epsilon_1\epsilon_2] = 0$.

(4) In the case where $l_1 = l_2$, the endpoints of l_1 and l_1' and l_1 and l_2' are distinct, and l_1' and l_2' share a common point (so there are only five independent random points involved in this case), let $E[\epsilon_1\epsilon_2] = u$. In the case where l_1 and l_2 share a common point, the endpoints of l_1 and l_1' and l_1 and l_2' are distinct, and l_1' and l_2' also share a common point (so there are four edges defined by six independent random points involved in this case), let $E[\epsilon_1\epsilon_2] = v$. Finally let $E[\epsilon_1\epsilon_2] = w$ in the case where l_1, l_2, l_1' and l_2' are consecutive (so in this case, there are four edges defined by five independent random points). Then we have $q' = 3p + 2(2u + v + w) > 0$, where p is as defined before.

Note that in comparison with lemma 1 by Arsuaga *et al* (2007), we have included the case in which some of the four or three edges involved are consecutive.

Proof.

(1) This is true since ϵ_1 and ϵ_2 are independent random variables in this case.

(2) For each configuration in which the projections of l_1' and l_2' both intersect the projection of l_1(since otherwise $\epsilon_1\epsilon_2 = 0$), there are eight different ways to assign the orientations to the edges. Four of them yield $\epsilon_1\epsilon_2 = -1$ and four of them yield $\epsilon_1\epsilon_2 = 1$. Since the joint density function of the vertices involved is simply $\frac{1}{V^6}$, where V is the volume of the confined space C^3, thus by a symmetry argument we have $E[\epsilon_1\epsilon_2] = 0$.

(3) This is true since in that case $\epsilon_1 = 0$ or $\epsilon_2 = 0$.

(4) Consider the case when the polygon has six edges. Let ϵ_{ij} be the crossing sign number ϵ between the edges l_i and l_j. Consider the variance of the summation $\sum_{1\leqslant i\leqslant 6}\sum_{\substack{j>i\\ j\neq i-1,i,i+1}}\epsilon_{ij}$

(the summation indices are taken modulo 6):

$$
\begin{aligned}
V\left(\sum_{1\leqslant i\leqslant 6}\sum_{\substack{j>i\\ j\neq i-1,i,i+1}}\epsilon_{ij}\right) &= E\left[\left(\sum_{1\leqslant i\leqslant 6}\sum_{\substack{j>i\\ j\neq i-1,i,i+1}}\epsilon_{ij}\right)^2\right] = \sum_{1\leqslant i\leqslant 6}\sum_{\substack{j>i\\ j\neq i-1,i,i+1}} E\left[\epsilon_{ij}^2\right] \\
&\quad +2\sum_{1\leqslant i\leqslant 6}\sum_{\substack{j>i\\ j\neq i-2,i-1,i,i+1}} E[\epsilon_{ij}\epsilon_{i(j+1)}] + 2\sum_{1\leqslant i\leqslant 6}\sum_{\substack{j>i\\ j\neq i-1,i,i+1,i+2}} E[\epsilon_{ij}\epsilon_{(i+1)j}] \\
&\quad +2\sum_{1\leqslant i\leqslant 6}\sum_{\substack{j>i\\ j\neq i-2,i-1,i,i+1,i+2}} (E[\epsilon_{ij}\epsilon_{(i+1)(j+1)}] + E[\epsilon_{i(j+1)}\epsilon_{(i+1)j}]) \\
&\quad +2\sum_{1\leqslant i\leqslant 6} E[\epsilon_{i(i+2)}\epsilon_{(i+1)(i+3)}]. \qquad (4)
\end{aligned}
$$

Since the ϵ_{ij} are identical random variables, i.e. they have the same distributions, each term in the first summation of the right-hand side in the above yields $2p$, each term in the second summation yields u (see lemma 1), each term in the third and fourth summation yields v and each term in the fifth summation yields w. There are 9 terms in the first summation, there are 12 terms in the second summation, 6 terms in the third summation, 6 terms in the fourth summation and 6 terms in the fifth summation. This leads to

$$
V\left(\sum_{1\leqslant i\leqslant 6}\sum_{\substack{j>i\\ j\neq i-1,i,i+1}}\epsilon_{ij}\right) = 9\cdot 2p + 2(12\cdot u + 6\cdot v + 6\cdot w) = 6(3p + 2(2u+v+w)). \qquad (5)
$$

Since $V\left(\sum_{1\leqslant i\leqslant 6}\sum_{\substack{j>i\\ j\neq i-1,i,i+1}}\epsilon_{ij}\right) > 0$, this implies that $3p + 2\,(2u+v+w) > 0$, as claimed. □

Proof of Theorem 3.2. Let us consider the orthogonal projection of P_n to a plane perpendicular to a random vector $\xi \in S^2$. Throughout the proof the averaging, E, is always over the space of configurations. The average, over the space of configurations, squared writhe of the projection of P_n to that plane is given by

$$
\begin{aligned}
E\left[\mathrm{Wr}_\xi^2(P_n)\right] &= E\left[\left(\sum_{1\leqslant i\leqslant n}\sum_{\substack{j>i\\ j\neq i-1,i,i+1}}\epsilon_{ij}\right)^2\right] = \sum_{1\leqslant i\leqslant n}\sum_{\substack{j>i\\ j\neq i-1,i,i+1}} E\left[\epsilon_{ij}^2\right] \\
&\quad +2\sum_{1\leqslant i\leqslant n}\sum_{\substack{j>i\\ j\neq i-2,i-1,i,i+1}} E[\epsilon_{ij}\epsilon_{i(j+1)}] + 2\sum_{1\leqslant i\leqslant n}\sum_{\substack{j>i\\ j\neq i-1,i,i+1,i+2}} E[\epsilon_{ij}\epsilon_{(i+1)j}] \\
&\quad +2\sum_{1\leqslant i\leqslant n}\sum_{\substack{j>i\\ j\neq i-2,i-1,i,i+1,i+2}} (E[\epsilon_{ij}\epsilon_{(i+1)(j+1)}] + E[\epsilon_{i(j+1)}\epsilon_{(i+1)j}]) \\
&\quad +2\sum_{1\leqslant i\leqslant n} [\epsilon_{i(i+2)}\epsilon_{(i+1)(i+3)}] \\
&= n^2(p + 2(u+v)) - n(3p + 2(4u + 5v - w)) \qquad (6)
\end{aligned}
$$

where p, u, v, w are defined as in lemma 3.3. By Arsuaga *et al* (2007) it has been proved that $p + 2(u + v) = q > 0$, thus $E\left[\mathrm{Wr}^2_{\xi}(P_n)\right] = qn^2 + O(n)$. Using lemma 3.3 we can see that $E\left[\mathrm{Wr}^2_{\xi}(P_n)\right]$ is bounded from below by $qn^2 - 6qn$.

Let us now take a partition of the surface of the 2-sphere $\Delta = \{I_1, I_2, \ldots, I_m\}$ such that the writhe of the projection of P_n is constant in each I_j, $1 \leqslant j \leqslant m$. By definition, we have that, for a sequence of partitions Δ_k such that $\mu(\Delta_k) \to 0$, the mean squared writhe of P_n is equal to

$$\begin{aligned}
E[\mathrm{Wr}^2(P_n)] &= E\left[\left(\frac{1}{4\pi}\lim_{\mu(\Delta_k)\to 0}\sum_{1\leqslant s\leqslant m_k}\mathrm{Wr}_{\xi_s}(P_n)\delta S\right)^2\right] \\
&= \frac{1}{16\pi^2}E\left[\left(\lim_{\mu(\Delta_k)\to 0}\sum_{1\leqslant s\leqslant m_k} Wr_{\xi_s}(P_n)\delta S\right)^2\right] \\
&= \frac{1}{16\pi^2}E\left[\lim_{\mu(\Delta_k)\to 0}\left(\sum_{1\leqslant s\leqslant m_k}\mathrm{Wr}^2_{\xi_s}(P_n)\delta S^2 + 2\sum_{1\leqslant s,t\leqslant m_k}\mathrm{Wr}_{\xi_s}(P_n)\mathrm{Wr}_{\xi_t}(P_n)\delta S^2\right)\right] \\
&= \frac{1}{16\pi^2}\lim_{\mu(\Delta_k)\to 0}E\left[\left(\sum_{1\leqslant s\leqslant m_k}\mathrm{Wr}^2_{\xi_s}(P_n)\delta S^2 + 2\sum_{1\leqslant s,t\leqslant m_k}\mathrm{Wr}_{\xi_s}(P_n)\mathrm{Wr}_{\xi_t}(P_n)\delta S^2\right)\right] \\
&= \frac{1}{16\pi^2}\cdot
\end{aligned}$$

$$\lim_{\mu(\Delta_k)\to 0}\left(\sum_{1\leqslant s\leqslant m_k}E\left[\mathrm{Wr}^2_{\xi_s}(P_n)\right]\delta S^2 + 2\sum_{1\leqslant s,t\leqslant m_k}E\left[\mathrm{Wr}_{\xi_s}(P_n)\right]E\left[\mathrm{Wr}_{\xi_t}(P_n)\right]\delta S^2\right), \tag{7}$$

where we use Lebesgue's theorem of dominated convergence, since the functions $S_k(P_n) = \left(\sum_{1\leqslant s\leqslant m_k}\mathrm{Wr}_{\xi_s}(P_n)\delta S\right)^2$, $S_k : \Omega \to \mathbb{R}$ are measurable functions, bounded above by $\left(\sum_{1\leqslant s\leqslant m} Cr_{\xi_s}(P_n)\delta S\right)^2$, where $Cr_{\xi_s}(P_n)$ is the number of crossings of the projection of P_n to the plane perpendicular to ξ_s and $\left(\sum_{1\leqslant s\leqslant m_k} Cr_{\xi_s}(P_n)\delta S\right)^2 \leqslant (24n^2)^2$.

