

Random knotting in very long off-lattice self-avoiding polygons

Jason Cantarella

Mathematics Department, University of Georgia, Athens, GA, USA

Tetsuo Deguchi

Department of Physics, Ochanomizu University, Bunkyo-ku, Tokyo, Japan

Henrik Schumacher

Institut für Mathematik, RWTH Aachen University, Aachen, Germany

Clayton Shonkwiler

Department of Mathematics, Colorado State University, Fort Collins, CO, USA

Erica Uehara

Graduate School of Informatics, Kyoto University, Sakyo-ku, Kyoto, Japan

(Dated: January 7, 2026)

We present experimental results on knotting in off-lattice self-avoiding polygons in the bead-chain model. Using Clisby’s tree data structure and the scale-free pivot algorithm, for each k between 10 and 27 we generated 2^{43-k} polygons of size $n = 2^k$. Using a new knot diagram simplification and invariant-free knot classification code, we were able to determine the precise knot type of each polygon. The results show that the number of prime summands of knot type K in a random n -gon is very well described by a Poisson distribution. We estimate the characteristic length of knotting as $656\,500 \pm 2500$. We use the count of summands for large n to measure knotting rates and amplitude ratios of knot probabilities more accurately than previous experiments. Our calculations agree quite well with previous on-lattice computations, and support both knot localization and the knot entropy conjecture.

I. INTRODUCTION

Self-avoiding polygons (SAPs) with knotted topologies have attracted much interest in various branches of polymer physics, chemistry, and biology because they provide a model for knotted ring polymers in nature and also in synthetic chemistry. For example, these polymers have been found in living organisms such as the circular DNA of *E. coli* [1–4].

As suggested by Edwards [5], it is nontrivial to study the statistical properties of a ring polymer in solution or melt. Such a polymer does not change its topology under thermal fluctuations since it cannot pass through itself. This topological constraint leads to a large reduction in the available volume or degrees of freedom in the configuration space of the ring polymer and hence affects the statistical properties of the ring polymer(s). For example, knotted ring polymer melts have a different crystallization behavior than unknotted ring polymer melts [6]. The reticulated structure of synthesized polymer networks may have self-entanglements, and it is suggested that they should enhance the elastic modulus [7].

In the 1960s, Frisch, Wasserman and Delbrück [8, 9] addressed the conjecture that nontrivial knotted ring polymers should be very common in large ring polymers. Knotting probability has been evaluated in numerical simulations by making use of knot invariants [10–19] and measured directly in biological experiments of DNA [2, 3]. Furthermore, it was rigorously shown that a self-avoiding

walk is exponentially likely to be knotted if it is very long [20, 21]. In all of these contexts, the flexibility and thickness of the polymer plays an important role.

The asymptotic knotting behavior of long and thick self-avoiding polygons (SAPs) with fixed topology is still not fully understood. Knots are generally considered localized and asymptotically independent events on a long-enough polymer [22–25], and there is strong numerical evidence for this idea [26], at least for relatively short polygons with a relatively small number of summands. In this paper, we consider the random variable

$$m_K^n = \text{number of prime summands of knot type } K \text{ in an } n\text{-gon.}$$

If prime summands are independent and localized, then m_K^n should be approximately Poisson-distributed; that is, we should have

$$P(m_K^n = m) \approx \frac{(\lambda_K(n))^m e^{-\lambda_K(n)}}{m!}, \quad (1)$$

where $\lambda_K(n)$ is the expected value of m_K^n [27]. This is an old observation in the random knotting community, probably dating back to the 1960's [28].

Assume that we are in a range of n , m , and K where (1) is a good approximation. Let $R_K(n) = \lambda_K(n)/n$ be the rate of production of summands of knot type K (per edge) in an n -gon. It would follow that the probability of finding a polygon with no prime summands of type K would be well-approximated by

$$P(m_K^n = 0) \approx \frac{(R_K(n)n)^0 e^{-R_K(n)n}}{0!} = e^{-R_K(n)n}.$$

Assuming that m_K^n and $m_{K'}^n$ are approximately independent when K and K' are distinct prime knot types, this immediately implies that the probability of an unknot

$$\begin{aligned} P_{0_1}(n) &\approx \prod_{K \text{ a prime knot type}} P(m_K^n = 0) \\ &= \prod_{K \text{ a prime knot type}} e^{-R_K(n)n} \\ &= e^{-(\sum_K R_K(n))n}, \end{aligned} \quad (2)$$

while the probability $P_K(n)$ that an n -gon has (prime) knot type K is

$$\begin{aligned} P_K(n) &= P(m_K^n = 1) \prod_{K' \neq K} P(m_{K'}^n = 0) \\ &\approx R_K(n) n e^{-R_K(n)n} e^{-\sum_{K' \neq K} R_{K'}(n)n} \\ &= R_K(n) n e^{-(\sum_{K'} R_{K'}(n))n} \\ &= R_K(n) n P_{0_1}(n). \end{aligned} \quad (3)$$

More generally, the *knot entropy conjecture* [29, 30] (cf. [31, 32]) says that for any (prime or composite) knot type \bar{K} the probability $P_{\bar{K}}(n)$ is well-approximated by

$$P_{\bar{K}}(n) \approx C_{\bar{K}} n^{m(\bar{K})} P_{0_1}(n) \left(1 + \frac{\beta_{\bar{K}}}{n^\Delta} + \frac{\gamma_{\bar{K}}}{n} \right), \quad (4)$$

where $m(\bar{K})$ is the number of prime factors of \bar{K} , we say $m(0_1) = 0$, and

$$P_{0_1}(n) \approx A_{0_1} n^{\alpha_{0_1}} \mu_{0_1}^n. \quad (5)$$

Here, the parameters $\beta_{\bar{K}}$ and $\gamma_{\bar{K}}$ describe the finite-size correction. The parameter α_{0_1} (the unknot critical exponent) and the dimensionless ratios $C_{\bar{K}}/C_{\bar{K}'}$ should be universal, while the other parameters $C_{\bar{K}}$, Δ , A_{0_1} , and μ_{0_1} are model-dependent.

