The Knot Spectrum of Random Knot Spaces

Abstract: It is well known that knots exist in natural systems. For example, in the case of (mutant) bacteriophage P4, DNA molecules packed inside the bacteriophage head are considered to be circular since the two sticky ends of the DNA are close to each other. The DNAs extracted from the capsid, without separating the two ends, can preserve the topology of the (circular) DNAs, and hence are well-defined knots. Furthermore, knots formed within such systems are often varied and different knots occur with different probabilities. Such information can be important in biology. Mathematically, we may view (and model) such a biological system as (by) a random knot space and attempt to obtain information about the system via mathematical analysis and numerical simulation. The question here is to find the probability that a randomly (and uniformly) chosen knot from this space is of a particular knot type. This is equivalent to finding the distribution of all knot types within this random knot space (called the knot spectrum in an earlier paper by the authors). In this paper, we examine the behavior of the knot spectrums for knots up to 10 crossings. Using random polygons of various lengths under different confinement conditions as the random knot spaces (model biological systems), we demonstrate that the relative spectrums of the knots, when divided into groups by their crossing numbers, remain surprisingly robust as these knot spaces vary. For a given knot type $\kappa$, we let $P_{\kappa}(L, R)$ be the probability that an equilateral random polygon of length $L$ in a confinement sphere of radius $R$ has knot type $\kappa$. We give a model for the family of functions $P_{\kappa}(L, R)$ and show that our model function fits the random polygon data we generated. For a fixed crossing number $Cr$, $3 \leq Cr \leq 10$, let $S_{Cr}$ be the subspace consisting of random polygons which form knots that have crossing number $Cr$. We study the relative distribution of all the different knot types within $S_{Cr}$ and illustrate how this distribution changes if we keep the length $L$ fixed (or the confinement radius $R$ fixed) and vary the confinement radius $R$ (or the length $L$). We observe that this distribution is quite robust and remains essentially unchanged under length and confinement radius variation, especially if one concentrates on subfamilies such as alternating prime knots, non-alternating prime knots, or composite knots.
Keywords: DNA packing; topology of circular DNA; random polygons; DNA knots; knot spectrum of random polygons

Mathematics Subject Classification 2010: 57M25, 92B99

10.1 Introduction

In the world of pure mathematics, or more specifically knot theory, one typically considers the space of knots as the entire collection of all knots, with each distinct knot type as an element of this discrete space. It is commonly agreed that some knots are in general “more complicated” than other knots, although there are many different ways to define knot complexity and sometimes a knot more complicated with one complexity measure may be simpler in terms of a different complexity measure. For example, some knots with high crossing numbers have small bridge numbers and some knots with relatively small crossing numbers have high bridge numbers. One could imagine that in an ideal knot space, each knot type is no more probable than any other knot type. However, this is not the case for knots occurring in the world of nature. For example, if one is to tie a knot with a rope of some fixed length, then one can only tie finitely many different types of knots ([4, 10, 11, 25]). Furthermore, if one is to tie a knot with a rope in a random fashion, then one will inevitably reach the conclusion that some knots are “easier” to tie than other knots (meaning some knots will be tied more often than other knots if the experiment is repeated). Some important biological problems concerning subjects such as DNA packing are also related to this topic. In DNA research, a relatively simple virus called bacteriophage is commonly used to study the DNA packing mechanism. This virus keeps its genome in a spherical protein container called a capsid. Although the packing of the DNA inside the capsids cannot be directly observed, DNA extracted from the capsids without being broken retains its topological information and this information is used as a probe in studying the DNA packing mechanism [2]. It is reported that many different DNA knots form within the capsids of bacteriophages, with certain knot types appearing with much higher frequencies [2].

Mathematically, a biological system such as the DNA knots within the bacteriophage head is really just a “knot space” whose elements are geometric closed curves with a certain (continuous) probability distribution. If one is to randomly and uniformly sample a knot from this space, the probability that one gets a particular knot type is the same as the percentage of knots of this knot type within the space. The distribution of the knot types within the space is called the knot spectrum in [12, 16]. In a sense, the knot spectrum (of a knot space) measures how easily a knot type can be realized by a randomly chosen knot in that space and provides a knot complexity measure. While it is quite conceivable that specific knot spaces can be constructed to favor a particular type of knots, we are interested in the following question: Are there
knots that are intrinsically easier to form in most random knot spaces without obvious topological biases? For example, in the space of random equilateral polygons with a fixed length, it is known that composite knots have higher frequencies than the prime knots when the polygons are long (in fact the probability of getting a prime knot goes to zero as the length of the polygon goes to infinity [9]).

In this paper, we study the above problem by examining the behavior of the knot spectrum for knots up to 10 crossings. Using random polygons of various lengths under different confinement conditions as the random knot spaces (model biological systems), we demonstrate that the relative spectrums of the knots, when divided into groups by their crossing numbers, remain surprisingly robust as these knot spaces vary. For a given knot type $\mathcal{K}$, we let $P_{\mathcal{K}}(L, R)$ be the probability that an equilateral random polygon of length $L$ in a confinement sphere of radius $R$ has knot type $\mathcal{K}$. We give a model for the family of functions $P_{\mathcal{K}}(L, R)$ and show that our model function fits data we generated from a large sample of random polygons. Furthermore, for a fixed crossing number $Cr$ ($3 \leq Cr \leq 10$) we consider the knots space $S_{Cr}$ of random polygons that form knots with $Cr$ crossings. We observe how the frequency distribution of knots in $S_{Cr}$ changes under various lengths and confinement conditions. One of the main findings in this article is that this distribution is quite robust and remains essentially unchanged under length and confinement radius variation, especially if one concentrates on subfamilies such as alternating prime knots, non-alternating prime knots, or composite knots. The support of the previous statement is given by the data we collected. The details of the data collection will be explained in Section 10.3. Even though our data set is quite large, it does not allow us to support any statements on the individual knot type distributions for knots with more than 10 crossings since there are very few samples per knot type for crossing numbers above 10. Furthermore, already for nine and 10 crossings, the natural sampling error is large enough to only weakly support any conjecture about the independence of knot type distributions on polygon length and confinement radius. As a final remark, we need to acknowledge that we do not see any way to prove any statement on knot distributions. At this point we cannot even prove much simpler statements like: The probability that a random polygon is knotted is larger if the polygon is confined when compared to an unconfined random polygon of the same length.

This article is organized as follows: In Section 10.2 we introduce some basic terminology in knot theory. In Section 10.3 we outline the size of our random polygon sample and how the knot types in our sample were determined. In Sections 10.4 and 10.5 we explain the choice of a model function for $P_{\mathcal{K}}(L, R)$ and show how well our data fits the model. Next, in Sections 10.6 and 10.7 we study how the confinement radius $R$ and the polygon length $L$ affect the distribution of the different knot types in $S_{Cr}$. Finally, in Section 10.8 we summarize our findings.
10.2 Basic mathematical background in knot theory

For the convenience of our reader, we outline and discuss briefly a few topological concepts that are most relevant to this paper. For a more detailed exposition, please refer to a standard text on knot theory such as [1, 6, 21, 26].

A knot $K$ is a simple closed curve in $\mathbb{R}^3$. Here we assume that such a curve is a piece-wise smooth curve (this includes a space polygon without self-intersections). Two knots are considered topologically equivalent if one can be continuously deformed, together with the entire $\mathbb{R}^3$ space surrounding it and without being broken or causing self-intersection in the process, to the other. The class of all knots equivalent to a knot is called a knot type. The knot type that contains the unit circle is called the trivial knot (type). For a fixed knot $K$, a regular projection of $K$ is a projection of $K$ onto a plane such that no more than two segments of $K$ cross at the same point in the projection. At each intersection point of a regular projection of $K$, it is usually marked which strand is over and which strand is under. These intersection points are called crossings. A regular projection with this over/under information marked at the crossings is also called a knot diagram. The minimum number of crossings among all possible knot diagrams of knots with the same knot type as $K$ is called the crossing number of $K$ (and is usually denoted by $Cr(K)$). A knot diagram is alternating if at the crossings under and over alternate as one travels along the knot projection. A knot type is alternating if it has an alternating diagram and is non-alternating if it does not have any alternating diagram. A knot is called a composite knot if it is realized by connecting two nontrivial knots as shown in Figure 10.1. If a knot is not a composite knot, then it is a prime knot.

