# Linking topology of tethered polymer rings with applications to chromosome segregation and estimation of the knotting length

John F. Marko
*Department of Physics and Astronomy and Department of Biochemistry, Molecular Biology and Cell Biology, Northwestern University, Evanston, Illinois 60208, USA*
(Received 9 February 2009; published 11 May 2009)

The Gauss linking number (Ca) of two flexible polymer rings which are tethered to one another is investigated. For ideal random walks, mean linking-squared varies with the square root of polymer length while for self-avoiding walks, linking-squared increases logarithmically with polymer length. The free-energy cost of linking of polymer rings is therefore strongly dependent on degree of self-avoidance, i.e., on intersegment excluded volume. Scaling arguments and numerical data are used to determine the free-energy cost of fixed linking number in both the fluctuation and large-Ca regimes; for ideal random walks, for $|\mathrm{Ca}|>N^{1/4}$, the free energy of catenation is found to grow $\propto|\mathrm{Ca}/N^{1/4}|^{4/3}$. When excluded volume interactions between segments are present, the free energy rapidly approaches a linear dependence on Gauss linking ($dF/d\mathrm{Ca}\approx 3.7k_BT$), suggestive of a novel "catenation condensation" effect. These results are used to show that condensation of long entangled polymers along their length, so as to increase excluded volume while decreasing number of statistical segments, can drive disentanglement if a mechanism is present to permit topology change. For chromosomal DNA molecules, lengthwise condensation is therefore an effective means to bias topoisomerases to eliminate catenations between replicated chromatids. The results for mean-square catenation are also used to provide a simple approximate estimate for the "knotting length," or number of segments required to have a knot along a single circular polymer, explaining why the knotting length ranges from $\approx$300 for an ideal random walk to $10^6$ for a self-avoiding walk.

DOI: 10.1103/PhysRevE.79.051905 PACS number(s): 87.14.gk, 87.16.Sr, 87.17.Ee, 82.35.Lr

## I. INTRODUCTION

Topological constraints ("entanglements") are key determinants of physical properties of materials composed of flexible polymers [1,2]. In biological systems, control of polymer topology is a matter of life and death, particularly in the case of double-stranded DNA (dsDNA). The process of DNA replication can be expected to leave a large number of remnant links, or *catenations* between "sister" replicated DNAs, given that initially the two parental strands start out linked together once every 10.5 base pairs [3,4]. For a $4.7\times10^6$ base pair (bp) *E. coli* bacterial chromosome, the parental strands are thus linked together roughly $4\times10^5$ times; for the largest human chromosome containing $\approx2.5\times10^8$ bp, the parental strands initially are linked roughly $2\times10^7$ times. Catenations remaining after DNA replication must be removed by type-II topoisomerases [5] which pass dsDNA through dsDNA, for sister chromosomal DNAs to condense and segregate to the two daughter cells during cell division.

This paper is concerned with the statistical mechanics of two flexible polymer rings, each $N$ statistical segments in length, attached to one another by a short (roughly one statistical segment-long) tether [Fig. 1(a)]. This situation is a simple, well-defined model for duplicated chromosomes during the late to terminal stages of DNA replication. Duplicated chromosomes may be held together during the late stages of DNA replication either by terminal regions of late-replicating DNA, chromosome domain structure [4], or by specialized protein linkers (e.g., "cohesin" complexes [6]). More generally, two circular polymers with a fixed distance between specified points on each chain is a basic model for analysis of random linking of polymers [7–14]. Tethering the polymers together eliminates the divergence of the unlinked state entropy with system volume and is appropriate for the biophysical applications of interest here.

### A. Distribution of Gaussian linking number as a tool for quantifying linking complexity

The Gauss linking invariant of two closed curves $\mathbf{r}_1$ and $\mathbf{r}_2$ is

$$\mathrm{Ca}=\frac{1}{4\pi}\oint\oint\frac{d\mathbf{r}_1\times d\mathbf{r}_2\cdot(\mathbf{r}_1-\mathbf{r}_2)}{|\mathbf{r}_1-\mathbf{r}_2|^3}. \quad (1)$$

Here this linking number is denoted Ca (catenation number), following molecular-biological notation for linking of two double-helix DNAs (the notation Lk is conventionally used to denote linking number of the two strands inside a double-helix DNA). The Gauss invariant can alternately be computed as simply the sum of signed crossings of one curve over the other. Ca is a topological invariant: its integer value does not change if the linking topology of the two curves is held fixed. Ca is easily calculated for piecewise smooth curves or for polygons [15], and its statistics are amenable to analytical computation within the framework of flexible polymer theory [7–14,16–23].

For a single snapshot of two finite-length polymers, Ca is a poor classifier of linking topology since it is highly degenerate [24]. Many topologically distinct linkages of two polymers have the same value of Ca, so there is no hope of uniquely identifying any link *via* Ca alone [25–27]. As an important example of this degeneracy problem, the value

 

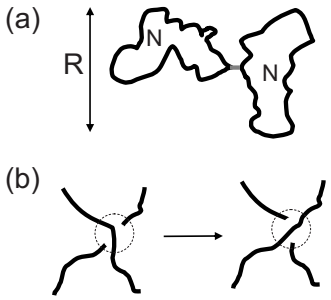


FIG. 1. (a) Tethered polymer rings. Two flexible polymers each $N$ segments in length are attached by a connector segment. (b) Polymers undergo close encounters where two segments are separated by approximately one segment length. At these points the sign of the crossing may be changed by a small shift in position of either segment, without changing the conformation of the remainder of the chains.

Ca=0 taken by unlinked rings is also shared by all other linkages that happen to have the same number of + and − signed crossings. Therefore, observation that Ca=0 for some particular configuration of two rings cannot be used by itself to decide whether that configuration is unlinked. Furthermore, one cannot use Ca to detect linking changes during numerical calculations of polymers where linking topology is to be held fixed. In such calculations one must use topological invariants with lower degeneracies (e.g., the two-variable Alexander polynomial invariant for links [25,26,28], or other polynomial linking invariants) and with much higher calculational complexities than the Gauss invariant (see Ref. [13]).

Despite the uselessness of Ca for precise classification of the topology of individual polymer conformations, the probability distribution of Ca can be useful for characterizing topologies in a statistical ensemble. The distribution of Ca is particularly useful when one is not interested in precise classification or constraint of the linking topology of two polymers, but instead in determining whether two polymers are more or less entangled together under conditions where topology is fluctuating [29].

An example of such a problem is the threshold at which substantial linking (or unlinking) begins to occur in an ensemble of topologies, as a function of some control parameter (e.g., solution conditions, degree of cross-linking, activity of DNA-condensing enzymes). For an ensemble with a high degree of linking complexity, i.e., with a large mean minimal crossing number, the distribution of Ca can be expected to be broad, even though the average and mode of the distribution may be small or zero (for achiral polymers, $\langle \mathrm{Ca} \rangle = 0$ by symmetry). Conversely, for an ensemble with a low degree of linking complexity, one can expect a small average Ca and also a narrow distribution of Ca. The probability distribution of Gaussian linking $P(\mathrm{Ca})$ is in this way useful for characterizing the linking complexity of an ensemble of conformational states.

In this paper the second moment of $P(\mathrm{Ca})$, the mean-square catenation number $\langle \mathrm{Ca}^2 \rangle$, will be given special attention since when $\langle \mathrm{Ca} \rangle$ and $\langle \mathrm{Ca}^2 \rangle$ are both near zero for an ensemble of randomly linked polymers, that ensemble must have a low average linking complexity. In the limit that both the average and mean-square catenation are zero, the ensemble must be unlinked, discounting the possibility that only nontrivially linked states with zero Gauss invariant are being visited during random fluctuations in topology. While logically possible, this would be quite unphysical since a single passage involving segments on different chains changes Ca by $\pm 1$. Fluctuations in topology generated by local passages of polymer segments through one another as is the case for any physically realizable situation (e.g., type-II topoisomerases acting on DNA molecules) cannot keep Ca fixed at any one value. In any physical situation one expects to observe a "random walk" of Ca values in time.

### B. Outline of this paper

Section II discusses the degree of mean-square Gaussian linking for two tethered polymer rings for achiral polymers of random topology. Section II A discusses a simple scaling argument indicating that for ideal random walks (IRWs), $\langle \mathrm{Ca}^2 \rangle \propto N^{1/2}$, a result first obtained by Tanaka for Gaussian polymers using a gauge field theory approach [7]. Section II B shows how Tanaka's result can be extracted from a direct calculation of mean-square Gauss invariant for Gaussian polymers due to Otto [14].

Section II C shows how the case of self-avoiding walks (SAWs) is quite different, presenting a scaling argument suggesting that $\langle \mathrm{Ca}^2 \rangle \propto \ln N$. This qualitative change in topology complexity is a far stronger effect than the small change in coil geometry (i.e., coil radius exponent $\nu$) resulting from introduction of self-avoidance.

Section II D presents Monte Carlo (MC) results for a flexible polymer model, for increasing levels of segment-segment excluded volume (self-avoidance). For IRWs there is excellent agreement with the result of Sec. II B, but for even small amounts of segment excluded volume (thin segments), the logarithmic scaling behavior of Sec. II C is seen.

Section III discusses the implications of the results of Sec. II for the free-energy cost of entanglements. For IRWs, a scaling argument indicates that for large Ca, the free energy should scale as $\sim |\mathrm{Ca}/N^{1/4}|^{4/3}$. For SAWs, the small magnitude of catenation fluctuations translates into a sufficiently large free-energy cost of catenation that in almost all physically realizable situations, catenations should "condense" together, leading to a novel, nearly linear increase in free energy with catenation [30], rather than the quadratic dependence expected from conventional elasticity. In accord with this, a numerical computation of free energy beyond the thermal regime and into the heavily linked regime reveals a linear dependence on linking number, with $dF/d\mathrm{Ca} \approx 3.7 k_B T$ over a wide range of Ca.

Section IV examines how the free energy of catenation can play a role in driving disentanglement of the catenated DNAs which are an intermediate product of DNA replication. A simple kinetic model is introduced to study relaxation of catenation by random stand passages, providing a crude model of the effect of type-II topoisomerases [5]. The result is that thermal forces can be expected to suppress catenation to the level of a link per roughly 100 persistence lengths for

moderate levels of segment excluded volume interaction, if a mechanism to allow strand exchange (topoisomerases) is available. This effect is then used to show that condensation of entangled polymers along their length, so as to thicken and stiffen them, is alone sufficient to drive their mutual and self-disentanglement, in times comparable to those observed *in vivo* for prokaryote and eukaryote cells.

In Sec. V the results of Sec. II are used to make an approximate estimate of the "knotting length," or number of segments $N_0$ of topology-annealed ring polymer required to have an appreciable (near 50%) probability of knotting. For IRWs, $N_0 \approx 300$ [31,32]. The knotting length rapidly increases with excluded volume; for cylindrical segments of diameter-to-length ratio of 0.2, $N_0 \approx 2 \times 10^4$ [32], and for segments which are about as thick as they are long, $N_0 \approx 10^6$ [31]. This dramatic increase in $N_0$ with segment thickness can be understood in terms of the quenching effect of excluded volume interactions on catenation of Sec. II, based on the idea that simple toroidal knots which appear at the onset of random knotting can be thought of as reconnections at near encounters of catenated rings.

## II. MEAN-SQUARE CATENATION OF TWO TETHERED POLYMER RINGS

Consider two flexible polymer rings, each $N$ segments in length, tethered together at one point along their length by a connector of length 0.3 times the polymer segment length [Fig. 1(a)]. This section focuses on calculating the mean-square value of Gauss linking number for freely fluctuating topology, $\langle \mathrm{Ca}^2 \rangle_0$ (the subscript 0 will refer to averages without constraint of topology).

### A. Scaling for ideal random walk: $\langle \mathrm{Ca}^2 \rangle_0 \propto N^{1/2}$

Rings made of segments that have zero thickness, i.e., without excluded volume interactions, have statistics of ideal (non-self- and-mutually-avoiding) random walks (IRW), or at large scales, Gaussian polymers. The average radius of each chain is $R \approx N^{1/2}$ (the RW step length is taken to be unity). To estimate Ca, one must consider the number and signs of crossings of one polymer over the other in a given projection. There are two distinct types of crossings of segments: "distant" crossings where the segments are separated in the projection direction by more than a segment length and "near" crossings where the segments involved are a segment length or less from one another.

The number of near crossings may be estimated using the segment concentration in each coil $c \approx N/R^3 \approx 1/N^{1/2}$, indicating that the number of times two segments are closer than a segment length from one another is $cN \approx N^{1/2}$. These close encounters each contribute $\pm 1$ to Ca, since small rearrangements of the chain geometry which do not cost appreciable entropy can flip the sign contributed by each of them [Fig. 1(b)]. Since all possible signs of these crossings occur with equal probability in the ensemble, the near crossings contribute $\approx N^{1/2}$ to $\langle \mathrm{Ca}^2 \rangle_0$.