The second term in (7) is equal to zero, because

$$E\left[\mathrm{Wr}_{\xi}(P_n)\right] = E\left[\sum_{1\leqslant i\leqslant n}\sum_{\substack{j>i\\ j\neq i-1,i,i+1}}\epsilon_{ij}\right] = \sum_{1\leqslant i\leqslant n}\sum_{\substack{j>i\\ j\neq i-1,i,i+1}}E[\epsilon_{ij}] = 0. \tag{8}$$

But we proved that $E\left[\mathrm{Wr}^2_{\xi}(P_n)\right] = qn^2 + O(n)$, and $E\left[\mathrm{Wr}_{\xi}(P_n)\right] = 0$, $\forall \xi \in S^2$ so

$$\begin{aligned}
E[\mathrm{Wr}^2(P_n)] &= \frac{1}{16\pi^2}\lim_{\mu(\Delta_k)\to 0}\left(\sum_{1\leqslant s\leqslant m_k}(qn^2 + O(n))\delta S^2\right) \\
&\Rightarrow \frac{1}{16\pi^2}(qn^2 + O(n))\left(\int_{\xi\in S^2}\mathrm{d}S\right) \\
&\leqslant E[\mathrm{Wr}^2(P_n)] \leqslant \frac{1}{16\pi^2}(qn^2 + O(n))\left(\int_{\xi\in S^2}\mathrm{d}S\right)^2 \\
&\Rightarrow \frac{1}{4\pi}(qn^2 + O(n)) \leqslant E(\mathrm{Wr}^2(P_n)) \leqslant qn^2 + O(n).
\end{aligned} \tag{9}$$

Note that in the case of a uniform random walk R_n, one has to add to the writhe the crossing between the first and last edges, otherwise the proof is the same and we obtain the same result, i.e. $E\left[\mathrm{Wr}(R_n)^2\right] \approx qn^2 + O(n)$. □

Remark 3.4. In Janse van Rensburg *et al* (1993), the mean absolute writhe of a self-avoiding polygon in $\mathbb{Z}^3$ is shown to have a lower bound of the form $O(\sqrt{n})$. The behavior of random walks on a lattice compared to the off-lattice random walks can have significant differences from walks of comparable lengths due to the special constraints of the lattice. In the case of uniform random walks confined to a cube, there are additional confounding influences that increase the effective density due to the uniform measure on the cube and the confinement. These result in striking differences in scaling of the writhe, that we have observed here, compared to the lattice scaling.

3.2. The mean squared linking number of two oriented uniform random walks in confined space

The proof of theorem 3.2 can be easily adapted in order to provide an analysis of the rate of scaling of the mean squared linking number of open chains. Specifically, we have the following theorem, generalizing theorem 3.1 by Arsuaga *et al* (2007).

Theorem 3.5. *The mean squared linking number between two oriented uniform random walks X and Y of n edges, contained in C^3, is of the order of $O(n^2)$. Similar results hold if C^3 is replaced by a symmetric convex set in $\mathbb{R}^3$.*

Proof. For a fixed orthogonal projection of the walks to a plane perpendicular to a vector $\xi \in S^2$, adapting theorem 3.1 of Arsuaga *et al* (2007) to the case of open walks, we have $E[lk_\xi(X, Y)] = \frac{1}{2}n^2 q + O(n)$ where $q > 0$. Then, following the proof of our theorem 3.2, we have that $E[Lk(X, Y)] = O(n^2)$. □

Note that $q = p + 2(u + v)$ has the same value in all theorems. By Arsuaga *et al* (2007), it has been estimated to be $q = 0.0338 \pm 0.024$. Our numerical results confirm this estimation.

3.3. The mean squared self-linking number of an oriented uniform random walk or polygon

The self-linking number was introduced to model two stranded DNA and is defined as the linking number between a curve l and a translated image of that curve l_ϵ at a small distance ϵ, i.e. $Sl(l) = L(l, l_\epsilon)$. This can be expressed by the Gauss integral over $[0, 1]^* \times [0, 1]^* = \{(x, y) \in [0, 1] \times [0, 1] | x \neq y\}$ by adding to it a correction term, so that it is a topological invariant of closed curves (Banchoff 1976):

$$\mathrm{SL}(l) = \frac{1}{4\pi}\int_{[0,1]^*}\int_{[0,1]^*} \frac{(\dot\gamma(t), \dot\gamma(s), \gamma(t) - \gamma(s))}{|\gamma(t) - \gamma(s)|^3}\,\mathrm{d}t\,\mathrm{d}s + \frac{1}{2\pi}\int_{[0,1]} \frac{\left(\gamma'(t) \times \gamma''(t)\right) \cdot \gamma'''(t)}{|\gamma'(t) \times \gamma''(t)|^2}\,\mathrm{d}t. \tag{10}$$

The first term, in the above, is the writhe of the curve which we studied in the last section. The second term is the total torsion of the curve, $\tau(l)$, divided by 2π. This measures how much the curve deviates from being planar. The torsion of a curve can be expressed as

$$\tau(l) = \sum_{1 \leqslant i \leqslant n} \phi_i(l), \tag{11}$$

where $\phi_i(l)$ is the signed angle between the binormal vectors B_i and B_{i+1} defined by the edges $i-1, i, i+1$ (Banchoff 1976).

The following theorem concerns the mean squared self-linking number of an oriented uniform random walk or polygon.

Theorem 3.6. *The mean squared self linking number of an oriented uniform random walk or polygon of n edges, contained in C^3 is of the order $O(n^2)$. Similar results hold if C^3 is replaced by a symmetric convex set in $\mathbb{R}^3$.*

Proof. We will use the definition of the self-linking number given by (10), i.e. $\mathrm{SL}(l) = \mathrm{Wr}(l) + \frac{1}{2\pi}\sum_i \phi_i$. The proof is based on the fact that the torsion angles ϕ_i $\forall i \neq 1, n$ and that the products $\epsilon_{ij}\phi_i$ $\forall j \neq i+2, k \neq i+1$ are independent in the URP model.

Let P_n denote a uniform random polygon in the confined space C^3. We project P_n to a fixed plane defined by a normal vector $\boldsymbol{\xi} \in S^2$. For each pair of edges of the uniform random polygon l_i and l_j, we define a random variable ϵ_{ij} as we did in the precious section. Then $E[\epsilon] = 0$ and $E[\epsilon^2] = 2p$. For each edge we define a random variable ϕ_i such that ϕ_i is equal to the signed angle between B_i and B_{i+1}, the normal vectors to the planes defined by the edges $i, i+1$ and $i+1$, $i+2$, respectively. Then $\phi_i \in [-\pi, \pi], \forall i$. Since each vertex of the uniform random polygon is chosen with respect to the uniform distribution, ϕ_i has equal probability of being positive or negative, thus $E(\phi_i) = 0, \forall i$. Now let $E\left[\phi_i^2\right] = w$ and $E[|\phi_i|] = w'$. For each pair of edges ϕ_i, ϕ_j, $i \neq 1$, $j \neq n$ we have that $E[\phi_i\phi_j] = 0$, since ϕ_i, ϕ_j are independent random variables in that case. We will now compute the mean squared self-linking number:

$$
\begin{aligned}
&E\left[\left(\sum_{\substack{1\leqslant i<j\leqslant n\\ j\neq i-1,i,i+1}} \epsilon_{ij} + \sum_{1\leqslant i\leqslant n}\phi_i\right)^2\right] \\
&\quad = E\left[\left(\sum_{\substack{1\leqslant i<j\leqslant n\\ j\neq i-1,i,i+1}} \epsilon_{ij}\right)^2 + \left(\sum_{1\leqslant i\leqslant n}\phi_i\right)^2 + 2\sum_{\substack{1\leqslant i<j\leqslant n\\ j\neq i-1,i,i+1}}\sum_{1\leqslant k\leqslant n}\epsilon_{ij}\phi_k\right] \\
&\quad = E\left[\left(\sum_{\substack{1\leqslant i<j\leqslant n\\ j\neq i-1,i,i+1}} \epsilon_{ij}\right)^2\right] + E\left[\left(\sum_{1\leqslant i\leqslant n}\phi_i\right)^2\right] + 2E\left[\sum_{\substack{1\leqslant i<j\leqslant n\\ j\neq i-1,i,i+1}}\sum_{1\leqslant k\leqslant n}\epsilon_{ij}\phi_k\right]. \qquad (12)
\end{aligned}
$$

In the previous section, we proved that $E\left[\left(\sum_{\substack{1\leqslant i<j\leqslant n\\ j\neq i-1,i,i+1}} \epsilon_{ij}\right)^2\right] = O(n^2)$. For the second term we have that

$$
E\left[\left(\sum_{1\leqslant i\leqslant n}\phi_i\right)^2\right] = \sum_{1\leqslant i\leqslant n} E\left[\phi_i^2\right] + 2\sum_{1\leqslant i<j\leqslant n} E[\phi_i\phi_j] = wn + 2E[\phi_1\phi_n] = O(n). \qquad (13)
$$

For the third term, we proceed as follows. If $j \neq i-2, i+2$, then ϵ_{ij}, ϕ_k are independent random variables for all k, thus $E[\epsilon_{ij}\phi_k] = 0$. If $j = i+2$ then ϵ_{ij}, ϕ_k are independent random variables for all $k \neq i+1$. For $k = i+1$, then there are eight different cases that can occur such that $E[\epsilon_{ii+2}\phi_{i+i}] \neq 0$ (see figure 2). All of them give $+|\phi_{i+1}|$. Since the vertices of the polygon

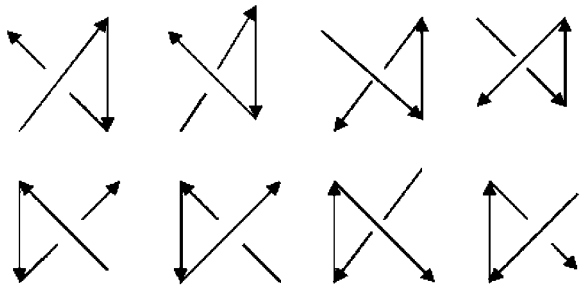

Figure 2. For all the possible configurations of the edges i, $i+1$ and $i+2$ such that $\epsilon_{i,i+2} \neq 0$, we have that $\epsilon_{i,i+2}\phi_{i+1} = |\phi_{i+1}|$.

are chosen with respect to the uniform distribution, all the cases have the same probability, thus $E[\epsilon_{ii+2}\phi_{i+1}] = E[|\phi_{i+1}|] = w'$. So, finally, we have that

$$E[Sl^2(P_n)] = qn^2 + O(n). \tag{14}$$

In the case of a uniform random walk R_n the self-linking number is not a topological invariant and one has to follow a similar averaging procedure as in the proof of theorem 3.2. Finally, the mean squared self-linking number of a uniform random walk R_n is $E[Sl^2(R_n)] = O(n^2)$. □

Remark 3.7. In the case of a uniform random polygon, theorem 3.6 can be proved using the following thinking. For a polygon P_n, $E[Sl^2(P_n)] = E\left[Lk\left(P_n, P_{n_\epsilon}\right)^2\right]$, where P_{n_ϵ} is the polygon that results by substituting every circular arc of the normal polygon of P_n (Banchoff 1976) by a straight segment. We can then apply the same method as for the mean squared linking number of two uniform random polygons used in theorem 3.1 by Arsuaga *et al* (2007). We note that small changes in this method are necessary due to the structure of P_{n_ϵ}. This does not change the rate of the scaling of the mean squared linking number and we obtain $E[Sl^2(P_n)] = O(n^2)$.

