CHAPTER 14

UNIVERSAL CHARACTERISTICS OF POLYGONAL KNOT PROBABILITIES

Kenneth C. Millett

Department of Mathematics,
University of California, Santa Barbara
Santa Barbara, CA 93016, USA
Email: millett@math.ucsb.edu
Homepage: http://www.math.ucsb.edu/˜millett/KM.html

Eric J. Rawdon

Department of Mathematics and Computer Science,
Duquesne University
Pittsburgh, PA 15282, USA
Email: rawdon@mathcs.duq.edu
Homepage: http://www.mathcs.duq.edu/˜rawdon

There is a striking qualitative similarity among the graphs of the relative probabilities of corresponding knot types across a wide range of random polygon models. In many cases one has theoretical results describing the asymptotic decay of these knot probabilities but, in the finite range, there is little theoretical knowledge and a variety of functional models have been used to fit the observed structures. In this paper we compare a selection of these models and study the extent to which each provides a successful fit for five distinct random knot models. One consequence of this study is that while such models are quite successful in this finite range, they do not provide the theoretically predicted asymptotic structure. A second result is the observed similarity between the global knot probabilities and those arising from small perturbations of three ideal knots.

1. Introduction

The probability that a random curve in 3-space is unknotted as a function of its length has long been the subject of experimentation, conjecture, and

mathematical analysis in the context of lattice polygons, random 3-space polygons, and smooth curves. The graphs of these probability functions, independent of the context, have many fundamental similarities, only some of which are often captured in proposed equations to calibrate the dependence of the knot probability on the length of the curve. For example, Figs. 1-5 show the graphs of length to probability for five different knot models. In view of the fundamental similarity between the overall shape of these and other graphs, we wish to capture the key features of the functional dependence whose universal nature is illustrated in these graphs and which is the subject of this paper. One purpose is to explore a variety of proposed functional models seeking a single one that has the property of fitting the data at "finite scales" as well as in the "asymptotic scales". In addition, we seek a single functional model that works for unknots as well as any non-trivial knot type.

For example, we propose that in the trivial knot probability data, the largest number of edges for which only trivial knots occur, within a given physical model, plays a role analogous to the number of edges for which a non-trivial knot attains its maximal probability. The objective of determining the features of the model from finite scale data has potential experimental importance due to the computational complexity associated with the generation and precise analysis of the knot data in the asymptotic regime.

We will first describe the various models for knots we consider in this paper, geometric, equilateral, and perturbations of ideal trivial, trefoil and figure-eight knots. Next we describe the Monte Carlo exploration of these knot ensembles or knot spaces and review the methods of analysis that are employed. We will then give a qualitative description of the knot probability data with attention to the critical features of the data. Many functional models have been proposed for such data, some which are intended to apply only in certain regimes. The data considered here principally concerns the "finite range", that is for numbers of edges in the polygonal models which are less than those necessary for the known asymptotic exponential decay to account for the data values. We will describe some of these functional models and compare the application of these models to our data sets in the finite scales and their relationship to expected asymptotic behavior. We propose two simple models that appear to be sufficiently robust, capturing the key features of the data for both trivial and non-trivial knots provided by our models in this finite range. These finite scales are the regime of many important scientific applications and, as a consequence, these model func-

tions may capture important aspects of their structure. A clear distinction between the finite and asymptotic ranges is visible in the structure of the best fit model functions.

2. Models of Knotting

A geometric polygonal knot K is defined by an ordered set of *vertices* $\{v_1, ..., v_n\}$ in $\mathbb{R}^3$ which are cyclically connected by straight line segments, called the *edges* of K. The set of all such knots is denoted $Geo(n)$. The subset of $Geo(n)$, where one requires that the length of each edge is equal to one, is called the space of equilateral knots, denoted $Equ(n)$. These spaces are considered as subsets of $3n$-dimensional Euclidean space and, as a consequence, inherit a natural distance function between two knots given by the maximum distance between the corresponding vertex coordinates. A third set of knots that we will consider in this paper consists of those knots in $Geo(n)$ each of whose vertices are, respectively, within a given distance r of the corresponding vertex of a knot K in $Equ(n)$. These sets are denoted $N(K, r)$. In particular, we will analyze the sets $N(K, r)$ where K is ropelength-minimized and r is the value of the thickness radius for K.

The thickness radius $r(K)$ of a polygonal knot employed in this paper is that defined by Rawdon in Refs. [22,23,24] which corresponds to the radius determined by polygonal analogs of the radius of curvature and the doubly-critical self-distance for smooth knots [15]. It provides an embedded solid torus neighborhood, whose radius we call the *thickness of* K, having K as its central curve. The *ropelength* is the ratio of the length of the polygon to the thickness radius. Given a fixed number of vertices and topological knot type that can be realized with the given number of edges, one can seek the minimum ropelength configuration of this type. These thickest knots have also been called *ideal knots* [13,18,19,25]. They are useful representatives for models as they provide spatial information that model physical phenomena. The family of knots created by a perturbation of the vertices of such an ideal knot K bounded by the radius of the thickest tube $r(K)$ determines the subset of $Geo(n)$ which we denote by $N(K, r(K))$. In particular, we will be analyzing $N(K, r(K))$ where K is a ropelength-minimized unknot, trefoil, or figure-eight knot. We call these spaces $Tube(0_1)$, $Tube(3_1)$, and $Tube(4_1)$, respectively, and suppress the number of edges n in the notation. Note that every polygon in $N(K, r(K))$ lies within the thick solid torus of radius $r(K)$ surrounding K and therefore are satellites of K.

3. Generation and Analysis of Knot Probability Data

The estimation of the relative proportion of a subset of a knot space, for example $Geo(n)$, consisting of a given knot type requires the uniform generation or sampling of the members of the entire space. Then the occurance of the knot type can be identified and the number of these in the sample gives an estimation of their frequency in $Geo(n)$. In this section we will describe how this sampling was accomplished for each of the knot spaces in which we are interested and, in addition, describe how the analysis of the knot type was achieved.

We define $Geo(n)$ to be the space of knots determined by the ordered set of n vectors uniformly distributed within the unit ball. These vectors are the vertices of the random knot. We propose this definition because the true set of geometric polygonal knots in 3-space, even those rooted at the origin, is a non-compact space. As a consequence, it is not possible to determine a proportion of the entire knot space, per se, determined by a given knot type. One must employ a secondary method. For the purposes of the present project, we have employed a spatial contraction that requires that the most distant vertex from the origin lie within the unit ball. The result is equivalent to a compactification of the knot space and allows one to estimate the knot probabilities in this knot space. We note, however, that this strategy to estimate geometric knot probabilities disrupts the spatial and physical structure of the knots and the average physical characteristics of the population. As a consequence, these knots may not provide a model for any biological or physical situation such as DNA knots or ring polymers in solution or melt. Rather, they are proposed as a vehicle to respond to the mathematical requirements of estimating the relative proportion of knots of various topological types.

The generation of equilateral knots requires a different strategy. Beginning with a regular polygon of n edges, we perform a sequence of random crankshaft rotations, which do not preserve the knot type, given by selecting two of the n vertices randomly, selecting one of the two arcs of segments of the polygon spanned by these vertices, and performing a random rotation of the segment about the axis determined by the two selected vertices. The number of pivot transformations, corresponding to steps in a random walk in $Equ(n)$, vary from 31,207,000, for the octagonal knots, to approximately 1,000,000 for some of the 150 estimates of knotting probabilities for n between 6 and 500.

While the regular equilateral polygon of n edges is the thickest unknot

of n edges, the corresponding question as to what is the thickest configuration for a given number of edges and knot type is still an open question. There are, however, excellent candidates for these thickest knots. Our "thick" knots were computed by TOROS [21], which uses simulated annealing to minimize the ropelength (i.e. the ratio of length to thickness radius) of polygonal knots. We cannot assume that the resulting knots are truly ideal due to the finite capabilities of computers and the nature of simulated annealing. Therefore, we will refer to these knots as *tight* or *thick*, although we do believe that they are excellent approximations of the ideal conformations. The perturbations of one of the resulting tight knots K is determined by adding to each vertex a perturbation vector that is uniformly distributed in a ball of radius $r(K)$.