There are many more distant crossings. Projection of the polymer into the plane gives $c_{2\mathrm{d}} \approx N/R^2 \approx 1$, and $c_{2\mathrm{d}} N \approx N$ distant crossings, suggesting that they should overwhelm the near-crossing contribution. However, almost all of the contributions from distant crossings exactly cancel out, since they alternate in sign as one chain crosses back and forth over the other.

There will be a slow buildup of contributions from the distant crossings, from distant "wrapping" of one chain around the other, i.e., correlated-sign distant crossings separated by sufficient chain contour for the chains to exchange their order in the projection direction, without near crossings. Near the tether point (monomer number $n \approx 0$), these exchanges occur relatively frequently but as $n$ increases, they become rarer; their rate with $n$ scales $\approx 1/n$. Since these nonlocal wraps occur with random signs, their contribution to $\langle \mathrm{Ca}^2 \rangle_0$ is $\approx \int_1^N dn/n = \ln N$ [30]. Therefore for IRWs the dominant contribution to random catenation is from near crossings, with $\langle \mathrm{Ca}^2 \rangle_0 \propto N^{1/2}$ for large $N$ [33].

### B. Gaussian polymer calculation of $\langle \mathrm{Ca}^2 \rangle_0$

The result $\langle \mathrm{Ca}^2 \rangle_0 \approx N^{1/2}$ was first obtained by Tanaka [7] in a computation where the Gauss invariant was expressed as an interaction of two polymers mediated by a gauge field. The same result may be extracted from Otto's direct computation of $\langle \mathrm{Ca}^2 \rangle_0$ for two circular Gaussian polymers constrained to have two points along their contours separated by displacement $\mathbf{r}$ [14],

$$\begin{aligned}\langle \mathrm{Ca}^2 \rangle_0 = \frac{1}{Z} \int \mathcal{D}\mathbf{r}_1 \int \mathcal{D}\mathbf{r}_2 \delta^3[\mathbf{r}_1(N) - \mathbf{r}_1(0)] \\ \times \delta^3[\mathbf{r}_2(N) - \mathbf{r}_2(0)] \delta^3[\mathbf{r}_1(0) - \mathbf{r}_2(0) - \mathbf{r}] \\ \times \left[ \frac{1}{4\pi} \oint \oint \frac{d\mathbf{r}_1 \times d\mathbf{r}_2 \cdot (\mathbf{r}_1 - \mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|^3} \right]^2 \\ \times \exp\left[ - \sum_{j=1,2} \frac{3}{2} \int_0^N dn_i \left( \frac{d\mathbf{r}_i}{dn_i} \right)^2 \right], \end{aligned} \tag{2}$$

where $Z$ is the partition function of the polymer rings,

$$\begin{aligned} Z = \int \mathcal{D}\mathbf{r}_1 \int \mathcal{D}\mathbf{r}_2 \delta^3[\mathbf{r}_1(N) - \mathbf{r}_1(0)] \delta^3[\mathbf{r}_2(N) - \mathbf{r}_2(0)] \\ \times \delta^3[\mathbf{r}_1(0) - \mathbf{r}_2(0) - \mathbf{r}] \exp\left[ - \sum_{j=1,2} \frac{3}{2} \int_0^N dn_i \left( \frac{d\mathbf{r}_i}{dn_i} \right)^2 \right], \end{aligned} \tag{3}$$

and where the polymer segment length has been taken to be unity. Otto's calculation leads to Eq. 25 of Ref. [14],

$$\langle \mathrm{Ca}^2 \rangle_0 = -8 \int \frac{d^3q}{(2\pi)^3} \int \frac{d^3u}{(2\pi)^3} \frac{\mathbf{q} \cdot (\mathbf{u} - \mathbf{q})}{q^2 |\mathbf{u} - \mathbf{q}|^2} \frac{e^{i\mathbf{u}\cdot\mathbf{r}}}{u^4} (1 - e^{-Nu^2/6})^2, \tag{4}$$

where terms generating subleading contributions in $N$ have been dropped.

Being based on a continuum model, this result requires a short-distance (large-$|\mathbf{q}|$ and $|\mathbf{u}|$) cutoff. This amounts to putting an upper limit on the magnitudes of the integration variables, $u \leq \Lambda\pi$ and $q \leq \Lambda\pi$. The order-unity constant $\Lambda$ can

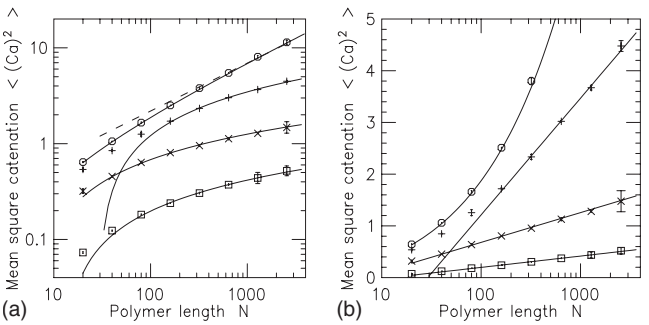


FIG. 2. (a) Monte Carlo (MC) simulation results for mean-square catenation $\langle \mathrm{Ca}^2 \rangle$ of tethered rings compared with analytical and scaling formulas. Circles show results for ideal random walks (IRW, segment diameter $d=0$); for $N>500$ the results converge to $\propto N^{1/2}$ (straight line). The smooth curve through the data points is the result of an exact calculation for the Gaussian polymer model (see text), which passes through all the MC data. Squares indicate self-avoiding walk (SAW) results obtained by increasing the segment diameter to $d=0.2$; the magnitude of $\langle \mathrm{Ca} \rangle$ is strongly suppressed, and follows a logarithmic behavior (smooth curve through the data is $a \ln N+b$, see text). Intermediate values of segment diameter ($d=0.01$, +; $d=0.05$, $\times$) give results where IRW-like ($\approx N^{1/2}$) behavior is seen for small $N$, but which for larger $N$ gives way to a slowdown toward SAW ($\approx \ln N$) behavior. (b) Same data as in (a), plotted on linear-log scale, showing that for nonzero excluded volume ($d=0.01$, +; 0.05, $\times$; 0.2, $\square$) mean-square catenation grows logarithmically with $N$.

be determined by comparison with numerical data for a specific polymer model.

The angular parts of the three-dimensional integrals may be computed exactly

$$\langle \mathrm{Ca}^2 \rangle_0 = \frac{1}{2\pi^4} \int \frac{du}{u^3} (1-e^{-Nu^2/6})^2 \frac{\sin(ur)}{ur} \times \int dqq \left[ \frac{2}{x} + \frac{1-x^2}{x^2} \ln \left| \frac{1-x}{1+x} \right| \right], \tag{5}$$

where $x=q/u$. Then, changing the inner integration variable to $x$ gives

$$\langle \mathrm{Ca}^2 \rangle_0 = \frac{1}{2\pi^4} \int_0^{\Lambda\pi} \frac{du}{u} \frac{\sin(ur)}{ur} (1-e^{-Nu^2/6})^2 \times \int_0^{\Lambda\pi/u} dx \left[ 2 + \frac{1-x^2}{x} \ln \left| \frac{1-x}{1+x} \right| \right]. \tag{6}$$

The lower limits on the integrals may be set to zero as the integrands are finite near $x=0$ and $u=0$. Equation (6) may be simply numerically computed (smooth curve through circles in Fig. 2).

The asymptotic behavior of this integral for large $N$ and $r \approx 1$ can be demonstrated to be $\propto N^{1/2}$ as follows. The only source of divergence is the rescaled cutoff of the inner integral; the integrand goes to a finite limit (the square bracket approaches 4 for $x>1$), so the inner integral contributes $4\pi\Lambda/u+O(1)$, considered as an expansion in $u$. This leads to

$$\langle \mathrm{Ca}^2 \rangle_0 = \frac{1}{2\pi^4} \int_0^{\Lambda\pi} \frac{du}{u} \frac{\sin(ur)}{ur} (1-e^{-Nu^2/6})^2 \left[ \frac{4\pi\Lambda}{u} + O(1) \right]. \tag{7}$$

Defining a rescaled integration variable $x^2=Nu^2/3$,

$$\langle \mathrm{Ca}^2 \rangle_0 = \frac{2\Lambda (N/3)^{1/2}}{\pi^3} \int_0^{\Lambda\pi (N/3)^{1/2}} \frac{dx}{x^2} \frac{\sin px}{px} (1-e^{-x^2/2})^2, \tag{8}$$

where $p=r/(N/3)^{1/2}$ and where $O(1)$ corrections to the leading $N^{1/2}$ scaling are omitted. For $r \approx 1$ and $N \gg 1$ (the case of interest here) the $p \to 0$ limit may be taken, and the integral upper limit may be taken to infinity, to extract the leading $N$-scaling:

$$\langle \mathrm{Ca}^2 \rangle_0 = \frac{2\Lambda}{\sqrt{3}\pi^3} N^{1/2} \int_0^{\infty} \frac{dx}{x^2} (1-e^{-x^2/2})^2 = \frac{2(\sqrt{2}-1)}{\sqrt{3}\pi^{5/2}} \Lambda N^{1/2}. \tag{9}$$

Note that the $O(1)$ numerical prefactor in the final result (multiplying $\Lambda N^{1/2}$) is $0.02734\cdots$, substantially smaller than unity. The original calculation by Tanaka [7] obtains a quite similar numerical constant $[3/(4\pi)^2=0.01899\cdots]$ multiplying a $O(1)$ cutoff constant, suggesting that a numerically small prefactor is intrinsic to this scaling law.

### C. Scaling for self-avoiding random walk: $\langle \mathrm{Ca}^2 \rangle_0 \propto \ln N$

In the presence of segment-segment excluded volume interactions, the near-crossing estimate of Sec. II A breaks down due to correlations between segment positions. Remarkably, for a self-avoiding polymer of $2N$ segments, sites on the two $N$-segment-long halves of the polymer collide only a *finite* number of times, even for $N \to \infty$ [34]. Therefore for two tethered $N$-segment rings near crossings contribute only $O(1)$ to $\langle \mathrm{Ca}^2 \rangle$.

Thus, for SAWs, $\langle \mathrm{Ca}^2 \rangle_0$ is determined by the distant crossings. As in Sec. II A, almost all of the contribution to Ca of distant crossings cancels out, but there is still the $\approx \ln N$ contribution from random wrapping of one polymer around the other (the polymer statistics do not affect the requirement of $n$ segments to form the next wrap a contour distance of $n$ from the tether point [30]). The result for two tethered SAW rings is thus

$$\langle \mathrm{Ca}^2 \rangle_0 \propto \ln N. \tag{10}$$

### D. Numerical calculations of $\langle \mathrm{Ca}^2 \rangle_0$

The results outlined above require validation given that they involve approximations. Given the lack of exact solutions, a numerical approach is indicated. In this subsection the results of Secs. II B and II C are checked using MC calculations.

#### *1. Monte Carlo method*

MC calculations were done for an athermal hard cylinder model, with small fluctuations in segment lengths as de-

scribed below. MC moves were of two types, vertex and reversal moves. Vertex moves were random displacements of randomly chosen vertices along the simulated polymer, where vertex displacement was distributed randomly over a ball of radius 0.35. Vertex moves were rejected if they caused either collisions of segments, or if they caused the length of any segment to be outside of the range 0.9 to 1.1. The connector segment between two (fixed) vertices was treated as a segment whose length was allowed to fluctuate between 0.27 and 0.33.

Reversal moves amount to choosing two random points along either polymer ring and then reversing the order of segments along that region of the ring. Such moves were rejected if they caused segment overlaps.

The vertex moves lead to small-scale randomization of chain configurations and are computationally inexpensive. The reversals, while somewhat more costly computationally, provide rapid equilibration of the polymer configurations on large scales and also provide a mechanism for topology change in the presence of self-avoidance suitable for calculation of equilibrium averages.

Calculations were carried out by alternating between series of vertex and reversal moves. 100 vertex moves per vertex were done, followed by one reversal move. After each reversal move, Gauss invariant was computed using the method discussed in Ref. [15]. Between 8 and 12 long simulations were run for each choice of segment diameter and ring length. Standard errors were calculated from the standard deviation of measured quantities among these separate runs. Between $10^8$ and $10^9$ vertex moves, with between $10^6$ and $10^7$ reversals moves per vertex were used for each $(d,N)$ case studied.

#### 2. Results

Calculations for the mean-square linking number of two tethered rings each of $N$ segments and of zero thickness (i.e., ring IRWs) are shown in Fig. 2, circles. Note that $\langle Ca^2\rangle_0$ substantially exceeds unity only for $N>10^2$. For $N>500$, $\langle Ca^2\rangle_0 \to a_0 N^{1/2}$ with prefactor $a_0=0.23\pm0.01$, with appreciable negative corrections to the asymptotic power law for $N<500$. The dashed line shows a pure $N^{1/2}$ law.

