Knotting Fingerprints Resolve Knot Complexity and Knotting Pathways in Ideal Knots

David A. B. Hyde [*]
Joshua Henrich [†]
Eric J. Rawdon [†]
Kenneth C. Millett [‡]

May 18, 2015

Abstract

We use disk matrices to define knotting fingerprints that provide fine-grained insights into the local knotting structure of ideal knots. These knots have been found to have spatial properties that highly correlate with those of interesting macromolecules. From this fine structure and an analysis of the associated planar graph, one can define a measure of knot complexity using the number of independent unknotting pathways from the global knot type as the knot is trimmed progressively to a short arc unknot. A specialization of the Cheeger constant provides a measure of constraint on these independent unknotting pathways. Furthermore, the structure of the knotting fingerprint supports a comparison of the tight knot pathways to the unconstrained unknotting pathways of comparable length.

1 Introduction

Within the natural sciences, knotted, linked, and entangled macromolecules are encountered in a wide range of contexts and scales. Their presence has important implications for physical and biological properties. Understanding how their presence causes these observed properties is a matter of contemporary interest. In this research, we focus on the local structure of a robust family of knots, the "ideal" or "tight" knots [43], whose spatial properties have been found to correlate with those of interesting macromolecules [12, 23, 24, 29, 34, 41, 42, 49] and subatomic glueballs [5, 6, 7, 8].

[*]Institute for Computational & Mathematical Engineering, Stanford University, Stanford, CA 94305, USA

[†]Department of Mathematics, University of St. Thomas, St. Paul, MN 55105, USA

[‡]Department of Mathematics, University of California, Santa Barbara, CA 93106, USA

One can think of these configurations as starting from a single arc that then grows along the configuration to eventually be the whole configuration. During this growth, the open chain subarcs evolve from an unknotted segment through intermediate knotted states and ultimately to the base knot type. The intermediate knotted states depend on where one starts growing the configuration, and the voyage through intermediate knots from different starting positions reveals information about the spatial structure of the configuration. In this article, we analyze the evolution of the knotting in these subarcs through this growth and introduce quantities that measure the degree of complexity in obtaining the configurations. While we analyze ideal knots here, the techniques could be used to analyze any closed chain.

Alternately, one can think of the base configuration as being shrunk, i.e. begin with the entire configuration and then continuously digest the configuration from a given starting point. In this sense, our work also is an analysis of the complexity seen during decay.

We will focus on identifying the constituent local knotted arcs within a knotted ring, as expressed in the knotting fingerprint, and analyzing their interrelationships [37, 44]. To identify the fine-grain knotting structure of a complex knotted ring, we employ a slight alteration of the MDS method [30, 33] that defines the knot type of an open arc. In Section 3.3 we describe more specifically how to display this information in the form of a color-coded disk, a disk matrix, in which the color of each cell corresponds to the identified knot type of a corresponding subchain [36], see Figure 1. Briefly, the radial distance from the center expresses the length of the subchain, with short subchains near the center and the entire chain (minus one edge) giving the border of the disk. The angular coordinate expresses the middle point of the subchain. In Section 3.5 this knotting fingerprint is then translated into the planar graph associated to the fingerprint. The vertices in this graph correspond to the connected regions associated to the knot type. Among these vertices, the central unknot (as short segments are never knotted) and the peripheral region (corresponding to the global knot type) have special roles. We assess the complexity of the knot by measuring the constriction in minimal channels between the central unknotted region and the peripheral global region in the graph using the Cheeger constant [14]. We also analyze a measure of structural complexity using the number of edge- and vertex-independent paths in the graph that begin at the central unknot vertex and end at the global knot vertex. These measures provide information reflecting the spatial properties of the knot. For example, the number of independent paths is precisely the number of independent unknotting/knotting pathways associated to the given spatial conformation. Further analysis suggests ways, three-dimensional in character, by which one can measure the spatial complexity of the knot. These appear to capture information that is independent of classical knot invariants [13].

2
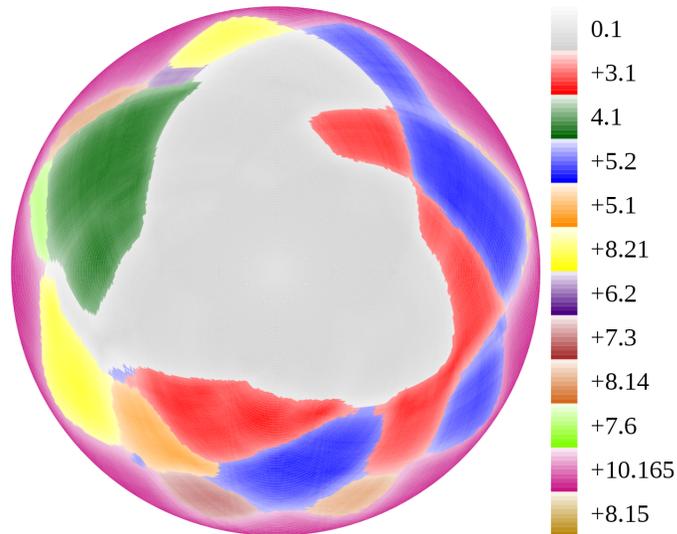
Figure 1: The knotting fingerprint of 10_{165}.

2 Ideal Knots

Ideal knots [4, 10, 21, 23, 26, 28, 43] are inspired by the result of tying a knot in some physical material (e.g. a piece of rope of some uniform thickness) and then seeking a conformation of the knot in which the length is the smallest possible. Thus, in the context of this study, we consider circular ropes of uniform thickness and minimal length among all conformations representing the same knot type. Such conformations are mathematically modeled by smooth curves, usually $C^{1,1}$ or C^2, for which one can define the radius of an embedded normal tube and the length of the curve. The *ropelength* of a knot is defined to be the minimum of ratios of the arclength and this thickness radius over all conformations of a given knot type. A curve realizing this minimum is then a *tight knot* or, equivalently, an *ideal knot*. Rigorous results for ideal knot conformations are very limited. For example, we only have very good estimates of both the lower and upper bounds for the minimal ropelength of the trefoil knot: it lies between 31.32 and 32.7429345 [17, 35]. As a consequence, we are limited to approximate conformations described by polygons resulting from computer simulations. In this research, we apply our analysis to the ideal prime and composite knot conformations resulting from the knot-tightening code *ridgerunner*, developed by Ashton, Cantarella, Piatek, and Rawdon [2, 11], as they appear to provide good upper bounds.
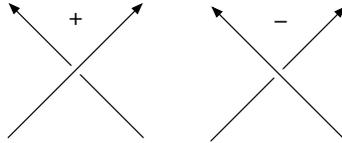
Figure 2: $+1$ and -1 algebraic crossing numbers at a crossing.

3 Knotting Fingerprints and the Associated Graphs

A closed chain in Euclidean 3-space is knotted if there is no ambient deformation of space taking the chain to the standard planar circle. The search for computationally efficient and effective methods to determine the specific structure of knotting for polygons is a continuing mathematical challenge. More importantly, the search for an appropriate formulation of knotting of open chains is even more challenging. From the classical topological perspective, knotting of open polygons is an artifact of a fixed spatial conformation because, if edge lengths and directions in the polygon are allowed to vary, each open polygon is ambient isotopic to a standard interval on the "x"-axis in 3-space (this is called the "light-bulb" theorem [38]). However, open polygons can be geometrically knotted if the edge lengths are fixed. This is demonstrated by the examples of Canteralla-Johnston and others [1, 9, 48].

3.1 Knot Identification

Here we study open, roughly equilateral polygons and require a robust method that will allow us to identify those that are "topologically" knotted. To do so, we compute the *HOMFLY* knot polynomial [20, 27] using the Ewing-Millett program [18]. This allows us to determine the chiral knot type, i.e. the spatial orientation of the knot in 3-space, with a high degree of reliability. A knot is said to be *chiral* if it is not equivalent to its mirror reflection. For many chiral knots, the *writhe* of a minimal crossing projection (defined as the algebraic sum of the crossing numbers, see Figure 2) is not zero, thereby defining a positive/negative instance dependent upon whether the writhe is positive/negative. If the specific knot, K, has positive writhe, it may be denoted by K, or by $+K$, depending upon the setting. If the negative instance is selected, it will always be denoted by $-K$. For alternating knots, the writhe of minimal crossing projections is an invariant [47] but is not for non-alternating knots. In addition, achiral knots, i.e. those equivalent to their mirror reflections, will have zero writhe in minimal crossing projections; however, note that this is not a sufficient condition for achirality. For knots with zero writhe minimum crossing projections, we use the standard presentations to identify which conformation will be $+$ and $-$.

