# A Brief Introduction to Hamilton Cycles in Random Graphs

Greg Brunet
Department of Computer Science
University of Toronto
Toronto, Canada

February 27, 2005

**Abstract**

We survey results concerning Hamilton cycles in random graphs. Specifically, we focus on existence results for general and regular graphs, and discuss algorithms for finding Hamilton cycles and solving related problems (that succeed with high probability).

## Contents

# 1 Introduction

Random graph theory is an active area of research that combines probability theory and graph theory. In its simplest form, the probabilistic method is used to prove the existence of combinatorial objects without constructing them. A probability model is defined on a large class of objects, and the occurrence of a desired structure is an event. If the event has positive probability, then some object with the desired structure exists. Erdös, one of the pioneers of this method, proved that for all naturals $g \geq 3$ and $k \geq 3$, there exists a graph with girth at least $g$ and chromatic number at least $k$. He did not construct the graphs, but showed that "most" graphs in a certain class could be altered to give the desired examples. This idea of applying the probabilistic method to extremal graph theory problems led to the theory of random graphs, introduced in a famous 1960 paper by Erdös and Rényi [6]. By defining appropriate probability spaces, they showed that for many monotone properties of random graphs, graphs of a size slightly less than a certain threshold are unlikely to have a given property, whereas graphs with slightly more edges are almost guaranteed to have it (this is referred to as *phase transition*). We can now answer questions such as "Is a typical graph connected?", "Is the chromatic number of a graph at least k?", or "Is a typical graph Hamiltonian?". The latter question was left open in [6] and is the subject of this survey.

More precisely, Erdös and Rényi asked the question: for what function $m(n)$ does the probability that a random graph with $n$ vertices and $m(n)$ edges contains a Hamilton cycle tend to 1 as $n$ tends to $\infty$? Now, if a graph has a Hamilton cycle, then it must be connected. Erdös and Rényi showed that $m(n) = \frac{1}{2}n \log n$ does not guarantee the connectivity of a graph or the existence of a 1-factor, with probability tending to 1. Years later, Komlós and Szemerédi showed that $m(n) = cne^{\sqrt{\log n}}$ edges guarantees the existence of a Hamilton cycle with probability tending to 1 [13]. However, Pósa improved their result by showing that for a sufficiently large constant $c$, $cn \log n$ edges suffices [17].

In this paper we will explore several improvements and variations of the early results mentioned above, as well as some algorithmic aspects. The purpose is to give a high-level overview of results related to the Hamiltonicity of random graphs. We focus more on ideas rather than details, as most of the results presented here require rather intricate and lengthy arguments. Specifically, Section 2 gives some preliminary definitions and describes notations, Section 3 presents existence results and some further properties of general random graphs and random regular graphs, and Section 4 presents some algorithms related to finding, counting, and generating Hamilton cycles in random graphs. Finally, in the Appendix we give an outline of the proof of Pósa's result mentioned above. Material that is not explicitly cited can be found in the text "Random Graphs" by Béla Bollobás [2].

# 2 Preliminaries

To discuss random graphs, we must first introduce the probability spaces (or models) of random graphs and give some subsequent definitions. All graphs are undirected unless otherwise stated. We assume familiarity with basic graph theory and probability theory.

The two models we focus on are $G(n, m)$ and $G(n, p)$. The model $G(n, m)$ consists of all graphs with vertex set $[n] = \{1, 2, \ldots, n\}$ having $m$ edges, where each graph is given the same probability. Defining $N = \binom{n}{2}$, for $0 \leq m \leq N$, $G(n, m)$ has $\binom{N}{m}$ elements and every element occurs with probability $\binom{N}{m}^{-1}$. We will write $G_{n,m}$ (or $G_m$ when $n$ is clear or unimportant) to denote a random element of this space. The model $G(n, p)$ (sometimes called the independent model) consists of all graphs with vertex set $[n] = \{1, 2, \ldots, n\}$ in which the edges are chosen independently with probability $p$, where $0 < p < 1$. It follows that each graph with $m$ edges has probability $p^m(1-p)^{N-m}$. We will write $G_{n,p}$ (or $G_p$) to denote a random element of this space. In most cases we use a sequence of probability spaces: for each natural $n$, there will be a probability space consisting of all graphs with exactly $n$ vertices. We are then interested in properties of this space as $n \to \infty$. Additional models will be defined as we use them.

We write $Pr[A]$ and $E[A]$ to denote the probability and expected value of an event $A$, respectively. Given a sequence of probability spaces, let $q_n$ be the probability that a property $Q$ (an event) holds in the $n$th space. Property $Q$ is said to *almost always* (a.a.) hold if $\lim_{n\to\infty} q_n = 1$. For random graphs, the $n$th space is a probability distribution over $n$-vertex graphs. When a property $Q$ almost always holds, we say that *almost every* (a.e.) graph has property $Q$.

A *monotone property* is a property that is preserved by the addition of edges. A *threshold function* for a monotone property $Q$ is a function $M(n)$ such that $p(n)/M(n) \to 0$ implies that almost no $G_p$ satisfies $Q$, and $p(n)/M(n) \to \infty$ implies that almost every $G_p$ satisfies $Q$ (defined similarly for $G_m$). For example, with respect to $G(n,p)$, $\frac{\ln n}{n}$ is a threshold function for $\delta(G) \geq 1$ (i.e., no isolated vertices).

A useful interpretation of a random graph is that of a living organism which evolves with time. It is born with a set of isolated vertices and grows by acquiring edges at random. The goal is then to determine at what stage of evolution a particular property of the graph is likely to occur. This idea will be treated more formally in section 3.2.