The solid curve is the analytical result for a Gaussian polymer, Eq. (5). Setting the cutoff parameter $\Lambda=8.5$ gives close quantitative agreement with MC data from $N=20$ to $N>2000$ (Fig. 2, upper curve). The value of $\Lambda=8.5$ used to describe the MC data is reasonable since one must have wavelengths somewhat less then the segment length resolved to take account of the connector length (0.3 segments in length) as well as the fluctuations in segment length ($\pm0.1$ segments).

If the polymer segments are made to be cylinders of diameter $d>0$ and are constrained to not overlap one another, the rings become self- and mutually avoiding. As $d$ increases from zero to order unity, segment-segment excluded volume is increased. As shown in Fig. 2, increase in $d$ from zero strongly suppresses $\langle Ca\rangle_0$ for any given ring length $N$. This suppression is not just a change in the prefactor or a small shift in exponent of the power law, but is a qualitative change in the scaling of catenation with ring length.

MC calculations for segments with $d=0.2$ (Fig. 2, squares) are in good accord with the expected logarithmic behavior with $\langle Ca^2\rangle_0 = a \ln N + b$, $a=0.095\pm0.002$, and $b=-0.24\pm0.01$. Note that for the largest rings studied, each $N=2560$ segments, $\langle Ca^2\rangle_0$ only reaches $\approx 0.5$; the logarithmic increase means that $\langle Ca\rangle_0$ reaches unity only for $N\approx 10^6$. Self-avoidance tremendously suppresses random catenation of two tethered rings.

Figure 2 also shows the variation in $\langle Ca^2\rangle_0$ with $N$ for two intermediate values of $d$, 0.01 (+) and 0.05 (×). Even for these rather small degrees of segment avoidance there is strong suppression of linking and a leveling off of $\langle Ca^2\rangle_0$ toward the logarithmic behavior seen for $d=0.2$. The curves drawn through these points in Fig. 2 are linear fits of the same form as those for the $d=0.2$ case. For $d=0.05$(×), $a=0.25\pm0.01$ and $b=-0.47\pm0.03$; for $d=0.01$(+), $a=0.98\pm0.05$ and $b=-3.3\pm0.5$.

While one might be tempted to construct a one-parameter scaling function description of the crossover of $\langle Ca^2\rangle_0$ from the power-law IRW case to the logarithmic SAW case, it should be noted that one-parameter scaling has long been known to be inapplicable even to simple geometrical properties of the self-avoiding walk [35].

## III. FREE ENERGY OF CATENATION

Given the fluctuation $\langle Ca^2\rangle_0$ of Sec. II, one is led to suppose that the free-energy cost of constraining catenation of two tethered ring polymers should display a quadratic Ca dependence near Ca=0 [14]:

$$\beta F(\mathrm{Ca}) = \frac{c}{2}\mathrm{Ca}^2, \tag{11}$$

where $c$ is a dimensionless effective "Gauss invariant rigidity," and where $\beta=(k_BT)^{-1}$. This expression is reminiscent of the small-deformation limit for the twist deformation energy of a uniform elastic filament. Recall that in that case one would expect $c\propto 1/N$ (the standard definition of twist elastic constant for an elastic filament would be $C=(k_BTNb/[4\pi^2])c\propto k_BTb$, indicating the $c\propto 1/N$ scaling [36]).

If free energy (11) held with no anharmonic corrections, then as long as $c\ll 1$, one would have $c=1/\langle Ca^2\rangle_0$. However, this "equipartition" relation depends on the form of the free energy for $\mathrm{Ca}>(\langle Ca^2\rangle_0)^{1/2}$. Furthermore, simply from the quantization of Ca there will be a threshold value of $N$ required for the quadratic behavior (11) to be observable.

### A. Ideal random walk polymers

#### 1. Scaling analysis

For IRWs, $\langle Ca^2\rangle_0 = a_0 N^{1/2}$ for $N\gg 1$. The lack of segment-segment correlations indicates that the probability distribution for Ca must be a function of only $Ca^2/(\langle Ca^2\rangle_0)$; the only scale for $Ca^2$ is its average value. Therefore

$$-\ln\frac{P(\mathrm{Ca})}{P(0)} = \beta F(\mathrm{Ca}) = f(\mathrm{Ca}/\langle \mathrm{Ca}^2\rangle_0^{1/2}), \tag{12}$$

where $f(x)$ is an even function. Note that the free energy of the Ca=0 state is taken to be zero.

For large $N$, $\langle Ca^2\rangle_0 = a_0 N^{1/2}$. Given $f(x) = kx^2/2$ for $x \ll 1$ (sign-reversal symmetry of Ca) one has, for sufficiently large $N$ and sufficiently small Ca,

$$\beta F(\mathrm{Ca}) = \frac{k\mathrm{Ca}^2}{2\langle \mathrm{Ca}^2\rangle_0} = \frac{k\mathrm{Ca}^2}{2a_0 N^{1/2}}, \quad |\mathrm{Ca}| < \langle \mathrm{Ca}^2\rangle_0^{1/2} = a_0^{1/2} N^{1/4}, \tag{13}$$

where the $N$ dependences hold for the $N \to \infty$ limit.

The rigidity of Eq. (11) is $c = k/\langle Ca^2\rangle_0$ or for large $N$, $c = k/(a_0 N^{1/2})$, which is a peculiar nonextensive form. The analogous twist rigidity of an elastic rod of length $\propto N$ would scale as $c \propto 1/N$ [36], giving an extensive free energy. In contrast, the effective rigidity associated with the Gauss invariant decays more slowly: $c \propto 1/N^{1/2}$. Conversion of this to a conventional filament elastic constant yields $C \propto Nc \propto N^{1/2}$, divergent with $N$.

One can estimate how large $N$ is required for the quadratic variation of free energy (11) to be observed. Given the $a_0 \approx 0.23$ determined in Sec. II, quadratic dependence of the free energy on Ca can be expected for $|\mathrm{Ca}| < 0.48 N^{1/4}$. Since Ca is quantized, one requires $0.48 N^{1/4} > 2$ (giving at least three values of Ca to fit a quadratic dependence to). Therefore quite long polymers, with $N > 300$, are required for the quadratic behavior (11) to be observable.

In the strong-catenation regime $|\mathrm{Ca}| > \langle \mathrm{Ca}^2\rangle_0^{1/2} \approx N^{1/4}$, one may estimate the form of the free energy by supposing that the polymers organize into correlation regions, each containing $n$ segments. With each correlation region there is associated free energy $\approx k_B T$ and a contribution to catenation $c \approx n^{1/4}$. Adding up the contributions from each correlation region (biased to have the same sign contribution to Ca) gives $|\mathrm{Ca}| = Nc/n = N/n^{3/4}$, giving $n = |N/\mathrm{Ca}|^{4/3}$, and

$$\beta F(\mathrm{Ca}) = k' \left| \frac{\mathrm{Ca}}{N^{1/4}} \right|^{4/3}, \quad |\mathrm{Ca}| > \langle \mathrm{Ca}^2\rangle_0^{1/2} = a_0^{1/2} N^{1/4} \tag{14}$$

for some $O(1)$ constant $k'$. This result is in accord with the scaling hypothesis (12). When $\mathrm{Ca} \approx N$, $\beta F \approx N$, since the forcing of the two chains to tightly associate quenches $O(N)$ degrees of freedom.

The structure of the $|\mathrm{Ca}| > N^{1/4}$ coil is a random walk of correlation regions each of size $\xi \approx n^{1/2}$, giving an overall coil size for the polymers of $R \approx (N/n)^{1/2} \xi \approx N^{1/2}$. The topological constraint $|\mathrm{Ca}| > N^{1/4}$ does not change the coil size at the level of scaling, i.e., by more than an $O(1)$ factor.

Nonlocal contributions to Ca from the $(N/n)^{1/2}$ collisions between distant correlation regions may be neglected. Each such collision generates $\approx n^{1/2}$ near encounters of segments, giving a total of $N^{1/2}$ nonlocal near encounters of segments along the polymers. These random-sign contributions to $|\mathrm{Ca}|$ are $\approx N^{1/4}$ and therefore may be neglected in the strong catenation regime where $|\mathrm{Ca}| > N^{1/4}$; the two polymers may be considered to be locally entangled with one another along their lengths.

#### 2. Monte Carlo calculation results

Figure 3 shows the free energy of catenation computed from the probability distribution of catenation using the Monte Carlo calculations of Sec. II D. Data are plotted versus catenation in units of the root-mean-squared catenation ($x \equiv \mathrm{Ca}/[\langle \mathrm{Ca}\rangle_0^2]^{1/2}$), the scaling variable of Eq. (12). All the data, for polymers with $N=20$ to $N=2560$, collapse onto one curve, supporting the scaling hypothesis (12). For $x<1$, the data are consistent with $f(x) \approx kx^2/2$ dependence (lower straight line in Fig. 3), with $k = 3.4 \pm 0.3$. For $x>1$, the data are well described by $f(x) \sim x^{4/3}$, in accord with scaling result (14).

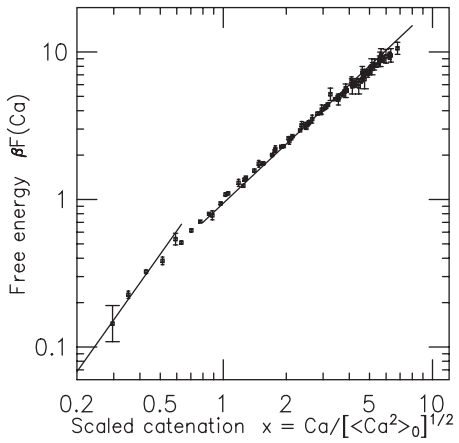


FIG. 3. Catenation free energy for tethered ideal random walks. Data are shown for $N=20$, 40, 80, 160, 320, 640, 1280, and 2560, in terms of the scaled catenation $x = \mathrm{Ca}/\langle \mathrm{Ca}^2\rangle_0^{1/2}$. Data for all lengths collapse onto a single curve; for $x<1$ the free-energy increase is consistent with a $x^2$ law (line, lower left); for $x>1$ the free-energy increase slows down and is consistent with $\approx x^{4/3}$ (line, upper right).

### B. Polymers with excluded volume interactions

If the polymers have excluded volume interactions, there are correlations in positions of segments and in catenation fluctuations along the chain. Fluctuations in Ca are much smaller, with $\langle \mathrm{Ca}^2\rangle_0 \approx a \ln N + b$ growing very slowly with $N$. The initial quadratic dependence of free energy (11) must have a very large effective elastic constant, $c \approx 1/(a \ln N + b)$. This decays with $N$ far more slowly than in the ideal random walk case.

A consequence of this large effective stiffness is that the free energy added per catenation $dF/d\mathrm{Ca}$ approaches $k_B T$ for catenation numbers of $\mathrm{Ca} \approx 1/c \approx \langle \mathrm{Ca}^2\rangle_0 \approx a \ln N + b$. Beyond this value of Ca one cannot expect to observe quadratic low-catenation free-energy behavior. To obtain $\langle \mathrm{Ca}^2\rangle_0 > 1$ for $d=0.2$ one needs $N > e^{(1-b)/a} = e^{1.24/0.095} \approx 4 \times 10^5$. For $d=0.2$ it is therefore unlikely that any conceivable computation (or experiment) could have large enough $N$ for quadratic free-energy dependence on Ca to be observed.

There is certainly no range of Ca for the $d=0.2$ case where quadratic free-energy behavior will be observable in the numerical calculations of this paper. Referring to Fig. 2 it is apparent that for $d=0.2$ catenation does not reach unity for even the largest polymers studied here ($N=2560$, $\langle \mathrm{Ca}^2\rangle_0 \approx 0.6$).