Once generated, the data is encoded so as to facilitate the calculation of the HOMFLY [8] polynomial using the Ewing-Millett program [7]. This polynomial is one of the very useful, but not entirely faithful, means to characterize the knot populations [10,11]. We employ them as a surrogates for the knot type and, thereby, create knot population distributions which are, in fact, distributions of the associated HOMFLY polynomials. For the range of edge numbers and knots we analyze here, the polynomial is an effective tool for measuring the probabilities of the trivial, trefoil, and figure-eight knots. One feature, however, that affects the quality of the resulting data for large numbers of edges derives from the complexity of the resulting knot presentations. We were unable to analyze presentations with more than 999 crossings in our preprocessing knot simplification program. Also, for those with less than 1,000 crossings, we were unable to analyze knots whose presentations could not be simplified to fewer than 250 crossings. These are the current limitations of the HOMFLY program. We believe that these two factors could lead to a reduced accuracy in the probability calculations for equilateral knots with larger numbers of edges.

4. Knot Probabilities and Associated Functional Models

In our study of knot probabilities, we begin with geometric polygonal knots, whose probability distribution as a function of the number of edges is shown in Fig. 1. The data starts at five edges since six edges are required for non-trivial knotting to occur. The proportion of unknots monotonically decreases from one to, asymptotically, zero with a characteristically exponential decay observed for large numbers of edges, i.e. in the aysmptotic regime. We note that, for the unknot data, the graph is concave down im-

mediately following the number of edges required to create a non-trivial knot and that a change of concavity occurs later in the graph. The initial concavity resembles the behavior that is observed at a local maximum. This suggests the potential of a structure analogous to that observed at the point of maximal probability in non-trivial knot probability. Only for "large n" does one observe the characteristic exponential decay.

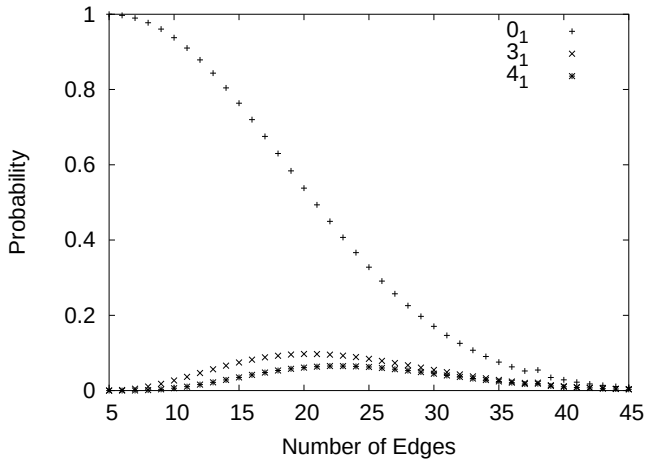

Fig. 1. Probability data for geometric polygonal knots.

For the trefoil and figure-eight knots, one observes a maximum and two changes in concavity: one between the initial data points and the point of maximum proportion and one following the maximum beyond which point the asymptotic range begins. The "asymptotic scale" dependence appears to be one of an exponential decay.

In Fig. 2, one can observe similarities and differences between data from equilateral polygonal knots and that of geometric polygonal knots. Figs. 3, 4, and 5 show the corresponding data for perturbations of the thick trivial, trefoil, and figure-eight knots respectively.

For $Equ(n)$ at larger numbers of edges (greater than 100), the data sets used to compute the proportions are sometimes relatively small, on the order of 1 million, compared to those employed for the smaller numbers of edges, on the order of 30 million for numbers of edges less than 20. At large numbers of edges, a generic projection has a large number of

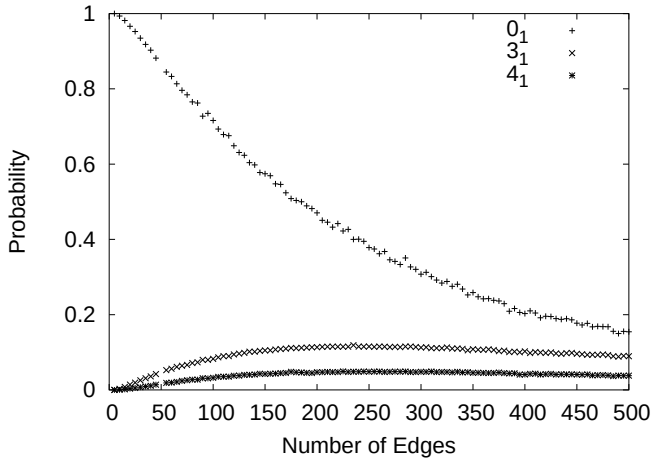

Fig. 2. Probability data for equilateral polygonal knots.

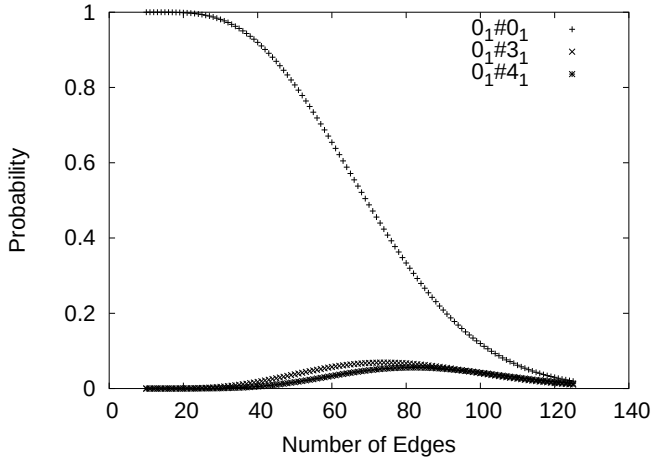

Fig. 3. Probability data for perturbations of the thick trivial knot.

crossings, often exceeding the current limitations of our program. While in theory the complexity of computing the HOMFLY polynomial increases exponentially on the number of crossings, for random knots there are many nugatory crossings which can be eliminated through careful preprocessing.

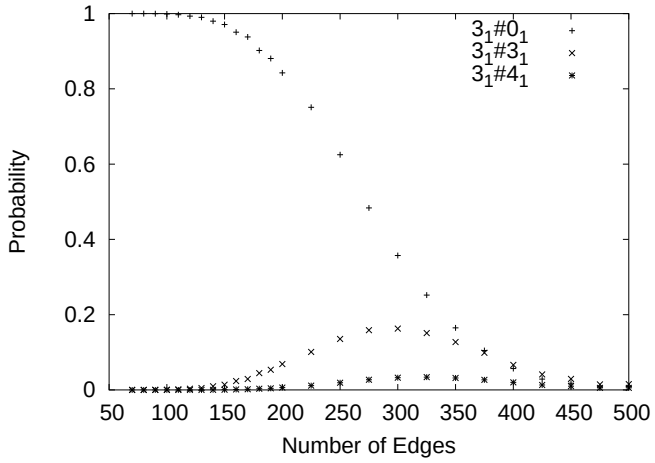

Fig. 4. Probability data for perturbations of the thick trefoil knot.

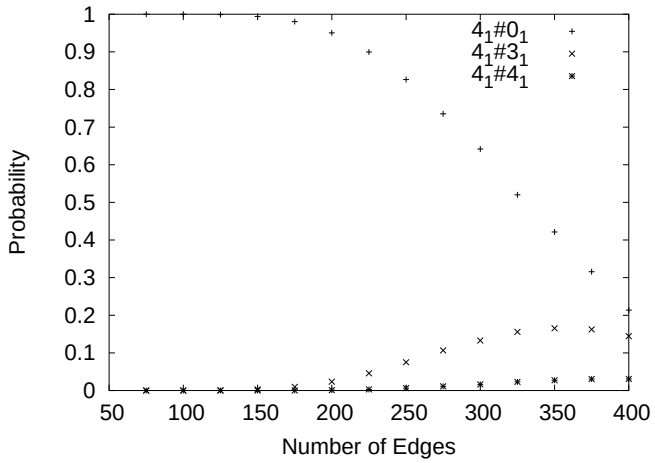

Fig. 5. Probability data for perturbations of the thick figure-eight knot.

Still, computing probabilities from a larger sample set for the larger numbers of edges in $Equ(n)$ was beyond our computational capacity [7].

For the probabilities within the thick trefoil tube, note that for $n < 200$, we have data points for every 10 edges. After $n = 200$, we have data points

separated by 25 edges. This may slightly bias the fit towards the "small n" regime, although it does not appear have created any anomalous behavior.