![Fig. 10.1: A composite knot from two non-trivial knots $K_1$ and $K_2$.](image)

One fundamental problem in knot theory is to determine the knot type of a given knot. The most common and powerful tools for this purpose are various knot polynomials. The well-known knot polynomials include the Alexander polynomial, the Jones polynomial, and the HOMFLYPT polynomial. These polynomials can be com-
puted from any knot diagram and remain unchanged when computed from a different knot diagram of the same knot type. In this paper, the authors rely on the HOMFLYPT polynomial and the knot table for knots up to 16 crossings for knot identification. For the definition of the HOMFLYPT polynomial and its properties, one may refer to [24]. It is important to note that the knot polynomials are not sufficient to distinguish all knot types and the tabulation of knots remains a difficult question in general. Furthermore, only prime knots up to 16 crossings have been completely tabulated [20].

10.3 Spaces of random knots, knot sampling and knot identification

In a series of papers, three authors of this paper have developed algorithms for several models to generate equilateral random polygons that are confined inside a sphere of fixed radius [13, 14, 15, 17]. The model presented in [14] is the one chosen for the study presented in this paper. The model can be described as follows: Consider equilateral random polygons that are “rooted” at the origin and assume that there is an algorithm that samples such objects with uniform probability. Now consider a confinement sphere \( S_R \) of radius \( R \geq 1 \) with its center at the origin. We keep those randomly generated equilateral polygons that are contained in the confinement sphere \( S_R \). Note that the algorithm used in [14] to generate polygons in confinement is not based on a direct accept-reject method (since such a method is extremely inefficient). Instead, it uses conditional probability density functions that can be explicitly formulated to guide the generation process. Each polygon is generated one edge at a time and there is no rejection involved. Furthermore, this algorithm generates polygons that are totally independent of each other so no de-correlation is necessary. Interested readers please refer to [14] for a detailed description of this algorithm. There is no biological or other reason for the polygons to be rooted at the center. It is rather a choice for simplicity: as it turns out, equilateral random polygons defined this way are much easier to generate due to the symmetry of the confining sphere (relative to the root) imposed on the equilateral random polygons. In order to obtain statistically significant results, it is necessary to use large samples and the simplicity of this model allows the computations to be feasible. We want to point out to the reader that, recently, a different and very efficient method to generate such confined polygons (also centered at the origin) was developed, see [7].

Our knot space consists of all random equilateral polygons of a given length and a given radius of confinement. By varying these two parameters, we obtain a family of knot spaces. Notice that the differences among these knot spaces are imposed by different geometric conditions and no explicit topological constraints exist. The data set used for our analysis in this paper comes from two sources.
The larger portion of the data is new and we describe how it is obtained here. To trace the effect of polygon length on knotting, we fixed the confinement radius at $R = 3$ and used a range of lengths of the random polygons from 10 to 90 in increments of 10. We used a maximal polygon length of 90 steps and a confinement radius $R = 3$ to ensure that, even for the longest sample, we can still identify the knot type for most polygons. This means that many of the knots that are generated must have at most 16 crossings so they are in the current knot table. From our past experience [12], we know that if we choose a confinement radius that is smaller than $R = 3$, then we would not be able to identify the knot types for the longer lengths. To trace the effect of the confinement radius, we fixed the polygon length at 30 steps and used radii of confinement ranging from $R = 1$ to $R = 4.5$. In more detail, between $R = 1$ and $R = 3$ the confinement radius increases with an increment of $1/10$, while from $R = 3$ to $R = 4.5$ the confinement radius increases with an increment of $1/2$. The reason for the wider spacing of the larger radii is that the confinement effect diminishes rather quickly for the polygons within the length-range of this study. The polygon length of 30 is chosen so that sufficiently many different knot types populate our knot spaces and that we are able to identify the knot type for most polygons. For each knot space (with a fixed confinement radius and a given polygon length), we sampled 100,000 different polygons. In total, this sample space consists of 32 sets, each containing 100,000 polygons (24 samples of length 30 with varying confinement and 9 samples of varying length with confinement $R = 3$), yielding a total of 3,200,000 random polygons (since one sample is in both sets).

A smaller portion of the data (1,640,000 polygons) is from one of our earlier papers [12]. We basically took the same set of radii (24 values) as described above and the same set of different length (9 values). For each combination of a radius and length 10,000 polygons where collected. However, for some of the larger lengths and smaller confinement radii, we could not identify the knot types reliably. The knots are too complex for our knot identification process to work – see the description below. Thus, some of these data sets were not used. We refer the reader to [12] for information about the exact description of the data. It is enough to know that this data set is grouped into subsets of size 10,000 for each fixed $(L, R)$ pair, where $1 \leq R \leq 4.5$ and $10 \leq L \leq 90$. Moreover, data points that are collected from this older data set have a larger error margin since the sample size is much smaller (10,000 versus 100,000 polygons).

As a final remark, our combined data set contains almost five million polygons. In the following sections we will refer to the two data sets as the old data (the 1,640,000 polygons) and the new data (the 3,200,000 polygons).

In the following we briefly outline our procedure of knot identification for a given polygon $P$. The polygon $P$ is projected onto a plane to obtain a knot diagram. Then unraveller [28] is used to (potentially) simplify the crossing information via a collection of simplification operations based on Reidemeister moves. The code unraveller produces two types of output: a DT-code which is used by knotfind [19] and crossing information
which is used to compute the HOMFLYPT polynomial $H$ using a program written by Ewing and Millett [18]. We have a table of HOMFLYPT polynomials for all chiral knot types, prime and composite, with 16 or fewer crossings (under the generally accepted assumption that the crossing number of composite knots is additive). For each $P$, we obtain a list of chiral knot types $\{K_1, \ldots, K_s\}$ from the table, all of which have $H$ as their HOMFLYPT polynomial. Note that if this list is empty then the polygon does not represent any prime or composite knot with fewer than 17 crossings. In addition to the HOMFLYPT polynomial calculation, for each polygon we also use knotfind to compute the (non-chiral) knot type. We use the simplified DT-code $D$ generated by unraveller, which might be the DT-code of a minimal diagram or might be the DT-code of a diagram close to the minimal diagram. Then knotfind takes $D$ and creates a “canonical” DT-code $D'$. If $D'$ represents a knot within the DT-code knot table (which only contains prime knot types) then that uniquely identifies that knot (up to chirality). If both methods agree then we claim that the knot type of $P$ is identified. If the two methods disagree, that is one method produces a knot that we should be able to identify, but we cannot confirm this with the other method, then we resolve this issue as follows: Either we start over using a different projection of the polygon $P$ or we run another diagram simplification program written by some of the authors. In this way, all such conflicts were resolved. Thus, we claim that with very high probability we identified all knots that are in the current knot table [16].

However there are still several additional issues which we address in the following paragraphs.

(i) The polygon $P$ might represent a knot that is not prime and therefore is not in the DT-code knot table. At every stage of the simplification process using DT-codes, knotfind attempts to identify factors of composite knots. If there is an “obvious” connected sum then a part of the DT-code maps onto itself. If such a situation is detected, the simplification process is applied separately to each DT-code of the two factors. In the end, we obtain a collection of simplified DT-codes from which it is possible to reconstruct the original composite knot type. If the HOMFLYPT polynomial of this reconstruction agrees with the originally computed HOMFLYPT polynomial of the polygon then we identify $P$ as the appropriate composite knot.

(ii) If the initial calculation of $H$ indicates that $P$ does not represent a knot in the table, then the simplified DT-code (or the set of simplified DT-codes in the case that an obvious connected sum was identified) provides an upper bound on the crossing number. In these cases, we double-checked by computing the HOMFLYPT polynomial from the simplified DT-code(s) to ensure that it matches $H$. We identify $P$ as having the crossing number provided by the simplified DT-code.