Remark 3.8. We call attention to the fact that the scalings of the mean squared writhe, the mean squared linking number and the mean squared self-linking number do not depend upon the size of the box in the URP model, i.e. this model is 'scale invariant'. The same data are randomly generated for any sized cube as the measure determining the random selection of points is uniform. If one wishes to increase the density of the chain, on average, it is necessary to increase the number of points that are selected in defining the chain. One defines the density of a chain contained in C^3 as $\rho = n/C^3$, as in Orlandini *et al* (2000) and Orlandini and Whittington (2004). One could determine the critical value of n at which the linking between the chains starts to become important. At the value for which linking becomes important, one would expect to observe correlated phase transitions of other polymeric properties.

3.4. The mean absolute value of the linking number of a uniform random walk or polygon with a simple closed curve in a confined space

In this section, following the proof of theorem 4 by Arsuaga *et al* (2007), we analyze the scaling of the absolute value of the linking number between a uniform random walk or polygon and a fixed simple closed curve in confined space.

Theorem 3.9. *Let R_n (or P_n) denote an oriented uniform random walk (or polygon, respectively) of n edges and S a fixed simple closed curve both confined in the interior of a symmetric convex set of $\mathbb{R}^3$. Then the mean absolute value of the linking number between R_n (or P_n) and S has a scaling with respect to the length of the walk (or the polygon) of the form*

$$E[|L(R_n, S)|] \approx O(\sqrt{n}). \tag{15}$$

In order to prove this theorem, we will need the following theorem from probability theory by Stein (1972). It is used to obtain a bound between the distribution of a sum of the terms of an m-dependent sequence of random variables (that is $X_1, X_2, \ldots, X_s$ is independent of $X_t, X_{t+1}, \ldots$, provided $t - s \geqslant m$) and a standard normal distribution.

Theorem 3.10. *Let $x_1, x_2, \ldots, x_n$ be a sequence of stationary and m-dependent random variables such that $E[x_i] = 0$, $E[x_i^2] < \infty$ for each i and*

$$0 < C = \lim_{n\to\infty} \frac{1}{n} E\left[\left(\sum_{1\leqslant i\leqslant n} x_i\right)^2\right] < \infty, \tag{16}$$

then $\frac{1}{\sqrt{nC}} \sum_{1\leqslant i\leqslant n} x_i$ converges to the standard normal random variable. Furthermore, if we let $\Phi(a) = \frac{1}{\sqrt{2\pi}} \int_{(-\infty,a]} \mathrm{e}^{-\frac{x^2}{2}}\, \mathrm{d}x$ be the distribution function of the standard normal random variable, then we have

$$\left| P\left(\frac{1}{\sqrt{nC}} \sum_{1\leqslant i\leqslant n} x_i \leqslant a\right) - \Phi(a)\right| \leqslant \frac{A}{\sqrt{n}} \tag{17}$$

for some constant $A > 0$

Proof of theorem 3.9. Note that the confined space can be any convex space and a simple closed curve may be of any knot type, but for simplicity we will assume that the confined space is the cube given by the set $C = \left\{(x, y, z) : -\frac{1}{2} \leqslant x, y, z \leqslant \frac{1}{2}\right\}$ and that the simple closed curve S is the circle on the xy-plane whose equation is $x^2 + y^2 = r^2$, where $r > 0$ is a constant that is less than $\frac{1}{2}$. Let ϵ_j be the sum of the ± 1's assigned to the crossings between the projections of the jth edge l_j of P_n and S, we need to take the sum since, in this case, the projections of l_j may have up to two crossings with S. It is easy to see that $\epsilon_j = 0, \pm 1, \pm 2$ for each j, the ϵ_j's have the same distributions and that, by symmetry, we have $E[\epsilon_j] = 0$ for any j. If $|i - j| > 1 \bmod (n)$, then ϵ_i and ϵ_j are independent; hence, we have $E[\epsilon_i\epsilon_j] = 0$. Let $p' = E\left[\epsilon_j^2\right]$ and $u' = E[\epsilon_i\epsilon_{i+1}]$. Then, if $n = 3$,

$$\begin{aligned} V\left(\sum_{1\leqslant i\leqslant 3} \epsilon_i\right) &= E\left[\left(\sum_{1\leqslant i\leqslant 3} \epsilon_i\right)^2\right] = \sum_{1\leqslant i\leqslant 3} E[\epsilon_i^2] + 2\sum_{1\leqslant i,j\leqslant 3} E[\epsilon_i\epsilon_j] \\ &= 3p' + 6u' = 3(p' + 2u') > 0. \end{aligned} \tag{18}$$

Thus, we have $p' + 2u' > 0$, where $p' = E\left[\epsilon_i^2\right]$ and $u' = E[\epsilon_i\epsilon_j]$. It follows that

$$0 < C = \frac{1}{n} E\left[\left(\sum_{1\leqslant j\leqslant n} \epsilon_j\right)^2\right] = p' + 2u' \tag{19}$$

for any n. If we ignore the last term ϵ_n in the above, then we still have

$$0 < C = \lim_{n\to\infty} \frac{1}{n-1} E\left[\left(\sum_{1\leqslant j\leqslant n-1} \epsilon_j\right)^2\right] = p' + 2u'. \tag{20}$$

Furthermore, the sequence $\epsilon_1, \epsilon_2, \ldots, \epsilon_{n-1}$ is a stationary and 2-dependent random number sequence since the ϵ_j's have the same distributions, and what happens to $\epsilon_1, \ldots, \epsilon_j$ clearly do not have any affect to what happens to $\epsilon_{j+2}, \ldots, \epsilon_{n-1}$ (hence they are independent).

By theorem 3.10, there exists a constant $A > 0$ such that

$$\begin{aligned}
&\left| P\left(\frac{1}{\sqrt{(n-1)(p'+2u')}} \sum_{1\leqslant i\leqslant n-1} \epsilon_i \leqslant \alpha\right) - \Phi(\alpha)\right| \leqslant \frac{A}{\sqrt{n-1}} \\
&\qquad \Rightarrow \left| P\left(\sum_{1\leqslant i\leqslant n-1} \epsilon_i \leqslant \alpha\sqrt{(n-1)(p'+2u')}\right) - \Phi(\alpha)\right| \leqslant \frac{A}{\sqrt{n-1}} \\
&\qquad \Rightarrow \left| P\left(\sum_{1\leqslant i\leqslant n-1} \epsilon_i \leqslant w\right) - \Phi\left(\frac{w}{\sqrt{(n-1)(p'+2u')}}\right)\right| \leqslant \frac{A}{\sqrt{n-1}},
\end{aligned} \tag{21}$$

where $w = \alpha\sqrt{(n-1)(p'+2u')}$.

Note that the linking number between the oriented uniform random polygon P_n and S is equal to the sum $\frac{1}{2}\sum_{1\leqslant i\leqslant n} \epsilon_i$.

Then as $n \to \infty$, $\frac{1}{2}\sum_{1\leqslant i\leqslant n-1} \epsilon_i \to Z$, where Z is a random variable that follows the normal distribution with mean 0 and variance $\sigma^2 = \frac{1}{4}(n-1)(p'+2u')$, i.e. $N\left(0, \frac{1}{4}(n-1)(p'+2u')\right)$. So the random variable $\left|\frac{1}{2}\sum_{1\leqslant i\leqslant n-1} \epsilon_i\right|$ follows the half normal distribution and $E\left[\left|\frac{1}{2}\sum_{1\leqslant i\leqslant n-1} \epsilon_i\right|\right] = \frac{1}{2}(2/\pi(n-1)(p'+2u'))^{1/2} = O(\sqrt{n})$.

Thus,

$$\left| E\left[\left|\frac{1}{2}\sum_{1\leqslant i\leqslant n-1} \epsilon_i\right|\right] - E\left[\left|\frac{1}{2}\epsilon_n\right|\right]\right| \leqslant E\left[|Lk(R_n, S)|\right] \leqslant E\left[\left|\frac{1}{2}\sum_{1\leqslant i\leqslant n-1} \epsilon_i\right|\right] + E\left[\left|\frac{1}{2}\epsilon_n\right|\right]. \tag{22}$$

But $E\left[\left|\frac{1}{2}\sum_{1\leqslant i\leqslant n-1} \epsilon_i\right|\right] = O(\sqrt{n})$ and $E[|\epsilon_n|]$ is a constant independent of n, so $E[|Lk(P_n, S)|] = O(\sqrt{n})$.

