

Generating 4-regular Hamiltonian Plane Graphs

O. Ascigil, Y. Diao, C. Ernst, D. High, and U. Ziegler

ABSTRACT. This paper describes the study of a special class of 4-regular plane graphs which are Hamiltonian. These graphs are of special interest in knot theory. An algorithm is presented that randomly generates such graphs with n vertices with a fixed (and oriented) Hamiltonian cycle in $O(n)$ time. An exact count of the number of such graphs with n vertices is obtained and the asymptotic growth rate of this number is determined. Numerical evidence is presented to show that the algorithm can be modified to generate these graphs with a near uniform probability. This can be considered as a first step in generating large random knots without bias.

1. Introduction

A *plane map* is a 2-cell decomposition of the oriented sphere into vertices (0-cells), edges (1-cells) and faces (2-cells). A specified edge in a map with a given orientation is called a *root* of the map and a map with a root is called a *rooted map*. The study of rooted maps can be dated back to Tutte [17, 18], where the number of Eulerian rooted plane maps were counted. In a more recent paper [13], Schaeffer gave a different proof of the above mentioned count together with a way that allows the uniform random generation of certain Eulerian plane maps with restricted vertex degrees. (The degree of a vertex is the number of edges incident to this vertex.) The special case of 4-regular rooted plane maps is studied in [15]. It is shown that if $a(n)$ is the number of distinct 4-regular rooted plane maps with n vertices, then

$$a(n) = 2 \frac{3^n (2n)!}{n!(n+2)!} \sim \frac{2}{\sqrt{\pi}} \frac{12^n}{n^{5/2}},$$

see [2, 15, 18].

The main question of this paper is: Is it possible to generate 4-regular Hamiltonian plane maps with an algorithm that has polynomial runtime with uniform or near uniform probability distribution? This question is motivated by an important problem in knot theory. Let K be a smooth knot (or link), that is, a smooth embedding of the unit circle (or circles) into \mathbb{R}^3 . The projection of K into a plane is a closed curve (or curves if K is a link) that may contain self-intersecting points, also called the *crossings* of the projection. The *multiplicity* of a crossing in the

1991 *Mathematics Subject Classification*. Primary 05A16, 05C85, Secondary 68R10, 05C10, 57M25.

Key words and phrases. plane graphs, 4-regular graphs, Hamiltonian cycle, knot diagrams.

projection is the number of strands that pass through that point. A projection is a *regular projection* if there are only finitely many crossings in the projection and all of them are of multiplicity 2. Furthermore, at each crossing in a regular knot projection, the strand that goes over and the strand that goes under are marked, see Figure 1. A regular projection of a knot K with the least number of crossings (among all regular projections of all knots that are topologically equivalent to K) is called a *minimum projection* of K and the number of crossings in that minimum projection is called the *crossing number* of K (which is often denoted by $\text{Cr}(K)$). If the crossings in the regular projection of K are treated as vertices and the arcs joining these crossings are treated as edges, then a regular projection of K can be viewed as a 4-regular plane map.

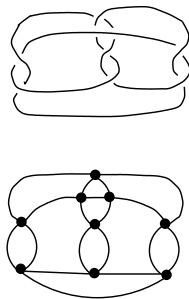


FIGURE 1. A regular projection of a nine crossing knot and its corresponding 4-regular plane map.

If \mathcal{K} is the collection of all (smooth) knots that are topologically equivalent to a given knot K , then the *ropelength* of \mathcal{K} can be thought of as the shortest length of a rope of unit radius that can be used to tie a smooth knot that is a member of \mathcal{K} . The concept of ropelength and its study are important not only in knot theory itself, but also in the applications of knot theory. See for example the collection of articles in the book [16] edited by Stasiak, Katritch, and Kauffman. In a series of papers [3, 4, 5, 6, 7, 8, 9], the authors have studied the ropelength of a knot and established several important results with regard to the asymptotic behavior of the ropelength of knots as a function of their crossing numbers. In particular, an algorithm has been developed and implemented that computes an upper bound on the ropelength of very large knots with thousands of crossings [7, 9]. These results are based on an important observation that any knot K with a crossing number $n = \text{Cr}(K)$ has a regular projection with at most $4n$ crossings that is Hamiltonian [7]. (Indeed, almost all 4-regular graphs are Hamiltonian [12]. Nonetheless, there are examples of knots that do not admit any Hamiltonian minimum projections [7].) It follows that numerical studies concerning large knots can be done by sampling knots from this more restricted and structured class of 4-regular knot projections, namely the 4-regular Hamiltonian plane graphs [8]. Thus it is necessary to develop an algorithm to generate such plane graphs in polynomial runtime with a uniform or near uniform probability distribution.

In this paper it is shown that such an algorithm exists with $O(n)$ runtime and near uniform distribution under the additional assumption that the graphs are rooted Hamiltonian. (For a precise definition see the next section.) In addition, a

precise formula for the number of such rooted 4-regular Hamiltonian plane graphs and an asymptotic estimation of it is given.

The next section contains some basic definitions. This is followed by a section devoted to the determination of the number $X(n)$ of rooted 4-regular Hamiltonian plane graphs (with n vertices). Section 4 proves a theorem that shows $X(n)$ behaves as $O(\frac{12^n}{n^3})$ asymptotically. Section 5 describes the development of an algorithm that generates such 4-regular Hamiltonian plane graphs with linear run time and near uniform distribution based on numerical evidence.

2. Definitions

DEFINITION 2.1. A *rooted Hamiltonian graph* G consists of a pair (G, H) satisfying the following conditions:

- (i) G is a 4-regular plane graph embedded in S^2 with a given Hamilton cycle H ;
- (ii) H contains a rooted edge, that is one edge of H has an orientation. This edge defines an orientation on H , that can distinguish the two disks bounded by H on S^2 as a disk on the right hand side and a disk on the left hand side of H .

It is easy to see that the word “rooted” here is consistent with the original meaning of a rooted graph.

DEFINITION 2.2. Two rooted Hamiltonian graphs (G, H) and (G', H') are *equivalent* if there exists a function $f : S^2 \rightarrow S^2$ such that:

- (i) f is an orientation preserving homeomorphism;
- (ii) $f(G) = G'$, $f(H) = H'$ and $f|_G$ is an isomorphism between the rooted graphs G and G' . That is f maps the rooted edge of G to the rooted edge of G' while preserving its direction.