The characteristic length of the unknot is given by $N_{0_1} := -1/\log \mu_{0_1}$ and the unknotting probability is approximately $\exp(-n/N_{0_1})$. If the number of segments n is much larger than N_{0_1} , this implies that the polygon is very likely to be knotted.

Comparing (2) and (5), we see that $A_{0_1} = 1$ and $\alpha_{0_1} = 0$ in this model, while

$$\mu_{0_1} = e^{-(\sum_K R_K(n))}. \quad (6)$$

Since $m_K^n = 1$ for a prime knot, combining (3), (4), and (5) yields

$$R_K(n) \approx C_K \left(1 + \frac{\beta_K}{n^\Delta} + \frac{\gamma_K}{n} \right). \quad (7)$$

It is difficult to test (4) directly because $P_K(n) \rightarrow 0$ exponentially fast as $n \rightarrow \infty$ for any prime knot K , making it difficult to gather data for large n (see [33, 34]). This has made it very hard to compute the values of the C_K accurately. However, we can compute the C_K and therefore estimate $P_K(n)$ by gathering data on $R_K(n)$. Further, as noted by Janse van Rensburg [35],

$$\frac{C_K}{C_{K'}} = \lim_{n \rightarrow \infty} \frac{P_K(n)}{P_{K'}(n)} = \lim_{n \rightarrow \infty} \frac{R_K(n)}{R_{K'}(n)} = \lim_{n \rightarrow \infty} \frac{\lambda_K(n)}{\lambda_{K'}(n)}.$$

That is, the amplitude ratio is the ratio of the mean number of prime summands of knot type K to the mean number of prime summands of knot type K' in a random self-avoiding n -gon. While the C_K values are model-dependent, we expect that this ratio should be universal (see Table I).

In this paper, we gather experimental data to test the approximations above in the off-lattice ‘‘string of pearls’’ model [36] for self-avoiding polygons. We focus on the medium-to-large n regime, using new computational tools which enable us to generate extremely large polygons and classify their knot types exactly.

We perform four experimental tests. First, we compute the empirical distribution of m_K^n for knot types through 6 crossings and $n = 2^k$ with $k \in \{10, 11, \dots, 27\}$. For 2^k -gons, we sampled at least 2^{43-k} polygons, so the total number of edges sampled was at least $2^{43} \approx 8.8 \times 10^{12}$ for each k . We check that the empirical distribution is very closely approximated by a Poisson distribution with the corresponding empirical mean $\lambda_K(n)$. These empirical means become quite large and are hence easy to sample accurately for large n . For example, $\lambda_{3_1}(2^{27}) \approx 193.82 \pm 0.09$ (the uncertainty is standard error).

Second, we compute knotting rates $R_K(n)$ for the same K and n and find that (7) explains the finite-size effects very well.

Third, we use these knotting rates to estimate the knotting probability for the unknot and compare the results with counts of unknots in our data. We see a good match to the ‘‘pure exponential’’ model given by (5), and we are able to estimate the characteristic length of knotting ($656\,500 \pm 2500$).

Last, we return to (4), and compute knotting probabilities $P_K(n)$ for simple knots, finding that (3) fits well and yields almost exactly the same value of C_K (and comparable values of β_K and γ_K) for trefoils. The remaining knot types show significant differences in all three parameters. We take this as an indication of the difficulty in computing C_K from $P_K(n)$ data.

Overall, our results are fully consistent with the hypothesis that the Poisson model (1) is a useful description of the knotting of random self-avoiding polygons. They support knot localization (and hence the knot entropy conjecture) in a range of n much larger than previous experiments have been able to test. Further, the standard model of the finite-size correction (3) is also supported by our data.

II. METHODOLOGY

To generate polygons, we used a modification of the classical polygonal fold method [37]. In each trial step, a vertex v_i is chosen uniformly on the self-avoiding polygon. For an n -gon, we then sample x uniformly on $[1, \log_2(n)]$, and round 2^x to the nearest integer k , as suggested by the “scale-free sampling” procedure of Clisby [38]. The “mobile arc” of the polygon is then defined to be the edges between vertices v_i and $v_{i\pm k}$ (with equal probability). The proposed move is uniformly sampled from the rotations around the line through v_i and $v_{i\pm k}$ and the reflections over planes containing v_i and $v_{i\pm k}$. In every case, a proposed step is accepted if the new configuration is still self-avoiding. We measured acceptance probability $\approx 0.63 n^{-0.057}$. This algorithm has detailed balance, so it samples uniformly. We have not proved that every bead-chain SAP is accessible from our starting configuration (the regular n -gon), though we conjecture that this is true.

We burned in each run for $20n$ attempted steps and then took samples at intervals of $n/30$ attempted steps. We used 64 parallel Markov chains with these parameters for all n except 2^{27} , where we took 128 parallel chains. We wrote a new, highly-performant implementation of Clisby’s tree [39, 40] (see also [41–43]), a data structure which provides $O(\log n)$ folds (including collision checking) while retaining $O(n)$ computation of all vertex positions. Our implementation requires considerable memory for very large polygons (≈ 30 gb for 2^{27} -gons), so we ran 2^{24} – 2^{27} -gons on the high-memory nodes of System C (Cinnamon) at Kyoto University and the remaining calculations on the Sapelo2 cluster at the Georgia Advanced Computing Resource Center (GACRC) at the University of Georgia. Our code `polyfold` is available on GitHub [44].

The self-avoiding polygons we generated had a large number of crossings in their planar projections. On average, projections of 2^{27} -gons had $\approx 2^{25}$ crossings. We constructed planar diagram codes for these knots with verified accuracy using extremely careful floating point computation for segment intersections (see [45, 46]). To identify the resulting knots we wrote a new code called `Knoodle` [44, 47] for very efficient combinatorial simplification of knot diagrams. The key idea is that an arc of a diagram that only crosses over other strands may be rerouted as desired to reduce the number of crossings in the diagram (this is called a “pass move” in knot theory, and we call this process “pass reduction”). `SnapPy` will perform similar simplifications, but requires (at least) $O(n)$ memory and time to find each move, so the entire simplification is $O(n^2)$. Our code finds simplifications in essentially constant time and memory, making the entire process roughly linear in time (see [47]). In practice, it is several orders of magnitude faster than `SnapPy`.