Finally, a *Hamilton path* is a path between two vertices of a graph that visits each vertex exactly once. A Hamilton path that is also a cycle is called a *Hamilton cycle.* A graph that contains a Hamilton cycle is said to be *Hamiltonian.*

# 3 Existence results

## 3.1 General graphs

The preliminary results proved by Pósa [17], and Komlós and Szemerédi [13] relied heavily on the notion of a *rotation.*

Suppose we have a path $P = x_0x_1...x_k$ in a graph $G$ and we wish to find a path of length $k+1$. If $x_0$ or $x_k$ has a neighbour not in $P$, then we can extend $P$ by adding the neighbour. If not, suppose $x_k$ has a neighbour $x_i$, where $0 \leq i \leq k-2$. If $i = 0$ and $G$ is connected, then there is an edge $e = (x_i, w)$ joining the cycle $x_0x_1...x_kx_0$ to the rest of the graph, and so the path $wx_ix_{i+1}...x_kx_0...x_{i-1}$ has length $k+1$. This is called a *cycle extension.* If $i \neq 0$, then we construct the path $x_0x_1...x_ix_kx_{k-1}...x_{i+1}$ of length $k$ with a different endpoint $x_{i+1}$ and look for further extensions. This is called a *rotation*, or a *simple transform.* An $x_0$-path is a *transform* of $P$ if it is obtained from $P$ by a sequence of simple transforms. These ideas will also be used in Section 4 to obtain algorithms for finding Hamilton cycles.

For a path $P = x_0x_1 \ldots x_k$, let $U$ be the set of endvertices of transforms of $P$ and let

$$N = \{x_i | 0 \leq i \leq k-1 \text{ and } \{x_{i-1}, x_{i+1}\} \cap U \neq \emptyset\}$$

and

$$R = V(P) - U \cup N.$$

Thus $N$ is the set of neighbours and $R$ is the rest of the vertex set of $P$. We can now prove the following lemma due to Pósa:

**Lemma 3.1** *There are no edges between $U$ and $R$.*

**Proof** Let $x \in U$ and $xy \in E(G)$. We will show that $y \notin R$. Let $P_x$ be a transform of $P$ ending at $x$. Since $xy \in E(G)$, $P_x$ has a simple transform $P_z$ ending in a vertex $z$ that is a neighbour of $y$ on $P_x$. Now, if $yz \in E(P)$, then $y \in N$. Otherwise, an edge $yw \in E(P)$ must have been deleted in a step of the sequence $P \to P' \to \ldots \to P_x$. When $yw$ was deleted the first time, then either $y$ or $w$ became an endvertex, so we have $\{y, w\} \subset U \cup N$. Therefore, $y \in U \cup N$, and hence $y \notin R$. ▮

This simple yet powerful lemma allowed Pósa to prove the following ground-breaking theorem.

**Theorem 3.2** *Let $m = \lfloor cn \log n \rfloor$. Then a random graph $G_{n,m} \in G(n,m)$ almost surely contains a Hamilton cycle.*

**Proof** Outlined in the Appendix. ▮

In other words, for some constant $c$, almost every labeled graph with $n$ vertices and at least $cn \log n$ edges is Hamiltonian. In fact, $c$ is chosen sufficiently large so as to guarantee that if $G$ is a graph with $n$ vertices and the edges are drawn independently with probability $\frac{c \log n}{n}$, then $Pr[G$ has a Hamilton cycle$] \to 1$ as $n \to \infty$.

We now turn to the task of determining the threshold function of a Hamilton cycle. Note that if $G$ is Hamiltonian, then $\delta(G) \geq 2$, so

$$Pr[G_m \text{ is Hamiltonian}] \leq Pr[\delta(G_m) \geq 2].$$

It can be shown that $Pr[\delta(G_m) \geq 2] \to 1$ iff $\omega(n) = 2m/n - \log n - \log\log n \to \infty$. Therefore, if a.e. $G_m$ is Hamiltonian, we must have $\omega(n) \to \infty$. It turns out that this condition is also sufficient, and was first proven by Komlós and Szemerédi using rotations and cycle extensions [14]. We give a sketch of a different proof found in [2].

**Theorem 3.3** *Let $\omega(n) \to \infty$, $p = (1/n)(\log n + \log\log n + \omega(n))$, and $m(n) = \lfloor (n/2)(\log n + \log\log n + \omega(n))\rfloor$. Then a.e. $G_p$ and a.e. $G_m$ are Hamiltonian.*

**Proof** (Sketch) Since the property of being Hamiltonian is monotone, it suffices to prove the result for $G_p$.

Let $k = \lfloor 4n/\log n \rfloor$, $p_i = 64 \log n/n^2$, $1 \leq i \leq k$, and $p_0 = p - kp_1 = p - (64 \log n/n^2)k$. Additionally, let $G_c(j) = G_c(n, p_0, \ldots, p_j)$ be the space of edge-coloured multigraphs whose edges are drawn independently, with probability $p_i$ of drawing a colour $i$ edge. The idea is to replace multiple edges by simple edges and ignore the colours. We can then identify $G_c(j)$ with $G(n, p'_j)$, where $p'_j \leq p_0 + jp_1$. Since $p'_k \leq p$, it then suffices to show that a.e. graph in $G_c(k)$ is Hamiltonian.

For $G_k \in G_c(k)$ and $0 \leq j \leq k$, let $G_j$ be the subgraph of $G_k$ formed by the edges of colours $0, 1, \ldots, j$. Then $G_k \to G_j$ defines a measure-preserving map of $G_c(k)$ into $G_c(j)$ (i.e., an endomorphism between probability spaces). To obtain a random element of $G_c(j)$ from a random element of $G_c(j-1)$, we add edges of colour $j$ with probability $p_j = p_1$.