We observed that the approach to knot simplification based on DT-codes is extremely reliable for the knots that are within the knot table. There is no reason to believe that the simplification of DT-codes becomes suddenly unreliable once the actual
crossing number exceeds 16. Thus, we believe that a simplified DT-code, while technically only providing an upper bound on the actual crossing number, gives a value that, in most cases, is actually the topological minimal crossing number. In particular, this is true for crossing numbers that are not far above 16 (the largest crossing number in the knot table). Therefore we report these approximated crossing numbers as if they are the actual crossing numbers. We also report a knot as prime if no composition was detected during the simplification process.

10.4 An analysis of the behavior of $P_{\mathcal{K}}$ with respect to length and radius

Let $P_{\mathcal{K}}(L, R)$ denote the probability that a randomly selected polygon in the set of equilateral random polygons of length $L$ in a confinement sphere of a fixed radius $R$ has knot type $\mathcal{K}$. In this section, we analyze the dependence of $P_{\mathcal{K}}(L, R)$ on the polygon length $L$ (for fixed $R$ values) and on the confinement radius $R$ (for fixed polygon lengths $L$). At the end of the section we view $P_{\mathcal{K}}(L, R)$ as a function of both $L$ and $R$.

10.4.1 $P_{\mathcal{K}}(L, R)$ as a function of length $L$ for fixed $R$

For each knot type $\mathcal{K}$, it is obvious that there is a minimal length $L_0$ such that $\mathcal{K}$ can be realized by an equilateral polygon of length $L_0$ but not by any equilateral polygon of length less than $L_0$ ($L_0$ is called the equilateral stick number of the knot type $\mathcal{K}$ [22, 27]). Thus, $P_{\mathcal{K}}(L, R) = 0$ for $L < L_0$. For all the data analyzed in this study, we have $2R < L$ which means that the polygons were generated under confinement pressure (though very light pressure for the data set for $L = 10$ and $R = 4.5$). Our numerical results strongly support the following conjecture: for any fixed $R$, as $L$ increases, $P_{\mathcal{K}}(L, R)$ increases to a single maximum at length $L_{\mathcal{K}}$, and then declines to zero as $L$ approaches infinity, see the left of Figure 10.2. Notice that a similar phenomenon has been observed in [8] in the absence of confinement. We further conjecture that the rate of the decline is exponential in terms of some positive power of $L$. This conjecture has been proven when the polygons are not confined [9]. In that proof [9], it is shown that as $L$ increases, the number of connected sum components in the equilateral random polygon increases with a probability that goes to one, hence the probability for the polygon to be of a given knot type goes to zero. Unfortunately, this argument may not be applicable to confined knots since a long confined random equilateral polygon may not have many connected sum components. However, it is quite plausible and intuitive that the overall knot complexity of a confined polygon increases as its length increases and thus knots with a lower complexity become less likely. This is strongly supported by our numerical results.
10.4.2 $P_K(L, R)$ as a function of confinement radius $R$ for fixed $L$

Here we assume that the fixed $L$ value is large enough so that a knot of type $K$ can be easily formed. Since we can only model confinement radii $R \geq 1$, we assume that $P_K(L, 1) > 0$ is a positive number. We can only speculate on the behavior of $P_K(L, R)$ for $R < 1$. Clearly no knots can exist for $R \leq 1/2$, so $P_K(L, R) = 0$ for $R \leq 1/2$. However, the behavior of $P_K(L, R)$ for $1/2 < R < 1$ could be different for different knot types. Extreme confinement may favor certain knot types, so it is conceivable that, for some knot type $K$, we have $P_K(L, R) = 0$ for all $R \leq R_0$ for some value $R_0 > 1/2$. Moreover, the value $R_0$ could depend on the knot type and have different values for different knot types. Since we have no data for $R < 1$, we only stipulate that $P_K(L, 1) > 0$ is a positive number in our model. Just as in the case of dependence on length, we propose that with increasing radius $P_K(L, R)$ rises to a single maximum at radius $R_K$, and then declines. However, this time $P_K(L, R)$ becomes a positive constant $U_L(K)$ once $R$ becomes large enough (say $R > L/2$). This is because, for large $R$, the effect of confinement disappears and $P_K(L, R)$ is simply the probability for an unconfined equilateral random polygon of length $L$ to be of knot type $K$. By our initial assumption on $L$, this probability is positive and we denote it with $U_L(K)$.

10.4.3 Modeling $P_K$ as a function of length and radius.

Based on known rigorous results and our numerical results, we propose the following model function for $P_K(L, R)$:

$$P_K(L, R) = a \left( d + \left( \frac{L - L_0(K)}{R - 0.6} \right)^e \right) \exp \left( -\frac{L}{bR - c} \right),$$

where $L_0(K)$ is the equilateral stick number of $K$, $L \geq L_0(K)$, $R \geq 1$, $R > c/b$, and $a$, $b$, $c$, $d$, $e > 0$ are positive constants depending on the knot type $K$. The condition $R > c/b$ insures that the exponential term decreases with $L$. In addition we set $P_K(L, R) = 0$
for $L < L_0(\mathcal{K})$. Notice that for a given fixed $R$ value, $P_{\mathcal{K}}(L, R)$ is a function of $L$ with the following properties:

1. $\lim_{L \to \infty} P_{\mathcal{K}}(L, R) = 0$; and
2. $P_{\mathcal{K}}(L, R)$ has a single maximum point.

On the other hand, for a fixed $L$ value, $P_{\mathcal{K}}(L, R)$ is a function of $R$ with the properties:

1. $\lim_{R \to \infty} P_{\mathcal{K}}(L, R) = ad > 0$; and
2. $P_{\mathcal{K}}(L, R)$ has a single maximum point.

### 10.5 Numerical results

#### 10.5.1 The numerical analysis of $P_{\mathcal{K}}(L, R)$ based on the old data

We fitted the above model function $P_{\mathcal{K}}(L, R)$ to knot data of several knots up to seven crossings. Here we used the older data set, where each data point is based on a sample of 10,000 polygons. Three examples are shown in Table 10.1 with the corresponding fitting parameters and $R^2$ values:

| Tab. 10.1: The parameters for the model $P_{\mathcal{K}}(L, R)$ for various knot types |
|---|
| parameter | $3_1$ | $5_2$ | $6_2$ |
| a | 0.000205 | $2.17277 \times 10^{-7}$ | $1.73334 \times 10^{-9}$ |
| b | 5.95573 | 4.28009 | 3.85288 |
| c | 2.2024 | 1.3623 | 1.31161 |
| d | 124.299 | 9736.84 | 335759. |
| e | 3.43966 | 5.66256 | 7.09841 |
| $R^2$ | .996 | .986 | .982 |

Figure 10.3 is a plot of the fitting function $P_{3_1}(L, R)$ together with the actual data points from two different view points. Here $L_0(3_1) = 6$. The maximal $z$-value of the data points is $\approx 0.21$ and the average value is $\approx 0.108$. If we compute the absolute value of the difference between the $z$-values of the data points and the best fit function, we get a maximal value of $\approx 0.026$ and an average value of $\approx 0.0067$.

In Figure 10.4 we show the fitting functions for $P_{5_2}(L, R)$ and $P_{6_2}(L, R)$ together with the actual data points. Note that for both of these knots $L_0(5_2) = L_0(6_2) = 8$ [27]. The maximal $z$-value of the data points for $5_2$ ($6_2$) is $\approx 0.059$ ($\approx 0.027$) and the
**Fig. 10.3:** Two different views of the surface $P_3(L, R)$ together with the actual data points. The slightly different coloring of the data points indicates if the points are above or below the surface.

**Fig. 10.4:** Left: The plot of the fitting function $P_5(L, R)$ together with the actual data points; Right: The plot of the fitting function $P_6(L, R)$ together with the actual data points. The slightly different coloring of the data points indicates if the points are above or below the surface.
average value is \( \approx 0.022 \) \((\approx 0.0088)\). The maximum difference between the \( z \)-values of the data points for \( 5_2 \) (62) and the best fit function is \( \approx 0.011 \) \((\approx 0.0055)\) while the average difference is \( \approx 0.0026 \) \((\approx 0.0011)\).