Further, portions of a knot diagram connected by only two arcs to the rest may be identified as connect summands and isolated immediately for further simplification. This process almost always results in a topologically minimal crossing diagram. When it does not, we use graph embedding techniques to construct a new lattice curve with the same knot type as the diagram, and rotate it to a different view, then resume the pass reduction process (again, see [47] for more detail). `Knoodle` is available on GitHub [44].

In total, we identified $\approx 2.5 \times 10^8$ prime summands in our $\approx 1.72 \times 10^{10}$ polygons. All but 17 had ≤ 13 crossings after our simplification process. We were able to classify these very rapidly without using knot invariants by matching them to a precomputed list of diagrams (see [48]). The remaining 17 knots were all hyperbolic, and we determined their knot types by recognizing their knot complements using `SnapPy`.

Although we were able to distinguish symmetry types of knots (such as 3_1 and its mirror image, 3_1^m) in our identifications of knot types, we grouped them together in our analysis to match the existing literature (for example writing $3_1/3_1^m$ to indicate all trefoils). With this convention, we had enough observations of the 7 nontrivial prime knot types with 6 or fewer crossings to perform statistical analyses. We made sporadic observations of *much* more complicated knots; for instance, the most complicated summands we observed were the 16-crossing knots $16a_{61059}$ and $16n_{521027}$.

These knot types are clearly much rarer than the ≤ 6 crossing knots, but our data on them is too sparse to quantify this precisely.

Both the raw data and summary statistics are archived on Dryad [49]. The summary statistics are also included as Supplemental Data.

III. RESULTS

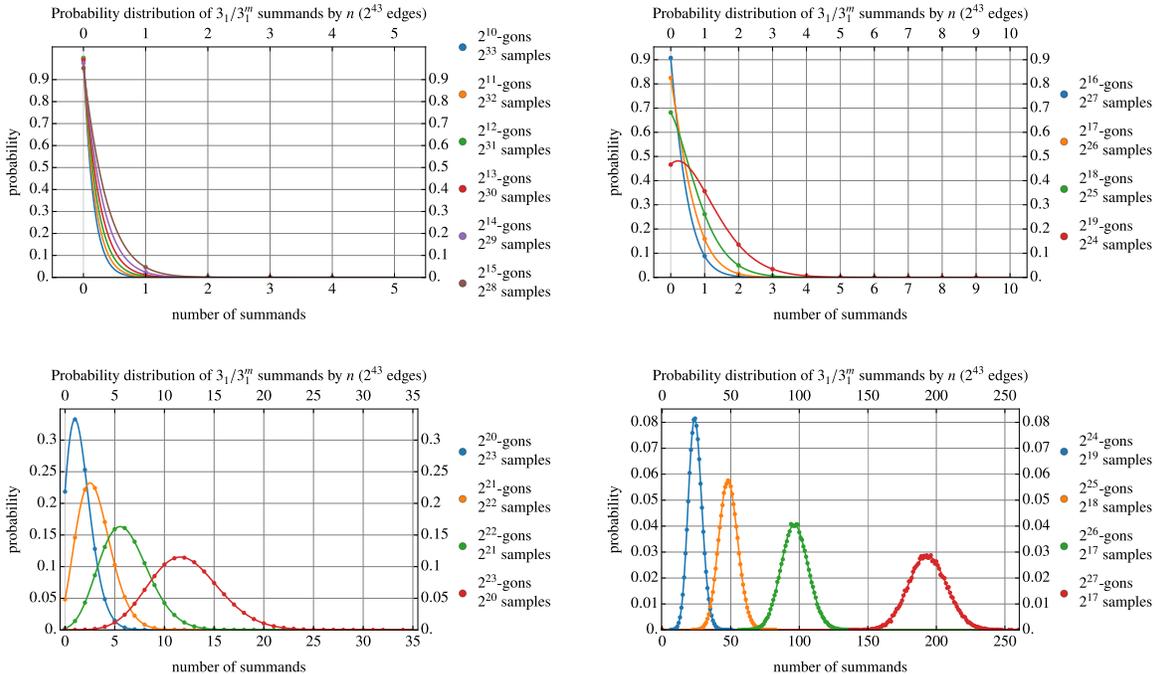


Figure 1. These plots show the probability $P(m_{3_1}(n) = m)$ of observing m trefoil summands in off-lattice self-avoiding polygons of length n , together with the Poisson distribution function $\frac{(\lambda_K(n))^m e^{-\lambda_K(n)}}{m!}$ from (1). This is defined only at integer m , but we draw connecting curves as guides for the eye. The number of n -gons sampled ($2^{43}/n$) decreases with n , so the data appears rougher for large n .

We restricted our analysis to K and n where we observed at least 10 summands of type K in each of the parallel Markov chains. This left us with the 7 knot types with ≤ 6 crossings. We used Geyer’s IPS estimate [50] for the sample mean and standard error of mean for m_K^n for each of the ≥ 64 Markov chains, then combined these estimates across parallel chains (assuming independence). The result was an empirical value for $\lambda_K(n)$ with an error estimate. According to these estimates, our maximum relative (standard) error in $\lambda_K(n)$ was less than 3%.

With these values of $\lambda_K(n)$, we were able to compute the Poisson distributions on the right-hand side of (1). Figure 1 and Figure 2 show the empirical distribution of m_K^n for trefoils, figure-8 knots, 5_1 , 5_2 , and 6_1 for various n , together with the corresponding Poisson distributions.

We computed the total variation distance (TV) [51] between the empirical distribution of m_K^n and the Poisson model. The results appear in Figure 3; they show that this distance is less than 0.01 (and often much smaller) except for a few values of $m_{3_1/3_1^m}^n$ and $m_{4_1}^n$ where n is very large. In these cases, TV was as large as 0.03. Our estimates for $\lambda_K(n)$ with estimated errors and the total variation distances are given in Supplemental Data and in the Dryad dataset [49].