4

3.2 A Variant of the *MDS Method*

In order to identify the knotting present in open chains, especially those used as models of protein structures, and being concerned with the uncertain features of some popular strategies, Millett, Dobay, and Stasiak [30] developed a stochastic method to identify and quantitatively measure the extent of knotting present in an open polygonal arc. This method has been employed in a study of knotting in random walks and tested against the previously identified knotting present in protein structures [33]. More recently, it has been employed to create the knotting fingerprint used in an extensive study of the presence and nature of knots and slipknots occurring in protein structures [22, 44].

To identify the knotting within open arcs, we use a slight variation of the *MDS Method*. Given an open polygonal arc, we close the configuration at infinity by extending rays in a common direction from each of the endpoints of the arc. Well beyond the convex hull of the polygonal arc, we connect the two rays to form a closed chain. We perform this new closure procedure 100 times per open arc using a roughly uniform distribution of directions. This choice is based upon extensive experimentation, during which we found that (1) a 'roughly uniform distribution of directions' is superior to a 'random distribution of directions' (consistent with the experience of researchers in numerical analysis [40]) and (2) a choice of 100 points provides data of sufficient quality for our purposes when compared to, for example, as many as 6400 points, based on our experience in earlier studies [30, 33, 36, 44]. The distribution of knot types on the two dimensional sphere of directions determines the *knotting spectrum* and provides a stochastic description of the knotting of the arc, see Figure 4 where we show the case of the DehI protein, PDB ID 3bjx [3, 44]. For all practical purposes, in a given open chain, this spectrum identifies a dominant knot type at the plurality level, see Figure 5. Here $10,000$ random walks of length 300 were generated and the occurrence of the most frequent knot type determined. In over 99% of the $10,000$ cases, one knot type appeared in more than 50% of the closures. Thus, when a single knot type occurs more than any other knot type in the closures, we identify this as the "knot type" of the segment and record the proportion of this knot type. The knot types of such knotted segments are called *subknots* of the chain. If a subknotted chain is contained in a larger unknotted segment, it is called an *ephemeral knot* and the unknotted chain is called a *slipknot* [31]. As a consequence, this approach provides a powerful method with which to analyze the knotting of open chains.

3.3 The Knotting Fingerprint

For open chains, specifically polygonal models of protein structures, triangular and square matrix arrays of colored cells have been employed to visually represent the knot types of the entire collection of subchains [25, 44, 46]. For a given knotted or unknotted polygonal ring of n edges, this triangular matrix fails to capture the periodic character of a closed ring. Therefore, one constructs a new *knotting fingerprint* given by a disk matrix structure consisting of $n - 1$
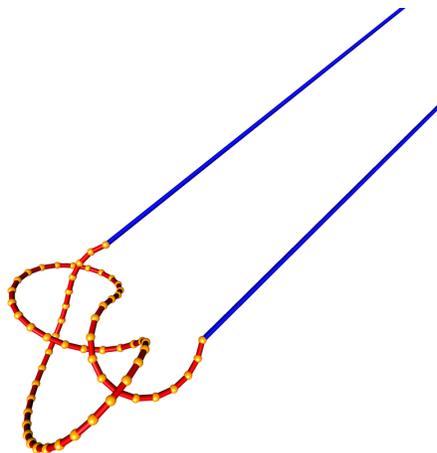
Figure 3: An example of one closure at infinity (two blue edges) for a polygonal configuration. For each open chain, 100 of these closures are created using a roughly uniform set of directions. The distribution of knot types of these closures is an approximation of the probability distribution over all closures and provides a stochastic description of the knotting of the open chain.
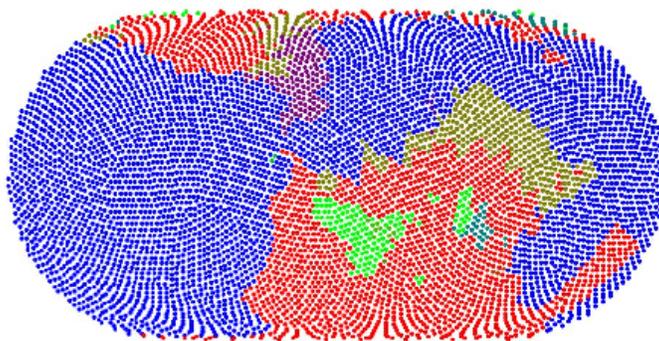
Figure 4: An Eckert IV area-preserving presentation of the spherical distribution of knot types in the DehI protein, PDB ID 3bjx [3, 44]. Each of the $64,000$ data points is coded to indicate the knot type of the closure from the given spherical closing direction. The Stevedore's knot, 6_1, is blue and represents 62% of the area, the unknot is red at 28%, the figure-8 knot, 4_1, is dark green at 6.0%, the trefoil, 3_1, is light green at 2.5%, 5_2 is brown at 0.8%, and 5_1 is blue-green at 0.1%, accounting for a total of 99.78% of data points. Five other knot types appear even less frequently: $3_1 \# 4_1$, 7_2, 8_2, 8_3, and 8_6.
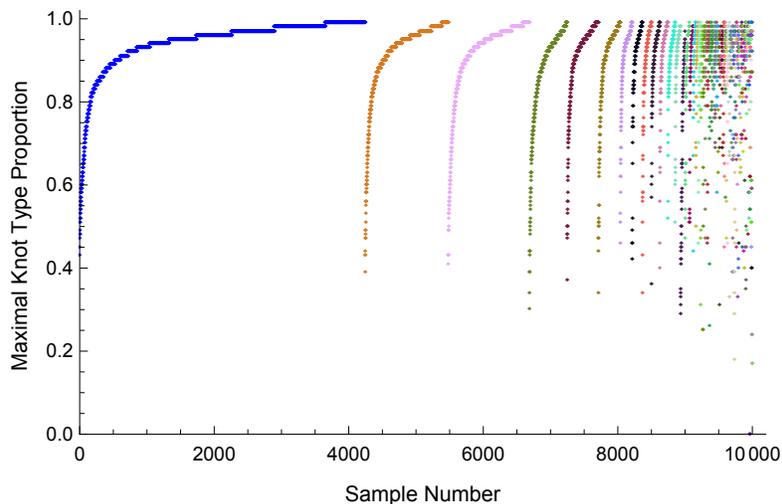
Figure 5: The largest knot type proportion for $10,000$ samples of 300-step random walks. Over 99% have a knot type appearing at 50% or more. The data is ordered by knot type, indicated by distinct colors, and by increasing proportion.

concentric rings, each of which is divided into n congruent colored cells [36]. Each of the colored cells is determined by first calculating the knotting spectrum of the associated subsegment of the chain and identifying the dominant knot type. This knot type has an assigned color giving the color of the cell, with the intensity of the color determined by the proportion of closures having the given knot type. The color of the unknot is always indicated by the color gray. The choice of color for other knot types is determined independently for each specific closed ring because the spectrum of knots appearing in it can change significantly with the choice of ring. These colored cells are arranged as follows: First, a base point and orientation of the chain is selected. For a given segment length, starting at one and increasing to $n - 1$, the colored segments are arranged at a constant radius corresponding to the proportion of the total length of the chain in a counter-clockwise fashion, with the angle from the initial cell indicating the middle point of the segment in the direction of the orientation. In Figure 6, we show the knotting fingerprint of an ideal 9_2 knot. The color bar to the right of the figure indicates the color code and intensity range for this knotting fingerprint. As very short segments of three edges or shorter must be unknotted, the central region of the knotting fingerprint is always gray. As the entire chain is always the global knot type, the cells in the outer ring of the knotting fingerprint are the color attributed to the knot type. As a result, each of the colored regions provides information about the knotting structure of the circular chain. For example, reading the color coding of rings of increasing radius, i.e. proportion of the total circular chain, one can determine the length

7
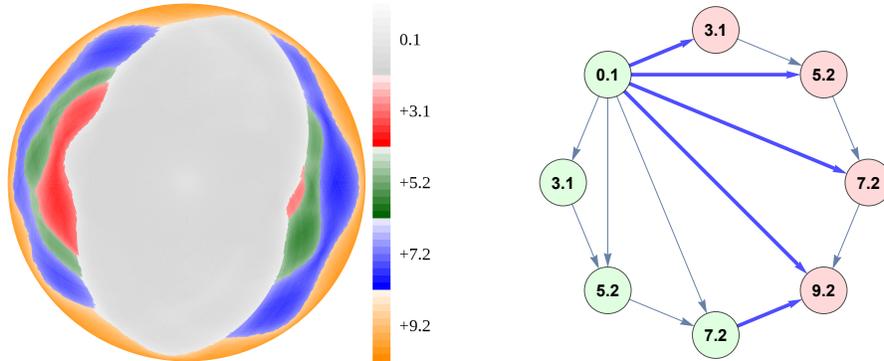
Figure 6: The knotting fingerprint of an ideal 9_2 and its associated knotting graph. The two different vertex colors identify the separation defining the Cheeger constant, here equal to $\frac{5}{4}$. The five thickened blue edges are those connecting the separated regions (each of size 4) that define the Cheeger constant.

of the shortest subsegment supporting the global knot type.