Fig. 10.5: Comparison of the fitting function \( P_K(L, R) \) obtained from the newer data and the older data for the knots \( 3_1, 4_1, 5_1, \) and \( 5_2 \). Left: Fixed \( R \); Right: fixed \( L \). The solid curves are based obtained by connecting the new data points by line segments. The circles are the old data and the dashed lines represent a cross section through the surface \( P_K(L, R) \) fitted to the old data.

### 10.5.2 The numerical analysis of \( P_K(L, R) \) based on the new data

Overall, the newer and larger data sets did not produce any surprises in terms of the behaviors of the \( P_K(L, R) \) functions, in fact there is an amazingly good agreement. Here we show two examples. In Figure 10.5 on the left we compare the new data points with the old data points and with the slices for \( R = 3 \) from the surfaces \( P_K(L, R) \) obtained from the old data. The figure shows the data for the first four non-trivial knot types in the knot table. Here each round data point is based on the old data set (from a sample of 10,000 polygons), whereas the solid lines are connecting the data points from the new data (from a sample of 100,000 polygons). Each of the two dashed curves shows the slice for \( R = 3 \) of \( P_K(L, R) \) when fitted to the 164 data points of the old data for knot types \( 3_1 \) and \( 4_1 \). Similarly, Figure 10.5 on the right shows a comparison for \( L = 30 \) based on the new data (solid curves) with their counterparts based on the old data (circles) and of two fitting functions \( P_K(L, R) \) (dashed curves) based on the old data. The \( R^2 \) values for the fitting functions based on the new data points are all \( \geq 0.999 \).

### 10.5.3 The location of local maxima of \( P_K(L, R) \)

The maximal points of \( P_K(L, R) \) for a fixed knot type \( K \) lie on a line in the \( (L, R) \) plane as shown in Figure 10.6 on the left using the example of the four knots \( 3_1, 4_1, 5_2, \) and
6_2. The points shown in the figure are computed from the model \( P_K(L, R) \) obtained from fitting the older data. The slope of these lines increases with knot complexity from about \( \approx 20.41 \) for the trefoil knot to about \( \approx 27.35 \) for the knot 6_2. Furthermore, for a fixed radius \( R \), it takes more and more length for the maximum to occur as the knot complexity increases.

In Figure 10.6 on the right, we show the relationship of the maximal value of \( P_K(L, R) \) and the \( (L, R) \) plane using the same four examples of 3_1, 4_1, 5_2, and 6_2. Here we see that the height of the maxima rapidly declines as the knot complexity increases. It appears that as the radius (and length) increases, the maximal value of \( P_K(L, R) \) initially drops slightly and then remains roughly constant. In Figure 10.7 we show each of these curves in a coordinate system by itself and now it appears that the value of each maximum of \( P_K(L, R) \) slightly increases for increasing \( R \) after the initial drop.

### 10.6 The influence of the confinement radius on the distributions of knot types

In this section we consider only the new data, and we would like the reader to recall that these are all polygons of length 30. A recurring observation made in this section is that the relative distribution of knots in a given group of knots (e.g. the group of 7-crossing alternating knots, or the 8-crossing composite knots) is remarkably steady for many different confinement radii. The last few subsections in this section illustrate this visually.

Many of the figures presented in Sections 10.6 and 10.7 contain error bars to indicate the reliability of the data. Usually the error bars are only added to some of the
curves of each graph so that the graph remains intelligible. The error bars were determined using the standard interval estimation for binomial proportions (see [3]). The standard confidence interval is computed as $CI_s = \hat{p} \pm z_{a/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ where $n$ is the sample size, $X$ is the number of successes (occurrences of knots of interest) in the sample and $\hat{q} = 1 - \hat{p}$. The value $z_{a/2}$ is the $100(1-a/2)$-th percentile of the standard normal distribution. In this article we use $a = 0.05$.

### 10.6.1 3-, 4-, and 5-crossing knots

Our data contains 501,838 knots with three, four, or five crossings. There are 322,609 31 knots, 85,134 41 knots, 35,184 51 knots, and 58,911 52 knots. There is a maximum of 21,291 knots at radius $R = 1.0$ and a minimum of 6,951 knots at radius $R = 4.5$. We note that, for all radii, the order of the knot types (from least frequent to most frequent) is consistent, as shown in Figure 10.5 on the right.

### 10.6.2 6-crossing knots

Our data contains 69,080 6-crossing knots (all polygons are of length 30). There are 17,940 61 knots, 21,077 62 knots, 12,620 63 knots, and 17,443 31#31 composite knots. There is a maximum of 7,065 knots at radius $R = 1.0$ and a minimum of 205 knots at radius $R = 4.5$. We note that we do not distinguish the square and the granny knot – instead the two different composite knots are grouped together as 31#31. The four functions $P_{3\kappa}(30, R)$ are shown in Figure 10.8 on the left. The figure shows that the order of knot types is not the same for all radii, unlike for the smaller knot types
shown in Figure 10.5 on the right. The order of the four knot types seems to be the same for the smaller radii \(1 \leq R \leq 1.5\), while it is hard to see what happens for the larger radii. In order to make this more visible, we show the relative percentage of the four different knot types, see Figure 10.8 on the right. Here, for each fixed \(R\)-value, the \(y\)-values of the four data points add up to one. This kind of normalization is done frequently in the following sections of this article. In order to make the differences of probabilities between different knot types more visible for a particular family of knots (for example knots with a fixed crossing number or knots with a particular property such as 9-crossing composite knots) we scale the actual percentages such that the sum of the scaled percentages adds up to one for all knots in this family. This new quantity is called relative percentage and in Figure 10.8 on the right we see this for the first time. Most noticeable in this figure is the relative decline of the composite knots as the confinement radius decreases. This is not really surprising as it has been shown in [12, 16] that tighter confinement suppresses composite knots. We also observe that the order of the alternating knots (from least frequent to most frequent) is consistent, except when the error bars become too large, and that the knot \(6_3\) occurs with the smallest frequency. This is due to the fact that \(6_3\) is achiral and achiral knots have been observed with a lower frequency than chiral knots of the same crossing number [12].

![Fig. 10.8: On the left: The percentage (or actual probability) of the four different 6-crossing knots. On the right: the relative percentage distribution of the four knot types. For a fixed \(R\), the frequencies of all data points for that \(R\) add up to one.](image)

### 10.6.3 7-crossing knots

Our data contains 40,071 7-crossing knots. There are 2, 587 \(7_1\) knots, 5, 125 \(7_2\) knots, 4, 763 \(7_3\) knots, 2, 835 \(7_4\) knots, 6, 497 \(7_5\) knots, 7, 665 \(7_6\) knots, 3, 717 \(7_7\) knots, and 6, 882 \(3_1 \# 4_1\) composite knots. There is a maximum of 5, 775 knots at radius \(R = 1.1\) (there are 5, 745 at radius \(R = 1.0\)) and a minimum of 47 knots at radius \(R = 4.5\). Since the values of \(P_K(30, R)\) are quite small, we only display the relative
percentages of all 7-crossing knots in Figure 10.9. As before, we observe that the relative percentage of the composite knot $3_1 \# 4_1$ declines with decreasing confinement radius. In addition, for radii $R < 2$, the order of the relative percentage of the 7-crossing prime knots seems to be independent of the radius. The knot $7_6$ is the most frequent, followed by $7_5$, $7_2$ or $7_3$, $7_7$, and $7_4$ or $7_1$. At this point, we have no explanation of why the data shows this ordering of the knot types. We have compared this order with the order of knot energies [23] and ropelength [5], however we found no correlation at all (see Table 10.2).