Define $l(G)$ to be the length of the longest path in $G$. If $G$ is Hamiltonian, then set $l(G) = n$. Now let $Q$ be the property that $G$ is connected and if $U \subset V(G)$ and $|U| \leq n/4$, then

$$|U \cup \Gamma(U)| \geq 3|U|,$$

where $\Gamma(U) = \{y | xy \in E(G) \text{ for some } x \in U\}$. Then, the proof reduces to showing the following two conditions:

$$P_0 = Pr[l(G_0) \geq n - k \text{ and } G_0 \text{ has } Q] = 1 - o(1), \tag{1}$$

and

$$P_j = Pr[l(G_j) \geq n - k + j - 1 | l(G_{j-1}) = n - k + j - 1 \text{ and } G_{j-1} \text{ has } Q] = \leq n^{-2}, \tag{2}$$

for $j = 1, 2, \ldots, k$. In particular, conditions (1) and (2) give

$$Pr[l(G_k) < n] \leq 1 - P_0 - \sum_{j=1}^{k} P_j = o(1).$$

However, showing conditions (1) and (2) relies on a series of lemmas, and so we omit the details. See [2] for a more detailed treatment. ∎

## 3.2 A stronger hitting time result

Recall that a random graph can be thought of as a living organism that acquires edges one by one at random. We can formalize this intuition using *graph processes.* Given $V = \{1, 2, \ldots, n\}$, a *graph process* is a sequence $\{G_t\}_{t=0}^{N}$ such that

- every $G_t$ is a graph on V,
- $G_t$ has $t$ edges for $t \in \{0, \ldots, N\}$, and

- $G_0 \subset G_1 \subset \dots$

Let $\widetilde{GP}$ be the set of $N!$ graph processes. We turn $\widetilde{GP}$ into a probability space by giving each element the same probability, and write $\tilde{G}$ for a random element of $\widetilde{GP}$. Additionally, $G_t$ is called the *state* of the process $\tilde{G} = \{G_t\}_{t=0}^{N}$ at time $t$.

Now let $Q$ be any monotone graph property. We define the *hitting time* $\tau$ of a property $Q$ to be the first time at which $Q$ appears:

$$\tau = \tau_Q(\tilde{G}) = \min\{t \geq 0 | G_t \text{ has } Q\}. \tag{3}$$

Specifically, consider the property of being Hamiltonian. Since it is required (but not sufficient) for a Hamiltonian graph $G$ to have minimum degree at least 2, it follows that the hitting time of being Hamiltonian is at least as large as the hitting time for having minimum degree at least 2. Amazingly, we can in fact prove equality: the edge that gives a graph $G$ minimum degree at least 2 also guarantees that $G$ is Hamiltonian (almost surely). This is stated in the following theorem, first proven by Bollobás in [1].

**Theorem 3.4** *Let $Q_{ham}$ be the property that a given graph is Hamiltonian, and $Q_{md}$ be the property that the minimum degree $\delta$ of a given graph is at least 2. For almost all graph processes $\tilde{G}$,*

$$\tau_{Q_{ham}}(\tilde{G}) = \tau_{Q_{md}}(\tilde{G}).$$

Theorem 3.4 tells us that the primary obstruction for the existence of Hamilton cycles is the existence of vertices of degree at most 1.

## 3.3 Regular graphs

We now turn our attention to existence results for random $r$-regular graphs.

Let $\Omega_{2n,r}$ be the set of all labeled $r$-regular graphs on $2n$ vertices. Turn $\Omega_{2n,r}$ into a probability space using the uniform distribution. For the rest of this section $r \geq 3$ is fixed unless otherwise stated. We are interested in the smallest value $K$ such that for all fixed $r \geq K$, almost all $r$-regular graphs are Hamiltonian. The long standing conjecture that $K = 3$ was settled by Robinson and Wormald in a series of two papers [18], [19]. The best previous result was $K = 85$.

**Theorem 3.5** *For fixed $r \geq 3$, almost every $r$-regular graph is Hamiltonian.*

The details of the proof of Theorem 3.5 are quite involved, so we will give a rough outline of the ideas behind the proof, as well as some technical details. The first step was to show that almost all cubic graphs ($r = 3$) are Hamiltonian. Let $H$ denote the number of Hamilton cycles of a graph $G \in \Omega_{2n,3}$.

**Theorem 3.6** *If $H$ is the number of Hamilton cycles in a cubic graph chosen uniformly at random from all labeled cubic graphs on $2n$ vertices, then $\lim_{n\to\infty} Pr[H > 0] = 1$.*

It can be shown that,

$$E[H] \sim \frac{e\sqrt{\pi}}{2\sqrt{n}} \left(\frac{4}{3}\right)^n, \tag{4}$$

$$Var[H] \sim \left(\frac{3}{e} - 1\right) (E[H])^2. \tag{5}$$

Using equations (4) and (5), it can be shown by the standard second moment method[1] that asymptotically at least $2 - e^{-1}$ of all cubic graphs are Hamiltonian. In fact, a stronger bound of $2 - 3e^{-13/12}$ can be shown by considering the space of triangle-free cubic graphs. However, the variance in eq. (5) is too large to deduce that almost all cubic graphs are Hamiltonian.

The key idea used in [18] is that the second moment method can still be used to prove Theorem 3.6 if $\Omega_{2n,3}$ is partitioned into some useful subsets. In particular, for $G \in \Omega_{2n,3}$, let $X_i(G)$ be the number of cycles of length $i$ in $G$. Then we have the following lemma found in [2] (originally proved in [21]):

[1] Informally, the second moment method consists of applying Chebyshev's inequality. Let $X$ be a random variable with mean $\mu$ and variance $\sigma^2$ (both finite). Chebyshev's inequality says that for all positive $k$, $Pr[|X - \mu| > k] \leq \frac{\sigma^2}{k^2}$

**Lemma 3.7** *For any fixed $k$ the variables $X_i$, $3 \leq i \leq k$, are asymptotically independent Poisson random variables, with*

$$E[X_i] \sim \frac{2^{i-1}}{i}. \tag{6}$$

The approach is then to divide the cubic graphs in $\Omega_{2n,3}$ into groups according to the values of the variables $X_i$, and refine the second moment method for these groups. It turns out that the variables $X_i$ for $i$ even have no effect asymptotically, and so the groups studied are the groups characterized by a sequence $c_1, \ldots, c_b$ for fixed $b$, where $c_i = X_{2i+1}$ for $i = 1, \ldots, b$. It is then shown that by taking $b$ to be large, the group means can be made arbitrarily close to the total variance $Var[H]$. Therefore, the mean of the group variances can be made arbitrarily small compared to $Var[H]$. Consequently, it is deduced that $Pr[H = 0]$ is bounded above by $\epsilon(b)$, where

$$\lim_{n \to \infty} \epsilon(b) = 0, \tag{7}$$

which establishes the desired result.