We fit these different data sets with several functional models which will be defined in the following sections. The fitting was done using a combination of two programs: `gnuplot` and random perturbation programs written in Maple and Mathematica. `Gnuplot` uses the Marquardt-Levenberg method to flow the fitting parameters to "better values". This is done to minimize the sum of the squared residuals (which we will refer to as the SSR), i.e. the sum of the squared distances between the data points and the function values at these points

$$SSR = \sum_{i=1}^{n} (f(x_i) - y_i)^2$$

where f is the fitting function and (x_i, y_i) are the data points. Like any sort of gradient-type descent, this algorithm is vulnerable to becoming stuck in local minima within the parameter space. Furthermore, the algorithms for nonlinear graph fitting can be extremely sensitive to the starting values for the parameters. We tried several different starting parameter values to minimize the impact of these effects. Also, we employed large-scale perturbations to confirm the Marquardt-Levenberg parameter values and to reduce the risk of these local minima. Furthermore, noting that the fitting could be improved by using the Maple or Mathematica perturbation programs, we used decreasing scales of perturbations to improve the final SSR values. The extent of improvement we saw varied with each case.

4.1. *The Exponential Decay Model (ED)*

Although somewhat similar in nature, apparently distinct functions have been employed as models for the kinds of data shown in these figures or, at least, certain aspects of such data. First, Hammersley [9] shows that the number of self-avoiding polygons grows as $e^{\kappa n + o(n)}$, where n is the number of edges. Sumners and Whittington [26] and Pippenger [20] show that the number of those that are unknotted grows as Eqn. (1) so that the probability of unknotted polygons decreases as $e^{-\alpha n + o(n)}$. Diao et al. extended this result to Gaussian random polygons [5]. This has been refined, numerically in Refs. [4,14,16], to propose that the probability of the unknot is given by Eqn. (1) with n_c being a model dependent characteristic length of random knotting and n_0 being the minimal number of edges required to construct

a non-trivial knot in this model. The exponential decay model is given by

$$P_0(n) = e^{-(n-n_0+1)/n_c} \ . \tag{1}$$

Because functions of this form are always decreasing and never equal to zero, this model will not achieve the goal of providing a single functional model that is valid for non-trivial as well as trivial knots. Furthermore, this functional model does not apply simultaneously to both the finite range and the asymptotic range as no change in concavity can occur for such functions. Parenthetically, we note that the location and size of the transition region between "finite range" and "asymptotic range" data appears uncertain. The data that serves as a basis for our study is consists principally in values within the "finite range". This plays a critical role in analysis of the fitting functions. As a consequence, we seek a functional model that addresses the change of concavity objective, an objective that appears to concern finite range behavior and not that of the asymptotic range. For an example of a study that might extend into the asymptotic range, we call attention to Moore, Lua, and Grosberg [17], who have reported an extensive analysis of data for knots from 15 to 3000 edges. Modulo large scale data collection and analysis concerns, their data fits this model very well at the larger scales. While not significant for their purposes, there is a less visible, but present, systematic departure from the model for small numbers of edges, i.e. in the finite range.

Fig. 6 shows a best fit of the probability that a knot in $Geo(n)$ is an unknot as a function of the number of edges for the exponential decay function model: $n_0 = 15.2901$, $n_c = 9.0695$, and the sum of the square residuals SSR is 0.3448.

4.2. *The Deguchi-Tsurusaki Model (DT)*

For non-trivial knots, Deguchi and Tsurusaki [1,2,3] have proposed Eqn. (2) where p_K is called the topological exponent of the knot, n_c is a variable to be fit, and n_K is the minimal number of edges required to create the knot. So long as p_K is greater than one, such functions satisfy the initial concavity requirement discussed earlier for non-trivial knots. They are, therefore, attractive functional model candidates with which to model this data. They also provide the desired asymptotic exponential decay and exhibit the changes in concavity observed in the data for both trivial and non-trivial knots. This model has been successfully used by Katritch et al. [12] (using a slightly different, yet equivalent, form) to model knotting

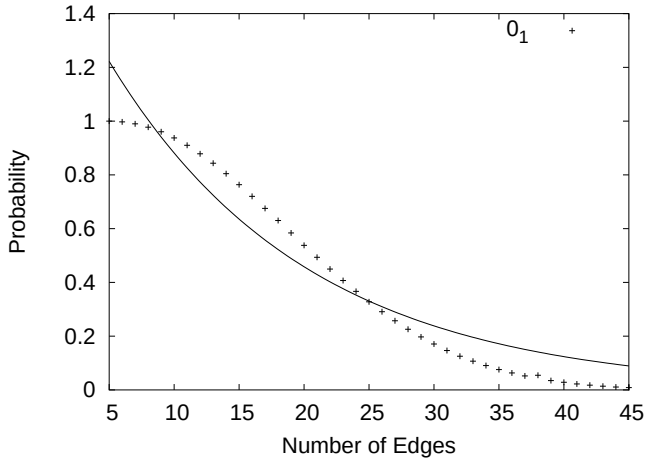

Fig. 6. Unknots in $Geo(n)$: fitting with the exponential decay model.

data beyond the finite range. The Deguchi-Tsurusaki model is given by

$$P_K(n) = C_K((n - n_K + 1)/n_c)^{p_K} e^{(-(n-n_K+1)/n_c)} \ . \tag{2}$$

While in fitting the data we did not require that p_K be greater than one, the fitting process often provides this property in the Deguchi-Tsurusaki model.

First, we ask if it is possible to model the trivial knot using the same functional model. Since the unknot has a different character than the non-trivial knots, we allow the n_K parameter to be a free parameter when fitting to the unknot data. Fig. 7 shows a best fit of the probability that a random knot in $Geo(n)$ is a trivial knot as a function of the number of edges for the Deguchi-Tsuruski model: $C_0 = 5.7752e - 31$, $n_0 = -53.4664$, $n_c = 2.0820$, $p_0 = 29.2866$, and $SSR = 0.002249$. While the Deguchi-Tsurusaki function fits this data quite well, there is a systematic deviation apparent in the graph. One might suspect that this is a consequence of the finite range constraint on the data set. An objective of the next models will be to provide a better fit, both numerically and visually, with the data.

4.3. *The Dobay et al. Model (DSDS)*

For non-trivial knots, Dobay, Sottas, Dubochet, and Stasiak [6] have proposed Eqn. (3) where p_K, called the topological exponent of the knot, n_c,

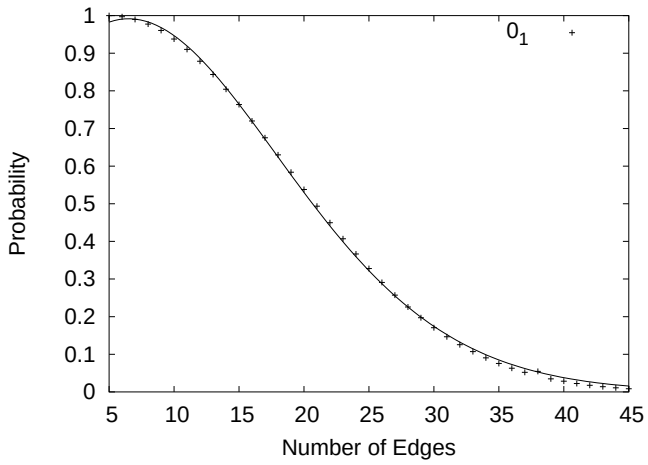

Fig. 7. Unknots in $Geo(n)$: fitting with the Deguchi-Tsurusaki model.

and q_K are variables to be fit, and n_K is the minimal number of edges required to create the knot. So long as p_K is greater than one and n_c and q_K are positive, such functions satisfy the initial concavity requirement discussed earlier for non-trivial knots and, therefore, are attractive functional model candidates with which to model this data. Dobay et al. [6] chose q_K to be an empirical constant equal to 0.18. The values of q_K arising from fitting our data do not suggest the existence of a single value that will work for all data. Functions of this nature do not provide the desired asymptotic exponential decay but do exhibit the changes in concavity observed in the data for both trivial and non-trivial knots. This model was proposed in the study of random walks in which the ends of the walk come in close vicinity (with distance less than or equal to two steps) to each other and thereby provides an opportunity for closure. The model is proposed to take into account the probability of closure in the course of a random walk. The Dobay et al. model is given by

$$P_K(n) = C_K(n - n_K + 1)^{p_K} e^{(-(n^{q_K})/n_c)} \ . \tag{3}$$

In fitting the data we did not require that p_K be greater than one although the fitting process often provides this property.