At first, the rise in TV with n might seem counterintuitive; the Poisson approximation should get better in the asymptotic limit. We think this an artifact of our experimental design, where

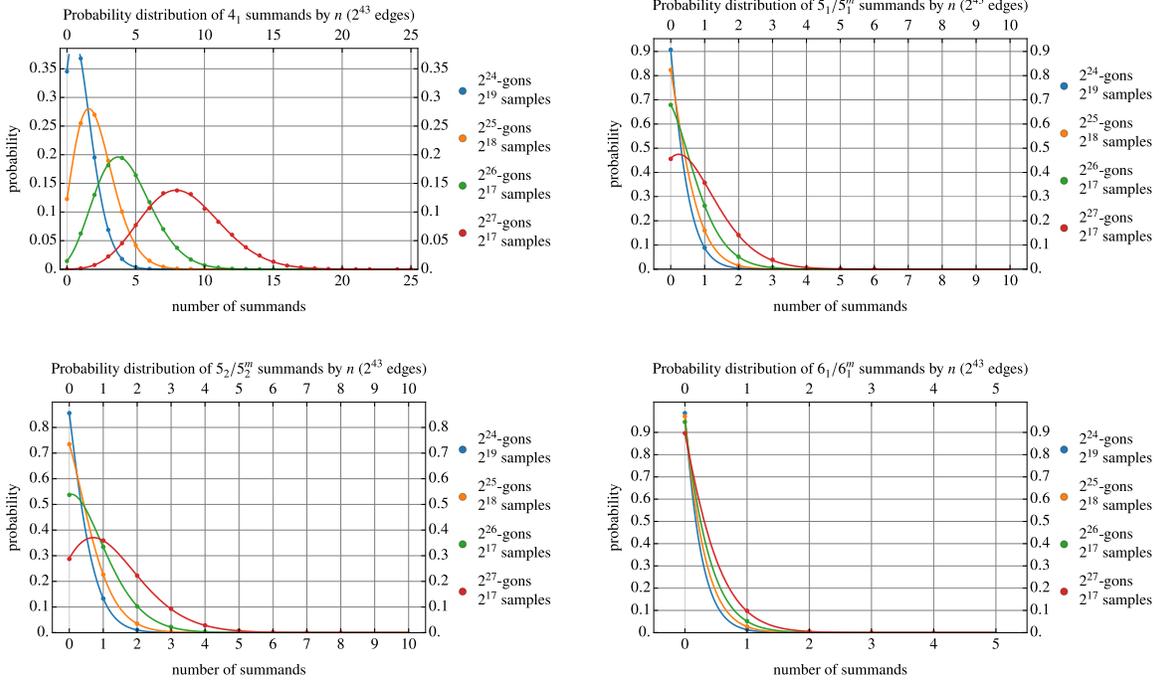


Figure 2. These plots show the probability $P(m_{4_1}(n) = m)$, $P(m_{5_1}(n) = m)$, $P(m_{5_2}(n) = m)$, and $P(m_{6_1}(n) = m)$ of observing m summands of knot type 4_1 , $5_1/5_1^m$, $5_2/5_2^m$, or $6_1/6_1^m$ in off-lattice self-avoiding polygons of length n , together with the Poisson distribution functions $\frac{(\lambda_K(n))^m e^{-\lambda_K(n)}}{m!}$ from (1). Since these knots are much less probable than trefoils, only large values of n are shown. The Poisson fit remains very good.

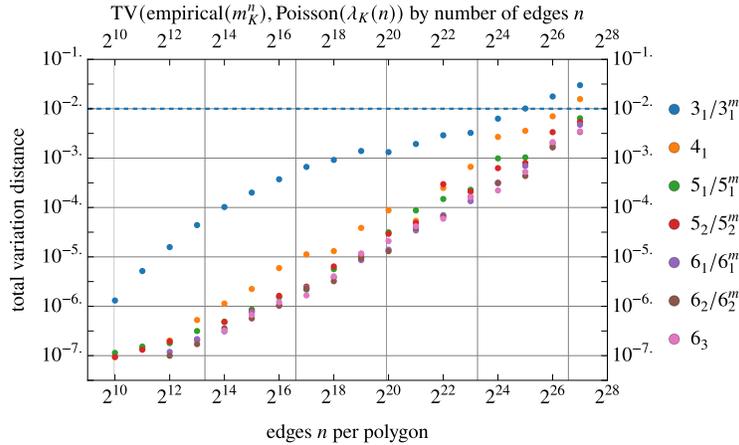


Figure 3. This figure shows the total variation distance $\text{TV}(p_1, p_2) = \sum_m |p_1(m) - p_2(m)|$ between the empirical distribution for the number of prime summands of knot type K in an off-lattice self-avoiding n -gon m_K^n and the corresponding Poisson model with mean $\lambda_K(n)$, using the estimate for $\lambda_K(n)$ from our dataset. We can see that almost all of these distances are very small, confirming that (1) is an accurate approximation for these K and n .

the number of polygons sampled decreases with n . To check this, we performed Pearson's χ^2 test to check the hypothesis that the large- n data for $m_{3_1/3_1^m}^n$ and $m_{4_1}^n$ was sampled from the Poisson distribution. As expected, for $K = 3_1/3_1^m$ and $2^{23} \leq n \leq 2^{27}$ and $K = 4_1$ and $2^{18} \leq n \leq 2^{27}$

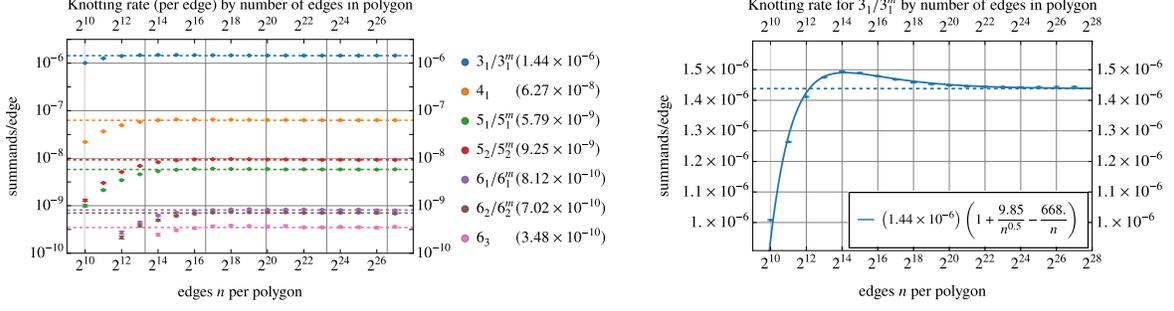


Figure 4. The log-log plot at left shows the rate of knotting per edge: $R_K(n) = \lambda_K(n)/n$ estimated by our experiment. 99% confidence intervals are included, but are too small to be visible on the plot. The fact that these values are becoming constant in n supports our conclusion that the asymptotic knotting rate $C_K = \lim_{n \rightarrow \infty} \lambda_K(n)/n$ exists. Our estimates for C_K appear with their knot types. The semilog plot at right shows this probability for trefoils only, along with our proposed fit to the finite-size correction (7).