To prove the general result for $r \geq 3$ of Theorem 3.5 turns out to be significantly more difficult, and is not just a simple extension of the method used to prove the cubic case. The difficulty stems from the fact that a subgraph induced by the union of two Hamilton cycles in $G$ has vertices of degree 2 or 3 in cubic graphs, whereas when $G$ is $r$-regular for $r \geq 4$ there can also be vertices of degree 4. This makes it more complicated to compute the variance of the number of Hamilton cycles. Instead, an approach involving perfect matchings gives a more elegant solution (although still quite difficult). In particular, let $M$ be the number of perfect matchings of a graph $G \in \Omega_{2n,r}$ (Note that $M = 0$ for odd order graphs, so this analysis is restricted to even order graphs). It can then be shown that

$$\frac{E[M^2]}{(E[M])^2} \sim e^{-(2r-1)/4(r-1)^2} \sqrt{\frac{r-1}{r-2}}. \tag{8}$$

The approach is again to divide $\Omega_{2n,r}$ into groups according to the values of $X_i$ for $i = 3, \ldots, b$, where $b \geq 3$, then refine the second moment method for $M$ using these groups. The goal is to show that $M$ is unlikely to be different from its mean by more than a constant factor. Hence, informally, the removal of a random perfect matching from $G$ produces a reasonably random $(r-1)$-regular graph. Then, Hamiltonicity of this graph implies Hamiltonicity of $G$.

In fact, the proof of Theorem 3.5 establishes much more. Define a *complete decomposition* of $G$ to be a decomposition of $G$ into a set of perfect matchings and the edges of a Hamilton cycle. Then the following Theorem is used to prove Theorem 3.5, and is the main result of [19]:

**Theorem 3.8** *For fixed $r \geq 3$, almost all $r$-regular graphs with an even number of vertices have a complete decomposition.*

**Proof** (Sketch) The proof proceeds by induction on $r$. The base case $r = 3$ follows from Theorem 3.6, because a cubic graph with a Hamilton cycle must have a complete decomposition. Thus let r $\geq 4$. As mentioned, the main objective is to show that for $G \in \Omega_{2n,r}$, $M$ is unlikely to be very small compared with its expectation. Once it is shown that there is a sequence $w(y) > 0$ such that

$$\liminf Pr[M \geq w(y)E[M]] \to 1 \text{ as } y \to \infty, \tag{9}$$

the following argument is given that links the spaces $\Omega_{2n,r}$ and $\Omega_{2n,r-1}$, which establishes the theorem (see [19] for the details of $w(y)$).

Let $R$ denote the event $\{M \geq w(y)E[M]\}$. Define a bicoloured graph $B$, in which the blue vertices are the elements of $\Omega_{2n,r}$ and the red ones are the elements of $\Omega_{2n,r-1}$, with an edge from a blue vertex $v_1$ to a red vertex $v_2$ iff $v_2$ can be obtained from $v_1$ by deleting the edges of a perfect matching.

Let $T_r$ be the event that $G \in \Omega_{2n,r}$ does not have a complete decomposition, and use $Pr_r$ for the probability in the space $\Omega_{2n,r}$. Choose an edge $(v_1, v_2)$ of $B$ uniformly at random with $v_1 \in \Omega_{2n,r}$ and

$v_2 \in \Omega_{2n,r-1}$. For this selection, let $P_1$ be the probability that $v_1 \in R \cap T_r$, and let $P_2$ be the probability that $v_2 \in T_{r-1}$. By definition of $R$, we have

$$P_1 \geq w(y) Pr_r[R \cap T_r]. \tag{10}$$

However, it can be shown that the maximum and minimum degrees of the red vertices in $B$ are asymptotically equal uniformly as $n \to \infty$. Therefore,

$$P_2 \leq Pr_{r-1}[T_{r-1}](1 + o(1)). \tag{11}$$

If $v_2$ has a complete decomposition then so does $v_1$, and so $P_1 \leq P_2$. Thus, equations (10) and (11) with the inductive hypothesis that $Pr_{r-1}[T_{r-1}] = o(1)$, imply that $Pr_r[R \cap T_r] = o(1)$ for $y$ fixed. Thus, by eq. (9), since $y$ can be chosen arbitrarily large, the theorem follows. ▮

Immediately from Theorem 3.8 we have the following corollary.

**Corollary 3.9** *For fixed $r \geq 3$, almost all $r$-regular graphs with an even number of vertices are $r$-edge-colourable.*

To complete the proof of Theorem 3.5, any graph with $2n + 1$ vertices is altered to obtain a graph with $2n$ vertices. By Theorem 3.8, the altered graph has a Hamilton cycle. The idea is then to lift this Hamilton cycle to the original graph, which establishes the theorem. See [19] for details.