As earlier, we ask if it is possible to model the trivial knot using the same functional model. Since the unknot has a different character than the non-trivial knots, we allow the n_K parameter to be a free parameter

when fitting to the unknot data. This is also true in the models proposed in the next two sections when applied to the trivial knot. Fig. 8 shows a best fit of the probability that a random knot in $Geo(n)$ is a trivial knot as a function of the number of edges for the Dobay et al. model: $C_0 = 0.3780$, $n_0 = -11.8976$, $n_c = 774.6570$, $p_0 = 0.3519$, $q_0 = 2.1763$, and $SSR = 0.0001348$.

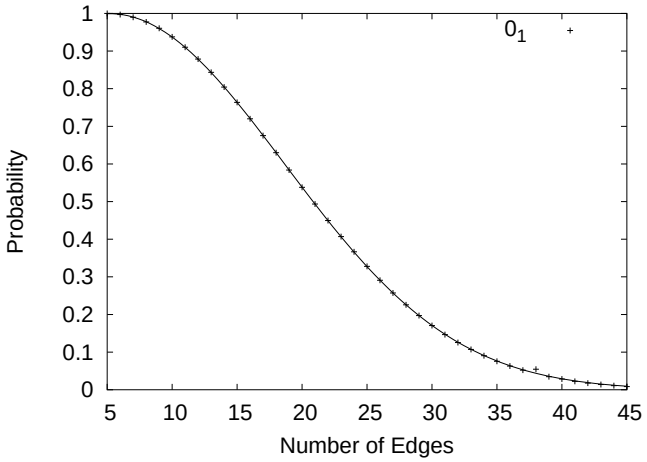

Fig. 8. Unknots in $Geo(n)$: fitting with the Dobay et al. model.

4.4. *The Quadratic Variation (QV)*

In the search for models that could better fit the observed finite range data, we use the Deguchi-Tsurusaki model as our point of departure by giving the function an equivalent but slightly different format and including quadratic factors to test the potential importance of second order effects in the exponential term in this range. The fact that theoretical results show that quadratic decay does not occur asymptotically does not preclude the presence of such effects in the finite range. The quadratic variation model is given by

$$P_K(n) = C_K((n - n_K + 1)^{p_K})e^{(-k_K n - l_K n^2)} . \tag{4}$$

As in the previous models n_K represents the number of edges required to create the knot type in the given model, except in the case of trivial knots when it is allowed to be a free parameter.

The remaining values are fitted to the data in an attempt to capture critical features of the data such as inflection points, the maximum, etc. For non-trivial knots, at $n = n_K - 1$, $P_K(n) = 0$, implying that p_K is greater than zero. In view of the initial concavity of the probability function data for non-trivial knots, one might also require that p_K be greater than one for non-trivial knots, although we have not done so here. For the unknot, we have allowed the values for n_0 and p_0 to be free parameters determined by the data. $P_0(n)$ is a monotonically decreasing function for $n \geq n_0$ with a single inflection point. Theoretical results, Sumners and Whittington [26] and Pippenger [20], show that the decrease is a linear exponential decay. To insure this, we require the parameters k_0 and l_0 to be positive. For non trivial knots, $P_K(n)$ is a monotonically decreasing function for n greater than the unique value for which $P_K(n)$ attains its maximum, and has a single inflection point. For the trivial geometric random knots, Fig. 9 shows the best fit for the quadratic variation model: $C_0 = 1.1614e - 51$, $n_0 = -53.4692$, $p_0 = 29.2908$, $k_0 = 0.4803$, $l_0 = 2.0825e - 40$, and $SSR = 0.002249$, roughly the same quality of fit achieved by the Deguchi-Tsurusaki function. The size of "l" may suggest an artifact of the fitting process rather than the presence of a significant second order term. A significant second order term would be inconsistent with theoretical results on the exponential decay and provide evidence of the domination of finite range data.

Further evidence that the data's dominant features reside in the finite range is clearly provided by allowing the fitting function to have fuller freedom in the exponent, for example by allowing k and/or l to have negative values. Let $QV\_$ denote this relaxed model. Allowing k and l to have negative values gives an order of magnitude improvement in the fit. This results in a rather different function having the parameter values: $C_0 = 0.9145614$, $n_0 = 4.0371$, $p_0 = 0.0001229$, $k_0 = -0.03179$, $l_0 = 0.002920$, and $SSR = 0.0001853$. This fit is of the same order of magnitude as that given by the Dobay et al. model where $SSR = 0.0001348$.

4.5. *The Full Variation Model (FV)*

Finally, we test for a second order variation in the power law factor by introducing an independent power term. The full variation model is given by

$$P_K(n) = C_K(n - n_K + 1)^{p_K}(1 + b_K(n - n_K + 1)^{q_K})e^{(-k_K n - l_K n^2)} . \quad (5)$$

For non-trivial knots, in order that P_K be defined at $n = n_K - 1$, we require that both p_K and q_K be greater than 0 and we invoke the same

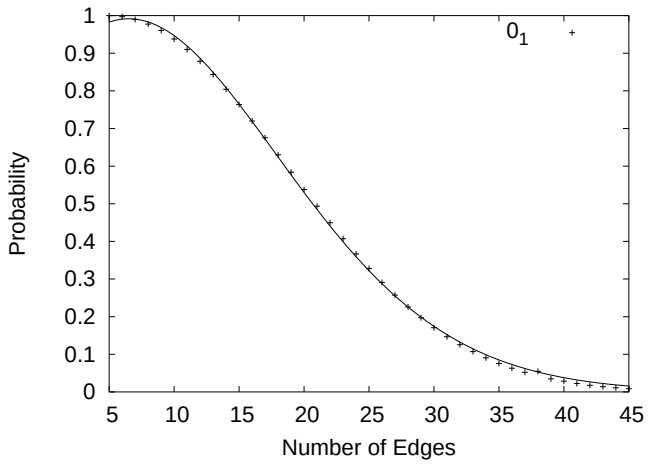

Fig. 9. Unknots in $Geo(n)$: fitting with the quadratic variation model.

constraint on the functional models of trivial knots.

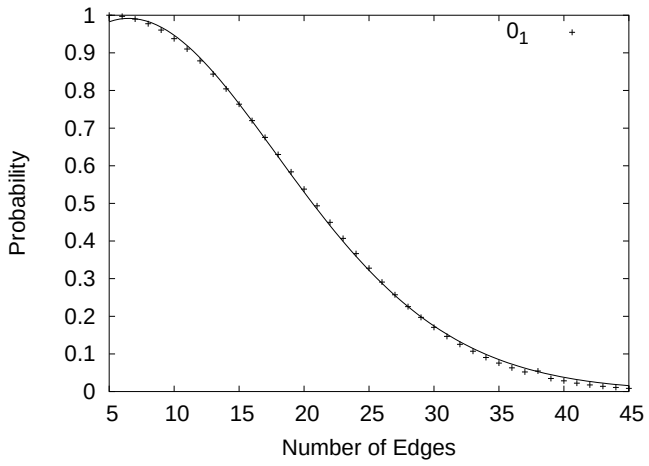

Fig. 10. Unknots in $Geo(n)$: fitting with the full variation model.

For the case of geometric random unknots, Fig. 10 shows the fit using the full variation model: $C_0 = 1.1600e - 51$, $n_0 = -53.4702$, $p_0 = 29.2910$,

$q_0 = 0.2263$, $b_0 = 0.000002700$, $k_0 = 0.4803$, $l_0 = 2.0823e - 40$, and $SSR = 0.002249$.

The evidence of the finite range character of the data appearing in the quadratic variation model is also present in the full variation. By allowing the fitting function to have negative k and l, we define the FV_{-} relaxed model. Fitting this model results in a rather different function having the parameter values: $C_0 = 0.9123$, $n_0 = 3.3827$, $p_0 = 0.00009720$, $q_0 = 2.2828$, $b_0 = 0.0002283$, $k_0 = -0.03465$, $l_0 = 0.003423$, and $SSR = 0.0001399$. This fit is essentially the same as that given by the Dobay et al. model where $SSR = 0.0001348$.