**Tab. 10.2:** The order of 7-crossing knots based on various measures.

| knot | $P_{\mathcal{K}}(30, R)$ | ropelength | Möbius energy approximation |
|------|-----------------|-------------|-----------------------------|
| $7_1$  | 7               | 1           | 1                           |
| $7_2$  | 3 or 4          | 3           | 2                           |
| $7_3$  | 3 or 4          | 2           | 5                           |
| $7_4$  | 5 or 6          | 4           | 4                           |
| $7_5$  | 2               | 5           | 5                           |
| $7_6$  | 1               | 7           | 6                           |
| $7_7$  | 5 or 6          | 6           | 7                           |

**Fig. 10.9:** The relative percentage of the different 7-crossing knot types on the vertical axis as a function of the confinement radius $R$.

### 10.6.4 8-crossing knots

In the following, we investigate the knot type distribution for 8-crossing knots. Our sample contains 42,711 knots with eight crossings, with a maximum of 7,301 knots
at radius $R = 1$ and a minimum of 38 knots at radius $R = 4.5$. As the number of knot types increases, the graphs become more and more complicated. There are 21 prime knot types and 3 composite knot types to consider. A figure similar to Figure 10.9 would be too crowded and so we display the information in a different way.

First, we claim that the knot type distributions are remarkably consistent regardless of the radius of confinement. In other words, while the exact value of the relative percentage is not the same, the order of the knot types (from smallest to largest relative percentage) is often the same. To support this, we no longer show a curve for each knot type versus the radius since this would be 24 curves (21 prime and 3 composite). Instead we combine the knots of 4 consecutive radii and show one curve for each group of radii reducing the number of curves to plot to six. That is we combine $R = \{1.0, 1.1, 1.2, 1.3\}$, $R = \{1.4, 1.5, 1.6, 1.7\}$, $R = \{1.8, 1.9, 2.0, 2.1\}$, $R = \{2.2, 2.3, 2.4, 2.5\}$, $R = \{2.6, 2.7, 2.8, 2.9\}$, and $R = \{3.0, 3.5, 4.0, 4.5\}$. For each of the groups of four radii, we tabulate the distribution of each of the 24 knot types by using relative percentages as before. Thus for each of the six curves the total sum of the values over all 8 crossing knot types equals to one. In effect, in Figures 10.10, 10.11, and 10.13 the information plotted vertically for each knot type is a compressed version of what Figure 10.9 shows (for 7-crossing knots) with each curve. Finally, we split the data into two groups, and display the data for the knot types on the $x$-axis, see Figures 10.10 and 10.11.

![8 crossings](image)

**Fig. 10.10:** The relative percentages of the different 8-crossing knots from $8_1$ to $8_{13}$ grouped into six groups by the different radii. In each group, the smallest radius is shown as a plot label.

The most striking feature of Figures 10.10 and 10.11 is how much more likely the non-alternating knot types $8_{19}$, $8_{20}$, and $8_{21}$ are when compared to all other knot types (prime and composite). This property is independent of the radius of confinement.
However, the larger confinement radii favor non-alternating knots even more than the smaller confinement radii. If one looks closely at Figure 10.11 then one can see that for the non-alternating knots $8_{19}$, $8_{20}$, and $8_{21}$ the order of the curves (from the smallest relative percentage to the largest relative percentage) is by the size of the radius. That is, the smallest group of confinement radii leads to the least relatively frequent occurrence of the knot type and the largest group of confinement radii leads to the most relative frequent occurrences of the knot types. For prime knots, the order of the curves is often the opposite, the curve representing the smallest group of confinement radii is on the top for many knots and the curve with the largest group of confinement radii is on the bottom. We suspect that this inconsistency in the order of curves maybe due the fact that our sample is not large enough, so that for knot types that have a much smaller frequency, the sample error could be large enough to create a different order. For the three composite knots $3_1 \# 5_1$, $3_1 \# 5_2$, and $4_1 \# 4_1$ in Figure 10.11, we notice that the curve for the largest radii is on the bottom (smallest relative percentage). However, the order of the other curves is not clear and the data points are very close to each other.

The least likely knot type is $8_{18}$ whose standard diagram is based on the basic Conway polyhedron $8^*$, which has already been observed in [12]. More precisely, the knot $8_{18}$ appears only 49 times in our sample, while the next lowest occurrence is an order of magnitude larger (the knot $8_3$ with 557 cases). The knot that occurs the most is $8_{20}$ with 6,964 cases. In order to get more insight, we look at some subsets of the knot types. Figure 10.12 on the left shows that $3_1 \# 5_2$ is more likely than $3_1 \# 5_1$, which in turn is more likely than $4_1 \# 4_1$. The curve for $3_1 \# 5_1$ has statistical error bars which are similar to those of the other two curves. This distribution of the composite types is expected since $5_2$ is more likely than $5_1$ and the composite $4_1 \# 4_1$ is achiral, and therefore less likely. The figures hint that this distribution is largely independent of
the confinement radius. Figure 10.12 on the right shows the frequencies of the three non-alternating knot types 8_{19}, 8_{20}, and 8_{21}. The figure shows that 8_{20} is more likely than the other two, which have about the same frequency, with 8_{21} being slightly more likely than 8_{19}.

![Graph](image)

**Fig. 10.12**: On the left, the relative percentages of 8-crossing composites versus the confinement radius. On the right, the relative percentages of 8-crossing non-alternating prime knots versus the confinement radius.

### 10.6.5 9-crossing knots

In Figure 10.13, we show the 9-crossing knot spectrum in the same way as we showed the knot spectrum for the 8-crossing knots. Since there are many more knot types (49 prime and 6 composite) the knot types are not labeled on the horizontal axis. However we indicate the extent of the three knot groups: prime alternating knot types, prime non-alternating knot types, and composite knot types. Our sample contains 26,742 knots with nine crossings, with a maximum of 5,665 random knots at radius \( R = 1 \) and a minimum of 18 knots at radius \( R = 4.5 \). The total of 26,742 random knots is about 63% of the number of 8-crossing random knots while the number of knot types is twice as large (24 versus 55 knot types). Thus as expected, the fluctuations due to the smaller random sample muddle the picture. Nevertheless, the curves (maybe with the exception of the curve for the largest radii) are still roughly parallel. Here, by roughly parallel, we mean that the six curves could be put into a relatively narrow band – it does not mean that the curves cannot intersect each other. The reader should also keep in mind that the smaller the radius the more actual knots are represented by such a curve. Thus the curve representing the largest radii has the largest statistical error, which is reflected by the fact that it often crosses through all the other curves. The smallest numbers of knots occur for 9_{40} (6 knots), 9_{41} (39 knots), 9_{35} (46 knots) and 9_{34} (60 knots). We also note that the standard diagrams of the knots 9_{40} and 9_{34} are based on the Conway basic polyhedra 9^* and 8^* respectively, see also [12]. The largest numbers of knots occur for 9_{42} (2, 604 knots), 9_{44} (2, 585 knots), 9_{45} (2, 028 knots),
and $9_{43}$ (1, 799 knots). The next knot in the frequency list drops by more than 600. As in Figure 10.11, there is a large increase in the number of non-alternating knots. The same observations we made for 8-crossing knot types about the order of the different curves according to size of the confinement radius could be true. However, the picture is less clear, which could be due to the larger sample variance caused by the larger number of knot types and the smaller overall knot sample.

![Graph showing relative percentages of different 9-crossing knots grouped into six groups by the different radii.](image)

**Fig. 10.13:** The relative percentages of the different 9-crossing knots grouped into six groups by the different radii. In each group, the smallest radius is shown as a plot label.