The proof carries through similarly in the case of a uniform random walk R_n and a simple closed curve S. Note that in that case one does not have to ignore the last term in (19) and has to carry through an averaging procedure over all possible projections as well as in the proof of theorem 3.2. The result is again $E[|Lk(R_n, S)|] = O(\sqrt{n})$. □

Remark 3.11. One can understand the above result, for the case of a uniform random polygon, and a simple closed curve using the following argument proposed by De Witt Sumners. We know that $Lk(R_n, S)$ equals the algebraic number of times the polygon R_n passes through the surface S_1 with S as perimeter. Let $m = O(n)$ be the number of times the polygon R_n passes through the surface S_1 with S as perimeter. We can assume that $k = O(n)$ of those edges are non-consecutive (note that consecutive edges cancel each other and do not contribute to the

linking number). Then we associate a variable $x_i = \pm 1$ to each one of those non-consecutive edges depending upon the orientation of the edge. Observe that these are independent random variables. Then $x_i, 1 \leqslant i \leqslant k$ is a one-dimensional random walk; thus, the distance of the starting point of the random walk of $k = O(n)$ steps is of the order $O(\sqrt{n})$, and thus $|Lk(R_n, S)| = O(\sqrt{n})$.

Remark 3.12. We stress that it is the mean absolute value of the linking number of two uniform random walks or polygons in confined space and the mean absolute writhe and self-linking number of a uniform random walk or polygon in confined space that are of greatest interest for us. Although their analysis is much more challenging, we propose that these provide a clearer picture of the scaling of the quantities associated with the geometry and topology of the chains or polygons. We study these numerically in the next section. It would be interesting for future work to prove the numerical scaling analytically.

4. Numerical results

In this section, we will describe results obtained by simulations of uniform random walks and polygons in confined space and of equilateral random walks. First we consider the scaling of the mean squared writhe, $E[\mathrm{Wr}^2]$, and the mean absolute value of the writhe, $E\,[|\mathrm{Wr}|]$, of an oriented uniform random walk and polygon of n edges in confined space. Then we study the scaling of the mean absolute value of the linking number $E[|Lk|]$ between an oriented uniform random walk or polygon of n edges and a fixed oriented simple closed curve; and the mean absolute value of the linking number between two oriented uniform random walks or polygons of n edges each. Finally we study the scaling of the mean absolute value of the linking number between two oriented equilateral random walks of n edges whose starting points coincide, $\langle \mathrm{ALN} \rangle$, and the scaling of the mean absolute value of the self-linking number of an oriented equilateral random walk of n edges, with respect to the number of edges, $\langle \mathrm{ASL} \rangle$.

4.1. Generation of data

To generate uniform random walks and polygons confined in C^3, each coordinate of a vertex of the uniform random walk was drawn from a uniform distribution on [0, 1], and to generate equilateral random walks, each edge vector was drawn from a uniform distribution on S^2.

For the computation of the linking number or the writhe of uniform or equilateral random walks or polygons, we used the algorithm by Klenin and Langowski (2000), which is based on the Gauss integral. For each pair of edges e_1, e_2, their linking number is computed as the signed area of two antipodal quadrangles defined by the two edges over the area of the 2-sphere.

We estimated the linking, the writhe and the self-linking numbers between oriented uniform random walks and polygons by analyzing pairs of 10 subcollections of 500 oriented uniform random walks or polygons ranging from 10 edges to 100 edges by a step size of 10 edges, for which we calculated the mean and then computed the mean of the 10 means for our estimate. We did the same for the study of the linking number and the self-linking number of equilateral random walks. For the computation of the scaling of the linking number between an oriented uniform random walk or polygon and an oriented simple closed planar curve, first we considered a fixed square and we analyzed 10 subcollections of 500 oriented uniform random walks or polygons ranging from 10 edges to 100 edges by a step size of 10 edges. In order to illustrate that the result holds for any fixed knot, we also considered a fixed trefoil and we did the same analysis.

4.2. Analysis of data

In this section, first we analyze our data on the mean squared writhe of an oriented uniform random walk and for an oriented uniform random polygon in confined space. We fit our data to a function of the form predicted in theorem 3.2 using the method of least-squares. Then we check the variance and the coefficient of determination. The coefficient of determination takes values between 0 and 1 and is an indicator of how well the curve fits the data. We continue further our study of the writhe of uniform random walks and polygons by analyzing our data on the mean absolute value of the writhe of a uniform random walk or polygon in confined space. We do not have an analytical result for this scaling, but the data are very well fitted to a linear function. Thus, we observe that the distribution of the writhe of a uniform random walk or polygon has the property that the scaling of the mean squared writhe is equal to the scaling of the squared mean absolute writhe. We know that if a random variable X follows the normal distribution, $X \sim N(0, \sigma^2)$, then $|X|$ follows the half-normal distribution and $E[|X|] = \sigma\sqrt{2/\pi}$. Then $|X|/\sigma$ follows the χ-distribution and X^2/σ^2 follows the χ^2-distribution with one degree of freedom and mean 1. Thus, $E[X^2/\sigma^2] = 1$, $E[X^2] = \sigma^2$; hence we have that $E[X^2] \approx E[|X|]^2$. This indicates that the writhe of a uniform random walk or polygon in confined space follows the normal distribution.

By Arsuaga *et al* (2007), it has been proved and confirmed numerically that the mean squared linking number of two oriented uniform random polygons in confined space has a scaling of the form $O(n^2)$. We stress that in the proof of this theorem they have used the fact that the linking number of two closed oriented curves is independent of the projection. For two oriented uniform random walks (open curves) in confined space we have proved that the mean squared linking number has a scaling of the form $O(n^2)$ (theorem 3.1). This is confirmed by our data. Furthermore, we study numerically the mean absolute value of the linking number between two uniform random walks or polygons in confined space. Our data are well fitted to a linear function. This matches the observation of Arsuaga *et al* (2007) that the linking number follows the normal distribution.

The analytic study of the mean absolute linking number is a very challenging problem, but in the same subsection, we study the scaling of the mean absolute linking number of the special case of a uniform random walk or polygon and a fixed simple closed curve in confined space. In theorem 3.9 we proved that it has a scaling of the form $O(\sqrt{n})$ and this agrees with our data. We note that this is different from the mean absolute linking number of two uniform random walks or polygons in confined space for which our numerical results indicate a scaling of the form $O(n)$.

Next we analyze numerical data for the mean squared self-linking number of uniform random walks and polygons. The data strongly support a scaling of the form $O(n^2)$ and, thus, confirm theorem 3.6. We analyze our data for the mean absolute value of the self-linking number of a uniform random walk or polygon in confined space. Again, we observe a scaling of the form $O(n)$ which suggests that the self-linking number of a uniform random walk or polygon in confined space follows the normal distribution.

We conclude that all the above measures follow the same type of distribution. Furthermore this distribution has the property that the square root of the mean squared random variable is equal to the mean absolute random variable. This strengthens our intuition that the linking number follows a normal distribution.

4.2.1. Mean squared writhe and mean absolute writhe of an oriented uniform random walk or polygon in a confined space. Our first numerical study concerns the writhe of a uniform random walk or polygon. By theorem 3.2 the mean squared writhe of an oriented uniform

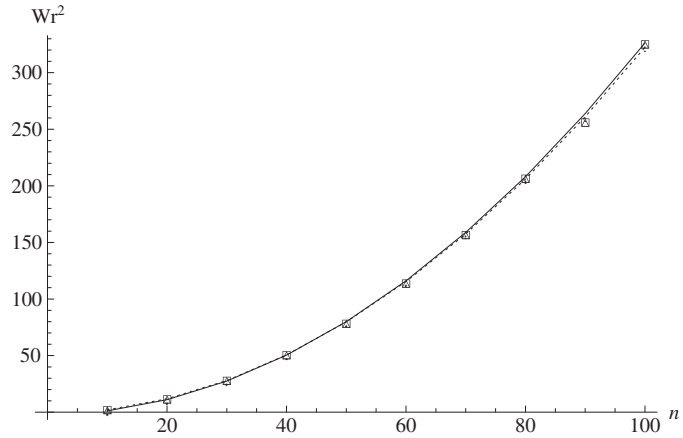

Figure 3. The mean squared writhe of uniform random walks and polygons. Values obtained by computer simulations are shown by triangles and squares, respectively. The black curve is the graph of the mean squared writhe of a uniform random polygon and the dashed curve is the graph of the mean squared writhe of a uniform random walk (open chain).

random polygon grows at a rate $E[\mathrm{Wr}^2] \approx qn^2 + O(n)$. For comparison with this analytical result we calculated the mean squared writhe of an oriented uniform random walk of varying length and of an oriented uniform random polygon of varying length.

Results are shown in figure 3. The black curve in the figure illustrates the scaling of the mean squared writhe of a uniform random polygon with respect to its number of edges and is fitted to a function of the form $qn^2 + a$ where q is estimated to be 0.0329 ± 0.0002 and a is estimated to be -2.1293 ± 0.9466, with a coefficient of determination $R^2 = 0.9998$. Thus, the estimate given in the theorem is strongly supported by the data. The dashed curve in the figure illustrates the scaling of the mean squared writhe of a uniform random walk with respect to its number of edges and is fitted to a function of the form $qn^2 + a$ where q is estimated to be 0.0324 ± 0.0002 and a is estimated to be -1.1499 ± 0.9948, with a coefficient of determination $R^2 = 0.9997$. Hence, the estimate given in the theorem is strongly supported by the data.

Note that q was estimated to be equal to 0.0329 ± 0.0002 and 0.0324 ± 0.0002, respectively, which coincides with the predicted value of q given by Arsuaga *et al* (2007), i.e. $q = 0.0338 \pm 0.024$.

We next estimate the mean absolute value of the writhe of an oriented uniform random walk or polygon. Results are shown in figure 4. The black curve in the figure represents the mean absolute writhe of a uniform random polygon contained in C^3 with respect to its number of edges. The curve is fitted to a function of the form $a + bn$ where a is estimated to be $-0.2372 \pm 0.042\,62$ and b is estimated to be 0.1460 ± 0.0007, with a coefficient of determination $R^2 = 0.9998$. The dashed curve in the figure represents the mean absolute writhe of a uniform random walk contained in C^3 with respect to its number of edges. The curve is fitted to a function of the form $a + bn$ where a is estimated to be -0.1958 ± 0.0630 and b is estimated to be 0.1460 ± 0.0010, with a coefficient of determination $R^2 = 0.9996$. Note that in both cases $b \approx \sqrt{q}$.