(covering all of the cases where $TV \geq 0.01$), this test could not reject the hypothesis at $p = 0.05$.

We then computed the summand production rate (per edge): $R_K(n) = \lambda_K(n)/n$ across our dataset. Again, we had enough data to compute these rates for knots with six and fewer crossings. Figure 4 shows the result; all of the knotting rates have a finite-size correction, which differs by knot type. Further, (7) with $\Delta = 0.5$ explains the finite-size corrections reasonably well. We notice that the asymptotic rate of production decreases very quickly as the knots become more complex, by more than an order of magnitude for each crossing number. This means that in this model the vast majority of knotting consists of trefoil and figure-8 summands, with all other knot types vanishingly rare in comparison. Baiesi, Orlandini, and Stella [52] and Rechnitzer and Janse van Rensburg [33] estimated the asymptotic ratios C_K/C'_K by computing $P_K(n)$, finding that these ratios did not depend on the lattice. We can make corresponding estimates for our off-lattice model, computing C_K from $R_K(n)$; the results appear in Table I.

-/-	4_1			$5_1/5_1^m$			$5_2/5_2^m$			$6_1/6_1^m$		$6_2/6_2^m$		6_3	
	(new)	BOS	JvRR	(new)	BOS	JvRR	(new)	BOS	JvRR	(new)	BOS	(new)	BOS	(new)	BOS
$3_1/3_1^m$	22.9	22.12	28	248	197	400	156	140	280	1770	2633	2050	1668	4140	2677
4_1	-	-	-	10.8	8.89	15	6.78	6.33	9	77.2	119	89.4	75.4	180	121
$5_1/5_1^m$	-	-	-	-	-	-	0.626	0.712	0.67	7.13	13.39	8.25	8.48	16.7	13.61
$5_2/5_2^m$	-	-	-	-	-	-	-	-	-	11.4	18.08	13.2	11.9	26.6	19.12
$6_1/6_1^m$	-	-	-	-	-	-	-	-	-	-	-	1.16	0.63	2.34	1.01
$6_2/6_2^m$	-	-	-	-	-	-	-	-	-	-	-	-	-	2.02	1.60

Table I. This table gives the amplitude ratios $C_K/C_{K'}$, where K is constant on rows, and K' is constant on columns. Our data appears in the columns marked (new), the columns marked BOS contain simple cubic lattice data for n up to 200 000 from Baiesi, Orlandini, and Stella [52]. The columns marked JvRR contain data for n up to 512 on the SCC, FCC, and BCC lattices from Janse van Rensburg and Rechnitzer [33]. The fact that our data is comparable to these lattice studies supports the idea that the bead-chain and lattice models all belong to the same universality class.

We can directly measure knotting rates $R_K(n)$ for only a few knots. According to (5),

$$\sum_K R_K(n) = -\frac{1}{n} \log P_{0_1}(n)$$

where the sum is over all prime knot types K . Since we can measure $P_{0_1}(n)$ for small n by counting unknotted polygons in our sample, we can use this to get a rough estimate of the production rate of

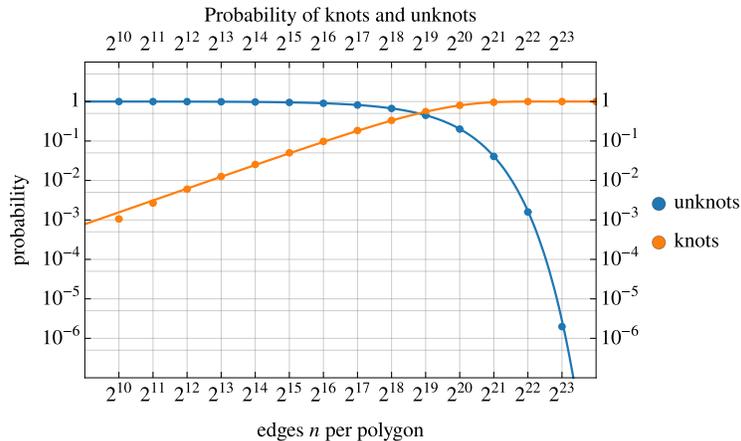


Figure 5. Fraction of knots and unknots for n -edge SAPs, together with plots of e^{-n/n_0} and $1 - e^{-n/n_0}$ with the characteristic length $n_0 = 656\,500$ estimated from knotting rates. The plot stops at $n = 2^{23}$ because we observed no unknotted SAPs for larger n .

more complex knots

$$R_{\text{complex}}(n) := \sum_{K: \text{cr}(K) > 6} R_K(n) = -\frac{1}{n} \log P_{0_1} - \sum_{K: \text{cr}(K) \leq 6} R_K(n)$$

Using this method, a conservative estimate is that for $2^{10} \leq n \leq 2^{21}$, we have $R_{\text{complex}}(n) < 10^{-8}$. Assuming that this bound holds for larger n , we can estimate the characteristic length $n_0 = 1/\sum_K C_K$ by $n_0 \approx 656\,500 \pm 2500$. Figure 5 shows that for large n , the approximation $P_{0_1}(n) \approx e^{-n/n_0}$ provides a quite useful estimate of the probability of unknotting and the probability of knotting. We note that [14] estimates $n_0 \approx 800\,000$ for this model based on simulations with ≤ 1000 beads.