In order to get more insight, we investigate some subsets of the knot types: composite knot types and non-alternating knot types. Figure 10.14 on the left shows that $3_1 \# 6_2$ is more likely than $3_1 \# 6_1$, which in turn is more likely than $3_1 \# 6_3$. This is as expected since $6_2$ is more likely than $6_1$, which in turn is more likely than $6_3$. We also see that $4_1 \# 5_2$ is more likely $4_1 \# 5_1$, which is as expected since $5_2$ is more likely than $5_1$. The curve for $3_1 \# 3_1 \# 3_1$ has statistical error bars which are similar to the other curves. The frequency of $3_1 \# 3_1 \# 3_1$ is about the same as that of $4_1 \# 5_1$. This shows that this distribution is largely independent of the confinement radius. Figure 10.14 on the right shows the relative frequencies of the eight non-alternating knots $9_{42}$ to $9_{49}$. It shows that $9_{47}$, $9_{48}$, and $9_{49}$ are much less likely than any of the others. We also note that a standard diagram of the knots $9_{47}$ is based on the Conway basic polyhedra $8^*$. 
10.6.6 10-crossing knots

In Figure 10.15 we show the relative percentages of the 10-crossing spectrum. Since there are many knot types (165 prime and 14 composite), the knots are not labeled on the x-axis. Our sample contains 24, 277 knots with 10 crossings, with a maximum of 6,021 knots at radius $R = 1$ and a minimum of 9 knots at radius $R = 4.5$. There are about as many 10-crossing random knots as 9-crossing random knots but the number of knot types is more than three times as large. Thus, we expect the fluctuations due to the smaller random sample to be even larger than it was for 9-crossing knots. There are very few knots for the larger radii, thus it makes very little sense to use the data for the large radii. As a result, in Figure 10.15 we only show three curves, the first for radii $R = 1.0 - 1.3$ contains 18,118 random knots, the second for radii $R = 1.4 - 1.7$ contains 4,647 random knots, and the third for radii $R = 1.9 - 2.1$ contains 1,004 random knots. (The fourth would only contain 310 knots and therefore does not yield a useful distribution.)

If we look at the five knots that have the lowest frequency, then four of these have a standard diagram based on Conway basic polyhedra: 10_{123} (2 knots and 10' polyhedra), 10_{115} (3 knots and 8' polyhedra), 10_{121} (4 knots and 9' polyhedra), and 10_{122} (4 knots and 9' polyhedra). In fifth place is 10_{99} (8 knots) which is not based on Conway basic polyhedra. As in the previous figures, we can see the large increase in the number of non-alternating knots. The non-alternating knots with the largest frequencies are as follows: 10_{132} (1007 knots), 10_{133} (762 knots), 10_{124} (596 knots), 10_{128} (596 knots), and 10_{137} (585 knots).

It is impossible to see in Figure 10.15 if the curves are still roughly parallel. To get a feeling of how similar the curves are, in Figure 10.16 we show a zoomed in version of the first two curves in Figure 10.15 using only the alternating knots of 10 crossings. Here we can see that there is still a lot of similarity between the two curves.

The error bars for most of the individual knot types are too large to allow definite conclusions. We illustrate this by just looking at the 10-crossing composites in Figure 10.17. The figure on the right shows the 10-crossing composite knots which include a $3_1$...
Fig. 10.15: The relative percentages of the different 10-crossing knots grouped into three groups by the different radii. In each group, the smallest radius is shown as a plot label.

Fig. 10.16: A zoomed in version of Figure 10.15 showing only the relative percentages of the different alternating 10-crossing knots using the first two groups of the different radii.
Confined knots

component. Their relative percentages seem to be determined by the relative percentages of the associated 7-crossing knots, with \(7_5\) and \(7_6\) the most likely, causing \(3_1\#7_5\) and \(3_1\#7_6\) to be the most likely composite knots. Similarly, \(7_1\) and \(7_4\) are the least likely, causing \(3_1\#7_1\) and \(3_1\#7_4\) to be the least likely composite knots in the graph (see Figure 10.9 and Table 10.2). On the right part in Figure 10.17, a similar observation holds for 10-crossing composite knots which contain a \(4_1\) component.

![Figure 10.17](image)

**Fig. 10.17:** On the left, the relative percentages of 10-crossing composites with \(3_1\), and, on the right, the relative percentages of 10-crossing composites with \(4_1\) versus the confinement radius.

### 10.7 The influence of polygon length on the distributions of knot types in the presence of confinement

In this section, we discuss the dependence of the distribution of knot types on length in the presence of confinement. We want to remind the reader that we consider only the new data, and all polygons are in a sphere with confinement radius \(R = 3\). This investigation is similar to that of Section 10.6. Intuitively, we expect that keeping the confinement radius fixed and increasing the length of the polygons has an analogous effect to keeping the length of polygons fixed and decreasing the confinement radius. Thus many of the results are similar to those of the last section. This includes the observations that the relative distribution of knots in a given group of knots is remarkably steady for many different confinement radii.

#### 10.7.1 3-, 4-, and 5-crossing knots

Our data contains 184,517 knots with three, four, or five crossings. There are 117,059 \(3_1\) knots, 31,827 \(4_1\) knots, 13,075 \(5_1\) knots, and 22,556 \(5_2\) knots. There is
a maximum of 32,711 knots at length $L = 70$ and a minimum of 666 knots at length $L = 10$. We note that for all lengths the order (from most frequent to least frequent) of knot types is consistent, as shown in Figure 10.5 on the left.

### 10.7.2 6-crossing knots

Our data contains 33,161 6-crossing knots (all in a confinement sphere of radius $R = 3$). There are 7,351 $6_1$ knots, 8,311 $6_2$ knots, 5,107 $6_3$ knots, and 12,392 $3_1#3_1$ composite knots. There is a maximum of 7983 knots at length $L = 90$ and a minimum of 2 knots at length $L = 10$. The four functions $P_K(L, 3)$ are shown in Figure 10.18 on the left. The figure shows that the order of the four knot types remains unchanged, with $3_1#3_1$ having the highest frequency, followed by $6_2$, $6_1$, and $6_3$ with the lowest frequency. Figure 10.18 on the right shows the relative frequency of the four different knot types. Here for each fixed value $L$, the $y$-values of the four data points add up to one.

The most noticeable difference to the case in the previous section is the relative stability of the knot distribution, that is for length $L \geq 30$ the four functions in Figure 10.18 on the right look basically constant. The previous section shows that stronger confinement causes a relative decline in the number of composite knots leading to the expectation of a decline of composite knots for increased polygon lengths (and fixed confinement radius). Figure 10.18 indicates that this is not the case. It has already been observed in [12, 16] that for composite knots, the analogy between decreasing the radius of the confinement sphere for polygons and increasing the length of the polygons in a fixed confinement sphere breaks down. We also observe that the relative order of the prime knot types is the same in both cases (see Figure 10.8 for the smaller confinement radii) and that the achiral knot $6_3$ occurs with the smallest frequency.

![Figure 10.18](image-url)  

**Fig. 10.18:** On the left: The relative percentages of the four different 6-crossing knots versus length. On the right: the relative percentages of the four knot types versus length.
10.7.3 7-crossing knots

Figure 10.19 shows the dependence of the distribution of 7-crossing knot types in the presence of a confinement sphere of radius $R = 3$. Our sample contains 19,971 7-crossing knots. There are no 7-crossing knots of length 10 in our sample and the data starts with 15 knots of length 20, rising to 5,587 knots of length 90. The least likely knot is $7_4$ which occurs 1063 times in the sample, while the composite knot $3_1#4_1$ occurs with the largest frequency of 5,769. It appears that the relative distribution of 7-crossing knots varies only slightly with the length of the polygons for $L \geq 40$. From the data for 6-crossing knots (see Figure 10.18), one can speculate that the relative distribution of 7-crossing knots is also independent of the lengths of the polygons. There are significant similarities between this relative percentage distribution and the data for the dependence of the relative percentage of 7-crossing prime knots on the radius of confinement as shown in Figure 10.9: The curves show that $7_6$ is the most frequent, followed by $7_5$. Next are $7_2$ and $7_3$ which are again approximately equally likely. This is followed by $7_7$ and at the end we have $7_1$ and $7_4$ (again approximately equally likely). Just as in the case of the 6-crossing knots, the difference between the two distributions shown in Figures 10.19 and 10.9 is the relative percentage of the composite knot $3_1#4_1$, which for small radii has a frequency similar to $7_5$ (or $7_6$), but for radius $R = 3$ is the most likely.

![Graph](image)

**Fig. 10.19:** The relative percentages of the different 7-crossing knot types versus length.

10.7.4 8-crossing knots

Next we look at 8-crossing knot types in the presence of a confinement sphere of radius $R = 3$. Our sample contains 20,921 8-crossing knots. There is a single knot of length 10, 12 knots of length 20, rising to 6,475 knots of length 90. The least likely knot is
8_{18} which occurs 19 times in the sample, while the composite knot $3_1 \# 5_2$ occurs with the largest frequency of 3,486. The prime knot with the largest frequency is the non-alternating knot $8_{20}$, which occurs 2,521 times.