4.6. *Applications to Unknot Data*

In addition to geometric polygonal unknot data, we have applied these same functional models to the case of equilateral polygonal trivial knots and the vertex perturbations of the tight unknot, trefoil knot, and the figure-eight knots which are less than the radius of the thick tube surrounding the core knot. In the case of the perturbations within the tight knots, one always gets a satellite knot, usually in the form of a connected sum with the core knot. Thus, trivial knots within these emsembles correspond to perturbations which result in the same knot type as the core knot. The results of our analysis are similar to that for geometric trivial knots. Table 1 reports the parameters for each of the four functional models applied to the trivial knots in the five data sets.

For the geometric unknots, shown in Fig. 11, where the SSR values are plotted on a logarithmic scale, we see significant increases in the quality of fit as one goes from the exponential decay to the Deguchi-Tsurusaki and Dobay et al. models to the quadratic variation model followed by a less significant improvement in the full variation model. This is not uniformly the case with all the unknot data sets as one observes in Table 2. Here we see that all models give equivalent fits for the full range of data of the equilateral unknots while this is less the case for the geometric unknots, and the perturbations of the thick trivial knot, trefoil knot, and figure-eight knot.

Since the Dobay et al. and the QV_{-} and FV_{-} (in which the exponential term is not linear or the linear part has a positive coefficient) most often give the best fit to the data, one is lead to conclude that the nature of the finite range is quite different from that of the asymptotic range where theoretical results prove a linear exponential decay. This conclusion is supported by

Table 1. The fitting parameters for the unknot probability.

Probability of 0_1								
Model	C_K	n_K	n_c	p_K	k_K	l_K	b_K	q_K
				$Equ(n)$				
ED		1.31e+01	2.50e+02					
DT	1.07e+00	4.44e-10	2.44e+02	9.24e-03				
DSDS	9.51e-01	5.95e-01	1.17e+02	4.75e-02				8.92e-01
QV	1.06e+00	6.00e+00		4.30e-03	4.10e-03	4.31e-29		
*QV*_	1.04e+00	5.98e+00		4.46e-03	4.06e-03	4.47e-29		
FV	1.04e+00	5.98e+00		4.27e-03	4.06e-03	4.45e-29	1.00e-08	8.65e-09
*FV*_	1.04e+00	5.98e+00		4.27e-03	4.06e-03	4.45e-29	1.00e-08	8.65e-09
				$Geo(n)$				
ED		9.07e+00	1.53e+01					
DT	5.78e-31	-5.35e+01	2.08e+00	2.93e+01				
DSDS	3.78e-01	-1.19e+01	7.75e+02	3.52e-02				2.17e+00
QV	1.16e-51	-5.35e+00		2.93e+01	4.80e-01	2.08e-40		
*QV*_	9.15e-01	4.04e+00		1.23e-04	-3.18e-02	2.92e-03		
FV	1.16e-51	-5.35e+01		2.93e+01	4.80e-01	2.08e-40	2.65e-06	2.26e-01
*FV*_	9.12e-01	3.38e+00		9.72e-05	-3.47e-02	3.42e-03	2.28e-04	2.28e+00
				$Tube(0_1)$				
ED		2.51e+01	5.61e+01					
DT	2.97e-13	-1.05e+02	8.09e+00	1.62e+01				
DSDS	1.07e-33	-4.25e+01	8.10e-01	2.06e+01				6.77e-01
QV	1.16e-33	-1.06e+02		1.62e+01	1.23e-01	2.60e-06		
*QV*_	8.32e-01	9.61e-01		3.87e-04	-1,67e-02	3.51e-04		
FV	1.15e-33	-1.06e+02		1.62e+01	1.23e-01	2.64e-06	7.14e-05	2.32e-02
*FV*_	1.17e-01	8.49e-01		3.36e-04	-1,67e-02	3.51e-04	6.07e+00	3.93e-07
				$Tube(3_1)$				
ED		1.15e+02	2.01e+02					
DT	2.01e+00	-8.46e+00	6.40e+01	1.97e+00				
DSDS	6.89e-21	-2.08e+02	4.65e+01	8.54e+00				1.03e+00
QV	1.69e-03	3.07e-02		1.71e+00	1.47e-02	6.54e-09		
*QV*_	6.36-01	7.88e-01		3.33e-07	-7978e-03	3.34e-05		
FV	9.53e-02	4.25e-03		5.59e-01	7.69e-06	2.06e-05	5.69e-05	4.32e-02
*FV*_	4.16e-01	8.75e-07		2.39e-10	-7.28e-03	3.11e-05	6.022-01	1.51e-14
				$Tube(4_1)$				
ED		1.12e+02	3.79e+02					
DT	2.85e-07	-3.77e+02	4.59e+01	1.09e+01				
DSDS	1.28e-29	-8.96e+01	5.59e-01	1.59e+01				4.87e-01
QV	4.04e-29	-3.78e+02		1.10e+01	2.19e-02	3.54e-09		
*QV*_	7.89e-01	-6.84e+00		2.23e-06	-4,42e-03	1.76e-05		
FV	3.01e-29	-3.79e+02		1.10e+01	2.19e-02	3.59e-09	6.43e-09	2.60e-02
*FV*_	1.27e-01	9.94e+00		1.93e-04	-4.41e-03	1.76e-05	5.21e+00	3.87-04

the corresponding SSR values for the trefoil and figure-eight knots in Table 2.

Deguchi and Tsurusaki [4] propose that the parameter, p_K "should be universal: it is independent upon the models of random polygon(s) and is

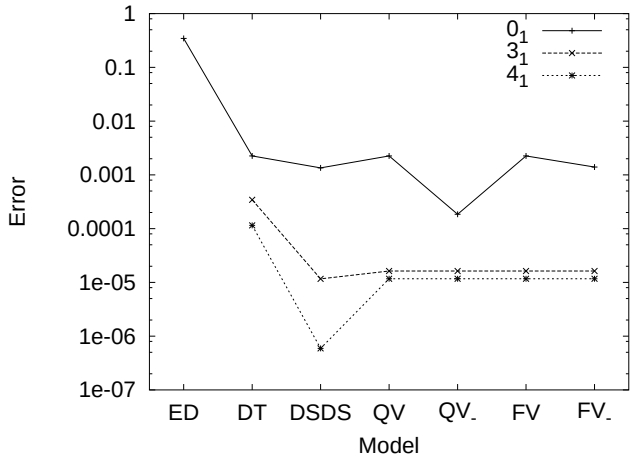

Fig. 11. Comparison of SSR values of the exponential decay, Deguchi-Tsurusaki, Dobay et al., quadratic variation, and full variation models for trivial, trefoil, and figure-eight knots in $Geo(n)$.

determined only by the knot K". While our data may concern models for random polygons outside the category they had in mind, we have explored the degree to which such universality holds. It appears that the different data sets require rather different values for p_K as shown in Fig. 12.

Comparison of p_K for the full variation model is shown in Fig. 13. Here also, one observes significant variation in the values of p_K for the four polygonal trivial knot models.

Tables 1, 3, and 4 report the parameter values for the different models and knot space data sets for the equilateral, geometric, and perturbations of the tight trivial knot, trefoil knot, and figure-eight knots. Table 2 gives the associated SSR values for each set of parameters.

5. Analysis of Functional Models of Non-trivial Knot Probability

With the exception of the exponential decay model, the functional models have been developed to study knotting probability of non-trivial knots, and therefore, one expects a better fit with the data will be possible. We will not include the exponential decay model in our analysis as it is of interest only in the asymptotic range. Furthermore, in this note, we explore the data for two nontrivial knot types, the trefoil knot and the figure-eight knot,

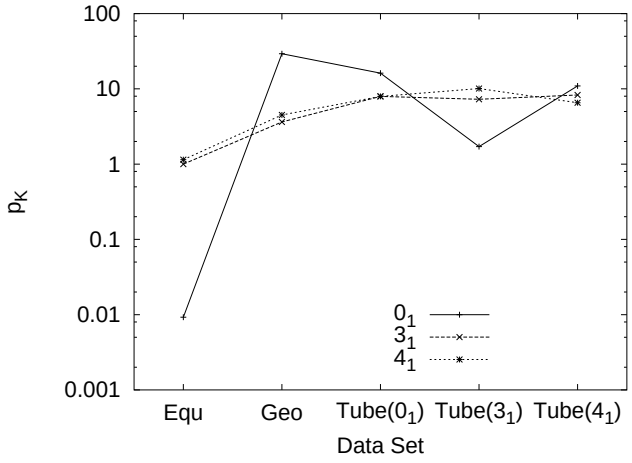

Fig. 12. Comparison of p_K of the Deguchi-Tsurusaki model for all knots and data sets.