Another way to look at the small-catenation free energy is to note that it has a very large stiffness. If one did have sufficiently long polymers to observe the quadratic free-

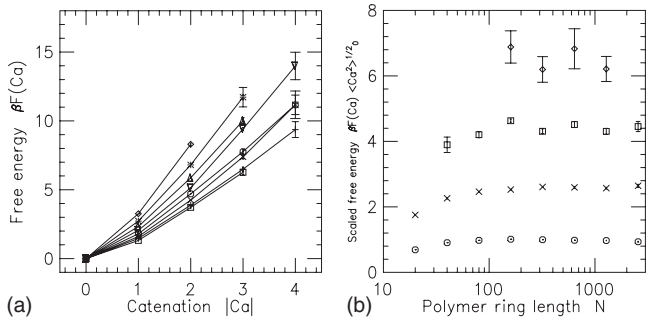


FIG. 4. (a) Catenation free energy versus Ca for tethered polymers with self-avoidance ($d=0.2$) in the small-Ca regime, obtained from unbiased Ca fluctuations ($\tau=0$). Data are shown for $N=20$ ($\diamond$), 40 (*), 80 ($\Delta$), 160 ($\nabla$), 320 ($\bigcirc$), 640 ($\times$), 1280 (+), and 2560 ($\square$). As $N$ increases, the free energies for given Ca go down, but only slowly. (b) Free energies of (a), multiplied by $(\langle \mathrm{Ca}^2\rangle_0)^{1/2} \approx 0.095 \ln N - 0.24$, for $|\mathrm{Ca}|=1$ ($\bigcirc$), 2 ($\times$), 3 ($\square$), and 4 ($\diamond$). For $N>100$, the scaled free energies are nearly constant, indicating that for $N$ up to $\approx 3000$, $\beta F \approx g(\mathrm{Ca})/(\langle \mathrm{Ca}^2\rangle)^{1/2}$, where $g(\mathrm{Ca})$ is an $N$-independent constant.

energy regime, the effective elastic constant would behave as $c \propto 1/(\ln N)$, or in elastic filament terms, a very divergent $C \propto N/\ln N$.

In the numerical calculations of Sec. II, for the case $d=0.2$, the large free-energy cost of catenation resulted in only a few values of Ca being observed, essentially between $\pm 4$. Minus one times the logarithm of the probability of each catenation case is the free energy of catenation, and this is plotted in Fig. 4(a) for $N=20$, 40, 80, 160, 320, 640, 1280, and 2560 (free energies are given relative to the Ca=0 state). As $N$ increases, the free energy comes down only slowly with $N$, e.g., much slower than $1/N$ or even $1/N^{1/2}$.

For large $N$ and small Ca one expects scaling behavior $\beta F(\mathrm{Ca},N) \approx f(\mathrm{Ca}/\langle \mathrm{Ca}^2\rangle_0^{1/2})$ where $f$ is an $N$-independent distribution, for consistency of the catenation probability distribution with its second moment $\langle \mathrm{Ca}^2\rangle_0$.

Figure 4(b) shows how the set of free energies for the case $d=0.2$, when multiplied by $\langle \mathrm{Ca}^2\rangle_0^{1/2} \approx a \ln N + b$ vary as a function of $N$. For large $N$, the free energies scaled in this way are nearly constant. Therefore, for $N>10^2$ and $<10^4$ and $|\mathrm{Ca}| \leq 4$ one has

$$\beta F(\mathrm{Ca}) = \frac{g(|\mathrm{Ca}|)}{\langle \mathrm{Ca}^2\rangle_0^{1/2}}, \tag{15}$$

where $g(|\mathrm{Ca}|)$ is an $N$-independent function of Ca. Figure 4(b) indicates that for $\mathrm{Ca}>0$, $g(|\mathrm{Ca}|) \approx A|\mathrm{Ca}|+B$ where $A=1.8\pm0.2$ and $B \approx -0.8 \pm 0.1$, approximately linear in $|\mathrm{Ca}|$ [note $g(\mathrm{Ca}=0)=0$].

### *Free energy for larger Ca*

Figure 4(b) hints that the free energy will grow nearly linearly for larger Ca; the data for $N \geq 320$ suggest a nearly linear increase in free energy for $\mathrm{Ca}>2$, with $dF/d\mathrm{Ca} \approx 4$. This cannot be convincingly demonstrated by calculations with freely fluctuating topology [Fig. 4(a)], simply because the probabilities of large-Ca states are prohibitively small.

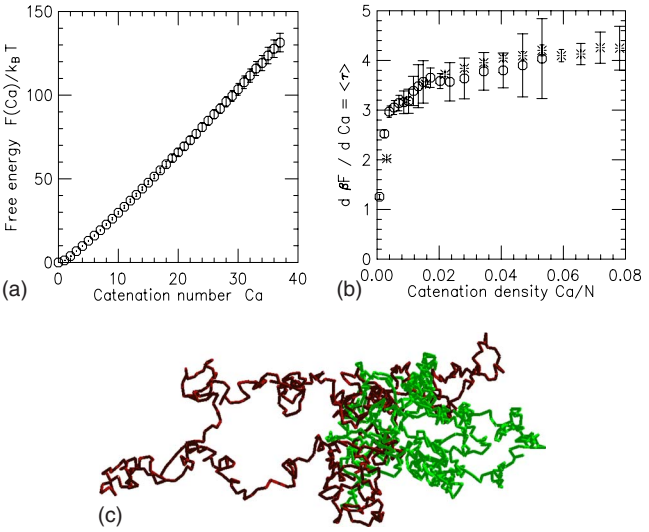


FIG. 5. (Color online) Free energy of catenation over wide range of Ca obtained from biased-Ca computations ($\tau>0$). (a) Free energy obtained from computations with $\tau=1$, 2, 3, 3.5, 3.6, and 3.75, for $N=640$. After an initial sublinear increase from zero, $F(\mathrm{Ca})$ grows nearly linearly with Ca. (b) Free-energy change per added catenation ($dF/d\mathrm{Ca}$) versus catenation density ($\mathrm{Ca}/N$). Results for $N=160$ (*) and $N=640$ ($\bigcirc$) rapidly jump to near $3.7k_BT$ per added catenation. The lack of $N$ dependence indicates that the free energy is dominated by local effects, i.e., by formation of a tightly entangled state. (c) Sample configuration of $N=640$ chains with $\tau=3.5$, showing separated disentangled and entangled domains.

In order to study the behavior of the free energy for larger Ca, MC calculations were done with the term $\beta \Delta E = -\tau \mathrm{Ca}$ added to the microscopic energy. The field $\tau$ biases the ensemble to have nonzero average Ca and can be thought of as a kind of torque, but driving Gauss invariant rather than simple toroidal catenation (if one were concerned with thinking about the work done per radian of Gauss invariant the analogous "torque" one should use is $k_BT\tau/[2\pi]$). By carrying out a series of calculations at different values of Ca one can obtain the relative free energies over a few different ranges of Ca (note $\beta F(\mathrm{Ca}') - \beta F(\mathrm{Ca}) = \ln[P(\mathrm{Ca})/P(\mathrm{Ca}')] + \tau[\mathrm{Ca}' - \mathrm{Ca}]$, where the $P$ are obtained from calculations with nonzero $\tau$). The results for different $\tau$ can then be combined together to compute $F(\mathrm{Ca})$ over a wide range of Ca.

Figure 5(a) shows results of this type of calculation, for the case $N=640$, chosen to be large enough to be out of the regime where there is strong variation in free energy with $N$ [see Fig. 4(a)] but where the calculations are still computationally tractable. Several calculations of roughly $10^8$ vertex moves and $10^6$ reversal moves were carried out for each of $\tau=1$, 2, 3, 3.5, 3.6, and 3.75. In each run roughly half the total time was used to reach equilibrium, while the remaining time was used to obtain the catenation distributions. Figure 5(a) shows the total free energy as a function of Ca; except for a small "foot" near Ca=0, the free energy is nearly linear in Ca. The results for $\mathrm{Ca} \leq 4$ are in accord with the $\tau=0$ results of Fig. 4.

The linear behavior is strikingly apparent in Fig. 5(b) which shows the "equation of state" $d\beta F/d\mathrm{Ca} = \langle \tau \rangle$ versus

catenation density Ca/$N$. The free energy per successive added catenation for $N=640$(○) rapidly increases and then is nearly constant at approximately $3.7k_BT$ per added catenation. The plateau behavior of $dF/d\mathrm{Ca}$ is nearly independent of $N$; Fig. 5(b) also shows results for $N=160$(*) which are nearly the same as those for $N=640$.

The linear dependence of $F$ on Ca in the regime where $|dF/d\mathrm{Ca}|>k_BT$ suggests that "catenation phase separation" is occurring [30]. In this regime spreading of entanglements along the chains can become thermodynamically less favorable than condensation of catenations together into a localized tangle. Once a localized entangled structure forms, additional catenations will simply add polymer length to it. This can be thought of as phase separation of catenation, or formation of catenation-dense and catenation-free "phases." In the entanglement condensation regime each additional catenation converts an additional bit of polymer into condensate, giving a free energy that grows linearly with Ca [30].

Supporting this condensation picture are results for the $N=640$ calculation with $\tau=4$; this value of Gauss invariant torque generates a large buildup of Ca (data not shown), indicating that $\tau=4$ is above the free energy per catenation associated with the condensate. Figure 5(c) shows a sample configuration for two $N=640$ rings with $\tau=3.5$; regions of densely entangled and relatively free chain can be observed.

This catenation condensation effect is not a true phase transition for any finite $N$, hence the slight slope on the plateaus of Fig. 5(c), with more slope in the smaller-$N$ case. However, in the $N\to\infty$ limit Fig. 5(c) will become a flat plateau with $d\beta F/d\mathrm{Ca}\approx 3.7$. The onset of condensation will occur for $\mathrm{Ca}\approx \ln N$, or for $\mathrm{Ca}/N\approx N^{-1}\ln N\to 0$ as $N\to\infty$ [30]. For $\mathrm{Ca}/N\approx 0.1$ the plateau will end at an abrupt upturn in $d\beta F/d\mathrm{Ca}$, corresponding the formation of the fully catenation-condensed state.

## IV. RELEASE OF CATENATIONS BY RANDOM SEGMENT EXCHANGES

The free-energy cost of catenation of Sec. III generates a thermodynamic driving force that can guide removal of linking between replicated DNA molecules by topoisomerases (enzymes which permit DNA to pass through itself) [37]. In principle, given sufficient time, thermal forces will guide Ca to $\approx(\langle \mathrm{Ca}^2\rangle_0)^{1/2}$, or to $\mathrm{Ca}\approx(a\ln N+b)^{1/2}$ for long and isolated self-avoiding polymers. For naked DNA (segment length 100 nm containing 300 bp, thickness including electrostatic effects is 5 nm, corresponding to $d=0.05$, Fig. 2, ×) $\langle \mathrm{Ca}^2\rangle$ does not pass unity until roughly $N=200$, or 60 kb. Thus even without supercoiling, action of proteins, or confinement effects, $\langle \mathrm{Ca}^2\rangle$ for large DNA plasmids or chromosomal loops will eventually drop to near zero simply via thermodynamic forces, if they are able to reach isolated-coil thermal equilibrium. However, the kinetics of this topological equilibration process is unclear. Below it will be shown that there is an initial rapid catenation removal regime for $|\mathrm{Ca}|>0.03N$ where the dynamics is essentially local ($N$ independent); thermal removal of Ca below this level will proceed more slowly due to strong thermal fluctuation effects.

### A. Kinetics of catenation release

To estimate the time required for random thermalization of catenation topology, numerical simulations were carried out, using the Monte Carlo approach discussed above with two changes. First, only local vertex moves were used, which in the presence of excluded volume interactions preserve polymer topology and provide diffusive, Rouse-type dynamics. Given the step size and unit segment length, the unit of real time corresponding to one move per vertex can be considered to be $\tau=\eta b^3/(k_BT)$, where $b$ is the segment length and $\eta$ is the viscosity. For naked DNA where $b=100$ nm and $\eta=10^{-3}$ Pa sec, $\tau\approx 3\times 10^{-4}$ s. For either bacterial or eukaryote chromatin where $b\approx 60$ nm, one has a somewhat shorter time scale of $\tau\approx 5\times 10^{-5}$ s.

Second, a topoisomerase-like mechanism was introduced to permit strand passage events as follows. When a vertex move was found to lead to overlap of two segments, it was permitted, with probability $p_{\mathrm{overlap}}=0.01$. Thus one in every $1/p_{\mathrm{overlap}}=100$ collisions involving only two segments was allowed to form an overlap. Then, for subsequent vertex moves, the collision was allowed to persist until the two segments involved no longer were overlapped. This procedure permitted random strand passages to occur via a local, topoisomerase-like mechanism, providing a pathway for topology changes, with $\Delta\mathrm{Ca}=\pm 1$. The value of $p_{\mathrm{overlap}}$ can be thought of as corresponding to the presence of one topoisomerase per $1/p_{\mathrm{overlap}}$ DNA segments; *in vivo* there is thought to be roughly one type-II topoisomerase per few kilobases of DNA, comparable to the choice of $p$.

Calculations were carried out where two ring polymers each of $N$ segments with $d=0.2$ were initially catenated together to have $\mathrm{Ca}/N=0.16$, as simple toroidal links (i.e., the two polymers were initially wound around one another $0.16N$ times). This initially entangled state was allowed to thermalize for $\approx 10^7$ steps per vertex, in the absence of topology-changing moves. Then, topology-changing moves were allowed to occur and Ca was recorded as a function of time steps per vertex. Four independent simulation runs were carried out for each of $N=122$, 152, 302, 452, and 602 segments.