As the number of edges of the polygons increases, we observe a growth at a rate $O(n)$. This suggests the conjecture

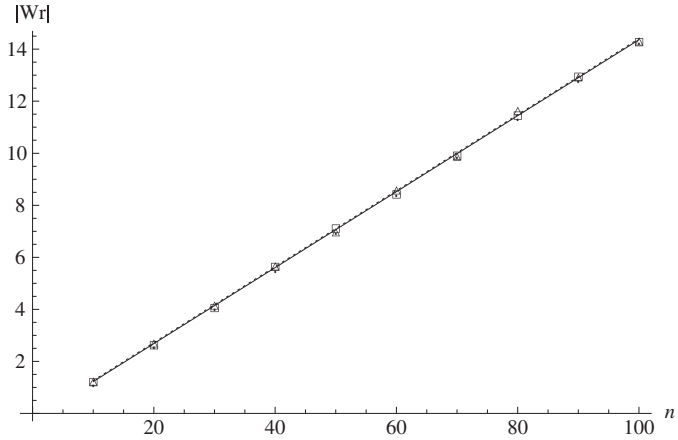

Figure 4. The mean absolute writhe of uniform random walks and polygons in confined space. Values obtained by computer simulations are shown by triangles and squares, respectively. The black curve corresponds to the mean absolute value of the writhe of a uniform random polygon and the dashed curve corresponds to the mean absolute value of the writhe of a uniform random walk (open chain).

Conjecture 4.1.

$$\sqrt{E[\mathrm{Wr}^2]} \sim E[\sqrt{\mathrm{Wr}^2}] = E[|\mathrm{Wr}|]. \tag{23}$$

The proof of conjecture 4.1 would strengthen our intuition that the mean writhe of a uniform random walk or polygon in confined space follows the normal distribution.

Remark 4.2. We observe a difference between the scaling of the mean absolute writhe of a uniform random walk or polygon confined in a cube and that of a self-avoiding polygon in $\mathbb{Z}^3$. In Janse van Rensburg *et al* (1993) the mean absolute writhe of a self-avoiding polygon in $\mathbb{Z}^3$ is proved to have a lower bound of the form $O(\sqrt{n})$. Furthermore, the numerical results in Janse van Rensburg *et al* (1993) and Orlandini *et al* (1994) show a scaling of the form $E[|\mathrm{Wr}|] \approx n^{0.52}$ in the case of self-avoiding walks and polygons in $\mathbb{Z}^3$. Indeed, considering the fact that a uniform random walk or polygon is confined in a cube, and each vertex is chosen with respect to the uniform distribution, as well as the differences between lattice and off-lattice models, we should expect that the mean absolute writhe of a uniform random walk or polygon would scale faster than that of a self-avoiding polygon in $\mathbb{Z}^3$.

Remark 4.3. Comparing our numerical results to those concerning phantom chains in confined space presented in Micheletti *et al* (2006), we notice a significant difference. In Micheletti *et al* (2006) it is shown that the mean absolute writhe of phantom chains in confined space has a scaling of the form $O(n^{0.75})$. This difference in the order of scaling may be due to the fact that in the URP model there is no fixed bond length. The phantom polygons considered in Micheletti *et al* (2006) are equilateral, so the polygon cannot get knotted as easily as in the case of the URP model, i.e. a polygon should have a greater length in order to reach the abilities of conformation of a uniform random polygon confined in a cube of fixed length. Furthermore, the spatial constraint considered in Micheletti *et al* (2006) is of a different nature than the one that we study. One more aspect that could influence the numerical

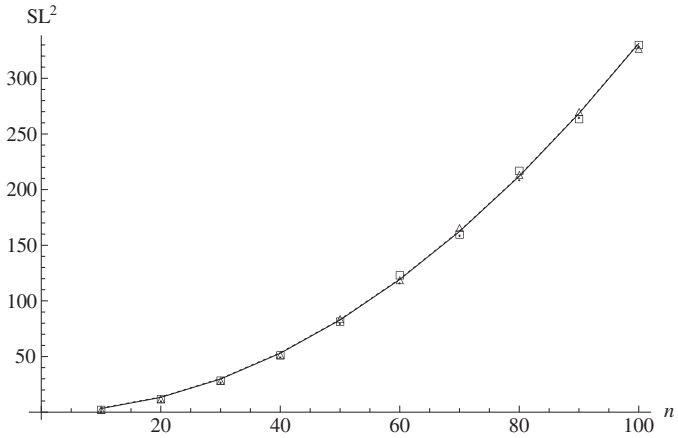

Figure 5. The mean squared self-linking number of uniform random walks and polygons. Values obtained by computer simulations are shown by triangles and squares, respectively. The black curve is the graph of the mean squared self-linking number of uniform random polygons contained in C^3. The dashed curve is the graph of the mean squared self-linking number of uniform random walks contained in C^3.

results presented in Micheletti *et al* (2006) is that the writhe is computed by using only 500 projections and taking their average.

4.2.2. Mean squared self-linking number and mean absolute self-linking number of an oriented uniform random walk or polygon in a confined space. Our second numerical study concerns the self-linking of an oriented uniform random walk or polygon in confined space.

The self-linking number is defined as (Banchoff 1976)

$$\mathrm{SL}(l) = \mathrm{Wr}(l) + \frac{1}{2\pi}\tau(l). \tag{24}$$

The first term is the writhe of the curve, which we studied in the previous subsection, and the second term is the total torsion of the curve, $\tau(l)$ divided by 2π which measures how much the curve deviates from being planar.

We calculated the mean squared self-linking number of an oriented uniform random walk of varying length and of an oriented uniform random polygon of varying length.

Results are shown in figure 5. The black curve in the figure illustrates the mean squared self-linking number of uniform random polygons in confined space and is fitted to a function of the form $qn^2 + a$ where q is estimated to be 0.0331 ± 0.0003 and a is estimated to be -0.0131 ± 1.5226, with a coefficient of determination $R^2 = 0.9993$. The dashed curve in the figure illustrates the mean squared self-linking number of uniform random walks in confined space and is fitted to a function of the form $qn^2 + a$ where q is estimated to be 0.0331 ± 0.0002 and a is estimated to be 0.3980 ± 1.1778, with a coefficient of determination $R^2 = 0.9996$. Thus the estimates given in the theorem are strongly supported by the data.

Note that the estimated value of q coincides with the predicted value of q given by Arsuaga *et al* (2007), i.e. $q = 0.0338 \pm 0.024$.

We next estimate the mean absolute value of the self-linking number of an oriented uniform random walk or polygon. Results are shown in figure 6. The black curve in the figure illustrates the mean absolute self-linking number of uniform random polygons contained in

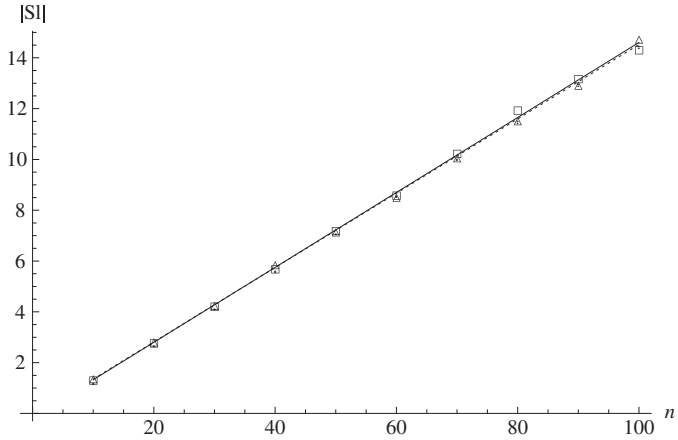

Figure 6. Mean absolute self-linking number of uniform random walks and polygons in confined space. Values obtained by computer simulations are shown by triangles and squares, respectively. The black curve is the graph of the mean absolute self-linking number of uniform random polygons and the dashed curve is the graph of the mean absolute self-linking number of uniform random walks.

C^3 and is fitted to a curve of the form $a + bn$ where a is estimated to be -0.1516 ± 0.1051 and b is estimated to be 0.1476 ± 0.0017, with a coefficient of determination $R^2 = 0.9990$. The dashed curve in the figure illustrates the mean absolute self-linking number of uniform random walks contained in C^3 and is fitted to a curve of the form $a + bn$ where a is estimated to be -0.1089 ± 0.0809 and b is estimated to be 0.1463 ± 0.0013, with a coefficient of determination $R^2 = 0.9994$. Note that in both cases $b \approx \sqrt{q}$. This suggests the conjecture

Conjecture 4.4.

$$\sqrt{E[SL^2]} \sim E[\sqrt{SL^2}] = E[|SL|]. \tag{25}$$

The proof of conjecture 4.4 would suggest that the mean self-linking number of a uniform random walk or polygon follows the normal distribution.

4.2.3. Mean squared and mean absolute linking number of oriented uniform random walks in a confined space. Our third numerical study concerns the linking between two oriented uniform random walks or polygons in confined space. By theorem 3.1, the mean squared linking number between two oriented uniform random polygons grows at a rate $O(n^2)$ with regard to the number of edges of the polygons.

We calculated the mean squared linking number between two oriented uniform random walks of varying lengths and the results are shown in figure 7. The curve in the figure is fitted to a function of the form $a + \frac{q}{2}n^2$ where a is estimated to be -0.5499 ± 0.9297 and q is estimated to be 0.0346 ± 0.0004, with a coefficient of determination $R^2 = 0.9991$. Clearly, as the number of edges of the walks increases, we observe a growth at a rate $O(n^2)$, as expected from theorem 3.5.

Next we calculated the mean absolute linking number between two oriented uniform random walks or polygons of varying length and results are shown in figure 8. The black curve in the figure illustrates the mean absolute linking number of two uniform random

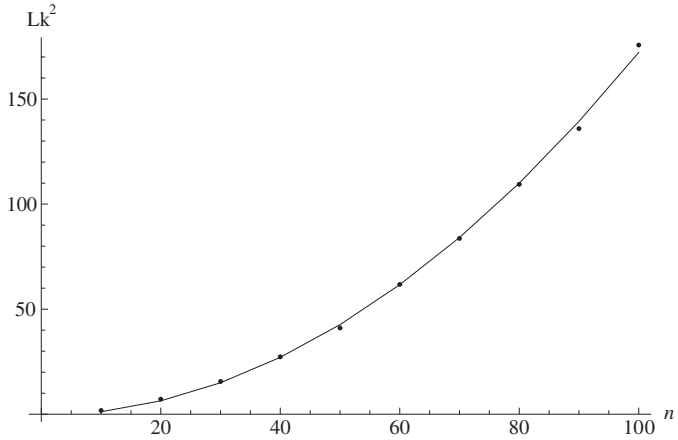

Figure 7. The mean absolute linking number of two uniform random walks contained in C^3.