Note that in [19] a similar definition of equivalence was used for 3-regular Hamiltonian graphs. The goal is to create rooted Hamiltonian graphs algorithmically and to count the total number of such graphs. The conditions that preserve the root and the Hamiltonian cycle in the definition of graph equivalence makes this count possible. The rooted 4-regular Hamiltonian plane graphs are allowed to have loop edges and multiple edges. Using the definition of equivalence, a *standard drawing* of a rooted Hamiltonian graph is defined below. There is exactly one standard diagram (except for deformation of edges) for each equivalence class of rooted 4-regular Hamiltonian plane graphs. An example is shown in Figure 2.

DEFINITION 2.3. A *standard diagram* of a rooted 4-regular Hamiltonian plane graph G is a realization of G such that H is drawn as a circle with its vertices numbered in clockwise order around H and its first vertex v_1 on top of the circle. The rooted edge is the edge from v_1 to v_2 along the Hamiltonian cycle.

Let G be a rooted 4-regular Hamiltonian plane graph with H being the given Hamilton cycle. Since G is 4-regular, each vertex v of G is incident to two edges of G that are not on H . It is possible that these two edges are the same, in other words, it is possible for v to be incident to a loop edge that is not on H . For each vertex v , there are the following three possibilities.

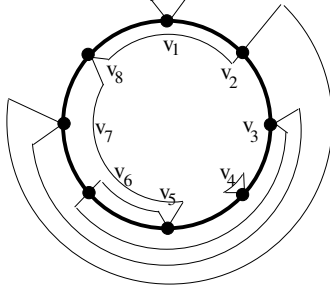


FIGURE 2. A standard diagram with $n=8$, inside string $((()()))$, and outside string $()((()))$. v_7 is a double-outside vertex, v_2 is a transition vertex, and v_5 is a double-inside vertex. The edge (v_8, v_5) is an inside edge and the edge (v_6, v_3) is an outside edge.

(a) v is incident to two edges (or a loop edge) contained in the unbounded region outside of H . In this case v is called a *double-outside vertex*, or a *DOV* for short.

(b) v is incident to two edges (or a loop edge) contained in the disk bounded by H . In this case v is called a *double-inside vertex*, or a *DIV* for short.

(c) v is incident to one edge contained in the disk bounded by H and one edge contained in the unbounded region outside of H . In this case v is called a *transition vertex*, or a *TV* for short.

Similarly, for an edge e of G that is not on H , it is either an *inside edge* (IE) or an *outside edge* (OE), depending on whether or not it is in the disk bounded by H .

3. The number of rooted 4-regular Hamiltonian plane graphs

In this section the exact number of the different rooted 4-regular Hamiltonian plane graphs with n vertices is determined. In this process an algorithm is developed to randomly generate standard diagrams of such rooted Hamiltonian graphs.

The following process can be used to (re-) draw a given standard diagram of a rooted 4-regular Hamiltonian plane graph. First H (the circle) with all the vertices is drawn. None of the edges not on H are added directly. Instead, two short line segments are attached tentatively to each DOV so that the line segments are in the outside region of H with an open end. These line segments are part of the edges not on H incident to the vertex. Each such open end is called an *outside edge end point* (OEEP). Similarly, two line segments bounded in the inside disk are attached to each DIV with the two open ends called inside edge end points, or IEEPs. Two line segments, one bounded in the inside disk and another bounded in the outside region of H , will be attached to each TV. See Figure 3 below.

After all the line segments have been added to the vertices, the following observation can be made regarding the inside edges. Since the IEEPs are paired by the inside edges (not yet drawn) that connect them, there is a natural order between

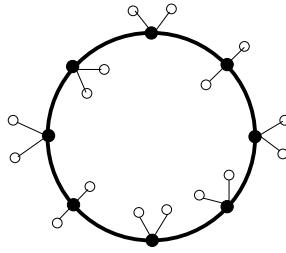


FIGURE 3. An illustration of the edge end points from the graph of Figure 2.

the two edge endpoints of each pair when traveling clockwise once around H starting at v_1 . For each such pair, one assigns the first edge endpoint encountered an open parenthesis ‘(’ and the second point encountered a closed parenthesis ‘)’. This generates a balanced string of parentheses. In such a balanced string the number of ‘(’ equals the number of ‘)’ due to the pairing of endpoints. In addition the fact that the inside edges do not intersect each other leads to a string where in every prefix of the string, the number of ‘)’ is never larger than the number of ‘(’. Similarly a string of balanced parentheses can be assigned to the outside edges, see Figure 2.

Conversely, for a given arrangement of IEEPs and OEEPs, a unique Hamiltonian graph G can be constructed from two such balanced strings of parentheses (of the appropriate length), see Figure 4. In other words, there is a one-to-one correspondence between pairs of balanced strings of parentheses and standard diagrams of rooted 4-regular Hamiltonian plane graphs for a given arrangement of IEEPs and OEEPs. These observations lead to the following algorithm for generating rooted 4-regular Hamiltonian plane graphs with n vertices.

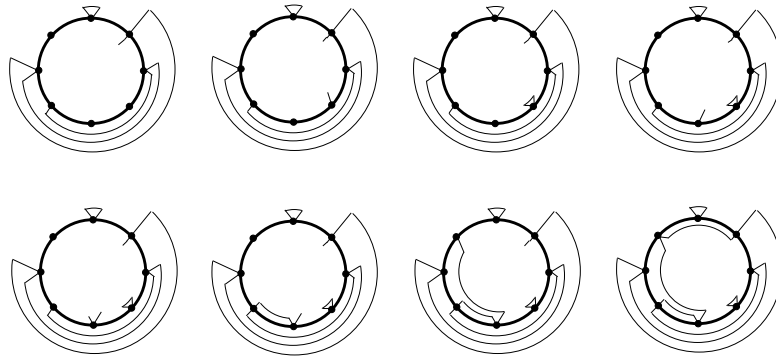


FIGURE 4. Given the balanced string $((()())$ for the IEEPs, the inside edges from the graph of Figure 2 are reconstructed.

1. **Arrangements of vertex types:** Starting with a Hamiltonian cycle H_n of n vertices, determine (choose) the number of vertices of each vertex type (TV, DIV or DOV) and pick a random arrangement of these vertices along the Hamiltonian cycle.

2. Arrangement of inside/outside edges: Assign non-Hamiltonian edges to the inside and separately to the outside edge-endpoints using randomly chosen balanced strings of the appropriate length.

The detailed descriptions of the algorithm follow.

Arrangements of vertex types.

The algorithm must first choose the number of vertices for each type and decide where to put them on H_n .