Results for the time evolution of Ca averaged over the four runs as a function of time steps per vertex are shown in Fig. 6(a). For the range of lengths studied, there is rapid removal of Ca during the time range $10^3$ and $3\times 10^4$ time steps; beyond $3\times 10^4$ time steps there is a marked slowdown in removal of Ca as thermal fluctuations start to establish the equilibrium topology distribution.

Over the whole time course, different realizations yield similar results. Figure 6(b) shows the time course of the mean-squared catenation $\langle \mathrm{Ca}^2\rangle^{1/2}$ averaged over four runs. The square root of this average is plotted to allow easy comparison of Figs. 6(a) and 6(b); the two plots are nearly the same over the rapid removal time regime (time $<3\times 10^4$) indicating that Ca follows essentially the same time course in the different runs. Differences between the runs only start to be noticeable in the fluctuation regime (time $>3\times 10^4$ steps) where the catenation number is approaching order unity.

Section III showed that for self-avoiding polymers the free energy grows linearly with Ca, indicating the free-

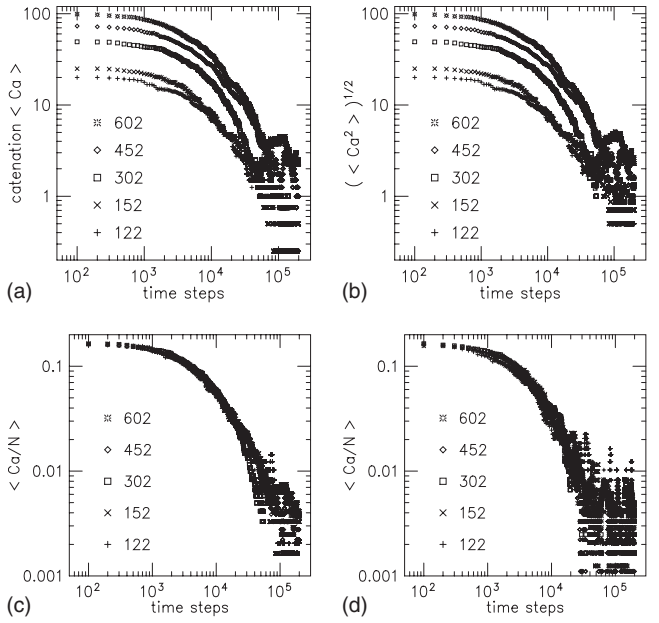


FIG. 6. Time course of removal of catenation in MC simulations with only local vertex diffusion, for various polymer lengths (legend gives $N$ values). In all cases $\mathrm{Ca}/N=0.16$ at $t=0$; at subsequent times Ca can change as a result of intermittently allowed segment overlaps. Each data set shows results averaged over four independent runs. (a) $\langle \mathrm{Ca} \rangle$ as a function of time, for $d=0.2$. Catenations are smoothly removed for $t<3\times10^4$; at later times fluctuations begin to dominate over free-energy-driven catenation reduction. (b) $\langle \mathrm{Ca}^2 \rangle^{1/2}$ as a function of time for the runs of (a). In the smooth relaxation regime the results are nearly the same as in (a) indicating that there is very little variation from run to run. (c) Catenation density $\langle \mathrm{Ca} \rangle$ of (a) and (b) as a function of time. In the smooth removal regime $t<3\times10^4$, the data collapse onto a single curve terminating near $\mathrm{Ca}/N=0.03$. The results indicate that the dynamics in the smooth removal regime are dominated by local relaxation processes. (d) Data for $\langle \mathrm{Ca} \rangle/N$ versus time for $d=0.05$ collapse for times $t<10^4$.

energy density is linear in catenation density. This suggests that at least the rapid-removal regime should be governed by local kinetics, i.e., catenation density should be an $N$-independent function of time. If the data of Fig. 6(a) are plotted in terms of catenation density $\mathrm{Ca}/N$ [Fig. 6(c)], the trajectories do indeed closely overlap during the rapid removal regime (for times up to $3\times10^4$ steps), indicating that the catenation removal dynamics are essentially local. There is a lag phase of roughly $10^3$ time steps due to the low probability of segment exchanges occurring per collision (recall the probability for forming an overlap of $10^{-2}$), and then between times of $5\times10^3$ and $3\times10^4$ catenations disappear at a roughly power law rate $\sim 1/t$. At the end of the rapid removal regime, catenation density is reduced to $\mathrm{Ca}/N \approx 0.03$.

These results show that a total time of $3\times10^4$ time steps per vertex reduces catenation density to near $\mathrm{Ca}/N=0.03$ for polymer rings of up to $N=600$ segments, with moderate excluded volume ($d=0.2$). The time scale of $10^4$ steps is physically plausible, given the Rouse relaxation time of a 100-segment region $\approx 10^4$ time steps.

The behavior during the rapid removal regime is not highly sensitive to the segment diameter to length ratio. Figure 6(d) shows results for $d=0.05$, fourfold thinner segments than those of Fig. 6(c). A fourfold smaller vertex displacement had to be used to eliminate topology changes by those displacements; at the same time the acceptance rate of displacements was higher. However, the overall dynamics during the rapid removal regime are nearly the same as for $d=0.2$, and terminate at roughly the same $\mathrm{Ca}/N \approx 0.03$. This indicates that in the rapid removal regime it is mainly the free energy of wrapping which is driving release of catenation, which is only weakly dependent on the segment thickness in the range $d=0.05$ to $d=0.2$. Once one is in the fluctuation time regime (times $>3\times10^4$), the fluctuations are larger for the smaller $d=0.05$ case than for $d=0.2$, again as expected from the larger equilibrium value of $\langle \mathrm{Ca}^2 \rangle$.

While hydrodynamic interactions have not been used in these dynamic calculations, their effect may well accelerate relaxation of catenation since hydrodynamics leads to nonlocal coupling of polymer modes. In a nuclear or cell interior, it is quite unclear on what length scales hydrodynamics applies, and a local friction model may well be most appropriate. Hydrodynamic effects will not alter the free-energy cost of Ca.

### B. Driving entanglements out of chromosomes by lengthwise condensation

The suppression of $\langle \mathrm{Ca}^2 \rangle_0$ with increasing excluded volume interaction (Fig. 2), plus the efficient removal of catenation down to levels of $\mathrm{Ca}/N=\mathrm{Ca}/N \approx 0.03$ [Fig. 6(c)] observed for $d=0.2$ indicate that an effective mechanism to eliminate links between long DNA molecules in the presence of topoisomerases can be based on lengthwise condensation [38]. Lengthwise condensation of a polymer is quite different from conventional polymer collapse (the coil-globule transition); in the latter, an isolated polymer forms a dense globule via indiscriminate two-body attractions. In contrast, during lengthwise condensation, long-ranged cross-linking is assumed to be absent; instead self-interactions of the polymer are assumed to fold it up locally along its length ("*in cis*") without generating long-range ("*in trans*") cross-links (Fig. 7).

As lengthwise-condensation proceeds, a polymer continues to act as a flexible self-avoiding polymer, but with values of segment length $b'$, segment thickness $d'$, and total contour length $L'=N'b'$ which are different from the values $b$, $d$, and $L=Nb$ characteristic of the original polymer. In general one can expect $d'>d$ and $L'<L$ as a result of lengthwise condensation; in addition one can expect $b' \geq b$, so that $N'=L'/b'$ should be less than the original value of $N=L/b$. A number of mechanisms could independently or in concert generate lengthwise condensation of chromosomes *in vivo*, including folding of DNA by bending and wrapping around proteins [39], short-range self-tethering of chromosomes, or formation of plectonemically supercoiled domains as occurs in many bacteria.

The case of protein-driven lengthwise condensation requires that those proteins act along a single stretch of DNA,

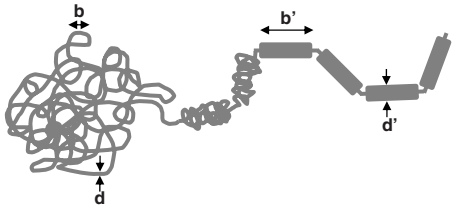


FIG. 7. Lengthwise condensation mechanism. A segment of length $b$ of polymer of thickness $d$ is lengthwise-compacted into a new segment of length $b'$ and thickness $d'$. Successive segments are connected by short flexible linkers of the original polymer. As the segments become longer and thicker, entanglements are driven out of the polymer, both with itself and with other nearby polymers undergoing the same type of condensation.

e.g., by wrapping or looping DNA, or by binding cooperatively along DNA so as to make a thick nucleoprotein filament. A familiar example of this is provided by eukaryote chromosomal DNA, where every $\approx 6$ kb of DNA is compacted into $\approx 30$ nucleosomes, which wrap into a 30-nm-thick and 60-nm-long "segment." This lowest level of eukaryote chromosome compaction accomplishes a roughly 30-fold length compaction from 3 bp/nm for naked DNA to 100 bp/nm for chromatin fiber.

To illustrate how lengthwise-compaction-driven entanglement removal works, consider tethered polymer rings of $N$ segments, each of length $b$ and diameter $d$. Suppose that $d/b=0.2$; the results of Fig. 6 indicate that in the presence of topoisomerases, catenation density will be reduced to $\mathrm{Ca}/N \approx 0.03$ in roughly $3\times 10^4\tau$. Now, suppose that condensation along length gradually occurs via a process which increases $b$ and $d$ both by a factor $\lambda$, so $b'=\lambda b$ and $d'=\lambda d$ (one might imagine the folding of a length $b$ of DNA or chromatin into a nucleoprotein complex of length $b'$, where there are flexible "hinges" between successive complexes, see Fig. 7). Assuming that volume is conserved, the number of segments contained in each polymer will be reduced to $N'=N/\lambda^3$; the total contour length will vary as $L'=N'b'=Nb/\lambda^2=L/\lambda^2$, giving a length compaction factor $L/L'=\lambda^2$.

According to the results shown in Fig. 6, at each stage of this condensation process the linking density will be driven to $\mathrm{Ca}/N' \approx 0.03$; therefore the catenation number will be driven to $\mathrm{Ca} \approx 0.03N'=0.03N/\lambda^3$. Catenation of the initially linked polymers can therefore be eliminated by lengthwise compaction with

$$\lambda_e \approx (0.03N)^{1/3} \tag{16}$$

This factor compacts the original $N$-segment polymer to one with about $N'=N/\lambda_e^3=30$ segments of length $b'=\lambda_e b$ and $d'=\lambda_e d$. The chromosome length compaction factor required to eliminate entanglements is therefore $L/L'=(0.03N)^{2/3}$.

For this mechanism to operate, stable long-ranged "cross-linking" of arbitrary chromatin segments should be delayed until after lengthwise condensation-driven entanglement-resolution occurs. The above calculation may be generalized to the case where the ratio $b'/d'$ changes with condensation, but without changes to the main results since one expects $b/d$ to remain comparable to unity during the time when one can treat a chromosome as a flexible polymer with thermalized topology, and given that the rapid decatenation dynamics are not strongly $b/d$ dependent (Fig. 6).

#### 1. Time scale for entanglement removal by lengthwise condensation

Precise kinetic estimates require detailed knowledge of the time course of $\lambda$, or equivalently geometrical details of folding intermediates. However, given the self-diffusion time scale for a compacted segment of $\tau'=\eta b'^3/k_BT=\lambda^3\tau$, where $\tau$ is the self-diffusion time for the original uncompacted segments, the time required for entanglement removal driven by condensation factor $\lambda_e$ may be roughly estimated to be $T_e=10^4\tau'=10^4\lambda_e^3\tau$.

One might suppose lengthwise condensation by a total factor $\lambda$ to occur *via* a series of intermediate stages, e.g., a series of $n$ steps each of compaction factor $\alpha$, with $\lambda_e=\alpha^n$. Summing over the time required for catenation equilibration at each step yields a geometric series giving essentially the same result for $T_e$, i.e., dominated by its last term.

Given that the lengthwise condensation must have $\lambda \approx (0.03N)^{1/3}$ to ensure complete disentanglement for the case $b'/b=d'/d$, the minimum time required for complete disentanglement of a chromosome region of $N$ statistical segments by lengthwise condensation is

$$T_e = 300N\tau. \tag{17}$$

This time scales slower with $N$ than does the relaxation time of the original polymer which is at first glance paradoxical. However, as lengthwise condensation proceeds internal conformational rearrangement modes are systematically eliminated, drastically reducing the relaxation time of the whole chain.

One may put an upper limit on the forces that must be applied along the polymer to condense it in this minimum time. Taking the worst case where the polymer is initially entirely stretched out, the force needed to translate the entire polymer along its entire length in time $T_e$ is at most $f \approx \eta(Nb)^2/T_e \approx (N/300)(k_BT/b)$ Given than $b\approx 100$ nm for DNA or chromatin, domains of up to $N\approx 10^4$ segments may undergo length condensation in time $T_e$ without exceeding tensions of a few piconewtons (recall $1k_BT/\mathrm{nm}\approx 4$ pN at 300 K). For a more realistic initial condition (e.g., an initial folded or coiled conformation) the forces required can be much lower than this worst-case estimate.