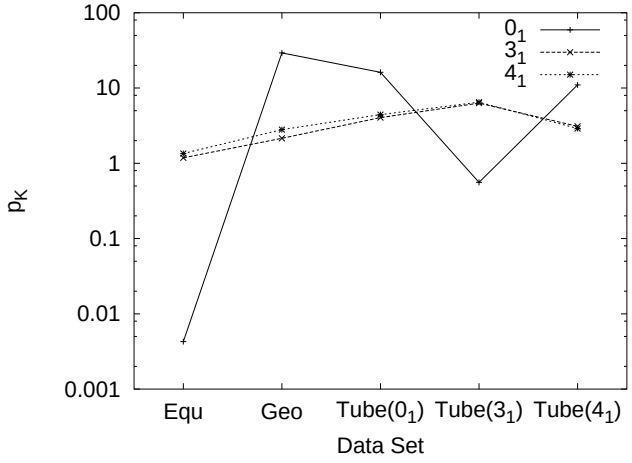

Fig. 13. Comparison of p_K of the full variation model for all knots and data sets.

as they occur in five polygonal knot contexts: equilateral, geometric and perturbations of the tight unknot, trefoil knot, and figure-eight knot.

Table 2. SSR values for all knots, data sets, and models.

	SSR values				
Model	$Equ(n)$	$Geo(n)$	$Tube(0_1)$	$Tube(3_1)$	$Tube(4_1)$
			0_1		
ED	1.2117e-02	3.4482e-01	2.0215e+00	6.1453e-01	2.5860e-01
DT	9.0422e-03	2.2490e-03	7.5266e-02	7.8516e-02	3.1228e-02
DSDS	5.1867e-03	1.3483e-03	2.4599e-02	3.6708e-02	1.8822e-02
QV	9.1813e-03	2.2488e-03	7.4494e-02	8.3402e-02	3.1181e-02
$QV\_$	4.9453e-03	1.8529e-04	2.1224e-02	2.5167e-02	1.6621e-02
FV	4.9450e-03	2.2488e-03	7.4469e-02	2.7480e-02	3.1149e-02
$FV\_$	4.9450e-03	1.3987e-03	2.1217e-02	1.2541e-02	1.6647e-02
			3_1		
DT	4.0408e-04	3.4432e-04	2.5416e-04	7.2692e-04	5.7876e-05
DSDS	1.4319e-04	1.1645e-05	1.1827e-06	1.2210e-04	3.0108e-05
QV	4.0422e-04	1.6298e-05	9.8789e-07	7.2692e-04	5.4966e-05
$QV\_$	4.0422e-04	1.6298e-05	9.8789e-07	1.1109e-04	1.1787e-05
FV	3.4764e-04	1.6298e-05	9.8788e-07	8.9632e-05	5.4966e-05
$FV\_$	3.4764e-04	1.6298e-05	9.8788e-07	8.9632e-05	1.1767e-05
			4_1		
DT	1.1416e-04	1.1504e-04	1.2672e-04	1.8650e-05	1.4368e-06
DSDS	1.8545e-05	5.9223e-07	2.9904e-06	9.7717e-06	5.3741e-07
QV	1.1750e-04	1.1693e-05	7.8559e-07	9.7804e-06	7.9669e-07
$QV\_$	1.1750e-04	1.1693e-05	4.3770e-07	9.7440e-06	5.1171e-07
FV	1.0769e-04	1.1692e-05	5.9955e-07	7.1188e-06	7.9141e-07
$FV\_$	1.0769e-04	1.1692e-05	4.3343e-07	7.1188e-06	5.1171e-07

5.1. *The Trefoil Knot*

As in the case of the equilateral unknots, the Deguchi-Tsurusaki, quadratic variation, and full variation functional models are essentially equally successful for trefoil knots. This is also the case for the perturbations of the tight figure-eight knot but not for the geometric trefoil knots nor for the perturbations of either the tight unknot or trefoil knot. The quadratic variation is essentially as good a model as the full variation and both are substantially better than the Deguchi-Tsurusaki model as shown in Fig. 14. Note that because the different data sets have different numbers of data points, one should compare the quality of the fitting models within a particular data set, not across the data sets. We include the connecting lines only to simplify the reading of the graph.

Comparison of p_{3_1} for the Deguchi-Tsurusaki model with that of the full variation model is shown in Figs. 12 and 13. Here also, one observes significant variation in the values of p_{3_1} for the five polygonal models.

Table 3. The fitting parameters for the trefoil probability.

Probability of 3_1								
Model	C_K	n_K	n_c	p_K	k_K	l_K	b_K	q_K
				$Equ(n)$				
DT	3.16e-01	6.00e+00	2.47e+02	9.99e-01				
DSDS	4.25e-04	6.00e+00	7.09e+00	1.55e+00				5.49e-01
QV	1.31e-03	6.00e+00		9.99e-01	4.04e-03	6.96e-10		
QV_-	1.31e-03	6.00e+00		9.99e-01	4.04e-03	6.96e-10		
FV	2.55e-03	6.00e+00		1.18e+00	3.43e-03	3.62e-12	-5.90e-01	6.43e-02
FV_-	2.55e-03	6.00e+00		1.18e+00	3.43e-03	3.62e-12	-5.90e-01	6.43e-02
				$Geo(n)$				
DT	3.44e-02	6.00e+00	4.05e+00	3.63e+00				
DSDS	1.30e-03	6.00e+00	2.25e+02	2.11e+00				1.92e+00
QV	1.36e-03	6.00e+00		2.15e+00	1.80e-02	2.97e-03		
QV_-	1.36e-03	6.00e+00		2.15e+00	1.80e-02	2.97e-03		
FV	1.36e-03	6.00e+00		2.15e+00	1.80e-02	2.97e-03	9.34e-16	1.11e-08
FV_-	1.36e-03	6.00e+00		2.15e+00	1.80e-02	2.97e-03	9.34e-16	1.11e-08
				$Tube(0_1)$				
DT	7.77e-05	1.10e+01	8.79e+00	7.09e+00				
DSDS	4.45e-08	1.10e+01	1.76e+03	4.00e+00				1.94e+00
QV	4.30e-08	1.10e+01		4.05e+00	5.53e-03	3.90e-04		
QV_-	4.30e-08	1.10e+01		4.05e+00	5.53e-03	3.90e-04		
FV	4.30e-08	1.10e+01		4.05e+00	5.53e-03	3.90e-04	9.34e-06	1.11e-02
FV_-	4.30e-08	1.10e+01		4.05e+00	5.53e-03	3.90e-04	9.34e-06	1.11e-02
				$Tube(3_1)$				
DT	1.28e-04	8.00e+01	2.84e+01	7.27e+00				
DSDS	3.55e-09	8.00e+01	7.97e+05	3.65e+00				2.51e+00
QV	5.73e-14	8.00e+01		7.27e+00	3.53e-02	1.22e-26		
QV_-	1.25e-08	8.00e+01		2.95e+00	-1.73e-02	5.27e-05		
FV	2.06e-11	8.00e+01		6.31e+00	4.24e-02	7.20e-06	5.28e-10	4.33e+00
FV_-	2.06e-11	8.00e+01		6.31e+00	4.24e-02	7.20e-06	5.28e-10	4.33e+00
				$Tube(4_1)$				
DT	1.54e-05	7.50e+01	3.38e+01	8.30e+00				
DSDS	6.53e-10	7.50e+01	5.96e+09	3.65e+00				3.86e+00
QV	7.01e-17	7.50e+01		8.09e+00	2.88e-02	1.16e-26		
QV_-	5.66e-10	7.50e+01		3.09e+00	-2.27e-02	4.75e-05		
FV	7.02e-17	7.50e+01		8.09e+00	2.88e-02	1.16e-26	9.37e-06	1.11e-02
FV_-	5.60-10	7.50e+01		3.10e+00	-2.26e-02	4.75e-05	9.99e-10 0	1.00e+00