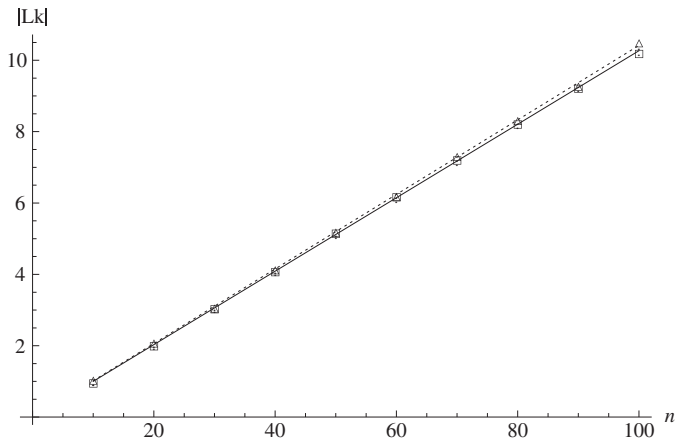

Figure 8. The mean squared linking number of two uniform random walks and polygons contained in C^3. Values obtained by computer simulations are shown by triangles and squares, respectively. The black curve is the graph of the mean absolute linking number of two uniform random polygons and the dashed curve is the graph of the mean absolute linking number of two uniform random walks.

polygons contained in C^3 and is fitted to a function of the form $a + bn$ where a is estimated to be -0.0300 ± 0.0256 and b is estimated to be 0.1030 ± 0.0004, with a coefficient of determination $R^2 = 0.9999$. The dashed curve in the figure illustrates the mean absolute linking number of two uniform random walks contained in C^3 and is fitted to a function of the form $a + bn$ where a is estimated to be -0.0251 ± 0.0350 and b is estimated to be 0.1044 ± 0.0006, with a coefficient of determination $R^2 = 0.9998$. Clearly, as the number of edges of the polygons increases, we observe a growth at a rate $O(n)$. This suggests the conjecture

Conjecture 4.5.

$$\sqrt{E[Lk^2]} \sim E[\sqrt{Lk^2}] = E[|Lk|]. \tag{26}$$

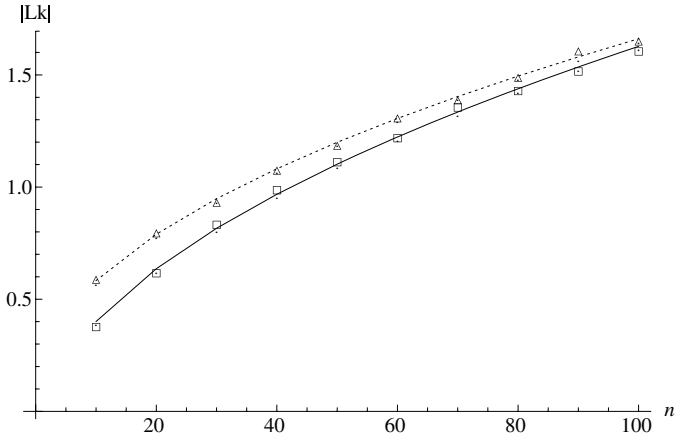

Figure 9. The mean absolute linking number of a uniform random walk or polygon and a fixed square in confined space. Values obtained by computer simulations are shown by triangles and squares, respectively. The black curve is the graph of the mean absolute linking number of a uniform random polygon and a fixed square and the dashed curve is the graph of the mean absolute linking number of a uniform random walk and a fixed square.

Remark 4.6. Note that (26) agrees with our intuition and with the numerical results obtained by Arsuaga *et al* (2007) that the linking number of two uniform random walks or polygons in confined space follows the normal distribution. Indeed, if $Lk \sim N(0, \sigma^2)$, where $\sigma^2 = O(n^2)$, then $|Lk|$ follows the half-normal distribution and $E[|Lk|] = \sigma\sqrt{2/\pi} = O(n)$. Then $\frac{|Lk|}{\sigma}$ follows the χ-distribution and Lk^2/σ^2 follows the χ^2-distribution with one degree of freedom and mean 1. Thus, $E[Lk^2/\sigma^2] = 1$, $E[Lk^2] = \sigma^2 = O(n^2)$, hence we have that $E[Lk^2] \approx E[|Lk|]^2$.

Our next numerical study concerns the linking between an oriented uniform random walk or polygon and a fixed simple closed curve in confined space. By theorem 3.9, the mean absolute value of the linking number between an oriented uniform walk or polygon of n edges and a fixed oriented simple closed curve in confined space has a scaling of the form $O(\sqrt{n})$. First we consider the oriented square S_1 defined by the sequence of vertices (0.1, 0.1, 0.5), (0.9, 0.1, 0.5), (0.9, 0.9, 0.5), (0.1, 0.9, 0.5), (0.1, 0.1, 0.5) and a uniform random walk or polygon. The results of our simulations can be seen in figure 9. The black curve shows the growth rate of the mean absolute linking number of a uniform random polygon and S_1. The data are fitted to a function of the form $a + b\sqrt{n}$, where a is estimated to be -0.1665 ± 0.0212 and b is estimated to be 0.1794 ± 0.0029, with a coefficient of determination $R^2 = 0.9980$. The dashed curve shows the growth rate of the mean absolute linking number of a uniform random walk and S_1. The data are fitted to a function of the form $a + b\sqrt{n}$, where a is estimated to be 0.0848 ± 0.0153 and b is estimated to be 0.1576 ± 0.0021, with a coefficient of determination $R^2 = 0.9986$. Thus, we can see that the data are consistent with theorem 3.9.

To illustrate that the growth rate of the mean absolute linking number does not depend on the knot type of the fixed simple closed curve, we consider the oriented trefoil S_2 defined by the sequence of vertices (0.9, 0.5, 0.5), (0.1, 0.5, 0.4), (0.5, 0.3, 0.9), (0.6, 0.3, 0.1), (0.2, 0.9, 0.6), (0.5, 0.2, 0.5), (0.9, 0.5, 0.5) and a uniform random walk or polygon. The results of our simulations can be seen in figure 10. The black curve shows the growth rate of the mean absolute linking number of a uniform random polygon and S_2. The data are fitted

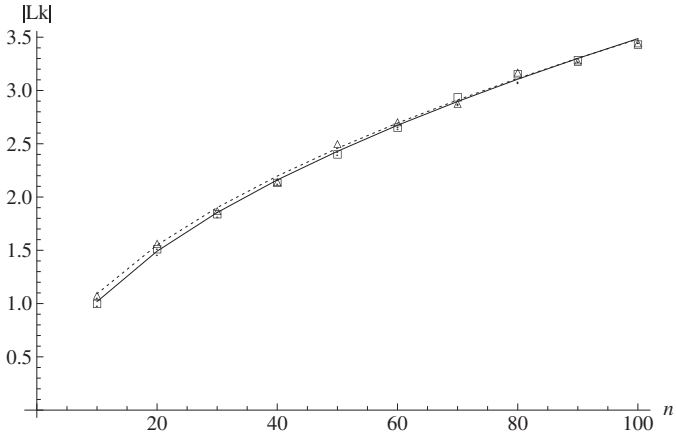

Figure 10. The mean absolute linking number of a fixed trefoil and a uniform random walk or polygon in confined space. Values obtained by computer simulations are shown by triangles and squares, respectively. The black curve is the graph of the mean absolute linking number of a fixed trefoil and a uniform random polygon. The dashed curve is the graph of the mean absolute linking number of a fixed trefoil and a uniform random walk.

to a function of the form $a + b\sqrt{n}$, where a is estimated to be -0.1232 ± 0.0381 and b is estimated to be 0.3610 ± 0.00514, with a coefficient of determination $R^2 = 0.9984$. The dashed curve shows the growth rate of the mean absolute linking number of a uniform random walk and S_2. The data are fitted to a function of the form $a + b\sqrt{n}$, where a is estimated to be -0.0142 ± 0.0432 and b is estimated to be 0.3496 ± 0.0058, with a coefficient of determination $R^2 = 0.9978$. Hence, the data confirm our analytical result in theorem 3.9 for any fixed simple closed curve.

4.3. Numerical results on equilateral random walks

Equilateral random walks are widely used to study the behavior of polymers under θ-conditions. It is of great interest to study the scaling of the linking number, the self-linking number and the writhe of equilateral random walks and polygons. These situations are also much more complex in comparison to the uniform models studied in previous sections, since the probability of crossing of two edges depends on their distance and the probability of positive or negative crossing is independent upon the previous edges. In this section, we present numerical results concerning equilateral random walks and polygons. It would be of great interest to have rigorous proofs of the scaling observed.

4.3.1. Mean absolute self-linking number of an equilateral random walk. In this section we discuss our numerical results on the scaling of the self-linking number of a random walk.