#### 2. Bacterial chromosome segregation

An *E. coli* chromosome consists of DNA coated by DNA-bending proteins including Fis and HU, which make the $4.7\times 10^6$ bp correspond to $N\approx 10^4$ (DNA-bending proteins of this type tend to reduce the segment length from its naked-DNA value of 100 nm containing 300 bp to a segment length $b\approx 60$ nm containing $\approx 500$ bp [40,41]), with $d/b\approx 0.1$. Entanglements between replicated chromosomes (note that before replication, the DNA strands start out with $\mathrm{Ca}\approx 400\,000$) must be removed before segregation can occur [42]. Local compaction by protein-DNA interactions (e.g., by proteins Fis, H-NS, or MukBEF), and by plectonemic super-

coiling generates lengthwise condensation and will promote decatenation by type-II topoisomerases (in *E. coli*, topo IV and DNA gyrase) according to the scheme outlined above.

Plugging $N=10^4$ into Eq. (16) indicates that $\lambda_e \approx 7$ is sufficient to drive complete decatenation of entire chromosomes. This corresponds to an increase in segment length from $b \approx 30$ nm to $b' \approx 200$ nm and a reduction in number of segments to $N' \approx 30$. The (random polymer) size of the bacterial chromosome at the end of this condensation-segregation process will be $\approx b' N'^{1/2} = 1100$ nm, comparable to the size of bacterial chromosomes inside the *E. coli* cell as they spontaneously segregate from one another before the cell itself divides.

The experimentally observed degree of condensation of the *E. coli* chromosome *in vivo* during rapid growth is sufficient to drive most if not all entanglements out of adjacent chromatids. Intriguingly, there have been observations of essentially fully condensed, and in some cases visibly circular or twisted-circular chromosomes in *E. coli*, suggestive of lengthwise condensation [43–45]. The observed organization of the chromosome at large scales has been argued to be consistent with packing of large condensed domains in the *E. coli* cell interior [46]. That latter (theoretical) study found that condensation of the chromosome into 22 large domains drove segregation, while a less complete condensation into 44 domains led to a failure of segregation, consistent with the estimate of this paper of thermal disentanglement requiring condensation to $N'=30$ segments. Finally, mutations or deletion of bacterial DNA-folding or compacting proteins *E. coli* MukB [47] (and its homologue in *B. subtilis*, bsSMC [48]), HU [49,50], H-NS [51], and Fis [52] have been observed to inhibit chromosome condensation or cause anuculate cell (failed chromosome segregation) phenotypes (observation of a high rate of anuculate cells was used to identify the bacterial SMC protein MukB; the name of that protein comes from the Japanese *mukaku*, for "anuculate" [53]).

The time scale for the condensation process for the 4.7 Mbp *E. coli* chromosome is given by Eq. (17) as $T_e = 160$ sec, a little less than 3 min. This is appreciably shorter than the 20-minute minimum time between successive cell divisions of *E. coli*, necessary to allow chromosome segregation to occur by the length-condensation mechanism.

Note that during rapid growth of *E. coli*, the 20-minute cell cycle is about half of the minimum 40 min required to fully replicate the chromosome [43]. Therefore, during rapid growth, one can expect to observe spatially separated subchromosomes in each half of the cell corresponding to partially complete rounds of replication.

The form of Eq. (17) suggests a characteristic "velocity," or polymer length per unit time that can be segregated by random strand passages, $v^* = N/T_e = 1/(300\tau)$ segments per unit time. Considering naked DNA (300 bp per segment, $\tau \approx 3 \times 10^{-4}$ sec) one finds $v^* \approx 3000$ bp/sec. If DNA is replicated more slowly than this, then lengthwise condensation will be able to segregate DNA as replication occurs. On the other hand, if DNA is replicated faster than this velocity, entanglements cannot be removed as replication proceeds, leading to a pileup of tangles which would delay DNA segregation, and possibly stall replication. Since the two diverging replication forks in *E. coli* progress at approximately $10^3$ bp/sec, removal of catenations can keep up with DNA replication, provided that mechanisms are in place to make sure that lengthwise condensation occurs promptly along newly replicated DNA.

The coincidence of replication velocity and segregation velocity suggests that in a rapidly dividing cell, one should be able to observe a gradient of condensation and corresponding disentanglement as one moves away from the replication fork. Recent live-cell experiments observe large-scale separation of labeled chromosome loci (few-hundred nanometer to micron scale) indicating that segregation occurs progressively with replication even during rapid growth [54]. Another study has very recently observed that colocalization of replicated loci along the *E. coli* chromosome separate roughly 15 min following their replication, and that impairment of topoisomerase IV (the type-II topoisomerase responsible for chromosome decatenation in *E. coli*) suppresses this separation, indicating that the colocalization is a result of sister chromatid entanglement (likely dominated by "precatenanes," essentially remnant linking of the mother cell DNA strands) [42]. Measurements of fluctuations of such loci during their separation might provide evidence that termination of locus cohesion is a result of chromosome condensation; in this case, separation might correlate with a reduction in locus position fluctuation.

Other effects including entropic separation of polymers due to confinement [55] may also contribute to chromosome segregation. However, the complete segregation that can be driven by condensation along the length of the chromosome explains why chromosome condensation and segregation proceed in a coupled manner. Direct study of mechanical properties of isolated bacterial chromosomes might be useful in analyzing the mechanism of condensation, but care must be taken to carry out this type of experiment in cytoplasmlike conditions since conventional chromosome isolation (essentially hypotonic bursting of cells) may lead to loss of relatively weak protein-DNA interactions that may be important to maintaining complete chromosome condensation *in vivo*.

Even more tricky to diagnose will be entropic depletion interactions between relatively large (10 nm) nucleoprotein bodies along the chromosome, which will tend to aggregate due to depletion interactions with soluble proteins a few nanometers in diameter [56]. Bursting a cell releases the soluble proteins responsible for maintaining the chromosome as a condensed "phase," so it expands to take on a nonphysiological (noncondensed) conformation [57–59]. Important cases of large complexes present in large numbers along the bacterial chromosome include RNA polymerase (which has been implicated in bacterial chromosome compaction [60,61]), and ribosomes attached to the chromosome through binding to partially transcribed mRNAs. It is important to note that although greatly expanded, extracted bacterial chromosomes do not swell to the size expected for a simple random coil of DNA, indicating that they are appreciably "cross-linked," most plausibly by DNA-binding proteins [58,62].

#### 3. Eukaryote chromosomes

Electron microscopy studies suggest that the next level of chromosome folding above the chromatin fiber is formation

of roughly 100-nm-diameter condensation domains ("chromomeres") [63,64] that organize into a roughly 100-nm-thick fiber which then folds into a chromatid through a series of intermediate folding steps [65]. Sister chromatids do not appear to fully separate until condensation is nearly complete at late prophase [64].

Crude application of the lengthwise condensation model considers a human chromosome of $10^8$ bp organized into "30 nm" chromatin fiber with roughly 6000 bp per statistical segment (100 bp/nm along a $b=60$ nm segment with $d/b \approx 0.5$; note that segment length estimates for chromatin vary quite widely, running up to $b \approx 300$ nm [66]). This gives $N \approx 10^4$, and again a $\lambda_e \approx 7$ necessary to entirely eliminate catenations from sister duplicate chromatids, as well as to segregate different chromosomes from one another. This level of compaction corresponds to a linear density of roughly 5000 bp/nm (about 1600 times that of naked DNA) and an increase in thickness to roughly 200 nm. This level of compaction roughly corresponds to that occurring during chromosome and sister chromatid separation in late prophase of mammalian cells [64] and is about 1/4 to 1/6 of the maximum compaction that occurs at metaphase.

This estimate is encouraging since it provides an explanation of how whole unreplicated chromatids of $\approx 10^8$ bp in length can disentangle from one another while they are lengthwise-condensing in the presence of topoisomerases, as occurs in free-solution experiments with *Xenopus* egg extracts [67]. Despite the lack of any cytoskeletal (mitotic spindlelike) structures, unreplicated chromatids in those experiments are able to start from a dense and entangled initial state, and then simultaneously length-condense while topologically resolving themselves from one another.

However, for eukaryote chromosomes in cells, it should be kept in mind that near the end of DNA replication chromatid pairs are thought to be organized into a series of replicated loop domains joined together along the length of the chromosome, by specialized "cohesin" chromatid-linking complexes [6]. The size of chromosome loop domains in human cells is thought to be $\approx 100$ kb. Each loop thus corresponds to $N \approx 20$ segments of chromatin fiber, so adjacent loops will thermally separate from one another with only minimal length condensation, e.g., by tightening up of nucleosome stacking in chromatin or by formation of chromomeres. The gradual removal of cohesin complexes known to occur in species with larger chromosomes can then allow larger-scale lengthwise-condensation-segregation [84]. In addition to lengthwise condensation, at late stages of chromosome condensation, addition of reversible cross-linking (condensin I) in the presence of topoisomerase II will generate bulk "chromatin gel" elasticity that will drive chromatids apart as discussed in Ref. [37]. It is conceivable that once long-ranged cross-linking starts to occur, self-entanglements within individual segregated chromosomes may be formed by topoisomerases.

## V. "KNOTTING LENGTH" FOR A RING POLYMER

Flexible polymers have a characteristic "knotting length" $N_0$, the number of segments at which the probability of having a knot along the chain becomes appreciable ($\approx 50\%$) in an ensemble of random topologies. This knotting length is more formally defined as the decay constant for the unknot probability for a given type of polymer with $N$, i.e., $P_{\text{unknot}} \propto e^{-N/N_0}$ [31,32,68–71]. The asymptotic exponential decay comes about due to the necessity for small knotted regions to appear along a very long polymer with random topology, as assured by the Kesten "pattern theorem" for random walks [72].

The knotting length is known to depend strongly on segment self-avoidance: ideal random walks have a knotting length $N_0 \approx 250$ [31], while strongly self-avoiding polymers have knotting lengths $\approx 10^6$ [31,32]. Here, this drastic increase on $N_0$ with self-avoidance will be argued to be related to the strong dependence of random linking on segment excluded volume interaction discussed in Sec. II.

If one considers a single circular polymer of $N$ segments, for each conformation there will be a nearest (in 3d space) encounter of some pair of segments of contour spacing $\approx N/2$ [Fig. 8(a)]. The chain can be reconnected at this point to form two subchains of length $N/2$, connected by a short tether, with only local changes to chain conformation, and therefore little free-energy cost. This heuristic mapping of the configurations of a $N$-segment loop to those of a pair of tethered $N/2$-segments loops suggests an approximate recursion relation for the unknotting probability, $P_{\text{unknot}}(N) = [P_{\text{unknot}}(N/2)]^2$, since given a sufficiently close encounter at the reconnection point, then if the original chain was an unknot, the subchains will be unknotted as well.

However, this neglects cases where neither subchain is knotted, but where the two halves of the original chain are "knotted together." Figure 8(b) illustrates such a case for the trefoil knot: neither subchain is knotted after reconnection. The recursion relation for the unknot probability can be modified to account for this effect:

$$P_{\text{unknot}}(N) = [P_{\text{unknot}}(N/2)]^2[1 - P_{\text{halves-knotted}}(N)], \quad (18)$$

where $P_{\text{halves-knotted}}(N)$ is the probability that the two halves of the chain form a knot.

Equation (18) states that an unknotted polymer must not have either subchain knotted, and must also not have the two subchains knotted together, and employs a mean-field-type approximation that subchain knot probabilities of chains of length $N/2$ are equal to the knotting of a whole chain of length $N/2$, and that knot states of subchains are uncorrelated. In addition, this formula ignores the possibility that the second half of the chain "unties" a knot present in the first half of the chain, another mean-field-like neglect of correlations. This latter possibility is likely rather improbable given the use of the closest half-chain encounter for the reconnection.