3.4 Analysis of the Knotting Fingerprint

The knotting fingerprints are limited by the resolution of the knot configurations they represent, i.e. corresponding to the number of segments in the chain. Hence there are certain scenarios where the apparent knotting fingerprints do not agree with what one might expect. In some cases, this is a matter of resolution, while in others, it may give evidence of an unanticipated evolution of the local structure. For example, we frequently observe tiny, e.g. single-cell, isolated regions of a certain type near but not contiguous to much larger regions of the same type. When this phenomenon occurs near the boundary of two regions in the knotting fingerprint, it suggests that the tiny regions are inadvertent artifacts, due to the limited resolution, and should not be considered as singular regions distinct from their larger neighbors of the same type. In such situations it may be appropriate to "smooth" the data so the boundaries between distinct regions are more regular.

In other situations, we observe features in the knotting fingerprint that appear to be inconsistent with one's interpretation of the consequences of knot theory. For example, there are several cases when the global knot has an unknotting number greater than one, but the unknot appears to connect to the global knot by the addition of a single segment. One might expect that the difference between unknotting numbers of adjacent regions must be no greater than one [16, 19], so these fingerprints may appear to be incorrect. For a single closure direction from the sphere of directions, the addition of a sufficiently small edge would account for no more than a single strand passage, but in our

case, there are two features that weaken this simplistic analysis. First, for a single closure direction, a single edge addition may cause more than one strand passage. This situation might be eliminated through higher resolution, i.e. by subdividing the edge segments of the chain. Second, our analysis concerns a stochastic process giving spherical regions representing the distinct knot types arising from the closures. The process of adding an edge causes an evolution of these regions. Thus our choice of the plurality knot type can lead to a jump of two or more in the strand passage difference between the competing knot types (see Figure 4). We will see that this represents a real artifact of the ideal knot presentation, not merely a question of its resolution. Therefore, although with greater resolution we would expect to see a more accurate knotting fingerprint, the strand passage difference between two adjacent regions may or may not reflect the structure of the ideal knot. For these analyses, we carefully account for this potential error by deleting the edge between the unknot and the global knot in the corresponding knotting graph when the ideal knot structure suggests that it is a resolution artifact.

3.5 The Knotting Graph

The knotting fingerprint (possibly smoothed) defines the planar *knotting graph* by associating a vertex to each of the connected knotting regions (the connected components of the knotting fingerprint) and associating edges between pairs of vertices whose knotting regions are contiguous, directed with increasing segment length. As described in the previous section, one may encounter spurious cells in the knotting fingerprint whose presence requires a "smoothing" of the regions. As a consequence, we employ the resulting smoothed fingerprints in constructing the knotting graph. In Figure 6, we show the knotting graph associated to the 9_2 fingerprint. In Figure 8, for example, note the presence of an edge from the gray unknot component, 0_1, to the green -3_1 component, and another edge going in the opposite direction due to the presence of a ray from the center to the outer edge traversing from gray to green and back to gray. Another instance of this structure occurs with the presence of slipknots, see Figure 10 where there are three such instances.

The presence of the edge connecting the vertices associated to two contiguous regions reflects the passage of one of the parallel supplementary edges added to the chain, see Figure 3, through one or more edges of the subsegment of the chain. If the resolution of the ideal knot is fine enough, one might expect that a single edge passage would occur and the unknotting numbers of the associated closed conformations would change, if at all, by at most one. The phenomena that the addition of a single small segment may result in a change of unknotting number of two or more is quite possible and actually occurs in practice. As noted earlier, this may be a resolution issue, i.e. one that could be resolved by a refinement of the chain structure or by an increase in the density of the uniform closure points. Thus one needs to look very closely at the possibility of a complex structural evolution as is illustrated by the case of composite knots (see Section 4.4 and Figure 11).

9

The knotting graph of 9_2, shown in Figure 6, has two distinguished vertices. The first corresponds to the component of small unknotted segments, labeled 0.1. The second corresponds to the knot type of the entire ring, labeled 9.2 in Figure 6, indicating that it is the 9_2 knot in the classical knot enumeration realized with the positive chirality. In addition, there are two components each of 3_1, 5_2, and of 7_2, as shown in Figure 6, each giving a vertex in the knotting graph. There are edges between the 3_1 vertices and the 0_1 and 5_2 vertices, as the red 3_1 components share common frontiers with the gray 0_1 and green 5_2 components, etc.

4   Analysis of the Knotting Graph

We employ the knotting graph associated with the knotting fingerprint of a given knot as the principal vehicle supporting our analysis of the spatial character of the knot. The unknot vertex and the global knot vertex anchor our analysis as we study the extent to which there are constraints inherent in the evolution between the unknot and the global knot, reflected in the structure of the knotting graph. Is there a small number of knotting states through which this evolution must pass? One powerful measure of such a constraint or "bottleneck" is provided by our specialization of the Cheeger constant.

4.1   The Cheeger Constant

In graph theory, the *Cheeger constant* [15] is a measure of whether a graph contains a "constriction" or "bottleneck." It is inspired by Cheeger's isoperimetric constant $h(M)$ for a compact Riemannian manifold, M, in terms of the area of a codimension one hypersurface, S, dividing the manifold into two disjoint pieces of equal volume [14]. For graphs, our modification of the Cheeger constant is defined as follows: Let G denote a connected graph, $V(G)$ be the vertices of G, and $E(G)$ be the edges of G. For a subset of vertices, S, containing either the initial unknot vertex or the global knot vertex (but not both), let ∂S denote the set of edges that have exactly one vertex in S, and let $|\partial S|$ be the number of such edges. We define the *Cheeger constant* by

$$h(G) = \text{minimum} \left\{ \frac{|\partial S|}{|S|} \ : \ 0 < |S| \leq \frac{|V(G)|}{2} \right\} .$$

This formulation of a Cheeger constant is designed to detect the presence of a constriction in the separation of the knotting graph that lies between the trivial knot and the global knot and, as such, represents a constriction in the growth of the knotting structure. In Figure 6, we show the set of vertices, $S = \{0.1, 3.1, 5.2, 7.2\}$, connected by five thick blue edges, ∂S, to the remaining vertices that include the global knot vertex, 9_2. This configuration has Cheeger constant equal to $\frac{5}{4}$, which, since it is greater than one, indicates that the 9_2 knotting fingerprint is not a constricted knotting formation.

4.2 Independent Knotting Pathways

Another measure of constriction is inspired by the *Max-Flow-Min-Cut Theorem* and the related theory of Menger [50]. We determine the maximum number of edge independent directed paths, i.e. no edge appears in more than one path, from the unknot vertex to the global knot vertex. In Figure 6, we observe that the maximum number of edge independent paths from the unknot vertex to the global knot vertex is three; in this case, the constraint is given by the degree of the global knot vertex. The maximal number of edge independent paths is bounded above by the degree of the unknot vertex, the degree of the global knot vertex, and the number of edges in the minimum edge cut set separating the unknot vertex from the global knot vertex [50]. The number of edges in the minimum edge cut set is related to the Cheeger constant as well as the Max-Flow-Min-Cut analysis. Specifically, the numerator of the unreduced Cheeger constant fraction is the number of edges in a cut edge set separating the trivial knot vertex from the global knot vertex and, therefore, is an upper bound for this number. We propose, therefore, to call the maximum number of undirected edge independent paths the *edge robustness index*, $ER(K)$, of the knotting graph. We observe, as is shown in Figure 6, that the specific set of paths is not unique.