It again appears that the relative distribution of 8-crossing knots varies little with the length of the polygons once the polygons are long enough. Data supporting this claim is presented in several graphs. For Figures 10.20 and 10.21, all polygons up to length 50 are combined into one group since, for the smaller lengths, there are not many 8-crossing knots. As before we show relative percentages, that is for each curve (representing different lengths) the sum of the percentages over all knot types adds up to one. Similar to the data for the dependence of the relative percentages of 8-crossing knots on the radius of the confinement sphere, there are many more non-alternating prime knots than alternating prime knots, although there are also many more composite knots. Just as in the case of 7-crossing knots, the composite knots behave differently with increasing polygon length than they do with decreasing confinement radius. We also note that for the non-alternating prime knots, the curves are strictly ordered by length, with the shortest length leading to the largest relative frequency (on top) and the longest length leading to the lowest relative frequencies (on the bottom). The behavior for the alternating prime knots and composite knots is inconsistent but the order of curves is mostly reversed. Thus with the exception of the composite knots, this behavior is similar to the behavior of the 8-crossing knot types with a changing confinement radius.

![Graph](image)

**Fig. 10.20:** The relative percentage of the 8-crossing knots from $8_1$ to $8_{13}$ grouped into six groups by the different lengths.

Figure 10.22 shows the relative percentages of non-alternating knots (right) and the composite knots (left). It is remarkable how close the relative percentages for composite knots of length 90 ($R = 3$) ($\{0.32, 0.58, 0.1\}$ rounded to the nearest per-
Fig. 10.21: The relative percentages of the 8-crossing knots from $8_{13}$ to $4_1 \# 4_1$ grouped into six groups by the different lengths.

Fig. 10.22: The relative percentages of the different 8-crossing knot types on the vertical axis and its dependence on the polygon length. On the left: composite knots; on the right: non-alternating knots. The frequencies are normalized so that adding all values for a fixed length yields one.

For alternating prime knots of eight crossings, the picture is not clear since the error bars are too large to draw reliable conclusions about the frequency of all knot types. This is shown in Figure 10.23. Clearly, the knots that are the most frequent, $8_8$ and $8_{14}$, are also on the top of the knot type distribution for a fixed length and varying confinement radius, see Figures 10.10 and 10.11.
10.7.5 9-crossing knots

Our sample contains 14, 406 9-crossing knots. There are no knots of length 10, 4 knots of length 20, rising to 4, 913 knots of length 90. The least likely knot is 9_{40} which occurs 6 times in the sample, while the composite knots 31\#6_1 and 31\#6_2 occur with the largest frequency of 1, 007 and 1, 115 respectively. The prime knot with the largest frequency is the non-alternating knot type 9_{44}, which occurs 971 times. The most frequent alternating prime knot is 9_{8}, which occurs 307 times. There are 5, 925 alternating prime knots (9_{1} to 9_{41}), 3, 836 non-alternating prime knots (9_{42} to 9_{49}), and 4, 645 composite knots. This shows again that the frequency of non-alternating knot types and composite knot types is larger than the frequency of alternating prime knot types. Similar to the last two subsections, the data suggests that the relative distribution of 9-crossing knots varies little with changes of the polygon length once the polygons are long enough (L \geq 50). Data to support this claim is presented in two graphs. Figure 10.24 shows the relative percentage of non-alternating knots (right) and of composite knots (left). It is remarkable how close the percentages for composite knots of length 90 (and radius R = 3.0), rounded to the nearest percent, are to the percentages for composite knots at radius R = 1 (and length L = 30) (see Table 10.3). Just like in the section about the dependence of the knot distribution on the confinement radius, the order of the relative frequencies of 9-crossing composite knots is strongly related to the relative frequencies of the 6- and 5-crossing prime knots. Similar results are obtained for the non-alternating prime knot types of length 90 in confinement of radius R = 3 compared with length 30 polygons in confinement of radius R = 1 (see Table 10.3).
Tab. 10.3: Similar relative percentages for polygons of length 90 (and radius $R = 3.0$) and polygons in confinement radius $R = 1.0$ (and length $L = 30$). On the left for composite knots and on the right for non-alternating knots.

| knot         | length 90 | $R = 3$ | $L = 30$ | knot         | length 90 | $R = 3$ | $L = 30$ |
|--------------|-----------|--------|----------|--------------|-----------|--------|----------|
| $3_1 \# 3_1$ | 0.17      | 0.07   |          | $9_{42}$     | 0.24      | 0.22   |          |
| $3_1 \# 6_1$ | 0.22      | 0.22   |          | $9_{43}$     | 0.18      | 0.16   |          |
| $3_1 \# 6_2$ | 0.25      | 0.27   |          | $9_{44}$     | 0.24      | 0.25   |          |
| $3_1 \# 6_3$ | 0.14      | 0.20   |          | $9_{45}$     | 0.18      | 0.21   |          |
| $4_1 \# 5_1$ | 0.08      | 0.08   |          | $9_{46}$     | 0.10      | 0.08   |          |
| $4_1 \# 5_2$ | 0.15      | 0.15   |          | $9_{47}$     | 0.01      | 0.02   |          |
|              |           |        |          | $9_{48}$     | 0.03      | 0.04   |          |
|              |           |        |          | $9_{49}$     | 0.02      | 0.02   |          |

Fig. 10.24: On the left the relative percentages of different 9-crossing composite knot types versus length. On the right, the relative percentages of different 9-crossing non-alternating prime knot types versus length.

10.7.6 10-crossing knots

Our sample of knots with 10 crossings contains 12,326 knots. There are no knots of length 10, one knot of length 20, rising to 4,739 knots of length 90. The least likely knot is $10_{123}$, which is the only knot that does not occur in the sample. A minimal standard diagram of $10_{123}$ is based on the $10^*$ basic Conway polyhedron. We note that this is the only knot type with less or equal to 10 crossings that was not observed in the old data set either [12]. This confirms the observation that knots whose minimal standard diagrams are based on basic Conway polyhedrons occur with a probability that is much lower than the probability of other knot types [12]. The composite knot $3_1 \# 3_1 \# 4_1$ occurs with the largest frequency of 457. The prime knot with the largest frequency is the non-alternating knot $10_{132}$, which occurs 379 times. The most frequent alternating prime knot is $10_{20}$, which occurs 81 times. There are 3,751 alternating prime knots ($10_1$ to $10_{123}$), 5,520 non-alternating prime knots ($10_{124}$ to $10_{165}$), and
3,055 composite knots. This shows, just as for other crossing numbers, that the frequency of non-alternating knot types and composite knot types is larger than the frequency of alternating prime knot types. The relative distributions of 10-crossing knots in our data seems to vary more than for other graphs with the length of the polygons (once the polygons are long enough), however this is due to the large error bars. Thus, the authors claim that for a large enough sample, the statements of the previous sections would still hold: The relative distributions vary only slightly with changes in the length of the polygons. To support this, we show the non-alternating knots and the composite knots. The reader should compare Figure 10.25 with Figure 10.17 to follow the discussion in this paragraph. In the figure on the left are the composite knots with a 3\_1 component. Their frequency seems to be determined by the frequency of the 7-crossing knots, with 7\_5 and 7\_6 the most likely knots, causing 3\_1#7\_5 and 3\_1#7\_6 to be the most likely composite knots. Similarly 7\_1 and 7\_4 are the least likely knots, causing 3\_1#7\_1 and 3\_1#7\_4 to be the least likely composite knots, see Figure 10.9. On the right are the results for composite knots with a 4\_1 component. Here we would expect that 4\_1#6\_1 occurs with the highest frequency – but this does not always happen. However, as the error bar on the knot 4\_1#6\_2 indicates, the curves do not allow any definite conclusions.

**Fig. 10.25:** On the left, the relative percentages of 10-crossing composites with a 3\_1-component versus length. On the right, the relative percentages of 10-crossing composites with a 4\_1-component versus length.