All these approximations are likely to be most severe for $N \approx N_0$ where subchain knotting is most tenuous and easily undone, and where knotting of chain halves plays a major role. Notably, studies of the size of a single knot in a circular polymer [73–76] indicate that the knotted region is relatively small, scaling as $N_{\text{knot}} \approx N^{0.7}$ for the SAW. This suggests that near the knotting threshold the knotted region will be local-

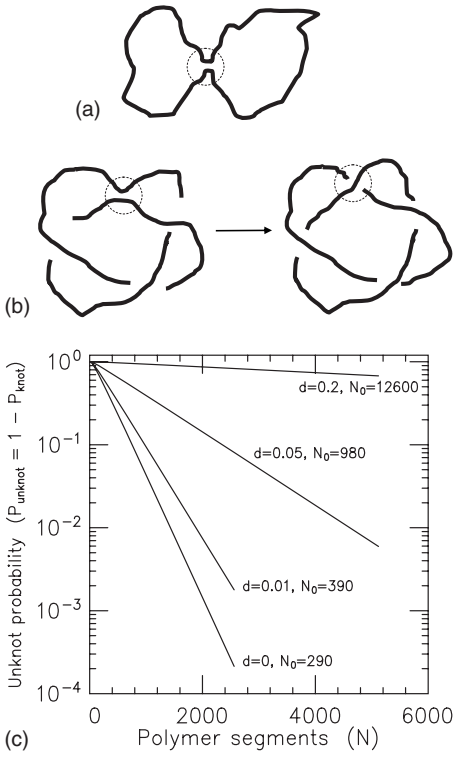


FIG. 8. (a) Close encounter of two segments spaced by roughly $N/2$ along a large circular polymer of $N$ segments allows it to be considered as two tethered rings. (b) A reconnection converting a trefoil knot on a single ring of $N$ segments to a Ca=2 link of two unknotted rings of $N/2$ segments. (c) Results for unknot probability $P_{\text{unknot}}$ as a function of $N$ for polymers with $d=0$ (IRW), $d=0.01$, $d=0.05$, and $d=0.2$. Exponential decays ($P_{\text{unknot}} \propto e^{-N/N_0}$) determine knotting lengths $N_0$.

ized, making the decomposition approach of Eq. (18) plausible.

The knotting probability $P_{\text{knot}}(N)=1-P_{\text{unknot}}$ satisfies a similar recursion equation:

$$P_{\text{knot}}(N) = 2P_{\text{knot}}(N/2) - P_{\text{knot}}(N/2)^2 + [1 - P_{\text{knot}}(N/2)]^2 P_{\text{halves-knotted}}(N). \quad (19)$$

It is to be stressed that these recursion formulas are heuristically derived and are an approximate first step towards understanding topology of a ring polymer in terms of topological properties of reconnected subchains.

Given an estimate for $P_{\text{halves-knotted}}$, Eq. (18) allows $P_{\text{unknot}}(N)$ to be estimated, starting from a small-$N$ cutoff $N_{\text{min}}$ where $P_{\text{unknot}}(N_{\text{min}})=1-P_{\text{halves-knotted}}(N_{\text{min}})$ (i.e., setting $P_{\text{unknot}}(N)=1$ for $N<N_{\text{min}}$). The resulting series for $P_{\text{unknot}}(N)$ approaches a large-$N$ behavior $P_{\text{unknot}} \propto e^{-N/N_0}$, with knotting length given by

$$\frac{1}{N_0} = \sum_{k=0}^{\infty} \frac{-\ln[1 - P_{\text{halves-knotted}}(2^k N_{\text{min}})]}{2^k N_{\text{min}}}. \quad (20)$$

Now $P_{\text{halves-knotted}}(N)$ will be estimated, again using the mapping of a reconnected $N$-segment loop to two tethered $N/2$-segment chains. The occurrence of a knot on the $N$-segment chain requires linking of the $N/2$-segment chains. Subchains with Ca=2 are easily reconnected to yield the simplest (trefoil) knot with its three correlated (same-sign) crossings [Fig. 8(b), right to left]. However, there is a second reconnection of the same close encounter on the Ca=2 link which yields an unknot [the "><" reconnection of the region in the dotted circle of Fig. 8(b), left] Estimating that when $|\text{Ca}| \geq 2$ configurations have their near collisions randomly reconnected one obtains a knot approximately half the time:

$$P_{\text{halves-knotted}}(N) = \frac{\sum_{\text{Ca}=\pm 2,\pm 3\cdots} \frac{1}{2} P(\text{Ca}, N/2)}{\sum_{\text{Ca}=0,\pm 1,\pm 2,\pm 3\cdots} P(\text{Ca}, N/2)}, \quad (21)$$

where $P(\text{Ca},N)$ is the probability of catenation Ca of two tethered polymers of $N$ segments. The possible but less likely reconnection of Ca=1 links to form knots is neglected here. It is to be emphasized that the estimation of $P_{\text{halves-knotted}}$ in terms of catenation of subchains is at best a rough approximation.

Figure 8(c) shows the resulting $P_{\text{unknot}}$ using Eq. (18), with unknot probabilities computed via Eq. (21) using the catenation probabilities (free energies) of Sec. III A cutoff $N_{\text{min}}=40$ is used as the mean-squared catenation of chains below $N=20$ is negligible for all strengths of self-avoidance (Fig. 3). The unknotting probability displays an exponential decay with $N$, $P_{\text{unknot}} \propto e^{-N/N_0}$, thanks to the product form of the solution to Eq. (18). For the IRW ($d=0$) case, $N_0=290$, while for SAWs with $d=0.01$, 0.05 and 0.2, $N_0=390$, 980 and 12600, respectively. Direct simulation calculation of $N_0$ for hard-cylinder polymers [32] gives values of $N_0=270$, 370, 990, and 16 000.

The good accord between the results for numerical studies of random knotting with this crude catenation-reconnection argument provides some insight into the origin of the polymer length scale $N_0$. The large value of the number $N_0$ in the IRW case arises *via* the numerically small prefactor in Eq. (7); the dramatic increase in $N_0$ to a value in the range $10^4$ for that $d=0.2$ SAW arises via the $\ln N$ scaling of $\langle \text{Ca}^2 \rangle_0$.

## VI. CONCLUSIONS

This paper has centered on the question of the statistics of the Gauss linking invariant of two tethered circular polymers. The key result is while $\langle \text{Ca}^2 \rangle_0 \propto N^{1/2}$ for IRWs, it is $\propto \ln N$ for SAWs. Consequences of this for the free energy of polymers as a function of Gauss invariant, entropy-driven separation of polymers, and the knotting length of polymers have been presented.

### A. $N^{1/4}$ scale for IRW catenation

The basic result for the IRW ($\langle \text{Ca} \rangle^2 \propto N^{1/2}$, or more loosely $|\text{Ca}| \approx N^{1/4}$) was first obtained by Tanaka [7] and can be extracted from the essentially exact Gaussian polymer calculations of Ref. [14] (Sec. II B). This result is also consistent with other related calculations involving $\langle \text{Ca}^2 \rangle_0$ for flexible polymers. Brereton and Shah [20,21] used a mean-field ap-

proach to the $n \to 0$-component field theory of two chains to show $\langle Ca^2 \rangle_0 \approx \rho N$ where $\rho$ was the segment concentration contributed by one of the chains in the vicinity of the other. Plugging in the result for the segment concentration $\rho \approx 1/N^{1/2}$ appropriate to describe the interior of an $N$-segment Gaussian polymer yields $\langle Ca^2 \rangle_0 \approx N^{1/2}$.

More detailed calculations aimed at computation of linking in many-ring systems obtained the same general result $\langle \Sigma_i Ca_i^2 \rangle_0 \approx \rho N$, where $Ca_i$ denotes the linking of one chain with the others in a dense melt, using a variety of theoretical methods [8,10–14]. In the dense many-chain context, $\rho \approx 1$, but each chain volume is shared by $\approx N^{1/2}$ interpenetrating chains, leading back to $\langle Ca^2 \rangle_0 \approx N^{1/2}$ for any two of the Gaussian chains in the system which are within a chain radius of one another.

### B. 4/3 law for IRW catenation free energy

In Sec. III A 1 scaling arguments indicated that for $|Ca| > N^{1/4}$, a transition to a locally entangled state should occur, with free energy $F \propto |Ca/N^{1/4}|^{4/3}$. This behavior was observed in numerical simulations. The 4/3 free-energy exponent ought to be exactly calculable since it is a property of the cumulants of the Gauss invariant for the simple Gaussian polymer; the framework of Otto [14] might be useful for this.

### C. SAW catenation free energy

When segment excluded volume is present, catenation fluctuations are greatly suppressed. Exact calculations are out of the question, but a simple scaling argument indicates that $\langle Ca^2 \rangle_0 \approx \ln N$. This growth law matches the similar logarithmic dependence found for a SAW encircling a point two dimensions [77,78], or a straight line in three dimensions [79]. Numerical results of this paper are consistent with this scaling behavior.

The small value of $\langle Ca^2 \rangle_0$ for the SAW leads to a quickly increasing free-energy cost of catenation. This is argued to lead to a catenation condensation effect [30], and a free energy linear in $|Ca|$. It remains to be convincingly determined whether there is clear segregation of catenations from uncatenated polymer, or whether fluctuations lead to alternating catenated and open regions on large scales in the $N \to \infty$ limit. Analogy with the localization of knots on large polymers [73–76,80,81] suggests the former.

The first-order-transition-like behavior observed in this paper is reminiscent of the first-order transition inferred to occur for a single self-avoiding polymer subject to a field coupled to its writhe [82]. Analogous RG calculations could be done for the case of two tethered self-avoiding rings subject to a field coupled to their Gauss invariant; the results of this paper suggest that a finite value of $\tau$ will drive a first-order transition.

### D. Applications to DNA and chromosomes

The results of this paper have application to problems involving segregation of chromosomes or chromosome loop domains. For naked DNA, one has a long and thin statistical segment ($b = 100$ nm, $d/b \approx 0.05$ under physiological solution conditions) that contains 300 bp. When considering bacterial chromosomes, one must note that the DNA is coated by a variety of nonspecifically bound DNA-bending proteins; these likely reduce the effective segment length to $b \approx 60$ nm while increasing the relative thickness to $d/b \approx 0.1$; the resulting statistical segment will contain somewhat more DNA $\approx 500$ bp simply due to the geometry of bending. Thus a $10^6$ bp bacterial chromosome can be considered to be a roughly $10^4$-segment polymer.

For eukaryote chromatin, the compaction by the lowest level of folding is large; formation of closely spaced nucleosomes yields a segment length of $b \approx 60$ nm, but now containing 6000 bp of DNA, and with a relative thickness $d/b = 0.5$. The result is that a $10^8$ bp chromosome should be considered to be a $\approx 10^4$ to $10^5$ segment polymer. According to this paper, the formation of chromatin fiber has the dramatic effect of keeping entanglement complexity in chromosomes of higher eukaryotes to levels comparable to that found in bacterial cells. The large $d/b$ for chromatin fiber further enhances the driving force for entanglement removal.

### E. Dynamics of chromosome segregation

Free energy of catenation was shown to provide a driving force for segregation of chromatin domains; its dependence on polymer segment properties was shown to provide a physical basis for lengthwise condensation to drive segregation of large chromosome segments, even for "blind" topoisomerases which pass whichever DNA segments happen to collide with them. In fact, it has been shown that type-II topoisomerases actually use ATP to preferentially remove entanglements and knots from DNA molecules *in vitro* [83]. This effect should further accelerate the lengthwise-condensation entanglement resolution process.

The segment-folding model analyzed in Sec. IV considered the case where the segment width-to-length $d/b$ ratio was held fixed during the lengthwise condensation process: this constraint was used mainly for convenience, and the same argument can be made for situations where $d/b$ changes during the condensation process. For low-level folding it is possible that $d/b$ may approach 1 (i.e., by formation of strings of beadlike folding intermediates [64]). In this limit excluded volume driven segregation is further enhanced; for $d/b \approx 1$ the knotting threshold for a polymer rises to $N_0 \approx 10^6$. This mechanism can drive the majority of the decatenations out of bacterial or eukaryote chromosomes at small scales early during chromosome segregation, by a rather mild lengthwise condensation. This mechanism is capable of separating replicated chromosome "loop domains" at a rate comparable to that of replication itself. Lengthwise condensation may be accomplished by relatively passive DNA-wrapping, -looping or -folding proteins, or by more active mechanisms such as the SMC protein complexes [84].

A key feature of the length-condensation mechanism that is that it can work perfectly well in open free solution. In eukaryote cells, chromosome segregation is known to be able to proceed rather well in open solution and in the absence of cytoskeletal, mitotic spindle, or other cell-structural components. For example, experiments with cell extracts are able to

take sperm DNA and condense it into chromatin, which then folds into prometaphase-like chromatids which segregate from one another (see Fig. 1 of Ref. [67]).

In the crowded confines of a cell, this mechanism will also work well. If condensation proceeds along length, without generating interchromatid cross-links, then entanglements will gradually be eliminated, first at small scales, and then later at large scales. As condensation proceeds, first sections of chromosomes and then finally chromatids and whole chromosomes will be forced apart. The only limitation on the mechanism is whether when condensation is complete, the chromosomes are compact enough to fit into the cell without running into one another. All cells appear to conform to this constraint during chromosome segregation.