First, choose the number of transition vertices. Note that the number of transition vertices cannot be odd, since this would lead to an odd number of IEEPs and OEPPs and that to an edge which has to cross a Hamiltonian edge. Let $2t$ be the number of transition vertices. From the remaining $n - 2t$ vertices, choose p DIVs. The remaining $j = n - 2t - p$ vertices all become DOVs. Once t , p and j are chosen, there are

$$\binom{n}{2t} \binom{n-2t}{p} \binom{n-2t-p}{n-2t-p} = \binom{n}{2t} \binom{n-2t}{p} = \frac{n!}{(2t)!p!j!}$$

different ways to arrange these vertices on H_n since the graph is rooted. Figure 2 is an example for the case $n = 8$, $t = 1$ and $p = 3$.

Arrangement of inside/outside edges.

An arrangement of IEs is achieved by selecting a balanced string of parentheses of length $2t + 2p$ (which is equal to the number of IEEPs). This string is called an *inside string*. Each parenthesis in an inside string corresponds to one unique IEEP. Starting at v_1 and traversing clockwise, the first IEEP encountered is called the *starting* IEEP. The process of assigning the edges moves through the symbols of the balanced inside string and through the IEEPs simultaneously, see Figure 4. Beginning at the starting IEEP and coming across the other IEEPs in clockwise order, if a '(' is the next symbol in the balanced string an edge is started at the current IEEP. If a ')' is the next symbol in the balanced string an edge is closed at the current IEEP with the most recently started but not-yet-closed edge, see Figure 4. The number of balanced strings of size $2x$ is $\frac{1}{x+1} \binom{2x}{x}$, which is also commonly called a Catalan number, see [13, 14, 15] for examples of this construction. Therefore, there are $\frac{1}{t+p+1} \binom{2t+2p}{t+p}$ ways to choose the inside string. Similarly, the outside edges are constructed using an outside string of size $2n - (2t + 2p)$. It follows that there are

$$\frac{1}{t+p+1} \binom{2t+2p}{t+p} \frac{1}{n-t-p+1} \binom{2n-2t-2p}{n-t-p}$$

ways to connect all the edges. This leads to the following theorem.

THEOREM 3.1. *The number of non-equivalent rooted 4-regular Hamiltonian plane graphs of n vertices is*

$$X(n) = \sum_{t=0}^{\frac{n}{2}} \sum_{p=0}^{n-2t} \binom{n}{2t} \binom{n-2t}{p} \binom{2t+2p}{t+p} \binom{2n-2t-2p}{n-t-p} \frac{1}{t+p+1} \frac{1}{n-t-p+1}.$$

The steps in the algorithm described above can be executed in linear time and with linear memory resources. This is obviously true for constructing H_n . Once t and p have been chosen, a unique arrangement of the vertices is selected by a uniform random permutation of an arbitrary vector containing $2t$ TVs, p DIVs,

and $j = n - 2t - p$ DOVs. (Well developed linear permutation algorithms can be easily found.) There are several different ways to generate random uniform balanced strings in $O(n)$ time, see for example [1]. 10,000 graphs with 2000 crossings are generated by the implementation of the algorithm in 24 seconds on a standard desktop PC.

4. The asymptotic behavior of $X(n)$

Let

$$g(n, t, p) = \binom{n}{2t} \binom{n-2t}{p} \binom{2t+2p}{t+p} \binom{2n-2t-2p}{n-t-p} \frac{1}{t+p+1} \frac{1}{n-t-p+1}$$

so that

$$X(n) = \sum_{t=0}^{\frac{n}{2}} \sum_{p=0}^{n-2t} g(n, t, p).$$

Because $X(n)$ is strictly increasing, investigating the asymptotic behavior for values of n that are divisible by 6 will suffice. The reason for this restriction on n is mainly for the simplicity in analyzing the maximum of $g(n, t, p)$. Let us start with the following lemma.

LEMMA 4.1. (i) $g(n, t, n/2-t+x) = g(n, t, n/2-t-x)$ for $x \in \{0, 1, \dots, n/2-t\}$.

(ii) Let x be a nonnegative integer divisible by 3, then $\frac{x!}{(\frac{x}{3})!(\frac{x}{3})!(\frac{x}{3})!} \geq \frac{x!}{a!b!c!}$ for any nonnegative integers a, b, c such that $a + b + c = x$.

PROOF. It can be seen that

$$\begin{aligned} & g(n, t, n/2-t+x) \\ &= \binom{n}{2t} \binom{n-2t}{\frac{n}{2}-t+x} \binom{n+2x}{\frac{n}{2}+x} \binom{n-2x}{\frac{n}{2}-x} \frac{1}{\frac{n}{2}+x+1} \frac{1}{\frac{n}{2}-x+1} \\ &= \binom{n}{2t} \binom{n-2t}{\frac{n}{2}-t-x} \binom{n+2x}{\frac{n}{2}+x} \binom{n-2x}{\frac{n}{2}-x} \frac{1}{\frac{n}{2}+x+1} \frac{1}{\frac{n}{2}-x+1} \\ &= g(n, t, n/2-t-x). \end{aligned}$$

This proves (i). To avoid fractions $(3x)!/a!b!c!$ is considered for nonnegative integers a, b, c such that $a + b + c = 3x$. It suffices to show that $x!x!x! \leq a!b!c!$ for such a, b, c . Notice that there are two cases: exactly one of a, b, c is greater than x or two of a, b, c are greater than x . For the first case, $x!x!x! \leq a!b!c!$ is equivalent to $x!x!x! \leq (x-k)!(x-l)!(x+k+l)!$ for nonnegative integers k, l such that $0 < k+l \leq x$. For the second case, $x!x!x! \leq a!b!c!$ is equivalent to $x!x!x! \leq (x+k)!(x+l)!(x-k-l)!$ for positive integers k, l such that $k+l \leq x$. The second case is proven below and the first case is left to the reader. This is equivalent to

$$\frac{x!}{(x-k-l)!} \leq \frac{(x+k)!}{x!} \frac{(x+l)!}{x!},$$

which reduces to

$$\underbrace{[x(x-1)\dots(x-k)(x-k-1)\dots(x-k-l+1)]}_{k+l \text{ factors}} \leq \underbrace{[(x+1)\dots(x+k)(x+1)\dots(x+l)]}_{k+l \text{ factors}}.$$

Both sides of the above contain $k + l$ factors, where each factor on the right is greater than any factor on the left. So the inequality holds. \square