We can now go back to the original knot entropy conjecture, and compare the observed probabilities $P_K(n)$ for various prime knot types with the predictions from (4). Our goal is to check that the prediction of knot probabilities is good, and that the fitted values for C_K , β_K and γ_K are comparable to those we obtained by $R_K(n)$ by fitting (7). Figure 6 shows the result for the trefoil knot. Here, where we have many observations of knots with knot type $3_1/3_1^m$, the results are all quite similar. For the remaining knot types, we have a lot less data for $P_K(n)$ and the results are different; fitting $R_K(n)$ predicts asymptotic values for C_K which are about 10% lower than the values obtained by fitting $P_K(n)$. The complete collection of results is shown in Table II.

IV. CONCLUSION AND FUTURE DIRECTIONS

We have presented new evidence consistent with knot localization and the knot entropy conjecture for very large off-lattice self-avoiding polygons in the “bead-necklace” model. Three things made our experiments possible: switching our focus from computing the probability of individual knot types to counting summands in the prime decomposition, which gave us a new observable to target, an implementation of Clisby’s tree for polygons, which gave us large polygons, and a new knot simplification and classification code, which allowed us to compute our observable.

Previous experiments have been essentially limited by the difficulty of computing knot invariants for large polygons. The limiting factor for our experiment turned out to be the time required to burn in the polygon-generating Markov chain, which consumed the majority of the computation time

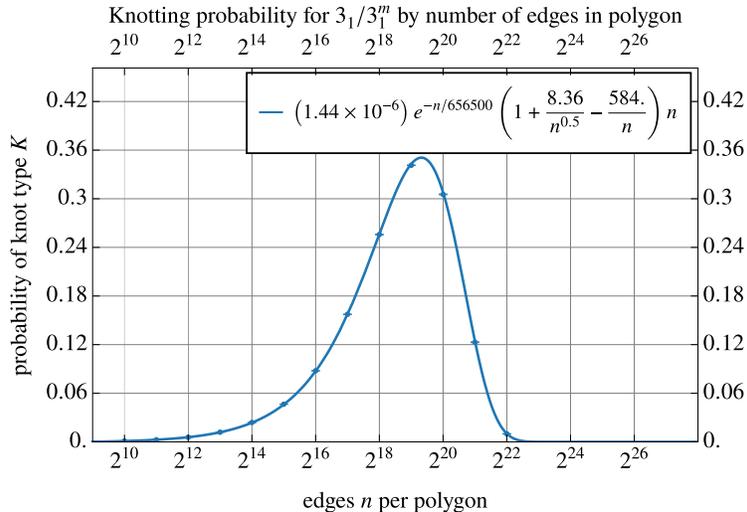


Figure 6. The probability of knot type $K = 3_1$ or $K = 3_1^m$, along with 99% confidence intervals, based on observed counts in our dataset (the errors were estimated with Geyer’s IPS estimate for each Markov chain, then combined across parallel chains assuming independence). The data fits well to the standard model (3). The coefficient C_K and the β_K and γ_K for the finite-size correction are comparable to those measured for $R_K(n)$ shown in Figure 4.

	C_K	β_K	γ_K	adjusted R^2
$3_1/3_1^m$	1.44×10^{-6} (1.44×10^{-6})	9.85 (8.36)	-668. (-584.)	1. (0.99991)
4_1	6.27×10^{-8} (6.71×10^{-8})	19.4 (-8.94)	-2390. (-424.)	0.99999 (0.99837)
$5_1/5_1^m$	5.79×10^{-9} (6.53×10^{-9})	17.1 (-29.8)	-3320. (67.8)	0.99992 (0.99651)
$5_2/5_2^m$	9.25×10^{-9} (1.06×10^{-8})	17.9 (-35.3)	-3780. (219.)	0.9999 (0.99662)
$6_1/6_1^m$	8.12×10^{-10} (9.04×10^{-10})	28.7 (-37.1)	-9370. (-520.)	0.99981 (0.99857)
$6_2/6_2^m$	7.02×10^{-10} (8.16×10^{-10})	45.5 (-48.1)	-12800. (44.8)	0.9998 (0.99862)
6_3	3.48×10^{-10} (3.89×10^{-10})	86.2 (-14.4)	-22600. (-4150.)	0.99962 (0.99915)

Table II. This table shows our final values for the amplitude C_K and the finite-size correction coefficients β_K and γ_K computed by fitting the observed rate of knotting $R_K(n)$ from our dataset to (7). The last column shows the adjusted R^2 value for the fit to $R_K(n)$ rate as a function of n . These fits are extremely good. In parentheses, we see the corresponding values predicted by fitting (3) to our much more limited data on $P_K(n)$, with the last column showing the adjusted R^2 for the fit to $P_K(n)$. For the trefoil, these agree quite well. In the remaining cases, we think that the $R_K(n)$ predictions are more accurate, as we have data for much larger n in these cases.

for 2^{26} - and 2^{27} -gons (cf. Section 4 of [39]). We don’t think that the number of steps for burn-in can be reduced from the current $20n$. So it would be very interesting to find a better strategy for generating starting polygons.

The difference between the amplitudes C_K computed by fitting R_K and P_K for non-trefoil knots is also an interesting issue. We think the difference is mostly a matter of data quality; a much larger experiment with many more polygons might observe enough “pure” figure-8 knots, for instance, to make the C_K figures agree. However, it is also possible that the finite-size effect model needs to be refined. Again, we leave this to future work.

Whittington suggests [28] that the knotting rate $R_K(n)$ should continue to increase with n , perhaps very slowly, as larger and larger knots of type K occur. This is a very interesting idea. We do see a very slight rise in $R_{3_1/3_1^m}$ for $n = 2^{26}$ and $n = 2^{27}$, though our error bars do not allow us to resolve this unambiguously. Ultimately, this question will have to be resolved by a theorem, or by

much larger experiments than we do here.

Our knot simplification and identification toolchain [44] works well for much larger polygons, and is independent of the model used to generate the polygons. We therefore hope that other authors will attempt similar on-lattice experiments, in particular to try to refine the estimates of amplitude ratios on the lattice. It would also be interesting to apply our toolchain to non-self-avoiding polygons, and we intend to do so. Preliminary tests indicate that these have many simple summands, but that they also seem to have one or two prime factors with much more complicated knotting.

Of course, now that we can generate very large example polygons containing many trefoil summands, it would be interesting to see if they can be geometrically isolated and their sizes measured directly. We leave this challenge to future work, as it is nontrivial, but it is encouraging that the Poisson model in (1) is quite compatible with the knot localization conjecture that most prime factors should be tight and isolated from each other.