5.2. *The Figure-Eight Knot*

As in the case of the equilateral unknots and the trefoil knot, the Deguchi-Tsurusaki, Dobay et al., quadratic variation, and full variation functional models are essentially equally successful for equilateral figure-eight knots. This is also the case for the perturbations of the tight figure-eight knot but not for the geometric trefoil knots nor for the perturbations of either the tight unknot or trefoil knot. For the geometric figure-eight knots and for

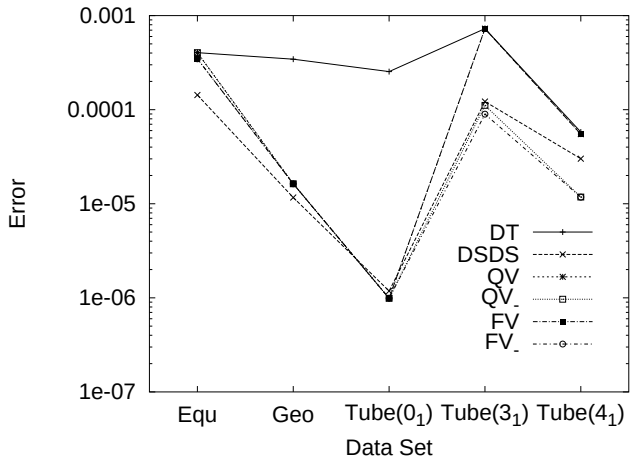

Fig. 14. Comparison of SSR values of the Deguchi-Tsurusaki, Dobay et al., quadratic variation, and full variation models for all trefoil knots.

the perturbations of the tight unknot and the trefoil knot, the quadratic variation is essentially as good a model as the full variation and both are substantially better than the Deguchi-Tsurusaki model as shown in Fig. 15.

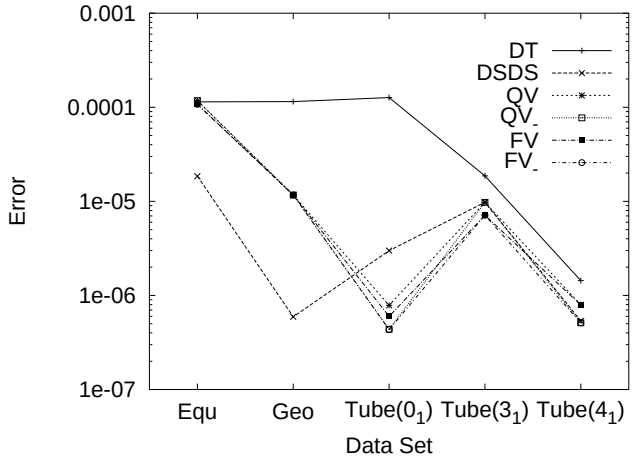

Fig. 15. Comparison of SSR values of the Deguchi-Tsurusaki, Dobay et al., quadratic variation, and full variation models for all figure-eight knots.

Table 4. The fitting parameters for the figure-eight probability.

Probability of 4_1								
Model	C_K	n_K	n_c	p_K	k_K	l_K	b_K	q_K
				$Equ(n)$				
DT	1.32e-01	7.00e+00	2.24e+02	1.15e+00				
DSDS	6.89e-05	7.00e+01	3.69e+00	1.92e+00				4.88e-01
QV	2.80e-04	7.00e+00		1.14e+00	4.43e-03	1.26e-10		
QV−	2.80e-04	7.00e+00		1.14e+00	4.43e-03	1.26e-10		
FV	1.28e-03	7.00e+00		1.35e+00	3.82e-03	2.04e-12	-8.27e-01	2.31e-02
FV−	1.28e-03	7.00e+00		1.35e+00	3.82e-03	2.04e-12	-8.27e-01	2.31e-02
				$Geo(n)$				
DT	6.37e-03	7.00e+00	3.54e+00	4.54e+00				
DSDS	2.08e-04	7.00e+01	2.14e+02	2.72e+00				1.93e+00
QV	2.34e-04	7.00e+00		2.80e+00	3.13e-02	3.02e-03		
QV−	2.34e-04	7.00e+00		2.80e+00	3.13e-02	3.02e-03		
FV	2.34e-04	7.00e+00		2.80e+00	3.13e-02	4.13e-04	4.79e-08	1.27e+00
FV−	2.34e-04	7.00e+00		2.80e+00	3.13e-02	4.13e-04	4.79e-08	1.27e+00
				$Tube(0_1)$				
DT	1.10e-05	1.60e+01	8.21e+00	7.95e+00				
DSDS	2.99e-06	1.60e+01	1.26e+03	4.78e+00				1.88e+00
QV	3.83e-09	1.60e+01		4.63e+00	2.93-03	4.07e-04		
QV−	5.82e-09	1.60e+01		4.43e+00	-4.26-03	4.31e-04		
FV	4.28e-09	1.60e+01		4.58e+00	1.01e-03	4.13e-04	8.31e-08	1.27e-03
FV−	5.73e-09	1.60e+01		4.44e+00	-3.97e-03	4.30e-04	9.97e-08	1.11e-03
				$Tube(3_1)$				
DT	6.06e-08	8.00e+01	2.36e+01	1.01e+01				
DSDS	9.77e-06	8.00e+01	1.95e+04	6.21e+00				1.96e+00
QV	4.32e-15	8.00e+01		6.17e+00	6.37e-04	3.85e-05		
QV−	1.45e-14	8.00e+01		5.79e+00	-3.25e-03	4.19e-05		
FV	2.43e-13	8.00e+01		6.49e+00	3.78e-02	1.06e-05	5.04e-10	4.49e+00
FV−	2.43e-13	8.00e+01		6.49e+00	3.78e-02	1.06e-05	5.04e-10	4.49e+00
				$Tube(4_1)$				
DT	1.01e-04	1.25e+02	4.27e+01	6.53e+00				
DSDS	5.37e-07	1.25e+02	2.41e+07	4.06e+00				2.96e+00
QV	1.84e-12	1.25e+02		4.84e+00	3.72e-04	2.20e-05		
QV−	1.47e-10	1.25e+02		2.90e+00	-2.60e-02	4.71e-05		
FV	1.93e-10	1.25e+02		4.82e+00	1.14e-04	2.22e-05	-7.55e-11	1.64e-07
FV−	1.47e-10	1.25e+02		290e+00	-2.60e-03	4.71e-05	-1.25e-11	1.06e-02

Comparison of p_{4_1} for the Deguchi-Tsurusaki model with that of the full variation model is shown in Figs. 12 and 13. Here also, one observes significant variation in the values of p_{4_1} for the five polygonal models.

6. Conclusions and Speculations

This search for a universal functional model for the probability of specific knot types in a range of models for knots has shown that there may exist

such functions, but only with some constraint on the range of application and the expectation that the model be effective in both the finite and asymptotic ranges. For example, the Dobay et al. model is usually the most successful in fitting the data in the range of the data used in this article. Similarly, the QV_- and FV_- models are more successful than the remaining models, but these three models suffer from the defect that are not consistent with the theoretical results giving the specific nature of the asymptotic decay. Among those models that more closely match the asymptotic requirements, the results of this project suggest that the quadratic variation of the exponential portion of the Deguchi-Tsurusaki function is sufficient to capture much of the structure of the data. One is still motivated to relax the expectation that the fitting function be consistent with the theoretically proven linear exponential decay in the asymptotic range. Even with some additional flexibility, however, the model is not entirely successful in capturing all the desired finite range structure of critical behaviour as is shown in Fig. 16. Here the model fails to closely follow the slow decay near the initial data points. This same problem is encountered with efforts to fit the trivial knot distributon in other data sets. Other challenging features, such as the concavity of the function for non-trivial knots, provide additional problems as illustrated in Fig. 17. Notice how the data values wander back and forth across the graph of the model function. This suggests a "higher order" behavior that is not captured in the model function assuming, of course, that the data are sufficiently accurate to demonstrate this structure.