The ideas used in the proof of Theorem 3.5 were extended by Molloy, Robalewska, Robinson, and Wormald [16] and independently by Janson [12] to apply to a related situation. Given two sequences $\{\Lambda_i\}$ and $\{\Lambda'_i\}$ of probability spaces where $\Lambda_i$ and $\Lambda'_i$ have the same underlying sets, $\Lambda_i$ and $\Lambda'_i$ are said to be *contiguous* when every event is almost sure in $\Lambda_i$ if and only if it is almost sure in $\Lambda'_i$. Now define $\Omega^*_{2n,r}$ to be the probability space with the same domain as $\Omega_{2n,r}$, in which each graph occurs with probability proportional to the number of ways in which its edge set can be partitioned into the edges of a $k$-regular graph and $r - k$ 1-factors, for $2 \leq k \leq r - 1$. The proof of Theorem 3.8 implies that for all $r \geq 3$, $\Omega_{2n,r}$ and $\Omega^*_{2n,r}$ are contiguous. Additionally, define $\Omega^+_{2n,r}$ to be the probability space with the same domain as $\Omega_{2n,r}$, in which each graph occurs with probability proportional to the number of ways in which its edge set can partitioned into the edges of $r$ 1-factors. Then the main result in [16] is that $\Omega_{2n,r}$ and $\Omega^+_{2n,r}$ are contiguous for all $r \geq 3$.

# 4 Algorithms

The problem of finding a Hamilton cycle in a graph is NP-complete, so we can not expect to find polynomial time algorithms that always succeed, unless $P = NP$. However, using random graphs as models for the average case analysis of graph algorithms banishes the pessimism of worst-case analysis. In particular, NP-completeness casts a much smaller shadow on problems related to Hamilton cycles, making them tractable, in a sense. In this section, we survey a few of these results.

## 4.1 HAM: finding Hamilton cycles in general random graphs

We describe a polynomial time algorithm due to Bollobás, Fenner and Frieze [3] for finding Hamilton cycles in undirected graphs. The main result is an algorithm HAM that, when run on a random graph, has asymptotic probability of success equal to the existence of such a cycle.

As mentioned in Section 3.1, Komlós and Szemerédi settled the existence question for Hamilton cycles. Their result was proven using rotations and cycle extensions and is stated in a stronger form than in Section 3.1 as follows:

**Theorem 4.1** *Let $m(n) = (n/2)(\log n + \log\log n + c_n)$ for some sequence $c_n$. Then*

$$\begin{aligned}\lim_{n\to\infty} Pr[G_{n,m} \textit{ is Hamiltonian}] &= \begin{cases} 0, & \textit{if } c_n \to -\infty; \\ e^{-e^{-c}}, & \textit{if } c_n \to c; \\ 1, & \textit{if } c_n \to +\infty. \end{cases} \\ &= \lim_{n\to\infty} Pr[G_{n,m} \textit{ satsifies } \delta(G_{n,m}) \geq 2].\end{aligned}$$

The best previous polynomial algorithm, call it HAM1, is due to Shamir [20]. It was shown to satisfy

$$\lim_{n\to\infty} Pr[\text{HAM1 finds a Hamilton cycle in } G_{n,m}] \to 1,\ \text{ if } c_n > (\frac{3}{2}+\epsilon)\log\log n \text{ for } \epsilon > 0 \text{ fixed,}$$

where $c_n$ is as in Theorem 4.1.

The algorithm HAM improves upon this result, giving (essentially) the best possible result by extending the ideas used in the proof of Theorem 4.1. As hinted at in Section 3.1, HAM works by maintaining a list of partial paths that it tries to extend until a Hamilton path is obtained or a dead end is hit (i.e., all partial paths are checked and no extension can be made). We can assume connectivity of the input graph, otherwise no Hamilton cycle exists. The initial partial path contains the first vertex in the graph and its first neighbour. To search for extensions of the current partial path, HAM tries each neighbour of both endpoints of the path. If any neighbour of one of the endpoints is unvisited, then we can clearly extend the current partial path by taking that neighbour.

If there is no edge from either of the current endpoints of the current partial path to a neighbour outside the path, then HAM checks if the partial path forms a cycle (i.e., if there is an edge between the two endpoints). If a cycle exists, then the algorithm scans through the vertices of the cycle for a vertex with a neighbour not on the cycle. By connectivity, such a vertex can always be found. The new vertex is made the endpoint of a new path, and one of its neighbours is made the other endpoint. So, HAM can add at least one vertex to the path, and thus using cycle extensions the partial path is extended. See Figure 1 for an example of a cycle extension.

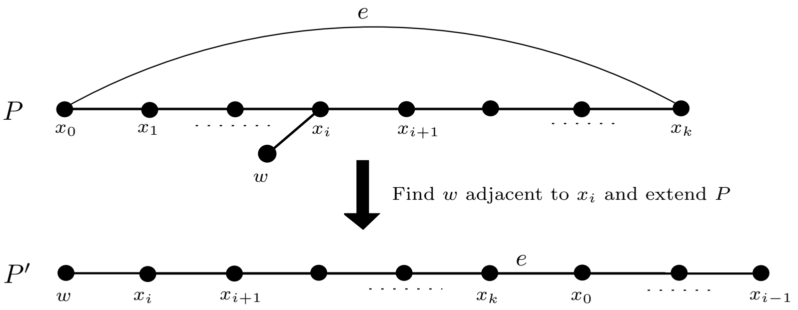

Figure 1: A cycle extension of $P = x_0x_1\dots x_k$ to $P' = wx_ix_{i+1}\dots x_kx_0x_1\dots x_{i-1}$

If no such cycle exists, then suppose that $x_k$ (an endpoint) has a neighbour $x_i$ (not an endpoint) on the path. HAM performs a rotation using the edge $(x_i, x_k)$ to obtain a new partial path, and adds it to the list of paths to check. See Figure 2 for an example of a rotation. After all neighbours of the initial partial path are checked, the algorithm will begin checking these other paths (breadth-first search). In other words, HAM backtracks through the possible sets of rotations of the initial partial path. The path extension search terminates by setting a limit on the search depth, which is equal to the number of rotations that were applied to the original partial path.
It was shown in [3] that HAM satisfies the following theorem:

**Theorem 4.2** *Let $m(n)$ be as in Theorem 4.1, then*

$$\lim_{n\to\infty} Pr[\textit{HAM finds a Hamilton cycle in } G_{n,m}] = \begin{cases} 0, & \textit{if } c_n \to -\infty; \\ e^{-e^{-c}}, & \textit{if } c_n \to c; \\ 1, & \textit{if } c_n \to +\infty. \end{cases}$$

*Furthermore, for $\epsilon > 0$, HAM runs in $o(n^{4+\epsilon})$ time.*

Now consider the model $G(n,p)$ with $p = 1/2$. Let $A$ be the property that the given graph is connected, has minimum degree at least 2, yet HAM terminates unsuccessfully. It is shown that $Pr[G_{n,1/2}$ has $A] = o(2^{-n})$.