The self-linking number of a random walk X is equal to the writhe plus the total torsion of the random walk:

$$\begin{aligned} E[|\mathrm{SL}(X)|] &= E[|\mathrm{Wr}(X) + \tau(X)|] \leqslant E[|\mathrm{Wr}(X)|] + E[|\tau(X)|] \\ &= E[|\mathrm{Wr}(X)|] + E\left[\left|\sum_i \phi_i\right|\right] \leqslant E[|\mathrm{Wr}(X)|] + E\left[\sum_i |\phi_i|\right] \end{aligned} \tag{27}$$

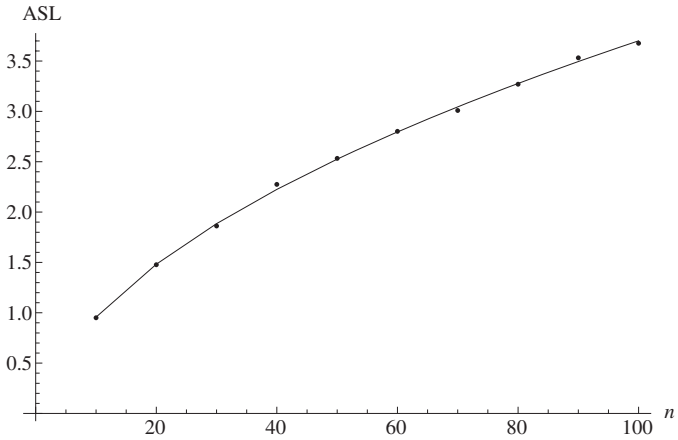

Figure 11. Mean absolute value of the self-linking number of an equilateral random walk.

where $E\left[\sum_i |\phi_i|\right]$ has been proved to be approximately equal to $n\frac{\pi}{2} - \frac{3\pi}{8}$ (Plunkett *et al* 2007, Grosberg 2008).

Previous numerical results (Orlandini *et al* 1994) suggest that the average of the absolute value of the writhe of an ideal walk increases as $\sqrt{n}$, where n is the length of the walk.

Figure 11 shows the $\langle|\mathrm{ASL}|\rangle$ values obtained in numerical simulations of ideal random walks.

We have fitted the computing data points with the function $a + b\sqrt{n}$, leaving the two parameters a and b free. Then a was estimated to be -0.3131 ± 0.0313 and b was estimated to be 0.4014 ± 0.0042, with a coefficient of determination $R^2 = 0.9991$.

Remark 4.7. We observe that our numerical results concerning the mean absolute self-linking number of an equilateral random walk are consistent with those observed for the mean absolute writhe of a self-avoiding polygon in $\mathbb{Z}^3$ (Janse van Rensburg *et al* 1993, Orlandini *et al* 1994). Recall that in Janse van Rensburg *et al* (1993), the mean absolute writhe of a self-avoiding polygon in $\mathbb{Z}^3$ is shown to have a lower bound of the form $O(\sqrt{n})$. Furthermore, the numerical results in Janse van Rensburg *et al* (1993) and Orlandini *et al* (1994) show a scaling of the form $E[|\mathrm{Wr}|] \approx n^{0.52}$ in the case of self-avoiding walks and polygons in $\mathbb{Z}^3$.

4.3.2. Mean absolute linking number of two equilateral random walks whose starting points coincide. In this section we discuss our numerical results on the scaling of the linking number between two equilateral random walks whose starting points coincide. Figure 12 shows the $\langle\mathrm{ALN}\rangle$ values obtained in numerical simulations of ideal random walks in a non-constrained linear form.

By the numerical results presented in the last section, we expect that the mean absolute value of the writhe of an oriented equilateral random walk X of n steps will have a power-law dependence on the length of the walk:

$$\langle\mathrm{AWR}\rangle \approx O(\sqrt{n}). \tag{28}$$

Let $X = (X_0, X_1, \ldots, X_n)$ and $Y = (Y_0, Y_1, \ldots, Y_n)$ denote two oriented equilateral random walks of length n whose starting points coincide, i.e. $X_0 = Y_0 = 0$. One can use the scaling of $\langle\mathrm{AWR}\rangle$ to give information concerning the scaling of $\langle\mathrm{ALN}\rangle$ as follows.

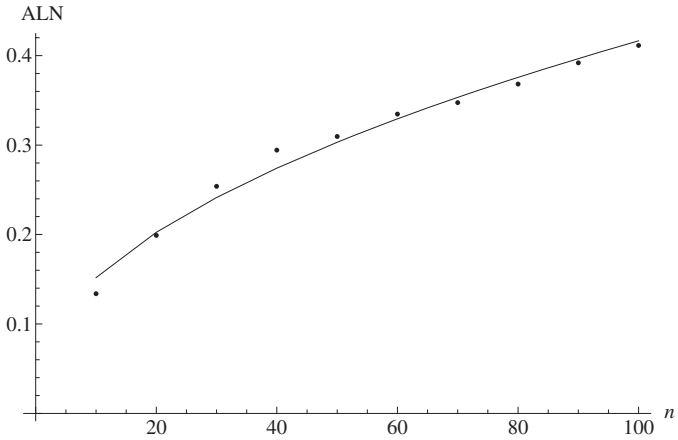

Figure 12. The mean absolute value of the linking number of two equilateral random walks whose starting points coincide.

We define X–Y to be the oriented equilateral random walk of $2n$ steps $(Y_n, \ldots, Y_1,\ Y_0 = X_0,\ X_1, \ldots, X_n)$. Its writhe then is

$$\begin{aligned}
\mathrm{Wr}(X-Y) &= \mathrm{Wr}(X) + \mathrm{Wr}(-Y) + 2L(X,-Y) \\
&\Leftrightarrow \mathrm{Wr}(X-Y) = \mathrm{Wr}(X) + \mathrm{Wr}(-Y) - 2L(X,Y) \\
&\Rightarrow 2L(X,Y) = -\mathrm{Wr}(X-Y) + \mathrm{Wr}(X) + \mathrm{Wr}(Y) \\
&\Rightarrow |L(X,Y)| \leqslant \tfrac{1}{2}(|\mathrm{Wr}(X-Y)| + |\mathrm{Wr}(X)| + |\mathrm{Wr}(Y)|) \\
&\Rightarrow |L(X,Y)| \leqslant \tfrac{1}{2}(O(\sqrt{2n}) + O(\sqrt{n}) + O(\sqrt{n})) \\
&\Rightarrow |L(X,Y)| \leqslant O(\sqrt{n}). \qquad (29)
\end{aligned}$$

We decided therefore to check whether the average of the absolute value of the linking number between two ideal walks increases as $\sqrt{n}$. We have fitted the computing data points with the function $a + b\sqrt{n}$, leaving the two parameters a and b free. Then a was estimated to be 0.0294 ± 0.0130 and b was estimated to be 0.0387 ± 0.0018 with a coefficient of determination 0.9839.

5. Conclusions

The measurement of the entanglement of open chains is of great interest for many applications, such as the study of the properties of polymer melts. In this paper, we focused our study in the case of uniform random walks (open chains) and polygons in confined volumes. In theorems 3.2, 3.1 and 3.6, we gave rigorous proofs that the scaling of the mean squared linking number, the mean squared writhe and the mean squared self-linking number of oriented uniform random walks and polygons in confined space, with respect to their length, is of the form $O(n^2)$.

Further, we are interested in the mean absolute value of the linking number of two uniform random walks or polygons in confined space. In this direction, we prove in theorem 3.9 that the mean absolute value of the linking number of an oriented uniform random walk or polygon and a fixed oriented simple closed curve in confined space is of the form $O(\sqrt{n})$. Our numerical results confirm the analytical prediction and furthermore suggest that for two oriented uniform

random walks or polygons in confined space $\sqrt{E[lk^2]} \sim O(n) \sim E[\sqrt{lk^2}]$. A possible direction for future work would be to prove these results analytically.

Ideal random walks are used to model the behaviour of polymers under θ-conditions. We have analyzed numerically the scaling of the mean absolute value of the linking number between two equilateral random walks of n steps and the mean absolute value of the self-linking number of an equilateral random walk of n steps. Both appear to scale as $O(\sqrt{n})$. An important direction for future work is to complete the analysis and provide proofs for the scaling of the self-linking number and that of the linking number of equilateral random walks and polygons.

Acknowledgments

We would like to express our thanks to D Theodorou and C Tzoumanekas for sharing with us the problem of measuring the entanglement of a polymer melt. Also, we would like to thank A Stasiak, D W Sumners and G Dietler for very useful discussions.

References

Agarwal P K, Edelsbrunner H and Wang Y 2004 Computing the writhing number of a polygonal knot *Discrete Comput. Geom.* **32** 37–53

Arsuaga J, Vazquez M, Trigueros S, Sumners D W and Roca J 2002 Knotting probability of DNA molecules confined in restricted volumes: DNA knotting in phage capsids *Proc. Natl Acad. Sci. USA* **99** 5373–7

Arsuaga J, Blackstone T, Diao Y, Karadayi E and Saito M 2007a The linking of uniform random polygons in confined spaces *J. Phys. A: Math. Theor.* **40** 1925–36

Arsuaga J, Blackstone T, Diao Y, Hinson K, Karadayi E and Saito M 2007b Sampling large random knots in a confined space *J. Phys. A: Math. Theor.* **40** 11697–711

Arsuaga J, Diao Y and Vazquez M 2009 Mathematical methods in DNA topology: applications to chromosome organization and site-specific recombination *Mathematics of DNA Structure, Function and Interactions* ed C J Benham, S Harvey, W K Olson, D W Sumners and D Swigon (New York: Springer Science + Business Media) pp 7–36

Arsuaga J, Borgo B, Diao Y and Sharein R 2009 The growth of the mean average crossing number of equilateral polygons in confinement *J. Phys. A: Math. Theor.* **42** 465202–11

Arteca G A 1997 Self-similarity in entanglement complexity along the backbones of compact proteins *Phys Rev.* E **56** 4516–20

Arteca G and Tapia O 2000 Relative measure of geometrical entanglement to study folding-unfolding transitions *Int. J. Quantum Chem.* **80** 848–55

Banchoff T 1976 Self linking numbers of space polygons *Indiana Univ. Math. J.* **25** 1171–88

Barbi M, Mozziconacci J and Victor J 2005 How does the chromatin fiber deal with topological constraints *Phys. Rev.* E **71** 031910

Bauer W R, Crick F H C and White J H 1980 Supercoiled DNA *Am. Sci.* **243** 118

Berger M A and Prior C 2006 The writhe of open and closed curves *J. Phys. A: Math. Gen.* **39** 8321–48

Buck G, Scharein R G, Schnick J and Simon J 2008 Accessibility and occlusion of biopolymers, ray tracing of radiating tubes and the temperature of a tangle *Phys. Rev.* E **77** 011803-6

Calugreanu G 1961 Sur les classes d'isotopie des noeuds tridimensionnels et leurs invariants *Czech. Math. J.* **11** 588–625