While the Cheeger constant numerator and $ER(K)$ are equal for the simplest knots, the case of the twist knot, 7_2, illustrates both the relationship and the possible difference between these values for a given knot, see Figure 7. In this case the number of distinct knotting pathways is reduced by one due to the fact that vertex 7_2 has degree 3.

Alternatively, one could additionally require that the connecting paths are both edge and vertex independent. In this case one defines the *edge vertex robustness index*, $EVR(K)$, of the knotting graphs. We note that there are cases in which these two indices of a knotting graph are different. The smallest crossing number example is the knot 8_7, whose knotting fingerprint and graph are shown in Figure 8. An analysis of the associated knotting graph shows that there are six edge independent paths while there are only five edge-vertex independent paths.

4.3 Second Order Pathway Independence

In our analysis of knotting graphs, we discovered a collection of knots for which $EVR(K)$ is equal to one due to the existence of a *bridge edge* in the knotting graph, i.e. an edge whose removal disconnects the unknot vertex from the global knot vertex [50]. The simplest examples of this structure are the $(2, 2n + 1)$-torus knots, for example see Figure 9. In this case, all edges are bridges. In other cases, for example 8_{19} or 9_{35}, this first measure of robustness does not fully capture the complexity of the knotting fingerprint. We propose, as a consequence, to create a second order measurement associated to the two connected components that result from the removal of the bridge edge. One of the bridge's vertices can be identified as a terminal vertex when it lies in the component containing the unknot vertex, while the other can be identified as an initial vertex of

11
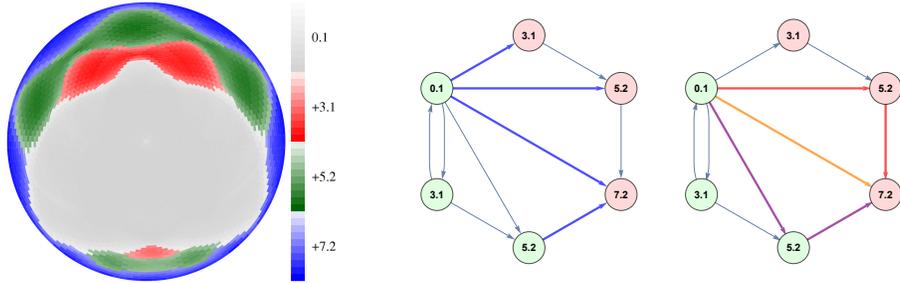
Figure 7: The 7_2 knotting fingerprint, on the left, and associated knotting graph illustrate the difference between the numerator of the Cheeger constant, 4, shown in the central graph and $ER(K) = 3$, shown in the right graph, due to the requirement that distinct paths employ disjoint edge sets. On the right, each bold path that contributes to $ER(K)$ is given its own color.
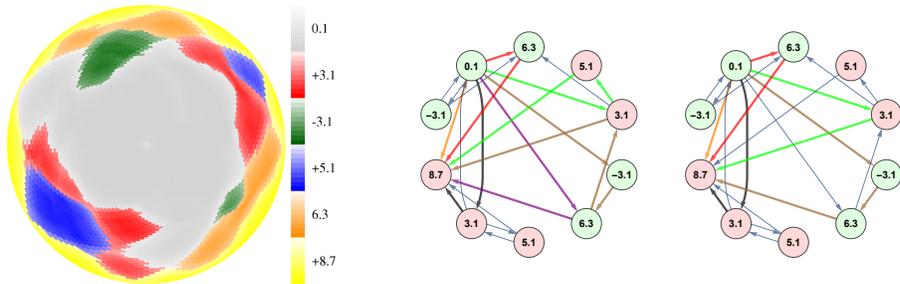
Figure 8: The 8_7 knotting fingerprint and associated knotting graphs illustrate the difference between $ER(K) = 6$ and $EVR(K) = 5$. In particular, the vertex independence restriction means that only one path may pass through the 3_1 node on the right of the graph when computing $EVR(K)$, while that vertex is used for two paths when computing $ER(K)$. Each bold path that contributes to $ER(K)$ or $EVR(K)$ is given its own color.

12
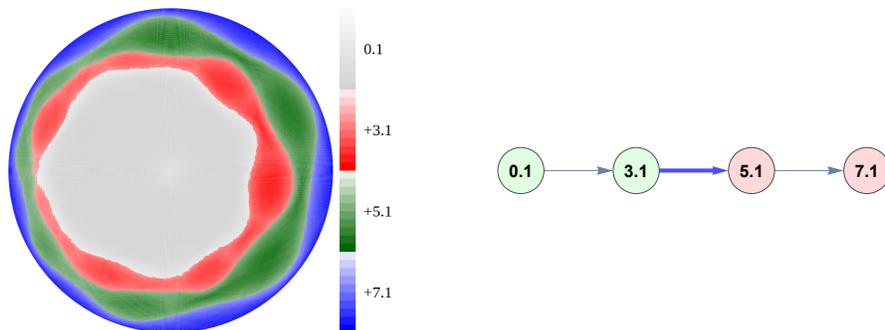
Figure 9: The 7_1 torus knot has a linear knotting graph. The two different vertex colors identify the separation defining the Cheeger constant, where the bold edge connects the separated regions.
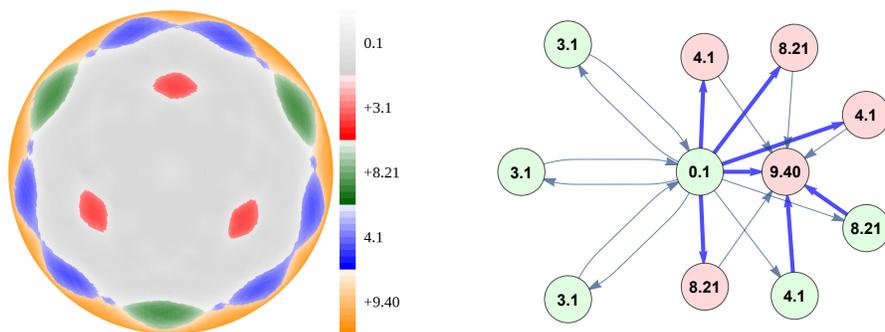
Figure 10: The 9_{40} knotting fingerprint and the associated knotting graph. The two different vertex colors identify the separation defining the Cheeger constant, where the bold edges are those connecting the separated regions.

the other component. We then determine the edge vertex robustness index for the two resulting subgraphs, thereby giving a pair of indices, (p, q), that define the second level of pathway independence. In some cases, the bridge is adjacent to the unknot vertex or the global knot vertex. In those cases, we would assign the index zero to the component consisting of the single vertex. We find the following second order indices for the prime knots, whose knotting graphs are non-linear but contain a bridge, through 10 crossings: $8_{19} : (0, 4)$, $9_{35} : (12, 0)$, $10_{120} : (0, 12)$, $10_{123} : (0, 10)$, $10_{124} : (0, 4)$, $10_{152} : (0, 6)$, and $10_{154} : (0, 9)$.

4.4 Prime Knots, Composite Knots, and Slipknots in Knotting Fingerprints

We have seen that the knotting fingerprints of ideal prime knots can be quite complex (see Figure 1) in that they contain a complex spectrum of subknot

13
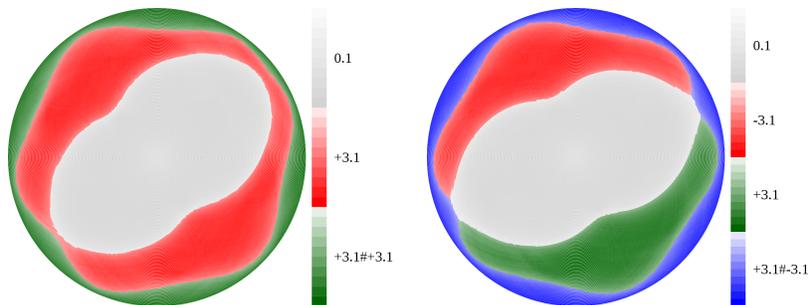
Figure 11: The knotting fingerprints of the connected sums $3_1\#3_1$ and $3_1\#-3_1$.

types. For example, one may find subknots of a prime knot that are ephemeral knots contained within slipknots, i.e. they are contained with larger segments that are unknotted (see Figure 10 in which there are three ephemeral trefoils which become unknotted as the number of segments is increased). In the associated knotting graph, see Figure 10, the associated ephemeral knot type regions correspond to vertices which, with the unknot vertex, support a loop reflecting the slipknot structure.