Finally, note that the “sum rules” of [17, 53] for $P_{K\#K'}(n)$ follow as immediate consequences of the Poisson model in (1). They are also verified by the empirical measurements of $P_{K\#K'}$ in our dataset.

ACKNOWLEDGMENTS

We are very grateful to our colleagues for many helpful discussions, especially Nick Beaton, Mark Dennis, Shura Grosberg, Esias Janse van Rensburg, Enzo Orlandini, Andrew Rechnitzer, Chris Soteros, and Stuart Whittington. Several important improvements in this paper came from BIRS workshop 25w5490, *Polymer Modeling and DNA Topology: The Interplay between Theory, Computation and Experiments* and some crucial work was done at BIRS during 25rit038, *Stick Numbers and Polygonal Knot Theory*. We appreciate the support of the Japan Science and Technology Agency (CREST Grant Number JPMJCR19T4) and the National Science Foundation (DMS–2107700). H.S. was supported by the Research Training Group “Energy, Entropy, and Dissipative Dynamics (EDDy)”, funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), project no. 320021702/GRK2326.

-
- [1] M. A. Krasnow, A. Stasiak, S. J. Spengler, F. Dean, T. Koller, and N. R. Cozzarelli, Determination of the absolute handedness of knots and catenanes of DNA, *Nature* **304**, 559 (1983).
 - [2] V. V. Rybenkov, N. R. Cozzarelli, and A. V. Vologodskii, Probability of DNA knotting and the effective diameter of the DNA double helix, *Proceedings of the National Academy of Sciences* **90**, 5307 (1993).
 - [3] S. Y. Shaw and J. C. Wang, Knotting of a DNA chain during ring closure, *Science* **260**, 533 (1993).
 - [4] R. K. Sharma, I. Agrawal, L. Dai, P. S. Doyle, and S. Garaj, Complex DNA knots detected with a nanopore sensor, *Nature Communications* **10**, 4473 (2019).
 - [5] S. F. Edwards, Statistical mechanics with topological constraints: II, *Journal of Physics A: General Physics* **1**, 15 (1968).
 - [6] K. Hagita, T. Murashima, N. Sakata, K. Shimokawa, T. Deguchi, E. Uehara, and S. Fujiwara, Molecular dynamics of topological barriers on the crystallization behavior of ring polyethylene melts with trefoil knots, *Macromolecules* **56**, 15 (2023).
 - [7] W. W. Graessley and D. S. Pearson, Stress–strain behavior in polymer networks containing nonlocalized junctions, *The Journal of Chemical Physics* **66**, 3363 (1977).
 - [8] H. L. Frisch and E. Wasserman, Chemical topology, *Journal of the American Chemical Society* **83**, 3789 (1961).
 - [9] M. Delbrück, Knotting problems in biology, in *Mathematical Problems in the Biological Sciences*, Proceedings of Symposia in Applied Mathematics No. 14, edited by R. E. Bellman (American Mathematical Society, 1962) pp. 55–63.

- [10] A. V. Vologodskii, A. V. Lukashin, M. D. Kamenetskii, and V. V. Anshelevich, The knot problem in statistical mechanics of polymer chains, *Soviet Journal of Experimental and Theoretical Physics* **39**, 1059 (1974).
- [11] J. des Cloizeaux and M. Mehta, Topological constraints on polymer rings and critical indices, *Journal de Physique* **40**, 665 (1979).
- [12] M. L. Bret, Monte carlo computation of the supercoiling energy, the sedimentation constant, and the radius of gyration of unknotted and knotted circular dna, *Biopolymers* **19**, 619 (1980).
- [13] J. P. J. Michels and F. W. Wiegel, Probability of knots in a polymer ring, *Phys. Lett. A* **90**, 381 (1982).
- [14] K. Koniaris and M. Muthukumar, Knottedness in ring polymers, *Physical Review Letters* **66**, 2211 (1991).
- [15] T. Deguchi and K. Tsurusaki, A statistical study of random knotting using the Vassiliev invariants, *Journal of Knot Theory and Its Ramifications* **3**, 321 (1994).
- [16] T. Deguchi and K. Tsurusaki, Universality of random knotting, *Physical Review E* **55**, 6245 (1997).
- [17] E. Uehara and T. Deguchi, Knotting probability of self-avoiding polygons under a topological constraint, *Journal of Chemical Physics* **147**, 094901 (2017).
- [18] A. Xiong, A. J. Taylor, M. R. Dennis, and S. G. Whittington, Knot probabilities in equilateral random polygons, *Journal of Physics A: Mathematical and Theoretical* **54**, 405001 (2021).
- [19] J. Cantarella, H. Schumacher, and C. Shonkwiler, A faster direct sampling algorithm for equilateral closed polygons and the probability of knotting, *Journal of Physics A: Mathematical and Theoretical* **57**, 285205 (2024).
- [20] D. W. Sumners and S. G. Whittington, Knots in self-avoiding walks, *Journal of Physics A: Mathematical and General* **21**, 1689 (1988).
- [21] N. Pippenger, Knots in random walks, *Discrete Applied Mathematics* **25**, 273 (1989).
- [22] V. Katrich, W. K. Olson, A. Vologodskii, J. Dubochet, and A. Stasiak, Tightness of random knotting, *Physical Review E* **61**, 5545 (2000).
- [23] B. Marcone, E. Orlandini, A. L. Stella, and F. Zonta, What is the length of a knot in a polymer?, *Journal of Physics A: Mathematical and General* **38**, L15 (2005).
- [24] B. Marcone, E. Orlandini, A. L. Stella, and F. Zonta, Size of knots in ring polymers, *Physical Review E* **75**, 041105 (2007).
- [25] E. Orlandini, A. L. Stella, and C. Vanderzande, The size of knots in polymers, *Physical Biology* **6**, 025012 (2009).
- [26] E. Orlandini, E. J. Janse Van Rensburg, M. C. Tesi, and S. G. Whittington, Entropic Exponents of Knotted Lattice Polygons, in *Topology and Geometry in Polymer Science*, The IMA Volumes in Mathematics and Its Applications No. 103, edited by Whittington, S. G., Sumners, D. W., and Lodge, T. (Springer, New York, 1998) pp. 9–21.
- [27] For very large m , we know that $P(m_K^n = m) = 0$, while $\frac{(\lambda_K(n))^m e^{-\lambda_K(n)}}{m!} > 0$, so the two sides of (1) cannot be exactly equal.
- [28] S. G. Whittington, (2025), personal communication.
- [29] N. R. Beaton, K. Ishihara, M. Atapour, J. W. Eng, M. Vazquez, K. Shimokawa, and C. E. Soteros, A first proof of knot localization for polymers in a nanochannel, *Journal of Physics A: Mathematical and Theoretical* **57**, 38LT01 (2024).
- [30] N. R. Beaton, K. Ishihara, M. Atapour, J. W. Eng, M. Vazquez, K. Shimokawa, and C. E. Soteros, Entanglement statistics of polymers in a lattice tube and unknotting of 4-plats, *Discrete Applied Mathematics* **379**, 242 (2026).
- [31] E. Orlandini, M. C. Tesi, E. J. Janse van Rensburg, and S. G. Whittington, Entropic exponents of lattice polygons with specified knot type, *Journal of Physics A: Mathematical and General* **29**, L299 (1996).
- [32] E. Orlandini, M. C. Tesi, E. J. Janse van Rensburg, and S. G. Whittington, Asymptotics of knotted lattice polygons, *Journal of Physics A: Mathematical and General* **31**, 5953 (1998).
- [33] E. J. Janse Van Rensburg and A. Rechnitzer, On the universality of knot probability ratios, *Journal of Physics A: Mathematical and General* **44**, 162002 (2011).
- [34] E. Uehara and T. Deguchi, Characteristic length of the knotting probability revisited, *Journal of Physics: Condensed Matter* **27**, 354104 (2015).
- [35] E. J. Janse van Rensburg, The probability of knotting in lattice polygons, in *Physical Knots: Knotting, Linking, and Folding Geometric Objects in R^3 : AMS Special Session on Physical Knotting and Unknot-*