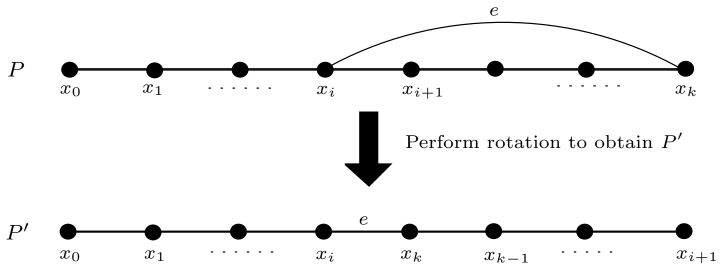

Figure 2: A rotation of $P = x_0 x_1 \ldots x_k$ to $P' = x_0 x_1 \ldots x_i x_k x_{k-1} \ldots x_{i+1}$ using the edge $e = (x_i, x_k)$.

Because this probability of failure is so small, it is possible to apply dynamic programming techniques [10] on the failed cases, which takes time $o(n^2 2^n)$, to obtain the following result.

**Theorem 4.3** *There is an algorithm for solving the Hamilton cycle problem with polynomial expected running time.*

Therefore, for $G(n, 1/2)$, we have a hybrid algorithm that *always succeeds* and has polynomial expected running time.

## 4.2 SparseHam: a variation of HAM

In [9], Frieze presents an algorithm SparseHam, which is a variation of the algorithm HAM discussed in the previous section. It is shown that SparseHam almost surely finds a Hamilton cycle in two classes of random graphs with constant average degree. SparseHam uses the same techniques as HAM (i.e. cycle extensions and backtracking through path extensions), but there are three major differences in implementation.

1. SparseHam uses depth-first search for backtracking through path extensions, rather than breadth-first search.

2. SparseHam limits the paths that are searched (HAM had no such criteria). It stores the edge from the current endpoint to the selected vertex (rotation edge) for each rotational transform. While the path has not been extended, future rotations are not allowed to use that edge as a rotation edge. This prevents repetition of partial paths when backtracking.

3. SparseHam only considers extending the current partial path from one of its endpoints, whereas HAM considers both. If SparseHam cannot find a path extension, it will search back through the list of paths and look for an extension from the other endpoint.

The remaining details are analogous to HAM.

## 4.3 Finding hidden Hamilton cycles

In this section we describe the "hidden structure" version of the search problem for Hamilton cycles.

Suppose we a given a graph $G$ that is known to contain a Hamilton cycle $H$, "hidden" among relatively few additional random edges that by themselves are unlikely to admit a Hamilton cycle. The goal is find a Hamilton cycle (not necessarily $H$) in $G$. Specifically, a random graph $G$ in this problem is defined by $n$ labeled vertices with an edgeset that is the union of a random graph $G_{n,d/n}$ and an $n$-edge Hamilton cycle ($d$ fixed). The class of graphs in $G(n, d/n)$ that are Hamiltonian has a negligible probability as $n$ tends to infinity. Effectively, we have boosted this probability to 1, and so can ask for an algorithm that will almost surely produce a Hamilton cycle in $G$. In [4], Broder, Frieze, and Shamir show that there exists an $O(dn^3)$ algorithm, say HIDDEN-HAM, that solves this problem and succeeds almost surely.

The algorithm HIDDEN-HAM essentially consists of two parts. The first part identifies a set $X$ of potentially "obstructive" vertices. Naturally, these are the vertices of low degree, because it is the low degree degree vertices that make it harder for Hamilton cycles to exist and be found. Then a set of disjoint paths is constructed such that each path consists of a least one $x \in X$, but not as a endpoint. In a sense, by finding disjoint paths that contain the low degree vertices, the graph is "reduced" to a graph of large minimum degree.

The second part consists of using a rotation-extension type algorithm, which attempts to join the disjoint paths constructed in part one to form a Hamilton cycle. The success of this part relies on the fact that the partial paths were constructed with endpoints of high degree. Therefore, the number of rotations performed is almost surely enough to find a path extension.

The algorithm HIDDEN-HAM has some implications for encryption and knowledge hiding protocols. Most modern encryption schemes are based on number theory. There have also been several attempts to hide bits via hidden structures on graphs. However, the efficient algorithm described here for finding hidden Hamilton cycles rules out this combinatorial candidate and further adds to the evidence that hidden combinatorial structures in general are more prone to failure than the number theoretic solutions.

## 4.4 Generating and counting Hamilton cycles in random regular graphs

Recall that in Section 3.3 we saw that for fixed $r \geq 3$, almost every $r$-regular graph is Hamiltonian. The proof of this theorem was non-constructive. The corresponding algorithmic questions were tackled by Frieze, Jerrum, Molloy, Robinson, and Wormald [7]. They stray from the rotation-extension approach and instead use an approach based on rapidly mixing Markov chains[2]. We give a brief outline of their results here.

The first problem tackled is that of finding a Hamilton cycle in a random $r$-regular graph.