Ideal composite knots, the connected sums of two or more prime knots, can exhibit another type of complex structure (see Figure 11). Although $3_1\#3_1$ contains two distinct 3_1 knot components, its knotting graph is linear, reflected by the single 3_1 component in the knotting fingerprint. In contrast, the knotting fingerprint for $3_1\#-3_1$ exhibits two distinct components, one for each of the summands, separated by an unknot region. In this knotting graph, see Figure 12, the unknot component is contiguous to the connected sum component, a knot of unknotting number two. Thus, the fine structure of the knotting evolution in this area must be much more complex, perhaps along the lines discussed earlier, in which there is an evolving proportion that includes the unknot and the two distinct summands. In the case where the two summands are the same type, their knot type is cumulative, thereby providing the ring separation observed in the knotting fingerprint and a cut vertex in the knotting graph.

5 Knot Complexity

In the following sections and associated tables, we present the results of the calculations of the measures of knot complexity.

5.1 Cheeger Constant Complexity

We note that the only observed n-vertex linear knotting graphs are those associated with the family of ideal $(2, 2n + 1)$-torus knots. When this occurs, the Cheeger constant is $\frac{1}{n}$, the smallest values observed for odd minimal n-crossing number prime knots. In our data, this is the case for 3_1, 5_1, and 7_1. For 9_1,
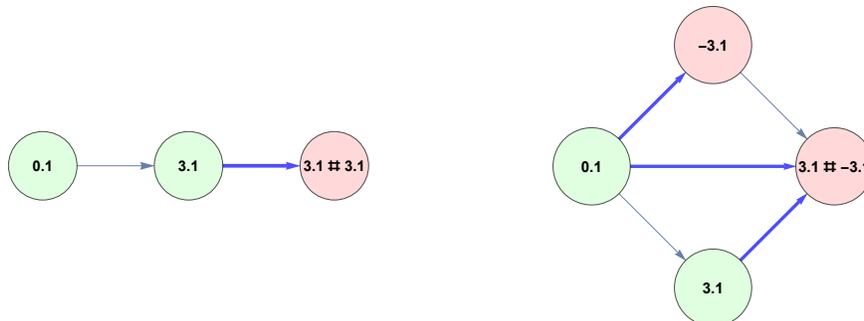
Figure 12: The knotting graphs of the connected sums: $3_1\#3_1$ and $3_1\#-3_1$. The two different vertex colors identify the separation defining the Cheeger constant, where the bold edges are those connecting the separated regions.

we see that this is no longer the case, although one can imagine that the relationship might hold with a higher quality ideal 9_1 and/or finer resolution (see Figure 13). On the other hand, one knows that knot length minimization can significantly disrupt the symmetry or regularity shown in classical knot presentations [11]. Although one might expect to see complete rings of each knot type in the 9_1 knotting fingerprint, as in the case of 7_1 (see Figure 9), here the 3_1 and 5_1 rings are incomplete, reflecting the possible disruption of the symmetry in the more complex ideal $(2, 2n + 1)$-torus knots.

This phenomenon is quite different from the one observed in the case of the connected sums of trefoil knots. There are two chirally distinct cases depending upon the the writhe, i.e. the $3_1\#3_1$ (granny knot) and $3_1\#-3_1$ (square knot). In Figure 12, one sees that $3_1\#3_1$ has a linear graph, as its knotting fingerprint consists of concentric rings similar to the torus knot case. The graph of $3_1\#-3_1$ is more complex because the unknot region and the global knot region are contiguous. As mentioned earlier, this contiguity is associated with the interplay between the two types of trefoils that prevents their knotting regions from contiguity, thereby forcing the surprising connection between the unknot and the connected sum, an unknotting number two knot. An even more complex example of this phenomenon is exhibited by the knotting fingerprint of the connected sum $3_1\#5_2$ and its associated knotting graph (see Figure 14).

As measured by the Cheeger constant, the most complex knotting fingerprints for knots through nine crossings belong to 8_{10} and 8_{20}, whose knotting graphs are shown in Figure 15.

5.2 Independent Path Complexity

Since the numerator of the Cheeger constant is the number of edges in an edge-cut set separating the global knot from the unknot, this numerator gives an upper bound on the number of edge independent paths connecting the unknot

15
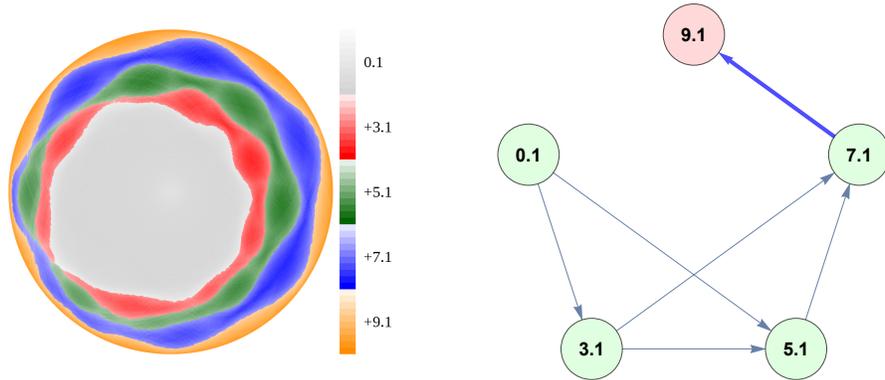
Figure 13: The ideal 9_1 knotting fingerprint and associated knotting graph. The two different vertex colors identify the separation defining the Cheeger constant, where the bold edge connects the separated regions.
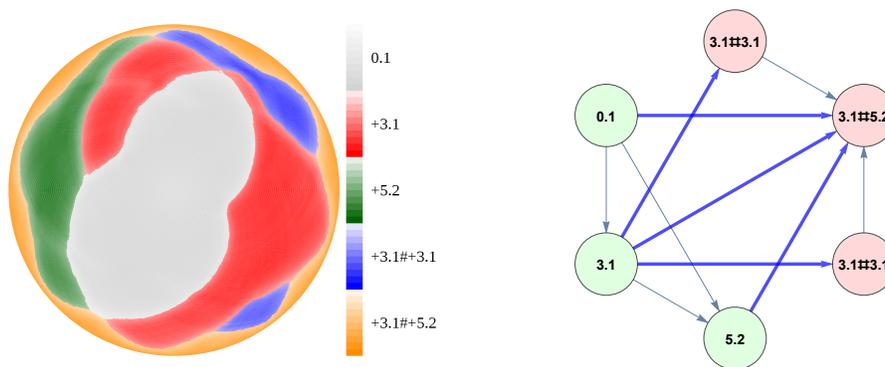
Figure 14: The ideal $3_1\#5_2$ knotting fingerprint and associated knotting graph. The two different vertex colors identify the separation defining the Cheeger constant, where the bold edges are those connecting the separated regions.

region to the global knot region. From the tables, one sees that EVR is often smaller than this numerator. Such is the case, for example, for the knot 7_2 in Table 1. If one employs $EVR(K)$ as a measure of complexity instead of the Cheeger constant, one finds that the $(2, 2n + 1)$-torus knots are identified as the simplest structures, along with 8_{19} and 9_{35}. The most complex, having 11 independent pathways, are 8_{16}, 9_{29}, and 9_{32}. The latter knots suggest that the independent pathway measure may capture a distinctly new dimension of knot complexity.