THEOREM 4.2. *If $n \pmod{6} \equiv 0$, then for fixed t , $g(n, t, p)$ is maximized at $p = n/2 - t$. Moreover the absolute maximum of $g(n, t, p)$ occurs at $t = n/6$ and $p = n/3$.*

PROOF. By Lemma 4.1 (i), it suffices to show that for $p' = \frac{n}{2} - t + x$ where $x \in \{1, 2, \dots, \frac{n}{2} - t\}$, $g(n, t, p') < g(n, t, \frac{n}{2} - t)$. After some simplification, this inequality reduces to

$$\begin{aligned} & \binom{n-2t}{\frac{n}{2}-t+x} \binom{n+2x}{\frac{n}{2}+x} \binom{n-2x}{\frac{n}{2}-x} \frac{1}{\frac{n}{2}+x+1} \frac{1}{\frac{n}{2}-x+1} \\ < & \binom{n-2t}{\frac{n}{2}-t} \binom{n}{\frac{n}{2}} \binom{n}{\frac{n}{2}} \frac{1}{\frac{n}{2}+1} \frac{1}{\frac{n}{2}+1}. \end{aligned}$$

Writing the binomials as factorials and using the fact that all quantities are positive, this inequality is further simplified to:

$$(4.1) \quad \frac{\left(\frac{n}{2}\right)!^2}{\left(\frac{n}{2}+x\right)!^2} \frac{\left(\frac{n}{2}\right)!^2}{\left(\frac{n}{2}-x\right)!^2} \frac{\left(\frac{n}{2}-t\right)!^2}{\left(\frac{n}{2}-t+x\right)! \left(\frac{n}{2}-t-x\right)!} \\ \times \frac{(n+2x)!(n-2x)!}{n!n!} \frac{\left(\frac{n}{2}+1\right)\left(\frac{n}{2}+1\right)}{\left(\frac{n}{2}+x+1\right)\left(\frac{n}{2}-x+1\right)} < 1.$$

Inequality (4.1) is proven by induction on x . For $x = 1$, (4.1) simplifies to

$$\frac{\left(\frac{n}{2}\right)!^2}{\left(\frac{n}{2}+1\right)!^2} \frac{\left(\frac{n}{2}\right)!^2}{\left(\frac{n}{2}-1\right)!^2} \frac{\left(\frac{n}{2}-t\right)!^2}{\left(\frac{n}{2}-t+1\right)! \left(\frac{n}{2}-t-1\right)!} \frac{(n+2)!(n-2)!}{n!n!} \frac{\left(\frac{n}{2}+1\right)\left(\frac{n}{2}+1\right)}{\left(\frac{n}{2}+2\right)\left(\frac{n}{2}\right)} < 1,$$

or equivalently

$$\frac{\frac{n}{2}\left(\frac{n}{2}-t\right)(n+1)(n+2)}{\left(\frac{n}{2}+2\right)\left(\frac{n}{2}-t+1\right)(n)(n-1)} = \frac{n^3 + 3n^2 + 2n - 2n^2t - 6nt - 4t}{n^3 + 5n^2 + 2n - 2n^2t - 6nt + 8t - 8} < 1.$$

The denominator is always positive for $n \geq 6$ and this inequality will be true whenever

$$(4.2) \quad n^3 + 3n^2 + 2n - 2n^2t - 6nt - 4t < n^3 + 5n^2 + 2n - 2n^2t - 6nt + 8t - 8.$$

But (4.2) is equivalent to $0 < 2n^2 + 12t - 8$, which is true for $n \geq 6$, $t = \{0, 1, \dots, \frac{n}{2}\}$. This proves the case $x = 1$. Assume now that the inequality is true for $x = k \geq 1$ and consider the case $x = k + 1$ (where $k \leq \frac{n}{2} - t - 1$). It needs to be shown that

$$\begin{aligned} & \frac{\left(\frac{n}{2}\right)!^2}{\left(\frac{n}{2}+k+1\right)!^2} \frac{\left(\frac{n}{2}\right)!^2}{\left(\frac{n}{2}-k-1\right)!^2} \frac{\left(\frac{n}{2}-t\right)!^2}{\left(\frac{n}{2}-t+k+1\right)! \left(\frac{n}{2}-t-k-1\right)!} \\ & \times \frac{(n+2k+2)!(n-2k-2)!}{n!n!} \frac{\left(\frac{n}{2}+1\right)\left(\frac{n}{2}+1\right)}{\left(\frac{n}{2}+k+2\right)\left(\frac{n}{2}-k\right)} < 1. \end{aligned}$$

Rewrite the above inequality as

$$\underbrace{\frac{\left(\frac{n}{2}-k\right)^2}{\left(\frac{n}{2}+k+1\right)^2} \frac{\left(\frac{n}{2}-t-k\right)}{\left(\frac{n}{2}-t+k+1\right)}}_A \underbrace{\frac{\left(\frac{n}{2}\right)!^2}{\left(\frac{n}{2}+k\right)!^2} \frac{\left(\frac{n}{2}\right)!^2}{\left(\frac{n}{2}-k\right)!^2} \frac{\left(\frac{n}{2}-t\right)!^2}{\left(\frac{n}{2}-t+k\right)! \left(\frac{n}{2}-t-k\right)!}}_B$$

$$\begin{aligned} & \times \underbrace{\frac{(n+2k+1)(n+2k+2)}{(n-2k)(n-2k-1)} \frac{(\frac{n}{2}+k+1)(\frac{n}{2}-k+1)}{(\frac{n}{2}+k+2)(\frac{n}{2}-k)}}_C \\ & \times \underbrace{\frac{(n+2k)!(n-2k)!}{n!n!} \frac{(\frac{n}{2}+1)(\frac{n}{2}+1)}{(\frac{n}{2}+k+1)(\frac{n}{2}-k+1)}}_D < 1. \end{aligned}$$

Now $BD < 1$ by the induction assumption, so it suffices to show that $AC \leq 1$ since it is easy to see that every factor in A and C is non-negative. After simplifying the following equation is obtained for AC:

$$AC = \frac{(n-2t-2k)}{(n-2t+2k+2)} \frac{(n-2k+2)}{(n-2k-1)} \frac{(n+2k+1)}{(n+2k+4)}$$

Furthermore, since $\frac{(n+2k+1)}{(n+2k+4)} \leq 1$, it suffices to show that

$$(4.3) \quad \frac{(n-2t-2k)}{(n-2t+2k+2)} \frac{(n-2k+2)}{(n-2k-1)} \leq 1$$

for $n \geq 6$, $t = \{0, 1, \dots, \frac{n}{2}\}$, $k = \{1, 2, \dots, \frac{n}{2} - t - 1\}$. Notice that (4.3) is equivalent to $n - 4nk + 8k^2 + 2k - 6t + 2 < 0$. Since $n - 4nk + 8k^2 + 2k - 6t + 2 \leq n - 4nk + 8k^2 + 2k + 2$, it suffices to show that $n - 4nk + 8k^2 + 2k + 2 < 0$. The maximum value of the quadratic function in the last inequality in terms of k for $1 \leq k \leq n/2 - 1$ is at the end points of the interval, so either at $k = 1$ or at $k = n/2 - 1$. The reader can verify that both are negative (for $n \geq 6$). Thus (4.3), and hence (4.1) hold.