During the later stages of eukaryote chromosome segregation long, relatively stiff condensed chromosomes are formed. In this case, $d/b$ may become very small, since for a uniform elastic medium, one would expect $b'/b \approx (d'/d)^4$ [36]. Whole metaphase eukaryote chromosomes do appear to behave as uniform elastic media insofar as their bending and stretching elasticities satisfy this relation [84,85], with $d/b \approx 10^{-3}$ (segment length of $10^3$ microns) for large animal (newt) chromosomes. Evidence for similar stiffness of chromosomes *in vivo* includes the circular shapes seen for meiotic metaphase chromosomes between crossover points [37] and circular shapes taken by circularized chromosomes [86]. In this situation, disentanglement of adjacent chromatin domains (chromatids) is driven by osmotic repulsion and elasticity of adjacent condensing domains [37], a mechanism distinct from the early-condensation disentanglement discussed in Sec. IV.

In the bacterial chromosome case, microscopy studies suggest large-scale chromosome condensation occurs during rapid cell growth [43,60,61], suggesting a role for a condensation-resolution mechanism as discussed in this paper. Jun and Mulder [55] noted that condensation-based mechanisms may not be necessary for segregation of bacterial chromosomes, given the tendency for confined polymers to separate from one another in a cylindrical pore. In fact, during slow growth of *E. coli* in poor growth media, chromosome segregation involves much less condensation than during rapid growth [61]. It is possible that a greater degree of lengthwise-condensation-resolution occurs during rapid growth in *E. coli* to facilitate rapid chromosome segregation; multiple partial copies of the chromosome are separated during rapid growth [54]. During slow growth slower segregation driven by a lesser degree of chromosome condensation might be sufficient.

### F. Further questions about knotting length

The qualitative change in the scaling of $\langle Ca^2 \rangle_0$ generated by excluded volume interactions suggests a mechanism to explain the drastic increase in the knotting length for polymers with self-avoidance (recall $N_0 \approx 300$ for IRWs, and $\sim 10^6$ for SAWs [31]). A heuristic correspondence between ensembles of ring polymers and tethered rings was used to convert the problem of knotting to one of linking.

A rough and approximate calculation based on the results of Sec. II is encouraging; by summing over occurrences of $|Ca| \geq 2$ links to estimate knotting of halves of one ring, one obtains estimates for $N_0$ in good accord with available numerical data [31,32]. Certainly the estimate of the probability of reconnection of links with $|Ca| > 2$ to knots (50%) made here is crude but this could be studied in much more detail by directly examining linking and knotting statistics of reconnections of knotted chains. The mechanism of Sec. V provides some insight into the large size of $N_0$ and its strong dependence on self-avoidance, in terms of scaling laws and order-unity constants.

## ACKNOWLEDGMENTS

An enlightening conversation with R. Everaers concerning the scaling behavior of randomly linked IRWs is gratefully acknowledged. This work was supported by NSF Grant No. DMR-0715099.

[1] P. G. de Gennes, *Scaling Concepts in Polymer Physics* (Cornell University Press, Ithaca, 1985).
[2] M. Doi and S. F. Edwards, *Theory of Polymer Dynamics* (Oxford University Press, New York, 1987).
[3] M. Delbruck, Proc. Symp. Appl. Math. **14**, 55 (1962).
[4] A. I. Alexandrov, N. R. Cozzarelli, V. F. Holmes, A. B. Khodursky, B. J. Peter, L. Postow, V. Rybenkov, and A. V. Vologodskii, Genetica **106**, 131 (1999).
[5] J. C. Wang, Q. Rev. Biophys. **31**, 107 (1998).
[6] K. Nasmyth and C. H. Haering, Annu. Rev. Biochem. **74**, 595 (2005).
[7] F. Tanaka, Prog. Theor. Phys. **68**, 148 (1982); **68**, 164 (1982).
[8] T. A. Vilgis and M. Otto, Phys. Rev. E **56**, R1314 (1997).
[9] M. Otto and T. A. Vilgis, Phys. Rev. Lett. **80**, 881 (1998).
[10] F. Ferrari, H. Kleinert, and I. Lazzizzera, Phys. Lett. A **276**, 31 (2000).
[11] F. Ferrari, H. Kleinert, and I. Lazzizzera, Eur. Phys. J. B **18**, 645 (2000).
[12] M. Otto, J. Phys. A **34**, 2539 (2001).
[13] F. Ferrari, Ann. Phys. **11**, 255 (2002).
[14] M. Otto, J. Phys. A **37**, 2881 (2004).
[15] K. Klenin and J. Langowski, Biopolymers **54**, 307 (2000).
[16] S. F. Edwards, Proc. Phys. Soc. Lond. **91**, 513 (1967).
[17] S. F. Edwards, J. Phys. A **1**, 15 (1968).
[18] S. F. Edwards and J. W. Kerr, J. Phys. C **5**, 2289 (1972).
[19] M. G. Brereton and S. Shah, J. Phys. A **13**, 2751 (1980).
[20] M. G. Brereton and S. Shah, J. Phys. A **14**, L51 (1981).
[21] M. G. Brereton and S. Shah, J. Phys. A **15**, 985 (1982).
[22] M. G. Brereton and T. A. Vilgis, J. Phys. A **28**, 1149 (1995).
[23] M. G. Brereton, J. Phys. A **34**, 5131 (2001).
[24] On the set of *all* links, Ca is *infinitely* degenerate; for two flexible polymers composed of a finite number of fixed-length segments the set of attainable links is finite making the degeneracy finite but large, and divergent with polymer length.

[25] A. V. Vologodskii, A. V. Lukashin, and M. D. Frank-Kamenetskii, Zh. Eksp. Teor. Fiz. **67**, 1875 (1974).
[26] M. D. Frank-Kamenetskii and A. V. Vologodskii, Usp. Fiziol. Nauk **134**, 641 (1981).
[27] J. P. J. Michels and F. W. Weigel, Proc. R. Soc. London, Ser. A **403**, 269 (1986).
[28] A. V. Vologodskii and N. R. Cozzarelli, J. Mol. Biol. **232**, 1130 (1993).
[29] E. Orlandini, E. J. van Rensburg, M. C. Tesi, and S. G. Whittington, J. Phys. A **27**, 335 (1994).
[30] J. F. Marko, Phys. Rev. E **55**, 1758 (1997).
[31] K. Koniaris and M. Muthukumar, Phys. Rev. Lett. **66**, 2211 (1991).
[32] M. K. Shimamura and T. Deguchi, Phys. Lett. A **274**, 184 (2000).
[33] R. Everarers (private communication).
[34] M. Baiesi, E. Orlandini, and A. L. Stella, Phys. Rev. Lett. **87**, 070602 (2001).
[35] J. J. Prentis, J. Phys. A **17**, 1723 (1984).
[36] L. D. Landau and I. M. Lifshitz, *Theory of Elasticity* (Pergamon, New York, 1986), Sec. II.18.
[37] J. F. Marko and E. D. Siggia, Mol. Biol. Cell **8**, 2217 (1997).
[38] J. F. Marko, in *Multiple Aspects of DNA and RNA from Biophysics to Bioinformatics, Les Houches Session LXXXII* (Elsevier, Amsterdam, 2005), pp. 211–270.
[39] S. Cocco, J. F. Marko, R. Monasson, A. Sarkar, and J. Yan, Eur. Phys. J. E **10**, 249 (2003).
[40] J. Yan and J. F. Marko, Phys. Rev. E **68**, 011905 (2003).
[41] D. Skoko, B. Wong, R. C. Johnson, and J. F. Marko, Biochemistry **43**, 13867 (2004).
[42] X. Wang, R. Reyes-Lamothe, and D. J. Sherratt, Genes Dev. **22**, 2426 (2008).
[43] H. Niki, Y. Yamaichi, and S. Hiraga, Genes Dev. **14**, 212 (2000).
[44] S. B. Zimmerman, J. Struct. Biol. **138**, 199 (2002).
[45] S. B. Zimmerman, J. Struct. Biol. **153**, 160 (2006).
[46] J. Fan, K. Tuncay, and P. J. Ortoleva, Comput. Biol. Chem. **31**, 257 (2007).
[47] H. Niki, A. Jaffé, R. Imamura, T. Ogura, and S. Hiraga, EMBO J. **10**, 183 (1991).
[48] A. Volkov, J. Mascarenhas, C. Andrei-Selmer, H. D. Urich, and P. L. Graumann, Mol. Cell. Biol. **23**, 5638 (2003).
[49] M. Wada, Y. Kano, T. Ogawa, T. Okazaki, and F. Imamoto, J. Mol. Biol. **204**, 581 (1988).
[50] A. Jaffé, D. Vinella, and R. D'Ari, J. Bacteriol. **179**, 3494 (1997).
[51] A. Kaidow, M. Wachi, J. Nakamura, J. Magae, and K. Nagai, J. Bacteriol. **177**, 3589 (1995).
[52] M. Filutowicz, W. Ross, J. Wild, and R. Gourse, J. Bacteriol. **174**, 398 (1992).
[53] S. Hiraga, H. Niki, T. Ogura, C. Ichinose, H. Mori, B. Ezaki, and A. Jaffé, J. Bacteriol. **171**, 1496 (1989).
[54] H. J. Nielsen, B. Youngren, F. G. Hansen, and S. Austin, J. Bacteriol. **189**, 8660 (2007).
[55] S. Jun and B. Mulder, Proc. Natl. Acad. Sci. U.S.A. **103**, 12388 (2006).
[56] D. Marenduzzo, C. Micheletti, and P. R. Cook, Biophys. J. **90**, 3712 (2006).
[57] S. Cunha, T. Odijk, E. Suleymanoglu, and C. L. Woldringh, Biochimie **83**, 149 (2001).
[58] S. Cunha, C. L. Woldringh, and T. Odijk, J. Struct. Biol. **136**, 53 (2001).
[59] L. D. Murphy and S. B. Zimmerman, J. Struct. Biol. **133**, 75 (2001).
[60] J. E. Cabrera and D. J. Jin, Mol. Microbiol. **50**, 1493 (2003).
[61] D. J. Jin and J. E. Cabrera, J. Struct. Biol. **156**, 284 (2006).
[62] S. Cunha, C. L. Woldringh, and T. Odijk, J. Struct. Biol. **150**, 226 (2005).
[63] O. V. Zatsepina, V. Yu. Polyakov, and Yu. S. Chentsov, Chromosoma **88**, 91 (1983).
[64] A. T. Sumner, Chromosoma **100**, 410 (1991).
[65] A. Belmont and K. Bruce, J. Cell Biol. **127**, 287 (1994).
[66] K. Bystricky, P. Heun, L. Gehlen, J. Langowski, and S. M. Gasser, Proc. Natl. Acad. Sci. U.S.A. **101**, 16495 (2004).
[67] A.-E. de la Barre, M. Robert-Nicoud, and S. Dimitrov, *Chromatin Protocols* (Humana, Totowa, NJ, 1999), pp. 219–229.
[68] D. W. Sumners and S. G. Whittington, J. Phys. A **21**, 1689 (1988).
[69] N. Pippenger, Discrete Appl. Math. **25**, 273 (1989).
[70] Y. Diao, N. Pippenger, and D. Sumners, J. Knot Theory Ramif. **3**, 419 (1994).
[71] Y. Diao, J. Knot Theory Ramif. **4**, 189 (1995).
[72] H. Kesten, J. Math. Phys. **4**, 960 (1963).
[73] E. Orlandini, M. C. Tesi, E. J. Janse van Rensburg, and S. G. Whittington, J. Phys. A **31**, 5953 (1998).
[74] V. Katritch, W. K. Olson, A. Vologodskii, J. Dubochet, and A. Stasiak, Phys. Rev. E **61**, 5545 (2000).
[75] B. Marcone, E. Orlandini, A. L. Stella, and F. Zonta, J. Phys. **38**, L15 (2005).
[76] B. Marcone, E. Orlandini, A. L. Stella, and F. Zonta, Phys. Rev. E **75**, 041105 (2007).
[77] M. E. Fisher, V. Privman, and S. Redner, J. Phys. A **17**, L569 (1984).
[78] B. Duplantier and H. Saleur, Phys. Rev. Lett. **60**, 2343 (1988).
[79] J. Rudnick and Y. Hu, Phys. Rev. Lett. **60**, 712 (1988).
[80] E. J. Janse van Rensburg and S. G. Whittington, J. Phys. **24**, 3935 (1991).
[81] A. Y. Grosberg, A. Feigel, and Y. Rabin, Phys. Rev. E **54**, 6618 (1996).
[82] J. D. Moroz and R. D. Kamien, Nucl. Phys. B **506**, 695 (1997).
[83] V. Rybenkov, C. Ullsperger, A. V. Vologodskii, and N. R. Cozzarelli, Science **277**, 690 (1997).
[84] J. F. Marko, Chromosome Res. **16**, 469 (2008).
[85] M. G. Poirier, S. Eroglu, and J. F. Marko, Mol. Biol. Cell **13**, 2170 (2002).
[86] A. Bajer, Chromosoma **14**, 18 (1963).