While there is only one minimal path taking the trefoil knot 3_1 to the unknot 0_1, the story is more complex for other $(2, 2n + 1)$-torus knots. For 5_1, an unknotting number two knot, if each edge implies a change in unknotting number of at most one, the shortest paths must have length two. However, these paths are no longer unique, as one may add any single strand passage resulting in an unknotting number one knot, creating another shortest path. Employing `TopoIce-X` within the KnotPlot software [39], we find that, in addition to $5_1 \to 3_1 \to 0_1$, one must also consider $5_1 \to 5_2 \to 0_1$, $5_1 \to 8_7 \to 0_1$, and $5_1 \to 9_{26} \to 0_1$, staying within the class of knots of crossing number no larger than 10. In the knotting fingerprint for 5_1, only the first unknotting pathway is observed. For 7_1, the situation is much more complex. In addition to the $7_1 \to 5_1 \to 3_1 \to 0_1$ pathway, the three other previous pathways occur. Adding even more pathways are those knots starting with $7_1 \to 7_3$, $7_1 \to 7_5$, and $7_1 \to 10_5$ since each of these is an unknotting number two knot with their own collections of unknotting pathways. Again, only the first occurs for the ideal 7_1. The constraint that the knotting pathway be supported by knotted subsegments of the ideal knot effectively limits the knotting pathway options to the "standard $(2, 2n+1)$-torus knot" pathway. We note that the knotting graph complexity that may arise for the larger $(2, 2n + 1)$-torus knots could increase the number of paths; for example 9_1 would have two (intersecting) pathways.

The twist knots provide another simple but interesting class to consider. After 3_1 and 4_1, the first of these is 5_2, another unknotting number one knot. Thus, its shortest path is $5_2 \to 0_1$ but, as 3_1 is a subknot of 5_2 there is a second, independent unknotting pathway $5_2 \to 3_1 \to 0_1$ within its knotting fingerprint. Furthermore, there are two disjoint 3_1 components giving rise to a second, independent $5_2 \to 3_1 \to 0_1$ unknotting pathway. As a consequence, both the Cheeger constant, $3/2$, and EVR, 3, provide a finer measure of the real structural complexity of the 5_2 knot than, for example, the unknotting number or the genus of the knot. This same complex structure is exhibited in the knotting fingerprints of 6_1, 7_2, 8_1, and 9_2, other twist knots with fewer than 10 crossings.

What can the number of independent knotting pathways tell us about more complex knots, e.g. 8_{20} shown in Figure 15? The Cheeger constant is $7/3$ and EVR is 5. It is an unknotting number one pretzel knot whose knotting fingerprint contains $\pm 3_1$, $\pm 5_2$ and $3_1 \# -3_1$ supporting five independent knotting pathways: $8_{20} \to 3_1 \to 0_1$, $8_{20} \to 3.1\#-3_1 \to -3_1 \to 0_1$, $8_{20} \to 5_2 \to 0_1$, $8_{20} \to -5_2 \to 0_1$, and $8_{20} \to 0_1$. Note that 8_{20} is an unknotting number one knot that contains an unknotting number two subknot, $3.1\# - 3.1$.

17
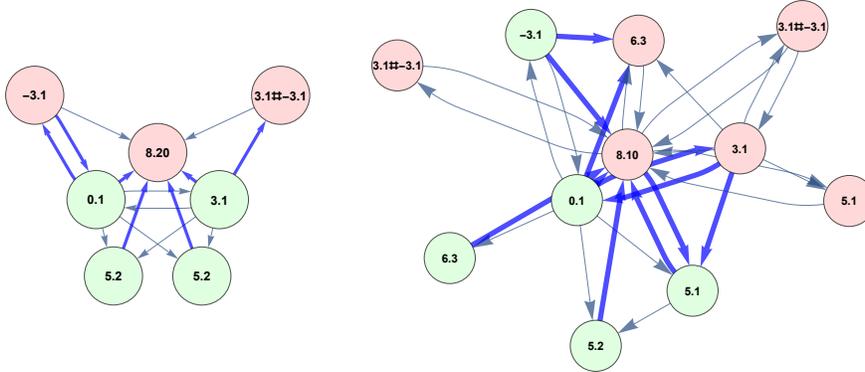
Figure 15: The largest Cheeger constant for knots through 9 crossings is given by 8_{20} with $h(8_{20}) = \frac{7}{3} \approx 2.33$. The knot 8_{10} has $h(8_{10}) = \frac{11}{5} = 2.2$. The two different vertex colors identify the separation defining the Cheeger constant, where the bold edges are those connecting the separated regions.

Another interesting knot is 8_{10} (see Figure 15), an unknotting number two knot whose Cheeger constant is $11/5$ and EVR is 7. The seven independent paths provide a substantial degree of complexity, although less than the Cheeger separation set of 11 edges. We note that 8_{10} also contains a composite subknot, two copies of $3_1\#-3_1$, which has unknotting number two, as does 8_{10}.

6   Discussion and Conclusions

In this paper we have presented the knotting fingerprint of a polygonal approximation of an ideal, or tight, knot, showing the structure of the knotting of subsegments of the knot. The associated structure of subknot types defines regions of the knotting fingerprint, i.e. a planar map, to which one can associate a planar graph with two distinguished vertices corresponding to the unknot and the global knot. We have proposed that the complexity of the knotting fingerprint and the associated knotting graph provides a new measure of the intrinsic complexity of the knot. Interested in the ways in which knots can be unknotted or, inversely, constructed from unknotted segments, we have proposed strategies by which one can quantitatively measure this complexity. The first strategy is analogous to the Cheeger constant, $h(K)$, of the graph whereby we partition the vertices of the graph (requiring the unknot to be a member of one subset and the global knot to be a member of the other subset) and take the minimum ratio of the number of edges connecting the two subsets and the number of vertices in the smaller of the two subsets over all such partitions. The second method, the edge robustness index $ER(K)$, is defined to be the number of edge independent paths in the graph connecting the unknot vertex to the global knot vertex. The third method, the edge vertex robustness index $EVR(K)$, is defined to be

Table 1: Analysis of Ideal Prime Knots (through 7 Crossings)

Knot	$h(K)$	$ER(K)$	$EVR(K)$
3_1	1/1	1	1
4_1	1/1	1	1
5_1	1/1	1	1
5_2	3/2	3	3
6_1	3/2	3	3
6_2	5/3	5	5
6_3	5/3	5	5
7_1	1/2	1	1
7_2	4/3	3	3
7_3	5/4	4	4
7_4	7/4	5	5
7_5	7/4	5	5
7_6	8/4	7	7
7_7	6/3	6	6

Table 2: Analysis of Ideal Prime Knots (8 Crossing Knots)

Knot	$h(K)$	$ER(K)$	$EVR(K)$
8_1	4/3	3	3
8_2	7/5	5	5
8_3	8/5	7	7
8_4	9/6	7	7
8_5	6/5	4	4
8_6	9/5	9	9
8_7	8/5	6	5
8_8	9/5	8	8
8_9	10/6	8	8
8_{10}	11/5	7	7
8_{11}	9/6	9	9
8_{12}	6/4	5	5
8_{13}	9/5	8	8
8_{14}	9/5	8	8
8_{15}	11/6	6	6
8_{16}	12/6	11	11
8_{17}	10/5	9	9
8_{18}	8/5	8	8
8_{19}	1/1	1	1
8_{20}	7/3	5	5
8_{21}	6/3	4	4

Table 3: Analysis of Ideal Prime Knots (9 Crossing Knots)

Knot	$h(K)$	$ER(K)$	$EVR(K)$		Knot	$h(K)$	$ER(K)$	$EVR(K)$
9_1	1/1	1	1		9_{26}	11/7	9	9
9_2	5/4	3	3		9_{27}	10/7	10	10
9_3	5/6	5	5		9_{28}	15/6	8	7
9_4	8/7	8	8		9_{29}	14/10	11	11
9_5	8/7	6	6		9_{30}	10/7	8	8
9_6	6/6	4	4		9_{31}	8/5	6	6
9_7	10/6	7	7		9_{32}	11/6	11	11
9_8	10/6	9	9		9_{33}	11/6	8	8
9_9	5/6	3	3		9_{34}	9/6	9	9
9_{10}	8/8	7	7		9_{35}	1/1	1	1
9_{11}	10/7	9	7		9_{36}	7/6	4	4
9_{12}	9/7	8	8		9_{37}	8/7	7	7
9_{13}	9/9	7	5		9_{38}	12/9	4	4
9_{14}	9/6	8	8		9_{39}	10/8	9	9
9_{15}	10/6	10	9		9_{40}	7/5	7	7
9_{16}	9/7	2	2		9_{41}	8/7	7	7
9_{17}	13/7	10	10		9_{42}	5/3	4	4
9_{18}	9/8	8	6		9_{43}	6/4	5	5
9_{19}	9/5	8	8		9_{44}	6/3	5	5
9_{20}	11/8	10	10		9_{45}	6/4	5	5
9_{21}	13/7	9	9		9_{46}	7/4	5	5
9_{22}	12/7	9	9		9_{47}	7/4	7	7
9_{23}	10/7	8	8		9_{48}	11/6	8	8
9_{24}	10/7	8	8		9_{49}	8/6	6	6
9_{25}	8/7	5	5					