Substituting $p = \frac{n}{2} - t$ into $g(n, t, p)$ results in:

$$g(n, t, \frac{n}{2} - t) = \underbrace{\binom{n}{2t} \binom{n-2t}{\frac{n}{2} - t}}_A \frac{\binom{n}{\frac{n}{2}}}{(\frac{n}{2}+1)} \frac{\binom{n}{\frac{n}{2}}}{(\frac{n}{2}+1)}.$$

This is maximized if A is maximized. However

$$A = \frac{n!}{(2t)!(\frac{n}{2}-t)!(\frac{n}{2}-t)!}.$$

By Lemma 4.1 (ii), A achieves its maximum at $2t = \frac{n}{2} - t = \frac{n}{3}$, i.e., $t = \frac{n}{6}$ and $p = \frac{n}{2} - t = \frac{n}{3}$. \square

The above result about the maximum value of $g(n, t, p)$ can be used to determine the asymptotic order of $X(n)$. The graph of $g(n, t, p)$ is single-peaked and all terms for which $t \approx n/6$ and $p \approx n/3$ are of similar order. In the following, all factorials are replaced with an approximation based on Stirling's formula

$$n! = \sqrt{2\pi n} n^n e^{-n} e^{\lambda_n},$$

(where $\frac{1}{12n+1} < \lambda_n < \frac{1}{12n}$), which does not affect the asymptotic order of $X(n)$.

THEOREM 4.3. $X(n) \sim O(\frac{12^n}{n^3})$.

PROOF. Substituting Stirling's Formula for factorials and simplifying leads to:

$$\begin{aligned}
g(n, t, p) &= \frac{n!(2t+2p)!(2n-2t-2p)!}{(2t)!p!(n-2t-p)!(t+p)!^2(n-t-p)!^2} \frac{1}{(t+p+1)} \frac{1}{(n-t-p+1)} \\
&\doteq \frac{\sqrt{2}}{4\pi^2(t+p+1)(n-t-p+1)} \sqrt{\frac{n}{tp(n-2t-p)(t+p)(n-t-p)}} \\
(4.4) \quad &\times \frac{n^n 4^n}{(2t)^{2t} p^p (n-2t-p)^{n-2t-p}}.
\end{aligned}$$

For $t \approx n/6$ and $p \approx n/3$, the combined contribution of all factors in the second line of formula (4.4) is of order $O(1/n^4)$ and thus $g(n, t, p)$ is mainly determined by the factors in the third line of (4.4). The behavior of those factors is investigated first in terms of p . Define $h(p) = p^p (n-2t-p)^{n-2t-p}$, then differentiating $\ln h(p)$ yields

$$\frac{h'(p)}{h(p)} = \ln \frac{p}{n-2t-p}.$$

This implies that $h'(p) = 0$ for $p = n/2 - t$, $h'(p) > 0$ if $p > n/2 - t$, and finally $h'(p) < 0$ if $p < n/2 - t$. This is consistent with the earlier result that $g(n, t, p)$ has a maximum in p for $p = n/2 - t$ (since $h(p)$ is in the denominator). Similarly, the behavior in t is investigated. Define $f(t) = (2t)^{2t} (n-2t-p)^{n-2t-p}$, then differentiating $\ln f(t)$ yields

$$\frac{f'(t)}{f(t)} = 2 \ln \frac{2t}{n-2t-p}.$$

This implies that $f'(t) = 0$ for $t = (n-p)/4$, $f'(t) > 0$ if $t > (n-p)/4$, and finally $f'(t) < 0$ if $t < (n-p)/4$. Thus defining

$$\hat{g}(n, t, p) = \frac{n^n 4^n}{(2t)^{2t} p^p (n-2t-p)^{n-2t-p}},$$

means $\hat{g}(n, t, p)$ is increasing for $p < \frac{n}{2} - t$ and decreasing for $p > \frac{n}{2} - t$ as a function of p , and is increasing for $t < \frac{n-p}{4}$ and decreasing for $t > \frac{n-p}{4}$ as a function of t .

The asymptotic order of $g(n, t, p)$ around $(t = n/6, p = n/3)$ (where it achieves its absolute maximum) is investigated next. Let $t = n/6 + x$ and $p = n/3 - x + y$ for $0 \leq |x| \leq n^{3/4}$ and $0 \leq |y| \leq n^{3/4}$. Note that the minus sign of the x term arises from the fact that for fixed $t = n/6 + x$, $g(n, t, p)$ is maximized at $p = n/2 - t = n/3 - x$, see Theorem 4.2. Varying this by y results in $p = n/3 - x + y$. Due to earlier remarks it suffices to investigate $\hat{g}(n, t, p)$. Considering one factor at a time the following is obtained:

$$(2t)^{2t} = \left(\frac{n}{3} + 2x\right)^{n/3+2x} = \left(\frac{n}{3}\right)^{n/3+2x} \left(1 + \frac{6x}{n}\right)^{n/(6x)} e^{2x+12x^2/n}.$$

For large n , $(1 + \frac{6x}{n})^{n/(6x)}$ is approximately e , since $0 \leq |x| \leq n^{3/4}$. So the asymptotic order is:

$$(2t)^{2t} \sim \left(\frac{n}{3}\right)^{n/3+2x} e^{2x+12x^2/n}.$$

Using similar arguments leads to

$$(p)^p = \left(\frac{n}{3} - x + y\right)^{n/3-x+y} \sim \left(\frac{n}{3}\right)^{n/3-x+y} e^{-x+y+3(y-x)^2/n}$$

and

$$(n - 2t - p)^{n-2t-p} = \left(\frac{n}{3} - x - y\right)^{n/3-x-y} \sim \left(\frac{n}{3}\right)^{n/3-x-y} e^{-x-y+3(y+x)^2/n}.$$