Calvo J A 1999 Geometric knot spaces and polygonal isotopy arXiv:math/9904037v2

Crick F H 1976 Linking numbers and nucleosomes *Proc. Natl Acad. Sci. USA* **73** 2639–43

de Gennes P G 1979 *Scaling Concepts in Polymer Physics* (Ithaca, NY: Cornell University Press)

Diao Y, Pippenger N and Sumners D W 1993 On random knots *J. Knot Theory Ramif.* **3** 419–29

Diao Y 1995 The knotting of equilateral polygons in $\mathbb{R}^3$ *J. Knot Theory Ramif.* **4** 189–96

Diao Y, Dobay A, Kushner R B, Millett K and Stasiak A 2003 The average crossing number of equilateral random polygons *J. Phys. A: Math. Gen.* **36** 11561–74

Diao Y, Dobay A and Stasiak A 2005 The average inter-crossing number of equilateral random walks and polygons *J. Phys. A: Math. Gen.* **38** 7601–16

Dobay A, Dubochet J, Millett K, Sottas P and Stasiak A 2003 Scaling behavior of random knots *Proc. Natl Acad. Sci. USA* **100** 5611–15

Edvinsson T, Elvingson C and Arteca G 2000 Variations in molecular compactness and chain entanglement during the compression of grafted polymers *Macromol. Theory Simul.* **9** 398–406

Edwards F 1967 Statistical mechanics with topological constraints: I *Proc. Phys. Soc.* **91** 513–9

Edwards F 1968 Statistical mechanics with topological constraints: II *J. Phys. A: Gen. Phys.* **1** 15–28

Ewing B and Millett K 1997 Computational algorithms and the complexity of link polynomials *Progress in Knot Theory and Related Topics* (Paris: Herman) pp 51–68

Freedman M H, He Z-X and Wang Z 1994 Mobius energy of knots and unknots *Ann. Math.: Second Ser.* **134** 189–229

Freyd P, Yetter D, Hoste J, Lickorish W, Millett K and Ocneanu A 1985 A new polynomial invariant for knots and links *Bull. Am. Math. Soc.* **12** 239–46

Fuller B 1978 Decomposition of the linking number of a closed ribbon: a problem from Molecular biology *Proc. Natl Acad. Sci. USA* **75** 3557–61

Gauss K F 1877 *Zur Mathematischen Theorie der Electrodynamischen Wirkungen* Werke vol 5 Königl. Ges. Wiss. Göttingen (Göttingen), p 605

Grassberger P 2001 Opacity and entanglement of polymer chains *J. Phys. A: Math. Gen.* **34** 99959–63

Grosberg Y 2008 Total curvature and total torsion of a freely jointed circular polymer with $n \gg 1$ segments *Macromolecules* **41** 4524–7

Hirshfeld A C 1997 Knots and Physics: old wine in new bottles *Am. J. Phys.* **66** 1060–6

Hoidn P, Kusner R B and Stasiak A 2002 Quantization of energy and writhe in self-repelling knots *New J. Phys.* **4** 20.1–11

Holmes V and Cozzarelli N R 2000 Closing the ring: links between SMC proteins and chromosome partitioning, condensation and supercoiling *Proc. Natl Acad. Sci. USA* **97** 1322–4

Iwata K and Edwards S F 1989 New model of polymer entanglement: localized Gauss integral model *J. Chem. Phys.* **90** 084567–15

Janse van Rensburg E J, Orlandini E, Sumners D W, Tesi M C and Whittington S G 1993 The writhe of a self-avoiding polygon *J. Phys. A: Math. Gen.* **26** L981–6

Jones V F R 1985 A polynomial invariant of knots via von Neumann algebras *Bull. Am. Math. Soc.* **12** 103–12

Katritch V, Bednar J, Michoud D, Scharein R G, Dubochet J and Stasiak A 1996 Geometry and Physics of knots *Nature* **384** 142–5

Kauffmann L H 2001 *Knots and Physics (Ser. Knots Everything* vol 1*)* (Singapore: World Scientific)

Klenin K and Langowski J 2000 Computation of writhe in modelling of supercoiled DNA *Biopolymers* **54** 307–17

Kung W and Kamien R D 2003 Topological constraints at the Theta-point: closed loops at two loops *Europhys. Lett.* **64** 323–9

Lacher R C and Sumners D W 1991 *Data Structures and Algorithms for the Computation of Topological Invariants of Entanglements: Link, Twist and Writhe Computer Simulations of Polymers* ed R J Roe (New York: Prentice-Hall) p 365

Laing C and Sumners D W 2006 Computing the writhe on lattices *J. Phys. A: Math. Gen.* **39** 3535–43

Laing C and Sumners D W 2008 The writhe of oriented polygonal graphs *J. Knot Theory. Ramifications.* **17** 1575–94

Liu Z and Chan H S 2008 Efficient chain moves for Monte Carlo simulations of a wormlike DNA model *J. Chem. Phys.* **128** 145104–31

Mansfield M L 1994 Knots in Hamilton cycles *Macromolecules* **27** 5924–6

Martinez-Robles M L, Witz G, Hernandez P, Schvartzman J B, Stasiak A and Krimer D B 2009 Interplay of DNA supercoiling and catenation during the segregation of sister duplexes *Nucl. Acids Res.* **37** 5126–37

McMillen T and Goriely A 2002 Tendril perversion in intrinsically curved rods *J. Nonlinear Sci.* **12** 241–81

Micheletti C, Marenduzzo D, Orlandini E and Sumners D W 2006 Knotting of random ring polymers in confined spaces *J. Chem. Phys.* **124** 64903.1–10

Millett K C 2000 *Knots in Hellas'98 (Delphi) (Ser Knots Everything* vol 24*)* (Singapore: World Scientific) pp 306–34

Millett K, Dobay A and Stasiak A 2004 Linear random knots and their scaling behavior *Macromolecules* **38** 601–6

Millett K and Sheldon B 2005 Tying down open knots: a statistical method for identifying open knots with applications to proteins *(Ser. Knots Everything* vol 36*)* (Singapore: World Scientific) pp 203–17

Orlandini E and Whittington S G 2007 Statistical topology of closed curves: some applications in polymer physics *Rev. Mod. Phys.* **79** 611–42

Orlandini E, Tesi M C, Whittington S G, Sumners D W and Janse van Resburg E J 1994 The writhe of a self-avoiding walk *J. Phys. A: Math. Gen.* **27** L333–8

Orlandini E, Tesi M C and Whittington S G 2000 Polymer entanglement in melts *J. Phys. A: Math. Gen.* **33** L181–6

Orlandini E and Whittington S G 2004 Entangled polymers in condensed phases *J. Chem. Phys.* **121** 12094–99

Pippenger N 1989 Knots in random walks *Discrete Appl. Math.* **25** 273–8

Plunkett P, Piatek M, Dobay A, Kern J, Millett K, Stasiak A and Rawdon E 2007 Total curvature and total torsion of knotted polymers *Macromolecules* **40** 3860–67

Przytycki J and Traczyk P 1987 Conway algebras and skein equivalence of links *Proc. Am. Math. Soc.* **100** 744–48

Randell R 1988a A molecular conformation space *Studies in Physical and Theoretical Chemistry* vol 54 ed R C Lacher (Amsterdam: Elsevier) pp 125–40

Randell R 1988b Conformation spaces of molecular rings *Studies in Physical and Theoretical Chemistry* vol 54 ed R C Lacher (Amsterdam: Elsevier) pp 141–56

Ricca R L 2000 *Knots in Hellas '98 (Delphi) (Ser on Knots Everything* vol 1*)* (Singapore: World Scientific) pp 361–79

Reimann G T, Arteca G and Tapia O 2002 Proteins in vacuo: a connection between mean overcrossing number and orientationally averaged collision cross section *J. Chem. Phys.* **4** 4058–64

Rogen P and Fain B 2003 Automatic classification of protein structure by using Gauss integrals *Proc. Natl Acad. Sci.* **100** 119–24

Rubinstein M and Colby R 2006 *Polymer Physics* (Oxford: Oxford University Press)

Sheng Y-J and Tsao H-K 2002 The mobility and diffusivity of a knotted polymer: topological deformation effect *J. Chem. Phys.* **116** 10523–8

Simon J 2009 Applications of Knot Theory *Am. Math. Soc. Proc. Symp. Appl. Math.* **66** 155–82

Stasiak A, Dubochet J, Katrich V and Pieranski P 1998 *Ideal Knots* ed A Stasiak, V Katritch and L H Kauffmann (Singapore: World Scientific) pp 1–19

Stein C 1972 *Proc. 6th Berkeley Symposium on Mathematical Statistics and Probability (Probability Theory)* (Berkeley, CA: University California Press) vol 2 pp 583–602

Stump D, Fraser W and Gates K 1998 The writhing of circular cross-section rods: undersea cables to DNA supercoils *Proc. R. Soc. Lond.* A **454** 2123–56

Sumners D W and Whittington S G 1988 Knots in self-avoiding walks *J. Phys. A: Math. Gen.* **21** 1689–94

Tesi M, Janse van Rensbourg E, Orlandini E and Whittington S G 1994 Knot probability for lattice polygons in confined geometries *J. Phys. A: Math. Gen.* **27** 347–60

Tzoumanekas C and Theodorou D 2006 Topological analysis of linear polymer melts: a statistical approach *Macromolecules* **39** 4592–604

Vologodskii A V and Cozzarelli N R 1993 Monte Carlo analysis of the conformation of DNA catenanes *J. Mol. Biol.* **232** 1130–40

Vologodskii A 1999 Cicular *DNA On-line Biophysics Textbook* ed V Bloomfield http://www.biophysics.org/education/vologodskii.pdf

Weber C, Stasiak A, Fleurant M, De Los Rios P and Dietler G 2006 Gel electrophoresis of DNA knots in weak and strong electric fields *Biophys. J.* **90** 3100–5

White J H and Bauer W R 1969 Self-linking and the Gauss integral in higher dimensions *Am. J. Math.* **91** 693–728

Whittington S G 1992 Topology of polymers *New Scientific Applications of Geometry and Topology* ed D W Sumners *et al Proc. Symp. Applied Math. Soc.* vol 145 (Providence, RI: AMS) pp 73–95