Overall, however, one is struck by the qualitative fit achieved by the quadratic variation in almost all cases when the requirement of proper asymptotic behavior is relaxed. This provides evidence of the existence of a phase transition separating the "finite scale" from the "asymptotic scale" in which the linear exponential decay is sufficient to give a good fit of the data, e.g. as given by the Deguchi-Tsurusaki function. Similarly, one observes a close fit of the data using the Dobay et al. model when we relax the condition of linear exponential decay. The various functions do not uniformly provide the same quality of fit in the finite range, even when one allows full freedom in the exponential coefficients. This can be observed through a comparison of the SSR values achieved for various models, even those which are not consistent with the asymptotic decay to those achieved which are consistent. Note that, in the case of the Dobay et al. model, there is an exceptional improvement over the fit achieved by the other models when considering fitting figure-eight knot data.

Thus, we are led to two conclusions. First, there is no known single

function that adequately models the data in both the finite and asymptotic scales. Second, while the Deguchi-Tsurusaki function may be the optimal fitting function in the asymptotic regime, our quadratic variation is a better option in the finite scale regime.

Consideration of the potential universality of p_K across the five models shows that there is good agreement between the equilateral and geometric trefoil and figure-eight knot values, but that this is no longer the case for the perturbations as shown in Figs. 12 and 13.

The data analysis undertaken here provides information to test the extent of universality of these significant functional parameters for each of the random knot models: the equilateral knots, the geometric knots, and the perturbations of the thick trivial, trefoil, and figure-eight knots. We propose that the similarity one observes in the nature of the knot probability distribution functions is a reflection of much larger similarity in the structure of the various knot spaces, perhaps at differing scales, that are reflected in the changing parameters. One word of caution is necessary, however, in that there are subtle artifacts of the numerical limitations of the accuracy of the data generation and the software employed to analyze the data to extract the parameters. The extraction of the key information is a very delicate process indeed as the number of edges increases.

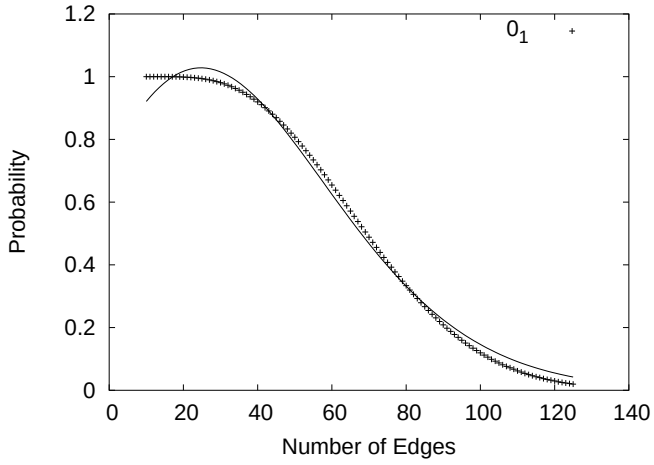

Fig. 16. Perturbation of the thick trivial knot: application of the full variation model to trivial knots.

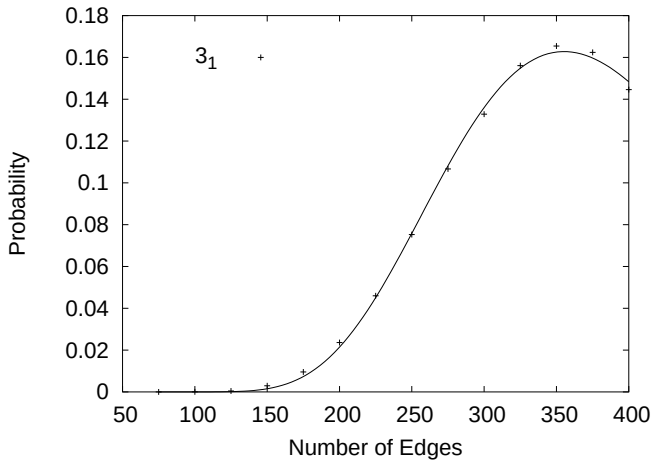

Fig. 17. Perturbation of the thick figure-eight knot: application of the full variation model to trefoil knots.

7. Acknowledgements

We wish to thank Akos Dobay, Tetsuo Deguchi, Yuanan Diao, Andrzej Stasiak, and Stuart Whittington for their review of an early draft and recommendations that have contributed to the work reported here. We also thank Michael Piatek for generating the tube perturbation data and Johnny Zarate for writing the program used to analyze this data. Rawdon was supported by NSF Grant No. 0311010.

References

1. Tetsuo Deguchi and Kyoichi Tsurusaki. Topology of closed random polygons. *J. Phys. Soc. Japan*, 62(5):1411–1414, 1993.
2. Tetsuo Deguchi and Kyoichi Tsurusaki. A statistical study of random knotting using the Vassiliev invariants. In *Random knotting and linking (Vancouver, BC, 1993)*, volume 7 of *Ser. Knots Everything*, pages 89–121. World Sci. Publishing, River Edge, NJ, 1994.
3. Tetsuo Deguchi and Kyoichi Tsurusaki. Universality of random knotting. *Phys. Rev. E*, 55(5):6245–6248, August 1997.
4. Tetsuo Deguchi and Kyoichi Tsurusaki. 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, 1998.
5. Yuanan Diao, Nicholas Pippenger, and De Witt Sumners. On random knots.

J. Knot Theory Ramifications, 3(3):419–429, 1994. Random knotting and linking (Vancouver, BC, 1993).

6. Akos Dobay, Pierre-Edouard Sottas, Jacques Dubochet, and Andrzej Stasiak. Predicting optimal lengths of random knots. *Lett. Math. Phys.*, 55(3):239–247, 2001. Topological and geometrical methods (Dijon, 2000).
7. 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.
8. P. Freyd, D. Yetter, J. Hoste, W. B. R. Lickorish, K. Millett, and A. Ocneanu. A new polynomial invariant of knots and links. *Bull. Amer. Math. Soc. (N.S.)*, 12(2):239–246, 1985.
9. J. M. Hammersley. The number of polygons on a lattice. *Proc. Cambridge Philos. Soc.*, 57:516–523, 1961.
10. J. Hoste, 2000. personal communication.
11. Jim Hoste, Morwen Thistlethwaite, and Jeff Weeks. The first 1,701,936 knots. *Math. Intelligencer*, 20(4):33–48, 1998.
12. Vsevold Katrich, Wilma K. Olsen, Jacques Vologodskii, Alexander Dubochet, and Andrzej Stasiak. Tightness of random knotting. *Phys. Rev. E*, 61(5):5545–5549, 2000.
13. V. Katritch, W. K. Olson, P. Pieranski, J. Dubochet, and A. Stasiak. Properties of ideal composite knots. *Nature*, 388:148–151, July 1997.
14. Kleanthes Koniaris and M. Muthukumar. Self-entanglement in ring polymers. *J. Chem. Phys.*, 95(4):2873–2881, May 1991.
15. R. A. Litherland, J. Simon, O. Durumeric, and E. Rawdon. Thickness of knots. *Topology Appl.*, 91(3):233–244, 1999.
16. J. P. J. Michels and F. W. Wiegel. On the topology of a polymer ring. *Proc. Roy. Soc. London Ser. A*, 403(1825):269–284, 1986.
17. Nathan T. Moore, Rhonald C. Lua, and Alexander Y. Grosberg. Topologically driven swelling of a polymer loop. *PNAS*, 101(37):13431–13435, 2004.
18. J. O'Hara. Energy of knots. In *Ideal knots*, pages 288–314. World Sci. Publishing, River Edge, NJ, 1998.
19. Piotr Pierański. In search of ideal knots. In *Ideal knots*, pages 20–41. World Sci. Publishing, River Edge, NJ, 1998.
20. Nicholas Pippenger. Knots in random walks. *Discrete Appl. Math.*, 25(3):273–278, 1989.
21. E. Rawdon. TOROS. http://www.mathcs.duq.edu/∼rawdon. Program for visualizing, manipulating, and thickness maximizing knots.
22. E. Rawdon. *The Thickness of Polygonal Knots*. PhD thesis, University of Iowa, 1997.
23. Eric J. Rawdon. Approximating the thickness of a knot. In *Ideal knots*, pages 143–150. World Sci. Publishing, River Edge, NJ, 1998.
24. Eric J. Rawdon. Approximating smooth thickness. *J. Knot Theory Ramifications*, 9(1):113–145, 2000.
25. 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, River Edge, NJ, 1998.

26. D. W. Sumners and S. G. Whittington. Knots in self-avoiding walks. *J. Phys. A*, 21(7):1689–1694, 1988.