Combining these three factors gives:

$$\hat{g}(n, t, p) = \frac{n^n 4^n}{(2t)^{2t} p^p (n - 2t - p)^{n-2t-p}} \sim \frac{n^n 4^n}{(n/3)^n e^{(18x^2+6y^2)/n}} = \frac{12^n}{e^{(18x^2+6y^2)/n}}.$$

Recall that the second line of formula 4.4 is of order $O(1/n^4)$. This results in the asymptotic estimate:

$$g(n, t, p) \sim \frac{12^n}{n^4} \frac{1}{e^{(18x^2+6y^2)/n}}$$

for $t = n/6 + x$ and $p = n/3 - x + y$ with $0 \leq |x|, |y| \leq n^{3/4}$. The following leads to a lower bound for $X(n)$:

$$\begin{aligned} X(n) &\geq \sum_{|x| \leq n^{3/4}} \sum_{|y| \leq n^{3/4}} g(n, t, p) \\ &\sim \frac{12^n}{n^4} \sum_{|x| \leq n^{3/4}} \sum_{|y| \leq n^{3/4}} \frac{1}{e^{(18x^2+6y^2)/n}} \\ &\sim \frac{12^n}{n^4} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-(18x^2+6y^2)/n} dx dy \\ &= \frac{12^n}{n^4} \frac{n\pi}{6\sqrt{3}} = \frac{\pi 12^n}{6\sqrt{3}n^3}. \end{aligned}$$

To show that $X(n)$ is at most of the same order $O(\frac{12^n}{n^3})$, one observes that

$$X(n) = \sum_{|x| \leq n^{3/4} \text{ and } |y| \leq n^{3/4}} g(n, t, p) + \sum_{|x| > n^{3/4} \text{ or } |y| > n^{3/4}} g(n, t, p).$$

It suffices to show that the second sum in the above is of an order at most $O(\frac{12^n}{n^3})$.

$$\sum_{|x| > n^{3/4} \text{ or } |y| > n^{3/4}} g(n, t, p) \sim \frac{1}{n^4} \sum_{|x| > n^{3/4} \text{ or } |y| > n^{3/4}} \hat{g}(n, t, p).$$

Here an upper bound can be found by assuming that either $x = n^{3/4}$ or $y = n^{3/4}$, since the earlier analysis of the behavior of $\hat{g}(n, t, p)$ shows that the terms $\hat{g}(n, t, p)$ are maximized under this condition. Under this assumption,

$$\hat{g}(n, t, p) \sim \frac{12^n}{e^{(18x^2+6y^2)/n}} \leq \frac{12^n}{e^{6\sqrt{n}}}.$$

Since there are at most n^2 such terms, the second sum is bounded above by a term of order

$$\frac{12^n}{n^4} \frac{n^2}{e^{6\sqrt{n}}} = \frac{12^n}{n^2 e^{6\sqrt{n}}}.$$

Therefore the asymptotic order of the upper bound is also $\frac{12^n}{n^3}$ which completes the proof. \square

REMARK 4.4. Numerical evidence indicates that

$$X(n) \approx 1.27 \frac{12^n}{n^3}.$$

For example, the quotient $\frac{X(n)n^3}{12^n}$ equals 1.269 for $n = 1000$ and 1.272 for $n = 2000$. Recall that in the introduction it is stated that the number of 4-regular rooted

planar graphs with n vertices is of the order $\frac{12^n}{n^{2.5}}$ which is a larger order. This makes perfect sense since each 4-regular rooted Hamiltonian graph gives rise to a 4-regular rooted graph by simply considering the edge from v_1 to v_2 on the Hamiltonian cycle as a root edge. However, a 4-regular rooted graph may not be Hamiltonian.

In [19] it is shown that the number of cubic rooted Hamiltonian plane graphs is of order $\frac{4^n}{n^3}$ which is - of course - much smaller than the order of $X(n)$.

5. Generating rooted 4-regular Hamiltonian graph with approximate uniformity

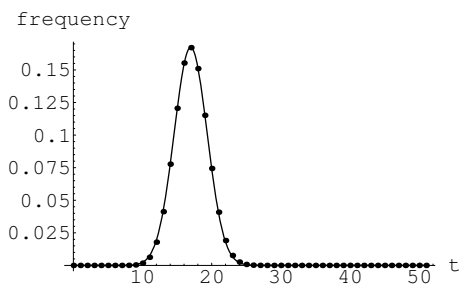


FIGURE 5. A typical example of the distribution $T_n(t)$ for $n = 102$ together with the fitted normal distribution with $\mu = 16.9565$ and $\sigma = 2.3846$.

Every rooted Hamiltonian graph has a non zero probability to be generated by the algorithm described in section 3. However two such graphs are not created with equal probability if the values of t and p are picked with uniform probability from their respective intervals $[0, n/2]$ and $[0, n - 2t]$. As was shown in the previous section, the number of graphs which can be generated varies dramatically for different t and p values. This section outlines steps that are based on numerical evidence and seem to generate the different rooted 4-regular Hamiltonian graphs with approximate uniformity probability. The basic idea is to choose the parameters t and p with a probability distribution that reflects the different numbers of graphs that are possible for given values of t and p .

If one defines $T_n(t) = \frac{\sum_{p=0}^{n-2t} g(n,t,p)}{X(n)}$ then the distribution generated over the interval $[0, n/2]$ can be approximated by a normal distribution $f_n(\mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ as shown in Figure 5. The quality of the non-linear fit of T_n to a normal distribution increases for increasing n , with an estimated variance of 2.810^{-9} for $n = 600$. T_n was approximated by a normal distribution for values of n up to 850. The μ and σ values of these approximations change with n as described by the functions $\mu = n/6 - 0.0526$ and $\sigma = 0.2355\sqrt{n} + 0.0063$. The fit for both functions is obtained with $R^2 \geq .999$. The $T_n(t)$ with $n \equiv 0 \pmod{6}$ is maximal for $t = n/6$. The -0.0526 in the function for the μ is indicative of the fact that there are more graphs for $t < n/6$ than for $t > n/6$. It is expected that the mean is $n/6$ asymptotically.