- ting, Las Vegas, Nevada, April 21-22, 2001*, Contemporary Mathematics No. 304, edited by Calvo, J. A., Millett, K. C., and Rawdon, E.J. (American Mathematical Society, Providence, R.I., 2002).
- [36] Here, a ring polymer is modeled by interior-disjoint unit diameter spheres centered at v_1, \dots, v_n where the spheres centered at v_i and v_{i+1} (as well as v_n and v_1) must be tangent [54–56].
 - [37] N. Madras and A. D. Sokal, The pivot algorithm: A highly efficient Monte Carlo method for the self-avoiding walk, *Journal of Statistical Physics* **50**, 109 (1988).
 - [38] N. Clisby, Calculation of the connective constant for self-avoiding walks via the pivot algorithm, *Journal of Physics A: Mathematical and Theoretical* **46**, 245001 (2013).
 - [39] N. Clisby, Efficient implementation of the pivot algorithm for self-avoiding walks, *Journal of Statistical Physics* **140**, 349 (2010).
 - [40] N. Clisby, Accurate estimate of the critical exponent ν for self-avoiding walks via a fast implementation of the pivot algorithm, *Physical Review Letters* **104**, 055702 (2010).
 - [41] N. Clisby and D. T. C. Ho, Off-lattice and parallel implementations of the pivot algorithm, *Journal of Physics: Conference Series* **2122**, 012008 (2021).
 - [42] S. Schnabel and W. Janke, Monte Carlo simulation of long hard-sphere polymer chains in two to five dimensions, *Macromolecular Theory and Simulations* **32**, 2200080 (2023).
 - [43] S. Schnabel and W. Janke, Accelerating polymer simulation by means of tree data-structures and a parsimonious Metropolis algorithm, *Computer Physics Communications* **256**, 107414 (2020).
 - [44] H. Schumacher and J. Cantarella, Knoodle, <https://github.com/HenrikSchumacher/Knoodle> (2025).
 - [45] S. Boldo, Kahan’s algorithm for a correct discriminant computation at last formally proven, *IEEE Transactions on Computers* **58**, 220 (2009).
 - [46] C.-P. Jeannerod, N. Louvet, and J.-M. Muller, Further analysis of Kahan’s algorithm for the accurate computation of 2×2 determinants, *Mathematics of Computation* **82**, 2245 (2013).
 - [47] J. Cantarella, H. Schumacher, and C. Shonkwiler, Hard unknots are often easy from a different perspective (in preparation).
 - [48] J. Cantarella, H. Schumacher, and C. Shonkwiler, A very fast code for knot classification up to 13 crossings (in preparation).
 - [49] J. Cantarella, T. Deguchi, H. Schumacher, C. Shonkwiler, and E. Uehara, *Data for: Random knotting in very long off-lattice self-avoiding polygons [Dataset]* (2025), Dryad.
 - [50] C. J. Geyer, Practical Markov chain Monte Carlo, *Statistical Science* **7**, 473 (1992).
 - [51] This is the maximum difference between the probabilities of any event (such as $m_K^n = m$) according to the two distributions.
 - [52] M. Baiesi, E. Orlandini, and A. L. Stella, The entropic cost to tie a knot, *Journal of Statistical Mechanics: Theory and Experiment* **2010**, P06012 (2010).
 - [53] T. Deguchi and E. Uehara, Topological sum rules in the knotting probabilities of DNA, in *Topology and Geometry of Biopolymers*, Contemporary Mathematics, Vol. 746, edited by E. Flapan and H. Wong (American Mathematical Society, Providence, RI, USA, 2020) pp. 57–83.
 - [54] S. D. Stellman and P. J. Gans, Efficient computer simulation of polymer conformation. I. Geometric properties of the hard-sphere model, *Macromolecules* **5**, 516 (1972).
 - [55] R. Kenyon and P. Winkler, Branched polymers, *American Mathematical Monthly* **116**, 612 (2009).
 - [56] A. McMullen, M. Holmes-Cerfon, F. Sciortino, A. Y. Grosberg, and J. Brujic, Freely jointed polymers made of droplets, *Physical Review Letters* **121**, 138002 (2018).