### 10.8 Conclusions

We finish this chapter with several conjectures. Here $R$ is the radius of confinement and $n$ is the length of a polygon $P$ in our sample. We also want to remind the reader that the evidence collected in this article is restricted to $1 \leq R \leq 9/2$ and $10 \leq L \leq 90$. Furthermore our sample does not contain enough knots for large crossing numbers $Cr > 10$ to draw conclusions about the frequencies of individual knot types.
• The distributions of different prime knot types are virtually independent of the confinement radius and length. To be more precise, we claim that if $K_1$ and $K_2$ are two prime knot types in a knot space $S_{Cr}$ (that is both knots have the same crossing number $Cr$) and if $P_{K_1}(L_0, R_0) \ll P_{K_2}(L_0, R_0)$ for some confinement radius $R_0$ and some length $L_0$ (where $L_0$ is big enough so that both knots can be formed) then $P_{K_1}(L, R) \ll P_{K_2}(L, R)$ for all confinement radii $R \geq 1$ and all lengths $L \geq L_0$. Similarly if $P_{K_1}(L_0, R_0) \approx P_{K_2}(L_0, R_0)$ for some confinement radius $R_0$ and some length $L_0$ (where $L_0$ is big enough so that both knots can be formed) then $P_{K_1}(L, R) \approx P_{K_2}(L, R)$ for all confinement radii $R \geq 1$ and all lengths $L \geq L_0$.

• The above property also holds if both knot types in $S_{Cr}$ are composite knots.

• The relative probabilities of a composite knot type versus the probabilities of a prime knot type are virtually independent of the length but not independent of the confinement radius. To be more precise, we claim that if $K_1$ is a prime knot and $K_2$ a composite knot in a knot space $S_{Cr}$ (that is both knots have the same crossing number $Cr$) and if $P_{K_1}(L_0, R_0) \ll P_{K_2}(L_0, R_0)$ for some confinement radius $R_0$ and some length $L_0$ (where $L_0$ is big enough so that both knots can be formed) then $P_{K_1}(L, R_0) \ll P_{K_2}(L, R_0)$ for all lengths $L \geq L_0$. Similarly if $P_{K_1}(L_0, R_0) \approx P_{K_2}(L_0, R_0)$ for some confinement radius $R_0$ and some length $L_0$ (where $L_0$ is big enough so that both knots can be formed) then $P_{K_1}(L, R_0) \approx P_{K_2}(L, R_0)$ for all lengths $L \geq L_0$.

• For a fixed crossing number $Cr$ the probability that a polygon represents a knot in $S_{Cr}$ eventually decreases in strong confinement. That is to say, provided that the length $L$ is long enough so that $K$ can easily be formed, the function $P_K(L, R)$ is initially an increasing function in $R$ considering $R = 1/2$ as a starting point. For composite knots $K$ this increase is more pronounced than for prime knots $K$.

**Acknowledgment:** Y. Diao was supported by Chinese NSF grant number 91324201. E. J. Rawdon was supported by USA National Science Foundation grant number 1418869.

**Bibliography**

[1] Adams, C. (2004). *The Knot Book: An elementary introduction to the mathematical theory of knots*. American Mathematical Society.

[2] Arsuaga, J., Vázquez, M., Trigueros, S., Sumners, D. W., and Roca, J. (2002). Knotting probability of DNA molecules confined in restricted volumes: DNA knotting in phage capsids. *Proc. Natl. Acad. Sci. USA*, 99:5373–5377.

[3] Brown, L. D., Cai, T. T., and DasGupta, A. (2001). Interval estimation for a binomial proportion. *Statistical science*, pages 101–117.
[4] Buck, G. and Simon, J. (1999). Thickness and crossing number of knots. *Topology Appl.*, 91(3):245–257.

[5] Bunyi, R. V., Cantarella, J., Kephart, T. W., and Rawdon, E. J. (2014). Tight knot spectrum in QCD. *Physical Review D*, 89(5):054513.

[6] Burde, G. and Zieschang, H. (2003). *Knots*, volume 5. Walter de Gruyter.

[7] Cantarella, J., Deguchi, T., and Shonkwiler, C. (2014). Probability theory of random polygons from the quaternionic viewpoint. *Comm. Pure Appl. Math.*, 67(10):1658–1699.

[8] Deguchi, T. and Tsurusaki, K. (1998). Numerical application of knot invariants and universality of random knotting. In *Knot theory (Warsaw, 1995)*, volume 42 of *Banach Center Publ.*, pages 77–85. Polish Acad. Sci., Warsaw.

[9] Diao, Y. (1995). The knotting of equilateral polygons in $\mathbb{R}^3$. *J. Knot Theory Ramifications*, 4(2):189–196.

[10] Diao, Y. and Ernst, C. (1998). The complexity of lattice knots. *Topology Appl.*, 90(1-3):1–9.

[11] Diao, Y., Ernst, C., and Janse van Rensburg, E. J. (1997). Knot energies by ropes. *J. Knot Theory Ramifications*, 6(6):799–807.

[12] Diao, Y., Ernst, C., Montemayor, A., Rawdon, E., and Ziegler, U. (2014). The knot spectrum of confined random equilateral polygons. *Molecular Based Mathematical Biology*, 2(1):19–33.

[13] Diao, Y., Ernst, C., Montemayor, A., and Ziegler, U. (2011). Generating equilateral random polygons in confinement. *J. Phys. A: Math. Theor.*, 44:405202.

[14] Diao, Y., Ernst, C., Montemayor, A., and Ziegler, U. (2012a). Generating equilateral random polygons in confinement II. *J. Phys. A: Math. Theor.*, 45:275203.

[15] Diao, Y., Ernst, C., Montemayor, A., and Ziegler, U. (2012b). Generating equilateral random polygons in confinement III. *J. Phys. A: Math. Theor.*, 45:465003.

[16] Diao, Y., Ernst, C., Rawdon, E., and Ziegler, U. (2017). Relative frequencies of composite, alternating and non-alternating prime knots in random knot spaces. *Experimental Mathematics*, 6:1–18.

[17] Diao, Y., Ernst, C., Saarinen, S., and Ziegler, U. (2015). Generating random walks and polygons with stiffness in confinement. *J. Phys. A: Math. Theor.*, 48(9):095202.

[18] Ewing, B. and Millett, K. C. (1997). Computational algorithms and the complexity of link polynomials. In *Progress in knot theory and related topics*, pages 51–68. Hermann, Paris.

[19] Hoste, J. and Thistlethwaite, M. (1999). Knotscape. http://www.math.utk.edu/~morwen/knotscape.html. Program for computing topological information about knots.

[20] Hoste, J., Thistlethwaite, M., and Weeks, J. (1998). The first 1,701,936 knots. *Math. Intelligencer*, 20(4):33–48.

[21] Kawauchi, A. (1996). *A Survey of Knot Theory*. Birkhäuser Verlag, Basel.

[22] Kim, H., No, S., and Oh, S. (2014). Equilateral stick number of knots. *Journal of Knot Theory and Its Ramifications*, 23(07):1460008.
[23] Kusner, R. B. and Sullivan, J. M. (1998). Möbius-invariant knot energies. In Ideal knots, pages 315–352. World Sci. Publishing, Singapore.

[24] Lickorish, W. R. (2012). An introduction to knot theory, volume 175. Springer Science & Business Media.

[25] Litherland, R. A., Simon, J., Durumeric, O., and Rawdon, E. (1999). Thickness of knots. Topology Appl., 91(3):233–244.

[26] Livingston, C. (1993). Knot theory, volume 24. Cambridge University Press.

[27] Rawdon, E. J. and Scharein, R. G. (2002). Upper bounds for equilateral stick numbers. In Physical knots: knotting, linking, and folding geometric objects in $\mathbb{R}^3$ (Las Vegas, NV, 2001), volume 304 of Contemp. Math., pages 55–75. Amer. Math. Soc., Providence, RI.

[28] Thistlethwaite, M. (1990). unraveller. Code provided through private communication.