If one defines $P_{n,t}(p) = \frac{g(n,t,p)}{\sum_{p=0}^{n-2t} g(n,t,p)}$ then the distribution generated over the interval $[0, n - 2t]$ can also be approximated by a normal distribution $f_{n,t}(\mu, \sigma)$. The

quality of the non-linear fit of $P_{n,t}$ to a normal distribution increases for increasing n , with an estimated variance of 6.610^{-11} for $n = 600$ and $t = 200$. $P_{n,t}$ was approximated by a normal distribution for values of n varying from 50 to 850 and for each t in $[0, n/2]$ for each n . The μ for all n fit the function $\mu = n/2 - t$. The σ fit functions of the type $\sigma = c_1\sqrt{n/2 - t} + c_2$, where the constants c_1 and c_2 depend on n as shown in Figure 6. Based on the σ values of the approximations, $c_1(n) = 0.7074 - \frac{5.31}{n^{1.5}}$ and $c_2(n) = 0.003 + \frac{2.47}{n^{3/4}}$ were obtained with $R^2 \geq .999$.

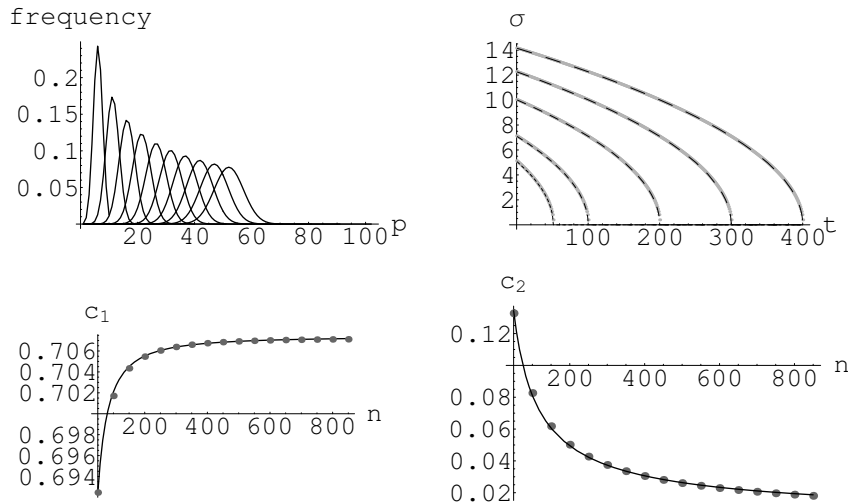


FIGURE 6. The left top graph shows $P_{n,t}(p)$ for $n = 102$ and for t values 5, 10, 15, . . . , 50. The top right graph shows the standard deviation data collected for $n = 102, 200, 400, 600,$ and 800 . together with functions of the form $\sigma = c_1\sqrt{n/2 - t} + c_2$. The bottom graphs show the approximation of the c_1 and c_2 values through their fitting functions both of which were obtained with $R^2 \geq .999$.

If t and p are chosen according to the algorithm outlined above then the question is how uniform is the distribution generated over the set of all rooted Hamiltonian graphs of size n ? One standard way to compare a sample distribution to a known distribution is through a χ -square test. In the given situation this is not appropriate since the distribution obtained for the graphs is known to be not uniform and thus the test failed - as expected - every time. In addition, for very large sample spaces, the test will fail almost always [10]. However, two other tests were conducted to answer this question which provide evidence that the distribution is approximately uniform. Both tests compare the distribution obtained through the described algorithm with a truly uniform distribution. Since the number $X(n)$ of graphs is growing very fast - for example $X(7) = 85, 278$ and $X(8) = 720, 448$ - the tests were only conducted for small values of n due to the limits on computing power.

The first test analyzed the probability that a values appears more than once in a sample. The function

$$p(k) = \binom{M}{k} \left(\frac{k}{M}\right)^h \left(1 + \sum_{i=1}^{k-1} (-1)^i \binom{k}{i} \left(\frac{k-i}{k}\right)^h\right)$$

gives the theoretical probability of picking a sample of size h with uniform probability from a space of size M and obtaining exactly k different values. Note that for larger values of n , the space M is very large and thus it is very difficult to compute $p(k)$ exactly and thus a small n value was picked. For 7-vertex standard diagrams ($M = 85,278$) a sample of size $h = 1000$ was used. Given the above distribution, the probability that such the sample contains at least 991 different numbers is calculated to be .928555. In 10 test runs with the described algorithm, only one failed this test with 989 different graphs. For 8-vertex diagrams and a sample of size $h = 1000$, the probability that such a sample contains at least 996 different numbers is .999263. In 5000 test runs 4 failed this test with 995 different graphs.

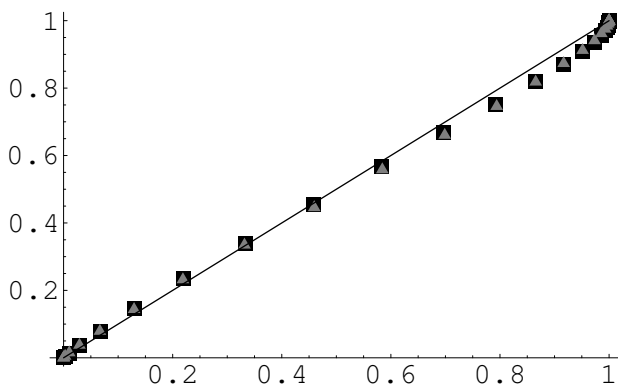


FIGURE 7. Comparing expected and actual cumulative probabilities for 7 (gray triangles) and 8 (black squares) vertex graphs. The difference between the points and the line $x = y$ indicates that the distribution is not quite uniform. Note that the maximal difference between x- and y-coordinate is 0.048.

The second test compared the cumulative probabilities for the generated data with the cumulative probabilities of data generated by a truly uniform distribution in a probability-probability plot, see [10]. To generate a truly uniform distribution a random number generator was used to pick positive integers in the interval $[1, X(n) = M]$. All random numbers for the emperical explorations were generated using an implementation of the Mersenner Twister random number generator which has been shown to pass several important statistical test for good random number generators [11].

As a total sample size for both distributions $10M$ was chosen. This again only allows small values of n . For example, for $n = 8$ the sample of size $10M$ consists of over 7.2 million generated graphs. Let $N(i)$ and $G(i)$ be the two sample distributions where $N(i)$ or $G(i)$ indicate how often the i -th number or the i -th

graph was generated where $i \in \{1, 2, \dots, M\}$. (For the positive integer there is a natural order and for the graphs a random order of all the graphs was chosen.) Let $Y_N(x)$ be the x -th smallest distinct value in $N(i)$ and $N_G(x)$ be the x -th smallest distinct value in $G(i)$. Figure 7 shows the results of this analysis for $n = 7$ and $n = 8$. Each data point in the plot compares cumulative probabilities. The x -coordinate uses the uniform random number distribution and the y -coordinate uses the distribution of random graphs generated by the described algorithm. To be more precise the k -th data point in Figure 7 is obtained in the following manner: For the x -coordinate compute $\frac{\text{number of } N(i) \leq Y_N(k)}{10M} - \frac{0.5}{10M}$ and for the y -coordinate compute $\frac{\text{number of } G(i) \leq Y_G(k)}{10M} - \frac{0.5}{10M}$ (see [10]). If the actual graph distribution is a uniform distribution, then all the data points will be on the line $x = y$. Figure 7 shows that the distribution obtained through the described algorithm is not uniform, however the data points are close to the line $y = x$.