Table 4: Analysis of Ideal Prime Knots (Selected 10 Crossing Knots)

Knot	$h(K)$	$ER(K)$	$EVR(K)$
10_1	5/4	3	3
10_2	11/7	6	4
10_3	9/8	5	5
10_4	18/9	10	10
10_5	9/8	7	7
10_{10}	10/8	7	7
10_{11}	7/9	6	6
10_{20}	8/7	6	6
10_{35}	5/6	4	4
10_{36}	9/7	7	6
10_{58}	4/3	2	2
10_{60}	10/7	8	8
10_{70}	8/8	5	4
10_{120}	2/1	1	1
10_{125}	5/5	3	3
10_{126}	6/6	5	5
10_{127}	7/6	5	5
10_{128}	5/4	2	2
10_{130}	9/6	9	7
10_{131}	8/6	7	7
10_{134}	7/5	2	2
10_{135}	9/6	5	5
10_{137}	8/4	7	7
10_{140}	9/6	8	7
10_{141}	7/4	4	4
10_{146}	7/5	6	6
10_{147}	7/6	6	6
10_{151}	13/6	6	6
10_{161}	16/8	6	4
10_{162}	10/8	9	9

Table 5: Analysis of Ideal Composite Knots (Selected Through 8 crossings)

Knot	$h(K)$	$ER(K)$	$EVR(K)$
$3_1\#3_1$	1/2	1	1
$3_1\#4_1$	3/2	2	2
$3_1\#5_1$	4/3	3	2
$3_1\#5_2$	5/3	2	2
$3_1\#-3_1$	3/2	3	3
$3_1\#-5_1$	3/3	3	2
$3_1\#-5_2$	6/4	4	4

the number of edge and vertex independent paths in the graph connecting the unknot vertex to the global knot vertex.

Our analysis of prime knots through 9 crossings, a sampling of 10 crossing knots, and a sampling of composite knots demonstrates that these measures are interesting tools for assessing the complexity of an ideal knot. The analysis further identifies instances of complex evolutionary structure, such as contiguous knotting regions representing knot types separated by more than one crossing change (as reflected in their unknotting numbers). While the $(2, 2n + 1)$-torus knots are clearly those of simplest structure, our analysis calls attention to the apparently simple structure of some other knots, e.g. 8_{19} and 9_{35}, among the knots with fewer than 10 crossings. Furthermore, we have provided a small sample of examples that demonstrate that the knotting pathways arising within ideal knots come from a quite specific set of options when compared with the shortest knotting pathways available, without being constrained to being supported within the ideal knot structure.

One may consider these quantities as defining knot invariants and, in this case, ask how they might be related to known knot invariants. We have sought correlations between the unknotting number, genus, braid index, super bridge number, Thurston-Benniquin number, average crossing number of ideal configurations, and determinant, see [13], and our values for the prime knots through nine crossings. These knot invariants were selected for their diversity and intrinsic three-dimensional nature. The conclusion of this facet of our investigation is simply the lack of apparent correlation between our measures and these knot invariants, leading one to ask whether they truly represent new three-dimensional characteristics of a prime knot or if there is a more complex relationship to be identified. To illustrate our analysis, we show several scatter plots arising from these data in Figures 16, 17, and 18. These scatter plots represent the most hopeful evidence in favor of correlation. To measure the extent of correlation between these data, we have employed the *distance correlation* [45]. A distance correlation of 0 implies independence of the data with stronger relationships, including nonlinear ones, being implied by larger values. In the end, these suggest that the Cheeger constant and the number of independent paths provide quantities that capture new dimensions of the structure of prime knots. The scatter plot in Figure 18 shows that our newly defined quantities are, among themselves, independent measures worthy of further exploration.

7   Acknowledgements
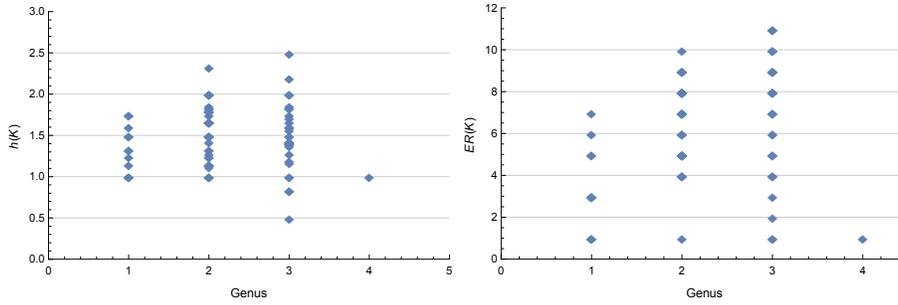
Figure 16: Scatter plots of the Cheeger constant and edge robustness index versus the genus for knots through nine crossings. The distance correlations are, respectively, 0.157 and 0.327.
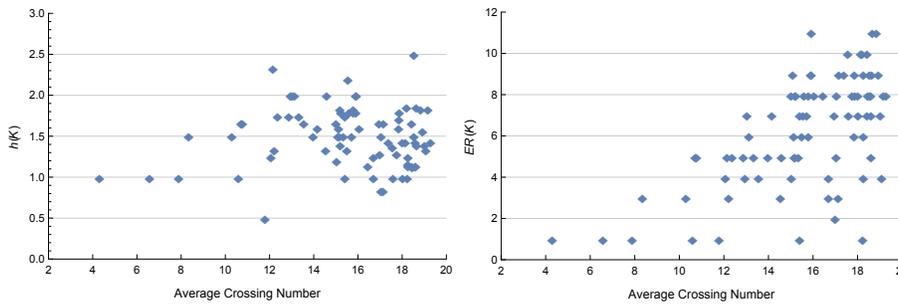
Figure 17: Scatter plots of the Cheeger constant and edge robustness index versus the average crossing number of the 32-edge ideal knot configuration [32] for knots through nine crossings. The distance correlations are, respectively, 0.261 and 0.569.
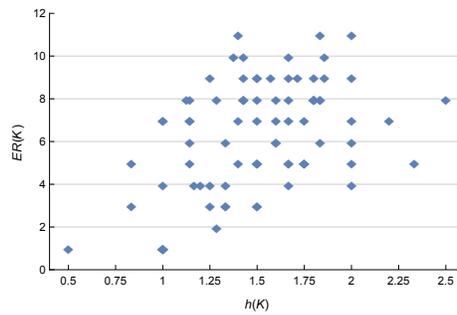
Figure 18: Scatter plot of the edge robustness index versus the Cheeger constant of the knot for knots through nine crossings. The distance correlation is 0.451.

References

[1] Greg Aloupis, Günter Ewald, and Godfried Toussaint. More classes of stuck unknotted hexagons. *Beiträge Algebra Geom.*, 45(2):429–434, 2004.

[2] Ted Ashton, Jason Cantarella, Michael Piatek, and Eric J. Rawdon. Knot tightening by constrained gradient descent. *Experiment. Math.*, 20(1):57–90, 2011.

[3] D. Bolinger, J. I. Sułkowska, H. P. Hsu, L. A. Mirny, M. Kardar, J. N. Onuchic, and P. Virnau. A Stevedore's protein knot. *PLoS Comput. Biol.*, 6(4):e1000731, Apr 2010.

[4] Gregory Buck and Jeremy Orloff. A simple energy function for knots. *Topology Appl.*, 61(3):205–214, 1995.

[5] Roman V. Buniy, Jason Cantarella, Thomas W. Kephart, and Eric J. Rawdon. The spectrum of tightly knotted flux tubes in QCD. *Journal of Physics: Conference Series*, 544(1):012025, 2014.