**Theorem 4.4** *Let $r \geq 3$ be fixed and let $G$ be chosen uniformly at random from $\Omega_{n,r}$. There is a polynomial time algorithm FIND that almost always constructs a Hamilton cycle in $G$.*

In fact, the procedure FIND can be used to solve one of the open problems in [4]. Consider a graph $G$ with vertex set $[n] = \{1, 2, \ldots, n\}$ obtained by adding a random perfect matching to a random Hamilton cycle $H$. Then the problem is similar to that of Section 4.3: find a Hamilton cycle in $G$ without knowing $H$. It is shown that if $n$ is even and $G$ is obtained as the union of a random perfect matching $M$ and a random (disjoint) Hamilton cycle $H$, then applying FIND to $G$ almost always leads to the construction of a Hamilton cycle. Again, this is a negative result for authentication protocols using hidden combinatorial structures.

Now for any graph $G$, let $H(G)$ be the set of Hamilton cycles in $G$. If $H(G) \neq \emptyset$, then a *near uniform generator* for $H(G)$ is a randomized algorithm which on input $\epsilon > 0$ outputs a cycle $H \in H(G)$ such that for any fixed $H_1 \in H(G)$

$$\left| Pr[H = H_1] - \frac{1}{|H(G)|} \right| \leq \frac{\epsilon}{|H(G)|}. \tag{12}$$

Note that the probabilities above are with respect to the random choices of the algorithm, because $G$ is fixed in eq. (12). Additionally, the algorithm is polynomial if it runs in time polynomial in $n$ and $1/\epsilon$. The authors of [7] give a procedure GENERATE that is a polynomial time generator for $H(G)$. More precisely,

**Theorem 4.5** *Let $r \geq 3$ be fixed. There is a procedure GENERATE such that if $G$ is chosen uniformly at random from $\Omega_{n,r}$, then GENERATE is almost always a polynomial time generator for $H(G)$.*

Polynomial time generators for a given set $X$ can often be used to estimate $|X|$. Thus, having a generator for the set of Hamilton cycles in a graph $G$ may lead to a good estimate of the number of Hamilton cycles in $G$. The authors of [7] give a procedure that is an FPRAS (Fully Polynomial Randomized Approximation Scheme) for $H(G)$. Formally, an FPRAS for $H(G)$ is a randomized algorithm which on input $\epsilon$, $\delta > 0$ produces an estimate $Z$ such that

$$Pr\left[ \left| \frac{Z}{|H(G)|} - 1 \right| \geq \epsilon \right] \leq \delta. \tag{13}$$

[2] For our purposes, it is enough to know that this is a technique that has applications in approximate counting.

The probabilities in (13) are with respect to the algorithms choices, and the running time of the algorithm must be polynomial in $n, \epsilon^{-1}$, and $\log(\delta^{-1})$.

**Theorem 4.6** *Let $r \geq 3$ be fixed. There is a procedure COUNT such that if $G$ is chosen uniformly at random from $\Omega_{n,r}$, COUNT is almost surely an FPRAS for $H(G)$.*

Roughly speaking, an FPRAS is a notion of an efficient (or "good") approximation algorithm (See [11]).

# 5 Closing remarks

We have seen several results relating to the existence of Hamilton cycles in random graphs, as well as algorithms for finding them (and in some cases generating and counting them) that succeed almost surely. This is by no means a summary of the topic, as such would require a much larger exposition.

Although we have focused on general and regular random graphs, it should be noted that many of the results presented here have counterparts for directed graphs. Recall that rotations and cycle extensions played a major role in undirected graphs. However, we no longer have this tool in directed graphs, as the orientation of part of the path is reversed. Thus, the problem of determining the threshold function for random directed graphs requires a different type of argument. See [15] and [8] for existence and algorithmic results. Additionally, in [5], Cooper, Frieze, and Molloy prove an analog of Theorem 3.6: almost every 3-regular digraph is Hamiltonian.

# 6 Appendix

## 6.1 Proof of Pósa's result

Here we present an outline of the proof of Theorem 3.2, that a random graph $G \in G(n,m)$, with $m = cn \log n$, almost surely contains a Hamilton cycle[3].

Recall the definitions of $U$ and $R$ from Lemma 3.1. Also recall that Lemma 3.1 (we will refer to this as Lemma 1 from now on) tells us that there are no edges between $U$ and $R$. In Lemma 2 of Pósa's proof, he considers a random graph $G \in G(n,p)$, with $p = (c \log n)/n$. Using Lemma 1, it is shown that the probability that for some $q \leq \frac{1}{4}n$ there exists a set $A$ of $q$ vertices and a set $B$ (disjoint from $A$) of $n - 3q - 1$ vertices such that no edges joins a vertex of $A$ to a vertex of $B$, tends to 0 as $n$ tends to infinity. The rest of the proof depends on two theorems.

**Theorem 6.1** *For a random graph $G \in G(n,p)$, with $p = (c \log n)/n$, $G$ almost surely contains a Hamilton path.*

**Proof** (Sketch) Let us denote by $G(x)$ the graph with $n-1$ vertices obtained from $G$ by erasing $x$. We want to estimate the probability of $L(x)$ , a path of maximum length in $G$ that passes through $x$.

Choosing an arbitrary path $P$ of maximum length in $G(x)$, we define the sets $U$ and $R$ (from Lemma 1) in $G(x)$. There are two possible cases. In the first case, $U$ is small enough that Lemma 2 can be used. The second case involves larger sizes of $U$. If $x$ is adjacent to any element of $U$, then the path $P$ can be transformed to a path $P'$ (using a rotation) with an endpoint adjacent to $x$. This allows the path $P$ to be extended to include $x$. However, $U$ is large enough that the probability of a vertex $x$ existing not adjacent to $U$ approaches zero.

Therefore, there are two possible events that prohibit the formation of a Hamilton path. The probability of the first event is calculated in Lemma 2, and the probability of the second event is also shown to approach 0. From this, we have that with probability tending to 1, every path of maximum length in $G$ passes through all points of $G$. ▮

[3] It is clear from the context when we mean $\lfloor cn \log n \rfloor$ rather than $cn \log n$.