6. Conclusion

We end this discussion with some open questions.

1. Is it possible to sample rooted 4-regular Hamiltonian graphs with uniform probability? In order to do this one needs to understand analytically the distributions $T_n(t)$ and $P_{n,t}(p)$.

2. Can this be used to randomly sample large knots with only a small amount of bias in the sampling method? If one is interested in sampling large knots randomly, several additional problems arise. First, there is no unique way of changing a 4-regular plane graph uniquely into a knot diagram. Many different assignments of over- and under- information at each crossing are possible. Second, the number of different Hamiltonian cycles in rooted Hamiltonian graphs varies greatly and thus this sampling method may be biased towards plane graphs that have a large number of Hamiltonian cycles. When considering these graphs as knot diagrams, this will introduce a bias towards knots that have diagrams with lots of Hamiltonian cycles.

3. Is it possible to consider only 4-edge connected rooted Hamiltonian plane graphs with uniform probability? In knot theory one is usually concerned with prime knots or links. Without defining this concept, we just note that any minimal diagram of a prime knot or link is a 4-regular plane graph that is also 4-edge connected. Therefore sampling these graphs is of special interest in knot theory.

References

- [1] D. B. Arnold and M. R. Sleep, *Uniform Random Generation of Balanced Parenthesis Strings*, ACM Transactions on Programming Languages and Systems, **2**,(1980), pp. 122–128.
- [2] E. Brézin, C. Itzykson, G. Parisi and J. B. Zuber, *Planar diagrams*, Commun. Math. Phys. **59**,(1978), pp. 35–51.
- [3] Y. Diao and C. Ernst, *Realizable Powers of Ropelengths by Nontrivial Knot Families*, JP Journal of Geometry and Topology **4** no. 2 (2004), pp. 197–208.
- [4] Y. Diao and C. Ernst, *Hamiltonian Cycles and Rope Lengths of Conway Algebraic Knots*, to J. Knot Theory and its Ramifications **15** (2006), pp. 121–142.
- [5] Y. Diao and C. Ernst, *Ropelengths of Closed Braids*, Topology and its Applications, **154** (2007) pp. 491–501.
- [6] Y. Diao, C. Ernst and M. Thistlethwaite, *The linear growth in the length of a family of thick knots*, J. of Knot Theory and its Ramifications, **12** (2003), pp. 709–715.
- [7] Y. Diao, C. Ernst and X. Yu, *Hamiltonian Knot Projections and Lengths of Thick Knots*, Topology and its Applications, **136**, no. 1 (2004), pp. 7–36.

- [8] Y. Diao, C. Ernst and U. Ziegler, *Generating Large Random Link Diagrams*, Physical and numerical models in knot theory, J. A. Calvo, K. C. Millett, E. J. Rawdon and A. Stasiak, editors, Series on Knots and Everything **36**, World Scientific (2005), pp. 473–494 .
- [9] Y. Diao, C. Ernst, R. Kavuluru and U. Ziegler, *Numerical Upper Bounds on Ropelengths of Large Physical Knots*, Journal of Physics A: Mathematical and General, Vol. 39, no. 18 (2006), pp. 4829–4843.
- [10] A. M. Law and W. D. Kelton, *Simulation modeling and analysis* , McGraw-Hill Series in Industrial Engineering and Management Science. McGraw-Hill Book Co., New York, 1982.
- [11] M. Matsumoto and T. Nishimura, *Mersenne Twister: A 623-Dimensionally Equidistributed Uniform Pseudo-Random Number Generator*, ACM Transactions on Modeling and Computer Simulation, **8** (1998), no. 1, pp. 3–30.
- [12] R. W. Robinson and N. C. Wormald, *Almost all regular graphs are Hamiltonian* , Random Structures Algorithms. **5** (1994), no. 2, pp. 363–374.
- [13] G. Schaeffer, *Bijective census and random generation of Eulerian plane maps with prescribed vertex degrees*, Electron. J. Combin. **4** (1997), no. 1, Research Paper 20, 14pp (electronic).
- [14] G. Schaeffer, *Random sampling of large planar maps and convex polyhedra*, in *Annual ACM Symposium on Theory of Computing* (Atlanta, GA, 1999), pp. 760–769, ACM, New York 1999.
- [15] P. Schaeffer, and P. Zinn-Justin, *On the asymptotic number of plane curves and alternating knots*, Experiment. Math. **13** (2004), no. 4, pp. 483–493.
- [16] A. Stasiak, V. Katritch and L. Kauffman, editors, *Ideal Knots*, World Scientific, 1998.
- [17] W. T. Tutte , *A Census of slicings*, Can. J. Math. **14** (1962), pp. 708–722.
- [18] W. T. Tutte , *A Census of planar maps*, Can. J. Math. **15** (1963), pp. 249–271.
- [19] W. T. Tutte , *A Census of Hamiltonian polygons*, Can. J. Math. **14** (1962), pp. 402–417.

DEPARTMENT OF COMPUTER SCIENCE, UNIVERSITY OF KENTUCKY, LEXINGTON, KY 40506
E-mail address: `onur.ascigil@uky.edu`

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF NORTH CAROLINA AT CHARLOTTE, CHARLOTTE, NC 28223, USA
E-mail address: `ydiao@uncc.edu`

DEPARTMENT OF MATHEMATICS, WESTERN KENTUCKY UNIVERSITY, BOWLING GREEN, KY 42101, USA
E-mail address: `claus.ernst@wku.edu`

DEPARTMENT OF MATHEMATICS, WESTERN KENTUCKY UNIVERSITY, BOWLING GREEN, KY 42101, USA
E-mail address: `david.high@wku.edu`

DEPARTMENT OF COMPUTER SCIENCE, WESTERN KENTUCKY UNIVERSITY, BOWLING GREEN, KY 42101
E-mail address: `uta.ziegler@wku.edu`