[6] Roman V. Buniy, Jason Cantarella, Thomas W. Kephart, and Eric J. Rawdon. The tight knot spectrum in QCD. *Phys. Rev. D*, 89:054513, 2014.

[7] Roman V. Buniy and Thomas W. Kephart. A model of glueballs. *Phys. Lett.*, B576:127–134, 2003.

[8] Roman V. Buniy and Thomas W. Kephart. Glueballs and the universal energy spectrum of tight knots and links. *Int. J. Mod. Phys. A*, 20:1252–1259, 2005.

[9] Jason Cantarella and Heather Johnston. Nontrivial embeddings of polygonal intervals and unknots in 3-space. *J. Knot Theory Ramifications*, 7(8):1027–1039, 1998.

[10] Jason Cantarella, Robert B. Kusner, and John M. Sullivan. On the minimum ropelength of knots and links. *Invent. Math.*, 150(2):257–286, 2002.

[11] Jason Cantarella, Al LaPointe, and Eric J. Rawdon. Shapes of tight composite knots. *J. Phys. A*, 45(22):225202, 2012.

[12] Corinne Cerf and Andrzej Stasiak. A topological invariant to predict the three-dimensional writhe of ideal configurations of knots and links. *Proc. Natl. Acad. Sci. USA*, 97(8):3795–3798, 2000.

[13] J. C. Cha and C. Livingston. Knotinfo: Table of knot invariants. http://www.indiana.edu/∼knotinfo, 2015.

[14] Jeff Cheeger. A lower bound for the smallest eigenvalue of the Laplacian. In *Problems in analysis (Papers dedicated to Salomon Bochner, 1969)*, pages 195–199. Princeton Univ. Press, Princeton, N. J., 1970.

[15] Fan R. K. Chung. *Spectral graph theory*, volume 92 of *CBMS Regional Conference Series in Mathematics*. Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 1997.

[16] Isabel Dazey Darcy and De Witt Sumners. A strand passage metric for topoisomerase action. In *KNOTS '96 (Tokyo)*, pages 267–278. World Sci. Publ., River Edge, NJ, 1997.

[17] Elizabeth Denne, Yuanan Diao, and John M. Sullivan. Quadrisecants give new lower bound for the ropelength of a knot. *Geom. Topol.*, 10:1–26, 2006.

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

[19] Alessandro Flammini and Andrzej Stasiak. Natural classification of knots. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 463(2078):569–582, 2007.

[20] P. Freyd, D. Yetter, J. Hoste, W. B. R. Lickorish, K.C. Millett, and A. Ocneanu. A new polynomial invariant of knots and links. *Bull. Amer. Math. Soc. (N.S.)*, 12(2):239–246, 1985.

[21] Oscar Gonzalez and John H. Maddocks. Global curvature, thickness, and the ideal shapes of knots. *Proc. Natl. Acad. Sci. USA*, 96(9):4769–4773, 1999.

[22] M. Jamroz, W. Niemyska, E.J. Rawdon, A. Stasiak, K.C. Millett, P. Sulkowski, and J.I. Sulkowska. Knotprot: a database of proteins with knots and slipknots. *Nucleic Acids Research*, 43(1):D306–D314, 2015.

[23] Vsevolod Katritch, Jan Bednar, Didier Michoud, Robert G. Scharein, Jacques Dubochet, and Andrzej Stasiak. Geometry and physics of knots. *Nature*, 384:142–145, 1996.

[24] Vsevolod Katritch, Wilma K. Olson, Piotr Pieranski, Jacques Dubochet, and Andrzej Stasiak. Properties of ideal composite knots. *Nature*, 388:148–151, 1997.

[25] N.P. King, E.O. Yeates, and T.O. Yeates. Identification of rare slipknots in proteins and their implications for stability and folding. *J. Mol. Biol.*, 373:153–166, 2007.

[26] Otto Krötenheerdt and Sigrid Veit. Zur Theorie massiver Knoten. *Wiss. Beitr. Martin-Luther-Univ. Halle-Wittenberg Reihe M Math.*, 7:61–74, 1976.

[27] W. B. R. Lickorish and Kenneth C. Millett. A polynomial invariant of oriented links. *Topology*, 26(1):107–141, 1987.

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

[29] Amos Maritan, Cristian Micheletti, Antonio Trovato, and Jayanth R. Banavar. Optimal shapes of compact strings. *Nature*, 406:287–290, 2000.

[30] Kenneth Millett, Akos Dobay, and Andrzej Stasiak. Linear random knots and their scaling behavior. *Macromolecules*, 38(2):601–606, 2005.

[31] Kenneth C. Millett. Knots, slipknots, and ephemeral knots in random walks and equilateral polygons. *J. Knot Theory Ramifications*, 19(5):601–615, 2010.

[32] Kenneth C. Millett and Eric J. Rawdon. Energy, ropelength, and other physical aspects of equilateral knots. *J. Comput. Phys.*, 186(2):426–456, 2003.

[33] Kenneth C. Millett and Benjamin M. Sheldon. Tying down open knots: A statistical method of identifying open knots with applications to proteins. In *Physical and numerical models in knot theory*, volume 36 of *Ser. Knots Everything*, pages 203–217. World Sci. Publishing, Singapore, 2005.

[34] Piotr Pieranski, Sandor Kasas, Giovanni Dietler, Jacques Dubochet, and Andrzej Stasiak. Localization of breakage points in knotted strings. *New J. Phys.*, 3(10):10.1–10.13, 2001.

[35] Sylwester Przybyl and Piotr Pieranski. High resolution portrait of the ideal trefoil knot. *Journal of Physics A: Mathematical and Theoretical*, 47(28):285201, 2014.

[36] Eric J. Rawdon, Kenneth C. Millett, and Andrzej Stasiak. Subknots in ideal knots, random knots, and knotted proteins. *Sci. Rep.*, 5:8928, 2015.

[37] Eric J. Rawdon, Kenneth C. Millett, Joanna I. Sulkowska, and Andrzej Stasiak. Knot localization in proteins. *Biochem. Soc. Trans.*, 41(2):538–541, 2013.

[38] Dale Rolfsen. *Knots and Links*. Publish or Perish, Inc., 1976.

[39] Robert G. Scharein. KnotPlot. http://www.knotplot.com, 1998. Program for drawing, visualizing, manipulating, and energy minimizing knots.

[40] I. H. Sloan and R. S. Womersley. Extremal systems of points and numerical integration on the sphere. *Advances in Computational Mathematics*, 21:102–125, 2004.

[41] A. Stasiak, V. Katritch, J. Bednar, D. Michoud, and J. Dubochet. Electrophoretic mobility of DNA knots. *Nature*, 384:122, 1996.

[42] Andrzej Stasiak, Jacques Dubochet, Vsevolod Katritch, and Piotr Pieranski. Ideal knots and their relation to the physics of real knots. In *Ideal knots*, pages 1–19. World Sci. Publishing, Singapore, 1998.

[43] Andrzej Stasiak, Vsevolod Katritch, and Louis H. Kauffman, editors. *Ideal knots*. World Sci. Publishing, Singapore, 1998.

[44] Joanna I. Sułkowska, Eric J. Rawdon, Kenneth C. Millett, Jose N. Onuchic, and Andrzej Stasiak. Conservation of complex knotting and slipknotting patterns in proteins. *Proc. Natl. Acad. Sci. USA*, 109(26):E1715–E1723, 2012.

[45] G.J. Szekely, M.L. Rizzo, and N.K. Bakirov. Measuring and testing dependence by correlation of distances. *Annals of Statistics*, 35(6):2769–2794, 2007.

[46] William R. Taylor. A deeply knotted protein and how it might fold. *Nature*, 406:916–919, 2000.

[47] Morwen B. Thistlethwaite. Kauffman's polynomial and alternating links. *Topology*, 27(3):311–318, 1988.

[48] Godfried Toussaint. A new class of stuck unknots in Pol_6. *Beiträge Algebra Geom.*, 42(2):301–306, 2001.

[49] Cedric Weber, Mathias Carlen, Giovanni Dietler, Eric J. Rawdon, and Andrzej Stasiak. Sedimentation of macroscopic rigid knots and its relation to gel electrophoretic mobility of DNA knots. *Sci. Rep.*, 3:1091, 2013.

[50] Robin J. Wilson. *Introduction to graph theory*. Longman, Harlow, 1996. Fourth edition.