**Theorem 6.2** *For a random graph $G \in G(n,p)$, with $p = (c_1 \log n)/n$, where $c_1$ is sufficiently large, $G$ almost surely has a Hamilton cycle.*

**Proof** (Sketch) Construct two graphs $G_1$ and $G_2$ on the same $n$ vertices with edge probabilities $(c \log n)/n$ and $(\log n)/n$, respectively. Let $G$ be the union of the two sets of edges. Then $G$ is also a random graph, with edge probability $((\log n)/n)(c + 1 - (c \log n)/n)$. By Theorem 6.1, the probability that $G_1$ contains a Hamilton path tends to 1. Again, the sets $U$ and $R$ are defined as in Lemma 1. If $G$ contains no Hamilton cycle, then one of the following three events of small probability must occur:

- There is no Hamilton path in $G_1$ (occurs with small probability by Theorem 6.1).
- $U$ is small enough so that Lemma 2 holds, and so the probability of this event tends to 0.
- $U$ is too large for Lemma 2. In this case $x_n$ (the endpoint of the Hamilton path in $G_1$) cannot be joined by edges in $G_2$ to elements of $U$ (for suppose that $x_n$ is joined to $h \in U$. By definition of $U$, there exists a path $P^*$ with endpoints $x_n$ and $h$ which can be obtained by rotations from $P$. But $P^*$ together with the edge $(x_n, h)$ is a Hamilton cycle). The probability of the event that there is no edge in $G_2$ between $x_n$ and elements of $U$ tends to 0.

Therefore, the overall probability of no Hamilton cycle existing tends to 0. ▮

To complete the proof, consider a random graph $G_1 \in G(n,p)$, with $p = (c_1 \log n)/n$. If the number of edges in $G_1$ is less than $c_1 n \log n$, then place additional edges in $G_1$ at random until the number of edges is equal to $c_1 n \log n$ (otherwise we have nothing to do). Call the resulting graph $G_2$. Now let $A$ be the event that $G_2$ contains a Hamilton cycle, and $B$ be the event that the number of edges in $G_1$ is less than $c_1 n \log n$. It is not hard to show that $Pr[B] \rightarrow 1$ using Chebyshev's inequality. Also, from Theorem 6.2, $Pr[A] \rightarrow 1$ since even $G_1$ contains a Hamilton cycle with probability tending to 1. Putting these together, we have $Pr[A|B] \rightarrow 1$, which is precisely the probability we wanted.

# References

[1] B. Bollobás. The evolution of sparse graphs. *Graph Theory and Combinatorics (Proc. Cambridge Combinatorics Conference in Honors of Paul Erdös), Academic Press*, pages 35–57, 1984.

[2] B. Bollobás. *Random Graphs.* Cambridge University Press, 2nd edition, 2001.

[3] B. Bollobás, T.I. Fenner, and A.M. Frieze. An algorithm for finding Hamilton paths and cycles in random graphs. *Combinatorica*, 7(4):327–341, 1987.

[4] A.Z. Broder, A.M. Frieze, and E. Shamir. Finding hidden hamilton cycles. *Proc. 23rd Annual ACM Symposium on Theory of Computing*, pages 182–189, 1991.

[5] C. Cooper, A. Frieze, and M. Molloy. Hamilton cycles in random regular digraphs. *Combinatorics, Probability, and Computing*, 3:39–50, 1994.

[6] P. Erdös and A. Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hungar. Acad. Sci.*, 5:17–61, 1960.

[7] A. Frieze, M. Jerrum, M. Molloy, R.W. Robinson, and N.C. Wormald. Generating and counting hamilton cycles in random regular graphs. *J. Algorithms*, 21:176–198, 1996.

[8] A.M. Frieze. An algorithm for finding hamilton cycles in random digraphs. *J. Algorithms*, 9:181–204, 1988.

[9] A.M. Frieze. Finding hamilton cycles in sparse random graphs. *Journal of Combinational Theory, Series B*, 44:230–250, 1988.

[10] M. Held and R.M. Karp. A dynamic programming approach to sequencing problems. *SIAM J. Appl. Math.*, 10:196–210, 1962.

[11] Dorit S. HochBaum, editor. *Approximation algorithms for NP-Hard problems.* PWS Publishing Company, 1997.

[12] S. Janson. Random regular graphs: asymptotic distributions and contiguity. *Combinatorics, Probability and Computing*, 4:369–405, 1995.

[13] J. Komlós and E. Szemerédi. Hamilton cycles in random graphs. *Colloq. Math. Soc. J. Bolyai*, 10:1003–1011, 1975.

[14] J. Komlós and E. Szemerédi. Limit distribution for the existence of hamilton circuits in a random graph. *Discrete Math*, 43:55–63, 1983.

[15] C.J.H. McDiarmid. General percolation and random graphs. *Adv. Appl. Probab.*, 13:40–60, 1981.

[16] M.S.O. Molloy, H. Robalewska, R.W. Robinson, and N.C. Wormald. 1-factorizations of random regular graphs. *Random Struct. Alg.*, 10:305–321, 1997.

[17] L. Pósa. Hamiltonian circuits in random graphs. *Discrete Math*, 14:359–364, 1976.

[18] R.W. Robinson and N.C. Wormald. Almost all cubic graphs are hamiltonian. *Random Struct. Alg.*, 3:117–126, 1992.

[19] R.W. Robinson and N.C. Wormald. Almost all regular graphs are hamiltonian. *Random Struct. Alg.*, 5:363–374, 1994.

[20] E. Shamir. How many edges are needed to make a random graph hamiltonian? *Combinatorica*, 1983.

[21] N.C. Wormald. *Some Problems in the Enumeration of Labelled Graphs.* Doctoral thesis, Newcastle University, 1978.